Heterogeneous Silicon Photonic Integrated Circuits

Tin Komljenovic, Michael Davenport, Jared Hulme, Alan Y. Liu, Christos T. Santis, Alexander Spott, Sudharsanan Srinivasan, Eric J. Stanton, Chong Zhang, and John E. Bowers, Fellow, IEEE

(Tutorial Review)

Abstract—We review recent breakthroughs in the silicon photonic technology and components, and describe progress in silicon photonic integrated circuits. Heterogeneous silicon photonics has recently demonstrated performance that significantly outperforms native III–V components. The impact active silicon photonic integrated circuits could have on interconnects, telecommunications, sensors, and silicon electronics is reviewed.

Index Terms—Heterogeneous silicon platform, integrated optoelectronics, optoelectronic devices, semiconductor lasers, silicon-on-insulator (SOI) technology, silicon photonics.

I. INTRODUCTION

HETEROGEOUS silicon photonics, due to its potential for medium- and large-scale integration, has been intensively researched recently. Recent developments have shown that heterogeneous integration not only allows for a reduced cost due to economy of scale, but also allows for same or even better performing devices than what has previously been demonstrated utilizing only III–V materials. Furthermore we believe that optical interconnects are the only way to avoid the scaling limitation in modern processors, and that heterogeneous silicon photonics with on-chip sources is the best approach in the long term as it promises higher efficiency. We address both claims in sections that follow.

In this paper we plan to briefly address heterogeneous silicon approaches, and point-out that the heterogeneous silicon platform is more than just III–V on silicon (Section II). We then present a short rationale on answering the question of why to use heterogeneous silicon photonics for interconnects (see Section III). We address passives (see Section IV), outlining low-propagation loss and tight confinement benefits for integration, briefly touch upon polarization control on-chip as it is crucial for many new applications including polarization-diversity transceivers (see Section IV-A) and we address tuning and temperature stability (see Section IV-B). After passives, we highlight some recent record-performing active devices (see Section V) and provide a brief overview of sensors-on-chip progress (see Section V-A). Finally we address narrow-linewidth lasers (see Section VI). We overview commonly used linewidth measurement techniques (see Section VI-A) and then present record results together with theoretical explanations for record breaking results; first for single-wavelength lasers (see Section V-B) and then for widely-tunable ones (see Section VI-C).

We give special attention to narrow-linewidth lasers, as we believe that this is one of the areas that perfectly demonstrates the potential and benefits of utilizing heterogeneous approaches as recent performance of heterogeneously integrated narrow-linewidth lasers has far surpassed linewidth results demonstrated with III–V semiconductor lasers. Another area where this platform clearly shows its advantage are heterogeneously integrated high-power, high-speed photodiodes. We should point out that, due to the vast scope of research in this area and the high volume of publications, this paper, although covering a wide range of heterogeneous silicon photonics, will not provide an exhaustive list of published papers in the field.

II. HETEROGENEOUS SILICON PHOTONICS

One of the main motivations behind silicon photonics lies in its potential for bringing the large wafer size, volume throughput, and cost reduction of silicon manufacturing to photonic components. The silicon-on-insulator (SOI) platform fabrication infrastructure is compatible with CMOS technology and is highly accurate and mature, leading to a robust, high yield and reproducible technology. Photonic integrated circuits (PICs) operating in the telecommunication windows around wavelengths of 1.3 and 1.55 μm are perfect candidates for the SOI platform, which offers excellent waveguide capabilities as described in Section IV about passive components. Prime applications for SOI-based PICs are telecommunication, interconnects, and lately sensors or sensor-systems on chip.

The SOI platform by itself offers an almost complete suite of photonic components, including filters, (de)multiplexers, splitters, modulators, and photodetectors. However, electrically pumped efficient sources on silicon remain a challenge due to this material’s indirect bandgap. A way to introduce efficient electrically pumped sources is to utilize III–V gain regions placed directly on silicon. There are three approaches to achieve this [1]. One approach [2] uses III–V chips bonded on silicon with coarse alignment and subsequently processed on the Si wafer scale (see Fig. 1). This has the advantage of minimizing the area requirement of III–V material, thus minimizing the

0733-8724 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
III–V epitaxial wafer cost. It has also the great advantage that very different epitaxial stacks (and not necessarily III–V materials alone) can be integrated together and processed simultaneously [5]. A second approach is the direct epitaxial growth of III–V layers on silicon or SOI using intermediate buffer layers, typically Ge and strained superlattices, to minimize dislocations propagating into the active region [6], [7]. The use of quantum dot (QD) laser gain material can minimize the effect of threading dislocations on threshold and output power, as efficient capture and localization of injected carriers by individual QDs greatly reduces non-radiative recombination at dislocations. Recent results show tremendous improvement in laser lifetime of epitaxially grown QD lasers on silicon over similar quantum well lasers [8]. Although the improvement is significant, the lifetime is not yet adequate for most applications, but further progress is expected. A third approach is to combine these approaches: one can grow InAs QD gain material on silicon and then bond to patterned SOI wafers for efficient waveguide coupling and PIC fabrication. This could be on a wafer scale, directly up to 300 or 450 mm diameter wafers. It solves the wafer size limit for wafer scale III–V to silicon bonding, which has been restricted to 150 mm diameter, the maximum size of InP wafers available to date. Further, growth wafer reclamation when using smart cut or highly selective etching to perform substrate removal should be possible for cost reduction [1].

For it to be widely used, the heterogeneous-silicon PIC platform has to offer low power, low cost, high capacity, high volume, high yield and high reliability. Silicon wafer scale processing can offer high capacity, high volume and low cost. Publicly available studies suggest that the heterogeneous silicon platform is suitable for making reliable active optical devices, including lasers, modulators and photodetectors [9] and low power devices have readily been demonstrated [10]–[14] addressing all of the above requirements.

Heterogeneous integration on silicon can accommodate not only III–V materials that provide gain, detection and modulation, but can also include more exotic materials (see Fig. 2) such as LiNbO$_3$ and Ce:YIG bringing the promise of high performance modulators, nonlinearities (second-harmonic generation, parametric amplification, and entangled photon generation) and magnetic properties. LiNbO$_3$ based modulators on heterogeneous silicon platform are demonstrated either by thin-film direct wafer bonding [21] or by thin film bonding via benzo-cyclobutene [22]. Performance, at this stage, is still limited at the few GHz range due to non-optimized designs with high RC constants and acousto-optic resonances, but substantial improvements are expected with more refined designs. Magneto-optical materials can be combined with silicon photonics to obtain optical non-reciprocity. Surface activated direct bonding [23] and direct deposition of magneto-optic material [24] have been demonstrated. The performance of the MZI silicon waveguide optical isolator was demonstrated with an isolation of 30 dB at a wavelength of 1548 nm. In addition, a maximum isolation of 15.3 dB was obtained in an optical circulator fabricated with a silicon MZI waveguide [25].

Light generation on silicon chips can also be achieved by integration of rare-earth-ion (RE) doped dielectric thin films. The advantages of RE approach would be direct deposition onto silicon substrates, amplification of higher bit-rate signals without patterning effects, higher temperature of operation and inherent narrow linewidth. Disadvantages are significantly lower gain in range of few dB/cm and a need for optical pumping [97].

Some intrinsic limitations of all-silicon devices can also be overcome by combining conventional SOI waveguides with organic cladding material in a concept of silicon-organic hybrid integration. This approach has demonstrated extremely efficient modulators consuming only 0.7 fJ/bit to generate 12.5-Gbps data stream [98].

III. Optical Interconnects

Although operating frequency and sequential processor performance has slowed down, the scaling of Moore’s Law continues by increasing the number of processor cores while keeping total power consumption constrained due to total power budget roughly assumed to be around 200 W [96]. This is achieved by describe this wafer scale approach to integration, and “hybrid” is more commonly used for bonding or soldering individual die onto a common substrate. 

For it to be widely used, the heterogeneous-silicon PIC platform has to offer low power, low cost, high capacity, high volume, high yield and high reliability. Silicon wafer scale processing can offer high capacity, high volume and low cost. Publicly available studies suggest that the heterogeneous silicon platform is suitable for making reliable active optical devices, including lasers, modulators and photodetectors [9] and low power devices have readily been demonstrated [10]–[14] addressing all of the above requirements.

Heterogeneous integration on silicon can accommodate not only III–V materials that provide gain, detection and modulation, but can also include more exotic materials (see Fig. 2) such as LiNbO$_3$ and Ce:YIG bringing the promise of high performance modulators, nonlinearities (second-harmonic generation, parametric amplification, and entangled photon generation) and magnetic properties. LiNbO$_3$ based modulators on heterogeneous silicon platform are demonstrated either by thin-film direct wafer bonding [21] or by thin film bonding via benzo-cyclobutene [22]. Performance, at this stage, is still limited at the few GHz range due to non-optimized designs with high RC constants and acousto-optic resonances, but substantial improvements are expected with more refined designs. Magneto-optical materials can be combined with silicon photonics to obtain optical non-reciprocity. Surface activated direct bonding [23] and direct deposition of magneto-optic material [24] have been demonstrated. The performance of the MZI silicon waveguide optical isolator was demonstrated with an isolation of 30 dB at a wavelength of 1548 nm. In addition, a maximum isolation of 15.3 dB was obtained in an optical circulator fabricated with a silicon MZI waveguide [25].

Light generation on silicon chips can also be achieved by integration of rare-earth-ion (RE) doped dielectric thin films. The advantages of RE approach would be direct deposition onto silicon substrates, amplification of higher bit-rate signals without patterning effects, higher temperature of operation and inherent narrow linewidth. Disadvantages are significantly lower gain in range of few dB/cm and a need for optical pumping [97].

Some intrinsic limitations of all-silicon devices can also be overcome by combining conventional SOI waveguides with organic cladding material in a concept of silicon-organic hybrid integration. This approach has demonstrated extremely efficient modulators consuming only 0.7 fJ/bit to generate 12.5-Gbps data stream [98].

III. Optical Interconnects

Although operating frequency and sequential processor performance has slowed down, the scaling of Moore’s Law continues by increasing the number of processor cores while keeping total power consumption constrained due to total power budget roughly assumed to be around 200 W [96]. This is achieved by
increasing the number of cores and transistors on die (modern processors are approaching 10 billion transistors on a single chip), making them more power efficient and utilizing more advanced power conservation technologies. The scaling of number of cores is expected to have a limit due to Amdahl’s Law that limits the performance gain per core as number of cores is increased, but currently we are still in the exponential growth phase. As the number of cores is increasing, the capacity of interconnects between individual cores and between the processor and outside world also has to increase. Currently this communication is done by electrical interconnects and once the wiring fills all the available space, it has been shown that the capacity cannot be increased by changing the system size [15]. The limit to the total number of bits per second of information that can flow in a simple digital electrical interconnection is set only by the ratio of the length of the interconnection to the total cross-sectional dimension of the interconnect wiring - the “aspect ratio” of the interconnection. This limit is largely independent of the details of the design of the electrical lines, and because it is scale-invariant, it cannot be changed by either by growing or shrinking the system. Performance can be improved by using repeaters, advanced coding and multilevel modulations, but these techniques also have limitations, one of which is the energy efficiency as total power has to be kept to around 200 W of thermal budget. Optical interconnects, as illustrated in Fig. 3, can solve this problem since they avoid the resistive loss physics of copper even sooner. The baud rate of a single channel will not likely scale much with the move to optics as the accompanying fast electronics increases power consumption and is likely to be around 25 Gb/s over the next decade [17]. As optics is also limited by the size of waveguides, there is a limit in the number of optical channels that can be packed around the edge of chip. To satisfy the bandwidth requirements, future systems will generally be wavelength-division multiplexed (WDM) ones. In order to generate a large number of wavelengths from a single source, comb lasers are considered [17]. A challenge with densely packed WDM channels is thermal control of individual components. Active control uses energy and may not be suitable for integration with devices generating around 200 W of heat in a small volume, while passive temperature control brings its own set of tradeoffs. We briefly overview thermal tuning and athermal operation in Section IV-B. To further increase available bandwidth, vertical stacking of waveguides on chip could be employed.

Close integration of electronics with PIC (so called “Smart Photonics”), as shown in Fig. 4, also benefits photonics as it allows low-power, high-impedance drivers avoiding low-efficiency 50 Ω terminations [10], and allows for smart photonic circuits with self-calibration and active-feedback control on-chip in real time with reduced power consumption [20].
To conclude the interconnect section, it is clear that electrical interconnects will not be able to keep up with bandwidth demand in the long term. Optical interconnects provide an alternative that could overcome this limitation once performance on all three key metrics is sufficiently better than what is available in copper. The trend is evident as optical interconnects continuously push copper links to shorter and shorter lengths due to their inherent limitations. Furthermore, the close integration of driving electronics with PIC allows added flexibility and reduced power consumption.

IV. SILICON PHOTONICS PASSIVES

The range of passive optical components demonstrated on silicon is extremely impressive and includes waveguides, couplers, multiplexers, polarization control devices, filters, resonators, etc., [29], [32]. Silicon ($n_{Si} = 3.48$) and its oxide ($n_{SiO_2} = 1.48$) form high-index contrast, high-confinement waveguides ideally suited for medium to high-integration and small passive devices in their transparency wavelength range, including the most important 1300 and 1550 nm communication bands (see Fig. 5). Tight bends with losses lower than 0.09 dB for bending radii of only 1 μm have been demonstrated more than ten years ago [33], and propagation losses below 1 dB/cm have been shown in both communication bands allowing for compact and excellent performing passive devices.

Record low losses in silicon, less than 3 dB/m at 1600 nm, have been reported in [34], calculated by the internal $Q$ factor of silicon ring resonators of $2.2 \times 10^7$. The authors further conclude that the measured loss is still limited by bend loss in this low-confinement configuration with very low sidewall interaction, so the pure propagation loss could be even lower. Such record values are probably not suitable for wide integration as the radius of said ring was 2.45 mm. By increasing the confinement, the bend loss goes down at the expense of higher propagation loss, mainly due to scattering at vertical sidewalls. Nevertheless, complex devices with much smaller bend radii and measured propagation loss <0.5 dB/cm in C-band [35] and <0.7 dB/cm in O-band [36] have been demonstrated.

For even lower losses, one can turn to the Si$_3$N$_4$ waveguide platform that offers more than two orders of magnitude lower propagation loss (as low as 0.045 dB/m) [37]. The Si$_3$N$_4$ waveguides can readily be integrated with the silicon platform with coupling losses between 0.4–0.8 dB per transition depending on taper bandwidth [38]. Another benefit in using Si$_3$N$_4$ waveguides is the absence of two-photon absorption and the resulting free-carrier absorption present in Si waveguides, but the downside is lower confinement and larger minimum bend radii.

The heterogeneous silicon platform has been shown to be usable with an extremely broad wavelength range. The platform is capable of combining optical frequency bands spanning 4.2 octaves from ultraviolet to mid-wave infrared into a single, low $M^2$ output waveguide as demonstrated in (see Fig. 6) [39]. Using two waveguide types (silicon nitride and silicon), the prohibitively high material losses that would be present in a single waveguide platform for UV to mid-IR wavelengths are avoided while providing a platform compatible with heterogeneous integration of laser sources covering the same spectral range (see Fig. 6). This concept shows that an integrated array of lasers spanning UV to mid-IR bands spectrally combined into a single output waveguide for high power and ultra-broadband applications is feasible.

A. On-Chip Polarization Handling Devices

A polarization beam splitter (PBS) is a key component for separating or combining two orthogonal polarization modes (i.e., TE and TM polarizations), which is very important for PIC, e.g., modern 100-G transceivers utilize dual-polarizations for reducing the baud-rate. One can also envision polarization discrimination based sensors whose front-end is fully integrated on chip.

PBSs have been reported using various structures including multimode interference structures, directional couplers (DC), Mach–Zehnder interferometers, and photonic-crystal (PhC)/grating structures [32]. Most of the realizations are quite long physically, making integration more difficult, or require introduction of large stress or the use of highly birefringent
materials such as LiNbO3, III–V semiconductor compounds or liquid crystal. Another approach is using PhC or out-of-plane gratings. An overview of PBS structures with typical lengths is given in [40]. The same paper also proposes an asymmetrical directional coupler that utilizes evanescent coupling between a strip-nanowire and a nanoslot waveguide. Although such a design allows for a very short structures (less than 10 μm), the difficulty is the fabrication of the nano-slot. A structure with relaxed fabrication tolerance based on an asymmetrical bent dc using SOI nanowires is introduced theoretically [41] and demonstrated experimentally [42]. The bent DC is designed to cross-couple TM polarized light completely while there is almost no coupling for TE polarization. This is done by designing the bent coupling section to be phase-matched only for TM polarization. Once polarizations are separated into two waveguides, one of the polarizations can be rotated so both waveguide arms have identical polarizations.

Realization of on-chip polarization rotation is challenging as planar waveguides usually maintain polarization. A polarization rotator using an SOI nanowire with a cut corner has been demonstrated [43]. The cut corner almost fully hybridizes two lowest-order modes, so light entering the polarization rotator excites these two hybridized modes and two-mode interference takes place. Polarization rotation is made possible in only ~7 μm of length. A modified polarization rotator configuration fabricated using a four etch-step CMOS-compatible process including layer depositions on a SOI wafer has been demonstrated experimentally [44].

These two functions can be integrated in a single structure as proposed in [45]. Here an ultra-compact polarization splitter-rotator (shown in Fig. 7) is proposed by utilizing a structure combining an adiabatic taper and an asymmetrical directional coupler. The input TE polarization does not change when it goes through the adiabatic taper structure and is not coupled to the adjacent narrow waveguide due to phase mismatch. Fundamental TM polarization launched at the same port is first efficiently converted to higher-order TE mode in the adiabatic taper structure and is then coupled to TE fundamental mode of the adjacent narrow waveguide. The total device length is less than 100 μm and the design utilizes only one etch step.

B. Tuning and Thermal Stability

Devices based on silicon are readily thermally tunable owing to the large thermo-optic coefficient (dn/dT) of silicon. A value of \(dn/dT\) of \(1.87 \times 10^{-4} /K\) at room temperature and 1.5 μm is generally reported and the coefficient increases at shorter wavelength to \(1.94 \times 10^{-4} /K\) at 1.3 μm [46]. The highly-efficient temperature tuning has been used to realize tunable filters, switches, add/drop multiplexers, and widely-tunable lasers among others. Due to being thermal in nature, tuning is relatively slow in the μs to ms range, as summarized for switches [47]. There is usually a trade-off in terms of tuning efficiency and speed, and the trade-off is related to thermal isolation of the tuning section. In the case of low thermal resistance, the heaters are less efficient in terms of power, but are faster as heat can dissipate more quickly. In terms of efficiency, switches have been demonstrated with switching powers of only 0.6 mW at the expense of switching speed that is 3.6 ms [13]. Tuning ranges of 20 nm (see Fig. 8) for a single ring structure have been demonstrated [48] and wider tuning range is easily obtained.
by utilizing a Vernier effect. We return to wide-tunability in Section VI-C.

The high thermo-optic coefficient can also be a disadvantage if one wants to realize athermal devices that do not require power hungry and inefficient thermoelectric cooling units. Various athermal designs have been studied [49]–[51]. The simplest athermal technology is the athermal waveguide. By co-integrating waveguides with different thermo-optic coefficients, and adjusting the length of each waveguide type, the phase difference between the neighboring waveguides can in principle be independent of temperature, over a range in excess of 50 K. Such an approach has extensively been used in e.g. athermal arrayed-waveguide-gratings [52]. Unfortunately, the same approach is less well suited for devices whose response depends on the absolute phase shift such as ring resonators or Bragg gratings. One method of thermal stabilization of such structures is to adopt an upper cladding made of a material with negative thermo-optic coefficient to compensate for the positive thermo-optic coefficient of silicon. By using polymer claddings with $dn/dT = -2.65 \times 10^{-4} /K$, prototypes with temperature dependent resonant wavelength shift of only 0.5 pm/K were demonstrated [53]. Titanium dioxide (see Fig. 9) also has a strong negative thermo-optic coefficient of $\sim (1-2) \times 10^{-4} /K$, and is compatible with CMOS processes offering superior reliability than polymers. Thermal sensitivities of only $-2.9$ pm/K were demonstrated [51]. Athermal laser designs are described in [49] and uncooled athermal WDM lasers with channel spacing of 200 GHz have been demonstrated operating from 20 to 80 °C without significant change of wavelength [5].

V. HETEROGENEOUS SILICON PHOTONICS ACTIVE DEVICES

Since the first report of the electrically pumped heterogeneous silicon laser [54], a full suite of heterogeneous silicon III–V components has been developed [29]. Here we shall briefly mention some of recent record performing heterogeneous-silicon devices.

Ridge waveguide lasers at 1.3 μm fabricated from InAs QDs grown on silicon by molecular beam epitaxy (see Fig. 10) have shown room temperature continuous wave thresholds as low as 16 mA, output powers exceeding 176 mW, and lasing up to 119 °C [6]. The reliability of epitaxially grown InAs/GaAs lasers on silicon has improved tremendously and is now quoted at 4600 h [8], bringing the promise of wafer scale growth on silicon and allowing direct bonding with wafer sizes of up to 450 mm once reliability is further improved.

A broadband superluminescent III–V-on-silicon light-emitting diode (LED) with 292 nm of 3 dB bandwidth and on-chip power of $-8$ dBm was demonstrated [55]. To achieve the large bandwidth, quantum well intermixing and multiple die bonding of InP on a silicon photonic waveguide circuit were combined. The device consists of four sections with different bandgaps, centered around 1300, 1380, 1460, and 1540 nm (see Fig. 11). The fabricated LEDs were connected on-chip in a serial way, where the light generated in the smaller bandgap sections travels through the larger bandgap sections.

Colliding-pulse mode-locked laser diodes on heterogeneous silicon platform were investigated and characterized [56]. Reduction in microwave linewidth using two techniques has been
demonstrated. First, reducing the number of quantum wells reduces the confinement factor, thereby lowering the spontaneous emission contribution to the linewidth. Second, a ~4 cm long on-chip passive feedback cavity is used to provide optical feedback to stabilize the laser and further reduce the linewidth. The linewidth achieved for passive mode locking at 17.36 GHz using the above two techniques is 29 kHz.

Low threshold (8.8 mA) and high speed (9.5 GHz) short cavity distributed feedback (DFB) heterogeneous silicon lasers were demonstrated [57]. The large direct modulation bandwidth of the heterogeneous short cavity DFB shows its potential as a low cost and low power laser source.

There has also been progress on low threshold heterogeneous silicon microring lasers [14]. The improvement in performance comes by selective reduction of the active region volume. This is done by appropriately undercutting the multiple quantum well active region to force carriers to flow toward the outer edge of the microring for better gain/optical mode overlap. A reduction of the threshold to 3.9 mA and up to 80% output power enhancement is observed, mainly due to the improved injection efficiency. Thermal management of heterogeneous silicon ring lasers was also explored [58]. By improving the thermal impedance of microring lasers with a novel double gold thermal shunt design (see Fig. 12), continuous-wave lasing up to 105 °C was demonstrated. This high temperature operation is important for these lasers are to be used in interconnect applications in uncooled environments, such as data centers. Further improvement is possible, by using metals with higher thermal conductivity, such as copper, which is also CMOS process compatible, to optimize the efficiency of the thermal shunts.

The silicon photonics transmitter chip was successfully integrated with a low power CMOS driver chip with a flip-chip bonding method [59]. The transmitter chip (see Fig. 13) with 16 channel electro-absorption modulator (EAM) array was aligned and bonded with high speed very large scale integration (VLSI) circuits with high yield. The heterogeneous EAM with 100 μm length had an extinction ratio larger than 6 dB for 1 V voltage swing with the optical bandwidth over 20 nm in C band. The integrated EAM had an open eye at 12.5 Gb/s with 3.5 dB extinction ratio at 1510 nm, with the maximum driving voltage swing from 0–2.4 V.

Extremely low thresholds of 31 μA were shown for continuous-wave operation of lambda-scale embedded active-region PHC lasers (LEAP, Fig. 14) at room temperature fabricated on a Si wafer [12]. As LEAP has emission in a direction normal to the wafer, the light was coupled via grating couplers. The maximum output power is 0.27 μW at an injected current of 200 μA, which includes around 10 dB of optical coupling loss of the measurement setup. The maximum output power from the device is estimated at few μW.

The heterogeneous silicon platform has also been extended to non-communication wavelengths such as 2 μm [64]. Room temperature, continuous wave 2.0 μm wavelength lasers were
heterogeneously integrated on silicon by molecular wafer bonding of InGaAs quantum wells grown on InP. These heterogeneous silicon lasers operate continuous-wave up to 35 °C, have threshold currents as low as 59 mA and emit up to 4.2 mW of single-facet CW power at room temperature. These lasers should enable the realization of a number of sensing and detection applications in compact silicon photonic systems.

Heterogeneously integrated waveguide-coupled photodiodes on SOI with 12 dBm output power at 40 GHz have been demonstrated [61]. The InP-based modified uni-traveling carrier photodiodes on SOI waveguides have internal responsivity of up to 0.95 A/W and have the highest reported output power levels at multi-GHz frequencies for any waveguide photodiode technology including native InP, Ge/Si, and heterogeneously integrated photodiodes. The reasons for the exceptional performance lie in the added flexibility introduced by heterogeneous process that allows independent change of the widths of the Si waveguide and III–V mesa. This allowed simultaneous reduction of current crowding and a tailored absorption profile to reduce saturation effects via mode and bandgap engineering.

A. Sensors on Chip

The ability to integrate multiple material systems on the heterogeneous silicon platform presents the possibility of designing fully-integrated sensors on chip. Recently there have been proposals for integrated waveguide optical gyroscopes [62] and magnetometers [63]. A highly integrated optical gyroscope using low loss silicon nitride waveguides and integration of all the required active and passive optical elements on a chip is possible with a detection limit on the order of 19°/h/√Hz. This is for an area smaller than 6.5 cm² with a propagation loss of 1 dB/m in a 10 m long waveguide [62]. The analysis of novel highly sensitive optical magnetometers using low-loss silicon nitride waveguides shows that with recent advances in Ce:YIG pulsed laser deposition on silicon nitride, sensitivities on the order of 20 fT/√Hz at room temperature in an area <1 cm² should be possible. All the required elements can be fully integrated on a chip using the heterogeneous silicon platform. By using materials with greater sensitivity to magnetic field, like BiₓCe₃₋ₓFe₅O₁₂, the minimum achievable sensitivity could be further improved by a factor of two [63].

The most complex heterogeneous PIC is a fully-integrated free-space beam-steering PIC (see Fig. 15) consisting of 164 optical components including lasers, amplifiers, photodiodes, phase tuners, grating couplers, splitters, and a photonic crystal lens. The PIC exhibited steering over 23° × 3.6° with beam widths of 1° × 0.6° giving a total of 138 resolvable spots in the far field with 5.5 dB background suppression [65].

VI. NARROW-LINewidth LASERS

Narrow-linewidth is becoming increasingly important in modern communications and sensors. Modern 100G transceivers utilize dual-polarization quadrature-phase-shift-keying in order to send 4 bits simultaneously (two in each polarization) and reduce the symbol speeds to 28 Gbd. Moving to higher transmission-speeds at a single wavelength such as 200G and 400G while keeping the same symbol speeds necessitates using even more advanced modulation formats, such as DP-16QAM or DP-256QAM where QAM stands for quadrature amplitude modulation. Such advanced modulations require lasers and local oscillators for demodulation with very low phase noise, or narrow-linewidth (Fig. 16). For square 16-QAM constellation, the linewidth should be <100 kHz and for square 64-QAM linewidth should be <1 kHz assuming 28 Gbd rate [67]. Traditional III–V lasers had linewidths in MHz range and only recently have been able to demonstrate sub-MHz and finally sub-100 kHz linewidth with careful optimization of the resonator and the gain sections [68]–[74]. It should be noted
that a direct comparison between quoted values is sometimes hard to make as methods to measure and quote linewidths differ. Heterogeneous-silicon lasers have been showing sub-MHz linewidths for some time and recently have shown results significantly surpassing the performance of pure III–V lasers [31], [35], [36], [75]–[82]. A comparison is shown in Fig. 17. In both, single-wavelength and widely-tunable lasers, a key enabler for narrow linewidth with heterogeneous designs lies in utilization of low-loss silicon waveguides and high-Q resonators.

A. Common Linewidth Measurement Techniques

Laser linewidth is often defined in terms of the full-width half-maximum (FWHM), or −3 dB points of the optical field power spectrum. Linewidth of narrow-linewidth lasers cannot be directly resolved with grating-based optical spectrum analyzers, and is usually measured in two ways. One approach is to measure a beat signal between two lasers (heterodyning) or a beat signal between the laser’s direct and sufficiently delayed beam (self-homodyning or self-heterodyning). The beat signal is generated on a fast photodetector used as a mixing element. The second approach uses a frequency discriminator to convert the FM noise to AM noise, which is then also measured by a fast photodetector.

Heterodyning or beating two lasers (see Fig. 18(A)) on a photodetector is probably the simplest method for measuring the linewidth, but requires a stable reference laser (local oscillator) with a linewidth much narrower than the linewidth of the laser to be measured [83]. Furthermore the wavelengths of the laser to be measured and one used as a local oscillator have to be very close in frequency (tens of gigahertz) so their beat spectra can be captured with a fast photodetector. It is assumed that the measured linewidth is the linewidth of the laser under test, as the local oscillator has a much narrower linewidth; theoretically approximated to be a delta-function. Another option is having two almost identical lasers and beating them together, assuming that their linewidths are comparable; then the linewidth can be assumed to be half of the measured value.

The delayed self-homodyne detection (see Fig. 18(B)) provides a method to perform laser linewidth measurement without using a separate local oscillator. Incident light is split into two paths by the interferometer. If the delay ($\tau_d$) of one path exceeds the coherence time ($\tau_c$) of the source, the two combining beams interfere as if they originated from two independent lasers. An advantage of using the same laser as signal and local oscillator is the automatic wavelength tracking/locking, so no additional tuning to align lasers close enough or feedback to stabilize them is necessary. The problem is that the beat signal is downconverted to baseband and electrical spectrum analyzers typically have problems measuring values close to dc and under a few kHz due to 1/f noise and spurs from the ac line. A solution is to introduce a frequency offset in one arm, usually via an acousto-optic
modulator. The offset is in 10 s of MHz to up to 200 MHz range and it shifts the beat signal to an offset frequency where there is no low frequency noise and impairments. Such a measurement configuration is called a delayed self-heterodyne method [84]. The self-heterodyne configuration is again automatically tracked/locked. The delay difference ($\tau_d$) places a limit on observation time, effectively limiting the influence of $1/f$ noise on linewidth measurement. In other words the fiber delay, depending on the FM noise spectrum of the laser, can significantly influence linewidth numbers [85]. As the delay has to be longer than the coherence length of the lasers, measuring very narrow linewidth lasers can be challenging [86], [87]. The described methods give linewidth as a single number often calculated by Lorentzian or Voigt fitting to measured data [85].

A frequency discriminator (see Fig. 18(C)) is another popular technique for measuring laser linewidth. This method directly measures the frequency noise of lasers by converting the frequency fluctuations to intensity fluctuations. This is done by an interferometer that is kept in quadrature giving highest FM to AM conversion [88], [89]. Locking in quadrature is usually done by applying a slow feedback either to the laser or the interferometer. Feedback can result in spurs in measurements. As an interferometer, Mach–Zehnder, Michelson, or Fabry–Perot configurations have commonly been used. By locking in quadrature and measuring with a fast photodetector, the frequency noise as a function of frequency is obtained. To get the linewidth as measured with homodyne or heterodyne methods outlined above, one has to integrate noise from infinity to the value limited by observation time or fiber delay. Many papers often cite so-called white-noise or Lorentzian linewidth by multiplying the (single-sided) power spectral density white-noise value with $\pi$. This value corresponds to so-called short term linewidth and is usually smaller than the one obtained by beating the laser on a photodetector. Direct comparison is therefore somewhat more complicated, but simple approximate formulas have been derived [90].

### B. Single-Wavelength Lasers

The linewidth of single-frequency semiconductor lasers is inherently broader than, e.g., that of solid-state lasers. In a semiconductor laser there are two mechanisms broadening the linewidth: 1) the spontaneous emission which alters the phase and intensity of lasing field and 2) the linewidth enhancement factor $\alpha$ that characterizes the coupling between intensity and phase noise and is specific to semiconductor lasers due to carrier density fluctuations. There has been continuous effort to improve the linewidth or coherence, and for DFB lasers methods included: long cavities, longitudinal mode engineering via multiple phaseshifts, optimization of the active medium (e.g., introducing strain), and wavelength detuning. The heterogeneous silicon photonics platform opens up a new possibility in improving the coherence by providing a mechanism to separate the photon resonator and highly-absorbing active medium [91] as shown in Fig. 19. The quality factor of conventional III–V semiconductor lasers is limited by free carrier absorption in the heavily doped p- and n-type cladding regions, as well as in the active region where photons, both spontaneous and induced, are generated. There is an inevitable compromise resulting from carrying out both photon generation and photon storage in the same III–V material. By creating high-$Q$ photon storage in silicon and using III–V for gain, the best of both worlds can be combined. The total $Q$ of a heterogeneous resonator can then be expressed as:

$$\frac{1}{Q} = \frac{1}{Q_{\text{III–V}}} + \frac{1}{Q_{\text{Si}}}$$

where $\Gamma$ is the mode confinement factor in III–V. By tailoring the transverse geometry, the modal confinement can be engineered and total $Q$ can be optimized for best performance. Removing light from III–V appears counterintuitive because it reduces the modal gain available to the laser; however, in the limit that the $Q_{\text{III–V}}$ term dominates the total $Q$, the reduction in modal gain is exactly balanced by a reduction in total modal loss, and thus the threshold carrier density remains constant. Increasing the quality factor of the laser cavity and keeping the same modal confinement provides a double benefit to phase noise by reducing the number of excited carriers needed to reach threshold, thus decreasing spontaneous photon generation while increasing photon storage.

By designing a high-$Q$ silicon resonator with $Q = 1.1 \times 10^4$ (Fig. 20) and having a $\Gamma = 15\%$, linewidths as low as 18 kHz were demonstrated. This is approximately 200\% improvement compared to previously reported results. The linewidth is measured by the FM discriminator method, and quoted number corresponds to the high-frequency white-noise floor limit. Specifics of the silicon resonator design are given in [92].

With further optimization and introduction of a spacer layer, the mode confinement factor was reduced to as low as 1.5\% in III–V and to only 0.2\% in QW region (see Fig. 21). Spacer is layer of low refractive index material (SiO$_2$) between the III–V and silicon, and was used to directly control the rate of spontaneous emission into the lasing mode. This further increased the $Q$ factor of the laser mode and reduced the linewidth below 1 kHz, as measured by the FM discriminator method quoting the white-noise limit [93]. This result shows that, not only the heterogeneous silicon platform can perform as good as a native III–V one, but can in fact offer more than order-of-magnitude better performance.
C. Widely-Tunable Lasers

Passive microring-resonator-coupled semiconductor lasers were proposed in 2001 [94]. In such a structure, an active region in the conventional Fabry–Perot cavity is coupled with a passive ring resonator. This is different from conventional ring lasers, where the active traveling wave ring resonator replaces the standing wave Fabry–Perot cavity. The ring inside the cavity improves side mode suppression ratio, linewidth, and decreases the frequency chirp. The concept can be extended to two or more rings, significantly improving the single-mode tuning range by utilizing the Vernier effect [95].

Using rings inside the cavity benefits the linewidth in two ways: 1) increasing the photon lifetime due to effective cavity length enhancement, and 2) providing negative optical feedback by slight detuning from the ring (resonator) resonance. Both mechanisms cannot be maximized at the same time, but there is an optimal point where the combined influence is maximized. The combination of ring resonators and cavity mirror (facet mirror, loop mirror, etc.) can be thought of as a frequency-dependent passive mirror with complex amplitude reflectivity $r_{\text{eff}}(\omega)$. The linewidth improvement [60] due to feedback from this frequency dependant mirror is given by factor $F^2$ where $\Delta \nu$ and $\Delta \nu_0$ are the linewidths with and without the $r_{\text{eff}}(\omega)$ mirror (see Fig. 22).

\[
\Delta \nu = \frac{\Delta \nu_0}{F^2}
\]

\[
F = 1 + A + B
\]

\[
A = \frac{1}{\tau_{\text{in}}} \Re \left\{ i \frac{d}{d\omega} \ln r_{\text{eff}}(\omega) \right\}
\]

\[
B = \frac{\alpha_H}{\tau_{\text{in}}} \Im \left\{ i \frac{d}{d\omega} \ln r_{\text{eff}}(\omega) \right\}
\]

where $\alpha_H$ is the linewidth enhancement factor. $\tau_{\text{in}} = 2n_{\text{eff}} L_a / c$ where $n_{\text{eff}}$ is the effective index of the gain section, $L_a$ is the length of active region and $c$ is the speed of light. The $A$ term, corresponding to the linewidth reduction from reduced longitudinal mode confinement, is often denoted as the ratio of the external (passive section) cavity path length to the gain section path length. As the effective length of a ring resonator is...
maximized at resonance, the $A$ factor is maximized when the ring is placed exactly at resonance. The effective length of the ring ($L_{\text{eff}} = -d\phi/d\beta$) can be approximated at resonance and with losses ignored by:

$$L_{\text{eff}} = \frac{1 - \kappa^2}{\kappa^2} L_{\text{ring}} \tag{6}$$

For a weakly coupled rings ($k \ll 1$), the effective length will be largely extended and can even dominate the total cavity length.

The $B$ term corresponds to the reduction from the negative feedback effect where a decrease in wavelength increases reflectivity (increasing photon density in the cavity) and hence decreases carrier density, which in turn causes the wavelength to increase due to the carrier plasma effect. The phase condition in the cavity can be used for a slight detuning of the laser oscillation with respect to the minimum cavity loss condition (resonator resonance). This negative feedback effect occurs only on the long wavelength side of the resonance and is optimum at the wavelength of highest slope in the transmission spectrum. At the ring resonance, i.e., the optimal condition for the $A$ term, it is equal to zero. On the short wavelength side of the resonance, the effect is reversed and operates in positive feedback, broadening the linewidth. The combined effect of $A$ and $B$ is at maximum when the laser is slightly detuned on the long wavelength side (lower frequency), as shown in Fig. 22.

We believe that these two mechanisms—the effective cavity length enhancement and the negative optical feedback—are responsible for the exceptional linewidth results shown by ring-coupled lasers. As the loss in rings ultimately limits the performance (obtainable $Q$), the low-loss silicon waveguide platform is the key enabler of the exceptional performance shown by recent devices. We now turn to some specific widely-tunable narrow-linewidth lasers realized by ring resonators. We make a distinction between heterogeneous (single-chip, monolithically-integrated) and hybrid (assembled using butt-coupling between Si and III–V regions, multiple-chips) designs.

Heterogeneously-integrated ring-based Vernier lasers [78] featured a linewidth of 330 kHz. An improved design, coupled-ring-resonator (CRR) lasers [31], [79] have shown linewidth of 160 kHz (see Fig. 23(A)). In both cases the linewidth was measured by delayed self-heterodyne technique. The advantage of using a CRR mirror design is a higher external differential quantum efficiency by avoiding having a drop port bus waveguide to form the necessary filter shape and having a simple two output-port laser. The threshold currents are around 50 mA with output powers in $>15$ mW range.

Further progress has been made with a monolithically integrated external cavity [36]. A $\sim4$ cm long on-chip cavity is made possible by a low-loss silicon waveguide platform. Tuning in excess of 54 nm in the O-band as well as significant reduction in laser linewidth due to controlled feedback from the external cavity were shown. The linewidth measured with the delayed self-heterodyne method in full tuning range is below 100 kHz and the best results are around 50 kHz. The threshold current is in the 30-mA range, output power is $>10$ mW and side-mode suppression ratio exceeds 45 dB across the entire tuning range.

The design of this laser is shown in Fig. 23(B). The gain section (SOA1) is inside a 2 mm long cavity formed by loop-mirrors. The lasing wavelength is determined in the tuning section comprising two ring resonators and a cavity phase section, all of which are controlled by thermal phase tuners. The front loop mirror (at the output of the laser) has a 10% power reflection and the output of the laser is terminated at the facet at an angle of $7^\circ$ to minimize reflections. The back loop mirror, after the wavelength tuning section, has a power reflection of 60%, which couples part of the light to the external cavity. In order to allow for a long external cavity, a low-loss waveguide platform is needed. Optimized silicon waveguides with a loss of 0.67 dB/cm at 1310 nm were used. The external cavity is $\sim4$ cm long and has its own phase adjustment section and gain section (SOA2). As the propagation loss in Si waveguides is very low ($\sim5$ dB loss per round-trip), there is quite strong feedback from the external cavity. The feedback is present even when SOA2 is reverse-biased, possibly from reflections at the tapers to the gain region or due to insufficient attenuation of the reverse-biased SOA2.

The SOA1 bias currents, for measured linewidths, were between 75 and 128 mA ($\sim2.5-4.2 \times I_{\text{th}}$), while the external SOA2 bias currents were between 0 and 13.06 mA. The bias for external SOA2 is typically below transparency, potentially limiting the linewidth improvement performance due to increased noise in the feedback signal owing to random spontaneous emission events. Even lower linewidths (below 20 kHz) were measured at higher currents supplied to external SOA2, but the laser would become multimode with mode spacing determined by the external cavity length ($\sim1$ GHz) as the ring tuning section could not filter out a single longitudinal mode. In this case the RIN and frequency noise spectrums have peaks at this mode.
cause it allows for better performance than native III–V devices, as has been recently demonstrated with narrow-linewidth lasers and high-power, high-speed photodiodes.

Furthermore, tight integration of electronics and photonics on-chip allows for better energy efficiency in terms of photonic components and provides a way to avoid interconnect scaling limitation in modern processors.

Heterogeneous silicon photonics may become the technology of choice, not only for future longer range, high speed communications, but also for future data centers, supercomputers, and sensors.

ACKNOWLEDGMENT

The authors would like to thank B. Koch, G. Fish, E. Norberg, and A. Fang of Aurrrion; M. Paniccia, R. Jones, and M. Sysak of Intel; D. Liang, G. Kurczveil and R. Beausoleil of Hewlett Packard; X. Zhang, J. Yao, J. E. Cunningham of Oracle; A. Yariv of Caltech; D. Dai of Zhejiang University; L. Theogarajan, A. Sohdi, L. Chen, D. Blumenthal, L. Coldren, and J. Peters of UCSB for helpful discussions.

REFERENCES


Tin Komljenovic received the M.Sc. and Ph.D. degrees in electrical engineering from the Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, in 2007 and 2012, respectively. During Ph.D. degree he was a Visiting Researcher at Institute of Electronics and Telecommunications, University of Rennes. His current research interests include optical access networks, wavelength-division multiplexed systems, and tunable optical sources. He has authored or coauthored more than 30 papers and holds two patents. He is currently working as a Postdoctoral Researcher at the University of California, Santa Barbara, CA, USA, on photonic integration. He received the EuMA Young Scientist Award for the work on spherical lens antennas and the Marie Curie Fellowship.

Michael Davenport received the undergraduate degree in optical engineering from the University of Alabama, Huntsville, AL, USA, in 2007, and the M.Eng. degree in electrical engineering from the University of California, Santa Barbara, CA, USA, in 2009, where he is currently working toward the Ph.D. degree in electrical engineering. His current research interests include low-noise mode-locked lasers for applications in optical networks microwave photonics and photonic integrated circuits.

Jared Hulme received the B.S. degree in electrical engineering from Brigham Young University, Provo, UT, USA, in 2008 and the M.Eng. degree in electrical engineering from the University of California, Santa Barbara, CA, USA, in 2012. He is currently working toward the Ph.D. degree in electronic engineering with an emphasis in electronics and photonics from the University of California. His research interests include hybrid III/V-silicon ring-based tunable lasers and their applications in photonic integrated circuits such as fully integrated 2-D beam-steering chips and integrated photonic microwave generators.

Alan Y. Liu received the B.S. degree in physics and mathematics with a minor in business from Tulane University, New Orleans, LA, USA, in 2011. He is currently working toward the Ph.D. degree in electronic and photonic materials at the University of California, Santa Barbara, CA, USA. His research interests include QD optoelectronics and silicon photonics.

Christos T. Santis received the B.Sc. degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 2005 and the M.Sc. and Ph.D. degrees in electrical engineering, from the California Institute of Technology, Pasadena, CA, USA, in 2008 and 2013, respectively. He is currently a Postdoctoral Researcher with the Institute for Energy Efficiency at the University of California, Santa Barbara, CA, USA. His current research interests include low-noise performance, physical scaling, and energy consumption of heterogeneously integrated semiconductor lasers on silicon.

Alexander Spott received the B.S. degree in physics from the University of Washington, Seattle, WA, USA, in 2011. He is currently working toward the Ph.D. degree in electrical engineering at the University of California, Santa Barbara, CA, USA. His research interests include silicon photonics and long-wavelength optoelectronic devices.

Sudharsanan Srinivasan received the B.S degree with specialization in engineering physics from the Indian Institute of Technology, Madras, India, in July 2009. He is currently working toward the Ph.D. degree at the University of California, Santa Barbara, CA, USA. His research interests are in silicon photonics.

Michael Holonyak received the M.S. degree from the University of California, Berkeley, in 1955, and the Ph.D. degree from the University of Illinois, Urbana, in 1959. He is a Cofounder of Aurrion, Aerius Photonics and Calient Networks. He holds the Fred Kavli Chair in Nanotechnology. He has published 10 book chapters, 600 journal papers, 900 conference papers, and has received 54 patents. He is a Member of the National Academy of Engineering, OSA, and the American Physical Society. He received the OSA/IEEE Tyndall Award, the OSA Holonyak Prize, the IEEE LEOS William Streifer Award, and the South Coast Business and Technology Entrepreneur of the Year Award. He and coworkers received the EE Times Annual Creativity in Electronics (ACE) Award for Most Promising Technology for the heterogeneous silicon laser in 2007. His research is primarily in optoelectronics and photonic integrated circuits.

John E. Bowers received the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA. He worked with AT&T Bell Laboratories and Honeywell before joining University of California Santa Barbara (UCSB). He is currently the Director of the Institute for Energy Efficiency and a Professor in the Departments of Electrical and Computer Engineering and Materials at the UCSB. He is a Cofounder of Aurrion, Aerius Photonics and Calient Networks. He holds the Fred Kavli Chair in Nanotechnology. He has published 10 book chapters, 600 journal papers, 900 conference papers, and has received 54 patents. He is a Member of the National Academy of Engineering, OSA, and the American Physical Society. He received the OSA/IEEE Tyndall Award, the OSA Holonyak Prize, the IEEE LEOS William Streifer Award, and the South Coast Business and Technology Entrepreneur of the Year Award. He and coworkers received the EE Times Annual Creativity in Electronics (ACE) Award for Most Promising Technology for the heterogeneous silicon laser in 2007. His research is primarily in optoelectronics and photonic integrated circuits.

Eric J. Stanton received the B.S. degree in electrical engineering from Cal Poly San Luis Obispo, San Luis Obispo, CA, USA, in 2012 and the M.S. degree in electrical and computer engineering from the University of California, Santa Barbara, CA, in 2014. He is currently working toward the Ph.D. degree in electrical engineering with emphasis in silicon photonics and he works on integrated devices for spectral beam combining.

Chong Zhang received the B.S. degree in electrical science and technology from the Harbin Institute of Technology, Harbin, China, in 2007, and the M.S. degree in optical engineering from Zhejiang University, Hangzhou, China, in 2010. He is currently working toward the Ph.D. degree at the University of California, Santa Barbara, CA, USA. His research interests include silicon photonics and long-wavelength optoelectronic devices.

Young Scientist Award for the work on spherical lens antennas and the Marie Curie Fellowship.