

Oversampled Detector for Transition-Noise-Limited Digital Storage Systems

Riccardo Pighi, Umberto Amadei and Riccardo Raheli
 Dept. of Information Engineering, University of Parma, Italy
 Email: {pighi,amadei}@tlc.unipr.it, raheli@unipr.it

Abstract—In the last decade, significant research has been performed on detection algorithms capable of mitigating the effects of colored Gaussian thermal noise and transition noise in digital storage systems. In this paper, we present a new Maximum A-Posteriori Probability (MAP) sequence detection scheme based on oversampling and linear prediction. The proposed solution improves the Bit Error Rate (BER) performance with respect to conventional systems and makes the detector more robust against transition noise. The results obtained for a magnetic channel can be also extended to optical storage systems.

I. INTRODUCTION

High-density longitudinal and perpendicular 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, is due to the magnetic interaction between data transitions in the information sequence stored on the medium: therefore transition noise is data-dependent and its power increases with the recording density.

In the literature, a few channel models have been proposed to enable the analysis and design of optimum detectors [1], [2]. In this paper the position jitter and width variation model [3], [4] is used. Based on this channel model, a multidimensional detection front-end based on a number of filters proportional to the modeling order of the transition noise was presented in [5], [6]. Extending previous work on optimal prefiltering [7], we now propose a new set of sufficient statistics which can be extracted from the continuous read back waveform at the output of the channel by means of oversampling.

With this new set of sufficient statistics and using linear prediction to estimate the realization of the transition noise process, we propose a detection scheme based on the Viterbi algorithm with modified branch metric for Maximum A-Posteriori Probability (MAP) sequence detection. Numerical analysis and simulation demonstrate good improvements in terms of Minimum Mean Square Prediction Error (MMSPE) and bit error rate (BER) of the proposed detector with respect to conventional state-of-the-art systems.

II. CHANNEL MODEL

In order to describe the proposed oversampled detector, we consider a longitudinal magnetic recording channel modeled by a first-order position jitter and width variation [4]. Our results can be also extended to perpendicular magnetic and

optical recording systems affected by transition noise in a straightforward manner. Let $h(t, w)$ denote the response to an isolated transition recorded in magnetic or optical media, where t is time and w is a parameter characterizing the pulse width. Let $a_k \in \{\pm 1\}$ represent 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 $\eta(t)$ with power spectral density $N_0/2$ can be expressed as

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

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. For the pulse response $h(t, w)$ we have adopted the well known Lorentzian approximation [8] for longitudinal recording, i.e.

$$h(t, w) = \sqrt{\frac{4E_t}{\pi PW_{50}}} \cdot \frac{1}{1 + (2t/PW_{50})^2}$$

where $PW_{50} = 2w$ is the pulse width at half the maximum amplitude and E_t is the energy of the isolated pulse response.

According to [3], taking an n -th order Taylor series expansion of the read back impulse, the signal at the channel output can be approximated as a linear sum of the noise-free response and residual responses due to deviations around the nominal position and width of the pulse. Limiting the series expansion to the first order, the read back impulse associated to this (first-order) channel model can be approximated as

$$\begin{aligned} h(t + \Delta t_k, w + \Delta w_k) &\simeq \\ &\simeq h(t, w) + \Delta t_k \frac{\partial h(t, w)}{\partial t} + \Delta w_k \frac{\partial h(t, w)}{\partial w}. \end{aligned} \quad (2)$$

Defining the impulse response of the filters modeling the position jitter and width variation noise process as

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

and using this first-order approximation (2) in (1), the continuous waveform at the output of the channel can be written as

$$r(t) \simeq y(t) + \eta(t)$$

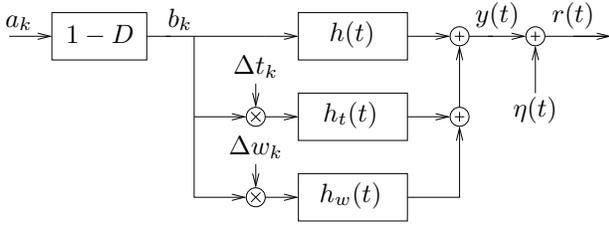


Fig. 1. Storage channel model with first-order media noise and additive white Gaussian thermal noise.

where we have defined $y(t)$ as¹

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

A block diagram descriptive of the first-order channel model is shown in Fig. 1.

III. SUFFICIENT STATISTICS AND DETECTOR STRUCTURE

It is a widespread belief that, in order to extract sufficient statistics from the continuous waveform $r(t)$ at the output of the channel, an optimal detector has to use a filter matched to the useful signal component, i.e., a filter with impulse response $h(-t)$. In [5], [6] it was shown that, in the presence of transition noise, the need for statistical sufficiency yield a detector front-end with a number of filters proportional to the modeling order of the transition noise: this approach, together with multidimensional linear prediction, was applied to MAP sequence detection and the improvement in terms of Signal-to-Noise Ratio (SNR) with respect to a conventional detector, shown in Fig. 2, was demonstrated by bit error rate simulations. Furthermore, in [9], the SNR improvement of the proposed multidimensional detector was also confirmed by an information rate analysis.

We now derive a new set of sufficient statistics, different from those proposed in [5], [6], and suggest a decoding algorithm based on this new quantities. First, it is possible to observe that the noiseless signal

$$y'(t) = r(t) - \eta(t)$$

has a bandwidth which exceeds half the signaling frequency $1/T$. Assuming a noiseless signal component $y'(t)$ with a one-side bandwidth B such that $B < \beta/2T$, with β a properly chosen integer, we can adopt as detector filter $p(t)$ one that has nonzero response over the bandwidth B and strictly bandlimited to B_p such that the following condition is satisfied

$$B_p < \frac{\beta}{T} - B. \quad (3)$$

Keeping in mind the results in [7] and the *reversibility principle* [10], which states that any reversible transformation carried out on the received signal does not modify the performance of an optimal detector, the set of samples at the

¹In the following, in order to simplify the notation, we omit the dependence of the various pulses on the pulse width w .

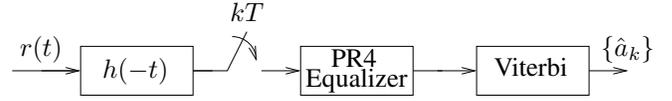


Fig. 2. Conventional detector.

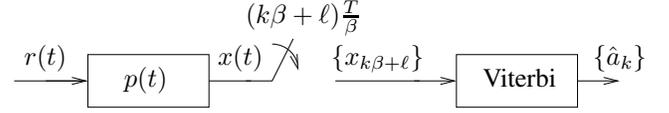


Fig. 3. Proposed detector.

output of the front-end filter $p(t)$ with rate β/T is a sufficient statistic for detection of the information sequence $\{a_k\}$. The samples $x_{k\beta+\ell}$ at the output of the sampler are indexed by k as to the symbol period and by ℓ , with $\ell = 0, 1, \dots, \beta - 1$, as the β samples extracted during the k -th symbol period. If the oversampling factor β is chosen in order to satisfy the condition (3), the sequence $\{x_{k\beta+\ell}\}$ is a sufficient statistic and can be used to enable MAP sequence detection based on the Viterbi algorithm [11]. In Fig. 3 the proposed detector is shown. In the following, we select $p(t)$ as a square-root raised cosine filter with 3-dB bandwidth equal to $\beta/2T$.

It should be pointed out that the sequence $\{x_{k\beta+\ell}\}$ represents a sufficient statistic, but this conclusion does not imply that the signal $x(t)$ could be reconstructed from its samples $\{x_{k\beta+\ell}\}$. In fact, the sampling theorem cannot be applied to the thermal noise component in $r(t)$. In other words, while $\{x_{k\beta+\ell}\}$ are sufficient to recover the information sequence, aliasing would be present in the filtered thermal noise component $\mu(t)$, where $\mu(t) \triangleq \eta(t) \otimes p(t)$. Therefore, the observable $x_{k\beta+\ell}$ at the output of the sampler can be expressed as

$$x_{k\beta+\ell} = \sum_n b_{k-n} g[n\beta + \ell] + \sum_n b_{k-n} \Delta t_{k-n} g_t[n\beta + \ell] + \sum_n b_{k-n} \Delta w_{k-n} g_w[n\beta + \ell] + \mu_{k\beta+\ell} \quad (4)$$

where $g[n\beta + \ell] \triangleq h(t) \otimes p(t)|_{t=n\beta+\ell}$, $g_t[n\beta + \ell] \triangleq h_t(t) \otimes p(t)|_{t=n\beta+\ell}$ and $g_w[n\beta + \ell] \triangleq h_w(t) \otimes p(t)|_{t=n\beta+\ell}$ are the oversampled impulse responses at the output of the filter $p(t)$ and $\mu_{k\beta+\ell}$ is the realization of the thermal noise process at time index $k\beta + \ell$. Note that an oversampling factor $\beta > 1$ is needed to satisfy (3): in this work we have used $\beta = 2$ and $\beta = 3$. Finally, note that the thermal noise samples $\{\mu_{k\beta+\ell}\}$ are white thanks to the selection of $p(t)$ as a square-root raised cosine filter.

IV. DETECTION STRATEGY BASED ON OVERSAMPLED LINEAR PREDICTION

Assuming a first-order channel model, we have shown that the quantities at the output of the receiver filter $\{x_{k\beta+\ell}\}$ are sufficient statistics for sequence detection. Collecting these

samples into a suitable vector \mathbf{x} , we can formulate the MAP detection strategy as

$$\hat{\mathbf{a}} = \underset{\mathbf{a}}{\operatorname{argmax}} P(\mathbf{a})f(\mathbf{x}|\mathbf{a}). \quad (5)$$

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

$$f(\mathbf{x}|\mathbf{a}) = \prod_{k=0}^{K-1} f(\mathbf{x}_{k\beta}^{(k+1)\beta-1} | \mathbf{x}_0^{k\beta-1}; \mathbf{a}_0^k) \quad (6)$$

where $\mathbf{x}_{k_1}^{k_2}$ is a shorthand notation for the vector collecting signal observations at the output of the sampler from time epoch k_1 to k_2 , K is the length of the recorded information sequence and \mathbf{a}_0^k are the information symbols² from time epoch 0 to k . The multidimensional conditional pdf in (6) can be further factorized as a product of β monodimensional conditional pdfs, i.e.

$$f(\mathbf{x}_{k\beta}^{(k+1)\beta-1} | \mathbf{x}_0^{k\beta-1}; \mathbf{a}_0^k) = \prod_{\ell=0}^{\beta-1} f(x_{k\beta+\ell}^{k\beta+\ell-1} | \mathbf{x}_0^{k\beta-1}; \mathbf{a}_0^k). \quad (7)$$

Using (7), we can now rewrite (6) as

$$\begin{aligned} f(\mathbf{x}|\mathbf{a}) &= \prod_{k=0}^{K-1} \prod_{\ell=0}^{\beta-1} f(x_{k\beta+\ell}^{k\beta+\ell-1} | \mathbf{x}_0^{k\beta-1}; \mathbf{a}_0^k) \\ &\simeq \prod_{k=0}^{K-1} \prod_{\ell=0}^{\beta-1} f(x_{k\beta+\ell}^{k\beta+\ell-1} | \mathbf{x}_{(k-\nu)\beta}^{k\beta+\ell-1}; \mathbf{a}_0^k). \end{aligned} \quad (8)$$

In the last step of (8), in order to limit the detector's memory, we have assumed Markovianity of order ν in the conditional observation sequences. Moreover we define a state of the system accounting for the "postcursors" and "precursors" of the impulse responses and the order of Markovianity ν as

$$\zeta_k = (a_{k-1}, a_{k-2}, a_{k-3}, \dots, a_{k-L}) \quad (9)$$

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 .

Since the thermal and transition noise processes have Gaussian distribution, the observation is Gaussian, given the data. The application of the chain factorization rule to (7) allows us to factor the multidimensional conditional pdf in (6) as a product of β one-dimensional conditional Gaussian pdfs, completely defined by the conditional mean

$$\hat{x}_{k\beta+\ell} = E \left\{ x_{k\beta+\ell} \mid \mathbf{x}_{(k-\nu)\beta}^{k\beta+\ell-1}; a_k, \zeta_k \right\}$$

and the conditional variance

$$\hat{\sigma}_{x_{k\beta+\ell}}^2 = E \left\{ (x_{k\beta+\ell} - \hat{x}_{k\beta+\ell})^2 \mid \mathbf{x}_{(k-\nu)\beta}^{k\beta+\ell-1}; a_k, \zeta_k \right\}$$

where $\hat{x}_{k\beta+\ell}$ can be interpreted as linear predictive estimate of $x_{k\beta+\ell}$ and $\hat{\sigma}_{x_{k\beta+\ell}}^2$ as the relevant MMSPE [12].

²Note that the size of the observation vector \mathbf{x} is β times the size of the data vector \mathbf{a} .

Note that, for a given value of ν , the number of prediction coefficients changes with respect to the number of past samples, indexed by ℓ , defined in the conditioning event and that the prediction coefficients are data-dependent [5], [6]. The solution of the Wiener-Hopf matrix equation for linear prediction based on an oversampled observable will be presented in Section V.

The detection strategy (5), the factorization (8) and linear prediction allow us to derive the branch metric to be used for sequence detection in a Viterbi algorithm. Taking the logarithm, assuming that the information bits are independent and identically distributed and discarding irrelevant terms, we can express the branch metric as

$$\lambda_k(a_k, \zeta_k) \propto \ln \prod_{\ell=0}^{\beta-1} f(x_{k\beta+\ell} | \mathbf{x}_{(k-\nu)\beta}^{k\beta+\ell-1}; a_k, \zeta_k).$$

The detection strategy (5) can be now formalized as

$$\hat{\mathbf{a}} = \underset{\mathbf{a}}{\operatorname{argmin}} \sum_{k=0}^{K-1} \lambda_k(a_k, \zeta_k)$$

where the branch metrics are expressed as

$$\lambda_k(a_k, \zeta_k) = \sum_{\ell=0}^{\beta-1} \left\{ \frac{(x_{k\beta+\ell} - \hat{x}_{k\beta+\ell})^2}{\hat{\sigma}_{x_{k\beta+\ell}}^2} + \ln \hat{\sigma}_{x_{k\beta+\ell}}^2 \right\}.$$

Finally, the state-complexity of a linear predictive detector can be limited by means of state-reduction techniques [13]–[16]. Let $Q < L$ denote the memory parameter to be taken into account in the definition of a "reduced" trellis state

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

The branch metric can be obtained by defining a "pseudo-state" [17]

$$\tilde{\zeta}_k(\omega_k) = \underbrace{(a_{k-1}, \dots, a_{k-Q})}_{Q \text{ bits}}; \underbrace{(\check{a}_{k-Q-1}(\omega_k), \dots, \check{a}_{k-Q-P}(\omega_k))}_{P \text{ bits}} \quad (10)$$

where P bits are chosen by a Per-Survivor Processing (PSP) technique [16]. Note that $\check{a}_{k-Q-1}(\omega_k), \dots, \check{a}_{k-Q-P}(\omega_k)$ are the P information bits associated with the survivor of ω_k . The branch metric $\tilde{\lambda}_k(a_k, \omega_k)$ in the reduced-state trellis can be defined in terms of the pseudo-state (10) according to

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

V. OVERSAMPLED LINEAR PREDICTION

In this section, we present the linear prediction approach applied to an oversampled observable collecting the sufficient statistics $\{x_{k\beta+\ell}\}$ and outline how to obtain an estimate of the transition noise samples at the output of the filter $p(t)$.

Given the sufficient statistics $\{x_{k\beta+\ell}\}$ and defining

$$s_{k\beta+\ell} = \sum_n b_{k-n} g[n\beta + \ell]$$

as the quantities collecting the noiseless signal component at the output of the front-end filter³, we can express the cost function to be minimized [12] as

$$J_l(\mathbf{p}_\ell, a_k, \zeta_k) = E \left\{ \left[(x_{k\beta+\ell} - s_{k\beta+\ell}) - \sum_{m=1}^{\nu} \sum_{i=0}^{\beta-1} p_{k-m,i}^{(\ell)} [x_{(k-m)\beta+i} - s_{(k-m)\beta+i}] - \underbrace{\sum_{j=0}^{\ell-1} p_{k,j}^{(\ell)} [x_{k\beta+j} - s_{k\beta+j}]}_{=0 \text{ if } \ell=0} \right]^2 \middle| a_k, \zeta_k \right\} \quad (11)$$

where $p_{n,i}^{(\ell)}$ is the prediction coefficient for index ℓ and observation sample $x_{n\beta+i}$. In fact, for a given prediction order ν , the number of predictors increases with ℓ since also the number of observables available at the detector increases. In other words, when $\ell = 0$ we can use the last $\nu\beta$ observables for linear prediction, when $\ell = 1$ we can use the same $\nu\beta$ observables, as done with $\ell = 0$, and the sample $x_{k\beta}$ obtained at time index $k\beta$. Generally speaking, at time index $k\beta + \ell$, the number of samples used in the evaluation of the predictor coefficient is $\nu\beta + \ell$. As a consequence, for every couple (a_k, ζ_k) we have β Wiener-Hopf systems to be solved in order to obtain, when ℓ reaches its maximum value $\beta - 1$, a number of prediction coefficients equal to

$$\nu\beta^2 + \sum_{j=0}^{\beta-1} j = \nu\beta^2 + \frac{\beta(\beta-1)}{2}.$$

Moreover, for a given prediction order ν , the proposed detector can operate on a wider set of observables with respect to a conventional detector operating at symbol rate, leading to a better estimate of the transition noise samples. Note finally that a similar approach was used in [5], [6], where a different set of sufficient statistics yields a detector with a multidimensional front-end and the detection strategy yields a formulation of a multidimensional linear estimate: the approach in this article can be viewed as multidimensional linear prediction applied not *in space* as in [5], [6] but *in time*.

In order to derive the Wiener-Hopf system, let us rewrite the vector of observations for linear prediction as

$$\mathbf{z}_n \triangleq [z_{n\beta}, z_{n\beta+1}, \dots, z_{(n+1)\beta-1}]^T$$

where $z_{n\beta+\ell} = x_{n\beta+\ell} - s_{n\beta+\ell}, \forall \ell = 0, \dots, \beta - 1$. With this definition we are now able to formulate the ℓ -th Wiener-Hopf equation as

$$\mathbf{R}_\ell(a_k, \zeta_k) \cdot \mathbf{p}_\ell = \mathbf{q}_\ell(a_k, \zeta_k) \quad (12)$$

where the system matrix is defined as

$$\mathbf{R}_\ell(a_k, \zeta_k) = \begin{bmatrix} \mathbf{R}_{k-\nu, k-\nu} & \dots & \mathbf{R}_{k-\nu, k-1} & \mathbf{R}_{k-\nu, k}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{R}_{k-1, k-\nu} & \dots & \mathbf{R}_{k-1, k-1} & \mathbf{R}_{k-1, k}^{(\ell)} \\ \mathbf{R}_{k, k-\nu}^{(\ell)} & \dots & \mathbf{R}_{k, k-1}^{(\ell)} & \mathbf{R}_{k, k}^{(\ell)} \end{bmatrix}$$

³I.e., the first term in (4).

in which $\mathbf{R}_{i,j} \triangleq E\{z_i z_j^H | a\}$ is a $\beta \times \beta$ correlation matrix, $\mathbf{R}_{i,j}^{(\ell)}$ is a correlation matrix whose dimension changes with respect to the oversampling index ℓ during the symbol interval $k\beta$ and the vector of known terms is

$$\mathbf{q}_\ell(a_k, \zeta_k) = \begin{bmatrix} E\{z_{k\beta+\ell} \cdot z_{k-\nu}\} \\ \vdots \\ E\{z_{k\beta+\ell} \cdot z_{k-1}\} \\ E\{z_{k\beta+\ell} \cdot z_k\} \end{bmatrix}.$$

The linear system defined in (12) can now be solved using Cholesky factorization [12], obtaining the prediction coefficient vector

$$\mathbf{p}_\ell = \mathbf{R}_\ell^{-1}(a_k, \zeta_k) \cdot \mathbf{q}_\ell(a_k, \zeta_k).$$

Since the transition noise sample we wish to estimate is data dependent, the system matrix $\mathbf{R}_\ell(a_k, \zeta_k)$ and the vector $\mathbf{q}_\ell(a_k, \zeta_k)$ are also data dependent. The prediction coefficients can be precomputed, given the state ζ_k and the current information bit a_k , and stored in a look-up table.

Finally, using the definition of \mathbf{p}_ℓ , $\mathbf{R}_\ell(a_k, \zeta_k)$ and $\mathbf{q}_\ell(a_k, \zeta_k)$, we can express the ℓ -th MMSPE as

$$\begin{aligned} J_\ell(\mathbf{p}_\ell, a_k, \zeta_k) &= \sigma_z^2 - 2 \cdot \mathbf{p}_\ell^T \cdot \mathbf{q}_\ell(a_k, \zeta_k) + \mathbf{p}_\ell^T \cdot \mathbf{R}_\ell(a_k, \zeta_k) \cdot \mathbf{p}_\ell \\ &= \sigma_z^2 - \mathbf{q}_\ell^T(a_k, \zeta_k) \cdot \mathbf{R}_\ell^{-1}(a_k, \zeta_k) \cdot \mathbf{q}_\ell(a_k, \zeta_k) \\ &= \sigma_z^2 - \mathbf{p}_\ell^T \cdot \mathbf{q}_\ell(a_k, \zeta_k) \end{aligned}$$

where σ_z^2 is the variance of the overall noise (thermal noise and transition noise).

VI. NUMERICAL RESULTS

A. MMSPE analysis

In order to assess the performance of the proposed oversampled detector with respect to other solutions, we adopt the reference conventional detector shown in Fig. 2 and based on a matched filter $h(-t)$, a sampler at symbol rate, a PR4 equalizer [18] and linear prediction applied to colored Gaussian thermal noise and transition noise. We compare its performance with that obtained using oversampling and linear prediction applied to transition noise.

In Fig. 4, it is shown the MMSPE as a function of the prediction order ν obtained both for the reference detector (continuous lines) and for the proposed oversampled detector (dashed and dotted lines) for $\beta = 2$. We consider a normalized density $D = 2.5$, defined as $D = PW_{50}/T$. In Fig. 4, note that the parameter ℓ is related to the number of noise samples used for linear prediction, i.e., with $\ell = 0$ the noise samples are $\nu\beta$ (MMSPE with dotted lines), with $\ell = 1$ the noise samples are $\nu\beta + 1$ (MMSPE with dashed lines).

Since media noise arises in transitions, two bit patterns were considered: (i) one characterized by continuous changes of the writing current's polarity, the so called "1T sequence" pattern $\{1, -1, 1, -1, \dots\}$, (ii) one characterized by a bit pattern like $\{+1, +1, -1, -1, +1, +1, \dots\}$, the so called "2T sequence". An average MMSPE is also shown, by averaging

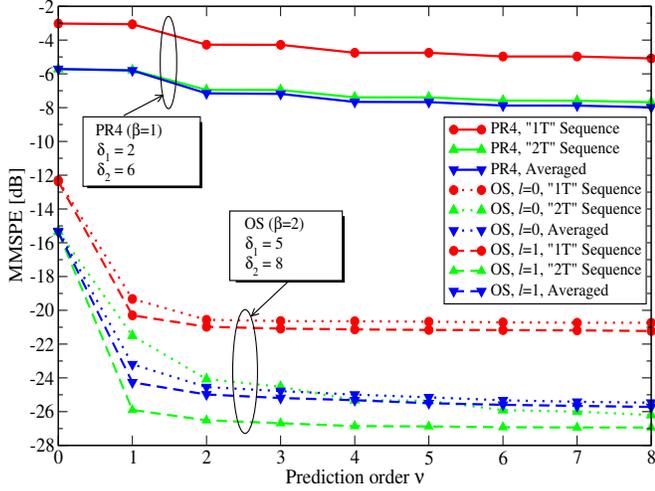


Fig. 4. *MMSPE versus the prediction order ν for longitudinal magnetic recording with $D = 2.5$, $\alpha = 95$, $\lambda = 50$ and $\text{SNR}_{95} = 10$ dB.*

over all possible bit patterns defining a trellis branch (a_k, ζ_k) (“averaged” pattern). In the reference detector the number of precursors and postcursors considered are, respectively, $\delta_1 = 2$ and $\delta_2 = 6$, while for the proposed oversampled detector these parameters assume the values $\delta_1 = 5$ and $\delta_2 = 8$. These values were selected in order to consider 99% of the energy of the channel impulse responses, for all considered detectors.

The SNR with transition noise [19] is defined at the output of the matched filter $h(-t)$ for both detectors as

$$\text{SNR}_\alpha = \frac{P_s}{\sigma_n^2 + \sigma_m^2} \quad (13)$$

where P_s is the signal power, σ_n^2 is the thermal noise power, σ_m^2 is the transition noise power and $\alpha = 100 \times [\sigma_m^2 / (\sigma_n^2 + \sigma_m^2)]$ denotes the percentage of transition noise with respect to thermal noise. This definition of SNR is motivated by the fact that we wish to compare the performance of all detectors under SNR equal conditions.

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 ($\lambda = 50$). Fig. 4 shows that, using the proposed detector working on an oversampled observable with a prediction order $\nu = 2$, it is possible to obtain a relative gain of almost 9.0 dB when $\ell = 1$ for the all-transition bit pattern, with respect to the MMSPE gain of 1.5 dB achievable by the reference detector. Note that this gain is obtained assuming that the estimate of transition noise is done by exploiting the full available set of samples, i.e., collecting data from $x_{(k-\nu)\beta}$ up to $x_{k\beta+(\beta-1)}$. Finally, averaging over all possible bit patterns, the MMSPE relative gain is 1.5 dB for the conventional detector and almost 10 dB for the proposed detector when $\ell = 1$.

B. BER analysis

Fig. 5 shows the BER performance for (i) a detector based on a matched filter with impulse response $h(-t)$, partial

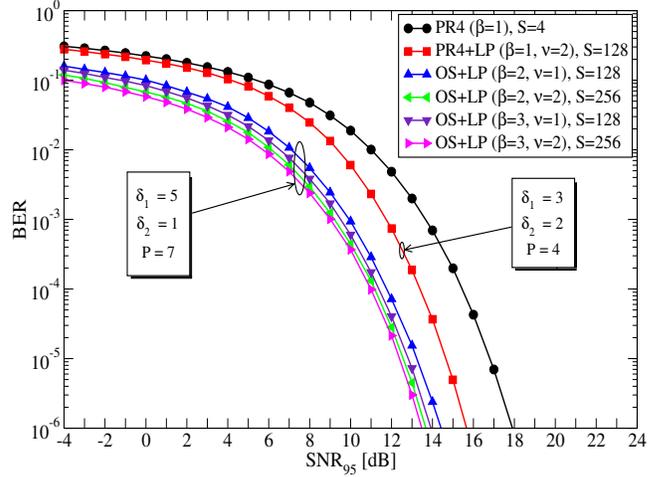


Fig. 5. *Bit error probability for the Lorentzian magnetic channel model with transition noise at $D = 2.5$, $\alpha = 95$ and $\lambda = 50$.*

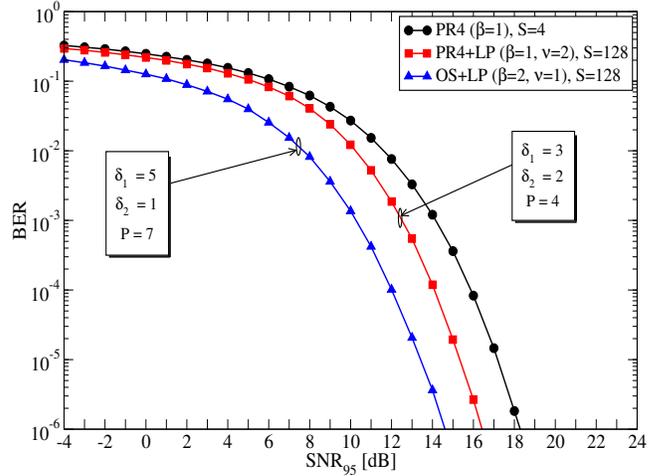


Fig. 6. *Bit error probability for the Lorentzian magnetic channel model with transition noise at $D = 3.0$, $\alpha = 95$ and $\lambda = 50$.*

equalization PR4 and Euclidean branch metric (curve labeled “PR4”), (ii) the reference detector (curve labeled “PR4+LP”) and (iii) the proposed oversampled detector (curve labeled “OS+LP”) assuming a transition noise consisting of 50% position jitter and 50% width variation ($\lambda = 50$). In Fig. 5 the state complexity for each detector (parameter S) and other relevant parameters, such as the prediction order ν , the PSP parameter P , the number of precursors and postcursors δ_1 and δ_2 and the number of samples considered for symbol period β are also shown: note that, in order to evaluate the robustness of the proposed detector, we present the BER curves with different combination of the parameters β, ν and S . The SNR gain of the proposed detector with respect to the reference one is approximately 1 dB with oversampling factor $\beta = 2$ and prediction order $\nu = 1$ and 2 dB with oversampling factor $\beta = 3$ and prediction order $\nu = 2$.

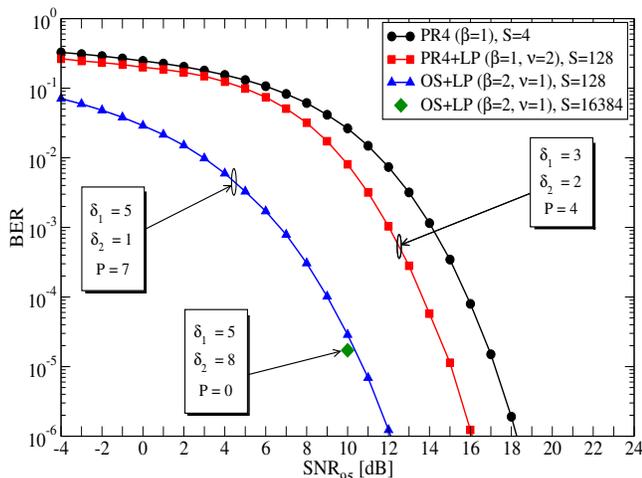


Fig. 7. Bit error probability for the Lorentzian magnetic channel model with transition noise at $D = 3.0$, $\alpha = 95$ and $\lambda = 100$.

Fig. 6 shows the BER performance for the same three detectors of Fig. 5 assuming a user density of $D = 3.0$ and a transition noise consisting of 50% position jitter and 50% width variation ($\lambda = 50$). Given the system parameters $\beta, S, \delta_1, \delta_2, \nu, P$, the proposed detector outperforms the reference detector of approximately 2 dB. It should be pointed out that the proposed detector shows an increasing SNR gain with respect to user density (the SNR gain in Fig. 5 was approximately 1 dB with the same parameters used in Fig. 6): these findings demonstrate the capability of our detector to copy well with transition noise also for a storage system working at higher user density.

Finally, Fig. 7 shows the BER curves obtained for the same three detectors of Fig. 5 operating at a user density $D = 3.0$ for a channel characterized by a transition noise consisting of 100% position jitter ($\lambda = 100$). The proposed detector outperforms the reference detector of approximately 4 dB. Fig. 7 also shows the performance of the oversampled detector at 10 dB with full state-complexity (16384 trellis states). This result emphasizes the limited penalty introduced by the state-reduction technique applied in this paper.

The results shown in Fig. 5, 6 and Fig. 7 confirm the findings in [5], [6]: i.e., when the channel is mainly limited by position jitter, the SNR gain achievable with the linear prediction approach operating on a multidimensional observable is increased with respect to that achievable for a channel affected by transition noise components in equal part ($\lambda = 50$).

VII. CONCLUSIONS

A new set of sufficient statistics for digital storage systems subject to data-dependent transition noise has been proposed. These sufficient statistics can be obtained through a detector operating on an oversampled sequence at the output of a suitable analog filter. Linear prediction based on an oversampled

observable can be used to modify the branch metrics of a Viterbi detector in order to improve the performance and make it more robust against transition noise.

The gain, in terms of SNR at a given bit error rate, achievable using the proposed detector is presented with respect to a state-of-the-art solution based on partial response equalization and linear prediction by means of numerical simulations. Finally, the proposed oversampled signal processing technique can be similarly applied to perpendicular magnetic recording systems and to optical storage system.

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] R. Pighi, R. Raheli, and U. Amadei, "Multidimensional receiver front-end for storage systems with data dependent transition noise," in *Proc. IEEE Int. Symp. Inform. Theory and its Appl.*, Parma, Italy, October 2004, pp. 885–890.
- [6] —, "Multidimensional signal processing and detection for storage systems with data-dependent transition noise," *IEEE Trans. Magnetics*, to be published.
- [7] H. Meyer, M. Oerder, and A. Polydoros, "On sampling rate, analog prefiltering, and sufficient statistics for digital receivers," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3208–3214, December 1994.
- [8] J. Proakis, *Digital communications*. McGraw-Hill, 2001.
- [9] R. Pighi, R. Raheli, and F. Cappelletti, "Information rates of multidimensional front-ends for digital storage channels with data-dependent transition noise," in *Proc. IEEE Symposium on Information Theory*, Adelaide, Australia, September 2005.
- [10] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. Waveland Press, 1990, (reprint of 1965 original from John Wiley and Sons).
- [11] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, March 1973.
- [12] S. Haykin, *Adaptive filter theory*. Prentice-Hall International Editions, 1991.
- [13] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, January 1988.
- [14] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.
- [15] P. R. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. Commun.*, vol. 37, no. 7, pp. 669–676, July 1989.
- [16] 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.
- [17] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection algorithms for wireless communications, with applications to wired and storage systems*. London: John Wiley & Sons, 2004.
- [18] R. D. Cideciyan, F. Dolivo, R. Hermann, W. Hirt, and W. Schott, "A PRML system for digital magnetic recording," *IEEE J. Select. Areas Commun.*, vol. 10, no. 1, pp. 38–56, January 1992.
- [19] 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.