DATA-AIDED TIMING RECOVERY IN THE PRESENCE OF DATA-DEPENDENT NOISE

J. Riani, S. Van Beneden, and J.W.M. Bergmans

Eindhoven University of Technology
Den Dolech 2, 5600 MB, Eindhoven, The Netherlands.
phone: + (31) 40 247 5082, fax: + (31) 40 246 6508, email: j.riani@tue.nl

ABSTRACT

This paper presents a new data-aided timing recovery algorithm for channels with data-dependent noise. Based on a data-dependent Gauss-Markov model of the noise, a maximum-likelihood timing recovery scheme is derived. The proposed timing recovery algorithm incorporates data-dependent noise prediction parameters in the form of linear prediction filters and prediction error variances. Moreover, because noise can be nonstationary in practice, an adaptive algorithm is proposed in order to estimate and track the noise prediction parameters. Simulation results, for a partial response maximum-likelihood system, show that our algorithm allows an important reduction in timing jitter whenever noise is dominantly data-dependent.

1. INTRODUCTION

Timing recovery is one of the critical functions for reliable data detection in digital synchronous communication systems. The key problem in timing recovery is the determination of time instants at which the received signal should be sampled for reliable data recovery. This problem has been a subject of investigation for many decades. Among the existing solutions [1], data-aided (DA) timing recovery schemes, e.g. [2][3][4], are known to be more powerful. DA schemes use the transmitted data sequence as side information to facilitate timing recovery. This information is available to the receiver either in the form of a known preamble pattern preceding the user data, or as decisions taken from the bit detector.

Existing timing recovery schemes assume that the noise at their input is stationary and that noise statistics are independent of the transmitted data. However in many communication systems noise is nonstationary and data-dependent. This data-dependent nature of the noise significantly deteriorates the performance of conventional timing recovery schemes. It increases timing jitter at a given loop bandwidth, resulting in an increased bit-error rate.

The motivations for this paper relate to digital recording systems where timing recovery becomes more critical as storage density increases because of bandwidth limitation and signal to noise ratio degradation on the one hand and noise nonstationarity and data-dependency on the other hand. Although the problem of data detection in such noise environments has received considerable attention, e.g. [5], much less attention has been devoted to the problem of timing recovery. For the simple case of additive white and Gaussian noise (AWGN) channels with a noise variance dependent only on the transmitted symbol, a timing recovery algorithm was proposed in [6]. This algorithm is not based on an optimal timing function but is derived as a modification of the Mueller and Müller algorithm [2].

In this paper we derive an optimal timing recovery algorithm for data-dependent correlated noise. We model the noise as a Gauss-Markov correlated noise whose statistics are data-dependent. Based on this model Maximum-Likelihood (ML) timing recovery is addressed. The resulting structure is a timing recovery scheme that incorporates data-dependent noise prediction. Moreover, because in practice noise can be nonstationary, an adaptive algorithm that tracks the prediction parameters is proposed. Simulation results for a partial response maximum-likelihood (PRML)[4] system show that our algorithm allows important reductions in timing jitter at low signal to noise ratios in the presence of data-dependent noise.

2. SYSTEM MODEL AND PROBLEM DEFINITION

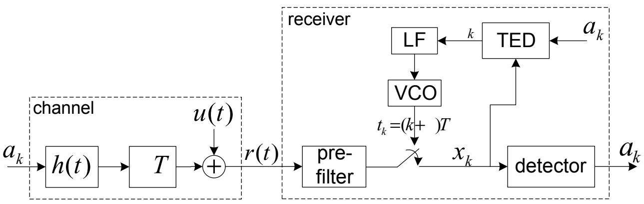

Figure 1: system model

In Fig. 1, a zero-mean data sequence $a_k \in \{\pm 1\}$ of length N, i.e. $a_1,a_2,...a_N$, of data rate $1/T$ is applied to a channel with symbol response $h(t)$, additive noise $u(t)$ and an a priori unknown and possibly time varying delay (in bit intervals T). Prior to detection, the receiver performs prefiltering that serves to suppress noise and may also condition intersymbol interference (ISI). The prefilter output is first sampled and then passed to a detector that produces bit decisions. For clarity of this paper, we assume that excess bandwidth at the prefilter output is negligible and consider only baud-rate sampling. The results of this paper can be easily extended to the oversampled case. The sampling instants are expressed as $t_k = (k+\ )T$ where is a sampling phase (normalized in units T). Based on the sampled sequence x_k, the receiver produces bit decisions $\hat{a}_k$ as well as a clock signal that indicates the sampling instants t_k. In order for the detector to operate properly, a timing recovery subsystem ensures that the sampling phase closely approaches . The timing recovery subsystem takes the form of a phase-locked loop (PLL) with a timing-error detector (TED), loop filter (LF), and a

This research was supported by the EU under project IST-2001-34168 (TwoDOS).

voltage controlled oscillator (VCO). The TED produces an estimate $_k$ of the sampling-phase error $\;=\;-\;$. In this paper we restrict attention to data-aided (DA) TEDs where a_k is assumed to be available to the receiver in the form of a known preamble, or as decisions when bit-error rates are small. PLL behavior depends also on the LF and VCO. A detailed description of this dependence can be found in [7].

To simplify the forthcoming analysis we assume, first, that the LF has a sufficiently high bandwidth to enable the variations of to be tracked. Under this assumption we can consider to be fixed. Second, the sampling-phase errors are restricted to a fraction of a symbol interval T (this reflects the situation when the PLL is in lock; PLL acquisition properties are beyond the scope of this paper). In this case, the equivalent discrete impulse response q_k at the detector input can be linearized as $q_k \simeq q_k^0 + \; q'_k$, where q'_k is the derivative of q_k with respect to at $\;= 0$. Both responses q_k^0 and q'_k are assumed to be known to the receiver. The detector input sequence can be written as

$$x_k = (q^0 * a)_k + \;(q' * a)_k + n_k, \tag{1}$$

where `$*$' denotes linear convolution and n_k is the equivalent noise sequence at the detector input. Unless specified otherwise, we assume that q_k^0 corresponds to the ideal ISI structure assumed by the detector. Any misequalization ISI at ideal sampling phase, i.e. due to a mismatch between q_k^0 and the ideal detector response, is embedded in the noise n_k. The noise n_k includes also channel noise that may be linearly or non-linearly data-dependent. The key to our timing recovery approach is our modeling of the noise. We assume the following statistical properties of n_k:

1. *Finite correlation length*: The noise n_k is assumed to be independent of past samples before some length $L \geq 0$ (Markov memory length). This independence implies that

$$p(n_k|n_{k-1},...,n_1,\underline{a}_1^N) = p(n_k|n_{k-1},...,n_{k-L},\underline{a}_1^N) \tag{2}$$

   where $p(.)$ denotes the probability density function (pdf) of n_k conditioned on the past noise samples and on the data $\underline{a}_1^N$ where $\underline{a}_{k_1}^{k_2} = [a_{k_1}, a_{k_1+1}, ..., a_{k_2}]$ for $k_2 \geq k_1$. The conditioning on $\underline{a}_1^N$ is meant to take into account the data-dependent correlation of the noise n_k.
2. *Finite data-dependent span*: The noise n_k depends only on its first K-neighbor symbols, i.e. $\underline{a}_{k-K_1}^{k+K_2}$, that we call symbol cluster, where $K = K_1 + K_2 + 1$. Eq. (2) becomes

$$p(n_k|n_{k-1},...,n_{k-L},\underline{a}_1^N) = p(n_k|n_{k-1},...,n_{k-L},\underline{a}_{k-L-K_1}^{k+K_2}) \tag{3}$$

3. *Joint Gaussian pdf's*: The joint pdf $p(n_k, n_{k-1}, ..., n_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2})$ is Gaussian with a covariance matrix $\mathbf{C}_k = \mathbf{C}(\underline{a}_{k-L-K_1}^{k+K_2})$ of size $(L+1)\times(L+1)$, i.e.

$$p(n_k,...,n_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2}) = \frac{\exp[-\underline{N}_k^{\mathrm{T}}\mathbf{C}_k^{-1}\underline{N}_k]}{\sqrt{(2\;)^{L+1}\det\mathbf{C}_k}}, \tag{4}$$

   where $[.]^{\mathrm{T}}$ denotes the transpose operation and $\underline{N}_k = [n_k, ..., n_{k-L}]^{\mathrm{T}}$.

3. MAXIMUM-LIKELIHOOD PHASE-ERROR ESTIMATOR

Data-aided ML timing recovery is optimum when no prior statistical knowledge about the phase-error is available. Before developing the DA ML TED for sample-by-sample timing recovery, let us first derive the one-shot ML estimator of the phase-error based on the observation of $x_1,...,x_N$. To this aim, we assume in this section that noise statistics are known and fixed during the transmission of the N symbols $\underline{a}_1^N$. The DA ML estimate of the phase-error is obtained by maximizing the likelihood function, i.e.

$$^{\mathrm{ML}} = \arg[\max p(x_1,...,x_N|\underline{a}_1^N,\;)], \tag{5}$$

over all possible phase-errors , where the likelihood function $p(x_1,...,x_N|\underline{a}_1^N,\;)$ is the joint probability density function of the received samples $x_1,...,x_N$ conditioned on the transmitted symbols $\underline{a}_1^N$ and the phase-error $\;=\;$.

In order to derive a practical criterion from (5) few steps are needed. We first apply Bayes rule and obtain

$$p(x_1,...,x_N|\underline{a}_1^N,\;) = \prod_{k=1}^{N} p(x_k|x_{k-1},...,x_1,\underline{a}_1^N,\;). \tag{6}$$

Upon invoking (1),(2) and (3) and applying Bayes rule once again, (6) can be factorized into

$$p(x_1,...,x_N|\underline{a}_1^N,\;) = \prod_{k=1}^{N} \frac{p(x_k,x_{k-1},...,x_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2},\;)}{p(x_{k-1},...,x_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2},\;)}. \tag{7}$$

The right-hand factors in (7) can be rewritten using (4) as:

$$\frac{p(x_k,..,x_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2},\;)}{p(x_{k-1},..,x_{k-L}|\underline{a}_{k-L-K_1}^{k+K_2},\;)} \propto \frac{\exp[(\underline{e}_k - \;\underline{s}_k)^{\mathrm{T}}\mathbf{c}_k^{-1}(\underline{e}_k - \;\underline{s}_k)]}{\exp[(\underline{E}_k - \;\underline{S}_k)^{\mathrm{T}}\mathbf{C}_k^{-1}(\underline{E}_k - \;\underline{S}_k)]}, \tag{8}$$

where the matrix $\mathbf{c}_k$ is the $L\times L$ lower principal submatrix of $\mathbf{C}_k$, i.e. $\mathbf{C}_k = \begin{bmatrix} _k & \underline{v}_k^{\mathrm{T}} \\ \underline{v}_k & \mathbf{c}_k \end{bmatrix}$, and where the column vectors:
$\underline{E}_k = [x_k - (q^0 * a)_k, ..., x_{k-L} - (q^0 * a)_{k-L}]^{\mathrm{T}}$,
$\underline{e}_k = [x_{k-1} - (q^0 * a)_{k-1}, ..., x_{k-L} - (q^0 * a)_{k-L}]^{\mathrm{T}}$,
$\underline{S}_k = [(q' * a)_k, ..., (q' * a)_{k-L}]^{\mathrm{T}}$,
$\underline{s}_k = [(q' * a)_{k-1}, ..., (q' * a)_{k-L}]^{\mathrm{T}}$.
The proportionality factor in (8) equals $\sqrt{\frac{(2\;)^L \det\mathbf{c}_k}{(2\;)^{L+1}\det\mathbf{C}_k}}$ which is independent of . It follows, by taking the log-likelihood in (7), that ML phase-error estimation is obtained by minimizing the following cost function:

$$(\;) = \sum_{k=1}^{N} (\underline{E}_k - \;\underline{S}_k)^{\mathrm{T}}\mathbf{C}_k^{-1}(\underline{E}_k - \;\underline{S}_k) - (\underline{e}_k - \;\underline{s}_k)^{\mathrm{T}}\mathbf{c}_k^{-1}(\underline{e}_k - \;\underline{s}_k). \tag{9}$$

This expression of $(\;)$ is still quite complex in that it involves inversions of the matrices $\mathbf{C}_k$ and $\mathbf{c}_k$ for all possible symbol clusters $\underline{a}_{k-L-K_1}^{k+K_2}$. A simplified expression of $(\;)$ can be derived via the matrix inversion lemma and reads

$$(\;) = \sum_{k=1}^{N} \frac{1}{_k}(\underline{w}_k^{\mathrm{T}}(\underline{E}_k - \;\underline{S}_k))^2, \tag{10}$$

where $\underline{w}_k = \begin{bmatrix} 1 \\ -\mathbf{c}_k^{-1}\underline{v}_k \end{bmatrix}$ (of size $(L+1)\times 1$) and $_k = \quad _k - \underline{v}_k^{\mathrm{T}}\mathbf{c}_k^{-1}\underline{v}_k$. The complexity to compute $(\ )$ is brought down to $O(N(L+1))$ in (10) instead of $O(N(L+1)^2)$ in (9). The vectors $\mathbf{c}_k^{-1}\underline{v}_k$ can be interpreted as data-dependent noise predictors and the values $_k$ as noise prediction variances. In fact, for a given symbol cluster $\underline{a}_{k-L-K_1}^{k+K_2}$, $\underline{w}_k$ acts to whiten the noise n_k by substracting from n_k the predicted component from the past noise samples. The variance of the whitened noise, i.e. $\underline{w}_k^{\mathrm{T}}\underline{N}_k$, equals $_k$.

The ML one-shot phase-error estimate $^{\mathrm{ML}}$ can be easily derived from (10) and is given by

$$^{\mathrm{ML}} = \frac{1}{\sum_{k=1}^{N} \frac{1}{_k}(\underline{w}_k^{\mathrm{T}}\underline{S}_k)^2} \sum_{k=1}^{N} \frac{1}{_k}(\underline{w}_k^{\mathrm{T}}\underline{E}_k)(\underline{w}_k^{\mathrm{T}}\underline{S}_k). \qquad (11)$$

The ML phase-error estimate (11) is a normalized average of an instantaneous timing error function given by $\frac{1}{_k}(\underline{w}_k^{\mathrm{T}}\underline{E}_k)(\underline{w}_k^{\mathrm{T}}\underline{S}_k)$. The ML TED can thus be simply written as

$$_k^{\mathrm{ML}} = \frac{1}{_k}(\underline{w}_k^{\mathrm{T}}\underline{E}_k)(\underline{w}_k^{\mathrm{T}}\underline{S}_k), \qquad (12)$$

where the vector $\underline{w}_k$ and the scalar $_k$ correspond to the cluster $\underline{a}_{k-L-K_1}^{k+K_2}$. Equation (12) presents two interesting properties. First, the division with $_k$ provides a weighing for every cluster of symbols $a_{k-L-K_1}^{k+K_2}$. The weight of a given cluster is inversely proportional to $_k$. More reliable symbol clusters that have smaller 'unpredictable' noise variance will be attributed higher gains in the extraction of timing information than noisy clusters. Second, the 'predictable' component of n_k from $n_{k-1},...,n_{k-L}$ is removed via the scalar product with $\underline{w}_k$, allowing thus less noise power to be sensed by the timing recovery subsystem. For example, in the extreme case where n_k is a deterministic linear combination of $n_{k-1},...,n_{k-L}$, the filtered noise $\underline{w}_k^{\mathrm{T}}\underline{N}_k$ is simply zero. These two properties together make up the strength of the proposed TED.

Remark: In the case of zero-mean additive white and data-independent noise with a variance ^{2}, one can show that $L = 0$, $_k = \ ^2$ and $\underline{w}_k = 1$. Equation (10) boils down to

$$(\ ) = \frac{1}{^2} \sum_{k=1}^{N} (e_k - \ (q' * a)_k)^2,$$

where $e_k = x_k - (q^0 * a)_k$. The optimum TED in this case is the Zero-Forcing (ZF) TED [1]. Its output, multiplied by ^{2}, is given by

$$_k^{\mathrm{ZF}} = e_k(q' * a)_k.$$

4. ADAPTIVE DATA-DEPENDENT NOISE CHARACTERIZATION

In the previous section, we assumed that $\underline{w}(\underline{a}_{k-L-K_1}^{k+K_2})$ and $(\underline{a}_{k-L-K_1}^{k+K_2})$ are known for all symbol clusters. However, the statics of the noise are not known in practice and need to be estimated from the data. Moreover, tracking these statistics adaptively is preferable in many applications because the noise may be nonstationary.

As mentioned earlier, the scalar product with $\underline{w}(\underline{a}_{k-L-K_1}^{k+K_2})$ is meant to whiten the noise samples $n_k,...,n_{k-L}$, for the symbol cluster $\underline{a}_{k-L-K_1}^{k+K_2}$, and $(\underline{a}_{k-L-K_1}^{k+K_2})$ is the variance of the whitened noise. Thus a simple scheme to estimate and track $\underline{w}(\underline{a}_{k-L-K_1}^{k+K_2}) = \begin{bmatrix} 1 \\ -\underline{\ }(\underline{a}_{k-L-K_1}^{k+K_2}) \end{bmatrix}$ and $(\underline{a}_{k-L-K_1}^{k+K_2})$ can be based on Fig. 2.

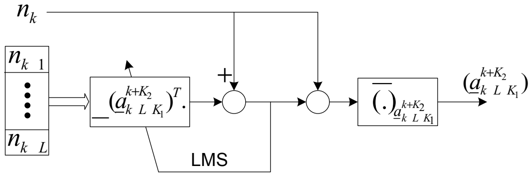

Figure 2: Adaptation of $\underline{\ }(\underline{a}_{k-L-K_1}^{k+K_2})$ and estimation of $(\underline{a}_{k-L-K_1}^{k+K_2})$. The averaging $\overline{(.)}_{\underline{a}_{k-L-K_1}^{k+K_2}}$ is symbol cluster dependent.

At every clock cycle, one prediction vector $\underline{\ }(\underline{a}_{k-L-K_1}^{k+K_2})$ and prediction variance $(\underline{a}_{k-L-K_1}^{k+K_2})$ are adapted. The adaptation of the prediction vector is based on the least mean square (LMS) technique and seeks to minimize $(\underline{w}_k^{\mathrm{T}}\underline{N}_k)^2$. The adaptation of $\underline{\ }(\underline{a}_{k-L-K_1}^{k+K_2})$ and estimation of $(\underline{a}_{k-L-K_1}^{k+K_2})$ are given by:

$$\begin{aligned} \underline{\ }^{\mathrm{new}}(\underline{a}) &= \underline{\ }^{\mathrm{old}}(\underline{a}) + \ (\underline{w}^{\mathrm{old}}(\underline{a})^{\mathrm{T}}\underline{N}_k)\underline{n}_k \\ ^{\mathrm{new}}(\underline{a}) &= (1 - \ )\ ^{\mathrm{old}}(\underline{a}) + \ (\underline{w}^{\mathrm{old}}(\underline{a})^{\mathrm{T}}\underline{N}_k)n_k, \end{aligned}$$

where denotes the adaptation constant and $\underline{n}_k = [n_{k-1},...,n_{k-L}]^{\mathrm{T}}$. For notational convenience, we indicate the symbol cluster by $\underline{a} = \underline{a}_{k-L-K_1}^{k+K_2}$.

In practice, n_k is not available to the receiver and the adaptation of the prediction parameters has to be based on the error signal e_k. In this case, the average TED gain, i.e. the average of $\frac{\ _k}{}(\ = 0)$ over all symbol clusters, may fluctuate around its ideal value because of a small phase-error for example. In order to ensure a fixed average TED gain, which is convenient for a proper dimensioning of the timing recovery loop, one can normalize the different values of such that the average TED gain equals a fixed value.

5. SIMULATION RESULTS FOR A PRML SYSTEM

Receivers for PRML systems typically use a linear equalizer followed by a Viterbi detector. The equalizer tries to shape the channel response h_k to an acceptably shorter target response g_k in order to limit the implementation complexity of the Viterbi detector (VD). A discrete-time model of a typical PRML system is shown in Fig. 3.

By way of illustration we consider run-length-limited data with run-length parameters $(\mathrm{d},\mathrm{k}) = (1,7)$ transmitted over an idealized optical storage channel according to the Braat-Hopkins model [8]

$$H(f) = \begin{cases} \frac{\sin(\ f)}{\ f}\left(\cos^{-1}|\frac{f}{f_c}| - \frac{f}{f_c}\sqrt{1-(\frac{f}{f_c})^2}\right), & |f| < f_c, \\ 0, & |f| \geq f_c \end{cases}$$

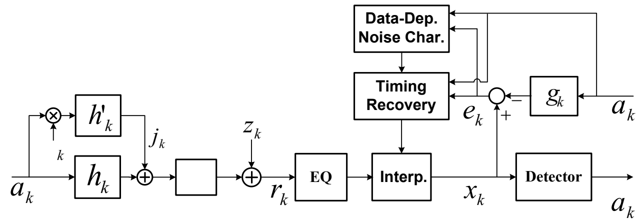

Figure 3: A discrete-time model of a PRML system including jitter noise.

where f_c denotes the optical cut-off frequency, fixed at $1/3$ in the sequel. These choices reflect the system described in [9]. The channel output is corrupted by two different noise components. The first one is the jitter noise j_k caused by a random phase shift $_k$, which is assumed to be independent and identically distributed (i.i.d) Gaussian with zero mean and variance ^{2}. The response h'_k denotes the derivative of h_k. The second noise component is additive white Gaussian noise z_k with zero mean and variance ^2_z. We define two SNR measures: a signal to jitter noise ratio (SJNR) and a signal to additive noise ratio (SANR) given by

$$\mathrm{SJNR} = \frac{_k h_k^2}{^2 \; _k h_k'^2} \quad \text{and} \quad \mathrm{SANR} = \frac{_k h_k^2}{^2_z}.$$

The channel output r_k is first filtered by the equalizer and then interpolated at a delay $-$ where is provided by the timing recovery subsystem. A six-tap Lagrange interpolator [10] is used. The equalized and interpolated signal x_k is subtracted from a reference signal $(g * a)_k$ to produce an error signal e_k. This error signal is used by the timing recovery subsystem to adjust the interpolation phase and by the noise characterization block to estimate the noise prediction parameters. The performance of the timing recovery scheme is measured as the variance of the interpolation phase-error, i.e. $^2 = E[(\; - \;)^2]$. A 5-tap target response $g = [0.17, 0.5, 0.67, 0.5, 0.17]$ and a 9-tap equalizer are used.

Before comparing our timing recovery algorithm with the ZF algorithm, a few steps are needed. First, the equalizer is trained using the LMS algorithm and then fixed. Second, noise characterization is achieved in the absence of any phase-error using $L = 1$ and $K_1 = K_2 = 3$. Third, a calibration process is used in order to ensure that both the ZF TED and our ML TED have the same open loop gain. The same loop filter is used for both timing recovery schemes. The gain of our algorithm compared to the ZF algorithm is measured as the ratio between the phase-error variance ^{2} of the ZF algorithm and that of our algorithm. Fig. 4 shows this gain as function of SJNR for different values of SANR. In order to indicate the range of practical interest of SJNR and SANR, VD bit-error rate (BER) values are shown for some specific SJNR and SANR values at ideal sampling, i.e. $=$. It is apparent that the gain of our timing recovery scheme is more pronounced at low values of SJNR compared to SANR, i.e. low BER. This is of practical interest because timing recovery performance becomes really critical at these signal to noise ratio values. If the noise is mainly data-independent, i.e. the additive noise is dominant compared to jitter noise, than the gain in timing jitter vanishes simply because the ZF scheme becomes optimal.

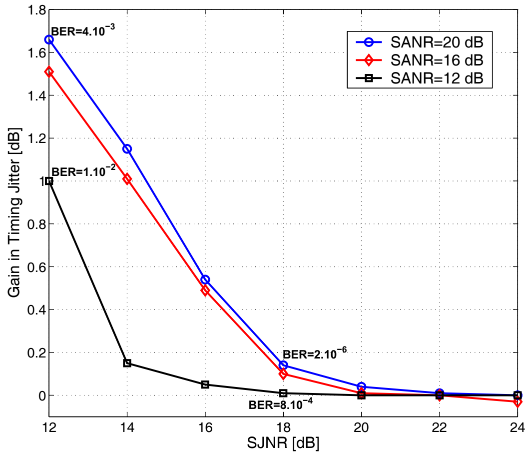

Figure 4: Gain in timing jitter relative to the ZF scheme for $L = 1$ and $K_1 = K_2 = 3$.

6. CONCLUSIONS

In this paper a new timing recovery algorithm for channels with data-dependent noise was presented. Based on a Gauss-Markov correlated noise model, a maximum-likelihood timing recovery algorithm was derived. The new algorithm incorporates data-dependent noise predictors. A simple adaptation scheme was proposed to estimate and track the noise prediction parameters. Simulation results for a partial response maximum-likelihood system show that significant improvements in timing jitter may be obtained at low signal to noise ratios in the presence of data-dependent noise.

REFERENCES

[1] J.W.M. Bergmans, *Digital Baseband Transmission and Recording*, Kluwer Academic Publishers, 1996.

[2] K.H. Mueller and M. Müller, "Timing Recovery in Digital Synchronous Data Receivers," *IEEE Trans. Commun.*, vol. COM-24, no. 5, pp. 516–531, May 1976.

[3] S.U.H. Qureshi, "Timing recovery for equalized partial-response systems," *IEEE Trans. Commun.*, vol. COM-24, no. 5, pp. 516–531, May 1976.

[4] R.D. Cideciyan *et al.*, "A PRML system for digital magnetic recording," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 38–56, January 1992.

[5] A. Kavcic and J. M. F. Moura, "The Viterbi Algorithm and Markov Noise Memory," *IEEE Trans. Magn.*, vol. 46, no. 1, pp. 291–301, Jan. 2000.

[6] A. Gameiro, "Baud sampling bit synchroniser for channels with data dependent noise," *Electronics Letters*, vol. 34, no. 21, pp. 2000–2002, 15th Oct. 1998.

[7] F.M. Gardner, *Phaselock Techniques. 2nd ed.*, New York: Wiley, 1979.

[8] G. Bouwhuis *et al.*, *Principles of Optical Disc Systems*, Adam Hilger Ltd, Bristol, UK, 1985.

[9] T. Narahara *et al.*, "Optical Disc System for Digital Video Recording," *Jpn. J. Appl. Phys.*, vol. 39, Pt. 1, no. 2B, pp. 912–919, Feb. 2000.

[10] M. Abramowitz and I. Stegun, *Handbook of mathematical functions (9th edition)*, Dover Publications, New York, 1970.