Eigenvalue-Domain Neural Network Demodulator for Eigenvalue-Modulated Signal

Ken Mishina, *Member, IEEE*, Shingo Sato, Yuki Yoshida, *Member, IEEE*, Daisuke Hisano, *Member, IEEE*, and Akihiro Maruta, *Member, IEEE, Member, OSA*

***Abstract*—Optical eigenvalue communication is a promising technique for overcoming the Kerr nonlinear limit in optical communication systems. The optical eigenvalue associated with the nonlinear Schrödinger equation remains invariant during fiber-based nonlinear dispersive transmission. However, practical applications involving use of such systems are limited by the occurrence of fiber loss and amplified noise that induce eigenvalue distortion. Thus, several time-domain neural-network-based approaches have been proposed and demonstrated to enhance receiver sensitivity toward eigenvalue-modulated signals. However, despite the substantial improvement in power margin realized using time-domain neural-network-based demodulators compared to their conventional counterparts, these devices require rigorous training for each transmission distance owing to changes in time-domain pulses during transmission. This paper presents a method for demodulation of eigenvalue-modulated signals using an eigenvalue-domain neural network and demonstrates its utility through simulation and experimental results. Simulation results obtained in this study reveal that the proposed demodulator demonstrates superior generalization performance compared to its time-domain counterpart with regard to the transmission distance. Moreover, experimental results demonstrate successful demodulation over distances from zero to 3000 km without training for each distance.**

***Index Terms*—Optical fiber communication, optical solitons, fiber nonlinear optics, machine learning, artificial neural networks.**

I. Introduction

OPTICAL eigenvalue communication [1] based on inverse scattering transform (IST) [2] has been extensively investigated as a means of overcoming the Kerr nonlinear limit [3]–[23]. IST has recently become well-known as the nonlinear Fourier transform (NFT). Although optical waveforms and frequency spectra change during signal propagation in nonlinear dispersive fibers, eigenvalues pertaining to the eigenvalue equation associated with the nonlinear Schrödinger equation (NLSE) remain invariant. In 1993, Hasegawa and Nyu proposed the first eigenvalue-based modulation scheme [1]. Recent advancements in digital coherent technologies have facilitated successful realization of eigenvalue-based communication [6], [7]. Over the past several years, many studies have been performed concerning use of advanced modulation and transmission techniques based on eigenvalue communication. Several eigenvalue-based modulation schemes, such as on–off encoding of multiple eigenvalues [8], [9] and phase-shift keying modulation of the spectral amplitude for multi-soliton pulses [10], have been proposed to enhance transmission capacity. In addition, several experimental demonstrations of eigenvalue transmission have been reported. Transmission of three eigenvalues located on the imaginary axis at 1.5-Gb/s has been demonstrated as the first experimental realization of on–off encoding of multiple eigenvalues [11]. Moreover, using 16APSK on the spectral amplitude, a 24-Gb/s signal transmission over a 1000-km distance has been experimentally demonstrated [12]. Recently, a 220-Gb/s transmission via dual-polarization b-modulation has been demonstrated with a spectral efficiency of 4-b/s/Hz [13].

To facilitate reliable demodulation of eigenvalue-modulated signals, eigenvalues and their nonlinear spectrum or the scattering coefficient b are detected from received time-domain signals using IST. Subsequently, a hard decision (HD) with a linear threshold on the complex eigenvalue plane or b-plane is performed using the typical method. However, a major concern here is that upon optimization of the decision threshold, deviations in the eigenvalue and coefficient b due to amplified noise and fiber loss are not independent and identically distributed (i.i.d.) with the circular Gaussian process, particularly in cases involving multi-eigenvalue systems [14], [15]. To facilitate improvement in the received power margin, use of an additional decision process based on the Euclidian minimum distance (MD) has been demonstrated [16]. In addition, machine learning-based demodulation methods for eigenvalue modulation, such as classification [17]–[20] and equalization [21], [22], have been investigated in recent years. Demodulation methods based on time-domain (TD) artificial neural networks (ANNs) for quadrature phase shift keying (QPSK) modulation of spectral amplitudes with two eigenvalues and

Manuscript received July 15, 2020; revised January 13, 2021; accepted April 15, 2021. Date of publication April 21, 2021; date of current version July 2, 2021. This work was supported in part by the Grants-in-Aid for Scientific Research under the Japan Society for the Promotion of Science (JSPS) under Grant JP19H02140. *(Corresponding author: Ken Mishina.)*

Ken Mishina, Shingo Sato, 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; s_sato@pn.comm.eng.osaka-u.ac.jp; hisano@comm.eng.osaka-u.ac.jp; maruta@comm.eng.osaka-u.ac.jp).

Yuki Yoshida is with the National Institute of Information and Communications Technology (NICT), Koganei, Tokyo 184-8795, Japan (e-mail: yuki@nict.go.jp).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2021.3074744.

Digital Object Identifier 10.1109/JLT.2021.3074744

on–off encoding of four-eigenvalue modulation have been proposed [17], [18]. TD-ANN-based demodulation is expected to address above-mentioned concerns, because simulation results obtained in extant studies have demonstrated the ANN method to outperform IST- and MD-based methods in terms of the bit error rate (BER) performance [17]. Moreover, in their extant study, the authors have experimentally demonstrated successful TD-ANN-based demodulation of multi-eigenvalue-modulated signals with on–off encoding [19]. However, although use of TD-ANN results in realization of a large power margin of 11 dB when compared against conventional IST+HD-based method, the former approach requires model training for each transmission distance considered owing to changes in TD pulses that occur during transmission. Furthermore, there have been no studies that discuss the generalization performance of ANN-based demodulators for eigenvalue-modulated signal.

This paper presents a scheme for demodulating eigenvalue-modulated signals based on an eigenvalue-domain (ED) ANN to enhance generalization performance. Feasibility of the proposed method has been demonstrated via numerical simulations and experimental data. As an extension of their extant research [20], the authors, in this paper, have described the demodulation method in detail along with the numerical investigation of basic characteristics pertaining to the proposed ED-ANN demodulator. The proposed method involves combined use of IST and ANN to retain benefits of invariant eigenvalues during transmission without the need for performing model training when considering different transmission distances. Simulation results obtained in this study demonstrate that the proposed ED-ANN demodulator offers higher generalization performance compared to its TD-ANN counterpart in terms of the transmission distance. Then, we apply the proposed ED-ANN demodulator to a multi-eigenvalue-modulated signal using on-off encoding of four eigenvalues, which has a complicated eigenvalue distribution and can take better advantage of the ANN demodulator. Furthermore, experiments performed in this study demonstrate successful demodulation of an eigenvalue-modulated signal with a bit rate of 2.5-Gb/s and BER $< 3.8 \times 10^{-3}$ using the proposed ED-ANN demodulator. Moreover, the trained ANN demodulator is valid for transmission distances from zero to 3000 km.

This remainder of this paper is organized as follows. Section II describes the eigenvalue-modulation method involving on–off encoding as well as compares the proposed demodulation against corresponding conventional approaches. Section III discusses results of numerical simulations performed in this study. Section IV discusses transmission experiments performed in this study to demonstrate the feasibility of the proposed method. Finally, Section V lists major conclusions drawn from this study.

II. Eigenvalue Modulation and Demodulation

A. Eigenvalue Modulation Based on On-Off Encoding

The complex envelope amplitude of electric fields propagating through optical fibers that exhibit anomalous dispersion and nonlinearity can be approximately expressed using the following NLSE [24]

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

where Z, T, and $u(Z,T)$ denote normalized values of the propagation distance, time moving with average group velocity, and complex envelope amplitude of the electric field, respectively. The eigenvalue equation associated with the above (1) can be expressed as

$$i\frac{\partial \phi_1}{\partial T} + u\phi_2 = \zeta\phi_1, -i\frac{\partial \phi_2}{\partial T} - u^*\phi_1 = \zeta\phi_2, \tag{2}$$

where ζ and $\phi_l(Z,T)$ $(l = 1, 2)$ denote the complex eigenvalue and eigen functions, respectively. So long as u satisfies (1), ζ remains invariant with distance Z.

Fig. 1 depicts a block diagram of the eigenvalue modulation–demodulation scheme based on on–off encoding. The on-off encoding is based on a one-to-one mapping between an N-bit input and the subsets of eigenvalues. If the value of the bit in the j-th position is 1 (or 0), the j-th eigenvalue is included (or excluded). Although the eigenvalue set is arbitrary, spectral efficiency and noise tolerance depend on it. The said scheme can be explained using an example with $N = 4$. First, the four information bits are encoded to a complex eigenvalue pattern; for example, consider bit sequences "1111" and "0110" to represent eigenvalue patterns of "$\zeta = \{0.25 + 0.5i, -0.25 + 0.5i, 0.25 + 0.25i, -0.25 + 0.25i\}$" and "$\zeta = \{-0.25 + 0.5i, 0.25 + 0.25i\}$," respectively. Subsequently, this encoded eigenvalue pattern can be mapped onto a pulse using IST. In this study, pulses were generated using IST with an N-soliton solution to avoid ill-conditioning problems [23]. The composed solution $u(Z,T)$ can be expressed as

$$u(Z,T) = \begin{cases} -2\sum_{n=1}^{N} \gamma_n^* \chi_{n2}^* \exp\left(-i\zeta_n^* T\right) (T < 0) \\ -2\sum_{n=1}^{N} \sigma_n \psi_{n1} \exp\left(-i\zeta_n T\right) (T \geq 0), \end{cases} \tag{3}$$

where $\gamma_n(Z)$ and $\sigma_n(Z)$ denote scattering parameters corresponding to each eigenvalue ζ_n; $\chi_{n2}(Z,T)$ and $\psi_{n1}(Z,T)$ represent eigen functions $\phi_l(Z,T)$ in (2) that satisfy boundary conditions corresponding to $T \to \infty$ and $T \to -\infty$, respectively. Subsequently, the complex IQ waveform supplied by a digital–analog converter (DAC) is fed into an IQ modulator for generating an optical signal. That is, N information bits are converted to optical pulses containing an on–off pattern of the N-tuple of eigenvalues. This optical eigenvalue-modulated signal can now be transmitted over an optical fiber transmission line.

At the receiver end, the transmitted optical signal is received by a generic coherent receiver that divides the signal into IQ components. These coherent-receiver outputs are converted to their digital form using an analog–digital converter (ADC). Subsequently, the digital signal waveform is divided into individual pulses, which in turn, are converted to corresponding eigenvalue patterns by solving the eigenvalue problem (2) in the frequency domain [7]. The detected eigenvalue pattern is decoded into an information bit sequence with a decision.

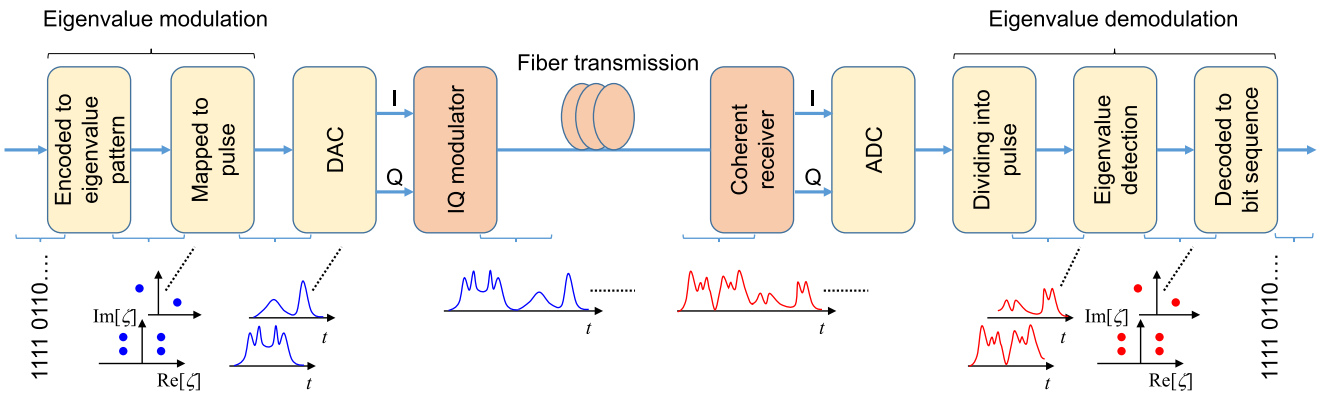

Fig. 1. Block diagram of eigenvalue modulation–demodulation scheme based on on-off encoding of four eigenvalues ($N = 4$).

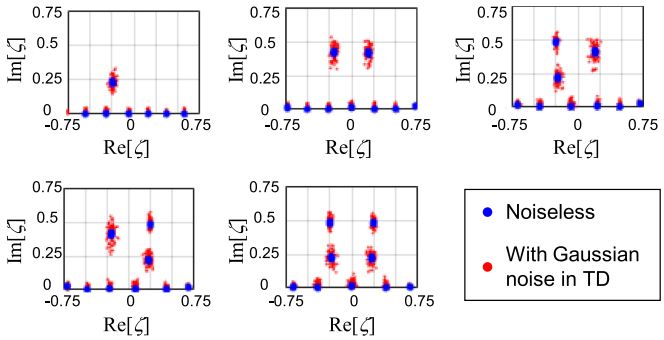

Fig. 2. Eigenvalue deviations of representative eigenvalue patterns ($N = 4$).

For a given decision, it is important to set the classification boundary accurately. During demodulation, eigenvalues are detected from a pulse via IST—a nonlinear transform. Even if we assume that deviation of each TD-pulse sampling point caused by amplified spontaneous emission (ASE) noise is i.i.d. with the said Gaussian process, deviations of detected eigenvalues are not i.i.d. with the circular Gaussian process. This is because soliton pulses consist of multi-sampling points in TD; this induces interaction of ASE noise when the TD-pulses are converted to eigenvalues using IST. This is especially true of multi-eigenvalue systems, wherein eigenvalue deviations bear a correlation with those of other eigenvalues [14]. Fig. 2 depicts eigenvalue deviations of representative eigenvalue patterns ($N = 4$). In the figure, white Gaussian noise is added on TD pulse so that the optical signal-to-noise ratio (OSNR) is 10.2 dB. As can be observed, eigenvalue deviation depends on eigenvalue position. Moreover, deviation changes may also be caused depending on the pattern while the eigenvalue remains at the same position. This phenomenon indicates that the linear threshold on the complex eigenvalue plane does not yield an optimum solution. To address this concern, this paper presents an ED-ANN-based demodulation method with an optimum nonlinear boundary plane to classify eigenvalue patterns.

B. ANN-Based Demodulation

Fig. 3 depicts three different demodulation schemes. In the conventional IST+HD scheme depicted in Fig. 3(a), the received pulse is converted to an eigenvalue pattern using IST. The detected eigenvalue pattern is decoded into an information bit sequence via appropriate setting of linear thresholds on the complex eigenvalue plane.

In the authors extant study [19], time-series data pertaining to the received pulse were provided as inputs to ANN and decoded bit-sequence information was obtained as a direct output, as illustrated in Fig. 3(b). In particular, IQ components of each pulse-shape sample point are considered ANN input, and probability factors of bit sequences corresponding to detected eigenvalue patterns are considered ANN output. The configuration of the ANN-based demodulator for four eigenvalues ($N = 4$) is depicted in Fig. 3 (b). For a 32-sample/pulse sampling rate, the number of input elements required equals 64 with 32 IQ components. The number of output elements equals 16 in accordance with the number of eigenvalue patterns—i.e., $2^4 = 16$. Because TD pulses are susceptible to change during transmission, the TD-ANN demodulator must be trained for each transmission distance.

Figure 3(c) depicts the proposed demodulation method based on ED-ANN. In this method, ANN inputs correspond to eigenvalue data converted from TD pulse using IST. The real and imaginary parts of each eigenvalue are input to ANN, which in turn, outputs the probability parameter for the bit sequence corresponding to detected eigenvalue patterns. Figure 3(c) depicts the configuration of the proposed ED-ANN demodulator for four eigenvalues ($N = 4$). For a sampling rate of 32 sample/pulse, the number of converted eigenvalues, including continuous spectrum, equals 32. There exist 64 input elements comprising 32 real and imaginary parts each of eigenvalues. The number of output elements equals 16 corresponding to the number of eigenvalue patterns (i.e., $2^4 = 16$) similar to TD-ANN.

III. Numerical Simulations

This section describes numerical simulations applying the proposed ED-ANN demodulator to eigenvalue modulation. First, we show demodulation results for 1- and 2-soliton transmissions using TD- and ED-ANN demodulators to compare their generalization performance on transmission distance.

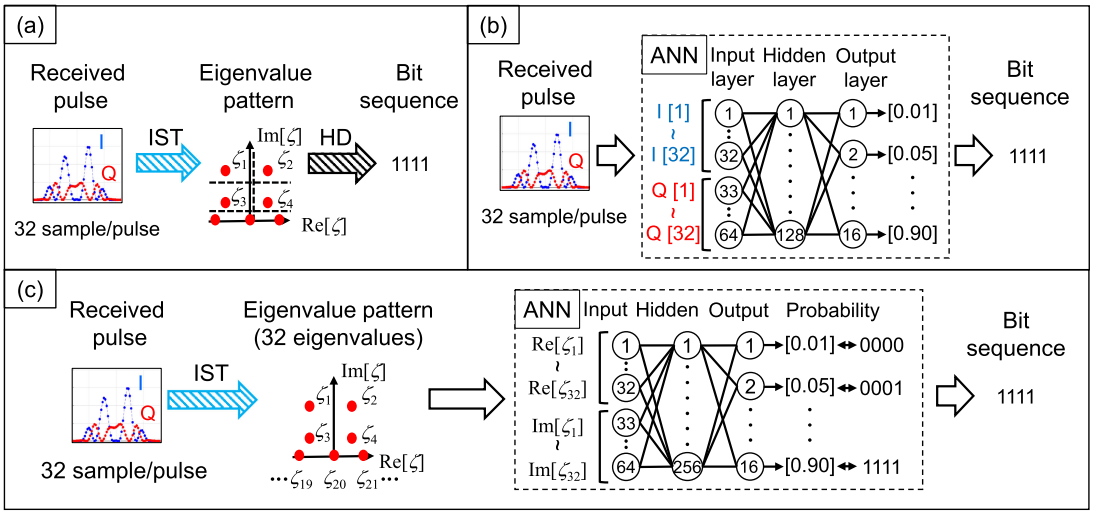

Fig. 3. Comparison between three demodulation schemes—(a) Inverse scattering transform (IST) + hard decision (HD) (conventional method); (b) time-domain artificial neural network (ANN) (previous method), and (c) eigenvalue-domain ANN (proposed method).

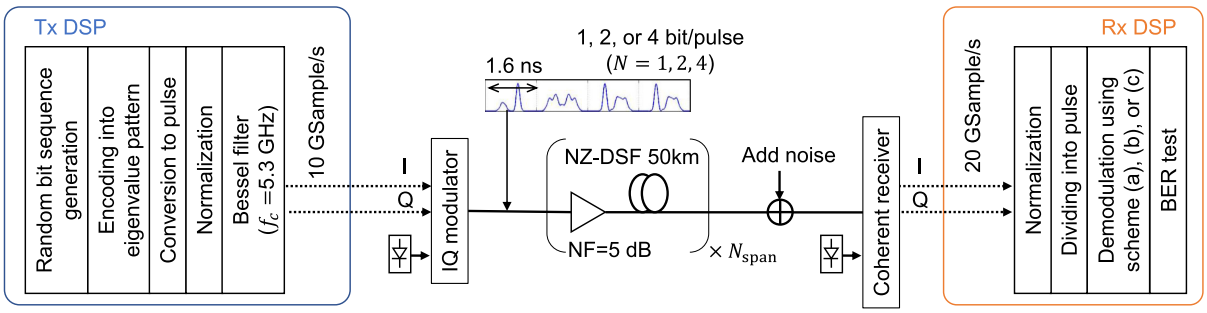

Fig. 4. Simulation model used in this study.

Then, demodulation results for on-off encoding of four eigenvalues are demonstrated to reveal the applicability and advantages of ED-ANN compared with the conventional IST+HD method.

A. Simulation Model

Fig. 4 illustrates the simulation model used in this study. For eigenvalue modulation, we prepared three modulated signals using on-off encodings of 1, 2, and 4 eigenvalues ($N = 1, 2, 4$). The mapping rules between bit sequences and eigenvalue subsets are summarized in Table I. A parameter value of $Z = -8$ and a normalized time window width of 32 ($-16 \leq T \leq 16$) were considered in (3) for the eigenvalue-modulated signal at the transmitter. Modulation was performed at 10 Gsample/s with a pulse duration of 1.6 ns. The bit rates were 0.625, 1.25, and 2.5 Gb/s for $N = 1, 2$, and 4, respectively. To facilitate transmission simulation, the nonlinear Schrödinger equation was numerically solved using the split-step Fourier method [25]. The transmission loop comprised a 50-km non-zero dispersion-shifted fiber (NZ-DSF) along with an Er-doped fiber amplifier (EDFA) with noise figure (NF) of 5 dB. Parameter values concerning NZ-DSF included dispersion $D = 4.4$ ps/nm/km, dispersion slope $S = 0.046$ ps/nm^2/km, nonlinear coefficient $\gamma = 2.1$ W^{-1}/km, and 0.2-dB/km fiber loss. The input signal powers were set to -13.7 dBm ($N = 1$), -9.0 dBm ($N = 2$), and -6.3 dBm ($N = 4$), which corresponded to the calculated ideal average power for eigenvalue modulation. At the receiver end, signals were demodulated using (a) conventional IST+HD, (b) TD-ANN, or (c) ED-ANN demodulator at 20-Gsample/s sampling rate. White Gaussian noise was added to eigenvalue-modulated signal in TD when evaluating BER. Ideal frequency characteristics were assumed for both the transmitter and receiver with the exception of Bessel filtering at the transmitter end.

The ANN configuration and parameters for the demodulation are described in Section II. We used a three-layer perceptron configuration and a rectified linear unit activation function. The number of hidden units of TD-ANN was set to 128 [18]. We set the number of hidden units of ED-ANN to 256, which is discussed in Section IV. The number of output units of TD- and ED-ANNs were 2, 4, and 16 for $N = 1, 2, 4$, respectively. The soft max and cross-entropy error functions was considered output and loss functions, respectively. We prepared two random

TABLE I
MAPPING RULES BETWEEN BIT SEQUENCES AND EIGENVALUE SUBSETS

(a) $N = 1$

Bit sequence	Eigenvalue subset
0	None
1	$0.25i$

(b) $N = 2$

Bit sequence	Eigenvalue subset
00	None
10	$0.5i$
01	$0.25i$
11	$0.5i,\ 0.25i$

(c) $N = 4$

Bit sequence	Eigenvalue subset
0000	None
1000	$0.25 + 0.5i$
0100	$-0.25 + 0.5i$
1100	$0.25 + 0.5i,\ -0.25 + 0.5i$
0010	$0.25 + 0.25i$
1010	$0.25 + 0.5i,\ 0.25 + 0.25i$
0110	$-0.25 + 0.5i,\ 0.25 + 0.25i$
1110	$0.25 + 0.5i,\ -0.25 + 0.5i,\ 0.25 + 0.25i$
0001	$-0.25 + 0.25i$
1001	$0.25 + 0.5i,\ -0.25 + 0.25i$
0101	$-0.25 + 0.5i,\ -0.25 + 0.25i$
1101	$0.25 + 0.5i,\ -0.25 + 0.5i,\ -0.25 + 0.25i$
0011	$0.25 + 0.25i,\ -0.25 + 0.25i$
1011	$0.25 + 0.5i,\ 0.25 + 0.25i,\ -0.25 + 0.25i$
0111	$-0.25 + 0.5i,\ 0.25 + 0.25i,\ -0.25 + 0.25i$
1111	$0.25 + 0.5i,\ -0.25 + 0.5i,$ $0.25 + 0.25i,\ -0.25 + 0.25i$

pulse sequences of 10 000 pulses for the training and 52 250 pulses for the validation and BER tests, respectively. The ANN was trained using the Adam optimizer [26] available within the MATLAB R2020a Deep Learning Toolbox. The said Adam optimizer was configured using step size $\alpha = 0.001$, exponential decay rates $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and coefficient $\epsilon = 10^{-8}$. Training data were uniformly extracted from available datasets with OSNR values in the -10–10 dB, -5–15 dB, and -2–17 dB ranges for $N = 1, 2, 4$, respectively. To avoid over-fitting, the training was terminated when the validation result (obtained once every 50 epochs) ceased to improve [27]. For the conventional demodulation method IST+HD, the linear threshold on complex eigenvalue plane was decided by maximum a posteriori (MAP) estimation assuming that each eigenvalue deviation is i.i.d. with the Gaussian process. The training data sets for the MAP estimation were extracted in the same way as the ANN demodulator.

B. Simulation Results

1) One and Two Eigenvalues ($N = 1, 2$): The on-off encoding of one eigenvalue ($N = 1$) corresponds to the on-off keying of the fundamental soliton. In this case, the TD-ANN does not require training for each transmission distance because the pulse shape of the fundamental soliton is invariant during the transmission, as shown in Fig. 5(a). On the other hand, the modulated signal using on-off encoding of two eigenvalue ($N = 2$) includes 2-soliton. The pulse shape of the 2-soliton changes during transmission, as shown in Fig. 5(a). Therefore, TD-ANN needs training for each transmission distance for multi-soliton transmission systems.

Fig. 6 shows the BER curves of one and two eigenvalue-modulated signals ($N = 1, 2$) using TD- and ED-ANN demodulators trained for a fixed transmission distance of 3000 km. For $N = 1$, TD-ANN can demodulate the eigenvalue-modulated signals of different transmission distances. However, TD-ANN is not valid for the eigenvalue-modulated signal when $N = 2$, including multi-soliton systems. On the other hand, ED-ANN can deal with multi-soliton systems in the case of $N = 2$ regardless of the transmission distance because the eigenvalue patterns are invariant during transmission, as shown in Fig. 5(b).

2) 4 Eigenvalues ($N = 4$): Fig 7(a) depicts BER curves for back-to-back (B-to-B) operation. As can be observed, the proposed ED-ANN demodulator outperforms the conventional IST+HD method by a power margin of 1.0 dB at a BER of 3.8×10^{-3} by assuming HD forward-error correction (FEC) limit. This is better compared to the conventional IST+HD method because the Euclidean distance in the 64-dimensional space of a 64-input ANN exceeds that in the two-dimensional complex eigenvalue plane considered in the conventional IST+HD method. The TD-ANN demodulator demonstrates superior B-to-B BER performance because the effect of noise in discrete ED exceeds that in TD. Fig. 7(b) and (c) depict the BER curves before and after transmission when using the TD- and ED-ANN-based demodulators trained for each transmission distances. After 5000-km transmission, occurrence of a large power penalty due to inter-symbol interference can be observed when employing TD-ANN. However, using the ED-ANN demodulator demonstrated successful demodulation with negligible power penalty when transmitting signals over 5000 km.

Fig. 7(d) and (e) depict BER curves obtained when using TD- and ED-ANN demodulators trained for a fixed transmission distance of 3000-km. Fig. 8 depicts received pulses and eigenvalue patterns detected by IST as representative patterns. The said pulses and eigenvalue patterns are plotted using 100 representative samples with OSNR values of approximately 16 dB. As observed in the TD-ANN case, as shown in Fig. 7(d), it is difficult to demodulate signals received over different transmission distances as well as in the case for $N = 2$. In contrast, as depicted in Fig. 7(d), the ED-ANN demodulator can easily deal with received signals transmitted over 5000 km even in the case for $N = 4$ owing to conservation of eigenvalue patterns during transmission, as illustrated in Fig. 8(b).

IV. EXPERIMENTS

A. Experimental Setup

Fig. 9 depicts the experimental setup used in this study comprising an offline ANN-based receiver. For eigenvalue modulation, the same eigenvalue subsets comprising four optical eigenvalues ($N = 4$) and initial parameters described in Section III were considered. An eigenvalue-modulated signal was generated using an offline digital signal processor (DSP). For the said eigenvalue-modulated signal, 62 250 random pulses were generated using the Mersenne twister algorithm. The set

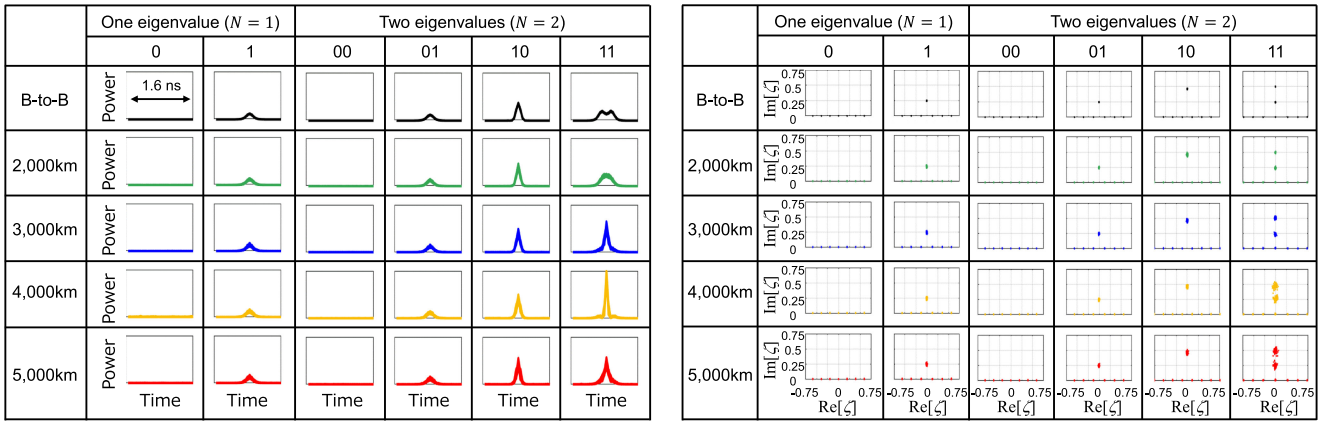

Fig. 5. (a) Received pulses and (b) detected eigenvalue patterns of one and two eigenvalue-modulated signals ($N = 1, 2$) before and after NZ-DSF transmission.

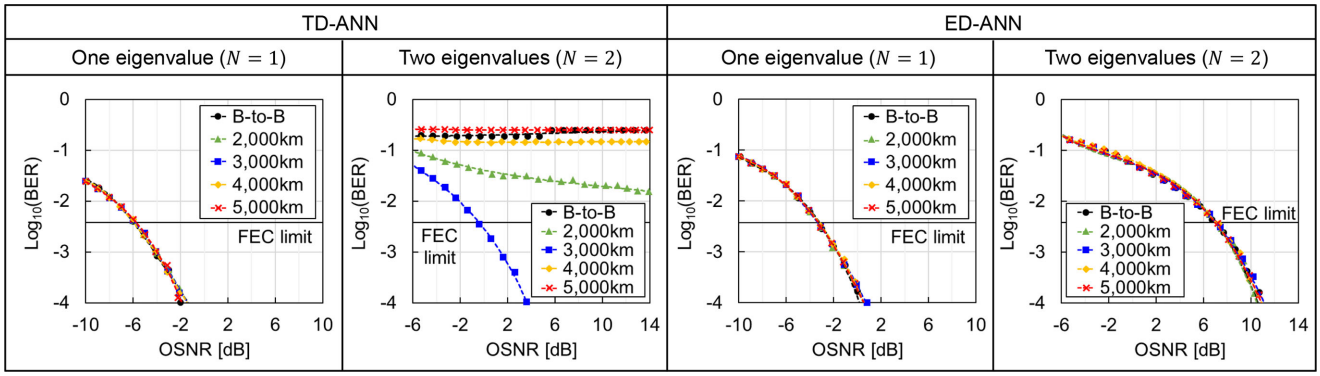

Fig. 6. BER curves of one and two eigenvalue-modulated signals ($N = 1, 2$) when using TD- and ED-ANN demodulators trained for fixed distance (3000 km).

pulse duration and resulting bit rate equaled 1.6 ns and 2.5 Gb/s, respectively. The optical signal was generated using an arbitrary waveform generator (AWG) and an IQ modulator. The said AWG operated at 10-Gsample/s sampling rate. Generated optical signals were launched into a transmission loop after signal amplification and trimming performed using an EDFA and an acoustic optical modulator (AOM) switch. The transmission loop comprised a 50-km NZ-DSF, EDFA, optical bandpass filter (OBPF), variable optical attenuator (VOA), and AOM switch. Parameter values concerning NZ-DSF were maintained identical to those considered during numerical simulations discussed in Section III. The input power was set to -3.0 dBm, which was adjusted average power of 62 250 pulses with due consideration of the loop loss, connection loss, and amplified spontaneous emission (ASE) noise from EDFA. The ASE noise source before the receiver was used when measuring BER curves.

At the receiver end, signals were analog–digital converted using a digital storage oscilloscope (DSO) operated at 40-Gsample/s sampling rate. The required digital signal processing for demodulation was performed offline at 20-Gsample/s sampling rate after downsampling acquired data. The ANN configuration and demodulation parameter values were maintained identical to those described in Section III. Training data were uniformly extracted from available data sets with OSNR values between -2 and 17 dB. We used two different random pulse sequences of 10 000 pulses for the training and 52 250 pulses for the validation and BER tests, respectively.

B. *Experimental Results*

Fig. 10 depicts BER curves obtained in this study by performing loop-transmission experiments employing different demodulation methods and training conditions. As observed, experimentally obtained BER results are worse compared to those obtained via numerical simulations under all conditions. This is because optical signals generated for use during experiments performed in this study differ from ideal signals considered during simulations owing to bandwidth limitation and nonlinear distortion by optical transmitter. Fig. 11 depicts received signal pulses and eigenvalue patterns detected by IST for representative patterns. The said pulses and eigenvalue patterns are plotted using 100 representative samples with OSNR values of approximately 16 dB. As can be realized, compared to simulation results (Fig. 8), experimental data reveal a shift in eigenvalue position. Moreover, observed trends concerning the said eigenvalue shift depend on the eigenvalue position and pattern. Fig. 10(a) depicts

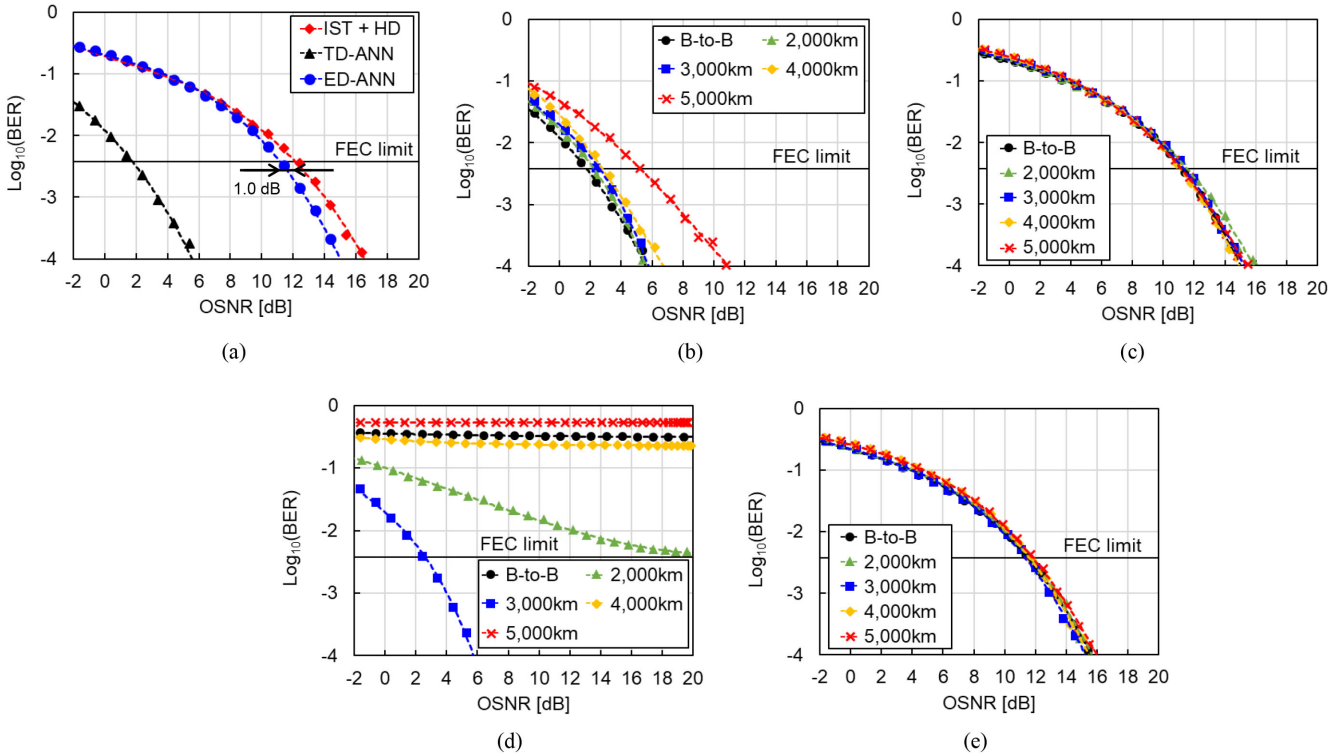

Fig. 7. BER curves by simulation for $N = 4$. (a) B-to-B; (b) TD-ANN trained for each distance; (c) ED-ANN trained for each distance; (d) TD-ANN trained for fixed distance (3000 km); (e) ED-ANN trained for fixed distance (3000 km).

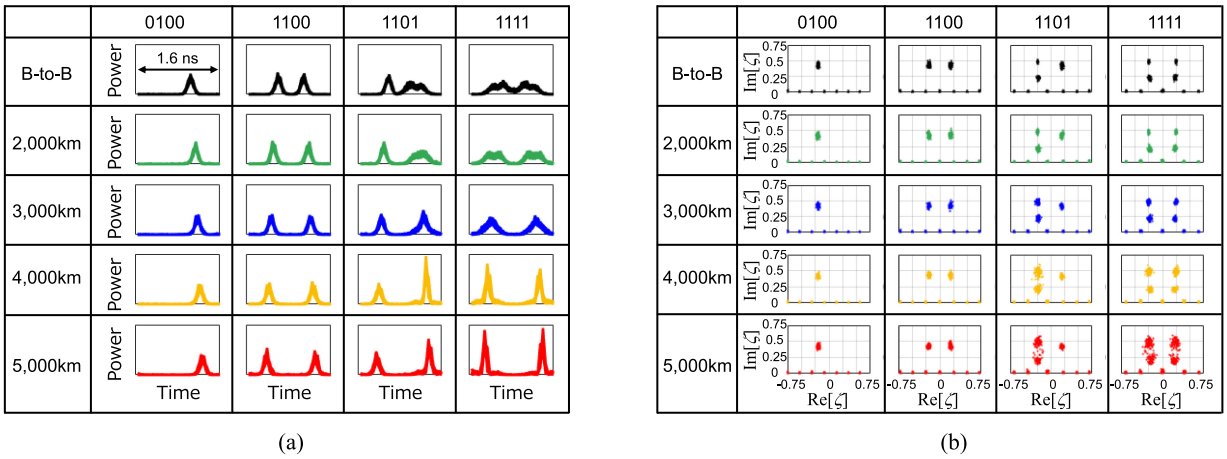

Fig. 8. (a) Representative received pulses and (b) detected eigenvalue patterns for $N = 4$ before and after NZ-DSF transmission.

BER curves for B-to-B operation, during which the ED-ANN demodulator outperforms its conventional IST+HD counterpart by a large power margin of 3.7 dB at BER of 3.8×10^{-3} assuming HD-FEC. Based on this results, it can be inferred that nonlinear eigenvalue shifts caused by optical transmitters emphasize BER improvement realized via use of ED-ANN demodulators. Furthermore, the TD-ANN demodulator achieves much better BER during B-to-B operation than the ED-ANN demodulator, as confirmed by both simulation and experimental results. Fig. 10(b) and (c) depicts BER curves before and after transmission when using the TD- and ED-ANN demodulators trained for different transmission distances. As can be observed, both TD-ANN and ED-ANN demodulators demonstrated occurrence of an power penalty at transmission distances exceeding 4000 km. Successful demodulation when transmitting over a 3000-km distance with BER $< 3.8 \times 10^{-3}$ can be confirmed for both TD- and ED-ANN demodulator cases. As illustrated in Fig. 11(b), eigenvalue patterns were conserved over the 3000-km distance. However, changes in eigenvalue position were found to exceed those observed during B-to-B operation and transmission over 2000-km distance. Compared to simulation results, experimentally obtained eigenvalue patterns demonstrated occurrence

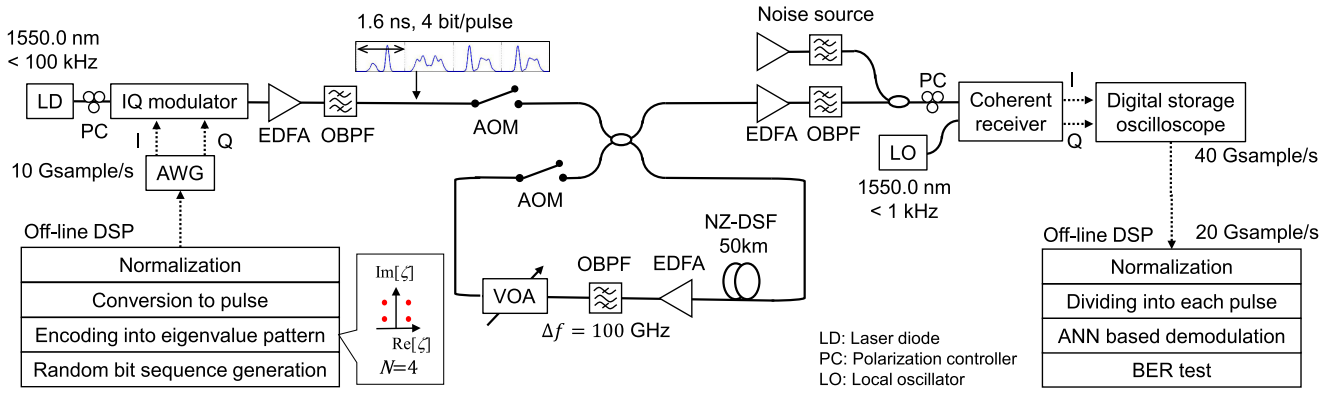

Fig. 9. Experimental setup used in this study.

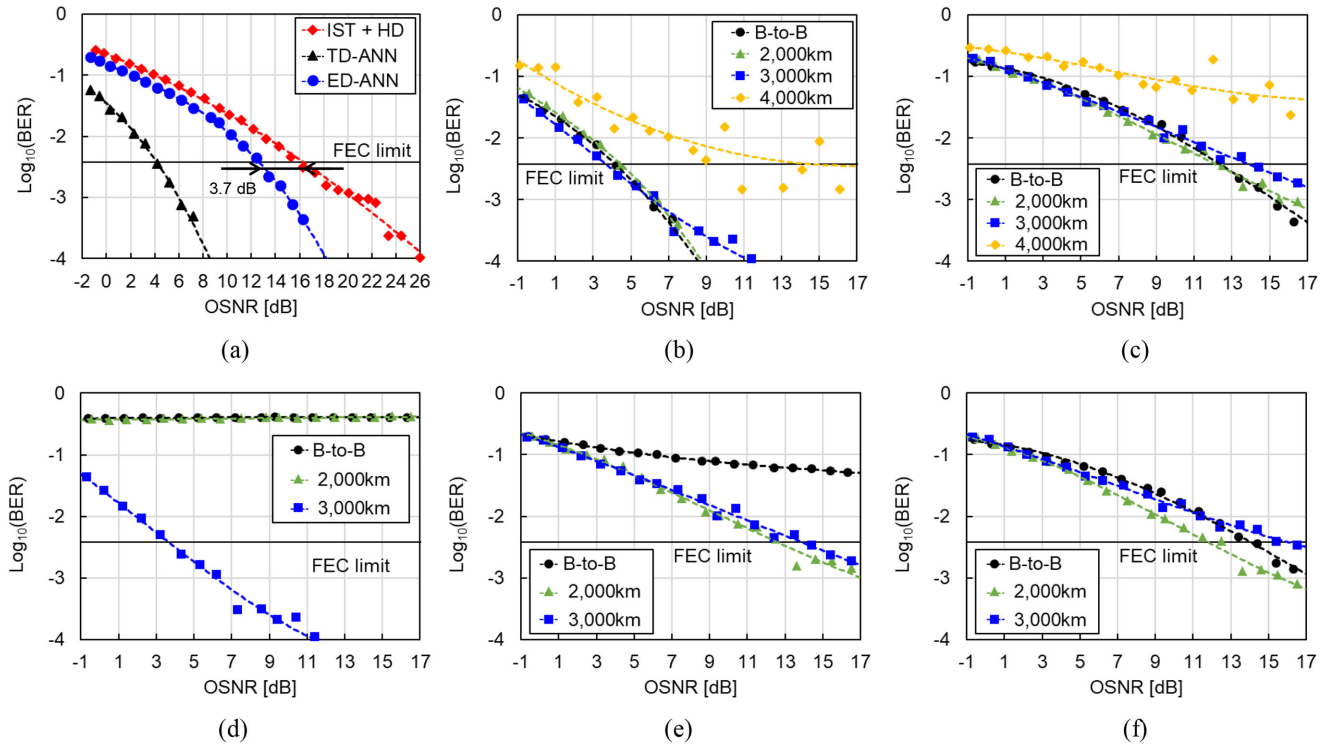

Fig. 10. BER curves obtained via loop-transmission experiments—(a) B-to-B operation; (b) TD-ANN trained for arbitrary transmission distance; (c) ED-ANN trained for arbitrary transmission distance; (d) TD-ANN trained for fixed transmission distance (3000 km); (e) ED-ANN trained for fixed transmission distance (3000 km); (f) eigenvalue-domain ANN trained with data for all 3 distances.

of severe distortion over shorter transmission distances. This could be attributed to not only the nonlinear distortion caused by the transmitter but also the effect of polarization mode dispersion and non-optimal input power supplied to the recirculating transmission loop. The optimum average input power to maintain constant eigenvalues depends on the number of loops required for signal transmission owing to accumulation of ASE noise from EDFA within recirculating transmission loops.

Fig. 10(d) and (e) depict BER curves obtained when employing trained TD- and ED-ANN for a fixed distance of 3000 km. As illustrated in Fig. 10(d), the TD-ANN demodulator struggles to demodulate the received signal for the different transmission distance owing to changes in the TD-pulse shape during signal transmission, as depicted in Fig. 11(a). In contrast, as shown in Fig. 10(e), the ED-ANN demodulator can deal with received signal transmitted over both 2000- and 3000-km distances. Simulation results demonstrated better generalization performances compared to experimental results owing to suppression of eigenvalue shifts during simulations due to generation of ideal signals sans nonlinear distortion as well as supply of optimum input power to the transmission loop. Fig. 10(f) depicts BER curves obtained using the ED-ANN demodulator trained using 10 000 data samples for each of the three distances of zero (i.e., B-to-B), 2000 km, and 3000 km. Realization of a BER

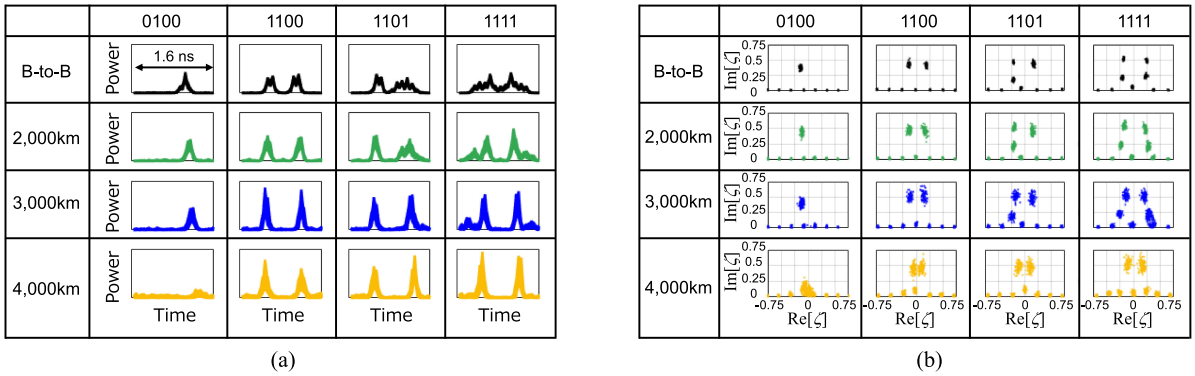

Fig. 11. (a) Representative pulses received and (b) detected eigenvalue patterns before and after NZ-DSF transmission.

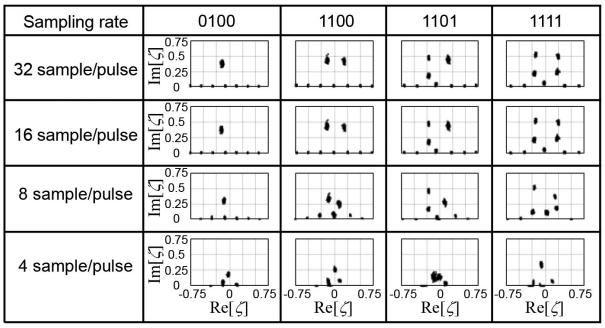

Fig. 12. Representative detected eigenvalue patterns during B-to-B operation with changing sampling rate.

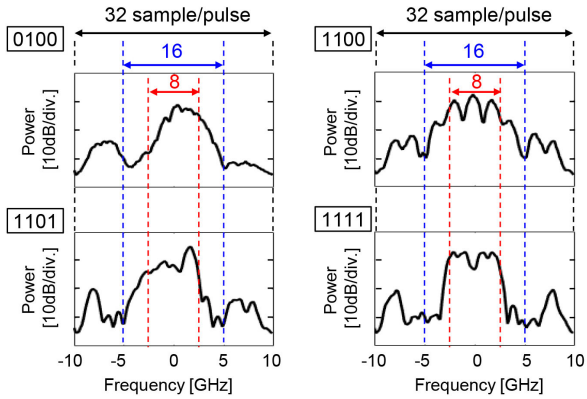

Fig. 13. Spectra of representative patterns during B-to-B operation.

result under the FEC limit for each distance between zero and 3000 km confirms the validity of trained ANN demodulators for transmission distances within the specified range.

C. Hyperparameters

Demodulation results obtained using an ANN-based demodulator depend on several hyperparameters, such as the number of input units, hidden units, and hidden layers. These hyperparameters can be discussed based on above-described experimental results. The number of input units is governed by the sampling rate. Fig. 12 depicts representative detected eigenvalue patterns during B-to-B operation with changing sampling rate. Fig. 13 shows spectra of representative patterns corresponding to those shown in Fig. 12. When the sampling rate equals 8 sample/pulse, a large shift in eigenvalue position is observed by limiting the bandwidth under the undersampling condition. In this case, the detected eigenvalue pattern changes depending on the transmission distance because the spectrum changes during transmission. This, in turn, results in degradation of the generalization performance of the ED-ANN demodulator. Therefore, a sufficiently high sampling rate of 32 sample/pulse was considered in this study by setting the number of input units to 64 to investigate the potential for generalization-performance enhancement in this study.

Fig. 14 depicts OSNR values at the FEC limit (BER$= 3.8 \times 10^{-3}$) as a function of the number of hidden units. The ANN

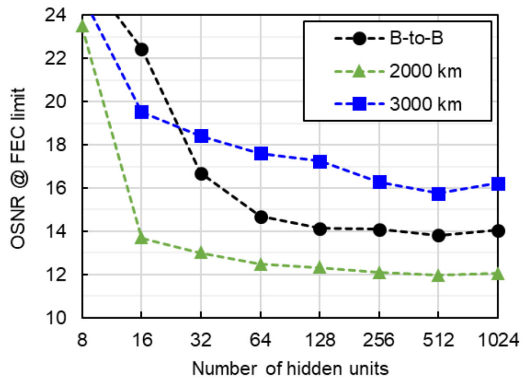

Fig. 14. OSNR values at FEC limit as a function of varying the number of hidden units.

training was performed using training data set for the three distances of zero, 2000 km, and 3000 km. The larger the number of hidden units, the more stable and lower the required OSNR was observed to be. From this result, 256 or 512 hidden units were considered sufficient in this study to cover all transmission distances of the eigenvalue-modulated signal.

Fig. 15 depicts OSNR values at the FEC limit as a function of the number of hidden layers. The number of hidden units of each

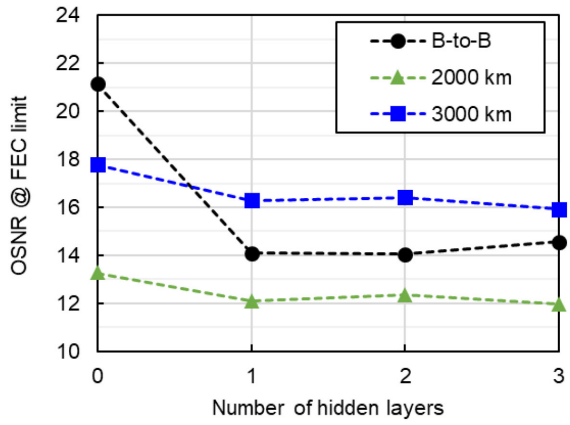

Fig. 15. OSNR values at FEC limit as a function of varying the number of hidden layers.

hidden layer was 256. No hidden layers ("0") means a linear classification of logistic regression in Fig. 15. The required OSNR was improved owing to a nonlinear classification by inserting hidden layers. However, we cannot see a clear improvement when increasing the number of hidden layers. One nonlinear hidden layer is enough to classify the eigenvalue patterns in this study.

V. Conclusion

This paper presents an eigenvalue-domain ANN-based demodulation scheme for optical eigenvalue-modulated signals meant for a large range of transmission distances. The utility of the proposed scheme has been demonstrated via both numerical simulations and experimental results. As observed, the proposed ED-ANN demodulator outperforms the conventional IST+HD demodulator by a large power margin of 3.7 dB during experiments. Furthermore, results of this study confirm realization of successful optical signal demodulation transmitted over distances between zero and 3000 km without the need for model training for each transmission distance. Thus, the findings of this study reveal the potential of ED-ANN demodulators to cover a large distance range in the point of generalization performance. To increase transmission capacity and spectral efficiency, the concept of the ED-ANN receiver is expected to be applied to further advanced eigenvalue modulation such as higher multilevel of on-off encoding, eigenvalue position encoding, and *b*-modulation in the near future.

References

[1] A. Hasegawa and T. Nyu, "Eigenvalue communication," *IEEE/OSA J. Lightw. Technol.*, vol. 11, no. 3, pp. 395–399, Mar. 1993.
[2] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. Philadelphia, PA, USA: SIAM, 1981.
[3] 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.
[4] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear fourier transform, part I: Mathematical tools," *IEEE Trans. Inform. 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. Inform. 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. Inform. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.
[7] H. Terauchi and A. Maruta, "Eigenvalue modulated optical transmission system based on digital coherent technology," in *Proc. 10th Proc. Conf. Lasers Electro-Opt. Pacific Rim, 18th Opto-Electron. Commun. Conf./Int. Conf. Photon. Switching*, Jul. 2013, Paper WR2-5.
[8] Y. Matsuda, H. Terauchi, and A. Maruta, "Design of eigenvalue multiplexed multi-level modulation optical transmission system," in *Proc. 19th OptoElectron. Commun. Conf./Australian Conf. Opt. Fibre Technol.*, Jul. 2014, pp. 1016–1018.
[9] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue communication," *IEEE/OSA J. Lightw. Technol.*, vol. 34, no. 13, pp. 3110–3117, Jul. 2016.
[10] V. Aref and H. Buelow, "Design of 2-soliton spectral phase modulated pulses over lumped amplified link," in *Proc. 42nd Eur. Conf. Opt. Commun.*, Sep. 2016, pp. 409–411.
[11] Z. Dong *et al.*, "Nonlinear frequency division multiplexed transmissions based on NFT," *IEEE Photon. Technol. Lett.*, vol. 27, no. 15, pp. 1621–1623, Aug. 2015.
[12] T. Gui, C. Lu, A. P. T. Lau, and P.-K. A. Wai, "High-order modulation on a single discrete eigenvalue for optical communications based on nonlinear fourier transform," *OSA Opt. Exp.*, vol. 25, no. 17, pp. 20286–20297, Aug. 2017.
[13] X. Yangzhang, S. T. Le, V. Aref, H. Buelow, D. Lavery, and P. Bayvel, "Experimental demonstration of dual-polarization NFDM transmission with b-modulation," *IEEE Photon. Technol. Lett.*, vol. 31, no. 11, pp. 885–888, Jun. 2019.
[14] W. Q. Zhang *et al.*, "Correlated eigen-values of multi-soliton optical communications," *Sci. Rep.*, vol. 9, no. 6399, pp. 1–17, Apr. 2019.
[15] 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. 2019.
[16] H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-Modulated spectral amplitudes," in *Proc. 42nd Eur. Conf. Opt. Commun.*, Sep. 2016, pp. 412–414.
[17] R. T. Jones, S. Gaiarin, M. P. Yankov, and D. Ziber, "Time-domain neural network receiver for nonlinear frequency division multiplexed systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 12, pp. 1079–1082, Jun. 2018.
[18] S. Yamamoto, K. Mishina, and A. Maruta, "Demodulation of optical eigenvalue modulated signal using neural network," *IEICE Commun. Exp.*, vol. 8, no. 12, pp. 507–512, Dec. 2019.
[19] K. Mishina, S. Yamamoto, T. Kodama, Y. Yoshida, D. Hisano, and A. Maruta, "Experimental demonstration of neural network based demodulation for on-off encoded eigenvalue modulation," in *Proc. 45th Eur. Conf. Opt. Commun.*, Sep. 2019, pp. 1–4, Paper W.1.B.4.
[20] K. Mishina, S. Sato, S. Yamamoto, Y. Yoshida, D. Hisano, and A. Maruta, "Demodulation of eigenvalue modulated signal based on eigenvalue-domain neural network," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2020, Paper W3D.1.
[21] Y. Wu *et al.*, "Robust neural network receiver for multiple-eigenvalue modulated nonlinear frequency division multiplexing system," *OSA Opt. Exp.*, vol. 28, no. 12, pp. 18304–18316, Jun. 2020.
[22] 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," *OSA Opt. Lett.*, vol. 45, no. 13, pp. 3462–3465, Jul. 2020.
[23] T. Kodama, T. Zuiki, K. Mishina, and A. Maruta, "Hyper multilevel modulation based on optical eigenvalue multiplexing," in *Proc. Photon. Switching Comput.*, 2018, Paper Th3C.4.
[24] A. Hasegawa and Y. Kodama, "Guiding-center soliton in optical fibers," *OSA Opt. Lett.*, vol. 15, no. 24, pp. 1443–1445, Dec. 1990.
[25] G. P. Agrawal, *Nonlinear Fiber Optics*, *6th Edition*. Cambridge, MA, USA: Academic Press, 2019.
[26] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representation*, May 2015, pp. 1–15, .
[27] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning* Cambridge, MA, USA: MIT Press, 2016.

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 Associate Professor with the Department of Information and Communication Technology, Division of Electrical, Electronic, and Information Engineering, Graduate School of Engineering, Osaka University. His research interests include optical fiber communication systems, all-optical signal processing, and photovoltaics. He is a Member of IEEE Photonics Society and IEICE, Japan.

Shingo Sato received the B.E. degree in electrical, electronic, and information engineering from Osaka University, Osaka, Japan, in 2019, where he is currently working toward the M.E. degree. He is a Student Member of IEICE, Japan.

Yuki Yoshida (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in informatics from Kyoto University, Kyoto, Japan, in 2004, 2006, and 2009, respectively. Between 2009 and 2016, he was an Assistant Professor with Osaka University, Osaka, Japan. Since 2016, he has been a Senior Researcher with Network System Research Institute, National Institute of Information and Communications Technology, Tokyo, Japan. He is currently a Visiting Associate Professor with Osaka University and the Graduate School for the Creation of New Photonics Industries (GPI), Japan. His research interests include digital signal processing for optical or wireless communications, optical or wireless access, and optical-wireless convergence. He is a Member of 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 NTT Access Network Service Systems Laboratories, Yokosuka, Japan. Since 2018, he has has been an Assistant Professor with the Department of Information and Communication Technology, Division of Electrical, Electronic, and Information Engineering, Graduate School of Engineering, Osaka University. His research interests include optical-wireless converged networks, optical communication, and all-optical signal processing. He is a Member of IEICE, Japan.

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