Distributed Detection in Large-Scale Sensor Networks with Correlated Sensor Observations

Wenjun Li and Huaiyu Dai

Department of Electrical and Computer Engineering,

North Carolina State University, Box 7911, Raleigh, NC 27695

wli5, huaiyu_dai@ncsu.edu*

Abstract

Detection of a deterministic signal in correlated Gaussian noise in large-scale sensor networks is considered. In contrast to existing works which assume that each sensor uses a dedicated channel for transmission, we explore the possibility of employing a shared multiple access channel (MAC) for detection, which significantly reduces the bandwidth requirement, or the detection delay. We assume that sensors transmit local decisions based on their respective local observations and a mapping rule through a MAC subject to an average power constraint. We consider two scenarios depending on whether the sensors have intelligence or not, and propose one mapping rule for each scenario. We show that while detection over a parallel access channel (PAC) always results in a loss in error exponent, the asymptotic performance of optimal centralized detection can be achieved with the first MAC mapping rule for intelligent sensors. For dumb sensors, the second MAC mapping rule has advantage in terms of energy efficiency compared with detection over PAC in certain applications. Numerical examples are given to illustrate how the results can be used to design optimal sensor spacing.

1 Introduction

Distributed detection of certain events or targets in the environment is an important application of sensor networks [1, 2]. Distributed sensors take observations and communicate local decisions with the fusion center, which makes a final decision based on the received information. The presence of noise in the communication channels as well as possible processing at sensors (local decision, quantization, etc.) result in a loss of detection performance as compared to centralized detection.

For a densely-deployed sensor field, we are usually interested in the trend of the detection performance as the number of sensors goes to infinity, which is provided by the measure of error exponent [1–4]. In [1], the performance of distributed detection strategies is assessed in terms of the error exponent for two types of channels: a parallel access channels (PAC) consisting of a bank of dedicated AWGN channels and an AWGN multiple access channel (MAC). As many other works [2], an important assumption made in [1] is that the sensor observations are independent and identically distributed. However, as sensors are packed more closely to each other, it is reasonable to expect that

*This research was supported in part by the National Science Foundation under Grant CCF-0515164.

their observations become more correlated [3]. Analysis of distributed detection with correlated observations is typically complicated even for a pair of Gaussian observations [5], yet in the asymptotic region, large deviation theory [6] provides an efficient way for computing the error exponent. Several recent works have employed the large deviation theory to study the problem of optimal sensor placement [3], [4]. The detection of a constant signal in correlated Gaussian noise and the detection of a Gaussian-Markov process in independent Gaussian noise are studied in [3], where it is assumed that each sensor transmits its observation directly through a dedicated AWGN channel. The latter problem is also studied in [4], based on the assumption that the local observations are perfectly available at the fusion center.

The major problem with using a PAC in large-scale sensor networks is that dedicated channels for each sensor results in either a large bandwidth requirement for simultaneous transmission or a large detection delay. In this paper, we explore the possibility of employing a shared multiple access channel for detection of a deterministic signal in correlated Gaussian noise. The deterministic signal is not limited to the constant signal as in [3], but is defined to admit a wide range of known signals. We assume that sensors transmit local decisions based on their respective local observations and a mapping rule through a MAC subject to an average power constraint. We consider two scenarios depending on whether the sensors are intelligent (having knowledge of source statistics) or not, and propose one mapping rule for each scenario. We show that the asymptotic performance of optimal centralized detection can be achieved with the first mapping rule for intelligent sensors. On the other hand, detection over PAC always results in a loss in asymptotic error exponent. For dumb sensors, the second MAC mapping rule has advantage in terms of energy efficiency compared with detection over PAC in certain applications. Numerical examples are given to illustrate the performance of various detection strategies and how the results can be used to design the optimal sensor spacing.

2 System Model and Preliminaries

2.1 System Model

We consider the binary hypothesis testing problem where the kth sensor's observation is given by

$$H_1 : x_k = m_k + v_k, \quad k = 1, 2, \cdots, n, \tag{1}$$

$$H_0 : x_k = v_k, \quad k = 1, 2, \cdots, n, \tag{2}$$

where $\{m_k\}_{k=1}^n$ is a real deterministic signal whose autocorrelation function is given by

$$r(k) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} m_{j+k} m_j = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\omega} dG(\omega), \tag{3}$$

where $G(\omega)$ is the spectral distribution, and if $G(\cdot)$ is absolutely continuous, its derivative $G'(\omega)$ is the spectral density of $\{s_k\}_{k=1}^n$ [7]. $\{v_k\}_{k=1}^n$ is a stationary Gaussian stochastic process with zero mean and covariance matrix $\mathbf{\Sigma}$. We assume that the local observations are first mapped through a function $U(\cdot) : y_k = U(x_k)$, subject to an average power constraint $\frac{1}{n}\sum_{k=1}^n E\{|y_k|^2\} \le P$, and may be transmitted over two types of channels:

1. A parallel access channel (PAC) consisting of n dedicated AWGN channels,

$$r_k = y_k + z_k, \quad k = 1, 2, \cdots, n, \tag{4}$$

where z_k's are i.i.d. zero mean Gaussian variables with unit variance.

2. A perfectly synchronized AWGN multiple access channel (MAC) given by

$$r = \sum_{k=1}^{n} y_k + z, \tag{5}$$

where z is zero mean Gaussian with unit variance.

The signal attenuation on the channel is assumed to be absorbed in $U(\cdot)$. In the following, we use the vector notation $\mathbf{m} = (m_1, \cdots, m_n)^T$, and similarly denote $\mathbf{v}$, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{r}$.

2.2 Preliminaries

The analysis in this paper requires two sets of mathematical tools, one is associated with the asymptotic properties of Toeplitz matrices, and the other is an important result in large deviation theory which characterizes the asymptotic behavior of non-i.i.d. sequences. For completeness, we briefly recall the necessary definitions and theorems.

Absolutely Summable Toeplitz Matrix: Let $\{\mathbf{\Sigma}^{(n)}\}$ be a sequence of $n \times n$ Toeplitz matrices with entries $t_k \in \mathbb{R}$ on the kth diagonal and dimension $n \to \infty$. If $\sum_{k=-\infty}^{\infty} |t_k| < \infty$, $\{\mathbf{\Sigma}^{(n)}\}$ has spectral density given by the Fourier Series of t_k: $S(\omega) = \sum_{k=-\infty}^{\infty} t_k e^{ik\omega}$ [8].

Toeplitz distribution theorem [8] and its extension [9, 10]: Let $\{m_k\}_{k=1}^n$ be a deterministic signal with spectral distribution $G(\omega)$. For an absolutely summable Toeplitz matrix $\mathbf{\Sigma}^{(n)}$ with spectral density $S(\omega)$, let $\{\lambda_k^{(n)}\}_{k=1}^n$ be the eigenvalues of $\mathbf{\Sigma}^{(n)}$ contained on the interval $[\delta_1, \delta_2]$, and $\{\boldsymbol{\phi}_k^{(n)}\}_{k=1}^n$ be the normalized eigenvectors of $\mathbf{\Sigma}^{(n)}$, then for any continuous function $h(\cdot)$ defined on $[\delta_1, \delta_2]$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} h(\lambda_k^{(n)}) = \frac{1}{2\pi} \int_0^{2\pi} h(S(\omega)) d\omega, \tag{6}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} h(\lambda_k^{(n)}) (\mathbf{m}^T \boldsymbol{\phi}_k)^2 = \frac{1}{2\pi} \int_0^{2\pi} h(S(\omega)) \, dG(\omega). \tag{7}$$

Gärtner-Ellis Theorem [6]: Let $Z_n \in \mathbb{R}$ be a sequence of random variables drawn according to the probability law μ_n, and define

$$\Lambda^{(n)}(\theta) = \log E[e^{\theta Z_n}]. \tag{8}$$

Assumptions: (1) For each $\theta \in \mathbb{R}$, the logarithmic moment generating function, defined as the limit $\Lambda(\theta) \doteq \lim_{n\to\infty} \frac{1}{n} \Lambda^{(n)}(n\theta)$ exists as an extended real number. (2) The interior of $D_\Lambda \doteq \{\theta \in \mathbb{R} : \Lambda(\theta) < \infty\}$, denoted by D_Λ^o, contains the origin. (3) $\Lambda(\cdot)$ is differentiable throughout D_Λ^o, and $\Lambda(\cdot)$ is steep, i.e., $\lim_{n\to\infty} \Lambda'(\theta_n) = \infty$ whenever $\{\theta_n\}$ is a sequence in D_Λ^o converging to a boundary point of D_Λ^o.
Under the above assumptions, the large deviation principle (LDP) satisfied by the sequence of $\{\mu_n\}$ can be characterized by the *Fenchel-Legendre* transform of $\Lambda(\theta)$:

$$\Lambda^*(x) = \sup_{\theta \in \mathbb{R}} \{\theta x - \Lambda(\theta)\}, \tag{9}$$

that is, for any closed set $F \subset \mathbb{R}$, $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \le -\inf_{x\in F} \Lambda^*(x)$, and for any open set $G \subset \mathbb{R}$, $\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \ge -\inf_{x\in G} \Lambda^*(x)$.

3 Distributed Detection

3.1 Optimal Centralized Detection

Optimal centralized detection, where the sensor observation vector $\mathbf{x}$ is perfectly available to the fusion center, serves as a performance baseline for distributed detection strategies. The optimal centralized detection is a threshold test on the normalized log-likelihood ratio [11]: Choose H_1 if

$$\frac{1}{n}\log\frac{\Pr(\mathbf{x}|H_1)}{\Pr(\mathbf{x}|H_0)} = \frac{1}{n}(\mathbf{m}^T\mathbf{\Sigma}^{-1}\mathbf{x} - \frac{1}{2}\mathbf{m}^T\mathbf{\Sigma}^{-1}\mathbf{m}) > \tau, \tag{10}$$

and choose H_0 otherwise. For the Bayesian problem with priori probabilities $P(H_0) = \pi_0$ and $P(H_1) = \pi_1$, the threshold $\tau = \frac{1}{n}\log\frac{\pi_0}{\pi_1} \to 0$ as $n \to \infty$. For the Neyman-Pearson problem, the threshold τ is chosen to minimize the type II error probability subject to a constraint on the type I error probability. We can write the test alternatively as

$$T_n = \frac{1}{n}\mathbf{m}^T\mathbf{\Sigma}^{-1}\mathbf{x} \gtrless T. \tag{11}$$

Let $\mathbf{\Sigma} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^T$, where $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$ is a diagonal matrix containing the eigenvalues of $\mathbf{\Sigma}$, and $\mathbf{\Phi}$ is a unitary matrix with eigenvectors of $\mathbf{\Sigma}$ as column vectors. Let $\mathbf{p} = \mathbf{\Phi}^T\mathbf{m}$ and $\mathbf{w} = \mathbf{\Phi}^T\mathbf{x}$. Then it is easily shown that under H_0, $\{w_k\}_{k=1}^n$ are i.i.d. $\mathcal{N}(0, \lambda_k)$, and under H_1, $\{w_k\}_{k=1}^n$ are i.i.d. $\mathcal{N}(p_k, \lambda_k)$. Thus we have

$$T_n = \frac{1}{n}\sum_{k=1}^{n}\frac{p_k w_k}{\lambda_k}. \tag{12}$$

The logarithmic moment generating functions under both hypotheses are given by

$$\Lambda_0^{(n)}(n\theta) = \log E_0\{e^{\theta\sum_{k=1}^n \frac{p_k w_k}{\lambda_k}}\} = \sum_{k=1}^{n}\frac{\theta^2 p_k^2}{2\lambda_k} = \frac{\theta^2}{2}\sum_{k=1}^{n}\frac{(\mathbf{m}^T\boldsymbol{\phi}_k)^2}{\lambda_k}, \tag{13}$$

$$\Lambda_1^{(n)}(n\theta) = \log E_1\{e^{\theta\sum_{k=1}^n \frac{p_k w_k}{\lambda_k}}\} = \sum_{k=1}^{n}\frac{(\theta^2+2\theta)p_k^2}{2\lambda_k} = \frac{\theta^2+2\theta}{2}\sum_{k=1}^{n}\frac{(\mathbf{m}^T\boldsymbol{\phi}_k)^2}{\lambda_k}. \tag{14}$$

Using extended Toeplitz distribution thoerem, we have

$$\Lambda_0(\theta) = \lim_{n\to\infty}\frac{1}{n}\Lambda_0^{(n)}(n\theta) = \frac{\theta^2}{4\pi}\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}, \tag{15}$$

$$\Lambda_1(\theta) = \lim_{n\to\infty}\frac{1}{n}\Lambda_1^{(n)}(n\theta) = \frac{\theta^2+2\theta}{4\pi}\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}, \tag{16}$$

where $S(\cdot)$ is the spectral density of $\mathbf{\Sigma}$ as defined in Section 2.2. It can be checked that the assumptions for Gärtner-Ellis theorem hold. Therefore T_n satisfies the large deviation principle with good rate functions

$$\Lambda_0^*(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \Lambda_0(\theta)\} = \frac{\pi x^2}{\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}}, \tag{17}$$

$$\Lambda_1^*(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \Lambda_1(\theta)\} = \frac{\pi}{\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}}\left(x - \frac{\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}}{2\pi}\right)^2. \tag{18}$$

When $0 \leq T \leq \frac{1}{2\pi}\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}$, the error exponent for type I and type II errors are given by

$$\lim_{n\to\infty} -\frac{1}{n}\log\alpha^{(n)} = \lim_{n\to\infty} -\frac{1}{n}\log\Pr\{T_n > T|H_0\} = \inf_{x>T}\Lambda_0^*(x) = \Lambda_0^*(T), \tag{19}$$

$$\lim_{n\to\infty} -\frac{1}{n}\log\beta^{(n)} = \lim_{n\to\infty} -\frac{1}{n}\log\Pr\{T_n \leq T|H_1\} = \inf_{x\leq T}\Lambda_1^*(x) = \Lambda_1^*(T). \tag{20}$$

For the Bayesian problem, the threshold is

$$T = \lim_{n\to\infty}\frac{1}{2n}\mathbf{m}^T\mathbf{\Sigma}^{-1}\mathbf{m} = \lim_{n\to\infty}\frac{1}{2n}\sum_{k=1}^{n}\frac{p_k^2}{\lambda_k} = \frac{1}{4\pi}\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}, \tag{21}$$

thus the error exponent for the average probability $P_e = \pi_0\alpha + \pi_1\beta$ is given by

$$\lim_{n\to\infty} -\frac{1}{n}\log P_e^{(n)} = \lim_{n\to\infty} -\frac{1}{n}\log\alpha^{(n)} = \lim_{n\to\infty} -\frac{1}{n}\log\beta^{(n)} = \frac{1}{16\pi}\int_0^{2\pi}\frac{dG(\omega)}{S(\omega)}. \tag{22}$$

3.2 Distributed Detection for Parallel Access Channel

For PAC, the correlation structure in the source is still retained at the receiver with each sensor directly transmitting an amplified version of the local observation, i.e., $y_k = ax_k$, where the signal amplification a is a constant (independent of the n) chosen to satisfy the average power constraint. Let

$$r'_k = \frac{r_k}{a} = x_k + \frac{z_k}{a}. \tag{23}$$

Then under H_0, $\mathbf{r}'$ is correlated Gaussian with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}' = \mathbf{\Sigma} + \frac{1}{a^2}\mathbf{I}$, and under H_1, $\mathbf{r}'$ is correlated Gaussian with mean $\mathbf{m}$ and covariance matrix $\mathbf{\Sigma}'$. The optimal test is given by

$$T'_n = \frac{1}{n}\mathbf{m}^T\mathbf{\Sigma}'^{-1}\mathbf{r}' \gtrless T. \tag{24}$$

Following similar analysis as in Section 3.1, the expressions of error exponents are the same as for centralized detection, except that $S(\omega)$ is replaced with $S'(\omega) = S(\omega) + \frac{1}{a^2}$. Consequently detection over PAC suffers from a loss in asymptotic performance which is dependent upon a.

3.3 Distributed Detection for Multiple Access Channel

In contrast to PAC which requires a bandwidth proportional to the number of sensors, the bandwidth required by MAC is independent of the number of sensors, hence MAC is attractive for large sensor network applications. With a MAC, however, the fusion center no longer have access to individual sensor observations. Therefore, the mapping rule should be carefully chosen so that the output of the MAC could yield a useful decision statistic for detection. In the following we consider two scenarios depending on whether the sensors have knowledge of source statistics or not.

3.3.1 Intelligent Sensors

In this scenario, we assume that the sensors have knowledge of source statistics. Specifically, each sensor knows $\mathbf{m}$, $\mathbf{\Sigma}$ and its index k. Observe that if we let $\boldsymbol{\gamma} = \mathbf{\Sigma}^{-1}\mathbf{m}$, the optimal decision statistic for centralized detection can be written as

$$T_n = \frac{1}{n}\mathbf{m}^T\mathbf{\Sigma}^{-1}\mathbf{x} = \frac{1}{n}\boldsymbol{\gamma}^T\mathbf{x} = \frac{1}{n}\sum_{k=1}^{n}\gamma_k x_k. \tag{25}$$

Consider *Mapping Rule 1*:

$$U(x_k) : y_k = a\gamma_k x_k. \tag{26}$$

where a is a constant independent of n.

Theorem 1: The threshold test on

$$\tilde{T}_n = \frac{1}{na}r = \frac{1}{n}\sum_{k=1}^{n}\gamma_k x_k + \frac{z}{na} \tag{27}$$

is asymptotically optimal, i.e., achieves the same error exponent as optimal centralized detection for mapping rule 1, and the error exponents do not depend on a.

Proof. For $\tilde{T}_n$, we have for H_0,

$$\Lambda_0^{(n)}(n\theta) = \log E_0\{e^{\theta(\sum_{k=1}^{n}\frac{p_k w_k}{\lambda_k}+\frac{z}{a})}\} = \sum_{k=1}^{n}\frac{\theta^2 p_k^2}{2\lambda_k} + \frac{\theta^2}{2a^2}, \tag{28}$$

Since $\Lambda_0(\theta) = \lim_{n\to\infty}\frac{1}{n}\Lambda_0^{(n)}(n\theta)$, the second term $\frac{\theta^2}{2na^2}$ vanishes asymptotically, and $\Lambda_0(\theta)$ is the same as for optimal centralized detection, and similarly for H_1. Therefore the error exponents are the same as for optimal centralized detection. □

In [1], it is shown that for i.i.d. observations, soft decision fusion (i.e., each sensor transmits the local log-likelihood ratio value) over MAC is asymptotically optimal. Theorem 1 suggests that for correlated observations, with a properly chosen mapping rule, detection over MAC is also asymptotically optimal, provided that the total transmit power scales with the number of sensors n. Mapping rule 1 employs the similar concept as distributed beamforming in cellular systems, where the signals at different antennas are weighted according their respective channel states at the transmitter so that they add constructively at the receiver. In contrast to PAC, the effect of channel noise on MAC is washed out asymptotically. Since the error exponents are independent of the signal amplification a, it is possible to achieve the optimal performance with little average power as long as the number of sensors is sufficiently large.

3.3.2 Dumb Sensors

In large-scale sensor networks, the knowledge of $\mathbf{m}$, $\mathbf{\Sigma}$ and index k may not be available at each sensor. Moreover, the assumption of dumb sensors implies symmetrical sensor functions, which provides the ease of implementation and scalability. In this case, detection over MAC can still be optimal if the observations are i.i.d. by employing type-based distributed detection [1]. For correlated observations, the correlation structure in the source is lost over MAC, and the performance of optimal centralized detection is no longer achievable. In this section we investigate the performance of a suboptimal mapping rule.

Consider *Mapping Rule 2*:

$$U(x_k) : y_k = a x_k^2, \tag{29}$$

where a is independent of n, and the decision statistic given by (recall that $\mathbf{w} = \mathbf{\Phi}^T \mathbf{x}$)

$$\hat{T}_n = \frac{1}{na} r = \frac{1}{n} \sum_{k=1}^{n} x_k^2 + \frac{z}{na} = \frac{1}{n} \sum_{k=1}^{n} w_k^2 + \frac{z}{na}. \tag{30}$$

As for mapping rule 1, we can ignore the effect of channel noise in asymptotic analysis and concentrate on the first term. For $\hat{T}_n$, we have

$$\Lambda_0^{(n)}(n\theta) = \log E_0\{e^{\theta \sum_{k=1}^{n} w_k^2}\} = -\frac{1}{2} \sum_{k=1}^{n} \log(1 - 2\theta\lambda_k), \tag{31}$$

$$\Lambda_1^{(n)}(n\theta) = \log E_1\{e^{\theta \sum_{k=1}^{n} w_k^2}\} = -\frac{1}{2} \sum_{k=1}^{n} \log(1 - 2\theta\lambda_k) + \theta \sum_{k=1}^{n} \frac{(\mathbf{m}^T \boldsymbol{\phi}_k)^2}{1 - 2\theta\lambda_k}. \tag{32}$$

According to Toeplitz distribution theorem and the extended theorem, we obtain

$$\Lambda_0(\theta) = -\frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\theta S(\omega)) \, d\omega, \tag{33}$$

$$\Lambda_1(\theta) = -\frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\theta S(\omega)) \, d\omega + \frac{\theta}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{1 - 2\theta S(\omega)}. \tag{34}$$

In general, the asymptotic logarithmic moment generating functions associated with quadratic functionals are not necessarily steep [12], so a direct application of the Gärtner-Ellis theorem is not possible. However, the results in [12] can be used to show that the rate functions governing the large deviation principle are still given by the Fenchel-Legendre transforms of (33) and (34):

$$\Lambda_0^*(x) = \theta_0 x + \frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\theta_0 S(\omega)) \, d\omega, \tag{35}$$

where θ_0 solves the equation $x = \frac{1}{2\pi} \int_0^{2\pi} \frac{S(\omega)}{1 - 2\theta_0 S(\omega)} \, d\omega$, and

$$\Lambda_1^*(x) = \theta_1 x + \frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\theta_1 S(\omega)) \, d\omega - \frac{\theta_1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{1 - 2\theta_1 S(\omega)}, \tag{36}$$

where θ_1 solves the equation $x = \frac{1}{2\pi} \int_0^{2\pi} \frac{S(\omega)}{1 - 2\theta_1 S(\omega)} \, d\omega + \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{[1 - 2\theta_1 S(\omega)]^2}$. Denote the noise power $\mathbf{\Sigma}_{1,1} = \sigma^2$. When the threshold T satisfies $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} E[v_k^2] \le T \le \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} E[m_k + v_k]^2$, or equivalently, $\sigma^2 \le T \le \sigma^2 + \frac{1}{2\pi} \int_0^{2\pi} dG(\omega)$, the error exponents for type I and type II errors are $\Lambda_0^*(T)$ and $\Lambda_1^*(T)$ given by (35) and (36). Since the average error probability is dominated by the smaller error exponent, the Bayesian error exponent is given by the value at the intersection of both rate functions.

4 Optimal Sensor Spacing: An Example

In this section, we illustrate the performance of various distributed detection strategies, and the application on optimal sensor spacing with an example.

4.1 Detection of a Sinusoid Signal

Consider the detection of a sinusoid signal over a straight line:

$$m(l) = \sqrt{2}m\cos(\omega_0 l), \quad 0 \le l < +\infty. \tag{37}$$

We assume that sensors are equally-spaced over the line with the location of the kth node $d_k = kd, k = 1, \cdots, n$. Therefore, the signal to be detected at the kth node is

$$m_k = \sqrt{2}m\cos(\omega_0 kd) = \sqrt{2}m\cos(\tilde{\omega}_0 k), \quad k = 1, \cdots, n, \tag{38}$$

where $\tilde{\omega}_0 = \omega_0 d$. The autocorrelation function of this signal is given by $r(k) = m^2\cos(k\tilde{\omega}_0)$, and the corresponding spectral density is $dG(\omega) = \pi m^2[\delta(\omega - \tilde{\omega}_0) + \delta(\omega - (2\pi - \tilde{\omega}_0))]$. Note that the constant signal $m(l) = m$ is a special case of this model corresponding to $\tilde{\omega}_0 = 0$, whose spectral density is given by $dG(\omega) = 2\pi m^2\delta(0)$. Assume that the zero mean stationary Gaussian observation noise process has covariance function $\rho(l_1, l_2) = \sigma^2\rho^{|l_1 - l_2|}$. Let $\tilde{\rho} = \rho^d$, then the covariance matrix of $\mathbf{v}$ is given by

$$\mathbf{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \tilde{\rho} & \cdots & \tilde{\rho}^{n-1} \\ \tilde{\rho} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{\rho} \\ \tilde{\rho}^{n-1} & \cdots & \tilde{\rho} & 1 \end{bmatrix}, \tag{39}$$

with spectral density $S(\omega) = \sigma^2 \frac{1-\tilde{\rho}^2}{1+\tilde{\rho}^2 - 2\tilde{\rho}\cos\omega}$.

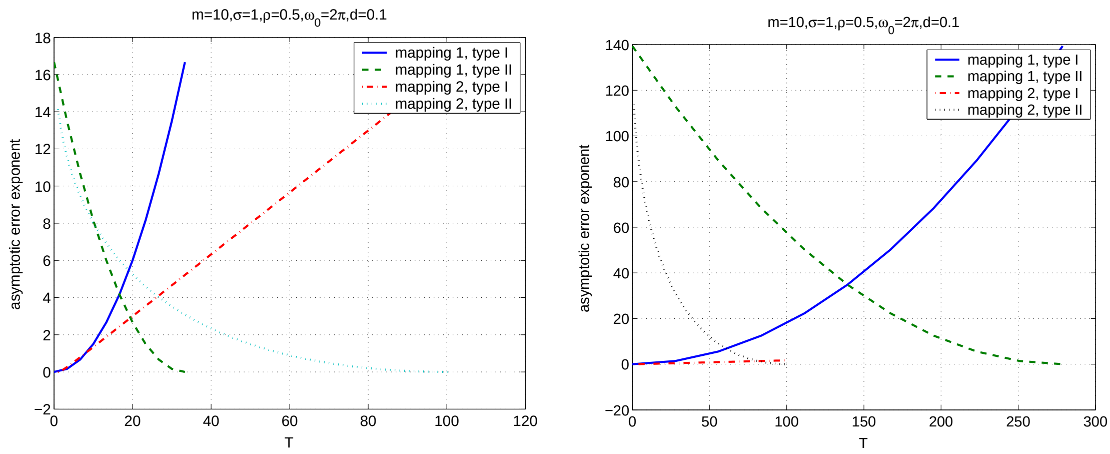
For MAC mapping rule 1, when $0 < T < \frac{m^2}{S(\tilde{\omega}_0)}$, the error exponents of type I and type II errors for the N-P problem are $\frac{S(\tilde{\omega}_0)}{2m^2}T^2$ and $\frac{S(\tilde{\omega}_0)}{2m^2}[T - \frac{m^2}{S(\tilde{\omega}_0)}]^2$ respectively, and the Bayesian error exponent is $\frac{m^2}{8S(\tilde{\omega}_0)}$. For detection over PAC, the error exponents have the same expressions as above except that $S(\tilde{\omega}_0)$ is replaced with $S'(\tilde{\omega}_0) = S(\tilde{\omega}_0) + \frac{1}{a^2}$. For MAC mapping rule 2, when $\sigma^2 < T < m^2 + \sigma^2$, the error exponents of type I and type II errors are $\theta_0 T + \frac{1}{2}\log\{\frac{1}{2}[1 + \tilde{\rho}^2 - 2\theta_0\sigma^2(1 - \tilde{\rho}^2) + \sqrt{[1 + \tilde{\rho}^2 - 2\theta_0\sigma^2(1 - \tilde{\rho}^2)]^2 - 4\tilde{\rho}^2}]\}$, with $T = \frac{(1-\tilde{\rho}^2)\sigma^2}{\sqrt{[1+\tilde{\rho}^2 - 2\theta_0\sigma^2(1-\tilde{\rho}^2)]^2 - 4\tilde{\rho}^2}}$, and
$\theta_1 T + \frac{1}{2}\log\{\frac{1}{2}[1 + \tilde{\rho}^2 - 2\theta_1\sigma^2(1 - \tilde{\rho}^2) + \sqrt{[1 + \tilde{\rho}^2 - 2\theta_1\sigma^2(1 - \tilde{\rho}^2)]^2 - 4\tilde{\rho}^2}]\} - \frac{m^2\theta_1}{1 - 2\theta_1 S(\tilde{\omega}_0)}$, with $T = \frac{(1-\tilde{\rho}^2)\sigma^2}{\sqrt{[1+\tilde{\rho}^2 - 2\theta_1\sigma^2(1-\tilde{\rho}^2)]^2 - 4\tilde{\rho}^2}} + \frac{m^2}{(1 - 2\theta_1 S(\tilde{\omega}_0))^2}$. In Fig. 1, the error exponents for MAC mapping 1 and mapping 2 v.s. threshold T are plotted for $m = 10, \sigma = 1, \rho = 0.5, \omega_0 = 2\pi$, and $d = 1$ and 0.1 respectively. The Bayesian error exponent achieved by mapping 2 is critically determined by the internode spacing d: it is very close to that of mapping 1 when $d = 1$, and much worse than that of mapping 1 when $d = 0.1$.

4.2 Optimal Sensor Spacing

In this section, we study the problem of designing the optimal sensor spacing d to maximize the Bayesian error exponent. For MAC mapping rule 1 and PAC, this is equivalent to minimize

$$S(\tilde{\omega}_0) = \sigma^2 \frac{1 - \rho^{2d}}{1 + \rho^{2d} - 2\rho^d\cos(\omega_0 d)}, \tag{40}$$

which can be solved numerically. For MAC mapping rule 2, the Bayesian error exponents for different values of d are first obtained, from which the optimal d that achieves the



(a) $d = 1$ (b) $d = 0.1$

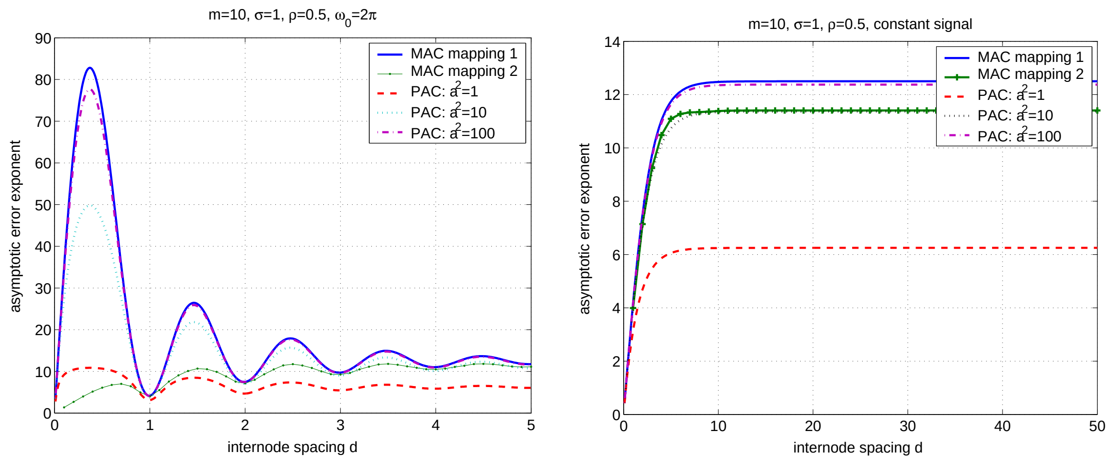
Figure 1: Error Exponents for MAC mapping 1 and mapping 2, $m = 10, \sigma = 1, \rho = 0.5$

global maximum is chosen. Fig. 2(a) shows the error exponents as functions of the inter-node spacing d for both MAC mapping rules as well as PAC with different values of a, where $m = 10, \sigma = 1, \rho = 0.5, \omega_0 = 2\pi$. We observe that in this case, the maximum Bayesian error exponent of mapping rule 2 over distance d is significantly smaller than that of mapping rule 1.

For the special case of a constant signal, we have $\tilde{\omega}_0 = 0$ and that for MAC mapping 1, the optimal d minimizes $S(0) = \sigma^2 \frac{1+\rho^d}{1-\rho^d}$, which is achieved as $d \to \infty$. The performance of various detection strategies when the signal is constant is given in Fig. 2(b), which suggests that for mapping rule 2, the error exponent is also maximized by increasing d. It can be shown that the performance of mapping 2 relative to mapping 1 improves as the frequency ω_0 decreases, and approaches mapping 1 for the constant signal. Note that the sensor field has a finite size and we need a sufficiently large number of sensors to achieve the asymptotic performance predicted by the large deviation theory. As can be seen from Fig. 2(b), the error exponent is almost constant when $d \geq d_0$, where d_0 is a constant depending on ρ. For example, when $\rho = 0.5$, an internode spacing of $d_0 = 7.64$ is sufficient to achieve an error exponent within 1% of the maximum error exponent for mapping rule 1. We finally remark that although the performance of MAC mapping 2 is worse than PAC for certain values of a, it can be achieved with an arbitrarily small average transmit power, hence is of interest for energy-aware sensor networks.

5 Conclusion

We study distributed detection of a deterministic signal in correlated Gaussian noise in large-scale sensor networks using a multiple access channel, which enjoys better bandwidth efficiency as compared to a parallel access channel. We show that for intelligent sensors, distributed detection over MAC can be asymptotically optimal by properly choosing a local mapping rule, while detection over PAC always incurs a loss in error exponent under the average power constraint. For dumb sensors, we propose a suboptimal mapping rule which is shown to perform closely to the intelligent sensor case for certain applications. The analytical results are also used to design optimal sensor spacing for maximizing the Bayesian error exponent.



(a) $\omega_0 = 2\pi$ (b) constant signal

Figure 2: Maximum Bayesian error exponents for various detection strategies, $m = 10, \sigma = 1, \rho = 0.5$

References

[1] K. Liu and A. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," in *Proc. 42nd Annual Allerton Conference on Communications, Control and Computing*, Monticello, IL, Oct. 2004, pp. 703 – 707.

[2] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detecion in power constrained sensor networks," *IEEE J. Selected. Areas Commun.*, vol. 22, pp. 1007–1015, Aug. 2004.

[3] ——, "How dense should a sensor network be for decentralized detection with correlated observations," submitted to IEEE Trans. Inform. Theory.

[4] Y. Sung, L. Tong, and H. V. Poor, "A large deviation approach to sensor scheduling for detection of correlated random fields," in *Proc. ICASSP 05*, Philadelphia, PA, Mar. 2005.

[5] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad, and ugly: distributed detection of a known signal in dependent gaussian noise," *IEEE Trans. Sig. Proc.*, vol. 48, no. 12, pp. 469–476, Dec. 2000.

[6] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer, 1998.

[7] T. Kawata, *Fourier Analysis in Probability*. New York: Academic Press, 1972.

[8] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*. New York: Chelsea Publish, 1958.

[9] R. K. Bahr, "Asymptotic analysis of error probabilities for the nonzero-mean gaussian hypothesis testing problem," *IEEE Trans. Inform. Theory*, vol. 36, no. 3, pp. 597–607, May 1990.

[10] G. R. Benitz and J. A. Bucklew, "Large deviation rate calculations for nonlinear detectors in gaussian noise," *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 358–371, Mar. 1990.

[11] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1994.

[12] B. Bercu, F. Gamboa, and A. Rouault, "Large deviations for quadratic functionals of gaussian processes," *Stochastic Processes and Their Applications*, vol. 71, pp. 75–90, 1997.