Multi Impairment Monitoring for Optical Networks

Trevor B. Anderson, Adam Kowalczyk, Ken Clarke, Sarah D. Dods, *Senior Member, IEEE*, Don Hewitt, *Member, IEEE*, and Jonathan C. Li, *Student Member, IEEE*

***Abstract*—As optical networks become more complex, the need for in-line monitoring of more than just channel wavelength, power and OSNR becomes compelling. In this paper we describe an asynchronous delay tap sampling technique coupled with statistical machine learning that enables a single monitor to measure multiple simultaneous impairments on multiple formats. We demonstrate the technique for simultaneous measures of CD and 1st order PMD on a 40 Gbit/s NRZ-DPSK signal.**

***Index Terms*—Asynchronous sampling, multi impairment monitoring, optical performance monitoring, polarization mode dispersion (PMD).**

I. Introduction

RECONFIGURABLE optical networks offer the potential for significant operational savings through automated path provisioning, enhanced fault management and optimization [1]. However, realizing this potential requires a real time picture of the optical impairments inside the network, and their distribution. In addition to OSNR, these impairments include PMD, four wave mixing, chromatic dispersion and slope, reflections, laser noise, and both inter-channel and interferometric crosstalk [2]. Q-factor and timing jitter are also useful as measures of the effects of these impairments on signal quality. In addition to fault diagnosis, monitoring is required for feedback to tunable compensation elements and signal quality assurances for networks carrying alien wavelengths.

Many techniques have been proposed to measure different subsets of these impairments. These techniques can be broadly classed as either spectrally- [3], [4] or sampling-based. The former include the use RF tones and measurement of RF clock power. These techniques are, however, format dependent. The laboratory eye diagram is the most familiar sampling technique for measuring signal quality and estimating the underlying causes [5], but requires clock recovery, is format and bit-rate dependent, and can be difficult to extract for strongly distorted signals.

Manuscript received February 21, 2009; revised May 21, 2009 and June 01, 2009. First published June 10, 2009; current version published August 07, 2009

T. B. Anderson, A. Kowalczyk, and J. C. Li are with Monitoring Division Inc., Level 1, East Melbourne, VIC 3002, Australia. They are also with National ICT Australia Victorian Research Laboratory, University of Melbourne, Parkville, VIC 3052, Australia (e-mail: trevor@monitoringdivision; adam@monitoringdivision; jonathanan.li@monitoringdivision).

K. Clarke and S. D. Dods are with Monitoring Division Inc., Level 1, East Melbourne, VIC 3002, Australia (e-mail: ken@monitoringdivision; sarah@monitoringdivision).

D. Hewitt is with National ICT Australia Victorian Research Laboratory, University of Melbourne, Parkville, VIC 3052, Australia (e-mail: don.hewitt@nicta.com.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2009.2025052

The asynchronous histogram technique has been proposed as an alternative sampling technique that does not require clock recovery [6]–[10]. Whilst this technique demonstrates sensitivity to multiple impairments, uniquely identifying a particular impairment in the presence of other impairments remains a challenge.

We have recently introduced a new sampling technique, known as "asynchronous delay tap sampling" in which multiple impairment measurements as well as signal quality are extracted from a two dimensional histogram of the signal [11], [12]. This histogram, known as a phase portrait, provides the information richness of an eye diagram without the requirement of clock extraction. The phase portraits contain unique impairment signatures that can be discovered using statistical pattern recognition techniques. With this approach we are able to not just classify impairments, but to quantitatively monitor simultaneous combinations.

Simulation results of 10 Gbit/s NRZ have shown the potential for simultaneous monitoring of OSNR, CD, DGD, crosstalk and optical filter detuning as well the ability to measure signal quality parameters Q and timing jitter [4]. The first experimental results (using a training set of 300 cases) were demonstrated for simultaneous combinations of OSNR, CD and DGD with RMS errors of 0.6 dB, 100 ps/nm and 3.1 ps respectively [12]. An important limitation of these earlier works was that the signal polarization was such that 1st order PMD had a fixed and equal power split between principal states. In recent work [13] the polarization restriction has been removed, with the first demonstration of CD (RMS error of 17 ps/nm) and 1st order PMD with random power splits on a commercial 10 Gbit/s NRZ WDM test channel. This work used an automated network emulator that enabled larger training sets (900 cases) to be created.

An advantage of the technique is that a simple direct detection receiver can be used to monitor a variety of bit rates and modulation formats without the need for demodulation of the signal or modification of the receiver bandwidth. For example the amplitude variations in both RZ-DPSK and RZ-DQPSK signals have been used to monitor OSNR and CD [14], [15] with delay tap sampling.

We note that monitoring of digital coherent systems using the information available in the digital equalizer at the receiver has recently been demonstrated for CD, PMD and OSNR and PDL [16], [17]. Whilst these techniques provide a cost effective solution at the receiver they cannot be readily used at intermediate points in the network. In contrast the potential for a distributed solution based on a simple direct detection receiver is a key attribute of the asynchronous delay tap technique.

In this paper we present a review of our technique and demonstrate its ability to simultaneously measure CD and PMD on a 40 Gbit/s NRZ-DPSK format.

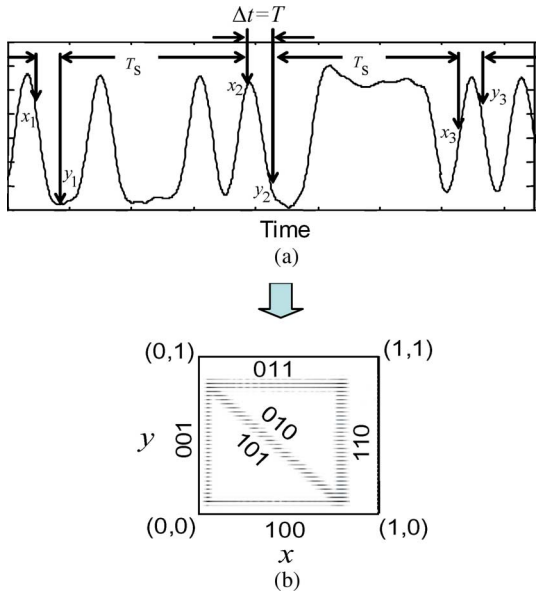

Fig. 1. Processing of delay tap sample pairs to create phase portraits. The labels on the phase portrait represent the sampled bit sequences.

The paper is organized as follows. In Section II, we describe the generation and interpretation of phase portraits and the use of pattern recognition techniques to identify impairments. We introduce a measure of DGD which takes into account arbitrary power splits between principal polarization states. Simulation results for a 40 Gbit/s NRZ-DPSK system are presented in Section III and in Section IV we describe the experimental implementation of our system which includes both the generation of training sets and the implementation of the monitor with results and discussion. We conclude the paper with Section IV.

II. Technique

There are two key components to the technique which we describe below. The first is the construction of the phase portrait. The second is the application of statistical pattern recognition techniques to extract impairment signatures from the phase portraits.

A. Phase Portrait[1]

Asynchronous delay-tap sampling is an alternative to the eye diagram that uses the joint probability density function (pdf) of a signal $x(t)$, and its delayed version $x(t+\Delta t)$ to characterize the signal. This pdf, known as a phase portrait, is sensitive to waveform distortion and noise and contains unique signatures of impairments. To generate the phase portrait the waveform is sampled in pairs separated by a known delay Δt, as shown for the NRZ signal in Fig. 1(a). The phase portrait is then created by binning the sample pairs into a two dimensional histogram as shown in Fig. 1(b), for a one-bit delay, $\Delta \mathrm{t} = \mathrm{T}$. We emphasize that the sampling is asynchronous, in that the time between the pairs, $\mathrm{T_s}$, is not related to the monitored signal bit rate, and can be many orders of magnitude longer.

A fundamental difference between the eye diagram and phase portrait is that the latter contains information about the correlation between samples on the time scale of a bit period. Alternatively for short time delays, for example 0.25 T, we can interpret the phase portrait as containing joint probability density information of the amplitude and slope of the waveform with respect to time. This information is absent in eye diagrams that are constructed from samples that are separated by long periods.

In Fig. 1(b), the corners of the phase portrait represent the sample pairs of well defined marks and spaces. The lines joining these points reflect the bit transitions between these states. For example the diagonal line represents the 010 and 101 transition. As we will show below, waveform distortions in the time domain manifest as variations in curvature and density of these lines.

We note that whilst the choice of 1 bit delay Δt is a useful tool for visualisation, it is not critical for pattern recognition. This enables a fixed delay of 25 ps for example, to be used for both 10 and 40 Gbit/s signals. To help understand the effects of different impairments on the phase portrait, we have simulated optical signal to noise ratio (OSNR), chromatic dispersion (CD), polarization mode dispersion (PMD) and interferometric crosstalk on a 10 Gbit/s NRZ signal. The resulting phase portraits and eye diagrams are shown Fig. 2. In all cases the tap delay for the phase portrait was chosen to be 1 bit period. Fig. 2(a) shows the results for no optical impairment (OSNR 35 dB), with a clean eye and a well defined geometric shape in the phase portrait. Fig. 2(b) shows the effects of reducing the OSNR to 25 dB. The predominant effect is to broaden the high power regions of both the eye and phase portrait, but the underlying geometric shapes are not affected. Fig. 2(c) shows OSNR of 35 dB with 800 ps/nm of CD, corresponding to 50 km of SMF. The eye diagram shows the characteristic narrowing of the peaks. In the phase portrait, the dispersion causes the diagonal to curve in towards the origin.

Fig. 2(d) shows an OSNR of 35 dB with 30 ps of first order PMD, with the power split equally between the principal polarization axes. The eye shows the characteristic 'triangularization', but still looks very clean. Interestingly at a first glance, the phase portrait does not show the effects of PMD. However a closer inspection shows significant differences in the distribution of points along the lines representing the 3 bit transitions. Fig. 2(e) shows OSNR of 35 dB with a single source of interferometric crosstalk at -25 dB. Both the eye and two-tap plots show similar broadening to the OSNR degradation, but different noise statistics. Finally Fig. 2(f) shows OSNR of 25 dB with the combined degradations from Figs. 2(a)–(f).

These initial results suggested that the phase portraits contained impairment signatures that could be exploited to enable us to distinguish between a variety of optical impairments. As we will show in the next section, this is indeed the case, in particular there exist features that enable us to identify and measure the individual impairments even in the cases where they occur simultaneously as in Fig. 2(f). However such a determination of individual impairment values in the case of mixed impairments requires more sophisticated approaches than hand-crafted measurement of image features. For instance,

[1]Unbeknown to the authors in the early development of this monitoring technique, delay tap sampling is used in the field of non-linear time series analysis where it is referred to as "time delay embedding". It is also where the term "phase portrait" originates and in this context the term "phase" bears no relation to optical phase.

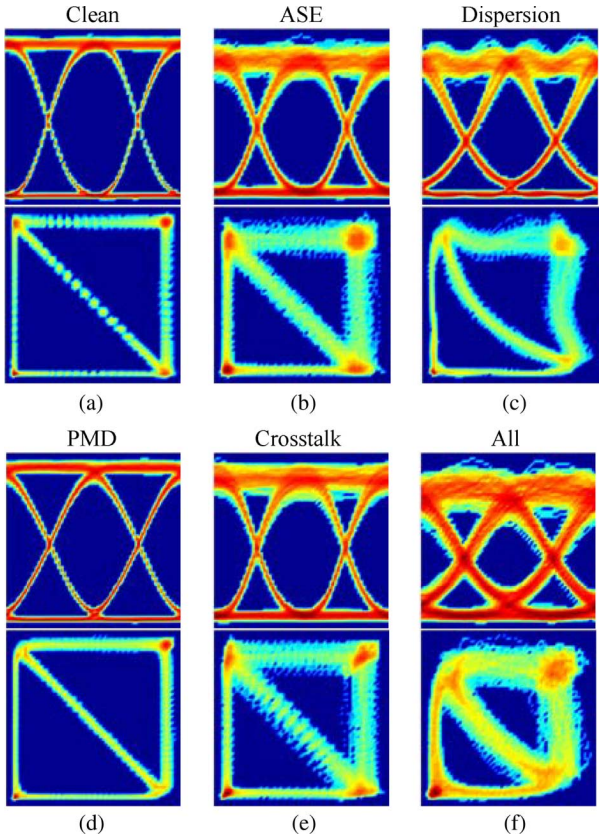

Fig. 2. Eye diagrams and phase portraits for NRZ, 1 bit delay (a) $\mathrm{OSNR} = 35$ dB and no impairment, (b) $\mathrm{OSNR} = 25$ dB, (c) $\mathrm{OSNR} = 35$ dB and $\mathrm{D} = 800\,\mathrm{ps/nm}$, (d) $\mathrm{OSNR} = 35$ dB and $\mathrm{PMD} = 40$ ps, (e) $\mathrm{OSNR} = 35$ dB and $\mathrm{crosstalk} = -25$ dB (f) $\mathrm{OSNR} = 25$ dB, $\mathrm{D} = 800\,\mathrm{ps/nm}$, $\mathrm{PMD} = 40$ ps, and $\mathrm{crosstalk} = -25$ dB.

in Fig. 2(f), the curvature of the diagonal line is obscured by ASE noise, and reduced by PMD. To address this issue in a principled manner, we used statistical learning techniques to automate feature selection.

Finally we note that the asynchronous sampling technique is applicable to a variety of signal formats [14], [15], [18]. For example, in Fig. 3 we show experimentally measured phase portraits for 10 Gbit/s RZ showing the effects of CD. The clear differences between the RZ and NRZ (from Fig. 2) portraits suggest that the technique may also be used to identify signal formats.

B. Pattern Recognition

Phase portraits can be treated as images, with the pixel intensity representing the number of hits in each bin. These pictures lend themselves well to pattern recognition techniques [5], [20]. Since, in this case, the impairment values are continuous variables; we used kernel based ridge regression [19], [20], rather than using classification into small discrete categories, as in [5] or hand-written digit recognition [20]. We have found, however, that a range of estimators, corresponding to various pre-processing of the data and kernels used, produced acceptable results.

As stated above, the first step in the training process is binning the sample pairs into a two dimensional histogram. For the

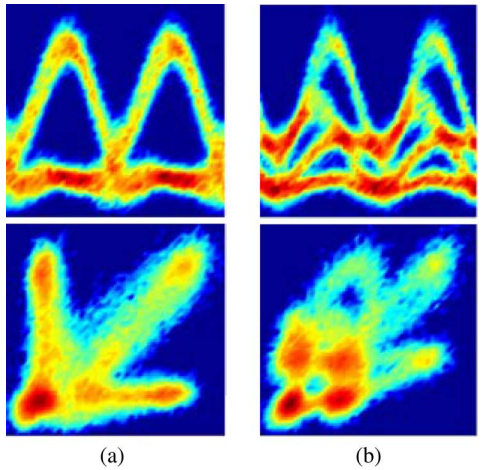

Fig. 3. Experimental eye diagram and phase portrait for RZ, 1 bit delay (a) clean signal (b) 850 ps/nm CD.

results shown in this work we use a 30 by 30 histogram. Features can then be extracted from these histograms to form the basis of the training sets. (For example in the approach adopted in [21] only six features are selected which represent the mean and variance of the bins in each of three quadrants). Alternatively, as in the case described here all 900 features are available for training.

The predictor for each impairment can, in general, be assumed to be a weighted non-linear combination of features. For example the predictor for abs(CD) can be written as

$$|CD| = \sum_{j=1}^{n} \alpha_j k(\vec{x}_j, \vec{x})$$

where k is in general a non linear kernel function of the phase portrait feature vector $\vec{x}$ to be measured and the n phase portraits in the training set. The training process determines the weights α_j which jointly minimises the sum of the squared errors and a regularisation term, the latter of which is used to avoid over fitting [20]. In practice, the training sets are created using a "network emulator" to add known quantities of the impairments to a clean signal.

Training is done independently for each impairment by minimising the impairment prediction error (CD for example) in the presence of the other background impairments (OSNR and DGD). This enables the CD predictor to be valid over the DGD and OSNR range included in training. For impairment levels outside this range the CD will be expected to suffer from OSNR and DGD dependence. We point out that the CD prediction does not require simultaneous DGD or OSNR predictions to be made as the predictors are independently trained.

It is currently assumed that monitor has access to knowledge of the formats and bit rates being monitored from the network management system and can select appropriate predictors from an on-board library. The requirements for training over different forward error correction (FEC) rates (eg 10.0 and 10.7 Gbit/s) can be met by either using a variable delay that is tuned to the bit period or by using a fixed delay and combining training sets from varying bit rates to produce a predictor. With this latter approach a single predictor could be used to monitor different

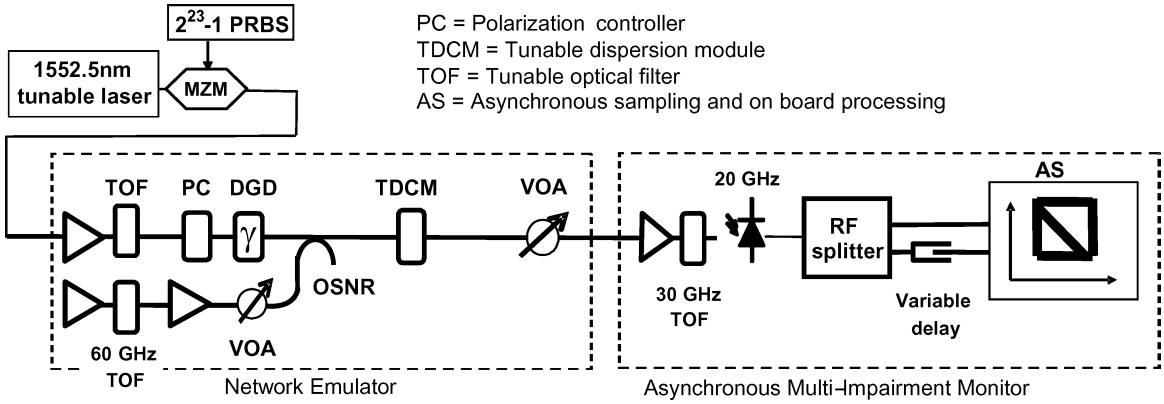

Fig. 4. Setup for multi-impairment monitoring using delay-tap asynchronous sampling. Network emulator adds known combinations of CD, 1st order PMD (DGD and γ) and OSNR to 40 Gbit/s NRZ-DPSK signal for generation of training and test cases. The polarization alignment of the signal with DGD emulator is varied between each measurement to ensure random γ.

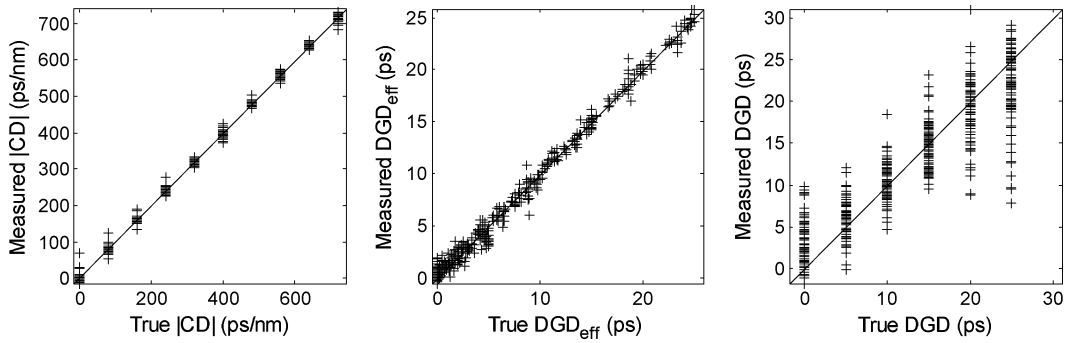

Fig. 5. Simulation results for 40 Gbit/s NRZ DPSK simultaneous measurements of $|\mathrm{CD}|$ (RMS error of 11 ps/nm), $\mathrm{DGD_{eff}}$ (RMS error of 0.75 ps) and DGD. The background OSNR ranged from 13 to 25 dB. 500 random test cases were selected. The power split between principle polarization states γ was assumed to be random. We have included the DGD results to illustrate advantage of using $\mathrm{DGD_{eff}}$ which includes the effect of γ.

FEC rates without recourse to the network management system. In principle, to make the technique more robust, one can extend the training set to include other effects which can potentially induce errors; these may include filter drift, delay variations or transponder variations. The challenge in this case is to select features that help prevent the training set from becoming too large.

C. 1st Order PMD

In our first demonstration of the technique [11], [12] the impairments included were OSNR, CD and DGD where the latter was restricted to worst case polarization alignment of the signal and the principal states. More recently we have demonstrated the technique at 10 Gbit/s taking into account both the differential group delay (DGD) and a random power split γ between principal states. To do this we define an effective differential group delay $\mathrm{DGD_{eff}} = 4\gamma(1-\gamma)\mathrm{DGD}$ that is proportional to first-order string length and is a measure of first-order PMD system penalty [22], [23]. When the signal is aligned with either principal state of polarization (PSP) ($\gamma = 0$ or 1) $\mathrm{DGD_{eff}} = 0$ and the signal remains undistorted regardless of the level of first-order PMD. The factor of 4 in $\mathrm{DGD_{eff}}$ is chosen so that when $\gamma = 1/2$, $\mathrm{DGD_{eff}} = \mathrm{DGD}$. The advantages of using $\mathrm{DGD_{eff}}$ are that it is directly related to the induced signal distortion and thus the 1st order PMD-induced system penalty, and it provides a dynamic measure for feedback for PMD compensation.

III. Simulations

In order to gauge the accuracy of the technique we present simulation results for a 40 Gbit/s NRZ-DPSK signal. A schematic of the setup of the network emulator and the multi impairment monitor used for both simulation and experiments is shown in Fig. 4. The key components of the monitor are a 30 GHz optical filter and a 20 GHz bandwidth receiver followed by a 1 bit tap delay. A set of 1700 phase portraits were produced by simulating combinations of CD from 0 to 700 ps/nm (70 ps/nm steps), DGD from 0 to 20 ps (5 ps steps), γ from 0 to 1 (random) and OSNR from 13 to 26 dB (2 dB steps). 1200 cases were randomly selected for training with the remaining 500 for testing.

Fig. 5 shows scatter plots of the predicted CD, $\mathrm{DGD_{eff}}$, and DGD values against the true impairment values. The RMS error for CD is ± 11 ps/nm over a range of 0 to 700 ps/nm and is ± 0.75 ps for $\mathrm{DGD_{eff}}$ over 0 to 25 ps. We have included the poor DGD result to highlight the advantage of the $\mathrm{DGD_{eff}}$ measure.

The large DGD error reflects the difficulty in separating the combined effects of DGD and γ on the phase portrait. For the cases where $\gamma = 0$ or 1 the phase portrait is independent of DGD and the DGD is therefore impossible to predict. In the more general case, the large variation in the distortion of phase portraits with identical DGD and varying γ poses a challenge for the sensitivity of this technique. In principle one may have expected an asymmetry in the phase portrait (about the line x =

y) to enable some discriminating ability for γ and hence DGD in the same way that one expects an asymmetry in the leading and trailing edges of eye diagram with varying power in the fast and slow principle states. Unfortunately however, intrinsic asymmetry in transponder characteristics have, to date, masked these effects.

A final comment on the DGD results concerns the distribution of errors. In particular, we note a positive bias in the predictions for 0 DGD. In the regression models used in this paper equal weighting is given to all training cases and DGD levels, a consequence of this is that in cases where there is little discrimination, (ie $\mathrm{DGD} = 0$ ps and $\mathrm{DGD} > 0$ ps, $\gamma = 0, 1$), the minimization of the overall error leads to positive bias at low DGD and negative bias at high DGD. By increasing the weight given to low DGD training cases we could potentially improve the predictor for low DGD. Despite the negative bias at high DGD's there are still cases for which the prediction is higher than the true DGD; these errors are due to the random variations in the number of hits in each bin and can be reduced by taking a larger number of samples or averaging over multiple measurements.

IV. Experiment

Fig. 4 shows a schematic of the experimental setup. The channel was accessed via a 1% monitoring port at the start of the link. The tap power level was around -20 dBm per channel. The OSNR at the tap point was measured with an OSA to exceed 25 dB, and the PMD and CD were assumed to be negligible. The tapped signal was connected to a custom built network emulation unit, which filtered the channel to be monitored and added known combinations of impairments for the training phase. The signal was then passed to the delay-tap Multi-Impairment Monitor, which used delay tap sampling to create and analyze the phase portraits, to provide performance monitoring. Control of the Network Emulator and the data processing of the Multi-Impairment Monitor were done via an external laptop.

A. Network Emulator

In the Network Emulator, the 40 Gbit/s signal was generated with a tunable DFB modulated with a dual drive Mach Zehnder and $2^{23} - 1$ PRBS. The signal was amplified with an EDFA to overcome emulator loss, and filtered with a tunable 100 GHz filter. A polarization controller was used to randomly select a polarization state and ensure a random distribution of γ. Various known combinations of DGD and CD were then added to the monitored signal. The DGD emulator had a range of 0 to 22.5 ps and true $\gamma = \mathrm{I_{fast}}/(\mathrm{I_{fast}} + \mathrm{I_{slow}})$ was derived from DC optical power measurements (averaged over the sampling time) using taps in the fast and slow axes of the DGD emulator. We note that the polarization controller was programmed to step through random polarization states and was held constant during the sampling of each phase portrait.

The OSNR was controlled by coupling in a variable amount of ASE, generated by filtering and amplifying the output of an EDFA. Varying levels of OSNR were used to ensure that the CD and PMD predictions were valid across normal operational ranges. The ASE source was placed after the PMD emulator to ensure that the measurement of γ was not compromised by the added ASE noise. CD was controlled using a tunable dispersion compensation module with a range of -400 to $+400$ ps/nm. The impaired signal was then passed to the MIM for sampling and analysis.

B. Multi Impairment Monitor

In the monitor, the signal was amplified to deliver a constant power of 0 dBm to the photodiode, after then filtering with a tunable optical filter with a 3 dB bandwidth of 30 GHz. The signal was then fed to a 20 GHz receiver followed by a 50:50 splitter with a tunable electrical delay in one arm. The signal and delayed ports were asynchronously sampled at 40 kS/s using a customized dual channel analogue to digital converter with a bandwidth of 20 GHz. For every Network Emulator setting, a phase portrait was generated from 40,000 sample pairs. Training and test sets were created by programming the emulator to step through combinations of CD from -400 to 400 ps/nm (20 ps/nm steps) and DGD from 0 to 22.5 ps (2.5 then 4 ps steps), while varying the OSNR between 15 and 25 dB (8 levels) giving a total of approximately 2000 phase portraits. We note that the polarization state and hence γ varied randomly across this phase portrait data set. In practice, allowing for emulator tuning set up times, the training time is approximately 3 hrs.

C. Data Processing

The set of 2000 phase portraits was then randomly divided, with 1500 becoming the training set. The training set was used to generate independent regression models for both CD and $\mathrm{DGD_{eff}}$. The accuracy of the resulting CD and $\mathrm{DGD_{eff}}$ models were then tested on the remaining 500 cases.

D. Results and Discussion

Fig. 6 shows a random selection of phase portraits taken from the training set. We note that the optical filter bandwidth of 30 GHz partially demodulates the signal. The effects of CD are seen in the loop structure of (a) and (b) and OSNR changes in the closure of (c) and (d). It is difficult to pick systematic DGD changes by eye.

Fig. 6 shows scatter plots of the predicted $|\mathrm{CD}|$ and $\mathrm{DGD_{eff}}$ results against the true impairment values. The RMS error for $|\mathrm{CD}|$ is ± 11 ps/nm over a range of 0 to 400 ps/nm and is ± 1.9 for $\mathrm{DGD_{eff}}$ over 0 to 22.5 ps. We emphasize these results are taken over all combinations of impairments. The measurement ranges are restricted by the Network Emulator, and do not represent an intrinsic limitation of the technique. The CD results are consistent with simulation results and show good accuracy across all values. A breakdown of the accuracy of $|\mathrm{CD}|$ as a function of $\mathrm{DGD_{eff}}$ and OSNR is given in Table I(a). As expected the CD error degrades with increasing $\mathrm{DGD_{eff}}$ and low OSNR. We expect that closer agreement between the experimental and simulation results for $\mathrm{DGD_{eff}}$ could be achieved with an improved measurement of γ. A breakdown of the accuracy of the $\mathrm{DGD_{eff}}$ results is given in Table I(b). As before the accuracy degrades with poor OSNR but surprisingly improves with increasing CD. A possible interpretation of this is that the CD induced phase to amplitude conversion provides a larger waveform (Fig. 6(a)) for $\mathrm{DGD_{eff}}$ to distort.

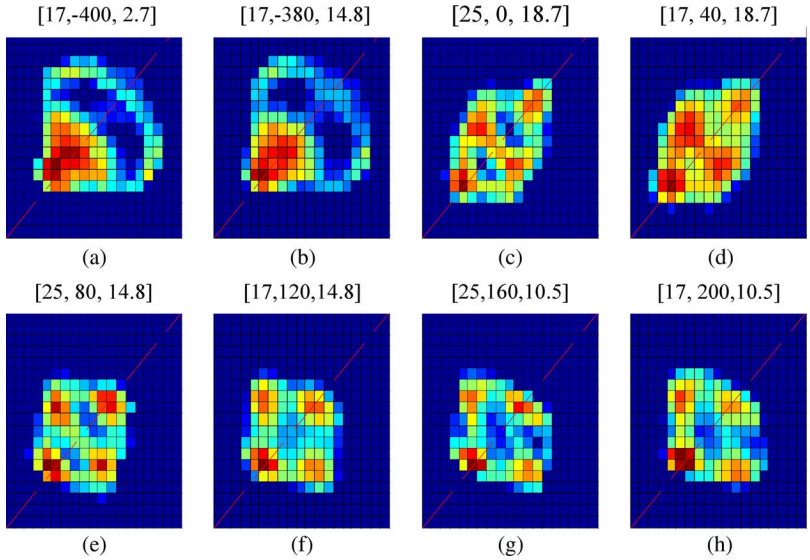

Fig. 6. Random selection of experimental 40 Gbit/s NRZ-DPSK 1 bit delay phase portraits taken from training set. We note that the optical filter bandwidth of 30 GHz partially demodulates the signal. Impairments are labeled as [OSNR (dB), CD (ps/nm), $\mathrm{DGD}_{\mathrm{eff}}$ (ps)]. The effects of CD are seen in the loop structure of (a) and (b) and OSNR changes in the closure of (c) and (d). It is difficult to pick systematic DGD changes by eye.

TABLE I
CONDITIONAL ERRORS FOR 40 Gbit/s NRZ-DPSK

Background $\mathrm{DGD}_{\mathrm{eff}}$ (ps)	Background OSNR (dB) 15	16-20	20-25
0-10	**9.4**	**8.7**	**7.2**
10-15	**10.8**	**8.5**	**8.9**
15-22	**22.5**	**20.1**	**16**

(a)

Background $\lvert\mathrm{CD}\rvert$ (ps)	Background OSNR (dB) 15	16-20	20-25
0 to 130	**3.8**	**2.3**	**1.9**
130 to 260	**2.3**	**1.4**	**1.6**
260 to 400	**1.7**	**1.6**	**1.9**

(b)

Conditional RMS errors for (a) $\lvert\mathrm{CD}\rvert$ (ps/nm) in the presence of background levels of OSNR and PMD (b) $\mathrm{DGD}_{\mathrm{eff}}$ (ps) in the presence of background. OSNR and CD. Results calculated from 500 test cases.

TABLE II
RMS ERRORS AS A FUNCTION OF THE NUMBER OF TRAINING CASES

Number of training cases	$\mathrm{DGD}_{\mathrm{eff}}$ (ps)	$\lvert\mathrm{CD}\rvert$ (ps/nm)
500	3.6	17.3
750	2.7	14.1
1000	2.5	13.0
1250	2.0	11.1
1500	1.9	10.4

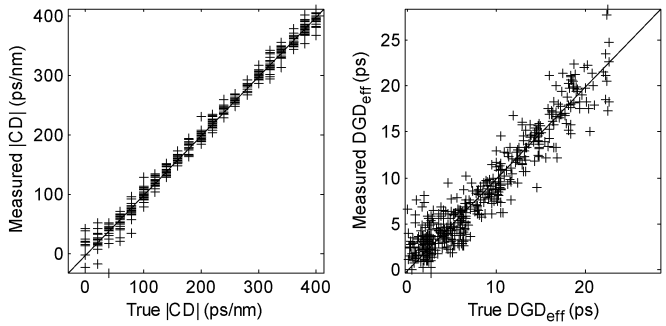

Fig. 7. Experimental results for 40 Gbit/s NRZ DPSK simultaneous measurements of $\lvert\mathrm{CD}\rvert$ (RMS error of 11 ps/nm), and $\mathrm{DGD}_{\mathrm{eff}}$ (RMS error of 1.9 ps). The background OSNR ranged from 15 to 25 dB. 500 random test cases were selected.

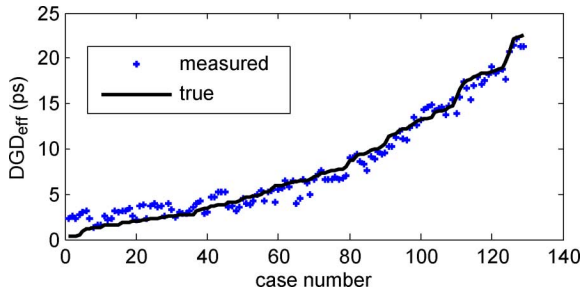

Fig. 8. Experimental results for order cases of 40 Gbit/s NRZ DPSK measurements of $\mathrm{DGD}_{\mathrm{eff}}$ using a running average of 5 cases.

Table II shows the improvement in accuracy obtained with increasing training set size. There is a relatively rapid improvement in performance for the first 1000 training cases which then plateaus for greater than 2000 cases. The required number of training cases will however, vary depending on the impairment range. We note that a significant component of the errors seen in the scatter plots of Fig. 7 are due to random variations in the phase portraits, the effects of which can be negated with a larger numbers of sample pairs or by averaging consecutive measurements. To illustrate this point Fig. 8 shows an ordered plot of true $\mathrm{DGD}_{\mathrm{eff}}$ and measured $\mathrm{DGD}_{\mathrm{eff}}$ for 40 Gbit/s DPSK. In this case we have taken a running average of 5 measurements. The RMS error is reduced by a factor of two and in this case the worst case error is 3 ps for $\mathrm{DGD}_{\mathrm{eff}}$. (A similar improvement for the RMS error for CD is obtained).

For the experimental setup described in this paper the measurement time was 1.2 s. This is dominated by the sampling time; the processing time required for prediction taking less than 0.2 s. In future implementations, cost effective solutions at sampling rates of 50 MS/s will be achievable with currently available sample and hold technology. This together with more efficient processing algorithms will enable measurement times of

better than 50 ms and provide an increased ability to exploit the improvement in accuracy that is achievable through averaging of multiple measurements.

V. Conclusion

The asynchronous delay tap sampling technique is a promising in-service monitoring technique capable of measuring simultaneous impairments for both ASK and PSK signals. The two key components are the characterization of the signal with an asynchronously sampled phase portrait and the extraction of impairment features using statistical machine learning techniques.

In this paper we demonstrate the technique on a 40 Gbit/s NRZ DPSK signal. Prediction models were trained on combinations of dispersion, OSNR and 1st order PMD with random power splits between principle states. Results show $|\mathrm{CD}|$ and $\mathrm{DGD}_{\mathrm{eff}}$ measurements with standard errors of ± 11 ps/nm and ± 1.9 ps respectively, with OSNR levels varying between 15 and 25 dB.

The application of the technique to advanced formats including polarization multiplexed and multi-level systems is under investigation. The authors believe that the phase to amplitude conversion provided by optical filtering together with automated learning techniques provide the potential to monitor these formats with simple direct detection receivers. A key benefit of our approach is that it can be applied without the need for fine tuning of features. For example, using the same algorithms as applied here to NRZ-DPSK, we find that simulations of 40 Gbit/s RZ-DQPSK and 100 Gbit/s 8 PSK give results, that are consistent with accuracies for simultaneous CD and $\mathrm{DGD}_{\mathrm{eff}}$ presented in this paper. Finally, our experience to date suggests that the ability of the machine learning techniques to discern patterns in very "unstructured" phase portraits holds promise for application of the technique to the monitoring of polarization multiplexed systems.

References

[1] D. C. Kilper *et al.*, "Optical performance monitoring," *J. Lightw. Technol.*, vol. 22, pp. 294–304, Jan. 2004.

[2] A. Willner *et al.*, "Optical performance monitoring," in *Optical Fiber Telecommunications, V11B*. New York: Academic, 2008.

[3] Q. Yu *et al.*, "Chromatic dispersion monitoring technique using sideband optical filtering and clock phase-shift detection," *J. Lightw. Technol.*, vol. 20, pp. 2267–2271, 2002.

[4] H. Ji *et al.*, "Optical performance monitoring techniques based on pilot tones for WDM network applications," *J. Opt. Netw.*, vol. 3, pp. 510–533, 2004.

[5] R. A. Skoog *et al.*, "Automatic identification of impairments using support vector machine pattern classification on eye diagrams," *IEEE Photon. Techol. Lett.*, vol. 18, pp. 2398–2400, Nov. 2006.

[6] N. Hanik *et al.*, "Application of amplitude histograms to monitor performance of optical channels," *Electron. Lett.*, vol. 35, pp. 403–404, 1999.

[7] I. Shake *et al.*, "Quality monitoring of optical signals influenced by chromatic dispersion in a transmission fiber using averaged Q-factor evaluation," *IEEE Photon. Techol. Lett.*, vol. 13, pp. 385–387, 2001.

[8] N. Kikuchi *et al.*, "Performance of chromatic dispersion monitoring using statistical moments of asynchronously sampled waveform histograms," *IEEE Photon. Techol. Lett.*, vol. 17, pp. 1103–1105, 2005.

[9] Z. Li and G. Li, "In-line performance monitoring for RZ-DPSK signals using asynchronous amplitude histogram evaluation," *IEEE Photonic Technol. Lett.*, vol. 18, pp. 472–474, 2006.

[10] B. Kozicki, O. Takuya, and T. Hidehiko, "Optical performance monitoring of phase-modulated signals using asynchronous amplitude histogram analysis," *J. Lightw. Technol.*, vol. 26, pp. 1353–1361, May 2008.

[11] S. D. Dods and T. B. Anderson, "Optical performance monitoring technique using delay tap asynchronous waveform sampling," presented at the Optical Fiber Communication Conf., Anaheim, CA, Mar. 2006, Paper OThP5.

[12] T. Anderson *et al.*, "Multi impairment monitoring in photonic networks," presented at the European Conf. Optical Communication, Berlin, Germany, Sep. 2007.

[13] T. Anderson *et al.*, "Experimental demonstration of multi-impairment monitoring on a commercial 10 Gb/s NRZ channel," presented at the Optical Fiber Communication Conf., San Diego, CA, Mar. 2009, unpublished.

[14] B. Kozicki, A. Maruta, and K. Kitayama, "Experimental demonstration of optical performance monitoring for RZ-DPSK signals using delay-tap sampling method," *Opt. Exp.*, vol. 16, pp. 3566–3576, Mar. 2008.

[15] B. Kozicki *et al.*, "Asynchronous optical performance monitoring of RZ-DQPSK signals using delay-tap sampling," presented at the ECOC, 2007, P060.

[16] J. C. Geyer *et al.*, "Channel parameter estimation for polarization diverse coherent receivers," *IEEE Photon. Techol. Lett.*, vol. 20, no. 10, pp. 776–778, 2008.

[17] M. Mayrock and H. Haunstein, "Performance monitoring in optical OFDM systems," presented at the OFC/NFOEC 2009, OWM3.

[18] K. Clarke, T. Anderson, and S. D. Dods, "Monitoring of multiple modulation formats using asynchronous delay-tap sampling," presented at the Australian Conf. Optical Fibre Technology, Melbourne, VIC, Jun. 2007.

[19] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[20] R. Duda *et al., Pattern Classification*. New York: Wiley, 2000.

[21] J. Jargon *et al.*, "Optical performance monitoring using artificial neural networks with features derived from asynchronous delay tap sampling," presented at the OFC/NFOEC, 2009, OTH1th.

[22] H. Kogelink *et al.*, "Polarization-mode Dispersion," in *Optical Fiber Telecommunications, IVB*, I. Kaminow and T. Li, Eds. New York: Academic, 2002.

[23] K. E. Cornick *et al.*, "Experimental comparison of PMD-induced system penalty models," *IEEE Photon. Techol. Lett.*, vol. 18, pp. 1149–1151, May 2006.

Trevor B. Anderson received the B.Sc. and M.Sc. degrees in physics from Auckland University, New Zealand, in 1981 and 1983, respectively.

He is CTO and co-founder of Monitoring Division Inc. (MDI) and is also principal researcher in Network Technologies at National ICT Australia (NICTA). Prior to this he was project leader for optical performance monitoring at NICTA where he developed the core IP for MDI. He has also held research positions at Telstra Research Laboratories and JDS Uniphase. His research interests are in the development of optical performance monitoring technologies and the application of machine learning techniques in Telecommunications.

Adam Kowalczyk is Chief Scientist Officer at Monitoring Division. He focuses on application of pattern recognition techniques to solve critical challenges in the Telecom and Medical space. He is also Principal Researcher at NICTA where he leads the cancer genomics project. He worked previously as Principal Scientist at Telstra Research Laboratories from 1984 to 2004, where he led several projects in artificial intelligence systems. Prior to this he was Research Fellow at Monash University, Australia and Lecturer at the University of Baghdad, Iraq and lecturer at Warsaw University of Technology.

Dr. Kowalczyk has received numerous awards and several patents over his diverse scientific career in the areas of applied mathematics, neural networks, pattern recognition and bioinformatics. He received his Master of Science and PhD from the Warsaw University of Technology, Poland.

Ken Clarke received the B.Sc.(hons.) in applied physics in 1985 from Heriot-Watt University, Edinburgh, Scotland.

Since that time he has worked in a wide range of photonic industries including military, medical, and telecommunications. He was the DWDM R&D Laboratory Manager for Telstra from the years 2000–2005. Since 2005 he worked on optical monitoring projects for NICTA at Melbourne University, and became a founding member of the spin-out company, mdi, in 2008.

Sarah D. Dods received the B.Sc. (hons) in physics and the Ph.D. degree in electrical engineering from the University of Melbourne, Australia.

She is a Graduate of the Australian Institute of Company Directors. She is the Chief Operating Officer of Monitoring Division Inc., She moved to her current business role as a co-founder of Monitoring Division as the company spun out from NICTA. Her prior experience includes over 12 years in photonics research and development, including her NICTA role as a Principal Researcher. She has also worked for VPIsystems, an optical systems simulation software company, as a senior optical engineer; and for Comalco, an aluminium producing subsidiary of Rio Tinto, in various technical roles. She has a broad conference and journal publications record, and has co-authored several patent applications.

Dr. Dods currently serves on the technical program committee for OFC.

Don Hewitt received the B.Sc. and M.E. in electrical engineering from The University of Melbourne, Australia, in 1954 and 1965, respectively.

From 1961 to 1995 he was an academic staff member of the Department of Electrical Engineering, University of Melbourne, with research interests in microwaves and photonics. He has held positions with GEC Applied Electronic Laboratories (UK) and Telstra Research Laboratories as well as University of Leeds. He recently spent a number of years in industry with VPI systems Inc, where he was responsible for developing a cable access product, before returning to the Photonics Research Laboratory, University of Melbourne working on projects with APCRC. His current research interests with NICTA include OFDM over optical networks, broadband access systems and optical monitoring.

Jonathan C. Li (S'07) received the B.E. (first-class hons) in electrical and electronic engineering in 2001, and the B.Sc. degree in computer science and information technology systems in 1999 from the University of Western Australia. He has been a Ph.D. degree student since 2006 working with the National ICT Australia (NICTA) Victoria Research Laboratory in the Managing and Monitoring the Internet (MAMI) project.

He also works as a Research Engineer in Monitoring Division, Inc. His research is currently focused on signal quality metrics in all-optical networks. Prior to this, Jonathan worked for Telstra Research Laboratories from 2002 in radio network performance optimization, conducting simulations into radio access MAC layer and TCP enhancement over wireless networks (an area in which he held a patent), as well as user perceived quality of service evaluation.