International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

Multidimensional Receiver Front-End for Storage Systems with Data-Dependent Transition Noise

Riccardo Pighi, Riccardo Raheli and Umberto Amadei

Università di Parma,
Dipartimento di Ingegneria dell'Informazione,
Parco Area delle Scienze 181A,
43100 Parma, Italy
E-mail: pighi@tlc.unipr.it, raheli@unipr.it

Abstract

In the last decade, research on detection algorithms capable of mitigating the effects of colored Gaussian thermal noise and transition noise in storage systems, has proposed a number of solutions. In this paper, we present a new detection scheme based on a multidimensional receiver front-end and linear prediction applied to Maximum A-Posteriory Probability (MAP) sequence detection, which improves the Bit Error Rate (BER) performance with respect to previous solutions.

1. Introduction

Recording densities in magnetic storage systems continue to increase at a considerable rate. These high recording densities require sophisticated detection schemes in order to preserve system reliability. High-density longitudinal and vertical magnetic recording systems based on thin-film media exhibit severe intersymbol interference, colored Gaussian thermal noise and signal-dependent transition noise. The last kind of noise, also known as media noise, increases with density and is due to the magnetic interaction between data transitions in the information sequence stored on the medium: therefore transition noise is data-dependent. In the literature, a few channel models have appeared to facilitate the analysis and design of the optimum detector, such as the microtrack model [1], the signal-dependent autoregressive channel model [2] and the position jitter and width variation model [3, 4]. The latter is used in this paper.

After the definition of a suitable channel model, many authors proposed detection schemes based on signal processing algoritms to reduce the effects of noise in magnetic recording channel: in [5] a detection scheme based on linear prediction was applied to colored thermal noise, and in [6] linear prediction was extended to signal-dependent transition noise. According to the model in [3, 4], the observable can be viewed as conditionally Gaussian, given the data, and one is enabled to exploit the principle of linear predictive detectors proposed for fading channels [7, 8].

In this paper, we extend the results in [6] to a receiver based on a multidimensional front-end for magnetic recording channels. In more detail, the presence of transition noise and the need for statistical sufficiency yield a receiver front-end with a number of filters proportional to the modelling order of the transition noise. Numerical analysis and simulations have demonstrated good improvements in terms of minimum mean square prediction error (MMSPE) and bit error rate (BER) of receivers using a front-end based on two or three filters, with respect to conventional receivers.

2. Channel Model

We consider a magnetic recording channel modeled by a first-order position jitter and width variation [4]. Let $h(t, w)$ denote the pulse response to an isolated magnetic transition recorded in a thin-film longitudinal or vertical media where t is time and w is half the width of the pulse at half height. Let $a_k \in \{\pm 1\}$ be the information bits to be stored. Assuming that transition noise can be decomposed into position jitter and width variation, the read back waveform $r(t)$ corrupted by additive white Gaussian thermal noise $w(t)$ can be expressed as

$$r(t) = \sum_k b_k h(t + \Delta t_k - kT, w + \Delta w_k) + w(t) \qquad (1)$$

where $b_k = a_k - a_{k-1} \in \{0, \pm 2\}$ denote transition symbols, Δt_k and Δw_k, modeled as independent Gaussian random variables with standard deviations $\sigma_{\Delta t}$ and $\sigma_{\Delta w}$, represent the effect of position jitter and width variation noise, respectively, and T is the symbol period. Obviously, when $\sigma_{\Delta t} = 0$ and $\sigma_{\Delta w} = 0$, the model reduces to a magnetic recording channel without transition noise. For the pulse response $h(t, w)$ we have adopted the well known Lorentzian approximation [9] for longitudinal recording, i.e.

$$h(t, w) = \frac{1}{1 + (2t/PW_{50})}$$

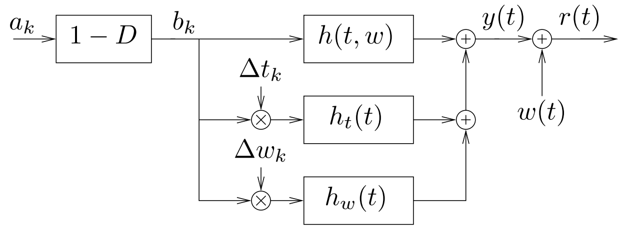

Figure 1: *Lorentzian channel model with first-order media noise and additive white Gaussian thermal noise.*

where PW_{50} is the pulsewidth at half the maximum amplitude and $PW_{50} = 2w$. We define the parameter $D = PW_{50}/T$ as the normalized density. According to the first-order channel model, the read back impulse can be approximated as

$$h(t + \Delta t_k, w + \Delta w_k) \simeq h(t,w) + \Delta t_k \frac{\partial h(t,w)}{\partial t} + \Delta w_k \frac{\partial h(t,w)}{\partial w} .$$

Defining the impulse response of the filters modelling the position jitter and width variation noise process as[1]

$$\frac{\partial h(t,w)}{\partial t} = h_t(t) \qquad \frac{\partial h(t,w)}{\partial w} = h_w(t)$$

and using this first-order approximation in (1), the continuous waveform at the output of the channel can be written as $r(t) = y(t) + w(t)$, where we have defined $y(t)$ as

$$y(t) = \sum_k \big[b_k h(t-kT) + b_k \Delta t_k h_t(t-kT) + b_k \Delta w_k h_w(t-kT) + w(t) \big] . \quad (2)$$

A block diagram descriptive of this channel model is shown in Fig. 1.

3. Sufficient Statistics

We now derive a set of sufficient statistics for the considered magnetic recording channel affected by transition noise. The signal at the output of the channel can be expressed as

$$r(t) = y(t, \mathbf{a}, \boldsymbol{\theta}) + w(t) \quad (3)$$

where $\mathbf{a}$ is the data information vector, $\boldsymbol{\theta}$ is a random vector collecting the unknown parameters affecting the observable, i.e. the sequences of random variables $\{\Delta t_k\}$ and $\{\Delta w_k\}$, $w(t)$ is an additive white Gaussian thermal noise process and $y(t, \mathbf{a}, \boldsymbol{\theta})$ is defined[2] as in (2). Given a probabilistic model of $\boldsymbol{\theta}$ with realizations in a suitable space Θ and noting that, for any finite number of transmitted bits, an information lossless discretization of signal $r(t)$ by expansion over an orthonormal finite-dimensional basis can be achieved, the detection strategy can be formulated as

$$\hat{\mathbf{a}} = \operatorname*{argmax}_{\mathbf{a}} P(\mathbf{a}) f(\mathbf{r}|\mathbf{a}) \quad (4)$$

where $P(\mathbf{a})$ is the a-priori probability of the information sequence $\mathbf{a}$ and $f(\mathbf{r}|\mathbf{a},)$ is the conditional probability density function (pdf) of the observation vector $\mathbf{r}$, given the information sequence $\mathbf{a}$. Under the assumption of statistical independence between $\boldsymbol{\theta}$ and $\mathbf{a}$, the conditional probability density function in (4) can be expressed as

$$f(\mathbf{r}|\mathbf{a}) = \int_\Theta f(\mathbf{r}|\boldsymbol{\theta}, \mathbf{a}) f(\boldsymbol{\theta})\, d\boldsymbol{\theta} \quad (5)$$

in which the integral is over the parameters space Θ and $f(\boldsymbol{\theta})$ is the probability density function of vector $\boldsymbol{\theta}$. Since the random variables Δt_k and Δw_k are Gaussian, the observation vector $\mathbf{r}$ is also conditionally Gaussian, given the data. Therefore it is possible to express the conditional pdf of the observation vector as

$$f(\mathbf{r}|\boldsymbol{\theta}, \mathbf{a}) = \frac{1}{(\pi N_0)^{\frac{N}{2}}} \mathrm{e}^{-\frac{1}{N_0} \|\mathbf{r} - \mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\|^2} \sim \frac{1}{(\pi N_0)^{\frac{N}{2}}} \mathrm{e}^{-\frac{1}{N_0} \left[\|\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\|^2 - 2\mathrm{Re}\{\mathbf{r}^T \mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\} \right]} \quad (6)$$

where $\|.\|$ denotes the Euclidean norm, the quantity $\|\mathbf{r}\|^2$ is irrelevant in the detection process and can be discarded, $\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})$ is the discretization of $y(t, \mathbf{a}, \boldsymbol{\theta})$ and the symbol $\sim$ denotes a monotonical relationship with respect to the variable on interest (i.e., the data sequence $\mathbf{a}$). Using (6) in (5), we obtain

$$f(\mathbf{r}|\mathbf{a}) \sim \frac{1}{(\pi N_0)^{\frac{N}{2}}} \cdot \int_\Theta \mathrm{e}^{-\frac{1}{N_0} \left[\|\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\|^2 - 2\mathrm{Re}\{\mathbf{r}^T \mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\} \right]} f(\boldsymbol{\theta})\, d\boldsymbol{\theta} .$$

The discrete correlation between the observation vector $\mathbf{r}$ and the signal vector $\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})$ can be equivalently expressed in the time domain, thanks to the optimal discretization procedure, as a correlation integral

$$\mathbf{r}^T \mathbf{y}(\mathbf{a}, \boldsymbol{\theta}) = \int_{-\infty}^{\infty} r(t) y(t; \mathbf{a}, \boldsymbol{\theta})\, dt . \quad (7)$$

Similarly, the square Euclidean norm pf $\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})$ is equal to the energy of the signal

$$\|\mathbf{y}(\mathbf{a}, \boldsymbol{\theta})\|^2 = \int_{-\infty}^{\infty} y(t; \mathbf{a}, \boldsymbol{\theta})\, dt .$$

[1] The subscript denotes the variable of differentiation.

[2] The used notation highlights the dependence of $y(t)$ on the random vector $\boldsymbol{\theta}$ and the information vector $\mathbf{a}$.

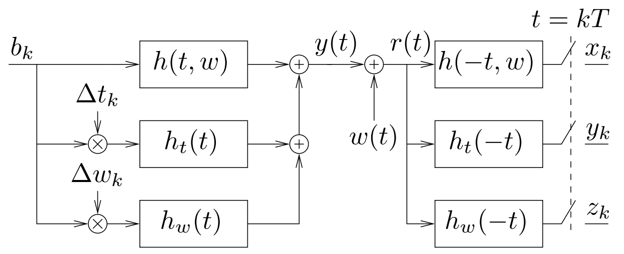

Figure 2: *Multidimensional receiver front-end for the Lorentzian channel model with first-order media noise and additive white Gaussian noise.*

Using (2), the correlation integral in (7) can be expressed as

$$\begin{aligned}\int_{-\infty}^{\infty} r(t) y(t; \mathbf{a}, \boldsymbol{\theta})\, dt \\ = \sum_k \Bigg[\int_{-\infty}^{\infty} b_k r(t) h(t - kT)\, dt \\ + b_k \Delta t_k \int_{-\infty}^{\infty} r(t) h_t(t - kT)\, dt \\ + b_k \Delta w_k \int_{-\infty}^{\infty} r(t) h_w(t - kT)\, dt \Bigg] . \quad (8)\end{aligned}$$

Defining the quantities

$$\begin{aligned} x_k &= \int_{-\infty}^{\infty} r(t) h(t - kT)\, dt \\ y_k &= \int_{-\infty}^{\infty} r(t) h_t(t - kT)\, dt \\ z_k &= \int_{-\infty}^{\infty} r(t) h_w(t - kT)\, dt \end{aligned}$$

equation (8) becomes

$$\begin{aligned}\int_{-\infty}^{\infty} r(t) y(t; \mathbf{a}, \boldsymbol{\theta})\, dt \\ = \sum_k b_k x_k + \sum_k b_k \Delta t_k y_k + \sum_k b_k \Delta w_k z_k . \quad (9)\end{aligned}$$

Equation (9) shows that $(x_k, y_k, z_k)^T$ is a vector of sufficient statistics for the detection process in a magnetic recording channel with data-dependent transition noise, according to the model of Fig. 1. These sufficient statistics can be obtained from the time-continuous received signal $r(t)$ by means of a multidimensional receiver front-end, as illustrated in Fig. 2. The proposed front-end is based on a bank of filters, each followed by a sampler at the symbol rate: the first filter is the usual matched filter, whereas the second and the third filters are matched to the second and third impulse responses modelling transition noise, respectively. We remark that commonly-used front-ends are based on the matched filter $h(-t, w)$ only, whereas in the presence of transition noise the discrete observation sequence $\{x_k\}$ is not a sufficient statistic. An intuitive explanation of the fact that $\{x_k\}$ is not a sufficient statistic in the presence of transition noise may be based on at its moltiplicative nature with respect to the transition sequence $\{b_k\}$. In fact, the transition noise waveform can be viewed as a noise signal corrupting information bits, but also as another kind of information-bearing signal superimposed to the useful signal. With the proposed front-end, we are able to extract this residual information from the observation of $r(t)$ and to use it in order to improve system reliability, i.e. to improve bit error rate performance.

It has to be highlighted that our derivation of the sufficient statistics can be straightforwardly extended to a magnetic channel model with transition noise of higher order: the number of filters in the multidimensional front-end is controlled by the degree of approximation of the transition noise process. For example, a second order channel model would result in a receiver front-end composed of a bank of 5 matched filters.

Linear prediction can be applied to estimate the realization of the transition noise process in order to incorporate its realization into a Viterbi algorithm and enable maximum a-posteriori sequence detection (MAP). With respect to the algorithms proposed in the literature [5, 6], which deal with colored Gaussian thermal noise and transition noise, respectively, and are based on the observation of $\{x_k\}$ only, we can now operate on a wider set of useful samples, allowing the proposed receiver to outperform the previously proposed detection schemes.

Since the relevant impulse responses are caracterized by a great amount of intersymbol interference, in order to reduce the complexity of the receiver, instead of adopting a partial response equalizer with the purpose of channel shortening, a bank of Whitening Filters (WF) matched to the impulse responses of the multidimensional front-end can be used. We found that the whitening process reduces the dispersion of the impulse response of the information-bearing signal as well as the length of the position jitter and width variation modelling impulses. These whitening filters decorrelate the thermal noise samples in the space domain at time k only (i.e. they decorrelate the samples at the output of the three matched filters), but they are still correlated in time and space. This is not a limiting factor because linear prediction can cope with this correlation.

Having obtained a set of sufficient statistics, we can now investigate how this quantities can be used to perform sequence detection: this issue is addressed in the next section.

4. Detection Strategy based on Linear Prediction

Assuming a first-order channel model, we have shown that the quantities (x_k, y_k, z_k) are sufficient statistics for sequence detection. Collecting these quantities into suitable vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ at the output of the multidimensional receiver front-end, we can reformulate the detection strategy (4) in equivalent form as

$$\hat{\mathbf{a}} = \operatorname*{argmax}_{\mathbf{a}} P(\mathbf{a}) f(\mathbf{x}, \mathbf{y}, \mathbf{z}|\mathbf{a}) . \qquad (10)$$

Assumying causality and applying the chain factorization rule to the multidimensional conditional pdf, we obtain

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}, \mathbf{z}|\mathbf{a}) &= \prod_{k=0}^{K-1} f(x_k, y_k, z_k|\mathbf{x}_0^{k-1}, \mathbf{y}_0^{k-1}, \mathbf{z}_0^{k-1}; \mathbf{a}_0^k) \\ &= \prod_{k=0}^{K-1} \Big[f(x_k|\mathbf{x}_0^{k-1}, \mathbf{y}_0^k, \mathbf{z}_0^k; \mathbf{a}_0^k) \\ &\qquad \cdot f(y_k|\mathbf{x}_0^{k-1}, \mathbf{y}_0^{k-1}, \mathbf{z}_0^k; \mathbf{a}_0^k) \\ &\qquad \cdot f(z_k|\mathbf{x}_0^{k-1}, \mathbf{y}_0^{k-1}, \mathbf{z}_0^{k-1}; \mathbf{a}_0^k) \Big] \\ &\simeq \prod_{k=0}^{K-1} \Big[f(x_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^k, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k) \\ &\qquad \cdot f(y_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1} \mathbf{z}_{k-\nu}^k; a_k, \zeta_k) \\ &\qquad \cdot f(z_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1} \mathbf{z}_{k-\nu}^{k-1}; a_k, \zeta_k) \Big] \end{aligned} \qquad (11)$$

where $\mathbf{x}_{k_1}^{k_2}$ is a shorthand notation for the vector collecting signal observations from time epoch k_1 to k_2 and K is the length of the transmission. In the last step of (11), in order to limit the receiver's memory, we have assumed Markovianity of order ν in the observation sequences. Moreover we have defined the state of the system accounting for the postcursors, precursors and the order of Markovianity ν as

$$\zeta_k = (a_{k-1}, a_{k-2}, a_{k-3}, \ldots, a_{k-L}) \qquad (12)$$

where $L = \delta_1 + \delta_2 + \nu$, with δ_1 and δ_2 denoting the number of precursors and postcursors in the impulse responses. The assumed Markovianity results in an approximation whose quality increases with ν.

Keeping in mind that, since the thermal and the transition noise processes have Gaussian distribution, the observation is Gaussian, given the data. The application of the chain factorization rule allows us to split the multidimensional conditional pdf in (10) in a product of $3K$ one dimensional conditional Gaussian pdf, completely defined by the conditional means

$$\begin{aligned} \hat{x}_k &= E\{x_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^k, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ \hat{y}_k &= E\{y_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ \hat{z}_k &= E\{z_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^{k-1}; a_k, \zeta_k\} \end{aligned}$$

and the conditional variances

$$\begin{aligned} \hat{\sigma}_{x_k}^2 &= E\{[x_k - \hat{x}_k]^2 \,|\, \mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^k, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ \hat{\sigma}_{y_k}^2 &= E\{[y_k - \hat{y}_k]^2 \,|\, \mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ \hat{\sigma}_{z_k}^2 &= E\{[z_k - \hat{z}_k]^2 \,|\, \mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^{k-1}; a_k, \zeta_k\} . \end{aligned}$$

It should be now clear that $\hat{x}_k, \hat{y}_k$ and $\hat{z}_k$ can be interpreted as linear predictive estimations of x_k, y_k, and z_k, respectively and $\hat{\sigma}_{x_k}^2, \hat{\sigma}_{y_k}^2$ and $\sigma_{z_k}^2$ as the relevant minimum mean square prediction error. It is also possible to express explicity the conditional mean values as

$$\begin{aligned} \hat{x}_k &= E\{x_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^k, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ &= s_{1,k}(a_k, \zeta_k) \\ &\quad + \sum_{i=1}^{\nu} p_{1,1,i}(a_k, \zeta_k)[x_{k-i} - s_{1,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=0}^{\nu} p_{1,2,i}(a_k, \zeta_k)[y_{k-i} - s_{2,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=0}^{\nu} p_{1,3,i}(a_k, \zeta_k)[z_{k-i} - s_{3,k-i}(a_k, \zeta_k)] \\ \hat{y}_k &= E\{y_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^k; a_k, \zeta_k\} \\ &= s_{2,k}(a_k, \zeta_k) \\ &\quad + \sum_{i=1}^{\nu} p_{2,1,i}(a_k, \zeta_k)[x_{k-i} - s_{1,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=1}^{\nu} p_{2,2,i}(a_k, \zeta_k)[y_{k-i} - s_{2,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=0}^{\nu} p_{2,3,i}(a_k, \zeta_k)[z_{k-i} - s_{3,k-i}(a_k, \zeta_k)] \\ \hat{z}_k &= E\{z_k|\mathbf{x}_{k-\nu}^{k-1}, \mathbf{y}_{k-\nu}^{k-1}, \mathbf{z}_{k-\nu}^{k-1}; a_k, \zeta_k\} \\ &= s_{3,k}(a_k, \zeta_k) \\ &\quad + \sum_{i=1}^{\nu} p_{3,1,i}(a_k, \zeta_k)[x_{k-i} - s_{1,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=1}^{\nu} p_{3,2,i}(a_k, \zeta_k)[y_{k-i} - s_{2,k-i}(a_k, \zeta_k)] \\ &\quad + \sum_{i=1}^{\nu} p_{3,3,i}(a_k, \zeta_k)[z_{k-i} - s_{3,k-i}(a_k, \zeta_k)] . \end{aligned}$$

In the definition of the conditional means, $s_{1,k}(a_k, \zeta_k)$, $s_{2,k}(a_k, \zeta_k)$ and $s_{3,k}(a_k, \zeta_k)$ are the information-bearing signal at the output of the first, the second and the third matched filter, respectively, and the coefficients $p_{\alpha,\beta,i}$ at time i are the solution of the Wiener-Hopf matrix equation [10], with index α denoting the branch number in the channel model of Fig.1 (from top to bottom) and β denoting the branch number in the receiver front-end.

Therefore, given the detection strategy (10) and the factorization (11) and taking the logarithm, we can express the branch metrics of a Viterbi detector as

$$\begin{aligned}\lambda_k(a_k,\zeta_k) \;=\;& \ln f(x_k|\mathbf{x}_{k-\nu}^{k-1},\mathbf{y}_{k-\nu}^{k},\mathbf{z}_{k-\nu}^{k};a_k,\zeta_k)\\ &+\ln f(y_k|\mathbf{x}_{k-\nu}^{k-1},\mathbf{y}_{k-\nu}^{k-1},,\mathbf{z}_{k-\nu}^{k};a_k,\zeta_k)\\ &+\ln f(z_k|\mathbf{x}_{k-\nu}^{k-1},\mathbf{y}_{k-\nu}^{k-1},,\mathbf{z}_{k-\nu}^{k-1};a_k,\zeta_k)\\ &+\ln P[a_k(\zeta_k)]\,.\end{aligned}$$

Assuming that the information bits are indipendent and equally distributed, the branch metrics can be finally expressed as

$$\begin{aligned}\lambda_k(a_k,\zeta_k) = &-\frac{[x_k-\hat{x}_k]^2}{\hat{\sigma}_{x_k}^2}-\ln\hat{\sigma}_{x_k}^2-\frac{[y_k-\hat{y}_k]^2}{\hat{\sigma}_{y_k}^2}-\ln\hat{\sigma}_{y_k}^2\\ &-\frac{[z_k-\hat{z}_k]^2}{\hat{\sigma}_{z_k}^2}-\ln\hat{\sigma}_{z_k}^2\,.\end{aligned}$$

The state-complexity of a linear prediction receiver can be naturally decoupled from the prediction order ν by means of state-reduction techniques. Let $Q < L$ denote the memory parameter to be taken into account in the definition of a reduced trellis state

$$\omega_k = (a_{k-1}, a_{k-2}, \ldots, a_{k-Q})\,.$$

The branch metric can be obtained by defining a pseudo state

$$\begin{aligned}\tilde{\zeta}_k(\omega_k) = (&\underbrace{\overbrace{a_{k-1},\ldots,a_{k-Q}}^{\omega_k}}_{Q\text{ bits}},\underbrace{\breve{a}_{k-Q-1}(\omega_k),\ldots,\breve{a}_{k-Q-P}(\omega_k)}_{P\text{ bits}},\\ &\underbrace{\hat{\hat{a}}_{k-Q-P-1},\ldots,\hat{\hat{a}}_{k-L}}_{L-Q-P\text{ bits}})\,.\end{aligned} \tag{13}$$

where P bits may be choosen by a per-survivor processing technique [11], and the last $L-Q-P$ bits can be defined as tentative (or preliminary) decisions $\hat{\hat{a}}_k$ at the detector output. Note that $\breve{a}_{k-Q-1}(\omega_k),\ldots,\breve{a}_{k-Q-P}(\omega_k)$ are the information bits associated with the survivor of ω_k. The branch metric in the reduced-state trellis can be defined in terms of the pseudo state (13) as

$$\tilde{\lambda}_k(a_k,\omega_k) = \lambda_k(a_k,\tilde{\zeta}_k(\omega_k))\,.$$

5. Simulation Results

In Fig. 3 we show curves of prediction error obtained both for a monodimensional front-end (indicated by 1D), for a bidimensional front-end (2D) and for a three dimensional front-end (3D), for a normalized density

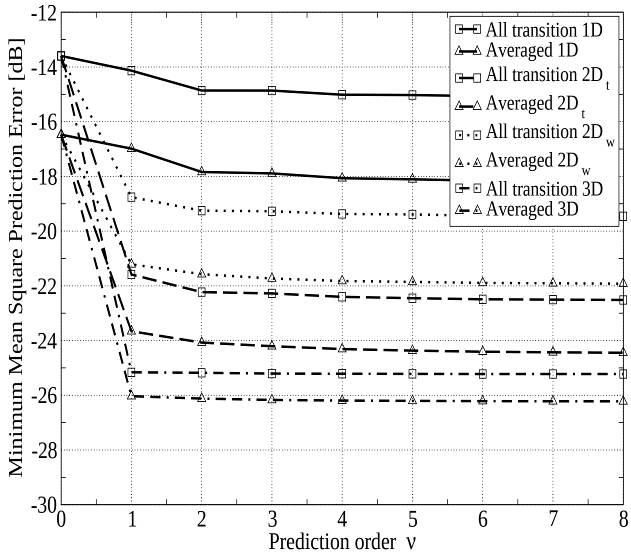

Figure 3: *Minimum Mean Square Prediction Error versus number of prediction coefficients* ν.

$D = 2.50$. In the bidimensional case, the second front-end filter was selected to be matched to either the time derivate of the Lorentzian pulse (2D$_t$ curves) or the width derivate of the Lorentzian pulse (2D$_w$ curves). Since media noise arises in transitions, a bit pattern characterized by continuous changes of the writing current's polarity, i.e. $\{1,-1,1,-1,\ldots\}$. An averaged MMSPE is also shown, by averaging over all possible bit patterns defining a trellis branch (a_k,ζ_k).

The SNR with transition noise [12] is defined at the input of the Viterbi detector for a one dimensional front-end as

$$\text{SNR}_\alpha = \frac{P_s}{\sigma_n^2+\sigma_m^2} \tag{14}$$

where P_s is the signal power, σ_n^2 is the thermal noise power, σ_m^2 is the transition noise power measured with an all-transition bit pattern $\pm\{+1,-1,+1,-1,\ldots\}$ and $\alpha = 100\times[\sigma_m^2/(\sigma_n^2+\sigma_m^2)]$. In order to evaluate the MMSPE, the signal to noise ratio was fixed at $\text{SNR}_{95} = 10$ dB, i.e. assuming a 95% transition noise consisting of 50% position jitter and 50% width variation, when measured with the "all-transition" bit sequence. Fig. 3 shows that, with the use of a bidimensional front-end and a prediction order $\nu = 2$, it is possible to obtain a gain in the MMSPE varying between 4.5 dB to 7.0 dB for the all-transition bit pattern, with respect to the gain obtained by a monodimensional receiver front-end (1D curves). Finally, looking at the MMSPE obtained for the complete multidimensional receiver, it is possible to outperform the results of the bidimensional front-end by at least 1.5 dB.

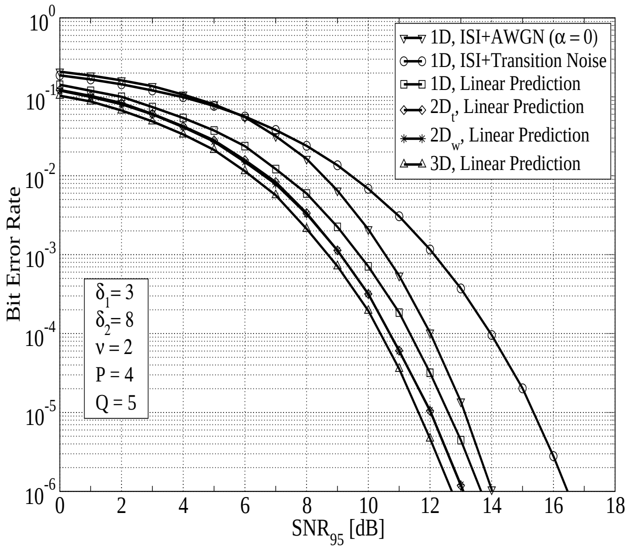

Figure 4: *Bit error probability for the Lorentzian channel model with transition noise.*

Fig. 4 shows the BER obtained for a monodimensional front-end without transition noise (ISI+AWGN curve), with transition noise but without linear prediction (ISI+Transition Noise curve) and with both transition noise and linear prediction (Linear Prediction curve), for a density $D = 2.50$ and a 95% transition noise (50% position jitter and 50% width variation). These curves are obtained with $\delta_1 = 3$ precursors, $\delta_2 = 8$ postcursors, a prediction order $\nu = 2$, per-survivor processing with state-reduction parameter $P = 4$ and 4 preliminar decisions ($L - P - Q = 4$). Therefore, the Viterbi algorithm searches a trellis diagram with $2^{\delta_1+\nu} = 32$ states. The BER curve obtained with linear prediction and a monodimensional front-end shows an SNR gain of almost 2.5 dB, with respect to the one obtained without linear prediction. With a bidimensional receiver front-end the BER (Linear Prediction 2D curves) outperforms the 1D linear prediction curve by approximately 0.6 dB (note that the 2D$_t$ curve lays upon the 2D$_w$ one), while with the multidimensional receiver the gain, with respect to the ISI+Transition Noise curve , is nearly 3.5 dB.

6. Conclusions

A set of sufficient statistics for the magnetic recording channel in the presence of data-dependent transition noise has been proposed. These sufficient statistics can be obtained through a multidimensional receiver front-end. Multidimensional linear prediction can be used to modify the Viterbi branch metric in order to improve the detector performance and make it very insensitive to transition noise.

References

[1] J. Caroselli and J. W. Wolf, "Application of a new simulation model for media noise limited magnetic recording channels," *IEEE Trans. Magnetics*, vol. 32, no. 5, pp. 3917–3919, September 1996.

[2] A. Kavcic and A. Patapoutian, "A signal-dependent autoregressive channel model," *IEEE Trans. Magnetics*, vol. 35, no. 5, pp. 2316–2318, September 1999.

[3] J. Moon, "Discrete-time modeling of transition-noise-dominant channels and study of detection performance," *IEEE Trans. Magnetics*, vol. 27, no. 6, pp. 4573–4578, November 1991.

[4] T. Oenning and J. Moon, "Modeling the lorentzian magnetic recording channel with transition noise," *IEEE Trans. Magnetics*, vol. 37, no. 1, pp. 583–591, January 2001.

[5] E. Eleftheriou and W. Hirt, "Improving Performance of PRML/EPRML through Noise Prediction," *IEEE Trans. Magnetics*, vol. 32, no. 5, pp. 3968–3970, September 1996.

[6] J. Moon and J. Park, "Pattern-dependent noise prediction in signal-dependent noise," *IEEE J. Select. Areas Commun.*, vol. 19, no. 4, pp. 730–743, April 2001.

[7] J. H. Lodge and M. L. Moher, "Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 38, no. 6, pp. 787–794, June 1990.

[8] G. Vitetta and U. Mengali, "Double-filtering receivers for PSK signals transmitted over Rayleigh frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 44, no. 6, pp. 686–695, June 1996.

[9] J. Proakis, *Digital Communications*. McGraw-Hill, 2001.

[10] J. Makhoul, "Linear Prediction: A Tutorial Review," *Proc. IEEE*, vol. 63, no. 4, April 1975.

[11] R. Raheli, A. Polydoros, and C. K. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 41, no. 2, pp. 354–364, February 1995.

[12] J. Moon, "Signal-to-Noise Ratio Definition for Magnetic Recording Channels with Transition Noise," *IEEE Trans. Magnetics*, vol. 36, no. 5, pp. 3881–3883, September 2000.