

Raman on-chip: Current status and future tracks

Nicolas Le Thomas^{*a,b}, Zuyang Liu^{a,b}, Chupao Lin^{a,b}, Haolan Zhao^{a,b}, Roel Baets^{a,b}

^aPhotonics Research Group, Ghent University-IMEC, 9052 Ghent, Belgium; ^bCenter for Nano- and Biophotonics, Ghent University, 9052 Ghent, Belgium

ABSTRACT

On-chip Raman sensing enabled by large-scale photonic integration is a promising technology for biological and healthcare applications. In this contribution we give a review of the current status of on-chip Raman sensing with a particular focus on the ultimate performances. We discuss the limitations in terms of detection limit and the different paths currently followed to get around them.

Keywords: Raman spectroscopy, integrated silicon photonics, thermodynamics fluctuation, extended irreversible thermodynamics, covariant thermodynamics.

1. INTRODUCTION

Monitoring the presence of specific molecules or micro-organisms in the environment or in human bodies is a major societal issue in terms of public health. The following key examples illustrate the need for efficient biological and chemical sensors. If the presence of pathogens, such as the Covid-19 virus, can be analyzed in wastewater in an efficient way, governments can take diligent measures to contain pandemics. If the level of drugs in patient body fluids is quantified throughout a treatment, the drug concentration can be optimized, which will result in saving life and cost by a proper usage of medication.

The analysis of the samples to probe should be as fast as possible to enable effective decisions. Currently, the samples are often transported toward analysis centers as in the case of wastewater collected in different urban areas. On-site analysis with handheld devices has the potential of drastically suppressing this time-consuming step. It requires a large-scale deployment of the sensors implying low-cost and large-scale manufacturing approaches. One solution that fulfills this need is silicon integrated photonics, namely photonic structures whose starting support are silicon wafers and that can be processed by well-established microelectronic foundries.

To address the need of molecular sensors for the specificity, scattering Raman spectroscopy is a very appropriate technique insofar as it probes the molecular vibrations frequencies that are fingerprints of a given molecule. This optical technique is non-invasive and can be implemented for probing several molecules at the same time. The number molecule species that can be analyzed depends on the complexity of their Raman spectra, of the signal to noise ratio, the resolution of the spectrometer and of the numerical methods used to disentangle the spectral information.

Although Raman spectroscopy is an appealing technique, its signal is weak: The spontaneous Raman scattering cross-sections is around ten orders of magnitude weaker than the fluorescence scattering cross-section. This weakness has several consequences as regards detection limit and cost. Until recently, the approach to cope with low Raman signals while pushing the molecular detection limit consisted in using bulky and expansive confocal Raman microscopy systems, namely about 1 m³ and one hundred thousand euros. Such systems are hardly compatible with an on-site sensing investigation.

Few years ago, we demonstrated that the implementation of Raman spectroscopy with photonics integrated circuits offers a compact and low cost, namely about 1 mm² and ten euros, that outperforms confocal Raman microscope for bulk solution sensing¹. The conversion efficiency of high-index contrast and single mode waveguides is more than two decades larger than for the Raman collection with a far-field microscope². The on-chip Raman approach is perfectly suited for field deployment; however, the detection limit still needs to be improved by at least one decade for targeting relevant applications. The ultimate reason for this limitation is the presence of a background in the collected optical spectrum that is intrinsic to waveguide itself.

*nicolas.lethomas@ugent.be; phone +32-9-264 3339; <https://www.photonics.intec.ugent.be>

The purpose of this paper is to discuss the current status of integrated waveguide-based Raman sensing with regards to their performances and our current understanding of the origin of the detection limit of these sensors. We will particularly focus on the physical origin of the background as it constitutes the main limitation and its description involves advanced physical theories at the border of our knowledge to model the interaction between light and solid-state matter.

The typical sizes of photonic structures are becoming so small that non-equilibrium thermodynamic effects start to play an important role for establishing the detection limit in integrated optical sensor. In the frequency range where molecular Raman sensing is interesting, i.e. for Stokes shifts at THz frequencies, the description of these non-equilibrium effects requires approaches that go beyond the conventional treatment of thermo-refractive noise.

Describing light matter interaction when irreversible non-equilibrium processes are at play is challenging, in particular at sub-picosecond time scale. Interestingly, phenomenological transient thermodynamics have been developed by cosmologists in the eighties of the last century. At that time, such theories were motivated by question of first principle and were of purely theoretical interest. It was expected that *“it will probably be some time before the state of the art in [astrophysical and cosmological situations] makes such refinements necessary”*³. Forty years later, the technological advances in integrated photonics makes such a nonequilibrium thermodynamics theory highly relevant to understand the origin of photonic noise in on-chip Raman sensors, and on-chip single photon and squeezed state sources. In this paper we will pinpoint the needs for further theoretical developments to adapt the existing transient thermodynamics theory to photonic engineering problems.

The article is organized as follows: After the introduction we make the state of the art of waveguide-based Raman sensors by articulating the discussion around the parameters that govern the intensity of the collected molecular Raman signal at the outcoupling port of the waveguide. The third section describes the current theoretical description of the shape of the background, whereas in the next section we pave the way for a more advanced and general description of the background noise. Before to conclude, we discuss different possible approaches to get around the presence of the background.

2. STATE OF THE ART OF THE DETECTION LIMIT

2.1 Parameters governing the intensity of the Raman signal collected at the waveguide output port

The principle of waveguide-based Raman sensing is schematically illustrated in figure 1. A pump beam propagates inside a single mode waveguide and is scattered by molecules surrounding the core of the waveguide. A large fraction of the spontaneous Raman signal that results from the scattering of the pump by the ensemble of interacting molecules is collected by the single guided mode. The total signal is spectrally analyzed at one of the output ports of the waveguide. The key asset of the waveguide compared to a confocal detection is the total length L of the interaction volume. It can reach several centimeters depending of the waveguide loss whereas it is limited to about one wavelength for a diffraction limited confocal detection. Note that the single mode nature of the waveguide or the diffraction limited property of the confocal detection are crucial to achieve the best spectral resolution in each case and enable a fair comparison of both techniques.

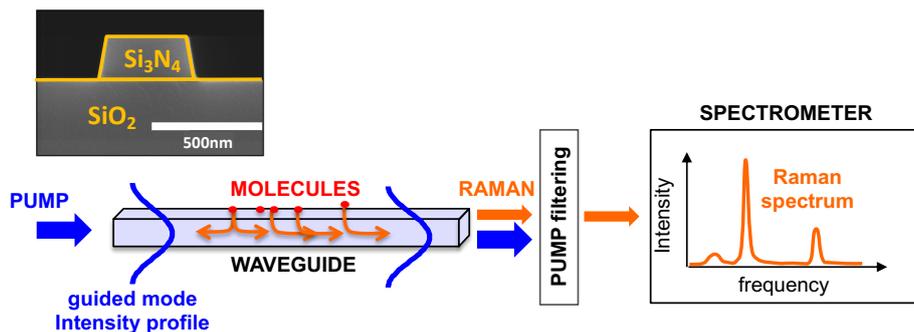


Figure 1. Schematic of the on-chip Raman sensing with a typical electron scanning micrograph of a silicon nitride waveguide

The function of the integrated photonic waveguide for Raman on-chip sensing is double. The waveguide acts as an efficient channel for an incident power P_{in} to excite a given spatial distribution of molecules surrounding the high-index core

material and to efficiently collect the molecular spontaneous Raman emission. Based on a classical electromagnetism model of the coupling between a dipole and a light field, the total molecular Raman emission at the outcoupling port of the waveguide $P_{\text{out}}^{\text{R}}$ can be expressed in a forward configuration as:

$$\frac{P_{\text{out}}^{\text{R}}}{P_{\text{in}}} = \rho \times \sigma \times \frac{\eta}{2} \times \gamma_{\text{in}} \times \gamma_{\text{out}} \times e^{-\alpha_{\text{P}}L} \times L \quad (1)$$

assuming that the losses α_{P} at the pump and α_{R} at the Stokes or Anti-Stokes Raman frequencies are identical.

The two first terms in the right part of this expression, namely the density of molecules ρ and the spontaneous Raman scattering cross-section σ are intrinsic properties of the analyte to probe. The third term $\eta/2$ can be defined as a light conversion factor between the pump beam and the Raman signal. It encompasses the light-matter interaction process between the guided mode and the molecules. The next parameters in Eq. (1) are all intrinsic characteristics of the photonic waveguide. These ones are the coupling efficiencies γ_{in} and γ_{out} of the light in and out of the photonic chip, the propagation loss coefficient α_{P} , and the waveguide length L . Note that L has to be replaced by $(e^{L\Delta\alpha} - 1)/\Delta\alpha$ when $\Delta\alpha = \alpha_{\text{P}} - \alpha_{\text{R}} \neq 0$. Besides, in a backward light-collection configuration, the factor $e^{-\alpha_{\text{P}}L} \times L$ is replaced by $(1 - e^{-(\alpha_{\text{P}} + \alpha_{\text{R}})L})/(\alpha_{\text{P}} + \alpha_{\text{R}})$.

To maximize the Raman signal for a given incident pump power at the input port of the waveguide, a first lever that is a common problematic to all integrated photonics structures is to minimize the propagation loss and maximize the in- and out-coupling efficiency. Minimizing the propagation loss is crucial to maximize the interaction length L between the guided mode and the analytes, i.e. to take all the benefit of the specificity of the waveguide-based sensing approach.

The second lever that is more specific to Raman sensing consists in maximizing the conversion factor η , namely the light-matter interaction, as detailed below. This factor contains the more advanced optimization step for the optical engineer in terms of waveguide design.

2.2 Raman conversion factor and its optimization

The conversion factor η can also be separated in contributions that are intrinsic to the analyte solution and to the waveguide, and in a mixed term involving both the guided mode (light) and the molecules (matter):

$$\eta = \frac{\lambda_{\text{R}}^2}{n} \times \frac{n_{\text{g}}(\omega_{\text{P}})}{A_{\text{eff}}(\omega_{\text{P}})} \times \frac{n_{\text{g}}(\omega_{\text{R}})}{A_{\text{eff}}(\omega_{\text{R}})} \times \frac{\iint |\vec{e}_{\text{d}}(\vec{r}) \cdot \vec{E}(\vec{r}, \omega_{\text{P}})|^2 |\vec{e}_{\text{d}}(\vec{r}) \cdot \vec{E}(\vec{r}, \omega_{\text{R}})|^2 f_{\text{d}}(\vec{r}) d\vec{r}}{\max\{\varepsilon(\vec{r}, \omega_{\text{P}}) |\vec{E}(\vec{r}, \omega_{\text{P}})|^2\} \max\{\varepsilon(\vec{r}, \omega_{\text{R}}) |\vec{E}(\vec{r}, \omega_{\text{R}})|^2\}} \quad (2)$$

The first contribution with n the index of the analyte solution and λ_{R} the wavelength at the Raman transition can in principle be maximized by increasing λ_{R} . The gain in η for the collected Raman emission $P_{\text{out}}^{\text{R}}$ is however counterbalanced by the Raman scattering cross section that varies as $1/\lambda_{\text{R}}^4$. In general, it is advantageous to work with λ_{R} as small as possible, i.e. with a large pump frequency ω_{P} .

The two terms involving the group index $n_{\text{g}}(\omega)$ and an effective interaction area $A_{\text{eff}}(\omega)$ of the guided mode at a frequency ω , either the pump or Raman frequency, depend only on the geometrical dielectric map of the cross-section of the waveguide. For a given dielectric map $\varepsilon(\vec{r}, \omega)$, these two parameters can be optimized with a sweep of the different dimensions of the waveguide cross-section. The effective interaction area that is given by

$$A_{\text{eff}}(\omega) = \frac{\iint \varepsilon(\vec{r}, \omega) |\vec{E}(\vec{r}, \omega)|^2 d\vec{r}}{\max\{\varepsilon(\vec{r}, \omega) |\vec{E}(\vec{r}, \omega)|^2\}} \quad (3)$$

provides a measure of the spatial distribution of the fields. Minimizing A_{eff} goes hand in hand with the enhancement of the conversion factor if the molecules are all located where the energy density $\varepsilon(\vec{r}, \omega) |\vec{E}(\vec{r}, \omega)|^2$ is maximum. This last condition can be implemented with plasmonic structures either in the form of nanoantennas⁴ or nanoplasmonic layers⁵ on top of the dielectric waveguide. The local concentration of the electromagnetic field boosts the Raman interaction with the molecules and can lead to a signal enhancement of more than six decades compared to a standard free space collection⁴. However, the use of nanoantenna strongly limits the detection volume and consequently the total number of molecules that can be probed. Besides, for solutions with a low concentration of analytes the low probability to reach the detection volume can be an obstacle for efficient sensing. Increasing the number of nanoantenna results in a limited improvement due to the absorption and scattering losses induced by the metal nanostructures. A compromise to benefit from the high-field enhancement of metal structures with a combination of a long sensing distance is the nanoplasmonic slot waveguide. For

a 15 μm long hybrid plasmonic waveguide, a conversion factor as large as $\eta = 10^{-9}$ has been reached with a monolayer of 4-nitrothiophenol molecule for which the Raman scattering cross-section is $\sigma = 1.8 \times 10^{-29} \text{ cm}^2/\text{sr}^5$. Currently, in the case of a monolayer and for the same analyte Raman scattering cross-section, the SiN dielectric slot waveguides outperform the hybrid plasmonic waveguides in terms of conversion factor due to their long interaction length L of few centimeters.

For bulk sensing, a large fraction of analyte molecules is far from the location where the energy density of the mode is maximal. In this case, the optimization of the conversion factor is the result of a compromise between minimizing A_{eff} and maximizing the last mixed term in Eq. (2) that takes into account the normalized spatial distribution $f_d(\vec{r})$ of molecules and the random orientation vector $\vec{e}_d(\vec{r})$ of their dipole moment. Maximizing this last term is equivalent to overlapping the field of the mode with the analyte solution as much as possible. Such an optimization problem is conveniently tackled with finite elements mode solver, e.g. COMSOL, and a sweep of the geometrical parameters defining the cross-section of the waveguide dielectric map⁶. The parameter sweep is constrained by the limitation of the fabrication process. Using deep UV lithography, a typical slot width minimum is 150 nm for silicon photonics. Using the atomic layer deposition technique (ALD), which provides a conformational deposition of thin layer with a monolayer control, the constraints imposed by deep UV lithography can be relaxed. It enables to reach slot width as narrow as 15 nm⁵.

In addition to the geometrical parameters of the cross-section of the waveguide, the optical index of the core material has a significant impact on the conversion factor. Larger index contrasts between the core and the cladding result in larger group indexes and smaller effective interaction areas, and consequently to larger Raman conversion factors. The advantage of high-index contrast waveguide can be illustrated with the four-fold η enhancement between alumina core ($n_{\text{Al}_2\text{O}_3}=1.6$) and SiN core ($n_{\text{SiN}}=1.9$) strip waveguides on thermal oxide, with ethanol analyte solution ($n_{\text{ethanol}}=1.36$) in the top cladding.

If the molecular Raman emission $P_{\text{out}}^{\text{R}}$ is the only optical signal present at the outcoupling port of the waveguide, the detection limit is simply provided by the signal to noise ratio (SNR) with the shot noise the ultimate noise contribution in a classical detection approach. In this ideal condition, the SNR is proportional to $\sqrt{\gamma_{\text{in}} \times \gamma_{\text{out}} \times L \times P_{\text{in}} / (\hbar\omega_R)}$, so maximizing the input power, the coupling efficiencies and the waveguide length sub linearly improves the detection limit.

In reality, the Raman emission from the analyte at a frequency ω_R is not the only signal collected at the waveguide outcoupling port: The matter that is used to make the waveguide also generates a signal at ω_R when excited at a pump frequency ω_p . This background signal of intensity $P_{\text{out}}^{\text{B}}$ is composed of molecular Raman features intrinsic to the atomic vibrations of the medium and of a spectrally broadband component, the origin of which is discussed below in the case of an amorphous material. The average value of the background can be subtracted from the total signal collected at the outcoupling port of the waveguide if the background varies much more slowly than the molecular Raman signal of the analyte. In this case, the signal to noise ratio can be written as $\text{SNR} = (P_{\text{out}}^{\text{R}} / \hbar\omega_R) / \sqrt{(P_{\text{out}}^{\text{R}} + P_{\text{out}}^{\text{B}}) / (\hbar\omega_R)}$. The SNR converges toward an ultimate figure of merit when $P_{\text{out}}^{\text{R}} \ll P_{\text{out}}^{\text{B}}$, which we call the signal to background noise ratio (SBNR), $\text{SBNR} = (P_{\text{out}}^{\text{R}} / \hbar\omega_R) / \sqrt{P_{\text{out}}^{\text{B}} / (\hbar\omega_R)}$. Due to the presence of the background, this figure of merit is currently the most relevant one to quantify the detection limit.

As the background varies linearly with the pump and the waveguide length, at least within the power range that has been experimentally investigated up to now, the SBNR is proportional to $\sqrt{\gamma_{\text{in}} \times \gamma_{\text{out}} \times L \times P_{\text{in}} / (\hbar\omega_R)}$, similarly to the SNR. Improving the coupling efficiencies and decreasing the optical losses to maximize the waveguide length can still lead to significant improvement of the detection limit even with the presence of a background as soon as the background signal does not overload the dynamic of the signal detector. The coupling efficiencies are currently of the order of -7dB per port due to the need for free-space coupling. The standard approach of lens fiber coupling is indeed not an option due to the extra background induced by optical fibers. Developing efficient (-0.5dB) grating coupler designed for the pump wavelength at the input port and for a specific Raman transition frequency at the output port could improve the SBNR by a factor 4.5. Grating couplers are an efficient approach at the telecom wavelength thanks to the high index of the silicon, they have not been demonstrated in the visible wavelength range with efficiency as high as -0.5dB yet.

3. FUNDAMENTAL LIMITATION OF THE DETECTION LIMIT

3.1 Origin of the background: Intrinsic thermal fluctuations

The background signal in amorphous waveguides deserves a special attention in view of its critical role for the detection limit of the Raman on-chip sensing. Neglecting molecular Raman features that are related to specific local symmetries of the amorphous medium, the overall shape of the background spectrum follows an exponential decay as unveiled in a logarithm plot, see Fig. 2. This exponential decay is common to all the dielectric amorphous media that we have currently investigated. In addition to silicon nitride (Si_3N_4), alumina (Al_2O_3) and tantala (Ta_2O_5) presented in Fig. 2, the background of titanium oxide (TiO_2) waveguides also exhibits the same exponential shape but with a larger signal as reported in⁷.

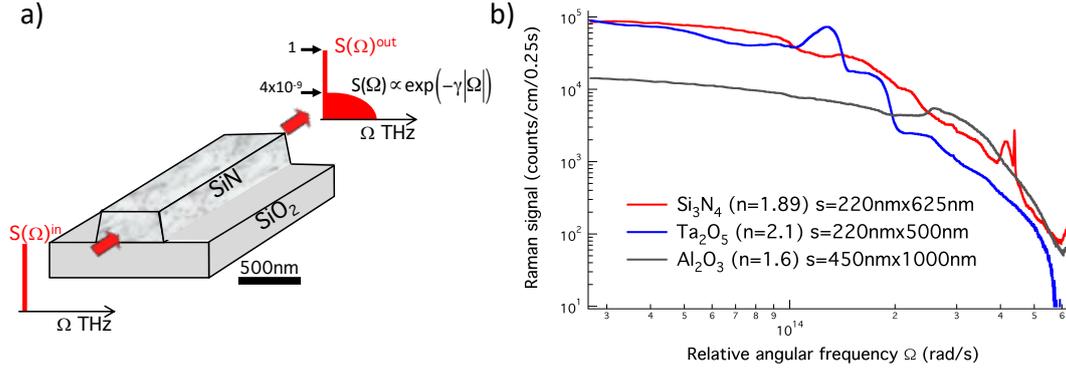


Figure 2. a) Illustration of the appearance of a frequency continuum when a single frequency pump propagates through a single mode waveguide. b) Optical spectra of the Raman background for three different materials of the waveguide core and a cross-section area s . The log-log plot highlights the exponential decay of the signal at high frequencies.

The presence of a continuous background is in itself not surprising at room temperature. Any medium at a temperature larger than the absolute zero Kelvin is subject to local thermal fluctuation at equilibrium. Standard theories⁸⁻¹⁰ of thermal fluctuations in continuous media assume that the thermal fluctuations at two different places are not correlated and that the constitutive equation connecting the heat flux \vec{q} to the local temperature is given by the Fourier heat transfer law: $\vec{q} = -\kappa \vec{\nabla} T$, with κ the heat conductivity. Moreover, the system is considered to be locally at equilibrium. Based on these two assumptions and implementing the fluctuation-dissipation theorem for distributed systems, the optical spectrum of a single frequency light beam propagating in the medium gains a background decaying as $1/\Omega^2$ at large frequencies. In the MHz regime such a thermo-refractive noise has been experimentally observed in optical fibers¹¹⁻¹³ and in optical microcavities¹⁴⁻¹⁶.

In the frequency range of interest for Raman spectroscopy, i.e. in the THz frequency range, the local equilibrium hypothesis is questionable as at such frequencies the dynamics is associated with sub-picosecond fluctuations. The system can be seen as a sea of acoustic phonons. Besides, the parabolic nature of the Fourier heat law leads to unphysical instantaneous response all over the medium when a heat perturbation is applied locally. Such an issue can be solved when the Fourier law is modified by taking into account the inertial effect of the thermal flux. Introducing a phenomenological relaxation time τ of the heat flux the Fourier law is replaced by the Cattaneo equation: $\tau \partial \vec{q} / \partial t = -(\vec{q} + \kappa \vec{\nabla} T)$. However, although it solves the problem of the instantaneous propagation, such an equation is not compatible with the assumption of a local equilibrium, and in particular with the request of an entropy production that is definite positive. To make the Cattaneo equation compatible with a thermodynamic description of the system, one approach consists to consider that the local “equilibrium” state of the system is defined by the thermal flux in addition to the specific energy (Note that for the time being we neglect the variation of the specific volume and the mechanical viscous flux) and that the specific entropy is a function of the heat flux in the same way as the specific energy¹⁷. This approach is at the heart of the extended irreversible thermodynamics (EIT) theory¹⁸. In particular it leads to a generalized Gibbs equation involving the thermal flux, from which we can determine the variation to the second order $\delta^2 s$ of the entropy induced by a perturbation of the equilibrium state. As $\delta^2 s$ is a quadratic function of the temperature fluctuation δT and of its gradients $\nabla(\delta T)$, the probability density

\mathcal{P} of the fluctuations that is linked to $\delta^2 s$ via the Einstein formula $\mathcal{P}(\delta T, \nabla(\delta T)) \propto e^{-|\delta^2 s|/k_B}$ leads to the spatial correlation of the thermal field fluctuation. In contrast to the standard model of thermo-refractive noise, the spatial fluctuations of the thermal field are characterized by a correlation length $\ell = \sqrt{\tau D_T}$, with D_T the thermal diffusivity of the medium.

We have shown in¹⁹ that the presence of a correlation length of the thermal field enables to retrieve the observed exponential decay of the background at high frequencies. More specifically, implementing EIT for the thermal field of the medium, the optical spectrum $S^B(\Omega)$ of the high-frequency background contribution at the output port of the waveguide is given by:

$$S^B(\Omega) = 4\pi^2 A_0^2 \langle \delta n^2 \rangle \frac{L\ell}{\lambda_0^2} \frac{\ell^2}{\ell^2 + 2W^2} \gamma e^{-\gamma|\Omega|}, \quad (4)$$

where A_0 is the amplitude of the electromagnetic mode propagating with a wavelength λ_0 , the transverse intensity profile of which is taken as a Gaussian function with radial half-width W . The parameter γ is a characteristic time of the thermal field dynamics taking into account the overlap between the electromagnetic field and the matter. The thermal field dynamic is coupled to the propagating electromagnetic mode via the electric susceptibility of the material, or equivalently here to the refractive index for a non-absorbing medium. A linear constitutive equation is assumed between the dipole moment field and the electromagnetic field. The fluctuation of the thermal field induces a refractive index fluctuation δn of the medium which impact the background spectrum via its second-order statistic moment $\langle \delta n^2 \rangle$.

3.2 Current limitation of the background model

Our approach is based on several simplifying approximations. A first one consists in assuming that the propagating electromagnetic field overlaps with all the medium. If the transverse profile of mode overlaps partially with the cross-section of the waveguide such as in slot waveguides, the background spectrum is expected to have a more complex dependence with W involving also the geometrical parameters of the cross-section of the waveguide. With the current approximation, the optical spectrum of the background varies as $1/W^2$ for $\ell \ll W$ as expected for a confined mode: The larger the size of the medium confining the thermal field, the smaller the thermal fluctuations⁸. Such a trend is supported by Fig.2b) where the alumina waveguide exhibits a background level lower by a factor height compared to the two other waveguides whereas its cross-section his four times larger.

As another drastic approximation, the viscous effects are neglected by cancelling the viscous pressure tensor of the medium. Such an approximation has the merit to lead to only one phenomenological parameter in the generalized Gibbs equation, namely the heat flux relaxation time τ . Within the current experimental conditions, this approximation does not seem to impact the overall shape of the background. It may however have an impact on the value of the correlation length of the thermal fields that is experimentally retrieved from Eq.4.

In¹⁹ we have implemented an approach where the self-consistent nature of the interaction between the medium and the electromagnetic field is disentangled by determining first the thermal fluctuations without the presence of the applied electromagnetic field. As a result, we have implicitly assumed that the dipole moment field resulting from the interaction between the propagating electromagnetic mode and the medium does not perturb the dynamics of the heat flux. Such an assumption is questionable for large field intensities as discussed below, at least for intensities larger than the coupled continuous power of few milliwatts currently used in our experiments.

4. IRREVERSIBLE RELATIVISTIC THERMODYNAMICS

4.1 Covariant non equilibrium thermodynamics

To gain a deeper understanding of the high-frequency background in photonics waveguides, it is enlightening to consider the phenomenological covariant non-equilibrium thermodynamics theory developed by Werner Israel²⁰ in the framework of relativistic hydrodynamics. This theory provides a clear justification for introducing the heat flux as a state variable in extended irreversible thermodynamics. In addition, it reveals the mutual coupling between the dipole moment field and the heat flux insofar as the covariant approach offers a unified framework to deal with the thermal field and the electromagnetic field.

The theory mainly considers a simple relativistic fluid from a Lorentz frame but can be extended to continuous media. The dynamics of an elementary part of the fluid, called here particle, is characterized by its world line x^μ and its four-momentum p^μ with μ running from 0 to 3 and indexing the axis of the four-dimensional space-time. We assume a metric tensor $g_{\mu\nu} = \text{diag}(-1,1,1,1)$. In the rest-frame, the fluid particle is located by $x^\mu = (ct, x, y, z)$. In equilibrium state, the particle flux four-vector N^μ and the energy-momentum tensor $T^{\lambda\mu}$ provide a complete description of a simple fluid. The particle flux four-vector is defined as $N^\mu = nu^\mu$ with n the particle density and u^μ an arbitrary four-velocity of an observer. The energy-momentum $T^{\lambda\mu}$ is defined as the tensor that determines the elementary four-momentum dp^λ that crosses an arbitrary oriented elementary three-area $d\Sigma_\mu$ ²¹. In particular, the observer measures an energy density $\rho = u_\lambda u_\mu T^{\lambda\mu}$ and an energy flux $q = -u_\lambda T^{\lambda\mu} n_\mu$ along a unit spatial vector n_μ in his frame. At equilibrium the heat flux $q = 0$ in the frame of the observer. The quantities N^μ and $T^{\lambda\mu}$ can be interpreted in the framework of a kinetic theory as the first and second moments of a distribution function $N(x^\mu, p_\nu)$ corresponding to a generalization of the phase density of the ensemble of fluid particles in a Riemannian space-time.

In Israel's approach an arbitrary local state of the fluid is specified by the quantities $N^\mu, T^{\lambda\mu}$ and the entropy flux four-vector S^μ , which are qualified as primary variables. The entropy flux four-vector is given by $S^\mu = Su^\mu$ where S is the entropy density. The theory postulates the usual conservation laws of the particles $N_{,\mu}^\mu=0$ and of the momentum and of the energy $T_{,\mu}^\mu=0$, and the positivity of the entropy production $S_{,\mu}^\mu \geq 0$, in which the coma indicates covariant differentiation and the conventional Einstein summation rule is used. For the equilibrium state that is assumed to be specified by n, ρ and u^μ , the standard equation of state $S(n, \rho)$ results in a covariant Gibbs relation $dS_{eq}^\mu = -\alpha dN_{eq}^\mu + \beta_\lambda dT_{eq}^{\lambda\mu}$ for an arbitrary displacement of the primary variables inside the space of equilibrium states. It follows that the space of equilibrium state $(N_{eq}^\mu, T_{eq}^{\lambda\mu}, S_{eq}^\mu)$ can be parametrized by a thermal potential α [(relativistic chemical potential)/T] and an inverse-temperature four vector $\beta_\lambda = (c^2/kT)u_\lambda$.

For small deviation from an equilibrium state (α, β_λ) and toward a state outside the space of equilibrium states, it is assumed that the covariant Gibbs relation is still valid for $(N^\mu, T^{\lambda\mu}, S^\mu)$. The Gibbs equation is not constrained to the space of equilibrium state anymore. Note that off-equilibrium the quantities $N^\mu, T^{\lambda\mu}, S^\mu$ specify any arbitrary state of the fluid but do not provide a complete description for which all the moments of the distribution function $N(x^\mu, p_\nu)$ are needed. Using a Taylor expansion of the off-equilibrium entropy flux four-vector and imposing the positivity of the entropy production, leads to dynamical equations for the small perturbation quantities appearing in N^μ and $T^{\lambda\mu}$, such as for instance the heat flux $q = -u_\lambda T^{\lambda\mu} n_\mu$, the bulk stress Π and the shear viscous stress tensor $\pi^{\lambda\mu}$. For a Taylor development to the first order, the standard constitutive laws, namely the Fourier law, the Navier-Stokes equations are retrieved²². The transient relativistic thermodynamics theory improves the description by extending the Taylor expansion to the second order:

$$S^\mu - S_{eq}^\mu = -\alpha(N^\mu - N_{eq}^\mu) - \beta_\lambda(T^{\lambda\mu} - T_{eq}^{\lambda\mu}) + Q^\mu + \dots, \quad (5)$$

where Q^μ is a quadratic function of not only the perturbation quantities appearing in N^μ and $T^{\lambda\mu}$ but of all the perturbation quantities appearing in higher order tensors required for a complete description of the state of the fluid. Considering small deviation from an equilibrium state, the term Q^μ can be written in a general quadratic form involving only the perturbation quantities appearing in N^μ and $T^{\lambda\mu}$ and five new phenomenological coefficients $\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3$ that depend on the specification (n, ρ) of the equilibrium state. The condition of the positivity of the entropy production results in the following new set of constitutive equations written here in the fluid particle frame:

$$\Pi = -\frac{1}{3}\zeta_V(u_{,\mu}^\mu + \beta_1\dot{\Pi} - \alpha_1 q_{,\mu}^\mu) \quad (6a)$$

$$q^\lambda = -\kappa T \Delta^{\lambda\mu} (T^{-1} \partial_\mu T + \dot{u}_\mu + \beta_2 \dot{q}_\mu - \alpha_1 \partial_\mu \Pi - \alpha_2 \pi_{\mu,\nu}^\nu) \quad (6b)$$

$$\pi_{\lambda\mu} = -2\zeta_S (u_{<\lambda|\mu>} + \beta_3 \dot{\pi}_{\lambda\mu} - \alpha_2 q_{<\lambda|\mu>}) \quad (6c)$$

with κ the thermal conductivity, ζ_V and ζ_S the bulk and shear viscosities, $\Delta^{\lambda\mu} = g^{\lambda\mu} + u^\lambda u^\mu$ the operator projecting onto the three-space orthogonal to an arbitrary unit time-like vector u^λ , and the angular brackets enclosing a pair of indices denoting the symmetrized trace-free spatial projection. The coefficients α_i are coupling constants between the heat flux and the viscous stresses whereas the coefficients $\frac{1}{3}\zeta_V\beta_1, \kappa T\beta_2$ and $2\zeta_S\beta_3$ are three relaxation times. When the heat flux is the predominant perturbation quantity, Eq. 6b is simply the Cattaneo equation with the heat relaxation time $\tau = \kappa T\beta_2$. There is an extra term \dot{u}_μ of purely relativistic nature that Carl Eckart was the first to pinpoint in the relativistic Fourier

law, i.e. in the first order relativist theory of a simple fluid²². The relaxation times have been determined only in the case of a simple gas. In particular, the heat relaxation time is given by $\tau = \frac{2}{5}(\kappa T/nmc^4)(mc^2/kT)^2$ in the non-relativistic regime, with c the speed of light, k the Boltzmann constant and m the masse of the particle³. In the general case of a continuous medium, we are not aware of any simple formula for estimating the phenomenological parameter, in particular in the case of amorphous solids. In¹⁹, measuring the value of the background signal, we have estimated the value of the heat relaxation time in silicon nitride waveguides between 0.4ns and 1.4ns based on the assumption that the bulk stress Π and the shear viscous stress tensor $\pi^{\lambda\mu}$ are negligible.

The Israel's theory leads to phenomenological relaxations times and coupling constants for determining the dynamics of the quantities that are necessary to describe a non-equilibrium state. It does not enable a quantitative comparison with the experimental signals. This point stands for any theory of non-equilibrium thermodynamics as there is currently no fundamental microscopic description of irreversible non-equilibrium processes. However, the phenomenological approach has the merit of predicting general physical behaviors such as the background that currently limiting the performances of on-chip waveguide-based Raman sensors.

4.2 Irreversible thermodynamics in dielectric continuous media subject to an optical field

Decoupling the heat flux and the perturbation of the electrical polarization P^λ for an off-equilibrium state as in¹⁹ makes the implicit assumption that the perturbation $P_{(1)}^\lambda = P^\lambda - P_{eq}^\lambda$ is negligible. It is however an open question for which applied electromagnetic field (E_μ, B_μ) this can be considered true. The relativistic approach of the transient thermodynamics effect enables to include electromagnetic field in a consistent way in continuous polarizable media²³ and to highlight the interplay between the heat flux and the electromagnetic field. Following a phenomenological approach similar to the case without electromagnetic field, the constitutive equations for the heat flux and $P_{(1)}^\lambda$ take the following form when the bulk stress and the shear viscous stress tensor and the perturbation of the magnetic polarization are negligible:

$$q^\lambda = -\kappa T \Delta^{\lambda\mu} (T^{-1} \partial_\mu T + \dot{u}_\mu + \beta_2 \dot{q}_\mu + \beta_4 \dot{P}_\mu^{(1)}) \quad (7a)$$

$$P_{(1)}^\lambda = -\chi \Delta^{\lambda\mu} (\mathcal{L}_u E_\mu - u_{(\mu|\nu)} E^\nu - \beta_2 \dot{q}_\mu - \beta_3 \dot{P}_\mu^{(1)}) \quad (7b)$$

with \mathcal{L}_u the Lie derivative, χ, β_3, β_4 extra phenomenological constants, and the parentheses enclosing a pair of indices denoting the antisymmetric part of the tensor. The constant χ is related to the classical electrical susceptibility. These equations reveal the local interplay between the electromagnetic field and the heat flux. Depending on values of the relaxation time constants and on the coupling constant for non-negligible bulk stress and the shear viscous stress tensor, it might be possible to find operating conditions, in particular applied optical field intensity, for which the background is minimal. To confirm this idea a direct connection of this theory of transient irreversible thermodynamics in polarizable continuous media with the optical spectrum of the mode propagating in the waveguide needs to be developed.

To the first order the divergence of the entropy is given by the sum of the product of the forces times the corresponding flux: $S_\mu^\mu = \text{sum of flux} \times \text{forces}$. The set of equations 6) or 7) assume linear relations between the ‘‘forces’’, e.g. $(T^{-1} \partial_\mu T + \dot{u}_\mu)$ or $(\mathcal{L}_u E_\mu - u_{(\mu|\nu)} E^\nu)$, and the ‘‘flux’’, e.g. q^λ or $P_{(1)}^\lambda$. From a certain intensity level of the optical pump field, such a linear approximation does not hold anymore, and non-linear optical effects have to be considered. We are not aware of any theory tackling this problem in a satisfying way. A deep understanding of the interplay between non-linear optical effects and irreversible thermodynamics is of interest for improving not only the performances of on-chip Raman sensors but also the ones of others on-chip devices such as squeezed states sources²⁴⁻²⁶.

5. APPROACHES TO GET AROUND THE BACKGROUND

5.1 Local enhancement of the analytes with mesoporous coatings

One method to boost the molecular Raman signal P_{out}^R is to enrich the analyte locally in the evanescent field with a sorbent layer. For example, mesoporous silica is a widely used, high-surface-area sorbent that has good versatility in surface chemistry. In²⁷, we combined silicon nitride slot waveguide with hexamethyldisilazane (HMDS) modified mesoporous silica. The methyl groups introduced to silica surface make the sorbent hydrophobic, leading to considerable enrichment

of non-polar compounds from aqueous environments. For instance, we have demonstrated 600x enrichment of 970 μ M cyclohexanone from water. From the nonlinear relation between enrichment factor and applied concentration²⁸, there should be even stronger enrichment for cyclohexanone concentration lower than 970 μ M in water.

Despite the perturbation introduced by the sorbent, the overall background intensity is not noticeably increased. Therefore, by using a suitable sorbent layer on top of the dielectric waveguide, we can boost the analyte signal without affecting $P_{\text{out}}^{\text{B}}$. The signal to background noise ratio can be increased by two orders of magnitude.

5.2 Resonant Raman scattering

The molecular Raman signal and the background collected at the output port of the waveguide vary both with the pump wavelength λ_0 according to $1/\lambda_0^2$. Consequently, in contrast to standard spontaneous scattering there is no significant advantage to minimize the pump wavelength, except when resonant Raman scattering kicks in.

A paramount of biological molecules, such as amino acids, absorb at UV wavelengths, which enables to boost the Raman scattering process by using resonant effects. Resonant Raman scattering cross-section are more than three order of magnitude large than the spontaneous Raman cross section. As the amorphous material constituting the waveguide is not expected to exhibit any resonant effects, operating the sensor at UV wavelengths should lead to at least a 10^3 -fold signal enhancement for many relevant molecules. The challenge lies in the fabrication of low-loss single mode waveguides in this wavelength range. We have recently fabricated single mode waveguides made of aluminum oxide, the propagation loss of which are 3dB/cm at a pump wavelength of 360nm. It is a promising platform for implementing on-chip resonant Raman sensing.

5.3 Crystalline waveguides

The extended irreversible thermodynamics theory is well suited for fluids or amorphous and polycrystalline solid. The only difference between an amorphous or polycrystalline solid and a fluid is the occurrence of a pressure tensor in the classical Gibbs equation for the first medium and a scalar pressure for the last one²⁹. It explains why hydrodynamic approaches are well-suited to describes thermal flux in amorphous waveguides. As regards crystalline structures, long-range interactions cause difficulties for using the concept of a local entropy. Crystalline media needs to be considered as a whole and consequently boundaries play an important role in contrast to amorphous and polycrystalline solid. Comparing the background of integrated waveguides made of crystalline material such as diamond or aluminum nitrides that are currently of technological interest and the background of amorphous material could help to better understand the dynamics of the thermal, mechanical and electromagnetic fields in photonics waveguides. We expected that the background of crystalline waveguides will be resonantly enhanced for some frequency bands due to the scattering by polariton modes resulting from the long-range interaction³⁰, for the benefit of other frequency bands if the boundary effects do not produce the opposite effect. Using crystalline waveguides might be an advantage for the detection of specific analytes in terms of signal to background noise ratio.

6. CONCLUSION

The strong electromagnetic field confinement in integrated single mode waveguide made of high-index contrast materials has unveiled the presence of an unexpected thermal noise contribution at Terahertz frequencies. This noise that results in an exponentially decaying background in the optical spectrum of response of the system sets the current detection limit of on-chip waveguide-based Raman sensors.

The covariant thermodynamics theory developed by cosmologists allows us to understand the physical mechanism at the origin of the background and hence the mechanism involved in the limitation of the performances of the sensor. Applying this theory to a propagating guided mode could reveal interesting operating regimes and help to identify the best settings and waveguide geometries to push the detection limit. Technological improvements are still possible to push further the detection limit toward sensors impacting society, not only in the framework of the analysis of liquid solutions but also for greenhouse-gas sensing³¹ and warfare agent detection³².

The technological maturity of integrated photonics has reached a physical regime where theories developed for cosmological objects such as neutron stars or high-energy matter start to be relevant to interpret the experimental data and gain a deep understanding of the fundamental principles at play. There is even a potential to explore advanced cosmological models with physical system that can be engineered. We think that the connection between the physics of relativistic fluids and the properties of nanophotonic circuits can benefit both fields. On one side all models developed in the former domain can allow us to quantify and predict the optical properties of high-frequency light field in strongly confined media. On the other side the very accessible and tractable experimental investigation of light matter interaction in integrated photonic systems in contrast to high-energy physics and cosmological objects can help to refine existing theories or developed advanced ones more in line experimental facts. It is finally interesting to note that a curiosity at the theoretical level could be crucial for reaching the ultimate performance of integrated sensors.

REFERENCES

- [1] Dhakal, A., Subramanian, A. Z., Wuytens, P., Peyskens, F., Le Thomas, N., Baets, R., “Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides,” *Opt. Lett.* 39, 4025 (2014).
- [2] Dhakal, A., Wuytens, P., Raza, A., Peyskens, F., Subramanian, A.Z., Clemmen, S., Le Thomas, N. and Baets, R., “Lab-on-a-chip Raman sensors out- performing Raman microscopes,” *CLEO, United States*, paper SM2O.3 (2016).
- [3] Israël, W., Stewart, J.M., “Transient Relativistic Thermodynamics and Kinetic Theory,” *Ann. Physics* 118, 341-371 (1979).
- [4] Peyskens, F., Dhakal, A., Van Dorpe, P., Le Thomas, N., Baets, R., “Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform,” *ACS Photonics*, 3(1), 102-108 (2016).
- [5] Raza, A., Clemmen, S., Wuytens, P.C., Muneeb, M., Daele, M. V., Dendooven, J., Detavernier, C., Skirtach, A., Baets, R., “ALD assisted nanoplasmonic slot waveguide for on-chip Enhanced Raman Spectroscopy,” *Applied Physics Letters - Photonics*, 3(11), 116105 (12 pages) (2018).
- [6] Dhakal, A., Raza, A., Peyskens, F., Subramanian, A., Clemmen, S., Le Thomas, N., Baets, R., “Efficiency of evanescent excitation and collection of spontaneous Raman scattering near high index contrast channel waveguides,” *Optics express*, 23(21), 27391-27404 (2015).
- [7] Raza, A., Clemmen, S., Wuytens, P.C., de Goede, M., Tong, A., Le Thomas, N., Liu, C., Suntivich, J., Skirtach, A., Garcia-Blanco, S., Blumenthal, D., Wilkinson, J., Baets, R., “High index contrast photonic platforms for on-chip Raman spectroscopy,” *Optics Express*, 27(16), p.23067-23079 (2019).
- [8] Landau, L. D., Lifshitz, E. M., [Statistical Physics, Part 1], 2nd ed. (Pergamon, 1980), Section 114.
- [9] Rytov, S. M., “Correlation theory of thermal fluctuations in an isotropic medium,” *J. Exp. Theor. Phys.* 33, 166-178 (1957).
- [10] Foster, S., Tikhomirov, A., Milnes, M., “Fundamental thermal noise in distributed feedback fiber lasers,” *IEEE J. Quantum Electron.* 43, 378-384 (2007).
- [11] Wanser, K. H., “Fundamental phase noise limit in optical fibres due to temperature fluctuations,” *Electron. Lett.* 28, 53-54 (1992).
- [12] Annovazzi-Lodi, V., Donati, S., Merlo, S., “Thermodynamic phase noise in fibre interferometers,” *Opt. Quantum Electron.* 28, 43-49 (1996).
- [13] Bartolo, R. E., Tveten, A. B., Dandridge, A., “Thermal phase noise measurements in optical fiber interferometers,” *IEEE J. Quantum Electron.* 48, 720-727 (2012).
- [14] Gorodetsky, M.L., Grudinin, I.S., “Fundamental thermal fluctuations in microspheres,” *J. Opt. Soc. Am. B* 21, 697-705 (2004).
- [15] Saurav, K., Le Thomas, N., “Probing the fundamental detection limit of photonic crystal cavities,” *Optica*, 4(7), 757-763 (2017).
- [16] Panuski, C., Englund, D., Hamerly, R., “Fundamental Thermal Noise Limits for Optical Microcavities,” *Phys. Rev. X* 10(4), 041046(28) (2020).
- [17] Müller, I., “Zum Paradoxon der Wärmeleitungstheorie,” *Zeitschr. f. Physik*, 198, 329-344 (1967).
- [18] D. Jou, J. Casas-Vázquez, and G. Lebon, [Extended Irreversible Thermodynamics], 4th ed. Springer, (2010).

- [19] N. Le Thomas, A. Dhakal, A. Raza, F. Peyskens, R. Baets, "Impact of fundamental thermodynamic fluctuations on light propagating in photonic waveguides made of amorphous materials," *Optica*, 5(4), 328-336 (2018).
- [20] Israël, W., "Nonstationary Irreversible Thermodynamics: A Causal Relativistic Theory," *Ann. Physics* 100, 310-331 (1976).
- [21] Synge, J.L., [Relativity: The Special Theory], North-Holland Publishing Company, Amsterdam (1956)
- [22] Eckart, C., "The Thermodynamics of Irreversible Processes III. Relativistic Theory of the Simple Fluid," *Phys. Rev.* 58, 919-924 (1940).
- [23] Israël, W., Stewart, J.M., "Progress in Relativistic Thermodynamics and Electrodynamics of Continuous Media," in *General Relativity and Gravitation* (ed. A. Held) Plenum, New York, Vol. 2, 491-525 (1980).
- [24] Vaidya, V. D., Morrison, B., Helt, L. G., Shahrokshahi, R., Mahler, D. H., Collins, M. J., Tan, K., Lavoie, J., Repington, A., Menotti, M., Quesada, N., Pooser, R. C., Lita, A. E., Gerrits, T., Nam, S. W., Vernon, Z., "Broadband quadrature-squeezed vacuum and nonclassical photon number correlations from a nanophotonic device," *Sci. Adv.*, 6 : eaba9186 (2020).
- [25] Cernansky, R., Politi, A., "Nanophotonic source of quadrature squeezing via self-phase modulation," *APL Photon.* 5, 101303 (2020).
- [26] Y. Zhao, Y. Okawachi, J. K. Jang, X. Ji, M. Lipson, A. L. Gaeta, "Near-Degenerate Quadrature-Squeezed Vacuum Generation on a Silicon-Nitride Chip," *Phys. Rev. Lett.* 124(19), 193601(7) (2020).
- [27] Liu, Z., Zhao, H., Baumgartner, B., Lendl, B., Stassen, A., Skirtach, A., Le Thomas, N, Baets, R., "Ultra-sensitive slot-waveguide-enhanced Raman spectroscopy for aqueous solutions of non-polar compounds using a functionalized silicon nitride photonic integrated circuit", *Opt. Lett.*, accepted (2021).
- [28] Foo, K.Y., Hameed, B.H., "Insights into the modeling of adsorption isotherm systems", *Chem. Eng. J.* 156, 2-10 (2010).
- [29] de Groot, S.R., Suttrop, L.G., [Foundations of Electrodynamics], North-Holland Publishing Company, Amsterdam (1972).
- [30] Barker, Jr., A.S., Loudon, R., "Response Functions in the Theory of Raman Scattering by Vibrational and Polariton Modes in Dielectric Crystals," *Rev. Mod. Phys.* 44(1), 18-47 (1972).
- [31] Holmstrom, S. A., Stievater, T. H., Kozak, D. A., Pruessner, M. W., Tyndall, N., Rabinovich, W. S., McGill, R. A., Khurgin, J. B., "Trace gas Raman spectroscopy using functionalized waveguides," *Optica* 3, 891 (2016).
- [32] Tyndall, N. F., Stievater, T. H., Kozak, D. A., Koo, K., McGill, R. A., Pruessner, M. W., Rabinovich, W. S., Holmstrom, S. A., "Waveguide-enhanced Raman spectroscopy of trace chemical warfare agent simulants," *Opt. Lett.* 43(19), 4803-4806 (2018).