PMD in High-Bit-Rate Transmission and Means for Its Mitigation

Reinhold Noé, *Member, IEEE*, David Sandel, and Vitali Mirvoda

Invited Paper

***Abstract*—Polarization-mode dispersion (PMD) prevents the cost-effective upgrading of fiber networks to 40 and sometimes even to 10 Gbit/s. This paper reviews recent progress in its mitigation and compensation and points out where more research is needed. Electronic PMD mitigation is preferable at 10 Gbit/s, due to its low cost, even though it is accompanied by a considerable residual penalty. A lot of work takes place in the field of optical PMD compensation. Among the numerous detection methods for first-order PMD, we prefer a purely electronic, hence low-cost, arrival time detection method, with a linear readout and ps-sensitivity. Surprisingly, the most easily detectable higher order of PMD is the third order, indicated by a slope steepness difference. Both methods rely on a polarization scrambler at the transmitter side, which can be shared. Regarding PMD compensators, $LiNbO_3$ devices are probably needed to guarantee a sufficient speed. A distributed PMD compensator allows to integrate a number of polarization transformers and differential group delay sections on one chip, thereby exactly emulating the way how the fiber accumulates, but in reverse order and orientation. We report on progress in using these devices, including their use for PMD compensation in a 40-Gbit/s carrier-suppressed return-to-zero differential phase-shift keying experiment. More work is needed to perfection the device, and to implement a fast endless polarization control. The theory of the distributed PMD compensator lends itself to a new definition of higher order PMD by a Fourier expansion of mode conversion, as an alternative to the familiar Taylor expansion of the PMD vector.**

***Index Terms*—Optical fiber communication, optical fiber polarization, polarization-mode dispersion (PMD).**

I. Introduction

POLARIZATION-mode dispersion (PMD) means in the simplest case that a nondichroic optical fiber exhibits a differential group delay (DGD) between two orthogonal principal states-of-polarization (PSPs) [1], [2]. PMD is a big obstacle for high-capacity long-haul optical communication systems. Most if not all telecom carriers with fiber plant dating from 1995 or earlier have problems transmitting 40-Gbit/s signals on that part of their network, sometimes even at 10 Gbit/s. While an upgrade to new, low-PMD fibers is possible, the underlying economics have directed a lot of effort toward PMD mitigation and compensation.

Manuscript received November 10, 2003. This work was supported in part by Deutsche Forschungsgemeinschaft.

The authors are with Optical Communication and High-Frequency Engineering, University of Paderborn, D-33098 Paderborn, Germany (e-mail: noe@upb.de).

Digital Object Identifier 10.1109/JSTQE.2004.827842

The use of solitons can suppress PMD to a certain degree [3]. However, since much of the transmission takes place in a fairly linear regime, PMD must also be handled actively.

Section II of this paper deals with electronic PMD mitigation. Implementation cost is moderate, at the expense of a considerable residual penalty. Electronic PMD mitigation is preferable at 10 Gbit/s, where optical PMD compensation would be too costly.

Optical PMD compensation has started (we believe) in 1994 [4], [5] and has rapidly expanded toward electrooptic [6] and 40-Gbit/s operation [7]. Important recent publications in the PMD, especially PMD, compensation field are, for example, [8]–[15]; many more are cited later.

Endless polarization control, i.e., the tracking of arbitrary paths that may circle the Poincaré sphere many times, is a much more demanding task than polarization control with limited control range. It seems that in many publications insufficient attention has been given to the problem of fast and endless PMD tracking. We believe this problem is usually more difficult to overcome than adaptation algorithm issues, such as the question of local minima.

This, and the interplay between polarization-dependent loss (PDL) and PMD, are outside the scope of this paper.

A limit of optical PMD compensation efforts is given by the onset of cross-phase modulation [16], [17], though this problem may be expected to be smaller in return-to-zero differential phase-shift keying (RZ-DPSK) systems.

The rest of this paper is organized as follows. Section III covers PMD detection, in particular electronic detection of first- and higher-order PMD. Section IV is devoted to optical PMD compensation (PMDC). We favor a distributed PMDC implemented in $LiNbO_3$, and present recent results on such a device. Section V shows that the distributed PMDC provides a new higher order PMD definition, which describes PMD accurately and is an interesting alternative to the current higher order PMD definition. A conclusion is also given.

II. Electronic PMD Mitigation

Most transmission systems where PMD is a concern are operated with optical amplification. At the receiver side there is a last optical amplifier with a gain so large that the shot noise can be neglected; only the amplifier-induced noise persists. In this case the amplifier noise can be modeled as Gaussian field quantities

1077-260X/04$20.00 © 2004 IEEE

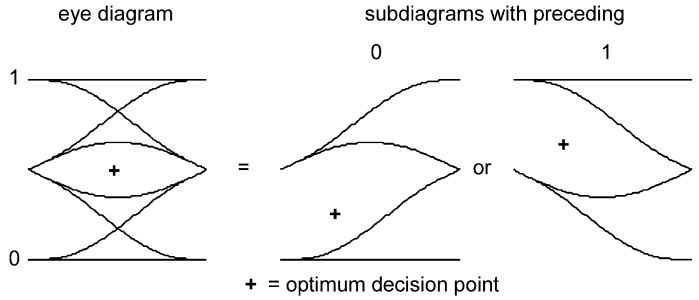

Fig. 1. Electrical PMD mitigation by quantized feedback. Note that the noise amplitude increases with the signal amplitude.

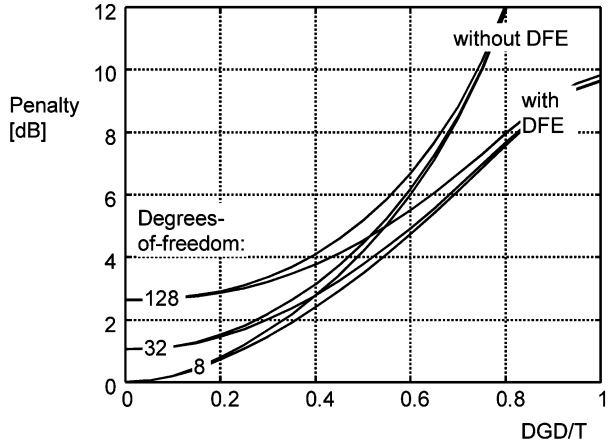

Fig. 2. Calculated OSNR penalty versus normalized DGD.

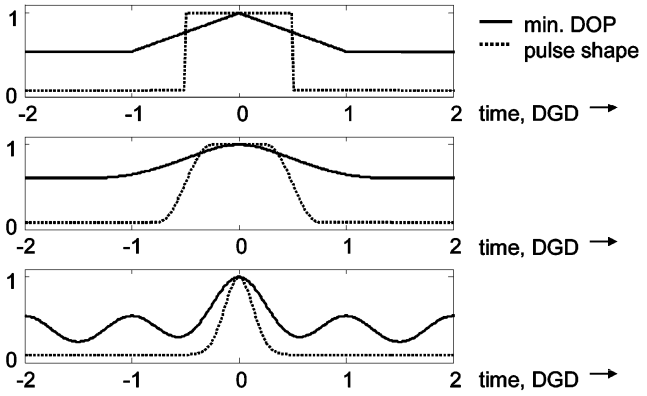

Fig. 3. Minimum DOP versus DGD for different pulse shapes.

in phase and in quadrature to the wanted signal. The photocurrent is proportional to the squared magnitude of the field, and has therefore a χ^2 probability density function. If noise is small and optical filtering is tight, then the decision threshold of the photocurrent is ideally near 1/4 of the full signal swing because the decision threshold of the optical field would be near 1/2 of the signal swing. Usually 1/3 is implemented because of nonideal modulator extinction. A compilation of the mathematical steps is given in [18].

If the eye diagram is fully closed, for example, because the transmitted power is launched with equal powers into both PSPs and the DGD is one bit duration, linear electronic equalizers cannot help. As an example, for the electronic PMD mitigation, consider therefore a decision feedback equalizer (DFE), which can help even in such cases. In Fig. 1 the left part schematically shows the received eye diagram. The decision margin is fairly small because the DGD is close to one bit duration. The right part shows the same eye diagram, but decomposed depending on whether the previous symbol was a zero or a one. The decision feedback can place the effective decision levels at or near the optimum positions.

While the eye opening penalty in the decomposed eye diagrams is less than 3 dB in Fig. 1, the optical signal-to-noise ratio (OSNR) penalty is a lot worse. Especially the preceding ones at the right side are difficult to work with. This is because only true zeros have low noise, and nonzero signals are accompanied by a lot of noise. Fig. 2 shows calculated OSNR penalties. When the DGD equals one bit duration T, the residual penalty with DFE can surpass 9 dB. The number of degrees-of-freedom (DOFs) of the χ^2 distributions is approximately equal to 2 (in-phase and quadrature noise) times 2 (polarizations) times the ratio of the optical bandpass filter width divided by twice the electrical lowpass filter width (at least equal to one). Much fewer than eight DOFs are not well realizable. The true residual penalty can be lower than predicted here, because a nonideal intensity modulator extinction ratio can cause a baseline penalty that masks some of the total penalty. For example, an extinction ratio of -13 dB causes a penalty of at least 1.5 dB.

Also, adding a transversal filter before the DFE improves the situation. Thirdly, thermal noise can never completely neglected and also adds a baseline penalty.

Not surprisingly, a number of experiments with electronic PMD mitigation circuits has shown penalties on the order of several decibels [19]–[25], especially when accompanied by EDFA noise. DFE operation beyond 10 Gbit/s is possible, but difficult [25]. In total, electronic PMD mitigation circuits are very attractive at 10 Gbit/s, where the cost of optical PMD compensators is too big an obstacle. Whether they can be implemented at 40 Gbit/s remains to be seen; this will at least be very difficult. The two next sections will therefore concentrate on PMD compensation in the optical domain.

III. PMD Detection

A. Optical PMD Detection Methods

Since PMD happens in the optical fiber, it is straightforward to detect it in the optical domain, before the photodetector destroys polarization information. Several related methods have mostly been employed: Kikuchi has proposed to monitor the degree-of-polarization (DOP) [26]. Later, Rosenfeldt *et al.* have added polarization scrambling at the transmitter side because this allows one to obtain a more informative readout, including higher order PMD [27]. Fig. 3 shows a simulated DOP readout

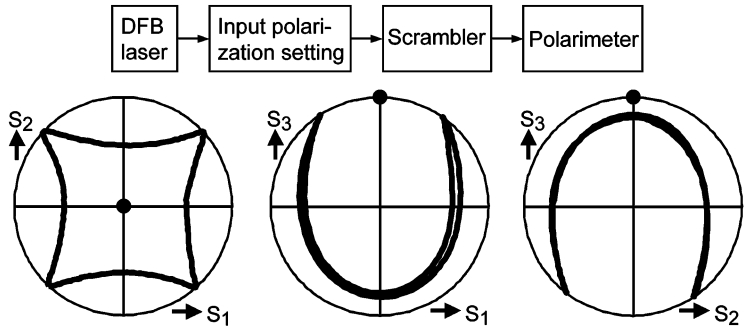

Fig. 4. Measurement setup and projections of measured output polarization trace of "tennis ball" polarization scrambler.

as a function of DGD for various pulse shapes. The readout is proportional to the DGD, but only if the pulse edges are shorter than the DGD. For RZ pulses the readout is somewhat ambiguous.

Rosenfeldt *et al.* have also reported on a polarimeter placed behind a tunable, frequency-swept optical bandpass filter [28]. Recently, Möller *et al.* have replaced the tunable filter by an angled fiber Bragg grating in view of polarimetry [29]. Photodetector arrays provide a parallel readout of the frequency-dependent Stokes parameters. In principle this is a low-cost solution that should allow one to construct such a spectral polarimeter, covering a whole WDM band. However, in practice it is more common to monitor the signal behind a PMD compensator that is dedicated to one WDM channel, and this would mean that one spectral polarimeter is needed for each WDM channel. Frequency-independent inline polarimeters can also be fabricated based on fiber Bragg gratings [30].

The efficiency of all these methods is of course increased if measurements take place at several positions, for example, at the input, at an internal tap, and at the output of an optical PMD compensator [27], [31], [15].

Since cost seems to be the most important decision criterion nowadays, the authors prefer electronic PMD detection methods. Some of them need polarization scrambling just like a few optical methods [27], [31].

B. Polarization Scrambling

Polarization scrambling is an ancillary science to PMD detection. A variety of optical PMD detection techniques [27], [31] need polarization scramblers. In most cases it is possible to share one scrambler by many WDM channels, so the cost is not prohibitive. Polarization-independent scramblers can be constructed [32], but the time-bandwidth product is so large that they are not very attractive: Either the PMD detection interval must be very long or the electronic processing bandwidth of the (optical or electronic) PMD detection device must be quite large.

We have therefore implemented a polarization scrambler with the lowest possible time-bandwidth product [32]. It works for a pair of fixed orthogonal input polarizations. This means that the WDM channel signals must be combined at the transmitter side with polarization-maintaining components.

Let the time-variable scrambler output polarization and a PSP of a subsequent fiber be denoted by the normalized Stokes vectors $\mathbf{S}_{\mathrm{SC}}(t)$. We drop the time (t) dependence and introduce the averaging operator $\langle\cdot\rangle$. The covariance matrix of the scrambler output polarization is

$$\mathbf{C} = \langle(\mathbf{S}_{\mathrm{SC}} - \langle\mathbf{S}_{\mathrm{SC}}\rangle)(\mathbf{S}_{\mathrm{SC}} - \langle\mathbf{S}_{\mathrm{SC}}\rangle)^T\rangle \tag{1}$$

where T denotes transposition. Ideally, $\langle\mathbf{S}_{\mathrm{SC}}\rangle$ should vanish and $\mathbf{C}$ should become equal to 1/3 times the identity matrix. Obviously its three eigenvalues are all equal to 1/3. One solution is the Stokes vector

$$\mathbf{S}_{\mathrm{SC}} = \begin{bmatrix} (1 \pm \sqrt{1/3})/2 \cdot \cos\omega t - (1 \mp \sqrt{1/3})/2 \cdot \cos 3\omega t \\ (1 \pm \sqrt{1/3})/2 \cdot \sin\omega t + (1 \mp \sqrt{1/3})/2 \cdot \sin 3\omega t \\ \sqrt{2/3} \cdot \cos 2\omega t \end{bmatrix} \tag{2}$$

where ω is the scrambling frequency. A single electrooptic waveplate [33] on a commercially available X-cut, Z-propagation $LiNbO_3$ chip has been used with circular input polarization, "rotating" orientation, and periodically time-variable retardation to implement such a scrambler. The upper signs in (2) were chosen. Fig. 4 shows projections of Poincaré sphere traces, taken at the device output at a low speed. They look like the vulcanization line of a tennis ball. The eigenvalues of $\mathbf{C}$ were very close to the ideal values of 1/3, with relative deviations from their mean of just $\pm 1.7\%$. When the wavelength was tuned, the lowest eigenvalue stayed above 0.28 over more than 40 nm bandwidth.

C. Electronic PMD Detection Methods

Among the electronic methods for first-order PMD detection, we favor the arrival time detection scheme [34], [36]. In the presence of PMD, the polarization modulation caused by the scrambler is accompanied by an arrival time modulation. If the slow PSP of a first-order PMD medium with a DGD τ is given by the normalized Stokes vector $\mathbf{S}_{\mathrm{PSP}}$, then the pulse arrival time is

$$\Delta\hat{t}(t) = (\tau/2)\mathbf{S}_{\mathrm{PSP}}^T\mathbf{S}_{\mathrm{SC}}(t).$$

The variance of $\Delta\hat{t}$ is

$$\sigma_{\Delta\hat{t}}^2 = (\tau/2)^2\mathbf{S}_{\mathrm{PSP}}^T \cdot \mathbf{C} \cdot \mathbf{S}_{\mathrm{PSP}}.$$

With (1) and (2), the result $\sigma_{\Delta\hat{t}}^2 = \tau^2/12$ is obtained. The rms value $\sigma_{\Delta\hat{t}}$ is directly proportional to the DGD. This is a great advantage because even small DGDs can be accurately detected.

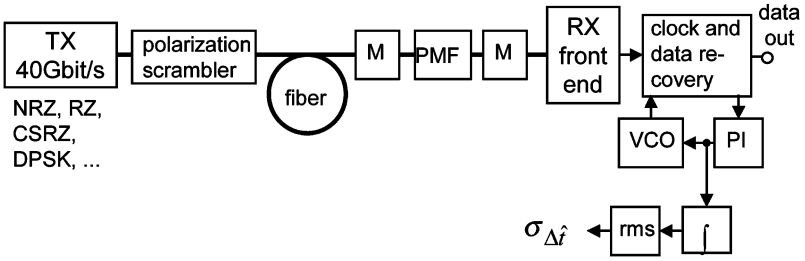

Fig. 5. First-order PMD detection using tennis ball scrambler and arrival time detection. M = motorized endless fiber-optic polarization transformers.

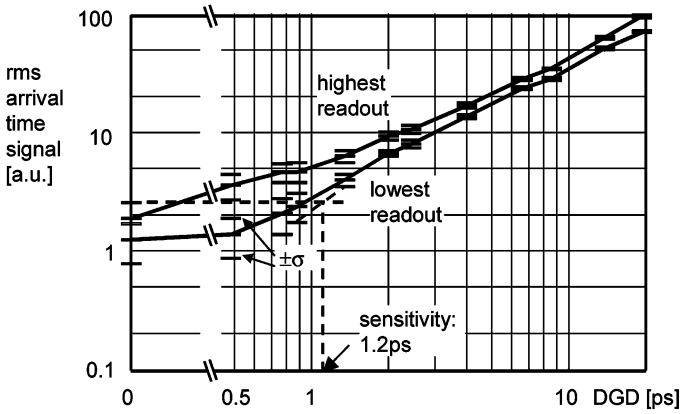

Fig. 6. PMD detection readout for tennis ball scrambler and arrival time detection scheme in a 40-Gbit/s CSRZ-DPSK setup. Solid curves correspond to lowest and highest readouts caused by motorized polarization transformers. Associated ± one standard deviation intervals are indicated by "−" and allow one to assess the sensitivity.

The purely electronic first-order PMD detection takes place in a setup according to Fig. 5. In our first experiments a 40-Gbit/s NRZ transmission system was used. The tennis ball scrambler was operated with a fundamental scrambling frequency of 417 kHz. Precharacterized polarization-maintaining fiber (PMF) pieces were inserted as PMD devices under test. In order to include any uncontrollable polarization influence including PDL, slow motorized polarization transformers (M) were added before and after the PMF. The clock and data recovery in the optical receiver features a clock phase detector. Together with a voltage-controlled oscillator (VCO) and a loop filter, a proportional-integral controller (PI), it forms a phase-locked loop (PLL). The clock recovery PLL tracks the PMD-induced arrival time fluctuations. If it is fast enough, then the signal at the VCO input is proportional to the temporal derivative of the arrival time modulation. It is sufficient to integrate this signal, and to measure the rms value of the integral, thereby obtaining $\sigma_{\Delta\hat{t}}$. The measurement interval was one scrambling period, 2.4 μs. The sensitivity, where the readout minus one standard deviation was more than the background readout plus one standard deviation, was about 0.88–1.35 ps [32]. Very similar readouts result for RZ modulation. A 210-km transmission and PMD compensation experiment was also conducted [36]. There was normal PDL in the link. No attempts to suppress it were made here, nor in the next experiment.

Recently the arrival time detection scheme was also tested for carrier-suppressed return-to-zero (CSRZ)-DPSK signals. In that case the receiver front end consisted of an interferometer and two photodetectors, and it was connected to the differential inputs of the clock-and-data recovery. The exact setup is shown later in Fig. 16. The obtained readouts are very similar to those obtained for NRZ intensity modulation in [32]. Fig. 6 shows a fairly linear readout from about 1 to 19 ps. The sensitivity was assessed to be 1.2 ps.

Quite interestingly, arrival time detection of PMD is also possible in polarization multiplexed transmission systems using the RZ signal format [35]–[37]. The big advantage here is that no scrambler is necessary at all because it is replaced by a differential phase modulation between the two orthogonal polarization channels, which in turn is realized by a small laser frequency modulation combined with a differential delay of the two polarizations before they are modulated. A PMD detection sensitivity of 150 fs has been achieved in a 4.8-μs measurement interval, and RZ polarization division multiplex transmission over 212 km with PMD compensation became possible.

Among the earliest reported electronic methods for PMD detection is the power measurement in bandpass portions of the photodetected electric signal [5]–[7]. Since PMD is generally a low-pass effect, it is sufficient for PMD compensation to try to maximize the high-frequency components of the electric signal. No scrambler is needed, but the readout for small DGDs rises with the square of the DGD.

PMD can also be detected by threshold scanning [38] in an extra decision circuit, or by reading the FEC raw errors. These digital methods are relatively easy to implement but suffer from speed penalties, since it takes extra time to filter out the quantization noise, compared to an analog detection technique.

Table I summarizes the properties of the most important optical and electrical first-order PMD detection methods. Arrival

TABLE I
HOW TO DETECT FIRST-ORDER PMD

Measurement	Threshold scanning or FEC	Electrical bandpass filters	Arrival time detection	Polarimetry
Polarization scrambler needed	no	no	yes	no **
Extra optics in each WDM channel	no	no	no	no ***
Extra RF electronics	yes (FEC: no)	yes	no	no
Readout is $\propto \mathrm{DGD}^n$; n =	2	2	1	1*
Speed	slow	fast	fast	fast ***

* as long as pulse rise and fall times are shorter than DGD
** in principle *** grating-based spectrometer

time detection possesses most advantages and can be realized with commercially available technology.

Higher order PMD can also be detected electrically, to some degree. Francia *et al.* have reported that a PMD vector that rotates in the Stokes space as a function of optical frequency generates asymmetric (leading or trailing) overshoot in detected NRZ signals [39].

According to our investigations [32], the asymmetric distortion is due to third-order PMD and is characterized by a steepness difference between rising and falling electrical signal slopes, while second-order PMD causes pulse distortions with a front-to-back mirror symmetry. To see this it is sufficient to pass an optical data signal through field transfer functions equal to $e^{-j\omega^2\beta'' L/2}$ (second-order chromatic dispersion, symmetric distortion) and $e^{-j\omega^3\beta'' L/6}$ (third-order chromatic dispersion, asymmetric distortion; see also [40]). In effect, the Taylor expansion of a circular trajectory must be of infinite order to be correct. So there is not only second- but also a lot of third-order PMD. The effect can only be understood to be of second order (but not to be due to second-order PMD) if one takes a PMD model with two cascaded DGD sections.

The important thing about the asymmetric distortions with slope steepness difference, caused by third-order PMD, is that they can be detected more sensitively than symmetric overshoot, caused by second-order PMD, see Fig. 21 of [32].

To discuss a very similar situation more in detail, we need to introduce now DGD profiles. The PMD vector Ω has a length equal to the DGD and points in the direction of a principal state-of-polarization in the three-dimensional normalized Stokes space. If n DGD sections are cascaded, the overall PMD vector is given the sum $\Omega = \sum_{i=1}^{n} \Omega_i$ of individual PMD vectors $\Omega_i = \mathbf{R}_{<i}^{-1}\Omega_{\mathrm{local},i}$. The Ω_i are referred to the input of the whole cascade, whereas $\Omega_{\mathrm{local},i}$ are the local individual PMD vectors. $\mathbf{R}_{<i}$ is the product of all 3×3 rotation matrices, which represent the retarders (including DGD sections) preceding the DGD section i. Plotting the sequence of Ω_i in such a way that the tail of Ω_{i+1} starts from the head of Ω_i results in a DGD profile. Its endpoint is given by $\mathbf{\Omega}$. More details and examples of measured DGD profiles are given in [32], [7], and [37].

After compensation of first-order PMD, the DGD profile origin and endpoint must coincide (at the optical carrier frequency) because this means that the first-order PMD is zero. The resulting DGD profile will most likely consist of a loop or a similar closed trajectory shown in Fig. 7(left). The nearly parabolic trajectory P is the position of the PMD vector as a function of frequency. The linear motion (LM) means second-order PMD. The quadratic (forth and back) motion (QM) is in this approximation due to third-order PMD. And here, as for any other DGD profiles, it is the projection (PQM) of the quadratic motion QM along the input polarization (arrow) that determines the amount of asymmetric overshoot, not second-order PMD. The corresponding simulated eye diagram is shown in Fig. 7(b). The auxiliary lines underline the slope steepness difference.

The presence of DGD profile loops, which exhibit second- and third- and all higher orders of PMD, has been detected in the 40-Gbit/s transmission setup of Fig. 8. The signal passed a PMD emulator (PMDE) and compensator (PMDC). The former consists of two mechanical polarization transformers (M1, M2), each of which is followed by a DGD section. The latter has the same structure but uses electrooptic x-cut, z-propagation $LiNbO_3$ polarization transformers (E1, E2). At the receiver side there was arrival time detection. The PMDC was controlled as to minimize first-order PMD. In addition, a slope steepness difference detector acted on the photodetected signal. It was built on a ceramic substrate with a differentiator and two oppositely poled one-way rectifiers. The usable frequency range was up to $\sim$20 GHz and turned out to be acceptable. The sum of the two rectified signals represents the slope steepness difference.

Experimentally obtained eye diagrams [41] are shown in Fig. 9, displayed with negative polarity, and recorded while the scrambler was stopped. For the eye diagrams at the bottom the third-order PMD QM is parallel or antiparallel to the input polarization.

The 2.4-μs scrambling period was also chosen for slope steepness difference measurement. In order to maximize higher order PMD, the controller was first set so that it tried to maximize the rms slope steepness difference by E1, E2. The manual polarization transformer M2 was also adjusted with the same aim. Six cases were investigated, where the subsequent sections had DGDs of 0 ps (only one manual, no electrooptic polarization transformer, essentially no PDL), 0+ 0 + 0 + 0 ps (back-to-back, with PDL just as in all following cases), 2.2 + 4 + 4 + 2.2 ps, 6.25 + 4 + 4 + 6.25 ps, 6.25 + 6.25 + 6.25 + 6.25 ps, and 6.25 + 19 + 22.8 + 6.25 ps, respectively. Nearly flat, rhomboid-shaped DGD profiles are expected, with areas of 0, 0, 8.8, 25, 39, and 261 ps^2, respectively. To give an example, the associated second-order PMD for the 6.25 + 6.25 + 6.25 + 6.25 ps case is 78 ps^2. Since the setup was also influenced by PDL ($\sim$1 dB), mainly from E1, E2, the worst case was assessed by adjusting M1 for smallest readout. This does not change the rhomboid; it just reorients it in the PMD vector space. The rms slope difference readouts were recorded as well as their standard deviations. After this the controller was switched in each case to minimize the rms slope difference readout, while M1, M2 were left unchanged. Value and standard deviation were again recorded. In Fig. 10 the solid lines are readout values and the "+" signs (sometimes hardly discernible) indicate $\pm$ one standard deviation error intervals.

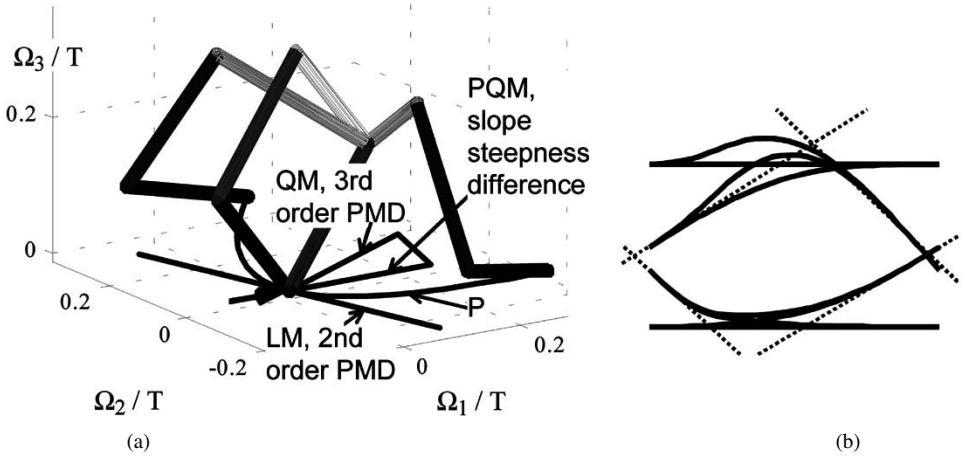

Fig. 7. DGD profile for vanishing first-order PMD. Output arrows of DGD profile are not shown for easier visibility. Corresponding eye diagram exhibits slope steepness difference.

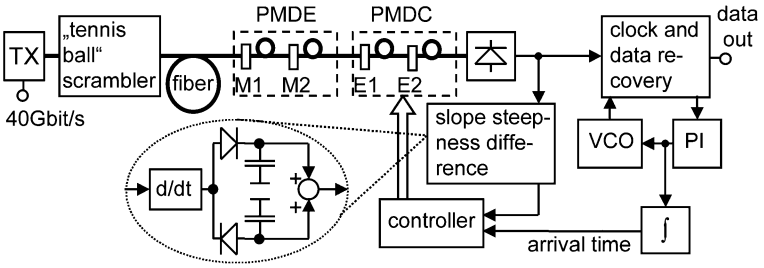

Fig. 8. 40-Gbit/s transmission setup with first- and third-order PMD detection and PMD compensation.

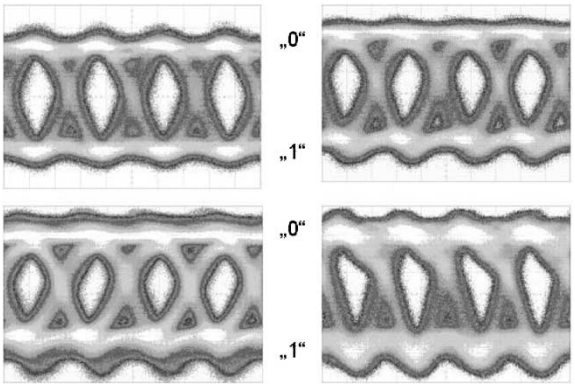

Fig. 9. 40-Gbit/s eye patterns back-to-back (top left), with unequal steepnesses of rising and falling slopes (bottom) and with unequal curvatures (all except back-to-back). Polarity is negative. Transmission took place in the setup of Fig. 8, which was configured to have a DGD profile like in Fig. 7, with 6.25 + 6.25 + 6.25 + 6.25 ps of DGD in four sections, but the scrambler was stopped.

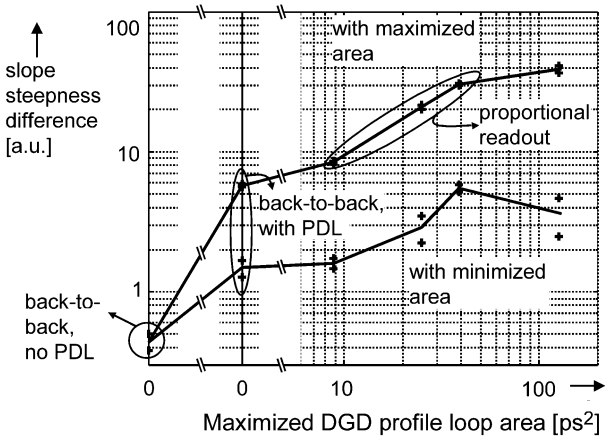

Fig. 10. Slope steepness difference readout versus DGD profile loop area. Lower curve results when controller tries to achieve zero DGD profile area.

Between maximized rhomboid areas of 8.8 and 39 ps^2, a rough proportionality of the readout was found as expected. It is seen that between the smallest of the maximized rhomboid areas (8.8 ps^2) and the maximized back-to-back case there is a readout difference of several standard deviations. The back-to-back case with reduced PDL and without PMDC (only $M1 + 0$ ps remained) was again a lot better. So, the large PDL had an influence on the setup but even in the worst case it was possible to detect the negligibly small DGD profile rhomboid $2.2 + 4 + 4 + 2.2$ ps within just 2.4 μs. The compensator was able to essentially close it, as well as all larger DGD profile loops. First-order PMD was also compensated at all times.

A simultaneous Q factor measurement was not tried because threshold scanning influences the arrival time detection, and thermal drift in the DGD sections of the PMDE made it difficult to maintain a certain measurement condition over a longer time. Nevertheless, the assessed PMD situations resulted in particularly distorted eye diagrams (under the condition that first-order PMD was essentially eliminated).

IV. Optical PMD Compensation

A. Devices and Methods

Before specializing on one PMD compensation device and method we review here some work in this field.

Inside an optical PMDC with cascaded DGD sections each of its polarization transformers must be able to endlessly transform any input polarization into a principal state-of-polarization of the subsequent differential group delay section [7]. Endless tracking means that polarization fluctuations may at any point in the fiber link go around the Poincaré sphere more than once, possibly an infinite number of times. This is a much more difficult task than polarization control with limited control range: For any polarization transformer, the tracking of certain small polarization fluctuations will require large drive signal changes. For example, when an electrooptic Soleil–Babinet compensator [33] (rotatable waveplate with adjustable retardation) with circular input polarization reaches a retardation angle of π while the output polarization marches straight through the point of circular retardation, then it must change its orientation angle by π; else the retardation angle and the required driving voltages will keep increasing, eventually destroying the device.

Polarization transformers with retarders having fixed eigenmodes are more complicated to control. Procedures are necessary where a retardation is typically reset by 2π under cooperation of other retarders to keep the output polarization unchanged [34]. Fiber squeezers and most kinds of liquid crystal cells belong to this class. The tracking speed that can be achieved is minimum if the absolute value of the retardation is not stable. Stability is usually not specified by manufacturers. Even when the retardation is stable the safe endless tracking speed is just 0.033 rad/iteration [42]. We therefore see the potential of such retarders for PMD compensation as quite critical. For combination with a high-speed controller, we believe the response time should be on the order of 5 μs or less.

All polarization transformers can be operated as endless devices even without stable retardation if there are many of them. However, the speed penalty paid for this simplification is dramatic.

A workhorse for a lot of experimental PMD compensation work are electrooptic waveplates in x-cut, Z-propagation $LiNbO_3$. These devices are not birefringent and have a wide optical bandwidth. The response is increased by a finite buffer layer isolation, but is therefore not fully instantaneous.

In addition to endless control of one variable polarizaiton with an electrooptic Soleil–Babinet compensator [33] two Soleil–Babinet compensators [43] or three waveplates [44] have been used to transform any input into any output polarization. Devices where the waveguide birefringence is tuned out completely by a slight off-axis propagation can be operated with low offset voltages. But the waveguide exhibits a reciprocal circular retardation. Another issue is dc drift [45]: Charges generated by the pyroelectric effect are separated under the influence of a static external electric field, thereby weakening this field inside and near the waveguide. Ion migration in the buffer layer and/or conductivity disturbances of the crystal cause a similar effect. DC drift limits are not known or at least not specified for commercial x-cut, z-propagation $LiNbO_3$ polarization transformers. The dc drift lets the retardation offset calibration drift slowly with time, and this is the biggest problem.

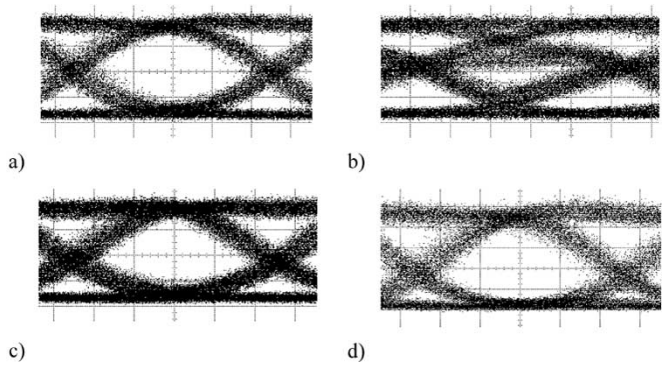

Fig. 11. 20-Gbit/s eye diagrams (a) back-to-back, (b) with emulator and idle PMDCs (unusually good here), (c) with emulator and working PMDCs, and (d) with PMDCs alone.

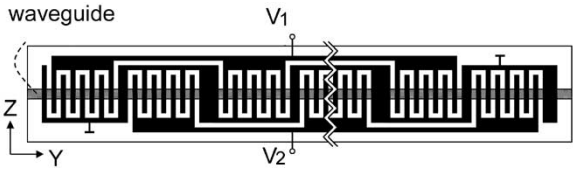

Fig. 12. Elementary in-phase and quadrature mode converter (one section) in x-cut, y-propagation $LiNbO_3$.

Polarization controllers made of any other materials than $LiNbO_3$ are less suitable in our opinion.

B. Distributed PMD Compensator

Regarding cost and space, the integratability of several polarization transformers and DGD sections into one device is decisive. This requirement, and most others, are best fulfilled in a distributed PMDC. In a preparatory attempt a weakly polarization-maintaining fiber with a total DGD of 77 ps was twisted at 64 positions. This resulted in successful PMD compensation at 20 Gbit/s [7].

Later the principle was implemented in an x-cut, y-propagation $LiNbO_3$ chip. At 20 Gbit/s, 10+20 ps of DGD were compensated by two cascaded distributed $LiNbO_3$ PMDCs with a total DGD of 43 ps [46]. Corresponding eye diagrams are shown in Fig. 11. The device was also tested at 40 Gbit/s [47].

In the following we will concentrate on recent progress with these devices [48], [49]. Fig. 12 shows a basic polarization transformer, an in-phase and quadrature mode converter as proposed by Heismann and Ulrich [50]. Several or many of them are cascaded in a distributed PMDC. Comb-shaped mode converter electrodes are distributed along a birefringent waveguide.

The electrode period equals the TE-TM beat length (22 μm at 1550 nm). This mode converter may be called a Soleil–Babinet analog SBA(φ, ψ) [7] because of the analogy to a Soleil–Babinet compensator: What the Soleil–Babinet compensator does with circular, the SBA can do with horizontal/vertical polarizations. It is described by a Jones matrix

or a 3×3 rotation matrix (=Mueller submatrix), shown in (3) at the bottom of the page, respectively. An SBA introduces an angle equal to its retardation φ between the PMD vector directions of a preceding and a following DGD section, thereby rotating the subsequent rest of the DGD profile by φ. Angle ψ defines the phase shift between an unconverted and a mode-converted wave and can be called an orientation angle. An SBA retardation of a DGD τ is represented by a phase shifter PS$(-\omega\tau)$ with Jones and rotation matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{j\omega\tau} \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega\tau & \sin\omega\tau \\ 0 & -\sin\omega\tau & \cos\omega\tau \end{bmatrix}$$

respectively, where ω is the optical angular frequency. The whole distributed PMDC is approximately described by the Jones or rotation matrix product

$$\text{PMDC} = \prod_{i=n}^{1} (\text{PS}(-\omega\tau_i)\text{SBA}(\varphi_i, \psi_i)). \tag{4}$$

The product must be executed from left to right with descending index i while the light passes the SBAs and PSs in ascending order i. A longitudinally variable pattern of in-phase voltages $V_{1,i}$ and quadrature voltages $V_{2,i}$ (Fig. 12) couples TE and TM modes. Using a constant G, we can write

$$\varphi_i e^{j\psi_i} = G(V_{1,i} + jV_{2,i}). \tag{5}$$

Note that it holds $\varphi \propto \sqrt{V_1^2 + V_2^2}$ and $\psi = \arg(V_1 + jV_2)$.

A pigtailed PMDC was electrooptically investigated. It had $n = 69$ pairs of in-phase and quadrature electrodes, each segment $i = 1 \ldots n$ being $\sim$1.27 mm long. The total PMD was $\sim$23 ps, which is good for 40-Gbit/s PMD compensation. The PMDC was operated with a horizontal input polarization in the waveguide. A polarimeter was connected to the PMDC output. Consider the case when all voltages except those of segment i are zero. With PMDC given by its Jones or its rotation matrix, the output polarization state can be expressed by the Jones or normalized Stokes vector

$$\begin{bmatrix} \cos\varphi_i/2 \\ je^{-j(\psi_i+\zeta_i)}\sin\varphi_i/2 \end{bmatrix} = \text{PMDC} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} \cos\varphi_i \\ \sin(\psi_i+\zeta_i)\sin\varphi_i \\ -\cos(\psi_i+\zeta_i)\sin\varphi_i \end{bmatrix} = \text{PMDC} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{6}$$

respectively. Here $\zeta_i = -\omega\sum_{k=i}^{n}\tau_k$ are phase shifts. If they are modulo 2π unequal, then the waveguide is nonuniform. Four voltage pairs $V_{1,i} + jV_{2,i} = 20\ \text{V} \cdot \{1, j, -1, -j\}$ were applied sequentially to each of the electrodes in section i, and the output polarization state was measured. Equations (5) and (6) allowed one to determine φ_i and ζ_i. Performance was best at

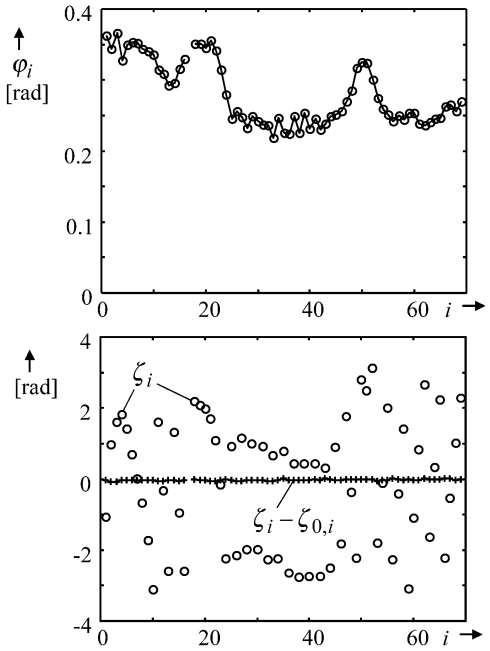

Fig. 13. (Top) Polarization change φ_i and its orientation ζ_i (o) and (b) orientation error $\zeta_i - \zeta_{0,i}$ (+, solid line) for $V_{1,i} + jV_{2,i} = 20$ V. Point $i = 17$, where both electrodes had internal shorts to ground, is omitted.

$\lambda = 1540$ nm, where φ_i varied somewhat as a function of i but not too much (Fig. 13).

Waveguide nonuniformity can decrease the conversion efficiency of long mode converters. Indeed, ζ_i varied unusually much due to errors which occurred in the production of this particular device. So, if a number of adjacent in-phase electrodes were connected in parallel, and likewise for quadrature electrodes, then a low mode conversion efficiency would result, especially for $i = 25 \ldots 45$ where adjacent ζ_i differ by about π. Apart from improved fabrication, the only way to combat this is to subdivide the mode converters into fairly short sections, which was done here. However, it has been argued that a large number of electrode voltages cannot be controlled sufficiently fast.

In PMD compensation algorithms, it is common to dither various electrode voltages, one at a time, and to optimize them by a multidimensional gradient search using one or more control criteria. For a maximum voltage of 55 V (=10 V/μm), a waveguide length of $\sim$4–6 mm will be needed for full mode conversion, and it is therefore possible to reduce the number of control variables below $2n$. A straightforward approach is to specify, say, four "discrete" polarization transformers at 0, $L/4$, $L/2$, and $3L/4$, where L is the chip length. Due to the previous calibration process, just two control variables are required per polarization transformation, which is the physically necessary minimum. Such a PMDC needs $\leq$48 voltages but just eight

$$\begin{bmatrix} \cos\varphi/2 & je^{j\psi}\sin\varphi/2 \\ je^{-j\psi}\sin\varphi/2 & \cos\varphi/2 \end{bmatrix}, \begin{bmatrix} \cos\varphi & -\sin\psi\sin\varphi & \cos\psi\sin\varphi \\ \sin\psi\sin\varphi & \cos^2\psi + \sin^2\psi\cos\varphi & \cos\psi\sin\psi(1-\cos\varphi) \\ -\cos\psi\sin\varphi & \cos\psi\sin\psi(1-\cos\varphi) & \sin^2\psi + \cos^2\psi\cos\varphi \end{bmatrix} \tag{3}$$

control variables, and is four times more variable than a PMDC featuring one commercial x-cut, z-propagation $LiNbO_3$ device. The latter requires up to 16 control variables for its eight waveplates, but a reduction of the number of control variables is also conceivable there.

An alternative approach employs a number of spatial Fourier coefficients

$$F_k = n^{-1} \sum \varphi_i e^{j(\psi_i + \zeta_{0,i})} e^{-j2\pi ik/n} \tag{7}$$

of retardation or mode coupling as control variables. The required inverse Fourier transform and multiplication can be implemented in an FPGA for fast execution

$$V_{1,i} + jV_{2,i} = G^{-1} e^{-j\zeta_{0,i}} \sum F_k e^{j2\pi ik/n}.$$

For a distributed PMD compensator, the DGD profile is a straight rod as long as no voltages are applied. The Fourier coefficient F_0 causes a bend or loop, and F_k with $k \neq 0$ a spiral of the DGD profile. The coefficients permit a precise control of the DGD profile shape. Optimizing the low-order coefficients more often than the high-order ones may improve the mean PMD tracking speed. Some 38 control variables or so (real and imaginary parts of $F_k, k = -9, -8, \ldots, 9$) may be sufficient for a 96-mm-long chip. If the center of gravity of the coefficient spectrum, rms-averaged over extended times, is not near $k = 0$, there may be a static phase mismatch. To overcome it, one may correct the chip temperature: A temperature change by $k \cdot (\pm 0.44\ \mathrm{K})$ will translate the center of gravity from k to zero for a 96-mm-long chip. The sign depends on the definition of in-phase versus quadrature.

Note, however, that at the chip input an extra polarization transformation, independent of the Fourier coefficients, should preferably be implemented, in order to orient the waveguide DGD as soon as possible in such a direction that PMD is compensated.

At the experimental side we have, for a start, optimized the control performance of one SBA. Six mode converter cells (sections) were used, four of which were sufficient for a halfwave SBA operation (full mode conversion). After calibration the sections form a single retarder, which depends only on total retardation and total orientation as parameters.

A PSP, say, TE, is fed into the polarization transformer. The device was initially operated as a rotating halfwave SBA, and the output polarization (near TM) was recorded with a polarimeter. After proper amplitude adjustment of the sinusoidal voltages, the output polarization stayed within a circle having a radius of just 0.05 rad on the Poincaré sphere. Fig. 14 shows output polarizations when the SBA retardation was varied from zero to almost π for various SBA orientations. They converge fairly well in one point, which is essential for endless polarization tracking. According to our experience, such a good accuracy was not achievable with electrooptic waveplates in x-cut, z-propagation $LiNbO_3$ from two vendors.

Next a moving target polarization, which the polarization controller has to track, was defined by software. Only the distance (on the Poincaré sphere) between the actual polarization, measured by a slow polarimeter, and the target was used

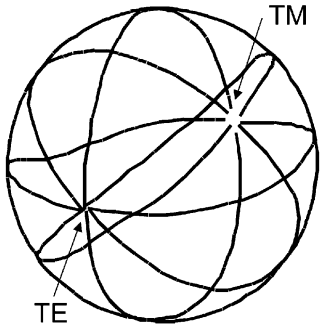

Fig. 14. Poincaré sphere meridians generated by SBA where the retardation varies from zero (uncritical "pole," "TE") to almost π (critical "pole," "TM") with various orientations. Coordinate system rotation is arbitrary.

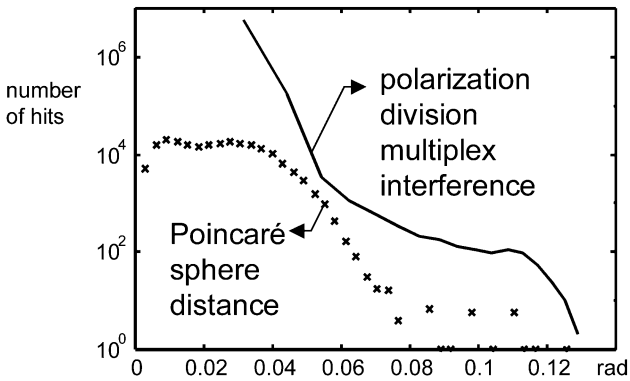

Fig. 15. Misalignment angle distributions during worst case polarization tracking with a polarimeter ($\times$) and during random polarization tracking using a polarization division multiplex interference signal (—).

in the control algorithm. Virtually all possible polarization states and trajectories were investigated but extreme care was taken to scrutinize the behavior at and near the "pole" of TM polarization. In 30 most critical great circle trajectories across the poles, the worst displacement from the wanted trajectory was <0.125 rad (Fig. 15). Most measurements had deviations <0.04 rad. The tracking speed was 0.012 rad/iteration.

In order to get rid of the slow polarimeter, a polarization division multiplex signal was set up, and the interchannel interference was monitored as apolarization error signal [51]. Fig. 15 (—) shows the misalignment angle distributions obtained while random polarization variations are being tracked. These were generated by motorized endlessly rotating fiber coils. The tracking speed was $\sim$1.3 rad/s or 0.002 rad/iteration (one iteration $=$ 1.5 ms). The worst case misalignment was <0.13 rad. Note, however, that endless tracking is a worst case scenario that happens rarely. The instantaneous (non endless) tracking speed is more than 100 times higher: signal acquisition from anywhere on the Poincaré sphere occurs in ≤ 5 iterations or so.

The slower permissible endless tracking per iteration, compared to the previous case and compared to the 0.033 rad/iteration in [42], is believed to be due to a nonrectangular electrooptic step response of the device. It is caused by a higher conductivity of the buffer layer compared to the $LiNbO_3$. More work needs to be done to shorten the electrooptic step response. Alternatively, the response may be electrically

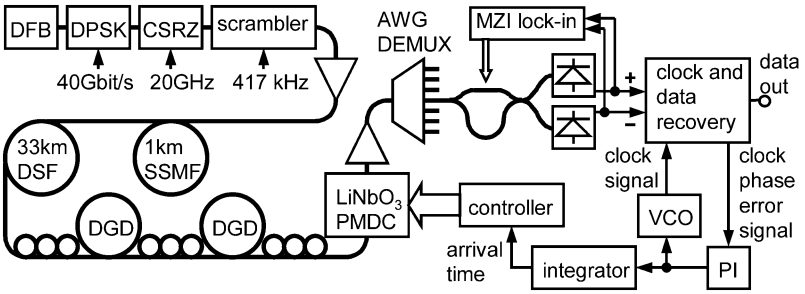

Fig. 16. 40-Gbit/s transmission setup with first- and third-order PMD detection and PMD compensation.

equalized. At present the tracking speed is heavily limited by deliberately imposed algorithmic constraints, which could be released if the step response were fast and rectangular. This is expected to be a key to faster tracking. The execution time per iteration can probably cut down to <50 μs by a suitable hardware effort.

The dc drift problem observed for x-cut, z-propagation devices must also be discussed here. For random polarization variations the voltages needed to control a distributed PMDC will have zero means. This should in principle avoid all dc drift. But in reality polarization varations may not be as random as required. However, if a halfwave SBA with slowly rotating orientation angle is added at the input of the PMDC, a differential phase TE-TM shift is generated. In order to maintain the desired DGD profile, the SBA orientations in the rest of the distributed PMDC must rotate at twice this rate. This can also be accomplished automatically during the control process. As a consequence, all driving voltages will be dc-free even if the PMD scenario is static [48].

Recently, this distributed $LiNbO_3$ PMDC was tested in a 40-Gbit/s CSRZ-DPSK transmission setup (Fig. 16), similar to that in [52]. Arrival time detection was implemented; for results see Fig. 6. For PMD compensation, the "tennis ball" scrambled signal was transmitted over available fiber, 33 km of DSF, and 1 km of SSMF. At the receiving end there were two DGD sections, inserted between three motorized fiber-coil-type endless polarization transformers. The distributed PMDC was operated as a four-section PMDC with polarization transformers at positions 0.1, 0.35, 0.55, and 0.75 L, where L is the full chip length ($\sim$23 ps of DGD). A controller processed the arrival time signal and adjusted the driving voltages for minimum residual PMD. A total of 48 voltages was used, which could in the future be controlled by just eight degrees-of-freedom as described above.

Measured spectra at the output of the integrator are displayed in Fig. 17, here with DGD sections equal to 8.6 and 6.6 ps, respectively. When PMD compensation is on, the arrival time signal is typically $>$30 dB lower than when it is stopped, indicating a $>$30-fold DGD reduction. The multiples one, two, and three of the fundamental frequency of 417 kHz are expected to be found for the "tennis ball" curve of the scrambled polarization on the Poincaré sphere. Higher harmonics were also observed here, but their amplitudes are unnecessarily increased by the fact that the integrator was not a true integrator and had a constant gain beyond 2 MHz. The higher harmonics present no significant disadvantage.

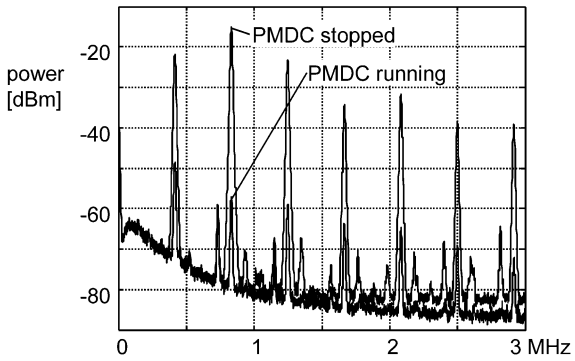

Fig. 17. 40-Gbit/s CSRZ-DPSK transmission. Spectra at integrator output with PMDC stopped or running.

The back-to-back receiver sensitivity (without scrambler, fiber, DGD sections, PMDC) at the input connector of an optical preamplifier equals -32.0 dBm at a BER equal to 10^{-9}. DPSK eye pattern monitoring using an oscilloscope was not possible in the PMD-compensated setup, because the photodetected signals need to be connected to the clock and data recovery. However, measured Q factors for various configurations are given in Fig. 18, using 2–10-s-long counting intervals for each BER value. The various Q values belonging to a particular configuration were obtained for different settings of the fiber polarization controllers, in the quest of best and worst cases. Note that the scrambler alone reduces the back-to-back Q factor from 29.5 dB (measured separately) to $\geq$23.7 dB. This is mainly due to PDL in the photodiodes: the longer the BER counting intervals were, the later the extrapolation began, and the better was the Q factor. So, most given Q factors are likely to because the noise probability density function falls off more sharply than that of Gaussian noise. Two DGD sections having 6.6 and 8.6 ps of DGD, in addition to scrambler, PMDC, and 34 km of fiber, resulted in Q factors $>$20 dB.

The PMDC was also tested dynamically by activating the motorized fiber polarization controllers. Operation was usually error-free. Occasional errors are believed to be due to higher order PMD, which was not detected here.

The bottom line is that the distributed PMDC:

- is a natural compensator for any PMD, as will also be seen in Section V;
- can be operated with a minimum number of control variables, even if the waveguide is unusually inhomogeneous;
- can be expected to solve the 40-Gbit/s PMDC problem once it is fabricated in a commercial fab.

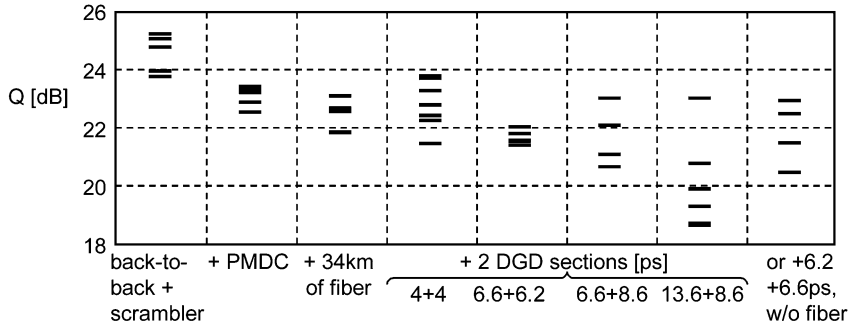

Fig. 18. 40-Gbit/s CSRZ-DPSK transmission with distributed PMD compensation: Q-factors measured for various configurations.

V. Higher Order PMD Definition

A. Taylor Expansion of PMD Vector (TEPV)

The definition of first-order PMD is undisputed. Poole and Wagner have searched those PSPs that will, to first order, not vary as a function of optical frequency at the output of a fiber, and have found that they exhibit a DGD [1]. Alternatively one can determine the polarization-dependent small-signal intensity modulation transfer function of a fiber, and will find maximum and minimum delays, hence, a DGD between orthogonal two polarizations, which are the PSPs [7], [53], [54].

Regarding higher order PMD, almost all previous work on its definition concerns the familiar truncated Taylor expansion of the PMD vector (TEPV). It is easy to calculate the PMD vector derivatives from the Jones matrix. The opposite path has recently been opened by Heismann [55]. The advantages of a higher order PMD definition by the TEPV are the following.

- Easy analytical calculation of higher PMD orders
- Addition/subtraction of second-order PMD to the fiber chromatic dispersion. The same holds for other orders of PMD and chromatic dispersion.
- Relation of third-order PMD to slope steepness differences of NRZ signals.

Its key disadvantages are the following.

- No direct relation to physical fiber parameters exists.
- The true frequency-dependent trajectory of the DGD vector in the Stokes space would be described by sums of sinusoids with arguments that depend linearly on frequency. But sinusoids are not well approximated by a Taylor series. Inevitably, an infinite DGD will be predicted far off the optical carrier frequency.

Second-order PMD is usually classified with respect to the direction of first-order PMD [56]. This is good to discuss statistics. However, like for the third-order PMD discussed in Section III-C, its effect on received eye patterns is better described when it is discussed with respect to the signal polarization: Depending on whether second-order PMD is parallel or antiparallel to the input polarization, it will add and/or subtract to fiber chromatic dispersion.

B. Sequence of DGD Sections (SDGD)

Most PMD simulation is carried out by assuming a sequence of cascaded DGD sections (SDGD), because it is widely accepted that an infinite number of randomly cascaded sections produces "natural" PMD. A finite SDGD can also be used for PMD description [7]. A restricted 3-section variant for effects up to the second order was proposed by Shtaif *et al.* [57]. Möller [58] has pointed out that the structure of a cascade of DGD sections can be obtained by a layer-peeling algorithm [59]. We have implemented this algorithm and have obtained experimental DGD profiles of simple PMD media, including our distributed PMD compensator in several configurations [32]. Important advantages of the SDGD method are the following.

- Easy graphical display of DGD profile.
- Building a PMD emulator with SDGD is much easier than building one for the TEPV with adjustable orders of first- and higher order PMD [60].
- It emulates what a real fiber typically does, which is not evident for TEPV-based higher order PMD emulators.

A disadvantage of the PMD definition by a SDGD is the potentially large number of unknowns. A modification of the scheme might involve choosing a number of DGD sections, for example, all with equal, initially unknown lengths, and then finding their length and orientations that describe the PMD medium best. Obviously a single DGD section describes first-order PMD alone. Two DGD sections describe this, plus the fairly typical interplay between higher orders of the Taylor-approximated PMD vector. The total DGD is constant, which is reminiscent of [56], where it has been shown that second-order PMD parallel to first-order PMD is a lot weaker than that perpendicular to it. The typical slope steepness difference, caused by third-order PMD, can easily be observed even in two DGD sections. More accurate PMD modeling is possible with three or more DGD sections.

While all this makes sense, we discuss in the following a related technique to describe higher order PMD.

C. Fourier Expansion of Mode Coupling (FEMC)

The direction change between adjacent DGD sections is the retardation of a mode converter. Seen by an observer who looks in the direction of the preceding DGD section, this direction change occurs up/down or right/left. This amounts to in-phase or quadrature mode coupling between local PSPs. If the number of DGD sections approaches infinity, mode coupling becomes continuous. We point out that it is possible to describe a PMD

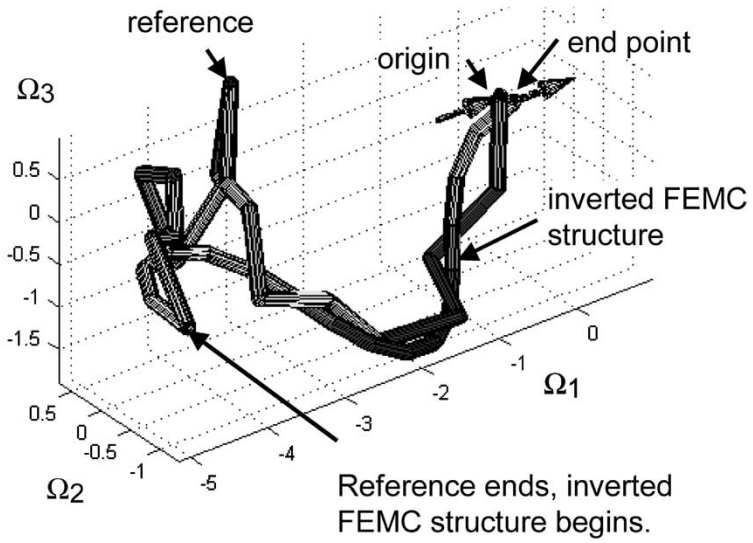

Fig. 19. DGD profile of reference (exemplary PMD medium) cascaded with inverted FEMC structure (which thereby forms a PMD equilizer). Scaling unit is one DGD section length of the reference structure.

medium by a Fourier expansion of mode coupling (FEMC) as follows.

1) A frequency-independent mode conversion at the fiber input. This is described by two parameters, for example, retardation and orientation.
2) A total DGD.
3) A frequency-independent mode conversion at the fiber output. In the general case a mode conversion (two parameters, as at the input) and a differential phase shift (one more parameter) are needed. In total this means that there is a frequency-independent elliptical retarder at the output.
4) Complex Fourier coefficients F_k (7) of mode coupling along the birefringent medium, which exhibits the above-mentioned total DGD only in the absence of mode conversion.

Among these four items, the first three simply describe first-order PMD, and this has been pointed out by many authors. Only the fourth item makes it a FEMC. If there is mode coupling, the DGD profile will bend. Bends at discrete positions would correspond to the SDGD model. Fourier coefficients describe DGD profile bending in a continuous manner (Fig. 19).

As mentioned, the zero-order coefficient F_0 coils the DGD profile. Whether coiling occurs up/down or right/left or in a mix of these cases depends on the phase angle of F_0. The coiling radius is inversely proportional to the magnitude of this coefficient. Other F_k will wind a spiral when they occur alone. F_k combined with F_{-k} can result in a forth-and-back bending of the DGD profile.

The number of real parameters needed to describe PMD by the three mentioned methods is listed in Table II. In all cases three extra parameters must generally be added to specify a frequency-independent elliptical retarder at the fiber output. Only the order 1, corresponding to first-order PMD, is identical for all methods.

The TEPV needs three vector components in the Stokes space for each PMD order. Maximum PMD order covered and method order are equal only here.

TABLE II
ORDER AND NUMBER OF REAL PARAMETERS IN HIGHER ORDER PMD DEFINITION METHODS

Method (below) and its order (right)	1	2	3	4
Taylor expansion of PMD vector (TEPV)	3	6	9	12
Sequence of DGD sections (SDGD)	3	5	7	9
Fourier expansion of mode coupling (FEMC)	3	5	**9**	13

Number of real parameters needed for different orders of a chosen method are given, excluding 3 extra parameters for frequency-independent output polarization transformation. One example of bold-faced case is given in Figs. 19, 20.

The SDGD needs the two parameters of an SBA, plus one total DGD, for the first order. Each additional order is defined by two parameters of an SBA. The method order here means how many DGD sections there are.

In the FEMC, no mode coupling occurs in the first-order PMD case. F_0 adds two real parameters. Each higher order of the method adds two Fourier coefficients $F_{\pm k}$, which amounts to four more real parameters.

In the following we give an FEMC example for method order 3 ($|k| \leq 1$). It needs nine real parameters, like third-order TEPV. A random PMD medium has been taken as a reference. It is composed of 16 DGD sections with equal lengths. The length of one DGD section defines the normalized unit length in Fig. 19. The first-order PMD was 5.1 units, the (first-order) PMD vector was $[-4.98, -1.24, -0.42]^T$. The reference is cascaded with a smoother DGD profile that is an inversion of the FEMC structure. It follows the jagged reference profile with gentle bends and more or less cuts through the "messy" left part of the reference profile. For convenience the FEMC structure was also represented by 16 DGD sections (instead of an infinite number) but this coincidence has no importance because their total DGD is only about half as high as that of the 16 reference DGD sections.

The FEMC coefficients were determined as follows. A Gaussian input pulse was assumed, with a width equal to the

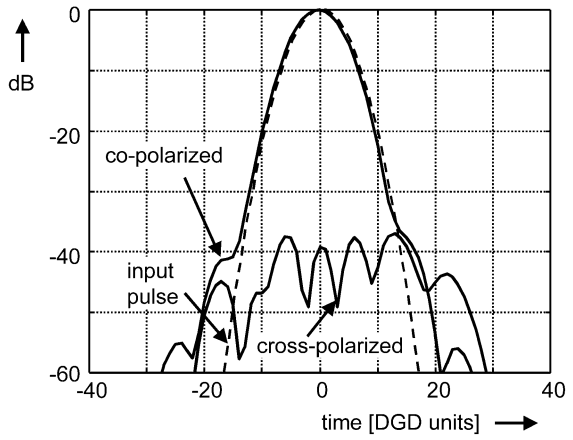

Fig. 20. Magnitudes of Gaussian input pulse and of output pulses resulting from cascaded reference and inverted FEMC structure. Time scale is the same as in Fig. 19. Input pulse width is 8.07 DGD units.

total DGD of the DGD profile used in the FEMC (assuming no mode conversion). This is not the only possible FEMC pulse shape and duration. However, it makes sense to choose the total DGD rather than the first-order DGD because the former is related to the overall complexity of the PMD situation, while the latter may even disappear. Pulse width and the identical total DGD were of course varied during the optimization. The PMD medium (reference) and the inverse of the structure defined by the FEMC were concatenated. The various parameters were adjusted so that the output signal was—as far as possible—in only one (co)polarization mode, and that the impulse in the other (cross-)polarization had its residual maximum amplitude near the time origin, not elsewhere like in the case of first-order PMD. Fig. 20 shows the magnitudes of the electric fields in co- and cross-polarized output pulses. The unwanted polarization is $\geq$37.2 dB down. The Gaussian input pulse is also shown.

As an alternative to the FEMC, the SDGD method could describe the reference profile exactly, but only if it comprised 16 sections equal to those of the reference.

For comparison, the TEPV method was also tested. The TEPV was calculated up to the third order from the Jones matrix of the reference (PMD medium). Then, the Jones matrix corresponding to this truncated TEPV was built as described in [57]. As a third candidate, the EMTY method, an exponential expansion of the Jones matrix by Eyal, Marshall, Tur, and Yariv [61],[62] was tested the same way. For all methods, the input pulse width was chosen identical to that after convergence of the FEMC. Table III shows orthogonal polarization suppression versus method and its order. In the given example, the FEMC holds an advantage over the TEPV ad the EMTY method. The extinction improvement in dB after addition of higher-order terms is several times larger for FEMC than for TEPV and EMTY. In a few more tested PMD examples, the FEMC also held an advantage. This is not surprising because the FEMC and SDGD models are closely related to natural PMD.

However, more work is needed to assess whether (and how much) the FEMC outperforms the TEPV (and the EMTY method) on average. If he answer is yes, then a PMDC that could, for example, compensate the traditional PMD orders 1 to 3, might not be the most efficient equalizer. Rather,

TABLE III
SUPPRESSION OF CROSS POLARIZATION BY EQUALIZERS DEFINED BY HIGHER-ORDER PMD DEFINITION METHODS

Method order	1	2	3
Chosen input pulse width [units]	5.1	5.3	8.1
Taylor expansion of PMD vector (TEPV)	–9.6 dB	–15.7 dB	–15.2 dB
Exponential Jones matrix expansion (EMTY)	–9.6 dB	–11.6 dB	–16.2 dB
Fourier expansion of mode coupling (FEMC)	–9.6 dB	–21.3 dB	–37.2 dB

All methods are equal for order 1. Results refer to example in Fig. 19.

a distributed PMDC could be preferable, with as sharp as possible a polarization transformation at its input, with one more at the output if a defined output polarization is needed, and controlled amounts of mode coupling along its length, for example, defined by Fourier coefficients.

In its realization the distributed PMDC has a fixed total DGD. So its DGD profile must be flexible enough to fold, in order to eliminate any excess DGD.

Note that for the calculations of Table III, three more parameters than shown in Table II were chosen because the "equalizer" needed to be aligned to the reference to separate the polarizations as shown in Fig. 20.

Several variations of the FEMC are conceivable. For example, instead of the Fourier coefficients F_k with orders $k = 0, \pm 1, \pm 2, \ldots$, one could use the orders $k = \pm 1/2, \pm 3/2, \pm 5/2, \ldots$. This would mean that the numbers in the last line of Table II would need to be replaced by 3, 7, 11, and 15. Alternatively, the input and output polarization transformers could be made part of the mode conversion process. This is also practically the case if a distributed PMDC is used for PMD compensation plus output polarization control.

Determining the F_k and other FEMC coefficients presently involves the described numerical minimization process. It may indeed be considered as the most important drawback of this PMD description method that an analytical solution is not known. Finding an easy solution might help in the control of distributed PMDCs—this is a question of not getting trapped in local optima during the PMD control process.

The SDGD method is conceptually similar to FEMC and can be an alternative to FEMC, maybe by keeping the DGD section lengths constant and varying their number.

VI. CONCLUSION

Research on PMD mitigation and compensation issues is still wide open, only with the stringent constraint of cost. Interesting topics are in particular:

- electronic PMD mitigation at 40 Gbit/s;
- higher order PMD detection for alternative signal and modulation formats;
- fabrication of distributed PMD compensators in $LiNbO_3$;
- fast endless polarization control.

Progress on these items is still needed for a fast, accurate, and sufficiently cheap PMD compensation at 40-Gbit/s speeds.

In addition, the lively discussion of PMD definition and description, to which we have contributed a Fourier expansion of mode coupling, is likely to continue.

ACKNOWLEDGMENT

The authors thank F. Abas Ismail, S. Bhandare, S. Chotchidjoum, A. Hidayat, S. Ibrahim, B. Milivojevic, and F. Wüst for their contributions to this paper. They gratefully acknowledge fabrication of distributed $LiNbO_3$ PMDCs by H. Herrmann, R. Ricken, and W. Sohler, Applied Physics, University Paderborn, Germany. They likewise acknowledge partial funding by Deutsche Forschungsgemeinschaft.

REFERENCES

[1] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, no. 19, pp. 1029–1030, May 1986.

[2] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," in *Proc. Nat. Academy Sciences USA*, vol. 97, Sept. 2000, pp. 4541–4550.

[3] H. Sunnerud, M. Karlsson, C. Xie, and P. A. Andrekson, "Polarization-mode dispersion in high-speed fiber-optic transmission systems," *J. Lightwave Technol.*, vol. 20, pp. 2204–2219, Dec. 2002.

[4] M. Yoshimura, T. Kudo, and T. Ozeki, "Polarization mode dispersion equalization," in *OEC*, Japan, pp. 258–259. 14E-12.

[5] T. Takahashi, T. Imai, and M. Aiki, "Automatic compensation technique for timewise fluctuating polarization mode dispersion in in-line amplifier systems," *Electron. Lett.*, vol. 30, no. 4, pp. 348–349, Jan. 1994.

[6] F. Heismann, D. A. Fishman, and D. L. Wilson, "Automatic compensation of first-order polarization mode dispersion in a 10-Gb/s transmission system," in *ECOC 1998*, Madrid, Spain, pp. 529–530. WdC11.

[7] R. Noé, D. Sandel, M. Yoshida-Dierolf, S. Hinz, V. Mirvoda, A. Schöpflin, C. Glingener, E. Gottwald, C. Scheerer, G. Fischer, T. Weyrauch, and W. Haase, "Polarization mode dispersion compensation at 10, 20, and 40 Gb/s with various optical equalizers," *J. Lightwave Technol.*, vol. 17, pp. 1602–1616, Sept. 1999.

[8] H. Sunnerud, C. Xie, M. Karlsson, R. Samuelsson, and P. A. Andrekson, "A comparison between different PMD compensation techniques," *J. Lightwave Technol.*, vol. 20, no. 2, pp. 368–378, 2002.

[9] I. T. Lima Jr., R. Khosravani, P. Ebrahimi, E. Ibragimov, C. R. Menyuk, and A. E. Willner, "Comparison of polarization mode dispersion emulators," *J. Lightwave Technol.*, vol. 19, pp. 1872–1881, Dec. 2001.

[10] PMD Workshop at ECOC-IOOC 2003. [Online]. Available: http://photonics.dei.unipd.it/~galta/

[11] H. Sunnerud, M. Westlund, J. Li, J. Hansryd, M. Karlsson, P.-O. Hedekvist, and P. A. Andrekson, "Long-term 160 Gb/s-TDM, RZ transmission with automatic PMD compensation and system monitoring using an optical sampling system," in *Proc. ECOC 2001*, vol. 6, Amsterdam, the Netherlands, Sept. 30–Oct. 4 2001, pp. 18–19. PD.M.1.9.

[12] N. Gisin, "Polarization effects in optical fibers: Measurement issues," in *Proc. OFC2002*. ThA5.

[13] P. C. Chou, J. M. Fini, and H. A. Haus, "Demonstration of a feed-forward PMD compensation technique," *IEEE Photon. Technol. Lett.*, vol. 14, no. 2, pp. 161–163, Feb. 2002.

[14] J. C. Rasmussen, A. Isomura, and G. Ishikawa, "Automatic compensation of PMD for 40 Gb/s transmission systems," *J. Lightwave Technol.*, vol. 20, no. 12, pp. 2101–2109, 2002.

[15] J. C. Rasmussen, A. Isomura, Y. Akiyama, Y. Qiao, and G. Ishikawa, "Automatic PMD-Compensation at 43 Gbit/s utilizing an accurate DOP-monitor," in *Proc. ECOC*, Copenhagen, Denmark, 2002. 11.1.4.

[16] L. Möller, L. Boivin, S. Chandrasekhar, and L. L. Buhl, "Setup for demonstration of cross channel-induced nonlinear PMD in WDM system," *IEE Electron. Lett.*, vol. 37, no. 5, pp. 306–307, May 2001.

[17] A. Bononi, A. Vannucci, A. Orlandini, E. Corbel, S. Lanne, and S. Bigo, "Degree of polarization degradation due to cross-phase modulation and its impact on polarization-mode dispersion compensators," *J. Lightwave Technol.*, vol. 21, pp. 1903–1913, Sept. 2003.

[18] R. Noé, "Optical amplifier performance in digital optical communication systems," *Elect. Eng.*, vol. 83, pp. 15–20, 2001.

[19] H. Bülow, F. Buchali, W. Baumert, R. Ballentin, and T. Wehren, "PMD mitigation at 10 Gbit-s using linear and nonlinear integrated electronic equaliser circuits," *Electron. Lett.*, vol. 36, no. 2, pp. 163–164, Feb. 2000.

[20] L. Möller, A. Thiede, S. Chandrasekhar, W. Benz, M. Lang, T. Jakobus, and M. Schlechtweg, "ISI mitigation using decision feedback loop demonstrated with PMD distorted 10 Gbit/s signals," *Electron. Lett.*, vol. 35, no. 24, pp. 2092–2093, Nov. 1999.

[21] F. Cariali, F. Martini, R. Chiappa, and R. Ballentin, "Electronic compensation of PMD and chromatic dispersion with an IC in 10 Gbit-s transmission system," *Electron. Lett.*, vol. 36, no. 10, pp. 889–890, May 2000.

[22] H. F. Haunstein, K. Sticht, R. Schlenk, W. Sauer-Greff, M. Lorang, A. Dittrich, and R. Urbansky, "Control of 3-tap electrical feed-forward equalizer by conditional error counts from FEC in the presence of PMD," in *Proc. Optical Fiber Communication Conf. Exhibit 2002 (OFC 2002)*, Mar. 17–22, 2002, pp. 307–308. WQ7.

[23] H. Bülow, "Electronic equalization of transmission impairments," in *Optical Fiber Communication Conf. Exhibit, 2002 (OFC 2002)*, Mar. 17–22, 2002, pp. 24–25. TuE4.

[24] G. S. Kanter, P. Capofreddi, S. Behtash, and A. Gandhi, "Electronic equalization for extending the reach of electro-absorption modulator based transponders," presented at the Optical Fiber Communication Conf. (OFC 2003), Atlanta, GA, Mar. 2003.

[25] L. Möller, Z. Gu, A. Thiede, S. Chandrasekhar, and L. Stulz, "20 Gbps electrical data recovery using decision feedback equaliser supported receiver," *Electron. Lett.*, vol. 39, no. 1, Jan. 2003.

[26] N. Kikuchi, "Analysis of signal degree of polarization degradation used as control signal for optical polarization mode dispersion compensation," *J. Lightwave Technol.*, vol. 19, pp. 480–486, Apr. 2001.

[27] H. Rosenfeldt *et al.*, "Automatic PMD compensation at 40 Gbit/s and 80 Gbit/s using a 3-dimensional DOP evaluation for feedback," in *Proc. OFC2001*, vol. 4, pp. PD27-1–PD27-3.

[28] H. Rosenfeldt, R. Ulrich, U. Feiste, R. Ludwig, H. G. Weber, and A. Ehrhardt, "PMD compensation in 10 Gbit/s NRZ field experiment using polarimetric error signal," *Electron. Lett.*, vol. 36, no. 5, pp. 448–450, Mar. 2000.

[29] L. Möller, P. Westbrook, S. Chandrasekhar, R. Dutta, and S. Wielandy, "SOP and PMD monitoring with WDM polarimeter," *Electron. Lett.*, vol. 38, no. 12, pp. 583–585, June 2002.

[30] J. Peupelmann, E. Krause, A. Bandemer, and C. Schäffer, "Fiber polarimeter based on grating taps," *Electron. Lett.*, vol. 38, no. 21, pp. 1248–1250, Oct. 2002.

[31] L.-S. Yan, Q. Yu, A. B. Sahin, Y. Wang, and A. E. Willner, "Simple bit-rate independent PMD monitoring for WDM systems," in *Proc. ECOC 2001*, Amsterdam, the Netherlands, Sept. 30–Oct. 4 2001. TU.A.3.2.

[32] D. Sandel, V. Mirvoda, S. Bhandare, F. Wüst, and R. Noé, "Some enabling techniques for polarization mode dispersion compensation," *J. Lightwave Technol.*, vol. 21, pp. 1198–1210, May 2003.

[33] R. Noé, H. Heidrich, and D. Hoffman, "Endless polarization control systems for coherent optics," *J. Lightwave Technol.*, vol. 6, pp. 1199–1207, July 1988.

[34] R. Noé, German Patent Application DE 100 50 266.0, Oct. 9, 2000.

[35] R. Noé, D. Sandel, V. Mirvoda, F. Wüst, and S. Hinz, "Polarization mode dispersion detected by arrival time measurement of polarization-scrambled light," *J. Lightwave Technol.*, vol. 20, pp. 229–235, Feb. 2002.

[36] D. Sandel, F. Wüst, V. Mirvoda, and R. Noé, "Standard (NRZ 1×40 Gbit/s, 210 km) and polarization multiplex (CS-RZ, 2×40 Gbit/s, 212 km) transmissions with PMD compensation," *IEEE Photonics Technol. Lett.*, vol. 14, Aug. 2002.

[37] R. Noé. Combatting and equalizing the effects of PMD in 40 Gb/s systems and beyond. presented at *Proc. 28th Eur. Conf. Optical Communication (ECOC 2002)*. [Online]. Available: http://ont.upb.de/publikationen/ecoc2002_noe_tut_add.pdf

[38] E. Buchali, W. Baumert, H. Bülow, and J. Poirrier, "A 40 Gb/s eye moitor and its application to adaptive PMD compensation," in *Proc. OFC 2002*, pp. 202–203. WE6.

[39] C. Francia, F. Bruyere, D. Penninckx, and M. Chbat, "PMD second-order effects on pulse propagation in single-mode optical fibers," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1739–1741, Dec. 1998.

[40] M. Amemiya, "Pulse broadening due to higher order dispersion and its transmission limit," *J. Lightwave Technol.*, vol. 20, pp. 591–597, Apr. 2002.

[41] F. Wüst, D. Sandel, V. Mirvoda, and R. Noé, "Electrical slope steepness difference indicates higher-order PMD at 40 Gbit/s," presented at the Optical Fiber Communication Conf. (OFC 2003), Atlanta, GA, Mar. 2003.

[42] R. Noé, E. Meissner, B. Borchert, and H. Rodler, "Direct modulation 565 Mb/s DPSK experiment with 62.3 dB loss span and endless polarization control," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 1151–1154, Oct. 1992.
[43] N. G. Walker and G. R. Walker, "Polarization control for coherent optical communications," *J. Lightwave Technol.*, vol. 8, no. 3, pp. 438–458, 1990.
[44] F. Heismann and M. S. Whalen, "Broadband reset-free automatic polarization controller," *Electron. Lett.*, vol. 27, no. 4, pp. 377–379, Feb. 1991.
[45] J. Nayyer *et al.*, "Fabrication of high-speed asymmetric Ti:$LiNbO_3$ Mach-Zehnder optical modulatiors with improved thermal-drift characteristics," in *Proc. IOOC'95*, vol. 2, Hong Kong, June 26–30, 1995. WD1-5.
[46] R. Noé, D. Sandel, S. Hinz, M. Yoshida-Dierolf, V. Mirvoda, G. Feise, H. Herrmann, R. Ricken, W. Sohler, F. Wehrmann, C. Glingener, A. Schöpflin, A. Färbert, and G. Fischer, "Integrated optical LiNbO3 distributed polarization mode dispersion equalizer in 20 Gbit/s transmission system," *Electron. Lett.*, vol. 35, no. 8, pp. 652–654, Apr. 1999.
[47] D. Sandel, S. Hinz, M. Yoshida-Dierolf, V. Mirvoda, R. Noé, G. Feise, H. Herrmann, R. Ricken, W. Sohler, H. Suche, F. Wehrmann, and R. Wessel, "Integrated-optical polarization mode dispersion compensation for 6-ps, 40-Gb/s pulses," in *Proc. ECIO'99*, vol. PDP, Turin, Italy, Apr. 14–16, 1999, pp. 17–19.
[48] R. Noé and D. Sandel, "Reduced number of control variables for fast control and zero-mean voltages for DC drift suppression in distributed $LiNbO_3$-based PMD compensators," presented at the Eur. Conf. Integrated Optics (ECIO 2003), vol. 1, Prague, Czech Republic, Apr. 2–4, 2003, pp. 87–90. WeA3.4.
[49] D. Sandel and R. Noé, "Truly endless polarization control with I&Q mode converters in X-cut, Y-propagation lithium niobate," in *Proc. ECOC-IOOC 2003*, Rimini, Italy. Mo 4.5.4.
[50] F. Heismann and R. Ulrich, "Integrated-optical single-sideband modulator and phase shifter," *IEEE J. Quantum Electron.*, vol. QE-18, pp. 767–771, Apr. 1982.
[51] R. Noé, S. Hinz, D. Sandel, and F. Wüst, "Crosstalk detection schemes for polarization division multiplex transmission," *IEEE J. Lightwave Technol.*, vol. 19, pp. 1469–1475, Sept. 2000.
[52] B. Milivojevic, D. Sandel, S. Bhandare, R. Noé, and F. Wüst, "40 Gbit/s CSRZ-DPSK transmission system with signed online chromatic dispersion detection," *Electron. Lett.*, vol. 39, no. 20, pp. 1455–1456, Oct. 2003.
[53] L. E. Nelson, R. M. Jopson, H. Kogelnik, and J. P. Gordon, "Measurement of polarization mode dispersion vectors using the polarization dependent signal delay method," *Opt. Expr.*, vol. 6, no. 8, pp. 158–167, Apr. 2000.
[54] A. Eyal, D. Kuperman, O. Dimenstein, and M. Tur, "Polarization dependence of the intensity modulation transfer function of an optical system with PMD and PDL," *IEEE Photonics Technol. Lett.*, vol. 14, pp. 1515–1517, Nov. 2002.
[55] F. Heismann, "Jones matrix expansion for second-order polarization mode dispersion," in *Proc. ECOC-IOOC 2003*. Th1.7.5.
[56] G. J. Foschini, R. M. Jopson, L. E. Nelson, and H. Kogelnik, "Statistics of polarization dependent chromatic fiber dispersion due to PMD," in *Proc. ECOC 1999*, vol. II, Nice, France, Sept. 26–30, 1999, pp. 56–59.
[57] M. Shtaif, A. Mecozzi, M. Tur, and J. A. Nagel, "A compensator for the effects of high-order polarization mode dispersion in optical fibers," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 434–436, Sept. 2000.
[58] L. Möller, "Filter synthesis for broad-band PMD compensation in WDM systems," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 1258–1260, Sept. 2000.
[59] S. E. Harris, E. O. Ammann, and I. C. Chang, "Optical network synthesis using birefringent crystals. I. Synthesis of lossless networks of equal-length crystals," *J. Opt. Soc. Amer.*, vol. 54, no. 10, pp. 1267–1279, Oct. 1964.
[60] H. Kogelnik, L. E. Nelson, and J. P. Gordon, "Emulation and inversion of polarization-mode dispersion," *J. Lightwave Technol.*, vol. 21, pp. 482–495, Feb. 2003.
[61] A. Eyal, W. K. Marshall, M. Tur, and A. Yariv, "Representation of second-order polarization mode dispersion," *Electron. Lett.*, vol. 35, no. 19, pp. 1658–1659, 1999.
[62] A. Eyal, Y. Li, W. K. Marshall, and A. Yariv, "Statistical determination of the length dependence of high-order polarization mode dispersion," *Opt. Lett.*, vol. 25, no. 12, pp. 875–877, 2000.

Reinhold Noé (M'93) was born in Darmstadt, Germany, in 1960. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Technische Universität München (Munich), Germany, in 1984 and 1987, respectively.

At that time he realized the first endless polarization control systems. Then he spent a postdoctoral year at Bellcore, Red Bank, NJ, to continue his work on coherent optical systems. In 1988 he joined Siemens Research Laboratories in Munich. Since 1992 he has been with the Department of Optical Communication and High-Frequency Engineering, University of Paderborn, Paderborn, Germany. Most of his recent experiments deal with PMD detection and compensation, polarization-division multiplex, chromatic dispersion detection, and DPSK transmission, all at 40 Gbit/s. He has (co)authored more than 120 publications. He is one of the Editors of *Electrical Engineering* and he is a frequent referee for *Electronics Letters*.

Prof. Noé is a frequent referee for IEEE publications.

David Sandel was born in Wangen/Allgäu, Germany, on December 26, 1966. He received the Dipl.-Ing degree from the Technische Hochschule Karlsruhe, Germany, in 1992 and the Dr.-Ing. degree from the University of Paderborn, Paderborn, Germany, in 1997.

In the same year, he joined the Department of Optical Communication and High-Frequency Engineering, University of Paderborn, Germany. His doctoral work on fiber Bragg gratings. His research interest now concentrates on polarization-mode dispersion compensation and polarization multiplex data transmission.

Vitali Mirvoda was born in Taganrog, Russia, in 1974. He received the M.S. degree in electrical engineering from Taganrog State University of Radio Engineering, Taganrog, Russia, in 1996. He is currently pursuing the Ph.D. degree in electrical engineering and information technology at the University of Paderborn, Paderborn, Germany.

Since 1998 he has been engaged in research and engineering of optical communication systems at the University of Paderborn. He is currently a postgraduate student with the Department of Optical Communication and High-Frequency Engineering. His research areas include polarization-mode dispersion compensation and chromatic dispersion.