Quantum Theory of Noise in Stokes Vector Receivers and Application to Bit Error Rate Analysis

Kazuro Kikuchi, *Fellow, IEEE, Fellow, OSA*

***Abstract*—Polarization-modulated optical signals can be demodulated via direct detection of Stokes vectors. In this article, we develop a quantum theory of noise in the Stokes vector receiver. Examining the conditions that enable simultaneous measurement of three elements of the Stokes vector, we demonstrate that an optical pre-amplifier in front of the Stokes vector receiver can reduce the power penalty associated with simultaneous measurements, down to the quantum limit of 3 dB. Then, the 3-D Gaussian noise model for the shot noise and optical pre-amplifier noise is discussed. Subsequently, based on the noise analyses, we derive bit error rate formulae for the cubic-lattice polarization modulation format and evaluate the sensitivity of the Stokes vector receiver.**

***Index Terms*—Optical fiber communication, optical noise, optical polarization, optical receivers, quantum theory.**

I. Introduction

A STOKES vector receiver with low-complexity digital signal processing (DSP) circuits can track random fluctuations in the state of polarization (SOP) of the optical signal, even if it is based on direct detection [1], [2]. Such a receiver facilitates the employment of the multi-level SOP-modulation format, in which constellations are designed in the 3-dimensional (3D) Stokes space. Therefore, it offers an effective solution for high-capacity, cost-effective optical communication systems [3]–[7].

Quantum noise characteristics of the Stokes vector receiver, however, have not yet been studied. For example, the following crucial problems regarding its performance remain unsolved [8], [9]:

1) Although the shot noise is the most fundamental quantum noise generated in direct detection receivers, it has not been demonstrated how the shot noise distributes on the 3D Stokes vector.
2) The effects of the optical pre-amplifier noise in the Stokes vector receiver have not been discussed.
3) When all three elements of the Stokes vector are measured simultaneously, sensitivity degradation is inevitable in the Stokes vector receiver from a quantum mechanical perspective. However, we do not know the minimum sensitivity degradation that the quantum theory predicts or how we could achieve it in a real system.
4) The theoretical sensitivity limit of the Stokes vector receiver for the most basic cubic-lattice SOP modulation format has not yet been demonstrated.

This study focuses on the fundamental aspects of the Stokes vector receiver to answer the above-mentioned problems. After developing a quantum theory of noise in Stokes vectors, we derive the conditions for enabling simultaneous measurement of all three Stokes parameters. One-to-three branching of the incoming signal is the most common technique, and the power penalty caused by branching loss is 4.8 dB. On the other hand, the use of an optical pre-amplifier can reduce the power penalty down to 3 dB, which is the minimum value that the quantum theory predicts.

Thereafter, the Gaussian noise approximation for the shot noise and optical pre-amplifier noise is discussed. We calculate the probability density functions (PDF) of the noises for the cubic-lattice SOP modulation format using the Gaussian noise model and computer simulations. Subsequently, the validity of the Gaussian noise model is demonstrated via a comparison between these. The bit error rate (BER) formulae for the cubic-lattice SOP modulation format are derived by using the Gaussian noise model. We calculate the BER as a function of the number of average photons per bit and examine the effects of the branching loss and optical pre-amplification on receiver sensitivity.

The remainder of this paper is organized as follows. Section II reviews the classical description of the Stokes vector, the configurations of the Stokes vector receiver, and the cubic-lattice SOP modulation format. Section III deals with the quantum theory of the Stokes vector. After discussing the characteristics of the shot noise on the Stokes vector, we derive the conditions for enabling simultaneous measurement of all three Stokes parameters. Section IV deals with the Gaussian noise model of the shot noise and pre-amplifier noise for the cubic-lattice SOP modulation format. In Section V, the BER formulae for the cubic-lattice SOP modulation format are derived based on the Gaussian noise model. Finally, Section VI concludes the paper.

II. Classical Description of the Stokes Vector

This section discusses the classical interpretation of the Stokes vector. Its definition, measurement system, and modulation format are described.

A. Definition of the Stokes Vector

Let E_x and E_y be the x- and y-polarization components of the signal, respectively. Then, the Stokes vector, $\mathbf{S}$, is defined

Manuscript received November 8, 2019; revised December 30, 2019; accepted January 14, 2020. Date of publication January 17, 2020; date of current version June 16, 2020.

The author is with the National Institution for Academic Degrees and Quality Enhancement of Higher Education, Kodaira-shi 187-8587, Japan (e-mail: kikuti@niad.ac.jp).

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

Digital Object Identifier 10.1109/JLT.2020.2967420

TABLE I
STOKES PARAMETERS EXPRESSED BY DIFFERENT BASES

	x and y bases	$45°$ and $135°$ bases	L and R bases
S_1	$\lvert E_x\rvert^2 - \lvert E_y\rvert^2$	-2 Re $\left(E_{45°} E^*_{135°}\right)$	2 Im $\left(E_L E^*_R\right)$
S_2	2 Re $\left(E_y E^*_x\right)$	$\lvert E_{45°}\rvert^2 - \lvert E_{135°}\rvert^2$	2 Re $\left(E_L E^\star_R\right)$
S_3	2 Im $\left(E_y E^*_x\right)$	-2 Im $\left(E_{45°} E^*_{135°}\right)$	$\lvert E_L\rvert^2 - \lvert E_R\rvert^2$

[10] as

$$\mathbf{S} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} |E_x|^2 - |E_y|^2 \\ 2\mathrm{Re}\left(E_x^* E_y\right) \\ 2\mathrm{Im}\left(E_x^* E_y\right) \end{bmatrix}. \tag{1}$$

The total power, $S_0 = |E_x|^2 + |E_y|^2$, is the norm of the Stokes vector. On the other hand, the Jones vector for the signal electric fields is written as

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \sqrt{S_0} \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\psi}{2}\right) \\ i\sin\left(\frac{\psi}{2}\right) \end{bmatrix}. \tag{2}$$

The electric fields, E_x and E_y, generally constitute an elliptical SOP. In Eq. (2), $\theta/2$ denotes the tilting angle of the ellipse, and $\psi/2$ represents the degree of ellipticity. Then, the Stokes vector corresponding to such an elliptical SOP is given from Eqs. (1) and (2) as

$$\mathbf{S} = S_0 \begin{bmatrix} \cos\psi\cos\theta \\ \cos\psi\sin\theta \\ \sin\psi \end{bmatrix}. \tag{3}$$

Equation (3) demonstrates that any SOP can be expressed as a vector, whose spherical coordinates are (S_0, θ, ϕ). Thus, we can design the constellation diagram for multi-level SOP modulation formats in the 3D vector space.

The Stokes vector given by Eq. (1) can be written in different ways according to the change of bases. For example, Table I shows the Stokes parameters using the $45°$- and $135°$-linear polarization components, $E_{45°}$ and $E_{135°}$, respectively, and the left-circular- and right-circular-polarization components, E_L and E_R, respectively. Therefore, we can also express the Stokes vector as follows:

$$\mathbf{S} = \begin{bmatrix} |E_x|^2 - |E_y|^2 \\ |E_{45°}|^2 - |E_{135°}|^2 \\ |E_L|^2 - |E_R|^2 \end{bmatrix}, \tag{4}$$

using different bases for each Stokes parameter.

B. Stokes-Vector Receivers

From Eq. (1), we obtain the configuration of the *coherent-detection-like* Stokes-vector receiver shown in Fig. 1 [4]. The signal light is separated into the x- and y-polarization components using a polarization beam splitter (PBS). Each polarization component is further split into two branches. The power ratio between the branches is 1:2. The light powers divided into $1/3$ are detected directly. Meanwhile, the lights, whose powers are divided into $2/3$, are led to the $90°$ optical hybrid after polarization matching using a half-wave ($\lambda/2$) plate. Balanced photodiode (BPD) outputs, I_1, I_2, and I_3, give us elements of the Stokes vector, $S_1/3$, $S_2/3$, and $S_3/3$, respectively. Note that the detected power is 1/3 of the incident power because of the one-to-three branching loss in the receiver required for simultaneous measurement of all three Stokes parameters.

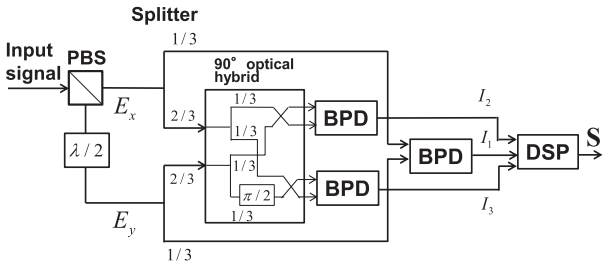

Fig. 1. Configuration of the *coherent-detection-like* Stokes-vector receiver based on Eq. (1).

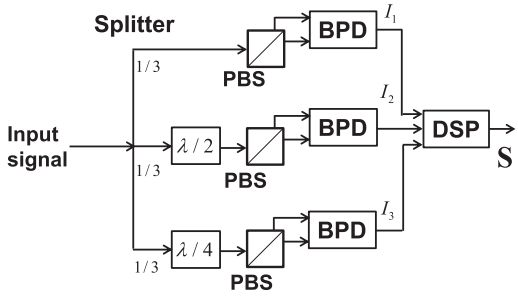

Fig. 2. Configuration of the *direct-detection-like* Stokes-vector receiver based on Eq. (4).

The configuration of the *direct-detection-like* Stokes-vector receiver is shown in Fig. 2, as obtained from Eq. (4) [5]. The incoming signal is equally split into three branches. In the first, a PBS splits the signal into its x- and y-polarization components. In the second, a PBS following a $\lambda/2$ plate splits the signal into its $45°$- and $135°$-linearly-polarized components. In the third, a PBS following a quarter-wave ($\lambda/4$) plate splits the signal into its right- and left-circularly-polarized components. BPD outputs, I_1, I_2, and I_3, give us elements of the Stokes vector, $S_1/3$, $S_2/3$, and $S_3/3$, respectively. In this receiver, the detected power is 1/3 of the incident power because of the one-to-three branching loss.

Regardless of the configuration, I_1, I_2, and I_3 are sent to the DSP circuit for polarization tracking and symbol discrimination [2], [3]. In the following sections, we assume that random fluctuations of SOP are fully compensated in the digital domain.

C. Modulation Formats

In this study, we consider the cubic-lattice SOP modulation format [5]. Figure 3 shows the 3D constellation map, where S_1, S_2, and S_3 are modulated in a binary manner. Hence, eight

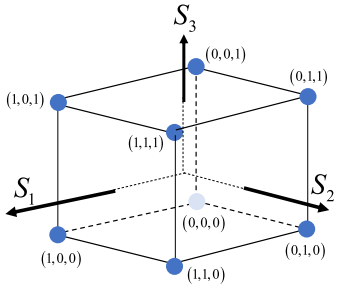

Fig. 3. Constellation map of the cubic modulation format in the Stokes space. The gray-code type assignment of the three bits is also indicated.

constellation points constitute a cube in the 3D Stokes-vector space. Hereafter, we call this modulation format POL8. The gray-code type assignment of the three bits is also indicated. When the modulation level is increased to four, the constellation map creates the cubic lattice. The number of constellation points is 64. Hereafter, this modulation format is called POL64.

Since each Stokes-vector element is independently modulated in these modulation formats, the BER performance can be evaluated by a 1-dimensional (1D) analysis, as shown in Section V. Such a cubic-lattice modulation format is the 3D extension of the 2D quadrature amplitude modulation (QAM) format, which is most commonly used in coherent optical communication systems.

III. Quantum Noise Characteristics of Stokes Receivers

This section deals with the quantum mechanical description of the Stokes vector. The quantum noise characteristics of the Stokes vector and the noise associated with the measurement are discussed.

A. Shot Noise on the Stokes Vector

Equation (1) yields the quantum mechanical Stokes parameters expressed as

$$\hat{S}_1 = \hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}, \tag{5}$$

$$\hat{S}_2 = \hat{a}^\dagger \hat{b} + \hat{a}\hat{b}^\dagger, \tag{6}$$

$$\hat{S}_3 = i\left(\hat{a}\hat{b}^\dagger - \hat{a}^\dagger \hat{b}\right), \tag{7}$$

where $\hat{a}^\dagger$ and $\hat{a}$ are the creation and annihilation operators for the x-polarization component, respectively, whereas $\hat{b}^\dagger$ and $\hat{b}$ are those for the y-polarization component [11]. These equations are obtained by replacing the electric fields, $E_x^\star$, E_x, $E_y^\star$, and E_x in Eq. (1) with the corresponding operators, $\hat{a}^\dagger$, $\hat{a}$, $\hat{b}^\dagger$, and $\hat{b}$, respectively.

To analyze the shot-noise characteristics of the Stokes parameters, we express the annihilation operators for the x- and y-polarization components as follows:

$$\hat{a} = A_0 + \Delta\hat{a}, \tag{8}$$

$$\hat{b} = \Delta\hat{b}, \tag{9}$$

where A_0 is a real constant, and $\Delta\hat{a}$ and $\Delta\hat{b}$ are zero-mean vacuum fluctuation operators. When the signal is in the coherent state [12], we have

$$\left\langle \hat{S}_1 \right\rangle = A_0^2 = n_0, \tag{10}$$

$$\left\langle \hat{S}_2 \right\rangle = 0, \tag{11}$$

$$\left\langle \hat{S}_3 \right\rangle = 0, \tag{12}$$

where $\langle \hat{x} \rangle$ means taking an average of $\hat{x}$, and n_0 denotes the average number of photons. Furthermore, Eqs. (5)–(9) yield the following equations for fluctuations in the Stokes parameters:

$$\Delta\hat{S}_1 = 2A_0\Delta\hat{a}_r + \left(\Delta\hat{a}^\dagger\Delta\hat{a} - \Delta\hat{b}^\dagger\Delta\hat{b}\right), \tag{13}$$

$$\Delta\hat{S}_2 = 2A_0\Delta\hat{b}_r + \left(\Delta\hat{a}^\dagger\Delta\hat{b} + \Delta\hat{a}\Delta\hat{b}^\dagger\right), \tag{14}$$

$$\Delta\hat{S}_3 = 2A_0\Delta\hat{b}_i + i\left(\Delta\hat{a}\Delta\hat{b}^\dagger - \Delta\hat{a}^\dagger\Delta\hat{b}\right). \tag{15}$$

The subscripts, r and i, in the first terms of Eqs. (13)–(15) represent the real and imaginary parts, respectively, of the fluctuation operators $\Delta\hat{a}$ and $\Delta\hat{b}$, which are given by

$$\Delta\hat{a}_r = \frac{\Delta\hat{a} + \Delta\hat{a}^\dagger}{2}, \tag{16}$$

$$\Delta\hat{a}_i = \frac{\Delta\hat{a} - \Delta\hat{a}^\dagger}{2i}, \tag{17}$$

$$\Delta\hat{b}_r = \frac{\Delta\hat{b} + \Delta\hat{b}^\dagger}{2}, \tag{18}$$

$$\Delta\hat{b}_i = \frac{\Delta\hat{b} - \Delta\hat{b}^\dagger}{2i}. \tag{19}$$

The vacuum fluctuation associated with the signal electric field in the coherent state has the two-dimensional Gaussian PDF on the complex plane, whose variances are given as

$$\left\langle \Delta\hat{a}_r^2 \right\rangle = \left\langle \Delta\hat{a}_i^2 \right\rangle = \left\langle \Delta\hat{b}_r^2 \right\rangle = \left\langle \Delta\hat{b}_i^2 \right\rangle = \frac{1}{4}. \tag{20}$$

From Eqs. (13)–(15) and (20), we find that

$$\left\langle \Delta\hat{S}_1^2 \right\rangle = \left\langle \Delta\hat{S}_2^2 \right\rangle = \left\langle \Delta\hat{S}_3^2 \right\rangle = n_0, \tag{21}$$

noting that the variances stem only from the first terms of Eqs. (13)–(15), and the second terms of these equations do not contribute to the variances. Equation (21) implies that the isotropic "3D shot noise" is associated with the average Stokes vector, $\left\langle \hat{\mathbf{S}} \right\rangle = [n_0, 0, 0]^{\mathrm{T}}$. This shot-noise formula holds regardless of the direction of the Stokes vector.

B. Uncertainty Relation for the Stokes Parameters

Using the commutation relation for $\Delta\hat{b}_r$ and $\Delta\hat{b}_i$ [12],

$$\left[\Delta\hat{b}_r, \Delta\hat{b}_i\right] \triangleq \Delta\hat{b}_r\Delta\hat{b}_i - \Delta\hat{b}_i\Delta\hat{b}_r = \frac{i}{2}, \tag{22}$$

we can derive the commutation relations for the Stokes parameters from Eqs. (13)–(15) as

$$\left\langle \left[\Delta\hat{S}_2, \Delta\hat{S}_3\right]\right\rangle = 4A_0^2[\Delta\hat{b}_r, \Delta\hat{b}_i] = i2n_0, \tag{23}$$

$$\left\langle \left[\Delta\hat{S}_1, \Delta\hat{S}_2\right]\right\rangle = 0, \tag{24}$$

$$\left\langle \left[\Delta\hat{S}_3, \Delta\hat{S}_1\right]\right\rangle = 0. \tag{25}$$

The following Robertson's inequality holds for any arbitrary operators and quantum states [13]:

$$\left\langle \Delta\hat{A}^2\right\rangle \left\langle \Delta\hat{B}^2\right\rangle \geq \frac{1}{4}\left|\left\langle \left[\Delta\hat{A}, \Delta\hat{B}\right]\right\rangle\right|^2. \tag{26}$$

Using Eq.(26), Eq. (23) leads to the Heisenberg uncertainty relation, expressed as

$$\left\langle \Delta\hat{S}_2^2\right\rangle \left\langle \Delta\hat{S}_3^2\right\rangle \geq n_0^2. \tag{27}$$

The 3D shot noise (Eq. (21)) associated with the coherent state satisfies the minimum uncertainty between $\left\langle \Delta\hat{S}_2^2\right\rangle$ and $\left\langle \Delta\hat{S}_3^2\right\rangle$, as determined from Eq. (27). However, using the squeezed vacuum for $\Delta\hat{b}$ [14], it is possible to reduce either $\left\langle \Delta\hat{S}_2^2\right\rangle$ or $\left\langle \Delta\hat{S}_3^2\right\rangle$ below the shot noise level, as far as Eq. (27) holds. On the other hand, $\left\langle \Delta\hat{S}_1^2\right\rangle$ can be smaller than the shot noise level. This is achieved by using the squeezed state for $\hat{a}$, where $\langle\Delta\hat{a}_r^2\rangle \leq 1/4$, instead of using the coherent state [14].

C. Quantum Noise Associated With Stokes-Vector Reception

The discussions presented in III-A and III-B are valid when the three Stokes parameters are measured individually. In other words, in Figs. 1 and 2, we measure the Stokes parameters without using the splitter on a one-by-one basis. However, in real systems, all the parameters must be measured simultaneously. From the simultaneously measured Stokes parameters, we estimate the Stokes parameters of the input signal, which are called the input-referred Stokes parameters. The fluctuation operators for the input-referred Stokes parameters, $\Delta\hat{S}_{m1}$, $\Delta\hat{S}_{m2}$, and $\Delta\hat{S}_{m3}$, must satisfy

$$\left\langle \left[\Delta\hat{S}_{m1}, \Delta\hat{S}_{m2}\right]\right\rangle = 0, \tag{28}$$

$$\left\langle \left[\Delta\hat{S}_{m2}, \Delta\hat{S}_{m3}\right]\right\rangle = 0, \tag{29}$$

$$\left\langle \left[\Delta\hat{S}_{m3}, \Delta\hat{S}_{m1}\right]\right\rangle = 0, \tag{30}$$

in order that each Stokes parameter can be measured independently.

Let the receiver noise for $\hat{S}_2$ be $\Delta\hat{N}_2$ and that for $\hat{S}_3$ be $\Delta\hat{N}_3$. Then, we have

$$\Delta\hat{S}_{m2} = \Delta\hat{S}_2 + \Delta\hat{N}_2, \tag{31}$$

$$\Delta\hat{S}_{m3} = \Delta\hat{S}_3 + \Delta\hat{N}_3. \tag{32}$$

From Eqs. (23), (29), (31), and (32), we obtain

$$\left\langle \left[\Delta\hat{N}_2, \Delta\hat{N}_3\right]\right\rangle = -i2n_0. \tag{33}$$

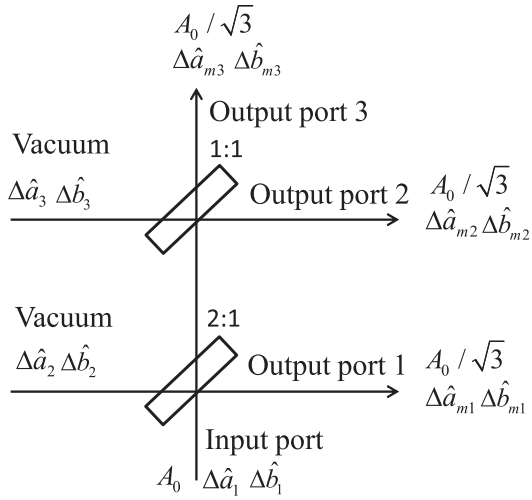

Fig. 4. A model of one-to-three beam splitting for simultaneous measurements of Stokes parameters. When the signal electric field, A_0, is divided, the vacuum fluctuations inevitably merge into each output port.

Using Robertson's inequality (Eq. (26)), we have

$$\left\langle \Delta\hat{N}_2^2\right\rangle \left\langle \Delta\hat{N}_3^2\right\rangle \geq n_0^2. \tag{34}$$

Then, the uncertainty relation between $\Delta\hat{S}_{m2}$ and $\Delta\hat{S}_{m3}$ is given as

$$\begin{aligned}
&\left\langle \Delta\hat{S}_{m2}^2\right\rangle \left\langle \Delta\hat{S}_{m3}^2\right\rangle \\
&\quad = \left(\left\langle \Delta\hat{S}_2^2\right\rangle + \left\langle \Delta\hat{N}_2^2\right\rangle\right)\left(\left\langle \Delta\hat{S}_3^2\right\rangle + \left\langle \Delta\hat{N}_3^2\right\rangle\right) \\
&\quad \geq \left(\left\langle \Delta\hat{S}_2^2\right\rangle + \left\langle \Delta\hat{N}_2^2\right\rangle\right)\left(\left\langle \Delta\hat{S}_3^2\right\rangle + \left\langle \Delta\hat{N}_3^2\right\rangle\right) \\
&\quad - \left(\sqrt{\left\langle \Delta\hat{S}_2^2\right\rangle \left\langle \Delta\hat{N}_3^2\right\rangle} - \sqrt{\left\langle \Delta\hat{S}_3^2\right\rangle \left\langle \Delta\hat{N}_3^2\right\rangle}\right)^2 \\
&\quad = \left(\sqrt{\left\langle \Delta\hat{S}_2^2\right\rangle \left\langle \Delta\hat{S}_3^2\right\rangle} + \sqrt{\left\langle \Delta\hat{N}_2^2\right\rangle \left\langle \Delta\hat{N}_3^2\right\rangle}\right)^2 \\
&\quad \geq 4n_0^2.
\end{aligned} \tag{35}$$

Comparing Eq. (27) with Eq. (35), we find that the uncertainty product increases by four times, owing to the noise generated in the measurement system. In the most common case, where $\left\langle \Delta\hat{S}_{m2}^2\right\rangle = \left\langle \Delta\hat{S}_{m3}^2\right\rangle$, we find that the variance of the noise in the input-refereed Stokes parameters, $\hat{S}_{m2}$ and $\hat{S}_{m3}$, is doubled.

D. Effect of the Power Branching in the Stokes-Vector Receiver

In actual Stokes-vector receivers, as shown in Figs. 1 and 2, the incident signal power is divided into three branches to measure the three parameters simultaneously. In what follows, we demonstrate that the one-to-three signal power branching realizes the conditions shown by Eqs. (28)–(30).

Figure 4 is the model of the branching circuit, where the input signal is equally split into three output ports using two beam

splitters. The ratio of the branched powers of the first beam splitter is 2:1, and that of the second beam splitter is 1:1. In output ports 1, 2, and 3, the Stokes parameters, S_1, S_2, and S_3, are measured, respectively. The input-referred Stokes parameters are obtained by multiplying the simultaneously measured Stokes parameters by three.

The signal electric field is equally divided into the three ports, and the divided electric field are given as $A_0/\sqrt{3}$. In such a case, the vacuum fluctuations of the x-polarization component from the three ports, $\Delta\hat{a}_1$, $\Delta\hat{a}_2$, and $\Delta\hat{a}_3$, are merged into the output ports, as indicated in Fig. 4. Then, the fluctuation operator of the x-polarization component in each port is given as

$$\Delta\hat{a}_{m1} = \frac{\Delta\hat{a}_1}{\sqrt{3}} + \sqrt{\frac{2}{3}}\Delta\hat{a}_2, \tag{36}$$

$$\Delta\hat{a_{m2}} = \frac{\Delta\hat{a}_1}{\sqrt{3}} - \frac{\Delta\hat{a}_2}{\sqrt{6}} + \frac{\Delta\hat{a}_3}{\sqrt{2}}, \tag{37}$$

$$\Delta\hat{a}_{m3} = \frac{\Delta\hat{a}_1}{\sqrt{3}} - \frac{\Delta\hat{a}_2}{\sqrt{6}} - \frac{\Delta\hat{a}_3}{\sqrt{2}}. \tag{38}$$

The fluctuation operators for the y-polarization component are obtained by replacing $\Delta\hat{a}$ with $\Delta\hat{b}$ in Eqs. (36)–(38). Therefore, from Eqns. (13)–(15), we have the following input-referred fluctuation operators for the Stokes vector:

$$\Delta\hat{S}_{m1} = 2\sqrt{3}A_0\left[\frac{\Delta\hat{a}_{1r}}{\sqrt{3}} + \sqrt{\frac{2}{3}}\Delta\hat{a}_{2r}\right], \tag{39}$$

$$\Delta\hat{S}_{m2} = 2\sqrt{3}A_0\left[\frac{\Delta\hat{b}_{1r}}{\sqrt{3}} - \frac{\Delta\hat{b}_{2r}}{\sqrt{6}} + \frac{\Delta\hat{b}_{3r}}{\sqrt{2}}\right], \tag{40}$$

$$\Delta\hat{S}_{m3} = 2\sqrt{3}A_0\left[\frac{\Delta\hat{b}_{1i}}{\sqrt{3}} - \frac{\Delta\hat{b}_{2i}}{\sqrt{6}} - \frac{\Delta\hat{b}_{3i}}{\sqrt{2}}\right], \tag{41}$$

where we ignore the second terms in Eqs. (13)–(15) because all commutations stemming from theses terms are zero. Consequently, we find that Eqs. (28)–(30) are satisfied in this power-branching system, using the commutation relations between the real and imaginary parts (see Eq. (22)). This result is valid for any arbitrary Stokes vector.

When the signal is in the coherent state, the variances of the fluctuations of the input-referred Stokes parameters are given from Eqs. (39)–(41) as

$$\left\langle\Delta\hat{S}^2_{m1}\right\rangle = \left\langle\Delta\hat{S}^2_{m2}\right\rangle = \left\langle\Delta\hat{S}^2_{m3}\right\rangle = 3n_0. \tag{42}$$

In this manner, the one-to-three power branching in the Stokes-vector receiver enables simultaneous measurements of the Stokes parameters. However, the input-referred shot noise increases by a factor of three (4.8 dB). The uncertainty product is given as

$$\left\langle\Delta\hat{S}^2_{m2}\right\rangle\left\langle\Delta\hat{S}^2_{m3}\right\rangle = 9n_0^2, \tag{43}$$

which is larger than the quantum limit of $4n_0^2$ given by Eq. (35).

It should be noted here that when we consider the measured Stokes vector, $\hat{\mathbf{S}}_m/3$, instead of the input-referred Stokes vector, $\hat{\mathbf{S}}_m$, its average is $n_0/3$, and its variance per dimension is $n_0/3$. Therefore, the signal-to-noise ratio is given by $n_0/3$ in each case.

E. Impact of Optical Pre-Amplifiers

When an optical pre-amplifier is employed in front of the one-to-three beam splitter, Eqs. (28)–(30) are satisfied. However, in such a case, the increase in the uncertainty product, $\left\langle\Delta\hat{S}^2_{m2}\right\rangle\left\langle\Delta\hat{S}^2_{m3}\right\rangle$, can be suppressed to the quantum limit, as shown below.

When an optical pre-amplifier has the spontaneous emission factor $n_{sp} = 1$ (i.e., the noise figure (NF) = 2 (3 dB)) and the gain G, the input signals, $\hat{a}$ and $\hat{b}$, are amplified as

$$\hat{a}_{out} = \sqrt{G}\hat{a} + \sqrt{G-1}\Delta c^\dagger, \tag{44}$$

$$\hat{b}_{out} = \sqrt{G}\hat{b} + \sqrt{G-1}\Delta d^\dagger, \tag{45}$$

where $\Delta\hat{c}^\dagger$ and $\Delta\hat{d}^\dagger$ are the creation operators of the vacuum fluctuations for the x- and y-polarization components, respectively, originating from the amplified spontaneous emission [14], [15]. Provided that $G \gg 1$, the input-referred annihilation operators are approximately given by

$$\hat{a}_{in} = \frac{\hat{a}_{out}}{\sqrt{G}} \simeq \hat{a} + \Delta\hat{c}^\dagger, \tag{46}$$

$$\hat{b}_{in} = \frac{\hat{b}_{out}}{\sqrt{G}} \simeq \hat{b} + \Delta\hat{d}^\dagger. \tag{47}$$

Using Eqs. (46) and (47), we can estimate the noise characteristics of the input-referred Stokes vector independently of the pre-amplifier gain and branching loss after the pre-amplifier.

In fact, Eqs. (39)–(41) are replaced with

$$\Delta\hat{S}_{m1} = 2A_0\left(\Delta\hat{a}_r + \Delta\hat{c}_r\right), \tag{48}$$

$$\Delta\hat{S}_{m2} = 2A_0\left(\Delta\hat{b}_r + \Delta\hat{d}_r\right), \tag{49}$$

$$\Delta\hat{S}_{m3} = 2A_0\left(\Delta\hat{b}_i - \Delta\hat{d}_i\right), \tag{50}$$

which, in turn, yield

$$\begin{aligned}\left[\Delta\hat{S}_{m2}, \Delta\hat{S}_{m3}\right] &= 4n_0\left(\left[\Delta\hat{b}_r, \Delta\hat{b}_i\right] - \left[\Delta\hat{d}_r, \Delta\hat{d}_i\right]\right)\\ &= 0.\end{aligned} \tag{51}$$

Furthermore, we can prove that Eqs. (28)–(30) hold for any arbitrary Stokes vector when we use an optical pre-amplifier.

Noting that in Eqs. (48)–(50),

$$\left\langle\Delta\hat{c}^2_r\right\rangle = \left\langle\Delta\hat{c}^2_i\right\rangle = \left\langle\Delta\hat{d}^2_r\right\rangle = \left\langle\Delta\hat{d}^2_i\right\rangle = 1/4, \tag{52}$$

we find that the input-referred noise exhibits variances given by

$$\left\langle\Delta\hat{S}^2_{m1}\right\rangle = \left\langle\Delta\hat{S}^2_{m2}\right\rangle = \left\langle\Delta\hat{S}^2_{m3}\right\rangle = 2n_0. \tag{53}$$

Thus, in this case, the increase in the input-referred noise is 3 dB, as compared with the 3D shot noise (Eq. (21)). Such a noise increase is understood to be a result of the NF of the pre-amplifier. The uncertainty product between $\Delta\hat{S}_2$ and $\Delta\hat{S}_3$

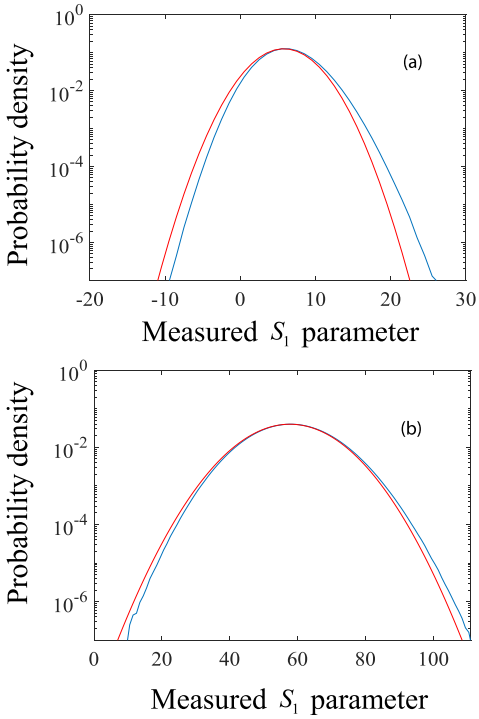

Fig. 5. Probability density functions for the shot noise: (a) $n_0/3 = 10$, (b) $n_0/3 = 100$. The red curves are calculated based on the Gaussian noise model. The blue curves are obtained by computer simulations.

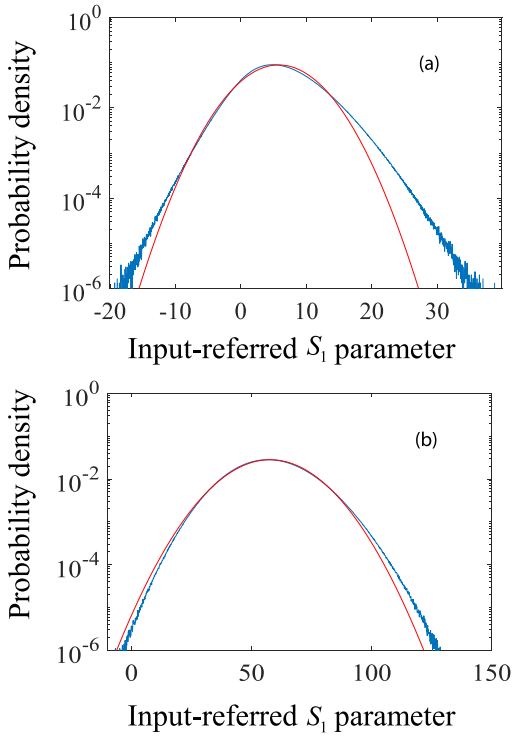

Fig. 6. Probability density functions for the pre-amplifier noise: (a) $n_0 = 10$, (b) $n_0 = 100$. The red curves are calculated based on the Gaussian noise model. The blue curves are obtained by computer simulations.

is given as

$$\left\langle \Delta \hat{S}_{m2}^2 \right\rangle \left\langle \Delta \hat{S}_{m3}^2 \right\rangle = 4n_0^2, \tag{54}$$

which is the minimum value required to make the Stokes parameters commutable, as shown by Eq. (35).

IV. Probability Density Functions of the Shot Noise and Optical Pre-Amplifier Noise

In this study, we calculate the PDFs of the Stokes parameters for the cubic polarization modulation format illustrated in Fig. 3. Through comparison with computer simulation results, we show the validity of the Gaussian noise model.

A. Shot Noise

We assume that the 3D shot noise after the one-to-three branching be approximated as the Gaussian noise. When the average number of input photons is n_0, the measured Stokes vector obeys a 3D Gaussian distribution with an average of $n_0/3$ and a variance of $n_0/3$ per dimension, as shown in III-D. The red curves in Fig. 5 illustrate the PDFs along the S_1 axis, calculated as a function of the measured S_1 value using the Gaussian noise model. We assume here that the measured Stokes vector is $\langle \hat{\mathbf{S}} \rangle = [1, 1, 1]^{\mathrm{T}} n_0/(3\sqrt{3})$, and $n_0/3$ =10, 100.

Moreover, the blue curves indicate the PDFs that were obtained by computer simulations. The simulation process is as follows. Let the average number of photons for the x polarization be n_{x0} and that for the y polarization be n_{y0}. To obtain the measured Stokes vector of $\langle \hat{\mathbf{S}} \rangle = [1, 1, 1]^{\mathrm{T}} n_0/(3\sqrt{3})$, n_{x0} and n_{y0} must be

$$n_{x0} = \frac{\sqrt{3}+1}{2\sqrt{3}} \frac{n_0}{3}, \tag{55}$$

$$n_{y0} = \frac{\sqrt{3}-1}{2\sqrt{3}} \frac{n_0}{3}, \tag{56}$$

and the phase difference between the x- and y-polarization components should be $\pi/4$. Quantum mechanically, it is well-known that the photon number distribution for the coherent state obeys the Poisson distribution. Therefore, in the simulations, the numbers of photons for the x- and y-polarization components are generated repeatedly according to Poisson distributions, whose averages are n_{x0} and n_{y0}. Taking the difference between these, we can obtain the S_1 value and its PDF using Eq. (1). We find that the PDFs of the Gaussian noise exhibit strong agreement with those obtained by the computer simulations.

B. Optical Pre-Amplifier Noise

When the optical pre-amplifier is employed, we assume that the input-referred Stokes vector obeys a 3D Gaussian distribution having an average of n_0 and a variance of $2n_0$ per dimension, as shown in III-E. The red curves in Fig. 6 indicate the PDFs along the S_1 axis, calculated as a function of the input-referred S_1 parameter based on the Gaussian noise model. We assume that the input-referred average Stokes vector is $\langle \hat{\mathbf{S}} \rangle = [1, 1, 1]^{\mathrm{T}} n_0/\sqrt{3}$, and n_0 =10, 100.

In Eqs. (46) and (47), the vacuum fluctuation parts are expressed as

$$\Delta\hat{a}_{in} = \Delta\hat{a} + \Delta\hat{c}^{\dagger}, \tag{57}$$

$$\Delta\hat{b}_{in} = \Delta\hat{b} + \Delta\hat{d}^{\dagger}. \tag{58}$$

Thus, we can demonstrate that

$$\left\langle \Delta\hat{a}_{in,r}^2 + \Delta\hat{a}_{in,i}^2 \right\rangle = \left\langle \Delta\hat{b}_{in,r}^2 + \Delta\hat{b}_{in,i}^2 \right\rangle = 1, \tag{59}$$

$$\left\langle \Delta\hat{a}_{in}^{\dagger}\Delta\hat{a}_{in} \right\rangle = \left\langle \Delta\hat{b}_{in}^{\dagger}\Delta\hat{b}_{in} \right\rangle = 1. \tag{60}$$

Equations (59) and (60) imply that the electric fields and directly detected signals exhibit fluctuations having one-photon energy. In such cases, we can treat Eqs. (57) and (58) as classical electric fields, the noise energies of which are one photon, and conduct computer simulations using the classical electric fields.

On the contrary, when the optical pre-amplifier is not used, the operators, $\Delta\hat{a}$ and $\Delta\hat{b}$, have the following quantum mechanical characteristics:

$$\left\langle \Delta\hat{a}_r^2 + \Delta\hat{a}_i^2 \right\rangle = \left\langle \Delta\hat{b}_r^2 + \Delta\hat{b}_i^2 \right\rangle = \frac{1}{2}, \tag{61}$$

$$\left\langle \Delta\hat{a}^{\dagger}\Delta\hat{a} \right\rangle = \left\langle \Delta\hat{b}^{\dagger}\Delta\hat{b} \right\rangle = 0. \tag{62}$$

Equations (61) and (62) indicate that the vacuum fluctuations have half-photon energy, whereas the vacuum fluctuations do not induce the photon-number fluctuations. In such a case, it is evident that the operators cannot be replaced with the classical electric fields.

The blue curves in Fig. 6 indicate the PDFs thus obtained by classical computer simulations, where the electric field is expressed as a constant amplitude and a Gaussian noise superimposed on it. We find that the PDFs of the Gaussian noise exhibit agreement with those obtained by the computer simulations, particularly when n_0 gets larger.

V. BER Characteristics

In this section, the BER characteristics for the cubic-lattice polarization modulation format are described. The BERs of the POL8 and POL64 formats are calculated using 1D analyses. The BER formulae are derived based on the Gaussian noise model.

A. POL8 Format

Let the coordinates of the binary-modulated signal on the S_1 axis be s_1 and s_2, and let the variances of the Gaussian noise associated with the signal be σ_1^2 and σ_2^2, respectively. The BER is then given as

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\frac{|s_1 - s_2|}{\sqrt{2}\left(\sqrt{\sigma_1^2} + \sqrt{\sigma_2^2}\right)}\right), \tag{63}$$

where $\mathrm{erfc}(*)$ represents the complimentary error function [16].

When the Stokes-vector receiver is operating in the shot-noise-limited state, we have the following input-referred S_1 values and the variances of the Gaussian noise:

$$s_1 = \frac{n_0}{\sqrt{3}}, \tag{64}$$

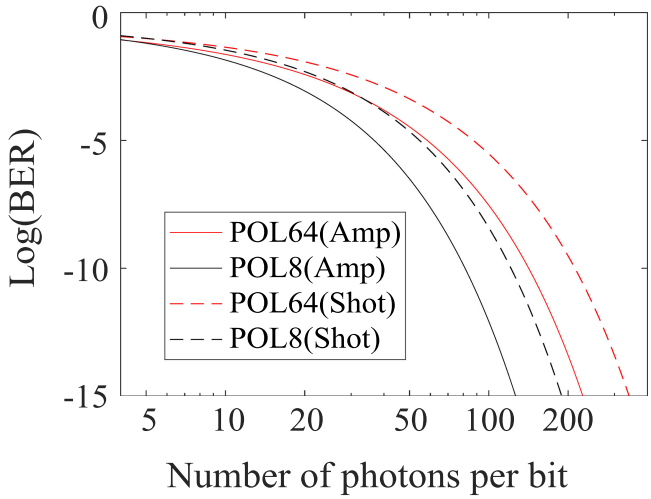

Fig. 7. BER characteristics for the cubic-lattice polarization modulation format. The solid curves are calculated when an optical pre-amplifier is used. The broken curves are obtained in the shot-noise-limited state. The red curves are the BERs for the POL64 format, and the black curves are the BERs for the POL8 format.

$$s_2 = -s_1, \tag{65}$$

$$\sigma_1^2 = \sigma_2^2 = 3n_0. \tag{66}$$

Subsequently, using Eqs. (63)–(66), the BER is expressed as

$$\mathrm{BER}_{shot} = \frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{n_0/3}}{\sqrt{6}}\right). \tag{67}$$

Note that $n_0/3$ denotes the number of input photons per bit.

On the other hand, when we employ an optical pre-amplifier, the input-referred signal has the same coordinates as Eqs. (64) and (65), but the variances of the Gaussian noise are given as

$$\sigma_1^2 = \sigma_2^2 = 2n_0. \tag{68}$$

Then, Eqs. (63)–(65), and (68) lead to the BER formula expressed as

$$\mathrm{BER}_{amp} = \frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{n_0/3}}{2}\right). \tag{69}$$

Comparing Eq. (69) with Eq. (67), we find that the optically pre-amplified receiver has a sensitivity higher than the shot-noise-limited receiver by a factor of 1.5 (1.8 dB). This is because the optical pre-amplifier compensates for the 4.8-dB branching loss of the receiver. Instead, the receiver sensitivity is degraded by 3 dB because of the NF of the optical pre-amplifier.

The black curves in Fig. 7 show the BER characteristics for the POL8 format. The solid curve is calculated from Eq. (69), when an optical pre-amplifier is used. The broken curve is obtained in the shot-noise-limited state, using Eq. (67). The horizontal axis is the number of photons per bit (i.e., $n_0/3$).

B. POL64 Format

We consider the POL64 format, where each Stokes parameter is 4-level modulated. The coordinates of the 4-level modulated

signal on the S_1 axis, s_1, s_2, s_3, and s_4, are given as

$$s_1 = \frac{y\overline{n}}{\overline{S}_0}, \tag{70}$$

$$s_2 = \frac{\overline{n}}{\overline{S}_0}, \tag{71}$$

$$s_3 = -s_2, \tag{72}$$

$$s_4 = -s_1, \tag{73}$$

where y (> 1) is an adjustable parameter to minimize the BER, and $\overline{S}_0$ is defined as

$$\overline{S}_0 = \frac{\sqrt{3y^2} + 3\sqrt{2y^2+1} + 3\sqrt{y^2+2} + \sqrt{3}}{8}, \tag{74}$$

so that $\overline{n}$ represents the average photon number for the POL64 format.

Then, the BER is approximately calculated as

$$\mathrm{BER} = \frac{\mathrm{BER}_1 + 2\mathrm{BER}_2 + \mathrm{BER}_3}{4}, \tag{75}$$

where

$$\mathrm{BER}_1 = \frac{1}{8}\mathrm{erfc}\left(\frac{\frac{\overline{n}}{\overline{S}_0}}{\sqrt{2}\sqrt{\sigma_2^2}}\right) + \frac{1}{4}\mathrm{erfc}\left(\frac{\frac{(y-1)\,\overline{n}}{\overline{S}_0}}{\sqrt{2}\left(\sqrt{\sigma_1^2}+\sqrt{\sigma_2^2}\right)}\right) \tag{76}$$

$$\mathrm{BER}_2 = \frac{1}{8}\mathrm{erfc}\left(\frac{\frac{\overline{n}}{\overline{S}_0}}{\sqrt{2}\sqrt{\sigma_3^2}}\right) + \frac{1}{4}\mathrm{erfc}\left(\frac{\frac{(y-1)\,\overline{n}}{\overline{S}_0}}{\sqrt{2}\left(\sqrt{\sigma_2^2}+\sqrt{\sigma_3^2}\right)}\right) \tag{77}$$

$$\mathrm{BER}_3 = \frac{1}{8}\mathrm{erfc}\left(\frac{\frac{\overline{n}}{\overline{S}_0}}{\sqrt{2}\sqrt{\sigma_4^2}}\right) + \frac{1}{4}\mathrm{erfc}\left(\frac{\frac{(y-1)\,\overline{n}}{\overline{S}_0}}{\sqrt{2}\left(\sqrt{\sigma_3^2}+\sqrt{\sigma_4^2}\right)}\right). \tag{78}$$

When the receiver is operating in the shot-noise-limited state, the variances of the input-referred Gaussian noise are given as

$$\sigma_1^2 = \frac{3\sqrt{3y^2}}{\overline{S}_0}\bar{n}, \tag{79}$$

$$\sigma_2^2 = \frac{3\sqrt{2y^2+1}}{\overline{S}_0}\bar{n}, \tag{80}$$

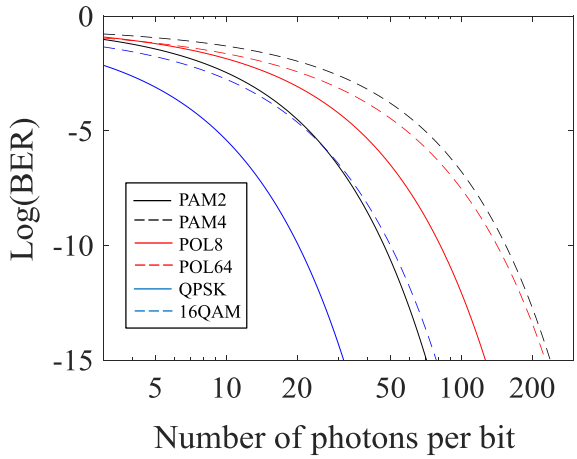

Fig. 8. Comparison of the BER characteristics among the PAM, QAM, and cubic-lattice SOP modulation formats. The black curves show the BER characteristics of the PAM format, the blue curves are those of the QAM format, and the red curves are those of the cubic-lattice SOP modulation format. The solid curves represent the case where the level of modulation per dimension is two, whereas the broken curves represent the case where the modulation level per dimension is four.

$$\sigma_3^2 = \frac{3\sqrt{y^2+2}}{\overline{S}_0}\bar{n}, \tag{81}$$

$$\sigma_4^2 = \frac{3\sqrt{3}}{\overline{S}_0}\bar{n}. \tag{82}$$

On the other hand, when an optical pre-amplifier is employed, the variances are obtained by replacing factor 3 with 2 in Eqs. (79)–(82).

The red curves in Fig. 7 show the BER characteristics for the POL64 format. The solid curve is calculated when an optical pre-amplifier is used. The broken curve is obtained in the shot-noise-limited state. The horizontal axis is the number of photons per bit (i.e., $\overline{n}/6$). The parameter, y, is adjusted to 5.2 to obtain the minimum BER. We find that for the POL64 format, the pre-amplified Stokes-vector receiver has a sensitivity higher than the shot-noise-limited receiver.

C. Comparison of the Receiver Sensitivity Among Various Multi-Level Modulation/Demodulation Schemes

In this section, we compare the receiver sensitivity among the pulse-amplitude modulation (PAM), QAM, and cubic-lattice SOP modulation. The PAM2 and PAM4 formats are demodulated by the pre-amplified direct-detection receiver, and the 4QAM (i.e., QPSK) and 16QAM formats are demodulated by the coherent receiver. Their sensitivities are compared with the POL8 and POL64 formats demodulated by the pre-amplified Stokes-vector receiver. The BER characteristics of the PAM and QAM formats can be analyzed by the method similar to that described in V-A and V-B.

In Fig. 8, the black curves show the BER characteristics of the PAM format. The blue curves are those of the QAM format, and the red curves are those of the cubic-lattice SOP modulation format. The solid curves represent the case where the level

of modulation per dimension is two (i.e., PAM2, 4QAM, and POL8), whereas the broken curves represent the case where the modulation level per dimension is four (i.e., PAM4, 16QAM, and POL64). The horizontal axis is the number of photons per bit.

We find that the coherent receiver has the best sensitivity. On the other hand, regarding the direct-detection receiver, the pre-amplified Stokes-vector receiver has a sensitivity higher than the pre-amplified direct-detection receiver when the level of modulation becomes four. Thus, the POL64 format seems to be an attractive modulation format because of its high capacity and high receiver sensitivity.

VI. Conclusion

We developed a quantum theory of noise for the Stokes-vector receiver. After deriving the conditions for enabling simultaneous measurement of all three Stokes parameters, we discussed the Gaussian noise model of the shot noise and optical pre-amplifier noise. The validity of the Gaussian noise model was verified using PDF calculations. The BERs for the cubic-lattice polarization modulation format were analyzed based on the Gaussian noise model.

References

[1] K. Kikuchi, "Electronic polarization-division demultiplexing based on digital signal processing in intensity-modulation direct-detection optical communication systems," *Opt. Express*, vol. 22, no. 2, pp. 1971–1980, Jan. 2014.

[2] K. Kikuchi, "Simple and efficient algorithm for polarization tracking and demultiplexing in dual-polarization IM/DD systems," presented at the Opt. Fiber Commun. Conf., Los Angeles, CA, USA, Mar. 22–26, 2015, Paper Th1E.3.

[3] K. Kikuchi and S. Kawakami, "Multi-level signaling in the stokes space and its application to large-capacity optical communications," *Opt. Express*, vol. 22, no. 7, pp. 7374–7387, Apr. 2014.

[4] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, "Stokes vector direct detection for linear complex optical channels," *J. Lightw. Technol.*, vol. 33, no. 3, pp. 678–684, Feb. 2015.

[5] K. Kikuchi, "Sensitivity analyses of stokes-vector receivers for cubic-lattice polarization modulation," presented at the Opto-Electron. Commun. Conf., Jeju, Korea, Jul. 2-6, 2018, Paper 5B1-2.

[6] K. Kikuchi, "Stokes-vector receivers and their performance analysis," presented at the IEEE Photon. Conf., Reston, VA, USA, Sep. 30- Oct. 4, 2018, Paper TuA2.1.

[7] S. Ghosh, T. Tanemura, Y. Kawabata, K. Katoh, K. Kikuchi, and Y. Nakano, "Decoding of multilevel stokes-vector modulated signal by polarization-analyzing circuit on InP," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 187–194, Jan. 2018.

[8] K. Kikuchi, "Noise characteristics of stokes vector receivers and bit-error-rate formulae for cubic polarization modulation," presented at the Opto-electron. Commun. Conf./ Photon. Switching Comput., Fukuoka, Japan, Jul. 7-11, 2019, Paper ThC3-1.

[9] K. Kikuchi, "Quantum theory of noise in stokes vector receivers and application to bit error rate analysis," presented at the Eur. Conf. Opt. Commun., Dublin, Ireland, Sep. 22-26, 2019, Paper P.8.

[10] C. Brosseau, *Fundamentals of Polarized Light*, Hoboken, NJ, USA: Wiley, 1998.

[11] P. Usachev, J. Sönderholm, G. Björk, and A. Trifinov, "Experimental verification of differences between classical and quantum polarization properties," *Opt. Commun.*, vol. 193, no. 1–6, pp. 161–173, Jun. 2001.

[12] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 157–179, Jan. 2016.

[13] H. P. Robertson, "The uncertainty principle," *Phys. Rev.*, vol. 34, no. 1, pp. 163–164, Jul. 1929.

[14] 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.

[15] S. Stenholm, "The theory of quantum amplifiers," *Physica Scripta*, vol. T-12, pp. 56–66, 1986.

[16] T. Okoshi and K. Kikuchi, *Coherent Optical Communication Systems*, Tokyo: KTK Scientific Publisher, Dordrecht, Boston, London, Tokyo: Kluwer Academic Publisher, 1988.

Kazuro Kikuchi (Fellow, IEEE) was born in Miyagi, Japan, in 1952. He received the B.S. degree in electrical engineering in 1974, the M.S. degree in electronic engineering in 1976, and the Ph.D. degree in electronic engineering in 1979, all from the University of Tokyo, Tokyo, Japan. He joined the University of Tokyo in 1979 and retired in 2016. He is currently a specially appointed Professor of the National Institution of Academic Degrees and Quality Enhancement of Higher Education, Japan. He has also worked with Bell Communications Research, NJ, USA from 1986 to 1987 as a Consultant and presently serves on the Board of Directors of Alnair Labs Corporation, Japan. Throughout his career, his research has focused on optical fiber communications, including optical devices and systems.

Dr. Kikuchi is a Fellow of the IEEE Photonics Society, a Fellow of OSA, and an Honorary Member and a Fellow of the Institute of Electronics, Information and Communication Engineers (IEICE). He is the recipient of numerous awards, including the IEICE Achievement Award, Japan IBM Science Prize, Hattori Hokosho Prize, Ericsson Telecommunications Award, Japanese Prime Minister's Award for the promotion of academy-industry collaboration, C&C Prize, and the John Tyndall Award.