

# CLOSED-FORM ERROR EXPONENT FOR THE NEYMAN-PEARSON FUSION OF TWO-DIMENSIONAL MARKOV LOCAL DECISIONS

*Jorge Plata-Chaves, and Marcelino Lázaro*

Department of Signal Theory and Communications  
Universidad Carlos III de Madrid, Leganés (Madrid), SPAIN

## ABSTRACT

We consider a distributed detection system formed by a large number of local detectors and a fusion center that performs a Neyman-Pearson fusion of the binary quantizations of the sensor observations. The aforementioned local decisions are taken with no kind of cooperation and transmitted to the fusion center over error free parallel access channels. Furthermore, the devices are located on a rectangular lattice so that sensors belonging to a specific row or column are equally spaced. For each hypothesis  $H_0$  and  $H_1$ , the correlation structure of the local decisions is modelled with a two-dimensional causal field where the rows and columns are outcomes of the same first-order binary Markov chain. Under this scenario, we derive a closed-form error exponent for the Neyman-Pearson fusion of the local decisions. Afterwards, using the derived error exponent we study the effect of different design parameters of the network on its overall detection performance.

## 1. INTRODUCTION

The design of sensor networks that undertake a distributed detection task has drawn much attention in last years because of the stringent constraints under which they operate. Initially, with the aim of exploiting all their detection potential, many researchers developed several data fusion techniques that maximize the detection performance of the fusion rule employed by the data fusion center and performed on the sensor observations. Among others, for the case where the sensor observations are conditionally independent under each hypothesis of the test, we recall the optimum fusion rule derived by Chair and Varshney [1] under the Bayesian set up and the corresponding one obtained by Thomopoulos *et al.* [2] under the Neyman-Pearson formulation. Afterwards, holding the assumption of independent local sensor observations given any hypothesis, a big number of studies obtain theoretical results that yield design guidelines aimed at optimizing the overall detection performance of a parallel fusion network. For instance, Tsitsiklis [3] showed that the optimal binary decentralized detection is achieved by identical local detection rules when the number of sensors is arbitrarily large. Additionally, in connection with the previous result Chamberland and Veeravalli [4] proved that having more sensors exceeds asymptotically the benefits of getting detailed information from each sensor.

More recently, the studies of the aforementioned sensor networks considered dependence among the local sensor observations under each one of the hypothesis involved in the inference problem. In a scenario where the sensors send conditionally dependent local decisions to the fusion center, one of the first results on this topic [5] showed that the optimum fusion rule under the Neyman-Pearson criteria consist in the joint likelihood ratio of the messages received by the data fusion center. Meanwhile, when a Bayes hypothesis test is performed and the sensor observations are dependent, Kam *et al.* [6] generalized the structure of the optimal data fusion rule. Later, given the aforementioned optimal fusion rules, as it happened when the sensor observations are conditionally independent, under the distributed detection paradigm various works have provided answers to different design aspects regarding the deployment of the network as well as the use of its limited resources. For instance, using the large deviation theory works such as [7]-[11] have obtained amenable tools that allow to study the effect of different physical

parameters of a one-dimensional network on its overall detection performance.

Although there are many results on the topic of decentralized detection with dependent observations, the literature that solve the cited problem in a two-dimensional (2-D) setting is less extensive. Some of the latest results that address this issue are [12] and [13]. Under the Neyman-Pearson formulation Anandkumar *et al.* [12] derive a closed-form error exponent when a test for independence is considered, and when the correlation structure of the sensor observations is given by a nearest-neighbour Gauss-Markov Random Field (G.M.R.F.). In this way, the authors characterize the detection performance of the employed Neyman-Pearson fusion rule with respect to different design parameters of the network, e.g. the density of the deployment or the signal to noise ratio of the observations performed by the devices. In [13] the authors investigate the amount of information obtainable from a sensor network where the devices are located on a 2-D lattice, and where under each one of the two hypothesis the observations received by the fusion center are distributed according to a 2-D hidden G.M.R.F. defined by a symmetric first-order conditional autoregression (S.F.C.A.R.) model.

However, up to now no work obtains analytically tractable expressions that allow the design of 2-D sensor networks where the devices are located on a rectangular grid and the Neyman-Pearson fusion rule is performed on dependent quantized summaries of the sensor observations. Motivated by this last fact, in the spirit of [7]-[10], and [12], we extend the work made in [11] in order to derive a closed-form error exponent for the Neyman-Pearson fusion rule performed by the fusion center of the sensor network shown in Figure 1. Specifically, modelling the correlation structure of the local decisions by means of a 2-D random process constructed from a first-order binary Markov chain, this closed-form expression is obtained for a scenario where the dependent local decisions are taken with no kind of cooperation, and they are noiseless transmitted to the fusion center. This way, we provide an amenable tool that links the detection performance of a 2-D version of the sensor network described in [5] with some of its physical and design features.

## 2. PROBLEM STATEMENT

We consider a network formed by a data fusion center and  $NL$  devices located on a 2-D lattice  $\mathcal{S}_{N,L}$  where, as it is shown in Figure 1, the sensors belonging to a specific row or column of the lattice are equally spaced. With the aim of deciding what state of the phenomenon is present,  $H_0$  or  $H_1$ , the distributed system undertakes the following steps. Firstly, each device performs a local observation of the environment,  $y_{i,j}$ . Secondly, they apply a binary detection rule to it,  $\gamma(y_{i,j})$ , not necessarily based on a log-likelihood ratio test (L.L.R.T.). This way, each device makes a local decision,  $u_{i,j} = \gamma(y_{i,j})$ , regarding the presence or absence of the phenomenon we want to detect. Thirdly, the binary local quantizations of the sensor observations are transmitted to the fusion center over error free parallel access channels. Finally, based on the  $NL$  local decisions taken by the devices of the network, the fusion center makes a global decision,  $u_{FC}$ , under the Neyman-Pearson formulation. Considering that the correlation structure of the local decisions is hypothesis dependent and modelled with a 2-D causal field where the rows and columns are outcomes of the same first-order binary Markov chain,

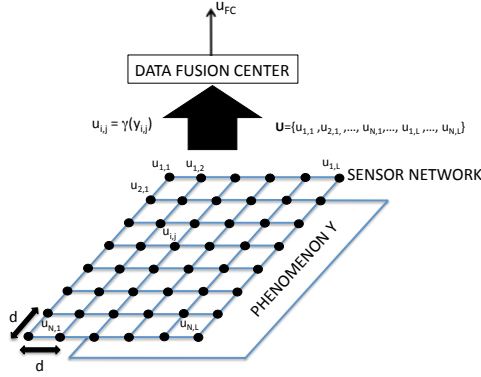


Figure 1: Block diagram of a 2-D parallel fusion network.

we have the following inference problem at the fusion center

$$\left. \begin{aligned}
 H_k: \quad & \mathbf{U} \sim \text{2-D process where the binary local quantizations} \\
 & \text{belonging to a specific row } i \in \{1, \dots, N\} \text{ or column} \\
 & j \in \{1, \dots, L\} \text{ form the same first-order Markov} \\
 & \text{chain, } M_k, \text{ with transition matrix} \\
 & \Pi_k = \begin{bmatrix} p_k^{(0,0)}(d, \mathcal{P}_k) & p_k^{(0,1)}(d, \mathcal{P}_k) \\ p_k^{(1,0)}(d, \mathcal{P}_k) & p_k^{(1,1)}(d, \mathcal{P}_k) \end{bmatrix}^T
 \end{aligned} \right\} \quad (1)$$

where, for  $k \in \{0, 1\}$ ,

- $\mathbf{U} = [U_{1,1}, U_{2,1}, \dots, U_{N,1}, \dots, U_{1,L}, \dots, U_{N,L}]^T$ , defined on  $\mathcal{U}^{NL}$  with  $\mathcal{U} = \{0, 1\}$ , denotes the  $NL$  local decisions performed by the  $NL$  devices that form the network.
- $p_k^{(0,1)}(d, \mathcal{P}_k)$  and  $p_k^{(1,0)}(d, \mathcal{P}_k)$  are the transition probabilities of the Markov chain,  $M_k$ , that extends along the rows and columns of  $\mathcal{S}_{NL}$  under  $H_k$ , i.e. for  $u_m, u_{m-1} \in \{0, 1\}$

$$p_k^{(u_{m-1}, u_m)}(d, \mathcal{P}_k) = P(U_m = u_m | U_{m-1} = u_{m-1}, d, \mathcal{P}_k, H_k) \quad (2)$$

- $d$  is the distance between two neighbour devices that belong to the same row or column of the lattice  $\mathcal{S}_{NL}$ .
- $\mathcal{P}_k$  is the set of physical and design parameters of the network that being known by the data fusion center are arguments of the transition probabilities associated with the Markov chain  $M_k$  [e.g. the employed local decision rules or the physical properties of the environment where the network is deployed].

Note that, due to the causality of the correlation model assumed under  $H_k$ , the local decision  $U_{i,j}$  is independent of the rest of the binary local quantizations when some of the ones taken by the neighbour nodes are known. Specifically, given the hypothesis  $H_k$ , with  $k \in \{0, 1\}$ , the local decisions responsible for the aforementioned conditional independence are established by the directions that the Markov chains  $M_k$  adopt when extending along the rows and columns of the 2-D lattice,  $\mathcal{S}_{NL}$ . For instance, as it happens in [14], if we assume that the Markov chains under both hypothesis extend from left to right along the rows and from top to bottom along the columns, under  $H_0$  and  $H_1$  the decision  $U_{i,j}$  is independent of the rest of the local detectors when the binary local quantizations  $U_{i,j-1}$  and  $U_{i+1,j}$  are known. Without loss of generality and from now and on, given the hypothesis  $H_k$  we are going to assume the previous direction in the flow of the Markov chains present in each column and row of  $\mathcal{S}_{NL}$ . Consequently, given the correlation structure assumed for  $\mathbf{U}$  under both hypothesis, if we also consider that the spacing between the nearest neighbours,  $d$ , is known by the data fusion center, the optimum Neyman-Pearson detection rule [5]

performed at the fusion center is given by

$$u_{fc} = \ln \left( \frac{P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|\mathcal{Q}_0, H_0)}{P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|\mathcal{Q}_1, H_1)} \right) \frac{H_0}{H_1} \geq \tau \quad (3)$$

where, for  $k \in \{0, 1\}$ ,  $\mathcal{Q}_k = d \cup \mathcal{P}_k$ ,  $\tau$  is a constant chosen to bound the overall false alarm probability to a given value  $\alpha \in (0, 1)$

$$P_{FA}(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) = \sum_{\mathbf{u} \in \mathcal{V}_{NL}^1(\tau)} P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|\mathcal{Q}_0, H_0) \leq \alpha \quad (4)$$

$\mathcal{V}_{NL}^1(\tau)$  is equal to the acceptance region of hypothesis  $H_1$  and, after omitting whatever kind of dependency on  $\mathcal{Q}_k$  for the sake of simplicity,  $P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|H_k)$  is the joint probability mass function (p.m.f.) of the local decisions under the hypothesis  $H_k$

$$\begin{aligned}
 P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|H_k) &= P_{U_{1,1}|H}(u_{1,1}|H_k) \times \prod_{i=2}^N P_{U_{i,1}|U_{i-1,1},H}(u_{i,1}|u_{i-1,1}, H_k) \\
 &\times \prod_{j=2}^L P_{U_{N,j}|U_{N,j-1},H}(u_{N,j}|u_{N,j-1}, H_k) \\
 &\times \prod_{i=N-1}^1 \prod_{j=2}^L P_{U_{i,j}|U_{i+1,j},U_{i,j-1},H}(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_k)
 \end{aligned} \quad (5)$$

with  $P_{U_{1,1}|H}(u_{1,1}|H_k)$  equal to the p.m.f. of the initial state associated with the 2-D random process present under  $H_k$ ,

$$P_{U_{i,j}|U_{i-1,j},H}(u_{i,j}|u_{i-1,j}, H_k) = \sum_{(r,s) \in \mathcal{U}^2} p_k^{(s,r)} \delta[r - u_{i,j}] \delta[s - u_{i-1,j}] \quad (6)$$

for all  $i \in \{2, 3, \dots, N\}$  and  $j \in \{1, 2, \dots, L\}$ ,

$$P_{U_{i,j}|U_{i,j-1},H}(u_{i,j}|u_{i,j-1}, H_k) = \sum_{(r,s) \in \mathcal{U}^2} p_k^{(s,r)} \delta[r - u_{i,j}] \delta[s - u_{i,j-1}] \quad (7)$$

for all  $i \in \{1, 2, \dots, N\}$  and  $j \in \{2, 3, \dots, L\}$ , and

$$\begin{aligned}
 &P_{U_{i,j}|U_{i+1,j},U_{i,j-1},H}(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_k) \\
 &= \sum_{(r,s,t) \in \mathcal{U}^3} \frac{p_k^{(t,r)} p_k^{(r,s)}}{\hat{p}_k^{(t,s)}} \delta[r - u_{i,j}] \delta[s - u_{i+1,j}] \delta[t - u_{i,j-1}]
 \end{aligned} \quad (8)$$

when  $i \in \{1, 2, \dots, N-1\}$ ,  $j \in \{2, 3, \dots, L\}$  and

$$\hat{p}_k^{(t,s)} = \sum_{v=0}^1 p_k^{(t,v)} p_k^{(v,s)} \quad (9)$$

### 3. ERROR EXPONENT

Here we provide a tool that allows the design and analysis of a sensor network as the one described in the previous section. To be specific, based on information theoretic results we derive a design tool that links the detection performance of the Neyman-Pearson test (3) with different physical and design parameters of the network shown in Figure 1. In order to accomplish this aim, we would want to characterize the overall probability of detection of the network

$$P_D(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) = \sum_{\mathbf{u} \in \mathcal{V}_{NL}^1(\tau)} P_{\mathbf{U}|\mathcal{Q},H}(\mathbf{u}|H_1) \quad (10)$$

when a fixed constraint is imposed on the overall false alarm probability, i.e. when  $P_{FA}(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) \leq \alpha \in (0, 1)$ . Taking into account that the derivation of a closed-form expression for (10) is

not feasible, our approach has focused on using more easy-to-use measures of performance provided by the large deviations theory corresponding with Neyman-Pearson hypothesis tests. In particular, following the trend of related publications as well as focusing on our distributed detection problem, our objective is to derive a closed-form expression of the exponential rate of decay in  $P_M(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) = 1 - P_D(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L)$  as  $N$  and  $L$  approach infinity and when  $P_{FA}(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) \leq \alpha \in (0, 1)$ . As it is proved in [15], this problem results in calculating the almost-sure limit under  $H_0$  (a.s. in  $H_0$ ) of the asymptotic Kullback-Leibler rate

$$\begin{aligned} K &\triangleq \lim_{N, L \rightarrow \infty} -\frac{1}{N \cdot L} \log(P_M(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L)) \\ &= \lim_{N, L \rightarrow \infty} \frac{1}{N \cdot L} \ln \left( \frac{P_{\mathbf{U}|H}(\mathbf{u}|H_0)}{P_{\mathbf{U}|H}(\mathbf{u}|H_1)} \right) \quad (\text{a.s. in } H_0) \end{aligned} \quad (11)$$

In this way, we arrive at the following theorem.

**Theorem 1.** *Suppose that  $0 < p_k^{(u_{m-1}, u_m)} < 1$  for all  $u_m, u_{m-1}, k \in \{0, 1\}$  and that the p.m.f. of the local decision  $U_{1,1}$  under  $H_0$  is absolutely continuous with respect to the corresponding one under  $H_1$ ,  $P_{U_{1,1}|H}(u_{1,1}|H_0) \ll P_{U_{1,1}|H}(u_{1,1}|H_1)$ . Then, given a fixed constraint  $P_{FA}(\mathcal{Q}_0, \mathcal{Q}_1, N \cdot L) \leq \alpha \in (0, 1)$ , the best Neyman-Pearson error exponent for the distributed detection problem given by Equation (1) is*

$$\begin{aligned} K &= \sum_{(r,s,t) \in \mathcal{U}^3} \pi_0^{(t)} p_0^{(t,r)} p_0^{(r,s)} \ln \left( \frac{p_0^{(t,r)} p_0^{(r,s)} \hat{p}_1^{(t,s)}}{p_1^{(t,r)} p_1^{(r,s)} \hat{p}_0^{(t,s)}} \right) \\ &= D(P(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_0) || P(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_1)) \end{aligned} \quad (12)$$

where

- $\mathcal{U}^3$  equals the cartesian product of the set  $\mathcal{U} = \{0, 1\}$  with itself, three times.
- $\pi_0^{(0)} = 1 - \pi_0^{(1)} = P(u = 0|H_0) = \frac{p_0^{(1,0)}}{p_0^{(1,0)} + p_0^{(0,1)}}$  is the unique stationary probability of deciding  $H_0$  when the 2-D random process constructed from the Markov chain,  $M_0$ , is present.
- $D(P(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_0) || P(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_1))$  is the conditional Kullback-Leibler divergence of

$$P_{U_{i,j}|U_{i+1,j}, U_{i,j-1}, H}(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_0)$$

and

$$P_{U_{i,j}|U_{i+1,j}, U_{i,j-1}, H}(u_{i,j}|u_{i+1,j}, u_{i,j-1}, H_1)$$

in the stationary regime.

*Proof.* Since the error exponent for the Neyman-Pearson detector with a fixed level  $\alpha \in (0, 1)$  is given in implicit form by (11), we focus on the calculation of this limit for the scenario described in Section 2. Taking into account the factorization of  $P_{\mathbf{U}|H}(\mathbf{u}|H_k)$  given in (5) as well as the characterizations provided in (6)-(9), Equation (11) can be written as follows

$$\begin{aligned} K &= \lim_{N, L \rightarrow \infty} \frac{1}{N \cdot L} \ln \left( \frac{P_{U_{1,1}|H}(u_{1,1}|H_0)}{P_{U_{1,1}|H}(u_{1,1}|H_1)} \right) \quad (\text{a.s. in } H_0) \\ &\quad \underbrace{\hspace{10em}}_{\triangleq A_1} \\ &+ \underbrace{\sum_{r,s \in \mathcal{U}^2} \ln \left( \frac{p_0^{(s,r)}}{p_1^{(s,r)}} \right) \lim_{N, L \rightarrow \infty} \frac{N_0(r, s | u_{2,1}^{N,1})}{N \cdot L}}_{\triangleq A_2} \\ &+ \underbrace{\sum_{r,s \in \mathcal{U}^2} \ln \left( \frac{p_0^{(s,r)}}{p_1^{(s,r)}} \right) \lim_{N, L \rightarrow \infty} \frac{N_0(r, s | u_{N,2}^{N,L})}{N \cdot L}}_{\triangleq A_3} \end{aligned}$$

$$+ \underbrace{\sum_{r,s,t \in \mathcal{U}^3} \ln \left( \frac{p_0^{(t,r)} p_0^{(r,s)} \hat{p}_1^{(t,s)}}{p_1^{(t,r)} p_1^{(r,s)} \hat{p}_0^{(t,s)}} \right) \lim_{N, L \rightarrow \infty} \frac{N_0(r, s, t | u_{1,2}^{N-1,L})}{N \cdot L}}_{\triangleq A_4} \quad (13)$$

where

- $N_0(u_{i,j}, u_{i-1,j} | u_{2,1}^{N,1})$  equals the number of times that the pair  $(u_{i,1}, u_{i-1,1})$  occurs in the sequence of local decisions  $u_{2,1}^{N,1} = \{u_{i,j} : 2 \leq i \leq N, j = 1\}$  performed by the corresponding devices under the hypothesis  $H_0$ .
- $N_0(u_{i,j}, u_{i,j-1} | u_{N,2}^{N,L})$  equals the number of times that the pair  $(u_{N,j}, u_{N,j-1})$  occurs in the sequence of local decisions  $u_{N,2}^{N,L} = \{u_{i,j} : i = N, 2 \leq j \leq L\}$  performed by the corresponding devices under the hypothesis  $H_0$ .
- $N_0(u_{i,j}, u_{i+1,j}, u_{i,j-1} | u_{1,2}^{N-1,L})$  equals the number of times that the triple  $(u_{i,j}, u_{i+1,j}, u_{i,j-1})$  occurs in the lattice of local decisions  $u_{1,2}^{N-1,L} = \{u_{i,j} : 1 \leq i \leq N-1, 2 \leq j \leq L\}$  performed by the corresponding devices under the hypothesis  $H_0$ .

Starting from the assumptions that ensure the existence of  $K$  we calculate each one of the terms that appear in the right hand side of  $K$ . In particular, we are going to prove that Equation (13) yields (12) when  $P_{U_{1,1}|H}(u_{1,1}|H_0) \ll P_{U_{1,1}|H}(u_{1,1}|H_1)$  and  $0 < p_k^{(u_{m-1}, u_m)} < 1$  holds for all  $u_m, u_{m-1}, k \in \{0, 1\}$ . Firstly, if we take into account that  $P_{U_{1,1}|H}(u_{1,1}|H_0) \ll P_{U_{1,1}|H}(u_{1,1}|H_1)$  we have that

$$\ln \left( \frac{P_{U_{1,1}|H}(u_{1,1}|H_0)}{P_{U_{1,1}|H}(u_{1,1}|H_1)} \right) < \infty \quad (14)$$

and therefore, we prove that  $A_1$  vanishes as  $N$  and  $L$  go to infinity. Secondly, we derive the terms  $A_2$  and  $A_3$ . Considering the assumptions that ensure the existence of  $K$ , we guarantee the regularity of the Markov chain,  $M_k$ , that extends along the rows and columns of  $\mathcal{S}_{N,L}$  under  $H_k$  with  $k \in \{0, 1\}$  (see [16]). Consequently, given the hypothesis  $H_k$  and knowing that the sequences  $u_{2,1}^{N,1}$  and  $u_{N,2}^{N,L}$  are outcomes of the same first-order binary Markov chain,  $M_0$ , as it is proved in Theorem 1 of [11] under the same initial assumptions, we obtain

$$\begin{aligned} &\sum_{r,s \in \mathcal{U}^2} \ln \left( \frac{p_0^{(s,r)}}{p_1^{(s,r)}} \right) \lim_{N \rightarrow \infty} \frac{N_0(r, s | u_{2,1}^{N,1})}{N} \\ &= \sum_{r,s \in \mathcal{U}^2} \ln \left( \frac{p_0^{(s,r)}}{p_1^{(s,r)}} \right) \lim_{L \rightarrow \infty} \frac{N_0(r, s | u_{N,2}^{N,L})}{L} \\ &= \sum_{r,s \in \mathcal{U}^2} \pi_0^{(s)} p_0^{(s,r)} \ln \left( \frac{p_0^{(s,r)}}{p_1^{(s,r)}} \right) < \infty \end{aligned} \quad (15)$$

From (15) we can easily show that  $A_2$  and  $A_3$  are equal to zero. Finally, we evaluate  $A_4$ . For this calculation, we need to obtain an asymptotic closed-form expression for the empirical joint probability of the triple,  $(u_{i,j}, u_{i+1,j}, u_{i,j-1}) \in \mathcal{U}^3$  given the hypothesis  $H_0$  and the lattice  $u_{1,2}^{N-1,L}$ . From the stationarity of the regular Markov chain,  $M_0$ , regarding the type of  $(u_{i,j}, u_{i+1,j}, u_{i,j-1})$  in  $u_{1,2}^{N-1,L}$  it is straightforward to prove that

$$\begin{aligned} &\lim_{N, L \rightarrow \infty} \frac{N_0(u_{i,j}, u_{i+1,j}, u_{i,j-1} | u_{1,2}^{N-1,L})}{N \cdot L} \\ &= \pi_0^{(u_{i,j-1})} p_0^{(u_{i,j-1}, u_{i,j})} p_0^{(u_{i,j}, u_{i+1,j})} \end{aligned} \quad (16)$$

if, as we assumed without loss of generality,  $M_0$  extends from left to right and from top to bottom along the rows and columns of  $\mathcal{S}_{N,L}$ , respectively. Lastly, a pointwise substitution of (16) into the definition of  $A_4$  results in the closed-form error exponent provided in Theorem 1. This way we conclude the proof.  $\square$

#### 4. CHARACTERIZATION OF THE ERROR EXPONENT

Through synthetic experiments based on the evaluation of  $K$  for a specific physical model of the transition probabilities, in this section we give some insights into the behaviour of  $K$  when the dependence strength among the local decisions varies. As a measure of the aforementioned dependence we use the parameter derived in [11]. To be more precise, this index captures the mean correlation strength among the random variables corresponding with neighbour steps of a first-order Markov chain with binary state space

$$\rho = \begin{cases} P(H_0)\rho_0 + P(H_1)\rho_1 & \text{When the local decisions are} \\ & \text{dependent under } H_0 \text{ and } H_1. \\ \rho_k & \text{When the local decisions are only} \\ & \text{dependent under } H_k. \end{cases} \quad (17)$$

where,  $\forall k \in \{0, 1\}$ ,  $P(H_k) \in (0, 1)$  denotes the prior probability of hypothesis  $H_k$  and

$$\rho_k = 1 - p_k^{(0,1)} - p_k^{(1,0)} \quad (18)$$

Regarding the physical model considered for the transition probabilities of the Markov chains,  $M_0$  and  $M_1$ , we have that

$$p_k^{(0,1)} = 1 - p_k^{(0,0)} = \xi_k(1 - m_k e^{-\gamma_k d}) \quad (19)$$

and

$$p_k^{(1,0)} = 1 - p_k^{(1,1)} = (1 - \xi_k)(1 - m_k e^{-\gamma_k d}) \quad (20)$$

where, for  $k \in \{0, 1\}$ ,

- $\xi_k$  is the probability of false alarm or detection probability of the local detectors when independence among the local decisions is assumed under  $H_0$  or  $H_1$  respectively.
- $\gamma_k$  is a strictly positive constant that indicates the exponential rate of growth of the transition probabilities  $p_k^{(0,1)}$  and  $p_k^{(1,0)}$  as  $d$  increases.
- $m_k$  is a strictly positive constant that, taking on values less than one, controls the correlation between the pair of local decisions  $(U_{i-1,j}, U_{i,j})$  and  $(U_{i,j}, U_{i,j+1})$  when the distance between the devices involved in each pair is zero, i.e  $d = 0$ .

Note that the physical model considered for  $p_k^{(0,1)}$  and  $p_k^{(1,0)}$  fulfills some regularity conditions that typically appear in a detection scenario. On the one hand, due to the fact that the transition probabilities are monotonically concave increasing functions of  $d$ , the correlation strength among the local decisions decays as the devices become farther apart. On the other hand, modelling the so-called nugget effect, according to the definition provided in [5] the local decisions are never going to be maximally dependent under  $H_k$  since the transition probabilities are greater than zero for  $m_k < 1$ . Hence, given  $H_k$  and the physical model described in (19) and (20), as imposing a minimum distance between two neighbour nodes belonging to a specific row or column of  $\mathcal{S}_{N,L}$ , the correlation strength between the corresponding local decisions is always less than one for  $m_k < 1$ . In addition to the previous regularity conditions, we can also appreciate that, in the model described by (19) and (20), the local decisions  $U_{i,j}$ , with  $i \in \{1, 2, \dots, N\}$  and  $j \in \{1, 2, \dots, L\}$ , are only independent when  $d$  equals infinity.

In Figure 2, we plot the error exponent of Equation (12) as a function of the mean correlation strength provided through Equations (17) and (18). Theoretical curves of this figure have been generated when the local decisions are only dependent under  $H_1$  and  $\xi_1$  is equal to  $\{0.8, 0.98, 0.998\}$ . As it can be appreciated, regardless

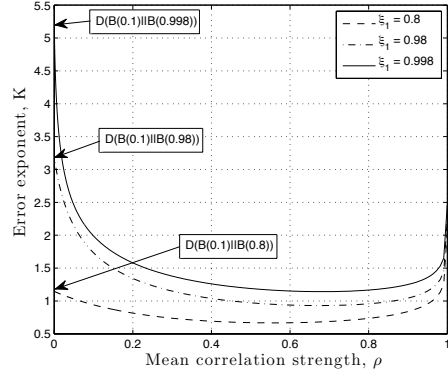


Figure 2: Neyman-Pearson optimum error exponent,  $K$ , as a function of the mean correlation strength among neighbour binary decisions,  $\rho$ , when they are only dependent under  $H_1$  and  $\xi_1 = \{0.8, 0.98, 0.998\}$ . Parameters:  $\xi_0 = 0.1$ ,  $\gamma_0 = 2$ ,  $\gamma_1 = 0.9$ , and  $m_1 = 1 - 10^{-4}$ .

the value of  $\xi_1$ ,  $K$  initially decreases as  $\rho$  increases, and after a specific value  $\rho^*$ , it increases as  $\rho$  approaches one. On the one hand, the initial decrease of  $K$  in  $\rho$  happens because there is a loss of discrimination between  $H_0$  and  $H_1$  when the information provided by each sensor is more and more correlated without improving the detection performance of the sensors. On the other hand, the increasing behaviour of  $K$  with respect to  $\rho$  occurs because excessive redundancy among the local decisions is informative when the local decisions are only dependent under  $H_1$ . Note that, under  $H_0$ , each sensor performs its decisions independently of the rest the network. However, due to the fact that  $\rho = \rho_1$ , under  $H_1$ , all the local detectors tend to decide the same hypothesis when  $\rho$  approaches one. When the dependence under  $H_1$  exceeds a specific threshold,  $\rho^*$ , the agreement that exists among the values taken by the major part of the binary local quantizations can be used in order to discriminate  $H_1$  against  $H_0$ . As it can be proved through a careful analysis, the aforementioned value,  $\rho^*$ , depends on physical parameters of the network such as  $\xi_0$  and  $\xi_1$ . In particular, confirming the behaviour observed in Figure 2 we can prove that  $\rho^*$  is shifted closer to one as the ratio  $\xi_1/\xi_0$  increases.

In order to conclude the analysis of Figure 2, we can check a pair of analytic results that, being omitted for the sake of brevity, show the consistency of Theorem 1 with similar studies performed when the dependence among the local decisions results in boundary values of  $\rho$ . Firstly, when  $\rho$  equals zero, we can appreciate that, for all the plotted curves,  $K$  collapses to the subsequent Kullback-Leibler divergence,  $D(B(\xi_0)||B(\xi_1))$ , where  $B(\xi_k)$  denotes a Bernoulli random variable with probability of success equal to  $\xi_k$ . Taking into account that the local decisions are independent and identically distributed (i.i.d.) when  $\rho$  equals zero, this behaviour is consistent with the Neyman-Pearson error exponent stated by the Stein's Lemma [17]. Subsequently, if we center on the evolution of the curves shown in Figure 2 when  $\rho$  approaches one, we can appreciate that the derived error exponent,  $K$ , diverges regardless the value of  $\xi_1$ . Due to the fact that  $p_1^{(r,s)} \notin (0, 1)$  for all  $r, s \in \{0, 1\}$  as long as  $\rho_1 = 1$ , a brief justification of the previous behaviour is based on the fact that some of the initial assumptions of Theorem 1 do not hold. In order to be more precise, when the local decisions are maximally dependent under  $H_1$ , the support of  $P_{\mathbf{U}|H}(\mathbf{u}|H_1)$  only consists of the events where all the devices decide the same hypothesis [5]. However, under conditional independence given the hypothesis  $H_0$ , the support of  $P_{\mathbf{U}|H}(\mathbf{u}|H_0)$  is formed by the two aforementioned events as well as the rest of the  $2^{NL}$  binary  $NL$ -tuples. Consequently,  $P_{\mathbf{U}|H}(\mathbf{u}|H_1) \ll P_{\mathbf{U}|H}(\mathbf{u}|H_0)$  holds, and therefore, the existence of  $K$  can not be ensured although

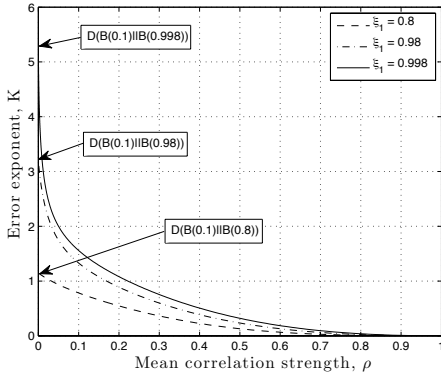


Figure 3: Neyman-Pearson optimum error exponent,  $K$ , as a function of the mean correlation strength among neighbour binary decisions,  $\rho$ , when they are dependent under  $H_1$  and  $H_0$ , and  $\xi_1 = \{0.8, 0.98, 0.998\}$ . Parameters:  $\xi_0 = 0.1$ ,  $\gamma_0 = 2$ ,  $\gamma_1 = 0.9$ ,  $P(H_k) = 0.5$  and  $m_k = 1 - 10^{-4}$  with  $k \in \{0, 1\}$ .

$P_{U_{1,1}|H}(u_{1,1}|H_0) \ll P_{U_{1,1}|H}(u_{1,1}|H_1)$  occurs.

Next, in Figure 3 we make the same analysis as in Figure 2 with the proviso that the local decisions are also dependent under  $H_0$ . For each one of the curves plotted in Figure 3 we can observe that  $K$  decreases as the correlation strength,  $\rho$ , increases from the case of independent local decisions,  $\rho = 0$ , to the case of maximally dependent local decisions,  $\rho = 1$ , as the definition provided in [5]. In addition, it can be seen that the amount of decrease in  $K$  becomes smaller as  $\rho$  increases. Nevertheless, unlike the case where a correlation structure is only present under  $H_1$ , different simulations show that the monotonicity of  $K$  with respect to  $\rho$  is not common to all the possible set of values that can take the parameters involved in (19) and (20). What is more, when the local decisions are dependent under both hypothesis the characterization of the monotonicity of  $K$  as a function of  $\rho$  is not analytically tractable.

Finally, we analyze the convergence of  $K$  shown in Figure 3 for  $\rho = 0$  and  $\rho = 1$ . Being consistent with the behaviour observed in Figure 2, when  $\rho = 0$  the error exponent reduces to  $D(B(\xi_0)||B(\xi_1))$ , i.e. the Stein's Lemma. Meanwhile, as it is expected from analytic studies, in Figure 3 when  $\rho$  goes to one the aforementioned error exponent converges to zero independently of the parameters that appear in the physical model described in (19) and (20). In addition to the analytic proof that we have omitted, an intuitive interpretation of the previous convergence can be found by realizing that excessive dependence among the local decisions under both hypothesis makes the reading of an extra device provide the fusion center with a lot of redundant information when discriminating  $H_1$  against  $H_0$ . Equivalently, if the local decisions are maximally dependent, once the fusion center has read a device, new readings do not provide additional information that improves the overall detection performance of the network.

## 5. CONCLUSIONS

In this paper we addressed the design of a distributed detection system formed by a large number of devices and a data fusion center. With no kind of cooperation and from its own observation each device performs a local decision regarding the underlying binary hypothesis testing problem. Afterwards, over an error free parallel access channel each sensor transmits its decision to a fusion center that makes a global decision,  $u_{FC}$ , under the Neyman-Pearson formulation. We considered that the devices are located on a rectangular lattice where the sensors belonging to a specific row or column are equally spaced. Additionally, the local decisions are assumed to be dependent under both hypothesis. In particular, given each one of the two hypothesis the dependence among the local deci-

sions is modelled by means of a 2-D causal field where the rows and columns are outcomes of the same first-order binary Markov chain. Under an arbitrary physical model that links the physical parameters of the network with the transition probabilities of the aforementioned Markov chains, we firstly derived a closed-form error exponent for the Neyman-Pearson test performed at the fusion center. Finally, after choosing a physical model for the transition probabilities corresponding with the Markov chains present under each of the two hypothesis, we perform several evaluations and analytical studies in order to know how the error exponent behaves as the dependence among the binary local quantizations vary.

## Acknowledgment

This work was supported in part by Ministerio de Educación y Ciencia of Spain (project MONIN, id. TEC2006-13514-C02-01), Ministerio de Ciencia e Innovación of Spain (project COMONSENS, id. CSD2008-00010) and Comunidad de Madrid (project PROMULTIDIS-CM, id. S0505/TIC/0223).

## REFERENCES

- [1] P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, 1986.
- [2] S. C. A. Thomopoulos, R. Viswanathan, and D. C. Bougoulas, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 23, no. 5, pp. 644–653, 1987.
- [3] J. N. Tsitsiklis, "Decentralized detection," *Adv. Stat. Signal Process.*, vol. 2, pp. 297–344, 1993.
- [4] J. F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, 2003.
- [5] E. Drakopoulos and C. C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–606, 1991.
- [6] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, 1992.
- [7] J. F. Chamberland and V. V. Veeravalli, "How dense should a sensor network be for decentralized detection with correlated observations," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5099–5106, 2006.
- [8] Y. Sung, L. Tong, and H. V. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: closed-form error exponent and properties," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1354–1365, 2006.
- [9] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, 2007.
- [10] S. Misra and L. Tong, "Error exponents for the detection of Gauss-Markov signals using randomly spaced sensors," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3385–3396, 2008.
- [11] J. Plata-Chaves and M. Lázaro, "Closed-form error exponent for the Neyman-Pearson fusion of Markov local decisions," in *Proc. IEEE/SP 15th Workshop Stat. Signal Process.*, 2009, pp. 533–536.
- [12] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random fields with nearest-neighbor dependency," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 816–827, 2009.
- [13] Y. Sung, H. V. Poor, and H. Yu, "How much information can one get from a wireless ad hoc sensor network over a correlated random field?" *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2827–2847, 2009.
- [14] J. Justesen, "Fields from Markov chains," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4358–4362, 2005.
- [15] I. Vajda, *Theory of statistical inference and information*. Kluwer Academic, Dordrecht, The Netherlands, 1989.
- [16] J. G. Kemeny and J. L. Snell, *Finite Markov chains*. Springer, 1976.
- [17] T. M. Cover and J. A. Thomas, *Elements of information theory*. Wiley-Interscience, New York, 2006.