eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Fusion Rules for Distributed Detection in Clustered Wireless Sensor Networks with Imperfect Channels

Sami A. Aldalahmeh, *Member, IEEE,* Saleh O. Al-Jazzar, Des McLernon, *Member, IEEE,*
Syed Ali Raza Zaidi, *Member, IEEE,* and Mounir Ghogho, *Fellow, IEEE*

Abstract—In this paper we investigate fusion rules for distributed detection in large random clustered-wireless sensor networks (WSNs) with a three-tier hierarchy; the sensor nodes (SNs), the cluster heads (CHs) and the fusion center (FC). The CHs collect the SNs' local decisions and relay them to the FC that then fuses them to reach the ultimate decision. The SN-CH and the CH-FC channels suffer from additive white Gaussian noise (AWGN). In this context, we derive the optimal log-likelihood ratio (LLR) fusion rule, which turns out to be intractable. So, we develop a sub-optimal linear fusion rule (LFR) that weighs the cluster's data according to both its local detection performance and the quality of the communication channels. In order to implement it, we propose an approximate maximum likelihood based LFR (LFR-aML), which estimates the required parameters for the LFR. We also derive Gaussian-tail upper bounds for the detection and false alarms probabilities for the LFR. Furthermore, an optimal CH transmission power allocation strategy is developed by solving the Karush-Kuhn-Tucker (KKT) conditions for the related optimization problem. Extensive simulations show that the LFR attains a detection performance near to that of the optimal LLR and confirms the validity of the proposed upper bounds. Moreover, when compared to equal power allocation, simulations show that our proposed power allocation strategy achieves a significant power saving at the expense of a small reduction in the detection performance.

Index Terms—Wireless sensor networks (WSNs), cluster head (CH), distributed detection, decision fusion, power allocation strategy, KKT conditions.

I. INTRODUCTION

W IRELESS sensor networks (WSNs) are becoming a mainstream technology constituting the backbone of several emerging technologies, such as the internet of things (IoT) [1] and smart cities [2] (see references therein). Indeed, the flexible nature of WSNs [3] enables them to invade such a wide spectrum of applications. However, several aspects of WSNs remain fertile research grounds, especially distributed detection (DD) in WSNs [4]. In such a scenario, battery-powered sensor nodes (SNs) monitor the region of interest (ROI), which are geographically distributed in a vast region in

Sami A. Aldalahmeh and Saleh O. Al-Jazzar are with the Faculty of Engineering and Technology in the Al-Zaytoonah University of Jordan, Amman Jordan (email: s.aldalahmeh@zuj.edu.jo, saleh.g@zuj.edu.jo).

Des McLernon, Syed Ali Raza Zaidi and Mounir Ghogho are with the School of Electronic and Electrical Engineering in Leeds University, Leeds, UK (email: d.c.mclernon@leeds.ac.uk, s.a.zaidi@leeds.ac.uk, m.ghogho@leeds.ac.uk).

Mounir Ghogho is also with Department of Computer Science in Université Internationale de Rabat, Rabat, Morocco.

order to detect any intruders. The locations of the SNs are best modeled as a random point process [5], because they might be out of communication range, out of power or might be even dropped from an airplane to form a network [6]. Due to constrained power and bandwidth, the collected data is often compressed into a single bit decision. Moreover, because the communication range is limited, providing adequate coverage for the large number of SNs is a challaging task. So, the WSN is divided into geographical clusters [7] and hierarchically into three tiers; SNs, cluster heads (CHs) and the fusion center (FC). The SNs in each cluster send their data to the CH, which usually has more power and a larger communication range. The CHs in turn report the collected data to the FC, thus acting as high-power relays. Often data is relayed in an amplify-and-forward (AF) or decode-and-forward (DF) fashion over imperfect communication channels [8].

In this paper, we investigate the decision fusion for distributed detection in a randomly deployed clustered-WSN operating with constrained power and over imperfect channels. In particular, the channels between SNs and CHs (termed SN-CH) and the channels between CHs and the FC (termed CH-FC) are assumed to suffer from additive white Gaussian noise (AWGN), in contrast to our previous work in [9], which assumes ideal channels. To the best of the a authors' knowledge, this is the first work that studies fusion rules in the above problem setting.

In the light of the previous framework, the main contributions of this paper are:

1) The optimal log-likelihood ratio (LLR) rule is presented first, which is analytically difficult to implement. Subsequently, a sub-optimal linear fusion rule (LFR) is derived. Intuitively, the LFR gives more weight to clusters with better detection and channel qualities.
2) We propose the approximate maximum likelihood estimator based fusion rule (LFR-aML) as a practical implementation of the LFR. The LFR-aML estimates the statistical parameters required for the detection fusion rule by solving a constrained maximum likelihood (ML) problem via the aid of Karush-Kuhn-Tucker (KKT) conditions.
3) To quantify the performance of the LFR and its derivatives, we derive Gaussian-tail upper bounds for the detection and the false alarm probabilities.
4) The optimal CH's transmission power is found in a closed-form manner while still adhering to a specific detection performance. This is achieved by also solving the KKT conditions of the related convex optimization

problem.

The rest of the paper is organized as follows. In Section II related work is reviewed. The adopted notation is explained in Section III. The system model is presented in Section IV. The proposed fusion rules are discussed in Section V, in which we begin by formulating the optimal DD fusion rule for the noisy clustered WSN. Then, the LFR algorithm is derived in the light of the previous optimal rule. Finally, the LFR-aML is proposed as the practical implementation of the LFR. The power allocation for the LFR is investigated in Section VI. Section VII discusses the practical implementation procedure of the LFR-aML algorithm. Section VIII presents the simulation results and their discussions. Finally, the conclusions are given in Section IX.

II. RELATED WORK

Since the seminal work by Tenney and Sandell [10], distributed detection has become, and still is, a rich research topic see for example [11], the recent tutorial [12] and references therein. Fundamental results have been attained for parallel, tandem and tree sensor paradigm [13], emphasizing the optimality of fusing and quantizing the local log-likelihood ratios (LRTs) of the distributed sensors. Further extensions of the classical problem were presented in [14], such as weak signal detection and robust detection.

For the case of single-bit quantization over perfect parallel networks, the opimal Chair-Varshney fusion rule (CVR) was derived in [15], which implicitly requires knowledge of the target parameters (location and power). A generalized likelihood ratio (GLRT) detector was proposed in [16] where the target parameters are estimated. However, the GLRT is computationally demanding and so the suboptimal counting rule (CR) was proposed in [17], which is simply the sum of positive local detections. Its performance on the other hand, was investigated in [18]. The weighted decision fusion (WDF) is introduced in [19] as an improvement of the CR, where the decisions are weighted first before being fused at the FC. In a different direction, a detector based on the Rao test was suggested in [20] and the generalized Rao test in [21] that both strike a trade-off between complexity and performance. In a similar effort, the generalized locally optimum detector was devised in [22] and [23]. However, the previous detectors in general suffered from the problem of spurious detection[1], especially in a large WSN. Scan statistics-based detection was proposed in [24] to overcome the previous problem but at the expense of a significant delay due to the sliding-window structure of the detector. A local vote decision fusion rule (LVDF) was proposed in [25], in which sensors use neighbouring decisions to correct their decisions locally and then integrate them globally.

DD over multi-hop (tree) sensor networks has also received considerable attention. Asymptotic results were presented in [26] and [27] for DD with Neyman-Pearson and Bayesian criteria, respectively. Mainly, it was shown that the error probability decays exponentially as the number of SNs increases.

Decision fusion rules over multi-hop networks were investigated in [28] with flat-fading noisy channels. The relaying SNs decode-and-forward the received data to the FC. The derived suboptimal rules in [28] de-emphasizes sensors with more hops. Similarly, authors in [29] investigated the same problem but with a binary-symmetric channel (BSC) model in the network, where it was shown that the optimal fusion rule is a weighted order statistic filter.

Clustered sensor networks were introduced for DD in [30], in which sensors report to CHs that in turn report to a FC[2]. Majority-like fusion (MLF) rules were used on both the cluster level and the FC level. Surprisingly, results there show that the detection performance of a clustered sensor network is worse than the performance of sensors reporting directly to the FC. This is due to employing the MLF rule in the CHs level, thus introducing additional errors in the decision process.

On the other hand, in our previous work [9] we have derived an optimal-cluster-based fusion rule (OCR) for clustered sensor networks. In this context, CHs collect the local SNs decisions and send this data to the FC over ideal channels. In this paper however, we consider a two-hop network in which the SN-CH and CH-FC channels are affected by AWGN. The CHs employ an AF scheme to send the collected data to the FC, in contrast to [29] that adopts sending hard decisions over multiple-hop BSCs. Moreover, the adopted AF scheme enables us to minimize the transmission power, provided a specific detection performance is satisfied.

III. NOTATION

In this paper we will generally refer to deterministic values by lowercase symbols (x), bold symbols refers to vector values $(\mathbf{x})$, whereas random values are referred to by uppercase symbols (X). However, we denote the number of CHs as (M), global detection threshold as (Γ) and the transmitted power as (P). The operator $\succcurlyeq$ refers to element-wise greater than or equal to. The probability of an event A is denoted by $\mathbb{P}(A)$. A normal distributed random variable (RV) X with mean μ and variance σ^2 is denoted $X \sim \mathcal{N}(\mu, \sigma^2)$ and Poisson RV Y with mean λ is denoted by $Y \sim \text{Pois}(\lambda)$. The expectation operator with respect to (w.r.t.) RV X is written as $\mathbb{E}_X[\cdot]$ and the moment generating function (MGF) for RV X is defined as $M_X(t) = \mathbb{E}_X[\exp(tX)]$. Finally, the estimate of any variable x is denoted by $\widehat{x}$.

IV. SYSTEM MODEL

The WSN is functionally divided into three tiers as shown in Fig. 1. In this section we present the sensing model, the stochastic geometry model for the SNs deployment, similar to [5] and [17], and the communication model between the three tiers.

A. Sensing and Sensor Deployment Models

Consider a WSN deployed over a region, $\mathcal{A} \subset \mathbb{R}^2$ where $\mathcal{A}$ is assumed to be significantly large. The WSN is modeled by a

---

[1] Spurious detection is defined in [24] as the event when a target is present and some sensors far from it declare the presence of a target in their vicinity.

[2] In [30] the CHs are refered to as FCs and the FC is refered to as an access point (AP).
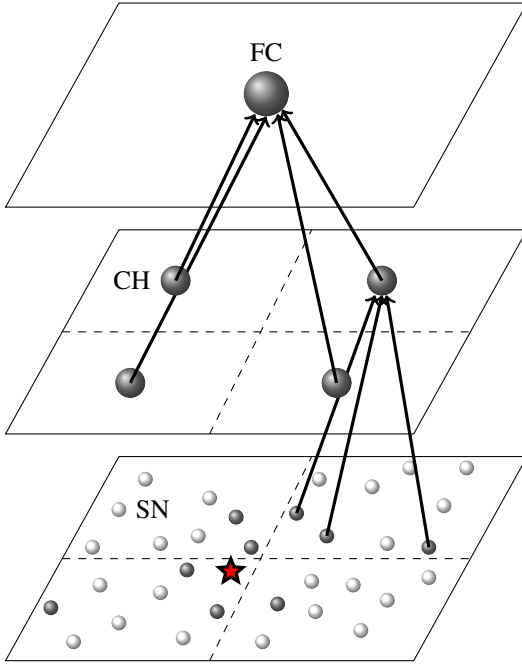
Fig. 1: The WSN topology, in which the star is the target, gray-shaded nodes are the detecting SNs and white-shaded nodes are the non-detecting SNs.

simple Poisson Point Process (PPP) $\Phi = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_N\}$ in $\mathcal{A}$ [31], where $\mathbf{X}_i \in \Phi$ is the the coordinate of the ith SN. This implies that the $\mathbf{X}_i$'s are random variables (RVs) and their number $N = |\Phi|$ is a Poisson RV having the distribution $N \sim \text{Pois}(\mathbb{E}[N])$ where $\mathbb{E}[N]$ is the average number of SNs. The SN intensity (λ) is defined as the average number of points (SNs) in a unit area. In general, the PPP might be non-homogeneous, i.e., the intensity is location dependent, described by $\lambda(\mathbf{x})$ where $\mathbf{x}$ is the location coordinates. This case might arise due to environmental or application specific constraints. However, using the non-homogeneous PPP model complicates the analysis. Thus we adopt a homogeneous PPP in our treatment ($\lambda(\mathbf{x}) = \lambda, \ \forall \mathbf{x} \in \mathcal{A}$), in which we can approximate the non-homogeneous case appropriately if the mth cluster is adequately small. In other words, the intensity in the mth cluster does not vary significantly with space, i.e. $\lambda_m(\mathbf{x}) \approx \lambda_m$ where λ_m is the mth cluster SN intensity. In the homogeneous case, the number of SNs follows the distribution $\text{Pois}(\lambda|\mathcal{A}|)$, where $|\mathcal{A}|$ is the area of $\mathcal{A}$. Fig. 1 shows a homogeneous random network deployment.

The WSN is tasked with the detection of any intruder or target entering the ROI. A target at location $\mathbf{X}_t \in \mathcal{A}$ leaves a signature signal sensed by the SNs, which might be thermal, magnetic, electrical, seismic or electromagnetic signal [32]. We adopt the sensing model in [18], in which the signature power in the far-field is assumed to decay quadratically with distance. The target's parameters are given in the vector $\boldsymbol{\theta} = [P_t, \mathbf{X}_t]^T$, where P_t is the target's signal power. The noise-free signal received at the ith SN located at a given $\mathbf{x}_i$ has the

following amplitude:

$$a(\mathbf{x}_i) = \frac{\sqrt{P_t}}{\max(d_0, d_i)} \quad (1)$$

where d_0 is the reference distance to the node's sensor and $d_i = \|\mathbf{x}_t - \mathbf{x}_i\|$ is the Euclidean distance between the target and the ith SN. Note that the measured signal is saturated if the distance to the target is smaller than d_0. The above model can adequately describe acoustic or electromagnetic signals.

For a given realization of Φ, each SN samples the environment to decide whether an intruder is present or not. Assuming conditional independence, the collected data at the ith SN under the null and alternative hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, takes the following form:

$$\mathcal{H}_1 : S(\mathbf{x}_i) = a(\mathbf{x}_i) + Q_i \quad (2)$$
$$\mathcal{H}_0 : S(\mathbf{x}_i) = Q_i \quad (3)$$

where Q_i is a white Gaussian noise at the ith SN with zero mean and variance σ_s^2. However, in practice the collected data are actually correlated [33]. The noise is assumed to be identically and independently distributed over all SNs and is not dependent on $\mathbf{x}_i$. The sensing SNR is defined as $\text{SNR}_s = P_t/\sigma_s^2$. Each SN computes its binary local decision, $I(\mathbf{x}_i) = \{0, 1\}$, by comparing the collected data with a local decision threshold τ, i.e.,

$$I(\mathbf{x}_i) = \begin{cases} 1, & g(S(\mathbf{x}_i)) \geq \tau \\ 0, & g(S(\mathbf{x}_i)) < \tau \end{cases} \quad (4)$$

where $g(\cdot)$ is the local detection function, e.g., matched filter or energy detector. Here, τ is the same for all SNs. Therefore, the local probabilities of false alarm and detection are given respectively by

$$P_{fa} = f_0(\tau, \sigma_s) \quad (5)$$
$$P_d(\mathbf{x}_i) = f_1(a(\mathbf{x}_i), \sigma_s, \tau) \quad (6)$$

where $f_0(\cdot)$ and $f_1(\cdot)$ are the false alarm and detection probabilities under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. These functions depend on the type of local detector used, such as a matched filter or an energy detector. Note, however, that the probability of detection in (6) also depends on the target parameters, P_t and $\mathbf{x}_t$ through (1).

B. Communication Model

Due to the large area of the ROI, the WSN is geographically divided into M disjoint zones: $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_M$, where $\mathcal{C}_m \in \mathcal{A}$ for $m = 1, \cdots, M$. For the sake of simplicity, the $\mathcal{C}_m$'s are assumed to be identical. Each zone is managed by a CH located at $\mathbf{x}_m \notin \Phi$. The number of clusters is fixed and their locations are also fixed and known to the WSN. SNs located at $\mathbf{x}_i \in \mathcal{C}_m$ send their decisions to the mth CH. The CHs in turn report the collected decisions back to the FC. The three-tier network is shown in Fig. 1.

Due to cost and bandwidth constraints, SNs use on-off keying (OOK) to transmit their binary local decisions to the CH over a shared multiple access (MAC) AWGN channel. These SNs transmit with the same power P_0 within the cluster

and are assumed to be synchronized to the same time slot. Hence, the received signal at the mth CH is

$$Y_m = \sqrt{P_0} \Lambda_m + W_m \quad (7)$$

where

$$\Lambda_m = \sum_{\mathbf{X}_i \in \mathcal{C}_m} I(\mathbf{X}_i), \ m = 1, \cdots, M \quad (8)$$

is the number of positive local decisions in the mth cluster and W_m is the AWGN at that CH with distribution of $\mathcal{N}\left(0, \sigma_{c,m}^2\right)$. Note that since Λ_m is actually the result of thinning of the PPP in the mth cluster, Λ_m is a Poisson RV distributed as [9]

$$\Lambda_m \sim \begin{cases} \text{Pois}(\lambda_{0,m}), & \mathcal{H}_0 \\ \text{Pois}(\lambda_{1,m}), & \mathcal{H}_1 \end{cases} \quad (9)$$

where $\lambda_{0,m}$ are $\lambda_{1,m}$ are the mean numbers of the detecting SNs (Λ_m) in the mth cluster under $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively and are given by

$$\lambda_{0,m} = \lambda P_{fa} |\mathcal{C}_m| \quad (10)$$
$$\lambda_{1,m} = \lambda \int_{\mathcal{C}_m} P_d(\mathbf{x}) d\mathbf{x}. \quad (11)$$

Note however that in the homogeneous case $\lambda_{0,m} = \lambda_0 \ \forall m$, since that $\mathcal{C}_m$'s are assumed to be identical.

However, in order to implement the models in (7) and (8), the CH controls the SNs' transmission power via a power control scheme. Further discussion is provided in Section VII. Each CH adopts the AF scheme to relay the gathered data to the FC over a dedicated AWGN wireless channel. The CHs are assumed to have more transmission power capabilities compared to the SNs. The received signals at the FC from the mth CH are

$$Z_m = \sqrt{P_m} Y_m + V_m, \ m = 1, \cdots, M \quad (12)$$

where P_m is the transmission power used by the mth CH and $V_m \sim \mathcal{N}\left(0, \sigma_{f,m}^2\right)$ is the AWGN receiver in the channel between the FC and the mth CH.

V. FUSION RULES IN CLUSTERED WSNs

In this section we present the fusion rules for clustered WSNs. The ideal channel case is presented first as a benchmark for comparison. Then we proceed to discuss fusion rules for noisy channels.

A. Decision Fusion in Ideal Channel Clustered WSN

For a clustered WSN with ideal communication channels, the majority-like fusion rule has been proposed in [30], in which the counting rule is implemented on the CH and the FC levels. However, since this rule showed a degraded performance in random WSNs so we will not discuss it further in this paper.

In [9], we proposed the OCR, in which CHs send the sum of the collected SNs' decisions, Λ_m, to the FC to be optimally fused, i.e.,

$$\Lambda_{\text{OCR}} = \sum_{m=1}^{M} c_m \Lambda_m \quad (13)$$

where the optimal weighing coefficient is $c_m = \log(\lambda_{1,m}/\lambda_{0,m})$. This weighing effectively suppresses the previous spurious detection problem, since the clusters containing these spurious decisions have small weighing coefficients. Note however, that $\lambda_{1,m}$ depends on the target's parameters, $\boldsymbol{\theta}$, through (11). Thus an exact implementation of (13) requires knowledge of $\boldsymbol{\theta}$. This problem has been circumvented in [9] by using a complexity-reduced GLRT, in which $\boldsymbol{\theta}$ is coarsely estimated.

It is interesting to note however, that the CR [17] is a special case of the OCR when there is only one global cluster encompassing the whole ROI. Thus the fusion rule in (13) reduces to

$$\Lambda_{\text{CR}} = \sum_{\mathbf{X}_i \in \Phi}^{N} I(\mathbf{X}_i). \quad (14)$$

The CR is also used as benchmark performance comparison of the fusion rules.

B. Optimal Decision Fusion in a Noisy Clustered WSN

In order to develop the optimal fusion rule in clustered WSNs with noisy channels, we investigate the received signals at the FC. By combining (7) and (12), the received signal is

$$Z_m = \sqrt{\widetilde{P}_m} \Lambda_m + \widetilde{V}_m \quad (15)$$

where $\widetilde{P}_m = P_m P_0$ and $\widetilde{V}_m = \sqrt{P_m} W_m + V_m$ is the aggregate noise at the mth CH-FC channel having a distribution of $\mathcal{N}\left(0, \widetilde{\sigma}_m^2\right)$ where $\widetilde{\sigma}_m^2 = P_m \sigma_{c,m}^2 + \sigma_{f,m}^2$. The likelihood-ratio-test (LRT) for the signal in (15) is

$$\Lambda_{\text{LRT}} = \prod_{m=1}^{M} \frac{p(z_m; \mathcal{H}_1)}{p(z_m; \mathcal{H}_0)} = \prod_{m=1}^{M} \frac{\mathbb{E}_{\Lambda_m}[p(z_m|\Lambda_m); \mathcal{H}_1]}{\mathbb{E}_{\Lambda_m}[p(z_m|\Lambda_m); \mathcal{H}_0]}$$
$$= \prod_{m=1}^{M} \frac{\mathbb{E}_{\Lambda_m}\left[\exp\left(-\frac{1}{2\widetilde{\sigma}_m^2}\left(z_m - \sqrt{\widetilde{P}_m}\Lambda_m\right)^2\right); \mathcal{H}_1\right]}{\mathbb{E}_{\Lambda_m}\left[\exp\left(-\frac{1}{2\widetilde{\sigma}_m^2}\left(z_m - \sqrt{\widetilde{P}_m}\Lambda_m\right)^2\right); \mathcal{H}_0\right]}. \quad (16)$$

Note that the expectations in the numerator and denominator are w.r.t. the distributions in (9). Therefore, $p(z_m; \mathcal{H}_1)$ is actually the convolution of the Poisson distribution of Λ_m and the Gaussian distribution of the noise leading to the fourth term in (16). Unfortunately, the corresponding log-likelihood ratio (LLR) is still not simpler:

$$\Lambda_{\text{LLR}} = \sum_{m=1}^{M} \log\left(\mathbb{E}_{\Lambda_m}\left[\exp\left(-\frac{s_m}{2}(\widetilde{z}_m - \Lambda_m)^2\right); \mathcal{H}_1\right]\right)$$
$$- \log\left(\mathbb{E}_{\Lambda_m}\left[\exp\left(-\frac{s_m}{2}(\widetilde{z}_m - \Lambda_m)^2\right); \mathcal{H}_0\right]\right) \quad (17)$$

where $\widetilde{z}_m = z_m/\sqrt{\widetilde{P}_m}$ and $s_m = \widetilde{P}_m/\widetilde{\sigma}_m^2$ is the mth CH-FC channel SNR.

C. Linear Decision Fusion Rule (LFR)

Although the fusion rule in (17) is optimal, unfortunately it is impractical and does not lend itself to analysis. In order to

derive a practical rule, we first need to find the distribution of Z_m. The MGF of Z_m in (15) is computed as:

$$
\begin{aligned}
M_{Z_m}(t) &= M_{\Lambda_m}\left(\sqrt{\widetilde{P}_m}t\right) M_{V_m}(t) \\
&= \exp\left(\lambda_{j,m}\left(e^{t\sqrt{\widetilde{P}_m}} - 1\right) + \frac{\widetilde{\sigma}_m^2}{2}t^2\right) \quad (18)
\end{aligned}
$$

where $j = 0, 1$ for $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. Unfortunately, the above MGF is also intractable. However, using a first order approximation of the exponential function ($e^x \approx 1 + x$) yields the following:

$$
M_{Z_m}(t) \approx \exp\left(\lambda_{j,m}t\sqrt{\widetilde{P}_m} + \frac{\widetilde{\sigma}_m^2}{2}t^2\right) \quad (19)
$$

which is the MGF of the Gaussian RV with $p(z_m) \sim \mathcal{N}\left(\lambda_{j,m}\sqrt{\widetilde{P}_m}, \widetilde{\sigma}_m^2\right)$. Therefore, we approximate the LLR in (17) as

$$
\begin{aligned}
\Lambda_{\text{LLR}} &\approx \sum_{m=1}^M \frac{1}{2\widetilde{\sigma}_m^2}\left(z_m - \lambda_{1,m}\sqrt{\widetilde{P}_m}\right)^2 \\
&\quad - \frac{1}{2\widetilde{\sigma}_m^2}\left(z_m - \lambda_{0,m}\sqrt{\widetilde{P}_m}\right)^2. \quad (20)
\end{aligned}
$$

Expanding the above and rearranging the terms gives

$$
\begin{aligned}
\Lambda_{\text{LLR}} &\approx \sum_{m=1}^M \frac{1}{2\widetilde{\sigma}_m^2}\left(z_m^2 - 2\lambda_{1,m}\sqrt{\widetilde{P}_m}z_m + \widetilde{P}_m\lambda_{1,m}^2\right) \\
&\quad - \sum_{m=1}^M \frac{1}{2\widetilde{\sigma}_m^2}\left(z_m^2 - 2\lambda_{0,m}\sqrt{\widetilde{P}_m}z_m + \widetilde{P}_m\lambda_{0,m}^2\right) \\
&= -\sum_{m=1}^M \sqrt{\widetilde{P}_m}\left(\frac{\lambda_{1,m} - \lambda_{0,m}}{\widetilde{\sigma}_m^2}\right)z_m \\
&\quad + \sum_{m=1}^M \widetilde{P}_m\left(\lambda_{1,m}^2 + \lambda_{0,m}^2\right). \quad (21)
\end{aligned}
$$

When comparing Λ_{LLR} with the detection threshold (Γ), the last term in (21) is absorbed by Γ since it is independent of z_m. So the resulting linear fusion rule (LFR) becomes:

$$
\Lambda_{\text{LFR}} = \sum_{m=1}^M d_m z_m \gtrless \Gamma, \quad (22)
$$

where the linear weighing coefficients are

$$
d_m = \frac{\sqrt{\widetilde{P}_m}}{\widetilde{\sigma}_m^2}\left(\lambda_{1,m} - \lambda_{0,m}\right). \quad (23)
$$

The LFR is essentially a weighted sum of the data provided by each cluster. The impact of each cluster is reflected by its weight d_m, which is a measure of the detection performance and the channel quality of that cluster.

Remark. *The LFR intuitively gives more weight to clusters with better detection, which is manifested in the mean difference term $(\lambda_{1,m} - \lambda_{0,m})$. Also, more weight is given to clusters with good channel quality, i.e., large $\sqrt{\widetilde{P}_m}/\widetilde{\sigma}_m^2$.*

Clearly the LFR is computationally simple, in contrast to the LLR in (17). In fact, its computational complexity amounts to $O(M)$ only.

D. Approximated Maximum Likelihood-Based LFR (LFR-aML)

Although the LFR is a analytically simple, its implementation requires $\lambda_{1,m}$'s to be known by the FC. Hence, we extend the LFR in this section to include an estimation phase prior to the detection phase.

1) Estimation phase: At first glance, the $\lambda_{1,m}$'s can be computed by initially estimating $\boldsymbol{\theta}$ through maximum likelihood estimation from the log-likelihood of Z_m's under $\mathcal{H}_1$ (the first term in (17)). However, such an estimation problem is also nonlinear and nonconvex, leading to high computational complexity. So we propose estimating $\lambda_{1,m}$'s directly. Still, attempting to do that from the current log-likelihood expression will not provide satisfactory results since each CH provides a single data point about Z_m, consequently, several instances of Z_m are required. Thus we extend the LFR to the multiple-sample case, in which each SN makes L independent decisions, $\{I_{i,l}\}_{l=0}^{L-1}$, that are relayed to the FC by the CHs. Further details of the implementation are provided in Section VII.

Given the set of collected data $\widetilde{z}_{l,m}$'s, which is the lth sample from the mth CH, then the corresponding likelihood function is $\prod_{m=1}^M \prod_{l=0}^{L-1} p(\widetilde{z}_{l,m}; \lambda_{1,m})$. It follows directly that the constrained ML problem using the related log-likelihood is

$$
\begin{aligned}
&\max_{\lambda_{1,m}} \sum_{m=1}^M \sum_{l=0}^{L-1} \log\left(p\left(\widetilde{z}_{l,m}; \lambda_{1,m}\right)\right) \quad (24) \\
&\text{s.t.} \quad \lambda_{1,m} \geq \lambda_0 \quad \forall m.
\end{aligned}
$$

Even though the ML problem above is separable in $\lambda_{1,m}$, it is still complicated. Therefore, we propose to solve a suboptimal version of (24) by using the corresponding lower bound provided by the following lemma:

Lemma 1. *The lower bound of the log-likelihood function in (24) is given as*

$$
\log\left(p\left(\widetilde{z}_{l,m}; \lambda_{1,m}\right)\right) \geq \widehat{\Lambda}_{l,m} \log \lambda_{1,m} - \lambda_{1,m} + C_1 \quad (25)
$$

where

$$
\widehat{\Lambda}_{l,m} = \frac{\sum_{k=0}^{\infty} k p\left(\widetilde{z}_{l,m}|k\right)}{\sum_{k=0}^{\infty} p\left(\widetilde{z}_{l,m}|k\right)} \quad (26)
$$

is the mean estimate of the lth received sample from the mth CH and $p(\widetilde{z}_{l,m}|k)$ is the conditional Gaussian distribution given k and

$$
C_1 = \log C_0 - \sum_{k=0}^{\infty} \pi_k \log k! \quad (27)
$$

where $C_0 = \sum_{k=0}^{\infty} p(\widetilde{z}_{l,m}|k)$ and $\pi_k = p(\widetilde{z}_{l,m}|k)/C_0$.

Proof. See Appendix A. □

Consequently, we have the surrogate optimization problem:

$$
\begin{aligned}
&\max_{\lambda_{1,m}} \sum_{m=1}^M \sum_{l=1}^L \left(\widehat{\Lambda}_{l,m} \log \lambda_{1,m} - \lambda_{1,m}\right) \quad (28) \\
&\text{s.t.} \quad \lambda_{1,m} \geq \lambda_0 \quad \forall m.
\end{aligned}
$$

The optimal solution is given in the following lemma.

Lemma 2. *The optimal solution of the constrained optimization problem in* (28) *is*

$$\widehat{\lambda}_{1,m} = \begin{cases} \widehat{\Lambda}_m, & \eta_m < 0 \\ \lambda_0, & \eta_m = 0 \end{cases} \tag{29}$$

$$\eta_m = 1 - \widehat{\Lambda}_m/\lambda_0 \tag{30}$$

where

$$\widehat{\Lambda}_m = \frac{1}{L} \sum_{l=0}^{L-1} \widehat{\Lambda}_{l,m} \tag{31}$$

is the average of the mth CH's estimates $\widehat{\Lambda}_{l,m}$ *given in* (26).

Proof. See Appendix B. □

The estimator $\widehat{\lambda}_{1,m}$ in (29) is intuitive in the sense that it is the average of all sample estimates when a target is sensed and is simply $\widehat{\lambda}_{1,m} = \lambda_0$ otherwise.

2) Detection phase: Given the estimates $\widehat{\lambda}_{1,m}$'s, we then propose the LFR-aML detector as

$$\Lambda_{\text{aML}} = \sum_{m=1}^{M} \widehat{d}_m \widehat{z}_m \gtrless \Gamma \tag{32}$$

where the weighing coefficients now are

$$\widehat{d}_m = \frac{\sqrt{\widetilde{P}_m}}{\widetilde{\sigma}_m^2} \left(\widehat{\lambda}_{1,m} - \lambda_0 \right) \tag{33}$$

and the $\widehat{z}_m$ is the averaged data for each CH defined as

$$\widehat{z}_m = \frac{1}{L} \sum_{l=1}^{L} \widetilde{z}_{l,m}. \tag{34}$$

The computational complexity of the LFR-aML, on the other hand, is $O(LM)$, which is relatively larger than that for the LFR.

E. Linear Fusion Rule Performance

Despite being a practical fusion rule, the LFR's detection and false alarm probabilities do not have closed forms that lend themselves to analysis. Thus we resort to finding the upper bound for those tail probabilities in the following theorem.

Theorem. *The upper bound for tail probability for the LFR is*

$$\mathbb{P}\left(\Lambda_{LFR} > z; \mathcal{H}_j\right) \leq \exp\left(-\frac{\left(z - \overline{\lambda}_{j,d}\right)^2}{2\widetilde{\sigma}_{j,d}^2} \right) \tag{35}$$

for $j = 0, 1$, *where*

$$\overline{\lambda}_{j,d} = \sum_{m=1}^{M} \lambda_{j,m} d_m \sqrt{\widetilde{P}_m} \tag{36}$$

$$\widetilde{\sigma}_{j,d}^2 = \sum_{m=1}^{M} d_m^2 \left(\lambda_{j,m} \widetilde{P}_m + \widetilde{\sigma}_m^2 \right). \tag{37}$$

Proof. See Appendix C. □

Clearly, the upper bound given above is a Gaussian-tail bound. It is interesting to note that the mean defined in (36) is actually a scaled version of the Λ_{LFR} mean defined in (22). Thus it is expected that if the LFR value is increased the detection probability will significantly improve.

VI. POWER ALLOCATION

WSNs are notorious for being power constrained. Hence, the power should be used wisely, especially if the application is critical such as intruder detection,. It is well established that the main source of power usage in a WSN is wireless communication [34]. Thus it is desired to minimize the transmission power used by SNs and CHs while jointly taking into consideration the minimum required detection performance.

Fortunately, the linear LFR structure facilitates the use of power allocation strategy, in contrast to the LLR. The mean-difference (MD) [35] is adopted as the detection performance criteria due to its desirable form. The MD is defined as

$$\begin{aligned} \text{MD} &= \overline{\lambda}_{d,1} - \overline{\lambda}_{d,0} \\ &= \sum_{m=1}^{M} d_m \sqrt{\widetilde{P}_m} \left(\lambda_{1,m} - \lambda_0 \right) \\ &= \sum_{m=1}^{M} \frac{P_0 P_m \left(\lambda_{1,m} - \lambda_0 \right)^2}{\sigma_{c,m}^2 P_m + \sigma_{f,m}^2}. \end{aligned} \tag{38}$$

In general the MD and the detection performance can be significantly improved by increasing the SNs deployment density, λ, without the need to increase the transmission power. So, attention should be directed at reducing the transmission power given a specific detection performance constraint[3]. Thus we wish to solve the following optimization problem:

$$\begin{aligned} \min_{\mathbf{P}} \quad & \|\mathbf{P}\|_1 \tag{39} \\ \text{s.t.} \quad & \mathbf{P} \succcurlyeq \mathbf{0} \\ & \text{MD} = P_0 \sum_{m=1}^{M} \frac{P_m \left(\lambda_{1,m} - \lambda_0 \right)^2}{\sigma_{c,m}^2 P_m + \sigma_{f,m}^2} > D_0 \end{aligned}$$

where $\mathbf{P} = (P_1, P_2, \cdots, P_M)^T$ is the lumped CH's transmission powers vector and D_0 is the minimum mean difference as specified by the network. The l_1-norm is adopted since it reduces the residuals leading to smaller component values in $\mathbf{P}$ and hence less transmission power.

Note however, that the above power allocation problem (39) is very similar to the water-filling problem [36], but here the power sum is minimized in contrast to the conventional water-filling formulation. Obviously, the above problem is convex, since the objective function is convex and the constraint is a linear-fractional function, which is also convex [36]. However, the l_1-norm is not differentiable and consequently a closed form solution cannot be attained. So we replace the $\|\mathbf{P}\|_1$ by the summation of the elements of $\mathbf{P}$'s (where each element is non-negative).

---

[3]Although P_0 is responsible for a large part of the used power, the P_m's on the other hand, play a more critical role in the WSN since if any CH runs out of power a significant part of the network is rendered useless

Theorem. *The optimal power allocation based on the formulation in (39) is given as*

$$P_m = \left(\frac{\lambda_{d,m} \sigma_{f,m} \sqrt{\nu}}{\sigma_{c,m}^2} - \frac{\sigma_{f,m}^2}{\sigma_{c,m}^2} \right)^+ \tag{40}$$

where $(x)^+ = \max(0, x)$ and

$$\sqrt{\nu} = \frac{\displaystyle\sum_{m=1}^{M} \frac{\lambda_{d,m} \sigma_{f,m}}{\sigma_{c,m}^2}}{\displaystyle\sum_{m=1}^{M} \frac{\lambda_{d,m}^2}{\sigma_{c,m}^2} - D_1}. \tag{41}$$

Proof. See Appendix D. $\square$

Intuitively, the allocated power is proportional to the cluster's detection performance manifested in the mean difference $\lambda_{d,m}$ and is inversely proportional to the SN-CH channel noise, $\sigma_{c,m}^2$. Of course, when using the power allocation algorithm practically, $\widehat{\lambda}_{1,m}$ is used to compute $\lambda_{d,m}$.

In this work, we will denote the LFR with power allocation strategy as LFR-PA, whereas the LFR using equal power allocation is just denoted by LFR. Similarly, the fusion rule using the estimates in (30) and (29) is called as LFR-aML.

VII. PRACTICAL CONSIDERATIONS

In this section we discuss how the LFR-aML algorithm is implemented in practice.

The LFR-aML is preceded by an initialization stage where the communication parameters are estimated. In order to ensure a constant received power at the CHs, thus validating the model in (8), a simple power control scheme is implemented. The CHs send pilot signals to the SNs in the cluster to be used to adapt the SNs' transmission power accordingly. Moreover, the CHs compute the SN-CH channel noise variances, $\sigma_{c,m}^2$, and likewise the FC computes the CH-FC channel noise variances, $\sigma_{f,m}^2$. Finally, λ_0 can be estimated off-line.

Then the LFR-aML is initiated where it performs estimation of the clusters' average number of detecting SNs, global distributed detection and optimal CH power allocation. Algorithm 1 illustrates the complete LFR-aML detection procedure.

VIII. SIMULATION RESULTS AND DISCUSSION

We simulate a WSN deployed in a 50×50 ROI. The intruder's power is $P_0 = 1$ located arbitrarily at $(4, 5)$. The sensing SNR is set to 0 dB. The SNs have a reference distance of $d_0 = 1$ units with a local probability of false alarm of 10^{-2}. The SNs adopt the matched filter as their local detector. The WSN has a SN deployment density of $\lambda = 2$ per unit area and is divided into 9 clusters. The system is simulated for 10^5 Monte Carlo iterations. Note however, that the simulation setting here is arbitrary but these results also hold for different scenarios.

Fig. 2 shows the effect of the SNRs of the SN-CH and the CH-FC channels on the detection performance through the ROC graphs. The OCR and CR are included as upper and lower bounds for the proposed algorithms. The case of

---

Algorithm 1 : LFR-aML

Initialization:
1: CHs send pilot signals to SNs.
2: SNs use pilot signal to adjust P_0.
3: CHs estimate the $\{\sigma_{c,m}^2\}_{m=1}^M$.
4: FC estimate the $\{\sigma_{f,m}^2\}_{m=1}^M$.
5: FC estimates λ_0 via (10), where $\lambda_{0,m} = \lambda_0 \, \forall m$.

LFR-aML:
1: $P_m = P_{tot}/M$.
2: **loop**
    Estimation:
3:     SNs compute $\{I_{i,l}\}_{l=0}^{L-1}$ via (4).
4:     SNs send $\{I_{i,l}\}_{l=0}^{L-1}$ to the CHs.
5:     CHs compute $\{\lambda_{1,m}\}_{m=1}^M$ via (29), (30) and (31).
6:     CHs compute $\{\widehat{z}_m\}_{m=1}^M$ via (34).
7:     CHs use P_m to send $\{\widehat{\lambda}_{1,m}, \widehat{z}_m\}_{m=1}^M$ to FC.
    Detection:
8:     FC computes Λ_{aML} via (32), (33) and (34).
9:     FC tests condition ($\Lambda_{\text{aML}} \gtrless \Gamma$) for global detection.
    Power Allocation:
10:     FC computes P_m's via (40) and (41).
11:     FC sends P_m's to CHs.
12: **end loop**

---

having equal high SNR for all SN-CH and CH-FC is shown in Fig. 2(a), in which the LLR and the LFR (employing equal power allocation, $P_m = 1 \, \forall m$) achieve the optimal performance provided by the OCR. By contrast, Fig. 2(b) illustrates the case of having equal low SNR for all the clusters' channels. Here, the LLR rule performance is as good as that of the CR, whereas the LFR rule performs worse than the latter. This behaviour is explained by the direct dependence on the SNR in the weighing coefficients, in (23). Fig. 2(c) shows the case of having better SN-CH channels compared to the CH-FC channels, while Fig. 2(d) shows the opposite case. The LFR performance virtually does not change whereas the LLR slightly degrades in the latter case. The LFR behaviour is explained by noting that $\widetilde{\sigma}_m^2$ is the weighted sum of $\sigma_{f,m}^2$ and $\sigma_{c,m}^2$. The LLR sensitivity might be attributed to its nonlinear form.

Next, Fig. 3 shows the effect of having a good channel quality specifically in the cluster containing the target. Interestingly, the results here show that it is sufficient to have a good SN-CH channel quality in the cluster containing the target to achieve good detection performance as is evident in Fig. 3(a) and Fig. 3(c). In a similar manner, a bad channel will significantly decrease the performance as can be seen in Fig. 3(b) and Fig. 3(d). Fig. 4 depicts the upper bounds for the P_{FA} and the P_D for the LFR presented in (35).

Then the performance of the LFR-aML is compared with the LLR, LFR and the CR in Figure 5. To make the comparison fair, all the rules use the received data average value in (34). In Figure 5a, the LFR-aML shows a satisfactory performance despite some loss of performance due to the inaccurate $\lambda_{1,m}$ estimation. On the other hand, Figure 5b shows the power allocation via (40) over the CHs. Notice that with exact $\lambda_{1,m}$
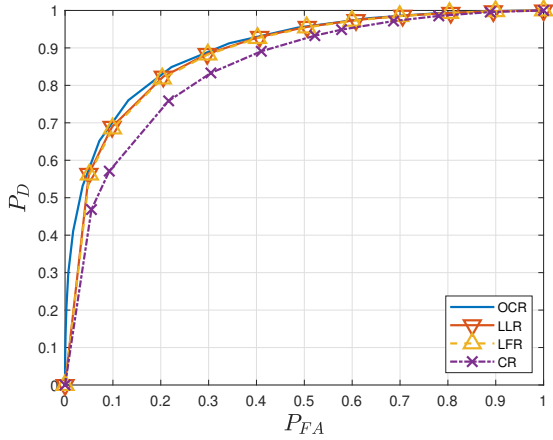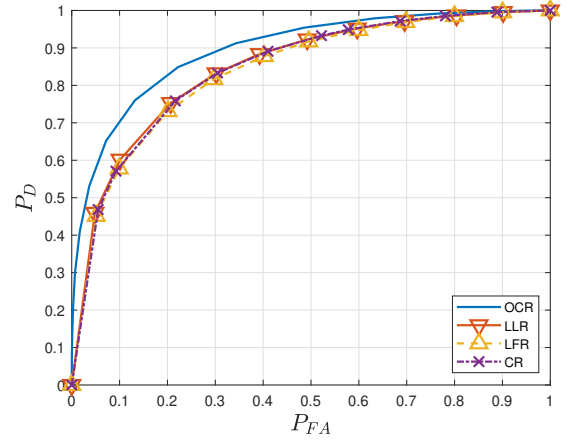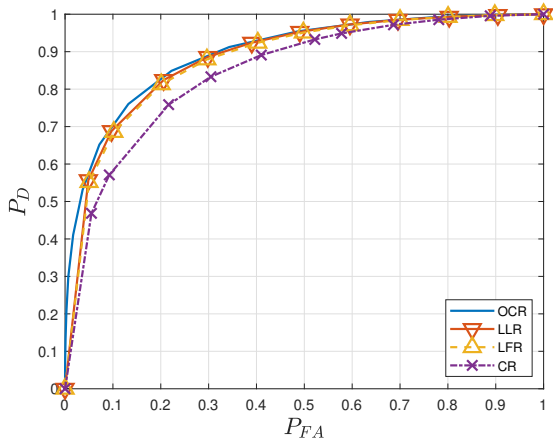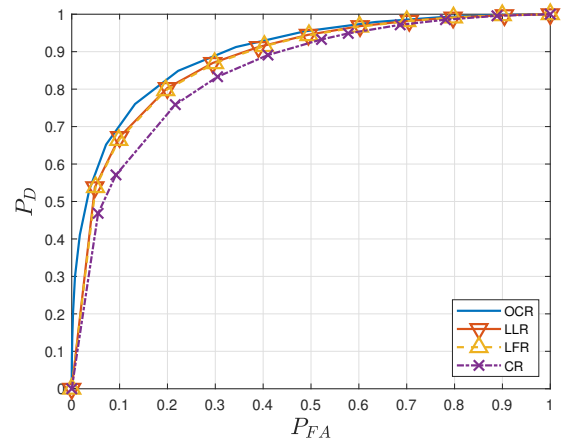
(a) $\text{SNR}_{f,m} = \text{SNR}_{c,m} = 5, \text{dB}, \forall m.$

(b) $\text{SNR}_{f,m} = \text{SNR}_{c,m} = -5\,\text{dB}, \forall m.$

(c) $\text{SNR}_{f,m} = 0\,\text{dB} < \text{SNR}_{c,m} = 5\,\text{dB}, \forall m.$

(d) $\text{SNR}_{f,m} = 5\,\text{dB} > \text{SNR}_{c,m} = 0\,\text{dB}, \forall m.$

Fig. 2: Effect of SN-CH and CH-FC channel SNRs on the detection performance. The SN transmission power is, $P_0 = 1$ and the CH transmission power is $P_m = 1, \forall m.$

values, the power is concentrated in the cluster containing the target. Other clusters may receive no power allocation at all. However, using estimated values leads to spreading the power across the cluster, while giving more power to the target's cluster.

Figure 6a shows P_D using the LFR-PA plotted with respect to D_1, which is the detection performance constraint. The P_D values are fitted with a third degree polynomial to emphasize the trend in a better way. In Figure 6b, the power saving using the LFR-PA rule is plotted against D_1 for different WSN conditions. The power saving is defined as

$$P_{\text{sav}} = \frac{1}{P_{tot}} \sum_{m=1}^{M} (P_{tot} - P_m) \times 100\% \qquad (42)$$

where P_{tot} is the power used by the LFR rule. The OCR, LFR, and CR are plotted for the sake of comparison. P_D and P_{sav} are plotted against D_1. It is clear that as D_1 increases, improved detection performance is achieved (in the range of 8%) at the expense of less saved energy. However, the power saving is significantly greater in the case of a better FC-CH

channel as shown. Indeed, at $D_1 = 5.5$ the power saving is 84% for the case of $\text{SNR}_{f,m} = 5\text{dB}$ and $\text{SNR}_{c,m} = 2\text{dB}$. Whereas it is 67% for both cases $\text{SNR}_{f,m} = \text{SNR}_{c,m} = 5\text{dB}$ and $\text{SNR}_{f,m} = \text{SNR}_{c,m} = 2\text{dB}$. Furthermore, the performance of the equal power version is attained with 64% power saving. This follows from the direct proportionality of P_m with $\sigma_{f,m}^2$ in (40).

IX. CONCLUSION

In this paper we have discussed fusion rules for DD in clustered-WSNs with communication channels experiencing AWGN. We have derived the optimal log-likelihood ratio fusion rule that turned out to be analytically intractable. A suboptimal linear fusion rule is subsequently derived, in which the cluster's data is linearly weighed. This rule, intuitively, gives more weight to clusters with better channels and better data quality. However, the LFR requires the knowledge of the mean value of detecting SNs. Thus we proposed the LFR-aML that employs an approximate constrained ML estimator to find the required parameters.
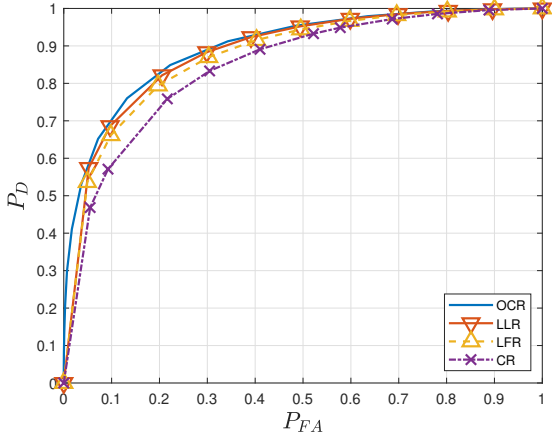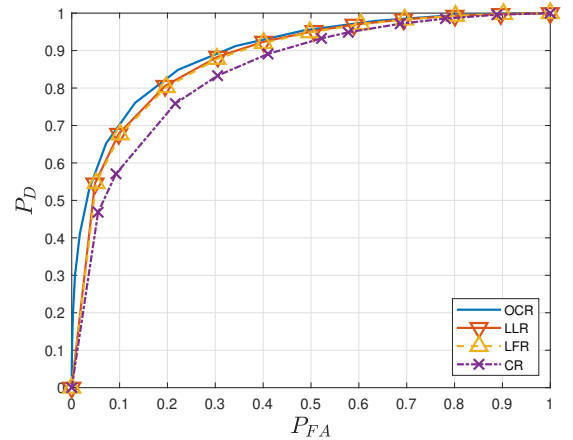
(a) Cluster containing target has $\mathrm{SNR}_{c,m} = 5\,\mathrm{dB}$, rest of cluster have $\mathrm{SNR}_{f,m} = \mathrm{SNR}_{c,m} = 2,\mathrm{dB}$.
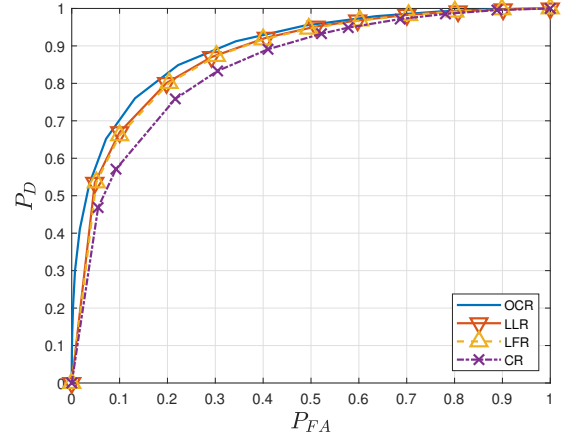
(b) Cluster containing target has $\mathrm{SNR}_{c,m} = -5\,\mathrm{dB}$, rest of cluster have $\mathrm{SNR}_{f,m} = \mathrm{SNR}_{c,m} = 2,\mathrm{dB}$.
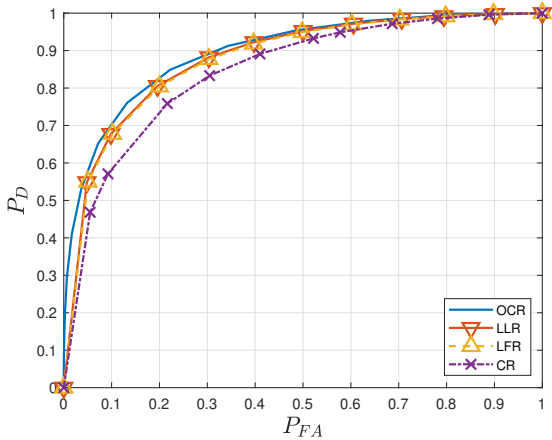
(c) Cluster containing target has $\mathrm{SNR}_{f,m} = 5\,\mathrm{dB}$, rest of cluster have $\mathrm{SNR}_{f,m} = \mathrm{SNR}_{c,m} = 2,\mathrm{dB}$.

(d) Cluster containing target has $\mathrm{SNR}_{f,m} = -5\,\mathrm{dB}$, rest of cluster have $\mathrm{SNR}_{f,m} = \mathrm{SNR}_{c,m} = 2,\mathrm{dB}$.

Fig. 3: Effect of SN-CH and CH-FC channel SNRs on the detection performance. The SN transmission power is $P_0 = 1$ and the CH transmission power is $P_m = 1$, $\forall m$.

In addition, Gaussian-tail upper bounds for the LFR's detection and false probabilities are derived using approximated moment generating functions. Moreover, a power allocation strategy is proposed that minimizes the total power used by the WSN. The resulting allocated power is proportional to the expected number of detecting SNs in the cluster.

Extensive simulations show that in order to achieve near optimal detection performance, it is sufficient to have a good channel quality for the cluster(s) containing the detecting SNs and moderate quality in the rest of the clusters. However, when using the LFR-aML, there is a performance gap with the ideal LFR due to the inherent estimation errors. It has been shown that the proposed power allocation strategy can achieve 84% power saving with only 5% performance reduction compared to the equal power scheme. Furthermore, the same detection performance as the equal power allocation version can be attained with a 14% power saving.

So in summary, for the first time a fusion rule has been derived for clustered WSN distributed detection with imperfect channels, and we have saved significant power usage with only a small reduction in detection performance. Several directions for future work can be pursued, including non-homogeneous PPP model and other types of channel imperfections such as path-loss, fading and channel failure. Furthermore, fusion rules for distributed detection can be investigated in which a decode-and-forward scheme (quantization) is employed at the CH.

APPENDIX A
PROOF OF LEMMA (1)

Recall that the likelihood function of the lth received sample from the mth cluster under $\mathcal{H}_1$ is actually the expectation of the conditional distribution $p\left(\widetilde{z}_{l,m}; \lambda_{1,m}\right) = \mathbb{E}_{\Lambda_m}\left[p\left(\widetilde{z}_{l,m} | \Lambda_m\right)\right]$. Since Λ_m is a Poisson RV, then the previous expectation becomes as follows:

$$p\left(\widetilde{z}_{l,m}; \lambda_{1,m}\right) = \sum_{k=0}^{\infty} p\left(\widetilde{z}_{l,m} | k\right) p\left(k; \lambda_{1,m}\right) \qquad (43)$$
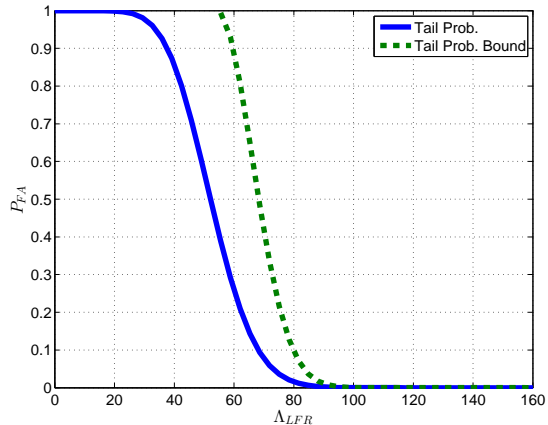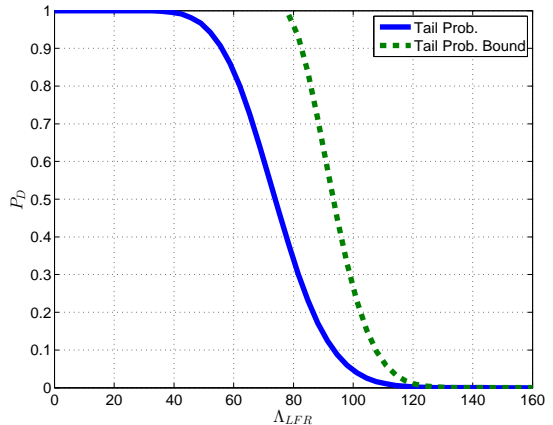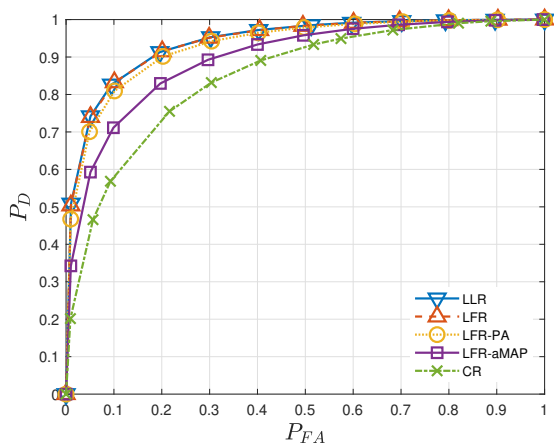
(a) P_{FA} Bound.

(b) P_D Bound.

Fig. 4: Tail Bounds for false alarm probability and detection probability. The SNs transmission power is $P_0 = 1$ and the CHs transmission power is $P_m = 1, \forall m$.
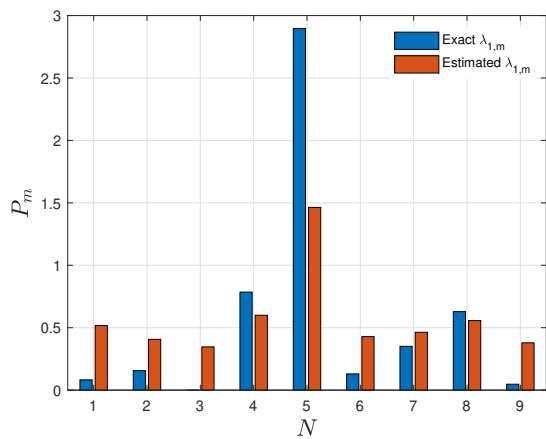

(a) ROC.

(b) Power used P_m by CHs.

Fig. 5: The effect of using the aML estimator on the ROC and the power used at $P_0 = 2$, data samples of $L = 5$, $\text{SNR}_{f,m} = 2$ and $\text{SNR}_{c,m} = 2, \text{dB}, \forall m$.

where Λ_m is replaced by k in order to simplify the notation. Note that (43) is a convex sum in terms of $p(\widetilde{z}_{l,m}|\Lambda_m)$ since $p(k; \lambda_{1,m})$ is a proper (Poisson) distribution[4]. A lower bound of $\log(p(\widetilde{z}_{l,m}; \lambda_{1,m}))$ can be attained via Jensen's inequality, however it is not useful in its current form. Instead, we proceed by treating the sum in (43) as a convex combination of $p(k; \lambda_{1,m})$ where $p(\widetilde{z}_{l,m}|k)$ are the weighing coefficients. But first, we need to normalize those coefficients as follows:

$$p(\widetilde{z}_{l,m}; \lambda_{1,m}) = C_0 \sum_{k=0}^{\infty} \pi_k p(k; \lambda_{1,m}) \qquad (44)$$

where $C_0 = \sum_{k=0}^{\infty} p(\widetilde{z}_{l,m}|k)$ and $\pi_k = p(\widetilde{z}_{l,m}|k)/C_0$ is a discrete distribution. Now we are ready to apply Jensen's

[4]For any probability mass function all the probabilities must be less than one and also sum to unity.

inequality on the log of (43) yielding

$$\log(p(\widetilde{z}_{l,m}; \lambda_{1,m})) \geq \sum_{k=0}^{\infty} \pi_k \log p(k; \lambda_{1,m}) + \log C_0$$

$$= \sum_{k=0}^{\infty} k\pi_k \log \lambda_{1,m} - \lambda_{1,m}\pi_k + C_1$$

$$= \widehat{\Lambda}_{l,m} \log \lambda_{1,m} - \lambda_{1,m} + C_1 \qquad (45)$$

where the second line above follows from the definition of the Poisson distribution and C_1 is a constant including the terms independent of $\lambda_{1,m}$. The last line results from the definition of the $\widehat{\Lambda}_{l,m}$ in (26) and the fact that π_k's sum up to unity.

APPENDIX B
PROOF OF LEMMA (2)

First, problem (28) is put in the following canonical form:
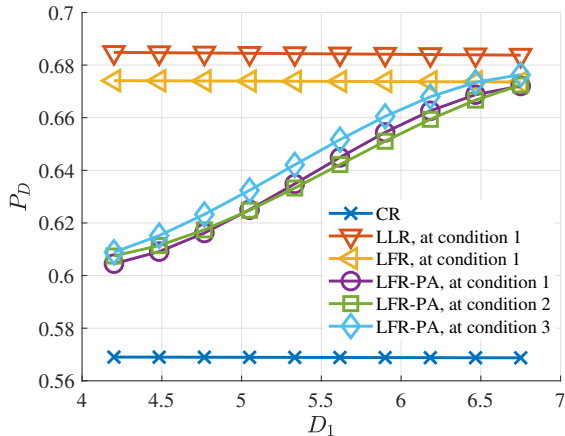
(a) P_D at $P_{FA} = 0.1$.
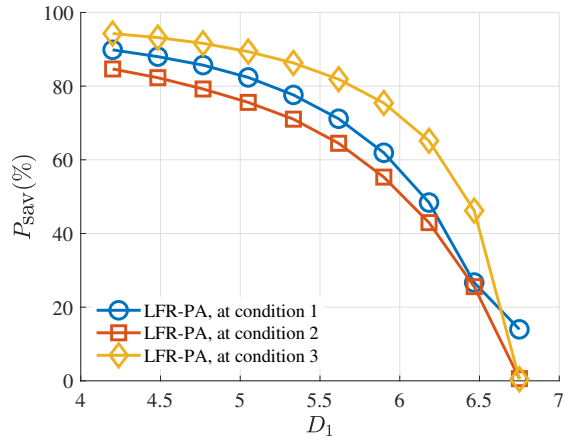
(b) Saved power percentage compared to the LFR rule.

Fig. 6: Detection performance and the power saving percentage achieved using the LFR-PA rule at $P_0 = 5$ and $P_m = 2$ and the following conditions: (1) $\text{SNR}_{f,m} = \text{SNR}_{c,m} = 2\text{dB}$, (2) $\text{SNR}_{f,m} = 2$ and $\text{SNR}_{c,m} = 5\text{dB}$ and (3) $\text{SNR}_{f,m} = 5$ and $\text{SNR}_{c,m} = 2\text{dB}$.

$$\min_{\lambda_{1,m}} \sum_{m=1}^{M} \left(\lambda_{1,m} - \widehat{\Lambda}_m \log \lambda_{1,m} \right) \tag{46}$$
$$\text{s.t.} \quad \lambda_0 - \lambda_{1,m} \leq 0 \quad \forall m.$$

where the problem (46) is scaled by $1/L$, and $\widehat{\Lambda}_m$ is defined in (31). The corresponding Lagrangian is

$$\mathcal{L} = \sum_{m=1}^{M} \lambda_{1,m} - \widehat{\Lambda}_m \log \lambda_{1,m} + \eta_m \left(\lambda_0 - \lambda_{1,m} \right) \tag{47}$$

where $\eta_m \geq 0 \ \forall m$ are the slack variables. The corresponding KKT conditions [36] are

$$\frac{\partial \mathcal{L}}{\partial \lambda_{1,m}} = 1 - \frac{\widehat{\Lambda}_m}{\lambda_{1,m}} - \eta_m = 0 \tag{48}$$

for all m and the slack conditions are

$$\eta_m \left(\lambda_{1,m} - \lambda_0 \right) = 0 \tag{49}$$

with $\eta_m \geq 0$. The complementary slack conditions in (49) dictate that $\lambda_{1,m} = \lambda_0$ when $\eta_m > 0$. Solving for the latter yields $\eta_m = 1 - \widehat{\Lambda}_m / \lambda_0$. Alternatively, when $\eta_m = 0$, that implies $\lambda_{1,m} = \widehat{\Lambda}_m$.

APPENDIX C
PROOF OF THEOREM (V-E)

The MGF of Λ_{LFR} in (22) under either hypotheses is given by the conditional independence as

$$M_{\text{LFR}} = \prod_{m=1}^{M} M_{Z_m} \left(t d_m; \mathcal{H}_j \right)$$
$$= \prod_{m=1}^{M} M_{\Lambda_m} \left(t d_m \sqrt{\widetilde{P}_m}; \mathcal{H}_j \right) M_{V_m} (t d_m). \tag{50}$$

From the MGFs of the Gaussian and Poison distributions we have

$$M_{\text{LFR}} = \exp \left(\sum_{m=1}^{M} \lambda_{j,m} \left(e^{t d_m \sqrt{\widetilde{P}_m}} - 1 \right) + \frac{d_m^2 \widetilde{\sigma}_m^2 t^2}{2} \right). \tag{51}$$

Using the second-order Taylor series for the exponential function the MGF becomes

$$M_{\text{LFR}} \approx \exp \left(t \sum_{m=1}^{M} \lambda_{j,m} d_m \sqrt{\widetilde{P}_m} + \frac{t^2}{2} d_m^2 \left(\lambda_{j,m} \widetilde{P}_m + \widetilde{\sigma}_m^2 \right) \right)$$
$$= \exp \left(\overline{\lambda}_{j,d} t + \frac{\widetilde{\sigma}_{j,d}^2}{2} t^2 \right), \tag{52}$$

where $\overline{\lambda}_{j,d}$ and $\widetilde{\sigma}_{j,d}^2$ are defined in (36) and (37) respectively. The Chernoff bound for Λ_{LFR} is

$$\mathbb{P} \left(\Lambda_{\text{LFR}} > z \right) < \inf_{t>0} \exp \left(-zt \right) M_{\text{LFR}}(t)$$
$$= \inf_{t>0} \exp \left(-zt + \overline{\lambda}_{j,d} t + \frac{\widetilde{\sigma}_{j,d}^2}{2} t^2 \right). \tag{53}$$

The infimum is found by taking the derivative of the exponential argument, equating to zero and solving for t. This gives the upper bound

$$\mathbb{P} \left(\Lambda_{\text{LFR}} > z; \mathcal{H}_j \right) < \exp \left(-\frac{\left(z - \overline{\lambda}_{j,d} \right)^2}{2 \widetilde{\sigma}_{j,d}^2} \right). \tag{54}$$

APPENDIX D
PROOF OF THEOREM (VI)

The Lagrangian of canonical optimization problem is

$$\mathcal{L} (P_m, \mu_m) = \sum_{m=1}^{M} P_m - \sum_{m=1}^{M} \mu_m P_m$$
$$+ \nu \left(D_1 - \sum_{m=1}^{M} \frac{P_m \lambda_{d,m}^2}{\sigma_{c,m}^2 P_m + \sigma_{f,m}^2} \right) \tag{55}$$

where $\lambda_{d,m} = (\lambda_{1,m} - \lambda_0)$ and $D_1 = D_0/P_0$. Then the related KKT conditions are

$$1 - \mu_m - \frac{\nu\lambda_{d,m}^2\sigma_{f,m}^2}{\left(\sigma_{c,m}^2 P_m + \sigma_{f,m}^2\right)^2} = 0 \tag{56}$$

$$\mu_m P_m = 0 \tag{57}$$

$$\nu\left(\sum_{m=1}^{M}\frac{P_m\lambda_{d,m}^2}{\sigma_{c,m}^2 P_m + \sigma_{f,m}^2} - D_1\right) = 0 \tag{58}$$

$$\mu_m \geq 0, \quad \nu \geq 0 \tag{59}$$

for all m, where μ_m's and ν are the slack variables. To solve for P_m, we first recognize that if $\nu > 0$ then having $\mu_m > 0$ leads to $P_m = 0$ that follows from (57), which violates the condition (58). Hence, this forces $\mu_m = 0$ and $\nu > 0$. The solution of P_m follows from the condition (56) as stated in (40). Finally, substituting the latter equation in (58) gives ν solution stated in (41).

REFERENCES

[1] N. Khalil, M. Abid, D. Benhaddou, and M. Gerndt, "Wireless sensors networks for internet of things," in *IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, April 2014, pp. 1–6.

[2] B. Rashid and M. H. Rehmani, "Applications of wireless sensor networks for urban areas: A survey," *Journal of Network and Computer Applications*, vol. 60, pp. 192–219, 2016.

[3] C. Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[4] J. Chamberland and V. V. Veeravalli, "Wireless sensors in distributed detection applications," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 16–25, May 2007.

[5] P. Zhang, I. Nevat, G. W. Peters, G. Xiao, and H. Tan, "Event detection in wireless sensor networks in random spatial sensors deployments," *IEEE Transactions on Signal Processing*, vol. 63, no. 22, pp. 6122–6135, Nov. 2015.

[6] W.-Z. Song, R. Huang, M. Xu, A. Ma, B. Shirazi, R.-Z. S. LaHusen, R. Huang, M. Xu, A. Ma, B. Shirazi, and R. LaHusen, "Air-dropped sensor network for real-time high-fidelity volcano monitoring," in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys2009)*, 2009, pp. 305–318.

[7] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3. IEEE, 2003, pp. 1713–1723.

[8] Y. Hong, W. Huang, F. Chiu, and C. J. Kuo, "Cooperative communications in resource-constrained wireless networks," *IEEE Signal processing magazine*, vol. 24, no. 3, pp. 47–57, 2007.

[9] S. A. Aldalahmeh, M. Ghogho, D. McLernon, and E. Nurellari, "Optimal fusion rule for distributed detection in clustered wireless sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 5, Jan 2016.

[10] R. Tenney and N. Sandell, "Detection with distributed sensors," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-17, no. 4, pp. 501 –510, 1981.

[11] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Phil. Trans. R. Soc. A*, vol. 370, no. 1958, pp. 100–117, 2012.

[12] S. H. Javadi, "Detection over sensor networks: a tutorial," *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 3, pp. 2–18, March 2016.

[13] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors part i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54 –63, Jan. 1997.

[14] R. Blum, S. Kassam, and H. Poor, "Distributed detection with multiple sensors i. advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64 –79, Jan. 1997.

[15] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions On Aerospace And Electronic Systems*, vol. AES-22, no. 1, pp. 98–101, 1986.

[16] R. Niu and P. K. Varshney, "Joint detection and localization in sensor networks based on local decisions," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, Nov. 2006, pp. 525 –529.

[17] ——, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP Journal on Wireless Communincation snd Networking*, vol. 2005, no. 4, pp. 462–472, 2005.

[18] ——, "Performance analysis of distributed detection in a random sensor field," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 339 –349, Jan. 2008.

[19] S. H. Javadi and A. Peiravi, "Fusion of weighted decisions in wireless sensor networks," *IET Wireless Sensor Systems*, vol. 5, no. 2, pp. 97–105, 2015.

[20] D. Ciuonzo, G. Papa, G. Romano, P. Salvo Rossi, and P. Willett, "One-bit decentralized detection with a Rao test for multisensor fusion," *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 861–864, Sep. 2013.

[21] D. Ciuonzo, P. Salvo Rossi, and P. Willett, "Generalized Rao test for decentralized detection of an uncooperative target," *IEEE Signal Processing Letters*, vol. 24, no. 5, pp. 678–682, May 2017.

[22] D. Ciuonzo and P. Salvo Rossi, "Distributed detection of a non-cooperative target via generalized locally-optimum approaches," *Information Fusion*, vol. 36, pp. 261–274, 2017.

[23] ——, "Quantizer design for generalized locally optimum detectors in wireless sensor networks," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 162–165, Apr. 2018.

[24] M. Guerriero, L. Svensson, and P. Willett, "Bayesian data fusion for distributed target detection in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3417 –3421, June 2010.

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

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

[27] ——, "Bayesian detection in bounded height tree networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 4042–4051, Oct. 2009.

[28] Y. Lin, B. Chen, and P. K. Varshney, "Decision fusion rules in multi-hop wireless sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 2, pp. 475–488, Apr. 2005.

[29] Q. Tian and E. J. Coyle, "Optimal distributed detection in clustered wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3892–3904, 2007.

[30] G. Ferrari, M. Martalo, and R. Pagliari, "Decentralized detection in clustered sensor networks," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 47, no. 2, pp. 959–973, 2011.

[31] R. Streit, *Poisson point processes imaging, tracking, and sensing*. New York: Springer, 2010.

[32] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda *et al.*, "A line in the sand: a wireless sensor network for target detection, classification, and tracking," *Computer Networks*, vol. 46, no. 5, pp. 605–634, 2004.

[33] A. Sundaresan, P. K. Varshney, and N. S. V. Rao, "Copula-based fusion of correlated decisions," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 454–471, January 2011.

[34] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal processing magazine*, vol. 19, no. 2, pp. 40–50, 2002.

[35] Z. Liu and A. Nehorai, "Detection of particle sources with directional detector arrays and a mean-difference test," *IEEE transactions on signal processing*, vol. 53, no. 12, pp. 4472–4484, 2005.

[36] S. Boyd, *Convex Optimization*. England: Cambridge, 2004.