The following publication 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," in Journal of Lightwave Technology, vol. 35, no. 9, pp. 1542-1550, 1 May1, 2017 is available at https://doi.org/10.1109/JLT.2017.2654493.

Alternative Decoding Methods for Optical Communications based on Nonlinear Fourier Transform

Tao Gui, Terence H. Chan, Chao Lu, Alan Pak Tao Lau and P.K.A. Wai

***Abstract*— Long-haul optical communications based on nonlinear Fourier Transform(NFT) has gained attention recently as a new communication strategy that inherently embrace the nonlinear nature of the optical fiber. For communications using discrete eigenvalues $\lambda \in \mathbb{C}^+$, information are encoded and decoded in the spectral amplitudes $q(\lambda) = b(\lambda)/(\frac{da(\lambda)}{d\lambda})$ at the root λ_{rt} where $a(\lambda_{rt}) = 0$. In this paper, we propose two alternative decoding methods using $a(\lambda)$ and $b(\lambda)$ instead of $q(\lambda)$ as decision metrics. For discrete eigenvalue modulation systems, we show that symbol decisions using $a(\lambda)$ at a prescribed set of λ values perform similarly to conventional methods using $q(\lambda)$ but avoids root searching and thus significantly reduce computational complexity. For systems with phase and amplitude modulation on a given discrete eigenvalue, we propose to use $b(\lambda)$ after for symbol detection and show that the noise in $\frac{da(\lambda)}{d\lambda}$ and λ_{rt} after transmission are all correlated with that in $b(\lambda_{rt})$. A linear minimum mean square error(LMMSE) estimator of the noise in $b(\lambda_{rt})$ is derived based on such noise correlation and transmission performance is considerably improved for QPSK and 16-QAM systems on discrete eigenvalues.**

***Index Terms*— Fiber Nonlinearity, Nonlinear Fourier Transform, Noise**

I. Introduction

Long-haul optical communications suffer from fiber Kerr nonlinearity effects which ultimately limits transmission capacity and reach. Recently, nonlinear Fourier transform(NFT) is proposed as a new theoretical framework that incorporate the nonlinear nature of the fiber in signal design and processing[1]-[6]. In NFT, independent information streams are encoded in parallel sub-carriers (or eigenvalues λ of the Lax operator governing the ideal Nonlinear Schrodinger Equation(NLS)[7]) and in a noiseless scenario, the signals will propagate through the nonlinear fiber in a 'linear' manner without mutual interference. In presence of noise, the works in [8]-[11] show that there is only a weak inter-mode coupling effect due to the noise action and deviation of the true fiber model from the pure ideal NLS. The eigenvalues λ can be divided into *continuous eigenvalues* where λ is real and *discrete eigenvalues* where λ lies on the upper-half complex plane. Such 'eigenvalue communication' was firstly introduced by Hasegawa and Nyu [12] and recent experiments can be categorized into two main directions depending on which part of the nonlinear spectrum is used for the modulation and transmission. Nonlinear Inverse Synthesis(NIS) proposed in [13] is based on modulation of the continuous part of nonlinear spectrum and experimental demonstrations of transoceanic transmissions are recently reported [14][15]. The other direction is discrete spectral modulation, such as eigenvalue-multiplexed multi-level modulation [16][17], OOK on 3 or 4 eigenvalues [18] as well as QPSK on 7 discrete eigenvalues[19].

The basic premise of communications using discrete eigenvalues is that information be encoded and decoded in the spectral amplitudes $q(\lambda) = b(\lambda)/\frac{da(\lambda)}{d\lambda} = b(\lambda)\,/\,a'(\lambda)$ at the root λ_{rt} where $a(\lambda_{rt}) = 0$. The terms $a(\lambda)$ and $b(\lambda)$ are the *nonlinear Fourier coefficients* defined in the next section. However, calculating λ_{rt} and the derivative $a'(\lambda_{rt})$ is computationally intensive and may prohibit its practical implementation. Another important issue of NFT-based communications is that despite some preliminary studies on the topic, a sufficient understanding of how noises will impact the nonlinear Fourier coefficients is generally lacking [9],[10],[20],[21]. Hence, signaling strategies exploiting the noise characteristics in the NFT domain to improve transmission performance are yet to be fully investigated.

In this paper, we attempt to address both issues by proposing to use $a(\lambda)$ and $b(\lambda)$ instead of the spectral amplitudes $q(\lambda)$ for symbol decoding. For discrete eigenvalue modulation systems (or On-Off Keying(OOK) on multiple eigenvalues, or general Frequency-Shift Keying(FSK)), we show that with an appropriate choice of λ, $a(\lambda)$ can be used as the metric for symbol decoding and provide comparable performance to using $q(\lambda)$, thus saving significant computational complexity by avoiding root searching. On the other hand, for phase and amplitude modulation (or general quadrature amplitude modulation(QAM)) systems on discrete eigenvalues, we show that $b(\lambda_{rt})$ contains all the information needed for symbol decoding and the noise in $b(\lambda_{rt})$, $a'(\lambda_{rt})$ and λ_{rt} are all correlated with each other. The correlations enable us to develop a linear minimum mean square error(LMMSE) estimator of the noise in $b(\lambda_{rt})$ and improve transmission performance for QPSK and 16-QAM systems on discrete eigenvalues using $b(\lambda_{rt})$ as decoding metric.

The rest of the paper is organized as follows. A brief introduction of NFT, its root-searching procedures and noise analysis to date is given in Section II. In section III, decoding methods using $a(\lambda)$ without root searching for discrete eigenvalue modulation will be discussed, followed by decoding methods using $b(\lambda_{rt})$ with root searching for QAM on discrete eigenvalues in Section IV. Conclusions will be drawn in Section V.

II. NONLINEAR FOURIER TRANSFORM

The NFT of a signal $q(t)$, supported in the interval $[T_1, T_2]$, is defined by solving the differential system:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \begin{pmatrix} -j\lambda & q(t) \\ -q^*(t) & j\lambda \end{pmatrix} v, v(T_1,\lambda) = \begin{pmatrix} v_1(T_1,\lambda) \\ v_2(T_1,\lambda) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda T_1} \tag{1}$$

where λ and $v(t,\lambda)$ are, respectively, the eigenvalue and eigenvector. Let

$$\begin{pmatrix} a(\lambda) \\ b(\lambda) \end{pmatrix} = \lim_{t\to\infty} \begin{pmatrix} v_1(t,\lambda)e^{j\lambda t} \\ v_2(t,\lambda)e^{-j\lambda t} \end{pmatrix} \tag{2}$$

The NFT is a function of λ defined as:

$$\mathrm{NFT(q)}(\lambda) = \begin{cases} \frac{b(\lambda)}{a(\lambda)} & \lambda \in \mathbb{R} \\ \frac{b(\lambda)}{a'(\lambda)} & \lambda \in S \subset \mathbb{C}^+ \end{cases} \tag{3}$$

where the prime denotes differentiation and S is the set of the (isolated) zeros of the analytic function $a(\lambda_{rt}) = 0$ in the upper half complex plane $\mathbb{C}^+$. The eigenvalues λ can also be thought of as complex-valued frequency. In this paper, we limit transmissions to signals with a small number of discrete eigenvalues with no continuous nonlinear spectral components. Jointly modulating N eigenvalues corresponds to N-soliton transmission.

If we denote $q(t)$ and $\hat{q}(t)$ as the input and output signal to a fiber channel with distance L respectively, let the corresponding nonlinear Fourier coefficients be $a(\lambda)$ and $\hat{a}(\lambda)$. In the ideal case of normalized NLS, it is well known that $a(\lambda) = \hat{a}(\lambda)$[4]. However, in practical systems with fiber loss, amplifier noise and other distortions, $\hat{a}(\lambda)$ is distorted and $a(\lambda) \neq \hat{a}(\lambda)$ in general. Thus, one needs to compute the root $\hat{a}(\lambda_{rt}) = 0$ for every received waveform $\hat{q}(t)$. A standard numerical technique for searching the root of a general function is the Newton-Raphson method[22]. If one starts from a reasonably good initial estimate, the Newton-Raphson method typically need a few iterations to calculate λ_{rt} accurately. However, when initialized randomly or when the function has a complex structure, more iterations will be required and sometimes it may not even converge[5]. Therefore, root searching algorithms are computationally intensive and constitute a major portion of the overall NFT operations for systems in presence of noise and distortions.

Unlike linear communications systems, additive noises do not decompose in a simple fashion in the nonlinear Fourier domain. In the simplest case of fundamental soliton (i.e., solitons with one discrete eigenvalue)[26] where the solution of the NLS is known, the noises/perturbations in the root due to distributed amplification is commonly modelled as Gaussian random process[23]-[25]. For general case of N-solitons (and also signals with continuous nonlinear spectral components), recent works [9],[20],[21] studied noises characteristics for 2-solitons and observed that the noises in the two roots of the 2-solitons are correlated with each other. However, the papers did not propose any analytical model describing the noise correlations and our understanding in the noise processes remains rather limited. Without a proper understanding of the noise characteristics, decoding must rely on the use of empirical statistics (e.g., obtained via the use of pilots), or resort to the use of suboptimal decoding algorithms.

III. DECODING USING $a(\lambda)$ WITHOUT ROOT SEARCHING FOR DISCRETE EIGENVALUE MODULATION

In this section, we consider an optical fiber transmission system based on modulating N discrete eigenvalues $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(N)}$ (in the nonlinear Fourier domain). One may also see this as an On-Off Keying (OOK) system where the presence (ON) or absence (OFF) of a discrete eigenvalue is modulated by the input. Alternatively, this can also be viewed as generalized frequency shift keying(FSK). In this case, there are 2^N possible input signals $q_m(t)$ $(m = 0,1, \dots 2^N - 1)$ with $a_m(\lambda)$ as the corresponding nonlinear Fourier coefficient of $q_m(t)$. As an illustrative example, the real part $Re\{a_m(\lambda)\}$ is shown in Fig. 1(a) for $\lambda \in [j, 2j]$ (where j is the imaginary number or imaginary unity) and m=1,2,3,4. It is clear that the numerical values of $Re\{a_m(\lambda)\}$ are generally different for different m. Consequently, one can choose a certain value of λ (such as $\lambda = 1.5j$ as highlighted by the orange vertical line) and use $Re\{\hat{a}(1.5j)\}$ as a decision metric for m and hence detect the signal $q_m(t)$. In a realistic transmission system with noise and other transmission distortions, Fig. 1(b) shows example empirical distributions of $Re\{\hat{a}_1(1.5j)\}$, $Re\{\hat{a}_2(1.5j)\}$, $Re\{\hat{a}_3(1.5j)\}$ and $Re\{\hat{a}_4(1.5j)\}$ obtained from simulations. The distributions are easily distinguishable from each other and one can expect that $q_m(t)$ can be detected reliably.

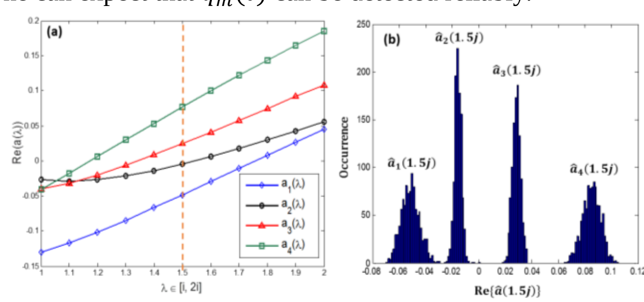

Fig. 1. (a) Re$\{a_m(\lambda)\}$ vs. λ for m=1,2,3,4 (note that each nonlinear Fourier coefficient $a_m(\lambda)$ corresponds to the signal $q_m(t)$). (b) Empirical distributions of $\hat{a}_1(1.5j)$, $\hat{a}_2(1.5j)$, $\hat{a}_3(1.5j)$ and $\hat{a}_4(1.5j)$ in presence of noise and other transmission distortions. The means and co-variances of the distributions can be estimated and used for ML detection.

More generally, we can choose a set of k eigenvalues $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_k\}$ and base our decoding on the vector of $\hat{\boldsymbol{a}}(\boldsymbol{\lambda}) = [\hat{a}(\lambda_1)\ \hat{a}(\lambda_2) \dots \hat{a}(\lambda_k)]$ calculated from the received signal $\hat{q}(t)$. Note that in this proposed algorithm, we are not calculating the root locations λ_{rt} and the k eigenvalues $\lambda_1, \dots, \lambda_k$ chosen are in general different from the modulated eigenvalues $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(N)}$. Let the joint probability density function(pdf) of $\hat{\boldsymbol{a}}(\boldsymbol{\lambda})$ given $\{a_m(\lambda_1), \dots a_m(\lambda_k)\}$ be $P(\hat{\boldsymbol{a}}(\boldsymbol{\lambda})|a_m(\lambda_1), \dots a_m(\lambda_k))$. In this case, the optimal detection strategy is maximum likelihood(ML) detection which is given by

$$m_{ML} = \mathrm{argmax}_m\, P(\hat{\boldsymbol{a}}(\boldsymbol{\lambda})|a_m(\lambda_1), \dots a_m(\lambda_k)). \tag{4}$$

While a full analytical derivation of $P\left(\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})|a_m(\lambda_1),\dots a_m(\lambda_k)\right)$ provides deep insights about noise statistics in NFT systems and are ultimately necessary, one can approximate the distributions as jointly Gaussian and with means μ_m and co-variances Σ_m empirically estimated through pilot symbols or offline simulations.

A. Simulation Results

We evaluate the proposed decoding method using $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ and compare with conventional root-searching methods using $\hat{q}(\lambda_{rt})$ through simulations. Each $q_m(t)$ is recursively computed using the Darboux transformation method[27] and is normalized (by P_0 and T_0) to have the pulse-width of 1 ns (i.e. 1 GBaud transmission). A random pulse train of $q_m(t)$ containing around 1 million bits are generated and launched into a fiber link. The link consists of spans of 25km standard single mode fiber (SSMF) with α = 0.2 dB/km, $\beta_2 = -20.41$ ps^{2}km^{-1} and γ = 1.31 W^{-1}km^{-1}. Each span is followed by an EDFA amplifier with noise figure 3.5dB to compensate the fiber loss. The received signal are then coherently detected and followed by the NFT operation to calculate the nonlinear Fourier coefficients.

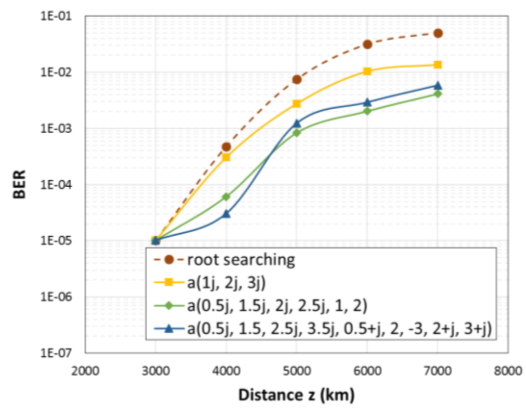

Fig. 2. BER vs. transmission distance for 3-eigenvalue OOK system using conventional root searching and the proposed decoding method using $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ for different sets of $\boldsymbol{\lambda}$.

We first study a 3-eigenvalue OOK system with $\{\lambda^{(1)},\lambda^{(2)},\lambda^{(3)}\} = \{1j,2j,3j\}$. The bit error ratio (BER) for the conventional root searching algorithm and our proposed method were studied for various propagation distances and the results are shown in Fig. 2. When using $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ with $\boldsymbol{\lambda} = \{1j,2j,3j\}$ as decision metrics, the BER is better than the conventional root searching algorithm and the performance gap increases with distance. We would like to point out that the choices of $\boldsymbol{\lambda}$ are not necessarily restricted to the set $\{1j,2j,3j\}$. Instead, one has the freedom to choose (and to optimize the choice of) $\boldsymbol{\lambda}$ as well. To illustrate, we consider two possible choices:

$$\boldsymbol{\lambda} = \{0.5j, 1.5j, 2j, 2.5j, 1, 2\}$$

and

$$\boldsymbol{\lambda} = \{0.5j, 1.5j, 2.5j, 3.5j, 0.5+j, 2, -3, 2+j, 3+j\}.$$

Our choices are hand-picked and are not necessarily optimized. They are selected merely to show that it is not necessary for our proposed algorithm to select $\boldsymbol{\lambda}$ corresponding to the discrete eigenvalues of the inputs and we are still able to obtain BER results better than root searching methods. This fact also suggests that our decoding algorithm may also be employed to inputs which have no discrete eigenvalues at all. It is of future interest (and beyond the scope of this paper) to study systematic approaches to optimize the choice of $\boldsymbol{\lambda}$ for decoding performance and achieve tradeoff between decoding complexity (which is related to the size of $\boldsymbol{\lambda}$) and decoding performance.

Next, we consider another modulation scheme in which all inputs are 2-solitons with the two discrete eigenvalues taken from the set $\lambda^{(1)} \in \{0.3j, 0.45j, 0.6j, 0.75j\}$ and $\lambda^{(2)} \in \{0.9j, 1.05j, 1.2j, 1.35j\}$ i.e. the signal set consists of 16 distinct 2-solitons. The resulting BER using different decoding methods are shown in Fig. 3. Again, decoding with $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ (even with a small set $\boldsymbol{\lambda} = \{j, 2j\}$) outperforms the root searching algorithm. We again want to point out that the choice of $\boldsymbol{\lambda}$ in our new decoding algorithms need not be the roots of the input signals. In this example, the eigenvalues $\boldsymbol{\lambda} = \{1j, 2j\}$ we pick are not discrete eigenvalues of any input. In fact, there are no inputs which have a discrete eigenvalue close to $2j$ at all.

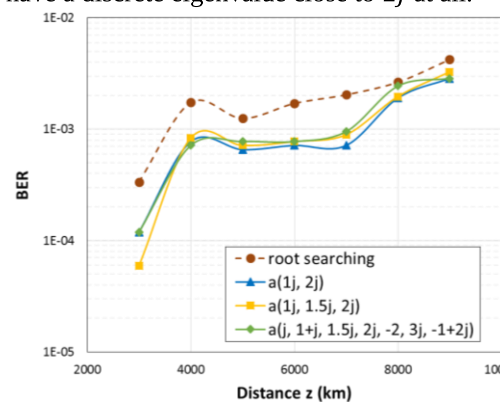

Fig. 3. BER vs. transmission distance for 2-eigenvalue OOK system with $\lambda^{(1)} \in \{0.3j, 0.45j, 0.6j, 0.75j\}$ and $\lambda^{(2)} \in \{0.9j, 1.05j, 1.2j, 1.35j\}$ conventional root searching method and the proposed decoding method using $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ for different sets of $\boldsymbol{\lambda}$.

B. Experimental Verifications

We also conducted experimental verifications of the proposed algorithm. The experimental setup is similar to that described in [18]. The fiber loss is 0.19 dB/km and each span of 50 km is amplified by Raman amplifiers and EDFA. For transmission of 1800km, the BER using the proposed decoding method with different sets of $\boldsymbol{\lambda}$ are shown in Table 1. With $\boldsymbol{\lambda} = \{j, 2j, 3j\}$, the BER is 0.0034, and can be further reduced to 0.0016 if a larger set $\boldsymbol{\lambda} = \{j, 2j, 3j, 1+j, -2\}$ is chosen. In contrast, the BER is at a much higher value of 0.0393 when conventional root searching algorithm is used.

$\boldsymbol{\lambda} = \{\boldsymbol{\lambda_1}, \dots, \boldsymbol{\lambda_k}\}$	BER
{0.5j}	0.1036
{-2}	0.0667
{1+j}	0.0254
{0.5j, 1+j}	0.0143
{0.5j, -2, 1+j}	0.0079
{j, 2j, 3j}	0.0036
{j, 2j, 3j, 1+j, -2}	0.002
{0.5j, 1.5j, 2j, 2.5j, 1, 2}	0.0016

Table 1: BER for 3-eigenvalue OOK transmission over 1800 km using $\widehat{\boldsymbol{a}}(\boldsymbol{\lambda})$ as decision metrics for different sets of $\boldsymbol{\lambda}$.

Similarly, we also conducted experiments for transmission of 16 distinct 2-solitons over 400 km and the BER are shown in Table 2. For the proposed decoding method using $\hat{\boldsymbol{a}}(\boldsymbol{\lambda})$, the BER is 0.0634 when we choose $\boldsymbol{\lambda} = \{0.5j, 1+j\}$ and can be further reduced to 0.0004 if $\boldsymbol{\lambda} = \{0.5j, 1.5j, 2j, 2.5j, 1, 2\}$. This is in contrast with a BER of 0.0086 when conventional root searching method is used. It is interesting to note that if the size of $\boldsymbol{\lambda}$ is too small, the BER performance will be worse than that of conventional root searching. Over all, the insights derived from the experiments generally agree with simulation predictions.

$\boldsymbol{\lambda} = \{\boldsymbol{\lambda_1}, \dots, \boldsymbol{\lambda_k}\}$	BER
{0.5j}	0.1752
{-2}	0.1662
{1+j}	0.1723
{0.5j, 1+j}	0.0634
{0.5j, -2, 1+j}	0.0076
{j, 2j, 3j}	0.0018
{j, 2j, 3j, 1+j, -2}	0.0008
{0.5j, 1.5j, 2j, 2.5j, 1, 2}	0.0004

Table 2: BER for 2-eigenvalue OOK transmission over 400 km with $\lambda^{(1)} \in \{0.3j, 0.45j, 0.6j, 0.75j\}$ and $\lambda^{(2)} \in \{0.9j, 1.05j, 1.2j, 1.35j\}$ using $\hat{\boldsymbol{a}}(\boldsymbol{\lambda})$ as decision metrics for different sets of $\boldsymbol{\lambda}$.

C. Computational Complexity Reduction

In addition to reducing the BER for discrete eigenvalue modulated systems, perhaps a more important advantage of the proposed decoding algorithm is its reduced computational complexity. In our simulations and experiments for the 2-solitons system described above, an average of 8-10 evaluations of $\hat{a}(\lambda)$ are needed to calculate the root λ_{rt} with enough precision. However, this is actually based on the unrealistic assumption that we know the eigenvalues of the original signal $q(t)$ in advance and initialize the root estimates there. If random initial estimates are chosen, the number of iterations will dramatically increase and sometimes the algorithm cannot converge. On the other hand, for our proposed decoding algorithm, the size of $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_k\}$ is at most 6 and hence no more than 6 evaluations of $\hat{a}(\lambda)$ are needed. In fact, even with 2 values of λ (and hence only 2 evaluations of $\hat{a}(\lambda)$), the BER performance is already comparable with the root searching method. This clearly indicates an order-of-magnitude complexity reduction using our proposed decoding method. Moreover, the results showed that there is a tradeoff between the complexity (in terms of the size of $\boldsymbol{\lambda}$) and BER, which is in agreement with expectations. Finally while it is fairly easy to come up with a random set of $\boldsymbol{\lambda}$ that gives good BER performance, we note that a systematic procedure to derive an optimal set of $\boldsymbol{\lambda}$ is yet to be fully investigated.

IV. Decoding using $b(\lambda)$ with Root Searching for QAM on Discrete Eigenvalues

For NFT systems employing phase and amplitude modulation (or generally QAM) on a discrete eigenvalue $\lambda = \lambda_R + j\lambda_I$, we hereby propose an alternative decoding method using the nonlinear Fourier coefficient $b(\lambda)$ defined in (3) that involves root searching. First consider a time waveform $q(t)$ with corresponding spectral amplitudes $q(\lambda)$. Since $e^{j\theta}q(t) \leftrightarrow e^{j\theta}q(\lambda)$ and $q(t-t_0) \leftrightarrow e^{-2j\lambda t_0}q(\lambda)$, modulating the phase and amplitude of $q(\lambda)$ by $Ae^{j\theta} = e^{\ln A + j\theta}$ results in

$$e^{\ln A + j\theta}q(\lambda) = e^{j(\lambda_R + j\lambda_I)\frac{-\ln A}{\lambda_I} + j\left(\theta + \frac{\lambda_R \ln A}{\lambda_I}\right)}q(\lambda)$$

$$\leftrightarrow e^{j\left(\theta + \frac{\lambda_R \ln A}{\lambda_I}\right)}q\left(t - \frac{\ln A}{2\lambda_I}\right)$$

i.e. a time shift and phase shift of the original waveform $q(t)$ as shown in Fig. 4(a) and (b). On the other hand, note that the nonlinear Fourier coefficient $a(\lambda)$ in (3) can be alternatively defined as [4]

$$a(\lambda) = \lim_{t\to\infty} y(t,\lambda)$$

where

$$\begin{cases} \dfrac{d^2 y(t,\lambda)}{dt^2} - \left(2j\lambda + \dfrac{q_t}{q}\right)\dfrac{dy(t,\lambda)}{dt} + |q|^2 y(t,\lambda) = 0, \\ y(-\infty,\lambda) = 1, \quad \dfrac{dy(-\infty,\lambda)}{dt} = 0. \end{cases} \tag{5}$$

Now, since a phase shift in $q(t)$ does not affect $|q(t)|^2$ and q_t/q, Equation (5) is unchanged and hence $a(\lambda)$ is independent of θ. In addition, a time shift $q\left(t - \frac{\ln A}{2\lambda_I}\right)$ results in a corresponding time shift $y\left(t - \frac{\ln A}{2\lambda_I}\right)$ while $a(\lambda) = \lim_{t\to\infty} y\left(t - \frac{\ln A}{2\lambda_I}\right)$ remains unchanged. Taken together, one can conclude that for QAM systems on a specific discrete eigenvalue λ, $a(\lambda)$ and hence $a'(\lambda)$ does not contain any information about the phase and amplitude information. In practical systems with noise and other distortions, $a(\lambda)$ is merely needed to compute the root location $\hat{a}(\lambda_{rt}) = 0$ and once λ_{rt} is obtained, $\hat{b}(\lambda_{rt})$ actually contains all the information in $\hat{q}(\lambda_{rt})$ for symbol detection purposes.

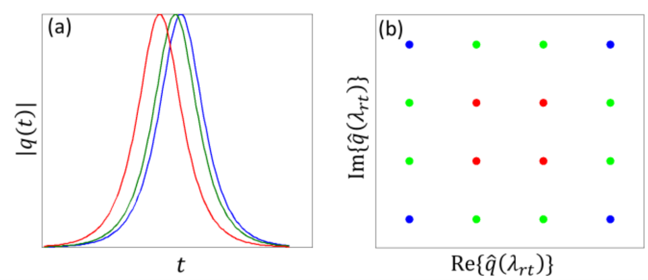

Fig. 4. (a),(b): Amplitude and phase modulation on a given discrete eigenvalue λ_{rt} translate into modulating the timing and overall phase of a specific waveform $|q(t)|$ (each pulse in (a) corresponds to the set of signal points in (b) with the same color). The nonlinear Fourier coefficient $a(\lambda)$ are insensitive to such modulation.

Simulations are conducted to study the characteristics of $\hat{a}'(\lambda_{rt})$, $\hat{q}(\lambda_{rt})$ and $\hat{b}(\lambda_{rt})$ in a 16-QAM system on $1j$. The simulation configuration is same as that in section III. The fiber link consists of spans of 25 km standard single mode fiber (SSMF) with $\alpha = 0.2$ dB/km, $\beta_2 = -20.41$ ps^{2}km^{-1} and $\gamma = 1.31$ W^{-1}km^{-1}. Each span is followed by an EDFA to compensate the fiber loss with noise figure 3.5 dB. After coherent detection at the receiver, $\lambda_{rt}, \hat{a}'(\lambda_{rt}), \hat{b}(\lambda_{rt})$ and $\hat{q}(\lambda_{rt})$ are computed for each received pulse. Note that 128 samples per pulse are used to minimize numerical errors.

Fig. 5 shows the received scatter plots of $\hat{a}'(\lambda_{rt})$, $\hat{q}(\lambda_{rt})$ and $\hat{b}(\lambda_{rt})$ for 16-QAM system on $1j$ transmitted over 4000 km. It can be seen that $\hat{a}'(\lambda_{rt})$ is pre-dominantly corrupted by amplitude noise. Therefore, the amplitude noise variance of $\hat{b}(\lambda_{rt})$ is smaller than that of $\hat{q}(\lambda_{rt})$. The phase noise however, will be largely unaffected. The division operation $\hat{q}(\lambda_{rt}) = \hat{b}(\lambda_{rt}) \;/\; \hat{a}'(\lambda_{rt})$ merely adds an unnecessary noise component and should be discarded. Therefore, for systems using QAM on a specific eigenvalue, we can use $\hat{b}(\lambda_{rt})$ for symbol decoding instead.

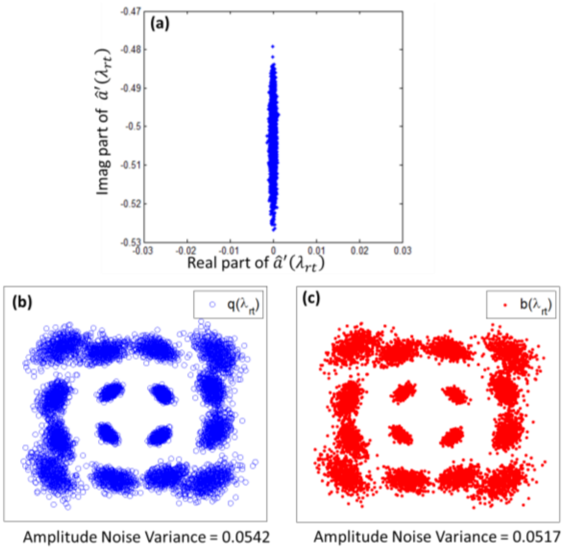

Fig. 5. (a) Received scatter plots of $\hat{a}'(\lambda_{rt})$, (b) $\hat{b}(\lambda_{rt})$ and (c) $\hat{q}(\lambda_{rt})$ for 16-QAM on 1j over 4000 km. As the received $\hat{a}'(\lambda_{rt})$ consists primarily of amplitude noise, the amplitude noise variance of $\hat{b}(\lambda_{rt})$ is smaller than that of $\hat{q}(\lambda_{rt})$ while the phase noise variance of remains largely the same.

Besides the fact that $\hat{a}'(\lambda_{rt})$ is unnecessary and worsen the detection quality, $\hat{a}'(\lambda_{rt})$ are actually highly correlated with noise of $\hat{b}(\lambda_{rt})$. Fig. 6 shows the empirical joint distribution of the phase noise $\triangle\theta_b$ of $\hat{b}(\lambda_{rt})$ vs. imaginary noise $\Delta a'_I$ of $\hat{a}'(\lambda_{rt})$. The correlation coefficient has a very high value of 0.825.

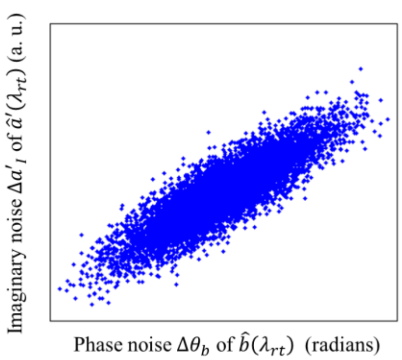

Fig. 6. Scatter plot of received phase noise $\Delta\theta_b$ of $\hat{b}(\lambda_{rt})$ and imaginary noise component $\Delta a'_I$ of $\hat{a}'(\lambda_{rt})$ for 16-QAM on 1j over 4000 km with inline amplifier noise. The empirical correlation coefficient between $\Delta\theta_b$ and $\Delta a'_I$ is 0.825.

In fact, Table 3 shows the empirical correlations between the phase and amplitude noise $\Delta\theta_b$, $\Delta\Gamma_b$ of $\hat{b}(\lambda_{rt})$ and the real and imaginary noise components $\Delta\lambda_R$, $\Delta\lambda_I$, $\Delta a'_R$ $\Delta a'_I$ of λ_{rt} and $\hat{a}'(\lambda_{rt})$ respectively. The correlation is most pronounced between phase noise $\Delta\theta_b$ and $\Delta a'_I$,$\Delta\lambda_I$ and between amplitude noise $\Delta\Gamma_b$ and $\Delta\lambda_R$. We note that a subset of correlation properties identified here is also reported by Hari and Kschischang in a recent contribution[9].

Correlation coefficients	Phase noise $\Delta\theta_b$	Amplitude noise $\Delta\Gamma_b$
$\Delta a'_R$	-0.015926	-0.121288
$\Delta a'_I$	0.864393	0.028622
$\Delta\lambda_R$	0.005491	0.799362
$\Delta\lambda_I$	0.860241	0.031317

Table 3: Empirical correlation coefficients between real and imaginary noise $\Delta a'_R$, $\Delta a'_I$ of $\hat{a}'(\lambda_{rt})$, real and imaginary eigenvalue noise $\Delta\lambda_R$ $\Delta\lambda_I$ with the amplitude and phase noise $\Delta\Gamma_b$, $\Delta\theta_b$ of $\hat{b}(\lambda_{rt})$.

The correlation between various noises $\Delta a'_R$, $\Delta a'_I$, $\Delta\lambda_R$, $\Delta\lambda_I$, and $\Delta\Gamma_b$, $\Delta\theta_b$ should be appropriately characterized and used to reduce the noise in $\hat{b}(\lambda_{rt})$. While a joint pdf of all the noise should ultimately be in place, one can derive a simple linear minimum mean-square estimate (LMMSE) of $\Delta\Gamma_b$ and $\Delta\theta_b$ from $\mathbf{n} = [\Delta a'_R \;\; \Delta a'_I \;\; \Delta\lambda_R \;\; \Delta\lambda_I]$ to reduce their variance and improve detection performance. In particular, we seek vectors **c**, **d** so that the residual amplitude and phase noise variance of $\hat{b}(\lambda_{rt})$ is minimized, i.e.

$$\begin{aligned} &\text{argmin}_{\mathbf{c}} \mathbf{E}\left[(\Delta\Gamma_b - \mathbf{c}\mathbf{n}^T)^2\right] \\ &\text{argmin}_{\mathbf{d}} \mathbf{E}\left[(\Delta\theta_b - \mathbf{d}\mathbf{n}^T)^2\right] \end{aligned} \tag{6}$$

where $\mathbf{E}[\cdot]$ denotes expectation. This is standard mean square estimation and linear prediction of correlated random variables [28] and **c**, **d** are given by

$$\begin{aligned} \mathbf{c} &= \mathbf{E}\left[\Delta\Gamma_b \mathbf{n}\right]\cdot \text{cov}(\mathbf{n})^{-1} \\ \mathbf{d} &= \mathbf{E}\left[\Delta\theta_b \mathbf{n}\right]\cdot \text{cov}(\mathbf{n})^{-1} \end{aligned} \tag{7}$$

cov(•) denotes the covariance matrix.

A. *Simulation Results*

With the above derivations, the received signal distributions using the proposed LMMSE method are shown in Fig. 7 for QPSK and 16-QAM transmissions on 1j. We also compared our results with the MMSE phase rotation method in [11]. It is clear that leveraging the noise correlations between $\hat{b}(\lambda_{rt})$ and $\hat{a}'(\lambda_{rt})$ and $\Delta\lambda$ results in much better received signal distributions. The evolution of phase and amplitude noise variance for decoding using $\hat{b}(\lambda_{rt})$ with different compensation methods are shown in Fig. 8. The solid lines represent the performance of 16QAM while the dash lines represent QPSK systems. The performance using the original $\hat{b}(\lambda_{rt})$ without additional signal processing is also shown as a reference. It is clear that both the MMSE phase method and our LMMSE phase and amplitude method can largely reduce the phase noise variance by more than half. However, our method can also decrease the amplitude noise variance, which is distinctly illustrated in Fig. 8(b). Therefore, our method should have better BER performance especially for phase and amplitude modulated systems such as 16-QAM.

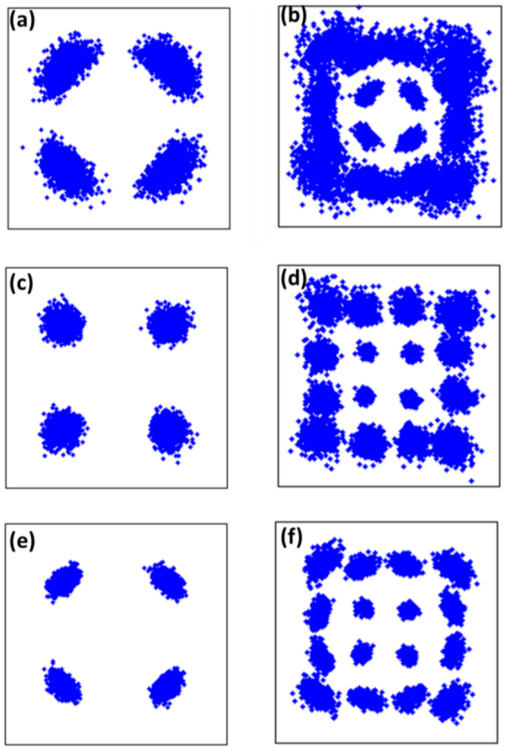

Fig. 7. Received distributions of $\hat{b}(\lambda_{rt})$ for QPSK and 16-QAM transmissions on $1j$ over 4000 km a),b) with no additional DSP. c),d) with the MMSE phase method[11] e),f) with the LMMSE phase and amplitude method.

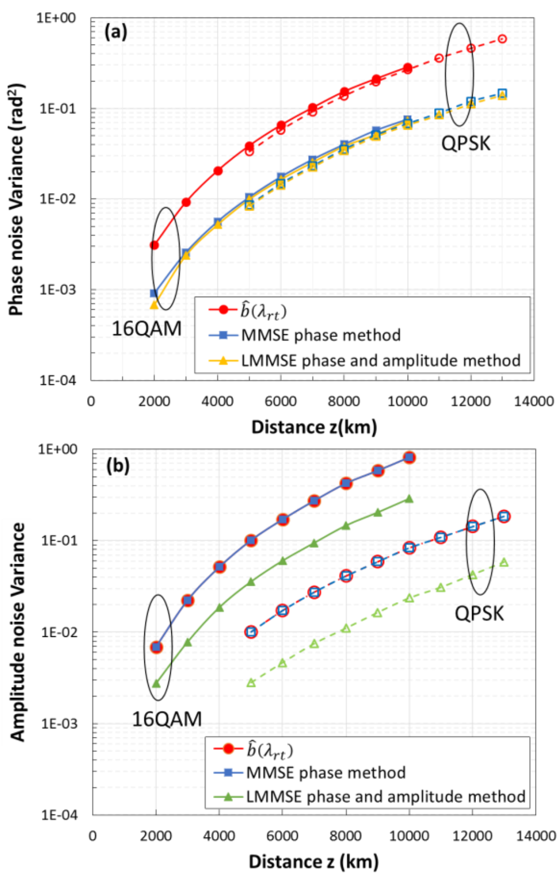

Fig. 8. a) Phase noise variance and b) amplitude noise for QPSK and 16-QAM transmissions on $1j$ for various distances.

Fig. 9 shows the corresponding BER for QPSK and 16-QAM systems on $1j$ as a function of distance using various compensation methods. For a BER threshold of 2e-2, our method achieves almost 1000km longer transmission distance than the MMSE method for 16QAM, while for QPSK a modest distance extension is obtained.

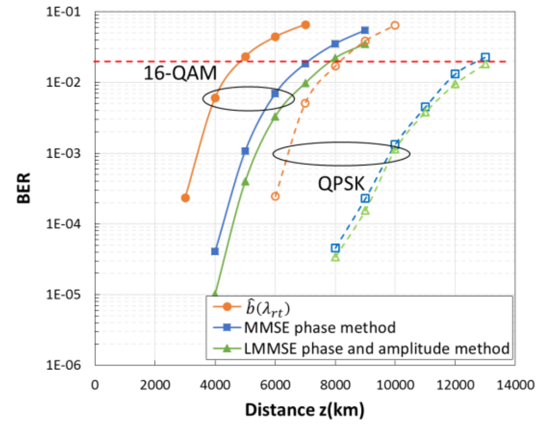

Fig. 9. BER vs. transmission distance for QPSK and 16-QAM transmissions on $1j$.

B. *Experimental Verifications*

Experiments are also conducted to verify the proposed decoding algorithm using $\hat{b}(\lambda_{rt})$. The experimental setup is similar to that described in [18] with the exception that the fiber recirculating loop is changed to NZ-DSF fiber (with α = 0.19 dB/km, $\beta_2 = -5.01$ ps^{2}km^{-1} and γ = 1.2 W^{-1}km^{-1}) with 50-km span length and lumped amplification by EDFA. We transmitted 2 GBaud 16-QAM signals on $0.5j$ in one polarization while a pilot tone from the other polarization is used to estimate and compensate the laser phase noise and frequency offset. The NFT operation and root searching algorithm is used to find the roots and calculate the nonlinear Fourier coefficients. Fig. 10 shows the BER as a function of distance using $\hat{q}(\lambda_{rt})$, $\hat{b}(\lambda_{rt})$ and $\hat{b}(\lambda_{rt})$ with the proposed LMMSE phase and amplitude method. The results clearly show that decoding using $\hat{b}(\lambda_{rt})$ outperforms decoding using $\hat{q}(\lambda_{rt})$ and the LMMSE method further improve the BER performance.

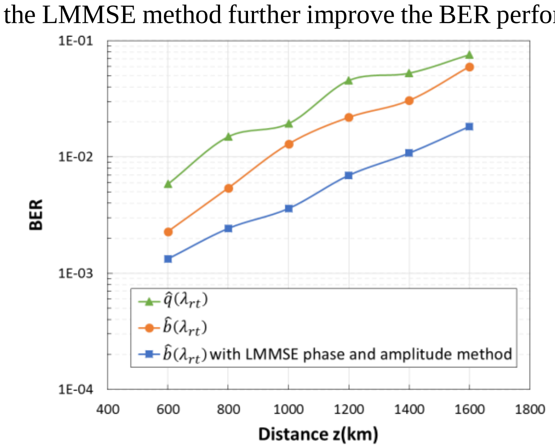

Fig. 10. BER vs. transmission distance for 2 GBaud 16-QAM transmissions on $0.5j$.

C. Extensions to QAM on multiple eigenvalues

The same noise compensation strategy can be extended to systems with independent QAM signals on multiple eigenvalues. We simulated a transmission system with independent QPSK on $1j$, $2j$ over 1500km. The empirical correlations between various amplitude and phase noises are shown in table 4. It should be noted that some of the noises across different eigenvalues are also strong correlated.

Correlation coefficients	$\Delta\theta_{b,1}$	$\Delta\Gamma_{b,1}$	$\Delta\theta_{b,2}$	$\Delta\Gamma_{b,2}$
$\Delta a'_{R,1}$	0.012	-0.798	-0.014	0.822
$\Delta a'_{R,2}$	-0.012	0.799	0.014	-0.818
$\Delta a'_{I,1}$	-0.899	-0.018	0.863	-0.002
$\Delta a'_{I,2}$	0.896	0.019	-0.847	-0.00007
$\Delta\lambda_{R,1}$	-0.013	0.840	0.016	-0.747
$\Delta\lambda_{R,2}$	0.011	-0.722	-0.011	0.832
$\Delta\lambda_{I,1}$	0.896	0.020	-0.828	0.001
$\Delta\lambda_{I,2}$	-0.799	-0.012	0.877	0.002

Table 4: Empirical correlation coefficients between real and imaginary noise $\Delta a'_{R,1(2)}$, $\Delta a'_{I,1(2)}$ of $\hat{a}'(\lambda_{rt})$ for 2 eigenvalues, real and imaginary eigenvalue noise $\Delta\lambda_{R,1(2)}$ $\Delta\lambda_{I,1(2)}$with the amplitude and phase noise $\Delta\Gamma_{b,1(2)}$, $\Delta\theta_{b,1(2)}$ of $\hat{b}(\lambda_{rt})$ for independent QPSK on $1j$, $2j$ over 1500km.

The received signal distribution results for independent QPSK on eigenvalues $1j$, $2j$ are shown in Fig. 11 (c) using different compensation methods. In comparison to the MMSE phase method (Fig. 7 (b)), we can easily observe that our proposed method results in much better received distributions. This is also reflected in the BER performance shown in Fig. 12. The MMSE method exceeds the 2e-2 BER threshold after transmission of 2250km, while the LMMSE method can substantially extend the transmission distance. The results suggest that our method is more robust and suitable for systems with independent QAM modulation on multiple eigenvalues.

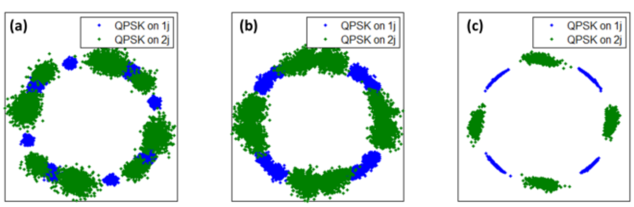

Fig. 11. Received distributions of $\hat{b}(\lambda_{rt})$ for independent QPSK transmissions on $1j$ and $2j$ over 1500 km a) with no additional DSP. b) with the MMSE phase method[11] c) with the LMMSE phase and amplitude method.

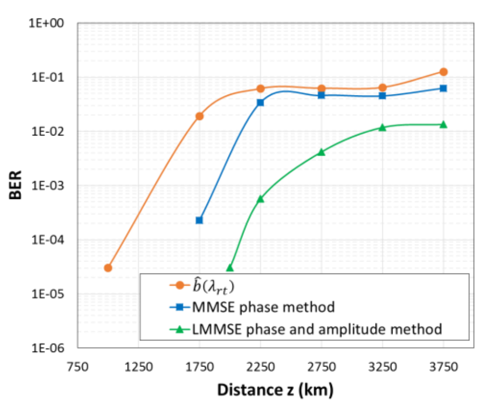

Fig. 12. BER vs. distance for independent QPSK on $1j$, $2j$ using various compensation methods.

Finally, it should be noted that since we studied systems with fiber loss in the paper, the numerical values of noise correlation coefficients will differ for lossless systems and are generally sensitive to other system parameters. However, the general performance trend of the proposed decoding method and conclusions presented herein will still apply. In a practical setting, one can first obtain the table of correlation coefficients offline through simulations or experiments before applying the LMMSE filter during data transmission.

V. Conclusions

Communications based on nonlinear Fourier Transform might hold key to further advances in transmission capacities for long-haul optical communication systems. For discrete eigenvalue modulated systems, we proposed to choose a set of eigenvalues $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_k\}$ around the roots $\hat{a}(\lambda_{rt}) = 0$ and use the spectral coefficients $\hat{\boldsymbol{a}}(\boldsymbol{\lambda})$ as decision metrics. We obtained comparable performance with conventional root searching techniques could be obtained with much lower computational complexity since root searching is avoided in the calculating $\hat{\boldsymbol{a}}(\boldsymbol{\lambda})$. For QAM systems on discrete eigenvalues, we proposed to use $\hat{b}(\lambda_{rt})$ as decision metrics. We showed show that the correlations between the noises in $\hat{a}'(\lambda_{rt})$, λ_{rt} and $\hat{b}(\lambda_{rt})$ can be used to develop a linear minimum mean square error(LMMSE) estimator of the noise in $\hat{b}(\lambda_{rt})$ and improve detection performance. More in-depth analytical studies of the noise correlations together with extensions of the proposed decoding methods to systems with joint discrete and continuous eigenvalue modulation will be important topics for future research.

VI. Acknowledgements

The authors would like to acknowledge the support of the Hong Kong Government General Research Fund under project number PolyU 152116/15E and Australian Research Council (ARC) under ARC DP150103658.

References

[1] E. G. Turitsyn et al, Digital signal processing based on inverse scattering transform, *Opt. Lett.* vol. 38, pp. 4186-4188, 2013.

[2] J. E. Prilepsky et al, “Nonlinear spectral management: Linearization of the lossless fiber channel” *Opt. Express*, vol. 21, pp. 24344-24367, 2013.

[3] H. Bulow, “Experimental Demonstration of Optical Signal Detection Using Nonlinear Fourier Transform," *IEEE J. Lightw.Technol*, vol. 33, no. 7, pp. 1433–1439, Apr. 2015.

[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.M. I.

[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] V. E. Zakharov and A. B. Shabat, “Exact theory of 2-dimensional self-focusing and onedimensional self-modulation of waves in nonlinear media,'" *Sov. Phys.-JETP*, vol .34, pp. 62-69, 1972.

[8] S. A. Derevyanko J. E. Prilepsky and S. K. Turitsyn, "Capacity estimates for optical transmission based on the nonlinear Fourier transform", *Nature Commun.*, vol. 7, 12710, 2016.

[9] S. Hari and F. R. Kschischang, "Bi-Directional Algorithm for Computing Discrete Spectral Amplitudes in the NFT", *IEEE J. Lightw.Technol*, vol. 34, no. 15, Aug. 2016.

[10] S. Hari and F.R. Kschischang, "Multieigenvalue communications," *IEEE J. Lightw.Technol*, vol. 34, pp. 3110-3117, 2016.

[11] V. Aref, H. Buelow, and K. Schuh, "On Spectral Phase Estimation of Noisy Solitonic Transmission", paper W3A.3, *OFC 2016*.

[12] A. Hasegawa and T. Nyu "Eigenvalue communication" *IEEE J. Lightw.Technol*, vol. 11, no. 3, pp. 395–398, Apr. 1993.

[13] J. E. Prilepsky et al, "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," *Phys. Rev. Lett.* vol. 113, 013901, 2014.

[14] S. T. Le et al., "First Experimental Demonstration of Nonlinear Inverse Synthesis Transmission over Transoceanic Distances," paper Tu2A.1, *OFC 2016*.

[15] S.T. Le, I.D. Philips, J.E. Prilepsky, P. Harper, A.D. Ellis, and S.K. Turitsyn, "Demonstration of Nonlinear Inverse Synthesis Transmission Over Transoceanic Distances," *IEEE J. Lightw.Technol*, vol. 34 (10), pp. 2459-2466, 2016.

[16] Y. Matsuda, H. Terauchi, and A. Maruta, "Design of Eigenvalue-multiplexed Multi-level Modulation Optical Transmission System,", Paper TH12B3, *OECC/ACOF 2014*.

[17] A. Maruta, A. Toyota, Y. Matsuda, and Y. Ikeda, "Experimental demonstration of long haul transmission of eigenvalue modulated signals," Tyrrhenian International Workshop on Digital Comminications (TIWDC), Florence, Italy, pp. 28-30, 2015.

[18] Z. Dong, S. Hari, T. Gui, K. Zhong, M. Yousefi, C. Lu, P.-K. Alexander Wai, F. R. Kschischang and A.P.T. Lau, "Nonlinear Frequency Division Multiplexed Transmissions based on NFT," *Photon. Technol. Lett*., Vol. 27, No. 15, 2015.

[19] H. Bülow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," preprint arXiv:1605.08069, 2016.

[20] W, Q. Zhang, Q. Zhang, K. Amir, T. Chan, and S. Afshar V, "Noise Correlation between Eigenvalues in Nonlinear Frequency Division Multiplexing," *Photonics and Fiber Technology Conference*, Sept. 2016.

[21] H. Buelow et al., "Experimental Nonlinear Frequency Domain Equalization of QPSK Modulated 2-Eigenvalue Soliton", paper Tu2A.3, *OFC 2016*.

[22] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd ed. Springer-Verlag, 1993.

[23] J. P. Gordon and H. A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission", *Opt. Lett.* 11 (10), 665 (1986).

[24] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks* (Wiley, New York, 1998)

[25] Q. Zhang and T. 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)*, June 2015, pp. 455–459

[26] Q. Zhang, T. H. Chan, and A. Grant, "Spatially periodic signals for fiber channels," in *2014 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2014.

[27] V. B. Matveev and M. A. Salle, *Darboux transformations and solitons*. Springer, 1991.

[28] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.

Tao Gui received his BS degree in 2011 from Xidian University and M.Sc. degree in 2014 from Jinan University, China. Now, he is a PhD student majoring in optical communication engineering at Hong Kong Polytechnic University. He worked at China State Key Laboratory of Optical Comm. Technologies and Networks at Wuhan in 2013 on long haul and large capacity optical transmission using superchannel OFDM techniques. His current research interests are DSP in optical transmission system and Nonlinear Frequency Division Multiplexing systems for fiber communication.

Terence Chan received his B.Sc (Math), Master's and Ph.D. degrees in Information Engineering in 1996, 1998 and 2000 respectively, all from The Chinese University of Hong Kong. In 2001, he was a visiting assistant professor in the Department of Information Engineering at the same university. From February 2002 to June 2004, he was a Post-doctoral Fellow at the Department of Electrical and Computer Engineering at the University of Toronto. He was an assistant professor in University of Regina from 2004-2006. He is currently an Associate Professor in Institute for Telecommunications Research at University of South Australia. He received the Croucher Foundation Fellowship in 2002 respectively. He respectively serves as the Technical and General Co-Chair for the 2011 and 2015 IEEE International Symposium On Network Coding.

Chao Lu obtained B.Eng. degree in electronic engineering from Tsinghua University, China in 1985, and M.Sc. degree and Ph.D. degree from University of Manchester, UK, in 1987 and 1990, respectively. He joined the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore as Lecturer in 1991 and has been an Associate Professor since January 1999. From June 2002 to December 2005, he was seconded to the Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore, as Program Director and Department Manager, helping to establish a research group in the area of optical communication and fibre devices. Since April 2006, he has been a Professor in the Department of Electronic and Information Engineering, Hong Kong Polytechnic University. His research interests are optical communication systems and networks, fibre devices for optical communication and sensor systems.

Alan Pak Tao Lau received his B.A.Sc in Engineering Science (Electrical Option) and M.A.Sc. in Electrical and Computer Engineering from University of Toronto in 2003 and 2004 respectively. He obtained his Ph.D. in Electrical Engineering at Stanford University in 2008 and has joined the Hong Kong Polytechnic University as an Assistant Professor and. He worked at NEC Labs America in summer 2006 on receiver structures for multi-mode fiber systems. He is now an Associate Professor and his current research interests include various aspects of coherent fiber-optic communication systems.

P. K. A. Wai received the B.S. (hons.) degree from the University of Hong Kong, Pokfulam, Hong Kong, in 1981, and the M.S. and Ph.D. degrees from the University of Maryland, College Park, in 1985 and 1988, respectively. In 1988, he joined Science Applications International Corporation, McLean, VA, USA, where he was a Research Scientist involved with the Tethered Satellite System project. In 1990, he became a Research Associate with the Department of Physics and the Department of Electrical Engineering, University of Maryland.

In 1996, he joined the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, where he became a Chair Professor of optical communications in 2005, and currently he is the Vice President (Research Development). He is an active con-tributor to the field of photonics and optical communications, having authored or coauthored more than 300 international refereed publications. His research interests include soliton, fiber lasers, modeling and simulations of optical devices, long-haul optical fiber communications, all-optical packet switching, and network theories. Dr. Wai is a Fellow of the IEEE, and the Optical Society of America. He serves as a Guest Editor of the IEEE Journal Of Selected Areas of Quantum Electronics. Currently, he is an Associate Editor of IEEE/OSA Journal Of Lightwave Technology.