Long-Haul Optical-Eigenvalue Transmission Using a Neural Network Demodulator and SD-FEC

Ken Mishina, *Member, IEEE*, Ryotaro Harada, Tsuyoshi Yoshida, *Member, IEEE*, Daisuke Hisano, *Member, IEEE*, and Akihiro Maruta, *Member, IEEE*

***Abstract*— Optical eigenvalue transmission based on inverse scattering transform (IST) has been studied for one of approaches to overcome the Kerr nonlinearity limit in fiber optic communications. In the recent decade, several multilevel modulation schemes that are based on IST, such as 16-ary and 64-ary signals using on–off encoding and b-modulation of multieigenvalues, have been proposed. To increase the transmission capacity and extend the transmission distance, applications of machine learning-based approaches to IST-based transmission have been proposed and demonstrated. Another prospective approach to increase the transmission capacity and extend the transmission distance in fiber optic communication involves the application of soft-decision forward error correction (SD-FEC). However, the applicability of SD-FEC to eigenvalue-modulated signals has not yet been investigated in detail because the distribution of the received signal is complicated for IST-based transmissions. In this paper, we describe in detail the theory of optical eigenvalue transmission, including the design of multilevel eigenvalue-modulated signals and the effects of noise and scaring parameters (coefficients for normalization of the nonlinear Schrödinger equation). We explain why neural network (NN) demodulators are advantageous for eigenvalue transmission systems. Consequently, we propose a combination of NN-based demodulators and SD-FEC decoding. A multilabel NN-based demodulator is employed to compute the L-value from the received eigenvalue pattern at the receiver. For a 16-ary eigenvalue-modulated signal, the proposed method outperformed a combination of the Gaussian approximation and SD-FEC in the simulation. Moreover, the experimental results show successful operation with error-free transmission through a 3000-km optical fiber line. In addition, we experimentally demonstrate the applicability of SD-FEC to a 4096-ary eigenvalue-modulated signal. The experimental results indicate that an achievable transmission distance can be extended to 1200 km using the NN demodulator and SD-FEC.**

***Index Terms*— Optical fiber communication, fiber nonlinear optics, neural networks, forward error correction.**

Received 4 April 2024; revised 13 September 2024; accepted 4 November 2024. Date of publication 19 February 2025; date of current version 14 May 2025. This work was supported by the National Institute of Information and Communications Technology (NICT), Japan, under Project J012368C08401. *(Corresponding author: Ken Mishina.)*

Ken Mishina, Ryotaro Harada, Daisuke Hisano, and Akihiro Maruta are with the Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan (e-mail: mishina@comm.eng.osaka-u.ac.jp; harada23@pn.comm.eng.osaka-u.ac.jp; hisano@comm.eng.osaka-u.ac.jp; maruta@comm.eng.osaka-u.ac.jp).

Tsuyoshi Yoshida is with the Information Technology Research and Development Center, Mitsubishi Electric Corporation, Kamakura 247-8501, Japan (e-mail: Yoshida.Tsuyoshi@ah.MitsubishiElectric.co.jp).

Digital Object Identifier 10.1109/JSAC.2025.3543512

I. Introduction

Digital transformation along with modern lifestyles and the advent of cyber-physical systems have resulted in a demand for high-capacity fiber optic communications. In the future, more high-capacity transmission systems will be required. In general, the transmission capacity can be increased by increasing the signal-to-noise ratio, i.e., increasing the power of the signal launched into the optical fiber line. However, under high-power signal conditions, Kerr nonlinearity in the optical fiber induces signal distortion, which is an important impairment that limits the transmission capacity and distance [1]. Several nonlinear compensation techniques have been proposed to reduce this limit, e.g., digital backpropagation (DBP) [2]. However, several problems remain in the implementation of these techniques in practical systems. For example, DBP requires high computational complexity because it computes signal backpropagation along the fiber longitudinal direction step by step or span by span in the digital domain. Furthermore, further computational resources are required to compensate for the effect of the cross-phase modulation and four-wave mixing with high accuracy in WDM transmission systems.

As a novel approach for overcoming the Kerr nonlinearity limit, optical eigenvalue communication [3], [4], [5], [6] based on the inverse scattering transform (IST) [7] has recently attracted considerable attention [8], [9]. IST is well-known as nonlinear Fourier transform (NFT). Eigenvalues that are associated with the lossless nonlinear Schrödinger equation (NLSE) are invariant during transmission in dispersive and nonlinear fibers despite changes in the waveform. In 1993, Hasegawa and Nyu proposed the concept of eigenvalue communication [3]. The development of digital coherent technology enables the implementation of the concept of eigenvalue communication [4], [5], [6], [10].

In the recent decade, various transmission schemes based on the NFT have been proposed and demonstrated. To increase the transmission capacity, the on–off encoding of multieigenvalues has been proposed [11], [12]. We experimentally demonstrated the 50-km transmission of a 4096-ary eigenvalue-modulated signal using a triangular-lattice-shaped configuration with 12 eigenvalues [13]. In another approach, several modulation schemes using the nonlinear spectrum and scattering coefficient b associated with eigenvalues were proposed [14], [15], [16], [17]. In [17], the transmission of 16QAM of the coefficient b using eigenvalues over 1200 km was experimentally demonstrated. Joint modulation using both continuous and discrete nonlinear spectra has also been reported [18], [19], for example, 3200-km transmission of dual-polarization QPSK signals on a continuous spectrum and two discrete spectra was reported in [19]. As with the DBP technique, problems of computational complexity for scattering data detection and application to WDM system persist; however, studies have focused on addressing these problems [20], [21].

To further improve the receiver sensitivity for multilevel modulation and long-haul transmission, machine learning-based approaches are valid even for NFT-based transmission systems, such as classification [22], [23] and equalization [26], [27]. Deviations in the eigenvalue and coefficient b due to amplified spontaneous emission noise are not independent and identically distributed (i.i.d.) using the circular Gaussian process, especially in cases involving multieigenvalue systems [28], [29]. In addition, the statistics of the eigenvalues and scattering coefficients are not completely understood when white Gaussian noise is added to time-domain signals of multieigenvalue transmission systems. Therefore, machine learning-based approaches are effective for demodulation, particularly in NFT-based transmission systems. Time-domain (TD) neural network (NN) [22] and eigenvalue-domain (ED) NN receivers [23] outperformed the conventional demodulation methods without NNs in terms of the bit error rate (BER) characteristics. The TD-NN receiver has a simple configuration; however, the TD-NN requires retraining of the NN when the transmission distance or receiver state changes, such as fluctuation of carrier frequency offset (CFO). By contrast, the ED-NN receiver exhibits high generalization performance in terms of transmission distance and receiver state [23], [24], although it requires more computation for the NFT compared to the TD-NN receiver. In [24], for a large CFO, a demodulation technique combining an eigenvalue-based CFO compensation and the ED-NN receiver has been demonstrated.

Furthermore, soft-decision forward error correction (SD-FEC) techniques have been applied to optical fiber communications to increase the transmission capacity and extend the transmission distance [30], [31]. The logarithmic ratio of *a posteriori* probabilities (L-value) is calculated from the received signals and generally utilized in SD-FEC. However, deriving the L-value from the received eigenvalue-modulated signal is complicated because the statistics of the eigenvalues and scattering coefficients are not completely understood. Therefore, the applicability of SD-FEC to eigenvalue-modulated signals has not yet been investigated in detail. Some previous studies on NFT-based transmissions have evaluated the BER characteristics before applying FEC with the FEC limit, assuming the use of hard-decision (HD) FEC or SD-FEC. Some studies have estimated achievable information rates using mutual information [25].

In this paper, we describe a comprehensive theory of optical eigenvalue transmission including the design of a multilevel eigenvalue-modulated signal and the effects of noise and scaling parameters. Consequently, we propose the combination of an NN-based demodulator and SD-FEC decoding. A multilabel NN-based demodulator is employed to compute the L-value from the eigenvalue input at the receiver. We investigated the applicability of SD-FEC to multilevel eigenvalue-modulated signals using both numerical simulations and experiments. An eigenvalue transmission using an on-off encoding is suitable for long-haul transmission because of its high tolerance to timing jitter and phase noise [5], [12]. Furthermore, because the noise distribution of eigenvalues becomes more complex when the number of eigenvalues is large, we studied the case that the proposed method is applied to multieigenvalue transmission with the on-off encoding, such as 4 and 12 eigenvalues. As an extension to our previous study [32], [33], in this paper, we describe in detail the demodulation method along with theoretical, numerical, and experimental investigations of basic characteristics and their applicability to 4096-ary eigenvalue-modulated signals. First, we describe the theory of eigenvalue modulation, eigenvalue-modulated signal design, and the concepts of the proposed method. By performing simulations, we demonstrate that the proposed method is valid for 16-ary eigenvalue-modulated signals, and shows a clear waterfall BER curve after SD-FEC. In addition, the combination of the NN demodulator and SD-FEC outperforms other methods using the Gaussian approximation without NN in terms of the BER after SD-FEC. Subsequently, we experimentally demonstrated error-free transmission through a 3000-km optical fiber. Finally, the applicability of the proposed method to another multilevel signal, i.e., a 4096-ary eigenvalue-modulated signal, was determined experimentally using 12 eigenvalues. Thus, the transmission distance of the 4096-ary eigenvalue-modulated signal can be extended to 1200 km using the NN demodulator and SD-FEC.

The remainder of this paper is organized as follows: Section II describes eigenvalue transmission with on-off encoding. Section III describes the NN-based demodulator for the SD-FEC proposed in this study. Section IV discusses the applicability of the proposed method to a 16-ary eigenvalue-modulated signal by performing numerical simulations and experiments. Section V discusses the experimental results obtained when the proposed method is applied to a 4096-ary eigenvalue-modulated signal. Finally, Section VI lists the major conclusions that were obtained from this study.

II. Eigenvalue Transmission

Sections II-A and B introduce the fundamental backgrounds of NLSE and IST, which were constructed by the related works. Sections II-C and D describes the eigenvalue transmission based on the on–off encoding and design of 4096-ary signal used in this paper. Sections II-E and F discuss the effects of noise and scaling parameter on the eigenvalue transmission.

A. Nonlinear Schrödinger Equation

Lightwave propagation in dispersive and nonlinear fiber can be described using NLSE [37]:

$$i\frac{\partial E}{\partial z} - \frac{\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + \gamma|E|^2 E = -i\alpha E, \tag{1}$$

where z, t, $E(z,t)$, β_2, γ, and α denote the propagation distance, time moving with average group velocity, complex envelope amplitude of the electric field, group velocity dispersion (GVD) parameter, nonlinear parameter, and loss coefficient, respectively. In (1), we assumed that the third-order dispersion and Raman scattering effects were negligible for a pulse width > 1 ps. Furthermore, a scalar propagation of single-polarization component is considered. We introduce the normalized distance Z, time T, and complex envelope amplitude q using base time t_0 as defined below

$$T = \frac{t}{t_0}, \quad Z = \frac{|\beta_2|}{t_0^2} z, \quad q = t_0 \sqrt{\frac{\gamma}{|\beta_2|}} E. \tag{2}$$

Considering the case of anomalous dispersion ($\beta_2 < 0$), NLSE (1) can be normalized as

$$i \frac{\partial q}{\partial Z} + \frac{1}{2} \frac{\partial^2 q}{\partial T^2} + |q|^2 q = -i\Gamma q, \tag{3}$$

where $\Gamma = \alpha t_0^2 / |\beta_2|$.

For a lossless fiber ($\Gamma = 0$), the normalized NLSE (3) can be analytically solved by IST. When $\Gamma \neq 0$, we applied an approximation of the guiding center soliton [38]. Considering the loss effect and periodic amplification, we further introduce the normalized amplitude u as

$$q = a(Z) u, \quad a(Z) = \sqrt{\frac{2\Gamma L_a}{1 - \exp(-2\Gamma L_a)}} \exp(-\Gamma Z), \tag{4}$$

where L_a is the normalized amplifier spacing. When L_a is sufficiently small, i.e., $L_a \ll 1$, the loss effect is negligible. Therefore, the normalized NLSE for $u(Z,T)$ is obtained as:

$$i \frac{\partial u}{\partial Z} + \frac{1}{2} \frac{\partial^2 u}{\partial T^2} + |u|^2 u = 0. \tag{5}$$

NLSE (5) can be solved by IST as well as the lossless case in (3). Note that the above path-average model can be applied to a single span between lumped optical amplifiers. As an alternative way to arrive at a lossless NLSE, schemes using distributed Raman amplification [34] and dispersion-decreasing fibers [35] have been reported.

When noise is added by optical amplifiers, (5) is modified by adding a noise term $n(Z,T)$ as follows

$$i \frac{\partial u}{\partial Z} + \frac{1}{2} \frac{\partial^2 u}{\partial T^2} + |u|^2 u = n(Z,T). \tag{6}$$

In this paper, we consider lumped periodic amplification using erbium-doped fiber amplifiers (EDFA), which provides random noise at $Z = mL_a (m = 1, 2, 3, \ldots, M)$ in M spans.

B. Inverse Scattering Transform

In 1972, Zakharov and Shabat found a Lax pair for the NLSE [36], and the details of the method to solve the NLSE were subsequently developed by Ablowitz and Segur [7]. Using the Ablowitz–Kaup–Newell–Segur (AKNS) formula, the scattering problem of the NLSE (5) is expressed by

$$\begin{cases} \dfrac{\partial \phi_1}{\partial T} = -i\zeta \phi_1 + iu\phi_2 \\[2ex] \dfrac{\partial \phi_2}{\partial T} = iu^* \phi_1 + i\zeta \phi_2 \end{cases} \tag{7}$$

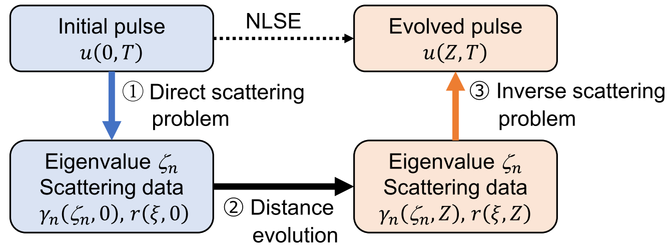

Fig. 1. Procedure to solve NLSE using IST.

where ζ and $\phi_l(Z,T)$ ($l = 1, 2$) denote the complex eigenvalues and eigenfunctions, respectively. (7) is called the Zakharov-Shabat eigenvalue problem. If u satisfies (5), ζ remains invariant with distance Z.

Fig. 1 depicts a block diagram of the procedure to solve the NLSE (5) using the IST [7]. The procedure is as follows: ① Direct scattering problem: Scattering data, that is, eigenvalue ζ_n, the norming constant $\gamma_n(\zeta_n, 0)$, and the reflection coefficient $r(\xi, 0)$ (for $\zeta = \xi \in \mathbb{R}$), are derived from the initial pulse $u(0,T)$. ② Distance evolution of scattering data: the evolution of the scattering data with distance Z is calculated. ③ Inverse scattering problem: Evolved pulse $u(Z,T)$ is constructed from the scattering data ζ_n, $\gamma_n(\zeta_n, Z)$, and $r(\xi, Z)$ at distance Z. In the field of optical fiber communications, processes ③ and ① are typically called NFT and inverse NFT (INFT), respectively.

In ①, the direct-scattering problem in (7) is considered treating $u(0,T)$ as a potential. In the case of vanishing boundary conditions, the reflected and transmitted waves were derived for incident waves from $T = \infty$. Assuming the solutions to satisfy the boundary conditions, the scattering coefficients a and b are obtained. The coefficients a and b respectively correspond to the amplitudes of the incident and reflected waves for $T \to \infty$. The norming constant γ_n and the reflection coefficient r are defined as follows:

$$\gamma_n = \frac{b(\zeta_n)}{a'(\zeta_n)}, \quad r(\xi) = \frac{b(\xi)}{a(\xi)}, \tag{8}$$

where

$$a'(\zeta_n) = \left. \frac{\partial a(\zeta)}{\partial \zeta} \right|_{\zeta = \zeta_n}, \quad a(\zeta_n) = 0. \tag{9}$$

In the NFT, γ_n and r are called the discrete and continuous nonlinear spectra, respectively. γ_n consists of the discrete eigenvalues $\zeta_n \in \mathbb{C}^+$, which corresponds to the soliton components. r consists of the continuous eigenvalues on the real axis $\xi \in \mathbb{R}$, which corresponds to the dispersive (radiation) wave components. A modulation scheme that uses both γ_n and r is known as a joint modulation scheme. The discrete eigenvalue ζ_n corresponds to the soliton components.

Next, we briefly describe the ② distance evolution of the scattering data. The evolution of the scattering coefficients is described by:

$$\begin{cases} a(\zeta_n, Z) = a(\zeta_n, Z = 0) \\ b(\zeta_n, Z) = b(\zeta_n, Z = 0) e^{2i\zeta_n^2 Z}. \end{cases} \tag{10}$$

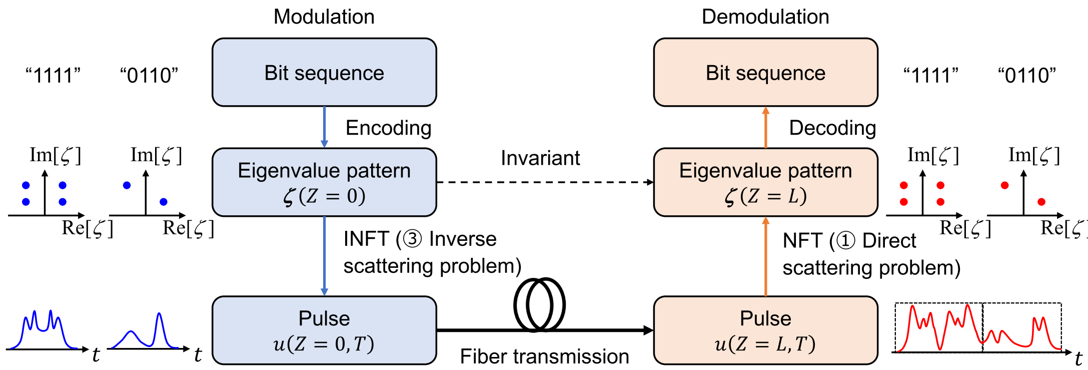

Fig. 2. Overview of eigenvalue modulation–demodulation scheme based on on–off encoding of four eigenvalues ($N = 4$).

From (8), the evolution of the spectral amplitude can be described as follows:

$$\begin{cases} \gamma_n(\zeta_n, Z) = \gamma_n(\zeta_n, Z = 0)e^{2i\zeta_n^2 Z} \\ r(\xi, Z) = r(\xi, Z = 0)e^{2i\xi^2 Z}. \end{cases} \tag{11}$$

As described in (10) and (11), the scattering coefficient b and nonlinear spectral amplitudes evolve linearly with Z. Therefore, b-modulation and the nonlinear spectral amplitude modulation prevent nonlinear distortion owing to the Kerr nonlinearity in a lossless ideal fiber.

In the ③ inverse scattering problem, the evolved pulse $u(Z, T)$ is constructed using ζ_n, $\gamma_n(\zeta_n, Z)$, $r(\xi, Z)$, and evolved eigen function. When no components exist in the nonlinear continuous spectrum, $u(Z, T)$ denotes an N-soliton solution. When $N = 1$, the solution corresponds to a fundamental soliton.

$$u(Z, T) = \eta \,\mathrm{sech}[(\eta T - T_s(Z)]e^{-i\kappa T - i\theta(Z)}, \tag{12}$$

where $\zeta = (\kappa + i\eta)/2$. The imaginary η and real κ parts of the eigenvalue correspond to the soliton amplitude (pulse width) and frequency, respectively. T_s and θ are related to the nonlinear spectral amplitude γ_n or scattering coefficient b.

In the recent decade, various transmission schemes that use scattering data, such as ζ, b, γ_n, and r, were proposed and demonstrated [5], [11], [12], [13], [16], [17], [18], [19]. These transmission schemes are known as NFT-based transmissions. It should be emphasized that the parameters evolve with distance Z linearly. In particular, the eigenvalue ζ is invariant during the transmission. Therefore, NFT-based transmission has the potential to overcome the Kerr nonlinearity limit by completely canceling the effects of nonlinear distortions under an ideal condition.

In the eigenvalue transmission, the information bits are encoded into the discrete eigenvalues ζ_n and the discrete nonlinear spectrum $\gamma_n(\zeta_n)$ (or $b(\zeta_n)$). In this method, the eigenvalue-modulated signal consists of only soliton components, not including dispersive wave. Because the soliton components are bounded in the time domain, especially when using the eigenvalues on the imaginary axis, it is less susceptible to inter-symbol interference effect from the neighbor symbols. Although the center position and phase of the soliton component respectively correspond to the amplitude and argument of the discrete nonlinear spectrum $\gamma_n(\zeta_n)$ (and $b(\zeta_n)$), the timing jitter and phase variation do not influence the position of the eigenvalue [5]. Therefore, transmission schemes that use only discrete eigenvalues for encoding are suitable for long-distance transmission because eigenvalue detection has a high tolerance to timing jitter and phase noise. By contrast, the spectral efficiency of the eigenvalue transmission is lower than that of joint modulation, which includes the continuous spectrum.

C. Eigenvalue Modulation Based on On–Off Encoding

In this study, we employed multieigenvalue transmission, which is based on the on–off encoding of multieigenvalues [12], [39]. Fig. 2 illustrates an overview of the multieigenvalue transmission system using on–off encoding. On–off encoding is based on a one-to-one mapping between an N-bit input and the subsets of N eigenvalues. If the value of the bit in the jth position is 1 (or 0), the jth eigenvalue is included (or excluded). An example of $N = 4$ is shown in Fig. 2. A sequence of four bits is mapped at the transmitter for the eigenvalue patterns. Note that on–off encoding is based on a one-to-one mapping between a four-bit input and the subsets of the four eigenvalues. In this scheme, γ_n (or b) is set to a constant value, not modulated. Consequently, the eigenvalue pattern was converted into a pulse using the inverse NFT (INFT) (③ inverse scattering problem). In this study, we used an INFT algorithm described in [39]. In [39], a pulse $u(Z, T)$ is generated by dividing the cases by $T \geq 0$ and $T < 0$ to avoid ill-conditioning problems. For the pulse generation, one eigenfunction is used for $T \geq 0$, whilst another eigenfunction is used for $T < 0$. By using this method, a stable pulse is obtained even for N-soliton with a large N such as 12. As another approach for the INFT, Darboux transformation-based algorithms are typically used, such as the proposed method in [6]. In the on–off encoding of N eigenvalues, the converted pulse corresponds to a symbol that carries N information bits. After the INFT, an eigenvalue-modulated signal is generated by connecting the converted pulses in series in the time domain. The optical eigenvalue-modulated signal is generated by an IQ modulator and transmitted into the fiber transmission line. As described in (7), the eigenvalue pattern is conserved during transmission in the absence of loss and noise.

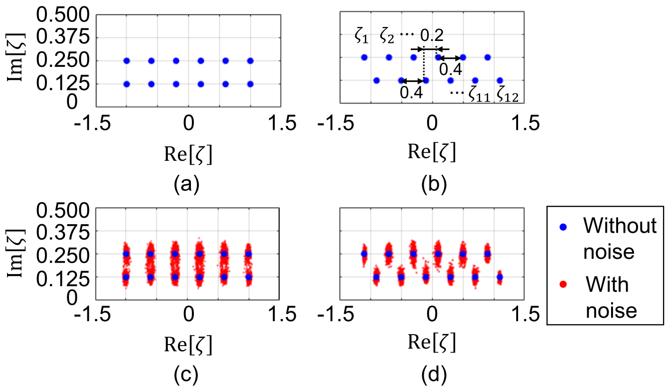

Fig. 3. Eigenvalue configurations: (a) Previous work without noise, (b) this work without noise, (c) previous work with noise, and (a) this work with noise.

An eigenvalue pattern is detected at the receiver using the NFT (① direct scattering problem) from the complex envelope amplitude obtained using a coherent receiver. When we use the Fourier collocation method [5], [10] for eigenvalue detection, the number of detected eigenvalues depends on the number of sampling points. For example, 32 eigenvalues are obtained when the sampling points per pulse is 32. Because the Fourier collocation assumes periodic boundaries, extra eigenvalues appear on the real axis, which correspond to the degenerate bands in the periodic NFT. Finally, the detected eigenvalue pattern is decoded into a bit sequence.

D. Design of 4096-Ary Eigenvalue-Modulated Signal

Fig. 3 shows eigenvalue configurations of 12 eigenvalues for a 4096-ary eigenvalue-modulated signal in previous work [23] ((a) and (c)) and this work ((b) and (d)) with/without noise. Figs. 3(c) and (d) shows the simulation results when white Gaussian noise was added to the time-domain pulse. In the previous work for a 16-ary eigenvalue-modulated signal using four eigenvalues, eigenvalues are lattice-shaped, such as the eigenvalue configuration depicted in Fig. 3(a). When a higher-level multilevel eigenvalue-modulated signal is generated, such as a 4096-ary signal using 12 eigenvalues, the eigenvalues must be allocated more densely because the available area on the complex eigenvalue plane is limited by the sampling rate and the bandwidth of the transmitter. When the sampling rate and bandwidth are limited, eigenvalues with large values in the real and imaginary parts are not detected accurately [40]. Because the real part of the eigenvalue expresses the soliton frequency, a pulse with a large real part of an eigenvalue requires a wideband transceiver. Similarly, because the imaginary part expresses the soliton amplitude, a pulse with a large imaginary part of an eigenvalue has a narrow pulse width and a wide spectral width and requires a wideband transceiver. It has been reported that insufficient sampling rates degraded the eigenvalue detection and the demodulation performance of the NN receivers [23]. By contrast, the dense eigenvalue allocation decreases the distance from the next signal points, which results in a decrease in noise tolerance. To improve noise tolerance, a triangular lattice configuration has been proposed for PSK modulation on the nonlinear spectral amplitudes of seven eigenvalues [15]. We applied the triangular lattice configuration, as depicted in Fig. 3(b), to the on-off encoding of multieigenvalue to emphasize noise tolerance [13]. Fig. 3(c) and (d) present the detected eigenvalue patterns after the Gaussian noise loading in the time domain for a signal-to-noise ratio (SNR) of 20 dB. The conventional lattice configuration exhibited evident eigenvalue errors because the deviation in the eigenvalue of the imaginary part exceeds that of the real part. In contrast, all 12 eigenvalues of the proposed configuration can be classified using a threshold on the complex eigenvalue plane. In this study, a triangular lattice-shaped eigenvalue pattern is used, as shown in Fig. 3(b) for a 4096-ary eigenvalue-modulated signal. The 12 eigenvalues are summarized as shown in Table I.

TABLE I

EIGENVALUES USED FOR 4096-ARY EIGENVALUE-MODULATED SIGNAL

$\zeta_1 = -1.1 + i0.25$	$\zeta_7 = -0.9 + i0.125$
$\zeta_2 = -0.7 + i0.25$	$\zeta_8 = -0.5 + i0.125$
$\zeta_3 = -0.3 + i0.25$	$\zeta_9 = -0.1 + i0.125$
$\zeta_4 = 0.1 + i0.25$	$\zeta_{10} = 0.3 + i0.125$
$\zeta_5 = 0.5 + i0.25$	$\zeta_{11} = 0.7 + i0.125$
$\zeta_6 = 0.9 + i0.25$	$\zeta_{12} = 1.1 + i0.125$

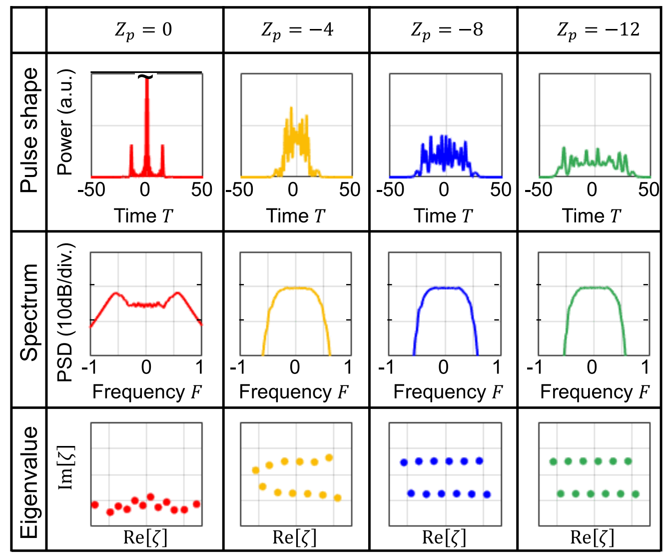

Fig. 4. Parameter Z_p dependence.

In addition, the initial pulse $u(Z,T)$ at the transmitter depends on the norming constants $\gamma_n(\zeta_n, Z)$ (or $b(\zeta_n, Z)$). It has been reported that a combination of $\gamma_n(\zeta_n, Z)$ was optimized to achieve better spectral efficiency [41] in multieigenvalue soliton transmission systems. However, considering a simple combination to evaluate the NN demodulator and the application of SD-FEC in this study, we used a combination of $b(\zeta_1) = b(\zeta_2) = \cdots = b(\zeta_{12}) = 1i$ at $Z = 0$ for all the eigenvalues.

Under the above condition, we introduced a pre-propagation parameter Z_p. The parameter Z_p expresses the distance of the pre-propagation at the transmitter. That is, we chose a soliton pulse of $u(T, Z_p)$ as a transmitting pulse. In other words, pre-propagated coefficients $b(\zeta_n, Z = 0)e^{2i\zeta_n^2 Z_p}$ are used for a transmitting pulse. Fig. 4 shows the pulse shapes, spectra, and detected eigenvalue patterns obtained by varying parameter Z_p when all 12 eigenvalues are in the "on" state. The eigenvalue patterns were detected under a sampling rate of 128 samples per pulse and a normalized time window width of 128 ($-64 \leq T \leq 64$). When $Z_p = 0$, the pulse has a high peak power and wide spectrum because 12 soliton components are multiplexed around $T = 0$. Consequently, the detected eigenvalue is degraded because of limited sampling points per pulse. We compressed the spectrum and peak-to-average power ratio by decreasing the value of Z_p. For $Z_p = -12$, although the eigenvalue pattern can be reconstructed, the necessary width of the time window increases. In this paper, we used $Z_p = -8$ with the balance of the temporal pulse width and spectral width [13]. In the experiments described in Section V, we set the window size to 64. Because the main pulse components were within the normalized time range of ± 30, the guard time was estimated to be approximately 4 (67 ps for $t_0 = 16.7$ ps) at the transmitter.

E. Effect of Noise

Amplified spontaneous emission (ASE) noise that is caused by optical amplifiers is one of the main factors that limit transmission capacity and distance in eigenvalue transmissions. Even if we assume that the deviation of each time-domain pulse sampling point caused by ASE noise is i.i.d. with the Gaussian process, deviations of detected eigenvalues are not i.i.d. with the Gaussian process. This is because soliton pulses consist of multiple sampling points in the time domain, which induces the interaction of ASE noise when the time-domain pulses are converted to eigenvalues via NFT. In other words, the Gaussian distribution of the noise in the time domain is converted into a non-Gaussian distribution using the nonlinear process. This is particularly complicated in multieigenvalue systems, where eigenvalue deviations are correlated with the other eigenvalues [42]. Fig. 5 presents the eigenvalue deviations for representative eigenvalue patterns. In the figure, white Gaussian noise was added to the time-domain pulse such that the optical signal-to-noise ratio (OSNR) is 3.6 dB. To evaluate the difference in the probability distribution, the Kullback–Leibler (KL) divergence D_{KL} between the actual distribution and Gaussian fitting is also shown in the figure. As shown in Fig. 5, the eigenvalue deviation depends on the position of the eigenvalue. Moreover, there may be changes in the deviation depending on the pattern, whereas the eigenvalues remained at the same position. In addition, the histogram of the eigenvalue deviation does not match the Gaussian distribution. For one eigenvalue such as $\zeta_1 = 0.4i$ and $\zeta_2 = 0.2i$, a steeper peak in the histogram is observed compared with the Gaussian fitting. For two eigenvalues of $\{\zeta_1, \zeta_2\} = \{0.4i, 0.2i\}$, the distributions are asymmetric and the values of D_{KL} are larger than those of one eigenvalue. This indicates that it is difficult to derive the L-value for SD-FEC from the detected eigenvalue pattern on the complex eigenvalue plane. To address this issue, this study presents an NN-based demodulator to compute the L-value.

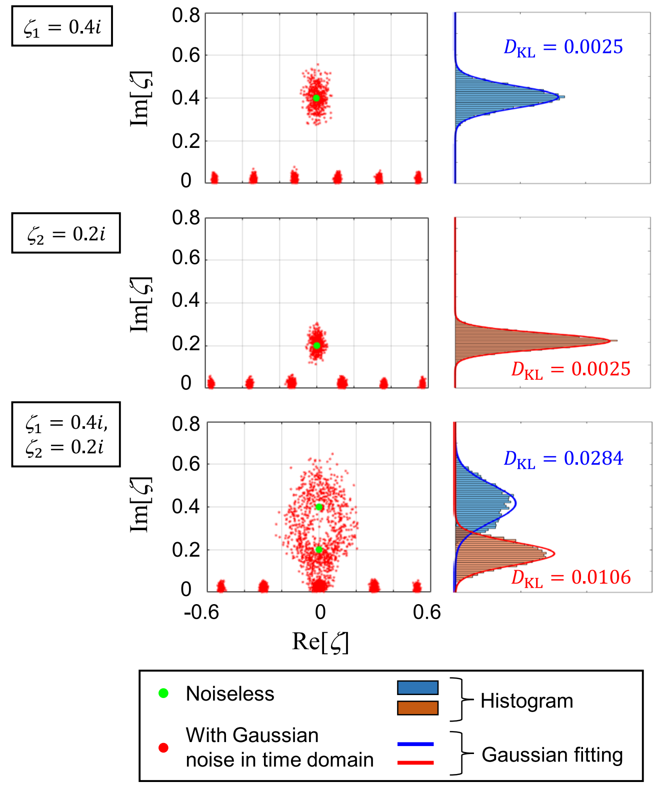

Fig. 5. Eigenvalue deviations of representative eigenvalue patterns with noise for OSNR = 3.6 dB.

The details of the proposed NN demodulator are described in Section III.

F. Effect of Scaling Parameter

The scaling parameter in the normalization (2) is another major factor that limits the transmission capacity and distance in the eigenvalue transmission. Although the eigenvalue is conserved during the transmission, the waveform and linear Fourier spectrum evolve with distance z. Here, z is the real distance, which is related to the normalized distance Z

$$z = \frac{t_0^2}{|\beta_2|} Z = z_0 Z \tag{13}$$

where t_0 is the basic time (time-scaling parameter) when used for normalization defined in (2) at the transmitter, and corresponds to the temporal pulse width for a fundamental soliton that has an eigenvalue of $\zeta = 0.5i$. Given the specified eigenvalues and initial parameters, the waveform and spectrum of the eigenvalue-modulated signal are determined by the normalized distance Z. Then, the real distance z is scaled by t_0 and β_2, as described in (13). The distance scaling parameter z_0 corresponds to $t_0^2/|\beta_2|$. To increase transmission capacity, the symbol rate can be increased by decreasing t_0. In this paper, we chose $t_0 = 16.7$ ps for the 4096-ary eigenvalue-modulated signal, which is a minimum value to achieve a stable modulation and demodulation for the back-to-back (B-to-B) configuration under the experimental environment in Sec. V. However, this approach decreases z_0 and induces a

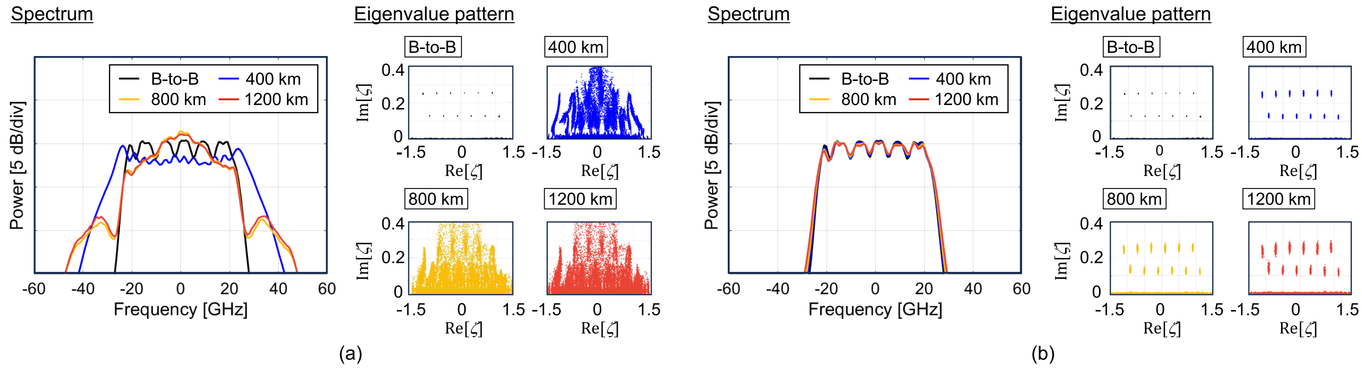

Fig. 6. Simulation results: Linear fourier spectra and detected eigenvalue patterns after the transmission for (a) $D = 4.4$ and (b) 0.7 ps/nm/km.

phenomenon in which the waveforms and spectrum change rapidly with real distance z.

Fig. 6 shows the linear Fourier spectra and detected eigenvalue patterns of the designed 4096-ary eigenvalue-modulated signal that was presented in the previous section for $t_0 = 16.7$ ps. The simulation results after transmission in non-zero dispersion-shifted fiber (NZ-DSF) with a dispersion parameter $D = 4.4$ ps/nm/km are shown in Fig. 6(a). In the simulation, the transmission loop consisted of a 50-km NZ-DSF and an EDFA with a noise figure of 5 dB. White Gaussian noise was added to the eigenvalue-modulated signal in the time domain at the end of each 50-km span. The lightwave propagation was calculated using the split-step Fourier method [37]. The distance scaling parameter z_0 was 50 km at a wavelength of 1550 nm. In this simulation, the pre-propagation parameter Z_p was set to -8 for the initial pulse at the time of transmission. Therefore, the pulse propagated in the reverse direction over a normalized distance of 8 was used as the initial pulse. At a transmission distance of 400 km, which corresponded to a normalized distance of 8, the soliton components of each eigenvalue were concentrated at $t = 0$, which resulted in a steep pulse and broadened spectrum. When sampling was limited to 60 GSa/s, the detected eigenvalue patterns were significantly degraded (as shown in Fig. 6(a)) because of the spectral broadening of the signal. Fig. 6(b) shows the simulation results after the transmission in an NZ-DSF with a dispersion parameter of $D = 0.7$ ps/nm/km. In this case, the distance scaling parameter z_0 was 311 km at a wavelength of 1550 nm. The spectrum changes slowly with the real distance z, and clear eigenvalue patterns were observed even after a transmission of 1200 km.

When the eigenvalues with different real parts are used for eigenvalue transmission, N-soliton solution is split to each soliton component in the time domain because of the different group velocities. In this case, inter-symbol interference from the neighbor symbols degrades the signal without enough guard time. By using an NZ-DSF with a small dispersion parameter that achieves a large z_0, the effect of the inter-symbol interference can also be relaxed because the waveform changes slowly with the real distance z.

Next, we discuss the relation between the loss and scaling parameter. In the simulation, the amplifier spacing $l_a (= z_0 L_a)$ and the fiber loss were set to 50 km and 0.2 dB/km, respectively. As described in Section II-A, the condition $l_a/z_0 \ll 1$ for the guiding center soliton [38] needs to be satisfied. However, for $D = 4.4$ ps/nm/km, the amplifier spacing l_a was comparable with the distance scaling parameter z_0. Therefore, the effect of the fiber loss is non-negligible, and eigenvalue deviations and spectral degradation occur after 400–1200 km, as shown in Fig. 6(a). On the other hand, for $D = 0.7$ ps/nm/km, the distance scaling parameter z_0 is 311 km and the normalized distance $Z = 8$ corresponds to the real distance $z = 2491$ km. The eigenvalue patterns show less degradation compared to the case with $D = 4.4$ ps/nm/km. The above simulation results indicate that the NZ-DSF with a small positive dispersion parameter is suitable, especially for the long-haul transmission of 4096-ary eigenvalue-modulated signals. Note that fibers (wavelength range) with zero dispersion and normal dispersion are inapplicable to the eigenvalue transmission in this study because bright solitons can not propagate such fibers.

III. Application of NN Demodulator and SD-FEC

In a previous study [13], [43], the eigenvalue pattern is decoded into a bit sequence using an NN-based classifier, assuming the use of an HD-FEC. In [13], an NN demodulator with multiclass classification was proposed, whereas in [43], an NN demodulator with multilabel classification was proposed. Fig. 7(a) provides an overview of the multilabel NN demodulator in [43]. The real and imaginary parts of all of the detected eigenvalues $\zeta_{\mathbf{r}}$ were provided as inputs to the NN. Note that extra eigenvalues on the real axis are obtained depending on the number of the sampling points per pulse when we use the Fourier collocation method. When the number of the sampling points is 32, we use the information of 32 eigenvalues including those on the real axis as the input to the NN. For a 16-ary eigenvalue-modulated signal, the number of output units is four, which corresponds to the number of eigenvalues for on–off encoding. The multilabel NN used a logistic sigmoid function for the output layer, which outputs *a posteriori* probability of the on-state of each bit corresponding to each eigenvalue. Assuming the use of an HD-FEC, the on–off state of the eigenvalues is identified using an appropriate threshold for the NN output. The input data were linked to the multilabel, namely the multi-on-state of the eigenvalues when more than one output exceeded

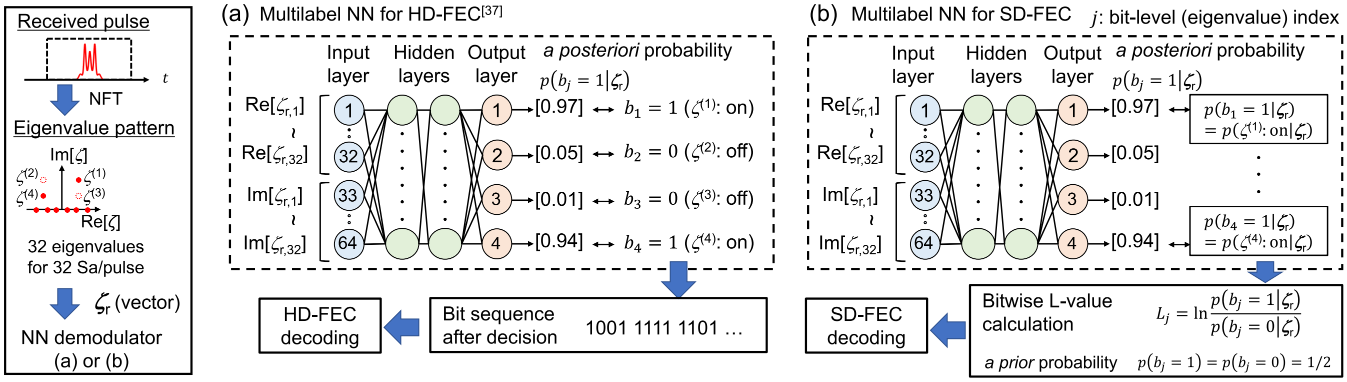

Fig. 7. Overview of NN demodulator for an eigenvalue-modulated signal using an on–off encoding of multieigenvalue: Multilabel NN demodulator for (a) HD-FEC (previous work) and (b) SD-FEC (proposed).

the threshold. The multilabel NN demodulator has a lower computational complexity and requires fewer training data when compared with multiclass classification NN [43]. Note that a typical feed-forward NN is sensitive to the order of its inputs. In other words, the order of the eigenvalues sent to the NN affects the demodulation performance. Therefore, in this study, the eigenvalues were sorted in an ascending order for the real part of eigenvalue before inputting to the NN.

In this study, we applied SD-FEC to the eigenvalue transmission system. The problem is that it is difficult to derive the L-value because the distributions of the eigenvalues obtained using NFT did not follow a Gaussian distribution [29]. Therefore, we propose the use of a multilabel NN-based demodulator to compute the L-value from the received eigenvalue pattern. Fig. 7(b) shows an overview of the proposed method. The NN configuration is identical to that of the multilabel NN demodulator described above. The input data are linked with the multilabel, in other words, *a posteriori* probability of the on-state of jth eigenvalue corresponding to $p(b_j = 1|\zeta_{\mathbf{r}})$. represents the output [44]. The bitwise L-value L_j can be calculated as follows:

$$L_j = \log \frac{p(b_j = 1|\boldsymbol{\zeta}_{\mathrm{r}})}{p(b_j = 0|\boldsymbol{\zeta}_{\mathrm{r}})} = \log \frac{p(b_j = 1|\boldsymbol{\zeta}_{\mathrm{r}})}{1 - p(b_j = 1|\boldsymbol{\zeta}_{\mathrm{r}})}. \tag{14}$$

The calculated L-value is input into the SD-FEC decoder. The SD-FEC decoder requires L-value in the rigorous definition [45], while it is sometimes called log-likelihood ratio (LLR) because those two are equivalent for uniformly distributed signals. Several studies that applied NN-based demodulators to the computation of the L-value or LLR for SD-FEC have been reported for other optical communication systems, such as coherent fiber-optic transmissions using QAM [46] and visible-light communication (VLC) [47].

IV. Application of NN Demodulator and SD-FEC to 16-Ary Eigenvalue-Modulated Signal

To demonstrate the feasibility of the proposed method, we performed numerical simulations and transmission experiments using a 16-ary eigenvalue-modulated signal with four eigenvalues.

A. Simulations

First, we investigate the validity of the multilabel NN demodulator for SD-FEC using numerical simulations. Fig. 8 shows the simulation model. We used the on–off states of the four eigenvalues, $\{\zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)}, \zeta^{(4)}\} = \{0.25 + i0.5, -0.25 + i0.5, 0.25 + i0.25, -0.25 + i0.25\} \in \mathbb{C}$ for the eigenvalue modulation. The scattering coefficients $b(\zeta^{(n)})$ were set to $1i$ at $Z = 0$ for all the eigenvalues. The initial soliton pulses at the transmitter were generated with a pre-propagation parameter of $Z_p = 8$. The basic time t_0 and the normalized time window size T_w were set to 50 ps and 32, respectively. The modulation was performed at 10 GSa/s, the pulse duration was 1.6 ns, and the bit rate was 2.5 Gb/s. The sampling rate of 10 GSa/s was sufficient to obtain clear eigenvalue patterns in this study [23] because 98% of the total power of the eigenvalue-modulated signal was within 5 GHz in the frequency spectrum at the transmitter. We confirmed the B-to-B operation to demonstrate the feasibility of the proposed method. Then, white Gaussian noise was added to the eigenvalue-modulated signal in the time domain.

At the receiver end, the eigenvalue patterns were detected from the received signal at 20 GSa/s. The number of sampling points per pulse was 32, which was sufficient to detect the eigenvalues accurately [23]. The bitwise L-value was computed using the multilabel NN-based demodulator, as discussed in the previous section. Further, a four-layer perceptron configuration was employed while using the rectified linear unit (ReLU) activation function; the number of hidden units was set to 256. For the loss function, we used a binary cross entropy function for each output of the NN-based demodulator and adopted the total loss of them. A total of 62,250 received pulses were divided into separate sequences of 10,000 and 52,250 pulses for training and BER test, respectively. Subsequently, the NN was trained using the Adam optimizer [48], and the following training patterns were examined: The NN was trained (i) for each OSNR using each OSNR dataset, (ii) using the OSNR data near the SD-FEC limit (OSNR=3.4 dB), and (iii) using training data with an approximately 3-dB lower OSNR value than the test data. Condition (i) means that a different NN receiver trained with a different condition was used for each OSNR

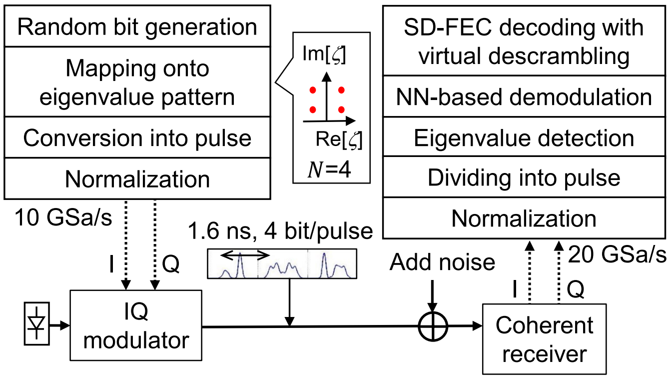

Fig. 8. Simulation model.

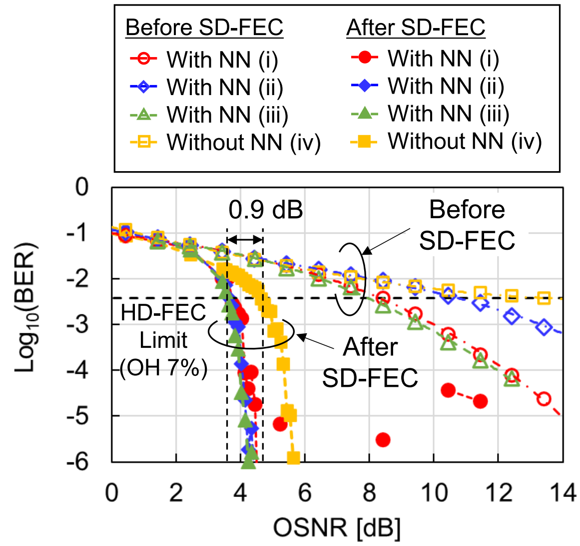

Fig. 9. BER curves before and after SD-FEC with/without NN-based demodulator in the simulation (B-to-B).

signal. For condition (ii), the same NN receiver was used for all OSNR signals, however, the target OSNR of 3.4 dB could be found through the analysis of condition (i). As well as condition (i), for condition (iii), we used a different NN receiver trained with a different condition was used for each OSNR signal. For example, a NN trained with data of OSNR $=$ 3.4 dB was used for a test of OSNR $=$ 6.4 dB, while a NN trained with data of OSNR $=$ 13.4 dB was used for a test of OSNR $=$ 16.4 dB. For comparison, condition (iv) without the NN demodulator was prepared as follows: The L-value was calculated assuming that the distribution of the eigenvalue pattern can be approximated as a Gaussian distribution following the 1-dimensional (1D) projection of the eigenvalue pattern using Fisher's linear discriminant analysis [49]. In this case, the optimum weights of the 1-D projection were determined using the training data. Consequently, the parameters of the Gaussian distribution following the 1-D projection for each bit, such as the mean and variance, were estimated using the training data. The L-value for the BER test was calculated using the optimized weights and the estimated 1-D distribution parameters. The BER before SD-FEC can be directly estimated from the L-value.

For SD-FEC, a simulation was performed using a random bit sequence, assuming the use of scramblers and descramblers at the transmitter and receiver, respectively [50]. We used the DVB-S2 low-density parity check code (LDPC) [51] with an overhead (OH) of 20%. The number of decoding iterations for SD-FEC was optimized in the range from 1 to 10. In this simulation, we used the L-values of the 210,000 bits data for the test, however, the decoding was performed changing the combination and the position of the bits (i.e., the L-values) randomly [50]. By using this method, we can say that we have evaluated many hypothetical cases of bit interleaving. The estimated spectral efficiencies were 0.125 and 0.104 b/s/Hz with and without the FEC OH, respectively, assuming the use of a 20-GHz bandwidth at the receiver.

Fig. 9 shows the BER curves before and after SD-FEC. It is clear that the BER was improved by SD-FEC under training conditions (i)–(iii). However, residual errors were observed when OSNR was within the range of 8–12 dB for condition (i). In contrast, clear waterfall curves were obtained without any residual errors for conditions (ii) and (iii). An error-free operation was achieved for OSNR values greater than 3.7 dB. In addition, the OSNR gain that was required to achieve an error-free operation was determined to be 0.9 dB when compared with the case without the NN (condition (iv)). Note that in this paper we considered an error-free operation when the BER after SD-FEC was below 3.8×10^{-3}, assuming the use of HD-FEC with an OH of 7% after performing SD-FEC. Moreover, we assumed that the errors at an input of the HD-FEC decoder were spread by using an interleaver. In the case without SD-FEC (before SD-FEC, with NN (ii)), a BER below the FEC limit for HD-FEC was achieved when the OSNR was greater than 10.8 dB. An OSNR gain for an error-free operation owing to SD-FEC was estimated at 7.1 dB for condition (ii).

Fig. 10 shows the histogram of the eigenvalue pattern distribution after the 1D projection of the eigenvalue pattern under condition (iv) (OSNR$=$ 3.4 dB). It is observed that the distribution does not match Gaussian fitting curves. The distributions for $b_1 = 0$, $b_2 = 0$, $b_3 = 1$, and $b_4 = 1$ were asymmetric. Moreover, the distributions of $b_3 = 0$ and $b_4 = 0$ exhibited steeper peaks than the Gaussian distribution. Therefore, large KL divergences D_{KL} between the actual distribution and Gaussian fitting were observed. These distribution mismatches resulted in an OSNR penalty under condition (iv), as shown in Fig. 9.

The system performance is better evaluated by BER, however, as long as we are dealing with SD-FEC, it is also essential to evaluate it by the decoder input information, which is the soft information that corresponds well to SD-FEC decoder output performance, not pre-FEC BER [52]. Therefore, we evaluated the Q-factor (Q_{soft}) that was calculated from the soft information (asymmetric information (ASI)) [53]. Fig. 11 shows Q_{soft} and the number of decoding iterations for SD-FEC. Q_{soft} and the number of iterations were unstable when residual errors were prevalent for condition (i). This implies an unsuccessful L-value computation because the NN model was not properly trained. Fig. 12 shows the histogram of the NN output for $b_1(\zeta^{(1)})$ and condition (i). When the NN was trained using the OSNR data of (a) 3.4 dB,

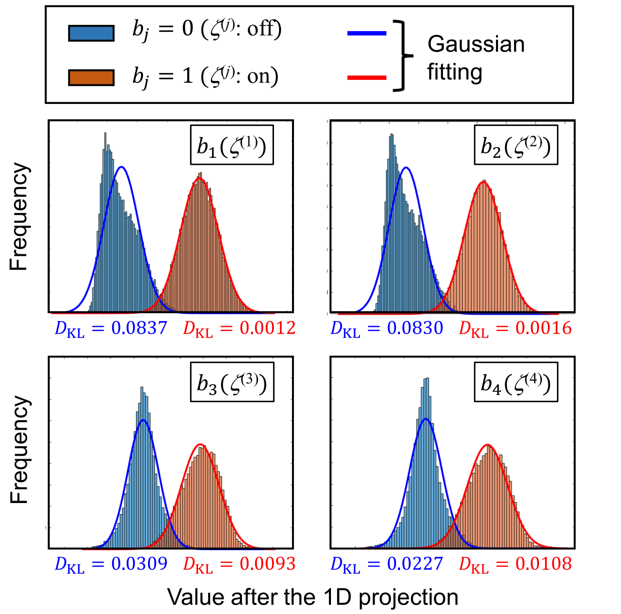

Fig. 10. Histogram of the detected eigenvalue pattern after the 1D projection (condition (iv)). The lines are the Gaussian fitting curves.

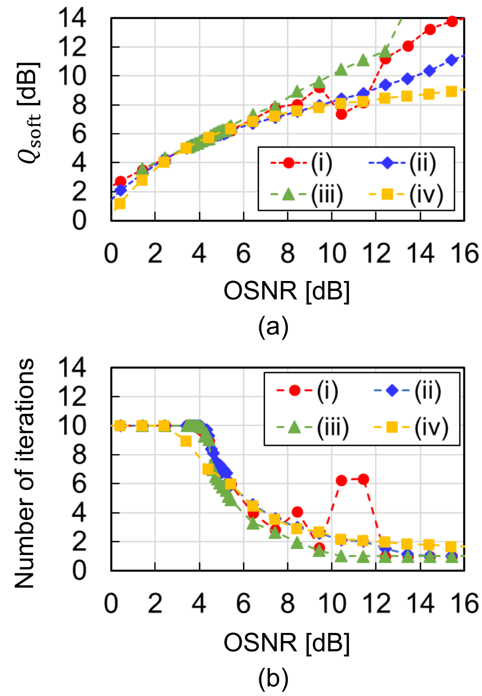

Fig. 11. (a) Q_{soft} calculated from the soft information and (b) number of iterations for SD-FEC in the simulation (B-to-B).

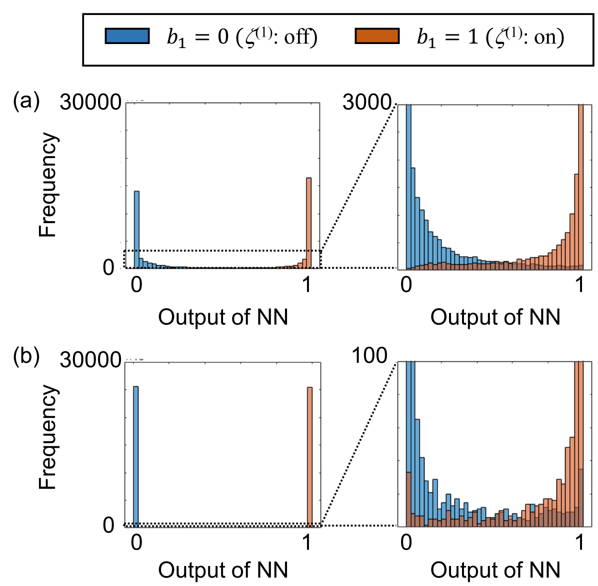

Fig. 12. Histogram of the NN output for $b_1(\zeta^{(1)})$. The NN was trained with data of (a) 3.4 dB and (b) 11.4 dB.

the frequency of the NN output smoothly changed between 0 and 1 because the data far from the ideal signal points were sufficiently included. However, when training the NN using data with a high OSNR such as (b) 11.4 dB, where the data far from the ideal signal points were not sufficiently included, processing noisy data and outliers was difficult, as shown in Fig. 12(b). By contrast, under conditions (ii) and (iii), the stable characteristics of Q_{soft} and iterations were obtained, as shown in Fig. 11. Condition (iii) exhibited a better Q_{soft} characteristic than condition (ii). However, condition (ii) has an advantage in terms of the general performance in practical systems because the NN does not require retraining for various types of OSNR test data under condition (ii). Therefore, condition (ii) was employed in subsequent experiments.

B. Transmission Experiments

To investigate the applicability of an NN-based demodulator and SD-FEC to the eigenvalue transmission, we performed a transmission experiment. Fig. 13 depicts the experimental setup with an offline NN-based receiver. For eigenvalue modulation, the same eigenvalue subsets and initial parameters used in the simulations were considered. An eigenvalue-modulated signal was generated using an offline digital signal processor (DSP) employing the same process as what was used in the simulation. An arbitrary waveform generator (AWG) and an IQ modulator were used to generate the optical signals. Subsequently, the optical signals were launched into a transmission loop comprising a 50-km NZ-DSF, EDFA, optical bandpass filter (OBPF), variable optical attenuator, and acoustic optical modulator (AOM) switch, after signal amplification and trimming were performed using an EDFA and an AOM switch. The NZ-DSF parameters were a dispersion parameter of $D = 4.4$ ps/nm/km, a dispersion slope of $S = 0.046$ ps/nm^2/km, a nonlinear coefficient of $\gamma = 2.1$ W^{-1}/km, and a fiber loss of 0.2 dB/km. The input power was set to -3.0 dBm, which was the adjusted average power of 62,250 pulses considering the loop loss, connection loss, and ASE noise from EDFA. Specifically, the input power was optimized so that the eigenvalue constellation after 2000 km transmission was the clearest by monitoring the detected eigenvalue. The ASE noise source before the receiver was used when measuring the BER curves. The OSNR was estimated by using an optical spectrum analyzer (OSA) after noise loading.

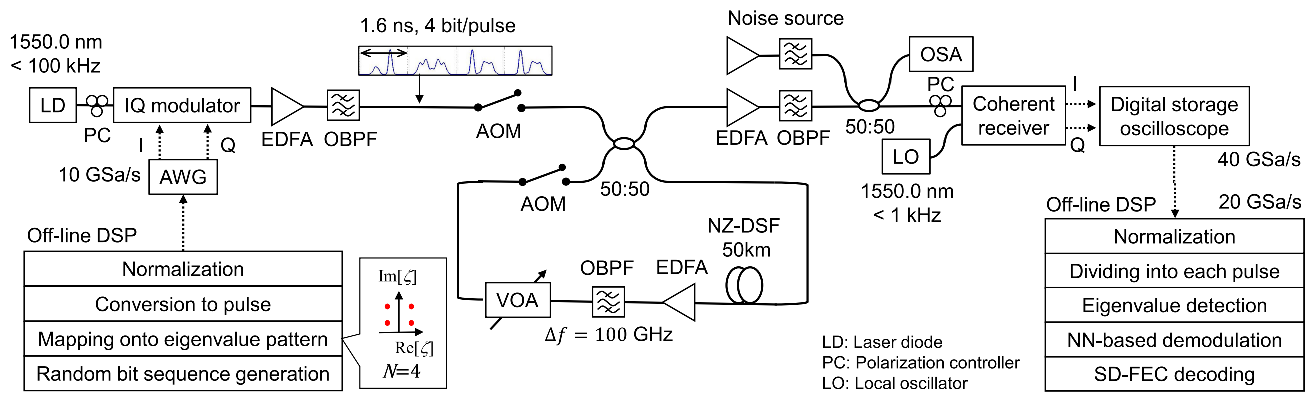

Fig. 13. Experimental setup.

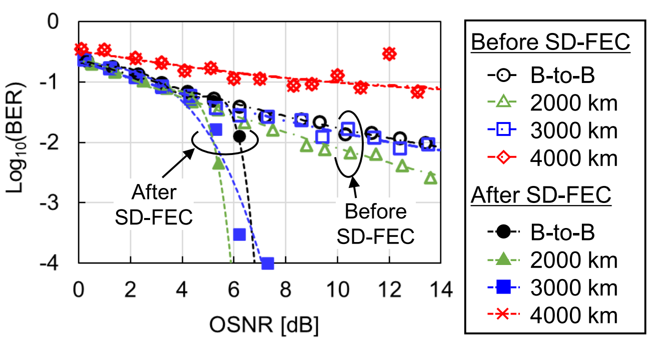

Fig. 14. BER curves after the transmission in the experiment.

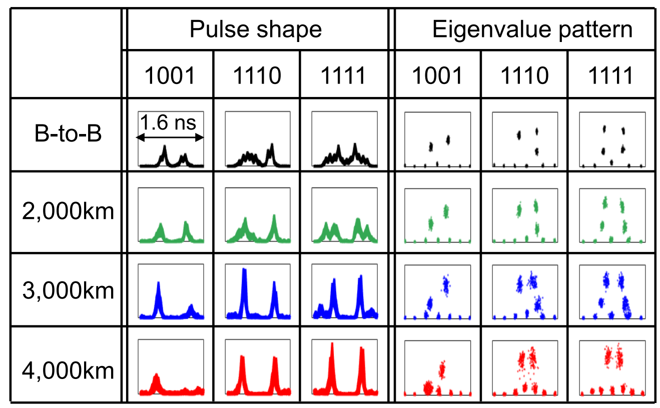

Fig. 15. Received pulses and eigenvalue patterns for representative patterns obtained in the experiment.

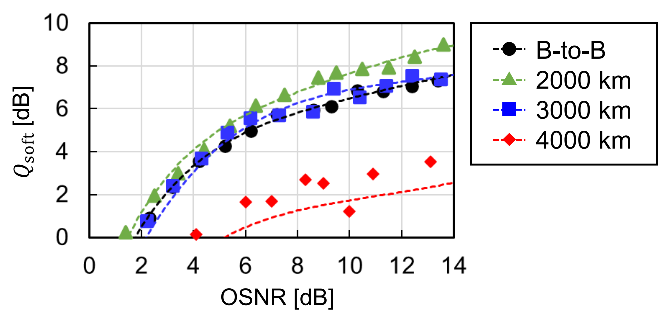

Fig. 16. Q_{soft} after the transmission in the experiment.

At the receiver end, the signals underwent analog-digital conversion using a digital storage oscilloscope (DSO). Digital signal processing for demodulation was performed offline at a rate of 20 GSa/s. The NN configuration, demodulation parameters, and SD-FEC parameters were kept identical to those used in the simulations. We employed condition (ii) for the NN training. The NN was trained with data obtained in the experiments. A total of 62,250 received pulses were divided into separate sequences of 10,000 and 52,250 pulses for training and BER test, respectively, as well as the simulation. For training condition (ii) in the experiment, we adopted the training data with the nearest OSNR to 3.4 dB, i.e., 3.2 dB for B-to-B, 3.4 dB for 2000 km, 3.2 dB for 3000 km, 3.2 dB for 4000 km, respectively. If training condition (ii) is applied to a practical system, it is considered that pseudo-random noise generated in the digital domain is added to the received signal to generate training data up to the threshold predicted in the simulation.

Fig. 14 shows the BER curves before and after SD-FEC. The BER was improved by SD-FEC and no residual errors were observed for OSNR values greater than 8 dB, even for transmission distances of 2000 and 3000 km. Fig. 15 depicts received signal pulses and eigenvalue patterns detected by NFT for representative patterns. The pulse and eigenvalue patterns were plotted using 100 representative samples with OSNR values of approximately 16 dB. After a 4000-km transmission, the eigenvalue patterns were significantly degraded owing to the inter-symbol interference and accumulated noise, which resulted in demodulation failure. Fig. 16 shows the Q_{soft} characteristics. The trend of the BER characteristics after SD-FEC corresponds well with that of Q_{soft}. From the above results, we demonstrate that SD-FEC is applicable to the eigenvalue-modulated signal with the multilabel NN-based demodulator, even in transmission experiments.

V. Transmission of 4096-Ary Eigenvalue-Modulated Signal

To extend the transmission distance, we experimentally investigated the applicability of the multilabel NN receivers and SD-FEC to a 4096-ary eigenvalue-modulated signal with 12 eigenvalues. The configuration of the experimental setup was identical to that shown in Fig. 13. However, several conditions for the eigenvalue-modulated signal, transmitter, and NZ-DSF were different from those described in the

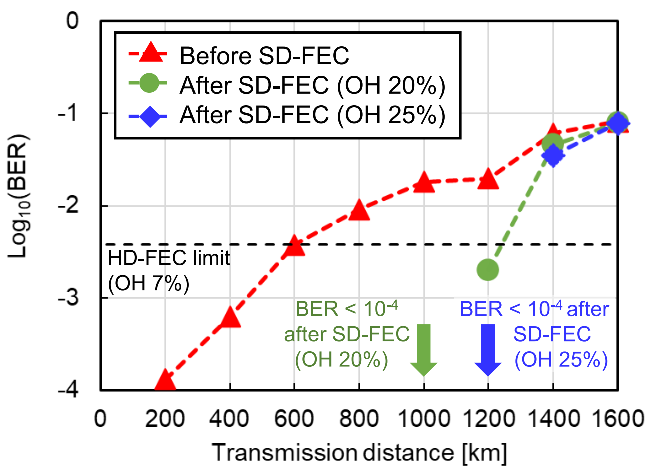

Fig. 17. BER before and after SD-FEC with varying transmission distances of the 4096-ary eigenvalue-modulated signal.

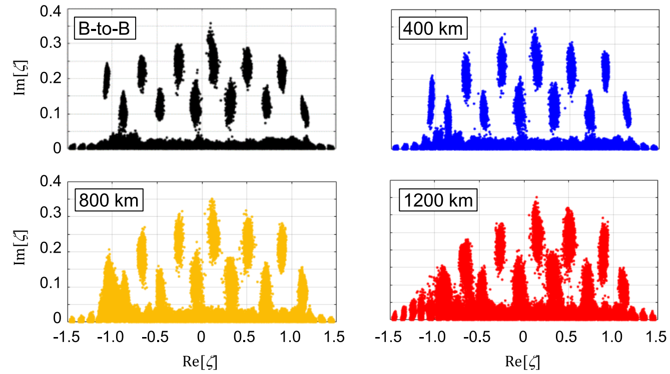

Fig. 18. Eigenvalue patterns detected from the 4096-ary eigenvalue-modulated signal after the transmission.

previous section. The on–off encoding of the 12 triangular lattice-shaped eigenvalues described in Sec. II was employed. For the eigenvalue-modulated signal, we prepared a pulse sequence by shuffling the 4096 pulses randomly. The base time t_0 and normalized window size were set to 16.7 ps and 64, respectively. The pulse duration was set to 1.07 ns, and the bit rate with an OH of the SD-FEC was 11.25 Gb/s. The eigenvalue-modulated optical signal was generated using an AWG operated at 120 GSa/s and an IQ modulator.

For the transmission loop, we employed a 40-km NZ-DSF with a low dispersion parameter of $D = 0.7$ ps/nm/km. The other NZ-DSF parameters were as follows: a dispersion slope of $S = 0.063$ ps/nm^2/km, a nonlinear coefficient of $\gamma = 2.4$ W^{-1}/km, and a fiber loss of 0.21 dB/km.

At the receiver end, the digital IQ signals were obtained using a coherent receiver and a DSO. The DSP for demodulation was performed offline at 60 GSa/s, which corresponds to 64 samples per pulse. The bit-wise L-value was computed using the multilabel NN-based demodulator, as described in Sec. III. The number of input and output units of the NN was 128 and 12, respectively. Two hidden layers with 512 hidden units and a rectified linear unit (ReLU) activation function were employed in the NN. The number of training and test samples was 16,384 and 65,536, respectively. The NN was trained using the Adam optimizer [48] and condition (ii). For SD-FEC, we used the DVB-S2 LDPC [51] with an OH of 20% or 25%. The estimated spectral efficiencies were 0.188 and 0.153 b/s/Hz with and without the FEC OH (20%), respectively, assuming the use of a 60-GHz bandwidth at the receiver.

Fig. 17 shows the BER values that were experimentally obtained before and after SD-FEC with varying transmission distances. Fig. 18 shows the detected eigenvalue patterns, which include all 4096 patterns after transmission. The BER and eigenvalue patterns gradually degraded owing to the effects of dispersion, noise, and bandwidth limitations of the transmitter as the transmission distance increased. In particular, the eigenvalues around the edges (Re[ζ] $=$ ± 1.1) were distorted because these eigenvalue components correspond to high-frequency components in the frequency domain and the edges of the pulse in the time domain in the case of the combination of the coefficients b provided above. Assuming that HD-FEC with an OH of 7% was used for decoding, the achievable transmission distance was approximately 600 km. However, owing to the multilabel NN and SD-FEC, error-free operation was achieved even after transmission for 1200 km. Moreover, using SD-FEC with an OH of 25%, the BER after SD-FEC was improved even after the 1200 km transmission. We expect that a transmission distance greater than 1200 km can be achieved by suppressing the eigenvalue distortions around the edges. e.g., using a highly dense eigenvalue pattern in the real-axis direction.

VI. Conclusion

In this study, we describe the theory of optical eigenvalue transmission including the design of a multilevel eigenvalue-modulated signal and the effects of impairments such as noise and chromatic dispersion. Consequently, we propose a multilabel NN-based demodulator to compute the L-value from the eigenvalue input at the receiver for SD-FEC decoding. This study investigates the applicability of SD-FEC for multilevel eigenvalue-modulated signals using both numerical simulations and experiments.

First, we described the theory of eigenvalue modulation, the design of the eigenvalue-modulated signal, and the concept of the proposed method. We explained that NN-based demodulators have an advantage in eigenvalue-transmission systems because the deviations in the eigenvalues are not i.i.d. with the circular Gaussian process even when white Gaussian noise is added to time-domain eigenvalue-modulated signals. By performing simulations, we demonstrated the validity of the proposed method for 16-ary eigenvalue-modulated signals, and showed a clear waterfall BER curve after SD-FEC. In addition, the combination of the NN demodulator and SD-FEC outperformed the case involving the use of a Gaussian approximation without NN in terms of BER after SD-FEC. Subsequently, we experimentally demonstrated an error-free transmission over a 3000-km optical fiber. Finally, we experimentally demonstrated the applicability of the proposed method to other multilevel signals, i.e., a 4096-ary eigenvalue-modulated signal using 12 eigenvalues. The transmission distance of the 4096-ary eigenvalue-modulated signal could be extended from 600 km to 1200 km using the NN demodulator and SD-FEC.

When multieigenvalue is used in the nonlinear spectrum modulation and b-modulation, these distributions are different from Gaussian distribution. Therefore, we believe that the combination of the NN and SD-FEC is effective for these modulation schemes as well. The proposed scheme is expected to be applied to more advanced NFT-based transmission such as further multilevel eigenvalue modulation, b-modulation, and joint modulation.

ACKNOWLEDGMENT

The authors would like to thank Takumi Motomura for useful discussions.

REFERENCES

[1] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fiber communications," *Nature*, vol. 411, pp. 1027–1030, Jun. 2001.

[2] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *IEEE/OSA J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 15, 2008.

[3] A. Hasegawa and T. Nyu, "Eigenvalue communication," *IEEE/OSA J. Lightw. Technol.*, vol. 11, no. 3, pp. 395–399, Mar. 1993.

[4] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform—Part I: Mathematical tools," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014.

[5] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform—Part II: Numerical methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, Jul. 2014.

[6] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform—Part III: Spectrum modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.

[7] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. Philadelphia, PA, USA: SIAM, 1981.

[8] A. Maruta, "Brief analysis of the impact of Hasegawa-Nyu's work," *Optik*, vol. 287, Sep. 2023, Art. no. 171036.

[9] S. K. Turitsyn et al., "Nonlinear Fourier transform for optical data processing and transmission: Advances and perspectives," *OSA Optica*, vol. 4, no. 3, pp. 307–322, Mar. 2017.

[10] H. Terauchi and A. Maruta, "Eigenvalue modulated optical transmission system digital coherence technology," in *Proc. 10th Conf. Lasers Electro-Opt. Pacific Rim, 18th Opto-Electron. Commun. Conf./Internaltional Conf. Photon. Switching*, Kyoto, Japan, Jul. 2013, pp. 1–3, Paper WR2-5.

[11] Y. Matsuda, H. Terauchi, and A. Maruta, "Design of eigenvalue-multiplexed multi-level modulation optical transmission system," in *Proc. OptoElectron. Commun. Conf. Austral. Conf. Opt. Fibre Technol.*, Melbourne, VIC, Australia, Jul. 2014, pp. 1016–1018, Paper Th3C.4.

[12] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue communication," *IEEE/OSA J. Lightw. Technol.*, vol. 34, no. 13, pp. 3110–3117, Jul. 1, 2016.

[13] K. Mishina, H. Takeuchi, T. Kodama, Y. Yoshida, D. Hisano, and A. Maruta, "4096-ary eigenvalue modulation using 12 triangular-lattice-shaped eigenvalues," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Washington, DC, USA, Jun. 2021, pp. 1–3, Paper M3I.1.

[14] V. Aref, H. Bülow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Valencia, Spain, Sep. 2015, pp. 1–3, Paper Tu.1.1.2.

[15] H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," in *Proc. ECOC 42nd Eur. Conf. Opt. Commun.*, Sep. 2016, pp. 412–414, Paper Tu.3.E.2.

[16] S. Wahls, "Generation of time-limited signals in the nonlinear Fourier domain via b-modulation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Gothenburg, Sweden, Sep. 2017, pp. 1–3, Paper W.3.C.6.

[17] T. Gui, T. H. Chan, C. Lu, A. P. T. Lau, and P. A. Wai, "Alternative decoding methods for optical communications based on nonlinear Fourier transform," *IEEE/OSA J. Lightw. Technol.*, vol. 35, no. 9, pp. 1542–1550, May 1, 2017.

[18] S. T. Le, V. Aref, and H. Buelow, "Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit," *Nature Photon.*, vol. 11, no. 9, pp. 570–576, Sep. 2017.

[19] F. Da Ros et al., "Dual-polarization NFDM transmission with continuous and discrete spectral modulation," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 10, pp. 2335–2343, May 15, 2019.

[20] P. Koster and S. Wahls, "Fast and reliable detection of significant solitons in signals with large time-bandwidth products," *IEEE/Optica J. Lightw. Technol.*, vol. 41, no. 20, pp. 6586–6598, Oct. 15, 2023.

[21] X. Chen, X. Fang, F. Yang, and F. Zhang, "10.83 Tb/s over 800 km nonlinear frequency division multiplexing WDM transmission," *IEEE/Optica J. Lightw. Technol.*, vol. 40, no. 16, pp. 5385–5394, Aug. 2022.

[22] R. T. Jones, S. Gaiarin, M. P. Yankov, and D. Zibar, "Time-domain neural network receiver for nonlinear frequency division multiplexed systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 12, pp. 1079–1082, Jun. 15, 2018.

[23] K. Mishina, S. Sato, Y. Yoshida, D. Hisano, and A. Maruta, "Eigenvalue-domain neural network demodulator for an eigenvalue-modulated signal," *IEEE/OSA J. Lightw. Technol.*, vol. 39, no. 13, pp. 4307–4317, Jul. 1, 2021.

[24] K. Mishina, T. Maeda, D. Hisano, Y. Yoshida, and A. Maruta, "Combining IST-based CFO compensation and neural network-based demodulation for eigenvalue-modulated signal," *IEEE/OSA J. Lightw. Technol.*, vol. 39, no. 23, pp. 7370–7382, Dec. 1, 2021.

[25] X. Yangzhang, V. Aref, S. T. Le, H. Buelow, D. Lavery, and P. Bayvel, "Dual-polarization non-linear frequency-division multiplexed transmission with b-modulation," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 6, pp. 1570–1578, Mar. 15, 2019.

[26] Y. Wu et al., "Robust neural network receiver for multiple-eigenvalue-modulated nonlinear frequency division multiplexing system," *Opt. Exp.*, vol. 28, no. 12, pp. 18304–18316, Jun. 8, 2020.

[27] O. Kotlyar, M. Pankratova, M. Kamalian-Kopae, A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, "Combining nonlinear Fourier transform and neural network-based processing in optical communications," *Opt. Lett.*, vol. 45, no. 13, pp. 3462–3465, Jul. 1, 2020.

[28] S. Wahls, "Second order statistics of the scattering vector defined the D-T nonlinear Fourier transform," in *Proc. 11th Int. ITG Conf. Syst. Commun. Coding*, Hamburg, Germany, Jun. 2017, pp. 1–6.

[29] F. J. García-Gómez, and V. Aref, "Statistics of the nonlinear discrete spectrum of a noisy pulse," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 14, pp. 3563–3570, Jul. 15, 2019.

[30] K. Sugihara, K. Ishii, K. Dohi, K. Kubo, T. Sugihara, and W. Matsumoto, "Scalable SD-FEC for efficient next-generation optical networks," in *Proc. ECOC 42nd Eur. Conf. Opt. Commun.*, Germany, Sep. 2016, pp. 1–3, Paper W.2.C.3.

[31] K. Onohara et al., "Soft-decision-based forward error correction for 100 Gb/s transport systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1258–1267, Sep. 2010.

[32] K. Mishina, T. Yoshida, D. Hisano, and A. Maruta, "Application of SD-FEC to optical eigenvalue-modulated signals using a neural network-based demodulator," in *Proc. 49th Eur. Conf. Opt. Commun.*, vol. 2023, Nov. 2023, pp. 274–277, Paper Th.B.5.4.

[33] R. Harada, T. Yoshida, D. Hisano, A. Maruta, and K. Mishina, "1200-km transmission of 4096-ary eigenvalue-modulated signal using a neural network-based demodulator and SD-FEC," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Diego, CA, USA, 2024, pp. 1–3, Paper M4K.7.

[34] S. Le, J. E. Prilepsky, P. Rosa, J. D. Ania-Castañón, and S. K. Turitsyn, "Nonlinear inverse synthesis for optical links with distributed Raman amplification," *IEEE/OSA J. Lightw. Technol.*, vol. 34, no. 8, pp. 1778–1786, Apr. 15, 2016.

[35] V. Bajaj, S. Chimmalgi, V. Aref, and S. Wahls, "Exact NFDM transmission in the presence of fiber-loss," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 11, pp. 3051–3058, Jun. 1, 2020.

[36] V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of wave in nonlinear media," *Sov. J. Exp. Theory Phys. (JETP)*, vol. 34, no. 1, pp. 62–69, Jan. 1972.

[37] G. Agrawal, *Nonlinear Fiber Optics*, 6th ed., Cambridge, MA, USA: Academic, 2019.

[38] A. Hasegawa and Y. Kodama, "Guiding-center soliton in optical fibers," *Opt. Lett.*, vol. 15, no. 24, p. 1443, Dec. 15, 1990.

[39] T. Kodama, K. Mishina, Y. Yoshida, D. Hisano, and A. Maruta, "Transmission of hyper-multi-level eigenvalue-modulated signal using arbitrary optical multi-eigenvalue," *Opt. Commun.*, vol. 546, Nov. 2023, Art. no. 129748.

[40] K. Nishino, S. Nakao, T. Maeda, K. Mishina, and A. Maruta, "Influence of bandwidth limitation in optical eigenvalue modulation," in *Proc. Int. Conf. Photon. Switching Comput. (PSC)*, Mantova, Italy, Sep. 2023, pp. 1–3, Paper Fr2B.5.

[41] A. Span, V. Aref, H. Bülow, and S. T. Brink, "Time-bandwidth product perspective for nonlinear Fourier transform-based multi-eigenvalue soliton transmission," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5544–5557, Aug. 2019.

[42] W. Q. Zhang et al., "Correlated eigenvalues of multi-soliton optical communications," *Sci. Rep.*, vol. 9, no. 1, pp. 1–17, Apr. 2019.

[43] H. Takeuchi, K. Mishina, Y. Terashi, D. Hisano, Y. Yoshida, and A. Maruta, "Eigenvalue-domain neural network receiver for 4096-ary eigenvalue-modulated signal," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, San Diego, CA, USA, Mar. 2022, pp. 1–3, Paper Th3F.3.

[44] O. Shental and J. Hoydis, "'Machine LLRning': Learning to softly demodulate," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, San Diego, CA, USA, Dec. 2019, pp. 1–7, Paper Th3F.3.

[45] L. Szczecinski and A. Alvarado, *Bit-Interleaved Coded Modulation: Fundamentals, Analysis and Design*. Chichester, U.K.: Wiley, 2015.

[46] T. Koike-Akino, Y. Wang, D. S. Millar, K. Kojima, and K. Parsons, "Neural turbo equalization: Deep learning for fiber-optic nonlinearity compensation," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 11, pp. 3059–3066, Jun. 1, 2020.

[47] Y. Onodera, D. Hisano, K. Maruta, and Y. Nakayama, "First demonstration of 512-color shift keying signal demodulation using neural equalization for optical camera communication," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, San Diego, CA, USA, Mar. 2023, pp. 1–3, Paper Th3H.7.

[48] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015, pp. 1–15.

[49] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[50] T. Yoshida, M. Karlsson, and E. Agrell, "Efficient offline evaluation of FEC codes based on captured data with probabilistic shaping," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, San Diego, CA, USA, Mar. 2018, pp. 1–3, Paper M4E.5.

[51] *Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding, and Modulation. Systems for Broadcasting, Interactive Services and News Gathering and Other Broadband Satellite Applications—Part 1: DVB-S2*," Standard ETSI, EN 302 307–1 V.1.4.1, 2014.

[52] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *IEEE/OSA J. Lightw. Technol.*, vol. 33, no. 20, pp. 4338–4352, Oct. 15, 2015.

[53] T. Yoshida, "Soft Q-factor in optical communication systems," *IEICE Commun. Exp.*, vol. 11, no. 12, pp. 871–876, Dec. 1, 2022.

Ken Mishina (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical, electronic, and information engineering from Osaka University, Osaka, Japan, in 2005, 2007, and 2012, respectively. In 2007, he joined Shimadzu Corporation, Kyoto, Japan. Since 2018, he has been an Associate Professor with the Department of Information and Communications Technology, Division of Electrical, Electronic, and Information Engineering, Graduate School of Engineering, Osaka University. His research interests include optical fiber communications, systems, all-optical signal processing, and photovoltaics. He is a member of the IEEE Photonics Society and IEICE, Japan.

Ryotaro Harada received the B.E. degree in electrical and electronic engineering from Kobe City College of Technology, Kobe, Japan, in 2023. He is currently pursuing the M.E. degree with Osaka University. He is a member of IEICE, Japan.

Tsuyoshi Yoshida (Member, IEEE) received the B.Sc. degree in electrical and electronic engineering and the M.Sc. degree in communications and computer engineering from Kyoto University, Kyoto, Japan, in 2005 and 2007, respectively, and the Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Osaka, Japan, in 2022. Since 2007, he has been with Mitsubishi Electric Corporation, Kamakura, Japan, where he has been engaged in the research and development of high-speed optical transmission, coding and modulation, digital signal processing, and algorithms for classic and quantum hybrid systems. From 2016 to 2017, he visited the Fiber-Optic Communications Research Center, Chalmers University of Technology, Gothenburg, Sweden. He is a member of Optica and IEICE, Japan.

Daisuke Hisano (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical, electronic, and information engineering from Osaka University, Osaka, Japan, in 2012, 2014, and 2018, respectively. In 2014, he joined the NTT Access Network Service Systems Laboratories, Yokosuka, Japan. Since 2018, he has been an Assistant Professor with the Department of Information and Communications Technology, Division of Electrical, Electronic, and Information Engineering, Graduate School of Engineering, Osaka University. His research interests include optical-wireless converged networks, optical communications, and all-optical signal processing. He is a member of the IEICE, Japan.

Akihiro Maruta (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University in 1988, 1990, and 1993, respectively. Then, he joined the Department of Communications Engineering, Osaka University, in 1993, where he has been a Professor with the Department of Information and Communication Technologies since 2016. His current research interests include optical fiber communication systems and all-optical signal processing. He is a member of the IEEE Photonics Society and Optica and a Senior Member of IEICE, Japan.