A 40-Gb/s SiGe-BiCMOS MZM Driver With 6-$V_{p\text{-}p}$ Output and On-Chip Digital Calibration

Leonardo Vera, *Member, IEEE* and John R. Long, *Fellow, IEEE*

***Abstract*— A wide-swing optical modulator driver is implemented in 0.13-μm SiGe-BiCMOS using a three-stage distributed amplifier with a digital input line. Measurements demonstrate 6-$V_{p\text{-}p}$ differential output, symmetric 6-ps (min) rise/fall times, 333-fs (rms) additive jitter, and better than 20-dB output return loss (S_{11}) below 58 GHz. Full output swing with adjustable (6–12 ps) rise/fall times is realized after a three-step calibration sequence facilitated by an on-chip energy detector and digital control. The circuit supports 28–48-Gb/s external data sources, or $2^{11}-1$ PRBS and 1-0 internal data generators for calibration and characterization respectively. The 3-mm^2 driver IC (1.8-mm^2 active area) consumes 1.92 W from +5/−2.5-V supplies.**

***Index Terms*— Digital phase shifter, digitally controlled input line, distributed amplifier (DA), dynastat divider, energy detector, half-rate-clock PRBS, injection-locked I/Q oscillator, Mach–Zehnder modulator driver, optical communication.**

I. Introduction

MULTI-Gb/s optical communication systems typically rely upon external modulation using a Mach-Zehnder modulator (MZM) [1] to transmit a modulated carrier with relatively low frequency chirp and high extinction ratio compared with direct modulation. The voltage required to switch the MZM optical output can range from 2 V to more than 10 V [2], [3].

Modulator drivers implemented as distributed amplifiers (DAs) have excellent gain-bandwidth performance, however, the analog input line in conventional DAs is a source of dispersion, attenuation, ringing, and pulse distortion. Our previous work [4] highlighted impairments of the input line, and demonstrated that a digital input line improves performance by driving each stage of a 10-Gb/s MZM driver with an identical replica of the retimed input data. Digital control of the retiming clock at each stage ensures that input data are aligned with the signal propagating across the output line, thereby minimizing phase distortion.

The digitally controlled DA concept is developed further in this paper, and applied to an MZM modulator driver aimed at 40-Gb/s communication. The driver is designed to produce a 6-$V_{p\text{-}p}$ differential output swing across a 100-Ω load with 6-ps rise/fall times. Calibration and on-chip test capabilities are added to the 40-Gb/s prototype. On-chip calibration of the DA is shown to minimize the output rise/fall times, and it also eliminates many sources of processing, temperature, and supply variations. Furthermore, the self-test capability built into the new prototype simplifies characterization of the driver circuits. A production 0.13-μm SiGe-BiCMOS technology [5] is used for the implementation because it enables cointegration of digital signal preprocessing, control, supervision, and self-test circuits onto the same chip.

Innovations developed for the 40-Gb/s driver include: 1) calibration of the digital input line facilitated by an on-chip energy detector and a three-step calibration algorithm; 2) a shielded output line in standard aluminum metal (not copper); 3) 10% less power consumption than the circuit reported in [4] despite a 4× increase in data rate; 4) operation across 28–48-Gb/s data rates enabled by a digitally controlled clock synchronizer; and 5) on-chip square wave (1-0) and $2^{11}-1$ PRBS data sources for built-in self-test (BiST), calibration, and characterization.

Section II of this paper describes the new MZM prototype and its circuit details. The energy detector circuit and the three-step sequence used to calibrate the driver are presented in Section III. Finally, measured results from characterization of the prototype, and a comparison with selected commercial and recently published modulator drivers are presented in Section IV.

II. Digitally Controlled Distributed Amplifier

A block diagram of the driver prototype is shown in Fig. 1. A −2.5-V dc source supplies all the circuit blocks. A +5-V supply biases the output stages via 50-Ω metal-film on-chip back-termination resistors, which are decoupled to ground by C_D (top-right in Fig. 1). The high-speed clock input (bottom-left) is terminated in 50 Ω on-chip, and buffered to an injection-locked oscillator (ILO via CB_1) data retiming flip-flop F_1 via CB_2, and a clock divide-by-two circuit [6]. Three data sources are possible: 1) external binary data via a differential input with 60-$mV_{p\text{-}p}$ sensitivity; 2) a $2^{11}-1$ length pseudorandom-binary sequence (PRBS) generated for 40-Gb/s test and characterization; and 3) an internal 1-0 pattern generated for calibration purposes.

The data source is selected via multiplexer M_1. Full retiming by F_1 provides margin against variations in processing, supply

Manuscript received March 23, 2016; revised June 29, 2016 and August 29, 2016; accepted August 31, 2016. Date of publication November 22, 2016; date of current version January 30, 2017. This paper was approved by Associate Editor Ichiro Fujimori.

L. Vera is with the Electronics Research Laboratory, Delft Institute of Microelectronics and Submicron Technology, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: l.vera.1980@ieee.org).

J. R. Long is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: jrlong@uwaterloo.ca).

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

Digital Object Identifier 10.1109/JSSC.2016.2606601

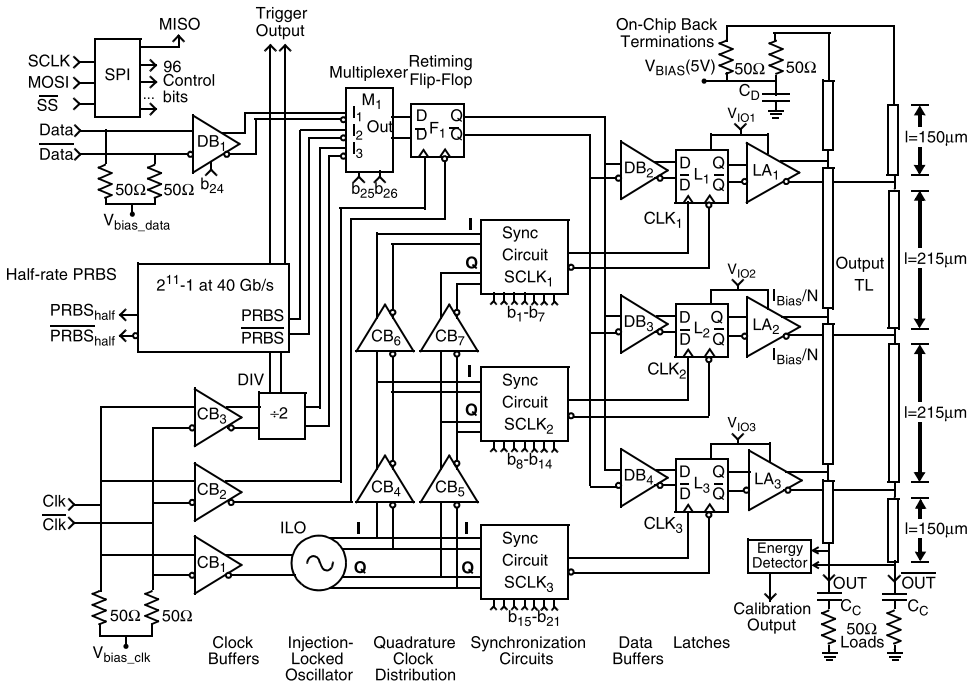


Fig. 1. 40-Gb/s M–Z modulator driver block diagram.

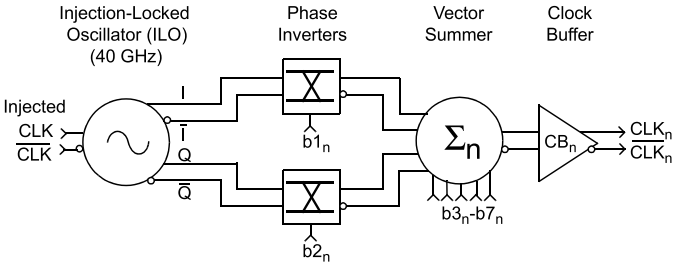


Fig. 2. Clock synchronization circuit for phase control of individual clocks ($n = 1, 2,$ or 3).

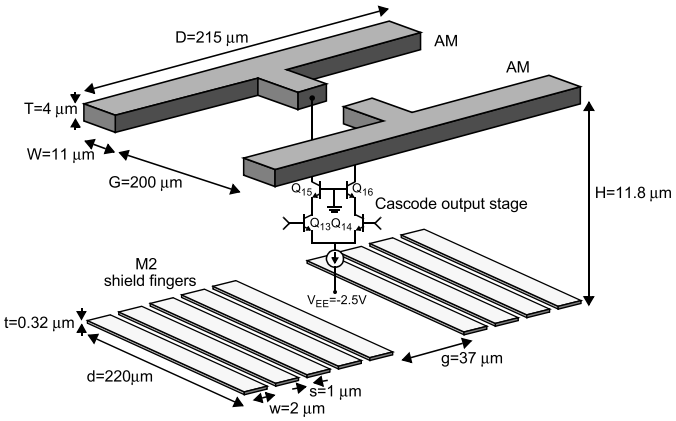


Fig. 3. Cross section of topmetal (AM) output line and M2 substrate shield.

voltage, and temperature (PVT). The retimed data are buffered and distributed to latches L_1–L_3 by DB_2–DB_4. Each latch output drives a limiting amplifier (LA_1–LA_3) in the DA with a replica of the input data timed to match the phase of the signal traveling along the 730-μm-long output transmission line (TL). An energy detector circuit measures the mean-square amplitude of the output. It is used during calibration to optimize the output amplitude, and thereby maximize rise/fall times at the output.

Quadrature (I/Q) clock streams are generated by a two-stage ILO that is injection-locked to the input clock. The ILO outputs are buffered and distributed to synchronization circuits (see Fig. 2). Each clock synchronizer consists of two phase inverters driving a vector phase shifter biased by current DACs (addressed by control bits $b3_n$–$b7_n$). It is used to adjust the phase of the clock that retimes data into latches L_1–L_3 (i.e., the inputs to each DA stage).

A. Output Transmission Line and Back-Termination Design

The 0.13-μm SiGe-BiCMOS technology used to implement the 40-Gb/s prototype has seven levels of metal available for interconnections, comprised of five (thin) copper layers beneath two (thick) aluminum metals at the top of the stack. The physical layout for each TL section connecting consecutive stages of the DA is shown in Fig. 3. The 4-μm-thick (T) and 11-μm-wide (W) topmetal lines have the lowest parasitic capacitance to the substrate and are therefore used to implement the output TL. The lines are separated by 200 μm (G). The cross-sectional area of the TL is chosen to satisfy dc and rms current restrictions imposed by electromigration requirements for the SiGe-8HP technology (i.e., 6-mA/μm width of topmetal) at 6-$V_{p\text{-}p}$ differential output swing across the 50-Ω loads. Attenuation is minimized by differential shielding implemented in the second copper layer (M2). The first metal (i.e., M1, below the shield) is reserved for supply and low-frequency wiring. The TL is synthesized in lumped-element (*L-C-L*) sections, where the series inductance (L_{line}) is derived from the two-wire topmetal line,

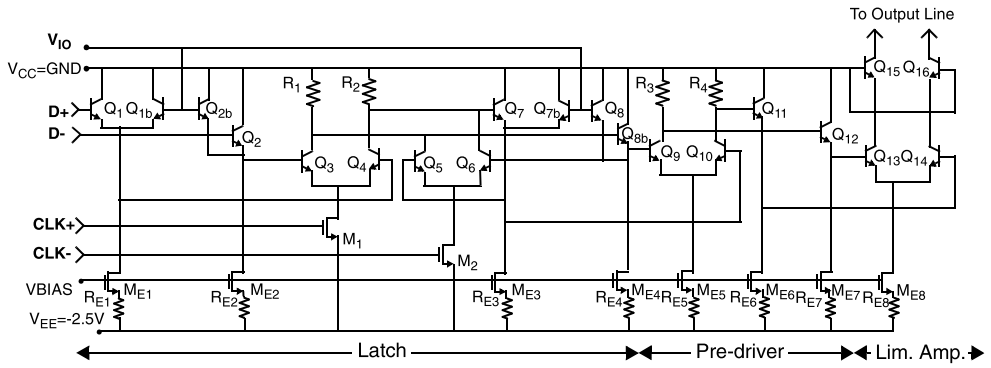


Fig. 4. Latch, predriver, and limiting amplifier schematics.

and the capacitance (C_{line}) arises from parasitics of the output amplifier and the substrate shield. The capacitive loading of each transistor on the line (C_{LA}) is 46 fF. Each section is 215 μm long (D), yielding a self-inductance of 160 pH and a parasitic capacitance of 18 fF for the topmetal lines over the shield. The characteristic impedance (Z_0) synthesized by the TL, and output loading of the limiting amplifier is $Z_o = (L_{\text{line}}/(C_{\text{line}} + C_{\text{LA}}))^{1/2}$, or 50 Ω. Each limiting amplifier connects to the TL at the center of each section (i.e., 107.5 μm from each end). TL sections connecting the back terminations and the outputs of the TL are therefore extended by 42.5 to 149.5 μm in total length. The 32-pH self-inductance of the extension compensates for 13 fF of parasitic capacitance added by the back terminations, bondpads, and associated wiring.

The substrate shield (see Fig. 3) is comprised of floating metal-2 fingers (i.e., not connected to any other conductors). The shield fingers reside 11.8 μm below the TL, and shield the top conductors from the substrate by electric induction [7]. Unwanted common-mode components in the output signal penetrate the floating shield and are attenuated by the semiconducting substrate. Digital control signals wired in M1 are therefore shielded from the differential output signal. Each shield finger is 0.32 μm thick (t), 220 μm long (d), and 2 μm in width (w) with 1-μm space between fingers (s). An opening of 37 μm in the shield (g in Fig. 3) facilitates connections between the amplifier stages and the TL in the layout. Large-signal simulations of the limiting amplifiers connected to the output TL predict a delay of $\sim$3 ps between consecutive output stages.

The 50-Ω back terminations are tantalum nitride (TaN) thin-film resistors. Each resistor conducts 60 mA of dc current continuously from the +5-V supply to the three DA stages. The desired differential output swing of 6 $V_{\text{p-p}}$ requires that 120 mA of current must be switched. This current is divided equally between the output stages, and therefore each stage switches 40 mA. Given the current limit of 0.5 mA/μm and sheet resistance of 60 Ω/sq for the TaN film, each back-termination resistor is sized at 120 μm in width and 100 μm in length.

B. Latch and Limiting Amplifier

Schematics of the latch, predriver, and limiting amplifier stages are shown in Fig. 4. Bipolar logic in the data path is cascoded onto the push–pull NMOS pair M_1 and M_2 (w/l = 5.2 μm/0.13 μm) in latches L_1–L_3 (from Fig. 1). Eliminating the tail current source normally used in a differential pair permits faster clocking because a significant source of parasitic capacitance is eliminated. A supply voltage of -2.5 V is possible because headroom for a bias current source is not required. However, to conduct a current (I_T) of 4 mA when switched "ON" and near zero when turned "OFF," the common-mode bias at the gates of M_1 and M_2 must be regulated. Gate bias for the transistors is set by the clock buffer preceding the latch (CB_n in Fig. 2). The common-mode output voltage of the buffer is controlled by an on-chip DAC that sets the voltage drop across R_9 in Fig. 7 (top-right).

Simulations predict full switching of transistors M_1 and M_2 across the anticipated process (best and worst cases), supply voltage (-2.3 to -2.7 V), and 0 °C–85 °C temperature range when a 400-$\text{mV}_{\text{p-p}}$ clock drives each latch input. The latch output swing is determined by the voltage drop across 100-Ω load resistors R_1 and R_2 (see Fig. 4) when conducting the full current. Bipolars Q_3–Q_6 are 0.13 μm wide, 2 μm in length, and biased to maximize the switching performance (i.e., 1.5 times the current density yielding peak f_T).

Emitter followers Q_7 and Q_8 (w = 0.13 μm, l = 6 μm) buffering the latch outputs are biased at 8 mA (peak f_T). They drive the limiting amplifier pair Q_9 and Q_{10} (w = 0.13 μm, l = 9.5 μm). Emitter followers Q_{11} and Q_{12} are 0.13 μm by 12 μm long and are also biased near peak f_T (17 mA). Driving amplifier transistors Q_{13}–Q_{16} are $2 \times 0.13 \times 9$ mm^2 in area. Cascode stage $Q_{15,16}$ reduces the Miller effect seen at the input of the differential driver, and raises the amplifier output impedance. A 3-$V_{\text{p-p}}$ output swing (6-$V_{\text{p-p}}$ differential) exceeds the collector-emitter breakdown voltage (open-base, $\text{BV}_{\text{CEO}} < 1.8$ V) of the npn in the 0.13-μm technology. However, the maximum voltage swing at the output before breakdown (i.e., BV_{CBO}) of $Q_{15,16}$ is actually 6 V, because a low impedance is used between the base and the ground for the cascode transistors [8]. The relatively large BV_{CBO} suggests that an output swing beyond 3 $V_{\text{p-p}}$ (single-ended) is possible. However, transistor beta degradation [9] as the base-collector swing approaches BV_{CBO} could affect reliability. In the absence of reliability data, a 3-V margin was adopted for this design.

Each output stage (LA_1–LA_3 in Fig. 1) may be turned ON/OFF independently using control voltages V_{IO1}–V_{IO3}

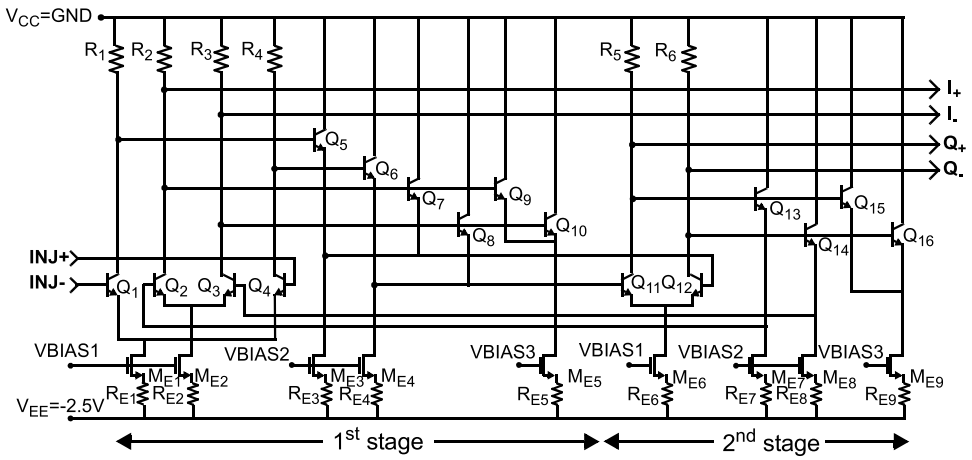


Fig. 5. Two-stage ILO.

(0 V for "ON" and −1.5 V "OFF"). Voltage V_{IO} steers current biasing transistors M_{E1}–M_{E4} away from the followers using $Q_{1b,2b}$ and $Q_{7b,8b}$ (see Fig. 4), thereby preventing ac signal from reaching the driver outputs. These controls are used to turn the limiting amplifiers "OFF" during system startup.

C. Injection-Locked Oscillator

The two-stage ILO produces a quadrature (I/Q) retiming clock from the input clock source.

An alternative for the generation of quadrature signals is a polyphase filter (PPF, e.g., [10]). However, the multiple PPF stages required to realize acceptable I–Q phase accuracy and wide operating bandwidth would suffer from insertion loss, and occupy large area.

The tunable ILO for quadrature clock generation developed in this paper has advantages in operating range (25–52 GHz), very good I/Q accuracy (±10°), and occupies just 90 μm × 90 μm of chip area.

Differential pair $Q_{2,3}$ and loads $R_{2,3}$ comprise the in-phase (I) signal gain stage (see Fig. 5). Emitter followers $Q_{7,8}$ buffer the voltage output to drive the second stage, which generates the quadrature (Q) signal via $Q_{11,12}$ and loads $R_{5,6}$. The ILO free-running frequency is determined partly by low-pass filtering from the load resistors and parasitic capacitances contributed by Q_5–Q_8, $Q_{13,14}$, $Q_{9,10}$, and $Q_{15,16}$. Emitter followers Q_5–Q_8 connect to the second-stage differential pair $Q_{11,12}$, and followers $Q_{13,14}$ feedback signal from the Q output to the I input.

The clock is buffered by transistors Q_1 and Q_4, and then injected into the ring oscillator by a wired-OR connection of emitter followers at the second-stage input (see Fig. 5). Simulations predict that injection of a 400-mV peak signal at these nodes increases the ILO lock range by 4 GHz compared with injecting the clock signal at the collector loads of Q_2 and Q_3 directly (i.e., summing the injected and first stage currents in loads R_2 and R_3). Injection via the emitter followers is less sensitive to capacitive loading of the injecting circuit.

The ILO lock range defines the frequency range for the entire DA. The range is extended by avoiding the use of

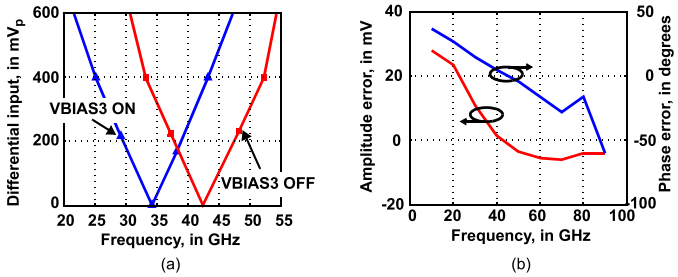


Fig. 6. Simulated sensitivity and I-Q error for the two-stage ILO. (a) Input sensitivity. (b) Amplitude and phase errors vs. frequency.

LC tuning and minimizing the Q-factor for the closed-loop circuit [11]. When $Q_{9,10}$ and $Q_{15,16}$ are biased "OFF" ($V_{BIAS3} = -2.5$ V), the self-oscillation frequency is 42 GHz, and the ILO locks to 400-mV_{pk} injected clocks between 34 and 52 GHz, as desired for the 40-Gb/s MZM driver. However, PVT simulations for supply voltage (−2.3 to −2.7 V), temperature (25 ºC–85 ºC), and best-to-worst case processing variations predict that the ILO self-oscillation frequency changes by ±7.5 GHz [200 Monte Carlo (MC) trials, and V_{BIAS3} fixed at −2.5 V]. Electronic tuning of the self-oscillation frequency was added to increase the locking range. It is realized by adjusting the input capacitance of emitter followers $Q_{9,10}$ and $Q_{15,16}$ via the transistor bias currents using V_{BIAS3}. This extends the low end of the lock range (postlayout) from 34 GHz down to 25 GHz when V_{BIAS3} is switched from −2.5 to −0.9 V (for a 400-mV_{pk} clock signal), as shown in Fig. 6(a). The extended tuning range accommodates PVT variations expected for the ILO circuit. The amplitude error is removed by limiting in the clock buffers.

D. Phase Inverters, Vectors Summer, and Clock Buffer

Data must be clocked into latches L_1–L_3 with an approximately 3-ps delay between consecutive stages to match the propagation delay of the signal along the output TL. Independent control of the retiming clocks (i.e., phases of CLK_1, CLK_2, and CLK_3 in Fig. 1) is realized using the I/Q phase selector and vector summing subcircuits shown in Fig. 7. The vector sum produced at the outputs of differential pairs

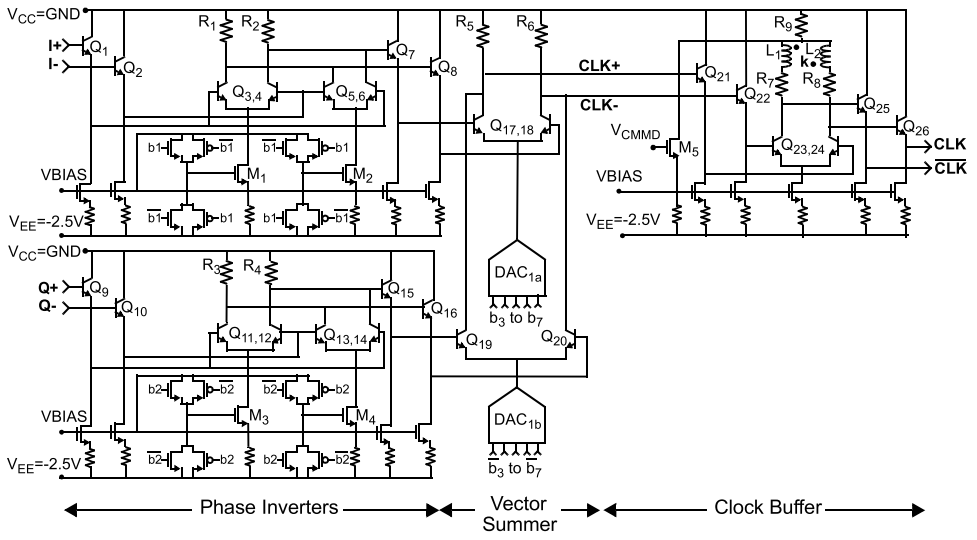


Fig. 7. I/Q clock phase inverter, vector summer, and clock buffer schematics.

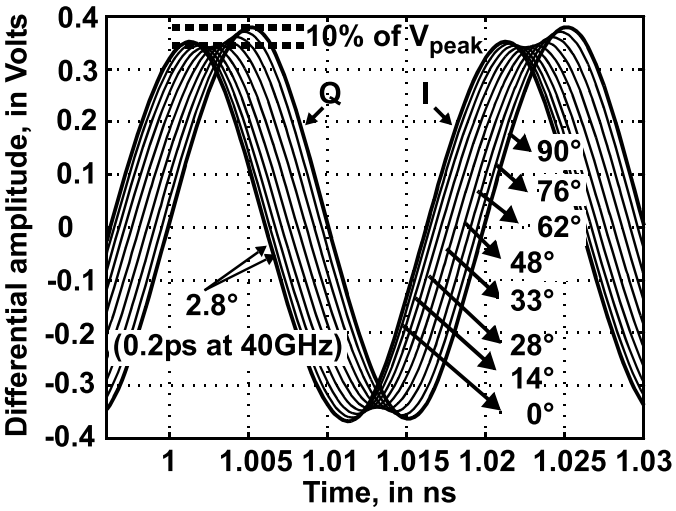


Fig. 8. Simulated vector summing output within one quadrant for ten code settings.

$Q_{17,18}$ and $Q_{19,20}$ varies with their tail currents. The tail currents are controlled by 5-b current D/A converters DAC_{1a} (I) and DAC_{1b} (Q), respectively. The total current supplied by the DACs is 4 mA, and a 340-mV_{pk} phase-variable clock is produced across 85-Ω loads ($R_{5,6}$) at the summer outputs. The output amplitude of the summer varies by 10% across the 90° phase control range (see Fig. 8).

The range of the phase shifter is extended to a full 360° by selecting from differential I or $\overline{I}$, and differential Q or $\overline{Q}$ clocks using control bits $b1_n$ and $b2_n$. Thus, a 7-b word controls each phase shifter, with the two least significant bits (LSBs) controlling the 0° or 180° selection of differential I and Q. The resolution of the vector summer determines the accuracy when adjusting the time delay between the stages of the DA along the output line. Matching the delay between stages to 3.0 ± 0.3 ps per stage (i.e., within 10% error) is possible when the minimum phase step is less than 4.3° for a 40-GHz clock. The 5-b current DAC biasing the vector summer has a resolution of 2.8° (i.e., $90° \div 2^5$), which corresponds to a 0.2-ps change in delay time per step at a 40-GHz clock. Note that jitter in the DA output signal is determined by the I/Q clock jitter (transferred to the DA output via retiming clocks CLK_1–CLK_3 from the synchronizers) and any jitter added by the data retiming circuitry.

The resolution of the DAC affects the DA output rise/fall times. The 20%–80% rise/fall times vary by ±3.2% at the 40-Gb/s data rate for a change of ±0.5 LSB around the optimal setting of the DAC current. The percentage error in the rise/fall times decreases as the clock frequency is increased, and vice versa (e.g., rise/fall time increases by 7.3% for a change of ±0.5 LSB at a 20-GHz clock). Furthermore, phase error between the ILO outputs causes variation in the resolution of the vector summer, which depends upon the quadrant selected by the phase inverter stages of the synchronizer. For example, an I–Q error of +10° (max. error across 30–50 GHz from simulation) in the first and third Cartesian quadrants (i.e., $f_{e-1,3} = f_Q - f_I - 90° = 10°$) yields a minimum phase step of 3.13° (i.e., 100°/32 states), or 217 fs at 40 GHz. The phase error if the second or fourth Cartesian quadrant is selected would be $f_{e-2,4} = f_1 - f_Q - 90° = -10°$, and the minimum phase step decreases to 2.5° (174 fs at 40 GHz). If the delay between stages is matched to within 217 fs (i.e., worst case for 10° I/Q error), then a maximum change of +/−3.5% in the 20-80% rise/fall time at 40-GHz clock is predicted from simulation.

E. Dynastat Frequency Divider

The near-dc minimum toggle frequency of a static divider and higher maximum toggling frequency of a dynamic implementation are combined in the dynastat divider (see Fig. 9) [6], [12]. For V_{MODE} equal to −2.5 V, the divider operates as a static divider. When V_{MODE} is set to −1.5 V, transistors M_2 and M_4 are turned off and the divider operates in the dynamic mode with a self-oscillation frequency

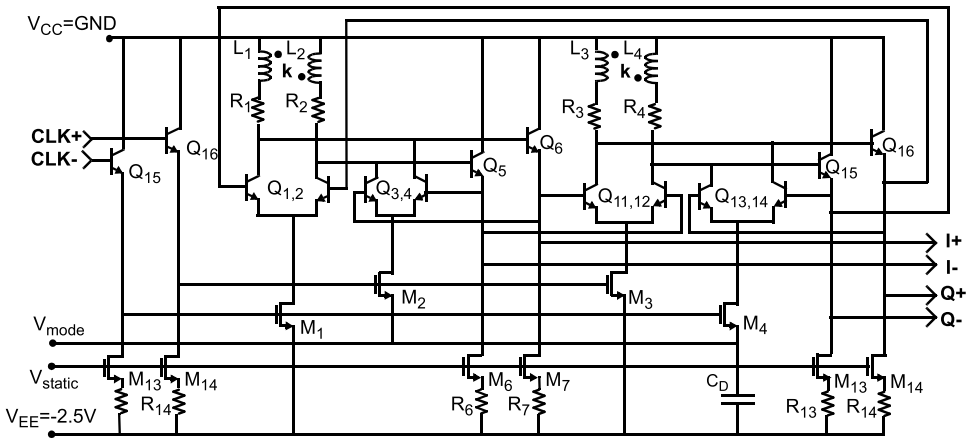


Fig. 9. Dynastat divide-by-2 circuit schematic [12].

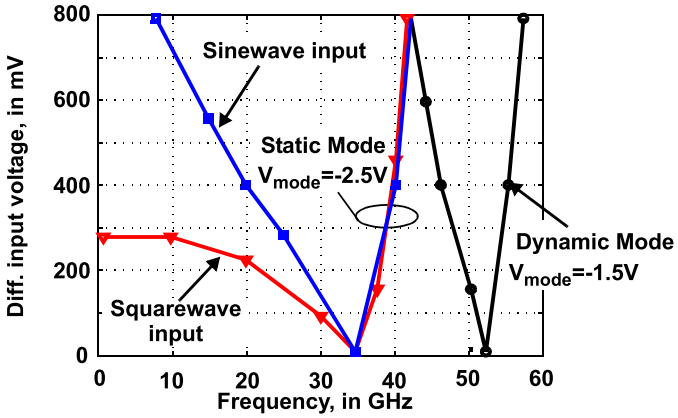


Fig. 10. Simulated input frequency sensitivity for the dynastat divider [12].

of 52 GHz. The clock frequency ranges for the two modes overlap, giving a higher maximum toggle frequency than a fully static circuit, and a minimum toggle frequency approaching dc. The clock amplitude sensitivity of the divider is also controlled using V_{MODE}. Junction capacitances C_{je} and C_{jc} (of $Q_{3,4}$ and $Q_{13,14}$ in Fig. 9) decrease as V_{MODE} increases from -2.5 V. Reduced loading due to lower C_{je} and C_{jc} increases the frequency for minimum input sensitivity. The simulated sensitivities to sinusoidal and square-wave inputs (5-ps rise/fall times) are plotted in Fig. 10. Postlayout simulations predict self-oscillation frequencies of 32 GHz in the static mode ($V_{\mathrm{mode}} = -2.5$ V), and 52 GHz in the dynamic mode ($V_{\mathrm{mode}} = -1.5$ V). The sensitivity of the divider is controlled via V_{mode} [12]. For example, simulations predict a self-oscillation frequency of 40 GHz (i.e., the desired nominal) when V_{MODE} is -2.35 V (i.e., dynastat in the static mode).

F. 40-Gb/s, $2^{11} - 1$ PRBS Generator and Trigger Countdown Output

The 40-Gb/s pseudorandom binary sequence (PRBS) generator [12] shown in Fig. 11 is integrated on-chip for self-test and characterization purposes. The PRBS circuit consists of two XOR gates and eleven registers clocked at one-half of the input clock rate. The two, half-rate outputs are multiplexed together to produce the full-rate data output. The generator produces a

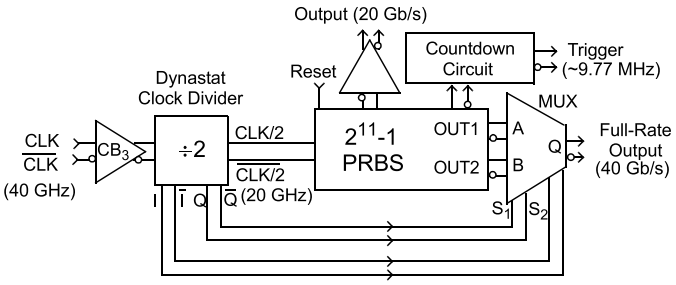


Fig. 11. Half-rate-clock $2^{11} - 1$ PRBS generator with trigger countdown output.

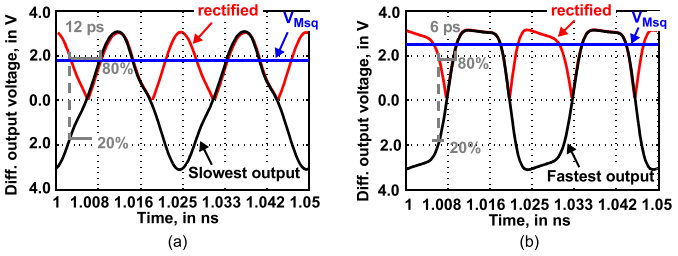


Fig. 12. DA output voltage waveforms for different interstage delays. (a) Slowest output waveform. (b) Fastest output waveform.

$2^{11} - 1$ length data sequence and the trigger signal required to synchronize external test and measurement equipment (e.g., an oscilloscope). The trigger is derived by counting down the half-rate sequence through a chain of nine divide-by-two stages to approximately 9.77 MHz for a 40-GHz input clock (i.e., divide by 512 in total).

III. MZM Driver Calibration

Mismatch in timing between the input and output lines of the DA results in distortion of the desired pulse shape (e.g., over- or undershooting) and suboptimal rise/fall times. Therefore, clocks driving the DA input latches are calibrated at startup to synchronize the retiming of data at each stage and match the propagation delay between stages across the output line. Any static errors caused by parameter drift during operation (e.g., caused by PVT variations) may also be

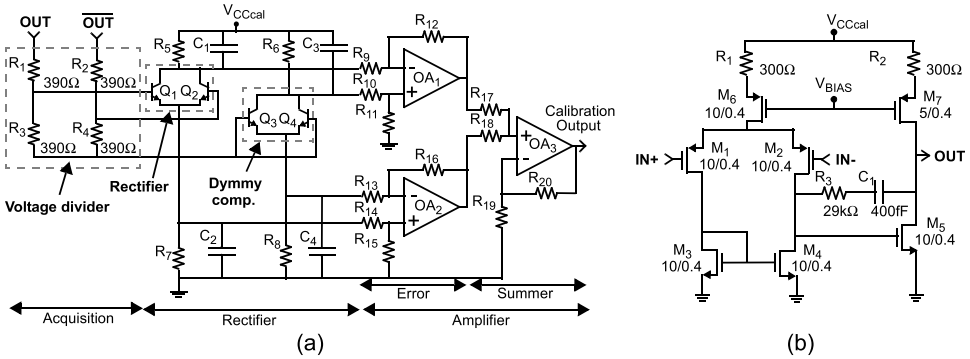


Fig. 13. Schematic of (a) the energy detector circuit, and (b) op-amp used for calibration.

corrected through (periodic) recalibration. MC simulations (200 trials) for supply settings −2.3/4.6 V and −2.7/5.4 V (i.e., ±10%), temperatures of 25 °C and 85 °C, and fast to slow process corners predict a ±30% variation in the rise/fall times from worst to best cases. However, changes in the rise/fall times after calibration caused by supply voltage and temperature variations may be reduced through recalibration. For example, the slowest rise/fall times (7.8 ps) predicted by the MC simulations at the slow process corner of 85 °C and low supply are reduced to 7.05 ps after recalibration. This is approximately 1 ps, or just 17% slower than predicted for the nominal process and supply at 25 °C.

Transient simulations of the DA for an alternating 1-0 input data pattern at 40 Gb/s are plotted in Fig. 12. The fastest edge rates (20%–80% rise–fall times of 6 ps) are realized for a uniform 3-ps delay between the latch clocks. Rise/fall times slow to 12 ps when the interstage delay between the clocks is extended to 9 ps. It is clear from observation of the waveshapes that the width and opening of the data eye is largest when the edge rates are as fast as possible. Therefore, calibration of the driver aims to realize the largest possible eye opening by adjusting the clock timing. Direct measurement of the rise/fall times at the output is avoided by noting that the mean-square output voltage (V_{Msq}) is proportional to the data eye, where $V_{\mathrm{MSq}} = (1/(T_2 - T_1)) \int_{T_1}^{T_2} V_{\mathrm{out}}^2$. In fact, V_{Msq} is largest when the interstage clock delay between the latch inputs matches the interstage propagation delay time across the output line. The proposed calibration circuit outputs a dc voltage proportional to the mean-square DA output voltage for a predetermined test pattern (e.g., a square wave).

A. Calibration Circuit

The calibration circuit is shown in Fig. 13. It consists of sensing, rectifying, and amplifying sections that measure the mean-square of the differential output voltage without affecting the signal quality adversely (e.g., due to capacitive loading). A peaking inductance of 35 pH in series with the output compensates for capacitive parasitics. Resistors R_1–R_4 (390 Ω each) attenuate the DA output amplitude by 6 dB before measurement. The single-ended parasitic capacitance to ground of the resistor and the 45×65 μm^2 output bondpad is 18 fF from postlayout extraction.

The attenuated output signal is rectified at the collector–emitter nodes of $Q_{1,2}$. Transistors Q_1 and Q_2 in Fig. 13 are biased using the 2-mA tail current flowing through R_7. Loading of the circuit on the driver 50-Ω outputs is negligible across dc-30 GHz at 1.5 kΩ. The collector–emitter bias of dummy pair $Q_{3,4}$ replicates the bias of rectifier $Q_{1,2}$. DC offset in the rectified signal is removed by summing amplifiers OA_1, OA_2, and OA_3 (overall gain of 14). The output from OA_3 is proportional to the mean-square of the DA output voltage.

Low-pass filters C_1-R_5, and C_2-R_7 (330 Ω||500 fF, see Fig. 13) filter the full-wave rectified signal at the respective collector and emitter nodes of $Q_{1,2}$. The ripple in the mean-square output (ΔV_{MSq}) can be estimated from $((\Delta V_{\mathrm{MSq}})/\Delta t) = -(I/C)$, where $C_1 = C_2 = C$, and I is the current biasing $Q_{1,2}$. The interval Δt is approximately one-half the period of the ac signal output by the rectifier [13]. Simulations predict that the ripple is less than 10% after filtering.

Frequency compensation of the OAs is realized by series-connected R_3 and C_1 (29 kΩ and 400 fF, respectively). Simulations predict a phase margin of 51° and gain margin of 12 dB when the op-amp output is loaded by 5 pF.

The simulated output of the rectifier for a 20-GHz sinewave input is plotted in Fig. 14. The error compared with the ideal response is just 4% for a 2-V input. Compared with an emitter follower peak detector [13], the linear range with respect to the input voltage of the differential full-wave rectifier ($Q_{1,2}$ in Fig. 13) is 50% wider, as shown in Fig. 14 (predicted from simulations).

B. Calibration Procedure

A timing diagram of the three-step calibration sequence developed in this paper is shown in Fig. 15. The 1-0 sequences (i.e., bit period = T_{CLK}) input to latches DB1–DB3 are assumed to be identical with zero skew. The phase of CLK_1 is adjusted in the first step (Step 1). Any violation of the setup time (t_{setup}) requirement for latch L_1 (e.g., CLK_1 rising before D input stable) is detected at the calibration circuit output by a decrease in voltage V_{MSq}. It should be noted that t_{setup} of the latches ($t_{\mathrm{setup}} = 2.4$-ps typ. from simulation) depends on circuit components, parasitics, data jitter, temperature variation, etc.

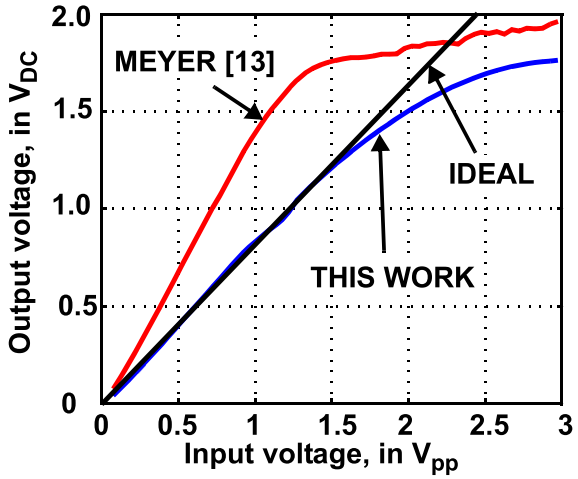


Fig. 14. Peak detector proposed in [13] compared to the energy detector in this paper.

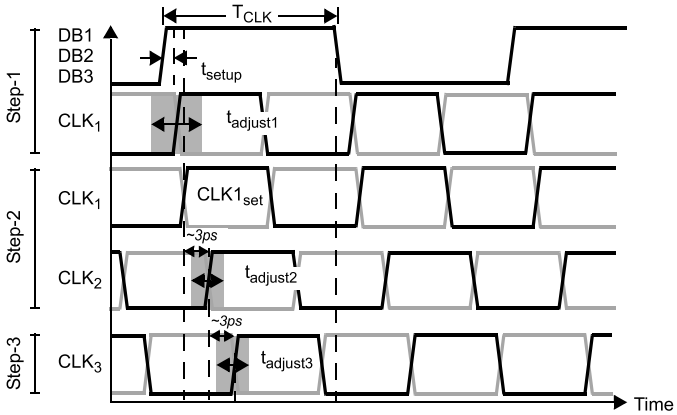


Fig. 15. Calibration steps and timing diagrams.

The time delay of CLK_2 with respect to CLK_1 is adjusted in the second calibration step (Step 2) to approximately 3 ps (i.e., the interstage delay across the output line). Finally, the third step (Step 3) adjusts CLK_3 to ~3-ps delay with respect to CLK_2. For example, the simulation results shown in Fig. 16 illustrates the effect of varying the delays (simultaneously) of CLK_2 with respect to CLK_1, and CLK_3 with respect to CLK_2 across the range from -6 ps to $+5$ ps. Clock timing for the first latch in the DA (i.e., CLK_1) is delayed by 2.4 ps with respect to the bit transition at the D-input of Latch 1 (as already determined from Step 1). Voltage V_{MSq} reaches a maximum of 1.55 V for 2.7-ps delay between clocks, which is in close agreement with the calculated propagation delay $t_{pd} = (L_{line} \cdot (C_{line} + C_{LA}))^{1/2} = 3.2$ ps.

A flowchart for the three-step calibration sequence is shown in Fig. 17. During calibration, data selector M1 (see Fig. 1) is set to select the 1-0 (sq-wave) pattern generated by the dynastat divider. Steps 1–3 adjust clocks CLK_1–CLK_3 in sequence. During Step 1, only the first DA stage is biased "ON" (i.e., $V_{IO1} = -1.5$ V in Fig. 1). Voltage V_{MSq} from the calibration circuit is monitored while bits b_1–b_7 controlling the synchronizer generating CLK_1 are incremented from 0 to 127 (7-b control word). The code t_{D1} corresponds to the minimum V_{MSq}. At t_{D1}, there is alignment between simultaneous transitions of the clock and data applied to latch L_1. As the latch captures $D = 0$ or $D = 1$ inputs with approximately equal probabilities, V_{MSq} approaches its lowest value. Once t_{D1} is identified, the synchronizer for CLK_1 is set to $t_{D1} + t_{setup}$. The calibration sequence then proceeds to Step 2, and the first and second output stages in the DA are biased "ON" ($V_{IO1} = V_{IO2} = -1.5$ V in Fig. 1). All 128 combinations for bits b_8–b_{14} addressing clock synchronizer 2 are evaluated while monitoring V_{MSq}. The code yielding the maximum V_{MSq} is identified, which corresponds to the fastest edge rates at the driver outputs. In the final step (Step 3), the third output stage is also biased "ON" ($V_{IO3} = -1.5$ V). Following a procedure similar to Step 2, bits b_{15}–b_{21} addressing clock synchronizer 3 are used to vary the phase of CLK_3. The code where V_{MSq} is maximum is then determined, which yields the fastest edge rates and the largest output amplitude from the driver.

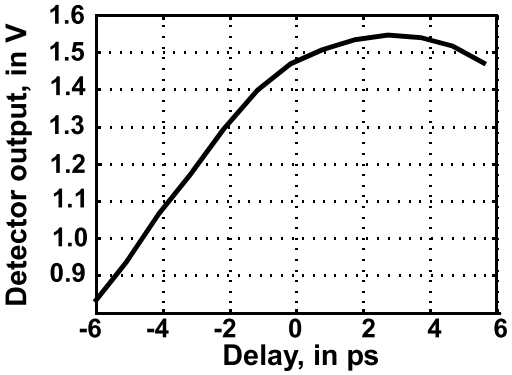


Fig. 16. Simulated detector output versus clock delay (i.e., CLK2 with respect to CLK1, or CLK3 with respect to CLK2).

IV. Prototype Characterization

A photomicrograph of the 1.5×2.0-mm^2 MZM driver prototype is shown in Fig. 18. It is fabricated in Global Foundries 0.13-μm SiGe-BiCMOS 8HP technology ($BV_{CEO} = 1.8$ V and $BV_{CBO} = 5.9$ V [5]). DC and RF on-wafer probes are used to characterize the driver. An external clock source generates the 40-GHz clock, which is fed to the chip via 40-GHz GSGSG probes. The 1-GHz trigger for eye pattern measurements is generated by mixing the 40-GHz clock and a second generator set at 39 GHz (i.e., 1 GHz below the clock synthesizer frequency). This trigger is distinct from the 9.77-MHz synchronization signal generated via the PRBS generator, which is used to measure the PRBS sequence. The three-step calibration sequence proposed in Section III was applied to set the clock timing for each driver stage using the internal 1-0 data pattern. The measured output of the calibration circuit and synchronizer codes during each step are shown in Fig. 19. The code corresponding to a 2.4-ps delay (t_{setup}) is added to t_{D1} at the end of Step 1. Note that the output of the calibration circuit steps up by ~0.1 V when output stages 2 and 3 of the DA are activated. The code words determined at the end of each step are also annotated in Fig. 19. All 384 combinations for the three control words addressing the clock synchronizers were tested. The time required to execute the entire calibration sequence is less than

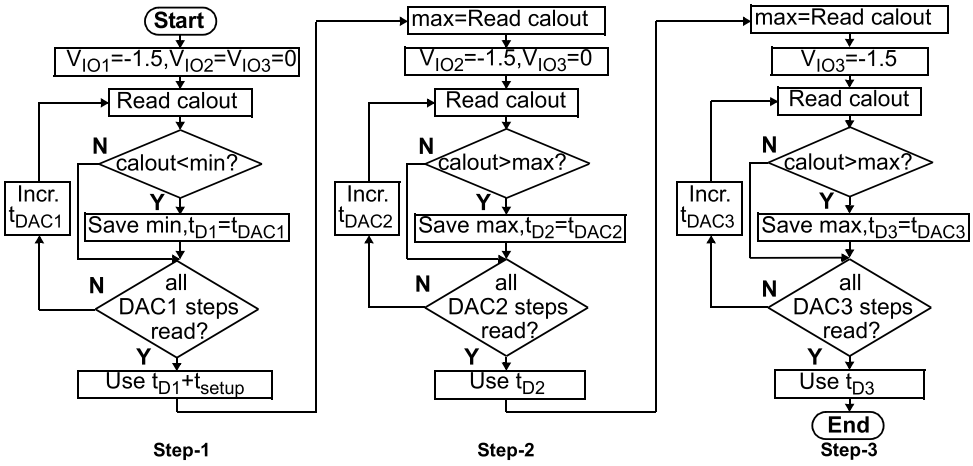


Fig. 17. Calibration sequence for DA cells during input clock phase adjustment.

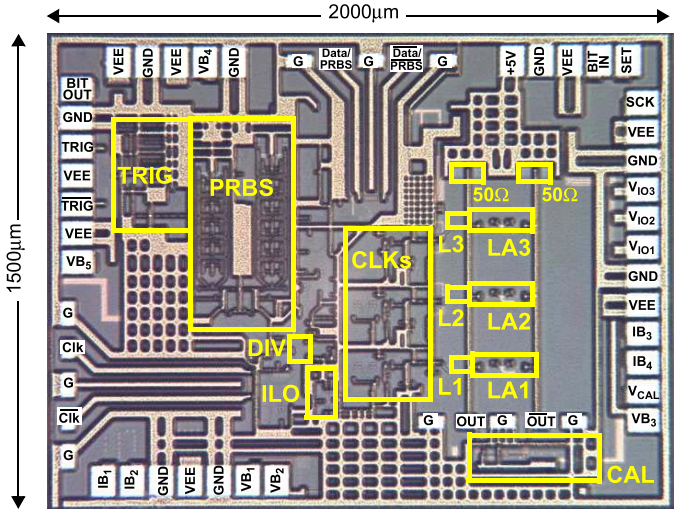


Fig. 18. 40-Gb/s M–Z modulator driver prototype photomicrograph.

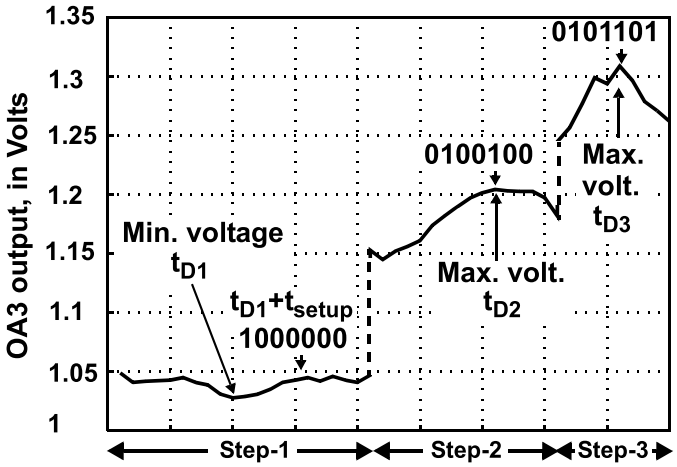


Fig. 19. Calibration sequence for DA cells during input line phase adjustment.

0.5 s using an external microcontroller and an ADC with a conversion time of 1 ms.

Digital control over the inputs to the DA opens up the possibility of generating other edge profiles at the output.

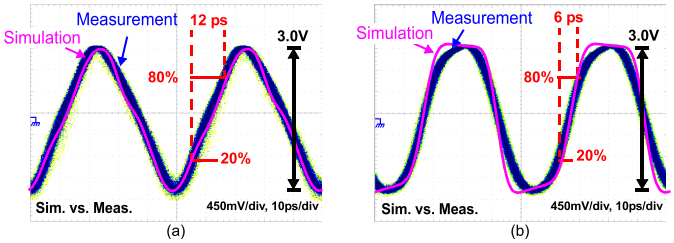


Fig. 20. Measured output waveforms for two input phase settings. (a) Slowest output waveform. (b) Fastest output waveform.

For example, slow and fast measured waveforms are shown in Fig. 20. The slowest output [see Fig. 20(a)] is triangular in waveshape, with symmetric 12-ps rise/fall times (20%–80%) and a single-ended output amplitude of 3 $V_{p\text{-}p}$. This output is obtained for clock synchronizer control codes corresponding to 9 ps of interstage delay. The fastest output with 6-ps rise/fall times [see Fig. 20(b)] is obtained for 3-ps interstage delay between timing clocks.

After calibration, the on-chip $2^{11} - 1$ PRBS data generator is used for characterization. The measured output eye at 40 Gb/s for the calibrated DA is shown in Fig. 21(a), and it is identical at both single-ended outputs (except for a phase inversion). The common-mode content in the output signal (aside from DC offset) is negligible, thus the differential output is simply twice the amplitude of the single-ended output signal, as shown in the simulation of the differential output in Fig. 21(b). The output path for the test setup includes on-wafer probes (V-connectored GSGSG probes), dc-blocking capacitors, 20-dB fixed attenuators, and Tektronix 80E09 sampling modules (60-GHz bandwidth). V-type (1.85 mm) connectors are used for all the connectorized components in the measurement path. The measured rise and fall times (20%–80%) for 6-$V_{p\text{-}p}$ differential output (3-$V_{p\text{-}p}$ single-ended) are identical at 6 ps each (including rise/fall times of the measurement path). The eye diagram rise/fall times can be adjusted in the same fashion as demonstrated in Fig. 20. To measure the PRBS sequence, the oscilloscope is synchronized to the BiST trigger output. The

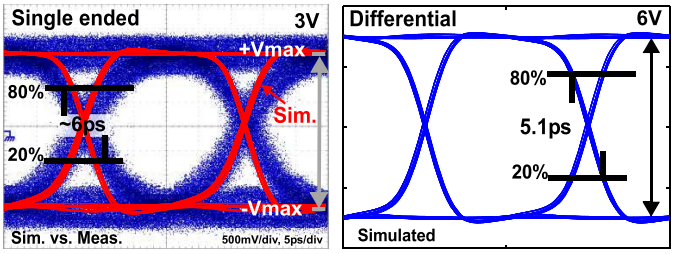


Fig. 21. Time-domain measured and simulated eye diagrams at 40 GHz using on-chip $2^{11} - 1$ PRBS. (a) Single-ended measured and simulated eye diagram. (b) Differential simulated eye diagram.

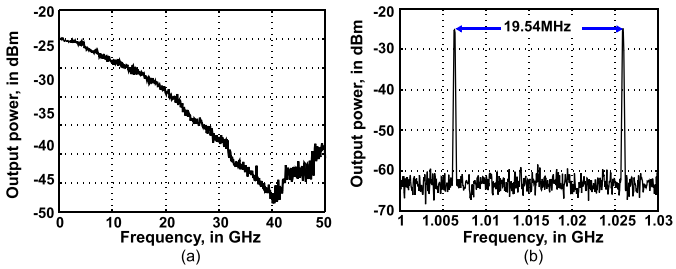


Fig. 23. Frequency-domain spectra of the outputs at 40 GHz with on-chip $2^{11} - 1$ PRBS. (a) Output voltage spectrum. (b) Discrete tones (40 GHz, $2^{11} - 1$ PRBS).

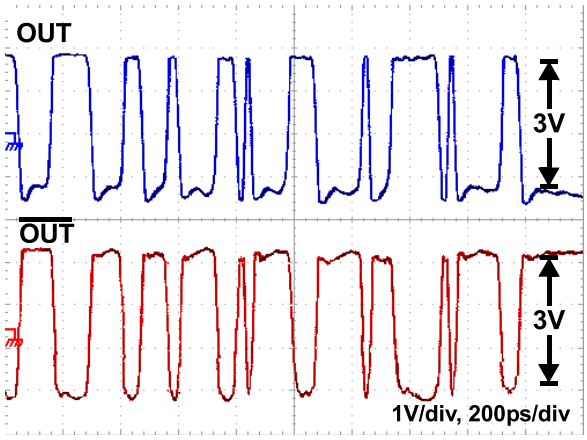


Fig. 22. Measured time-domain driver outputs at 40 GHz using on-chip $2^{11} - 1$ PRBS.

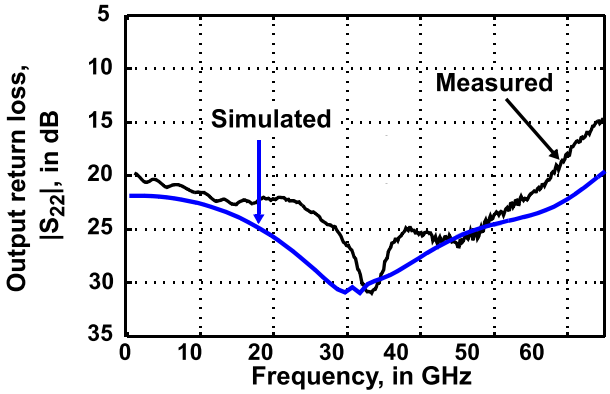


Fig. 24. Output return loss of the MZM driver (single-ended).

single-ended output measured after averaging 64 waveforms is shown in Fig. 22. The total measured time jitter at the output is 797 fs-rms, of which 334 fs-rms is added by the driver assuming that all the jitter sources are uncorrelated (i.e., jitter contributed by the driver, external clock source, oscilloscope, etc., add as the sum of squares).

Ringing observed in the 1-to-0 transitions arises from the second harmonic content in the output. The interconnect between the predrivers and limiting amplifiers crosses only one of the output lines, and the unbalanced parasitic creates a ~200-mV difference between the outputs. Resimulation of the DA including this coupling matches the measurements. The effect of the crossover in the layout can be compensated easily by balancing the interconnection parasitics on the output lines.

The output signal spectrum shown in Fig. 23(a) is measured with an R&S FSU-50 spectrum analyzer. The notch at 40 GHz is characteristic of a 40-Gb/s PRBS. The frequency spacing between the tones in Fig. 23(b) is $f_{\text{clk}}/(2^{11} - 1)$, or 19.54 MHz, as expected for a $2^{11} - 1$ length sequence generated from an f_{clk} of 40 GHz. The performance of the driver using a longer data sequence (e.g., $2^{31} - 1$ PRBS) requires increased decoupling of the common mode at the back termination (capacitor C_D in Fig. 1). Extra capacitance can be added externally to the IC to augment the on-chip portion of C_D.

Output return loss (ORL) for the driver with all output stages biased "ON" was measured using a network analyzer after TRL calibration. This measurement captures the behavior of the loaded TL, back-termination resistors, and bondpads. Excellent agreement is seen between measurement and simulation (seen in Fig. 24). The measured ORL is better than 20 dB from 1 to 58 GHz, and better than 15 dB up to 65 GHz.

The measured performance of the 40-Gb/s MZM DA prototype is compared with other drivers in Table I. Drivers fabricated in III-V technologies (e.g., GaAs or InP) generally produce greater output swing and consume less dc power than the SiGe prototype developed in this paper. For example, [18] is implemented in 0.15-μm pHEMT technology and produces a 25% higher output swing (i.e., 7.5 $V_{p\text{-}p}$). However, the rise/fall times (20%–80%) for the SiGe DA driver are faster. Moreover, they are adjustable across a range from 6 ps (min) to 12 ps (max). By comparison, the output rise/fall times of the pHEMT [18] and TGA-4942 (GaAs [16]) drivers are 4 ps greater at 10 ps. The SiGe prototype is the only driver with trimmable rise/fall times.

The ORL of commercially available drivers [14]–[16] exceeds 10 dB up to 40 GHz. The measured ORL for the digitally controlled SiGe-BiCMOS driver is 10 dB better across a 17-GHz wider frequency range.

Power consumption for the 40-Gb/s SiGe prototype is 1.92 W. If bias-Ts are used at the back-termination, it can be reduced to 1.55 W, which approaches the 1.4-W benchmark set by the driver [16]. Power consumption of the (single-ended) PSPL 5852 example [14] is 1.3 W, for an ac output/dc power-conversion efficiency of 2.8% compared with 4.7% for the SiGe-DA driver.

TABLE I
MODULATOR DRIVER PERFORMANCE COMPARISON

Ref.	Rate (Gb/s)	$V_{out\ p\text{-}p}$ (V)	$V_{in\ p\text{-}p}$ (V)	P_{DC} (W)	V_{DC} (V)	ORL (dB)	Rise/ Fall (ps)	Jitter rms (ps)	Area (mm x mm)	Application/ Technology
[14], PSPL5882 Tektronix	40	2.7 S.E.	0.6 S.E.	1.3	+8, -5	-10 <40 GHz	9/9 10-90%	-	-	Electro-absorption mod.
[15], 810 SHF Comm. Tech. AG	40	6.4 S.E.	0.33 S.E.	2.2	+10	-10 <40 GHz	9/9 20-80%	0.55	-	MZM
[16], TGA4942-SL Qorvo	43	8 S.E.	0.4 S.E.	1.4	+6	-10 <40 GHz	10/10 20-80%	0.4	-	Modulator driver
[17], GaAs IC 2004	40	5.5 Diff.	1.2 Diff.	1.7	+4.7, -4	-20 <45 GHz	8/8 20-80%	0.7	1.0 x 1.7 0.5 x 1.5 [b]	1.2µm In-GaAs-InP HBT
[18], RFIC 2003	40	7.5 Diff	1 Diff.	3	+1.5,+4 -4.5	-10 <20 GHz	10/10 20-80%	0.8	1.4 x 1.7 0.9 x 1.4 [b]	0.15µm PHEMT
This work	**40**	**6 Diff**	**0.3 Diff.**	**1.92**	**-2.5, +5**	**-20 <57 GHz**	**6/6[a] 20-80%**	**0.33**	**2.0 x 1.5 1.0 x 0.8** [b]	**MZM / 0.13µm SiGe BiCMOS**

a. Measured rise/fall times includes the rise/fall times of the 60GHz bandwidth measurement set-up.

b. Active area only

A unique aspect related to implementation in SiGe-BiCMOS is the presence of calibration and BiST circuitry on-chip to compensate for PVT variations at startup and (potentially) on the fly. These capabilities have been unavailable in high bit-rate modulator drivers reported to date. Complete built-in self-calibration and BiST functionality could be added to the driver in the future work by integrating the microcontroller, ADC, and associated control logic onto the BiCMOS IC. The calibration methodology demonstrated here uses a 1-0 pattern generated on-chip and a new calibration circuit. Self-calibration using live data traffic remains a topic for future work.

V. CONCLUSION

A digitally controlled, 40-Gb/s Mach–Zehnder modulator (MZM) driver prototype produces a 6-$V_{p\text{-}p}$ differential output voltage swing with excellent symmetry. The DA delivers edge rates trimmable from 6 ps (min. rise/fall time) to 12 ps (max.) under digital control of the input line, which is a unique feature of the circuit. The minimum rise/fall times realized by the DA prototype are faster than most drivers reported in silicon or III–V technologies. An on-chip energy detector circuit facilitates calibration using a simple three-step sequence, and the integrated $2^{11} - 1$ PRBS data generator enables characterization. Timing jitter added by the driver is 0.33 ps-rms, or just 1.3% of the 40-Gb/s period. The measured return loss is below -15 dB up to 65 GHz. Power consumption for the 40-Gb/s prototype (1.92 W) is 10% lower than a 10-Gb/s driver reported previously, despite the $4\times$ increase in data rate.

ACKNOWLEDGMENT

The authors would like to thank D. L. Harame and Global Foundries for technology access and fabrication, MOSIS for facilitating the MPW service, and A. Akhnoukh at TU Delft for the technical support.

REFERENCES

[1] A. K. Ghatak and K. Thyagarajan, *Optical Electronics*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 1989.

[2] K. Noguchi, O. Mitomi, and H. Miyazawa, "Millimeter-wave Ti:$LiNbO_3$ optical modulators," *J. Lightw. Technol.*, vol. 16, no. 4, pp. 615–619, Apr. 1998.

[3] L. Liao *et al.*, "40 Gbit/s silicon optical modulator for highspeed applications," *Electron. Lett.*, vol. 43, no. 22, Oct. 2007.

[4] Y. Zhao, L. Vera, and J. R. Long, "A 10 Gb/s, 6 $V_{p\text{-}p}$, digitally controlled, differential distributed amplifier MZM driver," *IEEE J. Solid-State Circuits*, vol. 49, no. 9, pp. 2030–2043, Sep. 2014.

[5] B. A. Orner *et al.*, "A 0.13 μm BiCMOS technology featuring a 200/280 GHz (f_T/f_{max}) SiGe HBT," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Sep. 2003, pp. 203–206.

[6] L. Vera and J. R. Long, "Dynastat frequency divider with DC-153 GHz range," *Electron. Lett.*, vol. 51, no. 12, pp. 908–910, Jun. 2015.

[7] T. S. D. Cheung and J. R. Long, "Shielded passive devices for silicon-based monolithic microwave and millimeter-wave integrated circuits," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1183–1200, May 2006.

[8] H. Veenstra, G. A. M. Hurkx, D. van Goor, H. Brekelmans, and J. R. Long, "Analyses and design of bias circuits tolerating output voltages above BV_{CEO}," *IEEE J. Solid-State Circuits*, vol. 40, no. 10, pp. 2008–2018, Oct. 2005.

[9] J. Kim, A. Sadovnikov, P. Menz, and J. Babcock, "Considerations for forward active mode reliability in an advanced hetero-junction bipolar transistor," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Sep./Oct. 2012, pp. 1–4.

[10] W. L. Chan, J. R. Long, and J. J. Pekarik, "A 56-to-65 GHz injection-locked frequency tripler with quadrature outputs in 90 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, San Francisco, CA, USA, Feb. 2008, pp. 480–629.

[11] B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1415–1424, Sep. 2004.

[12] L. Vera and J. R. Long, "A 40-Gb/s 2^{11}-1 PRBS with distributed clocking and a trigger countdown output," *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 63, no. 8, pp. 758–762, Aug. 2016.

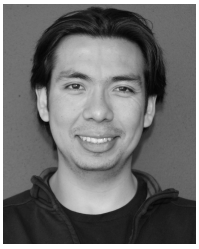
[13] R. G. Meyer, "Low-power monolithic RF peak detector analysis," *IEEE J. Solid-State Circuits*, vol. 30, no. 1, pp. 65–67, Jan. 1995.

[14] *40 Gb/s Broadband Amplifier, PSPL 5882 Datasheet*, Tektronix, Beaverton, OR, USA, Apr. 2013.

[15] *Datasheet SHF 810 Broadband Amplifier*, SHF Commun. Technol. AG, Berlin, Germany, Feb. 2007.

[16] *TGA4942-SL 43 Gb/s DPSK Modulator Driver, THA 4942 Datasheet*, Qorvo, Greensboro, NC, USA, Dec. 2012.

[17] Y. Baeyens *et al.*, "A high-gain InP D-HBT driver amplifier with 50 GHz bandwidth and 10 Vpp differential output swing at 40 Gb/s," in *Proc. IEEE GaAs IC Symp.*, Nov. 2003, pp. 153–156.
[18] Y. Baeyens *et al.*, "Compact high-gain lumped differential 40 Gb/s driver amplifiers in production 0.15 μm PHEMT technology," in *Proc. RFIC Symp.*, Jun. 2003, pp. 67–70.



Leonardo Vera (S'03–M'13) received the B.S. degree in electronics from the Faculty of Science and Technology, University of San Simon, Cochabamba, Bolivia, in 2004, and the M.Sc. and Ph.D. degrees in electrical engineering from the Delft University of Technology, Delft, The Netherlands, in 2008 and 2016, respectively.

From 2008 to 2010, he was a Research Engineer with the Electronics Research Laboratory, Delft University of Technology. He was with IBM, Essex Junction, VT (now Global Foundries) developing benchmark circuits in advance SiGe HBT BiCMOS technologies from 2012 to 2013. He is currently with Inphi Corporation, Westlake Village, CA, where he is involved in the design of modulator drivers. His current research interests include broadband amplifiers, frequency multipliers, dividers, and optical modulator drivers.

John R. Long (M'83–SM'14–F'15) received the B.Sc. degree in electrical engineering from the University of Calgary, Calgary, AB, Canada, in 1984, and the M.Eng. and Ph.D. degrees in electronics from Carleton University, Ottawa, ON, Canada, in 1992 and 1996, respectively.

He was with the Advanced Technology Laboratory, Bell-Northern Research, Ottawa, for 12 years, and then began an academic career at the University of Toronto, Toronto, ON, Canada, from 1996 to 2002. From 2002 to 2014, he was the Chair of the Electronics Research Laboratory with the Delft University of Technology, Delft, The Netherlands. In 2015, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His current research interests include low-power and broadband circuits for highly-integrated wireless transceivers, energy-efficient wireless sensors, millimeter-wave, and high-speed data communication IC design.