

Neyman-Pearson Detection of Gauss-Markov Signals in Noise: Closed-Form Error Exponent and Properties

Youngchul Sung and Lang Tong
 School of Electrical and Computer Engineering
 Cornell University
 Ithaca, NY 14850, USA
 Email: {ys87, ltong}@ece.cornell.edu

H. Vincent Poor
 Dept. of Electrical Engineering
 Princeton University
 Princeton, NJ 08544
 Email: poor@princeton.edu

Abstract—The performance of Neyman-Pearson detection of correlated stochastic signals using noisy observations is investigated via the error exponent for the miss probability with a fixed level. Using the state-space structure of the signal and observation model, a closed-form expression for the error exponent is derived, and the connection between the asymptotic behavior of the optimal detector and that of the Kalman filter is established. The properties of the error exponent are investigated for the scalar case. It is shown that the error exponent has distinct characteristics with respect to correlation strength: for signal-to-noise ratio (SNR) > 1 the error exponent decreases monotonically as the correlation becomes stronger, whereas for SNR < 1 there is an optimal correlation that maximizes the error exponent for a given SNR.

I. INTRODUCTION

We consider in this paper the detection of correlated stochastic signals using noisy observations y_i under the Neyman-Pearson formulation. The corresponding null and alternative hypotheses are given by

$$\begin{aligned} H_0 &: y_i = w_i, \quad i = 1, 2, \dots, n, \\ H_1 &: y_i = s_i + w_i, \quad i = 1, 2, \dots, n, \end{aligned} \quad (1)$$

where $\{w_i\}$ is independent and identically distributed (i.i.d.) $\mathcal{N}(0, \sigma^2)$ noise with a known variance σ^2 , and $\{s_i\}$ is a stationary signal process correlated in time. Specifically, we assume that $\{s_i\}$ is a Gauss-Markov process following a state-space model given by

$$\begin{aligned} s_{i+1} &= a s_i + u_i, \quad i = 1, \dots, n, \\ s_1 &\sim \mathcal{N}(0, \Pi_0), \\ u_i &\stackrel{i.i.d.}{\sim} \mathcal{N}(0, Q), \quad Q = \Pi_0(1 - a^2), \end{aligned} \quad (2)$$

where a and Π_0 are known scalars with $0 \leq a \leq 1$ and $\Pi_0 > 0$. We assume that the process noise $\{u_i\}$ is independent of the measurement noise $\{w_i\}$ and the initial state s_1 is uncorrelated with u_i for all i . Thus, the signal samples $\{s_i\}$ form an autoregressive sequence. Notice that the value of a determines the amount of correlation between signal samples. For $a = 0$ we have i.i.d. signal samples. On the other hand, $a = 1$ gives perfectly correlated signals. Notice also that the noisy observations $\{y_i\}$ are not autoregressive under H_1 ; they follow the hidden Markov model. Due to the stationarity of the signal, the signal-to-noise ratio (SNR) for the observations is constant and is given by

$$\text{SNR} = \frac{\Pi_0}{\sigma^2}. \quad (3)$$

This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564. Prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011.

The work of H. V. Poor was supported in part by the Office of Naval Research under Grant N00014-03-1-0102.

In this paper, we are interested in the large-sample error performance of the Neyman-Pearson detector for the hypotheses (1) with a given level. In many cases, the miss probability P_M with a fixed size decays exponentially as the sample size increases, and the *error exponent* is defined as the rate of exponential decay, i.e.,

$$K \triangleq \lim_{n \rightarrow \infty} -\frac{1}{n} \log P_M \quad (4)$$

under the given size constraint (i.e., the false alarm probability $P_F \leq \alpha$). The error exponent is a good performance index for detectors in the large sample regime since it gives an estimate of the number of samples required for a given detector performance; faster decay rate implies that fewer samples are needed for a given miss probability. For the case of i.i.d. samples where each sample is drawn independently from the common null probability density p_0 or alternative density p_1 , the error exponent under the fixed size constraint is given by the Kullback-Leibler distance $D(p_0||p_1)$ between the two densities p_0 and p_1 (C. Stein [23]). For more general cases, the error exponent is given by the asymptotic Kullback-Leibler rate defined as the almost-sure limit of

$$\frac{1}{n} \log \frac{p_{0,n}}{p_{1,n}}(y_1, \dots, y_n) \quad \text{as } n \rightarrow \infty, \quad (5)$$

under $p_{0,n}$, where $p_{0,n}$ and $p_{1,n}$ are the null and alternative joint densities of y_1, \dots, y_n , respectively, assuming that the limit exists [24]–[28]. However, the closed-form calculation of (5) is available only for restricted cases. One such example is the discrimination between two autoregressive (AR) signals with distinct parameters under the two hypotheses [26], [27]. In this case, the joint density, $p_{j,n}$, is easily decomposed using the Markov property under each hypothesis, and the calculation of the rate is straightforward. However, for the problem of (1) this approach is not tractable since the observation samples under the alternative hypothesis do not possess the Markov property due to the additive noise.

A. Summary of Results

Our approach to this problem is to exploit the state-space model. The state-space approach in detection is well established in calculation of the log-likelihood ratio (LLR) for correlated signals [6], [10]. With the state-space model, the LLR is expressed through the innovations representation [9] and the innovations are easily obtained by the Kalman filter. The key idea for the closed-form calculation of the error exponent for the hidden Markov model is based on the properties of innovations. Since the innovations process is independent from time to time, the joint density under H_1 is given by the product of marginal densities of the innovations, and the LLR is given by a function of the sum of squares of the innovations; this functional form facilitates the closed-form calculation of (5).

By applying this state-space approach, we derive a closed-form expression for the error exponent K for the miss probability of the Neyman-Pearson detector for (1) of fixed size α and interpret the derived error exponent by making a connection with the empirical distribution in a finite alphabet case. We also make a connection between the state-space approach and the spectral domain approach via the canonical spectral factorization.

We next investigate the properties of the error exponent via the obtained closed-form expression. We show that the error exponent K is a function of the SNR and the correlation, and has distinct behavior with respect to (w.r.t.) the correlation strength depending on the SNR. We show a sharp phase transition at SNR = 1: at high SNR, K decreases monotonically as a function of the correlation, while at low SNR, on the other hand, there exists an optimal correlation value that yields the maximal K .

We also make a connection between the asymptotic behavior of the Kalman filter and that of the Neyman-Pearson detector. It is shown that the error exponent is determined by the asymptotic (or steady-state) variances of the innovations under H_0 and H_1 together with the noise variance.

B. Related Work

The detection of Gauss-Markov processes in Gaussian noise is a classical problem. See [5] and references therein. Our work focuses on the performance analysis as measured by the error exponent, and relies on the connection between the likelihood ratio and the innovations process as described by Schweppe [6]. In addition to the calculation of the LLR, the state-space approach has been used in the performance analysis in this detection problem. Exploiting the state-space model, Schweppe obtained a differential equation for the Bhattacharyya distance between two Gaussian processes [8], which gives an upper bound on the average error probability under a Bayesian formulation [7].

There is an extensive literature on the large deviations approach to the analysis of the detection of Gauss-Markov processes [16]-[22]. Most of these results rely on the extension of Cramer's theorem by Gärtner and Ellis [13], [14], [15] and the properties of the asymptotic eigenvalue distributions of Toeplitz matrices [11], [12]. To find the rate function, however, this approach usually involves an optimization that requires nontrivial numerical methods except in some simple cases, and the rate is given as an integral of the spectrum of the observation process; closed-form expressions are difficult to obtain except for the case of a noiseless AR process in discrete-time and its continuous-time counterpart, the Ornstein-Uhlenbeck process [20], [21]. In addition, most results have been obtained for a fixed threshold for the normalized LLR test, which results in expressions for the rate as a function of the threshold. For ergodic cases, however, the normalized LLR converges to a constant under the null hypothesis and the false alarm probability also decays exponentially for a fixed threshold. Hence, a detector with a fixed threshold is not optimal in the Neyman-Pearson sense since it does not use the level constraint fully; i.e., the optimal threshold is a function of sample size.

II. ERROR EXPONENT AND PROPERTIES

In this section, we present a closed-form expression of the error exponent K of miss probability defined in (4) and examine the properties of the error exponent w.r.t. correlation strength and the SNR of observations.

Theorem 1 (Error exponent): For the Neyman-Pearson detector of the hypotheses (1, 2) with level $\alpha \in (0, 1)$ (i.e. $P_F \leq \alpha$) and $0 \leq$

$\alpha \leq 1$, the error exponent of miss probability is independent of α and is given by

$$K = -\frac{1}{2} \log \frac{\sigma^2}{R_e} + \frac{1}{2} \frac{\tilde{R}_e}{R_e} - \frac{1}{2}, \quad (6)$$

where R_e and \tilde{R}_e are the steady-state variances of the innovation process of $\{y_i\}$ calculated under H_1 and H_0 , respectively. Specifically, R_e and \tilde{R}_e are given by

$$R_e = P + \sigma^2, \quad (7)$$

$$\tilde{R}_e = \sigma^2 \left(1 + \frac{a^2 P^2}{P^2 + 2\sigma^2 P + (1-a^2)\sigma^4} \right), \quad (8)$$

$$P = \frac{\sqrt{[\sigma^2(1-a^2) - Q]^2 + 4\sigma^2 Q} - \sigma^2(1-a^2) + Q}{2}, \quad (9)$$

In the frequency domain,

$$K = \frac{1}{2\pi} \int_0^{2\pi} D(\mathcal{N}(0, \sigma^2) || \mathcal{N}(0, S_y(\omega))) d\omega, \quad (10)$$

where $D(\cdot || \cdot)$ is the Kullback-Leibler distance, and the spectrum $S_y(\omega)$ of $\{y_i\}$ under H_1 is given by

$$S_y(\omega) = \sigma^2 + \frac{\Pi_0(1-a^2)}{1-2a \cos \omega + a^2}. \quad (11)$$

Proof: See [29]. ■

Theorem 1 follows from the fact that the almost-sure limit (5) of the normalized log-likelihood ratio under H_0 is the error exponent for general ergodic cases [24]-[27]. To obtain the closed-form calculation of error exponent for the hidden Markov structure of $\{y_i\}$, we express the log-likelihood ratio through the innovations representation [6]; the log-likelihood ratio is given by a function of the sum of squares of the innovations on which the strong law of large numbers (SLLN) is applied. The calculated innovations are true in the sense that they form an independent sequence only under H_1 , i.e., when the signal actually comes from the state-space model. It is worth noting that \tilde{R}_e is the steady-state variance of the ‘‘innovations’’ calculated as if the observations result from the alternative, but it is actually from the null hypothesis. In this case, the ‘‘innovation’’ sequence becomes the output of a recursive filter driven by an i.i.d. process $\{y_i\}$ since the Kalman filter converges to the recursive Wiener filter for time-invariant stable systems.

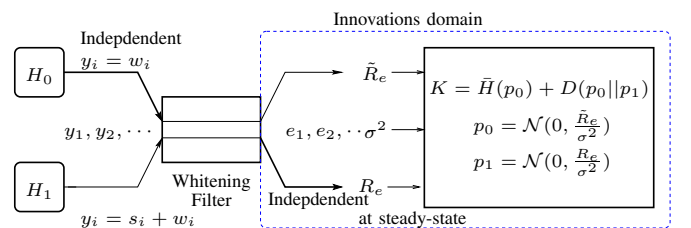


Fig. 1. Interpretation via whitening filter.

Our innovations approach provides an interpretation for the error exponent in Theorem 1. We can rewrite (6) as

$$\begin{aligned} K &= \frac{1}{2} \log \frac{\tilde{R}_e}{\sigma^2} + \left(-\frac{1}{2} \log \frac{\tilde{R}_e}{R_e} + \frac{1}{2} \frac{\tilde{R}_e}{R_e} - \frac{1}{2} \right), \\ &= \bar{H}(p_0) + D(p_0 || p_1), \end{aligned} \quad (12)$$

where the normalized entropy $\bar{H}(p_0) \triangleq \frac{1}{2} \log \frac{\tilde{R}_e}{\sigma^2}$ so that the differential entropy becomes zero when $\tilde{R}_e = \sigma^2$, and the normalized marginal null and alternative distributions in the steady-state innovations domain $p_0 = \mathcal{N}\left(0, \frac{\tilde{R}_e}{\sigma^2}\right)$ and $p_1 = \mathcal{N}\left(0, \frac{R_e}{\sigma^2}\right)$. Now, consider a finite alphabet case where X_i is drawn i.i.d. from a distribution

Q defined on a set $\mathcal{X} = \{a_1, \dots, a_{|\mathcal{X}|}\}$. It is well known that the probability that a sequence $\mathbf{x}_n \triangleq \{X_1, \dots, X_n\}$ has empirical distribution $P_{\mathbf{x}}$ is given by [4]

$$\Pr\{\mathbf{x}_n \text{ has empirical distribution } P_{\mathbf{x}}\} = 2^{-n(H(P_{\mathbf{x}}) + D(P_{\mathbf{x}}||Q))}. \quad (13)$$

Note that the i.i.d. assumption is required only for the true underlying distribution Q not for the empirical distribution $P_{\mathbf{x}}$ in (13). In our Neyman-Pearson problem the error exponent in Theorem 1 is for the probability of a miss event for which the true distribution for y_i is not i.i.d. In the innovations domain,¹ however, the observation process is whitened and the variance of the innovation converges eventually so that the innovations process becomes an i.i.d. sequence with variance R_e in the steady state under H_1 . Thus, the derived error exponent (12) expressed by the marginal distributions in the steady-state innovations domain matches the exponent in (13), which implies that the error of missing can be viewed simply as the event that the observation sequence (in the innovations domain) generated i.i.d. from $\mathcal{N}(0, \frac{R_e}{\sigma^2})$ (the alternative distribution) has empirical distribution $\mathcal{N}(0, \frac{R_e}{\sigma^2})$ (the null distribution).

The frequency-domain result is obtained by representing the log-likelihood ratio in terms of the eigenvalues of the autocovariance matrix of observations and using the convergence of the eigenvalues of Toeplitz matrices [11], [12]. The equivalence of the innovations approach and spectral domain approach is shown by the canonical spectral factorization. Note that the relationship between the error exponent and the asymptotic properties of the Kalman filter is evident in (6). The asymptotic variances of the innovations under the two hypotheses determine the error exponent for the Neyman-Pearson detector.

A. Properties of Error Exponent

First, it is easily seen from Theorem 1 that K is a continuous function of the correlation coefficient a ($0 \leq a \leq 1$) for a given SNR, and is positive for all values of SNR and $0 \leq a < 1$.

Theorem 2: The error exponent is positive for any SNR and $0 \leq a < 1$. Furthermore,

- (i) for i.i.d. observations ($a = 0$), the error exponent reduces to the Kullack-Leibler distance $D(p_0||p_1)$ where $p_0 \sim \mathcal{N}(0, \sigma^2)$ and $p_1 \sim \mathcal{N}(0, \Pi_0 + \sigma^2)$;
- (ii) for perfectly correlated signal ($a = 1$), the error exponent is zero for any SNR, and the miss probability is bounded by

$$\left(\frac{1}{\sqrt{2\pi}} - D\right) cn^{-1/2} \leq P_M \leq \frac{1}{\sqrt{2\pi}} cn^{-1/2} \quad (14)$$

for sufficiently large n , where c and $D \in (0, \frac{1}{\sqrt{2\pi}})$ are positive constants.

Proof: See [29]. ■

When $a = 0$ the theorem corresponds to Stein's lemma for the i.i.d. case. For the perfectly correlated case ($a = 1$), the miss detection does not decay exponentially; it decays with $\Theta(\frac{1}{\sqrt{n}})$.

Having obtained the behavior of the error exponent at two extreme correlation cases, we now investigate the error behavior for intermediate values of correlation, and show that the error exponent has distinct characteristics in different SNR regimes.

Theorem 3 (K vs. correlation): The error exponent as a function of correlation strength is characterized by the following:

¹The output of the whitening filter provides a sufficient statistic for the problem.

- (i) For $\text{SNR} > 1$, K decreases monotonically as the correlation strength increases (i.e. $a \uparrow$);
- (ii) For $\text{SNR} < 1$, there exists a non-zero value a^* of the correlation coefficient that achieves the maximal K , and a^* is given by the solution of the following equation.

$$[1 + a^2 + \Gamma(1 - a^2)]^2 - 2(r_e + \frac{a^4}{r_e}) = 0, \quad (15)$$

where $r_e = R_e/\sigma^2$. Furthermore, a^* converges to one as SNR decreases to zero.

Proof: See [29]. ■

We first note that Theorem 3 implies that an i.i.d. signal gives the best error performance for a given $\text{SNR} > 1$ with the maximal error exponent being $D(\mathcal{N}(0, 1)||\mathcal{N}(0, 1 + \Gamma))$. The intuition behind this result is that the innovations (new information about the signal process) provide more benefit to the detector than the noise averaging effect present for correlated observations since the signal component in the observation is strong at high SNR. Fig. 2 (left) shows the error exponent as a function of the correlation coefficient a . The monotonicity of the error exponent is clearly seen.

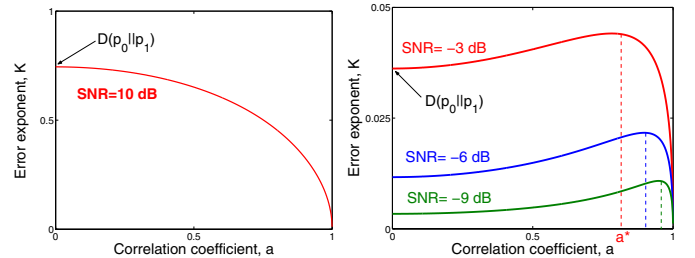


Fig. 2. K vs. a (left: $\text{SNR}=10$ dB, right: $\text{SNR}=-3,-6,-9$ dB)

In contrast, the error exponent does not decrease monotonically at $\text{SNR} < 1$, and there exists an optimal correlation as shown in Fig. 2 (right). It is seen that the i.i.d. case no longer gives the best error performance for a given SNR. The error exponent initially increases as a increases, and then decreases to zero as a approaches one. As the SNR further decreases (see the cases of -6 dB and -9dB) the error exponent decreases for a fixed correlation strength, and the value of a achieving the maximal error exponent is shifted closer to one. At low SNR the noise in the observations dominates. So, intuitively, making the signal more correlated provides the benefit of noise averaging. However, excessive correlation does not provide new information by observation, and the error exponent ultimately converges to zero as a approaches one, as predicted by Theorem 3 (ii). Notice that the ratio of the error exponent for the optimal correlation to that for the i.i.d. case becomes large as SNR decreases. Hence, the improvement due to optimal correlation can be large for low SNR cases. Fig. 3 shows

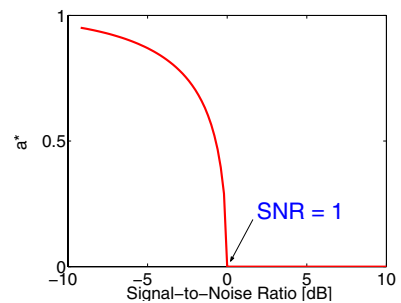


Fig. 3. Optimal correlation a^* vs SNR

the value of a that maximizes the error exponent as a function of SNR. As shown in the figure, unit SNR is a transition point between two different behavioral regimes of the error exponent with respect to correlation strength, and the transition is very sharp; the optimal correlation a^* approaches one rapidly when SNR becomes smaller than one.

Finally, we investigate the behavior of the error exponent with respect to SNR.

Theorem 4 (K vs. SNR): The error exponent K increases monotonically as SNR increases for a given correlation coefficient $0 \leq a < 1$. Moreover, at high SNR the error exponent K increases linearly with respect to $\frac{1}{2} \log(1 + \text{SNR}(1 - a^2))$.

Proof: See [29]. ■

The log SNR increase of K w.r.t. SNR is analogous to similar error-rate behavior arising in diversity combining of versions of a communications signal arriving over independent Rayleigh-faded paths in additive noise, since the signal component is stochastic in both cases. The log SNR behavior of the optimal Neyman-Pearson detector for random signals applies to general correlations as well. Comparing with the detection of a deterministic signal in noise, where the error exponent is proportional to SNR, the increase of error exponent w.r.t. SNR is much slower for the case of a stochastic signal in noise. Fig. 4 shows the error exponent with respect to SNR for a given correlation strength. The log SNR behavior is evident at high SNR.

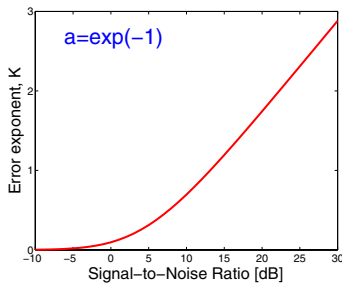


Fig. 4. K versus SNR ($a = e^{-1}$)

III. EXTENSION TO THE VECTOR CASE

In order to treat general cases in which the signal is a higher order AR process or the signal is determined by a linear combination of several underlying phenomena, we now consider a vector state-space model, and extend the results of the previous sections to this model. The hypotheses for the vector case are given by

$$\begin{aligned} H_0 &: y_i = w_i, & i = 1, 2, \dots, n, \\ H_1 &: y_i = \mathbf{h}^T \mathbf{s}_i + w_i, & i = 1, 2, \dots, n, \end{aligned} \quad (16)$$

where \mathbf{h} is a known vector and $\mathbf{s}_i \triangleq [s_{1i}, s_{2i}, \dots, s_{mi}]^T$ is the state of an m -dimensional process at time i following the state-space model

$$\begin{aligned} \mathbf{s}_{i+1} &= \mathbf{A} \mathbf{s}_i + \mathbf{B} \mathbf{u}_i, \\ \mathbf{s}_1 &\sim \mathcal{N}(\mathbf{0}, \mathbf{\Pi}_0), \\ \mathbf{u}_i &\stackrel{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{Q}), \quad \mathbf{Q} \geq 0. \end{aligned} \quad (17)$$

We assume that the feedback and input matrices, \mathbf{A} and \mathbf{B} , are known with the matrix \mathbf{A} being stable, and the process noise $\{\mathbf{u}_i\}$ independent of the measurement noise $\{w_i\}$. We also assume that

the initial state \mathbf{s}_1 is uncorrelated with \mathbf{u}_i for all i , and the initial covariance $\mathbf{\Pi}_0$ satisfies the following Lyapunov equation

$$\mathbf{\Pi}_0 = \mathbf{A} \mathbf{\Pi}_0 \mathbf{A}^T + \mathbf{B} \mathbf{Q} \mathbf{B}^T. \quad (18)$$

Thus, the signal sequence $\{\mathbf{s}_i\}$ forms a stationary vector process. In this case the SNR is defined similarly to (3) as $\frac{\mathbf{h}^T \mathbf{\Pi}_0 \mathbf{h}}{\sigma^2}$. The error exponent for the vector model is given by the following theorem.

Theorem 5 (Error exponent): For the Neyman-Pearson detector for the hypotheses (16, 17) with level $\alpha \in (0, 1)$ (i.e. $P_F \leq \alpha$) and a stable matrix \mathbf{A} , the error exponent of the miss probability is given by (6) independently of the value of α . The steady-state variances of the innovation process R_e and \tilde{R}_e calculated under H_1 and H_0 , respectively, are given by

$$R_e = \sigma^2 + \mathbf{h}^T \mathbf{P} \mathbf{h}, \quad (19)$$

where \mathbf{P} is the unique stabilizing solution of the discrete-time algebraic Riccati equation

$$\mathbf{P} = \mathbf{A} \mathbf{P} \mathbf{A}^T + \mathbf{B} \mathbf{Q} \mathbf{B}^T - \frac{\mathbf{A} \mathbf{P} \mathbf{h} \mathbf{h}^T \mathbf{P} \mathbf{A}^T}{\mathbf{h}^T \mathbf{P} \mathbf{h} + \sigma^2}, \quad (20)$$

and

$$\tilde{R}_e = \sigma^2 (1 + \mathbf{h}^T \tilde{\mathbf{P}} \mathbf{h}), \quad (21)$$

where $\tilde{\mathbf{P}}$ is the unique positive-semidefinite solution of the following Lyapunov equation

$$\tilde{\mathbf{P}} = (\mathbf{A} - \mathbf{K}_p \mathbf{h}^T) \tilde{\mathbf{P}} (\mathbf{A} - \mathbf{K}_p \mathbf{h}^T)^T + \mathbf{K}_p \mathbf{K}_p^T, \quad (22)$$

and $\mathbf{K}_p = \mathbf{A} \mathbf{P} \mathbf{h} R_e^{-1}$.

In spectral form, K is given by (10), where $S_y(\omega)$ is given by

$$S_y(\omega) = [\mathbf{h}^T (e^{j\omega} \mathbf{I} - \mathbf{A})^{-1} \mathbf{1}] \begin{bmatrix} \mathbf{Q} & 0 \\ 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} (e^{-j\omega} \mathbf{I} - \mathbf{A}^T)^{-1} \mathbf{h} \\ 1 \end{bmatrix}. \quad (23)$$

Proof: See [29]. ■

For this vector model, simple results describing the properties of the error exponent are not tractable since the relevant expressions depend on the multiple eigenvalues of the matrix \mathbf{A} . However, (6, 20, 22) provide closed-form expressions for the error exponent which can easily be explored numerically.

IV. SIMULATION RESULTS

We considered the Neyman-Pearson detection of a first-order autoregressive signal described by (2). We considered SNR values of 10 dB and -3 dB, and several values of a for each SNR. The probability of false alarm was set at 0.1% for all cases. Fig. 5 shows

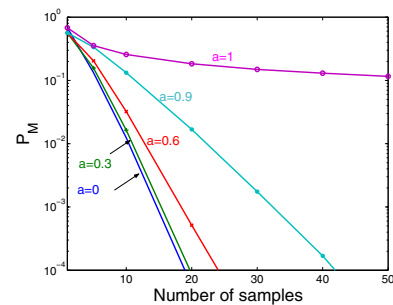


Fig. 5. P_M vs. number of samples (SNR=10dB)

the miss probability as a function of number of samples for 10 dB

SNR. It is shown that the i.i.d. case ($a = 0$) has the largest slope for error performance, and the slope of error decay is monotonically decreasing as a increases to one. Notice that the error performance for the same number of samples is significantly different for different correlation strengths even for the same SNR, and the performance for weak correlation is not much different from the i.i.d. case predicted by Fig. 2 (left). (We can see that the slope decreases suddenly near $a = 1$.) It is seen that the behavior of the miss probability for the highly correlated case ($a = 1$) deviates considerably from exponential decay. The error performance for SNR of -3 dB is shown in Fig. 6. It

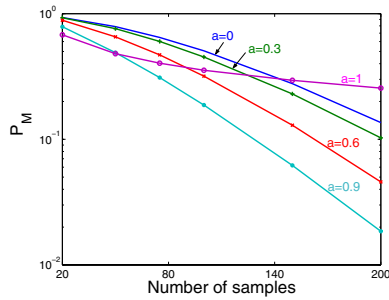


Fig. 6. P_M vs. number of samples (SNR=-3dB)

is seen that the slope increases as a increases from zero, and reaches a maximum with a sudden decrease after the maximum. Notice that the error curve is still not a straight line in the low SNR case due to the $o(n)$ term in the exponent. Since the error exponent increases only with $\frac{1}{2} \log \text{SNR}$, the required number of sensors for -3 dB SNR is much larger than for 10 dB SNR for the same miss probability. It is clearly seen that P_M is still larger than 10^{-2} for 200 samples whereas it is 10^{-4} with 20 samples for the 10 dB SNR case.

V. CONCLUSIONS

We have considered the detection of correlated signals in noisy observations. We have derived the error exponent for the Neyman-Pearson detector with a given level using the innovations and the spectral domain approaches. We have also provided the error exponent in closed form for a vector state-space model. We have further examined the properties of the error exponent. We have shown that the error exponent is a function of SNR and correlation strength, and the behavior of the error exponent w.r.t. correlation strength is sharply divided into two regions depending on SNR. For $\text{SNR} > 1$ the error exponent decreases monotonically as correlation becomes stronger, whereas there exists a nonzero correlation strength that gives the maximum slope when $\text{SNR} < 1$.

REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd Edition, Springer, New York, 1994.
- [2] C. W. Helstrom, *Elements of Signal Detection and Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1994.
- [3] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, Englewood Cliffs, NJ, 2000.
- [4] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.
- [5] T. Kailath and H. V. Poor, "Detection of stochastic processes," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2230-2259, Oct. 1998.
- [6] F. C. Schwappe, "Evaluation of likelihood functions for Gaussian signals," *IEEE Transactions on Information Theory*, vol. IT-1, pp. 61-70, 1965.
- [7] F. C. Schwappe, "State space evaluation of the Bhattacharyya distance between two Gaussian processes," *Information and Control*, vol. 11, no. 3, pp. 352-372, 1967.

- [8] T. Kailath, "The divergence and Bhattacharyya distance measures in signal detection," *IEEE Transactions on Communication Technology*, vol. COM-15, no. 1, pp. 52-60, Feb. 1967.
- [9] T. Kailath, "The innovations approach to detection and estimation theory," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 680-695, May 1970.
- [10] H. L. van Trees, "Applications of state-variable techniques in detection theory," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 653-669, May 1970.
- [11] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, 1958.
- [12] R. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. 18, no. 6, pp. 725-730, Nov 1972.
- [13] J. Gärtner, "On large deviations from an invariant measure," *Teor. Veroyatnost. i Primenen.*, vol. 22, no. 1, pp. 27-42, 1977.
- [14] R. S. Ellis, "Large deviations for a general class of random vectors," *Annals of Probability*, vol. 12, no. 1, pp. 1-12, Feb. 1984.
- [15] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, MA, 1993.
- [16] G. R. Benitz and J. A. Bucklew, "Large deviation rate calculations for nonlinear detectors in Gaussian noise," *IEEE Transactions on Information Theory*, vol. 36, no. 2, pp. 358-371, March 1990.
- [17] R. K. Bahr, "Asymptotic analysis of error probabilities for the nonzero-mean Gaussian hypothesis testing problem," *IEEE Transactions on Information Theory*, vol. 36, no. 3, pp. 597-607, March 1990.
- [18] R. K. Bahr and J. A. Bucklew, "Optimal sampling schemes for the Gaussian hypothesis testing problem," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1677-1686, Oct. 1990.
- [19] P. Barone, A. Gigli, and M. Piccioni, "Optimal importance sampling for some quadratic forms of ARMA processes," *IEEE Transactions Information Theory*, vol. 41, no. 6, pp. 1834 - 1844, 1995.
- [20] W. Bryc and W. Smolenski, "On the large deviation principle for a quadratic functional of the autoregressive process," *Statistics and Probability Letters*, vol. 17, pp. 281-285, 1993.
- [21] B. Bercu, F. Gamboa, and A. Rouault, "Large deviations for quadratic forms of stationary Gaussian processes," *Stochastic Processes and their Applications*, vol. 71, pp. 75-90, 1997.
- [22] J.-F. Chamberland and V. V. Veeravalli, "Design of sensor networks for detection applications via large-deviation theory," *Proceedings of the 2004 IEEE Information Theory Workshop*, San Antonio, TX, Oct. 2004.
- [23] R. R. Bahadur, *Some Limit Theorems in Statistics*, Conference Board of the Mathematical Sciences Regional Conferences Series in Applied Mathematics, no. 4. SIAM, Philadelphia, PA, 1971.
- [24] I. Vajda, *Theory of Statistical Inference and Information*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989.
- [25] I. Vajda, "Distances and discrimination rates of stochastic processes," *Stochastic Processes and their Applications*, vol. 35, pp. 47-57, 1990.
- [26] H. Luschgy, "Asymptotic behavior of Neyman-Pearson tests for autoregressive processes," *Scandinavian Journal of Statistics*, vol. 21, pp. 461-473, 1994.
- [27] H. Luschgy, A. L. Rukhin, and I. Vajda, "Adaptive tests for stochastic processes in the ergodic case," *Stochastic Processes and their Applications*, vol. 45, pp. 47-59, 1990.
- [28] Po-Ning Chen, "General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 316-323, Jan. 1996.
- [29] Y. Sung, L. Tong, and H. V. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponent and properties," preprint, Nov. 2004.