On Decision Aided Carrier Phase and Frequency Offset Estimation in Coherent Optical Receivers

Adaickalavan Meiyappan, *Student Member, IEEE*, Pooi-Yuen Kam, *Fellow, IEEE*, and Hoon Kim, *Senior Member, IEEE*

***Abstract*—We investigate carrier estimation (CE) for coherent optical receivers where the received signal is impaired by additive white Gaussian noise, laser phase noise, and frequency offset. Best practical 4-, 8-, and 16-point constellations are identified. A generalized differential encoding rule for signal constellations is presented. Performance of our complex-weighted decision-aided maximum-likelihood (CW-DA-ML) phase noise and frequency offset estimator is analyzed at low signal-to-noise ratio (SNR) and the optimal filter lengths are found. CW-DA-ML CE is put in perspective with respect to two fundamental estimators in the literature: (i) differential frequency estimator followed by block Mth power phase estimator (DiffFE-Mth CE), and (ii) fast Fourier transform based frequency estimator followed by block Mth power phase estimator (FFTbE-Mth CE), in terms of laser linewidth tolerance, frequency estimation range and speed, SNR threshold, and cycle slip probability. CW-DA-ML CE is 2.5 and 10.5 times faster than DiffFE-Mth CE in 4 phase-shift keying and 16 quadrature amplitude modulation signals, respectively, at a 1-dB system penalty for a bit-error rate of 10^{-3}. Our CE has lower cycle slip probability and transmission overhead than DiffFE-Mth and FFTbE-Mth CE. Hence, our CE is shown to be favourable in pilot-assisted (PA) systems. A PA CW-DA-ML CE is introduced and shown to be robust against time-varying frequency offset with minimal training overhead. Analog-to-digital convertor quantization error on our CE performance is also addressed.**

***Index Terms*—Block Mth power, cycle slip, differential encoding, fast Fourier transform, frequency offset, laser phase noise.**

I. Introduction

TODAY, 100-Gb/s Ethernet has become a commercial reality and research goals are geared towards next possible interface rates of 400 Gb/s and 1 Tb/s [1]. Given a constrained optical amplification bandwidth, maximizing spectral efficiency (SE) is of paramount importance to service the sustained growth of data traffic. Binary modulation, such as on-off keying and differential phase-shift keying (PSK), only achieve an SE of 0.8 bit/s/Hz per polarization [2]. Moving to nonbinary modulation, the asymptotic SE for an unconstrained intensity-modulated direct-detection (IMDD) system is $0.5\log_2(\gamma_s) - 0.5$ bits/s/Hz [3]. Here, γ_s is the signal-to-noise ratio (SNR) per symbol. However, the asymptotic SE for a constant-intensity-modulated coherent-detection system can reach $0.5\log_2(\gamma_s) + 1.10$ bits/s/Hz [4] . Although both IMDD and constant-intensity modulation has only one degree of freedom per polarization for encoding, the coherent system outperforms the noncoherent IMDD system by an SE of 1.6 bits/s/Hz at large SNR [5]. It has been suggested that 4-PSK is the most attractive modulation to achieve SE between 1 and 2 bits/s/Hz, whereas 8 or 16 quadrature amplitude modulation (QAM) are necessary for SE beyond 2 bits/s/Hz per polarization [6]. All these modulation formats demand coherent detection. Coherent detection is promising as it yields superior SE compared to noncoherent systems [7] and enables attainment of Shannon's capacity with the use of coding, such as turbo codes [8].

A major impediment in homodyne coherent detection is the synchronization of the local oscillator (LO) laser to the carrier of the received optical signal. The received signal can be perturbed by phase noise arising from nonzero laser linewidth, $\Delta\nu$, and frequency offset, Δf, between the transmitter and LO lasers. Laser linewidth can range from the order of 10 kHz for external-cavity tunable lasers [9] to 10 MHz for distributed feedback (DFB) lasers. The frequency misalignment can be as large as ± 2.5 GHz over the lifetime of a typical tunable laser [10]. Traditionally, phase-locked loops (PLL) have been employed for coherent demodulation of optical signals. However, the PLL is sensitive to loop propagation delay which can cause loop instability. Loop delay greater than the bit duration becomes nonnegligible and severely constraints the permissible laser linewidth [11]. Moreover, the PLL has a limited frequency offset estimation range [12]. Experimental results using a PLL in 16-QAM show constrained values of frequency offset per bit rate, $\Delta f T_b = 1.79 \times 10^{-3}$ at linewidth per bit rate, $\Delta\nu T_b = 3.57 \times 10^{-6}$ [13] and $\Delta f T_b = 6.25 \times 10^{-4}$ at $\Delta\nu T_b = 2.5 \times 10^{-6}$ [14] for reliable carrier estimation (CE).

Current interest lies in intradyne coherent detection using a free running LO laser, followed by sampling with high-speed analog-to-digital converter (ADC), and execution of CE in digital signal processors (DSP). Even when PLL may fail due to delay constraints, DSP based CE methods allow to use cost-effective DFB lasers by relaxing the laser linewidth and frequency offset tolerance. CE available in the literature comprise a two stage sequential process of frequency offset compensation followed by phase noise compensation, as phase estimators are only unbiased in the absence of frequency offset [15]. CE algorithms can be categorized into decision-aided (DA) algorithms which use symbol decisions from a data detector and non-de-

Manuscript received March 06, 2013; revised April 15, 2013; accepted April 25, 2013. Date of publication April 30, 2013; date of current version May 24, 2013. This work was supported by the Singapore MoE AcRF Tier 2 Grant MOE2010-T2-1-101.

The authors are with the Department of Electrical and Computer Engineering, National University of Singapore, 117576 Singapore (e-mail: adaickalavan@hotmail.com).

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

Digital Object Identifier 10.1109/JLT.2013.2260723

cision-aided (NDA) algorithms which do not use any symbol decisions.

The two prevalent fundamental NDA frequency estimators in the literature are the differential frequency estimator (DiffFE) which is a sample autocorrelation method and the fast Fourier transform based estimator (FFTbE) which is a periodogram method. DiffFE for MPSK presented in [16] estimates the frequency by computing the phase increment between two adjacent samples raised to the Mth power for modulation removal. The complexity of raising to the Mth power increases with modulation order. DiffFE was adapted for 16-QAM by only phase differencing consecutive Class I symbols with modulation phases $\pi/4 + i\pi/2$, $i = 0, 1, 2, 3$, and raising to the 4th power for modulation removal [17].

FFTbE [18], formed by exploiting the cyclostationary statistics of the received signal, involves a computationally intensive peak search in a periodogram. The frequency estimate, $\Delta\widehat{f}$, for MPSK is given by

$$\Delta\widehat{f} = \frac{1}{M} \underset{|\Delta\bar{f}|<1/2T}{\arg\max} \left| \sum_{k=0}^{N-1} r^M(k)\, e^{-j2\pi\Delta\bar{f}Tk} \right| \qquad (1)$$

where M is the modulation order, T is the symbol duration, N is the received sample size over which the frequency acquisition is performed, and $|\cdot|$ is the modulus operator. Here, the function $\arg\max_i x(i)$ selects the i that maximizes $x(i)$. For square 16-QAM, $\Delta\widehat{f}$ is obtained by setting $M = 4$ in (1). The frequency estimate resolution is limited by N to $1/MNT$. A larger N improves the frequency estimate accuracy, but increases the acquisition time and FFT complexity. These conflicting requirements call for a tradeoff but no automatic optimization method is known. FFTbE suffers from an undesirable SNR threshold effect, where a noise peak exceeding the true frequency peak causes a large frequency estimation error (known as an outlier) below some critical SNR value [19]. The likelihood of outliers increases with decreasing SNR, having a disabling effect on FFTbE at low SNR. DiffFE and FFTbE have limited modulation-format-dependent $\Delta f T_b$ estimation ranges of $\pm 1/2M\log_2 M$ for MPSK and $\pm 1/8\log_2 M$ for 16-QAM.

A prevalent fundamental NDA phase estimator is the block Mth power, presented for MPSK in [20], where a single phase estimate $\widehat{\theta}$ is used to phase-correct all the symbols in a block. Block Mth power was adapted for 16-QAM through a QPSK partitioning technique in [21]. In block Mth power, nonlinear operations of raising to the Mth power and taking $\arctan(\cdot)$ increases the system latency. Moreover, the phase estimate $\widehat{\theta}$ requires phase unwrapping. However, accurate phase unwrapping is difficult at low SNR [22] or at large laser linewidth, and may induce cycle slips which are a highly nonlinear phenomenon [23]. Differential encoding (DE) is generally applied to arrest cycle slips, but it is not possible to have a universal DE rule for all the signal constellations [24].

A DA phase estimator of interest is the decision-aided maximum-likelihood (DA-ML) estimator [25]. The interest in DA-ML estimator lies in its ML derivation and its near optimum phase estimation performance at high and medium SNR [26]. The DA-ML estimator avoids additive noise contribution from higher powers ($\geq$2) which are present in block Mth power estimator [25].

Recently, by modifying the filter coefficients of the DA-ML estimator, we formed a complex-weighted DA-ML (CW-DA-ML) CE for joint phase noise and frequency offset estimation [27]. CW-DA-ML CE is a DA least-squares based estimator, which is modulation format independent, achieves complete frequency estimation range, and requires no phase unwrapping. The main objective of this paper is to perform a comprehensive analysis of CW-DA-ML CE performance in a channel impaired by additive white Gaussian noise (AWGN), linear phase noise, and frequency offset. Prospective 4-, 8-, and 16-point constellations are examined, and the most favourable constellations are identified. The performance of CW-DA-ML CE is placed in perspective with respect to two fundamental NDA CE in the literature, namely, (i) DiffFE [16], [17], followed by block Mth power estimator [20], [21] (referred to as DiffFE-Mth), and (ii) FFTbE [18] followed by block Mth power estimator [20], [21] (referred to as FFTbE-Mth). The advent of advanced forward error correction (FEC) codes with a 7% overhead sets a low bit-error rate (BER) threshold of 3.8×10^{-3}, enabling the operation of modern systems at low SNR [28]. Furthermore, it is desirable to operate at low signal power to minimize nonlinear phase noise [29] and to save power. The CW-DA-ML CE is known to perform well at high SNR but remains untested at the low SNR [27]. Therefore, all analysis in this paper is performed at low SNR values corresponding to a BER of 10^{-3}, which is the target range of modern systems.

In Section II, we examine the merits of several 4-, 8-, and 16-point constellations in terms of their AWGN resilience, phase rotation tolerance, and transmitter implementation complexity. A generalized DE rule applicable to all signal constellations having greater than or equal to 2 rotationally symmetric positions with no dc signal point is presented. In Section III, we briefly derive CW-DA-ML CE and examine its optimum filter length. In Section IV, we compare the features of CW-DA-ML CE with DiffFE-Mth and FFTbE-Mth CE in terms of: (a) laser linewidth tolerance, (b) frequency offset estimation range, (c) acquisition time, accuracy, and SNR threshold, and (d) cycle slip probability. In Section V, pilot-assisted (PA) CW-DA-ML CE is introduced and performance gain over its DE counterpart is discussed. In Section VI, the robustness of CW-DA-ML CE against a time-varying frequency offset is studied. In Section VII, the necessary ADC resolution is considered. In Section VIII, a complexity analysis of CE algorithms is performed. Section IX concludes the paper. Throughout this paper, $\mathrm{E}[\cdot]$, $\lfloor x \rfloor$, superscript *, and superscript T denotes expectation, smallest integer larger than x, complex conjugate, and transpose, respectively.

II. Modulation Formats

A. Several 4-, 8-, and 16-Point Constellations

In Fig. 1, we consider several prospective 4-, 8-, and 16-ary discrete-point constellations which use both field quadratures. We compare the constellations in terms of: (i) minimum

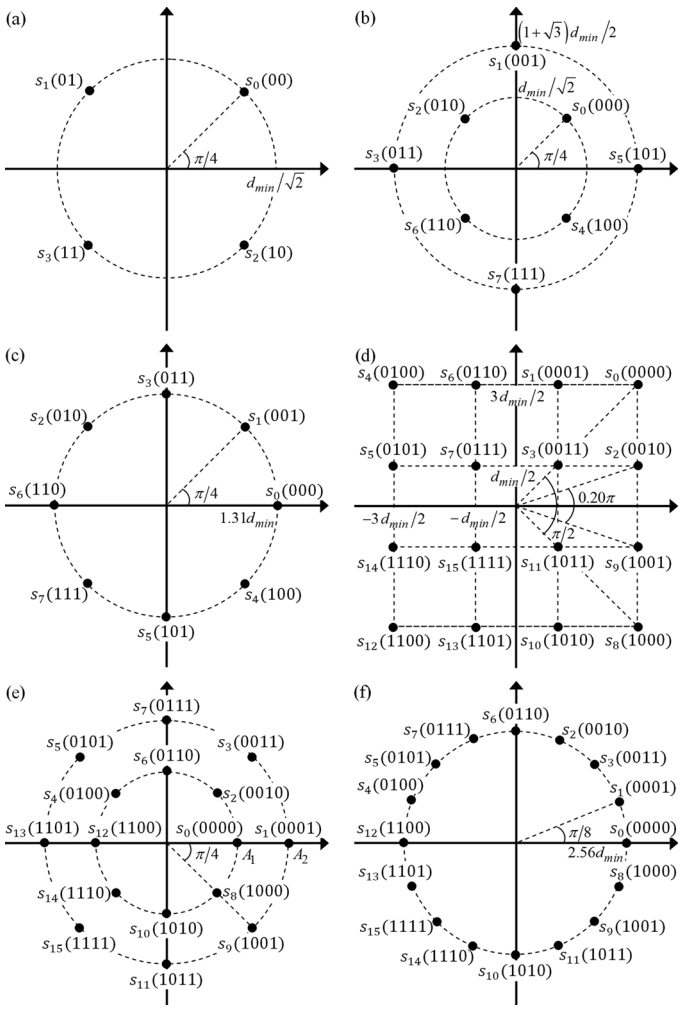

Fig. 1. Signal constellation and bits-to-symbol mapping for (a) 4-PSK, (b) 8-QAM, (c) 8-PSK, (d) 16-QAM, (e) 16-Star, and (f) 16-PSK.

Euclidean distance between adjacent points, $d_{\min}$, characterizing its resilience against AWGN, (ii) minimum angular separation between adjacent points with the same radius, $\phi_{\min}$, characterizing its phase rotation resilience against phase noise and frequency offset, and (iii) simplicity in transmitter implementation. The $d_{\min}$ is normalized to a constellation with a unity average symbol energy. Larger values of $d_{\min}$ and $\phi_{\min}$ imply greater AWGN and phase rotation resilience, respectively.

In 4-point constellations, we only consider 4-PSK as it is well established to achieve the best performance for AWGN channel with the largest $d_{\min} = \sqrt{2}$ among all 4-point constellations [30]. The 4-PSK signal is also easy to generate.

The 8-QAM is defined to be the constellation shown in Fig. 1(b). In an AWGN channel, 8-QAM ($d_{\min} = 0.919$) outperforms 8-PSK ($d_{\min} = 0.765$) but is marginally inferior to the optimum 8-point constellation, 8-Hex ($d_{\min} = 0.963$), by 0.35 dB [30], [31]. However, 8-QAM ($\phi_{\min} = \pi/2$) has better phase rotation tolerance than 8-PSK ($\phi_{\min} = \pi/4$) and 8-Hex ($\phi_{\min} < \pi/3$). Unlike 8-Hex, 8-QAM has a simple transmitter configuration realizable with Mach-Zehnder modulators and couplers [32], and has a simple DE technique as will be shown below. Hence, we only consider 8-QAM for its desirable properties as outlined above and 8-PSK for further analysis in this paper.

The optimum ring ratio, $RR = A_2/A_1$, for 16-Star in an AWGN channel maximizing the distance between adjacent points in the inner ring and that between the two rings is obtained when $RR = 1 + 2\cos(0.375\pi) \approx 1.77$ [33]. We have used $RR = 1.77$ in this paper, as the optimum RR with respect to phase rotations only deviate slightly from 1.77 [34]. In an AWGN channel, 16-QAM ($d_{\min} = 0.632$) outperforms 16-Star ($d_{\min} = 0.534$) and 16-PSK ($d_{\min} = 0.390$), but is second by 0.5 dB to the optimum 16-point hexagonal-like constellation [31]. 16-QAM is preferred, compared to the optimum 16-point constellation, due to its simple transmitter implementation where integrated 16-QAM modulators are already being developed [35] and simple DE technique as will be shown below. However, in terms of phase rotation tolerance, 16-Star ($\phi_{\min} = \pi/4$) outperforms 16-QAM ($\phi_{\min} = 0.20\pi$ & $\pi/2$) and 16-PSK ($\phi_{\min} = \pi/8$). Hence, we only consider 16-QAM for its desirable properties as outlined above, 16-Star for its phase rotation tolerance, and 16-PSK for further analysis in this paper.

B. BER of 4-, 8-, and 16-Point Constellations

The ML detector in an AWGN limited and phase rotation limited channel has a Euclidean metric with straight-line decision boundaries and a non-Euclidean metric with circular-line boundaries forming polar wedges, respectively [32]. Considering the implementation difficulty of a non-Euclidean metric with circular-line boundaries requiring lookup tables, advances in laser linewidth making DFB lasers with 10 kHz linewidth available [36], and low SNR operating region of modern systems where AWGN is dominant, we use the Euclidean metric with straight-line decision boundaries in this paper. The BER over an AWGN channel without DE for MPSK given by [37]

$$BER = \frac{2}{\log_2 M} \cdot Q\left[\sqrt{2\gamma_b \log_2 M}\sin\left(\frac{\pi}{M}\right)\right], \qquad (2)$$

8-QAM given by [2]

$$BER = \frac{22}{16} \cdot Q\left[\sqrt{\frac{6\gamma_b}{(3+\sqrt{3})}}\right], \qquad (3)$$

16-QAM given by [36]

$$BER = 1 - \left(1 - \left(\frac{2}{\log_2 M}\right)\left(1 - \frac{1}{\sqrt{M}}\right) \cdot Q\left[\sqrt{\frac{3\log_2 M}{M-1} \cdot \gamma_b}\right]\right)^2, \qquad (4)$$

and 16-Star obtained through Monte Carlo (MC) simulation, are shown in Fig. 2. Here, γ_b is the SNR per bit. Theoretical values without DE at BER $= 10^{-3}$ are given in the second column of Table I.

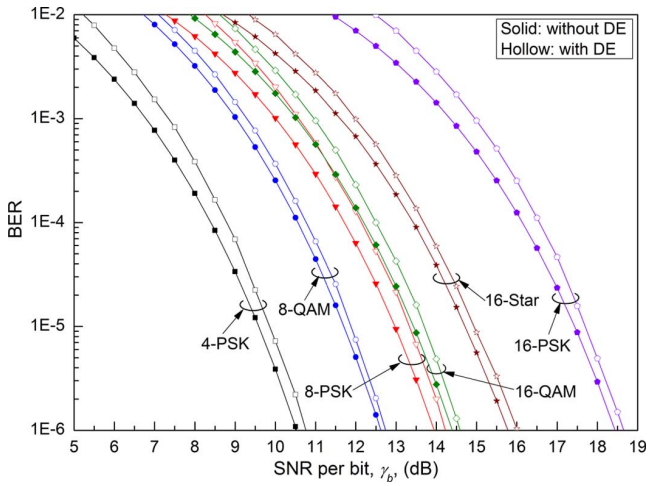

Fig. 2. BER versus SNR per bit in an AWGN channel with and without DE.

TABLE I
SNR PER BIT, γ_b, VALUES AT BER $= 10^{-3}$

Format	AWGN channel without DE (dB)	AWGN channel with DE (dB)	DE penalty (dB)
4-PSK	6.82	7.38 (MC)	0.56
8-QAM	9.04	9.33 (MC)	0.29
8-PSK	10.01	10.60 (MC)	0.59
16-QAM	10.53	10.97 (MC)	0.44
16-Star	11.64 (MC)	11.99 (MC)	0.35
16-PSK	14.37	14.97 (MC)	0.60

C. Differential Encoding Technique

We present a generalized sector based DE technique, following the idea in [24], applicable to all constellations having greater than or equal to 2 rotationally symmetric positions and no dc signal point (i.e., no signal point at the origin). In a q-sector rotationally symmetric MPSK and MQAM constellations, any constellation point can be obtained by rotating a corresponding constellation point from the first rotationally symmetric sector. Hence, the kth information signal point $s(k)$ can be represented by $s(k) = \rho(k)\, d(k)$. Here $\rho(k) = \exp(j2\pi i/q)$, $i \in \{0, \ldots, q-1\}$, is the appropriate sector-rotation term and $d(k)$ is the corresponding constellation point of $s(k)$ in the first rotationally symmetric sector. The kth DE transmitted symbol $m(k)$ is then obtained as $m(k) = \bar{\rho}(k)\, d(k)$, where $\bar{\rho}(k) = \rho(k)\, \bar{\rho}(k-1)$ simply represents the current sector in which $m(k)$ lies. Differential decoding of the kth symbol $m(k)$ at the receiver proceeds as $s(k) = \bar{\rho}(k)\, d(k) / \bar{\rho}(k-1)$. Initial sector $\bar{\rho}(-1) = 1$.

DE increases the BER as any symbol detection error manifests itself twice through DE and is depicted in Fig. 2. The DE induced penalty at BER $= 10^{-3}$ in an AWGN channel is summarized in column four of Table I. To minimize bit errors due to symbol errors, careful bits-to-symbol mapping is needed. For constellations where $q = 2^{\bar{q}}$ for some integer $\bar{q}$, we adopt the following bits-to-symbol mapping scheme. The first $\bar{q}$ bits of all symbols within a sector share the same bits to minimize bit errors due to adjacent symbol errors caused by AWGN. The first $\bar{q}$ bits of symbols are DE by sector. The last $\log_2 M - \bar{q}$ bits in each symbol are encoded to be rotationally invariant, thus making them immune to cycle slips. The bits-to-symbol mapping for DE signals are shown in Fig. 1.

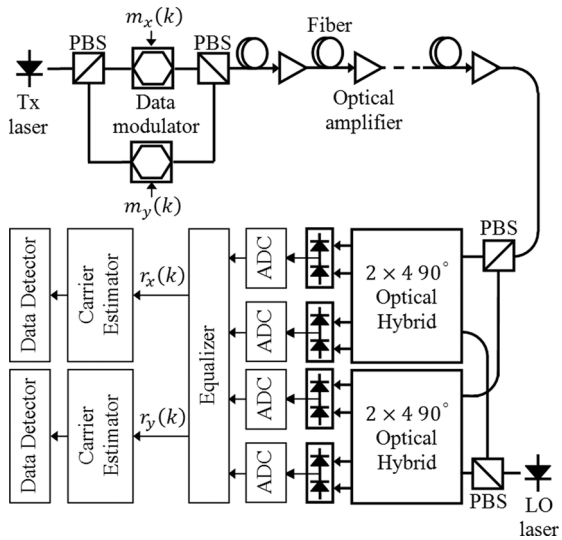

Fig. 3. Polarization multiplexed coherent optical system. PBS: polarization beam splitter, Tx: transmitter.

III. CARRIER ESTIMATION

A. System Model

We consider CE for a coherent optical transmission system shown in Fig. 3. The symbol sequence $\{m(k)\}$ is modulated onto the optical carrier by data modulators and transmitted through an optical fiber with periodic optical amplifiers. Optical amplifiers introduce amplified spontaneous emission noise which can be modeled as AWGN. At the receiver, the received optical signal is mixed with a LO laser, downconverted to an intermediate frequency (IF) through 2×4 90° optical hybrids, detected by balanced photodetectors, and sampled by ADCs. The IF is equivalent to the frequency offset between the transmitter and the LO laser. The received signal is then processed by an equalizer for chromatic dispersion compensation, polarization demultiplexing, and polarization mode dispersion compensation. Finally, the phase noise introduced by nonzero laser linewidth and frequency offset is estimated by CE, and a symbol decision is made by a symbol-by-symbol data detector. We assume the received signal samples $r_x(k)$ and $r_y(k)$ over the kth symbol interval $[kT, (k+1)T)$ have known symbol timing with one complex sample per symbol, and is free of intersymbol interference, nonlinear distortions, and polarization crosstalk. CE is performed independently for each channel and all equations hereafter refer to that of one polarization channel.

A canonical model of the received signal can be written as

$$r(k) = m(k)\, e^{j(\Delta\omega k + \theta(k))} + n(k) \tag{5}$$

where $m(k)$ assumes equiprobable symbols from the signal set $\{s_i\}_0^{M-1}$. Here, $\Delta\omega = 2\pi\Delta f T$ is the angular frequency offset, assumed static throughout the paper unless otherwise stated. The $\theta(k) = \eta(k) + \theta(k-1)$ is the laser phase noise modeled as a Wiener process, where $\{\eta(k)\}$ is a sequence of independent and identically distributed Gaussian random variables with mean zero and variance $\sigma_p^2 = 2\pi\Delta\nu T$ [38]. Here, $\Delta\nu$ is the combined linewidth of the transmitter and LO lasers. The set $\{n(k)\}$ is a sequence of circularly symmetric AWGN with zero mean and variance N_0. SNR per bit is defined as

$\gamma_b = E\left[|m(k)|^2\right]/N_0 \log_2 M$. All γ_b penalty values are referenced to their theoretical γ_b value without DE at BER $= 10^{-3}$, summarized in the second column of Table I. DE is employed throughout the paper to arrest cycle slips, unless otherwise stated.

B. CW-DA-ML Carrier Estimator

The laser phase noise process $\{\theta(k)\}$ is assumed to vary slower than the symbol rate such that we can approximate $\theta(k)$ to be time-invariant over an interval longer than LT, where L is an integer. In CW-DA-ML CE, we form a reference phasor (RP) $V(k+1)$ for the carrier at time $k+1$ by filtering the immediate past L samples, as [27]

$$V(k+1) = C(k) \sum_{l=1}^{L} w_l(k)\, r(k-l+1)\, \widehat{m}^*(k-l+1) \tag{6}$$

where each $w_l(k)$ is a complex weight, and $\widehat{m}(k)$ is the data detector's symbol decision on $r(k)$. The data detector declares the signal s_i from the signal set $\{s_i\}_0^{M-1}$ which maximizes $\mathrm{Re}\left[r(k)V^*(k)s_i^* - (1/2)|s_i|^2\right]$ as the symbol decision $\widehat{m}(k)$. In other words,

$$\widehat{m}(k) = \underset{0\le i\le M-1}{\arg\max}\ \mathrm{Re}\left[r(k)V^*(k)s_i^* - \frac{1}{2}|s_i|^2\right]. \tag{7}$$

In (6), $C^{-1}(k) = \sum_{l=1}^{L} |\widehat{m}(k-l+1)|^2$ normalizes the magnitude of $V(k+1)$, making CW-DA-ML CE applicable to both MPSK and MQAM formats.

Momentarily let $\{\theta(k)\} = 0$, $\{n(k)\} = 0$, and assume ideal decision feedback such that $\widehat{m}(k) = m(k)$. In the presence of frequency offset, consecutive symbols differ by a phase rotation of $\Delta\omega$ and the filter-input vector at time k appears as

$$\begin{aligned}\boldsymbol{y}(k) &= [r(k)\widehat{m}^*(k), r(k-1)\widehat{m}^*(k-1),\\ &\quad \ldots, r(k-L+1)\widehat{m}^*(k-L+1)]^T\\ &= [|m(k)|^2 e^{j\Delta\omega k}, |m(k-1)|^2 e^{j\Delta\omega(k-1)},\\ &\quad \ldots, |m(k-L+1)|^2 e^{j\Delta\omega(k-L+1)}]^T.\end{aligned} \tag{8}$$

The filter-weight vector at time k, $\boldsymbol{w}(k) = [w_1(k), w_2(k), \ldots, w_L(k)]^T$, is designed to rotate each filter-input term to have the same angular frequency offset of $\Delta\omega(k+1)$, matching that of the received sample $r(k+1)$ that we are trying to coherently demodulate. From (8), it is intuitively clear that $\arg(\boldsymbol{w}(k))$ should equal $[\Delta\omega, 2\Delta\omega, \ldots, L\Delta\omega]^T$, but $\Delta\omega$ is unknown in practice. Hence, we adaptively pick the $\boldsymbol{w}(k)$ at each time k to minimize the sum-of-error-squares cost function $J(k)$,

$$J(k) = \sum_{l=1}^{k} \left| \frac{r(l)}{\widehat{m}(l)} - C(l-1)\boldsymbol{w}^T(k)\boldsymbol{y}(l-1) \right|^2. \tag{9}$$

Minimizing $J(k)$ with respect to $\boldsymbol{w}(k)$, we obtain a least-squares solution $\boldsymbol{w}(k) = \boldsymbol{\Phi}^{-1}(k)\boldsymbol{z}(k)$, $k \ge 1$, where

$$\boldsymbol{\Phi}(k) = \boldsymbol{\Phi}(k-1) + C^2(k-1)\boldsymbol{y}^*(k-1)\boldsymbol{y}^T(k-1) \tag{10}$$

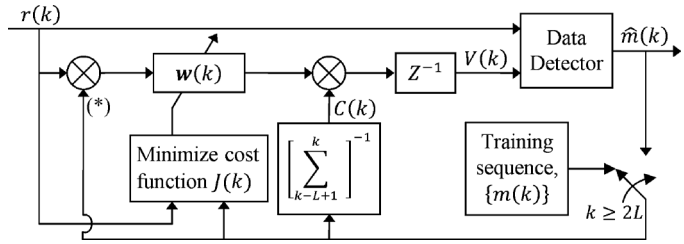

Fig. 4. Structure of CW-DA-ML CE.

is the time-average L-by-L autocorrelation matrix and

$$\boldsymbol{z}(k) = \boldsymbol{z}(k-1) + \frac{C(k-1)\boldsymbol{y}^*(k-1)r(k)}{\widehat{m}(k)} \tag{11}$$

is the time-average L-by-1 cross-correlation vector. Matrix $\boldsymbol{\Phi}(k)$ and vector $\boldsymbol{z}(k)$ can be formed recursively as shown in (10) and (11), respectively, minimizing the memory resources required. We initialize $V(0) = 1$, $\boldsymbol{w}(0) = [1, 0, \ldots, 0]^T$, and $\boldsymbol{\Phi}(0) = 0.01\boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix. The weight-vector $\boldsymbol{w}(k)$ can adapt to follow time-varying channels as it depends on the observed signal samples $\{r(k)\}$. The structure of CW-DA-ML CE is shown in Fig. 4.

CW-DA-ML CE uses a training sequence for the initial symbol decisions and subsequently uses the output of the data detector. A training sequence length of $2L$ has been shown to be sufficient [27], thus the overhead is kept low. It is crucial to note that unlike block Mth power based CE, CW-DA-ML CE does not require phase unwrapping as the RP has an unambiguous phase tracking range of $[0, 2\pi)$.

We analyze the steady-state filter weights $\{w_l\}_1^L$ in CW-DA-ML CE with filter length $L = 12$ for a 16-QAM signal. Each steady-state value is obtained by averaging over 500 runs the average of its value from $k = 50 \times 10^3$ to $k = 51 \times 10^3$ in each run. First, the magnitude, $|w_l|$, and phase, $\arg(w_l)$, of the weights are plotted in Fig. 5(a) and (b), respectively, as $\Delta\nu T_b$ and γ_b are varied while $\Delta f T_b$ is kept constant at 2×10^{-2}. As $\Delta\nu T_b$ and γ_b increases, the decay rate of $|w_l|$ with l increases and thus the recent received samples $\{r(l)\}$ are increasingly prioritised compared to those further back in time [see Fig. 5(a)]. This amounts to shortening of the filter's effective averaging length. Laser phase noise $\theta(k)$ becomes less related with $\theta(k-l)$ as l increases. Hence, samples $\{r(l)\}$ further back in time carry less useful information on the laser phase noise in sample $r(k+1)$ and thus get weighted down. Additionally, increasing γ_b reduces the interval over which additive noise smoothing needs to be performed. However, regardless of the variation in $\Delta\nu T_b$ and γ_b, the relation of $\arg(w_l)$ to l does not change [see Fig. 5(b)]. Second, the $|w_l|$ and $\arg(w_l)$ are plotted in Fig. 5(c) and (d), respectively, as $\Delta\nu T_b$ and $\Delta f T_b$ are varied while γ_b is kept constant at 11.53 dB. Regardless of the variation in $\Delta f T_b$, for a given $\Delta\nu T_b$, the relation of $|w_l|$ to l does not change [see Fig. 5(c)]. However, each $\arg(w_l)$ converges to $\Delta\omega l$ depending on the $\Delta f T_b$ value present, whereas the change in $\Delta\nu T_b$ has no discernible effect on $\arg(w_l)$ [see Fig. 5(d)]. We can conclude from Fig. 5 that the magnitude of the weight w_l responds to $\Delta\nu T_b$ and γ_b in weighing down less-relevant samples, whereas the phase of the weight w_l responds to $\Delta f T_b$ in correcting for the frequency

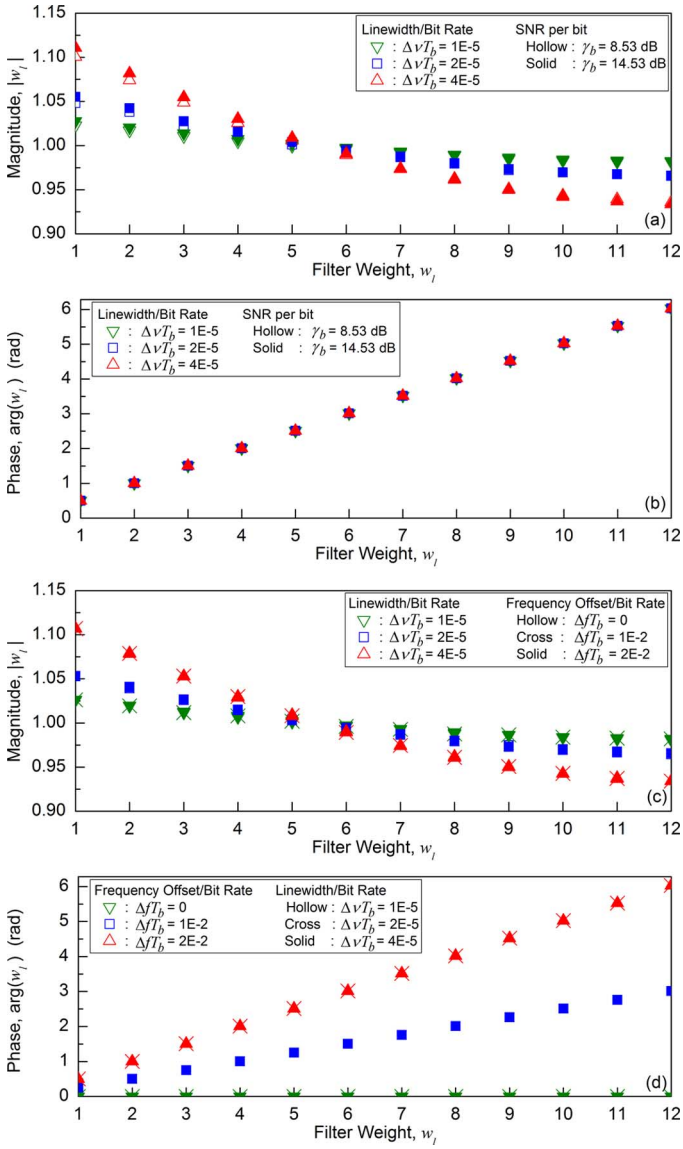

Fig. 5. Filter weights at fixed $\Delta f T_b = 2 \times 10^{-2}$: (a) $|w_l|$, and (b) $\arg(w_l)$. Filter weights at fixed $\gamma_b = 11.53$ dB: (c) $|w_l|$, and (d) $\arg(w_l)$. Example of legend types 'Hollow', 'Cross', and 'Solid' are ▽, ⊠, and ▼, respectively.

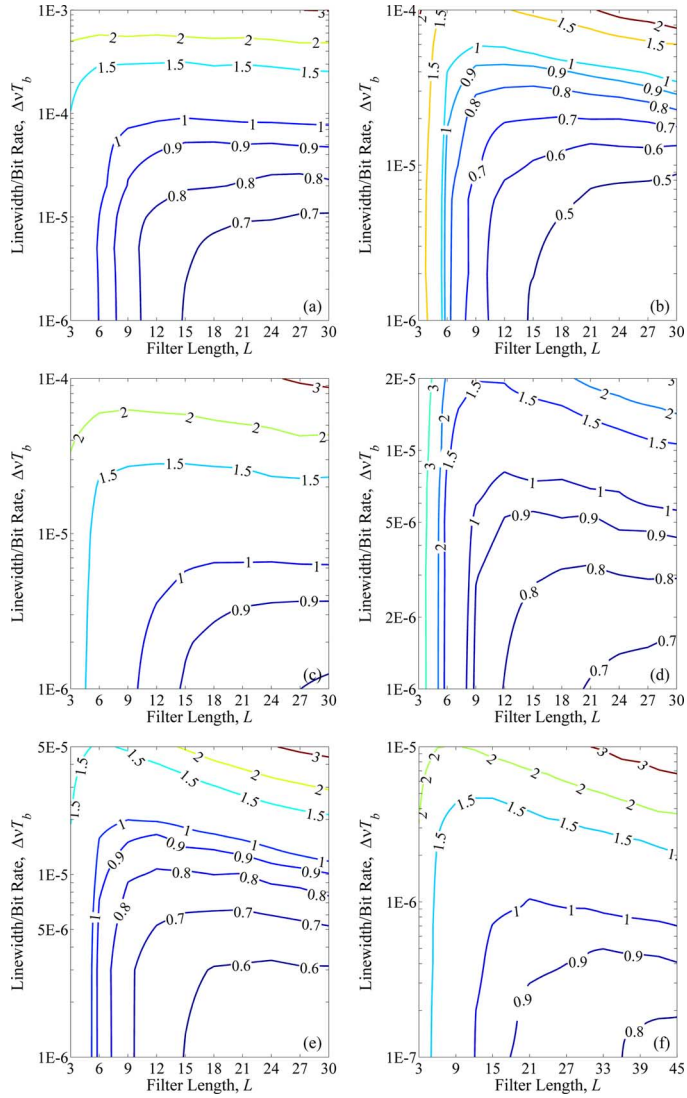

Fig. 6. SNR per bit penalty at BER $= 10^{-3}$ versus linewidth per bit rate and filter length in CW-DA-ML CE for (a) 4-PSK, (b) 8-QAM, (c) 8-PSK, (d) 16-QAM, (e) 16-Star, and (f) 16-PSK signals.

offset. Hence, given convergence of $\arg(w_l)$ to $\Delta\omega l$, the optimum filter length is only dependent upon $\Delta\nu T_b$ and γ_b.

In general, there is a trade-off between the need for long filter lengths to average over AWGN and the need for short filter lengths to ensure filter inputs are correlated. Contour plots of γ_b penalty at BER $= 10^{-3}$ on a $\Delta\nu T_b$ versus filter length map are drawn in Fig. 6 for CW-DA-ML CE. The contour plots confirm that the optimum filter length decreases with $\Delta\nu T_b$ and there is a minimum filter length even in the absence of phase noise. For comprehensiveness, we provide the contour plots of DiffFE-Mth CE in Fig. 7. Optimal filter lengths at a 1-dB γ_b penalty for CW-DA-ML and DiffFE-Mth CE are given in Table II. A sufficiently large sample size was used for frequency acquisition in DiffFE-Mth and CW-DA-ML CE to ensure convergence of their frequency offset estimates to the true value. Therefore, the optimal filter length of DiffFE-Mth CE will be equally applicable to FFTbE-Mth CE. Optimum filter lengths stated in Table II are used for all CE analysis in this paper.

IV. Performance Analysis

A. Linewidth/Bit Rate Tolerance

The SNR per bit penalty versus linewidth per bit rate for CW-DA-ML CE is plotted in Fig. 8. Without loss of generality, we set $\Delta f = 0$. The tolerable $\Delta\nu T_b$ values for a 1-dB γ_b penalty at BER $= 10^{-3}$ are summarized in Table III. Sorted in decreasing order of $\Delta\nu T_b$ tolerance, we have 4-PSK, 8-QAM, 16-Star, 16-QAM, 8-PSK, and 16-PSK. It is interesting to note that 16-QAM achieves similar $\Delta\nu T_b$ tolerance as 8-PSK and yet occupies $1/12T_b$ less spectral width. As argued extensively in [39], 16-Star ($\phi_{\min} = \pi/4$) has greater $\Delta\nu T_b$ tolerance than 16-QAM ($\phi_{\min} = 0.20\pi$ & $\pi/2$) by virtue of its larger $\phi_{\min}$. Since 16-Star ($d_{\min} = 0.534$) has poorer packing density than 16-QAM ($d_{\min} = 0.632$), it generally requires higher γ_b and thus only attractive for $\Delta\nu T_b \geq 2.75 \times 10^{-5}$ as depicted in Fig. 9. Similar conclusion holds for other 16-point ring constellations considered in [40] which offer moderate advantage in terms of $\Delta\nu T_b$ tolerance but at the expense of poor packing

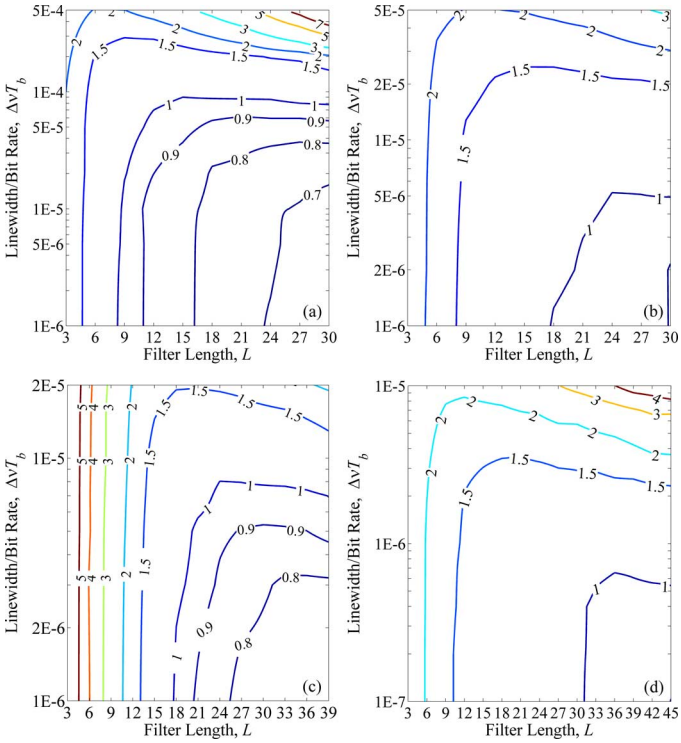

Fig. 7. SNR per bit penalty at BER $= 10^{-3}$ versus linewidth per bit rate and filter length in DiffFE-Mth CE for (a) 4-PSK, (b) 8-PSK, (c) 16-QAM, and (d) 16-PSK signals.

TABLE II
OPTIMAL FILTER LENGTH FOR 1-dB SNR PER BIT PENALTY AT BER $= 10^{-3}$

Format	DiffFE-Mth	CW-DA-ML
4-PSK	15	15
8-QAM	-	9
8-PSK	24	18
16-QAM	24	12
16-Star	-	9
16-PSK	36	21

density, increased transmitter complexity, and DE complexity, compared to 16-QAM.

From Fig. 8 and Table III, it is seen that CW-DA-ML CE equals or outperforms DA-ML estimator in terms of $\Delta\nu T_b$ tolerance even when $\Delta f = 0$. This is because DA-ML estimator weighs all filter inputs equally, i.e., $\boldsymbol{w}(k) = [1, 1, \ldots, 1]^T$, but CW-DA-ML CE weighs the samples in a decaying manner resulting in a better phase estimate. Although DiffFE-Mth CE has an inherent advantage of being a noncausal filter in its phase estimation, the entirely causal CW-DA-ML CE still equals DiffFE-Mth CE in 4-PSK and 16-QAM, whereas outperforms DiffFE-Mth CE in 8- and 16-PSK. As frequency estimation convergence was ensured in DiffFE-Mth CE, its $\Delta\nu T_b$ tolerance results equally apply to FFTbE-Mth CE.

B. Frequency Offset/Bit Rate Estimation Range

Frequency offset estimation range of block Mth power, DA-ML, DiffFE-Mth, FFTbE-Mth, and CW-DA-ML estimators are drawn in Fig. 10 while keeping $\Delta\nu = 0$. The theoretical maximum $\Delta f T_b$ tolerance of block Mth power in the absence of laser phase noise and AWGN is $\pm 1/2LM\log_2 M$. The

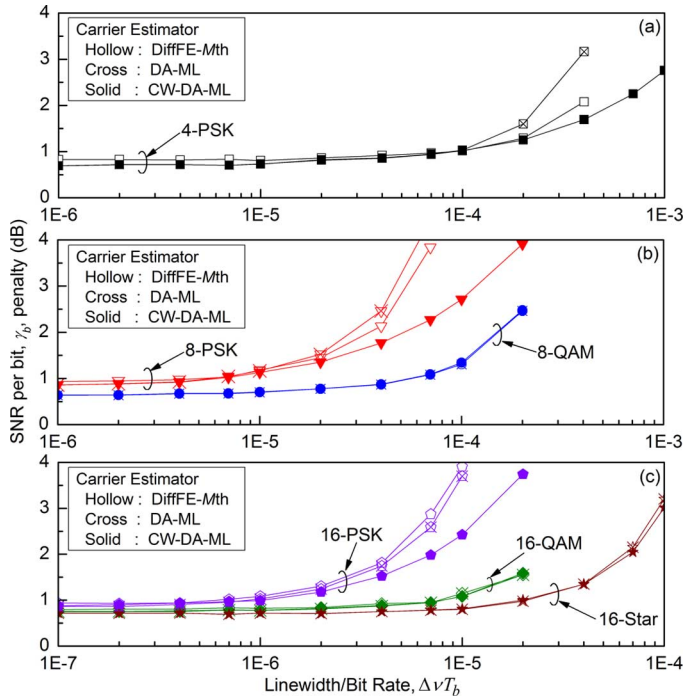

Fig. 8. SNR per bit penalty versus linewidth per bit rate for (a) 4-, (b) 8-, and (c) 16-point constellations. $\Delta f = 0$.

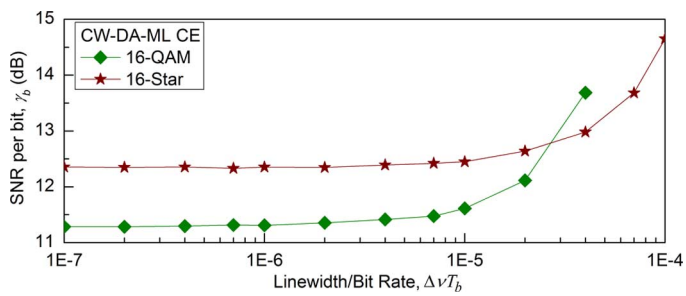

Fig. 9. SNR per bit versus linewidth per bit rate.

TABLE III
LINEWIDTH/BIT RATE TOLERANCE FOR 1-dB SNR PER BIT PENALTY AT BER $= 10^{-3}$

Format	DA-ML	DiffFE-Mth	CW-DA-ML
4-PSK	8.8×10^{-5}	9.0×10^{-5}	9.0×10^{-5}
8-QAM	5.8×10^{-5}	-	5.8×10^{-5}
8-PSK	5.8×10^{-6}	5.3×10^{-6}	6.5×10^{-6}
16-QAM	7.7×10^{-6}	8.0×10^{-6}	8.0×10^{-6}
16-Star	2.0×10^{-5}	-	2.0×10^{-5}
16-PSK	9.0×10^{-7}	6.5×10^{-7}	1.1×10^{-6}

limited frequency offset tolerance of block Mth power and DA-ML estimators at 1-dB γ_b penalty, summarized in Table IV, reiterates the need to incorporate a dedicated frequency offset estimation capability into the CE. Block Mth power and DA-ML estimators are only suitable for $\Delta f T_b \leq \pm 2.1 \times 10^{-3}$.

DiffFE-Mth and FFTbE-Mth CE raises the received signal to the Mth power in MPSK format to remove the information-bearing phase. This limits their $\Delta f T_b$ estimation range to a modulation-format-dependent $\pm 1/2M\log_2 M$. Likewise, their $\Delta f T_b$ estimation range in 16-QAM is limited to $\pm 1/8\log_2 M$ since the received signal is raised to the 4th power.

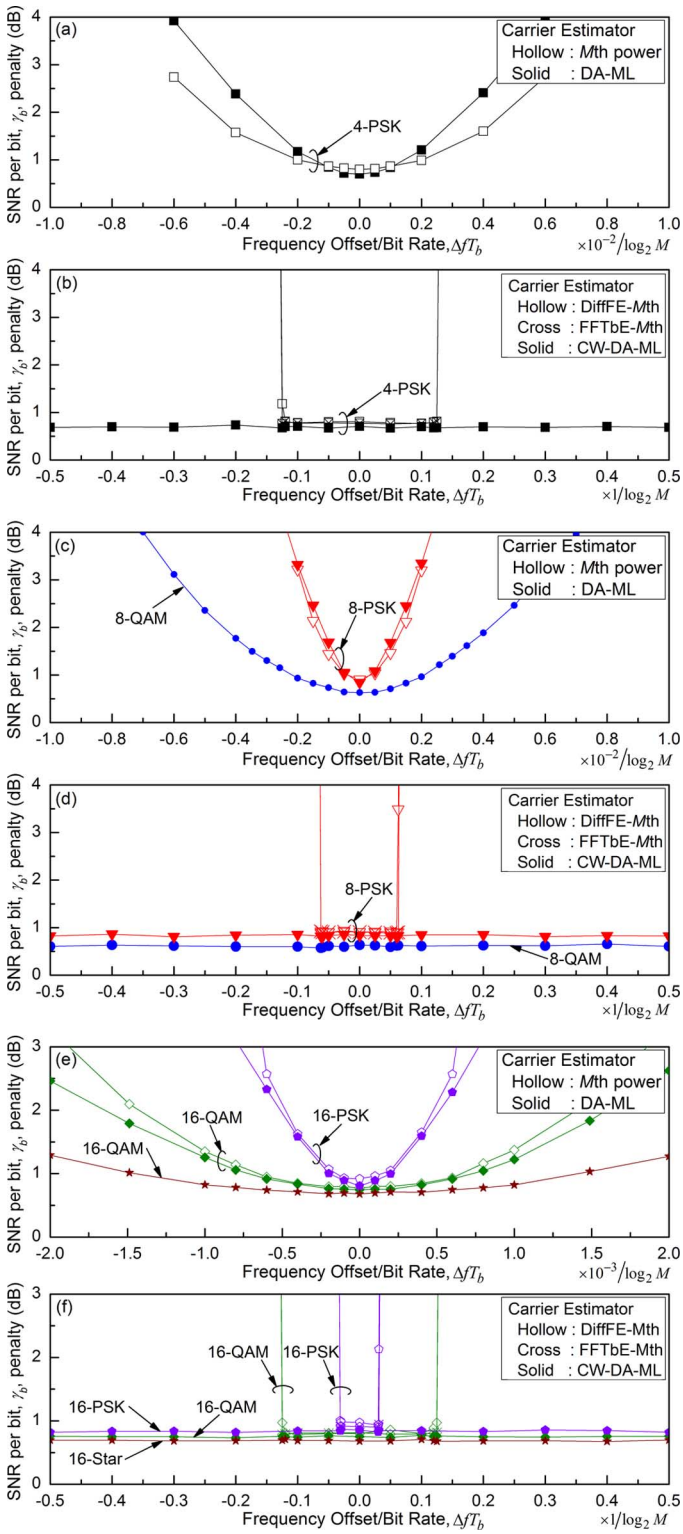

Fig. 10. SNR per bit penalty versus frequency offset per bit rate for (a)&(b) 4-point, (c)&(d) 8-point, and (e)&(f) 16-point constellations. $\Delta\nu = 0$.

On the other hand, CW-DA-ML CE achieves a complete ΔfT_b estimation range of $\pm 1/2\log_2 M$, as it uses an RP with a complete phase tracking range of $[0, 2\pi)$. Note that CW-DA-ML CE is modulation format independent, unlike DiffFE, FFTbE, and block Mth power estimator which are limited to MPSK and 16-QAM format. Moreover, DiffFE, FFTbE, and block Mth power estimators need to be modified separately according to the constellation used, rendering them less attractive in flexible optical systems using multiple modulation formats.

TABLE IV
ESTIMATION RANGE FOR 1-dB SNR PER BIT PENALTY AT BER $= 10^{-3}$ WITH $\Delta\nu = 0$

Format	FREQUENCY OFFSET/BIT RATE, ΔfT_b, $\times \log_2 M$				
	Block Mth power	DA-ML	DiffFE-Mth	FFTbE-Mth	CW-DA-ML
4-PSK	$\pm 2.0 \times 10^{-3}$	$\pm 1.5 \times 10^{-3}$	$\pm 1/8$	$\pm 1/8$	$\pm 1/2$
8-QAM	-	$\pm 2.1 \times 10^{-3}$	-	-	$\pm 1/2$
8-PSK	$\pm 0.4 \times 10^{-3}$	$\pm 0.4 \times 10^{-3}$	$\pm 1/16$	$\pm 1/16$	$\pm 1/2$
16-QAM	$\pm 6.5 \times 10^{-4}$	$\pm 0.7 \times 10^{-3}$	$\pm 1/8$	$\pm 1/8$	$\pm 1/2$
16-Star	-	$\pm 1.4 \times 10^{-3}$	-	-	$\pm 1/2$
16-PSK	$\pm 1.5 \times 10^{-4}$	$\pm 0.2 \times 10^{-3}$	$\pm 1/32$	$\pm 1/32$	$\pm 1/2$

C. Acquisition Time, Accuracy, and SNR Threshold

CE accuracy is determined by the total phase error, i.e., phase estimate error plus frequency estimate error, and not by the individual estimate errors. Total phase error variance against the frequency acquisition time in terms of the received sample size N for DiffFE-Mth, FFTbE-Mth, and CW-DA-ML CE of 4-PSK, 8-QAM, and 16-QAM with constant $\Delta fT_b = 0.1 \times \log_2 M$ are plotted in Fig. 11. The error variance at each sample size N was obtained as $\sigma_\varepsilon^2 = (1/h)\sum_{i=1}^{h} \left|\vartheta_i(N) - \widehat{\vartheta}_i(N)\right|^2$ by averaging over $h = 900$ independent realizations. Here, $\vartheta_i(N)$ is the true total phase and $\widehat{\vartheta}_i(N)$ is the corresponding estimate in the ith realization for sample size N. In DiffFE-Mth and FFTbE-Mth CE, the frequency estimator block uses the first N received samples $\{r(l), 0 \le l < N\}$ to produce a frequency estimate and is thereafter stopped. Subsequent received samples $\{r(l), l \ge N\}$ are immediately corrected using the frequency estimate and fed into the phase estimator block for phase noise estimation. We refer to this structure as the serial configuration. In CW-DA-ML CE, the filter-weight vector $\boldsymbol{w}(l)$ is adapted continuously over the first N received samples and is thereafter stopped. Subsequently, the same filter-weight vector $\boldsymbol{w}(N)$ is used.

First, error variance is plotted for a 1-dB system penalty at BER $= 10^{-3}$ in each modulation format: 4-PSK ($\gamma_b = 7.82$ dB, $\Delta\nu T_b = 9 \times 10^{-5}$), 8-QAM ($\gamma_b = 10.04$ dB, $\Delta\nu T_b = 5.8 \times 10^{-5}$), and 16-QAM ($\gamma_b = 11.53$ dB, $\Delta\nu T_b = 8 \times 10^{-6}$). The error variance decreases initially with N due to improving frequency estimate accuracy and later reaches an error floor limited by AWGN and laser phase noise. The carrier acquisition time, defined as the sample size required for the error variance to reach within 3% of the error floor, is summarized in Table V. CW-DA-ML CE is 2.5 and 10.5 times faster than DiffFE-Mth CE in 4-PSK and 16-QAM, respectively. This is due to the use of recursive least-squares technique in CW-DA-ML CE which is well known to have a short convergence time [37]. Additionally, CW-DA-ML CE uses all N samples for frequency estimation in 16-QAM unlike DiffFE-Mth CE which only uses $N/2$ samples as the probability of a pair of consecutive Class I symbols is 1/4. Although FFTbE has the fastest frequency acquisition, only requiring $N = 45$ and 360 in 4-PSK and

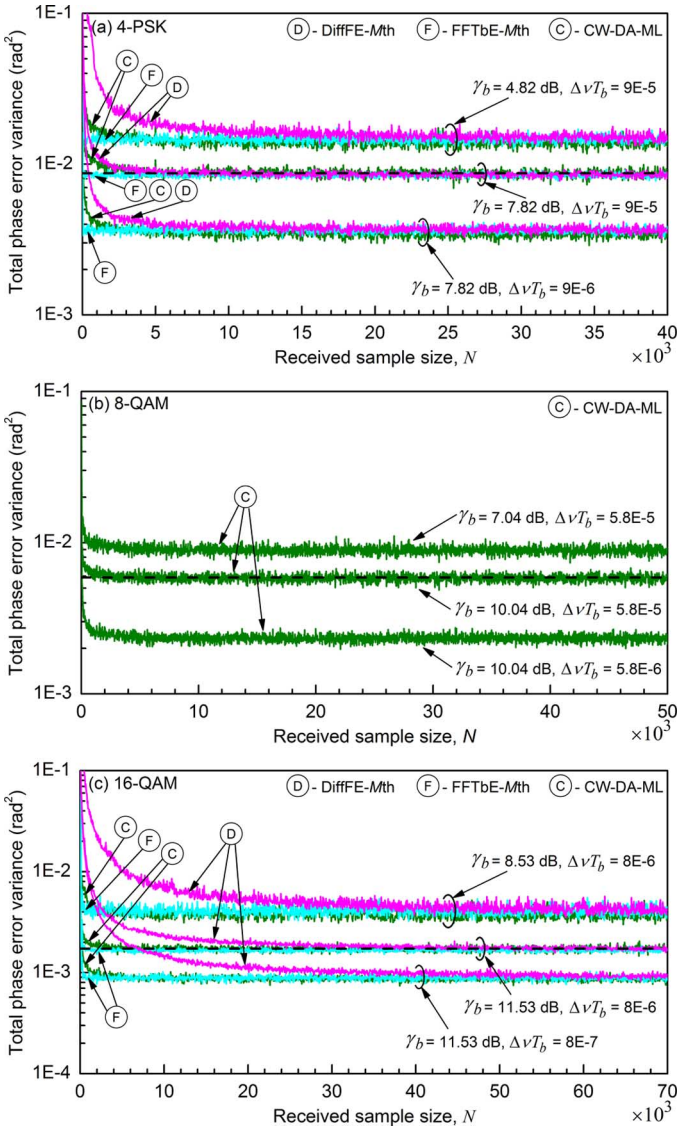

Fig. 11. Error variance versus received sample size N for (a) 4-PSK, (b) 8-QAM, and (c) 16-QAM signals. $\Delta f T_b = 0.1 \times \log_2 M$. The horizontal dashed line '--' is the theoretical error variance of DA-ML estimator with $\Delta f = 0$ at 1-dB system penalty.

16-QAM, respectively, we shall later see it is unsuitable for real-life time-varying frequency offset environment. Frequency acquisition time of CW-DA-ML CE remains about $N = 4 \times 10^3$ for all modulation formats due to its modulation format independence, unlike DiffFE-Mth and FFTbE-Mth CE which have significant modulation-format-dependent time values.

Second, the error variance is plotted at an SNR reduced by 3 dB from the 1-dB system penalty point for each modulation format in Fig. 11. Lower SNR gives rise to higher error floor and increased convergence time as averaging in frequency estimation need to be performed over a larger sample size to smooth out the AWGN. As SNR decreased, speed of CW-DA-ML CE over DiffFE-Mth CE increased to 6 and 11.6 times in 4-PSK and 16-QAM signals, respectively.

Third, the error variance is plotted at a $\Delta\nu T_b$ reduced by an order of magnitude from the 1-dB system penalty point for each modulation format in Fig. 11. Lower $\Delta\nu T_b$ results in lower error floor, and increased convergence time in DiffFE-Mth and

TABLE V
CARRIER ACQUISITION TIME

Format	SNR per bit, γ_b (dB)	Linewidth/ Bit Rate, $\Delta\nu T_b$	Received sample size, N		
			DiffFE-Mth	FFTbE-Mth	CW-DA-ML
4-PSK	4.82	9.0×10^{-5}	30×10^3	165	5×10^3
	7.82	9.0×10^{-5}	10×10^3	45	4×10^3
	7.82	9.0×10^{-6}	15×10^3	45	6×10^3
8-QAM	7.04	5.8×10^{-5}	-	-	4×10^3
	10.04	5.8×10^{-5}	-	-	3×10^3
	10.04	5.8×10^{-6}	-	-	4×10^3
16-QAM	8.53	8.0×10^{-6}	58×10^3	528	5×10^3
	11.53	8.0×10^{-6}	42×10^3	360	4×10^3
	11.53	8.0×10^{-7}	54×10^3	360	5×10^3

CW-DA-ML CE as better frequency estimate can be obtained by averaging over a larger sample size. However, convergence time of FFTbE-Mth CE remains unchanged. This is because the peak position in the FFT spectrum, and thus its frequency estimate accuracy, remains unaffected as variation in $\Delta\nu T_b$ merely alters the spectral width around the peak.

Horizontal dashed line in Fig. 11 depicts the theoretical error variance of DA-ML estimator at $\Delta f = 0$ given by [41]

$$\sigma^2_{\varepsilon,DA-ML} \approx \frac{2L^2 + 3L + 1}{6L}\sigma_p^2 + \frac{1}{2L\gamma_b \log_2 M} \qquad (12)$$

for a 1-dB system penalty at BER $= 10^{-3}$. A similar error floor achieved by CW-DA-ML CE implies near-ideal frequency estimation by our CE. The theoretical error variance given by (12) can thus be used to quickly obtain an approximate error floor achievable by CW-DA-ML CE.

Fig. 12 illustrates the error variance versus γ_b of DiffFE-Mth, FFTbE-Mth, and CW-DA-ML CE in 4-PSK, 8-QAM, and 16-QAM signals using different values of N with $\Delta f T_b = 0.1 \times \log_2 M$. CW-DA-ML CE achieves superior or equal frequency estimation accuracy compared to DiffFE-Mth CE at any given N and γ_b. This can be attributed to the L-sample lag autocorrelation used in CW-DA-ML CE [see (10)] being less affected by AWGN compared to the 1-sample lag autocorrelation used in DiffFE-Mth CE. CW-DA-ML CE tends to outperform FFTbE-Mth CE, and is therefore a better option, at low SNR and/or low N. Low SNR increases the occurrence of outliers and low N reduces the frequency estimate resolution, thus degrading FFTbE-Mth CE. Furthermore, CW-DA-ML CE does not exhibit sharp SNR threshold but rather a gradual deterioration of error variance with decreasing SNR. As the error variance is a decreasing function of N, CW-DA-ML CE can be made suitable for very low SNR operation by adequately increasing N.

In the serial configuration of DiffFE-Mth and FFTbE-Mth CE, the frequency estimate becomes available only at the Nth time point. Hence, the first N received samples need to be processed retrospectively at time point N resulting in a huge processing bottleneck. Alternatively, the first N samples may be treated as training sequence at the expense of a large overhead, e.g., 360 and 42×10^3 samples in 16-QAM for DiffFE-Mth and FFTbE-Mth CE at 1-dB system penalty point, respectively. In contrast, CW-DA-ML CE requires merely $2L$ samples as

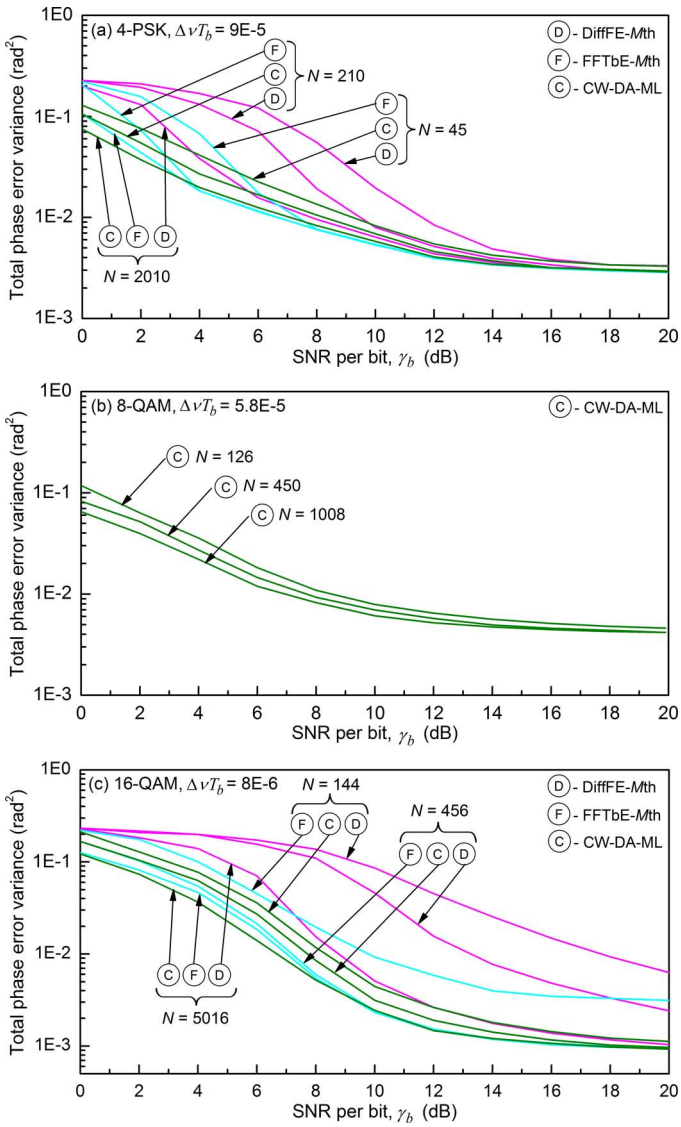

Fig. 12. Error variance versus SNR per bit using various received sample sizes for frequency estimation, N, in (a) 4-PSK, (b) 8-QAM, and (c) 16-QAM signals. $\Delta f T_b = 0.1 \times \log_2 M$.

training sequence and thus has a lower overhead, e.g., 24 samples in 16-QAM at 1-dB system penalty point.

Before proceeding further, it is instructive to consider whether it is feasible to run the frequency estimator and phase estimator block of DiffFE-Mth CE in a concurrent manner. In concurrent configuration, the intermediate frequency estimate after every L samples, $\{\Delta\widehat{f}(iL-1), 1 \le i \le \lfloor N/L \rfloor\}$, is used to frequency-correct the most recent L samples $\{r(l), (i-1)L \le l < iL\}$ followed by phase noise compensation using block Mth power phase estimator. After the Nth sample, the operation becomes the same as the serial configuration. Fig. 13 simulates BER curves of DiffFE-Mth CE using $N = 42 \times 10^3$ for 16-QAM signal at 1-dB system penalty point in serial configuration and concurrent configuration. The BER of the concurrent configuration (BER computed over the first 5×10^6 symbols) deviates from that of the serial configuration, indicating the presence of a large error burst. The error burst can be traced to occur during the first 42×10^3 symbols. When intermediate frequency estimates are used, the frequency

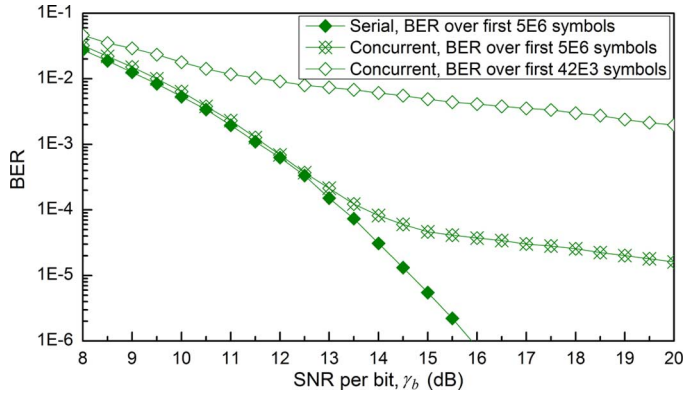

Fig. 13. BER versus SNR per bit of DiffFE-Mth CE in serial and concurrent configuration. $\Delta\nu T_b = 8 \times 10^{-6}$, $\Delta f T_b = 0.1 \times \log_2 M$, and $N = 42 \times 10^3$.

estimate of the $(i+1)$th block will differ from that of the ith block by $\Delta\widehat{f}((i+1)L-1) - \Delta\widehat{f}(iL-1)$. This difference in frequency estimate makes it likely for the first symbol decision of $(i+1)$th block, $\widehat{m}(iL)$, to be differentially decoded incorrectly. The error susceptibility of the first symbol of each block contributes to the error burst. Recovery of the first $N = 42 \times 10^3$ symbols using FEC is not attractive as the BER is higher than 2×10^{-3} even at $\gamma_b = 20$ dB. Moreover, a safe BER margin to the FEC threshold is required by many systems. Hence, the use of concurrent configuration to sidestep the processing bottleneck or large overhead in serial configuration is not practical. Results of Fig. 13 are equally applicable to FFTbE-Mth CE.

In practice, the frequency offset varies with time and needs to be tracked. DiffFE-Mth and FFTbE-Mth CE will incur a large processing bottleneck or a large overhead every time the frequency is re-estimated. It is desirable for a CE algorithm to track the time-varying frequency offset continuously, such as CW-DA-ML CE, to ensure the best performance. DiffFE-Mth and FFTbE-Mth CE can only produce periodically-updated static frequency estimates and is likely to incur some performance penalty as a result.

D. Cycle Slip Probability of Carrier Estimators

When AWGN, laser phase noise, and/or frequency offset pushes the estimate of CE from the true stable operating point into the domain of attraction of a neighbouring stable operating point, a cycle slip is said to have occurred. The estimate remains in the vicinity of the new stable operating point until another cycle slip occurs, causing a large error burst. Angular spacing of the stable operating points, φ, concur with that of the rotationally symmetric positions of the constellation. We have φ equal to $2\pi/M$ in MPSK, $\pi/4$ in 16-Star, and $\pi/2$ in 8- and 16-QAM constellations. Cycle slip, and thus the error burst, can be confined to the slip duration by using DE. However, DE is undesirable for it increases the BER through correlated errors and excludes the use of powerful soft decision FEC codes with high coding gain [42]. Alternatively, cycle slip can be mitigated by inserting pilot sequences at a frequency greater than the cycle slip probability. In PA system, a low cycle slip probability is preferred to minimize the required training overhead cost. Without loss of generality, we use sufficiently large N to ensure frequency estimation convergence and set $\Delta f = 0$ throughout

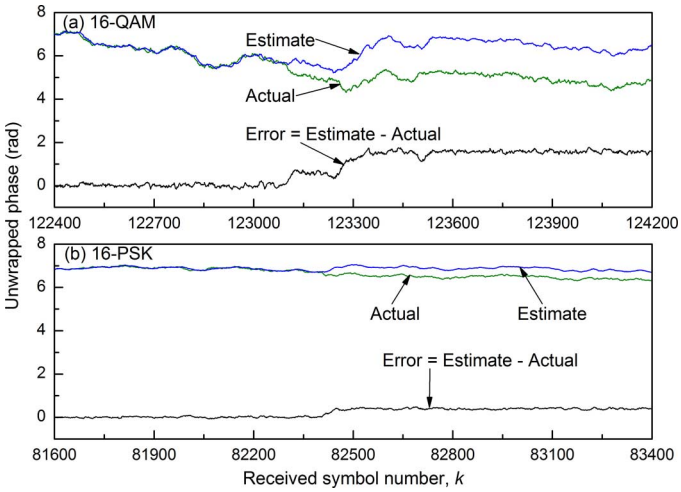

Fig. 14. Unwrapped actual and estimated total phase by CW-DA-ML CE of (a) 16-QAM, and (b) 16-PSK signals. $\Delta f = 0$.

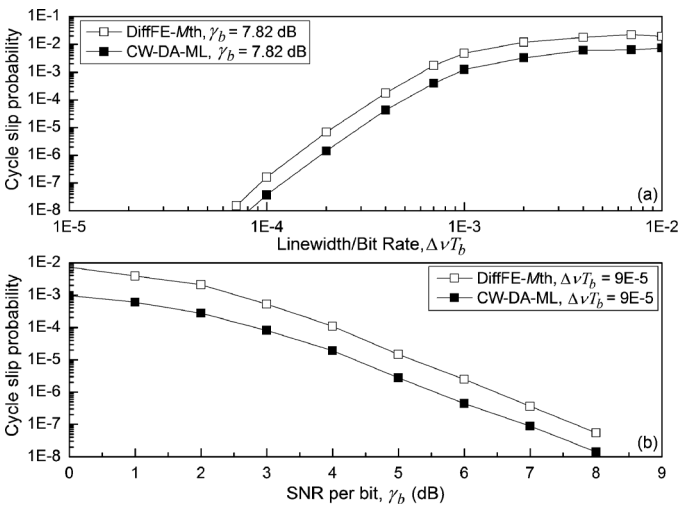

Fig. 15. Cycle slip probability in 4-PSK signal (a) at different $\Delta\nu T_b$ with a fixed $\gamma_b = 7.82$ dB, and (b) at different γ_b with a fixed $\Delta\nu T_b = 9 \times 10^{-5}$.

this subsection. Detection of cycle slip follows the technique of [43], where 11 or more consecutive symbol errors were assumed to be due to a cycle slip.

A cycle slip of $\pi/8$ and $\pi/2$ by CW-DA-ML CE in 16-QAM and 16-PSK signals, respectively, is shown in Fig. 14. It is key to note that the cycle slip in 16-QAM transitioned through an intermediate state of 0.20π rotation before settling at the stable point spaced away by $\pi/2$. An example of this trajectory could be from the domain of point s_9 to s_2 to s_1 in 16-QAM constellation [see Fig. 1(d)].

Fig. 15(a) plots the cycle slip probability versus $\Delta\nu T_b$, and Fig. 15(b) plots the cycle slip probability versus γ_b, for DiffFE-Mth and CW-DA-ML CE of 4-PSK signal. The γ_b was fixed at 1 dB above its theoretical value for BER $= 10^{-3}$ in Fig. 15(a), and the $\Delta\nu T_b$ was fixed at 9×10^{-5} in Fig. 15(b). The cycle slip probability approximates an exponentially increasing function with increasing laser phase noise and decreasing SNR. We see that CW-DA-ML CE has a lower cycle slip probability than DiffFE-Mth CE. For example, at $\gamma_b = 7.82$ dB and $\Delta\nu T_b = 9 \times 10^{-5}$, CW-DA-ML CE achieves a low cycle slip probability of 2×10^{-8} compared

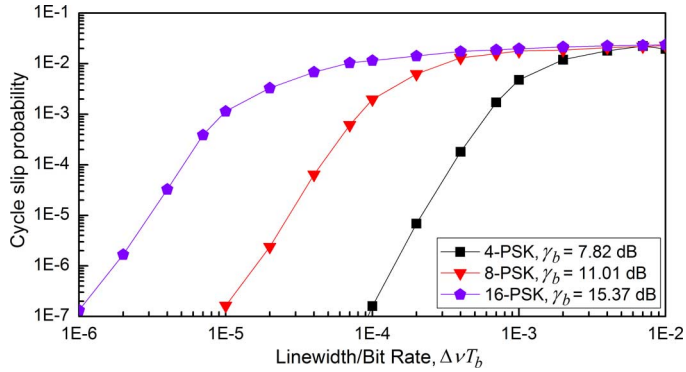

Fig. 16. Cycle slip probability versus linewidth per bit rate of DiffFE-Mth CE.

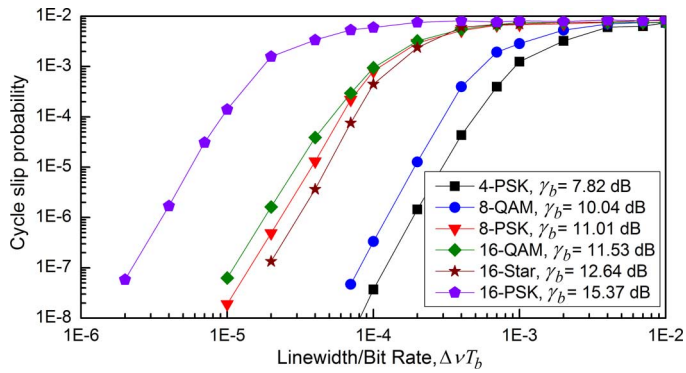

Fig. 17. Cycle slip probability versus linewidth per bit rate of CW-DA-ML CE.

to the 8×10^{-8} of DiffFE-Mth CE. Cycle slip is induced in CW-DA-ML CE by erroneous symbol decision feedback from the data detector during the formation of RP [see (6)]. However, cycle slip in DiffFE-Mth CE is caused by inaccurate phase unwrapping in the block Mth power phase estimator. Due to block Mth power estimator's modulo $2\pi/M$ operation, its phase estimate $\widehat{\theta}(k)$ needs to be unwrapped to track the true laser phase noise trajectory. Unwrap function selects $\widehat{\theta}(k) \pm 2\pi i/M$, $i \in \{0, 1, 2, \ldots\}$, such that $\widehat{\theta}(k) - \widehat{\theta}(k-1)$ is within $\pm\pi/M$. However, if the true $|\theta(k) - \theta(k-1)|$ was greater than π/M, a cycle slip will occur. Increased laser phase noise and reduced SNR contributes to such unwrapping errors as witnessed in Fig. 15.

Besides SNR and laser phase noise, the size of the basic unwrapping interval also contributes to the cycle slip probability. In Fig. 16, the cycle slip probability of DiffFE-Mth CE for 4-, 8-, and 16-PSK signals are plotted at 1 dB above their respective theoretical γ_b values for BER $= 10^{-3}$. As M increases at a given $\Delta\nu T_b$, it is more likely for the true $|\theta(k) - \theta(k-1)|$ to exceed π/M, thus increasing the cycle slip probability. Moreover, introduction of higher powers of noise with modulation order M by block Mth power estimator compounds the angular uncertainty of the received sample, making cycle slips more likely. Therefore, block Mth power based CE, such as DiffFE-Mth and FFTbE-Mth CE, is less desirable in practical PA systems than CW-DA-ML CE due to their higher cycle slip probability.

Suitability of various modulation formats in terms of cycle slip probability is investigated in Fig. 17 using CW-DA-ML CE. Each modulation format was simulated at 1 dB above its theoretical γ_b value for BER $= 10^{-3}$. Sorted in increasing order

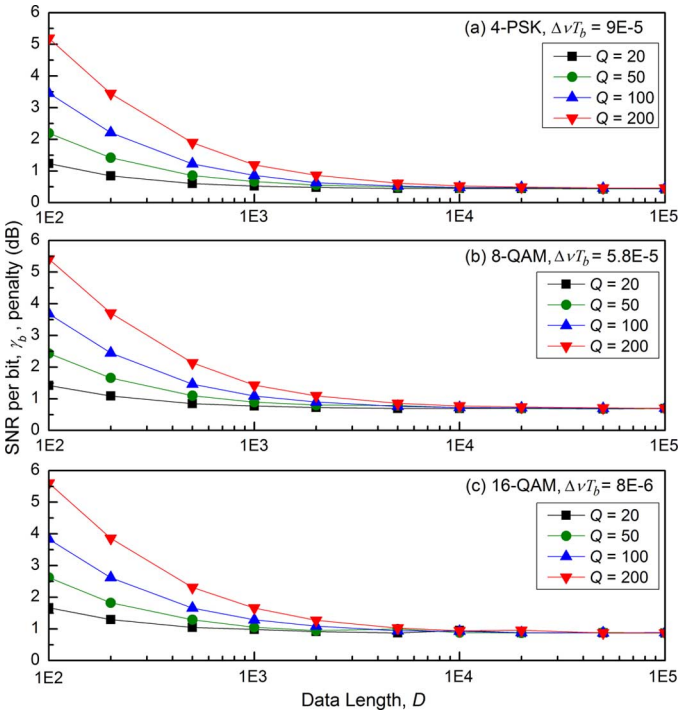

Fig. 18. SNR per bit penalty versus data length at different pilot lengths, Q, for (a) 4-PSK, (b) 8-QAM, and (c) 16-QAM signals. $\Delta f T_b = 0.1 \times \log_2 M$.

of cycle slip probability, we have 4-PSK ($\varphi = \pi/2$), 8-QAM ($\varphi = \pi/2$), 16-Star ($\varphi = \pi/4$), 8-PSK ($\varphi = \pi/4$), 16-QAM ($\varphi = \pi/2$), and 16-PSK ($\varphi = \pi/8$). This order follows the trend of constellations with larger angular separation of stable operating points, φ, having a lower cycle slip probability, but with the exception of 16-QAM. The irregularity can be explained by recalling that cycle slip in 16-QAM tends to occur through an intermediate state spaced apart by 0.20π which is smaller than $\varphi = \pi/4$ of 8-PSK. Hence, it is more likely for 16-QAM to incur cycle slips compared to 8-PSK. In terms of cycle slip tolerance, best 4-, 8-, and 16-point constellations are 4-PSK, 8-QAM, and 16-Star, respectively. However, we should remember that higher SNR and more complex transmitter is required for 16-Star than 16-QAM.

V. Pilot-Assisted Carrier Estimation

Cycle slips can be combated using DE or pilot sequences. PA CE is preferred in practice for it avoids the DE penalty listed in column four of Table I, enables the use of powerful soft decision FEC codes with high coding gain, and can be simultaneously used for fiber nonlinearity compensation. In PA CE, alternating D-symbol-long data and Q-symbol-long pilot sequences are transmitted. When data are transmitted in packet frames, headers containing protocols such as the physical addresses of the receiver and FEC information can be used as pilot symbols. Extra symbol energy of the pilots is accounted for in our PA CE simulation by computing the effective launched energy of each symbol, $E_{s,eff}$, as $E_{s,eff} = E_{s,actual} \cdot (D+Q)/D$. Here, $E_{s,actual}$ is the actual energy of each transmitted symbol. Fig. 18 shows the γ_b penalty of PA CW-DA-ML CE at BER $= 10^{-3}$ with various training overhead costs, defined as $Q/(D+Q)$. As the ratio

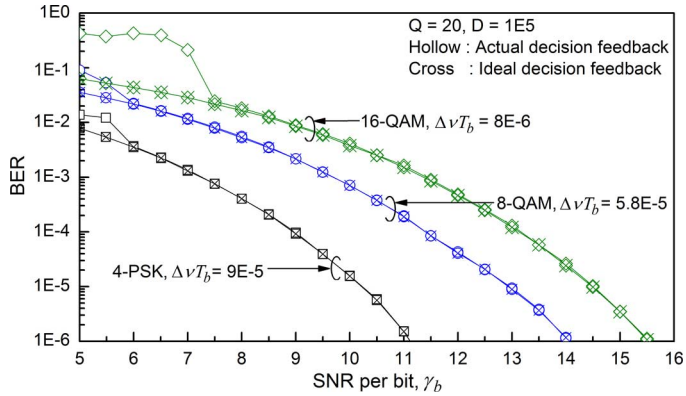

Fig. 19. BER versus SNR per bit for PA CW-DA-ML CE with ideal and actual decision feedback. $\Delta f T_b = 0.1 \times \log_2 M$.

$(D+Q)/D$ decreases towards 1, the $E_{s,actual}$ approaches $E_{s,eff}$, hence the receiver sensitivity improves. From Fig. 18, PA system incurs a γ_b penalty of 0.43, 0.70, and 0.87 dB in 4-PSK, 8-QAM, and 16-QAM, respectively, at $D = 10^4$ and $Q = 20$. Since the $\Delta\nu T_b$ used was set at the tolerance value for a 1-dB γ_b penalty in a DE system, a gain of up to 0.57 dB in 4-PSK, 0.30 dB in 8-QAM, and 0.13 dB in 16-QAM is achieved by the PA system compared to its DE counterpart while keeping the overhead costs as low as 0.2%.

In Fig. 18, the improvement in receiver sensitivity levels off for data sequence length $D \geq 10^4$, indicating negligible error propagation arising from a low cycle slip probability. We can infer that the mean time to lose lock (i.e., cycle slip) is greater than 10^5 symbols. This is proven in Fig. 19 by the negligible performance loss with actual, compared to ideal, decision feedback for BER $\leq 10^{-3}$ at $D = 10^5$ and $Q = 20$.

VI. Time-Varying Frequency Offset

In practice, laser frequency drifts over time in the MHz/s range due to aging or temperature variation and might also experience sudden frequency jumps due to mechanical disturbances to the laser cavity. Hence, the frequency offset needs to be continuously tracked for best BER performance in a symbol-by-symbol receiver. Fig. 20 evaluates the robustness of PA CW-DA-ML CE of a 14 Gbaud single polarization 16-QAM signal at $\gamma_b = 12$ dB, $\Delta\nu T_b = 8 \times 10^{-6}$, and a training overhead of 0.2% ($D = 10^4$, $Q = 20$) in a time-varying frequency offset environment. The stable BER, measured at 10 ms intervals, demonstrates the reliable tracking of frequency offset experiencing a continuous drift of 10 MHz/s and rapid jumps of 100 kHz every 10 ms. PA CW-DA-ML CE can continuously track the frequency offset, thanks to its observation dependent weight vector.

VII. ADC Resolution

In coherent receivers, the real and imaginary dimensions of each polarization are sampled and quantized to a discrete set of values by ADCs, whose resolution is determined in number of bits, B. In general, ADCs with higher sampling rates are limited to lower resolution [44]. Hence, higher quantization error is introduced in coherent systems requiring higher sampling rates.

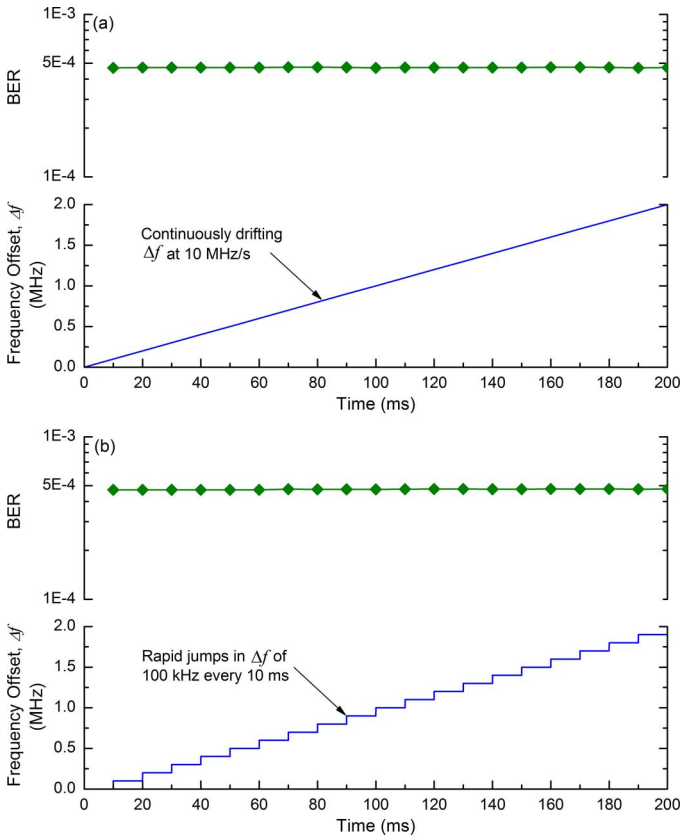

Fig. 20. BER of PA CW-DA-ML CE in time-varying frequency offset experiencing: (a) continuous drift, and (b) rapid jumps. $\gamma_b = 12$ dB, $\Delta\nu T_b = 8 \times 10^{-6}$, and training overhead of 0.2%.

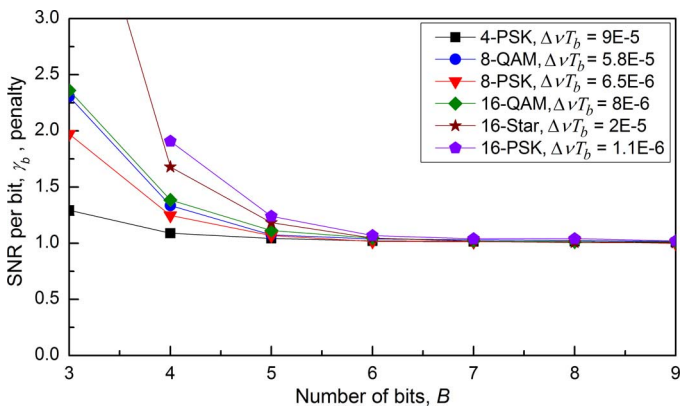

Fig. 21. SNR per bit penalty versus ADC resolution of DE CW-DA-ML CE. Time-invariant $\Delta f T_b = 0.1 \times \log_2 M$.

Impact of uniform quantization by ADC on DE CW-DA-ML CE is investigated in Fig. 21 under a time-invariant $\Delta f T_b$ of $0.1 \times \log_2 M$. The $\Delta\nu T_b$ is fixed at the tolerance value for a 1-dB γ_b penalty obtained in Table III. In an uniformly quantizing ADC, each dimension with a range of $[+A, -A]$ is divided into 2^B non-overlapping quantization intervals of equal width α. The midpoint of each interval is designated as a centroid. We have $A = 2^B \sqrt{E_{d,\max}} / \left(2^B - 1\right)$, where $E_{d,\max}$ is the maximum transmitted symbol energy in the respective dimension. For the constellations considered in Fig. 21, the $E_{d,\max}$ of the real and imaginary dimensions are equal. The received signal in each dimension is quantized to the nearest centroid. The quantization error can be treated as an additional Gaussian additive noise with variance $\alpha^2/12$ per dimension [45]. An ADC resolution greater than 5 bits is seen to be sufficient for all the modulation formats tested. The bits requirement may be reduced by using ADC with nonuniform quantization steps for nonequally spaced signal point constellations such as 8-QAM.

VIII. Complexity Analysis

The least-squares solution $\boldsymbol{w}(k) = \boldsymbol{\Phi}^{-1}(k)\,\boldsymbol{z}(k)$ of CW-DA-ML CE can be computed recursively using the matrix inversion lemma [27], [37]. Hence, $\boldsymbol{\Phi}^{-1}(k)$ is obtained in terms of $\boldsymbol{\Phi}^{-1}(k-1)$, avoiding any actual matrix inversion. The matrix $\boldsymbol{\Phi}^{-1}(k)$ is Hermitian, thus only the upper triangle needs to be computed and stored whereas the lower triangle is filled by diagonal reflection. This reduces the memory size required to $L^2 + 6L + 4$ buffer units. Each buffer unit is defined to hold one real value. Our transversal recursive least-squares CE has a complexity of $O\left(L^2\right)$ to produce the phase and angular frequency offset estimate, i.e., $V(k)$, per symbol. It needs $6L^2 + 14L + 10$ real multiplications and $6L^2 + 8L + 6$ real additions. However, the CW-DA-ML algorithm can be equally realized in the form of a recursive least-squares lattice filter, reducing the complexity to $O(L)$ real multiplications and additions per symbol [37]. Use of coordinate rotation digital computers to implement the recursive least-squares lattice filter is expected to further simplify the computation as it can perform vector rotations in the complex plane efficiently [46].

We consider the individual complexities of DiffFE, FFTbE, block Mth power, and the blind phase search (BPS) phase estimator of [36]. Complexity of the frequency and phase estimators are taken to be the computations required to estimate the angular frequency offset $\Delta\omega$ over N symbols and that to estimate the phase θ per symbol, respectively. Each $\arctan(\cdot)$ and phase unwrapping operation is expressed as one access to a read-only memory (ROM) and one unwrap function, respectively. The complexities of DiffFE, radix-2 FFTbE, block Mth power, and BPS in PSK and QAM formats are summarized in Table VI. Here, β is the number of test phase angles used in BPS estimator. Complexity of DiffFE and block Mth power estimators for 16-QAM was obtained by using the Class I symbol probability of 1/2. FFTbE incurs its total complexity stated in Table VI at time point N causing a processing bottleneck, since FFT is only performed after N symbols. Note that the required number of test phase angles β, and thus the complexity, of BPS increases with modulation order [36].

A simple straightforward complexity comparison among the algorithms may not be wholly justified. CW-DA-ML CE is a symbol-by-symbol carrier estimator whereas DiffFE-Mth and FFTbE-Mth CE only produces static frequency estimates after N symbols. Moreover, unlike CW-DA-ML CE, BPS is susceptible to frequency offsets. Compared to other estimators in general, CW-DA-ML CE has increased number of multiplications and additions but avoids any intermediate decision devices, comparators, ROM access, and unwrap functions. The increase in number of multiplications and additions is traded off with its fast, wide, and continuous frequency tracking features. It is notable that our CE reduces the required buffer units by a factor greater than 2.7 compared to FFTbE and BPS in 16-QAM. For 16-QAM, CW-DA-ML CE ($L = 12$) requires

TABLE VI
CARRIER ESTIMATION COMPLEXITY ANALYSIS

Estimator	Complexity to estimate	Format	Real multiplications	Real additions	Decision devices	Comparators	ROM access	Unwrap function	Buffer Units
DiffFE	$\Delta\omega$ over N symbols	MPSK	$4(N-1)(1+log_2M)+1$	$2(N-1)(1+log_2M)+2N-2$	0	0	1	0	0
		16-QAM	$7N+1$	$4N-2$	0	2	1	0	0
FFTbE	$\Delta\omega$ over N symbols	MPSK	$2Nlog_2N+4Nlog_2M+2N+2$	$3Nlog_2N+2Nlog_2M+N$	0	N	0	0	$2N$
		16-QAM	$2Nlog_2N+10N+2$	$3Nlog_2N+5N$	0	N	0	0	$2N$
Block Mth power	θ per symbol	MPSK	$4log_2M+1/L$	$2log_2M+2-2/L$	0	0	$1/L$	$1/L$	0
		16-QAM	$24+5/L$	$17-2/L$	0	$(L+4)/2L$	$1/L$	$1/L$	0
BPS	θ per symbol	MPSK/ MQAM	6β	$(L+4)\beta$	β	β	0	1	$L\beta$
CW-DA-ML	$V(k)$ per symbol	MPSK/ MQAM	$6L^2+14L+10$	$6L^2+8L+6$	0	0	0	0	L^2+6L+4

220 buffer units as opposed to the 720 buffer units of FFTbE ($N = 360$) and 608 buffer units of BPS ($L = 19$, $\beta = 32$).

IX. CONCLUSION

By considering the AWGN immunity, laser phase noise tolerance, ease of differential encoding, cycle slip probability, and transmitter implementation complexity, we conclude that the most viable 4-, 8-, and 16-point constellations for coherent optical communication are 4-PSK, 8-QAM, and 16-QAM, respectively. Our causal CW-DA-ML carrier estimator achieves a complete frequency offset estimation range and avoids phase unwrapping as it uses a reference phasor with an unambiguous phase tracking range of $[0, 2\pi)$. Having lower cycle slip probability than block Mth power based estimators, continuous carrier tracking feature, low training overhead, ability to operate at low SNR region, and being modulation format independent makes CW-DA-ML algorithm an attractive carrier estimator for flexible multi-modulation coherent receivers with laser frequency instabilities.

REFERENCES

[1] P. J. Winzer, "Beyond 100 G Ethernet," *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 26–30, Jul. 2010.

[2] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Exp.*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[3] A. Mecozzi and M. Shtaif, "On the capacity of intensity modulated systems using optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 13, no. 9, pp. 1029–1031, Sep. 2001.

[4] K.-P. Ho and J. M. Kahn, "Channel capacity of WDM systems using constant-intensity modulation formats," in *Proc. OFC/NFOEC*, Anaheim, CA, 2002, Paper ThGG85.

[5] K.-P. Ho, "Exact evaluation of the capacity for intensity-modulated direct-detection channels with optical amplifier noises," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 858–860, Apr. 2005.

[6] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 259–272, Mar. 2004.

[7] L. G. Kazovsky, G. Kalogerakis, and W.-T. Shaw, "Homodyne phase-shift-keying systems: Past challenges and future opportunities," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4876–4884, Dec. 2006.

[8] C. Berrou, "The ten-year-old turbo codes are entering into service," *IEEE Commun. Mag.*, vol. 41, no. 8, pp. 110–116, Aug. 2003.

[9] E. Ip, J. M. Kahn, D. Anthon, and J. Hutchins, "Linewidth measurements of MEMS-based tunable lasers for phase-locking applications," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2029–2031, Oct. 2005.

[10] Integrable Tunable Laser Assembly Multi Source Agreement 2005, OIF-ITLA-MSA-01.1.

[11] S. Norimatsu and K. Iwashita, "Linewidth requirements for optical synchronous detection systems with nonnegligible loop delay time," *J. Lightw. Technol.*, vol. 10, no. 3, pp. 341–349, Mar. 1992.

[12] K.-Y. Kim and H.-J. Choi, "Design of carrier recovery algorithm for high-order QAM with large frequency acquisition range," in *Proc. ICC*, Helsinki, Finland, 2001, pp. 1016–1020.

[13] P. J. Winzer, A. H. Gnauck, C. R. Doerr, M. Magarini, and L. L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 547–556, Feb. 2010.

[14] T. Kobayashi *et al.*, "160-Gb/s polarization-multiplexed 16-QAM long-haul transmission over 3,123 km using digital coherent receiver with digital PLL based frequency offset compensator," in *Proc. OFC/NFOEC*, San Diego, CA, 2010, Paper OTuD1.

[15] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sep. 2010.

[16] 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.

[17] 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.

[18] Y. Wang, E. Serpedin, P. Ciblat, and P. Loubaton, "Non-data aided feedforward cyclostationary statistics based carrier frequency offset estimators for linear modulations," in *Proc. Conf. Rec. GLOBECOM'01*, Paris, France, 2001, pp. 1386–1390.

[19] M. Morelli and U. Mengali, "Feedforward frequency estimation for PSK: A tutorial review," *Eur. Trans. Telecommun.*, vol. 9, no. 2, pp. 103–116, Mar. 1998.

[20] 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.

[21] 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.

[22] H. Fu and P.-Y. Kam, "MAP/ML estimation of the frequency and phase of a single sinusoid in noise," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 834–845, Mar. 2007.

[23] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. New York, NY, USA: Wiley, 1997.

[24] W. J. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.*, vol. 26, no. 3, pp. 385–391, Mar. 1978.
[25] P.-Y. Kam, "Maximum likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.*, vol. 34, no. 6, pp. 522–527, Jun. 1986.
[26] R. D. Gaudenzi, T. Garde, and V. Vanghi, "Performance analysis of decision-directed maximum-likelihood phase estimators for M-PSK modulated signals," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 3090–3100, Dec. 1995.
[27] A. Meiyappan, P.-Y. Kam, and H. Kim, "A complex-weighted, decision-aided, maximum-likelihood carrier phase and frequency-offset estimation algorithm for coherent optical detection," *Opt. Exp.*, vol. 20, no. 18, pp. 20102–20114, Aug. 2012.
[28] *Forward Error Correction for High Bit-Rate DWDM Submarine Systems*, ITU-T Recommendation G.975.1, 2004.
[29] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *J. Lightw. Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.
[30] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.
[31] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," *IEEE Trans. Commun.*, vol. 22, no. 1, pp. 28–38, Jan. 1974.
[32] E. Ip and J. M. Kahn, "Carrier synchronization for 3- and 4-bit-per-symbol optical transmission," *J. Lightw. Technol.*, vol. 23, no. 12, pp. 4110–4124, Dec. 2005.
[33] W. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Communications*. London, U.K.: Pentech Press, 1994.
[34] H. Zhang, P.-Y. Kam, and C. Yu, "Optimal ring ratio of 16-star quadrature amplitude modulation in coherent optical communication systems," in *Proc. OECC*, Kaohsiung, Taiwan, 2011, pp. 577–578.
[35] C. R. Doerr, L. Zhang, P. J. Winzer, and A. H. Gnauck, "28-Gbaud InP square or hexagonal 16-QAM modulator," in *Proc. OFC/NFOEC*, Los Angeles, CA, 2011, Paper OMU2.
[36] 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.
[37] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008.
[38] Y. Atzmon and M. Nazarathy, "Laser phase noise in coherent and differential optical transmission revisited in the polar domain," *J. Lightw. Technol.*, vol. 27, no. 1, pp. 19–29, Jan. 2009.
[39] H. Zhang, S. Zhang, P.-Y. Kam, C. Yu, and J. Chen, "Optimized phase error tolerance of 16-star quadrature amplitude modulation in coherent optical communication systems," in *Proc. OECC*, Sapporo, Japan, 2010, pp. 592–593.
[40] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "On the selection of a two-dimensional signal constellation in the presence of phase jitter and Gaussian noise," *Bell Syst. Tech. J.*, vol. 52, no. 6, pp. 927–965, Feb. 1973.
[41] S. Zhang, P.-Y. Kam, J. Chen, and C. Yu, "Bit-error rate performance of coherent optical M-ary PSK/QAM using decision-aided maximum likelihood phase estimation," *Opt. Exp.*, vol. 18, no. 12, pp. 12088–12103, Jun. 2010.
[42] E. Ip and J. M. Kahn, "Addendum to "Feedforward carrier recovery for coherent optical communications"," *J. Lightw. Technol.*, vol. 27, no. 13, pp. 2552–2553, Jul. 2009.
[43] 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.
[44] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.
[45] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.
[46] J. E. Volder, "The CORDIC trigonometric computing technique," *IRE Trans. Electron. Comp.*, vol. EC-8, no. 3, pp. 330–334, Sep. 1959.

Adaickalavan Meiyappan (S'13) received the B.E. degree in electrical engineering from the National University of Singapore, Singapore, in 2010, where he is currently working toward the Ph.D. degree in electrical engineering.

His research interest includes digital signal processing for coherent optical communication systems.

Pooi-Yuen Kam (F'10) was born in Ipoh, Malaysia, and educated at the Massachusetts Institute of Technology, Cambridge, Mass., USA where he obtained the S.B., S.M., and Ph.D. degrees in electrical engineering in 1972, 1973, and 1976, respectively.

From 1976 to 1978, he was a member of the technical staff at the Bell Telephone Laboratories, Holmdel, N.J., U.S.A., where he was engaged in packet network studies. Since 1978, he has been with the Department of Electrical and Computer Engineering, National University of Singapore, where he is now a professor. He served as the Deputy Dean of Engineering and the Vice Dean for Academic Affairs, Faculty of Engineering of the National University of Singapore, from 2000 to 2003. His research interests are in the communication sciences and information theory, and their applications to wireless and optical communications. He spent the sabbatical year 1987 to 1988 at the Tokyo Institute of Technology, Tokyo, Japan, under the sponsorship of the Hitachi Scholarship Foundation. In year 2006, he was invited to the School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada, as the David Bested Fellow.

Dr. Kam is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. Since September 2011, he is a senior editor of the IEEE Wireless Communications Letters. From 1996 to 2011, he served as the Editor for Modulation and Detection for Wireless Systems of the IEEE Transactions on Communications. He also served on the editorial board of PHYCOM, the Journal of Physical Communications of Elsevier, from 2007 to 2012. He was elected a Fellow of the IEEE for his contributions to receiver design and performance analysis for wireless communications. He received the Best Paper Award at the IEEE VTC2004-Fall, at the IEEE VTC2011-Spring, and at the IEEE ICC2011.

Hoon Kim (S'97–A'00–M'04–SM'11) received the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Taejon, Korea, in 1996 and 2000, respectively.

From 2001 to 2002, he was with Bell Laboratories, Lucent Technologies, where he worked on advanced optical modulation formats. From 2002 to 2007, he was with Samsung Electronics, Suwon, Korea, where he was involved in research on optical duobinary transmission systems, optical sources for access applications, and fiber-optic networks for wireless communications. Since 2007, he has been an Assistant Professor of Department of Electrical & Computer Engineering, the National University of Singapore.

Dr. Kim now serves as an Associate Editor of IEEE Photonics Technology Letters and has served on several international conferences as technical committee members, including OFC, IEEE LEOS, OECC, and APOC.