Invited Papers

Heterogeneous silicon light sources for datacom applications

Di Liang\textsuperscript{a,b,}, Geza Kurczveil\textsuperscript{a}, Xue Huang\textsuperscript{a}, Chong Zhang\textsuperscript{a}, Sudharsanan Srinivasan\textsuperscript{b}, Zhihong Huang\textsuperscript{a}, M. Ashkan Seyedi\textsuperscript{a}, Kate Norris\textsuperscript{a}, Marco Fiorentino\textsuperscript{a}, John E. Bowers\textsuperscript{b}, Raymond G. Beausoleil\textsuperscript{a}

\textsuperscript{a} System Architecture Lab, Hewlett Packard Labs, Palo Alto, CA 94304, United States
\textsuperscript{b} Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, United States

\textbf{A B S T R A C T}

Photonic interconnects are being quickly developed and adopted to be an indispensable large-bandwidth data transmission venue in high-performance computing (HPC) and datacenter, i.e., datacom, applications. The distinct emphasis on power, operational robustness, volume manufacturability and system cost posts a number of challenges in photonic material platform choice, device design, integration scheme, and packaging, etc. As the major roadblock to realize a fully integrated silicon (Si)-based photonic interconnect system, light sources have attracted significant research and development effort and driven great innovation in the past 12 years. We review our recent progresses towards building heterogeneous Si multi-wavelength lasers with exceptional tolerance to harsh application environment and novel controllability. Advanced quantum dot (QD) material is now successfully implemented in the heterogeneous platform to serve as the robust optical gain medium in a comb laser to deliver multiple low relative intensity noise (RIN) lasing wavelengths and seamlessly integrate with Si photonic circuits for dense wavelength-division multiplexing (DWDM) photonic interconnect architecture. A new heterogeneous metal-oxide-semiconductor (MOS) capacitor is also developed to introduce an independent carrier effect to this platform and enable essentially power-free device tunability and superior high-speed data modulation capability for both modulators and lasers. Finally the on-going effort to incorporate superior QD material, MOS capacitor and several previously developed functionalities into heterogeneous comb lasers and microring lasers for different datacom application scenarios is discussed.

1. Introduction

In the present Big Data era, data is growing in an unprecedented pace. The IT infrastructure to sustainably support this amount of data generation, computation and transportation has reached a physical limit where fundamentally revolutionary technologies are necessary. Optical interconnect is such a disruptive technology to break network bandwidth constraints and energy costs arising from architectural designs originally based on copper interconnects [1]. The concept of optical communication was developed several decades ago, and large data transmission bandwidth is always ideal. But compared with the conventional long-haul fiber optics, the system catering to the datacom applications has a distinctly different emphasis on power efficiency, reliable operation at a temperature-fluctuating environment, e.g., 30–85 °C, volume manufacturability and overall system cost in order to be technically and economically competitive [2].

The emerging silicon (Si) photonics in the past 15 years have significantly challenged the traditional III-V compound semiconductor counterparts for the undisputable advantages in matured complementary metal-oxide semiconductor (CMOS) technology [3]. This platform has numerous intrinsically better material properties, such as low material cost, low optical loss, excellent thermal conductivity, and lower two-photon absorption coefficient, etc. A large number of Si-based photonic components, such as high-speed modulators [4], Ge and SiGe photodetectors [5–7], and integrated circuits with comparable or even superior performance to their III-V compound semiconductor counterparts have been developed [8]. A laser source still remains to be the most challenging active component owing to intrinsic Si indirect bandgap and four key factors mentioned above [9,10]. Generally speaking, laser sources currently dominate the overall power consumption and component cost in the total optical link and largely affect another critical aspects like packaging, which eventually determines the total system cost and overall performance.

Several Si-based hybrid, heterogeneous or monolithic laser source integration schemes have been developed recently to provide a full, cost-effective Si photonic toolbox [11–14]. In this paper, we review our approaches and recent progress on developing high-performance Si heterogeneous lasers [15–17], aiming to address all those critical requirements for datacenter and high-performance computer (HPC) businesses.

2. Heterogeneous quantum dot comb lasers

Thermal management, regardless from internal Joule heating or
external environmental heating, is a long-term inevitable challenge against the diode laser operation. It drastically reduces the device efficiency, shifts the wavelength and shortens the lifetime. The cost constraint, physical size limit, and photonic integration with the driver chip further complicates the laser source implementation. Lasers with QD active regions have shown superior performance over traditional quantum well (QW) designs in their ability to maintain large optical gain at elevated temperatures. Due to the three dimensional confinement of carriers, QD lasers promise higher temperature stability and lower threshold current densities when compared to quantum well lasers. These promises have been realized by the demonstration of QD lasers that operate at a stage temperature of 220 °C [18], have negligible threshold variation between −40 and 100 °C [19], and have threshold current densities as low as 17 A/cm² [20–22]. In addition, due to the size distribution of the dots, QD lasers have a wide gain bandwidth [23], which allows for a larger channel count in a wavelength-division multiplexing (WDM) link. It has also been found that QD lasers have a low relative intensity noise (RIN) [24]. Single-section QD-based Fabry-Perot lasers with a RIN low enough for error-free operation at 10 Gb/s have been demonstrated [25,26].

The above advantages make QD lasers a particularly attractive candidate for on-chip, multi-wavelength sources in datacom applications. Before these devices can be deployed at a large scale, a commercially viable solution must be found. Using a pure III-V approach limits a foundry to using relatively small wafers ~150 mm in diameter at relatively old processing nodes of ~250 nm. The integration on Si substrates, on the other hand, would potentially allow the use of widely available CMOS foundries which would result in high-volume, low-cost integrated III-V lasers with high-quality and reliable Si photonics. Furthermore, the integration of the laser source directly on Si would reduce coupling losses, subsequently reducing the link budget and packaging costs inherent in multi-chip solutions. The most promising integration scheme will take advantage of high performance active devices while also taking advantage of the excellent optical properties of low optical losses in Si. Typical integration schemes include growth [13], flip-chip bonding [27], and wafer bonding [28,29]. Early wafer-bonded approaches use a metal layer between the Si and the GaAs wafers to facilitate bonding. This metal layer may increase optical losses if it overlaps with the optical mode. These approaches also used an electrical contact on the Si substrate, which required a highly doped Si layer. This further increased optical losses [30], and prevented the use of a buried oxide layer. Therefore, light could not be confined in the Si for waveguiding.

Wafer bonded QD lasers without a metal layer have recently been demonstrated [17,31,32]. In these solutions, similar to other heterogeneous Si approaches [14], the bonding interface between the Si and the III-V consists of a dielectric with low optical losses rather than a high-loss metal layer. Furthermore, both electrical contacts are formed on the III-V. This has two important consequences. First, it allows the use of undoped Si as the waveguide layer, further reducing optical losses. Second, it allows the use of a buried oxide layer which prevents the optical mode from leaking into the Si substrate. This platform allows the seamless and low-loss integration of heterogeneous Si quantum dot lasers (HSQDL) with high-quality Si photonic components, such as [de]multiplexers, grating couplers, and active components such as modulators and photodetectors.

2.1. Device design and fabrication

A cross-sectional diagram of our HSQDL is shown in Fig. 1(a). It consists of an SOI wafer with etched air trenches to form a Si rib waveguide. A p-i-n GaAs-based diode laser structure with an InAs QD active region and 8 layers of QDs totaling 320 nm in thickness was then transferred to the SOI substrate using an O₂ plasma-assisted direct bonding process [33]. The barriers between the QD layers are p-doped as this was shown to improve T₀, the differential efficiency, and the differential gain [34]. A cross-sectional scanning electron microscope (SEM) image of a fabricated device where the tapers have been polished off is shown in Fig. 1(b). The QD active region is highlighted in red.

Using a fixed mesa width and etch depth, lasers with varying QD-confinement factors can be realized as the overlap of the optical mode with the QD region can be tuned by adjusting the width ‘w’ of the Si waveguide. This is illustrated in Fig. 1(c) and (d), where Si waveguide widths of 1.8 and 0.7 μm yield QD confinement factors of 14% and 55%, respectively, with no change in the III-V mesa width or etch depth. This is an important aspect of the heterogeneous Si approach – the optical mode can be transferred between the Si waveguide and the III-V waveguide using mode converters (tapers) in the Si waveguide, thus relaxing the tolerance on the photolithographic alignment and resolution during the III-V mesa lithography and etch [35]. The mode converter in the present device consists of two taper levels as illustrated in Fig. 1(e). By using a narrower Si waveguide, the light is mostly confined in the III-V section for maximal optical gain. As the light is exiting the gain region, it passes through the first taper level where the Si waveguide width is changed linearly from 0.7 μm to 2 μm over a length of 120 μm. After this taper level, the overlap of the optical mode with the Si waveguide width is ~90%. In the second taper level, the III-V width is changed linearly from 6 μm to ~0.2 μm over a length of 72 μm. The length of the tapers were chosen to give ~90% mode conversion efficiency in each taper level or over 80% for a single transition between a passive Si waveguide and an active III-V waveguide.

The SOI and GaAs QD material were purchased from commercial vendors. Devices were fabricated on a 100 mm Si-on-Insulator (SOI) wafer with a top Si thickness of 400 nm and a buried oxide thickness of 1000 nm. Rib waveguides were defined using 248-nm DUV lithography and etched to a depth of 320 nm. All following lithography steps were performed on a 365 nm i-line stepper and the process follows the one outlined in Ref. [17]. The fabricated devices were diced into multiple laser bars and the facets were polished to form mirrors with reflectivity...
Fig. 2. (a) Device LI characteristic as a function of stage temperature. Lasers were measured on a temperature controlled stage using a DC current source and an integrating sphere. The integrating sphere measured the light out of one facet. (b) Room-temperature LI data for the device with the lowest threshold current. The mesa width, gain section length, and cavity length were 4 μm, 800 μm, and 1.7 mm respectively, and the device had a threshold current of 15 mA.

~ 34%.

2.2. Device characterization

Lasers with a 6-μm-wide mesa and a 0.7-μm-wide Si waveguide were tested. The Fabry-Perot (FP) laser bar was 3 mm long, and had a 2-μm-long gain region while the rest of the cavity was taken up by tapers and a passive Si waveguide. Using an integrating sphere, continuous wave (CW) light-current (LI) data as a function of stage temperature were measured as shown in Fig. 2(a). The room-temperature threshold current is 32.5 mA yielding a threshold current density of 271 mA/cm². CW lasing is seen up to 100 °C, at drive currents up to 200 mA, with no thermal roll-over being observed. Higher currents were not tried to avoid excessive joule heating as the devices exhibited a high series resistance due to a fabrication error in the metallization step.

While the threshold current is relatively stable up to 60 °C, we anticipate improved performance with a lower device series resistance by using a better metallization process as this minimizes joule heating. Further improvements are expected by incorporating thermal shunts [36,37] which connect the high temperature region, e.g., the p-metal, with the Si substrate, thus circumventing the thermally insulating buried oxide layer (BOX) of the SOI wafer. The LI-curve for the device with the lowest room-temperature threshold current is shown in Fig. 2(b). The device had a mesa width, gain section length, and cavity length of 4 μm, 800 μm, and 1.7 mm respectively. The Si waveguide under the III-V mesa had a width of 0.7 μm which resulted in a QD confinement factor of ~55%. The device had a threshold current of 15 mA and thermal rollover didn’t appear until drive current is 10 × higher than the threshold.

Typical optical spectra below and above the threshold current from a device with an 8-μm-wide mesa with QD active region confinement factor Γ_{QD} = 55% at a stage temperature of 20 °C are shown in Fig. 3(a). Since the FP laser cavity includes no narrow wavelength selective filter, multiple longitudinal cavity modes lase simultaneously as shown in Fig. 3(b). This is the result of the large gain bandwidth of the QD material along with the small channel spacing, corresponding to the cavity round-trip time (0.08 μm, 14.3 GHz). The mode spacing is consistent with the laser cavity being formed by the polished Si facets. This channel spacing is too small for a high speed (> 10 Gb/s) WDM link. By reducing the cavity length to ~860 μm a mode spacing of 50 GHz can be achieved. This is large enough to avoid crosstalk between neighboring channels, while it is also small enough to allow a large number of channels [38]. Near field images in Fig. 3(c) and (d) show the polished Si facet with the laser on and off, respectively, proving that the optical mode is coupled to the Si waveguide.

In transmission links that use amplitude modulation, the RIN of each individual comb line (not the whole spectrum) is one of the most important characteristics of the comb laser. In QW-based FP lasers, the RIN of each comb line is typically limited by the competition of multiple longitudinal modes for optical gain, typically defined as mode partition noise. This mode partition noise is large enough to prevent using individual comb lines for error-free amplitude modulated transmission. Thus, WDM-links would typically consist of multiplexed single wavelength lasers, e.g., distributed feedback (DFB) or microring lasers, where each laser needs separate wavelength and current control. While this is adequate for light sources with 16 or less number of channel, it is not the best approach for a 64- or 128-channel light source, where a single comb laser is more appropriate. Using QW gain material, such a comb laser would require a relatively elaborate cavity configuration [39] to achieve the required RIN. However, it was shown that mode partition noise in QD lasers is significantly smaller than in QW lasers [24]. This allows the use of simple QD-based FP lasers as comb laser sources, where each comb line has a low-enough RIN for error-free amplitude-modulated transmission.

Preliminary experiments on our lasers demonstrate the viability of this approach. Using the lasers described above, individual comb lines were filtered out using an external optical filter and modulated using an external modulator. With a 10 Gb/s 2^{23-1} PRBS NRZ signal, we observe open eye diagrams as shown in Fig. 4 and we measured a bit error ratio of 1 × 10^{-12} or better. These results confirm that using QD gain material, a relatively simple cavity configuration can be used in order to reach 64- and 128-channel comb arrays that are suitable for error-free amplitude modulated transmission in applications of datacom and beyond.

3. Heterogeneous MOS device platform

In addition to adopting the better QD optical gain material to largely minimize the thermal management demand and achieve a milestone towards demonstrating the large channel count DWDM on-chip comb lasers, we developed a heterogeneous metal-oxide semiconductor (MOS) platform for multiple photonic device applications and new functionalities for the first time [15,16,40].

This platform is based on a heterogeneous III-V/dielectric/Si MOS capacitor structure realized naturally by the molecule wafer bonding technique [33,41] which is the key fabrication step for all heterogeneous Si photonic devices. As shown in a transmission-electron microscope (TEM) image in Fig. 5(a) [42], a bonding oxide interface layer forms as a result of either O2 plasma-assisted bonding process itself [33] or intentional and controllable dielectric deposition [43,45]. The dielectric layer here, is usually referred as the gate oxide, conventionally named in MOS field-effect transistors. Fig. 5(b) shows simulated majority carrier concentration, i.e., free electron and hole in n-InP and p-Si, respectively, as a function of bias voltage. Due to the different...
doping levels ($1 \times 10^{18}$ cm$^{-3}$ in n-InP and $5 \times 10^{16}$ cm$^{-3}$ in p-Si), free carriers around the gate oxide (15 nm Al$_2$O$_3$ in this work) are depleted at zero bias indicated by the band diagram in Fig. 5(b) inset. As shown in Fig. 5(b), once forward bias is applied to p-Si, the heterogeneous MOS capacitor operates in accumulation mode. The hole concentration at the oxide/Si interface increases by over 9 orders of magnitude at 1 V bias. Both hole and electron concentrations at the oxide/semiconductor interface reach $10^6$ cm$^{-3}$ level at 8 V bias. Typical capacitance-voltage (CV) characteristic of such a heterogeneous MOS capacitor structure is measured at both low and high frequency in Fig. 5(c). It agrees well with the calculated value of 0.53 $\mu$F/cm$^2$ based on the classic parallel plate capacitor model.

Similar to previously demonstrated pure Si MOS modulators [4,46,47], this heterogeneous MOS capacitor is extremely useful to introduce the plasma dispersion effect [30] into the heterogeneous photonic device platform. The refractive index and free-carrier absorption loss in InP and Si materials change as a result of free carriers accumulating or depleting near the gate oxide under an external bias. Thus phase tuning or data modulation can be achieved. In order to realize a heterogeneous waveguide where the optical mode(s) overlap with the MOS capacitor, a “staircase” shape of the waveguide geometry is designed as shown in Fig. 6 inset. Single mode operation condition is determined by the total thickness of Si, dielectric and III-V layer and waveguide lateral dimension. It is similar to Si MOS-type waveguide based on the Poly-Si/gate oxide/crystalline Si structure [47] except that the Poly-Si layer is replaced by a layer of crystalline III-V material. A large variety of III-V compound semiconductor, primarily InP- or GaAs-based materials in our applications, can be flexibly chosen to form such a heterogeneous MOS capacitor. Several inherent material property advantages in III-V materials results in much better overall light manipulation. The change of material index $\Delta n$ and free carrier absorption (FCA) loss $\Delta \alpha$ due to plasma dispersion effect are described by the classic Drude model [30] in Eqs. (1) and (2) below:
\[ \Delta n = \frac{-q^2 \lambda^2}{8 \pi^2 \epsilon_0 n_0 c^2} \left( \frac{\Delta N_n}{m^*_e n_0} + \frac{\Delta N_p}{m^*_h n_0} \right) \]  
\[ \Delta \alpha = \frac{-q^2 \lambda^2}{4 \pi^2 \epsilon_0 n_0 c^2} \left( \frac{\Delta N_n}{m^*_e n_0^2 n_0} + \frac{\Delta N_p}{m^*_h n_0^2 n_0} \right) \]

where \( q, c, \epsilon_0, n \) and \( \lambda_0 \) are the elementary charge, speed of light in vacuum, permittivity of free space, unperturbed material refractive index and light wavelength in vacuum, respectively. \( \Delta N_n \) and \( \Delta N_p \) are the respective concentration change of electrons and holes. Two material parameters, conductivity effective mass \( m^*_e, m^*_h \) and mobility \( \mu_e, \mu_h \) for electron and hole, respectively, differentiate the plasma dispersion effect in different materials. Due to much smaller conductivity electron effective mass and larger electron mobility in InP (\( m^*_e = 0.07 m_0, \mu_e = 5400 \text{ cm}^2/\text{V-s} \)) and GaAs (\( m^*_e = 0.063 m_0, \mu_e = 8500 \text{ cm}^2/\text{V-s} \)) than these of crystalline Si (\( m^*_e = 0.28 m_0, \mu_e = 1400 \text{ cm}^2/\text{V-s} \)), plasma dispersion effect-induced index change in n-type InP and GaAs is much more effective and FCA is lower. However, similarity in those parameters for holes among them doesn’t give advantages to III-V materials over Si, and the hole-induced intervalence band absorption is much larger in InP and GaAs on the contrary. Coincidently, free holes can lead to larger index change than electrons in Si [30]. Therefore a heterogeneous capacitor with n-type III-V and p-type Si is a natural choice to maximize the carrier-induced index change.

Simulated plasma dispersion effect-induced modal index change and FCA for the TE00 mode at 1550 nm as a function of bias voltage to the MOS capacitor in the heterogeneous MOS-type waveguide (inset) is shown in Fig. 6. Three MOS-type waveguides share the same 250 nm thick p-Si with 150 nm single-side rib etch depth, 800 nm waveguide width and a doping level of \( 5 \times 10^{16} \text{ cm}^{-3} \), and 15 nm gate oxide of Al2O3. The top n-type material is either crystalline Si, InP or GaAs with a fixed thickness of 150 nm and doping level of \( 1 \times 10^{18} \text{ cm}^{-3} \). At a bias voltage of 6 V, the structure with n-InP or n-GaAs gains an index change \( \sim 2.6 \times 3.5 \times \) larger than that with n-type crystalline Si. The FCA, on the other hand, is respectively 56% and 44% smaller for GaAs and InP case compared with n-type crystalline Si. Even larger margin in material loss reduction for heterogeneous case is expected when compared with the commonly used lossy poly-Si (\( \sim 9 \text{ dB/cm} \) [48]) instead of crystalline Si in Fig. 6.

Moreover, there are more carrier-charge effects, namely band-filling and bandgap shrinkage, in InP- and GaAs-based materials, whose overall effect is to further enhance the negative index change with negligible absorption loss in the 1.31 and 1.55 \( \mu \text{m} \) wavelength windows [49]. A detailed calculation can be referred to [49,50] and is not conducted here due to space limitations.

### 3.1. Tunable microring resonators and modulators

The intrinsic larger material index change and lower material losses in this heterogeneous MOS structure over recently demonstrated pure Si MOS counterparts promise its usefulness in phase modulators, switches and general wavelength tuning. We integrated this heterogeneous MOS capacitor into microring resonators for the first demonstration [15]. Mach–Zehnder interferometer (MZI)-based modulators with similar heterogeneous MOS structure were reported by several other groups very recently as well [51,52,50]. Fig. 7(a) is the cross-sectional schematic of such a microring resonator with a simulated fundamental TE mode assuming a diameter of 40 \( \mu \text{m} \). The "staircase"-shape heterogeneous MOS-type waveguide is easily formed by etching an air trench to define the inner edge/diameter of the resonator prior to wafer bonding and then a self-aligned process to define the outer edge/diameter of the device after gate oxide formation and III-V layer transfer. This air trench serves the dual purpose of electrical and optical isolation to limit the MOS capacitor area and reduce the mode number, e.g., only fundamental transverse mode. The width of the air trench should be carefully chosen, normally 1 \( \mu \text{m} \) in our design, to avoid any mechanic stability issue in the free-standing III-V thin film above it. It is a much simpler process compared with monolithic process to make pure Si modulators [46,47], because the dielectric filling to the isolation trench and chemical-mechanic polished (CMP) processes that inevitably introduce non-uniformity in the p-Si layer are avoided. Fig. 7(b) shows the device schematic of such a heterogeneous MOS microring resonator. The electrodes are placed on n-type InP and p-type Si to apply electrical field across the gate oxide. Fig. 7(c) is simulated 2D hole distribution in Si layers under 0, 4 and −4 V biases. Clearly the air trench confines the carrier variation in the fundamental transverse mode region to effectively reduce the capacitor area, which potentially benefits the device high-speed performance. At −4 V reverse bias, the heterogeneous MOS capacitor operates in carrier inversion mode, and the plasma dispersion effect is a lot weaker than accumulation mode in this device design.

Critical device fabrication steps are illustrated schematically in Fig. 8(a)-(f). The 100 nm diameter SOI substrate with 250 nm-thick, p-type lightly-doped (\( 5 \times 10^{16} \text{ cm}^{-3} \)) Si device layer and 1 \( \mu \text{m} \)-thick BOX is used in this work. Upon patterning the grating couplers (Fig. 8(a)), the heavily doped p + region outside the resonator is defined in order to form the ohmic contact in Si. We target a Boron doping level around 1 \( \times 10^{19} \text{ cm}^{-3} \) after ion implantation process and doping activation annealing at 1050°C. Then another photolithography and dry etch process is used to form the ring-shaped air trench for electrical and optical isolation along with vertical outgassing channels [33] that ensure proper bonding (Fig. 8(b)). After a rigorous surface cleaning of the Si and III-V wafers, the native oxide on Si and InGaAs layer cap on III-V are removed in buffered HF and H3PO4:H2O2:H2O solutions, respectively, followed by the Al2O3 deposition in the atomic layer deposition (ALD) tool. After depositing 15 nm Al2O3 in total, the III-V sample is bonded onto the SOI substrate with a 45-min anneal at 300°C (Fig. 8(c)). The InP substrate is then selectively removed in HCl:H2O solution to leave the ~2 \( \mu \text{m} \)-thick diode laser epitaxial layers on the SOI substrate. Details of the III-V epitaxial structure can be found in [35]. Then solutions of H3PO4:H2O2:H2O and HCl:H2O are used sequentially to selectively remove InGaAs, p-InP and multiple quantum well (MQW) active region layers to leave only a total of 150 nm-thick n-InP layer with a doping level of 1 \( \times 10^{18} \text{ cm}^{-3} \) plus 2 periods of InP/InGaAsP superlattice close to the bonding interface. Upon depositing 100 nm SiO2 to be the hard mask material, microring resonators in diameters of 20 and 40 \( \mu \text{m} \) are patterned by electron-beam lithography (EBL) and transferred through SiO2, III-V and 210 nm into the Si layer by self-aligned dryetches (Fig. 8(d)). After selectively removing all the III-V material to expose the Si surface outside the microring mesa, the
same liftoff and metallization processes (Ni/Ge/Au/Ni/Au) are carried out to form the metal contacts on n-InP and p++-Si (Fig. 8(e)), followed by an encapsulation process of 700 nm SiO₂ deposition to the chip surface. Finally, the through-dielectric vias and thick metal probe pads are formed to finish the fabrication (Fig. 8(f)). Top-View microscopic image and SEM cross-sectional image of a fabricated device are shown in Fig. 8(g) and (h), respectively. The width of the waveguide, i.e., MOS capacitor region, is designed to be 0.8, 1 and 1.2 μm.

During the measurement, Si grating couplers are used for optical signal input and output and a broad-band optical semiconductor amplifier is used as the light source. Fig. 9 shows the resonance wavelength shift of a device with 40 μm in diameter, 0.8 μm waveguide width under (a) positive voltage and (b) negative voltage bias. As simulation indicates, the microring resonance wavelength shifts towards shorter wavelength noticeably under the forward bias. This blue shift confirms its carrier-charge effect rather than thermal effect that leads to red shift. The measured resonance wavelength shift under different forward bias agrees well with the simulation, as shown in Fig. 9(c). A relative low quality factor $Q \sim 2300$ is measured due to rough sidewall-induced scattering loss and large coupling coefficient between the resonator and bus waveguide. Nevertheless, over 10 dB extinction ratio, which is limited by the resolution of the optical spectrum analyzer, is achieved at 4 V bias. At biases ranging from 4 V to −4 V we measured extremely low leakage currents of 20–50 fA, i.e., leakage current density of 0.7–1.8 nA/cm², using a high-resolution Keithley current meter. The wavelength shift of 0.055 nm/V under forward bias can be translated into a tuning power efficiency of 8 kW/GHz or 1.6 nm/pW (assuming 35 fA average leakage current). This is 9 orders of magnitude better than conventional thermal tuning in the μW/GHz (sub-nm/mW) regime [53,54]. For practical purposes the DC tuning power consumption can be neglected. This also allows a pre-set bias to be applied in order to enable resonance tuning to both shorter and longer wavelengths.

---

**Fig. 8.** (a)–(f) Schematic process flow to fabricate heterogeneous MOS microring resonators; (g) Top-View microscopic image and (h) SEM cross-sectional image of a fabricated device.

**Fig. 9.** Spectra of heterogeneous MOS microring resonator under (a) positive (0–4 V) and (b) negative (0–−4 V) bias voltage; (c) Measured and simulated wavelength shift as a function of forward bias voltage; (d) Measured modulation response of the same device.
wavelength. As expected, zero or little shift is observed within $-4 \text{ V}$ bias in Fig. 9(b), indicating very small carrier-charge effect observed when the MOS capacitor is under inversion mode. Since the carrier-charge effect in the MOS capacitor is a fast electric field effect, as shown in numerous demonstrated pure Si MOS modulators with speed up to 40 Gbps [46,47], high-speed modulators can be built on this heterogeneous MOS platform [51,52]. Due to a large electrical resistance ($\sim 350 \mu \Omega$) limited by 50 nm-thick, 1 $\mu$m-wide Si pedestal between the p $+$ Si region and p-type Si waveguide edge, the device modulation bandwidth is seriously limited by the RC constant, which results in a 3 dB bandwidth of 0.92 GHz. It is close to the measured bandwidth of 1.2 GHz in Fig. 9(d). By reducing the series resistance (e.g., p-type waveguide background doping at $3 \times 10^{17} \text{ cm}^{-3}$, 100 nm-thick pedestal, 100 nm distance from waveguide edge to p $+$ Si region) and shrinking waveguide width to 0.6 $\mu$m, the resulting electrical resistance $\sim 14 \Omega$ and capacitance $\sim 0.37 \text{ pF}$ will boost the RC-limited bandwidth to $\sim 30 \text{ GHz}$.

We also note here that the layer thickness and doping level in III-V and Si here for both simulation (Fig. 6) and experiment (Fig. 9) are based on our current heterogeneous laser design for proof-of-concept purpose. It demonstrates manufacturing simplicity to co-integrate the heterogeneous MOS modulators/switches with other heterogeneous devices, e.g., lasers, amplifiers or photodetectors, in the same platform and fabrication process. A larger absolute index change, i.e., phase modulation efficiency, is attainable when reducing the gate oxide thickness and optimizing the n-type III-V and p-type Si layer accordingly to place the gate oxide to the center of the mode and maximize the optical confinement in the MOS capacitor [52]. FCA loss can also be reduced by lowering the doping level in III-V to improve resonator Q factor or reduce insertion loss in MZI devices.

3.2. Three-terminal heterogeneous lasers

With the successful demonstration of phase tuning and high-speed modulation potential in the heterogeneous MOS microring resonators, the same MOS capacitor is integrated in the heterogeneous lasers to realize a novel three-terminal diode laser structure for the first time [16]. A simplified Fabry-Perot (FP) laser model is shown in Fig. 10(a). In addition to the conventional current injection to the optical gain medium where lasing modes overlap to build up the optical intensity inside the resonance cavity, the integrated MOS capacitor introduces an independent plasma dispersion effect to the laser operation as a new control freedom for tuning and modulation. While it is convenient to be implemented in a heterogenous laser, the concept is generic to many other diode laser materials and structures. Fig. 10(b) shows the 3D device schematic. The laser resonator comprises an InAlGaAs MQW-based p-i-n diode laser epitaxial structure above a concentric Si disk resonator. A Si bus waveguide is placed adjacent to the microring to evanescently couple a fraction of power out of the laser cavity. Electrons and holes are injected through the terminals N and P1, respectively, and recombine in the MQW active region. A third terminal (P2) uniquely to our design is added to the Si layer outside the microring laser. As highlighted in the schematic device cross-section (Fig. 10(c)), the Si layer and n-InP layer with a dielectric layer in between form a MOS capacitor, as the previous section describes. When a bias is applied to this capacitor through terminal P2 and shared common ground terminal N the concentration of free carrier (electrons and holes) in InP and Si changes either by accumulation, depletion, or inversion. A subsequent change in the FCA loss and the refractive index of Si and InP is induced. Fig. 10(c) shows that the resonator’s fundamental transverse electric (TE) mode overlaps the capacitor region so that the laser threshold current, output power, and lasing wavelength will change as a result of the plasma dispersion effect.

Fig. 11(a) and (b) show the simulated 2D hole distribution under different bias conditions in such a laser geometry with a common ground terminal N. In Fig. 11(a) P1 and P2 are not biased, whereas they are set to 2 and 8 V, respectively, in Fig. 11(b). The same air trench is etched through the top Si layer to significantly reduce the total capacitance from 10.5 pF to 1.23 pF in a 50 $\mu$m-diameter device for a fast capacitor response, and eliminates the high-order transverse modes in the heterogeneous resonator. Without the bias in P2, the lightly-doped p-Si is largely depleted because of the highly-doped n-InP on the other side, introducing negligible background FCA in p-Si. Once a forward bias is applied to P2, free holes quickly accumulate in Si and their concentration increases exponentially when approaching the oxide/Si interface (inset of Fig. 11(b)). Simulations on this heterogeneous microring laser structure (50 $\mu$m-diameter, 15 nm Al$_2$O$_3$ as gate oxide) reveals that under 6 V bias, a resonance shift of 0.17 nm and FCA over 7 dB/cm. These changes in phase and cavity loss introduce a new degree of freedom to control the laser operation without adjusting the laser injection current or applying heat, which degrades diode laser performance in general.

Device fabrication [16] is largely identical to the process to make heterogeneous MOS resonators in Fig. 8 since it is a major advantage that two type of devices can be conveniently integrated together with the same III-V epitaxial material stack, bonding-resulted MOS capacitor structure, and many other common processing steps. Upon forming the passive waveguide components and doped p $+$ Si sections, the same gate oxide deposition and wafer bonding are used to transfer the laser epitaxial structure onto the SOI substrate, i.e., Fig. 8(a)–(c). Then a self-aligned process defines the microring mesa. After removing the excess III-V material outside the mesa to expose the Si surface, P1, P2 and N electrodes are formed sequentially. Similar dielectric encapsulation process and thick probe metal pad formation in Fig. 8(f) are conducted to finish the fabrication. Fig. 12(a) and (b) show respectively a top-view and cross-sectional SEM images of a fabricated device. 3 $\mu$m wide III-V mesa and 1.5 $\mu$m wide Si waveguide underneath allow good contact resistance at P1 terminal and small MOS capacitance.

Devices with diameters from 30 to 50 $\mu$m are fabricated in the same chip. Depending on the lasing direction, laser output light propagates through a Si bus waveguide to an on-chip heterogeneous photodetector (PD) at one end or to a Si grating coupler at the other end. Fig. 13(a) shows the CW light-current-voltage (LIV) characteristic of a 40 $\mu$m-diameter laser at room temperature (RT). When the MOS bias rises from 0 to 4 V we observe a threshold increase from 8 to 10 mA and a reduction of over 50% in output power (based on the photocurrent from the on-chip photodetector). The IV curves for laser p-i-n diode at different MOS biases are nearly identical with an average differential
Fig. 11. Simulated 2D hole distribution profiles with (a) no bias in all terminals and (b) forward bias of 2 V in P1 and 8 V in P2. Inset: zoom-in hole distribution profile around the MOS capacitor region; (c) resulted resonance shift and FCA in a 50 μm-diameter device under 0 to 6 V MOS bias voltage.

Fig. 12. Top-view and b. cross-sectional view SEM images of a fabricated laser. Highlighted color sections: red – a. microring laser, b. active region; purple – Si; blue – III-V; green – BOX; yellow – selected metal contacts. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this article.)

Fig. 13. (a) LIV characteristic and (b) spectra of a 40 μm-diameter device around 1560 nm lasing wavelength at different MOS bias voltage; (c) spectra of a 50 μm-diameter device around 1340 nm lasing wavelength at different MOS bias voltage; (d) Thermal chirp correction demonstration in MOS-type microring lasers (e) lasing wavelength vs. leakage power between the terminals P2 and N in device with 30, 40 and 50 μm diameters. Inset: leakage current density vs. capacitor bias voltage.

Fig. 14. Schematic of (a) a heterogeneous MOS-type microring laser with QD active region and integrated thermal shunt, and (b) an array of such microring lasers with a DBR at the bus waveguide to enable unidirectional operation, and (c) an on-chip heterogeneous QD mode-locked comb laser with front and back DBR mirrors and integrated MOS capacitor for wavelength tuning/locking.
resistance of 34.3Ω, indicating no interference to the normal carrier injection into the laser active region. A discontinuity in the Si bus waveguide due to a fabrication error results in an extremely low laser power reaching the on-chip photodetector or grating coupler. Due to failure to accurately determine the laser output power, we were unable to extract the MOS capacitor-induced FCA experimentally. The lasing wavelength blue-shifts from 1559.39 to 1559.21 nm in Fig. 13(b), confirming a result of the plasma dispersion. Similar wavelength shift is also seen in another chip with a different MQW composition and lasing around 1340 nm in Fig. 13(c). Good extinction ratio over 30 dB is observed in both devices and the laser linewidth is limited by the instrument.

This new lasing wavelength control is very useful to fine tune or lock the wavelength for fabrication imperfection [55] or environment temperature variation. The dashed lines in Fig. 13(d) show the common thermal chirp in diode lasers due to injection-current heating, similar to a result of ambient temperature rise. Laser chirp increases crosstalk with adjacent WDM channels, and exacerbates chromatic dispersion so that the useful reach of interconnects is reduced. With appropriate MOS bias, the wavelength shift from increased injection current can be offset completely (solid lines).

We measured a direct-current (DC) leakage current density as small as few μA/cm² with within 4 V MOS bias (Fig. 13(e) inset), but 100× higher leakage was also observed in some devices. This is much higher than the one measured in the previous MOS microring resonators, and is likely due to some poorly passivated spots between terminals due to much larger surface topography change. By plotting the leakage power (leakage current × MOS voltage) vs. wavelength shift in Fig. 13(e), we obtain a phase tuning efficiency up to 6 nm/nW. This is still an impressive improvement of seven orders of magnitude over previous results that used a more conventional thermal or carrier injection approach [56–58]. To the best of our knowledge, our result is the most efficient phase tuning demonstrated in any diode lasers.

Low output power prevents us from performing a dynamic characterization of the laser at this time. However, we expect a high-speed amplitude or frequency modulation to be achieved in our MOS-type laser simply by directly changing the MOS bias. Modulating the cavity loss is an efficient way to modulate a laser at high speeds [59]. Smaller devices with higher resistance-capacitance bandwidth limit will favor high-speed operation if the resistance increase is minimal through smart doping and material thickness control. The proportion of the MOS-induced FCA to the total laser cavity loss will determine the amplitude modulation extinction ratio. Such modulation would also result in an enhanced dynamic performance and a low power consumption. The same concept can be applied to heterogeneous or compound semiconductor material systems, and interface oxide can be formed by bonding or oxidation. And may lead to new photonic and electronic devices as well.

4. Summary and future works

In this paper we reviewed our recent progress to utilize a superior laser gain material and build a novel laser structure on Si substrates for on-chip, robust light sources in datacom applications. Fundamentally better material qualities in high-temperature optical gain, larger optical gain bandwidth, lower threshold current density and RIN make QD material a perfect diode laser material for large wavelength-count comb laser operation in a perfecta diode harsh environment. The heterogeneous integration platform enables much easier optical coupling to the Si photonic integrated circuits, and 300 mm or larger size wafers potentially in a CMOS production line promises a significant reduction in fabrication cost.

The heterogeneous MOS capacitor structure enables the electrical usefulness of the Si portion and interface oxide in this heterogeneous integration platform. Not only can the advantage of the III-V material property give enhanced device performance, e.g., larger phase change and lower FCA for optical modulator application, but it also enables a new three-terminal laser structure with additional control freedom, essentially power-free tuning and great potential for zero-chip high-speed direct laser modulation. The compatibility of the material and fabrication with conventional heterogeneous photonic devices result in little additional cost in fabrication and chip complexity. Furthermore, a thermal shunt design [37] we previously demonstrated can be readily formed by extending the metal P2 into a deeply-etched through-BOX trench outside the device mesa to route the joule heat to the Si substrate for effective thermal management as schematically shown in Fig. 14(a). The combination of QD laser material and thermal shunt design is expected to further reduce the impact of external environment and internal joule heating to laser efficiency and wavelength shift to the minimum in near future.

Our research focus is now to develop a heterogeneous QD microring laser array with a reflector, e.g., DBR, at one end of the bus waveguide for unidirectional operation (Fig. 14(b)) [60], and simultaneously an on-chip heterogeneous QD mode-locked comb laser with integrated MOS capacitor (Fig. 14(c)). The MOS-type wavelength fine tuning within a DWDM channel spacing (e.g., 80 GHz) perfectly enables a power-free channel locking mechanism to guarantee a large number of low-RIN light streams consistently coming out at the fixed wavelength grid. We are developing both the zero-chirp directly modulated microring laser arrays and mode-locked comb lasers with external modulators to meet the bandwidth, power consumption and data transmission distance requirement for different application scenarios and computing architecture designs in datacenters and HPCs.

Reference

Optical Fiber Technology 44 (2018) 43–52

D. Liang et al.


D. Liang, J.E. Bowers, Robust hybrid quantum dot laser with ultra-high tuning efficiency, CLEO, San Jose, CA, USA, 2016, paper SThI4.4.


