

## MIT Open Access Articles

*Error exponents for decentralized  
detection in feedback architectures*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Tay, Wee Peng, and John N. Tsitsiklis. Error Exponents for Decentralized Detection in Feedback Architectures. In 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2496-2499. Institute of Electrical and Electronics Engineers, 2011.

**As Published:** <http://dx.doi.org/10.1109/ICASSP.2011.5946991>

**Publisher:** Institute of Electrical and Electronics Engineers (IEEE)

**Persistent URL:** <http://hdl.handle.net/1721.1/79644>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Creative Commons Attribution-Noncommercial-Share Alike 3.0



# ERROR EXPONENTS FOR DECENTRALIZED DETECTION IN FEEDBACK ARCHITECTURES

Wee Peng Tay

Division of Communication Engineering  
Nanyang Technological University  
50 Nanyang Avenue, Singapore 639798

John N. Tsitsiklis

Laboratory for Information and Decision Systems  
Massachusetts Institute of Technology  
Cambridge, MA 02139

## ABSTRACT

We consider the decentralized Bayesian binary hypothesis testing problem in feedback architectures, in which the fusion center broadcasts information based on the messages of some sensors to some or all sensors in the network. We show that the asymptotically optimal detection performance (as quantified by error exponents) does not benefit from the feedback messages. In addition, we determine the corresponding optimal error exponents.

**Index Terms**— Decentralized detection, feedback, error exponent, sensor networks.

## 1. INTRODUCTION

We consider the binary decentralized detection problem, in which each sensor in a network makes an observation, quantizes it to a given alphabet, and transmits the result to a fusion center. The fusion center makes a final decision based on all the sensor messages. The objective is to design the sensor quantization functions and the fusion rule so as to minimize a cost function, such as the probability of an incorrect final decision.

The decentralized detection problem has been widely studied for the parallel configuration (see [1] and the references therein), tandem networks [2–4], and bounded height tree architectures [5–8]. A variety of feedback architectures, under a Bayesian formulation, have been studied in [9, 10]. These references show that it is person-by-person optimal for every sensor to use a likelihood ratio quantizer, with thresholds that depend on the feedback messages. However, because of the difficulty of optimizing these thresholds when the number of sensors becomes large, it is difficult to analytically compare the performance of networks with and without feedback. To better understand the asymptotics of the error probability, [11] studies the error probability decay rate under a Neyman-Pearson formulation for two different feedback architectures. For either case, it shows that if the fusion center also has access to the feedback messages, then feedback does not improve the optimal error exponent. References [12, 13] consider the Neyman-Pearson problem in the daisy-chain architecture (see Figure 1), and obtain a similar result.

In this paper, we consider the decentralized Bayesian detection problem in various feedback architectures. We study the daisy chain architectures in [12], under which the sensors are divided into two groups, and sensors in the second group have full or partial knowledge of the messages sent by the first group. Reference [12] dealt with the Neyman-Pearson formulation. In this paper, we turn to the Bayesian formulation and resolve several questions that had been left

open. In addition, we provide results for the Bayesian counterpart of the feedback architecture considered in [11].

The remainder of this paper is organized as follows. In Section 2 we formulate the problems that we will be studying. In Section 3, we analyze the performance of various feedback architectures. We summarize and conclude in Section 4.

## 2. PROBLEM FORMULATION

We consider a decentralized binary detection problem involving  $n$  sensors and a fusion center. Each sensor  $k$  observes a random variable  $X_k$  distributed according to a measure  $\mathbb{P}_j$  under hypothesis  $H_j$ , for  $j = 0, 1$ . Under either hypothesis  $H_j$ ,  $j = 0, 1$ , the random variables  $X_k$  are assumed to be i.i.d. We use  $\mathbb{E}_j$  to denote the expectation operator with respect to  $\mathbb{P}_j$ .

In the daisy chain architecture introduced in [12, 13], every sensor sends a single message to the fusion center, but some of the sensors have access to the messages of other sensors (see Figure 1). The first stage consists of  $m$  sensors and the second stage  $n - m$  sensors. All observations at the sensors are assumed to be conditionally i.i.d., given the hypothesis. Each sensor  $k$  in the first stage sends a message  $Y_k = \gamma_k(X_k)$  to an aggregator. The aggregator forms a message  $U$  that is broadcast to all sensors in the second stage and the fusion center. Each sensor  $l$  in the second stage forms a message  $Z_l = \delta_l(X_l, U)$ , which depends on its own observation and the message  $U$ . The fusion center makes the final decision using the fusion rule  $Y_f = \gamma_f(U, Z_{m+1}, \dots, Z_n)$ . We denote by  $\Gamma$  the set of allowable quantization functions for the first stage sensors. For simplicity, we assume that  $\Gamma$  is rich enough so that for any given realization of  $U = u$ , the quantization functions  $\delta_l(\cdot, u) = \delta_l^u(\cdot) \in \Gamma$ . We can also view the architecture just described as a parallel configuration, in which the fusion center feedbacks a message based on information from sensors  $1, \dots, m$ , to the rest of the sensors  $m + 1, \dots, n$ .

We consider two cases for how  $U$  is formed. In the first case, we let  $U = (Y_1, \dots, Y_m)$ , i.e., the second stage sensors and fusion center have the full information available at the first stage aggregator. We call this the full feedback daisy chain. In another form of feedback, we take  $U = \gamma_u(Y_1, \dots, Y_m) \in \{0, 1\}$  to be a preliminary decision made in the first stage. We call this the restricted feedback daisy chain. In this case, the architecture is equivalent to a parallel configuration, in which the fusion center makes a preliminary decision based on the messages from the first  $m$  sensors, broadcasts the preliminary decision, and forgets the messages sent by the first  $m$  sensors. The fusion center could be subject to memory or security constraints, and does not retain the first  $m$  messages.

In the two-message feedback architecture (see Figure 2), each sensor  $k$  sends a message  $Y_k = \gamma_k(X_k)$  to the fusion center. Sim-

This research was supported, in part, by the Nanyang Technological University Startup Grant, and by the NSF under grant ECCS-0701623.

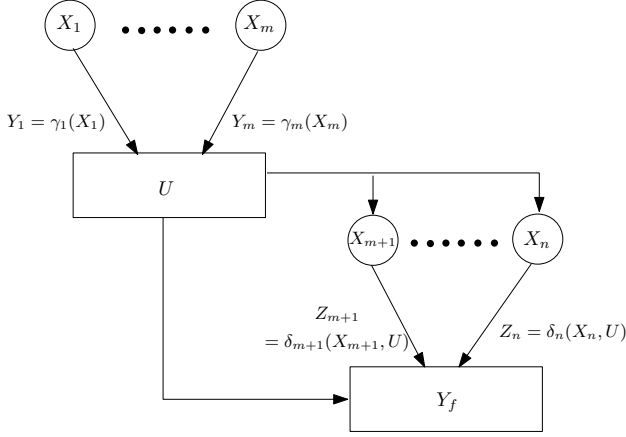


Fig. 1. The daisy chain architecture.

ilar to the restricted feedback for the daisy chain, the fusion center broadcasts the message  $U = \gamma_u(Y_1, \dots, Y_n)$  to all the sensors. Each sensor then sends a second message  $Z_k = \delta_k(X_k, U)$  to the fusion center. The final decision of the fusion center is made based on the received messages  $Y_1, \dots, Y_n$  and  $Z_1, \dots, Z_n$ .

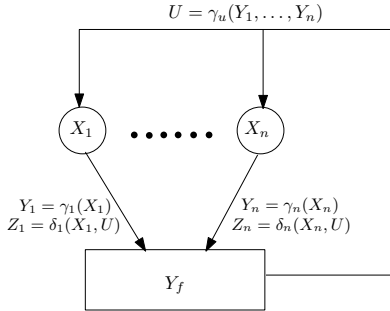


Fig. 2. A two-message architecture.

Let  $\mathbb{P}_i^X$  be the distribution of a random variable  $X$  under hypothesis  $H_i$ . Consider the Radon-Nikodym derivative  $d\mathbb{P}_i^X/d\mathbb{P}_j^X$  of the measure  $\mathbb{P}_i^X$  with respect to (w.r.t.) the measure  $\mathbb{P}_j^X$ . It is a random variable whose value is determined by  $X$ ; accordingly, its value should be denoted by a notation such as  $\ell_{ij}^X(X)$ . However, in order to avoid cluttered expressions, we will abuse notation and just write  $\ell_{ij}(X)$ . We also use  $\ell_{ij}(\gamma(X))$  to denote the Radon-Nikodym derivative of the random variable  $Z = \gamma(X)$ . Throughout this paper, we deal with various conditional distributions. Abusing notation as before, we let  $\ell_{ij}(X|Y)$  be the Radon-Nikodym derivative of the conditional distribution of  $X$  given  $Y$  under  $H_i$  w.r.t. that under  $H_j$ .

We make the following assumptions throughout this paper (see [14] for justifications).

**Assumption 1.** *The measures  $\mathbb{P}_0$  and  $\mathbb{P}_1$  are absolutely continuous w.r.t. each other. Furthermore, there exists some  $\gamma \in \Gamma$  such that  $-\mathbb{E}_0[\log \ell_{01}(\gamma(X_1))] < 0 < \mathbb{E}_1[\log \ell_{10}(\gamma(X_1))]$ .*

**Assumption 2.** *We have  $\mathbb{E}_j[\log^2 \ell_{01}(X_1)] < \infty$  for  $j = 0, 1$ .*

### 3. PERFORMANCE ANALYSIS

Let the prior probability of hypothesis  $H_j$  be  $\pi_j > 0$ ,  $j = 0, 1$ . Given a strategy, the probability of error at the fusion center is  $P_e = \pi_0 \mathbb{P}_0(Y_f = 1) + \pi_1 \mathbb{P}_1(Y_f = 0)$ . Let  $P_e^*$  be the minimum probability of error, over all strategies. We seek to characterize the optimal error exponent

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P_e^*.$$

From [14], the optimal error exponent for the parallel configuration without any feedback is given by

$$\mathcal{E}_p^* = \inf_{(\gamma, \delta) \in \Gamma^2} \min_{\lambda \in [0, 1]} \log \mathbb{E}_0[\exp(\lambda \log \ell_{10}(\gamma(X_1), \delta(X_1)))].$$

We first show that under the Bayesian formulation, the full feedback daisy chain and the two-message architecture both have the same optimal error exponent as a parallel configuration with  $n$  sensors. Let  $\mathcal{L}_{10}^{(n)}$  be the log likelihood ratio at the fusion center, and  $\psi_n(\lambda) = \log \mathbb{E}_0[\exp(\lambda \mathcal{L}_{10}^{(n)})]$  be the log moment generating function. The Fenchel-Legendre transform of  $\psi_n$  is given by  $\Psi_n(t) = \sup_{\lambda \in \mathbb{R}} \{\lambda t - \psi_n(\lambda)\}$ . The following lemma applies in both the daisy chain and two-message architectures. An outline of the proof is provided in the appendix.

**Lemma 1.** *Suppose Assumptions 1 and 2 hold.*

- (i) *For all  $s \in [0, 1]$ , we have  $\mathbb{E}_0[\log \ell_{10}(X_1)] \leq \psi_n'(s)/n \leq \mathbb{E}_1[\log \ell_{10}(X_1)]$ .*
- (ii) *Let  $t$  be such that for all  $n$ , there exists  $s_n \in (0, 1)$  with  $\psi_n'(s_n) = t$ . Then, there exists a constant  $C$  such that for all  $n$ , we have  $\psi_n''(s_n) \leq nC$ .*
- (iii) *For all  $s \in [0, 1]$ , we have  $\psi_n(s)/n \geq \mathcal{E}_p^*$ .*

**Theorem 1.** *There is no loss in optimality if all sensors in the full feedback daisy chain are constrained to using the same quantization function, with sensors in the second stage ignoring the feedback message. Similarly, there is no loss in optimality if all sensors in the two-message architecture ignore the feedback message, and use the same quantization functions (one for the first message and another for the second message). Moreover, the optimal error exponent in either architecture is  $\mathcal{E}_p^*$ .*

*Proof.* (Outline) Let the optimal error exponent for the full feedback daisy chain be  $\mathcal{E}^*$ . Since this architecture can simulate the parallel configuration, we have  $\mathcal{E}^* \leq \mathcal{E}_p^*$ . To show the reverse bound, let  $P_{e,j}$  be the conditional error probability under  $H_j$ . We use Lemma 1 in the following upper bound generalized from [15] to obtain

$$\max_{j=0,1} P_{e,j} \geq \frac{1}{4} \exp\left(\psi_n(s_n^*) - \sqrt{2\psi_n''(s_n^*)}\right) \geq \exp(n\mathcal{E}_p^* - C\sqrt{n}),$$

where  $C$  is a constant. Taking  $n \rightarrow \infty$ , we obtain the upper bound for  $\mathcal{E}^*$ . The proof for the two-message architecture is identical.  $\square$

In the following, we obtain the error exponent for the restricted feedback daisy chain architecture, and show that it is strictly worse than that of a parallel configuration with  $n$  sensors. We assume that  $\lim_{n \rightarrow \infty} m/n = r \in (0, 1)$ , otherwise the architecture is equivalent to a parallel configuration. Let  $\mathcal{E}_{dc}^*$  be the optimal error exponent. For  $\gamma \in \Gamma$ , and  $j = 0, 1$ , let  $\Lambda_j^*(\gamma, t) = \sup_{s \in \mathbb{R}} \{st - \log \mathbb{E}_j[\exp(s \log \ell_{10}(\gamma(X_1)))]\}$ . For  $i, j \in \{0, 1\}$ , let the rate of decay of the conditional probabilities be  $e_{ij} = -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_i(U = j)$ . We collect the decay rates into a vector  $\vec{e} = [e_{01}, e_{10}, e_{00}, e_{11}]$ .

**Lemma 2.** For any strategy for the restricted feedback daisy chain architecture, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P_e \geq -h(\vec{e}),$$

and

$$h(\vec{e}) = \min \left\{ (1-r) \sup_{\gamma \in \Gamma} \Lambda_0^* \left( \gamma, \frac{r}{1-r} (e_{10} - e_{00}) \right) + r e_{00}, \right. \\ \left. (1-r) \sup_{\gamma \in \Gamma} \Lambda_1^* \left( \gamma, -\frac{r}{1-r} (e_{01} - e_{11}) \right) + r e_{11} \right\}$$

*Proof.* The same argument as in the proof of Corollary 3.4.6 of [16] shows that it is sufficient to prove the lower bound for a strategy using a zero threshold log likelihood ratio test at the fusion center. Henceforth, we will assume that such a fusion rule is employed. Conditioning on the value of  $U$ , we have

$$P_1(Y_f = 0) = \mathbb{P}_1(Y_f = 0 | U = 0) \mathbb{P}_1(U = 0) \\ + \mathbb{P}_1(Y_f = 0 | U = 1) \mathbb{P}_1(U = 1).$$

Let  $\delta_i(\cdot, u) = \delta_i^u(\cdot) \in \Gamma$  be a function that depends on the value of  $u$ . Let  $\epsilon > 0$ . From the lower bound in Cramèr's Theorem [16], we have

$$\frac{1}{n} \log \mathbb{P}_1(Y_f = 0 | U = 0) \\ = \frac{1}{n} \log \mathbb{P}_1 \left[ \sum_{i=1}^{m_2} \log \ell_{10}(\delta_i^0(X_i)) \leq -\log \frac{\mathbb{P}_1(U = 0)}{\mathbb{P}_0(U = 0)} \right] \\ \geq -\frac{m_2}{n} \left[ \frac{1}{m_2} \sum_{i=1}^{m_2} \Lambda_1^* \left( \delta_i^0, -\frac{1}{m_2} \log \frac{\mathbb{P}_1(U = 0)}{\mathbb{P}_0(U = 0)} - \epsilon \right) - \epsilon \right] \\ + o(1) \\ \geq -\frac{m_2}{n} \left[ \sup_{\gamma} \Lambda_1^* \left( \gamma, -\frac{1}{m_2} \log \frac{\mathbb{P}_1(U = 0)}{\mathbb{P}_0(U = 0)} - \epsilon \right) - \epsilon \right] + o(1),$$

where  $o(1)$  is a term that goes to zero as  $n$  becomes large. Taking  $n \rightarrow \infty$  and then  $\epsilon \rightarrow 0$ , and using the uniform continuity of  $\Lambda_1^*(\gamma, \cdot)$ , we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_1(Y_f = 0 | U = 0) + \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_1(U = 0) \\ \geq -(1-r) \sup_{\gamma} \Lambda_1^* \left( \gamma, \frac{r}{1-r} (e_{10} - e_{00}) \right) - r e_{10} \\ = -(1-r) \sup_{\gamma} \Lambda_0^* \left( \gamma, \frac{r}{1-r} (e_{10} - e_{00}) \right) - r e_{00} \quad (1)$$

In the same way, it can be checked that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_1(Y_f = 0 | U = 1) + \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_1(U = 1) \\ \geq -(1-r) \sup_{\gamma} \Lambda_1^* \left( \gamma, -\frac{r}{1-r} (e_{01} - e_{11}) \right) - r e_{11}, \quad (2)$$

and we obtain  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_1(Y_f = 0) \geq -h(\vec{e})$ . A similar proof shows that  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_0(Y_f = 1) \geq -h(\vec{e})$ , and the lemma is proved.  $\square$

**Theorem 2.** The optimal error exponent for the restricted feedback daisy chain is

$$\mathcal{E}_{dc}^* = -(1-r) \sup_{\substack{\gamma, \delta \in \Gamma \\ t \in \mathbb{R}}} \min \left\{ \Lambda_0^* \left( \gamma, \frac{r}{1-r} \Lambda_1^*(\delta, t) \right), \right. \\ \left. \Lambda_1^* \left( \gamma, -\frac{r}{1-r} \Lambda_0^*(\delta, t) \right) \right\}. \quad (3)$$

Moreover, there is no loss in optimality if all sensors in the first stage are constrained to using the same quantization function; and sensors in the second stage ignore the feedback message, and are constrained to using the same quantization function.

*Proof.* (Outline) We skip most of the details and provide an outline of the proof here. Let the threshold of the first stage aggregator be  $t_n$ . Since  $t_n$  is in a bounded interval, we can choose a subsequence  $(n_k)_{k \in \mathbb{N}}$  such that  $t_{n_k} \rightarrow t$ . It suffices to prove the lower bound for this subsequence. For any fusion rule  $\gamma_u$  for the first stage, the lower bound in Lemma 2 can be achieved by letting all sensors in the second stage use the same quantization function that ignores the feedback message  $U$ . This implies that the restricted feedback architecture is equivalent to a tree architecture with two stages. Furthermore, it is optimal for the first stage fusion rule to be a log likelihood ratio test [1]. Consequently, it can be shown [12] that there is no loss in optimality if all the stage one sensors are restricted to the same quantization function  $\delta \in \Gamma$ . Applying Lemma 2 with  $e_{01} = \Lambda_0^*(\delta, t)$  and  $e_{10} = \Lambda_1^*(\delta, t)$ , we get the theorem.  $\square$

**Proposition 1.** Suppose that there exists  $\gamma, \delta \in \Gamma$  and  $t \in \mathbb{R}$  such that the supremum in (3) is achieved. The restricted feedback daisy chain performs strictly worse than the parallel configuration with the same total number of sensors, i.e.,  $\mathcal{E}_{dc}^* > \mathcal{E}_p^*$ .

*Proof.* We have

$$(1-r) \Lambda_0^* \left( \gamma, \frac{r}{1-r} \Lambda_1^*(\delta, t) \right) \\ = (1-r) \Lambda_1^* \left( \gamma, \frac{r}{1-r} \Lambda_1^*(\delta, t) \right) + r \Lambda_1^*(\delta, t) \\ < (1-r) \Lambda_1^*(\gamma, 0) + r \Lambda_1^*(\delta, t), \quad (4)$$

where the last inequality follows from  $\Lambda_1^*(\gamma, \cdot)$  being a decreasing function and  $\Lambda_1^*(\delta, t) > 0$ . Similarly,

$$(1-r) \Lambda_1^* \left( \gamma, -\frac{r}{1-r} \Lambda_0^*(\delta, t) \right) < (1-r) \Lambda_0^*(\gamma, 0) + r \Lambda_0^*(\delta, t). \quad (5)$$

Combining (4) and (5), we obtain

$$\mathcal{E}_{dc}^* > -(1-r) \Lambda_0^*(\gamma, 0) - r \min \{ \Lambda_1^*(\delta, t), \Lambda_0^*(\delta, t) \} \\ \geq -\Lambda_0^*(\gamma, 0) \geq -\sup_{\gamma \in \Gamma} \Lambda_0^*(\gamma, 0) = \mathcal{E}_p^*.$$

The proof is now complete.  $\square$

## 4. CONCLUSION

We have studied feedback architectures in which a group of sensors have access to information from sensors not in the group. We show that feedback does not improve the optimal error exponent. In

the case where the fusion center has only limited knowledge (a 1-bit summary) of the messages, the optimal error exponent is strictly worse than that of the parallel configuration. A similar result holds for the two-message architecture. This research is part of our ongoing efforts to quantify the performance of various network architectures. Future research directions include studying the impact of feedback on distributed multiple hypothesis testing and parameter estimation.

## 5. APPENDIX

*Outline Proof of Lemma 1* : The proof of claim (i) is similar to Proposition 3 of [14], and is omitted here. To prove (ii), we have

$$\begin{aligned}\psi_n''(s_n) &= \frac{\mathbb{E}_0[(\mathcal{L}_{10}^{(n)})^2 \exp(s_n \mathcal{L}_{10}^{(n)})]}{\mathbb{E}_0[\exp(s_n \mathcal{L}_{10}^{(n)})]} - (\psi_n'(s_n))^2 \\ &\leq C_0 \mathbb{E}_0[(\mathcal{L}_{10}^{(n)})^2 \exp(s_n \mathcal{L}_{10}^{(n)})],\end{aligned}\quad (6)$$

where the last inequality follows from the bound  $\mathbb{E}_0[\exp(s_n \mathcal{L}_{10}^{(n)})] \geq 1/C_0$ , for some constant  $C_0$  (proven in Proposition 3 of [14]). To bound the R.H.S. of (6), we have

$$\begin{aligned}&\mathbb{E}_0[(\mathcal{L}_{10}^{(n)})^2 \exp(s_n \mathcal{L}_{10}^{(n)})] \\ &= \mathbb{E}_0 \left[ (\mathcal{L}_{10}^{(n)})^2 e^{s_n \mathcal{L}_{10}^{(n)}} \mathbf{1}_{\{\mathcal{L}_{10}^{(n)} \leq 0\}} \right] \\ &\quad + \mathbb{E}_1 \left[ (\mathcal{L}_{10}^{(n)})^2 e^{-(1-s_n)\mathcal{L}_{10}^{(n)}} \mathbf{1}_{\{\mathcal{L}_{10}^{(n)} > 0\}} \right] \\ &\leq C_1 \left( \frac{1}{s_n^2} + \frac{1}{(1-s_n)^2} \right).\end{aligned}$$

The rest of the proof, which is technical, and is omitted because of space constraints, shows that both  $s_n$  and  $1-s_n$  are at least  $C_2/\sqrt{n}$  for some constant  $C_2$ . Therefore the claim holds.

In the following, we give an outline of the proof for claim (iii) for the two-message architecture; the proof for the daisy chain is similar, and is omitted. Let  $s = \lambda/n$  and  $Y_1^n = (Y_1, \dots, Y_n)$ . We have

$$\begin{aligned}&\mathbb{E}_0 \left[ \prod_{k=1}^n (\ell_{10}(Z_k | Y_1^n))^s \mid Y_1^n \right] \\ &= \mathbb{E}_0 \left[ \prod_{k=1}^n (\ell_{10}(\delta_k(X_k, U) | Y_k))^s \mid Y_1^n \right] \\ &\geq \prod_{k=1}^n \inf_{\delta^{Y_k} \in \Gamma} \mathbb{E}_0 \left[ (\ell_{10}(\delta^{Y_k}(X_k) | Y_k))^s \mid Y_k \right],\end{aligned}$$

where  $\delta^{Y_k}$  depends on the value of  $Y_k$ . We can define  $\xi_k \in \Gamma^2$  such that  $\xi_k(X_k) = (\gamma_k(X_k), \delta_k(X_k))$ , where  $\delta_k(X_k) = \delta_k^u(X_k)$  iff  $\gamma_k(X_k) = u$ . Therefore, we obtain

$$\begin{aligned}\psi_n(\lambda) &= \log \mathbb{E}_0 \left[ (\ell_{10}(Y_1^n))^s \mathbb{E}_0 \left[ \prod_{k=1}^n (\ell_{10}(Z_k | Y_1^n))^s \mid Y_1^n \right] \right] \\ &\geq \log \mathbb{E}_0 \left[ (\ell_{10}(Y_1^n))^s \prod_{k=1}^n \inf_{\delta^{Y_k} \in \Gamma} \mathbb{E}_0 \left[ (\ell_{10}(\delta^{Y_k}(X_k) | Y_k))^s \mid Y_k \right] \right] \\ &\geq n \inf_{\xi \in \Gamma^2} \log \mathbb{E}_0 [(\ell_{10}(\xi(X_1)))^s] \geq n\mathcal{E}_p^*,\end{aligned}$$

and the lemma is proved.

## 6. REFERENCES

- [1] J. N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.
- [2] R. Viswanathan, S. C. A. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 366–376, 1988.
- [3] J. D. Papastavrou and M. Athans, "Distributed detection by a large team of sensors in tandem," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 639–653, 1992.
- [4] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "On the sub-exponential decay of detection error probabilities in long tandems," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4767–4771, Oct. 2008.
- [5] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, "Optimization of detection networks: part II - tree structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 1, pp. 211–221, 1993.
- [6] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "On the impact of node failures and unreliable communications in dense sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2535–2546, Jun. 2008.
- [7] —, "Data fusion trees for detection: Does architecture matter?" *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4155–4168, Sep. 2008.
- [8] —, "Bayesian detection in bounded height tree networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4042–4051, Oct 2009.
- [9] S. Alhakeem and P. K. Varshney, "A unified approach to the design of decentralized detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 9–20, 1995.
- [10] S. Alhakeem and P. Varshney, "Decentralized Bayesian detection with feedback," *IEEE Trans. Syst., Man, Cybern.*, vol. 26, no. 4, pp. 503–513, Jul. 1996.
- [11] H. Shalaby and A. Papamarcou, "A note on the asymptotics of distributed detection with feedback," *Information Theory, IEEE Transactions on*, vol. 39, no. 2, pp. 633–640, Mar 1993.
- [12] S. I. Zoumpoulis, "Decentralized detection in sensor network architectures with feedback," Master's thesis, Massachusetts Institute of Technology, June 2009.
- [13] O. P. Kreidl, J. N. Tsitsiklis, and S. I. Zoumpoulis, "On decentralized detection with partial information sharing among sensors," *IEEE Trans. Signal Process.*, 2011, in press.
- [14] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Syst.*, vol. 1, pp. 167–182, 1988.
- [15] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels, I," *Information and Control*, vol. 10, pp. 65–103, 1967.
- [16] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. New York, NY: Springer-Verlag, 1998.