

## MIT Open Access Articles

*Decentralized detection in sensor  
network architectures with feedback*

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

**Citation:** Kreidl, O.P.; Tsitsiklis, J.N.; Zoumpoulis, S.I.; , "Decentralized detection in sensor network architectures with feedback," Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on , vol., no., pp.1605-1609, Sept. 29 2010-Oct. 1 2010

**As Published:** <http://dx.doi.org/10.1109/ALLERTON.2010.5707105>

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

**Persistent URL:** <http://hdl.handle.net/1721.1/71638>

**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



# Decentralized Detection in Sensor Network Architectures with Feedback

O. Patrick Kreidl, John N. Tsitsiklis, and Spyros I. Zoumpoulis

**Abstract**—We study a decentralized detection architecture in which each of a set of sensors transmits a highly compressed summary of its observations (a binary message) to a fusion center, which then decides on one of two alternative hypotheses. In contrast to the star (or “parallel”) architecture considered in most of the literature, we allow a subset of the sensors to both transmit their messages to the fusion center and to also broadcast them to the remaining sensors. We focus on the following architectural question: is there a significant performance improvement when we allow such a message broadcast? We consider the error exponent (asymptotically, in the limit of a large number of sensors) for the Neyman-Pearson formulation of the detection problem. We prove that the sharing of messages does not improve the optimal error exponent.

## I. INTRODUCTION

We consider a decentralized detection problem and study the value added (performance improvement) when feeding the messages (“preliminary decisions”) of some of the sensors to the remaining ones, so that the latter can take them into consideration, along with their own observations, to form their own messages. We carry out this comparison under a Neyman-Pearson formulation of the detection problem, and (for reasons of tractability) in the asymptotic regime, as the number of sensors increases. This work is part of a broader effort to understand the performance gains or losses associated with different sensor network architectures. Primarily because of analytical obstacles, this effort had been limited to the star (also called “parallel” architecture) [1], [2] and, somewhat more generally, to tree networks [3]. The present work is, to the authors’ knowledge, the first to provide an exact asymptotic performance analysis of a non-tree network.

### A. Background and Related Literature

We consider a binary hypothesis testing problem, and a number of sensors each of which obtains an observation whose distribution is determined by the true hypothesis. In a centralized system, every sensor communicates its observation to a fusion center that makes a final decision. In contrast, in a star decentralized detection architecture (introduced in the seminal work of Tenney and Sandell [4]), each sensor sends only a summary of its observation to a fusion center, in the form of a message that takes values in a finite alphabet. The fusion center then decides on one of the alternative hypotheses. The problem is to design rules through which each sensor can

form its message, and through which the fusion center can interpret these messages to make a final decision, in a manner that minimizes the probability of error.

Much research has followed [4]; for a review, see [5], [2]. For conditionally dependent observations (given the true hypothesis), the problem is NP-hard. Under the assumption of conditional independence, an optimal decision rule for each sensor takes the form of a likelihood ratio test, with a suitably chosen threshold. In turn, an optimization over the set of all thresholds can yield the desired solution. Numerical algorithms for optimizing sensor thresholds in the star network have been adapted to the series, or tandem, network and later extended to any singly-rooted tree network [6]. It is now known that likelihood ratio tests remain optimal (under the conditional independence assumption) in every directed acyclic network but, with regard to tractable computation, correctness and efficiency is retained only for sparsely-connected tree-structured networks: more specifically, threshold optimization scales exponentially with the maximal degree of the nodes [7].

Similar algorithmic developments have occurred for decentralized detection architectures with feedback [8], [9], [10], where the fusion center may communicate a preliminary decision to the peripheral sensors for them to take into consideration when forming a next message based on a next observation. While the sensors in these feedback architectures do remember all preceding messages from the fusion center, each sensor is forced to forget all preceding own observations; without this memory restriction, the design problem beyond the first stage of feedback essentially faces difficulties similar to those encountered in the intractable case of conditionally-dependent observations.

The complexity of threshold optimization for large, densely-connected decentralized detection networks has motivated the study of a more tractable, asymptotic formulation (as the number of sensors increases to infinity). Reference [11] focuses on optimizing the asymptotic error exponent, for the case of a star architecture with a large number of sensors that receive conditionally independent, identically distributed observations. The broader case of a network consisting of a large number of nodes arranged as a tree of bounded height is considered in [12], [3]. In the same asymptotic spirit, [13] studies decentralized binary detection in a wireless sensor network where each sensor transmits its data over a multiple access channel.

### B. Overview

We study the Neyman-Pearson decentralized detection problem (in the asymptotic regime) for a new network architecture

The authors are with the Laboratory for Information and Decision Systems, MIT Cambridge MA 02139 USA e-mail: opkreidl@alum.mit.edu, jnt@mit.edu, szoumpou@mit.edu. This research was partially supported by the NSF under grant ECCS-0701623.

that we refer to as a “daisy chain.” In this architecture, the second half of the sensors get to see the messages sent to the fusion center by the first half of the sensors, before forming their own messages. This is perhaps the simplest nontrivial non-tree architecture and it features at least partial feedback. While the study of non-tree or feedback architectures appears to be quite difficult in general, the daisy chain turns out to be amenable to asymptotic analysis. Indeed, we are able to prove that the additional information feedback that is present in the daisy chain does not result in a performance improvement: the optimal error exponent is the same as for the case of a star architecture with the same number of sensors.

To our knowledge, asymptotic analysis of decentralized detection networks with feedback has not been undertaken previously *except* in the portion of [14] devoted to concluding that the Neyman-Pearson performance improvement from a single stage of feedback diminishes fast with an increasing signal-to-noise ratio or an increasing number of sensors. For the daisy-chain architecture, our results strengthen those in [14], proving no asymptotic performance gain from a single stage of feedback regardless of the signal-to-noise ratio.

It is worth noting that there are certain cases where it is easily shown that feedback cannot improve asymptotic performance, for somewhat trivial reasons. One such example is binary hypothesis testing for the case of two Gaussian distributions with different means and the same variance, in the limit of a high signal-to-noise ratio. In this case, it turns out that a star architecture (with binary messages) achieves the same error exponent as a centralized architecture in which the observations are transmitted uncompressed to a fusion center [15]. By a sandwich argument, the performance of any architecture that involves binary messages plus some additional feedback falls in between and the error exponent remains the same.

The rest of the paper is organized as follows. We formulate the decentralized detection problem for the various architectures of interest in Section II. We present the main result and its proof in Section III. Finally, we summarize and discuss possible extensions and open problems in Section IV.

## II. PROBLEM FORMULATION

In this section, we introduce the classical star architecture and the daisy chain architecture. We define our notation, make the necessary probabilistic assumptions, define the performance measures of interest, and provide the necessary background.

### A. Probabilistic Assumptions

We assume that the state of the environment satisfies one of two alternative hypotheses  $H_0$  and  $H_1$ . There is an even number,  $n = 2m$ , of sensors, indexed  $1, \dots, n$ . Each sensor  $i$  observes the realization of a random variable  $X_i$ , which takes values in an observation set  $\mathcal{X}$ , endowed with a  $\sigma$ -field  $\mathcal{F}_{\mathcal{X}}$  of measurable sets. We assume that conditioned on either hypothesis  $H_j$ , the random variables  $X_i$  are independent and identically distributed (i.i.d.) according to a measure  $\mathcal{P}_j$  on  $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$ . In the sequel we use the notation  $\mathbf{E}_j[\cdot]$  to indicate an

expectation taken under hypothesis  $H_j$ , and  $\mathbf{P}_j(A)$  to denote the probability of an event  $A$  under  $H_j$ .

As in [11], we make the following technical assumption, which serves to facilitate the subsequent asymptotic analysis.

**Assumption II.1.** *The measures  $\mathcal{P}_0$  and  $\mathcal{P}_1$  are absolutely continuous with respect to each other, but not identical. Furthermore,  $\mathbf{E}_0[\log^2 \frac{d\mathcal{P}_0}{d\mathcal{P}_1}] < \infty$ , where  $\frac{d\mathcal{P}_0}{d\mathcal{P}_1}$  is the Radon-Nikodym derivative of the two measures.*

### B. The Star Architecture

Every sensor  $i$  forms a binary message  $Y_i$ , taking values in  $\{0, 1\}$ , by following a rule of the form  $Y_i = \gamma_i(X_i)$ , for some measurable function  $\gamma_i : \mathcal{X} \mapsto \{0, 1\}$ . Let  $\Gamma$  be the set of all possible such functions. Note that a particular choice of a function  $\gamma \in \Gamma$  results in particular distributions for the binary random variable  $\gamma(X_1)$ , under the two hypotheses. We define the Kullback-Leibler divergence  $D(\gamma)$  of these two distributions by

$$D(\gamma) = \mathbf{P}_0(\gamma(X_1) = 0) \cdot \log \frac{\mathbf{P}_0(\gamma(X_1) = 0)}{\mathbf{P}_1(\gamma(X_1) = 0)} + \mathbf{P}_0(\gamma(X_1) = 1) \cdot \log \frac{\mathbf{P}_0(\gamma(X_1) = 1)}{\mathbf{P}_1(\gamma(X_1) = 1)}. \quad (1)$$

The messages  $Y_1, \dots, Y_n$  are communicated to a fusion center which uses a fusion rule of the form  $\gamma_0 : \{0, 1\}^n \mapsto \{0, 1\}$  and declares hypothesis  $H_j$  to be true if and only if  $Y_0 = \gamma_0(Y_1, \dots, Y_n) = j$ . See Figure 1 for an illustration.

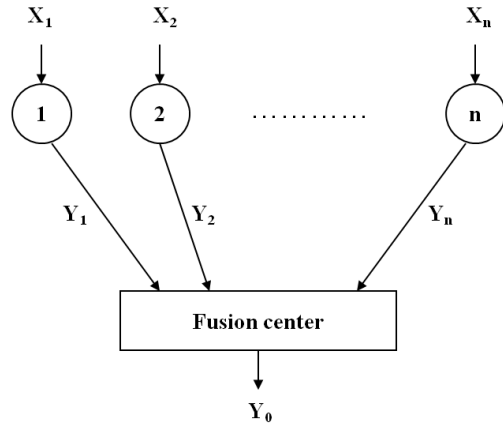


Fig. 1. The star architecture.

According to Theorem 2 in [11], under a Neyman-Pearson formulation, the optimal error exponent for the missed detection probability, in the limit as  $n \rightarrow \infty$ , is given by

$$g_p^* = - \sup_{\gamma \in \Gamma} D(\gamma). \quad (2)$$

Furthermore, given that the two measures  $\mathcal{P}_0$  and  $\mathcal{P}_1$  are not identical, it is easily seen that there exists some  $\gamma$  for which the distribution of  $\gamma(X_1)$  is different under the two hypotheses, so that  $D(\gamma) > 0$ . This implies that  $g_p^* < 0$ .

### C. The Daisy Chain Architecture

In the daisy chain architecture, the underlying probabilistic model and the sensor observations are the same as for the star architecture. What is different is that sensors  $m+1, \dots, 2m = n$  get to observe the messages sent by the first  $m$  sensors before forming their own messages. We use again  $Y_i$  to denote the message sent by sensor  $i$ , and let  $U = (Y_1, \dots, Y_m)$  be the additional information made available to sensors  $m+1, \dots, n$ . The equations that define this architecture are:

$$\begin{aligned} Y_i &= \gamma_i(X_i), & i &= 1, \dots, m, \\ Y_i &= \delta_i(X_i, U), & i &= m+1, \dots, n, \\ Y_0 &= \gamma_0(Y_1, \dots, Y_n). \end{aligned}$$

For  $i \leq m$ ,  $\gamma_i$  is as before the decision rule of sensor  $i$ , a measurable function from  $\mathcal{X}$  to  $\{0, 1\}$ . For  $i > m$ , the decision rule  $\delta_i$  of sensor  $i$  is a measurable function from  $\mathcal{X} \times \{0, 1\}^m$ . Finally, the decision rule  $\gamma_0$  of the fusion center is a function from  $\{0, 1\}^n$  to  $\{0, 1\}$ . Let  $\Gamma$ ,  $\Delta_n$ , and  $\Gamma_{0,n}$  be the sets of possible decision rules  $\gamma_i$ ,  $\delta_i$ , and  $\gamma_0$ , respectively. We use the shorthand notation  $\gamma^n = (\gamma_1, \dots, \gamma_m)$  and  $\delta^n = (\delta_{m+1}, \dots, \delta_n)$ , where  $m = n/2$ . See Figure 2 for an illustration.

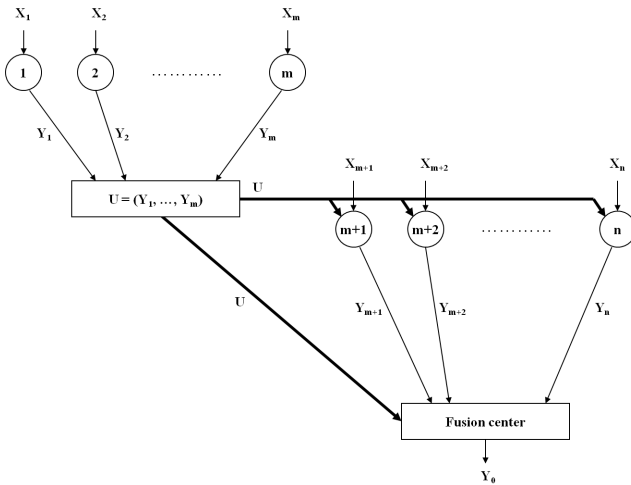


Fig. 2. The daisy chain architecture.

Similar to the star architecture, we define the probabilities of false alarm and missed detection associated with a collection  $(\gamma_0, \gamma^n, \delta^n)$  of decision rules by

$$\begin{aligned} H_n^I(\gamma_0, \gamma^n, \delta^n) &= \mathbf{P}_0(Y_0 = 1), \\ H_n^{II}(\gamma_0, \gamma^n, \delta^n) &= \mathbf{P}_1(Y_0 = 0). \end{aligned}$$

We define

$$q_n(\gamma_0, \gamma^n, \delta^n) = \frac{1}{n} \cdot \log H_n^{II}(\gamma_0, \gamma^n, \delta^n), \quad (3)$$

and, for every  $\alpha \in (0, 1)$ ,

$$Q_n(\alpha) = \inf_{\gamma_0, \gamma^n, \delta^n} q_n(\gamma_0, \gamma^n, \delta^n), \quad (4)$$

where the infimum is taken over all  $(\gamma_0, \gamma^n, \delta^n) \in \Gamma_{0,n} \times$

$\Gamma^n \times \Delta_n^n$  for which  $H_n^I(\gamma_0, \gamma^n, \delta^n) \leq \alpha$ . We finally define the optimal error exponent, denoted by  $g_d^*(\alpha)$ , by

$$g_d^*(\alpha) = \liminf_{n \rightarrow \infty} Q_n(\alpha).$$

It should be clear that the daisy chain architecture is “more powerful” than the star architecture: sensors  $m+1, \dots, n$  are free to ignore the additional information  $U$  that they receive and emulate any possible collection of decision rules for the star architecture. For this reason, for every finite  $n$ , the optimal missed detection probability  $Q_n(\alpha)$  in the daisy chain is no larger than the optimal missed detection probability  $R_n(\alpha)$  in the star configuration. By taking the limit, it follows that

$$g_d^*(\alpha) \leq g_p^*, \quad \forall \alpha \in (0, 1).$$

For any finite  $n$ , it will generically be the case that  $Q_n(\alpha) < R_n(\alpha)$ , because the additional information available in the daisy chain can be exploited to some advantage. On the other hand, our main result, proved in the next section, shows that the advantage disappears in the asymptotic regime.

### III. NO GAIN FROM FEEDBACK

Our main result asserts that the optimal error exponent (as the number of sensors increases) for the daisy chain architecture is no better than that of the star configuration.

**Theorem III.1.** *For every  $\alpha \in (0, 1)$ , we have  $g_d^*(\alpha) = g_p^*$ .*

*Proof:* As discussed at the end of the previous section, the inequality  $g_d^*(\alpha) \leq g_p^*$  is immediate. We only need to prove the reverse inequality. Toward this purpose, we consider an arbitrary choice of decision rules for the daisy chain architecture, and develop a lower bound on the probability of missed detection (for finite  $n$ ), in terms of the missed detection probability for the star architecture.

Throughout we fix some  $\alpha \in (0, 1)$ , and also an auxiliary parameter  $\epsilon > 0$ . Let us also fix  $n$  and decision rules  $\gamma_0, \gamma^n, \delta^n$ . Having done that, all of the random variables  $Y_i$  and  $U$  are well-defined. We assume that  $H_n^I(\gamma_0, \gamma^n, \delta^n) = \mathbf{P}_0(Y_0 = 1) \leq \alpha$  and we will derive a lower bound on  $H_n^{II}(\gamma_0, \gamma^n, \delta^n) = \mathbf{P}_1(Y_0 = 0)$ .

Let<sup>1</sup>

$$L = \log \frac{\mathbf{P}_0(U)}{\mathbf{P}_1(U)} = \sum_{i=1}^m \log \frac{\mathbf{P}_0(Y_i)}{\mathbf{P}_1(Y_i)}.$$

This is the log-likelihood ratio (a random variable) associated with the vector  $U$  of messages transmitted by the first  $m$  sensors. The first equality above is the definition of  $L$ , and the second follows from the definition  $U = (Y_1, \dots, Y_m)$  and the independence of the  $Y_i$ . Let  $\mu_0 = \frac{1}{m} \mathbf{E}_0[L]$ . By comparing with the definition (1), and using also Eq. (2), we have and note that

$$\mu_0 = \frac{1}{m} \sum_{i=1}^m D(\gamma_i) \leq -g_p^*. \quad (5)$$

<sup>1</sup>With some abuse of notation, we use  $\mathbf{P}_0(U)$  to denote the random variable that takes the numerical value  $\mathbf{P}_0(U = u)$  whenever  $U = u$ , and similarly for  $\mathbf{P}_0(Y_i)$ , etc.

We say that a possible value  $u$  of the random vector  $U$  is “normal” (symbolically,  $u \in N$ ), if

$$|L - \mu_0| \leq \epsilon m.$$

Because of Assumption II.1, and as pointed out in [11], the (conditionally independent) random variables  $\log(\mathbf{P}_0(Y_i)/\mathbf{P}_1(Y_i))$  have second moments that are bounded above (under  $\mathbf{P}_0$ ) by some absolute constant  $c$ . Thus, the variance of  $L$  (under  $\mathbf{P}_0$ ) is bounded above by  $cm$ . Chebyshev’s inequality then implies that

$$\mathbf{P}_0(U \notin N) \leq \frac{c}{\epsilon^2 m}.$$

We assume that  $m$  is large enough so that

$$\mathbf{P}_0(U \notin N) \leq \frac{1 - \alpha}{2(1 + \alpha)}. \quad (6)$$

Let us also say that a possible value  $u$  of the random vector  $U$  is “good” (symbolically,  $u \in G$ ) if

$$\mathbf{P}_0(Y_0 = 1 | U = u) \leq \frac{1 + \alpha}{2}.$$

We let  $B$  (for “bad”) be the complement of  $G$ .

Since  $\mathbf{P}_0(Y_0 = 1) \leq \alpha$ , we have

$$\begin{aligned} \alpha &\geq \mathbf{P}_0(Y_0 = 1) \\ &= \sum_u \mathbf{P}_0(U = u) \mathbf{P}_0(Y_0 = 1 | U = u) \\ &\geq \sum_{u \in B} \mathbf{P}_0(U = u) \mathbf{P}_0(Y_0 = 1 | U = u) \\ &\geq \frac{1 + \alpha}{2} \cdot \sum_{u \in B} \mathbf{P}_0(U = u) \\ &= \frac{1 + \alpha}{2} \cdot \mathbf{P}_0(U \in B). \end{aligned}$$

Thus,

$$\mathbf{P}_0(U \notin G) = \mathbf{P}_0(U \in B) \leq \frac{2\alpha}{1 + \alpha}. \quad (7)$$

Using Eqs. (6) and (7), we obtain

$$\begin{aligned} \mathbf{P}_0(U \in N \cap G) &\geq 1 - \mathbf{P}_0(U \notin N) - \mathbf{P}_0(U \notin G) \\ &\geq 1 - \frac{1 - \alpha}{2(1 + \alpha)} - \frac{2\alpha}{1 + \alpha} \\ &= \frac{1 - \alpha}{2(1 + \alpha)} > 0. \end{aligned} \quad (8)$$

We will now argue that whenever  $U$  takes a value in  $G$ , the missed detection probability admits a  $O(e^{mg_p^*})$  lower bound. Suppose that a certain value  $u$  of the random vector  $U$  has been realized. Conditioned on this event, and for this given value of  $u$ , the final decision is determined by a rule of the form

$$Y_0 = \gamma_0(u, \delta_{m+1}(X_{m+1}, u), \dots, \delta_n(X_n, u)).$$

Since  $u$  has been fixed to a constant, this is of the form

$$Y_0 = \bar{\gamma}_0(\bar{\delta}_{m+1}(X_{m+1}), \dots, \bar{\delta}_n(X_n)),$$

for suitable functions  $\bar{\gamma}_0$  and  $\bar{\delta}_i$ . (Of course these functions depend on the specific choice of  $u$ .) We recognize this as the expression for  $Y_0$  in a decentralized detection problem with a

star architecture and  $m$  sensors.

For the constant value of  $u$  under consideration, and since  $u \in G$ , the false alarm probability is bounded above by  $\alpha' = (1 + \alpha)/2 < 1$ . We now invoke the definition of  $g_p^*$ , suitably translated to the case of a finite number of sensors. It implies that for the given  $\alpha' \in (0, 1)$ , there exists some  $n_0$  (depending on  $\epsilon$  and  $\alpha'$ ) such that if  $J_m^I(\gamma_0, \gamma^m) \leq \alpha'$ , then

$$J_m^{II}(\gamma_0, \gamma^m) \geq e^{mg_p^* - m\epsilon}, \quad \forall m \geq n_0,$$

where  $m$  is the number of sensors in the star architecture. By applying this observation to the last  $m$  sensors of the daisy chain architecture, and conditioned on  $U = u$ , we obtain, for  $m \geq n_0$ ,

$$\mathbf{P}_1(Y_0 = 0 | U = u) \geq e^{mg_p^* - m\epsilon}, \quad \forall u \in G. \quad (9)$$

Let us now suppose that  $u \in N$ . From the definition of the log-likelihood ratio  $L$ , we have

$$\mathbf{P}_1(U) = e^{-L} \mathbf{P}_0(U).$$

When  $U \in N$ , we also have  $L \leq m\mu_0 + m\epsilon$ , and using also Eq. (5), we obtain

$$\begin{aligned} \mathbf{P}_1(U = u) &\geq e^{-m\mu_0 - m\epsilon} \mathbf{P}_0(U = u) \\ &\geq e^{mg_p^* - m\epsilon} \mathbf{P}_0(U = u), \quad \text{if } u \in N. \end{aligned} \quad (10)$$

We now apply the usual change of measure argument. We have, for  $m \geq n_0$ ,

$$\begin{aligned} \mathbf{P}_1(Y_0 = 0) &= \sum_u \mathbf{P}_1(U = u) \mathbf{P}_1(Y_0 = 0 | U = u) \\ &\geq \sum_{u \in N \cap G} \mathbf{P}_1(U = u) \mathbf{P}_1(Y_0 = 0 | U = u) \\ &\geq e^{mg_p^* - m\epsilon} \\ &\quad \cdot \sum_{u \in N \cap G} \mathbf{P}_0(U = u) \mathbf{P}_1(Y_0 = 0 | U = u) \\ &\geq e^{mg_p^* - m\epsilon} \sum_{u \in N \cap G} \mathbf{P}_0(U = u) e^{mg_p^* - m\epsilon} \\ &= e^{ng_p^* - n\epsilon} \mathbf{P}_0(U \in N \cap G) \\ &\geq e^{ng_p^* - n\epsilon} \frac{1 - \alpha}{2(1 + \alpha)}. \end{aligned}$$

Here, the second inequality follows from (10), and the third inequality from (9). The next equality is because  $n = 2m$ . The last inequality follows from (8).

Taking logarithms, dividing by  $n$ , and taking the limit as  $n \rightarrow \infty$ , we obtain that  $g_d^*(\alpha) \geq g_p^* - \epsilon$ . Since  $\epsilon$  was arbitrary, the desired result follows. ■

#### IV. CONCLUSION

We have proved that the daisy chain architecture introduced and analyzed in this paper performs no better (in the sense of the asymptotically optimal Neyman-Pearson error exponent) than a star architecture with the same number of sensors and observations. This is despite the fact that the daisy chain architecture provides substantially richer information to the last half of the sensors. To the authors’ knowledge, this is the first non-trivial non-tree architecture for which a precise

comparison and determination of the error exponent has been possible.

This work opens up a number of research questions. One involves the case of a Bayesian performance criterion: we are given prior probabilities for the two hypotheses (which do not change with the number of sensors) and wish to minimize the overall probability of error. From earlier works [12], [3], we know that Bayesian formulations can be qualitatively different. In particular, while for certain classes of trees the Neyman-Pearson error exponent is the same for tree and star architectures, this is not necessarily the case for the Bayesian error exponent. Determining whether feedback provides some added value (in terms of the optimal Bayesian error exponent) in the daisy chain architecture remains an open problem.

More generally, the range of possible non-tree or feedback architectures is vast. While such architectures tend to lead into intractable problems (as far as the optimal exponent is concerned), it may be possible to identify some that are tractable or to carry out some qualitative comparisons.

#### REFERENCES

- [1] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, Aug. 2004.
- [2] —, "Wireless sensors in distributed detection applications," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 16–25, May 2007.
- [3] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Data fusion trees for detection: does architecture matter?" *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4155–4168, Sep. 2008.
- [4] R. R. Tenney and J. N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, pp. 501–510, 1981.
- [5] J. N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.
- [6] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, "Optimization of detection networks: Tree structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, pp. 211–231, 1993.
- [7] O. P. Kreidl and A. S. Willsky, "An efficient message-passing algorithm for optimizing decentralized detection networks," *IEEE Trans. Autom. Control*, vol. 55, pp. 563–578, 2010.
- [8] 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, Jan. 1995.
- [9] —, "Decentralized bayesian detection with feedback," *IEEE Trans. Syst., Man, Cybern.*, vol. 26, no. 4, pp. 503–513, Jul. 1996.
- [10] P. F. Swaszek and P. Willett, "Parley as an approach to distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 447–457, Jan. 1995.
- [11] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signal, and Systems*, vol. 1, pp. 167–182, 1988.
- [12] W. P. Tay, "Decentralized detection in resource-limited sensor network architectures," PhD Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, February 2008.
- [13] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
- [14] D. A. Pantos, K. W. Halford, D. Kazakos, and P. Papanoni-Kazakos, "Distributed binary hypothesis testing with feedback," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, no. 1, pp. 21–42, Jan. 1995.
- [15] S. I. Zoumpoulis, "Decentralized detection in sensor network architectures with feedback," MEng Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, May 2009.