Fast Convergence by Machine Learning Optimizer for Adaptive MIMO Equalizer Used in SDM Transmission over Coupled-Core 4-Core Fiber and 4-Core EDFA

Manabu Arikawa* and Hidemi Noguchi

System Platform Research Laboratories, NEC Corporation, 1753 Shimonumabe, Nakahara-ku, Kawasaki, 211–8666, Japan *marikawa@nec.com

Abstract *We applied several optimizers from the machine learning field to an adaptive MIMO equalizer for SDM transmission. We experimentally compared their convergence properties in a SDM transmission over a 52-km coupled-core 4-core fiber and 4-core EDFA and showed over 22% faster convergence with Adam.*

Introduction

Space-division multiplexing (SDM) is a promising technology for overcoming the capacity limit of single-mode fibers (SMFs)[1]. Recent studies have demonstrated that coupled-core multi-core fibers (CC-MCFs) are attractive alternatives for ultra-long-haul transmission[2]-[6]. CC-MCFs can be designed to achieve low loss along with low spatial mode dispersion, which is closely related to the complexity of the multi-input multi-output (MIMO) equalizer for compensation of coupling among signals[7],[8]. The high spatial efficiency of CC-MCFs is also favorable in terms of integration of system components such as CC-MC optical amplifiers[9].

In SDM transmission systems with coupled spatial modes, MIMO digital signal processing (DSP) at the receiver side is necessary. Its filter coefficients must be adaptively controlled, similar to a 2×2 MIMO filter for polarization-division multiplexed (PDM) signals. Since adaptive MIMO equalizers for SDM transmission often have numerous filter coefficients, the convergence property should be carefully considered[10],[11]. To accelerate the convergence, a modified least mean square (LMS) algorithm for a frequency-domain MIMO equalizer using the average power within each frequency bin has been proposed[11].

For adaptive control of MIMO equalizers, stochastic gradient descent (SGD) is conventionally used in optical communication[12]. Interestingly, the machine learning field offers additional extended methods such as Momentum, RMSprop, and Adam, which are utilized for learning neural networks to improve the convergence property[13]. In this study, we applied these optimizers to an adaptive MIMO equalizer and evaluated their convergence properties in a transmission of 32-Gbaud PDM-QPSK/16QAM/64QAM over one span consisting of a 52-km coupled-core 4-core fiber (CC-4CF) and a coupled-core 4-core erbium-doped fiber amplifier (CC-4C-EDFA). The results show that using the root mean square (RMS) of gradients can accelerate the convergence.

Optimizers of adaptive MIMO equalizer

Conventional SGD updates a filter coefficient ξ^* of an adaptive MIMO equalizer as

$$\xi^* \to \xi^* - 2\alpha\frac{\partial\phi}{\partial\xi},$$

where ϕ is the loss to minimize and α determines the step size. In the case of the constant modulus algorithm (CMA), the loss is the magnitude of errors of the MIMO equalizer outputs y_i from the desired amplitude r, as

$$\phi = \sum_i \phi_i = \sum_i (r^2 - |y_i|^2)^2.$$

Momentum is a method that uses accumulated gradients. The filter coefficient update becomes

$$v \to \beta_1 v + (1-\beta_1)\frac{\partial\phi}{\partial\xi}, \quad \xi^* \to \xi^* - 2\alpha v,$$

where β_1 is the decay factor of the exponential weighted moving average.

RMSprop is a method that adjusts a learning rate with the RMS of each gradient, as

$$s \to \beta_2 v + (1-\beta_2)\left|\frac{\partial\phi}{\partial\xi}\right|^2, \quad \xi^* \to \xi^* - \frac{2\alpha}{\sqrt{s}+\epsilon}\frac{\partial\phi}{\partial\xi},$$

where β_2 and ϵ are the decay factor and a small amount to avoid zero division, respectively.

Adam combines Momentum and RMSprop.

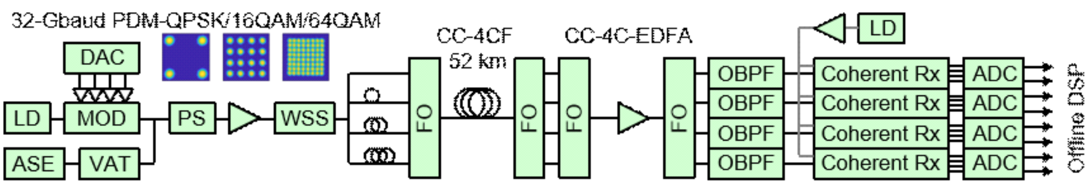

Fig. 1: Experimental setup for SDM transmission over one span of 52-km CC-4CF and CC-4C-EDFA.

The filter coefficient update is

$$v \to \beta_1 v + (1-\beta_1)\frac{\partial \phi}{\partial \xi}, \quad s \to \beta_2 s + (1-\beta_2)\left|\frac{\partial \phi}{\partial \xi}\right|^2,$$
$$\hat{v} = v/(1-\beta_1^t), \quad \hat{s} = s/(1-\beta_2^t),$$
$$\xi^* \to \xi^* - \frac{2\alpha}{\sqrt{\hat{s}}+\epsilon}\hat{v}.$$

Experimental setup

The experimental setup is shown in Fig. 1. 32-Gbaud PDM-QPSK/16QAM/64QAM with a roll-off factor of 0.1 was generated with a laser diode (LD) at the frequency of 193.3 THz having a linewidth of about 100 kHz and a four-channel digital-to-analog converter (DAC) at the sampling rate of 64 GS/s. The data consisted of eight (four in the case of QPSK) frames of the forward error correction (FEC) of low-density parity-check code for DVBS-2 with a frame length of 64,800 and a code rate of 4/5 while loading random bits to the payload. Amplified spontaneous emission (ASE) was added to the optical signal to set an OSNR. The signal was divided into four after low-speed polarization scrambling (PS), an EDFA, and a wavelength selective switch (WSS). The four signals were then decorrelated to emulate SDM signals.

The SDM signals were transmitted through one span of a CC-4CF and a CC-4C-EDFA. The length of the CC-4CF was 52 km. The cladding diameter and the effective area of each core were 125 μm and 109 μm^2. The loss coefficient was 0.172 dB/km and the total loss including the fanouts (FOs) at both ends was about 11 dB. The span input optical power was set to 0 dBm/core. The average noise figure of CC-4C-EDFA was 5.7 dB. In the case of no ASE added at the transmitter side, the received OSNR was 38.8 dB/0.1 nm.

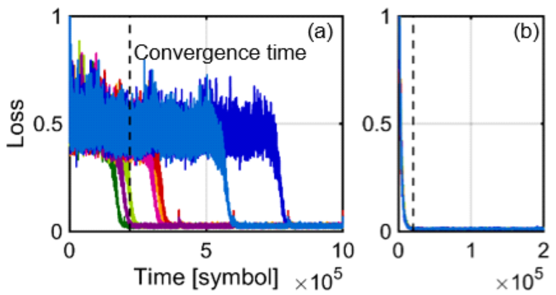

Fig. 2: Convergence of (a) SDG CMA and (b) SDG DALMS in SDM transmission of QPSK

After transmission over one span, each of the SDM signals was filtered by an optical bandpass filter (OBPF) and received by a polarization diversity coherent receiver. A common LD having a linewidth of about 4 kHz was used as a local oscillator for four coherent receivers. The outputs were sampled with a 16-channel digital oscilloscope as an analog-to-digital converter (ADC) at the sampling rate of 50 GS/s.

DSP was performed offline. After normalization and resampling to two-fold oversampling, chromatic dispersion (CD) compensation and matched low-pass filtering (LPF) were performed. Then, an adaptive 8×8 MIMO equalizer was applied. Finally, the pre/post-FEC bit error rate (BER) was evaluated.

MIMO DSP

The MIMO equalizer consisted of finite impulse response (FIR) filters with T/2 121 taps. In an adaptive MIMO equalizer, some of the outputs have a tendency to degenerate, known as the singularity problem[14]. This problem can be avoided by using data-aided LMS (DALMS)[5]. However, blind demodulation prior to DALMS is generally required to synchronize a training sequence due to spatial mode coupling and spatial mode dispersion in a channel.

Therefore, the MIMO DSP was performed as follows. After the filter coefficients were initialized, blind CMA and decision-directed LMS (DDLMS) were repeatedly applied to a part of the received signals. The goal of this stage is that at least one output of the MIMO equalizer be demodulated. It enables specifying the head of the transmitted pattern if different training sequences are used for different SDM signals. (Although the SDM signals were emulated from one signal in this experiment, this enables full-search.) Then, the filter coefficients were re-initialized, and DALMS was applied to a part of the received signals until all the outputs converged. Finally, DDLMS was applied to another part of the received signals. The pre/post-FEC BER was evaluated for the signals after DDLMS.

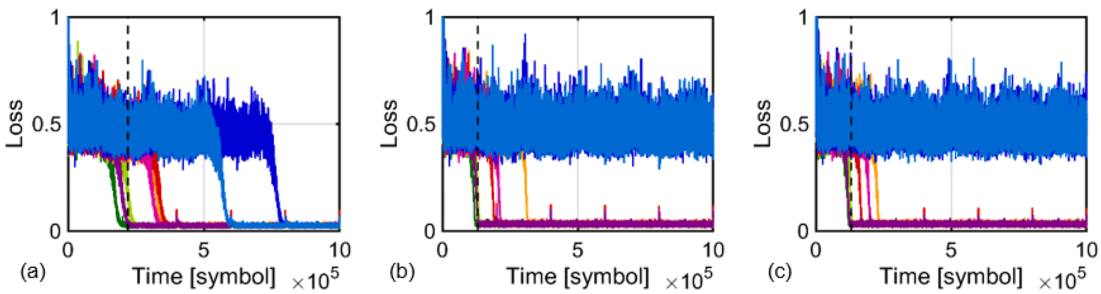

Fig. 3: Convergence of (a) Momentum CMA, (b) RMSprop CMA, and (c) Adam CMA.

Experimental results

We first compared the convergence of blind demodulation and DALMS with conventional SGD. Figure 2 shows the time evolution of the losses ϕ_i with blind CMA and DALMS for the SDM transmission of PDM-QPSK over one span of a 52-km CC-4CF and a CC-4C-EDFA without ASE loading. In the case of CMA, the losses gradually decreased and then at a certain timing each of them started to converge rapidly. A total of 2.2×10^5 symbols was required to converge, that is, until one output converged. In contrast, the losses converged from the beginning in the case of DALMS and it required 0.2×10^5 symbols to converge. These results demonstrate that the blind demodulation stage to detect a head of the sequence takes much more time to converge than DALMS does.

We thus applied several optimizers to the blind CMA stage and compared the convergence. Figure 3 shows the time evolution of the losses with Momentum, RMSprop, and Adam for the same received waveforms used in Fig. 2. The same α was used in this comparison. Momentum provided a similar result to SGD. In the case of RMSprop and Adam, they required 1.3×10^5 symbols to converge. These results demonstrate that using the RMS of gradients can accelerate the convergence. Figure 4 shows the time required to converge with blind SDG/RMSprop/Adam CMA for several transmitted OSNR conditions. Ten waveforms acquired at different timings were evaluated for each condition. The solid lines indicate the average and the error bars correspond to minimum to maximum. Adam improved the average required time to converge by 27 to 44% depending on the OSNR.

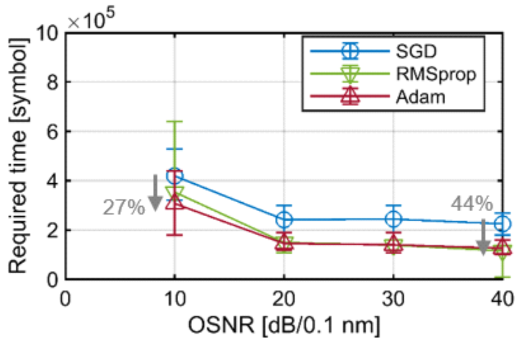

Fig. 4: Required time to converge with blind CMA.

Figure 5 shows the performances after DDLMS in the SDM transmission of PDM-QPSK/16QAM/64QAM over one span. Adam was used in the blind demodulation stage. OSNR penalty was evaluated by sweeping the transmitted OSNR. In Fig. 5, the pre/post-FEC BERs of eight signals are plotted as markers with different colors and the averages over signals are plotted as lines. The results under the back-to-back condition are also plotted, where four 2×2 MIMO equalizers with 41 taps were used instead of an 8×8 MIMO equalizer. In the cases of QPSK and 16QAM, only a small OSNR penalty (less than 1 dB) was observed after the 1-span transmission in the post-FEC BER. In the case of 64QAM, there was a high BER floor in the pre-FEC BER due to the limited bandwidth of the receivers, and the OSNR penalty was 2 dB in the post-FEC BER. These results demonstrate that the transmission over a CC-4CF, CC-4C-EDFA, and MIMO DSP incurred only a small penalty.

Conclusion

We compared the convergence property of an adaptive MIMO equalizer with SDG, Momentum, RMSprop, and Adam in the SDM transmission of 32-Gbaud PDM-QPSK/16QAM/64QAM over a 52-km CC-4CF and a CC-4C-EDFA. The results showed that using the RMS of gradients can accelerate the convergence by 27 to 44% compared to the conventional SDG, depending on the OSNR condition.

Acknowledgements

This work was partially supported by the Ministry of Internal Affairs and Communications, Japan (JPMI00316). We thank Kohei Hosokawa, Masaki Sato, Keiichi Matsumoto, and Hitoshi Takeshita for their support with the experiments.

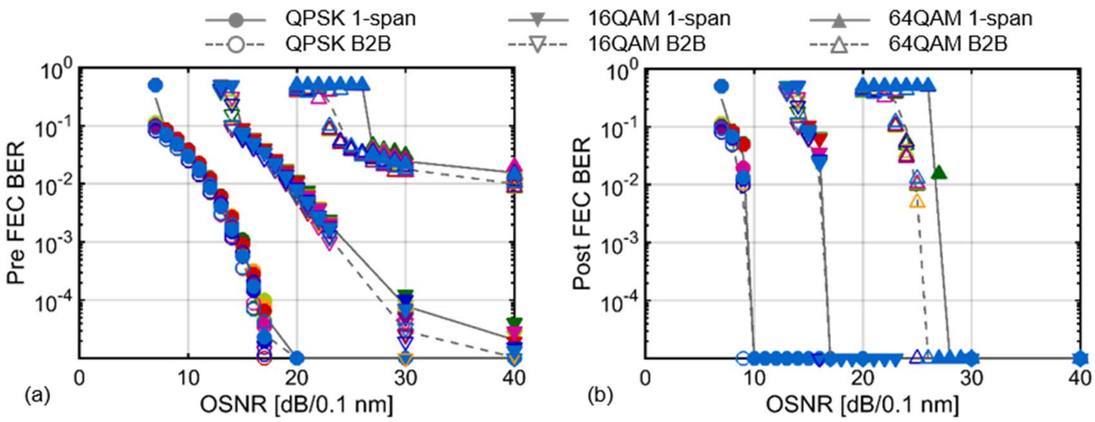

Fig. 5: Received (a) pre- and (b) post-FEC BER after transmission over one span of 52-km CC-4CF and CC-4C-EDFA.

References

[1] P. J. Winzer *et al.*, "Fiber-optic transmission and networking: the previous 20 and the next 20 years," Opt. Express **26**(18), 24190 (2018).

[2] C. Xia *et al.*, "Supermodes for optical transmission," Opt. Express **19**(17) 16653 (2011).

[3] R. Ryf *et al.*, "Space-division multiplexed transmission over 4200-km 3-core microstructured fiber," OFC 2012, PDP5C.2.

[4] R. Ryf, "Long-haul transmission over multi-core fibers with coupled cores," ECOC 2017, M.2.E.1.

[5] R. Ryf *et al.*, "Coupled-core transmission over 7-core fiber," OFC 2019, Th4B.3.

[6] C. Antonelli *et al.*, "Nonlinear interference noise in space-division multiplexed transmission through optical fibers," Opt. Express **25**(12), 13055 (2017).

[7] T. Sakamoto *et al.*, "Coupled multicore fiber design with low intercore differential mode delay for high-density space division multiplexing," J. of Lightwave Technol. **33**(6), 1175 (2015).

[8] T. Hayashi *et al.*, "Record-low spatial mode dispersion and ultra-low loss coupled multi-core fiber for ultra-ling-haul transmission," J. of Lightwave Technol. **35**(3), 450 (2017).

[9] N. K. Fontaine *et al.*, "Coupled-core optical amplifier," OFC 2017, Th5D.3.

[10] S. Randel *et al.*, "Complexity analysis of adaptive frequency-domain equalization for MIMO-SDM transmission," ECOC 2013, Th.2.C.4.

[11] X. He *et al.*, "A fast convergence frequency domain least mean square algorithm for compensation of different mode group delay in few mode fibers," OFC 2013, OM2C.4.

[12] S. J. Savory, "Digital filters for coherent optical receivers," Opt. Express **16**(2), 804 (2008).

[13] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," ICLR 2015.

[14] L. Liu *et al.*, "Initial tap setup of constant modulus algorithm for polarization de-multiplexing in optical coherent receivers," OFC 2009, OMT2.