# Multi-Eigenvalue Demodulation Using Complex Moment-Based Eigensolver and Neural Network

Yuhei Terashi, Daisuke Hisano ⓘ, *Member, IEEE*, Ken Mishina ⓘ, *Member, IEEE*, Yuki Yoshida ⓘ, *Member, IEEE*, and Akihiro Maruta ⓘ, *Member, IEEE*

*Abstract*— Optical eigenvalues originating in optical solitons are the potential for becoming information carriers not affected by chromatic dispersion and nonlinear effects in optical fibers. They are obtained by attributing the associated eigenvalue equations deduced by solving the nonlinear Schrödinger equation with inverse scattering transform (IST) to the matrix eigenvalue problem, and maintaining constant values regardless of the transmission distance. The eigenvalue communication systems require to solve the eigenvalues in soliton-by-soliton. While effective eigenvalue solution methods have not been studied well in telecommunication systems, one of the most well-known eigenvalue solution methods is the QZ decomposition-based method. However, the QZ algorithm requires a large complexity. To reduce the complexity, a method to demodulate the optical eigenvalues using a complex-moment eigenvalue solver (CME) was investigated. CME is a parallelizable eigenvalue solver that can extract any eigenvalue. This paper proposes a novel optical eigenvalue demodulation method that combines CME and an artificial neural network (ANN) based on employing an on-off encoded discrete eigenvalue modulation scheme. The ANN is sensitive to the input order of the units; therefore, the eigenvalues must be sorted. A lightweight sorting algorithm is hence required. In addition to the proposed scheme, this study proposes partial sorting using CME and ANN. Here, 2000 km fiber-transmission experiments for an on-off-encoded four-discrete-eigenvalue were conducted. The experimental results indicated that the proposed demodulation method obtained bit error rate (BER) characteristics comparable to conventional methods by devising an extraction range of eigenvalues in CME.

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

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

## I. INTRODUCTION

NONLINEAR effects in optical fibers limit the input optical power. This limitation is termed the nonlinear Shannon limit and provides upper limit of the transmission capacity [3], [4], [5]. Optical eigenvalue transmission has been proposed to overcome this limitation [6], [7]. The optical eigenvalues are invariant in the presence of chromatic dispersion and nonlinear effects in optical fibers. Optical eigenvalues are extracted by using the associated eigenvalue equations obtained when solving the nonlinear Schrödinger equation (NLSE) by inverse scattering transform (IST) [8]. Digital coherent technologies enable the extraction of optical eigenvalues because they can indirectly refer to the amplitude and phase of the signals. Thus, they have encouraged innovation in eigenvalue modulation techniques [1], [2], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19].

In the digital domain, the optical eigenvalues are obtained by attributing the associated eigenvalue equations to a matrix eigenvalue problem and then solving the eigenvalue problem. A stable algorithm solves the eigenvalue problem based on the QZ decomposition technique; hereinafter this is called "QZ." After conducting QZ, the optical eigenvalues are sorted. The sorted eigenvalues are input into the artificial neural network (ANN) for bit decision. A method has been proposed and its feasibility has been evaluated [14], [15]. However, the previous method had the following issues.

- QZ requires a large time complexity, $\mathcal{O}(n^3)$, where the matrix size is $n \times n$.
- The basic ANN requires sorting the eigenvalues before inputting them into ANN units because it is sensitive to the input order.

To reduce the computational complexity of the eigenvalues solution, a parallelizable eigenvalue solver based on CME methods [20], [21], which is characterized by its large scale parallel operation, hereinafter called "CME" was employed. CME methods draw multiple closed curves on a complex plane and extract the eigenvalues included in the curves. Furthermore, this paper proposes an optical eigenvalue demodulation method based on the combined use of CME and ANN. The contributions of this paper are summarized below.

- An error-free operation is achieved using CME and ANN for four eigenvalues extracted from the waveform with 2000 km fiber propagation.
- It is implied that the penalty for the optical signal and noise power ratio (OSNR) depends on the extraction range of

the continuous eigenvalues.[1] Since the modulated eigenvalues are only "discrete eigenvalues," there seems to be no need to extract "continuous eigenvalues." However, this paper shows that continuous eigenvalues also affect the BER characteristics. Therefore, we modify the range of eigenvalues extracted by CME (termed extraction range) to find the range of continuous eigenvalues that affect the BER characteristics.

- CME parameters are investigated to satisfy the accuracy required in optical eigenvalue communications.
- To reduce computational complexity, in addition to the proposed method, a partial sort method, that divides eigenvalues into small groups and employs sorting for each small group was proposed.

In the following sections, Section II introduces related works on optical eigenvalue communications. Section III describes the proposed method based on combined CME and ANN. The feasibility of employing the CME method for single-eigenvalue communication is simulated as a preliminary study in the case of on-off encoding in Section IV. The optimal parameters of the CME method are also evaluated for eigenvalue communication. In Section V, 2000 km fiber-transmission experiments are conducted to evaluate the proposed scheme in the case of an on-off encoded four-eigenvalue communication. Section VI is the conclusion.

## II. RELATED WORK

### A. Eigenvalue Communication

Optical eigenvalues derived by IST[2] have been widely studied as ideal modulation carriers that are unaffected by chromatic dispersion and nonlinear optical effects during optical fiber propagation. In 1993, a modulation technique using eigenvalues was proposed by Hasegawa and Nyu [23]. The details of the eigenvalue modulation scheme with the IST are presented in Appendix A.

With the advent of digital coherent technologies, many studies using optical eigenvalues have been proposed to increase the transmission capacity in recent years, such as the on-off encoding scheme [1], [2], [9], [10], [11], [12], [13], [14], [15], where multiple eigenvalues are associated on a complex plane, or the modulated spectral amplitude of multisoliton pulses [16], [17], [18], [19]. In the former scheme, the bit sequence corresponds to the on-off state of the eigenvalues, as shown in Fig. 1. When multiple eigenvalues are used, multiple bits can be transmitted by using a single pulse. The details of the on-off encoded multiple-eigenvalue modulation scheme and its demodulation method are presented in Appendices B and C, respectively.

However, the variance of the eigenvalues caused by bandwidth limitation and amplified spontaneous emission (ASE) noise is
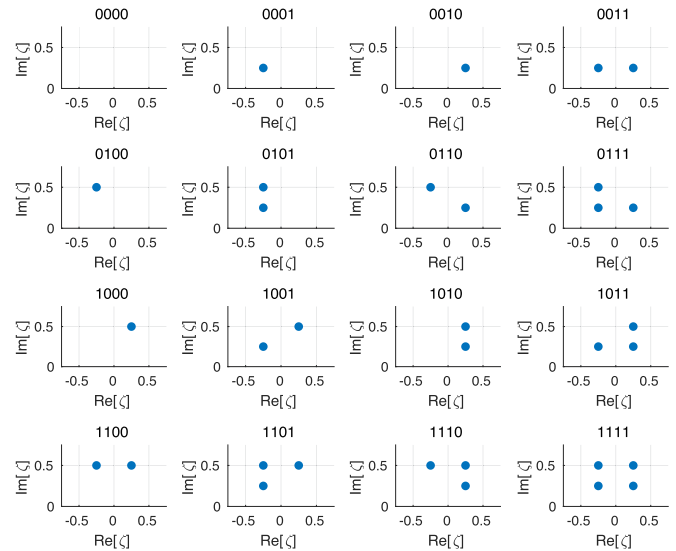


Fig. 1. Combination of four eigenvalues corresponding with 4-bit allocation.

not followed in a circularly symmetric complex Gaussian distribution [24]. In addition, the variance depends on the eigenvalue position and the combination of eigenvalues. Although eigenvalues were used in the same position, the variance is changed by eigenvalues with other positions. For example, the use of either eigenvalue set $\in \{i, -1 + i\}$ or $\{i, 1 + i\}$ is assumed for on-off encoded eigenvalue modulation. Eigenvalue $\{i\}$ was assigned to all eigenvalue sets; however, the amount of variance of eigenvalue $\{i\}$ was different between $\{i, -1 + i\}$ and $\{i, 1 + i\}$. This situation makes it extremely difficult to optimize the boundary of the symbol decision using linear thresholds on the complex plane. A decision process with the Euclidean minimum distance was demonstrated to set an appropriate threshold [17]. Machine learning-based demodulation methods (in particular, ANN) [14], [15], [19] and equalization [18], [25] have also been proposed.

The ANN cases include two types of typical demodulation methods: demodulation in the time domain (TD) [12], [14], [19] and in the eigenvalue domain (ED) [14], [15]. Our previous works [12], [14] using TD demodulation indicated a drastic power margin of 11 dB compared with the IST and hard decision demodulation methods. Meanwhile, previous TD works required the preparation of a training model of ANNs for each transmission distance. It is important to demodulate in the ED to maximize the inherent advantages of eigenvalue demodulation. This is because demodulation in the eigenvalue plane enables a distance-independent ANN demodulation. Our group has demonstrated an ED demodulation method that exhibits high generalization performance over propagation distances up to 3000 km for signals with on-off encoded eigenvalues [14] using four eigenvalues as shown in Fig. 1. To improve transmission capacity, a multilabel learning method that can demodulate on-off encoded twelve-optical-eigenvalue signals [13] over propagation distances of up to 200 km is also demonstrated [15].

However, to demodulate the signal in the eigenvalue plane, the receiver must perform eigenvalue computation and sort

---

[1]The continuous eigenvalues are defined as a part of the continuous spectrum. In the IST theory [22], they are also continuous real-valued eigenvalues. So, hereafter, we call them are also eigenvalues. Since the discrete Fourier transform (DFT) is used with a finite time window when the eigenvalues are calculated, the continuous spectrum was discretized on the real axis.

[2]IST is also well known as nonlinear Fourier transformation (NFT), e.g. see [7].

Fig. 2.    CME parallel computation.



Fig. 3.    Block diagram of eigenvalue demodulation scheme using QZ method [14]. (Coh. recv. : Coherent receiver, ADC : Analog-digital converter, IST : Inverse scattering transformation, ANN : Artificial neural network).
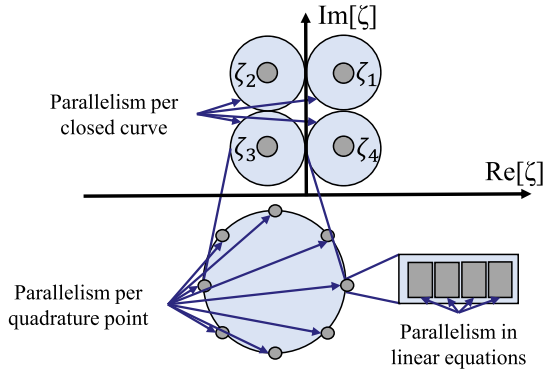
symbol-by-symbol (soliton pulse by soliton pulse). This requires enormous computational complexity; therefore, employing the CME [20], [21] to reduce the computational complexity required to extract the eigenvalues is considered.

### B. Eigensolver

Eigenvalues are generally obtained using QZ, which is a stable eigenvalue solving algorithm. QZ extends the QR algorithm for the standard eigenvalue problem ($\mathbf{Az} = \zeta\mathbf{z}$) to the generalized eigenvalue problem ($\mathbf{Az} = \zeta\mathbf{Bz}$) [26]. When using unitary matrices $\mathbf{Q}$ and $\mathbf{Z}$, square matrices $\mathbf{A}$ and $\mathbf{B}$ are transformed into upper triangular matrices $\mathbf{QAZ}$ and $\mathbf{QBZ}$. However, QZ requires large computational complexity $\mathcal{O}(n^3)$ where the matrix size is $n \times n$. In the eigenvalue communication, as discussed in the next section, $n$ is the two-fold number of samples per soliton pulse. When solving for eigenvalues with QZ, eigenvalues are obtained from the diagonal elements of the upper triangular matrix as $\{\zeta_1, \zeta_2, \ldots, \zeta_i, \ldots, \zeta_n\} = \mathrm{diag}(\mathbf{QAZ})$ where $i$ is the index of an element of the $i$-th row and column. The output is usually ordered by the magnitude of the absolute value $|\zeta_1| \geq |\zeta_2| \geq \cdots \geq |\zeta_n|$. This output order may cause bit errors when demodulating by combining it with an ANN. Details are provided in Section III.

The CME is considered to reduce the computational complexity when solving for the eigenvalues because QZ requires a large complexity. Fig. 2 shows an overview of the CME method [20], [21]. The CME sets arbitrary circular regions in a complex plane and solves the eigenvalues inside these circular regions. CME allows parallel computation for each circular region. In Fig. 2, three circles are drawn to extract five eigenvalues. To simulate the circle region in the computer, the circle is discretized and several quadrature points are prepared. Each quadrature point (eight points in Fig. 2) used in the numerical integration approximation in each circular region is parallelized, and each quadrature point is calculated in parallel. Thus, three parallel operations are performed. The mathematical details of CME are presented in Appendix D.

As a similar technique, a technique using contour integrals has been proposed [27]. This work can efficiently find discrete eigenvalues. The stable algorithm locating zero proposed by L. M. Delves and J. N. Lyness [28] is used in [27]. The related
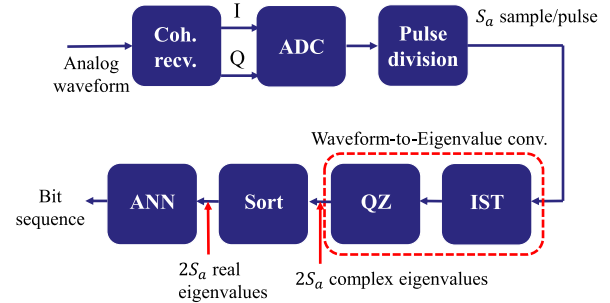
work [27] takes advantage that the poles of function $a(\zeta)$ described in NFT are discrete eigenvalues. CME also partially uses this algorithm [28]. As an advantage, CME can obtain the arbitrary eigenvalues in closed curves since it is a generic eigenvalue-seeking method. That is, CME can obtain continuous eigenvalues in addition to discrete ones. Our used ANN needs not only discrete eigenvalues but also continuous ones. Another vital advantage is the ability to perform parallelization massively.

Although the overall computational complexity of the CME is larger than that of the QZ, the computational complexity is $\mathcal{O}(Nn^3/N_p)$ because of hierarchical parallelism. Large-scale parallelization significantly reduces computing complexity compared to QZ $\mathcal{O}(n^3)$. Here, $N$ denotes the number of quadrature points prepared for the numerical integration approximation on each closed circle, and $N_p$ is the number of parallelizations. The larger the value of $N$, the higher the computational performance and the smaller the eigenvalue error. However, because the size of $N$ is directly related to the receiver scale, it is preferable to keep the size as small as possible when using this method in optical fiber communication systems. The next section describes the proposed method by combining CME with ANN, including the sorting algorithm, and comparing it with the QZ and ANN case.

## III. Proposed Scheme

### A. QZ Method With ANN (Baseline Method)

Fig. 3 shows an overview of the eigenvalue demodulation scheme. An analog optical signal is separated into IQ components by a coherent receiver and converted into a digital signal by an analog-to-digital converter (ADC). The obtained digital pulse train is divided into pulses, and the optical eigenvalues are extracted by IST. The pulse division must continue to be obtained with the same interval. That it, the periodic refresh operation is needed under the timing jitter. Note that this paper assumed the ideal condition for the jitter. Let the pulse division method under the jitter condition be a further issue. Optical eigenvalues are extracted using QZ. Let the sampling rate be $S_a$ samples/pulse and the number of extracted eigenvalues be $2S_a$. The typical number of $S_a$ is 64 or 128 [12], [13], [14], [15]. The number of samples is determined by the spectrum bandwidth. The modulated discrete eigenvalues mapped on large value in
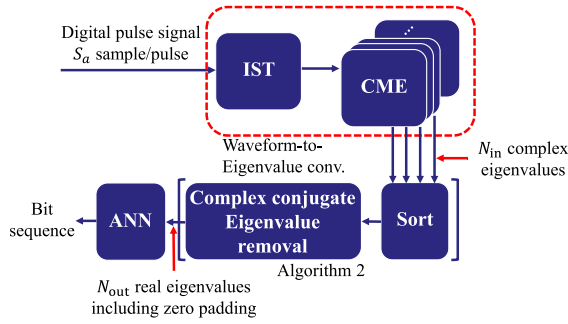
Fig. 4. Block diagram of eigenvalue demodulation scheme using CME method.

---

**Algorithm 1:** Sort in Fig. 3.

**Input**: $\zeta_{\mathrm{in}} \in \mathbb{C}^{2S_a}$       ▷ input eigenvalues
**Output**: $\zeta_{\mathrm{out}} \in \mathbb{R}^{2S_a}$      ▷output eigenvalues
1: **Sort** $\zeta_{\mathrm{in}}$ by their real parts $\Re\{\zeta_{\mathrm{in}}\}$
2: $\zeta_{\mathrm{tmp}} \leftarrow \mathbf{0} \in \mathbb{C}^{S_a}$        ▷initialization
3: **for** $i = 0$ to $S_a - 1$:
4:    **if** $\Im\{\zeta_{\mathrm{in}}^{(2i)} < 0\}$:    ▷element $\zeta_{\mathrm{in}}^{(2i)} \in \zeta_{\mathrm{in}}$
5:      $\zeta_{\mathrm{tmp}}^{(i)} \leftarrow \zeta_{\mathrm{in}}^{(2i)*}$
6: $\zeta_{\mathrm{out}} \leftarrow [\Re\{\zeta_{\mathrm{tmp}}\}, \Im\{\zeta_{\mathrm{tmp}}\}]$    ▷concatenate vectors

---

the direction of real axis have the larger value of $S_a$. That is, an increase in the number of modulated discrete eigenvalues brings that in the matrix size.

Algorithm 1 shows the sorting algorithm for eigenvalues. The eigenvalue set $\zeta_{\mathbf{in}} \in \mathbb{C}^{2S_a}$ extracted using QZ is prepared. In general, the eigenvalue set $\zeta_{\mathbf{in}}$ comprises complex eigenvalues and their conjugates. That is, half of the eigenvalues in the set do not contribute to bit information. Thus, the conjugate elements are removed using Algorithm 1.[3] First, the algorithm sorts the eigenvalues in the set in the real part direction (Line 1). The complex conjugate eigenvalues are removed from the sorted optical eigenvalues (Line 2-5). Subsequently, the real and imaginary parts of the eigenvalues $\zeta_{\mathbf{tmp}} \in \mathbb{C}^{S_a}$ are decomposed. The decomposed eigenvalues are connected to form a new eigenvalue set, $\zeta_{\mathbf{out}} \in \mathbb{R}^{2S_a}$ (Line 6). The eigenvalue set $\zeta_{\mathbf{out}}$ is used as the input to the ANN. In addition, when extracting optical eigenvalues, the QZ method requires a large computation time $\mathcal{O}(n^3)$, where the matrix size to solve for eigenvalues is $n$. Therefore it is necessary to reduce the computational complexity. An algorithm is proposed in [14]. The proposed algorithm is described in the following subsection. This paper compared the proposed method with this previous work as a baseline to determine its feasibility.

### B. Proposed CME With ANN

Fig. 4 shows a block diagram of demodulation using a CME combined with an ANN. The process before inputting the signal

---

---

**Algorithm 2:** Sort in Fig. 4.

**Input**: $\zeta_{\mathrm{in}} \in \mathbb{C}^{N_{\mathrm{in}}}$       ▷input eigenvalues
**Output**: $\zeta_{\mathrm{out}} \in \mathbb{R}^{N_{\mathrm{out}}}$      ▷output eigenvalues
1: **Sort** $\zeta_{\mathrm{in}}$ by their real parts $\Re\{\zeta_{\mathrm{in}}\}$
2: $\zeta_{\mathrm{tmp}} \leftarrow \mathbf{0} \in \mathbb{C}^{N_{\mathrm{out}}}, n \leftarrow 0$     ▷initialization
3: **for** $i = 0$ to $N_{\mathrm{in}} - 1$:
4:    **if** $\Im\{\zeta_{\mathrm{in}}^{(i)} > 0\}$:
5:      $\zeta_{\mathrm{tmp}}^{(n)} \leftarrow \zeta_{\mathrm{in}}^{(i)}$
6:      $n \leftarrow n + 1$
7: $\zeta_{\mathrm{out}} \leftarrow [\Re\{\zeta_{\mathrm{tmp}}\}, \Im\{\zeta_{\mathrm{tmp}}\}, \mathbf{0} \in \mathbb{R}^{N_{\mathrm{out}} - 2|\zeta_{\mathrm{tmp}}|}]$
     ▷concatenate vectors

---

into the IST is the same as in Fig. 3. The parallelized CME extracts eigenvalues in each circle from the matrix generated by the IST. Because the CME prepares multiple overlapped circles for eigenvalue extraction, duplicated eigenvalues can be extracted multiple times. The algorithm calculates the Euclidean distance between all eigenvalues in the overlapped region to remove the duplicated eigenvalues. When the Euclidean distance is less than the circle's radius, these eigenvalues are regarded as duplicated and removed. Thus, the duplicated eigenvalues in the CME algorithm are removed by referring to the overlap of the circles, and the unduplicated eigenvalues are input into the sorting algorithm.

Algorithm 2 shows the sorting algorithm for the eigenvalues extracted by CME. Eigenvalue set $\zeta_{\mathrm{in}} \in \mathbb{C}^{N_{\mathrm{in}}}$ where $N_{\mathrm{in}}$ is the number of extracted eigenvalues from the CME, is sorted in the real direction (Line 1). Eigenvalues with negative imaginary parts are removed from the sorted eigenvalue set (Lines 2–6). The complex eigenvalues output by the IST appear as pairs of their complex conjugate elements. The conjugate eigenvalues have the same information as the original eigenvalues; therefore conjugate eigenvalues are deleted. CME extracts only the eigenvalues inside the drawn circles, which makes it difficult to extract the eigenvalues if they are outside the circles. Because the number of input units of the ANN is fixed, it is necessary to match the number of eigenvalues to this number of units. Therefore, if the CME results in a shortage of extracted eigenvalues in this algorithm, zero-padding is employed (Line 7). In addition, the real and imaginary parts of the eigenvalue set $\zeta_{\mathrm{tmp}}$ are decomposed. The decomposed eigenvalues are concatenated to construct the eigenvalue set $\zeta_{\mathrm{out}}$ (Line 7). The eigenvalue set $\zeta_{\mathrm{out}}$ is used as the input to the ANN.

### C. CME With Partial Sort

As shown in Figs. 3 and 4, the eigenvalue set is sorted in the real part direction, and the eigenvalues are then divided into real and imaginary parts and input into the ANN. This sorting process requires $\mathcal{O}(n^2)$ of computational complexity using a simple algorithm. Even with a quicksort, which is well known as a fast sorting algorithm, the average time is $\mathcal{O}(n \log n)$ and the worst-case time is $\mathcal{O}(n^2)$. Thus, the complexity of the sorting process must be reduced.
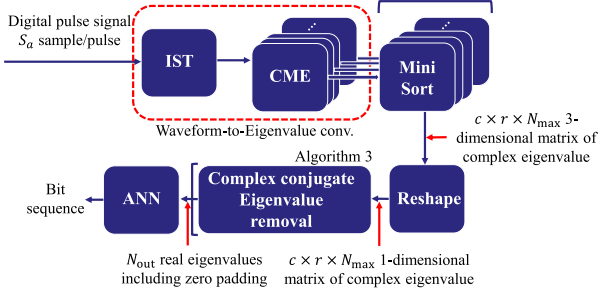
Fig. 5. Block diagram of eigenvalue modulation-demodulation scheme using partial sorting.



Fig. 6. Impact of the number of quadrature points $N$ on calculation accuracy.

---

**Algorithm 3:** Sort Used in Fig. 5.

**Input**: $\zeta_{\text{in}} \in \mathbb{C}^{c \times r \times N_{\max}}$      ▷input eigenvalues
**Output**: $\zeta_{\text{out}} \in \mathbb{R}^{N_{\text{out}}}$      ▷output eigenvalues
1: **for** $i = 0$ to $r - 1$ :
2:    **Sort** $\zeta_{\text{in}}(:, i, :)$ by their real parts $\Re\{\zeta_{\text{in}}(:, i, :)\}$
3: **Reshape** $\zeta_{\text{in}}$ from 3- to 1-dimension
4: **Remove** Null elements from $\zeta_{\text{in}}$
5: $\zeta_{\text{tmp}} \leftarrow \mathbf{0} \in \mathbb{C}^{N_{\text{out}}}, n \leftarrow 0$      ▷initialization
6: **for** $i = 0$ to $|\zeta_{\text{in}}| - 1$ :
7:    **if** $\Im\{\zeta_{\text{in}}^{(i)} > 0\}$ :
8:      $\zeta_{\text{tmp}}^{(n)} \leftarrow \zeta_{\text{in}}^{(i)}$
9:      $n \leftarrow n + 1$
10: $\zeta_{\text{out}} \leftarrow [\Re\{\zeta_{\text{tmp}}\}, \Im\{\zeta_{\text{tmp}}\}, \mathbf{0} \in \mathbb{R}^{N_{\text{out}} - 2|\zeta_{\text{tmp}}|}]$
   ▷concatenate vectors

---

This paper proposes a partial sort method that groups several circles of CME and employs the sort algorithm for each group. Fig. 5 shows a block diagram of the eigenvalue demodulation method using CME and partial sorting. The process before inputting the received signal into the IST is the same as in Fig. 3. Algorithm 3 presents an algorithm for partial sorting.

Eigenvalue set input into the algorithm denotes $\zeta_{\text{in}} \in \mathbb{C}^{c \times r \times N_{\max}}$ where $c, r \in \mathbb{Z}$ denote the identifiers of each circle of the CME in the vertical and horizontal directions, respectively. A circle at the origin of the complex eigenvalue plane is set as $(c, r) = (0, 0)$. The third dimension, $\zeta_{\text{in}}$ represents the number of eigenvalues in each region. $N_{\max}$ denotes the maximum number of eigenvalues. Nulls are stored when the number of eigenvalues in a circle is less than $N_{\max}$.

The eigenvalue set $\zeta_{\text{in}}$ extracted by the CME is grouped and sorted in the direction of the real part for circles with the same real part identifier $r$ (Lines 1–2). Next, the 3-dimensional eigenvalue set is reshaped to a 1-dimensional eigenvalue set (Line 3). Line 4 eliminates the null elements. The following operations are the same as those in Algorithm 3 (Lines 6–11). Depending on the number of circles, partial sorting reduces the complexity compared to sorting the entire eigenvalue set.

## IV. INVESTIGATION OF CME AVAILABILITY

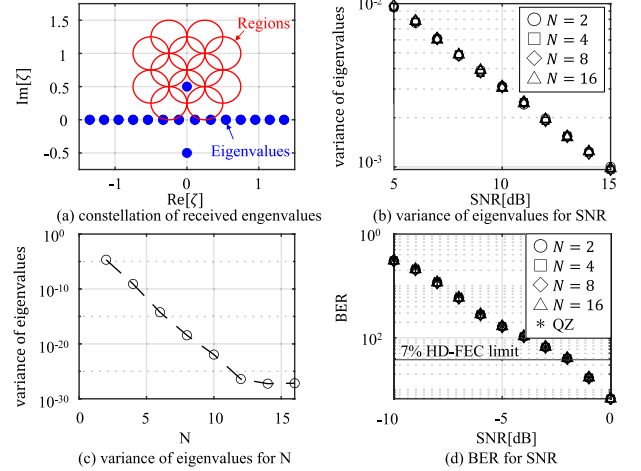We firstly employed CME on eigenvalue communication with a simulation in the case of a single eigenvalue withon-off

encoding $\zeta \in \{0.5i\}$. The simulation investigated the required number of quadrature points $N$ of the CME parameter in eigenvalue communication. In CME, while parallel operations are required for as many quadrature points as possible, $N$ should be as small as possible to reduce the complexity; there is a trade-off in which a small $N$ results in a large computational error. Therefore, the minimum $N$ parameter was determined.

In the simulation setup, the bit sequence generated by the transmitter was encoded into a single eigenvalue sequence. The IST transforms the eigenvalue sequence into a pulse train. The pulse train was input into the additive white Gaussian noise (AWGN) channel. We added the AWGN using a preset signal-to-noise power ratio (SNR). For the definition of SNR, the soliton pulse is discretized 32 samples per pulse. The simulation combined the multiple pulses and generated the pulse train. The signal power is calculated from the pulse train. Signal power divided by SNR determines the noise power.

At the receiver side, eigenvalues were extracted from the divided pulses using the IST. Fig. 6(a) shows the constellation of back-to-back eigenvalues and designated regions for the CME.

Fig. 6(b) shows the variance of the eigenvalues for the change in SNR. The variance was calculated between the ideal and received eigenvalue positions. The quadrature points were set as $N = 2, 4, 8,$ and 16. The effect of changing $N$ on variance is small. Fig. 6(c) shows the variance with $N$ in the back-to-back simulation without noise. Compared to Fig. 6(b), the variance caused by the CME calculation error was negligible. addFor example, when $N = 2$, the eigenvalue has the about $10^{-5}$ variance. This value is less than the variance caused by the noise. Althogh general CME applications focus on extracting exact eigenvalues, optical eigenvalue communication can use CME with a minimal number of $N$ because the noise in the channel is dominant.

To investigate the effect of the number of $N$ on the demodulation performance, the results of the bit error rate (BER) measurements are shown in Fig. 6(d). The simulation set the threshold for the hard decision to $0.25i$. Note that this simulation did not employ ANN. Even when $N$ was varied from 2 to 16, the BER results were the same as in the QZ case. In this simulation,
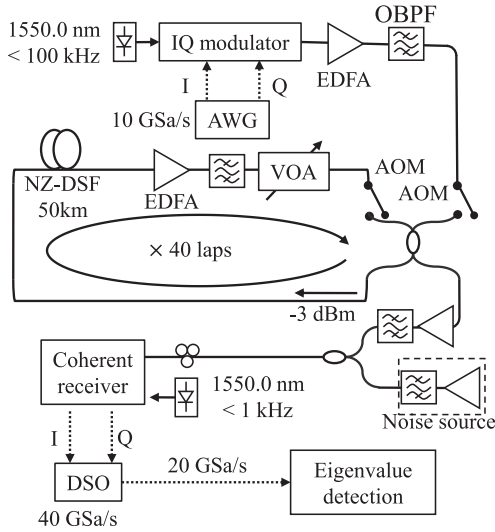
Fig. 7.    Experimental setup.

the use of a single eigenvalue was assumed. When multiple eigenvalues trsare used, several eigenvalues may exist in a circle. In this case, the CME requires higher number of $N$. Therefore, depending on the number of eigenvalues used, the receiver must determine the number of $N$.

## V. EXPERIMENT

### A. Setup

2000 km transmission experiments were conducted. Fig. 7 shows the experimental transmission system. The transmitter generates a random bit sequence and encodes every four bits into an eigenvalue pattern, as shown in Fig. 1. In total, 62250 pulses were generated. The pulse sequences were input into an arbitrary waveform generator (AWG) at 10 GSa/s (Sa/s means samples per second) with 32 samples/pulse and 5.3-GHz bandwidth limitation. The bit rate was 1.25 Gbps (= 0.315 Gpulse/s × 4 bit/pulse). The RF signal was modulated into an optical signal by an IQ modulator, which has 20-GHz bandwidth limit, and input into a 2000 km transmission system.

The transmission system comprises a 50 km non-zero dispersion-shifted fiber (NZ-DSF), an erbium-doped fiber amplitude (EDFA), a 100-GHz optical bandpass filter (OBPF), variable optical attenuator (VOA), and an acousto-optic modulator (AOM). After 2000 km transmission, the OSNR was adjusted by adding ASE noise using an EDFA.

For the transmission length, the dispersion length $L_D$ (nonlinear length $L_{NL}$) should be sufficient longer than the distance of the amplifier spacing to suppress the fiber loss effect based on the guiding-center soliton system [29]. In the experiment, the dispersion length $L_D$ could be lengthen to 446 km by using a NZ-DSF with a dispersion parameter $D$ of 4.4 ps/nm/km, which was sufficient longer than the distance of the amplifier spacing of 50 km. In addition, an inter-symbol interference is one of main impairments to limit the transmission distance in the eigenvalue transmission system [30]. This is the reason why many studies on experimental demonstrations of eigenvalue transmission and

nonlinear frequency division multiplexed (NFDM) system employed a NZ-DSF. Based on these backgrounds, the transmission distance was set at 2000 km.

The initial transmission power is set by balancing dispersion length ($L_D = t_0^2/|\beta_2|$) with nonlinear length ($L_{NL} = 1/\gamma P_0$) where $t_0$ is an input pulse width of fundamental soliton at $\zeta = 0.5i$, $\beta_2$ is dispersion of the group velocity, $\gamma$ is nonlinear coefficient, and $P_0$ is the peak power of fundamental soliton. That is, the peak power is expressed as,

$$P_0 = \frac{|\beta_2|}{\gamma t_0^2}. \tag{1}$$

Considering the guiding-center theorem [29], then the input power $P_{in}$ is given by,

$$P_{in} = \frac{2\Gamma Z_a}{1 - \exp(-2\Gamma Z_a)} P_0. \tag{2}$$

where $\Gamma = \alpha t_0^2/|\beta_2|$, $\alpha$ is the fiber loss, and $Z_a$ is the amplifier spacing. We multiplied $P_{in}$ by the normalized power $|q|^2$ described in Appendix A. Consequently, as a practical value, the input average power was set to $-3.0$ dBm taken into account the pattern sequence, the loop loss, connection loss, and amplified spontaneous emission (ASE) noise from EDFA. $t_0$ was set to 50 ps.

The optical signal was received by a coherent receiver of 20-GHz bandwidth limit, and the stored pulse sequences were resampled to 20 GSa/s using a 40-GSa/s and 50-GHz digital sampling oscilloscope (DSO), and then eigenvalue demodulation was performed.

### B. Parameters

Three algorithms were used for eigenvalue demodulation: (1) *QZ + Sort + ANN*, (2) *CME + Sort + ANN*, and (3) *CME + Partial sort + ANN*. The parameters for (2) and (3) are the same. The parameters for (1) and (2) are as follows.

*1) QZ + Sort + ANN:* The ANN used has a three-layer structure with one hidden layer consisting of 64 units for the input layer, 256 units for the hidden layer, and $16 (= 2^4)$ units for the output layer. The input unit layer has a 2-fold sampling rate to decompose the eigenvalues into real and imaginary components. For the activation function, a rectified linear unit (ReLU) function was applied to the hidden layer, and a softmax function was used for the output layer. The optimization method for learning the ANN was Adam [31], with an initial step size of $\alpha = 0.001$ and exponential decay rates of $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and a coefficient of $\epsilon = 10^{-8}$. *MATLAB Deep Learning Toolbox* released by Mathworks Inc. was used as the ANN framework. The BER measurement was performed assuming 10000 of the 62250 pulses were training data and the remaining 52250 pulses were test data.

*2) CME + Sort + ANN:* CME was implemented using z-Pares [32]. z-Pares is a package for solving parallel eigenvalues using the Sakurai-Sugiura method [20], [21]. Fig. 8 shows the designated region for eigenvalue solving using the CME method. There are two types of eigenvalues; discrete eigenvalues and
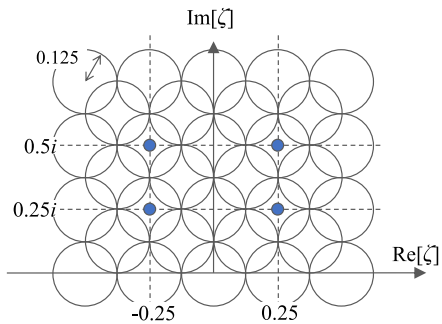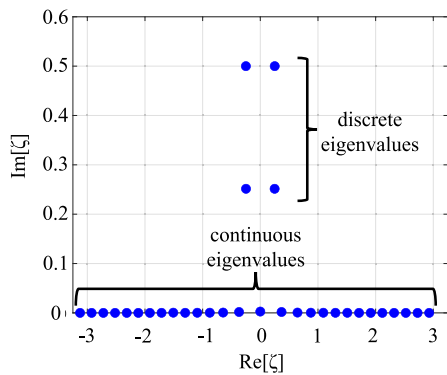
Fig. 8. Eigenvalue extraction regions.



Fig. 9. Transmission of mapped 32 eigenvalues.



Fig. 10. A part of constellation maps of eigenvalues extracted by (a) QZ and (b) CME in back-to-back transmission with OSNR = 16.3 dB.



Fig. 11. BER curves at back-to-back/2,000 km transmission and QZ/CME.



Fig. 12. Eigenvalue extraction regions when expanding the regions.

continuous eigenvalues. The example of the discrete eigenvalues appear in Fig. 1. Continuous eigenvalues have continuous spectrum [22]; however, when extracted from the sampled waveforms, they become discrete values and appear along the real axis. Fig. 9 shows the mapping of all 32 eigenvalues of 4 discrete and 28 continuous eigenvalues.

For ANN parameters, the input layer of ANN was 32 units, the hidden layer was 128 units, and the other parameters were the same as those in (1) *QZ + Sort + ANN*.

## C. Eigenvalue Demodulation Performance With CME Method

Fig. 10 shows the constellation of eigenvalues using (a) QZ and (b) CME. Although QZ in Fig. 10(a) extracts all eigenvalues existing in the complex plane, Fig. 10(b) obtains eigenvalues based on the region, as shown in Fig. 8. In the CME case, the four on-off encoded *discrete* eigenvalues were perfectly extracted. Fig. 11 shows the BER results for back-to-back and 2000 km transmissions. At 7% overhead, hard decision forward error correction (HD-FEC) limit $(= 3.8 \times 10^{-3})$ [33], the OSNR penalty between QZ and CME cases was approximately 1.6 dB for back-to-back and 2 dB after 2000 km transmission. The proposed scheme using CME and ANN was effective when these penalties were acceptable. Hypothetically, the OSNR penalty appeared because continuous eigenvalues other than the four discrete eigenvalues likely affected the demodulation performance. The following section investigats the effect of the continuous eigenvalues, for which only a part is obtained.
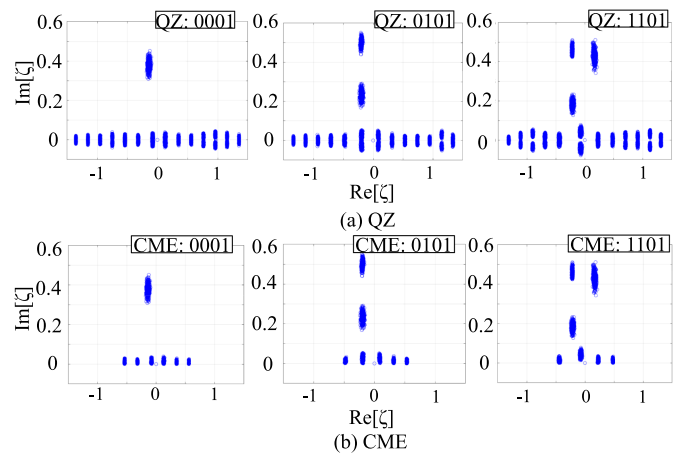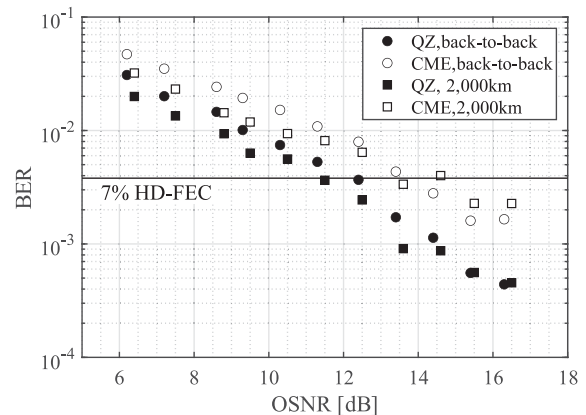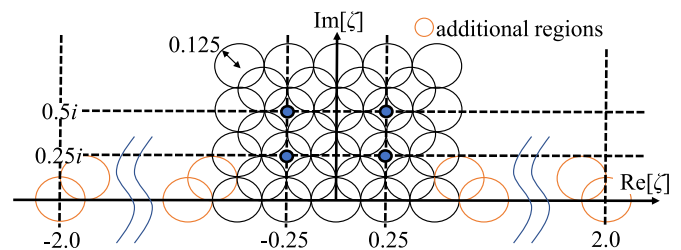
## D. Impact of Continuous Eigenvalues on BER Characteristics

The previous section discussed the reason of OSNR penalty between employing QZ and CME methods. A demodulation scheme using continuous eigenvalues has been reported [34]. To our best knowledge, there is no report that continuous eigenvalues are essential in the case of the on-off encoding scheme. In this section, we varied the number of circles in the CME to expand the extraction region of the eigenvalues. In particular, to extract the continuous eigenvalues, we added the circles on the real part axis. Fig. 12 shows the expanded circular region of the CME. Fig. 13 shows the constellation of eigenvalues extracted from the updated circular region. The additional region as shown
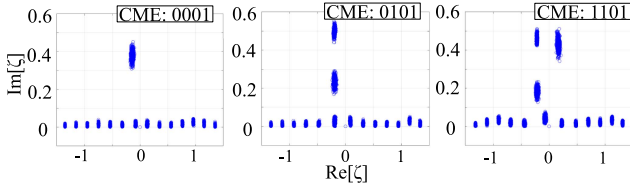
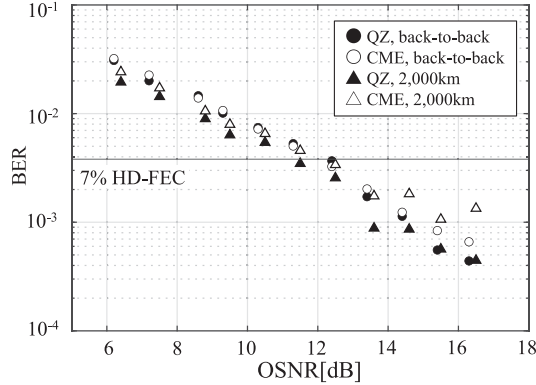Fig. 13. Eigenvalue constellation by CME method with expanding regions.



Fig. 14. BER curves with expanding CME region compared with QZ.
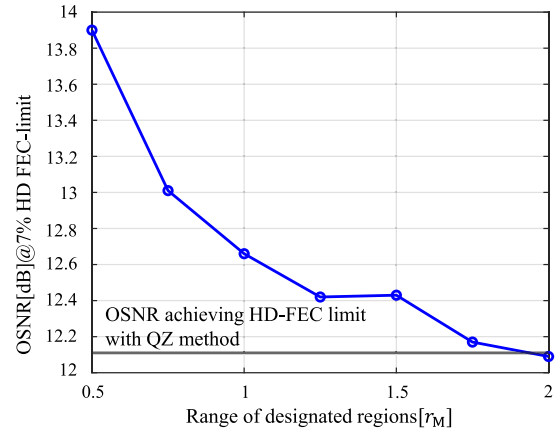


Fig. 15. Impact of eigenvalue extraction range on BER characteristics.



Fig. 16. Constellation of eigenvalues for varying the number of quadrature points $N$. (Back-to-back).

in Fig. 12 provides a wide range of continuous eigenvalues along the real axis. Fig. 14 shows the BER results for back-to-back and 2000 km transmission. By extracting continuous eigenvalues along the real axis, the OSNR penalty in Fig. 11 disappeares, and the demodulation performance is comparable to that of QZ. Therefore, ANN is likely to use the variation of discrete and continuous eigenvalues for class classification. Depending on the positions of the discrete eigenvalues, the variance of the continuous eigenvalues caused by noise is different; therefore, continuous eigenvalues are also likely to become features for bit decisions in the ANN.

The above analysis extracted all 32 eigenvalues using multiple circles of the CME. In the CME case, we can select any number of circles we use. In addition, a smaller number of circles reduces the receiver's complexity. The next analysis investigated the minimum number of circles of the CME to minimize the OSNR penalty. For the circlar positions of the CME, it was assumed that the circles were positioned with imaginary axis symmetry, as in Fig. 12. Here, the center coordinates of the maximum and minimum circles along the real axis were defined as $(c, r) = (0, \pm r_M)$. For example, in Fig. 12, $r_M$ was set to $\pm 2.0$. If the circles are removed at $r_M = \pm 2.0$; then, $r_M$ is reset to $\pm 1.75$. This procedure was used to adjust the number of circles. Fig. 15 shows the OSNR that satisfies the 7% HD-FEC limit by changing the $r_M$. The wider the eigenvalue extraction range, the higher the noise tolerance. By setting the eigenvalue extraction range to $|r_M| < 1.75$, the OSNR tolerance is close to that of the QZ case. In other words, when each discrete eigenvalue turned on or off according to the bit pattern, continuous eigenvalues in $|r_M| < 1.75$ affected the bit decision characteristics in the ANN. Meanwhile, increasing the number of circles to extract a wide range of continuous eigenvalues increases receiver complexity.
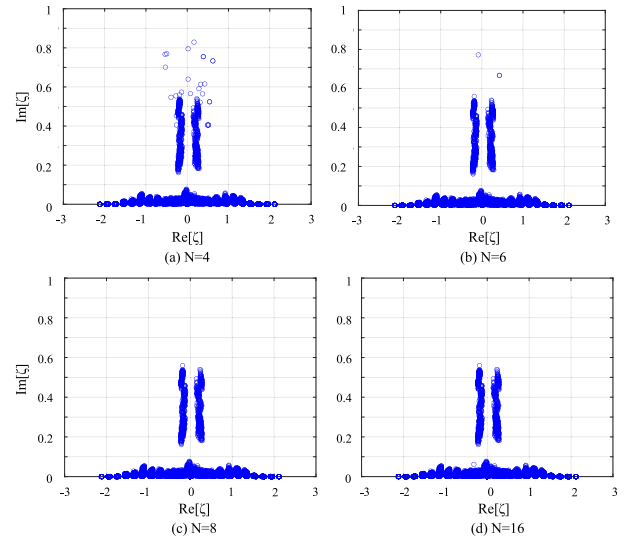
Therefore, it is necessary to determine the extraction region from the perspective of the required OSNR and available complexity.

### E. Impact of $N$ on the Accuracy

The smaller number of quadrature points $N$ in the CME is a critical factor that causes worse accuracy than that of QZ. In Section IV, better results were obtained with $N = 2$ in the case of a single eigenvalue communication; however, it is expected that the same results are not always obtained in the multiple eigenvalue cases. This subsection analyses the case with four eigenvalues. Fig. 16 shows the constellation maps on the complex eigenvalue plane with varying $N$: (a) $N = 4$, (b) $N = 6$, (c) $N = 8$, and (d) $N = 16$. In (a) $N = 4$, many outliers appear compared to (c) $N = 8$. When a CME is required to obtain several eigenvalues, accuracy depends on $N$. That is, the error in the numerical integration with the quadrature points corresponds to the accuracy of the eigenvalue extraction. When a larger error occurs, the eigenvalues move to random positions,
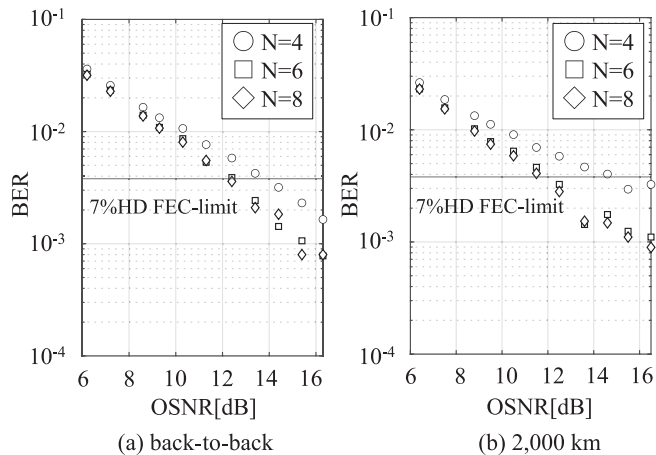
Fig. 17. Impact of the number of quadrature points $N$ on BER characteristics.



Fig. 18. BER characteristics with partial sort.

and thus outliers appear. In (b) $N = 6$, whereas there is a slight random outliers, it was more suppressed than in (a) $N = 4$. Note that with an increase in noise or the number of eigenvalues in a circle, the outliers are likely to increase, even though $N = 6$.

Fig. 17(a) shows the back-to-back BER measurement results for different values of $N$. For $N = 4$, an OSNR penalty of approximately 1.5 dB is observed at the 7% HD-FEC limit. The reason for the penalty is the outliers on the constellation map. The results for $N = 6$ were almost the same as those for $N = 8$. Outliers of the eigenvalues were observed for $N = 6$ as shown in Fig. 16(a). Meanwhile, there was a small number of pulses where outliers eigenvalue, that did not affect the BER performance.

Fig. 17(b) shows the BER measurement results after a 2000 km transmission with varying $N$. When the number of quadrature points is $N = 4$, the error floor appears at approximately the 7% HD-FEC limit. The error floor is expected to be caused by the random variance of the eigenvalues, as in the back-to-back case. The OSNR penalty was approximately 0.2 dB, even at $N = 6$. This is because random variance occurred, albeit only slightly. The higher the value of $N$, the higher the accuracy of the eigenvalue decision. From these results, we expect that $N$ needs to change depending on the number of eigenvalues within the closed curve. The fewer the eigenvalues in the closed curve, the smaller $N$ can be taken. In other words, the larger the noise is, the more the eigenvalues fluctuate, so $N$ must be larger. For four eigenvalues, it was possible to achieve the better BER results with $N = 6$ and 8. Although empirical, these results were a small value compared to the original CME application. Meanwhile, the larger the value of $N$, the larger the complexity. Therefore, it is necessary to set parameters in terms of accuracy for the required OSNR and complexity.

### F. Employing Partial Sort

To further reduce computational complexity, a proof-of-principle experiment was conducted for partial sorting. Using the
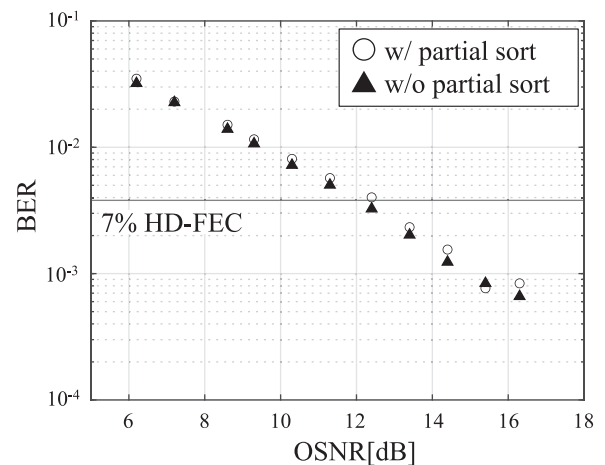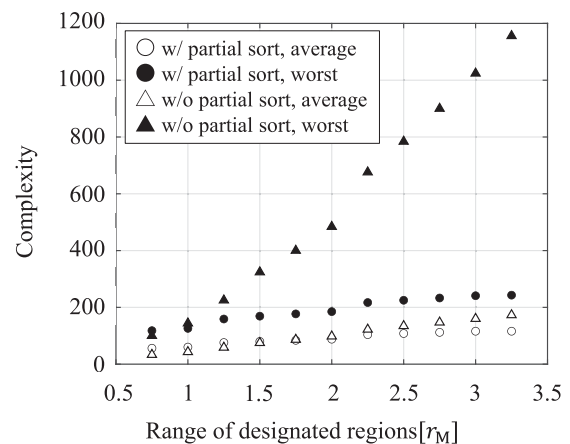


Fig. 19. Complexity characteristics. Y-axis indicates the order of complexity.

measurement data obtained from Fig. 7, we use the algorithm for eigenvalue demodulation, as shown in Fig. 5 and Algorithm 3. The parameters were the same as those described in Section V-D.

Fig. 18 shows the BER results when using the partial sort compared to the BER, as shown in Fig. 14. The BER characteristics were almost comparable for all measured points.

We confirmed the complexity reduction by employing the partial sort compared with the normal quick sort applied to all eigenvalues. Fig. 19 shows a comparison of sorting complexity. We varied the $r_M$ from $\pm 0.75$ to $\pm 3.25$. The vertical axis indicates the estimated computational complexity by comparing only the sort process. Another complexity is the same compared between CME with and without the partial sort.

Let the number of eigenvalues denote $N_\zeta$, we estimated the average complexity as $O(N_\zeta \log_2(N_\zeta))$, and the worst complexity was estimated as $O(N_\zeta^2)$. When using the normal quick sort without the partial sort, we referred the above estimation on the all extracted eigenvalues. That is, when $N_\zeta = 32$, the average and worst complexities were $32 \times \log_2(32) = 160$ and $32^2 = 1024$. Here note that, when the complexity is $O(f(n))$,

we refer to only the order $f(n)$. In partial sort, we estimated the complexity for each eigenvalue group indicated by Algorithm 3, and their complexities were summarized. We determined the merged complexities as the total complexity and plotted on Fig. 19.

The computational complexity of using the partial sort is smaller when the specified region by CME is set to $|r_M| \geqq 1.75$. For the average complexity at $|r_M| = 1.75$, both sorting methods are almost same values. At $|r_M| = 3.25$, the partial sort obtained 32.9% complexity reduction compared with normal quick sort. For the worst case at $|r_M| = 1.75$ and $|r_M| = 3.25$, the partial sort obtained 55.8% and 79.0% complexity reduction, respectively. In the worst complexity, the partial sort divided the sorting into smaller parts, so the complexity was suppressed even evaluated at $O(N_\zeta^2)$. Therefore, as a result, we use multiple eigenvalues and increase the sampling rate, then the partial sort is in significant effect.

## VI. Conclusion

This paper proposed joint demodulation of the CME and ANN for optical eigenvalue communication. In the discrete and continuous eigenvalues obtained by solving the IST, the CME extracted only the discrete and nearby continuous eigenvalues. After extracting the eigenvalues, they were input into the ANN for bit decision. As a result, a 2 dB OSNR penalty was obtained compared with the QZ case. This result indicates that the proposed method is practical for systems that can accept this penalty.

Furthermore, to investigate the reason for the 2 dB OSNR penalty, the CME expanded the circlar regions on the complex eigenvalue constellation map and increased the number of extracted continuous eigenvalues. The BER characteristics were comparable to those of QZ when the CME circle, which is the maximum distance between the center coordinate and the original point, was positioned at $|r_M| < 1.75$. In other words, We revealed that the variation of the discrete and continuous eigenvalues in $|\Re[\zeta]| < 1.875$ affects the bit decision in the ANN.

We investigated the number of quadrature points $N$, an internal parameter of the CME algorithm, that is of concern as another factor affecting BER characteristics. When using a single eigenvalue, even the minimum number of quadrature points of two did not affect the BER characteristics; however approximately six quadrature points were required when using multiple eigenvalues.

A partial sorting algorithm was proposed and validated with experiments. The BER characteristics have no OSNR penalty compared with using the normal quick sort applied to all eigenvalues while reducing complexity.

In a further study, FPGA implementation of the proposed method and evaluation of the throughput and latency of large-scale parallel operations are required. As a limitation of this study, the theoretical reason why continuous eigenvalues affect BER characteristics has not been revealed. Future work requires investigating the effect of continuous eigenvalues in the case of multiple discrete eigenvalue transmissions.

## APPENDIX A
### ASSOCIATED EIGENVALUE EQUATION

The behavior of the complex envelope amplitude of a light wave propagating in an optical fiber is described by [35]

$$i\frac{\partial E}{\partial z} - \frac{\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + \gamma|E|^2 E = -i\alpha E \tag{3}$$

where $z$ [m], $t$ [s], $E(z,t)$ [$\sqrt{W}$], $\beta_2$ [s$^2$/m], $\gamma$ [1/(m · W)], and $\alpha$ [1/m] are the propagation distance, time measured in a coordinate system moving at the group velocity, and the complex envelope amplitude of the electric field, dispersion of the group velocity, nonlinear coefficient, and fiber loss, respectively. (3) is conducted the following normalization,

$$T = \frac{t}{t_0}, \ Z = \frac{|\beta_2|}{t_0^2}z, \ q = t_0\sqrt{\frac{\gamma}{|\beta_2|}}E \tag{4}$$

where $T$ is a time normalized to the input pulse width $t_0$. Considering the normalization and guiding center theorem [29], the NLSE is given by,

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

The associated eigenvalue equations used to solve the NLSE analytically with the IST are expressed as follows:

$$\begin{cases} i\frac{\partial \psi_1}{\partial T} + q\psi_2 = \zeta\psi_1 \\ -i\frac{\partial \psi_2}{\partial T} - q^*\psi_1 = \zeta\psi_2 \end{cases} \tag{6}$$

where $\zeta$ and $\psi_l(Z,T)$ $(l = 1, 2)$ are complex eigenvalues and eigen functions, respectively. If $q(Z,T)$ satisfies (5) then, eigenvalues $\zeta$ of (6) are invariant with the distance $Z$. By mapping a bit sequence to the on-off states of the complex eigenvalues, multiple eigenvalues modulations are realized.

## APPENDIX B
### MULTIPLE EIGENVALUE MODULATION

In contrast to general multi-level modulation schemes such as quadrature amplitude modulation (QAM), multiple eigenvalues are mapped to a single symbol (one pulse) in multiple eigenvalue modulation schemes. Fig. 1 represents the optical eigenvalues in the complex plane and is an example of modulation that maps a bit sequence to the on-off states of the four eigenvalues. Turning multiple eigenvalues on and off achieved a 16 (= $2^4$) pattern of eigenvalue position [10]. This is called on-off encoding, showing that the number of bits directly corresponds to the number of multiple eigenvalues. Here, this pattern of eigenvalue placement is called the eigenvalue pattern.

After mapping the eigenvalues, they were converted into pulses using the IST described in APPENDIX A. The converted pulse is sent to a digital-to-analog conveter (DAC), and an IQ optical modulator generates an optical pulse.

## APPENDIX C
### EIGENVALUE DEMODULATION

On the receiver side, the IQ component of each optical pulse is received by a coherent receiver. The IQ components were converted into digital signals using an analog-to-digital converter

(ADC). This digital signal was divided into pulse waveforms using a window function, and the eigenvalues were extracted. To obtain the eigenvalues $\zeta$ of (6), (6) is Fourier-transformed into the integral equation shown below:

$$
\begin{cases}
\Omega\tilde{\psi}_1(Z,\Omega) + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\tilde{q}(Z,\Omega-\Omega')\tilde{\psi}_2(Z,\Omega')d\Omega' \\
\qquad = \zeta\tilde{\psi}_1(Z,\Omega) \\
-\Omega\tilde{\psi}_2(Z,\Omega) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\tilde{q}^*(Z,-\Omega-\Omega')\tilde{\psi}_1(Z,\Omega')d\Omega' \\
\qquad = \zeta\tilde{\psi}_2(Z,\Omega)
\end{cases}
\tag{7}
$$

where $\tilde{q}(Z,\Omega)$ and $\tilde{\psi}_l(Z,\Omega)$ are the Fourier transforms of $q(Z,T)$ and $\psi_l(Z,T)$, respectively. By discretizing $\tilde{q}(Z,\Omega)$ and $\tilde{\psi}_l(Z,\Omega)$ in the frequency domain and converting convolution integral of the second term on the left-hand side of (7) to the numerical integration, (6) can be converted to a matrix eigenvalue problem, as in the following equation [6]:

$$
\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline -\mathbf{B}^* & -\mathbf{A} \end{array}\right]\begin{bmatrix}\tilde{\psi}_1 \\ \tilde{\psi}_2\end{bmatrix} = \zeta\begin{bmatrix}\tilde{\psi}_1 \\ \tilde{\psi}_2\end{bmatrix}
\tag{8}
$$

where $\tilde{\psi}_l$ is a column vector whose elements are $\tilde{\psi}_l(\Omega_n)(n = 1, 2, \ldots, N)$ and $\Omega_n$ is the discretized frequency. $\mathbf{B}^*$ is the Hermitian transpose matrix, $\mathbf{B}$. $\mathbf{A}$ and $\mathbf{B}$ are square matrices of $N \times N$, and their elements are as follows:

$$
a_{jk} = \begin{cases} \Omega_j & (j=k) \\ 0 & (\text{otherwise}) \end{cases}
\tag{9}
$$

$$
b_{jk} = \begin{cases} \frac{1}{\sqrt{2\pi}}\tilde{u}(\Omega_{n_{jk}})\Delta\Omega & (1 \le n_{jk} \le N) \\ 0 & (\text{otherwise}) \end{cases}
\tag{10}
$$

where $n_{jk} = N/2 + j - k + 1$ and $\Delta\Omega(=\Omega_{n+1}-\Omega_n)$ is the frequency interval of the discretization.

## APPENDIX D
## CME METHOD

To reduce the computational complexity of solving the eigenvalue problem, the complex moment-based eigenvalue solving method is applied to optical eigenvalue communication. The Sakurai-Sugiura method (SSM), one of the complex moment-based eigenvalue solvers, was proposed by Sakurai and Sugiura in 2003, characterized by high parallel processing performance [20].

Here, the detailed algorithm of SSM is introduced. SSM is a parallel computation algorithm that solves only the eigenvalues within a specified region. First, a closed curve was drawn on the complex plane. The center of the closed curve is $\gamma$, the radius is $\rho$, and the quadrature point $z_j = \rho\xi_j + \gamma(j = 1, 2, \ldots, N)$ is on the closed curve where $\xi_j = \cos\theta_j + i\sin\theta_j$ and $\theta_j = (2\pi/N_s)(j - 1/2)$. QZ extracts all eigenvalues in the complex plane, whereas SSM only finds the eigenvalues in the specified region. In addition, each specified region can be calculated in parallel, which is the first parallelization. The following equation is then computed.

$$
\mathbf{f}_j = \mathbf{u}^{\mathsf{H}}(z_j\mathbf{I} - \mathbf{A})^{-1}\mathbf{v}
\tag{11}
$$

The eigenvalues within each region were obtained as poles by the spectral decomposition of the matrix bundle in the above (11). (11) can be computed in parallel at each integration point: and is the second parallelization step. The simultaneous calculation of (11) can also be performed in parallel. This is the third parallelization mrthod. Where $\mathbf{u}$ and $\mathbf{v}$ are non-zero random matrices $n \times L$, where $\mathbf{I}$ is the identity matrix. The $\mathbf{A}$ is the matrix of computed eigenvalues. $L$ is a parameter determined from the number of eigenvalues in the region. The number of eigenvalues in the domain is calculated as follows:

$$
m = \sum_{j=1}^{N} w_j\text{tr}\left((z_j\mathbf{I} - \mathbf{A})^{-1}\right)
\tag{12}
$$

where $w_j$ is the weight at each quadrature point in each region, and $w_j = (\cos\theta_j + i\sin\theta_j)/N_s$. To extract the poles, (11) is approximated by trapezoidal numerical integration of the contour integral along a closed curve.

$$
\tilde{\mu}_k = \sum_{j=1}^{N} w_j\xi_j^k\mathbf{f}_j \quad (k = 0, 1, \ldots, 2M - 1)
\tag{13}
$$

where $M$ is a parameter for $M < N_s$ and must satisfy the condition $n < ML$. $n$ denotes the number of eigenvalues within a certain region. Quadrature points were used to approximate the numerical integration when performing contour integration along a closed curve. Finally, the eigenvalues can be computed from the Hankel matrix that integrates the computations at each integration point, and is represented by the following equation:

$$
\tilde{\mathbf{H}}_M^{(0)} = \begin{pmatrix} \tilde{\mu}_0 & \cdots & \tilde{\mu}_{M-1} \\ \vdots & \ddots & \vdots \\ \tilde{\mu}_{M-1} & \cdots & \tilde{\mu}_{2M-2} \end{pmatrix}
\tag{14}
$$

$$
\tilde{\mathbf{H}}_M^{(1)} = \begin{pmatrix} \tilde{\boldsymbol{\mu}}_1 & \cdots & \tilde{\boldsymbol{\mu}}_M \\ \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{\mu}}_M & \cdots & \tilde{\boldsymbol{\mu}}_{2M-1} \end{pmatrix}
\tag{15}
$$

The singular value decomposition of the obtained Hankel matrix $\tilde{\mathbf{H}}_M^{(0)}$ is $\tilde{\mathbf{H}}_M^{(0)} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^{\mathsf{T}}$, and is performed as follows:

$$
\tilde{\mathbf{H}}_M^{(0)} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{\Sigma}_1 & O \\ O & \boldsymbol{\Sigma}_2 \end{pmatrix}\begin{pmatrix} \mathbf{W}_1 & \mathbf{W}_2 \end{pmatrix}^{\mathsf{H}}
\tag{16}
$$

where $\sigma_1, \sigma_2, \ldots, \sigma_M$ are the diagonal elements of $\Sigma$, such that where $\sigma_{m'} \ge \delta$, $\sigma_{m'+1} < \delta$, and $\delta$ is an arbitrarily small value, and in this paper $\delta = 10^{-12}$. Also, from $m'$, $\mathbf{U}_1 \in \mathbb{C}^{M \times m'}$, $\boldsymbol{\Sigma}_1 \in \mathbb{C}^{m' \times m'}$, $\mathbf{W}_1 \in \mathbb{C}^{M \times m'}$. From this, the eigenvalues $\tilde{\zeta}_i$ $(i = 1, 2, \ldots, m')$ can be determined as the eigenvalues of $\tilde{\mathbf{A}} = \boldsymbol{\Sigma}_1^{-1}\mathbf{U}_1^{\mathsf{T}}\tilde{\mathbf{H}}_M^{(-1)}\mathbf{W}_1$. In this case, eigenvalue $\zeta_i$ of the matrix $\mathbf{A}$ are obtained as $\zeta_i = \rho\tilde{\zeta}_i + \gamma$.

## REFERENCES

[1] Y. Terashi, D. Hisano, K. Mishina, and A. Maruta, "Impact of arbitrary parameters in parallel eigenvalue solving algorithm on optical eigenvalue communication," in *Proc. OptoElectronics Commun. Conf.*, 2021, pp. 1–3, Paper W1B.5.

[2] Y. Terashi, D. Hisano, K. Mishina, Y. Yoshida, and A. Maruta, "Joint multi-eigenvalue demodulation using complex moment-based eigenvalue solver and artificial neural network," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2021, pp. 1–3, Paper W6A.13.

[3] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[4] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," *IEEE J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[5] A. Maruta, "Nonlinear shannon limit in optical fiber transmission system," *IEICE Trans. Commun.*, vol. 101, no. 1, pp. 80–95, 2018.

[6] H. Terauchi and A. Maruta, "Eigenvalue modulated optical transmission system based on digital coherent technology," in *Proc. OptoElectronics Commun. Conf. Photon. Switching*, 2013, paper WR2.5.

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

[8] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 1981.

[9] V. Aref, Z. Dong, and H. Buelow, "Design aspects of multi-soliton pulses for optical fiber transmission," in *Proc. IEEE Photon. Conf.*, 2016, pp. 224–225.

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

[11] T. Kodama, T. Zuiki, K. Mishina, and A. Maruta, "Hyper multi-level modulation based on optical eigenvalue Multiplexing," in *Proc. Photon. Switching Comput*, 2018, pp. 1–3.

[12] 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.*, 2019, pp. 1–4.

[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.*, 2021, pp. 1–3, Paper M3I.1.

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

[15] 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.*, 2022, pp. 1–3, Paper Th3F.3.

[16] V. Aref and H. Buelow, "Design of 2-soliton spectral phase modulated pulses over lumped amplified link," in *Proc. 42th Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.

[17] H. Bülow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," in *Proc. 42th Eur. Conf. Opt. Commun.*, 2016, pp. 1–3.

[18] 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, 2020.

[19] 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. 2018.

[20] T. Sakurai and H. Sugiura, "A projection method for generalized eigenvalue problems using numerical integration," *J. Comput. Appl. Math.*, vol. 159, no. 1, pp. 119–128, 2003.

[21] T. Ikegami, T. Sakurai, and U. Nagashima, "A filter diagonalization for generalized eigenvalue problems based on the Sakurai–Sugiura projection method," *J. Comput. Appl. Math.*, vol. 233, no. 8, pp. 1927–1936, 2010.

[22] A. Shabat and V. Zakharov, "Exact theory of two-dimensional self-focusing and one-dimensionalself-modulation of waves in nonlinear media," *Sov. Phys. JETP*, vol. 34, no. 1, pp. 62–69, 1972.

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

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

[25] 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, 2020.

[26] C. B. Moler and G. W. Stewart, "An algorithm for generalized matrix eigenvalue problems," *SIAM J. Numer. Anal.*, vol. 10, no. 2, pp. 241–256, 1973.

[27] A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, "Contour integrals for numerical computation of discrete eigenvalues in the Zakharov–Shabat problem," *Opt. Lett.*, vol. 43, no. 15, pp. 3690–3693, 2018.

[28] L. Delves and J. Lyness, "A numerical method for locating the zeros of an analytic function," *Math. Computation*, vol. 21, no. 100, pp. 543–560, 1967.

[29] A. Hasegawa and Y. Kodama, "Guiding-center soliton in optical fibers," *Opt. Lett.*, vol. 15, no. 24, pp. 1443–1445, 1990.

[30] S. Civelli, E. Forestieri, and M. Secondini, "Why noise and dispersion may seriously hamper nonlinear frequency-division multiplexing," *IEEE Photon. Technol. Lett.*, vol. 29, no. 16, pp. 1332–1335, Aug. 2017.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[32] "Z-pares: Parallel eigenvalue solver." [Online]. Available: https://zpares.cs.tsukuba.ac.jp/

[33] "ITU-TG. 975.1, forward error correction for high bit-rate DWDM submarine systems," 2005. [Online]. Available: https://www.itu.int/rec/T-REC-G.975.1/en

[34] M. Kamalian-Kopae, A. Vasylchenkova, O. Kotlyar, M. Pankratova, J. Prilepsky, and S. Turitsyn, "Artificial neural network-based equaliser in the nonlinear Fourier domain for fibre-optic communication applications," in *Proc. Conf. Lasers Electro- Opt. Europe Eur. Quantum Electron. Conf.*, 2019, pp. 1–1.

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

**Yuhei Terashi** received the B.E. degree in electrical, electronic and information engineering in 2021 from Osaka University, Osaka, Japan, where he is currently working toward the M.E. degree with the Graduate School of Engineering. His research interests include optical fiber communication. He is a Member of IEICE of 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 October 2018, he has been an Assistant Professor with Osaka University. His research interests include optical-wireless converged networks, optical communication, all-optical signal processing, visible light communication, edge computing, and application of deep learning to optical communication. He is also a Member of the IEICE. He was the recipient of IEEE Kansai Section Young Professionals Awards in 2021.

**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 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.

**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. From 2009 to 2016, he was an Assistant Professor with Osaka University, Osaka, Japan. In 2016, he joined the Network Research Institute, National Institute of Information and Communications Technology, Japan, where he currently a Research Manager. He is also a Visiting Associate Professor with Osaka University. His research interests include digital signal processing for optical/wireless communications, optical/wireless access, and optical- wireless convergence. 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 research interests include optical fiber communication systems and all-optical signal processing. He is a Member of the IEEE Photonics Society and Optical Society of America (OSA), and a Senior Member of IEICE, Japan.