Viterbi Equalizer With Analytically Calculated Branch Metrics for Optical ASK and DBPSK

Torsten Freckmann and Joachim Speidel, *Member, IEEE*

***Abstract*—We analytically derive probability density functions of the sampled electrically filtered photo current for optical amplitude-shift keying and differential binary phase-shift keying using the Karhunen–Loève series expansion method. This allows us to determine exact branch metrics for a Viterbi equalizer. We then compare the performance of histogram-based Viterbi equalizers with those using analytically calculated branch metrics.**

***Index Terms*—Maximum likelihood sequence estimation (MLSE), probability density function (pdf), Viterbi equalizer (VE).**

I. Introduction

INCREASING bit rates for optical communications systems require equalizers that can efficiently mitigate fiber impairments like chromatic dispersion (CD) and polarization mode dispersion (PMD), especially beyond 10 Gb/s. Recently, equalizer concepts that have been successfully applied to electronic wireline and wireless systems are considered for optical communications [1]–[3]. Among these, the Viterbi equalizer (VE) is of special interest, as it can provide minimal bit-error probability (BEP) by performing a maximum likelihood sequence estimation (MLSE) using the Viterbi algorithm if the transmission bits are equally distributed [4].

The problem arising when MLSE is applied to the nonlinear and non-Gaussian optical systems (where in the scope of this letter the nonlinearity is caused by the square law detection) is the determination of the branch metrics (BM), i.e., the photo current statistics. Several approximations have been investigated, namely Gaussian or chi-square functions, e.g., in [5]. In this letter, we derive BM analytically and assess the performance of VE for optical amplitude-shift keying (ASK) and differential binary phase-shift keying (DBPSK) in the case of dominating amplified spontaneous emission (ASE) noise.

The derivation of the probability density function (pdf) of the sampled and electrically filtered photo current, and hence the determination of the BM, is based on the Karhunen–Loève series expansion (e.g., [6] and [7]) and will be called the KLSE method. It takes into account arbitrary impulse shaping at the transmitter (TX), fiber-induced impairments, and arbitrary receiver (RX) filters. Therefore, the correlation of the noise (due to the RX filters) and its signal dependency (due to the square law detector) are completely considered. As a consequence, the results presented here are lower bounds for the BEP for the systems described above.

Manuscript received July 18, 2005; revised September 30, 2005.

The authors are with the Institute of Telecommunications, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: freckmann@inue.uni-stuttgart.de; speidel@inue.uni-stuttgart.de).

Digital Object Identifier 10.1109/LPT.2005.860054

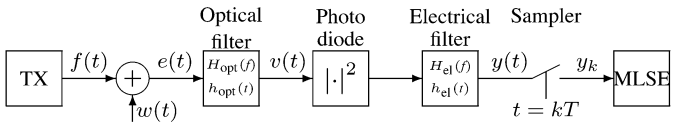

Fig. 1. Equivalent low-pass model of ASK system.

II. Calculation of Branch Metrics

In Fig. 1, the equivalent low-pass model of the investigated ASK system is shown. The TX consists of a continuous wave laser which is modulated externally by a Mach–Zehnder modulator, whose electrical drive signal is the output of an impulse shaper with given impulse response. The RX consists of an optical filter, a photo diode, and an electrical filter followed by the MLSE equalizer. For DBPSK, a precoder in the TX and a delay and add interferometer filter (DAF) followed by a balanced detector in the RX are needed in addition.

We now briefly sketch the KLSE method used to calculate the pdf of the sampled photo current y_k in Fig. 1, for ASK. We assume that the transmitted bit sequence, and hence all deterministic signals, are periodic with a period of K symbol durations and that ASE is the dominant noise source. The RX noise $w(t)$ in Fig. 1 can then be modeled as a complex additive white Gaussian noise with zero mean in front of the optical RX filter [6]. The optical field $v(t)$ at the input of the photo diode may thus be expressed as the sum of a signal component $x(t)$ and a noise component $n(t)$. We can now expand $x(t)$ and $n(t)$ in a series. However, instead of using the Fourier expansion as in [6], we use a Dirac-impulse basis as in [7]. In this case, the expansion coefficients x_m and n_m are just given by the sample values of $x(t)$ and $n(t)$ at time instants $t = m\Delta t$, where $\Delta t = T/M$ and M is the oversampling factor. Now, combining the expansion coefficients to vectors $\mathbf{x}$ and $\mathbf{n}$ containing $N = K \cdot M$ elements, the RX sample y_k may be written as a quadratic form [7]

$$y_k = \mathbf{v}_k^H \mathbf{H}_{\rm el} \mathbf{v}_k \tag{1}$$

where $\mathbf{v}_k = \mathbf{x}_k + \mathbf{n}_k$ is a random vector consisting of samples of $v(t)$ at time instants $t = (kM - m)\Delta t$ $(m = 0, 1, \ldots, N-1)$. $\mathbf{H}_{\rm el}$ is a diagonal matrix with samples of the electrical filter impulse response $h_{\rm el}(t)$ on the main diagonal.

We can now derive the moment generating function (MGF) $M_{y_k}(p)$ of the RX sample y_k, which is defined as the expected value of $\exp(p \cdot y_k)$. Using (1), we get

$$M_{y_k}(p) = E\left[e^{p \cdot \mathbf{v}_k^H \mathbf{H}_{\rm el} \mathbf{v}_k}\right] = \int_{-\infty}^{\infty} e^{p \cdot \mathbf{v}_k^H \mathbf{H}_{\rm el} \mathbf{v}_k} \cdot p_{\mathbf{v}_k}(\mathbf{v}_k) d\mathbf{v}_k. \tag{2}$$

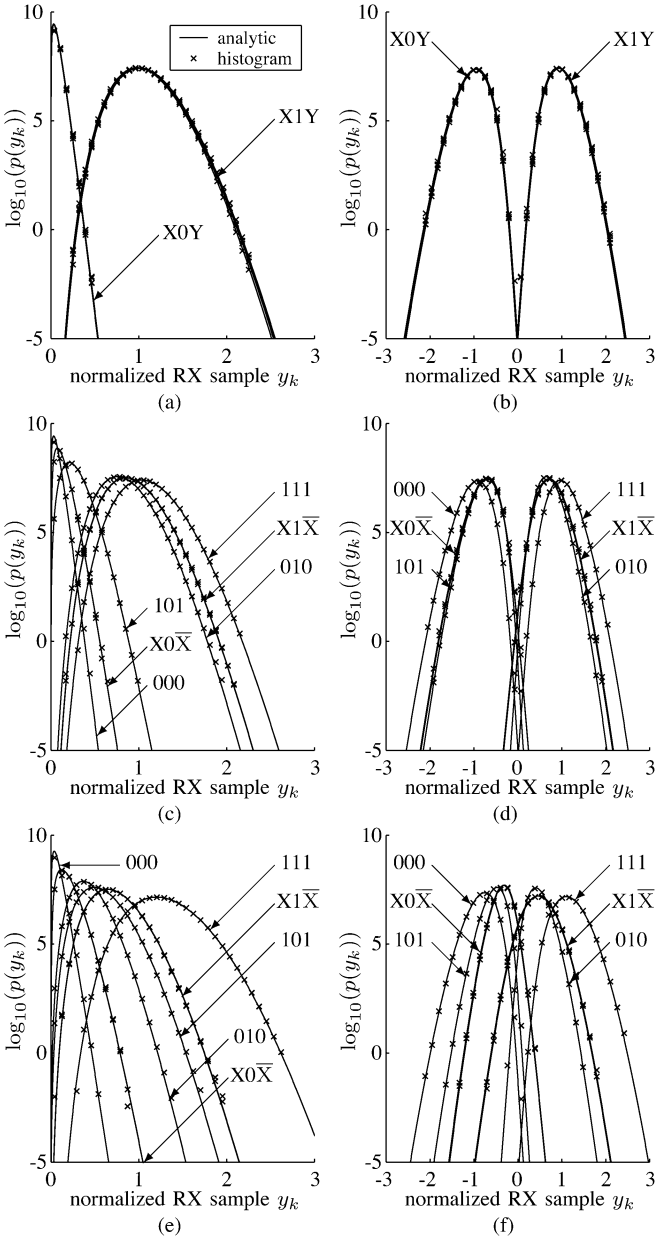

Fig. 2. PDF for (a) ASK and (b) DBPSK at $R_d = 0$ ps/nm. (c) ASK and (d) DBPSK at $R_d = 85$ ps/nm. (e) ASK and (f) DBPSK at $R_d = 170$ ps/nm.

Since the optical filter is linear and time invariant, we can conclude that the signal vector $\mathbf{v}_k$ is multivariate Gaussian distributed with the pdf [8]

$$p_{\mathbf{v}_k}(\mathbf{v}_k) = \pi^{-N}|\mathbf{C}|^{-1}e^{-(\mathbf{v}_k-\mathbf{x}_k)^H\mathbf{C}^{-1}(\mathbf{v}_k-\mathbf{x}_k)} \tag{3}$$

where $\mathbf{C}$ is the covariance matrix of $\mathbf{v}_k$. Their elements are given by the optical filter transfer function and the optical signal-to-noise ratio (OSNR) [7]. After inserting (3) into (2) and applying the Karhunen–Loève transform (KLT) as in [9], it can be shown [7] that

$$M_{y_k}(p) = \prod_{n=1}^{N} \frac{e^{p\lambda_n|z_n|^2/(1-p\lambda_n)}}{(1-p\lambda_n)^2}. \tag{4}$$

λ_n in (4) are the eigenvalues of $\mathbf{H}_{\mathrm{el}} \cdot \mathbf{C}$ and z_n are the elements of the vector $\mathbf{z}$, resulting when the KLT is applied to $\mathbf{x}$.

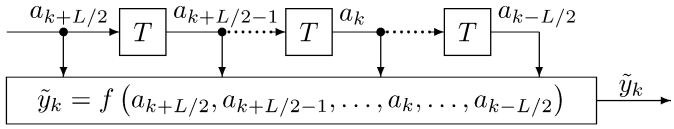

Fig. 3. Nonlinear state-based system model.

The pdf of the RX sample y_k is now given by the inverse Laplace transform of $M_{y_k}(-p)$. However, since we observed numerical problems in evaluating this transform, we first determine the cumulative distribution function (CDF) $F_{y_k}(y_k)$ of y_k by means of the integral

$$F_{y_k}(y_k) = -\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{M_{y_k}(p)}{p} e^{-py_k} dp \tag{5}$$

which is evaluated with the method of steepest descent [10]. We finally obtain the pdf as a first derivative of the CDF.

The extension of the KLSE method to DBPSK transmission is presented, e.g., in [7] and [11]. The main difference in the derivation is that the incorporated vectors and matrices are two times as large as for ASK, if balanced detection is assumed.

Fig. 2 shows some examples of the calculated pdf and the corresponding histograms for nonreturn-to-zero (NRZ)-ASK and NRZ-DBPSK for different values of the residual dispersion R_d and for the assumption that the memory length of the overall transmission system is restricted to three symbols. As a consequence, eight different pdf are possible, one for each binary sequence of length three. The graphs for the different sequences are labeled with the corresponding binary notation, where "X" and "Y" may be either "0" or "1" and "–" denotes logical NOT.

The analytical pdf were obtained with a 2^7 de Bruijn bit sequence. For the histograms, Monte Carlo (MC) simulations of 10^7 bits using a quantization of 5 bits were performed. The pdf are shown at OSNR $= 15$ dB for $R_d = 0, 85$, and 170 ps/nm in Fig. 2(a,b), (c,d), and (e,f), respectively. As expected, the pdf for all sequences with a one in the middle as well as those with a zero in the middle fall together for $R_d = 0$ ps/nm, since in this case we observe almost negligible intersymbol interference (ISI) only introduced by the RX filters. For increasing R_d, and hence increasing ISI, the pdf drift apart as shown in Fig. 2(c)–(f). The pdf for DBPSK are nonsymmetrical, since the two output signals of the DAF do not show a symmetry either.

III. Viterbi Equalizer for Nonlinear Systems

The VE is well known in communication theory, see, e.g., [4]. Applying it to the nonlinear optical system is straightforward as long as an overall finite memory system model can be derived [1]. Let a_k denote the binary TX symbols at symbol duration T and $y(t)$ the noiseless RX signal. Then, the discrete-time output sequence $\tilde{y}_k$ of the system model in Fig. 3 has to correspond to the transmission system output sequence y_k in Fig. 1. Assuming binary input and a model with memory of L symbols, the output $\tilde{y}_k$ of the state model is described, in general, by a finite-state table $f(\cdot)$ with 2^{L+1} entries [1]

$$\tilde{y}_k = f\left(a_{k+L/2}, \ldots, a_k, \ldots, a_{k-L/2}\right) = f\left(\mathbf{a}_k\right) \tag{6}$$

addressed by the current and $L/2$ preceding as well as $L/2$ succeeding symbols. The VE will then operate on a Trellis with 2^L states with 2^{L+1} branches. With the MLSE decision criterion,

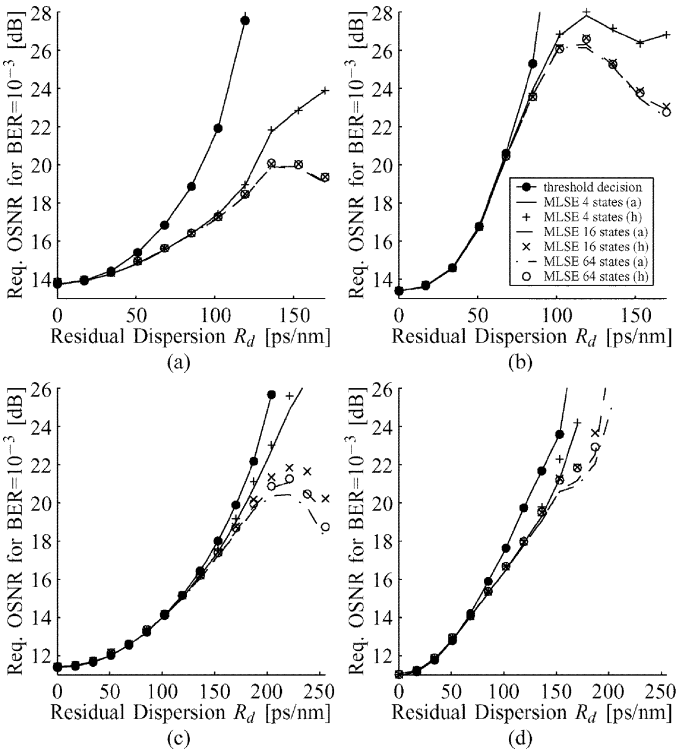

Fig. 4. Required OSNR for BER $= 10^{-3}$ versus residual dispersion for (a) NRZ-ASK, (b) RZ-ASK, (c) NRZ-DBPSK, and (d) RZ-DBPSK.

the most probable sequence $\hat{\mathbf{a}} = (\hat{a}_1, \ldots, \hat{a}_K)$ out of the set $\mathbf{A}$ of all possible sequences $\mathbf{a}$ is selected, i.e.,

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}\in\mathbf{A}} p(\mathbf{y}|\mathbf{a}) = \arg\max_{\mathbf{a}\in\mathbf{A}} \prod_k p(y_k|\mathbf{a}_k). \qquad (7)$$

The BM $p(y_k|\mathbf{a}_k)$ in (7) can be evaluated using (4) and (5).

IV. Simulation Results

We assess the performance of VE with different memories of $L=2$, $L=4$, and $L=6$ via MC simulations of 10^6 symbols. The bitrate of $R_b = 42.7$ Gb/s includes 6.75% forward-error correction overhead. For all simulations, rectangular impulse shaping is assumed. For NRZ [return-to-zero (RZ)] the RX use second (first)-order optical Gaussian bandpass filters with a full 3-dB bandwidth of $1.15 \cdot R_b$ ($1.35 \cdot R_b$) and third (third)-order electrical Bessel low-pass filters with a 3-dB cutoff frequency of $1.35 \cdot R_b$ ($0.85 \cdot R_b$).

It has been shown, e.g., in [1], that BEP can be minimized with respect to the sampling instant, which depends on CD. However, since our main focus is to compare VE performance for analytically calculated BM according to (5), to histogram-based metrics, we consider sampling at time instants $t = kT$.

Fig. 4 shows the required OSNR to achieve BER $= 10^{-3}$ versus residual dispersion R_d for ASK and DBPSK, respectively. From Fig. 4(a) and (b), one can observe that the results of the histogram-based approach agree with those of the KLSE method for both NRZ-ASK and RZ-ASK. It is, moreover, obvious that a VE with memory $L = 2$ is sufficient as long as $R_d \leq 100$ ps/nm; whereas, the high-complexity RX with $L = 6$ does not improve the performance even up to $R_d = 170$ ps/nm. For DBPSK in Fig. 4(c) and (d), in principle the same observations can be made. Again the four-state VE is sufficient as long as $R_d \leq 150$ ps/nm. However, the 64-state VE shows an about 0.5–0.7-dB better performance for NRZ than the 16-state VE if $R_d \geq 200$ ps/nm. Moreover, the KLSE method requires about 0.5–0.7-dB less OSNR than the histogram-based approach in this range. For RZ-DBPSK, the KLSE method outperforms the histogram-based VE as well by 0.7–1 dB for $R_d \geq 170$ ps/nm. However, the performance improvement of the 64-state VE compared to the 16-state VE is only about 0.3 dB.

The fact that the required OSNR may decrease while the residual dispersion increases can be explained by the average distance between different sequences, which also increases to some extent even though the eye is completely closed. The greater the distance between two possible sequences, the easier they can be distinguished by the VE.

V. Conclusion

We have evaluated analytical BM for optical ASK and DBPSK using the KLSE method. The derivation was neither restricted to specific impulse shaping at the TX nor to special pre- and postdetection filters at the RX. It turned out that the more practical histogram-based VE shows almost identical results as the one with analytic BM, whose determination is too complex for online calculation. Moreover, we observed that a state-based system model with memory two is sufficient to describe the nonlinear optical system up to $R_d = 100$ ps/nm for ASK and $R_d = 150$ ps/nm for DBPSK at 42.7 Gb/s. If more residual dispersion should be managed, the memory length of the system model has to be increased.

References

[1] H. F. Haunstein, W. Sauer-Greff, A. Dittrich, K. Sticht, and R. Urbansky, "Principles for electronic equalization of polarization-mode dispersion," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1169–1182, Apr. 2004.

[2] H. Bülow and G. Thielecke, "Electronic PMD mitigation—From linear equalization to maximum-likelihood detection," in *Proc. Opt. Fiber Commun. Conf.*, vol. 3, Anaheim, CA, Mar. 2001, Paper WAA3.

[3] M. Cavallari, C. R. S. Fludger, and P. J. Anslow, "Electronic signal processing for differential phase modulation formats," in *Proc. Opt. Fiber Comm. Conf.*, vol. 1, Los Angeles, CA, Feb. 2004, Paper TuG2.

[4] J. G. Proakis, *Digital Communications*, 4th ed, S. W. Director, Ed. New York: McGraw-Hill, 2001, Series Elect. Comput. Eng.

[5] F. Buchali, G. Thielecke, and H. Bülow, "Viterbi equalizer for mitigation of distortions from chromatic dispersion and PMD at 10 Gb/s," in *Proc. Opt. Fiber Comm. Conf.*, vol. 1, Los Angeles, CA, Feb. 2004, Paper MF85.

[6] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering," *J. Lightw. Technol.*, vol. 18, no. 11, pp. 1493–1503, Nov. 2000.

[7] T. Freckmann and J. Speidel, "Viterbi-equalizer with analytically calculated branch metrics for optical DBPSK and ASK with dominating ASE noise," in *Proc. 6th ITG Conf. Photonic Networks*, ITG Fachbericht, Leipzig, Germany, May 2005, pp. 169–175.

[8] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, no. 1/2, pp. 212–215, Jun. 1956.

[9] G. L. Turin, "The characteristic function of hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, no. 1/2, pp. 199–201, Jun. 1960.

[10] C. W. Helstrom, "Distribution of the filtered output of a quadratic rectifier computed by numerical contour integration," *IEEE Trans. Info. Theory*, vol. 32, no. 7, pp. 450–463, Jul. 1986.

[11] J. Wang and J. M. Kahn, "Impact of chromatic and polarization-mode dispersions on DPSK systems using interferometric demodulation and direct detection," *J. Lightw. Technol.*, vol. 22, no. 2, pp. 362–371, Feb. 2004.