Wavelength-Time Coding for Ultra Dense Wavelength Multiplexing

Frank Schaich and Joachim Speidel
Institut für Nachrichtenübertragung, Universität Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart
Email: frank.schaich@inue.uni-stuttgart.de

Abstract

By reducing channel spacing in dense wavelength division multiplexing (U-DWDM), the total bitrate can be significantly increased. To cope with the resulting strong interchannel interference (ICI), a novel coding scheme is presented, which encodes and decodes in both time and wavelength direction (wavelength-time coding, WTC). The scheme is described in quite some detail. Performance is compared to conventional U-DWDM with strict optical filtering to prevent ICI, however to the expense of increased intersymbol interference (ISI).

1 Introduction

Due to the rising demand for broadband internet access, spectral efficiency in $\frac{bit}{sHz}$ of the optical network has to be increased. Several schemes to reach this goal are under investigation in today's research. Multilevel modulation such as DQPSK [1], ASK-DPSK [2], ASK-DQPSK [3], 8-DPSK [4] can be used to transmit more than one bit per symbol. The lower symbol rates lead to smaller spectra, which can be placed more densely without increasing the spectral overlap. A way to double spectral efficiency is to use orthogonal polarizations for data transmission [5]. Also of great interest in today's research activities are more sophisticated coding schemes such as LDPC codes [6] and turbo schemes [7] to mitigate channel distortions.

Within this paper we introduce a novel coding scheme, which reduces the impact of ICI by jointly encoding the data of adjacent wavelength carriers. As a consequence, the channel spacing can be reduced without losing significant robustness against channel distortions. This scheme encodes both in time and wavelength direction and is therefore called wavelength-time coding (WTC).

In section 2 we investigate the ICI and motivate the use of WTC, which is described in more detail in section 3. The performance is presented in section 4. Section 5 concludes the paper.

2 Interchannel Interference (ICI)

When reducing channel spacing for a given bitrate per channel to increase spectral efficiency in U-DWDM systems, ICI becomes more and more significant. **Fig.1** shows the eye-diagrams after photo detection, in case of a single wavelength system (a) and U-DWDM with a channel spacing of 85 GHz (b) and 50 GHz (c). The bit rate is $R_b = \frac{1}{T_b} = 42.7\ \frac{Gbit}{s}$ and the 3 dB-bandwidth of the demultiplexer filter is choosen to 40 GHz. Obviously, the neigbouring channels cause a strong widening of the signal transitions and the eyes tend to close.

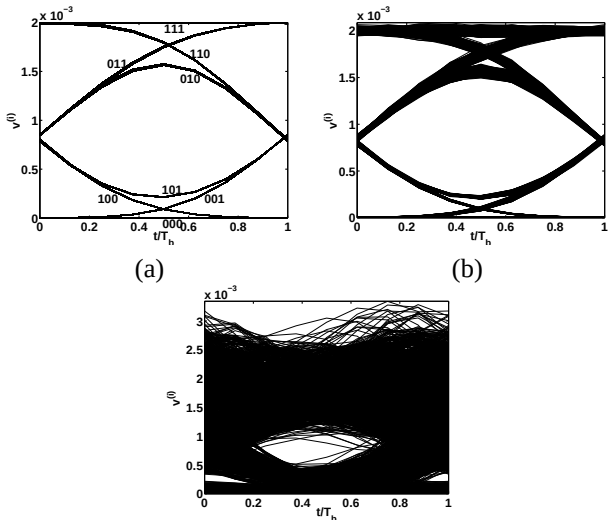


Fig. 1. Eye-diagrams after the photo diode. Single wavelength system (a), U-DWDM system with channel spacing 85 GHz (b) and 50 GHz (c). Bitrate is 42.7 $\frac{Gbit}{s}$ per channel. (The numbers in (a) reflect the underlying bit pattern.)

A detailed investigation and modeling of ICI is given in appendix I.

3 Wavelength-Time Coding (WTC)

As shown in **Fig. 2**, the bit sequence $a_k^{(i)}$ of channel i with wavelength λ_i is fed into the WT encoder ($i = 1, ..., K$; k discrete time). The output bits $c_m^{(i)}$ are separately

Schaich, F.; Speidel, J.: "Wavelength-Time Coding for Ultra Dense Wavelength Multiplexing", ITG-Fachtagung Photonische Netze, Leipzig, Germany, April 2006, pp. 113-120.

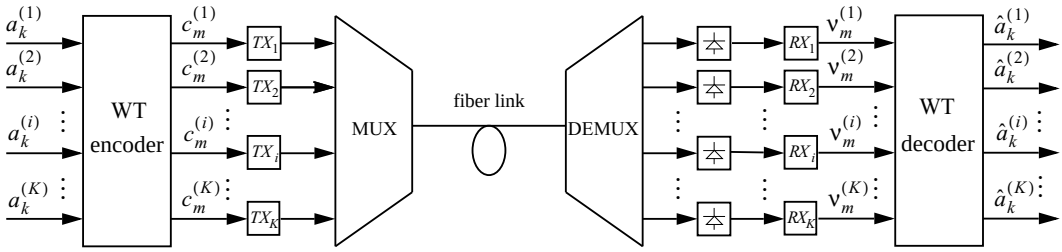


Fig. 2. WDM transmission system

transmitted using intensity modulation before being multiplexed and fed into the fiber link. Amplified spontaneous emission of the optical amplifier at the receiver is considered as the dominant noise source. The signals are demultiplexed and O/E-converted for further processing in the WT-decoder.

3.1 WT encoder

As the name indicates, data is not only encoded along time axis k but also in wavelength direction i. N consecutive bits of M adjacent channels build up an $M \times N$ input matrix A, where $M \leq K$. The WT encoder adds Q "parity bits" per channel yielding an $M \times (N+Q)$ output matrix C with elements $c_{m-\nu}^{(i)}$ ($\nu = 0, ..., N+Q-1$; $i = 1, ..., M$; $m \in \mathbb{Z}$). To show the principle, we assume in the following $M = 2$ adjacent channels. Thus the WT encoder executes the mapping:

$$\underbrace{\begin{pmatrix} a_{k-N+1}^{(i)} \cdots a_k^{(i)} \\ a_{k-N+1}^{(i+1)} \cdots a_k^{(i+1)} \end{pmatrix}}_{A} \mapsto \underbrace{\begin{pmatrix} c_{m-N-Q+1}^{(i)} \cdots c_m^{(i)} \\ c_{m-N-Q+1}^{(i+1)} \cdots c_m^{(i+1)} \end{pmatrix}}_{C}. \quad (1)$$

Let $\frac{1}{T}$ and $\frac{1}{T'}$ be the bitrate of $a_k^{(i)}$ and $c_m^{(i)}$, respectively. As input and output block duration are the same, i.e. $TN = T'(N+Q)$, output bitrate increases to $\frac{1}{T'} = \frac{1}{RT}$, where $R = \frac{N}{N+Q}$ is the coderate. As an example, for $N = 1$ and $Q = 1$ possible mappings are:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The cardinality (number of elements) of the sets A and C are as follows:

$$\begin{aligned} \text{card}\{A\} &= 2^{MN} \\ \text{card}\{C\} &= (2^M - x)^{N+Q}. \end{aligned} \quad (2)$$

x is the number of suppressed bit patterns. Obviously, the cardinality of A and consequently the complexity of the encoder and the decoder grows exponentially with the number of channels M. Thus, to keep complexity low not all K channels of a typical WDM system should be encoded jointly. Instead, we recommend to divide the set of K channels into $\frac{K}{M}$ sub-sets with M channels each and to use $\frac{K}{M}$ wavelength-time encoders and decoders in parallel. This leads to some channels beeing located side by side without beeing encoded together. The consequences will be addressed later in this chapter. In the following we choose $M = 2$. N, the number of bits per codeword in (1), has also an exponential impact on the complexity. Thus, N should be choosen as small as possible. In the following we choose $N = 3$. Finally, the number Q of parity bits per codeword in (1) has to be found. Of course card$\{C\} \geq$card$\{A\}$ has to hold. Thereof we conclude with (2):

$$Q \geq N(\frac{M\text{lg}(2)}{\text{lg}(2^M - x)} - 1). \quad (3)$$

Fig. 3 shows the resulting code rate R as a function of N, when the minimal $Q \in \mathbb{N}$ is choosen and $M = 2$. In this case $x = 1$. Obviously, there is a

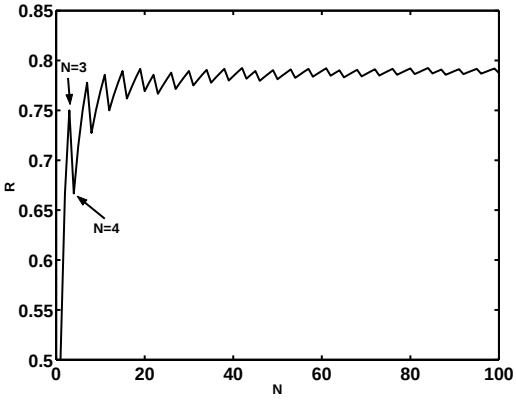


Fig. 3. Code rate R as a function of N for minimal Q

code rate R_{max}, which cannot be exceeded. With (3) and $M = 2$ we get $R_{max} = \frac{lg(3)}{lg(4)} = 0.7925$. For $N = 3$ and $Q = 1$ the code rate is $R = 0.75$ and a good compromise between performance, overhead and complexity is obtained.

In principle, a variety of mappings (1) exists. They should be selected such, that the bit error ratio (BER) at

the receiver output is minimized. We have card$\{C\}$ − card$\{A\}$ degrees of freedom for the mapping. They can be used for different optimization strategies:

I Maximize the Hamming distance.

II Minimize the number of 1's (Hamming weights) in each column of C to further reduce ICI.

III Choose C with the lowest number of '01' or '10' patterns in each row. These patterns are disadvantageous, as they increase ISI within the channel.

3.2 WT decoder

At the receiver in Fig. 2 the original bit sequences have to be recovered by finding estimates $\hat{a}_k^{(i)}$ of $a_k^{(i)}$. We propose a maximum a posteriori (MAP) symbol-by-symbol detector. For the given example $M = 2$ the set of probabilities for all possible codewords have to be calculated:

$$\begin{aligned} &P[c_{m-N-Q+1}^{(i)}, \ldots, c_m^{(i)}, c_{m-N-Q+1}^{(i+1)}, \ldots, c_m^{(i+1)} | \\ &v_{m-N-Q+1}^{(i)}, \ldots, v_m^{(i)}, v_{m-N-Q+1}^{(i+1)}, \ldots, v_m^{(i+1)}] \\ &= \prod_{n=-N-Q+1}^{0} \frac{p(v_{m+n}^{(i)} | c_{m+n}^{(i)}, c_{m+n}^{(i+1)}) P[c_{m+n}^{(i)}, c_{m+n}^{(i+1)}]}{p(v_{m+n}^{(i)})} \\ &\cdot \prod_{n=-N-Q+1}^{0} \frac{p(v_{m+n}^{(i+1)} | c_{m+n}^{(i)}, c_{m+n}^{(i+1)}) P[c_{m+n}^{(i)}, c_{m+n}^{(i+1)}]}{p(v_{m+n}^{(i+1)})}. \end{aligned} \quad (4)$$

The lower index of the received symbols $v_{\cdot\cdot}^{(\cdot\cdot)}$ now indicates discrete time. The MAP detector selects the symbol with the maximal probability as solution. In (4) statistically independent noise is assumed and $p(.|.)$ is the chi^2-distribution of the noisy signal after the photodiode, depending on the underlying bit pattern with the a priori probability $P[c_{m+n}^{(i)}, c_{m+n}^{(i+1)}]$. $p(v_{m+n}^{(i)})$ and $p(v_{m+n}^{(i+1)})$ are the overall probability density functions (pdf) of the photo diode output signal. Without going into detail, we can take ISI into account, which results in an extended version of (4).

3.3 System parameters

Several parameters have an impact on the performance of the system, e.g. channel spacing, 3 dB-bandwidth of the optical filters and the outer encoder, which are discussed in the following. As already mentioned encoding of all K channels of a WDM system jointly would increase complexity of the encoder and decoder too much. Instead, we propose to group the channels into pairs as shown in **Fig. 4** and apply on each pair WT encoding.

3.3.1 Outer encoder

As indicated in Fig. 4 we propose to enhance the scheme by an outer encoder, which is a Reed Solomon encoder RS($N_I + N_C, N_I, N_b$). N_I is the number of information symbols, N_C the number code symbols per RS frame, and N_b bits are used to form a symbol. We study three schemes (A)-(C) as shown in **Fig. 5** The framed elements represent the RS symbols. In (C) the pattern of the RS symbols match the pattern of the WTC symbols after decoding. In the following the symbol error ratio (SER) before the RS decoder is used as a performance figure.

3.3.2 Channel spacing

The smaller the channel spacing is the higher the bitrates per unit bandwidth are. However, ICI is increased. **Fig. 6** shows the required OSNR to achieve $SER = 0.005$ for different bitrates per unit bandwidth. The net bit rate per channel is $40\ \frac{Gbit}{s}$. WTC with mapping II in section 3.1 is used. The two schemes without WTC apply threshold detection and maximum-likelihood sequence estimation (MLSE), respectively. Obviously, WTC is superior for any bitrate under consideration. The higher the bitrate is, the larger the improvement by WTC is. Remember, that ICI increases with bitrate. In the following our investigation sticks to $0.8\frac{bit}{sHz}$. This corresponds to a channel spacing of 50 GHz.

When using the scheme proposed in Fig. 4, a non-uniform channel grid as shown in **Fig. 7** leads to the best performance. To minimize ICI the distance of Δf_{noWTC} between adjacent pairs of spectra has to be sufficiently large, while Δf_{WTC}, the spacing within the pairs, can be small. Beside the channel grid, the 3-dB bandwidths of the optical filters of the multiplexer $f_{3dB,mux}$ and demultiplexer $f_{3dB,demux}$ have an impact on ICI. In appendix II, we show how these parameters can be optimized. The results are given in **Table 1**.

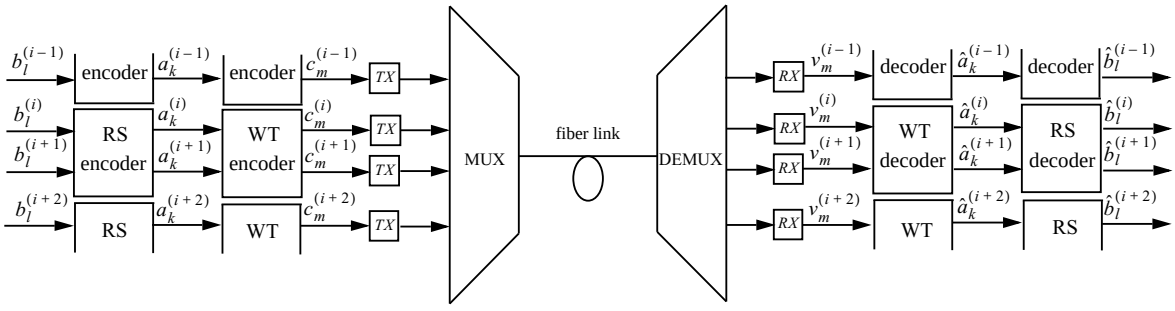


Fig. 4. Proposed transmission system (cut-out)

(A) RS (255, 239, 8) separate for both channels:

$$\cdots \quad \underbrace{b_{k-5}^{(i)} \; b_{k-4}^{(i)} \; b_{k-3}^{(i)} \; b_{k-2}^{(i)} \; b_{k-1}^{(i)} \; b_{k}^{(i)} \; b_{k+1}^{(i)} \; b_{k+2}^{(i)}}_{\text{RS}} \; \underbrace{b_{k+3}^{(i)} \; b_{k+4}^{(i)} \; b_{k+5}^{(i)}}_{\text{RS}} \quad \cdots$$

$$\cdots \quad \underbrace{b_{k-5}^{(i+1)} \; b_{k-4}^{(i+1)} \; b_{k-3}^{(i+1)} \; b_{k-2}^{(i+1)} \; b_{k-1}^{(i+1)} \; b_{k}^{(i+1)} \; b_{k+1}^{(i+1)} \; b_{k+2}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k+3}^{(i+1)} \; b_{k+4}^{(i+1)} \; b_{k+5}^{(i+1)}}_{\text{RS}} \quad \cdots$$

(B) RS (255, 239, 8) jointly for both channels:

$$\cdots \quad \underbrace{b_{k-5}^{(i)} \; b_{k-4}^{(i)} \; b_{k-3}^{(i)} \; b_{k-2}^{(i)}}_{\text{RS}} \; \underbrace{b_{k-1}^{(i)} \; b_{k}^{(i)} \; b_{k+1}^{(i)} \; b_{k+2}^{(i)}}_{\text{RS}} \; \underbrace{b_{k+3}^{(i)} \; b_{k+4}^{(i)} \; b_{k+5}^{(i)}}_{\text{RS}} \quad \cdots$$

$$\cdots \quad \underbrace{b_{k-5}^{(i+1)} \; b_{k-4}^{(i+1)} \; b_{k-3}^{(i+1)} \; b_{k-2}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k-1}^{(i+1)} \; b_{k}^{(i+1)} \; b_{k+1}^{(i+1)} \; b_{k+2}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k+3}^{(i+1)} \; b_{k+4}^{(i+1)} \; b_{k+5}^{(i+1)}}_{\text{RS}} \quad \cdots$$

(C) RS (64, 60, 6) jointly for both channels:

$$\cdots \quad \underbrace{b_{k-5}^{(i)} \; b_{k-4}^{(i)} \; b_{k-3}^{(i)}}_{\text{RS}} \; \underbrace{b_{k-2}^{(i)} \; b_{k-1}^{(i)} \; b_{k}^{(i)}}_{\text{RS}} \; \underbrace{b_{k+1}^{(i)} \; b_{k+2}^{(i)} \; b_{k+3}^{(i)}}_{\text{RS}} \; \underbrace{b_{k+4}^{(i)} \; b_{k+5}^{(i)}}_{\text{RS}} \quad \cdots$$

$$\cdots \quad \underbrace{b_{k-5}^{(i+1)} \; b_{k-4}^{(i+1)} \; b_{k-3}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k-2}^{(i+1)} \; b_{k-1}^{(i+1)} \; b_{k}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k+1}^{(i+1)} \; b_{k+2}^{(i+1)} \; b_{k+3}^{(i+1)}}_{\text{RS}} \; \underbrace{b_{k+4}^{(i+1)} \; b_{k+5}^{(i+1)}}_{\text{RS}} \quad \cdots$$

Fig. 5. Alternative schemes for RS encoding of two adjacent wavelength channels

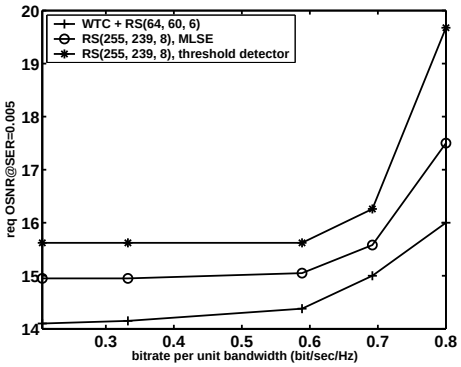


Fig. 6. Required OSNR@SER=0.005 as a function of the bitrate per unit bandwidth

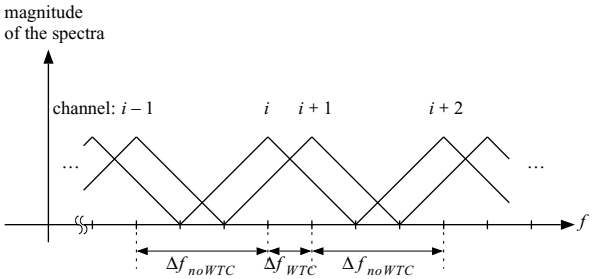


Fig. 7. Non-uniform channel grid

4 Performance of WTC

Performance of WTC is investigated by Monte-Carlo-Simulations. One pair of channels ($M = 2$) with $N = 3$ and $Q = 1$ is considered. The net bitrate is 40 $\frac{Gbit}{s}$ per channel. CD and PMD are assumed to be fully compensated. The received signals are corrupted by optical uncorrelated Gaussian noise with zero mean. Codeword mapping II of section 3.1 is used. The MAP

Table 1
Configurations used for simulations

	Δf_{WTC}	Δf_{noWTC}	$f_{3dB,mux}$	$f_{3dB,demux}$
WTC	45 GHz	55 GHz	70 GHz	40 GHz
no WTC	50 GHz	50 GHz	70 GHz	40 GHz

employs the extended version of (4). **Fig. 8** shows the SER before RS decoding when using WTC compared to systems without WTC. For the schemes without WTC two kinds of detectors are used, namely threshold detector and MLSE. All three schemes of the RS encoder given in Fig. 5 are compared. (C) leads to the best results. Obviously, the system with WTC and scheme C) for the RS encoding outperforms all the others.

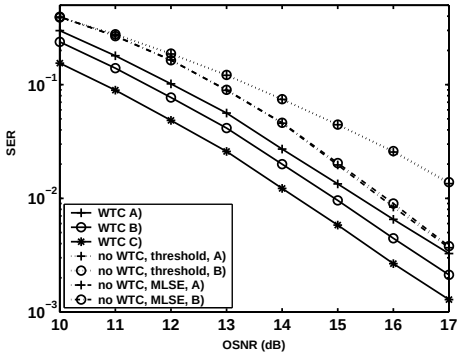


Fig. 8. SER before RS decoding as a function of OSNR for various coding schemes

In **Fig. 9** we have compared the three different WT code designs I, II and III from section 3.1. II performs best followed by III, both being superior to I.

Especially in WDM systems a major reason for signal degradation is the non-linearity of the fiber due to high power loading at the input. In appendix III

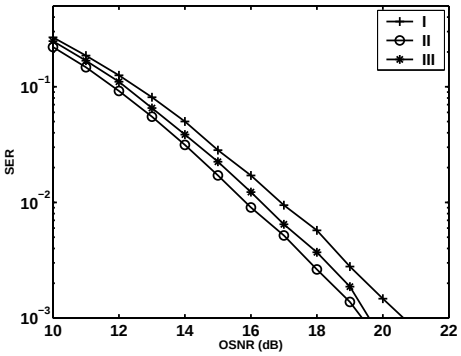


Fig. 9. Simulation results for the different codeword-designs I, II and III in section 3.1

we show, that the mean and the maximal number of '1's and thus the power load at the input of the fibre are reduced, when using WTC. **Table 2** shows the parameters of the distributions and the maximum, the mean and the variance of the random variables are listed (in the case of WTC codword design II is used):

Table 2

Parameters of the distributions

	P_1	P_0	$max[Z]$	$E[Z]$	σ_Z^2
WTC II	0.58	0.42	16	9.27	3.9
no WTC	0.5	0.5	32	16	8

Figs. 10 and **11** illustrate these distributions. Monte-Carlo simulations are used to confirm the results. Thus

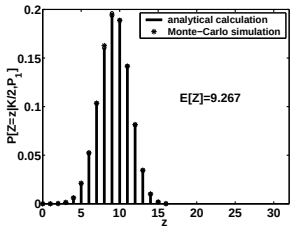


Fig. 10. Distribution of the number of '1's to be transmitted simultaneously when WTC is used

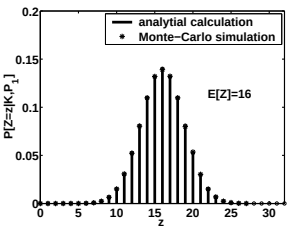


Fig. 11. Distribution of the number of '1's to be transmitted simultaneously when no WTC is used

the maximum and mean power fed into the fibre is reduced, when using WTC. This leads to a minor impact of the non-linearities compared to a system without WTC.

Another advantage over RS coding with MLSE is the fact that the WT decoder and encoder can be highly parallelized to lower hardware speed. This is possible, because encoding and decoding is symbol-by-symbol wise. The MLSE is based on sequence detection and can thus hardly be parallelized leading to higher hardware requirements.

5 Conclusion

We have presented a novel coding scheme for U-DWDM. It allows dense packing of wavelength channels to increase total bitrate. The resulting interchannel interference (ICI) is reduced by encoding along both the discrete-time and wavelength index. Different schemes for the Reed-Solomon (RS) coding are used. In all cases CD and PMD is assumed to be compensated and net bitrate is 40 $\frac{Gbit}{s}$ per channel. We show that the new wavelength-time coding (WTC) method can achieve lower SER at low OSNR. The WTC receiver operates with maximum a posteriori (MAP) symbol-by-symbol detection. We show, that WTC concatenated with an outer RS code performs superior as it provides an OSNR gain of 3 dB and 2 dB over the system without WTC with threshold detection and MLSE, respectively. We have also shown, that ISI per channel can be incorporated into the detection strategy for WTC. It is pointed out that WT encoder and decoder hardware can be parallelized, thus reducing speed requirements, which is crucial in the multi $\frac{Gbit}{s}$ range. As simultaneous '1's in adjacent channels are prohibited by WT encoding, instantaneous power is reduced and consequently the effects of fiber non-linearities are lowered. Degrees of freedom of the variety of WT codes can be used in such a way, that the number of outgoing '01' and '10' bit patterns in time direction is reduced, which minimizes ISI per channel.

APPENDIX I
MODELING OF ICI

To further investigate the ICI we calculate the signal after photo detection of a single channel within a WDM system. **Fig. 12** shows in principal the magnitudes of the spectra of the equivalent baseband representation of a WDM signal before demultiplexing. For simplification the spectra are approximated by triangles. The transfer function of the demultiplexer of channel i is assumed to be constant within the pass-band. Channel spacing is given by Δf.

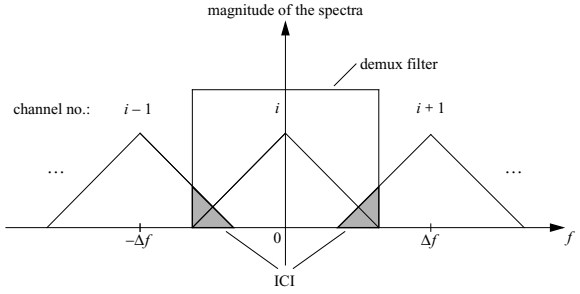


Fig. 12. Magnitudes of the equivalent baseband representation of the WDM system with respect to channel i

The equivalent baseband field signals after demultiplexing within the band of channel i are give by:

$$\vec{\mathrm{E}}_{(I)}^{(i)} = \begin{pmatrix} |\mathrm{E}_{(I),x}^{(i)}| e^{jarc\{\mathrm{E}_{(I),x}^{(i)}\}} \\ |\mathrm{E}_{(I),y}^{(i)}| e^{jarc\{\mathrm{E}_{(I),y}^{(i)}\}} \end{pmatrix} = \begin{pmatrix} A_{(I),x}^{(i)} e^{j\varphi_{(I),x}^{(i)}} \\ A_{(I),y}^{(i)} e^{j\varphi_{(I),y}^{(i)}} \end{pmatrix} \quad (5)$$

with: $I \in \{i-1, i, i+1\}$.

The underline indicates complex numbers and $j = \sqrt{-1}$. Let $\vec{\mathrm{E}}_{(i)}^{(i)}$ be the signal of channel i and $\vec{\mathrm{E}}_{(i-1)}^{(i)}$ and $\vec{\mathrm{E}}_{(i+1)}^{(i)}$ the interference signals coming from the adjacent channels $i-1$ and $i+1$, respectively. If we assume, that the contributions of the remaining channels of the WDM system are neglectable, the photo current at the output of the photo diode is:

$$v^{(i)} \propto |\vec{\mathrm{E}}_{(i-1)}^{(i)} + \vec{\mathrm{E}}_{(i)}^{(i)} + \vec{\mathrm{E}}_{(i+1)}^{(i)}|^2. \quad (6)$$

Inserting (5) in (6) and dropping the superscript (i) to simplify notation, we obtain:

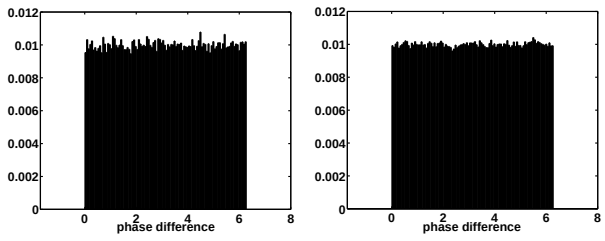
$$\begin{aligned} v^{(i)} \propto \quad & A_{(i-1),x}^2 + A_{(i),x}^2 + A_{(i+1),x}^2 \\ & + A_{(i-1),y}^2 + A_{(i),y}^2 + A_{(i+1),y}^2 \\ & + 2A_{(i-1),x}A_{(i),x}cos(\Delta\varphi_{(i-1,i),x}) \\ & + 2A_{(i),x}A_{(i+1),x}cos(\Delta\varphi_{(i,i+1),x}) \\ & + 2A_{(i-1),x}A_{(i+1),x}cos(\Delta\varphi_{(i-1,i+1),x}) \\ & + 2A_{(i-1),y}A_{(i),y}cos(\Delta\varphi_{(i-1,i),y}) \\ & + 2A_{(i),y}A_{(i+1),y}cos(\Delta\varphi_{(i,i+1),y}) \\ & + 2A_{(i-1),y}A_{(i+1),y}cos(\Delta\varphi_{(i-1,i+1),y}) \end{aligned} \quad (7)$$

with the phase differences:

$$\begin{aligned} \Delta\varphi_{(I1,I2),x} &= \varphi_{(I1),x} - \varphi_{(I2),x} \\ \Delta\varphi_{(I1,I2),y} &= \varphi_{(I1),y} - \varphi_{(I2),y} \end{aligned}$$

$$I1, I2 \in \{i-1, i, i+1\}.$$

Of course $v^{(i)}$ is a function of time t. In the following we consider $v^{(i)}$ at a certain point in time t, which we call a sample value $v^{(i)}$. t is dropped to simplify notation. With a closer look at (7), we see that the signal after the photo diode depends on the amplitudes and phase differences of the interfering fields. The phase differences can be modeled as random variables equally distributed between 0 and 2π. **Fig. 13** confirms this assumption as a result of computer simulation. Thus, $v^{(i)}$ becomes a random variable.



(a) (b)

Fig. 13. Histogram of the phase differences ($\Delta\varphi_{(i-1,i+1),x}^{(i)}(a)$, $\Delta\varphi_{(i,i+1),y}^{(i)}(b)$)

If only WDM channel i is active and the signals of the adjacent channels are zero, we get from (7):

$$v^{(i)} \propto \quad A_{(i),x}^2 + A_{(i),y}^2.$$

This signal is plotted as a function of time in Fig. 1(a). We see from this eye diagramm, that $v^{(i)}$ can take on 8 different values due to ISI for a fixed t. The values depend on the temporal bit sequence $c_{m-1}^{(i)} c_m^{(i)} c_{m+1}^{(i)}$ at the input of channel i ($t = kT_b$, T_b bit interval). These patterns 000, 001, ..., 111 are indicated in the eye diagram $v^{(i)}$ in Fig. 1(a). Throughout this paper we assume that those three consecutive bits are sufficient to approximate the eye diagram and thus the effect of ISI in channel i. An extension is straightforward. In Figs. 1(b) and (c), the adjacent channels $i-1$ and $i+1$ are active in addition with similar temporal bit patterns. Obviously, the resulting ICI causes a vast of values $v^{(i)}$ at a given time instant t. If only the interference signal from channel $i-1$ is active in channel i, the eye diagram at the output of the photo diode would look similar to Fig. 1(a). As an approximation, we also assume roughly 8 different levels depending on the temporal bit pattern. The same consideration holds, if only the interference signal from channel $i+1$ is active. Mathematically in summary, considering the 3 adjacent channels each with 8 levels at time instant t, we can define $8^3 = 512$ signal values $v^{(i)}$ within channel i. As the phase differences in (7) are random, we end up with 512 random variables. The expected value can be calculated as

$$\begin{aligned} E[v^{(i)}] \quad = \quad & A_{(i-1),x}^2 + A_{(i),x}^2 + A_{(i+1),x}^2 \\ & + A_{(i-1),y}^2 + A_{(i),y}^2 + A_{(i+1),y}^2 \end{aligned} \quad (8)$$

and the variance is

$$\begin{aligned} \sigma_{v^{(i)}} \quad = \quad & Var[v^{(i)}] \\ = \quad & 2A_{(i-1),x}^2 A_{(i),x}^2 \\ & + 2A_{(i),x}^2 A_{(i+1),x}^2 \\ & + 2A_{(i-1),x}^2 A_{(i+1),x}^2 \\ & + 2A_{(i-1),y}^2 A_{(i),y}^2 \\ & + 2A_{(i),y}^2 A_{(i+1),y}^2 \\ & + 2A_{(i-1),y}^2 A_{(i+1),y}^2. \end{aligned} \quad (9)$$

Obviously, the larger $\sigma_{v^{(i)}}$ the higher the distortion of the received signal and bit error ratio (BER) are. From (9) we see that the products of the squared amplitudes of the three channels determine $\sigma_{v^{(i)}}$. The higher these values are the higher the variance is. It is crucial to avoid the situation where both factors in each of the products in (9) are large, as this results in a high distortion. The amplitudes directly depend on the symbols to be transmitted. A '1' corresponds to a high amplitude, a '0' to a small one. Thus, the worst case scenario is, when a '1' is at each input of the two neighbouring channels at the same time. WTC can suppress such bit patterns.

APPENDIX II
OPTIMAL FREQUENCY PARAMETERS FOR WTC

To determine the optimal values of the channel spacing and the 3-dB bandwidths of the multiplexer and demultiplexer we use (10):

$$B^{(i)} = \frac{P_0(E[v_{10}^{(i)}]-E[v_{00}^{(i)}])+P_1(E[v_{10}^{(i)}]-E[v_{01}^{(i)}])}{P_{00}\sigma_{00}^{(i)}+P_{01}\sigma_{01}^{(i)}+P_{10}\sigma_{10}^{(i)}} \quad (10)$$

The lower index of the received symbols $v^{(..)}$ again reflects the underlying bit pattern. In (10) $E[v^{(i)}_{c_m^{(i)}\,c_m^{(i+1)}}]$ is the mean value, $\sigma^{(i)}_{c_m^{(i)}\,c_m^{(i+1)}}$ the standard deviation and $P_{c_m^{(i)}\,c_m^{(i+1)}}$ the a priori probability of $v^{(i)}_{c_m^{(i)}\,c_m^{(i+1)}}$. $P_{c_m^{(i+1)}}$ is defined as follows:

$$\begin{aligned} P_0 &= P_{00} + P_{10}\frac{P_{00}}{P_{00}+P_{01}} \\ P_1 &= P_{01} + P_{10}\frac{P_{01}}{P_{00}+P_{01}} \end{aligned} \quad (11)$$

Equation (10) can easiliy be expanded to take ISI into account. **Figs. 14**, **15** and **16** show $B^{(i)}$ for different channel spacing grids as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. OSNR is choosen to 16 dB to take the influence of the noise into account.The non-uniform grids perform better than the uniform one, as predicted earlier. When extending (10) in that way that the ISI is taken into account, the discrepancies get clearer (**Figs. 17**, **18** and **19**). The 3-dB bandwidth of the multiplexer must not be choosen too narrow (below 60 GHz), as the ISI introduced by the filters worsen the signal quality. This leads to the conclusion, that, when using such narrow channel spacings, it is not necessary to bandlimit the signals at the transmitter. The demultiplexer has to cut out the channels and reduce the impact of the optical noise. Bandwidths around 40..50 GHz produce the best results. Obviously the non-uniform grids reduce the sensitivity against the bandwidth of the demultiplexer. Additionally the optimal bandwidth of the demultiplexer drops, when ISI is taken into account. This is plausible as a more narrow demultiplexer reduces ICI while increasing ISI.

APPENDIX III
DISTRIBUTION OF THE POWER LOAD AT THE INPUT OF THE FIBER

Without WTC the bits at the input of the WDM channels do not depend on each other. Thus, these bit streams can be modeled as K independent random generators with output '0' or '1' (K is the number of channels). The number z of '1's being simultaneously at the channel inputs is then binomial distributed (P_1 is the a priori probability of the '1's, P_0 that of the '0's, $P_1 + P_0 = 1$):

$$P_{noWTC}[Z = z|K, P_1] = \left(\begin{array}{c} K \\ z \end{array} \right) P_1^z (1-P_1)^{K-z}.$$

When using WTC pairwise as in Fig. 4, the bits at the channel input within these pairs depend on each

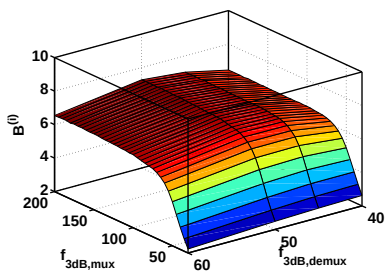


Fig. 14. $B^{(i)}$ as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 40\ GHz, \Delta f_{noWTC} = 60\ GHz$)

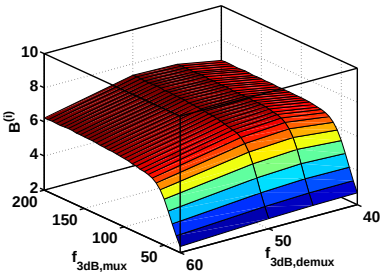


Fig. 15. $B^{(i)}$ as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 45\ GHz, \Delta f_{noWTC} = 55\ GHz$)

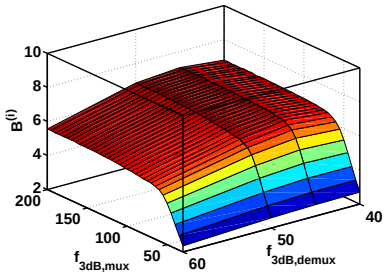


Fig. 16. $B^{(i)}$ as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 50\ GHz, \Delta f_{noWTC} = 50\ GHz$)

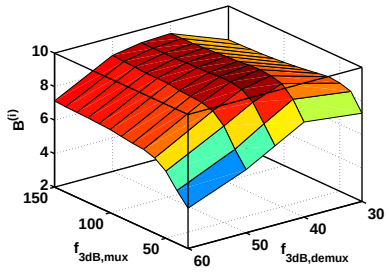


Fig. 17. $B^{(i)}$ including ISI as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 40\ GHz, \Delta f_{noWTC} = 60\ GHz$)

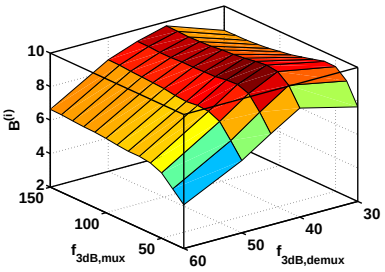


Fig. 18. $B^{(i)}$ including ISI as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 45\ GHz, \Delta f_{noWTC} = 55\ GHz$)

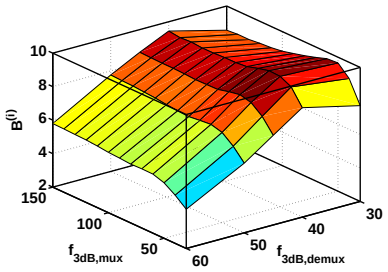


Fig. 19. $B^{(i)}$ including ISI as a function of the 3-dB bandwidths of the multiplexer and the demultiplexer, when using WTC. ($\Delta f_{WTC} = 50\ GHz, \Delta f_{noWTC} = 50\ GHz$)

other. Now the overall system can be modelled as $\frac{K}{2}$ independent random generators with output "no '1' is to be transmitted" corresponding to $(c_m^{(i)}, c_m^{(i+1)}) = (0,0)$ and "one '1' is to be transmitted" corresponding to $(c_m^{(i)}, c_m^{(i+1)}) = (0,1)$ and $(c_m^{(i)}, c_m^{(i+1)}) = (1,0)$. Now the number z of '1's simultaneously to be transmitted is again binomial distributed, but with different parameters (P_1 is now the a priori probability of the output "one '1' is to be transmitted", P_0 that of "no '1' is to be transmitted", $P_0 + P_1 = 1$):

$$P_{WTC}[Z = z|\tfrac{K}{2}, P_1] = \left(\begin{array}{c} \frac{K}{2} \\ z \end{array} \right) P_1^z (1-P_1)^{\frac{K}{2}-z}.$$

The number of '1's at the input of the fiber directly impacts the power load.

ACKNOWLEDGMENT

This work was funded by the German Ministry of Research within the joint MultiTeraNet project with Alcatel Stuttgart.

REFERENCES

[1] R. A. Griffin, A. C. Carter,*"Optical Differential Quadrature Phase-Shift Key (oDQPSK) for High Capacity Optical Transmission"*, OFC 2002, Anaheim, CA, March 2002, paper WX6.

[2] M. Ohm, J. Speidel,*"Quaternary Optical ASK-DPSK and Receivers With Direct Detection"*, IEEE Photon. Technol. Lett., vol. 15, no. 1, pp. 159-161, Jan. 2003.

[3] S. Hayase, N. Kikuchi, K. Sekine, S. Sasaki,*"Proposal of 8-state per symbol (binary ASK and QPSK) 30-Gbit/s optical modulation/demodulation scheme"*, ECOC 2003, Rimini, Italy, Sept. 2003, paper Th 2.6.4.

[4] M. Ohm,*"Optical 8-DPSK and Receiver with Direct Detection and Multilevel Electrical Signals"*, IEEE/LEOS Workshop on Advanced Modulation Formats, San Francisco, CA, July 2004, paper FC4.

[5] R. Fritsch, J. Speidel,*"OQ^{2}AM - Optical QAM scheme with orthogonal polarization"*, ITG-Fachtagung, Leipzig, Germany, May 2003, pp 169-176.

[6] B. Vasic, I. B. Djordjevic, *"Low-Density Parity Check Codes for Long-Haul Optical Communication Systems"*, IEEE photon. Technol. Lett., vol. 14, no. 8, pp. 1208-1210, Aug. 2002.

[7] M. Jäger, T. Rankl, J. Speidel, F. Buchali, H. Bülow,*"Performance of Turbo Equalizers for Optical PMD Channels"*, Journal of Lightwave Technologies, 2006, accepted for publication.