Digital Signal Processing for Coherent Transceivers Employing Multilevel Formats

Md. Saifuddin Faruk, *Member, OSA* and Seb J. Savory, *Fellow, IEEE, Fellow, OSA*

(Tutorial Review)

Abstract— **Digital coherent transceivers have revolutionized optical fiber communications due to their superior performance offered compared to intensity modulation and direct detection based alternatives. As systems employing digital coherent transceivers seek to approach their information theoretic capacity, the use of multilevel modulation formats combined with appropriate forward error correction becomes essential. Given this context, in this tutorial paper, we therefore explore the digital signal processing (DSP) utilized in a coherent transceiver with a focus on multilevel modulation formats. By way of an introduction, we open by discussing the photonic technology required to realize a coherent transceiver. After discussing this interface between the analog optical channel and the digital domain, the rest of the paper is focused on DSP. We begin by discussing algorithms that correct for imperfections in the optical to digital conversion, including IQ imbalance and timing skew. Next, we discuss channel equalization including means for their realization for both quasi-static and dynamic channel impairments. Synchronization algorithms that correct for the difference between the transmitter and receiver oscillators both optical and electrical are then discussed and issues associated with symbol decoding highlighted. For most of the cases, we start with polarization division multiplexed quadrature phase-shift keying (PDM-QPSK) format as a basis and then discuss the extension to allow for high order multilevel formats. Finally, we conclude by discussing some of the open research challenges in the field.**

Index Terms—Coherent detection, digital signal processing, optical fiber communication.

I. INTRODUCTION

DIGITAL coherent transceivers have emerged as the *de facto* solution for long-haul optical fiber communication systems due to their superior performance offered compared to intensity modulation and direct detection based alternatives. While historically systems operated at the limits of the electronic technology, data converters commensurate with the optical line rate emerged over the last decade, permitting the use of digital signal processing (DSP) for high speed optical fiber communication systems. The symbiotic combination of DSP, spectrally efficient modulation formats and coherent detection led to the advent of the digital coherent transceiver which now underpins 100 GbE systems.

The first application of DSP at rates commensurate with the current optical line rates was mitigating the impact of chromatic dispersion (CD) in direct detection systems [1, 2]. While this technique was well suited to a binary modulation format, as systems sought to employ spectral efficient modulation formats to increase their capacity, the complexity became prohibitive. Consequentially, research shifted to alternative approaches such as coherent detection since the associated linear receiver did not suffer the complexity scaling issues for multilevel formats associated with a direct detection receiver. The first demonstration of demodulation of 20-Gbits/s quadrature phase-shift keying (QPSK) signal using a phase diversity coherent detection followed by digital carrier-phase estimation was reported in [3]. The key advantage of a phase diverse intradyne receiver is in avoiding the need for an optical phase lock loop (OPLL) [4], which can be challenging to realize due to the finite loop delay. While intradyne detection allowed chromatic dispersion to be compensated digitally [5], the more critical impairment for the then 40 Gbit/s systems was polarization mode dispersion (PMD) which required dynamic receiver based compensation [6]. The realization that PMD could be mitigated digital provided a phase and polarization diverse coherent receiver was used [7] stimulated further research into digital coherent [8]. Once the DSP functions of equalization and synchronization where established, advanced functionalities such as spectral shaping and nonlinearity compensation were introduced into the DSP chain. Due to the time and cost associated with an application-specific integrated circuit (ASIC) for real time operation, much of the early research was based on offline processing. This facilitated ASIC designers to determine the feasibility of realizing key algorithms in a highly parallel CMOS ASIC with limited analog-to-digital (ADC) resolution. Having established the design space, this then allowed the first ASIC which incorporated DSP and ADC to be developed by Nortel which was commercially deployed in 2008 [9,10].

Most of the DSP algorithms for polarization division multiplexed QPSK (PDM-QPSK) are derived from the wireless communication counterparts. However, necessary modifications are required to cope with special nature of fiber optic channel. The DSP design become more challenging as

The authors are with the Electrical Engineering Division, Engineering Department, University of Cambridge, 9 JJ Thomson Avenue, Cambridge CB3 0FA, U.K. (e-mail: msf35@cam.ac.uk, sjs1001@cam.ac.uk).

the modulation format evolves from PDM-QPSK to multilevel quadrature amplitude modulation (QAM) formats. With an extension of [11], in this paper, we review the principle of fundamental DSP techniques used in coherent transceivers. We also explore the recent advancements of DSP algorithms especially for multilevel formats.

Rest of the paper is organized as follows: Sec. II gives an overview of DSP-based coherent transceivers. Compensation of different IQ imbalances arising from transmitter and receiver imperfections are discussed in Sec III. Section IV and V are concerned with equalization of different linear and nonlinear transmission impairments where both time and frequency domain approaches are considered. Synchronization algorithms are discussed in Sec. VI-VIII, after which symbol estimation and decoding are explained in Sec. IX. Finally, current research challenges are highlighted in Sec. X and the paper is concluded in Sec. XI.

II. DSP-Based Coherent Transceivers

The usual configuration of coherent transceiver for PDM transmission system utilizing advanced vector modulation formats and DSP techniques is shown in Fig.1. At the transmitter, first two independent data sequences for *x*- and *y*-polarizations are processed in the DSP which may include encoding, modulation, precompensation of linear and nonlinear transmission impairments and pulse shaping filtering. Such digitally processed signals are converted to analog signals using four digital-to-analog converters (DACs)[1] corresponding to the in-phase (*I*) and quadrature (*Q*) components of signals for *x*- and *y*- polarizations which are then used to drive two IQ modulators (IQMs). A single laser output split by polarization beam splitter (PBS) is used for two IQMs. Outputs of IQMs are then combined by polarization beam combiner (PBC) and transmitted through fiber.

At the receiver, the phase and polarization diversity architecture shown in Fig. 1(b) is often employed to linearly map incoming optical signal into four electrical signals, corresponding to the in-phase and quadrature field components for the two polarizations. The optical front-end is employed with two PBSs and a pair of 90° hybrids, one for each component of polarization [12]. The outputs of the optical hybrid are then detected with four balanced photodiodes whose outputs are given as [13]:

$$\begin{bmatrix} I_x \\ Q_x \\ I_y \\ Q_y \end{bmatrix} \propto \begin{bmatrix} \mathrm{Re}\{E_x E_{lo}^*\} \\ \mathrm{Im}\{E_x E_{lo}^*\} \\ \mathrm{Re}\{E_y E_{lo}^*\} \\ \mathrm{Im}\{E_y E_{lo}^*\} \end{bmatrix} \quad (1)$$

where Re $\{\cdot\}$ and Im $\{\cdot\}$ represent real and imaginary components, E_x and E_y denote electric fields of the dual polarization incoming optical signal with *x*- and *y*-polarization components, respectively, and E_{lo} is that from the local

---

[1] For QPSK format, *I* and *Q* data can directly be used to drive IQM without DACs if no pulse shaping or precompensation is used. Such a technique was used in the first generation of coherent transceivers [9] as well as in early coherent transmission experiments.

oscillator (LO). Afterwards, signals are amplified with trans-impedance amplifiers (TIAs) before digitizing by the ADCs.
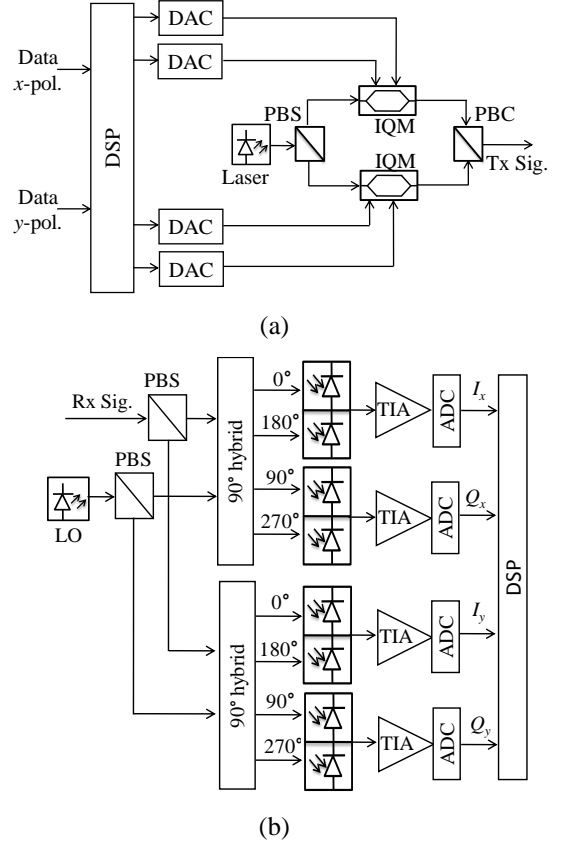

(a)

(b)

Fig. 1. Configuration of DSP-based (a): transmitter and (b): receiver. (DAC: digital-to-analog converter, IQM: IQ modulator, PBS: polarization beam splitter, PBC: polarization beam combiner, LO: local oscillator, TIA: trans-impedance amplifier, ADC: analog-to-digital converter, DSP: digital signal processing).
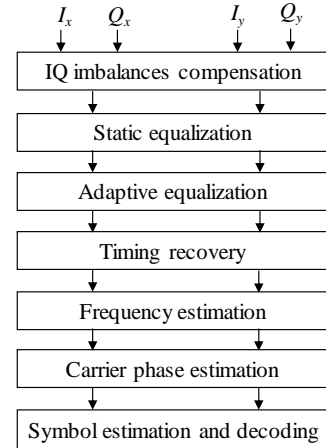

Fig. 2. Typical sequence of DSP operation in a coherent receiver to demodulate the data.

The received signal is contaminated by channel impairments and noise. The channel impairments may stem from linear effects such as CD, PMD etc. or fiber nonlinear effects such as self-phase modulation (SPM), cross phase modulation (XPM) etc. On the other hand, the sources of noise

are shot noise, thermal noise, amplified spontaneous emission (ASE) noise, residual beating of the LO with relative intensity noise (LO-RIN), quantization noise from ADC etc. [14, 15]. The dominant noise source for short reached unamplified systems is shot noise while that for long haul systems with inline amplifiers is ASE noise.

The digitized signal is passed through a chain of DSP operations to demodulate the data. While for a particular realization of a digital coherent receiver, the DSP operations may vary, the generic DSP procedures are shown in Fig. 2 and their functionalities can be summarized as below:

IQ Imbalances compensation: The mismatch of amplitude, phase and timing between *I* and *Q* component are compensated.

Digital equalization: The linear and nonlinear impairments of the channel is equalized.

Timing recovery: The timing errors are estimated and corrected from them.

Frequency and phase recovery: The frequency mismatch between transmitting laser and LO and the laser phase noise are estimated and compensated.

Symbols estimation and decoding: Error correction codes are employed and transmitted symbols are estimated.

III. IQ IMBALANCE COMPENSATION

The phase/gain mismatch between *I*- and *Q*- port of the received signal may arise at transmitter, for example from the improper biasing of IQ modulator or at receiver such as the imperfection in any of components of optical 90° hybrids, balanced photodiodes, or TIAs. In addition, timing mismatch between the I and Q ports may also induced by the difference in the physical path length of the circuit trace, which is known as the IQ delay skew.

There are different ways of modeling the IQ imbalances effect. One formalism is to relate a complex signal having imbalances $\{I'(t) + jQ'(t)\}$ with its ideal signal $\{I(t) + jQ(t)\}$ in the form of [16]

$$I'(t) + jQ'(t) = \left[g_I I\left(t - \frac{\tau_{IQ}}{2} \right) \right] + \left[g_Q Q\left(t - \frac{\tau_{IQ}}{2} \right) \right] e^{j\left(\frac{\pi}{2} + \phi_{IQ} \right)} \quad (2)$$

where, g_I and g_Q are the I/Q gains, τ_{IQ} is the timing skew and ϕ_{IQ} is the quadrature phase error. The IQ gain mismatch in dB is then defined by

$$\alpha_{IQ}(dB) = 20\log_{10}\left(\frac{g_Q}{g_I} \right). \quad (3)$$

TABLE-I
THE IQ IMBALANCES TOLERANCE FOR 1dB SNR PENALTY AT BER OF 10^{-2} WITHOUT ANY COMPENSATOR IN RECEIVER

Modulation format	α_{IQ} [dB]	ϕ_{IQ}[deg.]	τ_{IQ} [% of *symbol-period*]
4-QAM	3.8	28	26
16-QAM	1.6	10.5	11
64-QAM	0.6	4	4.2
256-QAM	0.2	1.4	1.5

The impact of receiver side IQ imbalance for different modulation formats are summarized in Table-I as a function of 1-dB sensitivity penalty at a bit-error rate (BER) of 10^{-2}. As can be seen in the results of Table-I, as the order of the modulation increases the requirements for the IQ imbalance become increasing more stringent. Thus, system performance can degrade significantly if for higher level modulation formats IQ imbalance is not compensated by the receiver DSP.

The well-known techniques for orthogonalization of two non-orthogonal vectors are the Gram-Schmidt orthogonalization procedure (GSOP) [17] and the Löwdin orthogonalization [18, 19]. The Gram-Schmidt orthogonalization is based on defining a new vector that is orthogonal to the initially selected vector. This makes two vectors orthogonal as shown in Fig. 3(a). This procedure can be shown mathematically as:

$$\begin{bmatrix} I(t) \\ Q(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix} \begin{bmatrix} I'(t) \\ Q'(t) \end{bmatrix} \quad (4)$$

where *a* is the inner product given as $a = \langle I' | Q' \rangle$.

On the other hand, Löwdin orthogonalization is a symmetric orthogonalization where both the vectors are rotated by the same angle to make them orthogonal as shown in Fig. 3(b). The transformation can be written as:

$$\begin{bmatrix} I(t) \\ Q(t) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \dfrac{1}{\sqrt{1+a}} + \dfrac{1}{\sqrt{1-a}} & \dfrac{1}{\sqrt{1+a}} - \dfrac{1}{\sqrt{1-a}} \\ \dfrac{1}{\sqrt{1+a}} - \dfrac{1}{\sqrt{1-a}} & \dfrac{1}{\sqrt{1+a}} + \dfrac{1}{\sqrt{1-a}} \end{bmatrix} \begin{bmatrix} I'(t) \\ Q'(t) \end{bmatrix}. \quad (5)$$

While the GSOP is widely used for receiver-side IQ phase/gain imbalance compensation in the literature, Löwdin orthogonalization offers a symmetric alternative.
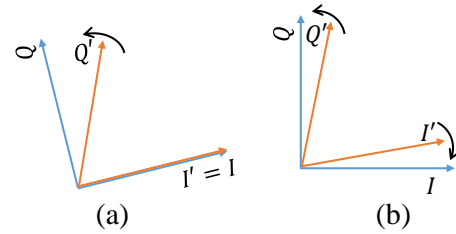

Fig. 3. Visual representation of the (a): Gram-Schmidt and (b): Löwdin orthogonalization algorithms.

The IQ delay skew can be compensated by interpolation which serves as a time shifting. The interpolator can be implemented using a finite-impulse response (FIR) filter. When the time window of interpolation is small enough, several polynomials of a lower degree can be used. For example, using interpolating function as a quadratic term, the skew compensator can be realized using a 3-tap FIR filter [20]. Alternatively, the skew can be estimated in frequency domain based on the principle of Godard phase detection and then estimated value can be used for compensation using time domain interpolator [21].

Recently it has been shown that with a modified configuration in the adaptive equalizer (generally used for polarization tracking) all the receiver-side IQ imbalances can

be compensated effectively [22, 23] which we discuss further in Sec. V.

While the DSP at receivers can fully compensate receiver-side imbalances for a sufficiently high ADC bit resolution, the transmitter-side imbalances only can be mitigated due to the noise which occurs between the transmitter and the receiver [16]. As an example, the performance of GSOP to compensate transmitter- and receiver-side phase imbalance compensation is shown in Fig. 4 considering 6-bits ADC resolution. The SNR penalty at BER 10^{-2} is measure when the 16-QAM signal is transmitted over additive white Gaussian noise (AWGN) channel. Clearly, the receiver-side imbalance is effectively compensated over a wide range while we must consider penalty for transmitter-side imbalance, albeit a considerable performance improvement can be achieved.
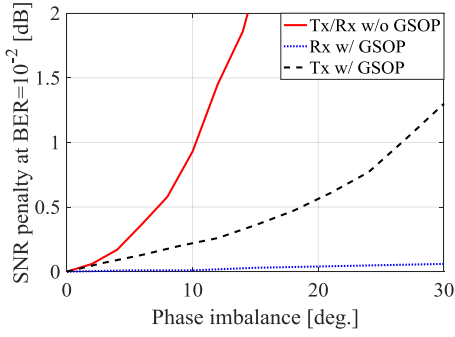

Fig. 4. Sensitivity penalty when GSOP is used to compensate for transmitter-/receiver-side IQ phase imbalance.

IV. STATIC EQUALIZATION

While in principle, equalization of all linear impairments could be realized with one digital filter, it is generally beneficial to partition the problem into static and dynamic equalizer. The rational for this partitioning is that static equalization typically requires large static filters, while dynamic equalization requires a set of relatively short adaptive filters to compensate for time-varying phenomena. The tracking speed of the adaptive filters depends on a number of factors including the number of filter taps, adaptation algorithm, the value of step size parameter, degree of parallelization etc. For example, tracking of maximum polarization rate of change of 50 kHz was demonstrated in [24] by five taps $T/2$-spaced filters adapted by least-mean-square (LMS) algorithm with an update in every 12 symbols. The fastest change of polarization states in an optical channel is generally in order of ms [25, 26] and thus can be tracked by the adaptive filters.

A. Chromatic Dispersion Compensation

The transfer function of the CD is given by [27]:

$$G(z,\omega) = \exp\left(-\frac{j\omega^2 \beta_2 z}{2}\right) \qquad (6)$$

where ω is the angular frequency, β_2 is the group delay dispersion parameter of the fiber and z is the transmission length. The dispersion compensating filter is therefore given

by the all-pass filter with transfer function $H_c(\omega) = 1/G(z,\omega) = G(-z,\omega)$, which can be approximated either in time-domain by using a FIR filter or in the frequency domain method.

1) Time Domain Equalizer:

The time domain equalizer (TDE) is preferably implemented by a FIR-filter whose direct realization can be viewed as in Fig. 5.
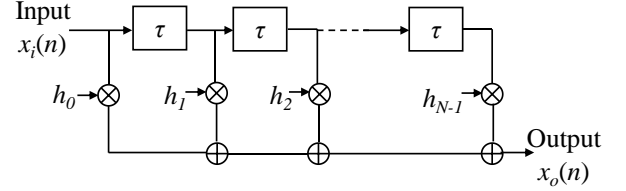

Fig. 5. Configuration of the FIR filters.

The input signal $x(n)$ is delayed by integer multiples of τ, where τ is a sample delay defined by the sampling rate. We define a vector using these delayed elements of the input as

$$\mathbf{x}_i(n) = [x_i(n), x_i(n-1), \cdots x_i(n-N+1)]^T \qquad (7)$$

where the superscript $(\cdot)^T$ denotes the transpose operation and N is the number of tap weights of the FIR filter. The tap weights also constitute a vector as

$$\mathbf{h}(n) = [h_0, h_1, \cdots h_{N-1}]^T. \qquad (8)$$

Then, the output of the FIR filter is given by

$$x_o(n) = \mathbf{h}^T \mathbf{x}_i(n) = \sum_{k=0}^{N-1} h_k x_i(n-k). \qquad (9)$$

To calculate the tap-weights vector for CD compensation, herein we consider two methods, namely, truncated impulse response and least-square method. The impulse function of the chromatic dispersion compensating filter $h_c(z,t)$ can be obtained by inverse Fourier transform of $G(-z,\omega)$ given as

$$h_c(z,t) = \frac{1}{\sqrt{-2\pi j\beta_2 z}} \exp\left(-j\frac{t^2}{2\beta_2 z}\right). \qquad (10)$$

The impulse response given by Eq. (10) presents a number of issues for digital implementation, not only it is infinite in duration but since it passes all frequencies for a finite sampling frequency, aliasing will occur. These problems can be solved by truncating the impulse response to a finite duration which can be express for sampling time T_s as [7, 28]:

$$h_{c_{TI}}(n) = \frac{1}{\sqrt{\rho}} \exp\left(-j\frac{\pi}{\rho}\left(n - \frac{N_c-1}{2}\right)^2\right) \qquad (11)$$

where, $\rho = 2\frac{\pi\beta_2 z}{T_s^2}$, $N_c = \lfloor \rho \rfloor$ and $n \in [0, 1, 2 \cdots N_c - 1]$.

where $\lfloor x \rfloor$ is the integer part of x rounded towards minus infinity.

The least squares criterion may be applied to the design of a

complex FIR filter, giving optimal tap weights in a least squares sense as in Eq. (12) [29], where $\mathrm{erfi}(x)$ is the imaginary error function given by $\mathrm{erfi}(x) = -j\,\mathrm{erf}(jx)$ given that $\mathrm{erf}(z) = \dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^z e^{-t^2}\,dt$.

For shorter link or higher order modulation format, the truncated impulse response method has significant performance limitation. Substantial performance improvement can be achieved by using the least square method. Nevertheless, the performance degradation can be mitigated by choosing number of filter taps beyond $\lfloor \rho \rfloor$.

The CD filter can alternatively be implemented using digital infinite-impulse response (IIR) filter with significantly reduced number of taps [30]. However, it requires large number of buffers for time reversal operations. More importantly, unlike FIR filter, it is a big challenge to realize the parallel structure of IIR filter which is one of the key requirement for real time implementation [31].

2) Frequency Domain Equalizer:
For long-haul dispersion uncompensated link, we need a large number of taps to compensate CD. As the number of taps, N_c, is increased, the complexity of filtering increases rapidly for time-domain implementations ($\mathcal{O}(N_c)$). In such a case, the filtering can be done more efficiently in frequency domain ($\mathcal{O}(\log N_c)$) [32]. To implement frequency domain equalizer (FDE) either overlap-save method [33, 34] or overlap-add method [35] has been investigated for CD compensation with former one having less complexity.
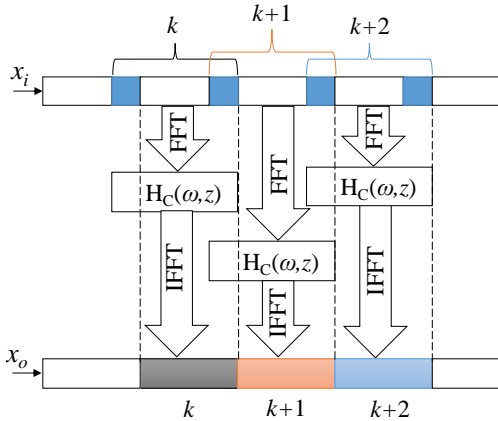

Fig. 6. Illustration of overlap-save implementation of CD-equalizer.

The CD compensating FDE based on overlap-save method is illustrated in Fig. 6. The incoming signal is first transformed into frequency domain with fast Fourier transform (FFT) with a block length $N_{\mathrm{FFT}} = N_o + N_c$, where N_o samples are part of last data block and N_c samples are from current data blocks. The frequency domain signal is then multiplied with $H_c(\omega)$ and then transformed into time domain using inverse FFT (IFFT). The first N_o samples are discarded for avoiding cyclic

properties of FFT and N_c equalized samples are retained. With an optimized FFT algorithms, using such implementing FDE can be beneficial over TDE even for $N_c > 4$ [33].

3) Subband Equalizer:
The total number of taps required for CD compensation can be greatly reduced with a subband equalizer [36, 37]. Fig. 7 shows the configuration of an oversampled subband equalizer with N_{sb} subbands, downsampled by a factor M ($M < N_{sb}$). The analysis filter bank is a uniformly modulated filter bank based on a prototype filter of $H(z)$. Then each subband is equalized with a CD equalizer $H_{c_k}(z)$ together with a delay δ_k ($k = 0, 1, \cdots, N_{sb}$-1). The synthesis filter bank is based on a prototype filter of $G(z)$ with an upsampling factor of K whose outputs are combined to produce the equalized output. Assuming the original signal of is oversampling by a factor of M/K, the total number of taps of all subbands is approximately $1/K$ of that for the full-band equalizer.
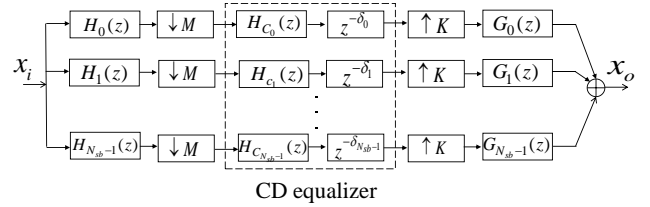

Fig. 7. Configuration of sub-band implementation of CD equalizer.

B. Matched Filtering

To enable future systems to minimize their spectral utilization for a given higher-order modulation format, Nyquist pulse shaping is required [38]. In the limit, Nyquist pulse shaping can generate near-rectangular spectra, lowing the wavelength division multiplexing (WDM) channel spacing to be reduced close to the symbol rate of each channel. Often raised-cosine (RC) spectral shaping is used to satisfy the Nyquist criterion and thus inter-symbol interference (ISI) free operation [39]. The RC response is given by:

$$H_{RC}(\omega) =$$
$$\begin{cases} T_s, & 0 \le |\omega| < \pi(1-\beta)/T_s \\ \dfrac{T_s}{2}\left(1 - \sin\left[\dfrac{T_s}{2\beta}\left(|\omega| - \dfrac{\pi}{T_s}\right)\right]\right), & \dfrac{\pi(1-\beta)}{T_s} \le |\omega| < \dfrac{\pi(1+\beta)}{T_s} \\ 0, & |\omega| \ge \pi(1+\beta)/T_s \end{cases} \quad (13)$$

where, β is the roll-off factor. It is well known that to maximize the signal-to-noise ratio, RC filter can be split to both transmitter and receiver side [40]. Therefore, a root-raised-cosine (RRC) spectral shaping at the transmitter side and a second RRC filter as a matched filter (MF) at the receiver side is used while RRC filter response is given as $\sqrt{H_{RC}(\omega)}$.

$$h_{c_{LS}}(n) = \frac{1}{2j}\left[\mathrm{erfi}\left(\sqrt{\frac{j\pi}{\rho}}\left(n - \frac{N_c-1}{2} + \frac{\rho}{2}\right)\right) - \mathrm{erfi}\left(\sqrt{\frac{j\pi}{\rho}}\left(n - \frac{N_c-1}{2} - \frac{\rho}{2}\right)\right)\right] \times \frac{1}{\sqrt{\rho}}\exp\left(-\frac{j\pi}{\rho}\left(n - \frac{N_c-1}{2}\right)^2\right). \quad (12)$$

A RRC filter with β close to zero may require a few hundreds of taps to implement; thus, MF and CD equalizer can be combined to reduce the computational burden [41]. The static filter implemented at receiver will then have a response:

$$H_{static}(\omega) = H_{RRC}(\omega)H_c(\omega, z). \quad (14)$$

It is important to note that, implementing the MF in the CD filter is sensitive to frequency offset and system performance can severely be degraded since at this point of DSP chain frequency offset compensation is difficult [41]. This problem however can be solved by allowing MF in the blind adaptive filtering with larger number of taps which can be efficiently implemented in frequency domain [42].

At the transmitter, the RRC filter can be combined with CD precompensation without any additional computational cost, allowing the CD compensation to be realized in the transmitter, receiver or a combination of the two.

C. Fiber Nonlinearity Compensation

The Kerr effect in optical fibers induces nonlinear waveform distortion and it is the major impairment of the coherent optical system that limits the maximum achievable transmission distance [43]. For long distances, the nonlinear transmission can be modeled by Manakov equation [44], which is of the form as:

$$\frac{\partial E_{x/y}}{\partial z} = -\frac{\alpha}{2}E_{x/y} - \frac{j}{2}\beta_2 \frac{\partial^2 E_{x/y}}{\partial t^2} + \frac{8j}{9}\gamma\left(|E_x|^2 + |E_x|^2\right)E_{x/y} \quad (15)$$

where α is the loss coefficient and γ represents the nonlinear coefficient. To mitigate the Kerr effect in the presence of loss and dispersion, one of the most investigated techniques is digital backpropagation (DBP) where the received signal is transmitted through virtual fibers and amplifiers to emulate the inverse characteristics of the actual transmission link [45, 46]. By reversing the signs of α, β_2, and γ of Eq. (15), the mathematical expression for this virtual propagation can be found as:

$$\frac{\partial E_{x/y}}{\partial z} = \frac{\alpha}{2}E_{x/y} + \frac{j}{2}\beta_2 \frac{\partial^2 E_{x/y}}{\partial t^2} - \frac{8j}{9}\gamma\left(|E_x|^2 + |E_x|^2\right)E_{x/y}. \quad (16)$$

The nonlinear partial differential equation given by Eq. (16) is solved numerically, generally using split step Fourier method (SSFM) where the fiber is divided into small sections. Fiber dispersion and nonlinearity is calculated independently; in one step only dispersion is considered ignoring nonlinearity and in the next step nonlinear term is calculated excluding dispersion. Alternatives to receiver based DBP is digital precompensation (DPC) [47] at the transmitter or a symmetric split of the nonlinearity compensation between transmitter and receiver [48].

The receiver based DPB, requires multiple computational steps per span and as such suffers from high computational complexity. To reduce the complexity, the perturbation-based nonlinear pre-compensation technique can be used with only one computation step with a symbol-rate sampling sequence [49]. The complexity can be further reduced by incorporating symmetric electronic dispersion compensation and RRC pulse shaping in perturbation based nonlinearity pre-compensation [50].

While single-channel DBP is effective to mitigate SPM, to combat with XPM for a WDM system, multi-channel DBP (MC-DBP) is required, realized by either high bandwidth receivers [51] or spectrally sliced receivers [52]. Recently, it has been demonstrated that for a fully loaded WDM system, the peak SNR gain[2] for a single-channel DBP is 0.5 dB while for joint 3- and 5-channel DBP the improvement is upper bounded to about 1 dB and 1.2 dB, respectively [53]. Each additional channel in the joint DBP process increases the peak-SNR by no more than 0.1 dB.

Other than the backpropagation based method, several other nonlinearity mitigation schemes have been proposed and demonstrated so far such as nonlinear Volterra equalizer [54], inverse scattering transform [55], machine learning method [56], maximum *a priori* (MAP) detection [57] etc.

V. ADAPTIVE EQUALIZATION

The adaptive equalizer is used to compensate for the dynamic impairments of channels. Generally, a two-by-two multiple-input multiple-output (MIMO) structured FIR filters as shown in Fig. 8 is used to estimate the inverse Jones matrix of the dynamic channel [8].

For simplicity of explanation, we consider that input sequence of the filter is symbol-spaced sampled with a sampling index n. Extension for two-fold oversampled input sequence is straightforward-tap weights are just updated for every two samples. Consider that for an N tap FIR filter, $\mathbf{h}_{xx}$, $\mathbf{h}_{xy}$, $\mathbf{h}_{yx}$, and $\mathbf{h}_{yy}$ are the column vector of length N representing the tap weights, with $\mathbf{x}_i$ and $\mathbf{y}_i$ representing a sliding block of N samples to which the filter is applied such that

$$\mathbf{x}_i(n) = [x_i(n), x_i(n-1) \cdots x_i(n-N)]$$

$$\mathbf{y}_i(n) = [y_i(n), y_i(n-1) \cdots y_i(n-N)]. \quad (17)$$
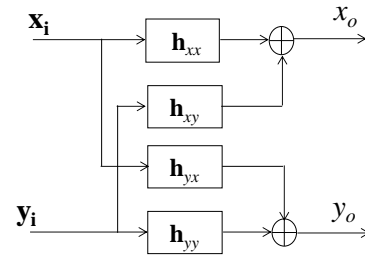

Fig. 8. Configuration of 2×2 MIMO structured adaptive FIR filters.

For simplification of mathematical expressions, consider that $\mathbf{u}_i(n) = [\mathbf{x}_i(n); \mathbf{y}_i(n)]$, $\mathbf{h}_x(n) = [\mathbf{h}_{xx}(n); \mathbf{h}_{xy}(n)]$ and $\mathbf{h}_y(n) = [\mathbf{h}_{yx}(n); \mathbf{h}_{yy}(n)]$ which yields the filter outputs in the form of

$$x_o(n) = \mathbf{h}_x^H(n)\mathbf{u}_i(n)$$

$$y_o(n) = \mathbf{h}_y^H(n)\mathbf{u}_i(n) \quad (18)$$

---

[2] Peak SNR gain is defined as the difference in peak SNR of SNR versus average signal power curve with and without DBP.

where superscript $(\cdot)^H$ denotes the conjugate transpose.

A. Adaptation Algorithms

In an adaptive equalizer, we frequently have a cost function whose gradient is stochastically estimated and used to update the tap weights. For details of underlying mathematics for such estimation, readers are refer to [28]. While enormous algorithms for filter tap adaption is available in signal processing literatures, herein we discuss the commonly used algorithms used in coherent optical systems.

1) Constant-Modulus Algorithm and Its Variants:

The constant modulus-algorithm (CMA) [58] attempts to minimize, in a mean square sense, the error ε_x and ε_y given by

$$\varepsilon_x(n) = R_2 - |x_o(n)|^2$$

$$\varepsilon_y(n) = R_2 - |y_o(n)|^2 \qquad (19)$$

where, R_2 is the real-valued constant depends on the ideal symbols $x_{sym.}$ and given by $R_2 = E|x_{sym.}|^4 / E|x_{sym.}|^2$. Correspondingly, the update equations for the filters are given as follows:

$$\mathbf{h}_x(n+1) = \mathbf{h}_x(n) + \mu\varepsilon_x(n)x_o^*(n)\mathbf{u}_i(n)$$

$$\mathbf{h}_y(n+1) = \mathbf{h}_y(n) + \mu\varepsilon_y(n)y_o^*(n)\mathbf{u}_i(n) \qquad (20)$$

where μ is the step size parameter and the superscript $(\cdot)^*$ denotes the complex conjugate operation. For initialization, all tap weights are set to zero with the exception of the central tap of $\mathbf{h}_{xx}$ and $\mathbf{h}_{yy}$, which are set to unity.

The CMA can be implemented reliably in a full-blind mode; however, since the equalizer is unconstrained with respect to its outputs, it is possible for the equalizer to converge on the same output, corresponding to the Jones matrix becoming singular. Nevertheless, there are many well-established methods to overcome such singularity problem [59–62].

The CMA equalizer tap weight updates can be simplified by discarding the information about the magnitude of the gradient term, keeping only the sign of the gradient [63, 64]. By taking the sign of both the error term and the output data symbols, such that only the sign of the gradient is used to update the filter taps, the update equations in Eq. (20) can be rewritten as:

$$\mathbf{h}_x(n+1) = \mathbf{h}_x(n) + \mu\,\mathrm{sgn}\{\varepsilon_x(n)\}\mathrm{csgn}\{x_o^*(n)\}\mathbf{u}_i(n)$$

$$\mathbf{h}_y(n+1) = \mathbf{h}_y(n) + \mu\,\mathrm{sgn}\{\varepsilon_y(n)\}\mathrm{csgn}\{y_o^*(n)\}\mathbf{u}_i(n) \qquad (21)$$

where we define the signum function of the real number x and complex number z as:

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\mathrm{csgn}(z) = \mathrm{sgn}\{\mathrm{Re}(z)\} + j\,\mathrm{sgn}\{\mathrm{Im}(z)\}.$$

Such adaptation is often known as sign-sign CMA. As shown

in Eq. (21), when coupled with a power-of-two step-size, a multiply-free tap update can be realized. This reduces the total number of multiplications for adaptive filtering by a factor of two.

While the CMA is well suited to constant modulus formats such at *M*-ary PSK, many of the formats considered for future optical networks are not constant modulus (e.g. PDM-16QAM) and as such the CMA error term can never converge to zero [65]. Nevertheless, the CMA can be adapted to a radially directed equalizer (RDE). In this case, the error signals are given as:

$$\varepsilon_x(n) = Q_r\left(|x_o(n)|^2\right) - |x_o(n)|^2$$

$$\varepsilon_y(n) = Q_r\left(|y_o(n)|^2\right) - |y_o(n)|^2 \qquad (22)$$

where Q_r (r^2) is a function that quantizes the radius according to the number of possible rings.

Note that for higher-order QAM, the relative probability of the radii, based on the number of constellation points in each radius can be included in the error functions of RDE [66] such that

$$\varepsilon_x(n) = P_r\left\{Q_r\left(|x_o(n)|^2\right) - |x_o(n)|^2\right\}$$

$$\varepsilon_y(n) = P_r\left\{Q_r\left(|y_o(n)|^2\right) - |y_o(n)|^2\right\} \qquad (23)$$

where P_r is the relative probability of receiving a radius *r*. The probabilistic RDE significantly improves performance in tracking fast changes in dynamic channels over conventional RDE for a higher QAM format beyond 16-QAM; however, the performance gain is negligible for a static channel.

2) Decision-Directed Least-Mean-Square Algorithm:

The error signal minimized for decision-directed LMS (DD-LMS) algorithm is given by [67]:

$$\varepsilon_x(n) = D\{x_o(n)\} - x_o(n)$$

$$\varepsilon_y(n) = D\{y_o(n)\} - y_o(n) \qquad (24)$$

where, $D\{x\}$ is the symbol closest to *x*. Correspondingly the update equations are given as:

$$\mathbf{h}_x(n+1) = \mathbf{h}_x(n) + \mu\varepsilon_x^*(n)\mathbf{u}_i(n)$$

$$\mathbf{h}_y(n+1) = \mathbf{h}_y(n) + \mu\varepsilon_y^*(n)\mathbf{u}_i(n). \qquad (25)$$

It is important to note that, blind startup of DD-LMS is not reliable since most of the decision at the startup may be wrong. Therefore, for reliable convergence, either training symbols based LMS algorithm or the CMA is used at the startup and then switch to blind DD mode.

Although CMA and DD-LMS algorithms require almost identical mathematical processing, one of the key difference between them is feedback latency. For a correct symbol decision to calculate the error signals for DD-LMS, the filtered outputs should pass through intermediate frequency offset compensation, carrier recovery and decoder blocks.

Thus, the feedback latency for DD-LMS includes all the latency from these blocks unlike the CMA where the error signals can be computed directly from filter outputs as shown in Fig. 9. Therefore, the LMS equalizer update has a much higher loop latency than the CMA update which is a key concern in hardware implementation [31].
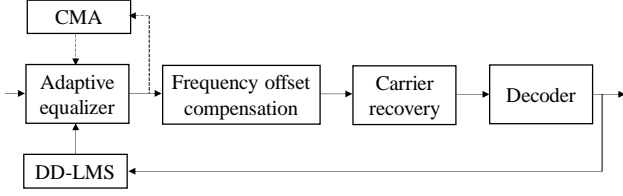

Fig. 9. DSP blocks for CMA and DD-LMS updates.

3) Recursive Least-Squares Algorithm

It is well established that the recursive least-squares (RLS) algorithm provides much faster convergence than the LMS algorithm with an expense of higher computational cost [68].

TABLE-II
SUMMARY OF STANDAR RLS ALGORITHM

Initialization	$\mathbf{h}_{x/y}(0)$-all zeros except a center spike at $\mathbf{h}_{xx}$ and $\mathbf{h}_{yy}$ $\mathbf{P}_{x/y}(0) = \delta\mathbf{I}_{2N}$; δ is small positive number, $\mathbf{I}$ identity matrix
Gain vector computation	$\mathbf{k}_x(n) = \dfrac{\lambda^{-1}\mathbf{P}_x(n-1)\mathbf{u}_i(n)}{1+\lambda^{-1}\mathbf{u}_i^H(n)\mathbf{P}_x(n-1)\mathbf{u}_i(n)}$ $\mathbf{k}_y(n) = \dfrac{\lambda^{-1}\mathbf{P}_y(n-1)\mathbf{u}_i(n)}{1+\lambda^{-1}\mathbf{u}_i^H(n)\mathbf{P}_y(n-1)\mathbf{u}_i(n)}$
Output calculation	$x_{0_{n-1}}(n) = \mathbf{h}_x^H(n-1)\mathbf{u}_i(n)$ $y_{0_{n-1}}(n) = \mathbf{h}_y^H(n-1)\mathbf{u}_i(n)$
Error estimation	$\varepsilon_x(n) = D\{x_{0_{n-1}}(n)\} - x_{0_{n-1}}(n)$ $\varepsilon_y(n) = D\{y_{0_{n-1}}(n)\} - y_{0_{n-1}}(n)$
Tap-weights adaptation	$\mathbf{h}_x(n) = \mathbf{h}_x(n-1) + \mathbf{k}_x(n)\varepsilon_x^*(n)$ $\mathbf{h}_y(n) = \mathbf{h}_y(n-1) + \mathbf{k}_y(n)\varepsilon_y^*(n)$
Correlation matrix update	$\mathbf{P}_x(n) = \lambda^{-1}\mathbf{P}_x(n-1) - \lambda^{-1}\mathbf{k}_x(n)\mathbf{u}_i^H(n)\mathbf{P}_x(n-1)$ $\mathbf{P}_y(n) = \lambda^{-1}\mathbf{P}_y(n-1) - \lambda^{-1}\mathbf{k}_y(n)\mathbf{u}_i^H(n)\mathbf{P}_y(n-1)$

In a standard RLS algorithm, we minimize the cost function

$$J_{x,y}(n) = \sum_{i=1}^n \lambda^{n-i}\left|e_{x/y}(i)\right|^2 \qquad (26)$$

where, λ is the forgetting factor ($\lambda<1$) and $e_{x,y}$ is same as that of LMS algorithm. The important feature of RLS algorithm is that it uses the information of input signal. Inverse auto correlation matrix $\mathbf{P}_{x/y}$ is used for such purpose which is recursively estimated using Woodbury's identity instead of direct calculation from input signal to reduce computational cost and then the standard RLS algorithm can be summarized as in Table-II [68, 69].

Modification of the standard RLS into RLS CMA is not straightforward. If we put standard definition of the CMA error function, the RLS cost function in Eq. (26) takes the form of, say for x-polarization as:

$$J_x(n) = \sum_{i=1}^n \lambda^{n-i}\left| R_2 - \left|\mathbf{h}_x^H(n)\mathbf{u}_i(i)\right|^2 \right|^2. \qquad (27)$$

Due to the presence of the modulus nonlinearity in Eq. (27), the cost function of interest here is nonquadratic in the array weight vector $\mathbf{h}_x(n)$, which prevents the application of the standard RLS optimization techniques. To overcome this limitation, we can consider an intermediate data vector $\mathbf{z}_{x/y}$ and then apply the standard RLS like update [65, 70] as:

$$\mathbf{z}_x(n) = \mathbf{u}_i(n)\mathbf{u}_i^H(n)\mathbf{h}_x(n-1)$$

$$\mathbf{z}_y(n) = \mathbf{u}_i(n)\mathbf{u}_i^H(n)\mathbf{h}_y(n-1) \qquad (28)$$

$$\mathbf{k}_x(n) = \frac{\lambda^{-1}\mathbf{P}_x(n-1)\mathbf{z}_x(n)}{1+\lambda^{-1}\mathbf{z}_x^H(n)\mathbf{P}_x(n-1)\mathbf{z}_x(n)}$$

$$\mathbf{k}_y(n) = \frac{\lambda^{-1}\mathbf{P}_y(n-1)\mathbf{z}_y(n)}{1+\lambda^{-1}\mathbf{z}_y^H(n)\mathbf{P}_y(n-1)\mathbf{z}_y(n)} \qquad (29)$$

$$\varepsilon_x(n) = R_2 - \mathbf{h}_x^H(n-1)\mathbf{z}_x(n)$$

$$\varepsilon_y(n) = R_2 - \mathbf{h}_y^H(n-1)\mathbf{z}_y(n) \qquad (30)$$

$$\mathbf{h}_x(n) = \mathbf{h}_x(n-1) + \mathbf{k}_x(n)\varepsilon_x^*(n)$$

$$\mathbf{h}_y(n) = \mathbf{h}_y(n-1) + \mathbf{k}_y(n)\varepsilon_y^*(n) \qquad (31)$$

$$\mathbf{P}_x(n) = \lambda^{-1}\mathbf{P}_x(n-1) - \lambda^{-1}\mathbf{k}_x\mathbf{z}_x^H\mathbf{P}_x(n-1)$$

$$\mathbf{P}_y(n) = \lambda^{-1}\mathbf{P}_y(n-1) - \lambda^{-1}\mathbf{k}_y\mathbf{z}_y^H\mathbf{P}_y(n-1). \qquad (32)$$

B. 4×4 MIMO Configuration Enabling IQ Imbalance Compensation

Because the 2×2 equalizer structure as shown in Fig. 8 enforces joint filtering of the in- phase and quadrature signal components, four complex-valued independent filters are unable to compensate for any imbalance between the in-phase and quadrature components. This constraint associated with the complex 2×2 MIMO, can be removed by employing sixteen real-valued independent filters connected as shown in Fig. 10 to enable tracking and compensation of any receiver-side IQ imbalances.

Considering the subscript r and i of a variable as the real and imaginary part of that variable the outputs of the filters are given as:

$$x_{o_r}(n) = \mathbf{h}_{x_r x_r}^T \mathbf{x}_{i_r} + \mathbf{h}_{x_r x_i}^T \mathbf{x}_{i_i} + \mathbf{h}_{x_r y_r}^T \mathbf{y}_{i_r} + \mathbf{h}_{x_r x_i}^T \mathbf{y}_{i_i}$$

$$x_{o_i}(n) = \mathbf{h}_{x_i x_r}^T \mathbf{x}_{i_r} + \mathbf{h}_{x_i x_i}^T \mathbf{x}_{i_i} + \mathbf{h}_{x_i y_r}^T \mathbf{y}_{i_r} + \mathbf{h}_{x_i y_i}^T \mathbf{y}_{i_i}$$

$$y_{o_r}(n) = \mathbf{h}_{y_r x_r}^T \mathbf{x}_{i_r} + \mathbf{h}_{y_r x_i}^T \mathbf{x}_{i_i} + \mathbf{h}_{y_r y_r}^T \mathbf{y}_{i_r} + \mathbf{h}_{y_r y_i}^T \mathbf{y}_{i_i}$$

$$y_{o_i}(n) = \mathbf{h}_{y_i x_r}^T \mathbf{x}_{i_r} + \mathbf{h}_{y_i x_i}^T \mathbf{x}_{i_i} + \mathbf{h}_{y_i y_r}^T \mathbf{y}_{i_r} + \mathbf{h}_{y_r y_i}^T \mathbf{y}_{i_i}. \qquad (33)$$
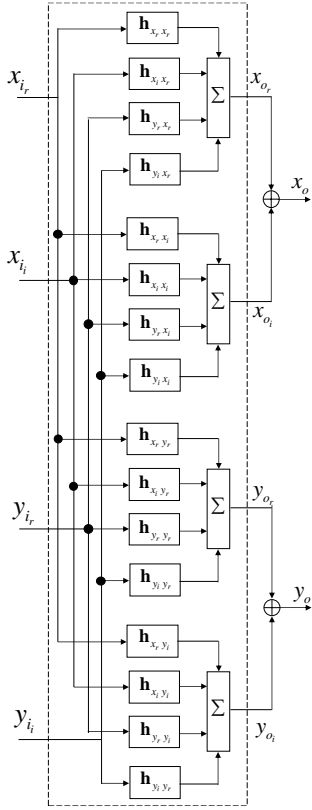
Fig. 10. Configuration of 4×4 MIMO real-valued FIR filters.
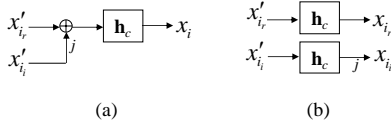

(a)   (b)

Fig. 11. Configuration of (a): conventional CD equalizer implementation causing IQ mixing, (b): IQ independent CD equalizer implementation.

Any standard adaption algorithm can be applied for updating sixteen filters. For example, in the case of CMA/RDE algorithm, which operates frequency offset and phase noise insensitively, the update equations can be written as:

$$\mathbf{h}_{x_r k}(n+1) = \mathbf{h}_{x_r k}(n) + \mu \varepsilon_x(n) x_{o_r}(n) k(n)$$

$$\mathbf{h}_{x_i k}(n+1) = \mathbf{h}_{x_i k}(n) + \mu \varepsilon_x(n) x_{o_i}(n) k(n)$$

$$\mathbf{h}_{y_r k}(n+1) = \mathbf{h}_{y_r k}(n) + \mu \varepsilon_y(n) y_{o_r}(n) k(n)$$

$$\mathbf{h}_{y_i k}(n+1) = \mathbf{h}_{y_i k}(n) + \mu \varepsilon_y(n) y_{o_i}(n) k(n) \qquad (34)$$

where $k \in \{x_r, x_i, y_r, y_i\}$ and ε_x and ε_y are given as in Eq. (19)/ Eq. (22) for the CMA/RDE.

While this 4×4 MIMO works well for short reach link or dispersion compensated link, for long-haul dispersion uncompensated link, we need to use a long CD filter prior to the adaptive filters which causes further IQ mixing which cannot be compensated with relatively shorter length adaptive filters. If we consider the classical CD compensation as convoluting the complex impulse response of CD filter $\mathbf{h}_c$ with its complex input x_i' as shown in Fig. 11 (a), in the complex

output x_i, further *IQ* mixing due to CD filter arises. However, such drawback can be avoided if the CD compensation is done independently on the real and imaginary part of the incoming signal as shown in Fig. 11 (b) [71].

C. Block Adaptive Filtering

For high-speed coherent communication systems, a high degree of parallelism is required to implement the DSP in the field-programmable gate array (FPGA) or CMOS ASIC. Therefore, it is desirable to implement the adaptive filters in block-by-block basis, rather than a sample-by-sample process, since this lends itself towards implementation in CMOS. Such block processing is applicable for optical communication since the block period is of the order of *ns* whereas the fastest change in optical channel is in *ms*.

In the subsequent subsection, to simplify the exposition, the adaptive filter for a single polarization case is described; however, it can easily be extended for dual polarization scenario considering the 2×2 MIMO structure as shown in Fig. 8.

1) Time Domain Implementation:

In the case of block adaptive filtering, the input sequence $x_i(n)$ is partitioned into *L*-point of blocks using a serial-to-parallel converter. The resulting block of input data is applied to an FIR filter **h** whose tap weights are kept constant over each block of data which produce a block of outputs. Output data samples $x_o(n)$ are then retained after parallel to serial conversion as shown in Fig. 12.
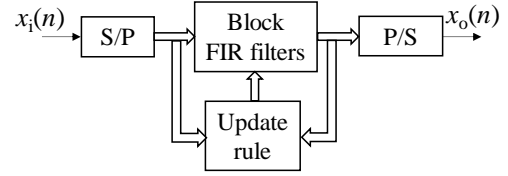

Fig. 12. Principle of block adaptive filtering.

If we consider the block index k, which relates to the sample index n as $n = kL + i$, where $i=1,2\cdots N$-1, the output is given as:

$$x_o(kL+i) = \mathbf{h}^H(k)\mathbf{x}_i(kL+i). \qquad (35)$$

The gradient to update the filter tap weights is an average of instantaneous gradients that are calculated over a block, and thus for the CMA the block update takes the form of

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \frac{\mu_B}{L}\sum_{i=0}^{L-1} e(kL+i)x_o^*(kL+i)\mathbf{x}_i(kL+i) \quad (36)$$

where μ is the convergence parameter and the error signal is calculated as

$$e(kL+i) = R_2 - \left|x_o(kL+i)\right|^2. \qquad (37)$$

2) Frequency Domain Implementation:

In some specific applications, such as to incorporate the MF for Nyquist pulse shaping in adaptive equalizer [42, 64] or to mitigate the differential mode group delay (DMGD) in mode-division multiplexed (MDM) systems [72-74], we require adaptive filters with large numbers of taps, causing a huge

computational burden. The key computational costs in time domain block adaptive filtering are the *linear convolution* of tap input and tap weights and the *linear correlation* between tap inputs and error signals. Both can be implemented in computationally efficient manner in frequency domain.

The most commonly used adaptive FDE is based on overlap-save method [32] as shown in Fig. 13. Considering 50% overlapping factor which is simple to implement, two input blocks of length N (N is number of tap weights) are concatenated and $2N$ point FFT operation is done for filter inputs $\mathbf{X}_i(k)$. Filters taps vector are padded by N zeros and transformed to frequency domain and multiplied with $\mathbf{X}_i(k)$ in element-by-element manner to get frequency domain output. The filtered output is then converted back to time domain using IFFT and as per overlap-save algorithm last half of the signal is saved as equalized output. Such a filtering operation can be summarized as below:

$$\mathbf{X}_i(k) = \text{FFT}\{[x_i(kN-N)\cdots x_i(kN+N-1)]\}$$

$$\mathbf{H}(k) = \text{FFT}\{[\mathbf{h}(k); \mathbf{O}_N]\}$$

$$\mathbf{X}_o(k) = \mathbf{X}_i(k) \circ \mathbf{H}(k)$$

$$\mathbf{x}_o(k) = \text{last } N \text{ elements of IFFT}\{\mathbf{X}_o(k)\} \quad (38)$$

where $\mathbf{O}_N$ is a column vector of length N with all elements zero and $\circ$ denotes the element-by-element multiplication.

The error signals for filter updates is then calculated in the time domain, N zeros are then appended after which it is transformed into the frequency domain. Considering the CMA, the error vector is computed as:

$$\mathbf{e}(k) = \{\mathbf{I}_N - \mathbf{x}_o(k) \circ \mathbf{x}_o^*(k)\} \circ \mathbf{x}_o(k)\}$$

$$\mathbf{E}(k) = \text{FFT}\{[\mathbf{O}_N; \mathbf{e}(k)]\} \quad (39)$$

where $\mathbf{I}_N$ is a column vector with all ones. The gradient vector is then calculated as:

$$\mathbf{s}(k) = \text{First } N \text{ elements of IFFT}\{\mathbf{X}_i^*(k) \circ \mathbf{E}(k)\}. \quad (40)$$

Finally, $\mathbf{s}(k)$ is appended with N zeros and filter is updated in frequency domain as

$$\mathbf{H}(k+1) = \mathbf{H}(k) + \mu \, \text{FFT}\{[\mathbf{s}(k); \mathbf{O}_N]\}. \quad (41)$$

Equations (40) and (41) place constraint on gradient vector which is required for a reliable linear convolution, albeit an additional FFT and IFFT are required.
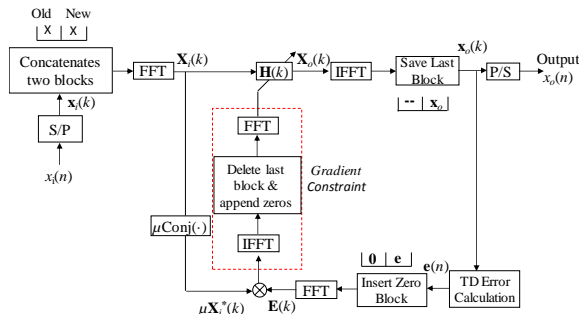

Fig. 13. Configuration of frequency-domain adaptive filtering with symbol-spaced sampled input sequences.

While the adaptive FDE of Fig. 13 works well for symbol-space input sequences a modification is required for a two-fold oversampled input sequence. For a two-fold oversampled input sequence, the samples are separated into two sequences corresponding to the even and odd samples (each sequence now can be considered symbol-spaced sampled) and splitting the filter into even and odd sub-filter with a length $N/2$ [75] as shown in Fig. 14. The filtering process can then be written as:

$$\mathbf{X}_o(k) = \mathbf{X}_{i_e}(k) \circ \mathbf{H}_e(k) + \mathbf{X}_{i_o}(k) \circ \mathbf{H}_o(k)$$

$$\mathbf{x}_o(k) = \text{last block of IFFT}\{\mathbf{X}_o(k)\}. \quad (42)$$

The error vector is calculated in similar manner as in Eq. (39) and the updates for even and odd filters are done as:

$$\mathbf{s}_{e/o}(k) = \text{First } N \text{ elements of IFFT}\{\mathbf{X}_{i_{e/o}}^*(k) \circ \mathbf{E}(k)\}$$

$$\mathbf{H}_{e/o}(k+1) = \mathbf{H}_{e/o}(k) + \mu \, \text{FFT}\{[\mathbf{s}_{e/o}(k); \mathbf{O}_N]\}. \quad (43)$$
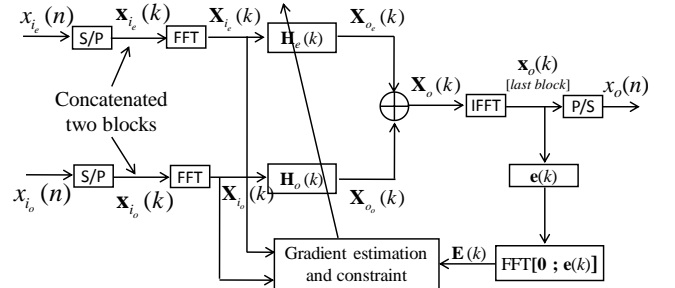

Fig. 14. Configuration of frequency-domain adaptive equalizer with even/odd sub-equalizer implementation concept enabling using two-fold oversampled input sequences.

Recently, frequency-domain RLS algorithm has been demonstrated for MDM systems [76-78] with much higher convergence rate than LMS algorithm; nevertheless, this requires additional complexity for calculating the Kalman gain vector and updating the correlation matrix.

It has been demonstrated that adaptive FDE can also be realized for rationally oversampled input sequence based on the sampling rate changing algorithm [79-81].

3) Hybrid Time-Frequency Domain Implementation:
While the adaptive FDE can significantly reduce the complexity, we cannot use the multiplier-free tap updates of the sign-sign CMA in frequency domain. This is because converting the time-domain error vector in frequency domain destroys the multiplier-free nature. However, a hybrid approach can be used combining an efficient frequency domain filtering and multiplier-free time domain filter tap weight updates using sign-sign CMA [82] as shown in Fig. 15. For a large number of taps, such hybrid approach reduces the complexity significantly even compare to FDE approach; because, we not only have multiplier-free updates but also FFT and IFFT used for implementing gradient constraint are no more required.
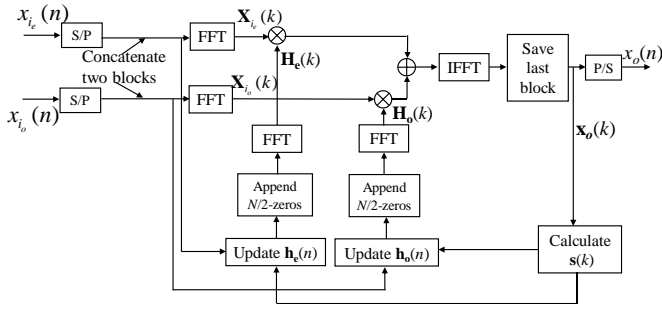
Fig. 15. Configuration of hybrid time/frequency domain implementation of adaptive filters.

While the frequency domain filtering process is same as in Eq. (42), the filter update is done in time domain as:

$$\mathbf{s}(k) = \text{sgn}\left\{\mathbf{I}_{N/2} - \mathbf{x_o}(k) \circ \mathbf{x_o^*}(k)\right\} \circ \text{csgn}\left\{\mathbf{x_o}(k)\right\}$$

$$\mathbf{h}_{e/o}(k+1) = \mathbf{h}_{e/o}(k) + \mu \sum_{i=0}^{N/2-1} \mathbf{x}_{i_{e/o}}^*(kN/2+i)s(kN/2+i). \quad (44)$$

The updated tap weights are then $N/2$ zeros padded and converted back to frequency domain for the filtering process.

Apart from low complexity implementations, the hybrid approach can also be effective for adaptive step-size implementation to improve the convergence rate as shown in [83].

VI. TIMING RECOVERY

Timing recovery circuits correct for the timing phase and frequency offset between the transmitter and receiver clocks. It can be done in feedback manner where the timing phase error signal is detected from sampled signals, which are then used to control voltage-controlled oscillator or to drive a digital interpolator or in a feedforward approach where the timing offset is estimated from blocks of samples, and then corrected by using digital interpolation.

The Gardner's method [84] is widely used due to its simplicity and independence of carrier phase. When the input complex signal $x_i(n)$ is two-fold oversampled, the classic Gardner's algorithm estimates the timing phase error which is given as:

$$\tau_{err} = \text{Re}\left\{\sum_{n=0}^{N/2-1} [x_i(2n-1) - x_i(2n+1)]x_i^*(2n)\right\} \quad (45)$$

where N is the number of samples used for estimation. Gardner's method can also be implemented in frequency-domain where Eq. (45) can be written as [85]:

$$\tau_{err} = \sum_{k=0}^{N/2-1} \text{Im}\left\{X_i(k)X_i^*(k+N/2)\right\} \quad (46)$$

where, $X_i(k)$ is DFT of $x_i(n)$. When frequency domain CD equalizer is used, this detector works directly on the frequency domain equalized signal. However, in the presence of PMD, data from both polarizations needs to be considered with further modifications in the timing error detection algorithm [86].

While Gardner's algorithm works satisfactory with classical NRZ/RZ signals, it fails for a Nyquist pulse shaped signal as

$\beta \rightarrow 0$. To overcome this shortcoming, the Gardner's phase detector can be modified where power of the incoming signal $P_i(n)$ is used instead of $x_i(n)$ as the input of the detector [87]. Since the input is a power signal, this phase detector is often referred as the fourth-power time-domain phase detector and the timing error is given as:

$$\tau_{err} = \sum_{n=0}^{N/2-1} [P_i(2n-1) - P_i(2n+1)]P_i(2n). \quad (47)$$

Frequency-domain implementation of fourth power phase detector is desirable, since the earliest point of DSP where timing can be estimated is the CD equalized signal which is in generally in frequency domain. The frequency domain implementation was developed in [88] which can be express mathematically as:

$$\tau_{err} = \sum_{k=-N'/2}^{N'/2} \text{Im}\left\{\left(\sum_{l=0}^{N'} X_i[l]X_i^*(l+k+N_{FFT}-N')\right)\right.$$
$$\left. \times \left(\sum_{l=0}^{N'} X_i[l+N_{FFT}-N']X_i^*(l+k)\right)^*\right\} \quad (48)$$

where N_{FFT} is the block length and N' is a power-of-two integer such that $N' < N$.

Regarding the jitter tolerance performance, the conventional Gardner's algorithm has too much jitter induced penalty for pulse shaping with $\beta < 0.1$ while 4th power time domain method works better in the range $0 < \beta < 0.6$. The 4th power frequency domain method shows superior performance and its performance is insensitive for any β.

The other classes of timing algorithm are square law algorithm [89] which operate on four-fold oversample sequence and Mueller-Müller algorithm [90] operating on symbol-spaced sampled signal. Mueller-Müller algorithm has an excellent jitter tolerance capability for all β, however since it needs to be implemented after carrier phase estimation in DSP chain, the significant latency imposed lower operating loop bandwidth.

VII. FREQUENCY OFFSET ESTIMATION

In an intradyne coherent receiver, the free running local oscillator and the transmitter lasers are not frequency locked. This results in some residual frequency offset in the received signal which should be estimated and compensated at DSP. When the input signal is in the form of $x_i(n) = x_{sym}(n)\exp\left(j\left[\theta(n) + n2\pi\Delta f T_{sym}\right]\right)$, the task of the system is to estimate Δf. The frequency estimation methods can be widely classified as blind and training-aided method.

A. Blind Estimation

1) Time-Domain Differential Phase Based Method:

The differential phase based method is illustrated in Fig. 16 where QPSK modulation format is considered as an example. Frequency offset is extracted from the average phase increment of two consecutives symbols and data modulation phase is removed with 4th power operation [91].
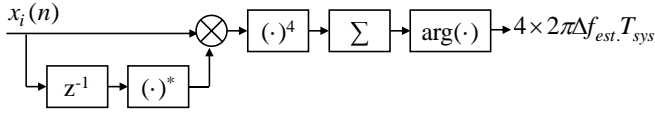
Fig. 16. Illustration of time-domain differential phase based frequency offset estimation scheme.

The estimated frequency offset can be expressed as:

$$\Delta f_{est.} = \frac{1}{8\pi T_{sys}} \arg \left\{ \sum_{n=1}^{L} \left[x_i(n) x_i^*(n-1) \right]^4 \right\} \quad (49)$$

where L is the averaging span. Though such time domain method can be easily used for M-ary PSK signals, the performance becomes poor for high-order QAM because only a small portion of the constellation points having equal phase spacing to extract frequency offset.

For accurate frequency offset estimation using Eq. (49), we need a large value averaging span (in some cases even several thousands). As an alternative, another estimator may be obtained by reversing the order of the operations, and iteratively, estimating the frequency offset with a convergence parameter μ as [92]:

$$\Delta f(n) = \Delta f(n-1)(1-\mu) + \mu \frac{\left[x_i(n) x_i^*(n-1) \right]^4}{8\pi T_{sys}}. \quad (50)$$

2) Frequency Domain Methods:

Since the spectrum of x_i^4 exhibits a peak at the frequency of four times the frequency offset, Δf can be estimated from the spectrum. The spectrum based method provides a better performance than time domain approach (for same block length); however, implementation complexity increases for high-order modulation formats. Moreover, a single discrete Fourier transform only gives only magnitude of Δf and an additional FFT is required to get the sign of Δf. Note that the complexity can be reduce for high order square QAM considering only outmost four constellation points combined with the use of linear interpolation and down sampling-based methods [93]. Alternatively, an iterative method can be employed to estimate Δf [94].

3) Blind Frequency Search Methods:

In this method, the frequency is scanned over a range and for each trial frequency, the carrier phase is first recovered and decisions of symbol is made and then the minimum mean square error is calculated and used as the frequency-selection criteria [95]. To reduce the implementation effort, the frequency offset is swept first in a coarse step with large deviation and then with a fine step with smaller step size as shown in Fig. 17.

While the methods discussed so far are essentially feedforward techniques, feedback techniques employing a frequency-controlled loop may also be used, having the advantage that they are agnostic to the modulation format [96-98].
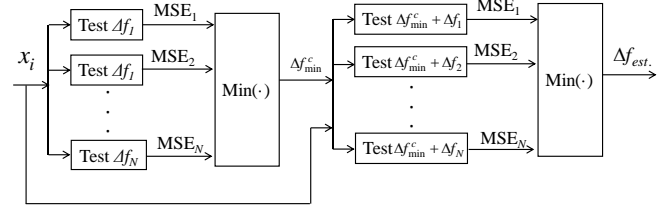

Fig. 17. Concept of two-stage blind frequency search method for frequency offset estimation.

B. Training-Aided Estimation

The 4th power operation can be avoided and frequency estimation can be done over a wide range (half of the symbol rate) for any modulation formats at an expense of training overhead [99]. The training symbols are periodically inserted with the payload. In the DSP, first the training symbols $x_t(n)$ is identified with a suitable method such as Schmidl-Cox algorithm [100] and then frequency offset is estimated as shown in Fig. 18. The modulation phase is removed using training symbols

$$x_i'(n) = x_i(n) x_t^*(n). \quad (51)$$

After that the frequency offset can be estimated in a similar manner as in Eq. (49) as:

$$\Delta f_{est.} = \frac{1}{2\pi T_{sys}} \arg \left\{ \sum_{n=1}^{L} \left[x_i'(n) x_i'^*(n-1) \right] \right\}. \quad (52)$$

Digital pilot symbols can be also used to estimate frequency offset by observing the location of pilot in the spectrum [101]. Such a method is modulation format independent and can accurately estimate Δf in presence of CD and PMD.
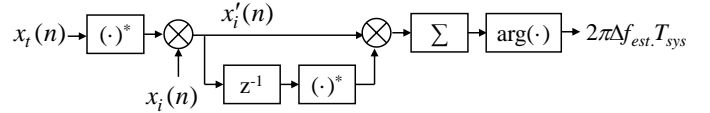

Fig.18. Block diagram of training symbol aided frequency offset estimation technique

VIII. CARRIER PHASE RECOVERY

Similar to frequency-offset estimation algorithms, carrier recovery algorithms can be classified as either blind or data-aided estimation techniques. Moreover, the algorithms can be implemented in feedforward manner or in feedback structure albeit feedforward is generally preferable for hardware implementation. When the input signal is in the form of $x_i(n) = x_{sym}(n) \exp(j\theta(n)) + w(n)$, the task of the system is to estimate θ.

A. Blind Estimation

For QPSK the carrier phase may be estimated using a fourth-order nonlinearity in order to remove the phase modulation, giving the estimate of the phase as follows

$$\theta_{est.}(n) = \frac{1}{4} \arg \left\{ \frac{1}{2L+1} \sum_{l=-L}^{L} f(l) x_i^4(l+n) \right\} \quad (53)$$

where $f(l)$ is a weighting function, which depends on the ratio of the additive white Gaussian noise to the laser phase noise [102]. With $f(l)=1$, and the estimator becomes that given by Viterbi and Viterbi [103]. The result of the weighting function is to apply a Wiener filter to estimate the phase noise [102, 104], which can approach the performance of an ideal MAP estimator of the phase.

One particularly hardware-efficient phase estimator is the Barycenter algorithm in which the order of the operations in Eq. (53) are reversed to give [105]:

$$\theta_{est.}(n) = \frac{1}{4(2L+1)} \sum_{l=-L}^{L} f(l) \arg\left\{x_i^4(l+n)\right\}. \quad (54)$$

However, once the modulation format moves to high order QAM, we need to modify the algorithm for its applicability. One of the straightforward technique is to use the only constellation points that belongs to QPSK like shape [106], for example constellation points of QPSK like class-I category of 16-QAM format as in Fig. 19. However, due to discarding a significant number of symbols, this method has less linewidth tolerance. The performance can be improved by QPSK portioning technique [107] where the symbols of the class-II is also used with some rotation operations. However, this technique gets complicated when modulation order is beyond 16-QAM.
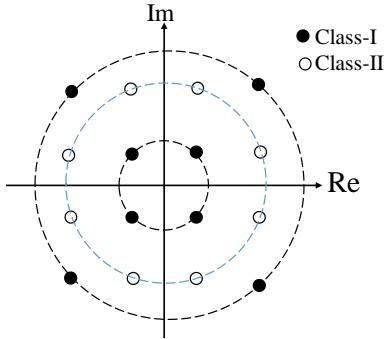

Fig. 19. Symbol classification for the 16-QAM signal.

Nevertheless, using conventional wireless approaches such as decision-directed phase locked loops (DD-PLL) can be applied for any QAM signals [65, 108], however the feedback presents challenges for CMOS-based parallel implementation [109].

As an alternative, binary phase search (BPS) algorithm can be used in a feedforward manner and realizable in parallel [109]. In this case, the carrier phase is scanned over a range (0 to $\pi/2$ for square QAM) and for each trial phase, mean square error is calculated using Euclidian distance between phase-rotated symbol and its decided symbol. Then the trial phase that gives minimum MSE is considered as the estimated phase. To overcome the ASE noise effect, distances of $2L$ symbols rotated by same angle are summed up to calculate MSE. Since the decoding is already done in test phase stage, decoded symbol can be selected using a switch selected by the index of minimum MSE. Though BPS shows a good tolerance to laser phase noise, with an increasing modulation order a large number of test phases are required which implied the computational burden.
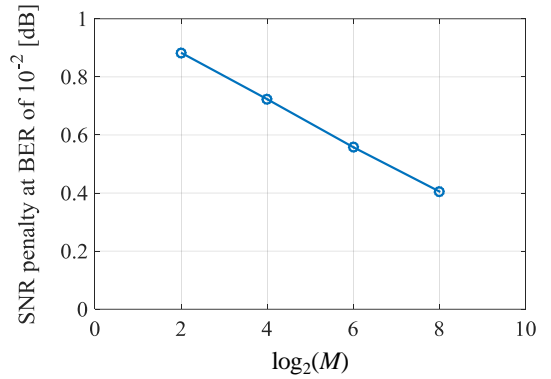

Fig. 20. SNR penalty at 10^{-2} BER due to different coding for different modulation formats. M is the modulation order.

Due to fourfold ambiguities of square QAM, the blind algorithms discussed so far may cause incorrect phase estimation by a multiple of $\pi/2$ causing cycle slip. However, this problem can be solved using differential encoding/decoding at an expense of a modest sensitivity penalty. However, since we apply the differential coding on quadrant basis instead of all the symbols of constellation points, the coding penalty reduces with increasing modulation order as shown in Fig. 20.

B. Pilot-Symbol Aided Estimation

Unlike the blind estimation techniques, pilot-symbol based methods exhibit more immunity to cycle slips, albeit at the expense of transmission overhead. In this technique, the pilot symbols are time multiplexed with data payload. Considering a pilot rate of $1/P$ (one pilot symbol is inserted periodically after P-1 payload symbols), we define the pilot sequence $p(n)$ which is a zero-padded sequence such that

$$p(n) = \begin{cases} x_{sym.}(n), & n = iP \\ 0, & n \neq iP \end{cases} \quad (55)$$

where $i = 1, 2, 3\cdots$ and $x_{sym.}(n)$ is considered as the unitary average power symbol sequence, including payload and pilot symbols.

First time synchronization of pilot symbol is obtained by correlating the received sequence with the pilot sequence. Then carrier recovery is done as shown in Fig. 21. The incoming sequence is multiplied with the complex conjugate of $p(n)$ and the resulting sequence is filtered by an interpolation filter $h_f(n)$ to estimate the phase noise [110, 111].
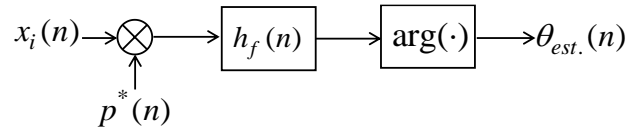

Fig. 21. Concept of training symbol aided carrier phase estimation.

Note that significant performance improvement can be achieved by using multi-pilot-symbols aided technique where for each phase estimation few preceding and following pilot symbols are used [112, 113].

C. Multi-Stage Estimation

Recently, significant research efforts have been made in the area of multi-stage carrier recovery algorithms suitable for high-order modulation formats to reduce the implementation complexity and/or improve the linewidth tolerance. The key idea behind these algorithms is to use a coarse-phase estimate at the first stage with a hardware-efficient but less-accurate phase estimator and a fine-phase estimator in the second or subsequent stages. The first stage algorithm can be data-aided [110] or non-data aided like BPS with coarse phase step [114-116], Viterbi-Viterbi algorithm with QPSK partitioning [117-119] or DD-PLL [120,121]. Once the first stage has provided a coarse estimate the fine-stage estimation can be realized using BPS with finer step size [115, 120, 121], maximum likelihood (ML) estimator [114, 118, 120] or QPSK constellation transformed based algorithm [117]. The choice of the combination is to be made based on complexity and performance requirements.

IX. Symbols Estimation and Decoding

Following carrier recovery, the signal may be decoded by the outer receiver. This could be done using a soft-decision forward error correction (FEC) with a Galois field corresponding to the symbol alphabet, or symbol estimation followed by hard-decision FEC. For hard decision decoding of binary data, symbol estimation and bit decoding is required. For rectangular constellations, such as QAM, this may be achieved by applying a series of decision thresholds to the in-phase and quadrature components separately and it can be done in a hardware efficient manner using logic circuits. While this technique works optimally for a system limited by AWGN, in presence of nonlinear phase noise non-rectangular decision boundaries can improve the performance [122] with an expense of hardware complexity. To address this issue, several techniques has been investigated such as ML based decision boundaries [123] or machine learning detector based on support vector machine (SVM) [124] or *k*-nearest neighbors [56].

X. Current Research Challenges

Herein we highlight some of the current research challenges relating to DSP for multilevel formats. The first area is associated with the modification of DSP algorithm from PDM-QPSK to PDM-QAM format. While many algorithms operated well for constant modulus format, new design methodologies are required for multi-level formats. For example, adaptive blind equalization while the CMA has been extensively used for PDM-QPSK an equivalent to CMA that is robust for high levels of QAM has yet to be determined. Likewise, Viterbi-Viterbi based carrier recovery algorithms works well with PDM-QPSK while its evolution for high order QAM is still challenging. The situation is further complicated by introducing stronger FEC with a higher net coding gain since this results in a reduced signal to noise ratio. This results in needing to design robust algorithms for equalization and synchronization that can recover the data from noisy, dispersed and nonlinearly distorted signals.

A second area of research is to develop DSP algorithms that can be realized in hardware efficiently. One of the key concerns in hardware realization is the reduction of computational complexity in order to reduce power consumption. Of the algorithms discussed in the paper, the key area of particular concern is the DSP algorithms associated with nonlinear compensation, since the computational complexity of many of the techniques explored thus far are prohibitive for implementation in a CMOS ASIC. More generally, since the CMOS DSP cannot operate at the sampling clock frequency of the ADC, the algorithms need to be developed that allow parallel processing and minimize the feedback latency in order to make real-time implementation practical.

A third area of future research arises with the development of dynamic elastic optical networking in which transceivers not only vary their rate but may be dynamically dropped and added, for example, to cope to the so called "elephant flows" between data centers. This requires developing both rate adaptive DSP and fast acquisition algorithms to minimize the time taken to establish a wavelength on demand service. Modulation format transparent DSP development is one of the key concerns for such networks. The area of intelligent transceivers for dynamic elastic optical networking also presents numerous challenges as DSP allows the transceivers to become intelligent agents that can dynamically utilize the available network resources, varying rate and bandwidth utilization in response to dynamical network demands.

A final area is that of co-designing the DSP and photonics in optical transceivers to trade performance against complexity, cost, and power consumption. For example, in cost sensitive applications, such as access networks, the challenge becomes to design DSP that can relax the requirements on the photonic components to reduce the overall cost. On the other hand, for performance critical applications, such as submarine systems, the requirement for DSP is to maximize point-to-point capacity.

XI. Concluding Remarks

DSP based coherent technology has caused a revolution in the design of optical core networks. Nevertheless, direct detection is often used in shorter-reach systems, such as metropolitan networks, due to power consumption and cost constraints associated digital coherent receivers. With an expectation that advancement of DSP algorithms will relax the photonics requirement significantly and the DSP itself will require less power, it is anticipated that digital coherent transceivers will become ubiquitous from core through to access networks. The key benefit of DSP over analog alternatives may be to improve the overall cost of optical transmission with the co-design of the DSP and photonics technology being particularly fruitful. Thus, it can be expected that within the next decade DSP with spectrally efficient modulation formats, will become standard for access and perhaps even data center networks, just as it has done for long-haul systems over the last decade.

References

[1] A. Färbert, S. Langenbach, N. Stojanovic, C. Dorschky, T. Kupfer, C. Schulien, et al., "Performance of a 10.7 Gb/s receiver with digital

equaliser using maximum likelihood sequence estimation," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2004, paper Th.4.1.5.

[2] J. McNicol, M. O'Sullivan, K. Roberts, A. Comeau, D. McGhan, and L. Strawczynski, "Electronic domain compensation of optical dispersion," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2005, paper OThJ3.

[3] D.-S. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature phase-shift keying signals with carrier phase estimation," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 12–21, Jan. 2006.

[4] F. Derr, "Coherent optical QPSK intradyne system: Concept and digital receiver realization," *J. Lightw. Technol.*, vol. 10, no. 9, pp. 1290–1296, Sep. 1992.

[5] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.

[6] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "A novel multilevel coherent optical system: 4-quadrature signaling," *J. Lightw. Technol.*, vol. 9, no. 4, pp. 514–523, Apr. 1991.

[7] S. J. Savory and J. Whiteaway, "Digital compensation for optical transmission system," *U.S. patent 7266310,* published Sep. 2007.

[8] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no.2, pp. 804-817, Jan. 2008.

[9] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Exp.*, vol. 16, no.2, pp. 873-879, Jan. 2008.

[10] K. Roberts, S. H. Foo, M. Moyer, M. Hubbard, A. Sinclair, J. Gaudette, et al., "High capacity transport-100G and beyond," *J. Lightw. Technol.*, vol. 33, no. 3, pp. 563-578, Feb. 2015.

[11] S. J. Savory, "Digital signal processing for multilevel modulation formats," In *Proc. Eur. Conf. Opt. Commun.*, Sep. 2016, paper Tu.2.A.1.

[12] K. Kikuchi, "Fundamentals of Coherent Optical Fiber Communications," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 157-179, Jan. 2016.

[13] S. J. Savory, "Digital coherent optical receivers: algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164-1179, Sep./Oct. 2010.

[14] K. Kikuchi and S. Tsukamoto, "Evaluation of sensitivity of the digital coherent receiver," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1817-1822, Jul. 2008.

[15] B. Zhang, C. Malouin, and T. J. Schmidt, "Design of coherent receiver optical front end for unamplified applications," *Opt. Exp.*, vol. 20, no. 3, pp. 3225-3234, Jan. 2012.

[16] C. Fludger, T. Kupfer, "Transmitter impairment mitigation and monitoring for high baud-rate, high order modulation systems," in *Proc. Eur. Conf. Opt. Commun.*, Sept. 2016, paper Tu.2.A.2.

[17] I. Fatadin, S. J. Savory, and D. Ives, "Compensation of quadrature imbalance in an optical QPSK coherent receiver," *IEEE Photon. Technol. Lett.*, vol. 20, no. 20, pp. 1733–1735, Oct. 2008.

[18] P. O. Löwdin, "On the non- orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals," *J. Chem. Phy.*, vol. 18, no. 3, pp. 365-375, Mar. 1950.

[19] I. Mayer, "On Lowdin's method of symmetric orthogonalization," *Int. J. Quantum Chem.*, vol. 90, no. 1, pp. 63–65, Oct. 2002.

[20] T. Tanimura, S. Oda, T. Tanaka, T. Hoshida, Z. Tao, and J. Rasmussen, "A simple digital skew compensator for coherent receiver," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2009, paper 7.3.2.

[21] J. C. Diniz, E. P. da Silva, M. Piels, and D. Zibar, "Joint IQ skew and chromatic dispersion estimation for coherent optical communication receivers," in *Proc. Signal Process. Photon. Commun.*, Jul. 2016, paper SpTu2F.2.

[22] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front end for optical quadrature amplitude modulation," *IEEE Photon. J.*, vol. 5, no. 2, Art. 7800110, Apr. 2013.

[23] M. Paskov, D. Lavery, S. J. Savory, "Blind equalization of receiver in-phase/quadrature skew in the presence of Nyquist filtering," *IEEE Photon. Technol. Lett.*, vol. 25, no. 24, pp. 2446–2449, Dec. 2013.

[24] C. Laperle, B. Villeneuve, Z. Zhang, D. McGhan, H. Sun, and M. O'Sullivan, "WDM performance and PMD tolerance of a coherent 40-Gbit/s dual-polarization QPSK transceiver," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 168-175, Jan. 2008.

[25] P. M. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus, "Field trial results on statistics of fast polarization changes in long

haul WDM transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2005, paper OThT6.

[26] S. Salaun, F. Neddam, J. Poirrier, B. Rayguenes, and M. Moignard, "Fast SOP variation measurements on WDM systems are OPMDC fast enough?," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2009, paper P1.04.

[27] G.P. Agrawal, *Nonlinear Fiber Optics*, Academic Press, 2001.

[28] S. J. Savory, "Digital equalization in coherent optical transmission systems," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*, chap. 8, 2016.

[29] A. Eghbali, H. Johansson, O. Gustafsson, and S. J. Savory, "Optimal least-squares FIR digital filters for compensation of chromatic dispersion in digital coherent optical receivers," *J. Lightw. Technol.*, vol. 32, no. 8, pp. 1449-1456, Apr. 2014.

[30] G. Goldfarb and G. Li, "Chromatic dispersion compensation using digital IIR filtering with coherent detection," *IEEE Photon. Technol. Lett.*, vol. 19, no.13, pp. 969-971, Jul. 2007.

[31] T. Pfau, "Real-time implementation of high speed digital coherent transceivers," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*, Chap. 12, 2016.

[32] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no.1, pp. 14-37, Jan. 1992.

[33] J. C. Geyer, C. R. Fludger, T. Duthel, C. Schulien, and B. Schmauss, "Efficient frequency domain chromatic dispersion compensation in a coherent Polmux QPSK-receiver," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2010, paper OWV5.

[34] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, and Y. Miyamoto, "Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3721–3728, Aug. 2009.

[35] T. Xu, G. Jacobsen, S. Popov, M. Forzati, J. Mårtensson, M. Mussolin, J. Li, K. Wang, Y. Zhang, and A. T. Friberg, "Frequency domain chromatic dispersion equalization using overlap-add methods in coherent optical system", *J. Opt. Commun.*, vol. 32, no. 2, pp.131-5, Jan. 2011.

[36] K.-P. Ho, "Subband equaliser for chromatic dispersion of optical fibre,", *Electron. Lett.*, vol. 45, no. 24, Nov. 2009.

[37] I. Slim, A. Mezghani, L. G. Baltar, J. Qi, F. N. Hauske, J. A. Nossek, "Delayed single-tap frequency-domain chromatic-dispersion compensation," *IEEE Photon. Technol. Lett.*, vol. 25, no. 2, pp. 167-170, Jan. 2013.

[38] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers," *J. Lightw. Technol.*, vol. 29, no. 1, pp. 53-61, Jan. 2011.

[39] B. Châtelain, C. Laperle, K. Roberts, M. Chagnon, X. Xu, A. Borowiec, F. Gagnon, and D. V. Plant, "A family of Nyquist pulses for coherent optical communications," *Opt. Exp.*, vol. 20, no. 8, pp. 8397-8416, Apr. 2012.

[40] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: Mcgraw-Hill, 2001.

[41] J. Wang, C. Xie, and Z. Pan, "Matched filter design for RRC spectrally shaped Nyquist-WDM systems," *IEEE Photon. Technol. Lett.*, vol. 25, no. 23, pp. 2263–2266, Dec. 2013.

[42] C. Zhu, B. Corcoran, A. V. Tran, and A. J. Lowery, "Nyquist-WDM with low-complexity joint matched filtering and adaptive equalization," *IEEE Photon. Technol. Lett.*, vol. 26, no. 23, pp. 2323-2326, Dec. 2014.

[43] A. Mecozzi, Rene Essiambre, "Nonlinear Shannon Limit in Pseudolinear Coherent Systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011-2024, Jun. 2012.

[44] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

[45] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416-3425, Oct. 2008.

[46] D. S. Millar, S. Makovejs, C. Behrens, S. Hellerbrand, R. I. Killey, P. Bayvel, Seb J. Savory, "Mitigation of fiber nonlinearity using a digital coherent receiver," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no.5, pp. 1217-1226, Sep./Oct. 2010.

[47] K. Roberts, L. Chuandong, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photon. Technol. Lett.*, vol. 18, no. 2, pp. 403-405, Jan. 2006.

[48] D. Lavery, D. Ives, G. Liga, A. Alvarado, S. J. Savory and P. Bayvel, "The Benefit of Split Nonlinearity Compensation for Single-Channel Optical Fiber Communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 17, pp. 1803-1806, Sep. 2016.

[49] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightw. Technol.*, vol. 29, no.17, pp. 2570-2576, Sep. 2011.

[50] Y. Gao, J. C. Cartledge, A. S. Karar, S. S.-H. Yam, M. O'Sullivan, C. Laperle, et al., "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," *Opt. Exp.*, vol. 22, no.2, pp. 1209-1219, Jan. 2014.

[51] R. Maher, T. Xu, L. Galdino, M. Sato, A. Alvarado, K. Shi, et al., "Spectrally shaped DP-16QAM super-channel transmission with multi-channel digital back-propagation," *Sci. Rep.*, vol. 5, Art. 8214, Feb. 2015.

[52] N. K. Fontaine, X. Liu, S. Chandrasekhar, R. Ryf, S. Randel, P. Winzer, et al., "Fiber nonlinearity compensation by digital backpropagation of an entire 1.2-Tb/s superchannel using a full-field spectrally-sliced receiver," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2013, Mo.3.D.5.

[53] R. Dar and P. J. Winzer, "On the limits of digital back-propagation in fully loaded WDM systems," *IEEE Photon. Technol. Lett.*, vol. 28, no. 11, pp. 1253-1256, Jun. 2016.

[54] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," *Opt. Exp.*, vol. 20, no. 2, pp. 1360–1368, Jan. 2012.

[55] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Parts I–III," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4369, Jul. 2014.

[56] D. Wang, M. Zhang, M. Fu, Z. Cai, Z. Li, H. Han, et al., "Nonlinearity mitigation using a machine learning detector based on *k*-nearest neighbors," *IEEE Photon. Technol. Lett.*, vol. 28, no. 19, pp. 2102-2105, Oct. 2016.

[57] Y. Cai, D. G. Foursa, C. R. Davidson, J. X. Cai, O. Sinkin, M. Nissov, and A. Philipetskii, "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmission," *in Proc. Opt. Fiber Commun. Conf.*, Mar. 2010, paper OTuE1.

[58] D. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[59] K. Kikuchi, "Performance analyses of polarization demultiplexing based on constant-modulus algorithm in digital coherent optical receivers," *Opt. Exp.*, vol. 19, no. 10, pp. 9868-9880, May 2011.

[60] L. Liu, Z. Tao, W. Yan, S. Oda, T. Hoshida, and J. Rasmussen, "Initial tap setup of constant modulus algorithm for polarization de multiplexing in optical coherent receivers," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2009, paper OMT2.

[61] C. Xie and S. Chandrasekhar, "Two-stage constant modulus algorithm equalizer for singularity free operation and optical performance monitoring in optical coherent receiver," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2010, paper OMK3.

[62] M. S. Faruk, Y. Mori, C. Zhang, and K. Kikuchi, "Proper polarization demultiplexing in coherent optical receiver using constant modulus algorithm with training mode," in *Proc. 15th OptoElectron. Commun. Conf.*, Jul. 2010, pp. 768–769.

[63] D. Cardenas, D. Lavery, P. Watts, and S. J. Savory, "Reducing the power consumption of the CMA equalizer update for a digital coherent receiver," in *Proc. Opt. Fib. Commun. Conf.*, Mar. 2014, paper Th4D.5.

[64] D. Lavery, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Reduced complexity equalization for coherent long-reach passive optical networks [Invited]," *J. Opt. Commun. Netw.*, vol. 7, no. 1, pp. A16-A27, Jan. 2015.

[65] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, Aug. 2009.

[66] D. Lavery, M. Paskov, R. Maher, S. J. Savory, and P. Bayvel, "Modified radius directed equaliser for high order QAM," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2015, paper We.4.D.5.

[67] S. U. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no.9, pp. 1349-1387, Sep. 1985.

[68] S. Haykin, *Adaptive filter theory* Englewood Cliffs, NJ: Prentice-Hall, 2001.

[69] B. Farhang-Boroujeny, *Adaptive filters–theory and applications*, Wiley 1998.

[70] Y. Chen, T. Le-Ngoc, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Trans. Signal Process.*, vol. 52, no.5, pp. 1452-1456, May 2004.

[71] R. Rios-Müller, J. Renaudier, and G. Charlet, "Blind receiver skew compensation and estimation for long-haul non-dispersion managed systems using adaptive equalizer," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1315-1318, Apr. 2015.

[72] G. Li, N. Bai, N. Zhao, and C. Xia, "Space-division multiplexing: the next frontier in optical communication," *Adv. Opt. Photon.*, vol. 6, no. 4, pp. 413-487, Dec. 2014.

[73] S. Ö. Arik, D. Askarov, and J. M. Kahn, "Effect of mode coupling on signal processing complexity in mode-division multiplexing," *J. Lightw. Technol.*, vol. 31, no. 3, pp. 423-431, Feb. 2013.

[74] B. Inan, B. Spinnler, F. Ferreira, D. van den Borne, A. Lobato, S. Adhikari, et al., "DSP complexity of mode-division multiplexed receivers," *Opt. Exp.*, vol. 20, no. 10, pp. 10859-10869, May 2012.

[75] M. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent optical receivers," *Opt. Exp.*, vol. 19, no. 13, pp. 12789-12798, Jun. 2011.

[76] S. Ö. Arık, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 1841-1852, May 2014.

[77] Z. Yang, J. Zhao, N. Bai, E. Ip, T. Wang and G. Li, "Experimental demonstration of adaptive recursive least square frequency-domain equalization for long-distance mode-division multiplexed transmission," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2015, Th.1.6.2.

[78] Y. Weng, T. Wang, and Z. Pan, "Fast-convergent adaptive frequency-domain recursive least-squares algorithm with reduced complexity for MDM transmission systems using optical few-mode fibers," in *Proc. CLEO: Sci. Innov.*, Jun. 2016, paper SW4F.6.

[79] M. S. Faruk and K. Kikuchi, "Frequency-domain adaptive equalizer with rational oversampling rates in coherent optical receivers," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2014, paper P.3.22.

[80] M. Paskov, D. Millar, and K. Parsons, "A fully-blind fractionally-oversampled frequency domain adaptive equalizer," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2016, paper Th2A.33.

[81] S. F. A. Shah, L. Wang, C. Li, and Z. Zhang, "Low-complexity design of noninteger fractionally spaced adaptive equalizers for coherent optical receivers," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1289-1293, Sep. 2016.

[82] M. S. Faruk, D. Lavery, R. Maher, and S. J. Savory, "A low complexity hybrid time-frequency domain adaptive equalizer for coherent optical receivers," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2016, paper Th2A.39.

[83] K. Shi and B. C. Thomsen, "Sparse adaptive frequency domain equalizers for Mode-Group division multiplexing," *J. Lightw. Technol.*, vol. 33, no. 2, pp. 311-317, Jan. 2015.

[84] F. M. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May 1986.

[85] X. Zhou, "Efficient clock and carrier recovery algorithms for single-carrier coherent optical systems: A systematic review on challenges and recent progress," *IEEE Signal Process. Mag.*, vol. 2, no. 2, pp. 35-45, Mar. 2014.

[86] H. Sun and K.-T. Wu, "A novel dispersion and PMD tolerant clock phase detector for coherent transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2011, paper OMJ4.

[87] M. Yan, Z. Tao, L. Dou, L. Li, Y. Zhao, T. Hoshida, et al., "Digital clock recovery algorithm for Nyquist signal," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2013, paper OTu2I.7.

[88] K.-T. Wu and H. Sun, "Frequency-domain clock phase detector for Nyquist WDM systems," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2014, paper Th3E.2.

[89] M. Oerder, H. Meyr, "Digital filter and square timing recovery", *IEEE Trans. Commun.*, vol. 36, no. 5, May 1988.

[90] K.H. Mueller, M. S. Müller, "Timing recovery in digital synchronous data receivers," *IEEE Trans. Commun.*, vol. COM-24, pp. 516-531, May 1976.

[91] A. Leven, N. Kaneda, U. V. Koc, and Y. K. Chen, "Frequency estimation in intradyne reception," *IEEE Photon. Technol. Lett.*, vol. 19, no. 6, pp. 366–368, Mar. 2007.

[92] S. Hoffmann, S. Bhandare, T. Pfau, O. Adamczyk, C. Wordehoff, R. Peveling, M. Porrmann, and R. Noe, "Frequency and phase estimation for coherent QPSK transmission with unlocked DFB lasers," *IEEE Photon. Technol. Lett.*, vol. 20, no. 18, pp. 1569–1571, Sep. 2008.

[93] Y. Sun and X. Zhou, "Blind carrier frequency recovery methods for coherent receivers using QAM modulation formats," *U.S. patent,* application 13/109853, May 2011.

[94] M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2009, paper P. 3.08.

[95] X. Zhou, J. Yu, M. -F. Huang, Y. Shao, T. Wang, L. E. Nelson, P. D. Magill, M. Birk, P. I. Borel, D. W. Peckham, and R. Lingle, "64-Tb/s, 8 b/s/Hz, PDM-36QAM transmission over 320 km using both pre- and post-transmission digital signal processing," *J. Lightwave Technol.*, vol. 29, no. 4, pp. 571–577, Feb. 2011.

[96] K. Piyawanno, M. Kuschnerov, B. Spinnler, and B. Lankl, "Fast and accurate automatic frequency control for coherent receivers," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2009, paper 7.3.1.

[97] I. Fatadin and S. J. Savory, "Compensation of frequency offset for 16-QAM optical coherent systems using QPSK partitioning," *IEEE Photon. Technol. Lett.*, vol. 23, no. 17, pp. 1246-1248, Sep. 2011.

[98] Y. Mori and K. Kikuchi, "Dual-stage decision-directed phase estimator enabling perfect frequency-offset elimination in digital coherent optical receivers," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2013, paper OTu3I.7.

[99] X. Zhou, X. Chen, and K. Long, "Wide-range frequency offset estimation algorithm for optical coherent systems using training sequence," *IEEE Photon. Technol. Lett.*, vol. 24, no. 1, pp. 82-84, Jan. 2012.

[100] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[101] D. Zhao, L. Xi, X. Tang, W. Zhang, Y. Qiao, and X. Zhang, "Digital pilot aided carrier frequency offset estimation for coherent optical transmission systems," *Opt. Exp.*, vol. 23, no.19, pp. 24822-24832, Sep. 2015.

[102] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing," *J. Lightw. Technol.*, vol. 27, no. 7, pp. 901–914, Apr. 2009.

[103] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543–551, Jul. 1983.

[104] E. Ip and J. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, Sep. 2007.

[105] S. Hoffmann, R. Peveling, T. Pfau, O.Adamczyk, R. Eickhoff, and R.Noe, "Multiplier-free real-time phase tracking for coherent QPSK receivers," *IEEE Photon. Technol. Lett.*, vol. 21, no. 3, pp. 137–139, Feb. 2009.

[106] M. Seimetz, "Laser linewidth limitations for optical systems with high-order modulation employing feed forward digital carrier phase estimation," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2008, paper OTuM2.

[107] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photon. Technol. Lett.*, vol. 22, no. 9, pp. 631-633, May 2010.

[108] Y. Mori, C. Zhang, K. Igarashi, K. Katoh, and K. Kikuchi, "Unrepeated 200-km transmission of 40-Gbit/s 16-QAM signals using digital coherent receiver," *Opt. Exp.*, vol. 17, no. 3, pp. 1435-1441, Feb. 2009.

[109] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for *M*-QAM constellations," *J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[110] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, et al., "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 9, pp. 739-741, May 2012.

[111] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1966–1974, Jul. 2011.

[112] M. Pajovic, D. Millar, T. Koike-Akino, K. Kojima, V. Arlunno, and K. Parsons, "Multi-pilot aided carrier phase estimation for single carrier coherent systems," in *Proc. Signal Process. Photon. Commun.*, Jul. 2015, paper SpT4D.4.

[113] D. S. Millar, R. Maher, D. Lavery, T. Koike-Akino, M. Pajovic, A. Alvarado, *et al.*, "Design of a 1 Tb/s superchannel coherent receiver," *J. Lightw. Technol.*, vol. 34, no.6, pp. 1453-1463, Mar. 2016.

[114] X. Zhou, "An improved feed-forward carrier recovery algorithm for coherent receivers with *M*-QAM modulation format," *IEEE Photon. Technol. Lett.*, vol. 14, no. 14, pp. 1051-1053, Jul. 2010.

[115] Q. Zhuge, C. Chen, and D. V. Plant, "Low computation complexity two-stage feedforward carrier recovery algorithm for *M*-QAM," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2011, paper OMJ5.

[116] K. P. Zhong, J. H. Ke, Y. Gao, and J. C. Cartledge, "Linewidth-tolerant and low-complexity two-stage carrier phase estimation based on modified QPSK partitioning for dual-polarization 16-QAM systems," *J. Lightw. Technol.*, vol. 31, no.1, pp. 50-57, Jan. 2013.

[117] S. M. Bilal, C. R. Fludger, V. Curri, and G. Bosco, "Multistage carrier phase estimation algorithms for phase noise mitigation in 64-quadrature amplitude modulation optical systems," *J. Lightw. Technol.*, vol. 32, no.17, pp. 2973-2980, Sep. 2014.

[118] Y. Gao, A. P. T. Lau, S. Yan, and C. Lu, "Low-complexity and phase noise tolerant carrier phase estimation for dual-polarization 16-QAM systems," *Opt. Exp.*, vol. 19, no. 22, pp. 21717-21729, Oct. 2011.

[119] X. Zhou and Y. Sun, "Low-complexity, blind phase recovery for coherent receivers using QAM modulation," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2011, paper OMJ3.

[120] T. Pfau and R. Noé, "Phase-noise-tolerant two-stage carrier recovery concept for higher order QAM formats," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no.5, pp. 1210-1216, Sep./Oct. 2010.

[121] J. Li, L. Li, Z. Tao, T. Hoshida, and J. C. Rasmussen, "Laser-linewidth-tolerant feed-forward carrier phase estimator with reduced complexity for QAM," *J. Lightw. Technol.*, vol. 29, no. 16, pp. 2358-2364, Aug. 2011.

[122] K.-P. Ho and J. M. Kahn, "Detection technique to mitigate Kerr effect phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.

[123] A. P. T. Lau and J. M. Kahn, "Signal design and detection in presence of nonlinear phase noise," *J. Lightw. Technol.*, vol. 25, no. 10, pp. 3008-3016, Oct. 2007.

[124] D. Wang, M. Zhang, Z. Li, Y. Cui, J. Liu, Y. Yang, H. Wang, "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," in *Proc. Eur. Conf. Exhibit. Opt. Commun.,* Sep./Oct. 2015, paper p.3.16.

Md. Saifuddin Faruk received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Khulna University of Engineering & Technology, Khulna, Bangladesh, in 2003, and Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2006, respectively, and the Ph.D. degree in electrical engineering and information systems from the University of Tokyo, Tokyo, Japan, in 2011.

He joined as a Lecturer in Department of Electrical and Electronic Engineering, Dhaka University of Engineering and Technology, Gazipur, Bangladesh, in 2004 and promoted to Assistant Professor, Associate Professor, and Professor in 2006, 2013, and 2015, respectively. He joined in optical network groups of University College London as Marie Curie Research Fellow in 2015, and transferred to engineering department with the same position in 2016 at the University of Cambridge, U.K., where he is currently being employed. His research interests include coherent optical communications and digital signal processing.

Dr. Faruk received outstanding achievement award for Ph.D. thesis from the Department of Electrical Engineering and Information Systems, University of Tokyo in 2012 and best student paper award from Opto-Electronic and Communication Conference (OECC) in 2011.

Seb J. Savory (M'07−SM'11−F'17) received the M.Eng., M.A., and Ph.D. degrees in engineering from the University of Cambridge, U.K., in 1996, 1999, and 2001, respectively, and the M.Sc. (Maths) degree in mathematics from the Open University, Milton Keynes, U.K., in 2007.

His interest in optical communications began in 1991, when he joined STL (subsequently Nortel) in Harlow, U.K. Having been sponsored by Nortel through his undergraduate and postgraduate studies, he rejoined the Harlow Laboratories in 2000. In 2005, he moved to UCL where he held a Leverhulme Trust Early Career Fellowship from 2005 to 2007, before being appointed as a Lecturer (2007), Reader (2012) and Professor (2015). During 2014/15 he held a RAEng / Leverhulme Trust Senior Research Fellowship and was a joint

recipient of the RAEng Colin Campbell Mitchell Award in 2015. In October 2015, he was elected as a Fellow of Churchill College, Cambridge and in January 2016 moved to Cambridge as a University Lecturer.

Dr. Savory is the Editor-in-Chief of IEEE Photonics Technology Letters and serves on the Steering Committee of the Optical Fiber Communication conference having previously served as a General Chair (2015) and Program Chair (2013). He is a Chartered Engineer and Fellow of the Institution of Engineering and Technology, U.K. Recently he was elevated to Fellow of the OSA and the IEEE for contributions to digital coherent transceivers for optical fiber communication.