A Differential Coding Method for the Symmetrically Differential Polarization Shift-Keying System

Kuen-Suey Hou, *Member, IEEE,* and Jingshown Wu, *Senior Member, IEEE*

***Abstract*—This paper presents a novel differential coding method for the symmetrically differential polarization shift-keying system. The proposed differential coding method, which can overcome the slow polarization variation in the fiber, constructs the reference frame by setting the sum, the difference, and the cross product of previous two symbols as the first, second, and the last axes of the signal space. Obviously, each axis is determined by the previous two symbols together. Thus, at the receiver, the new scheme constructs a noisy reference frame more accurately than other schemes and yields better performance. The optimal constellation for the system is symmetric and easy to find. The analytical integral form for the bit error rate (BER) of the proposed system is derived, and the saddle point method is applied to obtain the analytical results. It is found that the performance in terms of BER is much better than that of other differential polarization shift-keying systems. The analytical results agree with the simulation very well.**

***Index Terms*—Communication system performance, modulation/demodulation, optical communication, polarization, polarization shift keying (PolSK).**

I. Introduction

THE overwhelming Internet traffic calls for a large amount of bandwidth. A lot of researchers believe that optical communication is the best solution [1]–[3], because of huge bandwidth available in optical fibers. In addition to the conventional intensity modulation/direct detection (IM/DD) techniques, the multilevel modulation is also attractive as they allow for reduction in the signal spectrum of the electrical transceiver for a given bit rate. The narrower signal spectrum also means a larger symbol duration. These two factors can lessen a lot of the impact of fiber dispersion effects on the system. The price to be paid for signal spectrum reduction is an increased sensitivity to noise. Of course, any system that minimizes the performance degradation would have great potential applications.

Conventional multilevel modulations spread its symbols on a plane. Recently, multidimensional modulations are available [4]–[8]. Among them, the polarization shift keying (PolSK) and Stokes parameters shift keying (SPSK) are most attractive [4]–[7], because of their phase-noise insensitive property and good performance. These schemes use the states of polarization (SOPs) as the modulation parameters. Since the SOP is fully represented by three Stokes parameters [9], the symbol constellation is spread over a three-dimensional (3-D) space instead of a conventional two-dimensional (2-D) plane. These systems outperform multilevel differential phase-shift keying (DPSK) systems at higher order constellations [5]. However, the random and slowly time-varying fluctuation of the SOP at the receiving end [10] requires the system to employ the polarization tracking circuits [11]. An alternative technique to solve the problem, which is also phase-noise insensitive and offers a potentially simple structure and low startup delay, is differential polarization shift keying (DPolSK). The family consists of the double differential PolSK (DDPolSK) [12] and the Gram–Schmidt DPolSK (GDPolSK) [13]. These systems encode the information in a relative position of the present symbol with respect to the reference frame constructed by the two previous symbols. Because polarization variation is very slow, the relative positions of three consecutive symbols are preserved while the signal propagates through the fiber. In this way, no SOP tracking circuits are ever needed. This system may be particularly attractive in the development of local area networks (LANs), which require a large number of relatively inexpensive transceivers.

The GDPolSK scheme generalizes the concept of conventional DDPolSK, and provides a more efficient and easier design method for the symbol constellation. However, the optimized constellation in the GDPolSK is difficult to find, because it is nonsymmetric and depends on the angle between the previous two symbols.

In this paper, we show that there is no longer an optimization problem when a novel differential coding scheme is applied in constructing the reference coordinates instead of the conventional Gram–Schmidt algorithm. The novel scheme sets the sum of the previous two symbols as the first axis, the difference of them as the second axis, and the cross product of them as the third axis. In the new scheme, each reference axis is determined by the previous two symbols together. So we can expect that the novel frame-constructing scheme provides a better estimation for a noisy reference frame than the conventional DPolSKs. In addition, the optimal constellation in the new scheme is symmetric. Therefore, we call it symmetrically DPolSK (SDPolSK).

This paper presents the analysis of SDPolSK systems. We assume that the product of the signal spectrum bandwidth and the transmission distance is smaller than the limit imposed by the

Paper approved by R. Hui, the Editor for Optical Transmission and Switching of the IEEE Communications Society. Manuscript received April 10, 2000; revised November 30, 2001. This work was supported in part by the National Science Council and Ministry of Education, Taiwan, R.O.C., under Grant NSC 89-2213-E-002-061 and Grant 89E-FA06-2-4-7.

K.-S. Hou was with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C. He is now with Mediatek Inc., Hsinchu 300, Taiwan, R.O.C.

J. Wu is with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C.

Digital Object Identifier 10.1109/TCOMM.2002.807611

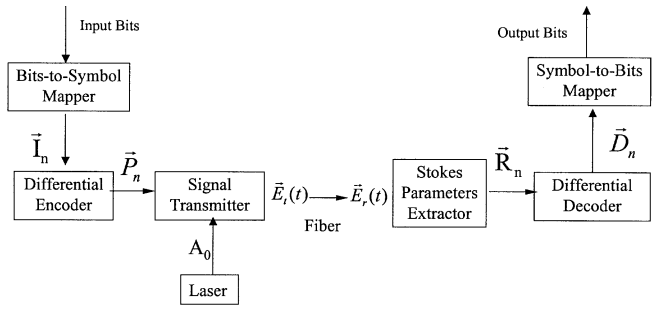

Fig. 1. Block diagram for DPolSK systems.

dispersion effects, including mainly chromatic dispersion and polarization mode dispersion (PMD), such that these effects can be neglected in the analysis. Usually, the chromatic dispersion effect can be reduced significantly by applying dispersion compensation techniques, such as linearly chirped Bragg gratings [14]. The PMD effect, which has emerged as one of the critical hurdles in high-speed, long-haul transmission systems [15], is expected to cause intersymbol interference (ISI) in the systems, and thus, degrades the performance. Note that in high-speed, long-haul SDPolSK systems, because the symbol duration at a given bit rate is longer than that of conventional binary IM/DD systems, the impact of dispersion effects on SDPolSK systems is much smaller. Surely, how robust the SDPolSK is to PMD is a meaningful extension of the work on high-speed, long-haul transmission systems, but in this paper, we concentrate on the new coding method and leave this topic for further study.

The remainder of this paper is organized as follows. Section II provides a brief review on the DPolSK system. The conventional and novel differential coding schemes are described in Section III. To make clear how the novel scheme performs, we derive the integral form for the bit error rate (BER) of the system in Section IV. As an example, an eight-symbol SDPolSK (8-SDPolSK) in cubic constellation is analyzed in Section V, where the saddle point method is applied to approximate the integral values for the BERs. The approximation results agree with the simulation very well. Section VI shows that the optimal constellation is symmetric and can be found easily. Finally, the conclusions are given in Section VII.

II. DPolSK

Generally speaking, the M-ary DPolSK system can be modeled as shown in Fig. 1. First, in the bits-to-symbol mapper, every $\log_2 \mathrm{M}$ input bits are transformed to a 3-D unit vector $\vec{I}$. Then, the following differential encoder block generates the reference coordinates, $\vec{e}_{tx}$, $\vec{e}_{ty}$, and $\vec{e}_{tz}$, and encodes the information vector $\vec{I}$ into its output symbol $\vec{P}$. Their relation can be expressed as

$$\vec{P} = [\vec{e}_{tx}, \vec{e}_{ty}, \vec{e}_{tz}]\,\vec{I}. \tag{1}$$

The symbol $\vec{P}$ is the Stokes parameters of the transmitted electrical field $\vec{E}_t(t)$, and the signal transmitter block is responsible for the transformation [4]–[6]. The detailed structure of the signal transmitter can be found in the literature [4]–[6]. Assuming that the linearly polarized lightwave propagates in the fiber along the z direction, we denote $\vec{E}_t(t)$ as

$$\begin{aligned}\vec{E}_t(t) =& [E_{tx}(t), E_{ty}(t)]^{\mathrm{T}} \\ =& e^{j\omega t}[A_{tx}\cos\varphi_{tx} + jA_{tx}\sin\varphi_{tx}, A_{ty}\cos\varphi_{ty} \\ &+ jA_{ty}\sin\varphi_{ty}]^{\mathrm{T}}\end{aligned} \tag{2}$$

where $E_{tx}(t)$ and $E_{ty}(t)$ are the electrical fields in the x and y directions, and (A_{tx}, φ_{tx}) and (A_{ty}, φ_{ty}) are the amplitudes and phases of $E_{tx}(t)$ and $E_{ty}(t)$, respectively.

The symbol $\vec{P}$ is given as

$$\vec{P} = \frac{\left[A_{tx}^2 - A_{ty}^2,\, 2\,A_{tx}\,A_{ty}\,\cos\delta,\, 2\,A_{tx}\,A_{ty}\,\sin\delta\right]^{\mathrm{T}}}{A_0} \tag{3}$$

where $\delta = \varphi_{tx} - \varphi_{ty}$ and A_0 is the amplitude of $\vec{E}_t(t)$, i.e. $A_0^2 = A_{tx}^2 + A_{ty}^2$.

Neglecting the nonlinear effect in the fiber and with the low PMD effect assumption, we can express the received electrical field $\vec{E}_r(t)$, which results from the transmitted $\vec{E}_t(t)$ propagating through the fiber, as [8]

$$\vec{E}_r(t) = e^{-[a+jb(\omega)]}\mathbf{J}\vec{E}_t(t) \tag{4}$$

and

$$\mathbf{J} = \begin{bmatrix} u_1 & u_2 \\ -u_2^* & u_1^* \end{bmatrix} \tag{5}$$

where a is the fiber attenuation, $b(\omega)$ is the fiber phase shift, $*$ denotes the complex conjugate, and $\mathbf{J}$ is the Jones matrix. $\mathbf{J}$ is a unitary operator which takes into account the polarization variation along the fiber due to coupling between the two polarization modes [10]. Under the assumption of low chromatic dispersion effect, in the interested range of ω, $b(\omega)$ is independent of ω. The parameter a does not influence the result of the analysis, and hereafter, we will ignore these terms.

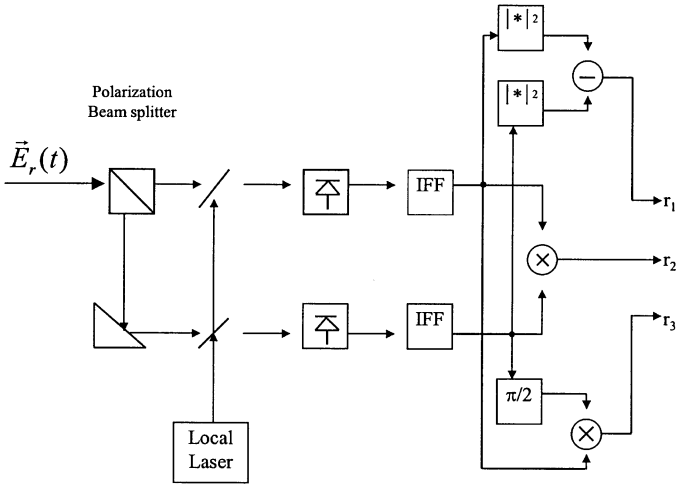

Fig. 2. Front-end receiver of the DPolSK system.

The front end of the DPolSK receiver, which is a Stokes parameters extractor for the received electrical $\vec{E}_r(t)$, is shown in Fig. 2 [3]–[6]. The intermediate frequency filter (IFF) is an intermediate frequency (IF) filter, and $\vec{R} \equiv [r_1, r_2, r_3]^{\mathrm{T}}$ is the Stokes parameter of $\vec{E}_r(t)$, whose direct-detection structure is

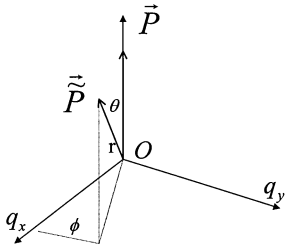

Fig. 3. Deviation of the received noisy vector $\vec{\tilde{P}}$ from the noise-free vector $\vec{P}$ in terms of (r, θ, ϕ).

also available [7]. We usually denote the noise induced in the detectors as $\vec{n} = \left[n_x^I + jn_x^Q,\, n_y^I + jn_y^Q\right]^{\mathrm{T}} \cdot n_x^I,\, n_x^Q,\, n_y^I$, and n_y^Q can be modeled as four independent identically distributed (i.i.d.) Gaussian noise processes with zero mean and variance σ^2.

After normalization, we assume that A_0 is equal to one and $\vec{E}_r(t)$ is $(\mathbf{J}\vec{E}_t(t) + \vec{n})$. For simplicity, we denote the Stokes parameters of $\mathbf{J}(\vec{E}_t(t) + \vec{n})$ and $(\vec{E}_t(t) + \vec{n}))$ as $\vec{\tilde{P}}^J$ and $\vec{\tilde{P}}$, respectively. Note that because the Jones matrix is unitary, the probability distribution functions of $\vec{\tilde{P}}^J$ and $\vec{R}$ are the same [4].

The multiplication of the Jones matrix $\mathbf{J}$ on the lightwave field implies a rotation of the reference axes in the Stokes space [4]. That is

$$\vec{\tilde{P}}^J = \mathbf{T}\vec{\tilde{P}} \tag{6}$$

where $\mathbf{T}$ is a three-by-three unitary matrix.

Similarly, we assume that the Stokes parameters of $(\mathbf{J}\vec{E}_t(t))$ is $\vec{R}^f$. Then, $\vec{R}^f$ is equal to $\mathbf{T}\vec{P}$. From the discussion above, we conclude that the distribution function of the deviation of $\vec{R}$ from $\vec{R}^f$ is the same as that of $\vec{\tilde{P}}$ from $\vec{P}$. In other words, the effect of the Jones matrix can be totally ignored in the analysis. Usually, we use the spherical coordinates $(r\ , \theta\ , \phi)$ to express the deviation of the received noisy vector $\vec{\tilde{P}}$ from the noise-free one $\vec{P}$, as shown in Fig. 3, where r is the magnitude of $\vec{\tilde{P}}$, θ is the angle between $\vec{\tilde{P}}$, and $\vec{P}$, ϕ is the angle made by $\overline{Oq_x}$ and the projection vector of $\vec{\tilde{P}}$ on the $q_x - \mathrm{O} - q_y$ plane, $q_x - \mathrm{O} - q_y$ plane is normal to $\vec{P}$, and $\overline{Oq_x}$ can be arbitrarily chosen on the plane [4]. In general, we use the subscript $n-i$ in the notations, $\vec{R}_{n-i}^f$, $\vec{R}_{n-i}$, $\vec{\tilde{P}}_{n-i}$, and $\vec{P}_{n-i}$, to represent the symbols in the time slot $n-i$. The associate deviation of $\vec{\tilde{P}}_{n-i}$ from $\vec{P}_{n-i}$ can be expressed by a function of (r_i, θ_i, ϕ_i). The probability density function of the deviation in terms of (r_i, θ_i, ϕ_i) is given by [4]

$$f_i(r_i, \theta_i, \phi_i) = \frac{1}{16\pi\sigma^4} r_i \sin\theta_i \times \exp\left(-\frac{S_{0i} + r_i}{2\sigma^2}\right) I_0\left(\frac{\sqrt{S_{0i}}}{\sigma^2}\sqrt{r_i}\cos\frac{\theta_i}{2}\right) \tag{7}$$

where $S_{0i} \equiv A_0^2 = 1$ in the system, and $\mathrm{I}_0(.)$ is the modified Bessel function of the first kind of order zero. Equation (7) is valid, irrelevant of the position of the original vector $\vec{P}_{n-i}$.

Because the polarization fluctuation of the fiber is very slow [10], the rotation matrix $\mathbf{T}$ in the Stokes space imposed by the fiber on any three consecutive vectors, $\vec{P}_{n-2}$, $\vec{P}_{n-1}$, and $\vec{P}_n$, can be assumed as the same. Therefore, the relative position of them is not altered during the propagation in the fiber. Due to this property, the differential decoder block can recover the estimated information vector $\vec{D}_n$ from the relative position of $\vec{R}_{n-2}$, $\vec{R}_{n-1}$, and $\vec{R}_n$. The estimated information vector $\vec{D}_n$ is exactly equal to the transmitted information vector $\vec{I}_n$ when the noises are absent. Of course, the noisy reference frame, $\vec{e}_{rx}$, $\vec{e}_{ry}$, and $\vec{e}_{rz}$, generated in this block should be constructed in the same way as that in the encoder. Their relation is given by

$$\vec{D}_n = [\mathrm{D_{nx}},\, \mathrm{D_{ny}},\, \mathrm{D_{nz}}]^{\mathrm{T}} = [\vec{e}_{rx},\, \vec{e}_{ry},\, \vec{e}_{rz}]^{\mathrm{T}}\, \vec{R}_n. \tag{8}$$

Finally, the position of the noisy estimated information vector determines the output estimated bits of the symbol-to-bits mapper.

III. Differential Encoding/Decoding Methods

A. Conventional GDPolSK

The encoder and decoder apply the well-known Gram–Schmidt algorithm [16]. At the encoder, the reference axes are given as

$$\vec{e}_{tz} = \vec{P}_{n-1} \tag{9}$$

$$\vec{e}_{tx} = \text{Normalized}\left(\vec{P}_{n-2} - \vec{e}_{tz}\left(\vec{P}_{n-2} \cdot \vec{e}_{tz}\right)\right) \tag{10}$$

$$\vec{e}_{ty} = \vec{e}_{tz} \cdot \vec{e}_{tx}. \tag{11}$$

Denote $\vec{I}_n$ as $[\mathrm{I_{nx}},\, \mathrm{I_{ny}},\, \mathrm{I_{nz}}]^{\mathrm{T}}$. Equations (1) and (9)–(11) can be visualized as follows. The vector $\vec{P}_{n-1}$ is put on the positive $\mathrm{I_{nz}}$ axis, and $\vec{P}_{n-2}$ is laid on the $\mathrm{I_{nz}} - \mathbf{I}_{\mathrm{nx}}$ plane to construct a frame of reference, and the vector $\vec{P}_n$ is located in a position such that its relative position with respect to the frame represents exactly the information symbol $\vec{I}_n$.

At the decoder, the same Gram–Schmidt algorithm is applied. The estimated reference frame is formulated as

$$\vec{e}_{rz} = \text{Normalized}\left(\vec{R}_{n-1}\right) \tag{12}$$

$$\vec{e}_{rx} = \text{Normalized}\left(\vec{R}_{n-2} - \vec{e}_{rz}\left(\vec{R}_{n-2} \cdot \vec{e}_{rz}\right)\right) \tag{13}$$

$$\vec{e}_{ry} = \vec{e}_{rz} \cdot \vec{e}_{rx}. \tag{14}$$

Note that if the possible $\vec{I}_n$'s are chosen as $[1,0,0]^{\mathrm{T}}$, $[-1,0,0]^{\mathrm{T}}$, $[0,1,0]^{\mathrm{T}}$, and $[0,-1,0]^{\mathrm{T}}$, and $\vec{P}_0$ and $\vec{P}_1$ are set to $[0,0,1]^{\mathrm{T}}$ and $[1,0,0]^{\mathrm{T}}$, respectively, then this 4-GDPolSK degenerates to the 6-DDPolSK system.

B. Proposed Symmetrical DPolSK

At the encoder, the reference axes are given as

$$\vec{e}_{tx} = \text{Normalized}\left(\vec{P}_{n-1} + \vec{P}_{n-2}\right) \tag{15}$$

$$\vec{e}_{ty} = \text{Normalized}\left(\vec{P}_{n-1} - \vec{P}_{n-2}\right) \tag{16}$$

$$\vec{e}_{tz} = \vec{e}_{tx} \cdot \vec{e}_{ty}. \tag{17}$$

The visualization for the reference frame is shown in Fig. 4.

The decoder applies the same frame-constructing algorithm. The estimated reference frame is formulated as

$$\vec{T}_{n-1} = \text{Normalized}\left(\vec{R}_{n-1}\right) \tag{18}$$

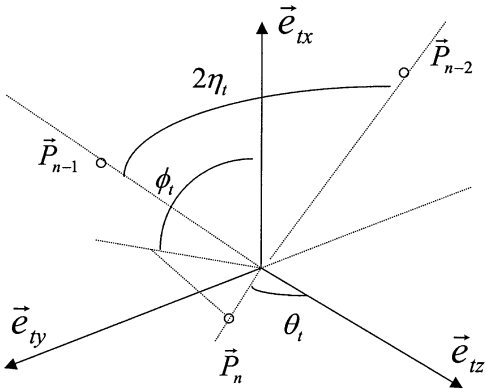

Fig. 4. Visualization of $(\eta_t, \theta_t, \phi_t)$.

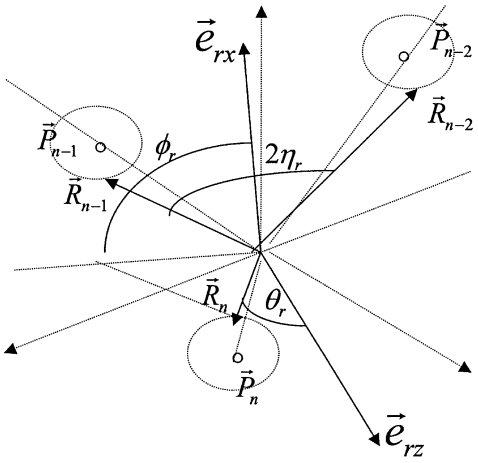

Fig. 5. Visualization of $(\eta_r, \theta_r, \phi_r)$.

$$\vec{T}_{n-2} = \text{Normalized}\left(\vec{R}_{n-2}\right) \tag{19}$$

$$\vec{e}_{rx} = \text{Normalized}\left(\vec{T}_{n-1} + \vec{T}_{n-2}\right) \tag{20}$$

$$\vec{e}_{ry} = \text{Normalized}\left(\vec{T}_{n-1} - \vec{T}_{n-2}\right) \tag{21}$$

$$\vec{e}_{rz} = \vec{e}_{rx} \cdot \vec{e}_{ry}. \tag{22}$$

IV. Analysis for the SDPolSK

To analyze the system performance, first we have to find the distribution of the relative positions of received signals. Note that from (6) and (7), and the assumption that the Stokes space rotation matrix $\mathbf{T}$ is unchanged during three symbol durations, the distribution of the relative position is independent of the channel Jones matrix. Therefore, we can totally ignore it in the analysis. That is, the discussion on the deviation of the relative position of $\vec{R}_{n-2}$, $\vec{R}_{n-1}$, and $\vec{R}_n$ with respect to $\vec{R}^f_{n-2}$, $\vec{R}^f_{n-1}$, and $\vec{R}^f_n$, is equivalent to that of $\vec{\tilde{P}}_{n-2}$, $\vec{\tilde{P}}_{n-1}$, and $\vec{\tilde{P}}_n$ with respect to $\vec{P}_{n-2}$, $\vec{P}_{n-1}$, and $\vec{P}_n$.

In general, whenever three consecutive noisy vectors, $\vec{\tilde{P}}_{n-2}$, $\vec{\tilde{P}}_{n-1}$, and $\vec{\tilde{P}}_n$ are received, the receiver can obtain six independent elements, $\vec{A}_r = (\mathrm{r}_2, \mathrm{r}_1, \mathrm{r}_0, \eta_\mathrm{r}, \theta_\mathrm{r}, \phi_\mathrm{r})$, where r_2, r_1, and r_0 are the amplitudes of $\vec{\tilde{P}}_{n-2}$, $\vec{\tilde{P}}_{n-1}$, and $\vec{\tilde{P}}_n$, respectively. The visualization of (η_r, θ_r, ϕ_r) is shown in Fig. 5. Note that in the general SDPolSK receiver, only $\mathrm{D_{nx}}$, $\mathrm{D_{ny}}$, and $\mathrm{D_{nz}}$ are used in the decision process, where $\mathrm{D_{nx}}$, $\mathrm{D_{ny}}$, and $\mathrm{D_{nz}}$ are equal to $r_2 \sin\theta_r \cos\phi_r$, $r_2 \sin\theta_r \sin\phi_r$, and $r_2 \cos\theta_r$, respectively. Similarly, for the noise-free case, $\vec{P}_{n-2}$, $\vec{P}_{n-1}$, and $\vec{P}_n$ give six relative position elements $A_\mathrm{t} = (\mathrm{S_{02}}, \mathrm{S_{01}}, \mathrm{S_{00}}, \eta_\mathrm{t}, \theta_\mathrm{t}, \phi_\mathrm{t})$, where $\mathrm{S_{0i}} \equiv A_0^2 = 1$ in the system, for i =1, 2, and 3. (η_t, θ_t, ϕ_t) are determined by the relative positions of $\vec{P}_{n-2}$, $\vec{P}_{n-1}$, and $\vec{P}_n$, as shown in Fig. 4. The distribution function of $\vec{A}_r$ given $\vec{A}_t$ is given as follows, and the detailed derivation is in the Appendix.

$$\mathrm{P}\left(\vec{A}_r | \vec{A}_t\right) = \iiint G(\vec{Q}) d\alpha d\beta d\gamma \tag{23}$$

where

$$\vec{Q} = (\mathrm{r}_2, \mathrm{r}_1, \mathrm{r}_0, \alpha, \beta, \gamma, \eta_\mathrm{r}, \theta_\mathrm{r}, \phi_\mathrm{r}) \tag{24}$$

$$\begin{aligned} \mathrm{G}(\vec{Q}) = {} & 2\sin\theta_2 \sin 2\eta_r \sin\theta_r \\ & \times \left[\prod_{1=0}^{2}\left[\frac{1}{16\pi\sigma^4} r_i \exp\left(-\frac{\mathrm{S_{0}}i + r_i}{2\sigma^2}\right)\right.\right. \\ & \left.\left. \times\ I_0\left(\frac{\sqrt{\mathrm{S_{0}}i}}{\sigma^2}\sqrt{r_i}\cos\frac{\theta_i(\vec{Q})}{2}\right)\right]\right] \end{aligned} \tag{25}$$

and

$$\theta_2(\vec{Q}) = \alpha - \eta_\mathrm{r} + \eta_t \tag{26}$$

$$\begin{aligned} \cos\theta_1(\vec{Q}) = {} & \\ & [\cos 2\eta_t \ \sin 2\eta_t \cos\beta - \sin 2\eta_t \sin\beta] \\ & \times \begin{bmatrix} -\cos\gamma\sin(2\eta_r)\sin(\theta_2) \\ +\cos(2\eta_r)\cos(\theta_2) \\ \cos\gamma\sin(2\eta_r)\cos(\theta_2) \\ +\cos(2\eta_r)\sin(\theta_2) \\ -\sin\gamma\sin(2\eta_r) \end{bmatrix} \end{aligned} \tag{27}$$

and

$$\begin{aligned} & \cos\theta_0(\vec{Q}) \\ & = [\sin\theta_t\cos(\phi_t+\eta_t) \ \ \sin\theta_t\sin(\phi_t+\eta_t) \ \ \cos\theta_t] \\ & \cdot \begin{bmatrix} \cos\theta_2 & -\sin\theta_2\cos\gamma & -\sin\theta_2\sin\gamma \\ \cos\beta\sin\theta_2 & \cos\beta\cos\theta_2\cos\gamma & \cos\beta\cos\theta_2\sin\gamma \\ & +\sin\beta\sin\gamma & -\sin\beta\cos\gamma \\ \sin\beta\sin\theta_2 & \sin\beta\cos\theta_2\cos\gamma & \sin\beta\cos\theta_2\sin\gamma \\ & -\cos\beta\sin\gamma & +\cos\beta\cos\gamma \end{bmatrix} \\ & \cdot \begin{bmatrix} \sin\theta_2\cos(\phi_r+\eta_r) \\ \sin\theta_r\sin(\phi_r+\eta_r) \\ \cos\theta_r \end{bmatrix}. \end{aligned} \tag{28}$$

Usually, the values of r_2, r_1, and r_0 are not used in the decision process, such that we can integrate (23) over the ranges of r_2, r_1, and r_0 to obtain the following formula:

$$\begin{aligned} \mathrm{P_e}\left(\eta_\mathrm{r}, \theta_\mathrm{r}, \phi_\mathrm{r} | \vec{A}_t\right) = {} & \frac{1}{(4\pi)^3}\int e^{\left[\lambda\left[\left(\sum_{i=0}^{2}\varphi_i\right)-3\right]\right]} \\ & \times 2\sin\theta_2\sin\theta_r\sin 2\eta_r \prod_{i=0}^{2} \\ & \times [1+\lambda(\varphi_i+1)] d\alpha d\beta d\gamma \end{aligned} \tag{29}$$

TABLE I
DECISION REGIONS FOR THE RECEIVED SYMBOL

Regions / Estimated bit	The Region for 1	The Region for 0
$\tilde{b}_0$	$D_{nx} \geq 0$	$D_{nx} < 0$
$\tilde{b}_1$	$D_{ny} \geq 0$	$D_{ny} < 0$
$\tilde{b}_2$	$D_{nz} \geq 0$	$D_{nz} < 0$

where $\varphi_0 = \cos\theta_0(\vec{Q})$, $\varphi_1 = \cos\theta_1(\vec{Q})$, $\varphi_2 = \cos\theta_2(\vec{Q})$, $\lambda = \mathrm{SNR_s}/2$, and $\mathrm{SNR}_s = 1/(2\sigma^2)$.

We define

$$\vec{X} = [\alpha, \beta, \gamma, \eta_r, \theta_r, \phi_r]^T \tag{30}$$

$$\mathrm{h}(\vec{X}) = \lambda\left[\left(\sum_{i=0}^{2}\varphi_i\right) - 3\right] \tag{31}$$

$$\mathrm{g}(\vec{X}) = \frac{2}{(4\pi)^3}\sin\theta_2 \sin\theta_r \sin 2\eta_r \prod_{i=0}^{2}[1 + \lambda(\varphi_i + 1)]. \tag{32}$$

Then (29) can be expressed as

$$\mathrm{P_e}(\eta_r, \theta_r, \phi_r|\vec{A}) = \int e^{\lambda h(\vec{X})} g(\vec{X}) d\alpha d\beta d\gamma. \tag{33}$$

V. EIGHT-SYMBOL SDPolSK (8-SDPolSK) IN CUBIC CONSTELLATION

In the bits-to-symbol mapper, the information symbol $\vec{I}_n$ are Gray coded as

$$\vec{I}_n = \frac{[(2\mathrm{b}_0 - 1), (2\mathrm{b}_1 - 1), (2\mathrm{b}_2 - 1)]^{\mathrm{T}}}{\sqrt{3}} \tag{34}$$

where b_2, b_1, and b_0 are the input bits.

At the receiver, the decision regions for the estimated information symbol $\vec{D}_n$ are shown in Table I, where $\tilde{b}_2$, $\tilde{b}_1$, and $\tilde{b}_0$ are the output estimated bits of b_2, b_1, and b_0, respectively. Table I implies that the decision circuits can be implemented by three folds of the threshold devices [17] with the threshold set to zero, for the decision boundaries are $\mathrm{D_{nx}} = 0$, $\mathrm{D_{ny}} = 0$, and $\mathrm{D_{nz}} = 0$ planes.

Based on (33), the integral forms of the BERs for the estimated b_2, b_1, and b_0, can be written down for different possibilities of $(\mathrm{b_2 b_1 b_0})$'s. For example, when $(\mathrm{b_2 b_1 b_0}) = (111)$ and the angle between $\vec{P}_{n-2}$ and $\vec{P}_{n-1}$ is $\pi/2$, we have $(\theta_t^{(n)}, \phi_t^{(n)}) = (\cos^{-1}(1/\sqrt{3}), \pi/4)$ and $\eta_t^{(n)} = \pi/4$, where the superscript (n) means that the parameter is associated to the symbol in the time slot n. In other words, $\vec{A}_t^{(n)}$ is equal to (1, 1, 1, $\theta_t^{(n-1)}$, $\cos^{-1}(1/\sqrt{3})$, $\pi/4$). The BER for bit b_{i} is given by

$$\mathrm{BER}(\mathrm{b_i}|\vec{A}_t^{(n)}) = \int_{\Omega_i} P_e(\eta_r, \theta_r, \phi_r|\vec{A}_t^{(n)}) d\eta_r d\theta_r d\phi_r = \int_{\Omega_i} e^{\lambda h(\vec{X})} g(\vec{X}) d\vec{X} \tag{35}$$

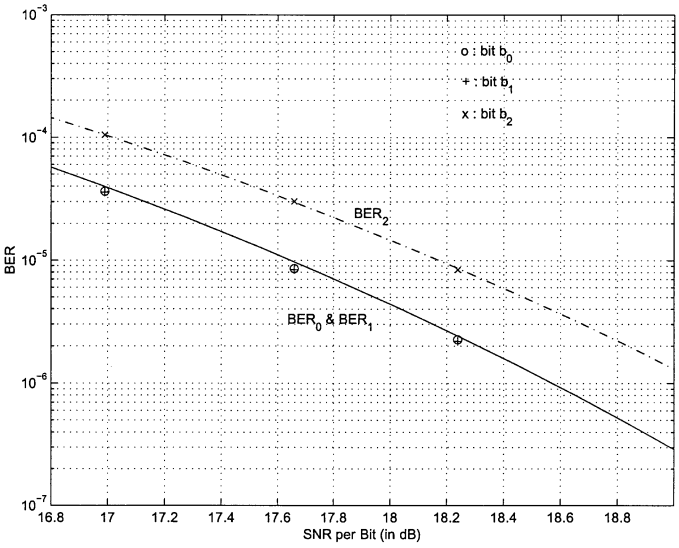

Fig. 6. Saddle-point approximation results and the Monte–Carlo simulation data when $\vec{A}_t^{(n)}$ is equal to (1,1,1, $\pi/4$, $\cos^{-1}(1/\sqrt{3})$, $\pi/4$), where the marks "o," "+," and "x" are the simulation data for bits b_0, b_1, and b_2, respectively.

where

$$\begin{aligned}
\mathrm{i} =& 0, 1, 2\\
\Omega_0 =& \{\vec{A}_r|\mathrm{D_{nx}} < 0\}\\
=& \{\vec{A}_r| -\pi < \phi_r < -\frac{\pi}{2} \mathrm{or} \frac{\pi}{2} < \phi_r < \pi\}\\
\Omega_1 =& \{\vec{A}_r \mathrm{D_{ny}} < 0\} = \{\vec{A}_r| -\pi < \phi_r < 0\}\\
\Omega_2 =& \{\vec{A}_r|\mathrm{D_{nz}} < 0\} = \{\vec{A}_r|\frac{\pi}{2} < \theta_r < \pi\}.
\end{aligned}$$

Because (35) satisfies the conditions of the Laplace-type integral [18]–[20] in the integrated region of $\vec{X}$, it can be well estimated by the saddle-point method as λ is large. The approximation BERs for the case $\vec{A}_t^{(n)}$ is equal to (1, 1, 1, $\pi/4$, $\cos^{-1}(1/\sqrt{3})$, $\pi/4$) are depicted in Fig. 6, where BER_2, BER_1, and BER_0 are the values of BER for the bits b_2, b_1, and b_0, respectively. The Monte–Carlo simulation result is also presented. It is found that they agree with each other very well.

VI. OPTIMIZED 8-SDPolSK

We take the eight possible $\vec{I}_n$'s as

$$\vec{I}_n = [(2b_0 - 1)\sin\theta_t \cos\phi_t, (2b_1 - 1) \times \sin\theta_t, \sin\phi_t, (2b_2 - 1)\cos\theta_t]^T \tag{36}$$

where b_2, b_1, and b_0 are the input bits of the bits-to-symbol mapper, $0 \leq \theta_t \leq \pi$, and $-\pi \leq \phi_t \leq \pi$. Note that the constellation in (36) is symmetric.

Because of the symmetrical nature of the novel frame-constructing algorithm [(15)–(22)], and the property that the deviation distribution of the noisy symbol from the noiseless one is independent of the angle ϕ_t [(7)], where i =0, 1, and 2, it can be proved that the mean BERs for all possible $(\mathrm{b_2 b_1 b_0})$'s are equal, given that the angle η_t is fixed. For example, the probability of receiving the consecutive three symbols, $\vec{R}_{n-2}$, $\vec{R}_{n-1}$, and $\vec{R}_n$, as in Fig. 7(a) is exactly the same as the probability of finding them located as in Fig. 7(b). Therefore, the BER for

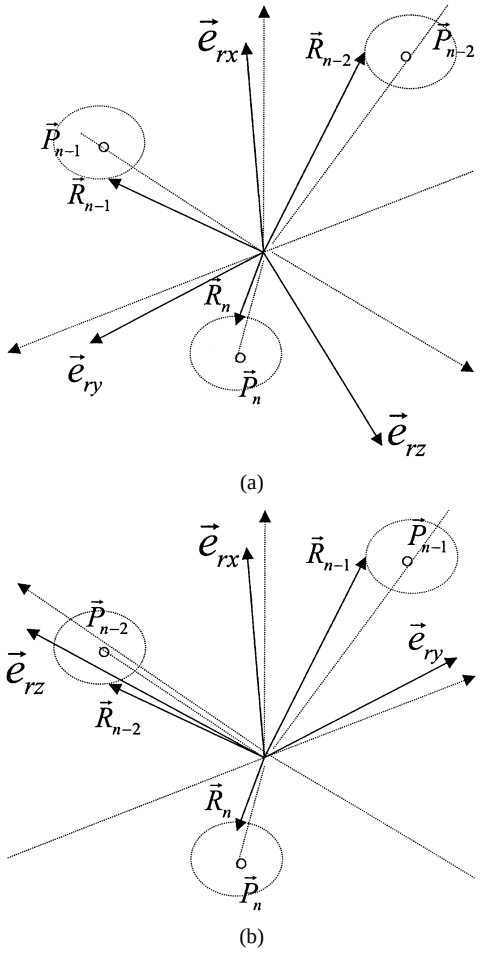

Fig. 7. The distribution probabilities are the same for following two cases. (a) $(b_2b_1b_0) = (111)$. (b) $(b_2b_1b_0) = (100)$.

$(b_2b_1b_0) = (111)$ is equal to that for $(b_2b_1b_0) = (100)$. Other cases can be proved in a similar way. In other words, if we can find the optimal position for the case $(b_2b_1b_0) = (111)$, (36) gives the optimal constellation for all cases.

When the angle between previous two symbols, $2\eta_t$, is equal to $\pi/2$, the optimal parameters (θ_t,ϕ_t) are found equal to (0.9314, $\pi/4$). The mean BER is plotted in Fig. 8. At BER of 10^{-9}, it is only about 2.3 dB worse than the ideal 8-PolSK, and 0.5 dB better than the conventional Gram–Schmidt-based 6-DDPolSK (four symbols available). The optimal parameters (θ_t, ϕ_t)'s for other η_t's can be found in the same way. However, when η_t is far from $\pi/4$, the mean BER of the optimized SD-PolSK degrades rapidly, as shown in Fig. 9. This phenomenon is due to the fact that when the angle η_t is too small (large), the reference axes except axis x (y) are too noisy to bear any information. It is not difficult to avoid the problem. For instance, making some observations on the angle between $\vec{P}_{n-1}$ and $\vec{P}_n$, we can find that when the angle $\eta_t^{(n)}$ is equal to $\pi/4$ and the cases $(b_2b_1b_0) = (101), (110), (001), (010)$ are chosen, it is guaranteed that the angle $\eta_t^{(n+1)}$ at the next time slot will still be $\pi/4$. Therefore, we can transmit two symbols, $\vec{P}_n^c$ and $\vec{P}_n$, at the same time, and $\vec{P}_n^c$ is a constraint in the four cases mentioned above. In addition, the two reference symbols at the next time slot are no longer $\vec{P}_{n-1}$ and $\vec{P}_n$, but $\vec{P}_{n-1}$ and $\vec{P}_n^c$. In this way, the angle η_t can be maintained $\pi/4$ constantly, and the

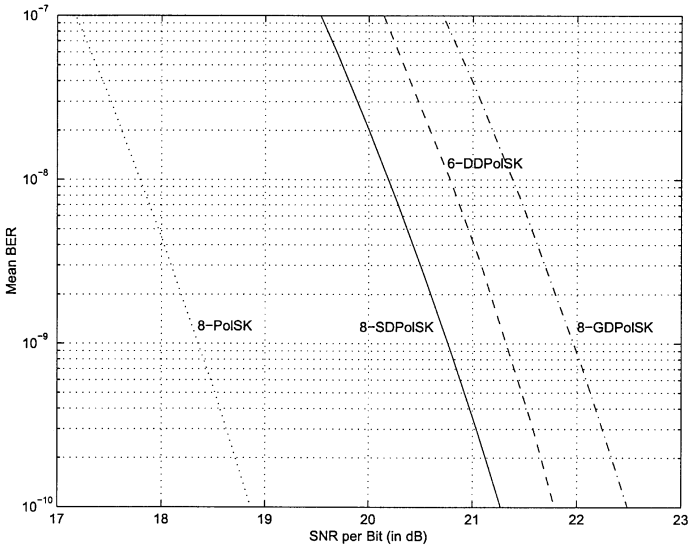

Fig. 8. Mean BERs of the ideal 8-PolSK, optimal 8-SDPolSK, 6-DDPolSK, and suboptimal 8-GDPolSK systems.

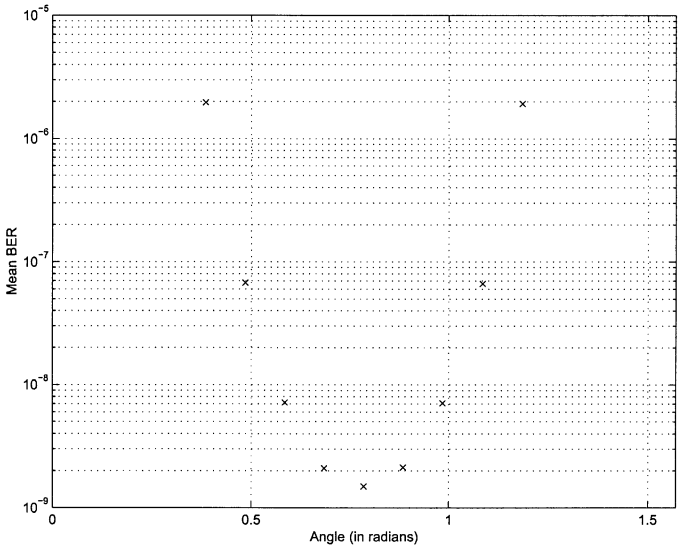

Fig. 9. Mean BER values for different values of the angle η_t.

b_0 bit in $\vec{P}_n^c$ is a stuffing bit, which does not carry information but makes the system automatically runlength-limited coded [21]. There is some coding gain provided in the symbol $\vec{P}_n^c$, but this aspect is not investigated here.

VII. Conclusion

A novel differential coding algorithm is proposed for the first time in constructing the reference frame for the DPolSK system, instead of the conventional Gram–Schmidt algorithm. The new scheme guarantees the symmetry of the optimal constellation. For example, the optimal constellation for the 8-SDPolSK can be easily found to be a rectangle. The analytic integral form for the BER of the system is derived, and the saddle-point method is applied to approximate the integral values which agree well with the simulation results. The performance in terms of BER is only about 2.3 dB worse than the ideal 8-PolSK, but 0.5 dB

better than the conventional 6-DDPolSK (four symbols available). This scheme presents the best performance among all DPolSKs so far.

APPENDIX

DEFINITIONS

For our convenience in the following descriptions, we introduce the rotation operator $\mathrm{Rot}_{(\vec{A},\theta)}$, where $\vec{A}$ is the rotation axis, and θ is the number of rotation degrees. As $\mathrm{Rot}_{(\vec{A},\theta)}$ operates on a vector $\vec{B}$, $\vec{B}$ is rotated by an angle θ counterclockwise about $\vec{A}$, and the resultant vector gives $\mathrm{Rot}_{(\vec{A},\theta)}\vec{B}$. It is known [20] that the operator $\mathrm{Rot}_{(\vec{A},\theta)}$ may be expressed as a vector operation form in a 3-D space, as follows:

$$\mathrm{Rot}_{(\vec{A},\theta)}\vec{B} = [\cos\theta \overleftrightarrow{I} + (1-\cos\theta)\vec{A}\vec{A}^T \cdot + \sin\theta \vec{A}\times]\vec{B} \quad \text{(A1)}$$

where $\overleftrightarrow{I}$ is a three-by-three identity matrix.

TRANSFORMATION RELATION BETWEEN $(\theta_1,\phi_1,\theta_2,\phi_2,\theta_3,\phi_3)$ AND $(\alpha,\beta,\gamma,\eta_\mathrm{r},\theta_\mathrm{r},\phi_\mathrm{r})$

Because the channel Jones matrix can be neglected, given the positions of the noise-free vectors $\overrightarrow{P_{n-2}}$, $\overrightarrow{P_{n-1}}$, and $\overrightarrow{P_n}$, the positions of noisy vectors $\overrightarrow{R_{n-2}}$, $\overrightarrow{R_{n-1}}$, and $\overrightarrow{R_n}$ in the $\mathrm{t_x}-\mathrm{t_y}-\mathrm{t_z}$ coordinate may be expressed in terms of $(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_0,\theta_1,\phi_1,\theta_2,\phi_2,\theta_0,\phi_0)$ or $(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_0,\alpha,\beta,\gamma,\eta_\mathrm{r},\theta_\mathrm{r},\phi_\mathrm{r})$, as shown in the following three equations, where the left side and the right side are in terms of $(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_0,\theta_1,\phi_1,\theta_2,\phi_2,\theta_0,\phi_0)$ and $(\mathrm{r}_1,\mathrm{r}_2,\mathrm{r}_0,\alpha,\beta,\gamma,\eta_\mathrm{r},\theta_\mathrm{r},\phi_\mathrm{r})$, respectively.

$$\begin{aligned}\overrightarrow{R_{n-2}} &= \mathrm{r}_2\mathrm{Rot}_{(\vec{P}_{n-2\perp},\pi/2)}[\sin\theta_2\cos\phi_2 \ \ \sin\theta_2\sin\phi_2 \ \ \cos\theta_2]^T \\ &= \mathrm{r}_2[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}]\,[\cos\eta_r \ \ -\sin\eta_r 0]^T \end{aligned} \quad \text{(A2)}$$

$$\begin{aligned}\overrightarrow{R_{n-1}} &= \mathrm{r}_1\mathrm{Rot}_{(\vec{P}_{n-1\perp},-\pi/2)}[\sin\theta_1\cos\phi_1 \ \ \sin\theta_1\sin\phi_1 \ \ \cos\theta_1]^T \\ &= \mathrm{r}_1\left[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}\right]\,[\cos\eta_r \sin\eta_r 0]^T \end{aligned} \quad \text{(A3)}$$

$$\begin{aligned}\overrightarrow{R_n} &= \mathrm{r}_0\mathrm{Rot}_{(\vec{e}_{tx},\phi_r)}\mathrm{Rot}_{(\vec{e}_{ty},\theta_r)} \\ &\quad\times[\sin\theta_0\cos\phi_0 \ \ \sin\theta_0\sin\phi_0 \ \ \cos\theta_0]^T \\ &= \mathrm{r}_0\left[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}\right]\,[\sin\theta_r\cos\phi_r \ \ \sin\theta_r\sin\phi_r \ \ \cos\theta_r]^T \end{aligned} \quad \text{(A4)}$$

where $\overrightarrow{P_{n-2}}_\perp$ and $\overrightarrow{P_{n-1}}_\perp$ are the noise-free vectors on the $\mathrm{t_x}-0-\mathrm{t_y}$ plane and orthogonal to $\overrightarrow{P_{n-2}}$ and $\overrightarrow{P_{n-1}}$, respectively; $[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}]$ is the coordinate frame constructed by $\overrightarrow{R_{n-2}}$ and $\overrightarrow{R_{n-1}}$; and α, β, and γ are the Euler angles [20] required to describe the coordinate of $[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}]$.

Equation (A2) is equivalent to the transformation matrix $[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}]$ rotates the vector, $[\cos\eta_r \ \ -\sin\eta_r \ 0]^T$, to $\mathrm{Rot}_{(\overrightarrow{P_{n-2}}_\perp,\pi/2)}[\sin\theta_2\cos\phi_2 \quad \sin\theta_2\sin\phi_2 \quad \cos\theta_2 \ 0]^T$ (i.e., $\overrightarrow{R_{n-2}}$). After some observations on the positions of $[\cos\eta_r \ \ -\sin\eta_r \ 0]^T$, $\overrightarrow{R_{n-2}}$, and $\overrightarrow{P_{n-2}}$, we find that $[\cos\eta_r \ \ -\sin\eta_r \ 0]^T$ can be rotated to $\overrightarrow{R_{n-2}}$ if $[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}]$ is set to the following form:

$$\begin{aligned}\left[\overrightarrow{e_{rx}} \ \overrightarrow{e_{ry}} \ \overrightarrow{e_{rz}}\right] &= \mathrm{Rot}_{(\vec{R}_{n-2},-\gamma)}\mathrm{Rot}_{(\vec{P}_{n-2},-\beta)} \\ &\quad\times\mathrm{Rot}_{(\vec{e}_{tz},\alpha)} \end{aligned} \quad \text{(A5)}$$

$$\alpha = \eta_\mathrm{r} - (\eta_t - \theta_2) \quad \text{(A6)}$$

$$\beta = \phi_2 - \left(\frac{\pi}{2} - \eta_t\right). \quad \text{(A7)}$$

Furthermore, γ is a constraint by (A3).

To simplify computing the Jacobian of the transformation, we rearrange (A3) and (A4) as follows:

$$[\sin\theta_1\cos\phi_1\sin\theta_1\sin\phi_1\cos\theta_1]^T = \left[\overrightarrow{K_{1X}}\overrightarrow{K_{1Y}}\overrightarrow{K_{1Z}}\right]\overrightarrow{u_1} \quad \text{(A8)}$$

$$[\sin\theta_0\cos\phi_0\sin\theta_0\sin\phi_0\cos\theta_0]^T = \left[\overrightarrow{K_{0X}}\overrightarrow{K_{0Y}}\overrightarrow{K_{0Z}}\right]\overrightarrow{u_0} \quad \text{(A9)}$$

where

$$\left[\overrightarrow{K_{1X}}\overrightarrow{K_{1Y}}\overrightarrow{K_{1Z}}\right] = \mathrm{Rot}_{(P_{n-1}\vec{}_{\perp},\frac{\pi}{2})}\left[\overrightarrow{P_{n-2}}\overrightarrow{P'_{n-2\perp}}\overrightarrow{e'_{rz}}\right] \quad \text{(A10)}$$

$$\overrightarrow{P'_{n-2\perp}} = \mathrm{Rot}_{(\vec{P}_{n-2},\beta)}\overrightarrow{P_{n-2\perp}} \quad \text{(A11)}$$

$$\overrightarrow{e'_{rz}} = \mathrm{Rot}_{(\vec{P}_{n-2},\beta)}\overrightarrow{e_{rz}} \quad \text{(A12)}$$

$$\overrightarrow{u_1} = \begin{bmatrix} -\cos\gamma\sin(2\eta_r)\sin(\theta_2)+\cos(2\eta_r)\cos(\theta_2) \\ \cos\gamma\sin(2\eta_r)\cos(\theta_2)+\cos(2\eta_r)\sin(\theta_2) \\ -\sin\gamma\sin(2\eta_r)\end{bmatrix} \quad \text{(A13)}$$

$$\left[\overrightarrow{K_{0X}}\overrightarrow{K_{0Y}}\overrightarrow{K_{0Z}}\right] = \mathrm{Rot}_{(\vec{e}_{ty},\theta_t)}\mathrm{Rot}_{(\vec{e}_{tz},\phi_t)}\left[\overrightarrow{e_{rx}}\overrightarrow{e_{ry}}\overrightarrow{e_{rz}}\right] \quad \text{(A14)}$$

$$\overrightarrow{u_0} = [\sin\theta_r\cos\phi_r \ \ \sin\theta_r\sin\phi_r \ \ \cos\theta_r]^T. \quad \text{(A15)}$$

Notice that $[\overrightarrow{K_{1X}} \ \overrightarrow{K_{1Y}} \ \overrightarrow{K_{1Z}}]$ is independent of γ and η_r, $[\overrightarrow{K_{0X}} \ \overrightarrow{K_{0Y}} \ \overrightarrow{K_{0Z}}]$ is independent of θ_r and ϕ_r, and both matrices are unitary.

Prior to calculating the Jacobian, we give two lemmas in advance.

Lemma 1:

$$\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\times\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r}+\frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right) = 2\sin 2\eta_r\overrightarrow{u_1} \quad \text{(A16)}$$

$$\frac{\partial\overrightarrow{u_0}}{\partial\theta_r}\times\frac{\partial\overrightarrow{u_0}}{\partial\phi_r} = \sin\theta_r\overrightarrow{u_0}. \quad \text{(A17)}$$

Lemma 2: If $[\sin\theta\cos\phi \ \ \sin\theta\sin\phi \ \ \cos\theta] = [\mathrm{A}(\mathrm{x}_1,\mathrm{x}_2) \ \mathrm{B}(\mathrm{x}_1,\mathrm{x}_2) \ \mathrm{C}(\mathrm{x}_1,\mathrm{x}_2)]$, then the partial derivatives are given by

$$\frac{\partial\theta}{\partial\mathrm{x}_1} = -\frac{\frac{\partial\mathrm{C}}{\partial\mathrm{x}_1}}{\sin\theta} \quad \text{(A18)}$$

$$\frac{\partial\phi}{\partial\mathrm{x}_1} = \frac{\mathrm{A}\frac{\partial\mathrm{B}}{\partial\mathrm{x}_1}-\frac{\partial\mathrm{A}}{\partial\mathrm{x}_1}\mathrm{B}}{\mathrm{A}^2+\mathrm{B}^2} = \frac{\mathrm{A}\frac{\partial\mathrm{B}}{\partial\mathrm{x}_1}-\frac{\partial\mathrm{A}}{\partial\mathrm{x}_1}\mathrm{B}}{\sin^2\theta} \quad \text{(A19)}$$

$$\frac{\partial\theta}{\partial \mathrm{x}_2} = -\frac{\frac{\partial \mathrm{C}}{\partial \mathrm{x}_2}}{\sin\theta} \tag{A20}$$

$$\frac{\partial\phi}{\partial \mathrm{x}_2} = \frac{\mathrm{A}\frac{\partial \mathrm{B}}{\partial \mathrm{x}_2} - \frac{\partial \mathrm{A}}{\partial \mathrm{x}_2}\mathrm{B}}{\sin^2\theta}. \tag{A21}$$

Remark: It is known that $\cos\theta = \mathrm{C}$, and $\tan\phi = \mathrm{B}/\mathrm{A}$. Applying chain rules yields the desired results.

The Jacobian J is given as

$$J = \begin{bmatrix} \frac{\partial r_2}{\partial r_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ x & \frac{\partial r_1}{\partial r_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ x & x & \frac{\partial r_0}{\partial r_0} & 0 & 0 & 0 & 0 & 0 & 0 \\ x & x & x & \frac{\partial\theta_2}{\partial\alpha} & 0 & 0 & \frac{\partial\theta_2}{\partial\eta_r} & 0 & 0 \\ x & x & x & 0 & \frac{\partial\phi_2}{\partial\beta} & 0 & 0 & 0 & 0 \\ x & x & x & \frac{\partial\theta_1}{\partial\alpha} & x & \frac{\partial\theta_1}{\partial\gamma} & \frac{\partial\theta_1}{\partial\eta_r} & 0 & 0 \\ x & x & x & \frac{\partial\phi_1}{\partial\alpha} & x & \frac{\partial\phi_1}{\partial\gamma} & \frac{\partial\phi_1}{\partial\eta_r} & 0 & 0 \\ x & x & x & x & x & \frac{\partial\theta_0}{\partial\gamma} & \frac{\partial\theta_0}{\partial\eta_r} & \frac{\partial\theta_0}{\partial\theta_r} & \frac{\partial\theta_0}{\partial\phi_r} \\ x & x & x & x & x & \frac{\partial\phi_0}{\partial\gamma} & \frac{\partial\phi_0}{\partial\eta_r} & \frac{\partial\phi_0}{\partial\theta_r} & \frac{\partial\phi_0}{\partial\phi_r} \end{bmatrix} \tag{A22}$$

where the mark "x" means the element is of no effect, and (A6) and (A7) are applied.

After some manipulations, the Jacobian J can be written as

$$|\mathrm{J}| = |\mathrm{J}_2||\mathrm{J}_1|\,|\mathrm{J}_0| \tag{A23}$$

where

$$J_2 = \begin{bmatrix} \frac{\partial\theta_2}{\partial\alpha} & 0 \\ 0 & \frac{\partial\phi_2}{\partial\beta} \end{bmatrix} = 1 \tag{A24}$$

$$J_1 = \begin{bmatrix} \frac{\partial\theta_1}{\partial\gamma} & \frac{\partial\theta_1}{\partial\eta_r} + \frac{\partial\theta_1}{\partial\alpha} \\ \frac{\partial\phi_1}{\partial\gamma} & \frac{\partial\phi_1}{\partial\eta_r} + \frac{\partial\phi_1}{\partial\alpha} \end{bmatrix} \tag{A25}$$

$$J_0 = \begin{bmatrix} \frac{\partial\theta_0}{\partial\theta_r} & \frac{\partial\theta_0}{\partial\phi_r} \\ \frac{\partial\phi_0}{\partial\theta_r} & \frac{\partial\phi_0}{\partial\phi_r} \end{bmatrix}. \tag{A26}$$

$|\mathrm{J}_1|$ and $|\mathrm{J}_0|$ are calculated in the following theorem.

Theorem 1:

$$\text{(a)}\ |\mathrm{J}_1| = \frac{2\sin 2\eta_r}{\sin\theta_1} \tag{A27}$$

$$\text{(b)}\ |\mathrm{J}_0| = \frac{\sin\theta_\mathrm{r}}{\sin\theta_0}. \tag{A28}$$

Notice that θ_1, θ_r, and $\theta_0 \in [0,\pi]$, and $\eta_\mathrm{r} \in [0,\pi/2]$.

Proof: (a) Applying *Lemma 2*, we have

$$\begin{aligned}
|-\mathrm{J}_1\sin^3\theta_1| &= \left|\left(\left(\frac{\partial\theta_1}{\partial\pi_r} + \frac{\partial\theta_1}{\partial\alpha}\right)\frac{\partial\phi_1}{\partial\gamma}\right.\right. \\
&\quad \left.\left. -\frac{\partial\phi_1}{\partial\gamma}\left(\frac{\partial\phi_1}{\partial\pi_r} + \frac{\partial\phi_1}{\partial\alpha}\right)\right)\sin^3\theta_1\right| \\
&= \left|\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\left\{\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\overrightarrow{u_1}\right)\left(\overrightarrow{K_{1Y}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\right.\right. \\
&\quad \left.\left. -\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\left(\overrightarrow{K_{1Y}}^{\mathrm{T}}\overrightarrow{u_1}\right)\right\}\right. \\
&\quad -\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\left\{\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\overrightarrow{u_1}\right)\left(\overrightarrow{K_{1Y}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\right. \\
&\quad \left.\left. -\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\left(\overrightarrow{K_{1Y}}^{\mathrm{T}}\overrightarrow{u_1}\right)\right\}\right| \\
&= \left|\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\left\{\overrightarrow{u_1}^{\mathrm{T}}\left[\left(\overrightarrow{K_{1Y}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\overrightarrow{K_{1X}}\right.\right.\right. \\
&\quad \left.\left.\left. -\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\overrightarrow{K_{1Y}}\right]\right\}\right. \\
&\quad -\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\left\{\overrightarrow{u_1}^{\mathrm{T}}\left[\left(\vec{K}_{1\gamma}{}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\overrightarrow{K_{1X}}\right.\right. \\
&\quad \left.\left.\left. -\left(\overrightarrow{K_{1X}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\overrightarrow{K_{1Y}}\right]\right\}\right| \\
&= \left|\overrightarrow{u_1}^{\mathrm{T}}\left\{\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\left[\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\times\left(\overrightarrow{K_{1Z}}\right)\right]\right.\right. \\
&\quad \left.\left. -\left(\overrightarrow{K_{1Z}}^{\mathrm{T}}\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\right)\left[\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\times\left(\overrightarrow{K_{1Z}}\right)\right]\right\}\right| \\
&= \left|\overrightarrow{u_1}^{\mathrm{T}}\left\{\left[\overrightarrow{K_{1Z}}\times\left(\frac{\partial\overrightarrow{u_1}}{\partial\gamma}\times\left(\frac{\partial\overrightarrow{u_1}}{\partial\eta_r} + \frac{\partial\overrightarrow{u_1}}{\partial\alpha}\right)\right)\right]\times\overrightarrow{K_{1Z}}\right\}\right| \\
&= \left|2\sin 2\eta_r\,\overrightarrow{u_1}^{\mathrm{T}}\left\{\left[\overrightarrow{K_{1Z}}\times\overrightarrow{u_1}\right]\times\overrightarrow{K_{1Z}}\right\}\right| \\
&= \left|2\sin 2\eta_r\,\overrightarrow{u_1}^{\mathrm{T}}\left\{\left(\overrightarrow{u_1}^{\mathrm{T}}\overrightarrow{K_{1Z}}\right)\overrightarrow{K_{1Z}} - \overrightarrow{u_1}\right\}\right| \\
&= \left|2\sin 2\eta_r\left(\cos^2\theta_1 - 1\right)\right| = 2\sin 2\eta,\sin^2\theta_1 \\
&\therefore\quad |\mathrm{J}_1| = 2\sin 2\eta_r/\sin\theta_1.
\end{aligned}$$

[Following similar procedures to (a), we can prove (b).]

SIMPLIFICATIONS

The expressions for θ_2, $\cos\theta_1$, and $\cos\theta_0$ are required in the integrand of (35). From (A6) and (A8), we can easily find the first two terms as

$$\theta_2 = \alpha - \eta_\mathrm{r} + \eta_t \tag{A29}$$

$$\begin{aligned}
\cos\theta_1 &= \overrightarrow{P_{n-1}}^{T}\left[\overrightarrow{P_{n-2}}\,\overrightarrow{P'_{n-2\perp}}\,\overrightarrow{e'_{rz}}\right]\overrightarrow{u_1} \\
&= \left[\cos 2\eta_t\ \sin 2\eta_t\ \cos\beta - \sin 2\eta_t\sin\beta\right]\overrightarrow{u_1}.
\end{aligned} \tag{A30}$$

The expression for $\cos\theta_0$ is given by

$$\cos\theta_0 = \overrightarrow{P_n}^T \mathrm{Rot}_{(\vec{R}_{n-2},-\gamma)}\mathrm{Rot}_{(\vec{P}_{n-2},\beta)}\mathrm{Rot}_{(\vec{e}_{tz},\alpha)}\overrightarrow{u_0}. \quad \text{(A31)}$$

It needs to be further simplified.

Lemma 3:

$$\begin{aligned}\left[\overrightarrow{e_{tx}}\ \overrightarrow{e_{ty}}\ \overrightarrow{e_{tz}}\right] &= \left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right] \\ &\times \begin{bmatrix}\cos\theta_2 & \sin\theta_2 & 0\\ -\sin\theta_2 & \cos\theta_2 & 0\\ 0 & 0 & 1\end{bmatrix} \times \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\beta & \sin\beta\\ 0 & -\sin\beta & \cos\beta\end{bmatrix} \\ &\times \begin{bmatrix}\cos\eta_t & -\sin\eta_t & 0\\ \sin\eta_t & \cos\eta_t & 0\\ 0 & 0 & 1\end{bmatrix}\end{aligned} \quad \text{(A32)}$$

where $\overrightarrow{R'_{n-2\perp}} = \overrightarrow{e'_{tz}} \times \overrightarrow{R_{n-2}}$.

Lemma 4:

$$\begin{aligned}\mathrm{Rot}_{(\vec{P}_{n-2},\beta)}\mathrm{Rot}_{(\vec{e}_{tz},\alpha)} &= \left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right] \\ &\times \begin{bmatrix}\cos\eta_r & -\sin\eta_r & 0\\ \sin\eta_r & \cos\eta_r & 0\\ 0 & 0 & 1\end{bmatrix}.\end{aligned} \quad \text{(A33)}$$

Lemma 5:

$$\begin{aligned}&\left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right]^T \left(\mathrm{Rot}_{(\overline{R}_{n-2},-\gamma)}\left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right]\right) \\ &\qquad = \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\gamma & \sin\gamma\\ 0 & -\sin\gamma & \cos\gamma\end{bmatrix}.\end{aligned} \quad \text{(A34)}$$

Applying the previous three lemmas, we have

$$\begin{aligned}\cos\theta_0 &= \overrightarrow{P_n}^T \begin{bmatrix}\cos\eta_t & -\sin\eta_t & 0\\ \sin\eta_t & \cos\eta_t & 0\\ 0 & 0 & 1\end{bmatrix}^T \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\beta & \sin\beta\\ 0 & -\sin\beta & \cos\beta\end{bmatrix}^T \\ &\times \begin{bmatrix}\cos\theta_2 & \sin\theta_2 & 0\\ -\sin\theta_2 & \cos\theta_2 & 0\\ 0 & 0 & 1\end{bmatrix}^T \\ &\times \left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right]^T \\ &\times \mathrm{Rot}_{(\overline{R}_{n-2},-\gamma)}\left[\overrightarrow{R_{n-2}}\ \overrightarrow{R'_{n-2\perp}}\ \overrightarrow{e'_{rz}}\right] \\ &\times \begin{bmatrix}\cos\eta_r & -\sin\eta_r & 0\\ \sin\eta_r & \cos\eta_r & 0\\ 0 & 0 & 1\end{bmatrix} \overrightarrow{u_0} \\ &= [\sin\theta_t\ \cos(\phi_t+\eta_t)\ \sin\theta_t\ \sin(\phi_t+\eta_t)\ \cos\theta_t] \\ &\times \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\beta & \sin\beta\\ 0 & -\sin\beta & \cos\beta\end{bmatrix}^T \\ &\times \begin{bmatrix}\cos\theta_2 & \sin\theta_2 & 0\\ -\sin\theta_2 & \cos\theta_2 & 0\\ 0 & 0 & 1\end{bmatrix}^T \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\gamma & \sin\gamma\\ 0 & -\sin\gamma & \cos\gamma\end{bmatrix} \\ &\times \begin{bmatrix}\sin\theta_r & \cos(\phi_r+\eta_r))\\ \sin\theta_r & \sin(\phi_r+\eta_r))\\ & \cos\theta_r\end{bmatrix} \\ &= [\sin\theta_t\ \cos(\phi_t+\eta_t)\ \sin\theta_t\ \sin(\phi_t+\eta_t)\ \cos\theta_t] \\ &\times \begin{bmatrix}\cos\theta_2 & -\sin\theta_2\cos\gamma & -\sin\theta_2\sin\gamma\\ \cos\beta\sin\theta_2 & \cos\beta\cos\theta_2\cos\gamma & \cos\beta\cos\theta_2\sin\gamma\\ & +\sin\beta\sin\gamma & -\sin\beta\cos\gamma\\ \sin\beta\sin\theta_2 & \sin\beta\cos\theta_2\cos\gamma & \sin\beta\cos\theta_2\sin\gamma\\ & -\cos\beta\sin\gamma & +\cos\beta\cos\gamma\end{bmatrix} \\ &\times \begin{bmatrix}\sin\theta_r & \cos(\phi_r+\eta_r))\\ \sin\theta_r & \sin(\phi_r+\eta_r))\\ & \cos\theta_r\end{bmatrix}.\end{aligned} \quad \text{(A35)}$$

ACKNOWLEDGMENT

The authors appreciate the reviewers' valuable comments and suggestions, and valuable discussions with M. Tsai.

REFERENCES

[1] S. Davidson, "Testing LANs optically to the gigabit Ethernet standard," *IEEE Spectr.*, vol. 34, pp. 86–90, Sept. 1997.
[2] D. Hanson, "Gibabit Ethernet optical fiber standard and its impact on the LAN industry," in *Proc. LEOS '98*, vol. 1, 1998, pp. 172–173.
[3] L. Raddatz, I. H. White, D. G. Cunningham, and M. C. Nowell, "Bandwidth enhancement of multimode fiber Gb/s networks using conditioned launch," in *CLEO'98 Tech. Dig.*, May 1998, pp. 396–397.
[4] S. Benedetto and P. Poggiolini, "Theory of polarization shift keying modulation," *IEEE Trans. Commun.*, vol. 40, pp. 708–721, Apr. 1992.
[5] ——, "Multilevel polarization shift keying: optimum receiver structure and performance evaluation," *IEEE Trans. Commun.*, vol. 42, pp. 1174–1186, Feb.–Apr. 1994.
[6] S. Betti, F. Curti, G. D. Marchis, and E. Iannone, "Multilevel coherent optical system based on Stokes parameters modulation," *J. Lightwave Technol.*, vol. 9, pp. 1314–1320, Oct. 1991.
[7] S. Benedetto, R. Gaudino, and P. Poggiolini, "Direct detection of optical digital transmission based on polarization shift keying modulation," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 531–542, Apr. 1995.
[8] S. Betti, F. Curti, G. D. Marchis, and E. Iannone, "A novel multilevel coherent optical system: four-quadrature signaling," *J. Lightwave Technol.*, vol. 9, pp. 514–523, Apr. 1991.
[9] M. Born and E. Wolf, *Principles of Optics*. Oxford, U.K.: Pergamon, 1975.
[10] S. C. Rashleigh, "Origins and control of the polarization effects in single-mode fibers," *J. Lightwave Technol.*, vol. LT–1, pp. 312–321, Feb. 1983.
[11] S. Benedetto, R. Gaudino, and P. Poggiolini, "Polarization recovery in optical polarization shift keying system," *IEEE Trans. Commun.*, vol. 45, pp. 1269–1279, Oct. 1997.
[12] R. J. Blaikie, D. P. Taylor, and P. T. Gough, "Multilevel differential polarization shift keying," *IEEE Trans. Commun.*, vol. 45, pp. 95–102, Jan. 1997.
[13] K.-S. Hou and J. Wu, "The differential four-quadrature coding scheme for applications in optical communications," *IEICE Trans. Commun.*, vol. E83-B, no. 7, pp. 1379–1388, 2000.
[14] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks*. New York: Wiley, 1998.
[15] A. E. Willner and Q. Yu, "Transmission limitations due to polarization mode dispersion," in *Proc. LEOS Summer Topical Meetings*, 2001, pp. 9–10.
[16] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[17] G. Einarsson, *Principles of Lightwave Communications*. New York: Wiley, 1996.
[18] D. Zwillinger, *Handbook of Integration*. Boston, MA: Jones and Bartlett, 1992.
[19] N. Bleistein and R. AA. Handelsman, *Asymptotic Expansions of Integrals*. New York: Dover, 1986.
[20] R. Wong, *Asymptotic Approximations of Integrals*. New York: Academic, 1989.
[21] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
[22] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Wesley, 1980.

Kuen-Suey Hou (M'02) was born in Yunlin, Taiwan, R.O.C. in 1973. He received the B.E. and the Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C. in 1995 and 2000, respectively.

Since 2000, he has been with the Mediatek Inc., Hsinchu, Taiwan, R.O.C., where he is involved in digital circuits and system design.

Jingshown Wu (S'73–M'78–SM'99) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C. in 1970 and 1972, and the Ph.D. degree from Cornell University, Ithaca, NY in 1978.

He joined Bell Laboratories in 1978, where he worked on digital network standards and performance, and optical fiber communication systems. In 1984, he joined the Department of Electrical Engineering, National Taiwan University, as Professor, and was the Chairman of the Department from 1987 to 1989. He was also the Director of the Communication Research Center, College of Engineering of the university from 1992 to 1995. From 1995 to 1998, he was the Director of the Division of Engineering and Applied Science, National Science Council, Taiwan, R.O.C., on leave from the university. From 1999 to 2002, he was the Chairman of the Commission on Research and Development, and the Director of the Center for Sponsor Programs of the university. Currently, he is the Vice-President of the university. He is interested in optical fiber communications, computer communications, and communication systems. He has published more than 100 journal and conference papers and holds 12 patents.

Dr. Wu is a Life Member of the Chinese Institute of Engineers, the Optical Society of China, and the Institute of Chinese Electrical Engineers. Since 1997, he has served as the Vice Chairman (1997–1998) and the Chairman (1998–2000) of IEEE, Taipei Section.