Channel-Aware Decision Fusion in Distributed MIMO Wireless Sensor Networks: Decode-and-Fuse vs. Decode-then-Fuse

Domenico Ciuonzo, *Student Member, IEEE,* Gianmarco Romano, *Member, IEEE,*
and Pierluigi Salvo Rossi, *Senior Member, IEEE*

***Abstract*—We study channel-aware binary-decision fusion over a shared Rayleigh flat-fading channel with multiple antennas at the Decision Fusion Center (DFC). We present the optimal rule and derive sub-optimal fusion rules, as alternatives with improved numerical stability, reduced complexity and lower system knowledge required. The set of rules is derived following both "Decode-and-Fuse" and "Decode-then-Fuse" approaches. Simulation results for performances are presented both under Neyman-Pearson and Bayesian frameworks. The effect of multiple antennas at the DFC for the presented rules is analyzed, showing corresponding benefits and limitations. Also, the effect on performances as a function of the number of sensors is studied under a total power constraint.**

***Index Terms*—Decision fusion, distributed detection, MIMO, wireless sensor networks.**

I. INTRODUCTION

A. Motivation

DECISION Fusion (DF) in a wireless sensor network (WSN) consists in transmitting local decisions about an observed phenomenon from individual sensors to a DF Center (DFC) for a final decision. For sake of simplicity, the usual architecture assumes that each sensor communicates through a parallel access channel (PAC), which has to be implemented through time, code or frequency division schemes, since the wireless channel is "naturally" a broadcast medium [1]–[4].

Recently it has been suggested to exploit the wireless medium as a multiple-access channel (MAC) for DF while coping with presence of intrinsic interference [5]. Furthermore, the problem of deep fading was addressed with multiple antennas at the DFC in order to ameliorate the fusion performances [5]–[7]. This choice demands only further complexity on DFC side and does not affect simplicity of sensors implementation. The result is a communication over a "virtual" Multiple-Input Multiple-Output (MIMO) channel between the sensors and the DFC, as shown in Fig. 1. Though appearing a stringent requirement, instantaneous channel state information (CSI) is often assumed [6], [7]. Design of channel-aware fusion rules and corresponding high performances motivates this assumption [2], [4], [8], [9], fulfilled in many scenarios[1]. For the same reason, channel-aware fusion rules for coherent, non-coherent, and differential modulation were already proposed for PAC [2], [4], [8]–[10].

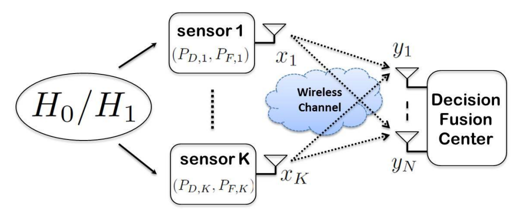


Fig. 1: The Decision Fusion model in presence of a (virtual) MIMO channel.

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

B. Related Work

A vast literature is present on DF, still growing in the context of WSNs. Several tutorial papers and books have been published on the topic, providing extensive references [3], [11]–[13]. Here we briefly discuss recent related work and focus on results needed for the work in this paper.

Sub-optimal rules for PAC scenario, presenting only issues (i) and (ii), were designed in [1], [2], [4], [8]–[10], [14]. More specifically, the 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

Manuscript received November 17, 2011; revised February 27, 2012; accepted May 17, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was G. Song.

The material in this paper was presented in part at the 2012 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'12), Stevens Institute of Technology, Hoboken (NJ), US, June, 2012.

The authors are with the Department of Information Engineering, Second University of Naples, via Roma, 29, 81031 Aversa (CE), Italy (e-mail: {domenico.ciuonzo, gianmarco.romano, pierluigi.salvorossi}@unina2.it).

Digital Object Identifier 10.1109/TWC.2012.061912.112049

[1]We focus on the case in which the sensors and the DFC have minimal movement and the environment changes slowly. More precisely, the coherence time of the wireless channel is assumed much longer than the time interval between two consecutive decisions made by the DFC, and instantaneous CSI can be obtained [7].

both suffer from a significant performance loss at moderate channel SNR [1], [2]. Instead EGC was shown to have robust performance for most SNR range [14]. Max-Log rule has been shown to outperform all the mentioned rules for all the SNR range [9]. Furthermore, robust sub-optimal fusion rules were obtained by fitting various non-linearities to the Log-Likelihood-Ratio (LLR) in [4]. The rules considered in [9] have also been derived and compared in the context of sensors differential encoding [8].

DF in distributed detection over MAC was first considered in [15], where an algorithm for the design of local-sensor quantizers was derived for the optimal fusion rule. Also, several advanced transmission schemes have been proposed for MAC. In [16], [17] a fully-loaded code-division multiple access (CDMA) communication between sensors and the DFC is proposed: DF exploiting instantaneous CSI is performed and sub-optimal alternatives, based on the partitioning of transmitted symbols decoding and DF rule, are compared with the optimum. Also, in [18] a CDMA protocol is considered, but no decision is performed locally by the sensors, since they employ an *amplify-and-forward* scheme. In [19], the idea of the direct sequence spread spectrum is exploited to perform a communication based on the On-Off Keying (OOK) modulation with censoring; at the DFC the optimal rule and sub-optimal alternatives are derived on the basis only of statistical CSI.

An alternative and novel communication scheme, based on the method of types, called Type-Based Multiple Access (TBMA), has been proposed for distributed detection in [20], [21]. The scheme relies only on statistical CSI and presents limited bandwidth requirements. However, it has been shown in [22] that TBMA suffers from significant loss of performances in i.i.d. zero-mean fading channels.

DF over MIMO channels was firstly proposed in [7], 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 statistical CSI knowledge and non-coherent modulation have been studied in [5]. Distributed detection over MIMO with instantaneous CSI at the fusion center is tackled with the use of *amplify-and-forward* sensors in [6]; the optimum (data) fusion rule is derived and performance improvement is demonstrated when using multiple antennas at the fusion center.

C. Main Results and Paper Organization

The main contributions of the paper are summarized as follows.

- We study the design of channel-aware DF rules over MIMO channels, to best of our knowledge, for the first time. We discuss advantages and drawbacks for each rule in terms of *complexity*, *system knowledge required* and *performances*. The rules are grouped under *Decode-and-Fuse* (DaF) and *Decode-then-Fuse* (DtF) approaches. In the former case fusion is performed on the received signal, while in the latter case fusion is performed through CV rule, which processes ML or Minimum Mean Square Error (MMSE) estimate of the transmitted symbols. Multiple antennas at the DFC are shown to be beneficial independently of the considered (optimal or sub-optimal) fusion rule, however a (rule-dependent) saturation effect is present.
- Optimality properties are analytically demonstrated (in low and/or high SNR regime) for CV-ML, Max-Log and MRC over MIMO channel, in perfect analogy to the PAC case. Also, we formulate efficient Generalized Sphere Decoder (GSD) [23] implementation to tackle exponential complexity of the first two rules.
- We show that CV-MMSE outperforms CV-ML in low-to-moderate SNR regime, since it accounts for the correlation among decisions in the decoding stage. Furthermore, the co-channel interference of the MIMO channel makes MRC to outperform EGC under a Neyman-Pearson framework, in *disagreement* with PAC case [2].
- We show that each discussed rule, except for MRC, present an optimal number of sensors to be employed, under a total power constraint, to minimize the system probability of error. Further addition of sensors in the WSN *counter-intuitively* degrades the overall performances. Instead the MRC benefits from an indefinite increase of the number of sensors employed.

The paper is organized as follows: Section II introduces the system model; Section III presents fusion rules based on DaF (including the optimum fusion rule) approach, while Section IV presents fusion rules based on DtF approach; Section VI presents an extensive set of simulations for performance comparison under different scenarios; some concluding remarks are given in Section VII; proofs and derivations are confined to Appendices.

Notation - Lower-case (resp. Upper-case) bold letters denote vectors (resp. matrices), with a_n (resp. $a_{n,m}$) representing 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 $\mathcal{A}^K$ representing 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; $\mathbb{E}\{\cdot\}$, $(\cdot)^*$, $(\cdot)^t$, $(\cdot)^\dagger$, $\Re(\cdot)$, $\angle(\cdot)$ and $\|\cdot\|$ denote expectation, conjugate, transpose, conjugate transpose, real part, phase and Frobenius norm operators; $P(\cdot)$ and $p(\cdot)$ are used to denote probabilities and probability density functions (pdf), in particular $P(A|B)$ and $p(a|b)$ represent the probability of event A conditioned on event B and the pdf of random variable a conditioned on random variable b, respectively; $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a circular symmetric complex normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; finally the symbols $\sim$ and $\propto$ mean "distributed as" and "proportional to" respectively.

II. System Model

We consider a distributed binary hypothesis test, where K sensors are used to discriminate between the hypotheses of the set $\mathcal{H} = \{H_0, H_1\}$. For example H_0 and H_1 may represent the absence and the presence of a specific target of interest, respectively. The *a priori* probability of hypothesis $H_i \in \mathcal{H}$ is denoted $P(H_i)$. The kth sensor, $k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$, takes a binary local decision $d_k \in \mathcal{H}$ about the observed phenomenon on the basis of its own measurements. The decision d_k is assumed independent of each other decisions d_ℓ, $\ell \in \mathcal{K}, \ell \neq k$, conditioned on $H_i \in \mathcal{H}$.

Each decision d_k is mapped to a symbol $x_k \in \mathcal{X} = \{-1, +1\}$ representing BPSK modulation[2]: without loss of generality we assume that $d_k = H_0$ maps into $x_k = -1$ and $d_k = H_1$ into $x_k = +1$. The quality of the kth sensor decisions is characterized by the conditional probabilities $P(x_k|H_j)$. 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 small-scale fading of the wireless medium; this configuration determines basically a distributed (or "virtual" [7]) MIMO channel, as shown in Fig. 1. Also, instantaneous CSI and perfect synchronization are assumed at the DFC as in [7]; 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 signal at the nth receive antenna of the DFC after matched filtering and sampling; $h_{n,k} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ the fading coefficient between the kth sensor and the nth receive antenna of the DFC; w_n the additive white Gaussian noise at the nth receive antenna of the DFC. The vector model at the DFC is the following:

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

where $\boldsymbol{y} \in \mathbb{C}^N$, $\boldsymbol{H} \in \mathbb{C}^{N\times K}$, $\boldsymbol{x} \in \mathcal{X}^K$, $\boldsymbol{w} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}_N, \sigma_w^2 \boldsymbol{I}_N)$ 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 ($K > N$). While 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 typically much larger than the number of antennas that could be employed at the DFC (i.e. $K \gg N$). Throughout this paper we will refer to the channel SNR as the ratio between the average total received energy from the WSN $\mathcal{E}_s = \mathbb{E}\left\{\|\boldsymbol{H}\boldsymbol{x}\|^2\right\}$ and the one-sided power spectral density of the continuous process noise σ_w^2, that is $SNR \triangleq \mathcal{E}_s/\sigma_w^2 = {}^{KN}/_{\sigma_w^2}$. Note that the corresponding individual channel SNR for the kth sensor will be $SNR_k = {}^{N}/_{\sigma_w^2}$.

III. Decode-and-Fuse

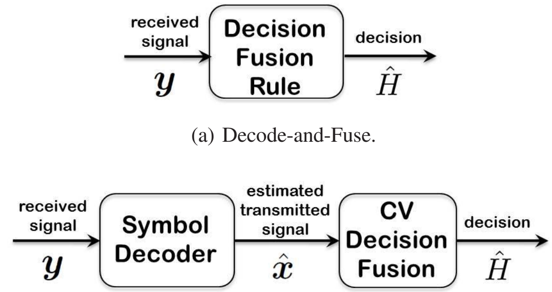
In this case, see Fig. 2a, the DFC aims at detecting the presence of the target directly from the received signal vector without any intermediate step to decode the transmitted vector.

A. Optimum Rule

The optimal test [24] for the considered problem can be formulated as

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

[2]Note that in case of 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 will refer only to BPSK, however the results presented in this paper apply straightforward to OOK.



(a) Decode-and-Fuse.

(b) Decode-then-Fuse.

Fig. 2: Fusion approaches.

where $\hat{H}$, Λ_{opt} and γ denote the estimated hypothesis, the LLR (i.e. the optimal fusion rule, referred also as the "optimum" in the following) and the threshold to which the LLR is compared to. The threshold γ can be determined to assure a fixed system false-alarm rate, if a Neyman-Pearson detection is employed, or can be chosen to minimize the probability of error in Bayes detection. An explicit expression of the LLR from Eq. (2) is given by

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

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

B. Maximum Ratio Combining (MRC)

The LLR of Eq. (3) can be simplified under the assumption of perfect sensors [9], [19], 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 the Eq. (3) reduces to:

$$\Lambda_{MRC} = \ln\left[\frac{\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}\right)}{\exp\left(-\frac{\|\boldsymbol{y}+\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}\right)}\right] \propto \Re(\boldsymbol{1}_K^t \boldsymbol{H}^\dagger \boldsymbol{y}) \tag{4}$$

where in the r.h.s. we have neglected the terms that can be incorporated in γ through Eq. (2). The following proposition states that as in PAC case [2] the MRC is the low-SNR approximation of the optimum of Eq. (3) when local performances of sensors are identical.

Proposition 1. *For low SNR, if* $(P_{D,k}, P_{F,k}) = (P_D, P_F)$, $k \in \mathcal{K}$*:* $\Lambda_{MRC} \approx \Lambda_{opt}$.

Proof: See Appendix A. ∎

C. Equal Gain Combining (EGC)

Motivated by the fact that Λ_{MRC} resembles a MRC statistics for diversity combining [25], cfr. Eq. (4), we propose a further rule in the simple form of an equal gain combiner:

$$\Lambda_{EGC} = \Re(\boldsymbol{z}^\dagger \boldsymbol{y}) \tag{5}$$

$$\boldsymbol{z} = e^{j\cdot\angle(\boldsymbol{H}\boldsymbol{1}_K)} \tag{6}$$

A similar expression was derived for MIMO beamforming and combining systems in [26].

D. Max-Log Rule

Let us first recall the Max-Log approximation known from turbo-codes literature [27]

$$\ln\left(\sum_{\ell=1}^{L} B_\ell e^{A_\ell}\right) = \ln\left(\sum_{\ell=1}^{L} e^{A_\ell+\ln(B_\ell)}\right) \approx \max_{\ell\in\{1,2,\ldots,L\}}[A_\ell+\ln(B_\ell)] \quad (7)$$

where $A_i \in \mathbb{R}$ and $B_i \in \mathbb{R}^+$. The approximation in Eq. (7) is accurate when one of the terms in the sum $\sum_{\ell=1}^{L} B_\ell e^{A_\ell}$ dominates over the remaining terms. LLR expression from Eq. (3) is in the same form of Eq. (7), thus using this approximation we obtain the following sub-optimal fusion rule:

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

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

Proposition 2. *Max-Log Properties:*

1) *For low SNR, if* $(P_{D,k} > \frac{1}{2}, P_{F,k} < \frac{1}{2}) : \Lambda_{Max-Log} \approx \Lambda_{MRC}$.
2) *For low SNR, if* $(P_{D,k}, P_{F,k}) = (P_D, P_F)$, $k \in \mathcal{K}$, *and* $(P_D > \frac{1}{2}, P_F < \frac{1}{2}) : \Lambda_{Max-Log} \approx \Lambda_{opt}$.
3) *For high SNR* : $\Lambda_{Max-Log} \approx \Lambda_{opt}$.

Proof: See Appendix B. ■

The above proposition states that under particular circumstances Max-Log approximates the optimum both in the low and high SNR regime.

IV. Decode-then-Fuse

In this case, see Fig. 2b, the DFC is based on the separation of decoding and fusing stages. Firstly, the decoding block computes an estimate of $\boldsymbol{x}$, denoted $\hat{\boldsymbol{x}}$ in the following, from $\boldsymbol{y}$. Finally, the global decision $\hat{H}$ is taken on the basis of $\hat{\boldsymbol{x}}$ using the Chair-Varshney (CV) rule [28], i.e. the optimal fusion rule for noiseless channels, whose expression is given by

$$\Lambda_{CV} = \sum_{k=1}^{K} \hat{u}_k \ln\left(\frac{P_{D,k}}{P_{F,k}}\right) + (1-\hat{u}_k)\ln\left(\frac{1-P_{D,k}}{1-P_{F,k}}\right), \quad (9)$$

where $\hat{u}_k \triangleq \frac{\hat{x}_k+1}{2}$, $k \in \mathcal{K}$. Note that when local sensor performances are identical, i.e. $(P_{D,k}, P_{F,k}) = (P_D, P_F)$, Eq. (9) reduces to a simple counting rule [28].

A. CV-ML

In this case $\hat{\boldsymbol{x}}$ is obtained through ML decoder [25] as

$$\hat{\boldsymbol{x}}_{ML} = \arg\min_{\boldsymbol{x}\in\mathcal{X}^K}\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2. \quad (10)$$

In analogy to PAC case [2], CV-ML is the high-SNR approximation of the optimum of Eq. (3), as stated through the following proposition.

Proposition 3. *For high SNR:* $\Lambda_{CV-ML} \approx \Lambda_{opt}$.

Proof: See Appendix C. ■

B. CV-MMSE

In this case the ML decoder is replaced with a sub-optimal one presenting reduced complexity, obtained via the MMSE solution [25]. The rule however needs to take into account the correlation between symbols x_k, $k \in \mathcal{K}$, since they observe the same phenomenon. This issue was addressed in [17] but restricted to the case $P_{D,k} = 1 - P_{F,k}$. In the more general case, the following MMSE decoder should be considered [29]:

$$\hat{\boldsymbol{x}}_{MMSE} = \mathrm{sign}\left[\overline{\boldsymbol{x}} + \boldsymbol{C}\boldsymbol{H}^\dagger(\boldsymbol{H}\boldsymbol{C}\boldsymbol{H}^\dagger + \sigma_w^2\boldsymbol{I}_N)^{-1}(\boldsymbol{y}-\boldsymbol{H}\overline{\boldsymbol{x}})\right] \quad (11)$$

where $\overline{\boldsymbol{x}} \triangleq \mathbb{E}\{\boldsymbol{x}\}$ and $\boldsymbol{C} \triangleq \mathbb{E}\{(\boldsymbol{x}-\overline{\boldsymbol{x}})(\boldsymbol{x}-\overline{\boldsymbol{x}})^\dagger\}$ are the mean vector and the covariance matrix of $\boldsymbol{x}$, respectively. Their explicit expression is derived and reported in Appendix D.

V. Implementation Issues

Practical implementation of LLR in Eq. (3) is difficult due to exponential functions with large dynamic range especially for moderate-to-high channel SNRs $^{KN}/_{\sigma_w^2} \gg 1$. This becomes a quite severe requirement for fixed point implementations [8], [9], [19]. All the proposed sub-optimal rules instead present numerical stability for realistic SNR values, although 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. More specifically, for each fusion rule, we report the system parameters required for implementation (first column), the complexity w.r.t. both the number of sensors K and antennas N (second column), numerical stability (third column) and finally optimality properties under the assumption of identical sensors performance (last column). Also, it is apparent that the fusion rules requiring less a priori information and lower computational complexity are EGC and MRC, while Max-Log fusion rule is the only one exhibiting optimality at both low and high SNR.

Note that CV-ML requires $(P_{D,k}, P_{F,k})$, $k \in \mathcal{K}$, only if local performances of sensors are different, cfr. Eq. (9); this is not the case for CV-MMSE because local sensor performances are always required in the decoding stage of Eq. (11). Finally, MRC and EGC require only reduced knowledge of channel matrix.

The dependence of the complexity with respect to N is moderate for all the presented rules and it is polynomial in

TABLE I: Comparison of the fusion rules.

Fusion Rule	required parameters	complexity	stability	optimality
Optimum	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$, σ_w^2	$\mathcal{O}(2^K N)$	no	*always*
MRC	$\boldsymbol{H}\boldsymbol{1}_K$	$\mathcal{O}(N)$	yes	*low SNR*, if $P_D > P_F$
EGC	$\angle(\boldsymbol{H}\boldsymbol{1}_K)$	$\mathcal{O}(N)$	yes	*never*
Max-Log	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$, σ_w^2	$\mathcal{O}(2^{(K-n_2)}N)$,$n_2 > 0$	yes	*low SNR*, if $(P_D > \frac{1}{2}, P_F < \frac{1}{2})$ - *high SNR*
CV-ML	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$	$\mathcal{O}(2^{(K-n_1)}N)$,$n_1 > 0$	yes	*high SNR*
CV-MMSE	$(P_{D,k}, P_{F,k})$, $\boldsymbol{H}$, σ_w^2	$\mathcal{O}(K(K+NK+N^2))$	yes	*never*

the worst-case (CV-MMSE[3]). This justifies the deployment of multiple antennas at DFC at the expenses of a slightly increase in the complexity burden. The dominant term of complexity for all the rules depends on the number of sensors K. The only exceptions are MRC and EGC, whose computational complexity, given by Eqs. (4) and (5), does not depend on the number of sensors K. In such a case, the dependence is only linear with respect to number of antennas, since during channel estimation step only the N-dimensional vector $\angle\boldsymbol{H}\boldsymbol{1}_K$ (corresp. $\boldsymbol{H}\boldsymbol{1}_K$), needs to be estimated.

Terms n_j, $j \in \{1, 2\}$, are inserted to underline that the *Exp*-complexity (with respect to K) of CV-ML and Max-Log can be mitigated by a GSD implementation [23]. 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. (10) can be efficiently solved, with $\boldsymbol{D}$ denoting the upper-triangular matrix deriving from the Cholesky Factorization of $\boldsymbol{G} \triangleq \boldsymbol{H}^\dagger\boldsymbol{H} + \beta\boldsymbol{I}_N$ (that is $\boldsymbol{G} = \boldsymbol{D}^\dagger\boldsymbol{D}$) and $\boldsymbol{\rho} \triangleq \boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{y}$.

Differently, GSD implementation of Max-Log rule requires slight modifications to the steps followed in [23]. The steps, reported in Appendix E, lead to

$$\Lambda_{Max-Log} = \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_0)\right] - \min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2} - \sum_{k=1}^{K}\ln P(x_k|H_1)\right] \quad (12)$$

The computation of Eq. (12) 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 [30]. 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 [23].

VI. Simulation Results

In this section we compare the performances of the presented fusion rules in a WSN with sensors of identical local performances $(P_{D,k}, P_{F,k}) = (P_D, P_F)$, $k \in \mathcal{K}$. Unless differently stated, we assume $(P_D, P_F) = (0.5, 0.05)$ as adopted in [2], [4], [9], [14] for fusion rules comparison in PAC. The global performances are analyzed in terms of system probabilities of false alarm, detection and error, defined respectively as

$$P_{F_0} \triangleq P(\Lambda > \gamma|H_0), \quad P_{D_0} \triangleq P(\Lambda > \gamma|H_1), \quad (13)$$

$$P_{E_0} \triangleq \min_{\gamma}\left\{[1 - P_{D_0}(\gamma)]\,P(H_1) + P_{F_0}(\gamma)P(H_0)\right\}, \quad (14)$$

with Λ representing the decision statistics of a generic fusion rule. In the following figures, for comparison purposes, we report the (upper) "observation bound" [5], i.e. the optimum performances over noiseless channel, given by:

$$P_{D_0}^{obs} = \sum_{i=K_\gamma}^{K}\binom{K}{i}(P_D)^i(1-P_D)^{K-i}, \quad (15)$$

$$P_{F_0}^{obs} = \sum_{i=K_\gamma}^{K}\binom{K}{i}(P_F)^i(1-P_F)^{K-i}. \quad (16)$$

where K_γ is a discrete threshold.

Receiver Operating Characteristic (ROC): In Fig. 3 we show the ROC (i.e. P_{D_0} vs P_{F_0}), for the presented rules in a WSN with $K = 8$ sensors and $N = 2$ antennas at the DFC, under a channel $(SNR)_{dB} = 15$ (corresp. $(SNR_k)_{dB} \approx 6$). It is apparent that Max-Log and Optimum ROCs are quite similar, but far from the observation bound. Instead ROCs of the MRC and EGC present a crossing point; the same happens between CV-ML and CV-MMSE. However while in the first case the result is independent of the specific channel SNR, in the latter case it depends on the poor performances of CV-ML statistics, due to the low channel SNR. This implies that when a fixed (low) P_{F_0} is imposed, as in typical Neyman-Pearson test [24], the MRC is a more attractive choice than EGC, in juxtaposition with the PAC case [2], [3]. The difference comes from the impossibility of vector $\boldsymbol{z}$ in Eq. (5) to sum the contributions of all the sensors coherently.

P_{D_0} vs $(SNR)_{dB}$: In Fig. 4 we show, for the presented rules, P_{D_0} as a function of the channel $(SNR)_{dB}$, under $P_{F_0} \leq 0.01$, in a WSN with $K = 8$ sensors (corresp. $(SNR_k)_{dB} = (SNR)_{dB} - 10\log_{10}K \approx (SNR)_{dB} - 9$); we plot the cases $N \in \{1, 2\}$ to investigate the effect on performances when two antennas are employed at the DFC. 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. Max-Log is optimal at high SNR and also it strictly approaches the same performances as the optimum over all the SNR range considered (i.e. $[0, 30]_{dB}$); the pay-off is a high requirement on system knowledge and computational complexity (cfr. Tab.

[3]It is worth remarking that CV-MMSE complexity in Tab. I has been derived under the assumption $K \gg N$.

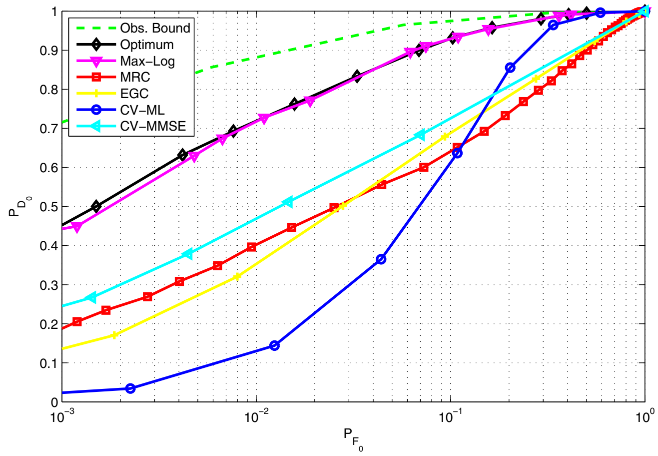


Fig. 3: ROC for the presented rules. WSN with $K = 8$ sensors, $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in \mathcal{K}$. $N = 2$, $(SNR)_{dB} = 15$.

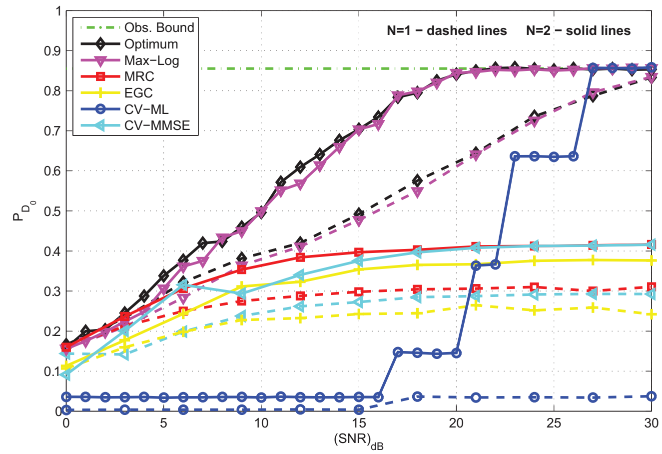


Fig. 4: P_{D_0} vs channel $(SNR)_{dB}$ for the presented rules; $P_{F_0} \leq 0.01$. WSN with $K = 8$ sensors, $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in \mathcal{K}$. $N \in \{1, 2\}$.

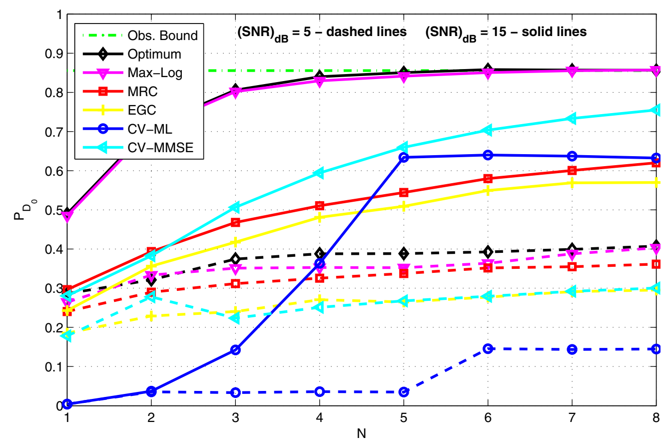


Fig. 5: P_{D_0} vs N for the presented rules; $P_{F_0} \leq 0.01$. WSN with $K = 8$ sensors, $(P_{D,k}, P_{F,k}) = (0.5, 0.05)$, $k \in \mathcal{K}$. $(SNR)_{dB} \in \{5,15\}$.

I). It is worth noticing that curves corresponding to DtF rules in Fig. 4 exhibit jumpy-step and non-monotonic behaviors in the case of CV-ML (as pointed out in [14] for PAC) and CV-MMSE rules, respectively. Such phenomena are not surprising when related to the discrete nature of CV decision statistics as well as the operation point on the corresponding ROC. More specifically, the operation point does NOT show a fixed probability of false alarm with respect to SNR, and both probabilities of detection and false alarm are sometimes lowered in order to meet the constraint on the maximum allowed probability of false alarm. Finally CV-MMSE performs fairly better than CV-ML at low-medium SNR, as it exploits the local sensor performances information in the decoding stage. All the rules significantly benefit from the presence of two antennas at DFC (cfr. solid with dashed lines in Fig. 4). 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. 4). Finally multiple antennas not only increase MRC, EGC and CV-MMSE performances at low-medium SNR, but also give better limiting performances.

P_{D_0} *vs* N: In Fig. 5, we show, for the presented rules, 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 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 chosen fusion rule. In particular, specific configurations achieve the observation bound (e.g. Max-Log with $N = 4$ at $(SNR)_{dB} = 15$) while others (e.g. MRC with $N = 6$ at $(SNR)_{dB} = 5$) exploit all the diversity gain.

P_{E_0} *vs* K*:* In Figs. 6 and 7 we show, for the presented rules, the system probability of error P_{E_0} (under the assumption $P(H_i) = 1/2$, $H_i \in \mathcal{H}$) as a function of the number of sensors K; we plot the case $(P_D, P_F) \triangleq (0.7, 0.05)$ to emphasize the results observed. We consider a WSN with $N = 1, 2$ (corresp. Fig. 6 and 7) antennas at DFC and a channel $(SNR)_{dB} = 15$. The latter assumption clearly represents a *total power constraint* (TPC) [18], [31] on the WSN, i.e. the individual channel SNR for the kth sensor is scaled as $SNR_k = \frac{SNR}{K}$. We also report the (upper) "communication" bound [5], i.e. P_{E_0} under the assumption of ideal sensors, i.e. $(P_D, P_F) = (1, 0)$. In this case the bound is represented by the symbol error probability of a N branch MRC combiner [32], that is $P_{E_0}^{comm} = \left(\frac{1-\mu}{2}\right)^N \sum_{l=0}^{N-1} \binom{N-1+l}{l} \left(\frac{1+\mu}{2}\right)^l$, where $\mu \triangleq \sqrt{\frac{SNR}{SNR+N}}$. Simulations show that all the rules, except for MRC, present a unimodal behaviour, i.e. there exists a finite number of sensors, under a fixed $(SNR)_{dB}$ value, which minimizes P_{E_0}. This implies that when fusing decisions using those rules, less sensors with a higher individual battery energy are more beneficial than several sensors with poor energy-powered batteries for improving performances. Instead, when

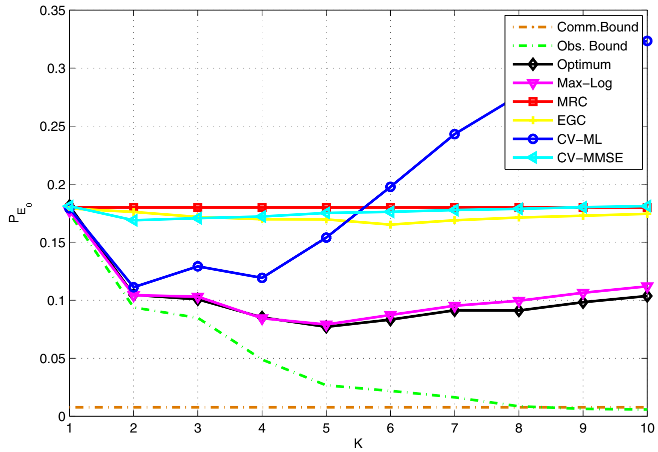


Fig. 6: P_{E_0} vs K for all the rules presented; WSN with $N = 1$ antenna at DFC and channel $(SNR)_{dB} = 15$; $(P_{D,k}, P_{F,k}) = (0.7, 0.05)$, $k \in \mathcal{K}$.

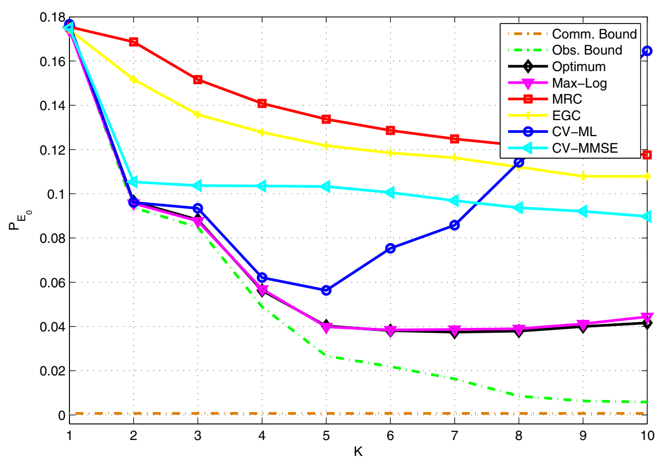


Fig. 7: P_{E_0} vs K for all the rules presented; WSN with $N = 2$ antennas at DFC and channel $(SNR)_{dB} = 15$; $(P_{D,k}, P_{F,k}) = (0.7, 0.05)$, $k \in \mathcal{K}$.

fusing with MRC, the suggested trend is to divide the available energy among as many sensors as possible. Furthermore, some specific facts needs to be clarified:

- $P_{E_0}^{EGC}$ is lower than $P_{E_0}^{MRC}$ in Figs. 6,7; this result does not contradict Figs. 4,5 because the ROCs of the two rules present an intersection point, cfr. Fig. 3, so that the EGC is better than MRC from a Bayesian point of view. The optimal point on ROC which minimizes P_{E_0} is on the right of the intersection point.
- An increase in number of antennas N gives a decrease in minimum P_{E_0} attainable by every rule; the minimum is typically obtained with a larger number of sensors K. For example with CV-ML when $N = 1$ the minimum $P_{E_0}^{CV-ML} \approx 0.11$ is obtained with $K = 2$ sensors; however when $N = 2$ the minimum $P_{E_0}^{CV-ML} \approx 0.05$ is obtained with $K = 5$ sensors. Finally, the increase of N affects slope and limiting value of $P_{E_0}^{MRC}$ for MRC; this means that the same $P_{E_0}^{MRC}$ when using multiple antennas can be obtained with less equal-battery sensors.

VII. Conclusions

In this paper we addressed the design of sub-optimal fusion rules, more suitable for practical implementation than the exact LLR, for a DF task performed over a virtual MIMO channel. The study was motivated by the need of multiple antennas at the DFC to obtain a dramatic improvement in performances with a reduced WSN energy budget. The presented alternatives solve the issues about fixed point implementations and present a wide spectrum of choices for reduced complexity and lower system knowledge. Max-Log, MRC and CV-ML, as in the PAC case, confirm also their asymptotic optimality properties. Nonetheless, all these rules still significantly benefit from the addition of multiple antennas at the DFC, with a saturation on performance depending on the specific rule and channel SNR. Finally, all these rules present an optimal number of sensors to be employed to minimize the system error probability, under a total power constraint. The exception is represented by MRC, which shows decreasing system probability of error as a function of the number of sensors. The latter property configures MRC as a more convenient choice when a huge number of low-powered sensors are available for the task.

VIII. Acknowledgement

The authors would like to thank Prof. R. Schober at University of British Columbia, Canada, for helpful discussions.

Appendix A
Proof of Proposition 1

From definition of exact LLR in Eq. (3) if we observe that for low SNR ($\mathbb{E}\{\|\boldsymbol{H}\boldsymbol{x}\|^2\}/\sigma_w^2 \ll 1$), $\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2}\right) \approx \exp\left(-\frac{\|\boldsymbol{y}\|^2}{\sigma_w^2}\right)\left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right)$, we get

$$\Lambda_{opt} \approx \ln\left[\frac{\sum_{\mathbf{x}\in\mathcal{X}^K}\left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right)\prod_{k=1}^K P(x_k|H_1)}{\sum_{\mathbf{x}\in\mathcal{X}^K}\left(1 - \frac{\|\boldsymbol{H}\boldsymbol{x}\|^2 - 2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}\right)\prod_{k=1}^K P(x_k|H_0)}\right] \tag{17}$$

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

$$\Lambda_{opt} \approx \frac{2\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\left(\mathbb{E}\{\boldsymbol{x}|H_1\} - \mathbb{E}\{\boldsymbol{x}|H_0\}\right)\}}{\sigma_w^2} + \alpha \tag{18}$$

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 $\mathbb{E}\{\boldsymbol{x}|H_0\} = \mathbf{1}_K(2P_F - 1)$, the LLR reduces to:

$$\Lambda_{opt} \approx (P_D - P_F)\frac{4\Re\{\boldsymbol{y}^\dagger \boldsymbol{H}\mathbf{1}_K\}}{\sigma_w^2} + \alpha \tag{19}$$

Eq. (19) represents the same statistics as Eq. (4), since the terms $(P_D - P_F)\frac{4}{\sigma_w^2}$ (recall that $P_D > P_F$) and α can be incorporated in the threshold γ of Eq. (2) .

APPENDIX B
PROOF OF PROPOSITION 2

Part 1): Starting from the expression of Max-Log formula of Eq. (8) let us define:

$$\hat{\boldsymbol{x}}_i \triangleq \arg\min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2}-\sum_{k=1}^{K}\ln P(x_k|H_i)\right],\quad H_i\in\mathcal{H}$$
$$=\arg\min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{H}\boldsymbol{x}\|^2-2\Re\{\boldsymbol{y}^\dagger\boldsymbol{H}\boldsymbol{x}\}}{\sigma_w^2}-\sum_{k=1}^{K}\ln P(x_k|H_i)\right] \tag{20}$$

where in the second line we have neglected the irrelevant term $\frac{\|\boldsymbol{y}\|^2}{\sigma_w^2}$. For low SNR ($\mathbb{E}\{\|\boldsymbol{H}\boldsymbol{x}\|^2\}/\sigma_w^2 \ll 1$), the first term of $\hat{\boldsymbol{x}}_i$ becomes irrelevant, so that the minimum is determined by $\sum_{k=1}^{K}\ln P(x_k|H_i)$. If $P_{D,k}>\frac{1}{2}$ and $P_{F,k}<\frac{1}{2}$, $k\in\mathcal{K}$, then $\hat{\boldsymbol{x}}_0\approx -\boldsymbol{1}_K$ and $\hat{\boldsymbol{x}}_1\approx \boldsymbol{1}_K$. Using these approximations in Eq. (8) provides:

$$\Lambda_{Max-Log}\approx\frac{\|\boldsymbol{y}+\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{1}_K\|^2}{\sigma_w^2}+\delta \tag{21}$$

where $\delta\triangleq\ln\frac{\prod_{k=1}^{K}P_{D,k}}{\prod_{k=1}^{K}(1-P_{F,k})}$ is not dependent on $\boldsymbol{y}$. Eq. (21) is identical to MRC, cfr. Eq. (4), since δ can be easily incorporated in threshold of Eq. (2).

Part 2): The property is easily demonstrated by combining Propositions 1 and 2.1 .

Part 3): Starting from the expression of Max-Log formula of Eq. (8), for high SNR ($\mathbb{E}\{\|\boldsymbol{H}\boldsymbol{x}\|^2\}/\sigma_w^2\gg 1$), we have that $\hat{\boldsymbol{x}}_i\approx\hat{\boldsymbol{x}}_{ML}$, since the first term in r.h.s. of Eq. (20) becomes dominant. Note that the term $\hat{\boldsymbol{x}}_{ML}$ is the same given by Eq. (10). Thus substituting the approximate expressions of $\hat{\boldsymbol{x}}_i$ in Eq. (8), we obtain the same rule as in Eq. (23), which is the CV-ML statistics. Since for high SNR, $\Lambda_{CV-ML}\approx\Lambda_{opt}$ (Proposition 3, Appendix C), then $\Lambda_{Max-Log}\approx\Lambda_{CV-ML}\approx\Lambda_{opt}$.

APPENDIX C
PROOF OF PROPOSITION 3

For high SNR ($\mathbb{E}\{\|\boldsymbol{H}\boldsymbol{x}\|^2\}/\sigma_w^2\gg 1$), if we denote the true transmitted vector as $\boldsymbol{x}_T$, the corresponding value in Eq. (3) will be a dominating term of the sums at numerator and denominator; the LLR is then well approximated by

$$\Lambda_{opt}\approx\ln\left[\frac{\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}_T\|^2}{\sigma_w^2}\right)\prod_{k=1}^{K}P(x_{k,T}|H_1)}{\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}_T\|^2}{\sigma_w^2}\right)\prod_{k=1}^{K}P(x_{k,T}|H_0)}\right] \tag{22}$$
$$=\ln\left[\prod_{k=1}^{K}P(x_{k,T}|H_1)\right]-\ln\left[\prod_{k=1}^{K}P(x_{k,T}|H_0)\right].$$

Also, for high SNR, the ML estimate $\hat{\boldsymbol{x}}_{ML}=\arg\min_{\boldsymbol{x}\in\mathcal{X}^K}\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2\approx\boldsymbol{x}_T$, i.e. the ML decoder works near perfectly. Thus Eq. (22) reduces to

$$\Lambda_{opt}\approx\ln\left[\prod_{k=1}^{K}P(\hat{x}_{k,ML}|H_1)\right]-\ln\left[\prod_{k=1}^{K}P(\hat{x}_{k,ML}|H_0)\right] \tag{23}$$

which can be rearranged easily to obtain Eq. (9).

APPENDIX D
CV-MMSE MOMENTS

We derive explicit expressions for $\overline{\boldsymbol{x}}=\mathbb{E}\{\boldsymbol{x}\}$ and $\boldsymbol{C}=\mathbb{E}\{(\boldsymbol{x}-\overline{\boldsymbol{x}})(\boldsymbol{x}-\overline{\boldsymbol{x}})^\dagger\}$ to be inserted in Eq. (11).

Proposition 4. *The explicit expressions of $\overline{\boldsymbol{x}}$ and $\boldsymbol{C}$ are given by:*

$$\overline{x}_k=P_{F,k}+P_{D,k}-1\quad k\in\mathcal{K} \tag{24}$$
$$c_{\ell,j}=r_{\ell,j}-\overline{x}_\ell\overline{x}_j\quad \ell,j\in\mathcal{K} \tag{25}$$
$$r_{\ell,j}=\begin{cases}1 & \ell=j\\ \frac{(2P_{F,\ell}-1)(2P_{F,j}-1)+(2P_{D,\ell}-1)(2P_{D,j}-1)}{2} & \ell\neq j\end{cases} \tag{26}$$

Proof: The kth element of the mean vector $\overline{x}_k$ can be expressed as :

$$\overline{x}_k=\mathbb{E}\{x_k\}=\sum_{i=0,1}\mathbb{E}\{x_k|H_i\}P(H_i) \tag{27}$$

Assuming equally likely priors (i.e. $P(H_i)=1/2$, $H_i\in\mathcal{H}$), observing that $\mathbb{E}\{x_k|H_0\}=2P_{F,k}-1$ and $\mathbb{E}\{x_k|H_1\}=2P_{D,k}-1$, and substituting these expressions in Eq. (27) gives Eq. (24). Define $r_{\ell,j}$ as the (ℓ,j)th element of the correlation matrix $\boldsymbol{R}=\mathbb{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\}$, i.e.

$$r_{\ell,j}=\mathbb{E}\{x_\ell x_j\}=\sum_{i=0,1}\mathbb{E}\{x_\ell x_j|H_i\}P(H_i) \tag{28}$$
$$=\begin{cases}\frac{1}{2}\sum_{i=0,1}\mathbb{E}\{x_\ell x_j|H_i\} & \ell\neq j\\ \frac{1}{2}\sum_{i=0,1}\mathbb{E}\{x_\ell^2|H_i\} & \ell=j\end{cases} \tag{29}$$

Exploiting the conditional independence (given H_i) of x_ℓ and x_j we have that $\mathbb{E}\{x_\ell x_j|H_i\}=\mathbb{E}\{x_\ell|H_i\}\cdot\mathbb{E}\{x_j|H_i\}$ and observing that $\mathbb{E}\{x_\ell^2|H_i\}=1$, $H_i\in\mathcal{H}$, we obtain Eq. (24). ■

APPENDIX E
MAX-LOG GSD DERIVATION

Proposition 5. *The Max-Log formula of Eq. (8) can be expressed in the equivalent form:*

$$\Lambda_{Max-Log}=\min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2}-\sum_{k=1}^{K}\ln P(x_k|H_0)\right]$$
$$-\min_{\boldsymbol{x}\in\mathcal{X}^K}\left[\frac{\|\boldsymbol{D}(\boldsymbol{\rho}-\boldsymbol{H}\boldsymbol{x})\|^2}{\sigma_w^2}-\sum_{k=1}^{K}\ln P(x_k|H_1)\right] \tag{30}$$

Proof: To prove this proposition we follow similar steps as in [23]. Starting from the expression in Eq. (8), adding and subtracting the constant term $\beta\frac{\boldsymbol{x}^\dagger\boldsymbol{x}}{\sigma_w^2}$ (recall that BPSK is a constant modulus modulation, i.e. $\boldsymbol{x}^\dagger\boldsymbol{x}=K$) and simplifying the term $\frac{\|\boldsymbol{y}\|^2}{\sigma_w^2}$ we get Eq. (31) at the top of the next page.

Denote $\boldsymbol{D}$ the upper-triangular matrix deriving from the Cholesky Factorization of $\boldsymbol{G}\triangleq\boldsymbol{H}^\dagger\boldsymbol{H}+\beta\boldsymbol{I}_N$ (that is $\boldsymbol{G}=\boldsymbol{D}^\dagger\boldsymbol{D}$) and $\boldsymbol{\rho}\triangleq\boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{y}$. We can add and subtract the term $\frac{\boldsymbol{\rho}^\dagger\boldsymbol{D}^\dagger\boldsymbol{D}\boldsymbol{\rho}}{\sigma_w^2}$ to both the minimum problems and use the relationships now defined to straightly obtain Eq. (30). The new squared radius choice r_{new}^2 for pruning is given by

$$r_{new}^2=r^2+\frac{\beta\boldsymbol{x}^\dagger\boldsymbol{x}+\boldsymbol{y}^\dagger\boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^\dagger\boldsymbol{y}-\boldsymbol{y}^\dagger\boldsymbol{y}}{\sigma_w^2} \tag{32}$$

$$\Lambda_{Max-Log} = \min_{\boldsymbol{x}\in\mathcal{X}^K} \left[\frac{\boldsymbol{x}^\dagger(\boldsymbol{H}^\dagger\boldsymbol{H} + \beta\boldsymbol{I}_N)\boldsymbol{x} - \boldsymbol{y}^\dagger\boldsymbol{H}\boldsymbol{x} - \boldsymbol{x}^\dagger\boldsymbol{H}^\dagger\boldsymbol{y}}{\sigma_w^2} - \sum_{k=1}^{K} \ln P(x_k|H_0) \right] - \min_{\boldsymbol{x}\in\mathcal{X}^K} \left[\frac{\boldsymbol{x}^\dagger(\boldsymbol{H}^\dagger\boldsymbol{H} + \beta\boldsymbol{I}_N)\boldsymbol{x} - \boldsymbol{y}^\dagger\boldsymbol{H}\boldsymbol{x} - \boldsymbol{x}^\dagger\boldsymbol{H}^\dagger\boldsymbol{y}}{\sigma_w^2} - \sum_{k=1}^{K} \ln P(x_k|H_1) \right] \quad (31)$$

where r represents the radius of the original test $\left[\frac{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2}{\sigma_w^2} - \sum_{k=1}^{K} \ln P(x_k|H_i)\right] < r^2,\ H_i \in \mathcal{H}$. ■

The computation of the radius in Eq. (32) can be avoided if depth-first Schnorr-Euchner enumeration strategy is adopted in the pruning process [33].

REFERENCES

[1] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Fusion of decisions transmitted over fading channels in wireless sensor networks," in *Proc. 2002 Asilomar Conference on Signals, Systems and Computers.*, vol. 2, pp. 1184–1188.

[2] ——, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[3] 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, July 2006.

[4] R. Jiang, S. Misra, B. Chen, and A. Swami, "Robust suboptimal decision fusion in wireless sensor networks," in *Proc. 2005 IEEE Military Communications Conference*, vol. 4, pp. 2107–2113.

[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] M. K. Banavar, A. D. Smith, C. Tepedelenlioglu, and A. Spanias, "Distributed detection over fading MACs with multiple antennas at the fusion center," in *Proc. 2010 IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 2894–2897.

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

[8] 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.

[9] ——, "Coherent max-log decision fusion in wireless sensor networks," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1327–1332, May 2010.

[10] 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.

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

[12] M. Gastpar, M. Vetterli, and P. Dragotti, "Sensing reality and communicating bits: a dangerous liaison," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 70–83, July 2006.

[13] 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.

[14] 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.

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

[16] S. K. Jayaweera, "Optimal Bayesian data fusion and low-complexity approximations for distributed DS-CDMA wireless sensor networks in Rayleigh fading," in *Proc. 2005 International Conference on Intelligent Sensing and Information Processing*, pp. 19–24.

[17] J. S. Dyer and B. Natarajan, "Improved low-complexity fusion receivers for synchronous DS-CDMA sensor networks," in *Proc. 2007 International Conference on Computer Communications and Networks*, pp. 1123–1126.

[18] S. K. Jayaweera, "Large system decentralized detection performance under communication constraints," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 769–771, Sep. 2005.

[19] 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.

[20] K. Liu and A. M. Sayeed, "Asymptotically optimal decentralized type-based detection in wireless sensor networks," in *Proc. 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 873–876.

[21] ——, "Optimal distributed detection strategies for wireless sensor networks," in *Proc. 2004 Allerton Conf. on Commun., Control and Comp.*

[22] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1081–1092, Mar. 2007.

[23] 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.

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

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

[26] D. J. Love and R. W. Heath Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, July 2003.

[27] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. 1995 IEEE International Conference on Communications*, vol. 2, pp. 1009–1013.

[28] 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.

[29] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Prentice Hall PTR, 1993.

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

[31] S. K. Jayaweera, "Bayesian fusion performance and system optimization for distributed stochastic Gaussian signal detection under communication constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1238–1250, Apr. 2007.

[32] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, 1st edition. Wiley, 2000.

[33] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

Domenico Ciuonzo (S'11) was born in Aversa (CE), Italy, on June 29th, 1985. He received the bachelor degree (*summa cum laude*) and the master's degree (*summa cum laude*) in computer engineering, respectively in 2007 and 2009, both from Second University of Naples, Aversa (CE), Italy. His work has been focused on wireless multiple-input multiple-output (MIMO) and code division multiple access (CDMA) channels. He is currently following the Ph.D. program in Electronics Engineering at the Department of Information Engineering, Second University of Naples, Aversa (CE), Italy. In 2011 he was involved in the Visiting Researcher Programme of NATO Underwater Research Center (NURC), La Spezia, Italy; he worked in the "Maritime Situation Awareness" project. He is currently a visiting scholar at the Electrical and Computer Engineering Department of University of Connecticut (UConn), Storrs, US. His research interests are mainly in the areas of Data and Decision Fusion, Model Learning and Probabilistic Graphical Models. Eng. Ciuonzo is a reviewer for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONICS SYSTEMS, and *Elsevier Physical Communications*.

Gianmarco Romano (M'11) is currently Assistant Professor at the Department of Information Engineering, Second University of Naples, Aversa (CE), Italy. He received the "Laurea" degree in Electronic Engineering from the University of Naples "Federico II" and the Ph.D. degree from the Second University of Naples, in 2000 and 2004, respectively. From 2000 to 2002 he has been Researcher at the National Laboratory for Multimedia Communications (C.N.I.T.) in Naples, Italy. In 2003 he was Visiting Scholar at the Department of Electrical and Electronic Engineering, University of Connecticut, Storrs, USA. Since 2005 he has been with the Department of Information Engineering, Second University of Naples and in 2006 has been appointed Assistant Professor. His research interests fall within the areas of signal processing and communications.

Pierluigi Salvo Rossi (SM'11) was born in Naples, Italy, on April 26, 1977. He received the "Laurea" degree in telecommunications engineering (*summa cum laude*) and the Ph.D. degree in computer engineering in 2002 and 2005, respectively, both from the University of Naples "Federico II", Naples, Italy. During his Ph.D. studies, he has been a researcher at the Inter-departmental Research Center for Signals Analysis and Synthesis (CIRASS), University of Naples "Federico II", Naples, Italy; at the Department of Information Engineering, Second University of Naples, Aversa (CE), Italy; at the the Communications and Signal Processing Laboratory (CSPL), Department of Electrical and Computer Engineering, Drexel University, Philadelphia, PA, US.; and an adjunct professor at the Faculty of Engineering, Second University of Naples, Aversa (CE), Italy. From 2005 to 2008 he worked as a postdoc at the Department of Computer Science and Systems, University of Naples "Federico II", Naples, Italy; at the Department of Information Engineering, Second University of Naples, Aversa (CE), Italy; at the Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway; and visited the Department of Electrical and Information Technology, Lund University, Lund, Sweden. Since 2008, he is an assistant professor (tenured in 2011) in telecommunications at the Department of Information Engineering, Second University of Naples, Aversa (CE), Italy; and has been a guest researcher at the Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway. His research interests fall within the areas of signal processing and communications.