Abstract—Signal transmission and information fusion in wireless sensor networks (WSNs) are conventionally assumed to operate over orthogonal channels, which makes the network bandwidth and throughput inefficient. To remedy this inefficiency and improve the performance of the WSNs, we consider complex field network-coded (CFNC) relay-assisted communications, which operates over nonorthogonal channels and provides both spatial and temporal diversity. We derive the optimal likelihood ratio test-based fusion rule for the considered system. To provide robustness against the multiaccess interference, each sensor in the CFNC-coded system is assigned to a unique predetermined signature. Hence, the signature selection and the relay power allocation become crucial factors affecting the performance of the WSNs. We also develop an analytical method to jointly adjust the sensor signatures and the relay power utilizing the average symbol error rate bound of the network together with some information theoretical results. Finally, we evaluate the detection performance of the proposed scheme and compare it with that of the conventional method. The simulation results suggest that the proposed signature selection and relay power allocation method considerably improves the network performance.

Index Terms—Distributed detection, fading, relaying, network coding, wireless sensor networks.

I. INTRODUCTION

RECENTLY, wireless sensor networks (WSN) have attracted much interest to develop for various applications such as environmental monitoring, health, surveillance and robotic exploration etc [1]. A WSN consists of low powered sensors, which have sensing, processing and communication capabilities. Typically, the sensors within the network work collaboratively to monitor an event by acquiring observations in the region of interest (ROI), from which they extract some useful information to send a special node named as “fusion center” (FC), which combines all the sensor outputs and forms a global situational assessment.

Distributed detection (DD) becomes attractive for WSNs with limited resources [2], because directly sending the raw data to the FC consumes more bandwidth and energy.

In the DD, the sensors first make their own decisions. Then, these decisions are transmitted to the FC to be fused optimally.

In the literature, the distributed detection has been extensively investigated. The optimal distributed signal detection has been analyzed with the assumption of conditional independence under both Bayesian and Neyman-Pearson (N-P) criteria in [3]–[5]. Moreover, the studies in [6]–[9] aimed to investigate the decision fusion, when the observations are correlated. It is important to note, however, that the communication channel between each sensor node and the FC is assumed to be error-free. Thomopoulos and Zhang [10] came up with the idea of distributed detection in non-ideal communication channels, where only the effect of the noise is considered by assuming that the communication channel between each sensor and the fusion center is a binary symmetric channel (BSC). Since fading in WSNs is another source of degradation in addition to the receiver electronics noise, Chen et al. [11] proposed the optimum and sub-optimum fusion rules for a noisy and Rayleigh fading WSN with parallel structure when the perfect channel state information (CSI) (i.e., fading coefficient) is assumed to be exactly known at the fusion center. This has later been extended in [12] to the case of the only knowledge of the channel envelope statistics (CS). Later, Eritmen et al. developed distributed decision fusion under both the perfect CSI and CS for WSNs with a hierarchical topology [13].

The distributed detection (DD) is very useful to combat the adverse effects of fading channels (e.g., multipath-fading, shadowing, noise) in wireless sensor networks, because the sensor nodes reaching the same decision provide spatial diversity since the FC is supplied with multiple copies of the transmitted signal over different channels that fade almost independently.

It is important to mention that in all the previously mentioned studies in [2]–[13], the signal transmission and the information fusion are assumed to be accomplished by using orthogonal communications (OC) to avoid multi-access interference (MAI), where only one sensor sends its decision at a certain time, while the others wait. This, however, decreases the throughput of the system, and turns out to be bandwidth inefficient particularly for large networks.

To improve the throughput efficiency of WSNs, the decision fusion with the use of non-orthogonal communications (NOC) has been analyzed in [15]–[18], where the local sensors directly send their decisions to the FC over a multiple-access channel (MAC).
The performance of the DD in WSNs can be also enhanced with the use of relaying, which has been proposed to achieve both spatial and time diversity in multi-user communications [14]. In the relaying process, the sensors first send their decisions over wireless medium, and because of the broadcast nature of the wireless channel both the relay node and the FC hear the information bearing signals of the sensors. Then, the relay node extracts the necessary information about the sensor decisions. Finally, in the subsequent time interval, the relay node forwards its signal to the FC by employing a relaying protocol [14].

Additionally, the use of relaying together with transmitting sensor decisions over a MAC channel improves both the reliability and throughput of WSNs, which can be further increased by employing a network-coding (NC) scheme [24] such as physical layer network coding (PNC) (a.k.a analog network coding), and complex field network coding (CFNC) [19]–[24] etc.

It is important to point out that although the works in ([15]–[18]) show that non-orthogonal signaling has a potential to improve error performance distributed detection and none of them has considered the use of relaying. Therefore, in this work, we consider CFNC coded relaying assisted communications over a MAC in a WSN with a parallel topology.

In the initial version of this work in [25], we proposed the idea of the use of the CFNC for relaying assisted communications over a MAC since the CFNC provides the highest throughput (1/2 symbol per user per channel-use) compared to PNC and XOR methods for communications over non-orthogonal channels since it uniquely allows decoding of user messages under multi-access interference (MAI) [24].

In CFNC, each sensor is assigned to a unique pre-determined complex number, which is referred as signature. Each signature is used to weight the signal of a particular sensor before the signal transmission, which provides robustness against the multi-access interference (MAI). Hence, the signature selection appears to become one of the important issues to improve the energy efficiency of the system considered. Wang et al. [24] selected signatures based on linear constellation precoding, which is purely complex exponential and distinctively rotates the constellation of each sensor. In contrast to [24], one can also employ signatures with non-unity magnitudes and optimize them according to a certain criterion to enhance the system performance while keeping the average transmit power of the network limited. In addition to signature optimization, the performance of the network can be further improved by appropriately allocating the relay power. In [25], we proposed to optimize the sensor signatures and the relay power by minimizing the symbol error rate (SER) bound of the network, which provided however only numerical results since the problem is highly non-linear and there is no closed form solutions for the parameters that need to be optimized.

Contrary to [25], in this paper, we aim at providing analytical solutions to the problem of jointly optimizing the sensor signatures and the relay power. Accordingly, we first derive the optimal LRT based fusion rule for a parallel WSN with a relay node that operates over non-orthogonal wireless channels. Then, we utilize the SER bound of the network together with information theoretical results, and make a series of approximations to determine sensor signature and the relay power. Finally, we have shown that the numerical experiments conducted with the proposed method significantly outperforms the classical approach.

In the next Section, we give some background on the classical distributed detection (CDD) for a parallel WSN. In Section III, the system model for CFNC coded relay assisted communications in WSNs is described, and subsequently, the LRT based optimum fusion rule is derived for the considered system. In Section IV, we present our analytical method for the selection of the sensor signatures and the relay power. Section V is devoted to investigate the performance of the proposed method through numerical simulations. Finally, our conclusions are included in Section VI.

II. OVERVIEW OF CLASSICAL DISTRIBUTED DETECTION
OVER ORTHOGONAL COMMUNICATION CHANNELS

In this part, we review the classical distributed detection (CDD) for orthogonal signaling (without a relay node) in WSNs with a parallel topology, where we focus on binary hypotheses: $H_1$ and $H_0$ (e.g., they may represent the existence and absence of a target, respectively) at the region of interest (ROI). A network of $N$ sensors and a FC (without a relay node) is considered as shown in Fig. 1, where the $k$th sensor, $s_k$, first acquires its measurement $m_k$ from the ROI and arrives at its decision $u_k$, which is later modulated to produce $x_k$. Then, the modulated signal $x_k$ is distorted by the fading and noise and produces the signal $y_k$ at the FC. Finally, the FC combines all of the signals it has received according to a fusion rule, and casts a final decision $u_o$. We assume throughout the manuscript that all fading coefficients are modeled as complex Gaussian random variables with zero mean and unit variance and the receiver electronics noise is modeled as an additive white Gaussian noise (AWGN) channel. Furthermore, we also assume that the channel state information (CSI) is available at the FC. Note that we just consider binary phase shift keying (BPSK) modulation for the modulated signal, but the extension of the results to other modulation schemes is straightforward.
Under the orthogonal signaling, the received signal at the
the fusion due to the transmission of the $k^{th}$ sensor can becomes:

$$y_k = \sqrt{\gamma_k} h_k x_k + z_k$$ \hspace{1cm} (1)

where $y_k$ is the path loss coefficient of the link between
the $k^{th}$ sensor and FC; $h_k$ is complex Gaussian channel
coefficient, $x_k$ is the BPSK modulated signal, which takes
values of $-1$ and $1$, respectively for $k^{th}$ sensor decision $u_k$
being $0$ and $1$, respectively; $z_k$ is zero-mean AWGN sample
with variance of $\sigma^2 = N_0/2$ per dimension.

For this network topology and orthogonal signaling model,
the optimal likelihood ratio (LRT) based fusion rule is given
in [11]$^1$ as:

$$\Lambda(y) = \prod_{k=1}^{N} \frac{f(y_k|H_1, h_k)}{f(y_k|H_0, h_k)}$$

$$= \prod_{k=1}^{N} \frac{P_{D_k} e^{-\frac{|y_k|^2}{2\sigma^2}} + (1 - P_{D_k}) e^{-\frac{|y_k|^2}{2\sigma^2}}}{P_{F_k} e^{-\frac{|y_k|^2}{2\sigma^2}} + (1 - P_{F_k}) e^{-\frac{|y_k|^2}{2\sigma^2}}}$$ \hspace{1cm} (2)

where $y = [y_1 \cdots y_N]^T$ is the received signal vector,
$P_{D_k}$ and $P_{F_k}$ are the probability of detection and the probability
of false alarm of the sensor $S_k$, respectively, which can be expressed as:

$$P_{D_k} = P(u_k = 1|H_1)$$

$$P_{F_k} = P(u_k = 1|H_0)$$ \hspace{1cm} (3)

The conventional strategy for the parallel network assumes
that the transmission from each sensor is accomplished over
orthogonal channels in time. As a result one-symbol information
regarding the decision of each sensor is transmitted by
$N$ channel-uses, each of which has duration of $T_0$ seconds.
Therefore, the information rate or throughput for the CDD
using orthogonal signaling can be obtained as:

$$R_{CDD} = \frac{1}{NT_0} \text{ symbols/sec per sensor per channel use.}$$ \hspace{1cm} (4)

III. DISTRIBUTED DETECTION FOR COMPLEX FIELD
NETWORK CODED RELAY ASSISTED COMMUNICATION
OVER MULTI-ACCESS CHANNELS

In this section, we first propose to use relaying over a
MAC for WSNs and then incorporate complex field
network coding (CFNC) to a WSN. Afterwards, we inves-
tigate distributed detection for CFNC coded relay assisted
distributed communication channel.

As mentioned earlier, the use of relay is beneficial to achieve
cooperative diversity in order to combat the detrimental effects
of fading channels. Hence, in this work, we propose to
incorporate CFNC to the parallel WSN with a relay node ($R$)
as depicted in Fig. 2 where all of the information signals are
transmitted over MAC channels.

In this method, each sensor $S_k$ is assigned to a unique
signature $\theta_k$. Then the modulated decisions of sensors, $x_k$’s,

$^1$The phase coherent detection formulation in reference [11] is equivalent
to the complex representation in Eq.(2)

are multiplied by the associated signature and the resultant
signals of sensors are sent over non-orthogonal channels
simultaneously in time slot 1, which causes interference both
at the relay and the FC. After that, based on the relaying policy
(e.g., amplify and forward, estimate and forward etc.), the
relay node sends its output to the FC in time slot 2. Therefore,
the signals resultant from the non-orthogonal communications
under the flat-fading can be written as

$$y_r = \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sr} \theta_k x_k + z_r$$ \hspace{1cm} (5)

$$y_{sd} = \sum_{k=1}^{N} \sqrt{\gamma_k} h_{skd} \theta_k x_k + z_d$$ \hspace{1cm} (6)

$$y_{rd} = \sqrt{\gamma_r} h_{rd} \sqrt{\alpha} x_r + z_d$$ \hspace{1cm} (7)

where $y_r$, $y_{sd}$, and $y_{rd}$ are the received signals at the relay node
and the FC in time slot 1, respectively; $y_{rd}$ is the signal acquired at the FC in time slot 2 due to the relaying; $h_{sr}$, $h_{skd}$, and $h_{rd}$
are the fading gains of $S_k$-FC, $S_k$-RC and R-FC links,
respectively, which are modeled as complex Gaussian random
variables with zero mean and unit variance; the parameter $\alpha$
determines the power allocated to the relay as a fraction of
the total transmit power of sensors; $\gamma_k$, $g_k$ and $g_r$ denote
the path-loss coefficients of $S_k$-FC, $S_k$-R, and R-FC links,
respectively; $z_r$, $z_d$ represent the noise samples at the relay node
and FC, respectively, which are modeled as additive white
Gaussian noise (AWGN) with zero mean and variance of
$N_0/2$ per dimension.

Hence, each sensor transmits one-symbol information in two
channel-uses each with a duration of $T_0$ seconds, for which the
information-rate becomes

$$R_{CFNC} = \frac{1}{2T_0} \text{ symbols/sec per sensor per channel use.}$$ \hspace{1cm} (8)

It is important to point out that the channel state information
(CSI) is assumed to be known at all receiving nodes
(i.e., the relay node and the FC), in this work. As in [24],
we employ estimate and forward type of relaying based on
maximum likelihood (ML) detection (i.e., ML relaying). For this relaying policy, the sensor messages are estimated as

\[ \hat{x}(1, \ldots, \hat{x}_N) = \arg \min_{x_1, \ldots, x_N} \left\| y_r - \sum_{k=1}^{N} \sqrt{g_k h_{x_k} \theta_k x_k} \right\|_2^2 \]  

(9)

Then, the relay signal is generated by incorporating the ML estimates of sensor messages with the sensor-signatures as

\[ x_r = \sum_{k=1}^{N} \theta_k \hat{x}_k \]  

(10)

After that, the relay signal \( x_r \) is forwarded to the FC according to (7) in time slot 2. Finally, the FC combines all of the signals it has received in time slot 1 and time slot 2 in Bayesian sense. Specifically, the optimal LRT based fusion rule is

\[ \Lambda(y_d) = \frac{\sum_{x} f(y_{sd}|H_0, x) f(y_{rd}|H_1, x) P(\hat{x}|x) P(x|H_1)}{\sum_{x} f(y_{sd}|H_0, x) f(y_{rd}|H_1, x) P(\hat{x}|x) P(x|H_0)} \]  

(11)

where \( y_d = [y_{sd}, y_{rd}] \), \( x = [x_1, \ldots, x_N]^T \). Also, \( P(\hat{x}|x) \) is the probability that the relay decides \( \hat{x} \), although \( x \) is transmitted.

To simplify the analysis, we assume that the decoding at the relay is perfect\(^2\) (i.e., \( \hat{x} = x \)), and thus, the following LRT rule is employed at the FC.

\[ \Lambda(y_d) = \frac{\sum_{x} f(y_{sd}|H_1, x) f(y_{rd}|H_1, x) P(x|H_1)}{\sum_{x} f(y_{sd}|H_0, x) f(y_{rd}|H_0, x) P(x|H_0)} \]  

(12)

Since the conditional probability density function of \( y_{sd} \) (\( y_{rd} \)) is independent of hypothesis, when \( x \) is given and sensor decisions are conditionally independent, the conditional distributions or probabilities in Eq. (12) can be written as

\[ f(y_{sd}|x) = \frac{1}{2 \pi \sigma^2} \exp\left\{ -\frac{|y_{sd} - \sum_{k=1}^{N} \sqrt{g_k h_{x_k} \theta_k x_k}|^2}{2 \sigma^2} \right\} \]  

(13)

\[ f(y_{rd}|x) = \frac{1}{2 \pi \sigma^2} \exp\left\{ -\frac{|y_{rd} - \sqrt{g h_{x_r} \theta x_r}|^2}{2 \sigma^2} \right\} \]  

(14)

\[ P(x|H_1) = \prod_{k=1}^{N} P_{D_{k}}(1 - P_{D_{k}})^{1-u_k} \]  

(15)

The FC generates its final decision \( u_0 \) as

\[ \Lambda(y_d) \gtrless_{u_0=0}^{u_0=1} \tau \]  

(16)

where \( \tau \) is the optimal threshold value used at the FC. In this paper, we consider minimum error probability detection at the FC. Hence, the optimal threshold value can be determined in terms of a priori probability of the event at the ROI as

\[ \tau = \frac{P(H_0)}{P(H_1)} \]  

(17)

where \( P(H_0) \) and \( P(H_1) \) are a priori probabilities of \( H_0 \) and \( H_1 \) respectively. The average probability of error of the network can be determined as

\[ P_e = P(H_0) P_F + P(H_1)(1 - P_D) \]  

(18)

where \( P_F \) and \( P_D \) are the false alarm and detection probability of the FC, respectively, which can be expressed as

\[ P_F = P(\Lambda(y_d) > \tau | H_0) \]  

(19)

\[ P_D = P(\Lambda(y_d) > \tau | H_1) \]  

(20)

Note that the complexity of the fusion in CDD is linear in \( N \), whereas its complexity in CFNC-DD is exponential in \( N \). Although the CFNC-DD seems to be more complex, it has a better detection performance than CDD as we shall show in Section V. This computation burden can be alleviated with the use of clustering [26], [27].
IV. OPTIMAL POWER ALLOCATION FOR COMPLEX FIELD NETWORK CODED RELAY ASSISTED COMMUNICATION OVER MULTI-ACCESS CHANNELS

In this section, we propose a way to select the sensor-signatures optimally by minimizing the ML bound on the symbol error probability of the network. We denote the vector of signatures by \( \theta = [\theta_1, \ldots, \theta_N]^T \). Since each sensor decision is binary, there are \( 2^N \) possibilities for the sensor decision vector which can be put in an ordered list. Let \( x_i \) be the \( i \)th possible decision vector in the list for \( 1 \leq i \leq 2^N \). The CFNC coded symbol for the decision vector \( x_i \) due to the non-orthogonal signaling becomes

\[
c_i = \theta^T x_i
\]

(21)

Hence, the CFNC coded symbol takes \( 2^N \) distinct values.

Assuming CSI is known at the relay and ML relaying is used, the pair-wise symbol error probability (PEP) of the relay node is

\[
\text{PEP}'(c_i, c_j) \triangleq P\left(c_i \rightarrow c_j \mid c_i, h_{sr}\right) = Q\left(\frac{\left|\sum_{k=1}^{N} \sqrt{g_k} h_{sr} \theta_k d_{ijk}\right|}{2\sigma}\right)
\]

(22)

where \( h_{sr} = [h_{sr}, \ldots, h_{sr}]^T \). \( P\left(c_i \rightarrow c_j \mid c_i, h_{sr}\right) \) is the probability of deciding symbol \( c_j \) given that symbol \( c_i \) is transmitted under the CSI vector \( h_{sr} \), \( Q(x) \equiv (1/\sqrt{2\pi}) \int_x^{\infty} \exp(-t^2/2) dt \), \( d_{ij} \) is the difference between the \( i \)th and \( j \)th decision vectors (i.e., \( d_{ij} = (x_i - x_j) \)), and \( d_{ijk} \) represents the \( k \)th component of \( d_{ij} \). Hence, the instantaneous CFNC symbol error rate (SER) at the relay can be bounded as

\[
P_{e}'(h_{sr}) \leq \frac{1}{2} \sum_{i=1}^{2^N} \sum_{j=1, j \neq i}^{2^N} P(c_i) Q\left(\frac{\left|\sum_{k=1}^{N} \sqrt{g_k} h_{sr} \theta_k d_{ijk}\right|}{2\sigma}\right)
\]

(23)

where \( P(c_i) \) is the probability of the CFNC coded symbol for \( 1 \leq i \leq 2^N \), which depends on the false alarm and detection probabilities of the sensors as follows

\[
P(c_i) = P(c_i|H_0) P(H_0) + P(c_i|H_1) P(H_1)
\]

\[
= P(x_i|H_0) P(H_0) + P(x_i|H_1) P(H_1)
\]

\[
= \prod_{k=1}^{N} P_{F_k}^{x_i} (1 - P_{F_k})^{1 - x_k} P(H_0)
\]

\[
+ \prod_{k=1}^{N} P_{D_{F_k}}^{x_i} (1 - P_{D_{F_k}})^{1 - x_k} P(H_1)
\]

(24)

By using the Chernoff-bound [26] (i.e., \( Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}} \)), the instantaneous PEP and SER of the relay node is further upper-bounded as:

\[
\text{PEP}'(c_i, c_j) \leq \frac{1}{2} e^{-\frac{N}{8\sigma^2} \sum_{k=1}^{N} \sqrt{g_k} h_{sr} \theta_k d_{ijk}^2}
\]

(25)

\[
P_{e}'(h_{sr}) \leq \frac{1}{2} \sum_{i=1}^{2^N} \sum_{j=1, j \neq i}^{2^N} P(c_i) e^{-\frac{N}{8\sigma^2} \sum_{k=1}^{N} \sqrt{g_k} h_{sr} \theta_k d_{ijk}^2}
\]

(26)

Consequently, bounds on the average PEP and SER can be obtained by averaging the upper-bounds in Eqs.(25) and (26) over fading gains of the sensors-to-relay links as:

\[
\text{APEP}'(c_i, c_j) \leq 0.5 + \frac{N}{8\sigma^2} \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2
\]

(27)

\[
P_{e}'(h_{sr}) \leq \frac{1}{2} \sum_{i=1}^{2^N} \sum_{j=1, j \neq i}^{2^N} P(c_i) \left(1 + \frac{N}{8\sigma^2} \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2\right)
\]

(28)

By employing ML sequence detection at the FC using the signals received in both time slots (i.e., \( y_{sd} \) and \( y_{rd} \)), the similar analysis results in following bounds on the average PEP and SER at the FC (please see Appendix II for details):

\[
\text{APEP}^D(c_i, c_j) \leq \frac{1}{2} \sum_{i=1}^{2^N} \sum_{j=1, j \neq i}^{2^N} P(c_i) \left(1 + \frac{N}{8\sigma^2} \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2\right)
\]

(29)

\[
\text{SER}^D \leq B_{e}^D(\theta, \alpha)
\]

\[
= \frac{1}{2} \left(1 + \frac{N}{8\sigma^2} \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2\right)
\]

(30)

where \( B_{e}^D(\theta, \alpha) \) denotes the SER upper bound at the FC. Note that, we do not consider any specific type of modulation while we are deriving this bound.

Authors in [25] proposed the determination of the sensor signatures and the relay power by minimizing the average SER bound in Eq. (30) under the constraints on the total transmit power and the network geometry as:

\[
\text{minimize } B_{e}^D(\theta, \alpha)
\]

such that

\[
Pt - \sum_{k=1}^{N} |\theta_k|^2 - \alpha \sum_{k=1}^{N} |\theta_k|^2 \geq 0
\]

\[
|\theta_k - \theta_l|^2 > 0 \quad \text{for } k \neq l
\]

\[
\alpha > 0
\]
where the first and second constraints are related to the total transmit power budget and the distinctiveness of the signatures, whereas the third constraint stems from the fact that the relay actively sends information.

As pointed out in [25], deriving closed form analytical results for the sensor signatures and the relay power are cumbersome since the Karush Khun Tucker (KKT) conditions for the convex program (please see the convexity proof in Appendix II) in Eq. (31) result in highly nonlinear equations.

Instead of pursuing this direction, we follow another approach, in which we have first expressed the signatures in the polar form as $\theta_i = \sqrt{P_i} e^{j\phi_i}$, where $P_i$ and $\phi_i$ represent the magnitude-square and phase of the $i^{th}$ sensor signature. It is important to note that the signature phases do not have any effect on the average pair-wise error probability (APEP) at the relay (due to the second term inside the sum of Eq.(28)), whereas they affect the average SER at the FC as seen from the second term inside the sum of Eq. (30). Following that, we heuristically select the phases of the signatures to increase the second term inside the sum of Eq. (30). Following that, we heuristically select the phases of the signatures to increase the separation among the constellations, which are generated by multiplying the modulated sensor decisions with the unit magnitude sensor signatures. This produces equally separated phases (using BPSK modulated decisions) as:

$$\phi_i = \frac{\pi}{N}(i - 1) \quad \text{for} \quad i = 1, \ldots, N$$

(32)

In order to optimize the signature magnitudes, we next consider a network of a sensor node called as super node (SN) network, which transmits CFNC symbol $\theta_{1X1} + \theta_{2X2} + \ldots + \theta_{NXN}$, a relay with a power control parameter $\alpha$, and a FC. The average transmit power, $P_{SN}$, of the super node should satisfy

$$P_{SN} = \sum_{k=1}^{N} P_k$$

(33)

which is because the energy of the decision symbols is unity.

By performing a similar analysis while deriving Eq.(29) and considering the worst symbol pair that gives the maximum APEP bound at the FC, the APEP of the SN network becomes

$$\text{APEP}_{SN}(c_i, c_j) \leq B_{SN}^D = \frac{0.5}{1 + \frac{\bar{g} P_{SN} d_{min}^2}{8\sigma^2}} + \frac{1}{1 + \frac{g_i a P_{SN} d_{min}^2}{8\sigma^2}} + \frac{0.5}{1 + \frac{\bar{g} P_{SN} d_{min}^2}{8\sigma^2}}$$

for all $i \neq j$

(34)

where $B_{SN}^D$ represents the maximum APEP bound of the SN network, $\bar{g}$ and $\bar{\gamma}$ represent the gains of the SN-R link and SN-FC link, respectively, and $d_{min}$ is the minimum distance in the constellation formed by $e^{j\phi_1 x_1} + e^{j\phi_2 x_2} + \ldots + e^{j\phi_N x_N}$, which is mathematically represented as:

$$d_{min} = \min_{i \neq j} |\tilde{c}_i - \tilde{c}_j|$$

(35)

where $\tilde{c}_i = [e^{j\phi_1}, e^{j\phi_2}, \ldots, e^{j\phi_N}]^T x_i$ (i.e., CFNC symbols obtained by normalizing the magnitude of each sensor signature to unity).

Assuming that the path gains of all sensor-to-relay links and sensor-to-FC links in the original $N$ node WSN are close to $\bar{g}$ and $\bar{\gamma}$, respectively (i.e., $g_1 \approx g_2 \approx \ldots \approx g_N \approx \bar{g}$ and $\gamma_1 \approx \gamma_2 \approx \ldots \approx \gamma_N \approx \bar{\gamma}$), the maximum APEP bound in Eq. (29) of the $N$ node network is close the maximum APEP bound of the SN network in Eq. (34), which implies

$$B_{SN}^D \approx B_{SN}^D$$

(36)

where $B_{SN}^D$ is the maximum of the APEP bound of the $N$ node network in Eq. (29).

Therefore, approximating the $N$-node WSN by the SN network is accurate when path gains of all sensor-to-relay links and sensor-to-FC links are close to $\bar{g}$ and $\bar{\gamma}$, respectively. For a general setting, in which these gains may vary, the average of path gains in dB scale is used to determine the path gains of SN-R link and SN-FC link, respectively, as:

$$\bar{g} \quad \text{(in dB)} = \frac{1}{N} \sum_{i=1}^{N} g_i \quad \text{(in dB)} \quad \text{and} \quad \bar{\gamma} \quad \text{(in dB)} = \frac{1}{N} \sum_{i=1}^{N} \gamma_i$$

(37)

where $\bar{g} = \frac{10}{\text{dB}}$ and $\bar{\gamma} = \frac{10}{\text{dB}}$. After determining the parameters of SN network in terms of parameters of $N$ node WSN, we are ready to determine optimal transmit powers of the SN and the relay node under the total transmit power constraint stated as:

$$(1 + \alpha) \left( \sum_{k=1}^{N} P_k \right) = (1 + \alpha) P_{SN} = P_T$$

(38)

The total power constraint in Eq. (38) can be used to replace $\alpha P_{SN}$ by $P_T - P_{SN}$ for the bound in Eq.(34), which results in

$$B_{SN}^D = \frac{0.5}{1 + \frac{\bar{g} P_{SN} d_{min}^2}{8\sigma^2}} + \frac{1}{1 + \frac{g_i a (P_T - P_{SN}) d_{min}^2}{8\sigma^2}} + \frac{0.5}{1 + \frac{\bar{g} P_{SN} d_{min}^2}{8\sigma^2}}$$

(39)

When $\frac{d_{min}^2}{\sigma^2}$ is high, a further simplification can be obtained as:

$$B_{SN}^D \approx \frac{32\sigma^4}{\bar{g} P_{SN} d_{min}^4} \frac{4\sigma^2}{g_i a (P_T - P_{SN}) d_{min}^2} \quad \text{(40)}$$

The optimal total power reserved for the transmission of all sensors, which minimizes the maximum pair-wise error probability bound in Eq. (40), is obtained by the derivative of Eq. (40) with respect to $P_{SN}$, and then by equating to zero as:

$$-\frac{32\sigma^4}{\bar{g} P_{SN}^2 g_i a (P_T - P_{SN}) d_{min}^4} + 32\sigma^4 \bar{g} P_{SN} g_i a (P_T - P_{SN})^2 d_{min}^4 - 4\sigma^2 \frac{g_i a^2 (P_T - P_{SN})^2}{\bar{g} P_{SN} d_{min}^2} = 0$$

(41)

Multiplying and dividing both sides of Eq. (41) by $P_{SN}^2 (P_T - P_{SN})^2 / 4$ leads to the following expression:

$$-8\sigma^4 \frac{(P_T - P_{SN})}{\bar{g} g_i a d_{min}^2} + \frac{8\sigma^4 P_{SN}}{\bar{g} g_i a^2} - \frac{\sigma^2 (P_T - P_{SN})^2}{\bar{g} d_{min}^2} = 0$$

(42)
The roots of the quadratic relationship in Eq.(42) with respect to \( P_{SN} \) can be found as:

\[
P_{SN} = P_T + \frac{(8\sigma^2/d_{\text{min}}^2) 2\bar{g} \pm \sqrt{(2\bar{g} g_r P_T + (8\sigma^2/d_{\text{min}}^2) 2\bar{g})^2 - 4\bar{g}^2 g_r^2 P_T^2}}{2\bar{g} g_r} \tag{43}
\]

The only root satisfying the power constraint \( 0 < P_{SN} < P_T \) is

\[
P_{SN} = P_T + \frac{(8\sigma^2/d_{\text{min}}^2) \bar{g}}{\sqrt{4\bar{g}^2 g_r P_T (8\sigma^2/d_{\text{min}}^2) + 4\bar{g}^2 (8\sigma^2/d_{\text{min}}^2)^2}} \tag{44}
\]

The approximate relay power control parameter is determined using the total power constraint in Eq.(38) as:

\[
\alpha \approx \frac{P_T}{P_T + \frac{(8\sigma^2/d_{\text{min}}^2) \bar{g}}{\sqrt{4\bar{g}^2 g_r P_T (8\sigma^2/d_{\text{min}}^2) + 4\bar{g}^2 (8\sigma^2/d_{\text{min}}^2)^2}}} - 1 \tag{45}
\]

Therefore, the SN network approximation allows us to determine the relay power parameter and the total transmit power of all sensors in Eq. (44), which does not, however, specify the individual transmit power of each sensor. The next sub-section is devoted to include an information theoretical power allocation method for each sensor by using the total transmit power result obtained in Eq. (44).

A. Information Theoretical Determination of Individual Sensor Powers

As mentioned above, the relay power and the total transmit power allocated to all sensors are found through the super node (SN) approximation. Unfortunately, this does not specify the individual sensor powers within the WSN. In this part, we develop an information theoretical approach to determine the individual sensor powers from the total power budget, which is obtained by the SN approximation.

While allocating an optimal power to each of the sensors is important, maintaining the fairness among sensors is also very crucial in realizing a practical communication network [33], which ensures that the access of any sensor to the network is not denied or overly penalized [34]. For resource allocation in communication systems, various fairness criteria are considered in the literature such as max-min fairness [35], proportional fairness [36] and fairness in information rate (a.k.a symmetric capacity) [37]. Since the symmetric capacity represents the fairest maximum common rate [37], we consider a fairness criterion based on the notion of symmetric capacity and aim to develop a fair power allocation policy for sensor nodes in this study, which ensures fairness among sensors in terms of their average rates and is referred as “average-rate fairness.”

For this purpose, we first consider the users-to-relay channel, which is a MAC channel so that the achievable average sum-rate of sensors (the average rates of sensor \( S_i \) is denoted by \( R_i \)) at the relay is upper bounded as:

\[
R_1 + R_2 + \cdots + R_N \leq \frac{1}{2} \log \left(1 + \frac{\sum_{i=1}^{N} g_i |h_{i,r}|^2}{2\sigma^2} \right) \tag{46}
\]

Note that the first line of Eq. (46) is because of the ergodic sum-capacity of the MAC channel, while the second line follows from the Jensen’s inequality since the ergodic sum-capacity is a log-concave function, and the third line is a result from the assumption of independently identically distributed (i.i.d) fading gains with unit energy. Also, the factor of \( 1/2 \) in the rate calculations results from the normalization due to the use of two time-slots during the communications.

As a result, the average (or ergodic) single user rate bounds can be obtained as:

\[
R_i \leq \frac{1}{2} \log \left(1 + \frac{g_i P_i}{2\sigma^2} \right) \quad \text{for } i = 1, 2, \ldots, N \tag{47}
\]

Assuming that the relay decodes the user messages perfectly (which can be achieved by adjusting the user powers appropriately), the signals received at the destination in both time slots (see Eq.(6) and Eq.(7)) can be written as:

\[
\begin{bmatrix} y_{sd}[n] \\ y_{rd}[n] \end{bmatrix} = \begin{bmatrix} \sqrt{\gamma_1 h_{s1d}} & \sqrt{\gamma_2 h_{s2d}} & \cdots & \sqrt{\gamma_N h_{sN,d}} \\ \sqrt{g_{rd} \alpha h_{rd}} & \sqrt{g_{rd} \alpha h_{rd}} & \cdots & \sqrt{g_{rd} \alpha h_{rd}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + \begin{bmatrix} z_{sd}[n] \\ z_{rd}[n] \end{bmatrix} \tag{48}
\]

where \( y \) is the received signal vector at destination; \( H \) is the channel gain matrix; \( x \) is the scaled and signature multiplied user message vector, and \( z \) is the AWGN noise vector. Therefore, this channel can be modeled as a \( 2\times N \) Virtual-MIMO system, since we consider the dependency of the relayed signal on the sensor messages in Eq. (10) under the assumption of perfect relay decoding.

The average joint-sensor rate-bound at the FC is obtained as:

\[
R_1 + R_2 + \cdots + R_N \leq \frac{1}{2} \log \left(\det \left( I + \frac{H H^\dagger}{2\sigma^2} \right) \right) \tag{49}
\]

where \( \dagger \) denotes the conjugate-transpose operation, \( S = E[xx^T] \) is the input covariance matrix, which is a diagonal matrix with \( P_1, P_2, \ldots, P_N \) on its diagonal.
Using Eq. (49), the average rate of individual sensor can be derived as:

\[ R_i \leq \frac{1}{2} \log \left( 1 + \frac{\gamma_i P_i + g_r a P_i}{2 \sigma^2} \right) \]  

(50)

By combining Eqs. (46)-(47) with Eqs. (49)-(50), the average single-sensor and joint-user rate bounds become

\[ R_i \leq \min \left\{ \frac{1}{2} \log \left( 1 + \frac{\gamma_i P_i + g_r a P_i}{2 \sigma^2} \right), \frac{1}{2} \log \left( 1 + \frac{g_i P_i}{2 \sigma^2} \right) \right\} \times R_1 + R_2 + \cdots + R_N \]

\[ \leq \min \left\{ \frac{1}{2} \log \left( \det \left( I + \frac{E \{ \text{HSH}^H \}}{2 \sigma^2} \right) \right), \right\} \times \frac{1}{2} \log \left( 1 + \frac{\sum_{i=1}^{N} g_i P_i}{2 \sigma^2} \right) \]  

(52)

In order to realize the average-rate fairness, we equalize the maximum average rate bounds of sensors in Eqs. (51) as:

\[ \min \left\{ \frac{1}{2} \log \left( 1 + \frac{\gamma_i P_i + g_r a P_i}{2 \sigma^2} \right), \frac{1}{2} \log \left( 1 + \frac{g_i P_i}{2 \sigma^2} \right) \right\} = \min \left\{ \frac{1}{2} \log \left( 1 + \frac{\gamma_{i+1} P_{i+1} + g_r a P_{i+1}}{2 \sigma^2} \right), \frac{1}{2} \log \left( 1 + \frac{g_{i+1} P_{i+1}}{2 \sigma^2} \right) \right\} \]

\[ \times \frac{1}{2} \log \left( 1 + \frac{\sum_{i=1}^{N} g_i P_i}{2 \sigma^2} \right) \]  

for \( i = 1, 2, \ldots, N - 1 \)  

(53)

which can be further simplified as:

\[ \min \{ \gamma_i + g_r a, g_i \} P_{i+1} = \min \{ \gamma_{i+1} + g_r a, g_{i+1} \} P_{i+1} \]

for \( i = 1, 2, \ldots, N - 1 \)  

(54)

By considering Eq.(54), the total transmit power of the sensors in Eq.(44) and the solution for the relay control parameter in Eq.(45), the optimal power allocation with the average-rate fairness criterion can be derived as:

\[ P_i = \frac{P_{SN}}{\sum_{j=2}^{N} \min \{ \gamma_i + g_r a, g_i \}} \]  

\[ \times \min \{ \gamma_{i+1} + g_r a, g_{i+1} \} \]  

\[ \text{for } i = 1, 2, \ldots, N \]  

(55)

As a result of our analysis, we determine the signature phases, the sensor powers, and the relay power control parameter in Eq. (32), Eq.(55) and Eq. (45), respectively, which shall be used in CFNC-DD. In the next section, we compare the performance of the proposed CFNC-DD with that of the CDD.

V. SIMULATION RESULTS

In this part, we investigate the performances of CFNC-DD over non-orthogonal signaling and CDD over orthogonal signaling by obtaining and comparing their probability of error plots and their receiver operating characteristics (ROC) curves. It is critical to note that the probability of error plots are obtained with the assumption of equally likely priors for various SNRs, while each simulated points on ROC curves is resultant due to the use of a different prior probability. Throughout our discussion, we assume that sensors are identical in terms of their false alarm and detection probabilities as \( P_{F1} = 0.05 \) and \( P_{D1} = 0.5 \). To make a fair comparison between the Classical-DD (CDD) and CFNC-DD, we keep the average transmit power of the network per time slot as the same, which is assumed to be 2 units (i.e., \( P_T = 2 \)).

For all numerical simulations presented in this part, we have considered the topologies in Fig. 1 and Fig. 2 for CDD and CFNC-DD, respectively, where all wireless fading coefficients are taken to be complex Gaussian random variables with zero mean and unit variance, which are only known to the receiving nodes. Also, the samples of the receiver electronic noise are modeled as Gaussian random variables with zero mean and variance of \( \sigma^2 \) per dimension. Additionally, we assume that distance from the sensor \( S_1 \)-to-FC link is one and the corresponding path loss coefficient is unity (i.e., \( \gamma_1 = 1 \) or 0 dB) so that other path loss coefficients (i.e., \( g_k, g_r \) and \( \gamma_k \)) are interpreted as power gains or losses relative to the sensor \( S_1 \)-to-FC link. In this work, the path-loss exponent is taken to be 2 and the path-loss coefficients of the relay channel should satisfy the following triangle inequalities.

\[ |g_k^{-0.5} - g_r^{-0.5}| < |g_k^{-0.5} < g_k^{-0.5} + g_r^{-0.5} \]  

(56)

In our simulations, the signal-to-noise ratio is defined in dB as:

\[ \text{SNR (in dB)} = 10 \log_{10} \left( \frac{P_T T_0}{2 \sigma^2} \right) \]  

(57)

To compare both schemes under the information rate of unity (i.e., \( R = 1 \) symbol/sec/channel use), we set the duration of each channel use, \( T_0 \), to \( 1/N \) by using Eq.(4) and to \( 1/2 \) by using Eq.(8) for CDD or CFNC-CDD, respectively. Since the power budget is fixed, we realize different SNR values by changing the variance of electronics noise in Eq. (57).

Since we are performing Monte Carlo simulations, we run simulations many times and average the results to obtain the average probability of error in Eq. (18).

In order to see the benefit of CFNC-DD over CDD, we first obtain probability of error as a function of SNR for a WSN with 2 sensors (i.e., \( N = 2 \)) and \( \gamma_1 = \gamma_2 = 0 \) dB, \( g_1 = g_2 = 10.45 \) dB and \( g_r = 3.10 \) dB, in which the sensor nodes are equally separated from FC and the distance from the relay node to each sensor node is the same. One can see from Fig. 4 that CFNC-DD outperforms CDD for all SNR values considered. Specifically, the CFNC-DD provides an improvement up-to 20.74% in average error probability over CDD. In addition to average error probability plots, we also obtained ROC curves as depicted in Fig. 5 for the same network and different SNR values of \(-5, 0 \) and 5 dB. The proposed CFNC-DD method can lead to 10.34%, 11.76%, and 37.77% detection performance improvement over CDD for 5, 0 and \(-5 \) dB SNR, respectively.

From the observations made above, we can say that by employing the proposed method, the probability of error gets better especially in the low-SNR regime compared to
the use of CDD. In the CDD, each sensor signal is sent over an orthogonal channel and disturbed by one noise sample at the FC, which results in the availability of $N$ noisy measurements at the FC. Contrary to that, CFNC-DD allows interference of sensor signals at both the relay and FC but each of these interference signals for each time slot of CFNC-DD experiences a distortion due to one noise sample at the FC, and therefore, there are only two noisy measurements, which carry all sensor data, at the FC. Thus, the noise has a worse impact on the performance of CDD. Moreover, CDD provides only spatial diversity using the sensors that give the same decision, whereas CFNC-DD also achieves time diversity in addition to the spatial diversity, since the relay node in CFNC-DD sends its decision to the FC over a different time slot. This is very beneficial to reduce the negative effect of the interference and the noise on the detection performance of the network.

Secondly, we consider a scenario, where sensors are located asymmetric with respect to the FC, and the relay node is closer to $S_1$. Accordingly, parameters $g_1$, $g_2$ and $g_t$ are selected as 10.45 dB, 3.10 dB and 3.10 dB, respectively. The value of $\gamma_2$ is varied and decided to satisfy the triangle inequality in Eq. (56). Then, we obtain the probability of error results as a function of SNR for various values of $\gamma_2$, which are shown in Fig. 6. Again, CFNC-DD outperforms CDD using the same $\gamma_2$ value in all SNR regimes. For example, CFNC-DD has an error performance improvement of 8.43%, 20.92% and 22.98%, respectively for $\gamma_2$ value of 20 dB, 7.96 dB and −2.28 dB under the SNR of −5 dB. After that, the ROC curves are presented in Fig. 7 for different $\gamma_2$ values and SNR of −5 dB. Consistently, CFNC-DD has a better detection performance than CDD for each $\gamma_2$ value and a given false alarm probability. In particular, CFNC-DD results in detection performance improvements up to 32.14%, 42.10% and 105.50% for $\gamma_2$ value of 20 dB, 7.96 dB and −2.28 dB, respectively. Therefore, our proposed method performs better even if sensor nodes are located far from the FC. As pointed out earlier, this is because the CFNC-DD results in both spatial diversity and the time diversity.

Next, we increase the number of sensors to $N = 4$ and select $g_1 = g_2 = g_3 = g_4 = 13.98$ dB, in which the relay node is equally separated from the sensors. Also, the parameters $\gamma_1$, $\gamma_2$, $\gamma_3$, $\gamma_4$ and $g_t$ should be chosen to satisfy the triangle inequalities in Eq. (56), for which $\gamma_1 = 0$ dB, $\gamma_2 = 0.91$ dB, $\gamma_3 = 0.91$ dB, $g_t = 0.91$ dB are used and the value of $\gamma_4$ has been let to change during the numerical experiments without conflicting the triangle inequalities. Then, we obtain the probability of error versus SNR curves for various values of $\gamma_4$ as shown in Fig. 8. One can see from this figure that CFNC-DD decreases the error probability of CDD up to 29.10%, 34.40% and 31.61% for $\gamma_4$ of 20 dB, 6.02 dB and 0.91 dB, respectively, for SNR of 0 dB. Finally, we obtain the ROC curves in Fig. 9 by changing the a priori probability...
different values of $\gamma_4$ for $N = 4$, $\gamma_1 = 0$ dB, $\gamma_2 = 0.91$ dB, $\gamma_3 = 0.91$ dB, and $g_1 = g_2 = g_3 = g_4 = 13.98$ dB, $g_r = 0.91$ dB.

Fig. 9. ROC curves of CFNC-DD and CDD under various $\gamma_4$ values for $N = 4$, $\gamma_1 = 0$ dB, $\gamma_2 = 0.91$ dB, $\gamma_3 = 0.91$ dB, and $g_1 = g_2 = g_3 = g_4 = 13.98$ dB, $g_r = 0.91$ dB.

of the event in the region of interest for $\text{SNR}$ of 0 dB and different values of $\gamma_4$. CFNC-DD has a detection performance increase up to 34.73%, 54.23% and 57.86% detection performance improvement over CDD for $\gamma_4$ of 20 dB, 6.02 dB and 0.91 dB, respectively.

These results suggest that the proposed signature selection and relay power allocation in CFNC-DD is a powerful method to be employed in high speed WSNs.

VI. CONCLUSIONS

In this work, we considered the complex field network coded (CFNC) relay assisted communications in order to improve the performance of parallel wireless sensor networks (WSN) under fading and noise. We derived an optimal LRT based fusion rule for the considered system. Then, we proposed an analytical method to jointly determine the sensor signatures and the relay power by utilizing an upper bound on symbol error probability of the network together with some information theoretical results. Finally, we have shown with numerical simulations that the proposed method significantly outperforms the classical distributed detection (CDD) in terms of detection performance. Therefore, the proposed signature selection method in the considered system is a promising technique to be used in high performance and high throughput next generation wireless sensor networks, which operate over non-orthogonal channels.

APPENDIX I

DERIVATION OF SER UPPER BOUND AT FC

In this section, we will derive pair-wise error probability (PEP) and upper bound for PEP, which we used in our optimizations. Pair-wise error probability (PEP) at the fusion center (FC), which is denoted as $D$, can be written as

$$P\left(c_i \rightarrow c_j \text{ at } D \mid c_i\right) = P\left(c_i \rightarrow c_j \text{ at } R \mid c_i\right) P\left(c_i \rightarrow c_j \text{ at } D \mid c_i \rightarrow c_j \text{ at } R, c_i\right) + P\left(c_i \rightarrow c_j \text{ at } R \mid c_i\right) \left(1 - P\left(c_i \rightarrow c_j \text{ at } D \mid c_i \rightarrow c_j \text{ at } R, c_i\right)\right)$$

(58)

where $P\left(c_i \rightarrow c_j \text{ at } R \mid c_i\right)$ and $P\left(c_i \rightarrow c_j \text{ at } R \mid c_i\right)$ are correctly decoding probability and PEP respectively at the relay when $c_i$ is sent. Also, $P\left(c_i \rightarrow c_j \text{ at } D \mid c_i \rightarrow c_j \text{ at } R, c_i\right)$ is the PEP at the FC given that $c_i$ is sent and the relay correctly decoded, $P\left(c_i \rightarrow c_j \text{ at } D \mid c_i \rightarrow c_j \text{ at } R, c_i\right)$ is the probability of correctly decoding $c_i$ at the FC given that $c_i$ is sent and the relay erroneously decode $c_i$. Assuming CSI is known at the relay and ML relaying is used, the PEP of the relay node $R$ is

$$P\left(c_i \rightarrow c_j \text{ at } R \mid c_i, h_{sr}\right) = P\left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} + z_r - \sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} \left| z_r \right|^2 \geq \left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} \right)^2 \right)$$

$$= P\left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} + z_r - \sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} \left| z_r \right|^2 \geq \left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k (x_k)_k} \right)^2 \right)$$

$$= P\left(-\left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k d_{ijk}} \right)^2 \geq \left(\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k d_{ijk}} \right)^2 \right)$$

(59)

where \(\left|z_r\right|^2\) stands for the complex conjugate of a complex number and the distribution of random variable \(z_r\) will be a Gaussian distribution with zero mean and variance. Therefore, PEP at the relay can be written as follows

$$P\left(c_i \rightarrow c_j \text{ at } R \mid c_i, h_{sr}\right) = \mathcal{Q}\left(-\frac{\sum_{k=1}^{N} \sqrt{g_k h_{skr} \theta_k d_{ijk}}}{2\sigma}\right)$$

(60)

In addition to that, assuming CSI is available and ML estimator is used at the FC; PEP at the FC given that the relay correctly
\[ P \left( c_i \rightarrow c_j \text{ at D} \mid c_i \rightarrow c_j \text{ at R}, c_i, h_{rd}, h_{rd} \right) \]
\[ = P \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d - \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) \right)^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) \right|^2 \]
\[ \leq P \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d - \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) \right)^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) \right|^2 \]
\[ = P \left( \left| z_d \right|^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} + z_d \right|^2 \right) \geq \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{sd} \theta_k(x_k) \right)^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) \right|^2 \]
\[ = P \left( \left| z_d \right|^2 + \left| \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d \right|^2 \right) + \frac{2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) + 2 \text{Re} \left( z_d \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} \right)} {1 + \sqrt{\gamma_k} h_{sd} \theta_k(x_k)} \right) \]
\[ = Q \left( \frac{2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) + 2 \text{Re} \left( z_d \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} \right)} {1 + \sqrt{\gamma_k} h_{sd} \theta_k(x_k)} \right) \]

(63)

decoded \( c_i \) will be
\[ P \left( c_i \rightarrow c_j \text{ at D} \mid c_i \rightarrow c_j \text{ at R}, c_i, h_{rd}, h_{rd} \right) \]
\[ = P \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d - \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) \right)^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) \right|^2 \]
\[ \geq \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{sd} \theta_k(x_k) \right)^2 + \left| \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) + z_d - \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k(x_k) \right|^2 \]
\[ = P \left( \left| z_d \right|^2 + \left| \sqrt{\gamma_k} h_{sd} \theta_k(x_k) + z_d \right|^2 \right) + \frac{2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) + 2 \text{Re} \left( z_d \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} \right)} {1 + \sqrt{\gamma_k} h_{sd} \theta_k(x_k)} \right) \]
\[ = Q \left( \frac{2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) + 2 \text{Re} \left( z_d \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} \right)} {1 + \sqrt{\gamma_k} h_{sd} \theta_k(x_k)} \right) \]

(62)

Also, by a similar analysis as in Eq.(61), we can obtain the probability of decoding \( c_i \) correctly at the FC when it is given that the relay decodes the \( c_i \) erroneously, as in (63), shown at the top of the page.

Therefore, PEP at the FC, which is given in Eq.(58), can be written as follows
\[ P \left( c_i \rightarrow c_j \text{ at D} \mid c_i \rightarrow c_i \text{ at R} \right) \]
\[ = E_h \left[ 1 - Q \left( \frac{\sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk}} {2\sigma} \right) \right] \times \frac{2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) + 2 \text{Re} \left( z_d \sqrt{\gamma_k} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk} \right)} {1 + \sqrt{\gamma_k} h_{sd} \theta_k(x_k)} \right) \]

(61)

Since \( z_d \) has a complex Gaussian distribution with zero mean and \( 2\sigma^2 \) variance which can be denoted as \( CN(0, 2\sigma^2) \), distribution of random variable \( 2 \text{Re} \left( \sum_{k=1}^{N} \sqrt{\gamma_k} h_{sd} \theta_k d_{ijk} \right) \)
Each fading coefficient \( h_{st}, h_{sd}, h_{rd} \) is assumed to be a zero-mean complex Gaussian random variable with unit variance, which is denoted as \( \mathcal{CN}(0,1) \). Hence, distributions of the random variables \( T_1 = \left| \sum_{k=1}^{N} \sqrt{g_k} h_{st} d \theta_k d_{ijk} \right|^2 \) and \( T_2 = \left| \sum_{k=1}^{N} \sqrt{g_k} h_{sd} d \theta_k d_{ijk} \right|^2 \) and \( T_3 = \left| \sum_{k=1}^{N} \sqrt{g_k} h_{rd} d \theta_k d_{ijk} \right|^2 \) are exponential with a mean of \( \lambda_1 = \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2 \) and \( \lambda_2 = \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk} \) and \( \lambda_3 = g_r \sum_{k=1}^{N} |\theta_k|^2 d_{ijk}^2 \), respectively.

Also, pairwise error probability in Eq. (64) can be upper bounded as

\[
P(c_i \rightarrow c_j \text{ at } D[c_i]) \leq E_h \left[ Q \left( \frac{\sqrt{\sum_{k=1}^{N} \sqrt{g_k} h_{sd} d \theta_k d_{ijk}}^2 + \sqrt{g_r} h_{rd} \sum_{k=1}^{N} \theta_k d_{ijk}^2}}{2\sigma} \right) \right]
\]

Using union bound we obtain SER upper bound at FC as follows

\[
\hat{P}_e^D \leq B_e^D(\theta, \alpha)
\]

\[
= \sum_{i=1}^{2N} \sum_{j=1}^{2N} P(c_i) \left( \frac{0.5}{1 + \frac{\sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2}{8\sigma^2}} + \frac{1}{1 + \frac{N g_r |\theta_k|^2 d_{ijk}^2}{8\sigma^2}} \right)
\]

\[
+ \sum_{i=1}^{2N} \sum_{j=1}^{2N} P(c_i) \left( \frac{1}{1 + \frac{\sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2}{8\sigma^2}} \right)
\]

\[(66)\]

**APPENDIX II**

**CONVEXITY OF THE AVERAGE SER BOUND AT THE FUSION CENTRE**

We show here the convexity of the average SER-bound \( B_e^D(\theta, \alpha) \) in Eq. (66). Clearly, \( \sum_{k=1}^{N} g_k |\theta_k|^2 d_{ijk}^2 \) and \( \sum_{k=1}^{N} g_r |\theta_k|^2 d_{ijk}^2 \) are convex functions of the user signatures. Also, \( g_r \sum_{k=1}^{N} |\theta_k|^2 d_{ijk} \) is convex with respect to the parameter vector \( p = (\theta, \alpha)^T \) since the multiplication of two convex scalar functions is also convex, when both functions are non-decreasing (non-increasing) and positive [32]. Additionally, \( \left( 1 + \frac{g_r \sum_{k=1}^{N} |\theta_k|^2 d_{ijk}}{8\sigma^2} \right)^{-1} \) is convex since \( f(x) = \frac{1}{1+x} \) is convex for \( x \geq 0 \) and the composition of a convex scalar function with another convex scalar function is convex. Therefore, the average SER-bound \( B_e^D(\theta, \alpha) \) is convex in both signatures \( \theta_k \) and parameter \( \alpha \), and thus, Eq. (31) is a convex program.

**REFERENCES**


Kayhan Eritmen received the B.S. degree in telecommunication engineering from Istanbul Technical University, Istanbul, Turkey, in 2006, and the M.S. and Ph.D. degrees from the Electronics Engineering Program, Sabancı University, Istanbul, Turkey, in 2008 and 2014, respectively. During his graduate study, he has been with the Communication Theories and Technologies Group and supervised by Dr. M. Keskinoz. His research interests are in the areas of distributed detection, data fusion, MIMO communication, and network coding in wireless networks.

Mehmet Keskinoz received the M.S. and Ph.D. degrees from the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA, in 1997 and 2001, respectively. In 2001, he joined the Electronics Engineering Program, Sabancı University, Istanbul, Turkey, where he is currently an Associate Professor. His research interests include signal processing for wired and wireless communications, UWB communications, Multiband OFDM UWB systems, Wireless Mesh Networks, magnetic and optical data storage systems, distributed detection and data fusion for wireless sensor networks, Turbo and LDPC coding, synchronization, and digital watermarking. He was a recipient of Turkish NSF Research grant on distributed detection in wireless sensor networks and Career Award on wireless mesh networks in 2005. He was a Co-Guest Editor of the IEEE Communications Magazine in 2009 Special Issue on Advances in Signal Processing for Wireless and Wired Communications. He is also a member of the IEEE Communication Society, the IEEE Signal Processing Society, and the Optical Society of America.