



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/97002/>

Version: Accepted Version

Article:

Nurellari, E, McLernon, DC and Ghogho, M (2016) Distributed Two-Step Quantized Fusion Rules via Consensus Algorithm for Distributed Detection in Wireless Sensor Networks. IEEE Transactions on Signal and Information Processing over Networks, 2 (3). ISSN: 2373-776X

<https://doi.org/10.1109/TSIPN.2016.2549743>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Distributed Two-Step Quantized Fusion Rules via Consensus Algorithm for Distributed Detection in Wireless Sensor Networks

Edmond Nurellari, Des McLernon, *Member, IEEE*, and Mounir Ghogho, *Senior Member, IEEE*

Abstract—We consider the problem of distributed soft decision fusion in a bandwidth-constrained spatially uncorrelated wireless sensor network (WSN). The WSN is tasked with the detection of an intruder transmitting an unknown signal over a fading channel. Existing distributed consensus-based fusion rules algorithms only ensure equal combining of local data and in the case of bandwidth-constrained WSNs, we show that their performance is poor and does not converge across the sensor nodes (SNs). Motivated by this fact, we propose a two-step distributed quantized fusion rule algorithm where in the first step the SNs collaborate with their neighbors through error-free, orthogonal channels (the SNs exchange quantized information matched to the channel capacity of each link). In the second step, local 1-bit decisions generated in the first step are shared among neighbors to yield a consensus. A binary hypothesis testing is performed at any arbitrary SN to optimally declare the global decision. Simulations show that our proposed quantized two-step distributed detection algorithm approaches the performance of the unquantized centralized (with a fusion center) detector and its power consumption is shown to be 50% less than the existing (unquantized) conventional algorithm.

Index Terms—Distributed detection, soft decision, quantized weighted average consensus, wireless sensor networks.

I. INTRODUCTION

WIRELESS sensor networks (WSNs) are spatially deployed over a field to monitor certain physical or environmental phenomena, to detect relevant quantities and to perform decision making functions. Because of their relatively low cost and robustness to sensor node (SN) failures they are receiving significant attention. However, there are a number of different strategies as to how the test statistics from each SN will be used in order to arrive at a final decision. We will first give a brief review before introducing our proposed approach.

First consider the *centralized solution* where noisy observations collected from spatially distributed local SNs are sent (inter-sensor collaboration is not considered) to a global fusion center (FC) for a final decision [1]-[12]. Then, there are some recent publications [13]-[14] (in the context of estimation) that considered the effect of *inter-sensor collaboration* on the estimation performance. Here, the local SNs collaborate

through error-free, low cost transmission links (defined by the symmetric adjacency matrix). After the collaboration stage, the SNs (which in general can be a subset of all SNs) report to a FC where the final decision is made. Reference [13] proposes an *efficient* collaboration strategy in a distributed fashion (as opposed to [15] where this optimal collaboration strategy is computed at a FC) by means of using only local SNs observations. While the authors in [13] claim to reduce the FC control overhead, [14] derives the optimum power allocation scheme constraint on the maximum total network power budget in order to improve the quality of the estimation back at the FC. These two *hybrid approaches* (a SN collaboration stage followed by reporting to a FC), like the first approach (no collaboration stage and every SN reports directly to a FC), rely on the integrity of the FC. Furthermore, collecting information at the FC lacks scalability, and may require large amounts of energy and communication resources [16].

The second approach is a *fully distributed strategy* (i.e., without a FC) [17]-[26], where the SNs exchange local information iteratively among their neighbors and are capable of reaching a global optimum decision. The authors of [17] and [18] adopt the diffusion-based protocol and propose a new diffusion LMS algorithm while [19] develop a fully distributed consensus-based LMS algorithm that outperforms the existing (relying on information diffusion) alternatives. The authors of [20] design a bio-inspired algorithm that can achieve globally optimal distributed decisions while in [21] they investigate the consensus problem in the presence of propagation delays. Reference [21]-[27] employ the iterative distributed consensus algorithm [28] for distributed inference. But these approaches consider ideal exchange of information among all the SNs, and as the SNs are battery operated (i.e., with limited energy available on-board) this assumption is unrealistic. Furthermore, practical WSN scenarios suffer from channel impairments such as fading and attenuation. Recently, to address the problem of consensus algorithms with quantized communications, a number of different approaches have been proposed. The authors in [29] propose a probabilistic quantization scheme that is shown to reach a consensus (almost surely) to a random variable whose expected value is equal to the desired average. Unfortunately, it is shown that this scheme performs poorly at low bit rate. Another approach to mitigate the quantization error in the consensus algorithm is to use an iteration dependent step size as in [6] and [30]. Adapting the weight link sequence in order to guarantee convergence is shown to decrease the convergence rate and so introduces

Part of the results in this paper have been presented at IET International Conference on Intelligent Signal Processing (ISP), London, UK, December 1-2, 2015 [32].

E. Nurellari and D. McLernon are with the School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT, Leeds, U.K. (e-mail: elen@leeds.ac.uk; d.c.mclernon@leeds.ac.uk).

M. Ghogho is with the School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT, Leeds, U.K., and also with the International University of Rabat, Rabat, Morocco (e-mail: m.ghogho@leeds.ac.uk).

a delay to the detection algorithm. Even employing such decaying link weights satisfying a persistence condition (i.e., their sum over time diverges, while their square sum is finite) cannot guarantee the convergence to the target average [30]. Recently [31] introduced a progressive quantization scheme that is shown to achieve the true average solution even at a low communication rate. However, this scheme has a high computational complexity and relies on doubly stochastic weight matrix. Now, most of the existing works on quantized consensus assume that the communication topology is symmetric (which is not the case in our manuscript). Furthermore, all the above-mentioned algorithms either maintain the average value in the network but cannot reach a consensus effectively, or converge to a random variable that is not always the target average value.

So, the purpose of this paper is to develop a fully distributed detection framework for realistic WSN scenarios. The communication links among SNs are modeled as channels with path loss, flat fading and additive white Gaussian noise (AWGN). The assumption of flat fading is reasonable because most of the WSNs operates at both short distances and low bit rate due to resource limitations. We will show that this new distributed framework can approach the performance of a centralized optimum detector (i.e., with a FC).

A. Contributions & Organization

So, the main contributions are as follows:

(i) First, the (unquantized) consensus algorithm [28] is modified in such a way that the SNRs of the local SNs are taken into account in order to further improve the global detection performance. We re-state the necessary conditions for convergence to the (unquantized) optimum linear combining solution [10]. Based on this, we provide a distributed consensus-based detection framework with (weight combining) quantized test statistic exchange (SNs implement a low complexity uniform quantizer and the number of quantization bits is constrained to match the channel capacity of each link). Using the probability of detection and the probability of false alarm as metrics, we show that this approach: (a) does not converge to a global decision across the network, and (b) does not approach the optimum quantized centralized detector (i.e., with a FC) performance [10].

(ii) Second, motivated by the above, we propose a novel two-step quantized distributed weighted fusion algorithm that now: (a) converges to a global decision across the network, (b) approaches the optimum centralized detector performance, and (c) achieves the global decision in a finite number of iterations. The main idea of this proposed two-step distributed (quantized) fusion algorithm is to arrive at an optimum global decision at every SN by taking advantage of the spatially distributed information across the WSN while combating flat fading.

Now, the summary of the paper is as follows. In Section II we formulate the detection problem and recall some basic definitions from graph theory that we will be using. Section III describes two different approaches (i.e., the centralized approach (with a FC) and the fully distributed approach

(without a FC)). In Section IV we describe a consensus-based distributed detection framework and analyze the detection performance by proving that the quantized distributed detector performance does not converge across the SNs. Motivated by this, we then propose a two-step quantized weighted fusion algorithm with performance comparable to the centralized (unquantized) optimum detector. Finally, Section V presents simulation results that confirm our analytical findings and in Section VI we give conclusions.

II. PROBLEM FORMULATION

Consider the problem of detecting the presence of any intruders by a use of WSN consisting of M spatially uncorrelated distributed SNs. The intruder leaves a signature signal $s(n)$ that is unknown to the WSN but it is assumed to be deterministic. N samples of the observed signal are gathered and energy estimation is then performed by each SN. The measurement (of $s(n)$) at each SN $s_i(n)$ is further corrupted by AWGN $w_i(n) \sim \mathcal{N}(0, \sigma_i^2)$. In this paper, we consider two different schemes: a) the centralized approach (see Fig. 1), where each SN sends its test statistic (quantized to L_i bits) to the FC (see section III-A) where the FC combines them and makes the final decision; and b) is the decentralized approach (see Fig. 2 and section III-B), where SN i shares iteratively its current test statistic (quantized to q_i bits) across the set (Δ_i) of its neighbors (see Δ_i definition in section II-B). Next, we explain in more detail the local sensing model and some graph theory definitions.

A. Sensing

The measured signal at SN i is either:

$$\mathcal{H}_0 : y_i(n) = w_i(n) \quad (1)$$

$$\mathcal{H}_1 : y_i(n) = s_i(n) + w_i(n) \quad (2)$$

and energy estimation is performed at the i^{th} SN to give

$$T_i = \sum_{n=1}^N (y_i(n))^2, \quad i = 1, 2, \dots, M \quad (3)$$

which for large N has an approximately Gaussian distribution [33]. Furthermore, the noise samples are assumed to be identically and independently distributed (i.i.d.) across time and space. It is not difficult to show that

$$\begin{aligned} \mathbb{E}\{T_i|\mathcal{H}_0\} &= N\sigma_i^2, & \text{Var}\{T_i|\mathcal{H}_0\} &= 2N\sigma_i^4 \\ \mathbb{E}\{T_i|\mathcal{H}_1\} &= N\sigma_i^2(1 + \xi_i), & \text{Var}\{T_i|\mathcal{H}_1\} &= 2N\sigma_i^4(1 + 2\xi_i) \end{aligned} \quad (4)$$

where $\xi_i = \sum_{n=1}^N s_i^2(n) / N\sigma_i^2$.

B. Sensor Nodes Interaction Model

The interaction among SNs is according to the communication topology which is given by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \dots, M\}$ represents the set of M SNs and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges $\{i, j\}$. The

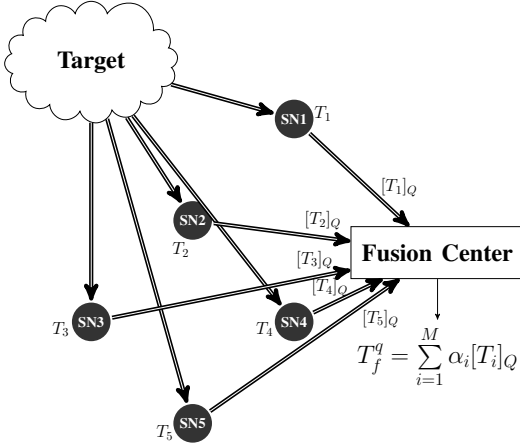


Fig. 1. Schematic communication architecture between peripheral SNs and the fusion center (FC). Each SN generates a test statistic (T_i) by observing the target and can communicate (using $[T_i]_Q$) with the FC only over an energy-constrained/bandwidth-constrained link.

graph properties can be represented by an *adjacency matrix* $\mathbf{E} \in \mathbb{R}^{M \times M}$ whose entries are defined as

$$e_{ij} = e_{ji} = \begin{cases} 1, & \text{if } j \in \Delta_i \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

We denote the i^{th} SN neighbor set as Δ_i and $|\Delta_i|$ is the number of neighbors. The definition of the graph *Laplacian* matrix ($\mathbf{L} \in \mathbb{R}^{M \times M}$) is $\mathbf{L} = \mathbf{D} - \mathbf{E}$ with $\mathbf{D} = \text{diag}(|\Delta_1|, \dots, |\Delta_M|)$. Next, we discuss the centralized and distributed detection approaches and provide an optimum distributed (i.e., without a FC) weight combining fusion rule framework.

III. CENTRALIZED VS. DISTRIBUTED

The first scheme¹ (see Fig. 1) is a WSN consisting of M spatially distributed SNs that report to a FC. Upon receiving the contributions from each individual local SN, the FC linearly combines them and then declares a global decision. We refer to this approach as a *centralized scheme*. In the other approach (see Fig. 2) the SNs collaborate among each other iteratively to come up to a global decision in a fully distributed fashion (i.e., without a FC). In this case, each SN i is able to perform a (global) decision. We refer to this approach as a *decentralized scheme*. Note that the results derived in this section will serve as the basis for developing the new optimum two-step quantized (weighted) fusion rule algorithm in section IV.

A. Centralized Approach

In order to better understand the fully distributed algorithm that we propose later in this paper we first describe two different centralized approaches: *quantized* and *unquantized*.

¹Now, $[T_i]_Q$ is the i^{th} SN quantized test statistic (see (7)), $\{\alpha_i\}_{i=1}^M$ are the optimum weights (see (12)) and the superscript “q” refers to “quantized”.

1) *Quantized Centralized Approach* [10]–[11]: Here, quantized linear² soft decision combining at the FC is proposed, where each individual SN has to quantize its observed test statistic (T_i) (prior to transmission to a FC) to L_i bits. So, to satisfy the capacity constraint on each SN to FC channel, we require:

$$L_i \leq \frac{1}{2} \log_2 \left(1 + \frac{p_i h_i^2}{\zeta_0} \right) \text{ bits/sample} \quad (6)$$

where p_i denotes the transmit power of SN i , h_i is the flat fading coefficient between SN i and the FC, and ζ_0 is the variance of the AWGN at the FC. The quantized test statistic ($[T_i]_Q$) at the i^{th} SN can be modeled as

$$[T_i]_Q = T_i + v_i \quad (7)$$

where v_i is the quantization noise independent of $w_i(n)$ in (1) and (2). Assuming uniform quantization with $T_i \in [0, 2U]$, then

$$\sigma_{v_i}^2 = \frac{U^2}{3 \times 2^{2L_i}}. \quad (8)$$

Linearly combining $\{[T_i]_Q\}_{i=1}^M$ at the FC gives¹

$$T_f^q = \sum_{i=1}^M \alpha_i [T_i]_Q. \quad (9)$$

For large M , T_f^q will be approximately Gaussian and we can show (10) and (11). Now, the optimum weights $\{\alpha_i\}_{i=1}^M$ are given as [10]:

$$\alpha = \left[\frac{N\sigma_1^2\xi_1}{2N\sigma_1^4(1+2\xi_1)+\sigma_{v_1}^2}, \dots, \frac{N\sigma_M^2\xi_M}{2N\sigma_M^4(1+2\xi_M)+\sigma_{v_M}^2} \right]. \quad (12)$$

So (12) establishes a relationship between the optimum weighting vector (α) and the SN transmit power (p_i) through the $\sigma_{v_i}^2$ quantity (see definition (6) and (8)). The FC then makes the following decisions:

$$\left. \begin{array}{l} \text{if } T_f^q < \Lambda_f, \text{ decide } \mathcal{H}_0 \\ \text{if } T_f^q \geq \Lambda_f, \text{ decide } \mathcal{H}_1 \end{array} \right\} \quad (13)$$

where Λ_f is the FC detection threshold. The probability of detection (P_d) for a fixed probability of false alarm (P_{fa}) is given as [34]:

$$P_d = Q \left(\frac{Q^{-1}(P_{fa}) \sqrt{\text{Var}\{T_f^q|\mathcal{H}_0\}} - \mathbb{E}\{T_f^q|\mathcal{H}_1\} + \mathbb{E}\{T_f^q|\mathcal{H}_0\}}{\sqrt{\text{Var}\{T_f^q|\mathcal{H}_1\}}} \right) \quad (14)$$

with appropriate quantities given in (10)–(11) (see [10]) and where $Q(\cdot)$ is the Q -function.

²The main motivation behind the linear combining rule consideration is that the probability of detection and the probability of false alarm metrics are obtained in a closed-form. This gives insight into the design of the system’s parameters, whereas for the LRT-based detector, analytically analyzing the detection performance is not tractable.

$$\mathbb{E}\{T_f^q|\mathcal{H}_0\} = \sum_{i=1}^M \alpha_i N \sigma_i^2, \quad \mathbb{E}\{T_f^q|\mathcal{H}_1\} = \sum_{i=1}^M \alpha_i (N \sigma_i^2 (1 + \xi_i)) \quad (10)$$

$$\text{Var}\{T_f^q|\mathcal{H}_0\} = \sum_{i=1}^M \alpha_i^2 (2N \sigma_i^4 + \sigma_{v_i}^2), \quad \text{Var}\{T_f^q|\mathcal{H}_1\} = \sum_{i=1}^M \alpha_i^2 (2N \sigma_i^4 (1 + 2\xi_i) + \sigma_{v_i}^2). \quad (11)$$

2) *Unquantized Centralized Approach*: Given the local test statistic T_i (see (3)) at the i^{th} SN, the optimum (unquantized) linear fusion rule³ has the structure [10]:

$$T_f^{uq} = \sum_{i=1}^M \alpha_i T_i \quad (15)$$

where the superscript “uq” refers to “unquantized” and $\{\alpha_i\}_{i=1}^M$ are the optimum weights given in (12) but now with $\{\sigma_{v_i}^2\}_{i=1}^M = 0$. The probability of detection (P_d) for a fixed probability of false alarm (P_{fa}) is given again as in (14) (replacing T_f^q by T_f^{uq}) by substituting the appropriate quantities given in (10)-(11) with $\{\sigma_{v_i}^2\}_{i=1}^M = 0$. This gives an upper bound on the receiver operating characteristic performance and we will refer later to this in the simulation results.

Now the limitation of the centralized approach is both the requirement of the FC to process a large amount of data (i.e., possible bottleneck) and the possible failure of the FC. Hence, distributed solutions are very attractive as the computational load splits across the network. The final decision can be taken at any arbitrary SN. As a result, the system is more robust against FC failure than in a centralized system.

B. Distributed Approach

Now we are after the fully distributed approach (see Fig. 2) and we propose a distributed quantized linear fusion rule. Even though there are different distributed algorithms in the literature (i.e., average consensus, diffusion, gossip-type algorithms, etc), we will use the consensus algorithm [28] as a basic tool to develop the distributed quantized linear fusion rule.

1) *Unquantized Distributed Equal Combining*: Now consider the conventional consensus-based [28] distributed equal combining scheme that fuses the contributions received among SNs (i.e., it does not accommodate properly the *more informative* and the *less informative* neighbors). At iteration $k+1$, each SN i updates its test statistic ($T_i^{eq}[k+1]$) as follows [28]:

$$T_i^{eq}[k+1] = T_i^{eq}[k] - \epsilon \sum_{j=1}^M e_{ij} (T_i^{eq}[k] - T_j^{eq}[k]), \quad k \geq 0, \quad (16)$$

for $i = 1, 2, \dots, M$

where the superscript “eq” refers to “equal combining”, $0 < \epsilon < 1/\Delta_{\max}$ with $\Delta_{\max} = \max(|\Delta_1|, \dots, |\Delta_M|)$, e_{ij}

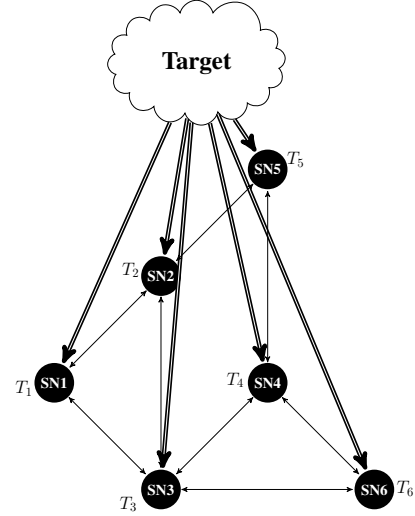


Fig. 2. Schematic for a distributed communication architecture among peripheral SNs. Each SN generates a test statistic (T_i) by observing the target (thick lines). The SNs have partial connectivity (thin lines) among themselves (i.e., not a complete graph), but only over an energy-constrained/bandwidth-constrained network.

is defined in (5) and $T_i^{eq}[0] = T_i$ in (3). The time evolution of (16) can be written as

$$\mathbf{T}^{eq}[k] = \mathbf{W}^k \mathbf{T}^{eq}[0], \quad k \geq 1 \quad (17)$$

where $\mathbf{W} = \mathbf{I} - \epsilon \mathbf{L}$ and $\mathbf{T}^{eq}[k] = [T_1^{eq}[k], T_2^{eq}[k], \dots, T_M^{eq}[k]]^T$. The decision can be taken locally at the i^{th} SN at the k^{th} iteration as follows:

$$\left. \begin{aligned} &\text{if } T_i^{eq}[k] < \Lambda_i[k], \text{ decide } \mathcal{H}_0 \\ &\text{if } T_i^{eq}[k] \geq \Lambda_i[k], \text{ decide } \mathcal{H}_1 \end{aligned} \right\} \quad (18)$$

where $\Lambda_i[k]$ is the threshold for the i^{th} SN at the k^{th} iteration. We can write:

$$\mathbb{E}\{T_i^{eq}[k]|\mathcal{H}_p\}_{p=\{0,1\}} = (\mathbf{W}^k \mathbb{E}\{\mathbf{T}^{eq}[0]|\mathcal{H}_p\})_i \quad (19)$$

$$\begin{aligned} \text{Var}\{T_i^{eq}[k]|\mathcal{H}_p\}_{p=\{0,1\}} &= (\text{Cov}\{\mathbf{T}^{eq}[k]|\mathcal{H}_p\})_{ii} \\ &= (\mathbf{W}^k \text{Cov}\{\mathbf{T}^{eq}[0]|\mathcal{H}_p\} \mathbf{W}^k)_{ii} \end{aligned} \quad (20)$$

where⁴ $(\mathbf{a})_i$ denotes the i^{th} element of vector \mathbf{a} and $(\mathbf{A})_{ij}$ denotes the (i, j) element of matrix \mathbf{A} . For a fixed probability of false alarm (i.e., $P_{fa}^i[k] = P_{fa}$, $\forall i$ and $\forall k$), the detection

³This is a special case assuming that the FC receives all the local test statistics $\{T_i\}_{i=1}^M$ without errors and in practice it is a strong assumption.

⁴For a random vector \mathbf{x} , $\mathbb{E}\{\mathbf{x}\}$ denotes expectation and $\text{Cov}\{\mathbf{x}\} = \mathbb{E}[(\mathbf{x} - \mathbb{E}\{\mathbf{x}\})(\mathbf{x} - \mathbb{E}\{\mathbf{x}\})^T]$ is the covariance matrix.

probability for the i^{th} SN at the k^{th} iteration can be written as

$$P_d^i[k] = Q\left(\frac{Q^{-1}(P_{fa})\sqrt{\text{Var}\{T_i^{eq}[k]|\mathcal{H}_0\}} + \Psi}{\sqrt{\text{Var}\{T_i^{eq}[k]|\mathcal{H}_1\}}}\right) \quad (21)$$

where $\Psi = \mathbb{E}\{T_i^{eq}[k]|\mathcal{H}_0\} - \mathbb{E}\{T_i^{eq}[k]|\mathcal{H}_1\}$. Now, (21) establishes a relationship between the probability of detection ($P_d^i[k]$) and the iteration number k at the i^{th} SN. It can be shown [28] that as k gets larger, the performance of the distributed detector (18) for a connected network⁵ approaches that of the (unquantized) equal combining centralized detector (13) (i.e., $\lim_{k \rightarrow \infty} P_d^i[k] = P_d$, $\forall i$ with $\alpha_i = 1$ and $\sigma_{v_i}^2 = 0$, $\forall i$ in (15)). However, this distributed fusion rule realizable via (16) (and also its centralized counterpart) is not optimum.

What we require now is a distributed approach that will converge to the equivalent of the optimum weighted linear combining FC solution in (15).

2) *Unquantized Distributed Weight Combining*: In our previous work [10] we have optimized the weights (α_i) such that the probability of detection is maximized. As can be seen from (12), the optimum weights are a function of local sensing quality (σ_i^2), received signal strength (ξ_i) and the SN transmit power (p_i) through the quantization noise ($\sigma_{v_i}^2$) (see (6) and (8)). So now using these optimal weights we derive a weighted exchange of information version of (16). Because the i^{th} SN does not know its neighbors' weights $\{\alpha_j\}_{j \in \Delta_i}$, we propose to weight the contributions received from the $|\Delta_i|$ neighbors by $f(\alpha_i)$, where f is the function that we elaborate later on. More specifically, the i^{th} SN updates its test statistic as follows:

$$T_i^w[k+1] = T_i^w[k] - \epsilon f(\alpha_i) \sum_{j=1}^M e_{ij} (T_i^w[k] - T_j^w[k]), \quad k \geq 0, \\ \text{for } i = 1, 2, \dots, M \quad (22)$$

where the superscript “ w ” refers to “*weighted*”, α_i are the centralized weighting coefficients in (12) with $\sigma_{v_i}^2 = 0$, $f(\alpha_i) \geq 0$, ϵ is defined for (16) and $T_i^w[0] = T_i$ in (3). The time evolution of (22) can be written as

$$\mathbf{T}^w[k] = \mathbf{W}^k \mathbf{T}^w[0], \quad k \geq 1 \quad (23)$$

with \mathbf{W} defined as

$$\mathbf{W} = \mathbf{I} - \epsilon \mathbf{\Gamma} \mathbf{L} \quad (24)$$

and $\mathbf{\Gamma} = \text{diag}(f(\alpha_1), f(\alpha_2), \dots, f(\alpha_M))$. We will now show that there exist a function f such that (23) (unquantized, distributed) converges to (15) (unquantized, centralized). First we prove two propositions.

Proposition 1 : Let \mathbf{W} be a matrix defined in (24) with $0 < \epsilon < 1/\Delta_{max}$. Then \mathbf{W} is a non-negative matrix (i.e., $\mathbf{W} \geq 0$) if $\mathbf{\Gamma} \leq 1$.

Proof : Note that from the definition of the Laplacian matrix (\mathbf{L}) defined in section II-B, (24) can be expressed as $\mathbf{W} = \mathbf{I} - \epsilon \mathbf{\Gamma} \mathbf{D} + \epsilon \mathbf{\Gamma} \mathbf{E}$. Now, by definition $\mathbf{\Gamma} \geq 0$, and so $\epsilon \mathbf{\Gamma} \mathbf{E}$ is also a non-negative matrix. The entries of the diagonal matrix

⁵A connected network is any network where there is a path (i.e., over one or more links) between every pair of SNs in the network.

($\mathbf{I} - \epsilon \mathbf{\Gamma} \mathbf{D}$) have to be non-negative, $\forall i$ (i.e., $1 - \epsilon f(\alpha_i) \Delta_i \geq 0, \forall i$). This can be achieved with $1 - \frac{f(\alpha_i) \Delta_i}{\Delta_{max}} \geq 0$ and since $f(\alpha_i) \leq 1, \forall i$ (i.e., $\mathbf{\Gamma} \leq 1$) $\implies 1 - \frac{f(\alpha_i) \Delta_i}{\Delta_{max}} \geq 0$. Then $\mathbf{\Gamma} \leq 1 \implies \mathbf{W} \geq 0$.

Proposition 2 : Let \mathbf{W} be a matrix defined in (24) with $0 < \epsilon < 1/\Delta_{max}$, $\mathbf{\Gamma} \leq 1$ and assuming a connected graph \mathcal{G} , then

$$\lim_{k \rightarrow \infty} \mathbf{W}^k \mathbf{T}^w[0] = \sum_{i=1}^M \frac{1}{f(\alpha_i)} \begin{bmatrix} \sum_{i=1}^M \frac{1}{f(\alpha_i)} T_i^w[0] \\ \sum_{i=1}^M \frac{1}{f(\alpha_i)} T_i^w[0] \\ \vdots \\ \sum_{i=1}^M \frac{1}{f(\alpha_i)} T_i^w[0] \end{bmatrix}_{M \times 1} \quad (25)$$

Proof : The proof is given in Appendix A.

Now, the convergence of (25) (with $T_i^w[0] = T_i$) to (15) up to a positive scaling can be only achieved if $f(\alpha_i) = \frac{1}{\alpha_i}$. It is worth mentioning that the condition ($f(\alpha_i) \leq 1, \forall i$) does not affect the optimality of the fusion rule defined in (15) for the structure considered in (14) and the condition can be satisfied by scaling the centralized weighting vector (α) by a positive constant c . Clearly, the distributed system (22) achieves the performance of the unquantized centralized approach in section III-A2.

We have now stated the necessary and sufficient conditions for the time evolution (22) to converge to the weighted centralized optimum linear fusion rule (15). The exchange of information between SNs is assumed error free and the bandwidth between two connected SNs is considered unlimited. In the next section, we relax these assumptions and provide a quantized distributed weighted linear fusion rule framework that operates over limited bandwidth fading channels.

IV. DISTRIBUTED DETECTION VIA TWO-STEP QUANTIZED DISTRIBUTED WEIGHTED FUSION RULE OVER FADING COMMUNICATION LINKS

Now, in section IV-A we develop a consensus-based quantized distributed weighted linear fusion framework. Next, in Section IV-B, using the probability of detection and the probability of false alarm as metrics, we analyze performance and give a proof that the *quantized distributed weighted linear fusion rule algorithm* does not converge across the SNs. Finally, in Section IV-C, based on the framework provided in Section IV-A, we propose a new two-step quantized distributed weighted fusion algorithm.

A. Quantized distributed weighted fusion rule

Here we propose a scheme, where each SN encodes the data (using a simple uniform quantizer with q_i bits) prior to information exchange with its neighbors. We also propose to establish a link between any two SNs i and j based on the (known) link SNR at node j , i.e.

$$\left. \begin{array}{l} \text{if } SNR_{ij} < \Upsilon, \quad e_{ij} = e_{ji} = 0 \\ \text{if } SNR_{ij} \geq \Upsilon, \quad e_{ij} = e_{ji} = 1. \end{array} \right\} \quad (26)$$

Now e_{ij} is defined in (5), Υ is a (link) SNR threshold parameter (see later) and SNR_{ij} is the received signal-to-noise ratio (at SN j) defined as:

$$SNR_{ij} = \frac{p_{ij}^t h_{ij}^2}{\zeta_0 d_{ij}^\gamma}. \quad (27)$$

Here p_{ij}^t denotes the i^{th} to j^{th} SN transmit power, h_{ij} is the flat-fading coefficient⁶ between the i^{th} and j^{th} SN, ζ_0 is the variance of the AWGN at each receiving SN (assumed to be the same for simplicity), γ is the path loss coefficient and d_{ij} is the physical distance between SN i and j (assumed to be known).

The thresholding operation (26) defines the communication topology. There are different approaches taken in the literature in order to define the topology of the network. In [36] a simplified relaxed (centralized) solution was presented, where the energy minimization problem was formulated as a convex-concave fractional programming. Another approach was followed in [37], where a distributed algorithm to decide which subset of communication links provides the optimum power consumption and the best network lifetime (i.e., minimizing simultaneously both the total power consumption and the maximum power consumption per SN) was developed. While both ([36] and [37]) improve the total power consumption and/or the whole network lifetime, they also assume that the exchange of information among SNs is ideal. But here we propose to quantize with q_i bits at SN i before transmitting to SN j and to satisfy the capacity constraint between SNs i and j we require:

$$q_i \leq \frac{1}{2} \log_2(1 + \Upsilon) \quad \text{bits/sample} \quad (28)$$

where we let $q_i = q, \forall i$. Now, Υ establishes a relationship between the number of bits that each SN has to transmit to its neighbors and also the topology of the network that defines the connections between the SNs (see (26)-(28)). A large Υ means fewer communication links (see (26)) resulting in slower information diffusion across the network. However, this will be counterbalanced by an increase in the number of bits that each SN can transmit to its neighbors (see (28)). As a consequence, the quantization noise variance (30) becomes negligible. Alternatively, a small Υ establishes a more connected graph and dictates a faster information diffusion across the network. However, this allows less transmission bits per iteration resulting in an increase in the quantization noise variance. It is now clear that Υ establishes a relationship between transmission bits and the graph connectivity. With quantization, the time evolution of (22) (by taking $f(\alpha_i) = \frac{1}{\alpha_i}$)

⁶We assume that the channel coefficients are varying slowly enough to be considered constant for the time interval necessary for the network to converge within a prescribed accuracy. This assumption is reasonable as our proposed algorithm converges rapidly.

now becomes:

$$\begin{aligned} \bar{T}_i^w[k+1] &= \bar{T}_i^w[k] - \frac{\epsilon}{\alpha_i} \sum_{j=1}^M e_{ij} (\bar{T}_i^w[k] - [\bar{T}_j^w[k]]_Q) \\ &= \bar{T}_i^w[k] - \frac{\epsilon}{\alpha_i} \sum_{j=1}^M e_{ij} (\bar{T}_i^w[k] - \bar{T}_j^w[k] - b_j[k]), \quad k \geq 0, \end{aligned}$$

for $i = 1, 2, \dots, M$ (29)

with $\bar{T}_i^w[0] = T_i$ in (3). (Note that the bar “-” differentiates from (22) where no quantization is used). Now $[\bar{T}_j^w[k]]_Q = \bar{T}_j^w[k] + b_j[k]$ represents quantization and $b_j[k]$ is the quantization noise independent of $w_i(n)$ in (1) and (2), $j = 1, 2, \dots, M, \forall i$ and $\forall n$. Assuming $\bar{T}_j^w[k] \in [0, 2U]$ and uniform quantization then:

$$\text{Var}\{b_j[k]\} = \sigma_{b_j}^2 = \frac{U^2}{3 \times 2^{2q}} \quad (30)$$

and we assume $\mathbb{E}\{b_j[k]\} = 0$ since the quantization noise is bipolar (i.e., it may take positive or negative values). We also assume that the i^{th} SN is capable to store its own *soft information* at the k^{th} iteration and communicate a quantized version to its neighbors. In the next $(k+1)^{th}$ iteration, every SN can update the test statistic (i.e., $\bar{T}_i^w[k+1]$) by using its own *soft information* and the *quantized information* received from other neighbors (i.e., it does not have access to their *soft information*).

Now, the power consumed by the whole network at a single iteration can be given as:

$$P_{throughout} = \sum_{i=1}^M \sum_{j=1}^M e_{ij} p_{ij}^t. \quad (31)$$

It is clear that establishing fewer communication links through (26) reduces $P_{throughout}$ and simultaneously imposes a slower information diffusion across the WSN. The number of bits that each SN can transmit to its neighbors will increase (see (28)). As a consequence, the quantization noise becomes negligible (see (30)). Alternatively, a smaller Υ (smaller quantization bits) dictates a more connected WSN and an increase in $P_{throughout}$ value. This results in an increase of quantization noise level that will tend to poor the detection performance.

It is now clear that Υ also establishes a trade-off between the quantization noise effect and the WSN total power⁷ consumption (P_T). In the simulation results section we will investigate the effect of the thresholding operation (26) on the P_T value as well as on the system detection performance. Therefore, the goal is to find an Υ_{opt} such that P_T and the detection performance are both improved. Next, we analyze the time evolution of (29) by using the probability of detection and the probability of false alarm as metrics.

B. Performance Analysis

Now we analyze the detection performance of the proposed distributed quantized (weighted) fusion rule (via the time evolution of (29)). Defining $\psi[k] = [\psi_1[k], \psi_2[k], \dots, \psi_M[k]]^T$

⁷The total power consumption is defined as $P_T = P_{throughout} K_T$, where K_T is the total number of iterations to run the time evolution (29) and (39) (i.e., $K_T = K_1 + K_2$) (see later section IV-C for details).

with $\psi_i[k] = \frac{1}{\alpha_i} \sum_{j=1}^M e_{ij} b_j[k]$ and so (29) can be written as:

$$\bar{\mathbf{T}}^w[k] = \mathbf{W}^k \bar{\mathbf{T}}^w[0] + \epsilon \sum_{z=1}^k \mathbf{W}^{z-1} \boldsymbol{\psi}[k-z], \quad k \geq 1 \quad (32)$$

where $\bar{\mathbf{T}}^w[k]$ is defined similarly to $\mathbf{T}^{eq}[k]$ in (17). The decision strategy for the i^{th} SN at the k^{th} iteration is again given in (18) (replacing $T_i^{eq}[k]$ by $\bar{T}_i^w[k]$), and the following also hold:

$$\mathbb{E} \{ \bar{T}_i^w[k] | \mathcal{H}_p \}_{p=\{0,1\}} = \left(\mathbf{W}^k \mathbb{E} \{ \bar{\mathbf{T}}^w[0] | \mathcal{H}_p \} \right)_i \quad (33)$$

$$\begin{aligned} \text{Var} \{ \bar{T}_i^w[k] | \mathcal{H}_p \}_{p=\{0,1\}} &= \left(\underbrace{\mathbf{W}^k \text{Cov} \{ \bar{\mathbf{T}}^w[0] | \mathcal{H}_p \} (\mathbf{W}^k)^T}_{(A)} \right)_{ii} \\ &+ \epsilon^2 \left(\underbrace{\sum_{z=1}^k \mathbf{W}^{z-1} \text{Cov} \{ \boldsymbol{\psi}[k-z] \} (\mathbf{W}^{z-1})^T}_{(B)} \right)_{ii} \end{aligned} \quad (34)$$

where $\text{Cov} \{ \boldsymbol{\psi}[k-z] \} = \frac{U^2}{3} \text{diag} \left(\frac{|\Delta_1|}{2^{2q}}, \frac{|\Delta_2|}{2^{2q}}, \dots, \frac{|\Delta_M|}{2^{2q}} \right)$. Now, the detection performance for the i^{th} SN at the k^{th} iteration can be evaluated using (21) (replacing $T_i^{eq}[k]$ by $\bar{T}_i^w[k]$) by substituting the expressions from (33) and (34). Note that as the dynamic system (32) evolves, the term (B) in (34) accumulates. Next we show how the detection performance for the i^{th} SN at the k^{th} iteration evolves by analyzing the variance term ($\text{Var} \{ \bar{T}_i^w[k] \}$) in (34).

Proposition 3: Assume that $\lambda_{max}(\mathbf{\Gamma}) \leq \frac{1}{\epsilon \lambda_{max}(\mathbf{L})(M-1)}$, where $\lambda_{max}(\mathbf{\Gamma})$ and $\lambda_{max}(\mathbf{L})$ are the maximum eigenvalues associated to $\mathbf{\Gamma}$ and \mathbf{L} respectively. From (34), the ‘‘scaled total variance’’

$$\begin{aligned} &\frac{1}{M-1} \sum_{i=1}^M \text{Var} \{ \bar{T}_i^w[k] \} \\ &\leq \text{Var}_k^{\max} \left(\frac{1}{M-1} + \lambda_2^k(\mathbf{W}) \right) + \epsilon^2 \sigma_{max}^2 \left(\frac{k}{M-1} + \frac{1 - \lambda_2^k(\mathbf{W})}{1 - \lambda_2(\mathbf{W})} \right) \end{aligned} \quad (35)$$

where $\text{Var}_k^{\max} = \max \{ \text{Var} \{ \bar{T}_1^w[k] \}, \dots, \text{Var} \{ \bar{T}_M^w[k] \} \}$, $\sigma_{max}^2 = \max \{ \text{Var} \{ \psi_1[k] \}, \text{Var} \{ \psi_2[k] \}, \dots, \text{Var} \{ \psi_M[k] \} \}$ and $\lambda_i(\mathbf{W})$, $i = 1, \dots, M$ are the eigenvalues of \mathbf{W} satisfying $\lambda_M \leq \lambda_{M-1} \leq \dots < \lambda_1 = 1$.

Proof: The proof can be found in Appendix B.

As k becomes large, it is clear that the second term of (35) grows and the performance of the distributed algorithm using quantized distributed weighted linear fusion does not approach the performance of the centralized quantized detector [10] (i.e., $\lim_{k \rightarrow \infty} P_d^i[k] \neq P_d$ in (13) of [10], $\forall i$).

Now, it is also clear that k establishes a trade-off between the local SNs test statistic improvement and the quantization error degradation. There is a finite optimum $k = K_1$ to stop the SNs collaboration (see later), but after that the quantization error overcomes the improvement gained from this collaboration. So, using this framework (i.e., the consensus algorithm with quantization matched to the channel capacity) we will now propose a two-step approach (still using quantized test statistics shared among SNs) that will perform comparable to the optimum unquantized centralized detector in section III-A2 (i.e., when using a FC and no quantization). And what

is more important, it converges across the network in a finite number of iterations.

C. Proposed two-step quantized distributed weighted fusion rule algorithm

(i) FIRST STEP: Run the quantized consensus algorithm in (32) to improve the local version of the test statistic at each SN. But then terminate the algorithm at $k = K_1$ (where the optimum value of K_1 is found later from simulation results and a sub-optimum solution to it is also proposed). We now have $\{ \bar{T}_i^w[K_1] \}_{i=1}^M$ from (32) and we will use this to generate a binary indicator random variable $I_i[0]$ as follows

$$\left. \begin{aligned} &\text{if } \bar{T}_i^w[K_1] < \Lambda_1, I_i[0] = 0 \\ &\text{if } \bar{T}_i^w[K_1] \geq \Lambda_1, I_i[0] = 1 \end{aligned} \right\} \quad (36)$$

where Λ_1 is a local (first step) detection threshold that is the same for all M SNs. We will now propose (for performance comparison purposes) two alternative *second step* decision rules:

$$1) \quad \left. \begin{aligned} &\text{if } \bar{T}_f^w[K_1] \neq M, \text{decide } \mathcal{H}_0 \\ &\text{if } \bar{T}_f^w[K_1] = M, \text{decide } \mathcal{H}_1 \end{aligned} \right\} \quad (37)$$

$$2) \quad \left. \begin{aligned} &\text{if } \bar{T}_f^w[K_1] = 0, \text{decide } \mathcal{H}_0 \\ &\text{if } \bar{T}_f^w[K_1] \neq 0, \text{decide } \mathcal{H}_1 \end{aligned} \right\} \quad (38)$$

where $\bar{T}_f^w[K_1] = \sum_{i=1}^M I_i[0]$ both in 1) and 2). But the problem is now how to evaluate $\bar{T}_f^w[K_1]$ in a distributed manner across SNs. This will be explained in the second step.

(ii) SECOND STEP:

1) *Second step defined in (37):* When the local individual SNs unanimously decide on the intruder presence, so decides this (global) decision *second step* (i.e., intruder is present). Otherwise, it decides that the intruder is not present. Here we will use [38] to show how to effectively evaluate (37) by first sharing $\{ I_i[0] \}_{i=1}^M$ and then iteratively updating across the SNs as follows:

$$\begin{aligned} I_i[k+1] &= I_i[k] \bigwedge \left(\bigwedge_{j \in \Delta_i} I_j[k] \right), \quad k = 0, 1, 2, \dots, K_2-1, \\ &\text{for } i = 1, 2, \dots, M \end{aligned} \quad (39)$$

where K_2 is the diameter of network⁸, ‘‘ \bigwedge ’’ denotes the logical ‘‘and’’ operation and Δ_i is defined for (5). Note that no quantization is needed and all $I_i[K_2]$ converge to either 1 or 0. So now we can easily show:

$$\left. \begin{aligned} &\text{If } I_i[K_2] = 0, \forall i \Rightarrow \bar{T}_f^w[K_1] \neq M, \text{decide } \mathcal{H}_0 \\ &\text{If } I_i[K_2] = 1, \forall i \Rightarrow \bar{T}_f^w[K_1] = M, \text{decide } \mathcal{H}_1. \end{aligned} \right\} \quad (40)$$

and so $I_i[K_2]$ (at any arbitrary i^{th} SN) can be used to implement the decision rule (37).

⁸The geodesic distance between two nodes in a (connected) graph is the number of the edges (i.e., links) in the shortest path connecting these two nodes. The diameter of a graph is the maximum geodesic distance taken over all possible pairs of nodes in the graph.

2) *Alternative second step defined in (38)*: Now, this alternative *second step* (global) decision fusion rule decides on the presence of the intruder if at least any arbitrary local SNs (at iteration $k = K_1$) has decided so. Again, $\bar{T}_f^w[K_1]$ can be evaluated in a distributer manner by first sharing $\{I_i[0]\}_{i=1}^M$ and then iteratively updating across the SNs using (39) (but now the “and” logical operation “ \wedge ” is replaced with the “or” logical operation “ \vee ”). Like before, all $I_i[K_2]$ converge to either 1 or 0 and we can easily show:

$$\left. \begin{array}{l} \text{If } I_i[K_2] = 0, \forall i \Rightarrow \bar{T}_f^w[K_1] = 0, \text{ decide } \mathcal{H}_0 \\ \text{If } I_i[K_2] = 1, \forall i \Rightarrow \bar{T}_f^w[K_1] \neq 0, \text{ decide } \mathcal{H}_1. \end{array} \right\} (41)$$

and so $I_i[K_2]$ (at any arbitrary SN) can be used to implement the decision rule (38). Overall, the proposed two-step fully distributed algorithm requires $(K_T = K_1 + K_2)$ iterations in total. Now, the two-step algorithm (with *second step* decision rule (38) can be summarized in **Algorithm 1**.

Algorithm 1: Distributed Detection via Two-Step Consensus Algorithm

STEP 1: Choose Υ and evaluate $\bar{T}_i^w[0] = T_i, \forall i$ in (3);

STEP 2: Choose an approximation model ((42) or (43)) to estimate K_1 and compute $\bar{T}_i^w[k], \forall i$ using (32) with $k = K_1$;

STEP 3: Generate the binary indicator random variable at each SN:
 if $\bar{T}_i^w[K_1] < \Lambda_1, I_i[0] = 0$
 if $\bar{T}_i^w[K_1] \geq \Lambda_1, I_i[0] = 1.$

STEP 4: Run (39) with $I_i[0]$ generated in step 3 for K_2 iterations to effectively perform the final test:

$$\left. \begin{array}{l} \text{if } \bar{T}_f^w[K_1] = 0, \text{ decide } \mathcal{H}_0 \\ \text{if } \bar{T}_f^w[K_1] \neq 0, \text{ decide } \mathcal{H}_1 \end{array} \right\}$$

$$\text{where } \bar{T}_f^w[K_1] = \sum_{i=1}^M I_i[0].$$

Next, in the simulation results, we will show that the first step spatial collaboration among SNs is crucial for the system detection performance and also for the network total power consumption. We will also show via simulations that there is an optimum K_1 (for both decision fusion (37) and (38)) such that the system detection performance is maximized. Then, we propose a sub-optimum but simple solution to find this optimum K_1 .

V. SIMULATIONS RESULTS

Here we will analyze the performance of our proposed two-step quantized (weighted) fusion rule algorithm for distributed detection deployment. First we have a WSN with M SNs with arbitrary SN geometry, where the distances d_{ij} in (27) between SNs i and j are assumed to be known. The other parameters in (27) are $p_{ij}^t = 300, \gamma = 2, \zeta_0 = 0.1$ and h_{ij}^2 is an exponential random variable (r.v.) with mean $\mu_{h_{ij}^2} = 30$. Using the r.v. SNR_{ij} in (27) in (26), we then construct two example topologies for different values of Υ (see Fig. 4). These topologies will be used later for Fig. 15 and Fig. 16. To provide results of more general validity, we also report the average performance where the average is carried out over 500 channel realizations unless otherwise stated. We now generate

the test statistics $\bar{T}_i^w[K_1]$ in (36), via (32) for $k = K_1$. As previously explained, any $I_i[K_2]$ in (40) or (41) can be used to decide either \mathcal{H}_0 or \mathcal{H}_1 , and this will define the new global detection and false alarm probabilities (i.e., P_d^g and P_{fa}^g respectively). Here we use 10^5 Monte-Carlo simulations. Finally, $\xi_a = 10 \log_{10} \left(\frac{1}{M} \sum_{i=1}^M \xi_i \right) = -9.5$ dB unless otherwise stated, where $\xi_i = \sum_{n=1}^N s_i^2(n) / N \sigma_i^2$. We will also refer to “equal weight” combining in (15) (i.e., $\alpha_i = 1, \forall i$) and use this as a benchmark. Finally, we choose L_i with equality in (6). The detection performance of the proposed two-step algorithm is also compared with the *centralized* soft Likelihood Ratio Test (LRT) based fusion rule in [11].

A. Validity of quantization noise assumption for low bit rate

Before we investigate the performance of the proposed two-step detection algorithm, we evaluate via simulations the mismatch between the assumed uniform quantization and the actual quantization for low bit rate. In Fig. 3, we show the probability distribution function (PDF) of the quantization error for $q = 2$ bits and $q = 3$ bits. The quantization error variance (σ_e^2) versus the number of quantization bits (q) is also plotted. In the case of $q = 2$ bits, the uniform (quantization

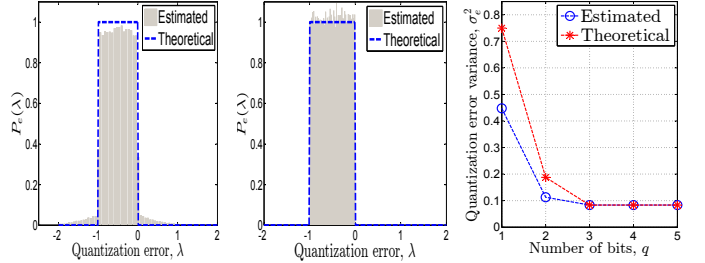


Fig. 3. Quantization error mismatch: (left/middle) probability distribution function (PDF) $P_e(\lambda)$ for $q = 2$ bits/ $q = 3$ bits; (right) quantization error variance (σ_e^2) mismatch versus number of quantization bits.

error) PDF is an approximation. However, in the case of $q = 3$ bits, this approximation is quite accurate. As a result, we conclude that the assumption of a uniform (quantization error) PDF is a valid assumption (or at least for the simulation set-up considered in this paper).

B. Impact of channel estimation on the network density

Now, we investigate the channel estimation error effect on the network density (ρ) versus the SNR threshold (Υ). We model the channel estimation error as a Gaussian random variable (i.e., $\hat{h}_{ij} = h_{ij} + e_h$) where $e_h \sim \mathcal{N}(0, \sigma_{e_h}^2)$ and \hat{h}_{ij} is the estimated flat fading channel coefficient.

In Fig. 4, we plot the network density¹⁰ (ρ) versus the SNR threshold (Υ) for different values of the estimation error variance ($\sigma_{e_h}^2$). For small $\sigma_{e_h}^2$, the network density is shown to be robust against the channel estimation error. That is not the case for relatively large $\sigma_{e_h}^2$ where a performance mismatch is observed.

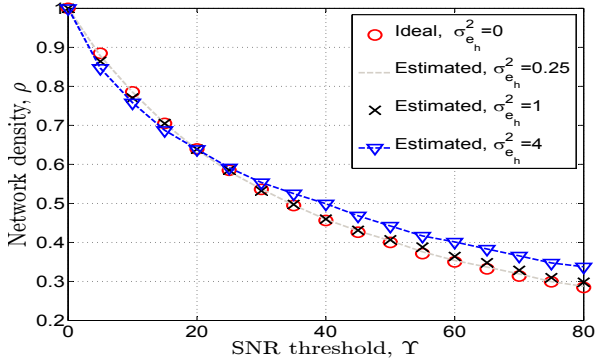


Fig. 4. Averaged (over 10000 \hat{h}_{ij}^2 realizations) network density (ρ) versus Υ in (26), with $U = 3$, $N = 20$, and $M = 17$.

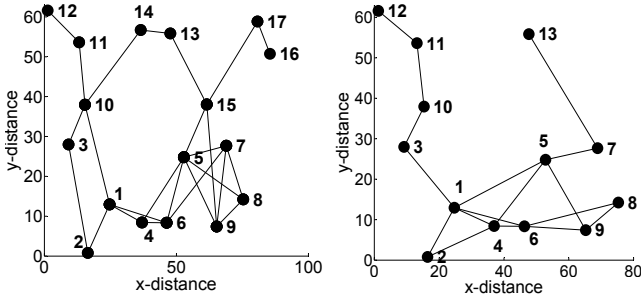


Fig. 5. Two different communication topologies (generated via ((26) and (27)), with $\sigma_{e_h}^2 = 0$ and the quantization bits following (28): (left) $M = 17$, $\Upsilon = 20$, $q = 2$ bits; (right) $M = 13$, $\Upsilon = 72$, $q = 3$ bits.

C. Impact of thresholding operation on the system detection performance and total power consumption

In section IV, we have shown that the link SNR threshold (Υ) parameter establishes a relation between the number of bits that each SN has to transmit to its neighbors and the topology of the network that defines the connections among them (see definitions (26)-(28)). It is then very important to investigate the impact of the Υ parameter on the system (global) detection performance (P_d^g) and on the total power consumption (P_T).

In Fig. 6 we plot the global probability of detection (P_d^g) versus Υ for different numbers of SNs (M) and for a fixed global probability of false alarm (P_{fa}^g) and K_1 . We observe that there is an optimum Υ that maximizes P_d^g for any arbitrary M .

Now, to give a more general validity on the results, in Fig. 7 we show the conventional (unquantized) consensus-based algorithm (22) (with the decision rule (18) by substituting $T_i^{eq}[k]$ with $T_i^w[k]$) and the proposed two-step (quantized) weighted fusion rule summarized in **Algorithm 1**: (upper plot) the average total power consumption $\mathbb{E}[P_T]$ (refer for its definition to (31) and below) versus the link SNR threshold (Υ); (middle plot) the global achievable⁹ probability of detection (P_d^*) versus link SNR threshold (Υ); (lower plot) the

⁹The global achievable probability of detection (P_d^*) (for a fixed Υ) is defined as the best global probability of detection (P_d^g) with respect to K_1 .

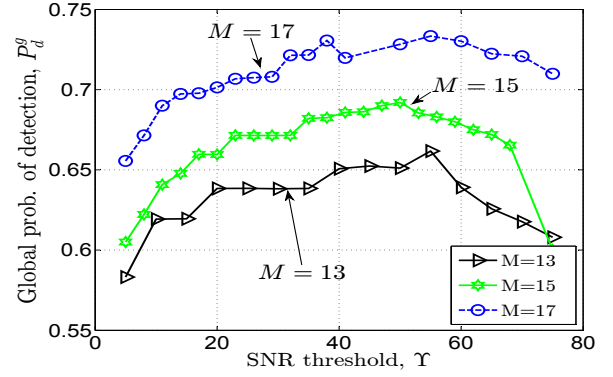


Fig. 6. Averaged (over 500 h_{ij}^2 realizations) global probability of detection (P_d^g) (using two-step approach) versus Υ in (26), $\sigma_{e_h}^2 = 0$, with decision fusion in (40), $P_{fa}^g = 0.2$, $U = 2$, $N = 20$, $K_1 = 10$ and $\alpha_i = 1, \forall i$ in (29).

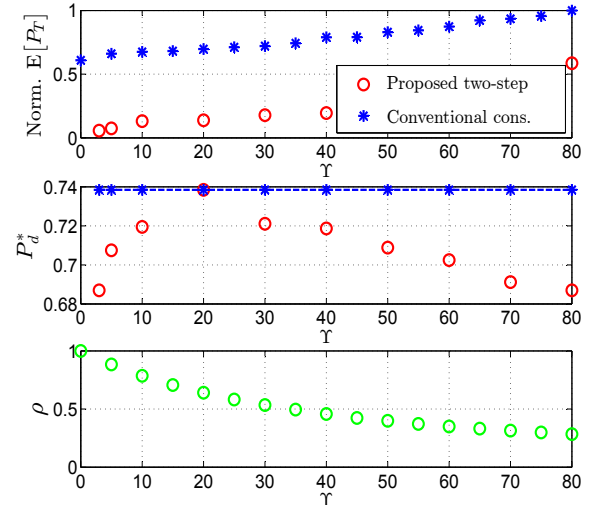


Fig. 7. Normalized average power consumption ($\mathbb{E}[P_T]$), achievable⁸ probability of detection (P_d^*) and the average communication link density (ρ) versus Υ in (26), with $\sigma_{e_h}^2 = 0$, decision fusion in (41), $P_{fa}^g = 0.1$, $U = 3$, $N = 20$, $M = 17$ and with α_i (scaled by M) in (12).

average network density¹⁰ (ρ) versus the link SNR threshold (Υ). Even-though the comparison made is not fair (i.e., for the proposed (quantized) two-step weighted fusion rule versus the (unquantized) conventional consensus-based fusion rule), clearly our proposed two-step fusion rule algorithm possesses the following: a) it requires much less power budget for all Υ compared to the (unquantized) conventional consensus-based algorithm, and b) converges across the WSN much faster and in a finite number of iterations ($K_T = K_1 + K_2$), whereas for the conventional consensus-based, the convergence holds in limit. Finally, in the lower plot we verify (as expected) that a smaller/larger Υ dictates a more/less connected WSN respectively. In the case of the conventional consensus-based algorithm, the convergence criteria we use here is the relative absolute difference: $\frac{\|\mathbf{T}^w[k+1] - \mathbf{T}^w[k]\|}{\|\mathbf{T}^w[k]\|} \leq \kappa$, where $\kappa = 10^{-7}$. The averages are performed over 500 (h_{ij}^2) realizations.

¹⁰The average network density ρ is defined as: $\rho = \mathbb{E} \left[\frac{\sum_{i=1}^M \sum_{j=1}^M e_{ij}}{M(M-1)} \right]$.

D. Impact of the K_1 parameter on the system detection performance

The first step *quantized collaboration* establishes a linear spatial collaboration among M SNs up to K_1 iterations for improving the overall detection performance. We have shown analytically (see proposition 3 and below) that the RHS of (35) diverges for $k = K_1$ (when K_1 is large) and the detection performance eventually declines. Next, we investigate (through simulations) the effect that (K_1) has on the global detection performance (P_d^g) and propose a sub-optimum (but simple) solution to evaluate K_1 .

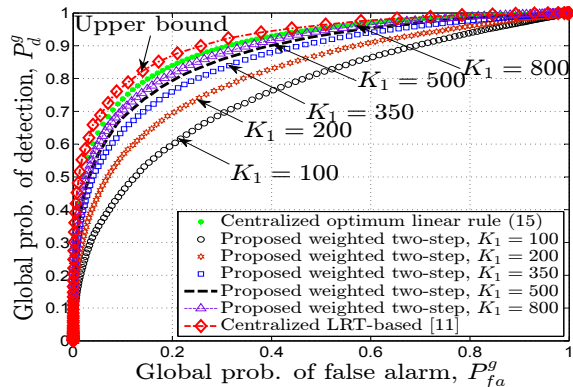


Fig. 8. Averaged (over 500 h_{ij}^2 realizations) ROC for the proposed two-step weighted algorithm with decision fusion in (40), $U = 3$, $N = 20$, $M = 17$, $K_2 = 3$, $\Upsilon = 30$, $\sigma_{e_h}^2 = 0$ and with α_i (scaled by M) in (12).

1) *Optimal numerical solution to K_1* : Now, in Fig. 8 we report the (averaged) receiver operating characteristic (ROC) against the first step number of iterations (K_1) for the proposed distributed two-step (weighted) algorithm with decision fusion in (40). As K_1 increases then P_d^g improves. In Fig. 9 we report

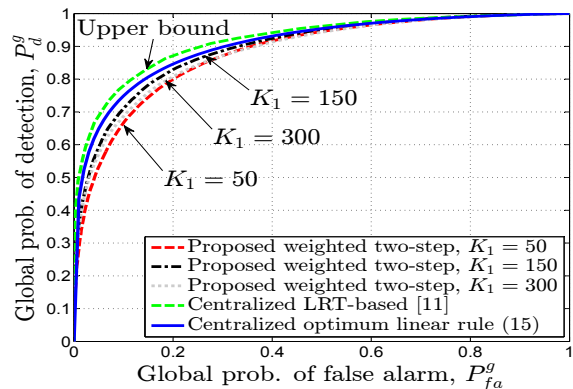


Fig. 9. Averaged (over 500 h_{ij}^2 realizations) ROC against first step iterations number (K_1), with decision fusion in (41), $K_2 = 2$, $U = 3$, $N = 20$, $M = 17$, $\Upsilon = 10$, $\sigma_{e_h}^2 = 0$ and with α_i (scaled by M) in (12).

the same for the proposed two-step (weighted) algorithm but now with the decision fusion in (41). As expected, the detection performance improves up to $K_1 = 150$ and after that it degrades. Then, in Fig. 10 we plot the (averaged) global detection performance (P_d^g) (for a fixed global probability of false alarm (P_{fa}^g)) versus first step number of iterations (K_1) for different link SNR thresholds (Υ). We observe that there

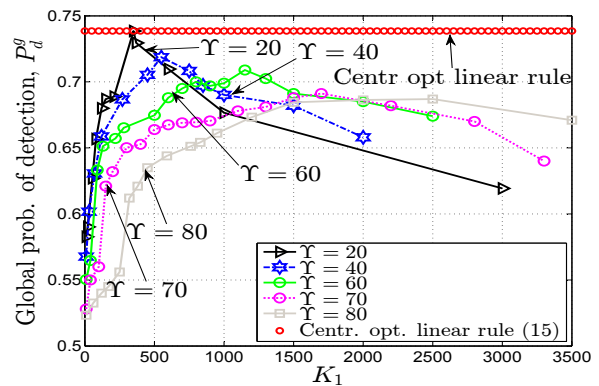


Fig. 10. Averaged (over 500 h_{ij}^2 realizations) global probability of detection (P_d^g) versus first step iterations number (K_1), with decision fusion in (41), $P_{fa}^g = 0.1$, $U = 3$, $N = 20$, $M = 17$, $\sigma_{e_h}^2 = 0$ and with α_i (scaled by M) in (12).

exists an optimum K_1 to run the first step time evolution (29) such that P_d^g is maximized for any arbitrary Υ . We also note that the best performance is attained for $\Upsilon = 20$.

Now, selecting the pair ($\Upsilon = 20$, $K_1 = 320$) (i.e., the Υ and K_1 that attain the best performance in Fig. 10 with the decision fusion in (41)), in Fig. 11 we examine the P_d^g performance against ξ_a for the proposed distributed two-step (weighted) algorithm assuming: (left) ideal channel estimation; (right) non-ideal channel estimation. Interestingly, (for the ideal channel case) the proposed two-step (weighted) algorithm performance (with decision fusion in (41)) attains its centralized counterpart's upper bound performance for all ξ_a . So, it is now clear that the optimum values of parameters Υ and K_1 are independent of ξ_a (i.e., the local ξ_i). This independence is important as it shows that the algorithm is robust against the local ξ_i and allows evaluating these parameters once at the beginning. We also observe that the proposed two-step performance with decision fusion (41) is the same (at low SNR) as that of decision fusion (40), but at high SNR it outperforms the latter. Now (for the non-ideal case), we can observe a slight detection performance degradation for the proposed two-step algorithm. Next, we propose (for the two-step algorithm with the second step decision rule (41)) a sub-optimum (but simple) solution to the optimum K_1 . Note that the extension with the second step decision fusion rule (40) is straight forward.

2) *Suboptimal solution to K_1* : Now, through simulation results shown in Fig. 10, we get an insight on how the optimum K_1 is related to the link SNR threshold (Υ). We also notice that an increase in Υ is translated into a corresponding increase in the optimum K_1 value (i.e., K_1 that corresponds to the maximum P_d^g). This result is not surprising and can be explained by the fact that a smaller Υ dictates a more connected graph (see (26)) and an increase in Υ dictates a sparse graph (hence more iterations are needed to diffuse the information across the SNs). Motivated by this fact, we now relate the first step iterations number (K_1) to the link SNR threshold parameter (Υ) with two fitting models:

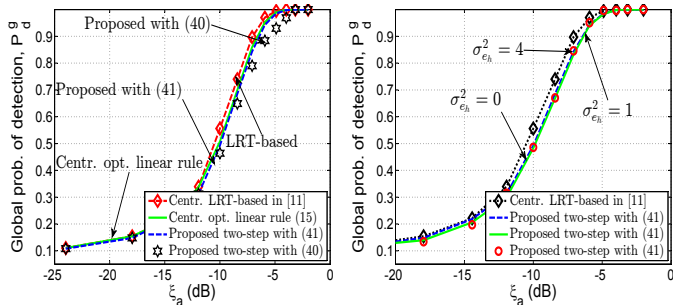


Fig. 11. Averaged (over 500 h_{ij}^2 realizations) probability of detection (P_d^g) against the signal to noise ratio (ξ_a) with $P_{fa}^g = 0.1$, $U = 3$, $N = 20$, $M = 17$, $K_1 = 320$, $\Upsilon = 20$, $\xi_i = \xi$, $\forall i$ in (4) and with α_i (scaled by M) in (12): (left) ideal, $\sigma_{e_h}^2 = 0$; (right) non-ideal, $\sigma_{e_h}^2 \neq 0$.

(i) *Exponential model* :

$$K_1 \approx g(\Upsilon) = \begin{cases} A \exp(b\Upsilon) & \text{type 1} \\ A \exp(b\Upsilon) + B \exp(c\Upsilon) & \text{type 2} \end{cases} \quad (42)$$

(ii) *Power model* :

$$K_1 \approx g(\Upsilon) = \begin{cases} A\Upsilon^b & \text{type 1} \\ A\Upsilon^b + C & \text{type 2} \end{cases} \quad (43)$$

where A , B , C , b and c are the coefficients given in Table I obtained using Matlab (Nonlinear Least Squares method and Trust-Region algorithm).

TABLE I
PARAMETERS FOR DIFFERENT FITTING MODELS

Model	Type	A	B	C	b	c	RMSE
Exponential	Exp 1	173.6	x	x	0.03319	x	49.82
	Exp 2	188.1	1.561e-014	x	0.03166	0.4584	53.09
Power	Pow 1	0.5976	x	x	1.89	x	136.27
	Pow 2	0.0079	x	2.853	338.9	x	63.65

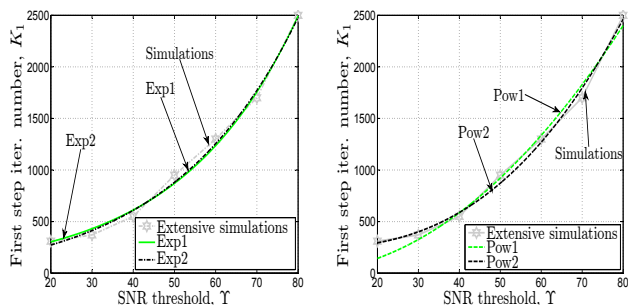


Fig. 12. First step iterations number (K_1) versus Υ in (26), with $U = 3$, $N = 20$, $M = 17$ and with α_i (scaled by M) in (12): (left) exponential fitting model; (right) power fitting model.

Now, in Fig. 12 we plot the first step number of iterations (K_1) versus the link SNR threshold (Υ) for two different fitting models (i.e., *exponential* and *power model*) and then compare these to the simulations. Clearly, the *exponential* of *type 1* is the best candidate as it attains the minimum RMSE (see Table I).

E. Detection Performance Comparison

We now compare the (averaged) global detection performance among/with: (a) the two-step (quantized) distributed weighted fusion rule algorithm with second step in (40) and (41), (b) the two-step (quantized) distributed equal combining fusion rule with second step in (40) and (41), (c) the optimum centralized (quantized) weighted fusion rule proposed in [10], and (d) the centralized (quantized) equal combining in [10].

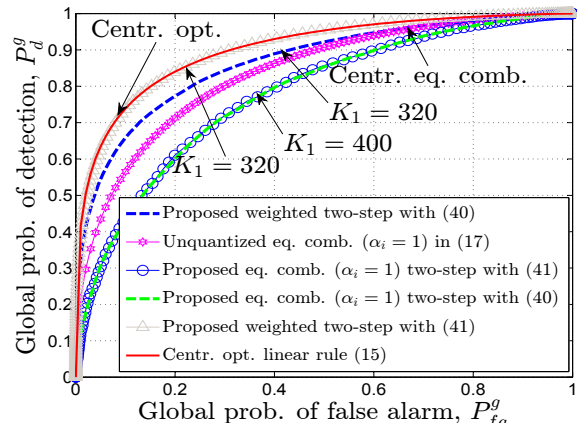


Fig. 13. Averaged (over 500 h_{ij}^2 realizations) ROC for the proposed (quantized) two-step weighted fusion rule with $U = 3$, $N = 20$, $\Upsilon = 20$, $M = 17$ and with α_i (scaled by M) in (12).

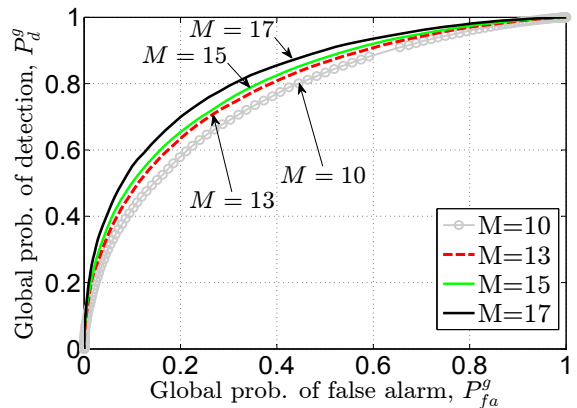


Fig. 14. ROC for the proposed (quantized) two-step distributed scheme with $\Upsilon = 20$ in (26), $U = 2$, $N = 20$, $K_1 = 10$ and $\alpha_i = 1$, $\forall i$ in (29).

In Fig. 13 we report the ROC for the two different schemes (i.e., centralized and distributed two-step). As can be seen, the distributed two-step algorithm approaches the upper bound (i.e., the centralized unquantized scheme performance in (15)). Now, we examine in Fig. 14 the ROC parametrized against M for the distributed (equal combining) two-step algorithm, illustrating how P_d^g improves as M increases. The ROC performance¹¹ among different (equal combining) schemes is illustrated in Fig. 15 and Fig. 16. In Fig. 15 we show the advantage of our proposed distributed two-step (equal combining) scheme over only the first step part (at SN 6). Also, if Υ is carefully chosen the distributed two-step

¹¹SN 6 in Fig. 15 and SN 3 in Fig. 17 were chosen for comparison purposes as they possess the best performances among M SNs for each case.

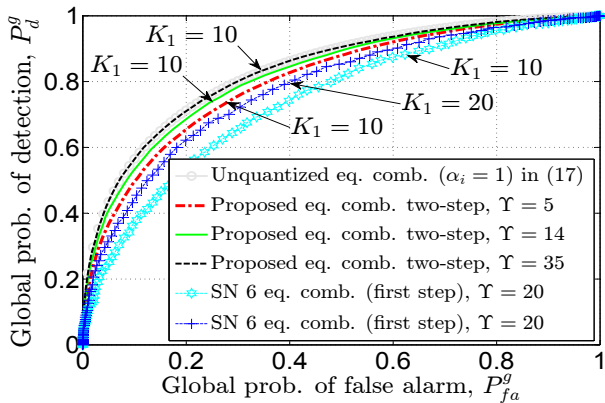


Fig. 15. ROC with $U = 2$, $N = 20$, $M = 17$ and topology given in left of Fig. 5 and $\alpha_i = 1, \forall i$ in (29).

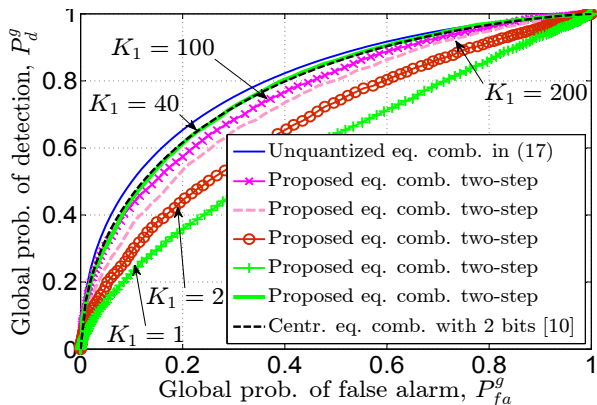


Fig. 16. ROC with $U = 2$, $N = 20$, $M = 13$ and topology given in right of Fig. 5 and $\alpha_i = 1, \forall i$ in (29).

(equal combining) scheme performance approaches that of the (equal combining) centralized detector (i.e., with FC and no quantization) in (18). Fig. 16 shows the ROC for the proposed quantized (3 bits) distributed (equal combining) two-step algorithm against K_1 compared to the quantized (2 bits) centralized (equal combining) scheme in [10]. As expected (similar to the weighted two-step), there is an optimum K_1 that maximizes P_d^g and after that P_d^g decreases.

Finally, Fig. 17 plots the P_d^g performance characterization against the average SNR (ξ_a) for 4 different (equal combining) schemes showing the performance improvement of our proposed distributed two-step algorithm.

VI. CONCLUSION

In this paper, we propose a fully distributed two-step consensus-based detection algorithm via SNs sharing with their neighbors a quantized version of the received energy test statistic. We relate the communication topology with the number of bits to be shared among SNs and through numerical results we show that there is an optimum topology (for a fixed first step number of iterations (K_1)) such that P_d^g (the global probability of detection) is maximized. In addition, we show that there is an optimum K_1 to terminate the first step SN collaboration (for any arbitrary topology) and after that the P_d^g

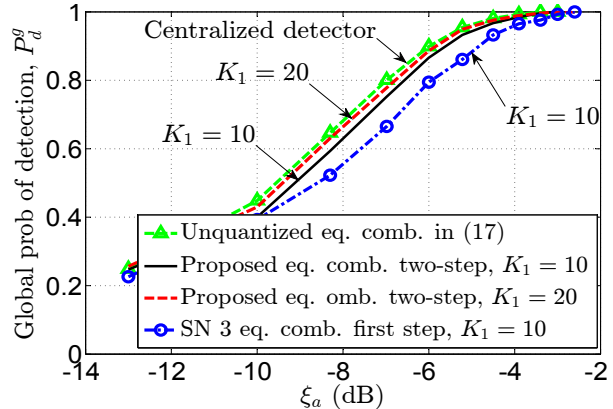


Fig. 17. Probability of detection (P_d^g) versus the signal to noise ratio (ξ_a) for $M = 13$, $\Upsilon = 72$, $U = 2$, $N = 20$, $P_{fa}^g = 0.1$, $\xi_i = \xi \cdot \forall i$ in (4) and $\alpha_i = 1, \forall i$ in (29). The topology used is given in right of Fig. 5.

performance declines. When parameters K_1 and Υ (the link SNR threshold in (26)) are appropriately chosen, the detection performance of the proposed quantized distributed two-step algorithm approaches the unquantized centralized optimum combining scheme performance of (15). Overall, the algorithm requires a finite number of iterations ($K_1 + K_2$). For example, targeting the optimum P_d^g (see Fig. 7 (middle plot) at $\Upsilon = 20$), our proposed two-step algorithm requires roughly 50% less power consumption (P_T) than the conventional consensus-based algorithm. Future work will investigate the analysis of the problem for time-varying SNs interaction topologies.

APPENDIX A PROOF OF PROPOSITION 2

Let $\mathbf{W} \geq 0$ be defined as in (24) with $0 < \epsilon < 1/\Delta_{max}$ and $\Gamma \leq 1$. Since we have assumed that the WSN forms a connected graph, then \mathbf{W} is irreducible and also primitive (i.e., the maximum eigenvalue has multiplicity one). So by the Perron Frobenius theorem [35], \mathbf{W} has unique left and right eigenvectors corresponding to the maximum eigenvalue¹² and also $\lim_{k \rightarrow \infty} \mathbf{W}^k = \mathbf{v}^r (\mathbf{v}^l)^T$, where $\mathbf{v}^l = [v_1^l, v_2^l, \dots, v_M^l]^T$ is the left and $\mathbf{v}^r = [v_1^r, v_2^r, \dots, v_M^r]^T$ is the right eigenvector corresponding to the maximum eigenvalue of \mathbf{W} . The problem is now finding these eigenvectors. Consider

$$\mathbf{W} \mathbf{v}^r = \mathbf{v}^r - \epsilon \Gamma \mathbf{L} \mathbf{v}^r = \mathbf{v}^r.$$

Now, the above relation is equivalent to $\epsilon \Gamma \mathbf{L} \mathbf{v}^r = \mathbf{0}$. It can be easily shown that if \mathbf{v}^r is in the right null space of \mathbf{L} (i.e., $\mathbf{L} \mathbf{v}^r = \mathbf{0}$), it is also true that \mathbf{v}^r is in the right null space of $(\Gamma \mathbf{L})$. Using this fact and the definition of \mathbf{L} (i.e., symmetric real matrix with rows and columns summing to zero), we can easily show that $\mathbf{v}^r = c_r [1, 1, \dots, 1]^T$ (where c_r is a positive constant (see later)). Similarly, we can find the left eigenvector (\mathbf{v}^l) by using the following:

$$(\mathbf{v}^l)^T \mathbf{W} = (\mathbf{v}^l)^T - \epsilon (\mathbf{v}^l)^T \Gamma \mathbf{L} = (\mathbf{v}^l)^T.$$

¹²For a connected graph the maximum eigenvalue of \mathbf{W} is unity (i.e., the zero eigenvalue associated to \mathbf{L} has multiplicity one).

Again, the above relation can be equivalently expressed as $\epsilon(\mathbf{v}^l)^T \mathbf{\Gamma} \mathbf{L} = \mathbf{0}$. Using the same analogy (like in the case of right eigenvector) we can also show that $(\mathbf{v}^l)^T$ is in the left null space of $\mathbf{\Gamma} \mathbf{L}$ if $(\mathbf{v}^l)^T \mathbf{\Gamma} = c_l[1, 1, \dots, 1]^T$ (i.e., if $v_i^l = \frac{c_l}{f(\alpha_i)}, \forall i$). Choosing $c_r = 1$ and $c_l = \frac{1}{\sum_{i=1}^M \frac{1}{f(\alpha_i)}}$ such that $(\mathbf{v}^r)^T \mathbf{v}^l = 1$, we can now easily show that:

$$\lim_{k \rightarrow \infty} \left(\mathbf{W}^k \mathbf{T}^w [0] \right)_i = \left(\mathbf{v}^r (\mathbf{v}^l)^T \mathbf{T}^w [0] \right)_i = \frac{\sum_{i=1}^M \frac{1}{f(\alpha_i)} T_i^w [0]}{\sum_{i=1}^M \frac{1}{f(\alpha_i)}}, \forall i. \quad (44)$$

This concludes the proof.

APPENDIX B PROOF OF PROPOSITION 3

Let \mathbf{W} be defined as in (24) with $0 < \epsilon < 1/\Delta_{max}$, $\mathbf{\Gamma} \leq 1$ and $f(\alpha_i) = \frac{1}{\alpha_i}, \forall i$. We complete the main proof as follows: 1) prove that the $(\mathbf{\Gamma} \mathbf{L})$ has both real and positive eigenvalues, and then 2) prove that the \mathbf{W} is a positive semi-definite matrix if $\lambda_{max}(\mathbf{\Gamma}) \leq \frac{1}{\epsilon \lambda_{max}(\mathbf{L})(M-1)}$, where $\lambda_{max}(\mathbf{\Gamma})$ and $\lambda_{max}(\mathbf{L})$ are the maximum eigenvalues associated to $\mathbf{\Gamma}$ and \mathbf{L} respectively, and finally 3) derive the upper bound on the ‘‘scaled total variance’’ at each SN.

Sub-proof 1: Consider the matrix multiplication $\mathbf{\Gamma} \mathbf{L}$ (which gives a non-symmetric matrix) with $\mathbf{\Gamma}$ defined below (24) and \mathbf{L} defined in section II-B. Note that $\mathbf{\Gamma}$ and \mathbf{L} are real diagonal and real symmetric matrices respectively by definition. It is not difficult to show that the eigenvalues of $\mathbf{\Gamma} \mathbf{L}$ are the same as the eigenvalues of $\mathbf{K} = \mathbf{\Gamma}^{-\frac{1}{2}} (\mathbf{\Gamma} \mathbf{L}) \mathbf{\Gamma}^{\frac{1}{2}}$. Now, \mathbf{K} can be simplified to $(\mathbf{\Gamma}^{\frac{1}{2}} \mathbf{L} \mathbf{\Gamma}^{\frac{1}{2}})$ (a real symmetric positive semi-definite matrix) which implies that the eigenvalues of $(\mathbf{\Gamma} \mathbf{L})$ are real and positive. This concludes the sub-proof 1.

Sub-proof 2: Now, to ensure that \mathbf{W} is positive semi-definite we require $\mathbf{z}^T \mathbf{W} \mathbf{z} \geq 0$ for $\mathbf{z} \neq \mathbf{0}$. Decomposing \mathbf{W} as:

$$2\mathbf{W} = \underbrace{(\mathbf{W} + \mathbf{W}^T)}_{\text{symmetric}} + \underbrace{(\mathbf{W} - \mathbf{W}^T)}_{\text{skew-symmetric}} \quad (45)$$

then, it can be shown that $\mathbf{z}^T \mathbf{W} \mathbf{z} \geq 0$ iff $\frac{\mathbf{z}^T (\mathbf{W} + \mathbf{W}^T) \mathbf{z}}{2} \geq 0$ (since $\frac{\mathbf{z}^T (\mathbf{W} - \mathbf{W}^T) \mathbf{z}}{2} = 0$). Now, $0 \leq \frac{\mathbf{z}^T (\mathbf{W} + \mathbf{W}^T) \mathbf{z}}{2} = \frac{\mathbf{z}^T \mathbf{W} \mathbf{z}}{2} + \frac{(\mathbf{z}^T \mathbf{W} \mathbf{z})^T}{2} = \mathbf{z}^T \mathbf{W} \mathbf{z} \implies \mathbf{W}$ is positive semi-definite iff $\left(\frac{\mathbf{W} + \mathbf{W}^T}{2} \right)$ is so. Now from [39], $\lambda_i(\mathbf{W} + \mathbf{W}^T) \geq 0, \forall i \implies \mathbf{z}^T \mathbf{W} \mathbf{z} \geq 0$ and from (24) we can easily show that: $\lambda_i(\mathbf{W} + \mathbf{W}^T) = 2 - \epsilon \lambda_i(\mathbf{\Gamma} \mathbf{L} + (\mathbf{\Gamma} \mathbf{L})^T)$. Now, it is clear that:

$$\lambda_i(\mathbf{\Gamma} \mathbf{L} + (\mathbf{\Gamma} \mathbf{L})^T) \leq \frac{2}{\epsilon} \implies \lambda_i(\mathbf{W} + \mathbf{W}^T) \geq 0 \implies \lambda_{max}(\mathbf{\Gamma} \mathbf{L} + (\mathbf{\Gamma} \mathbf{L})^T) \leq \frac{2}{\epsilon} \implies \lambda_i(\mathbf{W} + \mathbf{W}^T) \geq 0. \quad (46)$$

Using the result in sub-proof 1 and (46), then:

$$\lambda_{max}(\mathbf{\Gamma} \mathbf{L} + (\mathbf{\Gamma} \mathbf{L})^T) \leq 2 \sum_{i=1}^M \lambda_i(\mathbf{\Gamma} \mathbf{L}) \leq 2(M-1) \lambda_{max}(\mathbf{\Gamma} \mathbf{L})$$

So, $\lambda_{max}(\mathbf{\Gamma} \mathbf{L}) \leq \frac{1}{\epsilon(M-1)} \implies \lambda_i(\mathbf{W} + \mathbf{W}^T) \geq 0.$ (47)

Because of the structure of $\mathbf{\Gamma}$ and \mathbf{L} , then from [35]:

$$\lambda_{max}(\mathbf{\Gamma} \mathbf{L}) \leq \lambda_{max}(\mathbf{\Gamma}) \lambda_{max}(\mathbf{L}) \quad (48)$$

and from (47) and (48) we can show:

$$\lambda_{max}(\mathbf{\Gamma}) \leq \frac{1}{\epsilon \lambda_{max}(\mathbf{L})(M-1)} \implies \lambda_i(\mathbf{W} + \mathbf{W}^T) \geq 0 \quad (49)$$

and so \mathbf{W} is proved to be positive semi-definite.

Sub-proof 3: In [40], for any two $M \times M$ positive semi-definite matrices \mathbf{G} and \mathbf{H} , it was shown that:

$$\lambda_M(\mathbf{G}) \text{tr}(\mathbf{H}) \leq \text{tr}(\mathbf{G} \mathbf{H}) \leq \lambda_1(\mathbf{G}) \text{tr}(\mathbf{H}) \quad (50)$$

where $\lambda_i(\mathbf{G})$ is the i^{th} largest eigenvalue of \mathbf{G} . Using the condition on $\lambda_{max}(\mathbf{\Gamma})$ in (49) and the bound in (50) we get:

$$\begin{aligned} & \frac{1}{M-1} \sum_{i=1}^M \text{Var} \{ \bar{T}_i^w [k] \} = \frac{1}{M-1} \text{tr} \left((\mathbf{W}^k \text{Cov}(\bar{\mathbf{T}}^w [k] | \mathbf{H}_p) (\mathbf{W}^k)^T) \right) \\ & \leq \frac{1}{M-1} \left(\text{Var}_{\max} \text{tr}(\mathbf{W}^k (\mathbf{W}^k)^T) + \epsilon^2 \sigma_{max}^2 \text{tr} \left(\sum_{z=1}^k \mathbf{W}^{z-1} (\mathbf{W}^{z-1})^T \right) \right) \\ & \leq \frac{1}{M-1} \left(\lambda_1(\mathbf{W}) \text{Var}_{\max} \text{tr}(\mathbf{W}^k) + \epsilon^2 \sigma_{max}^2 \lambda_1(\mathbf{W}) \text{tr} \left(\sum_{z=1}^k \mathbf{W}^{z-1} \right) \right) \end{aligned} \quad (51)$$

where $\text{tr}(\cdot)$ denotes the trace operator, $\text{Var}_{\max} = \max(\text{Var} \{ \bar{T}_i^w [k] \}, \dots, \text{Var} \{ \bar{T}_M^w [k] \})$ and $\sigma_{max}^2 = \max(\text{Var} \{ \psi_1 [k] \}, \dots, \text{Var} \{ \psi_M [k] \})$. Now we can show that:

$$\begin{aligned} & \frac{1}{M-1} \left(\lambda_1(\mathbf{W}) \text{Var}_{\max} \text{tr}(\mathbf{W}^k) + \epsilon^2 \sigma_{max}^2 \lambda_1(\mathbf{W}) \text{tr} \left(\sum_{z=1}^k \mathbf{W}^{z-1} \right) \right) \\ & \leq \text{Var}_{\max} \left(\frac{1}{M-1} + \lambda_2^k(\mathbf{W}) \right) + \epsilon^2 \sigma_{max}^2 \left(\frac{k}{M-1} + \frac{1 - \lambda_2^k(\mathbf{W})}{1 - \lambda_2(\mathbf{W})} \right) \end{aligned} \quad (52)$$

where $\lambda_i(\mathbf{W})$, for $i = 1, 2, \dots, M$ are the eigenvalues of \mathbf{W} satisfying $\lambda_M \leq \lambda_{M-1} \leq \dots < \lambda_1 = 1$ and we have used

$$\text{tr}(\mathbf{W}) = \sum_{i=1}^M \lambda_i(\mathbf{W}) \text{ and}$$

$$\sum_{z=1}^k \lambda_i^z(\mathbf{W}) = \begin{cases} \frac{\lambda_i(\mathbf{W}) - \lambda_i^{k+1}(\mathbf{W})}{1 - \lambda_i(\mathbf{W})}, & \text{for } i = 2, 3, \dots, M \\ k, & \text{for } i = 1. \end{cases} \quad (53)$$

This concludes the proof.

REFERENCES

- [1] J. N. Tsitsiklis, "Decentralized detection," *In Advances in Statistical Signal Processing: vol. 2 - Signal Detection*, H. V. Poor, and John B. Thomas, eds., JAI Press, Greenwich, CT, pp. 297-344, Nov. 1993.
- [2] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, pp. 1018-1027, Mar. 2006.
- [3] J. F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1007-1015, Aug. 2004.
- [4] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131-1143, 2006.
- [5] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413-422, Feb. 2006.
- [6] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Distributed Detection and Estimation in Wireless Sensor Networks," In Rama Chellappa and Sergios Theodoridis eds., Academic Press Library in Signal Processing, vol. 2, Communications and Radar Sig. Process., pp. 329-408, 2014.
- [7] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics in Signal Processing*, vol. 2, no. 1, pp. 28-40, Feb. 2008.
- [8] J. Li and G. Alregib, "Rate-constrained distributed estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, pp. 1634-1643, May 2007.
- [9] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124-4140, Sep. 2008.
- [10] E. Nurellari, D. McLernon, M. Ghogho, and S. Aldalameh, "Optimal quantization and power allocation for energy-based distributed sensor detection," *Proc. EUSIPCO*, Lisbon, Portugal, 1-5 Sep. 2014.
- [11] E. Nurellari, S. Aldalameh, M. Ghogho, and D. McLernon, "Quantized Fusion Rules for Energy-Based Distributed Detection in Wireless Sensor Networks," *Proc. SSPD*, Edinburgh, Scotland, 8-9 Sep. 2014.
- [12] J. J. Xiao and Z. Q. Luo, "Universal Decentralized Detection in a Bandwidth-Constrained Sensor Network," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2617-2624, Aug. 2005.
- [13] S. Kar and P. K. Varshney, "A decentralized framework for linear coherent estimation with spatial collaboration," *Proc. ICASSP*, Florence, Italy, 4-9 May. 2014.
- [14] M. Fanaei, M. C. Valenti, A. Jamalipour, and N. A. Schmid, "Optimal power allocation for distributed blue estimation with linear spatial collaboration," *Proc. ICASSP*, Florence, Italy, 4-9 May 2014.
- [15] S. Kar and P. K. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Trans. Information Theory*, vol. 59, no. 6, pp. 3532-3553, 2013.
- [16] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," *Proc. ICASSP*, Salt Lake City, UT, vol. 4, pp. 2033-2036, May 2001.
- [17] F. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035-1048, Mar. 2010.
- [18] F. Cattivelli and A. H. Sayed, "Distributed Detection Over Adaptive Networks Using Diffusion Adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1917-1932, May 2011.
- [19] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 8, no. 6, pp. 2365-2381, Jun. 2009.
- [20] S. Barbarossa and G. Scutari, "Bio-Inspired Sensor Network Design: Distributed decisions through self-synchronization," *IEEE Signal Processing Magazine*, pp. 26-35, May 2007.
- [21] S. Barbarossa and G. Scutari, "Distributed Decision Through Self-Synchronization Sensor Networks in the Presence of Propagation Delays and Asymmetric Channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1667-1684, Apr. 2008.
- [22] A. Bertrand and M. Moonen, Distributed computation of the Fiedler vector with application to topology inference in ad hoc networks," *Signal Process.*, vol. 93, no. 5, pp. 1106-1117, May 2013.
- [23] A. Bertrand and M. Moonen, Topology-aware distributed adaption of laplacian weights for in-network averaging," in *Proc. EUSIPCO*, Marrakech, Morocco, 9-13 Sep. 2013.
- [24] S. Kar and J. M. F. Moura, "Topology for Distributed Inference on Graphs," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2609-2613, Jun. 2008.
- [25] W. Zhang, Z. Wang, Y. Guo, H. Liu, Y. Chen, and J. Mitola, "Distributed cooperative spectrum sensing based on weighted average consensus," *Proc. GLOBECOM*, Houston, Texas, USA, 5-9 Dec. 2011.
- [26] L. Xiao and S. Boyd, "Fast linear iteration for distributed averaging," *Sys. Contr. Lett.*, vol. 53, pp. 65-78, 2004.
- [27] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic optimality of running consensus in testing binary hypotheses," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 814-825, Feb. 2010.
- [28] R. O. Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. of the IEEE*, 95(1), pp. 215-233, Jan. 2007.
- [29] T. Aysal, M. Coates, and M. Rabbat, "Distributed Average Consensus with Dithering Quantization," *IEEE Trans. Automatic Control*, vol. 56, no. 10, Oct. 2008.
- [30] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: quantized data and random link failures," *IEEE Trans. Automatic Control*, vol. 58, no. 3, pp. 1383-1400, Mar. 2010.
- [31] D. Thanou, E. Kokiopoulou, Y. Pu, and P. Frossard, "Distributed average consensus with quantization re-finement," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 194-205, 2013.
- [32] E. Nurellari, D. McLernon, and M. Ghogho "Distributed detection in practical wireless sensor networks via a two step consensus algorithm," *accepted for publication in Proc. ISP*, London, United Kingdom, 1-2 Dec. 2015.
- [33] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523-531, Apr. 1967.
- [34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Englewood Cliffs, NJ: Prentice-Hall PTR, 1993.
- [35] R. Horn and C. R. Johnson, "Matrix Analysis," Cambridge University Press, 1985.
- [36] S. Sardellitti, S. Barbarossa, and A. Swami, "Optimal Topology Control and Power Allocation for Minimum Energy Consumption in Consensus Networks," *IEEE Trans. Signal Process.*, vol. 60, no. 1, Jan. 2012.
- [37] C. Asensio-Marco and B. Beferull-Lozano, "Network topology optimization for accelerating consensus algorithms under power constraints," *Proc. DCOSS*, 16-18 May 2012.
- [38] S. Zheng, X. Yang, and C. Lou, "Distributed consensus algorithms for decision fusion based cooperative spectrum sensing in cognitive radio," *Communications and Information Technologies (ISCIT)*, Hangzhou, China, 12-14 Oct. 2011.
- [39] C.R. Johnson, *Positive definite matrices*, American Mathematical Monthly, vol. 77, issue 3, pp. 259-264, Mar. 1970.
- [40] D. L. Kleinman and M. Athans, "The design of suboptimal linear time-varying systems," *IEEE Trans. Automatic Control*, vol. AC-13, pp. 150-159, Apr. 1968.

PLACE
PHOTO
HERE

Edmond Nurellari received the Diploma degree and the M.Sc. degree in Electrical and Electronic Engineering from Eastern Mediterranean University, Northern Cyprus, in 2010 and in 2012 respectively.

From September 2010 until February 2013, he served as a Research and Teaching Assistant in the department of Electrical and Electronic Engineering at Eastern Mediterranean University. In 2013, he was awarded the Leeds International Research Scholarship (LIRS) to pursue his Ph. D. at the School of Electronics and Electrical Engineering, University of

Leeds, U.K.

His research interests includes distributed signal processing, signal processing on graphs, resource allocations and distributed decisions in wireless sensor networks.

PLACE
PHOTO
HERE

Mounir Ghogho (SM'96) received the M.S. degree in 1993 and the Ph.D. degree in 1997 from the National Polytechnic Institute of Toulouse, France. He was an EPSRC Research Fellow with the University of Strathclyde, Glasgow (Scotland), from September 1997 to November 2001.

Since December 2001, he has been a faculty member with the school of Electronic and Electrical Engineering at the University of Leeds (UK), where he currently holds a Chair in Signal Processing and Communications. He is also a Professor and

Research Director at the International University of Rabat. He is currently an Associate Editor of the Signal Processing magazine. He served as an Associate Editor of the IEEE Transactions on Signal Processing from 2005 to 2008, the IEEE Signal Processing Letters from 2001 to 2004, and the Elsevier Digital Signal Processing journal from 2011 to 2012. He is currently a member of the IEEE Signal Processing Society SAM Technical Committee. He served as a member of the IEEE Signal Processing Society SPCOM Technical Committee from 2005 to 2010 and a member of IEEE Signal Processing Society SPTM Technical Committee from 2006 to 2011. He was the General Chair of the eleventh IEEE workshop on Signal Processing for Advanced Wireless Communications (SPAWC2010), General Chair of the 21st edition of the European Signal Processing Conference (EUSIPCO 2013), the Technical co-Chair of the MIMO symposium of IWCMC 2007 and IWCMC 2008, and a Technical Area co-Chair of EUSIPCO 2008, EUSIPCO 2009 and ISCCSP05.

His research interests are in signal processing and communication networks. He has published 3 book chapters, over 70 journal papers and 140 conference papers. He has supervised more than 25 PhD students. He held invited scientist/professor positions at many institutions including the US Army Research Lab (USA), Tlcom Paris-Tech (France), National Institute of Informatics (Japan), the University Carlos Third of Madrid (Spain), ENSICA (France), Technical University of Darmstadt (Germany), the University of Minnesota (USA), Beijing University of Posts and Telecommunication (China), and the University Mohamed V (Morocco). He is the EURASIP Liaison in Morocco. He was awarded the five-year UK Royal Academy of Engineering Research Fellowship in September 2000. He is also one of the recipients of the 2013 IBM Faculty award.

PLACE
PHOTO
HERE

Des McLernon (M'94) received his B.Sc in Electronic and Electrical Engineering and his M.Sc. in Electronics, both from Queens University of Belfast, N. Ireland.

He then worked on radar research and development with Ferranti Ltd. in Edinburgh, Scotland, and later joined the Imperial College, University of London, where he received his Ph.D. in signal processing. After first lecturing at South Bank University, London, UK, he moved to the School of Electronic and Electrical Engineering at the Univer-

sity of Leeds, UK, where he is a Reader in Signal Processing and Director of Graduate Studies.

His research interests are broadly within the domain of signal processing for communications, in which area he has published over 280 journal and conference papers.