Error Analysis for ML Sequence Detection in ISI Channels With Gauss Markov Noise

Naveen Kumar, Aditya Ramamoorthy, and Murti V. Salapaka

***Abstract*—We consider error probability analysis for a class of ISI channels with data-dependent Gauss–Markov noise and propose an efficiently computable BER upper bound that is tight in the high SNR regime.**

***Index Terms*—Inter-symbol-interference channels, flowgraph techniques, error-state diagram, performance analysis.**

I. Introduction

Inter-Symbol-Interference (ISI) channels with data-dependent noise are used for read channel modeling in many different domains. For instance, it is well-recognized that in magnetic recording, the statistics of percolation and nonlinear effects between transitions across adjacent magnetic domains [1], [2] result in noise that exhibits data-dependent time-correlation and significant ISI. More recently, it has been shown that nanotechnology based probe storage systems can be modeled in a similar manner [3]. Significant research work has addressed the design of high-performance, low-complexity detectors for such models, and has translated into gains in practical systems.

In a celebrated paper [4], Forney demonstrated a maximum likelihood sequence detection (MLSD) algorithm for the case of finite ISI channels with memoryless noise. Following this, the probability of error was analyzed using flowgraph techniques [5]–[7]. The work of Kavcic and Moura [8] considered finite ISI channels with data-dependent finite memory noise modeled by a Gauss-Markov process, and developed a ML sequence detection algorithm for them. However, estimating the probability of error for this algorithm turns out to be significantly more complicated. In [8] (Section V in [8]), certain approaches for computing estimates of the detector performance were presented (see also [9], [10]). However, these techniques are not based on flowgraph techniques. In particular, they require an enumeration of all the dominant error events of the detector, i.e., error events of relevant lengths. This step typically requires long simulations at the (high) signal-to-noise ratios of interest. Following the enumeration, an estimate can be obtained by first determining the probability of each error event and then computing a union bound over the placement of the error events at various locations in all possible transmitted blocks. The accuracy of this technique is contingent upon including all the relevant error events; in particular, there is no guarantee that it is an upper bound. It is important to note that an analytical technique for estimating detector performance is of great value, especially in the high SNR regime, where gathering the error events can be quite time-consuming.

There are two properties of ISI channels with memoryless AWGN that facilitate the application of flowgraph techniques for performance evaluation that we now discuss. Let $\bar{S}$ and $\hat{\bar{S}}$ be the true and estimated state sequences respectively. Let $P(\hat{\bar{S}} \mid \bar{S})$ denote the probability that the detector prefers $\hat{\bar{S}}$ to $\bar{S}$ (henceforth the pair-wise error probability (PEP)). Firstly, $P(\hat{\bar{S}} \mid \bar{S})$ can be upper-bounded by $P(\hat{\bar{S}} \mid \bar{S}) \leq \Pi_{k=0}^{N-1} h(\hat{S}_k^{k-1}, S_k^{k-1})$ where h is a function of current state and previous decoded states $\hat{S}_k^{k-1} = (\hat{S}_{k-1}, \hat{S}_k)$ and actual states $S_k^{k-1} = (S_{k-1}, S_k)$. Secondly, it can be seen that the PEP is symmetric, i.e., $P(\hat{\bar{S}} \mid \bar{S}) = P(\bar{S} \mid \hat{\bar{S}})$. These properties allow the application of error state diagrams for finding an asymptotically tight upper bound on the BER for such channels [7].

In contrast, for the ISI channel with data-dependent Gauss-Markov noise (considered in [8]), neither of these properties hold. The signal dependent and time-correlated noise makes the PEP asymmetric. Further the PEP does not factorize in a suitable manner as required for the application of flowgraph techniques. This makes the estimation of BER for such channels, quite challenging.

Main Contributions: In this correspondence, we consider a subset of the class of channels considered in [8], that continue to exhibit data-dependent noise. For these channels we use Gallager bounding techniques [11] to determine a PEP upper bound that can be expressed as a product of functions depending on current and previous states in the (incorrect) decoded sequence and the (correct) transmitted sequence. Based on this, we present an analytical technique for determining an upper bound on the BER using an error state diagram. Simulations results show that our proposed bound is tight in the high SNR regime.

Our work is a journal version of our conference article [12]. Specifically, in this work we demonstrate that an error state diagram can be constructed for an efficient computation of the transfer function for error-bound [7]. In the channels under consideration, there are two sources of memory: the inter-symbol-interference of length I and the noise memory of length L (we define these formally in the later discussion). In our conference article [12], we used the product state diagram for the bound computation which involves working with a matrix of order $M^2 \times M^2$ where $M = 2^{L+I}$. The key complexity in bound computation comes from the inversion of a matrix of order $(M^2 - M) \times (M^2 - M)$ which appears in the expression of the transfer function. The product state method is not desirable for bound computation for large values of M. In contrast, the reduced error state diagram has much fewer states. Let each state in the error state diagram take values from the set $\{0, +, -\}$; thus, there are 3^{L+I} number of states. The observation on the bound of PEP being symmetric can significantly reduce the number of states. In the symmetric PEP case, $(3^{L+I} - 1)/2$ pairs of states, which are equivalent, can be merged into a single equivalent state. This leads to a reduced error state diagram with $(3^{L+I} + 1)/2$ number of states. It is much easier to find the transfer function from the reduced error state diagram. For instance, even when $L + I = 4$, the product-state diagram has 256 states compared to 41 states in the error-state diagram (we refer the reader to [7] for further information on the reduced state diagram). The present manuscript presents an error state diagram based bounding technique, includes all the proofs and a comprehensive set of simulation results that show the effectiveness of the method.

The paper is organized as follows. Section II introduces the channel model and describes the corresponding Viterbi decoding algorithm. Section III presents an upper bound on the detector BER. Section IV demonstrates simulation results that confirm the analytical bounds.

Manuscript received January 26, 2012; revised May 17, 2012 and September 17, 2012; accepted April 26, 2013. Date of publication May 13, 2013; date of current version June 24, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Peter. J. Schreier. This work was supported in part by NSF grants CCF-1116322 and CCF-1149860. The material in this work was presented in part at the IEEE Global Telecommunication Conference (IEEE GLOBECOM) 2010.

N. Kumar was with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011USA. He is now with SK Hynix Memory Solutions, Inc., San Jose, CA 95134 USA (e-mail: nkumar@skhms.com).

A. Ramamoorthy is with Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: adityar@iastate.edu).

M. V. Salapaka is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: murtis@umn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2262274

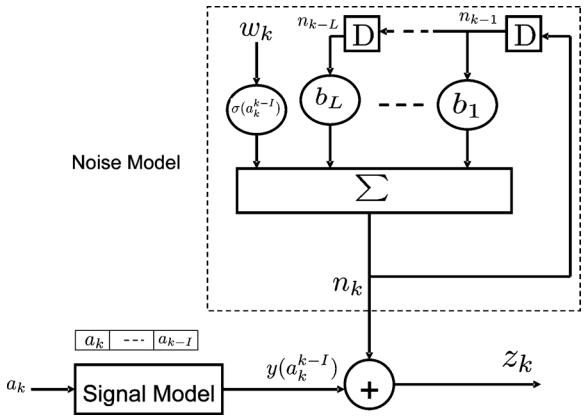

Fig. 1. Channel model with Gauss-Markov noise.

Section V summarizes the main findings of this paper and outlines future work.

II. Channel Model and Viterbi Detector

In the discussion below, the notation z_{k2}^{k1} denotes the column vector given by $z_{k2}^{k1} = [z_{k1}, \ldots, z_{k2}]^T$ where $k_1 \leq k_2$, and the notation $f(\cdot \,|\, \cdot)$ denotes a conditional pdf. The precise pdf under consideration will be evident from the context of the discussion.

A. Channel Model

Let a_k denote the kth source bit that is equally likely to be 0 or 1. The channel output shown in Fig. 1 with intersymbol interference (ISI) of length I is given by,

$$z_k = y\left(a_k^{k-I}\right) + n_k, \tag{1}$$

where $y(a_k^{k-I})$ is the noiseless channel output dependent only on the latest $I+1$ transmitted bits. The noise n_k is modeled as a signal dependent Gauss-Markov noise process with memory length L. Specifically,

$$n_k = \bar{b}^T n_{k-1}^{k-L} + \sigma\left(a_k^{k-I}\right) w_k, \tag{2}$$

where the vector $\bar{b}$ represents L coefficients of an autoregressive filter, $\sigma(a_k^{k-I})$ is a signal dependent parameter and w_k is a zero mean unit variance Gaussian random variable (random variables w_k are i.i.d.). Note that in the most general model (considered in [8]), even the autoregressive filter $\bar{b}$ would depend on the data sequence $\bar{a}$. However, in this work, we only work with models where $\bar{b}$ is fixed. We will revisit this point in Section III. The channel output z_k can be rewritten as,

$$z_k = y\left(a_k^{k-I}\right) + \bar{b}^T \begin{pmatrix} z_{k-L} - y\left(a_{k-L}^{k-L-I}\right) \\ \cdot \\ \cdot \\ \cdot \\ z_{k-1} - y\left(a_{k-1}^{k-1-I}\right) \end{pmatrix} + \sigma\left(a_k^{k-I}\right) w_k. \tag{3}$$

From the above analysis it can be concluded that (owing to space limitations we refer the reader to [8]),

$$f\left(z_k \,|z_{k-1}^0, \bar{a}\right) = f\left(z_k \,\middle|\, z_{k-1}^{k-L}, a_k^{k-L-I}\right). \tag{4}$$

B. Viterbi Detector

The maximum likelihood estimate of the bit sequence denoted $\hat{\bar{a}}$ is given by (see [8]),

$$\begin{aligned} \hat{\bar{a}} &= \arg \max_{\bar{a} \in \{0,1\}^N} f(\bar{z} \,|\, \bar{a}) = \arg \max_{\bar{a} \in \{0,1\}^N} \Pi_{k=0}^{N-1} f\left(z_k \,|z_{k-1}^0, \bar{a}\right) \\ &= \arg \max_{\bar{a} \in \{0,1\}^N} \Pi_{k=0}^{N-1} f\left(z_k \,\middle|\, z_{k-1}^{k-L}, a_k^{k-L-I}\right) \text{ (Using (4))} \\ &= \arg \max_{\bar{a} \in \{0,1\}^N} \Pi_{k=0}^{N-1} \frac{f\left(z_k^{k-L} \,|a_k^{k-L-I}\right)}{f\left(z_{k-1}^{k-L} \,|a_k^{k-L-I}\right)}. \end{aligned} \tag{5}$$

We define a state $S_k = a_k^{k-L-I+1}$ (there will be a total of M $(= 2^{L+I})$ number of states). With this definition, $f(z_k^{k-L} \,|\, a_k^{k-L-I}) = f(z_k^{k-L} \,|\, S_k^{k-1})$. Moreover, it is Gaussian distributed

$$f\left(z_k^{k-L} \,\middle|\, S_k^{k-1}\right) \sim N\left(\bar{\mu}\left(S_k^{k-1}\right), C\left(S_k^{k-1}\right)\right), \tag{6}$$

where $\bar{\mu}(S_k^{k-1})$ is the mean vector and $C(S_k^{k-1})$ is the covariance matrix. Here we assume the state S_0 is known and the initial L realizations of the observed output z_0^{-L} are also given.

With our state definition, we can reformulate the detection problem as the following MLSD problem [8].

$$\begin{aligned} \hat{\bar{S}} &= \arg \max_{\bar{S}} \Pi_{k=0}^{N-1} \frac{f\left(z_k^{k-L} \,|a_k^{k-L-I}\right)}{f\left(z_{k-1}^{k-L} \,|a_k^{k-L-I}\right)} \\ &= \arg \max_{\bar{S}} \Pi_{k=0}^{N-1} \frac{f\left(z_k^{k-L} \,|S_k^{k-1}\right)}{f\left(z_{k-1}^{k-L} \,|S_k^{k-1}\right)} \\ &= \arg \min_{\bar{S}} \sum_{k=0}^{N-1} \log \frac{\left|C\left(S_k^{k-1}\right)\right|}{\left|\mathbf{c}\left(S_k^{k-1}\right)\right|} \\ &\quad + \left(z_k^{k-L} - \bar{\mu}\left(S_k^{k-1}\right)\right)^T C\left(S_k^{k-1}\right)^{-1} \left(z_k^{k-L} - \bar{\mu}\left(S_k^{k-1}\right)\right) \\ &\quad - \left(z_{k-1}^{k-L} - \bar{\mu}'\left(S_k^{k-1}\right)\right)^T \mathbf{c}\left(S_k^{k-1}\right)^{-1} \\ &\quad \times \left(z_{k-1}^{k-L} - \bar{\mu}'\left(S_k^{k-1}\right)\right), \end{aligned} \tag{7}$$

where $\hat{\bar{S}}$ is the estimated state sequence, $\mathbf{c}(S_k^{k-1})$ is the upper $L \times L$ principal minor of $C(S_k^{k-1})$ and $\bar{\mu}'(S_k^{k-1})$ collects the first L elements of $\bar{\mu}(S_k^{k-1})$. It is assumed that the first state is known. With the metric given above, Viterbi decoding can be applied to get the ML state sequence [13] and the corresponding bit sequence.

The matrix $C(S_k^{k-1})$ is of dimension $(L+1) \times (L+1)$. For higher values of L, the complexity of detector increases as the decoding metric involves the inversion of the matrix $C(S_k^{k-1})$. However, the matrix inversion lemma can be used here to obtain [8]

$$\begin{aligned} C\left(S_k^{k-1}\right)^{-1} &= \begin{bmatrix} \mathbf{c}\left(S_k^{k-1}\right) & \bar{c} \\ \bar{c}^T & \lambda \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{c}\left(S_k^{k-1}\right)^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{\bar{w}\left(S_k^{k-1}\right) \bar{w}\left(S_k^{k-1}\right)^T}{\gamma\left(S_k^{k-1}\right)}, \end{aligned} \tag{8}$$

where

$$\bar{w}\left(S_k^{k-1}\right) = \begin{bmatrix} -\mathbf{c}\left(S_k^{k-1}\right)^{-1} \bar{c} \\ 1 \end{bmatrix} = \begin{bmatrix} -\bar{b} \\ 1 \end{bmatrix}, \quad \text{and}$$

$$\gamma\left(S_k^{k-1}\right) = \left(\lambda - \bar{c}^T \mathbf{c}\left(S_k^{k-1}\right)^{-1} \bar{c}\right) = \sigma^2\left(a_k^{k-I}\right).$$

As shown in [8] using (8), we can simplify the detector as follows.

$$\hat{\bar{S}} = \arg\min_{\bar{S}} \sum_{k=0}^{N-1} \log \sigma^2 \left(a_k^{k-I}\right) + \frac{\left([-\bar{b}^T, 1]\left(z_k^{k-L} - \bar{\mu}\left(S_k^{k-1}\right)\right)\right)^2}{\sigma^2\left(a_k^{k-I}\right)}. \quad (9)$$

As noted in [8] the above expression does not involve any matrix inversion; this reduces the complexity of the detector. Another observation is that the Viterbi decoding metric involves passing z_k^{k-L} through a filter $[-\bar{b}^T, 1]$ which is the inverse of the autoregressive filter of noise process n_k shown in Fig. 1. Thus, the metric first uncorrelates the noise with an FIR filter and then applies the Euclidean metric to the output of the filter.

III. Upper Bound on BER

It is evident from the discussion in Section II that for our channel model, the PEP is asymmetric. In particular, it does not factorize as a product of appropriate functions as required by flowgraph techniques. We now demonstrate that the Gallager upper bounding technique [11], coupled with a suitable change of variables can allow us to address this issue.

Denote an error event of length N as $\epsilon_N = (\bar{S}, \hat{\bar{S}})$ such that $\bar{S}$ and $\hat{\bar{S}}$ are valid state sequences and $S_j = \hat{S}_j$, $S_{j+N} = \hat{S}_{j+N}$, $S_{j+k} \neq \hat{S}_{j+k}$ for $1 \leq k \leq N-1$ and $S_{j+k} = \hat{S}_{j+k}$ for other values of k where $\hat{S}_k$ and S_k are the estimated and correct state respectively. Using this, an upper bound on the BER can be found as follows [5],

$$P_b(e) \leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{\bar{S}}:(\bar{S},\hat{\bar{S}})\in E_N} \nu(\bar{S}, \hat{\bar{S}}) P(\hat{\bar{S}} \mid \bar{S}), \quad (10)$$

where $\nu(\bar{S}, \hat{\bar{S}})$ is the number of erroneous bits along the sequences $\bar{S}$ and $\hat{\bar{S}}$ and E_N is the set of all error events ϵ_N of length N. The number of erroneous bits is given by [5]

$$\nu(\bar{S}, \hat{\bar{S}}) = \frac{d}{dZ}\left[\Pi_{k=1}^{N} Z^{\delta(a_{j+k}, \hat{a}_{j+k})}\right]\Big|_{Z=1} \quad (11)$$

where $\delta(a_{j+k}, \hat{a}_{j+k}) = 1$, if $a_{j+k} \neq \hat{a}_{j+k}$ and Z is a dummy variable. Using this the upper bound above can be expressed as

$$P_b(e) \leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{\bar{S}}:(\bar{S},\hat{\bar{S}})\in E_N} \frac{d}{dZ}\left[\Pi_{k=1}^{N} Z^{\delta(a_{j+k}, \hat{a}_{j+k})}\right]\Big|_{Z=1} \cdot P(\hat{\bar{S}} \mid \bar{S}) \quad (12)$$

where $P(\bar{S}) = P(S_j)P(S_{j+1} \mid S_j) \ldots P(S_{j+N} \mid S_{j+N-1}) = \frac{1}{M} \cdot \frac{1}{2^N}$ if $\bar{S}$ is valid state sequence, ($M(= 2^{L+I})$ is the number of states). The upper bound on the PEP can be simplified using Gallager's technique [11] as shown below.

Let $A(\bar{S}, \hat{\bar{S}}) = \{z_{j+N}^{j} : f(z_{j+N}^{j} \mid \hat{\bar{S}}) \geq f(z_{j+N}^{j} \mid \bar{S})\}$, i.e., the set of z_{j+N}^{j} where the detector prefers the incorrect sequence as compared to the correct sequence. Note that using previous arguments, we also have that $A(\bar{S}, \hat{\bar{S}}) = \{\bar{z}_j = z_{j+N}^{j+1} : \Pi_{k=1}^{N} \frac{f(z_{j+k} \mid \hat{S}_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L})}{f(z_{j+k} \mid S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L})} \geq 1\}$. Now,

$$\begin{aligned} P(\hat{\bar{S}} \mid \bar{S}) &= P(\hat{S}_j \ldots \hat{S}_{j+N} \mid S_j \ldots S_{j+N}) \\ &= \int_{A(\bar{S},\hat{\bar{S}})} \Pi_{k=1}^{N} f\left(z_{j+k} \,\middle|\, S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right) d\bar{z}_j \\ &\leq \min_{\forall \rho_k} \int \Pi_{k=1}^{N} f\left(z_{j+k} \middle| S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right) \\ &\Pi_{k=1}^{N} \left(\frac{f\left(z_{j+k} \middle| \hat{S}_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right)}{f\left(z_{j+k} \middle| S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right)}\right)^{\rho_k} d\bar{z}_j \\ &= \min_{\forall \rho_k} \int \Pi_{k=1}^{N} \left(f\left(z_{j+k} \,\middle|\, S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right)\right)^{1-\rho_k} \\ &\quad \times \left(f\left(z_{j+k} \middle| \hat{S}_{k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right)\right)^{\rho_k} d\bar{z}_j \end{aligned} \quad (13)$$

where $0 \leq \rho_k \leq 1$ for $k = 1, \ldots, N$.

The above integral can be simplified as shown in (14) at the bottom of the page, where $u_{j+k} = [-\bar{b}^T, 1] \cdot [z_{j+k-L}, \ldots, z_{j+k}]^T$, $\mathfrak{M}(S_{j+k}^{j+k-1})) = [-\bar{b}^T, 1] \cdot \bar{\mu}(S_{j+k}^{j+k-1})$, $\hat{\mathfrak{M}}(\hat{S}_{j+k}^{j+k-1})) = [-\bar{b}^T, 1] \cdot \hat{\bar{\mu}}(\hat{S}_{j+k}^{j+k-1})$. The Jacobian matrix for the change of variables has determinant equal to 1, since the corresponding matrix of partial derivatives has ones on the diagonal and is lower triangular. Note that the change of variables decouples the original expression, so that it can be expressed as the product of N independent integrals. Now we can simplify the PEP as follows.

$$\begin{aligned} P(\hat{\bar{S}} \mid \bar{S}) &\leq \Pi_{k=1}^{N} \min_{\rho_k} \int \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}\left(a_{j+k}^{j+k-I}\right)\hat{\sigma}^{\rho_k}\left(\hat{a}_{j+k}^{j+k-I}\right)} \\ &\quad \times \exp\left(-\frac{(1-\rho_k)}{2\sigma^2\left(a_{j+k}^{j+k-I}\right)}\left(u_{j+k} - \mathfrak{M}\left(S_{j+k}^{j+k-1}\right)\right)^2 \right. \\ &\quad \left. -\frac{\rho_k\left(u_{j+k} - \hat{\mathfrak{M}}\left(\hat{S}_{j+k}^{j+k-1}\right)\right)^2}{2\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)}\right) du_{j+k} \quad (15) \\ &= \Pi_{k=1}^{N} W\left(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1}\right) \quad (16) \end{aligned}$$

$$\int \Pi_{k=1}^{N} \left(f\left(z_{j+k} \middle| S_{j+k}^{j+k-1}, z_{j+k-1}^{j+k-L}\right)\right)^{1-\rho_k} \left(f\left(z_{j+k} \middle| \hat{S}_{j+k}^{j+k-1} \quad z_{j+k-1}^{j+k-L}\right)\right)^{\rho_k} d\bar{z}_j = \int \Pi_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma^{1-\rho_k}\left(a_{j+k}^{j+k-I}\right)\hat{\sigma}^{\rho_k}\left(\hat{a}_{j+k}^{j+k-I}\right)}$$

$$\times \exp\left(-\frac{(1-\rho_k)\left([-\bar{b}^T, 1]\left(z_{j+k}^{j+k-L} - \bar{\mu}\left(S_{j+k}^{j+k-1}\right)\right)\right)^2}{2\sigma^2\left(a_{j+k}^{j+k-I}\right)} \quad -\frac{\rho_k\left([-\bar{b}^T, 1]\left(z_{j+k}^{j+k-L} - \hat{\bar{\mu}}\left(\hat{S}_{j+k}^{j+k-1}\right)\right)\right)^2}{2\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)}\right) d\bar{z}_j = \Pi_{k=1}^{N} \int \frac{1}{\sqrt{2\pi}}$$

$$\times \frac{1}{\sigma^{1-\rho_k}\left(a_{j+k}^{j+k-I}\right)\hat{\sigma}^{\rho_k}\left(\hat{a}_{j+k}^{j+k-I}\right)} \exp\left(-\frac{\rho_k\left(u_{j+k} - \hat{\mathfrak{M}}\left(\hat{S}_{j+k}^{j+k-1}\right)\right)^2}{2\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)} \quad -\frac{(1-\rho_k)\left(u_{j+k} - \mathfrak{M}\left(S_{j+k}^{j+k-1}\right)\right)^2}{2\sigma^2\left(a_{j+k}^{j+k-I}\right)}\right) du_{j+k} \quad (14)$$

TABLE I
CHANNEL MODEL PARAMETERS

Type	$\bar{b}$	$\sigma^2(00)$	$\sigma^2(01)$	$\sigma^2(10)$	$\sigma^2(11)$	M
I	[0.1, 0.5]	0.11	0.12	0.13	0.14	8
	[0.1, 0.3, 0.5]	0.21	0.22	0.23	0.24	16
II	[0.2, 0.4]	0.21	0.22	0.23	0.24	8
	[0.2, 0.3, 0.4]	0.21	0.22	0.23	0.24	16

where the simplification of the integral in (15) and $W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1})$ are given in the Appendix.

Remark: It is important to note that the factorization of the PEP given by (16) for our channel model is possible because the autoregressive filter $\bar{b}$ is not dependent on the input bit sequence. Let

$$\Gamma_1 = -\frac{(1-\rho_k)([-\bar{b}^T, 1](z_{j+k}^{j+k-L} - \bar{\mu}(S_{j+k}^{j+k-1})))^2}{2\sigma^2(a_{j+k}^{j+k-I})}$$

and

$$\Gamma_2 = -\frac{\rho_k([-\bar{b}^T, 1](z_{j+k}^{j+k-L} - \hat{\bar{\mu}}(\hat{S}_{j+k}^{j+k-1})))^2}{2\hat{\sigma}^2(\hat{a}_{j+k}^{j+k-I})}.$$

These terms appear in the argument of the exponential in (14). If $\bar{b}$ is data independent, the change of variables in (14) allows us to interchange the order of integral and product. In contrast if $\bar{b}$ is data dependent the change of variables will not work simultaneously for both Γ_1 and Γ_2, i.e., in general we will be unable to express (14) as a product of appropriate terms.

Probability of bit error can now be further simplified as [5],

$$P_b(e) \leq \sum_{N=1}^{\infty} \sum_{\bar{S}} P(\bar{S}) \sum_{\hat{\bar{S}}:(\bar{S},\hat{\bar{S}})\in E_N} \frac{d}{dZ}\left[\Pi_{k=1}^{N} Z^{\delta(a_{j+k},\hat{a}_{j+k})}\right]\Big|_{Z=1}$$

$$\cdot P(\hat{\bar{S}} \mid \bar{S}) = \frac{d}{dZ} \sum_{N=1}^{\infty} \frac{1}{M} \cdot \frac{1}{2^N} \sum_{\bar{S}} \sum_{\hat{\bar{S}}} \Pi_{k=1}^{N} Z^{\delta(a_{j+k},\hat{a}_{j+k})} \quad (17)$$

$$\cdot P(\hat{\bar{S}} \mid \bar{S})\Big|_{Z=1} \leq \frac{d}{dZ} \sum_{N=1}^{\infty} \frac{1}{M} \cdot \frac{1}{2^N} \sum_{\bar{S}} \sum_{\hat{\bar{S}}} \Pi_{k=1}^{N} Z^{\delta(a_{j+k},\hat{a}_{j+k})}$$

$$\cdot W\left(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1}\right)\Big|_{Z=1} (\text{Using (16)}) = \frac{1}{M}\frac{d}{dZ} \sum_{N=1}^{\infty} \sum_{\bar{S}} \sum_{\hat{\bar{S}}}$$

$$\cdot \Pi_{k=1}^{N} \frac{1}{2} Z^{\delta(a_{j+k},\hat{a}_{j+k})} W\left(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1}\right)\Big|_{Z=1}$$

$$= \frac{1}{M}\frac{d}{dZ}(T(Z))|_{Z=1}, \quad (18)$$

where M is number of states and $T(Z) = \sum_{N=1}^{\infty} \sum_{\bar{S}} \sum_{\hat{\bar{S}}} \Pi_{k=1}^{N} \frac{1}{2} Z^{\delta(a_{j+k},\hat{a}_{j+k})} W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1})$. For computing $T(Z)$, we can exploit the symmetry property of the function $W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1}))$. It should be noted that the upper bounds on $P(\hat{\bar{S}} \mid \bar{S})$ and $P(\hat{\bar{S}} \mid \bar{S})$ are given by the same quantity $\Pi_{k=1}^{N} W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1})$. This property facilitates the use of error state diagram for computing $T(Z)$. The construction of error state diagram can directly provide the analytical expression for the upper bound on BER [7].

A. Upper Bound for Data Dependent $\bar{b}$ Channel Model

Strictly speaking, our approach does not apply for channels where $\bar{b}$ is data-dependent. However, in our simulations, we have observed that our technique continues to closely track the simulation BER, in the high SNR regime as long as the amount of variation in $\bar{b}(a_k^{k-I})$ is small. In such a scenario, we can make the assumption on u_{j+k} in (15) that $u_{j+k} = [-\bar{b}_{av}^T, 1] \cdot [z_{j+k-L}, \ldots, z_{j+k}]^T \approx [-\bar{b}(a_k^{k-I})^T, 1] \cdot [z_{j+k-L}, \ldots, z_{j+k}]^T$ where $\bar{b}_{av}$ is obtained by taking average over all a_k^{k-I}. While we cannot claim that this will result in an upper bound, in practice, as shown in the simulation results in Section IV, we still get a valid upper bound on the error rate obtained via simulation.

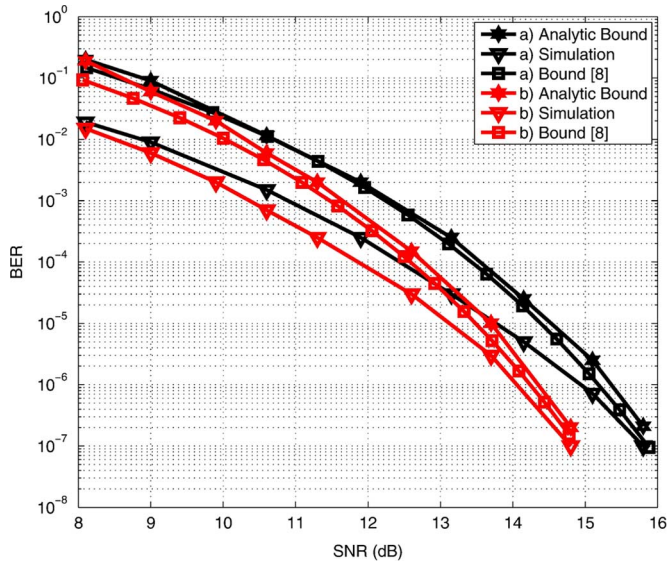

Fig. 2. BER with different SNR for the channel model with Type I parameters and a) 8 states in decoding and b) 16 states in decoding.

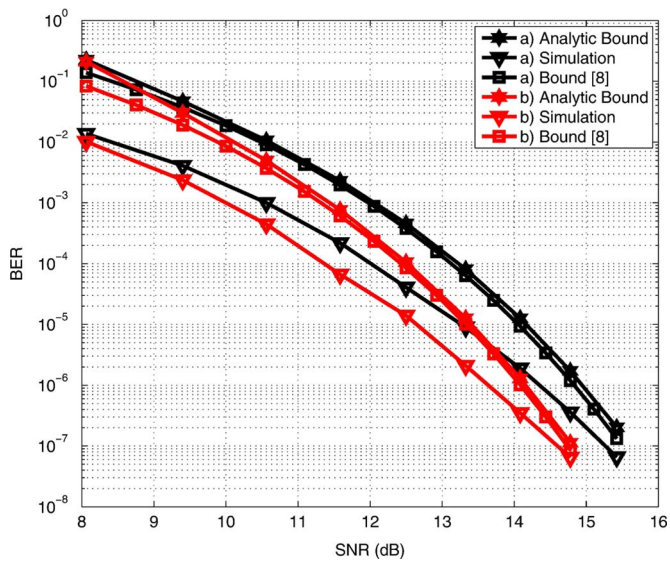

Fig. 3. BER with different SNR for the channel model with Type II parameters and a) 8 states in decoding and b) 16 states in decoding.

IV. SIMULATION RESULTS

We used BPSK signaling and a linear signal model, such that $y(a_{k-1}^{k}) = c(0.8a_k + 0.6a_{k-1})$ where the value of c can be varied to change the SNR. The SNR is defined as the signal power in $y(a_{k-I}^{k})$ divided by total noise variance. The specifications of the channel parameters are given in Table I. The simulations were performed by using trellis termination. It can be observed that in Figs. 2 and 3, the analytic bound follows the simulation BER in all cases. Precise values of the simulation BER and the proposed upper bound at certain high SNR points are given in Table II.

We note that even though the analytical bounds are tight in high SNR regime, the bounds do not match exactly with the simulation BER. This can be explained by considering the term $\Pi_{k=1}^{N} W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1})$

TABLE II
BER VALUES FOR DIFFERENT CHANNEL MODEL PARAMETERS

Type	M	SNR	BER	Analytic BER	Bound [8]
I	8	15.8 dB	1×10^{-7}	2.11×10^{-7}	1.01×10^{-7}
	16	14.8 dB	1×10^{-7}	2.02×10^{-7}	1.07×10^{-7}
II	8	15.42 dB	6.6×10^{-8}	2×10^{-7}	1.5×10^{-7}
	16	14.77 dB	6.5×10^{-8}	1.13×10^{-7}	1×10^{-7}

in (16). Our usage of the Gallager bound makes the term in (16) the same for $P(\hat{\bar{S}} \mid \bar{S})$ and $P(\bar{S} \mid \hat{\bar{S}})$. While this allows an efficient computation of the bound, it will in general be loose depending upon the difference between $P(\hat{\bar{S}} \mid \bar{S})$ and $P(\bar{S} \mid \hat{\bar{S}})$. In practice, however, it can be observed that the bound is quite tight and therefore useful.

It can be also be observed in simulation results that the KM (Kavcic and Maura) bound proposed in [8] gives a better bound compared to our bound. The KM bound computation involves the computation of pairwise error probability for each error event. Thus, it takes into account the asymmetric nature of PEP and provides a better bound. But we would like to clarify that the approach of KM is essentially a simulation-based heuristic and it is much more time-consuming for the class of channels that we consider in this work. Determining each error event in the KM bound requires extensive simulation. Moreover, computing the final bound requires summing over a set of error events that can be rather large. Thus, a subset of the error event set is considered in actually computing the bound. This simplification makes it a lower bound on the upper bound. Thus, one cannot claim that this is an upper bound. In contrast, our bound is easy to compute and considers all error events in the bound computation using the flowgraph techniques.

We also considered a channel with data dependent $\bar{b}(a_k^{k-I})$. The parameters are given by $\bar{b}(00) = [0.2, 0.4]$, $\bar{b}(01) = [0.21, 0.41]$, $\bar{b}(10) = [0.22, 0.42]$, $\bar{b}(11) = [0.23, 0.43]$, $\sigma^2(00) = 0.21$, $\sigma^2(01) = 0.22$, $\sigma^2(10) = 0.23$ and $\sigma^2(11) = 0.24$. For the upper bound computation, we used the average value of $\bar{b}$ given by, $\bar{b}_{av} = [0.215, 0.415]$ (cf. discussion in Section III-A). Once again in Fig. 4, it can be observed that the analytic bound follows the simulation bound. The KM bound works for the general channel model with data dependent $\bar{b}(a_k^{k-I})$ and takes into account the asymmetric PEP in the bound computation. Thus, it gives a better bound than our proposed bound.

V. CONCLUSIONS AND FUTURE WORK

We considered the problem of deriving an analytical upper bound for ML sequence detection in ISI channels with signal dependent Gauss-Markov noise. In these channels the pairwise error probability (PEP) is not symmetric. Moreover, it is hard to express the PEP as a product of appropriate terms that allow the application of flowgraph techniques. In this work, we considered a subset of these channels, and demonstrated an appropriate upper bound on the PEP. Using this upper bound along with an error state diagram based approach, we found analytical BER upper bounds that are tight in the high SNR regime. These bounds have been verified by our simulation results. It would be of interest to determine corresponding analytical lower bounds on the BER and to examine whether our techniques can be extended to address the general channel model.

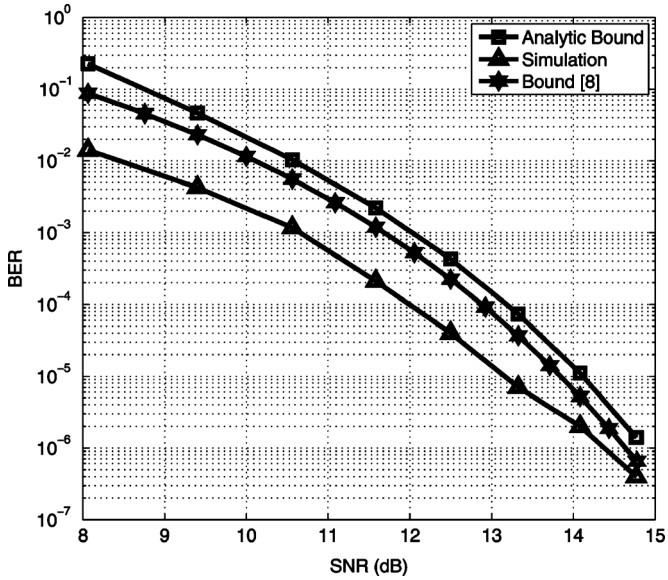

Fig. 4. BER with different SNR for the channel model with 8 states in decoding and data dependent $\bar{b}(a_k^{k-I})$.

APPENDIX

The integral in the (15) can be expressed in the following form,

$$\int \frac{1}{\sqrt{2\pi}\gamma} \exp\left(-\frac{1}{2}\left(\alpha(x-m)^2+\beta(x-\hat{m})^2\right)\right) dx = \frac{1}{\gamma\sqrt{\alpha+\beta}} \cdot \exp\left(-\frac{\alpha m^2+\beta\hat{m}^2}{2}+\frac{(\alpha m+\beta\hat{m})^2}{2(\alpha+\beta)}\right) \cdot \underbrace{\int \frac{\sqrt{\alpha+\beta}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha+\beta}{2}\left(x-\frac{\alpha m+\beta\hat{m}}{\alpha+\beta}\right)^2\right) dx}_{=1}$$
$$= \frac{1}{\gamma\sqrt{\alpha+\beta}} \exp\left(-\frac{\alpha m^2+\beta\hat{m}^2}{2}+\frac{(\alpha m+\beta\hat{m})^2}{2(\alpha+\beta)}\right)$$

where $\alpha = \frac{(1-\rho_k)}{\sigma^2(a_{j+k}^{j+k-I})}$, $\beta = \frac{\rho_k}{\hat{\sigma}^2(\hat{a}_{j+k}^{j+k-I})}$, $m = \mathfrak{M}(S_{j+k}^{j+k-1})$, $\gamma = \sigma^{1-\rho_k}(a_{j+k}^{j+k-I})\hat{\sigma}^{\rho_k}(\hat{a}_{j+k}^{j+k-I})$ and $\hat{m} = \hat{\mathfrak{M}}(\hat{S}_{j+k}^{j+k-1})$. Using the above equality, we can easily simplify the RHS of (15) and the term $W(S_{j+k}^{j+k-1}, \hat{S}_{j+k}^{j+k-1})$ as shown in the equation at the bottom of the page.

$$\min_{\rho_k} \frac{\sigma^{\rho_k}\left(a_{j+k}^{j+k-I}\hat{\sigma}^{1-\rho_k}\left(\hat{a}_{j+k}^{j+k-I}\right)\right.}{\sqrt{(1-\rho_k)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)+\rho_k\sigma^2\left(a_{j+k}^{j+k-I}\right)}} \exp\left(\frac{-\left((1-\rho_k)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)\mathfrak{M}^2\left(S_{j+k}^{j+k-1}\right)-\rho_k\sigma^2\left(a_{j+k}^{j+k-I}\right)\hat{\mathfrak{M}}^2\left(\hat{S}_{j+k}^{j+k-1}\right)\right.}{2\sigma^2\left(a_{j+k}^{j+k-I}\right)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)}\right.$$
$$\left.+\frac{\left((1-\rho_k)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)\mathfrak{M}\left(S_{j+k}^{j+k-1}\right)+\rho_k\sigma^2\left(a_{j+k}^{j+k-I}\right)\hat{\mathfrak{M}}\left(\hat{S}_{j+k}^{j+k-1}\right)\right)^2}{2\sigma^2\left(a_{j+k}^{j+k-I}\right)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)\left((1-\rho_k)\hat{\sigma}^2\left(\hat{a}_{j+k}^{j+k-I}\right)+\rho_k\sigma^2\left(a_{j+k}^{j+k-I}\right)\right)}\right)$$

REFERENCES

[1] J.-G. Zhu and H. Wang, "Noise characteristics of interacting transitions in longitudinal thin film media," *IEEE Trans. Magn.*, vol. 31, no. 2, pp. 1065–1070, Mar. 1995.

[2] A. Kavcic and J. M. F. Moura, "Correlation-sensitive adaptive sequence detection," *IEEE Trans. Magn.*, vol. 34, no. 3, pp. 763–771, May 1998.

[3] N. Kumar, P. Agarwal, A. Ramamoorthy, and M. Salapaka, "Maximum-likelihood sequence detector for dynamic mode high density probe storage," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1686–1694, Jun. 2010.

[4] G. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 363–378, 1972.

[5] E. Biglieri, "High-level modulation and coding for nonlinear satellite channels," *IEEE Trans. Commun.*, vol. 32, no. 5, pp. 616–626, May 1984.

[6] Y.-J. Liu, I. Oka, and E. Biglieri, "Error probability for digital transmission over nonlinear channels with application to tcm," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1101–1110, Sep. 1990.

[7] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York, NY, USA: McGraw-Hill, 1979.

[8] A. Kavcic and J. M. F. Moura, "The Viterbi algorithm and Markov noise memory," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 291–301, Jan. 2000.

[9] W. Zeng and J. Moon, "Modified Viterbi algorithm for a jitter-dominant 1-d2 channel," *IEEE Trans. Magn.*, vol. 28, no. 5, pp. 2895–2897, Sep. 1992.

[10] I. Lee and J. M. Cioffi, "Performance analysis of the modified maximum likelihood sequence detector in the presence of data-dependent noise," in *Proc. 26th Asilomar Conf.*, 1992, pp. 961–964.

[11] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[12] N. Kumar, A. Ramamoorthy, and M. Salapaka, "Performance evaluation of ML sequence detection in ISI channels with Gauss Markov noise," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–5.

[13] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967.