

Fundamentals of Coherent Optical Fiber Communications

Kazuro Kikuchi *Fellow, IEEE, OSA*

(Invited Tutorial)

Abstract—The recently developed digital coherent receiver enables us to employ a variety of spectrally efficient modulation formats such as M -ary phase-shift keying (PSK) and quadrature-amplitude modulation (QAM). Moreover, in the digital domain, we can equalize all linear transmission impairments such as group-velocity dispersion (GVD) and polarization-mode dispersion (PMD) of transmission fibers, because coherent detection preserves the phase information of the optical signal. This paper reviews the history of research and development related to coherent optical communications and describes the principle of coherent detection, including its quantum-noise characteristics. In addition, it discusses the role of digital signal processing (DSP) in mitigating linear transmission impairments, estimating the carrier phase, and tracking the state of polarization of the signal in coherent receivers.

Index Terms—coherent optical communications, optical fibers, optical receivers.

I. INTRODUCTION

Coherent optical fiber communications were studied extensively in the 1980s mainly because of the high sensitivity of coherent receivers that could enhance the unrepeated transmission distance [1]; however, since 1990, related research and development had been interrupted by the rapid advances in high-capacity wavelength-division multiplexed (WDM) systems, which employed the conventional intensity-modulation and direct-detection (IM-DD) scheme as well as the newly developed erbium-doped fiber amplifier (EDFA) to compensate for span loss of the link.

In 2005, the demonstration of digital carrier-phase estimation in coherent receivers led to the renewal of widespread interest in coherent optical communications [2]. The reason is that the digital coherent receiver enables us to employ a variety of spectrally efficient modulation formats such as M -ary phase-shift keying (PSK) and quadrature-amplitude modulation (QAM), relying upon stable carrier-phase estimation in the digital domain. In addition, because the phase information is preserved after detection, we can equalize linear transmission impairments such as group-velocity dispersion (GVD) and polarization-mode dispersion (PMD) of transmission fibers via digital signal processing (DSP). These advantages of the born-again coherent receiver afford considerable potential for innovating existing optical communication systems. Recently,

This work was supported in part by a Grant-in-Aid for Scientific Research (A) (25249038) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

K. Kikuchi is with the Department of Electrical Engineering and Information Systems, the University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8656, Japan (e-mail: kikuchi@ginjo.t.u-tokyo.ac.jp).

Manuscript received June 1, 2015; revised ****.

100-Gbit/s transmission systems, which employ quadrature PSK (QPSK) modulation, polarization-division multiplexing, and phase-diversity homodyne detection assisted with high-speed DSP at a symbol rate of 25 Gbaud, have been developed and introduced into commercial networks [3]. By placing WDM channels at 50-GHz-spaced grids, we can transmit a total capacity of up to 8.8 Tbit/s through a single fiber. Worldwide efforts are now underway to develop coherent receivers that can handle a bit rate of over 400 Gbit/s per WDM channel.

This paper reviews the history of coherent optical communications, the principle of coherent receivers, quantum-mechanical characteristics of coherent detection, multi-level modulation formats, and the newly developed digital coherent receiver, including the details of DSP algorithms. In other words, this paper focuses on the fundamental aspects of the coherent optical communications rather than introducing their cutting-edge technologies [4].

The remainder of this paper is organized as follows. Section II briefly reviews the over 30 years' history of coherent optical communications and the impact of DSP technology. Section III describes the principle of operation of the coherent receiver. Section IV discusses the quantum-mechanical properties of coherent detection. The difference between phase-sensitive and phase-insensitive coherent systems is emphasized from the viewpoint of Heisenberg's uncertainty relation. Next, Sec. V discusses optical modulators and multi-level modulation formats. Then, Sec. VI provides the details of DSP used in coherent receivers. We describe analog-to-digital conversion, adaptive equalizers based on finite-impulse-response (FIR) filters, algorithms for filter-tap adaptation, and the method of carrier-phase estimation. Finally, Section VII concludes the paper.

II. HISTORY OF COHERENT OPTICAL COMMUNICATIONS

This section briefly reviews the history of coherent optical communications. The technical details will be presented in the following sections.

A. Coherent Optical Communications in the Analog Era

Research and development related to optical fiber communication systems was initiated in early 1970s. Such systems used intensity modulation of semiconductor lasers, and the intensity of the optical signal transmitted through an optical fiber was detected by a photodiode, which acted as a square-law detector. This combination of transmitter and receiver is known as the intensity-modulation and direct-detection

(IM-DD) scheme, which is widely employed in optical communication even today. The IM-DD scheme has a significant advantage in that the receiver sensitivity is dependent neither on the carrier phase nor on the state of polarization (SOP) of the incoming signal, which randomly fluctuate in real systems.

On the other hand, a receiver, in which the signal is interfered with a local oscillator (LO) so as to extract the phase information of the signal, is called a coherent receiver; heterodyne and homodyne receivers are included in this category. Although coherent optical communications using heterodyne detection was first proposed by DeLange in 1970 [5], such a proposal did not attract much attention because the IM-DD scheme became mainstream in optical fiber communication systems during the 1970s. However, after Okoshi and Kikuchi [6] and Fabre and LeGuen [7] independently demonstrated precise frequency stabilization of semiconductor lasers in 1980, with the object of heterodyne detection for optical fiber communications, many demonstrations of coherent optical communication systems were reported until around 1990.

Table I compares coherent and IM-DD schemes. In the case of coherent receivers, we can restore full information on optical carriers, namely, in-phase and quadrature (IQ) components (or amplitude and phase) of the complex amplitude of the optical electric field and the state of polarization (SOP) of the signal. In order to gain such a significant advantage, coherent receivers are highly sensitive to random variations in the phase and SOP of the incoming signal. To deal with this problem, the configuration of coherent systems becomes much more complicated than that of IM-DD systems.

TABLE I
COMPARISON BETWEEN COHERENT AND IM-DD SCHEMES.

	Coherent	IM-DD
Modulation parameters	I & Q or Amplitude & Phase	Intensity
Detection method	Heterodyne or Homodyne detection	Direct detection
Sensitivity to carrier phase	Yes	No
Sensitivity to polarization	Yes	No

The applicability of semiconductor lasers to coherent systems was first investigated in the 1980s. The center frequency drift of semiconductor lasers for a transmitter and an LO could be maintained below 10 MHz, as shown in [6], [7]. Even when the frequency drift of the transmitter laser was suppressed, the carrier phase fluctuated randomly owing to the large phase noise of the semiconductor lasers. Therefore, narrowing the spectral linewidths of the semiconductor lasers was the crucial requirement for realizing stable heterodyne detection. A method for measuring laser linewidths, known as the delayed self-heterodyne method, was invented [8], and the spectral property of semiconductor lasers was studied extensively [9]–[12]. It was found that the linewidth of GaAlAs lasers was typically in the range of 10 MHz. Such a narrow linewidth together with the precisely controlled center frequency accelerated studies of coherent optical communication systems based on semiconductor lasers.

The SOP-dependence of the receiver sensitivity was over-

come by the polarization-diversity technique [13], [14]. Each polarization component of the transmitted signal was detected by orthogonally polarized LO lasers, and post processing of the detected signal could achieve SOP-independent receiver sensitivity. Current digital coherent receivers employ this polarization-diversity scheme. On the other hand, a linear polarization-maintaining fiber [15] and a circular polarization-maintaining fiber [16] were studied to stabilize the SOP of the transmitted signal, but they were not introduced into practical transmission systems. However, the former, known as the PANDA fiber, has been commercialized and applied to fiber pigtailed, fiber lasers, and optical sensing systems. A twisted fiber maintaining circular polarization has also been investigated for applications to all-optical signal processing [17].

Analyses of the receiver sensitivity of various modulation/demodulation schemes were performed [18], [19]. It was found that quantum-noise-limited receiver sensitivity (also called shot-noise-limited receiver sensitivity) could be achieved by injecting a sufficiently high LO power into the coherent receiver to combat the circuit noise of the receiver. The motivation for research and development related to coherent optical communications at this stage stemmed from the high receiver sensitivity that enhanced the unrepeated transmission distance.

Many demonstrations of heterodyne systems were reported around 1990. In such systems, the frequency-shift keying (FSK) modulation format was most commonly employed because the semiconductor-laser frequency could easily be modulated by directly modulating the bias current of the lasers. Such frequency modulation was demodulated by differential detection at the IF stage. For example, a field trial of undersea transmission at 2.5 Gbit/s was reported in [20], where advanced heterodyne technologies, such as continuous-phase FSK (CPFSK) modulation, polarization diversity, automatic frequency control (AFC) of semiconductor lasers, and differential detection, were introduced.

Homodyne receivers were also investigated in the 1980s. The advantage of homodyne receivers was that the baseband signal was directly obtained, in contrast to heterodyne receivers, which required a rather high intermediate frequency (IF). On the other hand, they required the optical phase-locked loop (OPLL) so that the LO phase tracked the carrier phase of the signal [21]. One of the most difficult issues for the OPLL-type homodyne receiver is the recovery of the carrier phase. The optical signal modulated in the PSK format does not have the carrier component; therefore, to recover the carrier phase, some kind of nonlinear signal processing is necessary. Then, the phase error between the recovered carrier phase and the LO phase is led to the frequency-controlling terminal of the LO so that the LO phase tracks the carrier phase. The OPLL bandwidth was usually limited below 1 MHz because of large loop delay, and it was difficult to maintain system stability when semiconductor lasers had large phase noise and frequency drift. This technical difficulty inherent in OPLL has not been resolved perfectly, even with state-of-the-art distributed-feedback (DFB) semiconductor lasers.

In the 1990s, the invention of erbium-doped fiber amplifiers

(EDFAs) made the quantum-noise-limited receiver sensitivity of the coherent receiver less significant [22]. This is because the signal-to-noise (S/N) ratio of the signal transmitted through the amplifier chain is determined from the accumulated amplified spontaneous emission (ASE) rather than the quantum noise. In addition, even in unrepeated transmission systems, an EDFA used as a low-noise pre-amplifier reduced the need for a coherent receiver with superior sensitivity. The coherent receiver actually had advantages besides high receiver sensitivity. For example, it could deal with multi-level modulation formats and had post-processing functions such as compensation for GVD and PMD of fibers. However, such advantages were neither urgent requirements for the system nor cost-effective solutions in the 1990s.

Technical difficulties in coherent receivers had not been resolved at that time. The heterodyne receiver required an IF that should be much higher than the signal bit rate. On the other hand, the homodyne receiver was essentially a baseband receiver; however, complexity in the stable locking of the carrier phase drift had prevented its use in practical applications.

The above-mentioned factors led to the interruption of further research and development activities related to coherent optical communications for more than ten years. On the other hand, the EDFA-based IM-DD system started to exploit WDM techniques to increase the transmission capacity of a single fiber. The hardware required for WDM networks became widely deployed owing to its simplicity and the relatively low cost associated with optical amplifier repeaters, where multiple WDM channels could be amplified simultaneously. The WDM technique marked the beginning of a new era in the history of optical communication systems and brought forth a thousand-fold increase in the transmission capacity in the 1990s.

B. Coherent Optical Communications in the Digital Era

With the increase in the transmission capacity of WDM systems, coherent technologies attracted a renewal of widespread interest after 2000. The motivation is to develop methods for meeting the ever-increasing bandwidth demand with multi-level modulation formats based on coherent technologies [23]. The first step in the revival of coherent optical communications research was triggered by the QPSK modulation/demodulation experiment featuring optical IQ modulation and optical delay detection [24]. In such a scheme, we can double the bit rate while maintaining the symbol rate, and a 40-Gbit/s (20-Gbaud) differential QPSK (DQPSK) system has been put into practical use.

The next stage emerged with high-speed DSP. In the field of radio communications, digital techniques have been widely applied to transmitters and receivers. Figure 1 shows configurations of (a) the DSP-based RF transmitter and (b) the DSP-based RF receiver. At the transmitter, after appropriate DSP, digital data are converted into two-channel analog signals by digital-to-analog converters (DACs), which modulate IQ components of the RF carrier. On the other hand, at the receiver, the transmitted RF signal is mixed with the LO, and IQ components are obtained at the baseband. Such IQ signals

are transformed into the digital domain by analog-to-digital converters (ADCs), and the symbols are decoded via DSP. This type of receiver is called the direct-conversion receiver.

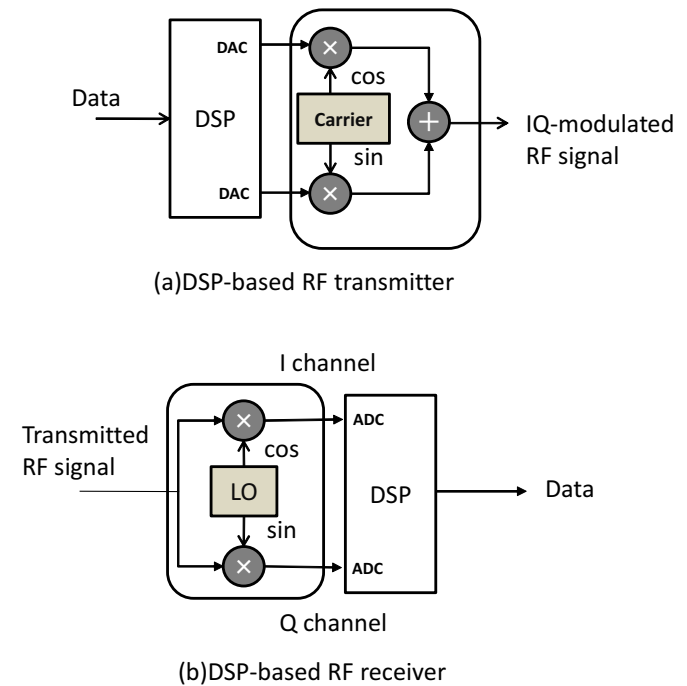


Fig. 1. Configurations of (a) the DSP-based RF transmitter and (b) the receiver. In (a), analog signals after DSP and DAC drive the RF modulator, and the IQ-modulated RF signal is generated. The receiver shown in (b) is called the direct-conversion receiver. The beat between the RF signal and the LO directly generates IQ components of complex amplitude, which are processed by ADC and DSP in the digital domain.

If the RF modulator in Fig. 1 (a) and the RF mixer in Fig. 1 (b) are replaced with optical counterparts, namely, the optical IQ modulator (Sec. V) [25] and the phase-diversity homodyne receiver (Sec. III) [2], respectively, we can easily conceive that the DSP-based optical transmitter and receiver should be those shown in Fig. 2. Such an optical transmitter and receiver were actually reported in [26] and [2], [27], [28], respectively.

This paper focuses on the coherent receiver. The DSP-based transmitter including functions such as Nyquist spectral shaping, signal pre-distortion/compensation, and forward error-correction (FEC) coding is commercially important but beyond the scope of this paper.

The recent development of high-speed digital integrated circuits has offered the possibility of processing high-speed electrical signals in a DSP core and retrieving the IQ components of the complex amplitude of the optical carrier from the homodyne-detected signal in a very stable manner. This concept appeared in [29], [30]; then, a 20-Gbit/s QPSK signal was demodulated with a phase-diversity homodyne receiver followed by digital carrier-phase estimation in [27], although BER measurements were still performed offline. Because the carrier phase is recovered after homodyne detection by means of DSP, this type of receiver is now commonly called the “digital coherent receiver.” Although an OPLL that locks the LO phase to the signal phase remains difficult to achieve owing to the loop delay problem, DSP circuits are become

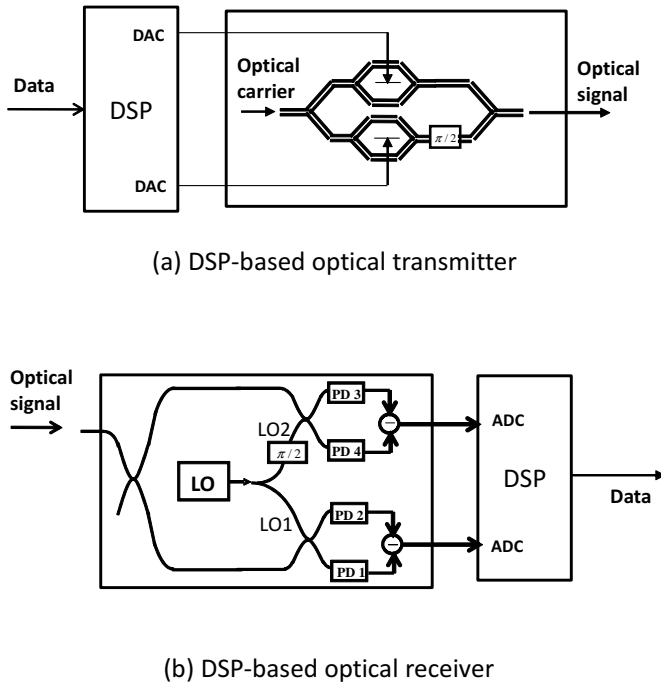


Fig. 2. Configurations of (a) the DSP-based optical transmitter and (b) the receiver. In (a), analog signals after DSP and DAC drive the optical IQ modulator. In (b), the outputs from the phase-diversity homodyne receiver are processed by DSP after ADC.

increasingly faster and provide us with simple and efficient means for estimating the carrier phase. Very fast tracking of the carrier phase improves system stability drastically as compared with the OPLL scheme.

Any type of multi-level modulation format can be introduced by using the coherent receiver [31], [32]. Whereas the spectral efficiency of binary modulation formats is limited to 1 bit/s/Hz/polarization, which is called the Nyquist limit, modulation formats with m_b bits of information per symbol can achieve a spectral efficiency of up to m_b bit/s/Hz/polarization. Although optical delay detection has been employed to demodulate the QPSK signal ($m_b = 2$), further increase in multiplicity can hardly be achieved with such a scheme.

Another important advantage of the digital coherent receiver is the post signal-processing function [33], [34]. The IQ demodulation by the digital coherent receiver is an entirely linear process; therefore, all the information on the complex amplitude of the transmitted optical signal is preserved even after detection, and optical signal processing functions acting on the optical carrier, such as optical filtering and dispersion compensation, can be performed at the electrical stage after detection. Electrical GVD compensation thus eliminates the need for the dispersion-compensation fiber (DCF).

Polarization alignment is also made possible after detection by introducing the polarization-diversity scheme into the homodyne receiver [35], [36]. The complex amplitude of the horizontal polarization and that of the vertical polarization are simultaneously measured and processed via DSP. Polarization demultiplexing and compensation for PMD have also been demonstrated with the digital coherent receiver [37], where

bulky and slow optical-polarization controllers as well as optical-delay lines are eliminated.

Because the digital coherent receiver requires high speed ADC and DSP, most of the experiments in the early years were conducted offline. Thus, after the transmitted data were stored in a computer, bit errors were analyzed offline. However, an application-specific integrated circuit (ASIC) designed for the 11.5-Gaud dual-polarization QPSK (DP-QPSK) signal was developed [38], and real-time operation of the digital coherent receiver at a bit rate of 46 Gbit/s was demonstrated by using such an ASIC [39]. This achievement is a milestone in the history of modern coherent optical communications. The combination of coherent detection and DSP has thus become a part of the next generation of optical communication systems and provides new capabilities that were not possible without the detection of the phase of the optical signal. Currently, an ASIC operating at 32 Gbaud has allowed for a single-channel transmission capacity of up to 127 Gbit/s including 27% overhead for FEC in commercial optical networks [40]. WDM signals are placed on 50-GHz-interval grids and the total capacity exceeds 8 Tbit/s.

III. PRINCIPLE OF COHERENT OPTICAL DETECTION

This section describes the basic operation principle of coherent optical detection [41]. We show how the coherent receiver measures the complex amplitude of the optical signal, using heterodyne detection, homodyne detection, and phase-diversity homodyne detection.

A. Coherent-Receiver Configuration

Figure 3 shows the configuration of the coherent optical receiver, where we assume that SOPs of the transmitted signal and the LO are aligned, controlling the polarization controller (PC). The fundamental concept underlying coherent detection is to take the product of electric fields of the modulated signal light and the continuous-wave (CW) LO. Let the complex electric field of the transmitted optical signal be

$$E_s(t) = A_s(t) \exp(i\omega_s t), \quad (1)$$

where $A_s(t)$ is the complex amplitude and ω_s is the angular frequency. This representation of the complex amplitude is common in the field of engineering. The real electric field is given by

$$\mathcal{E}_s(t) = \text{Re}\{E_s(t)\}, \quad (2)$$

where “Re” denotes the real part. Similarly, the complex electric field of the LO prepared at the receiver can be written as

$$E_\ell(t) = A_\ell \exp(i\omega_\ell t), \quad (3)$$

where A_ℓ is the constant complex amplitude and ω_ℓ is the angular frequency of the LO. The real electric field is given by

$$\mathcal{E}_\ell(t) = \text{Re}\{E_\ell(t)\}. \quad (4)$$

The complex amplitudes $A_{s,\ell}$ are related to the signal power P_s and the LO power P_ℓ as

$$P_s = k|A_s|^2/2, \quad (5)$$

$$P_\ell = k|A_\ell|^2/2. \quad (6)$$

Here, the constant k is given as

$$k = \frac{S_{eff}}{\zeta}, \quad (7)$$

where S_{eff} denotes the effective beam area and ζ is the impedance of the free space.

Balanced detection is usually introduced into the coherent receiver as a means to suppress the dc component and maximize the beat between the signal and the LO. The key is the use of a 3-dB optical coupler that adds a 180° phase shift to either the signal field or the LO field between the two output ports. When the signal and LO are co-polarized, the electric fields incident on the upper and lower photodiodes are given by

$$E_1 = \frac{1}{\sqrt{2}}(E_s + E_\ell), \quad (8)$$

$$E_2 = \frac{1}{\sqrt{2}}(E_s - E_\ell), \quad (9)$$

and the output photocurrents are written as

$$\begin{aligned} I_1(t) &= kR \left[\text{Re} \left\{ \frac{A_s(t) \exp(i\omega_s t) + A_\ell \exp(i\omega_\ell t)}{\sqrt{2}} \right\} \right]^{ms} \\ &= \frac{R}{2} [P_s(t) + P_\ell \\ &\quad + 2\sqrt{P_s(t)P_\ell} \cos\{\omega_{IF}t + \theta_{sig}(t) - \theta_\ell(t)\}], \quad (10) \end{aligned}$$

$$\begin{aligned} I_2(t) &= kR \left[\text{Re} \left\{ \frac{A_s(t) \exp(i\omega_s t) - A_\ell \exp(i\omega_\ell t)}{\sqrt{2}} \right\} \right]^{ms} \\ &= \frac{R}{2} [P_s(t) + P_\ell \\ &\quad - 2\sqrt{P_s(t)P_\ell} \cos\{\omega_{IF}t + \theta_{sig}(t) - \theta_\ell(t)\}], \quad (11) \end{aligned}$$

where “ms” denotes the mean square with respect to the optical frequencies, ω_{IF} is the IF given by $\omega_{IF} = |\omega_s - \omega_\ell|$, and $\theta_{sig}(t)$ and $\theta_\ell(t)$ are the phases of the transmitted signal and the LO, respectively. Further, R is the responsivity of the photodiode given by

$$R = \frac{q\eta}{\hbar\omega_s}, \quad (12)$$

where \hbar denotes Planck’s constant divided by 2π , q is the electron charge, and η is the quantum efficiency of the photodiode. The balanced detector output is then given by

$$\begin{aligned} I(t) &= I_1(t) - I_2(t) \\ &= 2R\sqrt{P_s(t)P_\ell} \\ &\quad \times \cos\{\omega_{IF}t + \theta_{sig}(t) - \theta_\ell(t)\}. \quad (13) \end{aligned}$$

The LO power P_ℓ is always constant and $\theta_\ell(t)$ includes only the phase noise that randomly varies with time. We find that only the beat between the signal and the LO is measured with balanced detection.

It should be emphasized that the photocurrent given by Eq. (13) is proportional to $\sqrt{P_\ell}$. Therefore, as the LO power increases, a gain for the photocurrent can be obtained; hence, the signal level overwhelms the thermal noise of the receiver circuit. In such a case, the S/N ratio of the received signal is determined from the intrinsic noise of the signal light and can reach the quantum-noise limit. Detailed analyses of quantum

noise associated with coherent detection will be given in Sec. IV.

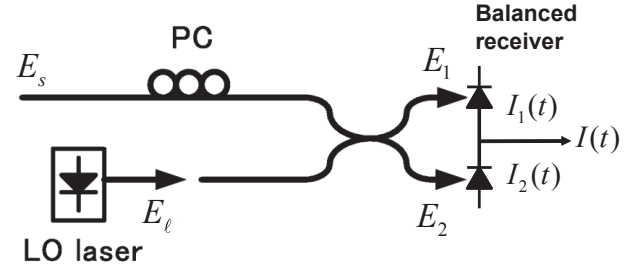


Fig. 3. Configuration of the coherent receiver that measures the beat between the signal and the LO. Balanced detection eliminates the dc component and maximizes the beat signal.

B. Heterodyne Detection

Heterodyne detection refers to the case that $|\omega_{IF}| \gg 2\pi/T$, where $1/T$ is the symbol rate determining the modulation bandwidth of the optical carrier. In such a case, Eq. (12) shows that the electric field of the signal light is down-converted to the IF signal including the amplitude information and the phase information without spectral folding, as shown in Fig. 4.

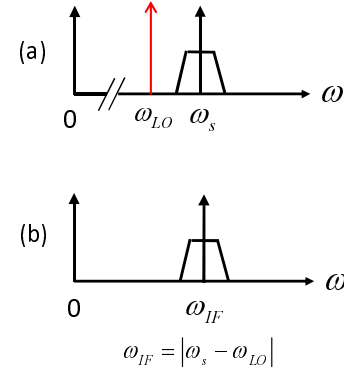


Fig. 4. Spectra of (a) the optical signal and (b) the down-converted IF signal. By setting the LO frequency close to the signal frequency, we can obtain the IF spectrum shown in (b).

The signal phase is given as $\theta_{sig}(t) = \theta_s(t) + \theta_{sn}(t)$, where $\theta_s(t)$ is the phase modulation and $\theta_{sn}(t)$ is the phase noise. The receiver output is given by

$$I(t) = 2R\sqrt{P_s(t)P_\ell} \cos\{\omega_{IF}t + \theta_s(t) + \theta_n(t)\}, \quad (14)$$

where $\theta_n(t)$ is the total phase noise given by

$$\theta_n(t) = \theta_{sn}(t) - \theta_\ell(t). \quad (15)$$

We can write the complex amplitude of the current in Eq. (14) as

$$I_c(t) = 2R\sqrt{P_s(t)P_\ell} \exp[i\{\theta_s(t) + \theta_n(t)\}]. \quad (16)$$

Note that Eq. (16) is equivalent to the complex amplitude $A_s(t)$ of the optical signal except for the phase noise increase stemming from the LO.

In order to synchronously demodulate the heterodyne-detected IF signal given by Eq. (14), we need to employ the phase-locked loop (PLL) circuit at the IF stage. Although the total phase noise $\theta_n(t)$ might vary with time, the electrical PLL circuit can estimate the phase noise and decode the symbol on the phasor given by Eq. (16).

C. Homodyne Detection

Homodyne detection refers to the case that $\omega_{IF} = 0$. The photodiode current from the homodyne receiver becomes

$$I(t) = 2R\sqrt{P_s(t)P_\ell} \cos\{\theta_{sig}(t) - \theta_\ell(t)\}. \quad (17)$$

Equation (17) implies that the homodyne receiver measures the inner product between the signal phasor and the LO phasor, as shown in Fig. 5. In order to decode the symbol correctly, the LO phase $\theta_\ell(t)$ must track the transmitter phase noise $\theta_{sn}(t)$ such that $\theta_n(t) = 0$. This function must be realized by OPLL; however, in practice, the implementation of such a loop is not simple and increases the complexity of configuration. In addition, Eq. (17) only gives the cosine component (in other words, the in-phase component with respect to the LO phase), and the sine component (the quadrature component) cannot be detected at the same time. Therefore, this type of homodyne receiver is not able to extract full information on the signal complex amplitude.

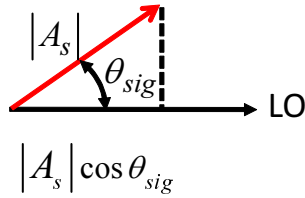


Fig. 5. Phasor diagram of the signal and LO for homodyne detection, which measures the inner product between the signal and LO phasors.

D. Phase Diversity

By preparing another LO, whose phase is shifted by 90° , in the homodyne receiver and branching the signal into two paths, we can detect both in-phase and quadrature components of the signal light simultaneously, as shown in Fig. 6. The optical circuit giving the 90° phase shift is called the 90° optical hybrid. The phasor diagram in Fig. 7 shows that we can measure the in-phase component with LO1 and the quadrature component with LO2.

Using the receiver shown in Fig. 6, we can obtain four outputs E_1 , E_2 , E_3 , and E_4 from the two inputs E_s and E_ℓ as

$$E_1 = \frac{1}{2}(E_s + E_\ell), \quad (18)$$

$$E_2 = \frac{1}{2}(E_s - E_\ell), \quad (19)$$

$$E_3 = \frac{1}{2}(E_s + iE_\ell), \quad (20)$$

$$E_4 = \frac{1}{2}(E_s - iE_\ell). \quad (21)$$

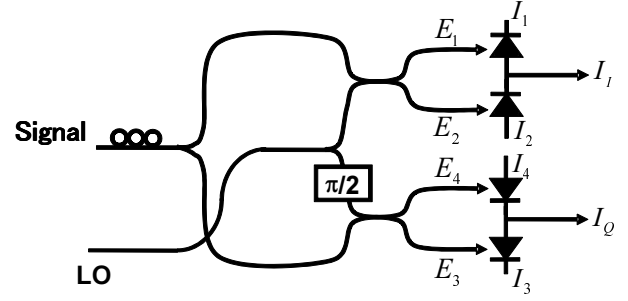


Fig. 6. Configuration of the phase-diversity homodyne receiver using a 90° optical hybrid. The in-phase component and the quadrature component of the signal complex amplitude are obtained simultaneously.

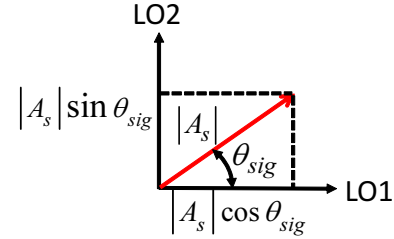


Fig. 7. Phasor diagram of the signal and LO for phase-diversity homodyne detection. Both the in-phase and quadrature components are measured at the same time.

Output photocurrents from balanced photodiodes are then given by

$$\begin{aligned} I_I(t) &= I_{I1}(t) - I_{I2}(t) \\ &= R\sqrt{P_s P_\ell} \cos\{\theta_{sig}(t) - \theta_\ell(t)\}, \end{aligned} \quad (22)$$

$$\begin{aligned} I_Q(t) &= I_{Q1}(t) - I_{Q2}(t) \\ &= R\sqrt{P_s P_\ell} \sin\{\theta_{sig}(t) - \theta_\ell(t)\}. \end{aligned} \quad (23)$$

Using Eqs. (22) and (23), we can restore the complex amplitude as

$$\begin{aligned} I_c(t) &= I_I(t) + iI_Q(t) \\ &= R\sqrt{P_s(t)P_\ell} \exp[i\{\theta_s(t) + \theta_n(t)\}], \end{aligned} \quad (24)$$

which is equivalent to the complex amplitude of the optical signal except for the phase noise increase.

Equation (24) shows that the electric field of the signal light is down-converted to the baseband. As shown in the down-converted spectrum of Fig. 8, we need to allow the negative frequency to express the complex amplitude at the baseband, which contains both the in-phase (or cos) and the quadrature (or sin) components. In contrast, the conventional homodyne receiver measures only the in-phase (or cos) component, because the spectrum is folded on the positive frequency side.

This type of receiver is commonly called the “phase-diversity homodyne receiver” [42], [43] or “intradyn receiver” [29]. The phase-diversity homodyne receiver emphasizes its ability of simultaneous IQ measurements, whereas the intradyne receiver emphasizes that the IF is not necessarily equal to zero as far as the condition $|\omega_{IF}| < 2\pi/T$ is satisfied. In this paper, we use the phase-diversity homodyne receiver.

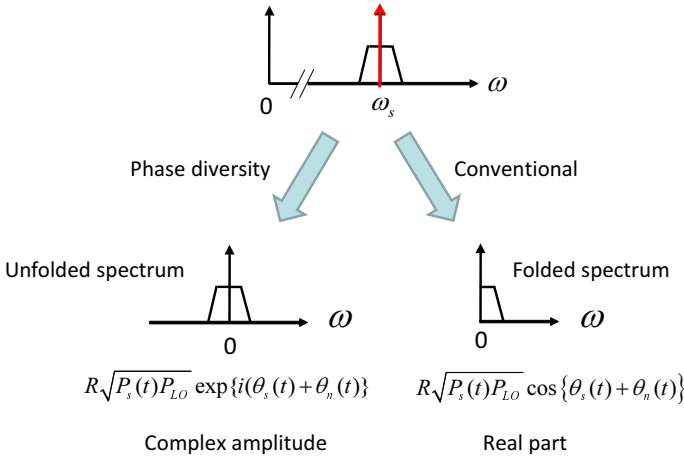


Fig. 8. Spectra of (a) the optical signal and (b) the homodyne-detected baseband signal. The conventional homodyne receiver measures only the real part of the optical complex amplitude, whereas the phase diversity receiver measures the complex amplitude itself, whose spectrum exists in both the positive and the negative frequency sides.

Eventually, the phase-diversity homodyne receiver and the heterodyne receiver can similarly restore the full information on the optical complex amplitude, as shown by Eqs. (16) and (24), respectively. However, because the phase-diversity homodyne receiver generates the baseband signal directly, it is more advantageous than the heterodyne receiver, which must deal with a rather high IF. Actually, we can construct the phase tracking circuit in the digital domain, as shown in Sec. VI.

E. Polarization Diversity

Thus far, it was assumed that the polarization of an incoming signal was always aligned to that of LO. However, in practical systems, the polarization of the incoming signal is unlikely to remain aligned to the SOP of the LO because of random changes in the birefringence of transmission fibers. To deal with this problem, the polarization-diversity receiver has been introduced into the coherent receiver.

The homodyne receiver employing polarization and phase diversities is shown in Fig. 9, where two phase-diversity homodyne receivers are combined in the polarization diversity configuration [35]. The incoming signal having an arbitrary SOP is separated into two linear polarization components with a polarization beam splitter (PBS); these components are homodyne-detected with the same LO. Even in the case of the polarization-multiplexed system, we can demultiplex the polarization tributaries from four outputs of the receiver using DSP. The DSP algorithm will be given in Sec. VI.

IV. QUANTUM-NOISE PROPERTIES OF COHERENT DETECTION

As shown in Eq. (13), the output current from the coherent receiver increases in proportion to $\sqrt{P_L}$. Therefore, when the LO power is sufficiently high, the output current can be much larger than the circuit noise in the receiver; in such a case, the receiver sensitivity is determined from the

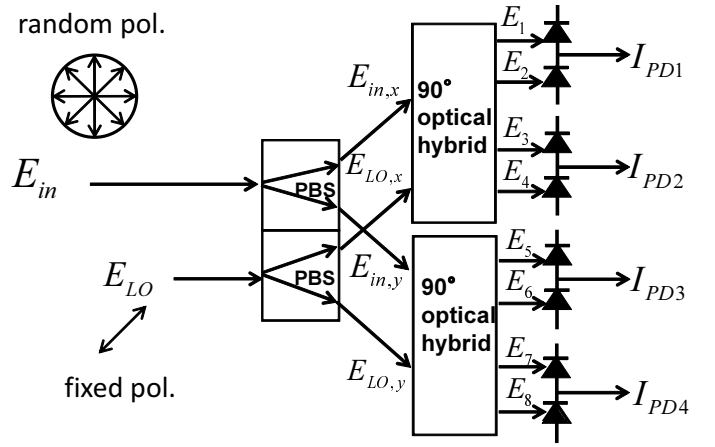


Fig. 9. Configuration of the homodyne receiver employing phase and polarization diversities. Two phase-diversity homodyne receivers are placed in the polarization-diversity configuration, where a common LO is employed.

quantum noise of the coherent light as well as the quantum-mechanical properties of coherent detection. In this section, we discuss the quantum-mechanical nature of coherent detection and elucidate the fundamental difference between heterodyne and homodyne detection.

A. Noise of Coherent State of Light

Let the quantum-mechanical operator for the in-phase component of the complex electric field be \hat{a}_1 , and let that for the quadrature component be \hat{a}_2 . These operators satisfy the commutation relation given by [44]

$$\begin{aligned} [\hat{a}_1, \hat{a}_2] &= \hat{a}_1 \hat{a}_2 - \hat{a}_2 \hat{a}_1 \\ &= \frac{i}{2}. \end{aligned} \quad (25)$$

Using Robertson's inequality [45] given by

$$\langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle \geq \left| \frac{1}{2} [\hat{a}_1, \hat{a}_2] \right|^2, \quad (26)$$

where $\langle \Delta \hat{a}_{1,2}^2 \rangle$ denote the variances of $\hat{a}_{1,2}$, Eq. (25) leads to Heisenberg's uncertainty relation between in-phase and quadrature components of the complex electric field given by

$$\langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle \geq \frac{1}{16}. \quad (27)$$

For the coherent state of light, we have

$$\langle \Delta \hat{a}_1^2 \rangle = \frac{1}{4}, \quad (28)$$

$$\langle \Delta \hat{a}_2^2 \rangle = \frac{1}{4}. \quad (29)$$

Equations (28) and (29) imply that the in-phase component and the quadrature component of the complex electric field in the coherent state are equally noisy and have the minimum uncertainty allowed by Eq. (27). These fluctuations have half of the one-photon energy, which stems from the quantum-mechanical vacuum fluctuation.

Figure 10 shows the phasor diagram for the light in the coherent state, which can be represented by the classical

electric field associated with the vacuum fluctuation. The average complex amplitude α_s is normalized as

$$|\alpha_s| = \sqrt{n_s}, \quad (30)$$

where n_s denotes the average number of photons in a symbol duration T . The photon number fluctuation stems from the beat between the constant amplitude and the in-phase component of the vacuum fluctuation, which is given by

$$\begin{aligned} \langle \Delta n^2 \rangle &= 4 |\alpha_s|^2 \langle \Delta \hat{a}_1^2 \rangle \\ &= n_s. \end{aligned} \quad (31)$$

From the classical interpretation, the photocurrent fluctuation is induced by the optical-to-electrical (OE) conversion process, and such noise is sometimes called shot noise. However, the photocurrent fluctuation is attributed to the nature of the coherent state of light, as shown in the following. The photocurrent obtained by direct detection is written as

$$\hat{I} = \frac{q\hat{n}}{T}. \quad (32)$$

Then, the average of the photocurrent is given by

$$\langle \hat{I} \rangle = \bar{I} = \frac{qn_s}{T}. \quad (33)$$

On the other hand, its variance is written as

$$\begin{aligned} \langle \Delta \hat{I}^2 \rangle &= \frac{q^2}{T^2} \langle \Delta n^2 \rangle \\ &= \frac{q\bar{I}}{T}. \end{aligned} \quad (34)$$

When the symbol duration of the signal is T , the sampling theorem determines the minimum required bandwidth B as

$$B = \frac{1}{2T}. \quad (35)$$

Equations (34) and (35) yield the well-known shot-noise formula [6]:

$$\langle \Delta \hat{I}^2 \rangle = 2q\bar{I}B. \quad (36)$$

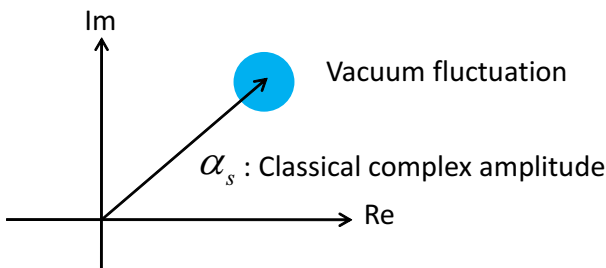


Fig. 10. Phasor diagram of the light in the coherent state. The electric field is expressed by a fixed classical complex amplitude associated with the vacuum fluctuation.

B. Quantum-mechanical Restriction on Simultaneous Measurements of IQ Components

Next, we consider the simultaneous measurement of the IQ components of the coherent state of light. Simultaneous measurement means that the measurement of the in-phase component does not affect the succeeding measurement of the quadrature component and vice versa. Such measurement is only achieved when the output operators \hat{x} and \hat{y} of the measurement system commute each other such that

$$[\hat{x}, \hat{y}] = 0. \quad (37)$$

In order for the output operators \hat{x} and \hat{y} to satisfy Eq. (37), they must be written as [44]

$$\hat{x} = \hat{a}_1 + \hat{B}_1, \quad (38)$$

$$\hat{y} = \hat{a}_2 + \hat{B}_2, \quad (39)$$

where \hat{B}_1 and \hat{B}_2 are noise operators associated with the measurement process. Using Eqs. (38) and (39), the commutation relation for \hat{x} and \hat{y} is given by

$$[\hat{x}, \hat{y}] = [\hat{a}_1, \hat{a}_2] + [\hat{a}_1, \hat{B}_2] + [\hat{B}_1, \hat{a}_2] + [\hat{B}_1, \hat{B}_2]. \quad (40)$$

Because there is no correlation between the inputs and the system noise, we have

$$[\hat{a}_1, \hat{B}_2] = 0, \quad (41)$$

$$[\hat{B}_1, \hat{a}_2] = 0. \quad (42)$$

Therefore, Eqs.(37), (40), (41), and (42) yield

$$[\hat{B}_2, \hat{B}_1] = [\hat{a}_1, \hat{a}_2]. \quad (43)$$

Using Robertson's inequality again, we have

$$\begin{aligned} \langle \Delta \hat{B}_1^2 \rangle \langle \Delta \hat{B}_2^2 \rangle &\geq \left| \frac{1}{2} \langle [\hat{a}_1, \hat{a}_2] \rangle \right|^2 \\ &= \frac{1}{16}. \end{aligned} \quad (44)$$

Then, the uncertainty relation between \hat{x} and \hat{y} is given by Eqs. (27), (38), (39), and (44) as

$$\begin{aligned} \langle \Delta \hat{x}^2 \rangle \langle \Delta \hat{y}^2 \rangle &= \left(\langle \Delta \hat{a}_1^2 \rangle + \langle \Delta \hat{B}_1^2 \rangle \right) \left(\langle \Delta \hat{a}_2^2 \rangle + \langle \Delta \hat{B}_2^2 \rangle \right) \\ &\geq \langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle + \langle \Delta \hat{B}_1^2 \rangle \langle \Delta \hat{B}_2^2 \rangle \\ &\quad + 2 \left(\langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle \langle \Delta \hat{B}_1^2 \rangle \langle \Delta \hat{B}_2^2 \rangle \right)^{\frac{1}{2}} \\ &\geq \frac{1}{4}. \end{aligned} \quad (45)$$

Note that Eq. (27) represents the uncertainty relation for the noise of the input light, Eq. (44) represents the uncertainty relation for the noise associated with the IQ measurement process, and Eq. (45) represents the uncertainty relation for the noise of the output from the measurement system. Thus, we find that for simultaneous IQ measurement, the additional quantum noise from the measurement process is inevitable; actually, the uncertainty product increases four-fold with the simultaneous IQ measurement.

On the other hand, when we measure either the in-phase component or the quadrature component of the complex electric field, such measurement is not governed by the uncertainty relation expressed by Eq. (45); in other words, noise-free measurement of either the in-phase component or the quadrature component is possible, although we can never eliminate the intrinsic quantum noise determined from the uncertainty relation, Eq. (27).

C. Quantum Noise Associated with Coherent Detection

From the classical description of the output of the coherent receiver given by Eq. (13), we can obtain its quantum mechanical description as

$$\hat{I} = \hat{I}_2 - \hat{I}_1 = \frac{q}{T} \left(\hat{E}_s^\dagger \hat{E}_\ell + \hat{E}_s \hat{E}_\ell^\dagger \right), \quad (46)$$

where $E_{s,\ell}^\dagger$ represents the time-dependent creation operator and $E_{s,\ell}$ denotes the annihilation operator. The subscripts s and ℓ denote the signal and LO, respectively. When the LO is the coherent light with a sufficiently large amplitude, we can ignore its fluctuation compared with its constant amplitude. In such a case, the LO annihilation operator is replaced with the classical complex amplitude A_ℓ , and the creation operator is replaced with A_ℓ^* :

$$\hat{E}_\ell \rightarrow A_\ell \exp(-i\omega_\ell t), \quad (47)$$

$$\hat{E}_\ell^\dagger \rightarrow A_\ell^* \exp(i\omega_\ell t). \quad (48)$$

Note that the use of the complex amplitude on $\exp(-i\omega_\ell t)$ in this section is different from that in Sec. III. This representation of the complex amplitude is common in the field in physics. We can assume that A_ℓ is real by letting the phase reference of the system be the LO phase. On the other hand, the annihilation and creation operators for the signal are written as

$$\hat{E}_s = \hat{a}_s \exp(-i\omega_s t), \quad (49)$$

$$\hat{E}_s^\dagger = \hat{a}_s^\dagger \exp(i\omega_s t), \quad (50)$$

where \hat{a}_s and \hat{a}_s^\dagger are time-independent annihilation and creation operators for the signal, respectively. Operators for the in-phase and quadrature components are given by

$$\hat{a}_{s1} = \frac{\hat{a}_s + \hat{a}_s^\dagger}{2}, \quad (51)$$

$$\hat{a}_{s2} = \frac{\hat{a}_s - \hat{a}_s^\dagger}{2i}. \quad (52)$$

Therefore, the photocurrent from the homodyne receiver, where $\omega_\ell = \omega_s$, is given by

$$\begin{aligned} \hat{I} &= \frac{q}{T} A_\ell (\hat{a}_s^\dagger + \hat{a}_s) \\ &= \frac{2q}{T} A_\ell \hat{a}_{s1}. \end{aligned} \quad (53)$$

Equation (53) shows that the homodyne receiver measures the in-phase component of the signal complex amplitude with respect to the LO phase, whereas we cannot determine the quadrature component simultaneously. In addition, it should be noted that homodyne detection is the noise-free measurement. Because the homodyne receiver determined only the in-phase

component, the noise-free measurement is allowed in contrast to the simultaneous IQ measurement, which inevitably generates the noise as shown in Eq. (45). When the signal phase is adjusted to 0 with the OPLL circuit, Eqs. (28) and (53) yield the S/N ratio of the received signal as

$$\frac{S}{N} = 4n_s. \quad (54)$$

From Eq. (53), it is evident that the fluctuation of the photocurrent stems from the quantum noise of the signal itself, and the signal quantum noise is multiplied by the LO amplitude due to the beat between the LO and the signal. Next, we consider the variance of the photocurrent. The variance is written from Eq. (53) as

$$\begin{aligned} \langle \Delta I^2 \rangle &= \frac{4q}{T} \frac{qn_\ell}{T} \langle \Delta \hat{a}_{s1}^2 \rangle \\ &= 2qI_\ell B \cdot 4\langle \Delta \hat{a}_{s1}^2 \rangle, \end{aligned} \quad (55)$$

where n_ℓ is the average LO photon number and $I_\ell (= qn_\ell/T)$ is the average photocurrent generated from the LO power. Using $\langle \Delta \hat{a}_{s1}^2 \rangle = 1/4$, the variance is rewritten as

$$\langle \Delta I^2 \rangle = 2qI_\ell B. \quad (56)$$

As shown by Eq. (36), this equation apparently implies that the noise of the photocurrent originates from shot noise due to the LO power (LO shot noise); however, the above-mentioned quantum-mechanical interpretation is correct.

To evaluate the bit-error rate performance of binary digital communication systems, we use the Q factor, defined as

$$Q = \frac{|d_0 - d_1|}{\sigma_0 + \sigma_1}, \quad (57)$$

where d_1 and d_0 denote signal levels for mark and space states, respectively, and σ_1 and σ_0 represent the standard deviations of noise for mark and space states, respectively. Using the Q factor, the bit-error rate (BER) is written as

$$\text{BER} = \frac{1}{2} \text{erfc} \left(\frac{Q}{\sqrt{2}} \right), \quad (58)$$

where $\text{erfc}(\cdot)$ denotes the complementary error function. In the binary PSK (BPSK) system using homodyne detection, where the phase is modulated between 0 and π in a binary manner, the Q factor is written as

$$Q = 2\sqrt{n_s}, \quad (59)$$

because $d_1 = \sqrt{n_s}$, $d_0 = -\sqrt{n_s}$, and $\sigma_1 = \sigma_0 = 1/2$. Then, the BER is given by

$$\text{BER} = \frac{1}{2} \text{erfc} (\sqrt{2n_s}). \quad (60)$$

The photon number/bit required to achieve $\text{BER} = 10^{-5}$, for example, is $n_s > 4.55$.

On the other hand, in the heterodyne receiver, where $\omega_s \neq \omega_\ell$, the beat between the signal and the LO appears at $\omega_{IF} = |\omega_s - \omega_\ell|$. It should be emphasized that the beat between the image-band signal at $\omega_i = 2\omega_\ell - \omega_s$ and the LO also appears at ω_{IF} . We do not inject the signal at the image band intentionally; however, the quantum-mechanical vacuum fluctuation at the image band is inevitably merged into the IF

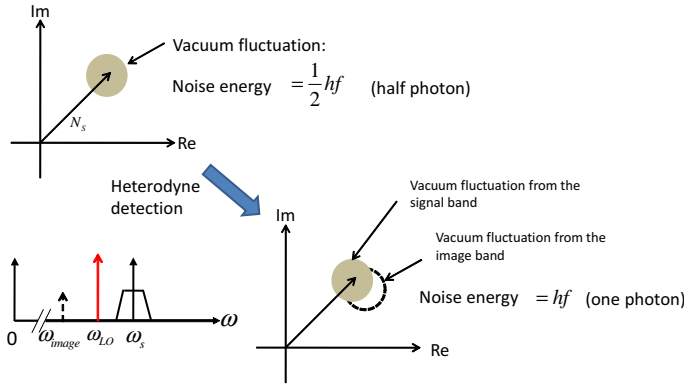


Fig. 11. Quantum-mechanical image of the noise associated with heterodyne detection. Vacuum fluctuations are merged into the IF signal from the image band as well as the signal band.

signal, as shown in Fig. 11. As a result, the energy of noise is doubled.

By taking these facts into consideration, Eq. (46) yields the output current of the heterodyne receiver as

$$\hat{I} = \frac{q}{T} \left(\hat{E}_s^\dagger \hat{E}_\ell + \hat{E}_s \hat{E}_\ell^\dagger + \hat{E}_i^\dagger \hat{E}_\ell + \hat{E}_i \hat{E}_\ell^\dagger \right), \quad (61)$$

where $\hat{E}_i = \Delta \hat{a}_i \exp(-i\omega_i t)$ represents the electric field at the image band including only the vacuum fluctuation. Equation (61) is transformed as

$$\begin{aligned} \hat{I} = & \frac{qA_\ell}{T} \left[(\hat{a}_s^\dagger + \Delta \hat{a}_i) \exp\{-i(\omega_\ell - \omega_s)t\} \right. \\ & \left. + (\hat{a}_s + \Delta \hat{a}_i^\dagger) \exp\{i(\omega_\ell - \omega_s)t\} \right], \end{aligned} \quad (62)$$

which is further rewritten as

$$\begin{aligned} \hat{I} = & \frac{2qA_\ell}{T} \left[(\hat{a}_{s1} + \Delta \hat{a}_{i1}) \cos(\omega_\ell - \omega_s)t \right. \\ & \left. + (\hat{a}_{s2} - \Delta \hat{a}_{i2}) \sin(\omega_\ell - \omega_s)t \right]. \end{aligned} \quad (63)$$

In Eq. (63), operators for in-phase and quadrature components at the image band $\Delta \hat{a}_{i1}$ and $\Delta \hat{a}_{i2}$ are given by

$$\Delta \hat{a}_{i1} = \frac{\Delta \hat{a}_i + \Delta \hat{a}_i^\dagger}{2}, \quad (64)$$

$$\Delta \hat{a}_{i2} = \frac{\Delta \hat{a}_i - \Delta \hat{a}_i^\dagger}{2i}. \quad (65)$$

We define operators \hat{x} and \hat{y} as

$$\hat{x} = \hat{a}_{s1} + \Delta \hat{a}_{i1}, \quad (66)$$

$$\hat{y} = \hat{a}_{s2} - \Delta \hat{a}_{i2}. \quad (67)$$

These are operators for measured IQ components and satisfy the commutation relation

$$[\hat{x}, \hat{y}] = 0, \quad (68)$$

because

$$\begin{aligned} [\hat{a}_{s1}, \hat{a}_{s2}] &= [\Delta \hat{a}_{i1}, \Delta \hat{a}_{i2}] = \frac{i}{2}, \\ [\hat{a}_{s1}, \Delta \hat{a}_{i2}] &= [\Delta \hat{a}_{i1}, \hat{a}_{s2}] = 0. \end{aligned} \quad (69)$$

Equation (68) shows that we can simultaneously measure the IQ components of the signal. Instead, vacuum noise is

merged into the IQ components from the image band, and they degrade the S/N ratio. When the signal is in the coherent state, variances for \hat{x} and \hat{y} are actually given by

$$\langle \Delta \hat{x}^2 \rangle = \frac{1}{2}, \quad (70)$$

$$\langle \Delta \hat{y}^2 \rangle = \frac{1}{2}. \quad (71)$$

When we measure the in-phase component of the signal after demodulation of the IF signal with the PLL circuit, the S/N ratio is given by Eq. (70) as

$$\frac{S}{N} = 2n_s. \quad (72)$$

The S/N ratio for heterodyne detection is degraded by 3 dB, compared with the homodyne case shown by Eq. (54). In other words, the heterodyne receiver incurs a 3-dB penalty to realize simultaneous IQ measurements. Note that Eqs. (70) and (71) satisfy the uncertainty relation given by Eq. (45).

Equation (63) yields the variance of the photocurrent as

$$\begin{aligned} \langle \Delta I^2 \rangle &= \frac{4q}{T} \frac{qn_\ell}{T} \frac{1}{2} (\langle \Delta \hat{a}_{s1}^2 \rangle + \langle \Delta \hat{a}_{s2}^2 \rangle + \langle \Delta \hat{a}_{i1}^2 \rangle + \langle \Delta \hat{a}_{i2}^2 \rangle) \\ &= 4qI_\ell B. \end{aligned} \quad (73)$$

Equation (73) apparently implies that the photocurrent fluctuation stems from LO shot noise because the required IF bandwidth is $2B$. However, the quantum-mechanical interpretation is that it is generated by the beat between the LO and vacuum noise at the signal band and the beat between the LO and vacuum noise at the image band, as we have discussed already.

On the other hand, we consider the BPSK system with heterodyne detection. Noting that $d_1 = \sqrt{n_s}$, $d_0 = -\sqrt{n_s}$, and $\sigma_1 = \sigma_0 = 1/\sqrt{2}$ in Eq. (57), we find that $Q = \sqrt{2n_s}$. Then, the BER is expressed as

$$\text{BER} = \frac{1}{2} \text{erfc}(\sqrt{n_s}). \quad (74)$$

From Eqs. (60) and (74), we find that the receiver sensitivity is degraded by 3 dB compared with homodyne detection. The photon number required to achieve $\text{BER} = 10^{-5}$, for example, is $n_s > 9.0$. We can construct the QPSK system with heterodyne detection, because simultaneous IQ measurements are feasible with heterodyne detection. The Q factor between the two closest constellation points is $Q = \sqrt{n_s}$, because $d_1 = \sqrt{n_s/2}$, $d_0 = -\sqrt{n_s/2}$, and $\sigma_1 = \sigma_0 = 1/\sqrt{2}$ in Eq. (57). Then, the symbol-error rate (SER) is approximately written as

$$\text{SER} = \frac{1}{2} \text{erfc}\left(\sqrt{\frac{n_s}{2}}\right). \quad (75)$$

This SER is approximately equal to the BER when symbols are encoded by the Gray code. Because the QPSK system sends two bits per symbol, the required number of photons per bit is the same between BPSK and QPSK systems.

In the case of the phase-diversity homodyne receiver, the input light is branched with the 3-dB coupler, as shown in Fig. 6. After power branching, we can achieve noise-free measurements of each of the in-phase and quadrature components using homodyne detection. However, the initial power

branching of the input coherent light induces a 3-dB S/N ratio reduction for each of the in-phase and quadrature components. This is because the signal power of each branch reduces by 3 dB, whereas the vacuum fluctuation is unchanged. Thus, the S/N characteristics of the phase-diversity homodyne receiver are identical to those of the heterodyne receiver.

D. Role of Optical Amplifiers in Coherent Systems

Optical amplifiers are introduced into coherent systems as optical pre-amplifiers and optical repeaters. In this subsection, we discuss the quantum-noise characteristics of optical amplifiers.

The optical amplifier evolves operators \hat{a}_1 and \hat{a}_2 for the in-phase and quadrature components of the input signal as [44]

$$\hat{b}_1 = \sqrt{G_1}\hat{a}_1 + \hat{F}_1, \quad (76)$$

$$\hat{b}_2 = \sqrt{G_2}\hat{a}_2 + \hat{F}_2, \quad (77)$$

where \hat{b}_1 and \hat{b}_2 are the output operators for the in-phase and quadrature components, respectively; G_1 and G_2 represent the in-phase gain and the quadrature gain, respectively; and \hat{F}_1 and \hat{F}_2 are the in-phase and quadrature noise from the amplifier, respectively.

As in the case of the input operators, the output operators must satisfy the commutation relation given by

$$[\hat{b}_1, \hat{b}_2] = \frac{i}{2}. \quad (78)$$

From Eqs. (76), (77), and (78), we find that noise operators must obey

$$[\hat{F}_1, \hat{F}_2] = \frac{i}{2} (1 - \sqrt{G_1 G_2}). \quad (79)$$

Next, we define equivalent input operators as

$$\hat{a}_{1in} = \frac{\hat{b}_1}{\sqrt{G_1}}, \quad (80)$$

$$\hat{a}_{2in} = \frac{\hat{b}_2}{\sqrt{G_2}}. \quad (81)$$

Using Eqs. (80) and (81), we can estimate the input operators from the output operators, \hat{b}_1 and \hat{b}_2 , and given amplifier gains, G_1 and G_2 . Note that $[\hat{a}_{1in}, \hat{a}_{2in}] \rightarrow 0$ when $G_{1,2} \rightarrow \infty$. This fact means that simultaneous IQ measurements become possible when amplifier gains for both in-phase and quadrature components are sufficiently high.

Instead, such simultaneous IQ measurements accompany the S/N reduction as shown in the following. The variances of equivalent input operators are given by

$$\langle \Delta \hat{a}_{1in}^2 \rangle = \langle \Delta \hat{a}_1^2 \rangle + \frac{\langle \Delta \hat{F}_1^2 \rangle}{G_1}, \quad (82)$$

$$\langle \Delta \hat{a}_{2in}^2 \rangle = \langle \Delta \hat{a}_2^2 \rangle + \frac{\langle \Delta \hat{F}_2^2 \rangle}{G_2}. \quad (83)$$

Robertson's inequality yields

$$\langle \Delta \hat{F}_1^2 \rangle \langle \Delta \hat{F}_2^2 \rangle \geq \frac{1}{16} (1 - \sqrt{G_1 G_2})^2. \quad (84)$$

From Eqs. (82), (83), and (84), the uncertainty product for equivalent input operators are calculated as

$$\begin{aligned} \langle \Delta \hat{a}_{1in}^2 \rangle \langle \Delta \hat{a}_{2in}^2 \rangle &= \left(\langle \Delta \hat{a}_1^2 \rangle + \frac{\langle \Delta \hat{F}_1^2 \rangle}{G_1} \right) \\ &\quad \times \left(\langle \Delta \hat{a}_2^2 \rangle + \frac{\langle \Delta \hat{F}_2^2 \rangle}{G_2} \right) \\ &\geq \langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle + \frac{\langle \Delta \hat{F}_1^2 \rangle \langle \Delta \hat{F}_2^2 \rangle}{G_1 G_2} \\ &\quad + 2 \left(\langle \Delta \hat{a}_1^2 \rangle \langle \Delta \hat{a}_2^2 \rangle \frac{\langle \Delta \hat{F}_1^2 \rangle \langle \Delta \hat{F}_2^2 \rangle}{\sqrt{G_1 G_2}} \right)^{\frac{1}{2}} \\ &\geq \frac{1}{16} \left(2 - \frac{1}{\sqrt{G_1 G_2}} \right)^2. \end{aligned} \quad (85)$$

When $G_1, G_2 \gg 1$, Eq. (85) yields

$$\langle \Delta \hat{a}_{1in}^2 \rangle \langle \Delta \hat{a}_{2in}^2 \rangle \geq \frac{1}{4}. \quad (86)$$

The uncertainty product is four times greater than the initial value of 1/16. Laser amplifiers such as the EDFA and the semiconductor optical amplifiers (SOA) belong to this category. Since $G_1 = G_2 \geq 1$ in this type of amplifier, we find that $\langle \Delta \hat{a}_{1in}^2 \rangle = \langle \Delta \hat{a}_{2in}^2 \rangle \geq 1/2$, which means that the minimum value of the noise figure (NF) of the amplifier is 3 dB. Further, note that the gain is insensitive to the carrier phase. The noise property of such an amplification process is equivalent to that of the heterodyne detection process.

On the other hand, the degenerate parametric optical amplifier satisfies the condition $G_1 \cdot G_2 = 1$. In such an amplifier, the uncertainty product is maintained after amplification as shown by Eq. (85), which means that noise-free amplification is achieved. However, if $G_1 \rightarrow \infty$, we cannot measure the quadrature component because $G_2 \rightarrow 0$; on the other hand, if $G_2 \rightarrow \infty$, we cannot measure the in-phase component because $G_1 \rightarrow 0$. This type of amplifier is called the phase-sensitive amplifier, because the amplification and de-amplification axes are determined from the phase of the pump wave for the parametric amplifier. The noise property of the phase-sensitive amplification process is equivalent to that of the homodyne detection process.

Such resemblance between optical amplification and detection is important to design the coherent transmission system shown in Fig. 12, where the fiber loss of each span $1/G$ is periodically compensated for by the repeater-amplifier gain G .

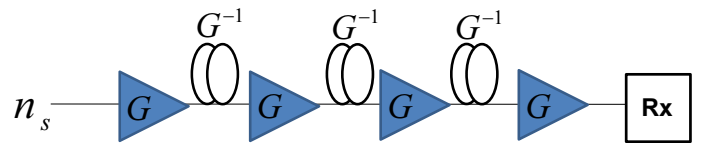


Fig. 12. Optical amplifier chain for coherent optical transmission. The optical amplifier gain of G compensates for the span loss of $1/G$.

Figure 13 shows possible combinations of the modulator, the repeater amplifier, and the receiver in the coherent transmission system. When we use both the IQ components for

modulation parameters, we need to employ laser amplifiers for repeaters and the heterodyne receiver (Fig. 13(a)) or the phase-diversity homodyne receiver (Fig. 13(b)). In such systems, the S/N ratio of the detected signal is given by $S/N=2n_s/m_a$, where m_a denotes the number of repeater amplifiers and we assume that the NF of each amplifier is ideally 3 dB. Note that when the optical pre-amplifier is used as shown in Fig. 12, we do not suffer from the S/N reduction of 3 dB in the receiver. This is because the optical-amplifier noise overwhelms the vacuum fluctuation associated with heterodyne detection or phase-diversity homodyne detection [36]. However, the optical-amplifier noise at the image band must be eliminated in the case of heterodyne detection; therefore, we can employ the optical pre-amplifier in the phase-diversity homodyne receiver much more easily than in the heterodyne receiver. It should also be emphasized that we can measure IQ components in an optically phase-insensitive manner; synchronous detection is done at the IF stage in the case of heterodyne detection (Fig. 13(a)), whereas it is done at the baseband in the case of phase-diversity homodyne detection (Fig. 13(b)).

On the other hand, when we use either the in-phase component or the quadrature component for modulation, the best alternative is to employ phase-sensitive repeater amplifiers and the homodyne receiver as shown in Fig. 13(c). In such a case, noise-free transmission and detection are feasible, and the S/N ratio of the received in-phase component is maintained at $S/N=4n_s$. However, the repeater amplifiers and the receiver are optically phase sensitive, which makes the system configuration more complex than the phase-insensitive scheme. In addition, we cannot use both the IQ components as modulation parameters at the same time.

Currently, only type (b) has been put into practical use. The advantages of type (b) are (1) simultaneous IQ modulation, amplification, and detection, (2) phase-insensitive amplification and detection, and (3) carrier-phase recovery with DSP after detection.

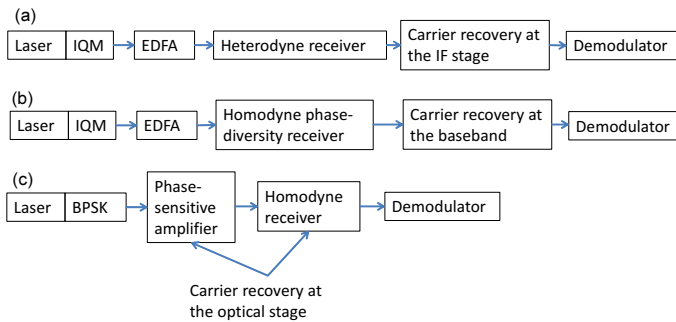


Fig. 13. Phase-insensitive and phase-sensitive coherent transmission systems. The optically phase-insensitive scheme includes (a) IQM, EDFAs, and the heterodyne receiver or (b) IQM, EDFAs, and the phase-diversity homodyne receiver. On the other hand, the optically phase-sensitive scheme (c) includes the BPSK modulator, phase-sensitive amplifiers, and the OPLL-based homodyne receiver.

V. MULTI-LEVEL MODULATION FORMATS

Using coherent receivers, we can demodulate any kind of multi-level modulation format. This subsection deals with

the method of such multi-level modulation, which aims at improvement of the spectral efficiency.

A. Optical IQ Modulator and Multi-level Modulation

Figure 14 shows a comparison of the device structure and the phasor diagram among phase modulation (PM), amplitude modulation (AM), and IQ modulation (IQM). The Pockels effect of non-centrosymmetric crystals induces a refractive-index change that is linearly proportional to the applied voltage. The light is phase-modulated (PM), passing through such a crystal. Optical amplitude modulation (AM) can be achieved by using phase modulators in a Mach-Zehnder configuration, which are driven in a push-pull mode of operation [46]. On the other hand, optical IQ modulation can be realized with Mach-Zehnder-type push-pull modulators in parallel, between which a $\pi/2$ -phase shift is given [25]. The IQ components of the optical carrier are modulated independently with the IQ modulator, enabling any kind of multi-level modulation format. IQ modulators integrated on LiNbO₃ substrates are now commercially available, and InP-based modulators have also been developed. In addition, by using two IQ modulators whose outputs are orthogonally polarized, we can generate the polarization-multiplexed signal.

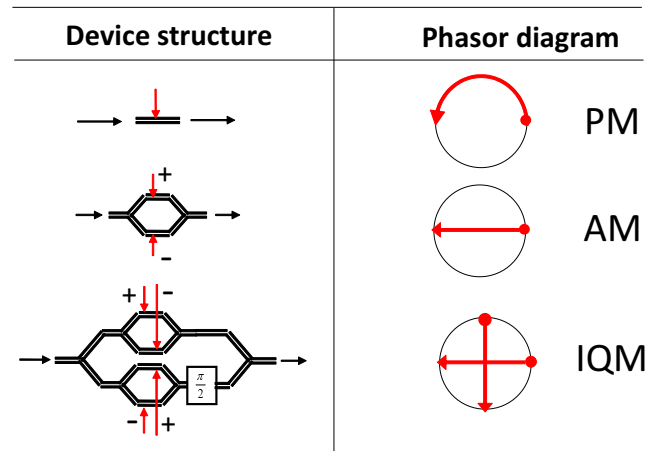


Fig. 14. Comparison of the device structure and phasor diagram of phase modulation (PM), amplitude modulation (AM), and IQ modulation (IQM).

While the spectral efficiency of binary modulation formats is limited to 1 bit/s/Hz/polarization, which is called the Nyquist limit, multi-level modulation formats with m_b bits of information per symbol can achieve a spectral efficiency of up to m_b bit/s/Hz/polarization. Figure 15 shows constellation maps for BPSK, QPSK, 8PSK, and 16QAM formats. These modulation formats can transmit 1 bit, 2 bits, 3 bits, and 4 bits per symbol, respectively. Each symbol is encoded by the Gray code to minimize the bit error. Thus, there is only one bit difference between the two nearest constellation points. In addition, we can double the bit rate using polarization-division multiplexing.

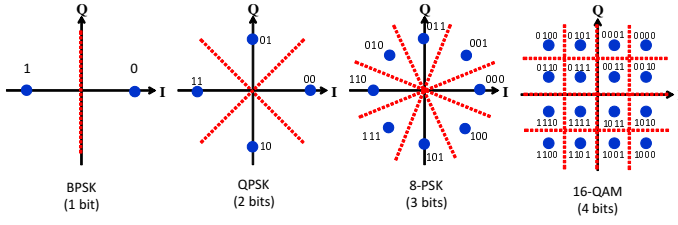


Fig. 15. Constellation maps of BPSK, QPSK, 8-PSK, and 16-QAM modulation formats. Each symbol is encoded by the Gray code. Namely, there is only one bit difference between the two nearest constellation points.

B. Spectral Efficiency

The channel capacity of the link C is given by [47] [48]

$$C = \frac{WD}{2} \log_2 \left(1 + \frac{S}{N} \right), \quad (87)$$

where W is the optical signal bandwidth, D is the dimension of the modulation format, S is the average signal power, and N is the average noise power. Because we always employ both the IQ components in the coherent optical communication system, $D = 2$ for single polarization transmission, and $D = 4$ for dual polarization transmission. The dimension can further be increased by using time/frequency channels for coding. Equation (87) can be rewritten as

$$\frac{C}{W(D/2)} = \log_2 \left[1 + \frac{C}{W(D/2)} \frac{E_b}{N_0} \right], \quad (88)$$

where N_0 is the power spectral density of noise. The parameter E_b/N_0 denotes the ratio of energy per bit to noise power spectral density, and $SE_{lim} = C/W(D/2)$ is the Shannon limit of the spectral efficiency per polarization and channel. Substituting SE_{lim} for $C/W(D/2)$ in Eq. (88), we have

$$SE_{lim} = \log_2 \left[1 + SE_{lim} \frac{E_b}{N_0} \right], \quad (89)$$

which is further transformed as

$$\frac{E_b}{N_0} = \frac{2^{SE_{lim}} - 1}{SE_{lim}}. \quad (90)$$

Thus, we can calculate E_b/N_0 as a function of SE_{lim} in the Shannon limit using Eq. (90) [49]. The colored region in Fig. 16 represents the region that obeys Shannon's theory.

We can obtain SE of a specific modulation format as

$$SE = \frac{\log_2 M}{D/2}, \quad (91)$$

where M is the number of symbols. On the other hand, by computer simulations, we can calculate the value of E_b/N_0 of the modulation format so that it gives BER = 10^{-5} , for example. In the cases of ideal heterodyne detection and phase-diversity homodyne detection of the coherent state of light, E_b/N_0 is equal to the number of photons per bit as

$$\frac{E_b}{N_0} = \frac{n_s}{m_b}, \quad (92)$$

because the number of noise photons in the signal is one, as shown in Fig. 11.

Then, SE and E_b/N_0 of the modulation format can be plotted in Fig. 16. The dots in Fig. 16 show SE as a function of E_b/N_0 for M -ary PSK, whereas the squares show the SE -versus- E_b/N_0 relation for M -ary QAM. We find that multi-level modulation formats improve the spectral efficiency SE at the expense of an increase in E_b/N_0 .

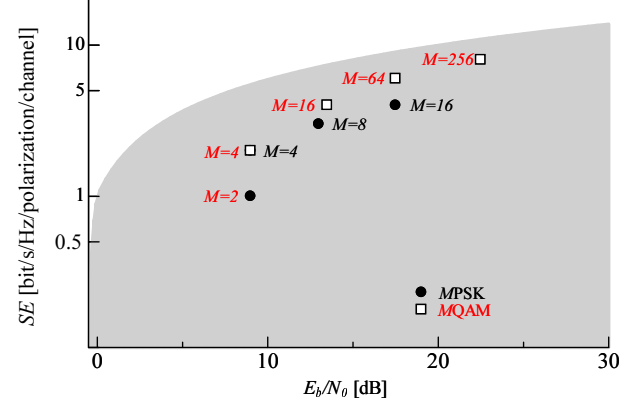


Fig. 16. Spectral efficiency SE as a function of E_b/N_0 . We plot $(E_b/N_0, SE)$ of M -ary PSK and QAM. The colored region is the region obeying Shannon's theory.

Recently, multi-dimensional modulation formats that achieve a trade-off between SE and E_b/N_0 have been studied intensively [48], [50]-[52]. Coherent optical communication technologies enable us to modulate an optical electric field in a four-dimensional (4D) manner, using two degrees of freedom of the complex amplitude of the electric field, that is, in-phase and quadrature components, and two degrees of the freedom of SOP. Then, the transmitted optical electric field \mathbf{E} is written as the following 4D vector:

$$\mathbf{E} = \begin{bmatrix} E_{x,I} \\ E_{x,Q} \\ E_{y,I} \\ E_{y,Q} \end{bmatrix}, \quad (93)$$

where the subscripts x and y represent x - and y -polarization components, respectively, and I and Q denote in-phase and quadrature components, respectively. A symbol of the conventional dual-polarization (DP) quadrature phase-shift keying (QPSK) format is transmitted by one of the 16 vectors given by

$$\mathbf{E}_{DP-QPSK} = \begin{bmatrix} \pm 1 \\ \pm 1 \\ \pm 1 \\ \pm 1 \end{bmatrix}. \quad (94)$$

On the other hand, well-designed constellations in the 4D vector space can improve E_b/N_0 and/or the spectral efficiency toward the Shannon limit. For example, the polarization-switched QPSK (PS-QPSK) format is the most power efficient format among all the 4D modulation formats. The following eight vectors are transmitted per symbol in this format [50],

[51]:

$$\mathbf{E}_{PS-QPSK} = \begin{bmatrix} \pm 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \pm 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix}. \quad (95)$$

Set-partitioning of 4D M -ary QAM constellations [52] and 4D permutation modulation [48] have been proposed as systematic approaches to design efficient 4D constellations.

Furthermore, FEC is a powerful tool for reducing E_b/N_0 toward the Shannon limit [53]. Basically, such improvement stems from the fact that the block code having a code length of n can be regarded as n -dimensional modulation using n time slots, and constellation points can be well aligned in the n -dimensional vector space. Thus, coding and multi-level/multi-dimensional modulation should be designed in a unified manner, which is exactly the concept of coded modulation.

VI. DIGITAL SIGNAL PROCESSING IN COHERENT RECEIVERS

This section describes the details of digital signal processing for coherent receivers. Outputs from the homodyne receiver comprising phase and polarization diversities are processed by DSP circuits, restoring the complex amplitude of the signal in a stable manner despite fluctuations of the carrier phase and the signal SOP. Fast control of such time-varying parameters in the digital domain can greatly enhance the system stability as compared with optical control methods. In addition, adaptive signal equalization is a powerful function of DSP.

A. Configuration of Digital Coherent Receivers

Figure 17 shows the basic configuration of the digital coherent receiver. First, the incoming signal is detected linearly with the homodyne receiver consisting of phase and polarization diversities. Using this receiver, we can obtain the complex amplitude of each polarization. The complex amplitudes measured by the receiver are converted into digital data with ADCs and processed by DSP circuits. Improvements in the performance, speed, and reliability of integrated circuits have made DSP an attractive approach for recovering the symbol from the homodyne-detected signal. In addition, the IQ information is preserved even after the OE conversion process; therefore, adaptive signal equalization becomes possible at the receiver.

The DSP circuit typically consists of the sequence of operations shown in Fig. 18. First, the four-channel ADC samples the received signal. In general, we need to use $\times 2$ oversampling at $2/T$ to remove the aliasing effect associated with ADC. For example, in 100-Gbit/s dual-polarization QPSK systems, the sampling rate is as high as 50 GS/s, because the symbol rate is 25 Gbaud. Let E_x and E_y be complex amplitudes in the digital domain, which are obtained from the x -polarization port and the y -polarization port of the polarization-diversity receiver, respectively. These are equalized by a fixed frequency-domain equalizer (FDE), which removes inter-symbol interference (ISI) stemming from fixed GVD of the link. The clock frequency is recovered from

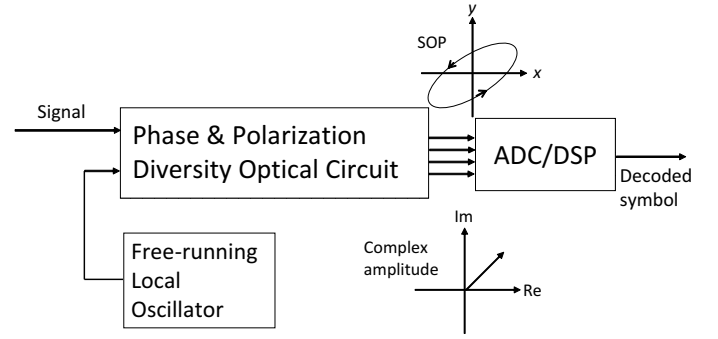


Fig. 17. Basic configuration of the digital coherent receiver. Four outputs from the homodyne receiver consisting of phase and polarization diversities are sent to the ADC/DSP circuit. These outputs include all the information of the complex amplitude of the signal, and linear impairments are compensated for in the DSP core.

the FDE output using DAC and an analog electrical circuit. Next, the dual-polarization signal is demultiplexed and PMD is compensated for with an adaptive equalizer. Clock-phase recovery can also be achieved at the adaptive equalization stage. Then, the carrier phase is estimated and the symbol is decoded. FEC decoding is an important function of the DSP circuit; however, it is outside the scope of this paper.

In the following subsections, we discuss DSP circuits in greater detail.

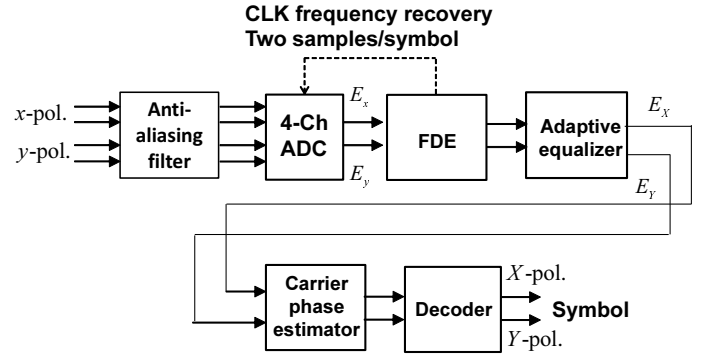


Fig. 18. Typical sequence of digital signal processing for decoding the symbol. After AD conversion, sampled complex amplitudes E_x and E_y are processed by the fixed GVD equalizer, the adaptive equalizer, the carrier-phase estimator, and the decoder. The clock frequency is recovered after FDE, using DAC and the analog circuit outside the DSP core.

B. Sampling of the Signal with ADCs and Clock Frequency Extraction

The Nyquist bandwidth of the optical signal with a symbol duration of T is given by $B = 1/T$. Then, outputs from the IQ ports of the phase-diversity homodyne receiver have a bandwidth of $B/2 = 1/(2T)$, as shown in Fig. 19. According to the sampling theorem, the sampling rate must be twice as high as the maximum frequency of the signal in order to avoid the aliasing effect. Therefore, the minimum required sampling rate for the Nyquist-filtered signal is $R_0 = 1/T$,

However, in general, when the bandwidth of the homodyne-detected signal is larger than $B/2$, we need oversampled

ADC. In the current 100-Gbit/s DP-QPSK receiver, the symbol rate is $1/T=25$ Gbaud and the sampling rate is $R = 50$ GS/s; therefore, the ADCs are operating in the two-times oversampling condition $R = 2/T = 2R_0$, which allows for a maximum signal frequency of $1/T$. The fractional sampling rate pR_0 can avoid the aliasing effect, as far as the signal bandwidth is $B = p/T$, where p is a fraction satisfying $1 < p < 2$. However, in such a case, digitized data are usually resampled to satisfy the two-times oversampling condition, which is necessary for adaptive equalization. This resampling process increases the computational complexity.

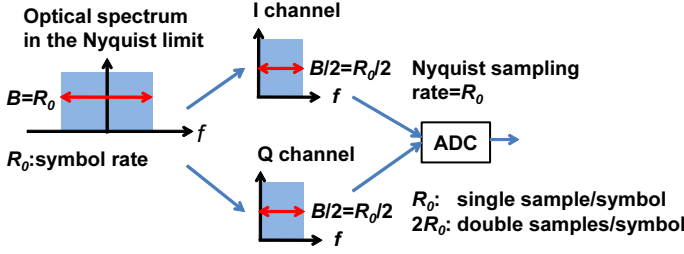


Fig. 19. Optical and electrical spectra of the signal shaped by the ideal Nyquist filter. The minimum required sampling rate is $R = R_0 = 1/T$, because the maximum frequency of IQ signals is $1/(2T)$.

Long-haul and high-bit-rate optical transmission systems suffer from linear transmission impairments such as GVD and PMD of optical fibers; however, they are efficiently compensated for by using fixed/adaptive equalizers implemented in the DSP core of coherent optical receivers. Such an efficient equalization function is mainly attributed to two-times oversampling in ADC. Two-times oversampling can avoid the aliasing effect associated with ADC, thereby allowing for precise signal equalization. In addition, the clock frequency can be extracted from the digital signal after fixed GVD compensation, whereas sampling phases for two polarization tributaries can separately be adjusted with adaptive equalizers, which results in clock-phase recovery and PMD compensation simultaneously.

In the remainder of this subsection, we analyze the clock-frequency recovery method through computer simulations. At the transmitter, we assume that the DP-QPSK signal is spectrally shaped with a Gaussian filter having a transfer function given by

$$G(f) = \exp \left[-2 \ln 2 \left(\frac{fT}{B_g T} \right)^2 \right], \quad (96)$$

where f denotes the optical frequency measured from the carrier frequency and B_g denotes the full-width at half maximum (FWHM) of the filter. In our simulations, we assume that $B_g T = 2$. On the other hand, the output from the phase/polarization-diversity homodyne receiver is filtered by an anti-aliasing filter before AD conversion. We assume that the Nyquist filter is used as the anti-aliasing filter, whose

transfer function is given by

$$F(f) = \begin{cases} 1 & \text{when } 0 \leq fT < \frac{1-\alpha}{2} \\ \frac{1}{2} \left\{ 1 - \sin \left[\frac{\pi}{2\alpha} (2fT - 1) \right] \right\} & \text{when } \frac{1-\alpha}{2} \leq fT < \frac{1+\alpha}{2} \\ 0 & \text{when } \frac{1+\alpha}{2} \leq fT \end{cases} \quad (97)$$

In Fig. 20, the broken curve (a) shows the amplitude response of a Nyquist filter with a roll-off factor of $\alpha = 0$. However, it is difficult to achieve such sharp cutoff characteristics using analog filters; hence, we use a Nyquist filter with a roll-off factor $\alpha = 1$ in our simulations, whose amplitude response is shown by the solid curve (b) in Fig. 20. Therefore, the maximum frequency of the received signal is $1/T$; then, the minimum sampling rate of AD conversion is $R = 2/T$, which implies two-times oversampling, in order to avoid the aliasing effect. In practical systems, anti-aliasing filtering can naturally be performed on the basis of the lowpass characteristics of photodiodes and front-end circuits.

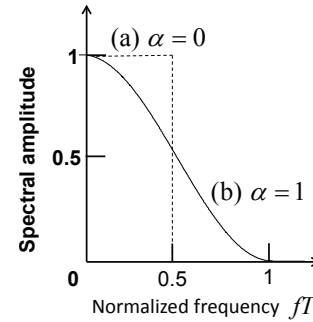


Fig. 20. Transfer function of Nyquist filters. Amplitude responses are shown for (a): $\alpha = 0$ and (b): $\alpha = 1$ in the linear scale.

The bandwidth of the sampled signal is $1/T$; therefore, the clock at the frequency of $1/T$ cannot be restored directly from the sampled signal. Figure 21 shows the circuit for recovering the clock frequency [54]. The total signal intensity is obtained in the analog domain after fixed GVD equalization and DAC. Although the homodyne-detected signal has a bandwidth of $1/T$, the bandwidth of the signal intensity is extended to $2/T$ through the squaring operation in the analog circuit. Then, using the PLL circuit, we can restore the clock frequency of $1/T$, which controls the sampling instances of AD converters.

Figure 22 shows the simulation result for the spectrum of the signal intensity obtained by GVD compensation, DAC, and squaring operation in the analog circuit, when E_b/N_0 is 10 dB. From the line spectrum at $f = 1/T$, we can extract the clock frequency using the PLL circuit. ADCs are driven with a clock whose frequency is doubled. However, the clock phase can hardly be determined from the PLL circuit, because PMD is usually involved in the received signal in long-haul

systems. The clock phase is actually recovered by using the adaptive equalizer together with PMD compensation.

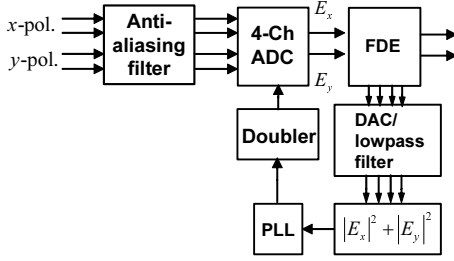


Fig. 21. Clock-frequency recovery circuit using $\times 2$ oversampled ADC. The digital signal after GVD compensation is converted into the analog signal with DA converters. The clock frequency is extracted from the signal intensity obtained in the analog domain.

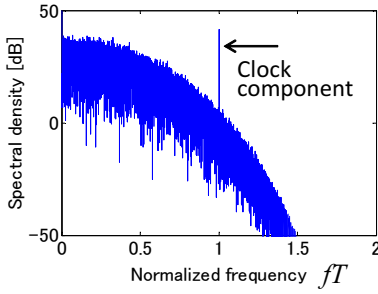


Fig. 22. Power spectrum of the signal intensity measured by the setup shown in Fig. 21. The clock frequency is extracted from the line spectrum at $f = 1/T$ using the PLL circuit.

C. Adaptive FIR Filter

Figure 23 shows the configuration of the finite-impulse-response (FIR) filter [55]. The input signal $x(n)$ is delayed by integer multiples of T/m , where m is an integer representing the oversampling ratio. We define a vector using these delayed elements of the input as

$$\vec{x}(n) = [x(n), x(n-1), \dots, x(n-k+1)]^T, \quad (98)$$

where the superscript “T” denotes the transpose of a matrix. The tap coefficients also constitute a vector as

$$\vec{c}(n) = [c_0(n), c_1(n), \dots, c_{k-1}(n)]^T. \quad (99)$$

Then, the output of the FIR filter is given by

$$\begin{aligned} y(n) &= \vec{c}(n)^T \vec{x}(n) \\ &= \sum_{j=0}^{k-1} c_j(n) x(n-j). \end{aligned} \quad (100)$$

When $\vec{c}(n)$ represents the impulse response of a system, the discrete Fourier transform (DFT) of $\vec{c}(n)$ gives the transfer function of the system. Therefore, by designing $\vec{c}(n)$, we can obtain the desired transfer function as long as the maximum delay is larger than the impulse response of the system. Because $\vec{c}(n)$ can be adaptively controlled, such an FIR filter plays an important role in equalizing the optical signal

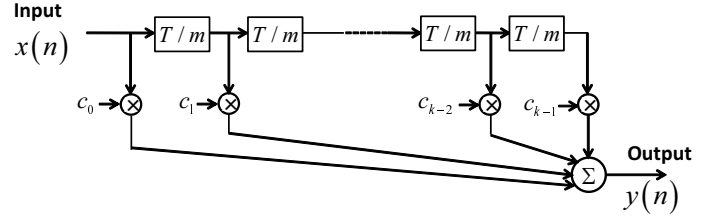


Fig. 23. Configuration of the finite-impulse-response (FIR) filter. By designing tap coefficients $\vec{c}(n)$, we can obtain any arbitrary transfer function when the tap length is sufficiently long.

transmitted through a fiber whose polarization characteristics fluctuate randomly.

Next, we discuss how such FIR filters are implemented in the digital coherent receiver to equalize linear transmission impairments. Let the Fourier transform of the complex amplitude of the signal at the input end of the transmission system be expressed as $\mathbf{E}_{in}(\omega) = [E_{in,x}(\omega), E_{in,y}(\omega)]^T$, where $E_{in,x}(\omega)$ and $E_{in,y}(\omega)$ denote Fourier transforms of the complex amplitude of the x -polarization and y -polarization, respectively. If the input power launched on a transmission link is sufficiently low to operate in the linear region, the Fourier transform of the received complex amplitude is given by

$$\begin{bmatrix} E_x(\omega) \\ E_y(\omega) \end{bmatrix} = \mathbf{H}_o(\omega) \begin{bmatrix} E_{in,x}(\omega) \\ E_{in,y}(\omega) \end{bmatrix}, \quad (101)$$

where $\mathbf{H}_o(\omega)$ represents the transfer function of the link and ω is the angular frequency measured from the carrier frequency. The transmitted signal suffers from distortion owing to the frequency dependence of the transfer function; in addition, the x - and y -polarization components are mixed together owing to nonorthogonal elements of the transfer matrix, which stem from the birefringence of transmission fibers.

In the linear region, the transfer function of the link $\mathbf{H}_o(\omega)$ can be modeled as a concatenation of the scalar GVD element $D(\omega)$, the two-by-two unitary matrix for PMD $\mathbf{U}(\omega)$, the two-by-two Hermite matrix for polarization-dependent loss (PDL) \mathbf{T} , and the unitary Jones matrix expressing the birefringence \mathbf{K} [56]:

$$\mathbf{H}_o(\omega) = D(\omega)\mathbf{U}(\omega)\mathbf{T}\mathbf{K}. \quad (102)$$

The scalar function $D(\omega)$ is given by

$$D(\omega) = \exp\left(-i\frac{\omega^2\beta_2 z}{2}\right), \quad (103)$$

where β_2 is the dispersion parameter, and z is the transmission distance. Because the GVD value is temporally fixed when the optical path is unchanged, we compensate for it with a fixed equalizer, and only the small remaining amount of GVD is adaptively controlled.

The PMD matrix $\mathbf{U}(\omega)$ is unitary and is given by

$$\mathbf{U}(\omega) = \mathbf{R}_1^{-1} \begin{bmatrix} \exp\left(i\frac{\omega\Delta\tau}{2}\right) & 0 \\ 0 & \exp\left(-i\frac{\omega\Delta\tau}{2}\right) \end{bmatrix} \mathbf{R}_1, \quad (104)$$

where \mathbf{R}_1 is a unitary matrix converting each principal state of polarization (PSP) into the x - or y -polarization component, and $\Delta\tau$ is the differential group delay (DGD) between two PSPs. The matrix \mathbf{K} is an Hermite matrix and is given by

$$\mathbf{T} = \mathbf{R}_2^{-1} \begin{bmatrix} \sqrt{\Gamma_{\max}} & 0 \\ 0 & \sqrt{\Gamma_{\min}} \end{bmatrix} \mathbf{R}_2, \quad (105)$$

where \mathbf{R}_2 is a unitary matrix converting each eigenmode for PDL into the x - or y -polarization component, and Γ_{\max} and Γ_{\min} are power transmission coefficients for the eigenmodes for PDL. The matrix \mathbf{K} is a frequency-independent 2×2 unitary matrix stemming from fiber birefringence, which mixes two polarization tributaries at the receiving end.

It should be noted that all the linear impairments stemming from the fiber are expressed as a single transfer matrix shown in Eq. (102). Then, the transfer matrix of the equalizer in the digital coherent receiver $\mathbf{H}_{eq}(\omega)$ should be the inverse matrix of $\mathbf{H}_o(\omega)$ as

$$\begin{aligned} \mathbf{H}_{eq}(\omega) &\simeq \mathbf{H}_o(\omega)^{-1} \\ &= \begin{bmatrix} h_{xx}(\omega) & h_{xy}(\omega) \\ h_{yx}(\omega) & h_{yy}(\omega) \end{bmatrix}. \end{aligned} \quad (106)$$

With such an inverse matrix, we can compensate for all the linear impairments at once.

Each element of the matrix $h_p(\omega)$ ($p = (xx, xy, yx, yy)$) in Eq. (106) can be realized by the FIR filter shown in Fig. 23 in the time domain; then, the matrix $\mathbf{H}_{eq}(\omega)$ in Eq. (106) is implemented in two-by-two butterfly-structured FIR filters shown in Fig. 24 [37].

Let the number of taps of FIR filters be k . The input column vectors $\vec{E}_x(n)$ and $\vec{E}_y(n)$ of the FIR filter are defined as

$$\vec{E}_x(n) = [E_x(n), E_x(n-1), \dots, E_x(n-k+2), E_x(n-k+1)]^T, \quad (107)$$

$$\vec{E}_y(n) = [E_y(n), E_y(n-1), \dots, E_y(n-k+2), E_y(n-k+1)]^T. \quad (108)$$

The tap-coefficient column vectors $\vec{h}_p(n)$ are written as

$$\vec{h}_p(n) = [h_{p,0}(n), h_{p,1}(n), \dots, h_{p,(k-2)}(n), h_{p,(k-1)}(n)]^T. \quad (109)$$

The filter outputs are then given by

$$E_X(n) = \vec{h}_{xx}(n)^T \vec{E}_x(n) + \vec{h}_{xy}(n)^T \vec{E}_y(n), \quad (110)$$

$$E_Y(n) = \vec{h}_{yx}(n)^T \vec{E}_x(n) + \vec{h}_{yy}(n)^T \vec{E}_y(n). \quad (111)$$

The required number of taps, k , is determined from the impulse response of the transfer function of the link.

D. Filter-tap Adaptation Algorithms

By controlling the tap coefficients, we can generate the inverse transfer matrix of the link, which can not only compensate for linear impairments such as GVD and PMD but also demultiplex polarization tributaries by removing the fiber birefringence. When we compensate for fixed GVD of the link, we may use a fixed equalizer where the tap coefficients are

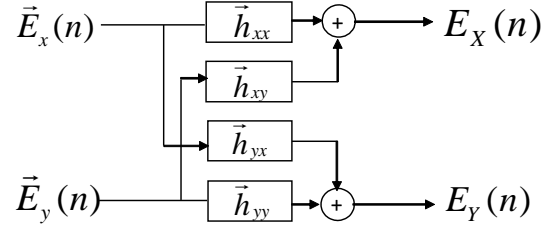


Fig. 24. Configuration of 2×2 butterfly-structured FIR filters. The DFT of the tap coefficients of each FIR filter generates elements of the transfer matrix given by Eq. (106).

predetermined because the GVD value is usually not time-varying. Furthermore, polarization-related impairments must be controlled in an adaptive manner. The fixed equalizer and the adaptive equalizer are usually concatenated.

Adaptive control of the filter-tap coefficients of each FIR filter can be achieved by using the decision-directed least-mean-square (DD-LMS) algorithm as follows [57]:

$$\vec{h}_{xx}(n+1) = \vec{h}_{xx}(n) + \mu e_X(n) \vec{E}_x(n)^*, \quad (112)$$

$$\vec{h}_{xy}(n+1) = \vec{h}_{xy}(n) + \mu e_X(n) \vec{E}_y(n)^*, \quad (113)$$

$$e_X(n) = d_X(n) - E_X(n), \quad (114)$$

$$\vec{h}_{yx}(n+1) = \vec{h}_{yx}(n) + \mu e_Y(n) \vec{E}_x(n)^*, \quad (115)$$

$$\vec{h}_{yy}(n+1) = \vec{h}_{yy}(n) + \mu e_Y(n) \vec{E}_y(n)^*, \quad (116)$$

$$e_Y(n) = d_Y(n) - E_Y(n). \quad (117)$$

In these equations, μ is a step-size parameter, $e_{X,Y}(n)$ is an error signal, and $d_{X,Y}(n)$ is either a training symbol in the training mode or a decoded symbol in the tracking mode for each polarization tributary. After updating the tap coefficients in the training mode so that the decoded symbol coincides with the desired symbol, we switch it into the tracking mode. In the case of $\times 2$ oversampling ($m = 2$), the tap coefficients are updated for every two samples.

When we cannot use the training sequence in the transmission system, we need to introduce an algorithm that operates in the blind mode such as the DD-LMS algorithm without the training sequence and the constant-modulus algorithm (CMA) [58].

In CMA, error signals for updating the tap coefficients are given instead of Eqs. (114) and (117) as

$$e_X(n) = (1 - |E_X(n)|^2) E_X(n), \quad (118)$$

$$e_Y(n) = (1 - |E_Y(n)|^2) E_Y(n). \quad (119)$$

This algorithm updates tap coefficients so that $|E_X(n)|^2$ and $|E_Y(n)|^2$ approach unity. Because this algorithm does not require the decision-feedback process, its computational complexity is lower than that of the DD-LMS algorithm.

In the case of M -ary PSK modulation formats, this algorithm works well because the signal has a constant envelope; therefore, all the linear impairments are compensated for at the same time, when the signal envelope becomes constant with CMA-based adaptive equalization. However, it is difficult to directly apply the CMA-based equalization method to M -ary QAM formats whose envelope has multiple levels. Another

problem inherent in CMA (and also in the blind-mode DD-LMS algorithm) is the singularity problem associated with polarization demultiplexing. When we use the blind-mode algorithm, it is likely that each output converges with the same polarization tributary. Such improper polarization demultiplexing occurs especially when the polarization-dependent loss (PDL) cannot be ignored. In [59], the modified CMA is proposed to prevent the singularity problem.

Note that CMA operates independently of the carrier phase fluctuation; therefore, the adaptive equalizer and the carrier phase estimator (Sec. VI-H) do not interfere with each other. On the other hand, because the DD-LMS algorithm tracks the carrier fluctuation, we need to separate the function of the equalizer from that of the carrier phase estimator [60]; otherwise, long delay taps will make the filtering performance unstable in the presence of fast phase fluctuations.

E. Equalization of Filtering Impairments

Transfer characteristics of optical filters in the link and electrical circuits in the transmitter and the receiver can be expressed by a scalar function $D_e(\omega)$. Even if $D_e(\omega)$ generates ISI for the signal, the adaptive FIR can compensate for it.

We examine such an equalization function of the FIR filter through computer simulations. Optical electric fields for x and y polarizations are independently modulated in the QPSK format, where each symbol is differentially encoded. The non-return-to-zero (NRZ) waveform includes 2 samples in one symbol duration and we apply $\times 2$ oversampled ADC. Additive white Gaussian noise (AWGN) is added to the filtered signals. The signal spectrum is shaped with either a root-Nyquist filter ($\alpha = 1$) or a Nyquist filter ($\alpha = 1$) at the transmitter. Note that the former filter generates ISI. We do not insert anti-aliasing filters because the optical signal spectrum is controlled within the bandwidth of $2/T$. Tap adaptation is done by the DD-LMS algorithm and the number of delay taps is 17. E_b/N_0 is determined at the input end of AD converters.

Figure 25 shows back-to-back BER characteristics as a function of E_b/N_0 . Red and black curves correspond to the x -polarization tributary and y -polarization tributary, respectively. When the signal spectrum is shaped with the root-Nyquist filter, the generated ISI is compensated for with another root-Nyquist filter, which is formed by the adaptive FIR filter. The root-Nyquist filter at the receiver acts as a matched filter and maximizes the S/N ratio after the filter. On the other hand, when the signal spectrum is shaped with the Nyquist filter, the FIR filter never has filtering characteristics because the input signal does not include any ISI. We find that the BER performance of the root-Nyquist filtering case (a) is better than that of the Nyquist filtering case (b). This is owing to the matched filtering and signal equalization functions of the FIR filter.

F. Clock-phase Recovery Function of the Adaptive Equalizer

A unique function of adaptive FIR filters is variable time delay with a resolution that is much higher than the sampling time interval. If the number of filter taps is sufficiently large,

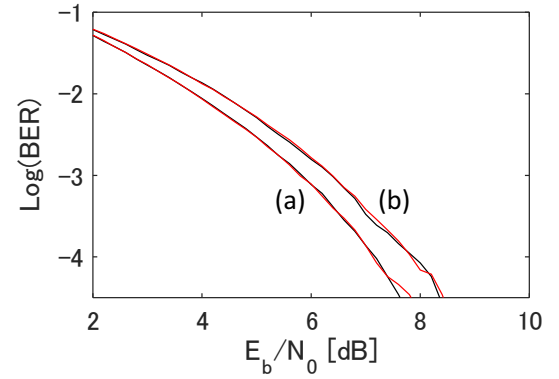


Fig. 25. Bit-error rates calculated as a function of E_b/N_0 when the input signal is shaped with the root-Nyquist filter (a) and with the Nyquist filter (b). The BER performance of (a) is better than that of (b). This is because the matched filter is created by the FIR filter at the receiver in the case of (a).

we can generate a quasi-continuously-variable time delay for the received signal by using discrete time-delay elements in the FIR filters. Owing to the quasi-continuously-variable time delay and the adaptive equalization algorithm, the signal waveform can be delayed so that the best time for symbol discrimination coincides with the sampling instant [61].

In fact, this process is clock-phase recovery. Note that this function is realized together with PMD compensation, and hence, the clock phase is determined for each of the polarization tributaries; on the other hand, in the clock-frequency recovery circuit shown in Fig. 21, the clock phase cannot be restored owing to PMD.

Figure 26 shows the principle of operation of clock-phase recovery with adaptive FIR filters. In Fig. 26 (a), the red arrows show the sampling instances for $\times 2$ oversampled ADC, and the initial sampling phase of AD conversion has not been optimized. On the other hand, after adaptive equalization using the algorithms shown in Sec. VI-D, the sampling phase is adjusted through quasi-continuous time shift; then, one out of every two sample points comes to the center position in one symbol interval, as shown by the red arrows in Fig. 26(b). The decision of symbols can be made at these optimum sampling points.

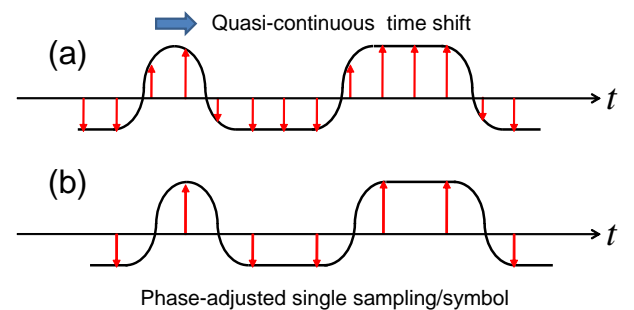


Fig. 26. Principle of clock-phase recovery with adaptive FIR filters. As shown in (a), quasi-continuous time shift is given to the sampled signal with the adaptive equalizer. Then, one out of every two sample points comes to the center position in one symbol interval, as shown in (b).

Next, we analyze the property of sampling-phase adjustment with FIR filters, ignoring other linear impairments. In the simulations, we assume that the clock frequency has already been recovered, as shown in Sec. VI-B, but that the initial sampling phase for AD conversion is uncontrolled.

Optical electric fields for x and y polarizations are independently modulated in the QPSK format, where each symbol is differentially encoded. The non-return-to-zero (NRZ) waveform includes 2^3 samples in one symbol duration. At the receiver, we assume $\times 2$ oversampling and the use of Nyquist filters (having a roll-off parameter of 0.2) as anti-aliasing filters. The initial sampling instant is set to $(t = 0, T/2)$, $(t = T/8, 5T/8)$, $(t = T/4, 3T/4)$, or $(3T/8, 7T/8)$ within one symbol interval. The step-size parameter μ in the DD-LMS algorithm is $1/(2^6 T)$.

Figure 27 shows bit-error rates calculated as a function of E_b/N_0 for the four initial sampling phases mentioned above. The red and black curves correspond to the x -polarization tributary and y -polarization tributary, respectively. Figures (a), (b), and (c) correspond to 1, 3, and 5 taps, respectively. The one-tap FIR filter is not capable of giving any time delay to the signal; therefore, the BER curves are strongly dependent on the initial sampling phase. On the other hand, those for the three-tap and five-tap filters become insensitive to the initial sampling phase. We find that the FIR filter with five delay taps can entirely remove the dependence of BER on the initial sampling phase, thereby achieving perfect clock-phase recovery.

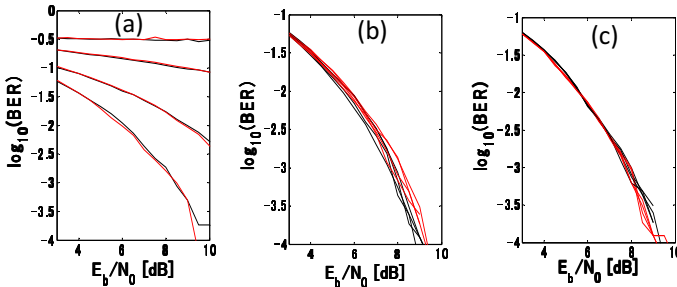


Fig. 27. Bit-error rates calculated as a function of E_b/N_0 for four initial sampling phases when we apply $\times 2$ oversampling [61]. The number of taps is 1 in (a), 3 in (b), and 5 in (c). In all the cases, the roll-off parameter of the anti-aliasing Nyquist filter is 0.2. The red and black curves correspond to the x -polarization tributary and y -polarization tributary, respectively.

G. Frequency-domain Fixed GVD Equalization

The computational complexity of FIR filters scales with the number of delay taps; therefore, it becomes difficult to implement FIR filters having a large number of delay taps in ASICs. On the other hand, frequency-domain equalization (FDE) can reduce this computational cost via block-by-block signal processing and efficient implementation of DFT. The single-carrier FDE concept is shown in Fig. 28 [62]. After serial data of the received complex amplitude are divided into blocks with the serial-to-parallel converter (S/P), each block in the time domain is transformed into the frequency domain with DFT. Then, the spectral components are controlled by

the frequency-domain equalizer, where the tap coefficients are adjusted so that the desired transfer function is obtained. In practical systems, the GVD value is monitored by a pilot tone, and then, the tap coefficients are determined from the monitored GVD value. The block in the frequency domain is transformed into the time domain with inverse DFT (IDFT) and converted into the serial data with the parallel-to-serial converter (P/S). The interference between blocks owing to GVD is usually addressed by the overlap-save method [63], [64].

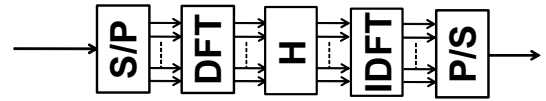


Fig. 28. Configuration of the frequency-domain equalizer. After serial data of the received complex amplitude are divided into blocks with the serial-to-parallel converter (S/P), each block in the time domain is transformed into the frequency domain with DFT. Each spectral component is then controlled by the transfer function H . The block equalized in the frequency domain is transformed into the time domain with inverse DFT (IDFT) and converted into serial data with the parallel-to-serial converter (P/S).

H. Carrier-phase Estimation

Because the linewidth of semiconductor DFB lasers used as the transmitter and LO typically ranges from 100 kHz to 10 MHz, the phase noise $\theta_n(t)$ varies much more slowly than the phase modulation $\theta_s(t)$. Therefore, by averaging the carrier phase over many symbol intervals, it is possible to obtain an accurate phase estimate. Figure 29 shows the effect of filtering the signal for phase estimation. In (a), the received signal is contaminated by ASE noise. In (b), optimum filtering of the signal improves the S/N ratio, whereas the phase fluctuation is unfiltered. On the other hand, in (c), excessively tight signal filtering eliminates the phase fluctuation.

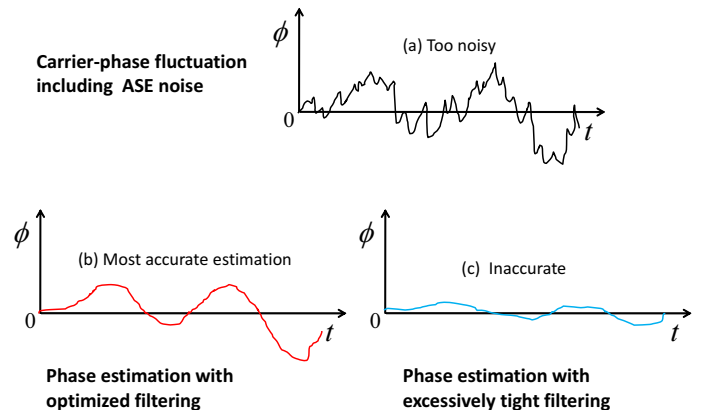


Fig. 29. Effect of filtering the signal for phase estimation. In (a), the received signal is contaminated by ASE noise. In (b), optimum filtering of the signal improves the S/N ratio, whereas the phase fluctuation is unfiltered. In (c), excessively tight filtering eliminates the phase fluctuation.

The phase of the complex amplitude obtained from Eq. (24) contains both the phase modulation $\theta_s(\ell)$ and the phase noise $\theta_n(\ell)$, where ℓ represents the number of samples. The

procedure for removing $\theta_s(\ell)$ and estimating $\theta_n(\ell)$ for M -ary PSK signals is shown in Fig. 30, where the case of QPSK is shown for simplicity [65]. We take the M th power of the measured complex amplitude $E(\ell)$, because the phase modulation is removed from $E(\ell)^M$ in the case of the M -ary PSK modulation format. Subtracting the phase noise thus estimated from the measured phase, we can restore the phase modulation. This is a feedforward algorithm, which is suitable for DSP implementation.

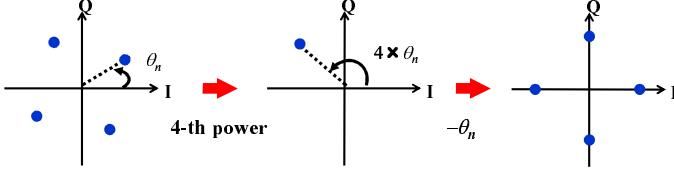


Fig. 30. Principle of the M th-power phase estimation method. For simplicity, the case of QPSK ($M = 4$) is shown. By taking the M th power of the received complex amplitude, we can eliminate the phase modulation and measure the phase noise.

However, note that this algorithm cannot directly be applied to high-order QAM formats. For example, in the case of 16QAM, we need to discriminate the level of each symbol and classify constellation points into two classes, as shown in Fig. 31 [66]. Class 1 represents the symbols on the diagonal lines of the constellation map, whereas Class 2 represents the other symbols. Because the symbols belonging to Class 1 are regarded as QPSK signals having two amplitude levels, phase estimation can be performed using the Class 1 symbols in the same manner as the QPSK signal. On the other hand, the Class 2 symbols are not used for phase estimation.

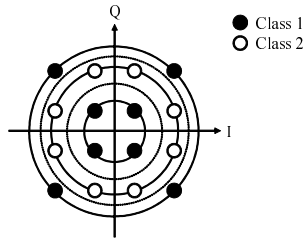


Fig. 31. Symbol classification for the 16QAM signal. The closed circles represent the symbols belonging to Class 1, which are used for phase estimation. Class 2 symbols, denoted by open circles, are disregarded in the phase-estimation process.

In the actual phase estimation, we average $E(\ell)^M$ over $(2k + 1)$ samples to improve the S/N ratio of the estimated phase reference. The estimated phase is thus given by

$$\theta_e(\ell) = \arg \left(\sum_{j=-k}^k E(\ell + j)^M \right) / M. \quad (120)$$

Using such a phase reference, the complex amplitude is rotated properly and discriminated among M symbols. Figure 32 shows the DSP circuit for such phase estimation. The optimum averaging span is dependent on the laser linewidth. When the linewidth is narrower, the optimum averaging span is longer, which gives us better BER performance.

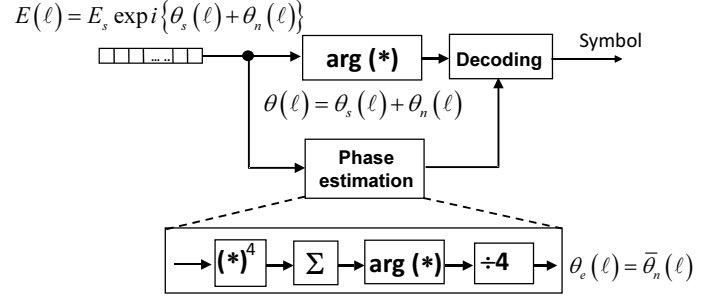


Fig. 32. DSP circuit for M -th power phase estimation. We average $E(\ell)^M$ to improve the S/N ratio of the estimated phase reference. Using the estimated phase reference, the complex amplitude is rotated properly and discriminated among M symbols.

The symbols thus obtained have phase ambiguity by $2\pi/M$. To cope with such phase ambiguity, the data should be differentially encoded at the transmitter. In such a case, by differentially decoding the discriminated symbol, we can solve the phase ambiguity problem, although the bit-error rate is doubled by error multiplication.

Figure 33 shows the simulation results of BER characteristics of QPSK signals, where the normalized laser linewidth $\delta f \cdot T$ is varied. The averaging span is optimized according to the linewidth to have the best BER characteristics and is shown in the inset. The power penalty at $\text{BER}=10^{-4}$ is less than 2 dB when $\delta f \cdot T$ is as large as 5×10^{-4} .

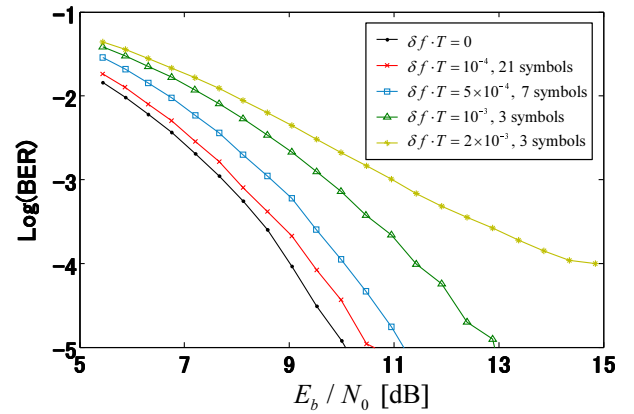


Fig. 33. Simulation results of BER characteristics of QPSK signals, where the normalized laser linewidth $\delta f \cdot T$ is changed. The averaging span is optimized to have the best BER characteristics.

Figure 34 shows the simulation results of BER characteristics of 16QAM signals, where the normalized laser linewidth $\delta f \cdot T$ is varied. The optimized averaging span is shown in the inset. The power penalty at $\text{BER}=10^{-4}$ is less than 2 dB when $\delta f \cdot T$ is 5×10^{-5} , which is one order of magnitude smaller than the linewidth required for the QPSK signal.

In the era of analog coherent optical communications, phase tracking in the homodyne receiver was achieved by OPLL, and the tracking speed was limited by the feedback delay. We needed to use narrow-linewidth lasers to obtain sufficient phase-locking characteristics with the limited feedback band-

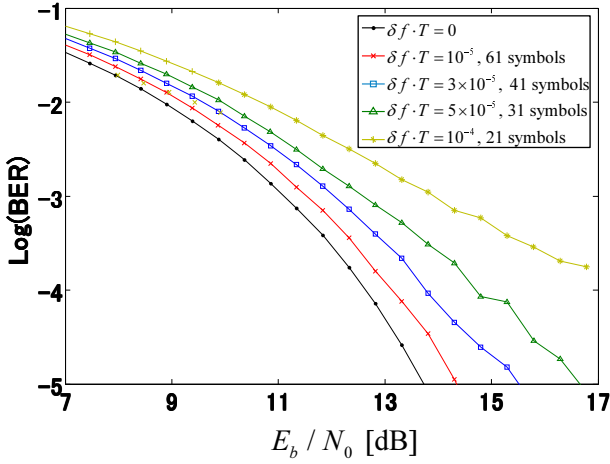


Fig. 34. Simulation results of BER characteristics of 16QAM signals, where the laser linewidth $\delta f \cdot T$ is changed. The averaging span is optimized to have the best BER characteristics.

width [67]. On the other hand, the digital coherent receiver improves the phase-tracking speed considerably, which relaxes the requirement for the laser linewidth.

I. Nonlinearity Compensation

The Kerr effect in optical fibers induces nonlinear waveform distortion [68] and strictly limits the maximum transmission distance of the high-order QAM signal [69]. Although compensation for the Kerr effect at the receiver is indispensable to enhance the transmission distance, the FIR-filter-based adaptive equalizer shown in Fig. 23 cannot deal with such nonlinear waveform distortion.

Basically, the digital coherent optical receiver can undo the Kerr-induced distortion by conducting virtual backpropagation in the digital domain, as shown in Fig. 35 [70], [71]. After coherent detection, the received signal is transmitted through virtual fibers and amplifiers, which are emulated in the digital domain so as to have the inverse characteristics of the actual transmission link. Signal propagation through a fiber is governed by the Manakov equation, given by

$$\begin{aligned} \frac{\partial E_{x,y}}{\partial z} = & -\frac{\alpha}{2} E_{x,y} - \frac{i}{2} \beta_2 \frac{\partial^2 E_{x,y}}{\partial T^2} \\ & + \frac{8i}{9} \gamma (|E_x|^2 + |E_y|^2) E_{x,y}, \end{aligned} \quad (121)$$

where E_x and E_y denote electric fields of the dual polarization signal with x - and y -polarization components, respectively, α is the loss coefficient, and γ represents the nonlinear coefficient. In these equations, randomly-distributed linear birefringence is assumed along the link [72]. On the other hand, by reversing the signs of α , β_2 , and γ , we can express the virtual backpropagation process as

$$\begin{aligned} \frac{\partial E_{x,y}}{\partial z} = & \frac{\alpha}{2} E_{x,y} + \frac{i}{2} \beta_2 \frac{\partial^2 E_{x,y}}{\partial T^2} \\ & - \frac{8i}{9} \gamma (|E_x|^2 + |E_y|^2) E_{x,y}. \end{aligned} \quad (122)$$

Note that the distributed gain is given to the fiber, whereas the lumped loss is given to the repeater amplifiers in the virtual

transmission system. Virtual backpropagation of the received signal with Eq. (122) in the digital domain can restore the input signal.

With such a scheme, deterministic transmission impairments such as GVD and the Kerr effect can be undone simultaneously. Because the virtual backpropagation equation is the nonlinear partial differential equation given by Eq. (122), the split-step Fourier method is used to solve the equation numerically [68]; however, its computational complexity is extremely high. Therefore, there is an urgent need for an improved algorithm that provides efficient compensation at a relatively low computational cost [73], [74].

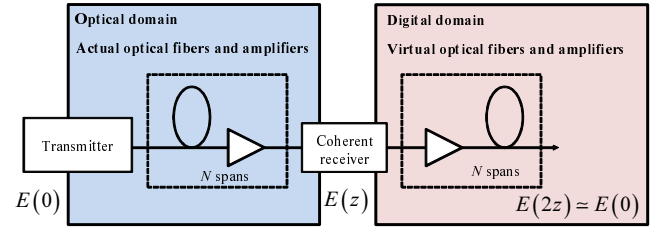


Fig. 35. Concept of the virtual backpropagation method for mitigating the nonlinear waveform distortion. The actually experienced loss/gain, GVD, and nonlinearity are undone in the DSP core.

Other than the backpropagation method, a number of nonlinearity equalizing schemes have been proposed and demonstrated. For example, the maximum a posteriori probability (MAP) detector [75] exploits the pattern dependence of nonlinear waveform distortions. The nonlinear Volterra equalizer is capable of compensating for nonlinear distortions [76]. The electronic predistortion at the transmitter can reduce nonlinear distortions while the signal travels through the fiber [77]. However, none of them has succeeded in achieving high performance with low computational cost for real-world applications.

VII. CONCLUSION

After reviewing the history of coherent optical communication systems, we have described the principle of coherent optical detection, including its quantum-noise characteristics. Next, we have shown how digital signal processing plays an important role in the newly developed digital coherent receiver, with special focus on FIR-filter-based adaptive equalization.

The combination of coherent detection and DSP provides us with new capabilities that were not possible without the detection of the phase of the optical signal. Now, 100-Gbit/s DP-QPSK digital coherent receivers are widely used in commercial core/metro networks. As the next step, bit rates as high as 400 Gbit/s have recently been demonstrated using the 16QAM format and two carriers. The developments of larger-scale photonic-integrated circuits and lower-power consumption ASICs may accelerate the application of digital coherent receivers to data-center networks and access networks in the near future.

REFERENCES

- [1] T. Okoshi and K. Kikuchi, *Coherent Optical Communication Systems*, Tokyo: KTK Scientific Publisher, Dordrecht, Boston, London, Tokyo: Kluwer Academic Publisher, 1988, ch. 1.
- [2] S. Tsukamoto, D.-S. Ly-Gagnon, K. Katoh, and K. Kikuchi, "Coherent demodulation of 40-Gbit/s polarization-multiplexed QPSK signals with 16-GHz spacing after 200-km transmission," presented at Optical Fiber Communication Conference (OFC 2005), Anaheim, CA, USA, Mar. 6-11, 2005, Paper PDP29.
- [3] K. Kikuchi, "Digital coherent optical communication systems: Fundamentals and future prospects," *IEICE Electron. Exp.*, vol.8, no.20, pp.1642-1662, Oct. 2011.
- [4] K. Kikuchi, "Coherent optical communication technology," presented at Optical Fiber Communications Conference (OFC 2015), Los Angeles, CA, USA, March 22-26, 2015, Paper Th4F.4.
- [5] O. E. DeLange, "Wide-band optical communication systems: Part II-Frequency-division multiplexing," *Proc. IEEE*, vol.58, no.10, pp.1683-1690, Oct. 1970.
- [6] T. Okoshi and K. Kikuchi, "Frequency stabilization of semiconductor lasers for heterodyne-type optical communication systems," *Electron. Lett.*, vol.16, no.5, pp.179-181, Feb. 1980.
- [7] F. Favre and D. LeGuen, "High frequency stability of laser diode for heterodyne communication systems," *Electron. Lett.*, vol.16, no.18, pp.709-710, Aug. 1980.
- [8] T. Okoshi, K. Kikuchi, and A. Nakayama, "Novel method for high resolution measurement of laser output spectrum," *Electron. Lett.*, vol.16, no.16, pp. 630-631, Jul. 1980.
- [9] C. H. Henry, "Theory of the linewidth of semiconductor lasers," *IEEE J. Quantum Electron.*, vol.QE-18, no.2, pp. 259-264, Feb. 1982.
- [10] Y. Yamamoto, "AM and FM quantum noise in semiconductor lasers-Part I: Theoretical analysis," *IEEE J. Quantum Electron.*, vol.QE-19, no.1, pp.34-46 Jan. 1983.
- [11] K. Vahala and A. Yariv, "Semiclassical theory of noise in semiconductor lasers-Part I," *IEEE J. Quantum Electron.*, vol.QE-19, no.6, pp. 1096-1101, Jun. 1983.
- [12] K. Kikuchi and T. Okoshi, "Measurement of FM noise, AM noise, and field spectra of 1.3 μm InGaAsP DFB lasers and determination of the linewidth enhancement factor," *IEEE J. Quantum Electron.*, vol.QE-21, no.11, pp.1814-1818, Nov. 1985.
- [13] T. Okoshi, S. Ryu, and K. Kikuchi, "Polarization-diversity receiver for heterodyne/coherent optical fiber communications," presented at International Conference on Integrated Optics and Optical Fiber Communication (IOOC'83), Tokyo, Japan, Jun. 27-30, 1983, Paper 30C3-2.
- [14] B. Glance, "Polarization independent coherent optical receiver," *J. Lightw. Technol.*, vol.LT-5, no.2, pp. 274-276, Feb. 1987.
- [15] N. Shibata, Y. Sasaki, K. Okamoto, and T. Hosaka, "Fabrication of polarization-maintaining and absorption-reducing fibers," *J. Lightw. Technol.*, vol.LT-1, no.1, pp.38-43, March 1983.
- [16] L. Jeunhomme and N. Monerie, "Polarization-maintaining single-mode fiber cable design," *Electron. Lett.*, vol.16, no.24, pp.921-922, Nov. 1980.
- [17] T. Tanemura and K. Kikuchi, "Circular-birefringence fiber for nonlinear optical signal processing," *J. Lightw. Technol.*, vol.LT-24, no.11, pp.4108-4119, Nov. 2006.
- [18] Y. Yamamoto, *IEEE J. Quantum Electron.*, "Receiver performance evaluation of various digital optical modulation-demodulation systems in the 0.5-10 μm wavelength region," *IEEE J. Quantum Electron.*, vol.QE-16, no.11, pp.1251-1259, Nov. 1980.
- [19] T. Okoshi, K. Emura, K. Kikuchi, and R. Th. Kersten, "Computation of bit-error rate of various heterodyne and coherent-type optical communication schemes," *J. Opt. Commun.*, vol.2, no.4, pp.134-141, Sep. 1981.
- [20] T. Imai, Y. Hayashi, N. Ohkawa, T. Sugie, Y. Ichihashi, and T. Ito, "Field demonstration of 2.5 Gbit/s coherent optical transmission through installed submarine fibre cables," *Electron. Lett.*, vol.26, no.17, pp.1407-1409, Aug. 1990.
- [21] S. Norimatsu, K. Iwashita, and K. Sato, "PSK optical homodyne detection using external cavity laser diodes in Costas loop," *IEEE Photon. Technol. Lett.*, vol.2, no.5, pp.374-376, May 1990.
- [22] R.J. Mears, L. Reekie, I.M. Jauncey, and D. N. Payne, "Low-noise Erbium-doped fibre amplifier at 1.54 μm ," *Electron. Lett.*, vol.23, no.19, pp.1026-1028, Sep. 1987.
- [23] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems," *IEEE J. Sel. Top. in Quantum Electron.*, vol.10, no.2, pp.259-272, Mar./Apr. 2004.
- [24] R. Griffin and A. Carter, "Optical differential quadrature phase-shift key (oDQPSK) for high capacity optical transmission," presented at Optical Fiber Communication Conference (OFC 2002), Anaheim, CA, USA, Mar. 17-22, 2002, Paper WX6.
- [25] S. Shimotsu, S. Oikawa, T. Saitou, N. Mitsugi, K. Kubodera, T. Kawanishi, and M. Izutsu, "Single side-band modulation performance of a LiNbO₃ integrated modulator consisting of four-phase modulator waveguides," *IEEE Photon. Technol. Lett.*, vol.13, no.4, pp.364-366, Apr. 2001.
- [26] D. MacGhan, C. Laperle, A. Savchenko, C. Li, G. Mak, and M. O'Sullivan, "5120 km RZ-DPSK transmission over G652 fiber at 10 Gb/s with no optical dispersion compensation," presented at Optical Fiber Communication Conference (OFC 2005), Anaheim, CA, USA, Mar. 6-11, 2005, Paper PDP27.
- [27] 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.LT-24, no.1, pp.12-21, Jan. 2006.
- [28] K. Kikuchi, "Phase-diversity homodyne detection of multi-level optical modulation with digital carrier phase estimation," *IEEE J. Sel. Top. in Quantum. Electron.*, vol.12, no.4, pp.563-570, Jul./Aug. 2006.
- [29] F. Derr, "Optical QPSK transmission system with novel digital receiver concept," *Electron. Lett.*, vol.27, no.23, pp.2177-2179, Nov. 1991.
- [30] R. Noé, "Phase noise tolerant synchronous QPSK receiver concept with digital I&Q baseband processing," presented at Opto-Electronics and Communications Conference (OECC 2004), Yokohama, Japan, Jul. 12-16, 2004, Paper 16C2-5.
- [31] S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent demodulation of optical multilevel phase shift-keying signals using homodyne detection and digital signal processing," *IEEE Photon. Technol. Lett.*, vol.18, no.10, pp.1131-1133, May 2006.
- [32] Y. Mori, C. Zhang, K. Igarashi, K. Katoh, and K. Kikuchi, "Unrepeated 200-km transmission of 40-Gbit/s 16-QAM signals using digital coherent receiver," *Opt. Exp.*, vol.17, no.32, pp.1435-1441, Feb. 2009.
- [33] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol.16, no.2, pp.674-676, Feb. 2004.
- [34] S. Tsukamoto, K. Katoh, and K. Kikuchi, "Unrepeated transmission of 20-Gbit/s optical quadrature phase-shift keying signal over 200-km standard single-mode fiber based on digital processing of homodyne-detected signal for group-velocity dispersion compensation," *IEEE Photon. Technol. Lett.*, vol.18, no.9, pp.1016-1018, May 2006.
- [35] S. Tsukamoto, Y. Ishikawa, and K. Kikuchi, "Optical homodyne receiver comprising phase and polarization diversities with digital signal processing," presented at European Conference on Optical Communication (ECOC 2006), Cannes, France, Sep. 24-28, 2006, paper Mo4.2.1.
- [36] K. Kikuchi and S. Tsukamoto, "Evaluation of sensitivity of the digital coherent receiver," *J. Lightw. Technol.*, vol.LT-20, no.13, pp.1817-1822, Jul. 2008.
- [37] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol.16, no.2, pp.804-817, Jan. 2008.
- [38] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40-Gb/s coherent system," *Opt. Exp.*, vol.16, no.2, pp.873-879, Jan. 2008.
- [39] L. E. Nelson, S. L. Woodward, M. D. Feuer, X. Zhou, P. D. Magill, S. Foo, D. Hanson, D. McGhan, H. Sun, M. Moyer, and M.O'Sullivan, "Performance of a 46-Gbps dual-polarization QPSK transceiver in a high-PMD fiber transmission experiment," presented at Optical Fiber Communication Conference (OFC 2008), San Diego, CA, USA, Feb. 24-28, 2008, Paper PDP9.
- [40] E. Yamazaki, S. Yamanaka, Y. Kisaka, T. Nakagawa, K. Murata, E. Yoshida, T. Sakano, M. Tomizawa, Y. Miyamoto, S. Matsuoka, J. Matsui, A. Shibayama, J. Abe, Y. Nakamura, H. Noguchi, K. Fukuchi, H. Onaka, K. Fukumitsu, K. Komaki, O. Takeuchi, Y. Sakamoto, H. Nakashima, T. Mizuochi, K. Kubo, Y. Miyata, H. Nishimoto, S. Hirano, and K. Onohara, "Fast optical channel recovery in field demonstration of 100-Gbit/s Ethernet over OTN using real-time DSP," *Opt. Exp.*, vol.19, no.14, pp.13139-13184, Jul. 2011.
- [41] K. Kikuchi, "Coherent optical communications: Historical perspectives and future directions," in *High Spectral Density Optical Communication Technology*, edited by M. Nakazawa, K. Kikuchi, and T. Miyazaki, Heidelberg, Dordrecht, London, New York: Springer Verlag, 2010, ch. 2.
- [42] H. Hodgkinson, R. A. Harmon, and D. W. Smith, "Demodulation of optical DPSK using in-phase and quadrature detection," *Electron. Lett.*, vol.21, no.19, pp. 867-868, Sep. 1985.
- [43] A. W. Davis, M. J. Pettitt, P. J. King, and S. Wright, "Phase diversity

- techniques for coherent optical receivers," *J. Lightw. Technol.*, vol.LT-5, no.4, pp.561-572, Apr. 1987.
- [44] Y. Yamamoto and H. Haus, "Preparation, measurement and information capacity of optical quantum states," *Rev. Mod. Phys.*, vol.58, no.4, pp.1001-1020, Oct. 1986.
- [45] H. P. Robertson, "The uncertainty principle," *Phys. Rev.*, vol.34, no.1, pp.163-164, Jul. 1929.
- [46] F. Koyama and K. Iga, "Frequency chirping in external modulators," *J. Lightw. Technol.*, vol.LT-6, no.1, pp.87-93, Jan. 1988.
- [47] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol.27, no.3, pp.379-423, Jul. 1948 and no.4, pp.623-656, Oct. 1948.
- [48] S. Ishimura and K. Kikuchi, "Multi-dimensional permutation-modulation format for coherent optical communications," *Opt. Exp.*, vol.23, no.12, pp.15587-15597, June 2015.
- [49] J. Proakis, *Digital Communications, 4th Ed.*, New York: McGraw Hill, 2007, ch. 7.
- [50] M. Karlsson and E. Agrell, "Which is the most power-efficient modulation format in optical links?" *Opt. Exp.*, vol.17, no.13, pp.10814-10819, Jun. 2009.
- [51] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol.LT-27, no.22, pp.5115-5126, Nov. 2009.
- [52] L. Coelho and N. Hanik, "Global optimization of fiber-optic communication systems using four-dimensional modulation formats," presented at European Conference on Optical Communication (ECOC 2011), Geneva, Switzerland, Sep. 18-22, 2011, Paper Mo.2.B.
- [53] T. Mizuochi, "Next generation FEC for optical communication," presented at Optical Fiber Communication Conference (OFC 2008), San Diego, CA, USA, Feb. 24-28, 2008, Paper OTuE5.
- [54] K. Kikuchi, "Performance analyses of digital coherent receivers with symbol-rate analog-to-digital conversion," presented at Signal Processing in Photonics Communications (SPPCom 2015), Boston MA, USA, June 27-July 1, 2015, Paper SpT3D.4.
- [55] S. Haykin, *Introduction to Adaptive Filters*, New York: Macmillan Publishing Company, 2001, ch. 4.
- [56] Md. S. Faruk, Y. Mori, C. Zhang, K. Igarashi, and K. Kikuchi, "Multi-impairment monitoring from adaptive finite-impulse-response filters in a digital coherent receiver," *Opt. Exp.*, vol.18, no.26, pp.26929-26936, Dec. 2010.
- [57] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol.73, no.9, pp.1349-1387, Sep. 1985.
- [58] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol.COM-28, no.11, pp.1867-1875, Nov. 1980.
- [59] K. Kikuchi, "Performance analyses of polarization demultiplexing based on constant-modulus algorithm in digital coherent optical receivers," *Opt. Exp.*, vol.19, no.10, pp.9868-9880, May 2011.
- [60] Y. Mori, C. Zhang, and K. Kikuchi, "Novel configuration of finite-impulse-response filters tolerant to carrier-phase fluctuations in digital coherent optical receivers for higher-order quadrature amplitude modulation signals," *Opt. Exp.*, vol.20, no.24, pp.26236-26251, Nov. 2012.
- [61] K. Kikuchi, "Clock recovering characteristics of adaptive finite-impulse-response filters in digital coherent optical receivers," *Opt. Exp.*, vol.19, no.6, pp.5611-5619, Mar. 2011.
- [62] F. Pancaldi, G. M. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Processing Magazine*, pp.37-56, Sep. 2008.
- [63] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, and Y. Miyamoto, "Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems," *J. Lightw. Technol.*, vol.LT-27, no.16, pp.3721-3718, Aug. 2009.
- [64] Md. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent receivers," *Opt. Exp.*, vol.19, no.13, pp.12789-12798, June 2011.
- [65] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol.IT-29, no.4, pp.543-551, Jul. 1983.
- [66] M. Seimetz, "Performance of coherent optical square-16-QAM-systems based on IQ-transmitters and homodyne receivers with digital phase estimation," presented at Optical Fiber Communication Conference (OFC 2006), Anaheim, CA, USA, Mar.5-10, 2006, Paper NWA4.
- [67] K. Kikuchi, T. Okoshi, M. Nagamatsu, and N. Henmi, "Degradation of bit-error rate in coherent optical communications due to spectral spread of the transmitter and the local oscillator," *J. Lightw. Technol.*, vol.LT-2, no.6, pp.1024-1033, Dec. 1984.
- [68] G. P. Agrawal, *Nonlinear Fiber Optics*, San Diego: Academic Press, 1989, ch. 2.
- [69] K. Kikuchi, "Analyses of wavelength- and polarization-division multiplexed transmission characteristics of optical quadrature-amplitude-modulation signals," *Opt. Exp.*, vol.19, no.19, pp.17985-17995, Sep. 2011.
- [70] K. Kikuchi, "Electronic post-compensation for nonlinear phase fluctuations in a 1000-km 20-Gbit/s optical quadrature phase-shift keying transmission system using the digital coherent receiver," *Opt. Exp.*, vol.16, no.2, pp.889-896, Jan. 2008.
- [71] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol.LT-26, no.20, pp.3416-3425, Oct. 2008.
- [72] S. G. Evangelides Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightw. Technol.*, vol.LT-10, no.1, pp.28-35, Jan. 1992.
- [73] L. Dou, Y. Zhao, Z. Tao, Y. Fan, T. Hoshida, and J. C. Rasmussen, "Advanced digital nonlinear distortion compensation," presented at Optical Fiber Communication Conference (OFC 2015), Los Angeles, CA, USA, Mar. 22-26, 2015, Paper Th3G.3.
- [74] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photonics*, vol.7, no.7, pp.560-568, May 2013.
- [75] Y. Cai, D. G. Foursa, C. R. Davidson, J. X. Cai, O. Sinkin, M. Nissov, and A. Philipetskii, "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmission," presented at Optical Fiber Communication Conference (OFC 2010), San Diego, CA, USA, Mar. 21-25, 2010, Paper OTuE1.
- [76] F. P. Guimar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," *Opt. Exp.*, vol.20, no.2, pp.1360-1368, Jan. 2012.
- [77] K. Roberts, Li Chuandong, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *IEEE Photon. Technol. Lett.*, vol.18, no.2, pp.403-405, Jan. 2006.