

Distributed Detection in Wireless Sensor Networks Using A Multiple Access Channel

Wenjun Li* and Huaiyu Dai

Abstract

Distributed detection in a one-dimensional sensor network with correlated sensor observations, as exemplified by two problems—detection of a deterministic signal in correlated Gaussian noise and detection of a first-order autoregressive signal in independent Gaussian noise, is studied in this paper. In contrast to the traditional approach where a bank of dedicated parallel access channels (PAC) is used for transmitting the sensor observations to the fusion center, we explore the possibility of employing a shared multiple access channel (MAC), which significantly reduces the bandwidth requirement or detection delay. We assume that local observations are mapped according to a certain function subject to a power constraint. Using the large deviation approach, we demonstrate that for the deterministic signal in correlated noise problem, with a specially-chosen mapping rule, MAC fusion achieves the same asymptotic performance as centralized detection under the average power constraint (APC), while there is always a loss in error exponents associated with PAC fusion. Under the total power constraint (TPC), MAC fusion still results in exponential decay in error exponents with the number of sensors, while PAC fusion does not. For the autoregressive signal problem, we propose a suboptimal MAC mapping rule which performs closely to centralized detection for weakly-correlated signals at almost all SNR values, and for heavily-correlated signals when SNR is either high or low. Finally, we show that although the lack of MAC synchronization always causes a degradation in error exponents, such degradation is negligible when the phase mismatch among sensors is sufficiently small. (EDICS: SSP-DETC)

Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695. (email: wli5, huaiyu.dai@ncsu.edu). This research was supported in part by the National Science Foundation under Grant CCF-0515164.

I. INTRODUCTION

Distributed detection of certain events or targets in the environment is an important application of sensor networks [1]. Distributed detection suffers a performance loss caused by the local processing at sensors (mapping, quantization, etc.) as well as the noise in the communication channel. As the advances in hardware technology enabled the dense deployment of low cost sensors, the trend of the detection performance as the number of sensors goes to infinity, as measured by the error exponent, has gained much research interest [2]–[7]. The error exponent gives an estimate of the number of sensors required to reach a certain error probability, and is therefore a useful performance index in the large sample regime [7].

The traditional approach of studying the distributed detection problem is to assume that sensors transmit their observations (possibly quantized versions of them) through a parallel access channel (PAC), which is independent across sensors [1]–[3]. For large-scale sensor networks, this assumption implies a large bandwidth requirement for simultaneous transmissions or a large detection delay. Alternatively, we can employ a multiple access channel (MAC), whose bandwidth requirement does not depend on the number of sensors, but due to the additive nature of the channel, the received signal at the fusion center is generally not sufficient for reliable detection. Recently, Type-Based Multiple Access (TBMA) has been proposed by Mergen and Tong [8]–[11] as well as by Liu and Sayeed [4], [5], which utilizes the MAC to perform distributed detection or estimation. With TBMA, each sensor transmits the waveform corresponding to its quantized observation over a MAC. For i.i.d. sensor observations and identical channel gain, the fusion center receives a noisy version of the type of the sensor observations, which is a sufficient statistic for detection, and TBMA achieves the same error exponent as centralized detection [4], [8]. For i.i.d observations and i.i.d. channel fading, the error probability still decays exponentially with the number of sensors if the channel fading has non-zero mean, but its behavior is sub-exponential if the channel has zero mean fading [11]. Anandkumar and Tong further proposed the Type-

Based Random Access (TBRA) to improve the performance of TBMA over the noncoherent channel by controlling the rate of random access [12]. These works mostly assumed i.i.d. sensor observations, and not much has been explored along the line for correlated observations, which would arise when detecting a random spatially-correlated signal in noise, or a deterministic signal in noise where the noise samples are correlated. Moreover, it is not known how the local quantization should be done to minimize the loss in detection performance.

In this paper, we investigate the possibility of performing distributed detection with correlated observations over a multiple access channel. We consider a one-dimensional (1-D) sensor network with equally-spaced sensors, and investigate two problems of interest: detection of a deterministic signal in correlated Gaussian noise where the noise correlation between each pair of neighboring sensors is the same, and detection of a first-order autoregressive (AR(1)) process in independent Gaussian noise. We consider the simple case where the MAC is perfectly-synchronized (the synchronization issue will be treated in Section IV) and has identical channel gain, as it is known that channel fading complicates the analysis and typically causes performance degradation [9]–[11]. Therefore, if sensors experience different pathlosses, we assume that sensors can perform transmission power control such that the overall channel gain is kept the same for all sensors. We assume that each sensor first maps its (unquantized) raw observation according to a continuous function which is subject to our design. The mapped signal of all sensors are simply amplitude-modulated and transmitted simultaneously, thus the fusion center receives a scalar output after match-filtering. Our approach is to design the local mapping rules such that the channel output comes close to the optimal decision statistic, i.e. the log-likelihood ratio, for centralized decision. No quantization is assumed before or after the local mapping, for two reasons: firstly, the optimal design of local quantizers and fusion rule for Neyman-Pearson and Bayesian detection is generally a difficult problem, and usually only a person-by-person-optimal solution can be found through some iterative methods [1], [13]; secondly, quantization makes it difficult to give a rigorous analysis through the large deviation theory [14], and our result can be regarded as an

estimate of the achievable detection performance if a fine quantization is performed on the after-mapping signals. For both MAC and PAC fusion, the transmitted symbols satisfy a global power constraint, which may be either an average power constraint (APC) or a total power constraint (TPC). Note that only the transmission power is considered in this paper. The extra power consumption required for synchronization and power control over MAC should be included to make a fair comparison between PAC and MAC fusion schemes when designing real systems.

A. Summary of Results and Related Works

For the deterministic signal in correlated Gaussian noise problem, it is observed that with an appropriate choice of mapping rule, the noise-free MAC output gives the optimal decision statistic. In Section III, we show that 1) under APC, our proposed MAC fusion scheme yields the same error exponents as optimal centralized detection, while PAC fusion always incurs a loss in error exponents; 2) under TPC, the proposed MAC scheme still results in exponential decay of the error probability with the number of sensors, while with PAC fusion, the error probability is not reduced by increasing the number of sensors.

Although perfect synchronization on MAC is a basic assumption throughout our paper, synchronization error often can not be completely removed in practical sensor networks. In Section IV, we describe a synchronization scheme motivated by [15], with which the synchronization error is resulted from the sensor placement error, and manifested by the phase mismatch among sensors. We then demonstrate that the phase mismatch always affects the error exponents, but when such mismatch is slight, close-to-optimal performance is still achievable.

The detection of an AR(1) process in independent Gaussian noise is treated in Section V. For this problem, a direct application of MAC does not readily yield the optimal decision statistic, and we propose a suboptimal mapping rule based on our observations of the high SNR and low SNR behavior of the optimal decision statistic. Simulations show that under APC, our proposed MAC fusion scheme results in similar performance as centralized detection for a large range of

correlation coefficients and signal-to-noise ratios, and significantly outperforms PAC fusion.

Several works have studied related detection problems considered in our work, assuming either centralized detection or PAC fusion is employed [2], [6], [7]. The distributed detection over PAC of a constant signal in correlated Gaussian noise, as well as an AR(1) process in independent Gaussian noise have been studied in [2]. The deterministic signal studied in this paper contains the constant signal as a special case. Centralized detection of an AR(1) process has also been studied in [6], [7], and closed-form Type-II error exponent subject to fixed Type-I error bounds for Neyman-Pearson formulation is obtained. Note that our results on error exponents for Neyman-Pearson centralized detection (see Section V-A.) assume a different form from theirs, because when obtaining the Type-II (Type-I) error exponent, we assume exponential rather than fixed error bounds on the Type-I (Type-II) error [16].

B. Notation and Organization

We will make use of the following notational conventions. Scalars are written in normal fonts; column vectors and matrices are in boldface with vectors in lowercases and matrices in capitals. $\text{Re}(z)$ denotes the real part of a complex number z . x_i denotes the i -th entry of the vector \mathbf{x} , and $\det(\mathbf{A})$ and $A_{i,j}$ denotes the determinant and the (i, j) -th element of the matrix \mathbf{A} respectively. $(\cdot)^T$ denotes the transpose. \mathbf{I} is reserved for the identity matrix. \mathbb{R}^k denotes the k -dimensional Euclidean space. $x \sim \mathcal{N}(\mu, \sigma^2)$ (or $x \sim \mathcal{CN}(\mu, \sigma^2)$) means that x is a Gaussian (or complex Gaussian) variable with mean μ and variance σ^2 . $\mathbb{E}(\cdot)$ denotes the expected value, and $\mathbb{E}_i(\cdot)$ with $i \in \{0, 1\}$ denotes the expected value under hypothesis i .

The rest of the paper is organized as follows. In section II, we describe the two detection problems, and provide some necessary theoretical background for our analysis. In section III, the detection of a deterministic signal in correlated noise problem is treated. The performance of MAC fusion under synchronization error is studied in Section IV. Section V deals with the detection of an AR(1) process. Finally, the concluding remarks are contained in Section VI.

II. SYSTEM DESCRIPTION AND PRELIMINARIES

A. Detection Problem

We consider a 1-D sensor network, where the sensors are equally-spaced over a straight line, with the location of the k th node $d_k = kd, k = 1, \dots, n$. We consider the binary hypothesis testing problem where the observation at the k th sensor is given by

$$H_1 : x_k = s_k + v_k, \quad k = 1, 2, \dots, n, \quad (1)$$

$$H_0 : x_k = v_k, \quad k = 1, 2, \dots, n, \quad (2)$$

where $\{s_k\}$ is the signal to be detected, and $\{v_k\}$ is a stationary Gaussian observation noise process independent of $\{s_k\}$.

1) *Detection of Deterministic Signal in Correlated Gaussian Noise*: In this problem, $\{s_k\}$ is a uniformly bounded deterministic signal, i.e., there exists a constant C_m such that $|s_k| \leq C_m, k = 1, 2, \dots, n$. The autocorrelation function of $\{s_k\}$ is recovered from the spectral distribution $G(\cdot)$ by a Riemann-Stieltjes integral

$$R(k) \doteq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n s_{j+k} s_j = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\omega} dG(\omega), \quad (3)$$

and if $G(\cdot)$ is absolutely continuous, its derivative $G'(\omega)$ is the spectral density of $\{s_k\}$ [17]. We assume that the stationary Gaussian noise process $\{v_k\}$ has zero mean and covariance function

$$\rho(k_1, k_2) \doteq \mathbb{E}\{v_{k_1} v_{k_2}\} = \sigma^2 \rho^{|k_1 - k_2|},$$

hence its covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho^{n-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{n-1} & \dots & \rho & 1 \end{bmatrix}. \quad (4)$$

2) *Detection of a First-Order Autoregressive Process in Independent Gaussian Noise:* We consider a first-order autoregressive signal given by [18]

$$\begin{cases} s_1 \sim \mathcal{N}(0, \Pi_0), \\ s_k = \rho s_{k-1} + \mu_k, \quad k = 2, \dots, n, \end{cases} \quad (5)$$

where $0 \leq \rho < 1$, and the innovation $\{\mu_k\}$ is i.i.d. $\mathcal{N}(0, \Pi_0(1 - \rho^2))$. Therefore, $\{s_k\}$ forms a stationary process with $s_k \sim \mathcal{N}(0, \Pi_0)$, $k = 1, \dots, n$. The zero mean stationary Gaussian observation noise process $\{v_k\}$ has covariance matrix $\sigma^2 \mathbf{I}$. The signal-to-noise ratio (SNR) is denoted by $\Gamma = \frac{\Pi_0}{\sigma^2}$.

B. Mapping Rule and Network Communication Channel

We assume that local observations are first mapped through a function $U(\cdot) : y_k = U(x_k)$. y_k is chosen to satisfy a global power constraint, which may be an *average power constraint (APC)*, given by $\frac{1}{n} \sum_{k=1}^n \mathbb{E}\{|y_k|^2\} \leq P_{av}$, or a *total power constraint (TPC)*, given by $\sum_{k=1}^n \mathbb{E}\{|y_k|^2\} \leq P_{tot}$. The mapped signal is then transmitted over one of the following channels:

- 1) **Parallel Access Channel (PAC):** A parallel access channel consists of n dedicated AWGN channels given by

$$r_k = y_k + z_k, \quad k = 1, 2, \dots, n, \quad (6)$$

where the communication noise z_k is i.i.d. $\mathcal{N}(0, 1)$.

- 2) **Multiple Access Channel (MAC):** Unless otherwise specified, in this paper we refer to MAC as a perfectly synchronized Gaussian multiple access channel given by

$$r = \sum_{k=1}^n y_k + z, \quad (7)$$

where $z \sim \mathcal{N}(0, 1)$.

Without loss of generality, we assume that the channel SNR is unity in the above channel models, but any other value of channel SNR can be incorporated with no difficulty, or we may as well think that it is absorbed in the mapping function $U(\cdot)$.

C. 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.

Definition 2.1 (Absolutely Summable Toeplitz Matrix [19], [20]): Let $\Sigma^{(n)}$ be an $n \times n$ Toeplitz matrix with entries $t_k \in \mathbb{R}$ on the k th diagonal and dimension $n \rightarrow \infty$. If $\{t_k\}$ is absolutely summable, i.e., $\sum_{k=-\infty}^{\infty} |t_k| < \infty$, the spectral density function of $\Sigma^{(n)}$ is given by

$$S(\omega) = \sum_{k=-\infty}^{\infty} t_k e^{-ik\omega}, \quad -\pi \leq \omega < \pi.$$

Theorem 2.1 [19]: The eigenvalues $\{\lambda_k\}$ of the absolutely summable Hermitian Toeplitz matrix $\Sigma^{(n)}$ with spectral density $S(\omega)$ are bounded by

$$m_f \leq \lambda_k \leq M_f,$$

where M_f and m_f denote the least upper bound and the greatest lower bound of $S(\omega)$.

Theorem 2.2 (Toeplitz distribution theorem [20] and its extension [21], [22]): Let $\{s_k\}$ be a deterministic signal with spectral distribution $G(\omega)$. For an absolutely summable Toeplitz matrix $\Sigma^{(n)}$ with spectral density $S(\omega)$, let $\{\lambda_k^{(n)}\}$ be the eigenvalues of $\Sigma^{(n)}$ contained on $[\lambda_{min}, \lambda_{max}]$, and $\{\phi_k^{(n)}\}$ be the normalized eigenvectors of $\Sigma^{(n)}$, then for any continuous function $h(\cdot)$ defined on $[\lambda_{min}, \lambda_{max}]$, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n h(\lambda_k^{(n)}) &= \frac{1}{2\pi} \int_0^{2\pi} h(S(\omega)) d\omega, \\ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n h(\lambda_k^{(n)}) (\mathbf{s}^T \phi_k)^2 &= \frac{1}{2\pi} \int_0^{2\pi} h(S(\omega)) dG(\omega). \end{aligned}$$

Theorem 2.3 (Gärtner-Ellis [14]): Let $\{\mathbf{Z}_n\} \in \mathbb{R}^k$ be a sequence of random variables, and define

$$\Lambda^{(n)}(\boldsymbol{\theta}) = \log \mathbb{E}[e^{\boldsymbol{\theta}^T \mathbf{Z}_n}], \quad \boldsymbol{\theta} \in \mathbb{R}^k.$$

Suppose that for each $\boldsymbol{\theta} \in \mathbb{R}^k$, the logarithmic moment generating function, defined as the limit $\Lambda(\boldsymbol{\theta}) \doteq \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda^{(n)}(n\boldsymbol{\theta})$ exists as an extended real number. If $\Lambda(\cdot)$ is an essentially smooth, lower-semicontinuous function as defined in [14], $\{\mathbf{Z}_n\}$ satisfies the large deviation principle (LDP) with rate function Λ^* given by the *Fenchel-Legendre* transform of $\Lambda(\boldsymbol{\theta})$:

$$\Lambda^*(\mathbf{x}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}^k} \{\boldsymbol{\theta}^T \mathbf{x} - \Lambda(\boldsymbol{\theta})\}, \quad \mathbf{x} \in \mathbb{R}^k. \quad (8)$$

That is, for any measurable set B ,

$$- \inf_{\mathbf{x} \in \text{int}(B)} \Lambda^*(\mathbf{x}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \Pr\{\mathbf{Z}_n \in B\} \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \Pr\{\mathbf{Z}_n \in B\} \leq - \inf_{\mathbf{x} \in B^c} \Lambda^*(\mathbf{x}),$$

where $\text{int}(B)$ denotes the interior of B and B^c denotes the closure of B .

Note that the sets of interest in hypothesis testing mostly satisfy the continuity property that $\inf_{\mathbf{x} \in \text{int}(B)} \Lambda^*(\mathbf{x}) = \inf_{\mathbf{x} \in B^c} \Lambda^*(\mathbf{x})$ [11], which implies

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr\{\mathbf{Z}_n \in B\} = - \inf_{\mathbf{x} \in B} \Lambda^*(\mathbf{x}). \quad (9)$$

For Neyman-Pearson hypothesis testing, the optimal detector is a threshold test on the normalized log-likelihood ratio [23]: Choose H_1 if

$$\frac{1}{n} \log \frac{\Pr(\mathbf{x}|H_1)}{\Pr(\mathbf{x}|H_0)} > \tau, \quad (10)$$

and choose H_0 otherwise. Let $\alpha = \Pr\{H_0 \rightarrow H_1\}$ and $\beta = \Pr\{H_1 \rightarrow H_0\}$ denote the type I and type II error probabilities respectively. We are interested in obtaining the error exponent for both types of errors for a given threshold τ :

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha^{(n)}, \quad \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta^{(n)}. \quad (11)$$

For Bayesian formulation with priors $P(H_0) = \pi_0$ and $P(H_1) = \pi_1$, the threshold $\tau = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{\pi_0}{\pi_1} = 0$. Note that the error exponent of the average error probability $\pi_0\alpha + \pi_1\beta$ is dominated by the larger error exponent of type I and type II error, which is therefore minimized when the error exponents for both types of error are equal, i.e.,

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha^{(n)} = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta^{(n)} = \lim_{n \rightarrow \infty} -\frac{1}{n} \log P_e^{(n)}. \quad (12)$$

III. DETECTION OF DETERMINISTIC SIGNAL IN CORRELATED NOISE

A. Optimal Centralized Detection

Optimal centralized detection, where the sensor observation vector \mathbf{x} is perfectly available at the fusion center, serves as a performance baseline for distributed detection strategies. The normalized log-likelihood ratio for centralized detection is given by

$$\frac{1}{n}(\mathbf{s}^T \Sigma^{-1} \mathbf{x} - \frac{1}{2} \mathbf{s}^T \Sigma^{-1} \mathbf{s}). \quad (13)$$

Since $\mathbf{s}^T \Sigma^{-1} \mathbf{s}$ is a constant, we can rewrite the test as

$$T_n = \frac{1}{n} \mathbf{s}^T \Sigma^{-1} \mathbf{x} \geq T. \quad (14)$$

The spectral density function for Σ is $S(\omega) = \sigma^2 \frac{1-\rho^2}{1+\rho^2-2\rho \cos \omega}$.

Proposition 3.1 (Optimal Centralized Detection): For the N-P formulation, when the threshold $0 \leq T \leq \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}$, the error exponents for type I and type II errors are respectively given by

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha^{(n)} = \frac{\pi T^2}{\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}}, \quad (15)$$

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta^{(n)} = \frac{\pi}{\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}} \left(T - \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} \right)^2. \quad (16)$$

For the Bayesian formulation, the error exponent for the average error probability is given by

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log P_e^{(n)} = \frac{1}{16\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}. \quad (17)$$

Proof: Let $\Sigma = \Phi \Lambda \Phi^T$, where $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ is a diagonal matrix containing the eigenvalues of Σ , and Φ is a unitary matrix with eigenvectors of Σ as column vectors. Let $\mathbf{p} = \Phi^T \mathbf{s}$ and $\mathbf{w} = \Phi^T \mathbf{x}$. Then it is easily shown that under H_0 , $\{w_k\}$ is i.i.d. $\mathcal{N}(0, \lambda_k)$, and under H_1 , $\{w_k\}$ is i.i.d. $\mathcal{N}(p_k, \lambda_k)$. Thus we have

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

Correspondingly, we get for the sequence $\{T_n\}$,

$$\begin{aligned}\Lambda_0^{(n)}(n\theta) &= \log \mathbb{E}_0 \left\{ e^{\theta \sum_{k=1}^n \frac{p_k w_k}{\lambda_k}} \right\} = \sum_{k=1}^n \frac{\theta^2 p_k^2}{2\lambda_k} = \frac{\theta^2}{2} \sum_{k=1}^n \frac{(\mathbf{s}^T \boldsymbol{\phi}_k)^2}{\lambda_k}, \\ \Lambda_1^{(n)}(n\theta) &= \log \mathbb{E}_1 \left\{ e^{\theta \sum_{k=1}^n \frac{p_k w_k}{\lambda_k}} \right\} = \sum_{k=1}^n \frac{(\theta^2 + 2\theta) p_k^2}{2\lambda_k} = \frac{\theta^2 + 2\theta}{2} \sum_{k=1}^n \frac{(\mathbf{s}^T \boldsymbol{\phi}_k)^2}{\lambda_k}.\end{aligned}$$

Using Theorem 2.2, the logarithmic moment generating functions under both hypotheses exist and are given by

$$\begin{aligned}\Lambda_0(\theta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_0^{(n)}(n\theta) = \frac{\theta^2}{4\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}, \\ \Lambda_1(\theta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_1^{(n)}(n\theta) = \frac{\theta^2 + 2\theta}{4\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)},\end{aligned}$$

with the Fenchel-Legendre transform

$$\begin{aligned}\Lambda_0^*(x) &= \sup_{\theta \in \mathbb{R}} \{\theta x - \Lambda_0(\theta)\} = \frac{\pi x^2}{\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}}, \\ \Lambda_1^*(x) &= \sup_{\theta \in \mathbb{R}} \{\theta x - \Lambda_1(\theta)\} = \frac{\pi}{\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}} \left(x - \frac{\int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}}{2\pi} \right)^2.\end{aligned}$$

It can be checked that the assumptions for Gärtner-Ellis theorem hold. Thus we have when $0 \leq T \leq \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}$, the error exponents for type I and type II errors are given by

$$\begin{aligned}\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha^{(n)} &= \lim_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{T_n > T | H_0\} = \inf_{x > T} \Lambda_0^*(x) = \Lambda_0^*(T), \\ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta^{(n)} &= \lim_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{T_n \leq T | H_1\} = \inf_{x \leq T} \Lambda_1^*(x) = \Lambda_1^*(T).\end{aligned}$$

The Bayesian error probability follows by setting the threshold to be

$$T = \lim_{n \rightarrow \infty} \frac{1}{2n} \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} = \lim_{n \rightarrow \infty} \frac{1}{2n} \sum_{k=1}^n \frac{p_k^2}{\lambda_k} = \frac{1}{4\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}.$$

■

Remark: We let $T \in [0, \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}]$ such that for both rate functions, the infimum over the set of interest is achieved at the boundary $x = T$. It can be shown that the error exponent for type I error is always 0 for $T \leq 0$, and that for type II error is always 0 for $T \geq \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}$.

B. Distributed Detection over PAC

For PAC, since the Gaussian source and the Gaussian channel is probabilistically matched for an individual sensor [24], least information loss can be ensured by letting the sensors transmit a scaled version of their observations. It is therefore intuitive to study the simple amplify-and-forward strategy for detection over PAC, as in [2]. Denote the average power of s_k by $P_s \doteq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n s_k^2 = \frac{1}{2\pi} \int_0^{2\pi} dG(\omega)$, we have $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(x_k^2) = \sigma^2 + \pi_1 P_s$.

1) *Average Power Constraint:* Under the average power constraint, the mapping rule can be written as $y_k = ax_k$, where $a = \sqrt{\frac{P_{av}}{\sigma^2 + \pi_1 P_s}}$ is a constant independent of n . Here and henceforth, we assume that the scaling factor controlling the transmission power is computed by the fusion center and broadcasted to all sensors prior to the detection.

Let $r'_k = \frac{r_k}{a} = x_k + \frac{z_k}{a}$. The optimal test becomes

$$T'_n = \frac{1}{n} \mathbf{s}^T \boldsymbol{\Sigma}'^{-1} \mathbf{r}' \geq T, \quad (18)$$

where $\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} + \frac{1}{a^2} \mathbf{I}$ is the covariance matrix of \mathbf{r}' . Following similar analysis as in Section III. A, 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 under the average power constraint, detection over PAC suffers from a loss in asymptotic performance which is dependent upon a .

2) *Total Power Constraint:* Under the total power constraint, the mapping rule becomes $y_k = \frac{a}{\sqrt{n}} x_k$, where $a = \sqrt{\frac{P_{tot}}{\sigma^2 + \pi_1 P_s}}$ is a constant independent of n . Then $r''_k = \frac{\sqrt{n} r_k}{a} = x_k + \frac{\sqrt{n} z_k}{a}$ has covariance matrix $\boldsymbol{\Sigma}'' = \boldsymbol{\Sigma} + \frac{n}{a^2} \mathbf{I}$ under both hypotheses. Note that $\boldsymbol{\Sigma}''$ is not absolutely summable, hence the theorems in Section II can not be applied to analyze this case. As a matter of fact, simulation shows that under the total power constraint, PAC fusion no longer results in exponential decay in the average error probability (see Section III. D). This result can be understood intuitively: as the number of sensors approaches infinity, the total signal power is constrained while the channel noise grows without bound, making it impossible for the fusion center to perform reliable detection.

C. Distributed Detection over MAC

With a MAC, the fusion center no longer have access to individual sensor observations. Therefore, the mapping rule should be carefully chosen so that the received signal yields a useful decision statistic for detection. Observe that if we let $\boldsymbol{\gamma} = \boldsymbol{\Sigma}^{-1}\mathbf{s}$, the optimal decision statistic for centralized detection can be written as

$$T_n = \frac{1}{n}\mathbf{s}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} = \frac{1}{n}\boldsymbol{\gamma}^T\mathbf{x} = \frac{1}{n}\sum_{k=1}^n\gamma_k x_k. \quad (19)$$

Assuming sensor k is informed of γ_k , if sensor k transmits a scaled version of $\gamma_k x_k$, the noise-free output of MAC readily yields the decision statistic (except for the factor $\frac{1}{n}$). Through some computations we obtain

$$\gamma_k = \frac{1}{\sigma^2(1-\rho^2)} \cdot \begin{cases} s_1 - \rho s_2, & k = 1, \\ (1 + \rho^2)s_k - \rho(s_{k-1} + s_{k+1}), & k = 2, \dots, n-1, \\ s_n - \rho s_{n-1}, & k = n. \end{cases} \quad (20)$$

This knowledge can be obtained, e.g., from the fusion center or neighboring sensors, at an initial stage.

1) *Average Power Constraint:* Under the average power constraint, the mapping rule is given by $y_k = a\gamma_k x_k$, where $a = \sqrt{\frac{P_{av}}{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(\gamma_k^2 x_k^2)}}^1$.

Theorem 3.1 (Asymptotic Optimality of MAC Fusion Under Average Power Constraint):

For the mapping rule $y_k = a\gamma_k x_k$, where a is a constant independent of n , the threshold test on

$$T_n^{APC} = \frac{1}{na}r = \frac{1}{n}\sum_{k=1}^n\gamma_k x_k + \frac{z}{na} \quad (21)$$

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

¹ $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(\gamma_k^2 x_k^2)$ exists because $\{s_k\}$ is uniformly bounded. The exact value of a can be computed for specific applications—see the example in Section III.D

Proof: For T_n^{APC} , we have for H_0 ,

$$\Lambda_0^{(n)}(n\theta) = \log \mathbb{E}_0 \left\{ e^{\theta(\sum_{k=1}^n \frac{p_k w_k}{\lambda_k} + \frac{z}{a})} \right\} = \sum_{k=1}^n \frac{\theta^2 p_k^2}{2\lambda_k} + \frac{\theta^2}{2a^2}, \quad (22)$$

Since $\Lambda_0(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_0^{(n)}(n\theta)$, 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. \blacksquare

Theorem 3.1 suggests that under the average power constraint, detection over MAC can be asymptotically optimal with each sensor transmitting a scaled version of its observation such that they add constructively at the fusion center. Intuitively, MAC has only one instantiation of the channel noise, but the total transmission power grows without bound under APC. Thus the signal-to-noise ratio at the fusion center approaches infinity asymptotically, which is essentially the centralized detection case. The optimal performance can be achieved irrespective of the average power as long as the number of sensors is large.

2) *Total Power Constraint:* Under the total power constraint, the mapping rule is given by $y_k = \frac{a}{\sqrt{n}} \gamma_k x_k$, where $a = \sqrt{\frac{P_{tot}}{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(\gamma_k^2 x_k^2)}}$.

Theorem 3.2 (MAC Fusion Under Total Power Constraint): For the mapping rule $y_k = \frac{a}{\sqrt{n}} \gamma_k x_k$, where a is a constant independent of n , the threshold test on

$$T_n^{TPC} = \frac{1}{\sqrt{na}} r = \frac{1}{n} \sum_{k=1}^n \gamma_k x_k + \frac{z}{\sqrt{na}} \quad (23)$$

yields suboptimal asymptotic performance: when the threshold $T \in [0, \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)}]$, the error exponents for type I and type II errors are $\frac{T^2}{2} \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} + \frac{1}{a^2} \right)^{-1}$ and $\frac{1}{2} \left(T - \frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} \right)^2 \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} + \frac{1}{a^2} \right)^{-1}$ respectively, and the Bayesian error exponent for average error probability is $\frac{1}{2} \left(\frac{1}{4\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} \right)^2 \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} + \frac{1}{a^2} \right)^{-1}$.

Proof: For T_n^{TPC} , we have for H_0 ,

$$\Lambda_0^{(n)}(n\theta) = \log \mathbb{E}_0 \left\{ e^{\theta(\sum_{k=1}^n \frac{p_k w_k}{\lambda_k} + \frac{\sqrt{n}z}{a})} \right\} = \sum_{k=1}^n \frac{\theta^2 p_k^2}{2\lambda_k} + \frac{n\theta^2}{2a^2},$$

TABLE I

BAYESIAN ERROR EXPONENTS FOR CENTRALIZED AND DISTRIBUTED DETECTION SCHEMES

Centralized Detection	PAC under APC	MAC under APC	MAC under TPC
$\frac{m^2}{8S(\bar{\omega}_0)}$	$\frac{m^2}{8 S(\bar{\omega}_0) + \frac{1}{a^2}}$	$\frac{m^2}{8S(\bar{\omega}_0)}$	$\frac{m^2}{8S(\bar{\omega}_0)} \frac{1}{1 + \frac{S(\bar{\omega}_0)}{m^2 a^2}}$

hence the logarithmic generating function

$$\Lambda_0(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_0^{(n)}(n\theta) = \frac{\theta^2}{4\pi} \int_0^{2\pi} \frac{dG(\omega)}{S(\omega)} + \frac{\theta^2}{2a^2},$$

and similarly for H_1 . The error exponents follow readily through similar analysis as in Section III. A. ■

Comparing the above result with proposition 3.1, we observe that the asymptotic optimality of MAC fusion is lost if each sensor is forced to use diminishing power as the number of sensors increases, and the loss is reflected in the term $\frac{1}{a^2}$. Nevertheless, under the total power constraint, MAC fusion is still preferable than PAC fusion as it results in exponential decay in error probability.

D. Numerical Example: Detection of A Sinusoid Signal

Consider a sinusoid signal over a straight line:

$$s_k = \sqrt{2}m \cos(\omega_0 k), \quad k = 1, \dots, n. \quad (24)$$

The spectral density of this signal is $G'(\omega) = \pi m^2 [\delta(\omega - \omega_0) + \delta(\omega - (2\pi - \omega_0))]$. Note that the constant signal $s_k = m$ is a special case of this model corresponding to $\omega_0 = 0$. The theoretical Bayesian error exponents using different detection strategies are summarized in Table I, where $S(\omega_0) = \sigma^2 \frac{1-\rho^2}{1+\rho^2-2\rho \cos \omega_0}$.

Fig. 1 and Fig. 2 plot the simulated Bayesian error probabilities and the corresponding error exponents for centralized detection and various distributed detection schemes, where we have assumed that the two hypothesis are equiprobable, i.e., $\pi_0 = \pi_1 = \frac{1}{2}$, and used the

TABLE II

SCALING FACTORS AND BAYESIAN ERROR EXPONENTS FOR CENTRALIZED AND DISTRIBUTED DETECTION SCHEMES,

$$\rho = 0.5, m = \sigma = 1, \omega_0 = \pi/4, P_{av} = 2, \text{ AND } P_{tot} = 10$$

	Centralized	PAC under APC	PAC under TPC	MAC under APC	MAC under TPC
a	N/A	1.15	2.58	1.47	3.30
error exponent	0.0905	0.0586	N/A	0.0905	0.0803

following parameters: $\rho = 0.5$, $m = \sigma = 1$, $\omega_0 = \frac{\pi}{4}$, $P_{av} = 2$, and $P_{tot} = 10$. To obtain the scaling factor for MAC fusion, we observe that for the periodic cosine function with $\omega_0 = \frac{\pi}{4}$, $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(\gamma_k^2 x_k^2) = \frac{1}{4} \sum_{k=1}^4 \gamma_k^2 (\frac{s_k^2}{2} + \sigma^2)$, where γ_k is given by (20). The corresponding scaling factors as well as the theoretical Bayesian exponents are given in Table II. It can be seen from Fig. 1 and Fig. 2 that the error exponents for various detection schemes approach the predicted values as the number of sensors becomes larger. The MAC fusion scheme achieves similar performance as centralized detection under APC, and the difference in error exponent becomes increasingly small as the number of sensors increases. Under both APC and TPC, the MAC fusion scheme significantly outperforms PAC fusion. In particular, under TPC, the error probability for MAC fusion still decays exponentially with the number of sensors, while the error probability for PAC fusion does not decrease with the number of sensors.

IV. MAC SYNCHRONIZATION AND THE IMPACT OF SYNCHRONIZATION ERROR

An important assumption for the proposed MAC fusion schemes to achieve the predicted asymptotic performance is the perfect synchronization among sensors, which is difficult to realize in large-scale sensor networks. In this section, we analyze the impact of synchronization error on the performance of MAC fusion.

The sensor synchronization required in our application can be achieved with essentially the same strategy as described in [15], but instead of taking a star topology, we assume that the

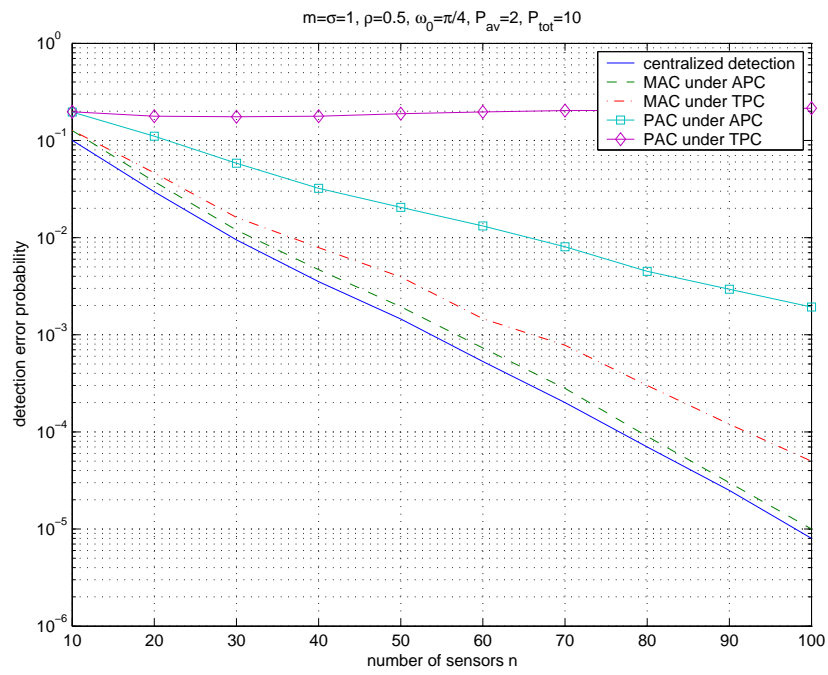


Fig. 1. Error probabilities for detection of a sinusoid signal, $m = 1$, $\sigma = 1$, $\rho = 0.5$, $\omega_0 = \pi/4$, $P_{av} = 2$, $P_{tot} = 10$

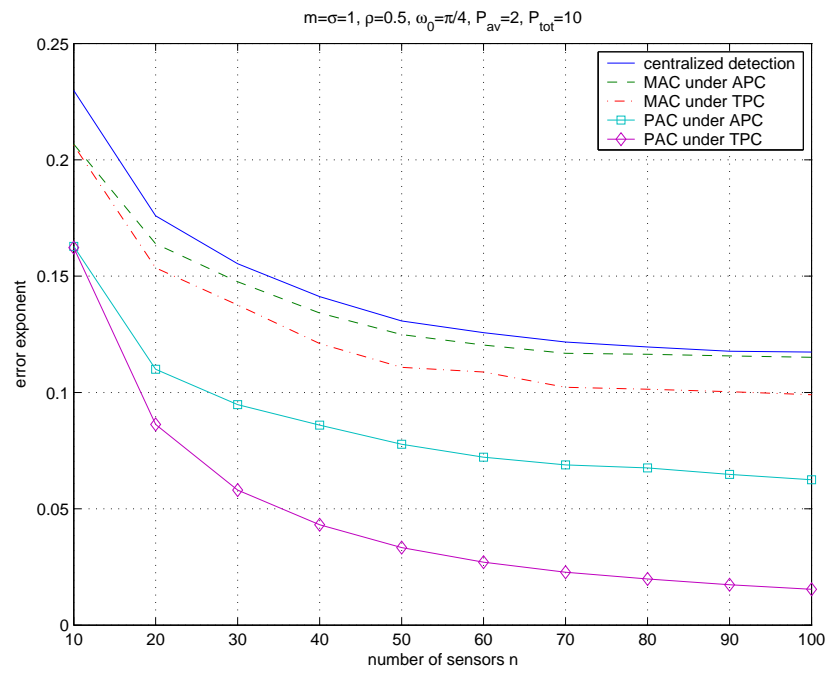


Fig. 2. Error exponents for detection of a sinusoid signal, $m = 1$, $\sigma = 1$, $\rho = 0.5$, $\omega_0 = \pi/4$, $P_{av} = 2$, $P_{tot} = 10$

synchronization head is a node located at coordinate 0 on the line. Prior to all transmissions, the synchronization head broadcasts a carrier signal and each sensor employs a phase-locked loop (PLL) to lock onto the carrier. If every sensor pre-compensate for the difference in their distances to the synchronization head by transmitting its signal with a proper delay and phase shift, the resultant timing error and phase error will assume the same form as described in [15]. Specifically, we assume the exact coordinate for the k th sensor node is $d_k = kd + \delta_k$, where δ_k is the placement error. Usually the placement error is small such that the effect of timing error, such as the intersymbol interference, can be neglected, and we only need to consider the effect of phase error [15]. Under such an assumption, the received baseband signal at the fusion center is

$$r = \text{Re} \left\{ a \sum_{k=1}^n y_k e^{j\varphi_k} + z \right\},$$

where $z \sim \mathcal{CN}(0, 1)$, and $\varphi_k = -\frac{4\pi f_0 \delta_k}{c}$ is the phase error for sensor k , with f_0 denoting the carrier frequency and c denoting the speed of light. In the foregoing analysis, we focus on the detection problem considered in Section III with average power constraint, in which case the decision statistic is given by

$$\tilde{T}_n = \frac{1}{na} r = \frac{1}{n} \sum_{k=1}^n \gamma_k x_k \cos \varphi_k + \frac{\text{Re}(z)}{na}. \quad (25)$$

A. General Analysis

In this section we show that under the general assumption that $\{\varphi_k\}$ is i.i.d. and independent of $\{x_k\}$, the presence of phase mismatch always results in a loss in the error exponent of the average Bayesian error probability. The logarithmic moment generating function is given by (note that the noise term vanishes under the average power constraint)

$$\tilde{\Lambda}_i(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_i^{(n)}(n\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_i \left\{ e^{\theta \sum_{k=1}^n \gamma_k x_k \cos \varphi_k} \right\}, \quad (26)$$

where $i \in \{0, 1\}$, and the expectation is with respect to both x_k 's and φ_k 's. Since e^x is convex, Jensen's inequality gives

$$\tilde{\Lambda}_i(\theta) \geq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_i \left\{ e^{\theta \mathbb{E}[\cos \varphi] \sum_{k=1}^n \gamma_k x_k} \right\}, \quad (27)$$

with the equality held when φ_k is a constant. It can be shown that the Fenchel-Legendre transform of the right-hand-side of (27), denoted by $\hat{\Lambda}_i^*(x)$, is

$$\hat{\Lambda}_i^*(x) = \Lambda_i^* \left(\frac{x}{\mathbb{E}[\cos \varphi]} \right), \quad (28)$$

where $\Lambda_i^*(x)$ is the rate function governing the LDP associated with T_n . It follows from the definition of the Fenchel-Legendre transform in (8) that

$$\tilde{\Lambda}_i^*(x) \leq \hat{\Lambda}_i^*(x). \quad (29)$$

Note that the Bayesian error exponent is equal to the value of both rate functions at their intersection, which remains the same under the scaling of x in both rate functions in (28). Denote the Bayesian exponent under perfect synchronization by B and that under phase mismatch by \tilde{B} . From (28) and (29) we obtain $B \geq \tilde{B}$, unless φ_k is a constant, i.e., sensors are perfectly synchronized.

B. Performance Subject to Small Phase Error

In this section, we assume that $\varphi_k \sim \mathcal{N}(0, \sigma_\varphi^2)$, and σ_φ is small. This is true when the placement error δ_k is a zero mean Gaussian variable with the standard deviation σ_δ much smaller than the wavelength $\lambda_0 = \frac{c}{f_0}$. For example, when $f_0 = 10\text{MHz}$, $\sigma_\delta = 0.1\text{m}$, we have $\sigma_\varphi = \frac{4\pi f_0 \sigma_\delta}{c} = 0.0133\pi$. In the following, we seek an upper bound on the performance loss due to phase mismatch. Taking expectation with respect to φ and using the approximation $\sin \varphi_k \approx \varphi_k$ for small φ_k , we obtain

$$\mathbb{E}_\varphi \left\{ e^{\theta \sum_{k=1}^n \gamma_k x_k \cos \varphi_k} \right\} \approx \mathbb{E}_\varphi \left\{ e^{\theta \sum_{k=1}^n \gamma_k x_k \left(1 - \frac{\varphi_k^2}{2} \right)} \right\} = \prod_{k=1}^n \frac{e^{\theta \gamma_k x_k}}{\sqrt{1 + \theta \gamma_k x_k \sigma_\varphi^2}}. \quad (30)$$

Since x_k is Gaussian, there exists a constant ϱ such that the probability of $|\theta\gamma_k x_k \sigma_\varphi^2| \geq \varrho$ is negligible. Then we can obtain a constant $M_\varrho \geq 1$ with which $(1 + \theta\gamma_k x_k \sigma_\varphi^2)^{-\frac{1}{2}} \leq M_\varrho e^{-\frac{1}{2}\theta\gamma_k x_k \sigma_\varphi^2}$ holds. Therefore we have

$$\tilde{\Lambda}_i(\theta) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_i \left\{ e^{\theta(1 - \frac{1}{2}\sigma_\varphi^2) \sum_{k=1}^n \gamma_k x_k} \right\} + \log M_\varrho, \quad (31)$$

which yields

$$\tilde{\Lambda}_i^*(x) \geq \Lambda_i^* \left(\frac{x}{1 - \frac{1}{2}\sigma_\varphi^2} \right) - \log M_\varrho. \quad (32)$$

Then we have $\tilde{B} \geq B - \log M_\varrho$, and the relative loss is bounded by

$$\epsilon \leq \frac{\log M_\varrho}{B}. \quad (33)$$

Generally, little performance loss can be ensured with a relatively small σ_φ . To illustrate this we take the detection of a constant signal $s_k = m$, $k = 1, \dots, n$ as an example. Note that for Bayesian detection $\theta_0 = -\theta_1 = 0.5$, and for a constant signal $\gamma_k = \frac{(1-\rho)m}{(1+\rho)\sigma^2}$, $k = 2, \dots, n-1$, and $B = \frac{m^2}{8\sigma^2} \frac{1-\rho}{1+\rho}$. If we ignore the probability of $|x_k| > |m| + 3\sigma$, we obtain

$$|\theta\gamma_k x_k \sigma_\varphi^2| \leq 0.5\sigma_\varphi^2 \left[\left(\frac{|m|}{\sigma} \right)^2 + 3 \frac{|m|}{\sigma} \right] \frac{1-\rho}{1+\rho} \equiv \varrho. \quad (34)$$

Under the following set of parameters: $\frac{|m|}{\sigma} = 2$, $\rho = 0.5$, $\sigma_\varphi = 0.1\pi$, we obtain $\varrho = 0.1648$, $M_\varrho = 1.0077$, $B = 0.1667$, and the relative loss $\epsilon \leq 4.6\%$. When σ_φ is decreased to 0.1 with other parameters kept the same, we have $\epsilon \leq 0.06\%$. Fig. 3 and Fig. 4 depict the simulated error probabilities and the corresponding error exponents with $\pi_0 = \pi_1 = 1/2$, $m = 1$, $\sigma = 0.5$, $\rho = 0.5$ and $P_{av} = 1$. We observe that when $\sigma_\varphi = 0.1\pi$, there is only a slight loss in the performance of MAC fusion compared with the perfect-synchronized case. When $\sigma_\varphi = 0.2\pi$, however, the performance degrades significantly, but MAC fusion still largely outperforms PAC fusion. When σ_φ is increased to 0.5π , the error probability of MAC fusion no longer decreases exponentially with the number of sensors.

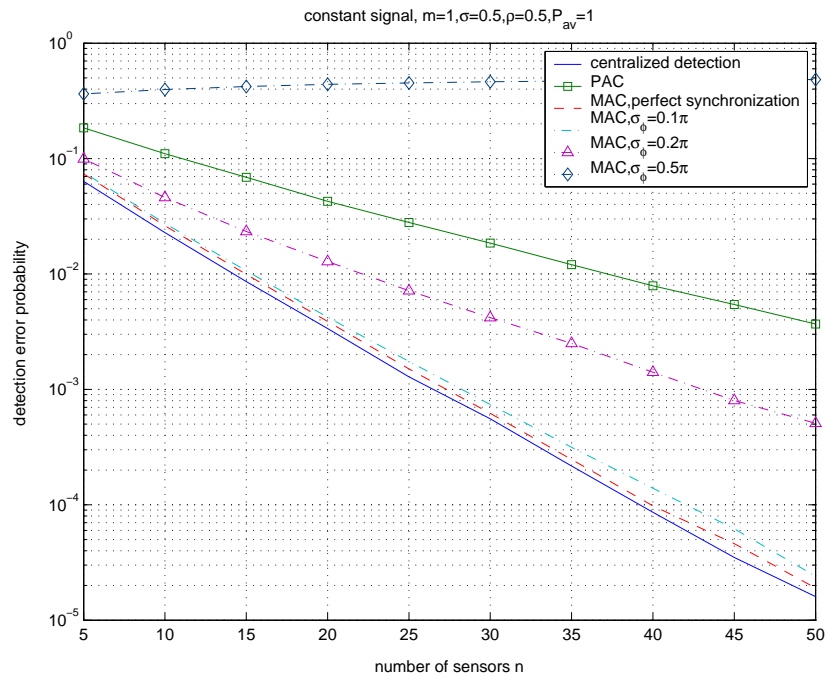


Fig. 3. Error probabilities for detection of constant signal under average power constraint, $m = 1$, $\sigma = 0.5$, $\rho = 0.5$, $P_{av} = 1$

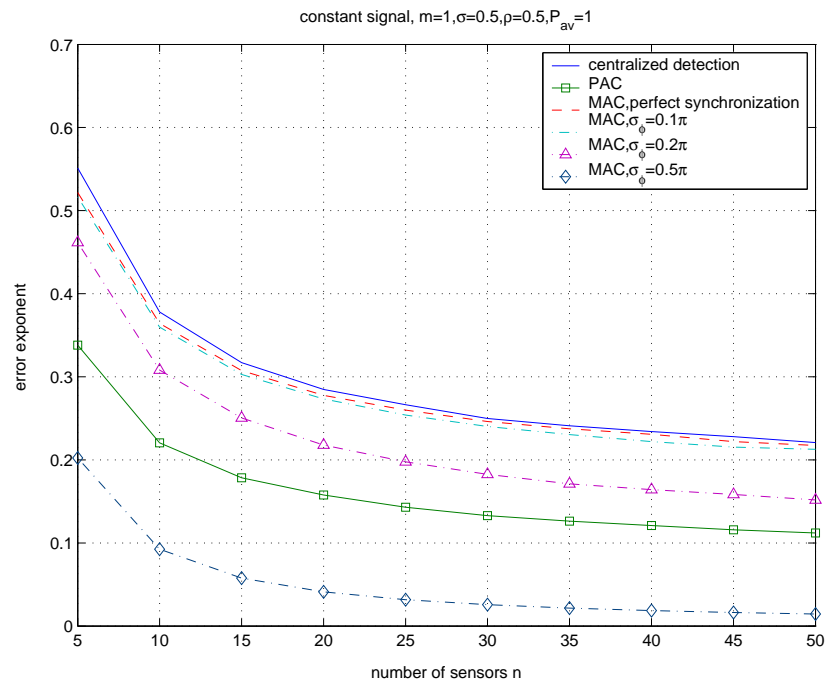


Fig. 4. Error exponents for detection of constant signal under average power constraint, $m = 1$, $\sigma = 0.5$, $\rho = 0.5$, $P_{av} = 1$

V. DETECTION OF A FIRST-ORDER AUTOREGRESSIVE PROCESS

A. Optimal Centralized Detection

For the first-order autoregressive signal described in Section II. A, the optimal centralized detection is a threshold test on the normalized LLR

$$\frac{1}{n} \mathbf{x}^T (\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1}) \mathbf{x} + \frac{1}{n} \log \frac{\det(\boldsymbol{\Sigma}_0)}{\det(\boldsymbol{\Sigma}_1)}, \quad (35)$$

where $\boldsymbol{\Sigma}_0 = \sigma^2 \mathbf{I}$, $\boldsymbol{\Sigma}_1 = \Pi_0$ $\begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{n-1} & \cdots & \rho & 1 \end{bmatrix} + \sigma^2 \mathbf{I}$ is the covariance matrix of \mathbf{x} under H_0 and H_1 . We can rewrite the test as

$$T_n = \frac{1}{n} \mathbf{x}^T (\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1}) \mathbf{x} \geq T. \quad (36)$$

Proposition 5.1 (Optimal Centralized Detection): For the N-P formulation, when the threshold

$\frac{\Gamma(1-\rho^2)}{\sqrt{[1+\rho^2+\Gamma(1-\rho^2)]^2-4\rho^2}} \leq T \leq \Gamma$, where $\Gamma = \frac{\Pi_0}{\sigma^2}$, the error exponents for type I and type II errors are respectively given by

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha^{(n)} = \frac{1}{2} (T - \xi_1 + \log \xi_2 - \log C), \quad (37)$$

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta^{(n)} = \frac{1}{2} (-\xi_1 + \log \xi_2 - \log 2), \quad (38)$$

where $C = 1 + \rho^2 + \Gamma(1 - \rho^2) + \sqrt{[1 + \rho^2 + \Gamma(1 - \rho^2)]^2 - 4\rho^2}$, $\xi_1 = \frac{\sqrt{\Gamma^2(1-\rho^2)^2+4\rho^2T^2-(1+\rho^2)T}}{\Gamma(1-\rho^2)}$, and $\xi_2 = \frac{\sqrt{\Gamma^2(1-\rho^2)^2+4\rho^2T^2+\Gamma(1-\rho^2)}}{T}$. The Bayesian error exponent is given by (37) or (38), with $T = \log \frac{C}{2}$.

Proof: See Appendix I.

We assume average power constraint in the following discussion of the distributed detection schemes, and the case of total power constraint can be analyzed by the same means as above and similar conclusions can be obtained.

B. Distributed Detection over PAC

Under the average power constraint, we have $y_k = ax_k$, where $a = \sqrt{\frac{P_{av}}{\sigma^2 + \pi_1 \Pi_0}}$. Let $r'_k = \frac{r_k}{a} = x_k + \frac{z_k}{a}$. Then the covariance matrix of $\mathbf{r}'_{\mathbf{k}}$ under H_0 and H_1 is $\Sigma'_0 = (\sigma^2 + \frac{1}{a^2}) \mathbf{I}$ and $\Sigma'_1 = \Pi_0 \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{n-1} & \cdots & \rho & 1 \end{bmatrix} + (\sigma^2 + \frac{1}{a^2}) \mathbf{I}$ respectively. The problem becomes the detection of the same autoregressive signal in Gaussian noise with covariance matrix $(\sigma^2 + \frac{1}{a^2}) \mathbf{I}$, and all the results in Section V.A follow with the SNR $\Gamma = \frac{\Pi_0}{\sigma^2}$ being replaced with $\Gamma' = \frac{\Pi_0}{\sigma^2 + \frac{1}{a^2}}$.

C. Distributed Detection over MAC

Denote $\mathbf{A} = \Sigma_0^{-1} - \Sigma_1^{-1}$. The decision statistic for optimal centralized detection involves the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i x_j$, where the double sum can not be accomplished through a direct application of MAC. In the following we explore the quadratic form in more detail and show that in the low or high SNR region, a suboptimal decision statistic which can be realized through MAC achieves similar performance as centralized detection.

Denote $\mathbf{B} = \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{n-1} & \cdots & \rho & 1 \end{bmatrix}$, then $\mathbf{B}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+\rho^2 & -\rho \\ 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}$. The spectral density of \mathbf{B} is $\frac{1-\rho^2}{1+\rho^2-2\rho\cos\omega}$, whose least upper bound and greatest lower bound are

$M_f = \frac{1+\rho}{1-\rho}$, and $m_f = \frac{1-\rho}{1+\rho}$. Thus by Theorem 2.1, the eigenvalue of \mathbf{B} , denoted as λ , satisfies $\frac{1-\rho}{1+\rho} \leq \lambda \leq \frac{1+\rho}{1-\rho}$.

Low SNR Behavior: We can write \mathbf{A} as $\mathbf{A} = \frac{1}{\sigma^2} (\mathbf{I} - (\Gamma \mathbf{B} + \mathbf{I})^{-1})$. When $\Gamma |\lambda| < 1$, or

$\Gamma < \frac{1-\rho}{1+\rho}$, Taylor expansion gives

$$\mathbf{A} = \frac{1}{\sigma^2} (\mathbf{I} - \mathbf{I} + \Gamma\mathbf{B} - (\Gamma\mathbf{B})^2 + \dots) = \frac{1}{\sigma^2} (\Gamma\mathbf{B} - (\Gamma\mathbf{B})^2 + \dots). \quad (39)$$

In particular, when $\Gamma \ll \frac{1-\rho}{1+\rho}$, $\mathbf{A} \approx \frac{1}{\sigma^2}\Gamma\mathbf{B}$, which is a Toeplitz matrix with the magnitude of the k th diagonal decreasing geometrically with k . In general, for $\Gamma < \frac{1-\rho}{1+\rho}$, it can be shown that $|A_{ij}|$ decreases with $|i-j|$ due to the dominance of $\frac{1}{\sigma^2}\Gamma\mathbf{B}$.

High SNR Behavior: We can alternatively write \mathbf{A} as $\mathbf{A} = \frac{1}{\sigma^2} \left(\mathbf{I} - \frac{1}{\Gamma}\mathbf{B}^{-1} (\mathbf{I} + \frac{1}{\Gamma}\mathbf{B}^{-1})^{-1} \right) = \frac{1}{\sigma^2} (\mathbf{I} + \frac{1}{\Gamma}\mathbf{B}^{-1})^{-1}$. When $\frac{1}{\Gamma}|\lambda^{-1}| < 1$, or $\Gamma > \frac{1+\rho}{1-\rho}$, we have

$$\mathbf{A} = \frac{1}{\sigma^2} \left(\mathbf{I} - \frac{1}{\Gamma}\mathbf{B}^{-1} + \left(\frac{1}{\Gamma}\mathbf{B}^{-1} \right)^2 - \dots \right), \quad (40)$$

where for $(\mathbf{B}^{-1})^k$, the $0, \dots, k$ th diagonals are nonzero, while the $k+1, \dots, n$ th diagonals are all zero. Therefore the magnitude of the entries on the k th diagonal of \mathbf{A} also decreases with k due to the coefficient $\frac{1}{\Gamma^k}$.

Note that $E(x_i x_j) = \pi_1 \rho^{|i-j|}$ also diminishes with $|i-j|$. The above demonstrates that when the SNR satisfies $\Gamma < \frac{1-\rho}{1+\rho}$ or $\Gamma > \frac{1+\rho}{1-\rho}$, the contribution of the terms corresponding to $|i-j| > 1$ in the quadratic form $\sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i x_j$ is small. Therefore, replacing these terms with their expected values results in little change in the value of the quadratic form. This motivates us to explore the suboptimal decision statistic

$$\frac{1}{n} \sum_{i=1}^n \sum_{|i-j| \leq 1}^n A_{ij} x_i x_j + \frac{1}{n} \pi_1 \left(\sum_{i=1}^n \sum_{|i-j| > 1}^n A_{ij} \rho^{|i-j|} \right). \quad (41)$$

The first term in (41) can be accomplished through a MAC with sensor k transmitting y_k , where

$$y_k = a \cdot \begin{cases} A_{1,1} x_1^2, & k = 1, \\ A_{k,k} x_k^2 + 2A_{k,k-1} x_{k-1} x_k, & k = 2, \dots, n. \end{cases} \quad (42)$$

The scaling factor a for controlling the transmission power is not straightforward, but can be obtained numerically. The above suggests that each sensor needs to obtain the local observation of its previous sensor. Thus each decision interval is made up of two phases:

- 1) Phase I (Local information exchange) consists of 4 time slots: in slot i ($i \in \{1, 2, 3, 4\}$), all sensors with index k satisfying $\text{mod}(k, 4) = i - 1$ transmit their own observations to their respective next neighbor on the line.
- 2) Phase II: Sensor k computes y_k using (42), and transmit simultaneously. The fusion center receives $r = \sum_{k=1}^n y_k + z$, and computes $T'_n = \frac{1}{n} \left[r + \pi_1 \left(\sum_{i=1}^n \sum_{|i-j|>1}^n A_{ij} \rho^{|i-j|} \right) \right]$.

We assume that local transmissions are done with power control such that the received observations are noiseless (this is a reasonable assumption given the fact that the distances from interfering transmissions are at least 3 times of the distance from the desired transmission, and the radiating power decays like d^{-4} with the distance d on the ground). Under the average power constraint, the channel noise causes no loss in error exponents achievable with MAC fusion, and the only loss comes from the use of the approximate rather than the exact decision statistic.

Fig. 5(a) shows the simulated error exponents when $\Pi_0 = 1, \sigma = 1, \rho = 0.5$, and $P_{av} = 1$. The scaling factor for MAC fusion is also obtained via simulation to ensure the power constraint is met. We observe that our proposed MAC fusion scheme performs closely to centralized detection, and significantly outperforms the PAC fusion scheme under the same average power constraint. Note that in this case, $\Gamma = 1$, while the condition under which our algorithm is guaranteed to work well is $\Gamma < 1/3$ or $\Gamma > 3$. It is seen that although our algorithm has been built based on the assumption of $\Gamma < \frac{1-\rho}{1+\rho}$ or $\Gamma > \frac{1+\rho}{1-\rho}$, when Γ is in between the two values, good performance can still be attained for weakly-correlated signals ($\rho \leq 0.5$), with which the contribution of $E(x_i x_j) = \pi_1 \rho^{|i-j|}$ vanishes fast with $|i - j|$. However, such a property is lost as signals becomes highly-correlated, and we need the SNR to be either low or high to guarantee that MAC fusion still achieves close-to-optimal performance. In Fig. 5(b)-(d) we plot the simulated error exponents when $\rho = 0.8, P_{av} = 1$, and the SNR Γ equals $1/9$ ($\Pi_0 = 1, \sigma = 3$), 1 ($\Pi_0 = 1, \sigma = 1$) and 3 ($\Pi_0 = 3, \sigma = 1$) respectively. It can be seen that when $\Gamma = 1/9$, our proposed MAC fusion scheme achieves roughly the same performance as centralized detection. When $\Gamma = 1$, MAC fusion performs considerably worse than centralized detection, but it still

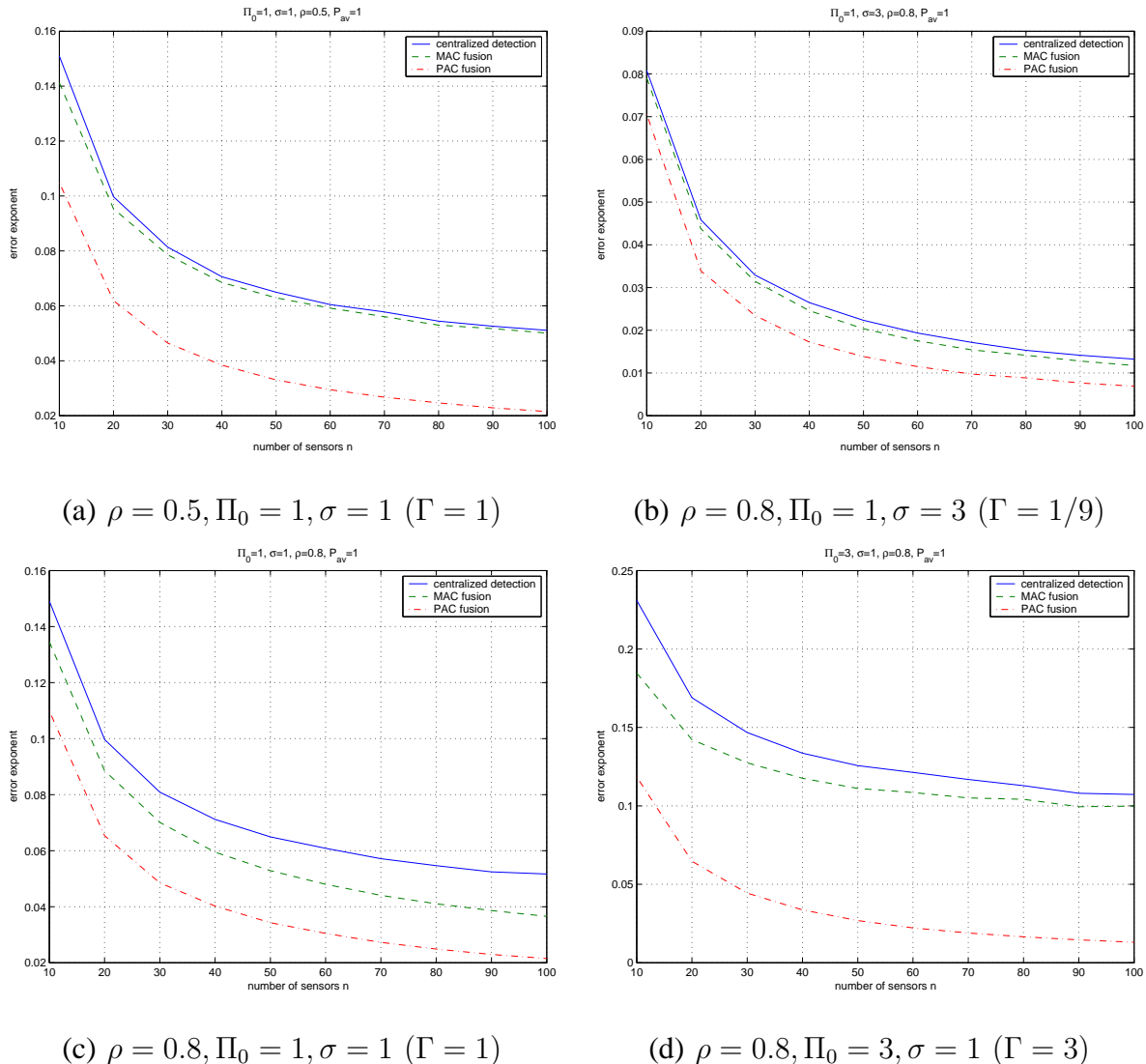


Fig. 5. Error exponents for detection of autoregressive signal under average power constraint, $P_{av} = 1$

largely outperforms PAC fusion. As SNR is increased to $\Gamma = 3$, the performance gap between MAC fusion and centralized detection becomes small again.

VI. CONCLUSION

In this paper, we study two distributed detection problems involving correlated sensor observations in a one-dimensional sensor network: detection of a deterministic signal in correlated Gaussian noise, and detection of a first-order autoregressive (AR(1)) signal in independent

Gaussian noise. In contrast to the traditional approach where a bank of dedicated parallel access channels (PAC) is used for transmitting the sensor observations to the fusion center, we explore the possibility of employing a shared multiple access channel (MAC), which significantly reduces the bandwidth requirement or detection delay.

Using the large deviation approach, we demonstrate that for the detection of a deterministic signal in correlated Gaussian noise, with a specially-chosen mapping rule, MAC fusion achieves the same asymptotic performance as centralized detection under the average power constraint (APC), while there is always a loss in error exponents associated with PAC fusion; under the total power constraint (TPC), MAC fusion still results in exponential decay in error probability with the number of sensors, while PAC fusion does not. For the detection of an AR(1) process in independent Gaussian noise problem, we propose a suboptimal MAC mapping rule which performs closely to centralized detection for weakly-correlated signals at almost all SNR values, and for heavily-correlated signals when SNR is either high or low. We also investigate the performance of MAC fusion under sensor synchronization error, and show that the performance degradation is negligible when the phase mismatch among sensors is sufficiently small.

Although MAC fusion enjoys better bandwidth efficiency and detection performance than PAC fusion, it does involve more overhead and its application is limited. Throughout the paper we have only considered the transmission power, and ignored the power consumption for MAC synchronization and for obtaining certain required information, such as the parameter γ_k in the deterministic signal in correlated noise problem. If the number of samples each sensor actually takes for testing a certain signal is small, the overhead for obtaining the required information would appear fairly large. PAC fusion is also more universal compared with MAC fusion. The fusion center can test for several different deterministic signals (or AR(1) processes with different values of ρ and Π_0) after receiving the noisy channel output using PAC fusion; for MAC fusion, however, each different deterministic signal (or AR(1) process) would require a reuse of the MAC. Furthermore, for a general detection problem, it is not clear whether or not there exists

a good mapping rule for MAC fusion that yields desirable performance.

APPENDIX I

PROOF OF PROPOSITION 5.1

Denote the spectral decomposition $\Sigma_1 = \Phi \Lambda \Phi^T$, where $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ contains the eigenvalues of Σ_1 , and Φ contains the normalized eigenvectors of Σ_1 , and let $\mathbf{w} = \Phi^T \mathbf{x}$. Then under H_0 , $\{w_k\}$ is i.i.d. $\mathcal{N}(0, \sigma^2)$, and under H_1 , $\{w_k\}$ is i.i.d. $\mathcal{N}(0, \lambda_k)$. Thus we have

$$T_n = \frac{1}{n} \sum_{k=1}^n \left(\frac{1}{\sigma^2} - \frac{1}{\lambda_k} \right) w_k^2.$$

Correspondingly, we get for $\{T_n\}$,

$$\begin{aligned} \Lambda_0^{(n)}(n\theta) &= \log \mathbb{E}_0 \left\{ e^{\theta \sum_{k=1}^n \frac{1}{\sigma^2 - \frac{1}{\lambda_k}} w_k^2} \right\} = -\frac{1}{2} \sum_{k=1}^n \log \left[1 - 2\theta \left(1 - \frac{\sigma^2}{\lambda_k} \right) \right], \\ \Lambda_1^{(n)}(n\theta) &= \log \mathbb{E}_1 \left\{ e^{\theta \sum_{k=1}^n \frac{1}{\sigma^2 - \frac{1}{\lambda_k}} w_k^2} \right\} = -\frac{1}{2} \sum_{k=1}^n \log \left[1 - 2\theta \left(\frac{\lambda_k}{\sigma^2} - 1 \right) \right]. \end{aligned}$$

Let the spectral density of Σ_1 be $S(\omega) = \Pi_0 \frac{1-\rho^2}{1+\rho^2-2\rho\cos\omega} + \sigma^2$. Using Theorem 2.2, we obtain the logarithmic generating functions

$$\begin{aligned} \Lambda_0(\theta) &= -\frac{1}{4\pi} \int_0^{2\pi} \log \left[1 - 2\theta \left(1 - \frac{\sigma^2}{S(\omega)} \right) \right] d\omega \\ &= -\frac{1}{2} \log \frac{1 + \rho^2 + (1 - 2\theta)\Gamma(1 - \rho^2) + \sqrt{[1 + \rho^2 + (1 - 2\theta)\Gamma(1 - \rho^2)]^2 - 4\rho^2}}{1 + \rho^2 + \Gamma(1 - \rho^2) + \sqrt{[1 + \rho^2 + \Gamma(1 - \rho^2)]^2 - 4\rho^2}} \end{aligned} \quad (43)$$

where $\theta \leq \frac{1}{2} \left(\frac{1-\rho}{\Gamma(1-\rho)} + 1 \right)$, and

$$\begin{aligned} \Lambda_1(\theta) &= -\frac{1}{4\pi} \int_0^{2\pi} \log \left[1 - 2\theta \left(\frac{S(\omega)}{\sigma^2} - 1 \right) \right] d\omega \\ &= -\frac{1}{2} \log \left\{ \frac{1}{2} \left[1 + \rho^2 - 2\theta\Gamma(1 - \rho^2) + \sqrt{[1 + \rho^2 - 2\theta\Gamma(1 - \rho^2)]^2 - 4\rho^2} \right] \right\}, \end{aligned} \quad (44)$$

where $\theta \leq \frac{1}{2\Gamma(1-\rho)}$.

The Fenchel-Legendre transforms of (43) and (44) are respectively given by:

$$\Lambda_0^*(x) = \theta_0 x - \Lambda_0(\theta_0), \quad (45)$$

$$\Lambda_1^*(x) = \theta_1 x - \Lambda_1(\theta_1), \quad (46)$$

where θ_0 solves the equation $x = \frac{\Gamma(1-\rho^2)}{\sqrt{[1+\rho^2+(1-2\theta_0)\Gamma(1-\rho^2)]^2-4\rho^2}}$, and θ_1 solves the equation $x = \frac{\Gamma(1-\rho^2)}{\sqrt{[1+\rho^2-2\theta_1\Gamma(1-\rho^2)]^2-4\rho^2}}$.

Denote $C = 1 + \rho^2 + \Gamma(1 - \rho^2) + \sqrt{[1 + \rho^2 + \Gamma(1 - \rho^2)]^2 - 4\rho^2}$, $\xi_1 = \frac{\sqrt{\Gamma^2(1-\rho^2)^2 + 4\rho^2 T^2 - (1+\rho^2)T}}{\Gamma(1-\rho^2)}$, and $\xi_2 = \frac{\sqrt{\Gamma^2(1-\rho^2)^2 + 4\rho^2 T^2 + \Gamma(1-\rho^2)}}{T}$. Using Theorem 2.3, we obtain for the N-P formulation, when the threshold T is chosen between

$$\lim_{n \rightarrow \infty} E(T_n | H_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left(1 - \frac{\sigma^2}{\lambda_k}\right) = \frac{1}{2\pi} \int_0^{2\pi} \left(1 - \frac{\sigma^2}{S(\omega)}\right) d\omega = \frac{\Gamma(1-\rho^2)}{\sqrt{[1+\rho^2+\Gamma(1-\rho^2)]^2-4\rho^2}},$$

$$\text{and } \lim_{n \rightarrow \infty} E(T_n | H_1) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left(\frac{\lambda_k}{\sigma^2} - 1\right) = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{S(\omega)}{\sigma^2} - 1\right) d\omega = \Gamma,$$

the error exponents for type I and type II errors are $\Lambda_0^*(T) = \frac{1}{2}(T - \xi_1 + \log \xi_2 - \log C)$ and $\Lambda_1^*(T) = \frac{1}{2}(-\xi_1 + \log \xi_2 - \log 2)$ respectively. For the Bayesian formulation, the threshold is set to be

$$T = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{|\sum_1|}{|\sum_0|} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \log \frac{\lambda_k}{\sigma^2} = \frac{1}{2\pi} \int_0^{2\pi} \log \frac{S(\omega)}{\sigma^2} d\omega = \log \frac{C}{2}.$$

ACKNOWLEDGEMENT

The authors wish to thank Professor Fulvio Gini and the anonymous reviewers for several helpful suggestions.

REFERENCES

- [1] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I-fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [2] J.-F. Chamberland and V. V. Veeravalli, "How dense should a sensor network be for decentralized detection with correlated observations," submitted to *IEEE Trans. Inform. Theory*.
- [3] —, "Asymptotic results for decentralized detection in power constrained sensor networks," *IEEE J. Selected. Areas Commun.*, vol. 22, pp. 1007–1015, Aug. 2004.
- [4] 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.
- [5] —, "Type-based decentralized detection in wireless sensor networks," submitted to *IEEE Transactions on Signal Processing*, Oct 2005.
- [6] 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.
- [7] —, "Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponent and properties," to appear in *IEEE Trans. Inform. Theory*, 2006.

- [8] G. Mergen and L. Tong, "Asymptotic detection performance of type-based multiple access in sensor networks," in *Proc. IEEE Sixth International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2005)*, New York City, NY, 2005.
- [9] —, "Estimation over deterministic multi-access channels," in *Proc. 42nd Annual Allerton Conference on Communications, Control and Computing*, Monticello, IL, Oct. 2004.
- [10] —, "Type-based estimation over multiaccess channels," to appear in *IEEE Transactions on Signal Processing*.
- [11] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," submitted to *IEEE Transactions on Signal Processing*.
- [12] A. Anandkumar and L. Tong, "Type-based random access for distributed detection over multiaccess fading channels," submitted to *IEEE Trans. Signal Processing*, Dec 2005.
- [13] T. M. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 2, pp. 469–476, Apr. 1998.
- [14] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer, 1998.
- [15] G. Barriac, R. Mudumbai, and U. Madhow, "Distributed beamforming for information transfer in sensor networks," in *Proc. 3rd Int. Symp. on Information Processing in Sensor Networks*, Berkeley, CA, Apr. 2004.
- [16] P.-N. Chen, "General formulas for the neyman-pearson type-II error exponent subject to fixed and exponential type-I error bounds," *IEEE Trans. Inform. Theory*, vol. 42, no. 1, pp. 316–323, Jan. 1996.
- [17] T. Kawata, *Fourier Analysis in Probability*. New York: Academic Press, 1972.
- [18] C. Chatfield, *The analysis of time series: An introduction*, 4th ed. London: Chapman and Hall, 1989.
- [19] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Stanford: Free book, 2002.
- [20] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*. New York: Chelsea Publish, 1958.
- [21] 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.
- [22] 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.
- [23] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1994.
- [24] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.