Kernel Affine Projection for Nonlinearity Tolerant Optical Short Reach Systems

Lu Zhang, *Member, IEEE*, Jiajia Chen, *Senior Member, IEEE*, Aleksejs Udalcovs, *Member, IEEE*, Xiaodan Pang, *Senior Member, IEEE*, Richard Schatz, Urban Westergren, Sergei Popov, *Senior Member, IEEE*, Shilin Xiao, *Member, IEEE*, and Oskars Ozolins, *Member, IEEE*

***Abstract*— Nonlinearity is one of the key issues that hinder the development of high-capacity optical short reach systems. This paper proposes three variants of kernel affine projection (KAP) algorithms, all of which combine kernel mapping and affine projection in a reproducing kernel Hilbert space for compensating nonlinear impairments in optical short reach systems. An intensity modulation/direct detection system with a single digital-to-analogue converter, a packaged externally modulated laser and a packaged InP photo-detector is used for experimental demonstration, achieving 238-Gbps (net rate 222-Gbps). Experimental results show that the KAP algorithms can mitigate nonlinear impairments in short-reach communications while maintaining low complexity in reproducing kernel Hilbert space.**

***Index Terms*— Equalizers, optical fiber communication, communication system nonlinearities.**

I. Introduction

THE increase in datacenter traffic drives to large capacity requirements for short reach interconnections [1]. The standardization group of next generation short-reach interfaces aims at 800-GbE and beyond [2], in which the data rate per lane is expected to be upgraded to 100-Gbps, 200-Gbps or even higher [2]–[4]. To meet such a high data rate, the intensity modulation/direct detection (IM/DD) systems with advanced modulation formats such as pulse amplitude modulation or discrete multi-tone (DMT) have been experimentally demonstrated [5]. However, the impairments from opto-electronic devices, such as the chirp characteristics from the direct modulated laser [6], nonlinear optical modulator transfer function [7], saturated power amplification from the optical/electrical amplifier, inter-subcarrier signal-to-signal beating interference at the photodetector [8], have become a major performance-limiting factor in high-speed optical short-reach systems. As a result, digital signal processing algorithms have been proposed for compensating the aforementioned nonlinear impairments [5].

The signal recovery process in the Hilbert space [14] is considered as a most straightforward method for compensating nonlinear impairments in optical short-reach communication systems. In the Hilbert space, any signal vector could be expanded on the basis vectors of the Hilbert space. The properties of the Hilbert space make it suitable for solving convex optimization problems, such as adaptive filtering. Several nonlinear adaptive filtering schemes that are always carried out in the Hilbert space, such as the Volterra filtering [9], [10] and machine learning [11], [12] based equalization schemes, have been introduced to realize 200-Gbps and beyond per lane for optical short-reach communication systems. However, the curse of dimension dilemma [13] induced high complexity in these methods hinders their implementation in optical short-reach systems. When the signal mapping order (i.e. signal dimension) increases, the performance can be improved but causing a significant increase in complexity. The Volterra filtering employs high-order mapping of the input signal X, such as X^2 (mapping order of 2) and X^3 (mapping order of 3) multiplications, the complexity increases exponentially with signal mapping orders. In the practical implementation, we can limit the mapping order up to 3, i.e. X^3. However, the complexity is still high and the performance to mitigate nonlinear impairments is degraded. Meanwhile, machine learning methods, such as neural network, compensate the nonlinear impairments through multiple neural layers. The complexity is even higher than the Volterra filtering and its implementation

Manuscript received December 11, 2019; revised May 11, 2020; accepted June 30, 2020. Date of publication July 7, 2020; date of current version October 16, 2020. The work is supported by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF), STINT Joint China-Sweden Mobility Programme, EU H2020 MSCA-IF Project NEWMAN (no. 752826), VINNOVA-funded SENDATE-EXTEND, SENDATE-FICUS (project no. 2017-01559), EU H2020 project TWILIGHT (no. 871741) and the Fundamental Research Funds for the Central Universities (no.2020QNA5012). The associate editor coordinating the review of this article and approving it for publication was M. Safari. *(Corresponding authors: Lu Zhang; Jiajia Chen; Oskars Ozolins.)*

Lu Zhang is with the College of Information Science and EE, Zhejiang University, Hangzhou 310027, China, and also with the Zhejiang Lab, Hangzhou 310000, China (e-mail: zhanglu1993@zju.edu.cn).

Jiajia Chen is with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: jiajiac@chalmers.se).

Aleksejs Udalcovs is with the NETLAB, RISE AB, 164 25 Stockholm, Sweden (e-mail: aleksejs.udalcovs@ri.se).

Xiaodan Pang is with the KTH Royal Institute of Technology, 164 40 Stockholm, Sweden, and also with Infinera Corporation, 117 43 Stockholm, Sweden (e-mail: xiaodan@kth.se).

Richard Schatz, Urban Westergren, and Sergei Popov are with the KTH Royal Institute of Technology, 164 40 Stockholm, Sweden (e-mail: rschatz@kth.se; urban@kth.se; sergeip@kth.se).

Shilin Xiao is with the State Key Laboratory of Advanced Optical Communication System and Networks, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: slxiao@sjtu.edu.cn).

Oskars Ozolins is with the NETLAB, RISE AB, 164 25 Stockholm, Sweden, and also with the Faculty of Optical Communication, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: oskars.ozolins@ri.se).

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

Digital Object Identifier 10.1109/TCOMM.2020.3007643

in the physical layer design of optical short-reach systems is challenging.

The aforementioned nonlinear impairment compensation methods are carried out in the Hilbert space, small performance improvement requires a dramatic increase of the calculation complexity when the signal mapping order is high. Our previous works in [14] for the first time introduced kernel mapping methods in optical short-reach communication systems, where a kernel function is used for signal mapping. As shown in Part A of Section II in this paper, such a method [14], [15] makes linear signal processing carried out in the high dimensional reproducing kernel Hilbert space yielding powerful extensions to compensate the nonlinear impairments. Two kernel mapping methods have been devised, namely kernel least-mean-square (KLMS) [15] and kernel recursive-least-squares (KRLS) [16]. On the one hand, the KLMS is with the linear number of multiplications in term of training data size, while the KRLS is with the quadratic number of multiplications in term of training data size. On the other hand, the KRLS greatly outperforms the KLMS on transmission performance. Both the KLMS and KRLS can achieve comparable performance with the Volterra filtering, and the KLMS shows a smaller number of multiplications than the high-order Volterra filtering.

In this paper, we present kernel affine projection (KAP) algorithms that combine kernel mapping and affine projection for channel equalization in optical short-reach communications. As an extension of our previous work in [17], three variants of the KAP equalizers, namely the gradient descent method, stochastic Newton's recursion method and regularized Newton's recursion method are introduced for coefficients convergence. In addition, error reuse criterion is included for complexity reduction. All three variants of the KAP can perform as good as the KRLS and common Volterra filtering considering nonlinear kernels up to the 4th order, while the required number of multiplications in the regularized Newton's recursion based KAP is lower than the KLMS and common Volterra filtering.

II. Operational Principles

In this section, we depict operational principles of affine projection algorithm and kernel affine projection algorithm. The description starts from a brief introduction of the basic kernel mapping. The main idea of the KLMS and KRLS is reviewed.

A. Kernel Mapping

Kernel mapping follows the classic methodology for linear adaptive filtering, while using Mercer kernel function as input signal mapping function. Mercer kernel $\kappa(x, x')$ is a continuous and symmetric basis function defined in the reproducing kernel Hilbert space, where x and x' are the signal vectors. In this paper, the Gaussian kernel function with inner production manner is utilized as the expression of Mercer kernel $\kappa(x, x')$:

$$\kappa(\boldsymbol{x}, \boldsymbol{x'}) = exp(-||\boldsymbol{x} - \boldsymbol{x'}||^2/2h^2), \tag{1}$$

where h is the Gaussian kernel bandwidth. Functions f and g are defined as follows:

$$f = \mathbf{\Sigma}_{i=1}^{N} \alpha_i \kappa(\boldsymbol{x_i}, \cdot), g = \mathbf{\Sigma}_{j=1}^{M} \beta_j \kappa(\boldsymbol{x_j}, \cdot), \tag{2}$$

where $\alpha_i$and β_j are the coefficients of the vectors $x_\mathbf{i}$ and $x_\mathbf{j}$, respectively. The bilinear form of f and g is defined as:

$$< f, g >= \mathbf{\Sigma}_{i=1}^{N} \mathbf{\Sigma}_{j=1}^{M} \alpha_i \beta_j \kappa(\boldsymbol{x_i}, \boldsymbol{x_j}). \tag{3}$$

The bilinear form in (3) satisfies the symmetry, linearity and positive definitive properties [14] and (3) can be considered as the inner product. Besides, it also meets the condition of the reproducing property [14], expressed as:

$$< f, \kappa(\cdot, \boldsymbol{x_j}) >= \mathbf{\Sigma}_{i=1}^{N} \alpha_i \kappa(\boldsymbol{x_i}, \boldsymbol{x_j}). \tag{4}$$

Thus, the kernel function meets the conditions of both the inner product and the reproducing property. Such a function is recognized as the reproducing kernel function, and the space composed of the kernel functions is called reproducing kernel Hilbert space. The vector signal x in the Hilbert space is mapped to the reproducing kernel Hilbert space by the kernel function $\kappa(., x)$, and the linear processing mechanism $f + g$ is transformed to an inner product calculation $< f, g >$. An important rule of the kernel function $\kappa(x, x')$ is expressed as (2).

$$\kappa(\boldsymbol{x}, \boldsymbol{x'}) =< \varphi(\boldsymbol{x}), \varphi(\boldsymbol{x'}) >= \varphi(\boldsymbol{x})^{\mathrm{T}} \varphi(\boldsymbol{x'}). \tag{5}$$

The kernel mapping function φ is the feature vector (i.e. basis vector) of the reproducing kernel Hilbert space. Our previous work in [14] demonstrates the Gaussian kernel function is capable of processing all the orders of the kernels, i.e., (1, x, $x^{\mathbf{2}}$, $x^{\mathbf{3}}$, …), where the higher orders ($x^{\mathbf{2}}$, $x^{\mathbf{3}}$, …) represent nonlinear impairments, the kernel mapping functions of (1) is:

$$\begin{aligned} &\varphi(x_i) \\ &= exp((-x_i^2/2h^2)^*(1, {a_1}^* x_i/h, {a_2}^* x_i^2/h^2, {a_3}^* x_i^3/h^3 \ldots), \end{aligned} \tag{6}$$

where x_i is an element of vector x. Therefore, after the kernel mapping the nonlinear effects are transformed into the inner production in (1). The usage of the Gaussian kernel function does not need the concrete high-dimensional expansion shown in (6). As a result, the strategy to employ kernel mapping is to formulate the classic linear signal processing in the reproducing kernel Hilbert space to iteratively mitigate nonlinear impairments. In our previous work [14], we have introduced the mathematical principles of kernel mapping and derived the KLMS and KRLS algorithms for compensating nonlinear impairments in optical short reach systems.

The main principle of KLMS is overviewed as follows.

$$e(i) = d(i) - \boldsymbol{w}(i-1)^{\mathrm{T}} \varphi(\boldsymbol{x}(i)), \tag{7}$$

$$\begin{aligned} \boldsymbol{w}(i) &= \boldsymbol{w}(i-1) + \eta e(i) \varphi(\boldsymbol{x}(i)) \\ &= \boldsymbol{w}(i-2) + \eta e(i-1) \varphi(\boldsymbol{x}(i-1)) + \eta e(i) \varphi(\boldsymbol{x}(i)) \\ &= \cdots\cdots \\ &= \boldsymbol{w}(0) + \mathbf{\Sigma}_{k=0}^{i} \eta e(i-k) \varphi(\boldsymbol{x}(i-k)). \end{aligned} \tag{8}$$

The transmitted training data signal at the i-th time sampling point is $d(i)$, and $x(i)$ is the received training signal vector

at the receiver side, which can be used to estimate $d(i)$. We denote the equalizer coefficient vector as w. η is the step-size factor for algorithm convergence. $e(i)$ is the error signal at the i-th iteration and $w(0)$ could be set to 0. According to (8), after the i-th training sample, the weight estimation is expressed as a linear combination of all the previous and present transformed inputs, weighted by the prediction errors (and scaled by η). When the estimated weights are used to compensate a new set of received vector x', the recovered signal is calculated by:

$$\begin{aligned}\boldsymbol{w}(i)\varphi(\boldsymbol{x}') &= \eta\boldsymbol{\Sigma}_{k=1}^{i}e(k)\varphi(\boldsymbol{x}(k))^{\mathrm{T}}\varphi(\boldsymbol{x}') \\ &= \eta\boldsymbol{\Sigma}_{k=1}^{i}e(k)\kappa(\boldsymbol{x}(k),\boldsymbol{x}').\end{aligned} \tag{9}$$

It is interesting to find the weight w absent in the equalized signals that are influenced by the nonlinear impairments. Instead, the sum of all past errors is multiplied by the kernel mapping on the previously received data (i.e., training data). It proves that the kernel mapping is a process without referring to the equalization coefficient w and the high-dimensional function φ. The KRLS also follows the same methodology as the aforementioned KLMS by transforming x to $\varphi(x)$ but has the target to minimize the recursive square errors. Therefore, the KRLS has much higher complexity than the KLMS. In this paper, we extend the kernel mapping by combining affine projection algorithm that reduces the gradient noise by multiple consecutive samples, which is referred to as kernel affine projection (KAP).

B. Affine Projection

Here, we present the operational principle of affine projection (AP) algorithm [19]. The training data set is $\{d(i), x(i)\}_{i=1,2,\ldots N}$, where N is the training data size. We denote the equalizer coefficient vector as w. The objective function of the equalizer is:

$$min_{\mathbf{w}}J(\boldsymbol{w}) = E|d-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}|^2. \tag{10}$$

The Wiener solution for (10) is:

$$\boldsymbol{w} = \boldsymbol{R}_{\mathbf{x}}^{-1}\boldsymbol{R}_{dx}, \tag{11}$$

where $R_{\mathbf{x}}$ is the positive-definite covariance matrix of $\mathbf{x}$ and R_{dx} is the cross-covariance vector of d and x. The LMS aims at minimizing the instantaneous squared estimation error, while the RLS minimizes the sum of squared estimation errors collected at the previous time points and the current time point. As a result, the LMS has the linear computational complexity and the RLS has the quadratic computational complexity with respect to N. Even after combining with the kernel mapping, such complexity properties remain for the KLMS and the KRLS. There are several methods reported to estimate w iteratively. Here, the gradient descent method [18], [19] and stochastic Newton's recursion method [18], [19] are first introduced. The update equation of w at the i-th iteration of gradient descent method is:

$$\boldsymbol{w}(i) = \boldsymbol{w}(i-1) + \eta[\boldsymbol{R}_{dx} - \boldsymbol{R}_{\mathbf{x}}\boldsymbol{w}(i-1)], \tag{12}$$

where η is the step-size factor for algorithm convergence. For stochastic Newton's recursion method to increase the convergence speed, the update equation of w is:

$$\boldsymbol{w}(i) = \boldsymbol{w}(i-1) + \eta(\boldsymbol{R}_{\mathbf{x}} + s\boldsymbol{I})^{-1}[\boldsymbol{R}_{dx} - \boldsymbol{R}_{\mathbf{x}}\boldsymbol{w}(i-1)], \tag{13}$$

where s is the smoothing factor. In the AP algorithm, $R_{\mathbf{x}}$ and $R_{\mathbf{dx}}$ are estimated by using the K most recent inputs and observations. Assuming:

$$\boldsymbol{x(i)} = [\boldsymbol{x}(i-K{+}1)\ldots\boldsymbol{x}(i)], \tag{14}$$

$$\boldsymbol{D}(i) = [d(i-K{+}1)\ldots d(i)]. \tag{15}$$

$R_{\mathbf{x}}$ and $R_{\mathbf{dx}}$can be estimated by:

$$\boldsymbol{R}_{\mathbf{x}} = K^{-1}\boldsymbol{X}(i)\boldsymbol{X}(i)^{\mathrm{T}}, \tag{16}$$

$$\boldsymbol{R}_{dx} = K^{-1}\boldsymbol{X}(i)\boldsymbol{d}(i). \tag{17}$$

Accordingly, the update equation of equalizer coefficient w in the AP using gradient descent method (referred to AP-G) is:

$$\boldsymbol{w}(i) = \boldsymbol{w}(i-1) + \eta\boldsymbol{X}(i)[\boldsymbol{D}(i) - \boldsymbol{X}(i)^{\mathrm{T}}\boldsymbol{w}(i-1)]. \tag{18}$$

The equalizer coefficient update equation of w in the AP using stochastic Newton's recursion method (referred to as AP-S) is:

$$\begin{aligned}\boldsymbol{w}(i) = \boldsymbol{w}(i-1) &+ \eta\boldsymbol{X}(i)[\boldsymbol{X}(i)^{\mathrm{T}}\boldsymbol{X}(i){+}s\boldsymbol{I}]^{-1} \\ &\times[\boldsymbol{D}(i) - \boldsymbol{X}(i)^{\mathrm{T}}\boldsymbol{w}(i-1)].\end{aligned} \tag{19}$$

Equitation (18) and (19) show how the equalizer coefficient is updated in the AP algorithm. It improves performance of the LMS algorithm by taking into accounting the K most recent inputs, while keeping linear operations and resulting in a low computational complexity.

C. Kernel Affine Projection

The KAP combines the kernel mapping with the AP algorithms. Its equalizer coefficient update equation shown in (18) and (19) needs to be modified by adding kernel mapping feature function $\varphi(\mathbf{x}(i))$. The objective function of equalizer is:

$$min_{\mathbf{w}}J(\boldsymbol{w}) = E|d-\boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{x})|^2. \tag{20}$$

The corresponding equalizer coefficient update equation at the i-th iteration of the KAP using gradient descent method (referred to KAP-G) is:

$$\begin{aligned}\boldsymbol{w}(i) &= \boldsymbol{w}(i-1) + \eta\Phi(\boldsymbol{x}(i))[\boldsymbol{D}(i) - \Phi(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{w}(i-1)] \\ &= \boldsymbol{w}(i-1) + \eta\Phi(\boldsymbol{x}(i))e(i),\end{aligned} \tag{21}$$

where $\Phi(\mathbf{x}(i)) = [\varphi(\mathbf{x}(\text{i-K}{+}1))\ \ldots\varphi(\mathbf{x}(i))]$ and $e(i)$ is the equalizer error at the i-th sampling point. (21) is expanded iteratively by using the same methodology as the KLMS and KRLS [14]:

$$\begin{aligned}\boldsymbol{w}(i) &= \boldsymbol{w}(i-1) + \eta\Phi(\boldsymbol{x}(i))e(i) \\ &= \boldsymbol{w}(i-2) + \eta\Phi(\boldsymbol{x}(i))e(i) + \eta\Phi(\boldsymbol{x}(i{-}1))e(i{-}1)\ldots \\ &= \boldsymbol{\Sigma}_{j=1}^{i}a_j(i)\Phi(\boldsymbol{x}(j)).\end{aligned} \tag{22}$$

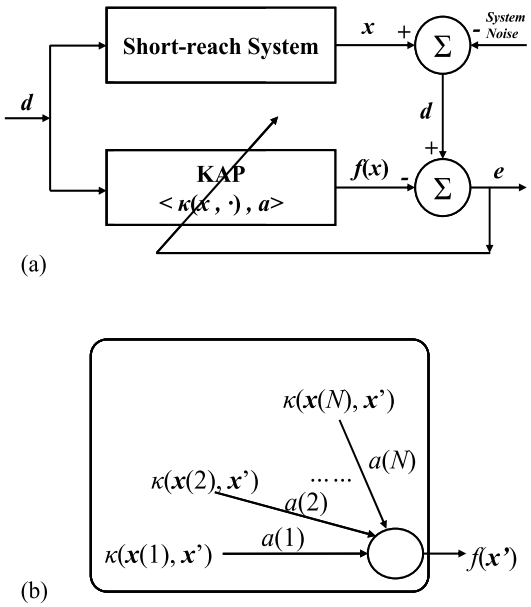

Fig. 1. (a) Schematic diagram of the KAP, (b) construction of the KAP output.

Algorithm 1 The signal processing flow of the KAP-G.

1. Set η, K and choose kernel function κ, $i = 1$;
2. $a_1(1)= \eta d(1)$;
3. while $i \leq N$ do
i. $a_i(i\text{-}1)=0$;
ii. for $k = \max(1, i\text{-}K + 1)$ to i do
a) $e_{K+k-i}(i) = d(k)\text{-}\ \mathbf{\Sigma}_{j=1}^{i-1}\{a_j(i\text{-}1)\kappa(x(k), x(j))\}$
b $a_k(i) = a_k(i\text{-}1)+\eta e_{K+k-i}(i)$
c) end for
iii. if $i>K$ do
a) for $k = 1$ to $i\text{-}K$ do
b) $a_k(i) = a_k(i\text{-}1)$
c) end for
iv. end if
v. $i = i+1$
vi. end while
4. Equalizer output of data signal x':
$f(x')= \mathbf{\Sigma}_{j=1}^{N}\{a(j)\kappa(x(j), x')\}$

The parameter $a_j(i)$ is the intermediate coefficient to get $w(i)$ and determines the final equalizer output. The relevant update function is shown in Algorithm 1.

It could be found that the convergence of the equalizer only depends on the kernel mapping feature function matrix $\Phi(x(i))$ and error signal e. By multiplication of $w(i)$ and received signal vector $\varphi(x')$, kernel function $\kappa(x, x')$ is utilized for equalization.

The signal processing flow of the KAP-G is shown in Algorithm 1 and the schematic diagram of the KAP is shown in Fig. 1. In Fig. 1(a), the transmitted data is d, after passing through the short-reach system, training vector x and testing vector x' are obtained. The training vector is used to train the coefficient a by calculating the error vector e. In Fig. 1(b), the trained coefficient a is used to equalize the testing vector x' by $f(x')= \mathbf{\Sigma}_{j=1}^{N}\{a(j)\kappa(x(j), x')\}$.

Regarding the KAP using stochastic Newton's recursion method (KAP-S), the update formula in (21) is transformed to (23). Compared with the KAP-G, the signal processing routine of the KAP-S is similar except normalizing the error by $[\Phi(x(i))^{\mathrm{T}}\Phi(x(i))+sI]^{-1}$ and thus

$$\begin{aligned}\boldsymbol{w}(i) &= \boldsymbol{w}(i-1)\\ &\quad +\eta\Phi(\boldsymbol{x}(i))[\Phi(\boldsymbol{x}(i))^{\mathrm{T}}\Phi(\boldsymbol{x}(i)) + s\boldsymbol{I}]^{-1}e(i). \quad (23)\end{aligned}$$

D. Error Reuse Criterion

The most complicated calculation in the KAP is error calculation, which consists of $(i\text{-}1)*K$ kernel calculations to compute $w(i)$, making its number of multiplications K times of the KLMS. By 'error reuse' criterion shown in (24), the number of multiplications could be squeezed.

$$\begin{aligned}\underline{e(K+k-i)} &= d(k) - \varphi(\boldsymbol{x}(k))^{\mathrm{T}}\boldsymbol{w}(i-1)\\ &= d(k) - \varphi(\boldsymbol{x}(k))^{\mathrm{T}}[\boldsymbol{w}(i-2)\\ &\quad +\eta\mathbf{\Sigma}_{j=i-K}^{i-1}\{e(K+j-(i-1))\varphi(\boldsymbol{x}(j))\}]\\ &= [d(k) - \varphi(\boldsymbol{x}(k))^{\mathrm{T}}\boldsymbol{w}(i-2)]\\ &\quad +\eta\mathbf{\Sigma}_{j=i-K}^{i-1}\{e(K+j-(i-1))\kappa(\boldsymbol{x}(k), \boldsymbol{x}(j))\}\\ &= \underline{e(K+k-(i-1))}\\ &\quad +\eta\mathbf{\Sigma}_{j=i-K}^{i-1}\{e(K+j-i+1)\kappa(\boldsymbol{x}(k), \boldsymbol{x}(j))\}.\end{aligned} \quad (24)$$

As a result, the number of kernel calculations is reduced from $(i\text{-}1)*K$ to $(i\text{-}1)$, and the computational complexity in terms of number of multiplications will be discussed in Section. IV.

E. Regularized KAP

To improve the convergence speed of convex optimization in the KAP, a regularization process is introduced in this paper. Here the regularized Newton's recursion method is used and the regularized KAP is referred to as KAP-R. The objective function of the KAP-R is updated as:

$$min_{\mathbf{w}} J(\boldsymbol{w}) = E|d - \boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{x})|^2 + \lambda||\boldsymbol{w}||^2, \quad (25)$$

where λ is the regularization parameter and its value is smaller than 1. In the KAP-R, the update formula in (21) and (23) is transformed to (26).

$$\begin{aligned}\boldsymbol{w}(i) &= (1-\eta)\boldsymbol{w}(i-1) + \eta\boldsymbol{\Phi}(\boldsymbol{x}(i))[\Phi(\boldsymbol{x}(i))^{\mathrm{T}}\Phi(\boldsymbol{x}(i))\\ &\quad +\lambda\boldsymbol{I}]^{-1}\boldsymbol{D}(i).\end{aligned} \quad (26)$$

Compared with the KAP-G and KAP-S, the KAP-R uses a scaling factor $1\text{-}\eta$ to multiply the previous weight, and imposes a forgetting mechanism so that the training data from the far past is not needed and the total amount of training data can be scaled down exponentially. Therefore, the data from the far past does not influence the equalizer output and the training of the coefficient is more accurate.

Considering the complexity, the KAP-R updates the weights without the error signal e, and uses matrix $\Phi(x(i))^{\mathrm{T}}\Phi(x(i))+\lambda I$

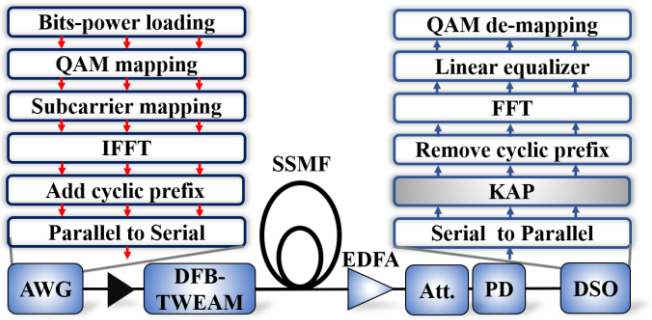

Fig. 2. Experimental setup.

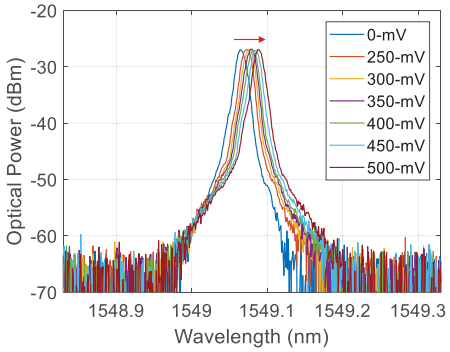

Fig. 3. Optical spectrum of TWEAM with different modulation amplitudes.

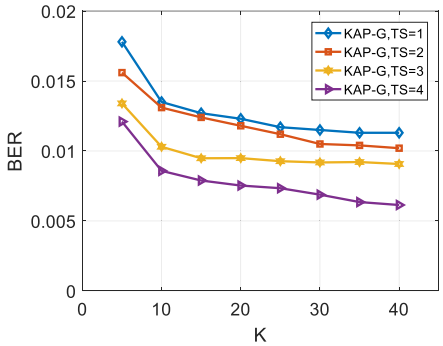

Fig. 4. BER performance of KAP-G versus K and *TS*.

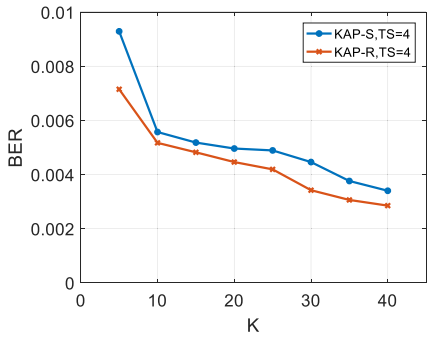

Fig. 5. BER performance of the KAP-S and KAP-R versus K.

instead. Using the sliding-window trick defined in [20], the complexity of KAP-R can be reduced to K^2, showing another implementation benefit.

III. Experimental Setup

The experimental setup and the signal processing flow scheme demonstrating the feasibility of the KAP in optical short-reach system are shown in Fig. 2. The DMT samples are generated offline and loaded into a 120-GSa/s DAC (Keysight M8194A). The signal from the DAC is amplified by an electrical amplifier with 11-dB gain before applying on the EML for modulation. The EML is composed of a monolithically integrated distributed feedback (DFB) laser with a travelling-wave electro absorption modulator (TWEAM) [21], [22]. The PD is a high-speed InP packaged prototype photo-detector with a responsibility of 0.5-A/W. The electrical signal from direct-detection is captured by one 256-GSa/s digital storage oscilloscope (DSO, Keysight UXR1102A). In the experiment, the length of inverse fast Fourier transform (IFFT) for DMT modulation is set to 2048, which corresponds to 58.6-MHz subcarrier spacing. The clipping ratio of the DMT signal is set to 0.75 to improve the signal to noise ratio (SNR) out of the DAC. The output amplitude of the DAC is set to 700-mV. The modulated optical signal from the EML ($\sim$0.5-dBm) is fed into a spool of 400-m standard single mode fiber (SSMF). Because of the lack of trans-impedance amplifier (TIA), one erbium doped fiber amplifier (EDFA) together with a variable optical attenuator (Att.) is used before PD to study the receiver sensitivity.

The optical spectrum of DFB-TWEAM with different modulation amplitudes (peak-to-peak voltage) is shown in Fig. 3. It displays a red shift with respect to a larger electrical signal amplitude due to the thermal regime variation caused by microwave absorption of impedance matching resistors [22]. This implies that memory nonlinearities are introduced when the DMT signal with a large swing is used for modulation. Besides the modulation nonlinearity of the EML, the nonlinearity of the short-reach DMT fiber communication system also comes from other sources, such as inter-subcarrier mixing in square-law detection, clipping of the signal and the nonlinearity from electrical amplifiers due to the high peak-to-average power ratio (PAPR) of the DMT signal [9].

IV. Experimental Results and Discussions

The BER performance in terms of K and training symbols (TS) in the KAP-G is shown in Fig. 4 and the BER performance in terms of K in the KAP-S and KAP-R is shown in Fig. 5. Here, the KAP is employed on top of the linear equalizer [24]. The received optical power (RoP) is 8-dBm. The size of length of each DMT training symbol is 2048 DMT samples (i.e., TS=1 corresponds to $N = 2048$), which corresponds to the IFFT size. In the experiment, there are 80 DMT symbols in total. The increase of K in the KAP improves the BER gradually. In Fig. 4, the same trend can be observed from the curves for the KAP-G with different numbers of training data size. In the experiments, the trend is the similar for the KAP-S and KAP-G, and their BER performance is only shown with TS=4 in Fig. 5. The BER improvement tends to get statured when K increases. With the

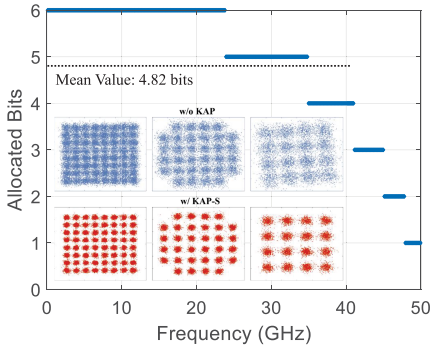

Fig. 6. Adaptive bits allocation at SSMF transmission and recovered constellation graphs with and without the KAP-S equalizer.

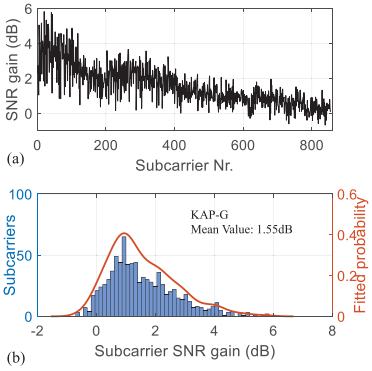

Fig. 7. (a) Subcarrier SNR gain of the KAP-G, (b) probability distribution of subcarrier SNR gains.

normalization factor, the KAP-S achieves better transmission performance compared with the KAP-G at different K values, and the KAP-R achieves better performance compared with the KAP-S because of its regularization operation. N of 8192 (TS=4) and K of 40 are selected for the experiment results here. It is worth noticing that periodical training could save the number of TSs. For example, carrying out training 4 times each using 1 TS is able to achieve the similar performance as one-time training but with 4 TSs. For latency-sensitive cases, the trade-off between the training time and the number of training symbols should be carefully taken into account.

The adaptive bits loading profile [23] at SSMF short-reach transmission case is shown in Fig. 6. The QAM orders ranging from BPSK to 64QAM are adopted in the system. The mean allocated bits is 4.82. The constellation graphs of 16QAM, 32QAM and 64QAM with and without the KAP-S are shown in the insets of Fig. 6. For minimizing linear impairments, the linear equalizer presented in [24] is used regardless of whether the KAP-S is implemented. It can be seen that the SNR has been obviously improved and the constellation points are clearer. It benefits nonlinear impairments compensation from the KAP-S since the nonlinear impairments makes the constellations scattered and the Euclidean distance between constellation points are reduced.

The SNR gain at different subcarriers of the KAP-G, KAP-S and KAP-R and their probability distributions are shown in Fig. 7, Fig. 8 and Fig. 9, respectively. The SNR gain is defined as the difference between the SNR of each subcarrier after and before the equalizer. There are SNR gains at most subcarriers in DMT signal. The SNR gain is higher at low-frequency range and lower at the high-frequency range for the KAP-G, KAP-S and KAP-R, which is caused by the low-pass filtering effect of the system response. The maximum subcarrier SNR gain of the KAP-G and KAP-S are around 6-dB, and the average SNR gain of the KAP-G, KAP-S and KAP-R are 1.55-dB, 2.26-dB and 2.68-dB respectively. The results are also in line with the observations in Fig. 4 and Fig. 5, where the KAP-R shows the best transmission performance.

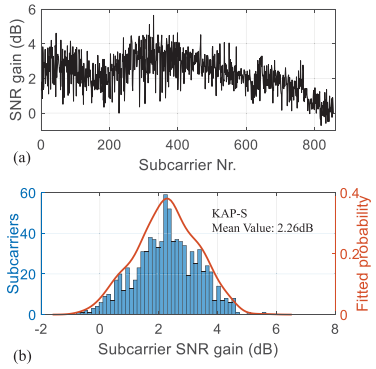

Fig. 8. (a) Subcarrier SNR gain of the KAP-S, (b) probability distribution of subcarrier SNR gains.

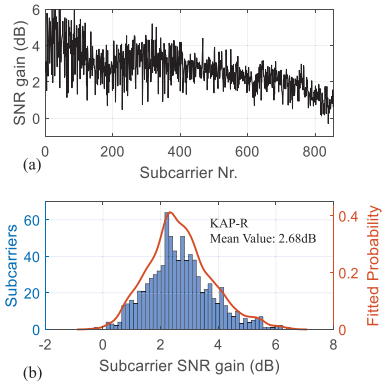

Fig. 9. (a) Subcarrier SNR gain of the KAP-R, (b) probability distribution of subcarrier SNR gains.

TABLE I
THE BER WITH DIFFERENT EQUALIZERS

Equalizers	*BER*
KLMS	7.10e-3
KAP-G	**6.14e-3**
Volterra(2nd)	6.90e-3
Volterra(3rd)	5.87e-3
CI-BCH FEC Limit	*4.52e-3*
KRLS	3.62e-3
KAP-S	**3.40e-3**
Volterra(4th)	3.24e-3

TABLE II
THE EQUALIZER COEFFICIENT
*w*UPDATE EQUATION AND NUMBER OF MULTIPLICATIONS OF DIFFERENT NONLINEAR EQUALIZATION METHODS

Methods	*Equalizer coefficient update equation*	*Number of multiplications*	
		Formula	*The tested case* (N=8096 *and* K=40)
Volterra	$\boldsymbol{w}(i)=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\varphi}(\boldsymbol{x}(i))[d(i)-\boldsymbol{\varphi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	N_L	8096
KLMS[14,30]	$\boldsymbol{w}(i)=[\eta\boldsymbol{I}+\boldsymbol{\Phi}(\boldsymbol{x}(i))\boldsymbol{\Phi}(\boldsymbol{x}(i))^{\mathrm{T}}]^{-1}\boldsymbol{\Phi}(\boldsymbol{x}(i))\boldsymbol{D}(i)$	N_L^2	65545216
KRLS[14,31]	$\boldsymbol{w}(i)=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{x}(i))[\boldsymbol{D}(i)\text{-}\boldsymbol{\Phi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	N_L+K^2	9696
KAP-G	$\boldsymbol{w}(i)=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{x}(i))[\boldsymbol{\Phi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x}(i))+s\boldsymbol{I}]^{-1}*[\boldsymbol{D}(i)\text{-}\boldsymbol{\Phi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	N_L+K^2	9696
KAP-S	$\boldsymbol{w}(i)=(1\text{-}\eta)\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{x}(i))[\boldsymbol{\Phi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x}(i))+\lambda\boldsymbol{I}]^{-1}\boldsymbol{D}(i)$	K^2	1600
KAP-R	$\boldsymbol{w}(i)=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\varphi}(\boldsymbol{x}(i))[d(i)-\boldsymbol{\varphi}(\boldsymbol{x}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	N_L	8096

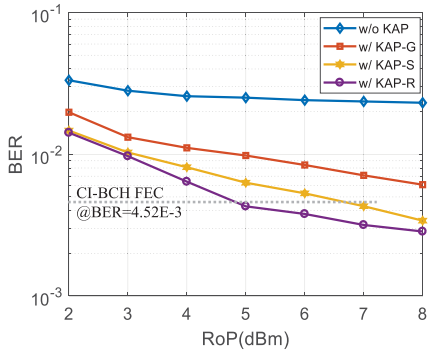

Fig. 10. System transmission performance.

The BER versus the RoP is shown in Fig. 10. We consider continuously-interleaved Bose–Chaudhuri–Hocquenghem FEC (CI-BCH (1020, 988)) [25], [26] with a BER-limit of 4.52E-3. The CI-BCH FEC BER-limit could be achieved only when the KAP is employed on top of the linear equalizer [24], and the KSP-R achieves the best performance compared with the KAP-G and KAP-S. Therefore, it demonstrates capability of the KAP to mitigate nonlinear impairments in optical short-reach systems. The maximum system line rate can achieve is 238-Gbps with the net rate of 222-Gbps.

The BER for different equalizers measured at the RoP of 8-dBm is compared in Table I. The Volterra filtering with a memory length of 11 is included for comparison purpose. All the compared equalizers are used on top of the linear equalizer [24]. The KAP-G outperforms the KLMS, and the KAP-S algorithm achieves similar performance as the KRLS and the Volterra filtering algorithm with nonlinear kernels up to the 4th order. The KAP-S obviously outperforms the KLMS and Volterra filtering algorithm with nonlinear kernels up to the 2nd/3rd order. The KAP-R achieves the best performance since it has retained the advantages of KAP-S while minimizing the negative effect of the far past samples on the decision sample.

A. Complexity Analysis

Considering the storage complexity, the KAP obtains coefficients just like the KLMS, and it avoids the least-square problem in the KRLS. The KAP has linear storage complexity in terms of the number of iterations (i.e. the length of training symbols, N, i.e. $\mathrm{O}(N)$ [18], [27].

The number of multiplications of the kernel methods mainly depends on the size of the reproducing kernel Hilbert space. A factor that influences this size is the number of training symbols N. For the KAP-G and KAP-S, their number of multiplications is equal to $N + K^2$, which shares the same order with the KLMS [14], [18], [27]. Comparatively, the number of multiplications of the Volterra series equalizer is $\sum_{p=1}^{P}(M_p)^p$, where M_p is the tap number (i.e. memory length) of the p-th order nonlinear kernels. Table II shows the number of multiplications of different algorithms. The N of 8192 and K of 40 are selected to compare the kernel methods with the Volterra filtering. Since the KAP-R does not need the calculation of the error signal e, and uses matrix $\Phi(x(i))^{\mathrm{T}}\Phi(x(i))+\lambda I$ for weight update. The required number of multiplications is reduced to K^2.

It is worth noticing that the sparsification techniques could be adopted for reducing the length of the training symbols, such as the novelty criterion (NC), the coherence criterion, the quantization criterion and surprise criterion [28]. In [29], the number of multiplications of the KLMS has decreased from 5000 to 150 after NC sparsification. Therefore, sparsification techniques could be an interesting tool to simplify the KAP for channel equalization in short-reach optical links. There are some research works on simplified Volterra filtering, such as reduced Volterra kernels [32], sparsifications [33], regularization [34], frequency transformation [35]. Moreover, the KAP provides a flexible and general way to transform to the other kernel mapping based equalization algorithms. For instance, the KAP-G transforms to the KLMS when K is equals to 1, which can be observed from Table II. When K is equal to N, the KAP-S transforms to regularization network [18], which shows a potential of implementing the KAP for machine learning based tasks.

V. CONCLUSION

The proposed kernel affine projection algorithms is efficient for compensating nonlinear impairments in optical short-reach systems. All the variants of the KAP algorithms have a

TABLE III
ABBREVIATIONS IN THE PAPER

Abbreviations	*Full Name*
AP	Affine Projection
BER	Bit Error Ratio
CI-BCH	Continuously-Interleaved Bose–Chaudhuri–Hocquenghem
DAC	Digital-to-Analogue Converter
DFB-TWEAM	Distributed Feedback Laser with a Travelling-Wave Electro Absorption Modulator
DMT	Discrete Multi-Tone
DSO	Digital Storage Oscilloscope
EDFA	Erbium Doped Fiber Amplifier
EML	Externally Modulated Laser
FEC	Forward Error Correction
IFFT	Inverse Fast Fourier Transform
IM/DD	Intensity Modulation Direct Detection
KAP	Kernel Affine Projection
KAP-G	Kernel Affine Projection using Gradient descent method
KAP-R	Kernel Affine Projection using Regularized Newton's recursion method
KAP-S	Kernel Affine Projection using Stochastic Newton's recursion method
KLMS	Kernel Least-Mean-Square
KRLS	Kernel Recursive-Least-Squares
LMS	Least-Mean-Square
PAPR	Peak-to-Average Power Ratio
PD	Photo-Detector
RLS	Recursive-Least-Squares
RoP	Received optical Power
SNR	Signal to Noise Ratio
SSMF	Standard Single Mode Fiber
TIA	Trans-Impedance Amplifier
TS	Training Symbols

comparable number of multiplications with the KLMS and are able to achieve comparable performance with the KRLS, having the merits of both algorithms while avoiding their drawbacks. We have experimentally demonstrated a 222-Gbps net rate IM/DD DMT system, where the KAP algorithms greatly improve the transmission performance.

APPENDIX

See Table III.

ACKNOWLEDGMENT

Keysight Technologies is acknowledged for the loan of DSO (Keysight UXR1102A) and AWG (Keysight M8194A). The authors would like to thank Hadrien Louchet, Thomas Dippon, and Markus Gruen from Keysight Technologies GmbH for the help with experimental setup.

REFERENCES

[1] *Cisco Global Cloud Index: Forecast and Methodology, 2016–2021*. Accessed: Oct. 1, 2019. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html
[2] *Ethernet Alliance, The 2020 Ethernet Roadmap*. Accessed: Oct. 8, 2019. [Online]. Available: https://ethernetalliance.org/technology/2020-roadmap/
[3] X. Zhou, R. Urata, and H. Liu, "Beyond 1Tb/s datacenter interconnect technology: Challenges and solutions," in *Proc. OFC*, San Diego, CA, USA, Mar. 2019, pp. 1–3.
[4] X. Pang *et al.*, "Beyond 200 gbps per lane intensity modulation direct detection (IM/DD) transmissions for optical interconnects: Challenges and recent developments," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Diego, Mar. 2019, pp. W4I–7.
[5] K. Zhong, X. Zhou, J. Huo, C. Yu, C. Lu, and A. P. T. Lau, "Digital signal processing for short-reach optical communications: A review of current technologies and future trends," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 377–400, Jan. 15, 2018.
[6] C.-C. Wei, H.-L. Cheng, and W.-X. Huang, "On adiabatic chirp and compensation for nonlinear distortion in DML-based OFDM transmission," *J. Lightw. Technol.*, vol. 36, no. 16, pp. 3502–3513, Aug. 15, 2018.
[7] H. Yang *et al.*, "Evaluation of effects of MZM nonlinearity on QAM and OFDM signals in RoF transmitter," in *Proc. Int. Topical Meeting Microw. Photon. Jointly Held Asia–Pacific Microw. Photon. Conf.*, Gold Coast, QLD, Canada, Sep. 2008, pp. 90–93.
[8] C.-C. Wei *et al.*, "Analysis of nonlinear distortion and SSII cancellation in EAM-based IMDD OFDM transmission," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 3069–3082, Jul. 15, 2015.
[9] L. Zhang *et al.*, "Nonlinearity-aware 200-Gbit/s discrete multi-tone transmission for C-band short-reach optical interconnects with a single packaged EML," *Opt. Lett.*, vol. 43, no. 2, pp. 182–185, Jan. 2018.
[10] H. Yamazaki *et al.*, "Transmission of 400-Gbps discrete multi-tone signal using $>$ 100-GHz-bandwidth analog multiplexer and InP Mach-Zehnder modulator," in *Proc. ECOC*, Rome, Italy, Sep. 2018, pp. 1–3.
[11] G. Chen *et al.*, "Nonlinear distortion mitigation by machine learning of SVM classification for PAM-4 and PAM-8 modulated optical interconnection," *J. Lightw. Technol.*, vol. 36, no. 3, pp. 650–657, Feb. 1, 2018.
[12] L. Sun, J. Du, and Z. He, "Machine learning for nonlinearity mitigation in CAP modulated optical interconnect system by using K-nearest neighbour algorithm," in *Proc. Asia Commun. Photon. Conf. (ACP)*, Wuhan, China, Nov. 2016.
[13] E. Novak and K. Ritter, "The curse of dimension and a universal method for numerical integration," in *Multivariate Approximation Splines*. Cham, Switzerland: Springer, 1997, pp. 177–187.
[14] L. Zhang *et al.*, "Kernel mapping for mitigating nonlinear impairments in optical short-reach communications," *Opt. Express*, vol. 27, no. 21, pp. 29567–29580, Oct. 2019.
[15] L. Zhang *et al.*, "Kernel adaptive filtering for nonlinearity-tolerant optical direct detection systems," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Rome, Italy, Sep. 2018, pp. 1–3.
[16] L. Zhang *et al.*, "Experimental demonstration of 503.61-Gbit/s DMT over 10-km 7-core fiber with 1.5-μm SM-VCSEL for optical interconnects," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Rome, Italy, Sep. 2018.
[17] L. Zhang *et al.*, "Kernel affine projection for compensating nonlinear impairments in optical direct detection systems," in *Proc. 45th Eur. Conf. Opt. Commun. (ECOC)*, Dublin, Sep. 2019, pp. 1–3.
[18] W. Liu, C. P. Jose, and H. Simon, *Kernel Adaptive Filtering: A Comprehensive Introduction*, vol. 57. Hoboken, NJ, USA: Wiley, 2011.
[19] K. Ozeki, *Theory of Affine Projection Algorithms for Adaptive Filtering*. Cham, Switzerland: Springer, 2016.
[20] S. Van Vaerenbergh, J. Via, and I. Santamaría, "A sliding-window kernel RLS algorithm and its application to nonlinear channel identification," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2006, pp. 789–792.
[21] O. Ozolins *et al.*, "100 GHz externally modulated laser for optical interconnects," *J. Lightw. Technol.*, vol. 35, no. 6, pp. 1174–1179, Mar. 15, 2017.
[22] M. Chacinski *et al.*, "Modulation and chirp evaluation of 100 GHz DFB-TWEAM," in *Proc. 36th Eur. Conf. Exhib. Opt. Commun.*, Sep. 2010, pp. 1–3.
[23] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 773–775, Feb. 1995.

[24] L. Zhang, S. Xiao, M. Bi, L. Liu, and Z. Zhou, "Channel estimation algorithm for interference suppression in IMDD–OQAM–OFDM transmission systems," *Opt. Commun.*, vol. 364, pp. 129–133, Apr. 2016.
[25] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," in *Proc. IEEE Photon. Conf. (IPC)*, Sep. 2018, pp. 1–5.
[26] M. Scholten, T. Coe, and J. Dillard, "Continuously-interleaved BCH (CI-BCH) FEC delivers best in class NECG for 40G and 100G metro applications," in *Proc. Nat. Fiber Optic Eng. Conf.*, Mar. 2010, pp. 1–3.
[27] S. Van Vaerenbergh and I. Santamaría, "A comparative study of kernel adaptive filtering algorithms," in *Proc. IEEE Digit. Signal Process. Signal Process. Educ. Meeting*, Aug. 2013, pp. 181–186.
[28] J. Platt, "A resource-allocating network for function interpolation," *Neural Comput.*, vol. 3, no. 2, pp. 213–225, 1991.
[29] U. Singh, R. Mitra, V. Bhatia, and A. Mishra, "Kernel LMS-based estimation techniques for radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2501–2515, Oct. 2019.
[30] W. Liu, P. P. Pokharel, and J. C. Principe, "The kernel least-mean-square algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 543–554, Feb. 2008.
[31] W. Liu, I. Park, Y. Wang, and J. C. Principe, "Extended kernel recursive least squares algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3801–3814, Oct. 2009.
[32] Y. Xu *et al.*, "260-Gb/s PAM-6 transmission using joint optical pre-equalization and a low-complexity volterra equalizer for short-reach optical interconnects," in *Proc. Asia Commun. Photon. Conf. (ACP)*, Oct. 2018, pp. 1–3.
[33] C.-Y. Chuang *et al.*, "Sparse volterra nonlinear equalizer by employing pruning algorithm for high-speed PAM-4 850-nm VCSEL optical interconnect," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, San Diego, CA, USA, Mar. 2019, p. M1F-2.
[34] W. Huang *et al.*, "93% complexity reduction of volterra nonlinear equalizer by ℓ_1-regularization for 112-Gbps PAM-4 850-nm VCSEL optical interconnect," in *Proc. OFC*, San Diego, CA, USA, Mar. 2018, pp. 1–3.
[35] J. Zhang *et al.*, "Low complexity frequency-domain nonlinear equalization for 40-Gb/s/wavelength long-reach PON," in *Proc. OFC*, San Diego, CA, USA, Mar. 2018, pp. 1–3.

Lu Zhang (Member, IEEE) received the bachelor's degree from Southeast University in 2014 and the Ph.D. degree from Shanghai Jiao Tong University in 2019. From 2016 to 2017, he was a Visiting Ph.D. Student at the KTH Royal Institute of Technology sponsored by the China Scholarship Council (CSC). Since 2018, he has been a Visiting Research Engineer with the KTH Royal Institute of Technology and the Kista High-Speed Transmission Lab, RISE AB. He is currently a Research Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University. His research interests include ultrafast THz communications, fiber-optic communications, and digital signal processing algorithms for optical and THz transmission systems.

Jiajia Chen (Senior Member, IEEE) received the Ph.D. degree from the KTH Royal Institute of Technology, Sweden, in 2009. She is currently a Professor with the Chalmers University of Technology, Sweden. Her research interests include optical network architecture design along with the supporting transmission techniques and resource allocation strategies, which contribute to improving capacity, reliability, energy efficiency and cost efficiency in access, core, and data center networks. She has been sharing her knowledge with the technical community with many peer-reviewed journal articles, conference papers, invited talks (more than 200 publications to-date), and 10 patent applications. She has organizing several IEEE international conferences in the field and serving on technical program committees, including the OFC (Subcommittee Chair 2019), ECOC, GLOBECOM (Symposium Co-Chair 2020), ICC, and so on.

Aleksejs Udalcovs (Member, IEEE) received the M.Sc. degree in telecommunications and the Dr.Sc.Ing. (Ph.D.) degree in electronics and telecommunications from Riga Technical University in 2011 and 2015, respectively. From 2012 to 2016, he worked at the KTH Royal Institute of Technology as a Visiting Researcher within the Swedish Institute Visby Program and a Post-Doctoral Researcher within the EU Project GRIFFON, Stockholm, Sweden. In 2015, he joined the Kista High-Speed Transmission Lab (HST-Lab) jointly operated by the KTH and Research Institutes of Sweden (RISE). In 2016, he moved to RISE after receiving a research grant from the Swedish ICT the Next Generation Consortium. Since 2019, he has been a Senior Scientist with RISE who is working in cooperation with industry and academia and providing his expertise in communication technologies. He is a (co)author of more than 100 papers in peer-reviewed international journals and conferences. His main research interests include the PHY-layer aspects in optical transport and photonic-wireless networks.

Xiaodan Pang (Senior Member, IEEE) received the Ph.D. degree from the DTU Fotonik, Technical University of Denmark, in 2013. He was a Post-Doctoral Researcher at the Research Institutes of Sweden (RISE) (former ACREO Swedish ICT) from October 2013 to March 2017 and then worked as a Researcher at the KTH Optical Networks Lab (ONLab) from March 2017 to February 2018. From March 2018 to February 2020, he worked as a Staff Opto Engineer and a Marie Curie Research Fellow at the Infinera Global HW Research Development Team, Stockholm. He has been working as a Senior Researcher at the Department of Applied Physics, KTH Royal Institute of Technology, Sweden, since March 2020. He is/has been the PI of the Swedish Research Council Starting Grant, the EU H2020 Marie Curie Individual Fellowship Project NEWMAN, and the Swedish SRA ICT-TNG Post-Doctoral Project. His research focuses on ultrafast communications with millimeter-wave, terahertz wave, free-space optics, and fiber-optics. He authored or coauthored over 170 publications in journals and conferences. He has been a TPC Member of in total 16 conferences, including OFC'20-21, ACP'18-20, and GLOBECOM'20. He is a Board Member of the IEEE Photonics Society Sweden Chapter.

Richard Schatz was born in 1963. He received the Ph.D. degree in photonics from the Laboratory of Photonics and Microwave Engineering, KTH Royal Institute of Technology, Stockholm, Sweden, in 1995. Since 1987, he has been conducting researches as a Senior Researcher and a Lecturer at the KTH Royal Institute of Technology. From 1992 to 1993, he was a Visiting Scientist with AT&T Bell Laboratories, Murray Hill, NJ, USA. He is currently a part of the Kista HST-Lab, Stockholm. He has developed laser simulation software for the photonics industry and authored or coauthored more than 150 journal articles and conference contributions. His research interests include modeling, design, and characterization of fiber-optical transmitters (edge emitter lasers, VCSELs, and modulators) and links, both for ON–OFF keying, and more advanced modulation formats.

Urban Westergren received the M.Sc., Ph.D., and Docent degrees from the KTH Royal Institute of Technology, Stockholm, Sweden, in 1984, 1992, and 2002, respectively. From 1984 to 1993, he was a Research Engineer in high-speed electronics with the Swedish Institute of Microelectronics, Stockholm. In 1994, he joined the Laboratory of Photonics and Microwave Engineering, KTH, Beijing, China, as a Research Associate. In 1996, he became an Associate Professor and in 2012, he was appointed to a Full Professor of optoelectronic integrated circuits at KTH. From 2000 to 2003, he was a Senior Expert in electronic design with Optillion Company. He has been a Teacher and an Examiner of more than 1000 bachelor's and master's level students and more than 100 Ph.D. students of electromagnetics, microwave engineering, photonics, and high-speed electronics. Since 2012, he has been the Deputy Director of international affairs KTH, regarding education. His research at KTH has included design and measurement of integrated receivers and transmitters for fiber optical communication at very high bitrates. In the later years, his research focused on design of electroabsorption modulators. He was an important applicant and the Coordinator of the EU Project HECTO from 2006 to 2010. HECTO was concluded with world-record field trials of a complete fiber optical communication system with 112 Gbit/s serial transmission using ON–OFF Keying over 42 km. He is the author or coauthor of more than 100 international publications, conference contributions, and book chapters. He holds two patents.

Sergei Popov (Senior Member, IEEE) received the M.Sc. degrees in applied physics and computer science in Russia in 1987 and 1989, respectively, and the Ph.D. degree in applied physics in Finland in 1999. He is currently a Professor with the Applied Physics Department, Royal Institute of Technology (KTH), Stockholm, Sweden. His expertise is in optical communications, laser physics, plasmonics, and optical materials. He has published over 300 articles and conference contribution. He stayed at Ericsson Telecom AB and Acreo AB, Sweden, before joining KTH. He is a fellow of OSA. He is the Editor-in-Chief of JEOS:RP Journal (EOS).

Shilin Xiao (Member, IEEE) received the M.S. degree from the University of Electronic Science and Technology of China in 1988 and the Ph.D. degree from Shanghai Jiao Tong University in 2003. From 1988 to 1999, he worked at the Guilin Institute of Optical Communications and obtained the technical titles as a Senior Engineer in 1995 and a Professor in 1999. Since 2000, he has been with the State Key Laboratory of Advanced Optical Communication Systems and Network, Shanghai Jiao Tong University. He is currently a Full Professor with Shanghai Jiao Tong University. His research interests include optical communications, especially optical switching, optical access networks, and components. He has published more than 200 papers in technical journals and conferences and has obtained more than 50 patents. He is a Senior Member of the Chinese Institute of Electronics and the Optical Society of China.

Oskars Ozolins (Member, IEEE) received the M.Sc. degree in telecommunications and the Dr.Sc.Ing. (Ph.D.) degree in optical communications from Riga Technical University, Riga, Latvia, in 2009 and 2013, respectively. He is currently a Senior Scientist and the Technical Lead with the Kista High-Speed Transmission Lab (Kista HST-Lab), RISE Research Institutes of Sweden (former ACREO) under the Swedish Research Council starting grant project Photonic-Assisted Signal Processing Techniques (PHASE). He is also appointed as Affiliated Faculty on Optical Communication at the Department of Applied Physics, KTH Royal institute of Technology. His research interests include digital and photonic-assisted signal processing techniques, high-speed short-reach communications and devices, optical and photonic-wireless interconnects, and single photon quantum communications. He is the author of around 170 international journal publications, conference contributions, invited talks, patents, and book chapters. He has regularly served as a Designated Reviewer for IEEE/OSA Journals, such as *Chinese Optics Letters*, *Applied Optics*, *Optics Express*, IEEE PHOTONICS TECHNOLOGY LETTERS, *Journal of Lightwave Technology*, and *Photonics Research*.