JOURNAL OF APPLIED PHYSICS **111**, 074504 (2012)

Fully coupled thermoelectromechanical analysis of GaN high electron mobility transistor degradation

M. G. Ancona,[a] S. C. Binari, and D. J. Meyer
Naval Research Laboratory, Washington, DC 20375, USA

(Received 1 November 2011; accepted 25 February 2012; published online 5 April 2012)

A fully coupled multi-dimensional continuum model of the thermoelectromechanics of GaN HEMTs is presented and discussed. The governing equations are those of linear thermoelectroelasticity, diffusion-drift transport theory, and heat conduction, with full coupling assumed, i.e., all mechanical, electrical, and thermal variables are solved for simultaneously. Apart from the known strains induced by epitaxy, plane-strain conditions are assumed, so that two-dimensional simulation suffices. Important aspects of the model are that it incorporates "actual" device geometries and that it captures field/stress concentrations that often occur near material discontinuities and especially at corners. The latter are shown to be especially important with regards to understanding the mechanisms of both electrical and mechanical degradation in GaN HEMTs. Various possible contributors to degradation are discussed, including electron injection, the inverse piezoelectric effect, thermal stress, SiN intrinsic stress, and device geometry. The possibilities of crack propagation and fracture of the AlGaN are also analyzed. [http://dx.doi.org/10.1063/1.3698492]

I. INTRODUCTION

As is well known, in epitaxial semiconductor heterostructures, it is possible to obtain levels of elastic deformation well beyond those achievable in conventional mechanical testing, e.g., reaching strains of 2% or more. These extreme elastic states are widely exploited to provide a variety of benefits, such as mobility enhancement through band structure distortion (e.g., in silicon[1] or the antimonides[2]) or enhanced channel charge via piezoelectrically induced polarization, as in the nitrides.[3] Of course, there are limits to how far such strain-based strategies can be pushed before material failure intercedes, with the primary restriction being that of the critical thickness at which the strained layer undergoes uniform, irreversible relaxation. In the case of AlGaN strained layers in conventional GaN HEMTs, there is another limit, observed following sustained operation at high current/voltage levels, wherein highly localized mechanical failures are seen to occur (see Fig. 1).[4,5] Localized mechanical stress/strain in the AlGaN has also been implicated in the generation of traps in GaN HEMTs that, in turn, produce electrical degradation.[6,7] The precipitating cause(s) of these localized mechanical/electrical degradation pathways has (have) yet to be definitively established, and the question of their origin is the underlying motivation for the present work, in which we develop a fully coupled, multi-dimensional thermoelectroelastic simulator for GaN devices. We then use this code to explore the thermoelectromechanics of GaN HEMTs to help understand their modes of degradation. In addition, we illustrate the simulator's potential value as a tool for designing new device configurations with enhanced reliability.

Understanding electrical/mechanical degradation and failure in GaN HEMTs is critical, because of the growing technological importance of these devices for both power and RF electronics. Indeed, a significant fraction of the overall research and development (R&D) effort in GaN electronics is devoted to understanding failure modes and to developing more robust devices. From a reliability perspective, the portion of a conventional GaN HEMT of most concern is the AlGaN barrier layer that, being centrally located, highly strained, and often sustaining large electric fields, tends to be most vulnerable to electrical and mechanical degradation. Also of concern is the SiN layer that is commonly added as a means of partially passivating traps that can cause gate lag and loss of power output.[8] Of the many reliability studies of GaN HEMTs, two sets of papers are most germane for our work. The first, by Joh and del Alamo,[6,7] focused on electrical damage and found that electron traps were generated in the AlGaN barriers of their GaN HEMTs, often preferentially on the drain side and when the gate bias exceeded a critical value. The second set of papers of interest used post-mortem TEM,[4] AFM,[5] or SEM[5] (see Fig. 1) to show that electrical degradation/failure in GaN HEMTs was often accompanied by mechanical/structural damage in the form of pits and/or cracks in the AlGaN layer at the edge of the gate metal, and again preferentially on the drain side. As to the trigger of the observed electrical and mechanical degradation, all of these papers favored a mechanism in which the localized piezoelectric stress induced by the applied voltages raises the total stress in the AlGaN above that needed to cause "lattice damage", which in turn can serve as electron traps; as paths for gate current, electromigration, or chemical reactivity; and/or as initiators for pit/crack formation. While certainly an intriguing idea, because of the complexity of the overall situation, this piezoelectric failure mechanism can hardly be regarded as definitively established.

Motivated by the aforementioned studies of highly localized mechanical/electrical degradation in GaN HEMTs, the present paper uses fully coupled numerical simulations to explore in detail the thermoelectromechanical conditions that prevail during electrical stressing, that could trigger

---

[a] Electronic mail: ancona@estd.nrl.navy.mil.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE
05 APR 2012

2. REPORT TYPE

3. DATES COVERED
00-00-2012 to 00-00-2012

4. TITLE AND SUBTITLE
Fully coupled thermoelectromechanical analysis of GaN high electron mobility transistor degradation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory,4555 Overlook Ave SW,Washington,DC,20375

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE
unclassified	**unclassified**	**unclassified**

17. LIMITATION OF ABSTRACT
Same as Report (SAR)

18. NUMBER OF PAGES
17

19a. NAME OF RESPONSIBLE PERSON

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

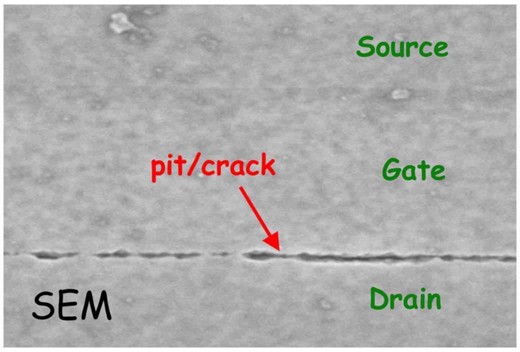

FIG. 1. SEM image of the AlGaN surface of a severely degraded GaN HEMT from which the electrodes and SiN have been stripped. Similar images appear in Ref. 5.

device failure, and that might be mitigated by appropriate device design. Our theoretical approach is the conventional one, based on linear electroelastic theory;[3,9–11] however, in contrast to most previous work, our treatment is not restricted to one dimension, instead taking full account (within a continuum framework) of the multi-dimensional device geometry and of localized behaviors, such as electric field or mechanical stress concentrations that may be crucial to understanding the observed damage. In addition, we include the coupling between the electromechanics and the electron transport (within the diffusion-drift approximation), as well as the effect of thermal stresses induced by Joule heating. Although quite general, it is important to recognize that these coupled continuum equations do not provide direct representations of the failure, as they include neither traps nor the possibility of stress-enhanced trap generation, and they assume elastic behavior and so do not encompass material failure. Given this, we proceed simply by monitoring specific metrics in order to judge where and when failure might be expected to occur. In particular, as an *electrical* metric, we use the threshold electric field at which significant Fowler-Nordheim injection into the AlGaN barrier can be expected, and as a *mechanical* metric, we use the size of the principal stresses as compared to the AlGaN tensile strength. One could also postulate a *thermal* metric for failure or conceivably even a *chemical* metric, but this is not done in this paper.

For the present paper, three types of mechanical deformations are of interest. First, there is the strain introduced by the growth of the heterostructures using the epitaxial techniques of MBE or metalorganic chemical vapor deposition (MOCVD) (with the layer thicknesses assumed to be below the critical value(s) at which relaxation occurs[12]). Such strains, which, as already noted, tend to be quite high, exist in three dimensions and consist of an in-plane biaxial strain imposed by the epitaxial growth plus an elastic response in the vertical direction associated with Poisson's ratio. The second type of deformation of interest enters when the as-grown pre-stressed heterolayers are further processed, e.g., when a silicon nitride passivation layer is added or etched. Because the devices of interest tend to be much wider than they are long, these additional strains are usually 2D plane-strains that are uniform across the width of the device. Finally, the third type of deformation of concern arises from the piezoelectric and thermoelastic nature of the materials and occurs when voltages are applied to the contacts of the device. Given the usual device configurations, these strains are also typically 2D plane-strains.

The organization of the paper is as follows: In Sec. II, we detail the equations of thermoelectroelasticity, upon which our work is based, and outline the numerical approach used for their solution. Following a brief discussion of verification in Sec. III, we then use thermoelectromechanical simulations in Sec. IV to examine GaN FETs under normal operating conditions, to investigate certain GaN HEMT failure scenarios under stress-bias conditions, and to study a few possible mitigation strategies. A final section provides some concluding remarks.

II. EQUATIONS OF THERMOELECTROELASTICITY

Our thermoelectromechanical model for AlGaN/GaN heterostructure devices is based on the continuum theory appropriate for piezoelectric semiconductors. It consists of the fully coupled equations of electroelasticity, diffusion-drift transport, and heat conduction, with steady-state conditions assumed. This means that, inside semiconducting regions (e.g., GaN), the governing equations are

$$\nabla\cdot\mathbf{D}=q(N-n)\quad \nabla\cdot\mathbf{J}=0\quad \mathbf{J}=qn\mu_n\nabla(\varphi_n-\psi)\quad [\text{semiconductor}],\tag{1a}$$

$$\nabla\cdot\tau=0\qquad \nabla\cdot\mathbf{q}=\eta\mathbf{J}\cdot\nabla(\varphi_n-\psi),\tag{1b}$$

where N is a possible bulk charge density (due to ionized impurities or fixed charge), τ is the mechanical stress tensor (which, by convention, is positive when tensile and negative when compressive), $\mathbf{q}$ is the heat flux vector, the right side of $(1a)_2$ being zero means we neglect electron generation-recombination effects, and the right side of the last equation in $(1b)_2$ gives the Joule heating associated with the diffusion-drift transport (with no contribution from recombination). The parameter η is an efficiency factor, normally equal to one, that has been introduced in order to represent the effect of RF output power delivery in one set of simulations in Sec. IV. All other variables and parameters have their usual meanings. Inside of solid insulators, which we assume to include the AlGaN, the lack of appreciable electron transport implies that the differential equations reduce to

$$\nabla\cdot\mathbf{D}=qN\quad \nabla\cdot\tau=0\quad \nabla\cdot\mathbf{q}=0\qquad [\text{insulator}].\tag{1c}$$

Inside metals, the equations simplify even further to

$$\nabla\cdot\tau=0\quad \nabla\cdot\mathbf{q}=0\qquad [\text{metal}],\tag{1d}$$

where it is assumed that the Joule heating is negligible. Finally, within the space overtop the structure or inside cracks, we ignore mass, momentum, and thermal transport, and so the only equation to be solved is that of electrostatics,

$$\nabla\cdot\mathbf{D}=0\qquad [\text{air/vacuum}].\tag{1e}$$

The system of equations in Eq. (1) is completed by supplying various constitutive equations that incorporate the specifics of material response. To represent the electroelastic behavior, in all solid materials we need expressions for the mechanical stress tensor τ and (except in metals) for the electric displacement vector $\mathbf{D}$. As noted earlier, we restrict our treatment to linear theory, which means that we assume these quantities depend only linearly on the electric field vector $\mathbf{E} = -\nabla\psi$ and on the linearized strain tensor $\mathbf{S} = (\nabla\mathbf{u} + \nabla\mathbf{u}^{\mathrm{T}})/2$, where $\mathbf{u}$ is the mechanical displacement vector. In indicial notation (with the Einstein summation convention), the appropriate constitutive equations are

$$\begin{aligned} \tau_a &= \tau_a^0 + c_{ab}\left[S_b - \alpha_b\left(T - T_p\right)\right] + e_{ka}\psi_{,k} \\ D_{\mathrm{i}} &= P_i^0 + e_{ia}S_a - \varepsilon_{ik}\psi_{,k}, \end{aligned} \tag{2a}$$

where the indices i, j, and k range over 1, 2, and 3 (corresponding to the directions x, y, and z), the standard shorthand indices a and b range from 1 to 6 (with the correspondences to the indices ij being $1 = 11$, $2 = 22$, $3 = 33$, $4 = 23$ and 32, $5 = 13$ and 31, and $6 = 12$ and 21), and the comma-subscripts denote partial differentiation. The quantities P_i^0 form the spontaneous polarization vector, the τ_a^0 account for any intrinsic or built-in stress, the ε_{ik} form the permittivity tensor, the c_{ab} are the elastic constants, the α_a are the thermal expansion coefficients, the e_{ia} are the piezoelectric constants, and T_p is the baseplate temperature. It should be noted that we ignore the pyroelectric effect (i.e., temperature dependence of the spontaneous polarization) in Eq. (2a), since it is believed to be small in the nitrides over the temperature range of interest.[13] Consistent with the linear approximation we also ignore Maxwell stresses which are readily shown to always be less than about 100 MPa. The values for the various material coefficients are listed in Table I,[14–16] with alloys treated by linear interpolation, except in the case of the thermal conductivity of AlGaN that is known to be much smaller than that for either of its constituent binaries.[17] Finally, we note that the numbers given in Table I for PECVD silicon nitride are "average" values.

TABLE I. Selected material constants used in the simulations.

Material	GaN	AlN	SiN	Au	W
Reference	3	3	14	15	16
$\alpha_{11} = \alpha_{22}$ **(10^{-6}K)**	5.59	4.2	1.5	14.2	4.3
α_{33} **(10^{-6}K)**	3.17	5.3	1.5	14.2	4.3
K **(W/cm-K, 300 K)**	1.38	2.85	0.15	3.18	1.73
C_{11} (GPa)	374	345	187	250	533
C_{12} (GPa)	106	125	69	181	204
C_{44} (GPa)	101	118	118	68	161
C_{13} (GPa)	70	120	69	181	204
C_{33} (GPa)	379	395	187	250	533
$\varepsilon_{11} = \varepsilon_{22}$	9.5	9.0	9.4	…	…
ε_{33}	10.4	10.7	9.4	…	…
e_{15} (C/m^2)	–0.3	–0.48	…	…	…
e_{31} (C/m^2)	–0.32	–0.38	…	…	…
e_{33} (C/m^2)	0.63	1.29	…	…	…
P_3 (C/m^2)	–0.029	–0.081	…	…	…

Regarding the electron gas in the GaN, we assume a Fermi-Dirac expression for the electron chemical potential $\varphi_n = \varphi_n(n)$, with temperature and strain dependences (that would produce thermoelectric and piezoresistive effects[18]) being neglected. Furthermore, we take the following form for the (scalar) electron mobility μ_n that incorporates velocity saturation:

$$\mu_n = \mu_{n0}\left(\frac{300K}{T}\right)^{3/2}\left[1 + \left(\frac{\mu_{n0}|\partial\psi/\partial x|}{v_n^{sat}}\right)^{\beta}\right]^{-1/\beta}, \tag{2b}$$

where the x direction is the predominant current flow direction, μ_{n0} is arbitrarily taken to be 1200 cm^2/V-sec, the saturation velocity is $v_n^{sat} \approx 2.3 \times 10^7$ cm/sec, and β is assumed to be one. The temperature dependence in Eq. (2b) is chosen to represent the mix of polar optical phonon and piezoelectric scattering believed to dominate,[19] and any effects of anisotropy and of strain[19] have been neglected. Lastly, for the thermal constitutive equation, we assume the Fourier law,

$$\mathbf{q} = -\kappa\nabla T, \tag{2c}$$

with the thermal conductivity tensor κ taken to be temperature-dependent in each region and a scalar everywhere, but in the SiC substrate.

The structures of interest are grown epitaxially, and as a result these heterostructures are orthotropic in nature, being composed of layers of $\mathrm{Al_xGa_{1-x}N}$ of various compositions, all assumed to be below critical thickness. In addition, we take all of these materials to be so-called Ga-face with wurtzite crystal symmetry and with the c-axis oriented in the growth direction which, in our coordinate system is the z-axis (see Fig. 2). Of the types of deformation that such structures can undergo (see Sec. I), the strains introduced by the epitaxy are unique in being three-dimensional, while the other deformations associated with materials processing and with electrical biasing can be taken to be plane-strains residing in the x-z plane, since the devices of interest are much wider than they are long. Because of this, it is convenient to split off the in-plane portion of the epitaxial strain as follows:

$$\mathbf{S} = \bar{\mathbf{S}}^{(k)} + \mathbf{s}, \tag{3}$$

where $\bar{\mathbf{S}}^{(k)}$ contains the in-plane (x-y) components of the epitaxial strain field in the kth layer. Since this strain results from the given layer accommodating itself to the unstrained GaN substrate, we have

$$\bar{\mathbf{S}}^{(k)} = \left(S_1^{(k)}, S_1^{(k)}, 0, 0, 0, 0\right), \quad \text{where} \quad S_1^{(k)} \equiv \frac{a_k - a_{sub}}{a_{sub}}, \tag{4}$$

where a_k and a_{sub} are the in-plane lattice constants of the kth layer and the substrate, respectively. With the in-plane epitaxial strain taken care of, $\mathbf{s}$ (unlike $\mathbf{S}$) will be a pure plane-strain in the x-z plane. In response to the imposed epitaxial strain, each overlayer will, of course, expand/contract in the third (z) direction, but because this response will depend on the material coefficients in Eq. (2), screening effects, etc., we simply include it in the overall calculation for $\mathbf{s}$.

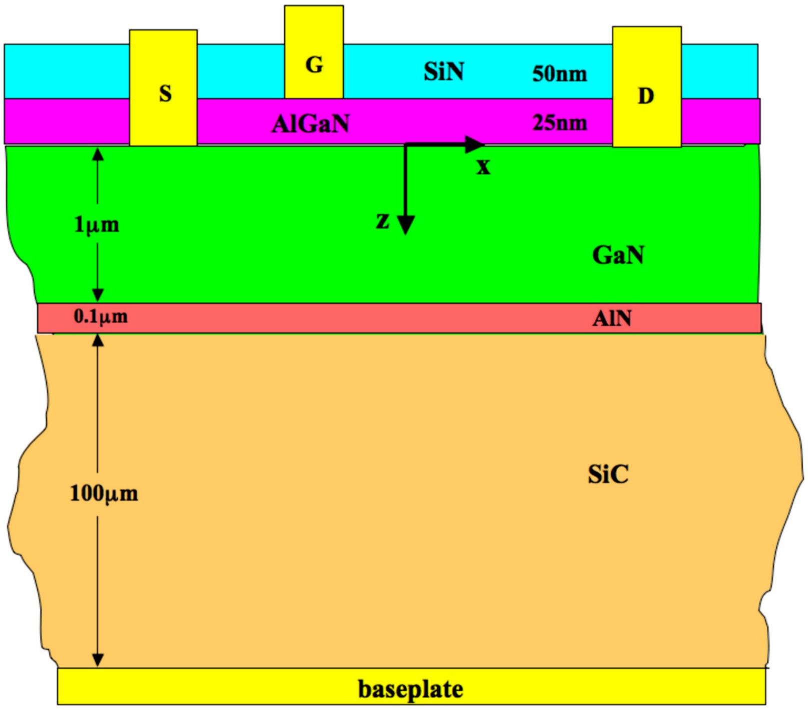

FIG. 2. Schematic showing the structure of the GaN HEMT simulated in this paper and the coordinate system used in its representation. The dimensions are not to scale.

Being under plane-strain conditions means that the electric field component in the y-direction, the strain component in the y-direction (apart from that in Eq. (4)), and any y-derivatives will be negligible. Furthermore, the electric displacement and stress components in the y-direction need not be computed, since they are simply the "forces" that serve to maintain the plane-strain. As a result, the only components in Eq. (2) of interest are D_1, D_3, τ_1, τ_3, and τ_5, which, for our oriented wurtzite crystal symmetry, take the following forms in layer k:

$$\begin{aligned}
D_1 &= e_{15}s_5 - \varepsilon_{11}\psi_{,1}\\
D_3 &= P_3^0 + e_{31}\left[s_1 + 2S_1^{(k)}\right] + e_{33}s_3 - \varepsilon_{33}\psi_{,3},\\
\tau_1 &= c_{11}\left[s_1 + S_1^{(k)} - \alpha_1(T - T_p)\right] + c_{12}\left[S_1^{(k)} - \alpha_1(T - T_p)\right]\\
&\quad + c_{13}\left[s_3 - \alpha_3(T - T_p)\right] + e_{31}\psi_{,3},\\
\tau_3 &= c_{13}\left(s_1 + 2S_1^{(k)} - 2\alpha_1(T - T_p)\right) + c_{33}\left[s_3 - \alpha_3(T - T_p)\right]\\
&\quad + e_{33}\psi_{,3}\\
\tau_5 &= c_{44}S_5 + e_{15}\psi_{,1}, \qquad (5)
\end{aligned}$$

where $S_1^{(k)}$ is non-zero only for the AlGaN layer, P_3^0 is the only non-zero component of the spontaneous polarization, and, for now, we ignore the possibility of intrinsic stress τ_a^0. Inserting Eq. (5) into Eq. (1) then gives five differential equations to be solved for the unknown fields u_x, u_z, ψ, n, and T.

To formulate and solve boundary value problems involving the foregoing differential equations, one must also supply a consistent set of boundary conditions. For the most part, these conditions are standard; however, a few deserve further elaboration. One such condition is the ordinary electrostatic one on the electric displacement, which at free surfaces must include the additional charges that accumulate from the air and that act to neutralize the polarization charge. The interfaces between GaN (or AlGaN) and deposited SiN are usually also taken to be electrically neutral (perhaps because, prior to the SiN deposition, the GaN surfaces are exposed to air). A second electrical boundary condition worth mentioning is the one representing source/drain contact resistance. Our approach is to take the contact resistance to be independent of temperature[20] and choose its value so that the simulated drain current in the ON-state (with $V_{GS} = 0$ V and $V_{DS} = 5$ V) is roughly 1 A/mm, so as to match measurements on state-of-the-art devices. An alternative approach not employed here would use the low-field mobility for this curve-fitting.

A more complicated boundary-related issue relates to heat-sinking and to the conditions to be applied at the lateral and bottom edges of the simulation region. In the actual device, the Joule heat is sunk largely through a metal baseplate upon which the chip sits, with very little heat flowing out of the topside contacts. Now, a complete treatment would entail modeling not only the chip itself, but also its packaging, an inherently three-dimensional problem. This is not pursued here because, in addition to being computationally intensive, it does not seem likely that such an analysis would yield new insights into the thermomechanics of GaN devices. Therefore, we simply ignore the 3D effects and, as discussed further below, introduce a small substrate boundary resistance in order to represent the effect of the packaging at the backside.

As depicted in Fig. 2, the substrate/device design considered in this paper consists of a 100-μm-thick 4H-SiC wafer, on top of which are grown a 0.1-μm-thick buffer layer of AlN, a 1-μm-thick GaN layer, and a 25-nm AlGaN layer, with the device defined in the latter two layers. To enhance power output, the full device structure is configured to have multiple gate fingers, each 125-μm-wide and spaced by 50 μm. Approximating this by an array of infinitely wide devices, we can then reduce our analysis to that of a single two-dimensional region that includes the cross-section shown in Fig. 2, is approximately 50×101 μm in size, and has periodic boundary conditions applied to its lateral boundaries. But even this reduced problem is computationally challenging if the fully coupled transport, stress, and heat equations are solved throughout. To improve efficiency, we therefore solve only for the thermal and mechanical variables in the full 50×101 μm region and confine the solution for the coupled electrostatic and electron transport variables to just a small 1×3-μm region that encompasses the active device (see Fig. 8). For the electrical boundary conditions on the small region, we employ "charge-neutral" conditions at all semiconductor edges, with any error incurred being unimportant, since these edges are electrostatically distant from the gate region.

In Secs. III and IV, we set up and solve boundary value problems within the theory outlined above that allow us to examine various aspects of GaN HEMT operation and reliability. Because of the complexity of both the device geometry and the governing differential equations, these problems do not admit analytical solutions and our solution approach is necessarily numerical. For this purpose, we employ the finite element method on unstructured grids, using the flexible implementation provided by the COMSOL software package.[21]

Although the foregoing equations encompass many important facets of GaN HEMT device physics, as noted in the Introduction, they do not provide a direct representation of electrical/mechanical degradation processes. For this reason, we instead rely on certain metrics to judge when and where one might expect traps to be generated and/or the material to yield. In particular, we assume that the primary *electrical* factor influencing degradation/failure is electron injection into the AlGaN barrier and so use as our electrical metric the threshold electric field at which significant Fowler-Nordheim injection can be expected to occur. Assuming a metal-AlGaN Schottky barrier height of 1.35 eV (Ref. 22) and a "minimum" tunneling distance of ∼1.5-2 nm, the threshold electric field (in the AlGaN directly beneath the gate electrode) for injection is 7-9 MV/cm. Of course, use of this metric ignores the questions of the type of damage that is created and with what efficiency. Furthermore, it seems likely that local strain and/or temperature would accelerate the damage from electron injection, and these synergistic factors are neglected by our simple metric-based approach.

With respect to the *mechanical* aspects, we take the primary factor affecting degradation/failure to be excessive tensile stress and then use as a metric the size of the maximum principal stress, as compared to AlGaN's tensile strength. Unfortunately, there seems no information in the literature on the strength of AlGaN, and even data for the component binaries is quite limited and inconclusive. To get a rough idea of the value to expect, we briefly review several lines of evidence, while keeping in mind that, for brittle materials there is generally no single "correct" value, but rather a Weibull distribution with the strength of any particular sample depending on its precise geometry, dislocation density, etc. First, a crude theoretical estimate comes from noting that the energy required to create two non-polar GaN surfaces (that would form the faces of a crack) is about 2×0.12 eV/Å^2,[23] and with a lattice constant of about 3 Å, this means the energy density needed to produce the break is about 80 meV/Å^3 or 13 GPa. Experimentally, nanoindentation tests by Nowak[24] suggest a similar value of ∼15 GPa for superb-quality GaN made by high-pressure crystallization; as one might expect, lower values (<7.5 GPa) have been reported for GaN nanowires.[25] Another experimental perspective is provided by the in-plane stresses and strains that are generated in GaN/Al$_x$Ga$_{1-x}$N heterostructures of varying composition, as plotted in Fig. 3 together with the experimentally estimated critical thicknesses.[12] In considering this plot, it is important to recall that critical thickness failure is driven by the relief of epitaxial strain energy (which grows with film thickness) and so surely differs from the mechanism behind the failures of interest in this paper, which, being localized, relieve very little of the film's strain energy. (It is for this reason that strain energy is not a good metric for gauging the potential for pit/crack formation in GaN HEMTs, though it has been used as such[7]). Nevertheless, that good quality AlN layers of a few nanometers can be grown successfully on GaN[26] indicates that their yield strength is greater than ∼10 GPa. Finally, it is reasonable to expect that the material would be weakened by synergistic effects, such as those associated with elevated temperatures and/or with damage induced by electron injection. In this regard, Yonenaga[27] reported significant loss of strength in GaN (down to 100-200 MPa) for temperatures around 1000 °C; however, it is doubtful that much weakening would occur at the temperatures of interest in operating GaN HEMTs (<400 °C). Even more uncertain is the effect that mechanical fatigue due to repeated stress cyclings might have.[28] In any event, as a nominal mechanical metric, we assume, on the basis of the foregoing, a conservative value of ∼10 GPa as the tensile strength of AlGaN.

One final complication regarding the theory that deserves mention is the fact that, just as in elasticity and electrostatics, when corners/cracks are treated as

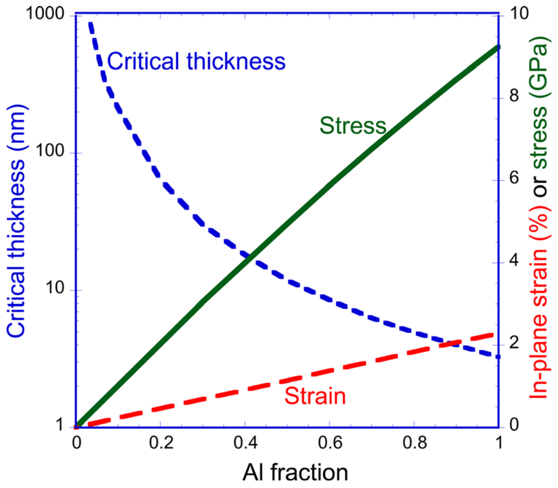

FIG. 3. The calculated stress and strain levels in an Al$_x$Ga$_{1-x}$N layer grown epitaxially on a GaN substrate as a function of the Al fraction. Also shown is the critical thickness as estimated in Ref. 12.

mathematically sharp (i.e., with vanishing radius of curvature), the solutions to the equations will often develop singularities. These infinities, which are of course spurious, come from the continuum theory ignoring the material's discreteness at the atomic level. As in dislocation theory, a rigorous analysis of the situation would join a continuum solution for the "far field" with a microscopic representation of the "core" region of the corner/crack. But because we do not know the precise geometry anyway, such an elaborate treatment seems unwarranted, and we instead simply ignore the microscopics entirely, truncating the continuum solutions with a cut-off distance of a few Angstroms. As illustration, for a GaN HEMT structure like that of Fig. 2, we plot in Fig. 4 the electric field profile along the gate/AlGaN interface as a function of the distance from the drain corner. The simulation finds that the field maximum at the corner, as computed numerically, grows without bound as the mesh is refined, appearing to become singular as $1/x^{1/3}$. As seen in the figure, the rapid increase in field is confined to positions within 1 Å of the corner, and it is inside this region that the continuum treatment is clearly invalid. Importantly, the electric field value just outside this "core" region depends only weakly on the choice of the cut-off distance (see Fig. 4), and this suggests that the error incurred (e.g., when estimating the maximum electric field) both in using such a cut-off and in its precise value will be small. On this basis, for this paper, we take the cut-off distance to be 3 Å or about the size of the unit cell.

III. VERIFICATION

Although the continuum theory set forth in Sec. II is the accepted description of piezoelectric semiconductors, the evidence for its applicability to GaN/AlGaN heterostructures is not as strong as one would like. For example, one publication reported a discrepancy between simulation and micro-Raman-based measurements of electric field of as much as a factor of 10.[29] One source of discrepancy could easily be that there is significant error in some of the values of the material constants for the ultra-thin layers of interest. More fundamentally, it may be that the linear constitutive theory on which the theory is based is inadequate. In any event, in this paper, we ignore all such issues and limit our attempts at "verification" to three consistency checks.

The first such check concerns a heterostructure formed of a GaN substrate with an epitaxial $Al_xGa_{1-x}N$ overlayer. In Fig. 5, we compare 1D analytical results from Ref. 3 (points) for the piezoelectric charge and the total polarization charge (piezoelectric plus spontaneous) at the AlGaN/GaN interface with results obtained by solving the equations of Sec. II numerically (dashed lines). Obviously, the agreement is excellent, thus demonstrating consistency with the work of Ref. 3.

The second check relates to the electrical characteristics of the conventional GaN HEMT depicted in Fig. 2. By solving an appropriate boundary value problem, the equations of Sec. II allow us to simulate the drain characteristics of the device, as shown in Fig. 6, where V_{GS} is stepped from 0 V to –4 V. (See Sec. IV for additional simulation results for this device). These characteristics are depletion-mode and look very much like those seen experimentally,[8] with an $I_{D\,\max}$ of about 1 A/mm (at $V_{GS} = 0$), a threshold voltage of about –4 V, and a slight drop in the current at high drain voltage due to Joule heating. As discussed earlier, the overall magnitude of the current depends on the low-field mobility and the contact resistance. The values used for these quantities in the simulation were 1200 cm^2/V-sec and ~ 0.7 Ω-mm, respectively, both of which are quite representative of actual devices.[8]

The final consistency check is one on temperature. An industrial vendor supplied us with information confirmed by micro-Raman measurements on the maximum temperatures in GaN test devices under a variety of conditions. Their estimated peak temperatures for a device like that of Fig. 2 when V_{DS} is ramped from 0 to 20 V (and the baseplate is at room temperature) are plotted in Fig. 7 (points). The corresponding simulation results with the back of the chip

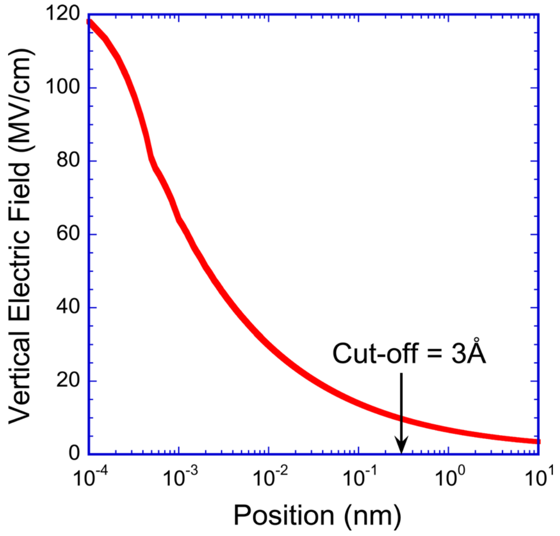

FIG. 4. Vertical profile along the AlGaN/gate interface just inside the drain corner showing the highly localized singularity in the electric field solution and justifying the choice of a cut-off distance of 3 Å.

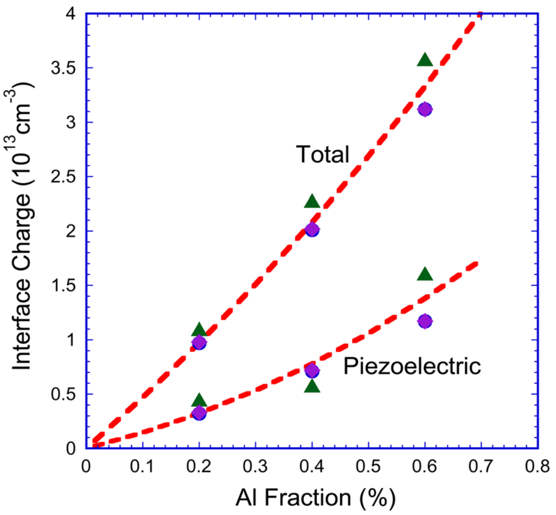

FIG. 5. Piezoelectric and total charge densities at a Ga-face GaN/AlGaN interface as computed using the equations in Sec. II and from Ref. 3.

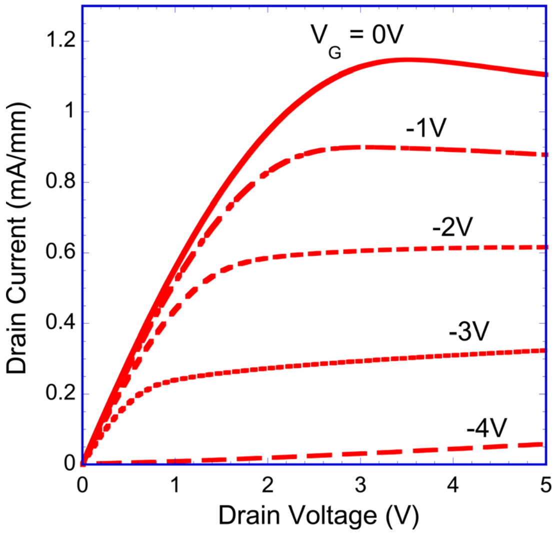

FIG. 6. Fully coupled thermoelectromechanical simulation of the drain characteristics of the GaN HEMT of Fig. 2.

assumed to be at room temperature agree to within about 30 °C (not shown), and given the many uncertainties in both the experiments and the modeling, this seems quite good. To further improve the agreement, we introduce a small thermal resistance at the substrate boundary that can be regarded as representing the packaging material that exists between the chip backside and the metal baseplate. Choosing the resistance so as to match the experiment at 8 W/mm gives the simulation result plotted in Fig. 7 (line). We believe that the main discrepancy that remains—the greater curvature in the simulated result, with deviations at the highest temperatures—originates in our neglect of 3D heat-sinking effects.

Obviously none of the consistency checks presented here constitutes true verification. Nevertheless, we regard them as sufficient justification for applying our thermoelectroelastic simulator to various GaN HEMT situations in hopes of gaining insight into the failure physics.

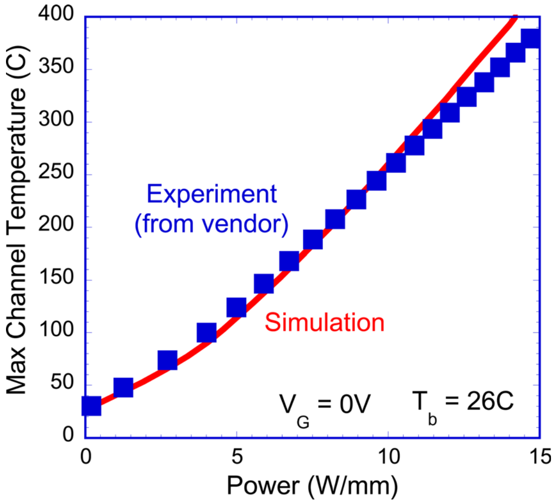

FIG. 7. Comparison between GaN HEMT peak temperatures as characterized experimentally by an industrial vendor and as estimated by numerical simulation with a small backside thermal resistance included.

IV. GaN HEMT SIMULATIONS

The GaN HEMT device examined in detail in this paper (and simulated above in Figs. 4, 6, and 7) is a simplified version of a transistor commonly used in various RF power amplifier applications. The specific structure studied is that shown in Fig. 2 with gold electrodes, a 25-nm barrier of $Al_{0.3}Ga_{0.7}N$, a 50-nm passivating layer of SiN, a gate length of 0.3 μm, a source-to-gate spacing of 0.5 μm, and a gate-to-drain spacing of 1.6 μm. Also, for convenience in the simulations, we dope the source and drain contact regions heavily in order to avoid any non-physical contact effects. (Practical devices differ further in including adhesion and diffusion barriers in the metallization, alloyed source/drain contacts, field plates, and packaging.) In the remainder of this section, a variety of simulation results for this device are presented.

A. Normal operating conditions

Before analyzing the GaN HEMT's response to electrical stressing, we begin with a brief account of the thermoelectromechanics of the transistor under normal operating conditions. As seen in Sec. III, such simulations are useful for verification. In addition, they provide a baseline understanding of the electric fields, stresses, strains, and temperatures in a normally operating device for later comparison with stressed devices.

Shown in Figs. 8(a)-8(e) are the simulated electron density, the electric field magnitude, the maximum principal stress, the in-plane strain, and the temperature in the ON-state ($V_{GS} = 0$ and $V_{DS} = 5$). In these plots, Fig. 8(a) shows the reduced-size electrical simulation region, while Fig. 8(e) shows the full thermoelastic simulation region, with the inset again showing where coupling to the electrical variables occurs. All the other plots in Fig. 8 are close-ups of the action occurring in the vicinity of the gate. Qualitatively, all of these plots look as expected, with Fig. 8(a) clearly showing the high electron density induced by the polarization in the access regions as well as the existence of a channel under the gate at zero bias. The maximum electric field of about 6.2 MV/cm, the maximum principal stress of about 3.6 GPa (tensile), the maximum strain of about 0.78%, and the maximum temperature of 128 °C all occur in the AlGaN at the drain-side corner of the gate. While these values are significant, in conformity with ON-state experiments, they are not so large as to make "rapid" device failure likely. For example, the electric field under the gate produces a Fowler-Nordheim injection barrier that is no thinner than ∼3.2 nm, and so, at most, we can expect long-term degradation via slow electron injection, with the temperature and stress possibly being contributing factors.

B. Bias-stress conditions

In evaluating device reliability, it is common practice to employ accelerated life-testing by operating the devices at elevated voltage/current levels and/or baseplate temperatures. Motivated by the intriguing proposal that piezoelectric stresses/strains could be a triggering mechanism for GaN HEMT degradation/failure, in this paper we focus on

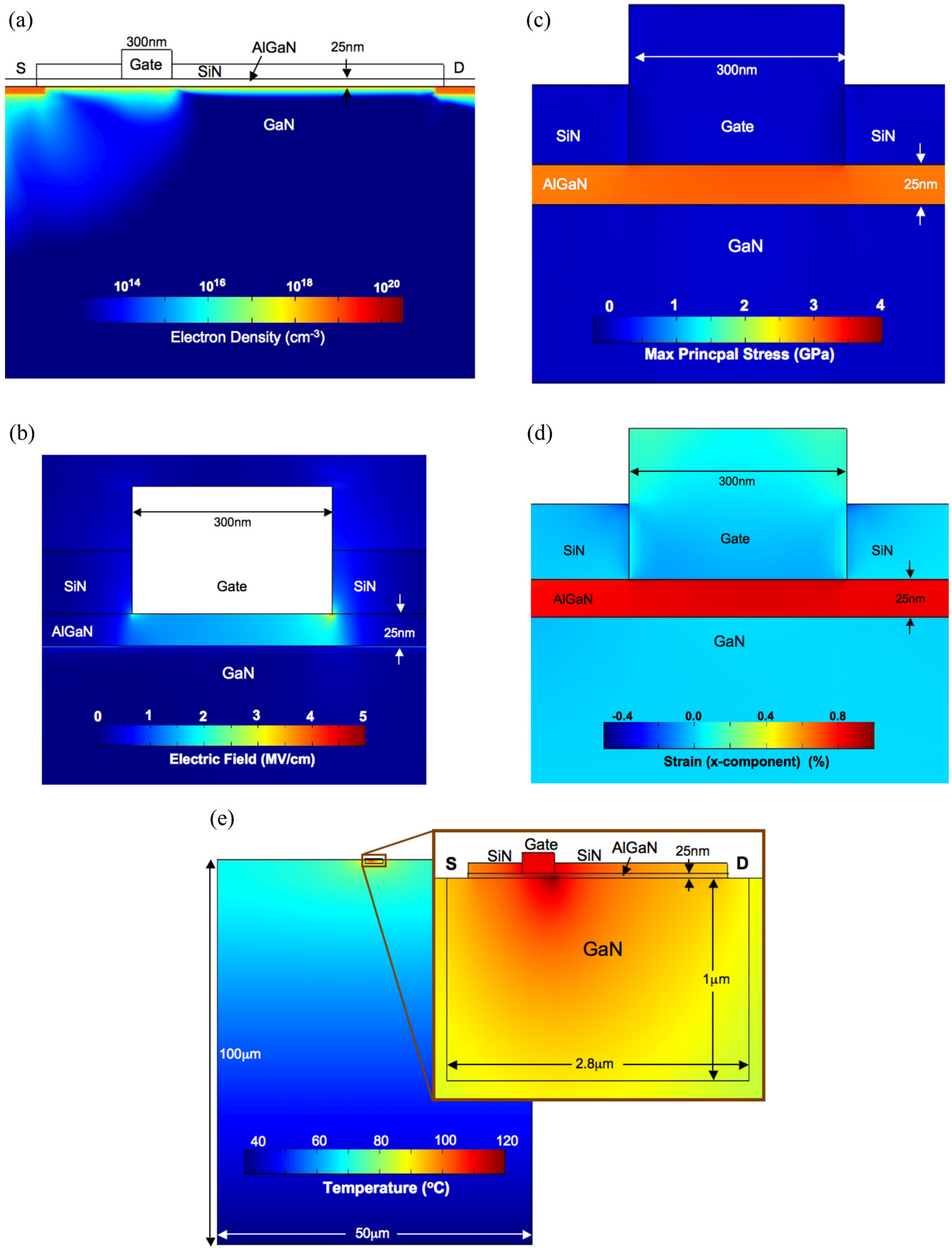

FIG. 8. (a) Electron density (log scale), (b) electric field, (c) maximum principal stress, (d) in-plane strain, and (e) temperature in a GaN HEMT in the ON-state (with $V_{DS} = 5$ V and $V_{GS} = 0$ V). The image in (e) shows the entire simulation region for the thermoelastic variables and, in the inset, the reduced region in which the electrostatic and electron transport variables are coupled.

electrical stressing and consider three basic stress-bias conditions: (i) An isothermal OFF state (with $V_{GS} = -5$ V and $V_{DS} = 15$ V) similar to that studied extensively by del Alamo and co-workers;[6,7] (ii) a dc high-power state with $V_{GS} = 0$ V, $V_{DS} = 20$ V, dc power (P_{DC}) of about 15 W/mm, and the baseplate at room temperature; and (iii) an RF stress state with $V_{GS} = -4.85$ V, $V_{DS} = 30$ V, $T_p = 225\,^{\circ}$C, and with the power dissipated in the device reduced by RF power conversion to about 3.5 W/mm. Of the three conditions, the high-power state is clearly the most extreme and is analyzed in the greatest detail.

Because of the high drain-to-gate bias, the electric field in the OFF-state will be quite high in the vicinity of the gate; indeed, as shown in the logarithmic contour plot in Fig. 9, these fields reach values in excess of 10 MV/cm. To be more quantitative, it is helpful to look at one-dimensional

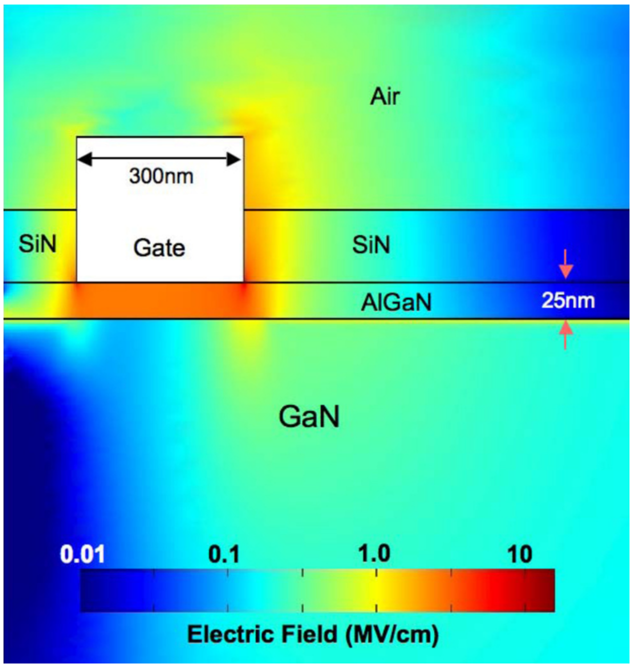

FIG. 9. Logarithmic contour plot of the magnitude of the electric field in the GaN HEMT in the OFF-state ($V_{DS} = 15$ V and $V_{GS} = -5$ V).

cutline profiles. One such plot for the OFF-state situation is presented in Fig. 10, where we show semi-log profiles of the electron energy (with respect to the gate Fermi level), the electric field, the maximum principal stress, and the contribution of the inverse piezoelectric effect to the total stress. The particular cut-line assumed in Fig. 10 is vertically across the AlGaN barrier starting from a point 3 Å just inside the gate corner in order to avoid the corner singularity, as discussed in relation to Fig. 4. The maximum electric field is seen to be about 12 MV/cm, and as the Figure indicates, this produces a triangular barrier about 2-nm wide, that assuredly means that the OFF-state biasing will induce strong electron injection into the AlGaN layer. As seen in the figure, the maximum stress is about 4 GPa, which, although piezoelectrically enhanced by about 0.5 GPa over that seen in the ON-state, is presumed to be still well short of the level needed to trigger direct mechanical failure. But it could easily be that this stress is high enough to augment the rate of trap formation by injected electrons, as suggested in Ref. 6, with the extra 13% provided by the biasing through the inverse piezoelectric effect possibly being significant. Obviously, temperature can play no role in accelerating the OFF state degradation, since the situation is isothermal.

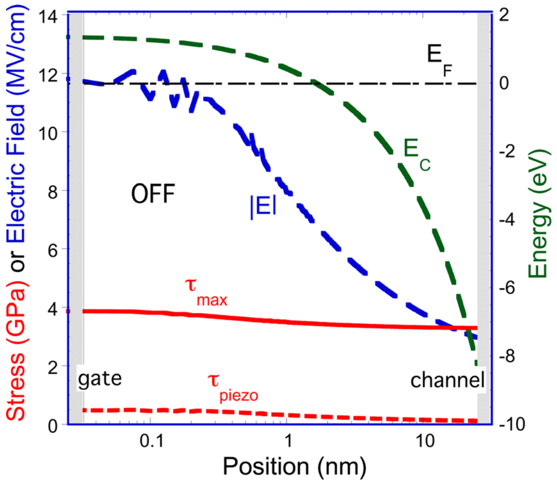

FIG. 10. Simulated profiles in the OFF-state ($V_{DS} = 15$ V and $V_{GS} = -5$ V) across the AlGaN layer on a cutline starting at the drain-side corner of the gate showing the maximum principal stress, the piezoelectric stress, the electric field magnitude, and the conduction band energy relative to the gate.

Next, we consider the more complicated situations in which the HEMT device is stressed with current flow and elevated temperatures. In Fig. 11, we plot the drain current for $V_{GS} = 0$ V as V_{DS} is ramped from zero up to the high-power state with $V_{DS} = 20$ V; also shown in the figure is the maximum channel temperature as generated by the current flow at each bias. The fall-off in the saturation current at high bias is again due to the drop in mobility that accompanies the Joule heating and elevated temperatures via Eq. (2b). At the highest drain voltage, the calculated maximum power dissipated in the device is about 14.9 W/mm, and for a baseplate temperature of 26 °C, this produces a simulated maximum channel temperature of around 426 °C, with a temperature distribution in the device region, as shown in Fig. 12. This maximum is very high, and it is reasonable to suggest that it might itself be sufficient to cause device degradation by activating "chemical" processes, such as impurity diffusion.[30] To better understand the electrical, mechanical, and thermal fields that accompany the current stressing, we next examine one-dimensional cutline profiles, much like those in Fig. 10. In Fig. 13, we show semi-log profiles for the high-power state of the electric field, the maximum principal stress, and the piezoelectric contribution to the total stress along two cutlines, both transecting the AlGaN barrier, with one at the drain corner of the gate as before (again starting at a point 3 Å just inside) and the other at the source corner of the gate. Having the same value of V_{DG} (20 V) as in the OFF-state, it is no surprise (see Fig. 13, left axis) that the maximum electric field is roughly the same, with the peak electric field around 12 MV/cm on the drain side; on the source side, the

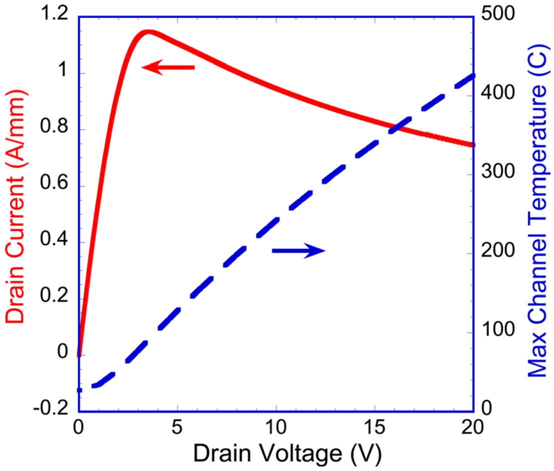

FIG. 11. Simulated drain characteristics and peak temperature in the GaN HEMT as a function of V_{DS} out to 20 V.

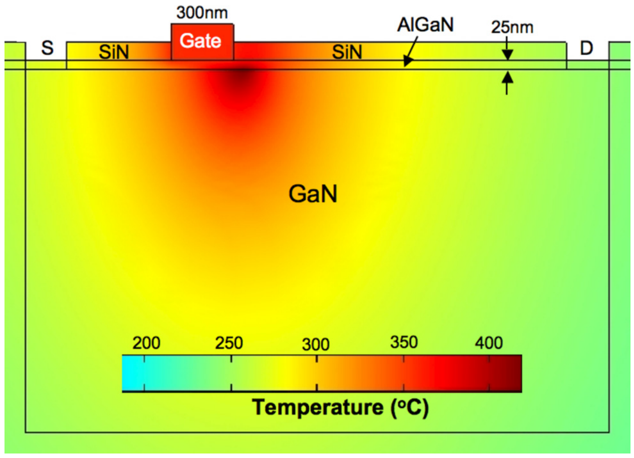

FIG. 12. Simulated temperatures in the GaN HEMT under high-power stressing conditions ($V_{DS} = 20$ V and $V_{GS} = 0$ V).

field is much lower (~3.4 MV/cm). As in Fig. 10, the high field at the drain side will produce a barrier about 2 nm wide, and so we expect strong electron injection there as well. Furthermore, through the inverse piezoelectric effect, the high drain field will induce a larger piezoelectric stress at the drain end of the gate (0.49 GPa) than at the source (0.14 GPa). Interestingly however, Fig. 13 also shows that the difference in the corresponding total stresses at these locations is even larger with 4.7 GPa versus 3.9 GPa. So if stress were the explanation for preferential structural failure at the drain end of the gate, in the case modeled in Fig. 13, the piezoelectric effect is providing only about half of that excess stress. As discussed in more detail below, the other half of the stress differential is thermal in origin. The thermal stresses are also apparent in the rapid rise in the total stress across the AlGaN, an increase that is clearly not explained by the smaller increase in the piezoelectric component (see Fig. 13). Finally, we observe that, even though the total stress level is elevated, it seems insufficient to trigger direct structural pit/crack formation (based on our discussion in Sec. II of the tensile strength of the AlGaN), unless the very high operating temperature and/or the strong electron injection are acting to lower the threshold for material failure.

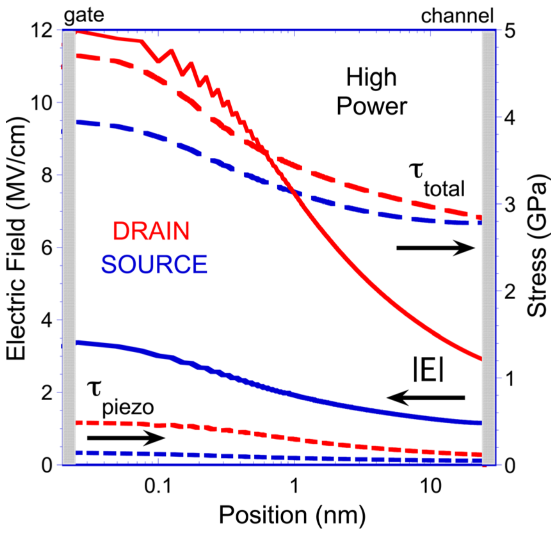

FIG. 13. Simulated profiles in the high-power state ($V_{DS} = 20$ V and $V_{GS} = 0$ V) across the AlGaN layer on cutlines at the source-side and drain-side corners of the gate showing the maximum principal stress, the piezoelectric stress, and the electric field magnitude.

To introduce our approach to quantifying the thermal stresses, it is useful to note that, in the case of the piezoelectric stresses, because the peak electric field is located within the most piezoelectric material in the HEMT structure (namely, the AlGaN), the resulting stress contribution will be generated almost entirely by the local electric field and so is readily calculated. In contrast, thermal expansion acts on all of the materials in the structure and is actually largest in the Au gate metal, because it is both very hot and has a large thermal expansion coefficient. Furthermore, the thermal stress is generated not by the expansion itself, but by *differential* expansion. Hence, a local calculation is not sufficient to estimate the total thermal stress in the AlGaN. (It is for this reason that a proper calculation of the thermal stress requires that the elastic variables be solved for in the full 50×101-μm simulation region of Fig. 8(e) rather than just in the restricted domain of the inset, where the electrical problem is solved). The approach used in this paper to find the total thermal stress is simply to subtract the results of two calculations: one with all thermal expansion terms included and the other with none included. As obtained in this way, the in-plane and vertical components of the normal thermal stress for the high-power state are shown in the contour plots in Figs. 14(a) and 14(b), respectively. Away from the gate region (and the S/D contacts, not shown), the basic effect of the thermal expansion is simply to exert an in-plane compressive stress on all layers, with less stress in the SiN, as a result of its smaller thermal expansion coefficient, and with no vertical component, because the layers are free to expand upward. Thus, away from the gate region (but in regions where the temperature is elevated), the presence of thermal stress moderates the epitaxial stress in the AlGaN; in particular, the average total in-plane stress (not shown) in the AlGaN drops by about 0.45 GPa from ~3.1 GPa to ~2.6 GPa. This also explains the drop in total stress near the channel in Fig. 13 as compared with Fig. 10. As is evident from Figs. 14(a) and 14(b), the thermal stresses in the vicinity of the gate are much more inhomogeneous and complicated, primarily because of the large thermal expansion of the gate metal. This origin, and the fact that the gate is essentially isothermal, means that the thermal stress distribution is fairly symmetric about the midline of the gate (which is, of course, not true of the piezoelectric stress). In the in-plane direction (Fig. 14(a)), the gate tends to stretch the AlGaN and, to some extent, the nearby channel region, thus reducing the compensation of the AlGaN's epitaxial stress and leaving the total stress in the vicinity of the gate at about 3.1 GPa at the midline. In the vertical direction, the main effect of the gate expansion arises from its interaction with the adjacent SiN. The mismatch in thermal expansion coefficients in this case is far larger than with the AlGaN, and so the SiN sidewalls are put strongly in tension (about +1 GPa), while the gate metal sidewalls are correspondingly strongly compressed (by about –1 GPa). Finally, and most importantly for this paper, is the action at the gate corners as emphasized by the insets in

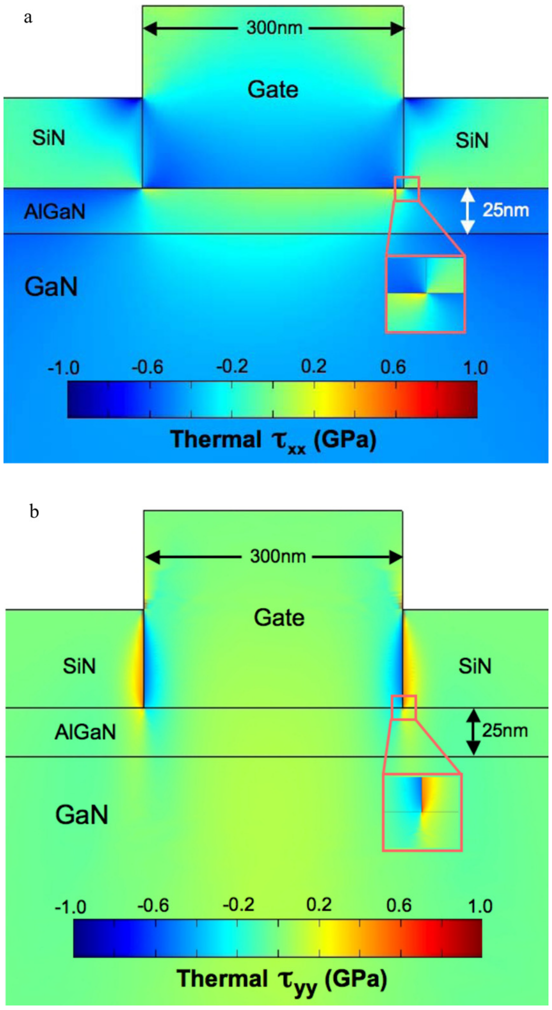

FIG. 14. Simulated (a) in-plane and (b) vertical components of the thermal stress in the GaN HEMT under high-power stressing conditions ($V_{DS} = 20$ V and $V_{GS} = 0$ V). The insets highlight the complicated stress fields in the vicinity of the drain-side corner of the gate.

Figs. 14(a) and 14(b) that magnify the drain corner regions. The thermal stress in these regions is a complex resultant of the in-plane and vertical effects already discussed. The small region of tensile thermal stress at the drain corner visible in both the in-plane (Fig. 14(a), inset) and vertical (Fig. 14(b), inset) components is what accounts for the enhancement of the total stress near the gate under high-power conditions, as seen in Fig. 13. Lastly, we note that, at the drain corners, the maximum shear stress (not shown) is quite high, reaching about 2.5 GPa, of which the thermal component adds about 1 GPa. Whether this plays a role in failure is not known.

The final stressing condition we consider is one that emulates a reliability test commonly applied to RF devices; the damage seen in Fig. 1 was obtained under just such conditions. For the simulation, we again treat the device of Fig. 2, set the dc drain voltage high ($V_{DS} = 30$ V), and choose a gate voltage ($V_{GS} = -4.85$V) so as to limit the dc power to $P_{DC} = 7.5$ W/mm. In an experiment of this type, an RF input is also applied, and the power added efficiency (PAE) of the device results in only some fraction of the dc power being dissipated; we assume a PAE of 50%. The reduced power dissipation means the temperature rise is also limited (~100 °C), and so for purposes of accelerated life-testing, one typically applies heating to the baseplate; for our example, we assume $T_p = 225$ °C. Also, since we perform dc simulations only, we mimic the action of the RF PAE by setting $\eta = 0.5$ in $(1b)_2$. With this assumption, we find $T_{\max} \cong 331\,^{\circ}$C, which corresponds quite well with measured values.

Profiles of the stresses and electron energy (relative to the metal Fermi level) in the device are plotted in Fig. 15 along the usual cutline at the drain-side edge of the gate. When compared with the similar profiles in Figs. 10 and 13, a number of differences are apparent, none of which are surprising. The higher voltage drop from gate to drain ($V_{DG} = 34.8$ V) leads to higher electric fields; indeed, the values are unrealistically high, because our device (Fig. 2) lacks the field plates that would be present in a practical RF device. In any event, in our case, the result is a narrow Fowler-Nordheim barrier of only about 1 nm, and so very strong electron injection from the gate would be expected. The higher electric field also leads to larger piezoelectric stresses (~0.7 GPa). That the temperature excursion is smaller than for the high power case (106 °C versus 400 °C) results in significantly reduced thermal stresses (~ 0.1 GPa). As a result, the overall maximum stress is somewhat smaller than in the high-power case (4.3 GPa). Lastly, although the temperature differences are smaller, having baseplate heating means the peak temperatures are still quite high. Therefore, electron injection and temperature would seem to be the principal factors triggering failure in the RF stressing situation.

C. Effect of gate metal

Although it is not clear from the discussion of Subsection IV B that mechanical stress plays a critical role in GaN HEMT failure, it is still of interest to note that one can influence these stresses through device design. Here, we examine

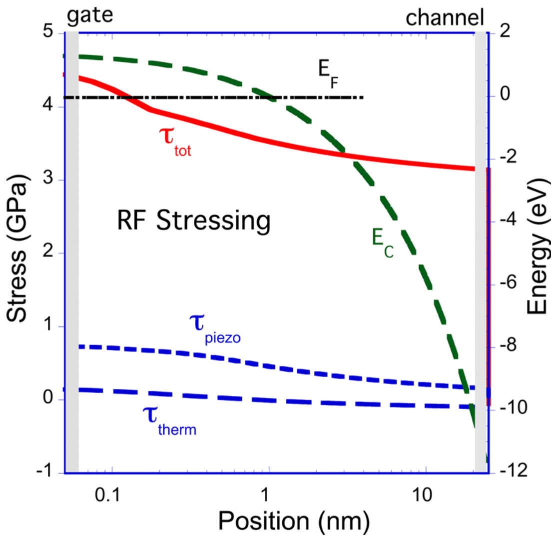

FIG. 15. Simulated profiles in the RF stress state ($V_{DS} = 30$ V, $V_{GS} = -4.85$ V, $T_p = 225$ °C, and RF power conversion efficiency of 50%) across the AlGaN layer on a cutline at the drain-side corner of the gate showing the maximum principal stress, the piezoelectric stress, the thermal stress, and the conduction band energy relative to the gate.

the effect of changing the gate material, considering an alternative tungsten gate for which the thermal expansion coefficient is about three times lower than it is for gold (and with elastic constants roughly doubled), as seen in Table I. A cutline plot across the AlGaN thickness (like that in Fig. 13) is shown in Fig. 16 that contrasts the total maximum principal stress and its piezoelectric and thermal components for Au- and W-gated GaN FETs under high-power stressing conditions. The two cases are quite similar on the channel side of the AlGaN, but become markedly different as the gate interface is approached, with the thermal stress for the Au gate actually reversing in sign. The thermal stress levels (under these high-power conditions) are roughly equal in magnitude to the piezoelectric stress near the gate corner, but in the case of the Au gate, the thermal contribution is additive and so essentially doubles the piezoelectric effect, whereas for the W gate, the two components nearly cancel. Consequently, the total stress near the gate corner of the Au gate is about +1 GPa larger than that of the W gate, a large difference that could be enough to make the tungsten gate device more reliable under accelerated life-testing conditions.

D. Effect of the gate shape

GaN HEMTs sometimes angle the SiN passivation so that, when the gate electrode is evaporated, it forms an overhanging or "slant field" gate.[31] A similar strategy commonly used is to modify the gate shape to include field plates. Generally speaking, the rationale behind all of these gate designs is to spread the electric field in the vicinity of the gate edge, thereby raising the breakdown voltage, reducing the gate current, and mitigating hot-carrier damage. In this paper, we do not take up a full simulation study of the impact of gate shape on the device thermoelectromechanics, but instead merely illustrate the consequence of one specific design, namely, with a slant field gate with a 45° slant. In general, the simulations meet expectation with the electric field at the drain-side gate corner dropping from about 12 MV/cm to 8.5 MV/cm. This reduction should significantly reduce the electron injection from the gate and thereby has the potential to improve reliability both with respect to trapped charge effects and perhaps even crack initiation. With regards to the stress, the lowered electric field does indeed reduce the piezoelectric stress (from 0.5 GPa to about 0.35 GPa), but because the thermal stress (for a gold gate under high-power conditions) remains essentially the same, the total stress at the gate corner improves only a little, dropping from 4.6 GPa to about 4.45 GPa.

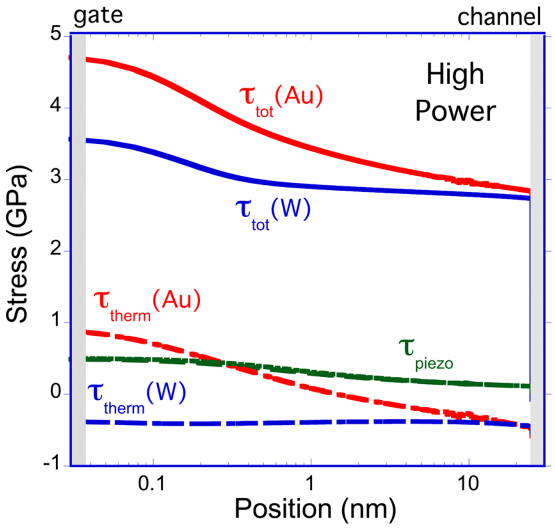

FIG. 16. Simulated profiles in the high-power state ($V_{DS} = 20$ V and $V_{GS} = 0$ V) as in Fig. 13, but comparing the total, piezoelectric, and thermal stresses for devices with gold and tungsten gates. The much lower thermal expansion coefficient of the latter has a significant impact on the peak stress.

E. Effect of SiN intrinsic stress

A SiN layer situated on top of the AlGaN, as depicted in Fig. 2, is a standard feature of GaN HEMTs that is included in order to reduce electron trapping and associated transient "gate-lag" effects.[8] We have already seen that such a SiN passivation layer plays a role in the understanding of thermal stresses in GaN HEMTs. In this section, we consider a second effect that SiN might have on the thermoelectromechanics of GaN HEMTs through the possibility of it incorporating intrinsic stresses. In other technological applications of SiN films, it is well known that by appropriate processing one can build in intrinsic stresses that can be either compressive or tensile, and with magnitudes as high as 3 GPa.[1] In order to assess the role, such pre-stresses might have on GaN HEMT reliability, we again study the high-power state and compare results when the SiN has an intrinsic stress level of either +2 GPa (tensile) or –2 GPa (compressive).

Some results are shown in the semi-log plot in Fig. 17, where, for high-power conditions, we show the profiles across the AlGaN layer of the total stress and the piezoelectric and thermal contributions for the cases of +2 GPa and –2 GPa intrinsic stress in the SiN. Clearly, the pre-stress in the SiN can have significant impact on the AlGaN stress, increasing its value at the drain corner of the gate from 4.6 GPa to 6.2 GPa in the tensile case or decreasing it to 3.3 GPa in the compressive case. As we have discussed elsewhere, it is not clear that even a stress as high as 6.2 GPa will directly fracture a pristine AlGaN layer, but it could

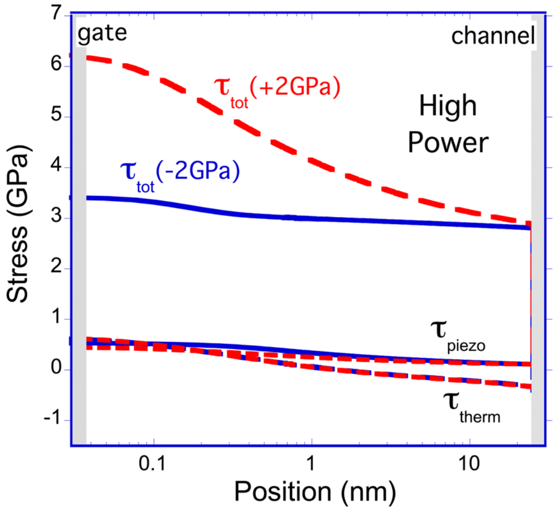

FIG. 17. Identical plot to Fig. 16, but comparing Au-gated GaN HEMTs, in which the SiN passivation layer has an intrinsic stress level of either +2 GPa (tensile) or –2 GPa (compressive).

well do so if the layer is weakened by injection-induced defects and/or by Joule heating.

F. Fracture

When sufficient mechanical stress is exerted, a covalent crystal, like AlGaN, will undergo brittle fracture. This phenomenon is generally analyzed in terms of two sequential steps, first crack initiation and then crack propagation, each of which deserves discussion. Of the two, judging whether and how crack initiation takes place is far more complicated for many reasons already presented. As we have seen, the actual stress levels in the AlGaN do not in themselves seem sufficient to initiate a crack. We therefore have to look to other factors, such as pre-existing material "flaws", weakening due to defect generation or localized heating, and/or sensitivity to electromigration or chemical reaction as possible contributing mechanisms. But whatever the cause, it is an experimental fact that stressing under high-power conditions often results in the formation of pits and/or incipient cracks in the AlGaN barrier layer.[4,5]

Regarding the crack propagation step of fracture, the classical picture is that of Griffith,[32] wherein an initial crack of some critical size will extend itself if the gain in energy from strain relief is greater than the cost of creating the two new surfaces of the crack. As before, we base our approach on comparing the maximum principal stress developed at the crack tip with the tensile strength of the material with the idea that crack propagation will occur so long as the stresses at the tip stay high enough to cause continued rupture of the lattice. For our analysis, we continue to use our 2D simulator, which means we assume the initiating crack to be a "groove" that extends across the width of the device (rather than a localized "pit" or "string of pits", as is seen initially in AFM images like that of Fig. 1 (Ref. 5)) and the crack propagation to be uniform across the device width.

A sample result from a 2D simulation of a "damaged" GaN HEMT that has a 2×2 nm "pit" introduced at the drain-side corner of the gate is shown in Fig. 18. The crack tip is circular, and because its radius of curvature is finite (1 nm), the continuum solution is not singular at its apex. Figure 18 shows the maximum principal stress, and we observe that the peak stress, located at the apex, is now much higher (13 GPa versus 4.7 GPa) than it was with no "pit" present (Fig. 13). The increase in stress is due to the concentrating effect of the "pit" and will be smaller if its radius is expanded. Based on our earlier discussion, this level of stress is likely above the tensile strength of the AlGaN, and so our analysis suggests that this very small crack will propagate. And as the crack deepens, simulation shows the stress continues to rise (to 35 GPa when the crack has almost fully traversed the AlGaN layer), implying that the crack will continue to propagate and will quickly traverse the AlGaN layer, as is seen most clearly in TEM cross-sections.[4]

An important aspect of the simulations just described is a decoupling of the mechanical and electrical fields that is brought about by the crack formation. This can be seen in the simple fact that the peak stress in Fig. 18 is essentially independent of the electrical biasing and thus the action of the "pit" is entirely one of focusing epitaxial strain. Furthermore, simulation shows that the presence of the initiating crack leads to a separation between the point of maximum electric field and maximum stress, which not only means that piezoelectric stress is not being concentrated at the crack tip, but it also implies that electron injection will not play a role in the crack propagation. Thus, while the crack initiation can potentially involve thermal, electrical, mechanical, and chemical factors, the AlGaN crack propagation is a much simpler and purely mechanical phenomenon.

A crucial question regarding the pits and cracks observed in stressed GaN HEMTs is one of causality: Does the mechanical damage *cause* the device's electrical degradation or

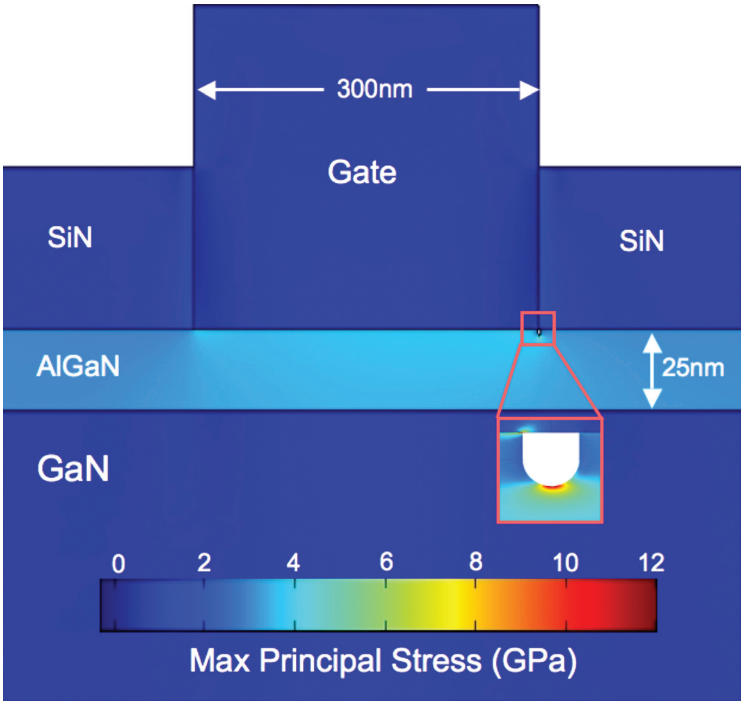

FIG. 18. Simulated maximum principal stresses in a GaN HEMT with a 2 nm × 2 nm "pit" in the AlGaN layer and situated at the drain-side corner of the gate. This "pit" acts to concentrate the epitaxial stress, raising the peak from 4.6 GPa to 13 GPa.

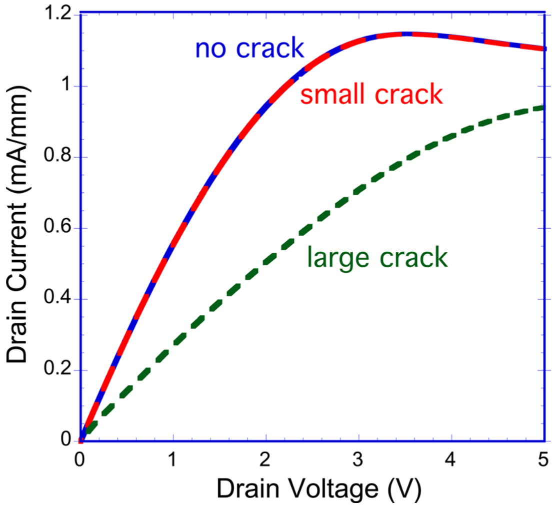

FIG. 19. Simulated drain characteristics of GaN HEMTs, comparing the situations with no damage (also plotted in Fig. 6) with the small "pit" shown in Fig. 18 and with a large crack that extends almost entirely across the AlGaN layer. Significant electrical degradation is seen only in the case of the large crack.

does it merely accompany a separate electrical damage pathway, e.g., via trap creation. Of course, there can be a synergistic combination of mechanical and electrical mechanisms, but as we have already discussed, the fact that the crack moves away from the high field region means that, while they might start together, their later evolution is essentially independent. One aspect of this causality issue that is readily explored with simulation is to ask to what extent the existence of the crack itself affects the device I-V characteristics. For the analysis, we assume the crack to be an open space and, because it is buried, that any polarization charges on its free surface remain unneutralized. Simulated I_{DS}-V_{DS} curves with $V_{GS} = 0$ are plotted in Fig. 19, where we compare the devices with no damage (as also appears in Fig. 6), with a small crack, as in Fig. 18, and with a crack that extends across the AlGaN layer. Clearly, the small crack has little effect on the current, whereas the large crack has substantial impact. This electrical degradation is produced by the polarization fields generated around the large crack that induce an electrostatic barrier in the channel (as may be seen in the electron density plot in Fig. 20) and that, thereby, reduce the current. Overall, the simulation of Fig. 19 demonstrates that *once initiation has occurred*, the experimental data is at least qualitatively consistent with a failure pathway in which the mechanical damage *causes* the electrical degradation.

V. SUMMARY AND FINAL REMARKS

In general, understanding degradation and failure in semiconductor devices like GaN HEMTs is very complicated because of the mix of materials, the non-trivial device geometries, and the many potential mechanisms. Further complicating matters is the fact that one wants not only to know what happened in the degradation/failure, but also to establish causality, i.e., what was the triggering mechanism(s). Given this complexity, mathematical modeling can be no more than one tool in the difficult process of reaching a full understanding. Of the many processes involved in GaN HEMT degradation, the most intractable from a modeling perspective relates to chemical phenomena, such as atomic diffusion that has been observed[30] (though it is unknown whether it is cause or consequence). For this reason, in this work, we have focused on electrical and mechanical degradation pathways and employed efficient continuum methods. In particular, we developed and applied a fully coupled, multi-dimensional thermoelectromechanical model that is appropriate for piezoelectric semiconductors like GaN and its alloys with aluminum. The continuum equations solved were those of linear thermoelectroelasticity, linear heat conduction, and diffusion-drift transport with full coupling assumed within the active region, i.e., all mechanical, electrical, and thermal variables were solved for simultaneously. By treating, analytically, the out-of-plane strain induced by the epitaxy, we showed that the rest of the problem reduces to one of plane-strain so that two-dimensional numerical simulation sufficed. We also discussed the important issue of

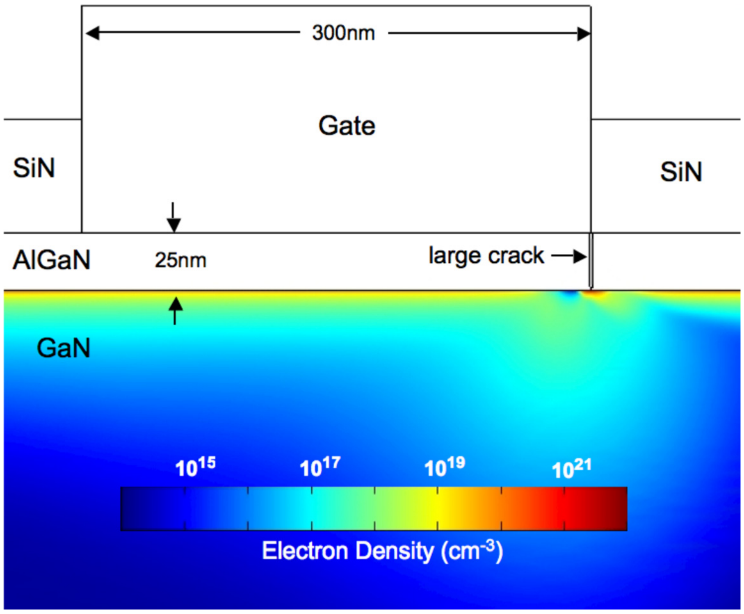

FIG. 20. Simulated electron density in a GaN HEMT with a large crack at the drain-corner of the gate and biased with $V_{DS} = 3$ V and $V_{GS} = 0$ V. The electrostatic perturbation on the channel carriers induced by the polarization charges associated with the crack is what is responsible for the I-V degradation seen in Fig. 19.

mathematical singularities that can appear at corners in the solutions and a simple procedure involving a cut-off distance for handling them. Finally, because of the limitations of the theory with respect to electron trapping and fracture mechanics, we introduced some simple "failure" metrics, namely, the threshold electric field for significant electron injection and the size of the maximal principal stress as compared to the yield strength of the material.

The specific simulations performed in this paper compared the normal ON state of a conventional GaN HEMT with its behavior under the high-bias conditions that are often used for accelerated life-testing. Not surprisingly, under the high-bias conditions, one saw considerable electron injection into the AlGaN, with obvious potential for creating traps and directly degrading the device. The stress/strain fields seen under these conditions were more complicated and include an appreciable piezoelectric contribution. Depending on the gate structure and composition, the thermal stresses induced when the device is ON and at high drain bias can easily exceed the piezoelectric stress. Pre-stress in the SiN can also have a large effect on the AlGaN stress/strain. However, in all cases, the stress levels seemed well below the expected tensile strength at room temperature. As a result, we expect direct mechanical/structural failure to be initiated only if there is synergistic weakening of the material by the injection-produced damage and/or by the Joule heating. By contrast, once failure has begun with the formation of a small "pit", we showed that the stress concentrating effect of the "pit" could readily lead to crack propagation and fracture of the AlGaN in a purely mechanical fashion.

Lastly, this work suggested several avenues for improving the reliability/robustness of GaN HEMTs. These strategies all related to the layers deposited on top of the III-N heterostructure. It is well known that gate shape modifications (including field plates) can be beneficial by spreading the electric field and thereby reducing its peak value. We showed that the gate shape, its material composition, and its abutting layers (e.g., SiN) can also be helpful in reducing the thermal stresses that develop under high-power conditions. One final design modification considered related to SiN pre-stress that, if compressive, can significantly mitigate the stress levels in the AlGaN. Given that it is still uncertain whether mechanical stress plays an important role in GaN HEMT reliability, at the very least, such design changes could provide an experimental path for investigating this basic question.

ACKNOWLEDGMENTS

The authors thank the Office of Naval Research for funding support.

[1] S. E. Thompson, G. Sun, Y. S. Choi and T. Nishida, "Uniaxial-process-induced strained-Si: Extending the CMOS roadmap," IEEE Trans. Electron Devices **53**, 1010 (2006).

[2] B. R. Bennett, M. G. Ancona, J. B. Boos, C. B. Canedy, and S. A. Khan, "Strained GaSb/AlAsSb quantum wells for p-channel field-effect transistors," J. Cryst. Growth **311**, 47 (2008).

[3] O. Ambacher, B. Foutz, J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, A. J. Sierakowski, W. J. Schaff, L. F. Eastman, R. Dimitrov, A. Mitchell, and M. Stutzmann, "Two dimensional electron gases induced by spontaneous and piezoelectric polarization in undoped and doped AlGaN/GaN heterostructures," J. Appl. Phys. **87**, 334 (2000); O. Ambacher *et al.*, J. Appl. Phys. **85**, 3222 (1999).

[4] U. Chowdhury, J. L. Jimenez, C. Lee, E. Beam, P. Saunier, T. Balistreri, S.-Y. Park, T. Lee, J. Wang, M. J. Kim, J. Joh, and J. A. del Alamo, "TEM observation of crack- and pit-shaped defects in electrically degraded GaN HEMTs," IEEE Electron Device Lett. **29**, 1098 (2008).

[5] P. Makaram, J. Joh, J. A. del Alamo, T. Palacios, and C. V. Thompson, "Evolution of structural defects associated with electrical degradation in AlGaN/GaN high electron mobility transistors," Appl. Phys. Lett. **96**, 233509 (2010).

[6] J. Joh and J. A. del Alamo, "Mechanisms for electrical degradation of GaN high-electron mobility transistors," Tech. Dig. – Int. Electron Devices Meet. **1**, 148 (2006).

[7] J. Joh and J. A. del Alamo, "Critical voltage for electrical degradation of GaN high-electron mobility transistors," IEEE Electron Device Lett. **29**, 287 (2008).

[8] S. C. Binari, K. Ikossi, J. A. Roussos, W. Kruppa, D. Park, H. B. Dietrich, D. D. Koleske, A. E. Wickenden, and R. L. Henry, "Trapping effects and microwave power performance in AlGaN/GaN HEMTs," IEEE Trans. Electron Devices **48**, 465 (2001).

[9] C. Rivera and E. Munoz, "The role of electric field-induced strain in the degradation mechanism of AlGaN/GaN high-electron-mobility transistors," Appl. Phys. Lett. **94**, 053501 (2009).

[10] A. Sarua, H. Ji, M. Kuball, M. J. Uren, T. Martin, K. J. Nash, K. P. Hilton, and R. S. Balmer, "Piezoelectric strain in AlGaN/GaN heterostructure field-effect transistors under bias," Appl. Phys. Lett. **88**, 103502 (2006).

[11] B. Jogai, J. D. Albrecht, and E. Pan, "Effect of electromechanical coupling on the strain in AlGaN/GaN heterojunction field effect transistors," J. Appl. Phys. **94**, 3984 (2003).

[12] S. R. Lee, D. D. Koleske, K. C. Cross, J. A. Floro, K. E. Waldrip, A. T. Wise, and S. Mahajan, "In situ measurements of the critical thickness for strain relaxation in AlGaN/GaN heterostructures," Appl. Phys. Lett. **83**, 6164 (2004).

[13] M.-A. Dubois and P. Muralt, "Properties of AlN thin films for piezoelectric transducers and microwave filter applications," Appl. Phys. Lett. **74**, 3032 (1999).

[14] H. Huang, K. J. Winchester, A. Suvorova, B. R. Lawn, Y. Liu, X. Z. Hu, J. M. Dell, and L. Faraone, "Effect of deposition conditions on mechanical properties of low-temperature PECVD silicon nitride films," Mater. Sci. Eng. A **435-436**, 453 (2006).

[15] D. Faurie, P. Renault, E. Le Bourhuis, P. Villain, and F. Badawi, "Measurement of thin film elastic constants by X-ray diffraction," Thin Solid Films **469-470**, 201 (2004).

[16] F. H. Featherston and J. R. Neighbours, "Elastic constants of tantalum, tungsten, and molybdenum," Phys. Rev. **130**, 1324 (1963).

[17] W. Liu and A. Balandin, "Temperature dependence of thermal conductivity of $\mathrm{Al_xGa_{1-x}N}$ thin films measured by the differential 3ω technique," Appl. Phys. Lett. **85**, 5230 (2004).

[18] M. Eickhoff, O. Ambacher, G. Krotz, and M. Stutzmann, "Piezoresistivity of $\mathrm{Al_xGa_{1-x}N}$ layers and $\mathrm{Al_xGa_{1-x}N}$/GaN heterostructures," J. Appl. Phys. **90**, 3383 (2001).

[19] M. Shur, B. Gelmont, and M. A. Khan, "Electron mobility in two-dimensional electron gases in AlGaN/GaN heterostructures and in bulk GaN," J. Electron. Mater. **25**, 777 (1996).

[20] A. Perez-Tomas, "Ohmic contact resistance dependence on temperature for GaN devices," Mater. Sci. Forum **679-680**, 816 (2011).

[21] See www.comsol.com for information about the Comsol software package.

[22] Z. Lin, W. Lu, J. Lee, D. Min, J. S. Flynn, and G. R. Brandes, "Barrier heights of Schottky contacts on strained AlGaN/GaN heterostructures: Determination and effect of metal work functions," Appl. Phys. Lett. **87**, 4364 (2003).

[23] J. E. Northrup and J. Neugebauer, "Theory of GaN($10\bar{1}0$) and ($11\bar{2}0$) surfaces," Phys. Rev. B **53**, R10477 (1996).

[24] R. Nowak, M. Pessa, M. Suganuma, M. Leszczynski, I. Grzegory, S. Porowski, and F. Yoshida, "Elastic and plastic properties of GaN determined by nano-indentation of bulk crystal," Appl. Phys. Lett. **75**, 2070 (1999).

[25] J. J. Brown, A. I. Baca, K. A. Berkness, D. A. Dikin, R. S. Ruoff, and V. M. Bright, "Tensile measurements of single crystal GaN nanowires on MEMS test stages," Sens. Actuators A **166**, 177 (2011).

[26] Y. Cao and D. Jena, "High-mobility window for two-dimensional electron gases at ultrathin AlN/GaN heterojunctions," Appl. Phys. Lett. **90**, 182112 (2007).

[27] I. Yonenaga and K. Motoki, "Yield strength and dislocation mobility in plastically deformed bulk single-crystal GaN," J. Appl. Phys. **90**, 6539 (2001).
[28] C. L. Muhlstein, S. B. Brown, and R. O. Ritchie, "High-cycle fatigue of single-crystal silicon thin films," J. Microelectromech. Syst. **10**, 593 (2001).
[29] T. Beechem, A. Christensen, S. Graham, and D. Green, "Micro-Raman thermometry in the presence of complex stresses in GaN devices," J. Appl. Phys. **103**, 124501 (2008).
[30] M. Tapajna, U. K. Mishra, and M. Kuball, "Importance of impurity diffusion for early stage degradation in AlGaN/GaN high electron mobility transistors upon electrical stress," Appl. Phys. Lett. **97**, 023503 (2011).
[31] Y. Pei, Z. Chen, D. Brown, S. Keller, S. P. Denbaars, and U. K. Mishra, "Deep-submicrometer AlGaN/GaN HEMTs with slant field plates," IEEE Electron Device Lett. **30**, 328 (2009).
[32] A. A. Griffith, Philos. Trans. R. Soc. London, Ser. A **221**, 163 (1921).