

High-order Statistical Equalizer for Nonlinearity Compensation in Dispersion-managed Coherent Optical Communications

Koike, Akino, T.; Parsons, K.; Kojima, K.; Duan, C.; Yoshida, T.; Sugihara, T.; Mizuocho, T.

TR2012-055 June 2012

Abstract

Fiber nonlinearity has become a major limiting factor to realize ultra-high-speed optical communications. We propose a fractionally-spaced equalizer which exploits a trained high-order statistics to deal with data-pattern dependent nonlinear impairments in fiber-optic communications. The computer simulation reveals that the proposed 3-tap equalizer improves Q-factor by more than 2 dB for long-haul transmissions of 5,230 km distance and 40 Gbps data rate. We also demonstrate that the joint use of a digital backpropagation (DBP) and the proposed equalizer offers an additional 1-2 dB performance improvement due to the channel shortening gain. A performance in high-speed transmissions of 100 Gbps and beyond is evaluated as well.

Optics Express

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

High-order statistical equalizer for nonlinearity compensation in dispersion-managed coherent optical communications

Toshiaki Koike-Akino,^{1,*} Chunjie Duan,¹
Kieran Parsons,¹ Keisuke Kojima,¹
Tsuoyoshi Yoshida,² Takashi Sugihara,² and Takashi Mizuochi²

¹ MERL, 201 Broadway, Cambridge, MA 02139, USA

² Mitsubishi Electric Corporation, 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan

*koike@merl.com

Abstract: Fiber nonlinearity has become a major limiting factor to realize ultra-high-speed optical communications. We propose a fractionally-spaced equalizer which exploits a trained high-order statistics to deal with data-pattern dependent nonlinear impairments in fiber-optic communications. The computer simulation reveals that the proposed 3-tap equalizer improves Q-factor by more than 2 dB for long-haul transmissions of 5,230 km distance and 40 Gbps data rate. We also demonstrate that the joint use of a digital backpropagation (DBP) and the proposed equalizer offers an additional 1–2 dB performance improvement due to the channel shortening gain. A performance in high-speed transmissions of 100 Gbps and beyond is evaluated as well.

© 2012 Optical Society of America

OCIS codes: (060.1660) Coherent communication; (060.4370) Nonlinear optics, fibers.

References and links

1. J. Renaudier, G. Charlet, P. Tran, M. Salsi, and S. Bigo, "A performance comparison of differential and coherent detections over ultra long haul transmission of 10Gb/s BPSK," OFC'07, OWM1 (2007).
2. A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," J. Lightwave Technol. **28**, 423–433 (2010).
3. X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," Opt. Express **16**, 880–888 (2008).
4. E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," J. Lightwave Technol. **26**, 3416–3425 (2008).
5. E. Ip, N. Bai, and T. Wang, "Complexity versus performance tradeoff in fiber nonlinearity compensation using frequency-shaped, multi-subband backpropagation," OFC'11, OThF4 (2011).
6. T. Yoshida, T. Sugihara, H. Goto, T. Tokura, K. Ishida, and T. Mizuochi, "A study on statistical equalization of intra-channel fiber nonlinearity for digital coherent optical systems," ECOC'11, Tu.3.A.1 (2011).
7. W. Yan, Z. Tao, L. Dou, L. Li, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Low complexity digital perturbation back-propagation," ECOC'11, Tu.3.A.2 (2011).
8. F. P. Guimar, J. D. Reis, A. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," ECOC'11, Tu.6.B.1 (2011).
9. N. Alić, G. C. Papen, R. E. Saperstein, L. B. Milstein, and Y. Fainman, "Signal statistics and maximum likelihood sequence estimation in intensity modulated fiber optic links containing a single optical preamplifier," Opt. Express **13**, 4568–4579 (2005).

10. Y. Cai, D. G. Foursa, C. R. Davidson, J. X. Cai, O. Sinkin, M. Nissov, and A. Pilipetskii, "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmissions," OFC'10, OTuE1 (2010).
11. T. Koike-Akino, C. Duan, K. Kojima, K. Parsons, T. Yoshida, T. Sugihara, and T. Mizuochi, "Fractionally-spaced statistical equalizer for fiber nonlinearity mitigation in digital coherent optical systems," OFC'12 OTh3C.3 (2012).
12. J. B. Anderson and S. Mohan, "Sequential coding algorithms: a survey and cost analysis," IEEE TCOM **32**, 169–176 (1984).
13. A. Azzalini and A. Capitanio, "Statistical applications of the multivariate skew normal distribution," J. Royal Stat. Society, **61**, 579–602 (1999).
14. G. H. Golub and C. F. Van Loan, *Matrix Computations, 3rd Ed.* (Johns Hopkins University Press, 1996).
15. C. Duan, K. Parsons, T. Koike-Akino, R. Annavajjala, K. Kojima, T. Yoshida, T. Sugihara, and T. Mizuochi, "A low-complexity sliding-window turbo equalizer for nonlinearity compensation," OFC'12, JW2A.59 (2012).
16. I. B. Djordjevic, L. L. Minkov, and H. G. Batshon, "Mitigation of linear and nonlinear impairments in high-speed optical networks by using LDPC-coded turbo equalization," IEEE J. Sel. Areas Commun. **26**, 73–83 (2008).
17. H. G. Batshon, I. B. Djordjevic, L. Xu, and T. Wang, "Iterative polar quantization based modulation to achieve channel capacity in ultra-high-speed optical communication systems," IEEE Photon. J. **2**, 593–599 (2010).

1. Introduction

Digital coherent optical transmissions have a potential to increase data rates with dual-polarized phase-shift keying (DP-PSK) and quadrature-amplitude modulation (DP-QAM). However, fiber nonlinearity can significantly degrade the advantage of coherent transmission systems over conventional non-coherent transmission systems as such spectrally efficient modulation formats require higher signal-to-noise power ratio whereas higher signal power causes more significant nonlinear distortion [1, 2]. Therefore, to mitigate fiber nonlinearity has been of great importance in optical communication researches.

Recently, it was shown that digital back-propagation (DBP) proposed in [3, 4] offers a substantial performance gain to compensate deterministic nonlinear effects. However, the DBP requires high-complexity processing based on split-step Fourier methods (SSFM). Although there exist several reduced-complexity methods [5–8], the performance is susceptible to the stochastic distortion including the amplified spontaneous emission (ASE) noise and the polarization mode dispersion (PMD). Moreover, the SSFM parameters used in DBP should be manually adjusted to achieve the best performance in general.

We focus on an alternative method based on statistical sequence equalizers (SSE), studied in [9, 10], which mitigates data-pattern-dependent nonlinearity. In the scheme presented in [10], the first-order statistics (i.e., the mean of the the nonlinear distortion) is trained first for some possible data patterns, and the pattern matching is performed to equalize the nonlinearity. The statistical sequence equalizer achieves good performance with low complexity for short-memory channels, and could be combined with other methods including DBP and frequency-domain equalizer (FDE).

In [11], we have extended the original SSE in several directions: i) we proposed the use of second-order statistics (i.e., not only mean but also covariance), ii) we adopted a fractionally-spaced signal processing, and iii) we used an excess window size for pattern matching. We have shown that the proposed SSE offers comparable performance to DBP in 40 Gbps DQ-DQPSK and differential QPSK (DP-DQPSK).

In this paper, we further introduce a cascaded equalizer which employs a reduced-complexity DBP to shorten the channel memory in conjunction with the SSE to suppress the residual nonlinearity. In addition, more detail performance analyses are provided: i) description of low-complexity statistics updating, ii) the impact of tap lengths, iii) the difference with and without fractionally-spaced processing iv) the case of 16QAM modulations, v) the effect of reduced-state SSE with M-algorithm [12], vi) the computational complexity comparisons, and vii) higher-speed transmissions of 100 Gbps and beyond.

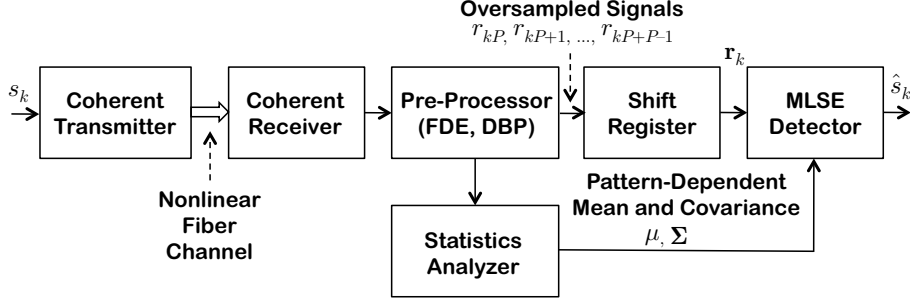


Fig. 1. Fractionally-spaced statistical sequence equalizer (FS-SSE) which exploits high-order statistics of nonlinearity distortion in coherent fiber-optic systems.

Through the computer simulations, we obtain more than 2 dB improvement using the proposed equalizer for coherent optical communications with 40 Gbps non-return-to-zero (NRZ) DP-QPSK or DP-DQPSK signals after 5,230 km transmissions. The achieved performance is better than that of the DBP in low local dispersion channels, and is comparable to the DBP in high local dispersion channels. We also demonstrate that the joint use of DBP and SSE enjoys an additional 1–2 dB gain. It is verified that M-algorithm is effective to drastically reduce the computational complexity.

2. Nonlinear equalizer

Figure 1 shows the schematic of the proposed statistical sequence equalizer in coherent optical communications. The digital data s_k at the time instance k is transmitted through the nonlinear fiber by using the coherent optical transceiver. The received data may be first processed by several pre-equalization units, including timing recovery, FDE for chromatic dispersion compensation, PMD compensation, DBP for nonlinear compensation, and so on. The pre-processed signal is fed into a shift register which accepts fractionally-spaced (or, oversampled) data. Those signals are also analyzed to obtain the statistics of the fiber nonlinearity. The oversampled data is then equalized by a maximum-likelihood sequence estimation (MLSE) detector which employs the Viterbi algorithm based on the fiber statistics.

2.1. Statistical sequence equalizer (SSE)

As discussed in [10], intra-channel nonlinear distortion highly depends on the transmitted data pattern. The statistical sequence equalizer (SSE) in [10] first acquires such data-pattern-dependent distortion characteristics by averaging the received sequence with training data or on-line learning process. The trained mean signals are then used to equalize the nonlinear distortion by searching for a pattern with the minimum Euclidean distance from the received sequence.

Figure 2 illustrates an example of the received I-Q (In-phase/Quadrature-phase) constellations distorted by the fiber nonlinearity for a launching power of -4 dBm over 5,230 km long-haul transmissions (corresponding to 8 dBQ with a 1-tap phase compensator). Here, we plot 4,000 sample-spaced receiving signals of a random DQPSK sequence. To show the data-pattern-dependent nonlinearity, we also present the averaged signal points of the k -th received signal $\mathbb{E}[r_k]$ conditioned on the consecutive 3-symbol data pattern $\mathbf{s}_k = [s_{k-1}, s_k, s_{k+1}]$, where $\mathbb{E}[\cdot]$ denotes the expectation. There are 64 points in the figure since the total number of the different data patterns is $4^3 = 64$ from $\mathbf{s}_k = [0, 0, 0]$ to $[3, 3, 3]$ for a 3-tap data sequence. It can

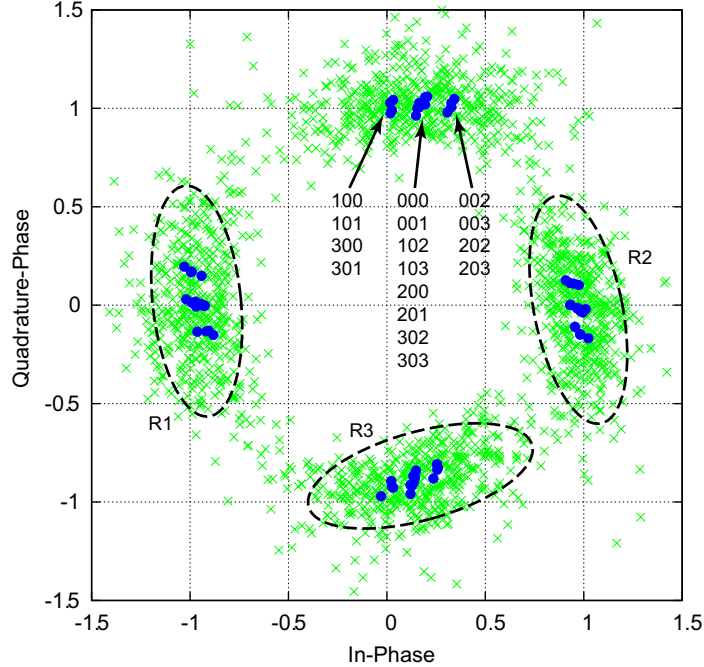


Fig. 2. Examples of received I-Q constellation with nonlinear distortion (launching power: -4 dBm, wavelength: 1551 nm, fiber distance: 5,230 km).

be seen from Fig. 2 that the mean of the received signal (e.g. for the data pattern $[0, 0, 0]$) differs from that for another (e.g. $[0, 0, 2]$). The SSE exploits such a data-pattern-dependent distortion to compensate the nonlinearity.

2.2. High-order statistics

In this paper, we propose the use of higher-order statistics (variance, skewness, *etc.*) in addition to the first-order statistics (mean) to mitigate residual nonlinear noise as well. As shown in Fig. 2, the residual (out-of-memory) nonlinear distortion performs as an effective noise which also depends on the data pattern; specifically, the distribution of the residual distortion around the region “R1” is different from that of the region “R3.” More importantly, the distribution is not circularly symmetric (i.e., ellipsoidal rather than circular).

To take the pattern-dependent residual nonlinearity into account properly, we introduce a model based on a circularly-asymmetric Gaussian distribution. We consider a window size of N receiving samples centered around the target transmission data to establish an empirical statistics. Let $\mathbf{r}_k = [r_{k-\lfloor(N-1)/2\rfloor}, \dots, r_k, \dots, r_{k+\lfloor N/2\rfloor}]^T \in \mathbb{C}^N$ denote the received signal sequence in the shift register of window size N , where $\lfloor \cdot \rfloor$, $[\cdot]^T$, and \mathbb{C} denote a floor function, a transpose operation, and a complex-number set, respectively. A statistics analyzer obtains the empirical mean vector and the covariance matrix of the received signals for each transmission data pattern $\mathbf{s} = [s_{k-\lfloor(M-1)/2\rfloor}, \dots, s_k, \dots, s_{k+\lfloor M/2\rfloor}] \in \mathbb{N}^M$, where \mathbb{N} denotes the natural number set (positive integers).

Letting \mathbf{s} be one of the 4^M possible data patterns (for 4-ary data), the empirical mean vectors

$\boldsymbol{\mu}(\mathbf{s}) \in \mathbb{R}^{2N}$ and the covariance matrices $\boldsymbol{\Sigma}(\mathbf{s}) \in \mathbb{R}^{2N \times 2N}$ are expressed as follows:

$$\boldsymbol{\mu}(\mathbf{s}) = \frac{1}{\mathcal{N}(\mathbf{s})} \sum_{j: \mathbf{s}_j = \mathbf{s}} \mathbf{r}'_j, \quad (1)$$

$$\boldsymbol{\Sigma}(\mathbf{s}) = \frac{1}{\mathcal{N}(\mathbf{s}) - 1} \sum_{j: \mathbf{s}_j = \mathbf{s}} (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s})) (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s}))^T, \quad (2)$$

where we define $\mathbf{r}'_j \triangleq [\Re[\mathbf{r}_j^T], \Im[\mathbf{r}_j^T]]^T \in \mathbb{R}^{2N}$. Here, $\mathcal{N}(\mathbf{s})$ is the total number of occurrences that the data pattern \mathbf{s} appeared in the past. The notations \mathbb{R} , $\Re[\cdot]$ and $\Im[\cdot]$ are a real-number set, the element-wise real part and imaginary part operations, respectively. The main reason to expand the complex-valued vector \mathbf{r}_j to a double-size real-valued vector \mathbf{r}'_j is to model the circular asymmetry illustrated in Fig. 2.

With the trained statistics, the expected likelihood of the received signals \mathbf{r}_k given a data pattern \mathbf{s} is calculated by the circularly-asymmetric Gaussian model as follows:

$$\Pr(\mathbf{r}_k | \mathbf{s}) = \frac{1}{\sqrt{\det[2\pi\boldsymbol{\Sigma}(\mathbf{s})]}} \exp\left(-\frac{1}{2}(\mathbf{r}'_k - \boldsymbol{\mu}(\mathbf{s}))^T \boldsymbol{\Sigma}(\mathbf{s})^{-1} (\mathbf{r}'_k - \boldsymbol{\mu}(\mathbf{s}))\right). \quad (3)$$

Its log-likelihood is then expressed as

$$\ln \Pr(\mathbf{r}_k | \mathbf{s}) = -\frac{1}{2}(\mathbf{r}'_k - \boldsymbol{\mu}(\mathbf{s}))^T \boldsymbol{\Sigma}(\mathbf{s})^{-1} (\mathbf{r}'_k - \boldsymbol{\mu}(\mathbf{s})) - \frac{1}{2} \ln \det[2\pi\boldsymbol{\Sigma}(\mathbf{s})]. \quad (4)$$

Note that it reduces to a simplified function of the Euclidean distance, $\|\mathbf{r}'_k - \boldsymbol{\mu}(\mathbf{s})\|$, when no information of the covariance is available. The benefit of the second-order statistics is twofold: i) less-noisy samples are prioritized via diagonal variance information and ii) correlated nonlinear noise is effectively whitened via off-diagonal correlation information. The extension to the use of the third-order statistics (skewness) is rather straightforward by using the multivariate skew-normal distribution [13].

2.3. Statistics updating

The empirical mean in Eq. (1) and the empirical covariance in Eq. (2) can be updated periodically or continuously with a pre-defined training sequence or a hard-decision data. To track the time-varying statistics, one can use an exponentially-weighted mean and covariance as $\boldsymbol{\mu}(\mathbf{s}) \leftarrow \nu \boldsymbol{\mu}(\mathbf{s}) + \mathbf{r}'_j$ and $\boldsymbol{\Sigma}(\mathbf{s}) \leftarrow \nu \boldsymbol{\Sigma}(\mathbf{s}) + (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s})) (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s}))^T$ with an appropriate normalization of $(1 - \nu)$ where $0 < \nu < 1$ is referred to as a forgetting factor. Note that the determinant and the inverse of the covariance matrix, those of which are required for the likelihood calculations as in Eq. (4), can be efficiently updated by the Sherman-Morrison formula [14] as follows:

$$\ln \det[\boldsymbol{\Sigma}(\mathbf{s})] \leftarrow \ln \det[\boldsymbol{\Sigma}(\mathbf{s})] + 2N \ln \nu + \ln c, \quad (5)$$

$$\boldsymbol{\Sigma}(\mathbf{s})^{-1} \leftarrow \frac{1}{\nu} \boldsymbol{\Sigma}(\mathbf{s})^{-1} - \frac{1}{c} \mathbf{b} \mathbf{b}^T, \quad (6)$$

where we define

$$\mathbf{b} \triangleq \frac{1}{\nu} \boldsymbol{\Sigma}(\mathbf{s})^{-1} (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s})), \quad c \triangleq 1 + \mathbf{b}^T (\mathbf{r}'_j - \boldsymbol{\mu}(\mathbf{s})). \quad (7)$$

It reduces the computational complexity from the cubic order $\mathcal{O}[(2N)^3]$ to the square order $\mathcal{O}[(2N)^2]$ for the matrix inversion, where $\mathcal{O}[\cdot]$ denotes the complexity order.

2.4. Excess window size

The window size M for the transmission data \mathbf{s}_k should be optimized to deal with the memory length of the fiber channel. However, the total number of possible patterns increases exponentially with the window size M . Hence, we may need to use a restricted window size in practice, for example $M = 3$ taps. On the other hand, the computational complexity just increases linearly with the window size N for the receiving data \mathbf{r}_k . We propose to use an excess window size, where we can use $N > M$ to enhance the performance. Doing so, we can keep the computational complexity low while a longer channel memory is considered with its cross correlation in $\mathbf{\Sigma}(\mathbf{s})$.

2.5. Fractionally-spaced processing

Furthermore, we introduce a fractionally-spaced processing to improve the performance by exploiting the correlation over the symbol transition. Let P be an oversampling factor. The received signal sequence is stored as $\mathbf{r}_k = [r_{kP-\lfloor(N-1)/2\rfloor}, \dots, r_{kP}, \dots, r_{kP+\lfloor N/2\rfloor}]^T \in \mathbb{C}^N$ in the N -sample shift register. The fractionally-spaced statistical sequence equalizer (FS-SSE) has an advantage especially when the transceiver filters have an inter-symbol interference due to the non-ideal Nyquist filtering. In addition, the symbol timing error is absorbed by oversampling.

2.6. Cascading with DBP

Since the proposed equalizer cannot use a large window tap size M due to the complexity issue in practice, an FDE to compensate the chromatic dispersion is useful as a pre-processing unit to shorten the effective channel memory. To shorten the channel memory more effectively, we can adopt some nonlinear compensation techniques including the DBP [3,4] as a pre-processor. For such a cascaded use, the DBP partly inverts the fiber nonlinearity, and the residual distortion with limited memory is dealt with by the FS-SSE. As we can see later, by using a reduced-complexity DBP with a small number of steps for SSFM computations, an equalizer cascaded with DBP and FS-SSE can enjoy a significant performance improvement while the overall complexity is maintained to be reasonably low for practical applications.

Although we focus on the hard-decision MLSE equalizer in this paper, the equalizer can be readily extended to a soft-decision equalizer including a soft-output Viterbi algorithm (SOVA) and a maximum *a posteriori* (MAP) equalizer. In [15], the authors extended the proposed statistical equalizer to a low-complexity MAP receiver for a turbo equalization [16, 17] in the system employing a soft-input soft-output decoder for forward error correction (FEC) codes.

3. Performance evaluations

3.1. Fiber plant configuration

For simulations, we use the fiber link configuration corresponding to the experimental setup used in [6]. The channel is a 10 GBd NRZ DP-QPSK or DP-DQPSK signal with a center wavelength of 1551.32 nm or 1561.01 nm. Figure 3 illustrates the dispersion maps of both channels, where we can see a low local dispersion for 1551 nm and high local dispersion for 1561 nm. After pre-dispersion compensation, the signal is propagated through 5 loops of 18 spans of non-zero dispersion shift fiber (NZ-DSF) and 3 spans of standard single-mode fiber (SSMF) with compensating erbium-doped fiber amplifiers (EDFAs) (5 dB noise figure), post-dispersion compensation and an optical filter (4-th order Gaussian filter with a bandwidth of 2.5×10 GHz). Coherent detection is performed using a hybrid mixer and balanced photo-detectors. The electric transmit filter uses a 25 ps rise-time Gaussian pulse, and the receiver uses a 1-st order Bessel filter with a cutoff of 75 % of the symbol rate. After digitizing to 8 samples per symbol the residual dispersion is removed using a linear frequency-domain equalizer (FDE) and the

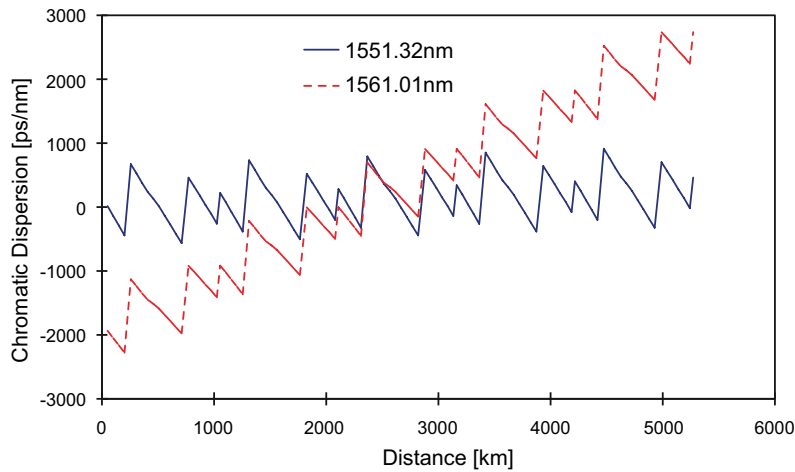


Fig. 3. Fiber channel dispersion maps (wavelength: 1551 nm and 1561 nm).

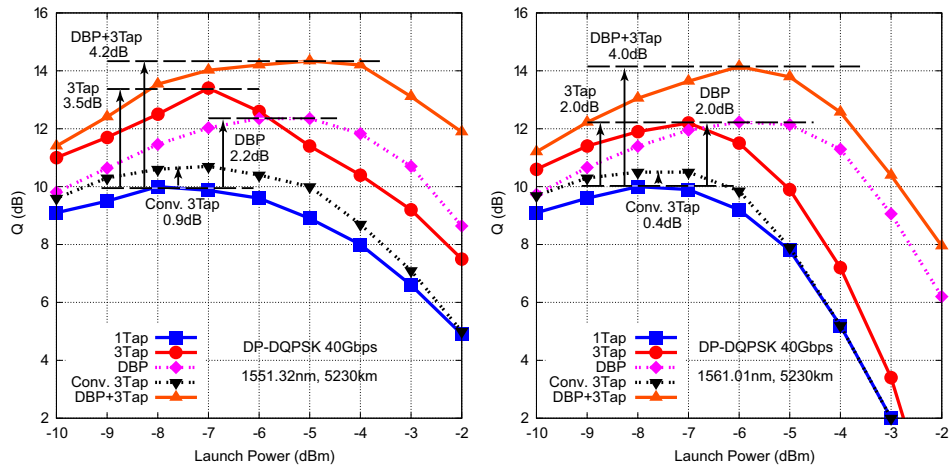


Fig. 4. Simulation results of Q-factor performance as a function of launching power for 40 Gbps DP-DQPSK (5,230 km). (a) Low local dispersion channel (1551.32 nm). (b) High local dispersion channel (1561.01 nm).

P -times oversampled signal is fed into the statistical sequence equalizer. The fiber distance per loop is 1,046 km. The Q-factor is calculated by bit error counting. We assume no PMD in simulations, and use a circular polarization basis so that two parallel equalizers for each polarization work individually. If it suffers from a strong PMD, we may need a standard polarization recovery such as constant modulus algorithms in a pre-processor block, or we can use the proposed scheme for joint polarization equalization with a complexity cost.

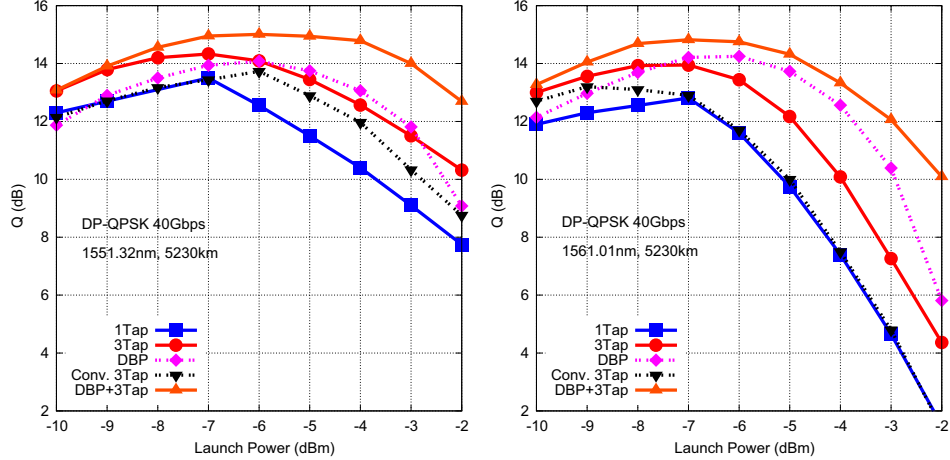


Fig. 5. Simulation results of Q-factor performance as a function of launching power for 40 Gbps DP-QPSK (5,230 km). (a) Low local dispersion channel (1551.32 nm). (b) High local dispersion channel (1561.01 nm).

3.2. Q versus launching power

Figures 4(a) and 4(b) show the simulation results of Q-factor for DP-DQPSK in low local dispersion channels (1551.32 nm wavelength) and high local dispersion channels (1561.01 nm wavelength) respectively, after 5 loops (5,230 km). Here, “1Tap” denotes a 1-tap equalizer which performs as a phase compensation filter with no memory. The notation of “3Tap” stands for the proposed fractionally-spaced statistical sequence equalizer with $M = 3$ taps, $N = 9$ excess window, and $P = 2$ oversampling. “DBP” denotes the DBP using manually optimized SSFM parameters. “Conv. 3Tap” denotes the conventional 3-tap statistical equalizer [10] without using the second-order statistics, excess window, and oversampling.

Comparing to 1-tap phase compensation method, one can see that an improvement of more than 2 dB in Q is achieved by the proposed fractionally-spaced equalizer with 3 taps. It should be noted that such a large gain is not obtained by a conventional statistical equalizer [10] with such a small number of taps, since the scheme does not exploit the second-order statistics. Moreover, the proposed equalizer can outperform the DBP (which uses 1 step per span, requiring 210 times Fourier transform operations over 5,230 km) by 1 dBQ in low local dispersion channels in Fig. 4(a). As shown in Fig. 4(b), the performance gain can be decreased in high local dispersion cases because the effective channel memory can be larger than 3-tap lengths. Nevertheless, even for such high local dispersion channels, the fractionally-spaced 3-tap equalizer achieves 2 dBQ improvement from 1-tap equalizers, and it is comparable to the DBP in a peak Q factor. More importantly, the proposed equalizer offers an additional 1–2 dBQ improvement when the DBP is used as a pre-processor. Since the DBP can shorten the effective channel memory, the proposed equalizer obtains a comparable performance gain of 4 dBQ against 1-tap equalizers for both high and low local dispersion channels.

An analogous behavior is seen in Figs. 5(a) and 5(b) where DP-QPSK signals are used. Comparing to the DP-DQPSK signals, the net gain of the proposed equalizer for DP-QPSK signals is limited to 1 dB in a peak Q factor. This is because the proposed equalizer can exploit the statistical correlation over symbols which are differentially demodulated. In fact, the performance of the 3-tap equalizer in Fig. 4(a) is almost comparable to the performance in Fig. 5(a) even

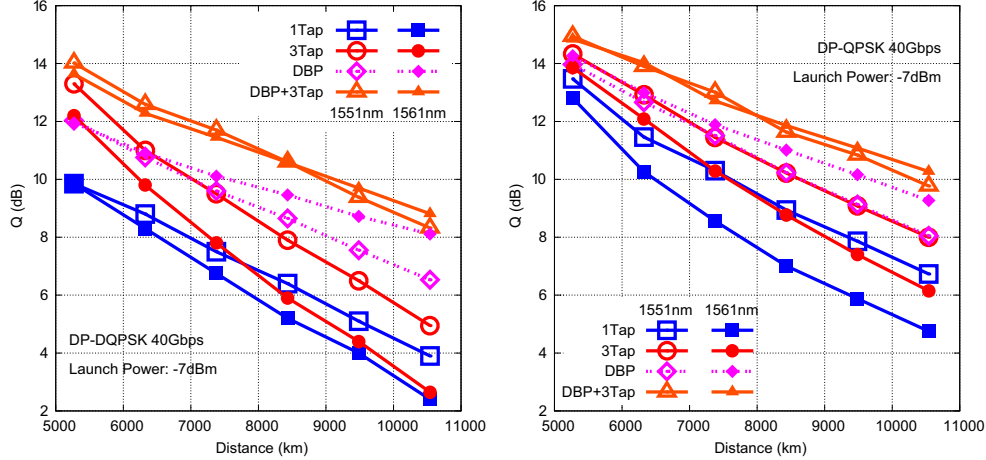


Fig. 6. Simulation results of Q-factor performance as a function of fiber distance for 40 Gbps DP-DQPSK/DP-QPSK (-7 dBm). (a) DP-DQPSK. (b) DP-QPSK.

though in Fig. 4(a) we used differential demodulations which cause approximately 3 dB loss in general. Hence, it implies that the proposed algorithm can partly mitigate the disadvantage of differential modulations.

3.3. Q versus fiber distance

Figures 6(a) and 6(b) show Q values as a function of the fiber distance, respectively for DP-DQPSK and DP-QPSK, at a launching power of -7 dBm, with increments of the fiber distance further from 5,230 km to 10,460 km by 1,046 km loop each. The Q performance degrades with the fiber distance. Compared to the 1-tap equalizer, the proposed statistical sequence equalizer with 3 taps still obtains 1 dBQ improvement even at 10,460 km for low local dispersion case, whereas the improvement is considerably reduced for high local dispersion case. It suggests that such a small tap equalizer should work with channel shortening methods for long-haul transmissions. The DBP itself works well for long-haul transmissions. And it is observed that the cascaded equalizer using both the DBP and FS-SSE achieves the best performance; more than 4 dB gain from the 1-tap equalizer at 10,460 km. The cascaded equalizer exhibits more robust performance against the local dispersion difference in long-haul transmissions; more specifically, two ‘DBP+3Tap’ curves in those figures agreed well whereas the other equalization schemes have dispersion-dependent performance.

3.4. Effect of tap length, modulation scheme, and reduced-complexity equalizer

Figure 7(a) shows the impact of the tap length where we use T -spaced equalizers not $T/2$ -spaced equalizers (T denotes the symbol time duration). It is seen that 2-tap equalizers can achieve more than 1 dBQ improvement compared to 1-tap equalizers. Whereas, 3-tap and 4-tap lengths offer moderate improvements; at most an additional 0.5 dBQ improvement respectively. It is because an effective channel memory is approximately 3-tap length.

In Fig. 7(b), we plot the Q performance of DP-16QAM transmissions at the same 10 GBd rate. The fractionally-spaced 3-tap equalizer based on 2nd-order statistics enjoys 1 dBQ improvement with and without DBP. The performance gain is less than the case of QPSK. This is because 16QAM is more sensitive to nonlinear phase noise. In this figure, we also present

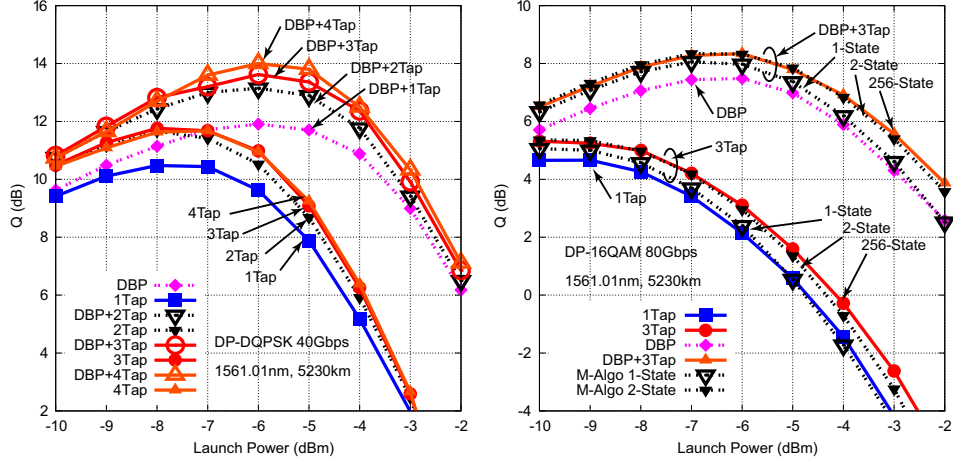


Fig. 7. Simulation results of Q-factor performance as a function of launching power for 10 Gbd DP-DQPSK/DP-16QAM (high local dispersion channel: 1561.01 nm, 5,230 km). (a) Tap length effect for DP-DQPSK. (b) M-algorithm effect for DP-16QAM.

the impact of reduced-complexity equalizers which employ the M-algorithm [12] to reduce the number of surviving states in a trellis diagram. For high-level modulation schemes, since the number of trellis-states increases significantly, such a complexity reduction scheme plays an important role for practical implementations. In fact, we can observe from Fig. 7(b) that 2-state equalizers can approach the full 256-state equalizers, and hence, approximately 78 % computational complexity can be reduced with almost no performance loss. It suggests that there are only a few dominant states to determine the performance and that an effective nonlinear channel memory is short.

3.5. 112 Gbps transmissions and beyond

We evaluate the performance for higher-speed data rates of 112 Gbps in Figs. 8(a) through 8(c), where we use 28 Gbd DP-DQPSK. Since the channel memory increases as approximately 3 times large as that of the 40 Gbps case in Fig. 4, the 3-tap FS-SSE itself has no performance gain against the 1-tap equalizer. Nevertheless, the 3-tap FS-SSE cascaded with DBP can improve the Q-factor by around 1 dB compared to the DBP alone. In this figure, we also plot the Q performance curve of the 3-tap T -spaced SSE cascaded with DBP. It is seen that FS-SSE has only a slight advantage against T -spaced SSE because there is no timing jitters in the simulation.

In Fig. 8(d), the performance for 224 Gbps DP-16QAM transmissions is presented. Although the 3-tap equalizer alone has only a marginal improvement, the cascaded FS-SSE with DBP offers an additional 0.8 dBQ improvement over the DBP alone for both high and low local dispersion channels. It is expected that a larger number of taps offers additional performance improvement at an expense of higher complexity.

In this paper, we only focused on intra-channel nonlinearity. In general, it is difficult, even with DBP, to compensate inter-channel nonlinearity because the nonlinear distortion depends on signal patterns of all channels. When pattern-independent nonlinearity dominates, it is expected that the advantage of the proposed SSE can be severely degraded. Nevertheless, it may remain some gains because the SSE still exploits pattern-independent statistics. To compen-

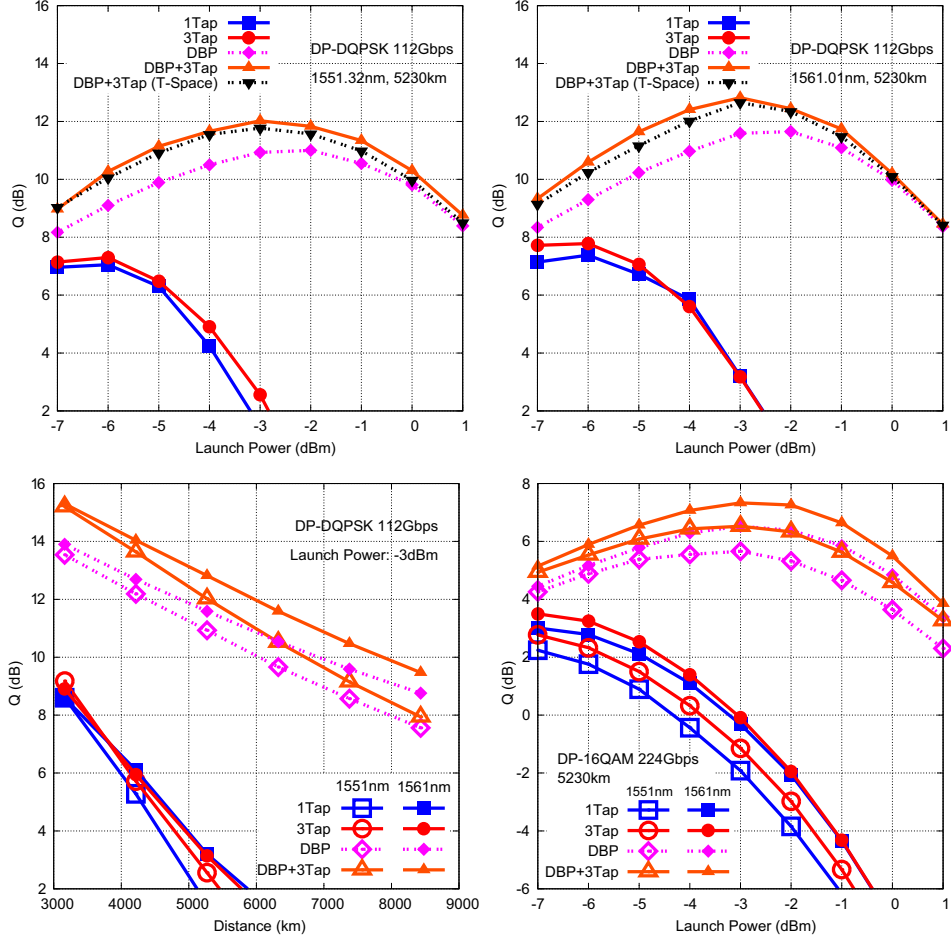


Fig. 8. Simulation results of Q-factor performance as a function of launching power and fiber distance for 28 GBd high-speed DP-DQPSK/DP-16QAM. (a) Q vs power for 1551.32 nm. (b) Q vs power for 1561.01 nm. (c) Q vs distance for -3 dBm. (d) Q vs power for DP-16QAM.

sate time-varying cross-polarization modulations (XPoIM), we may need a frequent statistics updating and the other XPoIM canceler.

3.6. Computational complexity

We here analyze the computational complexity of the proposed SSE in comparison to DBP. DBP requires 4-times Fourier transforms per step for dual polarization. Every step, phase and polarization rotations are performed in frequency domain and time domain. Let N , P , and \mathcal{M} be the size of the fast Fourier transform, the oversampling factor, and the number of steps, respectively. The computational complexity per symbol is in the order of $P\mathcal{M}(\log_2(N) + 16)$. In our simulations, we used $N = 256$, $P = 2$, and $\mathcal{M} = 105$. Hence, the computational complexity of (1 step per span) DBP becomes 5040 multiplication operations. When using reduced-step DBP of $\mathcal{M} = 5$, the computational complexity of DBP becomes 240.

Since the statistics does not require frequent updating and its complexity is low (a square-order of tap lengths), we focus on the computational complexity in MLSE detection for SSE. The proposed SSE has the computational complexity of the order $2N\mathcal{M}2^q(2N+1)$, where $\mathcal{M} \leq 2^{q(M-1)}$ is the number of survivors in M-algorithm, N is the window size, M is the tap length, and q is the number of bits per symbol. For $M = 3$ taps, $q = 2$ (QPSK), and $N = 3$ (without excess window), the computational complexity of the (full-state) SSE becomes 2688 multiplication operations. Therefore, the proposed SSE has lower complexity than 1-step per span DBP. When using M-algorithm of $\mathcal{M} = 2$, the reduced-state SSF has a complexity of 336. For 16QAM cases, the computational complexity of the full-state SSE becomes seriously large, more specifically, 172032 multiplications. However, as shown in the results, a reduced-state SSE with $\mathcal{M} = 2$ approaches the full-state SSE. With M-algorithm, the computational complexity of such a reduced-state SSE becomes 1344 even for 16QAM 3-tap cases.

4. Summary

We proposed the fractionally-spaced statistical sequence equalizer (FS-SSE) which exploits the second-order multivariate statistics of the fiber nonlinearity to mitigate nonlinear impairments that depend on a transmission data pattern. Through the computer simulations, a significant performance improvement of more than 2 dBQ was obtained for 40 Gbps coherent fiber-optic communication systems over 5,230 km. It was verified that oversampling with higher-order statistics can provide a significant gain compared to the existing statistical equalizer. More importantly, a short-memory equalizer with just 3 taps could achieve better performance than the DBP in low local dispersion conditions. We also demonstrated that the joint use of DBP and FS-SSE achieves an additional 1–2 dB gain due to the channel shortening effect. Even higher-order statistics including skewness and kurtosis can be readily introduced in the proposed method for further accurate nonlinearity modeling. The extension to the system with the inter-channel nonlinearity in wavelength-division multiplexing remains as a future work.

Acknowledgment

This research is in part supported by the National Institute of Information and Communication Technology (NICT) of Japan under “ λ Reach Project.”