KERNEL AFFINE PROJECTION FOR COMPENSATING NONLINEAR IMPAIRMENTS IN OPTICAL DIRECT DETECTION SYSTEMS

Lu Zhang[1,2], Jiajia Chen[1,5], Aleksejs Udalcovs[3], Xiaodan Pang[1,3,4], Richard Schatz[1], Urban Westergren[1], Sergei Popov[1], Shilin Xiao[2] and Oskars Ozolins[3*,1]*

[1]*KTH Royal Institute of Technology, Kista, Sweden*
[2]*Shanghai Jiao Tong University, Shanghai, China*
[3]*RISE Research Institutes of Sweden AB, Kista, Sweden*
[4]*Infinera, Stockholm, Sweden*
[5]*Chalmers University of Technology, Gothenburg, Sweden*
** jiajiac@chalmers.se; oskars.ozolins@ri.se*

Keywords: Optical fibre communication; optical direct detection systems; DMT modulation; digital signal processing; equalizer.

Abstract

We propose a novel kernel affine projection (KAP) algorithm, which combines kernel mapping and affine projection for channel equalization in optical direct detection systems. Experimental results show that the KAP algorithm can mitigate nonlinear impairments in short-reach communications while maintaining low computational complexity.

1 Introduction

Nonlinear impairments from opto-electronic devices have become a major performance-limiting factor in high-speed optical direct detection systems. Volterra filtering based schemes [1] and machine learning based schemes [2] have been proposed to suppress the nonlinear impairments in optical direct detection systems. However, these methods perform high-dimensional signal mapping, which causes the 'curse of dimension' [3] and hence normally result in high computational complexity.

Our previous works [4,5] for the first time introduced kernel mapping methods for channel equalization in optical short-reach communications, where 'Mercer kernel' is employed to map the signals from the Euclidean space to the reproducing kernel Hilbert space (RKHS). Such a 'kernel trick' [6] makes linear signal processing carried out in the high dimensional RKHS yield powerful extensions to compensate the nonlinear impairments. Two kernel mapping methods have been devised, namely kernel least-mean-square (KLMS) [4] and kernel recursive-least-squares (KRLS) [5]. On one hand, KLMS is with linear computational complexity in term of training data size, while the KRLS is with quadratic computational complexity. On the other hand, the KRLS outperforms the KLMS in terms of transmission performance.

In this paper, we propose another kernel mapping based method, referred to as kernel affine projection (KAP) for channel equalization in high-speed optical short-reach direct detection systems. The KAP can perform as good as the KRLS while maintains the computational complexity as low as that of the KLMS algorithm. We have experimentally demonstrated intensity modulation and direct detection (IM/DD) discrete multi-tone (DMT) system with 222 Gbps net rate, where the KAP improves the bit error ratio (BER) to satisfy the hard-decision forward error correction (HD-FEC) limit [6-7].

2 Kernel affine projection

We firstly present the operational principle of affine projection (AP) algorithm [8]. We assume the transmitted training data signal at the i-th time sampling point is $d(i)$, and $\boldsymbol{u}(i)$ is the received training signal vector at the receiver side used to estimate $d(i)$. The training data set is $\{d(i), \boldsymbol{u}(i)\}_{i=1,2,\dots N}$, where N is the training data size. We denote the equalizer coefficient vector as $\boldsymbol{w}$. The objective function of equalizer is:

$$\min_{\boldsymbol{w}} J(\boldsymbol{w})=\mathrm{E}|d-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{u}|^2. \quad (1)$$

The Weiner solution for Eq.1 is:

$$\boldsymbol{w}=\boldsymbol{R_u}^{-1}\boldsymbol{R}_{du}, \quad (2)$$

where $\boldsymbol{R_u}$ is the positive-definite covariance matrix of $\boldsymbol{u}$ and $\boldsymbol{R}_{du}$ is the cross-covariance vector of d and $\boldsymbol{u}$. 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 linear computational complexity and the RLS has quadratic computational complexity with respect to N. Even after combining with the kernel mapping, such complexity properties remain for the KLMS and the KRLS.

In the AP algorithm, $\boldsymbol{R_u}$ and $\boldsymbol{R}_{du}$ are estimated by using the K most recent inputs. Assuming:

$$\boldsymbol{U}(i) = [\boldsymbol{u}(i\text{-}K\text{+}1) \dots \boldsymbol{u}(i)] \quad (3)$$

$$\boldsymbol{D}(i)=[d(i\text{-}K\text{+}1) \dots d(i)], \quad (4)$$

$\boldsymbol{R_u}$ and $\boldsymbol{R}_{du}$ can be estimated by:

$$\boldsymbol{R}'_{\boldsymbol{u}}=K^{-1}\boldsymbol{U}(i)\boldsymbol{U}(i)^{\mathrm{T}} \quad (5)$$

$$\boldsymbol{R}'_{du}=K^{-1}\boldsymbol{U}(i)\boldsymbol{d}(i). \quad (6)$$

Accordingly, the equalizer coefficient is:

$$\boldsymbol{w}(i)= \boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{U}(i)[\boldsymbol{D}(i)\text{-}\boldsymbol{U}(i)^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)], \quad (7)$$

where η is the step-size factor for algorithm convergence. Eq. 7 shows 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 thus a low computational complexity.

For KAP by combining the kernel mapping with the AP algorithm, the equalizer coefficient update equation shown in Eq.7 needs to be modified by adding kernel mapping feature function $\varphi(\boldsymbol{u}(i))$. The corresponding equalizer coefficient update equation of the KAP is:

$$\boldsymbol{w}(i)=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{u}(i))[\boldsymbol{D}(i)\text{-}\boldsymbol{\Phi}(\boldsymbol{u}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$$
$$=\boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{u}(i))e(i), \tag{8}$$

where $\boldsymbol{\Phi}(\boldsymbol{u}(i))=[\varphi(\boldsymbol{u}(i\text{-}K+1)) \ \ldots\varphi(\boldsymbol{u}(i))]$ and $e(i)$ is the equalizer error at the i-th sampling point. By expanding Eq.8 iteratively, i.e., $\boldsymbol{w}(i) = \boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{u}(i))e(i) = \boldsymbol{w}(i\text{-}2)+\eta\boldsymbol{\Phi}(\boldsymbol{u}(i))e(i)+ \eta\boldsymbol{\Phi}(\boldsymbol{u}(i\text{-}1))e(i\text{-}1)\ldots$, we could find that the converged value of the equalizer coefficient only depends on the kernel mapping feature function $\varphi(\boldsymbol{u}(i))$ and error signal e. Defining the kernel function as κ, the kernel mapping rule [4, 9] is:

$$\kappa(\boldsymbol{u}(i), \boldsymbol{u}(j))= \varphi(\boldsymbol{u}(i))\ \varphi(\boldsymbol{u}(j)). \tag{9}$$

The signal processing flow of the KAP is shown as follows.

Algorithm 1. The signal processing flow of the KAP.

1. Set η, K and choose kernel function κ, i=1;
2. $\boldsymbol{a}_1(1)=\eta d(1)$;
3. while $i \leqslant N$ do
   - i. $\boldsymbol{a}_i(i\text{-}1)=0$;
   - ii. for k= max(1, i-K+1) to i do
     - a) $e_{K+k\text{-}i}(i)=d(k)\text{-} \Sigma_{j\text{-}1 \text{ to } i\text{-}1}\{a_j(i\text{-}1)\kappa(\boldsymbol{u}(k), \boldsymbol{u}(j))\}$
     - b) $a_k(i)= a_k(i\text{-}1)+ \eta e_{K+k\text{-}i}(i)$
     - c) end for
   - iii. $i=i+1$
   - iv. end while
4. Equalizer output of data signal $\boldsymbol{u}'$:
   $f(\boldsymbol{u}')= \Sigma_{1 \text{ to } i}\{a(i)\kappa(\boldsymbol{u}(i), \boldsymbol{u}')\}$

The equalizer coefficient $\boldsymbol{w}$ update equation and computational complexity of different kernel methods are summarized in Table. 1. The KLMS, KRLS and KAP algorithms aim at minimizing the instantaneous error function, the sum of squared estimation errors, and the sum of the K most recent inputs' errors, respectively. Subsequently, the KLMS has the minimal computational complexity of O(N), the KRLS has quadratic computational complexity of O(N^2) and the KAP also keeps linear computational complexity of O(N+K^2) where a constant K must be considered.

Comparing to the Volterra filtering algorithm [1], all the kernel methods show lower computational complexity. The infinite-order Volterra model by itself has non-polynomial computational complexity. With truncated Volterra model parameters and kernel combinations, its computational complexity can be decreased. For instance, the Volterra model has cubic computational complexity when up to the 3rd order nonlinear components are included. However, it is at the expense of performance degradation.

3 Experimental setup and results

The experimental setup and the signal processing flow scheme are shown in Fig. 1. The DMT samples are generated offline and loaded into a 120-GSa/s arbitrary waveform generator (AWG). The output amplitude of the AWG is set to 700-mV. The length of inverse fast Fourier transform (IFFT) for the DMT modulation is set to 2048. The signal from the AWG is amplified and modulated by an external modulated laser (EML). The measured central wavelength is ~1549.8-nm. After 400m length standard single mode fibre (SSMF) transmission, the DMT signal is detected by the photo-detector (PD). Because of the lack of trans-impedance amplifier, one erbium doped fibre amplifier (EDFA) together with a variable optical attenuator (Att.) is used before PD to study the sensitivity. The electrical signal from direct detection is captured by the 256-GSa/s digital storage oscilloscope (DSO).

The BER performance in terms of K in the KAP scheme is shown in Fig. 2. The size of length of each training symbol is 2048 DMT samples (i.e., N = 2048), which corresponds to the IFFT size. With the increase of K in the KAP, the BER

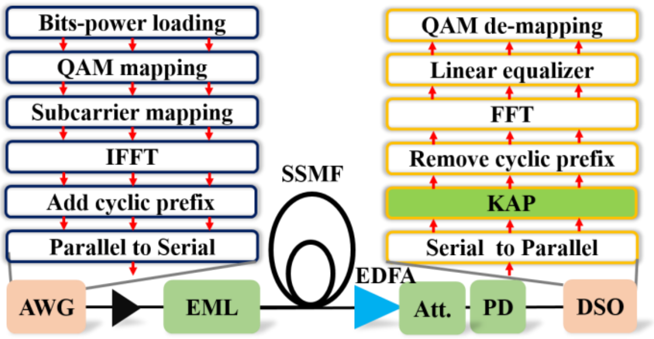

Fig. 1. Experimental setup.

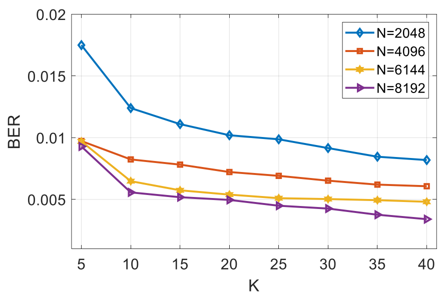

Fig. 2. BER performance versus K and N.

Table. 1. The equalizer coefficient $\boldsymbol{w}$ update equation and computational complexity of different kernel methods.

Kernel methods	Equalizer coefficient update equation	Computational complxity
KLMS[4]	$\boldsymbol{w}(i)= \boldsymbol{w}(i\text{-}1)+\eta\varphi(\boldsymbol{u}(i))[d(i)\text{-} \varphi(\boldsymbol{u}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	O(N)
KRLS[5]	$\boldsymbol{w}(i)= [\eta\boldsymbol{I}+\boldsymbol{\Phi}(i)\boldsymbol{\Phi}(i)^{\mathrm{T}}]^{-1}\boldsymbol{\Phi}(i)\boldsymbol{D}(\mathrm{i})$	O(N^2)
KAP	$\boldsymbol{w}(i)= \boldsymbol{w}(i\text{-}1)+\eta\boldsymbol{\Phi}(\boldsymbol{u}(i))[\boldsymbol{D}(i)\text{-}\boldsymbol{\Phi}(\boldsymbol{u}(i))^{\mathrm{T}}\boldsymbol{w}(i\text{-}1)]$	O(N+K^2)

improves gradually. The same trend can be observed from the curves with different numbers of training data size in Fig. 2. The BER improvement tends to get statured when K and N increased. The N of 8192 and K of 40 are selected for the experiment results presented later.

The BER versus the received optical power (RoP) is shown in Fig. 3. This experimental system cannot achieve HD-FEC with one-tap least square (LS) equalizer [10], which is often considered to compensate linear impairments. HD-FEC can be achieved when the KAP is employed on top of the LS equalizer. Therefore, the additional improvement achieved by the KAP demonstrates capability of the KAP to mitigate nonlinear impairments in high-speed optical direct detection systems, such as nonlinearities from the driving stage amplifiers, the chirp from EML [11] and square-law detection in PD. The constellation graphs after LS equalizer, LS+KAP equalizer are shown in Fig. 4. The system adaptively loads quadrature phase shift keying (QPSK), 8-quadrature amplitude modulation (8QAM), 16QAM, 32QAM and 64QAM with bits-power loading algorithm. Fig. 4 shows 16QAM, 32QAM and 64QAM. The system signal noise ratio is improved. The maximum system line rate that can achieve is 238Gbps with the net rate of 222Gbps showing potential of this low complexity algorithm.

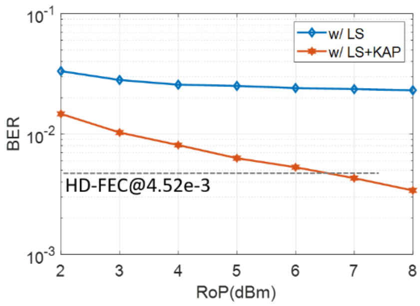

Fig. 3. System transmission performance.

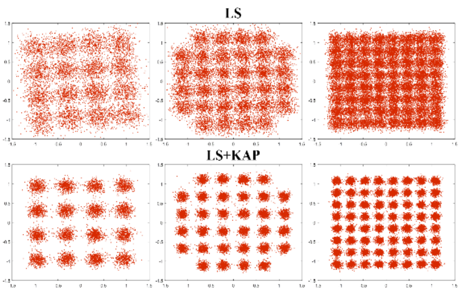

Fig. 4. Constellations graphs after LS equalizer and LS+KAP equalizer, respectively.

Table. 2. The BER with different equalizers.

Equalizers	BER
KLMS	7.10e-3
Volterra(2nd)	6.90e-3
Volterra(3rd)	5.87e-3
HD-FEC	*4.52e-3*
KRLS	3.62e-3
KAP	3.40e-3
Volterra(4th)	3.24e-3

The BER for different equalizers measured at the RoP of 8dBm (the best BER performance achieved in the experiments) is compared in Table. 2. The Volterra filtering with a memory length of 10 is also compared here. The KAP achieves similar performance as the KRLS and the Volterra filtering algorithm with nonlinear kernels up to the 4th order, whereas the KAP has much lower computational complexity. The KAP obviously outperforms the KLMS and Volterra filtering algorithm with nonlinear kernels up to the 2nd/3rd order.

4 Conclusions

A novel KAP algorithm is proposed for compensating nonlinear impairments in optical direct detection systems. A 222Gbps net rate IM/DD DMT system is experimentally carried out to verify the performance, the results show that the KAP algorithm can improve the transmission performance to reach HD-FEC with low computational complexity.

5 Acknowledgements

Keysight Technologies is acknowledged for the loan of DSO (Keysight UXR1102A) and AWG (Keysight M8194A). We would like to thank Hadrien Louchet, Thomas Dippon and Markus Gruen from Keysight Technologies GmbH for the help with experimental setup. The work is supported by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF), EU H2020 MSCA-IF Project NEWMAN (no. 752826), VINNOVA-funded SENDATE-EXTEND, and SENDATE-FICUS with project no. 2017-01559.

6 References

[1] L. Zhang, et al.: 'Nonlinearity-aware 200 Gbit/s DMT transmission for C-band short-reach optical interconnects with a single packaged electro-absorption modulated laser', Optics Letters, 2018, **43**, (2), pp. 182-185.

[2] G. Chen, et al.: 'Nonlinear Distortion Mitigation by Machine learning of SVM classification for PAM-4 and PAM-8 modulated optical interconnection', Journal of Lightwave Technology, 2017, **36**, (3), pp. 650-657.

[3] E. Novak, et al.: 'The curse of dimension and a universal method for numerical integration', Multivariate approximation and splines, 1997, pp. 177-187.

[4] L. Zhang, et al.: 'Kernel Adaptive Filtering for Nonlinearity-Tolerant Optical Direct Detection Systems', in Proc. ECOC, 2018.

[5] 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. ECOC, 2018.

[6] M. Scholten, et al.: 'Continuously-interleaved BCH (CIBCH) FEC delivers best in class NECG for 40G and 100G metro applications', in Proc. OFC, 2010, Paper NTuB3.

[7] E. Agrell, et al.: 'Information-Theoretic Tools for Optical Communications Engineers', in Proc. IPC, 2018, tutorial.

[8] Y. Li, et al.: 'Sparse adaptive channel estimation based on-norm-penalized affine projection algorithm', International Journal of Antennas and Propagation, 2014.

[9] W. Liu, et al.: 'Kernel adaptive filtering: a comprehensive introduction', John Wiley & Sons 2011, 57.

[10] L. Zhang, et al..: 'Channel Estimation Algorithm for Interference Suppression in IMDD-OQAM-OFDM Transmission Systems', in Optics Communications, 2016, 364, pp. 129-133.

[11] M. Chaciński et al.: 'Modulation and chirp evaluation of 100 GHz DFB-TWEAM', in Proc. ECOC, 2010, pp. 1-3.