Energy Detection for Decision Fusion in Wireless Sensor Networks over Ricean-Mixture Fading

Pierluigi Salvo Rossi[†‡], Domenico Ciuonzo[†], Torbjörn Ekman[‡], Kimmo Kansanen[‡]

[†]Department of Industrial and Information Engineering, Second University of Naples, Italy.
[‡]Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Norway.
email: {salvorossi, domenico.ciuonzo}@ieee.org; {torbjorn.ekman, kimmo.kansanen}@iet.ntnu.no

Abstract—In this paper we focus on the energy detector for decision fusion in wireless sensor networks over multiple access channels. More specifically, we derive analytical performance in terms of global probability of false alarm and detection (including asymptotic performance for large number of sensors) when the fading is a Ricean mixture, i.e. channel coefficients are sampled from a Gaussian mixture (GM) distribution. The motivation for the GM is the ability to model real-world scenarios while keeping mathematical tractability. Analytical results are confirmed through numerical simulations.

I. Introduction

Distributed detection represents a task of interest in many wireless sensor networks applications. Two architectures have been investigated: (i) *centralized* [1], the sensors transmit their local decisions to a fusion center which takes a (more reliable) global decision by appropriately combining the received information; (ii) *decentralized* [2], there is no fusion center and each sensor collects the information from the others in order to reach a reliable decision. Here, we focus on centralized architectures. Commonly, they are based on parallel access channels, where each sensor is provided with a non-interfering dedicated channel to communicate with the fusion center [3]. Near-optimal fusion rules with full channel state information available at the receiver have been proposed in [4].

Recently, the interfering nature of the wireless medium has been exploited in the context of distributed detection [5]–[7]. Also, multiantenna processing techniques at the fusion center have been investigated and compared in terms of performance, complexity, and knowledge requirements [8], [9]. Finally, energy detection was proven to be optimal in Rayleigh fading channels [10].

In this paper we consider a decision-fusion rule based on energy detection over Ricean-mixture fading channels. More specifically, we analyze the case in which each channel coefficient is modeled as a Gaussian mixture (GM). GMs can take into account multi-modality, asymmetry, heavy tails, and other characteristics that may be present in real-world scenarios, though exhibiting interesting properties in terms of mathematical tractability [11]. The main paper contribution is the derivation of analytical expressions for system performance in terms of global probabilities of false alarm and detection.

The outline of the paper is the following: Sec. II collects preliminaries on GMs; Sec. III describes the system model while Sec. IV the statistics for the decision with corresponding system performance; Sec. V compares the performance of simulated systems; final remarks are given in Sec. VI.

II. Preliminaries on GM

We use the following propositions for GM [11].

Proposition 1: Denote $u \sim \mathcal{N}_\mathbb{C}(\mu, \sigma^2)$, then $v = |u|^2$ is scaled non-central chi-squared, $v \sim \chi^2_2(|\mu|; \sigma^2)$, with pdf

$$p(v) = \frac{1}{2\sigma^2} \exp\left(-\frac{v + |\mu|^2}{2\sigma^2}\right) I_0\left(\frac{|\mu|}{\sigma^2}\sqrt{v}\right) .$$

where $I_0(x) = \frac{1}{\pi}\int_0^\pi \exp(x\cos(\theta))\, d\theta$ is the modified Bessel function of the first kind and zero order.

Proposition 2: Denote $u \sim \sum_{m=1}^M \rho_m \mathcal{N}_\mathbb{C}(\mu_m, \sigma_m^2)$ a complex GM with M components, i.e. such that its pdf is

$$p(u) = \sum_{m=1}^M \frac{\rho_m}{2\pi\sigma_m^2} \exp\left(-\frac{|u - \mu_m|^2}{2\sigma_m^2}\right) ,$$

with $\rho_m \geq 0$ and $\sum_{m=1}^M \rho_m = 1$. Using Prop. 1 and the total probability theorem, it is readily shown that $v = |u|^2$ is a scaled non-central chi-squared mixture, $v \sim \sum_{m=1}^M \rho_m \chi^2_2(|\mu_m|; \sigma_m^2)$, with its pdf being

$$p(v) = \sum_{m=1}^M \frac{\rho_m}{2\sigma_m^2} \exp\left(-\frac{v + |\mu_m|^2}{2\sigma_m^2}\right) I_0\left(\frac{|\mu_m|}{\sigma_m^2}\sqrt{v}\right) .$$

Denoting $Q(a; b) = \int_b^\infty x \exp\left(-\frac{x^2 + a^2}{2}\right) I_0(ax)\, dx$ the Marcum Q-function, its tail distribution is

$$\Pr(v > v_0) = \sum_{m=1}^M \rho_m Q\left(\frac{|\mu_m|}{\sigma_m}; \frac{\sqrt{v_0}}{\sigma_m}\right) . \qquad (1)$$

Proposition 3: Consider a set of N independent and identically distributed (*iid*) complex GMs

$$u_n \sim \sum_{m=1}^M \rho_m \mathcal{N}_\mathbb{C}\left(\mu_m, \sigma_m^2\right) , \quad n = 1, \ldots, N ,$$

then the linear combination $v = \sum_{n=1}^N c_n u_n$ is a complex GM with $\binom{M+N-1}{N}$ components, i.e.

$$v \sim \sum_{m_1=1}^M \cdots \sum_{m_N=1}^M \prod_{n=1}^N \rho_{m_n} \mathcal{N}_\mathbb{C}\left(\sum_{n=1}^N c_n \mu_{m_n}, \sum_{n=1}^N |c_n|^2 \sigma_{m_n}^2\right) . \qquad (2)$$

III. SYSTEM MODEL

We consider K sensors sensing a binary source, each taking autonomously a local decision. The two hypotheses are denoted $\mathcal{H}_0$ and $\mathcal{H}_1$. We assume that the local sensing and decision process is fully described by the local probability of false alarm (p_f) and the local probability of detection (p_d), both assumed to be stationary, identical, and conditionally independent given the specific hypothesis. Sensors, each with one single transmit antenna, communicate simultaneously their decision to a fusion center whose aim is to provide a robust decision on the basis of the multiple received information. All the sensors employ the same binary modulation, for energy saving purposes we consider OOK modulation, with identical parameters (transmission pulse, carrier frequency, etc.). We assume that the system is fully synchronized.

We denote $x_k \in \mathcal{X} = \{0, 1\}$ the symbol transmitted by the kth sensor encoding its local decision (we assume 0 for $\mathcal{H}_0$ and 1 for $\mathcal{H}_1$); h_k the fading channel coefficient on the link between the kth sensor and fusion center; y the signal received by the fusion center; and w the additive white Gaussian noise at the receiver. Channel coefficients from the various links are assumed *iid* complex GM with M components, i.e.

$$h_k \sim \sum_{m=1}^{M} \rho_m \mathcal{N}_{\mathbb{C}} \left(\mu_m, \sigma_m^2 \right) \ .$$

We will explicitly report results for two special cases: (i) one nonzero-mean component, namely *Rice fading*; (ii) two zero-mean components, here called *2ZM fading*. Those are the two simplest extensions for the case with one single zero-mean component analyzed in [10]. More specifically, in Rice fading (we omit the subscript for ease of notation) $h_k \sim \mathcal{N}_{\mathbb{C}} \left(\mu, \sigma^2 \right)$ where μ is real[1], while in 2ZM fading $h_k \sim \rho_1 \mathcal{N}_{\mathbb{C}} \left(0, \sigma_1^2 \right) + \rho_2 \mathcal{N}_{\mathbb{C}} \left(0, \sigma_2^2 \right)$.

The discrete-time model for the received signal is

$$y = \boldsymbol{h}^t \boldsymbol{x} + w \ . \tag{3}$$

where $\boldsymbol{x} = (x_1, \ldots, x_K)^t$ denotes the transmitted vector (of local decisions) and $\boldsymbol{h} = (h_1, \ldots, h_K)^t$ is the channel vector.

IV. DECISION FUSION

The decision is usually performed as a test comparing a signal-dependent statistic ($\lambda(y)$) and a fixed threshold (γ)

$$\lambda(y) \underset{\hat{\mathcal{H}}=\mathcal{H}_0}{\overset{\hat{\mathcal{H}}=\mathcal{H}_1}{\gtrless}} \gamma \ , \tag{4}$$

where $\hat{\mathcal{H}}$ denotes the estimated hypothesis. Performance is evaluated in terms of global probability of false alarm (q_f) and global probability of detection (q_d), defined as follows

$$q_f = \Pr\left(\lambda > \gamma | \mathcal{H}_0 \right) \ , \ q_d = \Pr\left(\lambda > \gamma | \mathcal{H}_1 \right) \ . \tag{5}$$

It is worth noticing that $\Pr(\lambda > \gamma | \mathcal{H}_i)$ describes both q_f and q_d (with $i = 0$ and $i = 1$, respectively). The threshold in Eq. (4) is usually selected according to Bayes or Neyman-Pearson criteria [12]. For system performance evaluation, we consider the behavior of the global probability of detection (q_d) versus the global probability of false alarm (q_f), commonly denoted *receiver operating characteristic* (ROC).

---

[1]Assuming real μ does not affect generality as explained in [6].

A. Optimal Test

The log-likelihood ratio (LLR) of the received signal under the two hypotheses provides the optimal test

$$\lambda(y) = \log \left(\frac{p(y|\mathcal{H}_1)}{p(y|\mathcal{H}_0)} \right) = \log \left(\frac{\sum_{\ell=0}^{K} p(y|\ell) \Pr(\ell|\mathcal{H}_1)}{\sum_{\ell=0}^{K} p(y|\ell) \Pr(\ell|\mathcal{H}_0)} \right) \ , \tag{6}$$

where $\ell = \boldsymbol{x}^t \mathbf{1}_K$ is the number of sensors transmitting 1. Eq. (6) is explained by noticing that: OOK is the modulation format; channel coefficients are *iid*.

From Prop. (3), it is straightforward to show that

$$p(y|\ell) = \sum_{\boldsymbol{m}^t \mathbf{1}_M = \ell} \binom{\ell}{\boldsymbol{m}} \frac{\exp(\boldsymbol{m}^t \log(\boldsymbol{\rho}))}{2\pi(\boldsymbol{m}^t \boldsymbol{\sigma}^2 + \sigma_w^2)} \times \\ \exp\left(-\frac{|y - \boldsymbol{m}^t \boldsymbol{\mu}|^2}{2(\boldsymbol{m}^t \boldsymbol{\sigma}^2 + \sigma_w^2)} \right) \ , \tag{7}$$

where $\boldsymbol{m} = (m_1, \ldots, m_M)^t$ is a vector of integers (with m_u representing the number of links experiencing the uth component of the complex GM), $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M)^t$ and $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_M^2)^t$ are vectors collecting the mean and the variance of each component. Eq. (7) clearly shows that the conditional received signal $(y|\ell)$ is a GM with $L(\ell) = \binom{M+\ell-1}{\ell}$ components, denoted for sake of simplicity as

$$y|\ell \sim \sum_{m=1}^{L(\ell)} \theta_m(\ell) \mathcal{N}_{\mathbb{C}} \left(\nu_m(\ell), \omega_m^2(\ell) \right) \ . \tag{8}$$

In the case of Rice fading, Eq. (8) simplifies in

$$p(y|\ell) = \frac{1}{2\pi(\ell\sigma^2 + \sigma_w^2)} \exp\left(-\frac{|y - \ell\mu|^2}{2(\ell\sigma^2 + \sigma_w^2)} \right) \ ,$$

while in the case of 2ZM fading we get

$$p(y|\ell) = \sum_{m=0}^{\ell} \binom{\ell}{m} \frac{(\rho_1)^m (\rho_2)^{\ell-m}}{2\pi(m\sigma_1^2 + (\ell-m)\sigma_2^2 + \sigma_w^2)} \times \\ \exp\left(-\frac{|y|^2}{2(m\sigma_1^2 + (\ell-m)\sigma_2^2 + \sigma_w^2)} \right) \ .$$

However, the optimal test is computationally expensive[2] (exponential with K) and additionally has high knowledge requirements (statistical channel state information (CSI), signal-to-noise ratio (SNR) and local sensor performance).

Finally, it is worth noticing that for conditionally *iid* sensors $\Pr(\ell|\mathcal{H}_i) = \binom{K}{\ell} p^\ell (1-p)^{K-\ell}$, where $p = p_f$ (resp. $p = p_d$) in the case $\mathcal{H}_0$ (resp. $\mathcal{H}_1$).

B. Energy Test

In the case of OOK, a common simpler alternative is obtained replacing the LLR with the energy of the received signal, i.e.

$$\lambda(y) = |y|^2 \ ,$$

---

[2]It is worth mentioning that in the specific case of Rice fading, a low-complexity test with optimal performance has been found [6].

which apparently requires little computational complexity and also has the advantage that neither CSI nor SNR nor local sensor performance are needed. Such a test has been proved to be optimal in Rayleigh fading scenarios [6], [10].

Using the total probability theorem and combining Eq. (1) in Prop. 2 with Eq. (5), we get

$$\Pr(\lambda > \gamma | \mathcal{H}_i) = \sum_{\ell=0}^{K} \sum_{m=1}^{L(\ell)} \Pr(\ell | \mathcal{H}_i) \theta_m(\ell) Q\left(\frac{|\nu_m(\ell)|}{\omega_m(\ell)}; \frac{\sqrt{\gamma}}{\omega_m(\ell)}\right).$$

(9)

In the case of Rice fading, Eq. (9) simplifies in

$$\Pr(\lambda > \gamma | \mathcal{H}_i) = \sum_{\ell=0}^{K} \Pr(\ell | \mathcal{H}_i) Q\left(\frac{\ell|\mu|}{\sqrt{\ell\sigma^2 + \sigma_w^2}}; \frac{\sqrt{\gamma}}{\sqrt{\ell\sigma^2 + \sigma_w^2}}\right).$$

(10)

while in the case of 2ZM fading we get

$$\Pr(\lambda > \gamma | \mathcal{H}_i) = \sum_{\ell=0}^{K} \sum_{m=0}^{\ell} \Pr(\ell | \mathcal{H}_i) \binom{\ell}{m} \rho_1^m \rho_2^{\ell-m} \times$$
$$\exp\left(-\frac{\gamma}{2\left(m\sigma_1^2 + (\ell-m)\sigma_2^2 + \sigma_w^2\right)}\right).$$

(11)

C. Asymptotic Performance of the Energy Detector

If the number of sensors is large ($K >> 1$), from Eq. (3)

$$\frac{y}{\sqrt{K}} | \mathcal{H}_i \approx \frac{1}{\sqrt{K}} \sum_{k=1}^{K} (h_k x_k | \mathcal{H}_i).$$

(12)

By using the *central limit theorem* [13], the real (resp. the imaginary) part of $\frac{y}{\sqrt{K}} | \mathcal{H}_i$ is Gaussian with mean $\sqrt{K}\tilde{\mu}_i = \sqrt{K} p \sum_{m=1}^{M} \rho_m \mu_m$ and variance $\tilde{\sigma}_i^2 = p \sum_{m=1}^{M} \rho_m (\sigma_m^2 + \mu_m^2) - \tilde{\mu}_i^2$ (resp. $\sqrt{K}\check{\mu}_i = 0$ and $\check{\sigma}_i^2 = p \sum_{m=1}^{M} \rho_m \sigma_m^2$), i.e. Eq. (12) describes an improper complex Gaussian.

In the case of Rice fading, and analogously if there is at least one nonzero mean component, the real part dominates and performance approach

$$\Pr(\lambda > \gamma | \mathcal{H}_i) = \mathcal{Q}\left(\frac{\sqrt{\gamma} - \sqrt{K}\tilde{\mu}_i}{\tilde{\sigma}_i}\right) + \mathcal{Q}\left(\frac{\sqrt{\gamma} + \sqrt{K}\tilde{\mu}_i}{\tilde{\sigma}_i}\right),$$

with $\tilde{\mu}_i = p\mu$ and $\tilde{\sigma}_i^2 = p\sigma^2 + p(1-p)\mu^2$. Differently, in the case of 2ZM fading, and analogously if all the components are zero mean, Eq. (12) describes a proper complex Gaussian ($\tilde{\sigma}_i^2 = \check{\sigma}_i^2$) and performance approach

$$\Pr(\lambda > \gamma | \mathcal{H}_i) = \exp\left(-\frac{\gamma}{2\tilde{\sigma}_i^2}\right),$$

with $\tilde{\sigma}_i^2 = p(\rho_1\sigma_1^2 + \rho_2\sigma_2^2)$.

V. SIMULATION RESULTS

Numerical results refer to Monte Carlo simulations with 10^5 runs using MATLAB. We considered sensor networks with $K = 5$ and $K = 10$ sensors, whose local performance are $(p_f, p_d) = (0.05, 0.5)$ unless differently specified.

Both Rice fading and 2ZM fading are considered. For both types of fading, we characterize the channels with respect to two parameters: (i) the ratio between the average power of the two components[3] (denoted ξ); (ii) the average total power (denoted ζ). More specifically, they are expressed as follows (left for Rice fading and right for 2ZM fading)

$$\begin{cases} \xi = \frac{|\mu|^2}{2\sigma^2} \\ \zeta = |\mu|^2 + 2\sigma^2 \end{cases}, \qquad \begin{cases} \xi = \frac{\rho_1\sigma_1^2}{\rho_2\sigma_2^2} \\ \zeta = 2\rho_1\sigma_1^2 + 2\rho_2\sigma_2^2 \end{cases}.$$

Also, only channels with unitary average power will be considered, i.e. $\zeta = 1$, both for Rice fading and 2ZM fading, and in the latter case with equally probable components ($\rho_1 = \rho_2 = 1/2$). ROC curves are labeled with respect to the SNR defined as $1/\sigma_w^2$.

Fig. 1 shows the (modulus) channel gain pdf for Rice fading and 2ZM fading with different ξ. It is apparent how increasing ξ makes the statistics more concentrated about the unit in the case of Rice fading, while more L-shaped with a peak close to zero in the case of 2ZM fading. Also, it is worth mentioning that the Rayleigh fading (in which the energy detector is the optimal receiver) is represented by Rice fading with $\xi = 0$ and by 2ZM fading with $\xi = 1$ ($\rho_1 = \rho_2 = 1/2$).

Fig. 2 shows the ROC curves for sensor networks at SNR$\in \{0, 5\}$ dB in the case of Rice fading and 2ZM fading. Solid and dashed lines refer to the analytical expressions, i.e. Eqs. (10) and (11), while circle and diamond markers refer to numerical simulations; a black asterisk represents the local performance. The improvement with respect to the SNR as well as the number of sensors (K) is apparent. Additionally, the impact of the fading statistics is reflected through the different shape of the ROC curves.

Fig. 3 shows the impact of the parameter ξ (in the case of Rice fading and 2ZM fading) on the performance for sensor networks at SNR= 5 dB and $q_f = 0.05$. It is apparent how, in the case of Rice fading, the presence of the line-of-sight component produces a beneficial effect in terms of absolute performance. However, it is worth noticing that the energy detector is optimal only in the case of $\xi = 0$ (Rayleigh fading). Differently, in the case of 2ZM fading, the absolute performance are maximum at $\xi = 1$ (Raylegh fading) when the energy detector is also optimal, and slowly decrease in other scenarios. Fig. 4 shows the impact of both ξ and K on the performance for both Rice fading and 2ZM fading. The effects of the channel statistics on system performance are apparent.

Finally, Figs. 5 and 6 show the validity of the asymptotic results in the case of Rice fading and 2ZM fading, respectively. In both cases $p_d = 0.5$, while $p_f = 0.3$ (resp. $p_f = 0.05$) for Rice fading (resp. 2ZM fading). It is worth noticing that the asymptotic performance: (i) in the case of Rice fading improve with K because of the *line-of-sight* and approach soon the ideal point $(q_f, q_d) = (0, 1)$; (ii) in the case of 2ZM fading are independent of K because of all components are zero mean. Also, the convergence to the asymptotic performance appears slower in the case of 2ZM fading than for Rice fading. Finally,

---

[3]In the case of Rice fading, the two components refer to the *line-of-sight* component (represented by the nonzero mean) and to the *non-line-of-sight* component (represented by the Gaussian-shaped random scattering). In the case of 2ZM fading, the two components refer to the two zero-mean components of the GM.
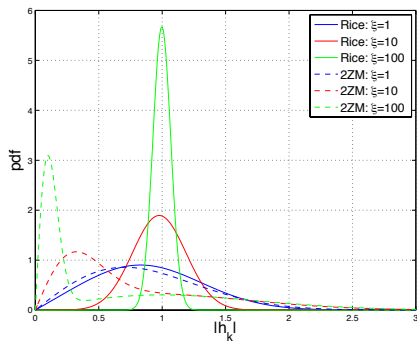
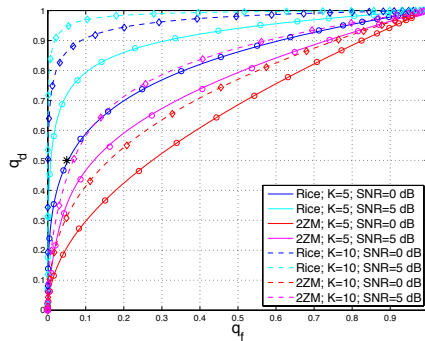Fig. 1. Impact of ξ on the pdf of $|h_k|$ for Rice and 2ZM fading.

Fig. 2. ROC for both Rice and 2ZM fading with $\xi = 10$. Lines and markers refer to analytical results and numerical simulations. The black asterisk denotes the local performance.
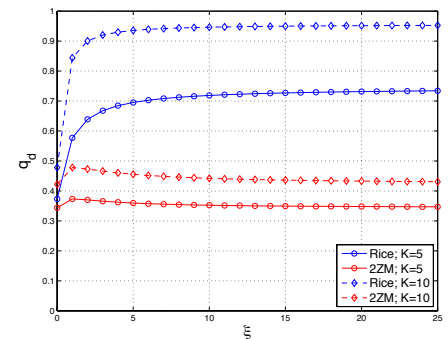

Fig. 3. Impact of ξ for Rice and 2ZM fading at SNR= 5 dB and $q_f = 0.05$ with $(p_f, p_d) = (0.05, 0.5)$.
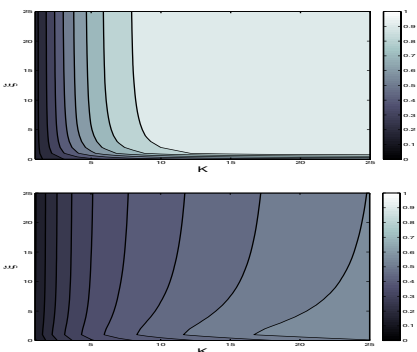

Fig. 4. Impact of ξ and K for both Rice (upper) and 2ZM (lower) fading at SNR= 5 dB and $q_f = 0.05$ with $(p_f, p_d) = (0.05, 0.5)$.
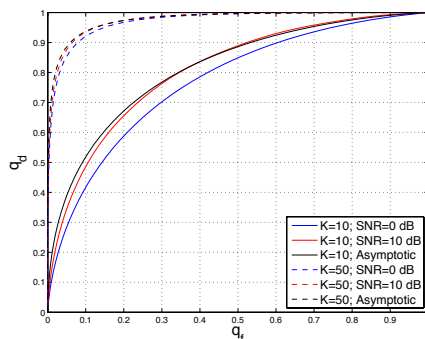

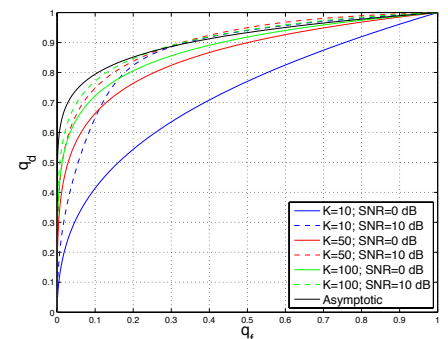Fig. 5. Asymptotic performance with $(p_f, p_d) = (0.3, 0.5)$ for Rice fading with $\xi = 10$.

Fig. 6. Asymptotic performance with $(p_f, p_d) = (0.05, 0.5)$ for 2ZM fading with $\xi = 10$.

it is worth noticing that asymptotic results, although obtained by neglecting the noise, are not upper bounds (apparently in both Figs. 5 and 6) because of the Gaussian approximation.

VI. CONCLUSION

Energy detector for decision fusion in wireless sensor networks over multiple access channels with Ricean-mixture fading (i.e. channel coefficients are GM distributed) has been analyzed. We derived analytical performance in terms of global probability of false alarm and detection. Analytical results, including asymptotic performance for large number of sensors, are confirmed through numerical simulations.

REFERENCES

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

[2] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Jul. 2010.

[3] 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.

[4] A. Lei and R. Schober, "Coherent max-log decision fusion in wireless sensor networks," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1327–1332, May 2010.

[5] 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.

[6] F. Li, J. S. Evans, and S. Dey, "Decision fusion over noncoherent fading multiaccess channels," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4367–4380, Sep. 2011.

[7] M. K. Banavar, A. D. Smith, C. Tepedelenlioglu, and A. Spanias, "On the effectiveness of multiple antennas in distributed detection over fading MACs," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1744–1752, May 2012.

[8] D. Ciuonzo, G. Romano, and P. Salvo Rossi, "Channel-aware decision fusion in distributed MIMO wireless sensor networks: decode-and-fuse vs. decode-then-fuse," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2976–2985, Aug. 2012.

[9] D. Ciuonzo, G. Romano, and P. Salvo Rossi, "Performance analysis and design of maximum ratio combining in channel-aware MIMO decision fusion," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4716–4728, Sept. 2013.

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

[11] J. T. Flåm, S. Chatterjee, K. Kansanen, and T. Ekman, "On MMSE estimation: a linear model under Gaussian mixture statistics," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3840–3845, Jul. 2012.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory.* Prentice Hall PTR, 1998.

[13] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.*, McGraw Hill, 2002.