Heterogeneous Integration on Silicon Photonics

Owen Marshall, Mark Hsu, Zhechao Wang, Bernardette Kunert, Christian Koos, Dries Van Thourhout

Abstract—To enhance the functionality of the standard silicon photonics platform and to overcome its limitations, in particular for light emission, ultrafast modulation and non-linear applications, integration with novel materials is being investigated by several groups. In this paper we will discuss, amongst others, the integration of silicon waveguides with ferroelectric materials such as PZT and BTO, with electro-optically active polymers, with 2D-materials such as graphene and with III-V semiconductors through epitaxy. We discuss both the technology and design aspects.

Index Terms—Integrated optics, silicon photonics, hybrid integration, III-V on silicon epitaxy, electro-optic modulation, laser

I. INTRODUCTION

The main driver for silicon photonics is the fact that standard processes and standard materials from the CMOS-industry can be reused for realizing complex photonic ICs with high yield and high volume. As extensively described in other articles of this issue, this has allowed the field to make tremendous progress over the last decade. However, in some cases the standard processes and standard materials of the CMOS-industry do not suffice. The most prominent example is the laser source. Despite considerable research efforts, thus far there is no viable path to realizing an efficient on-chip laser relying only on materials commonly used in the electronics industry. Also, continued improvement of “standard” modulators and switches relying on the carrier dispersion effect are increasingly hampered by fundamental trade-offs between modulation efficiency and losses (see the article by Witzens in this issue). Furthermore, it is not possible to decouple phase and amplitude modulation in these devices, while in many applications pure phase modulation or pure amplitude modulation is desired. Finally, while silicon exhibits a relatively large third order non-linearity, non-linear losses inhibit real breakthroughs in fields such as on-chip supercontinuum generation and frequency combs. Hence there is a growing need to integrate new materials into the silicon photonics platform. This includes direct bandgap III-V semiconductors, which can provide the gain needed for realizing integrated lasers and optical amplifiers. It also includes materials exhibiting a large electro-optic effect that can be used for realizing more efficient modulators, and interestingly often also exhibit a large second or third order nonlinearity, such as electro-optic polymers, ferro-electric materials, graphene and other 2D-materials.

As extensively described in the article by Bowers et al. in this issue, a first approach to integrate new materials with silicon waveguides relies on wafer bonding techniques. This approach is maturing rapidly and is extremely versatile in the sense that it can be extended to a wide range of different materials. It also preserves the high quality of the source material without introducing defects, which is particularly relevant for crystalline materials such as III-V semiconductors or lithium niobate. Nevertheless, in some cases there is the desire to use true wafer scale compatible processes such as physical vapor deposition, spin coating, chemical vapor deposition or epitaxy. These integration methods allow for a degree of integration not achievable with wafer bonding techniques, such as for the vertical and horizontal slot waveguides discussed below. In most cases however cost is the main driver, as these wafer scale processes can in principle be cheaper, especially for high volume applications. With each of these approaches, an important hurdle which must be overcome is the high development cost associated with introducing new materials and new processes into a silicon fab. This hurdle could be lower if the material is also already being investigated for new applications in the electronics industry. For example, this is the case for III-V semiconductors and for various 2D-materials, which are being considered for introduction in next generation electronic nodes.

Fig. 1 schematically illustrates how the new materials can be integrated on the silicon photonics platform. The most straightforward approach is to simply deposit the material on top of the waveguide (Fig. 1a, b). Planarizing the waveguide...
In this paper we will review recent progress in the integration of electro-active materials (blue) with silicon waveguides (orange). On top of a non-planarized (a) or planarized (b) waveguide. Embedded within a vertical (c) or horizontal (d) slot waveguide. Integrated in a hybrid silicon-plasmonic waveguide (e). Butt-coupling (f).

(Fig. 1b) by chemical mechanical polishing (CMP) is in certain cases essential for reaching high yield, e.g., for 2D-materials (see section IV). In other cases, in particular for lower index materials, this option leads to a very low confinement of the optical field in the optically active material, reducing the efficiency of the device. One option is therefore to embed the active material in the center of the waveguide core, as shown in Fig. 1c,d. In the case of a low index active material, embedded in a higher index waveguide, the boundary conditions for the surface normal electric field can further enhance the optical confinement. This is the so-called slot effect [30] and has been exploited both for EO-polymers, as discussed in section II and for barium titanate, as discussed in section III. The drawback of this embedding approach is that waveguide manufacturing is typically somewhat more involved. Even greater field enhancement can be obtained through the use of plasmonic waveguides and several different approaches have been proposed. The common disadvantage shared by these metallic waveguides are their high optical losses. The hybrid silicon-plasmonic waveguide shown in Fig. 1e exploits both plasmonic confinement and dielectric confinement, offering an interesting compromise between losses and field enhancement [1]. Finally, the active material can be deposited in a separate region. One example of this is the butt coupling scheme as shown schematically in Fig. 1f. When integrating silicon waveguides with III-V epitaxy this might be the preferred method (see section V). For a particular application, the choice of the most appropriate integration scheme may also depend upon other factors, such as the temperature budget of the processes involved and compatibility of a material with the fab.

In this paper we will review recent progress in the field of hybrid integration of novel materials with silicon photonics, using wafer scale processes. Specifically, we will discuss the integration with ferro-electric materials (section II), with electro-optical polymers (section III), with 2D-materials (section IV) and with III-V semiconductors (section V). For each material we discuss the integration approaches being investigated, current state-of-the-art and an outlook to the future maturation of the related platform.

II. INTEGRATION WITH FERRO-ELECTRIC MATERIALS

Lithium niobate (LiNbO$_3$ or LN) has long been the preferred material for building the high performance modulators used in telecom systems. They are also widely used in sensing applications and, for example, gyroscopes. They are very attractive because through the large Pockels effect in LN they can provide pure phase modulation, without spurious amplitude modulation, contrary to silicon modulators relying on the carrier dispersion effect or III-V modulators relying on the quantum confined Stark effect. The recent development of techniques for realizing thin film LN allows now for the integration of crystalline LN layers with silicon waveguides using wafer bonding approaches [2]. Using such an approach, the authors of [3] demonstrated Mach-Zhender modulators exhibiting a bandwidth beyond 100GHz, and a $V_dL = 2.2$ V cm (in a push-pull configuration). While this waferbonding based approach is very promising, thus far no practical method for integrating LN monolithically on silicon has been shown. Furthermore, the electro-optic coefficient of LN is relatively low. An interesting alternative is barium titanate (BaTiO$_3$ or BTO), which exhibits a very large electro-optic coefficient ($r > 1000$ pm/V, for bulk films). McKee et al. [4] showed that BTO can be grown on Si using a thin BaO buffer layer. In later work, it was shown that SrTiO$_3$ (STO) can be an efficient buffer layer for growing BTO on Si [5]. The STO buffer layer overcomes the chemical incompatibility between Si and STO and drives the desired 45 degree orientation between the lattices of both materials, which reduces the original mismatch in lattice constant between silicon and STO/BTO (>25%) to less than 2%. More recently, Abel et al. optimized the STO buffer layer to monolithically integrate BTO on Si for optical applications [6]. An effective electro-optic constant $r_{eff} = 148$ pm/V was extracted, establishing the potential of this approach for optical applications. However, BTO has a relatively low refractive index ($n_{BTO} = 2.38$) compared to silicon and therefore simply depositing it on top of a silicon waveguide (Fig. 1a,b) would lead to a low optical confinement. The authors of [7] therefore proposed to use a slot waveguide, fabricated by first growing a STO/BTO stack on an SOI substrate using molecular beam epitaxy, followed by an amorphous silicon layer. Only the top silicon layer was patterned, avoiding the need for etching the BTO layer. Several functional devices (grating couplers, ring resonator, MZI) were demonstrated, including modulators with $V_dL = 1.5$ V cm (with an applied bias $V_{bias} = 20$ V), equivalent with an effective Pockels coefficient $r_{eff} = 210$ pm/V. The largest modulation bandwidth obtained was 4.9 GHz, limited by electrical parasitics. Unfortunately, the waveguide losses were relatively high (~44 dB/cm). More recently, the authors of [8] used wafer bonding with a thin...
Al₂O₃ interlayer to form the desired horizontal slot waveguide structure, circumventing the need for amorphous silicon deposition. Again, a range of different passive devices with good performance were reported and tuning of a ring resonator resonance with very low power consumption (<4 µW/nm) was also demonstrated. The measured $V_Lp$ product was 1.35 V cm ($r_{eff} ≈ 300$ pm/V). However, waveguide losses were once again very high, varying from ~50 dB/cm up to ~600 dB/cm depending on the process conditions. An in-depth study of the origin of these losses was subsequently carried out and revealed hydrogenation of the STO buffer to be the primary culprit [9]. This study also showed that annealing at relatively low temperature (300°C to 400°C) can reduce these losses to 6 dB/cm, a very promising number for future applications. More recently a device working at high data rates was reported in the form of a BTO-plasmonic modulator, integrated on a silicon platform [10]. The plasmon waveguide ensures a very strong overlap between the optical and electrical field, resulting in a very compact device (10 µm). NRZ modulation at data rates up to 72 GBit/s was demonstrated. Also integration in the back end of line of a full silicon photonics platform was shown [11], with excellent performance (25 GBit/s operation, $V_L = 0.3$ V cm, loss 5.8 dB/cm). It is important to note that in order to optimize the electro-optical response of the integrated device the relative orientation of the BTO domains, the applied electric field and the polarization of the optical field is crucial [12]. In most cases the BTO layer is a-axis oriented (long tetragonal a-axis of the crystallographic unit cell parallel to the silicon surface), but it has been shown that through strain engineering c-axis orientation can also be obtained [13]. Recently, advanced in-plane piezo-force microscopy (PFM) was used to investigate in-plane domain dynamics during modulator operation [14].

Given the high cost of molecular beam epitaxy (MBE), the authors of [15] also investigated alternative deposition techniques such as pulsed laser deposition (PLD), RF-sputtering and chemical vapor deposition (CVD). However, in each case the Pockels coefficient of the resulting film was strongly degraded. An alternative approach was proposed in [16], relying on the use of a thin spin-coated lanthanide buffer layer on which the BTO layer is deposited using a sol-gel method. The transparent buffer layer can induce a preferential orientation in the BTO-layer. Later this method was extended to other ferro-electric materials such as PZT and an effective Pockels coefficient $r_{eff} ≈ 150$ pm/V was demonstrated. In [17] these layers were integrated on SiN ring resonators, demonstrating a $V_L ≈ 3.2$ V cm, low propagation losses ($ α = 1 $ dB/cm), bandwidths beyond 33 GHz and data rates of 40 Gbps.

The above results show that the monolithic integration of BTO on silicon using MBE is very promising: an effective Pockels coefficient considerably higher than that of bulk LN was demonstrated by several groups and the large waveguide losses originally observed seem now to be understood and resolvable. In addition, this approach is compatible with standard fabrication methods used in CMOS-foundries. That said, several important issues remain to be resolved. The best values for the effective Pockels coefficient were obtained under a large electrical bias. Furthermore, in some cases other effects (such as charge accumulation) that might affect practical use were observed. Finally, the effectiveness and stability of poling the domains in the BTO layer is not yet well studied. Consequently, thus far no practical device has been demonstrated, leaving considerable room for further research.

III. INTEGRATION WITH ORGANIC ELECTRO-OPTIC MATERIALS

Integration of organic materials with silicon waveguides allows one to combine the tremendous processing advantages of silicon photonics with the wealth of optical properties that can be obtained by molecular engineering of organic materials [18][19][20]. Silicon-organic hybrid (SOH) devices have been demonstrated to provide both second-order [21][22][23][24][25] and third-order [26][27][28] optical non-linearities, as well as optical gain [29]. For efficient devices, it is imperative to maximize the interaction of the optical silicon waveguide mode with the organic cladding materials, for example by exploiting discontinuous field enhancement in slot-waveguide structures [30]. The remainder of this section will focus on SOH electro-optic (EO) modulators, since these devices feature the largest application potential.

SOH EO modulators have been shown to outperform conventional all-silicon depletion-type $pn$-modulators both in terms of speed and energy efficiency [25][31][32][33]. The concept of an SOH Mach-Zehnder modulator (MZM) is sketched in Fig. 2a,b. Each arm of the modulator consists of a silicon photonic slot waveguide, comprising a pair of silicon rails. The doped rails are separated by a narrow slot with typical widths between 80 nm and 160 nm, depending on the resolution of the underlying fabrication technology. The electrical RF drive signal is applied via a ground-signal-ground (GSG) transmission line and co-propagates with the optical signal along the device. The transmission line is connected to the silicon rails of the slot waveguide by metal vias and conductive $n$-doped silicon slabs (typical thickness $h ≈ 70$ nm). The electrical RF voltage drops mainly across the narrow slot, to which the light is highly confined, thus leading to a strong overlap of the optical mode and the modulating RF field (see insets of Fig. 2b). This represents a key advantage in comparison to conventional organic electro-optic modulators [34][35]. The device is functionalized by filling and coating the slot waveguide with an EO organic cladding material. To activate the macroscopic $μ^2$-non-linearity, the microscopic molecular dipoles in the organic material need to be aligned in a dedicated one-time poling process. To this end, a poling voltage $U_{pol}$ is applied across the (floating) ground electrodes at an elevated temperature. After cooling down to ambient temperature, the poling voltage can be removed, and the molecule orientation remains in the aligned state (indicated by green arrows in Fig. 2b). The modulating field (red arrows) induced by the RF drive voltage $U_d$ is oriented parallel to the chromophore alignment in one phase modulator and antiparallel in the other phase modulator, thereby resulting in efficient chirp-free push-pull operation of the MZM. While fabrication
of SOH modulators for research experiments has relied on high-resolution electron-beam lithography [36][37][38], more recent demonstrations were based on deep-UV (DUV) lithography processes on a commercial platform, which offers the full portfolio of silicon photonic devices and SiGe detectors [39][40][41].

Over the last few years, the remarkable performance of SOH EO modulators has been demonstrated in a series of experiments, exhibiting efficiencies with voltage-length products down to $U_\pi L = 0.3$ V mm and energy consumptions of only a few fJ per bit [31][32][33]. Losses are typically in the order of 2 dB/mm, currently dominated by scattering losses in the slot waveguide. Improved lithography processes might reduce these at which point the doping in the silicon rails will become relevant as still some light is guided by the latter. The response of the electro-optic cladding materials is ultra-fast and enables small-signal modulation at frequencies beyond 100 GHz [41]. The viability of SOH devices has been demonstrated both for data transmission – using intensity modulation and direct detection (IM/DD) – and for coherent communications. Regarding IM/DD transmission, generation of 100 Gbit/s on-off keying (OOK) data signals and 120 Gbit/s PAM-4 was shown [40][42]. For coherent communications, generation of advanced modulation formats such as 16QAM was demonstrated at record-low energy consumptions and with symbol rates of up to 100 GBd transmitted on a single polarization [37][38][42][43]. The extraordinarily low drive voltage of SOH modulators allows operation of these devices directly from standard output ports of field-programmable gate arrays (FPGA), without external amplifiers or digital-to-analogue converters [41]. Such schemes can be used even if higher-order modulation formats such as 16QAM are to be generated [41]. SOH EO modulators have also proven useful for frequency comb generation [44] or as ultra-compact phase shifters for optical metrology applications [45].

Besides boosting the performance of SOH devices, research efforts have also concentrated on improved organic EO materials. Currently, in most material systems, less than 15% of the EO activity inherent in the chromophores is translated to macroscopic EO effects by poling. Improved theoretical methods [18][19][20] have led to new classes of organic EO materials and allowed systematic improvements of the poling efficiency. This resulted in macroscopic EO coefficients $r_{33}$ in excess of 500 pm/V in thin films [46][47][48][49][50] and of up to 390 pm/V in SOH devices [33] – the latter corresponding to the highest in-device EO coefficient ever achieved in an electro-optic modulator. Theoretical calculations also suggest that new chromophores with significantly improved molecular first-order hyperpolarizability are possible [46], which would enable SOH modulators with π-voltages of less than 100 mV.

Alongside the improvement of EO coefficients, the thermal and photochemical stability of EO cladding materials are also the subject of ongoing investigation and optimization. Thermal stability is defined by lattice hardness and usually quantified by the glass transition temperature $T_g$ of the material [47]. A glass transition temperature of $T_g = 150°C$ is adequate to satisfy Telcordia standards, and such temperatures may be achieved by crosslinking chemistry, for which values of $T_g = 200°C$ are routinely obtained [47]. A specific advantage of organic EO materials is that a variety of parameters such as EO activity, optical loss, dielectric permittivity, lattice hardness, material compatibility, and ease of material processing can be simultaneously optimized by systematic chemical modification of the material. SOH devices will hence immediately benefit from future advances in design and synthesis of functional organic optical materials.

**IV. INTEGRATION WITH 2D-MATERIALS**

Graphene and other 2D-materials, such as transition metal dichalcogenides, have attracted a lot of attention in the electronics industry in recent years. One reason for this interest is the fact that these 2D allotropes and compounds can possess electrical, mechanical and optical properties which differ significantly from their three dimensional relatives [51][52]. Furthermore they can be stacked in van der Waals heterostructures to extend their functionality [53].
Electrostatic gating was achieved with an electrical bias applied on graphene (SLG) with a Si waveguide, leading to a modulation depth of up to 0.1 dB/µm for quasi-TM polarized light, which has a higher interaction with the graphene layer than quasi-TE polarized light, for a detailed discussion see [62]). The same group went on to quickly demonstrate a double-layer graphene (DLG) modulator [61], where two graphene layers (again, separated by a dielectric spacer) were placed over the Si waveguide. In this case the gate bias was applied between the graphene layers, each sheet thereby gating the other, to achieve a modulation depth of over 0.16 dB/µm. The larger modulation depth per unit length in DLG modulators allows for smaller devices – crucial when electrical bandwidth is limited by the inverse of the device RC (resistance capacitance) product. It is worth noting that DLG modulators do not require a doped waveguide, simplifying one aspect of fabrication and removing another potential RC contribution, while also making them compatible with other waveguide materials such as silicon nitride [62][63].

Since their first demonstration, the performance of SLG-Si and DLG-Si modulators have continued to improve. Large modulation depths (>12 dB) have been achieved in SLG-Si microring resonators [64], albeit sacrificing the optically broadband aspect of non-resonant designs. On the other hand, SLG-Si modulators employing straight waveguides have been demonstrated with low driving bias (2.5 V), bit rates of 10 Gb/s, static power consumption <1×10^-4 mW and dynamic energy consumption as low as 35 fJ/bit [65]. Dalir et al. used DLG beneath an amorphous Si waveguide to achieve broadband (1500 nm to 1640 nm) athermal (25°C to 145°C) modulation at speeds of 35 GHz [66] (see Fig. 3), although with a relatively low modulation efficiency of ~0.1 dB/V (note that in all graphene-Si modulators a trade-off typically exists between modulation bias and speed; thinner gate dielectrics reduce the required gate bias but increase device capacitance). Many of the reported graphene-Si EAM performance figures approach or surpass those of state-of-the-art SiGe EAMs [65], although further work is still required to combine them in a single device.

To date, graphene-Si EAMs have been the most intensively studied. However, for some complex modulation formats phase modulators are necessary. These too can be achieved with graphene as its electrostatically controlled absorption can produce a large change in the effective refractive index of a SLG-Si waveguide (on the order of 10^-3) [67]. Recently this effect was utilized in SLG-Si phase modulators to achieve modulation depths and efficiencies of 35 dB and 0.28 dB/cm respectively [68].

As mentioned above, graphene is also a promising material for ultrafast integrated graphene-Si photodetectors (GPD). Its broadband absorption, extremely short photo-generated carrier lifetimes and high mobility (enabling fast carrier extraction) make it ideal for ultrafast operation. Pulsed laser-based measurement of GPD response times indicate bandwidths of over 260 GHz [69]. A number of GPD structures have already been demonstrated, relying on a variety of photodetection mechanisms such as the bolometric, photovoltaic, photothermal and photogating

Consequently, 2D-materials represent an exciting new platform which may be key to unlocking applications which are challenging (if not impossible) with conventional technologies and traditional materials [54]. While the library of available 2D-materials continues to grow [52], graphene is, at present, the most mature in terms of production, processing and integration with existing Si technology. Graphene’s appeal stems from its unique combination of properties, including high carrier mobility (room temperature \( \mu > 100\,000 \, \text{cm}^2/\text{V.s} \) at low carrier densities [55]) and the possibility of opening a band gap in this intrinsically zero-gap material [56], paving the way for high performance graphene-based FETs. However, for photonics applications, including integrated Si-based devices, it is graphene’s optical absorption which is the most fascinating. This single atomic monolayer of carbon, in the absence of significant doping, absorbs 2.3% of normally incident light from the visible to mid-IR, a value which is solely determined by fundamental constants (\( \pi \alpha, \) where \( \alpha = e^2/\hbar c \approx 1/137 \) is the fine structure constant) [58]. What’s more, by increasing the carrier concentration, for example by electrostatic gating, this absorption can be suppressed [59], making graphene highly appealing for use in both optical modulators and detectors. Graphene also lends itself naturally to Si waveguide integration, where the transferred graphene sheet lies in the plane of the waveguiding components (parallel to light propagation). If the waveguide mode has an electric field component parallel to the graphene surface then optical absorption can occur, with a magnitude which is then dependent upon the interaction length.

In 2011, Liu et al. reported the first integration of single-layer graphene (SLG) with a Si waveguide to produce a broadband (1.35 µm to 1.6 µm) electro-absorption modulator (EAM) [60]. Electrostatic gating was achieved with an electrical bias applied between the graphene and the underlying doped Si waveguide (separated by a thin dielectric spacer), leading to a modulation depth of up to 0.1 dB/µm (for quasi-TM polarized light, which has a higher interaction with the graphene layer than quasi-TE polarized light, for a detailed discussion see [62]).
effects [70][71][72][73][74][75]. It must be noted that multiple effects may occur simultaneously in devices or become dominant under varying operating conditions [76]. The underlying physics behind each of these effects is discussed in depth in a recent comprehensive GPD review [76]. In terms of device performance, the last few years have seen huge advances. Responsivity values of tens of mA/W at zero bias and up to hundreds of mA/W under applied bias have been reported [73], along with a steady increase in operating speeds, now reaching >76 GHz [75]. One common drawback of the lateral metal-graphene-metal architecture employed in many GPD structures is a high dark current under biased operation (necessary for high responsivity). An alternative architecture which avoids this issue is a metal-graphene-Si plasmonic Schottky photodetector [77]. Goykhman et al. showed that integration of graphene greatly enhance Schottky photodetector performance, giving an internal quantum efficiency of 7% and good responsivity (85 mA/W at 1 V reverse bias) with minimal dark current. Under higher bias avalanche multiplication gives even higher responsivities (~0.4 A/W at 3 V reverse bias).

The performance of graphene-Si modulators and photodetectors is expected to improve as 2D-material processing techniques continue to mature. Many of the results to date have relied upon small graphene flakes produced by mechanical exfoliation from graphite crystals, manually transferred to devices. Pristine (high µ) graphene is produced in this way, but sub-millimeter flake dimensions — and the time consuming nature of the process — limit the scalability of this approach. The most promising alternative is CVD-grown wafer scale monolayer graphene (typically grown on metal substrates and transferred to the target substrate using a polymer support [54]). Although this graphene is of lower quality than exfoliated material, recent high-speed GPD devices have been processed in a 6 inch semiconductor pilot line using large area CVD-grown graphene [75] (see Fig. 4). Single-crystal CVD-grown graphene [78], possibly in combination with graphene transfer printing techniques [79], may lead to even higher µ and therefore device performance (for example higher modulation efficiency and lower insertion loss in graphene-Si EAMs [62]). High contact resistances between graphene and metal, which limited early device operating speeds [60], can be overcome by using 1-dimensional edge contacts with CMOS compatible metals such as Ni (<100 Ω·µm [80]). Finally, encapsulation of graphene between thin layers of exfoliated hexagonal boron nitride (hBN) is known to dramatically improve graphene quality (giving flatter, higher mobility material, which is also protected from ambient moisture; a source of unintentional graphene doping). Ongoing advances in wafer scale CVD hBN growth are therefore expected to be hugely beneficial to graphene device performance. This nicely reinforces an important point when looking ahead to the future of heterogeneous integration of 2D-materials: Although the use of individual 2D-materials (such as graphene) can and will lead to novel, high performance devices, even greater accomplishments may be achieved when 2D-materials are combined to reveal their full potential.

V. III-V EPITAXY

Thus far, the use of III-V semiconductors provides the only viable path towards realizing practical laser sources for applications in integrated photonics. Wafer bonding techniques for integrating direct bandgap III-V semiconductors on silicon substrates have matured considerably over the last decade and
now provide a viable route towards large scale integration of efficient lasers on a silicon photonics platform. On the other hand, monolithic integration of III-V semiconductors on silicon using epitaxial techniques remains a very attractive approach to be pursued, as it would provide the ultimate integration density and the advantages of wafer scale integration. However, the large lattice constant mismatch between silicon and relevant III-V materials such as GaAs (4% mismatch) and InP (8%), and the formation of anti-phase domains (APD) at the polar/non-polar III-V/silicon interface form a big hurdle to achieving this goal. Nevertheless, very substantial progress has been made in the last decade, making commercial deployment of such technology in the coming years a distinct possibility.

Two main approaches can be discerned. In the first of these the III-V is grown over the full wafer, typically using a few micrometer thick metamorphic buffer layer optimized to suppress threading dislocations inside the device layer. Until now it is this technology which has been investigated most intensively and several groups have demonstrated electrically injected lasers operating at room temperature. However, at present this approach does not offer a straightforward path towards integration with standard silicon photonics devices. The second main approach focuses on the selective area growth of III-V materials on pre-patterned silicon substrates. This provides a clearer path towards integration with waveguides, but thus far only optically injected laser devices have been reported. Below we will discuss the main results obtained using these two approaches.

For realizing lasers operating at a wavelength of 1.3 µm, self-assembled InAs quantum dot layers (QD) show particular promise. III-V QD lasers have been shown to be less sensitive to non-radiative defects due to carrier localization within the individual dots. Fig. 5a shows a schematic and an SEM picture of a threading dislocation passing through a QD-based device layer (from [84]). Following the first demonstration of lasing at 1.3 µm under pulsed electrical injection [81] most efforts have focused on optimizing the buffer layer, aiming to suppress threading dislocations and APD-formation [82][83][84]. Using a Ge/Si substrate the authors of [83] demonstrated lasing with a low threshold current (16 mA), high output power (176 mW) and high continuous-wave (CW) operating temperature (119°C). The authors of [84] used a strained-layer superlattice (SLS) based on periodic In0.18Ga0.82As/GaAs layers as a defect filter to reduce the defect density to values below 10^12 cm^-2 (see Fig. 5b,c). This allowed the demonstration of a laser operating with a threshold as low as 62 A/cm^2 (12.5 A/cm^2 per QD-layer) up to temperatures of 75°C (Fig. 5d). The lasers were also operated (in CW) for up to 3100 hours.

While very impressive, all of the above results were obtained with off-cut Si substrates (4°-6° towards the [110] plane) in order to avoid the formation of antiphase domains. Unfortunately, such substrates are not compatible with standard CMOS processing. Recent research has therefore focused on exact (001) silicon substrates. In one approach, KOH is used to define v-shaped grooves along the [110] direction in the Si substrate, exposing the {111} surfaces that allow the growth of APD-free III-V layers. This allowed the demonstration optically pumped and later electrically pumped microdisk lasers [86]. In [85], a high-temperature Si substrate preparation step was used to promote the growth of an APD-free GaAs layer. However, the most successful approach to date relies on the use of a GaP/Si buffer layer [88][89]. The GaP layer is almost lattice matched to silicon and terminates APDs within ~40 nm of the interface. Using GaP/Si (commercially obtained from NASP III/V GmbH and grown by MOCVD), the authors of [89] developed an optimized GaAs-buffer layer (grown by SME) containing strained InGaAs dislocation filter layers to reduce the defect density to a few 10^6 cm^-2. A QD-based laser stack containing 4 QD layers was grown on top of this buffer (full stack shown in Fig. 6a). A record low threshold current for a III-V-on-silicon laser of 6.7 mA (Jth = 132 A/cm^2 or 36 A/cm^2 per QD layer), a wall plug efficiency of 38%, operation up to 80°C and output powers above 100 mW were obtained (see Fig. 6b). Follow-up reliability tests at 35°C for 1500 hours showed an extrapolated mean-time-to-failure of more than one million hours [90], a significant step forward, attributed to the low defect density.

Despite this enormous progress, reaching the important 1.55 µm wavelength is very challenging when relying on InAs-based QDs as the gain medium. A possible solution is to use...
Sb-based materials. It has been shown that using the correct growth conditions during MBE of III-Sb materials on silicon.

Fig. 8. GaAs nanoridge laser. (a) HAADF-STEM images. Theupper one shows zoomed-in QWs, barriers and the passivation layer while the lower high resolution (HR) TEM shows that defects are trapped in the V-shaped trench. (b) The guided TE-like mode calculated by FDE simulation. (c) Room-temperature spectra of a DFB laser under different pump powers (100 nm trench width, 170 nm grating period and 340 nm 2nd order grating coupler period). (d) L-L curve on logarithmic and linear (insert) scale of the measured DFB nanoridge laser. Black circles and solid line represent the experimental data and the rate equations fit respectively. Figure adapted from [95].

the high strain can be released through the formation of pure 90° misfit dislocations arranged in a two-dimensional network at the Si/III-Sb interface [91]. This leads to a very low threading-defect density without the need of a thick metamorphic buffer. Exploiting this effect, CW-lasing at 1.55 μm was demonstrated [92].

As previously mentioned, the second approach to overcome the large lattice mismatch between III-V semiconductors and silicon relies on selective area growth in narrow trenches patterned on silicon (001) substrates. The trenches are defined using an STI-process (shallow trench isolation) – a standard process in CMOS foundries – and then treated with a wet etch to expose the {111} faces of the silicon substrate. In initial work, bulk InP material was grown in 500 nm wide trenches [93]. The InP growth conditions (MOCVD) were optimised such that the lattice mismatch towards Si is accommodated via partial dislocations in a highly twinned region parallel to the {111} facets. This section of micro twins is only a few tens of nanometers thick at the Si/InP interface. Following definition of a DFB-grating in the top surface of the InP ridge, the silicon substrate was removed to avoid leakage of the optical field. Single mode lasing at a wavelength of ~900 nm was obtained under pulsed optical pumping. In later work a 50 nm InGaAs gain layer was grown on top of the InP material (after planarization using CMP), see Fig. 7. This allowed shifting of the lasing wavelength to 1.3 μm and a reduction of the lasing threshold to ±8 mW.

In more recent work GaAs nanoridges where grown, in this case starting from much narrower trenches (60 to 120 nm) [95]. Again, the v-shaped bottom of the trench suppresses APD formation, while the high aspect ratio of the trenches suppresses threading dislocations by trapping them at the sidewalls (see Fig. 8a). This results in high crystalline quality of the material growing out of the trenches and no threading locations are observed in the region above the substrate. For more details on the epitaxy process, see [96]. Changing the growth conditions allows one to control the ratio of the vertical and horizontal growth speed, thereby controlling the shape of the resulting nanoridge. In this way the nanoridge waveguide can be designed such that leakage of the optical mode towards the substrate is avoided and high optical confinement in the InGaAs quantum well gain region is obtained (see Fig. 8b). An InGaP capping layer was used to passivate the GaAs surface. Single mode lasing under pulsed optical pumping with a threshold of 37 mW was obtained, showing the high quality of the material (see Fig. 8c & d).

In a rather different approach, the so called template-assisted selective epitaxy (TASE) growth method, the III-V material is grown from a small silicon seed in a SiO2 cavity. In the example shown in Fig. 9a, the gain material is grown in 5 μm × 5 μm, hollow SiO2 growth templates [97]. SEM images reveal GaAs microdisks with diameters of approximately 3.1 μm and heights of 390 nm. Fig. 9b shows lasing under room temperature pulsed optical pumping of these devices.

The approaches based on selective area growth have the distinct advantage that in principle they can be integrated in a standard CMOS process flow (actually, in several cases they are derived from processes developed for future CMOS nodes). While integration with waveguides seems rather straightforward, the integration of the electrical contacts (required for electrical injection) without introducing excessive absorption losses is challenging. As such, at this point it is not clear which approach will dominate the field in the future.
VI. CONCLUSIONS

In this paper we reviewed the recent progress and status of the integration of novel materials, providing new or improved functionality, on the silicon photonics platform. Tremendous progress has been made in the field over the last few years. Quantum dot lasers, epitaxially grown on standard (001) wafers, now show extrapolated mean-time-to-failure of more than one million hours [90]. A bitrate of 72 Gbit/s was demonstrated in hybrid plasmonic-BTO modulators which were only 10 µm long [10]. SOH modulators exploiting new electro-optically active polymers now outperform standard silicon modulators relying on the carrier dispersion effect on all parameters [25]. On the other hand, graphene offers modulation and ultrafast detectors in a single platform, possibly leading to considerable cost savings. Nevertheless, there is still need for concerted research and development efforts before these new materials are offered in a standard platform. Thus far no integration of III-V lasers epitaxially grown on silicon with actual silicon waveguides has been shown. While the hybrid BTO-plasmonic modulator is proof of the materials intrinsic properties, successful demonstration of ferroelectric materials used in a more standard Mach-Zehnder modulator configuration is essential for their broader uptake. Processing of 2D-materials on 200 mm and 300 mm scales is under development but still a lot of progress is needed, especially with respect to the transfer of graphene from its growth substrate to the device wafer. And also for the silicon-organic hybrids a viable way to integrate these modulators in an industrial platform needs to be demonstrated. All of this will still require considerable resources. Fortunately, we can rely on parallel efforts carried out by the electronics industry, pursuing integration of often identical or at least very similar materials into next generation electronic chips. As a consequence, it is very likely that these efforts will be successful and that several of these materials will indeed be taken up by the photonics industry in a foreseeable timeframe.

REFERENCES


D. Jung et al., “High reliable low threshold InAs quantum dot lasers on on-axis (001) Si with 87% injection efficiency,” ACS Photonics 5, 3, 1094-1100


Owen P. Marshall received the MPhys degree in physics from the Victoria University of Manchester, UK, in 2004 and the Ph.D. degree in semiconductor physics from the Cavendish Laboratory, University of Cambridge, UK, in 2009. From 2009 to 2016, he was a Research Associate with the University of Manchester, UK, first in the School of Electrical and Electronic Engineering and later the School of Physics and Astronomy, where he worked on the development of novel terahertz quantum cascade lasers and graphene-based optoelectronic devices. Since early 2017, he has worked as a Research Associate in the Photonics Research Group, University of Gent-IMEC, Belgium. He has co-authored journal articles on a range of research topics, including liquid crystal defect dynamics, terahertz lasers, graphene plasmonics, molecular sieving and 2D-material-based optoelectronic modulators, amongst others. His current research interests primarily focus on the integration of graphene and other 2D-materials with silicon photonic integrated circuits.

Mark Hsu received the Ph.D. degree in photonics engineering from Photonics Research Group of Ghent University, Ghent, Belgium in Jul. 2017. His doctoral degree was focused on monolithic integration of functional oxides for optical modulator applications. Then, he joined Taiwan Semiconductor Manufacturing Company (TSMC) Europe B.V., Nederland. Currently, he is a principal engineer, working on silicon photonic device design and testing.

Zhechao Wang received the Bachelor degree in Optical Engineering and the Ph.D. degree in Applied Physics from Zhejiang University, China and Royal Institute of Technology, Sweden in 2005 and 2010 respectively. From March. 2011 to Dec. 2016 he was with the Photonics Research Group of Ghent University, Gent, Belgium. In March 2017 he joined the Keysight Laboratory, Keysight Technologies, Santa Clara, CA, USA, where currently he is a technical staff and a senior Photonics Scientist. His research topics mainly cover the design, fabrication and characterization of integrated photonic device, including silicon photonic devices, integrated optical comb laser, microwave photonics, monolithic/heterogeneous integration of III-V materials on silicon, etc. His current research interest is to explore applying revolutionary photonic
integration technology to optical/electrical metrology applications.

Bernardette Kunert biography not available at time of submission.

Christian Koos biography not available at time of submission.

Dries Van Thourhout received the degree in physical engineering and the Ph.D. degree from Ghent University, Ghent, Belgium in 1995 and 2000 respectively. From Oct. 2000 to Sep. 2002 he was with Lucent Technologies, Bell Laboratories, New Jersey, USA. In Oct. 2002 he joined the Department of Information Technology (INTEC), Ghent University, Belgium, where currently he is full professor. He is also associated to imec. His research focuses on the design, fabrication and characterization of integrated photonic devices. Main topics involve Silicon nanophotonic devices and the integration of novel materials (III-V, graphene, ferro-electrics, quantum dots, ...) on these waveguides to expand their functionality. He is working on applications for telecom, diatom, optical interconnect and sensing.