

Correspondence

Error Exponents for Neyman–Pearson Detection of Markov Chains in Noise

Alex S. Leong, *Student Member, IEEE*,
Subhrakanti Dey, *Senior Member, IEEE*, and
Jamie S. Evans, *Member, IEEE*

Abstract—A numerical method for computing the error exponent for Neyman–Pearson detection of two-state Markov chains in noise is presented, for both time-invariant and fading channels. We give numerical studies showing the behavior of the error exponent as the transition parameters of the Markov chain and the signal-to-noise ratio (SNR) are varied. Comparisons between the high-SNR asymptotics in Gaussian noise for the time-invariant and fading situations will also be made.

Index Terms—Error exponent, fading channel, hidden Markov model (HMM), Neyman–Pearson detection.

I. INTRODUCTION

The detection of random signals in noise is an important problem in engineering and signal processing. In general, performance analysis of detectors via closed-form expressions for the probability of error or related quantities is intractable except for very simple test statistics. An alternative is to study the rate of decrease of the probability of error as the number of samples increases, or in the Neyman–Pearson formulation, the probability of missed detection with a fixed false-alarm probability constraint.

Let P_M represent the probability of missed detection and P_{FA} the probability of false alarm. The error exponent for Neyman–Pearson detection is defined as the limit (log will refer to the natural logarithm unless stated otherwise)

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

and represents the rate at which P_M decays for a fixed constraint on P_{FA} as the number of samples $n \rightarrow \infty$. The error exponent K for the Neyman–Pearson detection problem can be shown (see [1] and the references therein) to be the following almost-sure limit (provided it exists) under H_0 ¹

$$K = \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{p_{0,n}}{p_{1,n}}(y^n) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log L_n(y^n) \quad (2)$$

where $p_{0,n}$ and $p_{1,n}$ are the null and alternative joint densities, respectively, of the measurements $y^n \equiv (y_1, \dots, y_n)$, and $L_n(y^n) =$

Manuscript received July 17, 2006; revised January 29, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Fredrik Gustafsson. This work was supported by the Australian Research Council. Parts of this work were presented at the Information, Decision and Control Symposium, Adelaide, Australia, February 2007.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN, an affiliated program of National ICT Australia), Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Vic. 3010, Australia (e-mail: a.leong@ee.unimelb.edu.au; s.dey@ee.unimelb.edu.au; j.evans@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/TSP.2007.897863

¹In the independent and identically distributed (i.i.d.) case, this is Stein's lemma, and (2) is equivalent to computing the Kullback–Leibler distance between the two densities.

$(p_{1,n}/p_{0,n})(y^n)$ is the likelihood ratio. Unfortunately, closed-form expressions for the error exponent are often still not possible except in rare cases. The subject of this correspondence is the numerical computation of the error exponent for Neyman–Pearson detection of a two-state Markov chain in noise. Such a detection problem can arise in the context of hidden Markov models (HMMs), which have applications in many areas, such as radar tracking and biology [2].

Some previous work, e.g., [3]–[6], have characterized the large sample behavior of detectors using results from large deviations theory such as the Gärtner–Ellis theorem to determine the rate functions, though evaluating these numerically usually required further optimization techniques. Moreover, often the minimum probability of error criterion or detection with fixed thresholds is used instead of the Neyman–Pearson criterion, so that $\lim_{n \rightarrow \infty} -(1/n) \log P_e$ is analyzed instead of (1). Neyman–Pearson detection of Gauss–Markov processes is, however, considered in [1]. By studying (2) and using properties of the innovations process and the state-space structure, the authors managed to obtain a closed-form expression for the error exponent and derived relationships between the error exponent and parameters of the system.

On the other hand, in detection problems involving Markov chains, the problem of trying to decide between two Markov chains with different transition parameters (but no noise) is treated in standard textbooks such as [7]. Likelihood ratios for detecting Markov chains in noise are derived in [8]. Classification of HMMs with empirically observed statistics is studied in [9]. Error exponents in HMM order estimation are considered in, e.g., [10]. To the best of our knowledge, however, the problem of determining the error exponent associated with detecting a Markov chain in noise, and how this error exponent behaves as system parameters vary, has not been previously treated in the literature.

Moreover, apart from [6] the impact of fading has not been addressed in these previous works, though likelihood ratios for various decentralized detection schemes over fading channels have been derived recently (see, e.g., [11] and [12]). Intuitively fading reduces detection performance, and it is of interest to know what the error exponents are in fading environments, for example in wireless sensor networks where there are limited resources and delay constraints such that waiting to collect too many data samples for a hypothesis test is not desirable.

In this correspondence, we will adopt a similar philosophy to [1] in studying the error exponent via (2) and properties of the likelihood ratio. While a closed-form expression cannot be obtained, numerical methods for computing the error exponent when the Markov chain has two states will be presented; this situation can occur when binary data (e.g., the presence or absence of a target) is being detected over a communication channel with noise. We first treat the case where the channel is time invariant and known in Section III, and then extend our methods to channels that are randomly time varying (e.g., a wireless link) in Section IV, for a Rayleigh fading channel with no instantaneous channel state information. We will numerically study the behavior of the error exponent as the transition parameters/probabilities of the Markov chain are varied, and also as the noise variance is varied. In particular, for Gaussian noise, we shall see that at high SNR, there is a marked contrast between the time-invariant and fading scenarios, with the scaling of the error exponent with SNR being linear and logarithmic, respectively.

II. SYSTEM MODEL

Let the null and alternative hypotheses be

$$\begin{aligned} H_0 : y_k &= w_k \\ H_1 : y_k &= h_k s_k + w_k \end{aligned} \quad (3)$$

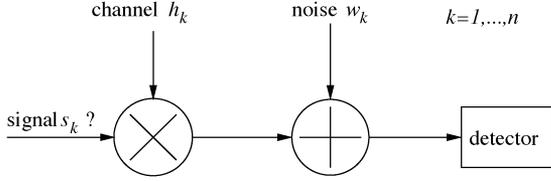


Fig. 1. System model.

for $k = 1, 2, \dots, n$ (see Fig. 1). The noise process $\{w_k\}$ is assumed to be independent and identically distributed (i.i.d.) and $N(0, \sigma^2)$.² The process $\{s_k\}$ that we wish to detect is a homogeneous two-state Markov chain, with state space $\{\psi_1, \psi_2\}$, and $\psi_2 = -\psi_1$. We use the convention that $a_{ij} = P(s_{k+1} = \psi_j | s_k = \psi_i)$ for the transition probabilities, with the assumption that $0 < a_{ij} < 1, \forall i, j$. In Section III, the process $\{h_k\}$ will be taken to be constant for all k , while in Section IV $\{h_k\}$ will be a fading process.

III. ERROR EXPONENT COMPUTATION FOR TIME-INVARIANT CHANNELS

We first consider time-invariant channels, where without loss of generality we take $h_k = 1, \forall k$ in (3). We will calculate the error exponent using the form (2), where $L_n(y^n) = (p_{1,n}/p_{0,n})(y^n)$ is the likelihood ratio. The likelihood ratio for the problem of detecting Markov chains in noise (3) has been previously derived in, e.g., [8]. One can write $L_n(y^n) = \prod_{k=1}^n l(y_k | y^{k-1})$, where

$$\begin{aligned} l(y_k | y^{k-1}) &= \frac{p_1(y_k | s_k = \psi_j, y^{k-1}) P(s_k = \psi_j | y^{k-1})}{p_0(y_k | y^{k-1}) P(s_k = \psi_j | y^k)} \\ &= \exp\left(\frac{\psi_j y_k}{\sigma^2} - \frac{1}{2} \frac{\psi_j^2}{\sigma^2}\right) \frac{P(s_k = \psi_j | y^{k-1})}{P(s_k = \psi_j | y^k)} \end{aligned}$$

for any $j \in \{1, 2\}$. Here, we choose $j = 1$. We have the recursive relations

$$P(s_k = \psi_j | y^{k-1}) = \sum_{i=1}^2 a_{ij} P(s_{k-1} = \psi_i | y^{k-1}) \quad (4)$$

$$P(s_k = \psi_j | y^k) = \frac{p_1(y_k | s_k = \psi_j) P(s_k = \psi_j | y^{k-1})}{\sum_{i=1}^2 p_1(y_k | s_k = \psi_i) P(s_k = \psi_i | y^{k-1})} \quad (5)$$

where $p_1(y_k | s_k = \psi_j)$ is the density of y_k under the alternative hypothesis, i.e.,

$$p_1(y_k | s_k = \psi_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_k - \psi_j)^2}{2\sigma^2}\right).$$

²The method presented here does not necessarily require the noise to be Gaussian [13], though some of the expressions will change for different noise distributions.

Thus, the error exponent can be written as

$$\begin{aligned} K &= \lim_{n \rightarrow \infty} -\frac{1}{n} \log \prod_{k=1}^n l(y_k | y^{k-1}) \\ &= \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{k=1}^n \log l(y_k | y^{k-1}) \\ &= \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{k=1}^n \left[\frac{\psi_1 y_k}{\sigma^2} - \frac{1}{2} \frac{\psi_1^2}{\sigma^2} + \log P(s_k = \psi_1 | y^{k-1}) \right. \\ &\quad \left. - \log P(s_k = \psi_1 | y^k) \right] \end{aligned}$$

provided that the limit exists. Now under H_0 , $\{y_k\}$ is i.i.d. zero-mean Gaussian. Thus

$$-\frac{1}{n} \sum_{k=1}^n \frac{\psi_1 y_k}{\sigma^2} \rightarrow 0 \text{ a.s.}$$

by the strong law of large numbers. Since $0 < a_{ij} < 1, \forall i, j$ by assumption, the transition matrix for $\{s_k\}$ will be primitive. From [14], it then follows that $P(s_k = \psi_1 | y^{k-1})$ and $P(s_k = \psi_1 | y^k)$ have invariant limiting distributions under H_0 as $k \rightarrow \infty$, by taking h_\bullet as the zero mapping in [14, Example 1.1] and considering our situation as a case of a misspecified hidden Markov model (HMM). Hence, by the arguments above and the ergodic theorem:

$$K = \frac{1}{2} \frac{\psi_1^2}{\sigma^2} - \mathbb{E} \left[\log(s_k = \psi_1 | y^{k-1}) - \log P(s_k = \psi_1 | y^k) \right] \text{ a.s.} \quad (6)$$

We also note that an alternative way of showing the existence of the almost-sure limit is by using [15, Proposition 3.2].

To calculate the above, it suffices to find the limiting distribution of $P(s_k = \psi_1 | y^k)$, since $P(s_k = \psi_1 | y^{k-1})$ is related by (4). We apply a method from [13], which was originally used for the problem of computing the probability of error in HMM filtering for two-state Markov chains.³ Let $q_k \equiv P(s_k = \psi_1 | y^k)$ and $f_k(q) dq \equiv P(q_k \in (q, q + dq))$. Then

$$\begin{aligned} P(q_k \in (q, q + dq), q_{k-1} \in (\tilde{q}, \tilde{q} + d\tilde{q})) \\ = P(q_k \in (q, q + dq) | q_{k-1} = \tilde{q}) f_{k-1}(\tilde{q}) d\tilde{q}. \end{aligned} \quad (7)$$

Using (5) and recalling that $\psi_2 = -\psi_1$, we have the relation shown in (8) at the bottom of the page. Thus

$$y_k = -\frac{\sigma^2}{2\psi_1} \ln \frac{[a_{11}q_{k-1} + a_{21}(1 - q_{k-1})](1 - q_k)}{[a_{12}q_{k-1} + a_{22}(1 - q_{k-1})]q_k} \equiv g(q_k, q_{k-1})$$

and

$$\left| \frac{dy_k}{dq_k} \right| = \left| \frac{\sigma^2}{2\psi_1 q_k (1 - q_k)} \right| = \frac{\sigma^2}{2|\psi_1| q_k (1 - q_k)}.$$

³This method unfortunately does not appear to be extendable to Markov chains with more than two states. The multi-state version of the problem considered in [13] is difficult; the best result so far appears to be [16], which requires certain transition parameters to be asymptotically small.

$$q_k = \frac{a_{11}q_{k-1} + a_{21}(1 - q_{k-1})}{a_{11}q_{k-1} + a_{21}(1 - q_{k-1}) + \exp\left(\frac{-2\psi_1 y_k}{\sigma^2}\right) [a_{12}q_{k-1} + a_{22}(1 - q_{k-1})]}. \quad (8)$$

Then, by a change of variable, again noting that y_k is $N(0, \sigma^2)$ under H_0 , we obtain

$$\begin{aligned} P(q_k \in (q, q + dq) | q_{k-1} = \bar{q}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{g(q, \bar{q})^2}{2\sigma^2}\right) \\ &\quad \times \frac{\sigma^2}{2|\psi_1|q(1-q)} dq \\ &\equiv S(q, \bar{q}) dq. \end{aligned}$$

Hence, from (7), we can see that in the limit as $k \rightarrow \infty$, $f_k(q)$ satisfies the following integral equation:

$$f(q) dq = \int_0^1 S(q, \bar{q}) f(\bar{q}) d\bar{q} dq. \quad (9)$$

This integral equation may be solved numerically using a procedure described in [17] (with slight modifications). Consider the eigenvalue problem $\lambda f(q) = \int_0^1 S(q, \bar{q}) f(\bar{q}) d\bar{q}$, which corresponds to (9) when $\lambda = 1$. We can solve this eigenvalue problem using the Nyström method [18]. Replacing the integral by an N -point quadrature rule, and defining

$$\mathbf{K} \equiv \begin{bmatrix} w_1 S(t_1, t_1) & \dots & w_N S(t_1, t_N) \\ \vdots & \ddots & \vdots \\ w_1 S(t_N, t_1) & \dots & w_N S(t_N, t_N) \end{bmatrix}$$

we obtain

$$\mathbf{K} [f(t_1) \dots f(t_N)]^T = \lambda [f(t_1) \dots f(t_N)]^T$$

where w_j represent the weights and t_j the quadrature points of the quadrature rule. The Gauss–Legendre quadrature rule is commonly used for the Nyström method [18, p. 110]. In this correspondence, we will use the Gauss–Legendre rule in a composite scheme [19, p. 113]; this allows a simple way to vary the precision by changing the number of subintervals. In the following sections, we use 200 subintervals, each subinterval using a fourth-order Gauss–Legendre rule. Then, in order to obtain an approximation for $f(q)$, we take the eigenvector that corresponds to the largest real eigenvalue of \mathbf{K} , and normalize so that $\int_0^1 f(q) dq = 1$ is satisfied. We may then calculate $\mathbb{E}[\log P(s_k = \psi_1 | y^k)]$, and hence the error exponent via (6). Some numerical properties of this method are described in [17].

A. Numerical Studies

First, we study how the error exponent varies when one of the transition parameters a_{12} is varied, while all other parameters are fixed.⁴ In Fig. 2, we show plots for four different fixed values of a_{21} , and a_{12} is varied. We fix $\psi_1 = -1$, $\psi_2 = 1$, $\sigma^2 = 1$. For a_{12} small, the error exponent K is large. As a_{12} increases, K decreases until a certain point at which it starts to increase again. Quantifying exactly where the point with minimum K occurs, however, seems rather difficult. An intuitive explanation as to why K is large when a_{12} is close to zero or one is that in either of these situations the signal will have a certain amount of “correlation,” which makes the signal easier to detect. This is because if a_{12} is close to zero then when it is in state 1, it tends to stay in the same state, whereas if a_{12} is close to one then when it is in state 1, at the next time instant it is very likely to switch to state 2.

Next, we look at how the error exponent behaves at different noise levels for fixed Markov chain transition parameters. In Fig. 3, we plot the error exponent as the noise variance σ^2 varies, for four different sets of parameter values. We again set $\psi_1 = -1$, $\psi_2 = 1$. Observe that the error exponent appears to scale linearly with the “signal-to-noise ratio” ψ_1^2/σ^2 . At high SNR, the error exponents also become very close to

⁴By the symmetry of the situation, we will get the same behavior when a_{21} is varied and the other parameters are fixed.

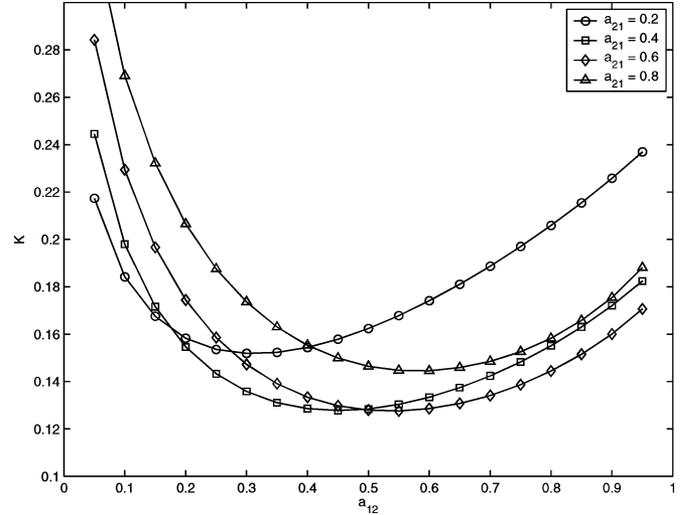


Fig. 2. Error exponent K for various a_{12} , with a_{21} fixed.

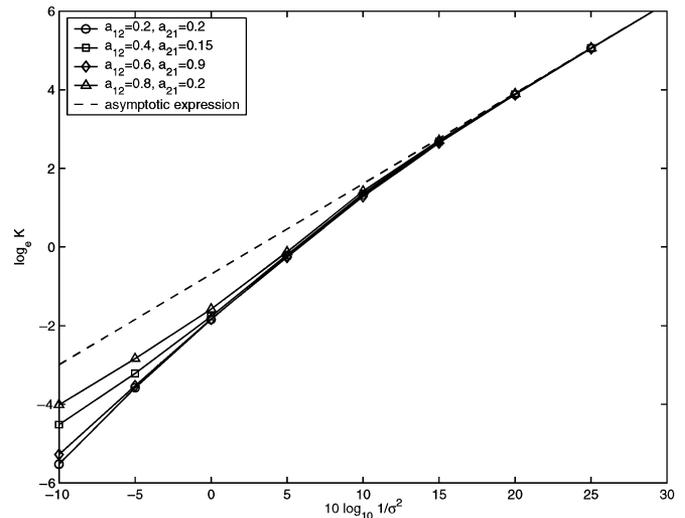


Fig. 3. Error exponent K for various σ^2 .

each other in a relative sense, for all four sets of parameter values used. We have the following result, which is proven in Appendix A.

Lemma 1: With Gaussian noise and as $\sigma^2 \rightarrow 0$

$$K \sim \frac{\psi_1^2}{2\sigma^2} \quad (10)$$

where $f \sim g$ means that $\lim_{\sigma \rightarrow 0} f(\sigma)/g(\sigma) = 1$ (see, e.g., [20, pp. 4–6]).

The dashed line in Fig. 3 is the asymptotic expression (10). By Lemma 1, the scaling at high SNR is indeed linear. Moreover, the asymptotic expression for K at high SNR does not depend on the parameters a_{12} and a_{21} , which completely specifies the two-state Markov chain. It is known that when detecting a constant signal in Gaussian noise, the error exponent scales linearly with SNR. In fact, it can be easily shown that for a constant signal of amplitude ψ_1 , the error exponent is exactly $\psi_1^2/2\sigma^2$. So from Lemma 1, the effect of Markov chain state transitions on the performance of detectors become less important at high SNR. Another potential use of Lemma 1 is to show that the high-SNR performance of a detector does not degrade too much even when the transition parameters may not be known exactly but only estimates are available, which will be a subject for further investigation.

IV. ERROR EXPONENT COMPUTATION WITH FADING

We now consider a randomly time-varying Rayleigh⁵ fading channel in the detection problem (3). Without loss of generality, we will let the process $\{h_k\}$ have unit power, with density $p(h_k) = 2h_k \exp(-h_k^2)$, $h_k \geq 0$ as in [12]. We assume that h_k and h_l are independent for $k \neq l$. This is a reasonable assumption if the sampling time is greater than the channel coherence time, as might occur in low data rate systems such as wireless sensor networks (see also the block fading models commonly used in the communications theory literature [21]). We assume that the receiver knows the distribution of the fading process, but the instantaneous values of h_k are unknown, i.e., has no instantaneous channel state information (CSI). This is because acquiring channel knowledge might be too expensive when resources are limited, and also because the assumption of full CSI may not be appropriate in the context of a signal detection problem, depending on the nature of the source of the signal.

An inspection of the derivations in [8] shows that one can still write the likelihood ratio for problem (3) in the form $L_n(y^n) = \prod_{k=1}^n l(y_k|y^{k-1})$, where

$$l(y_k|y^{k-1}) = \frac{p_1(y_k|s_k = \psi_j, y^{k-1}) P(s_k = \psi_j|y^{k-1})}{p_0(y_k|y^{k-1}) P(s_k = \psi_j|y^k)}$$

for any $j \in \{1, 2\}$ (the explicit expressions for the terms will now be different though). We have $p_1(y_k|s_k = \psi_j, y^{k-1}) = p_1(y_k|s_k = \psi_j)$ if the fades are independent. The recursive relations (4) and (5) will continue to hold. One also has, in a slight generalization of Lemma 1 in [12]

$$p_1(y_k|s_k = \psi_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{2\sigma^2}{\psi_j^2 + 2\sigma^2} e^{-y_k^2/2\sigma^2} \times \left[1 + \sqrt{2\pi} a_j y_k e^{(a_j y_k)^2/2} Q(-a_j y_k) \right] \quad (11)$$

where

$$a_j \equiv \psi_j / (\sigma \sqrt{\psi_j^2 + 2\sigma^2})$$

and

$$Q(x) = \int_x^\infty (1/\sqrt{2\pi}) \exp(-t^2/2) dt$$

is the Q function. For brevity, call $a \equiv a_1 = \psi_1 / (\sigma \sqrt{\psi_1^2 + 2\sigma^2})$. The error exponent can be written as

$$K = \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{k=1}^n \left[\log \frac{p_1(y_k|s_k = \psi_1)}{p_0(y_k|y^{k-1})} + \log P(s_k = \psi_1|y^{k-1}) - \log P(s_k = \psi_1|y^k) \right].$$

provided the limit exists. Note that under H_0 , $\{y_k\}$ is still i.i.d. zero-mean Gaussian. Then, by similar arguments as in Section III and again using results from [14], we can show that the almost-sure limit does exist and is equal to

$$-\int_{-\infty}^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k) \right] \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y_k^2/2\sigma^2} dy_k - \log \frac{2\sigma^2}{\psi_1^2 + 2\sigma^2} - \mathbb{E} \left[\log \frac{P(s_k = \psi_1|y^{k-1})}{P(s_k = \psi_1|y^k)} \right] \quad (12)$$

⁵Our method should also work for other fading distributions, provided that after "averaging" over h_k [12], $p_1(y_k|s_k = \psi_j)$ still has a tractable closed form such as (11).

where the first term (the integral) can be evaluated numerically. For calculation of the other terms, again let $q_k = P(s_k = \psi_1|y^k)$ and $f_k(q) dq = P(q_k \in (q, q + dq))$. The relation (7) still holds, but we now have from (5) after some rearranging that

$$\frac{1 - \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(a y_k)}{1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k)} = \frac{a_{11} q_{k-1} + a_{21}(1 - q_{k-1})}{a_{12} q_{k-1} + a_{22}(1 - q_{k-1})} \times \frac{1 - q_k}{q_k}. \quad (13)$$

Given q_k and q_{k-1} , (13) can be solved for y_k numerically. The solution obtained can be shown to be unique by the following argument. Since the right-hand side of (13) is positive, for $a > 0$,

$$\frac{1 - \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(a y_k)}{1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k)} \rightarrow \begin{cases} \infty, & y \rightarrow -\infty \\ 0, & y \rightarrow +\infty. \end{cases}$$

Moreover, we obtain after some calculations

$$\frac{d}{dy_k} \left(\frac{1 - \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(a y_k)}{1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k)} \right) = \frac{-\sqrt{2\pi} a e^{(a y_k)^2/2}}{[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k)]^2} < 0$$

so that the left-hand side of (13) is strictly decreasing in y_k , and hence a unique solution will be obtained. A similar argument holds if $a < 0$.

Next, using implicit differentiation on (13) and rearranging, we obtain

$$\left| \frac{dy_k}{dq_k} \right| = \frac{[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2/2} Q(-a y_k)]^2}{\sqrt{2\pi} |a| e^{(a y_k)^2/2} q_k^2} \times \frac{a_{11} q_{k-1} + a_{21}(1 - q_{k-1})}{a_{12} q_{k-1} + a_{22}(1 - q_{k-1})}. \quad (14)$$

By a change of variable

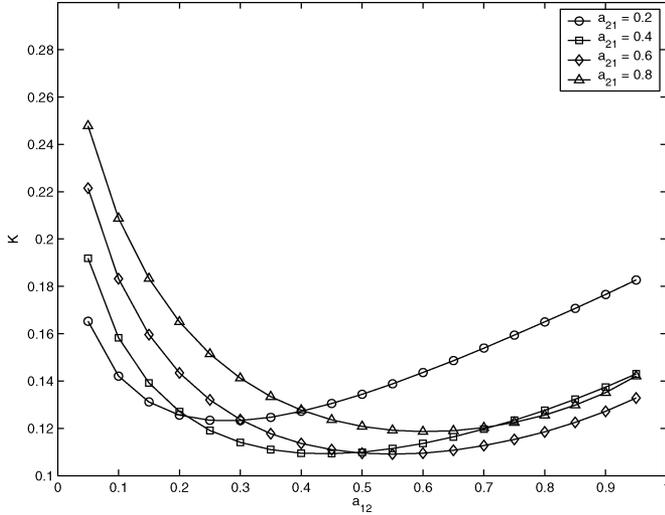
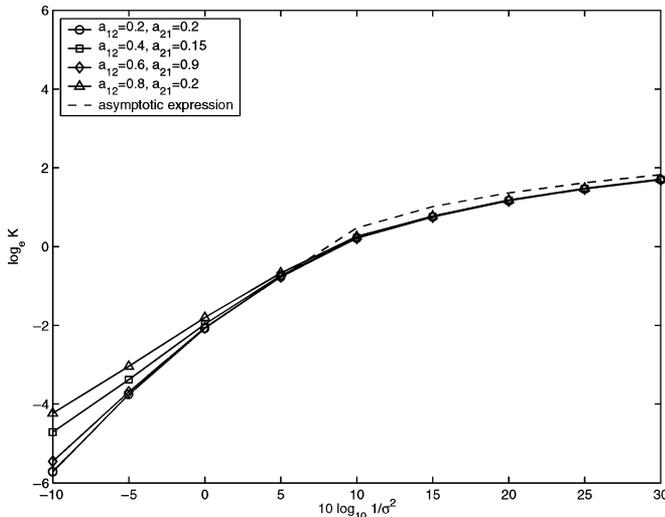
$$S(q, \tilde{q}) dq \equiv P(q_k \in (q, q + dq) | q_{k-1} = \tilde{q}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y_k^2/2\sigma^2} \left| \frac{dy_k}{dq_k} \right|_{q_k=q, q_{k-1}=\tilde{q}} dq$$

where y_k is the numerical solution to (13) and dy_k/dq_k is given by (14). We will then need to solve the integral equation $f(q) dq = \int_0^1 S(q, \tilde{q}) f(\tilde{q}) d\tilde{q} dq$ using the same numerical procedure as in Section III.

A. Numerical Studies

In Fig. 4, we show plots for four different fixed values of a_{21} , and a_{12} is varied. We fix $\psi_1 = -1$, $\psi_2 = 1$, $\sigma^2 = 1$. Similar behavior to the case without fading in Fig. 2 can be observed, though the error exponents are smaller due to the presence of fading.

In Fig. 5, we plot the error exponent with fading as the noise variance σ^2 varies, for four different sets of parameter values. We again set $\psi_1 = -1$, $\psi_2 = 1$. The error exponent now appears to scale logarithmically with the SNR. As in Fig. 3, at high SNR, the error exponents approach each other (in a relative sense) for all four sets of parameter values used. We have the following.


 Fig. 4. Error exponent \bar{K} with Rayleigh fading for various a_{12} , with a_{21} fixed.

 Fig. 5. Error exponent \bar{K} with Rayleigh fading for various σ^2 .

Lemma 2: With Gaussian noise and Rayleigh fading, as $\sigma^2 \rightarrow 0$

$$K \sim \log \frac{\psi_1^2}{2\sigma^2}. \quad (15)$$

The proof is in Appendix B. The dashed line in Fig. 5 is a plot of the expression (15). Note that we do not plot for smaller values of $1/\sigma^2$ as then the asymptotic expression takes on negative values. We see that the asymptotic expression is logarithmic, where again there is no dependence on the parameters a_{12} and a_{21} . The error exponent for detecting a constant signal of amplitude ψ_1 with the Rayleigh fading model used here can be shown to be asymptotic to $\log(\psi_1^2/2\sigma^2)$, so that as in the time-invariant case, state transitions of the Markov chain have a less significant effect on performance at high SNR.

The difference in behavior in the scaling of the error exponent with SNR in Figs. 3 and 5 can possibly be attributed to the discrete nature of the states in the signal $\{s_k\}$ in the case without fading, whereas with fading the received signal $\{h_k s_k\}$ (even without noise) can take on a continuous range. Other situations where logarithmic scaling in the error exponent has also been observed include (see, e.g., [1]) the detection of Gaussian signals in Gaussian noise, and the detection of Gauss–Markov systems in noise, where in both cases the signals to be detected have a continuous range.

V. CONCLUSION

We have presented numerical methods to compute the error exponent associated with Neyman–Pearson detection of a two-state Markov chain in noise, both with and without fading. Numerical studies relating the error exponent to the parameters of the system have been presented. In particular, with Gaussian noise, at high SNR, the error exponents scales linearly without fading but logarithmically with fading.

While the results presented in this correspondence have dealt with Gaussian noise, the methods to compute the error exponent will in principle still work with other noise distributions. However, in general, the high-SNR behavior of the error exponent in non-Gaussian noise is likely to be strongly dependent on the nature of the distributions. For instance, consider the generalized Gaussian distribution [22] with density

$$f(x) = \frac{c\eta(\sigma, c)}{2\Gamma(1/c)} \exp\{-[\eta(\sigma, c)|x|^c]\}$$

where $\eta(\sigma, c) \equiv \sigma^{-1}[\Gamma(3/c)/\Gamma(1/c)]^{1/2}$, σ^2 is the noise variance, and $c > 0$ is a parameter that controls the rate of decay of the tails. For $c = 2$, this reduces to the Gaussian density, and for $c = 1$ the Laplacian density. Using this density for the noise, preliminary results for the time-invariant case suggest a $O(1/\sigma^c)$ scaling of the error exponent at high SNR, so that for larger c (faster decaying tail probabilities), the rate of increase of the error exponent is greater than for distributions with heavier tails. Future work will include a more thorough investigation of the generalized Gaussian distribution and other types of noise and fading processes.

APPENDIX A PROOF OF LEMMA 1

Note that $\psi_1^2/2\sigma^2$ is just the first term on the right-hand side of (6), so it suffices to show that

$$\mathbb{E} \left[\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} \right] = o\left(\frac{1}{\sigma^2}\right)$$

for small σ . Denote $p_o(y_k) \equiv (1/\sqrt{2\pi\sigma^2}) \exp(-y_k^2/2\sigma^2)$. Suppose $P(s_k = \psi_1 | y^{k-1}) = p$. Then, from (5)

$$P(s_k = \psi_1 | y^k) = \frac{\exp\left(\frac{-(y_k - \psi_1)^2}{2\sigma^2}\right)p}{\exp\left(\frac{-(y_k - \psi_1)^2}{2\sigma^2}\right)p + \exp\left(\frac{-(y_k + \psi_1)^2}{2\sigma^2}\right)(1-p)}$$

and so

$$\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} = \log \left(p + (1-p) \exp\left(-\frac{2\psi_1 y_k}{\sigma^2}\right) \right).$$

Suppose $\psi_1 > 0$ (the calculations for $\psi_1 < 0$ are similar). For $y_k > 0$, one can easily show that

$$\log p \leq \log \left(p + (1-p) \exp\left(-\frac{2\psi_1 y_k}{\sigma^2}\right) \right) \leq 0$$

and so $\int_0^\infty \log(p + (1-p) \exp(-2\psi_1 y_k/\sigma^2)) p_o(y_k) dy_k = O(1)$.

For $y_k < 0$

$$\begin{aligned}
& \int_{-\infty}^0 \log \left(p + (1-p) \exp \left(-\frac{2\psi_1 y_k}{\sigma^2} \right) \right) p_0(y_k) dy_k \\
&= \int_0^{\infty} \log \left(p + (1-p) \exp \left(\frac{2\psi_1 y_k}{\sigma^2} \right) \right) p_0(y_k) dy_k \\
&= \int_0^{\infty} \log \left[e^{2\psi_1 y_k / \sigma^2} \left(p \exp \left(-\frac{2\psi_1 y_k}{\sigma^2} \right) + (1-p) \right) \right] \\
&\quad \times p_0(y_k) dy_k \\
&= \int_0^{\infty} \left(\frac{2\psi_1 y_k}{\sigma^2} + \log \left(p \exp \left(-\frac{2\psi_1 y_k}{\sigma^2} \right) + (1-p) \right) \right) \\
&\quad \times p_0(y_k) dy_k \\
&= \sqrt{\frac{2}{\pi}} \frac{\psi_1}{\sigma} + \int_0^{\infty} \log \left(p \exp \left(-\frac{2\psi_1 y_k}{\sigma^2} \right) + 1-p \right) p_0(y_k) dy_k
\end{aligned}$$

and the term $\int_0^{\infty} \log(p \exp(-2\psi_1 y_k / \sigma^2) + 1-p) p_0(y_k) dy_k$ is $O(1)$ by a similar argument to the case when $y_k > 0$. So

$$\mathbb{E} \left[\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} \middle| p \right] = O(1) + O\left(\frac{1}{\sigma}\right) + O(1) = o\left(\frac{1}{\sigma^2}\right)$$

for small σ .

Hence

$$\mathbb{E} \left[\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} \right] = \mathbb{E} \left[\mathbb{E} \left[\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} \middle| p \right] \right]$$

is also $o(1/\sigma^2)$.

APPENDIX B PROOF OF LEMMA 2

Referring to (12), for small σ we have

$$-\log \frac{2\sigma^2}{\psi_1^2 + 2\sigma^2} \sim \log \frac{\psi_1^2}{2\sigma^2}.$$

Denote $p_o(y_k) \equiv (1/\sqrt{2\pi\sigma^2}) \exp(-y_k^2/2\sigma^2)$. We first show that

$$\int_{-\infty}^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right] p_0(y_k) dy_k = o\left(\log \frac{1}{\sigma^2}\right).$$

Assume $\psi_1 > 0$ (again the calculations for $\psi_1 < 0$ are similar), which implies a is positive. We can write

$$\begin{aligned}
& \int_{-\infty}^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right] p_0(y_k) dy_k \\
&= \int_0^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right] p_0(y_k) dy_k \\
&\quad + \int_0^{\infty} \log \left[1 - \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(a y_k) \right] p_0(y_k) dy_k.
\end{aligned}$$

First, we have

$$\begin{aligned}
& \int_0^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right] p_0(y_k) dy_k \\
&\leq \int_0^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} \right] p_0(y_k) dy_k \\
&= \int_0^{\infty} \log \left[e^{(a y_k)^2 / 2} \left(e^{-(a y_k)^2 / 2} + \sqrt{2\pi} a y_k \right) \right] p_0(y_k) dy_k \\
&\leq \int_0^{\infty} \left(\frac{(a y_k)^2}{2} + \log(1 + \sqrt{2\pi} a y_k) \right) p_0(y_k) dy_k \\
&\leq \int_0^{\infty} \left(\frac{(a y_k)^2}{2} + \sqrt{2\pi} a y_k \right) p_0(y_k) dy_k \\
&= \frac{a^2 \sigma^2}{4} + a \sigma = O(1)
\end{aligned}$$

since $a \equiv a_1 = \psi_1 / (\sigma \sqrt{\psi_1^2 + 2\sigma^2})$. Next, note that

$$\begin{aligned}
& \left| \log \left[1 - \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(a y_k) \right] \right| \\
&\leq \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right]
\end{aligned}$$

using the result previously shown in Section IV that

$$(1 - \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(a y_k)) / (1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k))$$

is decreasing with y_k , and is equal to 1 when $y_k = 0$. So $\int_0^{\infty} \log[1 - \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(a y_k)] p_0(y_k) dy_k$ is also $O(1)$. Hence

$$\begin{aligned}
& \int_{-\infty}^{\infty} \log \left[1 + \sqrt{2\pi} a y_k e^{(a y_k)^2 / 2} Q(-a y_k) \right] p_0(y_k) dy_k \\
&= O(1) + O(1) \\
&= o\left(\log \frac{1}{\sigma^2}\right).
\end{aligned}$$

Using the bounds just derived, together with the same arguments as for the proof of Lemma 1, we can show that the remaining term in (12)

$$\mathbb{E} \left[\log \frac{P(s_k = \psi_1 | y^{k-1})}{P(s_k = \psi_1 | y^k)} \right] = O(1) = o\left(\log \frac{1}{\sigma^2}\right).$$

REFERENCES

- [1] Y. Sung, L. Tong, and H. V. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponent and properties," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1354-1365, Apr. 2006.
- [2] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518-1569, Jun. 2002.
- [3] G. R. Benitz and J. A. Bucklew, "Large deviation rate calculations for nonlinear detectors in Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 358-371, Mar. 1990.
- [4] R. K. Bahr and J. A. Bucklew, "Optimal sampling schemes for the Gaussian hypothesis testing problem," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1677-1686, Oct. 1990.
- [5] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407-416, Feb. 2003.

- [6] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1081–1092, Mar. 2007.
- [7] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer Verlag, 1998.
- [8] L. L. Scharf and L. W. Nolte, "Likelihood ratios for sequential hypothesis testing on Markov sequences," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 101–109, Jan. 1977.
- [9] N. Merhav, "Universal classification for hidden Markov models," *IEEE Trans. Inf. Theory*, vol. 37, no. 6, pp. 1586–1594, Nov. 1991.
- [10] E. Gassiat and S. Boucheron, "Optimal error exponents in hidden Markov models order estimation," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 964–980, Apr. 2003.
- [11] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.
- [12] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.
- [13] L. Shue, S. Dey, B. D. O. Anderson, and F. De Bruyne, "On state-estimation of a two-state hidden Markov model with quantization," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 202–208, Jan. 2001.
- [14] F. Le Gland and L. Mevel, "Exponential forgetting and geometric ergodicity in hidden Markov models," *Math. Control, Signals, Syst.*, vol. 13, pp. 63–93, 2000.
- [15] F. Le Gland and L. Mevel, "Asymptotic properties of the MLE in hidden Markov models," in *Proc. 4th Eur Control Conf.*, Bruxelles, Jul. 1997, Paper FrA-F6.
- [16] R. Khasminskii and O. Zeitouni, "Asymptotic filtering for finite state Markov chains," *Stoch. Process. Appl.*, vol. 63, pp. 1–10, 1996.
- [17] A. S. Leong, S. Dey, and J. S. Evans, "Probability of error analysis for hidden Markov model filtering with random packet loss," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 809–821, Mar. 2007.
- [18] K. E. Atkinson, *The Numerical Solution of Integral Equations of the Second Kind*. Cambridge, U.K.: Cambridge Univ. Press, 1997.
- [19] A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, 2nd ed. New York: McGraw-Hill, 1978.
- [20] F. W. J. Olver, *Asymptotics and Special Functions*. New York: Academic, 1974.
- [21] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.
- [22] J. H. Miller and J. B. Thomas, "Detectors for discrete-time signals in non-Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 18, no. 2, pp. 241–250, Mar. 1972.

A Generalized Subspace Approach for Mobile Positioning With Time-of-Arrival Measurements

H. C. So, *Member, IEEE*, and Frankie K. W. Chan

Abstract—The problem of locating mobile terminals has received considerable attention particularly in the field of wireless communications. In this correspondence, a simple subspace-based algorithm for mobile positioning with the use of time-of-arrival measurements deduced from signals received at three or more reference base stations is derived and analyzed. It is shown that the proposed approach is a generalization of the mobile localization method based on multidimensional similarity analysis. Computer simulations are included to contrast the estimator performance with Cramér–Rao lower bound.

Index Terms—Fast algorithm, mobile terminal, position estimation, range measurements.

I. INTRODUCTION

Mobile terminal (MT) positioning has been receiving considerable interest, especially after the Federal Communications Commission in the United States has adopted rules to improve the Emergency 911 (E-911) services by mandating the accuracy of locating an E-911 caller to be within a specified range, even for a wireless phone user [1]. Apart from emergency assistance, mobile position information is also the key enabler for a large number of innovative applications such as personal localization and monitoring, fleet management, asset tracking, travel services, location-based advertising, and billing [2].

Common positioning approaches [3] are based on time-of-arrival (TOA), received signal strength, time-difference-of-arrival, and/or angle-of-arrival measurements determined from the MT signal received at several reference base stations (BSs) with known locations. In this correspondence, we focus on two-dimensional (2-D) MT localization given the TOA information. In the TOA method, the one-way propagation time of the signal travelling between the MT and each of the BSs is measured. Each TOA measurement then provides a circle centered at the BS on which the MT must lie. With three or more BSs, the measurements are converted into a set of circular equations, from which the MT position can be determined with the knowledge of the BS geometry.

The optimum TOA-based localization approach involves solving the nonlinear circular equations in an iterative manner, and commonly used techniques [4] include linearization via Taylor-series expansion, the steepest descent method, and Newton-type iteration. However, this approach is computationally intensive and sufficiently precise initial estimates are required to obtain the global solution. On the other hand, computationally efficient but suboptimum position estimators, which allow real-time realization as well as ensure global convergence, have also been proposed in the literature [5]–[9]. In the least-squares (LS) calibration method [5], the nonlinear equations are reorganized into a set of linear equations via introduction of an extra variable, which is a function of the source position, and these linear equations are then

Manuscript received July 24, 2006; revised February 20, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Fredrik Gustafsson. The work described in this correspondence was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 119606).

The authors are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: hcs0@ee.cityu.edu.hk; k.w.chan@student.cityu.edu.hk).

Digital Object Identifier 10.1109/TSP.2003.896902