Decision Fusion in MIMO Wireless Sensor Networks with Channel State Information

Domenico CIUONZO*, Gianmarco ROMANO*, Pierluigi SALVO ROSSI*
* Department of Information Engineering, Second University of Naples, Italy

***Abstract*—We study channel-aware binary decision fusion over a Rayleigh flat fading shared channel with multiple antennas at the Decision Fusion Center (DFC). We derive the optimum and three sub-optimal fusion rules, namely the Max-Log, the Chair-Varshney - Maximum Likelihood and the Maximum Ratio Combining rules. Simulation results for the performance are presented in terms of Probability of Detection (under a fixed false alarm rate) vs SNR and number of antennas. The effect of multiple antennas at the DFC for the presented rules is analyzed, showing benefits and limitations of the choice.**

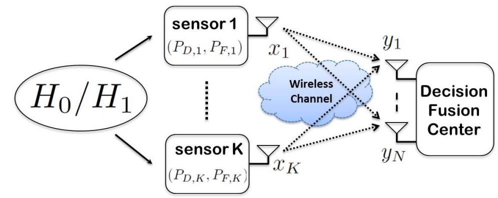

Figure 1. The Decision Fusion model in presence of a MIMO channel.

I. Introduction

In a wireless sensor network (WSN), Decision Fusion (DF) consists in transmitting local decisions about an observed phenomenon from individual sensors to a DF Center (DFC) for a final decision. The typical communication protocol between sensors and DFC is a parallel access channel (PAC), implemented through time, code or frequency division schemes. However the broadcast nature of the wireless medium can be exploited for DF as in [1].

Multiple antennas are employed at the fusion center in order to combat deep fading effects in [2], [3]. The result is a communication over a "virtual" Multiple-Input Multiple-Output (MIMO) channel between the sensors and the fusion center (see Fig. 1). The stringent assumption is the instantaneous channel state information (CSI) at the fusion center, which provides high performances via design of channel-aware fusion rules [4], [5], [6]. For this reason channel-aware fusion rules for coherent, non-coherent, and differential modulation were already proposed for PAC in [4], [5], [6], [7].

Unfortunately, the optimal DF rule over MIMO channels with instantaneous CSI presents several difficulties in the implementation: (i) complete knowledge of the channel parameters and sensors local performances; (ii) numerical instability of the formula, due to the presence of exponential functions with large dynamics; (iii) exponential growth of complexity with the number of sensors. This motivates the investigation of sub-optimal DF rules with simpler implementation and reduced system knowledge.

Sub-optimal rules for PAC scenario, presenting only the issues (i) and (ii), were designed in [4], [5], [6], [8]. More specifically optimal rule was compared to Maximum Ratio Combining (MRC), Chair-Varshney - Maximum Likelihood (CV-ML), Equal Gain Combining (EGC) and Max-Log. MRC and CV-ML fusion rules approach optimum performance at very low and very high channel SNRs, respectively and they both suffer from a significant performance loss at medium channel SNR [4]. EGC was shown to have robust performance for most SNR range [8]. Max-Log rule has been shown to outperform all the mentioned rules for all the SNR range [6]. The rules considered in [6] have also been derived and compared in the context of sensors differential encoding [5].

Zhang *et al.* [3] were the first to propose DF over MIMO channels, focusing on J-Divergence-optimal power allocation under non-identical local performances, which requires instantaneous CSI. DF rules over a matrix channel model, with only channel statistics knowledge and non-coherent modulation were studied in [1]. Distributed detection over MIMO with instantaneous CSI at the fusion center is tackled with the use of *amplify-and-forward* sensors in [2]; the optimum (data) fusion rule is derived and performance improvement is demonstrated when using multiple antennas at the fusion center.

In this paper we study channel-aware DF rules over MIMO channels, to best of our knowledge, for the first time. More specifically, we derive CV-ML, Max-Log and MRC fusion rules in this scenario (and suggest corresponding efficient implementations for the first two, through Sphere Decoder Algorithm), as appealing alternatives to the optimum in terms of performance-complexity tradeoff. We show that even the use of two antennas gives significative improvement to all the presented rules, letting them achieve the same performances as the single antenna case, but with a dramatic reduction of WSN energy consumption. Finally we illustrate that only a few antennas at the DFC are substantially beneficial as performance saturates with the number of antennas.

The paper is organized as follows: Sec. 2 introduces the system model; in Sec. 3 sub-optimal rules are derived in this scenario and implementation aspects are discussed; Sec. 4 shows performances via computer simulations in terms of system probability of detection (under a fixed false alarm rate) vs number of antennas and SNR; some concluding remarks are given in Sec. 5. *Notation*: Lower-case (resp. Upper-case) bold letters denote vectors (resp. matrices), with a_n (resp.

$a_{n,m}$) denoting the nth (resp. the (n,m)th) element of the vector $\boldsymbol{a}$ (resp. matrix $\boldsymbol{A}$); upper-case calligraphic letters denote discrete and finite sets, with $\mathcal{A}^K$ denoting the k-ary cartesian power of the set $\mathcal{A}$; $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix; $\mathbf{0}_N$ (resp. $\mathbf{1}_N$) denotes the null (resp. ones) vector of length N; $\mathbb{E}\{\cdot\}$, $(\cdot)^\dagger$, $\Re(\cdot)$, and $\|\cdot\|$ denote expectation, hermitian, real part, and Frobenius norm operators; $P(\cdot)$ and $p(\cdot)$ denote probabilities and probability density functions (pdf), with $P(A|B)$ and $p(a|b)$ their respective conditional counterparts; $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a circular symmetric complex normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; finally the symbols $\sim$, $\rightarrow$ and $\propto$ mean "distributed as", "maps to" and "proportional to" respectively.

II. System Model

We consider a distributed binary hypothesis testing task, where K sensors are used to discriminate between the hypotheses of the set $\mathcal{H} = \{H_0, H_1\}$, representing, e.g. (but not necessarily) the absence (H_0) or the presence (H_1) of a specific target of interest. The kth sensor, $k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$, takes a local binary decision $d_k \in \mathcal{H}$ about the observed phenomenon on the basis of its own measurements. The decision d_k is assumed independent on other decisions d_ℓ, $\ell \in \mathcal{K}, \ell \neq k$, conditioned on $H_i \in \mathcal{H}$. Each d_k is mapped to a symbol $x_k \in \mathcal{X} = \{-1, +1\}$ of a BPSK modulation[1]; w.l.o.g. we assume that $d_k = H_0 \rightarrow x_k = -1$ and $d_k = H_1 \rightarrow x_k = +1$. The quality of the kth sensor decisions is characterized by the conditional probabilities $P(x_k|H_i)$. More specifically, we denote $P_{D,k} \triangleq P(x_k = 1|H_1)$ and $P_{F,k} \triangleq P(x_k = 1|H_0)$, respectively the probability of detection and false alarm of the kth sensor. The sensors communicate with the DFC over a wireless flat-fading MAC, with i.i.d. Rayleigh fading coefficients of unitary mean power. The DFC is equipped with N receive antennas in order to exploit diversity and combat signal attenuation due to the wireless medium; this configuration determines basically a distributed (or "virtual" [3]) MIMO channel, as shown in Fig. 1. Also, instantaneous CSI and perfect synchronization are assumed at the DFC as in [3]; note that multiple antennas at the DFC do not make these assumptions harder to verify w.r.t. (single antenna) MAC. We denote: y_n the received discrete signal at the nth antenna of the DFC; $h_{n,k} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ the fading coefficient between the kth sensor and the nth antenna of the DFC; w_n the additive white Gaussian noise at the nth antenna of the DFC. The vector model at the DFC is the following:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{w} \tag{1}$$

where $\boldsymbol{y} \in \mathbb{C}^N$, $\boldsymbol{H} \in \mathbb{C}^{N \times K}$, $\boldsymbol{x} \in \mathcal{X}^K$, $\boldsymbol{w} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_N, \sigma_w^2 \boldsymbol{I}_N)$ are the received signal vector, the channel matrix, the transmitted signal vector and the noise vector, respectively. *Remarks:* The vector model in Eq. (1) can be underloaded ($K < N$), fully-loaded ($K = N$) or overloaded ($K > N$). Whether in MIMO communication systems all the three scenarios are of interest, in the specific case of WSN only the overloaded case is reasonable, as typically the number of sensors is larger than the number of antennas that could be employed at the DFC (i.e. $K > N$). Throughout this paper we will refer to the channel SNR as the ratio between the average total received energy from the WSN $\mathcal{E}_s = \mathbb{E}\{\|\boldsymbol{H}\boldsymbol{x}\|^2\}$ and the one-sided power spectral density of the continuous process noise σ_w^2, i.e. $SNR \triangleq \mathcal{E}_s/\sigma_w^2 = {}^{KN}/_{\sigma_w^2}$. Note that the corresponding channel SNR for the kth sensor is $SNR_k = {}^{N}/_{\sigma_w^2}$.

[1]Note that in an absence/presence task, where $\mathcal{H}_0$ is less probable, On-Off Keying (OOK) can be employed for energy efficiency purpose. In the following we refer only to BPSK, however the results presented in this paper apply straightforward to OOK.

III. Fusion Rules

Optimum Decision: The optimal test in Neyman-Pearson sense [9] for the considered problem can be formulated as

$$\Lambda_{opt} \triangleq \ln\left[\frac{p(\boldsymbol{y}|H_1)}{p(\boldsymbol{y}|H_0)}\right] \underset{\hat{H}=H_0}{\overset{\hat{H}=H_1}{\gtrless}} \gamma \tag{2}$$

where $\hat{H}$, Λ_{opt} and γ denote the estimated hypothesis, the Log-Likelihood Ratio (LLR, i.e. the optimal fusion rule, referred also as the "optimum" in the following) and the threshold chosen to assure a fixed system false-alarm rate, respectively. An explicit expression of the LLR from (2) is given by

$$\begin{aligned}\Lambda_{opt} &= \ln\left[\frac{\sum_{\mathbf{x}\in\mathcal{X}^K} p(\boldsymbol{y}|\boldsymbol{x}) \prod_{k=1}^K P(x_k|H_1)}{\sum_{\boldsymbol{x}\in\mathcal{X}^K} p(\boldsymbol{y}|\boldsymbol{x}) \prod_{k=1}^K P(x_k|H_0)}\right] \qquad (3)\\ &= \ln\left[\frac{\sum_{\mathbf{x}\in\mathcal{X}^\mathbf{K}} \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2}\right) \prod_{k=1}^K P(x_k|H_1)}{\sum_{\mathbf{x}\in\mathcal{X}^\mathbf{K}} \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2}\right) \prod_{k=1}^K P(x_k|H_0)}\right]\end{aligned}$$

where we have exploited the conditional independence among x_k (given H_i), and of $\boldsymbol{y}$ from H_i (given $\boldsymbol{x}$).

CV-ML Rule: In this case firstly an estimate of $\boldsymbol{x}$, denoted $\hat{\boldsymbol{x}}$ in the following, is computed with Maximum-Likelihood detector [10] from $\boldsymbol{y}$. Then the global decision $\hat{H}$ is taken on the basis of $\hat{\boldsymbol{x}}$ using the Chair-Varshney (CV) rule [11], i.e. the optimal fusion rule for noiseless channels. The expression of this two-stage fusion rule is given by

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathcal{X}^K} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 \tag{4}$$

$$\Lambda_{CV-ML} = \sum_{k=1}^K \hat{u}_k \ln\left(\tfrac{P_{D,k}}{P_{F,k}}\right) + (1 - \hat{u}_k) \ln\left(\tfrac{1-P_{D,k}}{1-P_{F,k}}\right) \tag{5}$$

where we denoted $\hat{u}_k \triangleq \frac{\hat{x}_k+1}{2}$, $k \in \mathcal{K}$. As in the PAC case [4] the CV-ML is the high-SNR approximation of the optimum of Eq. (3); the proof is reported in Appendix.

Max-Log Rule: Let us first recall the Max-Log approximation known from turbo-codes literature [12], and given by $\ln\left(\sum_{\ell=1}^L B_\ell e^{A_\ell}\right) \approx \max_{\ell\in\{1,2,\ldots,L\}}\{A_\ell + \ln(B_\ell)\}$, where $A_i \in \mathbb{R}$, $B_i \in \mathbb{R}^+$. This approximation is accurate when one of the terms in the sum $\sum_{\ell=1}^L B_\ell e^{A_\ell}$ dominates over the remaining terms. The expression of LLR from Eq. (3) has the same form and thus by using this approximation we obtain the

Table I
IMPLEMENTATION COMPARISON OF THE FUSION RULES.

Fusion Rule	required param.	complexity w.r.t. K - stability
Optimum	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$, σ_w^2	$\mathcal{O}(2^K)$ - unstable
CV-ML	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$	$\mathcal{O}(2^{K-n_1})$,$n_1 > 0$ - stable
Max-Log	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$, σ_w^2	$\mathcal{O}(2^{K-n_2})$,$n_2 > 0$ - stable
MRC	$\boldsymbol{H}\boldsymbol{1}_K$	$\mathcal{O}(1)$ - stable

following sub-optimal fusion rule:

$$\Lambda_{Max-Log} = \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_0)\right] - \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_1)\right] \quad (6)$$

which can be interpreted as the difference between *hypothesis prior-weighted* minimum distance searches.

MRC Rule: The LLR of Eq. (3) can be simplified under the assumption of perfect sensors [6], [13], i.e. $(P_{D,k}, P_{F,k}) = (1, 0)$, $k \in \mathcal{K}$. In this case the transmitted vector $\boldsymbol{x} \in \{\boldsymbol{1}_K, -\boldsymbol{1}_K\}$ and Eq. (3) reduces to:

$$\Lambda_{MRC} = \ln\left[\frac{\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}\right)}{\exp\left(-\frac{\|\boldsymbol{y}+\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}\right)}\right] \propto \Re(\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{1}_K) \quad (7)$$

where in the r.h.s. we have neglected the terms that can be incorporated in γ through the (2). As in the PAC case [4] the MRC is also the low-SNR approximation of the optimum of Eq. (3) when local performances of sensors are identical (i.e. $(P_{D,k}, P_{F,k}) = (P_D, P_F)$, $k \in \mathcal{K}$); again the proof is reported in Appendix.

Discussion on implementation: In the practice the LLR of Eq. (3) is difficult to compute as it contains exponential functions that have a large dynamic range especially for moderate-high channel SNRs $^{KN}/_{\sigma_w^2} \gg 1$; this becomes a quite severe requirement for fixed point implementations [5], [6], [13]. All the proposed sub-optimal rules instead present numerical stability, however they require a different degree of system knowledge and they also differ in computational complexity. In Tab. I we report a complete comparison of the aspects mentioned (note that CV-ML requires $(P_{D,k}, P_{F,k})$ only if sensors differ in local performances, cfr. Eq. (5)). Terms n_j, $j \in \{1, 2\}$, are inserted to underline that the *Exp*-complexity of CV-ML and Max-Log can be mitigated by implementing them through the Generalized Sphere Decoder (GSD) presented in [14]. In fact for CV-ML the equivalent problem $\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathcal{X}^K}\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2$ in place of Eq. (4) can be efficiently solved, with $\boldsymbol{D}$ denoting the upper-triangular matrix deriving from the Cholesky Factorization of $\boldsymbol{G} \triangleq \boldsymbol{H}^\dagger\boldsymbol{H} + \boldsymbol{I}_N$ (that is $\boldsymbol{G} = \boldsymbol{D}^\dagger\boldsymbol{D}$) and $\boldsymbol{\rho} \triangleq \boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{y}$. Instead the GSD implementation of Max-Log rule requires slight modifications to the steps followed in [14]. The steps, not reported here for sake of brevity, lead to

$$\Lambda_{Max-Log} = \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_0)\right] - \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_1)\right] \quad (8)$$

The computation of Eq. (8) can be easily performed through a double search with GSD (one for each hypothesis) or with a more efficient single search, following the same approach in [15]. In both cases the complexity of Max-Log is always higher than CV-ML, that is $n_1 > n_2$. Detailed results on the complexity reduction deriving from the GSD implementations of minimum distance searches can be found in [14].

IV. SIMULATION RESULTS

We compare the performances of the presented fusion rules in a WSN of $K = 8$ sensors with identical local performances $(P_{D,k}, P_{F,k}) = (P_D, P_F) \triangleq (0.5, 0.05)$, $k \in \mathcal{K}$, as adopted in [4], [6], [8] for rules comparison in PAC. The global performances are analyzed in terms of the system probabilities of false alarm and detection, defined respectively as $P_{F_0} \triangleq \Pr\{\Lambda > \gamma|H_0\}$ and $P_{D_0} \triangleq \Pr\{\Lambda > \gamma|H_1\}$, with Λ representing the decision statistics of a generic fusion rule.

P_{D_0} vs $(SNR)_{dB}$: In Fig. 2 we show the P_{D_0} as a function of the channel $(SNR)_{dB}$ (corresp. $(SNR_k)_{dB} \approx (SNR)_{dB} - 9$), under $P_{F_0} \leq 0.01$; we plot the cases $N \in \{1, 2\}$ to investigate the effect on performances when two antennas are employed at the DFC. In the same figure we also report the (upper) "observation bound" [1], i.e. the optimum performances over noiseless channel, given by $P_{D_0}^{obs} = \sum_{i=K_\gamma}^{K}\binom{K}{i}(P_D)^i(1-P_D)^{K-i}$, where K_γ is a discrete threshold. Firstly, numerical results confirm analytical derivations, i.e. CV-ML and MRC approach the optimum at high and low channel SNR, respectively, also in MIMO scenario (the "jumpy" behaviour of the CV-ML, given by the finite values assumed by Eq. (5), was already explained in [8] for the PAC case). Max-Log strictly approaches the same performances as the optimum over all the SNR range considered (i.e. $[0, 30]_{dB}$), but it requires complete system knowledge (cfr. Tab. I). All the rules significantly benefit from the presence of two antennas at DFC (cf. solid with dashed lines in Fig. 2). The Max-Log (as the optimum) has the best improvement in the range $[5, 20]_{dB}$ and reaches the observation bound at $(SNR)_{dB} \approx 20$, instead of $(SNR)_{dB} \approx 30$ when $N = 1$ at the DFC. CV-ML rule needs higher SNR to get acceptable performances, but the case $N = 2$ still needs less energy to reach the observation bound (in fact if $N = 1$ the bound is reached at $(SNR)_{dB} > 30$, not visible in Fig. 2). Finally multiple antennas not only increase MRC performances at low-medium SNR, but also give better limiting performances.

P_{D_0} vs N: In Fig. 3, we show the P_{D_0} as a function of the number of antennas N, under $P_{F_0} \leq 0.01$; we plot the cases $(SNR)_{dB} \in \{5, 15\}$ to investigate the performances when N increases under realistic channel SNR values. It is apparent

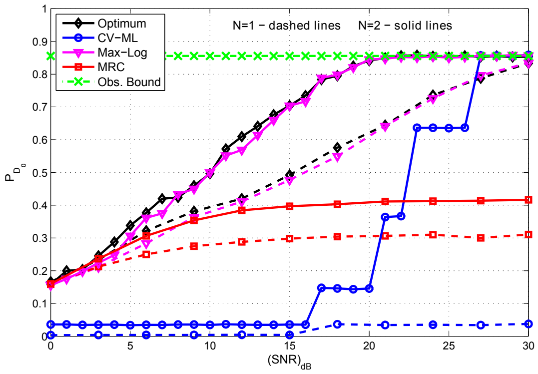

Figure 2. P_{D_0} vs channel $(SNR)_{dB}$; $P_{F_0} \leq 0.01$. WSN with $K = 8$ sensors, $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in \mathcal{K}$. $N \in \{1, 2\}$.

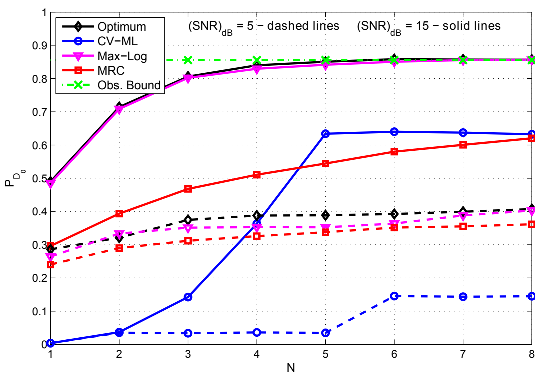

Figure 3. P_{D_0} vs N; $P_{F_0} \leq 0.01$. WSN with $K = 8$ sensors, $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in \mathcal{K}$. $(SNR)_{dB} \in \{5,15\}$.

that adding more antennas at the DFC is beneficial for all the rules presented, however a saturation effect is present. The saturation depends on the SNR and the fusion rule chosen; in particular specific configurations achieve the observation bound, e.g. Max-Log with $N = 4$ at $(SNR)_{dB} = 15$.

V. Conclusions

In this paper we addressed the design of sub-optimal fusion rules, more suitable for practical implementation than the optimum one, in a DF task over a virtual MIMO channel. The study was motivated by the need of multiple antennas at the DFC to obtain a dramatic increase of the performances with a reduced WSN energy budget. The proposed alternatives, Max-Log, CV-ML and MRC, solve the issues about fixed point implementations and present reduced requirements on complexity and partly on system knowledge. Furthermore they still greatly benefit from multiple antennas at the DFC.

Appendix

CV-ML: For small values of σ_w^2, if we denote the true transmitted vector as $\boldsymbol{x}_T$, the corresponding value in Eq. (3) will be a dominating term, thus the LLR is well approximated by Eq. (6). Also the following approximation holds:

$$\hat{\boldsymbol{x}}_i \triangleq \arg\min_{\boldsymbol{x} \in \mathcal{X}^K} \left[\frac{\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2} - \sum_{k=1}^{K} \ln P(x_k|H_i) \right] \approx \hat{\boldsymbol{x}} \quad (9)$$

where $\hat{\boldsymbol{x}}$ is given by Eq. (4), i.e. the second term becomes irrelevant. Thus, substituting Eq. (9) in Eq. (6), the LLR reduces to $\Lambda_{opt} \approx \ln \prod_{k=1}^{K} P(\hat{x}_k|H_1) - \ln \prod_{k=1}^{K} P(\hat{x}_k|H_0)$, which can be rearranged easily to obtain Eq. (5).

MRC: Starting from Eq. (3) and observing that for high values of σ_w^2, $\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2}\right) \approx \exp\left(-\frac{\|\boldsymbol{y}\|^2}{\sigma_w^2}\right)\left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right)$, we get

$$\Lambda_{opt} \approx \ln \left[\frac{\sum_{\mathbf{x} \in \mathcal{X}^{\mathbf{K}}} \left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right) \prod_{k=1}^{K} P(x_k|H_1)}{\sum_{\mathbf{x} \in \mathcal{X}^{\mathbf{K}}} \left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right) \prod_{k=1}^{K} P(x_k|H_0)} \right] \quad (10)$$

Exploiting the property $\sum_{\mathbf{x} \in \mathcal{X}^{\mathbf{K}}} P(\boldsymbol{x}|H_i) = 1$, and using the approximation $\ln(1 + x) \approx x$, when $x \ll 1$, we obtain:

$$\Lambda_{opt} \approx \frac{2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H} \left(\mathbb{E}\{\boldsymbol{x}|H_1\} - \mathbb{E}\{\boldsymbol{x}|H_0\}\right)\}}{\sigma_w^2} + \alpha \quad (11)$$

where α is a term not depending on $\boldsymbol{y}$. When local performances are identical, $\mathbb{E}\{\boldsymbol{x}|H_1\} = \mathbf{1}_K(2P_D - 1)$ and $\mathbb{E}\{\boldsymbol{x}|H_0\} = \mathbf{1}_K(2P_F - 1)$, which gives Eq. (7), except for α (irrelevant in Eq. (2)).

References

[1] 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, Sept. 2009.

[2] M. K. Banavar, A. D. Smith, C. Tepedelenlioglu, and A. Spanias, "Distributed detection over fading MACs with multiple antennas at the fusion center," in *IEEE ICASSP*, Mar. 2010, pp. 2894–2897.

[3] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124–4140, Sept. 2008.

[4] 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.

[5] A. Lei and R. Schober, "Multiple-symbol differential decision fusion for mobile wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 778–790, Feb. 2010.

[6] 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.

[7] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2668–2673, Nov. 2005.

[8] 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.

[9] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag New York, Inc., 1st edition, 1996.

[10] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press, 2007.

[11] 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.

[12] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *IEEE ICC*, June 1995, vol. 2, pp. 1009–1013.

[13] S. Yiu and R. Schober, "Nonorthogonal transmission and noncoherent fusion of censored decisions," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 263–273, Jan. 2009.

[14] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 423–425, May 2005.

[15] C. Studer and H. Bölcskei, "Soft-input soft-output single tree-search sphere decoding," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4827–4842, 2010.