Digital Signal Processing 123 (2022) 103427

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Maximum average entropy-based quantization of local observations for distributed detection ☆

Muath A. Wahdan *,[1], Mustafa A. Altınkaya

İzmir Institute of Technology, Faculty of Engineering, Department of Electrical & Electronics Engineering, İzmir, Turkey

ARTICLE INFO

Article history:
Available online 21 January 2022

Keywords:
Distributed detection
Decentralized detection
Multilevel quantization
Information theoretic distance measures
Wireless sensor networks

ABSTRACT

In a wireless sensor network, multilevel quantization is necessary to find a compromise between minimizing the power consumption of sensors and maximizing the detection performance at the fusion center (FC). The previous methods have been using distance measures such as J-divergence and Bhattacharyya distance in this quantization. This work proposes a different approach based on the maximum average entropy of the output of the sensors under both hypotheses and utilizes it in a Neyman-Pearson criterion-based distributed detection scheme to detect a point source. The receiver operating characteristics of the proposed maximum average entropy (MAE) method in quantizing sensor outputs have been evaluated for multilevel quantization both when the sensor outputs are available error-free at the FC and when non-coherent M-ary frequency shift keying communication is used for transmitting MAE based multilevel quantized sensor outputs over a Rayleigh fading channel. The simulation studies show the success of the MAE in the cases of both error-free fusion and where the effect of the wireless channel has been incorporated. As expected, the performance improves as the level of quantization increases and with six-level quantization approaches the performance of non-quantized data transmission.

1. Introduction

Wireless Sensor Networks (WSNs) have come into the spotlight recently due to a significant development in the Micro-Electro-Mechanical Systems (MEMS) [1],[2]. The recent development of WSNs has made this field a research focus of intensive researches. Researchers have been widely using it to monitor and characterize large physical environments and trace various environmental or physical conditions such as temperature, pressure, wind, and humidity. Apart from these, WSNs have vast fields to be applied in, such as harmful environmental exploration, wildlife monitoring, target tracking and smart cities established based on Internet of Things (IoT) [3–7]. Typically a WSN uses many comparatively inexpensive and low-energy sensors to collect observations and pre-process the observations. These sensors are generally deployed in the environment. Owing to strict energy and bandwidth restrictions, the sensors' observations are frequently needed to be quantized before transmitting them to a fusion center (FC) which makes a global decision [8,9]. This work concentrates on the distributed detection problem using a WSN and, particularly, how the local observations are quantized.

1.1. Pioneering studies on distributed detection

The additional requirement in detection because of the distributed nature is to jointly optimize both the sensors' processing and how their outputs are fused. The pioneering research on fusion rules was made by Tenney and Sandell [10] and Chair and Varshney [11]. In [10], a detection problem consisting of two sensors and one FC with a fixed fusion rule was considered to show that the optimum local decision rule is the likelihood ratio test (LRT) under the Bayesian criterion. However, the individual thresholds are coupled. Later, in [11], it was shown that the optimum fusion rule at the FC for multiple observations is also an LRT both under the Neyman-Pearson (NP) and the Bayesian criteria. Determining the optimal local decision rule is significantly more complicated. The optimality of LRT for each local decision rule was considered in [12] and [13] by assuming conditional independence of the observations under each hypothesis. But because of the coupling

☆ A preliminary version of this paper was presented in the 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24-26 April 2019.

* Corresponding author.
E-mail addresses: moath.wahdan@ptuk.edu.ps (M.A. Wahdan), mustafaaltinkaya@iyte.edu.tr (M.A. Altınkaya).

[1] This work was entirely carried out during Muath A. Wahdan's stay in IZTECH as Ph.D. student and he is now an assistant professor at Palestine Technical University – Kadoorie, Tulkarm, Palestine.

https://doi.org/10.1016/j.dsp.2022.103427
1051-2004/

between the LRT thresholds at the local detectors among themselves and with the one at the FC, solving the global optimization problem is complex, although not intractable [14]. This suggested determining the thresholds of the local detectors independently, that is, the threshold of each sensor is optimized for fixed decision rules at the other detectors and the FC, which is called person-by-person optimization (PBPO) [14–16]. Those works' adopted conditional independence assumption produces only locally optimal decisions, but even they become prohibitively complex for large sensor networks, and simpler solutions are needed. Additionally, the gain obtained by having more sensor nodes outperforms getting more information from each sensor in WSNs [17].

1.2. Studies on system models

The optimal fusion rule under the known probability of false alarm (p_{fa}) and probability of detection (p_{d}) of local sensors is given by the Chair-Varshney fusion rule in [11]. Many works are dedicated to studying the counting fusion rule, which achieved comparable performance to that optimum fusion rule [18–20]. Specifically, these works have suggested transmitting only the local decisions and using the counting rule at FC instead of transmitting the raw data, which is expensive, particularly for a typical WSN with limited bandwidth and energy.

Most of the previously mentioned WSN applications considered the case where signal-to-noise ratio (SNR) or the p_{d} and p_{fa} at local sensors are known to the FC. On the other hand, in many scenarios, an isotropic source of an unknown location is considered where the emitted signal is assumed to decay as a function of the distance from the source [20–22]. In [20], the detection performances, namely p_{fa} and p_{d}, were derived using numerous simplifying assumptions by several methods in order to simplify its computations for binary data transmissions.

On the contrary, in [21], a generalized likelihood ratio test (GLRT) detector and the Cramer-Rao lower bound (CRLB) are derived; however, GLRT needs a grid search on both the emitted power from the target and the location domains. A computationally simpler solution was given in [22], where one-bit distributed detection of an uncooperative target was considered, which assumes that both the target location and the emitted signal are unknown at FC. In that work, a Generalized Locally-Optimum Detector (G-LOD) test with nuisance parameters was proposed for a fixed value of local sensor thresholds. However, most of these mentioned works examined binary data transmissions, and the effect of Rayleigh fading was not considered. In contrast, this paper proposes a new type of quantization (MAE) for M-ary data transmissions by considering Rayleigh fading channel between the sensors and the FC. Consider the literature [23–26] for more details on the fusion rules for different network topologies such as parallel, serial, and tree topologies.

1.3. Studies on quantization for distributed detection

Optimum quantization levels in the sense of information-theoretic criteria for distributed detection systems were presented in [27–31]. In [27], the quantization based on Ali-Silvey distances between two simple hypotheses was investigated. After that, in [28], [29], the divergence was proposed as a distortion measure by considering a class of f-divergence measures which shows that the loss in divergence is quadratic with the quantization step size.

In [30,31], the authors considered that each local detector transmits a multiple-bit decision to the FC. The solution for partitioning the local decision space was derived by maximizing the distance between the mean values of the quantized hypotheses. It was shown that the global decision performance increases monotonically by increasing the number of partitions at the individual detector. This method is locally optimum in the sense of J-divergence (JD), but it does not necessarily yield a globally optimum solution. Even when four quantization levels are considered, the solution is given by complicated analytic expressions explaining the functional relationships between the detection probability and the false alarm probability of all detectors and their derivatives. Those works assumed that all local sensors are identical NP detectors observing the same SNR.

In [32,33], to perform optimum quantization in the sense of mean-error, deflection criterion (DC) and Chernoff information (CI) were defined for distributed detection systems consisting of one FC and multiple sensors by using the Bayesian detection criterion. DC and CI pose a nonlinear and non-convex problem, which mostly has more than one extreme. These optimization criteria are suitable for the case of known SNR, where the p_{d} and p_{fa} are known for each local detector.

There are also a few other information-theoretic methods of quantization that we can consider as not based on distance measures [34,35]. [34] called their approach as minimum equivocation detection, which is equivalent to maximizing mutual information (MI). Later, [35] used conditional-MI (CMI) method for distributed detection. The MI and CMI methods applied the PBPO approach to optimize both the sensor thresholds and the fusion rule; however, they were confined to binary sensor decisions and used for known SNR.

1.4. Some other works

In another work, a multi-bit Rao test was considered for fusing sensor outputs, and a quantizer design method was proposed to maximize the non-centrality parameter of the test statistic distribution [36]. That optimization problem was solved using a particle swarm optimization algorithm because of its nonlinear and non-convex nature.

Early distributed detection systems were developed by assuming an error-free communication between the local detectors and the FC [13]. Applying this theory to WSNs leads to a detection performance loss in the case of erroneous channels. Fusion rules for distributed detection systems considering fading channels were first discussed in [37] and later in [38–40] mainly for binary hypothesis testing, whereas [41–44] considered M-ary hypothesis testing problems.

Dependence between observations of spatially distributed sensors is a critical subject in DD problems. This dependence occurs for various reasons, such as sensing the same phenomenon and contamination by correlated noise. Some works have studied the influence of dependence on the performance of DD systems [45,46]. But, in many DD scenarios, this dependence is ignored and the sensor observations are considered to be independent.

1.5. Contributions of the paper

Inspired by quantizing signals using the Maximum Output Entropy (MOE) in [47], we propose an entropy-based method by maximizing the average entropy of observations under both hypotheses. We aim to determine the quantization intervals at distributed sensors to optimize the global binary decision at the FC about the existence of a point source under the NP criterion where sensors observe different signal levels which they do not know. Although maximizing the entropy is a well-known approach, only MI and CMI were used for maximization in decision problems [34,35]. The most probable reason for this is the widespread acceptance that an information-theoretic criterion for decision problems should concentrate on the distance of rival hypotheses. We consider scenarios with non-equivalently important hypotheses; that

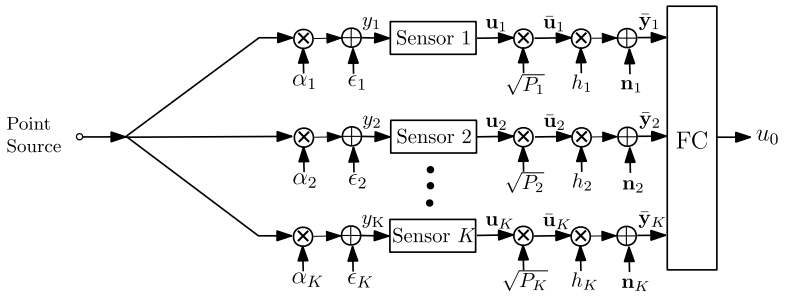

Fig. 1. Parallel distributed detection system.

is why the NP criterion is considered more suitable than the probability of error criterion in this work. This paper extends its preliminary version [48] in the following aspects. We compare the proposed method to DC [32,33], CI [33], MI [34,35], CMI [35] and JD [30,31] based methods, demonstrate its positively proportional relation with JD, include increased quantization levels resulting in a similar performance to non-quantized signaling. Instead of the binary symmetric channel as a simplified model for the channel from the sensors to the FC, we use a regular Rayleigh fading channel model for the wireless channel. Additionally, we utilize optimal and sub-optimal fusion rules modified from those for M-ary hypothesis testing in [44] to match the binary hypothesis testing problem with M-ary modulated data transmission.

1.6. Paper organization

The remaining part of this paper is organized as follows. First, we formulate the parallel distributed detection problem of a point source, including sensors to FC transmissions over a Rayleigh fading channel and various fusion rules in Section 2. Section 3 covers the development of the proposed average entropy-based quantization method, the JD-based method and their relation. Simulation results are given in Section 4, and conclusions are drawn in Section 5.

1.7. Manuscript notation

Boldface lower and upper case letters denote vectors and matrices, respectively. The symbol "$\sim$" stands for "distributed according to", whereas $\mathcal{N}(\mu, \sigma^2)$ denotes Gaussian probability density function (pdf) with mean μ and variance σ^2. $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ indicates complex Gaussian pdf with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$; p_{fa} and p_{d} denote the probability of false alarm and probability of detection, respectively; u_k represents the local M-ary decision at the k^{th} sensor and u_0 represents the global binary decision at the FC. ϵ_k refers to the additive white Gaussian noise at the k^{th} sensor; h_k denotes complex channel coefficient multiplying the k^{th} sensor's output. $\boldsymbol{\beta}_{\text{M}}$ is a vector of local thresholds for M-level quantization; η is the global threshold, $\hat{F}$ is entropy function; $\hat{E}(\cdot)$ is expectation operation, A is signal amplitude; y_k denotes the observation at sensor k; D_{KL} is Kullback–Leibler divergence; $\boldsymbol{\Lambda}(\cdot)$ is likelihood ratio; $\hat{u}_k$ is the estimate of the scalar u_k and $\mathbf{u}_k$ is M-FSK modulated sensor output.

2. System model

A binary hypothesis testing problem has been considered in this work, where a group of K sensors and one FC cooperate to detect the existence of a point source, as shown in Fig. 1. The hypothesis testing at each sensor node can be described as

$$\begin{aligned} &H_0 : y_k = \epsilon_k, \\ &\text{versus} \\ &H_1 : y_k = A_k + \epsilon_k, \end{aligned} \tag{1}$$

where y_k denotes the observation at the k^{th} sensor and ϵ_k denotes additive white Gaussian noise (AWGN) with variance σ^2 and zero mean. A_k denotes the received signal amplitude, which is equal to $\alpha_k A_{\max}$. Each sensor in the range of the point source measures a signal attenuated with a factor of α_k and makes an M-level local decision $u_k \in \{0, 1, ..., M-1\}$. The local decision is transmitted through multiplicative channel h_k to the FC, where the final decision u_0 is made. In Fig. 1, the sensor outputs $\{\mathbf{u}_k, k = 1, 2, ..., K\}$, the AWGNs in the channel from the sensors to the FC $\{\mathbf{n}_k, k = 1, 2, ..., K\}$ and the received signals $\{\bar{\mathbf{y}}_k, k = 1, 2, ..., K\}$ are shown as vectors following the M-dimensional signal model of Frequency Shift Keying (FSK) related modulated signal model, explained in detail in section 2.2.1. Additionally, $\sqrt{P_k}$ denotes the received power from the k^{th} sensor which includes the large scale fading effect.

2.1. The signal density with punctured disk simplification and optimal centralized detection

In this work, we consider an isotropic, static point event source for which the signal power decays with the square of the distance from the source [20]. We can equivalently say that the signal amplitude, A_k, received at the k^{th} sensor will be inversely proportional to the Euclidean distance, r_k, between the source and the k^{th} sensor. Practical receivers or sensors will have some sensitivity that is the minimum signal level, $A_{\min}$, which can be measured. Consequently, among the sensors uniformly deployed in a much larger area, only those located in a circle with radius $r_{\max}$, as given in Fig. 2, will be able to receive a signal. An upper limit for the received signal level should also be considered for practical sensors, which corresponds to saturation. Because of this, we preferred defining the smallest distance, $r_{\min}$, from the event location, which corresponds to the saturation signal level, $A_{\max}$, of the sensors. The signal amplitude at the k^{th} sensor can be given as

$$A_k = \alpha_k A_{\max}, \tag{2}$$

where $\alpha_k = r_{\min}/r_k$.

The pdf of A is inversely proportional with A. When we take its integral to find the area under the curve, we encounter a difficulty for large A. [20] solved this problem, limiting A to a maximum value corresponding to a small distance $r_{\min}$. Our choice in this paper is to neglect any sources in the source neighborhood defined by a distance of $r_{\min}$. So, dropping the index for the sensors, the pdf of the normalized signal amplitude, $A_n = A/A_{\max}$, at a sensor will have the form shown in Fig. 3 and will be given as:

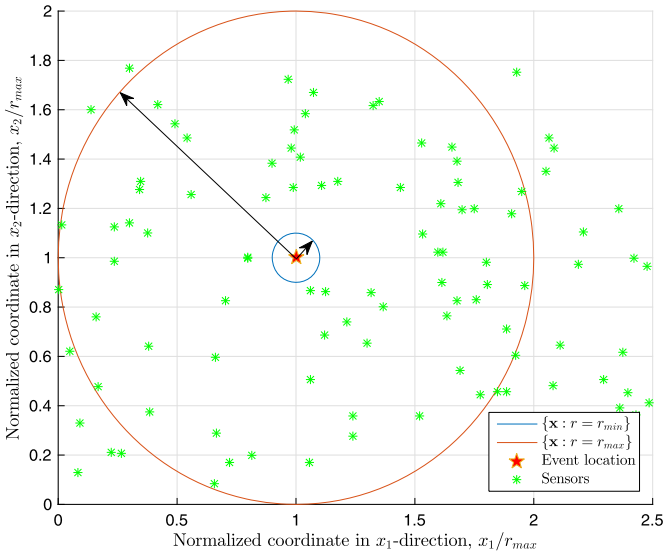

Fig. 2. Positions of the event location and uniformly distributed sensors in a scenario for detecting a point source.

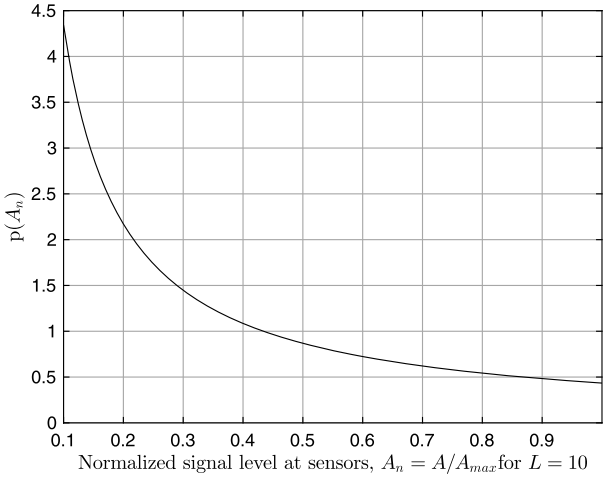

Fig. 3. The pdf of the signal amplitude observed at the sensors, $p(A_n)$.

$$p(A_n) = \frac{1}{A_n\left[\log(A_{\max}) - \log(A_{\min})\right]} = \frac{1}{A_n \log(L)}, \tag{3}$$

where $L = A_{\max}/A_{\min} = r_{\max}/r_{\min}$ and $\log(\cdot)$ is the natural logarithm. We define the SNR as the ratio between the maximum signal power, $A_{\max}^2$, and the noise power, σ^2. Let us assume that K of the sensors uniformly deployed in the area will be in the fat ring (or punctured disk) described by the radii $r_{\min}$ and $r_{\max}$. Then, the signal amplitudes at these sensors will be independent and come from the pdf given in (3) in the case of an event. Assuming that the sensor observations are available distortion-free at the FC, i.e., without transmission over a wireless channel, we have a centralized detection system, effectively, and the optimal Bayesian NP detector can be written as:

$$\mathbf{\Lambda}(\mathbf{y}) = \frac{\prod_{k=1}^{K} \int_{A_{\max}/L}^{A_{\max}} p(y_k|H_1; A_k) p(A_k) dA_k}{p(\mathbf{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{4}$$

Note that our approach differs from [20,22], where an averaging over the target's position is applied. Such an averaging is not required since only the sensors in the reception range of a signal are considered. A question arises about the selection of those sensors in the reception range. However, this should not be an issue since the FC, which knows the locations of sensors, can easily determine more or less such a circular region around the true target location based on the locations of sensors sending H_1 decisions. Since each A_k comes from the independent and identical pdf given in (3), we eliminate the index, k, and express the likelihood ratio as

$$\mathbf{\Lambda}(\mathbf{y}) = \frac{\prod_{k=1}^{K} \int_{A_{\max}/L}^{A_{\max}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_k - A)^2}{2\sigma^2}\right) \frac{1}{A\log(L)} dA}{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^K \exp\left(\frac{-\sum_{k=1}^{K}(y_k)^2}{2\sigma^2}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{5}$$

where $\mathbf{y} = [y_1, y_2, ..., y_K]^T$ denotes the $K \times 1$ column vector of observations from K sensors.

2.2. Fusion system: channel between sensors and FC

This section will investigate the complete model for the sensor to FC communication using a Rayleigh fading channel model and an M-ary frequency-shift keying (M-FSK) modulation scheme where for M different symbols carrier waves of M different frequencies are transmitted. M-FSK is a suitable modulation scheme for low-power low data rate transmission preferred by the majority of the sensor device equipment. It was shown in [44] that allocating some of the limited power at the sensors for training to make coherent reception of FSK signals at the FC results in worse performance in terms of error probability. Accordingly, non-coherent demodulation of M-FSK was adopted in this paper. Additionally, to concentrate on the fusion of sensor data with non-identical signal levels, firstly, we considered the case of error-free channels, i.e., when error-free sensor outputs are available at the FC, which we called direct data transmission (DDT). Once M-level quantized data from the sensors are at the FC, an equal gain fusion rule is applied since the relative reliability of sensor outputs is not evaluated.

2.2.1. Fading channel

This subsection considers the problem of fusing the data transmitted over a fading channel, as shown in Fig. 1. The FC has only information on the channel statistics. Non-coherent M-FSK modulation is employed for sending data to the FC. Let $\mathbf{u}_k$ denote the M-FSK modulated symbol at sensor k, where $\mathbf{u}_k \in \{\mathbf{e}_m, m = 1, ..., M\}$ and $\mathbf{e}_m$ is an $M \times 1$ column vector, all elements of which except the m^{th} one are zero. We refer to the transmit power of the data symbol as P_{t}. Assuming M-dimensional signal model for representing the orthogonal channels of M-FSK modulation scheme between the sensors and the FC [44] simplifies the analysis. Then, the signal at the FC received from the k^{th} sensor can be given as

$$\begin{aligned}\bar{\mathbf{y}}_k &= \sqrt{P_k} h_k \mathbf{u}_k + \mathbf{n}_k, \\ &= h_k \bar{\mathbf{u}}_k + \mathbf{n}_k,\end{aligned} \tag{6}$$

where P_k represents the received power which is a function of P_{t}, the wavelength, the path loss exponent and the distance between the k^{th} sensor and the FC [44], and it describes the effect of large-scale fading. The channel noise is denoted as $\mathbf{n}_k$ which is a zero-mean complex Gaussian vector $\mathbf{n}_k \sim \mathcal{CN}(0, \sigma_n^2\mathbf{I})$, where $\mathbf{I}$ is an $M \times M$ identity matrix. The complex channel coefficient h_k in (6) is modeled as $h_k \sim \mathcal{CN}(0, 1)$, which can be also represented as

$h_k = \xi_k e^{j\phi_k}$, where ξ_k represents the amplitude with Rayleigh distribution and ϕ_k represents the phase with uniform distribution. We adopt the NP criterion to find the optimal and a sub-optimal fusion rule at the FC to obtain a global decision $u_0 \in \{H_0, H_1\}$ as follows:

(i) The optimal fusion rule for the independent and identically distributed (i.i.d.) vectors, $\bar{\mathbf{y}}_k, k = 1, 2, ..., K$, is defined as

$$\log \mathbf{\Lambda}(\bar{\mathbf{Y}}) = \log \frac{p(\bar{\mathbf{Y}}|H_1)}{p(\bar{\mathbf{Y}}|H_0)} = \log \prod_{k=1}^{K} \frac{p(\bar{\mathbf{y}}_k|H_1)}{p(\bar{\mathbf{y}}_k|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{7}$$

where $\bar{\mathbf{Y}}$ is the matrix composed of row-wise stacking column vectors $\bar{\mathbf{y}}_k$.

Expanding $p(\bar{\mathbf{y}}_k|H_1)$ and $p(\bar{\mathbf{y}}_k|H_0)$ in (7) over the M-level sensor decisions, we obtain

$$\log \mathbf{\Lambda}(\bar{\mathbf{Y}}) = \sum_{k=1}^{K} \log \left(\frac{\sum_{m=1}^{M} p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m)) p(\mathbf{u}_k(m)|H_1)}{\sum_{m=1}^{M} p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m)) p(\mathbf{u}_k(m)|H_0)} \right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{8}$$

The conditional density $p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m))$ in (8) is a complex multi-variate Gaussian density, $\bar{\mathbf{y}}_k \sim \mathcal{CN}(0, \mathbf{C}_{\bar{\mathbf{y}}})$, $\mathbf{C}_{\bar{\mathbf{y}}}$ represents the diagonal matrix with entries $\mathbf{C}_{\bar{\mathbf{y}}}(j, j) = \sigma_n^2$ for $j \neq m$ and $\mathbf{C}_{\bar{\mathbf{y}}}(j, j) = P_k \sigma_h^2 + \sigma_n^2$ for $j = m$, where $j = 1, ..., M$. We can prove that

$$p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m)) = \frac{1}{\sqrt{\pi^M \det|\mathbf{C}_{\bar{\mathbf{y}}_m}|}} \times \exp\left\{ -(\bar{\mathbf{y}}_k - \boldsymbol{\mu})^H \mathbf{C}_{\bar{\mathbf{y}}_m}^{-1} (\bar{\mathbf{y}}_k - \boldsymbol{\mu}) \right\}, \tag{9}$$

which can be re-written as in [44]

$$p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m)) = \frac{1}{\sqrt{\pi^M \det|\mathbf{C}_{\bar{\mathbf{y}}_m}|}} \exp\left(\frac{P_k \sigma_h^2 |\bar{y}_k(m)|^2}{\sigma_n^2(\sigma_h^2 + \sigma_n^2)} \right) \times \prod_{j=1}^{M} \exp\left(\frac{|\bar{y}_k(j)|^2}{\sigma_n^2} \right). \tag{10}$$

Note that $|\bar{y}_k(m)|^2$ in (10) denotes M cross-correlator squared envelopes corresponding to non-coherent FSK detection. The values of $p(\mathbf{u}_k(m)|H_1)$ represent the probability masses under hypothesis H_1, which are estimated as

$$\overline{p_m^{H_1}} = \int_{A_{\max/L}}^{A_{\max}} p_m^{H_1}(A_n) p(A_n) dA_n, \tag{11}$$

where $p_m^{H_1}(A_n)$ represents the probability mass under H_1 as shown in Fig. 4 for an observed signal level A_n, the Gaussian signal's mean.

The values of $p(\mathbf{u}_k(m)|H_0)$ represent the probability masses under hypothesis H_0:

$$p(\mathbf{u}_k(m)|H_0) = p_m^{H_0}. \tag{12}$$

Fig. 4 shows a possible partitioning of a pdf under hypothesis H_i, $i = 0, 1$ and the probability masses for $M = 4$ corresponding to the areas under the pdf between successive thresholds.

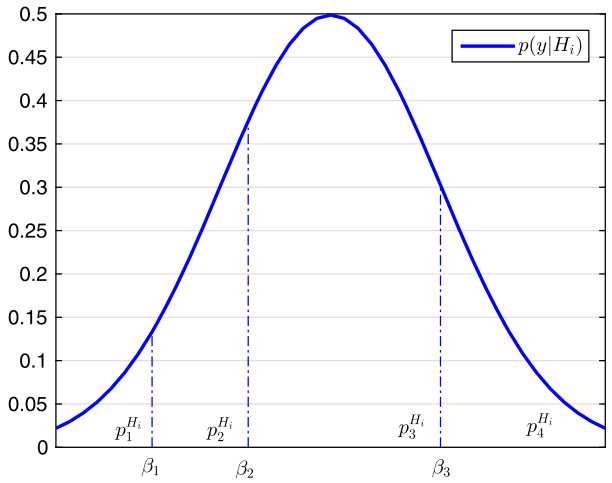

Fig. 4. A partitioning of the pdf for the observations at each sensor for 4-level quantization.

(ii) A sub-optimal fusion rule can be derived as follows: This is a type of decode-then-fuse class of fusion rule, as introduced in [49]. In (8), we see both the effects of fading channel and the local detection outputs to achieve the optimal performance. An alternate formulation could be used as a sub-optimal fusion rule by separating this into two steps. First, $\bar{\mathbf{y}}_k$ is used to infer about the local detector by applying the maximum likelihood (ML) estimate as an intermediate decision, $\hat{u}_k$. Then, the optimum fusion rule based on $\hat{u}_k$ is applied:

$$\hat{u}_k = \arg\max_m \quad p(\bar{\mathbf{y}}_k|\mathbf{u}_k(m)). \tag{13}$$

By substituting (10) in (13) after eliminating the terms which are irrelevant to m, we can re-write (13) as

$$\hat{u}_k = \arg\max_m \quad \exp\left(\frac{P_k \sigma_h^2 |\bar{y}_k(m)|^2}{\sigma_n^2(\sigma_h^2 + \sigma_n^2)} \right), \tag{14}$$

where $m = 1, ..., M$. The final decision rule is given as

$$u_0 = \sum_{k=1}^{K} \hat{u}_k \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{15}$$

3. Quantizer design

This paper aims to make a global decision at the FC under the NP criterion. Let us assume that each sensor will only make a single observation and will transmit this observation to the FC. Then, sensors will make i.i.d. observations under H_0, and none can estimate the signal level under H_1. Consequently, there is no additional information at the sensors to use different quantization thresholds under H_1. So, it is reasonable to use identical quantization thresholds at each sensor irrespective of their distance to the event location since it cannot be estimated. Definitely, the choice of the quantization thresholds affects the performance, making it desirable to choose the quantization thresholds that maximize the system performance. This paper proposes the maximum average entropy (MAE) method, that is, determining the quantization thresholds at the sensors to maximize the average entropy of the discrete information collected at the FC under both hypotheses without considering the effects of the succeeding wireless channel. Besides the MI and CMI methods [34,35], most entropy-based quantizers for detection problems are distance measures

[27,33]. Like MI and CMI, the proposed MAE method is not distance measure based and maximizes the transmitted information corresponding to both of the underlying probability mass functions (pmfs) jointly.

The optimum detector at the FC is based on likelihood ratios as given in (5). Equivalently, one can use log-likelihood (logarithm of likelihood) ratios. The log-likelihood ratio for the k^{th} sensor with an unknown signal amplitude can be calculated using the expected value of the signal amplitude, $\bar{A}$, as follows:

$$\log(\Lambda) = -\frac{\bar{A}^2}{2\sigma^2} + \frac{\bar{A}}{\sigma^2} y_k. \tag{16}$$

The linear (or more appropriately affine) transformation of observations in (16) to log-likelihood ratios is irrelevant in entropy-based quantization because that kind of transformation only results in translation and scaling of the underlying pdfs and will preserve the resulting probability masses corresponding to a vector of thresholds (such as β_1, β_2 and β_3 in Fig. 4). Consequently, the sensors will transmit quantized observation signals to the FC. A common information-based criterion for determining the quantization thresholds, which will be the best rival of the proposed MAE method as shown in Section 4.1, is the maximum JD (MJD) method which was used in the case of the constant signal level at sensors formerly [30]. We will first explain these criteria and, subsequently, their relation.

3.1. MAE method

An intuitive idea to have optimum performance at the FC is to maximize the entropy under both hypotheses, which we call as MAE method. So, we propose determining the quantization intervals at the sensors as resulting in MAE under both hypotheses. The entropy of a quantized sensor output can be calculated based on the partitioning of the observation pdf at each sensor, as shown in Fig. 4. In this figure, the number of quantization intervals is 4. For a general number of M quantization intervals, there will be $M-1$ thresholds, $\{\beta_1, \beta_2, ..., \beta_{M-1}\}$, and M partitions with corresponding probability masses of observations $\left\{p_1^{H_i}, p_2^{H_i}, ..., p_M^{H_i}\right\}$, where $i = 0, 1$. Under H_i, one can estimate the entropy of the observation as

$$\hat{F}_{H_i} = \hat{E}\left(-\sum_{m=1}^{M} p_m^{H_i} \log_2(p_m^{H_i})\right) \text{bit}. \tag{17}$$

The expectation, $\hat{E}(\cdot)$, is with respect to the distribution of the K sensor observations, and in the particular case of the scenario described in Fig. 2, this distribution is uniform in the sensing range of the sensors defined by a fat ring between radii $r_{\min}$ and $r_{\max}$ from the event location. $\mathbf{p}_M^{H_0} = \left[p_1^{H_0}, p_2^{H_0}, ..., p_M^{H_0}\right]$ denotes the vector of these probability masses, i.e., the probabilities of the partitions. In practice, one obtains an estimate of this expectation by averaging the sensors' information over the distribution of the sensor locations and AWGN realizations which is called a histogram method [50]. Fig. 5 shows the entropy functions $\hat{F}_{H_0}$, $\hat{F}_{H_1}$ and $\hat{F}_{\text{av}} = \frac{1}{2}\left(\hat{F}_{H_0} + \hat{F}_{H_1}\right)$ for binary quantization. For M-ary quantization, $\boldsymbol{\beta}_M^* = \left[\beta_1^*, \beta_2^*, ..., \beta_{M-1}^*\right]$ denotes the vector of optimum quantization thresholds in the sense of MAE which is found as

$$\boldsymbol{\beta}_M^* = \arg\max_{\boldsymbol{\beta}_M} \quad \hat{F}_{\text{av}}. \tag{18}$$

For binary quantization, the optimal quantization threshold is $\beta_2^* = 0.093$, as depicted in Fig. 5. Similarly, the entropy functions $\hat{F}_{H_0}$, $\hat{F}_{H_1}$ and $\hat{F}_{\text{av}}$ for three-level quantization are given as

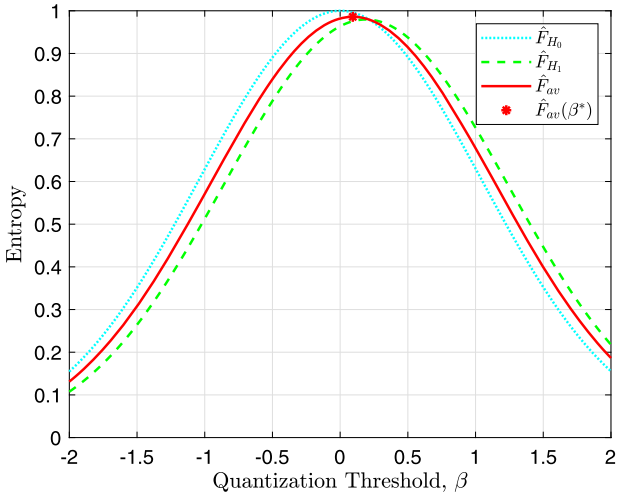

Fig. 5. The entropy functions $\hat{F}_{H_0}$, $\hat{F}_{H_1}$ and $\hat{F}_{\text{av}}$ for binary quantization.

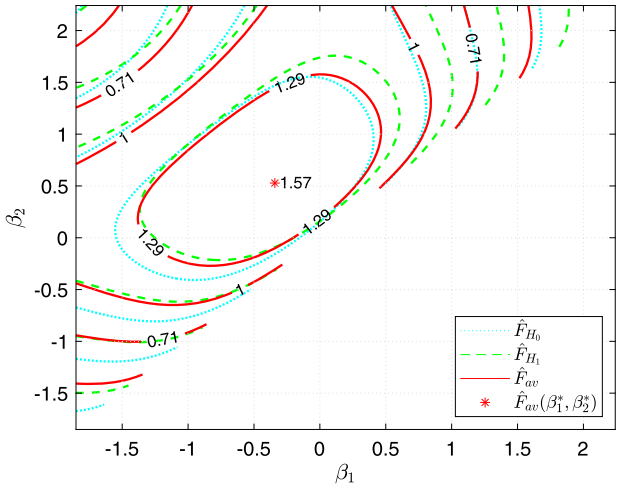

Fig. 6. The entropy functions $\hat{F}_{H_0}$, $\hat{F}_{H_1}$ and $\hat{F}_{\text{av}}$ for three level quantization.

contour plots in Fig. 6. The optimal quantization threshold vector is found as $\boldsymbol{\beta}_3^* = [-0.341, \quad 0.528]$, as shown in this figure. Similarly, the optimum thresholds in the 4-level and 6-level quantization cases are $\boldsymbol{\beta}_4^* = [-0.367, \quad 0.195, \quad 0.835]$ and $\boldsymbol{\beta}_6^* = [-1.08, \quad -0.572, \quad -0.060, \quad 0.4513, \quad 0.963]$, respectively. The pseudo-code of the MAE method for quantization is given in Algorithm 1.

3.2. MJD method

JD can be written in terms of the relative entropy for discrete probability distributions P and Q observed under the two hypotheses H_0 and H_1, respectively, as follows:

$$J = D_{KL}(P||Q) + D_{KL}(Q||P), \tag{19}$$

where the relative entropy between two pmfs $P(x)$ and $Q(x)$ is given as follows:

$$D_{KL}(P||Q) = \sum_{x\in\chi} P(x) \log_2\left(\frac{P(x)}{Q(x)}\right), \tag{20}$$

where χ denotes the alphabet of the pmfs for P and Q. In our context, JD measures the distributional distance, i.e., the dissimilarity between the distributions, of the observations under two

Algorithm 1 The procedure of performing MAE quantization.

Input $A_{\min}$, $A_{\max}$, δ_A, M, δ_β
Define $(M-1)$-vector $\boldsymbol{\beta}_m$ of thresholds with $\beta_{M-1} \geq \beta_{M-2} \geq \ldots \geq \beta_1$
for $\boldsymbol{\beta}_m = [\beta_{M-1_{\min}}, \beta_{M-2_{\min}}, \ldots \beta_{1_{\min}}] : [\beta_{M-1_{\max}}, \beta_{M-2_{\max}}, \ldots \beta_{1_{\max}}]$ ▷ $(M-1)$ nested loop with $\delta_\beta = 0.01$ step size in each dimension
  for $A = A_{\min} : \delta_A : A_{\min}$ **do** ▷ Averaging over the histogram of A
    $w_A = \frac{1}{A \log(L)}$ ▷ calculate the weight of each A
    for $m = 1 : M$ **do**
      $p_m^{H_1}(A) = \int_{\text{Area}_m} p(y|H_1)dy$
    end for
    $H_1(A) = w_A \times (-\sum_{m=1}^{M} p_m^{H_1}(A) \log_2(p_m^{H_1}(A)))$ bit.
end for
$\hat{F}_{H_1}(\boldsymbol{\beta}_m) = \sum_A (H_1(A))$
for $m = 1 : M$ **do**
  $p_m^{H_0} = \int_{\text{Area}_m} p(y|H_0)dy$
end for
$\hat{F}_{H_0}(\boldsymbol{\beta}_m) = -\sum_{m=1}^{M} p_m^{H_0} \log_2(p_m^{H_1})$bit. ▷ calculate the entropy under hypothesis H_0
$\hat{F}_{\text{av}}(\boldsymbol{\beta}_m) = \frac{1}{2}(\hat{F}_{H_0} + \hat{F}_{H_1})$
$\boldsymbol{\beta}_M^* = \arg\max_{\{\boldsymbol{\beta}_M\}} \quad \hat{F}_{\text{av}}$. ▷ find the optimal thresholds

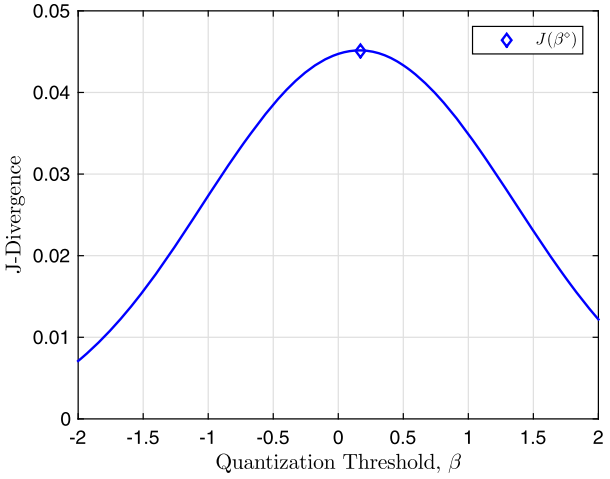

Fig. 7. The JD for binary quantization.

hypotheses, H_0, H_1. So, one can use it to find the local thresholds. The choice of local thresholds facilitates the design of local detectors, which determines the whole system's performance. One can estimate the expectation of JD by averaging the contribution to the JD over the distribution of sensor locations and noise realizations as performed for the entropy of the observations in (17) and can be written as:

$$\hat{J} = \hat{E}\left(\sum_{m=1}^{M}\left[p_m^{H_1}\log_2\left(\frac{p_m^{H_1}}{p_m^{H_0}}\right) - p_m^{H_0}\log_2\left(\frac{p_m^{H_1}}{p_m^{H_0}}\right)\right]\right). \tag{21}$$

It is evident that $\hat{J}$, as specified by (21), is a function of the probability masses corresponding to the partitions of the pdf. For M-ary quantization, $\boldsymbol{\beta}_M^\diamond = [\beta_1^\diamond, \beta_2^\diamond, \ldots, \beta_{M-1}^\diamond]$ denotes the JD optimized vector of quantization thresholds which can be given as

$$\boldsymbol{\beta}_M^\diamond = \arg\max_{\boldsymbol{\beta}_M} \quad \hat{J}. \tag{22}$$

Optimal quantization thresholds correspond to the maximum of $\hat{J}$, which is $\beta_2^\diamond = 0.17$ for binary quantization, as shown in Fig. 7. Similarly, we can estimate the optimal thresholds for 3-level quantization to be $\boldsymbol{\beta}_3^\diamond = [-0.444, \quad 0.784]$, as shown in Fig. 8. Likewise, the optimal thresholds are $\boldsymbol{\beta}_4^\diamond = [-0.725, \quad -0.699, \quad 0.6559]$ and $\boldsymbol{\beta}_6^\diamond = [-6.19, \quad -0.572, \quad -0.0603, \quad 0.9628, \quad 6.59]$ in the cases of 4-level and 6-level quantizations, respectively.

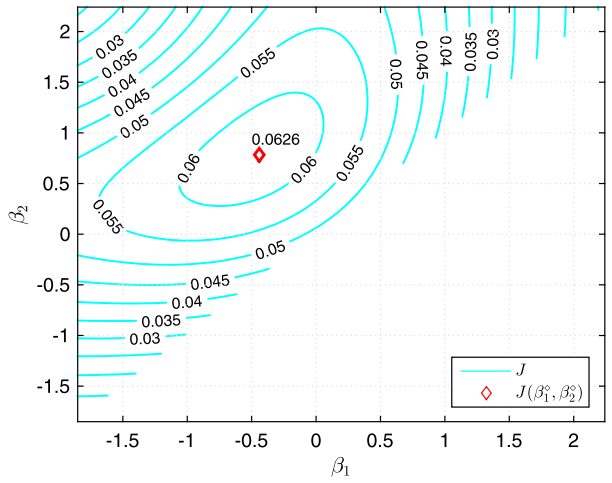

Fig. 8. The JD for three level quantization.

3.3. Relation of MAE and MJD methods

In this subsection, we will demonstrate that the information-based criteria, MAE and MJD, maximize similar quantities in showing that they are positively proportional. Let us first express $D_{KL}(P||Q)$ given in (20) as follows:

$$D_{KL}(P||Q) = \underbrace{\sum_{x\in\chi} P(x)\log_2\left(\frac{1}{Q(x)}\right)}_{R_1} + \underbrace{\sum_{x\in\chi} P(x)\log_2(P(x))}_{-F_{H_0}} \geq 0. \tag{23}$$

The equality holds only when $P = Q$. Similarly,

$$D_{KL}(Q||P) = R_2 - F_{H_1} \geq 0, \tag{24}$$

where $R_2 = \sum_{x\in\chi} Q(x)\log_2\left(\frac{1}{P(x)}\right)$. Substituting (23) and (24) into (19)

$$J = R_1 + R_2 - (F_{H_0} + F_{H_1}) \geq 0. \tag{25}$$

In (25), we consider R_1, R_2, F_{H_0} and F_{H_1} with their scalar values representing some information in bits. Then, defining $D_{KL}(P||Q) = c_1 F_{H_0}$ and $D_{KL}(Q||P) = c_2 F_{H_1}$, we can re-write the JD in (25) to show that there is a proportionality relation between the JD and the average entropy (AE):

$$\begin{aligned} J &= c_1 F_{H_0} + c_2 F_{H_1} \\ &= \min\{c_1, c_2\}\underbrace{(F_{H_0} + F_{H_1})}_{2F_{\text{av}}} + c_3 \end{aligned} \tag{26}$$

with

$$c_3 = \begin{cases} (c_1 - c_2)F_{H_0} & \text{for } c_1 \geq c_2, \\ (c_2 - c_1)F_{H_1} & \text{for } c_1 \leq c_2. \end{cases} \tag{27}$$

Obviously, $c_i \geq 0$ for $i = 1, 2, 3$. Therefore, AE and JD are positively proportional.

4. Simulation results

Monte Carlo simulations have been performed in order to evaluate the detection performance for the proposed method at SNR= 0 dB for $K = 25$ transmitting sensors and $L = A_{\max}/A_{\min} = 10$. First, we have performed simulations using the DDT method, that is, assuming the sensor outputs are available error-free at the FC.

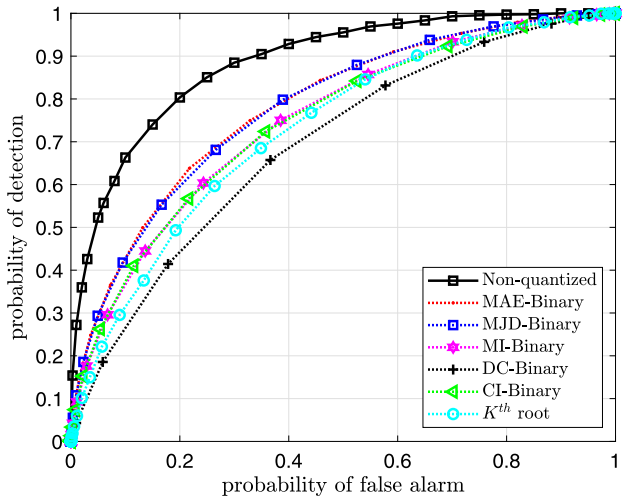

Fig. 9. Comparison between the ROC curves obtained using MAE, MJD, CI, DC, MI and K^{th} root methods for binary and Gaussian (non-quantized) DDT.

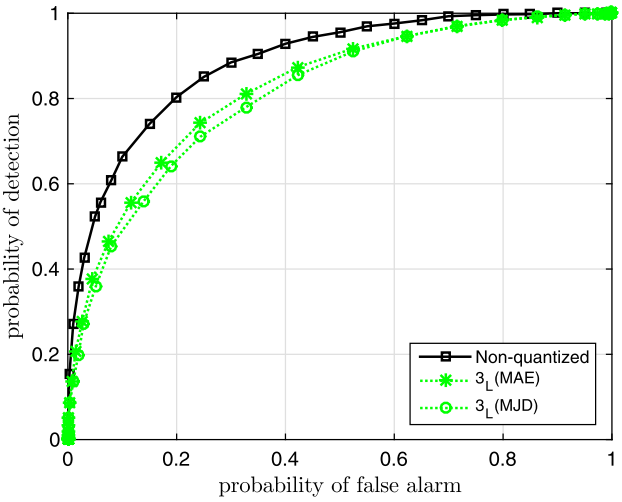

Fig. 10. Comparison between the ROC curves obtained using MAE and MJD methods for three-level quantized and Gaussian (non-quantized) DDT.

Then, a Rayleigh fading channel is considered to show the channel effect on the performance of our proposed quantization method, MAE. It is important to notice that, in distributed detection systems, randomized tests are not optimum when the local likelihood ratios contain no point masses of probability [51]. Nevertheless, dashed lines for ROCs of our simulation results correspond to randomization at the FC to achieve a particular p_{fa}.

4.1. Binary quantization and DDT

In Fig. 9, the Receiver Operating Characteristics (ROC), that is p_{d} versus p_{fa}, curves are plotted for the cases of using the quantization intervals from DC, CI, MI, K^{th} root, MJD and MAE methods for binary quantization with DDT and the corresponding non-quantized data transmissions. We also used Kth root quantization, which determines the local threshold assuming that AND-rule is applied at the FC. AND-rule corresponds to setting the false alarm threshold at the FC to all "one"s coming from K sensors, and for the global $p_{\text{fa}} = 0.1$, the local threshold will be the threshold, which makes the local pmf $p_2^{H_0} = 0.912$ at each sensor. However, we apply the fusion rule given in (15) as for all other quantization methods in the DDT scenario. In this figure, we observe that MAE and MJD performed significantly better than other methods. MI and CI performed similarly, sharing the third and fourth-best place in the performance. K^{th} root method showed the fifth-best performance, and DC was the worst-performing method. The figure also depicts a slightly better performance of MAE-based method compared to MJD-based one. Additionally, we observe that they are clearly inferior to the non-quantized case, which shows that there is quite a large space for gain in using higher quantization levels. Considering the clear superiority of MAE and MJD compared to other methods, we pursued the subsequent simulation studies only using MAE and MJD methods.

4.2. Performance of MAE and MJD with multilevel quantization and DDT

The simulation performances for the three-level, four-level and six-level quantizations by using the MAE and MJD methods are also obtained for DDT. ROC curves obtained using MAE and MJD methods for three levels of quantization and non-quantized data are shown in Fig. 10. This figure depicts that at global $p_{\text{fa}} = 0.2$, the p_{d}, attains the values $0.653, 0.684$ and 0.803 for the cases of three-level data transmissions with MJD, MAE and the non-quantized data transmission, respectively. Increasing the quantization level makes the MAE and MJD methods perform closer to the performance without quantization depicted in Tables 1 and 2.

Table 1 shows the p_{d} for 2, 3, 4 and 6 level MAE and MJD based quantized and non-quantized data transmissions for the values of $p_{\text{fa}} = 0.1, 0.2, 0.3$ and 0.4. The MAE method performs better at each quantization level than MJD, and the performance increases when the quantization level increases. At 6-level quantization p_{d} obtained by the MAE-based method is only slightly inferior to the limiting case with no quantization. Quantitatively, the difference in p_{d} is 0.022, 0.014, 0.018 and 0.002 for p_{fa} values of 0.1, 0.2, 0.3 and 0.4, respectively. Note that, at 6 level quantization for $p_{\text{fa}} = 0.4$, the performance with MAE is only 0.002 worse compared to the best possible p_{d}, which is the case with no quantization. The corresponding performance difference is 0.021 with MJD, approximately ten times the performance difference with MAE. Table 2 shows the achieved gain in p_{d} by using the MAE method with respect to MJD method and is given by $G = \left(p_{\text{d}}^{\text{MAE}_i} - p_{\text{d}}^{\text{MJD}_i}\right)$ with the resulting percentage gain $PG = (G \times 100\%)/p_{\text{d}}^{\text{MAE}_i}$, where $i = 2, 3, 4, 6$.

The previous tables show that MAE outperforms MJD for $M \geq 2$ levels. The achieved gain of MAE with respect to MJD is on average 0.0138 with a corresponding percentage gain of 2.13% for binary quantization. Likewise, the average gains are $= 0.0305$, 0.0318 and 0.0168 with corresponding average percentage gains of 4.31%, 4.12% and 2.09% for 3-level, 4-level and 6-level quantizations, respectively. In the same manner, the average difference in p_{d}, for $p_{\text{fa}} = 0.1, 0.2, 0.3$ and 0.4, between the 6-level quantizations achieved by MAE and no quantization equals 0.014 and 1.8%. The corresponding differences are 0.03 and 3.9% for MJD. These results show that with 6-level quantization, MAE performs nearly as well as no quantization, achieving better performance than the MJD method.

4.3. Multiple level quantization and Rayleigh fading channel

Fig. 11 shows the ROC curves for 2, 3, 4 and 6 levels MAE-based quantized and non-quantized data transmissions by using M-FSK modulation scheme with non-coherent demodulation over Rayleigh fading channels and using the optimal fusion rule in (8). The threshold, η, for each p_{fa} was estimated by running a Monte Carlo simulation under no event case. This figure shows that the

Table 1
The relation between p_d and p_{fa} for different levels of quantization obtained with MAE and MJD methods.

p_{fa} \ p_d	MJD_2	MAE_2	MJD_3	MAE_3	MJD_4	MAE_4	MJD_6	MAE_6	non-quantized
0.1	0.425	0.432	0.497	0.520	0.567	0.592	0.629	0.643	0.665
0.2	0.590	0.610	0.653	0.684	0.728	0.760	0.772	0.789	0.803
0.3	0.710	0.720	0.755	0.790	0.810	0.845	0.850	0.867	0.885
0.4	0.787	0.805	0.825	0.858	0.860	0.895	0.903	0.922	0.924

Table 2
Achieved gain in p_d by using the MAE method in quantization instead of MJD.

p_{fa} \ p_d	2-Level		3-Level		4-Level		6-Level	
	Gain	%Gain	Gain	%Gain	Gain	%Gain	Gain	%Gain
0.1	0.01	1.63	0.023	4.42	0.025	4.22	0.014	2.17
0.2	0.02	3.28	0.031	4.53	0.032	4.21	0.017	2.15
0.3	0.01	1.39	0.035	4.43	0.035	4.14	0.017	1.96
0.4	0.018	2.2	0.0330	3.85	0.035	3.914	0.019	2.06

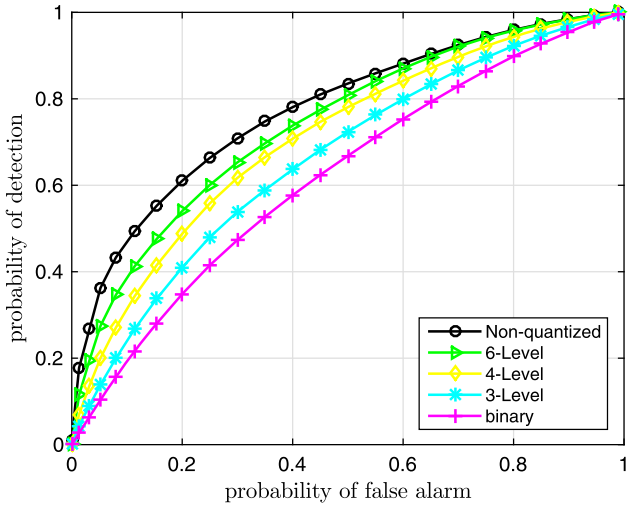

Fig. 11. ROC curves in the case of fading channel by using MAE based quantization and optimum fusion rule.

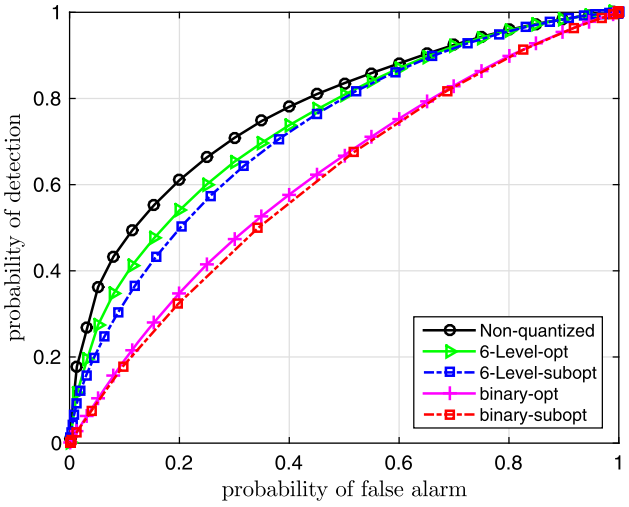

Fig. 12. A comparison between the ROCs of the optimal and sub-optimal fusion rule for binary and six level quantizations and the corresponding non-quantized data transmissions in the case of fading channel.

obtained p_d for 6-level quantization falls behind the limiting case of no quantization by 0.09 at $p_{fa} = 0.1$. This gain diminishes at $p_{fa} = 0.7$. When we compare the performances at different quantization levels, the achieved gain in p_d by transmitting 6-level quantized data instead of 2-level quantization is 0.21 for $p_{fa} = 0.1$ and this gain diminishes at $p_{fa} = 0.99$. Also, the sub-optimal fusion rule in (15) has been used to find the ROCs for the different levels of quantization. Fig. 12 shows a comparison between the optimal and sub-optimal fusion rule for 2 and 6 level quantization and compare them with the case of non-quantized data. The dashed line in ROCs for the sub-optimal fusion rule corresponds to randomization in the tests. This figure shows that the achieved gain using the optimum fusion rule with respect to the sub-optimal rule is 0.3 and 0.6 at $p_{fa} = 0.1$ for 2-level and 6-level quantizations, respectively.

5. Conclusion

In this study, we have proposed quantizing the sensor outputs by maximizing their average information in the presence and non-presence of an event in decentralized detection.

Although there are methods of maximizing information such as MI and CMI, quantization for decision processes has generally been based on distance measures such as JD and Bhattacharyya distance. Inspired by an intuitive idea of maximizing the information in the decision process MAE method of quantization is proposed.

One reason for suggesting another method like MAE instead of MJD is the non-symmetric nature of the considered problem, that is, probability of false alarm and probability of a miss are not equally important, and the fact that the advantage of Ali-Silvey type criteria [27] which MJD is a member of, is only valid for the symmetric performance measure probability of error. Although maximizing the transferred information under each hypothesis proposed by the MAE method is a conceptually different approach, we showed that average entropy and JD are positively proportional quantities. This means that one might expect comparable performances using either of them for determining the quantization levels, which was indeed the observation in the simulation results.

In order to spotlight the effects of how the sensor outputs are quantized on the system performance, we performed extensive simulation studies for the case that the sensor outputs are available error-free at the FC, which we called DDT. We adopted DC, CI, MI, CMI and MJD to compare the proposed method. In the simulation studies with binary quantization, the performances of DC, CI, MI and CMI were significantly inferior to MAE. A possible reason is that these methods are developed for known signal power or SNR at the sensors. Consequently, we continued further comparisons only with MJD.

The performances of considered information-based methods, namely MAE and MJD, gradually improved as the quantization

level was increased from binary to six-levels, and they approached the version of non-quantized data transmission. The proposed method, MAE, also performed significantly better than MJD for any quantization level. Additionally, the effects of the Rayleigh fading channel from the sensors to the FC have been investigated using the optimal and a sub-optimal fusion rule for MAE. Due to the power efficiency and small degradation in non-coherent communication, MFSK was adopted as the sensor to FC communication modulation scheme. In the case of the wireless channel model, similar results were obtained as in the case of DDT. Results with 6-level quantization were comparable to non-quantized data transmission.

This work showed that MAE is a valid and promising method in quantization for detection problems. A possible future work will be applying the MAE quantization method to discriminate between M-hypotheses, where $M > 2$. This type of problem is important, in which one of M signals needs to be detected. Also, it frequently arises in pattern recognition systems to distinguish between different patterns. Additionally, this work considers a parallel network topology, whereas, in WSNs, there are other network topologies to investigate, such as tree and serial topologies.

CRediT authorship contribution statement

Muath A. Wahdan: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **Mustafa A. Altınkaya:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Muath A. Wahdan was supported by Türkiye Bursları/Scholarship under grant number 13PS148104 for his Ph.D. studies at İzmir Institute of Technology.

References

[1] J. Hill, M. Horton, R. Kling, L. Krishnamurthy, The platforms enabling wireless sensor networks, Commun. ACM 47 (6) (2004) 41–46, https://doi.org/10.1145/990680.990705.
[2] P. Ragam, N.D. Sahebraoji, Application of mems-based accelerometer wireless sensor systems for monitoring of blast-induced ground vibration and structural health: a review, IET Wirel. Sens. Syst. 9 (3) (2019) 103–109, https://doi.org/10.1049/iet-wss.2018.5099.
[3] A. Rajput, V.B. Kumaravelu, Scalable and sustainable wireless sensor networks for agricultural application of internet of things using fuzzy c-means algorithm, Sustain. Comput.: Inform. Syst. 22 (2019) 62–74, https://doi.org/10.1016/j.suscom.2019.02.003.
[4] L. Muduli, D.P. Mishra, P.K. Jana, Application of wireless sensor network for environmental monitoring in underground coal mines: a systematic review, J. Netw. Comput. Appl. 106 (2018) 48–67, https://doi.org/10.1016/j.jnca.2017.12.022.
[5] H. Yu, J. Zhang, L. Zhang, S. Li, Polarimetric multiple-radar architectures with distributed antennas for discriminating between radar targets and deception jamming, Digit. Signal Process. 90 (2019) 46–53.
[6] M.A. Wahdan, M.F. Al-Mistarihi, M. Shurman, Static cluster and dynamic cluster head (scdch) adaptive prediction-based algorithm for target tracking in wireless sensor networks, in: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE, 2015, pp. 596–600.
[7] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, Internet of things for smart cities, IEEE Int. Things J. 1 (1) (2014) 22–32, https://doi.org/10.1109/JIOT.2014.2306328.
[8] M.A. Al-Jarrah, A. Al-Dweik, M. Kalil, S.S. Ikki, Decision fusion in distributed cooperative wireless sensor networks, IEEE Trans. Veh. Technol. 68 (1) (2018) 797–811, https://doi.org/10.1109/TVT.2018.2879413.
[9] D. Ciuonzo, G. Romano, P.S. Rossi, Optimality of received energy in decision fusion over Rayleigh fading diversity mac with non-identical sensors, IEEE Trans. Signal Process. 61 (1) (2012) 22–27, https://doi.org/10.1109/TSP.2012.2223694.
[10] R.R. Tenney, N.R. Sandell, Detection with distributed sensors, IEEE Trans. Aerosp. Electron. Syst. 4 (1981) 501–510, https://doi.org/10.1109/TAES.1981.309178.
[11] Z. Chair, P. Varshney, Optimal data fusion in multiple sensor detection systems, IEEE Trans. Aerosp. Electron. Syst. 1 (1986) 98–101, https://doi.org/10.1109/TAES.1986.310699.
[12] J. Tsitsiklis, Decentralized detection, Adv. Stat. Signal Process. 2 (2) (1993) 297–344.
[13] R. Viswanathan, P.K. Varshney, Distributed detection with multiple sensors part I. Fundamentals, Proc. IEEE 85 (1) (1997) 54–63, https://doi.org/10.1109/5.554208.
[14] J. Tsitsiklis, M. Athans, On the complexity of decentralized decision making and detection problems, IEEE Trans. Autom. Control 30 (5) (1985) 440–446, https://doi.org/10.1109/TAC.1985.1103988.
[15] R. Radner, Team decision problems, Ann. Math. Stat. 33 (3) (1962) 857–881.
[16] I.Y. Hoballah, P.K. Varshney, Distributed Bayesian signal detection, IEEE Trans. Inf. Theory 35 (5) (1989) 995–1000, https://doi.org/10.1109/18.42208.
[17] V.V. Veeravalli, P.K. Varshney, Distributed inference in wireless sensor networks, Philos. Trans. R. Soc. A, Math. Phys. Eng. Sci. 370 (1958) (2012) 100–117.
[18] R. Niu, P.K. Varshney, Distributed detection and fusion in a large wireless sensor network of random size, EURASIP J. Wirel. Commun. Netw. 2005 (4) (2005) 815873, https://doi.org/10.1155/WCN.2005.462.
[19] R. Niu, P.K. Varshney, Q. Cheng, Distributed detection in a large wireless sensor network, Inf. Fusion 7 (4) (2006) 380–394, https://doi.org/10.1016/j.inffus.2005.06.003.
[20] R. Niu, P.K. Varshney, Performance analysis of distributed detection in a random sensor field, IEEE Trans. Signal Process. 56 (1) (2007) 339–349, https://doi.org/10.1109/TSP.2007.906770.
[21] R. Niu, P.K. Varshney, Joint detection and localization in sensor networks based on local decisions, in: 2006 Fortieth Asilomar Conference on Signals, Systems and Computers, IEEE, 2006, pp. 525–529.
[22] D. Ciuonzo, P.S. Rossi, Distributed detection of a non-cooperative target via generalized locally-optimum approaches, Inf. Fusion 36 (2017) 261–274, https://doi.org/10.1016/j.inffus.2016.12.006.
[23] A. Tarighati, J. Jaldén, Bayesian design of tandem networks for distributed detection with multi-bit sensor decisions, IEEE Trans. Signal Process. 63 (7) (2015) 1821–1831, https://doi.org/10.1109/TSP.2015.2401535.
[24] R. Viswanathan, B. Ahsant, A review of sensing and distributed detection algorithms for cognitive radio systems, Int. J. Smart Sens. Intell. Syst. 5 (1) (2012).
[25] Z. Zhang, A. Pezeshki, W. Moran, S.D. Howard, E.K. Chong, Error probability bounds for balanced binary relay trees, IEEE Trans. Inf. Theory 58 (6) (2012) 3548–3563, https://doi.org/10.1109/TIT.2012.2187323.
[26] Z. Zhang, E.K. Chong, A. Pezeshki, W. Moran, S.D. Howard, Submodularity and optimality of fusion rules in balanced binary relay trees, in: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), IEEE, 2012, pp. 3802–3807.
[27] H. Poor, J. Thomas, Applications of Ali-Silvey distance measures in the design generalized quantizers for binary decision systems, IEEE Trans. Commun. 25 (9) (1977) 893–900, https://doi.org/10.1109/TCOM.1977.1093935.
[28] H.V. Poor, A companding approximation for the statistical divergence of quantized data, in: The 22nd IEEE Conference on Decision and Control, IEEE, 1983, pp. 697–702.
[29] H.V. Poor, Fine quantization in signal detection and estimation, IEEE Trans. Inf. Theory 34 (5) (1988) 960–972, https://doi.org/10.1109/18.21219.
[30] C.-C. Lee, J.-J. Chao, Optimum local decision space partitioning for distributed detection, IEEE Trans. Aerosp. Electron. Syst. 25 (4) (1989) 536–544.
[31] D. Warren, P. Willett, Optimum quantization for detector fusion: some proofs, examples, and pathology, J. Franklin Inst. 336 (2) (1999) 323–359, https://doi.org/10.1016/S0016-0032(98)00024-6.
[32] B. Picinbono, P. Duvaut, Optimum quantization for detection, IEEE Trans. Commun. 36 (11) (1988) 1254–1258, https://doi.org/10.1109/26.8934.
[33] C. Altay, H. Delic, Optimal quantization intervals in distributed detection, IEEE Trans. Aerosp. Electron. Syst. 52 (1) (2016) 38–48, https://doi.org/10.1109/TAES.2015.140551.
[34] I. Hoballah, P. Varshney, An information theoretic approach to the distributed detection problem, IEEE Trans. Inf. Theory 35 (5) (1989) 988–994, https://doi.org/10.1109/18.42216.
[35] Y.I. Han, T. Kim, Mutual and conditional mutual informations for optimizing distributed Bayes detectors, IEEE Trans. Aerosp. Electron. Syst. 37 (1) (2001) 147–157, https://doi.org/10.1109/7.913674.
[36] X. Cheng, D. Ciuonzo, P.S. Rossi, Multi-bit decentralized detection through fusing smart & dumb sensors based on Rao test, IEEE Trans. Aerosp. Electron. Syst. 56 (2) (2020) 1391–1405, https://doi.org/10.1109/TAES.2019.2936777.
[37] B. Chen, R. Jiang, T. Kasetkasem, P.K. Varshney, Fusion of decisions transmitted over fading channels in wireless sensor networks, in: Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002, vol. 2, IEEE, 2002, pp. 1184–1188.

[38] B. Chen, R. Jiang, T. Kasetkasem, P.K. Varshney, Channel aware decision fusion in wireless sensor networks, IEEE Trans. Signal Process. 52 (12) (2004) 3454–3458, https://doi.org/10.1109/TSP.2004.837404.
[39] B. Chen, P.K. Willett, On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels, IEEE Trans. Inf. Theory 51 (2) (2005) 693–699, https://doi.org/10.1109/TIT.2004.840879.
[40] R. Niu, B. Chen, P.K. Varshney, Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks, IEEE Trans. Signal Process. 54 (3) (2006) 1018–1027, https://doi.org/10.1109/TSP.2005.863033.
[41] J.H. Kotecha, V. Ramachandran, A.M. Sayeed, Distributed multitarget classification in wireless sensor networks, IEEE J. Sel. Areas Commun. 23 (4) (2005) 703–713, https://doi.org/10.1109/JSAC.2005.843539.
[42] B. Liu, A. Jeremic, K.M. Wong, Optimal distributed detection of multiple hypotheses using blind algorithm, IEEE Trans. Aerosp. Electron. Syst. 47 (1) (2011) 317–331, https://doi.org/10.1109/TAES.2011.5705678.
[43] N. Maleki, A. Vosoughi, Channel-aware m-ary distributed detection: optimal and suboptimal fusion rules, in: 2012 IEEE Statistical Signal Processing Workshop (SSP), IEEE, 2012, pp. 644–647.
[44] Z. Hajibabaei, A. Vosoughi, Impact of wireless channel uncertainty upon m-ary distributed detection systems, in: 2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), IEEE, 2014, pp. 692–696.
[45] H. Kasasbeh, L. Cao, R. Viswanathan, Soft-decision based distributed detection over correlated sensing channels, in: 2017 51st Annual Conference on Information Sciences and Systems (CISS), 2017, pp. 1–6.
[46] S. Zhang, L.N. Theagarajan, S. Choi, P.K. Varshney, Fusion of correlated decisions using regular vine copulas, IEEE Trans. Signal Process. 67 (8) (2019) 2066–2079, https://doi.org/10.1109/TSP.2019.2901379.
[47] D. Messerschmitt, Quantizing for maximum output entropy, IEEE Trans. Inf. Theory 17 (5) (1971) 612, https://doi.org/10.1109/tit.1971.1054681.
[48] M.A. Wahdan, M.A. Altınkaya, Optimal quantization in decentralized detection by maximizing the average entropy of the sensors, in: 2019 27th Signal Processing and Communications Applications Conference (SIU), IEEE, 2019, pp. 1–4.
[49] D. Ciuonzo, G. Romano, P.S. Rossi, Channel-aware decision fusion in distributed mimo wireless sensor networks: decode-and-fuse vs. decode-then-fuse, IEEE Trans. Wirel. Commun. 11 (8) (2012) 2976–2985, https://doi.org/10.1109/TWC.2012.061912.112049.
[50] J.-F. Bercher, C. Vignat, Estimating the entropy of a signal with applications, IEEE Trans. Signal Process. 48 (6) (2000) 1687–1694, https://doi.org/10.1109/78.845926.
[51] P. Willet, D. Warren, The suboptimality of randomized tests in distributed and quantized detection systems, IEEE Trans. Inf. Theory 38 (3) (1992) 355–361, https://doi.org/10.1109/18.119692.

Muath A. Wahdan received his B.S. degree in Electrical Engineering from Palestine Technical University - Kadoorie, Palestine in 2009 and M.S. degree in Electrical Engineering, specializing in Wireless Communications from Jordan University of Science and Technology, Jordan in 2013. Later he pursued his Ph.D. degree in Electronics and Communication Engineering from İzmir Institute of Technology, Turkey in 2021. From 2009 to 2012, he was with the Department of Electrical Engineering at Jordan University of Science and Technology, a Teaching Assistant, Jordan. From 2013 to 2014 he worked as a Lecturer in the Department of Communication Engineering at Palestine Technical University, Palestine. In 2021, he joined Electrical Engineering Department of Palestine Technical University - Kadoorie, where he is currently beholding the position of Assistant Professor. His research interest mostly involves Statistical Signal Processing and its applications in Telecommunications, distributed detection and WSNs.

Mustafa A. Altınkaya received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 1987, 1990, and 1996, respectively. From 1989 to 1990, he was a research associate with the TUBITAK Marmara Research Center, Kocaeli, Turkey. From 1991 to 1996, he was a research and teaching assistant with Boğaziçi University. He was a part-time instructor in Air Force Academy, Naval Academy, and Boğaziçi University, from 1995 to 1998. In 1998, he joined the Electrical and Electronics Engineering Department, İzmir Institute of Technology, where he is currently a professor. He was an associate editor of the Turkish Journal of Electrical Engineering and Computer Sciences from 2018 to 2019. His research interests are mainly on statistical signal processing and its applications in telecommunications, distributed detection, positioning and structural dynamics systems.