On-chip infrared photonics with Si-Ge-heterostructures: What is next?

Cite as: APL Photonics 7, 050901 (2022); https://doi.org/10.1063/5.0078608
Submitted: 14 November 2021 • Accepted: 04 April 2022 • Accepted Manuscript Online: 05 April 2022 • Published Online: 31 May 2022

I. A. Fischer, M. Brehm, M. De Seta, et al.

COLLECTIONS

This paper was selected as Featured

ARTICLES YOU MAY BE INTERESTED IN

Silicon nitride waveguide polarization rotator and polarization beam splitter for chip-scale atomic systems
APL Photonics 7, 046101 (2022); https://doi.org/10.1063/5.0077738

High-bandwidth Si/In2O3 hybrid plasmonic waveguide modulator
APL Photonics 7, 051301 (2022); https://doi.org/10.1063/5.0087540

High-speed, zero-biased silicon-germanium photodetector
APL Photonics 6, 041302 (2021); https://doi.org/10.1063/5.0047037
On-chip infrared photonics with Si-Ge-heterostructures: What is next?

I. A. Fischer, 1,a) M. Brehm, 2 M. De Seta, 3 G. Isella, 4 D. J. Paul, 5 M. Virgilio, 6 and G. Capellini 3,7

AFFILIATIONS
1 Experimental Physics and Functional Materials, Brandenburgische Technische Universität Cottbus-Senftenberg, 03046 Cottbus, Germany
2 Institute of Semiconductor and Solid State Physics, Johannes Kepler University Linz, Altenberger Strasse 69, 4040 Linz, Austria
3 Dipartimento di Scienze, Università Roma Tre, V.le G. Marconi 446, I-00146 Rome, Italy
4 LNESS Dipartimento di Fisica, Politecnico di Milano, 22100 Como, Italy
5 James Watt School of Engineering, University of Glasgow, Rankine Building, Oakfield Avenue, Glasgow G12 8LT, United Kingdom
6 Dipartimento di Fisica, Università di Pisa, L.go Pontecorvo 3, I-56127 Pisa, Italy
7 IHP–Leibniz-Institut für innovative Mikroelektronik, Im Technologiepark 25, 15236 Frankfurt (Oder), Germany

a) Author to whom correspondence should be addressed: inga.fischer@b-tu.de

ABSTRACT
The integration of Ge on Si for photonics applications has reached a high level of maturity: Ge photodetectors are available on the Si platform in foundry processes, and Si/Ge heterostructure multiple quantum-well photodiodes are rapidly progressing toward applications in light modulation. These successes result from decades of development of high-quality material growth and integration, which, more recently, has sparked an increasingly broad field of photonic device research based on Si/Ge heterostructures that extends from quantum cascade lasers to sensors. Here, we highlight selected recent structure and device developments as well as possible future trends that are enabled by the maturity of the SiGe material platform.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0078608

I. INTRODUCTION

Since its dawn in the early 2000s, the silicon photonics (SP) platform has been developed to a level of technological maturity at the present time comparable to other Si-based technologies, such as nano- and microelectromechanical systems (NEMS/MEMS). As a result, companies and industrial players of very different sizes, from fab-less start-ups to data/telecom giants, can exploit a highly differentiated offer of SP foundry services, ranging from prototyping lines to pilot lines, all the way to industrial fabs.1

The building blocks of the SP toolbox, such as waveguides (WGs), modulators, photodetectors (PDs), grating couplers, and so on, can be now routinely manufactured in multi-project wafer (MPW) runs relying on process design kits (PDKs) using procedures and methods borrowed from the “standard” microelectronics environment.2

This ecosystem has enabled the rapid and sustained diffusion of the SP technology, which initially clustered around SP transceivers (which became mainstream in data centers in the second half of the 2010s) with the on-field replacement of copper connections with active optical cables (AOCs) reaching 100G and beyond.1 SP transceivers for Datacom and long-haul communications are still dominating the SP market with a sales value of almost $1 \times 10^9$ USD in 2020 and a fourfold forecasted increase in the next five years.1 SP-based solutions, however, are finding their way in a plethora of other applications, such as environmental sensing, point-of-care medical tests, automotive light detection and ranging (LIDAR), and fiber-optic gyroscopes, further boosting the SP market size.4

This success on the market side has fostered the efforts of the scientific community toward new SP materials, devices concepts, and technological solutions. While SP as a research field builds on groundwork laid already in the 1980s, the term “silicon photonics”
began to diffuse in scientific literature in 2004–2005 (Fig. 1), in coincidence with the introductory book by Reed and Knights, and it became more and more popular, with an estimate of an average of about 1800 manuscripts published in the last 4 years on the subject. In the same plot, we can observe the appearance of manuscripts mentioning “germanium” in addition to SP, at about the same time and featuring a steady increase as well (albeit on a ×0.5 scale).

Germanium is the “next in line” of the group-IV elemental semiconductors and, for its desirable electronic characteristics, was at the base of the early age of solid-state electronics. The addition of germanium to the SP platform has been the key enabler for the manufacturing of active devices. Indeed, thanks to its almost ideal direct bandgap energy, Ge has been the material of choice for the realization of those Si-integrated waveguide-photodetectors, which have allowed the realization of the above-mentioned SP transceivers operating at near-infrared (NIR) wavelengths comprised in the L- and C-bands of telecom fiber optics. Thus, without heteroepitaxial germanium layers monolithically integrated on the Si-CMOS platform, the “SP revolution” would not have been possible, at least as we have seen it. After the first successful integration and commercialization of “Ge-inside” AOC from Luxtera in 2007, the Ge integration in Si has steadily evolved, and now Ge is recognized as a “standard” CMOS material, with institutes and foundries (such as HHP, CEA-Leti, IMEC, and AMF, to name a few) offering ePIC (electronic-photonic integrated circuits), MPWs, and PDKs.

Furthermore, Ge and SiGe alloys are living a “Renaissance” besides their use in SP, with, e.g., quantum computing, spintronic, thermoelectric, and microelectronic research communities turning their attention to this “new old” material.

This interest from a broad variety of realms, and for both “More than Moore” and “More Moore” applications, has contributed to the development of complex Ge/Si heterostructures. In particular, recent efforts on developing advanced epitaxial deposition processes have led to materials featuring enhanced structural, optical, and electrical quality. Moreover, the scientific community has deepened its understanding of the Ge(Si)/Si material systems, especially in the field of doping, defects, and lattice strain management.

This paper aims to give an overview of the future perspectives of selected SP integrated devices based on Ge/SiGe/Si heterostructures of different complexity. Clearly, our scope here is not to give a complete account of possible applications using Ge in the SP domain but rather to concisely report on some of the device concepts that we believe can have a potential impact in the short-medium term. In particular, we have individuated three main topical areas that are the subject of the three sections in which this article has been divided:

i. novel approaches on NIR light-emission and detection;
ii. sources, detectors, and non-linear devices for the mid-infrared (MIR)/far infrared (FIR) range;
iii. new concepts for on-chip wave-guiding and sensing.

II. NOVEL APPROACHES ON NIR LIGHT-EMISSION AND DETECTION

The achievement of an efficient and electrically pumped on-chip laser source that can operate above room temperature (RT) and that one can conveniently couple to other photonic components represents the Holy Grail of silicon photonics. Consequently, considerable efforts are being made to integrate laser sources on Si for on-chip optical interconnection, summarized in Refs. 18 and 19. Different approaches have been pursued to meet this ambitious goal, leveraging on GeSn, III-V quantum dots (QDs) on Si, or tensile strained Ge microbridges. In this latter case, in particular, spectacular optical properties have been reported in the recent years. However, at present, the boost in optical emission triggered by tensile strain is primarily used in conjunction with Sn alloying to further reduce the energy difference between the direct and indirect fundamental energy gap. All the above-mentioned approaches have already evolved to the extent that it is well worth a perspective paper of its own. Here instead, we prefer to focus only on the SiGe/Ge material system discussing unorthodox and future approaches both for SiGe-based light emission and detection that have recently emerged. As a matter of fact, also Ge/Si-based photodetectors have evolved dynamically from their dawn in the late 1990s to the high-performance devices now available, which, as mentioned in the Introduction, can be considered as “off-the-shelf” in modern SP technology.

A. SiGe nanostructures for NIR light emission

Owing to their detrimental impact on device performances, one typically aims at eliminating lattice defects in optoelectronic materials. On the contrary, we first discuss recent efforts to turn the crystal defects into opportunities.

It is well established that selected optical properties of semiconductors, such as single-photon emission, can be modified by intentionally created, isolated defects. Prominent examples include color centers in diamond or SiC and ZnO. Only recently, single-photon emission was demonstrated in commercial SOI wafers implanted with carbon atoms. Despite their characteristic radiative lifetimes of tens of nanoseconds, they show promising properties.

---

**FIG. 1.** This figure captures the time evolution of the SP “revolution” by displaying the number of scientific papers published per year featuring the words “silicon photonics” (red, left axis) and “silicon photonics germanium” (blue, right axis) in the title/abstract/keyword. Search on the Scopus database.
related to high Debye-Waller factors and high estimated intrinsic quantum efficiencies.\textsuperscript{32,33} Thus, these and other defect structures in isotopically enriched Si\textsubscript{28} can enable the realization of integrated, telecom quantum information networks based on spin–photon interfaces.\textsuperscript{34}

In the long term, a clear obstacle of these defects and their fabrication method through ion implantation in bulk Si or silicon-on-insulator substrates is the lack of deterministic lateral and vertical position control. SiGe(C) heterostructures and site-controlled top-down and bottom-up fabricated quantum dot and wire structures may help to improve the capability of at least lateral but maybe also vertical positioning of the radiative centers. Moreover, the spatial properties of the optical modes can be controlled by suitably designed SiGe multilayer stacks. Certainly, the integration of Ge-containing heterostructures will require combined experimental and theoretical efforts to shed light on the influence of different alloyed matrix materials, quantum confinement, interface effects, and strain effects on these quantum emitters.

A road toward enhanced optical properties by “defect-engineering” that already contains Ge and Ge/Si heterostructures is the ultra-low-energy implantation of Ge ions in epitaxial (Si)Ge quantum dots (defect-enhanced quantum dots—DEQDs) to increase their emission intensity.\textsuperscript{35} Thanks to low formation energies,\textsuperscript{36} the ion implantation triggers the formation of split-[110] self-interstitial structures (SSI)\textsuperscript{37} that are responsible for the enhanced photo- and electro-luminescence (PL, EL) emission intensity and for its robustness against thermal quenching, even above room temperature.\textsuperscript{38,40–42} This is in stark contrast to conventional epitaxial Ge/Si QDs, where PL/EL is usually observed at cryogenic temperatures only.\textsuperscript{43,44} Evidence for optically pumped lasing in devices based on Ge DEQDs has been recently reported.\textsuperscript{45}

These initial promising results deserve broader and concerted theoretical and experimental efforts to fully exploit the potential of this new class of defect-enhanced SiGe micro- and nanostructures for application in the realms of optoelectronic light emitters and detectors.

Indeed, in contrast to the more straightforward system of single defects in pure Si bulk material discussed above, there is not yet definitive agreement on the physical mechanism behind the defect-enhancement of the optical emission. Previous studies of the SSI’s electronic properties in bulk Si and Ge were restricted to the Γ point and indicated a decrease in the conduction band minimum energy by a few tens of meV (see Fig. 2).\textsuperscript{36,45,46} Therefore, it was suggested\textsuperscript{46} that electrons in the matrix material can tunnel into the QD where they recombine via SSI-induced states with holes, confined by the quantum dot-induced type II band alignment (Fig. 2). However, a more recent work\textsuperscript{47} expanded the electronic state calculations to a broader region of the BZ, revealing a more complex influence of SSI on the optical properties. In particular, in unstrained bulk Ge containing SSI, other dark and bright direct transitions, e.g., in X-direction of the BZ, are expected to contribute to light emission.\textsuperscript{47}

Considering the myriad of point defect types known in Si and Ge and the newly found significant impact on the optical properties,\textsuperscript{32,33} we believe that also optically active defects embedded in (Si)Ge/Si nanostructures demand more attention regarding exciting perspectives as Si-compatible (quantum)light-emitting devices.

Any deviation from the “ideal host lattice” conditions, common in epitaxial strained thin films,\textsuperscript{35–38} where, e.g., strain and Si–Ge intermixing take place, could, in principle, be engineered to boost the optical emission. As an example, SSI-based optically direct transitions along the X-direction surrounded by a strained or alloyed matrix could have a beneficial impact on momentum conservation when charge carriers are injected from the Si substrate.\textsuperscript{37}

In addition, on the experimental side, the field is in its infancy. In SiGe QDs, as used in Refs. 35–38 and 40–42, strain and alloying are always closely linked, while to date, no detailed examination of the separated effects of strain and alloying on the optical properties of such defect-containing QDs is available. We note that for Ge concentrations less than ∼85 at. %, the SiGe band structure changes into a Si-like one, thus making optically direct Γ-point transitions virtually impossible. Here, the addition of defects might be a game-changer if it turns out that the predicted optically direct transitions in other k-directions become feasible.\textsuperscript{37}

To better understand the role of alloying, experimental investigations of SSI in unstrained SiGe nanostructures should be carried out. Fully relaxed Ge and SiGe nanodots on nanopillars,\textsuperscript{48} co-implanted or post-growth ion-implanted, could represent a suitable testbed to address these fundamental questions.

On the other hand, the post-growth strain engineering of DEQDs using Si\textsubscript{1-x} stressor layers or piezo-actuated nanomembranes containing DEQDs might be a valuable tool to isolate the influence of strain on the SSI,\textsuperscript{49,50} as it was previously done for conventional SiGe nanostructures.\textsuperscript{51} Furthermore, for addressing the technologically relevant MIR spectral range, the incorporation of Sn into DEQDs represents an intriguing perspective.\textsuperscript{47}
For future technological development, one should also address issues related to the DEQD’s fabrication schemes. To date, published DEQDs are grown by molecular beam epitaxy (MBE), exploiting Ge ions in the chamber that are accelerating toward the substrate by a bias of ∼1–2 keV. Recently, DEQD-like emission at room-temperature was reported for GeSi nanoclusters created by Ge ion implantation on Si substrates. Although the resilience of the optical properties of these structures against thermal quenching is at the moment less pronounced than for DEQDs, this research presents an essential step toward industry compatibility.

To make use of the strong hole confinement in the QDs, the main hurdle concerning implanted GeSi nanoclusters is the small height requested to optimize their performances. Targeting a depth of a few nm with a low ion dose can be a real challenge when “standard” ion-implanter and implanting conditions are employed. Implantation through a dielectric absorber layer, which can subsequently be removed, might be a solution, leaving the surrounding of the dots unaffected. Other open issues are curing harmful implantation-caused defects in the matrix surrounding the DEQDs and the high-temperature annealing steps required to activate the implanted species. Nevertheless, DEQDs seem robust against thermal treatment, such as hour-long annealing at ∼600 °C or millisecond flash lamp annealing at ∼800 °C.

Finally, as for all Si-technology-compatible light emitters, strategies for low-loss electrical injection of charge carriers and efficient electrically pumped lasing at room temperature and above remain to be demonstrated in order to transform these first achievements into a disruptive technology. As an alternative way to enhance light-emission from SiGe, using Ge-rich SiGe alloys in hexagonal crystal phase has been pioneered. The formation of hexagonal Si (hex-Si), hex-Ge, and hex-SiGe alloys without detrimental crystal defects has been achieved by exploiting a conformal overgrowth of hex-group III–V core nanowires made, e.g., in GaP or GaAs. These nanowires, grown on Si(111) substrates by vapor–liquid–solid epitaxy using Au catalyst particles, serve as a seed crystal for SiGe layers that mimic the foreign Lonsdaleite crystal phase.

In Ref. 54, direct bandgap emission for hex-Si₃-xGeₓ alloys with Ge content x > 0.35 was demonstrated, recognized as a Physics World Breakthrough of 2020. The direct band edge can be tuned in the 0.35–0.68 eV range (λ = 3.5–1.85 μm), upon tuning the SiGe composition. Moreover, in hex-SiₓGe₁₋ₓ nanowires, the radiative lifetime was estimated to be ∼1 ns, similarly to core–shell, direct gap GaAs/AlGaAs nanowires.

Several hurdles are on the path of hex-group-IV before they can shine into real-world devices. The high aspect ratio of the nanowires, with lengths of several micrometers and sub-micrometer diameters, can naturally allow laser operation under optical pumping using the nanowire itself as an optical resonator. Although direct bandgap transitions were reported for hex-SiGe, optically pumped lasing was not demonstrated to date. It is hard to discriminate which physical process has prevented the achievement of laser action so far. Optical gain may be limited by the presence of surface imperfections, surface states, and, possibly, deep-trap states induced by remnants of the Au catalyst.

More importantly, Ref. 54 estimated an n-type background doping value in the SiGe shell of ∼10^{19} cm⁻³ due to As diffusion from the GaAs wire-core. Further investigations need to confirm that the excess carrier lifetime is dominated by the radiative channel as claimed in Ref. 54 to better understand the present gain limitations.

If single-wire-resonators are not sufficient for lasing to occur, it will be necessary to investigate other resonator geometries, e.g., photonic band edge cavities. Here, one has to accurately control the spatial distribution of the growth sites of a wire ensemble and their lateral dimension to form a photonic crystal structure featuring missing, smaller, or larger wires to enhance the electromagnetic field locally. For such an arrangement, the influence of the Au nanoparticle and material unintendedly grown between the wires on the optical properties will have to be examined theoretically and experimentally.

As an intermediate step, gain measurements in the spontaneous emission regime using wires of different lengths could provide valuable insights. Alternatively, post-growth length manipulation techniques, e.g., focused ion beam cutting, can be employed. Finally, we note that, besides the negative impact on the gain, the high As doping density represents a severe bottleneck in the path toward electrically pumped devices, where the achievement of efficient p-i-n diodes is required.

Another issue concerns the need for Si(111) substrates and the integrability of hex-SiGe with CMOS processes due to the now-used Au seeds. While Au-free wire-growth of InAs nanowires was reported, GaAs and GaP nanowires with the highest size homogeneity require Au catalysts.

Finally, a possible hex-SiGe nanowire laser needs to be coupled to other photonic components, e.g., detectors and waveguides. We note that demonstrators of group III–V nanowire lasers coupled to waveguides already exist, both defined on untypical SOI(111). To this aim, pick-and-place techniques might be used, which, however, are hard to combine with large-scale integration, and thus, the advantages of this cumbersome approach become questionable. Indeed, if one already needs a group III–V nanowire to fabricate hex-Si/SiGe/Ge, the advantage of using the “optically challenging” group-IV material instead of the direct bandgap group III–V material with a suitable emission wavelength seems not obvious.

For the above reasons, large-scale practical applications based on hex-SiGe nanowires remain far from being established. Nevertheless, the understanding of physical properties of Si, Ge, SiGe, and (Si)GeSn material in the Lonsdaleite crystal structure will undoubtedly benefit from the recent breakthroughs in hex-group-IV nanowires, which may also be heterostructured to obtain hex-group-IV quantum wells (QW) or quantum dots.

For targeting real-world applications, we believe that developing an efficient system to obtain a large region of planar hex-SiGe represents a game-changer. However, to date, only nano-sized hex-SiGe crystallites were reported using alternative technologies, which leverage on strain engineering or nanoindentation, and it is not clear if the size of the hex-nanocrystallites can be increased relying on these approaches.

To overcome this limitation, a possible alternative route could be the growth on (111)-substrates using templated self-assembly, where nanostructures deposited by CVD can be confined and stressed by predefined SiO₂ regions. Possibly, the hex-Si(111) core could then be expanded using the zipper-induced epitaxial lateral overgrowth, as demonstrated for wurtzite InP material.
B. SiGe nanostructures for NIR photodetection

As stated in the Introduction, photodetectors (PDs) have been the first Ge-based devices to be demonstrated within the SP technology and now feature bandwidth, sensitivities, and responsivities that put them at the forefront of the technology. Here, we report on other “degrees of freedom” in photodetector design utilizing nanostructuring.

Indeed, the polarization-dependent absorption spectral window of Ge/Si devices can be tailored by growth, i.e., by a proper selection of the composition and strain profile of the active layer and by making use of lateral nanostructuring. Both metallic and dielectric nanostructures have been widely investigated, aiming at engineering the light–matter interactions for applications such as imaging or sensing. Light–matter interaction can be modified both by choosing a particular geometry/material of the nanostructure and/or by arranging the nanostructures in arrays. As such, photonic crystals, plasmonic nanostructures, and optical metasurfaces are some of the research areas that have seen tremendous scientific activity over the last few decades. However, the incorporation of such structures as building blocks into active devices has a lot of potential for device applications that has not been fully realized to date. Here, we first highlight some existing approaches to integrate metallic or dielectric nanostructures with group-IV devices, starting with nanowires and then proceeding to subwavelength particles as 0D nanostructures (see Fig. 3 for an overview).

Dielectric nanowires support leaky modes as solutions to Maxwell’s equations, allowing them to efficiently trap light at certain excitation wavelengths. This can be used to achieve strong absorption in semiconductor nanostructures in wavelength ranges determined by their geometry. As such, by nano-patterning the Ge layer in photodetectors, one can obtain absorption in a spectral window vastly different from those of the unstructured material, as already observed in Ge-wire-based PDs. These nanoscale PDs offer other advantages. On the one hand, their reduced device capacitances (<1 fF) potentially enable very high operating speed, in excess of 40 GHz. On the other hand, applications such as optical interconnects can potentially benefit from a reduction in the size gap between μm-sized optoelectronic devices and nanoscale transistors.

Nonetheless, the reduced semiconductor volume in those nanoscale photodetectors can also lead to a reduction in absorption and low signal intensities: A responsivity of ~0.01 A/W at an illumination wavelength of 632 nm was, e.g., reported for a photodetector consisting of a single Ge nanowire with an absorbing volume of ~0.06 μm³. The absorbing volume can be increased by using, e.g., a large number of nanowires. If those are placed in regular, densely spaced arrays, optical properties of the arrays are shaped by collective excitations, such as Bloch photonic crystal modes.

In a different approach, metallic nanowires can guide surface plasmon-polariton (SPP) modes, i.e., electromagnetic waves coupled to collective oscillations of the electron gas within the metal at the interface with the surrounding dielectric. Compared to the leaky modes in dielectric nanowires, those SPP modes are strongly localized, i.e., at optical frequencies, the electromagnetic energy is confined within nanoscale regions. Combining metallic nanowires with Ge PDs has several interesting applications. As an example, Ge nanowire PDs have been exploited for SPP detection. Conversely, the excitation of SPPs at the interface between a metal grating and Ge has been utilized to extend the absorption spectrum of Ge photodetectors to energies below the bandgap of the semiconductor.

Moving to a further reduced dimensionality, we now discuss metallic and dielectric subwavelength particles, which can support resonances when excited by light. The collective oscillations of the electron gas within metallic nanostructures at specific excitation wavelengths, so-called local plasmon resonances, can be used to control and manipulate light in the visible and infrared spectrum. Such metallic nanoantennas have been utilized in combination with large-scale optoelectronic devices based on group-IV semiconductors to, e.g., selectively enhance light absorption in Si and Ge PDs. Regular arrangements of such plasmonic nanoparticles can provide wavelength-selective absorption enhancement for optoelectronic group-IV devices. However, it can also be used for polarization characterization in combination with Si photodetectors or sensing, as we will discuss in Sec. IV. For nanoscale photodetectors, in particular, plasmonic nanoparticles can be used to enhance absorption in Ge PDs in combination with a dipole antenna or subwavelength apertures.

Optical metasurfaces composed of arrays of dielectric or metallic subwavelength particles, whose resonances can be engineered to control amplitude, phase, and polarization of light interacting with the material, have enjoyed increasing attention in recent years, with potential applications ranging from lenses, holograms, and color and polarization filters to higher harmonic generation and optical frequency mixers. However, few attempts have been made to incorporate dielectric metasurfaces directly into devices so far. Nanostructuring has been used for, e.g., light-trapping and broadband efficiency enhancement of large-scale Si and Ge PDs. Going a step further by nanostructuring the entire active layer, the absorption in Ge metasurfaces on CaF₂ substrates was investigated experimentally and a configuration incorporating a contacting scheme for turning Ge metasurfaces into photodetectors was proposed.

The use of the metallic and dielectric nanostructures discussed above has a lot of untapped potential for future optoelectronic applications, e.g., for chip-sized spectrometers, if the concomitant...
challenges of device integration can be addressed, as we will discuss in the following.

First, we note that while the group-IV semiconductors Si, as well as Ge, are well suited for the fabrication of dielectric metasurfaces as a result of their comparatively high refractive indices, research on metallic nanoparticles has a strong focus on non-CMOS-compatible materials, such as Ag or Au. By comparison, CMOS-compatible metals, such as Al or TiN, have been far less investigated but have to be used for CMOS-compatible device fabrication.

Embedding optical metasurfaces into devices while preserving their properties poses additional challenges. While metallic nanoparticles can easily be combined with bulk semiconductor devices, shaping the semiconductor material itself into dielectric nanoparticles requires elaborate contacting schemes for the devices. The first major hurdle resides in the contact modules. Those that are usually used in CMOS foundries are opaque to all the IR spectrum. This could be solved by employing transparent conductive oxides, but unfortunately, these materials bring about relatively high contact resistances, which is particularly limiting for the performance of nanoscale devices. A different solution is the use of graphene layers as transparent contact materials. However, their integration into standard CMOS processes is still to be achieved. An alternate solution consists of making use of plasmonic nanoapertures that increase the PD efficiency for selected wavelengths and provide a contact resistance, which is particularly limiting for the performance of nanoscale devices. A different solution is the use of graphene layers as transparent contact materials. However, their integration into standard CMOS processes is still to be achieved. An alternate solution consists of making use of plasmonic nanoapertures that increase the PD efficiency for selected wavelengths and provide an efficient contacting scheme with low-resistance contacts made of CMOS-compatible metals (Fig. 4).

Since the influence of geometry and spatial arrangement on device absorption spectra is large, another challenging issue is the achievement of a large-scale precise nanofabrication approach that not only allows for well-defined optical properties but also ensures the repeatability of the results. This is of relevance not only for optical metasurfaces but also for the combination of both metallic and Ge nanostructures for the enhancement of nanoscale devices. In the strongly coupled regime, plasmonic and photonic resonances can hybridize with potential applications, e.g., for sensing, but only if the interplay of photonic and plasmonic modes can be tuned, and for this purpose, precise control of the geometry of the structures is particularly important.

![FIG. 4. Scanning electron microscope image (a) and schematic cross section (b) of a Ge nanopillar photodetector (diameter ~500 nm) with self-aligned nanoapertures in the contact metallization to enable light transmission to the photodetector and to provide plasmonic enhancement.](image)

C. Single-photon avalanche diodes

Single-photon avalanche diode (SPAD) detectors fabricated from a range of semiconductor materials that operate in the visible and short-wave infrared (SWIR) have become the standard optical detection approach for a range of existing and emerging applications. Superconducting single-photon detectors provide the highest performance of any detector in the visible and SWIR range, but the need for cryogenic operating temperatures (typically <3 K) precludes their use for many applications. As an alternative, CMOS SPAD imaging arrays have been commercially available since the early 2010s and can now be found in many products for range finding, including mobile phones, tablets, autonomous vacuum cleaners, and light detection and ranging (LiDAR) systems. CMOS SPADs are also used for biological and medical imaging applications, such as fluorescence lifetime imaging, optical tomography, positron emission tomography (PET), and Raman spectroscopy.

The single-photon detection efficiency (SPDE) at shorter wavelengths ($\lambda \sim 650$ nm) can be particularly high (up to ~70%), but the bandgap of silicon limits the operation wavelength to $\lambda < 1$ $\mu$m, a significant limitation for applications. The use of InGaAs SPADs single pixel and imaging arrays has allowed longer-$\lambda$ applications to be pursued, including quantum communications, quantum optics, quantum computing, and LiDAR at the telecommunication O–C bands. Longer wavelength operation is advantageous for LiDAR for multiple reasons: (i) the solar background reduces at longer $\lambda$; (ii) atmospheric and obscurant scattering (rain, mist, fog, and dust) reduces, as both Rayleigh and Mie scattering decrease for longer $\lambda$; and (iii) the optical safety threshold for the human eye (IEC-60825-1) increases the usable laser power for $\lambda > 1400$ nm.

The use of Ge and SiGe as an absorber of photons has the potential to extend the detection efficiency of silicon-based SPADs to longer wavelengths. Moreover, SiGe/Si SPAD devices can represent a much cheaper alternative to InGaAs SPADs, with the potential to scale to larger pixel arrays to target applications requiring cameras. The first Geiger mode Ge/Si avalanche photodetectors (APDs) were demonstrated by Lu et al. operating at $\lambda = 1310$ nm with dark count rates (DCRs) of $>3 \times 10^7$ counts per second (cps) at 180 K. Mesa etched Ge/Si SPADs were subsequently demonstrated with an SPDE of 4% at $\lambda = 1310$ nm, a DCR of 6 Mcps, and 300 ps jitters. These devices also recorded the first operation at $\lambda > 1550$ nm by any silicon-based device. Another significant achievement was reported by moving to planar device geometries, where $100 \mu$m wide Ge/Si SPADs demonstrated SPDEs of 38% at 125 K. Scaling these devices to 26 $\mu$m diameters resulted in SPDEs of 29.4% with low DCRs of 100 kcps and jitters of 134 $\pm$ 10 ps at 125 K.

The noise equivalent power (NEP) provides a useful figure of merit to compare different materials, devices, and technologies, and it is defined as

$$NEP = \frac{hv}{SPDE \sqrt{2DCR}},$$

where $h$ is Planck’s constant and $\nu$ is the detection frequency. In the NEP vs temperature plot of Fig. 5, Ge/Si SPADs and an all Ge APD operating at $\lambda = 1310$ nm are compared with InGaAs devices...
Nonetheless, the Ge/Si SPADs have demonstrated superior performances as compared to InGaAs SPADs in two realms that could be significant drivers to improve the technology (Table I). The first is a significantly lower afterpulsing probability. Afterpulsing is the false triggering by electrons or holes that are released from traps in the avalanche region of the device when there has been no photon trigger. The initial planar Ge/Si SPADs demonstrate afterpulsing probabilities, which were a factor of five lower than commercial InGaAs devices at 125 K, and all Ge APDs also have demonstrated lower afterpulsing than InGaAs SPADs. The activation energies for the afterpulsing in Ge on Si SPADs suggest that threading dislocations (TDs) originated by the heteroepitaxy and propagating into Si substrates may be responsible for the afterpulsing effect. Advanced heteroepitaxy leading to a reduction of these defects should therefore further reduce the afterpulsing probability and allow higher operating repetition rates than InGaAs SPADs and superconducting detectors. The second area of strength for Ge/Si SPADs is the injection efficiency from the absorber to the avalanche region. When using an InGaAs absorber and an InP avalanche region, there is a barrier of about 240 meV, which reduces the injection efficiency of holes from the InGaAs into the InP material. Graded injectors are therefore required to increase the injection efficiency. In the Ge/Si system, the Ge L-valley conduction band is shallower than commercial InGaAs devices, and Ge L-valleys, resulting in far higher injection efficiency into the avalanche region and nearly temperature-independent SPDE.

To consider the future perspective for Ge on Si SPADs, the initial markets for SWIR SPADs require devices to operate on Peltier coolers at temperatures (T) > 223 K, with a significant market share for L-band applications, if the present efficiency can be improved. The Ge/Si devices to date have had far higher DCRs than the mature InGaAs technology devices and operate at lower temperatures. One simple way to reduce DCR is to shrink the device area. Indeed, recent results have shown that the DCR scales directly with the area, leading to estimates of NEP as low as 5 × 10⁻⁶ WHz⁻¹/2 for a 1 μm thick Ge absorber in a 10 μm diameter SPAD with 8.5% excess bias at 223 K and 1310 nm wavelength.

![Figure 5. The NEP vs temperature for a 26 μm diameter mesa device (green circle), an all Ge APO (gray square), Ge on Si waveguide coupled SPAD (orange square), 100 μm diameter planar Ge on Si SPAD (blue circles), and 26 μm diameter planar Ge on Si SPAD (red circles) all at λ = 1310 nm. The NEP of a commercial 25 μm diameter InGaAs device at 1550 nm (purple triangles) is also plotted for comparison. The black triangles are early 20 μm InGaAs SPADs at 1550 nm. Commercial InGaAs SPADs from IdQuantique (green) and MPD (blue) at 1310 nm are also shown with predicted scaled Ge on Si SPAD devices to 10 μm diameter (blue line) and 5 μm diameter (ref line) and a 26 μm diameter Ge on Si SPAD with no Shockley–Read–Hall (SRH) generation (red dotted line).](image-url)

**TABLE I.** A comparison of SWIR SPAD devices from commercial InGaAs/InP and research devices of Ge on Si materials.

| Material | Id quantique ID230 | MPD pdm-ir | Ge on Si [113] | Ge on Si SPAD scaled | Ge on Si SPAD no SRH |
|----------|--------------------|------------|---------------|----------------------|----------------------|
| Wavelength (nm) | 975–1700 | 975–1700 | 600–1450 | 600–1550 | 600–1550 |
| Operating temperature (K) | 183 | 225 | 125 | 225 | 225 |
| Device diameter | 25 μm | 25 μm | 26 μm | 5 μm | 26 μm |
| SPDE (%) | 25.6% @ 1310 nm | 39% @ 1310 nm (at 6 V excess bias) | 29.4% @ 1310 nm (at 2.7 V excess bias) | 57% @ 1310 nm for 2 μm absorber | 76% @ 1310 nm for 4 μm absorber |
| NEP (W/Hz) @ 1310 nm | 1.2 × 10⁻¹⁷ | 3.4 × 10⁻¹⁶ | 7.7×10⁻¹⁷ | 2.6 × 10⁻¹⁶ | 1.9 × 10⁻¹⁷ |
| Dark count rate (counts per second) | 200 | 4 × 10⁵ | 1.0 × 10⁶ at 2.7 V excess bias | 4.7 × 10⁵ (area scaling) (no SRH) |
| Timing jitter FWHM (ps) | 200 | 83 | 157 ± 10 | 15 | 100 |
| Afterpulse probability | 5% at 20 μs | Not stated | 10% at 10 μs | 0.4% at 10 μs | <0.1% at 1 μs |
| Gating frequency (MHz) | Can be free-running | 100 or free running | 1% at 50 μs | 10 | ≥100 or free running | Free running |

at 1310 and 1550 nm. Here, the best planar Ge/Si SPADs are comparable to the early InGaAs devices. While these initial results are encouraging, Fig. 5 demonstrates that the Ge/Si devices have to improve significantly to compete in terms of performance with the commercial InGaAs SPADs.
There are a number of other approaches that can be considered to reduce the DCR further, such as the use of guard rings and the use of an improved surface passivation layer. Furthermore, the contribution of misfit and threading dislocations occurring during the heteroepitaxial growth on DCR is not yet understood, but they should contribute to higher DCR through Shockley–Read–Hall (SRH) and trap assisted tunneling processes active at high electric fields. Figure 5 demonstrates that the removal of Shockley–Read–Hall processes results in higher performance than current InGaAs devices. From this Perspective, limited area growth of the Ge absorbers with reduced diameters and waveguide coupled Ge on Si SPADs have the potential to significantly reduce the dislocation density, which will reduce Shockley–Read–Hall processes with the potential for ≤1 threading dislocation for device dimensions below 10 μm.

Another key feature is the optimization of the Ge absorber thickness since only 50% of photons are presently absorbed,120,121 with Ge/Si SPADs with 2 μm-thick Ge absorbers estimated to improve the NEP to 2.3 × 10⁻¹⁶ WHz⁻¹² (5 μm diameter SPADs with 8.5% excess bias at T = 223 K and λ = 1310 nm). Figure 5 demonstrates that a combination of reduced diameter Ge on Si SPADs with a reduction in defects has the potential to achieve InGaAs performance.

For what instead concerns the operation wavelength, if one considers the bandgap reduction for increasing temperature T and taking also into account that the tensile strain of 0.2% typically observed in Ge/Si heterostructures also contributes to red-shift the band-gap, an operating temperature of at least T = 223 K is required to achieve direct bandgap absorption at 1550 nm.122 Note that techniques that can further increase the tensile strain across the Ge absorber could also improve the SPDE of the device.123 1 μm of Ge at 1550 nm wavelength absorbs less than 10% of incident photons at 223 K. Therefore, much thicker Ge layers are essential for improved SPDE at 1550 nm. As an alternative, one approach that is being taken is the use of Ge microcrystals grown on Si pillars, which allow the growth of many μm thick low defect Ge.124 Another approach could be the incorporation of GeSn heterolayers to further red-shift the direct bandgap absorption and improve the SPDE further.

III. SOURCES, DETECTORS, AND NON-LINEAR DEVICES FOR THE MIR/FIR RANGE

A. Quantum cascade lasers

The development of quantum cascade lasers (QCLs) represents a promising strategy to circumvent the poor radiative efficiency related to the indirect bandgap, which hinders the realization of Si-based emitters. In a QCL, only a single type of carrier is responsible for the photon generation process since it involves electron (hole) transitions between quantized electronic states belonging to the conduction (valence) band, as schematically depicted in Fig. 6. The QCL structure comprises a stack of several identical modules, each made by different quantum wells (QW) coupled by tunnel barriers. The application of a vertical electric field of appropriate intensity aligns the subband energy levels, thus allowing the current flow via resonant tunneling and the achievement of population inversion between laser states, where carriers undergo a stimulated radiative intersubband transition (ISBT), resulting in photon emission. The carriers, now in the lower laser state, are then injected into the next module of the MQW stack, thus realizing the so-called quantum cascade. The optical gain is then proportional to the number N of MQW periods.

The first QCL device was demonstrated in 1994 at Bell Laboratories, with an emission wavelength in the mid-IR, sparking a rapid growth in the research field, further boosted by the subsequent extension of achievable emission wavelengths toward the THz region. Indeed, while THz radiation is potentially relevant for several applications in different fields (e.g., medicine, biology, ICT, quality control, and security), its broad diffusion is hampered by the lack of affordable, practical light sources, which instead are still expensive (∼1–10 k€), bulky, and/or requiring cryogenic operating temperature.125 From this perspective, the high cost and low operating temperatures of the III–V material system, traditionally

![FIG. 6. Periodic L-point band edge profile and intersubband states in the strain compensated Ge/Si₀.₁₅Ge₀.₈₅ MQW active region of a n-type Ge/SiGe QCL with a vertical electrical bias of 2.5 kV/cm. Starting from the SiGe injection barrier at the left side of the large Ge QW, the layer sequence, with thicknesses in nm, is 4.7/19.9/3/13.9/3/11.6/3/3.8/3/3.8/3/3.8/3/9.7; the Ge wells are in standard fonts and the Si₀.₁₅Ge₀.₈₅ barriers are in bold; the phosphorus-doped layer is underlined. The upper (2) and lower (1) laser levels are indicated in red and blue, respectively.](image-url)
used for QCL devices, pose severe limitations for large volume applications.\textsuperscript{130,131}

The first attempts toward a group IV-based QC emitter focused on p-type Si/Si-rich SiGe MQWs, confining holes in the valence band of Si. This hole-based material system had been chosen in view of its larger and better-established band offset, as compared to the conduction band, enabling the emission of photons with an energy of $\sim 100–200$ meV ($\lambda \sim 12–6 \mu$m). Although ISBT electroluminescence has been demonstrated already in the year 2000,\textsuperscript{129,133,134} the large non-parabolicity of the valence bands, which leads to subbands with non-parallel dispersions, and the still high hole effective masses ($m^* > 0.3m_0$) have made it difficult to engineer the optical gain above 10 cm$^{-1}$,\textsuperscript{129} a value well below the threshold required to overcome typical waveguide losses. Moreover, the engineering of the ISB transport dynamics was also complicated by the unavoidable interaction between heavy- and light-hole subbands.

More recently, n-type SiGe/Ge MQWs structures, where electron subbands are formed in the conduction band, have emerged as the most promising configuration\textsuperscript{134,135} and have driven an intense research effort in the last decade. Their main features are the relatively light effective mass along the confinement direction ($m^* = 0.13m_0$) and the relatively large L-point band offset ($\sim 120$ meV), enabling the design of ISBT in the THz range. Moreover, thanks to the weaker electron–phonon interaction featured by non-polar elemental semiconductor lattices, the optical gain is expected to be robust against increasing lattice temperature, and the recent state-of-the-art numerical calculations predict laser action persisting up to room temperature (RT).\textsuperscript{136}

Decisive advancements in different fields are necessary for the development of a SiGe/Ge QCL. The active material epitaxy and engineering, the waveguide design and fabrication, require substantial refinement. Since the THz radiation wavelength is in the $\lambda = 50–100$ $\mu$m range, the QCL active region has to be $\sim 10$ $\mu$m thick, thereby allowing for a high optical mode overlap and to have a large enough number of periods $N$ to achieve net optical gain. Structures of such thickness accumulate a large heteroepitaxial strain energy, arising from the lattice mismatch of the layers present in each module of the cascade, that can be potentially plasticly relaxed by the formation of highly detrimental dislocations. Consequently, a very fine strain balancing is the first step to be solved.

To this aim, the first step is the introduction and optimization of a relaxed SiGe virtual substrate (VS) deposited on the Si substrate, whose Ge content has to be intermediate in between that of the SiGe barrier and Ge well materials. In this way, the compressive strain of the Ge wells can be compensated by the tensile deformation of the barrier layers. A leap forward in the optimization of the VS has been achieved in 2008, with the introduction of the so-called reversed graded (RG) approach.\textsuperscript{137} In RG-VS, a rather thick (1–5 $\mu$m) relaxed Ge buffer layer is grown directly on the Si substrate, followed by layers featuring decreasing Ge contents, down to the desired SiGe composition. State-of-the-art Si$_{0.05}$Ge$_{0.95}$ relaxed VSs featuring a threading dislocation (TD) density as low as $2 \times 10^6$ cm$^{-2}$ have been demonstrated.\textsuperscript{138}

The adoption of the RG-VS has enabled the achievement of higher quality active material and the observation of ISBT in the 16–50 meV range from SiGe/MQWs samples.\textsuperscript{139,140} Note that the III–V Reststrahlen band in the THz range (24–41 meV) is entirely contained in this energy interval. FWHM values for the absorption peaks as small as 5 meV were measured and accounted for by different scattering mechanisms.\textsuperscript{141} Degenerate pump–probe absorption experiments with the same set of MQW samples demonstrated sub-band lifetimes larger than 10 ps. These promising values have also been observed at relatively high lattice temperatures and above the subband energy separation corresponding to the threshold for ISB relaxation via optical phonon emission.\textsuperscript{142,143} This result is to be attributed to the weaker electron–phonon interaction occurring in the non-polar layer of group-IV-based devices that make feasible the achievement of room temperature laser operation. These experimental data allowed to shed light on the peculiar role of the electron–phonon coupling in the intersubband dynamics of Ge-based 2D systems.

The second key aspect to be optimized is the accuracy of the epitaxial deposition of individual wells and barriers. Indeed, the few-atom thick layers comprised in a typical QCL design must be realized with ultra-high accuracy and reproducibility throughout the whole active layer stack to grant for the ISB alignment and tunneling rate. Furthermore, the quality of the hundreds of different heterointerfaces present in a QCL has a high impact on the optoelectronic performance.

Indeed, the complete understanding of the ISB dynamics requires a full-3D reconstruction of the SiGe/Ge heterointerfaces. Recently, by using atom probe tomography, it has been observed that Ge/SiGe interfaces feature a limited interdiffusion length ($W \sim 1$ nm) and rms roughness ($\Delta \sim 0.1$ nm).\textsuperscript{144} These experimental inputs have been used to further improve the accuracy of the QCL gain calculations that lead to a predicted 300 K gain of $\sim 20$ cm$^{-1}$. This value is compatible with lasing operation since numerical simulations for Ge-rich double metal SiGe WGs, performed using as input accurate experimental values for the dielectric constant of the alloy and metal layers, estimated optical waveguide cavity losses in the THz region <20 cm$^{-1}$.\textsuperscript{134} Although the growth of several $\mu$m thick strain compensated QCL structures has been already realized,\textsuperscript{146} the fabrication and the experimental characterization of a double plasmon SiGe waveguide in the THz range to confirm these predictions has not been reported yet.

All the advancements discussed above, together with the demonstration of efficient NiGe Ohmic contact modules with low specific contact resistance,\textsuperscript{147} very recently lead to the observation of ISB electroluminescence (EL) from an n-type SiGe/Ge cascade structure (Fig. 7),\textsuperscript{146} albeit with a relatively low efficiency (Table II). The reason for this sub-optimal performance was attributed to the large electron temperature associated with the characteristic weak electron–phonon interaction, to a not yet optimized doping profile, and to the possible detrimental role of threading dislocations.

The demonstration of ISB EL certainly represents a key intermediate milestone toward the realization of a laser device since it evidences that a high degree of control both in terms of growth and modeling of the active material has been achieved. Nevertheless, not all the physical phenomena underlying the complex intersubband dynamics at the basis of the n-type SiGe/Ge QCL are fully understood. At the same time, the design adopted so far for the EL experiments has been chosen for its simplicity and robustness, while it is clear that gain optimization requires more advanced architectures, which pose peculiar challenges. From this perspective, it would be interesting to explore QCL designs featuring larger periods...
and lower operation bias in the effort to limit the electron heating induced by the electrical field, as, for instance, suggested in Ref. 148. Another aspect that requires a deeper understanding is the possible role of parasitic subband states, originated by the $\Delta_2$ band edge, whose energy in the SiGe tensile strained barrier material is located below the L one. Nevertheless, thanks to the large confinement energy associated with the thin barriers, the energy of these $\Delta_2$ subbands can be designed to be larger than those of the L states involved in the optical transition.\(^{146}\) It has not been definitively clarified if, however, due to thermal excitation caused by the high electron temperature, carriers also populate these $\Delta_2$ subbands and, if this is the case, how the actual (i.e., including the Hartree contribution) band edge potential profile and the ISB transport are influenced.

Finally, also the control of the threading dislocation (TD) density in the active region and a better understanding of their role in the ISB transport are an open challenge toward the realization of the laser. TDs may introduce detrimental parallel vertical transport channels. At the same time, this kind of extended defects seems to be associated with a p-type background,\(^{149}\) which needs to be counter-balanced by means of donors. In particular, with the state-of-the-art TD density of about $2 \times 10^{10}$ cm\(^{-2}\), the extra-electron sheet density required to counteract the p-background has been roughly estimated to be at the same scale as the one typically used to dope the active region of the QCL ($\sim 10^{10}$ cm\(^{-2}\)). Therefore, the calibration of the optimum doping in the QCL structures would greatly benefit from a deeper understanding of the electrical properties of the TDs. We also note that the extra donor density required for compensation significantly increases the Coulomb scattering rates, thus lowering the subband lifetimes and the dephasing times. For these reasons, future research efforts should focus on a further suppression of the TD density, for which, leveraging on the optimization of the VS, it is realistic to target the $10^{8}$ cm\(^{-2}\) scale.\(^{150}\)

### B. Quantum well infrared photodetectors

Intersubband transitions (ISBTs) can also be exploited to absorb light at infrared wavelengths (IR) and generate photocurrents, as is the case in quantum-well infrared detectors (QWIPs).\(^{150–158}\)

By suitable band engineering, ISBT can be designed so that the QWIP can operate at wavelengths in the $\lambda = 3–20$ μm range. Photodetection occurs through photons that are absorbed within the doped QW regions (Fig. 8) by charge carriers undergoing a transition between two subbands (bound-to-bound transitions) or one subband as the initial state and one final state of the continuum above or at the barrier energy (bound-to-continuum or non-radiative upper laser subband lifetime $\tau_{NR}$ (ps)

| TABLE II. Comparison of relevant physical quantities estimated for n-type GaAs/AlGaAs and Ge/SiGe-based QCL devices with the same design. Gain data represent a theoretical estimation, refer to a four well bound-to-continuum architecture, and are taken from Ref. 136. The other reported values refer to a single quantum well active region design and have been reported in Ref. 148 using both theoretical inputs and experimental electroluminescence data. |
|---------------------------------|-------------|-------------|
| GaAs/AlGaAs | Ge/SiGe |
| Material gain @10 K (cm\(^{-1}\)) | 98 | 63 |
| Material gain @125 K (cm\(^{-1}\)) | 60 | 60 |
| Material gain @300 K (cm\(^{-1}\)) | −15 | 20 |
| Intersubband EL line broadening $\Delta f$ | 0.05 | 0.2 |
| Non-radiative upper laser subband lifetime $\tau_{NR}$ (ps) | 12.2 | 2–12 |
| Radiative efficiency $\eta_{rad}$ | $1.5 \times 10^{-6}$ | $1.2 \times 10^{-7} - 8 \times 10^{-7}$ |

![FIG. 7. An optical microscope image of an n-type Ge/SiGe THz quantum cascade emitter of 100 μm width.](image1)

![FIG. 8. Relevant subband states for the tensile strained p-type QWIP device discussed in Ref. 151, calculated with an electrical bias of 6 kV/cm. The MQW stack, grown on a Si\(_{0.78}\)Ge\(_{0.22}\) substrate, consists of 4.5 nm tick Si\(_{0.7}\)Ge\(_{0.3}\) wells, separated by 10 nm tick Si\(_{0.9}\)Ge\(_{0.1}\) barriers. The black (red) curves refer to the LH (HH) band profile and subband states (continuous levels in the energy region above the LH barrier are not reported). Due to the tensile strain field, the fundamental state has light-hole symmetry, and the LH–LH, LH–HH, and LH–SO-band transitions are associated with responsivity peaks in both the terahertz and mid-infrared regimes.](image2)
quasi-continuum transitions, respectively). In order to have sufficient absorption, the QWIP structures are typically made of 50–100 MQW.

As discussed in Sec. III A for the QCL, the use of ISBT decouples the QWIP’s operation wavelength range from the bandgap of the semiconductor materials in the well and barrier regions.

QWIPs based on the III–V semiconductor heterostructure (usually lattice-matched GaAs/AlGaAs MQW) have been successfully fabricated and commercialized, e.g., for thermal vision applications.

State-of-the-art III–V semiconductor QWIPs are fabricated using n-type doped quantum wells since the higher electron mobility leads to higher responsivity, and the smaller electron mass leads to higher optical absorption with respect to p-type QWIPs. However, for n-type QWIPs, the absence of off-diagonal elements in the effective mass tensor makes the presence of grating couplers unavoidable to activate ISBT with normal incident light. In contrast, p-type QWIPs are not affected by this limitation but typically exhibit lower dark currents.

Recent years have seen a growing need for QWIPs in applications such as sensing, medical diagnostics, free-space communications, and thermal imaging. As for other SP applications, a key feature would be the integration of SiGe based QWIPs with integrated I/O circuitry. In this case, since the mismatch of the thermal conductivity for SiGe QW on Si is reduced compared to III–V devices, limitations on array sizes, e.g., in thermal imaging applications, are also reduced.

For these reasons, SiGe-based QWIPs have always been in the spotlight, with first efforts dating back to the early 1990s (Table I). Early attempts were focusing on SiGe quantum wells grown on Si(001) substrates, and molecular beam epitaxy (MBE) was almost exclusively employed, with only one exception, where low pressure vapor phase epitaxy (LPVPE) was instead used. Most MBE-grown devices were limited in thickness, with the active region totaling up to ten wells. Consequently, device responsivities were limited to values below 0.1 A/W at 77 K. In one notable exception, where the structures comprised 50 quantum wells, larger responsivities of 0.3–0.6 A/W were reported. For all devices, detectivities at 77 K were about one order of magnitude lower than those of strain-free, n-type GaAs/AlGaAs QWIPs. For example, an average peak detectivity of $D^*$ = 8.3 × 10^10 cm Hz^1/2/W at 77 K was reported for a AlGaAs/GaAs QWPI with a response peak at ~8.4–8.55 μm.

As we discussed for QCL, the relatively large Ge–Si lattice mismatch makes the growth of a large number of quantum wells challenging, limiting the layer thickness and device performance. For all devices grown on Si(001) substrates, Si was used exclusively as the barrier layer material. Consequently, the realized QWIPs were all of p-type, since the large valence-band offset within SiGe/Si structures, as well as the higher absorption coefficient as a result of the small hole effective mass, proved to be advantageous. The quantum well regions were composed of $\text{Si}_{1-x}\text{Ge}_x$ with a Ge content $x$ up to 0.6, leading to devices operating in a wavelength range $\lambda = 2.5–11$ μm.

Aiming at overcoming strain-induced defect formation, especially for $\text{Si}_{1-x}\text{Ge}_x$ wells with high Ge content, later works explored the use of virtual substrates (VS), Si nanomembranes, or strained SOI substrates to enable the growth of high-quality, thick active layers.

As an alternative, the use of Ge wells or $\text{Si}_{1-x}\text{Ge}_x$ barriers was investigated, although with unsatisfactory device responsivities (<0.1 A/W).

Nonetheless, early theoretical calculations for SiGe QWIPs also predicted good device performance for n-type devices. This early

| Reference | Growth method | Substrate | Composition (well/barrier) | Approx. Thickness QW-region | Responsivity $(\text{mA/W})$ | Detectivity $D^*$ $(\text{cm Hz}^{1/2}/\text{W})$ |
|-----------|---------------|-----------|----------------------------|-----------------------------|-----------------------------|--------------------------------------------|
| 150       | MBE           | Si(001)   | $\text{Si}_{0.4}\text{Ge}_{0.6}/\text{Si}$ | 340 nm 8.1                  | ...                         | ...                                        |
| 151       | MBE           | Si(001)   | $\text{Si}_{0.35}\text{Ge}_{0.65}/\text{Si}$ | 2650 nm ~9 300–600 (77 K)    | ...                         | ...                                        |
| 153       | MBE           | Si(001)   | $\text{Si}_{0.75}\text{Ge}_{0.25}/\text{Si}$ | 340 nm 10.8 36 (77 K) 3.3 x 10^9 | ...                         | ...                                        |
| 154       | MBE           | Si(001)   | $\text{Si}_{1.4}\text{Ge}_{0.6}/\text{Si}$ | 170 nm 2.0 2.4 30 (77 K) 4 x 10^10 | ...                         | ...                                        |
| 156       | LPVPE         | Si(001)   | $\text{Si}_{0.865}\text{Ge}_{0.135}/\text{Si}$ | 2700 nm 7.2                  | ...                         | ...                                        |
| 157       | MBE           | Si(001)   | $\text{Si}_{1.5}\text{Ge}_{0.5}/\text{Si}$ | 2700 nm 7.7                  | ...                         | ...                                        |
| 158       | MBE           | Si(001)   | $\text{Si}_{1.4}\text{Ge}_{0.6}/\text{Si}$ | 2700 nm 9.5                  | ...                         | ...                                        |
| 159       | MBE           | Si(001)   | $\text{Si}_{0.85}\text{Ge}_{0.15}/\text{Si}$ | 2700 nm 10.2                 | ...                         | ...                                        |
| 160       | MBE           | Si(001)   | $\text{Si}_{0.75}\text{Ge}_{0.25}/\text{Si}$ | 520 nm 5 76 (77 K) 2 x 10^10 | ...                         | ...                                        |
| 161       | MBE           | Si(001)   | $\text{Si}_{0.77}\text{Ge}_{0.23}/\text{Si}$ | 1060 nm ~7.5 3 (10 K)        | ...                         | ...                                        |
| 162       | LEPECVD       | $\text{Si}_{0.2}\text{Ge}_{0.8}$ (001) VS | Varying comp. 280 nm ~4–5.5 1–1.5 (77 K) | ...                         | ...                                        |
| 164       | LPCVD         | Si nano-membrane | $\text{Ge}/\text{Si}_{0.5}\text{Ge}_{0.5}$ | 4–7 μm 6.5–8 3.7 (10 K)     | ...                         | ...                                        |
| 163       | MBE           | $\text{Si}_{0.85}\text{Ge}_{0.15}/\text{Si}$ | $\text{Si}_{0.38}\text{Ge}_{0.62}/\text{Si}$ | 150 nm ~5 73 (80 K)         | ...                         | ...                                        |

Overview of key parameters reported on SiGe-based QWIPs.
simulation work used (111) substrates to improve performance through an increased density of states from degenerate valleys, and recent Ge (111) epitaxial results could benefit from a similar scheme. As previously discussed, the non-polar crystal structure of group-IV semiconductors can lead to reduced phonon scattering and potentially large detection efficiency in Ge/SiGe QWIPs, even when compared to their III–V counterparts. Initial experimental results on Ge/SiGe quantum cascade structures have demonstrated strong THz intersubband absorption with the bandstructure potentially allowing operation up to 10 μm wavelength and the strain-balanced QWs now allowing large numbers of QWs to enable performance similar to group III–V based QWIPs.

In conclusion, the device performance of group-IV QWIPs has still to be improved if those devices are to be competitive. Device responsivity can be increased by increasing the total thickness of the active region. This requires careful control of strain-induced defects since, e.g., threading dislocations induced by growth can increase the dark current in vertical devices and limit device performance. The same recent progress that we believe can pave the way to the SiGe QCL (i.e., high-quality virtual substrates, advances in CVD growth on the precise control of layer thicknesses, low interface roughness, and very good composition control at large thicknesses) and could, therefore, enable the realization of high-performance SiGe QWIP devices. Alternatively, MBE can be used as a growth strategy with deposition of ultrathin Ge wells on Si, which allows the growth of a large number of quantum wells with high crystal quality in the active layers of the devices.

C. Non-linear effects

The development of SP has attracted great interest in the investigation of non-linear optics in Si-compatible structures. Indeed, crystalline Si features strong optical non-linearities, which can be further enhanced by means of appropriate material engineering using Si-rich silicon oxide and nitride. Currently, nonlinear effects, such as frequency generation and conversion, stimulated Raman scattering, self- and cross-phase modulation (SPM and XPM), four-wave mixing (FWM), and supercontinuum generation (SCG), are actively investigated.

Most of these reports focused on the NIR, while the extension of the electromagnetic spectral window usable for engineering non-linear interaction, from X-ray to mid-IR and even THz, has recently been driving a significant portion of the research efforts in the non-linear optics community, owing to possible impacts on several realms, including molecular sensing, early medical diagnosis, or secure communications. From this perspective, Ge and SiGe alloys can help to widen the field, thanks to their lower lower than Si and a much larger wavelength transparency window up to λ ∼ 14 μm. Furthermore, Ge exhibits a higher refractive and Kerr index (n2) than silicon. These properties can be exploited to design mid-IR supercontinuum generation (SCG) sources, which, for instance, can be integrated on Si wafer embedded with all the optical functionalities for on-chip broadband molecular absorption spectroscopy. In fact, radiation produced by SCG retains the intensity, directionality, and coherence properties of the laser pump while having a wide spectral window that can be exploited to carry out molecular spectroscopy based on their mid-IR optical fingerprint.

Supercontinuum generation across a broad mid-IR wavelength range (λ ∼ 3–8.5 μm) has been recently demonstrated in Ref. using an air clad SiGe/Si waveguide. The authors of Ref. optimized the waveguide design to achieve low anomalous dispersion over a large bandwidth and, thanks to the very low propagation losses (0.4 dB/cm), an active region 7 cm long. Indeed, it is worth noticing that there is a trade-off for the optimal transverse effective area Aeff of the waveguide (WG) employed for SC. In fact, smaller values of Aeff increase the nonlinear parameter γ = 2π/κAeff and allow for single-mode operation but at the same time negatively impact the low energy cutoff of the wavelength range for SC due to increased leakages in the absorbing Si substrate. Through a careful design of the WG geometry, a spectrally bright SC signal with more than 10 mW on-chip power, corresponding to ~50% power conversion efficiency in supercontinuum generation, was obtained.

The fine-tuning of γ and low anomalous dispersion can be achieved by vertically varying the refractive index as done in SiGe/SiGe–Si graded WGs. This recently led to the demonstration of two-octave spanning SCG between 3 and 13 μm achieved by pumping a SiGe/Si graded WG with an 8.5 μm laser and 8 mW input power.

Remarkably, this on-chip power and bandwidth are significantly larger than those measured in other CMOS-compatible platforms and well compare to data previously reported using the on-chip chalcogenide glass material platform. A little later, SiGe WGs have been engineered to support a TM mode featuring the low all-normal dispersion regime in the λ = 2.8–5.5 μm range. By pumping with a 4 μm source featuring a t = 205 fs width, the authors reported fully coherent SCG. Interestingly, the SCG in this system could be exploited to achieve pulse compression down to 22 fs. This opportunity represents an interesting perspective since it allows us to envisage the compression of picosecond pulses generated from an on-chip mode-locked laser down to the femtosecond scale. Such chip-based fully coherent SC sources can enable many practical applications, such as optical coherent tomography and coherent anti-Stokes Raman spectroscopy.

For SiGe waveguides featuring larger Ge concentrations, an in-depth characterization of their nonlinear properties has been recently reported in Ref. In that work, experimental data at λ = 1.58 μm from Si1−xGex/Si WGs with x > 0.7 have been used in combination with theoretical models to assess the wavelength-dependent value of n2 and the two phonon absorption coefficient βTPA in the λ = 1–10 μm range. In particular, upon investigating SiGe/Si WG samples featuring Ge content in the x = 0.7–0.9 range, the authors of Ref. have also clarified the impact of the Δ-L bandgap crossover on these key parameters.

Improvements in the SiGe heteroepitaxy on Si substrates have enabled the investigation of nonlinear optic effects in quantum multilayer nanostructures related to intersubband optical transitions between confined levels. In particular, in the III–V based multilayer structure, an impressive enhancement of the χ(2) coefficient has been observed, controlling the second harmonic generation (SHG) process, with reported giant values of the order of ~10⁵ pm/V, much larger than their “bulk” counterparts. Soon after, promising results have also been achieved with SiGe multilayers. Overcoming, in this way, the vanishing of χ(2) due to the centrosymmetric character of diamond lattices. We point here out...
that SHG is of great importance for many classical and quantum applications from high-speed optical modulation via the Pockels effect\textsuperscript{183} to $\nu$-$2\nu$ frequency-comb self-referencing,\textsuperscript{184} and even to direct frequency-comb generation by cascaded SHG events that can mimic third-order nonlinear effects.\textsuperscript{185}

From a more general perspective, different approaches have been proposed to reduce the Si and Ge symmetry, thus allowing the second harmonic generation process. These approaches exploit an electric field,\textsuperscript{186} a non-uniform strain field, or leverage on surface-induced asymmetry.\textsuperscript{187} Recently, relevant SHG power efficiencies of the order of $\sim$20\% have been obtained by leveraging electric bias-induced $\chi^{(2)}$ values, exploiting an optical cavity that amplifies the SH field produced by the non-linear material.\textsuperscript{188} Unfortunately, modest $\chi^{(2)}$ values of the order of tens of pm/V have been reported so far. Furthermore, the transparency range of the used SOI and Si$_3$N$_4$ materials limited the operating wavelength range to $\lambda > 4 \mu m$ due to strong phonon absorption.

In this context, the engineering of intersubband transitions in doped asymmetric quantum multilayers represents an elegant and efficient alternative to overcome most of these limitations. In these kinds of structures, quantum confinement effects can be tuned to control the subband energy spacing, thus making possible the realization of double resonances at $\nu E_{ISB}$, which greatly enhance the $\chi^{(2)}$ value at the corresponding photon frequency.\textsuperscript{189} Indeed, ISB transitions between these double resonant confined levels can display large oscillator strengths. This property, in combination with the parallel dispersion in the k-space of the 2D subbands associated with a $\delta(\nu E_{ISB})$-like joint density of states, is responsible for the observed giant values of the $\chi^{(2)} (\nu = E_{ISB})$ coefficient mentioned above.

Literature reports on SHG in SiGe structures focus on SiGe step quantum wells,\textsuperscript{190,191} or double asymmetric coupled quantum well structures.\textsuperscript{192,193} As in the case of the QCLs, those reports investigate p-type structures, targeting input photon energies around 100–150 meV ($\lambda = 12–8 \mu m$). Giant $\chi^{(2)}$ coefficients of the order of $10^2$ pm/V have been estimated, in good agreement with theoretical calculations and despite the typical deviations from parallel subband dispersion, which are triggered by the coupling of the different parent valence bands. Remarkably, this coupling brought about also positive effects for what concerns possible practical applications. In fact, we first note that $\chi^{(2)}$ is indeed a third-order tensorial quantity, usually indicated as $\chi^{(2)}_{\alpha\beta\gamma}$; the $i$ and $j$ indices refer to the polarization direction of the two input photon fields, while $k$ indicates the polarization of the 2h output photons. Different from what happens in n-type systems, where only the diagonal term $\chi^{(2)}_{zzz}$ associated with the growth direction $z$ is different from zero, hole-based p-type systems can feature large off-diagonal $\chi^{(2)}_{ijk}$ values. This fact facilitates the coupling geometry with normal incident input light or the surface-normal emission with in-plane input beams in photonic integrated circuits, as discussed in Refs. 181 and 192. Therefore, SiGe nanostructures engineered for SHG might be useful in the near future for spectroscopic applications. As a matter of fact, the MIR range targeted by SiGe p-type structures (with typical second harmonic in the $\lambda = 4–6 \mu m$ interval) is ideal for molecular sensing and occurs in a transparency window of Ge-rich SiGe integrated waveguides developed just for on-chip spectroscopic sensing, as will be discussed in detail in Sec. IV.\textsuperscript{174}

IV. NEW CONCEPTS FOR SENSING

Among the possible applications for optoelectronic devices, the field of biosensing, in particular, has seen a rapid growth. There is a growing need for miniaturized and ubiquitous sensors to detect gas molecules and biochemical substances. Such sensors have the potential to, e.g., drive personalized healthcare solutions and enable environmental monitoring for long-term sustainability. While the biosensor market is still dominated by electrochemical sensors, optical sensors have the advantage that they exhibit high sensitivities, are compact, and usually cheap to fabricate. Here, we focus on recent developments and future perspectives for (a) refractive index sensing (where the presence of the target molecule leads to a detectable modification of the refractive index close to the sensor surface) and (b) absorption sensing (i.e., detecting the characteristic absorption spectra of the target molecules in the 1–20 $\mu m$ wavelength range for their identification) as two of the main sensing principles.

Refractive index sensing offers a label-free detection method for molecular binding events: in these sensors, the surface is functionalized with analyte molecules that selectively bind to target molecules. Optical biosensors based on refractive index sensing typically utilize the sensitivity of either plasmonic or photonic structures to changes in the surrounding refractive index induced by analyte-ligand binding events.\textsuperscript{194} The change $\Delta \nu$ in the refractive index can be detected via a shift $\Delta \lambda$ in resonance wavelength of the structures, and the sensitivity $S = \Delta \nu / \Delta \lambda$ measured in nm/RIU (Refractive Index Unit) can benchmark the performance. We note that for biosensors, the relevant quantity is the sensitivity to near-surface rather than bulk refractive index changes. Another quantity used for sensor benchmarking is the LOD (Limit Of Detection), which corresponds to the minimum detectable refractive index change.

While both plasmonic and photonic structures themselves have been commercialized for biosensing,\textsuperscript{195} they usually require bulky instrumentation, such as spectrometers for readout. However, refractive index sensors have the potential for further miniaturization, and there is a growing need to realize new on-chip sensor concepts that can be easily combined with integrated circuitry to enable continuous monitoring of biochemical parameters and fast sensor data processing. To date, on-chip Si-based label-free optical biosensors have, e.g., been realized by combining passive, functionalized structures, such as interferometers, photonic crystals, and resonant microcavities\textsuperscript{196} with waveguides and on-chip photodetectors for read-out.\textsuperscript{197} The on-chip footprint of these structures is comparatively large, usually covering many hundreds of square micrometers. Other approaches consist of stacking plasmonic structures, such as nanohole-arrays,\textsuperscript{198} plasmonic nanostructures,\textsuperscript{199} or chirped gratings,\textsuperscript{200} on transparent or membrane substrates with CMOS cameras for read-out.

The smallest sensor footprint can be realized by combining functionalized plasmonic structures directly with on-chip transducers.\textsuperscript{201–208} While these approaches have large potential for on-chip biosensing, they also highlight the challenges that are present for on-chip concepts. Some of the existing concepts have low signal strength that requires the use of lock-in amplifiers,\textsuperscript{202,204} and other concepts exhibit sensitivities that are below those of the plasmonic structures alone.\textsuperscript{201} Compared to spectrometer-based concepts, on-chip approaches typically have much lower LODs.
On-chip approaches often rely on using a single wavelength for interrogation rather than analyzing the spectrum over a wavelength range. In combination with the higher noise present in compact read-out devices, this can lead to a significant loss in sensitivity. Those challenges have to be overcome for on-chip biosensors with low LODs that can be robustly fabricated. In fact, integrated on-chip concepts have the potential to realize hybrid plasmonic-photonic sensors, in which both the high field enhancement in the vicinity of plasmonic structures and the low damping of photonic modes are utilized for high-performance sensing. In this way, the optical device not only acts as a transducer but can be actively used to fine-tune and improve sensor performance. Finally, it has to be noted that the successful realization of on-chip biosensors also requires addressing further issues stemming from system integration, such as packaging or microfluidics integration.

Spectroscopic absorption/transmission sensing is a large field with a $2.33Bn market in 2020, ranging from forensic, traceable part per billion (ppb) levels from Fourier Transform InfraRed (FTIR) spectrometers to dollar level blackbody-surface alarm sensors used for large percentage measurements of $O_2$, $CO_2$, or $CO$ that are compulsory safety sensors in many homes and workplaces globally. For atmospheric sensing, the two main windows with low water absorption are at 3–5 μm and 8–12 μm. The former is used for gas sensing with strong absorption from most environmental gases ($CO$, $CO_2$, $NO$, $NO_2$, $N_2O$, $CH_4$, $SO$, $SO_2$, etc.). The later atmospheric window has significant applications for stand-off chemical and biological sensing, especially for security applications. The region 6.7–20 μm is known as the molecular fingerprint region as molecular absorptions in this region can provide label-free, unique identification of analytes and is especially important for healthcare biomarkers and analytes but also chemical and security applications. Spectroscopic approaches can therefore provide high levels of sensitivity (ppb), specificity (unique identification with high probability), and dynamic range (ppb to 100% saturation) depending on the analyte being sampled and the type of spectroscopic sensor.

All spectroscopic sensors require a mid-infrared light source, a mechanism to select the wavelength, a mechanism to interact with the analyte, and a mid-infrared detector.

Two main approaches have been taken with group-IV materials to enable interaction with the analyte. The first one uses an optical cavity or photonic bandgap to enable stronger interaction with the analyte. The second uses waveguides on a chip to guide the light with evanescent coupling to the analyte. Despite the optical modal overlap of evanescent coupling being only 1%–5%, this has been the most popular method for chip-scale systems as it is easier to integrate all the components required to deliver a sensor, and the technology is closest to the Si photonic integrated circuits that have been developed for telecoms applications—this is, therefore, at our main focus.

The optical loss in the mid-infrared is key to choose the best waveguide material, and a review of waveguide loss for Si, Ge, and SiGe platforms is shown in Fig. 9 above 5 μm wavelength. There are significant opportunities for SiGe heterostructures for mid-infrared spectroscopic sensors, especially for wavelengths greater than 7 μm, where the losses in Si/SiO$_2$ technology start to significantly increase (Fig. 9). In his seminal paper from 2010, Soref highlighted the relevance of Ge as a key material for the fabrication of mid-IR photonic circuits integrated on Si. Indeed, only a few years later, the wide mid-IR transparency window of Ge and its favorable refractive index step with respect to the Si substrate have been exploited to obtain SiGe core, Ge on-Si, and graded-index SiGe waveguides. After these first demonstrations, much effort has been devoted to understanding loss and pushing the operating wavelength further in the infrared. The use of Ge$_{1−x}$Sn$_x$ and Si$_{1−x}$Ge$_x$Sn$_y$, in particular, has the potential to extend the operating wavelength of waveguides out to 20 μm, possibly allowing the whole of the fingerprint region to be used.

Group-IV materials can also be used for signal amplification in the case of low analyte concentrations: The transparency of Ge in the mid IR range makes it an ideal plasmonic material, which can be used to amplify electric fields to amplify sensing signals once the plasma frequency has been moved into the 3–8 μm window by tuning the electron doping in the $10^{19}$–$10^{20}$ cm$^{-3}$ range. When compared to metal-based plasmonic structures operating in the same wavelength range, semiconductor-based plasmonics offer a clear advantage for the fabrication of microantennas and slits resulting from both a more mature microfabrication technology and larger field enhancements.

Concerning group-IV based approaches for wavelength selection and signal detection, components for integrated spectroscopic sensor systems have been demonstrated using SiGe heterostructures, including quantum well intersubband photodetectors (see Sec. III), SiGe microbolometers, polarization rotators, phase shifters, and Mach–Zehnder interferometers. Low-cost sensors will require the integration of all the components for a sensor, and this has yet to be achieved but is a key target for future research. To estimate component requirements for
FIG. 10. The minimum detectable concentration of (a) ammonia and (b) phosgene as a function of source power for chip-scale evanescent waveguide sensing for 4.34 cm long Ge on Si waveguides with modal overlap of 1.2% and 5%, the first having already been demonstrated and the second is easily achievable.

V. CONCLUSIONS

The successful integration of Ge-based devices on the Si platform continues to be a key driver of SP. Putting the devices and structures discussed in this perspective again into a broader context, these innovations are on the one hand enabled by progress in material research and nanostructuring and, on the other hand, absorption cross section of $2.78 \times 10^{-18}$ cm$^2$. The second application is the detection of the chemical weapon phosgene, which is also used in the manufacture of a number of polyurethanes and polycarbonate plastics by identifying the C–Cl bond absorption line at 11.764 μm wavelength with an absorption cross section of $3.993 \times 10^{-18}$ cm$^2$ at standard temperature and pressure. Already waveguides with ~1 dB/cm are available at these wavelengths, and a room temperature microbolometer detector with a NEP of $4 \times 10^{-9}$ W/√Hz could be integrated on the same chip-scale sensor. For 1 dB/cm Ge on Si waveguides at these wavelengths, the optimum sensing length of the waveguide can be shown to be 4.34 cm, and Fig. 10 compares modal overlaps with the analyte ($\eta$) of 1.2% and 5%, the first having already been demonstrated and the second is easily achievable. Figure 10(a) demonstrates that a source power of 50 mW with a measurement bandwidth of 1 Hz and a signal to noise of 3 is required with a modal overlap of 1.2% to be used for ammonia breath analysis for diagnosing chronic kidney disease, and this reduces to 12 mW for 5% modal overlap. In addition, as a sensor for detecting the U.S. National Institute of Occupational Safety and Health (NIOSH) permissible exposure limit of 35 ppm requires a source power of 2 mW with a modal overlap of 1.2% and 0.5 mW with a modal overlap of 5% with the analyte. For phosgene, the NIOSH permissible exposure limit is 100 ppb and evanescent Ge on Si waveguides with 1.2% modal overlap requires 200 mW of power while 5% modal overlap requires 35 mW of source power.

These simulations for ammonia and phosgene demonstrate that Ge on Si waveguides, polarization rotators and microbolometer detectors all have suitable performance for chip-scale sensor applications, but to date, no group-IV-based light source is available with suitable power and performance at these wavelengths. Heterogeneous integration of III–V QCLs would allow sensing systems at the $10k level addressing markets for hospital and doctor surgery applications. More research is required on mid-infrared group-IV-based light sources if cheaper sources to allow integration of such sensors into consumer products, such as mobile phones, are to be achieved (see, e.g., Sec. III on QCL). There is, therefore, still a significant amount of research and optimization required to deliver chip-scale sensors that can be mass-manufactured at low cost for healthcare, security, and environmental sensing applications. There have also been a number of demonstrations of non-linear wavelength conversion devices for third harmonic generation and supercontinuum generation using Si$_{1-x}$Ge$_x$ materials, and this area has enormous potential for future research to achieve broadband sources and detectors for sensors. Second harmonic generation in the mid-IR is also possible exploiting resonant inter-subband transitions in Ge/SiGe QWs (see Sec. III). Finally, we note that Ge$_{1-y}$Sn$_y$ and Si$_{1-x-y}$Ge$_x$Sn$_y$ also have the potential to enable the monolithic integration of LEDs, lasers, and photodetectors for chip-scale mid-infrared sensors at potentially out to 12 μm wavelength.
driven by the need to provide cost-effective solutions to an increasingly large range of applications covering wavelengths not only in the near-infrared but also in the far-infrared range.

From a material research point of view, improvements in SiGe heteroepitaxy continue to drive research, e.g., on quantum cascade laser structures and the use of non-linear effects but also have potential applications, e.g., in realizing quantum-well infrared photodetectors on Si. Here, an ongoing refinement in nanoscale control over interface roughness and defects can be expected to lead to further improvements in device characteristics. Material engineering on the nanoscale has, so far, led to the investigation of light emission from defect-engineered nanostructures but also to the realization of hexagonal SiGe for light emission. Developing these disruptive concepts further, however, will most likely require significant advances in material growth that are accompanied by advanced modeling and simulation strategies. It should be noted here that advances in SiGe material research are not exclusively related to advances in SiGe heteroepitaxy as research results in Ge-based semiconductor plasmonics show, which rely on significant improvements in doping.

Light–matter interaction cannot only be engineered by growth but also by lateral nanostructuring. Such nanostructuring is at the base of a broad range of current research directions encompassing plasmonics as well as optical metasurfaces. SiGe devices based on or incorporating metallic or dielectric nanoparticles, e.g., for wavelength and polarization filtering have possible applications in miniaturized spectrometers or hyperspectral imaging. Here, device integration of such nanostructures still poses challenges, reliable contacting being one of them.

From an application point of view, the cost-effective realization of device concepts on the Si platform continues to be one of the main drivers for SiGe based photonic device research. On-chip biosensors, for example, have potential applications ranging from point-of-care solutions to environmental analysis. However, their widespread application is contingent on the availability of low-cost devices. Research on compact refractive index sensors, for example, has already made important steps toward realization based on industrial technology. For other devices, such as quantum cascade lasers, single-photon avalanche diodes, and quantum well infrared photodetectors, for example, commercial solutions are exclusively based on III–V semiconductors at this time. Here, the main challenge resides not only in realizing such device concepts on the SiGe platform but also in demonstrating device efficiencies that are sufficient for application needs. In other areas, the properties of group-IV semiconductors, in particular, Ge, are already advantageous for the desired applications, one example being low losses in the mid-infrared as a prerequisite for waveguiding.

The investigation of new materials and their integration on Si continues to bring forward innovations in SP as demonstrated by recent successes in realizing Sn-based light sources on Si or in using ferroelectric oxides for electro-optic modulators. This perspective focuses on selected SP integrated devices that can be realized based solely on Ge/SiGe/Si heterostructures of different complexity. With this, our intention is to convince the reader that the SiGe material system will continue to provide a platform for an increasingly broad range of applications, for which a foundry-based technological realization, most importantly, is within close reach.

ACKNOWLEDGMENTS

L.A.F. acknowledges funding from the Federal Ministry of Education and Research of Germany (BMBF) within the iCampus Cottbus project (Grant No. 16ES1128K). M.B. acknowledges funding from the Austrian Science Fund (FWF) (Grant No. Y1238-N36) and the Linz Institute of Technology (LIT) (Grant No. LIT-2019-7-SEE-114). D.J.P. acknowledges the support of the Royal Academy of Engineering Chair (Grant No. CiET2021/123).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1 S. Y. Siew, B. Li, F. Gao, H. Y. Zheng, W. Zhang, P. Guo, S. W. Xie, A. Song, B. Dong, L. W. Luo, C. Li, X. Luo, and G.-Q. Lo, J. Lightwave Technol. 39, 4374 (2021).
2 W. Bogaerts and L. Chrostowski, Laser Photonics Rev. 12, 1700237 (2018).
3 L. Chrostowski and M. Hochberg, Silicon Photonics Design: From Devices to Systems (Cambridge University Press, Cambridge, UK, 2015).
4 See http://www.yole.fr/Si_Photonics_Datacom_Sensing.aspx for a report on the silicon photonics market; accessed 18 April 2022.
5 R. Soref, IEEE J. Sel. Top. Quantum Electron. 12, 1678 (2006).
6 G. T. Reed and A. P. Knights, Silicon Photonics: An Introduction (John Wiley & Sons New York, 2004).
7 See http://www.scopus.com for Analysis performed on the SCOPUS database; accessed 15 January 2022.
8 J. S. Kilby, Proc. IEEE 88, 109 (2000).
9 D. J. Paul, Semicond. Sci. Technol. 19, R75 (2004).
10 T. E. Pinquet et al., “Monolithically integrated high-speed CMOS photonic transceivers,” in 2008 5th IEEE International Conference on Group IV Photonics (IEEE, 2008), p. 362.
11 C. Gunn, G. Masini, J. Witzens, and G. Capellini, Electrochem. Soc. Trans. 3, 17 (2006).
12 See https://europractice-ic.com/technologies/photonics/ for examples for Multi-Project-Wafer (MPW) services; accessed 18 April 2022.
13 G. Scappucci, C. Kluseffil, F. A. Zwanenburg, D. Lous, M. Myronov, J. J. Zhang, S. De Franceschi, G. Katsaros, and M. Veldhorst, Nat. Rev. Mater. 6, 926 (2021).
14 P. Bottegoni, M. Celebrazio, M. Bollani, P. Biagioni, G. Isella, F. Ciccare, and M. Finazzi, Nat. Mater. 13, 790 (2014).
15 R. Yokogawa, H. Takeuchi, Y. Arai, I. Yonenaga, M. Tomita, H. Uchiyama, T. Watanabe, and A. Ogura, Appl. Phys. Lett. 116, 242104 (2020).
16 See http://www.itrs2.net/itrs-reports.html for the technical documents of the International Technology Roadmap for Semiconductors (ITRS); accessed 18 April 2022.
17 K. Tani, T. Okumura, K. Oda, M. Deura, and T. Ido, Opt. Express 29, 28021 (2021).
18 C. Cornet, Y. Léger, and C. Robert, Integrated Lasers on Silicon (ISTE Press, London, UK, 2016).
19 M. Brehm, “Light-emission from ion-implanted group-IV nanostructures,” in Silicon Photonics IV: Innovative Frontiers, Springer Series Topics in Applied Physics Vol. 139, edited by D. J. Lockwood and L. Pavesi (Springer Verlag, 2021), pp. 67ff.
B. Turker, H. Guner, S. Ayas, O. O. Ekiz, H. Acar, M. O. Guler, and A. Dána, Lab Chip 11, 282 (2011).

S. Patkovsky and M. Meunier, Ann. Phys. 525, 431 (2013).

M. Perino, E. Pasqualto, A. De Toni, F. Garoli, M. Scaramuzza, P. Zilio, T. Ongarello, A. Paccagnella, and F. Romanzoni, Appl. Opt. 53, 5969 (2014).

D. Sammito, D. De Salvador, P. Zilio, G. Biasiol, T. Ongarello, M. Massari, G. Ruoffato, M. Morpurgo, D. Silvestri, G. Maggioni, G. Bovo, M. Gaio, and F. Romanzoni, Nanoscale 6, 1390 (2014).

L. Augel, Y. Kawaguchi, S. Bechler, R. Körner, I. Schulze, H. Uchida, and I. A. Fischer, ACS Photonics 5, 4586 (2018).

Y. Ajiki, T. Kan, K. Matsumoto, and I. Shimoyama, Appl. Phys. Express 11, 022001 (2018).

L. Laplatine, O. A’Mrayat, E. Luan, C. Fang, S. Rezaiezadeh, D. M. Ratner, K. Cheung, Y. Dattner, and L. Chrostowski, Proc. SPIE 10061, 100610I (2017).

L. Augel, F. Berkmann, D. Latta, I. A. Fischer, S. Bechler, Y. Elogail, K. Kostecki, K. Potjé-Kamloth, and I. Schulze, Microfluid. Nanofluid. 21, 169 (2017).

Grand View Research, Global Gas Sensor Market Size and Share Report 2021-2028, 2021, https://www.grandviewresearch.com/industry-analysis/gas-sensors-market.

J. Chou, Hazardous Gas Monitors: A Practical Guide to Selection, Operation, and Applications (McGraw-Hill, 2000).

I. Soler Penades, A. Sánchez-Postigo, M. Nedeljkovic, A. Ortega-Monux, J. G. Wanguemert-Pérez, Y. Xu, R. Halir, Z. Qu, A. Z. Khokhar, Osman, W. Cao, C. G. Littlejohns, P. Cheben, I. Molina-Fernández, and G. Z. Mashanovich, Opt. Lett. 43, 795 (2018).

A. Sanchez-Postigo, A. Ortega-Monux, J. Soler Penades, A. Osman, M. Nedeljkovic, Z. Qu, Y. Wu, I. Molina-Fernandez, P. Cheben, G. Z. Mashanovich, and J. Gonzalo Wanguemert-Perez, Opt. Express 29, 16867 (2021).

A. Osman, M. Nedeljkovic, J. Soler Penades, Y. Wu, Z. Qu, A. Z. Khokhar, K. Deb Nath, and G. Z. Mashanovich, Opt. Lett. 43, 5997 (2018).

L. He, Y. Guo, Z. Han, K. Wada, L. C. Kimerling, J. Michel, A. M. Agarwal, G. Li, and L. Zhang, Opt. Lett. 42, 3454 (2017).

R. Soref, Nat. Photonics 4, 495–497 (2010).

M. Brun, P. Labeye, G. Grand, J.-M. Hartmann, F. Boulila, M. Carras, and S. Nicoletti, Opt. Express 22, 508 (2014).

Y.-C. Chang, V. Paeder, L. Hvozdara, J.-M. Hartmann, and H. P. Herzig, Opt. Lett. 37, 2883 (2012).

M. Nedeljkovic, J. S. Penades, C. J. Mitchell, A. Z. Khokhar, S. Stankovic, T. D. Bucio, C. G. Littlejohns, F. Y. Gardes, and G. Z. Mashanovich, IEEE Photonics Technol. Lett. 27, 1049 (2015).

K. Gallagher, R. W. Millar, U. Griškevičiūtė, L. Baldassarre, M. Sorel, M. Ortolani, and D. J. Paul, Opt. Express 26, 25667 (2018).

A. Malik, E. J. Stanton, J. Liu, A. Spott, and J. E. Bowers, IEEE J. Sel. Top. Quantum Electron. 24, 1 (2018).

M. Nedeljkovic, J. S. Penades, V. Mittal, G. S. Murugan, A. Z. Khokhar, C. Littlejohns, L. G. Carpenter, C. B. E. Gawith, J. S. Wilkinson, and G. Z. Mashanovich, Opt. Express 25, 27431 (2017).

J. Frigerio, A. Ballabio, G. Isella, E. Sakat, G. Pellegrini, P. Biagioni, M. Bollani, E. Napolitani, C. Manganelli, M. Virgilio, A. Grupp, M. P. Fischer, D. Brida, K. Gallagher, D. J. Paul, L. Baldassarre, P. Calvani, V. Giliberti, A. Nucara, and M. Ortolani, Phys. Rev. B 94, 085202 (2016).

C. A. Charvarin, E. Hardt, S. Gruessing, O. Sbrizzi, I. Costina, D. Spitzer, W. Seifert, W. Klesse, C. L. Manganelli, C. You, J. Fleisch, J. Pfeiler, M. Missori, L. Baldassarre, B. Witzigmann, and G. Capellini, Opt. Express 29, 7680 (2021).

G. Pellegrini, L. Baldassarre, V. Giliberti, J. Frigerio, K. Gallagher, D. J. Paul, G. Isella, M. Ortolani, and P. Biagioni, ACS Photonics 5, 3601 (2018).

K. Gallagher, R. W. Millar, V. Giliberti, E. Calandrini, L. Baldassarre, J. Frigerio, A. Ballabio, E. Sakat, G. Pellegrini, G. I. M. Ortolani, P. Biagioni, and D. J. Paul, in 2017 IEEE 14th International Conference on Group IV Photonics (IEEE, 2017), pp. 3–4.

V. Vakarin, W. N. Ye, J. M. Ramirez, Q. Liu, J. Frigerio, A. Ballabio, G. Isella, L. Vivien, C. Alonso-Ramos, P. Cheben, and D. Marris-Morini, Opt. Express 27, 9838 (2019).

K. Gallagher, R. W. Millar, U. Griškevičiūtė, M. Sinclair, M. Sorel, L. Baldassarre, M. Ortolani, R. Soref, and D. J. Paul, APL Photonics 5, 026102 (2020).

A. Malik, S. Dwivedi, I. Van Landschoot, M. Muneeb, Y. Shimura, G. Lepage, J. Van Campenhout, W. Vanopstal, R. Vanopstal, L. Goo, and P. Roelkens, Opt. Express 22, 28479 (2014).

M. Montesinos-Ballester, Q. Liu, V. Vakarin, J. M. Ramirez, C. Alonso-Ramos, X. L. Roux, J. Frigerio, A. Ballabio, E. Talamas, L. Vivien, G. Isella, and D. Marris-Morini, Sci. Rep. 9, 14633 (2019).

E. Gordon, L. S. Rothman, C. Hill et al., “The HITRAN2016 molecular spectroscopic database,” J. Quant. Specros. Radiat. Transfer 203, 3 (2017).

M. Fischer, A. Riede, K. Gallagher, J. Frigerio, G. Pellegrini, M. Ortolani, D. J. Paul, G. Isella, A. Leitenstorfer, P. Biagioni, and D. Brida, Light Sci. Appl. 7, 106 (2018).

M. A. Etlibib, L. Xu, A. Bogris, A. Kapsalis, M. Belal, E. Lorent, P. Labeye, S. Nicoletti, K. Hammami, D. Syvridis, D. P. Shepherd, J. H. V. Price, D. J. Richardson, and P. Petroopoulos, Opt. Lett. 40, 4118 (2015).

M. Montesinos-Ballester, V. Vakarin, Q. Liu, X. Le Roux, J. Frigerio, A. Ballabio, A. Barzagh, C. Alonso-Ramos, L. Vivien, G. Isella, and D. Marris-Morini, Opt. Express 28, 12771 (2020).