The following publication 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," in Journal of Lightwave Technology, vol. 39, no. 17, pp. 5459-5467, 1 Sept.1, 2021 is available at https://doi.org/10.1109/JLT.2021.3084825

Multi-symbol digital signal processing techniques for discrete eigenvalue transmissions based on Nonlinear Fourier Transform

Gai Zhou, Lin Sun, Chao Lu and Alan Pak Tao Lau

***Abstract*— Optical communications based on Nonlinear Fourier Transform (NFT) and digital coherent transceivers are proposed as a new theoretical framework for communications over the nonlinear optical fiber channel. For discrete eigenvalue transmissions (or soliton transmissions), one seeks to encode as much information as possible in each degree of freedom and shorten the distance between neighboring pulses to increase the overall bit rate. However, such attempts would result in nonlinear inter-symbol interference (ISI) across multiple symbols and significantly degrade transmission performance. In this paper, we investigated joint modulation of discrete eigenvalue λ and b-coefficents $b(\lambda)$ and developed a suite of multi-symbol digital signal processing (DSP) techniques to exploit the statistical correlations between the continuous and discrete eigenvalues and b-coefficents to mitigate nonlinear distortions and improve detection performance. This include 1) jointly modulating both λ and $b(\lambda)$ of pairs of 1-solitons so that the mean value of λ for solitons with odd index is $\alpha + 1i$ while it is $-\alpha + 1i$ for solitons with even index. This is followed by decoding superimposed received waveforms as 2-solitons with twice the INFT processing time window; 2) linear minimum mean squared error (LMMSE) estimation filters to mitigate noise in discrete eigenvalue λ using continuous eigenvalue; 3) multi-symbol (MS) LMMSE filters to mitigate noise in $b(\lambda)$ using discrete eigenvalue noise and 4) approximate the received signal distributions of λ and $b(\lambda)$ as Gaussians with mean and covariance matrices obtained empirically from experiments followed by Maximum Likelihood (ML) detection for each symbol or multi-symbol (MS)-joint ML detection of 2-soliton signals. We jointly modulate λ with 16-QAM and $b(\lambda)$ with 16-APSK and a record single-polarization discrete eigenvalue transmission of 64 Gb/s (net 54 Gb/s) over 1200 km is experimentally demonstrated with the proposed multi-symbol DSP algorithms.[1]**

***Index Terms*—optical communications, nonlinear optics**

I. Introduction

In long-haul optical fiber communications, Kerr nonlinearity induces complicated interactions between signal and noise and become the fundamental bottleneck to transmission speeds. Recently, Nonlinear Fourier transform (NFT) and nonlinear frequency division multiplexing (NFDM) framework decomposes the fiber channel characterized by the Nonlinear Schrodinger Equation (NLSE) into parallel channels and decomposes the signal into nonlinear spectral components, which can be obtained by solving the Zakharov-Shabat problem from the Lax pair operators corresponding to the NLSE [1]. The NFT/inverse-NFT (INFT) operators essentially transform the signal between time-domain and nonlinear frequency domain. In principle, one can encode information on nonlinear spectral components and they can be multiplexed together without mutual interference along fiber propagation. The nonlinear spectrum consists of discrete eigenvalues (or discrete spectrum) and continuous spectrum. From a communications perspective, continuous spectrum modulation resembles orthogonal frequency division multiplexing (OFDM) in linear systems while discrete eigenvalue modulation can be regarded as encoding information on physical parameters of (multi-) soliton pulses. Discrete eigenvalue modulation has been experimentally demonstrated in direct detection systems in the 1990s [2]. In recent years, coherent detection and advanced digital signal processing (DSP) techniques, or digital coherent transceivers, help revived NFT transmission research [3-6]. Since our first NFT transmission demonstration with 3 and 4 eigenvalues [7], there have been many promising experimental demonstrations of continuous-spectrum modulation [8-11], discrete-eigenvalue modulation [12-15] and both [16, 17].

In current discrete eigenvalue modulation systems, solitons are generated and processed in separated time windows. Neighboring solitons require certain guard time to prevent mutual interactions during propagation as small overlapping can result in significant distortions especially for high baud-rate and long-haul transmissions. However, larger guard time will reduce overall baud rate hence bit rate. To partially address this problem, we proposed [15] to encode information onto the

The authors acknowledge the support of National Key R&D Program of China (2019YFB1803502) and Hong Kong Government Research Grants Council General Research Fund (GRF) under Project PolyU 15220120 and 152757/16E. (Corresponding author: Lin Sun)

Gai Zhou, and Alan Pak Tao Lau are with the Photonics Research Center, Department of Electrical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong and The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China (e-mail: gai.zhou@connect.polyu.hk; eeaptlau@polyu.edu.hk).

Lin Sun is with The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China. (e-mail: linsanity.sun@polyu.edu.hk)

Chao Lu is with the Photonics Research Center, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong. (e-mail: enluchao@polyu.edu.hk).

phase and amplitude of the b-coefficient $b(\lambda)$ and design the eigenvalues of neighboring soliton pairs $\lambda = \pm\alpha + i\beta$ to have opposite signs in their real part as shown in Fig. 1 . In this case, solitons will walk toward each other, collide and then walk pass each other during transmission. At a particular distance after such collision, pairs of neighboring solitons will essentially swap their positions with each other and one can detect them as a regular received soliton pulse train. Despite the soliton interactions, we showed theoretically and experimentally that such encoding and decoding scheme significantly reduces noise and distortions compared to traditional non-interacting soliton pulse trains with the same eigenvalue across all soliton pulses. This can be viewed as a hybrid TDM/WDM in the Nonlinear frequency domain. This collision-separation process can be repeated and pairs of solitons will swap positions again at some further distances.

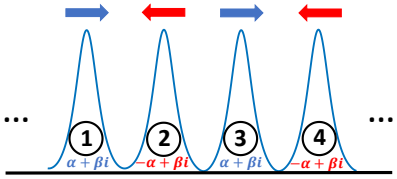

Fig. 1. For pulse train at transmitter, odd- and even-index solitons have $\lambda = \pm\alpha + \beta j$ respectively. Arrows indicate walk-off direction of each soliton during propagation.

A limitation of our previous scheme is that for a given α, there is only a few specific distances ℓ that solitons interacted during propagation but well-separated at receiver for 1-soliton pulse detection. Furthermore, the soliton interactions induce considerable noise correlations between different spectral components and across neighboring solitons. In this connection, we hereby propose a few multi-symbol DSP algorithms to exploit these statistical correlation structures to further improve detection performance. In particular, we propose to:

1. jointly modulate λ and $b(\lambda)$ of pairs of 1-solitons with the average eigenvalue $\mathbf{E}[\lambda] = \pm\alpha + 1i$ for odd- and even-indexed solitons across the transmitted pulse train. Here, $\mathbf{E}[\cdot]$ denotes the mean operator;
2. generate 1-solitons with symbol period (INFT processing time window) T but decode superimposed received waveforms as 2-solitons over a 2T time interval (NFT processing time window);
3. develop a linear minimum mean-squared estimation (LMMSE) filter to estimate and compensate the discrete eigenvalue noise $\Delta\lambda$ using the a-coefficient $a(\lambda)$ of the continuous spectrum (CS);
4. develop a multi-symbol (MS)-LMMSE filter to estimate and compensate the noise of $b(\lambda)$ using discrete eigenvalue noise $\Delta\lambda$ across neighboring solitons;
5. approximate the received signal distributions of λ and $b(\lambda)$ as Gaussians with joint mean and covariance matrices of each signal point obtained empirically from experiments. This is followed by Maximum Likelihood (ML) detection of λ and $b(\lambda)$ for each symbol or multi-symbol (MS)-ML detection of λ and $b(\lambda)$ for 2-soliton signals.

By incorporating all the DSP techniques, we experimentally demonstrated 64 Gb/s (net 54 Gb/s) transmission over 1200 km using 8 GBaud solitons with joint 16-Quadrature Amplitude Modulation (QAM) modulation on λ and 16-Amplitude Phase Shift Keying (APSK) modulation on $b(\lambda)$. To the best of our knowledge, this is a new bit-rate-distance record for single-polarization discrete eigenvalue transmissions and the multi-symbol DSP framework provides a new direction for DSP research to further improve NFT transmission performance.

The rest of the paper is organized as follows. Section II provides the background of NFT and describes the various proposed noise reduction filters and detection methods. Section III describes the experimental setup and details the transmission results. Section IV concludes the paper.

II. Nonlinear Fourier Transform and Multiple-Symbol DSP

A. Nonlinear Fourier Transform and Nonlinear Frequency Division Multiplexing (NFDM)

Without fiber loss and high order nonlinearity, optical signals propagating in optical fiber is governed by nonlinear Schrodinger Equation (NLSE):

$$j\frac{\partial}{\partial z}A(t,z) - \frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}A(t,z) + \gamma|A(t,z)|^2A(t,z) = 0 \tag{1}$$

where β_2, γ denotes group-velocity dispersion (GVD) coefficent and fiber nonlinearity coefficient. The variables t, z, A are physical time, distance and signal waveform. Those variables can be normalized as

$$\ell = \frac{z}{L_D}, \tau = \frac{t}{T_0}, L_D = \frac{2T_0^2}{|\beta_2|}, u = \sqrt{\gamma L_D}A \tag{2}$$

where τ, ℓ, u are normalized time, distance and signal envelope (T_0 is a free normalization parameter). The NFT of a normalized signal $u(\tau)$ ($\tau \in [T_1, T_2]$) is defined by solving the differential system (Zakharov-Shabat system)

$$\frac{\partial v}{\partial \tau} = \begin{pmatrix} -j\lambda & u(\tau) \\ -u^*(\tau) & j\lambda \end{pmatrix} \tag{3}$$

$$v(T_1,\lambda) = \begin{pmatrix} v_1(T_1,\lambda) \\ v_2(T_1,\lambda) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} exp\,(-j\lambda T_1) \tag{4}$$

where λ and $v(\tau,\lambda)$ are eigenvalue and eigenvector. The nonlinear spectral coefficients are defined as

$$\begin{aligned} a(\lambda) &= v_1(T_2,\lambda)\,exp(j\lambda T_2) \\ b(\lambda) &= v_2(T_2,\lambda)\,exp(-j\lambda T_2) \end{aligned} \tag{5}$$

The nonlinear spectrum of $u(\tau,\ell)$ is made of 1) continuous spectrum $q(\lambda) = b(\lambda)/a(\lambda)$ for $\lambda \in \mathbb{R}$. And 2) discrete spectrum $q(\lambda) = b(\lambda)/a'(\lambda)$ where $a(\lambda) = 0$ for $\lambda \in \mathbb{C}^+$. (The case when $a(\lambda)$ have multiple zeros at a particular eigenvalue are not considered in this paper and more details can be found in [18]).

A key property of NFT signals is that nonlinear spectrum propagates without mutual interference and they evolve with distance as

$$q(\lambda,\ell) = q(\lambda,0)exp(4j\lambda^2\ell) \tag{6}$$

with ℓ being the normalized distance. Similar to OFDM in linear systems, NFDM suggests that independent information be encoded on different parts of the nonlinear spectra so that they will not mutually interfere along propagation and at the receiver. Improving efficiency and accuracy of NFT [19-22]

and INFT [23-25] algorithms are important topics in current NFDM research. Meanwhile, encoding information in $b(\lambda)$ for discrete and continuous spectra is proposed to reduce noise and control signal time-duration [26, 27]. LMMSE algorithm is proposed to reduce noise on $b(\lambda)$ and λ [26, 28, 29]. Nonlinear Volterra filters are also studied [30]. Periodic NFT/INFT is proposed to mitigate inter symbol interference [31, 32]. New encoding and decoding schemes in nonlinear spectrum have been studied [26, 33, 34]. Some receivers are also proposed to detect NFDM signals instead of NFT. Detecting signals by similarity between received and transmitted waveform is studied [35, 36]. Several neural networks receivers are also exploited in NFDM signal detection [37-40] and end-to-end learning demo based on discrete spectrum modulation is also studied [41]. In our previous paper [26], we showed that amplitude and phase noise of b -coefficient $\Delta|b(\lambda)|$ and $\Delta\angle b(\lambda)$ is correlated to eigenvalue noise $\Delta\lambda$ and an LMMSE estimator was proposed to estimate and compensate noise on $b(\lambda)$ based on $\boldsymbol{\Delta}_\lambda = [\Re\{\Delta\lambda\} \quad \Im\{\Delta\lambda\}]$. Here, $\Re[\cdot]$ and $\Im[\cdot]$ denote the real and imaginary part. Specifically, we show that the LMMSE estimator

$$\begin{aligned} \boldsymbol{c_a} &= \mathbf{E}(\Delta|b(\lambda)|\boldsymbol{\Delta}_\lambda)cov^{-1}\left(\boldsymbol{\Delta}_\lambda^{\mathrm{T}}\boldsymbol{\Delta}_\lambda\right) \\ \boldsymbol{c_p} &= \mathbf{E}(\Delta\angle b(\lambda)\boldsymbol{\Delta}_\lambda)cov^{-1}\left(\boldsymbol{\Delta}_\lambda^{\mathrm{T}}\boldsymbol{\Delta}_\lambda\right) \end{aligned} \tag{7}$$

minimizes the mean squared error $\mathbf{E}[(\Delta|b(\lambda)| - \boldsymbol{c_a}\boldsymbol{\Delta}_\lambda)^2]$ and $\mathbf{E}[\left(\Delta\angle b(\lambda) - \boldsymbol{c_p}\boldsymbol{\Delta}_\lambda\right)^2]$ where $\mathbf{E}[\cdot]$ and $cov^{-1}(\cdot)$ denotes expectation and inverse of covariance matrix respectively. In the following, we will detail a few new DSP techniques that further leverage the statistical properties across neighboring solitonic pulses in discrete eigenvalue transmissions.

B. 1-soliton generation and 2-soliton detection

Consider a soliton pulse train where the i^{th} information symbol is drawn from an alphabet $m_{i,\lambda} \in \{1,2,\dots,M_\lambda\}$ and $m_{i,b} \in \{1,2,\dots,M_b\}$ are encoded λ and $b(\lambda)$ respectively. Each symbol carries $log_2(M_\lambda \cdot M_b)$ bits and $m_{i,\lambda}$ and $m_{i,b}$ uniquely maps to a point in a signal constellation through mapping functions $\lambda_i = \psi_\lambda(m_{i,\lambda})$ and $b_i(\lambda_i) = \psi_b(m_{i,b})$. It should be noted that in our paper, the modulation, detection and DSP for $b(\lambda)$ is performed in log scale. We have chosen the modulation format to be QAM and APSK for λ and $b(\lambda)$ respectively. Furthermore, we will design the mean of $\Re\{\lambda_i\}$ and $\Re\{\lambda_{i+1}\}$ to be α or $-\alpha$ so that pairs of neighboring pulses will walk past each other along propagation. The i^{th} soliton waveform is then calculated through the INFT operation. Typical soliton transmission systems generate and process individual waveform for a given $\{\lambda_i, b_i(\lambda_i)\}$ with separated non-overlapping time intervals $[(i-1)T, iT]$ where T is the symbol period in normalized time. At the receiver, we denote $\{\tilde{\lambda}_i, \tilde{b}_i(\tilde{\lambda}_i)\}$, $\{\bar{\lambda}_i, \bar{b}_i(\bar{\lambda}_i)\}$ and $\{\hat{\lambda}_i, \hat{b}_i(\hat{\lambda}_i)\}$to be the eigenvalue and b-coefficient of the received waveform before and after additional DSP and the final detected eigenvalue and b -coefficient respectively. For the rest of the paper, we will denote the corresponding eigenvalue noise as $\Delta\tilde{\lambda}_i = \tilde{\lambda}_i - \lambda_i$ and $\Delta\bar{\lambda}_i = \bar{\lambda}_i - \lambda_i$.

In our previous scheme, 1-solitons are generated with eigenvalues $\lambda = \pm\alpha + 1i$ for odd and even indexed solitons respectively (Fig. 2 (a)) and they periodically collided during transmission. For a given distance ℓ, α need to be carefully chosen so that solitons will be well-separated at receiver even though they have interacted during propagation. To relax this constraint between ℓ and α, we note that when solitons collide with each other, the super-imposed waveforms are still well-contained in a time window with duration 2T and well-separated from neighboring superimposed waveforms for any ℓ and α. Examples for $\ell = 3{,}4$ and $\alpha = 0.52$ are shown in Fig. 2 (b) and (c). Therefore, one can always divide the received signal into blocks of duration 2T and calculate the NFT to obtain pairs of eigenvalues and b-coefficients $\{\tilde{\lambda}_i, \tilde{b}_i(\tilde{\lambda}_i), \{\tilde{\lambda}_{i+1}, \tilde{b}_{i+1}(\tilde{\lambda}_{i+1})\}$ together. Consequently, the transmitter generates 1-soliton signals while the receiver detects and processes 2-soliton signals.

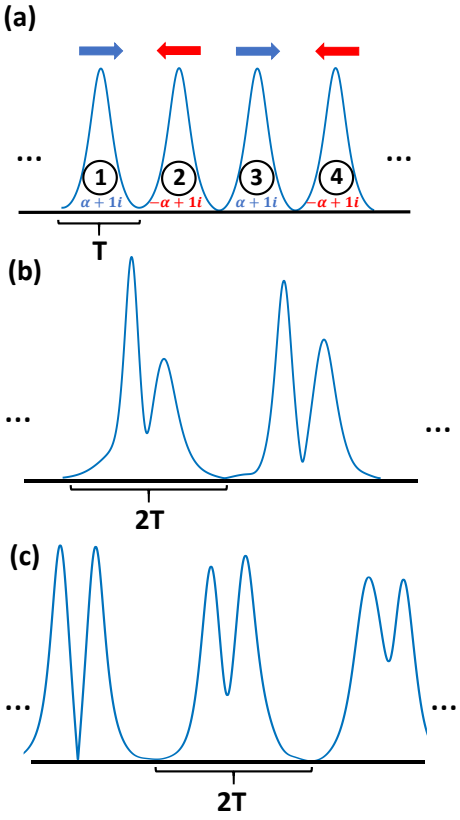

Fig. 2 (a) At the transmitter, individual 1-solitons are generated in separate non-overlapping time windows with duration T. The eigenvalue of the odd- and even-index solitons are $\alpha + 1i$ and $-\alpha + 1i$ respectively so that neighboring solitons have opposite group velocity and will walk towards each other. At distance of (b) $\ell = 3$ and (c) $\ell = 4$, pairs of solitons interact with each other but the superimposed waveforms are always well-contained in time windows with duration 2T. The received waveform is therefore divided into 2T intervals for NFT processing in our study.

C. LMMSE estimation of discrete eigenvalue from continuous spectrum (CS)

It has been shown that the noise in continuous spectrum (CS) of a signal is statistically correlated to the discrete eigenvalue noise $\Delta\tilde{\lambda}_i = \tilde{\lambda}_i - \lambda_i$ in the discrete spectrum [28]. Since we are studying discrete eigenvalue transmission where no

information is encoded in the CS, we therefore propose to use another set of LMMSE filters $\boldsymbol{c_c}$ to estimate and compensate $\Re\{\Delta\tilde{\lambda}_i\}$ and $\Im\{\Delta\tilde{\lambda}_i\}$ based on CS. Using our experimental data, we calculated and showed in Fig. 5 that correlations between $\Delta\tilde{\lambda}_i$ and CS is mainly between a-coefficient $a_i(\lambda)$ rather than b-coefficient $b_i(\lambda)$ or q-coefficient $q_i(\lambda)$ (Note: we dropped the subscript i in λ for CS to reflect that no information is encoded in the CS in our study). Hence, we only use $a_i(\lambda)$ to form an LMMSE estimate of $\Delta\tilde{\lambda}_i$. In our experiments, training symbols are used to calculate the corresponding means and covariance matrices to determine the LMMSE filter analogous to Eqn. (7). We also separate signals into two groups with $\mathbf{E}[\Re\{\lambda_i\}] = \alpha$ or $-\alpha$ and train a separate LMMSE estimator for each group. More detailed results and discussions are described in section IV and Fig. 5.

D. Multi-Symbol (MS)-LMMSE estimation of b-coefficient

In our proposed setup of 1-solition transmission and 2-soliton detection, neighboring solitons will interact along propagation and after eigenvalue detection, the noise and distortions in b-coefficients will exhibit correlations across multiple 2T time windows. To take advantage of these correlations, we propose a multi-symbol (MS) version of LMMSE estimators of the b-coefficient amplitude noise $\Delta|\tilde{b}_i(\hat{\lambda}_i)|$ and phase noise $\Delta\angle\tilde{b}_i(\hat{\lambda}_i)$ in which the inputs are a sliding window of eigenvalue noise $\Delta\tilde{\lambda}_{i-3}, \Delta\tilde{\lambda}_{i-2}, \cdots \Delta\tilde{\lambda}_{i+3}, \Delta\tilde{\lambda}_{i+4}$. Training symbols will be used to calculate the means and covariance matrices and derive the LMMSE estimator analogous to Eqn. (7). Note that as we also modulate λ_i in our study, we will train separate LMMSE estimators for each possible λ_i, resulting in $2 \times M_\lambda$ filters for $\Delta|\tilde{b}_i(\hat{\lambda}_i)|$ and another $2 \times M_\lambda$ filters for $\Delta\angle\tilde{b}_i(\hat{\lambda}_i)$.

E. Approximate Maximum Likelihood (ML) and MS-ML discrete eigenvalue and b-coefficient detection

Practical NFT transmissions suffer from fiber loss, amplified spontaneous emission (ASE) noise from optical amplifiers among other non-idealities. Fiber loss and lumped amplification are simple scalar effect in linear systems, but they become complicated and signal-depended distortions in the nonlinear spectrum. Also, ASE noise are no longer additive white Gaussian in nonlinear spectrum. The effects collectively result in complicated signal-dependent noise with memory that are correlated across the nonlinear spectrum [15]. Therefore, symbol detection based on simple Euclidian distance metric is clearly suboptimal. In principle, the optimal symbol detection strategy is maximum likelihood (ML) detection which require conditional probability density functions (pdf) of received signals given each transmitted symbol $P(\bar{\lambda}_i|\lambda_i)$ and $P(\bar{b}_i(\hat{\lambda}_i)|b_i(\lambda_i))$. As these pdfs are not available in full analytical forms at present, we will approximate ML detection by 1) separating λ and $b(\lambda)$ detection and 2) approximate $P(\bar{\lambda}_i|\lambda_i)$ and $P(\bar{b}_i(\hat{\lambda}_i)|b_i(\lambda_i))$ by a Gaussian distribution with the same mean and covariance matrix so that the ML estimates of $\{\lambda_i, b_i(\lambda_i)\}$ are given by

$$\hat{\lambda}_i = \underset{k\in\{1,\cdots M_\lambda\}}{argmax} \; -\left(\bar{\lambda}_i - \mu_{\lambda,k}\right)^H \Sigma_{\lambda,k}^{-1}\left(\bar{\lambda}_i - \mu_{\lambda,k}\right) \quad (8)$$

$$\hat{b}_i(\hat{\lambda}_i) = \underset{r\in\{1,\cdots M_b\}}{argmax} \; -\left(\bar{b}_i - \mu_{b,r}\right)^H \Sigma_{b,r}^{-1}\left(\bar{b}_i - \mu_{b,r}\right) \quad (9)$$

where $\mu_{\lambda,k}, \Sigma_{\lambda,k}, \mu_{b,r}, \Sigma_{b,r}$ are the means and covariance matrices of $\bar{\lambda}_i|\lambda_i = \psi_\lambda(k)$ and $\bar{b}_i(\hat{\lambda}_i)|b_i(\lambda_i) = \psi_b(r)$ respectively. In our experiments, they are empirically calculated from training symbols which can simultaneously be used to train the preceding LMMSE estimators as well.

Furthermore, in view of the proposed 2-soliton detection scheme, it is expected that noise in $\{\bar{\lambda}_i, \bar{b}_i(\hat{\lambda}_i), \bar{\lambda}_{i+1}, \bar{b}_{i+1}(\hat{\lambda}_{i+1})\}$ are highly correlated because they are grouped together at the receiver and affected by the same ASE noise and other distortions in the same 2T time window. Therefore, we also study multi-symbol (MS)-ML detection between pairs of overlapping soliton pulses. In this case, the ML estimates are given by

$$\{\hat{\lambda}_i, \hat{\lambda}_{i+1}\} = \underset{k,n\in\{1,\cdots M_\lambda\}^2}{argmax} \; P\left(\bar{\lambda}_i, \bar{\lambda}_{i+1}|\lambda_i = \psi_\lambda(k), \lambda_{i+1} = \psi_\lambda(n)\right) \quad (10)$$

$$\{\hat{b}_i(\hat{\lambda}_i), \hat{b}_{i+1}(\hat{\lambda}_{i+1})\} = \underset{r,s\in\{1,\cdots M_b\}^2}{argmax} \; P\left(\bar{b}_i, \bar{b}_{i+1}|b_i = \psi_b(r), b_{i+1} = \psi_b(s)\right) \quad (11)$$

and the $M_\lambda^2 + M_b^2$ possible pdfs will be approximated by multi-variate Gaussians with the same mean and covariance matrices. In principle, we can extend this approach to joint detection of 3 or more symbols in neighboring NFT time windows but to enumerate all possible combinations of the symbol values and ensure accurate estimates of mean and covariance matrices for each combination, the amount of training symbols increase exponentially. Therefore, we only study approximate MS-ML detection of $\{\bar{\lambda}_i, \bar{\lambda}_{i+1}\}$ and $\{\bar{b}_i(\hat{\lambda}_i), \bar{b}_{i+1}(\hat{\lambda}_{i+1})\}$ within the same 2T NFT time window. Fig. 3 summarize the overall DSP flow proposed in our study.

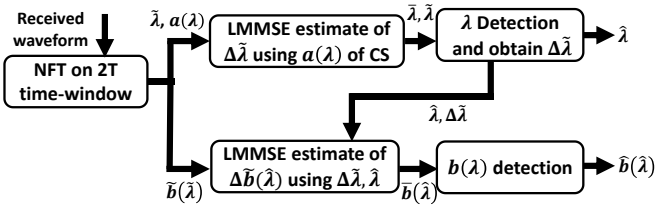

Fig. 3. Proposed multi-symbol DSP flow for discrete-eigenvalue transmissions. Firstly, the a-coefficient of the continuous spectrum is used to reduce the eigenvalue noise $\Delta\tilde{\lambda}_i = \tilde{\lambda}_i - \lambda_i$ through an LMMSE estimator followed by eigenvalue detection. The detected eigenvalue $\hat{\lambda}_i$ and noise $\Delta\tilde{\lambda}_i$ are then used in LMMSE estimators to reduce the noise in b-coefficient followed by detection. The detection method can be simple Euclidean-based detection, approximated ML or MS-ML detection.

III. EXPERIMENTAL SETUP

We conducted experiment with offline signal processing to investigate the practical performance of various DSP algorithms proposed in our study. Fig. 4 (a) shows the experimental setup and offline DSP structure. We jointly modulate λ and $b(\lambda)$ with 16-QAM and 16-APSK format

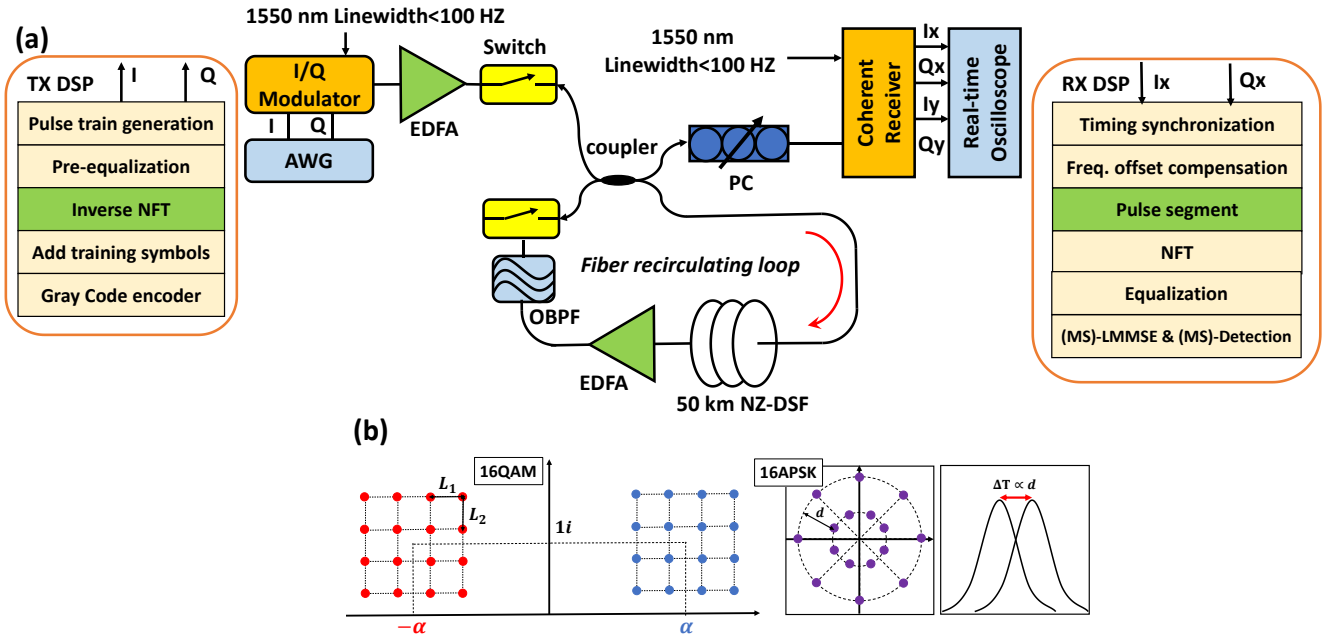

Fig. 4 (a) Experiment setup for discrete-eigenvalue transmission using multi-symbol DSP. (b) The 16-QAM constellation for λ modulation and 16-APSK $b(\lambda)$ modulation. ΔT is optimized to be 0.4 in experiment while L_1 and L_2 are 0.03 and 0.052.

respectively and the constellation shapes are shown in Fig. 4 (b). Gray coding is separately applied to the constellation points in λ and $b(\lambda)$. As shown in Fig. 2 (a), odd- and even-indexed solitons are generated with center of $\Re(\lambda) = \alpha$ and $-\alpha$ respectively. Because the normalization parameter T_0 and eigenvalue imaginary part is similar to our previous study [15], we chose the optimal value of α in that work ($\alpha = 0.8$) in our current study. In b-coefficient constellation, the magnitude corresponds to time position of soliton pulses and hence d determine the timing difference between solitons. For 1-soliton pulse, setting $|b(\lambda)| = 1$ locates the soliton at the center of the NFT time window and scaling $|b(\lambda)|$ with $\exp(2\Im(\lambda)\Delta T)$ leads to a time drift of ΔT [6, 42]. In our modulation scheme, we set $|b(\lambda)|$ in the two rings of 16-APSK so that their corresponding time-waveform are shifted by $\pm\Delta T/2$ from the center of the time window. Therefore, $|b(\lambda)|$ of two rings is set as $\exp(\pm\Im(\lambda)\Delta T)$ which depends on λ. Increasing ΔT will better distinguish the two types of soliton pulses in the time-domain but causes more overlapping with neighboring solitons and produce more distortions on the phase of $b(\lambda)$. In our study, we experimentally optimize ΔT to be 0.4 where the empirical symbol error ratios resulting from amplitude and phase noise of $b(\lambda)$ are similar. For λ constellation, we showed in our previous work [26] that noise variance on real and imaginary part of λ are different for 1-soliton. Consequently, we also optimized the horizontal and vertical distances L_1 and L_2 of the λ constellation to be 0.03 and 0.052, taking into account the trade-offs between separation of signal points in the λ-plane and inducing signal bandwidth expansion and nonlinear distortions to neighboring symbols i.e. nonlinear inter-symbol interference (ISI).

Pre-adaptation as [43] was applied to pre-compensate the transmitter component imperfections. For each jointly modulated symbol, we performed the INFT to calculate the corresponding time-domain waveform which are then loaded onto the arbitrary waveform generator (Keysight M8194A, 120GSa/s, 45 GHz) for electrical waveform generation. The waveform then goes into the I/Q modulator (Fujitsu 7962EP, 28GHz) to modulate the optical carrier, which are then amplified and launched into fiber. The loop consists of 1 span 50-km NZDSF (Dispersion: ~4e-6 s/m, loss: ~0.2 dB/km) and lumped amplified by an EDFA. A flat-top optical filter with a 3-dB bandwidth of 1 nm was used after the loop EDFA to suppress out-of-band amplified spontaneous emission (ASE) noise. Both the transmitter laser and local oscillator were from the same fiber laser source with very low laser phase noise (NKT Koheras ADJUSTIK Fiber laser with linewidth < 100Hz). After alignment by a polarization controller, the received signal was then coherently detected and sampled by a digital storage scope (Keysight, DSA-X-96204Q, 80 GSa/s, 33 GHz) and processed offline. The sampled signal passes through a series of digital signal processing (DSP) structure shown in Fig. 4 (a). After timing synchronization and frequency offset compensation, the received signal are separated into blocks of 2T described in Fig. 2. NFT processing is then applied to recover the nonlinear spectrum followed by the series of multi-symbol DSP discussed above. For detection, we first equalize the channel response through multiplying $\tilde{b}(\hat{\lambda})$ by $\exp(-4j\hat{\lambda}^2\ell)$. Since fiber loss and other unaccounted experimental factors breaks the integrability of the NLSE and effectively distorts the nonlinear channel response, additional training symbols are used to adjust the equalization term so that the distributions of $\tilde{b}(\hat{\lambda})$ after equalization for different $\hat{\lambda}$ are as consistent as possible before MS-LMMSE estimation of $\Delta\tilde{b}_i(\hat{\lambda}_i)$ and (MS)-ML detection.

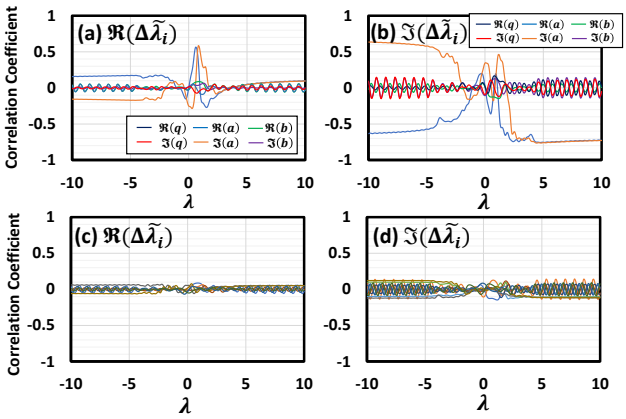

Fig. 5 (a, b) Empirical correlation coefficients between $\{\Re(\Delta\tilde{\lambda}_i), \Im(\Delta\tilde{\lambda}_i)\}$ and CS coefficients $\{q_i(\lambda), a_i(\lambda), b_i(\lambda)\}$ for 6.31 GBaud transmission over 1000 km, showing large correlation between $\Delta\tilde{\lambda}_i$ and $a_i(\lambda)$ compared to $q_i(\lambda)$ and $b_i(\lambda)$. (c, d) Empirical correlation coefficients between $\{\Re(\Delta\tilde{\lambda}_i), \Im(\Delta\tilde{\lambda}_i)\}$ and $\{q_{i\pm1}(\lambda), a_{i\pm1}(\lambda), b_{i\pm1}(\lambda)\}$ of signals in neighboring time windows.

IV. Results and Discussions

A. Characterization of correlations between different parameters and across neighboring symbols

We investigated the proposed multi-symbol DSP scheme over 1000 km NFT transmission and the empirical correlation coefficients between received discrete eigenvalue noise $\Delta\tilde{\lambda}_i$ and CS coefficients are shown in Fig. 5 for $\lambda \in [-10,10]$ for 6.31 GBaud transmission over 1000 km. 4096 pairs of solitons are used to calculate the correlation coefficients. It can be seen from Fig. 5 (a, b) that $a_i(\lambda)$ is much more correlated to $\Delta\tilde{\lambda}_i$ compared with $q_i(\lambda)$ and $b_i(\lambda)$. Moreover, Fig. 5 (c, d) show that such a large correlation does not extend across neighboring signals where the correlation between $\Delta\tilde{\lambda}_i$ and any CS coefficients of neighboring signals is quite low. This suggests that a multi-symbol version of LMMSE filter on $a_i(\lambda)$ is not necessary. Consequently, we only developed an LMMSE estimate of $\Delta\tilde{\lambda}_i$ using $a_i(\lambda)$ with 256 points uniformly distributed in $\lambda \in [-10,10]$.

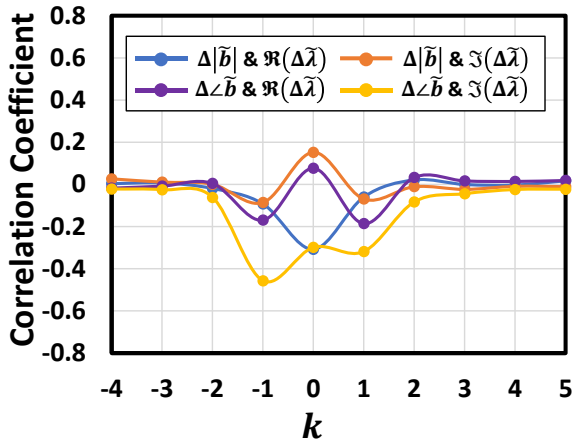

Fig. 6 Empirical correlation coefficients between the real and imaginary part of $\Delta\tilde{\lambda}_{i+k}$ and the amplitude and phase noise of b-coefficient $\Delta\tilde{b}_i(\hat{\lambda}_i)$ for 6.31 GBaud transmission over 1000 km.

After λ detection with the help of LMMSE filter using $a_i(\lambda)$, Fig. 6 shows the correlation coefficients between the real and imaginary part of $\Delta\tilde{\lambda}_{i+k}$ and the amplitude and phase noise of b-coefficient $\Delta\tilde{b}_i(\hat{\lambda}_i)$ for 6.31GBaud transmission over 1000 km. 4096 pairs of solitons are used to calculate the correlation coefficients. For this transmission distance, the solitons have collided once with their neighbors and therefore a strong correlation between $\Delta\tilde{b}_i(\hat{\lambda}_i)$ and $\Delta\tilde{\lambda}_{i\pm1}$ is observed. In addition, the empirical correlations extend towards three neighboring time windows ($\Delta\tilde{\lambda}_{i-3}$ to $\Delta\tilde{\lambda}_{i+4}$) before it become negligible. Therefore, we study a MS-LMMSE filter incorporating $\Delta\tilde{\lambda}_{i-3}, \Delta\tilde{\lambda}_{i-2}, \cdots \Delta\tilde{\lambda}_{i+3}, \Delta\tilde{\lambda}_{i+4}$ to minimize the phase and amplitude noise of $\Delta\tilde{b}_i(\hat{\lambda}_i)$.

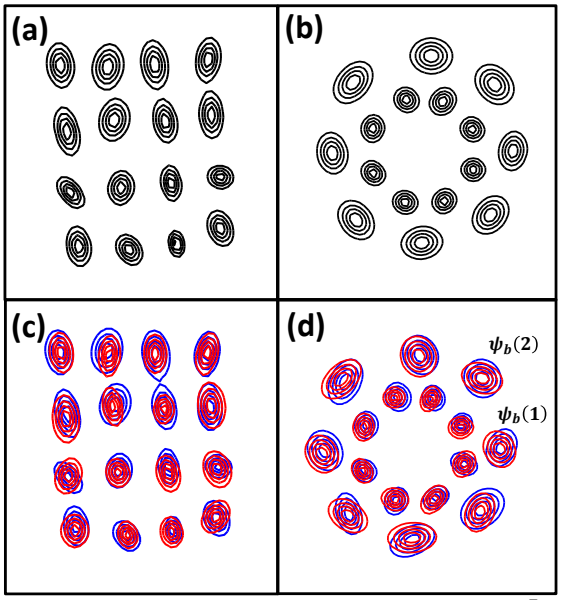

Fig. 7. Empirical received signal distributions (contour plots) of (a) $\bar{\lambda}_i$ and (b) $\bar{b}_i(\hat{\lambda}_i)$ for 6.31 GBaud transmission over 1000 km, showing the non-uniformity across different constellation points; Empirically received signal distributions (contour plots) of (c) $\bar{\lambda}_i|\lambda_{i+1} = -0.845 + 1.078i$ (blue) and $\bar{\lambda}_i|\lambda_{i+1} = -0.815 + 1.078i$ (red) and (d) $\bar{b}_i(\hat{\lambda}_i)|b_{i+1}(\lambda_{i+1}) = \psi_b(1)$ (blue) and $\bar{b}_i(\hat{\lambda}_i)|b_{i+1}(\lambda_{i+1}) = \psi_b(2)$ (red). The means and covariance matrices of $\bar{\lambda}_i$ and $\bar{b}_i(\hat{\lambda}_i)$ depends on $\lambda_{i+1}$and $b_{i+1}(\lambda_{i+1})$, thus illustrating the potential benefits of MS-ML detection. The amplitude of $\bar{b}_i(\hat{\lambda}_i)$ are shown in log scale.

The empirical received signal distributions (contour plots) of $\bar{\lambda}_i$ and $\bar{b}_i(\hat{\lambda}_i)$ is shown in Fig. 7 (a, b) for 6.31 GBaud transmissions over 1000 km. 4096 pairs of solitons are used to calculate the conditional pdfs. It can be seen that the distributions are generally different for each constellation points with different degree of correlations between $\Re(\Delta\bar{\lambda}_i)$ and $\Im(\Delta\bar{\lambda}_i)$. Also, eigenvalue noise is generally larger for imaginary part than real part. This is because solitons with larger $\Im(\lambda)$ induce more noise for the same ASE noise power [44]. On the other hand, the outer ring of $\bar{b}_i(\hat{\lambda}_i)$ generally have larger noise variances. As the conditional pdfs are different and non-uniform across different constellation points of $\bar{\lambda}_i$ and $\bar{b}_i(\hat{\lambda}_i)$, approximating them as jointly Gaussian using their respective empirical means and covariance matrices will certainly outperform simple Euclidean distance-based detection. This class-labelled approach of estimating means and covariance matrices will also be superior to unsupervised

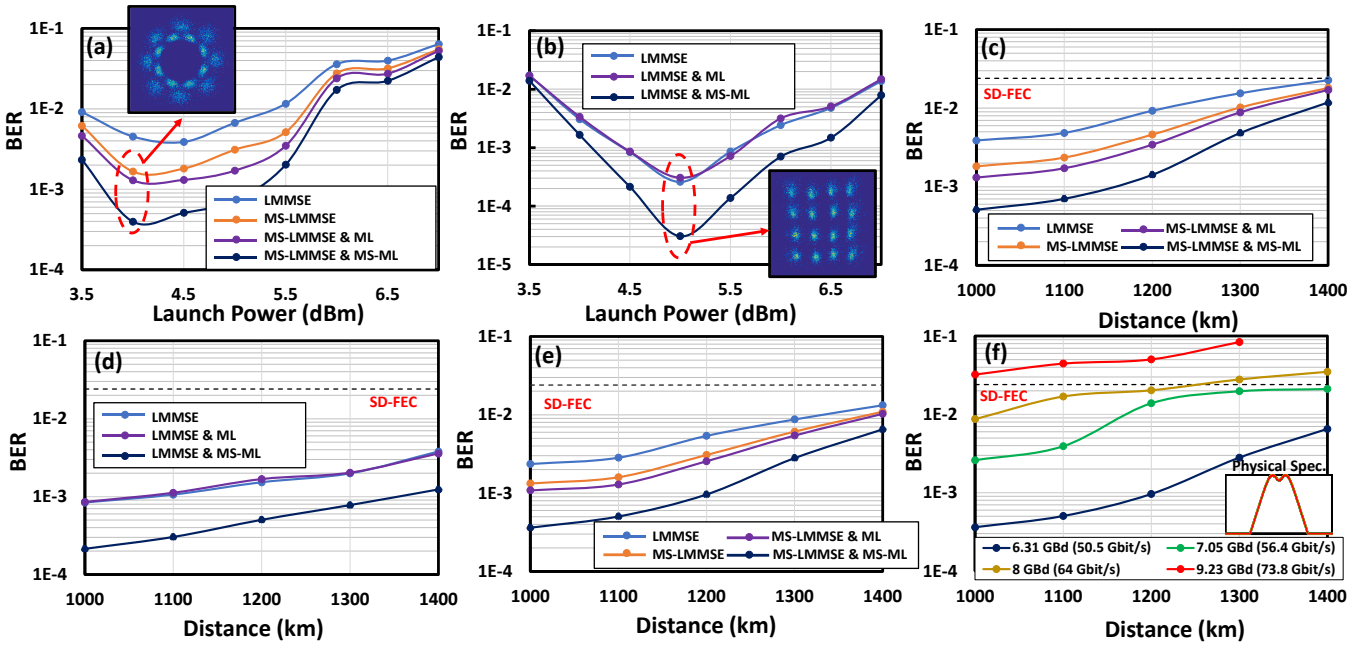

Fig. 8 BER of (a) $b(\lambda)$ and (b) λ detection vs. launch power using various LMMSE filters and detection methods. Received $\bar{\lambda}$ distribution after LMMSE and $\bar{b}(\hat{\lambda})$ distribution after MS-LMMSE are shown in insets. BER of (c) λ, (d) $b(\lambda)$ and (e) joint λ and $b(\lambda)$ detection vs. distance using different LMMSE and detection methods for 6.31 GBaud transmission are shown. The launched powers are optimized for each distance; (f) BER vs. distance with LMMSE on $\bar{\lambda}$, MS-LMMSE on $\bar{b}(\hat{\lambda})$ and MS-ML detection for different baud rates. The launched powers are optimized for each distance and baud rate. The signal physical spectra of all 4 baud rates are shown in inset and they are nearly the same.

clustering techniques such as EM algorithm [45] commonly used in machine learning literature. Fig. 7 (c) and (d) further show the empirical conditional pdf of $\bar{\lambda}_i|\lambda_{i+1} = -0.845 + 1.078i$ and $\bar{\lambda}_i|\lambda_{i+1} = -0.815 + 1.078i$, $\bar{b}_i(\hat{\lambda}_i)|b_{i+1}(\lambda_{i+1}) = \psi_b(1)$ and $\bar{b}_i(\hat{\lambda}_i)|b_{i+1}(\lambda_{i+1}) = \psi_b(2)$. As expected, the pdfs show clear dependence on λ_{i+1} and $b_{i+1}(\lambda_{i+1})$, which motivates the study of MS-ML detection by approximating the joint distributions of $\{\bar{\lambda}_i, \bar{\lambda}_{i+1}\}$ and $\{\bar{b}_i(\hat{\lambda}_i), \bar{b}_{i+1}(\hat{\lambda}_{i+1})\}$ as jointly Gaussian variables with the empirical means and covariance matrices of each possible combination of $\{\bar{\lambda}_i, \bar{\lambda}_{i+1}\}$ and $\{\bar{b}_i(\hat{\lambda}_i), \bar{b}_{i+1}(\hat{\lambda}_{i+1})\}$.

B. *Transmission Performance*

We compare the transmission performance of the various DSP proposed in our study. Fig. 8 (a, b) compare the BER of λ and $b(\lambda)$ detection with various LMMSE estimators and detection methods. Corresponding received signal distributions are also shown as insets. For λ detection, using approximate ML detection over simple Euclidean detection does not bring significant benefits. This might be because the conditional pdfs of $\bar{\lambda}_i$ is largely uniform for each dimension as can be seen in Fig. 7 (a). For $b(\lambda)$ detection, MS-LMMSE outperform LMMSE by exploiting correlations of $\Delta\tilde{\lambda}_i$ across neighboring symbols. ML detection on top of MS-LMMSE further improved the BER slightly as the shapes of the distributions are non-uniform across signal points (inset of Fig. 8 (a)). Finally, by capturing the correlations between the two λ and $b(\lambda)$ in a 2T NFT time window, MS-ML detection significantly improve BER performance for both λ and $b(\lambda)$ detection. This is expected since the two neighboring solitons interact with each other along propagation and the nonlinear distortions will be signal-dependent. This is evident by the difference in empirical pdf of $\bar{\lambda}_i$ and $\bar{b}_i(\bar{\lambda}_i)$ for different values of $\bar{\lambda}_{i+1}$ and $\bar{b}_{i+1}(\bar{\lambda}_{i+1})$ as shown in Fig. 6. It should be noted that the best launch power for λ detection is around 1 dB higher than that for $b(\lambda)$ detection.

To investigate the reach extension and bit rate improvements of the proposed suite of DSP algorithms, Fig. 8 (c-e) depict the BER of λ, $b(\lambda)$ and joint λ and $b(\lambda)$ detection for 1000 to 1400 km transmissions where neighboring solitons varied from fully separated to substantial collisions and overlapping at the receiver. The launch power is chosen to minimize the overall BER of λ and $b(\lambda)$ detection. Comparing the BER of $b(\lambda)$ in Fig. 8 (a, c) with BER of λ in Fig. 8 (b, d), it can be seen that the overall BER is mainly attributed by errors of $b(\lambda)$ detection. Note that In Fig. 8 (e), the MS-LMMSE results are the average BERs of λ detection after LMMSE and $b(\lambda)$ detection after MS-LMMSE. Improvement of MS-LMMSE on $b(\lambda)$ detection is also obvious in overall BER because errors in $b(\lambda)$ detection is the dominant factor. MS-ML detection provides substantial improvements for all distances, reflecting that strong correlations among λ and $b(\lambda)$ within the 2T NFT time window exists regardless of soliton collision status. We also studied the performance when the baud rate (hence bit-rate) is increased by shortening the time interval T between and the results are shown in Fig. 8 (f). For a SD-FEC threshold of 2.4e-2 (assuming an inner LDPC code with rate 9/10 and an outer HD-FEC staircase code with 6.25% overhead [46]), a total of 64 Gb/s (net 54 Gb/s) transmitted over 1200 km can be achieved. The BER further deteriorates with increasing baud rates as putting the solitons closer together induce more pulse

overlapping and nonlinear distortions. The physical spectra of the transmitted signals with various baud rates are shown as insets of Fig. 8 (f) with a 99% optical bandwidth of 35.7 GHz. It can be seen that signal bandwidth for 4 different baud rates are nearly the same. Therefore, increasing baud rate by putting solitons closer will not broaden the signal bandwidth. To the best of our knowledge, this is a new bit-rate-distance record for single-polarization discrete eigenvalue transmissions.

V. Conclusions

In this paper, we studied discrete-eigenvalue transmissions based on NFT and proposed various multi-symbol receiver DSP techniques that exploits the statistical correlations between eigenvalues λ and other nonlinear spectral components across neighboring symbols to improve detection performance. The DSP techniques include jointly modulate λ and $b(\lambda)$ of a 1-soliton pulse with the average eigenvalue $\mathbf{E}[\lambda] = \pm\alpha + 1i$ for odd and even indexed solitons across the transmitted pulse train, decode superimposed received waveforms as 2-solitons, a linear minimum mean-squared error (LMMSE) estimator of discrete eigenvalue noise using the a-coefficient $a(\lambda)$ of the continuous spectrum (CS), a multi-symbol (MS)-LMMSE estimator of $b(\lambda)$ noise using discrete eigenvalue noise across neighboring solitons, and approximate ML and MS-ML detection of λ and $b(\lambda)$ by Gaussian distributions with means and covariance matrices obtained empirically from experiments. We experimentally demonstrated a record 64 Gb/s (net 54 Gb/s) transmission over 1200 km using 16-QAM modulation on λ and 16-APSK modulation on $b(\lambda)$ with the aid of the DSP algorithms described above. The proposed multi-symbol DSP framework opens up a new dimension of algorithm research for NFT systems and more advanced nonlinear filters as well as maximum likelihood sequence detection (MSLD) techniques will be investigated in the future.

References

[1]. V. E. Zakharov and A. B. Shabat, "Exact theory of 2-dimensional selffocusing and one-dimensional self-modulation of waves in nonlinear media," Sov. Phys. J. Exp. Theor. Phys., vol. 34, no. 1, pp. 62–69, Jan. 1972.

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

[3]. E. G. Turitsyna and S. K. Turitsyn. "Digital signal processing based on inverse scattering transform," *Opt. Lett.*, vol. 38, no. 20, pp. 4186-4188, Oct. 2013.

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

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

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

[7]. Z. Dong, S. Hari, T. Gui, K. Zhong, M. Yousefi, C. Lu, P-K. A. Wai, F. Kschischang, and A. P. T. Lau, "Nonlinear frequency division multiplexed transmissions based on NFT," *IEEE Photon. Technol. Lett.*, vol. 99, no. 15, pp. 1621-1623, Aug. 2015.

[8]. S. T. Le, V. Aref, and H. Buelow, "Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit," *Nat. Photonics*, vol. 11 no. 9, pp. 570-576, Jul. 2017.

[9]. J. W. Goossens, M. I. Yousefi, Y. Jaoun, and H. Hafermann, "Polarization-division multiplexing based on the nonlinear fourier transform," *Opt. express*, vol. 25, no. 22 pp. 26437-26452, Oct. 2017.

[10]. W. A. Gemechu, T. Gui, J-W. Goossens, M. Song, S. Wabnitz, H. Hafermann, A. P. T. Lau, M. I. Yousefi, and Y. Jaouën, "Dual Polarization Nonlinear Frequency Division Multiplexing Transmission," *IEEE Photon. Technol. Lett.*, vol. 30, no. 18, pp. 1589-1592, Sep. 2018.

[11]. X. Yangzhang, S. T. Le, V. Aref, H. Buelow, D. Lavery and P. Bayvel, "Experimental Demonstration of Dual-Polarization NFDM Transmission With b-Modulation," *IEEE Photon. Technol. Lett.*, vol. 31, no. 11, pp. 885-888, Apr. 2019.

[12]. T. Gui, C. Lu, A. P. T. Lau, and P-K. A. Wai, "High-order modulation on a single discrete eigenvalue for optical communications based on nonlinear Fourier transform," Opt. Express, vol. 25, no. 17, pp. 20286-20297, Aug. 2017.

[13]. S. Gaiarin, A. M. Perego, E. P. D. Silva, F. Da Ros, and D. Zibar, "Dual-polarization nonlinear Fourier transform-based optical communication system," Optica, vol. 5 no. 3, pp. 263-270, Mar. 2018.

[14]. S. Gaiarin, F. D. Ros, N. D. Renzis, E. P. D. 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, Oct. 2018.

[15]. G. Zhou, T. Gui, C. Lu, A. P. T. Lau and P-K. A. Wai, "Improving soliton transmission systems through soliton interactions," *J. Lightwave Technol.*, vol 38 no.14, pp. 3563-3572, Jul. 2020.

[16] V. Aref, S. T. Le, and H. Buelow, "Modulation Over Nonlinear Fourier Spectrum: Continuous and Discrete Spectrum," *J. Lightwave Technol.*, vol. 36 no. 6, pp. 1289–1295, Mar. 2018.

[17]. F. D. Ros, S. Civelli, S. Gaiarin, E. P. D. Silva, N. D. Renzis, M. Secondini and D. Zibar, "Dual-polarization NFDM transmission with continuous and discrete spectral modulation," *J. Lightwave Technol.*, vol. 37 no. 10, pp. 2335-2343, May 2019

[18]. F. J. García-Gómez, "Communication using eigenvalues of higher multiplicity of the nonlinear Fourier transform," *J. Lightwave Technol.*, vol 36 no. 23, pp. 5442-5450, Dec. 2018

[19]. S. Hari, and F. R. Kschischang, "Bi-directional algorithm for computing discrete spectral amplitudes in the NFT," *J. Lightvave Technol.*, vol. 34, no. 15, pp. 3529-3537, Aug. 2016.

[20]. S. Wahls, and H. V. Poor. "Fast numerical nonlinear Fourier transforms," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6957-6974, Dec. 2015.

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

[22]. A. Span, V. Aref, H. Buelow and S. T. Brink, "Successive Eigenvalue Removal for Multi-Soliton Spectral Amplitude Estimation," *J. Lightvave Technol.*, vol 38 no. 17, pp. 4708-4714, Sept. 2020.

[23]. S. Wahls, and H. V. Poor, "Fast inverse nonlinear Fourier transform for generating multi-solitons in optical fiber," in *International Symposium on Information Theory (ISIT)*, Hong Kong, China, 2015, pp. 1676-1680.

[24]. L. L. Frumin, O. V. Belai, E. V. Podivilov, and D. A. Shapiro, "Efficient numerical method for solving the direct Zakharov–Shabat scattering problem," *JOSA B*, vol. 32, no. 2, pp. 290-296, Feb. 2015.

[25]. V. Aref, "Control and detection of discrete spectral amplitudes in nonlinear Fourier spectrum," *arXiv* preprint arXiv:1605.06328 (2016).

[26]. 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.

[27]. T. Gui, G. Zhou, C. Lu, A. P. T. Lau, and S. Wahls, "Nonlinear frequency division multiplexing with b-modulation: shifting the energy barrier," *Opt. Express*, vol. 26 no. 21, pp. 27978-27990 2018.

[28]. Q. Zhang, and F. R. Kschischang, "Improved soliton amplitude estimation via the continuous spectrum," *J. Lightwave Technol.*, vol. 37 no. 13, pp. 3087-3099, Jul. 2019

[29]. M. Shehadeh, Q. Zhang, and F. R. Kschischang, "Reduced-Complexity Nonlinear Soliton Amplitude Estimators," *IEEE Photon. Technol, Lett.*, vol. 31 no. 24, pp. 1933-1935, Dec. 2019.

[30]. J. Koch, K. Chan, C. G. Schaeffer and S. Pachnicke, "Signal Processing Techniques for Optical Transmission Based on Eigenvalue Communication," *IEEE J. Sel. Top. Quantum Electron.* Early Access.

[31]. M. Kamalian, J. E. Prilepsky, S. T. Le and S. K. Turitsyn "Periodic nonlinear Fourier transform for fiber-optic communications, Part I:

theory and numerical methods," *Opt. express*, vol. 24 no. 16, pp. 18353-18369, Aug. 2016

[32]. M. Kamalian, J. E. Prilepsky, S. T. Le and S. K. Turitsyn "Periodic nonlinear Fourier transform for fiber-optic communications, Part II: eigenvalue communication," *Opt. express*, vol. 24 no. 16, pp. 18370-18381, Aug. 2016.

[33]. L. L. Frumin., A. A. Gelash, and S. K. Turitsyn, "New approaches to coding information using inverse scattering transform," *Phys. Rev. lett.*, vol. 118, no. 22, pp. 223901, Jun. 2017.

[34]. A. Span, V. Aref, H. Buelow and S. T. Brink, "Efficient precoding scheme for dual-polarization multi-soliton spectral amplitude modulation," *IEEE Trans. Commun.*, vol. 67 no. 11, pp. 7604-7615, Nov. 2019.

[35]. S. Civelli, E. Forestieri, and M. Secondini, "Detection-feedback detection strategy for nonlinear frequency-division multiplexing," *Opt. express*, vol 26 no. 9, pp. 12057-12071, Apr. 2018.

[36]. H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," In *European Conference on Optical Communication*, Dusseldorf, German, 2016, pp. 1-3. VDE.

[37]. R. T. Jones, S. Gaiarin, M. P. Yankov, and D. Zibar, "Time-Domain Neural Network Receiver for Nonlinear Frequency Division Multiplexed Systems," *IEEE Photon. Technol. Lett.*, vol. 30, no. 12, pp. 1079-1082, Apr. 2018.

[38]. Y. WU, L. Xi, X. Zhang, Z. Zheng, J. Wei, S. Du, W. Zhang and X. Zhang, "Robust neural network receiver for multiple-eigenvalue modulated nonlinear frequency division multiplexing system," *Opt. Express*, vol. 28 no. 12. pp. 18304-18316, Jun. 2020.

[39]. K. Mishina, S. Yamamoto, T. Kodama, T. Yoshida, D. Hisano and A. Maruta, "Experimental demonstration of neural network based demodulation for on-off encoded eigenvalue modulation," In *45th European Conference on Optical Communication (ECOC)*, Dublin, 2019, pp. 1-4. IET.

[40]. S. Gaiarin, F. D. Ros, N. D. Renzis, R. T. Jones and D. Zibar, "Experimental demonstration of nonlinear frequency division multiplexing transmission with neural network receiver," *J. Lightwave Technol.*, vol. 38 no. 23, pp. 6465-6473. Dec. 2020.

[41]. S. Gaiarin, F. Daros, R. T. Jones and D. Zibar, "End-to-end optimization of coherent optical communications over the split-step Fourier method guided by the nonlinear Fourier transform theory," *J. Lightwave Technol.*, vol. 39, no.2, pp. 418-428, Jan. 2021.

[42] Aref, Vahid. "Nonlinear fourier transform of truncated multi-soliton pulses." SCC 2019; 12th International ITG Conference on Systems, Communications and Coding. VDE, 2019.

[43]. S. T. Le, and H. Buelow, "Nonlinear frequency division multiplexed transmissions with 64QAM," In *Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC)*, Singapore, 2017, pp. 1-2.

[44]. Q. Zhang, and T. H. Chan, "A Gaussian noise model of spectral amplitudes in soliton communication systems," In IEEE 16th International Workshop on *Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, 2015, pp. 455-459.

[45]. F. N. Khan, Q. Fan, C. Lu, and A.P.T. Lau, "An optical communication's perspective on machine learning and its applications," *J. Lightwave Technol*, vol. 37, no.2, pp. 493-516, Jan. 2019.

[46] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel. "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems". *J. Lightwave Technol.*, vol. 33, no. 20, pp. 4338-4352, OCT. 2015.