Fast-Tracking Optical Coherent Receiver Tolerating Transmitter Component Distortion

Etsushi Yamazaki, *Member, IEEE* and Shinya Sugiura, *Senior Member, IEEE*

***Abstract*—In this paper, we propose a receiver architecture capable of compensating for a time-varying signal distortion induced by the transmitter analog components. In the proposed architecture, the novel transmitter distortion compensating (TDC) block is implemented outside of the feedback control loop of the adaptive equalizer to suppress the feedback delay by optimizing the parameter used in the least mean square (LMS) algorithm. The proposed TDC block is designed for compensating for the DC offset, the IQ amplitude imbalance, the IQ orthogonality error, and the IQ timing skew. Our simulation results show that the proposed scheme significantly reduces the required signal-to-noise ratio penalty from the theoretical limit, which is imposed due to the transmitter components distortion. Furthermore, our theoretical analysis confirms that the delay induced in the feedback loop of the adaptive equalizer determines the upper bound of the LMS step size under the stable condition, hence allowing us to maximize the tracking speed of our receiver.**

***Index Terms*—Adaptive equalization, analog distortion, fast-tracking, least mean square, optical fiber communication, theoretical bound.**

I. Introduction

An optical fiber network acts as a lifeline platform, which is essential for economic activities. The capacity demand is further driven due to the expansions of inter/intra-data center communications and the evolution of wireless back-haul networks with a variety of applications, such as on-demand video streaming and cloud computing. Recent experiment reported that high capacity per wavelength, exceeding 1 Tbps, was achieved [1]–[7], where both the symbol rate and modulation size significantly increased. For example, the symbol rate increased from 32 Gbaud to more than 96 Gbaud [8]–[10]. This was achieved as the explicit benefits of the improved analog-to-digital (AD)/digital-to-analog (DA) converters [11]–[13], of with frequency-doubling techniques [2], [6], [14], [15], and of the improved analog components, such as optical modulators [16], [17], [10], drivers, photodiodes, and transimpedance amplifier (TIA) [18]. The modulation size was increased to more than 64-point quadrature amplitude modulation (QAM) constellations with the aid of capacity-approaching probabilistic constellation shaping [19]–[21].

In high-rate optical fiber communication, the signal distortion is induced by analog components at the transmitter and the receiver. To be more specific, upon increasing the modulation size, the detrimental effects of analog component distortions increase since the higher symbol rate causes more severe analog distortions. Several digital-domain equalization techniques at the receiver have been developed to compensate for the distortion imposed by the receiver-side components [22]–[24].

The transmitter-side analog components, such as the modulator, the driver, and electrical wiring on the printed circuit board, also distort signal because of their limited bandwidth, I/Q crosstalk, IQ skew, DC offsets, and nonlinearity. The sophisticated digital pre-distortion methods are proposed and transmitter-side distortions, including nonlinearity can be compensated [25]–[28]. Typically, the distortion on the transmitter side is estimated by the calibration process, and the optimum coefficients are set in a digital pre-distortion block on the transmitter side and post-compensation block on the receiver side. However, this calibration process generally needs special equipment or back-to-back configuration, which is challenging to be performed during the operation condition over the fiber channel. Thus, time-varying distortions caused by a variation in physical conditions, such as operating temperature or component aging, may not be tracked without relying on any additional feedback channel. Also, the digital pre-distortion method cannot compensate for DC offset because of the AC coupling of analog devices at the transmitter. Thus, the receiver has to tolerate residual transmitter-side distortion.

To this end, several receiver architectures were proposed to compensate for the time-varying transmitter-side distortion with the aid of receiver-side processing [29]–[31]. The compensating for the transmitter-side distortion at the receiver has advantages and disadvantages. More specifically, the digital pre-distortion tends to increase the peak-to-average power ratio in the DAC and to increase the quantization errors of the transmitter [27]. Thus, the partial compensation at the receiver benefits from suppressing the quantization noise of the transmitting. However, the receiver-side compensation for the transmitter nonlinearity and bandwidth limitation increases the associated noise components [32]. The paper shows that the receiver-side nonlinearity equalization using the Volterra filter remains a considerable amount of the SNR penalty. To reduce the effects, a more sophisticated technique, such as equalization in the likelihood ratio domain, is required. Thus, this paper assumes that most nonlinearity is compensated for at the transmitter. This allows us only to consider a small

E. Yamazaki is with the Network Innovation Laboratories, NTT Corporation, Hikarino-oka, Yokosuka, Kanagawa, 239-0847, Japan, and is also with the Institute of Industrial Science, The University of Tokyo, Meguro, 153-8505, Japan (e-mail: etsushi.yamazaki.wk@hco.ntt.co.jp).

S. Sugiura is with the Institute of Industrial Science, The University of Tokyo, Meguro, 153-8505, Japan (e-mail: sugiura@iis.u-tokyo.ac.jp).

amount of residual nonlinearity of a received signal.

Furthermore, in optical fiber transmission, a receiver has to track the state of polarization (SOP) and the polarization mode dispersion (PMD) even in a rapidly time-varying channel [33]–[35]. In the extreme case, it was reported that significantly fast SOP variation is caused by lightning strikes [36]–[38]. Typically, the receiver has to decompose the multiplexed polarization signals while eliminating intersymbol interference (ISI) caused by PMD and residual chromatic dispersion (CD), in an adaptive manner. The adaptive algorithms, such as the least-mean squares (LMS) algorithm and constant modulus algorithm (CMA), are used for tracking the channel coefficients and optimizing the parameters of an equalizer.

An upper limit of the step size of the LMS algorithm or CMA under the operation of stable convergence depends on the delay in the feedback control loop [39], [40], while a large step size allows us to increase the tracking speed. This implies that the delay in the LMS feedback control loop may be the limiting factor of an adaptive equalizer. Thus, to compensate for the signal distortion induced by the transmitter-side analog components at the receiver, the delay in the LMS feedback control loop has to be maintained sufficiently low.

Against the above-mentioned background, the novel contributions of this paper are as follows.

- We propose a novel fast-tracking adaptive optical receiver capable of compensating for a time-varying signal distortion induced by the transmitter analog components in the operation condition for transmission over the fiber channel. More specifically, we introduce a transmitter distortion compensating (TDC) block, which is capable of suppressing the feedback delay by optimizing the parameters used in the LMS-based adaptive equalizer. The proposed TDC block is capable of compensating for the DC offset, the IQ amplitude imbalance, the IQ orthogonality error, and the IQ timing skew. More specifically, the replica signal used in the LMS-based adaptive equalizer is modified, depending on the TDC coefficients, which allows our TDC block to be outside of the adaptive equalizer feedback loop, and hence is capable of suppressing the feedback delay.
- Moreover, we propose a theoretical analysis framework, which is used for optimizing the LMS step size in the presence of the feedback loop delay of an adaptive equalizer. Our simulation results show that our proposed scheme achieves a high bit-error-ratio (BER) performance close the theoretical bound while eliminating the detrimental effects of signal distortion. Furthermore, it is found that the proposed scheme's penalty steeply increases upon increasing the LMS step size beyond the derived upper limit, which verifies our theoretical framework.

The remainder of this paper is organized as follows. In Section II, the system model of our optical-fiber transceiver is presented. In Section III, the fast-tracking adaptive equalizer architecture compensating for the transmitter distortion is presented. In Section IV, the theoretical upper limit of the LMS step size of the adaptive equalizer is derived. In Section V, we provide performance results on penalty evaluation of the tracking speed of the proposed adaptive equalizer. Finally, Section VI concludes this paper.

II. System Model

Fig. 1 shows the system model for the optical-fiber transceiver considered in this paper. We assume the two orthogonal polarization dimensions of the single-mode optical fiber, each employing 2^M-sized QAM. In the IQ mapping block of each polarization, M bits are modulated onto 2^M-QAM symbols to generate $u(n)$ or $v(n)$, where n is the symbol index. Then, $u(n)$ and $v(n)$ are twice up-sampled with the aid of a root raised cosine (RRC) shaping filter to generate $x_{\mathrm{T}}(n)$ and $y_{\mathrm{T}}(n)$, respectively. Finally, the up-sampled signals are fed into the nested Mach-Zehnder (MZ) modulator, which consists of two inner Mach-Zehnder interferometers (MZIs) and a single outer MZI. More specifically, modulated baseband signal of the MZI x_{M} is modeled by

$$x_{\mathrm{M}} = j\frac{a_{\mathrm{O}}}{\sqrt{2}}\left(a_{\mathrm{I}}e^{j\frac{\pi}{2}\mathfrak{Re}\{x_{\mathrm{T}}\}} - \sqrt{1-a_{\mathrm{I}}^2}e^{-j\frac{\pi}{2}\mathfrak{Re}\{x_{\mathrm{T}}\}}\right) - \sqrt{\frac{1-a_{\mathrm{O}}^2}{2}}\left(a_{\mathrm{I}}e^{j\frac{\pi}{2}\mathfrak{Im}\{x_{\mathrm{T}}\}} - \sqrt{1-a_{\mathrm{I}}^2}e^{-j\frac{\pi}{2}\mathfrak{Im}\{x_{\mathrm{T}}\}}\right)\cdot e^{j\theta_{\mathrm{e}}}, \tag{1}$$

where the variables a_{I} and a_{O} are associated with the extinction ratios of the inner and outer MZIs, respectively. The extinction ratios are derived as the ratio of the maximum output power to the minimum power as [41]

$$R_{\mathrm{E}i} = \frac{(a_i + \sqrt{1-a_i^2})^2}{(a_i - \sqrt{1-a_i^2})^2} \quad (i \in \{\mathrm{O}, \mathrm{I}\}), \tag{2}$$

Then, (2) is rewritten as

$$a_i = \frac{1}{\sqrt{2}}\frac{\sqrt{R_{\mathrm{E}i}}+1}{\sqrt{R_{\mathrm{E}i}+1}} \quad (i \in \{\mathrm{O}, \mathrm{I}\}), \tag{3}$$

Furthermore, θ_e denotes the 90-degree phase errors of the outer MZI, and $\mathfrak{Re}\{x_{\mathrm{T}}\}$ and $\mathfrak{Im}\{x_{\mathrm{T}}\}$ correspond to I- and Q-components of the signal $x_{\mathrm{T}}(n)$ output from the RRC shaping block.

A low extinction ratio of the inner MZI R_{EI} causes the DC offsets in the $\mathfrak{Re}\{x_{\mathrm{T}}\}$ and $\mathfrak{Im}\{x_{\mathrm{T}}\}$. Also, a low extinction ratio of the outer MZI R_{EO} induces imbalance between $\mathfrak{Re}\{x_{\mathrm{T}}\}$ and $\mathfrak{Im}\{x_{\mathrm{T}}\}$, which is also caused by the gain differences of drivers amplifiers or passive path loss differences. The 90-degree phase errors originating from the control error of the outer MZI DC biases cause signal cross-talk between $\mathfrak{Re}\{x_{\mathrm{T}}\}$ and $\mathfrak{Im}\{x_{\mathrm{T}}\}$. Additionally, there typically exists the residual or dynamic skew between $\mathfrak{Re}\{x_{\mathrm{T}}\}$ and $\mathfrak{Im}\{x_{\mathrm{T}}\}$, even after the transmitter-side digital pre-equalization.These distortion factors cannot be perfectly compensated for only by the transmitter-side optimization, and combating such limitations are the targets of the proposed scheme. In this paper, we exclude compensation of nonlinear distortions, which is typically caused by the driver and the inner MZI, while the MZI nonlinearity distortion is included in our simulation.

As mentioned above, we assume that most nonlinear distortion is compensated for with the digital pre-distortion of the transmitter. Hence, the nonlinear distortion compensation

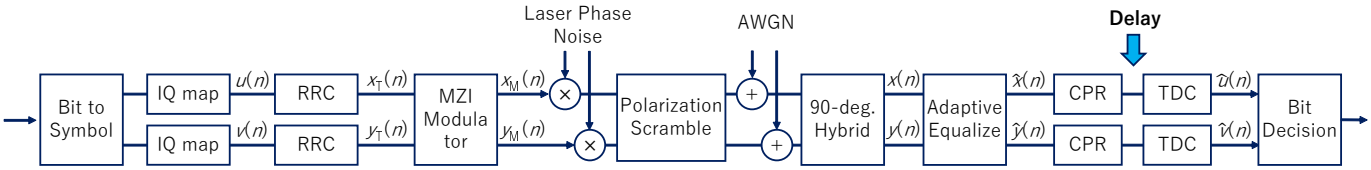

Fig. 1. System model of our optical-fiber transceiver.

is not our target at the receiver. Hence, only the nonlinearity induced by the inner MZI modulator, rather than the driver, is considered. To be specific, the RMS of the signal input to the inner MZI was set at $0.20\ \mathrm{V}\pi$. The SNR after the inner MZI was 34.5 dB, which is determined only by nonlinearity distortion. These are the reasonable conditions for the scenario where digital pre-distortion is used, and only the residual nonlinearity is present.

Transmit signals in each polarization are contaminated by laser phase noises (PNs), and the polarization scrambler rotates the polarization state. Then, the signal experiences the additive white Gaussian noises (AWGNs). Then, the AWGN-contaminated signals are fed into the 90-degree hybrid block to decompose I- and Q-phase orthogonal polarization signals. Note that the I- and Q-phase orthogonal polarization symbols at the receiver $x(n)$ and $y(n)$ are different from the I- and Q-lanes of each polarization at the transmitter $x_{\mathrm{T}}(n)$ and $y_{\mathrm{T}}(n)$ since the PNs rotate the carrier phases and the polarization scrambler mixes the two polarization symbols. The two complex-valued signals output from the 90-degree hybrid block are fed into the 2×2 complex-valued adaptive equalizer to de-multiplex two polarization symbols. The tap coefficients of the equalizer are calculated for minimizing the square error loss of the signal output from pilot-based feed-forward (PFF) carrier phase recovery (CPR), as defined by (11) and (12) in the next section, and the effects of the delay are included in the feedback control loop. The proposed TDC block is placed after the PFF-CPR block.

III. Proposed Compensation for Transmitter-Side Component Distortion

In this section, we propose a novel receiver architecture that is capable of tracking the rapidly time-varying channel and compensating for the signal distortion induced by the transmitter-side analog components. Our TDC block of Fig. 1 is designed for compensating for the DC offset, the amplitude imbalance, the 90-degree phase error, and the skew between I- and Q-lanes, as mentioned in Section II. The effects of nonlinear distortions are not compensated for in our scheme since it typically involves the noise enhancement issue. For detail, refer to [32].

Fig. 2 shows the proposed receiver architecture designed for single-mode optical fiber transmission.

A. FIR Block

In the 2×2 complex-valued finite impulse response (FIR) block of Fig. 2, the polarization cross-talk is canceled out, and the PMD distortion and ISI are compensated for to recover polarization-demultiplexed signals. Let us introduce two polarization-diversity input signals of

$$\boldsymbol{x}(n) = [x(n-L), ..., x(n), ..., x(n+L)]^T \in \mathbb{C}^{2L+1} \quad (4)$$

$$\boldsymbol{y}(n) = [y(n-L), ..., y(n), ..., y(n+L)]^T \in \mathbb{C}^{2L+1}. \quad (5)$$

where n is the time index of symbols, and the filter length is given by $2L+1$. Consider four complex-valued coefficient vectors of the 2×2 complex-valued FIR filters of $\boldsymbol{h}_{xx}(n) \in \mathbb{C}^{2L+1}$, $\boldsymbol{h}_{xy}(n) \in \mathbb{C}^{2L+1}$, $\boldsymbol{h}_{yx}(n) \in \mathbb{C}^{2L+1}$, and $\boldsymbol{h}_{yy}(n) \in \mathbb{C}^{2L+1}$. Then, we obtain the two de-multiplexed output symbols as follows:

$$\hat{x}(n) = \boldsymbol{x}(n)^T \boldsymbol{h}_{xx}(n) + \boldsymbol{y}(n)^T \boldsymbol{h}_{xy}(n) \quad (6)$$

$$\hat{y}(n) = \boldsymbol{x}(n)^T \boldsymbol{h}_{yx}(n) + \boldsymbol{y}(n)^T \boldsymbol{h}_{yy}(n), \quad (7)$$

which are fed back into the PFF-CPR block of Fig. 2.

B. PFF-CPR Block

In the PFF-CPR block, a carrier phase error is calculated by averaging the received and transmitted pilot symbols. More specifically, the estimated phase error of the x-polarization is given by

$$\phi(n) = \arg\left(\frac{1}{2N_{\mathrm{PA}}+1} \sum_{k=-N_{\mathrm{PA}}}^{N_{\mathrm{PA}}} \frac{\hat{x}(I_{\mathrm{p}}(n,k))}{f_{\mathrm{Dx}}(u(I_{\mathrm{p}}(n,k)))} \right), \quad (8)$$

where $f_{\mathrm{D}}(u)$ is the distortion function (DF) representing the inverse of TDC, which is formulated in the later part of this section using TDC coefficients. The DF contributes to reproducing the transmitter distortion. Also, $2N_{\mathrm{PA}}+1$ is the window size. Furthermore, $I_{\mathrm{p}}(n,k)$ denotes the symbol index of the kth nearest pilot symbol to the symbol index n. When k is a natural number, $I_{\mathrm{p}}(n,k)$ is larger than or equal to n. Similarly, when k is a negative number, $I_{\mathrm{p}}(n,k)$ is smaller than n. For example, suppose that the pilot symbol is inserted with a unique interval N_{PI} and that the time index of the pilot symbol is given as $N_{\mathrm{PI}} \times k$, then we have $I_{\mathrm{p}}(n,k) = n + kN_{\mathrm{PI}}$. Hence, (8) is the ratio of $\hat{x}(n)$ and $f_{\mathrm{Dx}}(u(n))$, averaged over a certain window size of $2N_{\mathrm{PA}}+1$. The notation $u(I_{\mathrm{p}}(n,k))$ denotes for the transmitted pilot symbol corresponding to the equalizer output $\hat{x}(I_{\mathrm{p}}(n,k))$. Note that the estimated carrier phase error $\phi(n)$ of (8) remains unchanged when the index n is between the two closest pilot symbols. This implies that the estimated carrier phase error $\phi(n)$ is updated every pilot interval. Similar to (8), the carrier phase error of the y-polarization $\psi(n)$ can be estimated. A symbol sequence typically has a frame structure where a synchronization sequence is inserted at the beginning of each frame. With the differential detection algorithm, the synchronization sequence is identified before the CPR block.

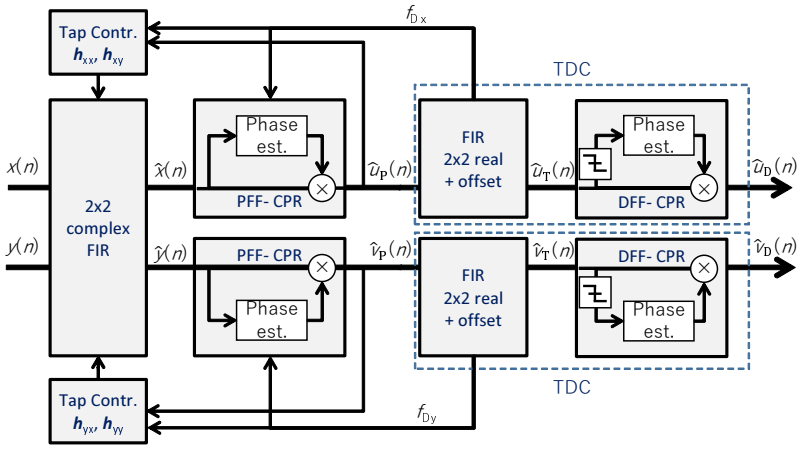

Fig. 2. Configuration of the proposed receiver architecture with transmitter-side distortion compensation.

This enables frame synchronization and the pilot position $I_{\mathrm{p}}(n,k)$ is identified.

Moreover, in the PFF-CPR block, the carrier phase rotation between two pilot symbols is output for each polarization symbol, according to the linear interpolation of two carrier phase errors estimated at the two pilot symbol positions as follows:

$$\hat{u}_{\mathrm{P}}(n) = \hat{x}(n)\exp\left[-jF_{\mathrm{I}}\big(n,\phi(n),\tilde{\phi}(n)\big)\right] \tag{9}$$

$$\hat{v}_{\mathrm{P}}(n) = \hat{y}(n)\exp\left[-jF_{\mathrm{I}}\big(n,\psi(n),\tilde{\psi}(n)\big)\right], \tag{10}$$

where $F_{\mathrm{I}}\big(n,\phi(n),\tilde{\phi}(n)\big)$ denotes the carrier phase error at symbol index n based on interpolation, and $\tilde{\phi}(n)$ and $\tilde{\psi}(n)$ denote the phase errors of the x- and y-polarizations, respectively, which are estimated in the previous iteration.[1] The above-mentioned method does not depend on the modulation format for pilot symbols. Then, the signal output from the PFF-CPR block is sent to the TDC block of Fig. 2.

C. Adaptive Tap Control Block

In the adaptive tap control block of Fig. 2, the FIR coefficients of our adaptive equalizer are calculated by the decision-directed LMS algorithm with the recovered data symbols when fast-tracking is required. Let us introduce square error loss functions of x- and y-polarizations, respectively, as follows:

$$J_x = \sum_n \left|f_{\mathrm{Dx}}(u(n)) - \hat{u}_{\mathrm{P}}(n)\right|^2 \tag{11}$$

$$J_y = \sum_n \left|f_{\mathrm{Dy}}(v(n)) - \hat{v}_{\mathrm{P}}(n)\right|^2, \tag{12}$$

where $u(n)$ and $v(n)$ denote the transmitted pilot symbols and the generated replica of data symbols. Then, the FIR coefficients of $\boldsymbol{h}_{xx}(n)$, $\boldsymbol{h}_{xy}(n)$, $\boldsymbol{h}_{yx}(n)$, and $\boldsymbol{h}_{yy}(n)$ are updated employing the steepest gradient descent method [42], [43] as follows:

$$\begin{aligned}\boldsymbol{h}_{xx}(n+1) &= \boldsymbol{h}_{xx}(n) + \mu\big\{f_{\mathrm{Dx}}(u(n-N_{\mathrm{D}})) - \hat{u}_{\mathrm{P}}(n-N_{\mathrm{D}})\big\} \\ &\times \exp\left[jF_{\mathrm{I}}\big(n-N_{\mathrm{D}},\phi(n-N_{\mathrm{D}}),\tilde{\phi}(n-N_{\mathrm{D}})\big)\right]\boldsymbol{x}^*(n-N_{\mathrm{D}})\end{aligned} \tag{13}$$

$$\begin{aligned}\boldsymbol{h}_{xy}(n+1) &= \boldsymbol{h}_{xy}(n) + \mu\big\{f_{\mathrm{Dx}}(u(n-N_{\mathrm{D}})) - \hat{u}_{\mathrm{P}}(n-N_{\mathrm{D}})\big\} \\ &\times \exp\left[jF_{\mathrm{I}}\big(n-N_{\mathrm{D}},\phi(n-N_{\mathrm{D}}),\tilde{\phi}(n-N_{\mathrm{D}})\big)\right]\boldsymbol{y}^*(n-N_{\mathrm{D}})\end{aligned} \tag{14}$$

$$\begin{aligned}\boldsymbol{h}_{yx}(n+1) &= \boldsymbol{h}_{yx}(n) + \mu\big\{f_{\mathrm{Dy}}(v(n-N_{\mathrm{D}})) - \hat{v}_{\mathrm{P}}(n-N_{\mathrm{D}})\big\} \\ &\times \exp\left[jF_{\mathrm{I}}\big(n-N_{\mathrm{D}},\psi(n-N_{\mathrm{D}}),\tilde{\psi}(n-N_{\mathrm{D}})\big)\right]\boldsymbol{x}^*(n-N_{\mathrm{D}})\end{aligned} \tag{15}$$

$$\begin{aligned}\boldsymbol{h}_{yy}(n+1) &= \boldsymbol{h}_{yy}(n) + \mu\big\{f_{\mathrm{Dy}}(v(n-N_{\mathrm{D}})) - \hat{v}_{\mathrm{P}}(n-N_{\mathrm{D}})\big\} \\ &\times \exp\left[jF_{\mathrm{I}}\big(n-N_{\mathrm{D}},\psi(n-N_{\mathrm{D}}),\tilde{\psi}(n-N_{\mathrm{D}})\big)\right]\boldsymbol{y}^*(n-N_{\mathrm{D}})\end{aligned} \tag{16}$$

where μ and N_{D} denote the step size of the LMS algorithm and the number of delay taps in the feedback control loop, respectively. Note that (13), (14) are derived from (11) while (15), (16) are derived from (12).

D. TDC Filter Block

The TDC filter block consists of the 2×2 real-valued FIR filter and the DC offset adder. The 2×2 real-valued FIR filter allows us to mitigate the IQ skew, the IQ amplitude imbalance, the 90-degree error, and the cross-talk between I- and Q-lanes.

Let us introduce a complex-valued vector with length $2L_{\mathrm{T}}+1$ as

$$\hat{\boldsymbol{u}}_{\mathrm{P}}(n) = [\hat{u}_{\mathrm{P}}(n-L_{\mathrm{T}}),...,\hat{u}_{\mathrm{P}}(n),...,\hat{u}_{\mathrm{P}}(n+L_{\mathrm{T}})]^T \in \mathbb{C}^{2L_{\mathrm{T}}+1} \tag{17}$$

Then, the symbols output from the TDC filter are given by

$$\begin{aligned}\hat{u}_{\mathrm{T}}(n) &= \mathfrak{Re}\{\hat{\boldsymbol{u}}_{\mathrm{P}}(n)^T\}\{\boldsymbol{w}_{\mathrm{II}} + j\boldsymbol{w}_{\mathrm{QI}}\} \\ &+ \mathfrak{Im}\{\hat{\boldsymbol{u}}_{\mathrm{P}}(n)^T\}\{\boldsymbol{w}_{\mathrm{IQ}} + j\boldsymbol{w}_{\mathrm{QQ}}\} + w_{\mathrm{I0}} + jw_{\mathrm{Q0}}\end{aligned} \tag{18}$$

where $\boldsymbol{w}_{\mathrm{II}}$, $\boldsymbol{w}_{\mathrm{IQ}}$, $\boldsymbol{w}_{\mathrm{QI}}$, and $\boldsymbol{w}_{\mathrm{QQ}}$ are the coefficient vectors of 2×2 real-valued FIR filter having the length of $2L_{\mathrm{TA}}+1$, while w_{I0} and w_{Q0} are the DC offset coefficients of I- and Q-lanes. Also, (18) corresponds to the process of x-polarization, and that of y-polarization is given by the symmetric copy of (18), hence omitted in this paper.

[1]Note that while in this paper, we employ a simple linear interpolation function, it may be readily possible to rely on a higher-order one. For example, we may employ $F_{\mathrm{I}}\big(n,\phi(n),\tilde{\phi}(n)\big) = (1-\frac{n'}{N_{\mathrm{PI}}})\tilde{\phi}(n) + \frac{n'}{N_{\mathrm{PI}}}\phi(n)$, where $n' = \mathrm{mod}(n, N_{\mathrm{PI}})$.

The distortion function $f_{\rm D}(u)$ is formulated as follows:

$$f_{\rm D}(u) = \mathfrak{Re}\{u\}(\tilde{w}_{\rm QQ} - j\tilde{w}_{\rm QI}) + \mathfrak{Im}\{u\}(-\tilde{w}_{\rm IQ} + j\tilde{w}_{\rm II}) - w_{\rm I0} - jw_{\rm Q0}, \quad (19)$$

where $\tilde{w}_{\rm II}$, $\tilde{w}_{\rm IQ}$, $\tilde{w}_{\rm QI}$, and $\tilde{w}_{\rm QQ}$ denote the middle taps of 2×2 real-valued FIR filter of the TDC block, and amplitude scaling is omitted. The distortion function is formulated as an inverse function of the TDC filter, where only the center taps are used. The signals output from the TDC block, i.e., $\hat{u}_{\rm T}(n)$ and $\hat{v}_{\rm T}(n)$, are input into the decision-directed feed-forward (DFF) CPR block to compensate for the residual carrier phase errors. The LMS algorithm is used for adapting the TDC coefficients based on the output from the DFF-CPR block.

E. TDC DFF-CPR Block

In the TDC DFF-CPR block, symbol replicas of data symbols $\hat{u}_{\rm R}(n)$ are generated based on the hard decision of $\hat{u}_{\rm T}(n)$. Then, the residual carrier phase error at the index n is estimated as

$$\phi_{\rm D}(n) = \arg\left(\frac{1}{2N_{\rm DA}+1}\sum_{k=-N_{\rm DA}}^{N_{\rm DA}} \frac{\hat{u}_{\rm T}(n-k)}{\hat{u}_{\rm R}(n-k)}\right), \quad (20)$$

which is averaged over $2N_{\rm DA}+1$ symbol intervals to exclude the effects of the white noise, similar to (8). Then, the symbol output from the TDC filters $\hat{u}_{\rm T}(n)$ is compensated for by the averaged carrier phase error $\phi_{\rm D}(n)$. Finally, the symbols output from the DFF-CPR block are used for optimizing the FIR coefficients of the TDC block. The four FIR coefficient vectors and the two DC offset coefficients are updated as follows:

$$\boldsymbol{w}_{\rm II} \leftarrow \boldsymbol{w}_{\rm II} + \mu\mathfrak{Re}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\}\mathfrak{Re}\{\hat{\boldsymbol{u}}_{\rm P}(n)\} \quad (21)$$

$$\boldsymbol{w}_{\rm IQ} \leftarrow \boldsymbol{w}_{\rm IQ} + \mu\mathfrak{Re}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\}\mathfrak{Im}\{\hat{\boldsymbol{u}}_{\rm P}(n)\} \quad (22)$$

$$\boldsymbol{w}_{\rm QI} \leftarrow \boldsymbol{w}_{\rm QI} + \mu\mathfrak{Im}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\}\mathfrak{Re}\{\hat{\boldsymbol{u}}_{\rm P}(n)\} \quad (23)$$

$$\boldsymbol{w}_{\rm QQ} \leftarrow \boldsymbol{w}_{\rm QQ} + \mu\mathfrak{Im}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\}\mathfrak{Im}\{\hat{\boldsymbol{u}}_{\rm P}(n)\} \quad (24)$$

$$w_{\rm I0} \leftarrow w_{\rm I0} + \mu\mathfrak{Re}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\} \quad (25)$$

$$w_{\rm Q0} \leftarrow w_{\rm Q0} + \mu\mathfrak{Im}\{(u(n)-\hat{u}_{\rm D}(n))e^{-j\phi_{\rm D}(n)}\} \quad (26)$$

Typically, the distortion condition of the transmitter-side components changes over time at least a couple of orders of magnitude slower than a symbol rate, and hence the requirement to suppress the feedback loop delay in the LMS algorithm for the FIR coefficients of the TDC block is lower than that imposed on the adaptive channel equalizer.

To elaborate a little further, one of our main contributions is on the feedback from the TDC block to the 2×2 complex-valued FIR equalizer and the PFF-CPR, which is characterized by the distortion function $f_{\rm D}(u)$ of (19). More specifically, $f_{\rm D}(u)$ reproduces the distortion of the transmitter components for the pilot symbols or the replica of data symbols, and modifies it from the ideal mapping position. This allows us to dispense with the explicit distortion compensation at the channel equalizer and the PFF-CPR blocks. Instead, the distortion is compensated for at the TDC block.

A further note is that as shown in (19), the effects of ISI caused by transmitter-side distortion are ignored, and the distortion-reproduced symbol $f_{\rm D}(u)$ is obtained without using the information of adjacent symbols. This feedback specific to our scheme significantly increases the resilience against the transmitter-side component distortion, which is evaluated in our numerical simulations of Section V.

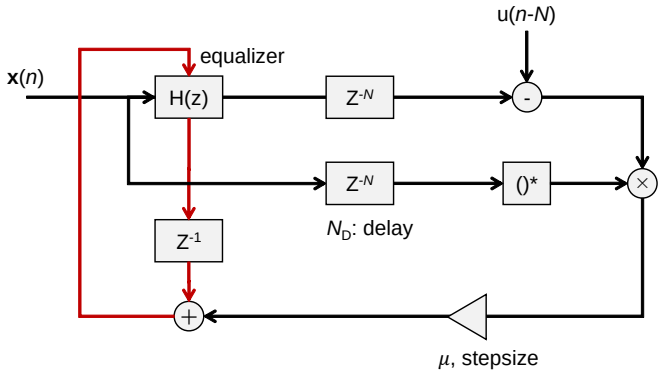

Fig. 3. The LMS equalizer model with the feedback loop delay.

IV. Stable Condition of LMS Equalizer in the Presence of Feedback Delay

In this section, we derive the stability condition of the LMS algorithm under the presence of the feedback loop delay. Note that the LMS step size that satisfies the stable condition decreases upon increasing the feedback loop delay. Fig. 3 shows the LMS equalizer model to consider the stability condition of the FIR filter coefficients. The input signal to the LMS equalizer $\boldsymbol{x}(n)$ is represented in a vector form as

$$\boldsymbol{x}(n) = [x(n-L), ..., x(n), ..., x(n+L)]^T \in \mathbb{C}^{2L+1}, \quad (27)$$

where n is the time index in a symbol interval, and the filter length is given by $2L+1$. [2] The input signal is convolved by the FIR filter coefficients $\boldsymbol{h}(n) \in \mathbb{C}^{2L+1}$ to output the equalized symbol $\boldsymbol{x}^T(n)\boldsymbol{h}(n)$. Then, the equalized symbol is used for updating the FIR filter coefficients with the aid of the LMS algorithm after adding $N_{\rm D}$-tap delay, where $N_{\rm D}$ represents the total delay in the feedback control loop.

The carrier phase recovery (CPR) has two main delay factors, i.e., the signal path imposed by averaging a sufficient number of symbols to estimate a stable carrier phase and the excess delay caused due to parallel-processing of multiple symbols per internal clock period. These delay factors caused by the CPR are taken into account in our scheme by the delay tap $N_{\rm D}$.

In the LMS algorithm, the FIR-filter coefficients $\boldsymbol{h}(n)$ are updated as follows:

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu\{u(n-N_{\rm D}) - \boldsymbol{x}^T(n-N_{\rm D})\boldsymbol{h}(n-N_{\rm D})\}\boldsymbol{x}^*(n-N_{\rm D}) \quad (28)$$

$$= \boldsymbol{h}(n) + \mu\{u(n-N_{\rm D})\boldsymbol{x}^*(n-N_{\rm D}) - \boldsymbol{x}^*(n-N_{\rm D})\boldsymbol{x}^T(n-N_{\rm D})\boldsymbol{h}(n-N_{\rm D})\}, \quad (29)$$

[2]Here, we assume a discrete time frame of a symbol interval for simplicity. However, our analysis is readily applicable to arbitrary over-sampling scenarios, as mentioned later in this paper.

where μ is the step size, while $u(n)$ is the nth information symbol that corresponds to the equalizer output $\boldsymbol{x}^T(n)\boldsymbol{h}(n)$. In our decision-directed LMS algorithm, the recovered information symbol $\hat{u}(n)$ is used as $u(n)$ of (29).

Since the time-varying term $\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)$ of (29) is hard to exactly obtain, we approximate it by a covariance matrix of $\boldsymbol{R}_x$=$\mathbb{E}_n[\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)]$. This approximation is reasonable when the step size μ is significantly low, i.e. $\mu \ll 1$, where the tap coefficient $\boldsymbol{h}(n)$ evolves slowly enough that the term $\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)$ can be approximated by its ensemble on n. Hence, (29) is further modified to

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu\big\{u(n-N_{\rm D})\boldsymbol{x}^*(n-N_{\rm D}) - \boldsymbol{R}_x\boldsymbol{h}(n-N_{\rm D})\big\}. \tag{30}$$

Let us introduce the z-transform of a correlation vector $u(n)\boldsymbol{x}^*(n)$ and that of the FIR coefficient vector $\boldsymbol{h}(n)$, respectively, as

$$\boldsymbol{C}(z) = \sum_{n=-\infty}^{\infty} z^{-n}u(n)\boldsymbol{x}^*(n) \tag{31}$$

$$\boldsymbol{H}(z) = \sum_{n=-\infty}^{\infty} z^{-n}\boldsymbol{h}(n) \tag{32}$$

With (31), (32), the z-transform of (30) is represented by

$$z\boldsymbol{H}(z) = \boldsymbol{H}(z) + \mu\{z^{-N_{\rm D}}\boldsymbol{C}(z) - z^{-N_{\rm D}}\boldsymbol{R}_x\boldsymbol{H}(z)\}. \tag{33}$$

Then, (33) is rewritten by

$$\boldsymbol{H}(z) = [(1-z^{-1})\boldsymbol{I} + z^{-N_{\rm D}-1}\mu\boldsymbol{R}_x]^{-1}[\mu z^{-N_{\rm D}-1}\boldsymbol{C}(z)] \tag{34}$$

Recalling that $\boldsymbol{R}_x$ is an Hermitian matrix, it is diagonalized with the aid of eigenvalue decomposition as follows: $\boldsymbol{R}_x = \boldsymbol{Q}^\dagger\boldsymbol{\Lambda}_x\boldsymbol{Q}$, where $\boldsymbol{Q}$ is a unitary matrix, having the relationship of $\boldsymbol{Q}^\dagger\boldsymbol{Q}$=$\boldsymbol{Q}\boldsymbol{Q}^\dagger$=$\boldsymbol{I}$, and $\boldsymbol{\Lambda}_x$ is a diagonal matrix composed of eigenvalues. Furthermore, $\boldsymbol{I}$ is the identity matrix. By applying this diagonalization to (34), we obtain

$$\begin{aligned}\boldsymbol{H}(z) &= [(1-z^{-1})\boldsymbol{I} + z^{-N_{\rm D}-1}\mu\boldsymbol{Q}^\dagger\boldsymbol{\Lambda}_x\boldsymbol{Q}]^{-1}[\mu z^{-N_{\rm D}-1}\boldsymbol{C}(z)]\\ &= [\boldsymbol{Q}^\dagger\{(1-z^{-1})\boldsymbol{I} + z^{-N_{\rm D}-1}\mu\boldsymbol{\Lambda}_x\}\boldsymbol{Q}]^{-1}[\mu z^{-N_{\rm D}-1}\boldsymbol{C}(z)]\\ &= \boldsymbol{Q}^\dagger[(1-z^{-1})\boldsymbol{I} + z^{-N_{\rm D}-1}\mu\boldsymbol{\Lambda}_x]^{-1}\boldsymbol{Q}[\mu z^{-N_{\rm D}-1}\boldsymbol{C}(z)].\end{aligned} \tag{35}$$

Since the z-transform $\boldsymbol{H}$(z) has to satisfy that all the poles are inside a unit circle, we arrive at the stable condition for feedback loop control of FIR coefficients $\boldsymbol{h}(n)$ as follows:

$$(1-z^{-1})+z^{-N_{\rm D}-1}\mu\lambda_x(k) = 0 \quad (k = 0, ..., 2L+1), \tag{36}$$

where $\lambda_x(k)$ denotes the kth diagonal component of $\boldsymbol{\Lambda}_x$. Finally, the condition that solution of (36) exists within a unit circle is equivalently given by [44]

$$\mu < \frac{2}{\max_k \lambda_x(k)} \sin\frac{\pi}{2(2N_{\rm D}+1)}. \tag{37}$$

Here, the highest eigenvalue of the covariance matrix $\boldsymbol{R}_x$ and the delay in the feedback control loop $N_{\rm D}$ determine the upper limit of the LMS step size μ for stable operation.

Note that the derivation of (37) may be generalized for an arbitrary oversampling rate by modifying the input signal vector $\boldsymbol{x}(n)$ to that supporting all input samples that are convoluted by the FIR coefficient vector.

V. Performance Results

In this section, we provide our performance results to evaluate the adaptive equalizer with the proposed TDC block and investigate the effects of the delay induced in the feedback control loop. In the simulations, we employed 64-QAM, i.e., $M = 6$ bits/symbol. The pilot symbol was mapped at one of four points $(\pm 3/8, \pm 3/8)$ in 64QAM constellation, which has the lower average power as the data symbol, and the pilot interval was set to $N_{\rm PI} = 32$.

Fig. 4 shows the Q-factor of the proposed and benchmark schemes, which is calculated as a normal distribution variable that gives a complementary cumulative probability equal to the observed BER, where we have the relationship of BER= $\int_Q^\infty \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})\,dx$. The step size of the decision-directed LMS equalizer was set to $\mu = 10^{-3}$. The delay parameter was set at $N_{\rm D} = 256$ symbols. The two green curves represent the scheme that deactivates our TDC block while employing the PFF-CPR in the presence of the PN. The two red curves correspond to the proposed schemes activating the TDC and PFF-CPR blocks. For the green and red dotted curves, the nested MZ modulator is activated, while the transmitter distortion is set negligible, i.e. $R_{\rm EI} = 99$ dB, $R_{\rm EO} = 99$ dB, and the 90-degree phase error of $\theta_e = 0$ deg. For the green and red solid curves, the transmitter distortion conditions of $R_{\rm EI} = 25$ dB, $R_{\rm EO} = 30$ dB, and $\theta_e = 5$ deg were considered. We also plotted the theoretical upper bound, the MMSE bound, and the ideal LMS bound. In the ideal LMS bound, the effects of the PN contamination were not considered, and the nested MZ modulator was deactivated. Note that while the MMSE equalizer exhibits theory-achieving performance in the static condition, it is not suitable in a rapidly time-varying channel.

The proposed scheme is evaluated in terms of the required SNR (RSNR) gap from the theoretical bound, where the RSNR is defined by a minimum SNR to achieve the BER of 1.0×10^{-2}. In the rest of this paper, the gap from the theoretical bound is referred to as *the RSNR gap*. Observe in Fig. 4 that the theoretical and the MMSE bounds nearly coincided. The RSNR gap of the LMS bound was less than 0.1 dB. When the PNs and the nested MZ modulator were considered, the RSNR gap increased to 0.9 dB. The laser PN linewidth was set at 200 kHz in the rest of this paper. This penalty includes the effects of the nonlinear characteristic of a sinusoidal transfer function in the inner MZI, despite the idealistic scenario of $R_{\rm EI} = R_{\rm EO} = 99$ dB and $\theta_e = 0$ deg., corresponding to the negligible imperfection condition. When the TDC block was activated in the proposed scheme with the optimum parameter $N_{\rm PA} = 7$, the RSNR gap was reduced to 0.6 dB, which is achieved as the explicit benefits of the DFF-CPR in our TDC block.

When the transmitter distortion was loaded with $R_{\rm EI} = 25$ dB, $R_{\rm EO} = 30$ dB, and $\theta_e = 5$ deg. in the benchmark scheme, the RSNR was not attained since the BER was higher

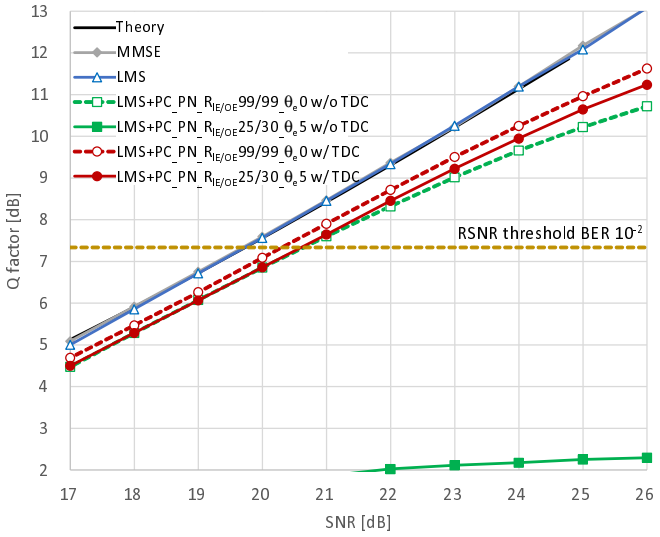

Fig. 4. Q-factor calculated from the BERs of the proposed scheme.

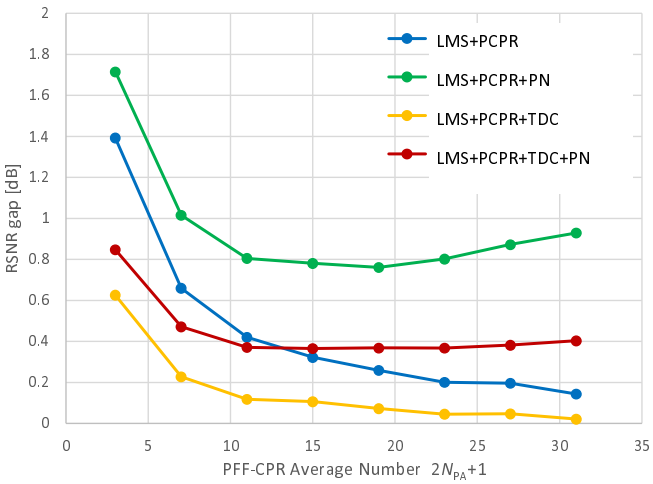

Fig. 5. The effects of window size $2N_{\mathrm{PA}}+1$ of the proposed scheme on the RSNR gap, where the four scenarios were considered, i.e., those with/without our TDC block in the presence/absence of the PN.

than 10^{-2}. By contrast, in the proposed scheme activating the TDC block with the distortion function, the BER was sufficiently low, and the RSNR gap was as low as 0.9 dB, even in the presence of the high transmitter distortion.

Fig. 5 investigates the effects of window size in the PFF-CPR N_{PA} on the RSNR gap of the proposed and benchmark schemes, where the nested MZ modulator was deactivated for simplicity, i.e., any transmitter-side impairments are not considered. We considered the four scenarios, i.e., 1) the deactivated TDC block without PN, 2) the deactivated TDC block with PN, 3) the activated TDC block without PN, and 4) the activated TDC block with PN. When the effects of PN were ignored in the scheme with and without our TDC block (i.e., the yellow and the blue curves), the RSNR gap monotonically decreased upon increasing N_{PA} and converged to the RSNR gap of 0 dB. This is because the PFF-CPR suffers from the estimated carrier phase error induced due to the AWGN when N_{PA} is small. For a high N_{PA}, such as $N_{\mathrm{PA}} > 5$, the estimated PN error was maintained to be sufficiently low for the ideal scenario of no PN. By contrast, when the effects of the PN were included in the benchmark scheme without our TDC block (corresponding to the green curve), the PFF-CPR, having significantly high N_{PA}, did not track the PN variation, where the optimum N_{PA} was $N_{\mathrm{PA}} = 9$, and the associated RSNR gap was as high as 0.76 dB. Furthermore, in the proposed scheme, including our TDC block (corresponding to the red curve), sufficiently suppressed the PN, where the RSNR gap was as low as 0.36 dB for the optimal window size of $N_{\mathrm{PA}} = 7$.

Fig. 6 shows the effects of the 90-degree phase error on the RSNR penalty, where the imperfection of the inner MZI was given by $R_{\mathrm{EI}} = 99$ dB, 30 dB, and 25 dB. For comparisons, we considered the three schemes, such as the proposed scheme with the TDC block and the DF, the proposed scheme with the TDC block without the DF, and the benchmark without either our TDC block or DF. The inset of Fig. 6 shows the Q-factor of the three schemes with $R_{\mathrm{EI}} = 25$ dB and the 90-degree phase error of 0 deg.

When R_{EI} and the 90-degree phase error were set at 99 dB and 0 deg., corresponding to no imperfection case, the RSNR gap was as low as 0.45 dB when the TDC block and the DF were deactivated. In the proposed scheme with the TDC block and the DF, the RSNR gap remains similarly low as 0.30 dB. In the benchmark without either our TDC block or DF, when the R_{EI} was set at 30 dB, the RSNR gap increased to more than 7 dB. When R_{EI} was further decreased to 25 dB, the BER floor was higher than 1.0×10^{-2}, as shown in the inset, and hence the associated RSNR cannot be calculated in our definition. By contrast, the RSNR gap of the scheme with the TDC blocks and the DF decreased to 0.35 dB and 0.45 dB, when the R_{EI} was set at 30 and 25 dB, respectively. When the 90-degree phase error was increased up to 5 deg., the RSNR gap increase was below 0.05 dB, which was common for $R_{\mathrm{EI}} = 99$, 30, and 25 dB. Furthermore, by activating the DF in the proposed scheme with our TDC block, the RSNR gap decreased by more than 1.0 dB for $R_{\mathrm{EI}} = 30$ dB. For $R_{\mathrm{EI}} = 25$ dB, the DF has the benefits of reducing the BER floor and improving the Q factor from 5.2 dB to 10.6 dB at the SNR of 24 dB, as shown in the inset of Fig. 6. This indicates that the gain achievable by the DF is significantly high.

We evaluated the achievable performance of the conventional scheme where the TDC FIR is inside the feedback control loop of the adaptive equalizer, and the tap coefficients are controlled by using $\hat{u}_{\mathrm{T}}(n)$. The RSNR gap of the conventional scheme for $R_{\mathrm{EI}} = 30$ dB and 25 dB were 0.32 dB and 0.43 dB, respectively. The difference in the RSNR gap was as low as 0.02–0.03 dB.

Fig. 7 shows the RSNR gap from the theoretical bound for two cases, with and without the proposed scheme, where the IQ skew at the transmitter was varied. When the IQ skew was set at 0.08 per symbol, the RSNR gap penalty decreased from 2.7 dB to 0.5 dB. The proposed scheme improves tolerance to IQ skew significantly. The RSNR penalty increase of 0.3 dB was observed with the TDC case at the IQ skew of 0.1 per symbol, which corresponds to 1 ps for the symbol rate

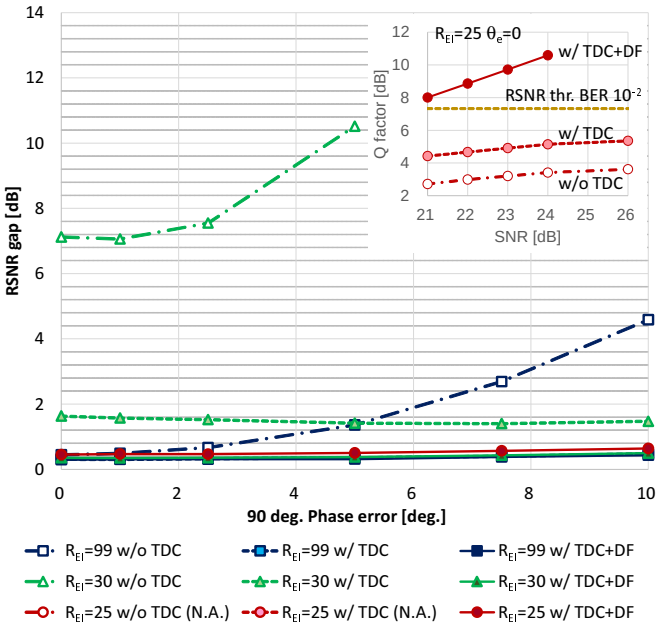

Fig. 6. RSNR gap of the proposed scheme, where imperfection of the extinction ratio of the inner MZ R_{EI} and the 90-degree phase error θ_{e} were considered. When R_{EI} was set at 25 dB, the BER floor was higher than 1.0×10^{-2} for the benchmark scenario without TDC and that without DF, hence the associated RSNR cannot be calculated in our definition, and the associated cases are shown by N.A. (not available).

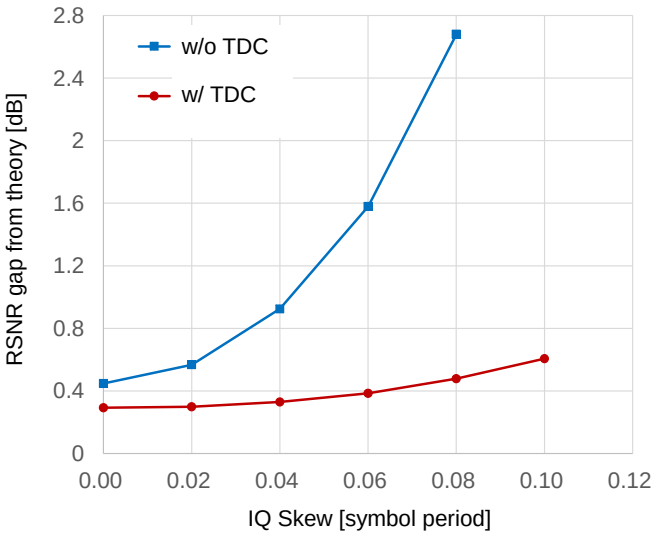

Fig. 7. The RSNR gap of the proposed scheme with the varied IQ skew.

of 100 GBaud. Since we assume that most of the transmitter IQ skew is compensated in the digital pre-distortion, a 10% range of symbol period is sufficient to tolerate the residual IQ skew caused by the variation in time and calibration error of the digital pre-distortion scheme.

Fig. 8 shows the proposed scheme's RSNR gap from the theoretical bound for BER of 10^{-2}, where we varied the speed of polarization scrambling normalized by the symbol rate from $2\pi \times 10^{-8}$ to $4\pi \times 10^{-5}$ rad/symbol. For example, when the normalized speed of polarization was set at $\pi \times 10^{-5}$, the

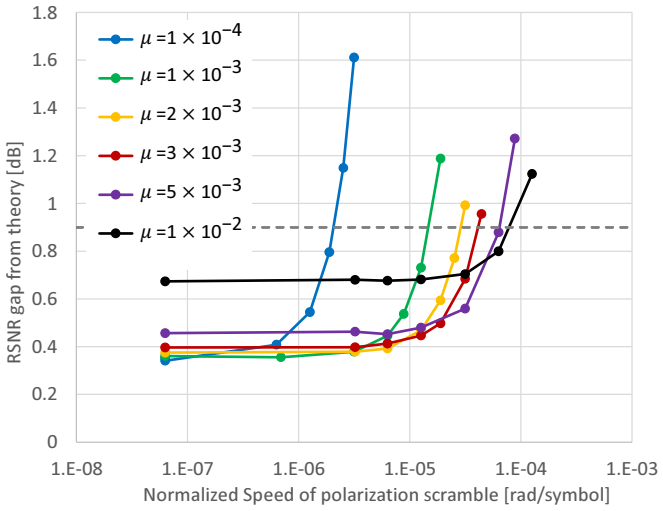

Fig. 8. The effects of the speed of polarization scramble on the RSNR gap for the different LMS step sizes.

state of signal polarization is rotated such that the x- and y-polarizations are gradually rotated and completely exchanged after 10^5 symbols, which corresponds to the rotation in the S_1–S_2 plane of Pointcare sphere [45]. Here, we considered the effects of the PN while ignoring those of the nested MZ modulator and transmitter imperfection. Also, both of the PFF-CPR and the TDC blocks were activated, while setting $N_{\mathrm{PA}} = 7$. The feedback delay in the adaptive equalizer was set to eight pilot-symbol intervals, which corresponded to the total delay of $N_{\mathrm{D}} = 8 \times 32$ symbol intervals. Observe in Fig. 8 that in each step size, the RSNR gap significantly increased when the normalized speed of the polarization scramble was higher than a specific threshold. Assuming that the RSNR penalty increase of 0.5 dB due to the polarization scramble is acceptable, the maximum tracking speed is given by a crossing point with the RSNR gap of 0.9 dB, as shown by the broken line in Fig. 8. For $\mu = 10^{-4}$, the maximum tracking speed was as low as 2.0×10^{-6} rad/symbol. When the step size was increased to $\mu = 10^{-3}$ and 10^{-2}, the maximum tracking speed were more than 1.5×10^{-6} and 1.8×10^{-5}, respectively. However, for the step size as high as $\mu = 10^{-2}$, the RSNR gap increased by 0.3 dB from the case of $\mu = 10^{-3}$. Hence, there is a tradeoff between the maximum tracking speed and the RSNR gap. Note that the step size of 10^{-3} was employed in our evaluations except for Figs. 8, 9, and 10, since the step size of 10^{-3} is the optimum value to track the scramble speed of 10^{-5} rad/symbol, maintaining the RSNR penalty sufficiently low in the stable condition.

Fig 9 shows the maximum tracking speed as a function N_{D}, when the RSNR gap increase of 0.5 dB is accepted. Upon increasing N_{D}, the maximum tracking speed decreased. Additional blocks in the feedback control loop tend to increase the feedback latency. Assuming the symbol rate of 100 GBaud, the operating clock frequency of the CMOS digital logic circuit is less than 1GHz, and the parallel expansion width is more than 100 symbols. Thus, for example, when the feedback control loop delay increases from 5 to 10 in the operating clock

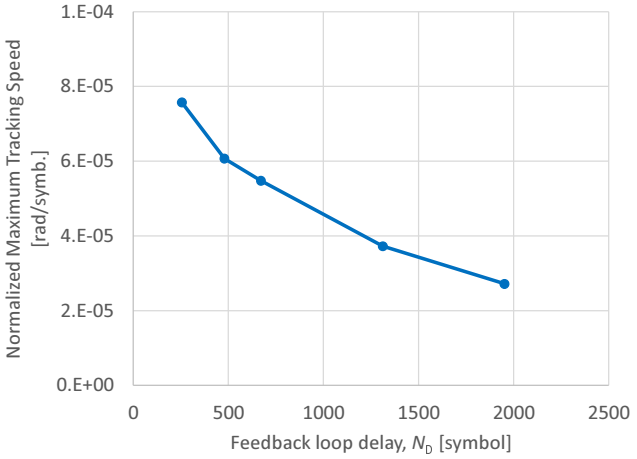

Fig. 9. The normalized maximum tracking speed for different feedback loop delay N_{D}.

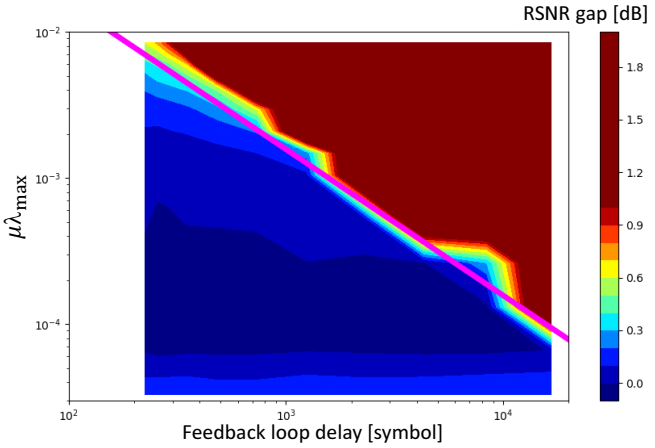

Fig. 10. Relationship between $\mu\lambda_{\max}$ and the delay in the feedback loop.

cycle, then N_{D} increases from 500 symbols to 1000 symbols, which degrades the maximum tracking speed from 6×10^{-6} to 4.5×10^{-6} rad/symbol based on Fig. 9. This corresponds to the decrease in the maximum tracking speed from 6 to 4.5 Mrad/second in the case of the 100 GBaud symbol rate. Consequently, our proposed architecture excluding the TDC block from the feedback control loop benefits from maintaining a fast-tracking feature of the adaptive equalizer.

Fig. 10 shows the contour plot of the RSNR gap, where the delay in the feedback control loop and the step size μ were varied. According to (37), the maximum step size that maintains the stable condition is determined by the delay in the feedback control loop,[3] while it is also the essential limiting factor of the maximum tracking speed. The vertical axis of Fig. 10 represents the product of the step size μ and the maximum eigenvalue $\lambda_{\max}$ of the covariance matrix $\boldsymbol{R}_x$ of the input signal vector $\boldsymbol{x}(n)$ as defined in Section IV. The red curve indicates the theoretical upper bound of (37). As shown in Fig. 10, the RSNR gap steeply increased when $\mu\lambda_{\max}$ became higher than the theoretical upper bound. Thus, it was confirmed that the theoretical bound is beneficial for striking the relationship between the delay and the step size.

VI. Conclusion

We proposed the optical coherent receiver architecture capable of fast-tracking the signals in a rapidly time-varying fiber channel while compensating for the residual signal distortion induced by the transmitter analog components. The TDC block in the proposed architecture has the explicit benefits by modifying the replica symbols used in the LMS tap optimizer while dispensing the increase in the feedback loop delay of the adaptive channel equalizer. Our simulation results show that the proposed scheme significantly reduces the RSNR gap from the theoretical bound. The RSNR gap caused by the low extinction ratio of the inner MZI was reduced by more than 6.5 dB, exhibiting the RSNR gap of as low as 0.4 dB. Furthermore, our theoretical analysis allows us to strike the balance of the feedback loop delay in the adaptive equalizer and the stable LMS step size, hence maximizing the tracking speed. Importantly, our theoretical analysis confirms the explicit advantage of implementing our TDC block outside of the feedback control loop for the sake of suppressing the feedback delay.

[3]The maximum step size also depends on the maximum eigenvalue of the covariance matrix of the input signal. In Fig. 10, only the step size μ and the feedback loop delay N_{D} were varied and it was found that μ and N_{D} unaffected the input signal of the 2×2 complex-valued FIR. Hence, in this paper, the eigenvalues of the covariance matrix of the input signal are assumed to be fixed.

Acknowledgement

We would like to thank Mitsuteru Yoshida, and Kyo Minoguchi at NTT Network Innovation Labs for the valuable discussion.

References

[1] G. Raybon, A. Adamiecki, J. Cho, P. Winzer, A. Konczykowska, F. Jorge, J.-Y. Dupuy, M. Riet, B. Duval, K. Kim, S. Randel, D. Pilori, B. Guan, N. Fontaine, and E. C. Burrows, "Single-carrier all-ETDM 1.08-Terabit/s line rate PDM-64-QAM transmitter using a high-speed 3-bit multiplexing DAC," in *IEEE Photonics Conference (IPC)*, 2015, pp. 1–2.

[2] M. Nakamura, F. Hamaoka, M. Nagatani, H. Yamazaki, T. Kobayashi, A. Matsushita, S. Okamoto, H. Wakita, H. Nosaka, and Y. Miyamoto, "104 Tbps/carrier probabilistically shaped PDM-64QAM WDM transmission over 240 km based on electrical spectrum synthesis," in *Optical Fiber Communication Conference (OFC)*, 2019, p. M4I.4.

[3] A. Matsushita, M. Nakamura, S. Yamamoto, F. Hamaoka, and Y. Kisaka, "41-Tbps C-band WDM transmission with 10-bps/Hz spectral efficiency using 1-Tbps/λ signals," *J. Lightwave Technol.*, vol. 38, no. 11, pp. 2905–2911, June 2020.

[4] F. Buchali, V. Lauinger, M. Chagnon, K. Schuh, and V. Aref, "CMOS DAC supported 1.1 Tb/s/λ DWDM transmission at 9.8 bit/s/Hz over DCI distances," *Journal of Lightwave Technology*, vol. 39, no. 4, pp. 1171–1178, 2021.

[5] F. Buchali, K. Schuh, R. Dischler, M. Chagnon, V. Aref, H. Buelow, Q. Hu, F. Pulka, M. Frascolla, E. Alhammadi, A. Samhan, I. Younis, M. El-Zonkoli, and P. Winzer, "1.3-Tb/s single-channel and 50.8-Tb/s WDM transmission over field-deployed fiber," in *45th European Conference on Optical Communication (ECOC)*, 2019, pp. 1–4.

[6] X. Chen, S. Chandrasekhar, G. Raybon, S. Olsson, J. Cho, A. Adamiecki, and P. Winzer, "Generation and intradyne detection of single-wavelength 1.61-Tb/s using an all-electronic digital band interleaved transmitter," in *Optical Fiber Communication Conference*, 2018, p. Th4C.1.

[7] F. Buchali, V. Aref, M. Chagnon, R. Dischler, H. Hettrich, R. Schmid, and M. Moeller, "52.1 Tb/s C-band DCI transmission over DCI distances at 1.49 $\mathrm{Tb/s/\lambda}$," in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.

[8] H. Sun, M. Torbatian, M. Karimi, R. Maher, S. Thomson, M. Tehrani, Y. Gao, A. Kumpera, G. Soliman, A. Kakkar, M. Osman, Z. A. El-Sahn, C. Doggart, W. Hou, S. Sutarwala, Y. Wu, M. R. Chitgarha, V. Lal, H.-S. Tsai, S. Corzine, J. Zhang, J. Osenbach, S. Buggaveeti, Z. Morbi, M. I. Olmedo, I. Leung, X. Xu, P. Samra, V. Dominic, S. Sanders, M. Ziari, A. Napoli, B. Spinnler, K.-T. Wu, and P. Kandappan, "800G DSP ASIC design using probabilistic shaping and digital sub-carrier multiplexing," *J. Lightwave Technol.*, vol. 38, no. 17, pp. 4744–4756, Sep 2020.

[9] M. Nakamura, F. Hamaoka, H. Yamazaki, M. Nagatani, Y. Ogiso, H. Wakita, M. Ida, A. Matsushita, T. Kobayashi, H. Nosaka, and Y. Miyamoto, "1.3-Tbps/carrier net-rate signal transmission with 168-GBaud PDM PS-64QAM using analogue-multiplexer-integrated optical frontend module," in *45th European Conference on Optical Communication (ECOC)*, 2019, pp. 1–3.

[10] X. Chen, G. Raybon, D. Che, J. Cho, and K. Kim, "Transmission of 200-GBaud PDM probabilistically shaped 64-QAM signals modulated via a 100-GHz Thin-film LiNbO3 I/Q modulator," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, 2021, pp. 1–3.

[11] H. Hettrich, R. Schmid, L. Altenhain, J. Wrtele, and M. Mller, "A linear active combiner enabling an interleaved 200 GS/s DAC with 44 GHz analog bandwidth," in *IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, 2017, pp. 142–145.

[12] F. Pittal, M. Schaedler, G. Khanna, S. Calabr, M. Kuschnerov, C. Xie, Z. Ye, Q. Wang, and B. Zheng, "220 GBaud signal generation enabled by a two-channel 256 GSa/s arbitrary waveform generator and advanced DSP," in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.

[13] F. Buchali, M. Chagnon, K. Schuh, and V. Lauinger, "Beyond 100 GBaud transmission supported by a 120 GSA/S CMOS digital to analog converter," in *45th European Conference on Optical Communication (ECOC)*, 2019, pp. 1–4.

[14] M. Nakamura, F. Hamaoka, M. Nagatani, Y. Ogiso, H. Wakita, H. Yamazaki, T. Kobayashi, M. Ida, H. Nosaka, and Y. Miyamoto, "192-Gbaud signal generation using ultra-broadband optical frontend module integrated with bandwidth multiplexing function," in *Optical Fiber Communication Conference*, 2019, p. Th4B.4.

[15] X. Chen, J. Cho, G. Raybon, D. Che, K. Kim, E. Burrows, P. Kharel, C. Reimer, K. Luke, L. He, and M. Zhang, "Single-wavelength and single-photodiode 700 Gb/s entropy-loaded PS-256-QAM and 200-GBaud PS-PAM-16 transmission over 10-km SMF," in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–4.

[16] Y. Ogiso, J. Ozaki, Y. Ueda, N. Kashio, N. Kikuchi, E. Yamada, H. Tanobe, S. Kanazawa, H. Yamazaki, Y. Ohiso, T. Fujii, and M. Kohtoku, "Over 67 GHz bandwidth and 1.5 V $V\pi$ InP-based optical IQ modulator with n-i-p-n heterostructure," *Journal of Lightwave Technology*, vol. 35, no. 8, pp. 1450–1455, 2017.

[17] M. Zhang, C. Wang, X. Chen, M. Bertrand, A. Shams-Ansari, S. Chandrasekhar, P. Winzer, and M. Lončar, "Ultra-high bandwidth integrated lithium niobate modulators with record-low $V\pi$," in *Optical Fiber Communication Conference*, 2018, p. Th4A.5.

[18] Y. Ogiso, J. Ozaki, Y. Ueda, H. Wakita, M. Nagatani, H. Yamazaki, M. Nakamura, T. Kobayashi, S. Kanazawa, T. Fujii, Y. Hashizume, H. Tanobe, N. Nunoya, M. Ida, Y. Miyamoto, and M. Ishikawa, "Ultra-high bandwidth InP IQ modulator for beyond 100-GBd transmission," in *Optical Fiber Communication Conference (OFC) 2019*, 2019, p. M2F.2.

[19] G. Bcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4651–4665, 2015.

[20] P. Schulte and G. Bcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2016.

[21] T. Yoshida, M. Karlsson, and E. Agrell, "Short-block-length shaping by simple mark ratio controllers for granular and wide-range spectral efficiencies," in *2017 European Conference on Optical Communication (ECOC)*, 2017, pp. 1–3.

[22] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front end for optical quadrature amplitude modulation," *IEEE Photonics Journal*, vol. 5, no. 2, pp. 7 800 110–7 800 110, 2013.

[23] M. Paskov, D. Lavery, and S. J. Savory, "Blind equalization of receiver in-phase/quadrature skew in the presence of Nyquist filtering," *IEEE Photonics Technology Letters*, vol. 25, no. 24, pp. 2446–2449, 2013.

[24] R. Rios-Mller, J. Renaudier, and G. Charlet, "Blind receiver skew compensation for long-haul non-dispersion managed systems," in *European Conference on Optical Communication (ECOC)*, 2014, pp. 1–3.

[25] V. Bajaj, F. Buchali, M. Chagnon, S. Wahls, and V. Aref, "Deep neural network-based digital pre-distortion for high baudrate optical coherent transmission," *Journal of Lightwave Technology*, vol. 40, no. 3, pp. 597–606, 2022.

[26] M. Sena, M. S. Erkilinc, T. Dippon, B. Shariati, R. Emmerich, J. K. Fischer, and R. Freund, "Bayesian optimization for nonlinear system identification and pre-distortion in cognitive transmitters," *Journal of Lightwave Technology*, vol. 39, no. 15, pp. 5008–5020, 2021.

[27] T. Sasai, M. Nakamura, E. Yamazaki, A. Matsushita, S. Okamoto, K. Horikoshi, and Y. Kisaka, "Wiener-Hammerstein model and its learning for nonlinear digital pre-distortion of optical transmitters," *Opt. Express*, vol. 28, no. 21, pp. 30 952–30 963, Oct 2020.

[28] P. W. Berenguer, M. Nlle, L. Molle, T. Raman, A. Napoli, C. Schubert, and J. K. Fischer, "Nonlinear digital pre-distortion of transmitter components," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1739–1745, 2016.

[29] C. R. S. Fludger and T. Kupfer, "Transmitter impairment mitigation and monitoring for high baud-rate, high order modulation systems," in *42nd European Conference on Optical Communication*, 2016, pp. 1–3.

[30] G. Bosco, S. M. Bilal, A. Nespola, P. Poggiolini, and F. Forghieri, "Impact of the transmitter iq-skew in multi-subcarrier coherent optical systems," in *Optical Fiber Communication Conference*, 2016, p. W4A.5.

[31] T. Kobayashi, M. Nakamura, F. Hamaoka, M. Nagatani, H. Wakita, H. Yamazaki, T. Umeki, H. Nosaka, and Y. Miyamoto, "35-Tb/s C-band transmission over 800 km employing 1-Tb/s PS-64QAM signals enhanced by complex 8×2 MIMO equalizer," in *Optical Fiber Communication Conference*, 2019, p. Th4B.2.

[32] E. Yamazaki, N. Farsad, and A. Goldsmith, "Low noise non-linear equalization using neural networks and belief propagation," 2019.

[33] P. M. Krummrich and K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," in *Optical Fiber Communication Conference*, 2004, p. FI3.

[34] P. M. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus, "Field trial results on statistics of fast polarization changes in long haul WDM transmission systems," in *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference*, 2005, p. OThT6.

[35] T. Matsuda, T. Kawasaki, T. Kotanigawa, A. Naka, and K. Oda, "Field trial of 43-Gbit/s RZ-DQPSK transmission in aerial fiber with rapidly changing SOP," in *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, 2009, p. NWD1.

[36] D. Charlton, S. Clarke, D. Doucet, M. O'Sullivan, D. L. Peterson, D. Wilson, G. Wellbrock, and M. Bélanger, "Field measurements of SOP transients in OPGW, with time and location correlation to lightning strikes," *Opt. Express*, vol. 25, no. 9, pp. 9689–9696, May 2017.

[37] F. Pittalà, C. Stone, D. Clark, M. Kuschnerov, C. Xie, and A. Haddad, "Laboratory measurements of SOP transients due to lightning strikes on OPGW cables," in *Optical Fiber Communication Conference*, 2018, p. M4B.5.

[38] M. Kurono, K. Isawa, and M. Kuribara, "Transient state of polarization in optical ground wire caused by lightning and impulse current," in *International Symposium on Polarization Analysis and Applications to Device Technology*, vol. 2873, 1996, pp. 242–245.

[39] G. Long, F. Ling, and J. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 9, pp. 1397–1405, 1989.

[40] P. Kabal, "The stability of adaptive minimum mean square error equalizers using delayed adjustment," *IEEE Transactions on Communications*, vol. 31, no. 3, pp. 430–432, 1983.

[41] M. Sotoodeh, Y. Beaulieu, J. Harley, and D. L. McGhan, "Modulator bias and optical power control of optical complex E-field modulators," *Journal of Lightwave Technology*, vol. 29, no. 15, pp. 2235–2248, 2011.

[42] W. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique," *Signal Processing*, vol. 6, no. 2, pp. 113–133, 1984.

[43] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, 1975.

[44] F. M. Gardner, *Phaselock techniques*, 3rd ed. John Wiley & Sons, 2005, appendix 4A.

[45] J. N. Damask, *Polarization Optics in Telecommunications*. Springer, 2005, ch. 1.