System Optimization of an All-Silicon IQ Modulator: Achieving 100 Gbaud Dual Polarization 32QAM

Sasan Zhalehpour, Mengqi Guo, Jiachuan Lin, Zhuhong Zhang, Yaojun Qiao, Wei Shi, and Leslie A. Rusch

IEEE/OSA Journal of Lightwave Technology, (accepted 30 Nov. 2019)

System Optimization of an All-Silicon IQ Modulator: Achieving 100 Gbaud Dual Polarization 32QAM

Sasan Zhalehpour, Mengqi Guo, Jiachuan Lin, Zhuhong Zhang, Yaojun Qiao, Wei Shi, and Leslie A. Rusch

***Abstract*—We experimentally demonstrate the highest, to the best of our knowledge, reported net rate in a SiP IQ modulator. At 100 Gbaud 32QAM (quadrature amplitude modulation), and assuming 20% FEC (forward error correction) overhead, we achieved a dual polarization net rate of 833 Gb/s. This record was achieved by adapting digital signal processing to the challenging pattern dependent distortion encountered in the nonlinear and bandwidth limited regime. First the Mach Zehnder modulator (MZM) operating point (trading off modulation efficiency and 3 dB bandwidth) and linear compensation (electrical and optical) are jointly optimized. Next, the key application of nonlinear pre- and post-compensation are explored. We show that nonlinear processing at the transmitter, in our case an iterative learning control (ILC) method, is essential as post-processing alone could not achieve reliable communications at 100 Gbaud. Nonlinear post-compensation algorithms pushed the performance under the FEC threshold with the introduction of structured intersymbol interference in post processing and a simple one-step maximum likelihood sequence detector. We provide detailed descriptions of our methodology and results.**

***Index Terms*—100 Gbaud 32QAM, All-silicon IQ modulator, Digital signal processing, Adaptive nonlinear pre-compensation, Blind nonlinear post-compensation.**

I. Introduction

FUTURE network technology must enable cloud infrastructure and data centers to meet the demand for bandwidth-hungry content, *e.g.*, video services, cloud applications, machine-learning applications, etc [1], [2]. This requirement has focused attention on new technologies that can scale up performance at lower cost and higher power efficiency.

Optical coherent systems can now operate at 100 Gbaud and beyond, as demonstrated with integrated optical Mach Zehnder modulators (MZMs) produced on several material platforms, including lithium niobate ($LiNbO_3$) [3], indium phosphide (InP) [4], and silicon photonics (SiP) [5]. Modulators based on each platform offer different trade-offs, which we review briefly for the most recent demonstrations at the highest reported baud rates.

S. Zhalehpour, M. Guo, L. A. Rusch, and W. Shi are with Department of Electrical and Computer Engineering, COPL, Université Laval, Quebec, Canada. (Corresponding auther: rusch@gel.ulaval.ca)

J. Lin and Z. Zhang are with Canada Research Center, Huawei Technologies Canada, Ottawa, Ontario, Canada

M. Guo and Y. Qiao are with the State Key Laboratory of Information Photonics and Optical Communications, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China.

This work was supported by Huawei Canada and NSERC (CRDPJ 486716-15).

$LiNbO_3$ has been the most frequently adopted technology in commercial single-carrier 400G/600G systems, typically running at 64 Gbaud with modulation formats of 16/64 quadrature amplitude modulation (QAM). Scaling to single carrier 800G or 1T for optical interfaces requires 100 Gbaud and beyond [3], [6]. A single carrier, dual polarization 64QAM case at 100 Gbaud was reported in [3]. For wavelength division multiplexing (WDM), 12 channels spaced 50 GHz apart at 120 Gbaud was reported in [6], using electronic time-division multiplexing (ETDM) and polarization division multiplexing (PDM) of 16QAM modulation.

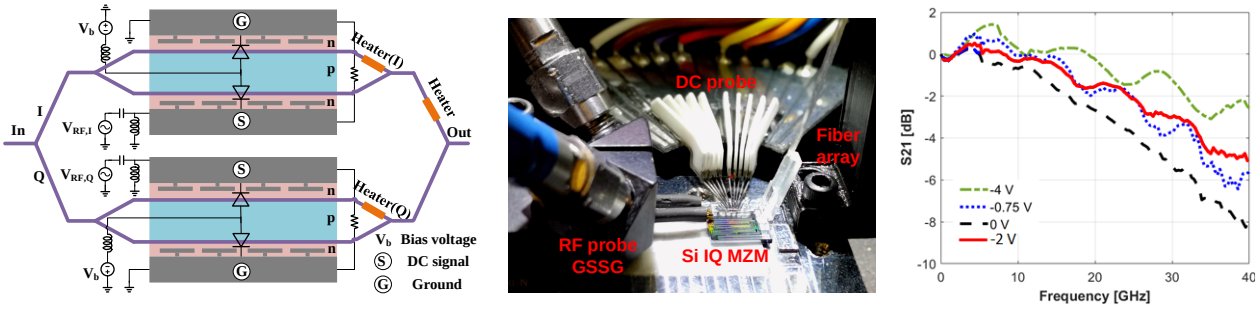
MZMs produced in InP can offer similar V_π, insertion loss, and baud rates as those produced in $LiNbO_3$, but at smaller size [7], [8]. A few excellent results have been reported including an InP-based transmitter running at 100 Gbaud 32QAM [4].

SiP offers some unique advantages missing in $LiNbO_3$ and InP, *e.g.*, complementary metal-oxide semiconductor (CMOS) compatibility for mass production, co-integration of electronics and photonics, and polarization manipulation on a single chip [7]. SiP also delivers a smaller footprint than $LiNbO_3$. However, SiP technology comes at the cost of higher insertion loss and V_π, and relatively lower bandwidth [7], [8]. Thus, due to the physical limitations on modulation efficiency and bandwidth, it is more challenging to achieve a high-quality QAM signal using a depletion-mode SiP MZM at 100 Gbaud and beyond. Silicon organic hybrid (SOH) technology has been used for 100 Gbaud 16QAM [9]. However, incorporating SOH into a standard CMOS-compatible silicon photonics process is not trivial, hence our focus is on all-silicon solutions.

From the first QAM experimental demonstration of quadrature phase shift keying (QPSK) at 50 Gb/s in 2012 [10], SiP IQ modulators continued to evolve in response to the demand for higher baud rates. Digital signal processing (DSP) techniques, in conjunction with improved SiP modulator design, is also boosting performance [11]. All-silicon modulators have been demonstrated at 85 Gbaud 16QAM and 64 Gbaud 64QAM [12], and 72 Gbaud 32QAM [13].

Recently, the authors reported a depletion mode all-silicon single carrier 32QAM at 100 Gbaud single polarization [5]. In this paper, we extend our preliminary results to present experimental results for dual polarization emulation and provide much greater detail on our DSP algorithms. We demonstrate 16QAM and 32QAM case at 100 Gbaud, corresponding to a line-rate of 400 Gb/s and 500 Gb/s per polarization, respectively. Several system operating parameters and reception al-

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15



(a) (b) (c)

Fig. 1: (a) Schematic diagram of SiP IQ modulator, (b) experimental set-up of optical and electrical connections on the chip, and (c) E/O S21 of SiP IQ modulator for several reversed DC bias voltages [13].

gorithms were optimized, including diode bias voltage, digital and optical pre-compensation, and receiver side DSP. We adopt iterative learning control (ILC) for nonlinear pre-compensation at the transmitter, as nonlinear post-processing alone could not achieve reliable communications at 100 Gbaud. To the best of our knowledge, this is the first demonstration of 100 Gbaud QAM generation using an all-silicon modulator that meets the forward error correction threshold at 16QAM and 32QAM.

The remainder of this paper is organized as follows. Section II describes MZM operation and our transceiver linear optimization methods in both optical and digital regimes. The experimental set-up for SiP IQ modulator signal generation and reception is described in section III, along with a discussion of DSP procedures to recover data and optimize performance. Section IV presents the nonlinear compensation at the transmitter and receiver sides. Section V presents the experimental results on dual polarization emulation with further discussion of optimization. We offer concluding remarks in section VI. Details of our single polarization results in [5] are presented in appendix as some DSP was routines varied.

II. Operating Point and Linear Filter Optimization

We describe our optimization, vis-à-vis the bandwidth and modulation efficiency, of the modulator operating point in the next subsection. We describe the transmitter linear optical and digital compensation methods in the second subsection. Note that receiver side linear signal processing is standard and described with the experimental setup in section III. Nonlinear compensation is described in section IV. A block diagram in Fig. 3 summarizes the equalization efforts covered in sections II-IV.

A. Operating point optimization

Our depletion-mode SiP IQ modulator has two nested MZMs with traveling-wave electrodes applied. A laterally doped p-n junction in a 220-nm-high silicon rib waveguide is used as the phase shifter. Three levels of dopant are applied to reduce the junction resistance without introducing excess optical loss with the series push pull configuration (Fig. 1a). Design details can be found in [13]. The V_π of the MZM is almost 8 V at -2 V reverse bias. The small-signal response (S_{21}) of the depletion-mode MZM shows a strong dependence on the DC bias. The 3-dB bandwidth is ∼22 GHz at zero bias, ∼26 GHz at -0.75 V, ∼30 GHz at -2 V, and ∼34 GHz at -4 V; see Fig. 1c. Furthermore, the modulation efficiency, in contrast to optical attenuation, decreases as the DC bias increases.

When operating the MZM at high baud rate, *e.g.*, 100 Gbaud, we need sufficient MZM bandwidth to avoid intersymbol interference (ISI). To ensure an appropriate trade-off between modulation efficiency and required bandwidth, we found the bit error rate (BER) at high signal to noise ratio for various bias voltages. The BER was obtained when using linear compensation, digital and optical pre-compensation methods will be discussed in section II-B. Results are reported in Fig. 2, where a clear minimum is visible at -2 V.

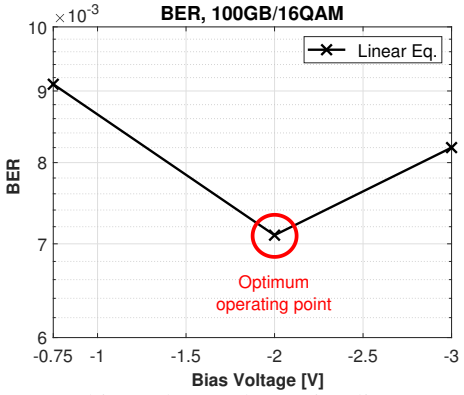


Fig. 2: BER vs. bias voltage when using linear compensation.

Contrast our result here with our previously published results for 84 Gbaud [13]. There, we used a relatively low reverse bias voltage of -0.75 V. Increasing the bias to -2 V reduces the RF loss at 50 GHz (3 dB bandwidth 30 GHz compared to 26 GHz), but also lowers modulation efficiency as the V_π was 7.25 V in the previous work at 84 Gbaud. As the BER curve in Fig. 2 varies with baud rate, the reverse bias operating point cannot be optimized by considering bandwidth alone.

B. Transmitter optimization of linear pre-compensation

The system level optimization is carried out by exploiting both optical and digital domains per the high level block diagram in Fig. 3. Optical pre-compensation is performed via

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15

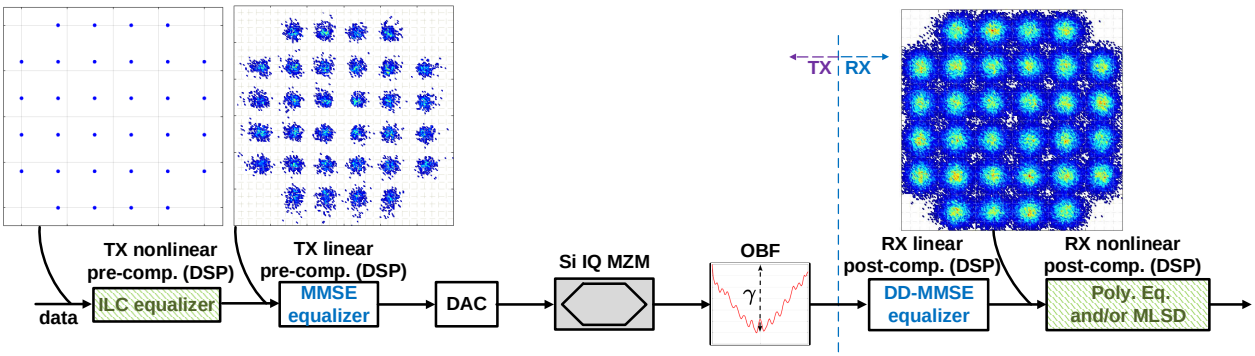


Fig. 3: High level block diagram of a signal processing (insets 32QAM at 100 Gbaud)

a programmable optical bandpass filter (OBF) and is strictly linear. In the digital domain, compensation is applied at both transmitter (this subsection) and receiver (next subsection), and can be linear alone, or linear combined with nonlinear techniques. We employ minimum mean square error (MMSE) linear digital compensation. Here we provide details on the transmitter-side linear compensation approaches; nonlinear techniques are covered in later sections.

Linear compensation is performed in two steps. We apply electrical compensation first, targeting inversion of the DAC alone. The MMSE filter compensation is found separately for in-phase (I) and quadrature (Q) branches. We use electrical back-to-back 16QAM transmission for MMSE estimation, treating it as two independent pulse amplitude modulation (PAM4) signals in I and Q. The MMSE filter has 250 taps.

Next, we linearly compensate the roll-off of the frequency response of the MZM in the optical domain. We transmit 16QAM at 100 Gbaud at high signal-to-noise ratio (SNR). The power spectral density (PSD) is estimated in the electrical domain after coherent detection. We design a pre-emphasis filter that inverts the PSD over a bandwidth of 100 GHz. The filter is smoothed to accommodate implementation in the optical domain via a programmable waveshaper, with smallest bin of 8 picometers. We again transmit 16QAM at 100 Gbaud at high SNR while sweeping the depth of the OBF, γ, to find γ that minimizes measured BER. We found that, when combined with the MMSE pre-compensation in the electrical domain, the optimal value is $\gamma = 17$ dB. This OBF is used for all further experiments.

III. Experimental set-up and signal processing

In this section, we present the experimental set-up along with offline DSP methods used in this work. The nonlinear processing is described in section IV.

A. Experimental set-up

The experimental set-up of the SiP IQ modulator system is shown in Fig. 4. The grey shading indicates the experimental set-up; all other paths are offline DSP work. The transmitter and receiver side DSP are summarized in tables to either side. The DSP path in the middle refers to the quasi-real-time adaptive pre-compensation entitled gain based iterative learning control (G-ILC).

The two outputs of a wideband DAC, carrying the I and Q components of an M-QAM signal, are de-skewed in time via tunable RF phase shifters (PS). The signals are then amplified with RF drivers (50 GHz, 24 dB gain). We used a ground signal-signal ground (GS-SG) configured RF probe, as seen in the photograph in Fig. 1b, to feed the RF signal ($\sim$5 V peak-to-peak) to the modulator.

The SiP IQ modulator configuration was discussed in section II-A. We drive the SiP IQ modulator at the null point; DC voltage of heaters (thermo-optic phase shifter) are accessed via a DC probe (see photographs in Fig. 1b) to set this operating point. We use a tunable external cavity laser (ECL) with linewidth less than 100 kHz at 1530 nm, coupled to the silicon chip via a fiber array (visible in Fig. 1b). Due to high losses in working with the SiP chip (9 dB coupling loss from the fiber array to the I/Q grating couplers, 6.8 dB modulator loss, and 3 dB splitting loss from the on-chip adiabatic 50:50 coupler), we boost the laser with a high power erbium doped fiber amplifier (EDFA) from 10 dBm to 26 dBm. Several other EDFAs are used throughout the optical channel to overcome the loss.

A commercial programmable waveshaper is used for optical pre-compensation, per section II. An OBF suppresses out-of-band amplified spontaneous emission noise after the second stage of EDFA. The SiP IQ modulator under test passes only TE mode, and a polarization controller (PC) is used.

We transmit a single polarization or employ a dual polarization emulator: a polarization beam splitter (PBS), a polarization beam combiner (PBC), and an optical delay link (ODL). The ODL is 2 m or 18 ns, which provides almost 1800 symbols delay between polarizations, while the MIMO DSP we use has only 65 taps avoiding any possible erroneous adaptation of the blind MIMO algorithm.

All experiments are back-to-back. We use a discrete coherent detector: a 90° hybrid coherent receiver module and four photodiodes (PDs) with 70 GHz bandwidth. As our wideband balanced PDs do not have transimpedance amplifiers, we use an EDFA as a pre-amplifier. The optical power before the 90° hybrid coherent receiver can be swept manually by means of a variable optical attenuator (VOA). The local oscillator (LO) has 16 dBm power and narrow linewidth (100 kHz). Signals are digitized by a 160 GSa/s, 63 GHz, real-time oscilloscope (RTO).

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15

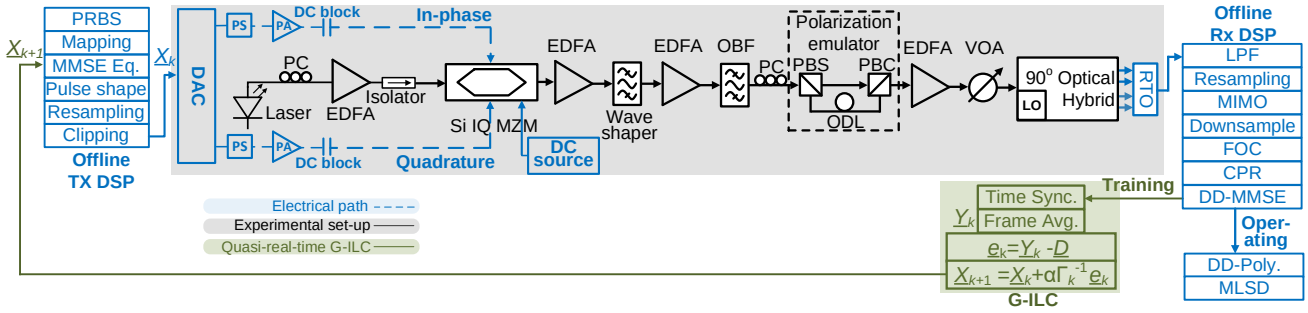


Fig. 4: Block diagram of the experimental set-up (grey shading) and offline DSP.

B. Standard signal processing

Consider the transmitter side DSP chain in Fig. 4. A pseudo random bit sequence (PRBS) of length $2^{19}-1$ (PRBS19) is Gray mapped and passed through the MMSE pre-compensation filter described in section II. The MMSE filter output is upsampled and shaped to a raised cosine pulse with roll-off of 0.01. The shaped pulse is resampled to accommodate the DAC sampling rate and baud rate (100 Gbaud). The resampled symbol sequence is clipped and quantized before being uploaded to the DAC.

At the receiver side, the signal is filtered by a 10^{th} order super Gaussian digital low-pass filter (LPF) and the captured 1.6 samples per symbol are upsampled to 2 samples per symbol. We use 4×4 multiple-input multiple-output (MIMO) for joint I/Q polarization demultiplexing; they are run at T/2 spacing to provide accurate timing, and then downsampled to one sample/symbol for all other processing. Frequency offset compensation (FOC) is performed blindly by a fast Fourier transform (FFT), carrier phase recovery (CPR) is a blind phase search with 64 test angles, and finally a decision-directed (DD) MMSE filter. Up to this step, the compensations employed are linear alone. These post-compensation methods are adapted blindly in the operating phase when evaluating system performance.

IV. Nonlinear Compensation

In this section we describe our nonlinear compensation. Transmitter side compensation is described first and involves adaptation with hardware-in-the-loop. Receiver side compensation combines two stages with parameters optimized offline.

A. Nonlinear digital pre-compensation

We use nonlinear compensation at the transmitter side to deal with the nonlinear response of the digital and optical components. We employ an adaptive method, G-ILC, which uses a fixed data sequence to optimize the transmit signal. The G-ILC method is well-known in control theory. It was initially introduced in [14] and recently employed in communications to linearize RF amplifiers [15], [16]. More detailed information for our application of G-ILC can be found in [17]. While high performance, the G-ILC method is limited to a specific sequence chosen for training. The G-ILC is a tool to estimate the best performance achievable by a system, or for calibration of the system.

A pre-determined sequence of data is transmitted. The error between transmitted and received signals is used to adaptively transform the transmitted signal to one with a clean received constellation via hardware-in-the-loop adaption. At each iteration, the transmitted signal is transformed until the received signal converges to the desired sequence $\underline{D}$, *i.e.*, standard IQ coordinates.

The G-ILC block diagram is shown in Fig. 4 in green shading. At the transmitter side, we build a 7000-symbol sequence and repeat this sequence multiple times until the DAC memory is full. At the receiver side, multiple copies of our transmitted sequence (of 7000 symbols) are captured and recovered through linear post-compensation alone.

Adaptation is done on a single polarization experiment. To reduce noise, we average over 14 time-synchronized sequences. In addition, we see at the top of Fig. 5 two approaches for linear compensation during adaptation. In the upper approach, the standard linear processing is applied - but with data-aided MMSE (400 taps) instead of the decision-directed MMSE used in BER assessment. In the lower approach we add a 2×2 MIMO step to undo correlations between I and Q and equalize each branch (65 taps). The MIMO is data-aided during adaptation instead of the decision-directed MIMO used in BER assessment. Note that in the lower approach data-directed MMSE is sufficient following the MIMO stage. By accessing a signal clear of linear distortion and additive noise, the G-ILC error signal can address the residual impairments more directly.

The averaged and linearly compensated output, $\underline{Y}_k$ in Fig. 4, is fed to the G-ILC controller block to update the transmitted signal for the next iteration. The G-ILC control block at the k^{th} iteration is expressed as

$$\underline{X}_{k+1} = \underline{X}_k + \alpha\Gamma_k{}^{-1}\underline{e}_k \tag{1}$$

where $\underline{X}_k$ and $\underline{X}_{k+1}$ are the transmitted sequences at the k^{th} and $(k+1)^{\text{th}}$ iteration, respectively; see Fig. 4. α is a variable step size to adjust the convergence speed. We define Γ_k as a diagonal learning gain matrix and $\underline{e}_k$ formed by the difference of $\underline{Y}_k$ and $\underline{D}$.

We report in Fig. 5 convergence results when using the two DSP approaches illustrated at the top of the plots. The upper method was the only approach used in single-polarization

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15

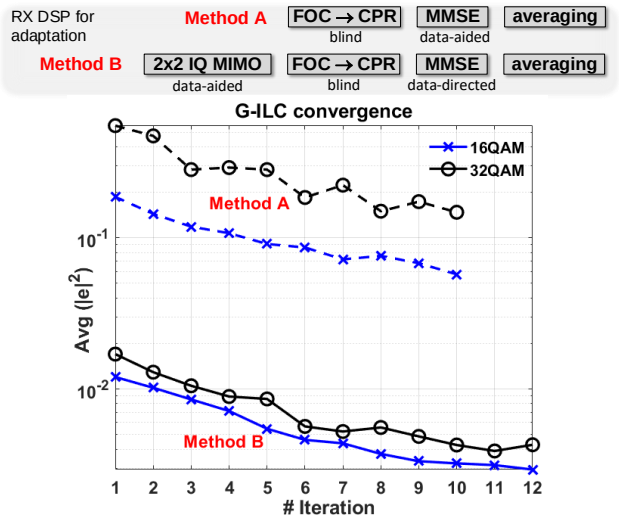


Fig. 5: Block diagrams of two DSP methods applied during training phase; convergence at 100 Gbaud for 16/32QAM: dashed lines using method A (upper) DSP, and solid lines using method B (lower) DSP.

[5] BER experiments. By incorporating the 2×2 MIMO we achieve an order of magnitude reduction in the mean squared error of the pre-distorted signal. This was a key enabler of our achievement of a line rate of 1 Tb/s with dual polarization emulation reported in the next section. By adjusting the adaptation step size, α, we can have a smooth convergence on the mean square absolute value of the error in each iteration.

In Fig. 5 we see that in using 2×2 MIMO the gap has closed between 16/32QAM error levels. The receiver offers a clear view of nonlinear impairments in this case, even for closely packed 32-QAM symbols. More importantly, the order of magnitude reduction means the G-ILC is a much more effective pre-compensation. Adaptation slope is not adversely affected, and convergence occurs in 12 iterations.

Figure 6 presents typical constellations before and after G-ILC block for single polarization transmission of 32QAM at 100 Gbaud at 2 dBm received power. All linear compensation are in place for both constellations, but no noise averaging. The G-ILC counters the pattern dependent distortion (PDD) introduced by memory effect in the system, with a visibly improved isolation of symbols.

B. Nonlinear post-compensation

While the G-ILC constellation in Fig. 6 is much improved, linear post-compensation was insufficient to bring BER below the FEC threshold; see section V. We therefore examined two blind nonlinear compensations approaches to deal with the residual nonlinear distortion. The first is a traditional polynomial based compensation, while the second is more involved.

1) Memory polynomial compensation: A Volterra based nonlinear equalizer introduces higher order terms into the linear filter structure. Due to the large computational complexity of the full-size Volterra equalizer, we use the simpler memory

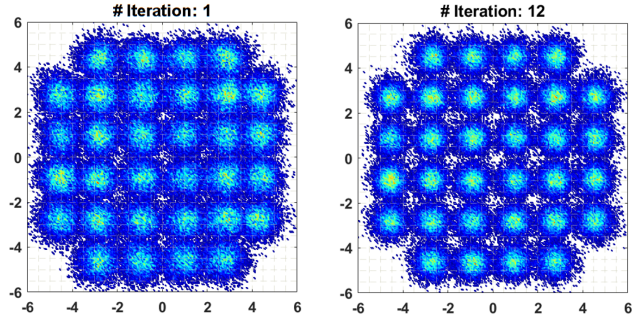


Fig. 6: 32QAM constellations (a) without and (b) with improved G-ILC.

polynomial compensation. In this case, only the diagonal terms of the Volterra kernels are retained [18]. In offline processing, we found that the memory polynomial post-compensation with 1^{st} and 3^{rd} order terms were able to improve performance. The complexity of including more terms was not justified given limited performance improvement. For input $x(n)$, the memory polynomial output $y(n)$ can be expressed as

$$y(n) = \sum_{l_1=0}^{L_1-1} w_1(l_1)x(n-l_1) \quad + \sum_{l_3=0}^{L_3-1} w_3(l_3)x^3(n-l_3), \tag{2}$$

where the tap coefficients $w_i(\cdot)$ are calculated with a DD-MMSE algorithm. We used filter lengths $L_1 = L_3 = 255$ to achieve BER results reported in the results section.

2) Post-filtering and one-symbol MLSD: In our bandwidth-limited system at 100 Gbaud, the linear equalizers, although minimizing the mean square error, enhanced the in-band additive noise. The nonlinear G-ILC pre-compensation and memory polynomial nonlinear post-compensation also leave residual pattern dependent distortion. We could combat these pattern dependent effects with a maximum likelihood sequence detector (MLSD), but they are prohibitively complex for systems with long memory and large constellation size. We turn to an MLSD approach with very limited complexity [19], [20].

Rather than trying to compensate for the true system mem-

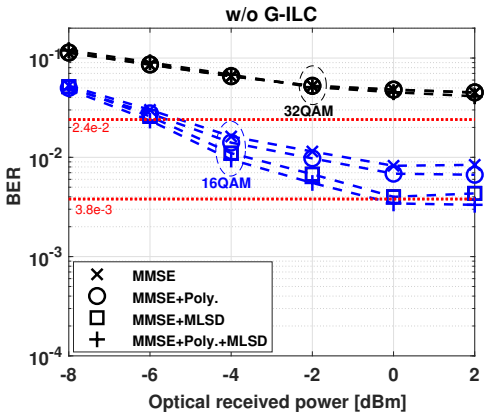
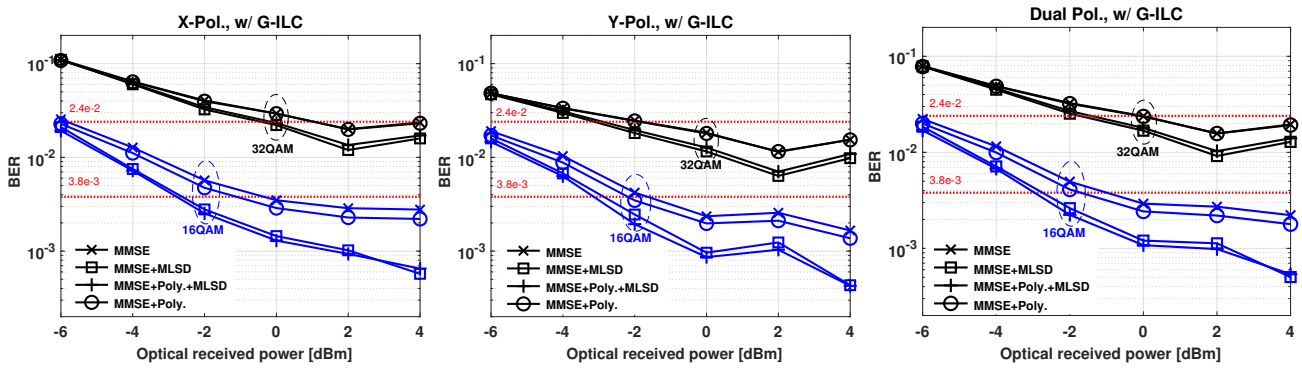


Fig. 7: BER sweep for 100 Gbaud DP-16/32QAM without G-ILC, but with post-compensation: DD-MMSE ($\times$), adding memory polynomial ($\circ$), adding MLSD ($+$).

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15



(a) (b) (c)

Fig. 8: BER vs. received power for dual polarization 16QAM and 32QAM at 100 Gbaud with nonlinear pre-compensation (trained G-ILC); post-compensation cases are linear DD-MMSE only ($\times$), linear and nonlinear polynomial alone ($\circ$), linear and nonlinear MLSD alone ($\square$), and all post-compensation together ($+$).

ory, we introduce structured intersymbol interference (ISI) in a final processing stage. Knowing the precise nature of that ISI, and holding the memory depth to a minimal single symbol duration, we can easily construct a low complexity MLSD. Intuitively, this approach gathers up residual impairments, adds simple, known structure (in the spirit of an error correcting code), and then uses a receiver optimal for the introduced ISI (not optimal for the system). This proved an effective boost to our system performance.

The transfer function of the post-filter introducing structured ISI has z-transform of

$$H(z) = 1 + \beta z^{-1}. \tag{3}$$

producing memory of one symbol duration - hence the simple MLSD [19]. The parameter β can be tuned to maximize performance improvement from the combined post-filter and MLSD. The MLSD considers path metrics

$$D(s(n)) = \sum_{i=n-TB}^{n} [z(i) - (s(i) + \beta s(i-1))]^2, \tag{4}$$

where $z(i) = y(i) + \beta y(i-1)$ is the output signal of the post-filter and $s(i)$ is the candidate hard-decision at each point for a given path; the summation is along the path leading to time n, $n > TB$, for some maximum traceback length TB. The MLSD is implemented via the computationally efficient Viterbi decoding that seeks the path with the minimal distance metric $D(s(n))$.

We chose received power of 2 dBm (one of least noisy values) to find the best β. We swept ten points in the range 0.1 to 1, and observed a bowl shaped BER. The optimal values were $\beta = 0.6$ for 16QAM and $\beta = 0.5$ for 32QAM. These values were then used when processing data for all received powers.

V. Results and Discussion

This section presents the experimental results for various combinations of the compensation approaches described previously. We examine only 100 Gbaud, but two modulation formats, 16QAM and 32QAM. Training of G-ILC is performed on a single polarization, and the same converged G-ILC is applied in both polarizations as we use polarization emulation. In section V.A. we present performance on the single polarization before the G-ILC is trained. In section V.B. we present dual polarization results. We provide the single polarization BER evaluation for the case of having G-ILC (part of our previous work [5]) in the appendix.

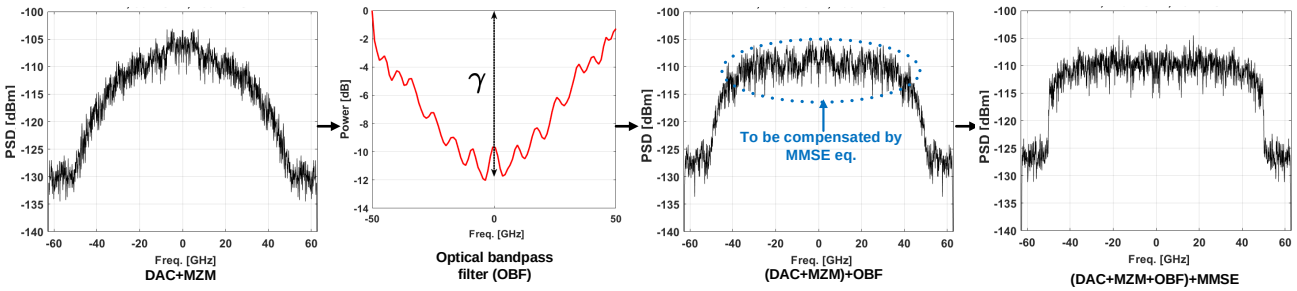
A. Prior to G-ILC training

We evaluate BER performance on a single polarization for 16QAM and 32QAM at 100 Gbaud before training the G-ILC pre-compensation. We examine baseline BER when linear combined (electrical/optical) pre-compensation at the transmitter side and various post-compensations at the receiver side. In Fig. 7, we present BER vs. received power with no G-ILC, *i.e.*, when no nonlinear pre-compensation is applied at the transmitter. For 16QAM, even with the help of two combined nonlinear post-compensation techniques, we can barely reach the 7% FEC threshold. In the 32QAM case, due to the high level of distortion, even the 20% FEC threshold is unattainable with all post-compensation efforts. The post-compensation improvement is worst for linear alone, and best for all methods (linear and nonlinear) combined. Among the two nonlinear approaches applied singly, the MLSD performs better than the memory polynomial filtering. The advantage is roughly $\sim$2 dB at the target BER of 10^{-2} when cascading DD-MMSE with MLSD, rather than DD-MMSE with memory polynomial filtering. From Fig. 7 we conclude that the performance improvement for 16QAM when including the memory polynomial does not justify the added complexity. Finally, we note that nonlinear approaches applied in post-compensation alone has no impact on 32QAM performance. The nonlinear post-compensation does not improve simple DD-MMSE for 32QAM.

B. Dual polarization case

We extend our 100 Gbaud 16QAM and 32QAM experiments in [5] to dual polarization by adding the polariza-

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15



(a) (b) (c) (d)

Fig. 9: Spectra at various points in linear compensation optimization: (a) after MZM, without any pre-compensation, (b) optical bandpass filter response with γ to control depth, (c) after OBF, and (d) after a MMSE compensation equalizer (*i.e.*, all linear techniques applied).

tion emulator to the optical path of the experimental set-up. The order of magnitude reduction in G-ILC adaptation error in incorporating 2×2 MIMO allowed us to meet the dual-polarization requirement for higher OSNR without any performance degradation.

BER results are reported in Fig. 8 for each polarization separately, X-pol and Y-pol, as well as for their average. In all plots, we have used nonlinear G-ILC plus linear combined (electrical/optical) pre-compensations at the transmitter side. We consider several post-compensation techniques, including DD-MMSE alone (cross markers), cascaded DD-MMSE and MLSD (square markers), cascaded DD-MMSE and memory polynomial (circle markers), and cascaded DD-MMSE and memory polynomial filtering and MLSD (plus markers). We observe the same behavior as that for single polarization. As can be seen in Fig. 8, the performance of Y-pol is slightly better than that of X-pol as a result of unequal sensitivity of the four PDs used. Unlike the single polarization, we swept the BER from -6 dBm to 4 dBm. In 16QAM case, we can achieve 7% FEC threshold as low as $\sim$-3 dBm and 32QAM BER reaches below 20% FEC threshold at less than $\sim$-2 dBm. As we already seen for the single polarization case, the memory polynomial stage does not offer much improvement on the performance compared to the MLSD stage.

The maximum back-to-back transmission rates reported to date on SiP platform offer net rate below 400 Gb/s per polarization. These include [9] using SOH component at 100 Gbaud 16QAM with a net rate of 333 Gb/s assuming 20% FEC overhead and [12] with a net rate of 333 Gb/s for 64 Gbaud 64QAM assuming 15% FEC overhead. Here, we experimentally demonstrated 100 Gbaud 32QAM; assuming 20% FEC overhead we achieved a single polarization net rate of 416 Gb/s and maintained that rate with a demonstration of dual polarization at 833 Gb/s. This is, to the best of our knowledge, the highest reported net rate in a SiP IQ modulator.

The G-ILC is not applicable to random data, but provides convincing evidence that nonlinear approaches to combat PDD are essential. A behavioral model of the system reflecting the G-ILC would be one solution, or a lookup table for nonlinearity could be helpful [21].

VI. Conclusion

We successfully demonstrated single carrier back-to-back 16QAM and 32QAM transmission at 100 Gbaud via a SiP IQ modulator with 3-dB bandwidth less than 34 GHz. We reported a net rate of 416 Gb/s on a single polarization, and maintained that performance in a (833 Gb/s dual polarization demonstration. We achieved the BER lower than 7% FEC (16QAM) with good margin at 747 Gb/s for dual polarization. To the best of our knowledge, this is the first demonstration of 100 Gbaud QAM operation using an all-silicon modulator.

We applied several optimization steps in both modulator operation and at the system level. Nonlinear pre-compensation was shown to be essential. The quality of adaptation was enhanced by using a 2×2 MIMO stage that lead to much reduced adaptation error. We demonstrated the effectiveness of introducing structured ISI in post-processing to increase OSNR margin.

VII. Appendix

In this appendix we provide information on differences in optimizing the linear compensation in our previous single polarization work in [5] vs. that used in section II-B. In [5] we jointly optimized the digital and optical linear compensation.

We first applied optical and no digital pre-compensation. We transmit 16QAM at 100 Gbaud at high SNR through the (uncompensated) DAC and MZM. The PSD is estimated in the electrical domain after coherent detection, Fig. 9a. We design a pre-emphasis filter that inverts the PSD over a bandwidth of 100 GHz. The filter is smoothed (see Fig. 9b) to accommodate implementation in the optical domain via a programmable waveshaper, with smallest bin of 8 pm.

The output spectrum following the OBF, (Fig. 9c), shows residual in-band ripples (see dashed blue oval). This was next compensated by designing an appropriate electronic MMSE filter with 500 taps. For the joint optimization, the depth γ of the OBF was swept. For each γ a new MMSE filter was found. The BER was measured for each γ, with lowest BER at γ_{opt}=12 dB.

In the SP case, the DAC was able to compensate the ripples and attain acceptable BER. The PSD after applying the joint digital/optical pre-compensation is shown in Fig 9d. The center frequencies are noticeably flatter, and the corners

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15

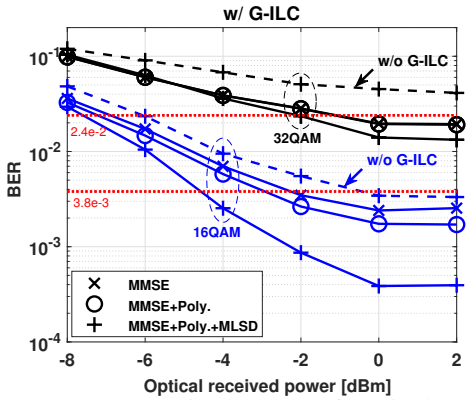


Fig. 10: BER vs. received power for single polarization 16QAM and 32QAM at 100 Gbaud, with nonlinear pre-compensation (trained G-ILC); post-compensation cases are linear DD-MMSE only ($\times$), linear and nonlinear polynomial alone ($\circ$), linear, and all post-compensation together ($+$).

of the Nyquist signal are more clearly defined. Note: in the dual polarization experiments reported in the body of the paper, both the technique described in this appendix and that described in section II-B were tested. Performance was best for the other approach in that case. The optimal OBF depth was quite different for the two approaches, 12 vs. 17 dB.

A. Single polarization with G-ILC

We add G-ILC to the transmitter side compensation and compare BER to the baseline case of only linear compensation at the transmitter (Fig. 7). This comparison establishes that nonlinear pre-compensation is essential; post-compensation, even nonlinear, is insufficient.

Adding G-ILC at the transmitter side significantly improves the performance, as seen in Fig. 10. We include the best case performance without G-ILC in dashed lines for comparison. The received power sensitivity for 16QAM (blue curves, plus marker) can be increased by more than 4 dB with G-ILC. We can now reach the 20% FEC threshold with 32QAM at received power as low as -2 dBm with using cascaded DD-MMSE, memory polynomial and MLSD post-compensations (black curve, plus marker). Even linear post-compensation alone is sufficient combined with G-ILC to surpass the performance when limiting nonlinear techniques to the receiver side. The PDD must be treated before transmission.

REFERENCES

[1] C. A. Thraskias, E. N. Lallas, N. Neumann, L. Schares, B. J. Offrein, R. Henker, D. Plettemeier, F. Ellinger, J. Leutho, and I. Tomkos, "Survey of Photonic and Plasmonic Interconnect Technologies for Intra-Datacenter and High-Performance Computing Communications," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2758-2783, 2018.

[2] Xiang Zhou, Hong Liu, and Ryohei Urata: '"Datacenter optics: requirements, technologies, and trends (Invited Paper)', Chin. Opt. Lett., 2017, 15, (5), pp. 120008.

[3] K. Schuh, F. Buchali, W. Idler, T. A. Eriksson, L. Schmalen, W. Templ, L. Altenhain, U. Dümler, R. Schmid, M. Möller, and K. Engenhardt, "Single Carrier 1.2 Tbit/s case over 300 km with PM-64 QAM at 100 GBaud," in *Proc. Opt. Fiber Commun. Conf.*, 2017, paper Th5B.5.

[4] R. W. Going, M. Lauermann, R. Maher, H. Tsai, A. Hosseini, M. Lu, N. Kim, P. Studenkov, S. W. Corzine, J. Summers, M. Anagnosti, M. Montazeri, J. Zhang, B. Behnia, J. Tang, S. Buggaveeti, T. Vallaitis, J. Osenbach, M. Kuntz, X. Xu, K. Croussore, V. Lal, P. Evans, J. T. Rahn, T. Butrie, A. Karanicolas, K. Wu, M. Mitchell, M. Ziari, D. Welch, and F. Kish,"1.00 (0.88) Tb/s per wave capable coherent multi-channel transmitter (receiver) InP-based PICs with hybrid integrated SiGe electronics," *IEEE J. Quantum Electron*, vol. 54, no. 4, pp. 8000310, 2018.

[5] S. Zhalehpour, J. Lin, M. Guo, H. Sepehrian, Z. Zhang, L. A. Rusch, and W. Shi, "All-Silicon IQ Modulator for 100 GBaud 32QAM cases," in *Proc. Opt. Fiber Commun. Conf.*, 2019, paper Th4A.5.

[6] J. Zhang, J. Yu, B. Zhu, Z. Jia, F. Li, X. Li, H. Chien, S. Shi, C. Ge, Y. Xia, and Y. Chen, "WDM case of Twelve 960 Gb/s Channels based on 120-GBaud ETDM PDM-16QAM over 1200-km TeraWave^{TM} Fiber Link," in *Proc. Opt. Fiber Commun. Conf.*, 2016, paper Tu3A.2.

[7] W. Heni, Y. Fedoryshyn, B. Baeuerle, A. Josten, C. B. Hoessbacher, A. Messner, C. Haffner, T. Watanabe, Y. Salamin, U. Koch, D. L. Elder, L. R. Dalton, and J. Leuthold, "Plasmonic IQ modulators with attojoule per bit electrical energy consumption," *Nat. Commun.*, vol. 10, pp. 1694, 2019.

[8] M. Jacques, A. Samani, D. Patel, E. El-Fiky, M. Morsy-Osman, T. Hoang, Md G. Saber, L. Xu, J. Sonkoly, M. Ayliffe, and D. V. Plant, "Modulator material impact on chirp, DSP, and performance in coherent digital links: comparison of the lithium niobate, indium phosphide, and silicon platforms," *Opt. Express*, vol. 26, no. 17, pp. 22471-22490, 2018.

[9] S. Wolf, H. Zwickel, C. Kieninger, M. Lauermann, W. Hartmann, Y. Kutuvantavida, W. Freude, S. Randel, and C. Koos, "Coherent modulation up to 100 GBd 16QAM using silicon-organic hybrid (SOH) devices," *Opt. Express*, vol. 26, no. 1, pp. 220-232, 2018.

[10] P. Dong, L. Chen, C. Xie, L. L. Buhl, and Y. Chen, "50-Gb/s silicon quadrature phase-shift keying modulator," *Opt. Express*, vol. 20, no. 19, pp. 21181-21186, 2012.

[11] J. Wang and Y. Long, "On-chip silicon photonic signaling and processing: a review," *Sci. Bulletin*, vol. 63, no. 19, pp. 1267-1310, 2018.

[12] J. Zhou, J. Wang, L. Zhu, Q. Zhang, and J. Hong, "Silicon Photonics Carrier Depletion Modulators Capable of 85Gbaud 16QAM and 64Gbaud 64QAM," in *Proc. Opt. Fiber Commun. Conf.*, 2019, paper Tu2H.2.

[13] J. Lin, H. Sepehrian, L. A. Rusch, and W. Shi, "Single-carrier 72 GBaud 32QAM and 84 GBaud 16QAM case using a SiP IQ modulator with joint digital-optical pre-compensation," in *Opt. Express*, vol. 27, no. 4, pp. 5610-5619, 2019.

[14] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Field Robot.*, vol. 65, no. 11, pp. 123-140, 1984.

[15] J. Chani-Cahuana, P. N. Landin, C. Fager, and T. Eriksson, "Iterative learning control for RF power amplifier linearization," *IEEE Trans. Microw. Theory Tech.* , vol. 64, no. 9, pp. 2778-2789, 2016.

[16] M. Schoukens, J. Hammenecker, and A. Cooman, "Obtaining the Preinverse of a Power Amplifier Using Iterative Learning Control," *IEEE Trans. Microw. Theory Tech.*, vol. 65, no. 11, pp. 4266-4273, 2017.

[17] S. Zhalehpour, J. Lin, H. Sepehrian, W. Shi, and L. Rusch, "Mitigating pattern dependent nonlinearity in SiP IQ-modulators via iterative learning control predistortion," *Opt. Express*, vol. 26, no. 21, pp. 27639-27649, 2018.

[18] L. Ding, G.T. Zhou, D.R. Morgan, Z. Ma, J.S. Kenney, J. Kim, and C.R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," in *IEEE Trans. on Communications*, vol. 52, no. 1, pp. 159-165, 2004.

[19] J. Zhang, J. Yu, B. Zhu, and H. Chien, "WDM case of Single-Carrier 120-GBd ETDM PDM-16QAM Signals Over 1200-km Terrestrial Fiber Links," in *IEEE J. of Lightwave Tech.*, vol. 35, no. 4, pp. 1033-1040, 2017.

[20] J. Li, E. Tipsuwannakul, T. Eriksson, M. Karlsson, and P. A. Andrekson, "Approaching Nyquist limit in WDM systems by low-complexity receiver-side duobinary shaping," *J. Lightw. Technol.*, vol. 30, no. 11, pp. 1664-1676, 2012.

[21] S. Zhalehpour, J. Lin, H. Sepehrian, W. Shi, and L. A. Rusch, "Experimental demonstration of reduced-size LUT predistortion for 256QAM SiP Transmitter," in *Proc. Opt. Fiber Commun. Conf.*, 2019, paper Th1D.3.

Sasan Zhalehpour received the B.Sc. and M.Sc. degrees in electrical engineering from Iran University of Science and Technology, Tehran, Iran and M.Sc. degree in electrical and computer engineering from Ozyegin University, Istanbul, Turkey. He is currently working toward the Ph.D. degree at the

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15

Centre for Optics, Photonics and Lasers (COPL), Department of Electrical and Computer Engineering, Université Laval, Québec, QC, Canada. His research interests include DSP algorithm and implementation, coherent optical communication, and compensation techniques.

Mengqi Guo received the B.S. degree in telecommunications engineering from Beijing University of Posts and Telecommunications (BUPT) and the B.S. degree (with first class honors) in telecommunications engineering from Queen Mary University of London in 2015. She is currently working toward the Ph.D. degree in School of Information and Communication Engineering, BUPT. From 2018 to 2019, she is a visiting Ph.D student with the Department of Electrical and Computer Engineering and the Centre for Optics, Photonics and Lasers (COPL), Université Laval. Her current research interests focus on digital signal processing for optical transmission systems.

Jiachuan Lin received the Ph.D. degree from the Institute of Information Photonics and Optical Communications (IPOC), Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2016. From 2016 to 2018, he was a Post-Doctoral Fellow with the Université Laval, Québec, QC, Canada. He is currently with Huawei Technologies Canada, Ottawa Research Center, Ottawa, ON, Canada. His research interests include digital signal processing, silicon photonics, optical frequency comb, and spatial division multiplexing system.

Zhuhong Zhang (M'01) received the B.S. degree in optical electronics from Wuhan University, Wuhan, China, in 1986, and the Ph.D. degree in laser physics and engineering from Shanghai Institute of Optics and Fine Mechanics, Shanghai, China, in 1991. In 1999, he joined Nortel, Ottawa, ON, Canada, where he was engaged in the research and development of signal processing and performance evaluation for optical coherent transmission system. Since 2009, he is a Coherent System Architect with Huawei Ottawa R&D Centre, Ottawa. His research interests include DSP algorithm and implementation, coherent modems, and system evaluation.

Yaojun Qiao is currently a professor at the School of Information and Communication Engineering at Beijing University of Posts and Telecommunications, China. He received the B.S. degree from Hebei Normal University, Shijiazhuang, China, in 1994, the M.S. degree from Jilin University, Jilin, China, in 1997, and the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2000. He was with Lucent and Fujitsu from 2000 to 2007. In 2007, he joined BUPT, his research interest includes optical fiber communication system and network.

Wei Shi (S'07–M'12) received the Ph.D. degree in electrical and computer engineering in 2012 from the University of British Columbia, Vancouver, BC, Canada, where he was awarded the BCIC Innovation Scholarship for a collaboration entrepreneurship initiative. He is currently an Associate Professor with the Department of Electrical and Computer Engineering and a member of Centre for Optics, Photonics and Lasers (COPL), Université Laval, Québec, QC, Canada. Before joining Université Laval in 2013, he was a Researcher with McGill University, Montreal, QC, Canada, where he held a Postdoctoral Fellowship with the Natural Sciences and Engineering Research Council of Canada (NSERC). He holds a Canada Research Chair in Silicon Photonics. His current research interests include integrated photonic devices and systems, involving silicon photonics, nanophotonics, CMOS-photonics co-design, high-speed optical communications, chip-scale lasers, and optical sensors.

Leslie Ann Rusch (S'91–M'94–SM'00–F'10) received the B.S.E.E. degree (Hons.) from the California Institute of Technology, Pasadena, CA, USA, in 1980 and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, USA, in 1992 and 1994, respectively. She currently holds a Canada Research Chair in Communications Systems Enabling the Cloud with the Department of Electrical and Computer Engineering, Université Laval (UL), Québec, QC, Canada. She is a member of the Centre for Optics, Photonics and Lasers at UL. From 1980 to 1990, she was a Communications Project Engineer with the Department of Defense. While on leave from Université Laval, she spent two years (2001–2002) with Intel Corporation creating and managing a group researching new wireless technologies. She has published more than 150 articles in international journals (90% IEEE/OSA) with wide readership, and contributed to more than 195 conferences. Her articles have been cited over 6500 times per Google Scholar. Her research interests include digital signal processing for coherent detection in optical communications using silicon photonic devices, spatial multiplexing using orbital angular momentum modes in fiber, radio over fiber and OFDM for passive optical networks; and in wireless communications, supporting 5G over passive optical networks and optimization of the optical/wireless interface in emerging cloud based computing networks. Dr. Rusch is the recipient of the IEEE Canada Fessenden award for Contributions to Telecommunications. She was also the recipient of numerous awards for graduate training, including the IEEE Canada Ham Award for Graduate Supervision. She was the Vice President of Technical Affairs on the Board of Governors of IEEE Photonics Society. She has served on multiple technical program committees for major international conferences, and as an Associate Editor for the IEEE/OSA Journal of Optical Communications Networks and the IEEE Communications Letters. She is Fellow of the OSA.

National Science and Engineering Research Council of Canada (NSERC) Funding: CRDPJ 486716-15