Interband cascade laser on silicon

ALEXANDER SPOTT,1,* ERIC J. STANTON,1 ALFREDO TORRES,1 MICHAEL L. DAVENPORT,1 CHADWICK L. CANEDY,2 IGOR VURGAFTMAN,2 MIJIN KIM,3 CHUL SOO KIM,2 CHARLES D. MERRITT,2 WILLIAM W. BEWLEY,2 JERRY R. MEYER,2 AND JOHN E. BOWERS1

1Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, California 93106, USA
2Code 5613, Naval Research Laboratory, Washington, DC 20375, USA
3KeyW Corporation, 7740 Milestone Parkway, Suite 150, Hanover, Maryland 21076, USA
*Corresponding author: spott@ece.ucsb.edu

Received 11 May 2018; revised 25 June 2018; accepted 23 July 2018 (Doc. ID 331503); published 16 August 2018

Mid-infrared (MIR) silicon photonic systems show great promise for miniaturizing a variety of sensing and detection technologies. Rapid progress has been made in recent years, and numerous passive and active MIR devices have now been constructed on various silicon-based platforms. We previously reported the heterogeneous integration on silicon of Fabry–Perot and distributed feedback quantum cascade lasers (QCLs) operating at 4.8 μm. Interband cascade lasers (ICLs) will be preferred for many on-chip sensing technologies because they operate in the 3–6 μm range with threshold drive powers 1–2 orders of magnitude lower than QCLs. In this work, we demonstrate the integration of ICLs on a silicon substrate. These lasers emit 3.6 μm light into silicon-on-insulator waveguides in pulsed mode at temperatures up to 50°C. This represents an important step toward MIR photonic integrated circuits on silicon that operate with much lower drive power and therefore an even smaller footprint. © 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

OCIS codes: (140.3070) Infrared and far-infrared lasers; (130.3120) Integrated optics devices; (130.0250) Optoelectronics; (140.5960) Semiconductor lasers.

https://doi.org/10.1364/OPTICA.5.000996

1. INTRODUCTION

Investigations of the broad mid-infrared (MIR) wavelength band spanning ~2–20 μm are being driven by a diverse array of potential applications. For example, numerous chemicals in gas form exhibit strong fundamental absorption features in the MIR [1]. Therefore, a variety of laser-based techniques are being developed to exploit the unique spectral signatures that allow trace amounts of a given species to be identified and quantified. Spectroscopic chemical sensing is now used in a broad range of scenarios that include medical sensing, biochemical analysis, environmental monitoring, industrial process control, and astronomy.

Integrated optics can address these sensing applications at the chip scale to provide ultracompact devices and systems. Because silicon is transparent to wavelengths as long as λ ≈ 8 μm, the silicon-based photonic integrated circuit (PIC) provides a promising low-cost platform for future integrated MIR sensors. The high-index contrast of silicon with dielectric materials such as Si3N4 and SiO2 is favorable for small devices and high-Q resonators formed by low-loss waveguides. While the optimal silicon waveguiding platforms for each spectral band and application are still under investigation, MIR devices have already been demonstrated using silicon-on-insulator (SOI), silicon-on-nitride (SON), silicon-on-nitride-on-insulator (SONOI), silicon-on-sapphire (SOS), and germanium-on-silicon (GOS) waveguides [2,3].

The optimal on-chip chemical sensing system will fully integrate one or more light sources in the PIC. However, the most appropriate MIR emitter depends on the wavelength range chosen for the given application, since no single semiconductor laser technology provides high performance throughout the MIR. In particular, interband GaSb- and InP-based diode lasers can operate with low threshold drive powers at MIR wavelengths as long as ≈6 μm [4–6], whereas InP-based quantum cascade lasers (QCLs), which derive gain from intersubband optical transitions, are capable of providing high output powers with high wall-plug efficiencies at wavelengths extending from 3 μm to the terahertz regime [7].

The interband cascade laser (ICL) derives gain from band-to-band electron-hole recombination, as in a conventional diode laser, but with multiple cascading stages to recycle the electrons as in a QCL [8,9]. These lasers function from 2.9–7 μm (where the longest wavelength results are integrated on InAs substrates) [10], with threshold drive powers 1–2 orders of magnitude lower than for QCLs [11], while still generating up to 500 mW of continuous wave (cw) output power [12]. GaSb-based ICLs are ideal for MIR spectroscopy in 3–6 μm applications that rarely require watt-level optical power. An ICL-based system with far lower energy demand will have a much smaller spatial footprint, for ready deployment in a hand-held or remotely deployed sensor powered by batteries or solar energy.
Heterogeneous integration, in which III-V wafer material is bonded to a prepatterned silicon waveguide platform, has enabled active III-V optoelectronic elements to be integrated with silicon photonic devices for PICs operating at both near-infrared [13,14] and MIR wavelengths [15]. Previously, we have heterogeneously integrated InP-based lasers on Si for wavelengths near 2.0 μm and 4.8 μm. When type-I laser diodes with InGaAs quantum wells were integrated with SOI waveguides, more than 8 mW of cw power were emitted at λ = 2.01 μm at room temperature [15,16]. Integrated semiconductor optical amplifiers (SOAs) with the same active material also provided >13 dB of peak gain at this wavelength [17]. More recently, QCLs designed for emission at λ = 4.8 μm were integrated on the broadband SONOI waveguide platform to provide the longest wavelength source yet available on silicon. Fabry–Perot lasers emitted 31 mW into silicon waveguides and operated in pulsed mode up to 60°C [18]. An advantage of heterogeneous integration is that multiple laser systems can be integrated on the same chip to achieve broadband operation [19].

Wavelength-selective feedback elements such as gratings or loop mirrors can also be implemented with heterogeneous integration [20]. By patterning first-order gratings on SONOI waveguides, we demonstrated distributed feedback (DFB) QCLs with threshold currents as low as 80 mA and threshold current densities below 1 kA/cm². These devices emitted >200 mW from polished III-V/Si facets and operated in pulsed mode up to 100°C [15,21].

Other demonstrations of MIR lasers fabricated above Si waveguides have also been reported. Type-II InP-based Fabry–Perot [22] and DFB [23] lasers were integrated on silicon for operation at wavelengths near 2.3 μm. A transfer-printing approach was quite recently used to integrate InP-based QCLs on SOS for operation at 4.7 μm [24].

While preliminary measurements of GaSb-based lasers operating in pulsed mode near 2.38 μm have been reported [25], the III-V active elements in the vast majority of lasers integrated on silicon have been grown on InP and GaAs substrates. GaSb-based device integration introduces additional fabrication and design challenges, because the selective etch chemistries required to define and pattern active antimonide devices are much less mature. The III-V layers of an ICL on silicon must be engineered to produce a high-gain optical mode and efficient III-V tapers to couple the hybrid and silicon waveguide sections with high yield.

In this work, we demonstrate the first integration of ICLs with silicon waveguides. These lasers emit at λ = 3.65 μm and operate in pulsed mode up to 50°C. To better understand the laser performance, each device was characterized both before and after the III-V taper was removed from one end of the laser.

2. DESIGN

The waveguide design of the heterogeneously integrated ICL is similar to that we used previously to integrate Fabry–Perot QCLs on silicon [18]. Figure 1(a) schematically shows a tilted top view of the device. The III-V mesa bonded to the silicon waveguide forms a 3100-μm-long hybrid III-V/Si gain region bounded by 320- and 350-μm-long passive silicon waveguide regions. Both ends of the III-V mesa narrow in 50-μm-long tapers, which are designed to adiabatically transfer the hybrid III-V/Si optical mode to the passive silicon waveguide. The silicon waveguide is terminated on both ends by polished silicon facets, which are designed to reflect light and form a Fabry–Perot laser cavity.

The lasers were constructed on the SOI platform. While a SONOI platform was used in our previous integration of QCLs, SOI waveguides have been shown to maintain propagation losses as low as 1.5 dB/cm at wavelengths up to 3.73 μm [2,26].

Figure 1(b) shows a cross section of the hybrid III-V/Si active region. A narrow, partially etched silicon rib waveguide sits underneath the wider III-V mesa. The thin n-InAs/AlSb bottom contact layer extends laterally to both sides of the mesa, where metal is placed to form anodes. Lasers were fabricated with silicon waveguide widths of 1, 2, or 3 μm, while the III-V mesa widths ranged from 5 to 11 μm. This paper focuses on the performance of three specific lasers, with the dimensions specified in Table 1.

The ICL wafer was grown by molecular beam epitaxy (MBE) on a GaSb substrate at the Naval Research Laboratory (NRL). Table 2 summarizes the III-V layer stack. The seven-stage “W” active region, which is similar to those employed in Refs. [27,12], is designed for emission near 3.6 μm. As in other heterogeneously integrated lasers, the surrounding layers are modified for integration on silicon. An etch stop layer and sacrificial layer are included to allow removal of the GaSb substrate. Alternating GaSb and AlSb layers comprise a bonding superlattice (SL) to prevent defects at the bonding interface from propagating into the

### Table 1. Laser Geometries

<table>
<thead>
<tr>
<th>Laser</th>
<th>Si Waveguide Width (μm)</th>
<th>III-V Mesa Width (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device A</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Device B</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Device C</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 2. III-V Layers

<table>
<thead>
<tr>
<th>Layers</th>
<th>Material</th>
<th>Thickness (nm)</th>
<th>Doping (cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate</td>
<td>GaSb</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Buffer</td>
<td>n-GaSb</td>
<td>400</td>
<td>1.0 x 10¹⁷</td>
</tr>
<tr>
<td>Etch stop</td>
<td>n-InAsSb</td>
<td>1000</td>
<td>5.0 x 10¹⁸</td>
</tr>
<tr>
<td>Sacrificial</td>
<td>n-GaSb</td>
<td>100</td>
<td>1.0 x 10¹⁷</td>
</tr>
<tr>
<td>Top contact</td>
<td>n-InAsSb</td>
<td>100</td>
<td>5.0 x 10¹⁸</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlInSb SL</td>
<td>33.8</td>
<td>5.0 x 10¹⁸</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlSb SL</td>
<td>31.9</td>
<td>5.0 x 10¹⁸</td>
</tr>
<tr>
<td>Top clad</td>
<td>n-InAs/AlSb SL</td>
<td>1400</td>
<td>1.5–5.0 x 10¹⁷</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlGaSb SL</td>
<td>24.5</td>
<td>2.0 x 10¹⁷</td>
</tr>
<tr>
<td>Top SCH</td>
<td>n-GaSb</td>
<td>600</td>
<td>0.05–1.0 x 10¹⁷</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlSb SL</td>
<td>20</td>
<td>2.0 x 10¹⁸</td>
</tr>
<tr>
<td>Active</td>
<td>InAs/AlSb/GaInSb</td>
<td>290</td>
<td>Various</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlSb SL</td>
<td>20</td>
<td>2.0 x 10¹⁷</td>
</tr>
<tr>
<td>Bottom SCH</td>
<td>n-GaSb</td>
<td>600</td>
<td>0.05–1.0 x 10¹⁷</td>
</tr>
<tr>
<td>Transition</td>
<td>n-InAs/AlSb SL</td>
<td>24.5</td>
<td>2.0 x 10¹⁷</td>
</tr>
<tr>
<td>Bottom contact</td>
<td>n-InAs/AlSb SL</td>
<td>200</td>
<td>5.0 x 10¹⁷</td>
</tr>
<tr>
<td>Bonding SL</td>
<td>GaSb/AlSb SL</td>
<td>30</td>
<td>–</td>
</tr>
<tr>
<td>Capping layer</td>
<td>GaSb</td>
<td>10</td>
<td>–</td>
</tr>
</tbody>
</table>

3. FABRICATION

Figure 2(a) shows an optical microscope image of the bar containing five integrated ICLs with various geometries (following all the fabrication). The 200-µm-wide by 70-µm-tall gold rectangles in between lasers are probe pads that are contacted by needle probes or wire bonds for testing. Figure 2(b) shows scanning electron microscope (SEM) images of the tapered III-V region of a representative device in the middle (left) and at the end (right) of the fabrication.

After the silicon processing is complete, the III-V epilayer is bonded above the silicon waveguides with a hydrophilic plasma-assisted bonding process similar to that used to heterogeneously integrate InP-based devices [29]. The surfaces of the silicon and III-V chips are subjected to an O₂ plasma and manually fused together. The bonded chip is then placed in a graphite bonding fixture and annealed overnight in an oven. While the anneal for heterogeneously integrating InP-based active materials with Si is typically at 300°C for 1 h, a lower temperature of 200°C was used for the ICL integration because the thermal expansion coefficient of GaSb is nearly 3 times larger than that of silicon. The GaSb substrate was then thinned by mechanical lapping to a thickness <100 µm. The remaining substrate material was chemically etched in a CrO₃:H₂O₂ solution that selectively stopped on the n-InAsSb etch stop layer. This etch stop layer was removed with a C₆H₅O₇:H₂O₂ solution, after which the underlying GaSb sacrificial layer was removed with a tetramethylammonium hydroxide (TMAH)-based developer.

![Figure 2](https://example.com/figure2.png)

**Fig. 2.** (a) Top-down optical microscope image of five fully fabricated ICLs integrated on silicon; (b) SEM images of the III-V taper tip of a representative device. Left, before SiN encapsulation; right, after deposition of the probe metal.
since plasma etch rates often vary by as much as 20% across a die. To achieve the necessary uniformity, a selective dry or wet etch can be used to planarize the etch field. Although a variety of highly selective wet chemical etches are available for processing InP-based devices, options for the GaSb material system are more limited. The III-V mesas were defined with an SiO₂ hard mask. The upper contact, cladding, SCH, and a portion of the active region were removed with a first BCl₃ ICP etch step, using a laser endpoint detection monitor. After the etch depth was confirmed, the exposed sidewalls were passivated with a 3s H₃PO₄:H₂O₂: Tartaric acid etch, 15 nm atomic layer deposition (ALD) of Al₂O₃, and plasma-enhanced chemical vapor deposition (PECVD) of SiN. The remainder of the active region was removed with a second ICP etch step, which stopped within the 600-nm-thick bottom GaSb SCH layer. Prior to metallization, the remaining III-V material was etched with a CH₄/H₂/Ar RIE to isolate individual devices, after which the lasers were encapsulated with 1125 nm of SiN deposited by PECVD. Finally, individual laser bars were diced, and the silicon waveguide facets were mechanically polished.

4. CHARACTERIZATION RESULTS

After fabrication, the laser bar on silicon was bonded with GE varnish to a copper submount, and the probe pads were contacted with wire bonds to inject current. The drive current source produced 250 ns pulses at a 1 kHz repetition rate. Digitized scope readings were averaged from 150–200 ns to measure the detector voltage. The output power emitted from the polished silicon waveguide facet was collected by an f/1 lens and imaged to overfill a fast 2 × 2 mm² HgCdTe detector element, which was not calibrated for quantitative determination of the peak powers of <1 mW.

Figure 4 shows the light intensity (in arbitrary units) and voltage versus drive current (I-V) characteristics of Devices A and B when operated at room temperature (20°C). Note first that the I-V curves show a rather soft turn-on. Although the precise output power could not be determined, the maximum output power collected from Devices A and B in this configuration was estimated to be near 0.5 and 0.1 mW, respectively, based on measurements performed on a different setup. Device A had a threshold current of 359 mA (1.0 kA/cm²), while Device B had a slightly lower threshold current of 359 mA (1.45 kA/cm²). The threshold current for Device C was 561 mA (3.01 kA/cm²). These threshold current densities are substantially higher than typical room-temperature values of ≈150–300 A/cm² for narrow-ridge ICLs processed on native GaSb substrates [6,27], although the results for Devices A and B are comparable to those observed previously for Fabry–Perot QCLs on silicon [18]. Devices A and B both operated to above 40°C.

Figure 5 shows the spectral characteristics for Devices A (green) and B (blue) with both tapers intact, as measured by a 0.5-m-long Digikrom monochromator with a resolution of 1.9 nm. Both spectra, with centroid wavelengths 3.64 μm, display a pronounced longitudinal mode spacing of ~5 nm. This relatively large free spectral range (FSR) indicates the presence of a cavity much shorter than the 3770-μm-long Fabry–Perot cavity formed by the two polished silicon facets, and also much shorter than the 3100 μm length of the hybrid gain region bounded by the two tapers. However, it corresponds closely to the calculated FSRs of 5.3/5.8 nm for the 350/320-μm-long...
passive silicon waveguide regions on opposite ends of the device (with simulated $n_e = 3.58$). Apparently, a coupled cavity that includes one or both of these passive waveguides forms due to reflections between the III-V tapers and the polished silicon facets. In studying the first generation of QCLs integrated on silicon, we observed that even though the active gain region was surrounded on both sides by III-V tapers, the primary lasing cavity was formed by reflections from the tapers rather than the silicon end facets [18]. The conjecture that the tapers in the present devices similarly display high reflection is consistent with the formation of a cavity between the taper and the end facet as well as with other observations that will be discussed below.

The same lasers were recharacterized after one taper was removed by mechanical polishing. This reduced the length of the active gain region (including the other taper) from 3100 to 2530 μm and allowed light to be collected from the polished hybrid III-V/Si facet. A 4-mm-diameter InSb detector with 5 μs time constant collected all the light transmitted by the f/1 imaging lens. In this configuration, the total output powers were reliably calibrated by analyzing the angular profile determined from far-field characterization of each output (Fig. 8, below). The reported total power is adjusted to account for light not collected by the f/1 lens.

Figure 6(a) shows the room-temperature L-I characteristics for emission from the hybrid III-V/Si facets of all three devices. The slope efficiencies for Devices A/B/C (ridge widths 11/8/6 μm) are 13.7/10.1/6.6 mW/A, while the maximum (unsaturated) output powers are 6.6/4.6/2.0 mW. These powers and efficiencies are dramatically increased over the same devices before one taper was removed but still much lower than the typical efficiencies of >500 mW/A for seven-stage ICL ridges processed on native GaSb substrates [27]. The most striking aspect of the L-I characteristics in Fig. 6(a) is that the threshold current is nominally independent of ridge width, i.e., all three values fall within the narrow range.

Fig. 4. Output power (in arbitrary units, left axis) and voltage (right axis) versus drive current at 20°C for Devices (a) A and (b) B.

Fig. 5. Emission spectra of Devices A and B with both tapers intact at 20°C.

Fig. 6. (a) Single-side peak output power versus drive current for all three devices at 20°C, as collected from a polished hybrid III-V/Si facet after one taper was removed. (b) Voltage versus drive current density for the three devices, along with analogous data for a 32-μm-wide seven-stage ICL ridge fabricated on a native GaSb substrate (for reference).
0.31–0.36 Å. This implies that the high thresholds are dominated by current leakage at the ridge sidewalls (the same for any ridge width), rather than by the current required to invert the carrier populations in the active quantum wells (increases linearly with ridge width). Sidewall leakage will be discussed further below.

It follows that the threshold current density increases rapidly with decreasing ridge width, with samples A/B/C displaying values of 1.1/1.7/1.9 kA/cm². The threshold current densities after one taper was polished off differ from the devices with both tapers by ~7% / ~19% / ~40%, respectively. That is, the variation was minimal for Devices A and B, whereas the threshold for Device C decreased more significantly (because its more marginal operation at room temperature was more sensitive to the magnitude of the loss), implying a nonnegligible loss induced by the taper. Again, these threshold current densities are much higher than typical values for seven-stage ICLs fabricated on native GaSb substrates [6,27].

Figure 6(b) shows I-V characteristics for all three lasers, along with analogous data for a seven-stage ICL reference laser with 32-μm-ridge width and similar emission wavelength (3.45 μm) that was fabricated on a native GaSb substrate (black curve). That the integrated ICLs show a noticeably softer turn-on than the reference device is again consistent with the presence of severe current leakage at the ridge sidewalls. This leakage may be attributed to the BCl₃ ICP used to etch through the active stages of the III-V mesas. Although native-substrate ICLs processed with the alternative CH₃H₂SiCl₂ RIE usually display minimal sidewall leakage (as in the reference device) [27,30], that dry-etch chemistry was currently unavailable for the fabrication of these integrated devices at the University of California, Santa Barbara (UCSB). Previous ICL ridges fabricated at NRL on native substrates using the BCl₃ ICP etch chemistry have sometimes displayed minimal sidewall leakage [30], although on other occasions the sidewall currents have been quite substantial, as in the present devices. When NRL characterized native-substrate ICL test devices, which were processed by UCSB using the BCl₃ ICP etch chemistry prior to processing the integrated devices on silicon, the results confirmed sidewall leakage currents much lower than those indicated by the present data for integrated devices.

The slopes of the three I-V curves above threshold correspond to series resistances in the range 3.1–3.8 Ω, which are clearly higher than for the reference device. Transmission line measurements during fabrication estimated that the bottom metal contact resistivity and sheet resistance are 4 × 10⁻⁴ Ω - cm² and 340 Ω/sq, respectively. These values correspond to 1.0 Ω of series resistance added by the contacts and lateral current path in the bottom n-InAs/AlSb layer. However, the top contact resistivity of 4 × 10⁻⁶ Ω - cm² adds <0.1 Ω to the series resistance.

Figure 7 shows the emission spectra at 20°C for all three devices. The dual lobes indicate emission primarily in the higher-order TE₁₀ mode. The profile for the narrowest ridge (Device C), with its symmetric shape and deepest central null, agrees well with the simulated TE₁₀ profile. The contrasting absence of a complete central null for Devices A and B may indicate additional emission in the TE₅₀ mode (approximately Gaussian with a single lobe). The notable horizontal asymmetry of the data for Devices A and B may be due to an offset of the emission angle of the two modes with different effective indices. This could result from a slight angle in the polished facet with respect to the waveguide ridge, or to an asymmetry in the waveguide’s refractive index profile at the facet associated with residue or the generally low quality of the polished facet. A SEM examination of the III-V/Si hybrid facets after one taper was polished off revealed a high degree of roughness, and even some separation between the III-V and silicon layers. This is likely to have contributed additional loss at the polished facet, although apparently not as much as the loss at the tapers, because the lasing threshold for Device C decreased substantially when one taper was replaced by a polished facet. It will be seen below that Devices A and B also lased to a higher temperature after the taper was removed.

These data imply a mismatch of the dominant modes on opposite sides of the tapers, since the far-field data show primarily a TE₁₀ mode in the hybrid gain region, whereas the simulations indicate that only the TE₅₀ mode is supported on the silicon waveguide side of the tapers. Due to the mode mismatch, we expect the tapers to reflect most of the light reaching them (in addition to scattering and absorption losses) rather than transmitting it. This is consistent with the very low differential efficiencies observed for emission from the silicon waveguide before a taper was removed. In this context, it seems somewhat surprising that the FSRs in the spectral data show evidence for coupling between the two cavities. It may be that mode mixing in the intermediate tapered regions induces the coupling.

A further question is why the TE₁₀ mode is selected for lasing according to the far-field data, when the TE₅₀ mode should have lower sidewall scattering loss as well as greater feedback from the silicon waveguides beyond the tapers (we pointed out above that the calculated optical confinement factors are effectively equal). This may be related to the excessive sidewall leakage that causes a large fraction of the current to flow down the edge of the ridge III-V/Si facets of all three devices. The dual lobes indicate emission primarily in the higher-order TE₁₀ mode. The profile for the narrowest ridge (Device C), with its symmetric shape and deepest central null, agrees well with the simulated TE₁₀ profile. The contrasting absence of a complete central null for Devices A and B may indicate additional emission in the TE₅₀ mode (approximately Gaussian with a single lobe). The notable horizontal asymmetry of the data for Devices A and B may be due to an offset of the emission angle of the two modes with different effective indices. This could result from a slight angle in the polished facet with respect to the waveguide ridge, or to an asymmetry in the waveguide’s refractive index profile at the facet associated with residue or the generally low quality of the polished facet. A SEM examination of the III-V/Si hybrid facets after one taper was polished off revealed a high degree of roughness, and even some separation between the III-V and silicon layers. This is likely to have contributed additional loss at the polished facet, although apparently not as much as the loss at the tapers, because the lasing threshold for Device C decreased substantially when one taper was replaced by a polished facet. It will be seen below that Devices A and B also lased to a higher temperature after the taper was removed.

These data imply a mismatch of the dominant modes on opposite sides of the tapers, since the far-field data show primarily a TE₁₀ mode in the hybrid gain region, whereas the simulations indicate that only the TE₅₀ mode is supported on the silicon waveguide side of the tapers. Due to the mode mismatch, we expect the tapers to reflect most of the light reaching them (in addition to scattering and absorption losses) rather than transmitting it. This is consistent with the very low differential efficiencies observed for emission from the silicon waveguide before a taper was removed. In this context, it seems somewhat surprising that the FSRs in the spectral data show evidence for coupling between the two cavities. It may be that mode mixing in the intermediate tapered regions induces the coupling.

A further question is why the TE₁₀ mode is selected for lasing according to the far-field data, when the TE₅₀ mode should have lower sidewall scattering loss as well as greater feedback from the silicon waveguides beyond the tapers (we pointed out above that the calculated optical confinement factors are effectively equal). This may be related to the excessive sidewall leakage that causes a large fraction of the current to flow down the edge of the ridge III-V/Si facets of all three devices. The dual lobes indicate emission primarily in the higher-order TE₁₀ mode. The profile for the narrowest ridge (Device C), with its symmetric shape and deepest central null, agrees well with the simulated TE₁₀ profile. The contrasting absence of a complete central null for Devices A and B may indicate additional emission in the TE₅₀ mode (approximately Gaussian with a single lobe). The notable horizontal asymmetry of the data for Devices A and B may be due to an offset of the emission angle of the two modes with different effective indices. This could result from a slight angle in the polished facet with respect to the waveguide ridge, or to an asymmetry in the waveguide’s refractive index profile at the facet associated with residue or the generally low quality of the polished facet. A SEM examination of the III-V/Si hybrid facets after one taper was polished off revealed a high degree of roughness, and even some separation between the III-V and silicon layers. This is likely to have contributed additional loss at the polished facet, although apparently not as much as the loss at the tapers, because the lasing threshold for Device C decreased substantially when one taper was replaced by a polished facet. It will be seen below that Devices A and B also lased to a higher temperature after the taper was removed.

These data imply a mismatch of the dominant modes on opposite sides of the tapers, since the far-field data show primarily a TE₁₀ mode in the hybrid gain region, whereas the simulations indicate that only the TE₅₀ mode is supported on the silicon waveguide side of the tapers. Due to the mode mismatch, we expect the tapers to reflect most of the light reaching them (in addition to scattering and absorption losses) rather than transmitting it. This is consistent with the very low differential efficiencies observed for emission from the silicon waveguide before a taper was removed. In this context, it seems somewhat surprising that the FSRs in the spectral data show evidence for coupling between the two cavities. It may be that mode mixing in the intermediate tapered regions induces the coupling.

A further question is why the TE₁₀ mode is selected for lasing according to the far-field data, when the TE₅₀ mode should have lower sidewall scattering loss as well as greater feedback from the silicon waveguides beyond the tapers (we pointed out above that the calculated optical confinement factors are effectively equal). This may be related to the excessive sidewall leakage that causes a large fraction of the current to flow down the edge of the ridge
rather than within its interior. It is well known the surface of an InAs-rich epitaxial layer pins within the conduction band, which induces band bending and strong surface conduction. The surface conduction is quite sensitive to the etching and/or passivation procedures used to form the mesa [31]. Both the active electron wells and electron injector in each active stage of an ICL contain InAs layers [6] that are prone to surface conduction. With much more current flowing at the ridge sidewalls, gain may occur at the outer edges of the ridge before sufficient current to produce gain flows within the interior (where there may still be a net loss). This double-lobed gain profile would favor lasing in the TE$_{10}$ mode, and the effect may be greatest in the narrowest ridge, where sidewall leakage is the most prominent.

All three devices were characterized in the thermoelectric-cooler range of 10–50°C, while Device A was also measured down to −100°C in a liquid-nitrogen-cooled cryostat. Figure 9(a) shows the pulsed L-I characteristics of Device A over the entire temperature range. Lasing was observed to 50°C, which is 10°C higher than before the taper was removed. At −100°C, it produced up to 12.0 mW before any appreciable rollover. For $T = −60°C$, the decreasing output at the highest current is probably associated with an optical mode hop rather than thermal rollover, since no analogous rollover is observed at either higher or lower temperatures in the same current range.

Using the threshold current density as a function of temperature in Fig. 9(b), we extract a characteristic temperature of $T_0 =$ 43°C from an exponential fit to the higher-temperature data (0–50°C). At cryogenic temperatures, there is less improvement in the threshold with decreasing $T$. Devices B and C lased up to 50°C and 40°C, respectively, and showed similar characteristic temperatures of $T_0 =$ 43°C and 37°C, respectively. These values are only slightly lower than the typical results of ∼45–50 K for optimized ICLs fabricated on native substrates. A fit to the slope efficiency versus temperature for Device A over the range 10°C–50°C yields $T_1 =$ 71 K. These characteristic temperatures indicate that cw operation should be feasible once the threshold current density is reduced.

Figure 10 plots the threshold voltage of Device A as a function of temperature. $V_{th}$ in an ICL typically increases at lower temperatures, where the minimum voltage per stage (the photon energy) is higher and there is also less thermally assisted transport. An increase is also expected at higher temperatures, where the
increasing threshold current drops across the series resistance. However, the minimum $V_{th}$ of $\approx 4.1$ V at intermediate temperatures is somewhat higher than the more typical value of $\approx 3.6$ V for a seven-stage ICL emitting at this wavelength. The series resistance of the bottom contacts is higher, combined with the elevated threshold current density associated with sidewall leakage.

The spectral characteristics of Device A were also measured at temperatures from $-100^\circ$C to $50^\circ$C. The results shown in Fig. 11(a) indicate the usual redshift with increasing temperature, which is quantified in the plot of centroid wavelength versus temperature in Fig. 11(b). The tuning rate of 2.4 nm/$^\circ$C is consistent with the expected shift of the gain peak, as governed by the bandgap. The periodic longitudinal modes with $\text{FSR} \approx 5$ nm are observed at all temperatures, with individual modes shifting at the thermo-optic rate of $< 1$ nm/$^\circ$C. We also note an apparent systematic clumping of the modes, which becomes more obvious when the spectra for $T = 10^\circ$C–$50^\circ$C are presented over a finer grid of temperatures in Fig. 11(c). The clump spacing of $\sim 40$–50 nm is consistent with the possibility that reflections at multiple points within the tapering region form an additional short cavity of length 20–50 $\mu$m. The spectra for Sample B show a similar clumping, with period $\approx 40$ nm.

5. CONCLUSION

We have demonstrated the first ICLs integrated on a silicon substrate. The heterogeneously integrated devices with ridge widths of 6–11 $\mu$m emit at $\lambda \sim 3.65$ $\mu$m and operate up to $50^\circ$C in pulsed mode. At 20°C, the threshold current density is as low as 1.1 kA/cm$^2$, and up to 6.6 mW (unsaturated) of peak power is emitted from a hybrid III-V/Si facet. For the widest ridge at temperatures above 10°C, the characteristic temperatures for threshold and slope efficiency are $T_0 = 43^\circ$C and $T_1 = 71^\circ$C, respectively. Based on the minimum threshold power density of 5 kW/cm$^2$ and these characteristic temperatures, it appears likely that cw operation will be possible once more robust thermal dissipation is provided by flip-chip bonding.

These promising results are obtained despite issues that will be mitigated in future generations of integrated devices. Most obvious is the fundamental mismatch between the TE$_{10}$ optical mode that dominates the lasing-in in the hybrid III-V/Si gain region, and the TE$_{00}$ mode supported in the silicon waveguides that ideally form part of the cavity. Because of the mode mismatch, the tapers induce reflection and loss rather than efficiently coupling light between the hybrid and silicon waveguide sections. When one of the tapers is removed and the emission is from a polished hybrid facet, the efficiency increases by more than an order of magnitude, despite very poor facet quality. The maximum pulsed operating temperature also increases, by 10°C–20°C.

The mode mismatch may be induced in part by the presence of severe sidewall leakage currents, which are evident in both the invariance of the threshold current with ridge width and the very soft turn-on in the I-V characteristics of all three devices. The much higher current density at and near the outer edges of the ridge appears to favor lasing in the TE$_{10}$ mode rather than the fundamental (which has less intensity near the ridge sidewalls). Presumably, the BCl$_3$ ICP dry etch used to define the mesas produced an unfavorable surface quality that induced the leakage. In the next generation of integrated ICLs, this etch will be modified or replaced with CH$_4$:H$_2$:Cl$_2$ RIE. A much lower
threshold current density will naturally make cw operation much more straightforward to attain.

Once the emitted light is primarily in the TE00 mode, it may be possible to further enhance the III-V V taper coupling efficiency by engineering the optical mode to reside more in the silicon portion of the hybrid III-V/Si waveguide. This can be accomplished by reducing the separate confinement layer thickness in the ICL design. The gain in an ICL is high enough that a reduced optical confinement factor is acceptable, as long as the loss in the silicon is sufficiently low.

There appears to be no fundamental reason why ICLs integrated on silicon, with flip-chip bonding to provide cw operation, should not exhibit performance rivaling that of the optimized devices fabricated on native substrates. The first QCLs on silicon already operate at threshold current densities of ~1 kA/cm² [18,21], which is roughly at the state of the art for native-substrate devices, although the slope and wall-plug efficiencies still require improvement.

Heterogeneous integration has now been applied successfully to MIR gain materials grown on both InP and GaSb substrates, with MIR diode lasers, QCLs, and ICLs all being integrated with silicon waveguides. As these technologies mature, future MIR PICs will flexibly combine lasers emitting in multiple spectral bands, along with other optical components such as detectors and sensors. This will lead to new capabilities such as multispecies chemical sensing on the same silicon chip.

**Funding.** Air Force Research Laboratory (AFRL) (FA8650-17-C-5402); Office of Naval Research (ONR) (N00014-13-C-0147); National Science Foundation (NSF) (DGE 1144085).

**REFERENCES**


