

A complex-weighted, decision-aided, maximum-likelihood carrier phase and frequency-offset estimation algorithm for coherent optical detection

Adaickalavan Meiyappan, Pooi-Yuen Kam,* and Hoon Kim

Department of Electrical and Computer Engineering, National University of Singapore, 117576, Singapore
*elekampy@nus.edu.sg

Abstract: We present a symbol-by-symbol coherent optical receiver, which employs a novel, complex-weighted, decision-aided, maximum-likelihood (CW-DA-ML) carrier phase and frequency offset estimator. The CW-DA-ML carrier estimator uses a CW transversal filter to generate a carrier reference phasor, and the filter weights are automatically adapted on-line by linear regression on the observed signals. A complete modulo- R reduced frequency offset estimation (FOE) range of $\pm R/2$ is achieved, independent of modulation format, where R is the symbol rate. Carrier phase and frequency tracking is achieved rapidly. The acquisition speed of frequency offset in quaternary phase-shift keying (4-PSK) signals is more than 5 times faster than that of differential FOE. A constant penalty of approximately 1 dB at bit-error rate of 10^{-4} is demonstrated for all frequency offsets in 4-PSK signals with laser-linewidth-symbol-duration product of 8×10^{-5} .

©2012 Optical Society of America

OCIS codes: (060.4510) Optical communications; (060.1660) Coherent communications.

References and links

1. F. Derr, "Coherent optical QPSK intradyne system: concept and digital receiver realization," *J. Lightwave Technol.* **10**(9), 1290–1296 (1992).
 2. 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. Lightwave Technol.* **24**(1), 12–21 (2006).
 3. P. Y. Kam, "Maximum likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.* **34**(6), 522–527 (1986).
 4. S. Zhang, P. Y. Kam, J. Chen, and C. Yu, "Decision-aided maximum likelihood detection in coherent optical phase-shift-keying system," *Opt. Express* **17**(2), 703–715 (2009).
 5. "Integrable Tunable Laser Assembly Multi Source Agreement," OIF-ITLA-MSA-01.1 (Optical Networking Forum, Los Angeles, 2005).
 6. A. Meiyappan, P. Y. Kam, and H. Kim, "Performance of decision-aided maximum-likelihood carrier phase estimation with frequency offset," in *Proc. OFC/NFOEC, Los Angeles, CA, 2012*, paper OTu2G.6.
 7. A. Leven, N. Kaneda, U.-V. Koc, and Y.-K. Chen, "Frequency estimation in intradyne reception," *IEEE Photon. Technol. Lett.* **19**(6), 366–368 (2007).
 8. W. J. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.* **26**(3), 385–391 (1978).
 9. K.-P. Ho, *Phase-modulated Optical Communication* (Springer, 2005).
 10. P. Y. Kam, S. S. Ng, and T. S. Ng, "Optimum symbol-by-symbol detection of uncoded digital data over the Gaussian channel with unknown carrier phase," *IEEE Trans. Commun.* **42**(8), 2543–2552 (1994).
 11. P. Y. Kam, K. H. Chua, and X. Yu, "Adaptive symbol-by-symbol reception of MPSK on the Gaussian channel with unknown carrier phase characteristics," *IEEE Trans. Commun.* **46**(10), 1275–1279 (1998).
 12. J. G. Proakis, *Digital Communications* (McGraw-Hill, 2008).
 13. U. Mengali and A. N. D' Andrea, *Synchronization Techniques for Digital Receivers* (Plenum Press, 1997).
 14. H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers* (John Wiley, 1997).
 15. T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Lightwave Technol.* **27**(8), 989–999 (2009).
-

1. Introduction

Traditionally, a major challenge in homodyne coherent receivers is to phase- and frequency-lock the local oscillator (LO) laser to the received optical signal. Recently, digital signal processing (DSP) algorithms have been widely used to perform post-detection carrier synchronization, thereby giving birth to the intradyne coherent receiver featuring an intermediate frequency (IF) which allows for a free-running LO laser [1]. A prevalent DSP based block M th power carrier estimation (CE) was proposed in [2], but is meant for M -ary phase-shift keying (MPSK) formats only and requires nonlinear operations which increase system latency. A computationally linear, decision-aided, maximum-likelihood (DA-ML) CE with comparable performance to block M th power CE was derived in [3] and proposed for use in coherent optical communications in [4]. In an intradyne receiver, the frequency offset, Δf , between LO and transmitter laser can be as large as ± 5 GHz [5]. However, block M th power and DA-ML CE were designed to estimate the carrier phase without considering any frequency offset. Block M th power and DA-ML CE have limited frequency offset tolerance of 10^{-3} times the symbol rate at laser-linewidth-symbol-duration product of 6.35×10^{-5} [2, 6]. Hence, it is imperative for a digital CE to incorporate frequency offset estimation (FOE) capability.

A differential FOE, using the sum of differenced signal phases over an observation interval, was suggested for the block M th power CE in [7]. It utilizes the nonlinear operations of $\arctan(\cdot)$ and M th power to remove data modulation. Consequently, it is only applicable to MPSK formats and the FOE range is limited to $\pm R/2M$, where R is the symbol rate. With higher order modulation formats, the FOE range shrinks while the complexity of raising to the M th power increases. This differential FOE has an R/M frequency estimate ambiguity, requiring differential encoding (DE) which leads to a DE-induced performance penalty [8].

Extending upon the DA-ML CE of [3, 4], we propose here a modulation-format-independent symbol-by-symbol receiver employing a novel, complex-weighted (CW), DA-ML CE to estimate the unknown carrier phase and frequency offset. A complete modulo- R reduced frequency offset range of $\pm R/2$ can be estimated for arbitrary modulation formats. CW-DA-ML CE achieves near-ideal FOE with a low training overhead cost and no *a priori* knowledge of the system statistics. Carrier phase and frequency are acquired rapidly. The FOE for 4-PSK signals is more than 5 times faster than that of the differential FOE of [7].

The paper is organized as follows. In Section 2, the received signal in a phase and polarization diversity intradyne coherent receiver is modeled. Section 3 provides the derivation of the symbol-by-symbol receiver structure consisting of CW-DA-ML CE and data detection (DD). Section 4 examines the CE performance via simulations, and includes a discussion on its complexity and parallel implementation. Section 5 concludes the paper. Throughout this paper, $E[\cdot]$ is the expectation, $|\cdot|$ is the modulus operator, boldface characters represent matrices, and δ is the Kronecker delta. Superscript $*$, T , and H denote conjugate, transpose, and complex conjugate transpose, respectively. All phase quantities are treated modulo- 2π to account for their circular nature and frequency quantities are treated modulo- R , which is justified later in Subsection 4.3.

2. Phase and polarization diversity intradyne coherent receiver

The opto-electronic front end followed by DSP modules in a phase and polarization diversity intradyne coherent receiver is shown in Fig. 1. The entire received optical field, $E_r(t)$, within the receiver bandwidth is mapped into x - and y -polarized, in-phase, I , and quadrature-phase, Q , photocurrents by mixing $E_r(t)$ with an LO using two 90° optical hybrids. The LO laser is free running, in contrast to a homodyne coherent receiver where the LO laser needs to be locked to the received optical signal. Each photocurrent branch $i(t)$ is sampled at symbol rate R which amounts to sampling the complex signal at $2R$, therefore satisfying the Nyquist theorem for fully preserving the optical signal and noise statistics.

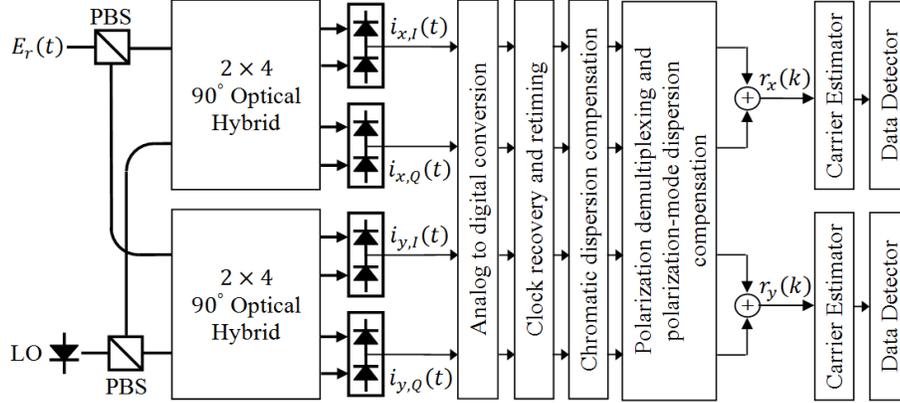


Fig. 1. Phase and polarization diversity intradyne coherent receiver. LO: local oscillator, PBS: polarization beam splitter

We assume ideal signal conditioning by the first four DSP blocks, namely, analog-to-digital conversion (ADC), clock recovery and retiming, chromatic dispersion compensation, and polarization demultiplexing and polarization-mode dispersion compensation. The received signal samples $r_x(k)$ and $r_y(k)$ over the k th symbol interval $[kT, (k+1)T)$ ($T =$ symbol duration) are clock-synchronized with one complex sample per symbol. They are assumed to be free of intersymbol interference (ISI), polarization crosstalk, and nonlinear distortions. The $r(k)$ in one of the polarizations can be shown to be [9]

$$r(k) = i_r(k) + ji_q(k) = \chi\sqrt{P_r P_{LO}} e^{j(2\pi\Delta f T k + \phi(k) + \theta(k))} + \chi\sqrt{P_{LO}} n'(k) e^{j(2\pi\Delta f T k - \theta_{LO}(k))} + i_{sh} \quad (1)$$

where χ , P_r , P_{LO} , and $\phi(k)$ denote responsivity of the photodiode, average received signal power, LO power, and data phase in the k th interval, respectively. Here, $\theta(k)$ is the combined laser phase noise of the transmitter and LO lasers, whereas $\theta_{LO}(k)$ is the LO phase noise. Most long-haul optical fiber systems are limited by amplified spontaneous emission (ASE) noise from inline optical amplifiers, which is represented by $n'(k)$ in Eq. (1). The i_{sh} is the LO shot noise, assuming $P_{LO} \gg P_r$. Thermal noise is ignored in Eq. (1). On the right-hand side of Eq. (1), the first term is the signal-LO beat term and the second term is the LO-ASE beat noise. Using a pair of balanced photodiodes to detect each photocurrent branch, compared to a single photodiode in [1], eliminates the signal-ASE and ASE-ASE beat noise terms.

A canonical model of the received signal in Eq. (1) can be written as

$$r(k) = m(k) e^{j(2\pi\Delta f T k + \theta(k))} + n(k), \quad k = 0, 1, 2, \dots \quad (2)$$

where $m(k) = \chi\sqrt{P_r P_{LO}} e^{j\phi(k)}$ is the k th data symbol. The LO-ASE beat noise and LO shot noise are modeled as additive white Gaussian noise (AWGN), and are combined to form $n(k) = \chi\sqrt{P_{LO}} n'(k) e^{j(2\pi\Delta f T k - \theta_{LO}(k))} + i_{sh}$. The set $\{n(k)\}$ is a sequence of circularly symmetric, AWGN with mean zero and covariance function $E[n(k)n^*(k-i)] = N_o \delta(i)$. The carrier phase $\theta(k)$ has a Gaussian random-walk model of $\theta(k) = \eta(k) + \theta(k-1)$,

where $\{\eta(k)\}$ is a set of independent, identically distributed, Gaussian random variables with mean zero and variance $\sigma_p^2 = 2\pi\Delta\nu T$ [9]. Here, $\Delta\nu$ is the combined 3-dB linewidth of the transmitter and LO lasers. The carrier phase $\theta(k)$ has a temporal correlation characterized by $E[\theta(i)\theta(j)] = \sigma_p^2 \min[i, j]$. Signal-to-noise ratio (SNR) per bit is defined as $\gamma_b = E[m(k)^2] / N_0 \log_2 M$. In an intradyne receiver, $E_r(t)$ is downconverted to an IF of Δf as seen in Eq. (2). Optical frequency bands around $f_{LO} + \Delta f$ and $f_{LO} - \Delta f$ map to the same IF, where f_{LO} is the LO frequency. In order to avoid crosstalk in a dense wavelength-division-multiplexing system and to avoid excess optical noise from unwanted image bands, an optical filter of a single channel bandwidth is required before the intradyne receiver.

3. Carrier estimation and data detection

In this paper, we consider the detection of an uncoded data symbol sequence $\{m(k)\}$ transmitted over an AWGN channel with unknown carrier phase and frequency offset, as modeled in Eq. (2). Because the data is uncoded, the symbols of $\{m(k)\}$ are independent, and each symbol can assume with equal probability any point S_i , $i = 0, \dots, M-1$, in the signal constellation. The elements of $\{\theta(k)\}$ are correlated, whereas the assumption of an AWGN channel with no ISI makes the elements of $\{n(k)\}$ independent. The elements of $\{r(k)\}$ are rendered independent when conditioned on given values of uncoded data symbol sequence $\{m(k)\}$, carrier phase process $\{\theta(k)\}$, and frequency offset, Δf . Hence, each data symbol $m(k)$ in the sequence $\{r(k)\}$ will be detected individually, i.e., symbol-by-symbol with minimum symbol error probability. We assume mutual statistical independence among $\{m(k)\}$, $\{\theta(k)\}$, and Δf , which leads to the separation of the CE problem from the DD problem [10]. Therefore, at high SNR, the optimum symbol-by-symbol receiver structure consists of a CE followed by a coherent DD, as illustrated in Fig. 1 [10].

We develop our new CE algorithm in Subsection 3.1 for carrier phase and frequency offset estimation. The estimated phase and frequency are treated as if they were the true values of the carrier and are used in (partially) coherent DD in Subsection 3.2. Separate symbol-by-symbol receiver structures, consisting of CE and DD, can be applied independently on each polarization channel as shown in Fig. 1. All equations and quantities expressed hereafter are thus meant for one polarization channel.

3.1 Complex-weighted DA-ML carrier estimation

In DA-ML CE [3, 4], the carrier phase is estimated using a complex reference phasor (RP), where the argument of the RP is the ML phase estimate of the carrier. It is assumed that the carrier phase process $\{\theta(k)\}$ is fluctuating slowly compared to the symbol rate such that it can be approximated to be piecewise constant over intervals longer than LT , where L is the estimator filter length. The RP $V(k+1)$ without considering any frequency offset is computed in [3, 4] using the immediate past L received signals as

$$V(k+1) = C(k) \sum_{l=1}^L r(k-l+1) \hat{n}^*(k-l+1) \quad (3)$$

where $\hat{m}(k)$ is the decision on k th symbol from DD and $C^{-1}(k) = \sum_{l=1}^L |\hat{m}(k-l+1)|^2$ is a factor to normalize the RP's magnitude in the event of a non-constant-energy signal constellation.

Considering the presence of an unknown carrier frequency offset, we propose here a new RP $V'(k+1)$ that is formed by

$$V'(k+1) = C(k) \sum_{l=1}^L w_l(k) r(k-l+1) \hat{m}^*(k-l+1) \quad (4)$$

where each $w_l(k)$ is a complex weight. Momentarily ignoring the carrier phase $\theta(k)$ and AWGN $n(k)$, the filter input terms with a carrier frequency offset appear as $\{r(k-l+1) \hat{m}^*(k-l+1) = |m(k-l+1)|^2 e^{j\Delta\omega(k-l+1)}, 1 \leq l \leq L\}$ where we have assumed $\hat{m}(k) = m(k)$, i.e., all receiver decisions are correct, and $\Delta\omega = 2\pi\Delta fT$ is the angular frequency offset. The complex weight $w_l(k)$ is designed to rotate $r(k-l+1) \hat{m}^*(k-l+1)$ by $\Delta\omega l$ such that all filter input terms are aligned with identical angular frequency offset of $\Delta\omega(k+1)$. This leads to the RP $V'(k+1)$ being angular-frequency-offset aligned to the next received signal $r(k+1) = m(k+1)e^{j\Delta\omega(k+1)}$, which enables an enhanced signal demodulation. The filter in Eq. (4) operates without the knowledge of SNR, carrier phase, and frequency offset statistics.

Ideally, one should have $w_l(k) = e^{j\Delta\omega l}$, but $\Delta\omega$ is unknown in practice. Thus, we propose to choose the complex weights automatically and adaptively at each time k based on the observations $\{r(l), 0 \leq l \leq k\}$ to minimize the cost function $J(k)$ given by [11]

$$\begin{aligned} J(k) &= \sum_{l=1}^k |e(l)|^2 \\ &= \sum_{l=1}^k \left| \frac{r(l)}{\hat{m}(l)} - C(l-1) \mathbf{w}^T(k) \mathbf{y}(l-1) \right|^2 \end{aligned} \quad (5)$$

where $\mathbf{w}(k) = [w_1(k) \ \cdots \ w_L(k)]^T$ is the L -by-1 filter-weight vector and $\mathbf{y}(k) = [r(k) \hat{m}^*(k) \ \cdots \ r(k-L+1) \hat{m}^*(k-L+1)]^T$ is the L -by-1 filter-input vector at time k . The error $e(l)$ is the difference between the desired response $r(l)/\hat{m}(l)$ and the RP output of Eq. (4) at time $l-1$ using the latest set of filter coefficients $\mathbf{w}(k)$. Minimization of the cost function $J(k)$ forces RP $V'(l)$ to track the normalized term $r(l)/\hat{m}(l)$, and thus forces $w_l(k)$ to track $e^{j\Delta\omega l}$. Adaptation of filter weights using a least-squares criterion, as opposed to a mean-square-error criterion $\tilde{J}(k) = \mathbb{E}[|e(k)|^2]$, requires no statistical information about the AWGN, carrier phase, or frequency offset.

Since $J(k)$ is a real valued function of $\mathbf{w}(k)$ and $\mathbf{w}^*(k)$, we solve the equation: $\partial J(k)/\partial \mathbf{w}^*(k) = 0$, for $\mathbf{w}(k)$ using the complex vector differentiation identity:

$\partial[a - \mathbf{w}^H \mathbf{b}] / \partial \mathbf{w}^* = -\mathbf{b}$, where a is a scalar and \mathbf{b} is a vector. The optimum $\hat{\mathbf{w}}(k)$ is obtained as the solution of a least-squares normal equation

$$\hat{\mathbf{w}}(k) = \Phi^{-1}(k) \mathbf{z}(k), \quad k \geq 1 \quad (6)$$

where $\Phi(k) = \sum_{l=1}^k C^2(l-1) \mathbf{y}^*(l-1) \mathbf{y}^T(l-1)$ is the L -by- L time-average autocorrelation matrix and $\mathbf{z}(k) = \sum_{l=1}^k C(l-1) \mathbf{y}^*(l-1) r(l) / \hat{m}(l)$ is the L -by-1 time-average cross-correlation vector. The $\Phi^{-1}(k)$ matrix and the least-squares $\hat{\mathbf{w}}(k)$ at each time k can be obtained recursively using the matrix inversion lemma [12]. Therefore, CW-DA-ML CE avoids any matrix inversion. The optimum $\hat{\mathbf{w}}(k)$ can respond to changing channel conditions, as it depends on the observations $\{r(k)\}$.

3.2 Symbol-by-symbol data detection

Treating $V'(k)$ as the true phasor $e^{j(\Delta\omega k + \theta(k))}$, we multiply $r(k)$ by $V'^*(k)$ to compensate for the carrier phase and frequency offset. The derotated $r(k)$ is plugged into the ML minimum-distance DD for an AWGN channel given by [12]

$$\hat{m}(k) = \arg \min_{0 \leq i \leq M-1} |r(k) V'^*(k) - S_i| \quad (7)$$

Assuming unconstrained symbol energy, the DD is simplified by expanding the equation, dropping terms that do not depend on the trial data symbol S_i , and rearranging them to yield

$$\begin{aligned} \hat{m}(k) &= \arg \min_{0 \leq i \leq M-1} \left(|r(k) V'^*(k)|^2 + |S_i|^2 - \operatorname{Re} [2r(k) V'^*(k) S_i^*] \right) \\ &= \arg \min_{0 \leq i \leq M-1} \left(|S_i|^2 - \operatorname{Re} [2r(k) V'^*(k) S_i^*] \right) \\ &= \arg \max_{0 \leq i \leq M-1} \operatorname{Re} \left[r(k) V'^*(k) S_i^* - \frac{1}{2} |S_i|^2 \right] \quad (8) \end{aligned}$$

Finally, our symbol-by-symbol receiver performs (partially) coherent DD by forming the decision statistic as in Eq. (8).

Our symbol-by-symbol receiver algorithm employing the novel CW-DA-ML CE with recursive updating of $\Phi^{-1}(k)$ matrix is outlined in Table 1. The $\Phi^{-1}(k)$ matrix is Hermitian symmetric, thus only the upper triangle of $\Phi^{-1}(k)$ needs to be computed whereas the lower triangle is obtained through diagonal reflection as signified by the $\operatorname{Tri}\{\cdot\}$ operator in Table 1. In operating the filter of Eq. (4), we set $V'(0) = 1$ and $\hat{\mathbf{w}}(0) = [1 \ 0 \ \dots \ 0]^T$ to give a maximum gain of one on the first filter input $r(0) \hat{m}^*(0)$. An initial preamble of N known symbols is used to enable $\hat{\mathbf{w}}(k)$ to settle to a steady state and for $V'(k)$ to acquire tracking of the phasor $e^{j(\Delta\omega k + \theta(k))}$. The filter operates in decision-directed mode subsequently.

Table 1. Symbol-by-symbol receiver employing CW-DA-ML CE

Initialize recursive algorithm at time $k = 0$

i. $V'(0) = 1$

ii. $\hat{\mathbf{w}}(0) = [1 \ 0 \ \dots \ 0]^T$

For each iteration, $k = 0, 1, 2, \dots$

1. Receiver decision

$$\hat{m}(k) = \arg \max_{0 \leq s \leq M-1} \text{Re} \left[r(k) V'^*(k) S_s^* - |S_s|^2 / 2 \right]$$
2. Preamble sequence, for $0 \leq k \leq N-1$

$$\hat{m}(k) = m(k)$$
3. Form filter input

$$\mathbf{y}(k) = [r(k) \hat{m}^*(k) \ \dots \ r(k-L+1) \hat{m}^*(k-L+1)]^T$$
4. Compute intermediate vector, $\boldsymbol{\psi}(k)$, for $k \geq 1$

$$\boldsymbol{\psi}(k) = C(k-1) \boldsymbol{\Phi}^{-1}(k-1) \mathbf{y}^*(k-1)$$
5. Compute gain vector, $\mathbf{g}(k)$, for $k \geq 1$

$$\mathbf{g}(k) = \frac{\boldsymbol{\psi}(k)}{1 + C(k-1) \mathbf{y}^T(k-1) \boldsymbol{\psi}(k)}$$
6. Compute *a priori* estimation error, $\xi(k)$, for $k \geq 1$

$$\xi(k) = \frac{r(k)}{\hat{m}(k)} - V'(k)$$
7. Recursively update tap-weight vector, $\hat{\mathbf{w}}(k)$, for $k \geq 1$

$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}(k-1) + \mathbf{g}(k) \xi(k)$$
8. Recursively update inverse correlation matrix, $\boldsymbol{\Phi}^{-1}(k)$, for $k \geq 1$

$$\boldsymbol{\Phi}^{-1}(k) = \text{Tri} \left\{ \boldsymbol{\Phi}^{-1}(k-1) - \mathbf{g}(k) \boldsymbol{\psi}^H(k) \right\}$$
9. Compute next reference phasor

$$V'(k+1) = C(k) \hat{\mathbf{w}}^T(k) \mathbf{y}(k)$$

4. Numerical results and discussion

In all subsequent Monte-Carlo simulations, a single-polarization received signal modeled after Eq. (2) is used with a fixed bit rate of 50 Gbit/s for all modulation formats. A filter length of $L = 5$ is maintained for all CE algorithms tested. The modulo- 2π reduced $\Delta\omega$ is assumed to have a probability density function given by $p(\Delta\omega) = 1/2\pi$ for $\Delta\omega \in [-\pi, \pi)$, where $\Delta\omega$ is random but time invariant. The models of the carrier phase, $\theta(k)$, and angular frequency offset, $\Delta\omega$, are not known to the receiver.

4.1 Operation of CW-DA-ML CE

In this subsection, we analyze the statistical performance of CW-DA-ML CE by simulating carrier recovery for a 4-PSK signal with a preamble of $N = 50$ and a laser linewidth of 2 MHz.

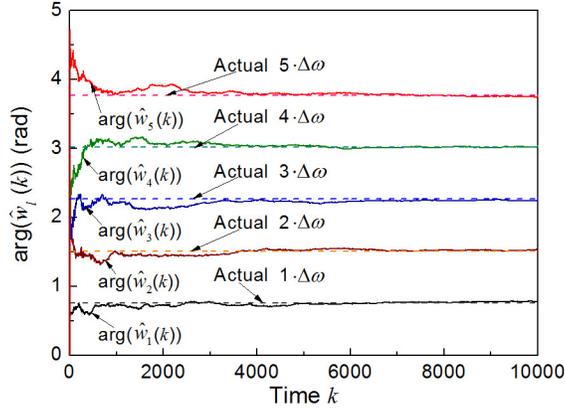


Fig. 2. Adaptation of $\arg(\hat{w}_l(k))$ from Eq. (6) in one sample simulation run.

Figure 2 plots the acquisition of the argument of the complex weights, $\{\arg(\hat{w}_l(k)), 1 \leq l \leq L\}$, from Eq. (6) at an γ_b of 10 dB and a frequency offset of 3 GHz. The $\arg(\hat{w}_l(k))$ converges quickly to the actual value of $l\Delta\omega$. Hence, $w_l(k)$ in Eq. (4) effectively removes the angular frequency offset difference between consecutive filter input terms $\{r(k-l+1)\hat{m}^*(k-l+1), 1 \leq l \leq L\}$ and helps to produce a more accurate RP compared to that produced by Eq. (3).

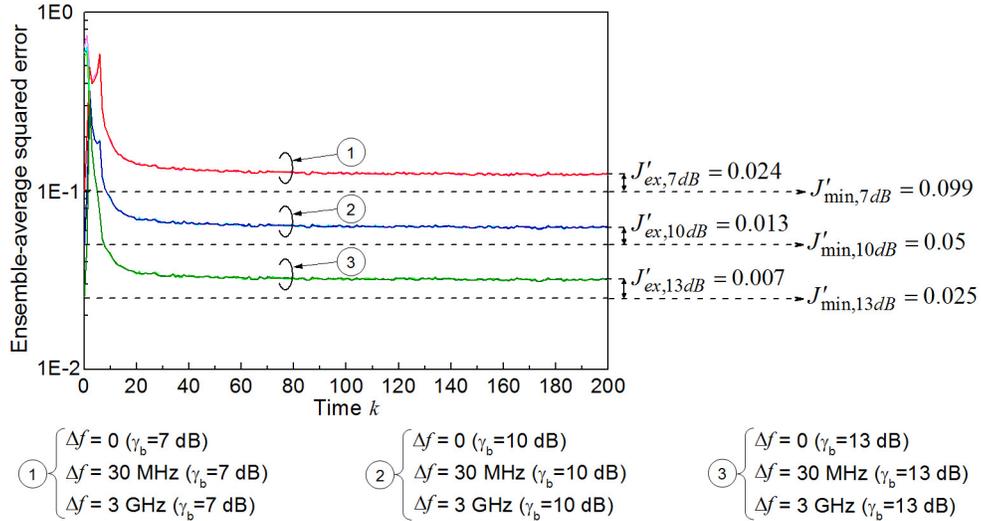


Fig. 3. Ensemble-average squared error curves with three different Δf and SNR values.

The mean-square error (MSE) learning curve of CW-DA-ML CE algorithm given by $J'(k) = E\left[\left|r(k)/\hat{m}(k) - V'(k)\right|^2\right]$ is empirically evaluated. The ensemble-average squared error curves averaged over 10^4 runs at frequency offsets of 0, 0.03, and 3 GHz are plotted for SNR values of 7, 10, and 13 dB in Fig. 3. Ideal decision feedback is assumed, i.e., $\hat{m}(k) = m(k)$ for all k . Using Eq. (2), the MSE $J'(k)$ can be expanded into

$$J'(k) = \mathbb{E} \left[\left| \frac{n(k)}{m(k)} \right|^2 + \left| e^{j(\Delta\omega k + \theta(k))} - V'(k) \right|^2 + \left(e^{j(\Delta\omega k + \theta(k))} - V'(k) \right) \left(\frac{n(k)}{m(k)} \right)^* + \left(e^{j(\Delta\omega k + \theta(k))} - V'(k) \right)^* \left(\frac{n(k)}{m(k)} \right) \right]. \quad (9)$$

The $V'(k)$ is independent of $n(k)$ and $m(k)$, as $V'(k)$ depends only on $\{n(l), m(l), l < k\}$. Moreover, Δf , $\theta(k)$, $n(k)$, and $m(k)$ are mutually independent. Thus, the third and fourth terms on the right hand side of Eq. (9) amount to zero because $\mathbb{E}[n(k)/m(k)] = 0$. Equation (9) then reduces to $J'(k) = J'_{\min}(k) + J'_{\text{ex}}(k)$, where $J'_{\min}(k) = \mathbb{E} \left[\left| n(k)/m(k) \right|^2 \right]$ is the minimum achievable MSE and $J'_{\text{ex}}(k) = \mathbb{E} \left[\left| e^{j(\Delta\omega k + \theta(k))} - V'(k) \right|^2 \right]$ is the excess MSE due to the MSE of the RP. It can be shown that $J'_{\min} = 1/\gamma_b \log_2 M$ for 4-PSK signals. Figure 3 shows the values of J'_{\min} , and also J'_{ex} which was obtained by averaging from time $k = 100$ to $k = 200$ the average of $\left| e^{j(\Delta\omega k + \theta(k))} - V'(k) \right|^2$ over 10^4 runs at each time point k . A J'_{ex} value less than 2.4×10^{-2} demonstrates the good tracking of $e^{j(\Delta\omega k + \theta(k))}$ by $V'(k)$, for SNR values above 7 dB. The excess MSE J'_{ex} decreases with increasing SNR but is indifferent to varying frequency offset, attesting that FOE by CW-DA-ML CE algorithm is unbiased to the frequency offset present.

CW-DA-ML CE algorithm converges quickly yielding rapid carrier phase and frequency tracking. We note that a short preamble of approximately twice the filter length, $N \approx 2L$, is sufficient to aid the complex weight $\hat{w}_i(k)$ in acquiring tracking of $e^{j\Delta\omega l}$, thus keeping the training overhead cost low. The MSE of CW-DA-ML CE for 4-PSK signals at SNR values above 7 dB settles to a steady state, implying acquisition of the frequency offset, within 100 observed samples. Whereas the differential FOE for 4-PSK signals in [7] used a minimum observation sample size of 500 for acquisition of the frequency offset. Hence, CW-DA-ML FOE is more than 5 times faster than that of the differential FOE of [7] for 4-PSK signals at SNR values above 7 dB.

4.2 Bit-error rate performance with CW-DA-ML CE

Bit-error rate (BER) curves of block M th power [2], DA-ML [3, 4], and CW-DA-ML CE are plotted in Fig. 4 for 4- and 8-PSK, as a function of SNR per bit, γ_b , in the presence of different frequency offsets. Filter length and laser linewidth was set to 5 and 0, respectively. A preamble of $N = 10$ known symbols is used for DA-ML and CW-DA-ML CE. DE according to [8] is employed to resolve phase ambiguity in block M th power CE, and to counter runaways due to decision errors in DA-ML and CW-DA-ML CE.

It is evident that block M th power and DA-ML CE are intolerant to frequency offsets. For instance, in block M th power CE of 4-PSK signals, as the frequency offset increases from 0 to 100 and 200 MHz, the γ_b penalty rises from 1.13 to 1.33 and 1.80 dB, respectively. Likewise for DA-ML CE, the γ_b penalty rises from 0.89 to 1.30 and 2.14 dB. SNR per bit penalties are referenced to the perfect coherent (ASE noise limited) receiver at $\text{BER} = 10^{-4}$. Moreover, block M th power and DA-ML CE become increasingly sensitive to frequency offset with increasing modulation order. For a fixed frequency offset of 100 MHz, as modulation format

increases from 4- to 8-PSK, block M th power CE incurs an increasing γ_b penalty of 1.33 and 3.18 dB, respectively. Likewise, DA-ML CE incurs an increasing γ_b penalty of 1.30 and 3.63 dB. Frequency offset intolerance of block M th power and DA-ML CE originates from the violation of their assumption that $\{r(l), k - L + 1 \leq l \leq k\}$ have identical angular frequency offsets.

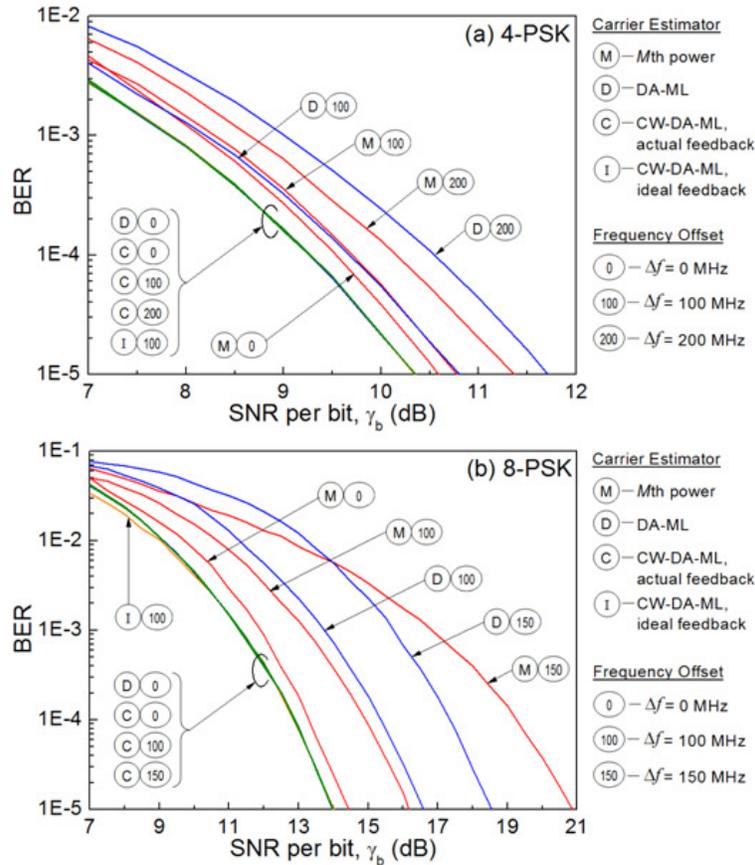


Fig. 4. BER curves of block M th power, DA-ML, and CW-DA-ML CE with different frequency offsets for (a) 4-PSK, and (b) 8-PSK.

On the contrary, the γ_b penalty for CW-DA-ML CE remains unchanged at 0.89 and 1.13 dB regardless of the frequency offset present in 4- and 8-PSK, respectively, at $\text{BER} = 10^{-4}$. Regardless of Δf , BER of CW-DA-ML CE over all tested SNR values closely replicate that of DA-ML CE at zero frequency offset. Hence, CW-DA-ML CE approaches ideal FOE, without requiring any statistical knowledge of AWGN, carrier phase, or frequency offset.

Figure 4 also shows CW-DA-ML CE at $\Delta f = 100$ MHz with ideal decision feedback, i.e., $\hat{m}(k) = m(k)$ for all k . In the range of SNR shown, the performance difference between actual and ideal decision feedback is not noticeable for 4-PSK, and is slightly more pronounced at the lower SNR region for 8-PSK. The effect of decision errors on the receiver performance becomes negligible for BER's below 10^{-3} .

4.3 FOE range of CW-DA-ML CE

SNR per bit, γ_b , penalty of block M th power CE [2], DA-ML CE [3, 4], differential FOE [7], and CW-DA-ML CE are plotted in Fig. 5 for 4- and 8-PSK, as frequency offset is varied between $\pm R/2$, in the presence of different laser linewidths, $\Delta\nu$. Any integer-multiple-of- R frequency offset, Δf , translates into the same modulo- 2π reduced $\Delta\omega$, thus having the same rotational effect on $r(k)$. Hence, it is sufficient to consider only the primary modulo- R reduced frequency offset range $\pm R/2$. The differential FOE utilizes a subsequent block M th power CE to correct any residual carrier phase, DE to overcome its R/M frequency estimate ambiguity, and a sample size of 500 to acquire the frequency offset, following that of [7].

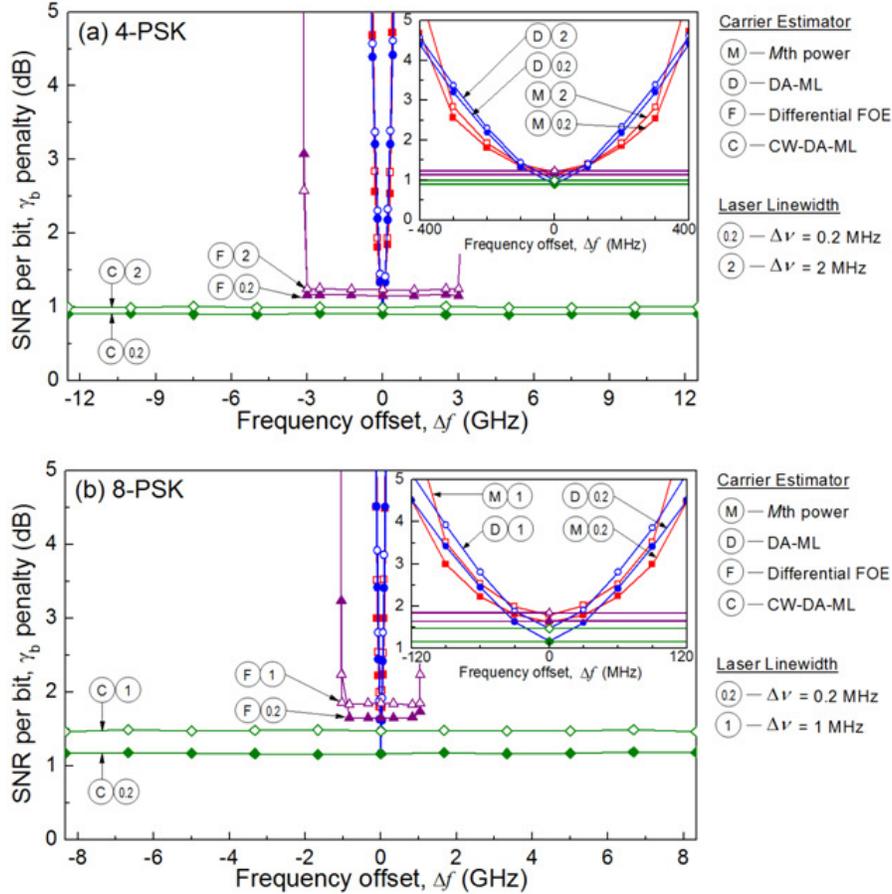


Fig. 5. FOE range of block M th power CE, DA-ML CE, differential FOE, and CW-DA-ML CE for (a) 4-PSK, and (b) 8-PSK. Insets show enlarged sections of the frequency offset axis.

CE techniques, such as block M th power and differential FOE, which erase data modulation by raising the received signal to the M th power, and extract the phase using the $\arctan(\cdot)$ operator are inherently disadvantaged in three ways. First, they are restricted to MPSK signals. Second, their FOE range is limited, modulation-format dependent, and decreases with increasing modulation order. For a fixed bit rate of R_b , block M th power CE has a maximum allowable frequency offset of $\pm R_b / 2LM \log_2(M)$ in the absence of carrier phase variations [7], and the differential FOE has an FOE range of $\pm R_b / 2M \log_2(M)$ [13]. Third, they require phase unwrapping in estimating the carrier phase and frequency offset,

which is difficult to perform accurately at low SNR and is prone to the highly nonlinear cycle slipping phenomenon [14].

On the other hand, our CW-DA-ML CE achieves a complete modulo- R reduced, modulation-format-independent, FOE range of $\pm R/2$ as witnessed in Fig. 5. This is attributed to the use of the RP, $V'(k)$, having an unambiguous phase tracking range of $[0, 2\pi)$. The use of the RP also eliminates the need for phase unwrapping in carrier phase and frequency offset estimation. The circular- 2π nature of the RP results in frequency-estimate duplicity of R , but this duplicity does not call for pilot assistance or DE to ensure correct DD. From Fig. 5(a), it is notable that a penalty of only approximately 1-dB is incurred at BER = 10^{-4} for a 50 Gbit/s 4-PSK signal having a laser linewidth of 2 MHz, regardless of the frequency offset present.

4.4 Complexity and parallel implementation of CW-DA-ML CE

The computational load, defined as the number of operations required to estimate $e^{j(\Delta\omega k + \theta(k))}$ per symbol, of the CW-DA-ML CE algorithm presented in Table 1 is $O(L^2)$ real multiplications and additions. However, by implementing Eq. (4) in a lattice structure, which is not presented here for brevity, the complexity reduces to simply $O(L)$ real multiplications and additions [12]. Differential FOE (inclusive of subsequent block M th power CE) requires $O(\log_2 M)$ real multiplications and additions, and $(2L + 1)/L$ accesses to a read-only-memory to map $\arctan(\cdot)$ operations. The $1/L$ factor arises from sharing of the common $\arctan(\cdot)$ operation among L symbols during the block M th power processing. The higher complexity of CW-DA-ML CE compared to differential FOE is traded off with CW-DA-ML CE's wider and faster frequency estimation, and its applicability to arbitrary modulation formats. The update of $w(k)$ need be performed only until the frequency offset is acquired, say, until the 100th iteration for 4-PSK signals according to Fig. 3. Thereafter, the same weight vector $w(k)$ can be used in Eq. (4) at each time point k , i.e., skip steps 4 to 8 in Table 1. Then, the complexity of CW-DA-ML CE becomes simply $4L + 6$ and $4L$ real multiplications and additions, respectively, per symbol.

Our CW-DA-ML CE is a D -symbol forward linear predictor, where the filter predicts $e^{j(\Delta\omega(k+D) + \theta(k+D))}$ as a weighted linear combination of past observed samples $\{r(l), l \leq k\}$ and symbol decisions $\{\hat{m}(l), l \leq k\}$ which are fed back. All of the above simulations used an ideal feedback delay of $D = 1$, which is the minimum in a feedback system. Practical coherent optical receivers supporting 100 Gb/s or higher data rates use parallel and pipelined processing to overcome the limited complementary metal-oxide-semiconductor processor speed of several Gb/s. In parallel processing, the effect of the observed sample on the feedback value is delayed by $D > 1$ [15]. Parallel implementation of CW-DA-ML CE, with a constant frequency offset, zero laser phase noise, and a feedback delay of D -symbols, is expected to incur marginal performance deterioration as it can estimate D -symbols ahead the linearly increasing angular frequency offset $\Delta\omega(k + D)$. However, in the additional presence of laser phase noise, performance penalty originating from feedback delay is expected as pointed out in [15]. This is due to the difficulty in estimating the accumulated Gaussian random-walk of the carrier phase over the previous D -symbols, for which the estimated carrier phase information is unavailable because of feedback delay. A detailed implementation and performance analysis of CW-DA-ML CE in parallel processing is left as part of our future work.

5. Conclusions

If the frequency-offset-symbol-duration product is small, separate frequency offset estimation (FOE) may not be necessary as block M th power and decision-aided, maximum-likelihood

carrier estimation (DA-ML CE) can cope with small frequency offsets. To combat larger frequency offsets, an FOE is necessary. Hence, we proposed a symbol-by-symbol receiver employing a novel complex-weighted (CW) DA-ML CE for optical channels with both carrier phase and frequency offset uncertainty. A near-ideal FOE over a complete modulo- R reduced frequency range of $\pm R/2$ (R = symbol rate) is demonstrated such that the bit-error rate performance is only limited by the carrier phase variance. CW-DA-ML CE is applicable to both M -ary phase-shift keying (MPSK) and quadrature amplitude modulation formats. The FOE for 4-PSK signals is more than 5 times faster than that of the differential FOE at signal-to-noise ratio values above 7 dB.

Acknowledgments

The financial support of the Singapore MoE AcRF Tier 2 Grant MOE2010-T2-1-101 is gratefully acknowledged.