2 NON LINEAR CANCELLATION AND MITIGATION TECHNIQUES

2.1 INTRODUCTION

As we have seen, PMD is a time-varying, frequency dependent phenomenon. It limits the upgrade of most of the installed systems. Since the deployment of 1 mile of fiber costs $100,000, keeping these links is dramatically cost efficient. However, techniques to overcome PMD are needed. They have to fulfill the following requirements:

- Channel by channel mitigation (high wavelength dependence)
- Adaptive technique (statistical time varying behavior)
- Able to handle large value of DGD (Maxwellian distribution)
- Able to handle, eventually, higher orders (high wavelength dependence)
- Inexpensive

This (enhanced by the increasing number of channels in WDM systems) leads naturally to the implementation of electrical compensation techniques in the receiver's integrated circuit. Such a solution addresses all the mentioned specifications, especially the first and the last. Furthermore, it provides some extra features, including transparency to the user

and the amelioration of other effects such as imperfect chromatic dispersion compensation or non-linearities. These methods are inexpensive and transparent in the sense that, integrated to the receiver IC, they are hidden in a receiver exhibiting all the features of a normal receiver, except its increased robustness to fluctuating ISI. We consider in this chapter some of the digital techniques that can be implemented at the receiver to process the signal.

First, we recall the principle of the NLC, including a discussion of its particularities and performance that we compare to simulations and experiments. To overcome some of its limitations, we analyze also its association with a Feed forwarD Equalizer (FDE). The third part addresses the issue of an adaptive NLC and the fourth part analyzes its impact on an optically amplified system, where spontaneous emission noise exists. The last section focuses on other mitigation techniques.

2.2 NON LINEAR CANCELLER

We present, in this section, the concept and the implementation of the adaptive Non Linear Canceller (NLC) for Fiber Optic Communication Systems (FOCS). Parts of these considerations are a summary of papers by Kasturia and Winters [18, 19].

2.2.1 Principle

The Non Linear Canceller (NLC) or Decision Feedback Equalizer (DFE) is a concept already widely accepted and used in the communication field especially for wireless links. However, as with most of the digital signal processing techniques, it had not been considered in fiber optic communication, mainly because of the restriction on the speed of electronics. Nevertheless with today's technology, a 10 Gb/s NLC does not seem prohibitively expensive.

The 2-bit NLC is a special receiver, consisting of four decision circuits, sharing in our example the same clock, but having independent thresholds. The decision of a circuit is output if the previous detected pattern corresponds to its associated pattern. The first circuit is activated only when the two previous bits are 00, another takes care of 10, and so on. According to the level of InterSymbol Interference (ISI), each threshold has its own value matching the centroid of the pattern-dependent expected values of 0 and of 1. Therefore, the threshold for 00 is expected to be smaller than for 11. To summarize, the NLC considers the previously received bits to estimate the ISI and then choose an adapted threshold.

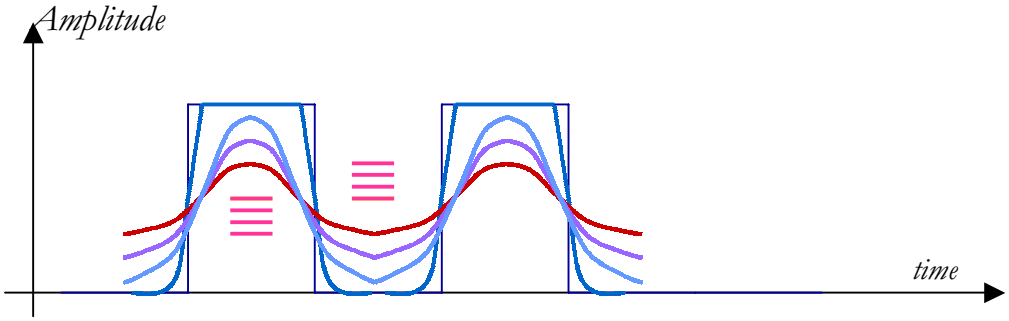

Fig. 32 Evolution of the decision threshold under fluctuating ISI

To understand the NLC's performance, let us recall some facts about ISI and about the Q factor. Then we will describe its principle and its impact.

2.2.1.1 ISI AND BER ESTIMATION

When receiving an On/Off keyed signal, at the sampling time k, the power received is not only due to the bit we want to estimate, but also due to the surrounding bits. If we call x_k the k^{th} symbol sent and s_k the signal received at the time k:

$$s_k = x_k + g(\ldots + x_{k-2} + x_{k-1} + x_{k+1} + x_{k+2} + \ldots)$$

Eq. 43 ISI expression

In the previous expression, g represents the contribution to the sampled signal of all the neighboring bits, called ISI. This undesired power can degrade the transmission quality by closing the eye, in other words by bringing the expected value of the 0 and of the 1 closer together. The decision becomes more sensitive to the noise.

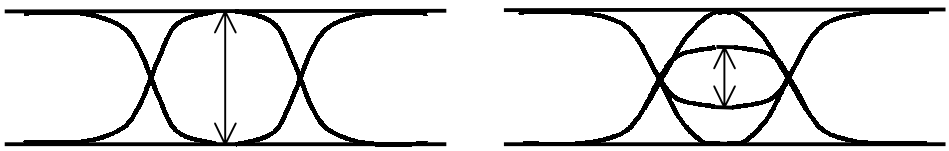

Fig. 33 Eye opening reduction

The BER is given by the probability to decide a 1 when a 0 was sent or to decide a 0 when a 1 was sent. As a consequence, assuming equal symbol probability:

$$BER = p(1)P(0/1) + p(0)P(1/0) = \tfrac{1}{2}\left[P(0/1) + P(1/0)\right]$$

Eq. 44 BER expression, P(0/1) is the probability of detecting a 0, given a 1 was sent

Since we consider Gaussian noise distributions of variance σ_0 and σ_1, the BER is given by:

$$BER = \frac{1}{4}\left\{erfc\left(\frac{I_1 - I_D}{\sigma_1\sqrt{2}}\right) + erfc\left(\frac{I_D - I_0}{\sigma_0\sqrt{2}}\right)\right\}$$

Eq. 45 BER expression (2)

There exists an optimal I_D giving the smallest BER. In this case, we can define a Q factor that gives an estimation of the system's performance. Q relates to the BER as stated in Eq. 46. This result relays on the common approximation of Gaussian noise distributions. One may notice that due to the noise statistics of optical systems, this analysis does not provide the exact optimum threshold. Nevertheless, the Q factor remains a good performance measure because the two approximations (on threshold and on noise statistics) balance each other.

$$BER = \frac{1}{2}erfc\left(\frac{Q}{\sqrt{2}}\right) \qquad I_D = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1} = \frac{I_1 + I_0}{2} \text{ if symmetrical noise}$$

$$Q \equiv \frac{I_1 - I_D}{\sigma_1} = \frac{I_D - I_0}{\sigma_0} = \frac{I_1 - I_0}{\sigma_1 + \sigma_0}$$

Eq. 46 Q factor expression for optimized threshold,

I_d: threshold, I_i: level, σ_i: noise variance, i = 0 or 1

A useful measure is Q=6 for a BER=10^{-9}. If the threshold is not optimized, one of the terms of Eq. 45 is limiting. Therefore, even in this case we can define an approximate Q factor providing reliable performance estimation.

2.2.1.2 NLC'S PRINCIPLE

Let us describe the NLC's function, using a special case where, first, a thermal noise limited system is assumed (constant, symmetrical noise distribution) and where the ISI is due only to the previous bit:

$$s_k = x_k + \alpha x_{k-1}$$
$$\tilde{s}_k = \left(x_k + \alpha x_{k-1}\right) \Big/ 1+\alpha \quad , \quad \text{after normalization}$$

Eq. 47 Received signal, ISI from the previous bit only

In the last expression, s_k represents the signal power before the receiver at the sampling time k, x_k is the k^{th} symbol sent. The second line presents the same entities but after normalization, e.g. right before decision.

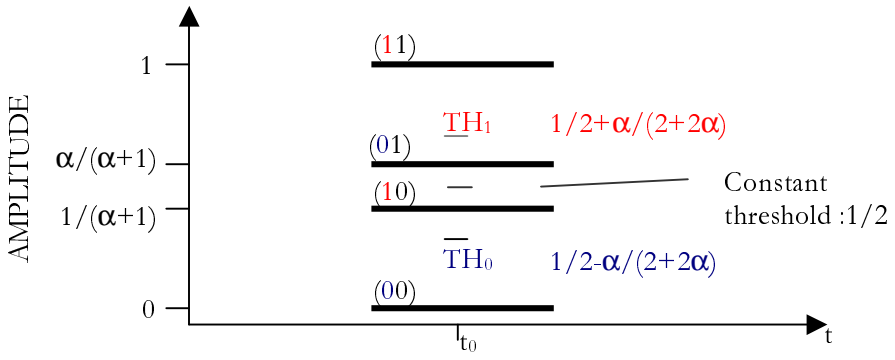

Fig. 34 4-level received eye

In a standard receiver, the threshold is fixed at 1/2, since this matches the ISI-free centroid of the noise distributions in the thermal noise limit. Consequently, the decision is

very sensitive to the noise when the pattern (01) or (10) is received. However, if we know the previous received bit, we can use a better threshold, set at the centroid (halfway, here) of the expected value of 0 and 1, given the previous bit.

So, we can see on Fig. 34 that if a 0 had been received, TH_0 provides a better performance. This is exactly the concept of the NLC. From the 2 previous decided bits, it selects, for the next decision, the best threshold (out of 4). The threshold is supposed to match the ISI and noise distributions estimated thanks to the previous bits.

2.2.1.3 PERFORMANCE

When ISI is severe enough to close the eye (possible when the splitting ratio γ = 0.5 and τ > a bit period, Fig. 25), a normal receiver cannot make an error-free decision anymore, even if the power is increased. On the other hand, the NLC because it takes the decision between 0 and 1/2 or 1/2 and 1 recovers the ISI free performance as soon as the power is doubled. This worst case PMD for a normal receiver results in an infinite power penalty, whereas the NLC exhibits only a 3dB penalty. The power penalty is defined as the power augmentation necessary to reach a reference performance (here the PMD-free situation).

This was a particular example. More generally, we see easily from Fig. 34 that an ISI of level α results in an eye closure of $1-\alpha$ for the NLC and $1-2\alpha$ for a normal receiver.

However, if ISI is due to future bits, as the NLC based its ISI estimation on previously detected bits, it does a poor job. In fact it does not act any better than a normal receiver. Therefore, NLC must be associated with a transversal filter that subtracts and

adds to the detection signal copies of this signal at different times. FDE is however not as efficient as a NLC. First, linear mitigation is poorer than non-linear mitigation (it can not ameliorate non-linear impairment). Second, it enhances the noise. The performance of these techniques is presented in 2.3. Let us now consider the high-speed implementation of the NLC.

2.2.2 Implementation

The very high speed of fiber optic systems, typically 10 Gb/s for the application we consider here, requires a smart implementation of the NLC.

2.2.2.1 A SIMPLE NLC

The common design of a NLC is given on Fig. 35. In such a design, the information on the ISI level is feedback by subtracting to the received signal a constant intensity (using the loop including the synchronization T_B). This intensity is determined by an adaptive loop (not represented) estimating ISI, providing B, the weighting coefficient.

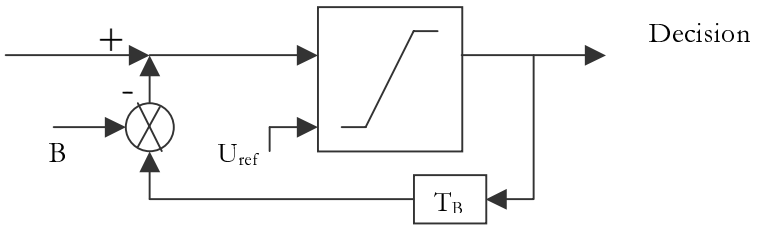

Fig. 35 Decision circuit with analog feedback to the signal

The circuit of Fig. 35 suffers from the distortion related to the subtraction of the feedback to the signal. It can be replaced by a distortion free design.

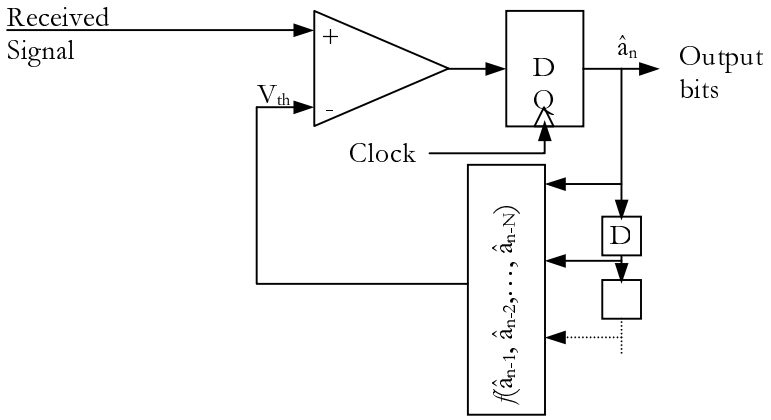

Fig. 36 Decision circuit with feedback to the threshold, the D-Q box represents a Delay logical gate

On the previous figure, *f* represents the correction function. It computes from the received bit pattern ($\hat{a}_{n-1}$, $\hat{a}_{n-2}$,…, $\hat{a}_{n-N}$) the threshold offset.

This network presents a severe time limitation. Basically, it can operate at most at the frequency, given by the inverse of the time delay required between two decisions. This delay has three origins: first, there exists a propagation delay before and after the computation unit (represented by *f*). Then a delay is associated with the computation itself. Finally, the adaptation delay of the comparator to the new threshold can be significant. Therefore, this design suffers from severe time restrictions. Let us examine another more complex but also faster scheme.

2.2.2.2 BOTTLENECK FEEDBACK

Time can be saved at the price of greater complexity. The first step is to avoid the constant modification of the threshold level; the second is to remove the computation from the feedback loop.

Instead of considering a function that outputs 4 possible levels for a 2-bit NLC, we can consider a function selecting a detector out of 4. In this case, each detector has its own fixed threshold. Fig. 37 shows a 1-bit canceller.

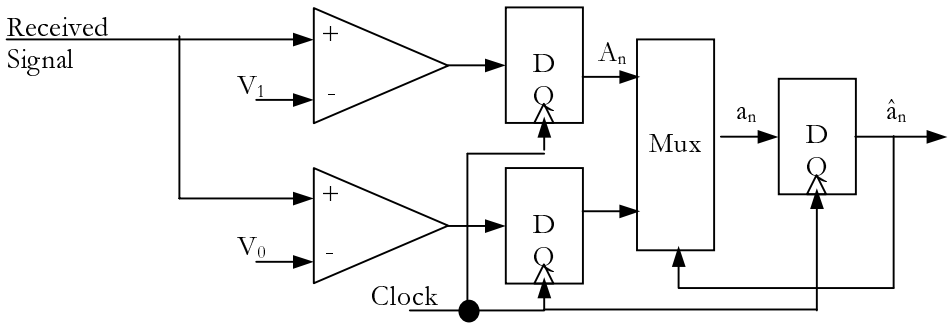

Fig. 37 removing the comparison adaptation time

In this configuration, the time delay reduces to the propagation through the last flip-flop gate and through the multiplexer, plus, of course, the propagation delay of the line from one component to the other.

2.2.2.3 LOOKAHEAD

On Fig. 37, the n^{th} bit is given by the following logical operation:

$$a_n = A_n a_{n-1} + B_n \overline{a}_{n-1}$$

Eq. 48 n th bit computation

where a_n represents the n^{th} sent bit, the overbar, the complement, A_n and B_n the output of each comparator, Fig. 37. We can again apply this equation at the rank n-1 to obtain an expression for a_n as a function of a_{n-2}. After computation:

$$a_n = \left(A_n A_{n-1} + B_n \overline{A}_{n-1}\right) a_{n-2} + \left(A_n B_{n-1} + B_n \overline{B}_{n-1}\right) \overline{a}_{n-2}$$

Eq. 49 n th bit as a function of n-2 th bit

The implementation of Eq. 49 instead of Eq. 48 increases the operating time from T to 2T. Therefore, the frequency of operation is half of the previous. This advantage would be worthless if the computation time of these two functions was greater than T. In fact, they can be extracted from the feedback loop. As shown on fig 7 of [19], all the computations are realized before the feedback loop, where speed is not critical. Our feedback loop, whose bottleneck is the multiplexer works at half the frequency allowing slower components. Basically, a 2-bit canceller working at 10 Gb/s uses only 5 Gb/s electronics. We gain a factor two but we can even hope for higher gain for higher order NLC.

Eq. 49 uses only even bits, if n is even. There exists an equivalent equation for odd bits. Therefore, every other bit is decided by a twin circuit. The two circuits work in parallel: one dedicated to the even bits the other to the odd bits. The very end of the receiver switches from one circuit to the other every bit.

Using a 2-bit canceller, with lookahead logic (process to transfer the computation stage out of the feedback loop), divides the speed by 4, but also multiplies by four the number of gates in the IC. Indeed, this last assertion does not hold for a smart integration

that shares gates between circuit, as presented on fig 10 and 11 of [19]. Pipelining is the main techniques used to realize this sharing.

2.2.2.4 GENERALIZATION

It is easy to understand that an N-bit canceller allows the speed of the feedback loop to be divided by 2^{N-1}. The generalization formulas are given in [19]. This reference provides also the final design of the 2-bit NLC. These formulas will not be discussed here because they do not provide any deeper insight into the NLC principle. Any further interest should be satisfied by the references, which are extremely readable and rigorous.

2.2.3 Adaptive NLC

PMD varies with time. So the propagation conditions change at a frequency of the order of seconds to hours. This requires the use of an adaptive device. Let us consider the concepts and possible implementation of an adaptive NLC.

2.2.3.1 PRINCIPLES

Reference [19] presents a clever way of achieving the adaptation, without knowing the system performance. It is based on the knowledge of the mean values of 0 and 1, knowing the previously received bit pattern. So, considering again Fig. 34, TH_0 is approximately the average of the signals level at t_k for the pattern (0 1) and (0 0). Symmetrically TH_1 is the average signal level for (1 1) and (1 0). This uses four extra detectors for a 1 bit NLC, two for each threshold. The adaptation part is activated only when the previous decided bit corresponds to the considered threshold. This part consists of a comparison of the input signal level to the actual threshold. It outputs the difference

between the two, by a feedback loop as shown on Fig. 38. The signal converges to the average level of the triggering pattern. A RC filter induces a low pass effect (that we expect to be in the range of the impairment change frequency) that softens and stores this average level.

The output of this stage is directly the level for, for example, (0 0). A twin network will compute the average for (0 1). The average of these value feeds the main comparator for a previous detected bit of 0.

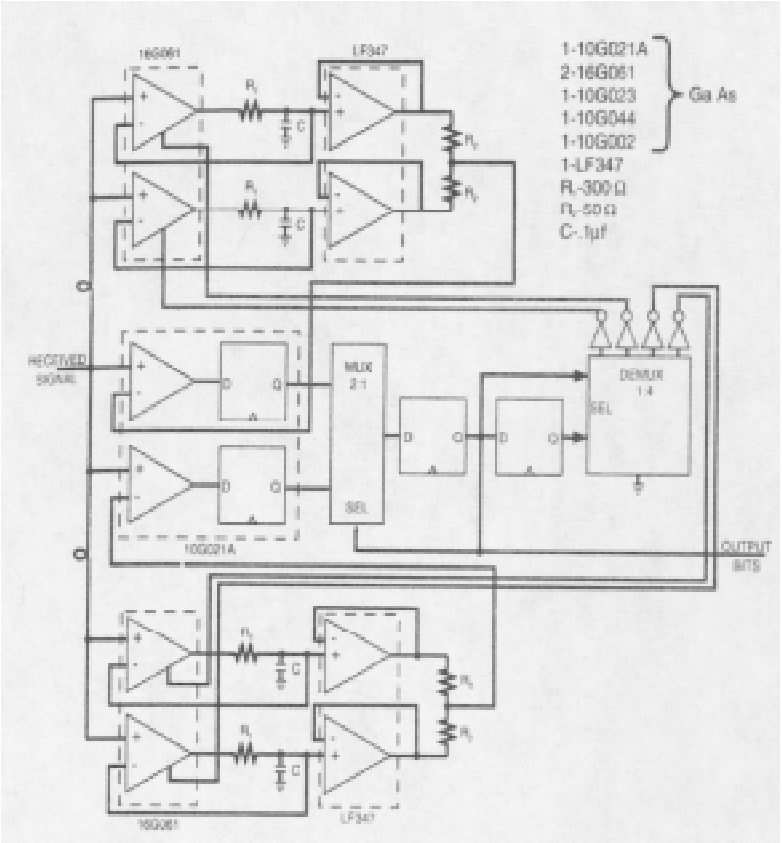

Fig. 38 Feedback loop from [20]

2.2.3.2 IMPLEMENTATION

For the implementation, only the detector (especially its triggering) has to follow the data rate. All other components can operate at reduced speed. The RC circuit tends to

average out the level of each signal and to hold this value when the pattern received does not correspond to the activating pattern of the circuit. Out of the RC circuit, we have the average signal level, however because of current drops due to the resistors, an Op-Amp, used as a follower is introduced. The adaptation speed is directly determined by the bandwidth of the RC filter.

$$f_{3dB} = \frac{1}{2\pi R_1 C}$$

Eq. 50 3dB cutoff frequency of the RC circuit

Two flip-flop gates are introduced in the main receiver part to synchronize the adaptation part to the occurrence of its related bit pattern.

2.2.3.3 EXPECTED ADAPTATION PERFORMANCE

The adaptation is a trade off between the required tracking speed and the stability of the initialization process, as well as stability of the voltage across the capacitor. According to the impairment we are considering, any frequency higher than 100 Hz seems reasonable and a frequency below 10 Mhz should allow this stability.

The implementation of the NLC of [19] exhibits a very good behavior; it adapts to abrupt change of ISI in less than 100ms. It acquires the signal level as soon as the main detector does not exhibit a too high BER.

2.2.3.4 CONCLUSION

We have seen that the NLC optimizes the decision threshold according to the ISI level, estimated from the decided bits. This results in a 3dB penalty when the eye is

closed. The implementation of its adaptation algorithm can be made without a control IC, which makes it even more attractive for tracking fluctuating impairments.

However as already mentioned, PMD exhibits a random, changing behavior and the NLC can take care of PMD only half of the time (ISI due to previous bits). As a consequence, it is necessary to consider another technique for these complementary periods.

2.3 FEED FORWARD EQUALIZER

A FDE, able to remove ISI due to future bits constitutes an appealing complement to the NLC. We recall its principle and study its implementation. We derive also the performance of its association with the NLC. Finally, we briefly mentioned some adaptation techniques.

2.3.1 Principle

The goal of equalization is to correct the frequency response of our channel to match a distortion free response, Fig. 39.

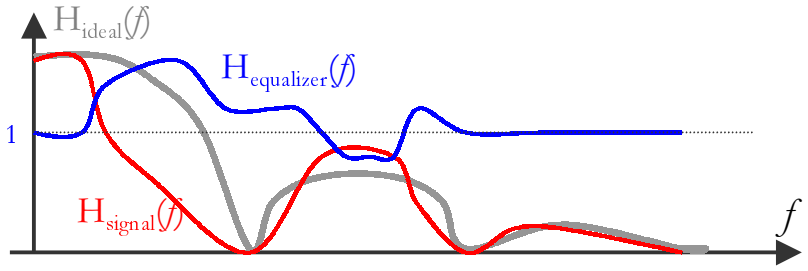

Fig. 39 Equalizer frequency response

The spectrum of the received signal arises from the modulation of the signal and from the distortion due to the channel. By comparing the ideal channel response and the received spectrum, one can deduce the frequency response of the link. This indicates the dispersion origin. For example, spectrum distortion from chromatic dispersion does not appear as the dual imaging spectrum more typical of PMD.

2.3.2 Implementation

The usual implementation of a FDE uses tap delay lines. In fact, its frequency response can be translated using Fourier analysis to an impulse response. This response appears as an infinite sum of Dirac impulses, equally spaced. Depending on the spectrum of the distortion (smaller or larger than the signal bandwidth), this time interval is equal or smaller than the bit period. Since an infinite tap delay line is not realistic and since the weight coefficients of the line are rapidly decreasing, the equalizer can be approximated using a few taps, Fig. 40.

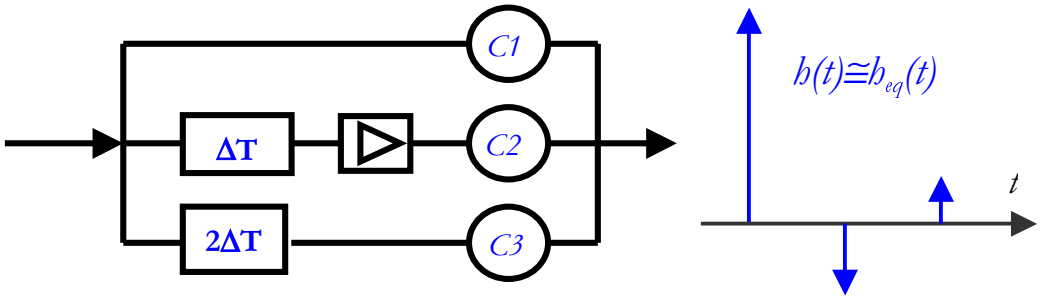

Fig. 40 3-taps transversal filter: scheme and impulse response

Realizing a 10 Gb/s FDE is not as challenging as implementing a DFE at this speed because there is no feedback. In our experiments, we realize the structure of Fig. 40 using different microwave components including splitters, attenuators (to reduce reflections), combiners, amplifiers and adjustable delays. Suppressing the reflections due to imperfect impedance matching was an issue; especially at the splitters/combiners, because of their resistive nature. This resulted in poorer performance than expected. Suppression of relative intensity noise due to these reflections requires SiGe technology [20,21]. The challenge in this case consists in building delay lines of 100 ps on a chip.

2.3.3 Performance

For the simple case of the dual imaging (first order PMD), we can translate the physical response into a signal response by considering the amount of power leaking from a bit to the next. This mapping from PMD to ISI depends on the DGD and the PSP as well as the pulse shape (including filters).

Assuming we know this mapping, providing the impulse response of Fig. 41(a), we see that a N-tap equalizer provide a reduction of the copy by α^{N-1} Fig. 41(b).

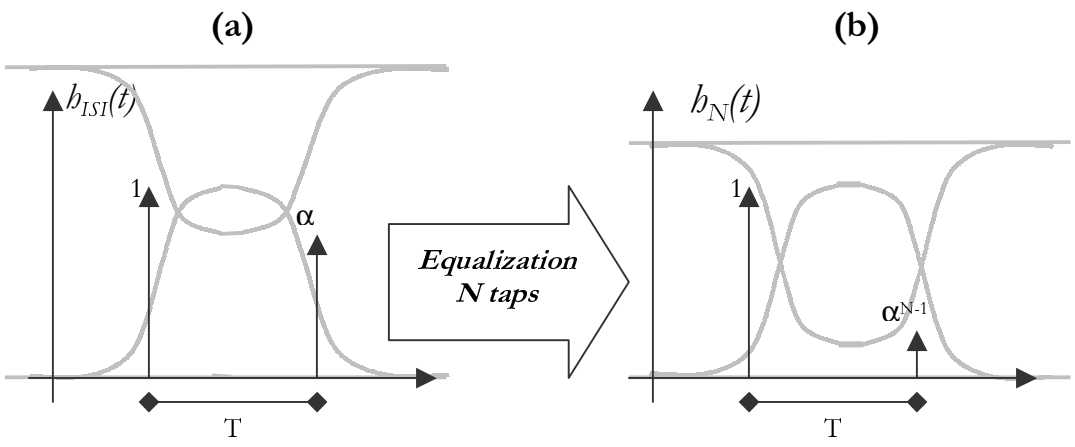

Fig. 41 ISI reduction, eye opening enhancement (gray)

If α is equal to 1 (closed eye) the FDE does not work. Fortunately, in such a case, as we mentioned earlier the NLC exhibits only a 3 dB penalty. Therefore the compound, NLC+FDEexhibits a 3 dB penalty as well.

The equalizer, as it adds and subtracts copies of the signal, enhances the noise variance before decision. Basically, in the assumption of constant symmetrical noise, the noise variance enhancement is given by the sum of the squares of the c(i).

$$\sigma^2_{post-eq} = \left(\sum_i c_i^2 \right) \sigma^2_{pre-eq}$$

Eq. 51 Noise enhancement for linear equalization

The NLC because it uses decided bits does not enhance noise. However, it may propagate errors. After an error, the NLC takes the output of the wrong decision circuit, increasing the risk of getting a new error. According to the expected BER (very low) for optical communications, this is not an issue.

To summarize, there exist mostly three possible situations:

- ISI due to future bits, removed by the FDE (little noise enhancement)
- ISI due to previous bits, cancelled by the NLC (negligible error propagation)
- Closed eye (worst case for a normal receiver), the NLC reduces the power penalty to 3 dB.

2.3.4 Adaptation

To be efficient an equalizer should be adaptive. We saw earlier one of the adaptation techniques for the NLC, having the great advantage that it does not require a control IC. Such an elegant solution has not been found yet to track ISI. Due to the very high speed of fiber optic systems, it is not likely that anyone will implement 'on the flight' training sequence. This would waste bandwidth as well as require the precise recognition of the waveform. This is not realistic for a 10 Gb/s system.

Therefore Peak distortion Criterion and the Zero Forcing Algorithms appear as the two main candidates of adaptation algorithm, [22]. Because of the necessary digitalization of the waveforms at such speed, we employ their discrete version. Thus, they converge to the same algorithm, which exhibits poorer performance. According to [23] it remains however within an acceptable range, especially thanks to the very low BER expected in optical communication that prevents them from diverging.

In this case, a control IC is required; it reduces the transparency of the device. Because of the peculiar implementation of the NLC, we discussed, the two techniques, FDE and NLC, should not interact. They can be independently optimized.

2.4 EXPERIMENTS/SIMULATIONS

In order to verify the expected performance of the compound NLC + FDE and to qualify our simulations, we ran experiments at 10 and 5 Gb/s for different cases of PMD.

2.4.1 Experimental setup

Our set up, same as in [24], used a CW laser, at 1552 nm, externally modulated (pull/push $LiNoB_3$, rise time 22 ps). The signal polarization was controlled before entering a PMD emulator (described later). An optical attenuator allowed us to vary the detected optical power. In the electrical domain, after amplification and filtering (Amp on Fig. 42), the electrical current was equalized. This corrected signal was fed into the NLC for decision. The NLC output ran to a Bit Error Test Set (BERTS) to measure the BER.

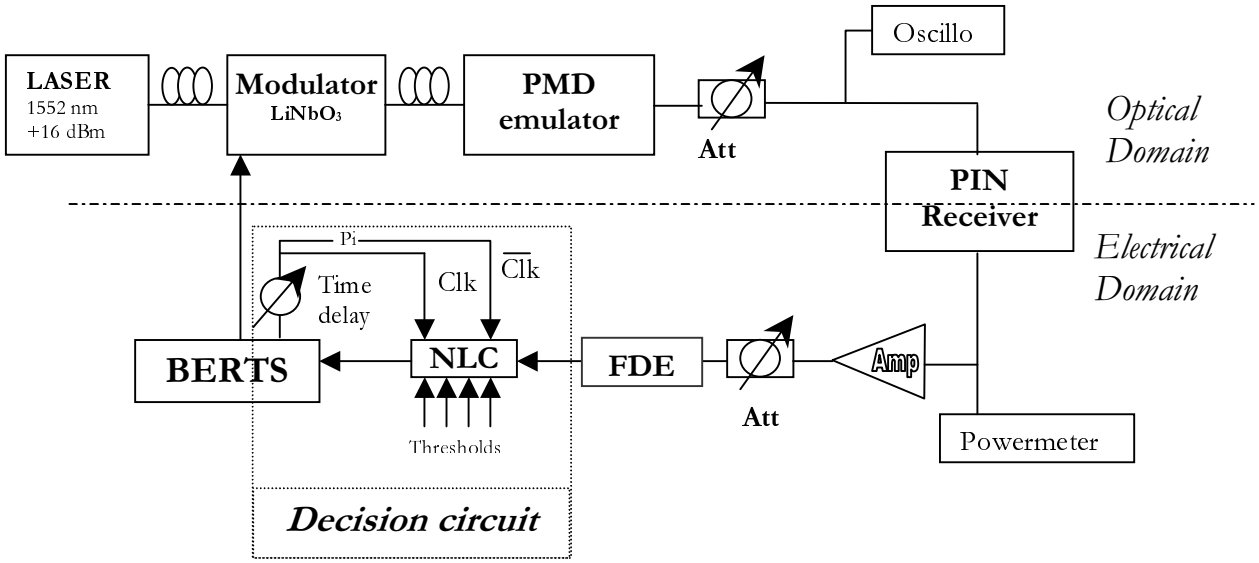

Fig. 42 Experimental set up: NLC+FDE

The clock and the threshold of the NLC are manually adjusted to optimal performance, so are the coefficients of the transversal filter. In order to compare receiver's performance, we use either a normal receiver of constant threshold (set to optimum of the PMD free case), a normal receiver having an adaptive threshold (threshold reset for each case to best performance) or the NLC. To take into account the possible limitations due to the device, we use; we keep the NLC chip to emulate all these cases. The NLC thresholds can be kept constant to obtain a typical system decision circuit. They also can be kept equal and this unique value optimized. This is equivalent to a normal receiver, using an adaptive threshold. Finally, the NLC was normally used with four independent optimized thresholds, providing the performance of decision feedback equalization.

2.4.2 Devices

Some of the 'boxes' of the previous graph deserve some extra attention. First the first order PMD emulator has a range of [-50ps;250 ps]. It can provide any splitting ratio (determined by the input polarization). It is consisted of two beam splitters (first: splitter, second combiner) separating the polarization that then encounters different path, Fig. 43. It had very low Polarization Dependent Loss (PDL), nominally 0.1 dB. Since we split orthogonal polarizations, there is not beating between the 2 beams at the combiner.

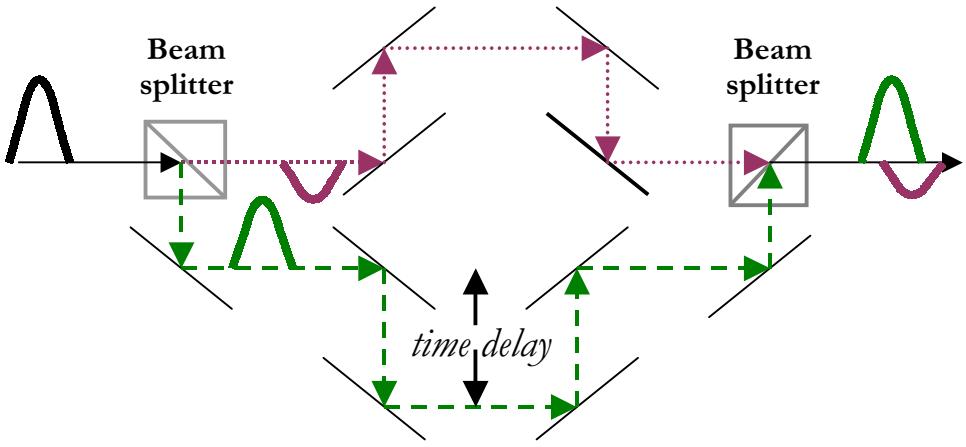

Fig. 43 PMD emulator structure

In the electrical domain, the box "FDE" hides a cumbersome transversal filter. Its design includes splitters, amplifiers and attenuators, Fig. 44. One may note the attenuators or amplifiers at the splitters. They help suppress the reflections.

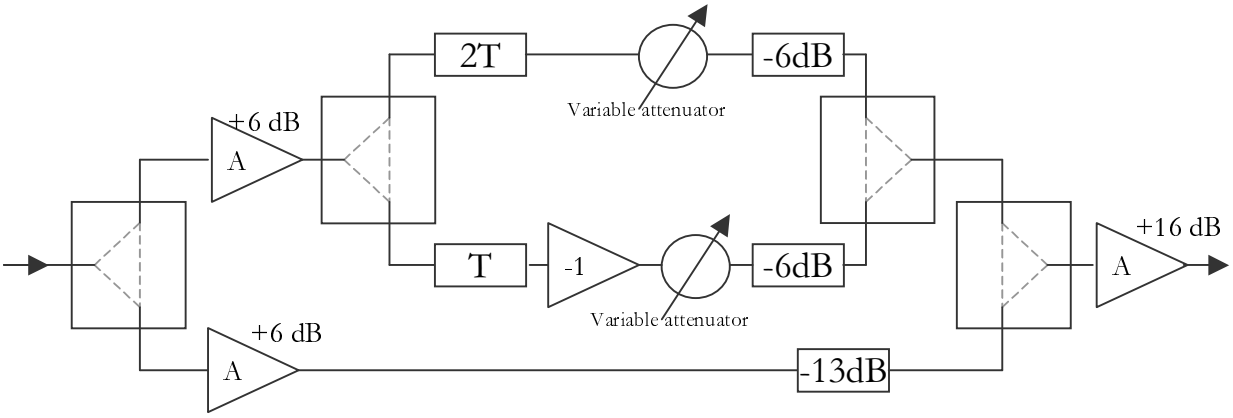

Fig. 44 FDE implementation

In the linear equalizer design, the boxes marked with T are delays. The delay T is usually a bit period (100 ps at 10 Gb/s). We also set it to half a bit period to check the fractional equalizer performance. The delays of an equalizer are usually taken constant. The modification of the tap coefficients tunes the equalization.

The NLC, we use is a 2-bit NLC. This means that it has four independent thresholds. They share the same clock. The decision of one of the circuits is output if its associated bit pattern was previously received. This technique is implemented on an IC[3].

The box called BERTS is a Bit Error Test Set. It generates Pseudo Random Bit Stream (PRBS), in our case, a PRBS of length 2^{31}-1. It receives the detected signal and decides what bits are received. Comparing its decision to the sent bits, it gives the BER of the system.

To verify the performance of the equalization technique in the thermal noise limit, we measure BER curves for different cases of first order PMD, generated by the emulator. A BER curve is the plot of the BER as a function of launched power. We obtain then sets of curves, from which we derive the receiver performance, presented in section 2.4.4.

2.4.3 Simulations

The simulations were written to match our experiments so that we can qualify them and use them for generalization, as we are looking for a performance map of the DSP techniques.

We model our signal in the base band as a NRZ bit stream. A pulse has a rise/fall time of 0.22T, T being the bit period. We create 64x2 bit streams representing all the patterns yxx0yyy0 and yxx1yy0 where xx are the previously received bits. We do not use

[3] This particular device had been realized by a commercial foundery after a design of Jack H. Winters. It was a 2 bit NLC without the adaptive setting of the thresholds.

pseudo random bit streams because they require case study at the receiver that is extremely computation inefficient. By classifying these streams, we avoid this time cost. The bit patterns experience PMD (impulse response linear in the electrical domain) and after detection a Bessel filtering at 3.4 Ghz for 5 Gb/s and 7Ghz at 10 Gb/s. We employ the impulse response of the transversal filter to model it. Finally the optimum BER, Eq. 44, is given for a power, a previously received pattern, a PMD case and an equalizer. This provides us with BER curves that we can exploit to derive power penalty curves.

We use Matlab for this simulation so that most of our bit streams are concatenated to save computation time. One may argue that we use convolutions in the time domain rather than multiplication in the frequency domain. This suits however better our setup. We compute, first the experimental situation and then derived a PMD mapping of the NLC and FDE expected performance.

2.4.4 Results/interpretations

In order to assess the DSP performance as a PMD mitigation technique, we need a tool computing the power penalty enhancement for any PMD case. This section presents the experimental validation of this tool and its use.

2.4.4.1 WORST CASE

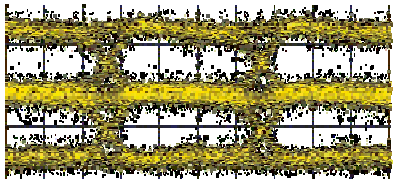

Fig. 45 Worst case PMD, experimental data

The PMD worst case results in the total closure of the eye (DGD = T, $\gamma = 0.5$). in such a case, a normal receiver exhibits an infinite penalty. We compare here experimental results to simulation results, first at 10 Gb/s, where T = 100 ps.

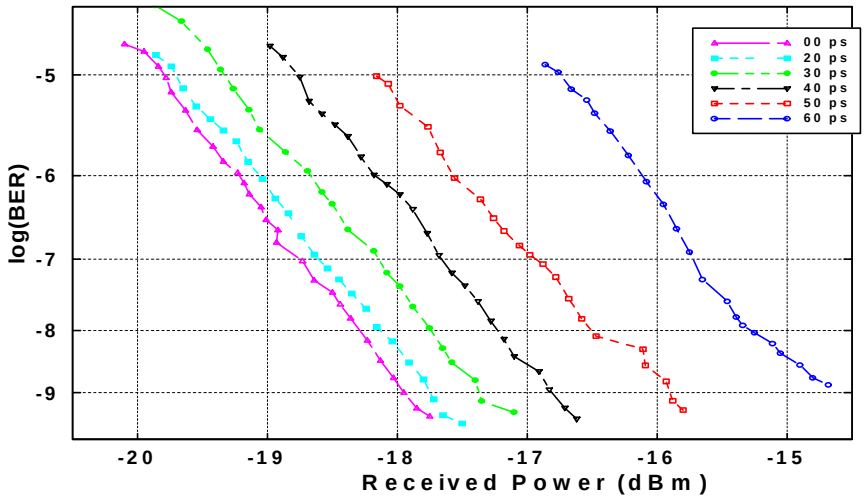

Fig. 46 BER curves, 10Gb/s, first order PMD,$\gamma = 0.5$, DGD<60ps, Normal receiver

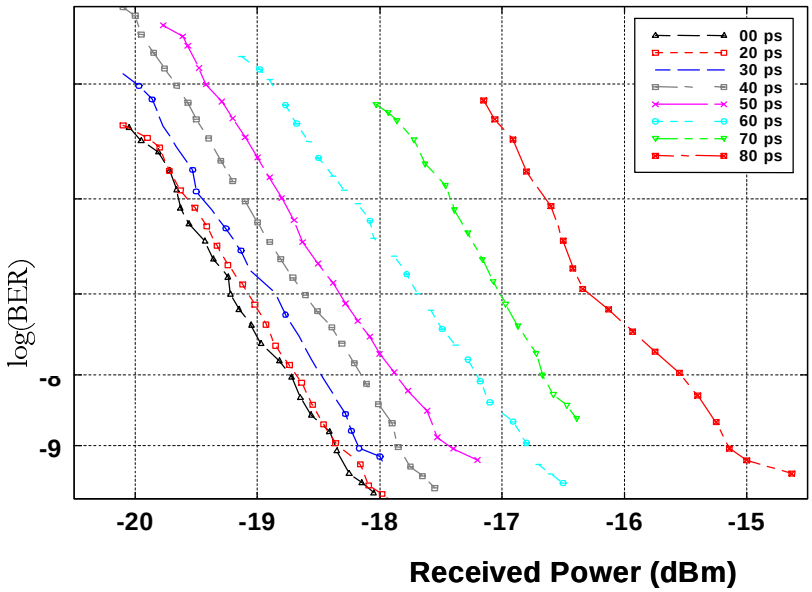

Fig. 47 BER curves, 10Gb/s, first order PMD,$\gamma = 0.5$, DGD<80ps, NLC

These sets of curves are obtained as the power dependence of the BER for different PMD parameters. They are for a given splitting ratio of ½ and for growing DGD (every 10 ps). The closure of the eye appears only as DGD = 100ps. If we cut the graphs at the BER = 10^{-9} value and measure the power from the PMD free (DGD = 0 ps) point to a PMD impaired point, we obtain the power penalty. This follows from the definition of power penalty.

It should be recalled that in this case, the FDE was turned off, so that the performance of the NLC alone is considered.

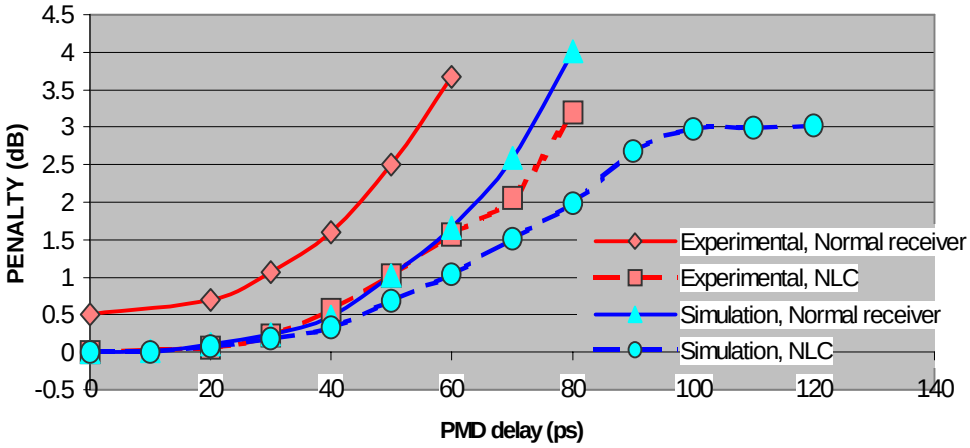

Fig. 48 Power penalty, 10 Gb/s, filter 7 Ghz, first order PMD, $\gamma = 0.5$

We see on the previous chart the dramatic improvement provided by the NLC (dashed) in both cases. However, whereas we expect the NLC to work even at the closure of the eye, it exhibits in the experiments an infinite penalty. There is significant disagreement between simulations and experiments. We attribute this mismatch to the small bandwidth of our device that acted as a low pass filter, inducing ISI.

Considering this assumption, it is no more surprising to see the NLC providing an extra 0.5 dB improvement in the absence of PMD (when DGD = 0ps).

To address the bandwidth issue, we conducted equivalent experiments at 5 Gb/s, where bandwidth should not be a limit anymore. As expected we obtain a better match of our simulation to the experiments (see Fig.51). We do not obtain the expected 3 dB penalty floor, but rather a 4 dB floor. This comes probably from other impairments present in the system and most likely to a non-perfect cancellation of ISI, due to the device, we use. Regardless of this small divergence on values, the NLC's behavior is very well modeled by the simulation.

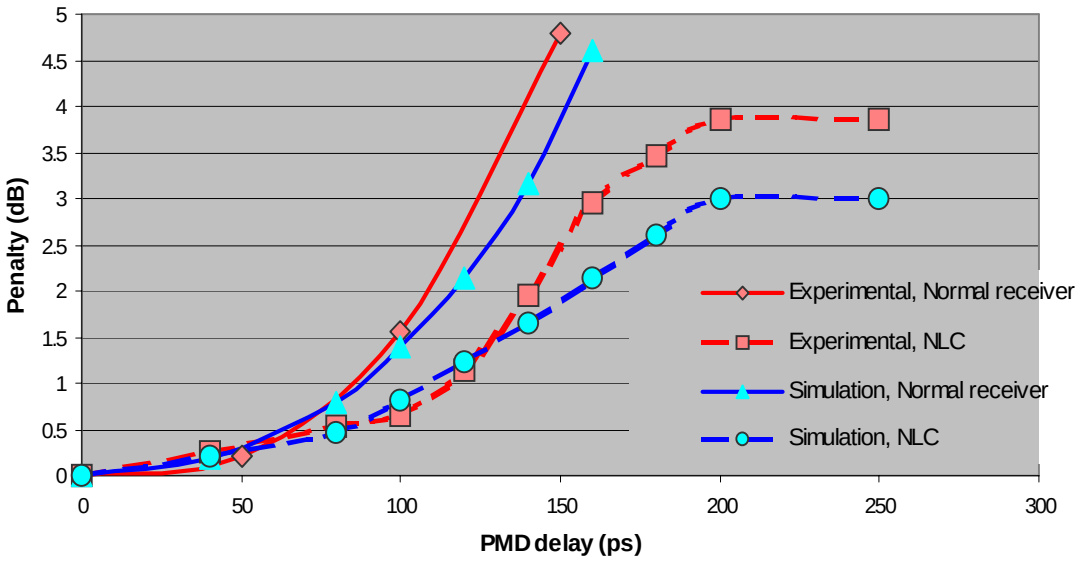

Fig. 49Power penalty, 5 Gb/s, filter 3.4 Ghz, first order PMD, $\gamma = 0.5$

2.4.4.2 PARTICULAR CASES

We verify experimentally other features of the 2-bit NLC. We emphasize here its asymmetrical behavior based on the fact that it cannot cancel ISI due to future bits. In

order to describe PMD we use a special formalism to characterize the splitting ratio γ. The notation 25/75 stands for an impulse response where 25 % of the total pulse energy arrives first. Similarly, 75/25 represents a case where most of the energy arrives first so that the rest can be considered as ISI due to the previous case. We expect then the NLC to work better. On the other hand, in the earlier mentioned case, the NLC does not work any better than a normal receiver. Whenever the ratio is less than 0.5 or the first percentage is the smallest, the NLC does not work better than a normal receiver does.

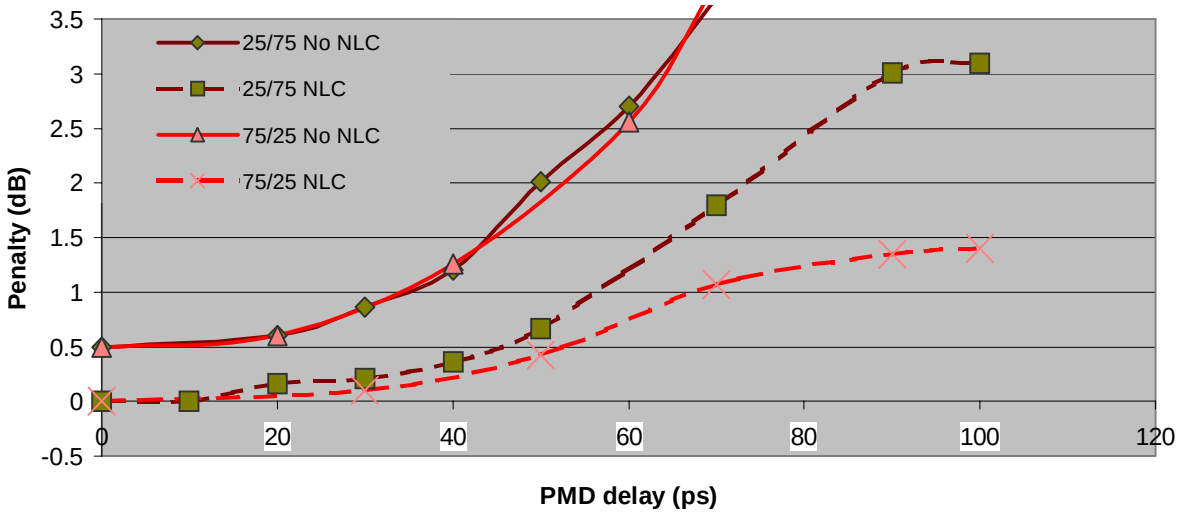

Fig. 50 Power penalty curves, NLC (dashed) vs Normal receiver (solid), for the cases 25/75 or 75/25 at 10 Gb/s, first order PMD

We verify on Fig. 50 the expected asymmetrical behavior of the NLC. Even if not represented, the experimental data for the NLC match very accurately our simulations. It is not the same for the normal receiver that suffers the remaining ISI, created by the device, we use. One may note the remaining 0.5 dB improvement as on Fig. 48. Of course, a better (more recent) device would not create such features.

We demonstrated 2-bit non-linear cancellation at 5 Gb/s to overcome PMD. We verified also our computer model of the NLC. Let us now study the performance of the FDE, for the range of PMD, where the NLC does not work.

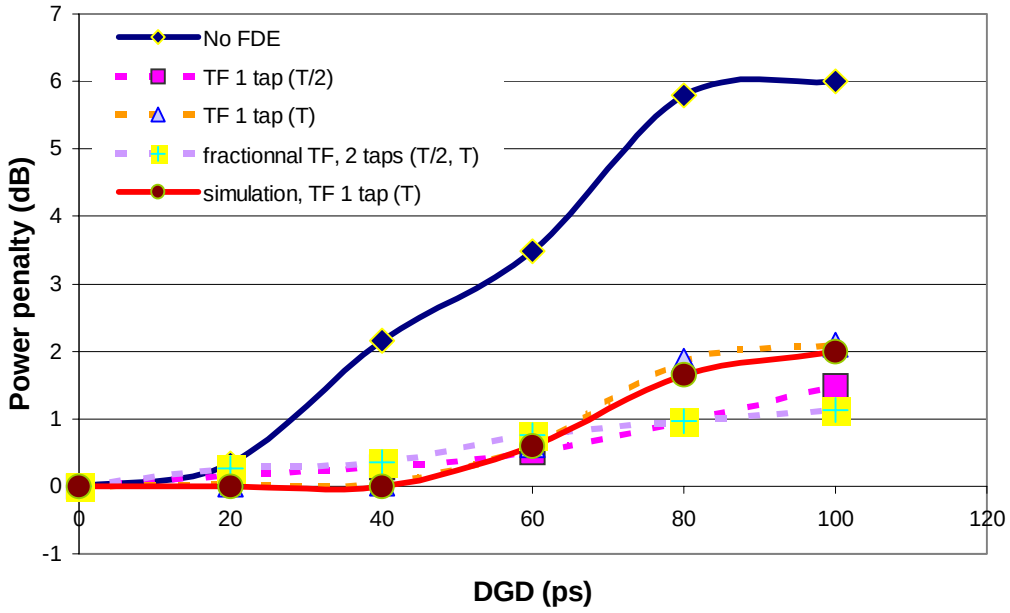

Fig. 51 Compound FDE+NLC versus NLC alone: experimental comparison at 10 Gb/s, the tap spacing is given in brackets in the legend, T stands for the bit period, here 100 ps

The previous graph represents the power penalty reduction due to linear equalization in the case 25/75. First, our experiments demonstrate at 10 Gb/s the efficiency of a FDE when the NLC cannot handle ISI (solid dark blue). The FDE of our set up reduces the power penalty from 6 dB to below 2 dB. There exist several implementations of this transversal filter (TF). The fractional space equalizers (dashed purple and pink) even reach the theoretical limit of 1.25 dB power penalty. We see that a normal 1 tap transversal filter (dashed orange) reduces consistently the power penalty even if it does not reach the performance of a fractionally spaced equalizer. We validate

the FDE+NLC simulation. There is a strong correlation between simulation (solid red) and experiments (dashed orange) for the first order FDE.

We have experimentally demonstrated at 10 Gb/s for most of the cases and at 5 Gb/s for all cases the electronic mitigation of first order PMD. The DGD can vary up to 2 bit periods before we reach infinite penalty. Furthermore, we validated our computer simulation so that we can derive a PMD-performance map of our techniques.

2.4.4.3 GENERALIZATION

Confident of our simulations, we generate a PMD mapping of power penalty for different type of receivers, Fig. 52.

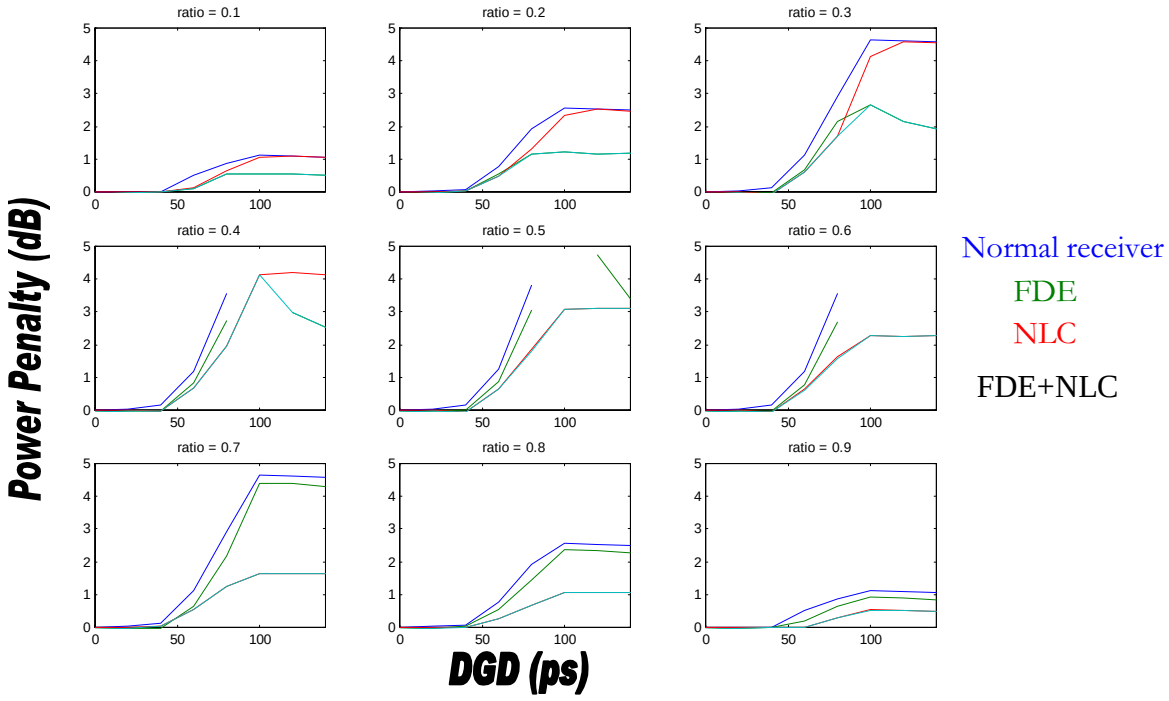

Fig. 52 PMD dependent power penalty mapping, ratio corresponds to γ, part of the energy arriving first

From the previous graph, we can deduce the worst case for the compound FDE+NLC. It appears for the splitting ratio 0.4 (40 % of the energy arrives first), Fig. 53. At 10 Gb/s, when the DGD equals 100 ps (a full bit period), the power penalty exhibited by the combination equalizer is 4 dB. This means that any type of first order PMD can be handled with less than 4 dB penalty if the range of the equalizer is broad enough.

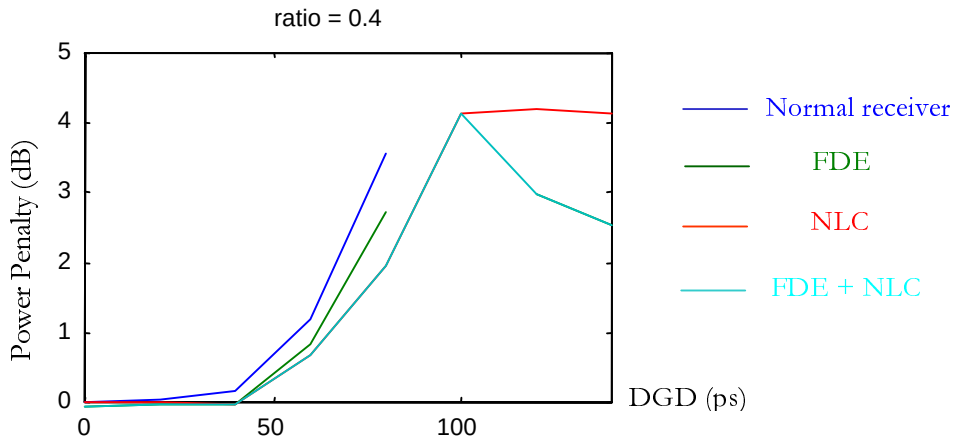

Fig. 53 PMD induced penalty, worst case of the compound NLC+FDE

The worst case for the compound is not the same as for the normal receiver because cancellation is not equivalent to linear equalization. This comes from the noise enhancement of linear equalization. Furthermore, we do not consider an optimal equalizer, since it has only 3 taps. Therefore, the worst case appears for a ratio of 0.4 when the NLC can not handle ISI but the ISI level is still high. High ISI level requires large equalization coefficients, which induces severe noise enhancement and creates a situation of very small margin.

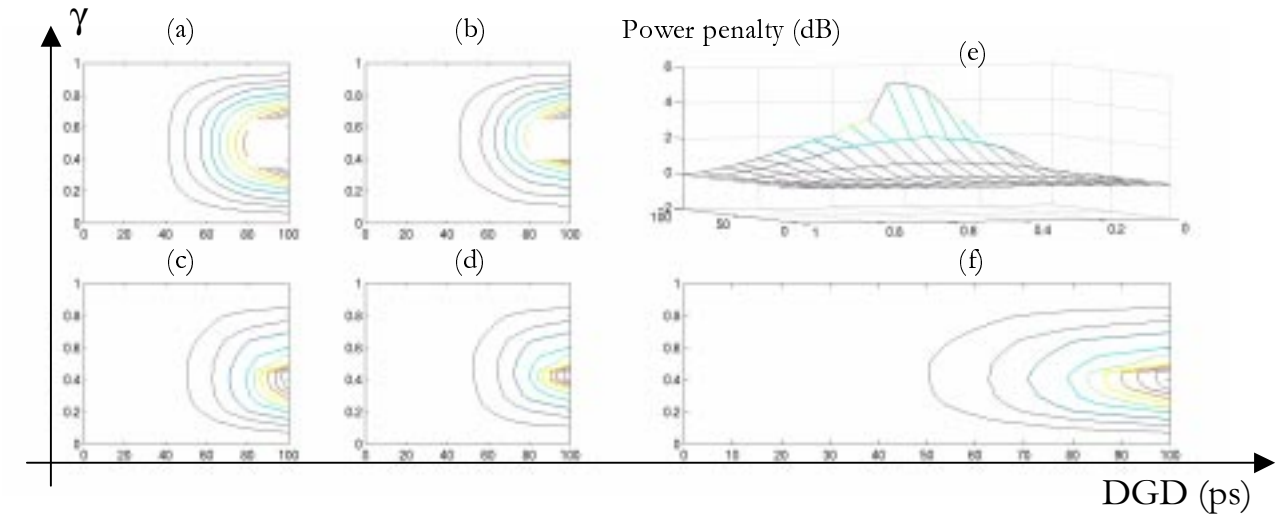

Fig. 54 Power penalty contour at 10 Gb/s: a line every 0.5 dB. 1 dB penalty: second blue line (a) Normal receiver, (b) FDE, (c) NLC, (d) compound, (e) mapping for the NLC, (f) associated contour

Fig. 54 represents the penalty contour of the different techniques we consider. Every line represents a 0.5 dB penalty. Consequently, the second line represents the 1 dB penalty edge. These contour were obtained using penalty maps similar to (e), where power penalty is given as a function of PMD (ratio, DGD).

If 1 dB is considered as the maximal accepted penalty, we see that we gain only 10-15 ps of actual DGD at 10 Gb/s. This would set a serious limit on electronic mitigation if systems were not only ISI limited but also power limited

We concentrated here on the PMD related issues, the performance of electronic mitigation relies however on some further parameters.

2.4.4.4 OTHER ISSUES

Our previous considerations seem to be system independent. Of course, they are not. We study qualitatively some consequences in the filtering and the clocking at the decision stage.

2.4.4.4.1 Filtering

We know that low pass filtering helps increase the SNR by suppressing the noise out of the signal bandwidth. This operation removes also some high frequency signal components. The transitions are not as sharp as without filtering, the eye closes in the transversal direction. For an incorrect filter design, it can even result in a vertical closure.

The performance of our devices is set by the level of ISI. Furthermore mapping between ISI and PMD is given by the pulse shape, highly dependent on the filters of our systems. Therefore, the filter designs affects directly the electronic mitigation performance, as on can see on

It should be recalled that in this case, the FDE was turned off, so that the performance of the NLC alone is considered.

Fig. 48 and Fig. 49.

Filtering because it smoothes the two replica of the pulse into a single one, can aggravate the eye closure. This would indicate that we want a filter as wide as possible. This is contradictory to our earlier elementary noise considerations. Moreover, filtering provides an improvement related to the decision clock.

2.4.4.4.2 Clock phase

PMD gives rise to three level bit patterns,

Fig. 55.As a consequence, there exist for a single bit two rising and two falling edges. As the clock results from a nonlinear operation on the signal, it may create 4 components where only two should exist (especially in case of RZ transmission). Furthermore, even if the clock generates only 2 clock components, whose time average triggered the decision phase, it may walk off the optimum point because of PMD pattern dependence. By filtering, we correlate the maximum eye opening with the middle of the edges. Because our decision circuits have a single clock phase, such symmetry provides better performance, closer to the perfect performance that a pattern dependent clock phase would provide.

Finally, the filter has to be designed to obtain the best tradeoff between timing jitter robustness, pulse shape and noise.

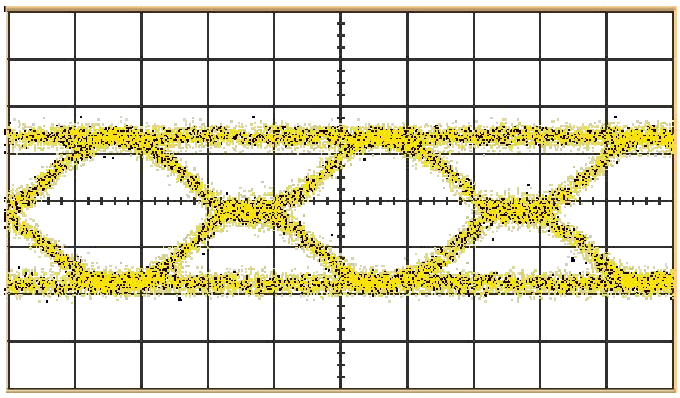

Fig. 55 3-level bit pattern, experimental data

These were the main factors that can influence strongly our derivation and demonstration during the integration into a system.

We demonstrated through simulations that any first order PMD can be handled within a 4 dB penalty margin, when using electronic techniques. This does not really fit in the 1 dB penalty requirement. However, old vintage systems exhibit lower loss than expected so that they are not power limited but ISI limited. Consequently, there is probably few dB of extra margin there. It would allow these techniques to provide the expected performance. However, one may notice that systems are no more designed in the thermal noise limit, which we used so far. Because of the optical amplifiers, they exhibit spontaneous emission that creates a noise, very different from the thermal noise. As derived in the next part, this modifies dramatically our expected performance. We obtain higher penalty than 3 dB for a closed eye and 4 dB for the worst case.

2.5 PERFORMANCE DEGRADATION IN OPTICALLY AMPLIFIED SYSTEMS

We study in this part the consequences of being no longer limited by the thermal noise but rather by the Amplified Spontaneous Emission (ASE). The beating between ASE and the signal (consequence of a square law detection[4]) results in signal dependent noise, Fig. 56.

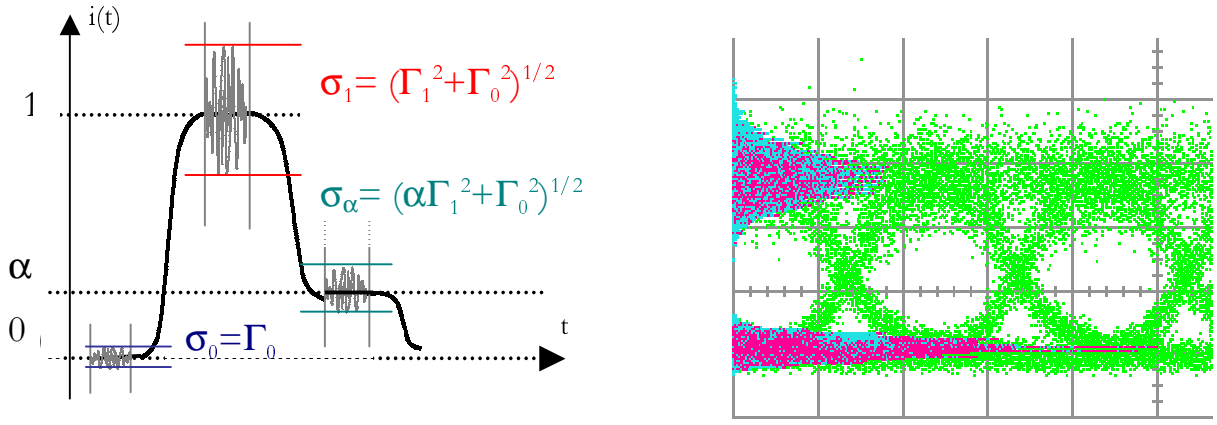

Fig. 56 Signal dependent noise, formalism and experimental data, α represents ISI level

We derive and verify the performance of electronic mitigation in the presence of ASE. We first recall the principal noise processes in fiber optic communication systems. We introduce a new formalism to obtain the performance bounds. We experimentally verify them. Finally, a simulation is conducted to estimate the global performance of the DSP techniques in this case.

2.5.1 *Noise processes in fiber optic communication systems*

A system exhibits different noise sources, red stars on Fig. 57.

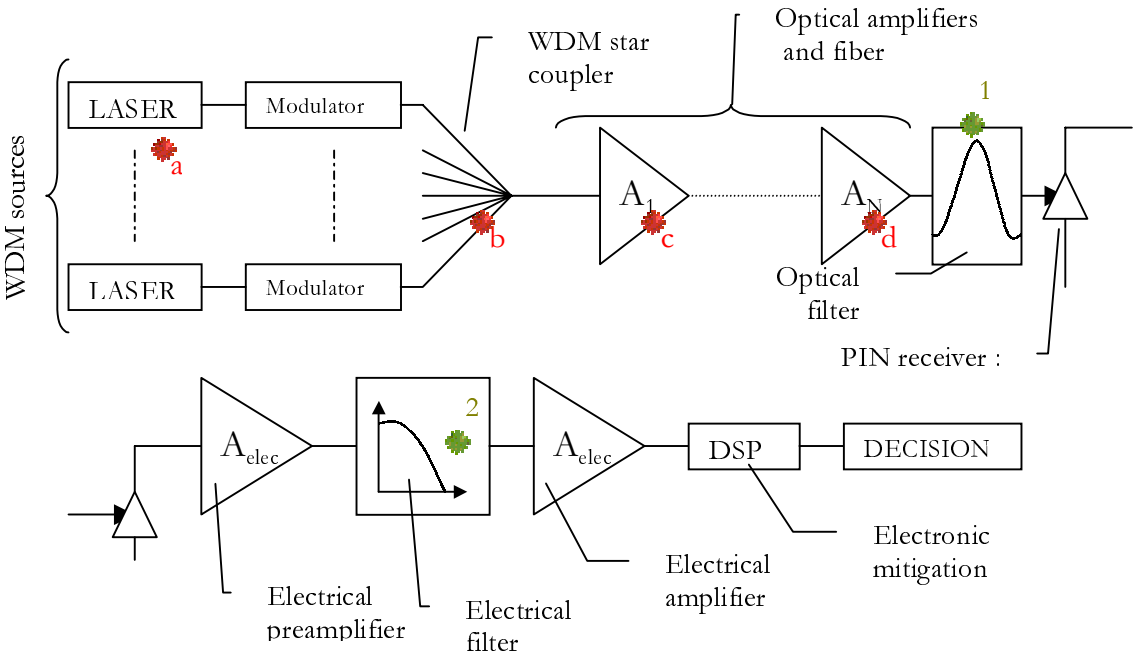

Fig. 57 Fiber Optic Communication Systems

The previous sketch does not give all the components of a real system. For example, the dispersion management elements are not represented. Are drawn, only the noise related devices. We classify them into four categories: thermal noise/dark current, shot noise, amplified emission noise and relative intensity noise.

2.5.1.1 THERMAL NOISE

Thermal noise appears in any electrical device. It results from the random movement of the electrons at finite temperature. The thermal agitation can be considered as a zero mean Gaussian noise, whose variance is given as:

[4] The receiver of an optical system transforms a photon into an electron. The photons of the signal beat with the noise photons. This results in a fluctuation of the received intensity.

$$\sigma_T^2 = \frac{4k_B T}{R_L} F_n \Delta f$$

with k_B the Boltzmann constant, T the temperature, R_L the load resistor, Fn the amplifier noise figure, Δf the band

Eq. 52 Thermal noise variance

It can be reduced by narrowing the electrical filter to the signal bandwidth. It depends also on the characteristics of the front end of the receiver by the noise figure of the amplifier and the load resistor. For this resistor, there exist a tradeoff between time response and noise cancellation.

One may note that this variance is constant and independent of the signal level.

2.5.1.2 SHOT NOISE

Shot noise appears also as an intensity fluctuation in the electrical signal. It is a manifestation of the fact that the electric current consists of a stream of electrons generated at random time. It can be approximated by a Gaussian random process. Its variance is a function of the received current.

$$\sigma_s^2 = 2q\left(I_p + I_d\right)\Delta f$$

q is the elementary charge, Ip the incident current, Id the dark current, Δf the receiver bandwidth

Eq. 53 Shot noise variance

It may be considered also as a consequence of the zero point fluctuation of vacuum. We mention in Eq. 53 the dark current. Because of the receiver design there exists always a negligible current (enhanced for avalanche photodiode) in the circuit even when no light is received. It is usually so small that it is not considered. We follow this approximation.

2.5.1.3 AMPLIFIED SPONTANEOUS EMISSION (ASE)

An optical amplifier amplifies the signal by stimulated emission, like a laser. Optical amplifiers are usually Erbium Doped Fiber Amplifiers (EDFA). From the rare earth element doping, we tailor the desired energy gap, giving the gain profile. Then population inversion is obtained by pumping the fiber with pump LASER at appropriate wavelengths. There exist of course other optical amplifiers, mostly Raman amplifiers. Their characteristics can be studied under the same formalism as EDFA, even if their respective performance are quite different. Consequently, we base our study on the example of EDFA. The main noise source in an EDFA is the spontaneous emission. The population inversion coefficient determines the number of photons spontaneously emitted.

$$F_n = 2n_{sp}\frac{G-1}{G} \approx 2n_{sp}$$

Eq. 54 Noise figure for EDFA,

n_{sp} represents the population-inversion factor, N2/(N2-N1), n_{sp}>1, G represents the gain

Moreover depending on the gain saturation of the amplifier, spontaneous photons are more or less amplified. The ASE co-propagates down the fiber with the signal and is amplified as much as the signal, so that the first amplifier in a chain of amplifiers constitutes the limiting one, as Eq. 55 states it.

$$F_n^{eff} = F_{n1} + \frac{F_{n2}}{G_1} + \frac{F_{n3}}{G_1 G_2} + \ldots$$

Eq. 55 Noise figure of cascaded amplifiers

At the receiver, the light is detected. The noise photons beat with the signal photons, mostly in the carrier. This creates signal dependent noise. The ASE beats also again itself, this can be reduced by filtering the light optically before detection. Of course, ASE beats with other sources of optical noise.

2.5.1.4 RELATIVE INTENSITY NOISE (RIN)

RIN is typically due to LASER fluctuations. It is characterized by its proportionality to the launched power. It may create a BER floor if it reaches an unacceptable value. As any system is designed away from BER floor, it can be neglected.

2.5.1.5 SUMMARY

[25], pp. 404-405 provides the proofs that in an optically amplified detected signal, the noise is due to five processes.

Thermal noise in a saturated optical amplifier regime, as it is a constant electrical noise would create a BER floor. Modern systems are usually not thermal noise limited. Since we assume also RIN, shot noise and dark current to be negligible, we are left with two contributions. They represent the beating of the spontaneous emission with itself and with the signal.

$$\textit{Spontaneous-Sp beating} \qquad \sigma^2_{sp-sp} = (q\eta G F_n)^2 \Delta\nu_{opt}\Delta f \propto \Gamma_0^2 G^2$$

$$\textit{Signal-spontaneous beating} \qquad \sigma^2_{sig-sp} = 2(q\eta G)^2 F_n P_s \Delta f / h\nu \propto \Gamma_1^2 P_{input} G^2$$

Eq. 56 Noise contributions at a pre-amplified receiver

The variances are expressed in terms of the receiver's electrical characteristics. This comes from the fact that, since the decisions take place in the electrical domain, the

noise processes have to be compared to a reference at the receiver, in the electrical domain. An optical SNR does not directly relate to a decision performance. It is, indeed, receiver dependent.

Summarizing, we deduce the total noise distribution expressions:

$$\sigma^2 = \sigma^2_{sp-sp} + \sigma^2_{sig-sp} \qquad \sigma_0^2 = \Gamma_0^2 \qquad \sigma_\alpha^2 = \Gamma_1^2 \alpha + \Gamma_0^2$$

Eq. 57 Approximated noise contribution at the receiver, α is the level of the received signal; $\sigma_{sp\text{-}sp}$ stands for the variance of ASE beating with itself; it is usually smaller than $\sigma_{sig\text{-}sp}$, beating of the signal with ASE

If the system considered is not saturated, we can in fact keep this formalism and introduce all other contributions except RIN. The signal dependent part of the noise can take care of the shot noise whereas the thermal noise appears as a constant noise. The dark current contributes to both variances. In order to consider RIN, we need to introduce a $\alpha^2\Gamma_2$ term.

The case of a saturated system is a bit more complex. The power received is constant. The noise created at the first amplifier decreases with the increase of the launched power (by reduction of the gain). This reduces the global SNR by reducing the optical SNR. Under saturation, the received optical power is constant; consequently all the non-optical noise sources have constant variances. Their associated SNR is constant since the power does not increase. This would result in BER floor. Therefore, we neglect non optical noise sources.

$\sigma^2_{sp\text{-}sp}$ is proportional to the square of the gain, consequently the relative sp-sp SNR decreases as the gain square(e.g. as the inverse of the input power), Eq. 56. $\sigma^2_{sig\text{-}sp}$

is proportional to the product of the square of the gain with the input power. Its relative SNR is therefore related to the inverse of the power. These last two assertions that can seem confusing are repeated using a Q formalism in section 2.5.2.2.

2.5.2 Theoretical bounds to digital electronic mitigation

2.5.2.1 MOTIVATION

System performance depends highly on the receiver design. We aim here at finding a way to reduce the number of parameters necessary to determine DSP performance estimation. First order PMD reduces to γ, DGD and pulse shape. We reduce here the noise receiver dependence to one parameter: ξ.

2.5.2.2 Q FACTOR FORMALISM

Consider a signal distorted by the ISI from the previous bit as in [19]:

$$s_k = x_k + \delta x_{k-1}$$

Eq. 58 ISI due to the previous bit, x_k is the symbol sent at the time k, x_k the distorted signal

This gives rise to the following four level eye,

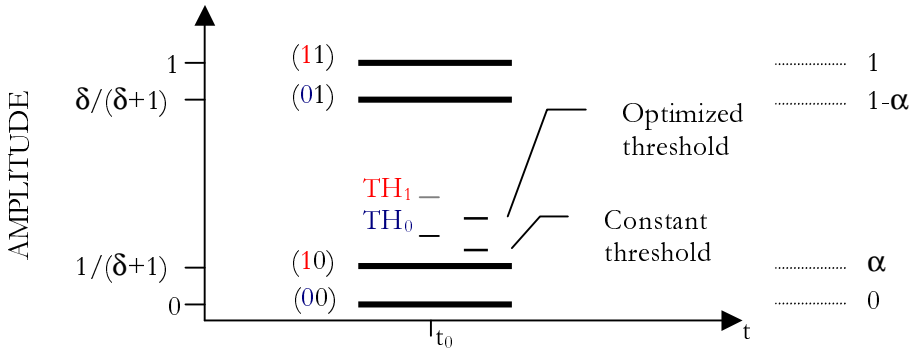

Fig. 58:

Fig. 58 Sketch of the four levels eye, TH1 and TH0 are the thresholds of the NLC

The previous sketch allow us to determine the limiting eye opening for each technique:

- For a threshold kept constant, the pattern (10) is limiting giving a Q factor found in Eq. 59 (b).
- For an optimized threshold, (10) and (01) are limiting. Q is close to Q_c given by the eye opening between these two levels. Eq. 59 (c)
- Finally a NLC, even for equal eye opening, has a limiting case, due to the noise distribution, that is 1 previously received, Eq. 59 (d).

Because of signal dependent noise, equal eye opening does not mean equal relative Q factors.

$$Q_a \equiv \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0} \qquad I_a \equiv \frac{\sigma_0\mu_1 + \sigma_1\mu_0}{\sigma_0 + \sigma_1} = \frac{\Gamma_0\mu_1}{\Gamma_0 + \left(\Gamma_0^2 + \Gamma_1^2\right)^{1/2}} = \frac{1}{1 + \left(1 + \xi^2\right)^{1/2}} \qquad \xi = \frac{\Gamma_1}{\Gamma_0}$$

$$Q_b \equiv \frac{I_a - I_\alpha}{\sigma_\alpha} = \frac{\left(\frac{1}{1 + \left(1 + \xi^2\right)^{1/2}} - \alpha\right)}{\left(\Gamma_0^2 + \alpha\Gamma_1^2\right)^{1/2}} \quad Q_c \equiv \frac{I_{1-\alpha} - I_\alpha}{\sigma_{1-\alpha} + \sigma_\alpha} = \frac{(1 - 2\alpha)}{\left(\Gamma_0^2 + (1-\alpha)\Gamma_1^2\right)^{1/2} + \left(\Gamma_0^2 + \alpha\Gamma_1^2\right)^{1/2}}$$

$$Q_d \equiv \frac{I_1 - I_\alpha}{\sigma_1 + \sigma_\alpha} = \frac{(1-\alpha)}{\left(\Gamma_0^2 + \Gamma_1^2\right)^{1/2} + \left(\Gamma_0^2 + \alpha\Gamma_1^2\right)^{1/2}}$$

Eq. 59 Q factor.(a) Q factor and threshold definition, no ISI (b) constant threshold, ISI (c) optimized threshold, ISI (d)NLC, ISI

We consider only the worst Q factor out of 2 cases because regardless of the other decision, the BER lies between the BER given by this Q and half this BER. We see in Eq. 59 that all the cumbersome parameters we had in Eq. 52-Eq. 56 are summarized under

the single parameter ξ. It represents the ratio of the standard deviation in the ones to the standard deviation in the zero, for an ISI free case of Q=6.

In Eq. 56, if we assume saturation, multiplying the input power by β, because it reduces the gain of the first amplifier (limiting noise source for a chain of amplifiers) by β, divides $\sigma^2_{sp\text{-}sp}$ by β^2 and $\sigma^2_{sp\text{-}sig}$ by β.

2.5.2.3 PROOF

The power penalty is the increase in power necessary to reach a reference performance (BER = 10^{-9} or Q=6). We look for β such that Q for the considered case and for a power increased by β is equal to the initial Q = 6. For example, Eq. 60 gives the power penalty for a constant threshold.

$$Q_a = \frac{1}{\Gamma_0 + \left(\Gamma_0^2 + \Gamma_1^2\right)^{1/2}} = Q_{a,\beta} = \frac{\left(\frac{1}{1+\left(1+\xi^2\right)^{1/2}} - \alpha\right)}{\Gamma_0/\beta + \left(\Gamma_0^2/\beta^2 + \alpha\Gamma_1^2/\beta\right)^{1/2}} \Rightarrow \left(1 - \left(1+\left(1+\xi^2\right)^{1/2}\right)\alpha\right)\beta = 1 + \left(1+\alpha\beta\xi^2\right)^{1/2}$$

Eq. 60 Power penalty equation for the constant threshold case

Using similar equalities, we find:

$$\beta_{TH=Cst} = \frac{\alpha\xi^2 + \left((\alpha\xi^2)^2 + 4(1-\alpha R)^2\right)^{1/2}}{2(1-\alpha R)^2}$$

$$\beta_{TH=opti} = \frac{\xi^2 + 2\left(R^2 + \alpha(1-\alpha)(\xi^4 - 4R^2)\right)^{1/2}}{R^2(1-2\alpha)^2} \qquad R = \left(1 + \left(1+\xi^2\right)^{1/2}\right)$$

$$\beta_{NLCi} = \frac{(1+\alpha)\xi^2 + 2\left(\alpha\xi^4 + (1-\alpha)^2\right)^{1/2}}{R^2(1-\alpha)^2}$$

Eq. 61 Power penalty versus α for constant threshold, optimized threshold and NLC

We can see on Fig. 59 the power penalty as a function of ISI, given as α.

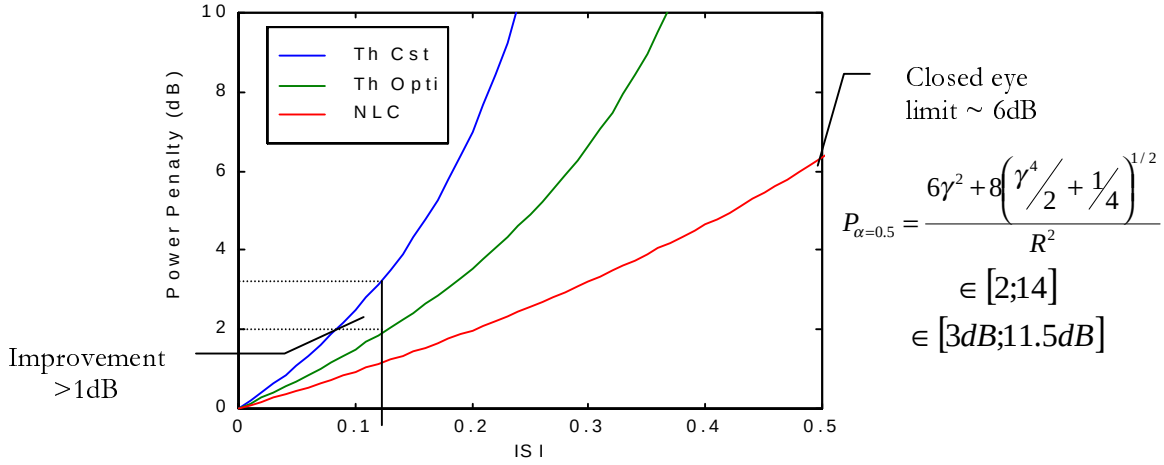

Fig. 59 Power Penalty (dB) as a function of ISI, given as α, γ^2=3

We confirm a consistent amelioration provided by an adaptive threshold versus a threshold kept constant, [26]. It improves the power penalty by more than 1 dB as soon as $\alpha>0.1$.

Moreover, since systems use a constant threshold, compensation skill performance cannot be estimated using only the eye opening. It requires an analysis considering the limit set by the constant threshold.

On the other hand, we find a theoretical power penalty for the NLC for a closed eye of 6 dB. This number depends upon the considered system. It varies from 3 dB to 11.5 dB; usually around 6 dB, which may be too large for a relevant mitigation.

This is not the only consequence of signal dependent noise. The optimum threshold does not lie halfway between expected value of 0 and 1, as it does for the thermal noise limit. Consequently, the algorithm of [19] that was based on the pattern dependent estimation of the average of these two values does not work. The optimum threshold is a non-linear function of these levels discarding any easy modification of this technique. This elegant adaptation method cannot optimally track fluctuating ISI.

2.5.3 Experimental verification

We verify our model validity by experiments and simulations.

2.5.3.1 ISI MAPPING

Our derivation of the power penalty uses α, the ISI level, rather than a PMD characterization. It can therefore be adapted to any impairment as soon as it is translated into ISI. It allows furthermore a system free approach, independent for the different filter and noise parameters.

We obtain the one to one mapping between ISI and PMD by extrapolating data. We applied to previous data [24] our equation in the thermal noise limit, ξ=0.

Assuming then ξ=3½ and saturation, we predicted the NLC performance by using Eq. 61.

2.5.3.2 SETUP

In order to verify the saturation model, we ran experiments at 5 Gb/s using the set up shown on Fig. 60. We used an external modulator and a first order PMD emulator, inducing delays between -50 and 250 ps and providing any given splitting ratio. The light was detected using a saturated preamplifier. We had 4 times more noise in the 1 than in the 0 for the reference case (value set by the bandwidth of our optical filter,), giving $\xi^2=3$.

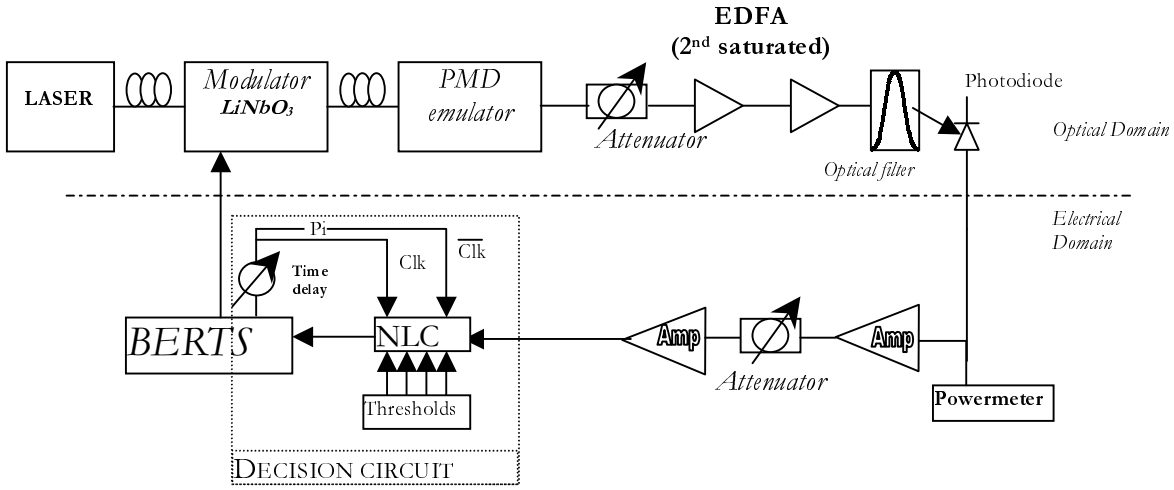

Fig. 60 Experimental set up at 5 Gb/s, PBRS 2^{31}-1

The current out of the photodiode was fed into a 2-bit NLC used either with all thresholds kept constant, all thresholds equal or as a decision feedback equalizer.

2.5.3.3 RESULTS

Using a mapping between PMD and ISI from previous data [24] (thermal noise limit), we predicted penalty versus PMD. This estimation matched our experimental results, Fig. 61.

The mismatch between the 6 dB limit presented on Fig. 59 and the 8 dB of Fig. 61 appears as a consequence of the non-perfect device we used. We took this factor into

account using experimental data rather than simulation data of [24]. Experimental data (dashed lines) lay above theoretical expectations but within the measurement uncertainty. The difference of 1 dB for the NLC at 200 ps and above may be due to a slightly incorrect input polarization into the PMD emulator.

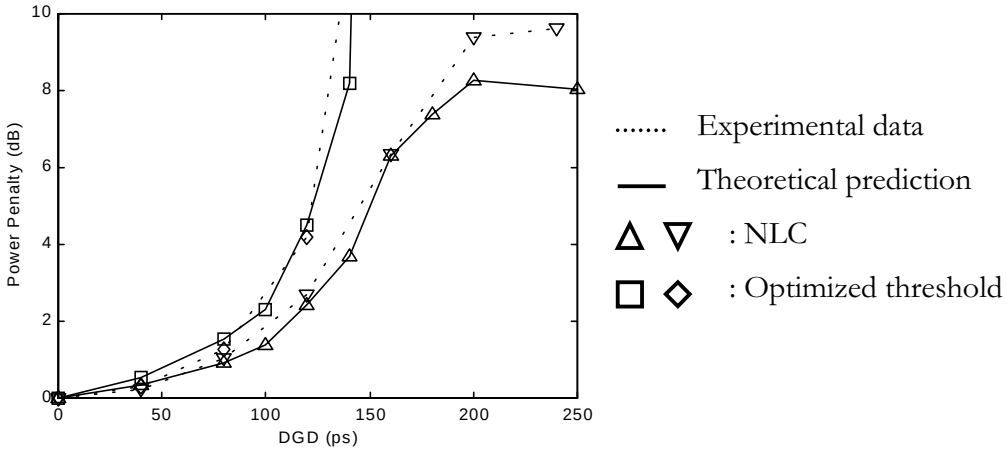

Fig. 61 Comparison of experimental data to theoretical expectation:

Power penalty versus DGD, for first order PMD, splitting ratio of ½ at 5 Gb/s

2.5.3.4 RESULTS

Our model being experimentally confirmed, we ran simulations to obtain a mapping of the DSP performance similar to the one we built for the thermal noise limit. We do not compare our techniques only to a normal receiver of constant threshold but also to an adaptive unique threshold receiver. It would have been possible, as in [26], to consider a threshold margin.

A threshold margin is based on the observation that the threshold is always set to the best performance for the ISI free case. Then, we allow a certain range of ISI equivalent to a 1 dB penalty. However the function relating the penalty to the ISI is

monotonic, Fig. 62 (a). We use only one side of the operation curve. By setting the threshold a bit higher than the optimal value, we obtain a lower ISI free performance but we gain robustness[5], Fig. 62 (b).

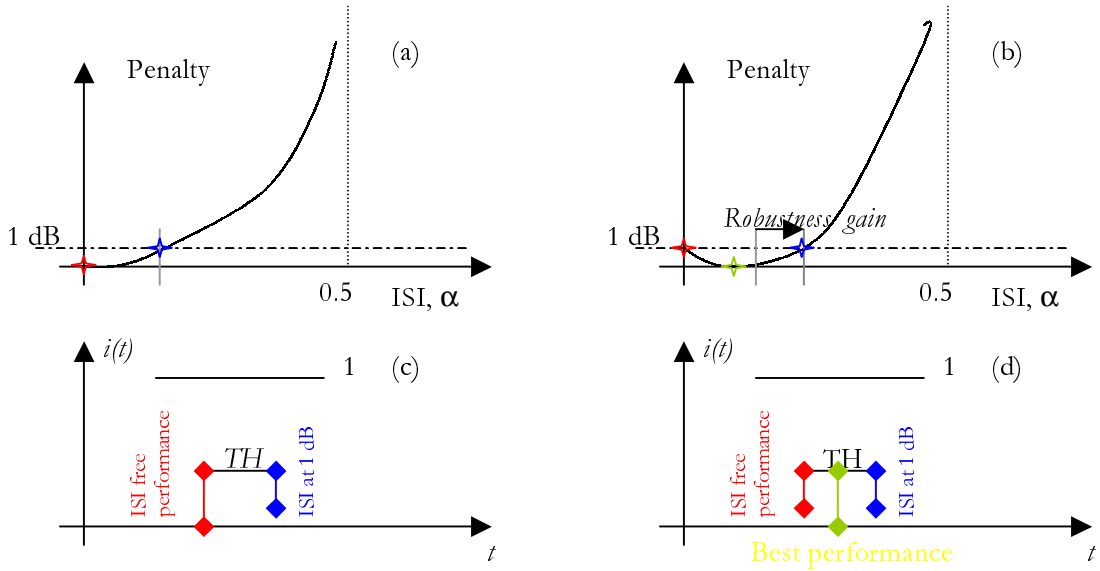

Fig. 62 Robustness enhancement of non-optimal threshold

[5] Greater robustness means that our system is less sensitive to ISI. It 'can take' a higher level of ISI before reaching unacceptable performance.

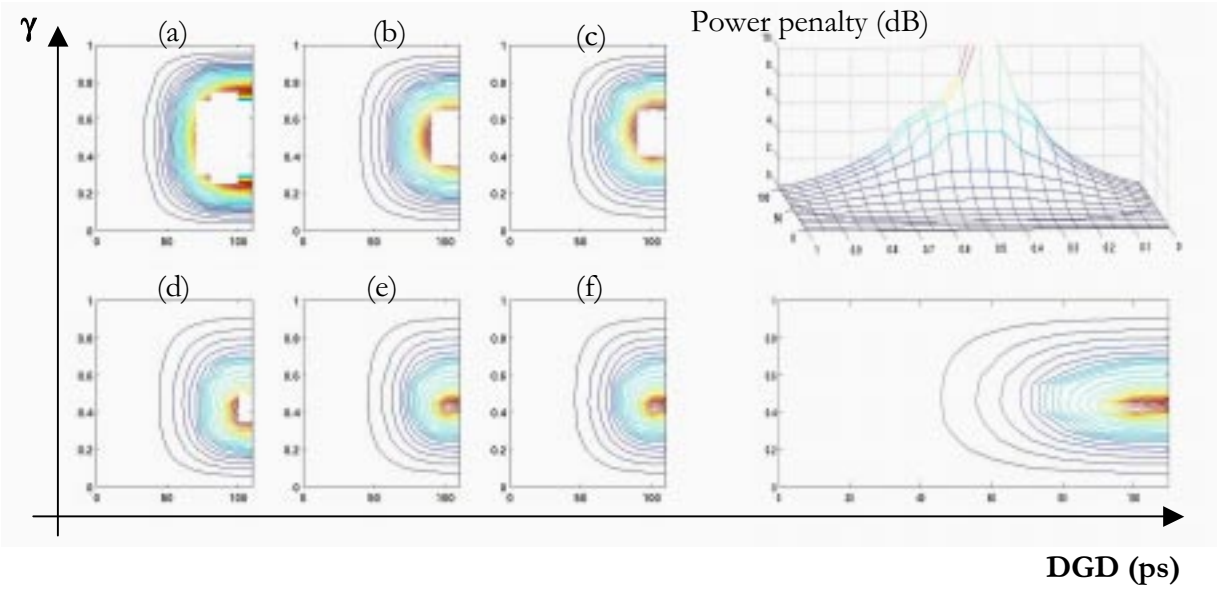

Fig. 63 First order PMD dependent penalty (line = +0.5 dB),

(a) Constant threshold (b)Adaptive single threshold (c) NLC (d) FDE (e) FDE+NLC (f)FDE+NLC, four independent clocks

Fig. 63 presents the first order PMD induced penalty for different decision techniques at 10 Gb/s. It shows that in its range of action (ratio>0.5), NLC exhibits a maximum 6dB penalty as expected. Over the global range of first order PMD, the compound exhibits a maximum 11 dB penalty. Again, this worst case appears for a ratio equal to 0.4. This very large penalty comes from, first the poor equalizer we consider (only 3 taps), second the harmful signal dependent noise enhancement due to the FDE and the asymmetry of the noise.

We can note that the DSP techniques provide a very strong improvement with respect to the normal receiver. We can double the acceptable DGD. However, there is little difference between a receiver using an adaptive threshold and even the compound NLC+ FDE, if we consider the 1 dB point. The difference increases dramatically as the

DGD approaches a full bit period but the penalty may be already unacceptable in this range.

2.5.4 Conclusion

We studied the performance of different DSP techniques for the case of an ASE noise limited system corresponding to most of the systems of interest. We obtained a closed form performance bound for the NLC as a function of ISI. The system dependence was reduced to the parameter ξ that represents the ratio of the noise standard deviations for the reference case.

We experimentally verified this bound for a specific case. We generalized these observations by a computer simulation. It provided us finally with a PMD induced penalty map for the different techniques.

It has been shown that DSP, the way it has been implemented so far, does not lead to acceptable performance over the range where it was expected to work. This results from the strong asymmetry of the noise. First, it increases the noise enhancement of the FDE. Second, in the case of a three level eye (closed), the two subsequent eyes have different associated Q, so that the NLC does not act optimally. Finally, the power penalty required to open an eye in case of ASE is nearly twice the power for the thermal noise limit, due to signal dependent noise. It induces a square root dependence of the Q factor on the power rather than a simple dependence.

Consequently, other solutions have to be sought in order to reach acceptable performance. It is possible to use a fractional equalizer to reduce the noise enhancement and to shape the signal eye to optimize the NLC action, as proposed in [20]. This does not appear sufficient since it cannot push the 1 dB penalty further than 0.7 T for the first order, where T stands for the bit period. Furthermore, it is not likely to give a better penalty than 6 dB for a closed eye. Finally, it requires a complex control IC because it cannot use the normal adaptation algorithms since the FDE is aimed at shaping the eye for the NLC and not for shaping the best eye. It seems to require a performance management based on the BER, which is not yet available.

Because the noise issues appear with the square law detection, it may be useful to consider a first mitigation in the optical domain, enhanced by an optimization of the receiver PMD robustness.

2.6 HYBRID SOLUTIONS

We consider now solutions that ameliorate PMD effect not only in the electrical domain but also in the optical domain. We expect from this association enhanced performance.

2.6.1 PSP launched and spectrum shaping

A common and demonstrated technique to overcome PMD is to launch the input SOP on one of the PSP. We suffer then only the effects of higher order PMD. It is possible to reduce the effects of higher order by reducing the spectrum width. This may be done using Single Side Band (SSB) transmission. A phase modulator associated with a usual amplitude modulator can approximate the Hilbert transform necessary to remove one side band.

The great improvement provided by SSB transmission has already been proved [27]. We propose to increase its efficiency by using the NLC. We have notice that according to the PSP, we launch, we can induce a time reversal of our pulse shape, (section 1.3.4.4). Consequently high order PMD induced ISI due to future bits can be transferred to previous bits. It means that we do not need anymore a compound FDE+NLC but a single NLC. The only extra degree of freedom is discrete and can be checked using the RF spectrum of the received signal. It should be possible to make the NLC adaptive without a control IC if we assume some range of ISI.

2.6.2 Diversity techniques

Based on the nature of PMD it is possible to use diversity to reduce either first order effects or second order effects:

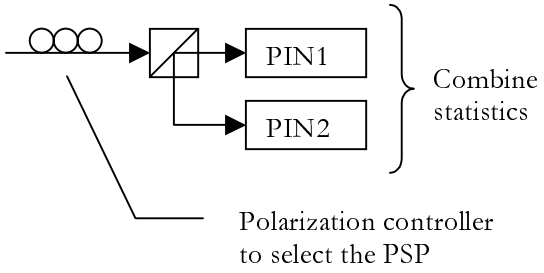

Fig. 64 First order diversity technique

Instead of matching input SOP and input PSP, as done in a PSP launched technique; we can try to get the signal in each output PSP. Because output PSP are orthogonally polarized, they can be split using a polarization controller and a beam splitter. Each light pulse is detected independently. We can apply on each signal a certain decision technique such as NLC and combine the statistics of each decision to get a better information estimation. Furthermore, it reduces the ISI related noise by reducing the detected optical level.

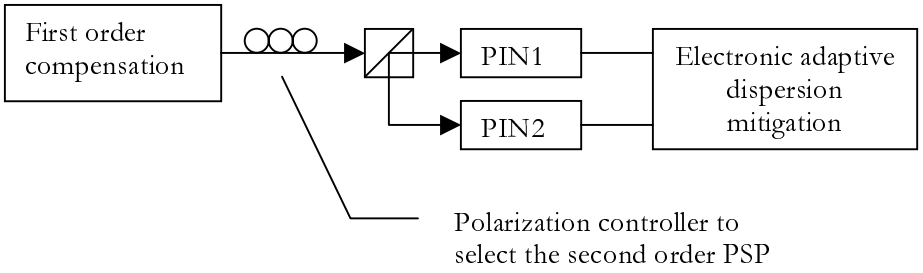

Fig. 65 Second order diversity technique

Using the higher order model of [13] presented in section 1.3.4.4, electronic mitigation of higher orders can be done. A first stage of the compensator mitigates first order. Then using again a polarization controller in association with a beam splitter, the signal in each 2^{nd} order PSP is detected. Each of these signals is impaired by an equivalent chromatic dispersion, whose effects can be ameliorated using electronic mitigation.