

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

# Optimal Nonlinear Spectral Back Rotation for Discrete Eigenvalue Transmission Systems

Chuang Xu and Alan Pak Tao Lau

**Abstract**—Nonlinear frequency division multiplexing (NFDM) provides interference-free transmission in lossless and noiseless fiber up to a phase rotation in the nonlinear spectrum and hence signal recovery using a simple phase back rotation will suffice. However, amplified spontaneous emission (ASE) noise distorts both the discrete eigenvalues and nonlinear spectrum of the signal, reducing the effectiveness of this simple back rotation method. We propose back rotating the nonlinear spectral phase by only *half* of the transmission distance as a computationally simple impairment compensation algorithm for discrete eigenvalue transmission systems. We analytically prove its optimality and show that it is equivalent to the linear minimum mean square error (LMMSE) algorithm in terms of its ability to eliminate the correlations between the discrete eigenvalues and b-coefficient and hence its performance but with lower complexity and no need to use training data to determine the filter coefficients in the LMMSE algorithm. Analytical, numerical, and experimental results are presented to verify the proposed algorithm. We also discussed the influence of possible non-ideal factors in practical systems affecting the proposed scheme.

**Index Terms**—Discrete NFDM, eigenvalue transmission, NFT, soliton.

## I. INTRODUCTION

NONLINEAR Frequency Division Multiplexing (NFDM) is a promising theory for addressing the fiber nonlinearity-induced capacity crunch. Using nonlinear Fourier transform (NFT) corresponding to the nonlinear Schrodinger equation (NLSE) that governs signal propagation in a single-mode fiber, information can be encoded into the signal's nonlinear spectrum and each nonlinear spectral component evolves linearly as a spectral phase rotation without mutual interference during propagation in a lossless fiber. Therefore, in principle, transmission impairment compensation and signal recovery can be achieved by back rotating the phase of the nonlinear spectrum.

As an extension of soliton communication, discrete eigenvalue communication was first proposed by Hasegawa and Nyu [1] in 1993, followed by a more general transmission scheme, NFDM proposed by Yousefi and Kschischang [2] in 2014, where information can be encoded in discrete and/or

This work was supported in part by the Hong Kong Government Research Grants Council General Research Fund (GRF) under Project PolyU 15220120, and in part by the National Natural Science Foundation of China under Grant 62205068. (Corresponding author: Chuang Xu).

Chuang Xu and Alan Pak Tao Lau are with the Department of Electrical and Electronic Engineering, Photonics Research Institute, The Hong Kong Polytechnic University, Hong Kong SAR, China. (e-mail: chuang.xu@connect.polyu.hk; eaptlau@polyu.edu.hk).

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

continuous nonlinear spectrum. Ordinary digital coherent transmission waveform combined with NFT detection was used in early experimental works for back-to-back [3] and long-haul [4], [5] transmission, while Dong et al [6] conducted the first experimental demonstration of discrete NFDM signal transmission containing up to four eigenvalues. Later, advanced modulation formats were applied to the discrete spectrum to carry information, but significant noise on eigenvalues was observed, so the authors used ideal noiseless eigenvalues in back rotation for recovery [7]. Encoding information on b-coefficient rather than discrete spectrum was found to be beneficial [8]. Furthermore, joint modulation of discrete and continuous spectrum [9] and joint modulation of eigenvalue and discrete b-coefficient [10] were investigated.

Although the integrability of NLSE is broken by fiber loss, NFDM scheme is still feasible with distributed amplification [11], or with discrete lumped amplification by adopting the path-averaging model [12], [13]. However, amplified spontaneous emission (ASE) noise from optical amplifiers distorts the nonlinear spectrum and discrete eigenvalues, and therefore discrete NFDM signal cannot be recovered by simply back rotating the nonlinear spectrum with noise-corrupted eigenvalues. To address the problem, the linear minimum mean square error (LMMSE) filter was proposed as a high correlation between the received eigenvalue and b-coefficient was observed [8]. Nonlinear filters and machine learning [14] have also been studied to reduce noise in b-coefficients but these schemes have higher computational complexity and require training symbols to learn the filter coefficients and/or weights.

In our previous work in [15], we built a simple but effective model of ASE noise accumulation on discrete eigenvalues and b-coefficients during propagation in a practical lossy discrete eigenvalue transmission system and proposed back rotating the received nonlinear spectrum by *half* of the full propagation distance to minimize the b-coefficient noise, achieving nearly the same level of performance as LMMSE filter. We hereby extend the work with detailed analytical derivations on the optimality of the half back rotation, performance comparisons between optimal half back rotation and LMMSE filter, and detailed analysis of non-ideal factors, as well as providing experimental verifications of the proposed signal processing algorithm.

The paper is organized as follows. In Section II, we briefly review the basics of NFT and NFDM, and then introduce the noise accumulation model and prove the optimality of half back rotation. In Section III, numerical results are presented, followed by discussions of practical non-ideal factors that may

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

affect the effectiveness of the half back rotation algorithm in Section IV. Section V contains experimental verification and Section VI concludes the paper.

## II. PRINCIPLE

### A. NFT and NFDM

The evolution of signal's complex envelope  $E(l, \tau)$  in a single-mode fiber is described by NLSE if fiber loss, polarization effect, higher-order dispersion, Raman effect, and noise are neglected:

$$i \frac{\partial E(l, \tau)}{\partial l} - \frac{\beta_2}{2} \frac{\partial^2 E(l, \tau)}{\partial \tau^2} + \gamma |E(l, \tau)|^2 E(l, \tau) = 0 \quad (1)$$

where  $\tau$  (s) is the temporal variable in a time-retarded frame,  $l$  (m) is the spatial variable,  $\beta_2$  is the second-order group velocity dispersion parameter ( $\text{m}^2/\text{s}$ ), and  $\gamma$  is the fiber nonlinearity parameter ( $1/\text{W}/\text{m}$ ).

It is conventional to convert (1) into a dimensionless form by changing variables:

$$q(z, t) = \frac{E(l, \tau)}{\sqrt{P}}, z = \frac{l}{L_D}, t = \frac{\tau}{T_0} \quad (2)$$

where  $P = 1/(\gamma L_D)$ ,  $L_D = T_0^2/|\beta_2|$  ( $L_D$  is known as dispersion length) and  $T_0$  is a freely chosen time normalization parameter. For the anomalous dispersion situation ( $\beta_2 < 0$ ), the normalized NLSE is given by

$$i \frac{\partial q(z, t)}{\partial z} + \frac{1}{2} \frac{\partial^2 q(z, t)}{\partial t^2} + |q(z, t)|^2 q(z, t) = 0 \quad (3)$$

The nonlinear spectrum of  $q(z = z_0, t)$ , is obtained by solving the eigenvalue problem of the Lax operator associated with NLSE,

$$\frac{\partial \mathbf{v}}{\partial t} = \begin{pmatrix} -j\lambda & jq(z_0, t) \\ jq^*(z_0, t) & j\lambda \end{pmatrix} \mathbf{v} \quad (4)$$

where  $\mathbf{v} = [v_1, v_2]^T$  ( $[\cdot]^T$  denotes transpose) is a  $2 \times 1$  vector that solves (4). The nonlinear Fourier coefficients  $a(z = z_0, \lambda)$  and  $b(z = z_0, \lambda)$  are given by (the spatial variable  $z = z_0$  is omitted for clear notation)

$$\begin{bmatrix} a(\lambda) \\ b(\lambda) \end{bmatrix} = \lim_{t \rightarrow +\infty} \begin{bmatrix} v_1 e^{j\lambda t} \\ v_2 e^{-j\lambda t} \end{bmatrix} \text{ with } \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \xrightarrow{t \rightarrow -\infty} \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{-j\lambda t} \quad (5)$$

The continuous and discrete spectral coefficient is defined as  $Q_c(\lambda) = b(\lambda)/a(\lambda)$ ,  $Q_d(\lambda_i) = b(\lambda_i)/a'(\lambda_i)$  where  $\lambda_i \in \mathbb{C}^+$  is the discrete eigenvalue satisfying  $a(\lambda_i) = 0$ , and  $\lambda \in \mathbb{R}$  is the continuous eigenvalue. During propagation, the eigenvalue stays unchanged, while the time-domain nonlinear signal evolution is translated into linear (forward) evolution of b-coefficient, given by  $b(z_0 + z, \lambda) = b(z_0, \lambda) e^{-2j\lambda^2 z}$  and  $b(z_0 + z, \lambda_i) = b(z_0, \lambda_i) e^{-2j\lambda_i^2 z}$ .

In fiber optics, a signal that only has with discrete nonlinear spectrum is also known as an optical soliton in the time domain, and a 1<sup>st</sup>-order soliton's temporal position and phase are determined by  $\ln|b(\lambda_1)|$  and  $\angle b(\lambda_1)$ , respectively. For transmission in the real-world lossy fiber with lump amplification by erbium-doped fiber amplifier (EDFA), path-average soliton should be used, which means that if span length  $L_A \ll L_D$ , increasing the theoretical soliton launch power by  $G \ln G / (G - 1)$  can make the soliton envelope well

retained during propagation even though the pulse power varies dramatically across spans, where  $G$  is the EDFA gain (also span loss).

### B. Nonlinear Spectral Back Rotation

Consider an  $N$ -span EDFA transmission system with each span of equal normalized length  $L$ , i.e.,  $L = L_A/L_D$  as shown in Fig. 1. At each EDFA, circularly symmetric complex white Gaussian ASE noise is added to the signal in the time domain. The ASE-induced noise to the discrete eigenvalues is generally not Gaussian but can be approximated by a conditional Gaussian distribution in the small noise limit [16]. We consider a signal with only one discrete eigenvalue,  $\lambda_0 = \alpha_0 + j\beta_0 = 0 + 0.5j$ , i.e., a 1<sup>st</sup>-order soliton in the time domain, and assume that at each EDFA independent Gaussian noise is added to the eigenvalues and b-coefficients and they are free from ASE noise during propagation between EDFAs. Their values right after the  $i^{\text{th}}$  EDFA are denoted by  $\alpha_i$ ,  $\beta_i$  and  $b_i$ , respectively (We will use the subscript to indicate span number hereafter, rather than the ordinal of eigenvalue as there is only one eigenvalue).

During propagation, instead of staying unchanged, the eigenvalue evolves in a random-walk manner and gradually deviates from its original value due to ASE noise loading, and the b-coefficient also deviates from its ideal evolution trace due to noise accumulation on itself, and more importantly, the deviated eigenvalue. Mathematically,  $b_{i+1} = b_i e^{-2j(\alpha_i + j\beta_i)^2 L} + n_{i+1}$ , where  $n_i$  stands for the noise added to the b-coefficient by the  $i^{\text{th}}$  EDFA. The received b-coefficient can be written as:

$$\begin{aligned} b_{N+1} &= ((b_0 + n_1)e^{-2j\lambda_1^2 L} + n_2)e^{-2j\lambda_2^2 L} \dots + n_{N+1} \\ &= b_0 e^{-2jL \sum_{k=1}^N \lambda_k^2} + \sum_{i=1}^N n_i e^{-2jL \sum_{k=i}^N \lambda_k^2} + n_{N+1} \end{aligned} \quad (6)$$

In practice, we noticed that  $|n_i| \ll |b_0|$ , which makes the first term dominant so the following terms can be safely dropped. In other words, the deviation of b-coefficient is mostly contributed by the deviated eigenvalue, rather than the noise on b-coefficient itself. Besides, note that during propagation,  $\alpha_i \approx 0$  and  $\beta_i \approx 0.5$ , which makes the deviation of  $\alpha$  and  $\beta$  contribute much more significantly to the deviation of  $|b_i|$  and  $\angle b_i$ , respectively. Thus, we can obtain  $|b_{i+1}| \approx |b_i| e^{4\alpha_i \beta_i L} \approx |b_i| e^{2\alpha_i L}$  and  $\angle b_{i+1} \approx \angle b_i - 2(\alpha_i^2 - \beta_i^2)L \approx \angle b_i + 2\beta_i^2 L$  (Appendix A), and it is more convenient to use the logarithm of  $|b|$ :

$$\ln|b_{N+1}| = \ln|b_0| + 2L \sum_{i=1}^N \alpha_i \quad (7)$$

$$\angle b_{N+1} = \angle b_0 + 2L \sum_{i=1}^N \beta_i^2 \quad (8)$$

The random evolution traces of  $\lambda_i$  and  $b_i$  are schematically drawn in Fig. 2. We noticed that within each span, the evolution of  $\angle b$  follows a curved trace, rather than the straight line given by the NFT evolution equation (while we plot it by straight line segments for simplicity). However, the trace curves back and coincides with the straight line at the end of each span, making the forward evolution model still work

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

well. The curved evolution is due to the unique property of

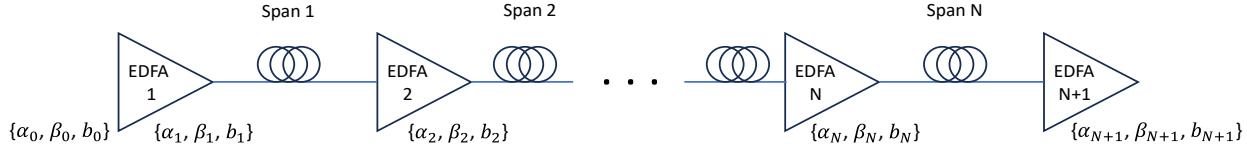


Fig. 1 Schematic diagram of the transmission link. Gaussian noise is independently added to  $\lambda$  and  $b$  at each EDFA.

path-averaged soliton, which will be discussed more in Section IV.

Without loss of generality, we set the initial  $b$ -coefficient  $b_0 = 1$  so  $\ln|b_0| = \angle b_0 = 0$ . We also assume the noise added to the eigenvalues at each EDFA is identical independently Gaussian distributed. That is to say,  $\alpha_i$  and  $\beta_i$  can be understood as a “discrete” Wiener process. More complex models can be found in [17], [18], [19], which argued that the noise in the discrete eigenvalues is generally not Gaussian.

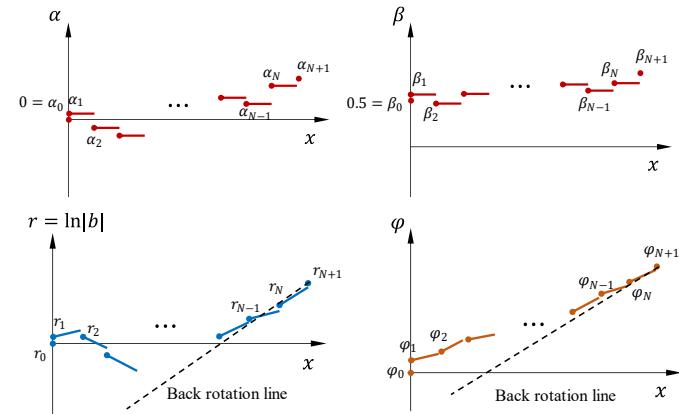


Fig. 2 Random evolution of  $\lambda_i$  and  $b_i$  due to the ASE noise added by each EDFA, and the x-axis is normalized to  $L$ ,  $x = z/L$ .

After transmission over the  $N$ -span system with  $N+1$  EDFA, the received signal is processed by NFT and noisy eigenvalue,  $\lambda_{N+1} = \alpha_{N+1} + j\beta_{N+1}$  and  $b$ -coefficient,  $b_{N+1}$  are obtained. We then back rotate  $b_{N+1}$  by  $e^{2j\lambda_{N+1}^2 z}$  with  $z$  ranging from 0 to  $NL$ , correspondingly referred to as zero to full back rotation. This is equivalent to moving back along the straight dash line in Fig. 2 that starts from point  $(N, b_{N+1})$  with a slope of  $2L\alpha_{N+1}$  or  $2L\beta_{N+1}^2$ , respectively. Note that in a noiseless case, the evolution and back rotation traces of  $b$ -coefficient will exactly coincide, so a perfect recovery of  $b$ -coefficient can be achieved. However, due to ASE noise loading, the eigenvalue evolves in a random-walk manner and the  $b$ -coefficient evolution trace is approximately the summation of the random walk of the eigenvalue. Therefore, the evolution trace and back rotation line will never be the same but just tangent to each other at point  $(N, b_{N+1})$ . More specifically, the back rotation line for  $\ln|b_{N+1}|$  is  $r(x) = 2L\alpha_{N+1}(x - N) + r_{N+1}$ , where we used  $r(x) = \ln|b(x)|$ ,  $r_{N+1} = \ln|b_{N+1}|$  for clear notation,  $x = z/L$  is the position coordinate normalized to  $L$ , and  $b(x)$  denotes the value of  $b$ -coefficient that back rotated to the position  $x$ . Note that  $r(x)$  is a random variable dependent on  $x$  since  $\alpha_{N+1}$  and  $r_{N+1}$  are random variables. The variance of  $r(x)$  is derived (in Appendix. B) to be

$$\text{Var}(r(x)) = 4L^2\sigma_\alpha^2(N+1) \left[ x(x-N) + \frac{N(2N+1)}{6} \right] \quad (9)$$

where  $\sigma_\alpha^2$  is the variance of the noise added to  $\alpha$  by each EDFA, which can be obtained by simulation or experiment. The  $N(N+1)(2N+1) \propto N^3$  term results from the variance of  $r_{N+1}$ , and this cubic dependence is consistent with the well-known Gordon-Haus effect [20], [21]. To minimize  $\text{Var}(r(x))$ , we set  $\frac{d}{dx} \text{Var}(r(x)) = 0$  and obtain  $x_{opt} = N/2$ .

Similarly, the back rotation line for  $\angle b_N$  is  $\varphi(x) = 2L\beta_{N+1}^2(x - N) + \varphi_{N+1}$ , where  $\varphi(x)$  represents  $\angle b(x)$  for clear notation. The variance of  $\varphi(x)$  is derived (in Appendix. C) to be

$$\text{Var}(\varphi(x)) \approx 4L^2\sigma_\beta^2(N+1) \left[ x(x-N) + \frac{N(2N+1)}{6} \right] \quad (10)$$

where  $\sigma_\beta^2$  is the variance of the noise added to  $\beta$  at each EDFA, and we also obtain  $x_{opt} = N/2$  that minimizes the variance. In short, both the variance of  $\ln|b|$  and  $\angle b$  are minimized when the back rotation length is *half* of the full transmission distance. Such a result has some resemblance to the optimal scale factor for nonlinear phase noise compensation in the chromatic-dispersion-free scenario [22]. Note that when a full back rotation of the nonlinear spectral phase is applied ( $x = 0$ ), both the variance of  $\ln|b|$  and  $\angle b$  become the same as those at the receiver, hence a full back rotation does *not* help to reduce the noise on  $b$ -coefficients at all. Besides, back rotating with the noiseless eigenvalue  $(0 + 0.5j)$  at the transmitter is effectively a constant operation across different symbols (back rotation lines are all parallel to each other) and hence does not reduce the variance of  $\ln|b|$  and  $\angle b$  too.

Theoretically, to make a perfect back rotation, we need all noisy eigenvalues at each EDFA and back rotate with them for each span in sequence, while in practice we only have access to the noisy eigenvalue at the receiver. However, the received noisy eigenvalue results from ASE noise accumulation along transmission, while the received noisy  $b$ -coefficient results from the accumulation of the accumulated ASE noise at each EDFA. Thus, the received noisy eigenvalue still contains considerable information about all the noise introduced along transmission and helps denoise the noisy  $b$ -coefficient to a certain extent. This fact is also reflected in their high correlation that has been experimentally reported [8] and will be theoretically analyzed in Section II. C. The optimality of half back rotation intuitively indicates that we should make some use of the received eigenvalue, but not fully trust it.

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

### C. Relation to the LMMSE Filter

We calculated the signal-to-noise ratio (SNR) of the signal after back rotation by  $SNR = 1/EVM_{RMS}^2$ , where  $EVM_{RMS}$  stands for root mean square error vector magnitude (EVM) commonly used in conventional coherent communications. It is noted from Fig. 5(b) that the signal SNR after half back rotation is close to that by LMMSE filter and applying an LMMSE filter to the half back rotated signal barely further improves the SNR. This motivates us to investigate the correlation between the b-coefficient and the eigenvalue during the back rotation process since LMMSE filter essentially denoises the b-coefficients by making use of their correlations with eigenvalues. With the noise model described above, the correlation coefficient between  $r(x)$  and  $\alpha_{N+1}$ ,  $\varphi(x)$  and  $\beta_{N+1}$  are derived (in Appendix D) to be

$$\rho_{\varphi(x),\beta_{N+1}} \approx \rho_{r(x),\alpha_{N+1}} = \frac{x - N/2}{\sqrt{x(x - N) + N(2N + 1)/6}} \quad (11)$$

Clearly, the correlation coefficient reaches zero right at half back rotation, achieving the same condition as the LMMSE filter and hence leading to the same SNR performance.

### III. NUMERICAL RESULTS

We conducted simulations on a  $40 \times 50$  km EDFA system with non-zero dispersion-shifted fiber (NZ-DSF), loss coefficient: 0.2dB/km, dispersion coefficient: 4 ps/km/nm) Ideal transmitters, EDFA and receivers were used in the simulation, while possible non-ideal factor's effect will be discussed in Section IV. 1024 path-averaged 1<sup>st</sup>-order soliton pulses with 16APSK-modulated b-coefficients were used to estimate the variance of  $\ln|b|$ ,  $\angle b$  and SNR during the forward evolution and back rotation process. The ratio of symbol interval to soliton's full width at half maximum (FWHM) was set to 8.8 to avoid interaction between the adjacent solitons so the effect of ASE noise could be isolated. The symbol rate was set to 3GBaud, leading to a dispersion length,  $L_D$  of 194.6 km. NFT was conducted to the power-normalized signal every 10 km to study the evolution of eigenvalues and b-coefficients of the soliton pulses. For back rotation, we evaluated the b-coefficients every 100 km backward from the receiver. Fig. 3 shows the forward evolution and back rotation traces of  $\ln|b|$  and  $\angle b$ , where the red curve and black dots represent the deviations  $\ln|b_i| - \ln|b_0|$  and  $\angle b_i - \angle b_0$  of each random pulse during forward evolution and back rotation, respectively. Interestingly, the evolution of  $\angle b$  is curved in each span and forms a stair-like trace over the full link.

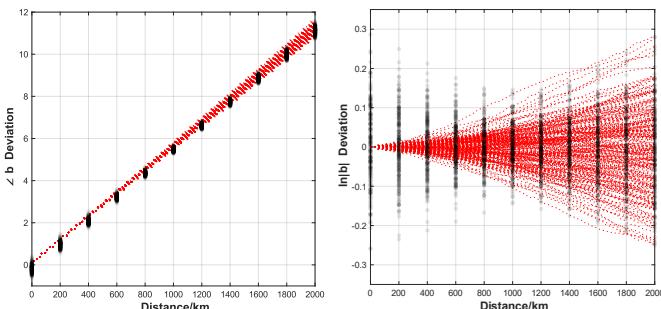


Fig. 3. Forward evolution (red curves) and back rotation (black dots) of  $\angle b$  and  $\ln|b|$ . Transparent dots are used to better show their dispersity during

back rotation, and back rotation dots are plotted every 200 km for a clearer view.

Fig. 4 shows that the variances of  $\ln|b|$  and  $\angle b$  indeed reaches their minimum with a half back rotation, and the SNR of b-coefficient reaches the maximum. The SNR for full back rotation is close to that at Rx, which is consistent with the theoretical analysis. Signal distributions at different back rotation ratios are inset for visual illustration of the optimality of the proposed half back rotation algorithm. In addition, the analytical expression of  $\rho_{\varphi(x),\beta_{N+1}}$  and  $\rho_{r(x),\alpha_{N+1}}$  match the simulated correlation coefficients quite well as shown in Fig. 5(a), and clearly, the correlation coefficient reduces to zero with half back rotation as expected and reaches its maximum of around  $\pm\sqrt{3}/2 \approx \pm 0.87$  (given  $N$  is not too small) with zero or full back rotation.

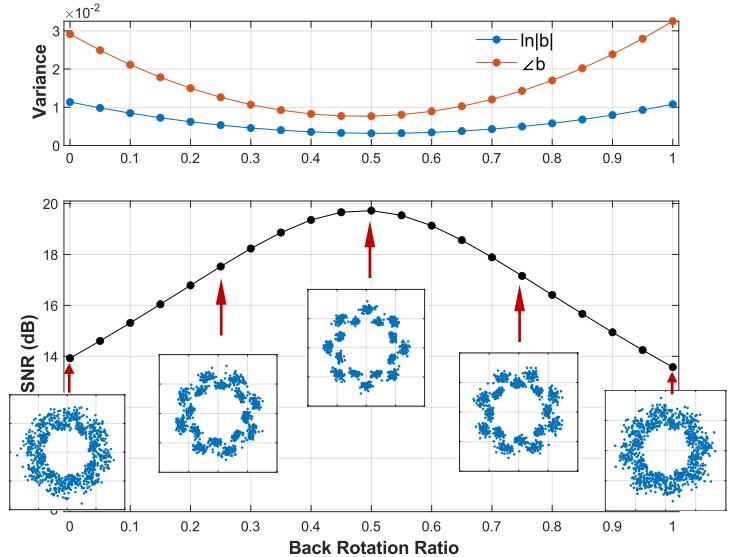


Fig. 4 Variance of  $\ln|b|$ ,  $\angle b$  and the SNR of b-coefficient for different ratios of back rotation.

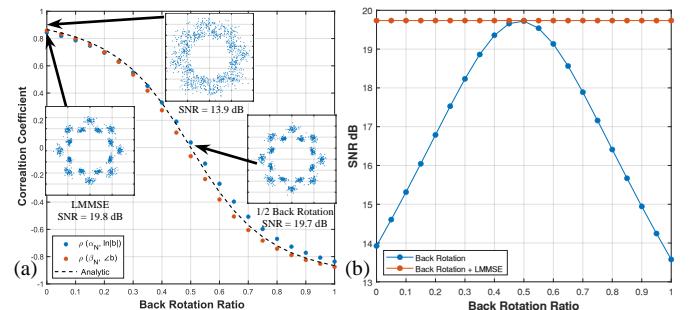


Fig. 5. (a) Simulated and analytic correlation coefficients between the  $b$  and  $\theta$  during back rotation. Constellations of the raw received signal, received signal processed by LMMSE filter or half back rotation are inset; (b) SNR of b-coefficient with back rotation only and with back rotation plus LMMSE filter.

### IV. EFFECT OF NON-IDEAL FACTORS

The analytical noise model and the corresponding numerical simulation is establish based on ideal transceivers and channel without other sources of noise and distortions. Here we discuss the impact of possible non-ideal factors in practice on the performance of the proposed half back rotation algorithm.

#### A. Non-ideal Bias Control

In a digital coherent system, the soliton signal was generated

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

by an IQ (in-phase and quadrature) modulator, which ideally should be biased at the null point. We noticed that non-ideal bias control can cause fluctuation in the soliton's amplitude right after modulation, which translates into fluctuation of  $\beta_0$ . Moreover, the amplitude fluctuation depends on the soliton's phase, hence the phase of b-coefficient, so the constellation clusters of  $b$  will rotate with different angular velocities that are determined by their corresponding  $\beta^2$ , resulting in an asymmetric constellation at the receiver side, which will also be experimentally shown in section V.

Specifically, let  $s(t) = \text{sech}(t) e^{i\theta}$  denote an ideal dimensionless 1<sup>st</sup>-order soliton, and  $\delta_{IQ} = \delta_I + j\delta_Q$  denote the bias on I and Q. Then the soliton shifted by IQ bias is  $s'(t) = s(t) + \delta_{IQ}$  with instantaneous power of  $P'(t) = |s(t) + \delta_{IQ}|^2 \approx \text{sech}^2(t) + 2 \text{sech}(t) R \cos(\theta - \phi)$ , where  $R = |\delta_{IQ}|$ ,  $\phi = \arctan(\delta_Q/\delta_I)$ . We may use Parseval's identity of nonlinear Fourier Transform [2] to obtain the  $\beta$  after the soliton is perturbated by the IQ bias:

$$\int_{-\infty}^{\infty} P'(t) dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 + |\tilde{q}(\lambda)|^2) d\lambda + 4\beta \quad (12)$$

which gives

$$\beta \approx 0.5 + \frac{\pi}{2} R \cos(\theta - \phi) \quad (13)$$

where we have assumed the number of eigenvalues is still 1, and the continuous nonlinear spectrum  $\tilde{q}(\lambda)$  can be ignored, given that the nonideal IQ bias is typically small in practice. Then, for a specific  $\delta_{IQ}$ ,  $\beta$  is determined by the soliton's phase  $\theta$ , and so as  $\angle b$  ( $= \pi - \theta$ ). Thus, different b-coefficient phases will lead to different initial values of  $\beta$  when IQ bias exists.

This can be illustrated by the following simulation example. 28 ideal 1<sup>st</sup>-order solitons are  $b$ -modulated by 14 different phases and 2 amplitudes, and  $\delta_{IQ} = 0.01 + 0.02j$  are applied to them. Fig. 6(a) shows the changes in amplitude due to the non-zero IQ bias, which are dependent on  $\angle b$ . Note that in Fig. 6(b), the IQ-bias-perturbed  $\beta$  follow the sinusoidal curve given by (13) quite well.

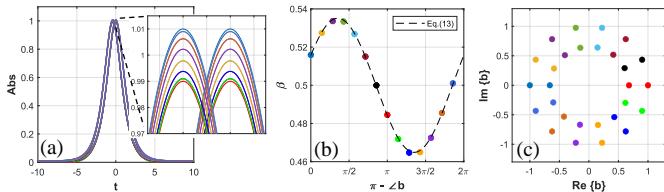


Fig. 6. (a) The amplitude of 16 PSK-modulated solitons perturbed by a given IQ bias; (b) the IQ-bias-perturbed  $\beta$ ; (c) the corresponding b-coefficients.

Let  $\Delta\beta_0$  denote the initial errors in  $\beta_0$ . The back rotation line becomes  $\tilde{\varphi}(x) = 2L\tilde{\beta}_{N+1}^2(x - N) + \tilde{\varphi}_{N+1}$ , where  $\tilde{\beta}_{N+1} = \beta_{N+1} + \Delta\beta_0$  and  $\tilde{\varphi}_{N+1} = 2L\sum_{i=1}^N(\beta_i + \Delta\beta_0)^2$  are the counterparts to  $\beta_{N+1}$  and  $\varphi_{N+1}$ , respectively. The optimal back rotation location for  $\angle b$  will be at (see Appendix E)

$$x_{opt} = \frac{N(N+1)\sigma_{\beta}^2}{2[(N+1)\sigma_{\beta}^2 + \sigma_{\beta_0}^2]} \quad (14)$$

where  $\sigma_{\beta_0}^2$  stands for the variance of  $\Delta\beta_0$ .

Thus, if the initial error is significant enough such that  $\sigma_{\beta_0}^2$

cannot be neglected (such as nonideal bias control or low OSNR at Tx EDFA), the optimal back rotation for phase will be greater than a half, shifting toward full back rotation, and it is reduced to the ideal case when  $\sigma_{\beta_0}^2 = 0$ . An intuitive interpretation of this phenomenon is that the initial error introduced by shifted bias affects the whole transmission process, and compared with the ASE noise gradually added along the transmission line, its effect is rather 'deterministic', so a full back rotation is necessary to undo the effect. Overall, the optimal back rotation ratio for  $\angle b$  stands between 1/2 and 1. The back rotation for  $\ln|b|$  is much less affected, and the optimal ratio is still 1/2 as  $\alpha$  is not affected by bias offset.

As an illustrative example, we offset a transmitted signal,  $s(t) = \sum_{k=-\infty}^{\infty} \text{sech}(t - T_k) e^{i\theta_k}$  by  $\delta_I = 1.5E-3$  in the dimensionless domain. Fig. 7 shows the corresponding forward evolution traces of b-coefficient and eigenvalue. The initial imaginary part  $\beta_0$  splits into various values at the transmitter (Fig. 7 (c)) due to the phase-dependent power fluctuation mentioned above, resulting in different phase evolution velocities indicated by the different trace slopes in Fig. 7 (a) and thus an asymmetric constellation at the receiver. Fig. 8 (a) and (c) shows the back rotation of  $\angle b$  and the corresponding variance versus different back rotation ratios, respectively. As predicted by (14), the lowest variance of  $\angle b$  happens with a back rotation ratio greater than 1/2, while that for  $\ln|b|$  is still 1/2.

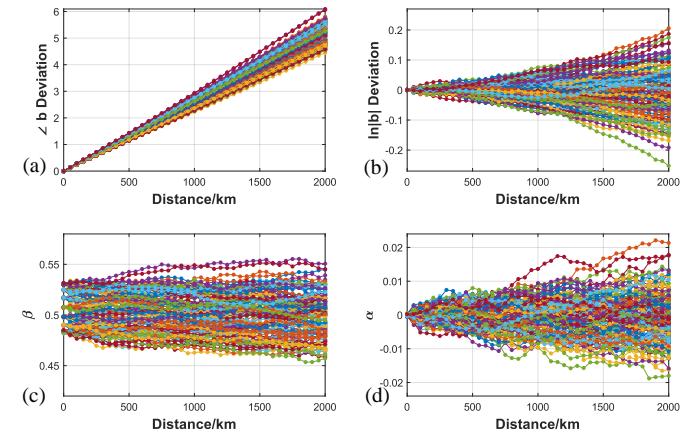
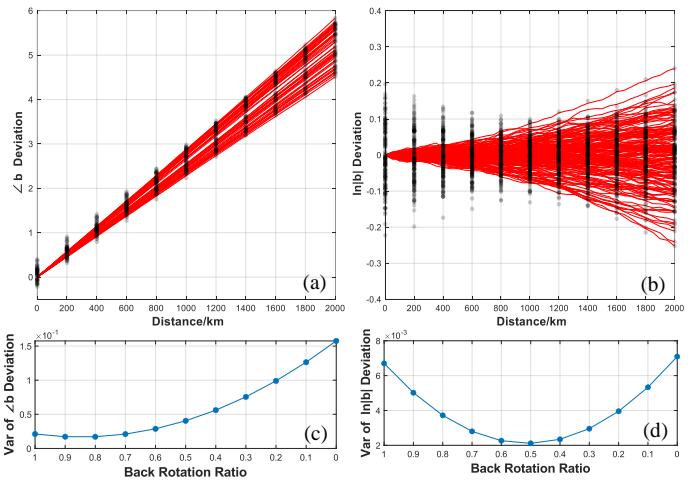


Fig. 7. Evolution of (a)  $\angle b$ , (b)  $\ln|b|$ , (c)  $\beta$  and (d)  $\alpha$  when  $\delta_I = 1.5E-3$  is added on  $s_I(t)$  at transmitter.



>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

Fig. 8. (a) Forward evolution and (b) back rotation of  $\angle b$  and  $\ln|b|$  with  $\delta_l = 1.5E-3$ . (c) and (d) show the corresponding variance versus the back rotation ratio.

### B. Gain-Loss Imbalance in Fiber Loop

In a strict sense, path-averaged soliton is not a ‘true’ soliton but just a soliton-like pulse. Although the path-averaged soliton’s power evolves periodically due to fiber loss and periodical discrete amplification, its envelope can remain almost unchanged in shape during propagation. The analytical expression of path-averaged soliton up to first-order correction [12] is

$$q(nL_a + \hat{L}, t) = a_0 e^{-\Gamma \hat{L}/2} (Q_0 + \tilde{Q}) \quad (15)$$

where  $\tilde{Q} = i\tilde{A}_1(\hat{L})|Q_0|^2 Q_0$ ,  $\hat{L}$  is the space coordinates between the  $i^{th}$  and  $(i+1)^{th}$  EDFA,  $0 < \hat{L} < L_a$ , and  $\tilde{A}_1(\hat{L})$  is a coefficient determined by fiber loss and span length.  $\Gamma$  is the normalized power loss coefficient given by  $\Gamma = \alpha L_D$ ,  $a_0$  the initial amplitude, and  $Q_0$  the ‘true’ soliton solution of NLSE.

Let  $Q_{PA}$  denotes  $Q_0 + \tilde{Q}$ , which is a true soliton  $Q_0$  perturbed by  $\tilde{Q}$ , and NFT was essentially performed on  $Q_{PA}$ . As shown in Fig. 9, we noticed that the evolution of  $\angle b(Q_{PA})$  does not follow the straight line of  $2(\beta^2 - \alpha^2)z$  given by the nonlinear spectral phase evolution from the NFT framework but evolves along a curved line. However, it eventually curves back and crosses the  $2(\beta^2 - \alpha^2)z$  straight line at the end of a span, which makes our forward evolution model and back rotation scheme still work.

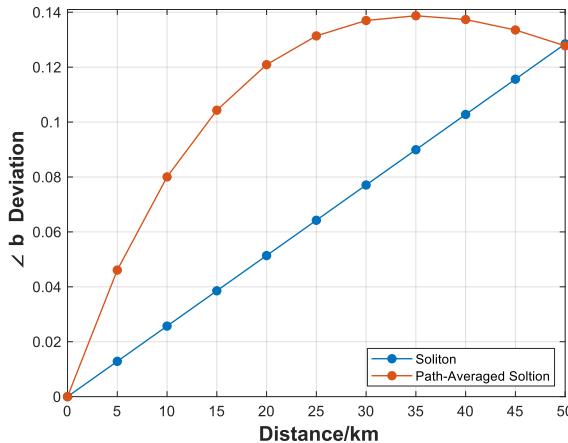


Fig. 9. Evolution of  $\angle b(Q_{PA})$  and  $\angle b(Q_0)$  within a span.

If the EDFA gain is slightly lower than the span loss, the pulse amplitude after each EDFA will gradually decrease, and we noticed the rotation of  $\angle b$  slows down during propagation, while  $\beta$  is still around 0.5 because we always conduct NFT to the power-normalized pulse,  $Q_{PA}$ . In such a case, we may consider the pulses as path-averaged solitons perturbated by (tiny) fiber loss. Actually, the normalized power loss coefficient  $\Gamma = \alpha L_D = 0.2 \times 194.6 / 4.343 = 8.96 > 1$ , which does not satisfy the loss perturbation condition of  $\Gamma \ll 1$  [23]. However, the periodic amplification essentially extend the distance over which the effect of fiber loss becomes significant to the path-averaged soliton. The gain-loss difference over a span may be interpreted as an “effective” perturbative loss (or gain) under the framework of path-averaged soliton. With the contribution of fiber loss and periodic EDFA amplification included, the normalized NLSE is

given by

$$i \frac{\partial q(z, t)}{\partial z} + \frac{1}{2} \frac{\partial^2 q(z, t)}{\partial t^2} + |q(z, t)|^2 q(z, t) \\ = i \left[ -\frac{\Gamma + \Gamma'}{2} + \frac{\Gamma}{2} L \sum_{n=1}^N \delta(z - nL) \right] q(z, t) \quad (16)$$

where  $\delta$  is the Dirac delta function modeling the EDFA amplification,  $\Gamma'$  represents the extra normalized loss/gain term originating from the loss-gain imbalance in the loop. Let  $q(z, t) = a(z)Q(z, t)$ , where  $a(z)$  satisfies

$$\frac{da(z)}{dz} = \left[ -\frac{\Gamma}{2} + \frac{\Gamma}{2} L \sum_{n=1}^N \delta(z - nL) \right] a(z) \quad (17)$$

then we have

$$i \frac{\partial Q(z, t)}{\partial z} + \frac{1}{2} \frac{\partial^2 Q(z, t)}{\partial t^2} + a^2(z)|Q(z, t)|^2 Q(z, t) \\ = -i \frac{\Gamma'}{2} Q(z, t) \quad (18)$$

$a^2(z)$  is fast-varying with respect to the dispersion length and may be replaced by its average value  $\langle a^2(z) \rangle$ , and for a path-average soliton, the launch power is increased such that  $\langle a^2(z) \rangle = \frac{1}{L} \int_0^L a^2(z) dz = 1$ , then (16) may be effectively interpreted as a path-average soliton transmission under perturbative fiber loss  $\Gamma'$ . Suppose the EDFA gain is, for example, 0.03 dB less than the span loss in the fiber loop, then  $\Gamma' = 0.03(\text{dB})L_D/L_A = 0.03/4.343 \times 194.6/50 = 0.027 \ll 1$ , which satisfies the criterion of perturbation. The phase evolution of soliton under perturbative attenuation is given by [24]

$$\varphi(z) = \varphi(0) + \frac{1 - \exp(-2\Gamma' z)}{4\Gamma'} \quad (19)$$

which exhibits a slowing-down trend and it matches quite well with the simulated phase evolution trace (ASE is turned off) as shown in Fig. 10.

As an illustrative example, Fig. 11(a) and (b) show the evolution and back rotation of  $\angle b$  and  $\ln|b|$ , respectively, with ASE turned on and loop gain set to be 0.03 dB less than loop loss. The evolution of  $\angle b$  slows down as predicted. However, the received eigenvalue is still around  $0 + 0.5j$ , which makes it “larger” than what is needed for back rotation and the back rotation line non-tangent to the forward evolution of  $\angle b$  any more. Thus, the optimal back rotation ratio for it will be less than 1/2, while the back rotation of  $\ln|b|$  is less affected because gain-loss imbalance does not affect the temporal position of a soliton. Similarly, if the gain is slightly higher than the loss in each span, the optimal back rotation ratio for  $\angle b$  will be more than 1/2.

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

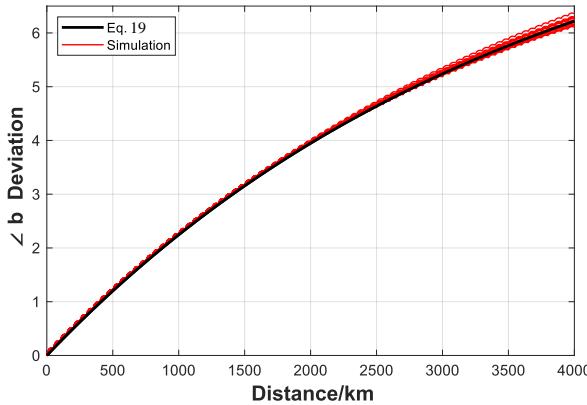


Fig. 10. Comparison of Eq. 14 and simulated  $\angle b(Q_{PA})$  trace with loop gain less than loop loss by 0.03dB, ASE free.

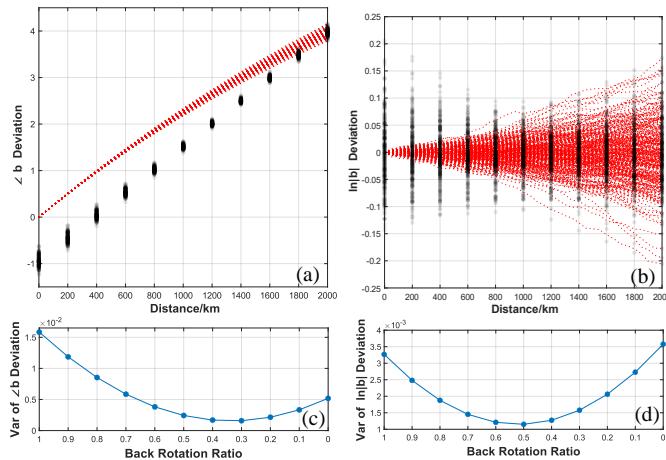


Fig. 11. (a) Forward evolution and (b) back rotation of  $\angle b$  and  $\ln|b|$  with loop gain set to be 0.03 dB less than loop loss. (c) and (d) show the corresponding variance versus the back rotation ratio.

## V. EXPERIMENTAL SETUP AND RESULTS

Experiment with offline digital signal processing (DSP) was conducted to verify the effectiveness of half back rotation. Fig. 12 shows the experimental setup and the offline DSP structure. 16APSK constellation was used for the b-coefficient modulation, with ring radii of  $\exp(\pm\beta_0 T_\Delta)$ . Thus, their corresponding normalized time-domain waveforms were centered at  $\pm T_\Delta/2$  respectively, where  $T_\Delta = 0.8$  was used in the experiment. At the transmitter, the corresponding time-domain waveform was obtained by INFT on the 16APSK symbols. After being added with synchronization head, it was loaded into the arbitrary waveform generator (Keysight M8196A, 90GSa/s, 3dB bandwidth of 35 GHz). The electrical waveform was then sent into an integrated coherent digital modulator (3dB bandwidth of 35 GHz) to modulate the optical carrier, and the modulated signal was then amplified by the Tx-EDFA and launched into the fiber loop. To generate path-averaged soliton, the launched power was increased by  $G \ln(G)/(G - 1)$  with respect to the theoretical power for generating an optical soliton, where  $G$  is the total loss of the 50-km fiber. The loop consisted of one span of 50-km NZ-DSF with a dispersion coefficient of 4 ps/km/nm and power loss coefficient of 0.2 dB/km, and a loop-EDFA to

compensate for the overall loss (about 19 dB) of the NZ-DSF, optical switch, and 50:50 coupler. A flat-top optical filter (3-dB bandwidth of 1 nm centered at  $\sim 1550$  nm) was used at the loop EDFA output to suppress out-of-band ASE noise. Both the transmitter laser and local oscillator shared the same fiber laser source (NKT Koheras BASIK Fiber laser, linewidth  $< 100$ Hz). After polarization alignment by a manual polarization controller, the received signal was coherently detected and then sampled by a digital oscilloscope (Keysight, UXR-0804A, 80 GSa/s, 63 GHz) and processed offline. The sampled signal passed through a series of DSP stages as shown in the inset of Fig. 12. After frequency offset (FO) compensation for the AOM-induced 100 MHz frequency shift and timing synchronization, the received signal was sliced into individual pulses and power-normalized, and NFT was then applied to obtain the eigenvalue and b-coefficient for each pulse, followed by back rotation and/or LMMSE filter. The baud rate was 2 Gbaud and the pulse FWHM was set to be 1/8 of the symbol period, leading to a dispersion length of 246 km that ensures path-averaged soliton can stably exist with the span length of 50 km. The normalization time parameter  $T_0$  was calculated to be 35.45 ps with the designed eigenvalue  $\lambda_0 = 0 + 0.5j$  and FWHM.

We conducted soliton transmission experiment over 1500 km with the parameters described above and back rotated the received noisy b-coefficients with received noisy eigenvalues. A fully trained LMMSE filter was also used after back rotation for performance comparison. Fig. 13(a) shows the BER at different ratios of back rotation with/without further processing by an LMMSE filter. The lowest BER happened at around 0.6 back rotation, but the LMMSE filter further decreased the BER after back rotation and returned the lowest BER at a back rotation ratio of 1. This is because in the experiment, the eigenvalue was perturbed by nonideal factors at the transmitter, which results in asymmetric constellation distributions during the process of forward evolution and also back rotation as discussed in Section IV. A. Note that in such a situation, back rotation has double effects, one is reducing the noise variance in each constellation cluster, for which a half back rotation will be optimal, while the other is bringing the mean of each constellation cluster back to the original symmetrically distributed position, for which a full back rotation will be optimal. These double effects make the overall optimal back rotation ratio greater than 1/2, which is a result of the trade-off between the variance within each cluster and the variance among clusters. Besides, we noticed that in the experiment,  $|b|$  also became asymmetry, which indicates a split of  $a$  at the transmitter. Although the cause needs further investigation, the back rotation effect on  $\ln|b|$  is similar to that on  $\angle b$ , so a full back rotation restores circular symmetry to the constellation both in radial and angular direction, while the noise in each cluster is raised again due to the ‘excessive’ back rotation. However, since the LMMSE filter can remove the noise within each cluster resulting from the ‘excessive’ back rotation, the BER can be further reduced when the back rotation ratio is over 1/2. To isolate the influence of

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

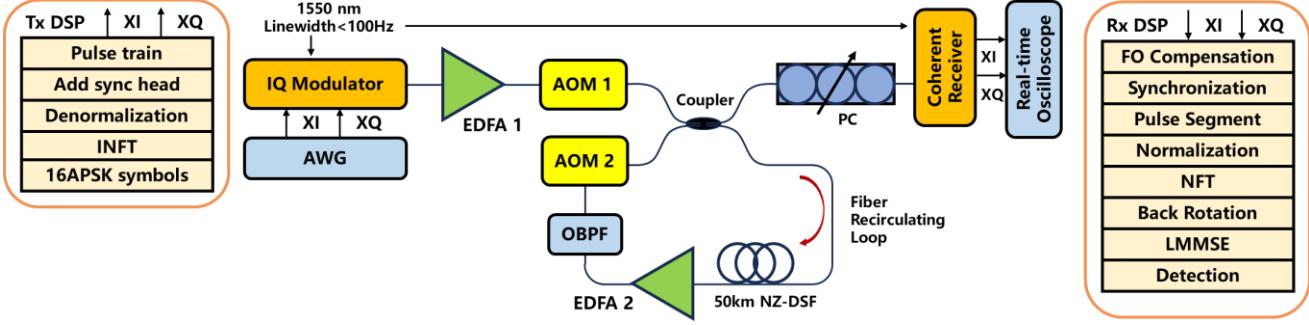


Fig. 12. Experimental setup for discrete eigenvalue transmission.

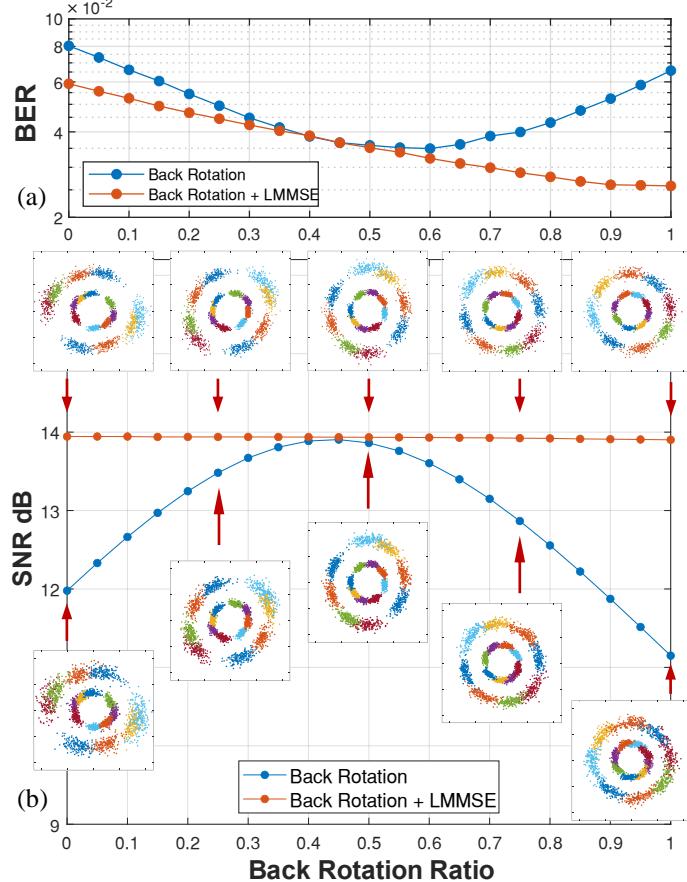


Fig. 13. (a) BER at different ratios of back rotation with/without further applying an LMMSE filter; (b) SNR at different ratios of back rotation with/without further applying an LMMSE filter. Insets are corresponding signal distributions.

constellation asymmetry, we calculated the signal SNR with respect to each cluster's mean point and obtained a flat SNR curve versus different back rotation ratios when further applying an LMMSE filter as shown in Fig. 13(b), which is consistent with the simulation results in Section III.

## VI. CONCLUSIONS

We propose back rotating the received noisy b-coefficient by half of the transmission distance with the received noisy eigenvalues for discrete eigenvalue transmission. The effectiveness of half back rotation was derived analytically and verified by simulation and experiment. It gives the same level of performance compared to the LMMSE filter but with less calculation complexity and no need to use training data.

We also derive and discuss the sensitivity of the proposed back rotation algorithm to nonideal practical factors such as non-ideal bias control and gain-loss imbalance. Extensions to distributed amplification systems and dual-polarization systems are expected to be natural and straightforward.

## APPENDIX

### A. Approximation of $(\alpha_i + j\beta_i)^2$

Let  $\tilde{\alpha}_i = \alpha_i - \alpha_0$ ,  $\tilde{\beta}_i = \beta_i - \beta_0$  denote the accumulated noise  $\alpha$  and  $\beta$  at the  $i^{\text{th}}$  EDFA. We have

$$b_{i+1} = b_i e^{-2j(\alpha_i + j\beta_i)^2 L} = b_i e^{4\alpha_i \beta_i L} e^{-2j(\alpha_i^2 - \beta_i^2)L} \quad (20)$$

for the index term, since both  $\tilde{\alpha}_i, \tilde{\beta}_i \approx 0$ , while  $\beta_i \approx 0.5$ , we approximately have

$$\begin{aligned} 4\alpha_i \beta_i &= 4(\tilde{\alpha}_i + \alpha_0)(\tilde{\beta}_i + \beta_0) \\ &= 4(\tilde{\alpha}_i \tilde{\beta}_i + \alpha_0 \tilde{\beta}_i + \tilde{\alpha}_i \beta_0 + \alpha_0 \beta_0) \\ &\approx 4(\tilde{\alpha}_i \beta_0) = 2\tilde{\alpha}_i = 2\alpha_i \end{aligned} \quad (21)$$

$$\begin{aligned} \alpha_i^2 - \beta_i^2 &= (\tilde{\alpha}_i + \alpha_0)^2 - \beta_i^2 \\ &= \tilde{\alpha}_i^2 - 2\tilde{\alpha}_i \alpha_0 + \alpha_0^2 - \beta_i^2 \\ &\approx -\beta_i^2 \end{aligned} \quad (22)$$

where  $\tilde{\alpha}_i \tilde{\beta}_i$  is negligible.

### B. Variance of $\ln|b(x)|$

Let  $\Delta\alpha_i = \alpha_i - \alpha_{i-1}$  denote the independent Gaussian noise with a variance of  $\sigma_\alpha^2$  added to  $\alpha$  at the  $i^{\text{th}}$  EDFA. We have

$$\begin{aligned} \text{Var}(r(x)) &= \text{Var}(2L\alpha_{N+1}(x-N) + r_{N+1}) \\ &= 4L^2(x-N)^2 \text{Var}(\alpha_{N+1}) + \text{Var}(r_{N+1}) \\ &\quad + 4L(x-N)\text{Cov}(r_{N+1}, \alpha_{N+1}) \end{aligned} \quad (23)$$

where

$$\begin{aligned} \text{Var}(\alpha_{N+1}) &= \text{Var}\left(\sum_{i=1}^{N+1} \Delta\alpha_i\right) \\ &= \sum_{i=1}^{N+1} \text{Var}(\Delta\alpha_i) + \sum_{i \neq j} \text{Cov}(\Delta\alpha_i, \Delta\alpha_j) \\ &= (N+1)\sigma_\alpha^2 \end{aligned} \quad (24)$$

$$\begin{aligned} \text{Var}(r_{N+1}) &= \text{Var}\left(\ln|b_0| + 2L \sum_{i=1}^N \alpha_i\right) \\ &= 4L^2 \text{Var}\left(\sum_{i=1}^N \alpha_i\right) \\ &= 4L^2 \sum_{i,j=1}^N \text{Cov}(\alpha_i, \alpha_j) \end{aligned}$$

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

$$= 4L^2\sigma_\alpha^2 \sum_{i,j=1}^N \min(i,j) \\ = 4L^2\sigma_\alpha^2 \frac{N(N+1)(2N+1)}{6} \quad (25)$$

$$\begin{aligned} \text{Cov}(r_{N+1}, \alpha_{N+1}) &= \mathbf{E}[r_{N+1}\alpha_{N+1}] - \mathbf{E}[r_{N+1}]\mathbf{E}[\alpha_{N+1}] \\ &= \mathbf{E}\left[2L \sum_{i=1}^N \alpha_i \alpha_{N+1}\right] \\ &= 2L \sum_{i=1}^N \mathbf{E}[\alpha_i \alpha_{N+1}] \\ &= 2L\sigma_\alpha^2 \sum_{i=1}^N i \\ &= N(N+1)L\sigma_\alpha^2 \end{aligned} \quad (26)$$

Thus,

$$\text{Var}(r(x)) = 4L^2\sigma_\alpha^2(N+1) \left[ x(x-N) + \frac{N(2N+1)}{6} \right] \quad (27)$$

### C. Variance of $\angle b(x)$

$\beta_i$  may be regarded as  $\beta_i = w_i + 0.5$ , where  $w_i$  is a zero-mean ‘discrete’ Wiener process, and we obtain  $\beta_i^2 = w_i^2 + w_i + 0.25 \approx w_i + 0.25$ . Let  $\Delta w_i = w_i - w_{i-1}$  denote the independent Gaussian noise with a variance of  $\sigma_\beta^2$  added to  $\beta$  at the  $i^{\text{th}}$  EDFA. We have

$$\begin{aligned} \text{Var}(\varphi(x)) &= \text{Var}(2L\beta_{N+1}^2(x-N) + \varphi_{N+1}) \\ &= 4L^2(x-N)^2\text{Var}(\beta_{N+1}^2) + \text{Var}(\varphi_{N+1}) \\ &\quad + 4L(x-N)\text{Cov}(\varphi_{N+1}, \beta_{N+1}^2) \end{aligned} \quad (28)$$

where

$$\begin{aligned} \text{Var}(\beta_{N+1}^2) &\approx \text{Var}(w_{N+1} + 0.25) \\ &= \text{Var}\left(\sum_{i=1}^{N+1} \Delta w_i\right) \\ &= (N+1)\sigma_\beta^2 \end{aligned} \quad (29)$$

$$\begin{aligned} \text{Cov}(\varphi_{N+1}, \beta_{N+1}^2) &= \text{Cov}\left(2L \sum_{i=1}^N \beta_i^2, \beta_{N+1}^2\right) \\ &\approx 2L \sum_{i=1}^N \text{Cov}(w_i + 0.25, w_{N+1} + 0.25) \\ &= 2L \sum_{i=1}^N \text{Cov}(w_i, w_{N+1}) \\ &= 2L\sigma_\beta^2 \sum_{i=1}^N i \\ &= N(N+1)L\sigma_\beta^2 \end{aligned} \quad (30)$$

$$\begin{aligned} \text{Var}(\varphi_{N+1}) &= 4L^2\text{Var}\left(\sum_{i=1}^N \beta_i^2\right) \\ &\approx 4L^2 \sum_{i,j=1}^N \text{Cov}(w_i, w_j) \\ &= 4L^2\sigma_\beta^2 \sum_{i,j=1}^N \min(i,j) \\ &= 4L^2\sigma_\beta^2 \frac{N(N+1)(2N+1)}{6} \end{aligned} \quad (31)$$

Thus,

$$\text{Var}(\varphi(x)) \approx 4L^2\sigma_\beta^2(N+1) \left[ x(x-N) + \frac{N(2N+1)}{6} \right] \quad (32)$$

### D. Correlation Coefficient, $\rho(r(x), \alpha_{N+1})$ , $\rho(\varphi(x), \beta_{N+1})$

$$\begin{aligned} \rho(r(x), \alpha_{N+1}) &= \frac{\text{Cov}(r(x), \alpha_{N+1})}{\sqrt{\text{Var}(r(x)) \cdot \text{Var}(\alpha_{N+1})}} \\ &= \frac{\text{Cov}(2L\alpha_{N+1}(x-N) + r_{N+1}, \alpha_{N+1})}{\sqrt{\text{Var}(r(x)) \cdot \text{Var}(\alpha_{N+1})}} \\ &= \frac{2L(x-N)\text{Var}(\alpha_{N+1}) + \text{Cov}(r_{N+1}, \alpha_{N+1})}{\sqrt{\text{Var}(r(x)) \cdot \text{Var}(\alpha_{N+1})}} \\ &= \frac{x-N/2}{\sqrt{x(x-N) + N(2N+1)/6}} \end{aligned} \quad (33)$$

$$\begin{aligned} \rho(\varphi(x), \beta_{N+1}) &= \frac{\text{Cov}(\varphi(x), \beta_{N+1})}{\sqrt{\text{Var}(\varphi(x)) \cdot \text{Var}(\beta_{N+1})}} \\ &= \frac{\text{Cov}(2L\beta_{N+1}^2(x-N) + \varphi_{N+1}, \beta_{N+1})}{\sqrt{\text{Var}(\varphi(x)) \cdot \text{Var}(\beta_{N+1})}} \\ &= \frac{2L(x-N)\text{Cov}(\beta_{N+1}^2, \beta_{N+1}) + \text{Cov}(\varphi_{N+1}, \beta_{N+1})}{\sqrt{\text{Var}(\varphi(x)) \cdot \text{Var}(\beta_{N+1})}} \\ &\approx \frac{x-N/2}{\sqrt{x(x-N) + N(2N+1)/6}} \end{aligned} \quad (34)$$

where we used

$$\begin{aligned} \text{Cov}(\beta_i^2, \beta_{N+1}) &\approx \text{Cov}(w_i + 0.25, w_{N+1} + 0.5) \\ &= \text{Cov}(w_i, w_{N+1}) = i\sigma_\beta^2 \end{aligned} \quad (35)$$

and

$$\begin{aligned} \text{Cov}(\varphi_{N+1}, \beta_{N+1}) &= \text{Cov}\left(2L \sum_{i=1}^N \beta_i^2, \beta_{N+1}\right) \\ &= 2L \sum_{i=1}^N \text{Cov}(\beta_i^2, \beta_{N+1}) \\ &\approx N(N+1)L\sigma_\beta^2. \end{aligned} \quad (36)$$

### E. Non-ideal Bias Control Case

With split  $\beta_0$  due to non-ideal bias control and/or low OSNR at the transmitter, the back rotation line becomes

$$\tilde{\varphi}(x) = 2L\tilde{\beta}_{N+1}^2(x-N) + \tilde{\varphi}_{N+1} \quad (37)$$

where (with seconder order terms neglected)

$$\begin{aligned} \tilde{\beta}_{N+1}^2 &= (\beta_{N+1} + \Delta\beta_0)^2 \\ &\approx \beta_{N+1}^2 + 2\beta_{N+1}\Delta\beta_0 \\ &= w_{N+1}^2 + w_{N+1} + 0.25 + 2(w_{N+1} + 0.5)\Delta\beta_0 \\ &\approx w_{N+1} + \Delta\beta_0 + 0.25 \end{aligned} \quad (38)$$

$$\begin{aligned} \tilde{\varphi}_{N+1} &= 2L \sum_{i=1}^N \tilde{\beta}_i^2 \\ &\approx 2L \sum_{i=1}^N (w_i + \Delta\beta_0 + 0.25) \\ &= 2L \sum_{i=1}^N w_i + 2NL\Delta\beta_0 + NL/2 \end{aligned} \quad (39)$$

Thus,

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

$$\begin{aligned} \text{Var}(\tilde{\varphi}(x)) &= \text{Var}(2L\tilde{\beta}_{N+1}^2(x-N) + \tilde{\varphi}_{N+1}) \\ &= 4L^2(x-N)^2\text{Var}(\tilde{\beta}_{N+1}^2) + \text{Var}(\tilde{\varphi}_{N+1}) \\ &\quad + 4L(x-N)\text{Cov}(\tilde{\varphi}_{N+1}, \tilde{\beta}_{N+1}^2) \end{aligned} \quad (40)$$

with

$$\begin{aligned} \text{Var}(\tilde{\beta}_{N+1}^2) &\approx \text{Var}(w_{N+1} + 0.25 + \Delta\beta_0) \\ &= \text{Var}(w_{N+1}) + \text{Var}(\Delta\beta_0) \\ &= (N+1)\sigma_\beta^2 + \sigma_{\beta_0}^2 \end{aligned} \quad (41)$$

$$\begin{aligned} \text{Cov}(\tilde{\varphi}_{N+1}, \tilde{\beta}_{N+1}^2) &= 2L \sum_{i=1}^N \text{Cov}(\tilde{\beta}_i^2, \tilde{\beta}_{N+1}^2) \\ &\approx 2L \sum_{i=1}^N \text{Cov}(w_i + \Delta\beta_0, w_{N+1} + \Delta\beta_0) \\ &= 2L \sum_{i=1}^N (\text{cov}(w_i, w_{N+1}) + \text{Var}(\Delta\beta_0)) \\ &= N(N+1)L\sigma_\beta^2 + 2NL\sigma_{\beta_0}^2 \end{aligned} \quad (42)$$

let  $\frac{d}{dx} \text{Var}(\tilde{\varphi}(x)) = 0$ , we obtain

$$x_{opt} = N - \frac{\text{Cov}(\tilde{\varphi}_{N+1}, \tilde{\beta}_{N+1}^2)}{\text{Var}(\tilde{\beta}_{N+1}^2)2L} = \frac{N(N+1)\sigma_\beta^2}{2[(N+1)\sigma_\beta^2 + \sigma_{\beta_0}^2]} \quad (43)$$

#### ACKNOWLEDGMENT

The authors would like to thank the reviewers for their valuable feedback and comments, thorough review, and constructive suggestions, which improved the quality and clarity of our work. We appreciate their time and effort in helping us improve our research work.

#### REFERENCES

- [1] A. Hasegawa and T. Nyu, "Eigenvalue communication," *J. Lightwave Technol.*, vol. 11, no. 3, pp. 395–399, Mar. 1993.
- [2] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Part I: Mathematical Tools," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014.
- [3] H. Terauchi and A. Maruta, "Eigenvalue Modulated Optical Transmission System Based on Digital Coherent Technology," in 2013 18th OptoElectronics and Communications Conference held jointly with 2013 International Conference on Photonics in Switching, Kyoto: OSA, 2013, p. WR2\_5.
- [4] A. Maruta, A. Toyota, Y. Matsuda, and Y. Ikeda, "Experimental demonstration of long haul transmission of eigenvalue modulated signals," in 2015 Tyrrhenian International Workshop on Digital Communications (TIWDC), Florence, Italy: IEEE, Sep. 2015, pp. 28–30.
- [5] H. Bulow, "Experimental Demonstration of Optical Signal Detection Using Nonlinear Fourier Transform," *J. Lightwave Technol.*, vol. 33, no. 7, pp. 1433–1439, Apr. 2015.
- [6] Z. Dong et al., "Nonlinear Frequency Division Multiplexed Transmissions Based on NFT," *IEEE Photon. Technol. Lett.*, vol. 27, no. 15, pp. 1621–1623, Aug. 2015.
- [7] V. Aref, H. Bulow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in 2015 European Conference on Optical Communication (ECOC), Valencia, Spain: IEEE, Sep. 2015, pp. 1–3.
- [8] T. Gui, T. H. Chan, C. Lu, A. P. T. Lau, and P.-K. A. Wai, "Alternative Decoding Methods for Optical Communications Based on Nonlinear Fourier Transform," *J. Lightwave Technol.*, vol. 35, no. 9, pp. 1542–1550, May 2017.
- [9] F. Da Ros et al., "Dual-Polarization NFDM Transmission With Continuous and Discrete Spectral Modulation," *J. Lightwave Technol.*, vol. 37, no. 10, pp. 2335–2343, May 2019.
- [10] G. Zhou, L. Sun, C. Lu, and A. P. T. Lau, "Multi-Symbol Digital Signal Processing Techniques for Discrete Eigenvalue Transmissions Based on Nonlinear Fourier Transform," *J. Lightwave Technol.*, vol. 39, no. 17, pp. 5459–5467, Sep. 2021.
- [11] S. Gajarin, F. Da Ros, N. De Renzis, E. P. da Silva, and D. Zibar, "Dual-Polarization NFDM Transmission Using Distributed Raman Amplification and NFT-Domain Equalization," *IEEE Photon. Technol. Lett.*, vol. 30, no. 22, pp. 1983–1986, Nov. 2018.
- [12] A. Hasegawa and Y. Kodama, "Guiding-center soliton in optical fibers," *Opt. Lett.*, vol. 15, no. 24, p. 1443, Dec. 1990.
- [13] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," *Opt. Express*, vol. 23, no. 7, p. 8317, Apr. 2015.
- [14] J. Koch, K. Chan, C. G. Schaeffer, and S. Pachnicke, "Signal Processing Techniques for Optical Transmission Based on Eigenvalue Communication," *IEEE J. Select. Topics Quantum Electron.*, vol. 27, no. 3, pp. 1–14, May 2021.
- [15] C. Xu and A. P. T. Lau, "Optimal Nonlinear Spectral Back-Rotation for Discrete Eigenvalue NFT Transmission Systems," in Optical Fiber Communication Conference (OFC) 2024, San Diego California: Optica Publishing Group, 2024, p. Th2A.10.
- [16] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Part III: Spectrum Modulation," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.
- [17] Q. Zhang and T. H. Chan, "A Gaussian noise model of spectral amplitudes in soliton communication systems," in 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden: IEEE, Jun. 2015, pp. 455–459.
- [18] Q. Zhang and T. H. Chan, "A spectral domain noise model for optical fibre channels," in 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, Hong Kong: IEEE, Jun. 2015, pp. 1660–1664.
- [19] S. A. Derevyanko, S. K. Turitsyn, and D. A. Yakushev, "Non-Gaussian statistics of an optical soliton in the presence of amplified spontaneous emission," *Opt. Lett.*, vol. 28, no. 21, p. 2097, Nov. 2003.
- [20] J. P. Gordon and H. A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission," *Opt. Lett.*, vol. 11, no. 10, p. 665, Oct. 1986.
- [21] D. Marcuse, "An alternative derivation of the Gordon-Haus effect," *J. Lightwave Technol.*, vol. 10, no. 2, pp. 273–278, Feb. 1992.
- [22] K.-P. Ho and J. M. Kahn, "Electronic Compensation Technique to Mitigate Nonlinear Phase Noise," *J. Lightwave Technol.*, vol. 22, no. 3, pp. 779–783, Mar. 2004.
- [23] G. P. Agrawal, *Nonlinear fiber optics*, Fifth edition. Amsterdam: Elsevier/Academic Press, 2013.
- [24] A. Hasegawa and Y. Kodama, "Signal transmission by optical solitons in monomode fiber," *Proc. IEEE*, vol. 69, no. 9, pp. 1145–1150, 1981.

**Chuang Xu** was born in Xiaogan, China in 1993. He received his Bachelor and Master degrees of Engineering in Material Science and Engineering from Wuhan University of Technology, Wuhan, Hubei, China, in 2014 and 2017, respectively. He worked as a fiber communication engineer at FiberHome from 2017 to 2021 and currently is a PhD student of the Department of Electrical and Electronic Engineering at The Hong Kong Polytechnic University. His research interest includes fiber nonlinearity mitigation in digital coherent communication systems.

**Alan Pak Tao Lau** received the B.A.Sc. degree in engineering science (electrical option) and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2003 and 2004, respectively, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA, USA, in 2008. He then joined The Hong Kong Polytechnic University as an Assistant Professor, and he is currently Professor and Associate Head (Promotions and Global Relations) of the

>REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT)<

Department of Electrical and Electronic Engineering. He was also a Visiting Professor at Stanford University in 2019. His current research interests include Digital Signal Processing and Machine Learning techniques for various aspects of optical communication systems and networks, optical performance monitoring as well as optical fiber sensing. He collaborates extensively with industry and serves as an invited/tutorial speaker, organizing committee chair for Optical Fiber Communications (OFC), and Associate Editor of the Journal of Lightwave Technology (JLT) among others. He is a fellow of OPTICA and currently serves as a member of the Publication Council of the IEEE Photonics Society.