Viterbi-Equalizer with Analytically Calculated Branch Metrics for Optical DBPSK and ASK with Dominating ASE Noise

Torsten Freckmann and Joachim Speidel, Institut für Nachrichtenübertragung, Universität Stuttgart, Pfaffenwaldring 47, D-70569 Stuttgart, Email: freckmann@inue.uni-stuttgart.de

Abstract

We show how to evaluate the analytically exact probability density functions of the sampled electrically filtered photo current in optical amplitude-shift keying and differential binary phase-shift keying transmission systems. This allows to determine exact branch metrics for the application of a Viterbi equalizer to mitigate fiber impairments. The method to calculate these probability density functions is based on the Karhunen-Loeve series expansion. It is applicable to receivers with direct detection where the photo current can be expressed in terms of a quadratic form of multivariate normal distributed random variables. We asses the performance of Viterbi equalizers with different numbers of states via Monte Carlo simulations and compare the results obtained with the analytically calculated branch metrics to a histogram based Viterbi equalizer as well as to systems without equalization and symbol-by-symbol hard decision.

1 Introduction

The interest in increasing bit rates for optical communication systems requires equalizer concepts that can efficiently mitigate fiber impairments like chromatic dispersion (CD) or 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]-[4]. Among these the Viterbi equalizer (VE) is of special interest, as it is known from theory that it can provide minimal bit error probability (BEP). If the transmission bits are equally distributed the VE can achieve this goal by performing a maximum likelihood sequence estimation (MLSE) using the Viterbi algorithm (VA) and intersymbol interference (ISI) is mitigated [5].

The problem arising when applying MLSE and the VA to the nonlinear and non-Gaussian optical channel is the determination of the branch metrics, i.e. the photo current statistics. In [6] and [7] several different approximate solutions were investigated. Among those are the approximation of the probability density function (PDF) of the electrically filtered photo current by Gaussian or chi-square functions. However, none of these methods is exact for arbitrary pre- and post-detection filtering. In this paper we show how to evaluate the exact branch metrics and asses the performance of VE for optical amplitude-shift keying (ASK) and differential binary phase-shift keying (DBPSK) in case of dominating amplified spontaneous emission (ASE) noise.

The evaluation of the PDF of the sampled and electrically filtered photo current and hence the determination of the branch metrics for the VE is based on the Karhunen-Loeve series expansion (e.g. [8],[9]) and will be called the KLSE method in the following. It allows taking account of arbitrary transmitter (TX) and receiver (RX) filters as well as arbitrary impulse shaping at the TX and fiber induced impairments. Therefore, both the correlation of the noise (due to the RX filters) and its signal dependency (due to the square law operation of the photo detector) are completely considered. As a consequence the results presented here are lower bounds for the BEP for systems described above.

The paper is organized as follows. In section 2 modeling of system and noise is described and the PDF for the samples of the photo current are calculated. The nonlinear state model of the transmission system, which is the basis for the implementation of the VE, is shown in section 3. Numerical results obtained for ASK and DBPSK with VE with different complexity are presented in section 4. A summary in section 5 concludes the paper.

2 Calculation of the PDF of the sampled photo current

2.1 System Model

In **Fig. 1** the basic system setup, which is investigated in this paper, is shown. The dashed blocks have only to be considered for DBPSK and are obsolete in case of ASK transmission. The transmitter in Fig. 1 (a) consists of a pseudo random bit sequence generator (PRBSG) and a continuous wave laser which is modulated externally by a Mach-Zehnder modulator (MZM). The electrical drive signal of the MZM is generated from the bit sequence by impulse shapers with raised-cosine (RC) impulse responses

Freckmann, T.; Speidel, J.: ITG-Fachtagung Photonische Netze, Leipzig, Germany, May 2005, pp. 169-175.

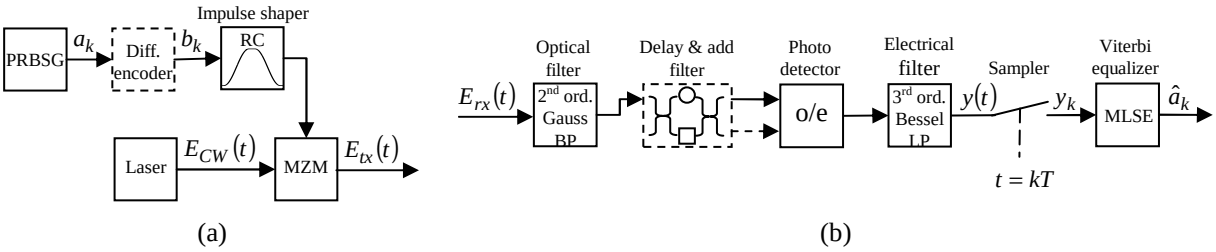


Fig. 1. (a) Transmitter and (b) receiver setup for optical ASK/DBPSK transmission

$$h(t)=\begin{cases}1 & ,|t|\le\frac{T}{2}(1-\alpha)\\ \cos^2\left[\frac{\pi}{4}\frac{2|t|-T(1-\alpha)}{\alpha T}\right] & ,\frac{T}{2}(1-\alpha)<|t|<\frac{T}{2}(1+\alpha)\\ 0 & ,|t|\ge\frac{T}{2}(1+\alpha)\end{cases} \quad (1)$$

$T=1/R_s$ is the symbol or bit duration, α is the roll-off factor. R_s is the symbol rate. It is worth to mention that the following derivation of the PDF of the photo current is not restricted to the assumption of RC impulse shaping, which is considered here as an example only. Return-to-zero (RZ) impulse shaping with 50% duty cycle is achieved by modulating the amplitude of the non-return-to-zero (NRZ) signal in a subsequent MZM with a periodic sequence of electrical Gaussian impulses with a full width at half maximum of $T/2$.

The receiver shown in Fig. 1 (b) consists for ASK of an optical pre-detection and an electrical post-detection filter with one photo diode in between followed by the VE. For DBPSK a delay & add interferometer filter (DAF) with delay $\tau=T$ in one arm and a phase shift $\psi=0$ in the other arm is needed in addition which is followed by a balanced detector using two photo diodes. For NRZ (RZ) both RX use 2nd (1st) order optical Gaussian band-pass and 3rd (3rd) order electrical Bessel low-pass filters. For more details of the transmitter and receiver see e.g. [10].

2.2 PDF calculation

The calculation of the PDF of the sampled photo current, i.e. of the receiver sample y_k in Fig. 1 (b), is based on the Karhunen-Loeve series expansion, as already mentioned above.

2.2.1 Series Expansion of Signal and Noise

In the following we briefly sketch the KLSE method for the ASK receiver whose equivalent low-pass model is shown in **Fig. 2**. It is assumed that the transmitted bit sequence and hence all deterministic signals are periodic with a period of K symbol durations.

Since we assume that ASE noise is the predominant noise source we can model the receiver noise $w(t)$ as an additive white Gaussian noise (AWGN) in front of the optical receiver band-pass filter as indicated in Fig. 2 [8]. The optical field in front of the photo detector $v(t)$ may thus be expressed as the sum of a signal-component $x(t)$ and a noise component $n(t)$ as follows:

$$v(t)=x(t)+n(t)=f(t)*h_{opt}(t)+w(t)*h_{opt}(t), \quad (2)$$

where $*$ denotes convolution and $h_{opt}(t)$ is the optical receiver filter impulse response. As in [8] we can now expand $x(t)$ and $n(t)$ in a series. However, instead of using the Fourier basis for this expansion , we use a dirac-impulse basis as in [9]. In this case the expansion coefficients

$$x_\nu=x(\nu\Delta t) \quad \text{and} \quad n_\nu=n(\nu\Delta t) \quad (3)$$

are just given by the function values of $x(t)$ and $n(t)$ at time instants $\nu\Delta t$, 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, it can be shown that the receiver sample y_k may be written as a quadratic form [9]:

$$y_k=\mathbf{v}_k^H\mathbf{H}_{el}\mathbf{v}_k, \quad (4)$$

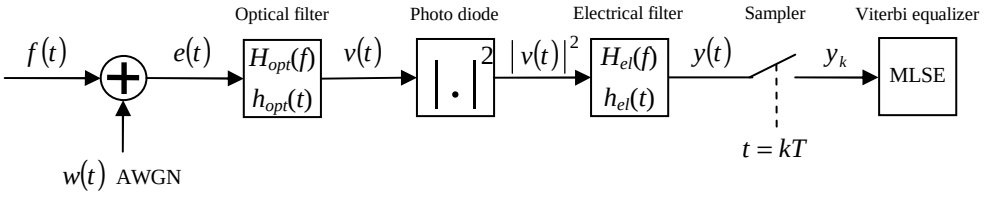


Fig. 2. Equivalent low-pass model of the ASK receiver

where $\mathbf{v}_k = \mathbf{x}_k + \mathbf{n}_k$ is a random vector consisting of the function values of $v(t)$ at time instants $(kM - \nu)\Delta t$. The filter matrix $\mathbf{H}_{el} = \mathrm{diag}[\mathbf{h}_{el}]$ is a diagonal matrix with the expansion coefficients of the electrical filter impulse response $h_{el}(t)$ on the main diagonal. Defined such, (4) equivalently describes the photo detector square-law operation and the convolution of $|v(t)|^2$ with $h_{el}(t)$ after sampling.

2.2.2 Moment Generating Function

We can now derive the moment generating function (MGF) $M_{y_k}(p)$ of the receiver sample y_k which is defined as the expected value of $\exp(p \cdot y_k)$:

$$M_{y_k}(p) = \mathrm{E}\left[e^{p \cdot y_k}\right] = \int_{-\infty}^{\infty} e^{p \cdot y_k} \cdot p_{y_k}(y_k)\,\mathrm{d}y_k \,. \tag{5}$$

Using the generally known property for calculating an expected value of a function of a random variable we can rewrite (5) using (4) as

$$M_{y_k}(p) = \mathrm{E}\left[e^{p \cdot \mathbf{v}_k^H \mathbf{H}_{el} \mathbf{v}_k}\right] = \int_{-\infty}^{\infty} e^{p \cdot \mathbf{v}_k^H \mathbf{H}_{el} \mathbf{v}_k} \cdot p_{\mathbf{v}_k}(\mathbf{v}_k)\,\mathrm{d}\mathbf{v}_k \,. \tag{6}$$

We know from theory that the output of a linear time invariant system is Gaussian distributed if its input is Gaussian distributed. Thus we can conclude that the signal vector $\mathbf{v}_k$ is multivariate Gaussian distributed with the PDF [11]

$$p_{\mathbf{v}_k}(\mathbf{v}_k) = \pi^{-N} |\mathbf{C}|^{-1} e^{-(\mathbf{v}_k - \mathbf{x}_k)^H \mathbf{C}^{-1} (\mathbf{v}_k - \mathbf{x}_k)}, \tag{7}$$

where the elements of the covariance matrix $\mathbf{C}$ are given by the optical filter transfer function [9]:

$$c_{\mu\nu} = r_{nn}(\mu\Delta t - \nu\Delta t) = \mathrm{F}^{-1}\left\{ N_0 \left| H_{opt}(f) \right|^2 \right\}_{|\mu\Delta t - \nu\Delta t} \tag{8}$$

$\mathrm{F}^{-1}\{\cdot\}$ denotes the inverse Fourier-Transform, $H_{opt}(f)$ is the optical filter transfer function and N_0 the spectral noise power density. As we assume noise not only along the signal polarization but along both polarizations, we define the optical signal-to-noise ratio $\mathrm{OSNR} = 10 \cdot \log_{10}[\overline{P}_s / (2 \cdot N_0 \cdot B)]$ with average signal power $\overline{P}_s$ and reference bandwidth $B = 12.5\,\mathrm{GHz}$. Even though it is not explicitly shown in (7) we should mention that the PDF $p_{\mathbf{v}_k}(\mathbf{v}_k)$ is dependent on the transmitted symbol sequence a_k.

After inserting (7) into (6) and applying a coordinate transform as in [12], that diagonalizes the quadratic form in the exponent of the resulting expression and hence transforms the statistically dependent random variables to independent ones, it can be shown [9] that

$$M_{y_k}(p) = \prod_{n=1}^{N} \frac{e^{\frac{p\lambda_n |\eta_n|^2}{1 - p\lambda_n}}}{(1 - p\lambda_n)^m} \,. \tag{9}$$

λ_n in (9) are the eigenvalues of $\mathbf{H}_{el} \cdot \mathbf{C}$ and η_n are the elements of the vector $\boldsymbol{\eta}$ resulting when the coordinate transform mentioned above is applied to the signal vector $\mathbf{x}$. If there is noise in both orthogonal polarizations $m = 2$ (otherwise $m = 1$).

The PDF of the RX sample is now given by the inverse Laplace transform of $M_{y_k}(-p)$

$$p_{y_k}(y_k) = \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} M_{y_k}(p) e^{-py_k}\,\mathrm{d}p \,. \tag{10}$$

However, since we observed numerical problems in evaluating the integral in (10) we first determine the cumulative distribution function (CDF) $F_{y_k}(y)$ of the receiver sample y_k by means of the similar integral

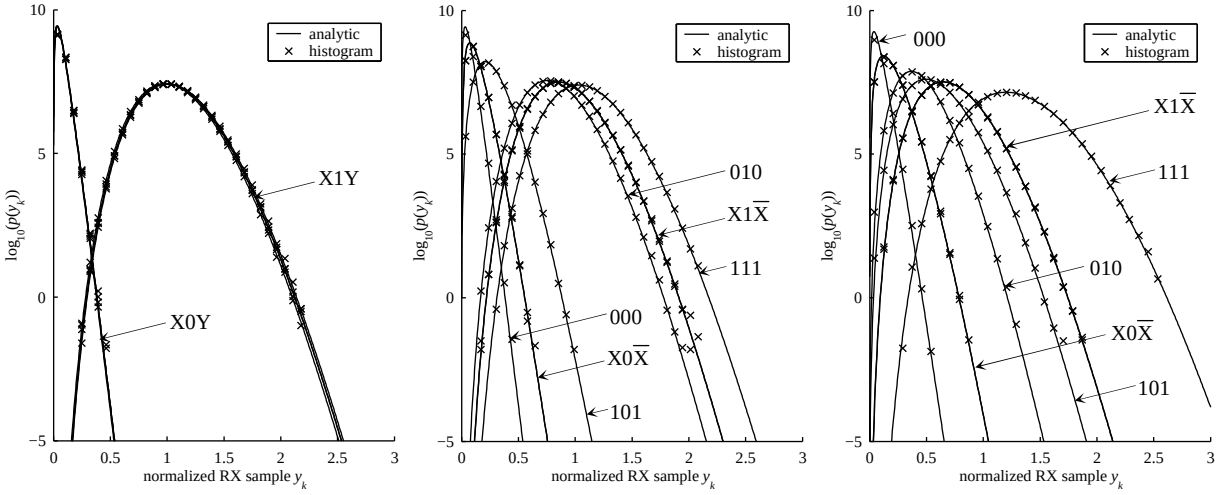
$$F_{y_k}(y_k) = -\frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{M_{y_k}(p)}{p} e^{-py_k}\,\mathrm{d}p \,. \tag{11}$$

This contour integral is evaluated using the method of steepest descent as described e.g. in [13]. We finally obtain the PDF as first derivative of the CDF.

The extension of the KLSE method to DBPSK transmission is presented e.g. in [9] and [14] and is straightforward. The main difference is that the incorporated vectors and matrices are two times as large if – as we do – balanced detection is assumed.

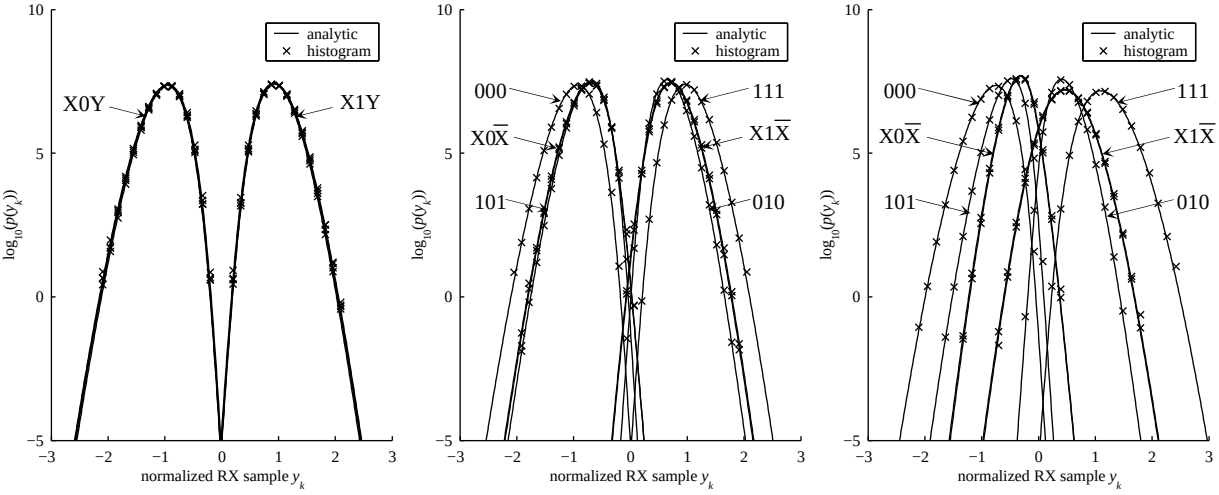
Figs. 3 and **4** show some examples of the calculated PDF and the corresponding histograms for NRZ-ASK and NRZ-DBPSK transmission for different values of the residual dispersion R_d and for the assumption that the memory length of the overall transmission system is restricted to three symbols. As a 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 ‘$\overline{\ }$’ denotes logical NOT.

The analytical PDF were obtained by the KLSE method using a 2^7 De-Brujin bit sequence which may be generated from a $2^7 - 1$ PRBS by inserting an additional zero and which contains every possible bit pattern of length seven exactly once. For the histograms Monte Carlo (MC) simulations of 10^7 symbols using a quantization of 5 bits were performed. The PDF are shown at $\mathrm{OSNR} = 15\,\mathrm{dB}$ for $R_d = 0\,\mathrm{ps/nm}$, $R_d = 85\,\mathrm{ps/nm}$ and $R_d = 170\,\mathrm{ps/nm}$ in Figs 3, 4 (a), (b) and (c), respectively.



(a) (b) (c)

Fig. 3. PDF for ASK at OSNR $=15\,\mathrm{dB}$ and (a) $R_d = 0\,\mathrm{ps/nm}$, (b) $R_d = 85\,\mathrm{ps/nm}$, (c) $R_d = 170\,\mathrm{ps/nm}$



(a) (b) (c)

Fig. 4. PDF for DBPSK at OSNR $=15\,\mathrm{dB}$ and (a) $R_d = 0\,\mathrm{ps/nm}$, (b) $R_d = 85\,\mathrm{ps/nm}$, (c) $R_d = 170\,\mathrm{ps/nm}$

As expected, the PDF for all sequences of length 3 bit with a '1' in the middle as well as those with a '0' in the middle fall together for $R_d = 0\,\mathrm{ps/nm}$ since in this case we observe almost negligible ISI only introduced by the receiver filters. For increasing R_d and hence increasing ISI this is no longer the case and the PDF drift apart as shown in Figs. 3, 4 (b) and (c). The PDF for DBPSK are not symmetrical, which is due to the fact that the eye diagrams of the two output signals of the DAF do not show a symmetry either. For both ASK and DBPSK one can observe that the PDF for the sequence '111' is the rightmost whereas the one for the sequence '000' is the leftmost. This is plausible since a neighboring '1' will induce more ISI than a neighboring '0'. The succession of the PDF for all other sequences must not necessarily stay the same for changing R_d.

3 Viterbi-Equalizer for nonlinear channels

The MLSE receiver and the VA are well known in communication theory, see e.g. [5]. Applying it to the nonlinear optical channel is straightforward as long as an overall finite memory channel model can be derived [1]. Let $\tilde{y}(t)$ denote the noiseless output of the considered system transmitting binary symbols a_k at symbol duration T. Then the discrete-time output sequence $\tilde{y}_k$ of the channel model in **Fig. 5** has to correspond to the overall deterministic transmission channel output sequence $\tilde{y}(kT)$ at sampling instants kT. We assume a model with binary input and a memory of L symbols. In this case the output of the state model is in its most general form described by a

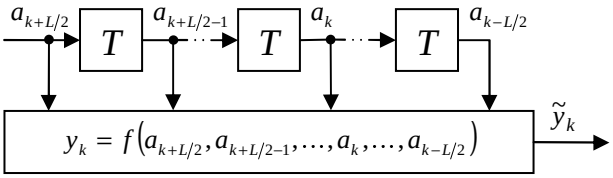


Fig. 5. Nonlinear state-based channel model

finite state table $f(\cdot)$ with 2^{L+1} entries [1]

$$\tilde{y}_k = f(a_{k+L/2}, \ldots, a_k, \ldots, a_{k-L/2}) = f(\mathbf{a}_k) \tag{12}$$

addressed by the current and $L/2$ preceding as well as $L/2$ succeeding symbols, combined in the vector $\mathbf{a}_k$ if the ISI is assumed to spread symmetrically as it is approximately the case for CD. If e.g. $L=2$ only the two neighboring symbols are assumed to contribute to the ISI observed by the symbol between them.

The VE will operate on a Trellis with 2^L states with 2^{L+1} branches in-between them. The MLSE decision criterion is to select the most probable sequence $\hat{\mathbf{a}} = (\hat{a}_1, \ldots, \hat{a}_K)$ out of the set $\mathbf{A}$ of all possible sequences $\mathbf{a}$, i.e.

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}\in\mathbf{A}} p(\mathbf{y} \mid \mathbf{a}) = \arg\max_{\mathbf{a}\in\mathbf{A}} \prod_k p(y_k \mid \mathbf{a}_k). \tag{13}$$

Or equivalently using the log-likelihood function

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}\in\mathbf{A}} \sum_k \log[p(y_k \mid \mathbf{a}_k)], \tag{14}$$

where $\log[p(y_k \mid \mathbf{a}_k)]$ are the metric increments or branch metrics. As already shown, these can be evaluated using (9)–(11).

Table 1
Optimal filter bandwidths for ASK and DBPSK

Format	ASK		DBPSK	
	NRZ	RZ	NRZ	RZ
$\Delta f_{3\text{dB},opt}/R_s$	1.15	1.35	1.15	1.35
$f_{3\text{dB},el}/R_s$	1.05	0.85	1.35	0.85

4 Simulation Results

We asses the performance of MLSE receiver with memory $L=2$, $L=4$ and $L=6$, i.e. with 4, 16 and 64 states respectively, via MC simulations of 10^6 symbols. The considered bitrate of 42.7 Gb/s includes 6.75% FEC overhead, as e.g. for a standard Reed-Solomon RS(255,239) code. Thus the raw bitrate is 40 Gb/s. The roll-off factor of the RC impulse shapers is in the following set to zero, i.e. ideal rectangular impulse shaping is assumed. We consider both NRZ and RZ transmission. The receiver filter bandwidths have been optimized for symbol-by-symbol hard decision (SSHD) with respect to BEP for a fixed OSNR in [9]. In **Table 1** the results for the optimal bandwidths are summarized.

It has been shown in [1] and [15] that optimizing the sampling instant dependent on the amount of CD yields lower BEP performance for the MLSE RX. However, since our main focus is to compare MLSE RX with analytically calculated branch metrics according to (11) to those using histogram-based metrics, we consider sampling at time instants kT.

Figs. 6 and **7** show required OSNR vs. residual dispersion R_d at $\text{BER} = 10^{-3}$ for ASK and DBPSK respectively. The graphs for SSHD are given as a

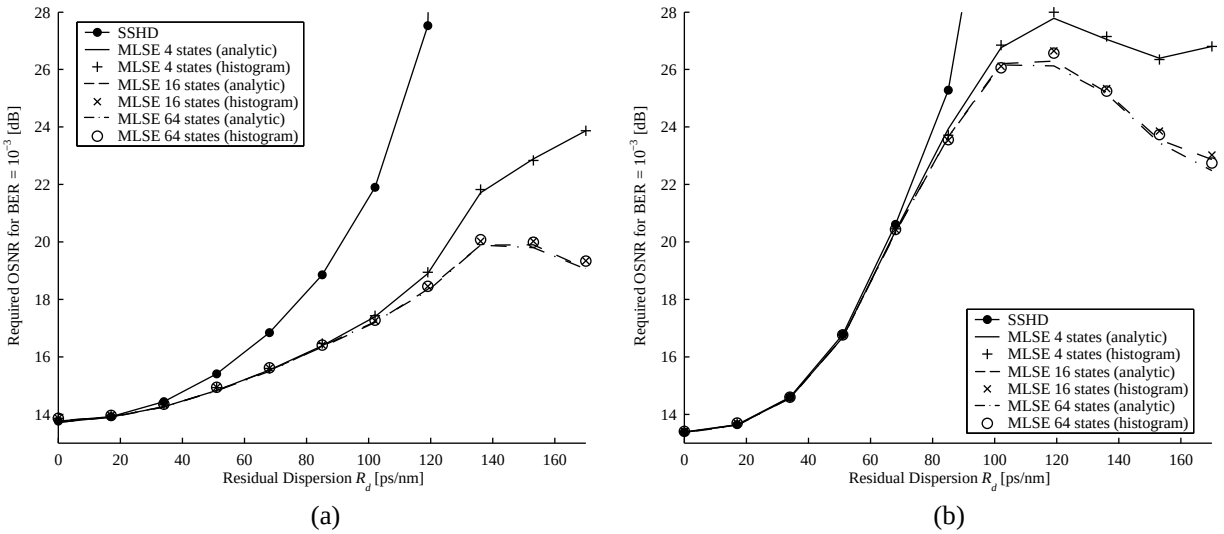


Fig. 6. Required OSNR for $\text{BER} = 10^{-3}$ vs. residual dispersion for (a) NRZ-ASK and (b) RZ-ASK

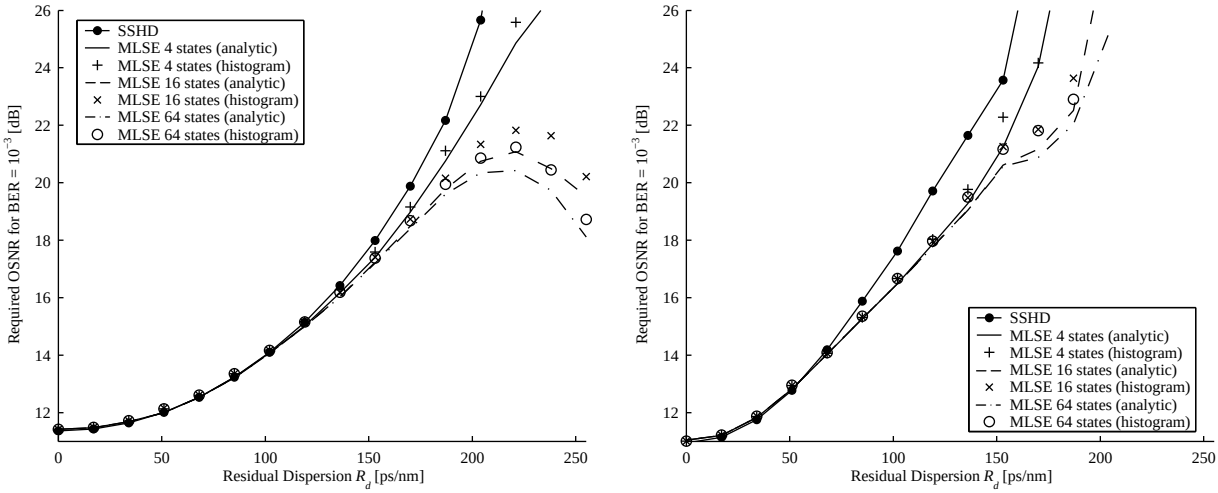


Fig. 7. Required OSNR for $\mathrm{BER} = 10^{-3}$ vs. residual dispersion for (a) NRZ-DBPSK and (b) RZ-DBPSK

reference. From Fig. 6 one can observe that the results obtained for the histogram-based approach agree with those of the KLSE method for NRZ-ASK as well as for RZ-ASK. It is moreover obvious that an MLSE RX with memory $L = 2$ is sufficient until the residual dispersion exceeds $100\,\mathrm{ps/nm}$, whereas the high complexity RX with $L = 6$ does not improve the performance even up to $R_d = 170\,\mathrm{ps/nm}$. For DBPSK in Fig. 7 in principle the same observations as for ASK can be made. Again the MLSE with memory $L = 2$ is sufficient as long as $R_d \le 150\,\mathrm{ps/nm}$. However, now the 64 states MLSE shows an about 0.5–0.7 dB better performance for ASK than the 16 states MLSE if the residual dispersion is larger than $200\,\mathrm{ps/nm}$. Additionally one can observe that in this range the KLSE method needs about 0.5–0.7 dB less OSNR than the histogram-based approach. For RZ-DBPSK the KLSE method outperforms the histogram-based MLSE as well by 0.7–1dB for $R_d \ge 170\,\mathrm{ps/nm}$. However, the performance improvement of the 64 states MLSE compared to the 16 states MLSE is only about 0.3 dB.

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

5 Conclusion

We showed how to evaluate analytically exact branch metrics for optical ASK and DBPSK using the KLSE method. The derivation of the PDF of the photo current was neither restricted to specific impulse shaping at the transmitter nor to special pre- and post-detection filters at the receiver. Therefore the results present lower bounds for BEP. It turned out, that the more practical histogram-based MLSE receiver shows almost identical results as the VE with analytically calculated branch metrics. Moreover we observed that a state-based channel model with memory two is sufficient to describe the nonlinear optical channel up to $R_d = 100\,\mathrm{ps/nm}$ for ASK and $R_d = 150\,\mathrm{ps/nm}$ for DBPSK at $42.7\,\mathrm{Gb/s}$. If more residual dispersion shall be managed, the memory length of the system model should be increased.

6 References

[1] H. F. Haunstein, W. Sauer-Greff, A. Dittrich, K. Sticht, and R. Urbansky, "Principles for Electronic Equalization of Polarization-Mode Dispersion," in *J. Lightwave Technol.*, vol. 22, no. 4, April 2004, pp. 1169-1181.

[2] H. Bülow, and G. Thielecke, "Electronic PMD Mitigation – from Linear Equalization to Maximum-Likelihood Detection," in *Proc. Optical Fiber Conf.*, 2001, pp. WAA3-1–WAA3-3.

[3] M. Cavallari, C. R. S. Fludger, and P. J. Anslow, "Electronic Signal Processing for Differential Phase Modulation Formats," in *Proc. Optical Fiber Conf.*, 2004, pp. TuG2-1–TuG2-3.

[4] T. Rankl, J. Speidel, H. Bülow, and F. Buchali, "Turbo Equalization of Optical Transmission Systems," in *Workshop der ITG-Fachgruppe 5.3.1 Flexible Netzwerkarchitekturen*, Dortmund, February 2005.

[5] J. G. Proakis, *Digital Communications*, 4th Ed., New York: McGraw-Hill, 1995.

[6] F. Buchali, “Anwendung von Viterbi-Entzerrern in optischen Systemen mit dominierendem ASE Rauschen,” in *Workshop der ITG-Fachgruppe 5.3.1 Modellierung photonischer Komponenten und Systeme*, Munich, Mai 2004.

[7] F. Schaich, *Zwischenbericht zum BMBF Projekt MultiTeraNet*, Institut für Nachrichtenübertragung, Universität Stuttgart, 2003.

[8] E. Forestieri, “Evaluating the Error Probability in Lightwave Systems with Chromatic Dispersion, Arbitrary Pulse Shape and Pre- and Postdetection Filtering,” in *J. Lightwave Technol.*, vol. 18, no. 11, November 2000, pp. 1493-1503.

[9] T. Freckmann, *Bit Error Probability Evaluation for Optical Transmission Systems with Direct Detection*, Diploma Thesis, Institut für Nachrichtenübertragung, Universität Stuttgart, 2004.

[10] S. R. Chinn, D. N. Boroson, and J. C. Livas, “Sensitivity of optically preamplified DPSK receivers with Fabry-Pérot filters,” in *J. Lightwave Technol.*, vol. 14, no. 3, March 1996, pp. 370-376.

[11] R. A. Wooding, “The Multivariate Distribution of Complex Normal Variables,” in *Biometrika*, vol. 43, no. 1/2, June 1956, pp. 212-215.

[12] G. L. Turin, “The Characteristic Function of Hermitian Quadratic Forms in Complex Normal Variables,” in *Biometrika*, vol. 47, no. 1/2, June 1960, pp. 199-201.

[13] C. W. Helstrom, “Distribution of the Filtered Output of a Quadratic rectifier Computed by Numerical Contour Integration”, in *Trans. on Information Theory*, vol. 32, no. 4, July 1986, pp. 450-463.

[14] J. Wang, and J. M. Kahn, “Impact of Chromatic and Polarization-Mode Dispersion on DPSK Systems Using Interferometric Demodulation and Direct Detection,” in *J. Lightwave Technol.*, vol. 22, no. 2, February 2004, pp. 362-371.

[15] H. Griesser, J.-P. Elbers, and C. Glingener, “On MLSE Reception of Chromatic Dispersion Tolerant Modulation Schemes,” in *Proc. Tyrrhenian Int. Workshop on Digital Communication*, Pisa, October 2004, pp. 205-212.