Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip [Invited]

Ananth Z. Subramanian,1,2,* Eva Ryckeboer,1,2 Ashim Dhakal,1,2 Frédéric Peyskens,1,2 Aditya Malik,1,2 Bart Kuyken,1,2 Haolan Zhao,1,2 Shibnath Pathak,1,2 Alfonso Ruocco,1,2 Andreas De Groote,1,2 Pieter Wuytens,1,2* Daan Martens,1,2 François Leo,1,2 Weiqiang Xie,1,2 Utsav Deepak Dave,1,2 Muhammad Muneeb,1,2 Pol Van Dorpe,1,2 Joris Van Campenhout,1,2 Wim Bogaerts,1,2 Peter Bienstman,1,2 Nicolas Le Thomas,1,2 Dries Van Thourhout,1,2 Zeger Hens,1,2 Gunther Roelkens,1,2 and Roel Baets1,2

1Photonics Research Group INTEC Department, Ghent University-imec, Ghent 9000, Belgium
2Centre for Nano and Biophotonics, Ghent University, Ghent, Belgium
3University of California, Davis, California 95616, USA
4Department of Molecular Biotechnology, Ghent University, Ghent, Belgium
5imec, Kapeldreef 75, B-3001 Leuven, Belgium
6Physics and Chemistry of Nanostructures, Ghent University, B-9000 Ghent, Belgium
*Corresponding author: ananth.subramanian@intec.ugent.be

Received May 21, 2015; revised July 19, 2015; accepted July 21, 2015; posted July 23, 2015 (Doc. ID 241314); published August 28, 2015

There is a rapidly growing demand to use silicon and silicon nitride (Si₃N₄) integrated photonics for sensing applications, ranging from refractive index to spectroscopic sensing. By making use of advanced CMOS technology, complex miniaturized circuits can be easily realized on a large scale and at a low cost covering visible to mid-IR wavelengths. In this paper we present our recent work on the development of silicon and Si₃N₄-based photonic integrated circuits for various spectroscopic sensing applications. We report our findings on waveguide-based absorption and Raman and surface enhanced Raman spectroscopy. Finally, we report on-chip spectrometers and on-chip broadband light sources covering very near-IR to mid-IR wavelengths to realize fully integrated spectroscopic systems on a chip. © 2015 Chinese Laser Press

OCIS codes: (250.5300) Photonic integrated circuits; (130.3120) Integrated optics devices; (130.6010) Sensors; (190.4390) Nonlinear optics, integrated optics; (300.0300) Spectroscopy; (230.5590) Quantum-well, -wire and -dot devices.

http://dx.doi.org/10.1364/PRJ.3.000B47

1. INTRODUCTION

Optical spectroscopy is a technique to detect the presence and/or the concentration of a substance via its interaction with light. The interaction of light with the substance can lead to absorption, emission, and scattering of light, which, when described as a function of wavelength, is termed as a spectrum. The nature of the spectrum and the amount of information it is dependent not only on the chemical composition, temperature, and physical state of the substance, but also on the surrounding environment and the wavelength region being probed. The key advantage of spectroscopic techniques over other techniques lies in the selectivity they offer by probing the “fingerprint spectrum” of the substance. The traditional spectroscopic systems rely on expensive and bulky instrumentation prohibiting their widespread use especially in outside-of-the-lab environments. But in recent years, there has been a surge in demand for hand-held devices that can be used in the field and that are capable of accurate, sensitive, and in situ spectroscopic detection of a variety of substances. Such devices can serve applications ranging from physics and chemistry to biology and environmental sciences. This has led to the development of a range of optical biosensors based on emission, absorption, and refractometry. Although optical sensing can be performed in different platforms such as optical fibers [1], waveguides [2], and bulk optics [3], an integrated photonics approach has significant advantages. First, integrated photonics enables miniaturization and integration of active and passive optical components. Second, the integration leads to devices that are robust, rugged, reliable, mass producible, and low-cost. Finally, the integration allows for an opportunity for massive parallelism and multiplexing.

Evanescent wave sensing uses an optical waveguide or a fiber as the sensing element. Most of the light is confined in the core of the waveguide, but depending on the index contrast of the waveguide, a significant part of the light (the evanescent mode tail) can extend into the upper and side cladding of the waveguide. For evanescent sensing, the analyte can be sensed by bringing the analyte in direct contact with the core of the waveguide or by first depositing a receptor layer on the core of the waveguide and then bringing the analyte in contact with the receptor layer. In both scenarios the guided mode of the waveguide interacts with the analyte. This can then be correlated to the concentration or to the physical/chemical composition of the analyte. In refractive-index-based sensing, the presence of the analyte induces a change of the effective refractive index of the guided mode, which can be detected easily through interferometric methods. Examples include sensors based on fiber Bragg
gratings [4] and ring resonators [5]. The method only becomes specific for particular analytes if there is specific surface chemistry involved. In spectroscopic sensing, the output light has a spectrum that is a direct fingerprint for the analyte and therefore allows unambiguous identification of the analyte. Examples include absorption spectroscopy [6] and Raman spectroscopy [7].

Of all the possible photonic integrated circuit (PIC) platforms that have been reported, those that are based on high index contrast (HIC) have gained the most attention in the past decade. HIC waveguides—with a high difference between the refractive index of the core and that of the cladding—allow making strongly miniaturized circuits, because one can make tight bends with very low loss. Furthermore HIC waveguides have tightly confined guided modes, implying that the electric field strength of the evanescent tail is very strong for a given optical power, leading to intense interaction between light and the analyte.

The most prominent example of a HIC waveguide platform is definitely the silicon photonics waveguide platform. In this platform silicon-on-insulator (SOI) wafers are used to form waveguides with a core of silicon (with a refractive index of 3.5) and a cladding of silica (with a refractive index of 1.45) or of air. In recent years silicon photonics has reached a considerable level of maturity by leveraging the same technologies and tools that the microelectronics industry uses in a CMOS fab, to fabricate high-quality PICs on 200 and 300 mm SOI wafers [8,9] with a high level of process control and yield. Given the HIC, very tight bend radii (<5 μm) are possible. This allows for an extremely small footprint for PICs designed in silicon photonics. A potential downside of HIC waveguides is the fact that any surface or sidewall roughness will induce large scattering losses. Therefore, HIC only makes sense if the technology is capable of providing very smooth core-cladding interfaces. In the case of silicon photonics, low-loss waveguides (<1 dB/cm) have been demonstrated at 1550 nm wavelength [10]. Apart from that, high-quality PICs including passive components such as high-efficiency fiber-to-chip grating couplers, MMIs, and ring resonators, active components such as modulators, and hybrid III-V-on-silicon light sources have been demonstrated [11–14]. The major thrust for developing silicon photonics came from telecom and datacom applications. But increasingly, silicon photonics is also finding important applications in sensing and biosensing. Silicon is transparent between 2.1 and 8 μm, but the underlying oxide begins to absorb light from about 4 μm onwards. As a result SOI waveguides are relevant only in the 1.1–4 μm wavelength range.

While silicon photonics takes immense advantage of the capabilities of CMOS technology, it also suffers from the high cost of accessing CMOS fabs, especially at the research and low-volume manufacturing level. In order to alleviate this problem multi-project-wafer (MPW) approaches such as ePIXfab/Europractice [15] have been set up. The MPWs are built on a cost-sharing model so as to allow easy and low-cost access to standardized processes for rapid and proof-of-concept prototyping.

In order to extend silicon photonics to shorter wavelengths (below 1.1 μm)—in other words to make use of CMOS tools at these shorter wavelengths while maintaining a sufficiently HIC—silicon nitride (Si₃N₄) is probably the most widely used material. It has a refractive index of 2.0 and allows for bend radii down to about 40 μm. Although Si₃N₄ is transparent from 0.4 to 5 μm, it is particularly important for the visible and very near-IR wavelengths (<1 μm). This is because this wavelength window covers the therapeutic window (0.75–0.93 μm) for biological media leading to minimal photo-damage to cells and negligible water absorption. Furthermore in this wavelength range there is a large availability of light sources with high performance and low cost as well as high-performance silicon-based detectors. Recently we demonstrated low-loss (<1.5 dB/cm) nitride wires for the visible—very near-IR wavelength range [16].

At the other extreme, the mid-IR region (4 to 10 μm) also has a huge significance for spectroscopy. For example, most gases and molecules have very strong and fingerprint absorption peaks in this wavelength range that can be used for unambiguous sensing and identification [17]. As silicon itself is transparent up to 8 μm and very low-loss until 10 μm, to extend the advantages of silicon photonics to these longer wavelengths the most obvious choice would be to replace the underlying oxide in SOI by a mid-IR transparent material. Several candidates have been proposed in the literature to tackle this issue, such as free standing Si, Si-on-Si₃N₄, chalcogenide-based waveguides, and Si-on-sapphire waveguides [18]. However, all these waveguide platforms either have a limited transparency window or have a fabrication process containing many failure steps, which makes mass manufacturing difficult. Germanium-on-silicon offers a waveguide platform that 1) is widely transparent from 2 to 14 μm (appropriate waveguide design can ensure high mode confinement and large bend radius that allows us to overcome higher loss due to silicon underneath at wavelengths >10 μm), 2) has a straightforward fabrication scheme, and 3) can be mass manufactured in CMOS or MEMS foundries. Several passive and active functionalities [19–21] have been demonstrated on this waveguide platform.

Different forms of spectroscopic methods exist for sensing application, namely absorption, emission, and Raman scattering spectroscopy. The difference lies in the way the spectrum is collected and in the information contained in the spectrum itself. For any spectroscopic method, the main elements used are a light source (a broadband source or a laser), a probe, a filter, a spectrometer, and a detector. Various schemes used in spectroscopic systems are shown in Fig. 1. In absorption

![Fig. 1. Schematic of different spectroscopic methods and configurations: absorption spectroscopy (top), emission (fluorescence) spectroscopy (middle), and Raman spectroscopy (bottom).](image-url)
spectroscopy, a broadband light source can be used to probe the analyte, and the transmitted light passes through a spectrometer, which disperses the light into different spectral components that is detected by a detector array.

The transmission is a measure of extinction induced by the analyte in a particular wavelength range. When a broadband light source is used in conjunction with a tunable filter or when a tunable laser is used, the spectrometer is not necessary as shown in Fig. 1 (top). In emission spectroscopy such as fluorescence spectroscopy, a laser is used to probe an analyte and the resulting output consists of an emitted spectrum centered at a new wavelength. This is then detected using a spectrometer and a detector array as shown in Fig. 1 (middle).

Lastly for Raman spectroscopy, a laser at the pump wavelength induces scattering of light from the analyte molecules. The scattering is not only at the pump wavelength (Rayleigh) but also at longer (Stokes) and shorter (anti-Stokes) wavelengths that are characteristic of a particular molecular bond present in the analyte. The scattered light passes through a filter to block the pump wavelength, and the remaining light passes through a spectrometer and is detected by a detector array. Raman scattering is an inherently weak process (typically only one out of every 10^7 photons that interact with a molecule are scattered) and can be further enhanced by several orders of magnitude by using a probe functionalized with metallic nanostructures exploiting plasmonic resonances [7]. At resonance, the electric field at the metal interface is enhanced by several orders of magnitude, which boosts the Raman signals, a technique known as surface enhanced Raman scattering (SERS). For SERS, the measurement scheme is the same as that for conventional Raman spectroscopy except that the analyte is in contact with a substrate that is functionalized with metallic nanostructures.

In this paper, various theoretical and technological aspects of chip-based absorption and Raman spectroscopy are discussed. Section 2 is devoted to the waveguide-based evanescent sensing techniques and in particular discusses the theoretical principles and optimization approaches for absorption and Raman spectroscopy—including SERS—using waveguides. This is illustrated by examples. In order to realize spectroscopic systems on a chip, different elements (as shown in Fig. 1) need to be realized on a chip. Sections 3 and 4 are dedicated to two of the important building blocks: spectrometers and light sources. In Section 3, the design and performance of on-chip spectrometers covering different wavelength ranges and technology platforms are discussed. Section 4 discusses the technological advances made in integrating broadband light sources covering the visible to IR.

2. WAVEGUIDE-BASED EVANESCENT SENSING

A. Absorption Spectroscopy

Each material shows a signature spectrum whose strength and specificity depend on its chemical composition. Every molecule, owing to its vibrational and rotation energy levels, possesses a very strong fingerprint absorption spectrum in the mid-IR region [17]. This can be utilized for analytical chemical sensing. But accessing those wavelengths remains a challenge due to the lack of cheap sources, detectors, and optical fibers. One way to overcome this problem is to access their overtones, which, although weaker, lie in the near-IR region. In this wavelength range SOI technology, light sources, detectors, and optical fibers are very well developed. To probe the analyte through the waveguide, the evanescent field that typically extends up to ~100 nm into the cladding layer is utilized. The evanescent field is attenuated due to absorption in the cladding. This process is well described by Beer–Lambert’s law: \( T = e^{-\left(\alpha_s + \alpha_w\right) L} = 10^{-A}. \) This law relates the waveguide transmission \( T \), after propagating a distance \( L \), taking into account the losses in both waveguide (\( \alpha_w \)) and cladding (\( \alpha_s \)). \( 'A' \) is the absorbance. The fill factor \( \Gamma \) is the fraction of the optical mode that extends into the cladding. It is highly dependent on the waveguide geometry, but can be easily calculated. It is usually a small number. This also means that \( L \) needs to be large to have an appreciable change in transmission. In order to reduce the footprint while maintaining long interaction length and low losses, HIC waveguides can be routed into a very compact spiral shape.

Typically in such a sensing scheme, measurements are performed with and without the analyte using two identical waveguide sensors whilst keeping all other conditions similar. Then the difference in the power levels between the two measurements yields the absorbance by the analyte. To optimize the sensitivity, \( L \) needs to be optimized. If the interaction length is chosen too short, the absorption features are not evident whereas a too-long interaction length decreases the detected power too strongly. The optimal interaction length \( L_{opt} \) is given by \( L_{opt} = 1/(\Gamma \alpha_s + \alpha_w) \) [22]. But this optimal length is wavelength dependent through the confinement factor and the loss terms. For optimum operation a low-loss waveguide with a large confinement factor in the cladding is required. Given the higher confinement factor in the cladding and the lower impact of sidewall roughness, the transverse magnetic (TM) mode is better suited for evanescent sensing than the transverse electric (TE) mode. The waveguide losses can also be reduced by using a rib waveguide (partially etched) instead of a standard wire waveguide.

Using a lab-on-a-chip absorption spectroscopy approach we successfully demonstrated the in vitro detection of glucose in an aqueous solution in the clinically relevant concentration ranges [22]. Glucose has absorption features over a wide wavelength range, but in order to exploit the well-developed silicon photonics technology, sources, and detectors, the design was optimized for sensing in the telecom wavelength range (1540–1610 nm). The probe was a SOI spiral-shaped rib waveguide chip. A rib waveguide structure was chosen because of the reduced modal overlap with the sidewalls, lower propagation losses, and lower distributed backscattering [22]. Distributed backscattering leads to an unwanted interference pattern in the transmission spectrum that is sensitive to environmental parameters and that can mask a weak absorption signature. Figure 2(a) depicts the complete SOI design architecture for glucose sensing. The SOI chip comprises a partially etched 220 nm thick Si waveguide layer on a 2 μm buried oxide layer [Fig. 2(b)]. The optical circuit comprises both a signal and a reference sensor. The signal spiral is immersed in the glucose solutions, while a reference path accounts for the possible noise sources such as source power fluctuations and temperature variations. The photonic chip has a finish of a planarized 1.35 μm thick oxide cladding at the top. To allow evanescent sensing, the top oxide above
Since water absorption dominates the absorption spectra of glucose solutions with different glucose concentrations by dividing the measured glucose absorption spectra from the water absorption spectra. For the lowest concentration of 1 mM was chosen to show the detection limit. From the signal and reference spectral sweeps, the valve is switched between deionized water and a glucose solution. The lowest concentration of 1, 5.5, 16, 24, and 36 mM are used in the glucose measurement experiments. The spiral signal transmission was measured simultaneously. The fiber array is glued to the optofluidic chip for mechanical stability. The final fabricated device is shown in Fig. 2(c). The inlet of the fluidic channel is connected to a microfluidic pump loaded with the syringes to pump the glucose solutions of different concentrations at a steady rate of 9 μL/min. For the optical readout, a fiber array was used to couple light in and out of the optical chip and both the signal and the reference spiral transmission were measured simultaneously. The fiber array is glued to the optofluidic chip for mechanical stability. Polarization-maintaining fiber was used throughout the whole optical path to reduce the detected power fluctuations caused by variations in mode coupling in the fiber, which are potentially different from signal to reference path. Finally, the sample is placed onto a temperature controller for thermal stability. Five different glucose concentrations (1, 5.5, 16, 24, and 36 mM) are used in the glucose measurement experiments. The spiral signal transmission spectra are continuously acquired, and after every 16 spectral sweeps, the valve is switched between deionized water and a glucose solution. The lowest concentration of 1 mM was chosen to show the detection limit. From the signal transmission we obtain the glucose absorption spectra for different glucose concentrations by dividing the measured glucose solution spectra with the water absorption spectra [22]. Since water absorption dominates the absorption spectrum of the aqueous glucose solution, it needs to be eliminated from the measured signal to reveal the true glucose spectrum. Water absorption curves are obtained from the water applied before and after each glucose solution. In addition, various glucose absorption spectra are also averaged to obtain a high SNR.

The result is plotted in Fig. 3(a). When the glucose concentration is increased, the glucose absorption is higher and the transmission decreases accordingly. The replicate measurements (1 and 5.5 mM) are also shown in the Fig. 3(a). Due to the weak glucose absorption, the transmission is expected to be linearly dependent on the glucose concentration (\(T = e^{-\alpha_{\text{glucose}} C_{\text{glucose}}} \approx 1 - \alpha_{\text{glucose}} C_{\text{glucose}}\)). Therefore, we derive a linear regression model and do a least-squares fitting between the measured absorption spectra for every applied concentration and for a range of possible glucose concentrations between 0 and 50 mM. The predicted glucose concentration is the one for which the least-squares sum is minimal. The result is plotted in Fig. 3(b). The linear trend is clearly visible. The error-of-fitting for our model using the predictions of all the six measurements except 36 mM is 1.14 mM. Based on the predicted glucose concentrations, we also calculated the theoretically expected absorption spectra based on the reference data from [23]. These are plotted as gray curves in Fig. 3(a). We can see that our measurements are in good agreement with the theoretically expected shape.

### B. Raman Spectroscopy

The aspects of spectroscopic sensing that are really attractive are specificity and label-free sensing. Raman spectroscopy gains significance as a tool for providing label-free and unambiguous detection of chemical and biological molecules by directly probing the molecular vibrations. In comparison with absorption spectroscopy, Raman spectroscopy allows working in a wavelength range different from those corresponding with the vibrational absorption bands. This, for example, allows avoiding the influence of water absorption. However, this technique is limited in its use because of the extremely small cross section of the Raman scattering process, hence requiring a high-power pump laser, sensitive detectors, and long integration time. Traditionally, Raman microscopes are well suited for extracting information from a variety of molecules, but there are several advantages to moving towards on-chip Raman spectroscopy. First, the integrated photonics approach provides an opportunity to integrate crucial functionalities such as filtering, multiplexing, light sources, and detectors. Second, there are performance-related benefits—improving the overall efficiency of Raman spectroscopy through HIC waveguides with optimal length as will be discussed later. Recently we have demonstrated a lab-on-a-chip approach to Raman spectroscopy based on HIC...
The excitation of a dipole also depends on the modal overlap of the dipole field and the modal field at the dipole location. The excitation of a dipole also depends on the modal field. From Fig. 4 one can conclude that the overall efficiency of the absorption and subsequent reemission (or inelastic scattering) by a particle at the core-cladding interface of a slab waveguide depends quadratically on the index contrast and is optimal for a slab thickness equal to half the wavelength. The coupling is evidently dependent on the core index and thickness, the polarization of the mode, and the orientation of the dipole, and follows the overlap of the dipole field and the modal field at the dipole location. The excitation of a dipole also depends on the modal field.

In photonic channel waveguides, the conversion from a guided excitation to a guided emission can be expressed by a parameter we call specific conversion efficiency \( \eta_0 \), which is defined as the ratio of the collected emission to the pump power. Equation (2) allows us to quantify and measure the efficiency of the collected emission for different channel waveguide geometries. \( \eta_0 \) depends on the dielectric environment that underpins the role of index contrast and waveguide geometry for efficient excitation and collection.

For lossy waveguides of length \( L \), Eq. (1) can be remodeled to define a function \( \zeta(L) \) such that (see [24] for details)

\[
\zeta(L) \equiv \frac{P_{\text{col}}(L)}{P_{\text{in}}(L)} = \frac{\rho \sigma \eta_0}{\rho_0} \left[ \frac{e^{\alpha L} - 1}{\Delta \sigma} \right].
\]

Fig. 4. Power coupled to the fundamental TE and TM modes of a slab waveguide normalized to total free-space emission in all directions, for Si (left) and Si-N4 (right) cores on silica substrate and air cladding (scattering) by a particle at the core. The excitation of a dipole also depends on the modal overlap of the dipole field and the modal field at the dipole location. The excitation of a dipole also depends on the modal field.

For lossy waveguides of length \( L \), Eq. (1) can be remodeled to define a function \( \zeta(L) \) such that (see [24] for details)

\[
P_{\text{col}}/P_{\text{in}} = \rho \sigma \eta_0 L.
\]

Fig. 5. Schematic of experimental setup. H, half-waveplate; BF, band-pass filter; AL, aspheric lens; O, objective; EF, long-pass edge filter; PM, parabolic mirror; SMF, single-mode fiber.

\[
\zeta(L) \equiv \frac{P_{\text{col}}(L)}{P_{\text{in}}(L)} = \frac{\rho \sigma \eta_0}{\rho_0} \left[ \frac{e^{\alpha L} - 1}{\Delta \sigma} \right].
\]

\( \rho_{\sigma} \) is the transmitted pump power, and \( \Delta \sigma \) is the difference in wavenumber. Reproduced from [25].

For several waveguide geometries and polarization of the modes, Figure 5 shows \( \eta_0 \) calculated for Si-N4 (thickness, \( h = 220 \text{ nm} \)) strip and slot waveguides (slot width, \( s = 150 \text{ nm} \)) for various widths. The waveguide width \( w \) for slots includes the slot width. In Fig. 5 we clearly see that \( \eta_0 \) decreases as the waveguide width is increased, because the confinement of the mode in the core increases, thereby reducing the interaction volume in the upper cladding region where the analyte molecules are located. Due to the enhanced field in the slot, we also see as much as a 6.4-fold increment in \( \eta_0 \) for slot waveguides compared to strip waveguides of the same width (for example, for \( w = 700 \text{ nm} \)).

For experimental verification, isopropyl alcohol (IPA) droplets were applied on top of the Si-N4 spiral waveguide, and a pump at 785 nm was coupled into the waveguide. The transmission was recorded using a Raman spectrometer after filtering the pump at the output of the waveguide. The schematic of the experimental setup is shown in Fig. 5.

In Fig. 6(b), the experimentally recorded points for strip and slot waveguides with different spiral lengths follow the trend predicted by Eq. (2) nicely. Figure 7 shows the background observed from the waveguide before application of IPA and the Raman spectrum of IPA observed on it after application of IPA. The background was removed by subtracting the truncated polynomial fit to the background [24]. The IPA peak corresponding to the C-C-O vibration at 819/cm wave-numbers was used to compare the values of \( \eta_0 \) for several waveguides.
A conventional Raman confocal system is optimized to collect light from a diffraction-limited spot so only the molecules that are present in the focal volume of the spectrometer contribute to the Raman efficiency. An increase in the input power will help, but beyond a certain point it will also lead to damage of the sample. In our waveguide-based approach, the molecules under study over the length of the waveguide are evanescently excited, and the corresponding Raman signal is also evanescently collected into a single guided waveguide mode. This provides the optimum scenario for achieving the best spectral resolution of the Raman spectrum. The length of the waveguide is the main asset of this configuration. Meanwhile, this technique further exploits the enhancement effects inherent to the HIC waveguide, which leads to at least two orders of magnitude higher signal than in confocal microscopes for a 1 cm long waveguide. Given the benefits of integrated photonics based spectroscopy and HIC waveguides, an on-chip based approach promises to be an excellent platform for Raman spectroscopy.

C. Integrated Nanoplasmonic Antennas for On-Chip SERS

The Raman signal can be further enhanced by several orders of magnitude in a metallic nanostructure due to the collective oscillation of conduction electrons (plasmons) upon light excitation (plasmon resonance). This can occur due to two phenomena: surface plasmon polaritons (SPPs) or localized surface plasmons (LSPs) [28]. SPPs are propagating plasmons at the interface between a metal and a dielectric, whereas LSPs are the localized plasmons at the edges of metallic (e.g., gold) nanostructures. In both cases the electric field at the metal–dielectric interface is enhanced. This eventually boosts Raman signals, an effect also known as SERS [29]. As compared to a typical evanescent wave (∼100 nm), SPPs extend further into the dielectric (∼150 nm); therefore they can probe further into the top cladding. LSPs, on the other hand, remain tightly confined to the surface (∼25 nm). The field enhancement that can be achieved with SPPs is generally much smaller (10–100 times) than that obtained from LSPs (∼10⁴). For probing molecules, stronger fields are more beneficial to enhance the very weak Raman signals they emit. Hence, the LSPs are better suited for probing individual molecules. The strong increase in the signal allows for extremely sensitive measurements down to single molecule level. The required nanoplasmonic structures can be easily integrated on top of the waveguides to combine the advantages of waveguides and SERS. Plasmon resonance occurs at the so-called plasmon resonance wavelength $\lambda_{\text{res}}$. The resonance wavelength can be tuned by varying the nanoparticle’s size, shape, composition, and dielectric surroundings [29]. Most of the research to date has focused on nanoplasmonic structures on top of general substrates or in solution, which are excited using bulky and expensive free-space optics. The generated signals from such substrates are also collected by the same optics. Recently there have been several research activities towards integrating metallic nanoparticles on top of PICs [30]. Integrated nanoplasmonic structures can be aligned with the desired polarization by proper design and fabrication and therefore allow for the optimal excitation of the plasmon resonance. Furthermore one can excite and collect the relevant SERS signals (e.g., from analytes near the metal surface) through the same waveguide mode. Such nanoplasmonic PICs provide a potentially low-cost and versatile alternative for the currently used plasmonic biosensing devices.

The 700–1000 nm wavelength region is particularly relevant for on-chip SERS since pumping in this near-IR window (e.g., at 785 nm) can reduce the unwanted fluorescence background and allow low-loss propagation of both pump and Stokes light in a Si₃N₄ waveguide [16]. To this end we investigated several antenna types patterned on top of a 220 nm high and 700 nm wide single-mode Si₃N₄ waveguide with an oxide under-cladding [Figs. 7(a) and 7(b)] [31]. On each waveguide, N identical gold antennas with a period $\Lambda$ are patterned using e-beam lithography. The thickness of gold for the antennas was fixed at 35 nm, whereas across different waveguides the geometrical parameters (height $H$, gap $\Delta$, and apex angle $\alpha$) of each antenna type (single rod, double rod, and bowtie antennas) are varied. All antennas are excited with a TE-polarized waveguide mode [see Fig. 8(a)]. The single antenna extinction $E(\lambda)$ is calculated through

$$E(\lambda) = (T_{\text{ref}}(\lambda) - T_{\text{ant}}(\lambda))/N.$$  

$T_{\text{ref}}(\lambda)$ is the modal transmission of the reference waveguide and $T_{\text{ant}}(\lambda)$ the transmission of the patterned waveguide (with

![Image](https://example.com/image.png)

Fig. 7. (Left) Raw spectrum measured from a 1.6 cm waveguide without IPA (blue) and with IPA (red). The spectra are shifted vertically for clarity. (Right) Extracted IPA Raman spectrum. Reproduced from [24].
In general, the geometrical changes are, however, in good agreement. The differences among the theoretical and experimental data can be attributed to the difference between the actual metal refractive index and the one used in the simulation, and to the error in measured geometrical parameters using SEM.

The simulated and experimental sensitivity $S = d\lambda_{res}/dH$ to the geometrical changes are, however, in good agreement. In general $\lambda_{res}$ red-shifts when the height $H$ (along the TE direction) increases, while $\lambda_{res}$ blue-shifts when the gap $\Delta$ increases. For the gapped antennas $S$ also depends on the gap. By increasing the apex angle $\alpha$ of a bowtie antenna it is possible to excite multipolar (bright and dark) plasmon resonances [31]. Furthermore it can be shown that, depending on the specific antenna geometry, the waveguide cross section can be tuned to optimize the interaction with the antenna [31]. From Fig. 8 it is clear that on-chip nanoplasmonic antennas with resonances in the near-IR region can be efficiently produced on the Si$_3$N$_4$ platform, which is of importance for the design of on-chip SERS-based sensors.

A further miniaturization of the photonic chips decorated with nanoplasmonic elements allows the use of these chips for intracellular applications. The strong interaction of light with the surrounding cellular environment in optical or plasmonic cavities provides new opportunities for real-time intracellular sensing. We recently demonstrated the use of Si$_3$N$_4$ chips coated with gold nanostructures for intracellular SERS [32]. These chips offer the advantage of having a fixed gold nano-pattern during live-cell experiments, thereby avoiding spatiotemporal dynamics often encountered when using colloidal gold nanoparticles for intracellular SERS. In a proof-of-concept experiment, gold nanoparticle coated microchips were used for the label-free, live-cell detection of extraneous molecules. Furthermore, micrometer-sized dielectric chips can be coated with top-down fabricated plasmonic nano-antennas. This way, tunable and reproducible SERS chips can be fabricated similar to extra-cellular SERS substrates. For example, we showed that coating microdisks with a nano-triangular pattern through nanosphere lithography (NSL) (Fig. 10) provides a cost-effective method for the fabrication of millions of identical chips for intracellular applications [33]. Because the photonic microchips are planar, the total cell volume consumed by the photonic chips is limited. Therefore, the photonic chips do not seem to have a strong influence on cell viability [34], although their diameter of a few micrometers is reasonably large when compared to typical mammalian cells.

![Fig. 9. Resonance wavelength $\lambda_{res}$ behavior as a function of the geometry. (a) Single rod ($W = 55$ nm), (b) double rod ($W = 55$ nm), and (c) bowtie antenna ($\alpha = 60^\circ$). Reproduced from [31].](image1)

![Fig. 10. SEM image of a Si$_3$N$_4$ chip coated with adsorbed gold nanoparticles (left) and NSL fabricated gold triangles (right). Reproduced from [32].](image2)
we demonstrated a significant improvement in both crosstalk (3 to 6 dB improvement) and cumulative crosstalk (1 to 5 dB improvement) depending on the channel spacing and the number of output channels [37]. For PCGs, a distributed feedback reflector (DBR) is used to have high-reflectivity grating facets over a wide wavelength range. The phase errors accumulate in the slab region of the PCG as the variations in optical path length can be largely attributed to local thickness variations of the slab region, which is the largest source of crosstalk in PCGs. Grating facets induce undirected scattering that is inherent to PCGs. Making them larger reduces the problems but increases the size of the device. DBRs used as reflectors also induce excess insertion loss. The linewidth of crosstalk in PCGs. Grating facets induce undirected scattering that is inherent to PCGs. Making them larger reduces the problems but increases the size of the device. DBRs used as reflectors also induce excess insertion loss. The linewidth of the waveguides in the straight sections. The footprint of the device was larger than for SOI because of the lower index contrast and a combination of lower index contrast and wider waveguides in Ge-on-Si, both necessitating a larger bend radius. In SOI, we have reported PCGs covering the C-band [36,43] and much longer wavelengths up to 3.8 μm [39], we used an AWG with five outputs and four time-multiplexed inputs. The inputs were positioned such that their spectral response at the outputs was shifted according to a Vernier scale. This multiplied the resolution of the device, and also removed the ‘dead zones’ between the AWG output channels. The device is shown in Fig. 13. The inputs are each equipped with a Mach–Zehnder modulator to ‘label’ the input in a unique way. In this example we did this by sequentially switching the inputs on and off, but it is also possible to use modulations on a unique frequency. We demonstrated this device as a wavelength meter with 12 pm accuracy, even though the channel spacing of the AWG outputs is 4 nm.

AWGs were also designed in Si$_3$N$_4$ [41] for very near-IR (900 nm) and in Ge-on-Si [42] for mid-IR (5 μm) wavelengths. Both of them showed performance that was comparable to the more mature SOI AWGs, with a CT of 20 dB and insertion loss <2.5 dB. While both of these platforms are not as mature as its SOI counterpart, a lot of design adaptations can simply be mimicked, in particular those dealing with the high sensitivity to phase errors caused by the HIC. These include the double etch step and gradually broadening waveguides in the apertures of the star coupler, as well as widening of the waveguides in the straight sections. The footprint of the device was larger than for SOI because of the lower index contrast in Si$_3$N$_4$ and a combination of lower index contrast and wider waveguides in Ge-on-Si, both necessitating a larger bend radius. In SOI, we have reported PCGs covering the C-band [36,43] and much longer wavelengths up to 3.8 μm [44]. For example, in [43], we reported PCGs centered at 1550 nm with a channel spacing of 3.2 nm and 16 output channels.

AWG designed to be centered at 1550 nm with a channel separation (CS) of ∼3.2 nm, a free spectral range (FSR) of ∼57.6 nm, and 16 output channels [38]. The footprint of the device was 530 μm × 435 μm. The spectral resolution of AWGs and echelle gratings (PCGs) is determined by their channel spacing. However, it is possible to improve on that by multiplexing the same device with multiple inputs. In [40], we used an AWG with five outputs and four time-multiplexed inputs. The inputs were positioned such that their spectral response at the outputs was shifted according to a Vernier scale. This multiplied the resolution of the device, and also removed the ‘dead zones’ between the AWG output channels. The device is shown in Fig. 13. The inputs are each equipped with a Mach–Zehnder modulator to ‘label’ the input in a unique way. In this example we did this by sequentially switching the inputs on and off, but it is also possible to use modulations on a unique frequency. We demonstrated this device as a wavelength meter with 12 pm accuracy, even though the channel spacing of the AWG outputs is 4 nm.

Typical metrics used to characterize an AWG and PCGs are insertion loss, channel insertion loss, nonuniformity, and crosstalk. Depending on the requirements of the application, crosstalk can be described in two different ways. First, conventional crosstalk is defined as the difference between the maximum power at peak wavelength and the sum of all added crosstalk within a given bandwidth [39].

We have investigated both AWGs and PCGs for different wavelength ranges across different platforms. This is summarized in Table 1. Figure 12 shows the measured spectrum of an AWG designed to be centered at 1550 nm with a channel separation (CS) of ∼3.2 nm, a free spectral range (FSR) of ∼57.6 nm, and 16 output channels [38]. The footprint of the device was 530 μm × 435 μm. The spectral resolution of AWGs and echelle gratings (PCGs) is determined by their channel spacing. However, it is possible to improve on that by multiplexing the same device with multiple inputs. In [40], we used an AWG with five outputs and four time-multiplexed inputs. The inputs were positioned such that their spectral response at the outputs was shifted according to a Vernier scale. This multiplied the resolution of the device, and also removed the ‘dead zones’ between the AWG output channels. The device is shown in Fig. 13. The inputs are each equipped with a Mach–Zehnder modulator to ‘label’ the input in a unique way. In this example we did this by sequentially switching the inputs on and off, but it is also possible to use modulations on a unique frequency. We demonstrated this device as a wavelength meter with 12 pm accuracy, even though the channel spacing of the AWG outputs is 4 nm.

AWGs were also designed in Si$_3$N$_4$ [41] for very near-IR (900 nm) and in Ge-on-Si [42] for mid-IR (5 μm) wavelengths. Both of them showed performance that was comparable to the more mature SOI AWGs, with a CT of 20 dB and insertion loss <2.5 dB. While both of these platforms are not as mature as its SOI counterpart, a lot of design adaptations can simply be mimicked, in particular those dealing with the high sensitivity to phase errors caused by the HIC. These include the double etch step and gradually broadening waveguides in the apertures of the star coupler, as well as widening of the waveguides in the straight sections. The footprint of the device was larger than for SOI because of the lower index contrast in Si$_3$N$_4$ and a combination of lower index contrast and wider waveguides in Ge-on-Si, both necessitating a larger bend radius. In SOI, we have reported PCGs covering the C-band [36,43] and much longer wavelengths up to 3.8 μm [44]. For example, in [43], we reported PCGs centered at 1550 nm with a channel separation (CS) of ∼3.2 nm, a free spectral range (FSR) of ∼57.6 nm, and 16 output channels [38]. The footprint of the device was 530 μm × 435 μm. The spectral resolution of AWGs and echelle gratings (PCGs) is determined by their channel spacing. However, it is possible to improve on that by multiplexing the same device with multiple inputs. In [40], we used an AWG with five outputs and four time-multiplexed inputs. The inputs were positioned such that their spectral response at the outputs was shifted according to a Vernier scale. This multiplied the resolution of the device, and also removed the ‘dead zones’ between the AWG output channels. The device is shown in Fig. 13. The inputs are each equipped with a Mach–Zehnder modulator to ‘label’ the input in a unique way. In this example we did this by sequentially switching the inputs on and off, but it is also possible to use modulations on a unique frequency. We demonstrated this device as a wavelength meter with 12 pm accuracy, even though the channel spacing of the AWG outputs is 4 nm.

AWGs were also designed in Si$_3$N$_4$ [41] for very near-IR (900 nm) and in Ge-on-Si [42] for mid-IR (5 μm) wavelengths. Both of them showed performance that was comparable to the more mature SOI AWGs, with a CT of 20 dB and insertion loss <2.5 dB. While both of these platforms are not as mature as its SOI counterpart, a lot of design adaptations can simply be mimicked, in particular those dealing with the high sensitivity to phase errors caused by the HIC. These include the double etch step and gradually broadening waveguides in the apertures of the star coupler, as well as widening of the waveguides in the straight sections. The footprint of the device was larger than for SOI because of the lower index contrast in Si$_3$N$_4$ and a combination of lower index contrast and wider waveguides in Ge-on-Si, both necessitating a larger bend radius. In SOI, we have reported PCGs covering the C-band [36,43] and much longer wavelengths up to 3.8 μm [44]. For example, in [43], we reported PCGs centered

### Table 1. Performance Comparison between AWGs and PCGs Across Different Platforms

<table>
<thead>
<tr>
<th>Device/Technology</th>
<th>Central Wavelength (μm)</th>
<th>Footprint (μm$^2$)</th>
<th>No. of Channels/Channel Spacing (nm)</th>
<th>FSR (nm)</th>
<th>Insertion Loss (dB)</th>
<th>Crosstalk (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWG/Si$_3$N$_4$ [41]</td>
<td>1.55</td>
<td>530 × 435</td>
<td>16/3.2</td>
<td>57.6</td>
<td>&lt;3.0</td>
<td>&gt;25.0</td>
</tr>
<tr>
<td>AWG/Ge-on-Si [42]</td>
<td>0.89</td>
<td>450 × 750</td>
<td>12/2.0</td>
<td>30</td>
<td>&lt;1.5</td>
<td>20.0</td>
</tr>
<tr>
<td>S-AWG/Si$_3$N$_4$ [36]</td>
<td>5.3</td>
<td>1000 × 1000</td>
<td>5/19.0</td>
<td>148</td>
<td>2.5</td>
<td>19</td>
</tr>
<tr>
<td>PCG/Si$_3$N$_4$ [36]</td>
<td>1.55</td>
<td>305 × 260</td>
<td>4/30</td>
<td>144</td>
<td>&lt;2</td>
<td>19</td>
</tr>
<tr>
<td>PCG/Ge-on-Si [44]</td>
<td>3.8</td>
<td>1800 × 1700</td>
<td>8/10</td>
<td>100</td>
<td>&lt;1.5</td>
<td>&lt;20</td>
</tr>
<tr>
<td>PCG/Ge-on-Si [45]</td>
<td>5.0</td>
<td>1500 × 1200</td>
<td>6/25</td>
<td>170</td>
<td>&lt;5</td>
<td>22</td>
</tr>
</tbody>
</table>

Fig. 11. (a) Optical image of an AWG with zoomed-in images of (b) star coupler, (c) delay lines, and (d) aperture.
at 1550, 1650, and 2125 nm wavelengths with an intention to use them for the glucose spectroscopic sensing application. These PCGs with 8–16 output channels and FSR in excess of 110 nm exhibited crosstalk better than 18 dB and insertion loss of ∼5 dB. Recently, we demonstrated, to the best of our knowledge, the first complex mid-IR PIC functionality in an SOI platform in the form of AWGs and PCGs [44]. The PCGs were centered at 3.8 μm and showed crosstalk in the range of 15–19 dB and low insertion loss <2 dB. The higher crosstalk is mainly attributed to the nonoptimized layout parameters and air top clad instead of oxide clad. We also reported, to the best of our knowledge, the first PCG in Ge-on-Si for the 5 μm wavelength range with six output channels and 25 nm channel spacing [45]. The input and output apertures were widened to 4 μm. To increase the reflectivity to ∼50%, we employed a DBR mirror that lowered the insertion loss to 4.9/4.2 dB for TE/TM polarization, and the crosstalk was found to be 22 dB/23 dB as shown in Fig. 14. The insertion loss is higher than that reported for AWGs. The main reason behind it is the low reflectivity from the grating facets whose minimum linewidth is limited by our fabrication tool. For achieving a reflectivity of ∼95% a DBR with 320 nm gap needs to be fabricated, which is possible by using standard CMOS tools. The difference in the design implementation means that AWGs and PCGs fulfill different sets of requirements, and one can make a choice based on the required specs. PCGs perform better for large FSR and low channel spacing, while AWGs are a better choice for small FSR and large channel spacing [36].

4. ON-CHIP BROADBAND SOURCES

An efficient light source that can be integrated with PICs (e.g., silicon photonics) is an important building block to realize the goal of on-a-chip spectroscopic systems. In spite of all the unique advantages silicon has to offer it is an indirect bandgap material, due to which light generation in silicon remains a huge challenge. Therefore, to get light into the silicon chips there are three options: 1) by coupling an external light source to the SOI chip, but this is not an attractive solution due to the high coupling losses and high packaging costs involved; 2) by heterogeneous integration of III-V materials such as indium phosphide (InP) that are excellent light emitters, in the form of LEDs and lasers on to SOI [14]; 3) light generation in silicon itself by making use of the nonlinear effects in silicon (e.g., supercontinuum sources). For spectroscopic sensing, there is particular interest for an efficient broadband or tunable source that can not only reveal the relevant absorption features of molecules for sensing applications but also be easily integrated with PICs.

A. Heterogeneous Integration of Thin InP Membranes on SOI for Efficient On-Chip LEDs

We have been investigating the integration of thin InP membranes on SOI to create an efficient optically pumped waveguide coupled LED. The main focus of the work has been to optimize the SOI designs and to simplify the post-processing methods of InP devices to make reliable and efficient (in terms of yield and power coupled) LEDs on SOI. The main approach we are following is to make thin membranes by removing the layers necessary for electrically pumped devices, which will then couple the spontaneous emission more efficiently to a single-mode SOI waveguide. This is illustrated in Fig. 15 [46]. It consists of a SOI photonic wire, a III-V ridge waveguide, and a tapering section connecting the two. The membrane is chosen such that the optical pump (injected through the silicon wire) is heavily absorbed and the resulting spontaneous emission is efficiently collected. The high refractive index contrast of the membrane also allows for short and fabrication tolerant tapers. Figure 16 shows a simulation of the confinement of the pump light in the quantum wells and the spontaneous emission collection efficiency as a function of the membrane thickness. The vertical line shows the membrane chosen. One can see that around 115 nm there is a clear optimum in confinement. Moreover, this is virtually independent of the number of quantum wells, so we can put in as many as possible. We chose four quantum wells, which gives a total confinement of 18.4%. The spontaneous
emission collection efficiency is optimal at 96 nm, where 12% of the radiated light is captured in the fundamental waveguide mode (another 12% is captured in the other direction). The device was fabricated using wet etching and photoresist masks (inset in Fig. 15). Figure 16 (right) shows the very first spontaneous emission spectrum in the silicon waveguide when pumped with 0.5 mW [46]. The signal power spectrum in transmission is ±15 dB lower than the one in reflection. Also the peak wavelength and the 3 dB bandwidth are different: for the signal in reflection these are 1565 and 227 nm, while in transmission they are 1613 and 100 nm, respectively. The blue-shift, the difference in 3 dB bandwidth, and the power spectrum are due to the fact that the LED is not pumped completely. The unpumped part of the LED reabsorbs the generated light, where the shorter wavelengths are absorbed much more efficiently.

B. Colloidal Quantum Dot Integration on Si₃N₄ PICs

Si₃N₄ devices are thus far limited to passive functionalities, and an active Si₃N₄ photonics platform is highly desired. Achieving this requires combining an active medium with Si₃N₄ building blocks such as waveguides and resonators [47,48]. To this end, we have been investigating colloidal quantum dots (QDs), a novel class of materials with efficient and broadly tunable luminescence. To realize such an active platform, the most important aspect is to develop an integration scheme that is capable of preserving the optical properties of QDs while maintaining low-loss Si₃N₄ circuits. In addition, to enable effective coupling between the emission and the optical modes, more specifically in optical cavities, the light field in the Si₃N₄ components should maximally overlap with the QD emitters. In our study, we have developed a fully inorganic and low-loss Si₃N₄-QD hybrid photonic platform by using a low-temperature (120°C–270°C) plasma enhanced chemical vapor deposition process and an optimized dry etching step, and demonstrated loss as low as 2.7 dB/cm at 900 nm wavelength for waveguides with embedded QDs [49]. Based on this platform, we designed and fabricated freestanding Si₃N₄ microdisks with colloidal QDs embedded inside. The disks are vertically coupled with on-chip Si₃N₄ bus waveguides [Figs. 17(b) and 17(c)]. Efficient coupling of the QD photoluminescence (PL) to disk resonance modes is experimentally demonstrated in the visible wavelength range [50]. A spectrum for a disk of radius 3.5 μm is shown in Fig. 16(a). The coupling strength and selectivity for different order modes can be controlled by tuning the parameters of the geometric configuration. These high-performance active Si₃N₄-QD building blocks offer a promising platform for developing optical sources operating over a very broad wavelength range and open up new opportunities for both QD-based quantum optics and Si₃N₄-based photonic applications.

C. Supercontinuum Sources

Silicon combines a high linear refractive index with a very high nonlinear refractive index with the latter 20 times higher than the value found in Si₃N₄. As a consequence, high confinement SOI waveguides, where even moderate powers lead to high intensities, show a very strong nonlinear response. The nonlinear parameter, a parameter that takes into account both the nonlinear refractive index of the material as well as the confinement of the waveguide, is found to be two orders of magnitude higher than for Si₃N₄ waveguides. On top of that, the dispersion of the waveguides can be engineered through the strong waveguide dispersion of the photonic wire waveguides. And although silicon suffers from nonlinear absorption, two-photon absorption, at telecom wavelengths this unwanted effect disappears in the mid-IR. Indeed, using photon energies with energies close to or less than half the bandgap energy (λ = 2200 nm) reduces or even eliminates the two-photon absorption effect. In short, SOI waveguides are ideal candidates for performing nonlinear optics in the mid-IR—even more so because of the wide transparency band of silicon (from 1.1 to 8 μm).

We have demonstrated broadband supercontinuum generation (SCG) in the mid-IR, by pumping 2 cm long silicon waveguides with 2 ps pulses centered at 2120 nm [51]. These results were obtained with pump peak powers as low as 12.7 W. In the process of SCG, the narrow bandwidth input pulses (∼30 nm) are broadened, and their spectrum spans from 1535 nm down to 2525 nm. Figure 18 shows the spectrum of the output pulses. The main driving force behind the
broadening in the waveguides here is found to be the amplification of background noise, due to spontaneous four-wave mixing. As a result, the pulse-to-pulse variation can be very high. And although the averaged spectrum is very stable, the source is relatively noisy because of the high pulse-to-pulse variation caused by the noise amplification.

In a more recent experiment [52], we have demonstrated a supercontinuum that preserves the coherence of the input pulses such that the pulse-to-pulse variation is low. In this experiment a 390 nm thick, 1600 nm wide, 2 cm long waveguide is pumped with ultrashort pulses (~70 fs) centered at 2290 nm. Indeed it is known that SCG by ultrashort pulses tends to preserve the coherence [53]. The spectrum of the output pulses spans over more than an octave from 1540 up to 3200 nm. Because the nonlinear process preserves the coherence, the pulse-to-pulse variation of the broadened pulses after the waveguide is very limited. In turn this means that the spectrum (the Fourier transform of the periodic output pulse train) is a frequency comb. This has been verified experimentally [46]. In this experiment, the generated supercontinuum is combined with a narrowband continuous wave laser at 1586 nm and sent to a photodetector. The RF spectrum of the electrical output of the photodetector is shown in Fig. 19. Two beat notes of the lines adjacent to the CW laser in the supercontinuum/frequency comb with the CW laser can be seen in Fig. 19 at 28 and 72 MHz. Since the phase coherent supercontinuum is generated with a 100 MHz pulse train, an RF tone at this frequency can be seen as well. The broadband spectrum consisted of evenly spaced lines with a linewidth of <50 kHz. This shows how silicon waveguides can be used to help in the challenge of making mid-IR supercontinuum and frequency comb sources.

A broadband source covering red to near-IR wavelengths where tissue and cells possess low absorption and scattering loss is of particular interest to biological applications such as bio-imaging [54], optical coherence tomography [55], or Raman spectroscopy [56]. Si$_3$N$_4$ is a promising platform to realize a visible broadband source through SCG due to its transparency in visible wavelengths, relatively large index contrast with oxide cladding (~0.5), reasonably strong effective nonlinearity [57], and low intrinsic loss [16]. Ultrabroadband SCG has already been demonstrated [58] on the Si$_3$N$_4$ platform. The main hurdle for efficient SCG in Si$_3$N$_4$ waveguides is to achieve the anomalous group velocity dispersion (GVD) required for phase matching. At visible wavelengths, Si$_3$N$_4$ possesses strong normal material dispersion, which is difficult to compensate by waveguide dispersion with only tuning the dimensions of the waveguide core. This hurdle can be circumvented by pumping at longer wavelengths [57], yet the power fraction of the generated supercontinuum falling inside the biological window is limited. Cascading $\chi(3)$ and $\chi(2)$ processes [59] could also relax the requirement for phase matching, yet it requires careful dispersion management for each step. Recently we demonstrated a LPCVD Si$_3$N$_4$ waveguide with anomalous GVD through under-etching [60]. As shown in the inset of Fig. 20, we start from a waveguide that is 300 nm high and 500 wide and dip it into diluted hydrofluoric acid to under-etch 150 nm of the oxide cladding. The removal of the oxide cladding helps increase the index contrast and increases the waveguide dispersion. The total GVD of the engineered waveguide is shown as the green curve in Fig. 20, which is anomalous between 740 and 907 nm. The reduced mode area also enhances the effective nonlinearity of the waveguide to 7 $W^{-1}$ $nm^{-1}$. A pulse train from a MaiTai Tsapphire source (100 fs, 80 MHz) is coupled to this dispersion engineered waveguide, through a microscope objective (40x, NA = 0.5), and at a peak coupled power of 874 W we observed an octave spanning SCG ranging from 488 to 978 nm, as shown in Fig. 20 as the blue curve. The dispersion tuning by under-etching and the SCG demonstrated on this platform constitute the first step towards an on-chip compact broadband source for biological applications.

5. CONCLUSIONS

On-a-chip implementations of spectroscopic sensing functions hold immense potential for applications requiring large
manufacturing volumes, small footprints, and moderate cost. Examples include medical point-of-care devices, bodily implants for monitoring of glucose levels, and sensing devices integrated in smartphones. The field of spectroscopy on a chip is still very immature as compared to free-space spectroscopy systems, but progress is very rapid. Technological platforms are being developed for an array of wavelength ranges from the visible to the mid-IR. A variety of proof-of-concept on-a-chip demonstrations have been reported for absorption spectroscopy and for Raman spectroscopy. The underlaying building blocks—sources, spectrometers, and detectors—are maturing. In spite of the promise there are still many challenges. Optical path lengths on a chip are necessarily short, and that sets limits to the limit of detection for low-concentration analytes. The use of evanescent sensing implies that the light beam interacts not only with the analyte but also with the waveguide material, which can lead to background signals. Single-mode broadband light sources with high power spectral density and high power conversion efficiency are of utmost importance in this field, but most implementations reported so far still have poor performance. But all of these tough challenges are worth being tackled given that the rich variety of field enhancement mechanisms that can be deployed on a chip—throughwaveguide confinement, Purcell effects, cavity effects, and plasmonic effects—can lead to performance that outperforms conventional spectroscopic sensing systems.

ACKNOWLEDGMENT

The authors would like to acknowledge ERC-InSpectra Advanced Grant, ERC-MIRACLE, ERC-ULPPIC and Methusalem (Smart Photonics Chips) for their support. The authors would like to thank funding agencies IWT and FWO that helped in carrying out various parts of the work presented in the paper. The authors acknowledge ePIXfab for the silicon chips fabrication. The authors would like to acknowledge imec for Si3N4 chips fabrication and Ge-on-Si epi wafers. Finally, the authors would like to thank Steven Verstuyft and Liesbet Van Landschoot for cleanroom related work and Michael Vanslembrouck for measurements in the lab.

REFERENCES


