

# Design Curves and Information-Theoretic Limits for Perpendicular Recording Systems

Zheng Wu, Paul H. Siegel, *Fellow, IEEE*, H. Neal Bertram, *Fellow, IEEE*, and Jack K. Wolf, *Fellow, IEEE*

Center for Magnetic Recording Research, University of California at San Diego, La Jolla, CA 92093-0401 USA

To optimize a high-density magnetic recording system, one needs to know the tradeoffs between various components of the system including the read/write transducers, the magnetic medium, and the read channel. In this paper, we consider a channel model characterized by three parameters: the replay pulse width  $T_{50}$ , the transition jitter noise standard deviation  $\sigma_J$ , and the signal-to-electronic-noise ratio  $\text{SNR}_{\text{WG}}$ . We utilize information-theoretic tools to determine the acceptable region for the channel parameters  $T_{50}$  and  $\sigma_J$  when optimal detection and linear coding techniques are used. This paper is an extension of a similar analysis for a system that utilized a minimum mean-squared error (MMSE) equalizer, a Viterbi detector, and a Reed–Solomon (RS) code. Our main conclusion is that there is a considerable potential gain to be achieved by using improved detection and coding schemes as compared with the present system.

*Index Terms*—Information theory, jitter noise, parameter optimization, perpendicular magnetic recording.

## I. INTRODUCTION

THE optimization of high-density magnetic recording systems requires the analysis and evaluation of tradeoffs among the many system parameters. For example, code rate optimization for different channel models has been addressed in several recent studies [1]–[3]. In particular, Chaichanavong *et al.* [1] proposed a quasi-analytical methodology to optimize the system for a channel model with three parameters, the replay pulse width  $T_{50}$ , the transition jitter noise standard deviation  $\sigma_J$ , and the signal-to-electronic-noise ratio (SNR<sub>WG</sub>). The signal processing scheme included a minimum mean-squared error (MMSE) equalizer, a maximum-likelihood detector, and a Reed–Solomon (RS) error-correcting code. For a specified user bit spacing  $B_{\text{user}}$  and SNR<sub>WG</sub>, the results in [1] take the form of a design curve that defines the acceptable  $(T_{50}, \sigma_J)$  pairs for a specified sector error rate (SER). Each point on the curve has a corresponding optimal code rate  $R$  and an optimal channel bit spacing  $B = R B_{\text{user}}$ .

In this paper, for the same channel model, we apply information-theoretic tools to derive design curves that shed light on the acceptable region of  $(T_{50}, \sigma_J)$  pairs when optimal detection and linear coding techniques are used.

There has been much previous research reported on the computation of information rates and capacity for magnetic recording channels [5]–[8]. As an application of these techniques, Ryan *et al.* [4] determined the symmetric information rate (SIR) for a Lorentzian channel corrupted by additive white Gaussian noise (AWGN), and then used this result to maximize the user density.

We consider a perpendicular recording channel with both AWGN and jitter noise. Using upper bounds on the SIR, we compute information-theoretic design curves that, for a given

user density, provide insight into the tradeoff among the three parameters  $T_{50}$ ,  $\sigma_J$ , and SNR<sub>WG</sub>. The results are compared to those of the RS-coded system described in [1], and they show the possibility of obtaining considerable performance improvement with the help of advanced detection and coding techniques.

The paper is organized as follows. Section II gives a brief introduction to the channel model that is used throughout the paper. Section III describes our method for approximating the SIR. In Section IV, we derive design curves based upon the evaluation of the SIR and compare them with the analogous curves for the RS-coded system in [1]. Section V concludes the paper. Some derivations of results needed in Section III are presented in the Appendix.

## II. CHANNEL MODEL

The basic channel model used throughout this paper is the same as that used in [1] except for a slight change in notation.

The channel isolated transition response is given by

$$s(t) = V_{\text{max}} \text{erf} \left( \frac{0.954t}{T_{50}} \right) \quad (1)$$

where  $\text{erf}(\cdot)$  is the error function defined by

$$\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-u^2} du. \quad (2)$$

The constant 0.954 comes from the definition of  $T_{50}$ , which is the width of the transition response at half of the maximum amplitude, i.e.,  $s(T_{50}/2) = V_{\text{max}}/2$ .

The channel dipulse response is defined as

$$h(t) = \frac{1}{2} (s(t) - s(t - B)) \quad (3)$$

where  $B$  is the channel bit spacing.

The channel input is assumed to be an independent identically distributed (i.i.d.) equiprobable binary sequence  $\{x_i\}$ , where  $x_i \in \{+1, -1\}$ . The noiseless channel output  $y(t)$  can be

written as the convolution of the channel input and the dipulse response

$$y(t) = \sum_i x_i h(t - iB). \quad (4)$$

Jitter noise is modeled as

$$n_J(t) = \sum_i a_i \left( \frac{x_i - x_{i-1}}{2} \right) s'(t - iB) \quad (5)$$

where the  $\{a_i\}$  are independent Gaussian random variables with zero mean and variance  $\sigma_J^2$  and  $s'(t)$  is the first derivative of the transition response  $s(t)$ .

Electronic noise and a small amount of stationary medium noise, not included in the jitter noise, can be modeled as AWGN represented by  $n_W(t)$ . Therefore, the output of the channel can be written as

$$z(t) = y(t) + n_J(t) + n_W(t). \quad (6)$$

After low-pass filtering and sampling at the channel bit spacing  $B$ , the discrete-time channel output is

$$z_k = \sum_i h_i x_{k-i} + \sum_i \tilde{a}_{k-i} \left( \frac{x_{k-i} - x_{k-i-1}}{2} \right) s'_i + w_k \quad (7)$$

where  $h_i = h(iB)$ ,  $s'_i = B s'(iB)$ , and  $\{\tilde{a}_i\}$  are independent, zero-mean Gaussian random variables with variance  $(\sigma_J/B)^2$ , and  $\{w_k\}$  are independent, zero-mean Gaussian random variables with variance  $\sigma_W^2$ . The signal-to-electronic-noise ratio is given by  $\text{SNR}_{\text{WG}} = 10 \log_{10} (V_{\text{max}}^2 / \sigma_W^2)$ .

Assume  $h_i \approx 0$  when  $i < -A_1$  and  $i > A_2$  where  $A_1, A_2 > 0$ . Similarly, assume that  $s'_i \approx 0$  when  $i < -D_1$  and  $i > D_2$ , where  $D_1, D_2 > 0$ . The channel is, therefore, modeled as a finite-memory intersymbol interference (ISI) channel with data-dependent, finite-memory noise and AWGN.

### III. INFORMATION RATE COMPUTATION

The SIR of a channel is the mutual information rate corresponding to independent, identically distributed, equiprobable binary inputs. The SIR is generally interpreted as the maximum achievable rate for which reliable recording is possible using a binary, linear code. For the channel described previously, we use the SIR to examine the tradeoffs among system design parameters assuming optimal detection and linear coding. In this section, the method we used to evaluate the SIR is introduced.

In [5] and [6], independently, a simulation-based method to evaluate the information rate of a Markovian channel was proposed. This method was extended in [7] and [8] to evaluate the information rate and the capacity for the magnetic recording channel, assuming linear ISI and data-dependent, colored Gaussian noise. We use a similar method in this paper.

In the following analysis, we denote a column vector  $(x_i, x_{i+1}, \dots, x_j)^T$  by  $\bar{x}_i^j$ . The information rate between the channel input  $X$  and the channel output  $Z$  can be written as

$$I(X; Z) = \lim_{n \rightarrow \infty} \frac{1}{n} I(\bar{x}_1^n; \bar{z}_1^n) = h(Z) - h(Z | X) \quad (8)$$

where  $h(Z) = \lim_{n \rightarrow \infty} (1/n) h(\bar{z}_1^n)$  is the entropy rate of the channel output and  $h(Z | X) = \lim_{n \rightarrow \infty} (1/n) h(\bar{z}_1^n | \bar{x}_1^n)$  is the entropy rate of the channel output conditioned on the channel input. We proceed to evaluate these two values separately.

According to [5],  $-(1/n) \log \Pr(\bar{z}_1^n)$  converges to  $h(Z)$  with probability one. Since we are considering a finite memory Markovian channel, it can be computed by the forward recursion of the BCJR algorithm [9]. Let  $S_k$  be the trellis state at time  $k$ . Then, the probability of the channel output sequence is

$$\Pr(\bar{z}_1^n) = \sum_{m \in \Omega} \Pr(S_n = m, \bar{z}_1^n) \quad (9)$$

where  $\Omega$  is the set of all states.

Define the forward state metric as

$$\alpha_k(m) = \Pr(S_k = m, \bar{z}_1^k). \quad (10)$$

Using Bayes' rule

$$\alpha_{k+1}(m) = \sum_{m' \in \Omega} \Pr(S_k = m', \bar{z}_1^k) \cdot \Pr(z_{k+1}, S_{k+1} = m | S_k = m', \bar{z}_1^k). \quad (11)$$

If we define the branch metric to be

$$\gamma_k(m', m) = \Pr(z_k, S_k = m | S_{k-1} = m', \bar{z}_1^{k-1}) \quad (12)$$

then, we have the following recursion relationship:

$$\alpha_{k+1}(m) = \sum_{m' \in \Omega} \alpha_k(m') \gamma_{k+1}(m', m). \quad (13)$$

Since  $\alpha_k(m)$  is very close to zero when  $k$  is large, we normalized the sum of all forward state metrics at each time  $k$  to be 1 in order to avoid losing precision in the computation. Therefore, denoting the normalized forward state metric at time  $k$  and state  $m$  by  $\tilde{\alpha}_k(m)$ , and the normalization factor by  $\lambda_k$ , the new recursion relationship takes the form

$$\tilde{\alpha}_{k+1}(m) = \lambda_{k+1} \sum_{m' \in \Omega} \tilde{\alpha}_k(m') \gamma_{k+1}(m', m). \quad (14)$$

Setting the same initial values for recursions (13) and (14), we can show that

$$\tilde{\alpha}_k(m) = \left( \prod_{i=1}^k \lambda_i \right) \alpha_k(m). \quad (15)$$

Thus

$$\begin{aligned} -\frac{1}{n} \log \Pr(\bar{z}_1^n) &= -\frac{1}{n} \log \left[ \sum_{m \in \Omega} \alpha_n(m) \right] \\ &= -\frac{1}{n} \log \left[ \frac{\sum_{m \in \Omega} \tilde{\alpha}_n(m)}{\prod_{i=1}^n \lambda_i} \right] = \frac{1}{n} \sum_{i=1}^n \log(\lambda_i). \end{aligned} \quad (16)$$

The computation of the branch metric depends on the channel model. According to the derivations in the Appendix, the branch metric is

$$\gamma_k(m', m) = c_1 \exp \left\{ c_2 \left[ (\bar{z}_{k-L}^k - \bar{y}_{k-L}^k)^T \bar{w} \right]^2 \right\} \quad (17)$$

where  $c_1$  and  $c_2$  are edge-dependent constants and  $\bar{w}$  is an edge-dependent column vector, as defined in (30)–(32). Here,  $L$  is the memory of the jitter noise, defined as  $L = D_1 + D_2$ , with  $D_1$  and  $D_2$  as specified at the end of Section II.

Using the branch metrics, the initial values for the forward state metrics, and the recursion relationship (14), we can compute the probability of a long channel output realization, and, thus, get an estimate of the output entropy rate.

Computing the conditional entropy rate  $h(Z | X)$  is quite straightforward for this channel model [7]. The conditional entropy rate is

$$h(Z | X) = \lim_{n \rightarrow \infty} \frac{1}{n} E [h(\bar{z}_1^n | \bar{x}_1^n = \bar{x}^*)] \quad (18)$$

where  $\bar{x}^*$  is a realization of the input sequence  $\bar{x}_1^n$  and the expectation is over all possible input sequences.

Since we are considering the SIR, we assume the input sequences are equiprobable. For a random input sequence realization  $\bar{x}^*$ ,  $(1/n)h(\bar{z}_1^n | \bar{x}_1^n = \bar{x}^*)$  converges to  $h(Z | X)$  when  $n$  goes to infinity. With the knowledge that the channel output sequence obeys a joint Gaussian distribution when conditioned on a specific channel input sequence, the conditional entropy rate is

$$h(Z | X) \approx \frac{1}{2} \log(2\pi e |\mathbf{R}_{\bar{x}^*}|) \quad (19)$$

where  $\mathbf{R}_{\bar{x}^*}$  is the covariance matrix of the output sequence, given the input sequence  $\bar{x}^*$ . The length of the input sequence should be large.

The number of states in the trellis grows exponentially with the sum of the memory length of the channel ISI and the memory length of the jitter noise. Thus, if these quantities are large, the complexity in computing the SIR is dominated by the channel output entropy rate calculation.

One method to reduce the computational complexity is to reduce the number of states in the trellis by truncating the channel dipulse response  $\{h_i\}$  and the sampled first derivative of  $s(t)$ ,  $\{s'_i\}$ . In [7], it is stated that if we use an approximation of the real channel in the BCJR forward recursion, the computed output entropy rate is an upper bound on the true value of  $h(Z)$ . However, although this method can reduce the computational complexity significantly, it generates a loose upper bound.

Another method is to try to reduce the number of branch metric computations in the recursion. Recall that, for each state, the new state metric is obtained by calculating the branch metric for each incoming edge, multiplying each of these by the state metric of the corresponding state from the previous stage, and then summing up these products. However, the normalized forward state metric  $\tilde{\alpha}_k(m')$ , which equals the conditional probability  $\Pr(S_k = m' | \bar{z}_1^k)$ , may be very close to zero if the current state  $m$  is unlikely to be the actual state at time  $k$ , given

the past channel output sequence. Since the branch metric is a multivariate Gaussian density function, which is bounded, we can ignore terms in the summation (14) corresponding to sufficiently small values of  $\tilde{\alpha}_k(m')$ . Therefore, we set a threshold value  $\Delta$ , and if the forward state metric  $\tilde{\alpha}_k(m')$  is smaller than  $\Delta$ , it is set to zero and the branch metrics on the edges starting from state  $m'$  are not calculated. This approximation results in forward state metrics, and, therefore, estimated output sequence probabilities, that are smaller than their true values. It follows that the computed estimate of the output entropy rate is an upper bound for the true value.

Simulations confirmed that this method significantly reduces the required amount of computation, while producing a tight upper bound on the output entropy rate. For example, for a 512-state trellis and a threshold value  $\Delta = 10^{-3}$ , the computation time consumed with this approximation is about one-fifth of that required by the exact computation. We observed that the number of nonzero forward state metrics was reduced to no more than 100 after a few steps, and sometimes dropped to less than 10. At the same time, the entropy rate upper bound differed only slightly from the exact value. Therefore, in the simulations used to generate the design curves, we chose a reasonably large number of states, namely  $2^{13}$ , and applied this approximation method to compute our upper bounds on the SIR.

#### IV. DESIGN CURVES FOR SIR

For a specified user density and  $\text{SNR}_{\text{WG}}$ , the information-theoretic design curves computed in this paper essentially determine the region of  $(T_{50}, \sigma_J)$  pairs that are acceptable in a system utilizing a code that achieves the SIR.

As in [1], we simplify the calculation of the design curves by normalizing with respect to the user bit spacing, defined by  $B_{\text{user}} = B/R$ , where the code rate  $R$  is set equal to the SIR. The method for deriving the normalized curve is modeled after the approach used in [1]. Specifically, for a given  $\text{SNR}_{\text{WG}}$  and various values of the ratio  $T_{50}/\sigma_J$ , we first compute the SIR bound as a function of  $\sigma_J/B$ . We then determine the corresponding values of  $T_{50}/B_{\text{user}} = R T_{50}/B$ , where  $R$  is the calculated SIR bound.

Fig. 1(a) and (b) illustrates the numerical results for these two steps for the case where  $\text{SNR}_{\text{WG}} = 17$  dB. Note that for each curve in Fig. 1(b), there is a point representing a maximum value of  $T_{50}/B_{\text{user}}$ . The corresponding points in the  $(T_{50}/B_{\text{user}}, \sigma_J/B_{\text{user}})$ -plane form the normalized information-theoretic design curve for  $\text{SNR}_{\text{WG}} = 17$  dB, shown in Fig. 2 along with the normalized design curves for  $\text{SNR}_{\text{WG}}$  values of 14 and 20 dB.

For purposes of comparison, Fig. 2 also shows the normalized design curves derived in [1] for the RS-coded system at the same values of  $\text{SNR}_{\text{WG}}$ . (We remark that, in order to facilitate the comparison to the curves taken from [1], the information-theoretic results plotted in Fig. 2 were computed for different values of the ratio  $T_{50}/\sigma_J$  than those used to generate the curves in Fig. 1).

It is clear that there is a significant gap between the curves representing the same  $\text{SNR}_{\text{WG}}$ . Thus, we can infer that better detection methods and coding schemes might extend the acceptable region of system parameters considerably.

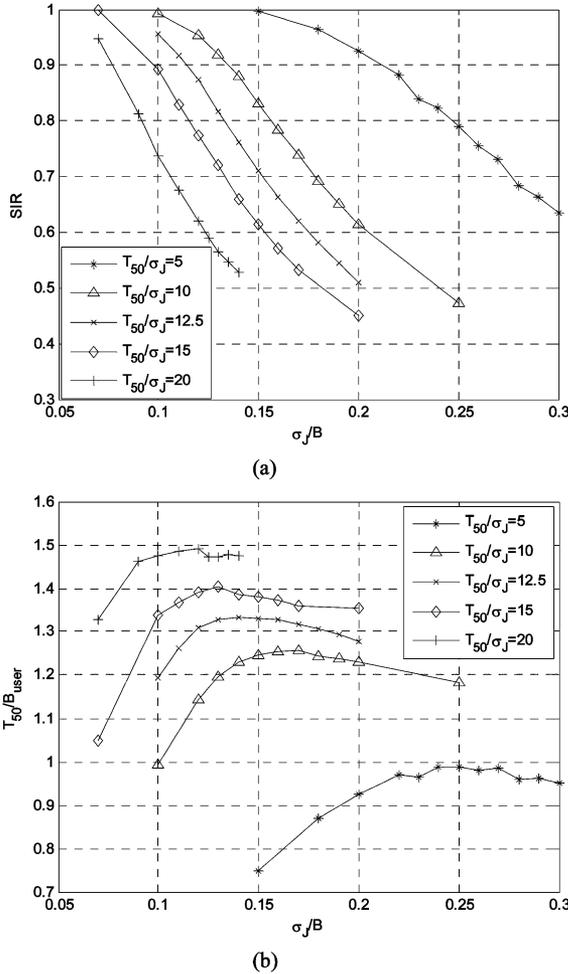
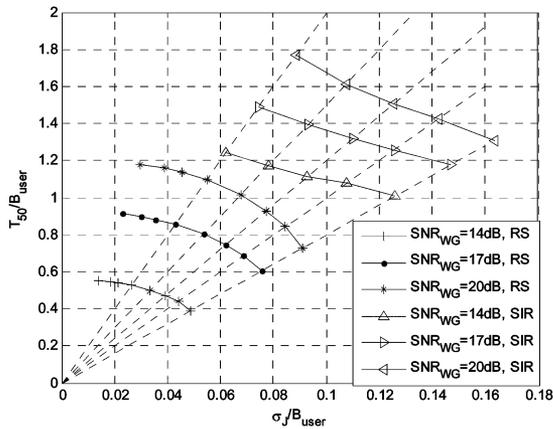
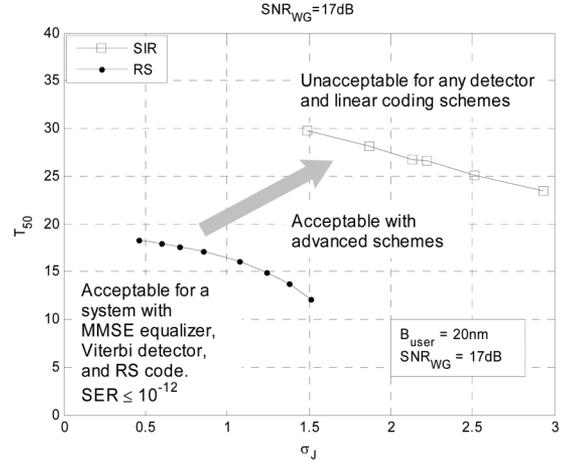
Fig. 1. SIR curves and normalized user density curves for  $\text{SNR}_{\text{WG}} = 17$  dB.

Fig. 2. Normalized design curves.

Also, it is interesting to note that the shapes of the two sets of curves are not the same. This suggests that the tradeoff between the system parameters might be different for SIR-achieving codes than for the optimized RS codes examined in [1].

In Fig. 3, we plot the design curves in the  $(T_{50}, \sigma_J)$ -plane for  $\text{SNR}_{\text{WG}} = 17$  dB and for user bit spacing  $B_{\text{user}} = 20$  nm. The region below the lower design curve is the acceptable region of  $(T_{50}, \sigma_J)$  pairs for the RS-coded system in [1]. The region

Fig. 3. Comparison of the design curves for optimal RS code and SIR at  $\text{SNR}_{\text{WG}} = 17$  dB,  $B_{\text{user}} = 20$  nm.TABLE I  
COMPARING  $T_{50}/B_{\text{user}}$ 

$T_{50}/\sigma_J$	RS-coded system			SIR		
	14dB	17dB	20dB	14dB	17dB	20dB
8	0.39	0.60	0.73	1.01	1.17	7.31
10	0.44	0.69	0.84	1.08	1.26	1.43
12	0.47	0.75	0.93	1.11	1.33	1.50
15	0.50	0.81	1.02	1.17	1.40	1.61
20	0.53	0.86	1.10	1.24	1.49	1.77

TABLE II  
COMPARING CODE RATE

$T_{50}/\sigma_J$	RS-coded system			SIR		
	14dB	17dB	20dB	14dB	17dB	20dB
8	0.55	0.71	0.77	0.66	0.70	0.71
10	0.60	0.76	0.79	0.63	0.74	0.75
12	0.61	0.78	0.80	0.71	0.69	0.78
15	0.63	0.77	0.82	0.65	0.72	0.72
20	0.63	0.81	0.86	0.54	0.62	0.77

above the upper design curve can be interpreted as  $(T_{50}, \sigma_J)$  pairs for which no detection method and linear coding scheme can guarantee reliable data retrieval. The region between these two curves gives insight into head and media parameters that are acceptable with sufficiently powerful detection and coding schemes.

The numerical results represented by the curves in Fig. 2 are shown in Tables I and II. Table I gives the  $T_{50}/B_{\text{user}}$  values for points on the normalized design curves for both the RS-coded system and the information-theoretic limiting case. Table II shows the corresponding optimal RS code rates and the SIR upper bound values. Note that the SIR-achieving code almost doubles the user density  $T_{50}/B_{\text{user}}$  relative to the optimized RS code, whereas the optimal RS code rates and SIR upper bounds are relatively close.

## V. CONCLUSION

In this paper, we determined information-theoretic design curves for a perpendicular magnetic recording channel model. We computed a tight upper bound for the SIR and presented design curves that demonstrate some tradeoffs among system

parameters when using an SIR-achieving coding scheme. Given a set of channel parameters ( $T_{50}$ ,  $\sigma_J$ , and  $\text{SNR}_{\text{WG}}$ ), we can use the design curves to determine—approximately—the maximum user density we can achieve using any detection method and linear coding scheme. Our results show a substantial gap between these information-theoretic design curves and those computed in [1] for an optimized RS-coded system. This suggests that advanced detection and coding schemes have the potential to significantly improve system capacity and performance.

#### APPENDIX

Referring to the channel model in Section II, we define  $I = \max\{A_1, D_1\}$  and  $T = \max\{A_2, D_2 + 1\}$ . We also recall that  $L = D_1 + D_2$ , so the state of the channel is given by  $S_k = \{x_{k-T-L+1}, \dots, x_{k+I}\}$ . With these definitions, the Markovian channel satisfies the relations

$$\Pr(S_k | \bar{S}_1^{k-1}, \bar{z}_1^{k-1}) = \Pr(S_k | S_{k-1}) \quad (20)$$

and

$$\Pr(z_k | \bar{S}_1^n, \bar{z}_1^{k-1}) = \Pr(z_k | S_{k-1}, S_k, \bar{z}_1^{k-1}). \quad (21)$$

The branch metric defined in Section III can, therefore, be written as

$$\begin{aligned} \gamma_k(m', m) &= \Pr(z_k, S_k = m | S_{k-1} = m', \bar{z}_1^{k-1}) \\ &= \Pr(z_k | S_k = m, S_{k-1} = m', \bar{z}_1^{k-1}) \\ &\quad \cdot \Pr(S_k = m | S_{k-1} = m', \bar{z}_1^{k-1}) \\ &= \Pr(z_k | S_k = m, S_{k-1} = m', \bar{z}_{k-L}^{k-1}) \\ &\quad \cdot \Pr(S_k = m | S_{k-1} = m'). \end{aligned} \quad (22)$$

The transition probability from state  $S_{k-1}$  to state  $S_k$  is either 0, if there is no edge between them, or 0.5, assuming i.i.d. equiprobable binary channel inputs.

The channel output satisfies

$$\bar{z}_{k-L}^k = \bar{y}_{k-L}^k + \mathbf{Q} \bar{a}_{k-L-B_2}^{k+B_1} + \bar{w}_{k-L}^k \quad (23)$$

where  $\mathbf{Q}$  is an  $L \times (2L + 1)$  matrix defined as

$$\mathbf{Q} = \begin{pmatrix} s'_{B_2} & s'_{B_2-1} & \cdots & s'_{-B_1} & 0 & \cdots & 0 \\ 0 & s'_{B_2} & s'_{B_2-1} & \cdots & s'_{-B_1} & 0 & 0 \\ 0 & 0 & \ddots & \cdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & s'_{B_2} & s'_{B_2-1} & \cdots & s'_{-B_1} \\ \delta_{k-L-B_2} & 0 & 0 & & & & \\ 0 & \ddots & 0 & & & & \\ 0 & 0 & 0 & \delta_{k+B_1} & & & \end{pmatrix} \quad (24)$$

and  $\delta_i$  is defined as  $\delta_i = (x_i - x_{i-1})/2$

The conditional joint density function  $\Pr(\bar{z}_{k-L}^k | S_{k-1}, S_k)$  is multivariate Gaussian with covariance matrix

$$\begin{aligned} \mathbf{R}^{(L)} &= E \left\{ \left( \mathbf{Q} \bar{a}_{k-L-B_2}^{k+B_1} + \bar{w}_{k-L}^k \right) \left( \mathbf{Q} \bar{a}_{k-L-B_2}^{k+B_1} + \bar{w}_{k-L}^k \right)^T \right\} \\ &= \sigma_J^2 \mathbf{Q} \mathbf{Q}^T + \sigma_W^2 \mathbf{I}. \end{aligned} \quad (25)$$

Similarly, the conditional density of  $\bar{z}_{k-L}^{k-1}$  given  $S_{k-1}$  and  $S_k$  is multivariate Gaussian with covariance matrix  $\mathbf{R}^{(L-1)}$  which is obtained from  $\mathbf{R}^{(L)}$  by eliminating the last row and column; that is

$$\mathbf{R}^{(L)} = \begin{pmatrix} \mathbf{R}^{(L-1)} & \bar{r}^{(L)} \\ (\bar{r}^{(L)})^T & r_{LL} \end{pmatrix}. \quad (26)$$

Noting that

$$\begin{aligned} \Pr(z_k | S_k = m, S_{k-1} = m, \bar{z}_{k-L}^{k-1}) \\ = \frac{\Pr(\bar{z}_{k-L}^k | S_k = m, S_{k-1} = m)}{\Pr(\bar{z}_{k-L}^{k-1} | S_k = m, S_{k-1} = m)} \end{aligned} \quad (27)$$

we can express the branch metric as

$$\begin{aligned} \gamma_k(m', m) &= \Pr(S_k = m | S_{k-1} = m') \cdot \sqrt{\frac{|\mathbf{R}^{(L-1)}|}{2\pi|\mathbf{R}^{(L)}|}} \\ &\quad \cdot \exp \left\{ -\frac{1}{2} (\bar{z}_{k-L}^k - \bar{y}_{k-L}^k)^T \right. \\ &\quad \cdot \left( \mathbf{R}^{(L)} \right)^{-1} (\bar{z}_{k-L}^k - \bar{y}_{k-L}^k) \\ &\quad + \frac{1}{2} (\bar{z}_{k-L}^{k-1} - \bar{y}_{k-L}^{k-1})^T \left( \mathbf{R}^{(L-1)} \right)^{-1} \\ &\quad \left. \cdot (\bar{z}_{k-L}^{k-1} - \bar{y}_{k-L}^{k-1}) \right\}. \end{aligned} \quad (28)$$

The Gaussianity of the conditional joint density functions mentioned previously implies that

$$\gamma_k(m', m) = c_1 \exp \left\{ c_2 \left[ (\bar{z}_{k-L}^k - \bar{y}_{k-L}^k)^T \bar{\omega} \right]^2 \right\} \quad (29)$$

where

$$c_1 = \begin{cases} \frac{1}{2} \sqrt{\frac{|\mathbf{R}^{(L-1)}|}{2\pi|\mathbf{R}^{(L)}|}}, & \text{if there is an edge from } S_{k-1} \text{ to } S_k \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

$$c_2 = -\frac{1}{2[r_{LL} - (\bar{r}^{(L)})^T (\mathbf{R}^{(L-1)})^{-1} \bar{r}^{(L)}]} \quad (31)$$

and

$$\bar{\omega} = \begin{pmatrix} -(\mathbf{R}^{(L-1)})^{-1} \bar{r}^{(L)} \\ 1 \end{pmatrix}. \quad (32)$$

All these three parameters are edge-dependent.

#### ACKNOWLEDGMENT

This work was supported in part by the Information Storage Industry Consortium (INSIC) EHDR Program.

## REFERENCES

- [1] P. Chaichanavong, H. N. Bertram, and P. H. Siegel, "Design parameter optimization for perpendicular magnetic recording systems," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2549–2554, Oct. 2006.
- [2] W. E. Ryan, "Optimal code rates for concatenated codes on a PR4-equalized magnetic recording channel," *IEEE Trans. Magn.*, vol. 36, no. 6, pp. 4044–4049, Nov. 2000.
- [3] R. D. Cideciyan, E. Eleftheriou, and S. Tomasin, "Performance analysis of magnetic recording systems," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, 2001, vol. 9, pp. 2711–2715.
- [4] W. E. Ryan, F. Wang, R. Wood, and Y. Li, "Optimal code rates for the Lorentzian channel: Shannon codes and LDPC codes," *IEEE Trans. Magn.*, vol. 40, no. 6, pp. 3559–3565, Nov. 2004.
- [5] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channel with memory," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, 2001, vol. 9, pp. 2692–2695.
- [6] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite state ISI channels," in *Proc. IEEE Global Commun. Conf.*, San Antonio, TX, 2001, vol. 5, pp. 2992–2996.
- [7] Z. Zhang, T. M. Duman, and E. Kurtas, "Information rates of binary-input intersymbol interference channels with signal-dependent media noise," *IEEE Trans. Magn.*, vol. 39, no. 1, pp. 599–607, Jan. 2003.
- [8] S. Yang, A. K. Kavčić, and W. E. Ryan, "Optimizing the bit aspect ratio of a recording system using an information-theoretic criterion," *IEEE Trans. Magn.*, vol. 39, no. 5, pp. 2228–2230, Sep. 2003.
- [9] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

Manuscript received July 7, 2006 (e-mail: z2wu@ucsd.edu).