IET Communications

Research Article

Distributed *M*-ary hypothesis testing for decision fusion in multiple-input multiple-output wireless sensor networks

ISSN 1751-8628
Received on 18th October 2019
Revised 19th March 2020
Accepted on 19th August 2020
doi: 10.1049/iet-com.2019.1053
www.ietdl.org

Ali Jamoos[1] ✉, Rushdi Abuawwad[1]
[1]Department of Electronic, Communication Engineering, Al-Quds University, Jerusalem, Palestine
✉ *E-mail: ali.jamoos@staff.alquds.edu*

Abstract: In this study, the authors study binary decision fusion over a shared Rayleigh fading channel with multiple antennas at the decision fusion centre (DFC) in wireless sensor networks. Three fusion rules are derived for the DFC in the case of distributed *M*-ary hypothesis testing, where *M* is the number of hypothesis to be classified. Namely, the optimum maximum a posteriori (MAP) rule, the augmented quadratic discriminant analysis (A-QDA) rule and MAP observation bound. A comparative simulation study is carried out between the proposed fusion rules in-terms of detection performance and receiver operating characteristic (ROC) curves, where several parameters are taken into account such as the number of antennas, number of local detectors, number of hypothesis and signal-to-noise ratio. Simulation results show that the optimum (MAP) rule has better detection performance than A-QDA rule. In addition, increasing the number of antennas will improve the detection performance up to a saturation level, while increasing the number of the hypothesis will deteriorate the detection performance.

1 Introduction

Wireless sensor networks (WSNs) have generated intensive interest from the research community in the last two decades [1–7]. WSNs usually include a large number of nodes where each node is equipped with a sensor to detect physical phenomena such as light, heat, pressure, temperature etc. WSNs are widely being used for sensing in smart environments by using emerging wireless communication techniques. Typical applications include battlefield surveillance, environment monitoring and structure monitoring, among others.

Decision fusion enables sensors to improve classification accuracy while reducing energy consumption and bandwidth demand for data transmission. Each sensor collects and possibly processes data about the phenomenon and transmits its observation or local decisions to decision fusion centre (DFC) for a final decision. The DFC makes a global decision about the state of the phenomenon based on the received local decisions from the sensors and possibly triggers an appropriate action [8–12].

Decentralised detection in a bandwidth-constrained sensor network with binary decisions made by sensor nodes has been investigated in [13]. Universal detectors for the decision fusion problem have also been considered in [14]. In [15], the authors have studied the formulation of fusion rule that minimise the error probability at the fusion centre, which achieve the maximum probability of the correct global classification. For parameter detection or estimation problems in WSNs, an important question is how to exploit a multi-antenna DFC to improve the probability of detection or reduce estimation error. Several recent papers have studied the benefit provided by multiple antennas in the WSN context [15, 16]. The well-known architecture for WSN assumes that each sensor node communicates over a parallel access channel, where each sensor can exploit a channel to communicate with the DFC. In [17, 18], the authors have recommended to exploit the wireless air interface as a multiple-access channel (MAC) for DFC, where several sensors communicate with a single DFC through a common channel.

In [19–21] the authors studied channel-aware binary-decision fusion with multiple antennas at the DFC. They presented several sub-optimal fusion rules, where decode-and-fuse and decode-then-fuse methods are compared by simulation. Among these sub-optimal fusion rules are the maximum ratio combining (MRC), log-likelihood ratio, equal gain combining and Chair–Varshney (CV) maximum likelihood. In [19], to limit complexity, the authors assume that the sensors make independent local decisions on the hypotheses based on their respective observations and forward these decisions over a multiple-input multiple-output (MIMO) channel to a DFC which forms a final decision on the hypothesis. The authors in [20] show that when the number of DFC antennas is very large, low complexity algorithms can asymptotically achieve an upper bound on detection performance even using a linear receiver with imperfect channel state information (CSI). In [21], the authors have presented a theoretical analysis of the MRC rule for decision fusion through MIMO fading channels with dependent and independent local sensor decisions. A stochastic geometry approach is followed in [22] for deriving the optimal fusion rule at the cluster level, which has been shown to approach the performance of CV rule

The studies in [23, 24] investigate channel-aware *M*-ary distributed detection, where the communication channels are modelled as additive white Gaussian noise (AWGN) and Rayleigh fading with perfect CSI available at the DFC. Most studies of parallel distributed detection for *M*-ary hypothesis testing assume that for each observation the local detector (LD) transmits at least $\log_2 M$ bits to the DFC, where M is the number of hypotheses to be classified. The authors in [25] assumed that it is possible to transmit using less than $\log_2 M$. In addition, they develop conditions for asymptotic detection of the correct hypothesis by the DFC, formulate the optimal decision rules for the DFC, and derive expressions for the performance of the system.

In this paper, we propose to extend the analysis in [19] to include distributed *M*-ary hypothesis testing [25]. Particularly, we have considered channel-aware decision fusion in distributed MIMO WSN with *M*-ary hypothesis testing and binary local decisions as shown in Fig. 1. Three fusion rules are designed and analysed for the classification task: the optimum maximum a posteriori (MAP) rule, the augmented quadratic discriminant analysis (A-QDA) rule and the MAP observation bound.

The rest of the paper is organised as follows. Section 2 presents the proposed system model. In Section 3, we introduce and compare three fusion rules for the classification task. In Section 4, we compare the presented fusion rules through simulations. In Section 5, some conclusions are drawn. Finally, the Appendix contains proofs and derivations.

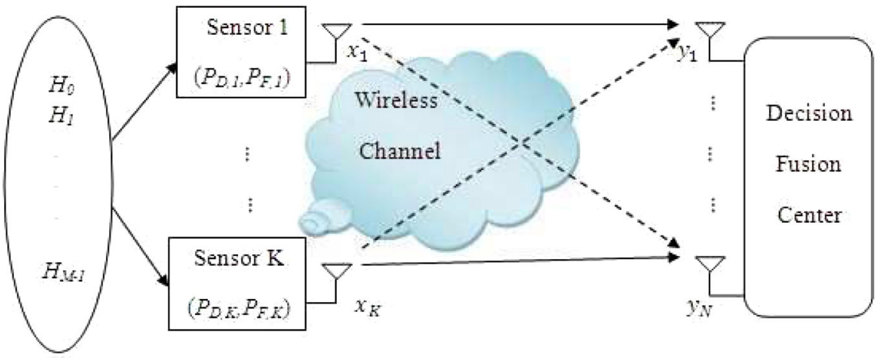

Fig. 1 *Proposed decision fusion model with distributed M-ary hypothesis testing and MIMO fading channel*

2 System model

We study a distributed M-hypotheses test, where K sensor nodes are employed to discriminate among the hypotheses of the set $\mathscr{H} = \{H_1, \ldots, \mathscr{H}_M\}$. The a priori probability of hypothesis $\mathscr{H}_i \in \mathscr{H}$ is denoted by $P(\mathscr{H}_i)$. The kth sensor, $k \in \mathscr{K} \triangleq \{1, 2, \ldots, K\}$, takes a binary decision $d_k \in \mathscr{X}$, where $\mathscr{X} \triangleq \{-1, 1\}$, about the perceived phenomenon based on its own measurements. Our distributed detection system employs K LDs to survey a common volume for evidence of one of the M hypotheses within $\mathscr{H}$. These LDs are restricted to make a single binary decision per observation, i.e. they have to compress each observation into either '1' or '−1', representing a BPSK modulation. The DFC uses the vector of local decisions $\boldsymbol{d} \in \{-1, 1\}^K$ to form a (potentially more accurate) global decision $\hat{H}$ in favour of one of the M hypotheses. The generic decision d_k is assumed to be independent of the other decisions d_ℓ, $\ell \in \mathscr{K}, \ell \neq k$, conditioned on $\mathscr{H}_m \in \mathscr{H}$.

In this context, it is appropriate to model the marginal pmf of kth sensor decisions through a set of transition probabilities $\rho_{k,m}$, $m = 1, \ldots M$, where $\rho_{k,m}$ is the probability that the kth sensor transmits $d_k = 1$ to the DFC when the phenomenon $\mathscr{H}_m$ is present, namely

$$\rho_{k,m} \triangleq \Pr\{d_k = 1 | \mathscr{H}_m\}\ . \tag{1}$$

The above probabilities are summarised for kth sensor in the vector $\boldsymbol{\rho}_k \triangleq [\rho_{k,1} \ \ \ldots \ \ \rho_{k,m}]^{\mathrm{T}}$. The sensors transmit their decisions to the DFC through flat-fading MAC, with i.i.d. Rayleigh fading processes of unitary mean power. The DFC has N receiving antennas that can be used for diversity reception to combat signal fading of the wireless channel. This arrangement can be seen as a distributed (or 'virtual' [9]) MIMO channel. In addition, it is assumed that synchronisation and instantaneous CSI are maintained at the DFC as in [9].

Let y_n denotes the received signal from the nth receiving antenna of the DFC; $h_{n,k} \sim \mathscr{N}_{\mathbb{C}}(0, 1)$ is the fading process between the kth sensor and the nth receiving antenna; w_n is the AWGN process at the nth receiving antenna. The vector model of the received signal at the DFC can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{d} + \boldsymbol{w} \tag{2}$$

where $\boldsymbol{y} \in \mathbb{C}^N$, $\boldsymbol{H} \in \mathbb{C}^{N \times K}$, $\boldsymbol{d} \in \mathscr{X}^K$, $\boldsymbol{w} \sim \mathscr{N}_{\mathbb{C}}(\boldsymbol{0}_N, \sigma_w^2 \boldsymbol{I}_N)$ denote the received signal vector, the fading channel matrix, the transmitted signal vector and the AWGN vector, respectively. It is not difficult to show that the received signal, under hypothesis $\mathscr{H}_m$, is distributed as

$$\boldsymbol{y} \Big| \mathscr{H}_m \sim \sum_{\boldsymbol{d} \in \mathscr{X}^K} \mathscr{N}_{\mathbb{C}}(\boldsymbol{H}\boldsymbol{d}, \sigma_w^2 \boldsymbol{I}_N)\, P(\boldsymbol{d} \Big| \mathscr{H}_m), \tag{3}$$

Remarks: The received signal model in (2) is usually classified as underloaded ($K < N$), fully-loaded ($K = N$) or overloaded ($K > N$). All of these three cases are of interest for MIMO wireless systems. In particular, the overloaded case is the only reasonable scenario for WSNs, due to the fact that the number of sensor nodes is usually much larger than the number of receiving antennas at the DFC (i.e. $K \gg N$). The signal-to-noise ratio (SNR) is referred to as the ratio between the average received power from the WSN $\mathscr{E}_{\mathrm{s}} = \mathbb{E}\{\| \boldsymbol{H}\boldsymbol{x} \|^2\}$ and the AWGN process variance σ_{w}^2, that is $\mathrm{SNR} \triangleq \mathscr{E}_{\mathrm{s}}/\sigma_{\mathrm{w}}^2 = KN/\sigma_{\mathrm{w}}^2$. Note that the SNR for the individual kth sensor node will be $\mathrm{SNR}_k = N/\sigma_{\mathrm{w}}^2$.

Proposition 1: The second order characterisation of the received vector under hypothesis $\mathscr{H}_m$ (i.e. $\boldsymbol{y} | \mathscr{H}_m$) is given by:

$$\mathbb{E}\{\boldsymbol{y} | \mathscr{H}_m\} = \boldsymbol{H}\, \mathbb{E}\{\boldsymbol{d} | \mathscr{H}_m\} \tag{4}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{y}|\mathscr{H}_m} = \boldsymbol{H}\, \boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m} \boldsymbol{H}^{\dagger} + \sigma_w^2 \boldsymbol{I}_N \tag{5}$$

$$\bar{\boldsymbol{\Sigma}}_{\boldsymbol{y}|\mathscr{H}_m} = \boldsymbol{H}\, \boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m} \boldsymbol{H}^{\mathrm{T}} \tag{6}$$

which denote the mean vector, the covariance and the pseudo-covariance, respectively.

Proof: The proof is provided in Appendix. □

The *augmented covariance* of $\boldsymbol{y} | \mathscr{H}_m$ is given in closed form as

$$\boldsymbol{\Sigma}_{\underline{\boldsymbol{y}}|\mathscr{H}_m} = \underline{\boldsymbol{H}}\, \boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m} \underline{\boldsymbol{H}}^{\dagger} + \sigma_w^2 \boldsymbol{I}_{2N}\ . \tag{7}$$

3 Fusion rules

3.1 Optimum MAP rule

The optimal criterion [26] for the suggested issue is that minimising the fusion error-probability, that is the MAP test, formulated as

$$\widehat{\mathscr{H}}_{\mathrm{map}} \triangleq \arg\max_{\mathscr{H}_m} P(\mathscr{H}_m \Big| \boldsymbol{y}) \tag{8}$$

$$= \arg\max_{\mathscr{H}_m} \frac{p(\boldsymbol{y} | \mathscr{H}_m)\, P(\mathscr{H}_m)}{p(\boldsymbol{y})} \tag{9}$$

$$= \arg\max_{\mathscr{H}_m} \frac{p(\boldsymbol{y} | \mathscr{H}_m)\, P(\mathscr{H}_m)}{p(\boldsymbol{y})} \tag{10}$$

$$= \arg\max_{\mathscr{H}_m} p(\boldsymbol{y} \Big| \mathscr{H}_m)\, P(\mathscr{H}_m) \tag{11}$$

$$= \arg\max_{\mathscr{H}_m} \ln p(\boldsymbol{y} \Big| \mathscr{H}_m) + \ln \pi_m\,. \tag{12}$$

where $\widehat{\mathscr{H}}$ and $\pi_m \triangleq P(\mathscr{H}_m)$. An explicit expression of the log-likelihood $\ln p(\boldsymbol{y} | \mathscr{H}_m)$ from (8) can be expressed as

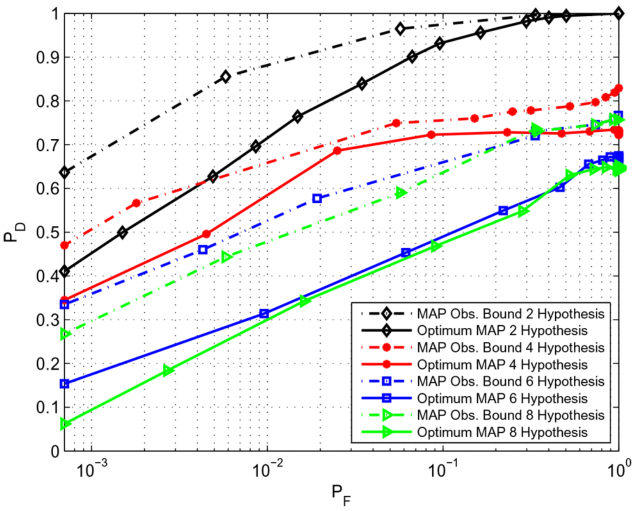

Fig. 2 *ROC for the optimum MAP and observation bound rules. Channel SNR* $= 15$ *dB*, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, *number of LDs* $K = 8$ *and number of antenna* $N = 2$

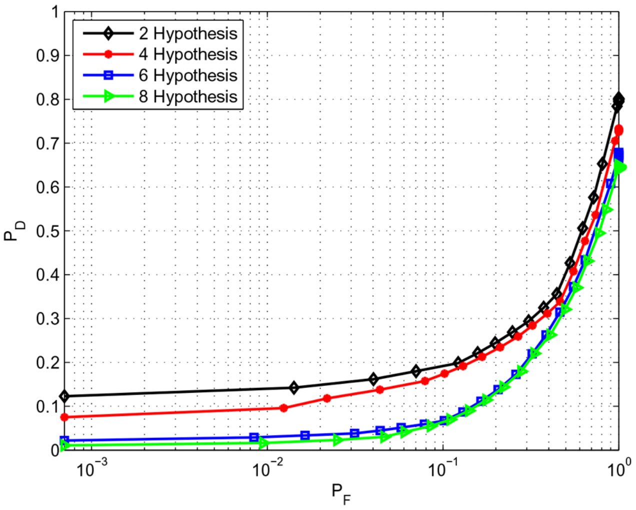

Fig. 3 *ROC for the A-QDA rule. Channel SNR* $= 15$ *dB*, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, *number of LDs* $K = 8$ *and number of antenna* $N = 2$

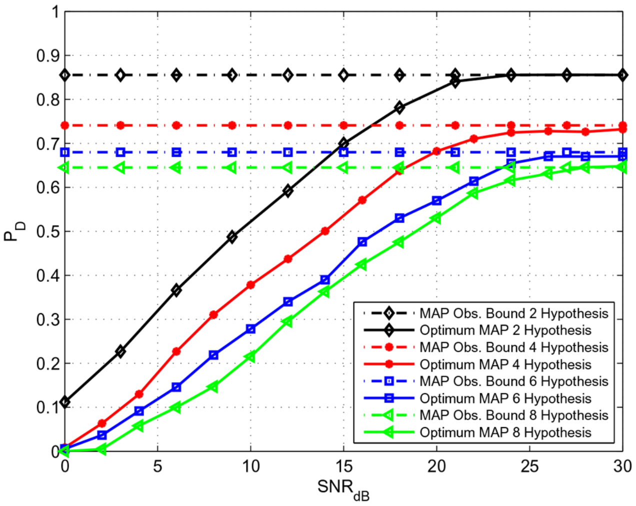

Fig. 4 P_D *versus channel* (SNR) *dB for the optimum MAP and observation bound rules*, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, *number of LDs* $K = 8$ *and number of antenna* $N = 2$

$$\ln p(\boldsymbol{y} \mid \mathscr{H}_m) = \ln\left[\sum_{\boldsymbol{d} \in \mathscr{X}^K} p(\boldsymbol{y} \mid \boldsymbol{d})\, P(\boldsymbol{d} \mid \mathscr{H}_m)\right] = \ln\left[\sum_{\boldsymbol{d} \in \mathscr{X}^K} \frac{1}{\sigma_w^2} \exp\left(-\frac{\| \boldsymbol{y} - \boldsymbol{H}\boldsymbol{d} \|^2}{\sigma_w^2}\right) P(\boldsymbol{d} \mid \mathscr{H}_m)\right] \quad (13)$$

where we have exploited the conditional independence of $\boldsymbol{y}$ from $\mathscr{H}_m$ (given $\boldsymbol{d}$). Nevertheless, the optimal rule in (13) have some difficulties in its implementation. Namely, (i) availability of $\hat{\boldsymbol{G}}$, $P(\boldsymbol{x} \mid \mathscr{H}_i)$ and σ_w^2 and (ii) instability of the mathematical expression containing exponential terms with large dynamics [19, 27]. In addition, the computational complexity growth exponentially with K which prevent practical implementation. To reduce complexity, several sub-optimal DF rules with simpler implementation is presented and compared in [19].

3.2 MAP observation bound

For comparison purposes, let us recall the observation bound (upper bound) [18] which yields the optimum performances over noiseless channel. It can be expressed by the following classifier:

$$\widehat{\mathscr{H}}_{\text{obs}} \triangleq \arg\max_{\mathscr{H}_m} P(\mathscr{H}_m \mid \boldsymbol{d}) \quad (14)$$

$$= \arg\max_{\mathscr{H}_m} \ln p(\boldsymbol{d} \mid \mathscr{H}_m) + \ln \pi_m . \quad (15)$$

Clearly, the MAP observation bound rule should be intended as an optimistic upper bound on the classification performance which can be achieved over a virtual MIMO channel.

3.3 A-QDA rule

In this subsection, the second order characterisation provided in (4)–(6) are fully exploited. Indeed, after fitting $\underline{\boldsymbol{y}} \mid \mathscr{H}_m$ to an improper complex Gaussian, a classifier based on a complex version of quadratic discriminant analysis can be obtained as [28]

$$\widehat{\mathscr{H}}_{\text{qda}} \triangleq \arg\min_{\mathscr{H}_m} \left\{(\underline{\boldsymbol{y}} - \mathbb{E}\{\underline{\boldsymbol{y}} \mid \mathscr{H}_m\})^{\dagger} \boldsymbol{\Sigma}_{\underline{\boldsymbol{y}}|\mathscr{H}_m}^{-1} (\underline{\boldsymbol{y}} - \mathbb{E}\{\underline{\boldsymbol{y}} \mid \mathscr{H}_m\}) + \ln\det\left(\boldsymbol{\Sigma}_{\underline{\boldsymbol{y}}|\mathscr{H}_m}^{-1}\right) + \ln \pi_m\right\} \quad (16)$$

where $\mathbb{E}\{\underline{\boldsymbol{y}} \mid \mathscr{H}_m\} = \underline{\boldsymbol{H}}\, \mathbb{E}\{\boldsymbol{d} \mid \mathscr{H}_m\}$ and $\boldsymbol{\Sigma}_{\underline{\boldsymbol{y}}|\mathscr{H}_m}$ is given in (7).

4 Simulation results

In this section, a comparative simulation study is carried-out between the suggested fusion rules applied at the DFC in the proposed WSN system model. The detection performance and receiver operating characteristic (ROC) curves are obtained for the derived fusion rules: the optimum MAP rule, the MAP observation bound, and the A-QDA rule. In addition, we studied the effect of various parameters on the detection performance of the fusion rules such as the channel SNR, the number of the antenna at the DFC (i.e. N), number of hypothesis (i.e. M) and the local sensors performance indices (i.e. P_{dk} and P_{fk}).

ROC curves: In Figs. 2 and 3 we show the ROC (i.e. P_{D} versus P_{F}), for different fusion rules in a WSN with $K = 8$ sensors and $N = 2$ antennas at the DFC, with channel $(\mathrm{SNR})_{\mathrm{dB}} = 15$ for different number of hypothesis. From a detection performance point of view, it can be noticed from Figs. 2 and 3 that the optimum MAP fusion rule and its observation bound provide better performance than A-QDA rule. In addition, increasing the number of hypothesis will decrease the detection performance.

Probability of detection P_{D} *versus* $(\mathrm{SNR})_{\mathrm{dB}}$: Figs. 4 and 5 illustrate the probability of detection P_{D} versus the channel $(\mathrm{SNR})_{\mathrm{dB}}$ for the presented fusion rules with different number of hypothesis $\mathscr{H} \in \{H_2, H_4, H_6, H_8\}$. We consider a WSN with a fixed number of sensor nodes $K = 8$ and antennas $N = 2$. It can be

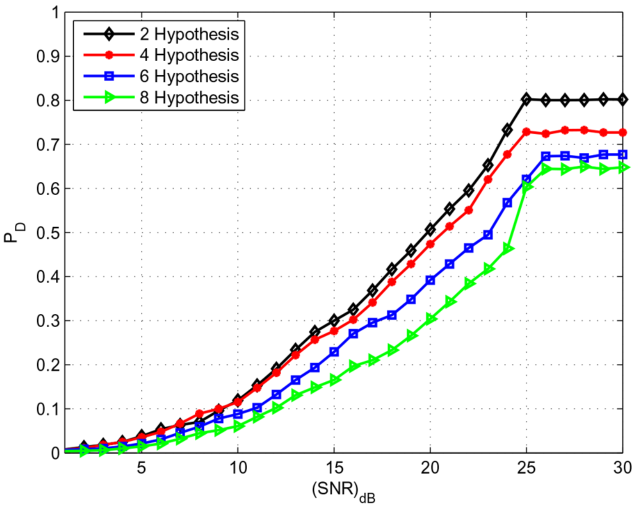

Fig. 5 *P_D versus channel (SNR) dB for the A-QDA rule, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, number of LDs $K = 8$ and number of antenna $N = 2$*

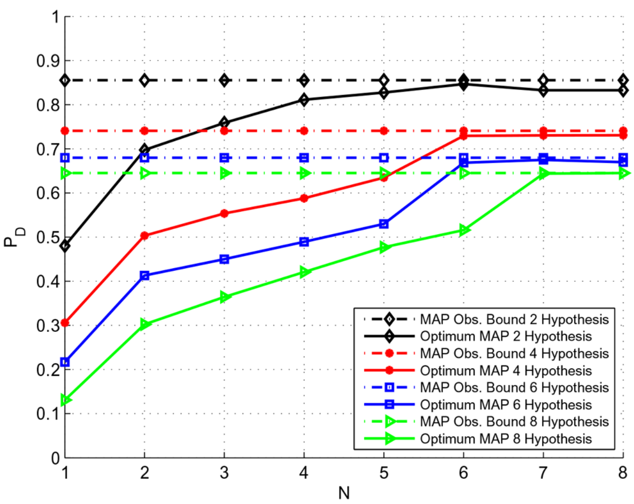

Fig. 6 *P_D versus N for the optimum MAP and observation bound rules, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, number of LDs $K = 8$ and (SNR) = 15 dB*

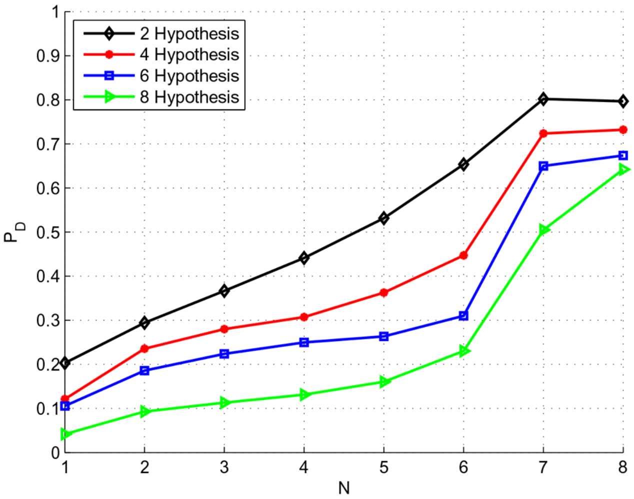

Fig. 7 *P_D versus N for the A-QDA rule, $P_{D,k} = (0.5)$, $P_{F,k} = (0.05)$, number of LDs $K = 8$ and (SNR) = 15 dB*

noticed from Figs. 4 and 5 that at higher channel SNR we can achieve higher detection performance. In Addition, the MAP observation bound yields an upper bound for the optimum MAP rule. Furthermore, the optimum MAP rule and its observation bound yields much better detection performance than the A-QDA rule for a wide range of channel SNR and for different hypothesis scenarios. Moreover, Fig. 5 shows that the proposed WSN system model could significantly improve the detection performance of A-QDA fusion rule at high channel SNR for different hypothesis scenarios.

Probability of detection P_D versus N: Figs. 6 and 7 illustrate the probability of detection P_D versus the number of antenna N for the presented fusion rules with different hypothesis $\mathscr{H} \in \{H_2, H_4, H_6, H_8\}$ and the case $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in K$. We consider WSN with $K = 8$ sensor nodes and a channel $(\text{SNR})_{\text{dB}} \simeq 15$. It is observed that increasing the number of antennas at the DFC yields better performance for the presented rules, however a saturation effect is noticed. The saturation level not only depends on the number of antennas and SNR, but also on the chosen fusion rule and number of hypothesis. In addition, the optimum MAP rule and its observation bound result in much better detection performance than the A-QDA rule. In particular, specific arrangements achieve the upper bound (optimum MAP with $N = 4$ at $(\text{SNR})_{\text{dB}} = 15$) and (A-QDA with $N = 7$ at $(\text{SNR})_{\text{dB}} = 15$).

5 Conclusions

In this paper, distributed M-ary hypothesis testing for decision fusion in MIMO WSNs over Rayleigh fading channel is addressed. Three fusion rules are derived and analysed for the classification task. Namely, the optimum MAP rule, the MAP observation bound and the A-QDA rule. The simulation study we carried-out between the suggested fusion rules showed that the optimum MAP rule and its observation bound outperform the A-QDA rule in-terms of the detection performance for the various scenarios considered. In addition, better detection performance is obtained for the various rules when increasing the number of antennas or decreasing number of hypothesis. Furthermore, it was observed that the detection performance reaches a saturation level when increasing the number of antennas or SNR for the various fusion rules with different hypothesis scenarios.

6 References

[1] Pottie, G.J., Kaiser, W.J.: 'Wireless integrated network sensors', *Commun. ACM*, 2000, **43**, (5), pp. 51–58

[2] Kumar, S., Zhao, F., Shepherd, D.: 'Collaborative signal and information processing in microsensor networks', *IEEE Signal Process. Mag.*, 2002, **19**, (2), pp. 13–14

[3] Estrin, D., Girod, L., Pottie, G., *et al.*: 'Instrumenting the world with wireless sensor networks'. Proc. 2001 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, May 2001, pp. 2033–2036

[4] Guo, X., He, Y., Atapattu, S., *et al.*: 'Power allocation for distributed detection systems in wireless sensor networks with limited fusion center feedback', *IEEE Trans. Wirel. Commun.*, 2018, **66**, (10), pp. 4753–4766

[5] Panigrahi, S.R., Bjorsell, N., Bengtsson, M.: 'Distributed detection with non-identical wireless sensors for industrial applications'. 2019 IEEE Int. Conf. on Industrial Technology (ICIT), Melbourne, Australia, February 2019, pp. 1403–1408

[6] Al-Jarrah, M.A., Al-Dweik, A., Kalil, M., *et al.*: 'Decision fusion in distributed cooperative wireless sensor networks', *IEEE Trans. Veh. Technol.*, 2019, **68**, (1), pp. 797–811

[7] Al-Jarrah, M.A., Yaseen, M.A., Al-Dweik, A., *et al.*: 'Decision fusion for IOT-based wireless sensor networks', *IEEE Internet Things J.*, 2020, **7**, (2), pp. 1313–1326

[8] Chen, B., Tong, L., Varshney, P.K.:'Channel-aware distributed detection in wireless sensor networks', *IEEE Signal Process. Mag.*, 2006, **23**, (5), pp. 16–26

[9] Zhang, X., Poor, H.V., Chiang, M.: 'Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks', *IEEE Trans. Signal Process.*, 2008, **56**, (9), pp. 4124–4140

[10] Zhang, W., Zhang, Z.: 'Belief function based decision fusion for decentralized target classification in wireless sensor networks', *Sensors*, 2015, **15**, (8), pp. 20524–20540

[11] Hussain, B.M., Jamoos, A.: 'Fusion of likelihood ratio test based decisions in wireless sensor networks'. Int. Conf. on Communications, Signal Processing, and their Applications (ICCSPA), Sharjah, May 2015

[12] Jamoos, A.: 'Improved decision fusion model for wireless sensor networks over Rayleigh fading channels', *Technologies*, 2017, **5**, (1), pp. 1327–1332

[13] Xiao, J., Luo, Z.: 'Universal decentralized detection in a bandwidth-constrained sensor network', *IEEE Trans. Signal Process.*, 2005, (53), pp. 2617–2624

[14] Ciuonzo, D., Maio, A., Rossi, P.S.: 'A systematic framework for composite hypothesis testing of independent bernoulli trials', *IEEE Signal Process. Lett.*, 2015, (22), pp. 1249–1253

[15] Niu, R., Chen, B., Varshney, P.K.: 'Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks', *IEEE Trans. Signal Process.*, 2006, **54**, (3), pp. 1018–1027

[16] Banavar, M.K., Smith, A.D., Tepedelenlioglu, C., *et al.*: 'On the effectiveness of multiple antennas in distributed detection over fading MACS', *IEEE Trans Wirel. Commun.*, 2012, **11**, (5), pp. 1744–1752
[17] Li, W., Dai, H.: 'Distributed detection in wireless sensor networks using a multiple access channel', *IEEE Trans. Signal Process.*, 2007, **55**, pp. 822–833
[18] Berger, C.R., Guerriero, M., Zhou, S., *et al.*: 'PAC vs. MAC for decentralized detection using noncoherent modulation', *IEEE Trans. Signal Process.*, 2009, **57**, (9), pp. 3562–3575
[19] Ciuonzo, D., Romano, G., Rossi, P.S.: 'Channel-aware decision fusion in distributed MIMO wireless sensor networks: decode-and-fuse vs. decode-then-fuse', *IEEE Trans. Wirel. Commun.*, 2012, **11**, pp. 4716–4728
[20] Ciuonzo, D., Rossi, P.S., Dey, S.: 'Massive MIMO channel-aware decision fusion', *IEEE Trans. Signal Process.*, 2015, **63**, (3), pp. 4716–4728
[21] Ciuonzo, D., Romano, G., Rossi, P.S.: 'Performance analysis of maximum ratio combining in channel-aware MIMO decision fusion', *IEEE Trans. Wirel. Commun.*, 2013, **12**, (9), pp. 4716–4728
[22] Aldalahmeh, S.A., Ghogho, M., Mclernon, D., *et al.*: 'Optimal fusion rule for distributed detection in clustered wireless sensor networks'. *EURASIP J. Adv. Signal Process.*, 2016
[23] Kotecha, J.H., Ramachandran, V., Sayeed, A.M.: 'Distributed multitarget classification in wireless sensor networks', *IEEE J. Sel. Areas Commun.*, 2005, **23**, (4), pp. 703–713
[24] Maleki, N., Vosoughi, A.: 'Channel-aware M-ary distributed detection: optimal and suboptimal fusion rules'. IEEE Statistical Signal Processing Workshop (SSP), Ann Arbor, MI, USA, August 2012, pp. 644–647
[25] Rorres, C., Zhu, X., Yuan, Y., *et al.*: 'Distributed *M*-ary hypothesis testing with binary local decisions', *Inf. Fusion*, 2004, **5**, (3), pp. 157–167
[26] Kay, S.M.: '*Fundamentals of statistical signal processing*' (Prentice-Hall PTR, Upper Saddle River, NJ, USA, January 1998)
[27] Lei, A., Schober, R.: 'Coherent max-log decision fusion in wireless sensor networks', *IEEE Trans. Commun.*, 2010, **58**, (5), pp. 1327–1332
[28] Schreier, P.J., Scharf, L.L.: '*Statistical signal processing of complex-valued data*' (Cambridge, 2010)

7 Appendix

7.1 Second-order characterisation of $\boldsymbol{y}|\mathscr{H}_m$

In this appendix we provide a second-order characterisation for $\boldsymbol{y}|\mathscr{H}_m$. First, We recall that the exact pdf is

$$\boldsymbol{y}\Big|\mathscr{H}_m \sim \sum_{\boldsymbol{d} \in \mathcal{X}^K} P(\boldsymbol{d}\Big|\mathscr{H}_m)\,\mathcal{N}_{\mathbb{C}}(\boldsymbol{H}\boldsymbol{d},\, \sigma_w^2 \boldsymbol{I}_N), \tag{17}$$

which is recognised as a mixture of 2^K *proper* complex-valued Gaussian vectors. Then, we evaluate the mean vector of $\boldsymbol{y}|\mathscr{H}_m$ as follows:

$$\mathbb{E}\{\boldsymbol{y}\Big|\mathscr{H}_m\} = \sum_{\boldsymbol{d} \in \mathcal{X}^K} \mathbb{E}\{\boldsymbol{y}\Big|\boldsymbol{d}\}\, P(\boldsymbol{d}\Big|\mathscr{H}_m) = \tag{18}$$

$$\boldsymbol{H} \sum_{\boldsymbol{d} \in \mathcal{X}^K} \boldsymbol{d}\, P(\boldsymbol{d}\Big|\mathscr{H}_m) = \boldsymbol{H}\,\mathbb{E}\{\boldsymbol{d}\Big|\mathscr{H}_m\} \tag{19}$$

It is worth noting that (19) was obtained by exploiting $\mathbb{E}\{\boldsymbol{w}\} = \boldsymbol{0}_N$. Differently, the covariance matrix is evaluated as

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{y}|\mathscr{H}_m} &= \mathbb{E}\{[\boldsymbol{H}(\boldsymbol{d} - \mathbb{E}\{\boldsymbol{d}|\mathscr{H}_m\}) + \boldsymbol{w}] \\ &[\boldsymbol{H}(\boldsymbol{d} - \mathbb{E}\{\boldsymbol{d}|\mathscr{H}_m\}) + \boldsymbol{w}]^{\dagger}\Big|\mathscr{H}_m\Big\} =\end{aligned} \tag{20}$$

$$\boldsymbol{H}\,\boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m}\boldsymbol{H}^{\dagger} + \mathbb{E}\{\boldsymbol{w}\boldsymbol{w}^{\dagger}\} = \tag{21}$$

$$\boldsymbol{H}\,\boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m}\boldsymbol{H}^{\dagger} + \sigma_w^2 \boldsymbol{I}_N \tag{22}$$

since: (i) $\boldsymbol{x}$ and $\boldsymbol{w}$ are *mutually independent* and (ii) $\mathbb{E}\{\boldsymbol{w}\} = \boldsymbol{0}_N$. Similarly, we obtain the *complementary covariance* [28] as

$$\begin{aligned}\bar{\boldsymbol{\Sigma}}_{\boldsymbol{y}|\mathscr{H}_m} &= \mathbb{E}\{[\boldsymbol{H}(\boldsymbol{d} - \mathbb{E}\{\boldsymbol{d}|\mathscr{H}_m\}) + \boldsymbol{w}] \\ &[\boldsymbol{H}(\boldsymbol{d} - \mathbb{E}\{\boldsymbol{d}|\mathscr{H}_m\}) + \boldsymbol{w}]^{\mathrm{T}}\Big|\mathscr{H}_m\Big\} =\end{aligned} \tag{23}$$

$$\boldsymbol{H}\,\boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m}\boldsymbol{H}^{\mathrm{T}} + \mathbb{E}\{\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\} = \tag{24}$$

$$\boldsymbol{H}\,\boldsymbol{\Sigma}_{\boldsymbol{d}|\mathscr{H}_m}\boldsymbol{H}^{\mathrm{T}} \tag{25}$$

where the last equality follows from $\mathbb{E}\{\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\} = \boldsymbol{O}_{N\times N}$ (indeed $\boldsymbol{w}$ is a *proper* random vector). Therefore, we conclude that $\boldsymbol{y}|\mathscr{H}_m$ is an *improper random vector*, since its complementary covariance matrix does not vanish, thus motivating the potential for WL processing.