# EXPONENTIAL ERROR BOUNDS FOR BINARY DETECTION USING ARBITRARY BINARY SENSORS AND AN ALL-PURPOSE FUSION RULE IN WIRELESS SENSOR NETWORKS

*John A. Gubner*

University of Wisconsin–Madison
Dept. of Electrical and Computer Engineering
Madison, WI 53706-1691 USA

*Louis L. Scharf and Edwin K. P. Chong*[*]

Colorado State University
Dept. of Electrical and Computer Engineering
Fort Collins, CO 80523-1373 USA

## ABSTRACT

Wireless sensor networks are considered in which sensors convey binary decisions over fading channels to a common fusion center. The fusion center first takes each received signal and makes an estimate of the transmitted bit. The average of the estimated bits is compared to a threshold to make a global decision. Exponential error bounds are derived that allow one to trade off signal-to-noise ratio versus the number of sensors to achieve desired average error levels. An attractive feature of the bounds is that they do not require exact knowledge of the wireless channel statistics; approximations are sufficient.

***Index Terms*—**Channel state information, distributed detection, fusion center, ultra-wideband, wireless sensor network.

## 1. INTRODUCTION

The parallel-architecture wireless sensor network shown in Fig. 1 is to be used to make a decision about a binary hypothesis. In this system, each sensor uses its measurement to generate a binary message to be sent over its own channel to the fusion center. We consider the case in which the fusion rule is required not to depend on exact knowledge of the performance of the individual sensors or on exact knowledge of channel-state information (CSI). Our goal is to provide exponentially decaying bounds on the average global false-alarm probability and the average global miss probability, where the average is over the CSI available to the decoders. Such averaging appears, for example, in [4, eq. (4)]. There, such averaging is regarded as infeasible for computation of the average global error probability itself. Here, since we focus on bounds rather than the error probability itself, we do not encounter such difficulties.

[*]The work of E. K. P. Chong was supported in part by NSF Grant ECCS–0700559.

## 2. SYSTEM MODEL

Writing $V_k := (U_k, Y_k, C_k, L_k)$ for $k = 1, \dots, n$, we assume that the $V_k$ are conditionally independent and identically distributed (conditionally i.i.d.) given either hypothesis. The random variables $U_k$ and $L_k$ are binary-valued, while $Y_k$ and $C_k$ may be discrete or continuous random variables (or even random processes), depending on the situation at hand. In particular, $C_k$ denotes the CSI of the $k$th channel, and is available to the $k$th decoder, but not to the fusion rule itself. We further assume that for each $k$, $C_k$ is independent of $(U_k, Y_k, L_k)$ and that the distribution of $C_k$ does not depend on the hypothesis or on $k$.

Although each $U_k$ is binary-valued, we do not require or assume that it is the result of a local likelihood-ratio test (LRT). Even so, we call $U_k$ a **local decision**, and define the local false-alarm and detection probabilities,

$$p_\text{F} := \mathsf{P}_0(U_k = 1) \quad \text{and} \quad p_\text{D} := \mathsf{P}_1(U_k = 1),$$

where $\mathsf{P}_0$ and $\mathsf{P}_1$ are probabilities computed under hypothesis $H_0$ and $H_1$, respectively.

Conditioned on the CSI $C_k$ of the $k$th channel, we put

$$\alpha_k := \mathsf{P}_0(L_k = 1|C_k) \quad \text{and} \quad \beta_k := \mathsf{P}_1(L_k = 1|C_k).$$

If we group the sensor, channel, and decoder into a **virtual sensor** and call $L_k$ the $k$th ***virtual*** **local decision**, then $\alpha_k$ can be viewed as the virtual local false-alarm probability of the $k$th virtual sensor. Similarly, $\beta_k$ can be viewed as the virtual local detection probability of the $k$th virtual sensor.

## 3. THE FUSION RULE

Let $\mathbf{C} := [C_1, \dots, C_n]'$. Conditioned on knowing $\mathbf{C}$, and assuming the fusion-rule designer knows the distribution of $V_k$ under each hypothesis, the optimum fusion rule under the Neyman–Pearson or Bayesian setups is given in terms of the likelihood ratio of $(L_1, \dots, L_n)$. Since the $L_k$ are conditionally independent under each hypothesis, it is a straightforward calculation, e.g., [1], to show that the required log likelihood

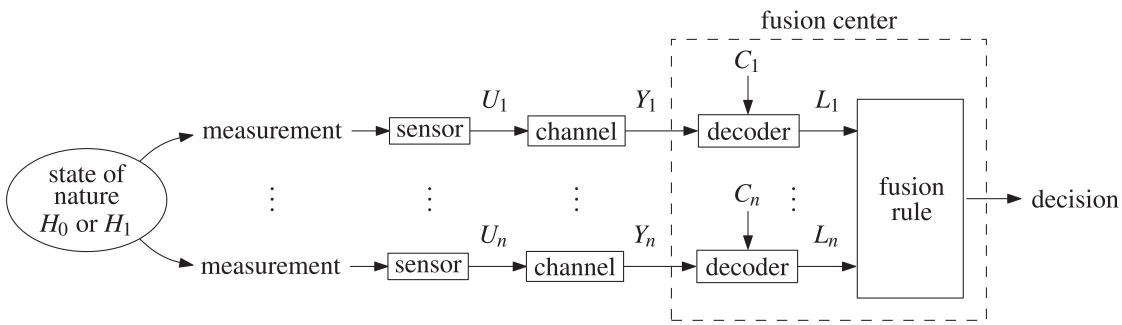


**Fig. 1**. A parallel-architecture wireless sensor network for a binary detection problem.

ratio can be reduced to

$$\sum_{k=1}^{n} L_k \ln \frac{\beta_k(1-\alpha_k)}{\alpha_k(1-\beta_k)},$$

which is a linear combination of the $L_k$. As this requires more information than is likely to be available [3], [11], [12], we propose instead the **average-of-virtual-local-decisions** statistic,

$$X_n := \frac{1}{n}\sum_{k=1}^{n} L_k,$$

to be compared with a threshold $\eta$. Specifically, the fusion rule declares $H_1$ to be true of $X_n > \eta$, and $H_0$ otherwise. Notice that the only way in which this fusion rule can depend on sensor or channel distributions is in the choice of the threshold. As we will see, the choice of suitable thresholds does not require exact knowledge or calculation. In this sense, our fusion rule is "all purpose."

## 4. PERFORMANCE ANALYSIS

We now bound the global probabilities of false alarm and miss conditioned on the CSI $\mathbf{C}$. A simple Chernoff-bound argument shows that for $s > 0$,

$$\mathsf{P}_0(X_n > \eta|\mathbf{C}) \le e^{-s\eta}\prod_{k=1}^{n}(1-\alpha_k+\alpha_k e^{s/n})$$

and

$$\mathsf{P}_1(X_n \le \eta|\mathbf{C}) \le e^{s\eta}\prod_{k=1}^{n}(1-\beta_k+\beta_k e^{-s/n}).$$

Defining

$$\alpha := \mathsf{E}[\alpha_k] \quad \text{and} \quad \beta := \mathsf{E}[\beta_k],$$

and taking expectations of the two preceding inequalities, we find that

$$\mathsf{E}[\mathsf{P}_0(X_n > \eta|\mathbf{C})] \le \exp[-s\eta + n\ln(1-\alpha+\alpha e^{s/n}]$$

and

$$\mathsf{E}[\mathsf{P}_1(X_n \le \eta|\mathbf{C})] \le \exp[s\eta + n\ln(1-\beta+\beta e^{-s/n})].$$

Minimizing the right-hand sides over $s > 0$ yields, assuming $\alpha < \eta < \beta$,

$$\mathsf{E}[\mathsf{P}_0(X_n > \eta|\mathbf{C})] \le e^{-nD(\eta\|\alpha)} \tag{1}$$

and

$$\mathsf{E}[\mathsf{P}_1(X_n \le \eta|\mathbf{C})] \le e^{-nD(\eta\|\beta)}, \tag{2}$$

where, for $0 < p < 1$, the error exponent $D(\eta\|p)$ is the Kullback–Leibler informational divergence,

$$D(\eta\|p) := \eta\ln\frac{\eta}{p} + (1-\eta)\ln\frac{1-\eta}{1-p}.$$

Thus, (1) and (2) give exponentially decaying upper bounds on the average global probabilities of false alarm and miss at the fusion center output.

### 4.1. Discussion

The bounds in (1) and (2) depend on $\alpha$ and $\beta$, which we may not know. However suppose we have an upper bound on $\alpha$, say $\alpha_{\max}$, and a lower bound on $\beta$, say $\beta_{\min}$, such that

$$\alpha \le \alpha_{\max} < \beta_{\min} \le \beta.$$

If we then choose $\alpha_{\max} < \eta < \beta_{\min}$, we can loosen the bounds in (1) and (2) to

$$\mathsf{E}[\mathsf{P}_0(X_n > \eta|\mathbf{C})] \le e^{-nD(\eta\|\alpha_{\max})}$$

and

$$\mathsf{E}[\mathsf{P}_1(X_n \le \eta|\mathbf{C})] \le e^{-nD(\eta\|\beta_{\min})},$$

where we have used the easily-verified fact that as a function of $p$, $D(\eta\|p)$ is increasing for $0 < p < \eta$ and is decreasing for $\eta < p < 1$.

In a Bayesian context with prior probabilities $\pi_0$ and $\pi_1$ of $H_0$ and $H_1$, respectively, we have the expected global probability of error,

$$\begin{aligned}\overline{P}_e &:= \mathsf{E}\big[\pi_0\mathsf{P}_0(X_n > \eta|\mathbf{C}) + \pi_1\mathsf{P}_1(X_n \le \eta|\mathbf{C})\big] \\ &\le \pi_0 e^{-nD(\eta\|\alpha)} + \pi_1 e^{-nD(\eta\|\beta)}.\end{aligned}$$

To make this go to zero as fast as possible as $n \to \infty$, we choose $\eta$ to equalize the divergences. The required value of $\eta$ is easily found to be

$$\eta = \frac{\ln \dfrac{1-\alpha}{1-\beta}}{\ln \dfrac{\beta}{\alpha} + \ln \dfrac{1-\alpha}{1-\beta}}.$$

If we only have bounds $\alpha_{\max}$ and $\beta_{\min}$, we can similarly equalize the divergences in

$$\overline{P}_e \le \pi_0 e^{-nD(\eta\|\alpha_{\max})} + \pi_1 e^{-nD(\eta\|\beta_{\min})}.$$

### 4.2. Analysis of $\alpha$ and $\beta$

Since $\alpha := \mathsf{E}[\alpha_k]$, we begin by evaluating $\alpha_k$. It is helpful to use the notation

$$\widehat{\alpha}(c,u) := \mathsf{P}_0(L_k = 1 | C_k = c, U_k = u),$$

where, for convenience, we take the binary values of $U_k = u$ to be $u = \pm 1$. Using properties of conditional expectation and conditional probability, we have

$$\begin{aligned} \alpha_k &:= \mathsf{P}_0(L_k = 1|C_k) \\ &= \mathsf{E}_0[\mathsf{P}_0(L_k = 1|C_k, U_k)|C_k] \\ &= \mathsf{E}_0[\widehat{\alpha}(C_k, U_k)|C_k] \\ &= \widehat{\alpha}(C_k, 1)p_{\mathrm{F}} + \widehat{\alpha}(C_k, -1)(1-p_{\mathrm{F}}), \end{aligned} \tag{3}$$

where the last step uses the independence of $C_k$ and $U_k$.

Similarly, since $\beta := \mathsf{E}[\beta_k]$, with

$$\widehat{\beta}(c,u) := \mathsf{P}_1(L_k = 1 | C_k = c, U_k = u),$$

we have

$$\begin{aligned} \beta_k &:= \mathsf{P}_1(L_k = 1|C_k) \\ &= \widehat{\beta}(C_k, 1)p_{\mathrm{D}} + \widehat{\beta}(C_k, -1)(1-p_{\mathrm{D}}). \end{aligned} \tag{4}$$

The formulas for $\widehat{\alpha}(c,u)$ and $\widehat{\beta}(c,u)$ depend only on the channel and the decoder, while $p_{\mathrm{F}}$ and $p_{\mathrm{D}}$ depend only on the measurement statistics and on how the sensor converts the measurement into $U_k$. The formulas (3) and (4) show how the channel and decoder are coupled with the sensor to compose $\alpha_k$ and $\beta_k$.

## 5. EXAMPLE

Similar to [3], [11], we assume independent, flat-fading channels of the form

$$Y_k = C_k U_k + Z_k,$$

where the $Z_k$ are i.i.d. Gaussian with zero mean and variance $\sigma^2$ under each hypothesis, and $U_k = \pm 1$. The $k$th decoder outputs $L_k = 1$ if $Y_k > 0$ and $L_k = 0$ otherwise. To carry out the program of the previous section, we compute $\widehat{\alpha}(c,1) = 1 - Q(c/\sigma)$ and $\widehat{\alpha}(c,-1) = Q(c/\sigma)$, where $Q$ denotes the standard normal complementary cumulative distribution function, $Q(z) := \int_z^\infty e^{-t^2/2}/\sqrt{2\pi}\,dt$. It follows from (3) and (4) that

$$\alpha_k = [1 - Q(C_k/\sigma)]p_{\mathrm{F}} + Q(C_k/\sigma)(1-p_{\mathrm{F}})$$

and

$$\beta_k = [1 - Q(C_k/\sigma)]p_{\mathrm{D}} + Q(C_k/\sigma)(1-p_{\mathrm{D}}).$$

To compute $\alpha := \mathsf{E}[\alpha_k]$ and $\beta := \mathsf{E}[\beta_k]$, it suffices to compute $\mathsf{E}[Q(C_k/\sigma)]$. If $C_k$ is a Rayleigh random variable with second moment $\mathsf{E}[C_k^2] = \rho^2$, then we have from [13, eq. (3.61)] or [7, p. 226] that

$$S := \mathsf{E}[Q(C_k/\sigma)] = \frac{1}{2}\left[1 - \left(1 + \frac{2}{\rho^2/\sigma^2}\right)^{-1/2}\right].$$

It follows that

$$\alpha = (1-S)p_{\mathrm{F}} + S(1-p_{\mathrm{F}})$$

and

$$\beta = (1-S)p_{\mathrm{D}} + S(1-p_{\mathrm{D}}).$$

Note that

$$\begin{aligned} \beta - \alpha &= (p_{\mathrm{D}} - p_{\mathrm{F}})(1-2S) \\ &= (p_{\mathrm{D}} - p_{\mathrm{F}})\left(1 + \frac{2}{\rho^2/\sigma^2}\right)^{-1/2}, \end{aligned} \tag{5}$$

which is positive whenever $p_{\mathrm{D}} > p_{\mathrm{F}}$ as would typically be the case. Furthermore, $\beta - \alpha$ is the length of the interval in which the admissible threshold $\eta$ must lie as noted above (1); we can now see how this length depends on the signal-to-noise ratio (SNR) $\rho^2/\sigma^2$. At infinite SNR, $\beta - \alpha = p_{\mathrm{D}} - p_{\mathrm{F}}$; as the SNR decreases to zero, $\beta - \alpha$ shrinks to zero according to the factor multiplying $p_{\mathrm{D}} - p_{\mathrm{F}}$ in (5).

### 5.1. Extensions

Conceptually it is straightforward to carry out the foregoing calculations for a multipath channel. In this case, $\widehat{\alpha}(c,\pm 1)$ and $\widehat{\beta}(c,\pm 1)$ will involve quantities roughly of the form $Q\big(\sqrt{\|c\|^2/\sigma^2}\big)$, where $c$ is now a vector. The challenge then is to compute $\mathsf{E}\big[Q\big(\sqrt{\|c\|^2/\sigma^2}\big)\big]$. Even when the length of the vector $c$ is random, as in some ultra-wideband channel models [6], [10], such expectations can sometimes be computed numerically, e.g., [8].

## 6. CONCLUSION

Our results allow one to trade off SNR versus number of sensors to achieve a given average global false-alarm probability, miss probability, and/or error probability. To achieve a longer network lifetime by lowering the SNR but still maintain desired probabilities, we can use our exponential bounds to determine the number of sensors required.

Our results make no assumptions about the design of the sensors. For example, design of optimal sensors that take into account the channel has been considered in [2], [5], [9]. Notice also that our fusion rule depends on the sensors and the channel only through bounds on the threshold $\eta$. Design of fusion rules that take optimal account of the channel has been considered in [3], [11].

## 7. REFERENCES

[1] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, Jan. 1986.

[2] J.-F. Chamberland and V. V. Veeravalli, "The impact of fading on decentralized detection in power constrained wireless sensor networks," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'04)*, May 17-21, 2004, vol. 3, pp. 837–840.

[3] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[4] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 16–26, July 2006.

[5] B. Chen and P. K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 693–699, Feb. 2005.

[6] J. Foerster, Ed., "Channel modeling sub-committee report final," IEEE, Document IEEE P802.15-02/490r1-SG3a, 2003.

[7] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge, UK: Cambridge Univ. Press, 2006.

[8] J. A. Gubner and K. Hao, "A computable formula for the average bit error probability as a function of window size for the IEEE 802.15.3a UWB channel model," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 4, pp. 1762–1768, Apr. 2006.

[9] B. Liu and B. Chen, "Decentralized detection in wireless sensor networks with channel fading statistics," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'06)*, May 14-19, 2006, Toulouse, France, vol. 4, pp. 857–860.

[10] A. F. Molisch *et al.*, "IEEE 802.15.4a Channel Model – Final Report," Tech. Rep., Document IEEE 802.15.04-0662-02-004a, 2005.

[11] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.

[12] R. Niu, P. K. Varshney, M. Moore, and D. Klamer, "Decision fusion in a wireless sensor network with a large number of sensors," in *7th Int. Conf. Inform. Fusion*, 2004.

[13] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge Univ. Press, 1998.