Data Fusion in the Air With Non-Identical Wireless Sensors

Smruti Ranjan Panigrahi, *Student Member, IEEE*, Niclas Björsell, *Senior Member, IEEE*, and Mats Bengtsson, *Senior Member, IEEE*

***Abstract*—In this paper, a multi-hypothesis distributed detection technique with non-identical local detectors is investigated. Here, for a global event, some of the sensors/detectors can observe the whole set of hypotheses, whereas the remaining sensors can either see only some aspects of the global event or infer more than one hypothesis as a single hypothesis. Another possible option is that different sensors provide complementary information. The local decisions are sent over a multiple access radio channel so that the data fusion is formed in the air before reaching the decision fusion center (DFC). An optimal energy fusion rule is formulated by considering the radio channel effects and the reliability of the sensors together, and a closed-form solution is derived. A receive beamforming algorithm, based on a modification of Lozano's algorithm, is proposed to equalize the channel gains from different sensors. Sensors with limited detection capabilities are found to boost the overall system performance when they are used along with fully capable sensors. The additional transmit power used by these sensors is compensated by the designed fusion rule and the antenna array gain. Additionally, the DFC, equipped with a large antenna array, can reduce the overall transmit energy consumption without sacrificing the detection performance.**

***Index Terms*—Wireless Sensor Network, Multiple hypotheses, Non-identical local detector, MAC, Data Fusion in the air, Optimal power fusion rule, Large antenna array.**

I. Introduction

A. Background

DECENTRALIZED distributed detection has been an active research topic since the seminal work by Tenney and Sandell [1]. In those days, to design an optimal fusion rule to detect the event accurately was the primary objective [2]. Moreover, the local observations from the different sensors were typically sent to the DFC through parallel access channels (PAC). Later, Duman and Salehi proposed a decentralized detection technique over multiple access channels (MAC) [3]. In the case of MAC distributed detection, the DFC observes a superposition of the signals, sent from the individual sensors. Distributed detection techniques in wireless sensor network (WSN) has been an active research topic over the last few years [4]. Analysis and design of fusion rules for both the PAC [5], [6] and the MAC [7]–[10] based schemes have been investigated since then. In these research papers, it is commonly assumed that the DFC is equipped with a single antenna.

The received data at the DFC is mostly distorted due to the presence of noise, co-channel interference, and fading in the radio channel. In order to reduce this distortion and subsequently to improve the probability of detection at the DFC, various diversity techniques (e.g., time, frequency, or space) can be employed. A possible solution is to employ a large antenna array at the DFC, which can increase both the spectral and the energy efficiency. This set up of spatially distributed wireless sensors and the DFC resembles a virtual multiple input multiple output (MIMO) system. Distributed detection over a MIMO channel in WSN was first introduced in [11]. After widespread popularity of massive MIMO, MAC distributed detection techniques with a large number of antennas at the DFC have also been studied in several recent publications [12]–[15].

B. Our Contribution

In a majority of the existing literature, the sensors or the local detectors are assumed to be identical with regard to the number of hypotheses observed by them. In other words, the distributed detection problem is formulated as either a binary or an M-ary hypotheses test, where the 'M' is equal for all the local detectors. The case of non-identical sensors is explored only in few publications, like [14]–[19] where, a binary hypotheses testing scenario is investigated with different probability of detections and false alarms among the sensors. However, to the best of our knowledge, we have not found any research, where the network consists of sensors with miscellaneous detection capabilities. For example, for a certain observation or a global event, some of the sensors in the network choose between two hypotheses, whereas the remaining sensors can see multiple hypotheses. To present this idea in a simple way, let us assume a scenario, where a robot is cooking. It will put different ingredients into a pan when the temperature reaches a certain level, i.e., 25, 50, 75, 100, 125, 150, 175, and 200 degrees centigrade. Here, the heating element has a very sophisticated sensor, which can detect whether the temperature has reached 25, 50, 75, 100, 125, 150, 175 or 200 degrees centigrade. However, the sensor in the pan can only detect 50, 100, 150 or 200 degrees

Manuscript received August 20, 2018; revised March 1, 2019; accepted July 1, 2019. Date of publication July 10, 2019; date of current version October 22, 2019. This work was supported in part by the European Commission within the European Regional Development Fund, through the Swedish Agency for Economic and Regional Growth, and in part by Region Gävleborg. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Cedric Richard. *(Corresponding author: Smruti Ranjan Panigrahi.)*

S. R. Panigrahi and N. Björsell are with the Department of Electronics, Mathematics and Natural Sciences, University of Gävle, 801 76 Gävle, Sweden (e-mail: smruti.panigrahi@hig.se; niclas.bjorsell@hig.se).

M. Bengtsson is with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: mats.bengtsson@ee.kth.se).

Digital Object Identifier 10.1109/TSIPN.2019.2928175

centigrade. Moreover, the sensor in the spoon can only detect the temperature above or below 100 degrees centigrade. The robot formulates the hypotheses based on the information available from all these sensors. Let us consider another possible use case from an industrial automation. To detect wear and tear of a ball bearing in a motor, the operator may exploit different sensors that would detect the speed (motion sensor), the temperature, and the high frequency sound waves (ultrasonic sensor). The more sophisticated sensors are able to sense both speed, temperature and sound, whereas other low-cost sensors can only detect the speed and the sound waves. Let us further assume; there are also few standalone temperature sensors in the environment. A central server or a programmable logic controller (PLC) analyzes the information from all these sensors and formulates the hypotheses accordingly. In this research work, we have investigated distributed detection with non-identical wireless sensors to handle the above types of scenarios.

As the set up between the sensors and the DFC resembles a virtual MIMO system, the local decisions can be sent to the DFC using either a spatial multiplexing or a spatial diversity based technique. In literature [11]–[15], spatial multiplexing based techniques are mostly employed so that all the local decisions are available at the DFC. Then, the DFC formulates an optimal fusion rule. The computational complexity of the fusion rule based on any spatial multiplexing scheme increases exponentially as the number of sensors increases. Moreover, it will not be possible to get as many orthogonal spatial radio channels as the number of sensors, when the number of sensors is higher than the number of antennas at the DFC, and this scheme would need more time/frequency resources. On the other hand, this will not be a problem for any technique based on spatial diversity, where the DFC receives a scalar signal, which is a superposition of all the transmitted signals (e.g., the local decisions). This approach is classified as fusion in the air in this paper [20]–[22]. To decode the received signal, a simple detection technique (e.g., maximum likelihood detection) with a much lower computational complexity can be utilized. In this research work, an optimal energy fusion rule based on spatial diversity is proposed such that data fusion will take place in the air (the radio channel), and a closed form solution is derived.

To achieve a high detection performance with a MAC based scheme, the channel state information (CSI) is essential both at the transmitters and the receiver [10], [23] to compensate for the phase of the effective radio channel. Due to multi-path propagation, a large variation in effective path loss between the sensors and the DFC can be observed, where only a few sensors with favorable propagation conditions dominate the detection outcome. To overcome this situation, a suboptimal receiver (Rx) beamformer, based on a modification of Lozano's algorithm [24] is proposed in this paper.

C. Paper Organization

The contents of this manuscript is organized as follows. Section II presents all the theoretical aspects to understand this paper. In Section III, simulation results are presented. Section IV concludes the paper. Proofs and the derivations are contained in a dedicated Appendix.

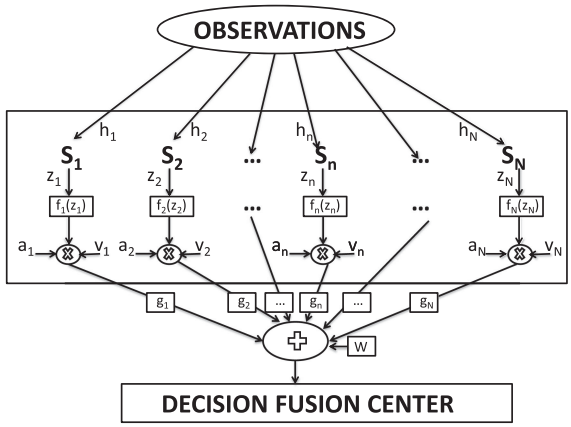

Fig. 1. Fusion in the air distributed detection.

D. Notations

In this paper, $|\bullet|$ represents the absolute value of a complex number. $\bullet^*$ is the conjugate of a complex number, $\odot$ is the Hadamard product, $[\bullet]^*$ is the conjugate transpose of a vector or a matrix and $[\bullet]^T$ is the transpose of a vector or a matrix. $\mathrm{E}[\bullet]$ is the expectation operator. Moreover, throughout this paper, lower case bold letters indicate vectors, whereas uppercase bold letters denote matrices.

II. System Description and Data Fusion Rule

A. The Observations and the Sensors' Decisions

A decentralized wireless sensor system is considered in this paper, as depicted in Fig. 1. The maximum number of global events/hypotheses H, observed by the sensors in the network is M, and the prior probability of the mth hypothesis is, $p_m = \Pr[H = m]$. There are N sensors, and they do not necessarily have identical observational capabilities. Each sensor makes the judgment based on its local detection rule. The number of hypotheses observed by the sensor, n is M_n and generally, $M_n \leq M$. When $M_n = M$, the sensor, n can potentially identify all the global events. However, if $M_n < M$, the sensor, n infers more than one global event as a single hypothesis. The local observations at each sensor are assumed to be conditionally independent of each other and not fixed to any particular distribution. The noisy local observation at the sensor, n is h_n. The local decision rule at the nth sensor is denoted symbolically as

$$z_n = \theta_n(h_n) \in \{1...M_n\}. \tag{1}$$

Conditioned that the global hypothesis is m, the probability that sensor n takes the decision m_n is $p_{n,m_n,m} = \Pr[z_n = m_n | H = m]$. These sensors communicate with the DFC over a coherent multiple wireless access channel. Each sensor is equipped with a single antenna, whereas the DFC has N_D antennas. The DFC needs only the prior probability of the hypotheses, p_m and the conditional detection probabilities, $p_{n,m_n,m}$ from each of the sensors to design an efficient fusion rule. Additionally, it is assumed that the channel information is available both at the DFC and at the sensors, using a suitable piloting scheme.

B. MAC Fusion Model

The decision at each sensor is first mapped to a complex modulation scheme and it is symbolically denoted by,

$$s_n = f_n(z_n). \tag{2}$$

Moreover, $s_n \in \{ s_{n,1},\ s_{n,2},\ \dots\ s_{n,M_n} \}$, where $s_{n,m_n} = f_n(z_n = m_n)$. The modulation scheme can be e.g., amplitude modulation (AM), phase shift keying (PSK) or quadrature amplitude modulation (QAM). Based upon each sensor's own detection capability, a well-designed signal constellation is selected to contribute towards the intended modulation point for the global detection at the DFC. Then, the modulated signals at each sensor are amplified (a_n at the sensor, n) and phase shifted (v_n at the sensor, n) to compensate for the phase of the effective channel, as shown in Fig. 1. The observation by the DFC is

$$\mathbf{y} = \mathbf{G}(\mathbf{v} \odot \mathbf{A}\mathbf{s}) + \mathbf{w}, \tag{3}$$

where,

$$\mathbf{v} = [v_1 \cdots v_N]^T, \tag{4a}$$

$$\mathbf{a} = [a_1 \cdots a_N]^T, \tag{4b}$$

$$\mathbf{A} = \mathrm{diag}(\mathbf{a}), \tag{4c}$$

$$\mathbf{s} = [s_1 \cdots s_N]^T, \tag{4d}$$

$$\begin{aligned}\mathbf{G} &= \widetilde{\mathbf{G}}\,\mathbf{D^{1/2}} \\ &= [\mathbf{g_1} \cdots \mathbf{g_N}],\end{aligned} \tag{4e}$$

$$\mathbf{D} = \mathrm{diag}\left([d_1 \cdots d_N]\right). \tag{4f}$$

$\mathbf{g_n} \in \mathbb{C}^{N_D \times 1}$ is the channel gain between the nth sensor and the DFC. The path loss between the sensors and the DFC is captured in the matrix, $\mathbf{D}$. We do not assume any specific distribution for the elements of the matrix, $\widetilde{\mathbf{G}} \in \mathbb{C}^{N_D \times N}$. It can, for example, be Rayleigh, Rician, or any other relevant distribution. The zero mean circularly symmetric complex Gaussian noise (AWGN), introduced by the channel between the sensors and the DFC is $\mathbf{w} \in \mathbb{C}^{N_D \times 1}$, where the distribution of $\mathbf{w}$ is $\mathcal{CN}(0, \sigma^2_{\mathbf{w}}\mathrm{I})$.

A normalized receiver (Rx) beamformer, $\mathbf{u} \in \mathbb{C}^{N_D \times 1}$ and $\|\mathbf{u}\| = 1$ is implemented at the DFC, with the resulting received signal is

$$\begin{aligned}\tilde{y} &= \mathbf{u}^*\mathbf{y} \\ &= \mathbf{u}^*\left(\mathbf{G}(\mathbf{v} \odot \mathbf{A}\mathbf{s}) + \mathbf{w}\right) \\ &= \left(\sum_{n=1}^{N} \mathbf{u}^*\mathbf{g_n} v_n a_n s_n\right) + \tilde{w}.\end{aligned} \tag{5}$$

The design of Rx beamformer is discussed in Section II.F. Once the Rx beamformer is determined, the phase shift factor at the transmitter (Tx), $\mathbf{v}$ can be calculated as

$$v_n = e^{-j\angle\{\mathbf{u}^*\mathbf{g_n}\}} = \frac{(\mathbf{u}^*\mathbf{g_n})^*}{|\mathbf{u}^*\mathbf{g_n}|}. \tag{6}$$

The conditional probability distribution (PDF) of $\tilde{y}$ given $H = m$, $\Pr[\tilde{y}\,|H = m]$ is a Gaussian mixture. As shown in Appendix A, the conditional mean ($\mu_{\tilde{y}|H=m}$) and variance ($\sigma^2_{\tilde{y}|H=m}$) are

$$\mu_{\tilde{y}|H=m} = \mathrm{E}[\tilde{y}\,|H = m] = \mathbf{a^T}\boldsymbol{\psi}_\mathbf{m}, \tag{7}$$

and

$$\sigma^2_{\tilde{y}|H=m} = \mathrm{var}[\tilde{y}\,|H = m] = \sigma^2_{\mathbf{W}} + \mathbf{a^T}\boldsymbol{\Gamma}_\mathbf{m}\mathbf{a}, \tag{8}$$

where the nth element of the vector $\boldsymbol{\psi}_\mathbf{m} \in \mathbb{C}^{N\times 1}$ is $\mathbf{u}^*\mathbf{g_n}$ $v_n \sum_{m_n=1}^{M_n} p_{n,m_n m} s_{n,m_n}$, whereas, $\boldsymbol{\Gamma}_\mathbf{m} \in \mathbb{C}^{N\times N}$ is a diagonal matrix, whose nth diagonal entry is $|\mathbf{u}^*\mathbf{g_n}v_n|^2(\sum_{m_n=1}^{M_n} p_{n,m_n,m}|s_{n,m_n}|^2 - |\sum_{m_n=1}^{M_n} p_{n,m_n,m}s_{n,m_n}|^2)$.

C. Signal Decoding at the DFC

The DFC detects the global hypothesis based on a maximum a posteriori probability (MAP) criterion,

$$\begin{aligned}\hat{y} &= \arg\max_{\forall m} \Pr[H = m\,|\tilde{y}] \\ &= \arg\max_{\forall m} p_m \Pr[\tilde{y}\,|H = m] \\ &= \arg\max_{\forall m} \left(p_m \sum_{\mathbf{s}} \Pr[\tilde{y}\,|\mathbf{s}] \Pr[\mathbf{s}\,|H = m]\right).\end{aligned} \tag{9}$$

The computational effort of (9) grows with N and M, and there are $\prod_{n=1}^{N} M_n$ permutations possible during the summation for each hypothesis. The PDF of $\Pr[\tilde{y}\,|H = m]$ is also not Gaussian. However, with an asymptotically large number of sensors, the variance, $\sigma^2_{\tilde{y}|H=m}$ approaches $\sigma^2_{\mathbf{W}}$. Therefore, the PDF of $\Pr[\tilde{y}\,|H = m]$ can be approximated as Gaussian with the same mean and variance, e.g., $\mathcal{CN}(\mu_{\tilde{y}|H=m}, \sigma^2_{\tilde{y}|H=m})$. Then, (9) can be further simplified to

$$\hat{y} = \arg\max_{\forall m} \left(\Delta_m - \frac{\left|\tilde{y} - \mu_{\tilde{y}|H=m}\right|^2}{\sigma^2_{\tilde{y}|H=m}}\right), \tag{10}$$

where

$$\Delta_m = \log_e p_m - 0.5\log_e \sigma^2_{\tilde{y}|H=m}. \tag{11}$$

D. Optimal Energy Decision Fusion Rule

The DFC allocates the transmit power to each sensor based on its reliability and the radio channel conditions. The fusion rule is formulated to minimize the overall transmit power while maintaining the average signal to interference plus noise ratio (SINR) at the DFC after beamforming, above a certain threshold, γ. Here, the average SINR is defined as the average power of the desired signal component over the average interference (local mis-detections) plus noise. This gives an optimal energy fusion rule,

$$\begin{aligned}&\min \|\mathbf{a}\| \\ &\text{subject to } \frac{\sum_{m=1}^{M} p_m \left|\mathbf{a}^T\boldsymbol{\psi}_\mathbf{m}\right|^2}{\sigma^2_{\mathbf{W}} + \sum_{m=1}^{M} p_m \mathbf{a^T}\boldsymbol{\Gamma}_\mathbf{m}\mathbf{a}} \geq \gamma.\end{aligned} \tag{12}$$

The above optimization problem can be solved using a convex optimization software such as CVX after a semidefinite

relaxation.

$$\begin{aligned} &\min\ \mathrm{Tr}\{\mathbf{A}\} \\ &\text{subject to} \\ &\mathrm{Tr}\{\mathbf{A}\Psi\} \geq \gamma\left(\mathrm{Tr}\left\{\mathbf{A}\left(\sum_{m=1}^{M} p_m \mathbf{\Gamma_m}\right)\right\} + \sigma_{\mathbf{W}}^2\right) \\ &\mathbf{A} \succeq 0, \end{aligned} \tag{13}$$

where

$$\begin{aligned} &\mathbf{A} = \mathbf{a}\mathbf{a}^T, \\ &\text{and } \mathbf{\Psi} = \sum_{m=1}^{M} p_m \boldsymbol{\psi}_{\mathbf{m}} \boldsymbol{\psi}_{\mathbf{m}}^*. \end{aligned} \tag{14}$$

In order to find a closed-form solution with lower computational complexity, we now show how to approximate $\mathbf{\Psi}$ by a rank-one matrix. As a first step, using Jensen's inequality, the average SINR in (12) can be bounded by,

$$\begin{aligned} \mathrm{SINR}_{\mathrm{avg}} &= \frac{\mathrm{E}\left[\left|\mathbf{a^T}\boldsymbol{\psi}_{\mathbf{m}}\right|^2\right]}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}} \\ &\geq \frac{\left(\mathrm{E}\left[\left|\mathbf{a^T}\boldsymbol{\psi}_{\mathbf{m}}\right|\right]\right)^2}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}} \\ &= \frac{\left(\sum_{m=1}^{M} p_m \left|\mathbf{a^T}\boldsymbol{\psi}_{\mathbf{m}}\right|\right)^2}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}}. \end{aligned} \tag{15}$$

In order to linearize $\left|\mathbf{a^T}\boldsymbol{\psi}_{\mathbf{m}}\right|$, an approximation, $\mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{m}}$ is assumed. Since, a well designed choice of constellation will be assumed at different sensors, the "true" detection should produce a constellation point that points reasonable well in the "right" direction in the complex plane. In other words, the phase of the complex number $\sum_{m_n=1}^{M_n} p_{n,m_n,m} s_{n,m_n}$ should be similar for all the sensors, for a given $H = m$. One way of formulating $\tilde{\boldsymbol{\psi}}_{\mathbf{m}}$ is by defining

$$\mu_{seq|H=m} = \sum_{n=1}^{N} \sum_{m_n=1}^{M_n} p_{n,m_n,m} s_{n,m_n}. \tag{16}$$

This is proportional to what we receive if all effective radio channel gains are equal, for all the sensors. Let ϕ_m be the phase of $\mu_{seq|H=m}$. Then $e^{-j\phi_m}\sum_{m_n=1}^{M_n} p_{n,m_n,m} s_{n,m_n}$ should be approximately aligned with the positive real axis, so that $\left|\mathbf{a^T}\boldsymbol{\psi}_{\mathbf{m}}\right| \approx \mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{m}}$, where

$$\tilde{\boldsymbol{\psi}}_{\mathbf{m}} = \mathrm{Re}\left[e^{-j\phi_m}\boldsymbol{\psi}_{\mathbf{m}}\right] \tag{17}$$

Therefore, (15) can be written as,

$$\begin{aligned} \mathrm{SINR}_{\mathrm{avg}} &\approx \frac{\left(\sum_{m=1}^{M} p_m \mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{m}}\right)^2}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}} \\ &= \frac{\left(\mathbf{a^T}\left(\sum_{m=1}^{M} p_m \tilde{\boldsymbol{\psi}}_{\mathbf{m}}\right)\right)^2}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}}. \end{aligned} \tag{18}$$

The optimal energy fusion rule in (12) is thereby approximated as

$$\begin{aligned} &\min \|\mathbf{a}\| \\ &\text{subject to } \frac{\left(\mathbf{a^T}\left(\sum_{m=1}^{M} p_m \tilde{\boldsymbol{\psi}}_{\mathbf{m}}\right)\right)^2}{\sigma_{\mathbf{W}}^2 + \sum_{m=1}^{M} p_m \mathbf{a^T \Gamma_m a}} \geq \gamma. \end{aligned} \tag{19}$$

E. Derivation of Closed Form Solution From (19)

The optimal fusion rule presented in (19) can be written as a second order cone program (SOCP).

$$\begin{aligned} &\min \|\mathbf{a}\|^2 \\ &\text{subject to } \mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\mathbf{a} \geq \gamma\left(\sigma_{\mathbf{W}}^2 + \mathbf{a^T \Gamma_{av} a}\right), \end{aligned} \tag{20}$$

where

$$\tilde{\boldsymbol{\psi}}_{\mathbf{av}} = \sum_{m=1}^{M} p_m \tilde{\boldsymbol{\psi}}_{\mathbf{m}} \tag{21}$$

and

$$\mathbf{\Gamma_{av}} = \sum_{m=1}^{M} p_m \mathbf{\Gamma_m}. \tag{22}$$

The convex optimization problem in (20) can be solved in closed form. Introducing a Lagrange multiplier λ, the Lagrangian is

$$\begin{aligned} &L(\mathbf{a};\lambda) \\ &= \|\mathbf{a}\|^2 - \lambda\left(\mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\mathbf{a} - \gamma\left(\sigma_{\mathbf{W}}^2 + \mathbf{a^T \Gamma_{av} a}\right)\right) \\ &= \mathbf{a^T}\left(\mathbf{I} + \lambda\gamma\mathbf{\Gamma_{av}} - \lambda\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\right)\mathbf{a} + \lambda\gamma\sigma_{\mathbf{W}}^2. \end{aligned} \tag{23}$$

The dual is given by $\max_{\lambda}\min_{\mathbf{a}} L(\mathbf{a};\lambda)$, which is equivalent to

$$\begin{aligned} &\max \lambda \\ &\text{subject to } \mathbf{I} + \lambda\gamma\mathbf{\Gamma_{av}} - \lambda\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}} \succeq 0. \end{aligned} \tag{24}$$

The solution of the above optimization problem is given by the inverse of the non-negative eigenvalue of $\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}} - \gamma\mathbf{\Gamma_{av}}$. There exists only one non-negative eigenvalue since, $\mathbf{\Gamma_{av}}$ is positive semidefinite. Once λ is calculated, the optimal $\mathbf{a}$ should fulfill

$$\left(\mathbf{I} + \lambda\gamma\mathbf{\Gamma_{av}} - \lambda\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\right)\mathbf{a} = 0, \tag{25}$$

which gives

$$\mathbf{a} = \xi(\mathbf{I} + \lambda\gamma\mathbf{\Gamma_{av}})^{-1}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}, \tag{26}$$

for some constant ξ, which is determined such that the constraint of (20) holds with equality. This gives

$$\xi = \sqrt{\frac{\gamma\sigma_{\mathbf{W}}^2}{\boldsymbol{\vartheta}^{\boldsymbol{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}} - \gamma\mathbf{\Gamma_{av}}\right)\boldsymbol{\vartheta}}}, \tag{27}$$

where,

$$\boldsymbol{\vartheta} = (\mathbf{I} + \lambda\gamma\mathbf{\Gamma_{av}})^{-1}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}. \tag{28}$$

The problem in (20) is only feasible when γ satisfies

$$\gamma < \max_{\mathbf{a}} \frac{\mathbf{a^T}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\mathbf{a}}{\mathbf{a^T \Gamma_{av} a}} = \tilde{\boldsymbol{\psi}}_{\mathbf{av}}^{\mathbf{T}}\mathbf{\Gamma_{av}^{-1}}\tilde{\boldsymbol{\psi}}_{\mathbf{av}} = \gamma_{\max}. \tag{29}$$

The signal to noise ratio (SNR) goes to infinity when $\gamma \to \gamma_{\max}$.

F. Designing Rx Beamformer

In order to give all the sensors a good chance to influence the joint decision, the Rx beamformer, $\mathbf{u}$ should be designed in such a way that all the resulting effective channels from the sensors have similar amplitude. This can be achieved by finding the optimal solution to

$$\begin{aligned} &\max_{\mathbf{u}} \min_{n} |\mathbf{u}^* \mathbf{g_n}| \\ &\text{subject to} \quad \|\mathbf{u}\| \leqslant 1, \end{aligned} \tag{30}$$

which equivalently can be written as

$$\begin{aligned} &\max_{\rho,\mathbf{u}} \rho \\ &\text{subject to} \\ &|\mathbf{u}^* \mathbf{g_n}| \geqslant \rho, \quad \forall n \in \{1...N\}, \\ &\|\mathbf{u}\| \leqslant 1. \end{aligned} \tag{31}$$

This problem is mathematically equivalent to the multicast beamforming problem [25], which is known to be NP-hard, but where high quality approximate solutions can be calculated via convex (semidefinite) approximation. In order to keep the computational complexity low, we instead propose a modified version of Lozano's algorithm [24].

G. Modified Lozano's Algorithm

Lozano's algorithm [24] solves (31) by iteratively taking a projected gradient step to improve the weakest remaining link, after first having discarded a portion of the weakest links. In [25], both strengths and drawbacks of this algorithm are discussed further and some new improvements are also proposed. Here, we propose a related idea, based on the observation that the weakest link gain is shared by most links at the optimum of (31). Therefore, instead of only considering the single weakest link in each iteration, we calculate the gradient of

$$\sum_{n \in \mathcal{I}} |\mathbf{u}^* \mathbf{g}_n|^2 , \tag{32}$$

where the index set $\mathcal{I}$ denotes the weakest links given the current value of $\mathbf{u}$. For simplicity, we use a fixed proportion, ρ, $0 < \rho < 1$, i.e. in each iteration the values of $|\mathbf{u}^* \mathbf{g}_n|$ are sorted and the indices of the $\lfloor \rho N \rfloor$ weakest values are included in $\mathcal{I}$. Similarly to [25], we use the principal singular vector of $\mathbf{G}$ to initialize $\mathbf{u}$ and a diminishing step size η_i. After each gradient update, $\mathbf{u}$ is projected onto the constraint $\|\mathbf{u}\| \leq 1$. The resulting Algorithm 1 does not come with any convergence guarantees, but our numerical experiments show that it is able to bring the effective channel gain $|\mathbf{u}^* \mathbf{g}_n|$ of all the sensors to a similar order, after a small number of iterations. This is good enough for the current application, since the goal is to ensure that the contribution from all the sensors to have a good chance to influence the DFC's decision. The algorithm stops when all $|\mathbf{u}^* \mathbf{g}_n|$ are equal within a 99% margin or the number of iterations exceeds $i_{\max}$. Numerical experiments show good performance with $\rho = 0.5$, step size rule $\eta_i = N^2/(i(\sum_{n=1}^{N} \|\mathbf{g}_n\|)^2)$ and $i_{\max} = 100$.

Algorithm 1: Modified Lozano's Algorithm.

$\mathbf{u} \leftarrow$ dominating singular vector of $\mathbf{G}$. $\triangleright$ Initialization
$i \leftarrow 0$
repeat
  $i \leftarrow i + 1$
  $\tilde{\gamma}_n \leftarrow \mathbf{g}_n^* \mathbf{u}$, $n = 1, \ldots, N$
  $\mathcal{I} \leftarrow$ indices of $\lfloor \rho N \rfloor$ smallest $|\tilde{\gamma}_n|$
  $\mathbf{u} \leftarrow \mathbf{u} + \eta_i \sum_{n \in \mathcal{I}} \tilde{\gamma}_n \mathbf{g}_n$ $\triangleright$ Gradient step
  $\mathbf{u} \leftarrow \frac{\mathbf{u}}{\|\mathbf{u}\|}$ $\triangleright$ Projection step
until $\frac{\min\{\tilde{\gamma}_n\}}{\max\{\tilde{\gamma}_n\}} > 0.99$ or $i > i_{\max}$.

H. Computational Complexity

Primarily, the DFC solves four different problems, e.g., determining Rx beam-former using Algorithm 1, calculating phase shift factor at Tx (6), allocating transmit power to all the sensors (26) and decoding the hypothesis from the received signal, $\tilde{y}$ (10). The computational complexity of Algorithm 1 can be maximum up to $O\left(N i_{\max}\left(N_D + \log N + \rho\right)\right)$ as the dominating singular vector of $\mathbf{G}$ can be determined within $O\left(NN_D\right)$ [26]. The phase shift factor at Tx (6) can be solved within $O\left(NN_D\right)$. In order to solve the transmit power allocation in (26), the set of equations (21), (22), (24), (27), (28), and (29) need to be solved first. Therefore, the computational complexity of transmit power allocation is $O\left(N^2 + NM\right)$ since $\mathbf{\Gamma_{av}}$ is a diagonal matrix. These first three problems are calculated once as long as the radio channel remains constant, e.g., once in a coherence time of the radio channel. In order to decode the received signal by the DFC (10), the set of equations (7), (8), (11) must be solved first. However, these three equations are updated once in every coherence time and can be solved within $O\left(MN\right)$, $O\left(MN\right)$ and $O\left(M\right)$ respectively. One of the inputs to (7) and (8) are $\boldsymbol{\psi}_{\mathbf{m}}$ and $\mathbf{\Gamma_m}$ respectively, and both of them can be determined within $O(N(M + N_D))$. Ideally, they remain constant over a longer duration. Therefore, the majority of the time when the sensors report an observation, the computational complexity of the decoder is $O\left(M\right)$.

III. Simulation and Results

A. Simulation Assumptions

The sensors are assumed to be distributed uniformly on a $10\,\mathrm{m} \times 10\,\mathrm{m}$ two dimensional plane. The DFC is elevated 25 meters above the center of the plane. A different realization of sensor positions was drawn for each Monte Carlo iteration, only in case of the first two sub-scenarios of scenario 1. The path loss matrix, $\mathbf{D}$ is calculated based on Friis free space path-loss [27] formula at 2.4 GHz frequency, with path loss exponent 2. The receiver and the transmitter antenna gain is 1. In all the simulated scenarios, the local observation at each sensor is assumed to be corrupted by AWGN. The observation noise power at the sensors is uniformly distributed in the interval $[0.1, 0.25]$, and a single realization of the noise power is used for all the Monte Carlo iterations of a particular scenario. However, a different realization of the noise power was drawn for each scenario.

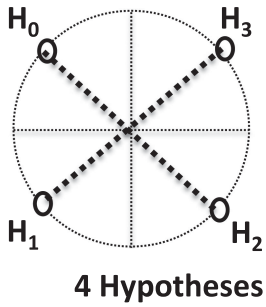

Fig. 2. Mapping of Hypotheses to QPSK signal constellations for scenario 1.

At each sensor, the local decisions are decided by a maximum likelihood detector. The variance of AWGN at the DFC, $\sigma_{\mathbf{w}}^2$ is assumed as 1. For each scenario, the number of channel realizations is 100 and the number of Monte Carlo iterations for each channel realization is 10^4.

The use of sensors is prevalent in industries. Moreover, the radio propagation characteristics in an industrial environment are much different from any home or office environments [28]. The radio channel model in these environments broadly follows Rician distribution. Therefore, we evaluate our proposed fusion rule in both Rayleigh and Rician fading channel. Figs. 3, 5, 7, 8, 12, 13, 16 and 17 capture the simulation results for a Rayleigh fading channel, i.e., the distribution of each elements in the matrix, $\widetilde{\mathbf{G}}$ is considered as $\mathcal{CN}(0,1)$. Whereas Figs. 4, 6, 9, 10, 14 and 18 capture the simulation results of a Rician channel. During the simulations, Rician K factor was $K = 12$ [28]. In all these plots, the legend "PAC" represents the simulation results of the optimal parallel fusion rule [29] in an ideal radio channel without any noise, where all the original local decisions from the sensors are available at the DFC to formulate the fusion rule. Therefore, the detection performance of the optimal "PAC" fusion rule is fairly comparable to any MAC scheme based on MIMO spatial multiplexing in a noiseless radio channel. Moreover, in these plots, the SNR at the DFC's antennas is captured before the receiver beam-forming.

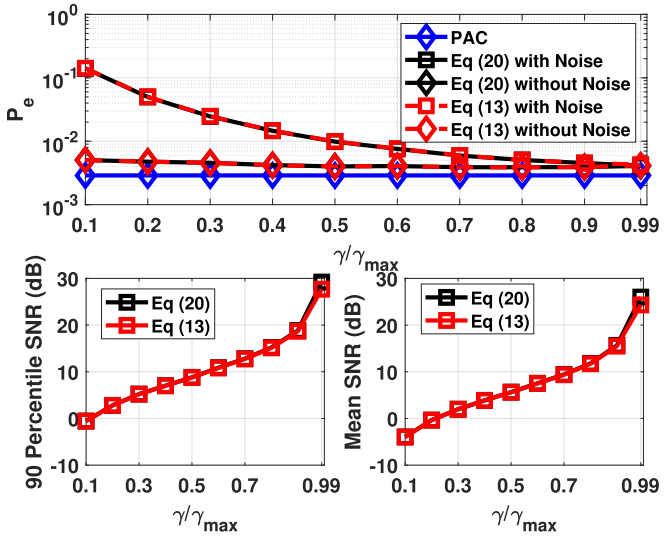

Fig. 3. Scenario-1 SubScenario-1 in a Rayleigh channel. Both, the number of sensors and the number of DFC's antennas are 8.

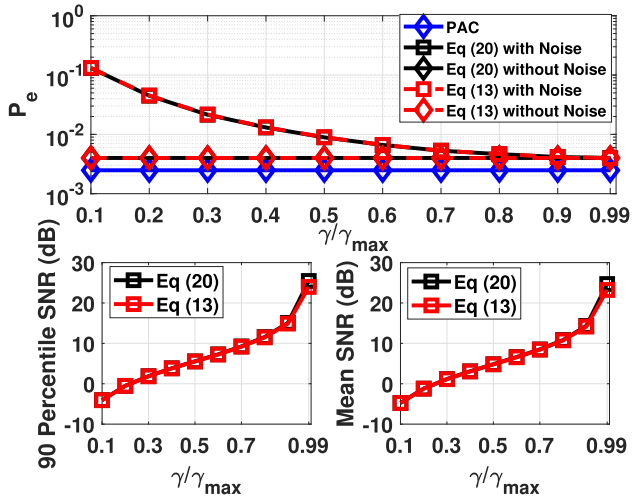

Fig. 4. Scenario-1 SubScenario-1 in a Rician channel. Both, the number of sensors and the number of DFC's antennas are 8.

B. Scenario-1: Performance With Identical Sensors

In this scenario, all the sensors are assumed to have similar detection capabilities. They can observe and decide between four different real numbers, e.g., $\{0, 1, 2, 3\}$. These numbers are defined as four different hypotheses, e.g., H_0, H_1, H_2, and H_3 respectively. The prior probabilities of these hypotheses are randomly assigned. The local decisions are mapped to quadrature PSK (Q-PSK) modulation, as depicted in Fig. 2. Figs. 3 to 10 capture the different simulation results of this scenario. In all these eight figures, the top plot captures the performance of the DFC in terms of detection error probability (P_e).

Here, the performance is evaluated in three different sub-scenarios, i.e., 1) the impact of γ with a fixed number of sensors and DFC's antennas, 2) the impact of the number of sensors with a fixed number of antennas at the DFC, and 3) the impact of the number of antennas at the DFC with a fixed number of sensors.

1) The Impact of γ: Figs. 3 and 4 capture the results of the impact of different values of γ when the number of sensors and the number of DFC's antennas both is 8. A convex solver is used to solve the optimization problem in (13) whereas our proposed formulation in (20) has a closed form solution. The "with Noise" and "without Noise" legend represent a noisy and a noiseless radio channel respectively to give an idea about the amount of error introduced by AWGN. Looking at the top plot in both the figures, we can conclude that (13) and (20) have similar performance for both the detection error probability and the SNR. This shows that the considered approximations to deduce (20) from (13) do not yield any performance degradations. The detection error probability of our proposed method is roughly around 1.5 times higher than "PAC" at a higher γ. Moreover, at a higher γ, the error introduces by the radio channel is almost zero. As γ increases gradually, the SNR increases at the receiver, and the transmit power consumption increases. For further simulations, we only consider the fusion rule, formulated in (20).

2) Impact of the Number of DFC's Antennas When the Number of Sensors Is 8: Figs. 5 and 6 capture the results of the impact of the number of antennas at the DFC. During the

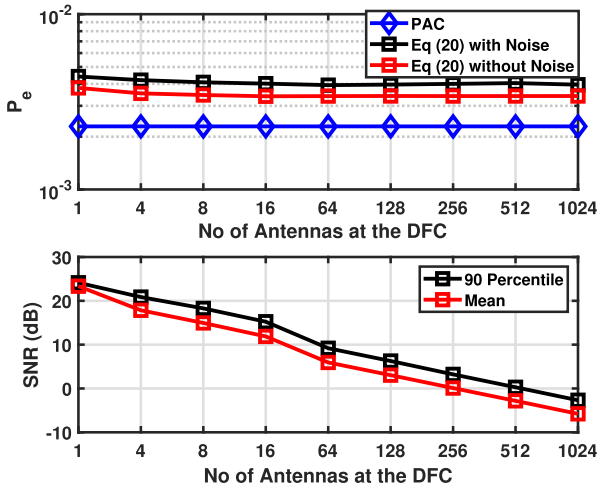

Fig. 5. Scenario-1: Subscenario-2 in a Rayleigh channel. The number of Sensors is 8 and $\gamma = 0.9\gamma_{max}$.

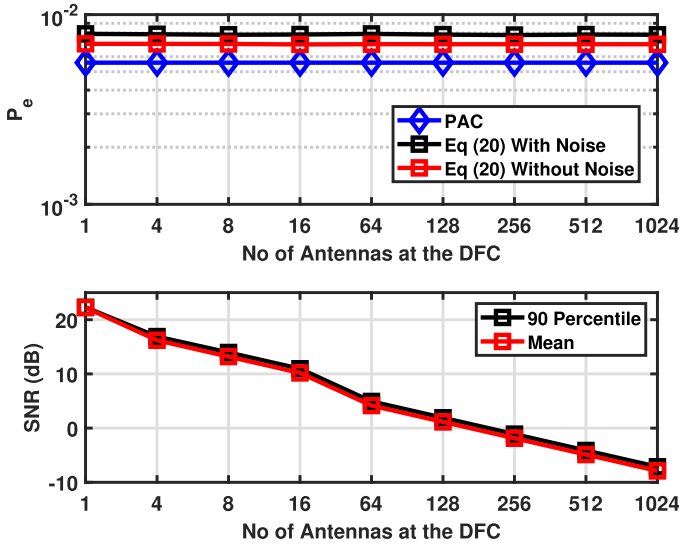

Fig. 6. Scenario-1: Subscenario-2 in a Rician channel. The number of Sensors is 8 and $\gamma = 0.9\gamma_{max}$.

simulation $\gamma = 0.9\gamma_{max}$ and the number of sensors is 8. As mentioned earlier, the "with Noise" and "without Noise" legend represent a noisy and a noiseless radio channel respectively. The proposed MAC fusion rule has slightly worse performance than the optimal parallel fusion rule. Moreover, the larger antenna configuration provides a similar detection performance at a lower SNR due to the proposed optimal energy fusion rule. The lower SNR is a consequential effect on the lower transmit power.

3) Impact of the Number of Sensors With 8 Antennas at the DFC: Figs. 7 to 10 capture the results of the impact of the number of sensors on the performance. Looking at Figs. 7 and 9, γ_{max} increases with the number of sensors as $\mathbf{a}^{\mathrm{T}}\boldsymbol{\Gamma}_{\mathbf{m}}\mathbf{a}$ decreases proportionally. In the top plot, the detection performance of our proposed method is almost identical to the optimal PAC fusion rule with a lesser number of sensors, when $\gamma = 0.9\gamma_{max}$. The PAC fusion rule has a better performance with more number of sensors, but the computational complexity of the PAC fusion

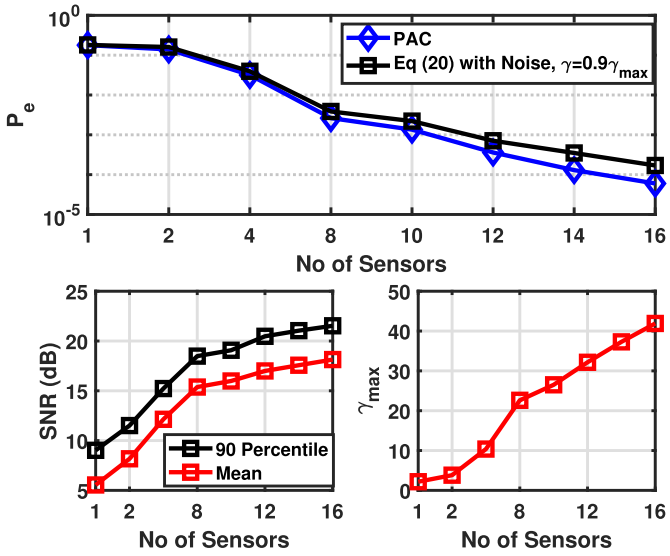

Fig. 7. Scenario-1: Subscenario-3 in a Rayleigh channel. The number of antennas at the DFC is 8.

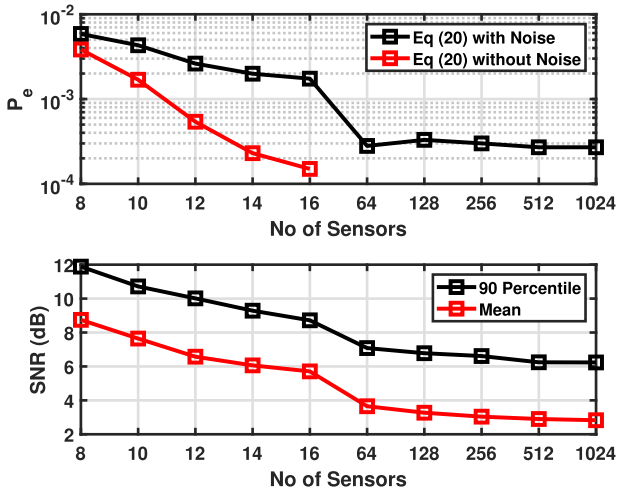

Fig. 8. Scenario-1: Subscenario-3 in a Rayleigh channel. The number of antennas at the DFC is 8 and $\gamma = 15$.

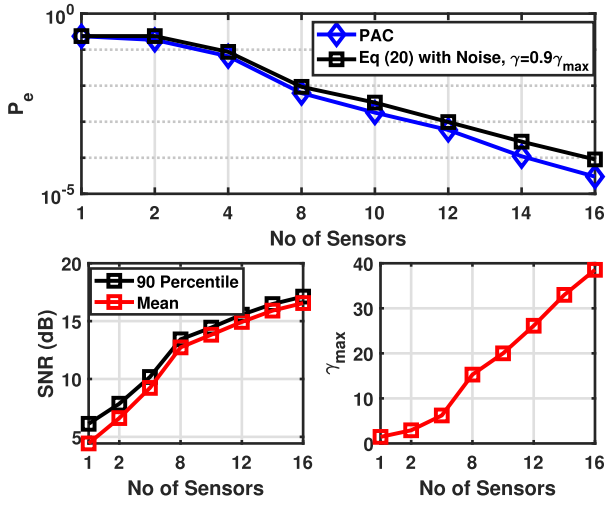

Fig. 9. Scenario-1: Subscenario-3 in a Rician channel. The number of antennas at the DFC is 8.

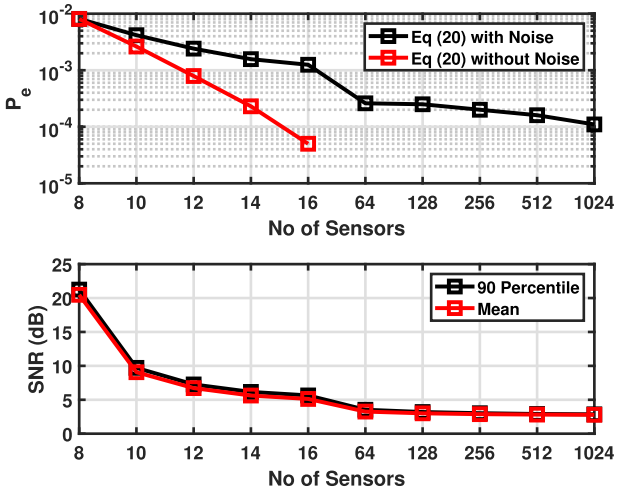

Fig. 10. Scenario-1: Subscenario-3 in a Rician channel. The number of antennas at the DFC is 8 and $\gamma = 15$.

rule increases exponentially with the number of sensors. With an increase in the number of sensors, the SNR increases at the antennas of the DFC due to increase in γ_{max}, as shown in Figs. 7 and 9.

In Figs. 8 and 10, γ is kept fix at 15. When the number of sensors is less, γ_{max} is lower than 15 and it causes an infeasible scenario. This is why the number of sensors more than or equal to 8 was considered during the simulation. As the number of sensors grows, more degrees of freedom is achieved, and the PDF of $\Pr[\tilde{y}\,|H = m]$ becomes Gaussian. Therefore, P_e gradually attains a minimum value, and this error is only due to AWGN at the DFC. When the number of sensors is in between 8 to 64, the PDF of $\Pr[\tilde{y}\,|H = m]$ is a Gaussian mixture. As the receiver is designed assuming $\Pr[\tilde{y}\,|H = m]$ as Gaussian, P_e decreases as the number of sensors increases. The optimal energy fusion rule and the transmit antenna array gain help to reduce the SNR required to achieve the desired detection performance when increasing the number of sensors.

C. Scenario-2: Performance of Non-Identical Sensors Keeping the Number of Sensors Fixed

In this scenario, two different examples are discussed. Figs. 12, 14, 16, and 18 capture the simulation results of this scenario. In all these figures, the top plot compares the detection error probability with respect to the number of antennas at the DFC. The left and right plots below capture the mean of the SNR, and the average of the sum transmitted power, respectively. The legend in the top plot is also applicable to the plots below. Moreover, Figs. 13 and 17 depict the confusion matrices to capture the detection performance in Rayleigh channel, when the number of antennas at the DFC is 8 and $\gamma = 0.9\gamma_{max}$. The confusion matrices for Rician channel are not included in this paper as they give similar information as Rayleigh channel.

1) Sub-Scenario-1: Let us extend the previous scenario, where the sensors in the network are not identical. They are expected to observe and decide between eight different numbers, $\{0, 1, 2, 3, 4, 5, 6, 7\}$. These set of numbers are defined as eight

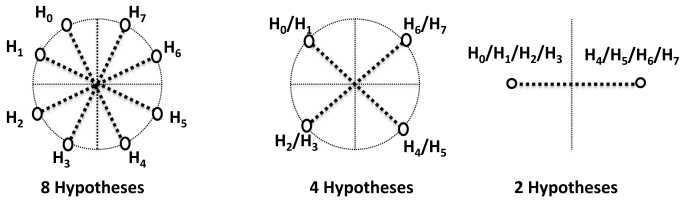

Fig. 11. Mapping of Hypotheses to PSK signal constellation for scenario-2 Sub-scenario-1.

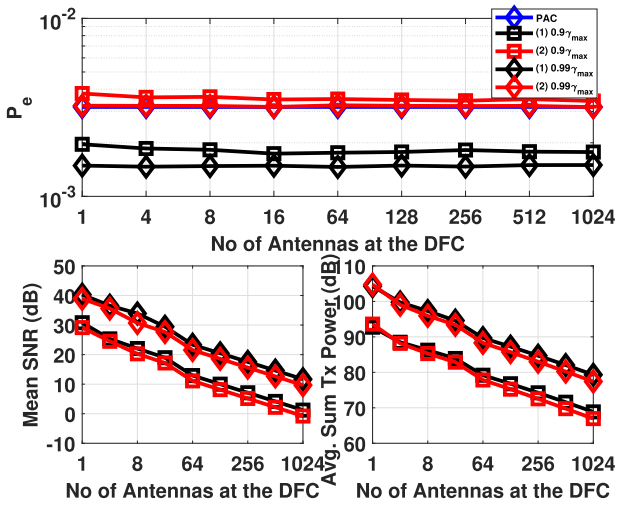

Fig. 12. Scenario-2: Sub scenario-1 in a Rayleigh channel.The legend in the top plot is also applicable to the plots below.

different hypotheses, e.g., $H_0, H_1, \ldots, H_7$ respectively. The prior probabilities of these hypotheses are randomly assigned. Let us assume; there are three different categories of sensors in the network. Category-1 sensors can observe all of the eight hypotheses and apply 8-PSK modulation scheme. Category-2 sensors cannot identify all the hypotheses and they consider groups of $\{H_0, H_1\}$, $\{H_2, H_3\}$, $\{H_4, H_5\}$ and $\{H_6, H_7\}$ as individual hypothesis. Therefore, these sensors can only observe four hypotheses and apply Q-PSK modulation scheme while forwarding the local decisions. Moreover, Category-3 sensors can only observe two hypotheses and consider groups of $\{H_0, H_1, H_2, H_3\}$ and $\{H_4, H_5, H_6, H_7\}$ as individual hypothesis. They apply a binary PSK (BPSK) modulation scheme to forward their local decisions. The mapping of hypotheses to the signal constellation is depicted in Fig. 11. All these three categories of sensors use the same radio access channel to communicate their local observation. During the simulation of this scenario, the number of sensors in the network is kept fixed at sixteen. Out of these 16 sensors, eight sensors are from Category- 1, four sensors are from Category-2, and the remaining four sensors are from Category-3. The results of the legend "PAC", "$(2)0.9\gamma_{max}$" and "$(2)0.99\gamma_{max}$" are generated when Category-1 sensors are only present in the network. The results of the legend "$(1)0.9\gamma_{max}$" and "$(1)0.99\gamma_{max}$" are generated when all three categories of the sensors are present in the network.

Looking at Figs. 12 and 14, the presence of Category-2 and 3 sensors in the network helps to bring the detection error down. Even though these sensors cannot observe the full set

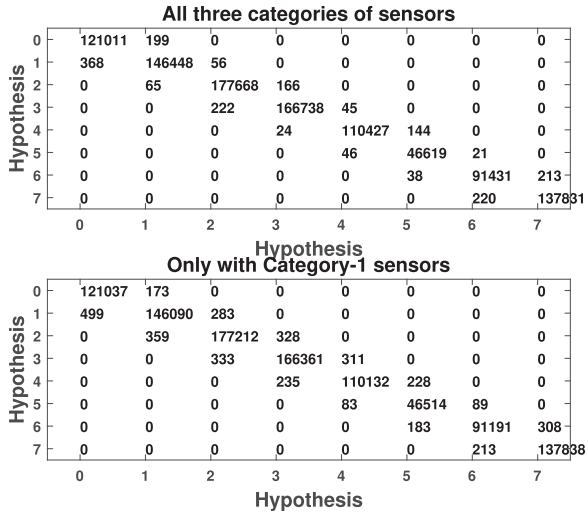

Fig. 13. Scenario-2: Sub scenario-1 in a Rayleigh channel.Confusion Matrix.

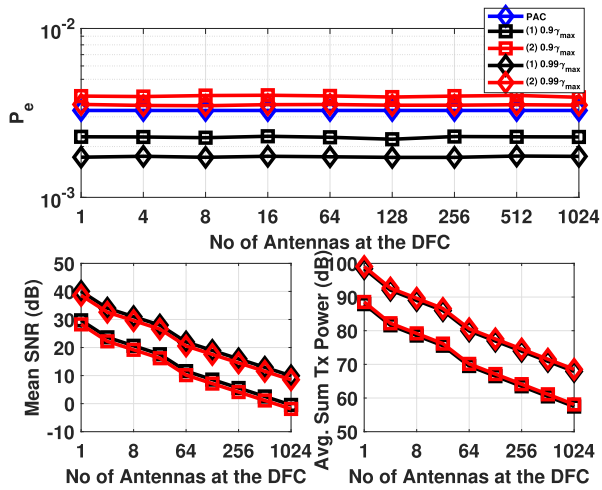

Fig. 14. Scenario-2: Sub scenario-1 in a Rician channel. The legend in the top plot is also applicable to the plots below.

of hypotheses, they have a positive impact on the performance. The transmit antenna array gain is more when Category-2 and 3 sensors are present in the network. Therefore, the performance of the sum transmitted power or the average SNR is almost similar to the scenario, where Category-2 and 3 sensors are present in the network. This signifies the transmit power saving at the individual sensors.

In Fig. 13, the top plot depicts the confusion matrix when all three categories of the sensors are present in the network whereas the below plot captures the confusion matrix when only the sensors of Category-1 are present. When all three categories of sensors are present, the error in choosing another hypothesis other than the true hypothesis has been decreased significantly.

2) Sub-Scenario-2: Let us discuss a different scenario with non-identical sensors, where a global event is defined with the combination of three parameters, i.e., temperature, pressure, and moisture. Moreover, these three parameters are assumed to be independent of each other. If these parameters are not independent of each other, then this scenario will be quite similar to the previous sub-scenario. All these sensors are binary sensors, and they measure the value between 0 and 1. Therefore, there are eight different hypotheses possible. The prior probabilities of these hypotheses are considered to be equal. Category-1 transmitters have all three sensors. Category-2 transmitters have pressure and moisture sensors. Category-3 transmitters have only temperature sensors. Therefore, Category-1 sensors map local decisions from all three sensors to an 8-PSK modulation scheme while sending the local decisions to the DFC. Similarly, Category-2 and 1 sensors map their local decisions to QPSK and BPSK modulation schemes, respectively. The mapping of hypotheses to the signal constellation is depicted in Fig. 15. All these three categories of sensors use the same radio access channel to communicate their local observation. During the simulation of this scenario, the number of sensors in the network is kept fixed at sixteen. Out of these 16 sensors, eight sensors are from Category-1, four sensors are from Category-2, and the remaining four sensors are from Category-3. The results of the legend "$(2)0.9\gamma_{max}$" and "$(2)0.99\gamma_{max}$" are generated when Category-1 sensors are only present in the network. The results of the legend "$(1)0.9\gamma_{max}$"

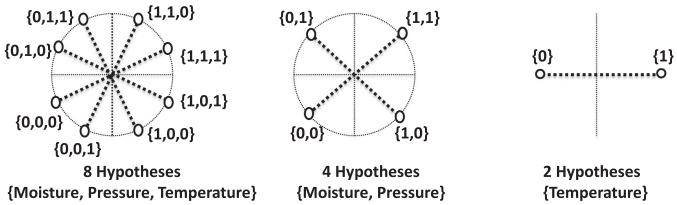

Fig. 15. Mapping of Hypotheses to PSK signal constellation for scenario-2 sub-scenario-2.

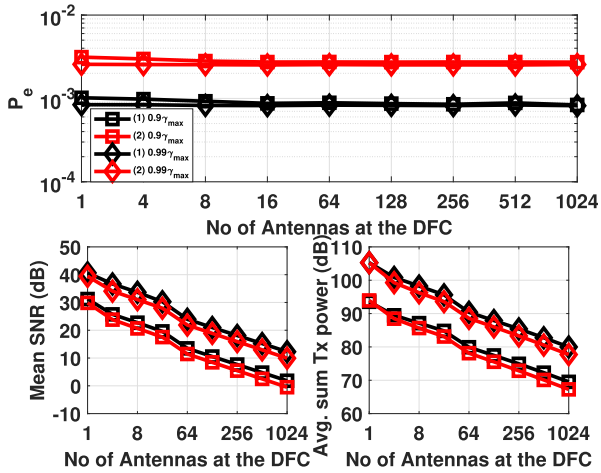

Fig. 16. Scenario-2: Sub scenario-2 in a Rayleigh channel. The legend in the top plot is also applicable to the plots below.

and "$(1)0.99\gamma_{max}$" are generated when all three categories of the sensors are present in the network.

Looking at Figs. 16 and 18, the presence of Category-2 and 3 sensors in the network improves the detection performance, like the previous sub-scenario. Due to the designed optimal fusion rule and the transmit antenna array gain, the sum transmitted power, and the average SNR are almost similar in both the cases, i.e., with and without Category 2 and 3 sensors. This signifies the lower transmit power consumption at the sensors. Like the previous sub-scenario, when all three categories of sensors are present, the error in choosing another hypothesis other than the

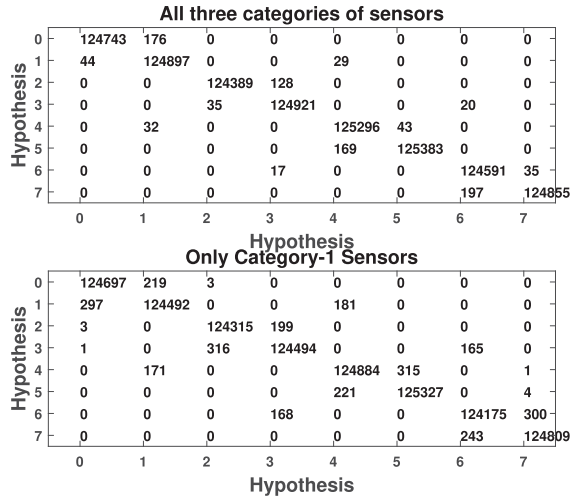

Fig. 17. Scenario-2: Sub scenario-2 in a Rayleigh channel. Confusion Matrix.

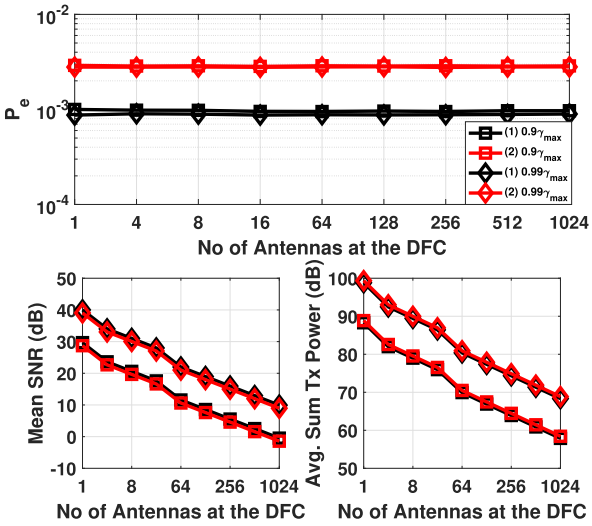

Fig. 18. Scenario-2: Sub scenario-2 in a Rician channel. The legend in the top plot is also applicable to the plots below.

true hypothesis has been decreased significantly, as shown in Fig. 17.

IV. Conclusion

In this paper, a fusion in the air distributed detection technique with non-identical local detectors/sensors is investigated, and an energy efficient fusion rule is proposed. The local decisions are sent to the DFC over a MAC wireless channel. The sensors are equipped with a single antenna whereas the DFC can be equipped with a large antenna array. A modified version of Lozano's algorithm is also proposed to design the receiver beamformer so that all the sensors in the network get a good chance to influence the decision. The detection performance of the fusion in the air technique is fairly comparable with the PAC scheme with a lesser number of sensors. Both schemes utilize similar radio frequency resources as long as the number of sensor is less than equal to the number of antennas at the DFC. On the other hand, if the number of sensors exceeds the number of receive antennas, the PAC scheme would need more time/frequency resources, whereas fusion in the air keeps providing improved performance without using more RF resources. Additionally, there are some sensors in the network, having limited detection capabilities. They cannot observe the whole set of hypotheses; rather, they can only see a few aspects of the global events. However, by considering the local decisions from these sensors at the DFC, the overall system performance can be augmented. These additional sensors do not increase the net transmit power consumption due to the designed optimal energy fusion rule and the antenna array gain. Moreover, with a higher number of antennas at the DFC, the overall transmit energy consumption can be reduced while achieving a similar detection performance.

Appendix
Derivation of Conditional Mean and Variance of $\tilde{y}$

$$\begin{aligned}
\mu_{\tilde{y}|H=m} &= \mathrm{E}\left[\tilde{y}\,|H=m\right] \\
&= \mathrm{E}\left[\sum_{n=1}^{N}\mathbf{u}^{*}g_n v_n a_n s_n + \tilde{w}\,|H=m\right] \\
&= \sum_{n=1}^{N}\mathbf{u}^{*}g_n v_n a_n \mathrm{E}\left[s_n\,|H=m\right] + \mathrm{E}\left[\tilde{w}\,|H=m\right] \qquad \text{(A.1)} \\
&= \sum_{n=1}^{N}\mathbf{u}^{*}g_n v_n a_n \left(\sum_{m_n=1}^{M_n} p_{n,m_n,m}s_{n,m_n}\right) \\
&= \mathbf{a^T}\psi_{\mathbf{m}},
\end{aligned}$$

where, $\mathbf{u}^{*}\mathbf{g_n}v_n\sum_{m_n=1}^{M_n}p_{n,m_n m}s_{n,m_n}$ is the nth element of the vector $\psi_{\mathbf{m}}$.

$$\begin{aligned}
\sigma^2_{\tilde{y}|H=m} &= \mathrm{var}\left[\tilde{y}\,|H=m\right] \\
&= \mathrm{E}\left[\left|\tilde{y}-\mu_{\tilde{y}|H=m}\right|^2|H=m\right] \\
&= \mathrm{E}\left[|\tilde{y}|^2+\left|\mu_{\tilde{y}|H=m}\right|^2-\tilde{y}^{*}\mu_{\tilde{y}|H=m}-\tilde{y}\mu^{*}_{\tilde{y}|H=m}\,|H=m\right] \\
&= \mathrm{E}\left[|\tilde{y}|^2\,|H=m\right]-\left|\mu_{\tilde{y}|H=m}\right|^2 \\
&= \mathrm{E}\left[\left|\left(\sum_{n=1}^{N}\mathbf{u}^{*}\mathbf{g_n}v_n a_n s_n\right)+\tilde{w}\right|^2|H=m\right]-\left|\mu_{\tilde{y}|H=m}\right|^2 \\
&= \mathrm{E}\left[\left|\left(\sum_{n=1}^{N}\mathbf{u}^{*}\mathbf{g_n}v_n a_n s_n\right)\right|^2|H=m\right]+\sigma^2_{\mathbf{W}}-\left|\mu_{\tilde{y}|H=m}\right|^2 \\
&= \sigma^2_{\mathbf{W}}+\sum_{n=1}^{N}\left(\mathbf{u}^{*}\mathbf{g_n}v_n a_n\right)^2 \\
&\quad\times\left(\sum_{m_n=1}^{M_n}p_{n,m_n,m}|s_{n,m_n}|^2-\left|\sum_{m_n=1}^{M_n}p_{n,m_n,m}s_{n,m_n}\right|^2\right) \\
&= \sigma^2_{\mathbf{W}}+\mathbf{a^T}\mathbf{\Gamma_m}\mathbf{a}, \qquad \text{(A.2)}
\end{aligned}$$

where, the $\mathbf{\Gamma_m}$ is a diagonal matrix and the nth diagonal entry of this matrix is $|\mathbf{u}^*\mathbf{g_n}v_n|^2(\sum_{m_n=1}^{M_n}p_{n,m_n,m}|s_{n,m_n}|^2 - |\sum_{m_n=1}^{M_n}p_{n,m_n,m}s_{n,m_n}|^2)$.

REFERENCES

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

[2] P. K. Varshney, *Distributed Detection and Data Fusion*. New York, NY, USA: Springer, 1996.

[3] T. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 2, pp. 469–476, Apr. 1998.

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

[5] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[6] B. Chen, L. Tong, and P. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 16–26, Jul. 2006.

[7] H. S. Kim and N. A. Goodman, "Power control strategy for distributed multiple-hypothesis detection," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3751–3764, Jul. 2010.

[8] C. R. Berger, M. Guerriero, S. Zhou, and P. Willett, "PAC vs MAC for decentralized detection using noncoherent modulation," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3562–3575, Sep. 2009.

[9] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, Mar. 2007.

[10] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.

[11] X. Zhang, H. 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.

[12] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. Lopez-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3005–3017, Jun. 2015.

[13] A. Shirazinia, S. Dey, D. Ciuonzo, and P. Salvo Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2499–2512, May 2016.

[14] D. Ciuonzo, A. Aubry, and V. Carotenuto, "Rician MIMO channel- and jamming-aware decision fusion," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 3866–3880, Aug. 2017.

[15] D. Ciuonzo, P. Salvo Rossi, and S. Dey, "Massive MIMO channel-aware decision fusion," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 604–619, Feb. 2015.

[16] D. Ciuonzo, G. Romano, and P. Salvo Rossi, "Optimality of received energy in decision fusion over Rayleigh fading diversity MAC with non-identical sensors," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 22–27, Jan. 2013.

[17] 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, no. 3, pp. 1018–1027, Mar. 2006.

[18] A. Patel, H. Ram, A. K. Jagannatham, and P. K. Varshney, "Robust cooperative spectrum sensing for MIMO cognitive radio networks under CSI uncertainty," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 18–33, Jan. 2018.

[19] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.

[20] L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Over-the-air computation for IoT networks: Computing multiple functions with antenna arrays," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5296–5306, Dec. 2018.

[21] G. Zhu and K. Huang, "MIMO Over-the-air computation for high-mobility multi-modal sensing," *IEEE Internet Things J.*, vol. 4662, pp. 1–14, Sep. 2018, doi: 10.1109/JIOT.2018.2871070.

[22] M. Goldenbaum, S. Stańczak, and H. Boche, "Interference-aware analog computation over the wireless channel: Fundamentals and strategies," in *Communications in Interference Limited Networks*. Berlin, Germany: Springer, 2016, pp. 97–118.

[23] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Estimation over fading channels with limited feedback using distributed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 414–425, Jan. 2010.

[24] A. Lozano, "Long-term transmit beamforming for wireless multicasting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2007, pp. III–417–III–420.

[25] E. Matskani, N. D. Sidiropoulos, Z. Q. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4882–4894, Dec. 2009.

[26] Y. Chahlaoui, K. Gallivan, and P. V. Dooren, "Recursive calculation of dominant singular subspaces," *SIAM J. Matrix Anal. Appl.*, vol. 25, no. 2, pp. 445–463, 2003.

[27] A. F. Molish, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2012.

[28] E. Tanghe *et al.*, "The industrial indoor channel: Large-scale and temporal fading at 900, 2400, and 5200 MHz," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2740–2751, Jul. 2008.

[29] W. Baek and S. Bommareddy, "Optimal m-ary data fusion with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1150–1152, Jul. 1995.

Smruti Ranjan Panigrahi received the B.Tech. degree in computer science engineering from the National Institute of Technology, Rourkela, India, in 2005 and the MSc. degree in communication system from Lund University, Lund, Sweden, in 2013. He is currently working toward the Ph.D. degree with the Royal Institute of Technology (KTH), Stockholm, Sweden, and is with the University of Gävle, Gävle, Sweden. He has few years of industrial work experience and was actively involved in many research and development projects in UMTS and LTE. His current research interests include signal processing in wireless communication and sensor networks.

Niclas Björsell received the Ph.D. degree in telecommunication from the Royal Institute of Technology, Stockholm, Sweden, in 2007. He is currently an Associate Professor with the University of Gävle, Gävle, Sweden, where he is the Head of the Electronics Group and Program Director for Automation Engineering. He has authored or co-authored more than 80 papers in international peer-review journals and conferences. His research interests include radio frequency measurement technology, analog-to-digital conversion, non-linear systems, wireless communication, and automation. He is an Associate Editor for the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.

Mats Bengtsson (M'00–SM'06) received the M.S. degree in computer science from Linköping University, Linköping, Sweden, in 1991, and the Tech. Lic. and Ph.D. degrees in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 1997 and 2000, respectively.

From 1991 to 1995, he was with Ericsson Telecom AB, Karlstad, Sweden. He currently a Professor in Signal Processing with the Information Science and Engineering Department, School of Electrical Engineering and Computer Science, KTH. His research interests include statistical signal processing and optimization theory and its applications to communications, multi-antenna processing, cooperative communication, radio resource management, and propagation channel modelling. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, Technical Chair for IEEE SPAWC 2015, and was a member of the IEEE SPCOM Technical Committee 2007–2012.