Digital Longitudinal Monitoring of Optical Fiber Communication Link

Takeo Sasai, *Member, IEEE*, Masanori Nakamura, *Member, IEEE*, Etsushi Yamazaki, *Member, IEEE*, Shuto Yamamoto, *Member, IEEE*, Hideki Nishizawa, and Yoshiaki Kisaka

(Top-Scored Paper)

***Abstract*—Optical transmission links are generally composed of optical fibers, optical amplifiers, and optical filters. In this paper, we present a channel reconstruction method (CRM) that extracts physical characteristics of multiple link components such as longitudinal fiber losses, chromatic dispersion (CD), multiple amplifiers' gain spectra, and multiple filters' responses, only from receiver-side (Rx) digital signal processing (DSP) of data-carrying signals. The concept is to reconstruct a virtual copy of an actual transmission channel in the digital domain, where optical fibers and amplifiers are modeled as the split-step Fourier method for the Manakov equation while optical filters are emulated as complex-valued finite impulse response filters. We estimate the model parameters such as losses, CD, gains, and filter responses from boundary conditions, i.e., transmitted and received signals. Experimental results show that, unlike traditional analog testing devices such as optical time-domain reflectometers and optical spectrum analyzers, CRM visualizes multi-span characteristics of fibers, amplifiers, and filters in Rx DSP, and thus localizes anomaly components among multiple ones without direct measurement.**

***Index Terms*—Channel reconstruction, chromatic dispersion, digital longitudinal monitoring, gain spectrum, optical fiber loss, passband narrowing, split-step Fourier method.**

I. Introduction

DIGITAL system identification (SI) has been a core idea that has supported recent progress in optical communication. The task of SI is to find a simplified model or identify internal system parameters from a set of input and output data of a system. Here, in this paper, a system refers to an optical communication link component (e.g., an optical fiber, optical transceiver, optical amplifier, and optical filter), a combination thereof, or the whole. In general, applications of SI are classified into two types:

(i) Predicting system outputs for arbitrary inputs (and vice versa), often with a lower computational cost.

(ii) Monitoring the system status by identifying system parameters.

In optical communication, numerous studies have been devoted to application (i), which is equivalent to compensating for system impairments and restoring (predicting) transmitted signals since such equalization techniques achieve high-capacity and long-haul transmission. Adaptive equalizers for linear impairments [1], [2] and the compensation of transceivers' bandwidth limitations [3]–[5] are good examples, and state-of-the-art research accounts for component nonlinearity, such as nonlinear digital pre-distortion of optical transmitters [6]–[8] and fiber nonlinearity compensation (NLC) enhanced by optimizing parameters in existing models [9]–[13]. As an extreme example, end-to-end deep learning of the entire optical communication system [14] can also be classified here.

Meanwhile, application (ii), which is the target of this study, corresponds to the health monitoring of optical transmission systems, enabled by identifying system parameters from received signals (and sometimes transmitted signals or pilot signals); this includes total chromatic dispersion (CD) [15], fiber nonlinearity noise, and optical signal-to-noise ratio (SNR) [16]. These are mostly cumulative quantities in a link but *fiber longitudinal* parameters, such as optical fiber loss profiles, chromatic dispersion (CD) maps, and individual responses of multiple link components (amplifiers and filters), are also essential for reliable and capacity-maximized systems, and thus, all components are generally tested at the beginning of and even during operation. However, these component characteristics have so far been tested by direct measurement with analog instruments such as optical time-domain reflectometers (OTDR) and optical spectrum analyzers (OSA). Such analog approaches can be time-consuming and costly since testing instruments should be placed on-site in a span-by-span and fiber-by-fiber manner to diagnose all link components. If longitudinal multi-span characteristics of link components are obtained at once via a digital SI approach, we can expect not only a reduction in CAPEX and OPEX but also "smart optical transmission systems," in which all necessary measurement tests, connection establishment, and fault detection are conducted autonomously, and even self-maximization of system capacity by auto-adjusting unoptimized devices (e.g., amplifiers' gain tilt and filters' center frequency detuning) can be achieved [17]. Accordingly, Tanimura *et al.* have demonstrated the signal power (loss) profile extraction of

Manuscript received July 14, 2021; revised October 31, 2021 and December 8, 2021; accepted December 15, 2021. Date of publication December 29, 2021; date of current version April 18, 2022. *(Corresponding author: Takeo Sasai.)*

The authors are with NTT Network Innovation Laboratories, Nippon Telegraph and Telephone Corporation, Kanagawa 239-0847, Japan (e-mail: takeo.sasai.cp@hco.ntt.co.jp; masanori.nakamura.cu@hco.ntt.co.jp; etsushi.yamazaki.wk@hco.ntt.co.jp; shuto.yamamoto.at@hco.ntt.co.jp; hideki.nishizawa.zw@hco.ntt.co.jp; yoshiaki.kisaka.dc@hco.ntt.co.jp).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2021.3139167.

Digital Object Identifier 10.1109/JLT.2021.3139167

a 260-km multi-span link from only received signals using a correlation method (CM) [18]. Our previous studies and other groups [19]–[23] have shown that not only the power profile but also CD maps, amplifiers' gain spectra, and the passband narrowing (PBN) at optical filters can be estimated. Since these methods reveal optical components' characteristics in the fiber propagation direction with digital signal processing (DSP), we call these techniques *digital longitudinal monitoring* (DLM) here. We will review these and other related works in detail in Section II-E.

In this paper, we present a channel reconstruction method (CRM) that extracts component-wise characteristics of a whole multi-span link (optical fibers, optical amplifiers, and optical filters) only from the receiver-side (Rx) DSP of data-carrying signals. The concept is to virtually reconstruct a transmission link in the digital domain from boundary conditions, i.e., transmitted and received signals. In our CRM, optical fibers and amplifiers are modeled as the asymmetric split-step Fourier method (SSFM) for the Manakov-PMD equation, while optical filters are emulated as complex-valued finite impulse response (FIR) filters. As a result, longitudinal optical power (loss) and CD profiles, multiple optical amplifiers' gain spectra, and multiple optical filters' responses along multi-span links, are obtained without any testing instruments such as OTDR and OSA. Experimental results show that CRM successfully detects anomaly losses, anomaly fiber launch power, unexpected fibers with different CD, amplifiers with anomaly gains, and PBN effects of optical filters, which are manually inserted in multi-span links. The essential property commonly lying behind these successful localizations of anomalies is the "non-commutativity" of linear and nonlinear operations.

This paper is an invited extension of our work presented in [20], where an optical fiber power profile and individual frequency responses of multiple optical filters were successfully obtained. This study extends [20] and our previous works [19], [21], [22] from the following perspectives:

- We provide a more detailed description of CRM in terms of system modeling, problem formulation, algorithm, localization principle, and performance limiting factors. We qualitatively discuss "the nonlinearity-to-noise ratio (NLNR)" limits the performance of DLM. (Section II and Appendix)
- A detailed experimental analysis of estimated power profiles is provided by applying CRM to a long-haul link ($\sim$2080 km) under various types of link conditions: excess loss inserted, different fiber launch powers, and non-uniform optical level diagrams. Not only inserted losses but also different fiber launch powers are successfully identified. (Section III)
- Absolute values of the estimated parameters (power and CD) are shown, while relevant works [18]–[22] provided only relative values. We experimentally demonstrate that the estimated absolute signal power evolution in the propagation direction shows an excellent agreement on the theoretical power profiles. (Section III)
- The minimum detectable limit of the anomaly loss is discussed with various transmission distances and fiber launch powers, which indicates the maximum measurement range and the required launch powers. (Section III)
- Regarding optical amplifiers, we review our previous experiments [21] and show that CRM can reveal Raman-amplified power profiles and reconstruct multiple Raman gain spectra (RGS) in multi-span distributed Raman amplification (DRA) systems. (Section IV)
- As for optical filters, we experimentally show that the CRM identifies a frequency-detuned optical filter among multiple cascaded ones [20], and the shifted frequency estimation scheme is also presented and evaluated. (Section V)
- The potential applications of CRM are discussed in detail, which include system health monitoring at the beginning of and during operation and fiber nonlinear noise estimation. This part is not limited to CRM but also applicable to other DLM techniques.

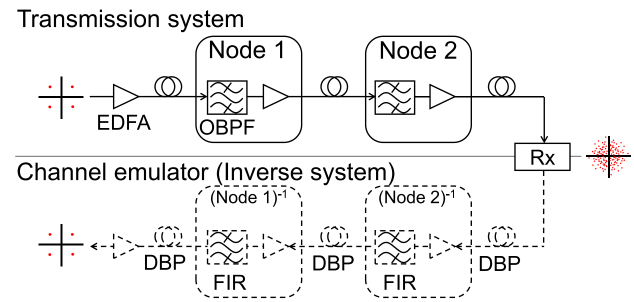

Fig. 1. Channel reconstruction for optical link component monitoring. Optical fibers, optical amplifiers, and optical filters are emulated in digital domain as split-step Fourier method and complex-valued FIR filter. Forward system model can also be adopted, though inverse system is used in this paper.

II. Channel Reconstruction for Link Component Monitoring

A. System Modeling

Our goal is to obtain optical fiber loss and CD profiles, amplifier gains, and optical filter responses on the receiver side in the manner of SI. For this purpose, as shown in Fig. 1, we reconstruct a channel emulator—an inverse system of an actual transmission link, including optical fibers, optical amplifiers, and optical filters—from boundary conditions, i.e., transmitted and received signals. Note two points: (i) although the method requires transmitted signals, it can be implemented fully in a blind manner; no special training sequence or channel for monitoring is required. This is because the transmitted signals can be fully restored in Rx DSP with a standard demodulation process (see Section III-A). This potentially enables the in-service monitoring of link components. (ii) One may find the forward system instead of the inverse system by reversing the transmitted and received signals. In this subsection, we describe the forward and inverse systems of each component we assume.

1) Optical Fibers: First, let us revisit the fiber propagation model and its inverse described in [24]–[26]. The propagation of polarization-multiplexed signals in an optical fiber is governed by the Manakov-PMD equation:

$$\frac{\partial \boldsymbol{E'}}{\partial z} = \left(-\frac{\alpha\left(z\right)}{2} - \frac{j}{2}\beta\left(z\right)\frac{\partial^2}{\partial t^2}\right)\boldsymbol{E'} + j\frac{8}{9}\gamma\left(z\right)\left\|\boldsymbol{E'}\right\|^2\boldsymbol{E'}, \tag{1}$$

where $\boldsymbol{E'} = [E_h{}'(z,\ t),\ E_v{}'(z,\ t)]^T$ denotes Jones vectors of baseband optical signals at position $z \in [0, L]$ from a transmitter and at time t, while norm $\|\boldsymbol{E'}\|^2 = |E_h{}'(z,\ t)|^2 + |E_v{}'(z,\ t)|^2$ is the instantaneous power of signals. Here, $[\cdot]^T$ is the transpose, and L is the total transmission distance. $\alpha(z)$, $\beta(z)$, and $\gamma(z)$ are loss (and amplification), group velocity dispersion, and nonlinear coefficients, respectively. By substituting the normalized amplitude $\boldsymbol{E}\left(z,\ t\right) = \boldsymbol{E'}(z,\ t)\exp(\frac{1}{2}\int_0^z \alpha(z')dz')$, which has a constant power during propagation, we can merge $\alpha(z)$ and $\gamma(z)$ as:

$$\begin{aligned}\frac{\partial \boldsymbol{E}}{\partial z} &= \left(-\frac{j}{2}\beta\left(z\right)\frac{\partial^2}{\partial t^2}\right)\boldsymbol{E} + j\gamma'\left(z\right)\left\|\boldsymbol{E}\right\|^2\boldsymbol{E} \\ &\equiv \left(\boldsymbol{D}\left(z\right) + \boldsymbol{N}\left(z\right)\right)\boldsymbol{E},\end{aligned} \tag{2}$$

where

$$\gamma'\ (z) = \frac{8}{9}\,\gamma\left(z\right)\exp\left(-\textstyle\int_0^z \alpha\left(z'\right)dz'\right). \tag{3}$$

Now, in turn, signal power variation is governed only by $\gamma'(z)$. In this study, $\gamma'(z)$ and $\beta(z)$ are the targets of estimation. If $\gamma(z)$ is constant, then $\gamma'(z)$ represents the accumulated loss and amplification and thus indirectly gives the signal power profile along a transmission link; meanwhile, $\beta(z)$ directly gives the dispersion distribution.

Eq. (2) is usually solved by the SSFM, which numerically emulates signal propagation by computing the linear part $\boldsymbol{D}(z)$ and nonlinear part $\boldsymbol{N}(z)$ independently and iteratively [26]. Standard digital backpropagation (DBP) performs the asymmetric SSFM backwards and restores the transmitted signals from received signals [25] as:

$$\begin{aligned}\hat{\boldsymbol{E}}\ (z_{k-1}, nT) &= \exp\left(-\boldsymbol{N}\left(z_k\right)\Delta z\right) \\ &\exp\left(-\boldsymbol{D}\left(z_k\right)\Delta z\right)\hat{\boldsymbol{E}}\left(z_k, nT\right),\end{aligned} \tag{4}$$

where $z_k\ (k \in \{0,\ 1, \ldots, K\})$ denotes the sampled position (thus $z_0 = 0$ and $z_K = L$), $\Delta z = z_k - z_{k-1}$, $n \in \{0,\ 1, \ldots, N_s - 1\}$ is the sample number, T is the sampling period, and N_s is the total number of samples. DBP consists of the cascaded iteration of CD compensation (CDC) and nonlinear phase rotation (NLPR) blocks. In this study, we use DBP as the inverse system of an optical fiber and attempt to identify $\gamma'(z_k)$ in NLPR blocks of DBP.

Though the selection rule of the spatial step size Δz is beyond the scope of this paper, it has been shown that the adequate step size selection enhances the estimation performance [27]. There, the selection rule $\Delta z(z_k) \propto 1/P(z = z_k)$ provided an excellent agreement between the estimated $\gamma'(z_k)$ and the theoretical one in CRM, where $P(z = z_k)$ is the signal power at $z = z_k$. Also, the modified versions of SSFM such as symmetric SSFM are not considered in this work since our asymmetric SSFM employs a sufficient number of spatial steps and at such a finer step size, the symmetric SSFM can be approximated as the asymmetric SSFM [26]. The estimation of $\gamma'(z_k)$ and $\beta(z_k)$ will be experimentally analyzed in Section III.

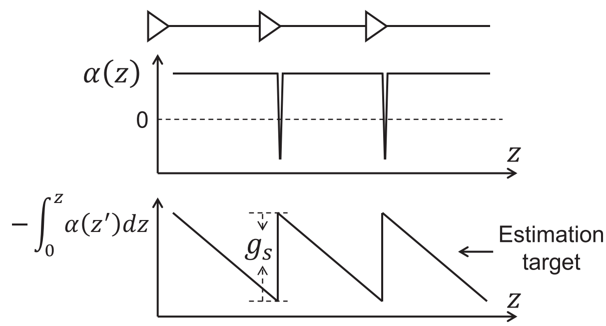

Fig. 2. $\alpha(z)$ and its integral over multi-span link. Channel reconstruction method (CRM) indirectly estimates integral version of $\alpha(z)$ from $\gamma'(z_k)$ and thus reveals optical amplification gains.

2) Optical Amplifiers: Optical amplifier gain can be included in $\alpha(z)$ (and thus in $\gamma'(z)$) as:

$$\alpha\left(z\right) = \alpha_0 + g\left(z\right), \tag{5}$$

where $\alpha_0 > 0$ is a constant fiber loss, and $g(z) < 0$ is the pointwise amplification gain. Fig. 2 illustrates the $\alpha(z)$ and its integral for a lumped amplification system. In this case, $g(z) = \sum_{s=1}^S g_s\delta(z - z_s)$, where g_s and z_s are the s-th amplifier's gain and position ($s \in \{1, 2, \ldots, S\}$), respectively. Thus, its integral inside the exponential in (3) becomes $\alpha_0 z + \sum_{s=1}^S g_s u(z - z_s)$, where $u(z)$ denotes a step function. Consequently, identifying $\gamma'(z)$ unveils both fiber loss and amplifier gain. DRA can also be included by modifying $g(z)$ in a distributed manner.

As we have mentioned, we extract $\gamma'(z)$ from data-carrying signals, which implies that the estimated gain is channel-specific. Thus, by obtaining $\gamma'(z)$ from all wavelength division multiplexed (WDM) channels and observing them in the frequency direction, we can reconstruct the gain spectrum of an amplifier. Notably, multiple amplifiers' gain spectra in a multi-span link can be separately obtained at once since $\gamma'(z)$ reflects distance-wise signal power. This separation of multiple gain spectra reveals anomaly amplifiers and is experimentally shown in Section IV for DRA systems.

3) Optical Filters: In reconfigurable optical add-drop multiplexing (ROADM) systems, signals often suffer from PBN [28], which occurs at optical filters in ROADM nodes (e.g., wavelength selective switch, WSS) and shaves off the higher-frequency component of signals as shown in Fig. 3. This impairment comes from misalignment or center frequency detuning of optical filters. We attempt to identify such anomaly filters by estimating their inverse response with an FIR filter that compensates for the PBN. Here, let $H\left(f\right) = F[h'(t)]$ be a fully-tuned filter response at baseband with no frequency deviation $\Delta f = 0$, where F denotes the Fourier operator and $h'(t)$ the time-domain impulse response. Then, the detuned response with deviation Δf is represented as follows.

$$H\left(f - \Delta f\right) = F\left[h'\left(t\right)\cdot e^{j2\pi\Delta f t}\right] \tag{6}$$

Due to center frequency deviation, the symmetry of the frequency response is broken, so the time-domain expression must be complex-valued. Thus, complex-valued FIR filters $\mathbf{h} = [h(0), h(1), \ldots, h(M-1)]^T$ are adopted for optical filter

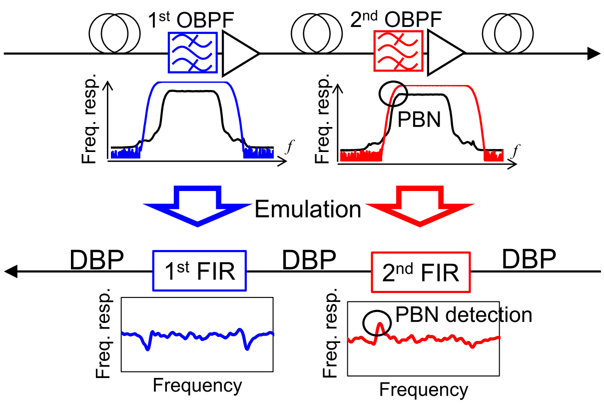

Fig. 3. Passband narrowing (PBN) emulation by complex-valued FIR filters. Corresponding FIR filter optimized in minimum mean square error criterion compensates for PBN at optical bandpass filters (OBPF).

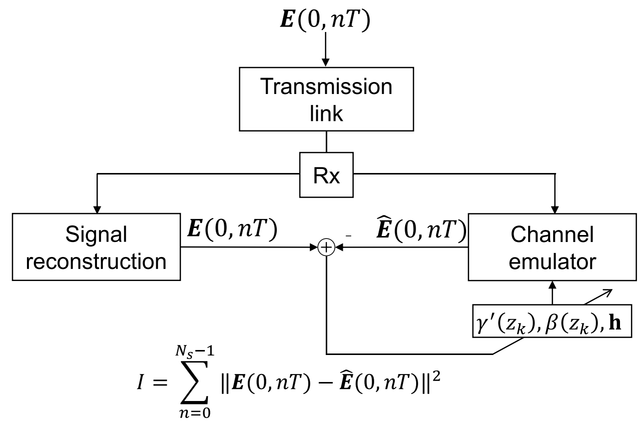

Fig. 4. Learning architecture for channel reconstruction. $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$ are optimized from channel emulator output $\hat{\boldsymbol{E}}(0, nT)$ and transmitted signals reconstructed in Rx DSP $\boldsymbol{E}(0, nT)$ as least square problems.

emulation in our inverse system, where M denotes the tap length. As shown in Fig. 3, if PBN occurs at an optical filter, the estimated frequency response of the corresponding FIR filter will show a peak at attenuated frequency components. We insert these FIR filters between DBP blocks at corresponding node positions and estimate their responses simultaneously. We present experiments on filter response separation in Section V.

B. Problem Formulation and Algorithm

Again, our estimation target is $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$. We find these parameters as optimal values that best emulate the propagated link, in which the received signals $\boldsymbol{E}(L, nT)$ return to the transmitted signals $\boldsymbol{E}(0, nT)$. This problem can be formulated as a classical least square problem in SI as:

$$\hat{\gamma}'(z_k), \hat{\beta}(z_k), \hat{\mathbf{h}} = \operatorname{argmin} I = \operatorname{argmin} \sum_{n=0}^{N_s-1} \left\| \boldsymbol{E}(0, nT) - \hat{\boldsymbol{E}}(0, nT) \right\|^2. \tag{7}$$

where I is a square-error cost function. Eq. (7) can be solved by a gradient descent using error backpropagation [29] as the system we assume is composed of cascaded differentiable blocks (e.g., CDC, NLPR, and FIR-filter blocks). As shown in Fig. 4, the parameters $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$ are optimized so that the mean square error between output signals of the channel emulator $\hat{\boldsymbol{E}}(0, nT)$ and transmitted signals $\boldsymbol{E}(0, nT)$ is minimized. To minimize it, the error backpropagation calculates the partial derivatives $\partial I/\partial\gamma'(z_k)$, $\partial I/\partial\beta(\underline{z_k})$, and $\partial I/\partial\bar{\mathbf{h}}$, which are used for coefficient updates, where $\overline{(\cdot)}$ denotes the complex conjugate. Here, Wirtinger calculus is adopted to deduce the partial derivatives since each block is complex-value-based. To calculate the cost function I, the transmitted signals $\boldsymbol{E}(0, nT)$ are required but can be generated by the normal demodulation process in the Rx DSP; no training sequence or dedicated channel for monitoring is needed. One can add a regularization term in (7) to enhance the performance. In [30], the CRM with a regularization term, which constraints the power fluctuation, is discussed. This regularization may suppress the fluctuation of the obtained $\gamma'(z_k)$ by sacrificing the sensitivity to the anomaly loss. No such regularization term was added in the following experiments to clearly observe the inserted anomaly loss.

Fig. 5 shows the detailed algorithm we use. As shown in Fig. 5(a), the entire algorithm is composed of data preprocessing, channel emulators, and cost function I. Here, an example of the data size used for 280-km transmission is shown. For data preprocessing, the input vectors are polarization demultiplexed signals, and 1024 samples ($=$ one waveform) are taken out of each polarization every two samples. Then, these synchronized *h*- and *v*-polarization signals are stored in a batch ($=60000$ waveforms $=30000$ waveforms $\times$ 2 pol.). During the optimization iterations, a mini-batch ($=100$ waveforms) is randomly chosen from a stored batch for every iteration to perform the stochastic gradient descent. Since the CD response in the first and the last 1/4 parts in a waveform is not sufficiently compensated for during the DBP due to the absence of adjacent samples, these uncompensated samples are eliminated before the cost function *I*.

After data preprocessing, a mini-batch is fed into the channel emulator (Fig. 5(b)), which consists of the iteration of linear CDC and NLPR blocks, complex-valued FIR filters, and residual phase rotation. Here, we assume that the channel emulator has the same S spans as an actual transmission link and, in the s-th span ($s \in \{1, 2, \ldots, S\}$), the K_s steps of DBP are implemented, and S-1 FIR filters for ROADM node filter emulation are placed at corresponding node positions.

Fig. 5(c)–(f) shows the gradient descent calculation of the fast Fourier transform (FFT) and the inverse FFT (IFFT), CDC, NLPR, and complex-valued FIR-filter blocks, respectively. The details of gradient calculations are described in Appendix and one can find all these calculations very simple. After the entire inverse system is passed through, the residual phase is compensated for and the signals are fed into the cost function. By using obtained partial derivatives, the coefficients $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$ are updated. For $\mathbf{h}$, an optimizer, Adam, is used [29].

C. Why is Localization Successful?: Non-Commutativity

From estimated $\gamma'(z_k)$ and $\beta(z_k)$ profiles, we can localize the position of anomaly loss, dispersion, and amplifiers. Similarly,

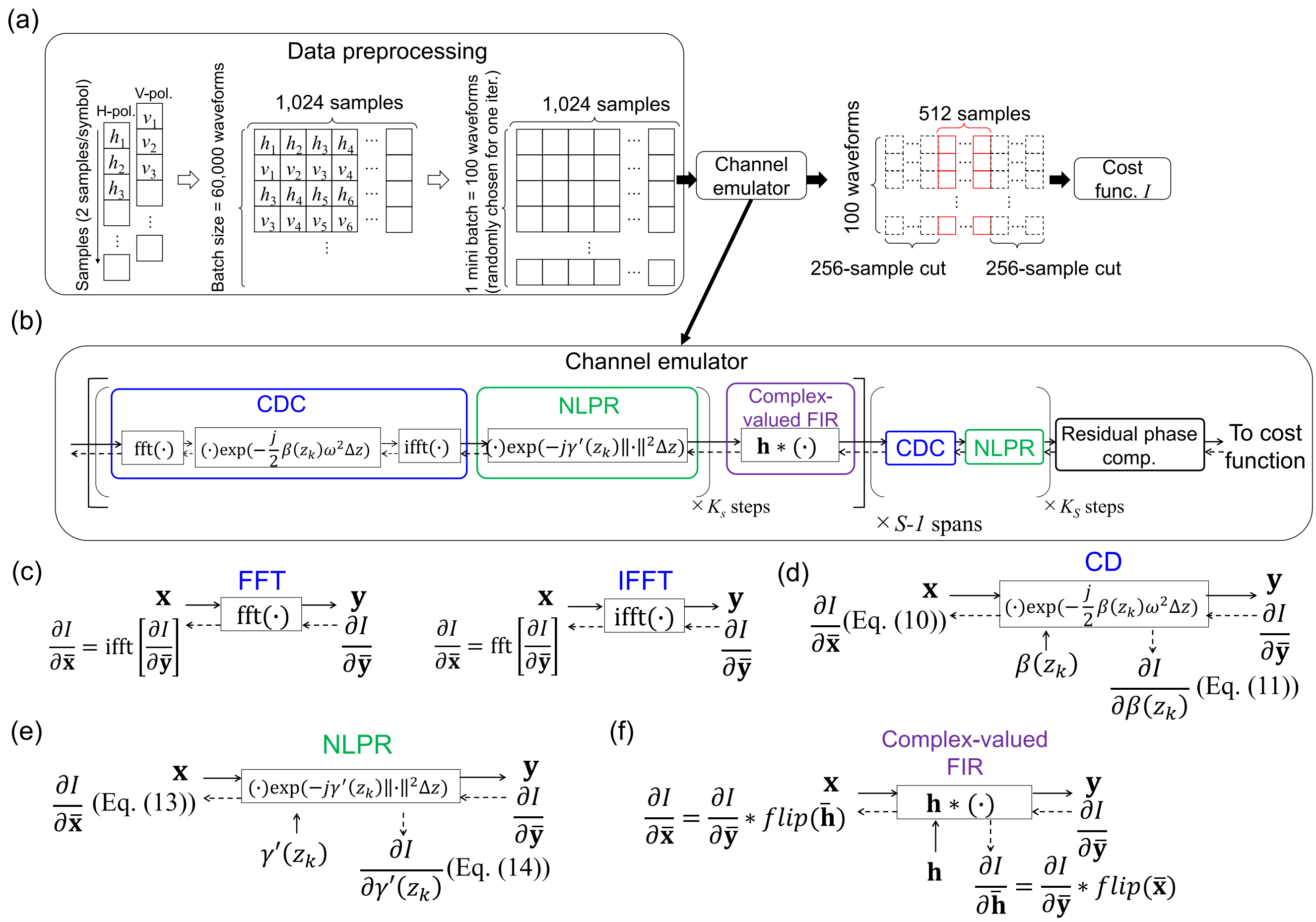

Fig. 5. Algorithms of CRM. (a) Algorithm and data structures for CRM, whose estimation targets are $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$. (b) Channel emulator whose structure fully reflects actual transmission channel. As channel parameters $\gamma'(z_k)$, $\beta(z_k)$, and $\mathbf{h}$ are optimized, virtual link is reconstructed. Backpropagation algorithm of (c) FFT, IFFT, (d) CDC, (e) NLPR, (f) complex-valued FIR filters. Wirtinger calculus is used for these complex-valued gradient calculations.

from $\mathbf{h}$, the responses of two different optical filters can be separately obtained, and anomaly filters can be identified. A notable commonality among these localization principles is that they leverage the non-commutativity of linear and nonlinear operations to identify the position of physical characteristics. For example, in this study, we can estimate pointwise dispersion $\beta(z_k)$ and nonlinear coefficients $\gamma'(z_k)$ because (linear) CD and NLPR operators ($\exp(\boldsymbol{D}(z_k)\Delta z)$ and $\exp(\boldsymbol{N}(z_k)\Delta z)$) are not commutative. Likewise, the responses of two optical filters are basically not separable at the receiver side when the system is completely linear. Their responses can be distinguished by leveraging fiber nonlinearity between optical filters since linear and nonlinear phenomena are non-commutative. For signals that have passed through multiple non-commutative systems with no noise, a reverse-order operation is a good (but may not be a unique) candidate for the inverse that minimizes the cost function in the sense of the zero-forcing (ZF) criterion. Therefore, we can determine the order and positions of the systems.

In general, the same notion is also applicable to other non-commutative systems. In [8], it was demonstrated that the modulator driver nonlinearity in an optical transmitter can be effectively compensated for by treating the transmitter structure—i.e., the digital-to-analog converter, drivers, and modulators—as a Wiener-Hammerstein model, (cascaded FIR filter, nonlinear function, and another FIR filter), which is an L-N-L system and thus non-commutative.

D. Performance Limiting Factor

However, as we will see later in the results section, we rarely find a unique inverse of a system for following reasons. (i) Finding system parameters of differential equations from boundary conditions is mathematically the inverse problem, in which there is no guarantee that the solution is unique. (ii) The discussion in Section II-C assumes the system has no noise and thus the inverse can be found on the basis of the ZF criterion. However, CRM inherently works on the basis of the minimum mean square error (MMSE) criterion, and, in a realistic scenario, the received signal accompanies a certain amount of noises such as amplified spontaneous emission (ASE) noise, transceiver noise, and fiber nonlinear noise. These noises disturb tracking the true NLPR in Rx DSP, and the optimized parameters (e.g., $\gamma'(z_k)$) are likely to fall into local minima. Especially in a lower signal power area (more linear area), the reduced NLPR is more vulnerable to noises and the non-commutativity will be broken; the false estimation of NLPR in the Rx DSP is likely to occur. Thus, qualitatively speaking, the estimation precision at z_k is (at least partly) determined by the following nonlinearity-to-noise ratio (NLNR)

$$\mathrm{NLNR}\left(z_k\right)=\frac{P_{NL}\left(z=z_k\right)}{N}\,, \tag{8}$$

where $P_{NL}(z)$ is the power of nonlinearity within a processed bandwidth at measurement location z while N is that of noises at the receiver, including ASE noise, transceiver noise, incompletely compensated Kerr nonlinear term [12], DBP-enhanced noise (as will be discussed in Section III), and even control error of gradient optimization of CRM. We will see this NLNR universally explains the results of power profiles in the following experiments. From (8), higher signal power is desirable to achieve better estimation precision. If repetitive signals are available, N of NLNR can be reduced by averaging multiple synchronized signals, though this approach is prohibitive after the communication operation begins. Even if they are not, the estimation precision can be enhanced by averaging multiple profiles or filter responses by extracting and processing received signals multiple times.

E. Related Works

1) Fiber Nonlinearity Compensation: Other works [9]–[13] use the same fiber-only model in the context of fiber NLC, explained as Learned-DBP or a neural-network (NN) based DBP, where the coefficients in time-domain CD filters and NLPR are optimized and thus the (communication theoretical) SNR gain is achieved. Although these studies showed a significant NLC gain and indicated a new direction for NLC, they fall under the scope of application (i), impairment compensation, introduced in Section I, and they focused on NLC. Consequently, it has not been confirmed whether the optimized parameters actually reflect the transmission line status in detail, such as the fiber loss, fiber launch power, and fiber types. In [12] and [13], the learned coefficients (CDC filter taps and $\gamma'(z_k)$) are shown for 1-step-per-span DBP (we will see that our results match those of these works in Section III-B), but the power variation in a span and fibers with different CDs are not evaluated. This study demonstrates such link parameter optimization to be more capable at revealing even the transmission-link status (anomaly loss, CD, even gains, PBN, etc.) in detail, paving the way for application (ii), system status monitoring. This is achieved by looking into the optimized parameters, not the output of NN and its accuracy.

2) Digital Longitudinal Monitoring: For DLM, several methods have been proposed so far. In [18], a correlation method (CM) has been proposed and the power profile estimation and the anomaly loss detection for a 260-km multi-span link have been demonstrated. The calibration of the estimated anomaly loss based on CM is also discussed in [32]. In CM, a partial CDC, a fixed amount of NLPR at location z_k, and the residual CDC are applied to the received signals, and by taking the correlation between this partially NL-compensated signal and transmitted signal, the (relative) signal power at z_k is estimated. Since CM avoids an iterative optimization, it seemingly requires less computational complexity than CRM. However, the estimated profile is vulnerable to channel noise and a sufficient number of averaging is required. In [18], 33000 averaging was applied for 260-km transmission, which accumulates a large number of FFTs. Regarding the functionality, CM assumes a constant CD map over the entire link, and thus $\beta(z_k)$ profile estimation is currently challenging. Very recently, as another DLM method, a Volterra-based method has been proposed in [23] and the filter frequency detuning estimation as well as the longitudinal power profile estimation are numerically verified. A fair evaluation should be performed in future work to compare these algorithms.

Implementation aspect of DLM has been discussed in [33]. There are two implementation scenarios for DLM: the edge-side implementation near transceivers and the cloud-side implementation near a network control plane. The former has the advantages of low latency and data security. Sending a waveform data to the central network controller may raise security concerns since the DLM reconstructs the transmitted signal during its process. The latter can utilize rich computational resources and accumulated data for performance enhancement and long-term monitoring.

III. Optical Fiber Loss and CD Profile Extraction

A. Proof of Concept: Metro-Reach Links

We first conducted experiments to examine the effectiveness of CRM for a straight-line 4-span link. To focus on $\gamma'(z_k)$ and $\beta(z_k)$ profile extraction only, we did not insert any optical filters at nodes and eliminated FIR filters from our CRM.

Fig. 6 shows the experimental setup and a block diagram of the offline DSP. A probabilistically shaped (PS) 64-QAM 64-GBd signal (400 G) with a period of 65536 symbols is generated with an information rate of 3.305 bits and an entropy of 4.347 bits, assuming a 21% forward error correction (FEC) overhead [34]. The signal is Nyquist-pulse-shaped by using a root-raised-cosine filter with a roll-off factor of 0.2; then, the transmitter's frequency response (FR) is compensated for [5], followed by a resampling block for emission from a 120-GSa/s arbitrary waveform generator (AWG). The generated electrical signal is converted into an optical signal by 65-GHz drivers and a dual-polarization IQ modulator (DP-IQM). A micro-integrable tunable laser assembly (μITLA) with a linewidth of 40 kHz and a carrier wavelength of 1555.752 nm was used for the transmitter laser and the Rx local oscillator. The fiber launch power was set to +5 dBm. The tested fibers are standard single-mode fibers (SSMF, α = 0.186 dB/km, β_2 = −21.7 ps^2/km), dispersion shifted fibers (DSF, α = 0.230 dB/km, β_2 = −0.486 ps^2/km), and non-zero DSF (NZ-DSF, α = 0.225 dB/km, β_2= −3.33 ps^2/km).

On the receiver side, the signal is post-amplified by an erbium-doped fiber amplifier (EDFA), filtered by a 5-nm optical bandpass filter (OBPF), and converted into electrical signals by a coherent receiver composed of a 90° hybrid and 100-GHz-bandwidth balanced photodetectors (BPDs). The received signals are digitized by a 256-GSa/s digital sampling oscilloscope (DSO) and demodulated offline in the Rx DSP. There,

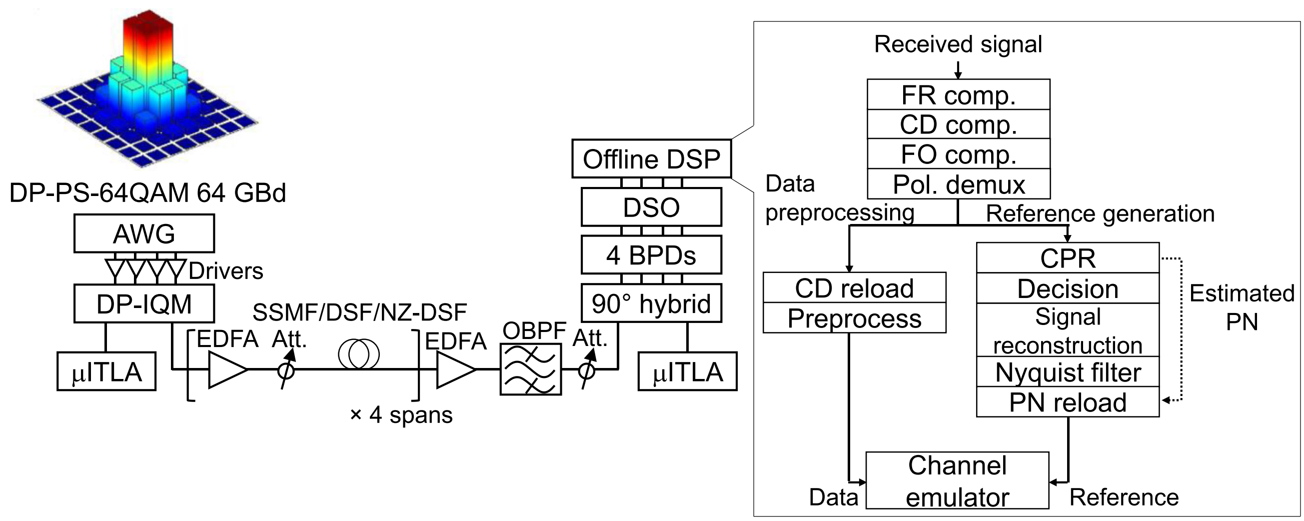

Fig. 6. Experimental setup for straight-line 280-km transmission. Details of transmission links are depicted with result figures. FR: frequency response; FO: frequency offset; PN: phase noise.

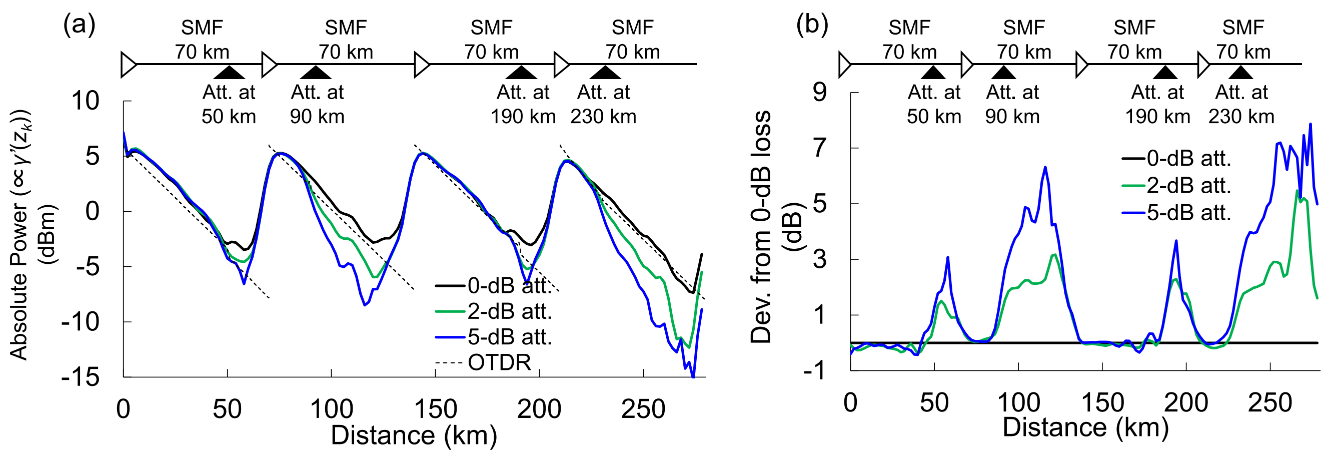

Fig. 7. (a) Obtained power profiles along 70 km $\times$ 4 SSMF span link. OTDR loss profile is also shown for reference. (b) Differences between normal (0-dB loss) and anomalous (2- and 5-dB losses) state profiles. Twenty profiles are averaged for each line.

the Rx FR estimated in advance is compensated for [5]; this is followed by frequency offset (FO) compensation [35]. After CDC, polarization demultiplexing is applied by an adaptive equalizer with a butterfly configuration.

Then, the signal is divided into two paths. One is for signal preprocessing to restore the dispersion and reshape the demodulated vectors, and the other is for transmitted signal reconstruction for gradient descent. The latter path consists of carrier phase recovery (CPR) [36], symbol decision, transmitted signal reconstruction, the same Nyquist filtering as on the transmitter side, and reloading of the phase noise estimated in CPR. Regarding the channel emulator, the length per DBP step (distance granularity) Δz is a constant 2 km and the mini-batch size is 100. No regularization term is added to the cost function. The initial values of $\gamma'(z_k)$ and $\beta(z_k)$ were set to 0 and the average CD (i.e., total CD/total DBP steps), respectively. All the algorithms were coded in MATLAB. For a $\gamma'(z_k)$ profile, the learning process was finished within 1 minute on a single GPU with 11 GB of memory. 20 profiles are averaged to enhance the profiles' SNR.

First, we obtained $\gamma'(z_k)$ (loss or signal power) profiles for the four spans of 70 km, as shown in Fig. 7(a). The vertical axis is the absolute power calculated from the estimated $\gamma'(z_k)$ as:

$$P(z_k) = \frac{9}{8}\frac{\gamma'(z_k)}{\gamma}, \tag{9}$$

where γ is the nonlinear coefficients of used SSMFs (1.07 1/W/km, measured by combining the method in [37] and SSFM simulation). Here, we assume that, the signal power $\|\boldsymbol{E}\|^2$ is normalized to 1 in the Rx DSP. Note that, in a practical use of CRM, γ is usually unknown and thus only the relative power (but absolute $\gamma'(z_k)$) is estimated. To see whether CRM can extract the true link parameters, we show the absolute power, assuming γ is known. The theoretical power calculated from the OTDR loss profile is also shown for reference. The iteration of the gradient descent was performed 15 times for these profiles. In all cases, the estimated power profiles clearly reflect the fiber loss and the amplification by the EDFA. More importantly, their absolute power shows an excellent agreement on the theoretical

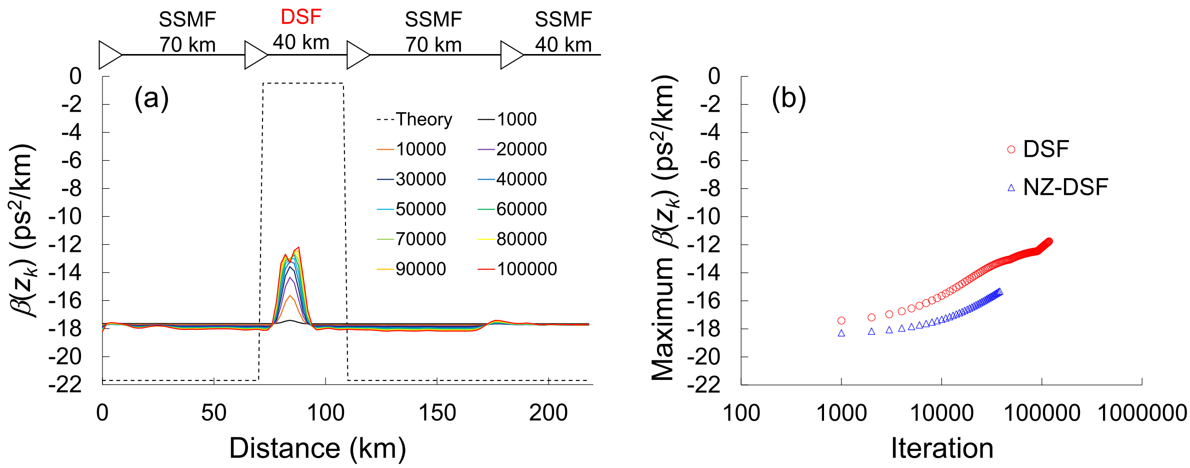

Fig. 8. (a) Obtained $\beta(z_k)$ profiles with a DSF in the second span at various gradient descent iterations (1000 to 100000). (b) Maximum value of $\beta(z_k)$ as a function of gradient descent iterations.

absolute power. When 2-dB (green) and 5-dB (blue) attenuators were inserted at 50, 90, 190, and 230 km, reduction of $\gamma'(z_k)$ was observed due to excess loss at the insertion points. Fig. 7(b) shows the difference between the normal (0-dB attenuation) and abnormal conditions. The inflection points of the rising part of the difference correspond to the anomaly loss positions, and the relative attenuation level is successfully obtained. However, the peaks of the first and third spans are lower than expected. This is presumably because the nonlinearity is too weak and NLNR in (8) is low there, which led to reduced sensitivity for estimated $\gamma'(z_k)$ to respond to the discontinuous loss event.

Next, we discuss $\beta(z_k)$ profiles. Fig. 8(a) shows the estimated $\beta(z_k)$ profiles for a 4-span link with a DSF in the second span and SSMFs in the other spans. We stopped gradient descent iterations for $\gamma'(z_k)$ optimization at 20. A peak was observed in the DSF span, indicating the presence of a lower-dispersion fiber in the link. As the gradient descent iteration increases, $\beta(z_k)$ in the DSF span gradually approaches the theoretical values, but the convergence of $\beta(z_k)$ is very slow. The reason is that, since the dispersion is dominant in the system with some SSMF spans, the cost function in (7) is mainly determined by the total dispersion and thus each $\beta(z_k)$ cannot change rapidly to keep the total dispersion constant. Once the total dispersion ($=\sum_{k=1}^{K}\beta(z_k)$) is deviated from its optimum value during the optimization of each $\beta(z_k)$, the waveform collapses and the cost function rapidly diverges. For the $\gamma'(z_k)$ optimization, this slow convergence does not occur since the nonlinearity is not the dominant factor for the cost function. In Fig. 8(b), we plot the maximum value of $\beta(z_k)$ in the DSF span. The case with the second span replaced with NZ-DSF is also shown. Since the convergence was slow, we stopped the iterations at 118000 but the maximum values are still increasing, so $\beta(z_k)$ may find the theoretical values with more iterations.

Fig. 9(a) shows the estimated $\beta(z_k)$ profiles with a DSF or NZ-DSF inserted in each span. The iterations were stopped at 10000. For all cases, peaks are again observed in spans with a DSF or NZ-DSF. Even flat areas of each line (i.e., SSMF spans) indicate different dispersion values for different span configurations. This is because the total dispersion varies among tested scenarios with different non-SSMF-span locations. Thus, to see the difference between DSF and NZ-DSF more clearly, we plot in Fig. 9(b), (c), (d), and (e) the relative $\beta(z_k)$ in each span with respect to the minimum value over the whole link. Note that two $\beta(z_k)$ profiles for two different captured data are shown. The peaks with NZ-DSF are smaller than those with DSF, indicating that the NZ-DSF spans have higher dispersion. However, the NZ-DSF in the second span and DSF in the third span show similar peaks, which implies that distinguishing NZ-DSF will be difficult without the reference in the same span. Furthermore, the fourth span exhibits a far higher peak than other spans. Though these span dependencies should be clarified in future work, at least the presence of DSF or NZ-DSF spans can be distinguished from SSMF spans, which is sufficiently convenient for a practical system with SSMF-dominant spans since it enables us to take such measures as fiber launch power adjustment or a DSF-span detour to maximize the transmission capacity.

B. *Performance Evaluation: Long-Haul Links*

Because CRM relies on fiber nonlinearity for power profile extraction, a higher fiber launch power is desirable to increase NLNR (Eq. (8)) and thus the accuracy; this is prohibitive in practical systems because increased fiber nonlinearity degrades the SNR. Therefore, the launch power dependency of the accuracy and transmission distance should be evaluated. To address this point, we further conducted long-haul transmission using a circulating loop in which the tested distance ranged from 260 (one circulation) to 2080 km (eight circulations). Fig. 10 shows the setup. In this experiment, no DSF was inserted, and a dispersion profile was not obtained, which we leave for future work. The basic configuration of the transmitter and receiver is the same as in the previous subsection. Uniform 16-QAM signals are generated and then chopped with an acousto-optic modulator (AOM1) with a passing period of one circulation (260 km). The transmission line consists of five spans of SSMF (α = 0.186

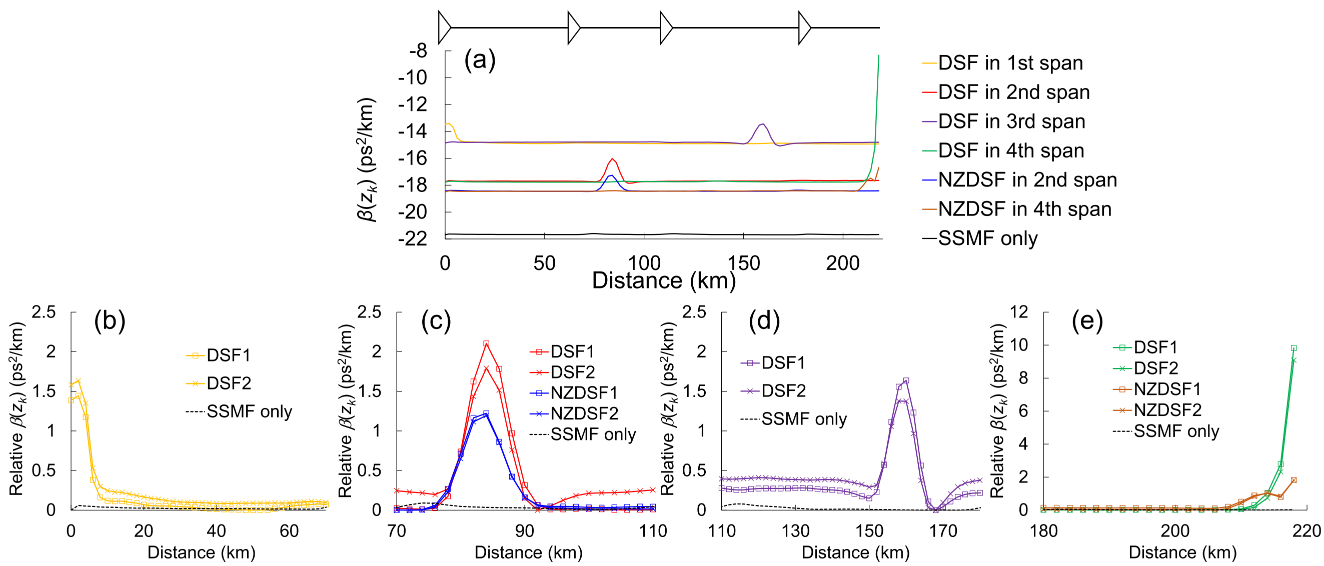

Fig. 9. (a) Obtained $\beta(z_k)$ profiles for four-span links with DSF and NZ-DSF spans mixed in SSMF-link. Relative values from the minimum $\beta(z_k)$ when DSF and NZ-DSF spans are inserted in (b) first, (c) second, (d) third, and (e) fourth span of link.

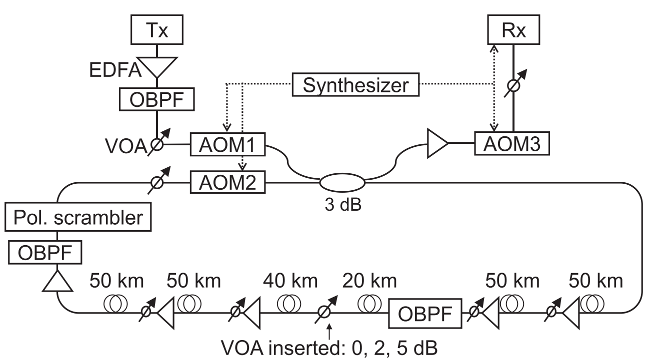

Fig. 10. Experimental setup for circulating-loop transmission. One circulation corresponds to five spans (260 km) of SSMF. VOA is inserted 20 km from beginning of third span to emulate anomaly loss (0, 2, and 5 dB). AOM: acousto-optic modulator.

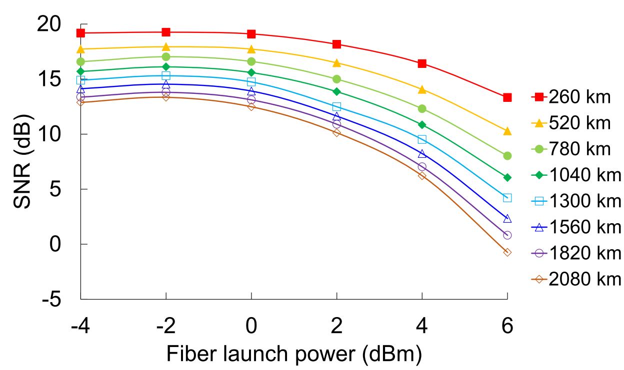

Fig. 11. Launch power dependency of SNR for various transmission distances.

dB/km, $\beta_2 = -21.7$ ps^2/km), and a variable optical attenuator (VOA) is inserted 20 km from the beginning of the third span to emulate an anomaly loss. The fiber launch power was varied by VOAs from -4 to 6 dBm. OBPFs cut the ASE noise and flatten the gain tilt. After a loop-synchronized polarization scrambler, AOM2 opens the gate for the duration of an arbitrary number of circulations and closes it while the signals are newly input to the circulating loop. AOM3 selects the desired circulated signals and cuts instantaneous surges from the EDFA, and a 256-GSa/s DSO captures the signals, synchronized to AOM3. For the channel emulator, we assumed that the transmitted signal was known as the reference signal for gradient descent because, in the long-haul case, an input power of 6 dBm is too high to reach the FEC limit due to the strong fiber nonlinearity. The mini-batch size was set to 20.

We first find the optimum launch power of the system in terms of SNR to see the true operation point (Fig. 11). Here, the SNR is calculated after CPR in the standard demodulation process in terms of the mean square error between the demodulated and reference signals. The optimum launch power of the system is about -2 to 0 dB regardless of all transmission distances. Note that this experiment is a single-channel transmission and that the expected optimum power for practical WDM systems will be lower.

With this result in mind, we obtained the power profiles for 1040-km transmission (four circulations) for launch powers of 6, 0, and -2 dBm, as shown in Fig. 12(a), (b), and (c), respectively. The attenuation by the VOA at 20 km from the beginning of the third span was varied among 0, 2, and 5 dB. One circulation corresponds to five spans (260 km), and the third span of each circulation exhibits loss due to the VOA for all three cases. As the launch power decreases, the profiles' agreement with OTDR clearly degrades since NLNR in (8) degraded. Nevertheless, for all three cases, the insertion of the attenuator was successfully observed, which supports the applicability of CRM for practical systems. However, the second span (at 50 km from the transmitter) can be recognized as an anomaly span if the "normal" 0-dB attenuated profile is not available. This leads to a false

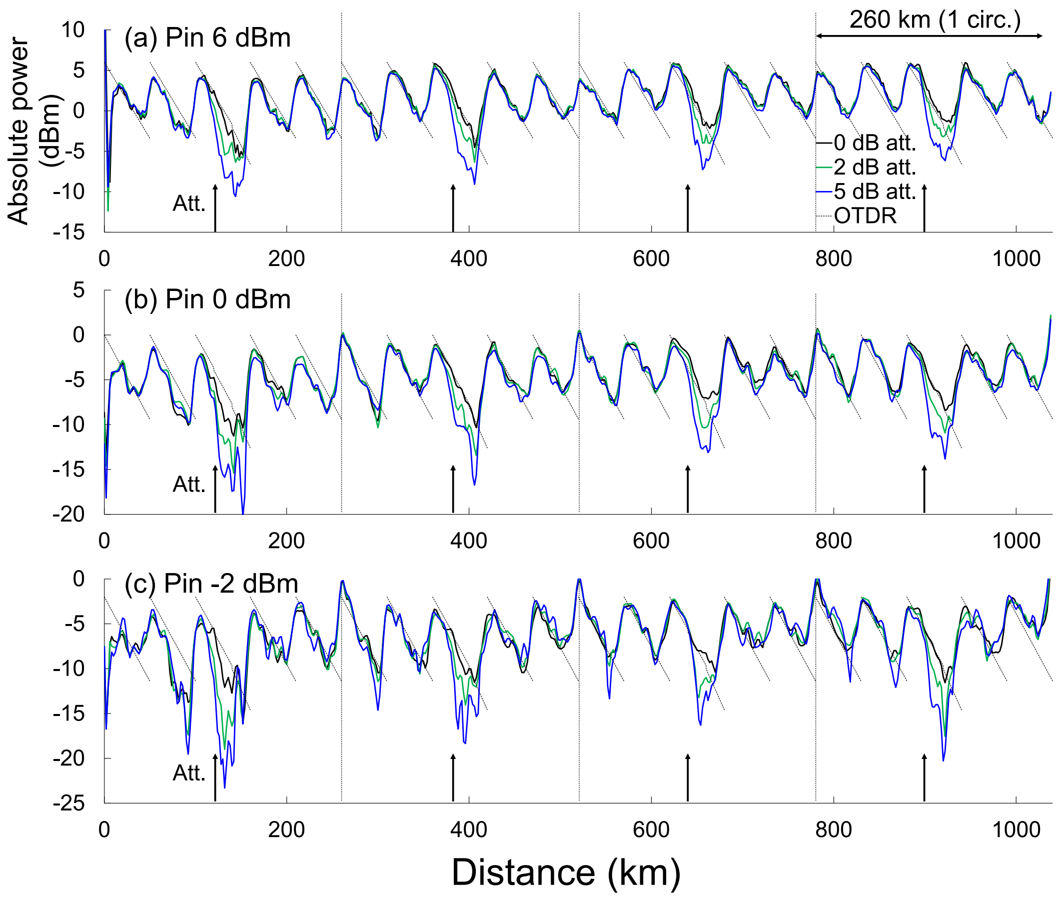

Fig. 12. Obtained power profiles for 1040-km transmission with fiber launch powers of (a) 6, (b) 0, and (c) −2 dBm. Attenuators with 0-, 2-, and 5-dB losses were intentionally inserted at 120 km (at 20 km in third spans) of one circulation (5 spans, 260 km).

detection of an anomaly span, and thus either the performance improvement of the obtained profiles or a premeasurement of a normal profile is needed. In the case of the latter, we need to check that a targeted transmission system contains no anomaly before the operation starts by using other measures such as OTDR.

Farther transmission distances bring another problem. Fig. 13 shows the power profiles for 2080-km transmission (eight circulations) for a launch power of 2 dBm. Though the inserted attenuation of 0, 2, and 5 dB is observed, as the DBP steps approaches the transmitter end (from 0 to 700 km), the estimated power becomes lower than expected. We can explain this drop as the suppression of noise enhancement. As signals backpropagate to the transmitter end with cascaded DBP operations, the noises accompanying the signals, such as ASE noise, the incompletely compensated Kerr nonlinear term, and even Rx electrical noise, are enhanced (N of NLNR in (8) is increased). As a result, the optimized $\gamma'(z_k)$ near the transmitter ($z_k \approx 0$) tends to be small to avoid further noise enhancement by NLPR, and thus, a weak $\gamma'(z_k)$ is observed. A similar explanation can be seen in [12] and the tendency of the estimated $\gamma'(z_k)$ in [13] matches this result though they focused on NLC with 1-step-per-span DBP. Another reason for the reduced accuracy in the long-haul case is that the number of optimized parameters $\gamma'(z_k)$ is increased for a longer transmission distance (here, Δz is a constant 2 km), making parameters more likely to fall into local minima.

Next, we evaluate the launch power dependency of the anomaly loss detection. Fig. 14(a) and (b) show the difference between intentionally attenuated profiles and the 0-dB attenuated profile for launch powers of 2 and 0 dBm, respectively. For reference, theoretical lines calculated from OTDR profiles are also shown. Again, note that the inflection points of the rising part of the deviation correspond to the inserted anomaly position. Here, we assume the inserted attenuations are estimated as deviation peaks. The inserted loss was clearly observed for a launch power of 2 dBm. However, in the case with 0-dBm launch power, the deviation profiles look noisy and several 2-dB peaks in non-attenuated area (i.e., false detections) are observed. To quantify these false detections, we show histograms of deviated values from 0-dB profiles *except for attenuated locations* in Fig. 14(c)−(f). The peak values in attenuated locations are shown as solid lines. Green and blue lines correspond to the case with 2-dB and 5-dB attenuations, respectively. In both Fig. 14(c) and (d) (a launch power of 2 dBm), the 99% point of the deviations is around 1 dB, which implies false detections of 1-dB attenuation are likely to occur. Since the peak values are higher than 1 dB, we can at least distinguish the inserted attenuations. However, when the launch power decreases to 0 dBm, we cannot

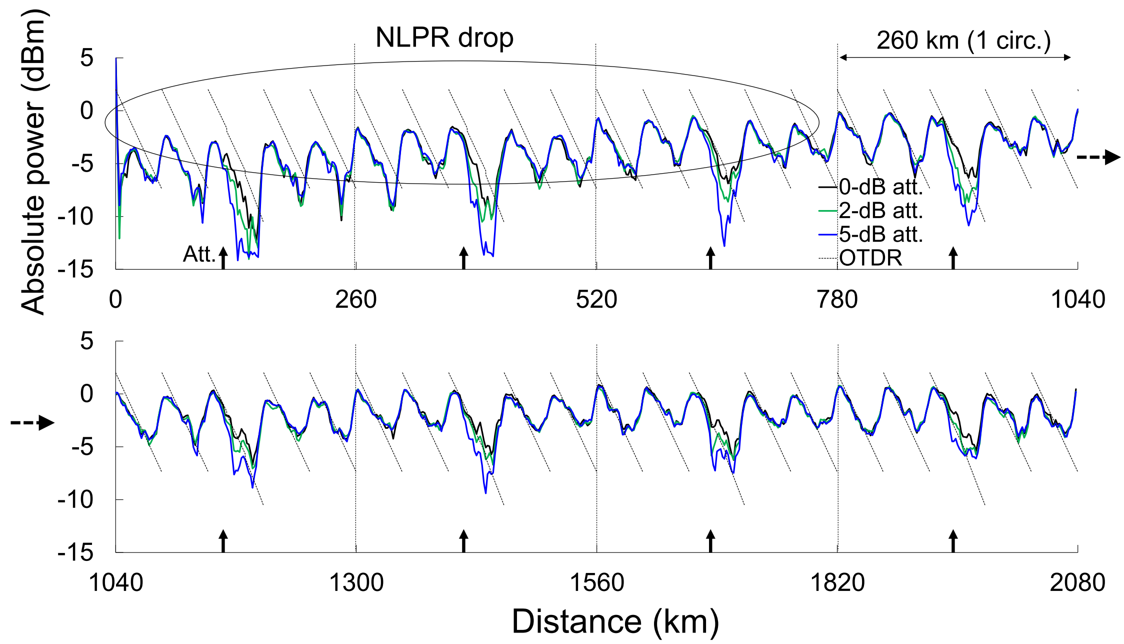

Fig. 13. Obtained power profiles for 2080 km transmission with fiber launch power of 2 dBm.

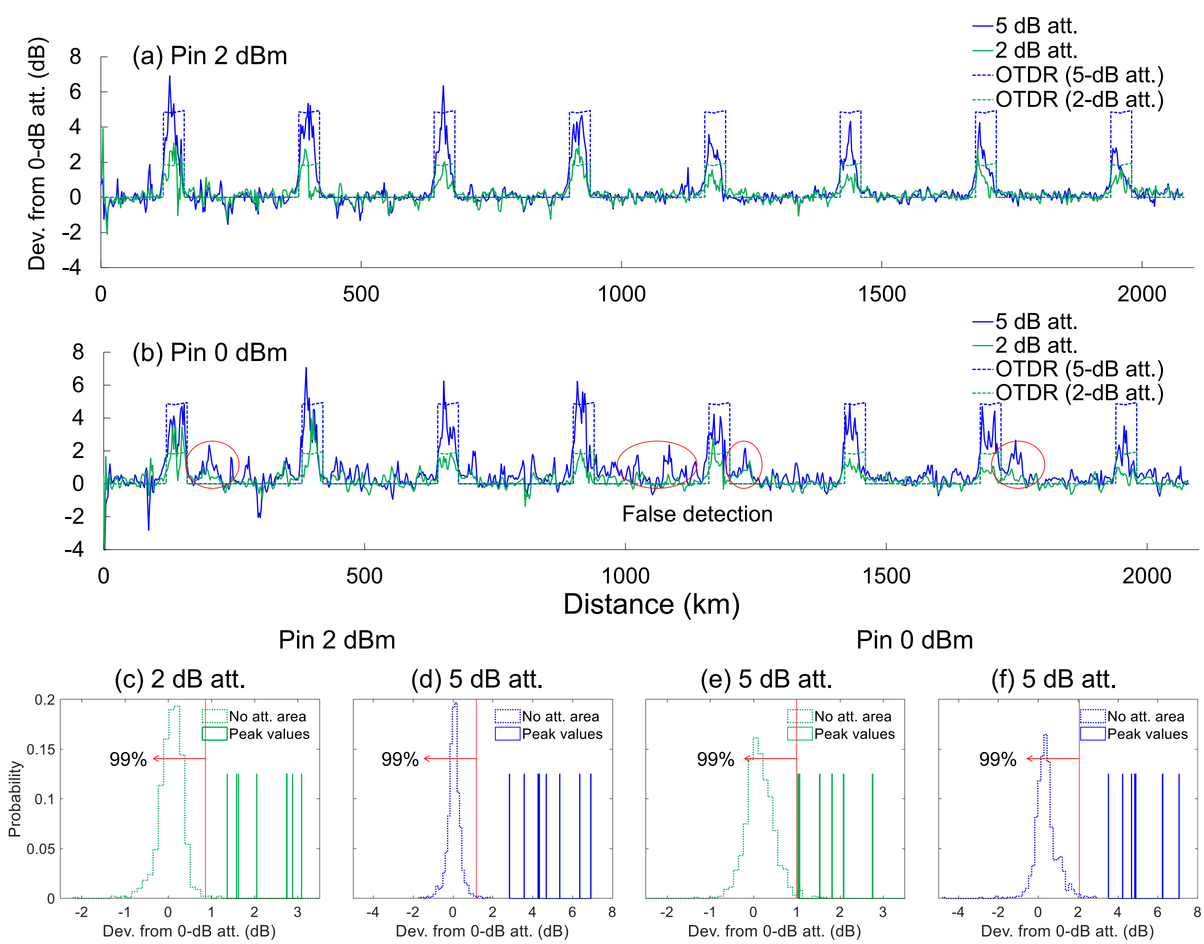

Fig. 14. Differences between normal (0-dB loss) and anomalous (2- and 5-dB losses) state profiles for 2080 km transmission with fiber launch powers of (a) 2 and (b) 0 dBm. Histograms of difference between normal and anomalous state except for attenuated spans (dashed line) for fiber launch power of 2 dBm with inserted attenuations of (c) 2 and (d) 5 dB. Peak values in attenuated spans (solid lines) are also shown. Same histograms for fiber launch power of 0 dBm with inserted attenuations of (e) 2 and (f) 5 dB.

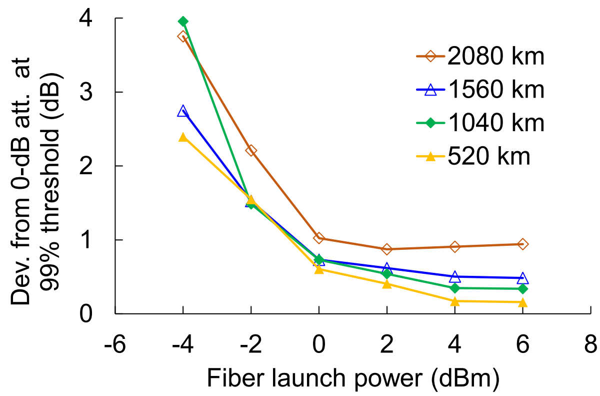

Fig. 15. 99% point of histograms for 2-dB inserted attenuation as a function of fiber launch power.

trust 2-dB deviation peaks in the following points. First, when 2-dB attenuators are inserted (Fig. 14(e)), their minimum peak values are 1 dB, which overlaps the histogram. Second, with 5-dB attenuators (Fig. 14(f)), the histogram becomes wide and the 99% point is around 2 dB. This comes from the fact that the nonlinearity is too weak due to the launch power reduction and the large attenuations. Thus, true 2-dB attenuations cannot be distinguished from false detections, while 5-dB attenuations can.

To quantify the fiber launch power required to detect a certain amount of inserted loss, we plot the 99% point of the histograms as a function of fiber launch power for various transmission distances (Fig. 15). The inserted attenuations are fixed to 2 dB. As we have already observed, the 99% point (false-detected attenuation) increases as the fiber input power decreases, which means the deviation profiles become unstable like Fig. 14(b). Also, as the transmission distance increases, the instability grows because of the accumulated link noise, noise enhancement suppression in CRM, and the increased number of optimization parameters. From this figure, we can determine the minimum detectable anomaly loss for a given transmission distance and a fiber launch power. For example, suppose that our targeted system is a 1040-km link with an operated launch power of -2 dBm. In that case, the 99% point of the deviation is around 1.5 dB. Assuming the minimum deviation peaks lie approximately 1/2 of the actual anomaly loss (see peak values in histograms), the detectable loss is 3 dB or higher. From another point of view, if we should detect at least 2-dB loss, then the 99% point should be around 1 dB, and thus a -2-dBm launch power is prohibitive for any transmission distances. For a 0-dBm launch power, we can achieve a measurement range of 1560 km. If more input power is allowed (e.g., 2 dBm or higher), then a 2080-km measurement range is achievable. However, the performance improvement from powers of 0 to 6 dBm is small for longer distances, though the difference between launch powers of -2 and 0 dBm is significant. This indicates that the 0-dBm fiber launch power is already close to the maximum performance for long-haul applications.

We note that in this evaluation the attenuated position is fixed but as we observed in Fig. 7(b), the sensitivity to inserted attenuation can vary with inserted positions, which requires a more comprehensive study of performance prediction. We also note that these experiments were conducted in a single-channel scenario. In actual WDM systems, inter-channel nonlinearity (i.e., cross-phase modulation, XPM) exists and will work as a stochastic noise source, which degrades the estimation precision if no information on adjacent channels is available. We note that in [16], the power profiles were successfully obtained in a WDM scenario even though a correlation method was used. We also note that, if our backpropagation includes not only the targeted channel but also adjacent channels (i.e., multi-channel DBP [38]), XPM, which has so far been a noise source, will turn into a source to estimate the total power of the entire signal field in the fiber, and thus the estimation performance will be improved.

C. Optical Level Diagram Extraction

For the results discussed so far, we have assumed that the fiber launch power is constant for all spans. However, in a realistic system, the launch power of each span is usually set to different values in accordance with span lengths to maximize the total SNR. If profiles obtained by CRM can automatically detect an anomalous input power for a span (i.e., non-uniform optical level diagram), then we can remotely adjust the input power of such a span without direct measurement. Fig. 16(a) shows power profiles with different fiber launch powers (-4 to 6 dBm) at the third span of the circulating loop for a 2080-km transmission. The input powers of the other spans are uniformly set to 2 dBm. The profiles clearly exhibit a non-uniform fiber input power at the third span of every five (one circulating loop), and the power at that span can be distinguished, while the power at the other spans is at the same level in all profiles. Fig. 16(b) shows the difference from the profile with uniform 2-dBm launch powers (light green). Dashed lines are theoretical values, and we can observe that the profile differences match these theoretical lines for larger launch powers. This indicates that CRM can reveal the optical level diagram and the input power of each span from the Rx DSP. However, near the transmitter side, the estimated power is lower (Fig. 16(a)) and the large gaps from theoretical lines are observed for smaller launch powers (Fig. 16(b)). These reduced sensitivities can be explained as the consequence of a lower NLNR (nonlinearity to noise ratio) in (8) with smaller NL and larger N. As we mentioned in the previous subsection, DBP-enhanced noise increases N near the transmitter side and estimated $\gamma'(z_k)$ tends to be small to avoid further noise enhancement. This comes from the fact that CRM is based on the MMSE criterion. On the other hand, as the launch power increased, such a drop diminished (e.g., red line). This is because NLPR in those spans is strong enough to overcome the above noises, and thus the receiver-side estimation by CRM can correctly detect the NLPR. These results imply that "NLNR" in (8) determines the estimation performance.

We note that this experiment is a particular test case with a single span in a circulation having different power and a different span length. In a more realistic scenario, each span usually has a different length and thus a different launch power. Though these scenarios are beyond the scope of this paper, we

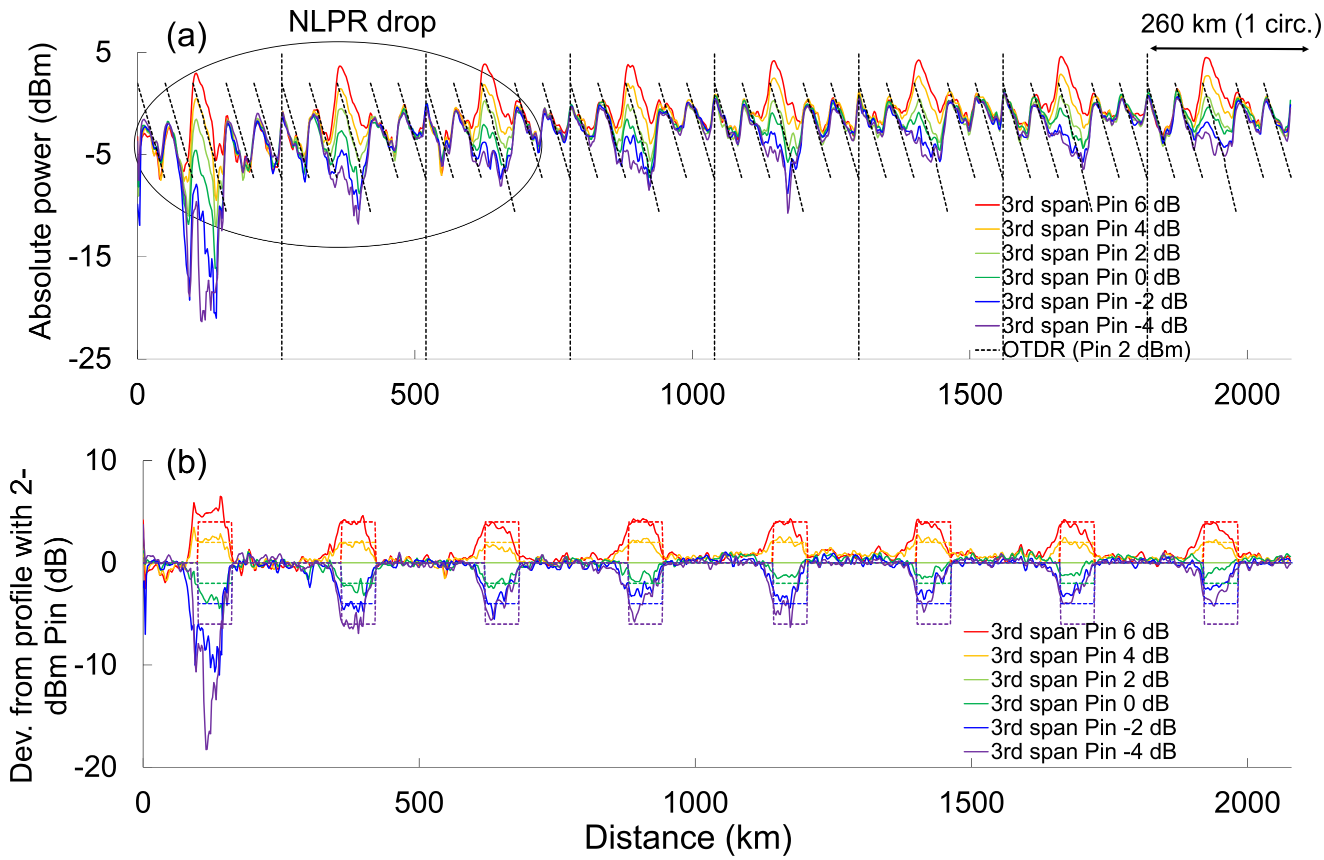

Fig. 16. (a) Non-uniform optical level diagram for 2080-km transmission. Launch power of third span of circulating loop was intentionally varied from −4 to 6 dBm, and that of other spans was uniformly set 2 dBm. (b) Difference between the non-uniform profiles and uniform profile (third span 2-dBm launch power).

can discuss that the transmission scenario that degrades the estimation precision of non-uniform $\gamma'(z_k)$ profiles would be when the power difference between the spans with the highest and lowest launch power is large (such as when long and short spans coexist in a transmission line). In such a case, the global error (i.e., the cost function $\|E(0,nT) - \hat{E}(0,nT)\|^2$) is mainly determined by the highest power span's NLPR. In other words, the change in $\gamma'(z_k)$ in the lowest span has less impact on the global error. Thus, it would be difficult to correctly estimate the $\gamma'(z_k)$ profiles in the lowest power span though a detailed evaluation is required.

D. Applications

1) *At Beginning of System Operation*: When we build a transmission system, all link components must be tested in advance. For this purpose, the operator needs to prepare testing devices such as OTDR and a dispersion analyzer and optimize launch powers for every span, increasing CAPEX and OPEX. The proposed CRM will facilitate this procedure since it extracts multi-span fiber characteristics at once. CD profiles are helpful when establishing a network with unknown fibers (e.g., dark fibers) and identifying the fiber type (SMF or DSF, etc.) and its span. For example, we can lower the fiber input power of identified DSF spans to avoid fatal nonlinearity or can even detour such a fiber to maximize the transmission capacity (if C-band transmission is assumed). Power profiles reveal multi-span optical level diagrams. We can perceive anomalous fiber launch power spans and can adjust them without direct measurement, enabling automated launch power optimization. However, CRM requires a high fiber launch power for the enhanced performance. Furthermore, without a normal state reference, the anomaly loss detection can be challenging. These issues should be overcome in future work.

2) *Under Operation*: During system operation, the fiber conditions and signal power should be monitored to check for physical damage to fibers. Usually, power monitors can be placed at the end of spans for this purpose, but this does not localize the anomaly position. Furthermore, it requires the signals to be branched off, which leads to additional loss. CRM can identify lossy points with a normal state reference and thus accelerate the cause analysis and link restoration without suffering from the additional loss since power profiles are extracted directly from received digitized signals. However, the required fiber launch power must be improved for CRM to work around the optimum launch power of a system.

3) *Fiber Nonlinear Noise Estimation*: The amount of fiber nonlinear noise can be estimated by feeding power profiles to an analytical model, such as a Gaussian noise (GN) model [39], [40] or the corresponding Python application, GNPy [41].

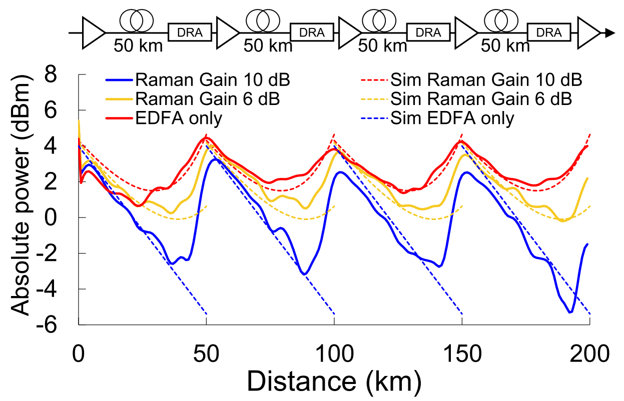

Fig. 17. Obtained Raman-amplified power profiles for 50 km × 4 spans with different Raman gains.

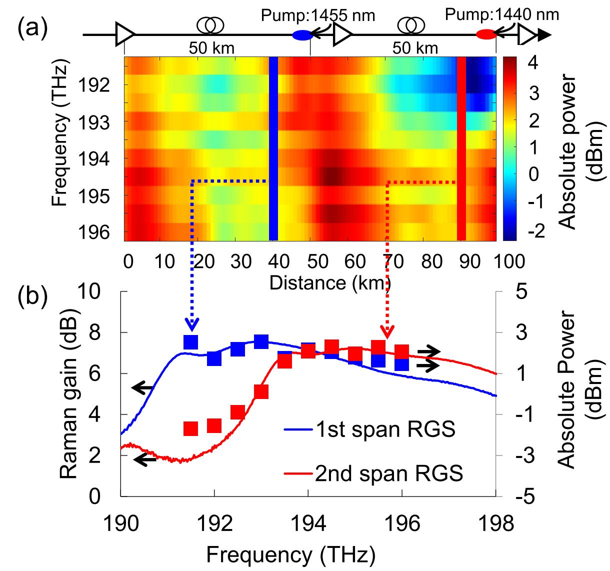

Fig. 18. (a) Frequency dependency of Raman-amplified power profiles for 50 km × 2 spans. Raman gain spectra of two backward Raman amplifiers at nodes are intentionally set differently by varying pump wavelength. (b) Estimated Raman gain spectra (square plots) reconstructed from frequency dependency of estimated power at 40 and 90 km. For reference, Raman gain spectra captured by an optical spectrum analyzer are shown (solid lines).

This application is more suitable for DRA systems as will be mentioned in Section IV-C.

IV. Optical Amplifier Gain Spectra Reconstruction

In the previous section, we observed that obtained profiles reflected the optical level diagram and thus the different amplifier gains. In this section, it is shown that CRM can separately reconstruct the gain spectra of multiple cascaded optical amplifiers in multi-span transmission links. This study aims to identify anomaly amplifiers, which have a narrower, non-flat, weaker or stronger gain spectrum.

Gain spectra are acquired in two steps: obtaining a signal power profile for all WDM channels and observing the obtained powers in the frequency direction at the amplifier position. The method is applicable to any types of amplifiers since the estimation mechanism does not rely on the amplification process but only on the signal power itself. Applicability to EDFA has partially been shown in the previous section. Thus, we here focus on DRA systems because the method can effectively offer several benefits for DRA systems. These points are discussed in Section IV-C.

A. Raman-Amplified Power Profiles

The experimental setup is basically the same as in Section III-A, except that the span amplifiers are replaced with hybrid Raman and EDFA. Details on the transmission lines are directly shown in the result figures.

First, we show the Raman-amplified power profile for multi-span links and investigate the Raman gain dependency. Fig. 17 shows the estimated power profiles for 50 km × 4 spans. The Raman on-off gain was varied among 0, 6, and 10 dB. The signal frequency was set to 192.7 THz (1555.752 nm), the Raman pump wavelength is 1455 nm. The maximum pump power was 520 mW at a 10-dB gain and the span loss was 9.39 dB, and thus, the 10-dB gain DRA fully compensates for the span loss. The fiber launch power was set to 4 dBm. The solid line is the estimated power, while the dashed line is the simulated power [42], which is calculated from the fiber launch power, fiber loss, span length, and gain. As the Raman gain increases, the power in the latter half of the span grows, which indicates that Raman-amplified profiles are successfully observed. The estimated profile fits the simulated line well. In the 10-dB gain case, in-span symmetry is observed, so nonlinear mitigation techniques such as optical phase conjugate (OPC) are expected to perform well if employed [43].

B. Raman Gain Spectra Reconstruction

Next, to confirm whether CRM can separately obtain two different gain spectra of cascaded amplifiers, we obtained the frequency dependency of the signal power profiles in a 50 km × 2 span link (Fig. 18(a)). The tested frequency is the entire C-band from 191.5 to 196.0 THz with a 0.5-THz granularity. To test the capability of distinguishing two different RGSs, the pump wavelengths of two DRAs were intentionally varied (1455 and 1440 nm for the first and second spans, respectively). Their RGS is shown by a solid line in Fig. 18(b), which was acquired with an OSA with a 1-nm resolution. Note that, for a fair comparison, these reference spectra are recalculated from the raw data obtained from the OSA using the simulated profiles (in Fig. 17) so that their gains correspond to 10 km before the span ends. The pump powers of both spans were fixed to 520 and 482 mW, respectively. In Fig. 18(a), a high signal power area is shown in red, and as the color turns blue, the power decreases. As expected, the latter half of each span shows different powers since two spans have different RGSs. To look closely into the frequency dependence, we extracted the signal powers at 40 and 90 km and plot them with squares in Fig. 18(b). We avoided choosing 50 and 100 km since the powers at these positions reflect the next fiber input power, which does not vary with frequency. The estimated signal power shows good consistency with the actual RGS, which means that the RGS can be reconstructed via the obtained signal power profile. Furthermore, two different gain spectra are successfully separated. Thus, without testing instruments like OTDR and OSA, we can obtain multi-span power evolution and even individual RGSs from only the signal of interest with DSP.

C. Applications

1) *Health Monitoring of Gain*: Since the method can separately reconstruct different RGSs of cascaded amplifiers, we can identify anomaly amplifiers with undesirable gain spectra, which would come from the degradation of pump lasers due to aging and temperature variation. Furthermore, since the gain spectra and fiber launch power for all WDM channels can be monitored, automatic adjustment of gain tilt or flatness using WSS can be expected for maximizing WDM system capacity. Usually, in ROADM systems, optical channel monitors (OCM) are implemented at ROADM nodes, which serve as OSA and monitor the output spectra of nodes. However, OCM requires signals to be tapped from the main transmission line and thus leads to component loss and optical SNR degradation. CRM potentially eliminates OCM and thus can avoid such loss and save on CAPEX.

2) *Fiber Nonlinear Noise Estimation*: In DRA, the fiber nonlinear noise is a more critical quantity since the amplified in-line signal power accompanies more nonlinear noise. By combining the GN model [39] and the estimated power profile from CRM, nonlinear noise can be estimated. OTDR generally does not work for power profile extraction in DRA because backward Raman scattering acts as forward Raman scattering for backscattered Rayleigh light. Although modified OTDR can measure profiles [44] and RGS [45], these measurements require pulses to be set near the signal wavelength to measure the gain for the signal of interest, and to be wavelength-swept for RGS measurement, which is not practical in WDM systems.

3) *Power Profile Symmetry Monitoring*: The power profile symmetry in a span enables the nonlinear mitigation performance of OPC to be maximized [43]. Ideally, such symmetry must be maintained for all WDM channels. Thus, monitoring power profiles and RGS with CRM will ease the performance maximization of such nonlinear mitigation techniques.

V. Optical Filter Response Separation

In this section, we show that CRM can separately estimate optical filter responses and thus identify the center frequency detuning of individual filters. In dense WDM (DWDM) systems, where less guard band is left, the signals are vulnerable to PBN due to the misalignment or bandwidth deviation of optical filters (e.g., WSSs) in ROADM systems [28]. Recent WSSs can tune their bandwidth and central frequency to as fine as a few GHz [46], and the PBN penalty can be minimized if these parameters are tailored to the incoming signals. However, it is challenging to identify a frequency-detuned filter among multiple cascaded filters and estimate its frequency shift and shape because concatenated linear systems (the overall responses of two optical filters) are not separable. CRM utilizes the fiber nonlinearity in between optical filters so that the overall system becomes a concatenation of linear and nonlinear systems and thus noncommutative. Thus, we can distinguish two filter responses.

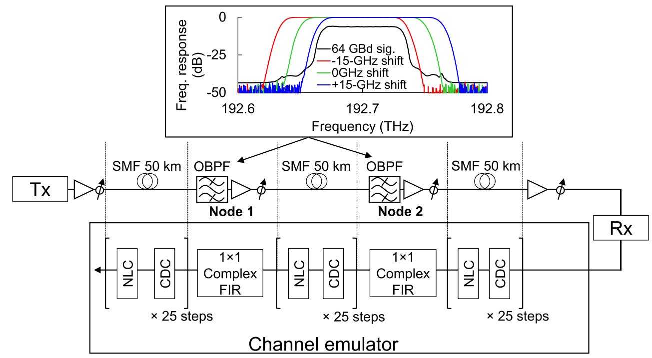

Fig. 19. Experimental setup and channel emulator for detecting PBN locations. Complex FIR filters that emulate PBN at OBPFs in nodes are inserted in channel emulator.

A. Multiple Filter Response Identification

First, we show the estimated whole inverse system characteristics of a 50 km $\times$ 3 span SSMF transmission link. Fig. 19 shows the experimental setup. Again, the transceiver setup and signal processing flow are basically the same as in Section III-A. As depicted, each node had an optical filter, and their center frequencies were varied from –20 to +20 GHz. Their 3-dB bandwidth was set to approximately 100 GHz, which covers the signal bandwidth (64 GBd with a roll-off factor of 0.2). No intentional attenuation was inserted, and all fiber launch powers were set to 10 dBm in this experiment. The channel emulator consists of three spans of 25-step DBP (thus $\Delta z = 2$ km) with two FIR-filter blocks $\mathbf{h}_1$ and $\mathbf{h}_2$ inserted at positions that correspond to the first and second nodes, respectively. Each filter had 201 taps. In this experiment, $\beta(z_k)$ was fixed to the total CD/total DBP steps. Initial values for $\gamma'(z_k)$ and $\mathbf{h}$ were set to 0 and the impulse function, respectively.

Fig. 20 shows the identified power profile and identified inverse frequency responses of optical filters. The estimated power profile and the OTDR reference are shown in Fig. 20(a). Fig. 20(b) and (c) show the responses of $\mathbf{h}_1$ and $\mathbf{h}_2$ for the case when the center frequency of the optical filters in the first and second nodes were shifted by +15 GHz, respectively. All responses are shown for two-sample/symbol operation. In Fig. 20(b), the first FIR filter $\mathbf{h}_1$ (blue) has a peak that compensates for the PBN, while the second one $\mathbf{h}_2$ (red) exhibits a flat response, which clearly means that the PBN occurred in the first node, not in the second node. The same tendency can be observed in Fig. 20(c), where only $\mathbf{h}_2$ (red) has a peak, while $\mathbf{h}_1$ (blue) does not, and thus, we can tell that the PBN occurred only in the second node. Thus, the two different optical filter responses were successfully separated, and we could detect the anomaly filter among multiple ones. By monitoring these responses, we can modify or optimize the center frequency of the analog optical filter settings, and thus avoid excess filtering penalties.

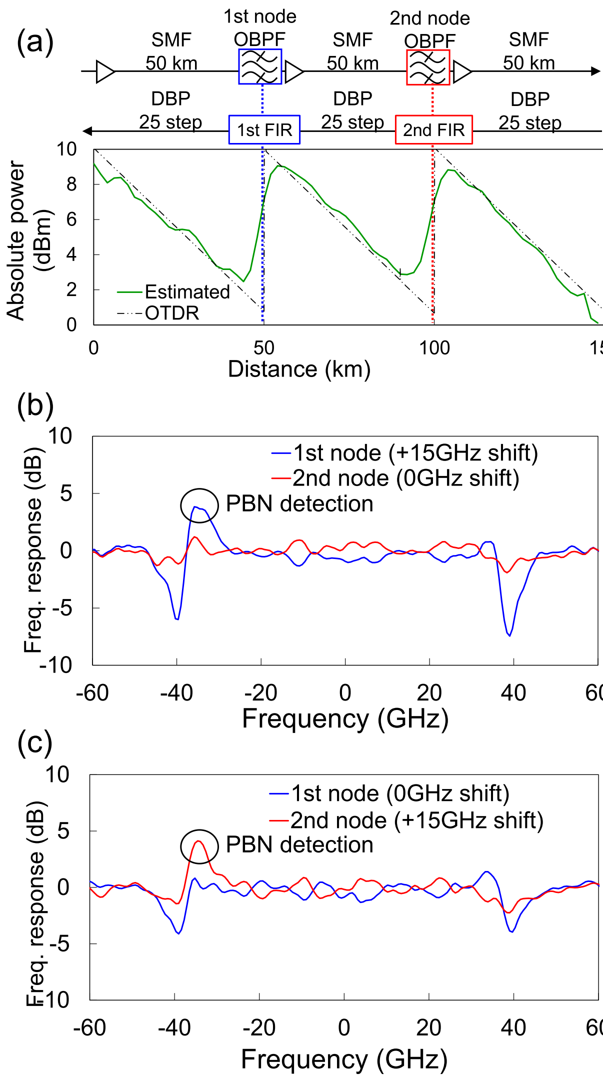

Fig. 20. (a) Obtained profiles for 50 × 3 spans. Estimated frequency response of FIR filters $\mathbf{h}_1$ and $\mathbf{h}_2$ when only (a) first and (b) second OBPF have center frequency shifted by +15 GHz.

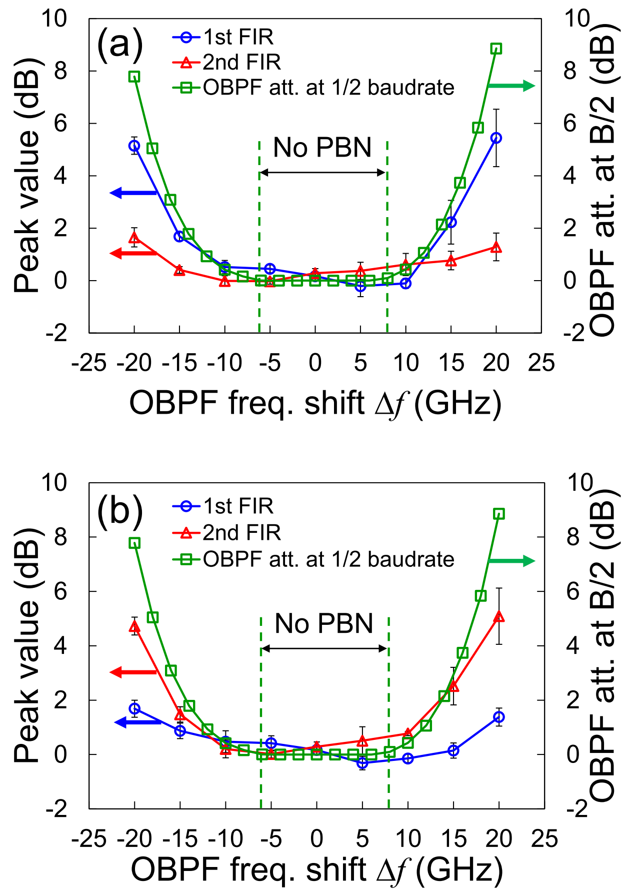

Fig. 21. Peak values of frequency responses of first FIR filter (blue) and second FIR filter (red) and measured attenuation level at half baudrate (green) when (a) first and (b) second node of OBPF were intentionally frequency-shifted by Δf and other node had 0-GHz shift. ($n = 5$).

B. Shifted Frequency Estimation

From the estimated FIR filters, the shifted frequency Δf of the optical filters can also be estimated since the attenuation at the edge of signal spectrum due to PBN $\alpha_{PBN}(\Delta f)$ corresponds to the peak in frequency responses $\max(|F[\mathbf{h}]|^2)$. We measured the optical attenuation level of OBPF at a half baudrate (B/2) as a function of Δf and compared it with the peaks of the estimated frequency responses of $\mathbf{h}$. Fig. 21 shows the peak values ($\max(|F[\mathbf{h}_1]|^2)$ and $\max(|F[\mathbf{h}_2]|^2)$) (blue, red) and $\alpha_{PBN}(\Delta f)$ (green) as a function of Δf. Note that, since the estimated responses were fluctuating as seen in Fig. 20, a low-pass filter was applied to $F[\mathbf{h}]$ before plotting the maximum values to reduce the fluctuation. Fig. 21(a) and (b) correspond to when only the first and second OBPFs were shifted, respectively. In both cases, no attenuation was observed for a shift from -6 to 8 GHz since the filter bandwidth is approximately 100 GHz and no PBN occurred in that region. As Δf increases, the measured attenuation level at B/2 grows due to PBN. When only the first OBPF was shifted (Fig. 21(a)), the peak values for $\mathbf{h}_1$had a good agreement with the measured loss. On the other hand, $\mathbf{h}_2$ and the measured loss are less correlated, which implies the successful discrimination of the anomaly and normal OBPFs. In Fig. 21(b), where only the second OBPF has a frequency shift, the peak values for $\mathbf{h}_2$ also have the same tendency as the measured loss, while the peak values for $\mathbf{h}_1$ show a flatter characteristic over Δf. When Δf was -6 GHz $< \Delta f <$ 8 GHz, where no PBN occurred, the two lines (red and blue) showed reversed characteristics. For example, when the peak values for $\mathbf{h}_1$ were smaller, those for $\mathbf{h}_2$ in turn exhibited a higher peak. This is because the total amount of compensation of two FIR taps must be 1 (0 dB) when there is no PBN. Thus, $\mathbf{h}_1$ and $\mathbf{h}_2$ try to balance each other so that their sum becomes constant.

VI. Conclusion

As a digital longitudinal monitoring (DLM) technique, we demonstrated a channel reconstruction method (CRM) that extracts multi-span fiber loss and CD profiles, multiple amplifiers' gain spectra, and multiple optical filters' responses only from data-carrying signals. From boundary conditions, CRM reconstructs the whole optical fiber link in the digital domain, where optical fibers and amplifiers are modeled as SSFM while optical filters are emulated by complex-valued FIR filters. Consequently, anomaly loss, unexpected CD, insufficient gain, and PBN were detected and localized. For a proof of concept, we observed the absolute power profiles ($\propto \gamma'(z_k)$) estimated by CRM matched theoretical values well for a 4-span metro-reach link. In our performance evaluation for a long-haul scenario ($\sim$2080 km), we observed that a reduced signal power and more channel noise degraded the estimation precision of $\gamma'(z_k)$, which implies "the nonlinearity to noise ratio (NLNR)" plays a large role in determining the performance of CRM. Also, the minimum detectable anomaly loss was discussed. Regarding the estimation of $\beta(z_k)$, CRM clearly detected anomaly fiber types (DSF or NZ-DSF) inserted in a SSMF link, though its convergence was slow and DSF and NZ-DSF are currently challenging to distinguish without reference profiles in the targeted span. As for optical amplifiers, we applied CRM to multi-span DRA systems and found that the algorithm successfully reconstructed two different DRA gain spectra at cascaded nodes at the same time. Finally, CRM separately estimated two different optical filter responses, which enabled the localization of anomaly filters with PBN. Unlike conventional analog approaches such as OTDR and OSA, DLM digitally characterizes multiple components in multi-span links at once, which reduces the CAPEX and OPEX of network operation. By further enhancing its accuracy and sensitivity, CRM will facilitate smart and cognitive optical transmission systems, in which system design, establishment, and management will be autonomously conducted for reliable and capacity-maximized networks.

Appendix
Gradient Calculation of CRM

Gradient calculation of FFT, IFFT, CDC, NLPR, and complex-valued FIR-filter blocks are described here in accordance with Fig. 5.

In the *k*-th linear CDC block, linear operations are implemented with FFT and IFFT, i.e., $\mathbf{y} = \mathrm{ifft}[\mathrm{fft}[\mathbf{x}] \cdot \exp(-\frac{j}{2}\beta(z_k)\omega^2\Delta z)]$, where $\mathbf{x} = [\mathbf{x}_h, \mathbf{x}_v]$ and $\mathbf{y} = [\mathbf{y}_h, \mathbf{y}_v]$ denote the input and output vectors of a block with *h*- and *v*- polarizations. As shown in Fig. 5(c), the backpropagation of the FFT is equivalent to the IFFT and vice versa. Also, the backpropagation of the frequency-domain CDC and the partial derivative for $\beta(z_k)$ update is calculated as (Fig. 5(d)):

$$\begin{aligned}\frac{\partial I}{\partial \bar{\mathbf{x}}} &= \frac{\partial I}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \bar{\mathbf{x}}} + \frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \\ &= 0 + \frac{\partial I}{\partial \bar{\mathbf{y}}} \cdot \exp\left(\frac{j}{2}\beta\left(z_k\right)\omega^2\Delta z\right), \end{aligned} \tag{10}$$

$$\begin{aligned}\frac{\partial I}{\partial \beta} &= \frac{\partial I}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \beta} + \frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \beta} \\ &= 2Re\left(\frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \beta}\right) \\ &= Re\left[\left(j\omega^2\Delta z\right)\frac{\partial I}{\partial \bar{\mathbf{y}}}\bar{\mathbf{x}}\exp\left(\frac{j}{2}\beta\left(z_k\right)\omega^2\Delta z\right)\right], \end{aligned} \tag{11}$$

where $Re[\cdot]$ takes the real part. Thus, the backpropagation of an entire CDC block is simply implemented as:

$$\frac{\partial I}{\partial \bar{\mathbf{x}}} = \mathrm{ifft}\left[\mathrm{fft}\left[\frac{\partial I}{\partial \bar{\mathbf{y}}}\right] \cdot \exp\left(\frac{j}{2}\beta\left(z_k\right)\omega^2\Delta z\right)\right]. \tag{12}$$

Fig. 5(e) shows the k-th NLPR block, where the signal phase is nonlinearly derotated as $\mathbf{y} = \mathbf{x} \cdot \exp(-j\gamma'(z_k)\|\mathbf{x}\|^2\Delta z)]$, $\|\mathbf{x}\|^2 = |\mathbf{x}_h|^2 + |\mathbf{x}_v|^2$. Its backpropagation is implemented as:

$$\begin{aligned}\frac{\partial I}{\partial \bar{\mathbf{x}}} &= \frac{\partial I}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \bar{\mathbf{x}}} + \frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \bar{\mathbf{x}}} \\ &= \left(-j\gamma'\Delta z\right)\frac{\partial I}{\partial \mathbf{y}}\mathbf{x}^2\exp\left(-j\gamma'\|\mathbf{x}\|^2\Delta z\right) \\ &\quad + \left(1 + j\gamma'\|\mathbf{x}\|^2\Delta z\right)\frac{\partial I}{\partial \bar{\mathbf{y}}}\exp\left(j\gamma'\|\mathbf{x}\|^2\Delta z\right), \end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial I}{\partial \gamma'} &= \frac{\partial I}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \gamma'} + \frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \gamma'} = 2Re\left(\frac{\partial I}{\partial \bar{\mathbf{y}}}\frac{\partial \bar{\mathbf{y}}}{\partial \gamma'}\right) \\ &= 2Re\left[\left(j\Delta z\right)\frac{\partial I}{\partial \bar{\mathbf{y}}}\bar{\mathbf{x}}\|\mathbf{x}\|^2\exp\left(j\gamma'\|\mathbf{x}\|^2\Delta z\right)\right]. \end{aligned} \tag{14}$$

In the *s*-th complex-valued FIR-filter blocks (Fig. 5(f)), which emulates optical filters, $\mathbf{h}$ is applied to signals as:

$$y\left(n\right) = \sum_{m=0}^{M-1} h\left(m\right)x\left(n-m\right), \tag{15}$$

where $x(n)$ and $y(n)$ ($n \in \{0, 1, \ldots, N-1\}$) denote the input and output of FIR blocks. Thus, its conjugate is

$$\overline{y\left(n\right)} = \sum_{m=0}^{M-1}\overline{h\left(m\right)}\;\overline{x\left(n-m\right)}. \tag{16}$$

By using (15) and (16), the backpropagation of FIR blocks is calculated as:

$$\begin{aligned}\frac{\partial I}{\partial \overline{x\left(n\right)}} &= \sum_{n'=0}^{N-1}\left(\frac{\partial I}{\partial y\left(n'\right)}\frac{\partial y\left(n'\right)}{\partial \overline{x\left(n\right)}} + \frac{\partial I}{\partial \overline{y\left(n'\right)}}\frac{\partial \overline{y\left(n'\right)}}{\partial \overline{x\left(n\right)}}\right) \\ &= \sum_{n'=0}^{N-1}\left(0 + \frac{\partial I}{\partial \overline{y\left(n'\right)}}\overline{h\left(n'-n\right)}\right), \end{aligned} \tag{17}$$

$$\begin{aligned}\frac{\partial I}{\partial \overline{h\left(m\right)}} &= \sum_{n'=0}^{N-1}\left(\frac{\partial I}{\partial y\left(n'\right)}\frac{\partial y\left(n'\right)}{\partial \overline{h\left(m\right)}} + \frac{\partial I}{\partial \overline{y\left(n'\right)}}\frac{\partial \overline{y\left(n'\right)}}{\partial \overline{h\left(m\right)}}\right) \\ &= \sum_{n'=0}^{N-1}\left(0 + \frac{\partial I}{\partial \overline{y\left(n'\right)}}\overline{x\left(n'-m\right)}\right). \end{aligned} \tag{18}$$

Note that the terms $\overline{h(n'-n)}$ and $\overline{x(n'-m)}$ are set to 0 for out of domain. Eqs. (17) and (18) mean that the backpropagation of convolution is the convolution of the backpropagated derivatives $\partial I/\partial \bar{\boldsymbol{y}}$ and flipped conjugate of the filters $flip(\bar{\mathbf{h}})$ or the inputs $flip(\bar{\mathbf{x}})$, respectively (see Fig. 5(f)).

References

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

[2] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, no. 2, pp. 804–817, Jan. 2008.
[3] A. Napoli *et al.*, "Digital compensation of bandwidth limitations for high-speed DACs and ADCs," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3053–3064, Jul. 2016.
[4] D. Rafique *et al.*, "Digital preemphasis in optical communication systems: On the DAC requirements for terabit transmission applications," *J. Lightw. Technol.*, vol. 32, no. 19, pp. 3247–3256, Oct. 2014.
[5] A. Matsushita *et al.*, "High-spectral-efficiency 600-Gbps/carrier transmission using PDM-256QAM format," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 470–476, Jan. 2019.
[6] P. W. Berenguer *et al.*, "Nonlinear digital pre-distortion of transmitter components," *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1739–1745, Apr. 2016.
[7] G. Paryanti *et al.*, "A direct learning approach for neural network based pre-distortion for coherent nonlinear optical transmitter," *J. Lightw. Technol.*, vol. 38, no. 15, pp. 3883–3896, Aug. 2020.
[8] T. Sasai *et al.*, "Wiener-Hammerstein model and its learning for nonlinear digital pre-distortion of optical transmitters," *Opt. Express*, vol. 28, no. 21, pp. 30952–30963, Oct. 2020.
[9] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Proc. Opt. Fiber Commun. Conf. Expo.*, San Diego, CA, USA, Mar. 2018, Art. no. W3A.4.
[10] V. Oliari *et al.*, "Revisiting efficient multi-step nonlinearity compensation with machine learning: An experimental demonstration," *J. Lightw. Technol.*, vol. 38, no. 12, pp. 3114–3124, Jun. 2020.
[11] R. M. Bütler, C. Häger, H. D. Pfister, G. Liga, and A. Alvarado, "Model-based machine learning for joint digital backpropagation and PMD compensation," *J. Lightw. Technol.*, vol. 39, no. 4, pp. 949–959, Feb. 2021.
[12] Q. Fan *et al.*, "Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning," *Nature Commun.*, vol. 11, pp. 3694–3704, Jul. 2020.
[13] B. I. Bitachon *et al.*, "Deep learning based digital backpropagation demonstrating SNR gain at low complexity in a 1200 km transmission link," *Opt. Express*, vol. 28, no. 20, pp. 29318–29334, Sep. 2020.
[14] B. Karanov *et al.*, "End-to-end deep learning of optical fiber communications," *J. Lightw. Technol.*, vol. 36, no. 20, pp. 4843–4855, Oct. 2018.
[15] R. A. Soriano, F. N. Hauske, N. G. Gonzalez, Z. Zhang, Y. Ye, and I. T. Monroy, "Chromatic dispersion estimation in digital coherent receivers," *J. Lightw. Technol.*, vol. 29, no. 11, pp. 1627–1637, Jun. 2011.
[16] Z. Wang *et al.*, "OSNR and nonlinear noise power estimation for optical fiber communication systems using LSTM based deep learning technique," *Opt. Express*, vol. 26, no. 16, pp. 21346–21357, Aug. 2018.
[17] B. Spinnler *et al.*, "Autonomous intelligent transponder enabling adaptive network optimization in a live network field trial," *J. Opt. Commun. Netw.*, vol. 11, no. 9, pp. C1–C9, Sep. 2019.
[18] T. Tanimura *et al.*, "Fiber-longitudinal anomaly position identification over multi-span transmission link out of receiver-end signals," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2726–2733, May 2020.
[19] T. Sasai *et al.*, "Simultaneous detection of anomaly points and fiber types in multi-span transmission links only by receiver-side digital signal processing," in *Proc. Opt. Fiber Commun. Conf. Expo.*, San Diego, CA, USA, Mar. 2020, Art. no. Th1F.1.
[20] T. Sasai *et al.*, "Digital backpropagation for optical path monitoring: Loss profile and passband narrowing estimation," in *Proc. Eur. Conf. Opt. Commun.*, Brussels, Belgium, Dec. 2020, Art. no. Tu2D.1.
[21] T. Sasai *et al.*, "Revealing Raman-amplified power profile and Raman gain spectra with digital backpropagation," in *Proc. Opt. Fiber Commun. Conf. Expo.*, San Francisco, CA, USA, Jun. 2021, Art. no. M3I.5.
[22] T. Sasai *et al.*, "Physics-oriented learning of nonlinear Schrödinger equation: Optical fiber loss and dispersion profile identification," 2021. *arXiv: 2104.05890, [eess.SP]*.
[23] S. Gleb *et al.*, "Fiber link anomaly detection and estimation based on signal nonlinearity," in *Proc. Eur. Conf. Opt. Commun.*, Bordeaux, France, Sep. 2021, Art. no. Tu2C2.5.
[24] D. Marcuse, C. R. Menyuk, and P. K. A.Wai, "Applications of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.
[25] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightw. Technol.*, vol. 28, no. 6, pp. 939–951, Mar. 2010.
[26] G. P. Agrawal, "Pulse propagation in fibers," in *Nonlinear Fiber Optics*, 5th ed. Cambridge, MA, USA: Academic, 2012, ch. 2, sec. 2.4, pp. 47–51.
[27] T. Sasai *et al.*, "Precise longitudinal power monitoring over 2,080 km enabled by step size selection of split step Fourier method," in *Proc. Opt. Fiber Commun. Conf. Expo.*, San Diego, CA, USA, Mar. 2022, Art. no. Th1C.4.
[28] Y. Hsueh *et al.*, "Passband narrowing and crosstalk impairments in ROADM-enabled 100G DWDM networks," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3980–3986, Dec. 2012.
[29] R. H. Nielsen, "Theory of the backpropagation neural network," in *Neural Networks for Perception*, vol. 2, 1st ed. Cambridge, MA, USA: Academic, 1991, ch. 3, sec. 3.3, pp. 65–93.
[30] L. S. Schanner *et al.*, "Power profile estimation of optical transmission links based on machine learning," *IEEE Photon. Technol. Lett.*, vol. 33, no. 19, pp. 1089–1092, Oct. 2021.
[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980v9 [cs.LG]*.
[32] A. May, F. Boitier, E. Awwad, P. Ramantanis, M. Lonardi, and P. Ciblat, "Receiver-based experimental estimation of power losses in optical networks," *IEEE Photon. Technol. Lett.*, vol. 33, no. 22, pp. 1238–1241, Nov. 2021.
[33] T. Tanimura *et al.*, "Concept and implementation study of advanced DSP-based fiber-longitudinal optical power profile monitoring toward optical network tomography," *J. Opt. Commun. Netw.*, vol. 13, no. 10, pp. E132–E141, Oct. 2021.
[34] M. Nakamura *et al.*, "Spectrally efficient 800 Gbps/carrier WDM transmission with 100-GHz spacing using probabilistically shaped PDM-256QAM," in *Proc. Eur. Conf. Opt. Commun.*, Rome, Italy, Sep. 2018, Art. no. We3G.5.
[35] U. Mengali and M. Morelli, "Data-aided frequency estimation for burst digital transmission," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 23–25, Jan. 1997.
[36] T. Sasai *et al.*, "Laser phase noise tolerance of uniform and probabilistically shaped QAM signals for high spectral efficiency systems," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 439–446, Jan. 2020.
[37] A. Boskovic, S. V. Chernikov, and J. R. Taylor, "Direct continuous-wave measurement of n_2 in various types of telecommunication fiber at 1.55 μm," *Opt. Lett.*, vol. 21, no. 24, pp. 1966–1968, Dec. 1996.
[38] E. Ip *et al.*, "Interchannel nonlinearity compensation for $3\lambda\times 114$ Gb/s DP-8QAM using three synchronized sampling scopes," in *Proc. Opt. Fiber Commun. Conf. Expo.*, Los Angeles, CA, USA, Mar. 2012, Art. no. OM3A.6.
[39] P. Poggiolini *et al.*, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.
[40] A. Carena *et al.*, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16335–16362, Jun. 2014.
[41] A. Ferrari *et al.*, "GNPy: An open source application for physical layer aware open optical networks," *J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. C31–C40, Mar. 2020.
[42] N. Takachio *et al.*, "Application of Raman-distributed amplification to WDM transmission systems using 1.55-μm dispersion-shifted fiber," *J. Lightw. Technol.*, vol. 19, no. 1, pp. 60–69, Jan. 2001.
[43] I. D. Phillips *et al.*, "Exceeding the nonlinear-Shannon limit using Raman laser based amplification and optical phase conjugation," in *Proc. Opt. Fiber Commun. Conf. Expo.*, San Francisco, CA, USA, Mar. 2014, Art. no. M3C.1.
[44] J. D. Ania-Castañón *et al.*, "Simultaneous spatial and spectral transparency in ultra-long fiber lasers," *Phys. Rev. Lett.*, vol. 101, no. 12, Sep. 2008, Art. no. 123903.
[45] K. Toge, K. Hogari, and T. Horiguchi, "Measurement of Raman gain distribution in optical fibers," *IEEE Photon. Technol. Lett.*, vol. 14, no. 7, pp. 974–976, Jul. 2002.
[46] D. Xie *et al.*, "LCoS-based wavelength-selective switch for future finer-grid elastic optical networks capable of all-optical wavelength conversion," *IEEE Photon. J.*, vol. 9, no. 2, Apr. 2017, Art. no. 7101212.

Takeo Sasai (Member, IEEE) received the B.E. and M.E. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 2014 and 2016, respectively. Since 2021, he has been working toward the Ph.D. degree in applied physics in the university. In 2016, he joined NTT Network Innovation Laboratories, Yokosuka, Japan. He is engaged in the modeling of optical fiber communication systems from both behavioral and data-driven approaches, particularly optical fiber nonlinearity and transceiver nonlinearity. He is a Member of the Optica and Institute of Electronics, Information and Communication Engineers (IEICE) of Japan. He was the recipient of the 2021 IEICE Communications Society Optical Communication Systems Young Researchers Award.

Masanori Nakamura (Member, IEEE) received the B.S. and M.S. degrees in applied physics from Waseda University, Tokyo, Japan, in 2011 and 2013, respectively, and the Ph.D. degree in electrical, electronic, and infocommunications engineering from Osaka University, Osaka, Japan, in 2021. In 2013, he joined NTT Network Innovation Laboratories, Yokosuka, Japan, where he is engaged in research on high-capacity optical transport networks. He is a Member of IEICE, Tokyo, Japan. He was the recipient of the 2016 IEICE Communications Society Optical Communication Systems Young Researchers Award.

Etsushi Yamazaki (Member, IEEE) received the B.E. and M.E. degrees in electronic engineering from the University of Tokyo, Tokyo, Japan, in 2000 and 2002, respectively. He is currently with Nippon Telegraph and Telephone Corporation (NTT) Network Innovation Laboratories, Yokosuka, Japan, where he is engaged in R&D for optical transmission systems and photonic networks. From 2018 to 2019, he was a Visiting Researcher with Stanford University, Stanford, CA, USA. Mr. Yamazaki is a Member of the IEICE of Japan. In 2012, he was the recipient of the Sakurai-Kenjiro Memorial Award from the Optoelectronics Industry and Technology Development Association (OITDA) of Japan.

Shuto Yamamoto (Member, IEEE) was born in Tokyo, Japan. He received the B.S. and M.S. degrees in physics and the Ph.D. degree in electronic engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2003, 2005, and 2018, respectively. He is currently a Senior Research Engineer with Network Innovation Laboratories, Nippon Telegraph and Telephone (NTT) Corporation, Yokosuka, Japan. His research interests include photonic transport network systems. He is a Member of the IEEE Photonics Society and a Member of the IEICE of Japan. Since 2018, he has been a Member of the Technical Program Committee of the Optical Networking and Communication Conference & Exhibition (OFC). He was the recipient of the Best Paper Award at the 16th OptoElectoronics and Communications Conference (OECC), Young Researcher's Award from IEICE, and Achievement Award from IEICE in 2011, 2012, and 2016, respectively. He was also the recipient of the NTT Presidential Award and the TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2014 and 2019, respectively.

Hideki Nishizawa received the B.E. and M.E. degrees in physics from Chiba University, Chiba, Japan, in 1994 and 1996, respectively. In 1996, he joined NTT Network Service System Laboratories, Musashino, Tokyo, Japan, where he is engaged in research work on photonic switching systems. In 2001, he moved to NTT Phoenix Communications Inc. (now NTT BizLink Inc.), where he worked on the development of visual communication services. In 2007, he moved to NTT Research and Development Planning Department, where he was involved in planning R&D activities for future networks. Since 2014, he has been with NTT Network Innovation Laboratories. His current research interests include disaggregated optical transport networks that use digital coherent technology and virtualization technology. He is a Member of IEICE.

Yoshiaki Kisaka received the B.E. and M.E. degrees in physics from Ritsumeikan University, Kyoto, Japan, in 1996 and 1998, respectively. In 1998, he joined NTT Optical Network Systems Laboratories, Yokosuka, Kanagawa, Japan, where he was engaged in R&D on high-speed optical communication systems, including 40-Gbit/s/ch WDM transmission systems and mapping/multiplexing schemes in optical transport networks (OTNs). From 2007 to 2010, he was with NTT Electronics Technology Corporation, where he was engaged in the development of 40/100-Gbit/s OTN framer LSIs. His current research interests include high-speed and high-capacity optical transport networks that use digital coherent technology. Since 2010, he has been contributed to R&D on digital coherent signal processers at 100 Gbit/s and beyond. He is a Member of IEICE. He was the recipient of the Young Engineer's Award from IEICE in 2001.