A Review of Raman Thermography For Electronic and Opto-Electronic Device Measurement with Sub-micron Spatial and Nanosecond Temporal Resolution

M. Kuball, Member, IEEE, and J. W. Pomeroy.

Abstract—We review the Raman thermography technique, which has been developed to determine the temperature in and around the active area of semiconductor devices with submicron spatial and nanosecond temporal resolution. This is critical for the qualification of device technology, including for accelerated lifetime reliability testing and device design optimization. Its practical use is illustrated for GaN and GaAs based high electron mobility transistors, and opto-electronic devices. We also discuss how Raman thermography is used to validate device thermal models, as well as determining the thermal conductivity of materials relevant for electronic and opto-electronic devices.

Index Terms—Thermography, thermal simulation, GaN, GaAs, HEMT, thermal management, reliability.

I. INTRODUCTION

Electronic and opto-electronic semiconductor devices have advanced greatly in recent decades by incorporating new materials and designs to achieve increased speed, decreased size and improved efficiency. This has enabled higher power densities at higher frequencies for electronic devices, and further wavelength ranges or increased optical output power for opto-electronic devices. The heterogeneous integration of complimentary materials has also increased device performance and functionality, giving a wide range of technological design options [1]. However, an ever increasing power dissipation density has resulted in thermal management challenges. It is critically important to be able to accurately determine device temperatures in order to assess the reliability of new technologies, because temperature is one of the dominant drivers of device degradation and influences device performance [2-7]. Phonon transport at the nanoscale and at interfaces impacts how waste heat is transported away from the active region in devices. The thermal properties of some “new” and “old” materials (e.g. SiC, GaN, diamond, graphene) are also often not as well-known as commonly thought, depending on the growth conditions or growth methods used. Relying on device thermal simulations alone, without experimental verification, can therefore potentially limit the accuracy achieved when predicting operating temperatures in electronic and opto-electronic devices, negatively impact device design or reliability.

Long term reliability is a concern with any new technology and qualification is one part of the process towards commercialization. For example, GaN high electron mobility transistors (HEMTs) are presently being developed and commercialized for both microwave and power electronics applications. A corresponding processes has happened before for Si, GaAs and InP, and will happen again for future materials and device systems. Part of this qualification is temperature accelerated lifetime testing which relies on detailed knowledge of channel temperatures which are used to predict the mean time to failure (MTTF) at the designed operating temperature. Current standards, such as JEDEC [7] for the determination of channel temperature, still mostly rely on “legacy” thermal characterization approaches, such as IR thermography, which often can no longer be applied with high accuracy in high power density devices due to their intrinsically limited spatial resolution and the resulting underestimation of the device peak temperature [8]. For example, temperature gradients as high as 100 °C/μm can be present close to the device channel of GaN-based RF HEMTs [9].

The question also arises - Which temperature is relevant for device reliability assessment? This has not been discussed extensively in the past because the spatial resolution limitations of existing thermography techniques has made it impossible to measure close to the device channel and the temperature profile directly. For example, it is important to identify that not only the peak channel temperature near the gate contact in a HEMT is a critical parameter, but also other lower temperature regions

This work is in part supported by the UK Engineering and Physical Sciences Research Council (EPSRC), the European Space Agency (ESA), the European Defense Agency (EDA), the US Office for Naval Research (ONR) and the Defense Advanced Research Projects Agency (DARPA) under Contract FA8650-15-C-7517 monitored by Dr. Avram Bar Cohen, supported by Dr. John Blevins, Dr. Joseph Maurer and Dr. Abirami Sivananthan. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA.

M. Kuball and J. W. Pomeroy are with H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, U.K. (e-mail: martin.kuball@bristol.ac.uk).

Copyright © 2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained by sending a request to pubs-permissions@ieee.org.
of the device may result in device degradation, such as of the ohmic source and drain contacts. Pulsed operation is also used in many applications, so any generally applicable thermography technique must combine high spatial resolution with high temporal resolution to be able to record fast temperature transients, which cannot be achieved using conventional IR thermography. In this paper, we review the Raman thermography technique which enables submicron spatial resolution and nanosecond temporal resolution measurements in semiconductor electronic and opto-electronic devices. Following a discussion of existing thermography techniques, the general principles of Raman thermography are given together with examples of applications.

II. SEMICONDUCTOR DEVICE TEMPERATURE MEASUREMENT TECHNIQUES

Various electrical characterization based methods have been used extensively to assess the channel temperature in electronic devices. The equipment needed for these temperature assessments is available in standard electrical testing laboratories. Different methods are used, all based on measuring changes in temperature dependent electrical parameters, including: Saturated drain current [10-14]; gate leakage current and threshold voltage, among other parameters. These techniques offer no spatial resolution as such, but instead are sensitive to the entire periphery of the device channel [14]. As a consequence, electrical methods tend to underestimate the peak device temperature. For example in a HEMT, temperature may be averaged not only along the whole gate width, but also typically between source and drain contact. Because the temperature measurement location is not well defined, it is can be difficult to relate the measured temperatures to a thermal simulation in a meaningful way. For opto-electronic devices, light emitting diodes and laser diodes, the emission wavelength is temperature dependent, following the semiconductor bandgap temperature dependence. Measuring the emission wavelength is therefore a convenient method for temperature estimation in the active region of opto-electronic devices [15]. Naturally this technique can only provide temperature information in areas that emit light, whereas obtaining the temperature distribution in other parts of the device structure may also be required to assess the full benefits of the device design. Alternatively, the temperature-induced shift in lasing threshold voltage can be used to measure the temperature in opto-electronic devices [16], although this technique lacks spatial resolution, in the same way as when electrical methods are applied to transistors.

IR thermography is the most commonly used technique for semiconductor device and circuit technology temperature measurement in industry [8, 17, 18]. This technique is based on measuring the thermal radiation emitted from the surface of a device, typically detected in the 3 – 10 µm spectral range, depending on the detector used. The radiated light intensity (radiance) scales as $T^4$, following the Stephan-Boltzmann law, where $T$ is the surface temperature in Kelvin. Therefore, temperature can be determined by measuring radiance, after calibrating the surface emissivity of the device under test (DUT) to take into account that the measured surface is not a perfect black body (emissivity = 1). Emissivity can be calibrated by measuring the radiance of the DUT at single or multiple well defined temperatures. Calibration must be performed on a pixel-by-pixel basis because the emissivity typically varies across the surface of a device, which has regions of exposed semiconductor, passivation and different metal contacts. Sample movement during the emissivity calibration is one potential source of artifacts in IR thermography temperature maps, particularly for transistor structures which have features of a similar or smaller length scale than the IR camera resolution; such artifacts can be particularly apparent at the edges of structures.

The achievable spatial resolution limit of IR thermography is typically on the order of the wavelength measured. Therefore, the long wavelength of the measured IR emission is a limitation when measuring devices with micron or sub-micron scale feature sizes. The effect of lateral spatial averaging on measured device temperatures must be carefully considered when interpreting results. For example, temperature is often averaged over the whole source–drain gap when a HEMT is measured. Furthermore, many semiconductor materials are transparent at the wavelength of the emitted IR thermal radiation (unless heavily doped) and radiation originating from below the surface contributes to the thermal image, including from the die attach/wafer backside. The resulting lateral and depth temperature averaging can cause the temperature measured using IR thermography to be significantly lower than the actual peak temperature [8, 19], as illustrated in Figure 1 for a GaN HEMT. Another challenge is that metal surfaces tend to have rather low emissivity which makes temperature determination of metal areas such as contacts challenging, unless measurements are performed at elevated background temperatures. Low emissivity or transparent surfaces can be addressed by covering the device surface with an opaque high emissivity black ‘paint’. However, relatively thick paint may not truly represent the surface temperature and may be impossible to fully remove from the device after the measurement [8]. It should be possible in theory to compare the
data obtained by IR thermography to a device thermal simulation, considering all the factors previously described. Attempts have been made to do this [20], although whether this approach can be used to accurately determine the peak temperature in all devices is controversial, particularly in devices with very high local temperature gradients such as GaN RF HEMTs. Although peak temperature accuracy may be limited by spatial averaging in IR thermography, it is a very efficient method for mapping the temperature over large surface areas, making this technique especially suitable for comparative screening of identical devices where exact knowledge of peak temperatures may not be essential. Problem areas or “hot spots” can be identified quickly, such as those caused by local leakage currents or power imbalances in monolithic microwave integrated circuits (MMICs). Integrating fast low resolution IR temperature mapping with slower high resolution serial temperature measurement approaches can therefore be beneficial and a combined Raman-IR thermography measurement system has been developed for this purpose, as described in [8].

Other optical techniques have been demonstrated for electronic and opto-electronic device thermography. Micro-photoluminescence (PL) measurements can be used to measure the semiconductor surface temperature in active devices [21], exploiting the temperature dependence of the band edge emission. PL is a laser-based technique offering submicron spatial resolution, although it should be noted that electron hole pairs generated by the above bandgap laser illumination can modify the current/field distribution in the device being measured. This additional photo-induced current can cause a significant additional device heating, which often dominates even the direct laser heating induced temperature rise [22]. Reflectivity changes can also be used to optically probe the device temperature, either considering only the change in reflectivity [23, 24], or additionally by measuring the phase shift of the reflected light [25, 26]. Calibration of the thermoreflectance coefficient can be challenging as the change in reflectivity with temperature is often small, however it is possible by performing careful measurements of specially prepared samples [27]. Liquid crystal thermography is another optical technique and relies on the known phase transition temperatures of liquid crystals [28]. However, it is less convenient and no longer commonly used; multiple depositions of liquid crystals with different phase transition temperatures are needed to map each temperature contour.

Raman thermography is the focus of this review paper. The use of Raman spectroscopy for temperature measurements was first demonstrated on Si devices [29], but has been more fully exploited following the advancement of GaN electronic devices [30], and has since been demonstrated extensively by several groups [31-35]. This technique is often combined with finite element (FE) thermal simulation to determine the channel temperature, as well as the measurement of material thermal properties used for device thermal simulations [36-38]. Figure 1 shows a comparison of temperatures measured in the same device using Raman thermography and IR thermography. The improvement in spatial resolution is clearly apparent, with a lateral resolution of 0.5 µm for Raman thermography and about 7 µm in this case for IR thermography. This clearly illustrates the benefit of using a higher spatial resolution technique for the thermal analysis of transistors in particular, where the heat generating regions tend to be on the micrometer or smaller length scale.

In the following sections we will discuss this technique in detail, including its practical application with examples. This optical technique exploits the temperature dependent properties of quantized lattice vibrations (phonons) of the materials in electronic or opto-electronic devices. As it is a laser based microscopy technique, it can provide temperature mapping with sub-micron lateral spatial resolutions as high as 0.5 µm (Figure 1), and even smaller if combined with advanced optical focusing concepts such as solid immersion lenses (SILs). One advantage of Raman thermography in comparison to PL is that sub-band gap lasers can be used, avoiding light absorption in the device which could impact its operation and also enabling three dimensional mapping through transparent materials. In addition to spatial mapping, the temperature transients can also be measured by modulating the probing laser, with nanosecond or better temporal resolution depending on the modulation scheme used.

III. PRINCIPLES OF RAMAN THERMOGRAPHY

A. Spatial Resolution

Raman thermography is a laser-based microscopy measurement technique. Laser light is focused through a microscope objective lens onto the region of interest and backscattered light is usually collected for analysis using the same lens, as illustrated in Figure 2(a). By the scanning focused laser spot over the DUT, temperature maps can be obtained by measuring each point sequentially. For example, the channel temperature line profile in a HEMT can be measured by scanning laterally in one dimension (Figure 1), whereas by scanning in two dimensions (2D) surface temperature maps can be obtained which may be used to probe the temperature distribution around defects or other features [39]. As with any
measurement technique, it is important to take into account spatial averaging when analyzing temperatures measured by Raman thermography, which is determined by the optical resolution of the measurement system. Two scenarios need to be distinguished when assessing the axial measurement resolution of Raman thermography: 1) When the probing laser photon energy is below the bandgap of the materials in the active area of the device; 2) when above-bandgap excitation is used. We consider the former case first, when photon absorption, heating and photo-induced current is negligible. This is discussed for the example of an AlGaN/GaN HEMT, as due to its bandgap of 3.4 eV, the GaN layer and substrate, with the exception of silicon, are transparent for most of the laser sources used in Raman systems. In this case, Raman thermography measures a spatially averaged temperature in a finite volume in the GaN layer and substrate, as illustrated in Figure 2(a). The transparency of the measured materials can also be exploited by focusing the probing laser below the surface, allowing 3D temperature mapping.

The diffraction limited spatial resolution of a scanning confocal laser microscope is defined as the elliptical half intensity radius of the point spread function (PSF), which depends on the numerical aperture (NA) of the objective and the laser wavelength ($\lambda$) and is independent of the magnification. The lateral resolution is given by

$$r_{lateral} = \frac{0.51 \cdot \lambda}{NA},$$

and the axial resolution, related to the depth of field (DoF), is given by

$$r_{axial} = \frac{0.88 \cdot \lambda}{(n - \sqrt{n^2 - NA^2})},$$

where $n$ is the refractive index of the focusing medium [40]; for $NA < 0.5$, (2) can be approximated by $1.77 \cdot n \cdot \lambda / NA^2$. Although these values can be improved slightly by using a highly confocal measurement system with a reduced aperture size [40], the NA of the objective lens is the main parameter that can be increased to enhance spatial resolution [41], and especially to improve the axial resolution. The maximum NA limit is equal to the refractive index of the immersion medium, which is $NA_{\text{max}} = n_{\text{air}} = 1$ for a conventional air immersion microscope, although the practically achievable NA is lower than the theoretical limit due to practical restrictions such as the objective lens minimum working distance. It is also important to note that (1) and (2) are only applicable for diffraction limited light focusing, which is the case close to the semiconductor/air interface where optical aberration can be neglected [42].

Raman spectroscopy is material selective, so that in a transparent layered structure, the characteristic phonons of each material can used to simultaneously probe the temperature in these layers. For example, a typical Raman thermography measurement using a 488 nm laser focused near the surface of a GaN device ($N = 2.4$), through an 0.5 NA objective lens, will achieve a resolution of $r_{lateral} \approx 0.5 \mu m$ and $r_{axial} \approx 8 \mu m$, calculated using (1) and (2). In this case, the depth of field is much larger than the 1-2 $\mu m$ thickness of the GaN layer and the measured temperature in that layer can be considered to be a simple depth average through this layer. A separate average through the upper portion of the SiC substrate may also be measured by considering the SiC Raman modes of the substrate, as illustrated in Figure 2(a), i.e., 3D temperature information can be obtained. This also applies if other transparent materials are being probed. Figure 3 shows the volume of GaN averaged in a typical Raman thermography measurement, compared with the simulated temperature distribution in an AlGaN/GaN RF HEMT. Although it is important to consider the effect of spatial averaging on the measured temperature, the Raman measured average temperature can be relatively close to the peak value.

As well as using the material specific phonon modes to measure the temperature in different layers of a structure separately, confocal Raman thermography can be used to map the 3D temperature distribution inside transparent materials. Focusing the laser through a thick transparent substrate, from the back side of the device, can also be advantageous because it enables the temperature to be probed underneath metal contacts, areas which would otherwise be obscured when measuring from the top side of the device (Figure 2(b)). However, spatial resolution may be aberration limited owing to refraction at the air/semiconductor interface and then (1) and (2) do not apply. In this scenario, marginal and on axis rays are focused at different depths, which affects both the lateral resolution and depth of focus, but especially the latter. The refraction limited DoF is then given by [42, 43]

$$r_{axial} = \frac{\lambda}{n} \left[ \left( \frac{NA^2(n^2-1)}{1-NA^2} + n^2 \right)^{1/2} - n \right],$$

which increases linearly with depth ($z$). In a GaN-on-SiC device for example, $r_{axial}$ can be as large as 30 $\mu m$ when focusing through a 300 $\mu m$ thick SiC substrate. This increased spatial averaging needs to be considered when analyzing Raman temperature measurement results.
induced by refractive index mismatch can be compensated for to some extent by using an objective lens with an adjustable coverslip correction, adaptive optics or by using solid immersion lenses (SILs) [41].

When Raman thermography measurements are performed using above bandgap laser wavelengths, the axial spatial resolution is determined by the absorption depth of the incoming and scattered light, rather than the DoF, and is given by

\[ r_{axial} = \frac{1}{2\alpha}, \]

where \( \alpha \) the absorption coefficient of the material in the device; for example when GaAs is probed using a 488nm wavelength laser, the depth probed is around 50 nm. Raman thermography can be used in this way to perform surface sensitive measurements, with the proviso that the effect of laser light absorption on the measurement should be carefully minimized.

Raman thermography exploits the measurement of phonons and therefore the tested device should obviously contain materials with Raman active phonon and optical access. The direct temperature measurement of metal is not easily possible, such as the contacts, gate and field plate present on the topside of a HEMT for example; metals either exhibit no Raman active phonons or are experimentally very difficult to measure. Therefore, this technique is limited to measuring device areas where optical access is not blocked by metal contacts, either in openings between contacts or alternatively measurements can be performed by focusing the laser through the substrate backside, as discussed above, since the measurement can be performed using a sub bandgap laser, in contrast to PL spectroscopy. Although metal surfaces cannot be measured directly using Raman thermography, the application of Raman active micro-thermometers can be used to sense the temperature of these surfaces, including, but not limited to, diamond nanoparticles, TiO\(_2\) nanoparticles or silicon nanowires [44, 45].

\[ \Delta \omega = \Delta \omega_T + \Delta \omega_c. \]

The simplest and most commonly used Raman thermography temperature measurement approach is to use an empirical expression to describe the phonon frequency shift as a function of temperature [48]:
\[ \Delta \omega = \omega_0 - \frac{A}{\exp\left(\frac{\hbar \omega}{k_B T}\right) - 1}, \]

where \( \omega_0 \) is the phonon frequency at temperature \( T = 0 \)K and \( k_B \)

is the Boltzmann constant. \( A \) and \( B \) are empirical fitting parameters unique to each phonon and material; these are adjusted to match calibration data obtained by placing a sample in a temperature controlled cell and performing Raman scattering measurements. Figure 4(b) shows the GaN \( E_2 \) phonon temperature calibration results for epitaxial layers on different substrates, compared to freestanding bulk GaN. Bulk GaN exhibits the “true” phonon temperature dependence in the absence of mechanical strain, which is the result of the lattice expansion and phonon decay contributions which vary with temperature [46]. The phonon temperature dependence deviates from bulk values for hetero-epitaxial layers, such as GaN grown on different substrates, owing to the additional mechanical strain induced by the CTE mismatch between the epitaxial layer and the substrate. Because the calibration curves shown in Figure 4(b) include the effect of these additional strain contributions, it is recommended that this calibration is performed for each device epitaxy. Typical calibration parameters for GaN HEMT wafers are given in Table I. We note that the effect of additional strain induced during device operation, for example due to high temperature gradients or the converse / inverse piezoelectric effect in piezoelectric materials, may also need to be taken into account to ensure accurate temperature measurements, as discussed later.

So far we have only considered the use of the phonon energy (Raman shift) for temperature measurements. It is also possible to use the intensity of Raman peaks or their linewidth. Temperature can be determined from the relative intensity of the Stokes (I\( S \)) and Anti-Stokes (I\( AS \)) Raman peaks [49]:

\[ \frac{I_S}{I_{AS}} = \left( \frac{\lambda_S}{\lambda_{AS}} \right)^4 \exp \left( \frac{\hbar \omega}{k_B T} \right), \]

where \( \lambda_S \) and \( \lambda_{AS} \) are the wavelength of the Stokes and anti-Stokes scattered light, respectively. This approach is easier to apply to materials with lower phonon energy, such as GaAs or InP, than for higher phonon energy materials such as GaN, SiC and diamond which have a low anti-Stokes Raman scattering intensity in the temperature range relevant for device measurements. In practice, differences in light absorption, system throughput, and detector sensitivities at different frequencies/wavelength need to be taken into account when applying (7) [50]. The Stokes/anti-Stokes temperature measurement method has the advantage that it is rather insensitive to mechanical strain. As this technique is time consuming, requiring long integration times to measure the weaker anti-Stokes Raman line intensity, it is not widely used for device temperature measurements. It also possible to obtain the temperature from the linewidth of the Raman peaks as this is also mainly affected by temperature, but not greatly by strain [51]. However, the temperature dependence of the linewidth is rather small and higher signal-to-noise ratio spectra are therefore typically needed with respect to the phonon peak position temperature determination method. The phonon linewidth is also broadened by the finite spectral resolution of the measurement system used, which needs to be considered in such a measurement. For these reasons, measuring the phonon peak center position is the most commonly used Raman thermography method.

### C. Mechanical Strain and Accuracy

To ensure accurate Raman thermography measurements, we must consider how to separate the contributions of temperature and strain to the measured total energy shift of the Raman modes. In particular, when an additional mechanical strain is caused by device operation, but was not present during the temperature calibration. The phonon frequency shift \( \Delta \omega \) resulting from strain in GaN and other wurtzite structure crystals is given by [52]

\[ \Delta \omega = a \left( \varepsilon_{xx} + \varepsilon_{yy} \right) + b \cdot \varepsilon_{zz} + c \left( \varepsilon_{xx} - \varepsilon_{yy} \right)^2 + 4 \varepsilon_{xy}^2, \]

where \( a, b \) and \( c \) are the phonon deformation potentials [53], \( \varepsilon_{xx}, \varepsilon_{yy} \), and \( \varepsilon_{zz} \) are components of the strain tensor. We note that terms in (8) relating to shear strain are modified for other crystal symmetries [52]. Alternatively, (8) can be formulated in terms of mechanical stress \( \sigma \) by using the phonon deformation potentials per unit stress, \( a', b' \) and \( c' \). Only considering the phonon shift induced by biaxial strain and neglecting the second and third terms in Eq. (8), a valid assumption for most device measurements, this can be expressed as [52]:

\[ \Delta \omega = 2a' \varepsilon_{xx}. \]

where \( \varepsilon_{xx} = \varepsilon_{yy} \). In piezoelectric materials, including GaN and other III-nitride materials, bias voltages applied to a device generate an internal strain through the so-called converse or inverse piezoelectric effect. This bias voltage dependent strain causes an additional phonon frequency shift (\( \Delta \omega_{piezo} \)) which can result in inaccurate Raman thermography measurements if not appropriately accounted for. Using the clamped piezomechanical model, assuming that \( \varepsilon_{xx} = \varepsilon_{yy} = 0 \), the \( \varepsilon_{zz} \) strain component can be written as [54]

<table>
<thead>
<tr>
<th>Material/Phonon</th>
<th>Temperature Dependence Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN</td>
<td>( TDMR ) @ 25°C</td>
</tr>
<tr>
<td>Bulk-GaN</td>
<td>( \varepsilon_{0} ) cm(^{-1} )</td>
</tr>
<tr>
<td>Bulk-GaN</td>
<td>E(_2) high</td>
</tr>
<tr>
<td>Bulk-GaN</td>
<td>A(_1) (LO)</td>
</tr>
<tr>
<td>GaN on SiC(^{\ast})</td>
<td>GaN E(_2) high</td>
</tr>
<tr>
<td>GaN on SiC(^{\ast})</td>
<td>GaN A(_1) (LO)</td>
</tr>
<tr>
<td>GaN on SiC(^{\ast})</td>
<td>SiC PTO</td>
</tr>
<tr>
<td>GaN on Si</td>
<td>GaN E(_2) high</td>
</tr>
<tr>
<td>GaN on Si</td>
<td>GaN A(_1) (LO)</td>
</tr>
<tr>
<td>GaN on Si</td>
<td>Si</td>
</tr>
</tbody>
</table>

Measured GaN HEMT wafer parameters, referred to in (6). Please note that these parameters may by specific to the growth conditions used.

\(^{\ast}\) 4H SiC polytype.
where $d_{33}$ and $d_{31}$ are components of the piezoelectric modulus, $E_z$ is the vertical (aligned with the c-axis in GaN) component of electrical field, $S_{11}$, $S_{12}$ and $S_{31}$ are components of the elastic compliance tensor, and $b$ the phonon deformation potential. Figure 5(a) shows a simulated $E_z$ field distribution for an AlGaN/GaN HEMT, highlighting that the peak $E_z$ field is located at the gate foot. However, the inverse piezoelectric strain induced phonon shift contribution in Raman measurements is proportional to the electric field averaged through the GaN layer depth, which is highest under the drain contact.

The most practical way to compensate for inverse piezoelectric strain is to perform a measurement with the HEMT in the ON state, and subtract the phonon frequency obtained in the pinch-OFF state at the same drain voltage (no current flow or self-heating), rather than the OFF state without any applied bias. This correction approach is valid because the electric field distribution, and resulting piezo-electric strain, does not differ significantly between the ON and pinched-OFF state at a fixed source drain bias [54]. This method may also be used to correct for any additional phonon frequency shift induced directly by the applied electric field [55]. $\Delta \omega_{\text{piezo}}$ values measured in the source drain opening of a GaN HEMT as a function of drain bias are shown in Figure 5(b), illustrating that the pinch off reference should be applied point-by-point and for each phonon mode used in the Raman thermography analysis. This piezoelectric strain compensation method was used for determining the temperature of the GaN device shown in Figure 6(a). When measuring transmission line model (TLM) structures, which have two ohmic contacts without a gate contact, it is important to realize that subtracting the inverse piezo-electric strain contribution is not possible using the approach describe here, because the channel current cannot be pinched off. Therefore, the pinch off reference correction technique does have some limits of applicability.

A large temperature gradient is generated near the gate contact during operation, whereas the sample temperature is uniform during calibration. The temperature gradient is particularly high for GaN HEMTs operated at very high power dissipation densities, for example as illustrated in Figure 3 and Figure 6(a). This can cause thermomechanical strain/stress as the hotter expanding region is constrained by the cooler surrounding material. A simulated thermomechanical stress profile in a GaN HEMT is plotted in Figure 6(b) and this will impact Raman temperature measurements [56].

Thermomechanical strain, if not corrected for, is apparent as a discrepancy between the temperatures derived from different phonons in a Raman thermography measurement, as illustrated...
in the inset of Figure 6(a) for the E₂ and A₁(LO) phonons of GaN. We note that the largest effect is close to the drain edge of the gate contact, although this region may not always be optically accessible. It is important to understand that the temperature gradient and resulting thermomechanical strain in the device channel scales with power dissipation density. The thermomechanical strain effect can often be neglected because typical device operating power densities are much lower than the 25 W/mm² value used to illustrate this effect in Ref. [56]. Another possible source of thermomechanical strain is the CTE mismatch between the die, die attach and package. Therefore it is important that temperature calibration is performed on freestanding die or wafers, rather than packaged devices. OFF-state phonon frequencies reference values should also be measured when the wafer chuck or fixture temperature is changed.

The GaN E₂ mode is relatively more sensitive and the A₁(LO) mode relatively less sensitive to strain with respect to temperature, as is apparent from Table II. Therefore, a simplified approach to minimize the effect of thermomechanical strain on GaN HEMT Raman thermography measurements is to use the A₁(LO) phonon shift, instead of the E₂ phonon, to obtain the temperature. Raman thermography measurements with a conservative temperature accuracy of ±5°C can be achieved in the worst case by using the A₁(LO) phonon alone at a measurement location >0.5µm from the gate edge [56]. We note that the A₁(LO) phonon mode is dependent on free carrier concentration, through the so-called phonon-plasmon coupled A₁(LO) mode [47], although in a HEMT the high density of carriers is concentrated in a thin sheet at the AlGaN/GaN interface, so can typically be neglected in the Raman thermography analysis. Despite the higher accuracy of the temperature obtained from the A₁(LO) phonon analysis, the E₂ mode intensity is greater (Figure 4(a)), so would be still be used in cases when the stress is negligible to shorten measurement acquisition times.

The most accurate way to perform Raman thermography measurements is to separate the temperature (Δω₁) and strain (Δω₂) phonon shift contributions given in (5). This can be achieved in materials where multiple phonons can be measured by considering that each phonon has a different temperature and stress/strain dependence, in GaN based devices for example, relating the measured E₂ and A₁(LO) phonon energies, to the actual temperature and stress. Additionally, direct electric field contributions to the phonon frequency shift may need to be considered at high bias voltages [55]. Writing (5) in full, by combining (6) and (9), the expression for the total measure phonon energy shift can be written as

\[ Δ\omega = \omega_0 - \frac{A}{\exp[b\omega_0/k_BT]-1} + 2a \cdot ε_{xx}, \]  

assuming a biaxial mechanical strain. Solving (11) simultaneously for multiple phonons allows the unknown actual temperature \( T \) and biaxial strain \( (ε_{xx} = ε_{yy}) \) induced during operation to be determined in the measured device [56]. Figure 6(a) shows the corrected measured temperature for a GaN HEMT obtained by applying (11) simultaneously to the E₂ and A₁(LO) phonon modes, closely matching with the simulated temperature profile. The resulting strain is also shown in Figure 6(b). A simplified stress/strain correction approach has also been proposed, using a linear approximation for the phonon temperature dependence [57]. Figure 7 is a flow chart illustrating the application of Raman thermography to any material where multiple phonons can be measured, in this example GaN. It is important to realize that the most complex analysis approach need not be applied in all cases and the more simplified approach of analyzing the less strain dependent A₁(LO) phonon has sufficient temperature accuracy in most cases [56].

### IV. EXPERIMENTAL SETUP

Raman thermography requires a standard micro-Raman spectrometer, in our case a Renishaw InVia confocal Raman microscope was used, although other systems can equally be used. Laser sources with a wavelength of 488nm or 532nm were used mostly, although other laser wavelengths are also available, ranging from the UV to the infrared spectral range, chosen to suit the material being studied. Either a notch or edge filter is used to suppress the elastically scattered laser light, enabling the much weaker Raman scattered light to be detected using a charge-coupled device (CCD), after dispersion by the spectrometer grating. The laser light filter rejection method offers higher light collection efficiency when compared with a double or triple monochromator approach. The devices are mounted on a computer controlled XYZ mapping stage, with an 0.1 µm step size, enabling translation of the DUT underneath the focused laser beam. The laser beam is focused through a long working distance (5-8 mm) objective lens with a typical NA of 0.5-0.7 and 50x magnification; this allows sufficient space to accommodate electrical probes, bond wires or device packaging. The Raman scattered light is collected through the same objective in the backscattering geometry. As previously mentioned, the scattering geometry used influences which Raman modes can be observed for the material being studied [47].

A continuous wave (CW) laser is used for steady state measurements, whereas an acousto-optic modulator chops the CW laser for time resolved Raman thermography measurements, generating laser pulses with adjustable pulse length, frequency and duty cycle. This approach gives complete flexibility in the pulse width and duty cycle, in contrast to a pulsed laser source. The shortest laser pulses achieved using our system were 10 ns [58], although shorter pulses could be generated if required by using a suitable optical modulator or a pulsed laser. The laser pulse and electrical pulsing of the DUT

<table>
<thead>
<tr>
<th>Phonon mode</th>
<th>a [cm⁻¹/°C]</th>
<th>b [GPa]</th>
<th>Temperature dependence: ( \alpha ) (°C⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E₂</td>
<td>-742</td>
<td>-715</td>
<td>0.87</td>
</tr>
<tr>
<td>A₁(LO)</td>
<td>-664</td>
<td>-881</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.3</td>
</tr>
</tbody>
</table>

* [53].
* [54].

Calculated using the GaN-on-SiC parameters given in Table I.
a serial measurement and each point in a spatial or temporal map is recorded sequentially, e.g., rastered for a line scan, a 2D or 3D map. To some degree 2D temperature mapping speed can be increased if line focusing is used [59], although at some expense of spatial resolution. Time resolved Raman thermography measurements are also not “single shot” and require integration over many pulses. Typically, a compromise needs to be made between point, line scan or map temperature measurements in terms of acquisition time. However, temperature point measurements are often adequate for the thermal analysis of devices when combined with thermal simulations, as will be discussed later.

V. RAMAN THERMOGRAPHY APPLICATIONS

A. Temperature measurement for accelerated lifetime testing

Knowledge of the exact peak temperature and temperature distribution is very important, especially in GaN RF HEMTs where very high temperature gradients are generated during operation. Peak channel temperature is often referred to in accelerated life testing, although the relevant temperature used to obtain an activation energy actually depends on the dominant degradation mechanism, which could be associated with the gate contact or ohmic source/drain contact, although the former location is the most common high temperature failure point [60]. Measuring the “peak” gate channel temperature would require nanometer scale resolution. Furthermore, the peak temperature location is usually covered by the metal Gamma or T-gate, or field plates. As discussed previously, spatial averaging must be taken into account when comparing thermography results to simulation. 3D FE thermal models must therefore be used to extrapolate from the measured temperature to the relevant temperature for reliability analysis, e.g., peak gate temperature. Of course, it is always preferable to measure as close as practically possible to the peak temperature location in order to minimize the uncertainty when performing such an extrapolation. Compared to other techniques, Raman thermography has the advantage that the temperature is measured in a well-defined volume within GaN layer, close to the peak temperature location, as illustrated in Figure 3; this minimizes the uncertainty when comparing measurement and FE modelling results. A typically data set of GaN-on-SiC HEMT thermal modelling parameters is given in Table III. The verified FE model can then be used to derive a scaling factor to extrapolate from the Raman thermography measured temperatures to the peak channel temperature or at another location in the channel. This scaling factor is dependent of different parameters, including the size of the device. Figure 8 shows that the relative difference between the Raman measured temperature and peak gate temperature decreases with increasing device size, because the $\Delta T$ across the GaN layer accounts for a smaller proportion of the total temperature rise with respect the substrate, die attach and packaging. In the limit of very large power switching devices, the difference between the Raman measured and peak simulated temperature can be negligible [61].
Apart from the material properties input into a thermal model, the peak channel temperature is dependent on the Joule heating distribution in the device channel, which is the scalar product of electric field and current density. To a good approximation, Joule heating is proportional to the in-plane (source-drain) component of the applied electric field. For transistors without field plates, or with field plates but operated at low source-drain voltages, self-heating can be well approximated by an 0.3–0.5µm-long, <100nm-thick heater block at the AlGaN/GaN interface, adjacent to the drain edge of the gate contact [62]. This approach was used to simulate the GaN HEMT temperature profile shown Figure 6(a), closely matching the measured temperature profile. Conversely, the measured temperature profile can be used to estimate the electric field distribution [62]. It should be noted that the exact size and position of the Joule heating region in the channel depends on the gate and field plate geometry, bias voltage and whether the surface is well passivated.

A more precise device thermal modelling approach is to combine the Joule heating distribution derived from a 2D drift diffusion simulation with a 3D FE transistor model [4, 63]. Figure 9 illustrates the simulated Joule heat distribution in the channel of a GaN HEMT operated over a range of drain bias voltages. At less than or equal to 30V_Ds the Joule heating distribution can be approximated by a block heater, as previously described. Above a certain threshold voltage the field plate increasingly spreads the electric field, and the resulting heating, away from the gate [4, 64], protecting the gate against damage. The Joule heat dissipation spreading illustrated in Figure 9 lowers the peak channel temperature and also shifts the peak temperature location away from the gate foot towards the drain, as has been observed experimentally in [62]. In effect, this reduces the difference between the peak channel temperature and the temperature measured by Raman thermography close to the field plate edge. The more sophisticated combined drift diffusion/3D FE thermal modelling approach may not be required in all cases. Normally the channel power distribution can be approximated by two block heaters, one representing heating close to the gate and the other heating under the field plate.

Generally, the question arises: How many temperature measurements are needed to validate a device thermal simulation? This depends on how many of the material thermal parameters are known or unknown. If all the input parameters are known in a thermal model before a Raman thermography measurement, apart from the thermal resistance between the epitaxial layer/substrate interfacial resistance (TBR_eff) and die attach thermal resistance, a minimum of two point temperature measurements are required: At a well-defined distance from the edge of the gate contact or field plate of a HEMT, normally as close as possible (0.5 µm), as well as an edge of chip (EoC) temperature, e.g., 300–500 µm laterally away from the active area. First, the die attach thermal conductivity (or wafer to chuck thermal resistance) is adjusted to fit the measured EoC temperature, then the measured GaN temperature in the channel can be used to determine the TBR_eff value. It is important to check these parameters because they can vary depending on growth/fabrication methods, as will be illustrated later.

Generally, it has been debated whether lifetime testing should be performed using the actual RF operating conditions for microwave devices, or whether the more cost-effective approach of DC lifetime test is adequate. Firstly, degradation mechanisms can be different under DC and RF conditions [65,

<table>
<thead>
<tr>
<th>Material</th>
<th>Thermal conductivity κ [W/mK]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN</td>
<td>160x(300/T)^1.4 a</td>
</tr>
<tr>
<td>SiC</td>
<td>420x(300/T)^1.5 a</td>
</tr>
<tr>
<td>TBR_eff</td>
<td>3—20x(300/T)^0.9 a</td>
</tr>
</tbody>
</table>

a [38]; b from [80], assuming a 100nm thick interface layer.
However, it is also important to establish whether the actual channel temperatures under RF and DC operation are equivalent. This is not self-evident because during RF operation the device electrically traverses the IV plane, and therefore the device exhibits voltages and currents in extreme parts of the IV plane. For RF devices operating in the in the GHz range, an operating cycle is much shorter than the heat diffusion time constant, and therefore the power dissipation “seen” by the transistor channel can be averaged over all points on the load line. The temperature will be constant during CW RF device operation, although this many not necessarily correspond to any DC temperature distribution [4]. The Joule heating distribution is strongly influenced by the source drain bias, particularly at high voltages, as shown in Figure 9 and discussed before. Figure 10(a) shows the resulting simulated channel temperature profiles in DC versus RF operation. At drain voltages up to 50V there is no clear difference between temperature under DC and RF operation, when channel power dissipation is matched in both cases. This is consistent with RF versus DC Raman thermography measurement comparisons performed at typical bias voltages, for example as shown in Figure 10(b). However, Figure 10(a) also illustrates that significant differences are expected to emerge between the simulated DC and RF temperature profile at very high drain bias voltages, this is due to the large voltage swings influencing the electric field distribution in the device. Power added efficiency (PAE) also needs to be considered in any comparison of measured DC and RF temperatures, because only a proportion of the total input electrical power is dissipated in the device as Joule heating, the rest is emitted as RF power.

It is also important to characterize the temperature of devices during transient operation, as operation of actual RF devices, is often pulsed. Figure 11(a) illustrates the transient thermal response of a GaN-on-SiC HEMT measured using time resolved Raman thermography. It is apparent that during the first 50-70 ns, there is little heat diffusion beyond the GaN layer, since the device channel temperature is independent of the substrate material – termed quasi-adiabatic heating. After which, the substrate thermal diffusivity does influence the temperature transient. A temperature profile is clearly apparent in the channel after 100 ns (Figure 11(b)) and it takes from several microseconds to milliseconds, depending on the size of the device, to reach the DC steady-state temperature. Time resolved Raman thermography has also provided valuable

Fig. 10: A comparison of GaN RF HEMT temperatures during RF and DC operation: (a) Simulated channel temperature profiles. Reprinted from [4], Copyright 2013, with permission from Elsevier; (b) GaN temperature measured in a discrete L-band microwave amplifier using Raman thermography at a distance of 0.5 μm from the field plate edge. The temperature rise measured during DC and RF operation is compared for a range of chuck temperatures. Solid lines are shown as guide to the eye. Reproduced with the permission of Tesat Spacecom GmbH.

Fig. 11: Time resolved Raman thermography measurement results showing: (a) Transient thermal response of AlGaN/GaN TLM structures on SiC and sapphire substrates. Reproduced from [58], with the permission of AIP Publishing; (b) Temperature profile measured between the source and drain contact of a single-finger, 125 μm-wide AlGaN/GaN HEMT on a SiC substrate, operated at 20 W/mm, shown for a range of times after switch-on. Reprinted from [86], Copyright 2008, with permission from Elsevier.
information about dynamic temperatures in GaN and GaAs-based microwave devices during actual pulsed RF operation, closely matching the conditions of intended applications [67]. Investigating transistors that exhibit current collapse is another example of where time resolved Raman thermography measurements can provide additional information: Simms et al. have used this technique to visualize the effect of the virtual gate mechanism on the temperature profile during pulsed operation [68].

So far Raman thermography was predominantly demonstrated for electronic devices, although naturally it can also be applied to opto-electronic devices. As opto-electronic devices emit light, the laser wavelength chosen for the Raman experiment needs to be as far as possible away from the light emission wavelength of the DUT. Examples in the literature include GaAs [69], GaN laser diodes [70], UV LEDs [71], amongst other device types. Figure 12 illustrates a Raman line temperature profile measured for a GaAs based laser diode. Similarly to electronic devices, all aspects related to modeling heat transport within layers and at interfaces need to be carefully considered. As it is not possible suppress current in an opto-electronic devices, in contrast to electronic devices, care also needs to be taken to consider effects such as the inverse piezo-electric field effect.

B. Characterization of Material Thermal Properties

The thermal conductivity of materials and the thermal resistances of interfaces between them within the device structure are critical input parameters needed for accurate device thermal simulations. These material parameters depend on the specific growth conditions used and doping of epitaxial layers. Because there isn’t a standard set of material parameters, they must be obtained by measurement. A variety of material thermal characterization techniques exist including 3-omega [72], laser flash [73] and transient domain time reflectance measurements (TDTR) [74]. It is however always preferable to determine the “device relevant” thermal material values which are obtained by measuring temperature rise in operating devices. The ability to map the temperature distribution in 3D through transparent materials is a unique benefit of Raman thermography. Using a suitable heat source, including a transistor or TLM structure, thermal conductivities and interfacial thermal resistances can obtained by measuring the $\Delta T$ across each layer in the device structure. When existing device structures can be used, no special sample preparation is needed. For example, if the thermal conductivity of a transparent substrate is unknown, such as the SiC of a GaN HEMT, the thermal conductivity can be obtained from the vertical temperature gradient measured below the active device as illustrated in Figure 13(b). When the substrate material is not transparent, it is possible to obtain the same information by mapping the surface temperature laterally away from the active area of the device [75]. Raman thermography measurements also allow the measurement of the “effective” thermal conductivity of polycrystalline diamond substrates used for GaN-on-diamond wafer technology, considering the variation in grain size present along the diamond growth direction [75]. The thermal conductivities of silicon, diamond, 4H and 6H SiC determined from Raman thermography measurements are given in Table IV. Correspondingly, Power et al. have measured the $\Delta T$ across the GaN layer in HEMTs to investigate the effect of different dopants and doping concentrations on the thermal

Fig. 12: Temperature determined by Raman thermography on the facet of a GaAs-based laser diode. Reproduced from [69] with the permission of AIP Publishing.

Fig. 13: Temperatures measured using Raman thermography for a 125 μm-wide AlGaN/GaN-on-SiC TLM: (a) Depth temperature profile through the GaN layer and substrate during steady state operation at a power dissipation of 2.1 mW/μm²; (b) transient temperature in the GaN layer and substrate (as inset) during pulsed operation at a peak power dissipation of 7.0 mW/μm². Thermal simulation (overlaid) is used to determine the interfacial thermal resistance (TBR$_{eff}$) from the measurement results. Copyright 2009 IEEE. Reprinted, with permission, from [81].
conductivity of the buffer layer [76]. Hodges et al. have used Raman thermography to demonstrate that the thermal conductivity of GaN layers within AlGaN/GaN power HEMTs can be significantly reduced due to phonon scattering effects when the layer thickness is less than 100nm [77]. It is possible to determine material thermal properties by performing Raman thermography measurements on simple resistor test structures, such as metal line heaters or ring heaters deposited on top of the material of interest, or by monitoring temperature changes as function of laser power. This has been demonstrated for measuring the thermal conductivity of diamond [45], graphene [78] and other materials.

Effective thermal boundary resistance (TBR$_{\text{eff}}$) is the thermal resistance between two dissimilar materials and is another significant contributor to the total thermal resistance in heteroepitaxial devices, which can account for up to 30% of the total thermal resistance in a GaN-on-SiC HEMT for example (Figure 8(b)). This thermal resistance includes the contribution of phonon mismatch across the interface between different materials, which is a true thermal boundary resistance (TBR), e.g., GaN/AlN and AlN/SiC in a GaN-on-SiC HEMT. This can be treated using the acoustic mismatch model (AMM) or diffuse mismatch model (DMM) [74, 79]. However, in many cases the TBR$_{\text{eff}}$ is mainly caused by the “bulk” properties of interfacial layers. For example, defects in the AlN nucleation layer used for GaN devices grown on SiC substrates [80], but also the extensive defect structure in the GaN near to the nucleation layer interface or the diamond microstructure and dielectric interlayers in GaN-on-diamond integrated devices [75]. Two Raman thermography based approaches have been used to determine TBR$_{\text{eff}}$. Ungated transistors (TLM) are used for this analysis because the heat flow is closer to 1D than it is in a gated transistor, increasing the measurement sensitivity to TBR$_{\text{eff}}$ as illustrated in Figure 14. The most straightforward method is to measure the vertical temperature profile through an ungated transistor and determine the TBR$_{\text{eff}}$ from the temperature discontinuity between the GaN layer and substrate; an example of this analysis is shown in Figure 13(b). However, it should be noted that this method is difficult to apply to ungated devices where heating is not uniform in the channel because of surface charge trapping [38]. Well-passivate devices should therefore be used when applying this technique. The second method for determining TBR$_{\text{eff}}$ is to record and analyze the transient temperature in an ungated transistor operated with nanosecond electrical pulses [81], illustrated in Figure 13(a). The interfacial thermal resistance influences the peak and amplitude of the temperature modulation in the device. The TBR$_{\text{eff}}$ can then be obtained from the measured temperature trace by fitting a thermal model.

A selection of commercial and research institute GaN-on-SiC wafers were characterized in [80] using a combination of these measurement approaches, producing the results shown in Figure 15. The large spread of TBR$_{\text{eff}}$ values measured emphasizes that growth parameters determine the magnitude of the interfacial thermal resistance and that TBR$_{\text{eff}}$ can be reduced

![Figure 14](image)

*Fig. 14: (a) Schematic representation of heat generation and heat spreading in a single finger HEMT versus a TLM structure, the former exhibiting 2D-like heat flow in the GaN layer and the latter 1D-like heat flow. (b) 3D FE simulation of the temperature rise across each component in the thermal stack, using a thermal model adapted from [41], including a temperature dependent TBR$_{\text{eff}}$ with a value of 2.5x10$^3$ m$^2$/W at 150 °C. The device width is 100 μm and power dissipation was set at 1 W (10 W/mm) and 2.24 W (2.24 mW/μm$^2$), for the HEMT and TLM, respectively, in order to achieve the same peak temperature rise.*

![Figure 15](image)

*Fig. 15: Effective thermal boundary resistance (TBR$_{\text{eff}}$) for nucleation layers in different AlGaN/GaN on SiC wafers, HEMTs, measured for devices from university / research institutes (URI) and commercial sources (CS). Copyright 2010 IEEE. Reprinted, with permission, from [80].

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>MEASURED GaN SUBSTRATE THERMAL CONDUCTIVITIES AND HEMT THERMAL RESISTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate Material</td>
<td>Thermal conductivity @25°C [W/mK]</td>
</tr>
<tr>
<td>Sapphire *</td>
<td>24</td>
</tr>
<tr>
<td>Silicon b</td>
<td>130</td>
</tr>
<tr>
<td>SiC c</td>
<td>420</td>
</tr>
<tr>
<td>Polycrystalline diamond d</td>
<td>1200 d</td>
</tr>
</tbody>
</table>

* [30]; b [36]; c [75]; d effective thermal conductivity value, representing a depth average through the substrate; e single finger HEMT.
through optimization of the nucleation layer [82]. It also illustrates that a device simulation cannot simply assume a thermal interface resistance as this may differ from one epitaxy to another. Ultimately, the effect of TBR_{eff} on the thermal resistance of a typical GaN RF HEMT can be calculated, which is in the range of an additional 10-30% temperature rise, as also shown in Figure 15. It is therefore crucial to consider its detailed impact on device temperature in any thermal simulation. The contribution of TBR_{eff} to the total device thermal resistance of GaN HEMTs decreases relatively with increasing device size, as illustrated in Figure 8. A greater proportion of the temperature rise is accounted for across the substrate, die attach and packaging in large devices. Figure 15 also shows that TBR_{eff} is temperature dependent, increasing with temperature due to the fact that the thermal conductivity of the AlN nucleation layer decreases with increasing temperature. We also note that the opposite can occur, for example, GaN-on-diamond contains an amorphous interlayer between the GaN and diamond, which increases in thermal conductivity with increasing temperature [83].

C. Raman thermography for device optimization

Many layout, materials, interfaces and packaging schemes. Raman thermography measurements have been used in many cases to explore these effects and validate the thermal models used to aid device design. RF HEMTs are designed with many gate fingers operating in a parallel multi-finger arrangement to achieve high RF output powers. This introduces an additional thermal resistance induced by the “cross-talk” between individual gate fingers [84]. Factors determining the magnitude of the thermal cross talk effect include the gate width and gate pitch (separation between fingers). Figure 16(a) illustrates the temperature measured in two multi-finger GaN-on-SiC layouts, comparing 8 fingers with a 25 µm pitch versus 4 fingers with a 50 µm pitch. The measured thermal resistance of the larger device is ~15 °C•mm/W, double that of the smaller device; this trend is also reproduced in the overlaid thermal simulation results. Experimental validation of thermal models for existing device layouts enables other design spaces to be simulated with higher confidence. Such a map of thermal resistance versus finger width and finger pitch is shown in Figure 16(b), which highlights the importance of carefully considering the effect of device layout on thermal resistance. For this reason, the gate pitch in a GaN-on-SiC RF HEMTs is typically >30µm, although for lower or high thermal conductivity substrates, this limit is increased or decreased correspondingly, e.g., a gate pitch of <10 µm is possible when substituting SiC with high thermal conductivity diamond substrates [85]. Manoi et al. have also shown that multi-finger devices have microsecond-scale thermal time constants associated with the transient thermal cross talk between gate fingers [86], which is an important consideration for intermodulation distortion in RF amplifiers.

The substrate material has a key heat spreading role due to its close proximity to the channel region in a GaN HEMT. High thermal conductivity substrate materials are therefore very beneficial for thermal management and reducing thermal resistance. Table IV shows the thermal resistance measured for single finger GaN HEMTs using Raman thermography, comparing sapphire, Si, SiC and polycrystalline diamond substrates. However, we note that the values given in Table IV are obtained for small periphery transistors and are not necessarily representative of larger devices; the thermal resistance improvement would be clearer for large devices because the total power dissipation is higher and the substrate thermal conductivity is more critical [85]. SiC with its relatively high thermal conductivity of ~420 W/mK is the current material of choice for high power applications, whereas silicon is attractive for lower cost applications. For ultra-high power applications, even SiC substrates can become a thermal bottle neck. Polycrystalline CVD diamond is being developed for GaN ultra-high-power applications to address this [87, 88]. For example, GaN-on-diamond thermal models, using material parameters obtained from Raman thermography measurements, have been used to demonstrate that GaN-on-diamond can offer a ~3× reduction in thermal resistance when scaled to large devices, with respect to GaN-on-SiC [85].

After the substrate, effective heat removal through packaging is increasingly important for high power dissipation devices. One approach is to substitute conventional packaging materials with high thermal conductivity materials. For example, the benefit of replacing CuW with a silver-diamond composite is demonstrated by Raman thermography measurements performed in Ref. [89]. Alternatively, a flip-chip package can be used by attaching a high thermal conductivity heat spreader.
such as AlN to the top side of the device for heat extraction [43]. 3D Raman thermography is particularly suitable for assessing flip chip packages. One example of this application is shown in Figure 17, in which the temperature gradient is measured in 3D through a GaN-on-sapphire HEMT mounted on a flip chip package, clearly showing that the device temperature is lower close to the solder bumps where heat is extracted.

D. Surface sensitive measurements

We have previously explored how Raman thermography can be used to probe the temperature in transparent layers. In the case of thin layers, this means that a depth average is measured through each layer. Sometimes it is advantageous to probe the surface temperature only, for example when measuring a device fabricated homo-epitaxially, such as a GaN-on-GaN device. This can be achieved by using an above bandgap laser to perform the Raman measurement, e.g. UV for GaN (3.4 eV bandgap) or visible for GaAs (1.4 eV bandgap). The depth probed is then given by (4) and is much less than the depth-of-field of the focusing optics. For direct bandgap semiconductors the absorption depth is typically 10’s nm (eg. GaN), while it is ~100’s nm for indirect bandgap semiconductors. However, special consideration does need to be taken to minimize the effect of laser heating on the measurement, similar to the PL method for determining temperature discussed earlier. This can be tested by measuring the device temperature as function of laser power. It is critically important to not perform these checks with the unbiased device “off”, but also with voltage applied, as the heating caused by the photo-induced current can often dominate over the direct laser heating of the material being measured [22]. Because of the necessity to use a low laser power, typically 10’s μW or less, above bandgap Raman thermography measurements are more time consuming and challenging than sub-bandgap measurements. GaAs devices are normally probed in Raman thermography using above bandgap laser sources. Figure 18 shows an example of a surface temperature profile measured in the gate-drain opening of an AlGaAs/GaAs pHEMT, obtained by measuring the Raman active LO phonon. Despite being experimentally challenging, time resolved Raman thermography measurements have also been applied to GaAs-based HEMTs [90]. For GaN electronic devices UV lasers have been used to measure the temperature in the GaN layer close to the 2DEG where Joule heating is generated [31, 91].

VI. OPTICAL RESOLUTION ENHANCEMENT

Raman thermography offers sub-micron lateral spatial resolution, although the axial spatial resolution is larger than the thickness of active layers in a typical device, e.g., greater than the 1-2µm thickness of the GaN layer in a GaN RF HEMT, as discussed in Section 3.1. The axial resolution is further degraded by refraction induced aberration when back-side measurements are performed through transparent substrates [42], and in particular for high refractive index materials such as SiC (n=2.6). Overcoming this limitation and improving the axial spatial resolution to the nanometer scale would make true 3D temperature mapping of devices a possibility, reducing the reliance on thermal modelling. One possible solution for increasing the lateral and axial spatial resolution of Raman thermography, beyond what is achievable with standard optics, is to apply the solid immersion lens (SIL) technique. A SIL is essentially a lens which is integrated into or attached to the device being measured, having the same or very similar refractive index as the device materials. Fortunately, the refractive index of the SiC substrate (n=2.6) is very similar to the GaN layer (n=2.4) in a GaN-on-SiC HEMT. There are two types of SIL, a hemispherical SIL (h-SIL), and a super-spherical SIL (s-SIL) in which the focal point is offset from the center of the sphere to from a truncated spherical lens. The h-SIL has negligible chromatic or birefringence induced aberration and therefore is the most appropriate approach for Raman thermography measurements. It increases the NA and magnification by a factor of n, with respect to a standard air immersion objective and introduces no optical aberration for a perfect SIL. The NA limit is equal to the refractive index of the focusing medium and therefore maximum benefit is obtained when using high refractive index SIL’s: Lateral resolution scales as 1/NA, whereas depth resolution scales more...
dramatically as \( 1/\sqrt{A} \).

Two approaches are possible, firstly the attachment of an external lens onto the device with a nanometer-sized gap (commercial or semi-commercial SILs exist for Si and diamond), and the etching or milling of the SIL into the devices being measured. For example SiC SILs have been fabricated on the backside of a GaN-on-SiC device structure in [41], as illustrated in Figure 19(a). By taking advantage of SILs fabricated in high-refractive index SiC, resolutions of \( r_{lateral} = 0.15 \mu m \) and \( r_{axial} = 0.5 \mu m \) at a wavelength of 532 nm are possible at \( NA=2.0 \) [41]. Figure 19(b) illustrates temperature profiles measured at two depths within the GaN layer of an AlGaN/GaN-on-SiC transistor achieved using a SIL, showing that the vertical temperature gradient within the GaN device layer can be probed. An axial resolution of 1.7 \( \mu m \) was achieved with \( NA = 1.3 \) in [41]; further improvements are possible by fabricating larger diameter SILs and introducing aberration correction.

VII. CONCLUSION

The Raman thermography technique was reviewed, explaining the principles of its application and giving examples of its practical use. The technique enables temperature determination in electronic and opto-electronic devices with a lateral resolution of \(-0.5 \mu m\) using standard optics and sub-micron axial resolution when combined with a solid immersion lens. Time resolved Raman thermography can be used to measure temperature transients with nanosecond temporal resolution. Combining Raman thermography measurements with finite element modeling analysis enables a detailed understanding of thermal management in devices and packages.

REFERENCES

GaN electronic devices, "Superior thermal con
re in pulsed
Time,
M. J. Uren, et al., "Temperature dependent thermal resistance of
D. Babic, F. Ejeckam, A. Nurmikko, and H.
Raman spectroscopy using a cooled CCD imaging


