

# Cross-Layer Design of Sequential Detectors in Sensor Networks

Stefano Marano, Vincenzo Matta, Peter Willett, *Fellow, IEEE*, and Lang Tong, *Fellow, IEEE*

**Abstract**—A network of sensors polled by a mobile agent (the SENMA paradigm) is used for detection purposes, with both the remote nodes and the mobile agent implementing Wald’s sequential tests. When polled, each remote node transmits its local decision (if any) to the agent, and two network/agent communication schemes are considered. One of these is designed with specific care to the network’s energy consumption. In both cases, collisions over the common communication channel are precluded by the sequentiality of the sensors’ query. The system performances in terms of average decision time, error probability, and network energy consumption are derived in exact analytical form. A tradeoff exists between the *amount* and the *reliability* of the information that the rover may collect: At optimality, the decentralized system overcomes a single supernode by orders of magnitude in terms of decision time, while only 30% of the sensors encountered by the mobile agent spend energy to reveal themselves. The remaining sensors contribute to the detection process by their silence.

**Index Terms**—Cross-layer design, SENMA, Wald’s sequential test, wireless sensor networks.

## I. INTRODUCTION

THE pervasive advent of sensor networks is leading to a strong trend toward using the aggregate of many cheap and possibly unreliable sensors as opposed to single “supernodes” (see [15] and references therein). The recent literature is rich with contributions addressing the new technical challenges posed by this novel technology (examples are offered in, e.g., [3] and [29]), and in particular with detection problems in such decentralized architectures (see, for instance, [1], [2], [17], [23], [25], and [26]).

As to the latter issue, a recent paper by Willett and Tong [28] considers a detection problem in a sensor network under severe communication constraints. They propose a cross-layer design in which the remote nodes of the network observe a common feature and implement a likelihood ratio (LR) test. A traveling “rover” (or mule, see [18]) polls the sensors, collects binary output variables, and fuses these into a final decision about two

mutually exclusive statistical hypotheses. The network architecture is known as SENMA (Sensor Network with Mobile Agents) [24] and is attracting considerable interest due to its simplicity, scalability, robustness and energy efficiency.

To mitigate even SENMA’s communication burden, it was assumed in [28] that the local decisions (binary variables) are transmitted to the rover only if the pertinent LR exceeds some suitably optimized thresholds. That is, as long as the local (remote) LR falls in the region between two thresholds (perhaps the *dead-zone*), the transmission towards the rover is inhibited even if the node is polled for. If, instead, the LR exceeds an upper (respectively, lower) threshold, then a symbol  $+1$  (respectively,  $-1$ ) is sent. Actually, in complying with (a simple modification of) the ALOHA protocol, the  $\pm 1$ s are sent only with a certain probability  $p_t$ , to further reduce the occurrence of data collisions. The tradeoff between the amount of information (hence, detector performances) collected by the rover and the corresponding communication burden is accounted for by optimizing over the local thresholds and over the choice of  $p_t$ . Key features of the design in [28] are the presence of a dead-zone at the local sensors (an approach that can be traced back to [16]), and the choice of the sensor to remain silent, with probability  $1 - p_t$ , even though data are believed informative (i.e., the LR is out of the dead-zone). In this paper, we build on the above.

The dead-zone assigned to the local likelihood ratios is reminiscent of sequential testing, first formalized by Wald in 1947 [27]. Wald’s sequential detection is known to be optimal in the sense of achieving a given performance (detection and false alarm probabilities) with, on the average, fewer samples than required by any other procedure, including the fixed-sample-size (FSS) test (see, for instance, [14] and [19]). Such average optimality holds under nominal conditions (for instance, for completely specified signal and noise models), and does not exclude the occurrence of occasional very long runs. However, previous literature addressing practical cases of signals with unknown parameters, and the possibility of truncating the test to avoid excessively long runs (see, e.g., [20] and [21]), do suggest that modification of the original sequential procedure may be appealing in some practical contexts. In this paper, we refer to a rather idealized scenario in which no unknown parameters are considered, leaving the investigations of possible generalizations for future study.

Even though we pay no explicit attention to that portion of the network’s energy consumption that is due to the processing burden at the sensors, it is clear that the sequential procedure *at the remote-node level* allows for energy saving, in that it avoids data processing beyond necessity. Further, since the sequential procedure automatically terminates when an appropriate

Manuscript received January 12, 2005; revised December 19, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Feng Zhao. This work was partially supported by the Office of Naval Research. Part of this work was presented at the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), New York, June 5–8, 2005.

S. Marano and V. Matta are with the Department of Information and Electrical Engineering, University of Salerno, I-84084 Fisciano (SA), Italy (e-mail: marano@unisa.it; vmatta@unisa.it).

P. Willett is with Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT 06269 USA (e-mail: willett@engr.uconn.edu).

L. Tong is with the Electrical and Computer Engineering Department, Cornell University, Ithaca, NY 14853 USA (e-mail: ltong@ece.cornell.edu).

Digital Object Identifier 10.1109/TSP.2006.880254

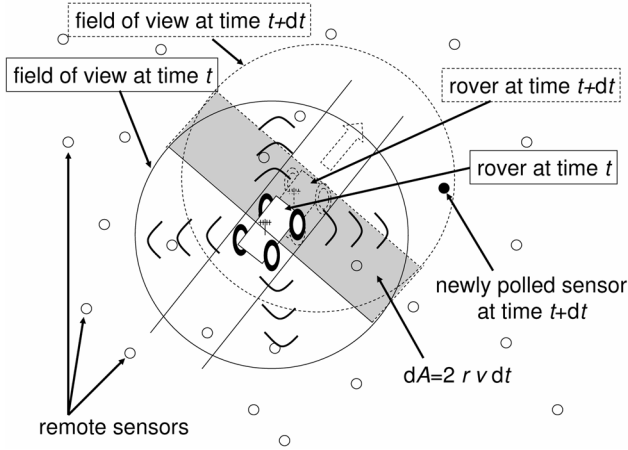


Fig. 1. Notional sketch of the scenario that we consider, of a traveling rover that polls sensors inside its circular field of view. In an interval  $dt$ , an incremental area of size  $dA$  is explored.

amount of information has been collected, implementing a sequential test *at the rover level* provides us with time savings.

The above motivates us to design a decentralized detector implementing sequential tests at the remote nodes, as well as at the rover. A distinctive feature of our setup, which differs from the approach in [28], is the rover/sensors communication protocol (to be detailed shortly), designed to avoid any channel collision concern and, in one case, with care to the sensors' energy-consumption. We maintain the analysis at an idealized level, with simplifying assumptions including the following.

- All the sensors observe the same feature and all of them start their tasks at a common time.
- The hypothesis test is of a shift-in-mean in Gaussian noise.
- The noise processes at different sensors are independent of each other, so that remotely collected observations are conditionally independent given the hypothesis.
- Sensors are randomly located in an unlimited area according to a homogeneous Poisson field model.

We warn the reader that many of these could be relaxed, and in fact they *should be* for modeling practical systems. But, as often happens, this is paid in the coin of analytical tractability: We have nice results, and we fear that not all their elegance would be preserved to more general situations.

The next brief section introduces the basic model, to be more deeply detailed separately in Sections III–V. Section VI deals with the system optimization, and concluding remarks are offered in Section VII. Appendix I summarizes results from the literature that are relevant to our analysis, for self-consistency; finally, Appendix II addresses a comparison with a fixed-length detection scheme. A short version of this work can be found in [10].

## II. MODEL

With reference to Fig. 1, we have a large network of sensors monitoring a common feature. The random sensor-location process obeys a homogeneous Poisson field law whose intensity per unit of area is denoted by  $S$ . The rover inspects the surveyed region, cumulatively a larger and larger swath. Specifically, during its “walk” across the area, a polling signal is contin-

uously emitted<sup>1</sup> to *awaken* sensors lying inside its field of view, here taken as circular. Accordingly, the total number of polled sensors at time  $t$  is a Poisson counting process with intensity per unit of time  $\lambda = 2Srv$ , where  $r$  is the radius of the rover's field of view,  $v$  is the constant velocity of the mobile agent, and  $dA = 2rv dt$  the incremental area.

Assume now that the  $s$ th sensor observes a realization of the process  $\{n_t^{(s)} + i\theta, t \in (0, \infty)\}$ , where  $n_t^{(s)}$  is a zero-mean Gaussian white random noise with power spectral density  $\sigma^2$ , and  $\theta$  is a known parameter ruling the shift-in-mean. Under the assumption that noises at different sensors are independent of each other, it is desired to decide which of the two mutually exclusive statistical hypotheses,  $i = 0, 1$ , is in force. A more convenient setup is obtained by considering the equivalent problem in which the observations at time  $t$ , say  $x_t^{(s)}$ , are integrated versions of the above:  $x_t^{(s)} = \int_0^t [n_\tau^{(s)} + i\theta] d\tau$ . Formally

$$H_i : x_t^{(s)} = w_t^{(s)} + i\theta t, \quad t \in (0, \infty), \quad s = 1, 2, \dots,$$

where  $i = 0, 1$ , and  $w_t$  is a Brownian motion with variance  $\sigma^2 t$ .

The final decision about  $H_i$ ,  $i = 0, 1$ , is requested from the rover. As said, each remote node implements its own sequential test; when polled, the sensor may or may not transmit a symbol, depending on the strength of its local observations. The rover bases its final decision upon a sequential procedure, using as observables these received symbols. The following three sections detail in due depth the proposed design.

## III. REMOTE SENSORS' SPRTS

We simplify the notation by omitting the superscript  $(s)$  used above. The generic node observes  $\{x_t, t \in (0, \infty)\}$  and, at each time instant, builds up the corresponding log-LR, as provided by the Cameron–Martin formula (see, e.g., [14])

$$l_t = \theta x_t - \theta^2 t / 2. \quad (1)$$

This is used to implement a (local) Sequential Probability Ratio Test (SPRT), which can be formalized as follows:

$$l_t \begin{cases} \geq \gamma_1 \Rightarrow \text{choose (locally) } H_1 \\ \leq \gamma_0 \Rightarrow \text{choose (locally) } H_0 \\ \in ]\gamma_0, \gamma_1[ \Rightarrow \text{go ahead.} \end{cases} \quad (2)$$

The local (random) time  $T_{1oc}$  at which the test terminates—eventually, with probability one—is a *stopping time* for the sequential test [19].

Let us denote by  $p_d$  and  $p_f$  the probabilities that a local test (i.e., one running over a generic sensor of the network) terminates with a choice in favor of  $H_1$ , in the case that the true state of the nature is  $H_1$  and  $H_0$ , respectively. We have

$$\gamma_0 = \log \frac{1 - p_d}{1 - p_f}, \quad \gamma_1 = \log \frac{p_d}{p_f}. \quad (3)$$

In discrete-time testing an “excess over the boundary” may exist. In that case the above relationships hold true only approximately (whence *Wald's approximations*). In our case, as

<sup>1</sup>This is in contradistinction to the “usual” SENMA concept of [24] in which the poll is periodic over non-intersecting regions.

the random process  $l_t$  is time-continuous and with continuous paths, the (3) are exact.

Let us compute some key quantities. First, let<sup>2</sup>  $F_i(t) = \Pr[T_{\text{loc}} \leq t | H_i]$  be the probability of having reached a *local* decision by time  $t$ , under hypothesis  $H_i$ ,  $i = 0, 1$ . This cumulative distribution function (henceforth, CDF) of the stopping time  $T_{\text{loc}}$ , can be found by resorting to the following result due to Tartakovsky [22]: for  $i = 0, 1$ ,

$$F_i(t) = 1 - \sum_{k=1}^{\infty} B_{ki} \exp \left\{ -\frac{\rho t}{2} \left[ A_i^2 + \left( \frac{\pi k}{\gamma_1 - \gamma_0} \right)^2 \right] \right\} \sin \frac{\pi k \gamma_0}{\gamma_0 - \gamma_1} \quad (4)$$

where  $\rho \triangleq \theta^2 / \sigma^2$ , and

$$B_{ki} = \frac{2\pi k}{(\gamma_1 - \gamma_0)^2} [\exp(A_i \gamma_0) + (-1)^{k+1} \exp(A_i \gamma_1)] \left[ A_i^2 + \left( \frac{\pi k}{\gamma_1 - \gamma_0} \right)^2 \right]^{-1} \quad (5)$$

with  $A_0 = -0.5$  and  $A_1 = 0.5$ . In spite of its apparent bulkiness, Tartakovsky's formula is exact, simple to implement numerically and converges reasonably fast: Some hundreds of terms make the truncation error negligible for our purposes.

Consider then the probability  $\pi_{ji}(t)$  that a generic remote sensor has reached its decision at time  $t$ , and that this decision is in favor of  $H_j$  while  $H_i$  is actually true; for instance,  $\pi_{01}(t) = \Pr\{l_t \leq \gamma_0, T_{\text{loc}} \leq t | H_1\}$ . Along the same lines as in the original development by Wald [27], one obtains the following relationships:

$$\begin{aligned} \pi_{11}(t) &= e^{\gamma_1} \pi_{10}(t) \\ \pi_{01}(t) &= e^{\gamma_0} \pi_{00}(t) \end{aligned}$$

where obviously  $\sum_{j=0}^1 \pi_{ji}(t) = F_i(t)$ . As a sanity-check, note that in the limit of  $t \rightarrow \infty$ , the above reduce to (3). Solving the previous system, and accounting for (3), we obtain

$$\begin{aligned} \pi_{10}(t) &= p_f G(t) \\ \pi_{11}(t) &= p_d G(t) \end{aligned} \quad (6)$$

( $\pi_{01}$  and  $\pi_{00}$  immediately follow from these), wherein

$$G(t) = \frac{F_1(t)(1 - p_f) - F_0(t)(1 - p_d)}{p_d - p_f}.$$

For simplicity of analysis, in the following we confine our attention to the symmetric case, that is  $p_f = 1 - p_d \triangleq p_e$ . Accordingly, the relevant formulas reduce to

$$\pi_{00}(t) = \pi_{11}(t) = (1 - p_e)F(t) \quad (7)$$

$$\pi_{10}(t) = \pi_{01}(t) = p_e F(t) \quad (8)$$

<sup>2</sup>Time dependence is indicated equivalently as a subscript or in parentheses, for notational convenience.

where

$$F(t) \triangleq F_1(t) = F_0(t) = G(t). \quad (9)$$

For later use, it is convenient to have a single symbol, say  $\beta$ , to denote alternatively  $p_e$  or  $1 - p_e$ . Thus, we define

$$\beta_i \triangleq p_e^i (1 - p_e)^{1-i} \quad (10)$$

with  $i = 0, 1$ , denoting the true hypothesis.

#### IV. ROVER/SENSORS COMMUNICATION

We investigate two communication/detection strategies. In the first, we assume that each remote sensor, when polled by the traveling rover, transmits a symbol  $+1$  (respectively,  $-1$ ) if it has reached a decision in favor of  $H_1$  (respectively,  $H_0$ ). If no decision has been taken when polled, nothing is transmitted.

Before making its decision, a node is in fact in *sleep* mode with its transceiver inactive. The energy consumption is only due to the sensing/LR-computation processes. If a polling signal is received, the sensor is simply *turned off*, thus never revealing itself to the rover.

Conversely, once a local decision is made, the local sensing/LR-computation tasks are inhibited and the node persists in its sleep mode, waiting for being eventually polled. Now, if it is polled for, the transmission toward the rover is activated. This may happen only once because, after sending its local decision, the sensor is turned off and it never contributes more: There is only *one chance* to transmit.

In what follows we assume i) that sensing/signal processing requires substantial less energy than communications, and ii) that the energy cost of the sleep mode can be neglected with respect to that required for communications (see also [24]). Under these hypotheses, the main source of energy consumption arises from revealing the local decisions to the rover, and the network energy burden can be accordingly computed in terms of the number of such deliveries.

The above communication/detection strategy can be regarded as a three-level quantization of the log-LR. Alternatively, we also consider the case in which, as before, a  $+1$  is transmitted if the decision taken is in favor of  $H_1$ . Otherwise, if the sequential test is still running when polled, or if the decision has been taken in favor of  $H_0$ , the node remains silent. This scheme is referred to as the two-level-quantization scheme, and the idea behind it is that of exploring potential benefits arising from silencing an additional fraction of nodes. Clearly, as before, once being polled by the rover, the queried sensors are turned off.

As to the rover, it collects the symbols emitted by the polled nodes. Specifically, at given time  $t$  and position  $(x_t, y_t)$  in a given system reference, the mobile agent polls sensors inside its field of view, and waits for their responses (symbols). Then, at time  $t + dt$ , it moves to a new position  $(x_{t+dt}, y_{t+dt})$ , polls again, and so on. The rover detection strategy relies upon the collected symbols up to the current time, as detailed later. As for the remote sensors, we find it convenient to adopt a continuous model for the rover, in which the polling and receiving times are ignored. This amounts to assume that the query/listening times are substantially smaller than the time interval  $dt$  needed to the

rover for moving to a new position,<sup>3</sup> which, in turn, is much smaller than  $1/\lambda$ .

It is to be emphasized that, accordingly to the Poisson assumption, in a time interval of  $dt$  at most *one* new sensor falls inside the field of view. Furthermore, as said earlier, only newly encountered sensors are *not turned off* and may communicate with the mobile agent. That is to say, for any given instant of time in the designed system at most one sensor is allowed to communicate: We are clearly avoiding any collision concern pertaining to the channel random access.

Finally, we would like to emphasize that transmitting the local decisions is only one of the many conceivable system designs. An advantage of this scheme is the possibility of employing *non-standard* analog transmission protocols, where sensors simply emit positive/negative short pulses, or *beeps*, as proposed in [11] and [12]. If standard digital communication schemes are to be used, the proposed scheme inherently embodies savings in terms of bandwidth and sensors/rover processing. However, if the transmission process invokes a standard wireless communications protocol—with packets already burdened by tens of header bits—then, presumably, more refined information (e.g., finer quantized version of the local LRs, amount of threshold overshoot, local stopping times) really ought to be sent to the rover. Taking this approach to an extreme, one could conceive a system in which the unquantized LR is sent, or nothing (see, e.g., [16]). These possible extensions are left for future work: Here, we have single-bit information flow between sensor and rover.

## V. ROVER'S SPRT

### A. Three-Level Quantization

1) *The Test*: Recall that the probability of encountering a sensor in an infinitesimal interval of time  $dt$  is  $\lambda dt + o(dt)$ . Assuming  $H_i$  is true, the probability of receiving  $+1$  at the rover is  $\lambda(1 - \beta_i)F(t)dt + o(dt)$ , and that of receiving  $-1$  amounts to  $\lambda\beta_i F(t)dt + o(dt)$ . Thus, denoting with  $X_t$  (respectively,  $Y_t$ ) the total number of  $+1$ s (respectively,  $-1$ s) received up to time  $t$ , a little thought reveals that these are two independent inhomogeneous Poisson processes.<sup>4</sup> Thus, defining

$$\omega(t) = \int_0^t F(\xi)d\xi \quad (11)$$

and using  $X \sim \text{Poi}(\chi)$  as a shortcut to state that  $X$  is a Poisson counting processes with mean value  $\chi$ , the statistical hypothesis test formalizes as follows:

$$H_i : [X_t \sim \text{Poi}(\lambda(1 - \beta_i)\omega(t)), Y_t \sim \text{Poi}(\lambda\beta_i\omega(t))]. \quad (12)$$

<sup>3</sup>In adopting a continuous-time model for the processes at the remote sensors, we have also inherently assumed that the nodes' sampling interval to acquire data is conveniently small.

<sup>4</sup>Intuitively, assume that a discrete model be in force. Then, the occurrence of the three symbols  $0$  (no signal),  $\pm 1$ , would obey a multinomial distribution, whose continuous limit is just the product of two Poisson densities (see, e.g., [13]).

Since  $X_t$  and  $Y_t$  are independent of each other, the pertinent log-LR takes the simple form

$$\begin{aligned} L_t &= X_t \log \frac{1 - \beta_1}{1 - \beta_0} - \lambda\omega(t)(\beta_0 - \beta_1) \\ &\quad + Y_t \log \frac{\beta_1}{\beta_0} - \lambda\omega(t)(\beta_1 - \beta_0) \\ &= (X_t - Y_t) \log \frac{1 - p_e}{p_e}. \end{aligned} \quad (13)$$

The statistical hypothesis test running at the rover is hence

$$L_t \begin{cases} \geq \Gamma_1 \Rightarrow \text{choose } H_1 \\ \leq \Gamma_0 \Rightarrow \text{choose } H_0 \\ \in ]\Gamma_0, \Gamma_1[ \Rightarrow \text{go ahead} \end{cases} \quad (14)$$

where  $\Gamma_0$  and  $\Gamma_1$  represent the two thresholds. The above test implicitly introduces the rover's stopping time  $T = \inf\{t : L_t \notin ]\Gamma_0, \Gamma_1[ \}$ .

2) *Performance*: Let  $P_d$  and  $P_f$  denote the system-level (i.e., at the rover) detection and false alarm probabilities. To complete the design and the analysis of our system we would derive now the relationships among the thresholds  $\Gamma_1$  and  $\Gamma_0$ , the relevant probabilities  $P_d$  and  $P_f$ , and the average stopping times  $E[T|H_i]$ . Another key performance figure—related to the system energy consumption—is the average number of symbols  $\pm 1$  globally emitted by the ensemble of sensors. We next compute these performance figures.

The statistical test (14) involves nonstationary observations, since the stopping times of the sensor-level SPRTs are not uniformly distributed in time, but instead occur according to (11). However, Wald's approximations such as (3) still hold true allowing, as usual, for possible threshold overshoots. Better, as the statistics in (13) take steps of size  $\Delta = \log(1 - p_e)/p_e$ , this latter concern can be avoided by an appropriate choice of the thresholds  $\Gamma_1 = m_1\Delta$  and  $\Gamma_0 = m_0\Delta$ , with  $m_i$  integers. This constrains the achievable pairs of  $(P_d, P_f)$  to a countable set, implicitly defined through

$$\Gamma_0 = \log \frac{1 - P_d}{1 - P_f}, \quad \Gamma_1 = \log \frac{P_d}{P_f} \quad (15)$$

and now the no-overshoot condition is fulfilled.

The computation of the average stopping time is more involved [6], [22]. A nice approach is to introduce a suitable time transformation such that the pertinent random processes, with respect to the new time scale, become homogeneous. In the case at hand, this transformation is suggested by (11). Applying the time substitution method [22] we get a new time axis  $\omega \triangleq \omega(t)$  and the associated homogeneous Poisson processes  $\tilde{X}_\omega = X_t$  and  $\tilde{Y}_\omega = Y_t$ , with intensities

$$I_{X_i} \triangleq \lambda(1 - \beta_i) \quad \text{and} \quad I_{Y_i} \triangleq \lambda\beta_i \quad (16)$$

respectively. The analog of (13), with respect to the new time scale, is

$$\tilde{L}_\omega = (\tilde{X}_\omega - \tilde{Y}_\omega) \log \frac{1 - p_e}{p_e}. \quad (17)$$

At this point, one is allowed to invoke the classical results by Wald for computing the expressions of the average (transformed) stopping time,  $E[\Omega]$ , where  $\Omega = \omega(T)$  is the stopping time in the new time scale.

To this end, let us introduce the formal derivative of the counting process  $\tilde{X}_\omega - \tilde{Y}_\omega$ : that is the *point* process  $\Xi(\omega) = \sum_k a_k \delta(\omega - \omega_k)$ , where  $a_k = \pm 1$ ,  $\omega_k$  is the time at which the  $k^{\text{th}}$   $\pm 1$  appears, and  $\delta(\cdot)$  is the Dirac distribution (impulse). We can formally rewrite (17) as an integral involving  $\Xi(\omega)$

$$\tilde{L}_\omega = \int_0^\omega \left[ \Xi(\xi) \log \frac{1-p_e}{p_e} \right] d\xi.$$

Evaluating the above at the stopping time  $\omega = \Omega$ , and assuming that a time-continuous version of Wald's equation holds,<sup>5</sup> one gets

$$E[\tilde{L}_\Omega | H_i] = (I_{X_i} - I_{Y_i}) \log \frac{1-p_e}{p_e} E[\Omega | H_i]. \quad (18)$$

Recognizing that the value assumed by the log-LR at the stopping time equals one of the two thresholds, and recalling the definition of the intensities  $I_{X_i}$  and  $I_{Y_i}$  (see (16)), (18) yields

$$\begin{aligned} E[\Omega | H_1] &= \lambda^{-1} \frac{D(P_d | P_f)}{D(1-p_e | p_e)} \\ E[\Omega | H_0] &= \lambda^{-1} \frac{D(P_f | P_d)}{D(p_e | 1-p_e)} \end{aligned} \quad (19)$$

where  $D(p_1 || p_2)$  denotes the Kullback–Leibler distance (or *divergence*) between the two *binary* probability mass functions  $[p_1, 1-p_1]$  and  $[p_2, 1-p_2]$  (see [4]). The two denominators above are equal, in fact  $D(p_e || 1-p_e) = D(1-p_e || p_e)$  and, for later use, we accordingly denote  $\mathcal{D}(\xi) \triangleq D(\xi || 1-\xi)$ ,  $0 < \xi < 1$ .

As to the stopping time  $E[T]$  (recall that we have mapped  $t \rightarrow \omega(t)$  according to (11)): It is tempting simply to invert the relation (11), but there is the problem that we must relate the expected value of a nonlinearly transformed variable to its original expectation. Fortunately, we can use Jensen's inequality to obtain the bound [4]

$$E[T | H_i] \leq \omega^{-1}(E[\Omega | H_i]) \quad i = 0, 1 \quad (20)$$

since  $\omega(t)$  in (11) is the integral of a CDF, and is hence a convex function. The reasonable tightness of this bound has been shown by extensive computer simulations, an instance of which is offered in Fig. 2. Left panel compares the simulation results with the formula  $E[T] \approx \omega^{-1}(E[\Omega])$ , showing a good fit. The tightness is due to the moderate curvature of  $\omega(t)$  in the region of relevance for the density of  $T$ ; this is illustrated in the right panel of the figure.

The last performance figure to be computed is the average number of symbols  $E[N]$  globally emitted by the ensemble of

<sup>5</sup>With “Wald's approximations” we refer to equations like (3) and (15). Instead Wald's *equation* quoted here is as follows: Consider  $Z_M = \sum_{k=1}^M W_k$ , where  $W_k$ s are independent and identically distributed random variables, and  $M$  a related stopping time. Under certain regularity conditions  $E[Z_M] = E[M]E[W_k]$ . We use an integral version of this identity, whose formal justification relies upon limit arguments (see, for instance, [9]).

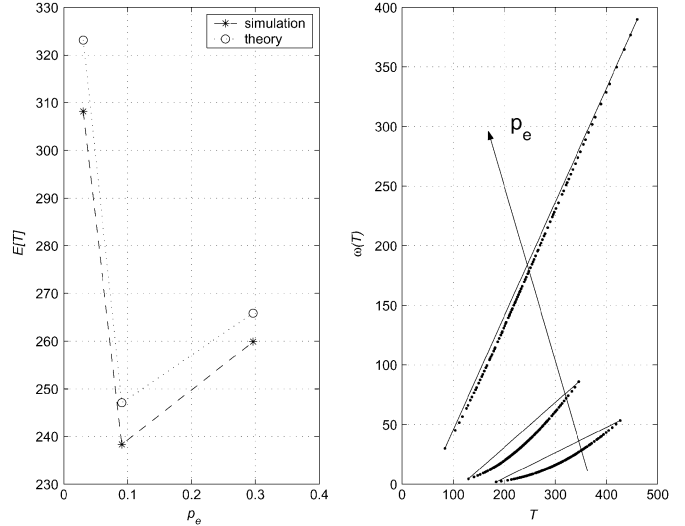


Fig. 2. *Left panel*: Simulated values of  $E[T]$  (denoted by  $*$ ) compared with the analytical bound  $\omega^{-1}(E[\Omega])$  (circles) for three sample values of  $p_e$ . Here it is assumed  $P_f = 1 - P_d = 10^{-3}$ ,  $\lambda = 10^{-1}$ ,  $\rho = 10^{-2}$ , and the results are valid under both the hypotheses. *Right panel*: The function  $\omega(T)$  computed at the values of  $T$  resulting from computer simulations, with values exceeding two times the standard deviation dropped (dots): We want to emphasize the region where the density of  $T$  is more relevant. The solid lines serve simply as reference and stress the moderate curvature of  $\omega(t)$  in the range of interest. Note that the larger  $p_e$  is (the same values of the left panel are considered), the tighter the bound is expected to be. This is confirmed by the numerical values given in the left panel.

sensors when the rover's detection process ends—it seems reasonable to assume that this is proportional to the total amount of energy spent by the network, given that transmitting  $\pm 1$  requires the same energy. We have  $E[N] = E[X_T + Y_T] = E[\tilde{X}_\Omega + \tilde{Y}_\Omega]$ , which is found to be<sup>6</sup>

$$E[N | H_i] = \lambda E[\Omega | H_i]. \quad (21)$$

This last equation has a nice interpretation if we assume that  $P_f = 1 - P_d \triangleq P_e$ , meaning that the rover's SPRT is symmetric. In this case (21) reduces to

$$E[N | H_i] = \frac{\mathcal{D}(P_e)}{\mathcal{D}(p_e)} \quad (22)$$

under both hypotheses. Given that  $P_e < p_e < 1/2$  (in practice one has  $P_e \ll p_e$ ), the smaller is  $P_e$  with respect to  $p_e$ , the larger is  $E[N | H_i]$ . This is quite obvious in that, if the global performances to be ensured by the network are much tighter than that at the single node's level, then the rover must collect a large number of local decisions. For  $P_e$  closer to  $p_e$  a few symbols collected by the rover allows for taking the final decision: In the limit case of  $P_e = p_e$ , just *one* symbol is necessary, as expected.

<sup>6</sup>If we were in a discrete setup the arguments would be as follows. According to its definition, the (assumed discrete) stopping time  $\Omega$  has the property that the sequence  $\{\tilde{X}_\omega - \tilde{Y}_\omega, \omega < \Omega\}$  specifies whether  $\{\Omega \geq \omega\}$  or not. As a consequence (being  $\tilde{X}_\omega$  and  $\tilde{Y}_\omega$  independent of each other) the event  $\{\Omega \geq \omega\}$  is independent of the sequence  $\{\tilde{X}_\omega, \omega \geq \Omega\}$  as well as of  $\{\tilde{Y}_\omega, \omega \geq \Omega\}$ . This implies that  $\Omega$  is a valid stopping time for each of the sequences  $\tilde{X}_\omega$ ,  $\tilde{Y}_\omega$  and  $\tilde{X}_\omega + \tilde{Y}_\omega$ , according to the more general stopping-time characterization, as given for instance by Gallager [8, p. 66]. We conclude that Wald's equation is applicable to the sum process and (21) follows, provided that similar arguments apply to the continuous-time case, as we assume here.

Before ending this section, we stress that the only approximations in the whole derivation so far is assuming equality in (20); as we have verified numerically, this is an excellent approximation. Consequently, we can state that *all the relevant performance figures of the designed system admit exact analytical expressions*: System performances at any degree of confidence can be computed exactly, and without resorting to computer simulations.

### B. Two-Level Quantization

Let us now assume that +1 is emitted by a remote node if and only if i) a local decision has been taken, and ii) such decision is in favor of  $H_1$ . The total number of received symbols at the rover by the time  $t$  is the previously defined process  $X_t$ . This is a Poisson counting process under both hypotheses and is presumably more parsimonious with sensor communication resource than is the three-level scheme. Let us see.

As before, it is appropriate to apply the time transformation in (11) for getting homogeneity. The log-LR of the transformed  $\tilde{X}_\omega$  is<sup>7</sup>

$$\tilde{L}_\omega = \tilde{X}_\omega \log \frac{1-p_e}{p_e} - \lambda(1-2p_e)\omega. \quad (23)$$

At this point, one may be tempted to invoke the classical Wald's approximations both for computing the thresholds<sup>8</sup> and for deriving the expressions of the average (transformed) stopping time, and as before this would conclude the analysis.

Unfortunately, expression (23) involves a subtlety. Indeed, as pointed out by DeLucia and Poor [5]

Unlike the analogous problem in sequential testing between Wiener processes, the primary indexes of performances . . . associated with sequential test between Poisson processes do not admit simple expressions.

The issue may be traced back to 1953 [7] (see also [9]) and is the reason why we cannot simply resort to Wald's approximations for analyzing the rover's SPRT in the case at hand.

The basic difficulty with an SPRT of Poisson counting processes is simply understood by inspection of the log-LR structure: The event of a lower threshold crossing is necessarily due to the second term on the RHS of (23), which is a negative continuous function of  $\omega$ . Accordingly, the no-overshoot condition, required for Wald's approximation to be true exactly, is certainly met. Conversely, an upper threshold crossing is ruled by the first addend on RHS of (23), which increments at *steps* of  $\log(1-p_e)/p_e$ . As a consequence, occasional overshoots are unavoidable: Neglecting the excess over the upper boundary would require  $p_e$  extremely close to 1/2.

DeLucia and Poor in [5] addressed and solved the problem of deriving the exact performance figures for SPRTs of Poisson counting processes. In the following we resort to their powerful results that, not to fragment the exposition, are collected in Appendix I. Exploiting these results we can easily compute

<sup>7</sup>We use the same symbol  $\tilde{L}_\omega$  as in (17).

<sup>8</sup>As it should be clear, also (15) could be equivalently computed on the transformed log-LR (17), rather than on the original log-likelihood (13): Time warping (11) does not affect the *values* taken by the process, but only the time instants at which specific values are attained.

$P_f$  and  $P_d$  ((37) and (38)), as well as  $E[\Omega|H_i]$  ((39) and (40)), as functions of  $p_e, \Gamma_0, \Gamma_1$ , and  $\lambda$ . As before, Jensen's inequality provides us with the bound

$$E[T|H_i] \leq \omega^{-1}(E[\Omega|H_i]) \quad i = 0, 1 \quad (24)$$

and numerical investigations (not shown) have ensured its tightness.

Note that the computed  $E[\Omega]$  is again intimately related to the energy consumption (average number of symbols sent) of the network. In fact, in this case we have  $E[N] = E[X_T] = E[\tilde{X}_\Omega]$ , which at the light of the arguments in footnote 6, yields  $E[N|H_i] = \lambda(1-\beta_i)E[\Omega|H_i]$ , or

$$\begin{aligned} E[N|H_1] &= \lambda(1-p_e)E[\Omega|H_1] \\ E[N|H_0] &= \lambda p_e E[\Omega|H_0]. \end{aligned} \quad (25)$$

It is worth noting that the comments in the last paragraph of Section V-A apply also to the two-level scenario: Jensen's inequality is tight and all the other formulas are exact, even if a bit more involved.

Comparing (25) with (21), it seems that  $E[N]$  is now reduced by a factor of  $p_e$  or  $1-p_e$ , as a natural consequence of the two-level communication scheme. However, the values of  $E[\Omega]$  in these formulas are different, thus leading to a less trivial comparison which is addressed numerically in the next section.

## VI. SYSTEM OPTIMIZATION

In this section we first optimize the threshold parameters, and then derive the performances, of a sequential SENMA scheme. Such performances, although approximate, are quite accurate and, even better, admit dramatically simple analytical expressions. Both the optimization and the performance evaluation are easiest in the symmetric case where  $p_f = 1-p_d = p_e$  and  $P_f = 1-P_d = P_e$ , as it is assumed henceforth—this is not a necessity but merely a restriction that offers simplified presentation. Furthermore, we assume that the signal-to-noise ratio (SNR)  $\rho$  and the sensor density  $\lambda$  are given; also prescribed is a certain value of  $P_e$ , while the parameter to be optimized over is the local sensors' performance  $p_e$ . Among the many optimization options, we find of particular interest and practical relevance the following: minimize the average time  $E[T]$  for taking a decision.<sup>9</sup>

It is convenient to introduce two dimensionless quantities: a normalized time  $\rho t$  and a normalized density  $\lambda/\rho$ , where  $\rho = \theta^2/\sigma^2$ , already defined after (4), is the SNR *per unit of time*.<sup>10</sup> For notational convenience we further append a ' to denote averaged and/or normalized quantities. Thus, we let

- $T'_i \triangleq \rho E[T|H_i]$  (normalized average stopping time, at the rover);
- $T'_{loc,i} \triangleq \rho E[T_{loc}|H_i]$  (normalized average stopping time, at the sensors);

<sup>9</sup>It is perhaps worth noting that an optimization over  $E[N]$  would lead, as we see from (22), to the less interesting result of  $p_e = P_e$ .

<sup>10</sup>Physical relevance of the two normalized quantities follows from observing that i) (4) is only a function of  $p_e$  (through the local thresholds  $\gamma_1$  and  $\gamma_0$ ) and of the product  $\rho t$ ; and ii) in terms of the normalized time  $\rho t$ , the density  $\lambda$  of encountered sensors scales just as  $\lambda/\rho$ .

- $N'_i \triangleq E[N|H_i]$  (average number of transmissions  $\propto$  energy demand for communications);
- $\lambda' \triangleq \lambda/\rho$  (normalized sensor density).

In the above, the subindex  $i$  denotes the hypothesis in force and will be omitted when inessential, while  $T$  and  $N$  are the previously introduced stopping time and number of symbols; this latter, we recall, is proportional to the energy consumption.

We also introduce

- $T'_{\text{single}} \triangleq \rho E[T_{\text{single}}] = 2D(P_e)$ ,

that is the normalized average time that a single sensor would require to end up with a decision at the performance level  $P_e$ . Thus,  $T'/T'_{\text{single}}$  is the decision time for the SENMA system, with respect to a *single sensor "system"* that implements an SPRT at the same error probability. It is obvious that we want to make this ratio as small as possible.

We are now ready for system optimization: to minimize  $T'$  with respect to the design parameter  $p_e$ , for any prescribed global performance  $P_e$ . As discussed earlier, in the optimization procedure we consider the bounds (20) and (24) as attained with equality.

#### A. Best Sensor-Level $P_e$ for Three-Level Quantization

From (11), also accounting for (4), we have that

$$\omega(t; \rho, p_e) = \frac{\omega(\rho t; p_e)}{\rho} \quad (26)$$

where the functional dependence on the system parameters is explicit. Then, assuming that (20) holds true with equality, and in view of (26) and (21), one readily gets

$$\frac{\omega(T'; p_e)}{N'(p_e, P_e)} = \lambda'^{-1} \quad (27)$$

which implicitly defines a curve  $T'(p_e)$ , for prescribed  $\lambda'$  and  $P_e$ . The above relationship allows us to perform numerically the optimization by selecting a value of  $p_e$ , say  $p_e^*$ , as that attaining the minimum of  $T'$ . The top panel of Fig. 3 illustrates the procedure by depicting the curve  $T'(p_e)$ , for several  $\lambda'$  and a specific value of  $P_e$ . For clarity, we consider the ratio  $T'/T'_{\text{single}}$  that emphasizes the relative behavior with respect to a single sensor. Note that, in complying with the earlier discussion about integer threshold jumps (see comments preceding (15)), for any given  $P_e, p_e$  is constrained to belong to a countable set, viz.,

$$p_e = \frac{1}{1 + \left(\frac{1-P_e}{P_e}\right)^{1/J}}, \quad J = 1, 2, \dots,$$

so that in figure we have discrete points denoted by circles.

As stated, the analytical expressions of the system performances can be dramatically simplified by resorting to some approximation. In fact, we note the following:

- the optimal  $p_e^*$  minimizing the average global decision time is attained at values of  $T'$  always near the expected local decision time  $T'_{\text{loc}}$ , that is  $T' \approx T'_{\text{loc}}$ ;
- meaningful values of the ratio  $T'/T'_{\text{single}}$  (say,  $< 10^{-1}$ ), where using the network gives larger gains, are met in the range  $p_e^* \geq 0.1$ .

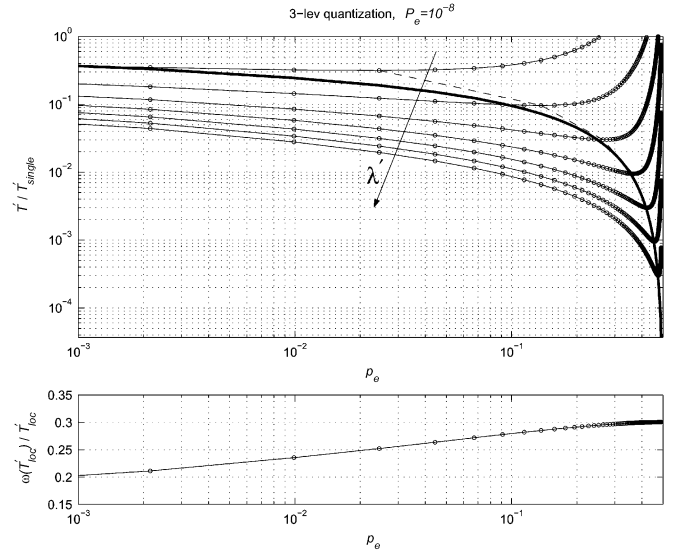


Fig. 3. *Top panel*: Circles represent  $T'/T'_{\text{single}}$  versus  $p_e$ , with  $\lambda'$  as parameter, where  $\lambda' = 10^m$ ,  $m = 0, 1, \dots, 6$ , grows as indicated by the arrow. Here  $P_e = 10^{-8}$ , but similar results are obtained with other small  $P_e$ 's. The abscissa attaining the minimum is selected as  $p_e^*$ , for each prescribed  $\lambda'$ , and the dashed line emphasizes the locus of such minima. The bold curve represents  $T'_{\text{loc}}/T'_{\text{single}}$ : We see that a minimum of  $T'/T'_{\text{single}}$  is approximately attained at the crossing point  $T' = T'_{\text{loc}}$ , and the approximation improves at large  $\lambda'$ . This motivates our analytical approximation (see main text). *Bottom panel*:  $\omega(T'_{\text{loc}}(p_e); p_e)/T'_{\text{loc}}(p_e)$  versus  $p_e$ . We see that, in the range of interest  $p_e > 0.1$ , an almost constant value of  $\approx 0.3$  arises, thus supporting the linear approximation, see discussion yielding (30).

Approximation i) allows us to equate  $T' = T'_{\text{loc}}$  in (27), getting

$$\frac{\omega(T'_{\text{loc}}(p_e^*); p_e^*)}{N'(p_e^*, P_e)} = \lambda'^{-1}. \quad (28)$$

From ii) and the bottom panel of Fig. 3, it follows that, in the specified range for  $p_e^*$ , the linear approximation

$$\omega(T'_{\text{loc}}(p_e^*); p_e^*) \approx \eta T'_{\text{loc}}(p_e^*) \quad (29)$$

holds for some  $\eta$ . Using this approximation, recalling  $T'_{\text{loc}}(p_e^*) = 2D(p_e^*)$  and accounting for (22), we get

$$D(p_e^*) = \sqrt{\frac{D(P_e)}{2\eta\lambda'}}. \quad (30)$$

It follows that

$$T' \approx 2D(p_e^*) = \sqrt{\frac{2D(P_e)}{\eta\lambda'}} \quad (31)$$

$$N' \approx \frac{D(P_e)}{D(p_e^*)} = \sqrt{2D(P_e)\eta\lambda'} \quad (32)$$

$$T'N' \approx 2D(P_e), \quad \text{or} \quad T' \approx \frac{T'_{\text{single}}}{N'}. \quad (33)$$

It is also noticeable that  $\eta$  is just the average number of *transmitting* sensors  $N'$  divided by the average number of *encountered* sensors  $\lambda E[T] \equiv \lambda'T'$ .

Removing the normalization, we have

$$E[T] \approx \sqrt{\frac{2\mathcal{D}(P_e)}{\eta}} \sqrt{\frac{1}{\lambda\rho}} \quad (34)$$

$$E[N] \approx \sqrt{2\mathcal{D}(P_e)\eta} \sqrt{\frac{\lambda}{\rho}} \quad (35)$$

$$E[T]E[N] \approx \frac{2\mathcal{D}(P_e)}{\rho} \quad \text{or} \quad E[T] \approx \frac{E[T_{\text{single}}]}{E[N]} \quad (36)$$

summarizing the system-level performances. Some comments follow.

- The chosen  $p_e^*$  minimizes the time to decision. At  $p_e = p_e^*$ , with good approximation, it holds that  $E[T] \approx E[T_{\text{loc}}]$ . This latter exposes many features of the system and emphasizes the role of the parameters involved.
- With reference to top panel in Fig. 3, if one chose  $p_e \ll p_e^*$  the system would employ high-performing local sensors but just a few of them would finish the task when polled. Conversely, if one chose  $p_e \gg p_e^*$ , then the sensors would be *fast* but their decisions would be less reliable. There is an *optimal local error-rate* that minimizes the overall decision time, that approximately corresponds to the balanced situation  $E[T] \approx E[T_{\text{loc}}]$ . Once decisions are made locally, the global decision quickly ensues.
- Such time balancing guarantees that, on the average, all the network sensors end their processing tasks about when the final decision is taken: There is little overhang of wasted computation at the sensors after the final decision has already been made. Even though our focus is more on the *transmission* energy burden, it is clear that data *processing* energy is also low in the optimized scheme.<sup>11</sup>
- The average time to take the final decision is proportional to  $1/\sqrt{\lambda\rho}$ : Each doubling of the SNR reduces the decision time of a factor  $\approx 0.7$ . The same holds true for the sensor density:  $\rho$  and  $\lambda$  play a similar role (34). Also, as expected, better-performing systems (i.e., lower  $P_e$ ) require larger processing times:  $E[T]$  varies as  $\sqrt{\mathcal{D}(P_e)}$  (see again (34)).
- The behavior of the number of transmissions  $E[N]$  with respect to  $\rho$  and  $P_e$  is the same as for  $E[T]$ . The density  $\lambda$  has a different impact: The larger is the density, the larger is  $E[N]$  (35).
- Assume that you increase the node density  $\lambda$ . From (34) and (35), we get

$$\frac{E[T]}{E[N]} = \frac{1}{\eta\lambda}$$

revealing that the processing time, normalized with respect to the average number of symbols really collected by the rover, decreases as the inverse of the *talking-node density*  $\eta\lambda$ , as one might expect.

<sup>11</sup>Accordingly one should *turn off* all the sensors in some way, when the rover takes the decision: In this respect, the condition  $E[T] \approx E[T_{\text{loc}}]$  ensures an automatic shutdown. Occasionally SPRTs exhibit very long runs (see, e.g., [14]), implying that a few rare sensors will continue computing their local likelihoods for a long time beyond  $E[T]$ . In fact, if a network node notices that its decision appears to be taking considerably longer than average, it can probably safely terminate its test.

- The average value  $E[T]$  is an important metric of the system; however, useful insight is also gained considering a properly normalized processing time, the ratio  $E[T]/E[T_{\text{single}}]$ . In fact, this ratio varies as  $\sqrt{\rho/\lambda}$ : While the absolute processing time decreases with increasing  $\rho$ , the ratio grows. In this respect, the lower the SNR, the greater advantage accrues to the decentralized structure. Similarly, we find that the more demanding the error probability (i.e., lower  $P_e$ ), the larger the gain of sequential SENMA over a single sensor “system”; this follows from  $E[T]/E[T_{\text{single}}] \propto 1/\sqrt{\mathcal{D}(P_e)}$  (see (35) and (36)).
- Equation (36) tells us that in the optimal system the product between the energy consumption and the (relative) processing time is constant

$$\frac{E[T]}{E[T_{\text{single}}]} E[N] = 1.$$

Regardless of the desired performance  $P_e$ , reducing the energy can only be achieved by increasing the processing time by an equal fraction, and vice versa.

- Another *invariant* of the designed system is  $\eta$ , the ratio between the number of transmissions and the number of *potential* transmissions (i.e., the number of sensors encountered by the rover). In a sense  $1/\eta$  is the energy saving that the designed sequential SENMA may achieve, with respect to a system in which *all* polled sensors send their current decisions. Furthermore, from the linear approximation it is straightforward to get

$$\eta = \frac{E[\Omega]}{E[T]}$$

which provides another interpretation of  $\eta$  as a *warping factor* ruling the time transformation (11). In our setup  $\eta$  is in the order of 0.3, that is  $\approx 30\%$  of the encountered sensors actually communicate their decisions to the rover; or, the transformed time axis is scaled by a factor of  $\approx 0.3$ .

- Due to symmetry (which we have assumed), all the above is true both under  $H_0$  and under  $H_1$ .

## B. Best Sensor-Level $P_e$ for Two-Level Quantization

Let us now consider the two-level quantization strategy. Fig. 4 reproduces Fig. 3 for the two-level quantization scenario under the two hypotheses, and is the starting point of our system optimization.<sup>12</sup> Here we choose to optimize the system under  $H_0$ , so that  $p_e^*$  is selected according to the minima of the left panel. Having fixed the only free parameter  $p_e$  to  $p_e^*$ , (24), (25), and (37)–(40) give the sought system performances.

The general behavior of the performance figures is similar to the three-level case, and many of the previous comments apply to the two-level scheme as well. However, here the statistical hypothesis actually in force makes a difference, due to the lack of symmetry. Further, the system analysis for the two-level quantization yields considerably more sophisticated analytical expressions than that of the three-level case, now involving DeLucia and Poor’s formulas. As a consequence, we limit the analysis

<sup>12</sup>Note that as opposed to the three-level scheme, here the achievable pairs  $(P_e, p_e)$  are *not* constrained to some countable set, so that the curves of Fig. 4 are continuous.



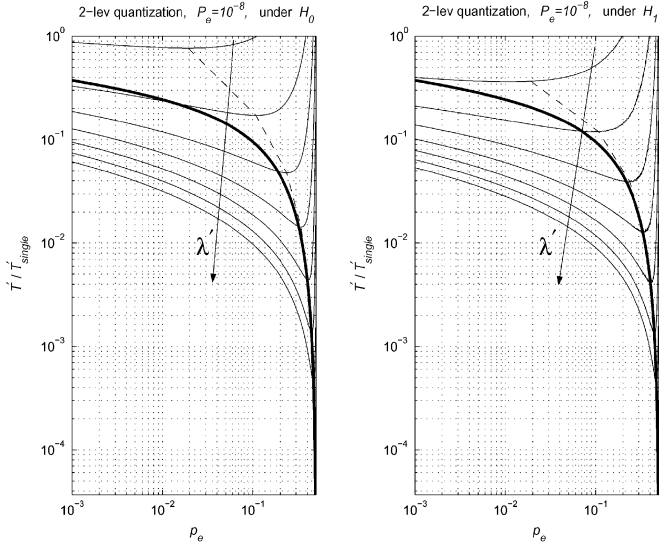


Fig. 4.  $T'/T'_{\text{single}}$  versus  $p_e$ , with  $\lambda'$  as parameter, where  $\lambda' = 10^m$ ,  $m = 0, 1, \dots, 6$ , grows as indicated by the arrow. Left panel refers to  $H_0$  and right panel to  $H_1$ . Here  $P_e = 10^{-8}$ , but similar results are obtained with other  $P_e$ 's. We select  $p_e^*$  as the abscissa of the minimum in the  $H_0$  case, and the dashed curve in the left panel is just the locus of such minima. The dashed curve in  $H_1$  simply represents  $T'/T'_{\text{single}}$  corresponding to the  $p_e^*$  computed under  $H_0$ . The bold curves represent  $T'_{\text{loc}}/T'_{\text{single}}$ : We see that the optimized values are approximately attained where  $T' = T'_{\text{loc}}$ , under both the hypotheses.

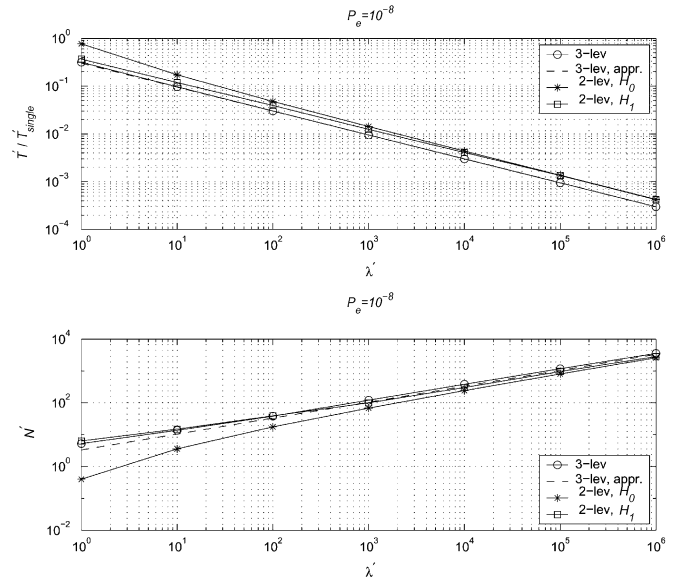


Fig. 6. *Top panel:* The ratio  $T'/T'_{\text{single}}$  versus  $\lambda'$ . *Bottom panel:*  $N'$  versus  $\lambda'$ . The global error probability is  $P_e = 10^{-8}$ . For the three-level quantization strategy, the approximation (appr.) is compared to the numerical evaluation of the exact formulas in Section V-A. In the two-level scenario, the two hypotheses  $H_0$  and  $H_1$  are separately considered.

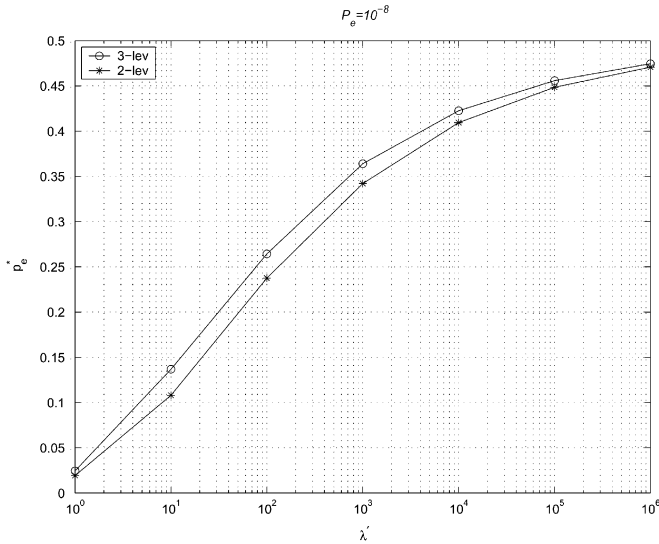


Fig. 5. Values of  $p_e^*$  as resulting from the optimization procedure, as function of the normalized sensor density  $\lambda'$ , in the three-level (3-lev) and two-level (2-lev) quantization schemes. The case study is  $P_e = 10^{-8}$  but, qualitatively, the same behavior is obtained with different total error probabilities.

to the numerical computation of the derived analytical relationships. This is done in the next section.

### C. Optimal Performance Portrait

Let us focus now on the numerical evaluation of the analytical formulas derived so far. For the three-level case, we also compare such formulas with the approximations computed in Section VI-A.

Fig. 5 reports the optimizing  $p_e^*$ . With larger normalized node density  $\lambda' = \lambda/\rho$ , sensors with lower local performance are

appropriate, and the sequential SENMA takes advantage from the larger number of data collected, even if these data are less reliable.

The top panel of Fig. 6 depicts the ratio  $T'/T'_{\text{single}} = E[T]/E[T'_{\text{single}}]$  as function of  $\lambda'$ . This is a measure of the time saving that SENMA provides with respect to a single sensor working at the same error probability. For the three-level scheme the approximated formula  $T'/T'_{\text{single}} = 1/N'$  is compared with the exact relationships derived in Section V-A, and the accuracy is very good. It is worth noting that the three-level scheme always outperforms the two-level one; the difference, however, is moderate. We see also that, for the two-level scenario, the ratio  $T'/T'_{\text{single}}$  under  $H_1$  is always lower than that under  $H_0$ , with the difference becoming negligible at large  $\lambda'/\rho$ . This is natural: The rover in the two-level scheme feeds from local decisions for  $H_1$ , and these are denser under  $H_1$  than under  $H_0$ , with the difference shrinking as local sensor accuracy worsens.

It is particularly of interest to investigate the energy required to the ensemble of sensors for data transmission. This is proportional to the average number of local decisions sent  $N'$ , which is depicted in the bottom panel of Fig. 6. An opposite behavior with respect to  $T'/T'_{\text{single}}$  is observed: Here, the larger is  $\lambda'$ , the larger the energy consumption. The three-level scheme requires more or less the same energy as the two-level quantization working under  $H_1$ . The two-level communication scheme operating under  $H_0$  is, not unexpectedly, less energy consuming than any other scheme or regime; what is perhaps surprising is the minor difference—recall that in the two-level scheme a sensor does not “bother” with a report unless there is a decision for  $H_1$ —and this is presumably reflective of the benefits of the system-level local-decision-quality optimization done.

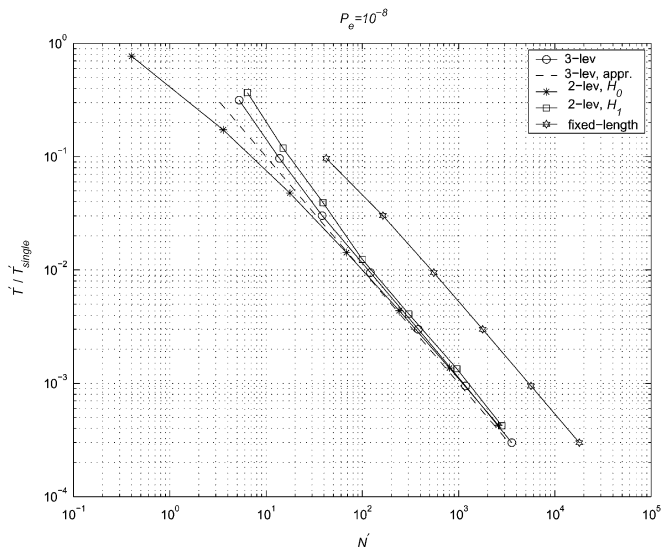


Fig. 7. Ratio  $T'/T'_{\text{single}}$  versus  $N'$ , for  $P_e = 10^{-8}$ ; as before, the label “appr.” refers to the approximation of the three-level quantization scheme. For comparison, the curve of a fixed-length *parallel* scheme is also shown; in this case the abscissa is  $N_{\text{fixed}}$  and the ordinate represents  $T'_{\text{fixed}}/T'_{\text{single}}$  (see also Appendix II).

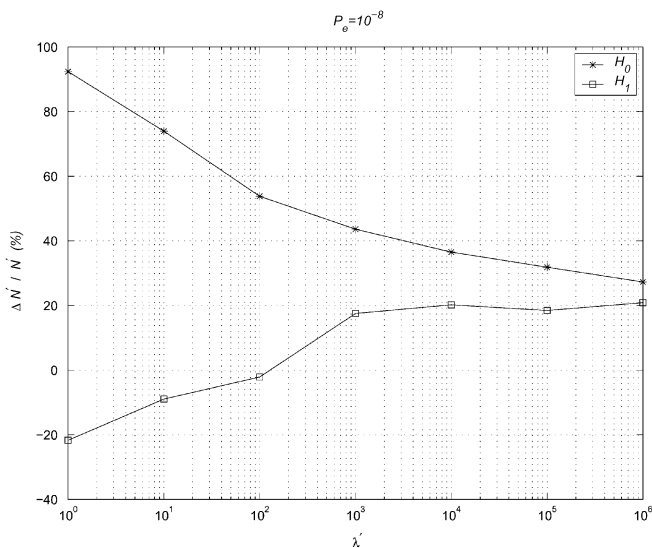


Fig. 8. Energy saving that the two-level quantization scheme may achieve over the three-level scheme, with respect to  $\lambda'$ , for a sample value of  $P_e = 10^{-8}$ . Precisely, the vertical axis reports the relative error (in percentage) between  $N'(3\text{-lev})$  and  $N'(2\text{-lev})$ , that is  $[N'(3\text{-lev}) - N'(2\text{-lev})]/N'(3\text{-lev})$ .

The two plots of Fig. 6 may be combined by resolving for  $\lambda'$ . This results in Fig. 7 that depicts the achievable pairs  $(T'/T'_{\text{single}}, N')$ , evidencing an obvious tradeoff: Time saving may be bought in the coin of energy consumption. In the plane of Fig. 7, the two communication schemes tend to be equivalent in the region of small computation times and large  $N'$ . (The curve labeled with “fixed-length” in Fig. 7 is commented on below.)

For very low  $N'$ , the differences become worth investigating, as is stressed in Fig. 8, in which the relative difference between  $N'(3\text{-lev})$  and  $N'(2\text{-lev})$ , is plotted as function of  $\lambda'$ . The figure reveals that the two-level communication strategy is more energy efficient than the three-level alternative: assuming large  $\lambda'$ ,

under  $H_1$  about 20% savings, while under  $H_0$  the percentage is larger. At the opposite extreme of small  $\lambda'$ , the gains become dramatic under  $H_0$ , but this effect is partially balanced by a *loss* under  $H_1$ . In practical scenarios where the system is expected to operate more often under  $H_0$  than under  $H_1$  (rare presence of targets), such asymmetry might be suitably exploited.

In Fig. 7 we also compare our sequential SENMA approach to a simple *parallel* (and nonsequential, meaning *fixed-length*) decentralized detection approach, in which we have  $N_{\text{fixed}}$  sensors and each observes for a time  $T_{\text{fixed}}$ , makes a decision and immediately transmits this to the fusion center. It must be noted that such a scheme would be rather impractical, since there is a sudden communication event in which all sensors communicate at the same time. Sequential SENMA does not have this disadvantage; but, nonetheless, comparison is worthwhile. The derivation of the performance for the parallel scheme is given in Appendix II, and the challenge here is to relate the quantities to those that are meaningful for a sequential test. At any rate, from Fig. 7, it can be seen that even against the idealized and impractical (in the sense of communication) parallel approach, the sequential SENMA approach performs well. The improvement is reminiscent of the (approximate) factor of three reduction in average time to decision enjoyed by sequential versus fixed testing in centralized problems [14].

## VII. CONCLUSION

We consider a SENMA network architecture for detection purposes. Elaborating on a recent paper [28], it is assumed that the sensors belonging to the network do not communicate with each other, and send to a common fusion center neither data as it arrives nor decisions therefrom. Instead, they mull on the locally available measurements, while waiting to be polled by a mobile agent that is required to fuse the data that are recorded in successive pollings. As opposed to [28], we avoid any collision concern by exploiting a time-continuous polling scheme. This, along with a Poisson random field model for the sensor locations (which are completely unknown to the rover), ensures that at most one sensor is allowed to access the channel at any given time instant.

Each remote sensor employs a Wald sequential test (SPRT) so that, when eventually polled it may have, or not, reached a (local) decision; if it has, then there are two possibilities considered: send that decision to the rover regardless of what it is; or send the decision only if it is for  $H_1$  (the rover must infer the number of “missing”  $H_0$  decisions from that). The rover *also* implements an SPRT using as observables the local decisions (amounting to some counting processes). The idea behind is that of reducing the time needed to end up with a final decision by the rover, with respect to similar, but nonsequential, schemes. A remarkable feature of our sequential SENMA is that all the relevant performance figures can be computed in closed analytical forms. These allow us to set, for example, a simple system optimization in which the free parameter  $p_e$  (i.e., the error rate of the local SPRTs) is chosen to achieve the *minimum average global decision time*.

One possible system design is to use local tests with very small error probabilities (say *intelligent* sensors). This ensures that, if a decision has been reached when polled, the information

is quite valuable for the rover. However, the average local decision time grows with decreasing  $p_e$ , so that a few sensors will respond to the rover polling signal. On the other extreme, we may conceive a network made of *dumb* sensors—that is, very unreliable or with large local error probabilities. In this case, information transferred to the rover is of minor value, but the amount of such information collected by the rover grows considerably. We have found that optimality is at neither extreme: An *optimal dumbness* exists and is at some intermediate value of the local error probability  $p_e$ , for prescribed overall error probability, sensor network and SNR.

The main features of the optimized system are the following. The network exploits the presence of many sensors for reducing the total error probability, not the decision time (with respect to any single sensor of the network). Also, there is little wasted computation, on the average, due to local sensors that continue to gather data after the final decision has already been made.

Finally, in the *three-level* scheme (report any ended test to the rover when it polls), the *effective ratio of reporting sensors* (the proportion of sensors that *do* report to those that are polled)—we have called this  $\eta$ —seems to be close to 30%. Reducing the communication (energy) load is possible with the proposed *two-level* (report only if the local decision is for  $H_1$ ) quantization scheme. Savings in the order of 20% and more are easy to obtain.

Our analysis has relied on a Gaussian shift-in-mean model for detection; generalization to other models is a current research line. For instance, Tartakovsky's formula could be adapted to non-Gaussian scenarios, resorting to a central limit theorem approximation for the pertinent log-likelihood. However, it may be that the simplicity and formal elegance we have obtained here are not preserved. Another line for future work could be the extension to richer hierarchies of sequential testing.

## APPENDIX I

### DELUCIA AND POOR'S FORMULAS

For self-consistency, we summarize some of the results obtained in [5] that are relevant to our problem.

Consider an SPRT built on the log-LR as given in (23). It results that

$$P_f = 1 - \frac{U_{\alpha_0}(J)}{U_{\alpha_0}(J-K)} \quad (37)$$

$$P_d = 1 - \frac{U_{\alpha_1}(J)}{U_{\alpha_1}(J-K)}. \quad (38)$$

In the above

$$\begin{aligned} K &= \frac{\Gamma_0}{\log[(1-p_e)/p_e]} \\ J &= \frac{\Gamma_1}{\log[(1-p_e)/p_e]} \\ \alpha_0 &= \tau^{-1} \log(1+\tau) \\ \alpha_1 &= (1+\tau)\tau^{-1} \log(1+\tau) \end{aligned}$$

and  $\tau = (1-2p_e)/p_e$ . The functions  $U_{\alpha_i}(\zeta)$ ,  $i = 0, 1$ , result from recursive equations. However, in the regime  $\zeta \gg 1$ , simple closed formulas are available

$$\begin{aligned} U_{\alpha_0}(\zeta) &= \frac{1}{1-\alpha_0} + \frac{(1+\tau)^{-\zeta}}{1-\alpha_1} \\ U_{\alpha_1}(\zeta) &= \frac{1}{1-\alpha_1} + \frac{(1+\tau)^\zeta}{1-\alpha_0}. \end{aligned}$$

Notably, the above asymptotic expressions become extremely accurate already for values of the parameter  $\zeta$  in the order of few units, a condition met in all the cases addressed in this paper; hence we use such asymptotics.

As to the related average stopping time  $E[\Omega]$ , in [5], it is shown that

$$\begin{aligned} \lambda p_e E[\Omega|H_0] &= P_f [1 - \alpha_0 Y_{\alpha_0}(J-K)] \\ &\quad + \alpha_0 [Y_{\alpha_0}(J-K) - Y_{\alpha_0}(J)] \quad (39) \end{aligned}$$

$$\begin{aligned} \lambda(1-p_e) E[\Omega|H_1] &= P_d [1 - \alpha_1 Y_{\alpha_1}(J-K)] \\ &\quad + \alpha_1 [Y_{\alpha_1}(J-K) - Y_{\alpha_1}(J)]. \quad (40) \end{aligned}$$

As for before, asymptotic expressions for  $Y_{\alpha_i}(\zeta)$ ,  $i = 0, 1$ , quite accurate even for moderately small  $\zeta$ , can be found

$$\begin{aligned} Y_{\alpha_0}(\zeta) &= -\frac{1}{2} \frac{\alpha_0}{(1-\alpha_0)^2} + \frac{\zeta-1}{1-\alpha_0} \\ &\quad - \frac{(1+\tau)^{1-\zeta}}{\log(1+\tau)(1-\alpha_1)} \\ Y_{\alpha_1}(\zeta) &= -\frac{1}{2} \frac{\alpha_1}{(1-\alpha_1)^2} + \frac{\zeta-1}{1-\alpha_1} \\ &\quad - \frac{(1+\tau)^{\zeta-1}}{\log(1+\tau)(1-\alpha_0)}. \end{aligned}$$

## APPENDIX II

### PERFORMANCE OF A PARALLEL SCHEME WITH FIXED-LENGTH LOCAL TESTS

In Fig. 7 we compare the performance of our sequential SENMA scheme to a simple parallel decentralized detection architecture—recognizing that the latter is not practical from a communication perspective—and here we show how. From Section VI we have that  $E[T_{\text{single}}] = 2D(P_e)/\rho$ , where  $E[T_{\text{single}}]$  is the un-normalized single-sensor time to decision via a sequential detector and  $P_e$  is the global (fused) probability of error. From this we can use the unit-normal exceedance probability  $Q(\cdot)$  to write

$$\begin{aligned} p_{(e\text{-fixed})} &= Q\left(\frac{1}{2}\sqrt{\rho T_{\text{fixed}}}\right) \\ &= Q\left(\sqrt{\frac{\mathcal{D}(P_e)T_{\text{fixed}}}{2E[T_{\text{single}}]}}\right) \\ &= Q\left(\sqrt{\frac{1}{2}\mathcal{D}(P_e)\frac{T'_{\text{fixed}}}{T'_{\text{single}}}}\right) \quad (41) \end{aligned}$$

for the probability of error for each local fixed-length test, where we have defined  $T'_{\text{fixed}} = \rho T_{\text{fixed}}$ . The fused fixed-length test's probability of error is

$$P_e = \sum_{k=\lfloor N_{\text{fixed}}/2 \rfloor + 1}^{N_{\text{fixed}}} \binom{N_{\text{fixed}}}{k} p_{(e-\text{fixed})}^k \cdot (1 - p_{(e-\text{fixed})})^{N_{\text{fixed}} - k} \approx Q \left( \frac{N_{\text{fixed}}(1 - 2p_{(e-\text{fixed})})}{2\sqrt{N_{\text{fixed}}p_{(e-\text{fixed})}(1 - p_{(e-\text{fixed})})}} \right) \quad (42)$$

for large  $N_{\text{fixed}}$ , or

$$N_{\text{fixed}} \approx \frac{4p_{(e-\text{fixed})}(1 - p_{(e-\text{fixed})})(Q^{-1}(P_e))^2}{(1 - 2p_{(e-\text{fixed})})^2} \quad (43)$$

in which  $p_{(e-\text{fixed})}$  is given in (41).

## REFERENCES

- [1] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II—Advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.
- [2] 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.
- [3] C.-Y. Chong and S. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.
- [4] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [5] J. DeLucia and H. V. Poor, "Performance analysis of sequential tests between Poisson processes," *IEEE Trans. Inf. Theory*, vol. 43, no. 1, pp. 221–238, Jan. 1997.
- [6] V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests—Part I: Asymptotic optimality," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2448–2461, Nov. 1999.
- [7] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Sequential decision problems for processes with continuous time parameter. testing hypotheses," *Ann. Math. Stat.*, vol. 24, pp. 254–264, 1953.
- [8] R. Gallager, *Discrete Stochastic Processes*. Norwell, MA: Kluwer, 1996.
- [9] D. Kazakos and P. Papantoni-Kazakos, "Sequential detection between poisson processes," *IEEE Trans. Inf. Theory*, vol. 26, no. 1, pp. 116–120, Jan. 1980.
- [10] S. Marano, V. Matta, P. Willett, and L. Tong, "SPRTs in sensor networks with mobile agents," presented at the IEEE Workshop Signal Processing Advances Wireless Communications (SPAWC), New York, Jun. 5–8, 2005.
- [11] S. Marano, V. Matta, P. Willett, and L. Tong, "Support-based and ML approaches to DOA estimation in a dumb sensor network," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1563–1567, Apr. 2006.
- [12] S. Marano, V. Matta, P. Willett, and L. Tong, "DOA estimation via a network of dumb sensors under the SENMA paradigm," *IEEE Signal Process. Lett.*, vol. 12, no. 10, pp. 709–712, Oct. 2005.
- [13] A. Papoulis, *Probability, Random Variable and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
- [14] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.
- [15] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, pp. 51–60, Mar. 2002.
- [16] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low communication rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.
- [17] C. K. Sestok, M. R. Said, and A. V. Oppenheim, "Randomized data selection in detection with applications to distributed signal processing," *Proc. IEEE*, vol. 91, no. 8, pp. 1184–1198, Aug. 2003.
- [18] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks," in *Proc. 1st IEEE Int. Workshop Sensor Network Protocols Applications*, May 2003, pp. 30–41.
- [19] A. N. Shiryaev, *Statistical Sequential Analysis*. Providence, RI: American Mathematical Soc., 1973.
- [20] S. Tantarana, "Some Recent Results on Sequential Detection," in *Advances in Statistical Signal Processing*. Greenwich, CT: JAI Press, 1993, vol. 2, pp. 265–296.
- [21] S. Tantarana and H. Poor, "Asymptotic efficiencies of truncated sequential tests," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 911–923, Nov. 1982.
- [22] A. G. Tartakovsky, "Sequential composite hypothesis testing with dependent nonstationary observations," *Problems Inf. Transmission*, vol. 17, pp. 18–28, 1981.
- [23] R. R. Tenney and N. R. Sandell Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501–510, 1981.
- [24] L. Tong, Q. Zhao, and S. Adireddy, "Sensor networks with mobile agents," presented at the Military Communications Conf. (MILCOM), Boston, MA, Oct. 2003.
- [25] J. N. Tsitsiklis, "Decentralized detection," *Adv. Statist. Signal Process.*, vol. 2, pp. 297–344, 1993.
- [26] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [27] A. Wald, *Sequential Analysis*. New York: Dover, 1947.
- [28] P. Willett and L. Tong, "One aspect to cross-layer design in sensor networks," presented at the Military Communications Conf. (MILCOM), Monterey, CA, Oct. 2004.
- [29] F. Zhao, J. Liu, J. Liu, L. Guibas, and J. Reich, "Collaborative signal and information processing: An information-directed approach," *Proc. IEEE*, vol. 91, no. 8, pp. 1199–1209, Aug. 2003.



**Stefano Marano** received the Laurea degree in electronic engineering (*cum laude*) and the Ph.D. degree in electronic engineering and computer science both from the University of Naples, Italy, in 1993 and 1997, respectively.

Currently, he is a Professor at the University of Salerno, Italy, where he was formerly Assistant Professor. His areas of interest include statistical signal processing with emphasis on inference, sensor networks, and information theory. He has coauthored approximately 70 papers about these and related

topics, including some invited, mainly on international journals/transactions and proceedings of international conferences.

Prof. Marano was corecipient of the S. A. Schelkunoff Transactions Prize Paper Award of the IEEE Antennas and Propagation Society for the best paper published in the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION in 1999. Recently, he is/was in the Scientific Committee of the Remote Sensing Laboratory for Environmental Hazard Monitoring (ReSLEHM), University of Salerno, and in the Organizing Committee of the Ninth International Conference on Information Fusion (FUSION 2006).



**Vincenzo Matta** received the Laurea degree in electronic engineering and the Ph.D. degree in information engineering from University of Salerno, Fisciano, Italy, in 2001 and 2005, respectively.

He is currently an Assistant Professor with the University of Salerno. His main research interests include detection and estimation theory, signal processing, wireless communications, multiterminal inference and sensor networks.



**Peter Willett** (S'83–M'86–SM'97–F'03) received the B.A.Sc. degree from the University of Toronto, Toronto, ON, Canada, in 1982 and the Ph.D. degree from Princeton University, Princeton, NJ, in 1986.

Previously, he was with the University of Toronto and Princeton University. Currently, he is a Professor of Electrical and Computer Engineering at the University of Connecticut, Storrs. He has written, to date, 86 journal papers, 202 conference papers, and seven book chapters about signal processing, detection, information theory, CDMA, learning from

data, and target tracking, among other topics.

He is Editor-in-Chief for the *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS* and Associate Editor for the *International Society for Information Fusion Journal of Advances in Information Fusion (ISIF-JAIF)* and was Associate Editor for the *IEEE Aerospace and Electronics Systems Magazine* and for *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS (both parts A and B)*. He is a member of the editorial board for *IEEE Signal Processing Magazine*. He was a track organizer for Remote Sensing at the IEEE Aerospace Conference from 2001 to 2003 and was Co-Chair of the Diagnostics, Prognosis, and System Health Management SPIE Conference, Orlando, FL. He also served as Program Co-Chair for the 2003 IEEE Systems, Man and Cybernetics Conference, Washington, DC, and is General Co-Chair of ISIF's 2006 Fusion Conference, Florence, Italy.



**Lang Tong** (S'87–M'91–SM'01–F'05) received the B.E. degree from Tsinghua University, Beijing, China, in 1985 and the M.S. and Ph.D. degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1987 and 1991, respectively.

In 1991, he was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, Stanford, CA. Currently, he is a Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY. He was also the 2001

Cor Wit Visiting Professor at the Delft University of Technology, Delft, The Netherlands. His areas of interest include statistical signal processing, wireless communications, communication networks and sensor networks, and information theory.

Dr. Tong received Young Investigator Award from the Office of Naval Research in 1996, the Outstanding Young Author Award from the IEEE Circuits and Systems Society in 1991, the 2004 IEEE Signal Processing Society Best Paper Award (with M. Dong), and the 2004 Leonard G. Abraham Prize Paper Award (with P. Venkatasubramanian and S. Adireddy) from the IEEE Communications Society. He serves as an Associate Editor for the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* and *IEEE SIGNAL PROCESSING LETTERS*.