Tomlinson-Harashima Precoding for Dispersion Uncompensated PAM-4 Transmission with Direct-Detection

Roi Rath, Dennis Clausen, Simon Ohlendorf, Stephan Pachnicke, *Senior Member, IEEE,* and Werner Rosenkranz, *Senior Member, IEEE, Fellow, OSA*

***Abstract*—Combating the destructive effects of power-fading in fiber-optic systems using direct-detection is the key for increasing the transmission rates and/or reach. Some of the solutions used in the past include additional components, e.g. a dispersion-compensating fiber or an IQ-modulator, which increase the costs of the entire system. Another solution, the compensation of the channel distortions using decision-feedback equalization, is limited to its error-propagation effect. In this paper, we introduce Tomlinson-Harashima precoding for the compensation of power-fading, which allows to use the efficiency of feedback equalization for channels with spectral zeros without the disadvantage of error propagation. In an experimental investigation with different scenarios, including transmission over 25 km up to 100 km of standard single-mode fiber with gross data rates ranging from 56 to 560 Gb/s, we show that Tomlinson-Harashima precoding significantly outperforms decision-feedback equalization in all cases. Using Tomlinson-Harashima precoding, power-fading can be efficiently compensated for using digital signal processing only, keeping the system flexible and at the same time cost-effective.**

***Index Terms*— decision-feedback equalization, direct detection, digital signal processing, optical fiber communications, power-fading, pulse amplitude modulation, Tomlinson-Harashima Precoding.**

I. Introduction

Recently, there exists a growing interest in fiber-optic systems using direct-detection, due to their cost-efficiency compared to systems using coherent detection. Specifically, systems integrating four-symbol pulse-amplitude modulation (PAM-4) and direct-detection are considered as a possible solution for intra and inter data center communications [1]–[3].

Despite the aforementioned cost-efficiency, the performance of such a system is subject to penalties due to the nonlinear behavior of the photo-diode used in direct-detection. The most severe contribution stems from the combination of direct-detection and chromatic dispersion (CD), which leads to a power-fading effect, where the detected signal spectrum can contain spectral zeros, depending on the transmission distance and signal bandwidth [4]. One could overcome this problem using a dispersion-compensating fiber (DCF) at the transmitter or receiver side [2], [3]. However, a DCF increases the costs of the entire system. Moreover, due to its relatively high loss, it may require an additional optical amplification at the receiver, further increasing the costs and reducing the optical signal-to-noise ratio (OSNR) of the system. Another solution is to filter the signal at the receiver side to create a vestigial-side-band signal, as was shown in [5]. Nevertheless, tunability restrictions in low-cost lasers [6] and the use of an arrayed waveguide grating with constant center frequencies and a constant bandwidth may affect the performance of this method, unless an additional tunable optical filter is used. One could choose to eliminate dispersion using complex-valued pre-compensation at the transmitter side [7], which, however, requires the use of a more expensive IQ-modulator. Therefore, combating the effects of power-fading using digital signal processing (DSP) only is of great interest, since the system can be kept flexible and simultaneously cost-effective.

The destructive effect of the spectral zeros cannot be compensated easily using a linear equalizer [8], but could be efficiently eliminated using a decision-feedback equalizer (DFE), as will be explained later on in this paper. Unfortunately, DFE is subject to possible error propagation, and cannot be integrated straight-forward with channel-coding schemes [9]. Alternatively, it is possible to use the Tomlinson-Harashima precoding (THP) scheme, which can eliminate the effect of the spectral zeros as effectively as DFE without suffering from error-propagation, since it is applied at the transmitter side. THP was already investigated for transmission in plastic optical fibers [10], [11]. Very recently, and perhaps in parallel to this work, a first trail on THP for an intensity-modulation/direct-detection system was conducted, allowing the transmission of (gross) 30 GBd PAM-4 over 30 km of standard single-mode fiber (SSMF) [12]. However,

Manuscript received April 05, 2017; revised May 15, 2017 and June 23, 2017; accepted June 30, 2017.

R. Rath, D. Clausen, S. Ohlendorf, S. Pachnicke and W. Rosenkranz are with the Chair of Communications, Christian-Albrechts-Universität zu Kiel, Kaiserstr. 2, 24143 Kiel, Germany (e-mail: roi.rath@tf.uni-kiel.de; dennis.clausen@tf.uni-kiel.de; simon.ohlendorf@tf.uni-kiel.de; stephan.pachnicke@tf.uni-kiel.de; wr@tf.uni-kiel.de).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

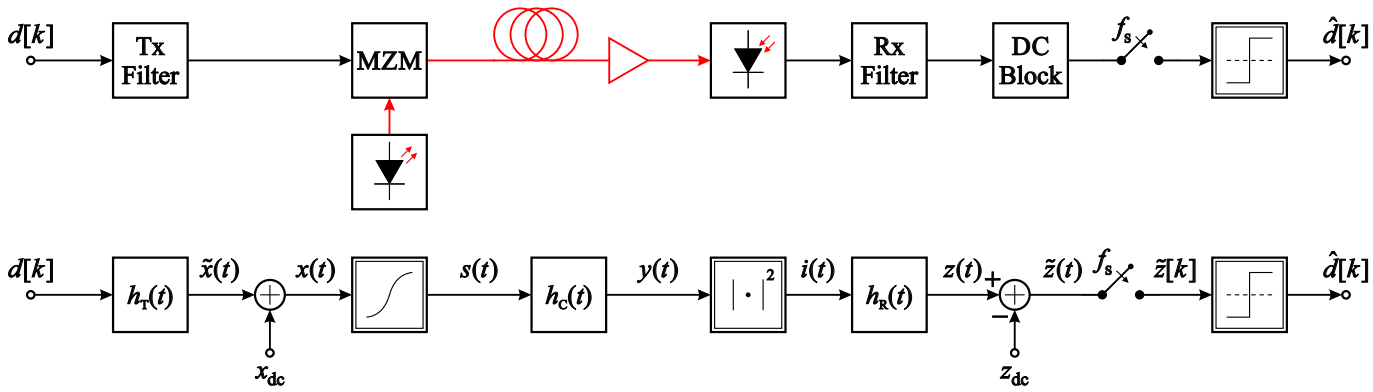

Fig. 1. A general block diagram of the system under investigation (upper subfigure), and its mathematical description in the equivalent baseband (lower subfigure). Here, the transmit and receive filters are described by the impulse responses h_T and h_R, respectively, the laser and the MZM are given as a dc addition followed by a cosine characteristic function, the optical fiber and amplifier are denoted by the impulse response h_C, the photo-diode by an absolute square law, and the dc block by dc subtraction. Note that discrete sequences are denoted with square brackets and the discrete time index k, dc-free signals are marked with a tilde, and dc signals with a "dc" subscript. In addition, f_s denotes the sampling frequency.

and to the best of the authors' knowledge, our contribution shows for the first time the usage of THP to achieve gross data rates of up to 560 Gb/s over a reach up to 100 km of SSMF.

The rest of the paper is organized as follows: In Section II we summarize and illustrate the problematic effect of power-fading in the system. The idea of the optical signal-to-carrier power ratio (OSCR) as an important optimization parameter is discussed in Section III, followed by an elaboration of the expected performance of equalizers using a feed-forward or feedback structure in Section IV. The experimental setup and DSP scheme for the investigation is described in Section V, followed by an analysis and discussion of the results in sections VI and VII, respectively. Finally, the paper is concluded in Section VIII.

II. Direct-Detection, Chromatic-Dispersion and Power-Fading

A block diagram describing a fiber-optic system with direct-detection is shown in Fig. 1, where the lower subfigure is the system's mathematical description in the equivalent baseband. Clearly, due to the magnitude-square operation of the photo-diode [13], a real-valued unipolar modulation (here PAM-4) is required. Switching from bipolar to unipolar modulation can be easily achieved by adding a dc value to the transmitted signal. The complex envelope of the transmitted signal can be written as

$$s(t) = \tilde{s}(t) + s_{\mathrm{dc}}\,, \tag{1}$$

where $\tilde{s}(t)$ is the dc-free information signal, corresponding to the original, bipolar sequence, and s_{dc} is the dc value added to the signal. Assuming no channel distortions, the signal after the photo-diode is

$$i(t) = s_{\mathrm{dc}}^2 + \tilde{s}^2(t) + 2s_{\mathrm{dc}}\tilde{s}(t)\,. \tag{2}$$

We consider transmission over an SSMF with a low enough launch-power, such that the fiber can be considered as a linear system (considering 1st order dispersion only and neglecting attenuation). In addition, the Mach-Zehnder modulator (MZM) nonlinearities are neglected for the sake of simplicity. Considering now direct-detection, and a *real* signal $s(t)$ with a structure as in (1), it can be shown that the photo-current is

$$i(t) = y_{\mathrm{dc}}^2 + \left|\tilde{s}(t) * h_{\mathrm{c}}(t)\right|^2 + 2y_{\mathrm{dc}}\tilde{s}(t) * \mathcal{F}^{-1}\left\{\mathrm{Re}\left\{H_{\mathrm{c}}(f)\right\}\right\}, \tag{3}$$

where $h_{\mathrm{c}}(t)$ is the inverse Fourier transform of $H_{\mathrm{c}}(f)$ and

$$\mathrm{Re}\left\{H_{\mathrm{c}}(f)\right\} = \cos\left(2\pi^2\beta_2 L f^2\right), \tag{4}$$

where β_2 is the group velocity parameter in s^2/m, and L is the fiber length. The photo-detected signal is the sum of three components. The first two are the dc component, which neither contains information, nor distorts the signal, and a nonlinear interference, also referred to in the literature as *inter-modulation distortion* [4] or *signal-signal beat interference* (SSBI) [14], [15]. The third component contains the information signal distorted by a *linear* system: a cosine function of a quadratic frequency. This is the power-fading effect, which may lead to spectral zeros in the spectrum of the signal. The locations of the spectral zeros on the frequency axis can be calculated by setting the cosine function to zero. The first five spectral zero frequencies for three different fiber lengths are summarized in Table I.

Neglecting the SSBI, the system is *linear*, and can be therefore represented by means of its frequency response or its impulse response. In order to better understand the system properties, we plotted the magnitude response, impulse response and pole-zero map of a 28 GBd system for different fiber lengths, namely 15, 50 and 100 km, as shown in Fig. 2. For that purpose, we further assumed that the impulse response of the transmit and receive filter is a sinc function, ideal sampling at the symbol rate, and the removal of the dc component using a dc block as shown previously in Fig. 1. From the magnitude responses (Fig. 2, left), we can clearly see how the number of spectral zeros changes from none to one and

TABLE I
First Five Spectral Zero Frequencies in GHz vs. Fiber Length

L [km]	1	2	3	4	5
15	15.642	27.093	34.976	41.385	46.926
50	8.567	14.839	19.157	22.667	25.702
100	6.058	10.493	13.456	16.028	18.174

Note: The gray values denote spectral zeros that are outside of the 28 GHz signal bandwidth in case of a sinc transmit and receive-filter.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

three by increasing the transmission distance. In Table I, we have marked in addition the out-of-band spectral zero frequencies for this system in grey. Notice that we have truncated the impulse response to 25 coefficients to allow an easier comparison of the pole-zero maps, where this truncation also fulfills a ratio of at least 30 dB from the maximum coefficient and the outmost coefficient. An observation of the impulse responses shows a symmetric combination of pre- and post-cursor inter-symbol interference (ISI) components. This affects the choice of the equalizer structure, as will be explained later on in Section IV. The pole-zero maps show clearly the zeros on the unit circle, corresponding to the spectral zeros observed in the frequency responses.

III. OSCR as a Crucial Optimization Parameter

In the last section, we have neglected the SSBI. In reality, this interference can be reduced by minimizing the OSCR. This action will have a direct effect on the effective OSNR of the system, as will be explained in this section. There were attempts to mitigate the SSBI using its iterative estimation and cancellation [14], [15]. This method, however, increases the complexity of the DSP unit as well as the system delay, and is not considered in this contribution.

The OSCR is defined as the ratio of the power of the information-carrying signal P_s (the effective power) and carrier power P_c at the output of the MZM:

$$\mathrm{OSCR} = \frac{P_s}{P_c} \,. \tag{5}$$

It is possible to show [16] that the (effective) signal to nonlinear interference power ratio ($\mathrm{SIR_{NL}}$) at the receiver is

$$\mathrm{SIR_{NL}} \propto \frac{1}{\mathrm{OSCR}} \,, \tag{6}$$

i.e. by reducing the OSCR we can mitigate the effects of the unwanted nonlinear contribution.

Let us observe now the effect of this action on the effective OSNR, i.e. the OSNR with respect to the information-carrying signal. The optical power at the output of the MZM is

$$P_L = P_s + P_c \,. \tag{7}$$

Using the definition of the OSCR, the effective power can be expressed as:

$$P_s = P_L \frac{\mathrm{OSCR}}{1+\mathrm{OSCR}} \,, \tag{8}$$

Assuming an additive white Gaussian noise with a power P_n right after the MZM (e.g. using noise loading in a back-to-back scenario), the effective OSNR can be determined:

$$\mathrm{OSNR_{eff}} = \frac{P_s}{P_n} = \frac{P_L}{P_n}\frac{\mathrm{OSCR}}{1+\mathrm{OSCR}} = \mathrm{OSNR}\frac{\mathrm{OSCR}}{1+\mathrm{OSCR}} \,. \tag{9}$$

Notice that the OSNR in (9) assumes a reference bandwidth equal to the signal bandwidth. It differs from the conventionally used OSNR, which is measured with a 0.1 nm reference bandwidth, by a constant factor that depends on the signal bandwidth. The conventional OSNR is denoted in this paper as $\mathrm{OSNR_{0.1}}$. For low OSCR values, the relation in (9) can be approximated as follows:

$$\mathrm{OSNR_{eff}}\big|_{\mathrm{OSCR}\ll 1} \approx \mathrm{OSNR}\cdot\mathrm{OSCR} \,. \tag{10}$$

Observing (6) and (10), there exists a trade-off between the amount of nonlinear mitigation and noise in the system. As an

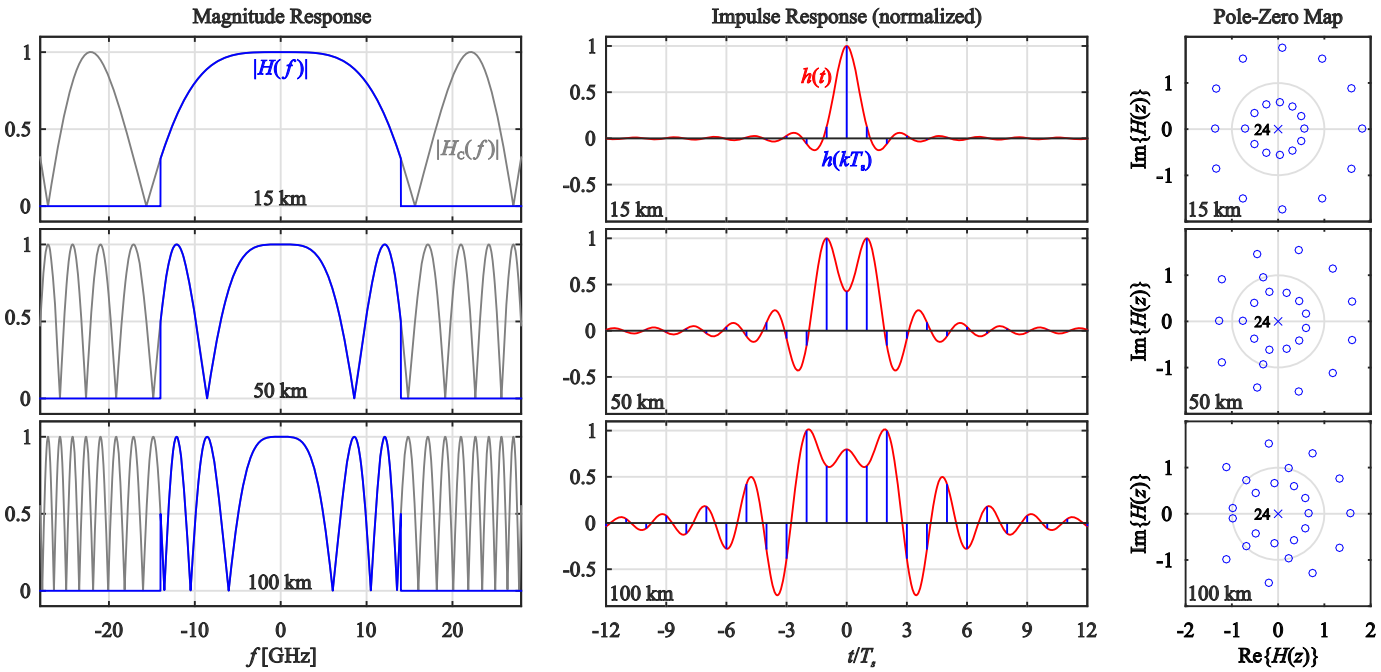

Fig. 2. Illustration of system properties for accumulated dispersion and direct detection. Left column: Magnitude responses showing the power-fading effect. Here, $H(f)$ is the total magnitude response including the transmit and receive filters, and $H_C(f)$ is the magnitude response of the fiber. Middle column: Impulse responses showing a symmetric combination of pre- and post-cursor ISI components. Right column: Pole-zero maps, showing (in case of 50 km and 100 km) the spectral zeros resulting from zeros lying directly on the unit-circle.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

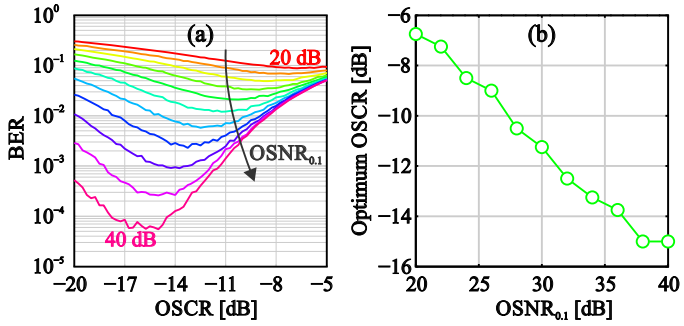

Fig. 3. OSCR optimization. (a) BER vs. OSCR for different $OSNR_{0.1}$ values from 20 dB to 40 dB. (b) Optimum OSCR. i.e. the minima of the curves in (a) vs. $OSNR_{0.1}$

outcome, the OSCR has to be optimized in order to maximize the system performance. This optimization parameter can be adjusted by varying the MZM's bias voltage v_b and/or the peak-to-peak voltage of the driving signal v_{pp} [17]. An example for such an optimization is given in Fig. 3, in case of transmission of 28 GBd over 25 km of SSMF, a zero roll-off and a compensation with an error-propagation free DFE (this hypothetical variant of the DFE will be clarified in the next section). In the left subfigure the bit-error ratio (BER) is plotted vs. the OSCR; altogether 11 curves for different OSNR values from 20 dB to 40 dB. By locating the minimum point (and hence the optimum) in each curve, one can differentiate between the regime dominated by noise (to the left of the minimum) and the regime dominated by SSBI (to the right of the minimum). The OSCR values that minimize the BER for each OSNR value were mapped to the graph in the right subfigure. As expected, for low $OSNR_{0.1}$ the OSCR has to be kept high to allow a high $OSNR_{eff}$. As the $OSNR_{0.1}$ increases, the OSCR can be further reduced to increase the SIR_{NL}, hence to allow a better compensation of linear ISI with the equalizer.

IV. FFE, DFE AND THP

In order to equalize ISI, we consider the use of either a *moving-average* (MA) filter or a *transversal* filter, an autoregressive (AR) filter, or a combination of both, i.e. an autoregressive moving-average (ARMA) filter [18]. The purpose of the equalizer is to alter the system's transfer function in the Z-plane such that the total transfer function is (ideally) constant on the unit-circle, i.e. an all-pass filter. An MA filter is an all-zero filter, i.e. it can only add zeros to the channel's transfer function, where an AR filter is an all-pole filter and therefore can only add poles. In many cases, an MA filter can equalize efficiently for ISI, at the expense of additional noise enhancement. However, what happens if the channel's transfer function contains zeros very close to or, in worst case, directly on the unit circle (as can be clearly seen in Fig. 2)? In this case, a mere insertion of zeros using an MA filter can neither cancel the spectral zeros, nor efficiently "pull" the system's magnitude response in the locations affected by a near zero. An AR filter, however, can successfully do so using pole-insertion. In the rest of the paper we consider a feed-forward equalizer (FFE), which corresponds to a linear MA filter, and a DFE or THP, which correspond both to an AR filter. Since an AR filter has a

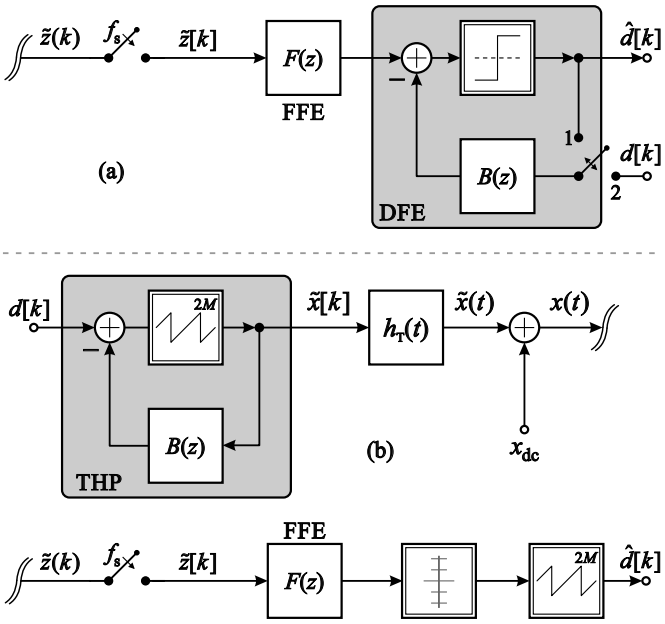

Fig. 4. Investigated equalizer structures, as a possible addition to the system illustrated in Fig. 1. (a) an FFE-DFE. The feedback structure can have two states: In state 1 it is a conventional DFE with error propagation, and in state 2 the transmitted symbols are fed to the feedback filter, i.e. the DFE operates error-propagation-free. (b) THP-FFE (separated between transmitter and receiver). In order to add the FFE-DFE to the general system description shown in Fig. 1, one should place it after the sampler. In case of precoding, THP should be placed between the symbol generator and the Tx filter, whereas its complementary FFE should be placed after the sampler.

feedback structure, it may become unstable, if its poles are on or outside the unit-circle. In order to allow stability in a bounded-input bounded-output sense, a nonlinear element is inserted in the upper arm of the AR filter. In case of DFE, it is the decision-device. In case of THP, a modulo operation is introduced, confining its output between $-M$ and $+M$, where M is the cardinality of the modulation format (in case of PAM-4, M=4) [9]. Because of their structure, DFE and THP can equalize spectral zeros more efficiently than an FFE. Nevertheless, since DFE (or THP) has poorer noise performance in the case the channel impulse response contains pre-cursor components [19] (which is also the case in this work, as can be seen in Fig. 2), the best choice for equalization is a combination of a feed-forward and a feedback structure equalizer. The structure of the different equalizers under investigation is illustrated in Fig. 4. First, we consider the DFE structure in Fig. 4a. A real DFE feeds back the symbol decisions (state 1 in the figure). Feeding wrong symbols due to false decisions may lead to error-propagation. Assuming that only correct decisions are fed back in the AR filter $B(z)$ of the DFE, no error-propagation will occur (state 2 in the figure). This so-called error-propagation-free DFE will be used later to determine the lower bound for the system's BER

The structure of THP is given in Fig. 4b. Observing the figure, we see that THP requires not only modifications at the transmitter side, but at the receiver side as well, as will be explained later on. Although being closely related to the DFE, THP possesses two different signal properties. The precoded sequence, i.e. the sequence after THP at the transmitter side, is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

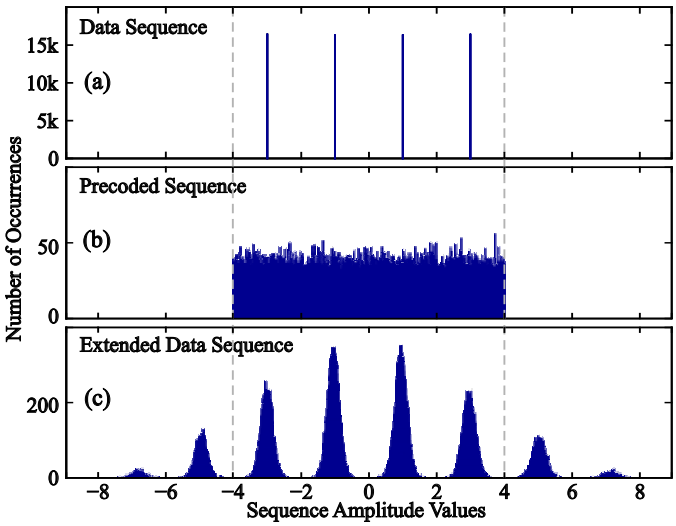

Fig. 5. Sequence histograms showing the distribution of difference sequences along a system using THP. (a) PAM-4 pseudo-random data sequence of length 2^{16} (a combination of two pseudo-random binary sequences, such that all four symbols are equally distributed). (b) Precoded sequence. (c) Extended data sequence (containing noise and residual ISI). All histograms are distributed over 2000 bins. The minimum and maximum of the modulo $2M$ function (–4,+4 in case of PAM-4) are marked with grey dashed lines.

almost independent and identically distributed with a uniform distribution between $-M$ and $+M$. These values are the minimum and maximum of the modulo $2M$ function used in the upper arm of the THP. The approximation of the sequence as uniformly distributed becomes more exact if the number of signal levels M increases [9]. As a result, the electrical power of the precoded sequence is higher than the electrical power of the data sequence by a factor of $M^2/(M^2-1)$, corresponding to 0.28 dB in case of PAM-4. This power penalty is referred to in the literature as "precoding loss" [9]. Assuming a linear system, the precoding loss is the same in both the electrical and the optical domain. However, since the MZM has a nonlinear cosine characteristic, the precoding loss will vary, depending on the driving voltage: Decreasing the driving voltage will decrease the MZM nonlinearities, and will make the actual precoding loss to converge to the theoretical one. The sequence after the FFE at the receiver side, i.e. after the complete equalization, is not yet the data sequence, but a so-called "*extended* data sequence" [9], drawn from an *extended* data set, which has more levels than the data sequence. The size of the extended data set is affected by the power of the channel tap weights and the size of the symbol set M. This sequence is first detected using a slicer designed to detect the extended data sequence and then modulo-decoded back to the original data sequence using an additional modulo operator. In Fig. 5, we illustrate the distributions of the data sequence, the precoded sequence, and the extended data sequence using histograms. As can be seen, the extended data sequence has the size of a PAM-8, where the two outermost symbols from each side should be decoded using the modulo $2M$ operation. For a profounder understanding of THP, the reader is referred to [9], [20], [21].

V. System Setup

A schematic of the experimental setup is given in Fig. 6. Starting from the transmitter DSP scheme, a two-dimensional pseudo random de Bruijn sequence of length 2^{15} was mapped to PAM-4 symbols. In case of precoding, the data sequence was processed with THP to generate the precoded data sequence (gray sub-block). The sequence was then resampled and shaped with a root raised-cosine (RACOS) filter with a roll-off factor of 0.3, followed by a linear pre-compensation filter to equalize the transmitter distortions originated in the digital-to-analog converter (DAC) and electrical amplifier. The resulting sequence was converted to an analog signal using the Keysight M8196A arbitrary waveform generator (AWG), operating at 84 or 90 Gsamples/s with a vertical resolution of 8 bits (ENOB of about 5.5). A tunable laser (100 kHz linewidth) operating at a wavelength λ_c =1549.411 nm and an MZM with a bandwidth of 31.7 GHz and a half-wave switching voltage of V_π = 4.6 V were used for the optical up-conversion. The analog electrical signal was amplified by 14 dB to allow reaching the maximum peak-to-peak voltage of V_π, i.e. the MZM's full driving range. The bias voltage v_b of the MZM was varied between $0.05V_\pi$ and $0.5V_\pi$, whereas the MZM was driven with a peak-to-peak voltage v_{pp} ranging from $0.1v_b$ to $2v_b$. The optical band-pass signal was boosted with an Erbium-doped fiber amplifier (EDFA) to 10 dBm to allow a constant OSNR regardless of the choice of bias and/or driving voltage. In case of WDM transmission, four other neighboring channels were generated using four lasers with 100 kHz linewidth and a second MZM with a bandwidth of 25.7 GHz and V_π = 1.5 V as described

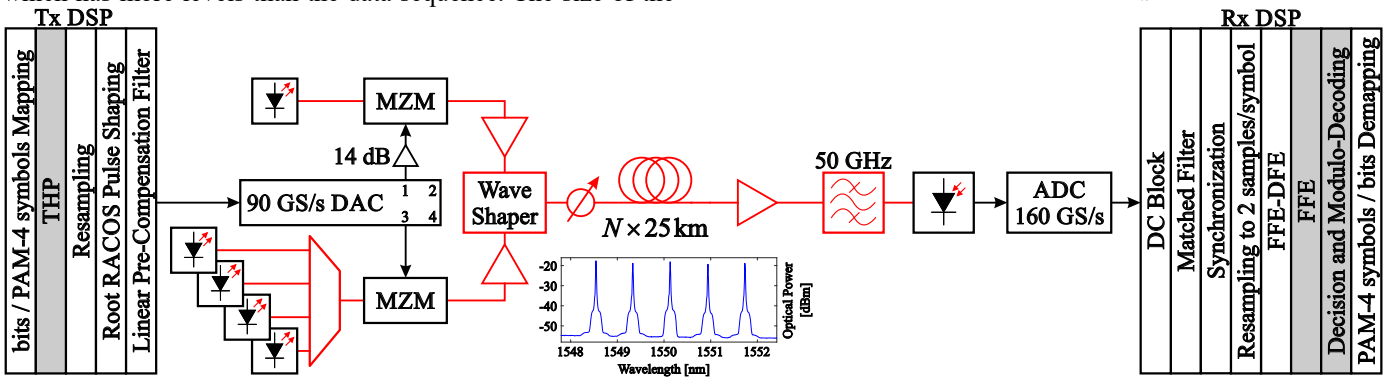

Fig. 6. Experimental setup. The black blocks and connections represent *electrical* devices and connections, whereas the red are *optical*. The numbers inside the DAC block represent the output channels of the real device. The variable N is an integer number and can get the values 1 to 4. In addition, the grey blocks in the DSP schemes are executed only if THP is applied. The measured spectrum in case of 5×28 GBd transmission is given as an inset.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

schematically in Fig. 6. All five channels, spaced 100 GHz apart, were then combined using a FINISAR® wave shaper, such that the investigated channel is the middle channel of the five. We have made all the necessary adjustments such that all channels have the same optical power as well as the same bias voltage and driving voltage relative to V_π. A variable optical attenuator (VOA) was used to control the launch-power into the fiber, simultaneously varying the OSNR. The motivation was to keep the launch-power as low as possible for a given OSNR, and thereby to reduce the fiber nonlinearities as much as possible. The signal was then transmitted through an SSMF with a length of 25.2, 50.4, 75.6 or 100.8 km. At the receiver side, the signal (or middle channel in case of WDM) was filtered with an optical filter (a second order Gaussian filter) with a 3 dB bandwidth of 50 GHz, and was then detected with a 50 GHz photo-detector. The power into the photo-diode was kept as high as possible (between 11.5 and 5.5 dBm, depending on the OSNR), since no transimpedance amplifier was used after photo-detection. After detection, the signal was sampled with the Keysight Infiniium DSAZ634A real-time oscilloscope operating at a rate of 160 Gsamples/s with a vertical resolution of 8 bits (ENOB of about 5.5).

The sampled signal was processed off-line according to the Rx DSP schematic given in Fig. 6. After the removal of the dc, the signal was filtered with a root RACOS filter (i.e. a matched filter). Synchronization was applied using cross-correlation with a training sequence to find the beginning of the received sequence, followed by a resampling of the sequence to two samples/symbol. In case of precoding, the sequence was equalized with an FFE for the elimination of pre-cursor ISI. The output of the FFE, i.e. the extended data sequence, was detected using an extended decision-device and modulo-decoded back to a PAM-4 signal (gray sub-blocks). Otherwise, the signal was equalized and detected with an FFE-DFE. The equalizer coefficients were estimated using the minimum mean square error criterion. The number of equalizer coefficients was set to 30 for the feed-forward and 30 for the feedback structure. We have chosen this number in order to avoid the optimization of the number of equalizer coefficients over different Baud rates and transmission lengths. This number of coefficients was found to be too high in many cases and could be decreased without degradation in performance. For example, in case of 28 GBd over 50 km, 16 feed-forward and 10 feedback coefficients are sufficient. The decision thresholds of the decision devices were optimized to avoid suboptimal performance. The decided PAM-4 symbols were de-mapped into bits followed by error counting and BER determination with a minimum of 100 errors for a BER of 10^{-5}.

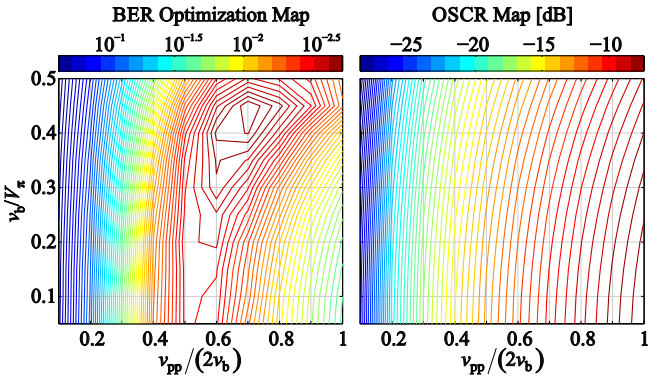

Fig. 7. BER optimization map (left subfigure) and the corresponding OSCR map in dB (right subfigure).

VI. Experimental Results

Prior to the experimental measurements, we conducted a thorough simulative investigation to obtain the lower bound for the system performance, and more importantly, the optimization maps to be used later-on in the lab for setting the MZM's bias voltage v_b and the driving peak-to-peak voltage v_{pp}. An example of such an optimization map is shown in Fig. 7, where the BER values for each (v_b , v_{pp}) pair are illustrated using a colored contour map in case of a 28 GBd transmission over 25 km of SSMF and equalization with DFE at an OSNR of 32 dB. We have included a contour map of the corresponding OSCR for each (v_b , v_{pp}) pair in Fig. 7 as well. One can clearly see the similar direction of the contour lines of both maps in the v_b axis, indicating the relation between the OSCR and the system performance. Moving from lower to higher v_{pp} values,

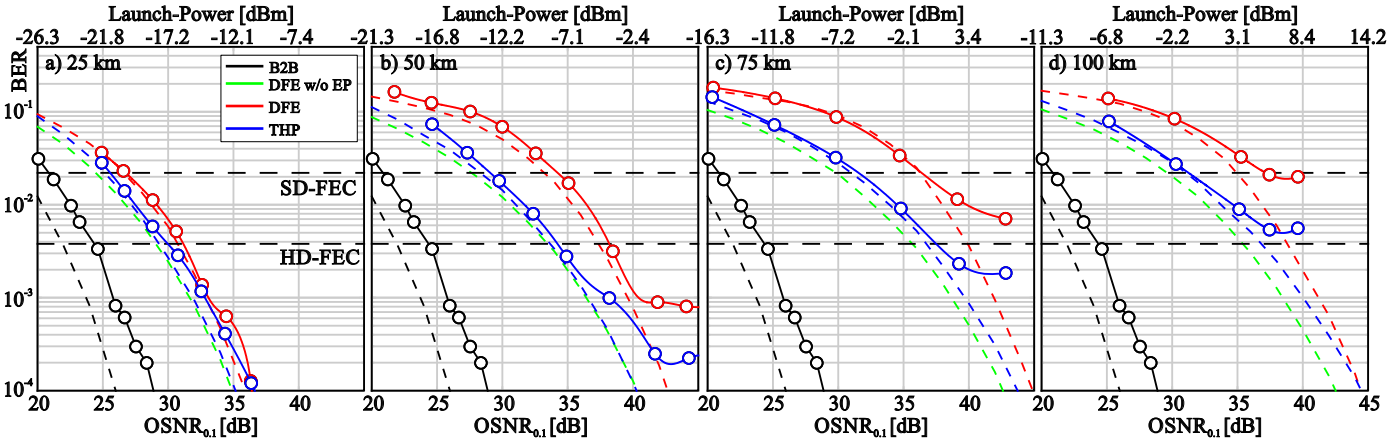

Fig. 8. BER curves for a of 28 GBd transmission over (a) 25 km (b) 50 km (c) 75 km and (d) 100 km. The simulation and experimental results are illustrated with dashed and solid lines, respectively, where the markers are the measured BER values. The green curve (DFE without error propagation) is the lower bound for the equalizers' performance. The back-to-back results are shown in black.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

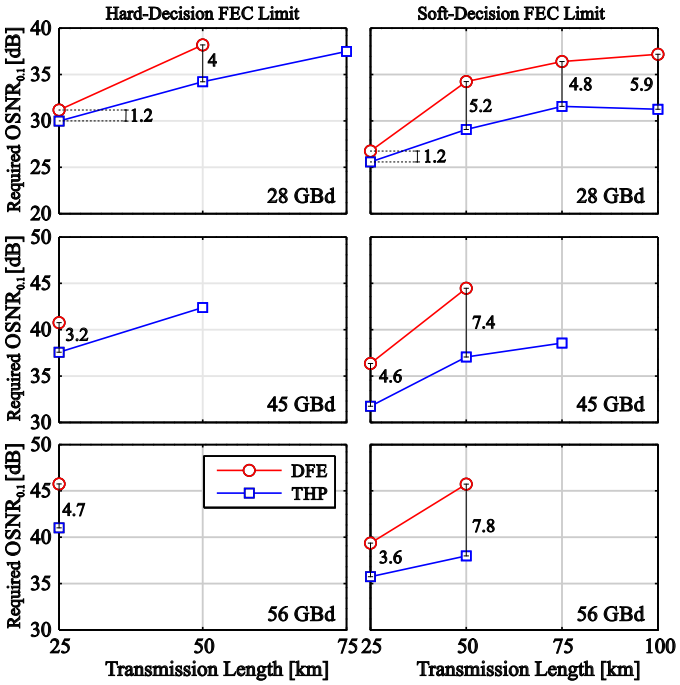

Fig. 9. Required $OSNR_{0.1}$ vs. transmission length for hard and soft-decision FEC limits. THP's OSNR gain over DFE is marked and given in dB for all possible scenarios.

the OSCR increases monotonically, whereas the BER decreases till it reaches its minimum and then increases again, as was already seen in Fig. 3. For this scenario, we find a unique minimum BER for the (normalized) pair (0.45 , 0.7).

Fig. 8 shows the BER vs. $OSNR_{0.1}$ curves of FFE-DFE and THP-FFE in case of 28 GBd for each of the investigated transmission distances. The back-to-back (B2B) results are given as a reference. Notice that the launch-power values corresponding to the OSNR values are given as well. The simulation results are plotted as dashed lines, whereas the measured values are given as round markers, and the solid lines are the spline interpolation of the results. The green curves correspond to a DFE without error-propagation (see Section IV), and are given as a lower bound for the achievable BER. Notice that the simulation model included MZM nonlinearities, but no fiber nonlinearities.

Observing first the simulation results for BER values lower than 10^{-3}, THP and DFE without error-propagation show similar performance, as expected, for 25 and 50 km. For 75 and 100 km, however, THP shows worse performance than the error-propagation-free DFE by 1 to 2 dB. This could be explained as follows: As the transmission distance increases, the power of the ISI components becomes higher. This results in a larger extended data sequence constellation [9]. The larger the constellation, the more severely it is affected by the signal-spontaneous beat noise, which is the dominant noise contribution in this scenario. Compared with THP or error-propagation-free DFE, the performance degradation of the (real) DFE is clearly visible in all scenarios. This is the result of the error-propagation effect. We observe in general a performance degradation when increasing the transmission distance. As the number of spectral zeros increases, all equalizers need the system to have less SSBI in order to better compensate for the linear ISI. This is achieved, as already explained in Section III, by reducing the OSCR, which in turn reduces the effective OSNR. This reduction is the reason for this degradation.

Examining the experimental results, we can recognize an excellent agreement between the simulation and the measured results, supporting the correct use of the optimization maps. In case of 25 km, THP shows no superiority over DFE at BER values lower than 10^{-3}. However, around both the hard-decision (HD) and soft-decision (SD) forward error correction (FEC) limits (which are $3.8 \cdot 10^{-3}$ and $2.2 \cdot 10^{-2}$, respectively [22]) a significant OSNR gain of 1.2 dB can be observed. In case of 50 km, THP clearly outperforms the DFE structure, registering an OSNR gain of 3.97 and 5.15 dB at the HD and SD-FEC limits, respectively. The results for 75 km and 100 km show similar OSNR gain values for the THP. However, compared with the 50 km scenario, the fiber nonlinearities are more dominant,

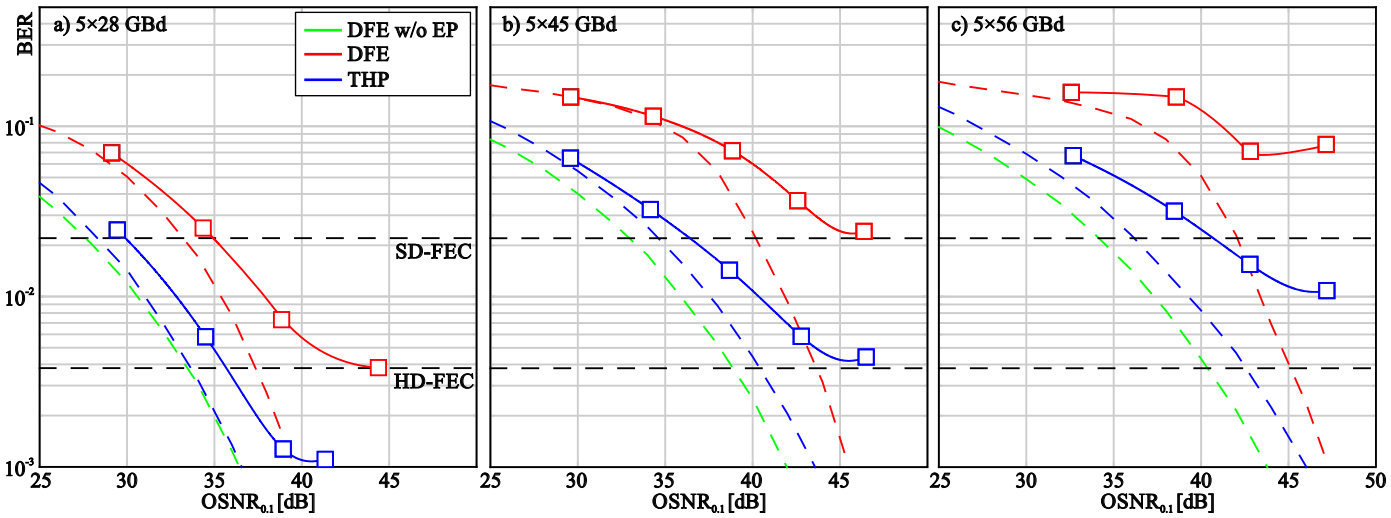

Fig. 10. Experimental results for a five-channel WDM transmission over 50 km in case of (a) 28 GBd (b) 45 GBd and (c) 56 GBd per channel. The simulation results for single-carrier transmission are illustrated with dashed curves, and the experimental results with solid curves. The markers are the measured BER values. The green curve (DFE without error propagation) is the lower bound for the equalizers' performance.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2724032, Journal of Lightwave Technology

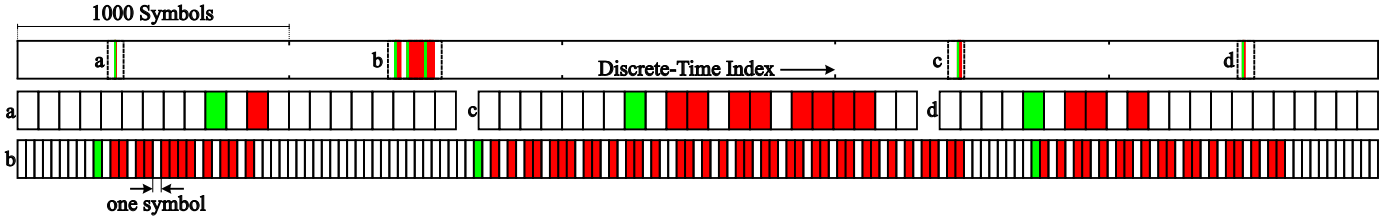

Fig. 11. Symbol error distribution vs. the sequence discrete-time index. The upper subfigure shows the error distribution over a sequence window of 5000 symbols. The four different discrete-time windows labeled from 'a' to 'd', which contain the symbol errors, are enlarged and illustrated separately, and are marked with the corresponding window letters. The correct symbols are illustrated as white bars, the symbol errors of an error-propagation-free DFE as green bars, and the additional errors due to error-propagation as red bars.

causing the measured curve to diverge from the simulation curve, where in the case of 100 km only the SD-FEC limit can be reached. The rest of the results are summarized in Fig. 9. We plotted the required $OSNR_{0.1}$ to achieve the HD or SD-FEC limit vs. the transmission length for each equalizer structure for all three investigated Baud rates. THP outperforms DFE with $OSNR_{0.1}$ gain ranging from 1.2 up to 7.8 dB.

We repeated the measurements for a five-channel WDM transmission over 50 km SSMF. Here, we aim to examine the performance degradation due to the Kerr effect as we increase the number of channels from one to five, and by that increasing the data rate by five as well. The results given in Fig. 10 show, as expected, similar perfomance curves as for single-carrier transmission at low $OSNR_{0.1}$ values, whereas the performance degrades earlier with the increasing of the $OSNR_{0.1}$ due to the stronger fiber nonlinearities. THP manages to go at least under the SD-FEC limit for all cases, whereas DFE fails to do so in case of 45 GBd and 56 GBd.

VII. The Error-Propagation Effect

As mentioned in the previous sections, THP outperforms DFE due to the error-propagation effect. One can quantify the error-propagation as follows: For a given OSNR, the average length of the error burst caused by a single symbol error due to error propagation is defined as

$$\varepsilon_{\mathrm{EP}}(\mathrm{OSNR}) = \frac{\mathrm{SER}_{\mathrm{EP}}(\mathrm{OSNR})}{\mathrm{SER}_{\mathrm{EPF}}(\mathrm{OSNR})} - 1, \qquad (11)$$

where SER_{EP} and SER_{EPF} is the symbol error ratio of a DFE and an error-propagation-free DFE, respectively. Assuming that a single symbol-error leads to a single bit error, one can use the BER values to obtain the same results. For example, in the scenario of 28 GBd over 50 km at an $OSNR_{0.1}$ of 40 dB, the error propagation factor defined in (11) reads

$$\varepsilon_{\mathrm{EP}}(40\ \mathrm{dB}) = \left(8\cdot 10^{-3}\right) \Big/ \left(1\cdot 10^{-3}\right) - 1 = 7\,. \qquad (12)$$

The actual error burst length depends on the state of the DFE filter $B(z)$, i.e. on the combination of symbols inside the DFE's shift register at the time the error has occurred. Some symbol combinations could lead to an extreme destructive interference, causing a long error burst, whereas some result in a constructive interference, causing a rather short error burst. Some examples for possible error burst lengths are given in Fig. 11, showing the symbol error distribution over a discrete-time span of 5000 symbols (upper subfigure). Four areas of interest, where symbol errors have occurred, are marked from 'a' to 'd', and are enlarged and illustrated separately, labeled with the corresponding letter. In the enlarged areas of interest, each bar corresponds to a single symbol. The correct symbol decisions marked as white bars, the green bars are symbol errors for an error-propagation-free DFE, and the red bars are the additional errors due to error propagation. As can be seen from subfigures 'a' to 'd', only six symbol errors (green bars) initiated by the channel. Each symbol error, however, results in a different error burst length varying from a single error (Fig. 11a) up to 35 errors (Fig. 11b). Notice that the error-propagation is affected by many system parameters, such as the length of the channel impulse response, the amplitude and sign of the different DFE coefficients, and the size of the modulation format.

VIII. Conclusion

In this contribution, we experimentally demonstrated the use of THP as a flexible and cost-effective solution for equalization of power-fading caused by the combination of chromatic dispersion and direct detection. THP was compared with DFE for transmission of gross data rates from 56 to 112 Gb/s over up to 100 km of SSMF for a single-carrier transmission. THP's error-propagation-free equalization was shown to outperform DFE, showing an OSNR gain ranging from 1.2 dB to 7.8 dB, depending on the target BER. The same behavior was observed in a five-channel WDM transmission trial over 50 km, yielding a gross data rate of 280 to 560 Gb/s. The great advantage of THP over DFE was confirmed again in the scenarios with high launch-power: THP managed to successfully transmit 5×45 GBd and 5×56 GBd over 50 km, whereas DFE couldn't reach even the SD-FEC limit, due to the fiber's Kerr-effect nonlinearities. To the best of the authors' knowledge, this is the first time THP was successfully used to achieve these data rates and reach for an intensity-modulation/direct-detection system. Furthermore, the importance of OSCR optimization on the performance of the entire system was explained mathematically and proven experimentally. In addition, the error-propagation effect, the main reason for THP's advantage over DFE, was discussed and characterized, and was illustrated for a given scenario.

Acknowledgment

R. Rath thanks J. Meusel for fruitful discussions. The authors thank Keysight for the loan of the AWG and digital oscilloscope.

REFERENCES

[1] A. Dochhan, N. Eiselt, H. Griesser, M. Eiselt, J. J. V. Olmos, I. T. Monroy, and J.-P. Elbers "Solutions for 400 Gbit/s inter data center WDM transmission," in *Opt. Commun. (ECOC), 2016 42nd Eur. Conf. and Exhib. Proc.*, pp. 680–682.

[2] N. Eiselt, J. Wei, H. Griesser, A. Dochhan, M. H. Eiselt, J.-P. Elbers, J. J. V. Olmos, and I. T. Monroy, "Evaluation of real-time 8×56.25 Gb/s (400G) PAM-4 for inter-data center application over 80 km of SSMF at 1550 nm," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 955–962, Feb. 2017.

[3] N. Eiselt, A. Dochhan, H. Griesser, M. Eiselt, J. J. V. Olmos, and I. T. Monroy, "Experimental comparison of 56 Gbit/ s PAM-4 and DMT for data center interconnect applications," in *Photon. Netw., 17th ITG Symp. Proc.*, 2016, pp. 134–138.

[4] D. J. F. Barros and J. M. Kahn, "Comparison of orthogonal frequency-division multiplexing and on-off keying in amplified direct-detection single-mode fiber systems," *J. Lightw. Technol.*, vol. 28, no. 12, pp. 1811–1820, Jun. 2010.

[5] J. Lee, N. Kaneda, Y.-K. Chen, "112-Gbit/s intensity-modulated direct-detect vestigial-sideband PAM4 transmission over an 80-km SSMF link," in *Proc. Opt. Commun. (ECOC), 2016 42nd Eur. Conf. and Exhib.*, pp. 136–138.

[6] A. Dochhan, H. Griesser, N. Eiselt, M. H. Eiselt, and J.-P. Elbers, "Solutions for 80 km DWDM systems," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 491–499, Jan. 2016.

[7] K. Petermann, C. Weber, and S. Warm, "Perspectives for direct detection optical fiber transmission systems with electronic pre-compensation," in *Proc. 15th OptoElectron. and Commun. Conf. (OECC)*, 2010, pp. 142-143, paper 7B3-1.

[8] N. Eiselt, H. Griesser, A. Dochhan, M. Eiselt, J. J. V. Olmos, and I. T. Monroy, "Experimental investigation and comparison of different equalizers for four level pulse amplitude modulation," in *Photon. Netw., 16th ITG Symp. Proc.*, 2015, pp. 140–144.

[9] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2002.

[10] S. A. Cheema, M. Wolf, Ö. Tolay, and M. Haardt, "Efficient techniques for multi-Gigabits/s transmission over plastic optical fiber," presented at the 17th Int. Conf. Transparent Opt. Netw. (ICTON), Budapest, Hungary, Jul. 5-9, 2015, paper Tu.B1.5.

[11] M. Schueppert, R. Kruglov, and C.-A. Bunge, "A comparison of one- and higher-dimensional modulation schemes for optical Gigabit transmission," in *Photon. Netw., 15th ITG Symp. Proc.*, 2014.

[12] K. Matsumoto, Y. Yoshida, A. Maruta, A. Kanno, N. Yamamoto and N. Yamamoto, "On the impact of Tomlinson-Harashima precoding in optical PAM transmissions for intra-DCN communication," presented at the Opt. Fiber Commun. Conf. (OFC), San-Diego, CA, USA, Mar. 11-15, 2017, paper Th3D.7.

[13] S. E. Miller and I. P. Kaminov, *Optical Fiber Telecommunications II*. Academic Press INC., 1988.

[14] W.-R. Peng, X. Wu, K.-M. Feng, V. R. Arbab, B. Shamee, J.-Y. Yang, L. C. Christen, A. E. Willner, and S. Chi, "Spectrally efficient direct-detected OFDM transmission employing an iterative estimation and cancellation technique," *Opt. Express*, vol. 17, no. 11, pp. 9099–9111, May 2009.

[15] Z. Li, M. S. Erkilinç, S. Pachnicke, H. Griesser, R. Bouziane, B. C. Thomsen, P. Bayvel, and R. I. Killey, "Signal-signal beat interference cancellation in spectrally-efficient WDM direct-detection Nyquist-pulse-shaped 16-QAM subcarrier modulation," *Opt. Express*, vol. 23, no. 18, pp. 23694–23709, Aug. 2015.

[16] S. Randel, D. Pilori, S. Chandrasekhar, G. Raybon, and P. Winzer, "100-Gb/s discrete-multitone transmission over 80-km SSMF using single-sideband modulation with novel interference-cancellation scheme," presented at the Opt. Commun. (ECOC), 2015 41st Eur. Conf. and Exhib., Valencia, Spain, Sep. 27 – Oct. 01, 2015.

[17] J. Leibrich, A. Ali, H. Paul, W. Rosenkranz, and K.-D. Kammeyer, "Impact of modulator bias on the OSNR requirement of direct-detection optical OFDM," *IEEE Photon. Technol. Lett.*, vol. 21, no. 15, pp. 1033-1035, Aug. 2009.

[18] S. Haykin, "Stationary processes and models," in *Adaptive Filter Theory*, New Jersey: Prentice-Hall, Inc., 1996, pp. 108–114.

[19] K.-D. Kammeyer, *Nachrichtenübertragung*, 4th ed. Wiesbaden: Vieweg+Teubner, 2008.

[20] R. Rath, C. Schmidt, and W. Rosenkranz, "Is Tomlinson-Harashima Precoding suitable for fiber-optic communication systems?," in *Photon. Netw., 14th ITG Symp. Proc.*, 2013.

[21] R. Rath and W. Rosenkranz, "Tomlinson-Harashima precoding for fiber-optic communication systems," presented at the Opt. Commun. (ECOC), 2013 39th Eur. Conf. and Exhib., London, UK, Sep. 22–26, 2013, paper We.2.C.2.

[22] "Forward error correction for high bit-rate DWDM submarine systems," ITU-T Recommendation G.975.1, 2004.