Time resolved Raman thermography analysis of transient heating in a pulsed GaN HEMT RF high power transistor

James W Pomeroy,[1] Filip Gucmann,[1] John Walker, [2] James Custer[2], Gabriele Formicone[2], and Martin Kuball[1]

[1]Center for Device Thermography and Reliability, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.
[2]Integra Technologies, Inc., 321 Coral Circle, El Segundo, CA, 90245-4620, USA.

ABSTRACT

Measurement is a vital step in thermal device model verification. We present a static/transient measurement and simulation study of a GaN high power pulsed transistor. Time resolved Raman thermography provides high spatial (0.5 µm-lateral) and temporal resolution (>10ns), enabling the temperature within the GaN transistor to be measured close to the peak channel temperature. This technique is therefore an ideal reference measurement technique for GaN HEMTs, which generate high temperature gradients close to the gate. We demonstrate this process of thermal model verification. Complementary IR thermography and transient thermo-reflectance measurements are also performed for comparison. The resulting experimentally validated finite element thermal simulation is used to derive a thermal equivalent circuit model, representing the thermal impedance (Z_{TH}) of the tested device.

I. INTRODUCTION

A critical parameter for any RF power amplifier is its reliability, which is usually measured by its Mean Time To Failure (MTTF). Although there are many components in an RF power amplifier, its overall MTTF is usually dominated by the MTTF of the RF power transistors used inside it. The MTTF of a transistor depends on how hot it gets and, very roughly, a 30°C increase in temperature results in the MTTF falling by a factor of 10. Given this very high sensitivity of MTTF to temperature, it is clearly very important to have an accurate knowledge of the operating temperature. However, the transistor is not at a uniform constant temperature as it would be if placed in an oven with no DC power dissipated within it, rather the temperature varies both laterally and vertically within the transistor chip. But the place where the temperature can be measured, by IR thermography for example, is not the highest temperature within the transistor. That occurs inside the channel region of the semiconductor material (GaN on SiC in this work), underneath the metallization where it can't be optically accessed in a packaged device.

The channel temperature is often termed the junction temperature, although channel temperature is a more accurate description for a HEMT device. The semiconductor industry has standardized on referring MTTF to the highest temperature within the device and not the measured temperature, and so it is necessary to know the relationship between measured temperature and the highest temperature inside the semiconductor material. There is a proportional relationship between the measured temperature and the temperature within the semiconductor, and in some very simple structures such as a circular PN junction diode this can be calculated analytically, but not in the case of a multi-finger GaN-on-SiC HEMT. Accordingly, it is necessary to resort to numerical computer simulation, which forms the basis of much of the work reported here. However, potential uncertainties in simulation parameters mean that modelling should not be relied upon alone, without experimental verification. The high temperature gradient present in GaN HEMTs poses a temperature measurement challenge, requiring a high spatial resolution technique to avoid severe spatial averaging. In this study high spatial resolution static and transient Raman thermography measurements, complemented by IR thermography and transient thermo-reflectance measurements, are used to validate a finite element method (FEM) thermal model of a high power GaN HEMT, from which a lumped element thermal equivalent circuit is derived.

Transistor manufacturers undertake accelerated life-tests to determine the MTTF of their transistors, but these tests are almost always undertaken under CW conditions. Under CW operation the channel temperature depends on the thickness of the various layers and their thermal conductivity, as well as on geometrical dimensions and the dissipated power density. However, many applications for RF power transistors, such as radar systems, require that the devices operate under pulsed conditions. Unlike CW operation, the channel temperature is not constant but increases during the pulse. Accordingly, it is essential to determine the channel temperature as a function of time, and this depends on the pulse length in comparison to the transistor's thermal time constant. The thermal time constant is determined by the thermal diffusivity rather than by the thermal conductivity of the various layers.

It is often assumed that the MTTF under pulsed conditions is simply the value determined from CW measurements divided by the duty cycle as a worst case situation - worst case because the peak temperature is only reached at the end of the pulse. Because of this apparent improvement in MTTF by operating under pulsed conditions, manufacturers of pulsed transistors often run the transistors at a higher channel temperature (by driving the transistor harder to get a higher RF power output) than they would if the device had to operate in CW. However, the underlying assumption in all of this is that the transistor cools down to ambient during the off period, but this depends on the thermal time constant of the transistor in comparison to the length of time that no RF signal is applied to the transistor. Much has been published on the peak channel temperature of transistors under CW operation, but comparatively little on either the temperature measurement or the calculation of the peak channel temperature under pulsed conditions, and this forms the basis of the work reported in this paper. An accurate thermal model for a pulsed RF power transistor is needed not just for accurate MTTF predictions but also for accurate prediction of RF performance when using a non-linear electro-thermal model.

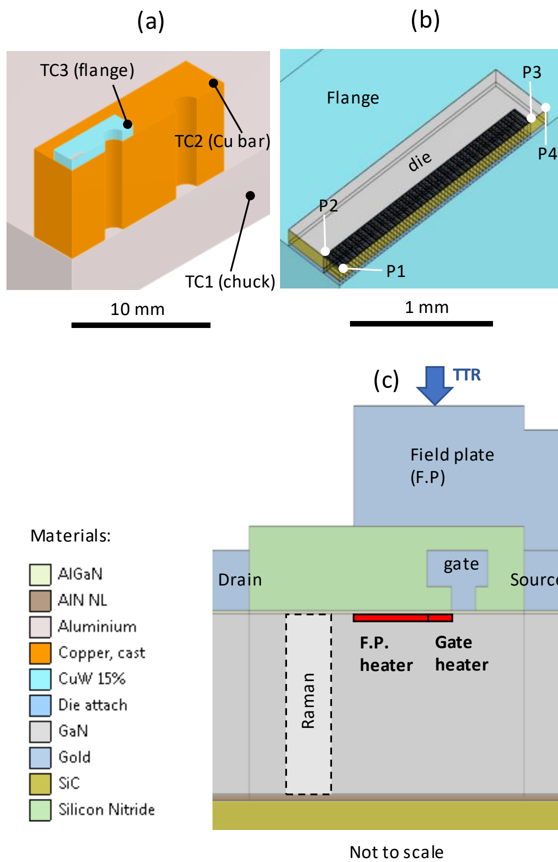

Figure 1: (a) Illustration of ¼ of the measured device and fixture. Thermocouple locations are labeled TC1-3. (b) Drawing of a ¼ of the die with Raman thermography measurement locations labelled P1-P4. (c) Schematic cross section of the transistor channel with the Raman measurement and TTR measurement locations indicated. The "gate" and "field plate (F.P.)" represent the volumetric heat loads applied in the F.E. thermal model.

II. Methodology

The device under test (DUT) consists of a 36 mm total gate periphery AlGaN/GaN-on-SiC HEMT, with the die brazed onto a CuW flange using AuSn solder. The flange is mounted onto a copper bar heat spreader, which in turn is mounted onto an aluminum thermoelectric chuck with active PID temperature control. Thermocouples were used as additional temperature reference points, for comparison to thermal simulation, and placed on the chuck (TC1), at the end of the copper bar (TC2) and end of the flange (TC3); these locations are illustrated in Fig. 1(a). The chuck is fixed to a three axis (XYZ) Prior motorized microscope stage with optical encoders and an 0.1 µm step precision, which enables the region of interest to be precisely positioned under the microscope lens. The transistor was operated at V_{ds}=50V: P_{diss}=42.5 W (1.18 W/mm) for static measurements and 180 W (5 W/mm) for pulsed measurements, with a 100 µs pulse and 10% duty cycle.

GaN HEMTs operate at high local power dissipation densities, up to several W/mm, and Joule self-heating occurs in a sub-micron high electric field region in the two-dimensional electron gas (2DEG) close to the drain edge of the gate. The high heat flux, combined with relatively high thermal conductivity materials (κ_{GaN}=160 W/m·K, κ_{SiC}=440 W/m·K) in the thermal dissipation path, can generate local temperature gradients of 10's °C per micron in the channel region [1]. The spatial resolution of static IR thermography measurements is

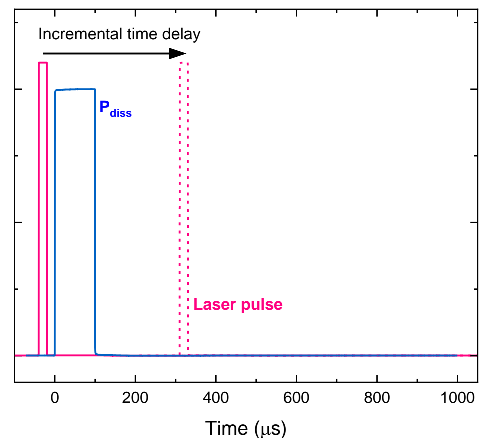

Figure 2: A diagram showing timing of the transistor power dissipation (P_{diss}) and the laser pulse used for the time resolved Raman thermography measurement.

typically diffraction limited to >5µm, a limitation due to the wavelength of the detected thermal emission. The GaN and SiC materials are also IR transparent, which further reduces the measured temperature. In the worst case, the static temperature measured by IR thermography can be 50% lower than the actual peak channel temperature for a single finger HEMT [1]. IR cameras are limited to millisecond image acquisition times and while single element IR detectors can achieve microsecond temporal resolution, spatial resolution is sacrificed (~20 µm). Raman thermography is a spectroscopic optical technique which exploits the temperature dependence of lattice vibrations (phonon frequency) in a material to measure temperature at a precise location. Sub-micron spatial resolution can be achieved because visible wavelength laser light is used for this measurement. A pulsed laser source is used for time resolved Raman thermography and a temporal resolution as short as 10 ns has been demonstrated [1].

A combined IR/Raman microscope was used [2], which integrates a QFI Infrascope camera for the IR thermography measurements and a Renishaw inVia spectrometer for Raman measurements. The IR camera is primarily used for screening devices, whereas Raman thermography is used for quantitative temperature measurements. A 15× 0.5 numerical aperture (NA) IR objective lens was used for the IR thermography measurements and 50× 0.5 NA microscope objective lens was used for the Raman measurements, having measured optical resolutions of 7-8 µm and 0.5 µm, respectively [2]. The sample surface emissivity was calibrated at 85°C before undertaking the IR thermography measurements. For the Raman measurements, a sub-bandgap laser is focused onto the region of interest, ensuring that the material is transparent and that there is negligible light absorption and laser heating; A continuous wave (CW) 532 nm diode pumped solid state laser (DPSS) was used for this study. Static measurements are made using the CW laser output directly, whereas for time resolved measurements the laser was chopped using an acousto-optic modulator (AOM) to generate 20 µs-duration pulses (2% laser duty cycle with respect to the HEMT electrical pulse period). The laser pulses are synchronized with the device electrical pulses and a time sequence is recorded by incrementing the delay between the optical and electrical pulses (Fig. 2) – in effect performing a boxcar average at each time point. The reported measured Raman temperatures are an average of at least 4 acquisitions in order to improve the signal/noise ratio. Each acquisition represents 45s and 240s long integration of Raman

Table 1: FEM thermal model parameters. *Temperature dependent specific heat capacities were included for the semiconductor materials.

Material	Thermal conductivity [W·m^{-1}·K^{-1}]	Density [kg/m^3]	Specific heat capacity [J·kg^{-1}K^{-1}]
Si_xN_x passivation	3	2500	170
AlGaN barrier	10	6150	420
GaN	$160\times(300/T)^{1.4}$	6150	420*
AlN nucleation layer	$2.5\times(300/T)^{0.9}$	6150	490
SiC	$440\times(300/T)^{1.15}$	3211	660*
AuSn die attach	24	14500	1000
CuW flange	190	2700	900
Copper bar	394	8940	385
Aluminium chuck	210	3260	780

scattered light during static and transient measurement, respectively.

Raman measurements were recorded at four locations: Middle of the central gate finger (P1); edge of the central gate finger (P2); middle of the outer finger (P3) and at the edge of the die (P4), as illustrated in Fig. 1(b). The probing laser was focused at closest optically accessible region to the peak channel temperature location (P1–3), which in this case is 0.5 µm away from the drain edge of the field plate, as illustrated in Fig. 1(c). The GaN temperature was determined from the measured A_1(LO) high phonon peak position shift, with reference to the device in the off state (channel pinched off, I_{ds}=0A), and applying a predetermined phonon temperature shift calibration function. Raman thermography probes a volumetric average through the GaN layer thickness, which must be accounted for when making comparison to a thermal simulation. The GaN A_1(LO) temperature measurement error is typically <5-10°C. More details about the Raman thermography method and analysis can be found in Ref. 1.

Complementary transient thermoreflectance (TTR) measurements were also recorded. TTR exploits the fact that the reflectivity of metals is proportional to the change in temperature. The CW laser was focused onto the top of the field plate metal and the modulation in reflectivity is recorded by photodiode and oscilloscope. TTR is a quick measurement, but the thermo-optic coefficient, which relates the reflectivity change to temperature change, is surface dependent [3]. For example, impurities or surface passivation layers will affect the thermo-optic coefficient. TTR must therefore be calibrated at each measurement point. This is conventionally achieved by placing the DUT onto a hot chuck and modulating temperature, but this introduces thermal expansion drift, which makes calibration of small targets such as the field plate challenging. In this case we simply normalize the TTR signal for comparison to the time resolved Raman thermography measurement, using a similar approach to Ref. 3.

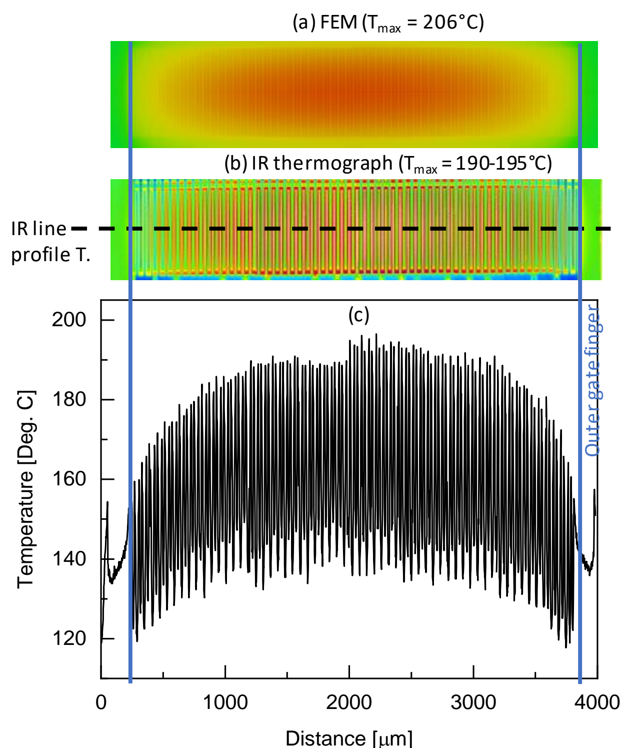

Figure 3: Plan view static temperature distribution across the transistor gate fingers for operation at V_{ds}=50V: P_{diss}=42.5 W (1.18 W/mm) and T_{chuck} = 85°C, showing (a) FEM simulation and (b) IR thermography measurement. (c) shows the IR measured line temperature profile along the center of the gate fingers.

III. Numerical Simulation

A transient finite element method (FEM) thermal model of the DUT was implemented in ANSYS. The FEM model accurately represents the geometry of the HEMT channel (including nm-thick AlGaN barrier), die, flange, fixture and thermal chuck. Joule self-heating, the product of electric field and current density in the channel, is bias dependent. The Joule heating distribution simulated using a drift diffusion (TCAD) model [4] is approximated by two 25 nm–thick volumetric heat loads in each gate finger of the FEM model. The heater locations are shown schematically in Fig 1(c), located adjacent to the drain edge of the gate, under the gate overhang and under the field plate, with a 2/3 and 1/3 split of the total power dissipation in the gate heaters and field plate, respectively. Reported thermal conductivity, density and specific heat capacity values were used and are summarized in Table 1. However, interfaces between different layers present the main uncertainties in the simulation parameters and should be checked/adjusted based on measured device temperatures, e.g., the effective interfacial thermal boundary resistance between the GaN layer and the SiC substrate (TBR_{eff}), the effective thermal conductivity of the die attach and the heat transfer coefficient (HTC) between the flange and heat sink. The FEM mesh was carefully refined to obtain a mesh-independent solution. With 3.7 million elements, the static and transient solutions

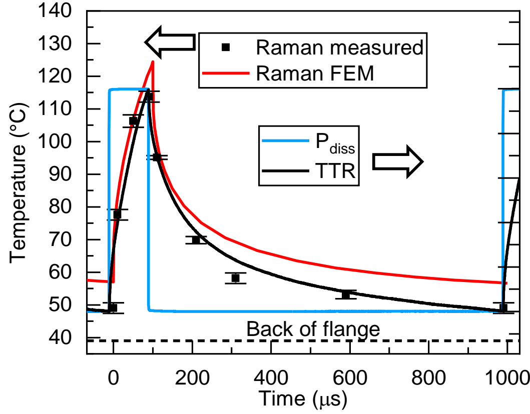

Figure 4: Time resolved Raman temperature measurement at location P1 (center of middle gate finger) for 100µs duration, P_{diss} = 180W pulse (5 W/mm), at T_{chuck}=25°C. The FEM simulated temperature transient at the Raman measurement location and normalized TTR signal measured on the field plate metal are also plotted. The simulated temperature at the back of the flange is displayed as a horizontal dashed line.

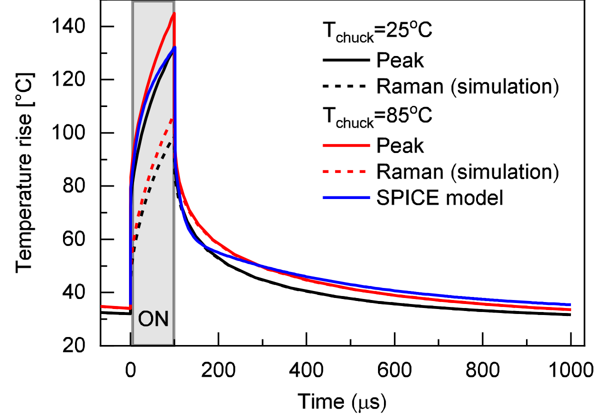

Figure 5: FEM thermal simulation of the transient temperature rise during pulsed operation, P_{diss}=180W, 100µs, 10% duty cycle, at a chuck temperature of 25°C and 85°C. The thermal equivalent circuit (SPICE model) result plotted for comparison.

take ~10 minutes and several hours, respectively, to compute on a desktop PC.

IV. Results and Discussion

Figure 3(b) shows a thermal image of the transistor measured using the IR camera. No unusual hot spots or anomalies were observed and the device had a typical symmetric temperature distribution, with a highest temperature at the center of the device, which is evident in Fig. 3(c). Next, static Raman and thermocouple measurements where recorded at the locations indicated in Fig.1(a) and (b). The measured temperatures, summarized in Table 2, were used to adjust specific parameters in the thermal model, following these steps: 1) The heat transfer coefficient between the copper bar and chuck was set to 50 $kWm^{-2}K^{-1}$ to match measured TC2 thermocouple value (edge of copper bar); 2) The die attach thermal conductivity was adjusted to 24 $Wm^{-1}K^{-1}$ matching the measured edge of die temperature (P4). We have used a typical value for the effective thermal boundary resistance, TBR_{eff} = 2×10^{-8} $m^2\cdot K/W$ (converted into a thermal conductivity value in Table 2) [5], which is associated with the nucleation layer present between the GaN and SiC substrate. After making these adjustments, the FEM simulated temperatures lie within the measurement range of accuracy. The simulated temperature distribution across the transistor gate fingers is shown Fig. 2(a), which qualitatively agrees with the IR thermal image. We note that IR measured temperature is only 10-15°C lower than the peak temperature in the thermal simulation. However, the static measurement is made at a low power dissipation density (1.18 W/mm) and the difference between the spatially averaged IR temperature and peak channel temperature will scale with the local power dissipation density, which in this case is 5 W/mm during pulsed operation.

Figure 4 shows the transient temperature measured by time resolved Raman thermography at the middle of the central gate finger (P1); A lower chuck temperature, 25°C, was used to minimize thermal expansion drift during the lengthy data acquisition. The measured temperature reaches 114°C at the end of the heating pulse, cooling to approximately ~50°C at the end of the cooling phase (ΔT=65°C). The recorded TTR trace qualitatively agrees with the shape of the temperature transient measured by time resolved Raman thermography. The temperature predicted at the same location by the transient FEM thermal model is slightly higher (10°C) than the Raman measurement. The thermal model therefore gives us a conservative estimate of the pulsed operating temperature. The peak channel temperature is compared to the simulated temperature at the Raman measurement location in Fig. 5, which shows that the peak channel temperature is predicted to be 30% higher than measured temperature for this operating condition. Increasing the chuck temperature from 25°C to 85°C is predicted to result in an additional 10°C temperature rise, which is attributed to the temperature dependence of the GaN, TBR_{eff} and SiC model parameters.

The experimentally verified finite element model produces an accurate prediction of the transient temperature during pulsed operation, but it is computationally demanding and unsuitable for circuit simulations. For this reason, it is convenient to represent the transient thermal response $Z_{TH}(t)=\Delta T/P_{diss}$ (impedance) as a lumped RC-network thermal equivalent circuit. The Foster model is most commonly used, which consists of a number (n) of RC elements in

Table 2: Summary of the measured steady state operating temperature at the locations labelled in Fig. 1: Gate fingers and die (P1–4, Raman thermography) and of the chuck, fixture and flange (TC1-3, thermocouple), compared to the FEM simulated values. The device was operated at V_{ds}=50 V and P_{diss}=42.5 W (1.18 W/mm), at two chuck temperatures. IR measured values are given at T_{chuck}=85°C for comparison, although it should be noted that a large area is probed and is not equivalent to the Raman measured GaN temperature.

	Temperature [°C]				
	At T_{chuck}=25°C		At T_{chuck}=85°C		
Location	Raman	FEM	Raman	IR	FEM
Peak		136			206
P1	131±10	128	206±5	190-195	197
P2	133±5	115	191±5		181
P3	97±10	97	166±5		162
P4	86±5	82	147±5	135	145
TC1	25	25	85	85	85
TC2	29	28.3	89	89	88.3
TC3	31.8	32.7	93	93	92.6

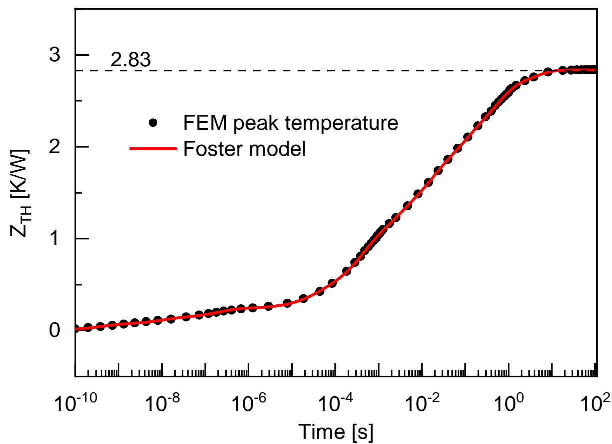

Figure 6: FEM simulated thermal transient at T_{chuck} 85 °C and stepping P_{diss} from 0 to 42.5 W at T=0s. The Foster thermal model fitted curve is overlaid.

series, which can be expressed as

$$Z_{th}(t) = \sum_{i=1}^{n} R_i \left[1 - \exp(-\frac{t}{\tau_i})\right] \quad , \quad (1)$$

where $\tau_i = R_i \cdot C_i$ is the time constant of element *i*. In order to derive the thermal equivalent circuit, the FEM model is used to simulate the response to a step change in power. Figure 6 shows the simulated transient, initially with the device at the ambient temperature (T_{chuck} 85 °C) and with P_{diss} stepped from 0 to 42.5 W (matching the static measurement condition) at T=0s. We note that Z_{th} increases gradually until 10μs, after which the slope steepens as heat reaches the die attach interface. After ~10 seconds, Z_{th} approaches the steady state value of 2.83 K/W. Equation (1) was fitted to the simulated temperature transient, incrementing the number of terms *n* until the fit converged at *n*=10; the resulting best fit is shown in Fig. 6. To ensure self-consistency, the Foster network was modelled using SPICE at the measurement operating condition (T_{chuck}=85°C), which is compared the FEM result shown in Fig. 3. The temperature transient produced by the thermal equivalent circuit is a reasonable approximation of the FEM model result, although underpredicts the peak temperature rise by ~10°C. The Foster thermal equivalent circuit method does not account for the temperature dependence of the *RC* elements, which is the reason for this discrepancy.

V. Conclusion

A finite element transient thermal model of a 36mm GaN HEMT has been developed and validated using static and time resolved Raman thermography measurements. The resulting experimentally validated finite element thermal simulation is used to derive a thermal equivalent circuit model, representing the thermal impedance (Z_{TH}) of the tested device.

References

[1] M. Kuball and J. W. Pomeroy, “A Review of Raman Thermography for Electronic and Opto-Electronic Device Measurement With Submicron Spatial and Nanosecond Temporal Resolution”, *IEEE Trans. On Device and Materials Reliability*, Vol. 16, no. 4, pp 667-683 (2016).

[2] Sarua et al., “Integrated micro-Raman/infrared thermography probe for monitoring of self-heating in AlGaN/GaN transistor structures”, *IEEE Trans. On Electron Devices*, Vol. 53, No. 10, pp. 2439–2447 (2006).

[3] S. Martin-Horcajo et al., “Transient Thermoreflectance for Gate Temperature Assessment in Pulse Operated GaN-Based HEMTs”, *IEEE Electron Device Letters*, Col. 37, No. 9, pp 1197–1200 (2016).

[4] J.W. Pomeroy et al., “Operating channel temperature in GaN HEMTs: DC versus RF accelerated life testing”, *Microelectronics Reliability*, Vol. 55, No 12, pp. 2505–2510 (2015).

[5] A. Manoi et al., “Benchmarking of Thermal Boundary Resistance in AlGaN/GaN HEMTs on SiC Substrates: Implications of the Nucleation Layer Microstructure”, *IEEE Electron Device Letters*, Vol. 31, No. 12, pp. 1395-1397 (2010).