Probability of Error Analysis for Hidden Markov Model Filtering With Random Packet Loss

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

***Abstract*—This paper studies the probability of error for maximum *a posteriori* (MAP) estimation of hidden Markov models, where measurements can be either lost or received according to another Markov process. Analytical expressions for the error probabilities are derived for the noiseless and noisy cases. Some relationships between the error probability and the parameters of the loss process are demonstated via both analysis and numerical results. In the high signal-to-noise ratio (SNR) regime, approximate expressions which can be more easily computed than the exact analytical form for the noisy case are presented.**

***Index Terms*—Hidden Markov model, observation losses, probability of error, state estimation.**

I. Introduction

WIRELESS sensor networks have received huge interest in the research community recently, due to the many technical challenges which have to be overcome in order to realise their full benefits. Algorithms for signal processing and their performance in unreliable environments is an important aspect in the design of such networks. This paper considers the error probabilities associated with the state estimation of Markov chains when measurements can be lost, with the loss process modelled by another Markov chain.

Estimation with lossy measurements was considered for linear systems in [1], for the Kalman filtering problem with losses modelled by an independent and identically distributed (i.i.d.) Bernoulli process. They showed that for an unstable system, there exists a threshold such that if the probability of reception exceeds this threshold then the expected value (with respect to the loss process) of the error covariance matrix (which is a random quantity due to random losses) will be bounded, but if the probability of reception is lower than this threshold then the error covariance diverges. In a slightly different context, [2] extends these results to Markovian loss processes, which allows modelling of more "bursty" types of errors. Estimation with Markovian packet losses was also studied in [3], and suboptimal estimators were derived which can be used to provide upper bounds on the estimation errors of the optimal estimator.

Manuscript received November 9, 2005; revised May 4, 2006. This work was supported by the Australian Research Council. The associate editor coordinating the review of this paper and approving it for publication was Dr. Yonina C. Eldar.

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

Color versions of figures available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2006.888056

The purpose of this paper is to use some of the ideas relating to lossy measurement processes, but to apply it to the different problem of state estimation for Markov chains. Hidden Markov models (HMMs) have found numerous applications [4], and in fields such as radar tracking and biology, sensor networks could potentially provide additional benefits. For HMM estimation problems, the state space is often finite, thus, notions of estimation stability as in [1] are not appropriate. Instead here we will use the probability of estimation error as our measure of performance, also see [5], [6] for mean square error (mse) criterions. Obtaining analytical expressions for the error performance associated with filtering for HMMs (even without any loss process) is a difficult problem however, where few general results are known. Some results for the continuous time case may be found in [5]. In discrete time, asymptotic formulae for "slow" Markov chains with finite state space were obtained in [7] for the probability of error, and [6] for a mean square error criterion. A general characterization of the error probability for the two-state hidden Markov model in discrete time was derived in [8], and a numerical method to calculate it was proposed.

The organization of the paper is as follows. We will first study the simpler problem with observation losses but no noise in Section III. Analytical expressions for the error probabilities will be derived and some special cases presented in Sections III-A and B, respectively. Some relationships between the error probability and the parameters of the Markovian loss process are established in Section III-C and numerical studies presented in Section III-D. Multi-state Markov chains are considered in Section III-E. We will then shift our attention to the more general HMM problem with noise in Section IV. In Sections IV-A and B, we will characterize the error probabilities for the two-state Markov chain, though using quite different methods. Section IV-C will present some numerical studies for the noisy case. It is more difficult here to prove properties similar to the ones in the noiseless case, and some conjectures which will require further investigation are stated. Section IV-D studies the situation when the signal is i.i.d. In Sections IV-E and F, high SNR approximations for the two-state and multi-state Markov chains respectively are derived and numerical comparisons made.

II. Model and Notational Conventions

The main model of interest is

$$Y_k = \gamma(Z_k)h(X_k) + v_k.$$

Here, $\{Y_k\}$ is the observation process, $\{v_k\}$ is the noise process which will be i.i.d. and $N(0,\sigma^2)$.[1] $\{X_k\}$ and $\{Z_k\}$ are homogeneous two-state Markov chains which are assumed to be independent of each other, with $h(1)=-1$, $h(2)=1$, $\gamma(1)=0$, $\gamma(2)=1$. $h(X_k)$ can be interpreted as the signal that we wish to estimate, and $\gamma(Z_k)$ as the correlated loss process, with the correlation modeled by a Markov chain. We will assume that $\gamma(Z_k)$ is known to us at each time instant. In Sections III-E and IV-F we will look at the situation where $\{X_k\}$ is a multi-state Markov chain.

In this paper, we will use the convention $a_{ij}=P(X_{k+1}=i|X_k=j)$ and $g_{ij}=P(Z_{k+1}=i|Z_k=j)$ for the transition probability matrices $A=(a_{ij})$ and $G=(g_{ij})$, (the columns will then sum to one). Assuming that a_{12}, a_{21}, g_{12}, $g_{21}\neq 0$, unique stationary probabilities will then exist and are given by $P(X_k=1)=a_{12}/(a_{12}+a_{21}), P(X_k=2)=a_{21}/(a_{12}+a_{21})$, $P(Z_k=1)=g_{12}/(g_{12}+g_{21}), P(Z_k=2)=g_{21}/(g_{12}+g_{21})$.[2]

Denote the conditional probability vector for the HMM filter by $\Pi_{k|k}$, with the ith entry being $\Pi^i_{k|k}=P(X_k=i|Y_0=y_0,\ldots,Y_k=y_k,Z_0=z_0,\ldots,Z_k=z_k)$. The MAP estimate of X_k, which is well-known to minimize the probability of error, is for the two-state case

$$\hat{X}_k=\begin{cases}1, & \Pi^1_{k|k}>\Pi^2_{k|k}\\ 2, & \text{otherwise.}\end{cases}$$

III. Noiseless Case

We first study the simpler model $Y_k=\gamma(Z_k)h(X_k)$ which does not have the noise term v_k. The probability of estimation error will be given in terms of an infinite series, which is a more explicit form than that which will be derived for the noisy case in Section IV-A. The noiseless case considered here is quite suitable for the noisy situation at high SNR, since (roughly speaking) at high SNR the errors due to the packet loss process tend to dominate the errors due to the noise term, see, for example, the discussion at the end of Section IV-E. Indeed, the derivation in Section IV-E of an approximation for the error probability at high SNR will be based on some of the techniques of this section.

A. Derivation of Probability of Error

For this simple noiseless model, whenever there is no packet loss, i.e., $Z_k=2$, the estimate (of $h(X_k)$) will be the same as the measurement. The probability vectors $\Pi_{k|k}$ are, therefore, updated as

$$\Pi_{k+1|k+1}=\begin{cases}A\Pi_{k|k}, & \gamma(Z_{k+1})=0\\ [1\ \ 0]^T, & \gamma(Z_{k+1})=1, Y_{k+1}=h(1)\\ [0\ \ 1]^T, & \gamma(Z_{k+1})=1, Y_{k+1}=h(2).\end{cases}$$

[1]Other noise types such as noise with state-dependent variances are possible, but some derivations will be more complicated.

[2]Strictly speaking, this is true if the initial state of the Markov chain has the same distribution as the stationary distribution, otherwise this holds only in the limit as $k\to\infty$.

So whenever there is no packet loss, the probability vector will "reset" to either $[1\ 0]^T$ or $[0\ 1]^T$, a fact we will exploit in our derivation of the probability of error. When a packet is received, no errors will be made, so

$$P(\text{Error})=P(\text{Error},Z_k=1).$$

This can be further split up as follows:

$$\begin{aligned}P(\text{Error},Z_k=1)&=P(\text{Error},Z_{k-1}=1,Z_k=1)\\&\quad+P(\text{Error},Z_{k-1}=2,Z_k=1)\\&=P(\text{Error},Z_{k-2}=1,Z_{k-1}=1,Z_k=1)\\&\quad+P(\text{Error},Z_{k-2}=2,Z_{k-1}=1,Z_k=1)\\&\quad+P(\text{Error},Z_{k-1}=2,Z_k=1)\\&=\ldots=\sum_{n=1}^{\infty}p(n)\end{aligned}$$

where $p(n)\equiv P(\text{Error},Z_{k-n}=2,Z_{k-n+1}=1,Z_{k-n+2}=1,\ldots,Z_k=1)$.

Expressions for each term $p(n)$ can be derived. Define $\Pi_{1,n}\equiv A^n[1\ \ 0]^T$ and $\Pi_{2,n}\equiv A^n[0\ \ 1]^T$, with $\Pi^j_{i,n}$ representing the jth element of $\Pi_{i,n}$. For brevity, also let $Z\equiv(Z_{k-n}=2,Z_{k-n+1}=1,Z_{k-n+2}=1,\ldots,Z_k=1)$. Then

$$\begin{aligned}p(n)&=P\left(Z,X_k=1,\arg\max_j\Pi^j_{1,n}\neq 1,X_{k-n}=1\right)\\&\quad+P\left(Z,X_k=2,\arg\max_j\Pi^j_{1,n}\neq 2,X_{k-n}=1\right)\\&\quad+P\left(Z,X_k=1,\arg\max_j\Pi^j_{2,n}\neq 1,X_{k-n}=2\right)\\&\quad+P\left(Z,X_k=2,\arg\max_j\Pi^j_{2,n}\neq 2,X_{k-n}=2\right)\\&=\frac{g_{21}}{g_{12}+g_{21}}g_{12}(1-g_{21})^{n-1}\\&\quad\times\sum_{r,s=1}^{2}P(X_{k-n}=r)a^{(n)}_{sr}I_{\operatorname{argmax}_j\Pi^j_{r,n}\neq s}\end{aligned}\tag{1}$$

where I is the indicator function and $a^{(n)}_{ij}$ is the (i,j)th entry of the matrix A^n. For a more explicit expression for $p(n)$, note that

$$A^n=\begin{bmatrix}\frac{a_{12}+a_{21}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}} & \frac{a_{12}-a_{12}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}}\\ \frac{a_{21}-a_{21}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}} & \frac{a_{21}+a_{12}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}}\end{bmatrix}\tag{2}$$

which may be verified using induction. Also define

$$\begin{aligned}q^1_n&\equiv\frac{a_{12}-a_{21}+2a_{21}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}}\\ q^2_n&\equiv\frac{a_{12}-a_{21}-2a_{12}(1-a_{12}-a_{21})^n}{a_{12}+a_{21}}.\end{aligned}\tag{3}$$

TABLE I
SIMULATION AND ANALYTICAL COMPARISON OF NOISELESS ERROR PROBABILITIES

a_{12}	a_{21}	g_{12}	g_{21}	simulation	analytical
0.67	0.51	0.30	0.06	0.3591	0.3596
0.34	0.31	0.97	0.65	0.2210	0.2208
0.91	0.43	0.38	0.18	0.2178	0.2178
0.20	0.74	0.59	0.69	0.0982	0.0981
0.09	0.04	0.69	0.55	0.0507	0.0506

Then it is easily shown that (1) can also be written in the form shown in (4).

$$p(n)=\begin{cases} \frac{g_{21}}{g_{12}+g_{21}}g_{12}(1-g_{21})^{n-1} \\ \quad\times\left[\frac{a_{12}}{a_{12}+a_{21}}a_{11}^{(n)}+\frac{a_{21}}{a_{12}+a_{21}}a_{12}^{(n)}\right], & q_n^1\le 0, q_n^2\le 0 \\ \frac{g_{21}}{g_{12}+g_{21}}g_{12}(1-g_{21})^{n-1} \\ \quad\times\left[\frac{a_{12}}{a_{12}+a_{21}}a_{11}^{(n)}+\frac{a_{21}}{a_{12}+a_{21}}a_{22}^{(n)}\right], & q_n^1\le 0, q_n^2> 0 \\ \frac{g_{21}}{g_{12}+g_{21}}g_{12}(1-g_{21})^{n-1} \\ \quad\times\left[\frac{a_{12}}{a_{12}+a_{21}}a_{21}^{(n)}+\frac{a_{21}}{a_{12}+a_{21}}a_{12}^{(n)}\right], & q_n^1> 0, q_n^2\le 0 \\ \frac{g_{21}}{g_{12}+g_{21}}g_{12}(1-g_{21})^{n-1} \\ \quad\times\left[\frac{a_{12}}{a_{12}+a_{21}}a_{21}^{(n)}+\frac{a_{21}}{a_{12}+a_{21}}a_{22}^{(n)}\right], & q_n^1> 0, q_n^2> 0. \end{cases} \tag{4}$$

Hence

$$P(\text{Error})=\sum_{n=1}^{\infty}p(n)$$

where $p(n)$ is given by (4), $a_{ij}^{(n)}$ is the (i,j)th entry of A^n in (2), and q_n^1 and q_n^2 are given by (3). Numerical computation of such infinite series can be easily handled using computer algebra software such as Mathematica.

In Table I, we compare the derived expression with simulation results for a selection of different parameter values. We set the distribution of the initial states of the Markov chains to be equal to the stationary distributions (though by the exponential forgetting property [9] the effect of the initial state should not have a major effect for long runs), and then run Monte Carlo simulations of the filtering updates to obtain the probability of error. The simulations results were averaged over 10 runs, each run of length 100 000. It may be seen that, not suprisingly, there is very close agreement at all values considered.

B. Special Cases

In certain cases, the expression for the error probability can be further simplified. We present two examples.

(i) If $a_{12}=a_{21}<0.5$ so that A is symmetric, then $q_n^1>0$ and $q_n^2\le 0$ always, so that

$$\begin{aligned}P(\text{Error})&=\sum_{n=1}^{\infty}\frac{1}{2}\frac{g_{21}g_{12}}{g_{12}+g_{21}}(1-g_{21})^{n-1}\left[1-(1-2a_{12})^n\right]\\&=\frac{g_{12}a_{12}}{(g_{12}+g_{21})(2a_{12}+g_{21}-2a_{12}g_{21})}.\end{aligned}$$

(ii) Suppose the signal is i.i.d., i.e., $a_{21}=1-a_{12}$. If $a_{12}\le 0.5$, then $q_n^1\le 0$, $q_n^2\le 0$, $\forall n$, and

$$P(\text{Error})=\sum_{n=1}^{\infty}\frac{g_{21}g_{12}}{g_{12}+g_{21}}(1-g_{21})^{n-1}a_{12}=\frac{a_{12}g_{12}}{g_{12}+g_{21}}.$$

If $a_{12}>0.5$, then $q_n^1>0$, $q_n^2>0$, $\forall n$, and

$$P(\text{Error})=\frac{(1-a_{12})g_{12}}{g_{12}+g_{21}}.$$

Fig. 5 of Section III-D contains some simulation results. The linear dependence on the stationary probability $g_{12}/(g_{12}+g_{21})$ when the data is i.i.d. also holds in the noisy case, see Section IV-D.

C. Theoretical Properties

We now demonstrate some relationships between the probability of error and the parameters of the packet loss process. Proofs of Theorems 1–3 may be found in the Appendix.

Theorem 1: For fixed A and g_{21}, the probability of error is monotonically increasing in g_{12}.

Theorem 2: For fixed A and g_{12}, the probability of error is monotonically decreasing in g_{21}.

Theorem 1 states that the error probability increases as g_{12} increases, when all other parameters are fixed. Intuitively this is reasonable, since g_{12} is the probability that the next packet is lost given that the current packet has been received, so we are more likely to drop packets and do worse at estimation when this parameter is increased. Theorem 2 is also quite intuitive, as g_{21} is the probability that the next packet will be received correctly given that the current packet has been lost, so increasing this parameter should improve our estimation performance.

For the third result, let p_0 be the stationary probability that a measurement is not received, i.e., $p_0=P(Z_k=1)=g_{12}/(g_{12}+g_{21})$. Theorem 3 shows that in general p_0 alone does not uniquely determine the error probabilities, but also depends on the sizes of g_{12} and g_{21}, which can be interpreted as how quickly/slowly the packet loss process is varying in time. For example, when both g_{12} and g_{21} are small, transitions from one state to the other are rare, so that we can regard the Markov chain as being slow. Essentially, Theorem 3 says that for a given p_0, slower dynamics are worse for state estimation in that we will get a higher probability of error (except perhaps when the signal is i.i.d. as noted in example (ii) of Section III-B).

Theorem 3:

(i) For fixed A and p_0, the probability of error is nonincreasing in g_{21} (equivalently in g_{12}).

(ii) The probability of error converges, for fixed p_0 and as $g_{21}\to 0$, to p_0E_0, where $E_0\equiv\min(a_{12}/(a_{12}+a_{21}),a_{21}/(a_{12}+a_{21}))$ is the probability of error in the complete absence of observations.

By Theorem 3 (ii), p_0E_0 is therefore an upper bound on the error probability. A lower bound can also be derived, by noting that for $p_0\le 1/2$, the largest possible values for g_{12} and g_{21} are $g_{21}=1$ and $g_{12}=p_0/(1-p_0)$, and for $p_0>1/2$, the largest possible values are $g_{12}=1$ and $g_{21}=(1-p_0)/p_0$. Substituting

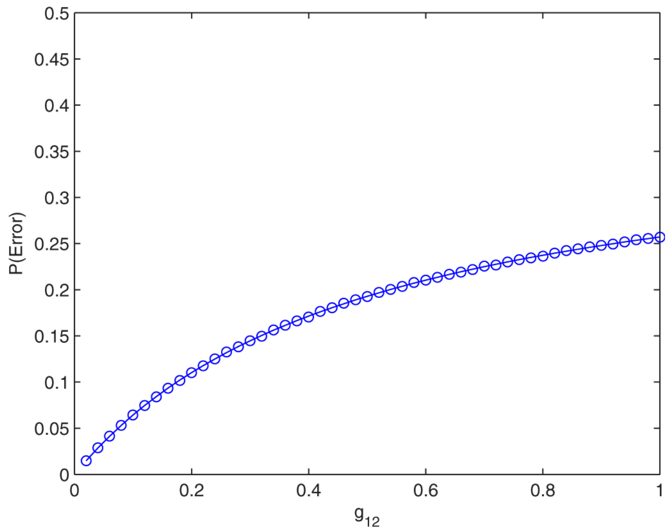


Fig. 1. Noiseless probability of error for various g_{12}.

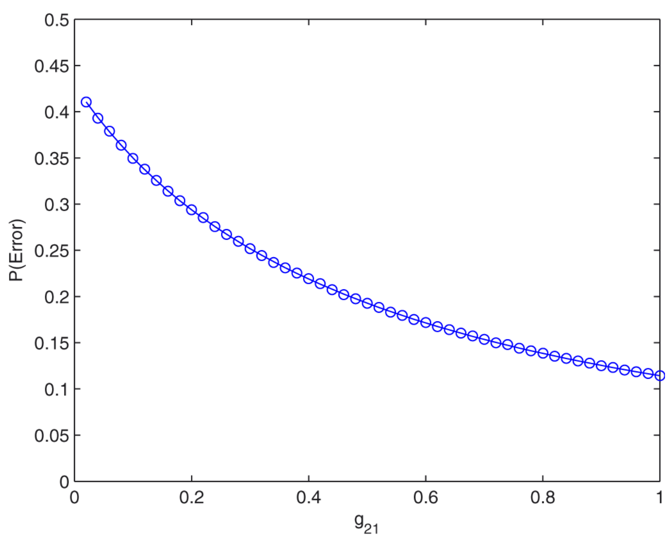


Fig. 2. Noiseless probability of error for various g_{21}.

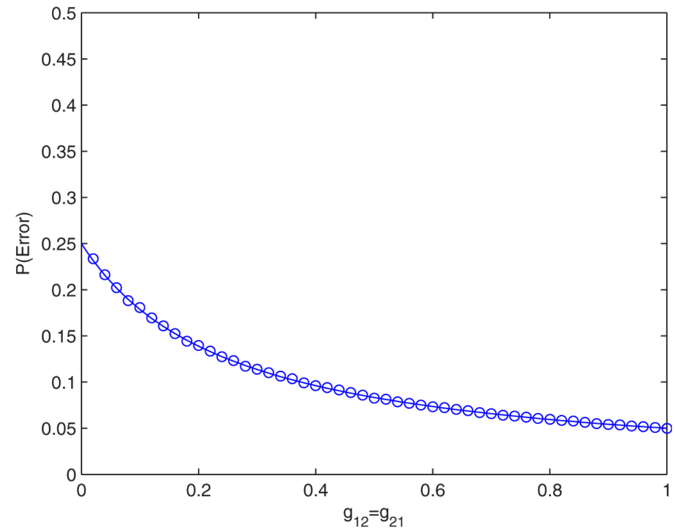


Fig. 3. Noiseless probability of error for various g_{12}, with p_0 fixed.

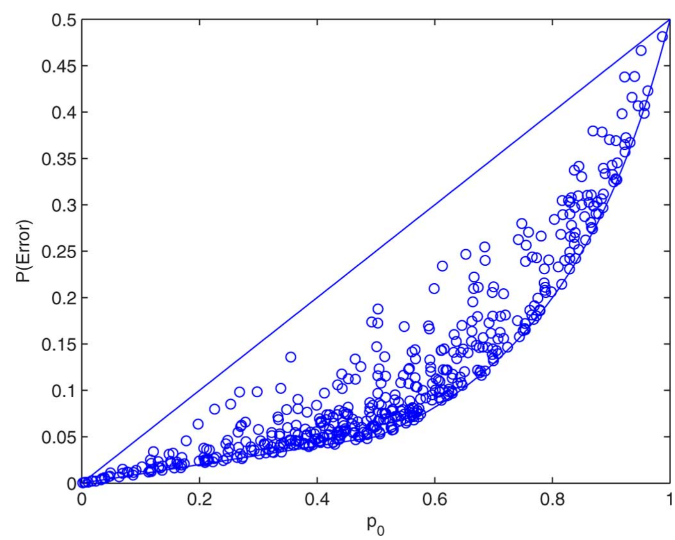


Fig. 4. Noiseless probability of error and bounds for various p_0, symmetric A.

these values into the formula for $P(\text{Error})$ and using Theorem 3 (i), we then obtain the lower bound

$$P(\text{Error}) \geq \begin{cases} p_0 t(1), & p_0 \leq 0.5 \\ \sum_{n=1}^{\infty}(1-p_0)\left(\frac{2p_0-1}{p_0}\right)^{n-1} t(n), & p_0 > 0.5 \end{cases}$$

where $t(n)$ is defined as (17) in the Appendix. As a special case, when $a_{12} = a_{21} < 0.5$ we have the more explicit expression

$$P(\text{Error}) \geq \begin{cases} a_{12}p_0, & p_0 \leq 0.5 \\ \frac{a_{12}p_0^2}{1-2a_{12}+(4a_{12}-1)p_0}, & p_0 > 0.5. \end{cases}$$

For the i.i.d. signal case, it can also be shown that the upper and lower bounds coincide.

D. Numerical Studies

For the simulations in the first three graphs, the length of each run is one million. The solid lines represent the theoretical error probability. In Fig. 1 we plot the simulated probability of error for 50 values of g_{12}, with $a_{12} = 0.4$, $a_{21} = 0.3$, and $g_{21} = 0.5$. In Fig. 2 we plot the simulated probability of error for 50 values of g_{21}, with $a_{12} = 0.4$, $a_{21} = 0.3$, and $g_{12} = 0.5$. In Fig. 3 we plot the simulated probability of error for 50 values of g_{12}, with $g_{21} = g_{12}$ (i.e., $p_0 = 0.5$ is fixed), $a_{12} = 0.1$ and $a_{21} = 0.1$. We can see that the results are in agreement with Theorems 1–3, respectively.

For the next two graphs, we randomly generate both g_{12} and g_{21}, and then form $p_0 = g_{12}/(g_{12} + g_{21})$. The length of each simulation run is 100 000. In Fig. 4 we plot the simulated probability of error for 500 values of p_0, with $a_{12} = a_{21} = 0.1$, i.e., A is symmetric. The solid lines are plots of the upper and lower bounds on the error probability mentioned after the statement of Theorem 3. The simulation results can be seen to lie within the bounds. In Fig. 5 we plot the simulated probability of error for 500 values of p_0, with $a_{12} = 0.7$ and $a_{21} = 0.3$, i.e., signal is i.i.d. The linear dependence on the probability of not receiving a packet, agrees with example (ii) of Section III-B.

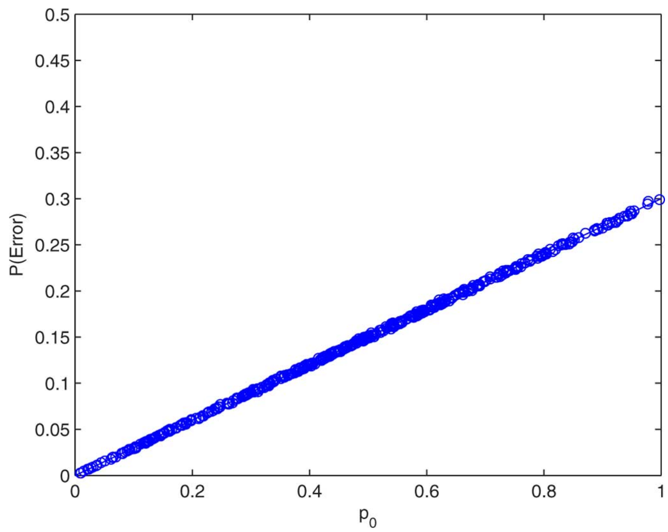


Fig. 5. Noiseless probability of error for various p_0, signal i.i.d.

E. Multiple States

We now consider $\{X_k\}$ as a Markov chain having $M > 2$ states, with the assumption that it has a unique stationary distribution. The packet loss process $\{Z_k\}$ is still the same two-state Markov chain.

The probability vectors are updated as

$$\Pi_{k+1|k+1} = \begin{cases} A\Pi_{k|k}, & \gamma(Z_{k+1}) = 0 \\ [1\ 0\ \ldots\ 0]^T, & \gamma(Z_{k+1})=1, Y_{k+1}=h(1) \\ [0\ 1\ \ldots\ 0]^T, & \gamma(Z_{k+1})=1, Y_{k+1}=h(2) \\ \vdots & \vdots \\ [0\ 0\ \ldots\ 1]^T, & \gamma(Z_{k+1})=1, Y_{k+1}=h(M) \end{cases}$$

and MAP estimates of the states are obtained as $\hat{X}_k = \arg\max_{j=1,\ldots,M} \Pi^j_{k|k}$.

Define $\Pi_{i,n} \equiv A^n[0\ \ldots\ 0\ 1\ \ldots\ 0]^T$ where the 1 is in the i-th position. Then in a straightforward extension of the two-state case (1), we can derive

$$p(n) = \frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1} \times \sum_{r,s=1}^{M} P(X_{k-n}=r) a^{(n)}_{sr} I_{\mathrm{argmax}_j \Pi^j_{r,n} \neq s} \quad (5)$$

with the probability of error given by $P(\mathrm{Error}) = \sum_{n=1}^{\infty} p(n)$. More explicit expressions (in terms of the Markov chain parameters) for the stationary probabilities $P(X_{k-n} = r)$ and the elements $a^{(n)}_{sr}$ would be very complicated to write down without additional structure in the transition matrix A, since for M states, the A matrix would have $M(M-1)$ free parameters in general. However, given a set of parameters these quantities can be evaluated on a computer quite easily.

As an example, let $g_{12} = 0.26$, $g_{21} = 0.48$ and

$$A = \begin{bmatrix} 0.18 & 0.56 & 0.13 \\ 0.76 & 0.36 & 0.27 \\ 0.06 & 0.08 & 0.60 \end{bmatrix}.$$

Computation of the error probability using (5) is 0.1527. Simulation results averaged over 10 runs, each of length 100 000, gives an error probability of 0.1523, which is very close to the analytical expression.

IV. Noisy Case

A. Derivation of Probability of Error

In this section we derive expressions for the probability of error in the noisy case. Unfortunately, the methods used for the noiseless case don't seem to extend to the situation here. We will use a different method, whose analysis is based in part on [8]. Recall the model from Section II, $Y_k = \gamma(Z_k)h(X_k) + v_k$. Given X_k and $\gamma(Z_k) = 1$, the observations Y_k are conditionally distributed as

$$P(Y_k \in dy | X_k = i, Z_k = 2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_k - h(i))^2}{2\sigma^2}\right) dy \equiv b_i(y_k)dy.$$

It is well known that the probability vectors can be updated recursively as follows:

$$\Pi_{k+1|k+1} = \begin{cases} A\Pi_{k|k}, & \gamma(Z_{k+1}) = 0 \\ \frac{B_{y_{k+1}} A\Pi_{k|k}}{[1\ \ 1]B_{y_{k+1}} A\Pi_{k|k}}, & \gamma(Z_{k+1}) = 1 \end{cases} \quad (6)$$

where

$$B_{y_{k+1}} = \begin{bmatrix} b_1(y_{k+1}) & 0 \\ 0 & b_2(y_{k+1}) \end{bmatrix}.$$

Using the definition $q_k \equiv \Pi^1_{k|k} - \Pi^2_{k|k}$ (noting that $-1 \le q_k \le 1$), the probability of filtering error can then be written as

$$\begin{aligned} P(\mathrm{Error}) &= P(X_k = 1, \hat{X}_k = 2) + P(X_k = 2, \hat{X}_k = 1) \\ &= P(X_k = 1, Z_k = 1, q_k \le 0) \\ &\quad + P(X_k = 1, Z_k = 2, q_k \le 0) \\ &\quad + P(X_k = 2, Z_k = 1, q_k > 0) \\ &\quad + P(X_k = 2, Z_k = 2, q_k > 0) \\ &= \int_{-1}^{0} f_k^{1,1}(q)dq + \int_{-1}^{0} f_k^{1,2}(q)dq \\ &\quad + \int_{0}^{1} f_k^{2,1}(q)dq + \int_{0}^{1} f_k^{2,2}(q)dq \end{aligned} \quad (7)$$

where $f_k^{i,I}(q)dq \equiv P(X_k = i, Z_k = I, q_k \in (q, q+dq))$.

To find a recursive relation which will allow us to characterize the densities $f_k^{i,I}(q)$, first consider

$$\begin{aligned} &P(X_{k+1} = i, Z_{k+1} = I, q_{k+1} \in (q, q+dq), \\ &\quad X_k = j, Z_k = J, q_k \in (\tilde{q}, \tilde{q}+d\tilde{q})) \\ &\quad = P(q_{k+1} \in (q, q+dq) | X_{k+1} = i, Z_{k+1} = I, q_k = \tilde{q}) \\ &\quad\quad \times P(X_{k+1} = i, Z_{k+1} = I | X_k = j, Z_k = J) \\ &\quad\quad \times P(X_k = j, Z_k = J, q_k \in (\tilde{q}, \tilde{q}+d\tilde{q})) \\ &\quad = S_{i,I}(q, \tilde{q})dq \times a_{ij} g_{IJ} \times f_k^{j,J}(\tilde{q})d\tilde{q} \end{aligned}$$

where we have used the Markov property and the independence of $\{X_k\}$ and $\{Z_k\}$, and defined

$$S_{i,I}(q,\tilde{q})dq \equiv P\left(q_{k+1} \in (q, q+dq) | X_{k+1}=i, Z_{k+1}=I, q_k=\tilde{q}\right).$$

For $\gamma(Z_{k+1}) = 1$, i.e., $Z_{k+1} = 2$, we can derive in a similar manner to [8] (which only contained expressions for symmetric A) the recursion shown in (8) at the bottom of the page, and so

$$S_{i,2}(q,\tilde{q})dq = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(g(q,\tilde{q})-h(i))^2}{2\sigma^2}\right] \frac{\sigma^2}{1-q^2}dq$$

with $g(q,\tilde{q}) = -(\sigma^2/2)\ln((1+q)[1-(a_{12}-a_{21})-(1-a_{12}-a_{21})\tilde{q}])/((1-q)[1+(a_{12}-a_{21})+(1-a_{12}-a_{21})\tilde{q}])$.

For $\gamma(Z_{k+1}) = 0$, i.e., $Z_{k+1} = 1$, it is straightforward to show that the recursion for q is now

$$q_{k+1} = a_{12} - a_{21} + (1 - a_{12} - a_{21})q_k. \tag{9}$$

Thus

$$\begin{aligned} S_{i,1}(q,\tilde{q})dq &= P\left(a_{12}-a_{21}+(1-a_{12}-a_{21})\tilde{q} \in (q, q+dq)\right) \\ &= I_B(q,\tilde{q},dq) \end{aligned}$$

where I is the indicator function and $B \equiv \{(q,\tilde{q},dq) : a_{12} - a_{21} + (1-a_{12}-a_{21})\tilde{q} \in (q,q+dq)\}$.

Hence in the steady state we have the following relations for the densities, which is a system of four Fredholm integral equations

$$f^{i,I}(q)dq = \sum_{j=1}^{2}\sum_{J=1}^{2} a_{ij} g_{IJ} \int_{-1}^{1} S_{i,I}(q,\tilde{q}) f^{j,J}(\tilde{q}) d\tilde{q} dq, \quad i=1,2, I=1,2. \tag{10}$$

B. Numerical Method

To compute the error probability we will need to numerically solve the system of integral equations (10). We will present an existing method which is slightly different from that of [8], where convergence analysis is perhaps more readily obtained.[3]

From p. 151 of [10], we can transform (10) into a single integral equation as follows. Define

$$\Phi(q) = \begin{cases} f^{1,1}(q), & -1<q<1 \\ f^{1,2}(q-2), & 1<q<3 \\ f^{2,1}(q-4), & 3<q<5 \\ f^{2,2}(q-6), & 5<q<7 \end{cases}$$

$$K(q,\tilde{q}) = \begin{cases} a_{11}g_{11}S_{1,1}(q,\tilde{q}), & -1<q<1,\ -1<\tilde{q}<1 \\ a_{11}g_{12}S_{1,1}(q,\tilde{q}), & -1<q<1,\ 1<\tilde{q}<3 \\ a_{12}g_{11}S_{1,1}(q,\tilde{q}), & -1<q<1,\ 3<\tilde{q}<5 \\ a_{12}g_{12}S_{1,1}(q,\tilde{q}), & -1<q<1,\ 5<\tilde{q}<7 \\ \vdots & \\ a_{21}g_{21}S_{2,2}(q,\tilde{q}), & 5<q<7,\ -1<\tilde{q}<1 \\ a_{21}g_{22}S_{2,2}(q,\tilde{q}), & 5<q<7,\ 1<\tilde{q}<3 \\ a_{22}g_{21}S_{2,2}(q,\tilde{q}), & 5<q<7,\ 3<\tilde{q}<5 \\ a_{22}g_{22}S_{2,2}(q,\tilde{q}), & 5<q<7,\ 5<\tilde{q}<7. \end{cases} \tag{11}$$

Then it can be seen that (10) is equivalent to the homogeneous Fredholm equation

$$\Phi(q) - \int_{-1}^{7} K(q,\tilde{q})\Phi(\tilde{q})d\tilde{q} = 0 \tag{12}$$

with $\Phi(q)$ also satisfying the normalising condition $\int_{-1}^{7}\Phi(q)dq = 1$. For the numerical solution of (12), consider the related eigenvalue problem [11]

$$\gamma\Phi(q) = \int_{-1}^{7} K(q,\tilde{q})\Phi(\tilde{q})d\tilde{q} \tag{13}$$

which corresponds to (12) when $\gamma = 1$. We will solve (13) using the Nyström method [11], [12]. Replacing the integral by a $4N$-point quadrature rule,[4] and defining

$$\mathbf{K} \equiv \begin{bmatrix} w_1 K(t_1,t_1) & \dots & w_{4N}K(t_1,t_{4N}) \\ \vdots & \ddots & \vdots \\ w_1 K(t_{4N},t_1) & \dots & w_{4N}K(t_{4N},t_{4N}) \end{bmatrix}$$

we obtain

$$\mathbf{K}\begin{bmatrix} \Phi(t_1) \\ \vdots \\ \Phi(t_{4N}) \end{bmatrix} = \gamma \begin{bmatrix} \Phi(t_1) \\ \vdots \\ \Phi(t_{4N}) \end{bmatrix}$$

where w_j represent the weights and t_j the quadrature points of the quadrature rule. In the results presented in this paper, the midpoint rule is used, though other alternatives such as composite Gauss-Legendre quadrature are possible [12, p.110.].

To obtain an approximation for $\Phi(q)$, we then take the eigenvector that corresponds to the largest real eigenvalue of $\mathbf{K}$, and normalise it so that $\int_{-1}^{7}\Phi(q)dq = 1$ is satisfied. Using a "weak" version of the Perron-Frobenius theorem [13, p. 28] on $\mathbf{K}$ shows that this eigenvector will have nonnegative entries, which is required if it is to approximate a probability density. The probability of error can then be calculated from (7) and (11).

[3]The method presented here and its analysis can also be applied with slight modifications to the problem in [8]

[4]We call it a $4N$-point rather an N-point quadrature rule for convenience, since $\Phi(q)$ is a combination of four densities

$$q_{k+1} = \frac{\exp\left(-\frac{2y_{k+1}}{\sigma^2}\right)[1+a_{12}-a_{21}+(1-a_{12}-a_{21})q_k] - 1 + [a_{12}-a_{21}+(1-a_{12}-a_{21})q_k]}{\exp\left(-\frac{2y_{k+1}}{\sigma^2}\right)[1+a_{12}-a_{21}+(1-a_{12}-a_{21})q_k] + 1 - [a_{12}-a_{21}+(1-a_{12}-a_{21})q_k]} \tag{8}$$

We would like the largest eigenvalue to be close to one. As $N \to \infty$, $\mathbf{K}$ will tend to a stochastic matrix, in the sense that the sum of each column will converge to one. For example, for the first column

$$\sum_{i=1}^{4N} w_1 K(t_i, t_1)$$
$$\to \int_{-1}^{1} (a_{11}g_{11}S_{1,1}(q,t_1) + a_{11}g_{21}S_{1,2}(q,t_1)$$
$$+a_{21}g_{11}S_{2,1}(q,t_1) + a_{21}g_{21}S_{2,2}(q,t_1))\,dq$$
$$= a_{11}g_{11} + a_{11}g_{21} + a_{21}g_{11} + a_{21}g_{21} = 1$$

where convergence is achieved for any reasonable composite quadrature scheme, such as the midpoint rule [14, p.116]. Similarly this holds for the other columns of $\mathbf{K}$. Since a stochastic matrix has largest eigenvalue one, and by the continuity of eigenvalues [15], it then follows that the largest eigenvalue of $\mathbf{K}$ can be made arbitrarily close to one for N sufficiently large. We can also show that the eigenvector corresponding to the largest eigenvalue is unique. Note that we can partition $\mathbf{K}$ as

$$\mathbf{K} = \begin{bmatrix} a_{11}g_{11}S_{11} & a_{11}g_{12}S_{11} & a_{12}g_{11}S_{11} & a_{12}g_{12}S_{11} \\ a_{11}g_{21}S_{12} & a_{11}g_{22}S_{12} & a_{12}g_{21}S_{12} & a_{12}g_{22}S_{12} \\ a_{21}g_{11}S_{21} & a_{21}g_{12}S_{21} & a_{22}g_{11}S_{21} & a_{22}g_{12}S_{21} \\ a_{21}g_{21}S_{22} & a_{21}g_{22}S_{22} & a_{22}g_{21}S_{22} & a_{22}g_{22}S_{22} \end{bmatrix} \quad (14)$$

with each element representing an $N \times N$ matrix. Referring to (14) and the definition of $\mathbf{K}$, it may be easily shown that the blocks $a_{ij}g_{IJ}S_{12}$ and $a_{ij}g_{IJ}S_{22}$ contain strictly positive entries. The blocks $a_{ij}g_{IJ}S_{11}$ and $a_{ij}g_{IJ}S_{21}$ can each be further divided into the form

$$\begin{bmatrix} 1 \\ \hline 2 \\ \hline 3 \end{bmatrix}$$

where blocks 1 and 3 contain all zeros, while block 2 has at least one positive entry in each row and column, with the sizes of these blocks being the same for all $a_{ij}g_{IJ}S_{11}$ and $a_{ij}g_{IJ}S_{21}$. The matrix $\mathbf{K}$ is thus reducible since it will have a number of rows which consist entirely of zeros, and the Perron-Frobenius theorem [13] is not directly applicable. We can however form a submatrix by deleting the all-zero rows and the associated columns, without changing the largest eigenvalue. Due to the structure in the blocks $a_{ij}g_{IJ}S_{11}$ and $a_{ij}g_{IJ}S_{21}$, it can be seen that this submatrix is primitive. The Perron-Frobenius theorem may then be applied to conclude that there is an eigenvalue with magnitude greater than any other, with a unique eigenvector (up to constant multiples). By the comments above, this will also hold for the original matrix $\mathbf{K}$.

In Table II we compare the numerical method with simulation results for a selection of different parameter values. Here we use $N = 1000$, and fix $\sigma^2 = 1$. Simulations results were again averaged over 10 runs, each run of length 100 000. It may be seen that there is very close agreement at the values considered here. However, for smaller values of σ, it has been observed (see Table III, also [8]) that the accuracy of the numerical method is not so good when using $N = 1000$. By the previous statements, the accuracy should increase with N, but due to memory limitations increasing N substantially is currently not feasible. In Section IV-E we will derive an approximation for the error probability which is more computationally tractible and provides close agreement with simulations for small σ.

TABLE II
SIMULATION AND ANALYTICAL COMPARISON OF NOISY ERROR PROBABILITIES, WITH $\sigma^2 = 1$

a_{12}	a_{21}	g_{12}	g_{21}	simulation	analytical
0.30	0.89	0.93	0.26	0.2241	0.2248
0.97	0.86	0.78	0.06	0.3882	0.3887
0.68	0.14	0.10	0.07	0.1415	0.1411
0.88	0.07	0.87	0.71	0.0658	0.0654
0.50	0.55	0.50	0.22	0.3793	0.3789

TABLE III
COMPARISON OF ERROR PROBABILITIES FOR VARIOUS VALUES OF σ

σ	simulation	approximation	method of IV-B	no packet loss
1.0	0.2126	0.2279	0.2123	0.1351
0.8	0.1754	0.1859	0.1763	0.0889
0.6	0.1350	0.1400	0.1540	0.0395
0.4	0.1064	0.1071	0.2193	0.0052
0.3	0.1024	0.1025	0.3520	0.00037

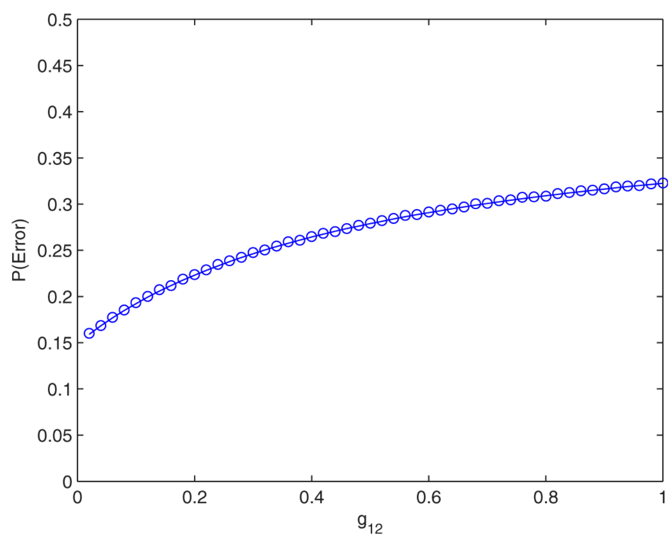


Fig. 6. Noisy probability of error for various g_{12}.

C. Numerical Studies

We now show some plots which are analogues of Figs. 1–5, with an additional noise term of variance $\sigma^2 = 1$. The simulation runs are of length one million for the first three graphs. In Fig. 6 we plot the simulated probability of error for 50 values of g_{12}, with $a_{12} = 0.4$, $a_{21} = 0.3$, and $g_{21} = 0.5$. In Fig. 7 we plot the simulated probability of error for 50 values of g_{21}, with $a_{12} = 0.4$, $a_{21} = 0.3$, and $g_{12} = 0.5$. In Fig. 8 we plot the simulated probability of error for 50 values of g_{12}, with $g_{21} = g_{12}$, $a_{12} = 0.1$ and $a_{21} = 0.1$. The solid lines represents the analytical calculation using $N = 200$. In Fig. 8 there is a slight but noticeable discrepancy between the simulation and analytical results at small values of g_{12}. This is due to the use of $N = 200$ in the numerical calculation, as it is found that increasing N to say 500 would give much closer agreement. For Figs. 9 and

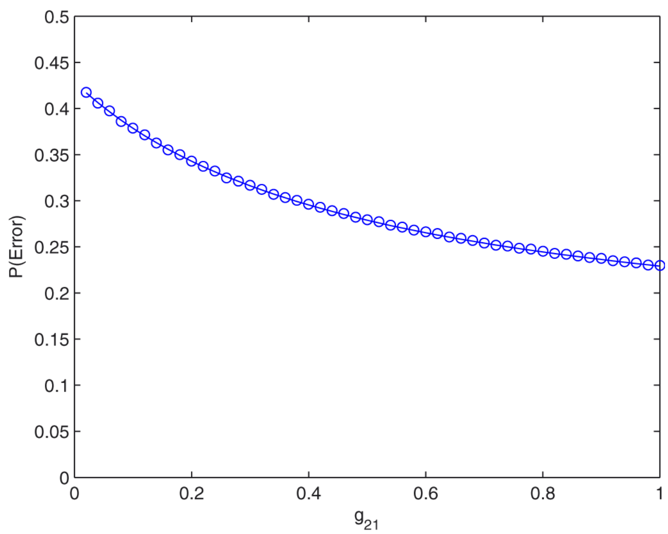


Fig. 7. Noisy probability of error for various g_{21}.

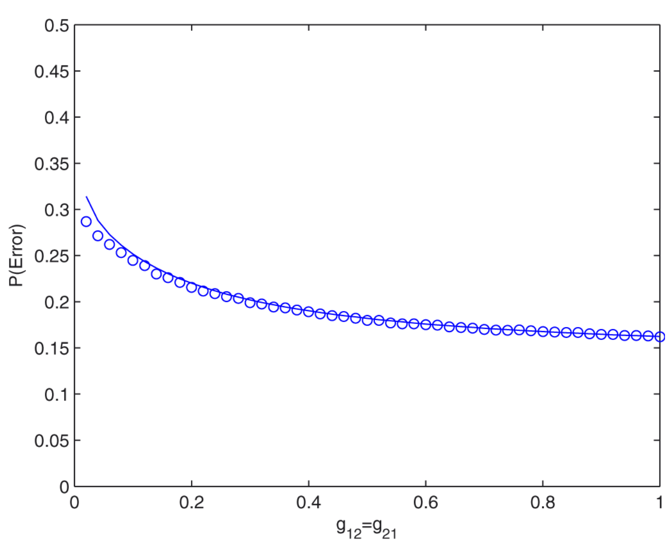


Fig. 8. Noisy probability of error for various g_{12}, with p_0 fixed.

10 the simulations runs are of length 100 000. In Fig. 9 we plot the simulated probability of error for 500 random values of p_0, with $a_{12} = a_{21} = 0.1$, i.e., A is symmetric. In Fig. 10 we plot the simulated probability of error for 500 random values of p_0, with $a_{12} = 0.7$ and $a_{21} = 0.3$, i.e., signal i.i.d.

Comparing these graphs with Figs. 1–5, we can see that there is a noise floor introduced which tends to shift the graphs upwards. The noiseless Figs. 1–5 also seem to cover a larger range of values as the parameters vary, one reason could be that there is a greater sensitivity to packet loss in the noiseless case, since losing packets is the only way in which one can make an error there.

Proving results similar to Theorems 1–3 in the noisy case seems to be very difficult. We will state these as conjectures

Conjectures

1) For fixed A, σ and g_{21}, the probability of error is monotonically increasing in g_{12}.

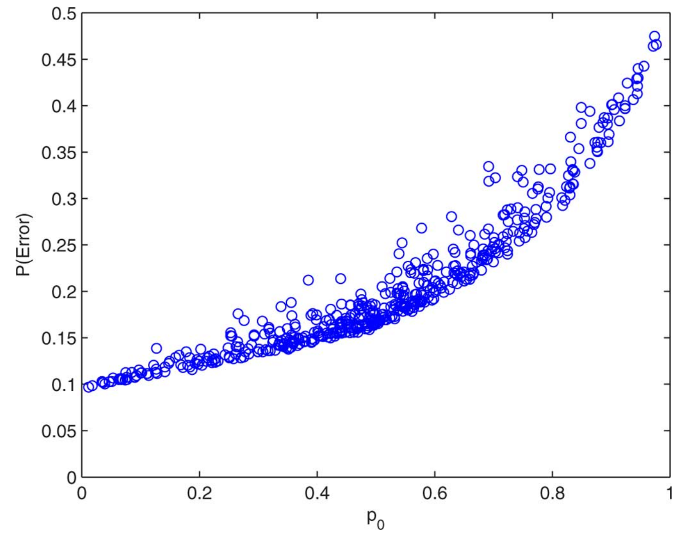


Fig. 9. Noisy probability of error for various p_0, symmetric A.

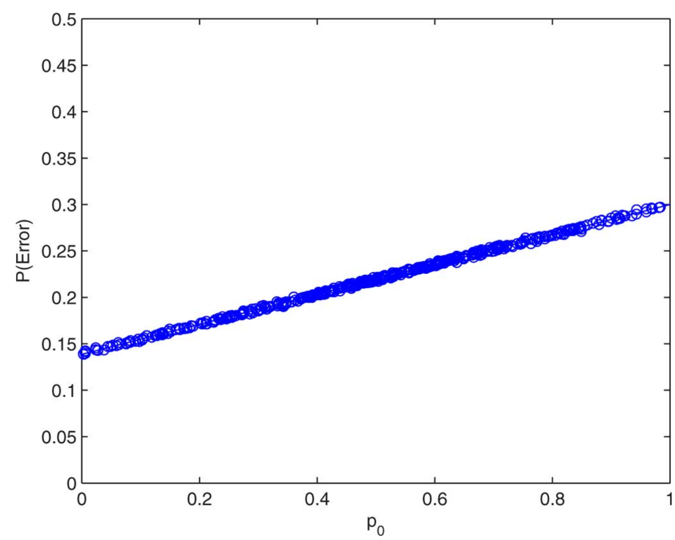


Fig. 10. Noisy probability of error for various p_0, signal i.i.d.

2) For fixed A, σ and g_{12}, the probability of error is monotonically decreasing in g_{21}.
3) Let p_0 be the stationary probability that a measurement is not received, i.e., $p_0 = P(Z_k = 1) = g_{12}/(g_{12} + g_{21})$. Then for fixed A, σ and p_0, the probability of error is nonincreasing with g_{12}. Furthermore, this probability of error converges, for fixed p_0 and as $g_{12} \to 0$, to $p_0 E_0 + (1 - p_0) E_1$, where E_1 is the probability of error obtained when measurements are always received, and $E_0 = \min(a_{12}/(a_{12} + a_{21}), a_{21}/(a_{12} + a_{21}))$ is the probability of error in the complete absence of observations.

The intuition behind conjecture 3 can be explained as follows: In [9], the exponential forgetting property of HMM filters is demonstrated. For slowly varying dynamics in the loss process, during the (usually) long periods where we always receive measurements, the probability of error there would be close to the error probability when we assume that measurements are always received. Similarly, for the periods when we don't obtain any observations, the error probability would be close to the error in the complete absence of observations. The overall error probability should then be able to be averaged over these two situations, giving $p_0 E_0 + (1-p_0)E_1$ as the conjectured limit.

D. Signal Is i.i.d.

Signals which are i.i.d. (or close to i.i.d.) are commonly encountered in digital communications. In Fig. 10, we saw that when the signal $\{X_k\}$ is i.i.d., there seems to be a linear relationship between p_0 and the probability of error even when the loss process is Markovian. In this section, we will show that this is indeed the case. If the signal is i.i.d., the A matrix has the form

$$A = \begin{bmatrix} a_{12} & a_{12} \\ 1-a_{12} & 1-a_{12} \end{bmatrix}$$

and the probability vector updates (6) become

$$\Pi_{k+1|k+1} = \begin{cases} \begin{bmatrix} a_{12} \\ 1-a_{12} \end{bmatrix}, & \gamma(Z_{k+1}) = 0 \\ \begin{bmatrix} \frac{a_{12} b_1(y_{k+1})}{a_{12} b_1(y_{k+1}) + (1-a_{12}) b_2(y_{k+1})} \\ \frac{(1-a_{12}) b_2(y_{k+1})}{a_{12} b_1(y_{k+1}) + (1-a_{12}) b_2(y_{k+1})} \end{bmatrix}, & \gamma(Z_{k+1}) = 1 \end{cases}$$

which depends on the values Y_{k+1} and Z_{k+1} but which importantly does not depend on values at previous times. Using this fact, an explicit expression for the error probability can then be derived quite easily.

Given $Z_k = 1$, the recursion (9) for q is just $q_k = 2a_{12} - 1$. If $a_{12} > 0.5$, then $q_k > 0$, and

$$P(\text{Error}|Z_k = 1) = P(X_k = 2, q_k > 0 | Z_k = 1) = 1 - a_{12}.$$

If $a_{12} \leq 0.5$, then $q_k \leq 0$, and so

$$P(\text{Error}|Z_k = 1) = P(X_k = 1, q_k \leq 0 | Z_k = 1) = a_{12}.$$

This can be written more compactly as $P(\text{Error}|Z_k = 1) = \min(a_{12}, 1 - a_{12})$. Next, we have

$$\begin{aligned} &P(\text{Error}|X_k = 1, Z_k = 2) \\ &\quad = P\left(a_{12} b_1(Y_k) \leq (1-a_{12}) b_2(Y_k) | X_k = 1\right) \\ &\quad = P\left(a_{12} \exp\left(\frac{-(Y_k+1)^2}{2\sigma^2}\right) \right. \\ &\qquad\qquad \left. \leq (1-a_{12}) \exp\left(\frac{-(Y_k-1)^2}{2\sigma^2}\right) | X_k = 1\right) \\ &\quad = P\left(Y_k \geq -\frac{\sigma^2}{2} \ln\left(\frac{1-a_{12}}{a_{12}}\right) | X_k = 1\right) \\ &\quad = P\left(v_k \geq 1 - \frac{\sigma^2}{2} \ln\left(\frac{1-a_{12}}{a_{12}}\right)\right) \\ &\quad = P\left(v'_k \leq -\frac{1}{\sigma}\left(1 - \frac{\sigma^2}{2} \ln\left(\frac{1-a_{12}}{a_{12}}\right)\right)\right) \end{aligned}$$

where v'_k is $N(0,1)$. Similarly, $P(\text{Error}|X_k = 2, Z_k = 2) = P(v'_k \leq -(1/\sigma)(1 + (\sigma^2/2)\ln((1-a_{12})/a_{12})))$, and so

$$\begin{aligned} &P(\text{Error}|Z_k = 2) \\ &\quad = P(\text{Error}|X_k = 1, Z_k = 2) P(X_k = 1) \\ &\qquad + P(\text{Error}|X_k = 2, Z_k = 2) P(X_k = 2) \\ &\quad = a_{12} P\left(v'_k \leq -\frac{1}{\sigma}\left(1 - \frac{\sigma^2}{2} \ln\left(\frac{1-a_{12}}{a_{12}}\right)\right)\right) \\ &\qquad + (1-a_{12}) P\left(v'_k \leq -\frac{1}{\sigma}\left(1 + \frac{\sigma^2}{2} \ln\left(\frac{1-a_{12}}{a_{12}}\right)\right)\right). \end{aligned} \tag{15}$$

The probability of error in the case of i.i.d. data is therefore

$$P(\text{Error}|Z_k = 1) P(Z_k = 1) + P(\text{Error}|Z_k = 2) P(Z_k = 2)$$

where we use the expressions for $P(\text{Error}|Z_k = 1)$ and (15), together with the stationary probabilities for the loss process $P(Z_k = 1) = g_{12}/(g_{12} + g_{21})$ and $P(Z_k = 2) = g_{21}/(g_{12} + g_{21})$. Since $P(Z_k = 2) = 1 - P(Z_k = 1)$, this probability of error is linear in $P(Z_k = 1)$, for fixed A and σ. The solid line in Fig. 10 is a plot of this linear expression, which can be seen to coincide with the simulation results.

E. High SNR Approximation

Here we will derive an approximate expression for the error in the high SNR, or small σ regime. When σ is small, we can see from (8) that for $Y_{k+1} > 0$, $q_{k+1} \approx -1$, and for $Y_{k+1} < 0$, $q_{k+1} \approx 1$. So at high SNR, the probability updates can be approximated by the simpler suboptimal scheme

$$\Pi_{k+1|k+1} = \begin{cases} A\Pi_{k|k}, & \gamma(Z_{k+1}) = 0 \\ [1\ \ 0]^T, & \gamma(Z_{k+1}) = 1,\ Y_{k+1} \in D_1 \\ [0\ \ 1]^T, & \gamma(Z_{k+1}) = 1,\ Y_{k+1} \in D_2. \end{cases}$$

where we have defined the sets $D_1 \equiv (-\infty, 0]$ and $D_2 \equiv (0, \infty)$. To derive the probability of error using such a scheme, first note that $P(Y_k \in D_2 | X_k = 1, Z_k = 2) = P(v'_k > (1/\sigma))$ and $P(Y_k \in D_1 | X_k = 2, Z_k = 2) = P(v'_k > (1/\sigma))$, where v'_k is $N(0,1)$. Using D_m^c to denote the complement of D_m,

$$\begin{aligned} &P(\text{Error}, Z_k = 2) \\ &\quad = P\left(Y_k \in D_1^c, X_k = 1, Z_k = 2\right) \\ &\qquad + P\left(Y_k \in D_2^c, X_k = 2, Z_k = 2\right) \\ &\quad = P\left(Y_k \in D_1^c | X_k = 1, Z_k = 2\right) P(X_k = 1) P(Z_k = 2) \\ &\qquad + P\left(Y_k \in D_2^c | X_k = 2, Z_k = 2\right) P(X_k = 2) P(Z_k = 2) \\ &\quad = P\left(v'_k > \frac{1}{\sigma}\right) P(Z_k = 2) \end{aligned}$$

where we have again used the independence of $\{X_k\}$ and $\{Z_k\}$.

The derivation of the term $P(\text{Error}, Z_k = 1)$ is similar to the noiseless error probability derived in Section III-A, we will use the same notation and point out the main differences. We can still write $P(\text{Error}, Z_k = 1) = \sum_{n=1}^{\infty} p(n)$ where $p(n) \equiv$

$P(\text{Error}, Z_{k-n} = 2, Z_{k-n+1} = 1, Z_{k-n+2} = 1, \ldots, Z_k = 1) \equiv P(\text{Error}, Z)$, but now

$$\begin{aligned} p(n) &= P\left(Z, X_k = 1, \arg\max_j \Pi^j_{1,n} \neq 1\right) \\ &\quad + P\left(Z, X_k = 2, \arg\max_j \Pi^j_{1,n} \neq 2\right) \\ &\quad + P\left(Z, X_k = 1, \arg\max_j \Pi^j_{2,n} \neq 1\right) \\ &\quad + P\left(Z, X_k = 2, \arg\max_j \Pi^j_{2,n} \neq 2\right) \\ &= P\Big(Z, X_k = 1, \arg\max_j \Pi^j_{1,n} \neq 1, \\ &\qquad Y_{k-n} \in D_1, X_{k-n} = 1\Big) \\ &\quad + P\Big(Z, X_k = 1, \arg\max_j \Pi^j_{1,n} \neq 1, \\ &\qquad Y_{k-n} \in D_1, X_{k-n} = 2\Big) + \ldots + \\ &\quad + P\Big(Z, X_k = 2, \arg\max_j \Pi^j_{2,n} \neq 2, \\ &\qquad Y_{k-n} \in D_2, X_{k-n} = 1\Big) \\ &\quad + P\Big(Z, X_k = 2, \arg\max_j \Pi^j_{2,n} \neq 2, \\ &\qquad Y_{k-n} \in D_2, X_{k-n} = 2\Big) \\ &= \frac{g_{21}g_{12}}{g_{12}+g_{21}}(1-g_{21})^{n-1} \\ &\quad \times \sum_{r,s,t=1}^{2} \Big[P(X_{k-n} = r)a^{(n)}_{sr} \times I_{\operatorname{argmax}_j \Pi^j_{t,n} \neq s} \\ &\qquad \times P(Y_{k-n} \in D_t | X_{k-n} = r, Z_{k-n} = 2)\Big]. \end{aligned}$$

As in the noiseless case, we can write this in a more explicit form as shown in (16) at the bottom of the page. The probability of error of this scheme is, therefore

$$\begin{aligned} P(\text{Error}) &= P(\text{Error}, Z_k = 1) + P(\text{Error}, Z_k = 2) \\ &= \sum_{n=1}^{\infty} p(n) + P\left(v'_k > \frac{1}{\sigma}\right)\frac{g_{21}}{g_{12}+g_{21}}. \end{aligned}$$

In Table III we compare simulation results of the optimal filter together with the "exact" analytical calculation of Section IV-B, and the suboptimal approximation just derived. We use $a_{12} = 0.2$, $a_{21} = 0.3$, $g_{12} = 0.5$, $g_{21} = 0.8$ and various values of σ. The computation using the method of Section IV-B was done with $N = 1000$. For a further comparison, in the final column we also include simulation results when there is no packet loss. The simulation runs are of length 10 million.

Firstly, we can see that for values of σ smaller than approximately 0.8, the numerical method of Section IV-B does not give accurate results when using $N = 1000$, in fact the accuracy worsens the smaller σ is. Improving the accuracy would involve increasing N substantially, which in turn increases the computation time and memory requirements dramatically. We can also see that the suboptimal expression gives very good agreement with simulations for small values of σ, moreover it can be computed very easily with current computer algebra software. Indeed, the noiseless probability of error can be computed to be 0.1022, so that even for $\sigma = 0.3$, the difference between the noisy and noiseless error probabilities are almost negligible. Since our approximation (16) converges to the noiseless expression (4) as $\sigma \to 0$, this is one reason why the approximation performs so well at high SNR.

Comparing the noiseless error probability of 0.1022, the error probabilities with no packet loss and the error probabilities with both packet loss and noise, it appears that for σ smaller than around 0.4–0.5, the packet loss starts to dominate for this example. In general, the σ value where the packet loss term starts to dominate will depend on other parameters such as g_{12} and g_{21}, and is an issue that requires further study. However through our numerical investigations, we have found that σ around 0.4–0.5 seems to be a reasonable figure for most (randomly generated) sets of parameter values.

F. Multiple States—High SNR

For the noisy case with multiple states (and no packet loss), asymptotic results for the error performance of slow Markov chains exist in the literature, e.g., [6], [7], but general expressions for arbitrary Markov chains are not known. In this section, we will treat arbitrary Markov chains with packet loss at high SNR. We will choose the signal levels $h(X_k)$ to be of M-ary Pulse-amplitude-modulation (PAM) type. Without loss of generality, we let these levels be situated at $1, 3, \ldots, 2M-1$, i.e.

$$h(m) = 2m - 1, \quad m = 1, 2, \ldots, M.$$

$$p(n) = \begin{cases} \frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{11} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{12}\right], & q^1_n \leq 0, q^2_n \leq 0 \\ P\left(v'_k < \frac{1}{\sigma}\right)\left(\frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{11} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{22}\right]\right) & \\ +P\left(v'_k > \frac{1}{\sigma}\right)\left(\frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{21} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{12}\right]\right), & q^1_n \leq 0, q^2_n > 0 \\ P\left(v'_k < \frac{1}{\sigma}\right)\left(\frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{21} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{12}\right]\right) & \\ +P\left(v'_k > \frac{1}{\sigma}\right)\left(\frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{11} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{22}\right]\right), & q^1_n > 0, q^2_n \leq 0 \\ \frac{g_{21}}{g_{12}+g_{21}} g_{12}(1-g_{21})^{n-1}\left[\frac{a_{12}}{a_{12}+a_{21}}a^{(n)}_{21} + \frac{a_{21}}{a_{12}+a_{21}}a^{(n)}_{22}\right], & q^1_n > 0, q^2_n > 0. \end{cases} \tag{16}$$

TABLE IV
COMPARISON OF NOISY ERROR PROBABILITIES FOR VARIOUS VALUES OF σ, 3-STATE EXAMPLE

σ	simulation	approximation
1.0	0.2905	0.3276
0.8	0.2444	0.2691
0.6	0.1948	0.2054
0.4	0.1580	0.1595
0.3	0.1524	0.1531

Define the sets

$$D_m = \begin{cases} (-\infty, 2], & m = 1 \\ (2m-2, 2m], & m = 2, 3, \dots, M-1 \\ (2m-2, \infty), & m = M. \end{cases}$$

The optimal way to update the probabilities is the obvious generalization of (6), but which appears to be very difficult to analyze. Motivated by the high SNR approximation in the 2-state case, consider the following suboptimal scheme:

$$\Pi_{k+1|k+1} = \begin{cases} A\Pi_{k|k}, & \gamma(Z_{k+1}) = 0 \\ [1\ 0\ \dots\ 0]^T, & \gamma(Z_{k+1}) = 1, Y_{k+1} \in D_1 \\ [0\ 1\ \dots\ 0]^T, & \gamma(Z_{k+1}) = 1, Y_{k+1} \in D_2 \\ \vdots & \vdots \\ [0\ 0\ \dots\ 1]^T, & \gamma(Z_{k+1}) = 1, Y_{k+1} \in D_M. \end{cases}$$

Similar to the two-state case in Section IV-E, we can derive

$$\begin{aligned} &P(\text{Error}, Z_k = 2) \\ &\quad = P\left(Y_k \in D_1^c | X_k = 1, Z_k = 2\right) P(X_k = 1) P(Z_k = 2) \\ &\qquad + \dots + P\left(Y_k \in D_M^c | X_k = M, Z_k = 2\right) \\ &\qquad \times P(X_k = M) P(Z_k = 2) \\ &\quad = P\left(v_k' > \frac{1}{\sigma}\right) P(Z_k = 2) \\ &\qquad \times [2 - P(X_k = 1) - P(X_k = M)] \end{aligned}$$

and

$$\begin{aligned} p(n) = \frac{g_{21} g_{12}}{g_{12} + g_{21}} (1 - g_{21})^{n-1} \sum_{r,s,t=1}^{M} \Big[P(X_{k-n} = r) a_{sr}^{(n)} \\ \times I_{\operatorname{argmax}_j \Pi_{t,n}^j \neq s} P(Y_{k-n} \in D_t | X_{k-n} = r, Z_{k-n} = 2) \Big] \end{aligned}$$

with $P(\text{Error}, Z_k = 1) = \sum_{n=1}^{\infty} p(n)$. As in the noiseless multi-state case, more explicit expressions would be very complicated to write down in general.

In Table IV we compare this approximation with simulation results of the optimal filter, using the 3-state example of Section III-E with an additional noise term. We can again see that the suboptimal expression gives good agreement with simulations for small values of σ.

V. CONCLUSION

In this paper, we have derived analytical expressions for the error probability of HMM filters in the presence of Markovian packet losses, with emphasis on two-state Markov chains. Performance analysis of such systems are important when operating in unreliable environments such as wireless sensor networks. A number of relationships between the error probabilities and the parameters of the loss process have been shown via numerical studies, and theoretical justification has also been obtained in some cases.

APPENDIX

For convenience in our proofs, let us define

$$s(n) \equiv \frac{g_{21} g_{12}}{g_{12} + g_{21}} (1 - g_{21})^{n-1}$$

$$t(n) \equiv \begin{cases} \frac{a_{12}(a_{12}+a_{21})}{(a_{12}+a_{21})^2}, & q_n^1 \le 0, q_n^2 \le 0 \\ \frac{a_{12}^2 + 2a_{12}a_{21}(1-a_{12}-a_{21})^n + a_{21}^2}{(a_{12}+a_{21})^2}, & q_n^1 \le 0, q_n^2 > 0 \\ \frac{2a_{12}a_{21} - 2a_{12}a_{21}(1-a_{12}-a_{21})^n}{(a_{12}+a_{21})^2}, & q_n^1 > 0, q_n^2 \le 0 \\ \frac{a_{21}(a_{12}+a_{21})}{(a_{12}+a_{21})^2}, & q_n^1 > 0, q_n^2 > 0 \end{cases} \tag{17}$$

so that $p(n) = s(n)t(n)$.

A. *Proof of Theorem 1*

Proof:

$$\frac{dP(\text{Error})}{dg_{12}} = \sum_{n=1}^{\infty} \frac{ds(n)}{dg_{12}} t(n) = \sum_{n=1}^{\infty} \frac{g_{21}^2 (1 - g_{21})^{n-1}}{(g_{12} + g_{21})^2} t(n).$$

Term-by-term differentiation of the infinite series can be justified by using the Weierstrass M-test, see, e.g., [16]. It is easy to see from (17) or (4) that the $t(n)$ terms are all positive. As each term in the sum is greater than zero, $(dP(\text{Error})/dg_{12}) > 0$. ■

B. *Proof of Theorem 2*

While the statement of Theorem 2 looks similar to that of Theorem 1, the proof is not as straightforward. Before we prove this, we need the following technical result.

Lemma 1: The terms $t(n)$ given by (17) form a nondecreasing sequence.

Proof: We will only outline the method of proof. One can assume throughout that $a_{12} \le a_{21}$, as the arguments when $a_{12} > a_{21}$ are almost identical. We consider the situations $a_{12} + a_{21} < 1$, $a_{12} + a_{21} > 1$ and $a_{12} + a_{21} = 1$ separately. Each of the situations $a_{12} + a_{21} < 1$ and $a_{12} + a_{21} > 1$ is further divided into the four cases: 1) $q_n^1 \le 0$ and $q_n^2 \le 0$, 2) $q_n^1 \le 0$ and $q_n^2 > 0$, 3) $q_n^1 > 0$ and $q_n^2 \le 0$, 4) $q_n^1 > 0$ and $q_n^2 > 0$ (some of these cases may not be possible). It is then a tedious but relatively straightforward verification that $t(n+1) \ge t(n)$ holds in all the possible cases, using the definitions (3) and (17). The remaining situation with $a_{12} + a_{21} = 1$ is straightforward, only the case $q_n^1 \le 0$ and $q_n^2 \le 0$ can occur, and $t(n) = (a_{12}/(a_{12} + a_{21})), \forall n$. ■

Now we can prove Theorem 2.

Proof: First

$$\frac{dP(\text{Error})}{dg_{21}} = \sum_{n=1}^{\infty} \frac{ds(n)}{dg_{21}} t(n) = \sum_{n=1}^{\infty} -\frac{g_{12}(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right]}{(g_{12}+g_{21})^2} \times t(n). \tag{18}$$

Showing that this quantity is negative is equivalent to showing that

$$\sum_{n=1}^{\infty} -(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right] t(n) < 0.$$

Unlike the proof of Theorem 1, not every term in the summation here is negative. However, it is not difficult to show that for $m \equiv \lfloor (g_{21}^2+g_{12})/(g_{21}^2+g_{12}g_{21}) \rfloor$, the first m terms will be positive, while the rest will be negative. We may then use Lemma 1 to obtain the bound

$$\sum_{n=1}^{\infty} -(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right] t(n) \le t(m)\sum_{n=1}^{\infty} -(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right].$$

To complete the proof, we note the following closed form expression, which can be verified using induction:

$$\sum_{n=1}^{k} -(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right] = -1+(1-g_{21})^{k-1}\left[1+g_{21}(k-1)+g_{12}k\right].$$

(Note that this expression also allows us to apply the Weierstrass M-test to justify the term-by-term differentiation (18).) Hence, $\sum_{n=1}^{\infty} -(1-g_{21})^{n-2}[g_{21}^2(n-1)+g_{12}(ng_{21}-1)] = -1$ and, therefore

$$\sum_{n=1}^{\infty} -(1-g_{21})^{n-2}\left[g_{21}^2(n-1)+g_{12}(ng_{21}-1)\right] t(n) \le -t(m) < 0.$$

■

C. Proof of Theorem 3

Proof: That E_0 actually is the error probability in the complete absence of observations is not difficult to show. For example, one can use the fact that A^n converges to a rank 1 matrix (with the stationary probabilities in the columns) as $n \to \infty$, so that without observations, one would choose the state estimate which on average is more likely to occur.

(i) The proof of this part is similar to that of Theorem 2. Since $P(\text{Error}) = \sum_{n=1}^{\infty} p_0 g_{21}(1-g_{21})^{n-1} t(n)$, we have

$$\frac{dP(\text{Error})}{dg_{21}} = \sum_{n=1}^{\infty} -p_0(1-g_{21})^{n-2}(ng_{12}-1)t(n).$$

It can be easily seen that for $m \equiv \lfloor 1/g_{21} \rfloor$, the first m terms in the series will be positive, while the rest will be negative. Using Lemma 1, we obtain the bound

$$\sum_{n=1}^{\infty} -p_0(1-g_{21})^{n-2}(ng_{12}-1)t(n) \le t(m)\sum_{n=1}^{\infty} -p_0(1-g_{21})^{n-2}(ng_{12}-1).$$

We have the following closed form expression:

$$\sum_{n=1}^{k} -p_0(1-g_{21})^{n-2}(ng_{12}-1) = (1-g_{21})^{k-1}p_0$$

so $\sum_{n=1}^{\infty} -p_0(1-g_{21})^{n-2}(ng_{12}-1) = 0$, and therefore

$$\sum_{n=1}^{\infty} -p_0(1-g_{21})^{n-2}(ng_{12}-1)t(n) \le 0.$$

(ii) We consider the cases $a_{12} \neq a_{21}$ and $a_{12} = a_{21}$ separately. First assume that $a_{12} \neq a_{21}$. From (3) it can be seen that there exists an N such that either $q_n^1 > 0, q_n^2 > 0, \forall n > N$ (when $a_{12} > a_{21}$), or $q_n^1 < 0, q_n^2 < 0, \forall n > N$ (when $a_{12} < a_{21}$). Hence, $t(n) = E_0, \forall n > N$ and

$$P(\text{Error}) = \sum_{n=1}^{N} p_0 g_{21}(1-g_{21})^{n-1}t(n) + E_0 \sum_{n=N+1}^{\infty} p_0 g_{21}(1-g_{21})^{n-1}.$$

Applying Lemma 1, we can obtain the bounds

$$t(1)\sum_{n=1}^{N} p_0 g_{21}(1-g_{21})^{n-1} + E_0 \sum_{n=N+1}^{\infty} p_0 g_{21}(1-g_{21})^{n-1} \le P(\text{Error}) \le E_0 \sum_{n=1}^{\infty} p_0 g_{21}(1-g_{21})^{n-1}$$

or $t(1)[p_0 - (1-g_{21})^N p_0] + E_0(1-g_{21})^N p_0 \le P(\text{Error}) \le E_0 p_0$. Taking the limit as $g_{21} \to 0$ then gives the result for $a_{12} \neq a_{21}$.

Now assume that $a_{12} = a_{21}$. We further divide into three cases.

1) For $a_{12} = 0.5$, we have $q_n^1 \le 0, q_n^2 \le 0, \forall n$, and

$$P(\text{Error}) = p_0 \times \frac{1}{2}\sum_{n=1}^{\infty} g_{21}(1-g_{21})^{n-1} = \frac{p_0}{2} = p_0 E_0$$

irrespective of g_{21}.

2) For $a_{12} < 0.5$, we have $q_n^1 > 0, q_n^2 < 0, \forall n$, so

$$P(\text{Error}) = p_0 \sum_{n=1}^{\infty} g_{21}(1-g_{21})^{n-1} \times \frac{1}{2}\left[1-(1-2a_{12})^n\right] = \frac{a_{12}p_0}{2a_{12} - 2a_{12}g_{21} + g_{21}}$$

which converges to $p_0/2$ as $g_{21} \to 0$.

3) For $a_{12} > 0.5$, we have $q_n^1 < 0$, $q_n^2 > 0$ for n odd, and $q_n^1 > 0$, $q_n^2 < 0$ for n even, so

$$\begin{aligned} P(\text{Error}) &= p_0 \sum_{m=1}^{\infty} g_{21}(1-g_{21})^{2m-2} \times \frac{1}{2}\left[1+(1-2a_{12})^{2m-1}\right] \\ &\quad + p_0 \sum_{m=1}^{\infty} g_{21}(1-g_{21})^{2m-1} \times \frac{1}{2}\left[1-(1-2a_{12})^{2m}\right] \\ &= p_0 \sum_{n=1}^{\infty} g_{21}(1-g_{21})^{n-1} \times \frac{1}{2}\left[1-(2a_{12}-1)^{n}\right] \\ &= \frac{(1-a_{12})p_0}{2(1-a_{12})-2(1-a_{12})g_{21}+g_{21}} \end{aligned}$$

which also converges to $p_0/2$ as $g_{21} \to 0$. ■

ACKNOWLEDGMENT

The authors thank the reviewers for their comments and suggestions, which significantly improved the quality of the manuscript.

REFERENCES

[1] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[2] A. Tiwari, M. Jun, D. E. Jeffcoat, and R. M. Murray, "Analysis of dynamic sensor coverage problem using Kalman filters for estimation," in *Proc. IFAC World Congress*, Prague, Czech Republic, Jul. 2005.

[3] S. C. Smith and P. Seiler, "Estimation with lossy measurements: jump estimators for jump systems," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2163–2171, Dec. 2003.

[4] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518–1569, Jun. 2002.

[5] W. M. Wonham, "Some applications of stochastic differential equations to optimal nonlinear filtering," *SIAM J. Contr.* , vol. 2, no. 3, pp. 347–369, 1965.

[6] G. K. Golubev, "On filtering for a hidden Markov chain under square performance criterion," *Probl. Inf. Transmiss.*, vol. 36, no. 3, pp. 213–219, 2000.

[7] R. Khasminskii and O. Zeitouni, "Asymptotic filtering for finite state Markov chains," *Stochastic Processes Their Applicat.*, vol. 63, pp. 1–10, 1996.

[8] 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.

[9] L. Shue, B. D. O. Anderson, and S. Dey, "Exponential stability of filters and smoothers for hidden Markov models," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2180–2194, Aug. 1998.

[10] F. G. Tricomi, *Integral Equations*. New York: Interscience, 1957.

[11] W. Hackbusch, *Integral Equations: Theory and Numerical Treatment*. Basel, Switzerland: Birkhauser Verlag, 1995.

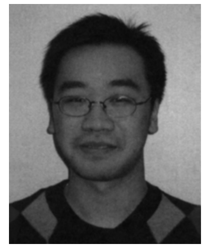
[12] K. E. Atkinson, *The Numerical Solution of Integral Equations of the Second Kind*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[13] E. Seneta, *Non-Negative Matrices and Markov Chains*, 2nd ed. New York: Springer-Verlag, 1981.

[14] A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, 2nd ed. New York: McGraw-Hill, 1978.

[15] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*. San Diego, CA: Academic , 1990.

[16] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.



Alex S. C. Leong (S'03–M'03) was born in Macau in 1980. He received the B.S. degree in mathematics and the B.E. degree in electrical engineering from the University of Melbourne, Melbourne, Australia, in 2003.

He is currently pursuing the Ph.D. degree in the Department of Electrical and Electronic Engineering, University of Melbourne. His research interests include statistical signal processing and wireless communications.

Mr. Leong was the recipient of the L. R. East Medal from Engineers Australia in 2003.

Subhrakanti Dey (S'94–M'96–SM'06) was born in Calcutta, India, in 1968. He received the Bachelor of Technology and Master of Technology degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, in 1991 and 1993, respectively. He received the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, in 1996.

He is currently appointed an Associate Professor with the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Australia, where he has been since 2000. During 1995–2000, he was a Postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University (ANU), and a Postdoctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include signal processing for telecommunications, wireless communications and networks, networked control systems, and statistical and adaptive signal processing. He is currently an academic staff member of the ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, University of Melbourne. He has held numerous research grants and authored and coauthored more than 60 internationally refereed papers.

Dr. Dey is an Associate Editor for the *Systems and Control Letters* and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He has also been on several major international conference program committees.

Jamie S. Evans (S'93–M'98) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, Newcastle, Australia, in 1992 and 1993, respectively, and the M.S. and Ph.D. degrees from the University of Melbourne, Australia, in 1996 and 1998, respectively, both in electrical engineering.

From March 1998 to June 1999, he was a Visiting Researcher with the Department of Electrical Engineering and Computer Science, University of California, Berkeley. He returned to Australia to take up a position as Lecturer at the University of Sydney, Sydney, Australia, where he stayed until July 2001. Since that time, he has been with the Department of Electrical and Electronic Engineering, University of Melbourne, where he is now an Associate Professor. His research interests are in communications theory, information theory, and statistical signal processing with current focus on wireless communications networks.

Dr. Evans was the recipient of the University Medal of the University of Newcastle and the Chancellor's Prize for Excellence for his Ph.D. dissertation.