Nanophotonic lithium niobate electro-optic modulators

CHENG WANG,1,4 MIAN ZHANG,1,4 BRIAN STERN,2,3 MICHAL LIPSON,3 AND MARKO LONČAR1,*

1John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
2School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
3Department of Electrical Engineering, Columbia University, New York, NY 10027, USA
4These authors contributed equally to this work
*loncar@seas.harvard.edu

Abstract: Since the emergence of optical fiber communications, lithium niobate (LN) has been the material of choice for electro-optic modulators, featuring high data bandwidth and excellent signal fidelity. Conventional LN modulators however are bulky, expensive and power hungry, and cannot meet the growing demand in modern optical data links. Chip-scale, highly integrated, LN modulators could offer solutions to this problem, yet the fabrication of low-loss devices in LN thin films has been challenging. Here we overcome this hurdle and demonstrate monolithically integrated LN electro-optic modulators that are significantly smaller and more efficient than traditional bulk LN devices, while preserving LN’s excellent material properties. Our compact LN electro-optic platform consists of low-loss nanoscale LN waveguides, micro-ring resonators and miniaturized Mach-Zehnder interferometers, fabricated by directly shaping LN thin films into sub-wavelength structures. The efficient confinement of both optical and microwave fields at the nanoscale dramatically improves the device performances featuring a half-wave electro-optic modulation efficiency of 1.8 V·cm while operating at data rates up to 40 Gbps. Our monolithic LN nanophotonic platform enables dense integration of high-performance active components, opening new avenues for future high-speed, low power and cost-effective communication networks.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement
OCIS codes: (130.3120) Integrated optics devices; (130.3730) Lithium niobate; (130.4110) Modulators; (230.2090) Electro-optical devices.

References and links


1. Introduction

Data centers, metropolitan and long-haul communication networks demand scalable and high-performance electro-optic modulators to convert electrical signals to modulated light waves at high speed [1–3]. For decades, lithium niobate (LiNbO\(_3\), LN) has been the material of choice owing to its excellent properties - namely large electro-optic response, high intrinsic bandwidth, wide transparency window, exceptional signal quality and good temperature stability [1, 2, 4]. Existing LN modulators however are not scalable due to the difficulty in nanostructuring LN [2]. As a result, they remain bulky (~10 cm long), discrete, expensive, and require high-power electrical drivers [1, 2].

To address the need of a scalable and high-performance electro-optic modulator for modern data communications, tremendous efforts have been made towards a variety of platforms that feature small footprints and high data bandwidths, including Si [5–8], InP [9–11], AlN [12], plasmonics [13], graphene [14–16] and polymers [17, 18]. However, both Si and InP modulators rely on switching mechanisms (carrier injection and quantum-confinement Stark effect, respectively) that are intrinsically nonlinear, absorptive, and sensitive to temperature fluctuation, posing increasing challenges for future ultra-high speed data transmission [5–11]. Limitations of other platforms include low switching efficiency (AlN) [12], high optical loss (plasmonics) [13], challenging scalability (graphene) [15] and compromised long term stability (polymer) [18].

LN remains the preferred material for high-performance electro-optic modulators due to its wide bandgap (high transparency) and large second order (\(\chi^2\)) electro-optic coefficient (30 pm/V) [4]. In contrast to Si and InP, the \(\chi^2\) process in LN changes its index of refraction linearly with an applied electrical field, at femtosecond timescale. The efficiency of this process is determined by the overlap of the optical and the electrical fields. Conventional ion-diffused LN waveguides suffer from the low refractive index contrast (\(\Delta n < 0.02\)) between core and cladding [Fig. 1(a)], resulting in large optical mode areas and bending radii [19]. As a result, the photonic structures are large and the radio-frequency (RF) electrodes have to be placed far away from the optical mode to prevent detrimental waveguide propagation loss, thus increasing the required voltage for electro-optic switching [1, 2].

Thin-film LN on insulator has recently emerged as a promising material platform to shrink the optical mode volume and boost the electro-optic efficiency [20]. While progress has been made towards chip-scale LN devices [20–33], realization of a monolithic LN nanophotonic circuit that can address the challenges of high-speed modern telecommunication remains challenging. This is partly due to the difficulty in LN dry etching which typically results in optical waveguides with high optical loss. For this reason, hybrid platforms have been pursued as an alternative, where an easy-to-etch material (e.g. Si and SiN) is used as a device layer and is bonded to non-etched thin LN films [22–26]. While these heterogeneous platforms have shown promising results, they may ultimately be limited by performance trade-off between electro-optic bandwidth, modulation efficiency, optical confinement and operating wavelength.

In this work we demonstrate an electro-optic platform that uses single-crystalline LN thin films as device layer. Photonic waveguides with submicron optical confinement, low propagation loss and small bending radii are realized by directly nanostructuring LN [Fig. 1(a)]. This allows us to place micrometer scale RF electrodes close to the waveguide which results in good overlap between optical and microwave fields and large electro-optic modulation efficiency [7 pm/V for micro-ring modulator and 1.8 V•cm for Mach-Zehnder interferometers (MZI)]. We perform high-speed data characterization of our devices, and find that they support data rates up to 40 Gbps.
2. Device design and fabrication

Figure 1(b) shows the schematic view of our nanophotonic LN modulators. The index contrast between the LN core and the SiO₂ cladding is $\Delta n = 0.67$, over an order of magnitude higher than ion-diffused LN waveguides. As shown in Fig. 1(c), devices are fabricated from a commercially available x-cut LN-on-insulator (LNOI) substrate (NANOLN), in which a LN thin film (700 nm thick) is bonded on top of silica (2 µm thick). An 800 nm thick amorphous silicon (a-Si) layer acting as a hard etching mask is first deposited on top of the LNOI using plasma enhanced chemical vapor deposition (PECVD). Waveguide and resonator patterns are defined by electron-beam lithography (EBL) writing on hydrogen silsesquioxane (HSQ) resist, and are aligned such that light propagation is perpendicular to the crystal z-axis. The patterns are then transferred into the a-Si layer using standard reactive ion etching (RIE), and subsequently into the LN thin film using Ar⁺ plasma etching in RIE. The Ar⁺ plasma physically bombards the LN thin film at a bias voltage of ~270 V, resulting in an etch rate of ~30 nm/min. The LN etch depth is ~400 nm, leaving a 300 nm slab across the
chip. The residual mask materials are removed in 30% KOH at 80 °C. Bottom metal electrodes (15 nm Ti/300 nm Au, deposited using an electron-beam evaporator) are formed along the MZI waveguides and racetrack straight arms using a standard PMMA/MMA double layer lift-off process. A 1.5 μm thick silica cladding layer is then deposited on top of the waveguides and electrodes using PECVD. A new PMMA resist is spun coated and written with EBL to open windows at corresponding via locations. Vias are formed by 5 min wet etching in buffered oxide etch (BOE). A second lift-off and evaporation process (under the same conditions) is then performed to produce the top electrodes including probe contact pads and metal strips connecting relevant vias. Due to the isotropic nature of the BOE process, the vias have slanted sidewalls that ensure the electrical conductance between top and bottom electrodes. The final devices are diced and polished to ensure good coupling from and to optical fibers.

Fig. 2. Fabricated optical devices and electrical contacts. (a-b) False-color scanning electron microscope (SEM) images of the fabricated racetrack and ring resonator based modulators (a) and Mach-Zehnder interferometer based modulators (b). (c) Close-up SEM image of the coupling region of the racetrack resonator. (d) Close-up SEM image of the metal electrodes and the optical waveguide. (e) Cross-section view of the simulated optical TE mode profile ($E_z$ component) and RF electrical field (shown by arrows). The x-cut LN used here is most sensitive to the horizontal component of the electric field ($E_x$). $h$: LN waveguide height; $w$: LN waveguide width; $s$: LN slab thickness; $g$: metal electrode gap.

Figures 2(a)-2(e) show a range of fabricated nanophotonic LN devices including nanowaveguides, ring resonators, racetrack resonators and MZIs. The electro-optic devices
presented in this work feature propagation loss of 3 dB/cm, which is limited by sidewall roughness transferred from the etch mask [33], and can be further reduced to ~0.03 dB/cm [34]. The ultimate material absorption limit of LN is < 0.002 dB/cm [35]. The resulting MZI and racetrack structures here have low on-chip insertion loss of 2 dB and 1.5 dB respectively (total waveguide lengths = 7 mm and 5 mm respectively), with additional 5 dB/facet coupling loss.

The highly confined optical mode allows us to place gold micro-RF electrodes close to the LN waveguide [Figs. 2(d) and 2(e)], enabling low operating voltages. Our devices make use of an x-cut LN configuration, where transverse-electric (TE) optical modes and in-plane electric fields ($E_z$) interact through the highest electro-optic tensor component ($r_{33}$) of LN. We design the waveguide geometry and the micro-RF electrode positions to achieve the optimal overlap between the optical and electric fields, while minimizing the bending loss and the metal-induced absorption loss. Figure 2(e) shows the numerically simulated overlap between the corresponding optical and electric fields. The mode confinement factor, defined as the percentage of electromagnetic energy confined within the LN region, is 94%. The rib optical waveguides have a top width $w = 900$ nm, a rib height $h = 400$ nm, and a slab thickness $s = 300$ nm [Fig. 2(e)]. To maximize the in-plane electric field ($E_z$), we sandwich the optical waveguide between the signal and ground electrodes with a gap $g = 3.5$ μm. A silica cladding is used to further enhance this overlap by increasing the dielectric constant of the surrounding media to match the high dielectric constant of LN ($\varepsilon = 28$) [4].

3. Measurement results

3.1 DC electro-optic response

We show efficient and linear electro-optic tuning in a racetrack modulator and a MZI modulator. Light from a tunable telecom laser (Santec TSL-510, 1480 – 1580 nm) is coupled into and collected from the waveguide facets using tapered lensed fibers. A 3-paddle fiber polarization controller is used to ensure TE input polarization. Figure 3(a) shows a typical transmission spectrum of a racetrack resonator with a loaded quality ($Q$) factor of 50,000. When a voltage is applied, the change of refractive index modifies the effective optical path length of the resonator, resulting in a resonance frequency shift. The electrical fields on the two racetrack arms are aligned to the same direction so that the modulation on the two arms adds up [Fig. 1(b)]. The measured electro-optic efficiency is 7.0 pm/V with good linearity and no observable changes in resonance extinction ratio and linewidth [Figs. 3(a) and 3(b)]. The microring modulators also show similar electro-optic efficiencies ~7 pm/V. The MZI modulator is a balanced interferometer with two 50:50 Y-splitters and two optical paths. The applied voltage induces a phase delay on one arm and a phase advance on the other, which in turn change the output intensity at the Y-combiner by interference. The minimum voltage that is needed to completely switch the output between on and off is defined as the half-wave voltage ($V_{1/2}$). We measure a $V_{1/2}$ of 9 V from a 2 mm long MZI modulator, with 10 dB extinction ratio [Fig. 3(c)]. The 10 dB extinction here is not fundamental and is limited by the excitation of spurious optical mode through our multimode optical waveguide. Much higher extinction ratio (> 30 dB) can be realized using a narrower, single mode optical waveguide. The measured modulator response translates to a voltage-length product of 1.8 V·cm, nearly an order of magnitude lower than bulk LN devices [1] and significantly better than previously reported LN thin-film devices [21–29].

3.2 High-speed measurements

We measure high electro-optic bandwidths ($S_{21}$ parameter) of our miniaturized devices using the setup shown in Fig. 4(a). Here an erbium-doped fiber amplifier (EDFA) and a band-pass filter (BPF) are used after the chip to increase the optical intensity in the detector. A 50 GHz microwave probe (GGB Model 50A) is used to connect the on-chip metal contacts and the 50
A 50 GHz vector network analyzer (VNA, Agilent E8364B) and a 40 GHz photodiode (New Focus 1014) are used. RF cable losses are subtracted from the measured frequency responses.

Fig. 3. DC electrical and optical characterization. (a) Measured transmission spectra of a high Q (~50,000) racetrack resonator exhibits large frequency shift with applied DC voltage. (B) Linear resonant wavelength shift as a function of DC voltage. The measured tuning efficiency is 7.0 pm/V. Error bars show the uncertainty in determining the resonance positions. (C) Optical transmission of a 2 mm long MZI modulator versus DC voltage applied, indicating a half-wave voltage ($V_{\pi}$) of 9 V and a voltage-length product of 1.8 V·cm.

For a racetrack resonator modulator featuring a Q factor of 8,000, we measure a 3 dB electro-optic bandwidth of 30 GHz [Fig. 4(b)]. This value is limited by the cavity-photon lifetime of the resonator (6 ps). We confirm that the bandwidth is limited by photon lifetime by testing additional resonators with Qs of 5,700 and 18,000. The resulting 3 dB bandwidths are 40 GHz and 11 GHz respectively (data not shown). The Q factors are engineered from the intrinsic value by controlling the distance between the RF electrodes and the optical waveguide. The intrinsic RC bandwidth limit of the racetrack modulator is estimated to be over 100 GHz. For the 2 mm long MZI device with direct capacitive modulation, the measured electro-optic 3 dB bandwidth is 15 GHz [Fig. 4(c)]. This is limited by the RC constant due to a larger capacitance (0.2 pF) induced by the longer RF electrode used. The RC limits of both schemes are confirmed from electrical return loss ($S_{11}$ parameter) measurements shown in the insets of Figs. 4(b) and 4(c). The $S_{11}$ responses generally show large electrical power reflection except the frequency where the capacitor is impedance matched with the 50 Ω RF circuit. The MZI modulator shows a clear $S_{11}$ dip at 15 GHz, while the racetrack modulator does not show an $S_{11}$ dip within our equipment frequency range, which agree well with our electro-optic measurement results. Note that since our on-chip electrical resistance is rather small (< 10 Ω), the measured bandwidth is currently limited by the 50 Ω impedance of the network analyzer drive, and can be pushed to a much higher value using a RF transmission line design (discussed in Section 4).

Our platform supports data transmission rates as high as 40 Gbps. Figure 4(d) displays non-return-to-zero (NRZ) open eye diagrams for both racetrack and MZI modulators at various data rates, obtained with 2^7-1 pseudo-random binary sequence (PRBS, CentellaxTG1P4A). The signals are amplified (Centellax OA4MVM3) to 5.66 V peak to peak ($V_{pp}$) before being sent to the modulators. Eye diagrams are obtained by sending the modulated and amplified light into the 30 GHz optical module of a sampling scope (Tektronix CSA8000). Because of the high signal quality, our devices can generally operate at data rates 50% higher than their 3 dB bandwidth, which translates to 40 Gbps and 22 Gbps for the racetrack and MZI devices, respectively. At maximum frequencies, the measured extinction ratios of the racetrack and MZI modulators are 3 dB and 8 dB with power consumptions (CV^2/4) of 240 fJ/bit and 1.6 pJ/bit, respectively, which is already better than silicon photonic modulators that reach the same bandwidth (30 pJ/bit) [36]. The insulating nature of LN also ensures that no DC holding energy is required in the system.
Fig. 4. Bandwidth and high-speed data operation. (a) Detailed high-speed measurement setup. VNA and PD (40 GHz) are used for electro-optic bandwidth test. PRBS, RF amplifier and optical sampling scope (with a 30 GHz optical module) are used to obtain the eye diagrams. (b–c) Electro-optic bandwidths ($S_{21}$ parameter) of a race-track resonator with $Q$ of 8,000 (b) and a 2 mm long MZI (c). The corresponding 3 dB bandwidths are 30 GHz and 15 GHz respectively. The measurements have been calibrated to exclude the RF cable losses. Insets show the electrical return loss measurement ($S_{11}$ parameter) for both schemes. The measured RC bandwidth limits (> 30 GHz for racetrack and ~15 GHz for MZI) agree well with the electro-optic bandwidth measurement results. (d) Eye diagrams of the racetrack (blue) and MZI (purple) modulator with data rates up to 40 Gbps and 22 Gbps. All eye diagrams are measured with 2^-1 PRBS in a non-return-to-zero scheme with a 5.66 $V_{pp}$ electrical drive. (e) Eye diagrams of the MZI modulator at 12.5 Gbps and 22 Gbps with the device heated up by 20°C. The extinction ratios in (d) and (e) are 3 dB and 8 dB for race-track resonator and MZI respectively. BPF, band-pass filter; DUT, device under test; EDFA, erbium doped fiber amplifier; FPC, fiber polarization controller; PD, photodiode; PRBS, pseudo-random binary sequence source; VNA, vector network analyzer.
We confirm that our MZI modulator maintains the stable thermal properties of their bulk counterparts, due to the low thermo-optic coefficient of LN ($3.9 \times 10^{-5} \text{K}^{-1}$) [4]. We use a Peltier controller to regulate the temperature of our chip within a $\Delta T = 20 \degree \text{C}$ range, limited by the setup, and record the eye diagrams. We find that the MZI modulator is able to maintain an open eye diagram at the maximum data rate of 22 Gbps without any feedback to compensate for temperature drifts (open loop configuration) [Fig. 4(e)]. During the test, laser wavelength, input polarization and bias voltage are kept constant, while the lensed fiber positions are fine tuned to compensate for the relative motion between the chip and the fibers.

4. Conclusions and Outlook

Our micrometer scale LN modulators feature high bandwidth, excellent linearity, low operating voltage and good temperature stability, potentially enabling a paradigm shift in electro-optic modulator scalability, performance and design. For over a decade, existing LN modulator performance has been capped due to the high dielectric constant of LN ($\varepsilon_{\text{RF}} = 28$) [4], which dictates that RF fields in LN propagate much slower than optical fields ($\varepsilon_{\text{opt}} = 4.6$). In order to realize velocity matching between the two waves, a silica buffer layer has to be inserted between the RF electrodes and the device layer, compromising the electro-optic efficiency [1]. In our thin-film LN approach, instead, phase matching can be achieved without sacrificing electro-optic overlap, since the electrical field primarily resides in the low dielectric silica substrate ($\varepsilon_{\text{RF}} \sim 4$) and readily propagates at nearly the same group velocity as light [26, 37]. Combining the high electro-optic efficiency and low propagation loss in our thin-film LN platform with a phase-matched RF transmission line could further increase the MZI arm length to $> 2 \text{ cm}$ while maintaining a high electro-optic bandwidth. This will lead to devices with ultra-high bandwidth ($> 60 \text{ GHz}$) and low modulation voltage ($< 1 \text{ V}$) simultaneously, that could be directly driven with CMOS circuitry. In addition, the dense integration, high linearity and transparency of our LN platform could enable coherent modulation technique to be efficiently implemented. The active micro-resonators and low loss waveguides could enable a chip-scale photonic circuit that is densely integrated with novel switches, filters and nonlinear wavelength sources and operates in a wide wavelength range (from visible to mid-IR). Furthermore, the ultra-compact footprint (as small as $30 \mu\text{m} \times 30 \mu\text{m}$) of micro-ring modulators is attractive for data center applications where real estate is at a premium. Our high-performance monolithic LN nanophotonic platform could become a practical cost-effective solution to meet the growing demands of next-generation data centers and long-haul optical telecommunications.

Funding

National Science Foundation (NSF) (ECCS-1609549); Office of Naval Research MURI (N00014-15-1-2761); Army Research Laboratory Center for Distributed Quantum Information (W911NF1520067); CIAN ERC (EEC-0812072).

Acknowledgments

The authors thank J. Kahn for valuable discussions, H. Majedi and D. Ham for the use of their equipment. Device fabrication was performed at the center for nanoscale systems (CNS) at Harvard University.