Transient Stress Characterization of AlGaN/GaN HEMTs due to Electrical and Thermal Effects

Jason P. Jones\textsuperscript{1}, Eric Heller\textsuperscript{2}, Donald Dorsey\textsuperscript{2}, Samuel Graham\textsuperscript{1}

Georgia Institute of Technology\textsuperscript{1}
Woodruff School of Mechanical Engineering

Air Force Research Laboratory\textsuperscript{2}
Materials and Manufacturing Directorate

ABSTRACT
In this paper, we present finite element simulation results of the transient stress response of an AlGaN/GaN high electron mobility transistor (HEMT). The modeling technique involves a small-scale electro-thermal model coupled to a large-scale mechanics model to determine the resulting stress distribution within a device operated under radio frequency (RF) conditions. The electrical characteristics of the modeled device were compared to experimental measurements and existing simulation data from literature for validation. The results show critical regions around the gate Schottky contact undergo drastically different transient stresses during pulsed operation. Specifically, stress profiles within the AlGaN layer around the gate foot print (GFP) undergo highly tensile electro-thermal stresses while stresses within the AlGaN outside the gate connected field plate (GCFP) towards the drain contact undergo highly tensile electrical stress and compressive thermoelastic stress. It is shown AlGaN/GaN HEMTs undergo large amounts of cyclic loading during typical transient operation. Based on these findings, transient failure mechanisms may differ from those previously studied under DC operation due to large amount of cyclic loading of a device around the gate structure.

KEY WORDS: AlGaN/GaN HEMTs, transient stress, pulsed devices, electro-thermo-mechanical simulation

NOMENCLATURE

\begin{align*}
V_g & \quad \text{Applied gate-to-source voltage} \\
V_{ds} & \quad \text{Applied drain-to-source voltage} \\
I_{ds} & \quad \text{Drain-to-source current} \\
\text{GFP} & \quad \text{Gate foot print} \\
\text{GCFP} & \quad \text{Gate connected field plate}
\end{align*}

INTRODUCTION

Gallium nitride (GaN) based HEMTs have recently been under intense research and are becoming attractive devices for high voltage and high-power applications at RF operating conditions. GaN is a wide band gap (~3.4 eV at room temperature) semiconductor with a promising combination of material properties including a high electric breakdown field, good electron mobility, high saturation velocity, relatively high thermal conductivity, and is stable at high operating temperatures [1]; all of which contribute to making these devices very suitable for RF devices where high power and high frequency operation are needed [2, 3]. Development and fabrication of reliable AlGaN/GaN HEMTs has significantly advanced in recent years to enable the production of high quality, commercially available devices in a wide variety of high power and high frequency applications. To further study these devices, however, it is essential to investigate the reliability physics associated with AlGaN/GaN HEMTs, especially under the transient operating regime where these devices are predicted to excel over existing technologies.

Several competing stress effects occur within an AlGaN/GaN HEMT under normal operating conditions that may lead to degradation. Several researchers have outlined these effects within devices, but the underlying mechanisms for degradation has yet to be fully understood [4-6]. Electrical degradation is characterized by the onset of gate leakage in the device, loss of power added efficiency, current collapse, change in transconductance, and gate current noise [4, 7, 8]. In addition, physical mechanical degradation in the form of cracking or pitting of devices around the edges of the gate contact has been reported in Refs. [4, 7-10]. Electric fields and inverse piezoelectric generated mechanical stress contributions will typically peak here [7] as well as temperature gradients from non-uniform and bias dependent heating [11] which generate stress through thermal expansion. Complicating the issue, this same region is usually at the junction of low coefficient of thermal expansion (CTE) SiN\textsubscript{x}, high coefficient gold (Au) gate metal, and middle valued III-N semiconductor. In addition, residual stresses exist due to fabrication processes [6, 12-15]. SiN\textsubscript{x} passivation, in particular, can be deposited in a highly stressed state [16, 17] and the AlGaN layer has large amounts of intrinsic tensile stress...
of the order of 1-3 GPa [6, 12] from pseudomorphic growth on GaN.

With these many stress effects, it is critical to be able to separate the distinct contributions to the stress and strain so as to be able to understand the physics of failure within the device. While temperature, stress, and strain profiles have been extensively studied using numerical multi-physics coupled simulations [6, 18-20] and experimentally during DC electrical testing [4, 21], few have studied the transient stress development even though these devices have numerous applications in the RF regime. Of those that have studied the transient response of AlGaN/GaN HEMTs, their focus was primarily on either electrical or thermal performance [12, 22, 23]. Ref. [24] presents a hybrid time-domain finite-element method (TD-FEM) model that couples the electro-thermal response, but only accounts for thermal stress effects. Under RF operation, the combination of electrical bias, frequency, duty cycle, and device structure will induce varying amounts of cyclic stress that varies from stress profiles seen in DC operation and must therefore be fully accounted for when reporting transient stress values.

Understanding transient failure modes requires an understanding of transient stress profiles. In this work, an electro-thermo-mechanical model is used to simulate the complex stress distributions within a device under RF conditions. In doing so, the magnitude and distribution of stress present in a device at critical regions such as the gate foot print (GFP) and gate connected field plate (GCFP) while under cyclic bias conditions are revealed. Our modeling effort allows transient components of stress to be convincingly decoupled to reveal individual and/or combined contributions due to electrical and thermal effects.

**MODEL DESCRIPTION**

Both the electro-thermal and mechanics model are based upon 2-finger, gate centered devices described in Ref. [25]. Due to symmetry through the drain ohmic contact, only one finger of the device is modeled for electro-thermo-mechanical characterization. Figure 1 depicts a representative structure for the modeled device with labeled thermal and structural boundary conditions.

**Electro-Thermal Model**

The electro-thermal behavior in 2 dimensions was captured using drift-diffusion electron and hole carrier equations using the commercially available Sentaurus Device simulator [26]. This software couples the electron and hole transport equations with the thermodynamic model to obtain the current densities within a semiconductor device. In doing so, it is possible to determine the electric potential distribution and Joule heating profile within the device as a function of applied gate and drain electrical bias ($V_{gs}$ and $V_{ds}$, respectively).

![Figure 1. Representative device geometry with structural and thermal boundary conditions. Two critical regions within the device are the gate foot print (GFP) where high amounts of stress develop and the gate connected field plate (GCFP) which is probed optically for stress characterization [6, 27-31]. Vertical dashed lines correspond to line positions labeled on Figure 3.](Image)

Due to the computational intensity associated with transient semiconductor physics, only a small portion of a full, packaged device is modeled for the electro-thermal simulation. The size of the electro-thermal modeling is adapted from Ref. [20] and contains 20 µm wide epilayers with the gate centered within the domain. The height of the AlGaN and GaN layers are 20 nm and 1.8 µm [24], respectively. A thermal boundary resistance is applied at the GaN/SiC interface to account for the thermal resistance of the AlN nucleation layer [6, 29]. A portion of the 6H-SiC is included as the bottom half of a multi-faced polygon with an imposed thermal boundary resistance to account for the remaining substrate and device package. This shape is based on heat radiating away from the drain side of the gate structure in nearly circular isothermal lines within the substrate, as described in Refs. [20] and [32]. The material properties for the electro-thermal model including temperature dependent thermal conductivities, heat capacities, and temperature- and mole-fraction-dependent semiconductor band gaps have been taken from [20] and [29]. A more detailed description of the electro-thermal model is described in Ref. [25]. The validity of using this software to accurately model these devices has been demonstrated in Refs. [6, 11, 20, 25, 27, 29].

**Mechanics Model**

To calculate the transient mechanical response of the modeled device, the electrical potential and the thermal heat generation distributions are transferred to a large-scale, COMSOL Multiphysics model. Here, the resulting combined stress distribution can be determined, or the individual stress contributions due to piezoelectric and thermoelastic effects can be decoupled. Within the mechanics model, anisotropic thermal and structural material properties were used wherever possible (see Table 1).
Table 1. Thermal and Structural Properties of Materials.

<table>
<thead>
<tr>
<th>Material</th>
<th>Coefficient of Thermal Expansion ($\times 10^{-6}$ K$^{-1}$)</th>
<th>Thermal Conductivity (Wm$^{-1}$K$^{-1}$)</th>
<th>Stiffness matrix (GPa)</th>
<th>Young's Modulus (GPa)</th>
<th>Poisson's Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN [33,36]</td>
<td>$a_d = -1.44 \times 10^{-5}T^2 + 1.7 \times 10^{-2}T + 0.553$</td>
<td>150 x $\left(\frac{T}{300}\right)^{-1.4}$</td>
<td>C$<em>{11}$ 390 C$</em>{12}$ 145 C$<em>{44}$ 105 C$</em>{13}$ 106 C$_{15}$ 398</td>
<td>- - - - -</td>
<td>- -</td>
</tr>
<tr>
<td>$a_c = -1.39 \times 10^{-5}T^2 + 1.64 \times 10^{-2}T + 0.216$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiC [33, 37, 38]</td>
<td>$a_d = -1.36 \times 10^{-6}T^2 + 3.99 \times 10^{-3}T + 2.28$</td>
<td>387 x $\left(\frac{T}{293}\right)^{-1.49}$</td>
<td>501 111 163 52 553</td>
<td>- - - - -</td>
<td>- -</td>
</tr>
<tr>
<td>$a_c = -8.51 \times 10^{-7}T^2 + 2.94 \times 10^{-3}T + 2.44$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Si$_3$N$_4$ [39, 40]</td>
<td>3.3</td>
<td>4.5</td>
<td>- - - - -</td>
<td>195 0.25</td>
<td>- -</td>
</tr>
<tr>
<td>Au [41]</td>
<td>14.2</td>
<td>317</td>
<td>- - - - -</td>
<td>97 0.42</td>
<td>- -</td>
</tr>
<tr>
<td>AlGaN [6]</td>
<td>4.34</td>
<td>30</td>
<td>393 142 126 105 395</td>
<td>- - - -</td>
<td>- -</td>
</tr>
</tbody>
</table>

In addition to the properties found in Table 1, density and specific heats for the listed materials are taken [24]. Unlike bulk materials, thin film properties are difficult to determine and a wide range of values exists for the elastic modulus, thermal expansion coefficient, and Poisson ratio [33, 42]. Such a large range of values can have a significant effect on the overall simulation, and therefore material properties must be selected with care. For this work, material properties were chosen from literature based upon values that have been previously verified through numerical simulations and experimental results [6, 29].

The large-scale mechanics model includes the full width and height of a device including a 100 x 100 $\mu$m SiC substrate [43]. AlGaN and GaN heights are matched to those in the small-scale electro-thermal model. In addition, the GaN layer extends the full width of the SiC substrate. The topography of the passivating Si$_3$N$_4$ layer is determined through AFM cantilever scans of an experimental device. This topography is a conformal layer covering the device metallizations (drain, gate, and source contacts) and extends the full width of the GaN and SiC substrate.

Similar to the Sentaurus Device model, a thermal boundary of 300 K was applied to the bottom of the SiC substrate and adiabatic boundary conditions were placed on all other surfaces, including the symmetry line through the center of the drain (see Figure 1). A fixed mechanical boundary condition was placed on the bottom of the package and an x-direction fixed mechanical boundary condition was placed along the symmetry line, which is consistent with [29], and all other surfaces are unrestrained. A thermal boundary resistance of 60 m$^2$KGW$^{-1}$ is placed at the GaN/SiC interface to account for the AlN nucleation layer [44] (domain not physically modeled for computational speed). Finally, the reference strain temperature is set to 300 K. Therefore, the calculated stresses are only due to electro-thermal operating effects.

RESULTS AND DISCUSSION

Prior to detailing the transient response of a device, the model is compared to experimental DC electrical data corresponding to a 2-finger AlGaN/GaN HEMT and previously presented bias dependent Joule heating data available from literature. Next, the transient stress response under pulsed mode operation is determined.

DC Results

The numerical model was validated by comparing simulated and experimental $I_D$-$V_{DS}$ responses under DC bias (Figure 2). Good agreement is seen between the numerical model and the experimental data across a broad range of bias conditions ($V_{DS}$ increases from -3 V to +1 V).

![Figure 2. Comparison of numerical and experimental device characteristics showing model accuracy across a wide range of bias conditions. Numerical model (solid line) captures the electro-thermal characteristics of a device including the current droop caused by self-heating.](Image)

Solid lines and symbols correspond to the numerical model and experimental results, respectively. In addition, the model effectively accounts for the current “droop” due to the Joule heating within the device [21]. Next, three DC bias conditions of equal power dissipation (6 Wmm$^{-3}$) are compared to demonstrate bias dependent electric field and Joule heating. Previous research has shown the temperature in these devices will
have a strong bias dependence due to the confinement or spreading of the heat generation region [27, 28]. In our work, \(V_{ds}\) is set to 10, 28, and 48 V while \(V_{gs}\) is -0.197, -2.072, and -2.539 V, respectively. Figure 4 shows the electrical bias impact on the electric field (top row) around the gate contact and the resulting Joule heating profile (bottom row) for the three bias conditions (columns) on the drain side of the gate contact. These results are consistent with similar numerical models from literature [27]. By varying the \(V_{ds}\) and \(V_{gs}\) bias conditions, the electric field around the gate structure is altered, resulting in a change in the Joule heating profile within the device. For GCFP devices, two peaks in electric field exist: one on the drain side of the gate at the GFP corner and another under the GCFP around the AlGaN/GaN interface on the drain side. The 48 V\(_{ds}\) case has the largest electric field at the GCFP corner and extends to the AlGaN/GaN interface in the device channel. A larger electric field under the GCFP causes an elongation in the Joule heating profile towards the drain (located to the left of the gate structure in Figure 4). This elongation drastically changes the thermal profile along the AlGaN/GaN interface, as demonstrated in Figure 3. At the interface, the elongation in Joule heating profile results in a shift in peak temperature. The GCFP position is denoted as the left dashed line, and the GFP position is denoted by the right dashed line (these correspond to the vertical dashed lines from Figure 1). Each case has nearly the same peak temperature, but the 10 V\(_{ds}\) case has a temperature distribution locally confined to the GFP. As \(V_{ds}\) is increased to 48 V, the increase in electric field around the GCFP causes the peak temperature location to shift by \(\sim 200\) nm towards the drain and into the channel of the device. This is on the order of the length of the GCFP and will change with varying device structures.

Proper handling of the electro-thermal coupling of an AlGaN/GaN HEMT is necessary for proper stress characterization. The total stress within a device is a summation of the electrical, thermal, and residual contributions. In this work, just the stress contributions due to electro-thermal effects are accounted for, but it is important to note the device will have additional residual stress due to device fabrication effects. For the 10, 28, and 48 V\(_{ds}\) cases, the electrical stress due to the inverse piezoelectric effect (IPE) is found to be 230, 323, and 365 MPa, while the thermoelastic stress is 249, 247, and 238 MPa, respectively.

![Figure 3. Temperature profile along the AlGaN/GaN interface for \(V_{ds}\) = 10, 28, and 48 V. Peak temperature position shifts towards the drain with increase drain bias due to increased electric field at the gate connected field plate (GCFP). Symbols are added for line identification purposes only.](image)

![Figure 4. Electric field distribution (top row) around the drain side of the gate structure and resulting Joule heating (bottom row) profile. As drain bias increases (from left to right column), the electric field at the GCFP increases causing an elongation in Joule heating.](image)
AC Results

Next, the response of the AlGaN/GaN HEMT to pulse conditions is determined. The simulated device is pulsed at 28 V\textsubscript{ds}, a constant gate bias of -2.072 V\textsubscript{gs}, and duty cycle of 50% at 100 kHz frequency from a starting temperature of 300 K to characterize the development of stress in the device. The model utilizes a transition period length of 100 ns from OFF-state, where V\textsubscript{ds} = 0 V and V\textsubscript{gs} = -2.072 V, to ON-state, where V\textsubscript{ds} = 28 V and V\textsubscript{gs} = -2.072 V. This creates a rapid shift in voltage to simulate actual operating conditions. These bias conditions were chosen based upon representative operating conditions for the device, and dissipate 6 Wmm\textsuperscript{-1} under DC operation. Figure 5 represents the waveform signal for the applied drain bias (V\textsubscript{ds}), the resulting dissipated power (P\textsubscript{ds}), and peak temperature (T\textsubscript{max}) during transient operation of the device.

The applied V\textsubscript{ds} value changes rapidly compared to the thermal response of the device. Because of this, the bulk of the piezoelectric portion of stress is expected to develop almost instantaneously due to the sudden voltage potential rise, while the thermal contribution to stress will increase throughout the ON-state portion due to the Joule heating in the device. The electrical and thermal response change during the ON-state because of the thermo-electrical coupling. This is seen as a reduction in dissipated power as peak operating temperature increases (Figure 5).

To characterize the transient development of stress during cyclic operation, two locations were chosen on the drain side of the gate: the gate foot print (GFP) edge and within the device channel near the gate connected field plate (GCFP). For both locations, the stress is averaged over a small area (5 x 1 nm) within the AlGaN layer. Both areas contain a high mesh density and are located away from material corners where stress concentrations can occur (particularly near the GFP where SiN\textsubscript{x}, Au, and AlGaN intersect). At the GFP, large amounts of stress have been shown to develop due to the inverse piezoelectric effect (IPE) and large thermal gradients. A point near the GCFP is chosen because this area can be probed relatively easily using optical techniques. Figure 6 and Figure 7 represent the transient stress response of the device at the GFP and GCFP regions, respectively. See Figure 1 for respective locations.

At the GFP edge (above), both the IPE and thermoelastic stresses are tensile, combining to an even higher state of stress. Because of the proximity to the Joule heating region, both contributors to stress develop rapidly with applied drain bias. The IPE stress, however, remains constant with applied bias while the thermoelastic stress builds throughout the ON-state of the device. This is due to the rapidly applied and constant V\textsubscript{ds}/V\textsubscript{gs} electrical bias during the ON-state. At the OFF-state, the IPE stress is reduced immediately as V\textsubscript{ds} returns to 0 V, and the thermoelastic stress relaxes until the next ON-state. IPE stress is non-zero during the OFF-state due to the V\textsubscript{gs} bias, which is held constant during pulsing.

Different transient stress profiles are seen within the channel at the GCFP. Here, the IPE stress is still tensile, but the thermoelastic stress is compressive (Figure 7).

![Figure 5. Pulsed-mode drain voltage with resulting power and peak temperature. Rapid transition between OFF-state (V\textsubscript{as} = 0 V) and ON-state (V\textsubscript{as} = 28 V) mimics real device behavior. Symbols are added for line identification purposes only.](image1)

![Figure 6. Transient stress values at the gate foot print (GFP) during pulsing. Both IPE and thermoelastic contributions to stress are highly tensile, and both portions of stress develop rapidly with applied bias due to proximity to the Joule heating region. Symbols are added for line identification purposes only.](image2)

![Figure 7. Transient stress values at the gate connected field plate (GCFP). IPE stress is tensile and remains constant during the ON-state while the thermoelastic stress is compressive and builds throughout the ON-state. Symbols are added for line identification purposes only.](image3)
In this location, a rapid, initial increase in IPE stress is seen at the start of the ON-state prior to a sharp decrease in total stress due to the onset of compressive thermoelastic stress. As before, the IPE stress remains constant with applied bias while the thermoelastic stress builds throughout the ON-state. At the OFF-state, the IPE stress releases nearly instantly, and the total stress state shifts to match the compressive thermoelastic stress. Again, the thermoelastic stress reduces throughout the duration of the OFF-state as the device temperature decreases.

Figure 8. Resulting stress profile around the gate structure at the end of the ON-state. Large stress gradients are seen between the GFP and GCFP areas.

The difference of stress values between the two locations is attributed to the GFP position’s proximity to highly varying CTE materials and the Joule heating profile. Areas around the gate structure undergo complex cyclic loading, and it is important to understand the stress distribution and how it changes during cyclic operation. Figure 8 illustrates the spatial distribution of stress around the gate structure for the time corresponding to just before the end of the first ON-state \((t = 1.0 \times 10^5 \text{ seconds})\). Horizontal lines within the AlGaN and GaN layers act as interpolation lines for the electric field and Joule heating profiles from Sentaurus Device and aid in generating a highly dense mesh scheme in this area. In addition to the large peak stress at the GFP, a high gradient of stress exists around the gate structure, especially within the AlGaN layer. During pulsed operation, these values are rapidly applied and released, which could induce fatigue-like failure mechanisms in devices.

CONCLUSION

Two dimensional electro-thermo-mechanical simulations were conducted to determine the magnitude and development of stress within an AlGaN/GaN HEMT under pulsed bias conditions as a function of position around the gate structure. It was shown complex stress profiles occur within AlGaN/GaN HEMTs under transient operating conditions. The electrical stress develops rapidly with applied bias, while the thermoelastic stress builds during the ON-state of a device. Areas of interest within the AlGaN layer including the gate foot print (GFP) and outside of the gate connected field plate (GCFP) undergo vastly different transient stresses due to the electrical and thermal profiles. At the GFP, the electrical and thermoelastic stresses are tensile, resulting in an extremely localized concentration of tensile stress that is suspected to cause reliability issues for these devices. At the GCFP, however, the electrical stress is tensile while the thermoelastic is compressive, leading to an overall relaxed stress state in this area of the device. Under transient operation, these points undergo heavy amounts of cyclic loading, which could induce failure mechanisms not seen under DC operating conditions, where most of the reliability studies for AlGaN/GaN HEMTs have been performed.

ACKNOWLEDGEMENTS

This work was supported by the Air Force Research Laboratory High-Reliability Electronics Virtual Center (HiREV) team and AFOSR.

REFERENCES


