DISTANCE ENHANCING CONSTRAINTS FOR NOISE PREDICTIVE MAXIMUM LIKELIHOOD DETECTORS

Bruce E. Moision and Paul H. Siegel, *Fellow,IEEE*
Department of Electrical and Computer Engineering
University of California at San Diego
La Jolla, California 92093

Abstract—**Using performance analysis of Reduced State Sequence Estimators (RSSE), we characterize dominant error events for a Noise Predictive Maximum Likelihood (NPML) detector. The error event characterization may be used to determine distance enhancing constraints that improve the reliability of NPML/RSSE detection. An example of a constraint that provides approximately .8 dB asymptotic coding gain for an NPML detector operating at a user bit density of 2.54 is illustrated.**

I. Introduction

Many high density magnetic recording devices employ equalization to a partial response target (PR) and maximum-likelihood (ML) sequence detection using a Euclidean distance metric, a combination known as PRML. To simplify the implementation of the receiver, the target polynomial is chosen to be short in duration and constrained to have integer coefficients. At higher recording densities, the intersymbol interference (ISI) arising from a single pulse affects a larger number of adjacent symbol periods. Continuing to equalize to a short target increases the mismatch of the channel and target, and the subsequent equalization required to alter the overall response leads to noise coloration, noise enhancement and a resulting performance degradation. On the other hand, increasing the length of the target leads to an undesirable increase in the complexity of the detector.

The Noise Predictive Maximum Likelihood (NPML) detector [7],[5] improves the reliability of the PRML detector by adding a noise whitening filter at the input to the Viterbi detector and performing sequence estimation based on a longer effective target, with possibly non-integer coefficients. The potential added complexity of the detector is compensated for by using Reduced State Sequence Estimation (RSSE) [6],[8], which limits the number of states in the detector trellis.

Using results from the analysis of RSSE detectors [6],[12] we examine the performance of an NPML detector. Dominant error event lists for the NPML detector are generated and used to design distance enhancing constraints. We give an example of a constraint with Shannon Capacity $C \approx .914$ that yields a .8 dB asymptotic coding gain for an NPML detector with front end target $h(D) = 1 - D^2$, a 4-tap predictor, and 18 detector trellis states operating on a Lorentzian channel at a user bit density of 2.54.

The paper is organized as follows. Section II introduces the channel model used and briefly describes a PRML detector; Section III describes the addition of noise prediction; Section IV discusses the performance of an NPML detector; Section V discusses error event characterization; Section VI discusses distance enhancing constraints; and Section VII gives simulation results.

This work was supported in part by the UC Micro Project 97-160 in cooperation with Marvell Semiconductor Incorporated and NSF grant NCR-9612802

II. Channel Model

Figure 1 illustrates the model for the channel and NPML detector. The upper branch of the detector is a PRML detector, included for comparison. The

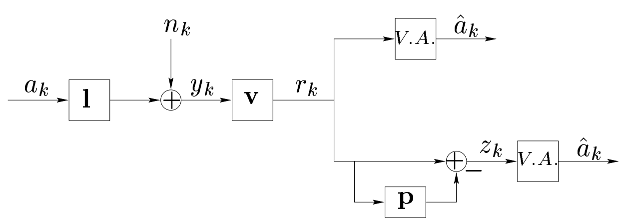

Fig. 1. Channel model

input to the channel is a binary sequence, $\{a_k\}_{k=-\infty}^{\infty}$, $a_k \in \{0,1\}$. The output is given by:

$$\begin{aligned} y_k &= \sum_{i=-N_l}^{N_l} l_i a_{k-i} + n_k \\ &= \mathbf{l}^T \mathbf{a} + n_k. \end{aligned}$$

The channel is modeled as a finite-length, discrete-time transversal filter, $\mathbf{l}$, plus additive noise, n_k. The additive noise samples are assumed uncorrelated, identically distributed and Gaussian with variance σ_n^2. For magnetic recording channels, this model could arise from a Lorentzian pulse followed by a sampled ideal low-pass filter or a sampled whitened matched filter, truncated to a finite duration in either case.

The equalizer $\mathbf{v}$, a finite-length transversal filter, is chosen to equalize to a front end target, $\mathbf{h}^T = [h_0, \cdots, h_\nu]$. The taps $\{v_k\}$ will be chosen to minimize the mean-square-error between the actual output of the equalizer, $(\mathbf{a} * \mathbf{l} + n_k) * \mathbf{v}$, and the desired output, $\mathbf{a} * \mathbf{h}$, where $*$ denotes discrete-time convolution. The output of the equalizer consists of the desired response, residual intersymbol interference from misequalization, ψ_k, and colored noise, η_k:

$$r_k = \mathbf{h}^T \mathbf{a}_{k:k-\nu} + \psi_k + \eta_k$$

Throughout the remainder of the analysis, we assume interference from misequalization is non-existent, $\psi_k = 0$, although simulation results include misequalization.

In a conventional PRML detector, an estimate of the transmitted sequence is formed by choosing the input sequence which, when passed through the desired impulse response, yields the sequence closest in Euclidean distance to the received sequence, $\{r_k\}$.

$$\hat{a}(D) = \arg\min_{b(D)} \|r(D) - b(D)h(D)\|$$

where $a(D)$ is the D transform of the sequence $\mathbf{a}$, $a(D) = \sum_{n=-\infty}^{\infty} a_n D^n$, and $\|a(D)\| = \sqrt{\sum a_i^2} = |\mathbf{a}|$. If the transmitted sequence was $a(D)$, an error in the estimate occurs when $\|r(D) - \hat{t}(D)\| < \|r(D) - t(D)\|$ where $t(D) = a(D)h(D)$, $\hat{t}(D) = \hat{a}(D)h(D)$, and $\hat{a}(D) \neq a(D)$.

Figure 2 illustrates the decision process between two candidate sequences in the signal space. An error

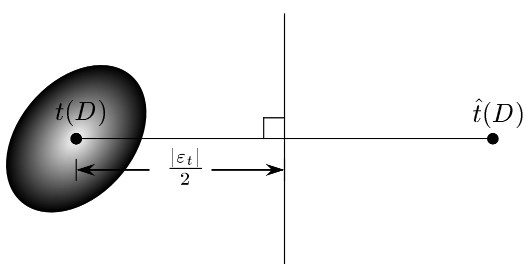

Fig. 2. Euclidean distance decision

occurs when the colored noise pushes $t(D) + \eta(D)$ over the decision boundary. The probability of this event is given by [3]

$$Pr[a(D) \to \hat{a}(D)] = Q\left(\frac{\frac{1}{2}|\varepsilon_t|}{\sigma_N}\right)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} dy$, $\varepsilon_t(D) = t(D) - \hat{t}(D)$ and σ_N^2 is the variance of the magnitude of the projection of the colored noise, $\eta(D)$, onto the vector $\varepsilon_t(D)$,

$$\begin{aligned} \sigma_N^2 &= E[|\mathrm{proj}_{\varepsilon_t} \eta|^2] \\ &= \frac{\varepsilon_t^T \mathbf{R}_{\eta,\eta} \varepsilon_t}{|\varepsilon_t|^2} \end{aligned}$$

where $[\mathbf{R}_{\eta,\eta}]_{i,j} = E[\eta_i \eta_j]$, the autocorrelation matrix of the noise process $\eta(D)$. If the additive noise were white (components of the noise vector are uncorrelated and identically distributed), the Euclidean decision would be optimal. However, in colored noise, this criterion is suboptimal, and may lead to a significant performance degradation.

III. Noise Predictive Maximum Likelihood Detector

An NPML detector seeks to improve the performance by whitening the noise sequence prior to the Viterbi Detector. In the process, it creates a new effective target. An estimate of the current noise sample, $\hat{\eta}_k$, is formed and subtracted from r_k. The estimate is the one-step linear predictor of the sequence, based on estimates of the previous N_p noise samples, $\{\tilde{\eta}\}_{k-1:k-N_p}$.

$$\hat{\eta}_k = \sum_{i=1}^{N_p} \tilde{\eta}_{k-i} p_i$$

The noise prediction coefficients, $\{p_i\}_{i=1}^{N_p}$, are chosen to minimize the error in the estimate $E[|\eta_k - \hat{\eta}_k|^2]$, where the estimates of the prior noise samples are formed from estimates of the data sequence, $\tilde{\eta}_{k-i} = r_{k-i} - \mathbf{h}^T \hat{\mathbf{a}}_{k-i:k-i-\nu}, 0 < i \leq N_p$. For computation of $\{p_i\}_{i=1}^{N_p}$, we assume the prior noise estimates are correct, $\tilde{\eta}_{k-i} = \eta_{k-i}$.

The NPML sequence estimator chooses the input sequence which, when passed through a channel with impulse response $g(D) = h(D)[1 - p(D)]$, yields the sequence closest in Euclidean distance to the equalized received sequence, $\gamma(D) = r(D)[1 - p(D)]$. However, the new target, $g(D)$, may be of a length whose full incorporation in the detector is prohibitively complex. A solution is to vary the complexity of the detector by limiting the number of distinct candidate paths considered. A set of 2^K states are retained, fewer than the $2^{N_p+\nu}$ states required to track all distinct candidate paths. The metrics in this Reduced State Sequence Estimator (RSSE) between states s_j and s_k are given by

$$\begin{aligned} \lambda_n(s_j, s_k) = \; & [\gamma_n - \textstyle\sum_{i=0}^{K} \hat{a}_{n-i}(s_j) g_i - \\ & \textstyle\sum_{i=K+1}^{N_p+\nu} \tilde{a}_{n-i}(s_j) g_i]^2 \end{aligned}$$

where the $\hat{a}_k$ are given by the 2^K states and the $\tilde{a}_k$ are given by the path histories. In [1], methods of choosing the path histories were investigated. Here we assume the use of full local feedback, such that each of the 2^K trellis states has a distinct path history.

IV. Performance of NPML Detector

In [6],[12], the performance of an RSSE detector is analyzed. An RSSE detector with $K < N_p + \nu$

suffers a significant performance loss relative to the full state detector, where $K = N_p + \nu$, by forcing a decision in the detector after fewer than $N_p + \nu$ consecutive steps with $\hat{a}_k = a_k$, thereby decreasing the effective minimum distance.

In addition, the residual interference due to errors in the path history increases the probability of an error event. The input to the NPML detector is $\gamma(D)$,

$$\begin{aligned} \gamma(D) &= a(D)g(D) + e(D) \\ &= \tau(D) + e(D) \end{aligned}$$

where $e(D)$ is the 'whitened' noise sequence $\eta(D)[1-p(D)]$. The Viterbi Algorithm on a trellis with $K = N_p + \nu$ makes a decision between $\tau(D) = a(D)g(D)$ and $\hat{\tau}(D) = \hat{a}(D)g(D)$ as illustrated in Figure 2. However, on the reduced state trellis a decision is made between $q(D)$ and $\hat{q}(D)$,

$$\begin{aligned} q(D) &= \tau(D) - \rho(D) \\ \hat{q}(D) &= \hat{\tau}(D) - \hat{\rho}(D) \end{aligned}$$

where $\rho(D)$ and $\hat{\rho}(D)$ are residual signal error sequences, reflecting the impact of errors prior to the start of the error event, i.e., errors in the path history $\{\tilde{a}_k\}$. If we assume that the error event extends from $k = k_1$ to $k = k_2$, then the residual error sequences are given by

$$\rho_k = \begin{cases} \sum_{i=k-k_1+1}^{N_p+\nu} (a_{k-i} - \tilde{a}_{k-i}) g_i & , k_1 \le k \le k_1 + N_p + \nu - 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$\hat{\rho}_k = \begin{cases} \sum_{i=k-k_1+1}^{N_p+\nu} (\hat{a}_{k-i} - \tilde{a}_{k-i}) g_i & , k_1 \le k \le k_1 + N_p + \nu - 1 \\ 0 & , \text{otherwise} \end{cases}$$

The effect of the residual interference in the signal space is illustrated in Figure 3. The residual interference effectively offsets the two candidate signal points such that the decision region is shifted. The probability of an error in the case of residual interference is

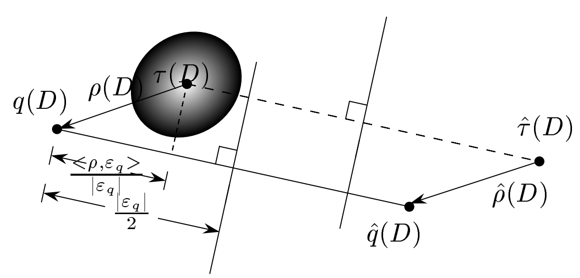

Fig. 3. Decision regions, RSSE ($K < N_p + \nu$)

$$Pr[a(D) \to \hat{a}(D) | \rho(D)] = Q\left(\frac{\frac{1}{2}|\varepsilon_q| + \frac{<\rho,\varepsilon_q>}{|\varepsilon_q|}}{\sigma_{N'}} \right)$$

where $\varepsilon_q(D) = q(D) - \hat{q}(D)$, $<\rho, \varepsilon_q> = \sum \rho_k \varepsilon_k$, and $\sigma_{N'}$ is the variance of the magnitude of the projection of the 'whitened' noise, $e(D)$, onto the vector $\varepsilon_q(D)$,

$$\begin{aligned} \sigma_{N'}^2 &= E[|\mathrm{proj}_{\varepsilon_q} \mathbf{e}|^2] \\ &= \frac{\varepsilon_q^T \mathbf{R}_{e,e} \varepsilon_q}{|\varepsilon_q|^2} \end{aligned}$$

where $[\mathbf{R}_{e,e}]_{i,j} = E[e_i e_j]$, the autocorrelation matrix of the noise process $e(D)$.

V. Error Event Characterization

Duel-Hallen and Heegard [6] illustrated that a good estimate of performance when K is close to $N_p + \nu$ is obtained by assuming $< \rho, \varepsilon_q > = 0$, i.e., no residual interference. We can identify the set of dominant error events on the RSSE detector trellis under this assumption, i.e., the set $\{\mathbf{a} - \hat{\mathbf{a}} \,|(|\varepsilon_q|^2/\sigma_{N'}^2) < \alpha\}$, where α is an appropriate constant. Search algorithms to accomplish this are described in [2],[1],[13]. The search uses an error state diagram with $3^{N_p+\nu}$ states, although closed error events on the RSSE trellis occur after K consecutive steps with $a_n - \hat{a}_n = 0$. Table I lists the results of a search for a sampled, ideal low-pass filtered Lorentzian channel operating at a density $\frac{PW_{50}}{T} = 2.86$ with a front-end target $h(D) = 1 - D^2$, a 4-tap noise predictor, 16-state detector trellis ($K = 4$), and normalized additive white noise power, $\sigma_n^2 = 1$. A shorthand vector notation is used to represent error events, where $\{+, 0, -\}$ denote $a_n - \hat{a}_n = +1, 0, -1$. A subset of the set $\{\mathbf{a} - \hat{\mathbf{a}} \,|(|\varepsilon_q|^2/\sigma_{N'}^2) < \alpha\}$ is determined by searching over all events up to length 20. We chose α to be the value of $|\varepsilon_q|^2/\sigma_{N'}^2$ corresponding to the Matched Filter Bound, i.e. the event $\mathbf{a} - \hat{\mathbf{a}} = \cdots 000 + 000 \cdots$. Zero cycles are denoted by $(\cdot)$, where (x) denotes insertion of any non-negative number of x's at that location. The value $|\varepsilon_q|^2/\sigma_{N'}^2$ listed for sets of events with a zero cycle corresponds to the shortest event of the set. Although this does not necessarily represent the minimum $|\varepsilon_q|^2/\sigma_{N'}^2$ for the set, the minimum is typically in the same range, and distinguishing the minimum would not have affected the design of a constrained code.

VI. Distance Enhancing Constraints

The performance of the detector may be improved by forbidding the occurrence of the lowest distance error events [13]. This is accomplished by constraining the binary input sequences such that the minimum distance pairwise error events are eliminated. We denote constrained sets of sequences by $\mathsf{X}_{\mathcal{F}}^{\mathcal{A}}$, where $\mathcal{A}$ is the sequence alphabet and $\mathcal{F}$ denotes the set of forbidden sequences over the alphabet. We are interested in two sets of constrained sequences and the relationships between them: the

TABLE I

DOMINANT ERROR EVENTS, LORENTZIAN CHANNEL, $\frac{PW_{50}}{T} = 2.86$, $h(D) = 1 - D^2$, $N_p = 4$, $K = 4$

$\left(\lvert\varepsilon_q\rvert^2/\sigma_{N'}^2\right)\rvert_{\sigma_n=1}$	$\mathbf{a} - \hat{\mathbf{a}}$
.34	+-+0000
.40	+-+-+0000
.44	+-+-+-+(-+)0000
.47	+0000
.47	+-0000
.47	+-+-+-(+-)0000

set of ternary error events (given by the difference of pairs of data sequences), $\mathsf{X}_{\mathcal{F}}^{\{+1,0,-1\}}$, and the set of constituent binary data sequences, $\mathsf{X}_{\mathcal{F}}^{\{0,1\}}$. We would like a constraint over the binary sequences with the highest code rate possible that forbids the dominant error events. Table I implies that a constraint $\mathsf{X}_{\{0+-+0,+-+-\}}^{\{+1,0,-1\}}$ is sufficient to guarantee that the constrained minimum distance is given by the Matched Filter Bound. One can show that

$$\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}} \Rightarrow \mathsf{X}_{\mathcal{F}}^{\{+1,0,-1\}} \subseteq \mathsf{X}_{\{0+-+0,+-+-\}}^{\{+1,0,-1\}}$$

Figure 4 illustrates the state diagram for the constraint $\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}}$. The Shannon Capacity of

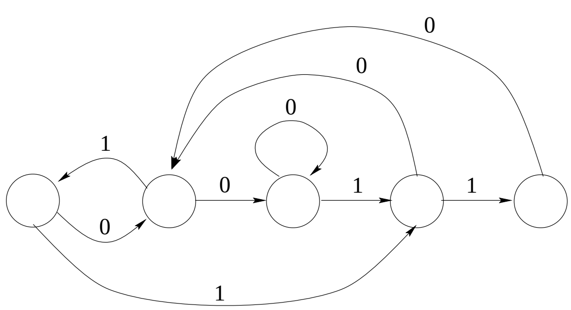

Fig. 4. State diagram, $\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}}$

this constraint, which is the limit on the encoding rate, is

$$\text{Shannon Capacity}[\mathsf{X}_{\{1111,00111\}_{\text{NRZI}}}^{\{0,1\}}] \approx .914$$

From Table I, it follows that this constraint should yield an asymptotic coding gain (ignoring the rate loss) of

$$\text{ACG} = 10\,\log(.47/.34) \approx 1.4\ dB$$

The dominant error events in Table I represent a subset of the dominant events in extended PRML systems [2],[10],[11]. Therefore, some of the distance-enhancing constraints for extended PRML systems, such as time-varying MTR constraints [4],[9], are potentially applicable in the NPML setting. Figure 5 illustrates the detector trellis for the $\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}}$ constrained system with $K = 4$. States are labeled $a_{k-4}a_{k-3}a_{k-2}a_{k-1}$ or $a_{k-4}a_{k-3}a_{k-2}a_{k-1}.a_k$ (if the next edge label is fixed), with input edge labels a_k. Output edge labels are given by $\sum_{i=0}^{4} \hat{a}_{n-i}(s_j)g_i - \sum_{i=5}^{N_p+\nu} \tilde{a}_{n-i}(s_j)g_i$. The structure of the $\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}}$ constrained trellis is not unique, and a variation could be used in implementation.

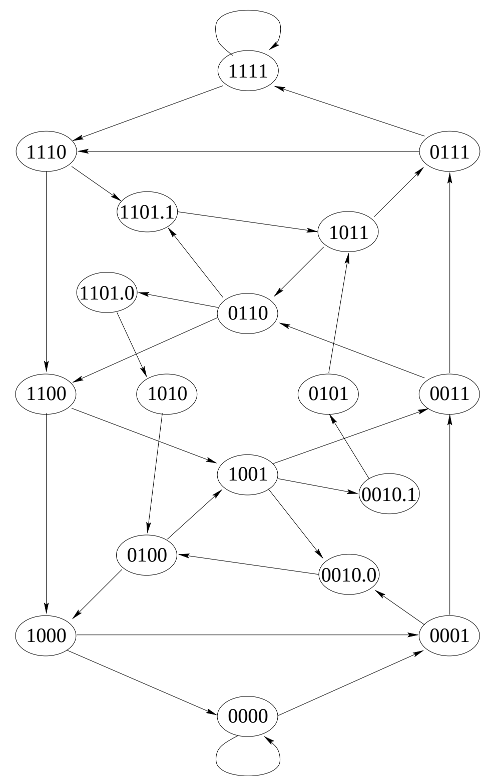

Fig. 5. $\mathsf{X}_{\{1111,00111\}_{NRZI}}^{\{0,1\}}$ constrained detector trellis, $K = 4$

VII. RESULTS

We simulated symbol-error-rate performance for the coded and uncoded NPML detectors. As the basis of the discrete-time model, the simulations used an ideal low-pass filtered Lorentzian channel model as illustrated in Figure 6. The input to the channel

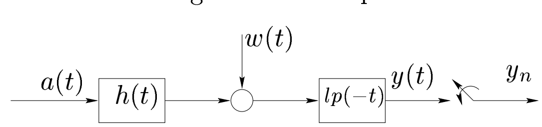

Fig. 6. Channel model

is a binary sequence, $\{a_n\}_{n=-\infty}^{\infty}, a_n \in \{0,1\}$, modulated by a unit pulse of duration T, $p_T(t)$,

$$p_T(t) = \begin{cases} 1 & ,0 \le t \le T \\ 0 & ,otherwise \end{cases}$$

$$a(t) = \sum_{k=-\infty}^{\infty} a_k p_T(t - kT).$$

The channel is defined in terms of its isolated step response, $g(t)$,

$$\begin{aligned} g(t) &= \frac{1}{1+(\frac{2t}{PW_{50}})^2} \\ h(t) &= g(t)-g(t-T) \end{aligned}$$

and the additive noise $w(t)$ is assumed zero mean, Gaussian with autocorrelation $E[w(t)w(s)] = \sigma_n^2\delta(t-s)$. The front-end filter, $lp(t)$, is an ideal low-pass filter, with cutoff frequency chosen such that there is no aliasing in the frequency response of the sampled process y_n.

The discrete channel model uses a 41-tap model of the ideal low-pass filtered Lorentzian channel ($N_l = 20$), a 21-tap MMSE front-end filter $\mathbf{v}$, a 4-tap noise predictor and truncation depth 20 in the NPML detector. The truncation depth was chosen in analogy with the rule of thumb from convolutional codes as $4K$. The constrained source inputs for the coded system are generated from the constraint graph with maxentropic transition probabilities, under the assumption that a highly efficient code will have similar statistics. The detectors are evaluated at a constant user bit density, where user bit density is the channel bit density divided by the rate, penalizing the coded systems by the rate loss. The rate is the highest $R = 4n/q$, with n, q integers, $n \leq 8$, that supports a '$0-1$' approximate eigenvector [15], thus guaranteeing a finite-state encoder based on a subgraph of the qth power of the constraint graph. 'Uncoded' detectors are penalized by a $R = 16/17$, which assumes the use of a run-length-limited code which provides negligible coding gain.

Figure 7 illustrates the simulation results. The SNR is defined as $E[(\mathbf{a} * \mathbf{l})_k^2]/\sigma_n^2$. We first note that the simulation results closely match estimates formed from the error event analysis. The uncoded NPML detector with $K = 4$ yields approximately 1.2 dB gain at a symbol error rate (SER) of 10^{-5} over a baseline EPR4 system, where the EPR4 system is a PRML detector with target $h(D) = (1-D)(1+D)^2$. By adding the $\mathsf{X}^{\{0,1\}}_{\{1111,00111\}_{NRZI}}$ constraint to the inputs of the channel, the coded NPML detector yields an additional .8 dB gain (at SER=10^{-5}) relative to the uncoded NPML detector.

Figure 8 illustrates the asymptotic gains of several coded and uncoded detectors relative to a baseline uncoded EPR4 detector. The NPML detector is the $K = 4$ detector described earlier, and the E^{2}PR4 detector is a PRML detector with target $h(D) = 1 + 2D - 2D^3 - D^4$. The gains are predicted from error event analysis and reflect SNR gain with SNR defined as above. Coded systems are evaluated at a higher channel bit density to account for the rate loss as mentioned above. The coding gain provided by the constraint is a function of the operating density. Outside the range of densities over which the constraint eliminates the minimum distance event, the rate loss of the constraint leads to a performance loss of the coded system relative to its uncoded counterpart. The figure illustrates that taking coding into account may affect the choice of a preferred target at a given density.

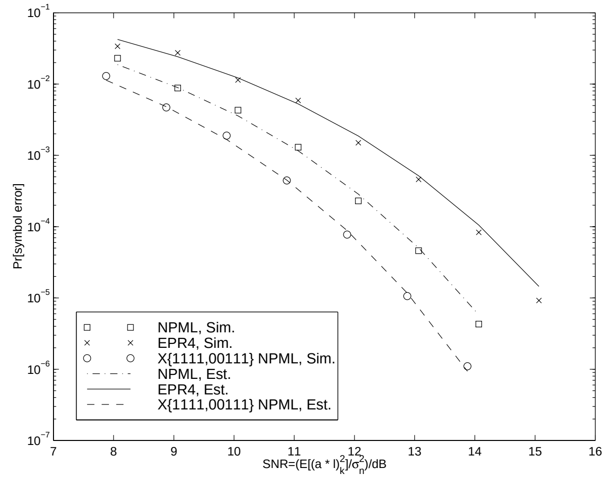

Fig. 7. Simulation results: Uncoded NPML detectors, K=4 and K=6, EPR4 detector, and coded NPML detector, $\frac{PW_{50}}{RT} = 2.54$

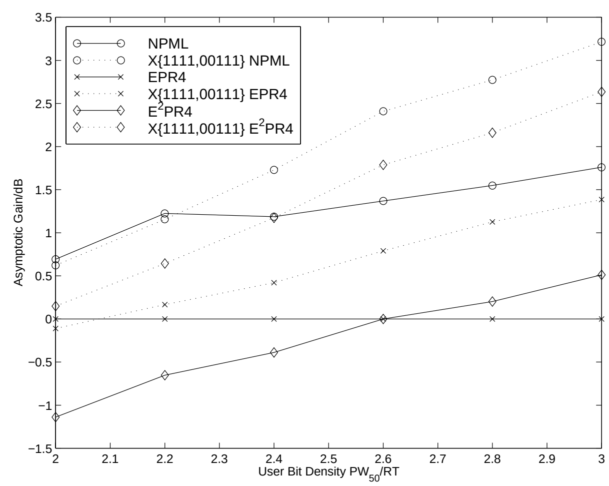

Fig. 8. Asymptotic gains relative to uncoded EPR4

VIII. Conclusions

The performance of an NPML/RSSE detector for K close to $N_p + \nu$ may be well approximated by characterizing the dominant error events for the detector under the assumption of no residual interference from errors in the path estimates. Distance-enhancing constraints may be determined in a straightforward manner from these error event lists. We have demonstrated through simulation the potential benefits of using distance-enhancing codes for an NPML/RSSE system with front-end target $h(D) = 1 - D^2$ and a 4-tap noise predictor. Taking such distance-enhancing coding into account affects the choice of the preferred target at a given density.

References

[1] S.A. Altekar, "Detection and coding techniques for magnetic recording channels," *Ph.D. dissertation*, University of California, San Diego, August, 1997.

[2] S. A. Altekar, M. Berggren, B. E. Moision, P. H. Siegel, and J. K. Wolf, "Error-event characterization on partial-response channels," *Proc. 1997 IEEE Int. Symp. on Info. Theory*, Ulm, Germany, June 29 – July 4, 1997, to appear in *IEEE Trans. Inform. Theory.*

[3] L. C. Barbosa, "Maximum likelihood sequence estimators: a geometric view," *IEEE Trans. Inform. Theory*, vol. 35, no. 2, pp. 419–427, March, 1989.

[4] W. Bliss, "An 8/9 rate time-varying trelis code for high density magnetic recording," *IEEE Trans. Magn.*, vol. 33, no. 5, pp. 2746–2748, Sept. 1997.

[5] P.R. Chevillat, E. Eleftheriou, and D. Maiwald, "Noise predictive partial-response equalizers and applications," *IEEE Conf. Records ICC'92*, pp.942–947, June 14–18, 1992.

[6] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.

[7] E. Eleftheriou and W. Hirt, "Noise-predictive maximum-likelihood (NPML) detection for the magnetic recording channel," *Proc. ICC '96*, vol. 1 , pp. 556–560, June 1996.

[8] V. M. Eyuboglu and S. U. Quereshi, "Reduced-state sequence estimation with decision feedback and set partitioning," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.

[9] K. Knudson Fitzpatrick and C.S. Modlin, "Time-varying MTR codes for high density magnetic recording," *1997 IEEE Global Telecommun. Cof. (GLOBECOM '97)*, Phoenix, AZ, Nov. 4–8, 1997, Conference Record, vol. 3, pp. 1250–1253.

[10] R. Karabed, P. Siegel, and E. Soljanin, "Constrained coding for binary channels with high intersymbol interference," accepted for publication, *IEEE Trans. Inform. Theory.*

[11] R. Karabed and P.H. Siegel, "Coding for higher-order partial-response channels," in *Coding and Signal Processing for Information Storage,* Mysore R. Raghuveer, Soheil A. Dianat, Steven W. McLaughlin, Martin Hassner, Editors, *Proc. 1995 SPIE Int. Symp. on Voice, Video, and Data Communications*, Philadelphia, PA, Oct. 1995, vol. 2605, pp. 115–126.

[12] Won U. Lee and F.S. Hill, "A maximum-likelihood sequence estimator with decision-feedback equalization," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 971–979, Sept. 1977.

[13] B. E. Moision, P. Siegel, and E. Soljanin, "Error event characterization and coding for the Lorentzian channel," to be presented at *1998 IEEE Int. Symp. Inform. Theory*, Cambridge, MA, Aug. 1998.

[14] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Trans. Magn.*, vol. 32, no. 5, pt. 1, pp. 3992–3994, Sept. 1996.

[15] B. H. Marcus, P. H. Siegel, and J. K. Wolf, "Finite-state modulation codes for data storage," *IEEE J. Sel. Areas Commun.* vol. 10, no. 1, pp. 5–37, Jan. 1992.