PHOTONICS Research

Single-drive high-speed lumped depletion-type modulators toward 10 fJ/bit energy consumption

Xinbai Li,[1] Fenghe Yang,[1] Fang Zhong,[1] Qingzhong Deng,[1] Jurgen Michel,[2] and Zhiping Zhou[1,*]

[1]*State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China*
[2]*MIT Microphotonics Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
Corresponding author: zjzhou@pku.edu.cn

Reduction of modulator energy consumption to 10 fJ/bit is essential for the sustainable development of communication systems. Lumped modulators might be a viable solution if instructed by a complete theory system. Here, we present a complete analytical electro-optic response theory, energy consumption analysis, and eye diagrams on absolute scales for lumped modulators. Consequently the speed limitation is understood and alleviated by single-drive configuration, and comprehensive knowledge into the energy dependence on structural parameters significantly reduces energy consumption. The results show that silicon modulation energy as low as 80.8 and 21.5 fJ/bit can be achieved at 28 Gbd under 50 and 10 Ω impedance drivers, respectively. A 50 Gbd modulation is also shown to be possible. The analytical models can be extended to lumped modulators on other material platforms and offer a promising solution to the current challenges of modulation energy reduction. © 2017 Chinese Laser Press

1. INTRODUCTION

As optical communication and interconnects rapidly march toward 400 G/1 T capacity, the energy consumption per bit needs to be drastically reduced to make the development trend sustainable. Total energy consumption (including electrical circuits) of current 100 G coherent transceivers and short-reach interconnect transceivers are on the level of >15 pJ/bit, and need to be improved by at least an order of magnitude to outperform electrical interconnects [1]. Specifically, the energy budget for modulators is on the order of ~10 fJ/bit. Although many silicon-resonator-based modulators have been reported [2–4], even realizing sub-fJ modulation energy [5], they require substantial thermal stabilization power of ~1 pJ/bit. Balanced interferometers eliminate the need for thermal control, but conventional silicon traveling wave modulators (TWMs) are generally energy-hungry physically because they use driving voltages inefficiently and dissipate all remaining power at termination resistors. Their energy consumption to realize a phase shift of $\Delta\varphi = 0.1\pi$ per arm is usually 1–10 pJ/bit [6–11]. This phase shift is the typical demonstration in the literature and is close to the short-reach interconnect requirement of 3.5 dB extinction ratio at quadrature point ($\Delta\varphi = 0.125\pi$). With state-of-the-art traveling wave electrode designs, TWM energy consumption as low as 32.4 and ~200 fJ/bit were

reported [12–14]. On the other hand, lumped silicon modulators have the potential of low energy consumption, but modulation speeds reported in the literature were primarily limited to 10 Gbps in the early phase of research [8,12]. For that reason, the research interest in silicon-lumped interferometer-based modulators has been less than that of high-speed TWMs. The bandwidth limitation for lumped modulators was later found to be impedance mismatch, and can be drastically improved with low characteristic impedance driving methods to above 25 or even 40 Gbps [15,16], which rekindled the hope of its practical use with specially designed integrated circuit (IC) drivers. However, current research efforts are still on the level of device prototype demonstrations without systematic methodology, and several issues need to be resolved before further performance improvements can take place: a) theories on electro-optical (EO) response are still incomplete, and knowledge of the approach and extent of energy consumption lowering has not been reported. b) It is known that, in TWMs, the appropriate voltage–length product for a π phase shift ($V_\pi L_\pi$) is around 1.8 V·cm (higher or lower both being detrimental). For lumped modulators, relations between modulation performance, junction efficiency, and phase shifter length remain unclear. c) A 3 dB EO bandwidth cannot exceed 20 GHz (0 V) when powered with standard 50 Ω radio-frequency (RF) cables.

In this work, we present a complete analytical model of small-signal EO analysis, time-domain large-signal EO response, quantitative energy consumption analysis, and eye diagrams on absolute value scales. Both Mach–Zehnder and Michelson interferometer types are investigated. In addition, the paper introduces a single-drive configuration to effectively enhance device EO bandwidth regardless of driving characteristic impedance. As a result, EO bandwidth can now reach 24 GHz with a 50 Ω driver. Finally, analysis of the dependence between dynamic phase change, phase shifter length, and doping concentration is performed. The lowest energy can be 80.8 and 21.5 fJ/bit for 50 and 10 Ω drivers at 28 Gbps, respectively, which are much lower than the energy consumption of conventional TWMs as well as previous lumped modulators [8,17]. This work can be extended to other lumped modulators including organic [18–20] and plasmonics [21,22] material platforms.

2. TERMINOLOGY

We first define the terms used in the following analysis to avoid ambiguity. "Lumped modulator" in this work refers only to interferometer-based modulators, excluding resonator-based ones. Mature theory is already available for the latter [23,24]. Only depletion-type modulation is investigated due to its fast EO response. "Lumped modulator" and "TWM" traditionally differ by whether device length is much less than RF wavelength so that voltage can be considered uniform and "instantaneous" along the device. Additionally, it is important that no termination resistors be used (otherwise, RF waves are still "traveling"). Preferably, voltage is applied at the center of the electrodes to accelerate RF dynamics. Compared to TWM theory, where RF velocity β_e and its mismatch with optical velocity $\Delta\beta$ is essential, β_e and $\Delta\beta$ play no role in lumped modulators. Instead, the optical wave interacts with uniform voltage and the interaction time inside lumped electrodes now exerts substantial impact on modulation speed. Furthermore, impedance mismatch is due to the difference in characteristic impedance between source/termination and traveling-wave electrodes in TWMs, and total impedance in lumped modulators, which is both physically and quantitatively different. The latter usually cause stronger mismatch and reflection as a result of very low resistance (normally <20 Ω) and high reactance (on the order of pF). From the energy perspective, the RF wave is severely attenuated on traveling-wave electrodes and dissipated upon termination, while more energy-efficient capacitor charging/discharging is dominant in lumped modulators.

"Double-drive" push–pull configuration requires two driving signals for the two electrically isolated arms widely used in TWMs with GSGSG electrodes [7–9,25]. For double-drive push–pull driving, the RF voltages applied to the two arms have a relative π-phase difference [Fig. 1(a)]. In "single-drive" push–pull configuration, the two arms are electrically connected and only one driving signal is needed [Fig. 1(a)] [12,26,27]. The consequence is half the voltage on each arm, half the total capacitance, and double the series resistance. "Mach–Zehnder" modulators (MZMs) and "Michelson" modulators (MIMs) are named by their optical interferometers. Light travels forward and backward in MIMs, and the equivalent modulation length

is doubled. It is obvious that MIMs are only effective for lumped devices without directionality, and will severely degrade the dynamic response of TWMs [28].

3. EO SMALL-SIGNAL ANALYSIS

Small-signal analysis assumes small voltage on junction so that the EO response is linear [$V_{eff} \ll V_B - V_b$, with V_{eff}, V_B, and V_b being the actual voltage on each junction, junction built-in potential, and bias voltage, respectively, per Eq. (A10)]. The theory is described in the following three parts: electrical, PN junction, and optical.

Electrical description: Voltage is considered to be instantaneously uniform along the lumped electrodes. The resistance–inductance–capacitance (RLC) equivalent circuit model has been reported to effectively describe the electrical properties of lumped modulators [8,29], shown in Fig. 1(a) inset, where C_1, R_1, and L_1 are junction capacitance, series resistance, and electrode inductance, respectively, and C_0 is the pad parasitic capacitance. Due to impedance mismatch with input characteristic impedance, the RF voltage transmission Γ_t can be written as

$$\Gamma_t(\omega) = \frac{2Z_l}{Z_l + Z_0}$$
$$\frac{1}{Z_l} = \frac{1}{j\omega L_1 + \frac{1}{j\omega C_1} + R_1} + \frac{1}{\frac{1}{j\omega C_0}}, \quad \textbf{(1)}$$

where Z_l represents the ($R_1 L_1 C_1 \| C_0$) load, Z_0 is the input characteristic impedance, and ω is the angular frequency of the electric signal. The system transfer function (voltage on junction C_1) is expressed as

$$H(j\omega) = \frac{\frac{1}{j\omega C_1}}{j\omega L_1 + R_1 + \frac{1}{j\omega C_1}}. \quad \textbf{(2)}$$

Relationship to input voltage is given by

$$V_{eff}(\omega) = \frac{1}{2} V_{in} \Gamma_t(\omega) H(j\omega), \quad \textbf{(3)}$$

where V_{eff} and V_{in} represent the voltage amplitude on each PN junction and of the input, respectively. The coefficient is 1/2 for single drive and 1 for double drive.

PN junction description: Junctions are completely described by its capacitance C_1, resistance R_1, and $\Delta n_{eff}(V)$, and hence any types of junction or doping models are applicable. As an example, lateral junction is used (see Appendix A):

$$\Delta n_{eff}(V) = K(AN + BN^{0.8})$$
$$\times \left[w_D(V) - w_D(0) + \frac{\sin(k_x w_D(V)) - \sin(k_x w_D(0))}{k_x} \right],$$
$$w_D(V) = 2\sqrt{\frac{\varepsilon}{q}\frac{1}{N}(V_B - V)}, \quad \textbf{(4)}$$

where N is the doping concentration for both P and N types, $A = 8.8 \times 10^{-22}$ cm^3 and $B = 8.5 \times 10^{-18}$ cm$^{2.4}$ are plasma dispersion effect coefficients [30], w_D is the depletion width, and ε and q are silicon permittivity and elementary charge. K and k_x are constants independent of voltage and doping concentration. For 450 nm × 220 nm rib waveguides, $K = 7.14 \times 10^5$ m^{-1} and $k_x = 4.46 \times 10^6$ m^{-1} (Appendix A).
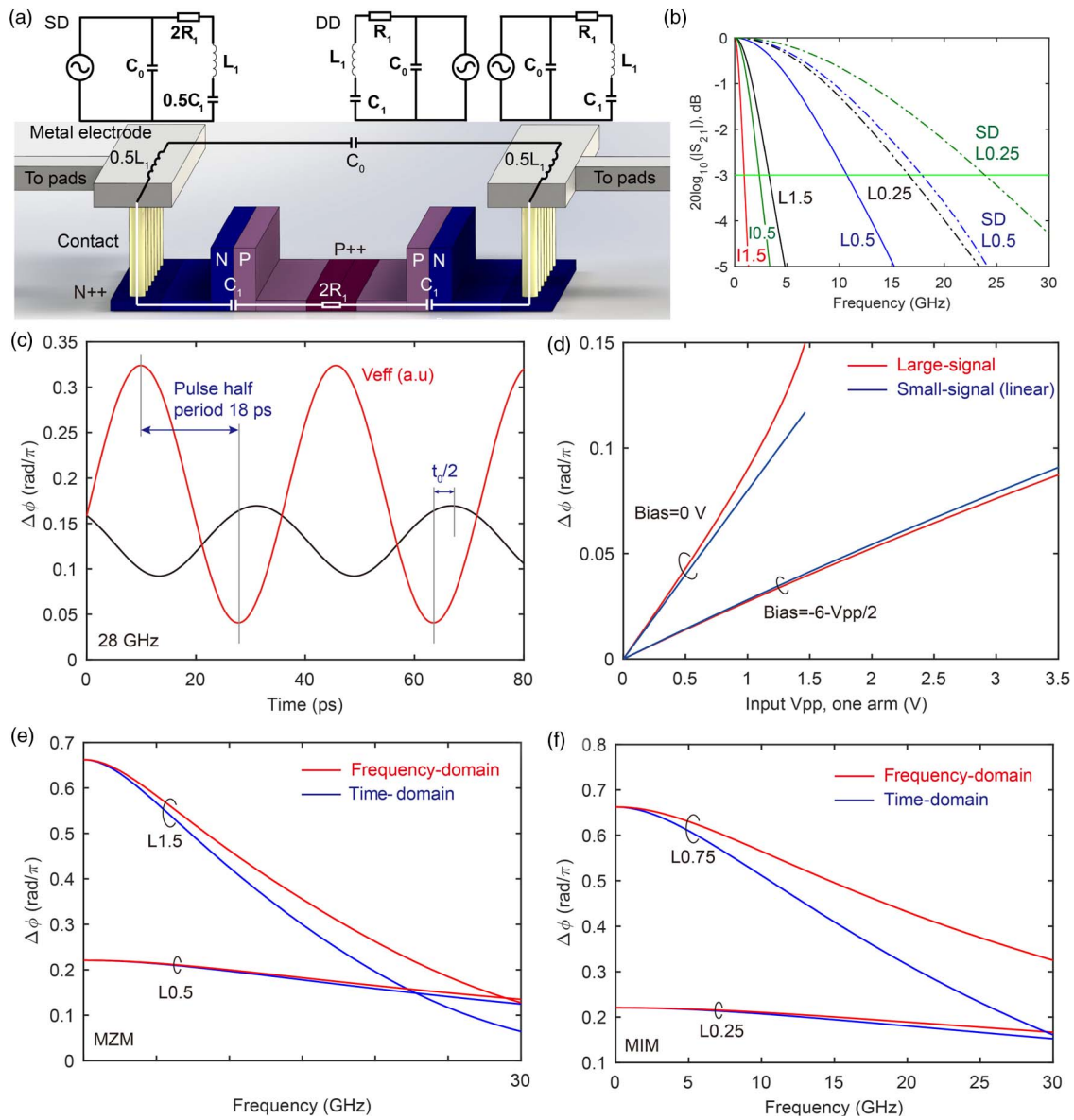
Fig. 1.   Schematic and the basic properties of single-drive lumped silicon modulator. (a) Schematic of single-drive configuration. Inset: equivalent circuit model of single-drive and double-drive configurations. (b) Small-signal EO response of double-drive configuration in literature and single-drive configuration in this paper. Equivalent circuit parameters are from Ref. [8]. "L" stands for lateral junction and "I" stands for interleaved junction. For example, I0.5 and L0.5 represent interleaved junction and lateral junction of 0.5 mm length. (c) Effective voltage on each junction (V_{eff}) and phase change in time domain, calculated at 28 GHz, MZM type. (d) Comparison between small-signal and large-signal models. Time-domain large-signal analysis and frequency-domain approximation for (e) MZM and (f) MIM. Calculation uses $V_\pi L_\pi = 1 \text{ V} \cdot \text{cm}$ (at −1 V). SD: single drive; DD: double-drive.

Optical description: Optical phase changes as a result of real-time interaction with junction voltage:

$$\Delta\phi(t) = k_0 v_g \int_{t-t_0}^{t} \Delta n_{eff}(V_{eff}(t)) \mathrm{d}t, \qquad (5)$$

where $k_0 = 2\pi/\lambda$, v_g represents the optical group velocity, and the related time inside the phase shifter is $t_0 = L/v_g$ for MZM and $t_0 = 2L/v_g$ for MIM. Here, no velocity mismatch term is present since the electrical signal is uniform on the electrodes. Combining Eqs. (4) and (5) and expanding the sinusoidal term to the first order when N remains unchanged yields

$$\Delta\phi(t) = K'\left[-t_0 \int_{t-t_0}^{t} \sqrt{1 - \frac{V_{eff}(t)}{V_B}} \, \mathrm{d}t\right]$$

$$\approx -\frac{K'}{2V_B}\left[\int_{t-t_0}^{t} V_{eff}(t)\mathrm{d}t\right], \qquad (6)$$

where constant K' is given as

$$K' = 4K\sqrt{V_B}k_0 v_g \sqrt{\frac{\varepsilon}{qN}}(AN + BN^{0.8}). \qquad (7)$$

A sinusoidal wave of angular frequency ω_s used in small-signal analysis will yield the following voltage on junctions:

$$V_{\text{eff}}(t) = \text{Re}\, F^{-1}\{F[V_{\text{in}} \exp(j\omega_s t)]H(j\omega)\Gamma_t(\omega)\}$$

$$= \text{Re}\, V_{\text{in}}H(j\omega_s)\Gamma_t(\omega_s)\exp(j\omega_s t), \qquad \textbf{(8)}$$

where F stands for Fourier transform. Frequency-domain EO S_{21} (biased at quadrature) is defined as

$$S_{21}(\omega_s) = F[E_{\text{out}}(t)] \approx -\frac{1}{2} \cdot F[\Delta\phi(t)]. \qquad \textbf{(9)}$$

Combining Eqs. (6), (8), and (9) yields

$$S_{21}(\omega) = K' \frac{\pi V_{\text{in}}}{4 V_B} \frac{1 - \exp(-j\omega t_0)}{j\omega} H(j\omega)\Gamma_t(\omega). \qquad \textbf{(10)}$$

Equation (10) gives the analytical model of the EO small-signal analysis. To confirm its validity, we first use this theory to reproduce experimentally measured S_{21} in published literature, where equivalent circuit parameters were given [8]. The results are shown in Fig. 1(b). "L0.5" and "I0.5" represent lateral junction and interleaved junction of 0.5 mm length. The left four lines show 3 dB bandwidth of [0.93, 2.4, 3.3, 10.7] GHz, while the experimentally measured values are [1.0, 2.6, 3.6, 10.0] GHz, showing good accordance with the analytical model. Similar agreement has also been achieved for baseline lumped modulators provided by some commercial foundries. Based on the established validity, the single-drive configuration is analyzed. Compared with the lateral 0.5 mm long device (blue line, 10.7 GHz bandwidth), simply employing single drive boosts its bandwidth to 17.9 GHz. Cutting device length in half (MIM) also helps, showing 16.6 GHz bandwidth with conventional double-drive configuration and 23.9 GHz bandwidth in combination with single drive. The bandwidth limitation is definitely not resistance–capacitance (RC) constant limited (exceeding 40 GHz). Equation (10) shows (Fig. 2) that the main contributor is Γ_t due to impedance mismatch, which is especially severe for double-drive modulators. Note that source characteristic impedance does not affect the equivalent circuit (RC constant) but solely the RF mismatch. To mitigate this issue, 50% reduction of capacitance brought by single drive allows a higher percentage of voltage to transmit to phase shifters and consequently enhances the high-frequency response.

For this reason, only single-drive lumped modulators will be evaluated in the large-signal modeling that follows.

4. EO LARGE-SIGNAL ANALYSIS

Large-signal modeling is similar to small-signal modeling except that Taylor expansion to the first order is no longer accurate. Additionally, sinusoidal terms in Eq. (4) are retained to facilitate analysis involving doping concentration. Silicon conductivity $\sigma(N)$ dependence is included in the calculation. The time-domain phase change is given by Eq. (5) and the peak-to-peak phase change $\Delta\varphi(\omega) = \max[\Delta\varphi(t)] - \min[\Delta\varphi(t)]$ is chosen to represent the modulation performance at a certain frequency. In push–pull driving, the total phase difference is thus $2\Delta\varphi$.

To quantitatively show the difference between small-signal and large-signal models, Fig. 1(d) depicts the two curves at two extreme bias conditions. The analytical expression for small-signal phase shift is provided in Eq. (A10). The large-signal model exhibits a hyperlinear relationship at zero bias (also known as differential drive), and is sublinear while operating at high reverse voltage. This can be understood by the improved modulation effect near the $0 - V_B$ voltage region and saturation [square root term in Eq. (4)] at high voltage, reflecting the modulation properties of PN junctions. From the small-signal derivation, the premise of Taylor expansion is $V_{\text{eff}} \ll V_B - V_b$, V_b being the bias voltage. Therefore, the small-signal approximation is valid in a wider V_{eff} range under reverse bias. This comparison further consolidates the validity of the large-signal model.

Another approximation contributes to the further understanding of the EO interaction. When the modulation speed is sufficiently low for RF voltage to be seen as invariant for a light pulse, frequency-domain approximation is valid:

$$\Delta\phi(\omega) = k_0 \int_0^L \Delta n_{\text{eff}}(V_{\text{eff}}(\omega))\mathrm{d}l = k_0 L \Delta n_{\text{eff}}(V_{\text{eff}}(\omega)).$$

$$\textbf{(11)}$$

Figures 1(e) and 1(f) illustrate large-signal analysis of MZM and MIM. The blue and red lines coincide initially and discrepancy grows as frequency increases, as predicted from the assumption of Eq. (11). The discrepancy gets smaller for
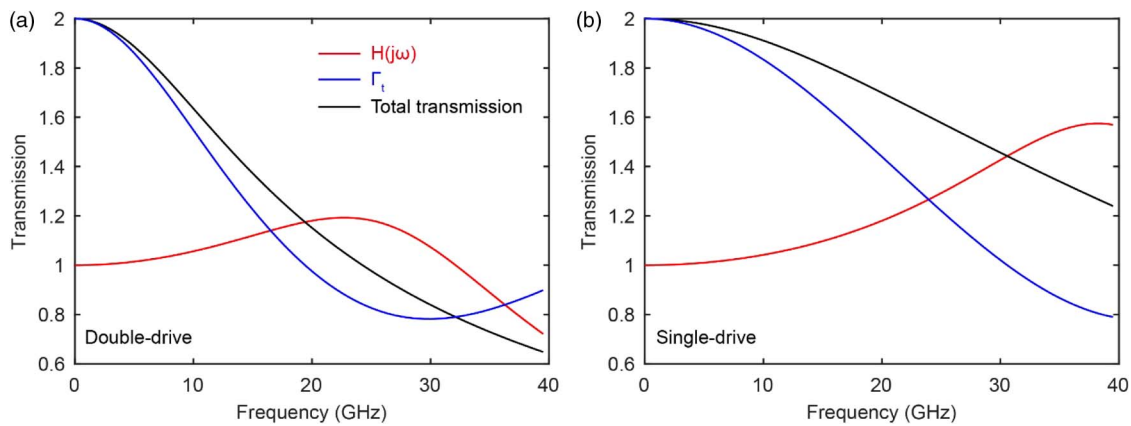

Fig. 2. Frequency response of lumped modulators of (a) double-drive and (b) single-drive. Red, blue, and black lines represent the modulus of system transfer function $H(j\omega)$, voltage transmission due to RF reflection Γ_t, and their product. Calculation uses $V_\pi L_\pi = 1$ V · cm.

shorter devices. It indicates that the real-time nature of applied voltage $V_{in}(t)$ during light traveling through the device is substantial, and the upper limit of device length is determined by symbol rate. For 28 Gbd modulation, the corresponding pulse half-period is 18 ps (1.3 mm physical length). Devices longer than this will suffer from crosstalk exerted by the succeeding RF pulse. Figure 1(c) shows this process in the time domain, where V_{eff} and $\Delta\varphi$ have a t_0-related retardation originating from the integration interval. This mechanism in lumped modulators is unique as opposed to velocity-matched TWMs. Considering t_0 alone, the closer it is to the upper limit, the larger the integral interval. However, the rise/fall time will be shown later to grow with device length and weaken the modulation effect, causing a tradeoff. From this perspective, the small-signal normalized S_{21}, although extensively reported in the literature, is not

instructive regarding accumulated modulation effect nor high-speed capabilities because the S_{21} bandwidth monotonously rises with decreasing device length while $\Delta\varphi$ does not (and is likely to be single-peaked with L). Figures 1(e) and 1(f) also reflect steeper roll-off for long phase shifters. Despite their impressive DC modulation effect, the high-speed performance is by no means favorable. Comparatively, due to shorter physical length, MIM exhibits more gradual roll-off, rendering MIM more suitable for high-speed operation.

5. ENERGY CONSUMPTION INVESTIGATION

We first specify the operation condition for energy analysis. Electrical bias voltage is chosen to utilize the efficient modulation range of the PN junction ($0 - -V_{pp,eff}$, Fig. 3 caption),
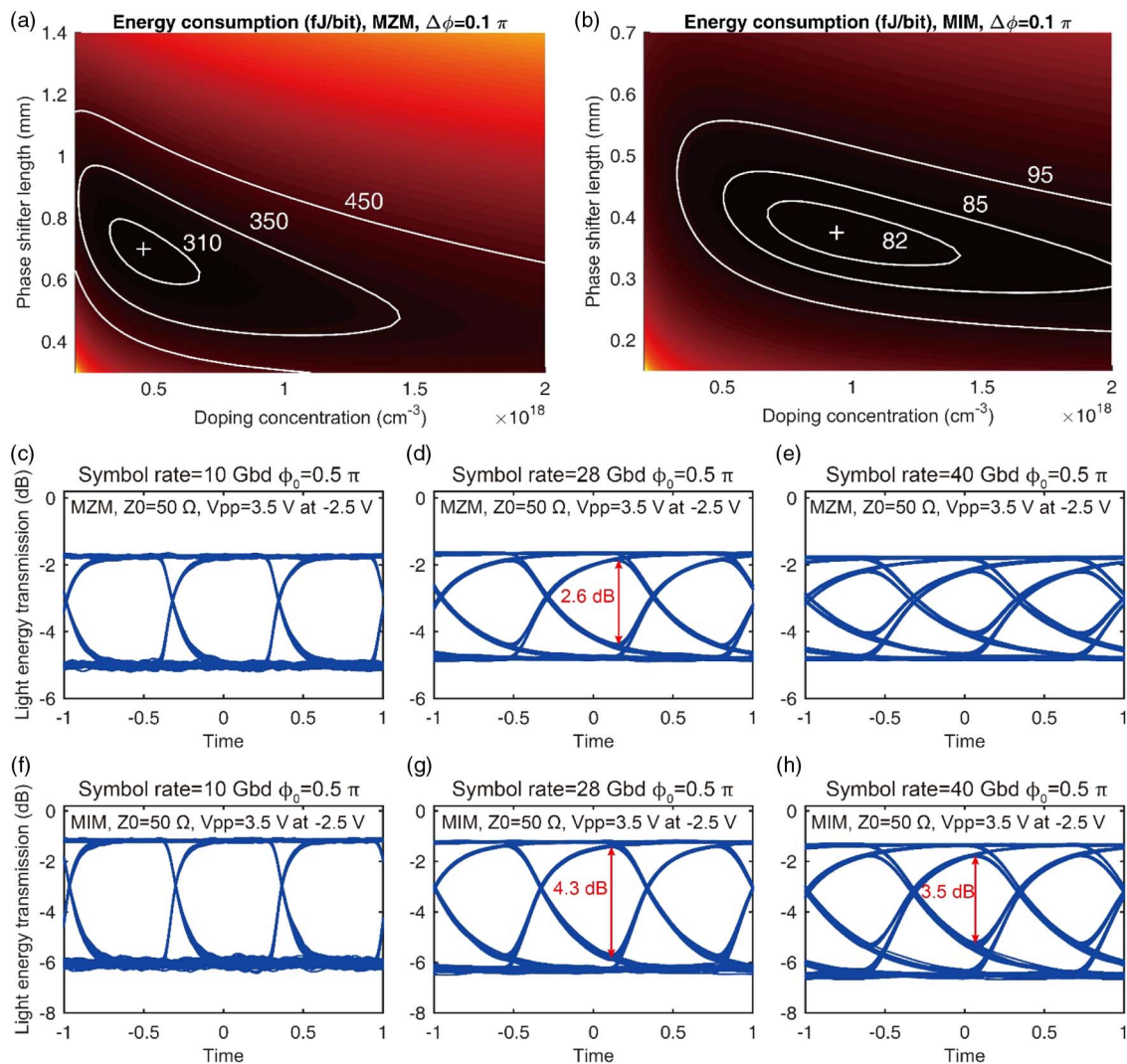

Fig. 3. Large-signal characterization under 50 Ω standard characteristic impedance driver. Energy consumption relation with phase shifter length and doping concentration of (a) MZM and (b) MIM. The shown energy is obtained when $\Delta\varphi = 0.1\pi$. Bias voltages are chosen for each structural parameter so that the actual voltage on each PN junction is $0 - -V_{pp,eff}$. The lowest energy point (marked by "+") in (a) is $[L, N, E, V_{in}] = [0.350 \text{ mm}, 4.6 \times 10^{17} \text{ cm}^{-3}, 303.8 \text{ fJ/bit}, 6.12 \text{ V}]$; $[L, N, E, V_{in}] = [0.375 \text{ mm}, 9.4 \times 10^{17} \text{ cm}^{-3}, 80.8 \text{ fJ/bit}, 1.95 \text{ V}]$ in (b). Color represents $\log_{10}(E)$ for better visual contrast. Eye diagrams of single-drive lumped modulators of (c)–(e) MZM and (f)–(h) MIM. The vertical axis is absolute value scale (not normalized). Doping concentration and phase shifter length for MZM and MIM are the optimal values in (a) and (b), respectively. Static working point: $\varphi_0 = 0.5\pi$ (biased at quadrature point). The driving "V_{pp}" specified in the figure is peak-to-peak voltage on each junction. Total input voltage is twice the value.

and the optical operation point is always on quadrature. The energy consumption is obtained by solving the driving voltage from $\Delta\phi = 0.1\pi$ and using the following relation [31]:

$$E = \frac{1}{4}C_1 V_{\text{in,pp}}^2 = C_1 V_{\text{in}}^2. \qquad (12)$$

RF reflection is already accounted for in V_{in}. An alternative method is reported from the perspective of "Q_π", the needed charge for π phase shift [32], where bit energy can be expressed as $E = (1/4)0.1Q_\pi \cdot V_{\text{in}}$ for $\Delta\varphi = 0.1\pi$ and $Q_\pi \approx 5.9$ pC. V_{in} can be solved with the proposed method in this work.

The numerical results for MZM and MIM are illustrated in Figs. 3(a) and 3(b). The vertical axes are chosen so that they have the same equivalent modulation length. For both types, energy consumption contours indicate that low doping concentration allows for a longer phase shifter since the junction capacitance per unit length is smaller. The static insertion loss (IL) is approximately linear with the doping concentration multiplied by the length ($N \cdot L$), and the low-doping side yields lower IL (for more accurate analysis, see Fig. 4). Three different slopes of energy contours are observed in Fig. 3(a) with different limitations. The steep slope to the left (low doping) is mainly $\Delta\phi$–limited, and therefore it rises rapidly with increasing length. For the other two slopes, one can notice that two lengths share the same energy consumption at each doping concentration. The upper contour is therefore capacitance (RF reflection) limited, and the lower one reveals the combined effect of these factors. Intriguingly, the bottom curve is tolerant to doping (Δn_{eff}) deviations while possessing the smallest footprint, desirable in practical applications. Although only one globally lowest energy condition exists at $[L, N, E] = [0.350 \text{ mm}, 4.6 \times 10^{17} \text{ cm}^{-3}, 303.8 \text{ fJ/bit}]$, the $E – L$ relationships are single-peaked at each doping concentration, which makes numerical analysis necessary in finding the appropriate parameters. The MIM energy consumption is much lower, the appropriate doping concentration shifts higher owing to shorter physical length, and $E = 80.8$ fJ/bit is realized. Optical properties of MIM and MZM are expected to be identical because of optical equivalence (MIM being electrically shorter). The only shortcoming of MIM practically is probably that protections against back-direction light might be needed.

The large-signal modeling proposed in this paper also allows us to obtain eye diagrams in absolute scale to reveal the response time or large-signal speed of the device. To better compare MZM and MIM, both modulators are driven by 3.5 V_{pp} at -2 V bias. A random RF square wave with noise is used as excitation source. As shown in Fig. 3, wide-open eyes can be obtained at 10 Gbd while, at 28 Gbd, the rise/fall time becomes longer than the pulse half-period, leading to an inner eye. This phenomenon is further aggravated at 40 Gbd. Comparatively, MIM provides better eye quality and both types are fully usable in 28 Gbd communication systems.

6. DISCUSSION

The speed limitation is known to be the impedance mismatch between input characteristic impedance and the load. If a specially designed, low characteristic impedance IC driver is used, this limitation can be removed [16,17] and RLC-limited performance will be dominant.

Figures 5(a) and 5(b) show the $E – L – N$ relationship when $Z_0 = 10$ Ω, and significant energy reduction up to ~20% compared to $Z_0 = 50$ Ω can be observed. The MZM lowest energy point slightly shifts toward longer device length, while the MIM energy contours undertake greater change, preferring a relatively high doping concentration of 1.98×10^{18} cm^{-3}—a result of the interplay between higher junction EO effect and higher RF reflection. Likewise, the opposite effect can be expected when using junction types with stronger EO effect and higher capacitance (e.g., interleaved junctions). The lowest energy consumption (21.5 fJ/bit at 28 Gbd) is a meaningful step toward ultra-low energy consumption. With further optimization of Z_0 and bias condition, the energy can potentially be reduced to below 10 fJ/bit [33].

MIM eye diagrams are shown in Figs. 5(c)–5(e). The applied voltage has to be reduced to keep a similar extinction ratio to the eye diagrams under 50 Ω driver. Surprisingly, the rise/fall time remains sufficiently short up to 40 Gbd, and only shows slight degradation at 50 Gbd speed. Such findings indicate that, with a properly designed IC driver, lumped modulators can well meet the requirement for 40–50 Gbd high-speed applications, as demonstrated by Cisco's SISCAP (MOS-type) modulators with dedicated IC drivers [16,34]. $Z_0 = 10$ Ω is just a proof-of-concept demonstration. Optimization of Z_0 and the overhead of the electric circuit remains the subject of further investigation.

Loss is another aspect of the analysis, but practically static IL is already much smaller in lumped modulators because of their short length, and is usually less than the dynamic loss incurred by inadequate voltage swing. Quantitative analysis of static IL is provided in Fig. 4.

Finally, we briefly comment on the larger phase-shift scenario and the significance to reverse-design in engineering. This work analyzes $\Delta\varphi = 0.1\pi$ per arm. For larger phase shift, the optimal length is expected to be longer since PN junction response is inefficient for large voltage. Additionally, low doping concentration (e.g., $N \leq 2 \times 10^{17}$ cm^{-3}) may not be
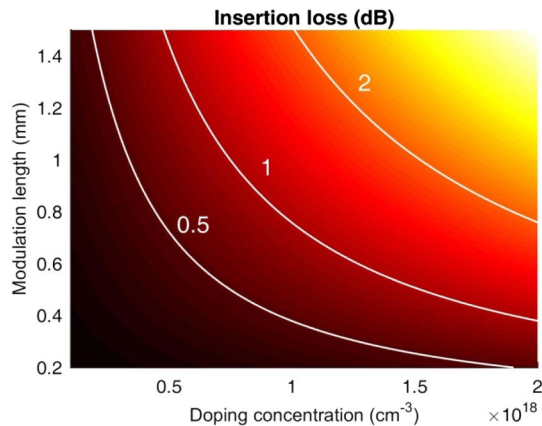

Fig. 4. Static IL under zero bias in the parameter space of interest. Modulation length is equal to the physical length of MZM and twice the physical length of MIM. The data used to obtain the coefficient is 1.12 dB/mm ($N = 1 \times 10^{18}$ cm^{-3}), not including propagation loss of the passive waveguide (typical value 0.2 dB/mm).
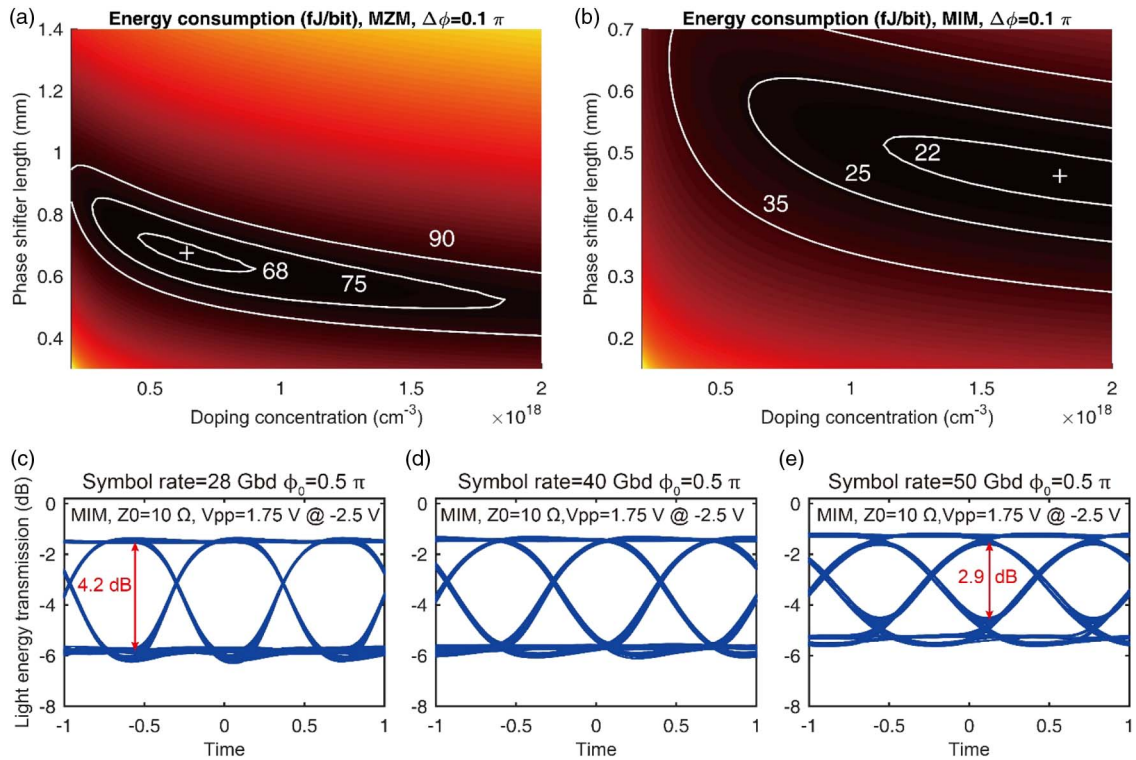
Fig. 5. Large-signal characterization under 10 Ω low-characteristic impedance driver. Energy consumption relation with phase shifter length and doping concentration of (a) MZM and (b) MIM. The shown energy is obtained when $\Delta\varphi = 0.1\pi$. Bias voltages are chosen for each structural parameter so that the effective voltage on each PN junction is $0 - -V_{pp,eff}$. The lowest energy point (marked by "+") in (a) is $[L, N, E, V_{in}] = [0.675 \text{ mm}, 6.4 \times 10^{17} \text{ cm}^{-3}, 66.9 \text{ fJ/bit}, 1.44 \text{ V}]$; $[L, N, E, V_{in}] = [0.4625 \text{ mm}, 1.80 \times 10^{18} \text{ cm}^{-3}, 21.5 \text{ fJ/bit}, 0.774 \text{ V}]$ in (b). (c)–(e) Eye diagrams of MIM modulators of the optimal parameters. Color represents $\log_{10}(E)$ for better visual contrast. Static working point: $\varphi_0 = 0.5\pi$ (biased at quadrature point). The driving "V_{pp}" specified in the figure is peak-to-peak voltage on each junction. Total input voltage is twice the value.

capable of large phase shift with any device length (limited by symbol rate). This work establishes the connection between structural and performance parameters, thus allowing reverse-design, one of the bases for customizable process design kits. The unimodal dependence in Fig. 5 also makes the reverse design computationally efficient.

7. CONCLUSIONS

This work proposes a complete set of theory and modeling for lumped silicon modulators. Validation is established through comparison with published experimental results. The roles of device length and doping concentration are analyzed. Our results suggest that PN junctions in lumped modulators cannot be separately characterized independent of device length, as opposed to the case of TWMs. The lowest energy achievable for MIM is 80.8 fJ/bit at 28 Gbd under a standard 50 Ω characteristic impedance driver, and 21.5 fJ/bit under a 10 Ω driver. Additionally, the quantitative model allows us to obtain eye diagrams with absolute scale values. Both MZMs and MIMs are found to be operable at 28 Gbd (50 Ω driver) and MIMs can satisfy 40–50 Gbd modulation under a lower impedance driver. This important step forward is the result of the proposed theory in two ways: the speed limitation is understood and alleviated with a single-drive method accordingly, and comprehensive

$E–L–N$ analysis enabled by the large-signal modeling reveals the roles of each parameter and yields the lowest reachable energy. Such findings are meaningful toward the development of 10 fJ/bit, high-speed datacom systems.

APPENDIX A

Analytical expression for Δn_{eff}. This section determines the effective refractive change in depletion-type silicon lateral PN junctions. Two variables are considered: $\Delta n_{eff}(V, N)$, where V is applied voltage and N is doping concentration. The following assumption is used to simplify the case: P-type and N-type doping concentration is the same ($N_A = N_D = N$), and the junction is placed at the center of the waveguide. The latter assumption is reasonable since the optimal offset is quite close to zero when $N_A = N_D$, e.g., optimal offset = 10 nm when $N = 5 \times 10^{17}$ cm^{-3}.

According to the definition, n_{eff} is written as

$$n_{eff} = \frac{\iint_{\infty} n(x, y)|E(x, y)|^2 dS}{\iint_{\infty} |E(x, y)|^2 dS}$$
$$= \frac{\iint_{WG} n(x, y)|E(x, y)|^2 dS}{E_0^2} + C, \qquad \textbf{(A1)}$$

where C represents the integral outside of the waveguide core, which is independent of voltage and doping concentration.

E_0 represents a normalizing constant. The origin is set at waveguide center. To obtain an analytical expression, the rib waveguide is approximated with the strip waveguide, where the electric field inside the waveguide can be written as

$$E(x, y) = \cos(k_x x) \cos(k_y y),\qquad \textbf{(A2)}$$

where k_x and k_y are the optical propagation constants in transverse directions.

For lateral junctions, the material index inside waveguide $n(x, y)$ is uniform along the y direction, and hence can be reduced to $n(x)$

$$
n_{\mathrm{eff}} = \frac{1}{E_0^2} \int_{-h/2}^{h/2} \left[\int_{-w_D/2}^{w_D/2} n_{\mathrm{Si}} \cos^2(k_x x) \mathrm{d}x \right.
$$
$$
\left. + 2 \int_{w_D/2}^{w/2} n_{PN} \cos^2(k_x x) \mathrm{d}x \right] \cos^2(k_y y) \mathrm{d}y + C, \quad \textbf{(A3)}
$$

where h is the height of the waveguide, $w_D(V)$ is depletion width, w is waveguide width, n_{Si} is intrinsic silicon refractive index, and n_{PN} is the average silicon refractive index of P-type and N-type silicon. Integrating and collecting the terms relating to w_D yields

$$
n_{\mathrm{eff}} \propto \Delta n \left[w_D + \frac{\sin(k_x w_D)}{k_x} \right]
$$
$$
+ 2 \left[C' + n_{PN} w + n_{PN} \frac{\sin(k_x w)}{k_x} \right], \qquad \textbf{(A4)}
$$

where n_{PN} is given by the plasma dispersion effect

$$
\Delta n = n_{\mathrm{Si}} - n_{PN} = \frac{A N_D + B N_A^{0.8}}{2}. \qquad \textbf{(A5)}
$$

Here, coefficients $A = 8.8 \times 10^{-22}$ cm³ and $B = 8.5 \times 10^{-18}$ cm$^{2.4}$ (at 1550 nm wavelength) are only dependent on wavelength. The change of effective index, $\Delta n_{\mathrm{eff}}(V) = n_{\mathrm{eff}}(V) - n_{\mathrm{eff}}(0)$, can therefore be written as

$$
\Delta n_{\mathrm{eff}}(V) = K(AN + BN^{0.8})
$$
$$
\times \left[w_D(V) - w_D(0) + \frac{\sin(k_x w_D(V)) - \sin(k_x w_D(0))}{k_x} \right]
$$
$$
w_D(V) = 2 \sqrt{\frac{\varepsilon}{q} \frac{1}{N} (V_B - V)}, \qquad \textbf{(A6)}
$$

where V_B is the built-in potential of the PN junction. This derivation is under the premise that PN junction-induced material refractive index change causes only perturbation to the optical field distribution, so that constants C, k_x, and K are independent of V and N. The assumption is true for silicon PN junctions since material refractive index change is only on the order of 10^{-3}, causing less than 1% confinement factor change.

Equation (A6) is the analytical model, with variables being V and N. Fitting to simulation results (verified with published experimental results without strip waveguide approximation) shows that $K = 7.138 \times 10^5$ m^{-1} and $k_x = 4.46 \times 10^6$ m^{-1} can give perfect matching (waveguide dimension: 450 nm width, 220 nm height, junction interface located in the center of the waveguide). Likewise, optical absorption can be derived and is shown in Fig. 4.

Small-signal phase change. This section gives the analytical form of small-signal phase change. Based on Eq. (A6), substituting the expression of V_{eff} yields

$$
\Delta\phi(t) = \mathrm{Re} \, \frac{K' V_{\mathrm{eff}}}{2 V_B} \frac{\exp(j\omega t)[1 - \exp(-j\omega t_0)]}{j\omega}. \qquad \textbf{(A7)}
$$

Still using the definition of $\Delta\phi(\omega) = \max[\Delta\phi(t)] - \min[\Delta\phi(t)]$, Eq. (A7) becomes

$$
\Delta\phi(\omega) = \frac{K' V_{\mathrm{eff}}}{V_B \omega} \sin\left(\omega \frac{t_0}{2} \right). \qquad \textbf{(A8)}
$$

The above-mentioned phase shift is under zero bias (differential drive). We can also expand it to the non-zero bias case,

$$
\Delta\phi(t) = \mathrm{Re}\, K' \sqrt{1 - \frac{V_b}{V_B}} \left\{ \left(1 - \left(1 - \frac{V_b}{V_B} \right)^{-1/2} \right) t_0 \right.
$$
$$
\left. - \frac{1}{2} \frac{V_{\mathrm{eff},ac}(t)}{V_B - V_b} \frac{\exp(j\omega_s t)}{j\omega_s} [1 - \exp(-j\omega_s t_0)] \right\}, \qquad \textbf{(A9)}
$$

where V_b is bias voltage and $V_{\mathrm{eff},ac}$ is the high-frequency component of V_{eff}. A constant (non-zero DC component) is present here because of our definition of $\Delta n_{\mathrm{eff}}(V) = n_{\mathrm{eff}}(V) - n_{\mathrm{eff}}(0)$, and naturally $\Delta n_{\mathrm{eff}}(V_b) \neq 0$. For $\Delta\varphi(\omega)$, it is independent of this constant:

$$
\Delta\phi(\omega) = \frac{K' V_{\mathrm{eff},ac}}{\omega \sqrt{V_B (V_B - V_b)}} \sin\left(\omega \frac{t_0}{2} \right). \qquad \textbf{(A10)}
$$

Essentially, small-signal analysis assumes the linearity of the system response and therefore will inevitably deviate from the true value when system nonlinearity rises. For example, the comparison between small-signal and large-signal theories in the main text shows discordance when bias voltage nears 0 V and is strongly reversely biased, which corresponds to the nonlinear response of $n_{\mathrm{eff}}(V)$. The accordance between large- and small-signal models at sufficiently small voltage is also well explained by the underlying physics and provides validity for the large-signal model.

REFERENCES

1. D. Miller, "Device requirements for optical interconnects to silicon chips," Proc. IEEE **97**, 1166–1185 (2009).
2. X. Li, Q. Deng, and Z. Zhou, "Low loss, high-speed single-mode half-disk resonator," Opt. Lett. **39**, 3810–3813 (2014).
3. J. C. Rosenberg, W. M. Green, S. Assefa, D. M. Gill, and T. Barwicz, "A 25 Gbps silicon microring modulator based on an interleaved junction," Opt. Express **20**, 26411–26423 (2012).
4. S. Manipatruni, K. Preston, L. Chen, and M. Lipson, "Ultra-low voltage, ultra-small mode volume silicon microring modulator," Opt. Express **18**, 18235–18242 (2010).
5. E. Timurdogan, C. M. Sorace-Agaskar, J. Sun, E. Shah Hosseini, A. Biberman, and M. R. Watts, "An ultralow power athermal silicon modulator," Nat. Commun. **5**, 4008 (2014).
6. H. Xu, X. Li, X. Xiao, Z. Li, Y. Yu, and J. Yu, "Demonstration and characterization of high-speed silicon depletion-mode Mach–Zehnder modulators," IEEE J. Sel. Top. Quantum Electron. **20**, 23–32 (2014).
7. Y. Yang, Q. Fang, M. Yu, X. Tu, R. Rusli, and G. Lo, "High-efficiency Si optical modulator using Cu travelling-wave electrode," Opt. Express **22**, 29978–29985 (2014).

8. H. Yu, M. Pantouvaki, J. Van Campenhout, D. Korn, K. Komorowska, P. Dumon, Y. Li, P. Verheyen, P. Absil, L. Alloatti, D. Hillerkuss, J. Leuthold, R. Baets, and W. Bogaerts, "Performance tradeoff between lateral and interdigitated doping patterns for high-speed carrier-depletion based silicon modulators," Opt. Express **20**, 12926–12938 (2012).

9. A. Brimont, D. J. Thomson, F. Y. Gardes, J. M. Fedeli, G. T. Reed, J. Martí, and P. Sanchis, "High-contrast 40 Gb/s operation of a 500 μm long silicon carrier-depletion slow wave modulator," Opt. Lett. **37**, 3504–3506 (2012).

10. H. Yu and W. Bogaerts, "An equivalent circuit model of the traveling wave electrode for carrier-depletion-based silicon optical modulators," J. Lightwave Technol. **30**, 1602–1609 (2012).

11. R. Becker, "Broad-band guided-wave electrooptic modulators," IEEE J. Quantum Electron. **20**, 723–727 (1984).

12. J. Ding, H. Chen, L. Yang, L. Zhang, R. Ji, Y. Tian, W. Zhu, Y. Lu, P. Zhou, R. Min, and M. Yu, "Ultra-low-power carrier-depletion Mach–Zehnder silicon optical modulator," Opt. Express **20**, 7081–7087 (2012).

13. T. Baehr-Jones, R. Ding, Y. Liu, A. Ayazi, T. Pinguet, N. Harris, M. Streshinsky, P. Lee, Y. Zhang, A. Lim, T. Liow, S. Teo, G. Lo, and M. Hochberg, "Ultralow drive voltage silicon traveling-wave modulator," Opt. Express **20**, 12014–12020 (2012).

14. J. Ding, R. Ji, L. Zhang, and L. Yang, "Electro-optical response analysis of a 40 Gb/s silicon Mach–Zehnder optical modulator," J. Lightwave Technol. **31**, 2434–2440 (2013).

15. M. R. Watts, W. A. Zortman, D. C. Trotter, R. W. Young, and A. L. Lentine, "Low-voltage, compact, depletion-mode, silicon Mach–Zehnder modulator," IEEE J. Sel. Top. Quantum Electron. **16**, 159–164 (2010).

16. M. A. Webster, K. Lakshmikumar, C. Appel, C. Muzio, B. Dama, and K. Shastri, "Low-power MOS-capacitor based silicon photonic modulators and CMOS drivers," in *Optical Fiber Communication Conference*, OSA Technical Digest (Optical Society of America, 2015), paper W4H.3.

17. D. Patel, V. Veerasubramanian, S. Ghosh, A. Samani, Q. Zhong, and D. V. Plant, "High-speed compact silicon photonic Michelson interferometric modulator," Opt. Express **22**, 26788–26802 (2014).

18. S. Koeber, R. Palmer, M. Lauermann, W. Heni, D. L. Elder, D. Korn, M. Woessner, L. Alloatti, S. Koenig, P. C. Schindler, H. Yu, W. Bogaerts, L. R. Dalton, W. Freude, J. Leuthold, and C. Koos, "Femtojoule electro-optic modulation using a silicon-organic hybrid device," Light Sci. Appl. **4**, e255 (2015).

19. L. Alloatti, R. Palmer, S. Diebold, K. P. Pahl, B. Chen, R. Dinu, M. Fournier, J. Fedeli, T. Zwick, W. Freude, C. Koos, and J. Leuthold, "100 GHz silicon-organic hybrid modulator," Light Sci. Appl. **3**, e173 (2014).

20. J. Clark and G. Lanzani, "Organic photonics for communications," Nat. Photonics **4**, 438–446 (2010).

21. C. Haffner, W. Heni, Y. Fedoryshyn, J. Niegemann, A. Melikyan, D. L. Elder, B. Baeuerle, Y. Salamin, A. Josten, U. Koch, C. Hoessbacher, F. Ducry, L. Juchli, A. Emboras, D. Hillerkuss, M. Kohl, L. R. Dalton, C. Hafner, and J. Leuthold, "All-plasmonic Mach–Zehnder modulator enabling optical high-speed communication at the microscale," Nat. Photonics **9**, 525–528 (2015).

22. A. Melikyan, L. Alloatti, A. Muslija, D. Hillerkuss, P. C. Schindler, J. Li, R. Palmer, D. Korn, S. Muehlbrandt, D. Van Thourhout, B. Chen, R. Dinu, M. Sommer, C. Koos, M. Kohl, W. Freude, and J. Leuthold, "High-speed plasmonic phase modulators," Nat. Photonics **8**, 229–233 (2014).

23. B. Pile and G. Taylor, "Small-signal analysis of microring resonator modulators," Opt. Express **22**, 14913–14928 (2014).

24. H. Yu, D. Ying, M. Pantouvaki, J. Van Campenhout, P. Absil, Y. Hao, J. Yang, and X. Jiang, "Trade-off between optical modulation amplitude and modulation bandwidth of silicon micro-ring modulators," Opt. Express **22**, 15178–15189 (2014).

25. T. Li, J. Zhang, H. Yi, W. Tan, Q. Long, Z. Zhou, X. Wang, and H. Wu, "Low-voltage, high speed, compact silicon modulator for BPSK modulation," Opt. Express **21**, 23410–23415 (2013).

26. J. Wang, L. Zhou, H. Zhu, R. Yang, Y. Zhou, L. Liu, T. Wang, and J. Chen, "Silicon high-speed binary phase-shift keying modulator with a single-drive push-pull high-speed traveling wave electrode," Photon. Res. **3**, 58–62 (2015).

27. L. Chen, P. Dong, and Y.-K. Chen, "Chirp and dispersion tolerance of a single-drive push-pull silicon modulator at 28 Gb/s," IEEE Photon. Technol. Lett. **24**, 936–938 (2012).

28. X. Li, X. Xiao, H. Xu, Z. Li, T. Chu, J. Yu, and Y. Yu, "Highly efficient silicon Michelson interferometer modulators," IEEE Photon. Technol. Lett. **25**, 407–409 (2013).

29. D. Samara-Rubio, U. D. Keil, L. Liao, T. Franck, A. Liu, D. W. Hodge, D. Rubin, and R. Cohen, "Customized drive electronics to extend silicon optical modulators to 4 GB/s," J. Lightwave Technol. **23**, 4305–4314 (2005).

30. G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," Nat. Photonics **4**, 518–526 (2010).

31. G. T. Reed, G. Z. Mashanovich, F. Y. Gardes, M. Nedeljkovic, Y. Hu, D. J. Thomson, K. Li, P. R. Wilson, S. Chen, and S. S. Hsu, "Recent breakthroughs in carrier depletion based silicon optical modulators," Nanophotonics **3**, 229–245 (2014).

32. T. Usuki, "Robust optical data transfer on silicon photonic chip," J. Lightwave Technol. **30**, 2933–2940 (2012).

33. Z. Zhou, B. Yin, Q. Deng, X. Li, and J. Cui, "Lowering the energy consumption in silicon photonic devices and systems [invited]," Photon. Res. **3**, B28–B46 (2015).

34. M. Webster, P. Gothoskar, V. Patel, D. Piede, S. Anderson, R. Tummidi, D. Adams, C. Appel, P. Metz, S. Sunder, B. Dama, and K. Shastri, "An efficient MOS-capacitor based silicon modulator and CMOS drivers for optical transmitters," in *IEEE 11th International Conference on Group IV Photonics (GFP)*, August 27–29 (2014), pp. 1–2.