Polarization-mediated remote surface roughness scattering in ultrathin barrier GaN high-electron mobility transistors

Yu Cao,[a] Huili Xing, and Debdeep Jena
Department of Electrical Engineering, University of Notre Dame, Indiana 46556, USA

(Received 26 October 2010; accepted 5 November 2010; published online 2 December 2010)

Ultrathin AlN/GaN heterojunctions are highly attractive for high-frequency transistor applications. In this work, remote surface roughness (RSR) scattering mediated by the high polarization is studied as a new scattering mechanism in such structures. In both depletion-mode and enhancement-mode devices with ultrathin AlN barriers, RSR scattering can be of the same order as polar optical phonons, the dominant scattering mechanism at room temperature. The study indicates that to achieve high-performance high-electron mobility transistors, the surface roughness during processing is as critical as maintaining a sharp heterojunction. © *2010 American Institute of Physics*. [doi:10.1063/1.3521258]

III-V nitride-based high-electron mobility transistor (HEMT) technology has made rapid progress over the last decade. For future high-speed device applications, both lateral and vertical scalings are necessary. In traditional AlGaN/GaN HEMTs, when the AlGaN barrier thickness is reduced, the 2DEG density decreases, and the access resistance increases, which degrades high-frequency performance. AlN/GaN heterojunctions provide extremely high-electron densities at the heterointerface covered by only 3–4 nm AlN barrier,[1] making them attractive for high-frequency applications.

Transport properties of the two dimensional electron gases (2DEGs) at such heterojunctions have been studied extensively to understand the performance limits of nitride devices. In this paper, a new scattering mechanism caused by remote surface roughness (RSR) is discussed. A similar concept has been studied for silicon metal-oxide-semiconductor field-effect transistors (MOSFETs),[2,3] where the variation of the oxide thickness causes the degradation of electron mobility in the channel under gate bias. In AlN/GaN HEMTs, AlN is an epitaxial gate dielectric. RSR scattering originates from the polarization in the barrier material and exists even when no gate bias is applied. The origin of such scattering is due to a rough surface coupled with polarization. Surface roughness is common in epitaxy; this scattering mechanism is valid for 2DEGs generated in all polar semiconductor heterostructures including II-VI oxide ZnO/MgZnO (Ref. 4) and especially perovskite oxides heterostructures[5,6] with very high polarization. Al(Ga)N/GaN HEMTs are a specific case treated in this work. Owing to the very large bandgap, AlN barriers allow a high degree of vertical scaling than silicon MOSFETs that incorporate high-K dielectrics; thus this scattering mechanism is especially severe for ultrathin AlN barrier HEMTs.

In polar semiconductor heterostructures, the 2DEG density is a function of the barrier thickness. A Schottky contact with a gate metal makes the surface of the barrier an equipotential plane. However, thickness variations in the barrier result in variations of the electric potential in the channel plane; this random potential variation is the origin of the scattering mechanism. To understand the effect of RSR scattering on device performance, it is necessary to study RSR scattering in the channel region under the gate. In depletion-mode (D-mode) HEMTs, the AlN barrier thickness exceeds the minimum thickness for 2DEG formation at zero gate bias. Due to remote surface roughness, the 2DEG density is not uniform along the channel. This will lead to different threshold voltages V_{th} across large wafers. Enhancement-mode
(E-mode) HEMTs are desired for high power switches and high temperature integrated circuits. Gate recessing is one of the most commonly used approaches to fabricate E-mode HEMTs.[7,8] An ultrathin AlN interlayer is widely used in AlGaN or InAlN HEMTs to remove alloy scattering. With gate recessing, dry etching techniques have been developed to selectively remove the top barrier, such as InAlN, and the gate metal is directly deposited on the ultrathin AlN interlayer.[9] The RSR scattering model derived in this paper is especially important in such E-mode HEMTs with recessed AlGaN or InAlN barriers.

Since the scattering originates from the thickness variation of the AlN barrier, it is necessary to find the functional dependence of the 2DEG density n_s in the gated channel on the barrier thickness t and the gate bias V_g. To do so, from the energy band diagram of an AlN/GaN heterojunction (Fig. 1), the electric field in the barrier can be expressed as

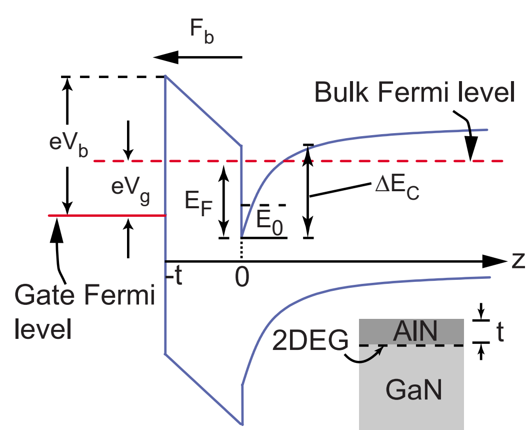

FIG. 1. (Color online) Band diagram of the AlN/GaN heterojunction calculated by 1D Poisson–Schrödinger solver.

[a]Electronic mail: ycao1@nd.edu.

$$F_b = -\frac{e(\sigma_\pi - n_s)}{\epsilon_0\epsilon_b} = -\frac{e(V_b - V_g) + E_F - \Delta E_c}{et}, \quad (1)$$

where e is the electron charge, σ_π is the polarization sheet charge at the AlN/GaN heterojunction, n_s is the 2DEG density in the well, eV_b is the surface barrier height, ΔE_c is the conduction band offset between the barrier and the well,[10] ϵ_b is the dielectric constant of the barrier material, t is the barrier thickness, and E_F is the Fermi energy shown in Fig. 1. The electron wave function in the growth direction z is well approximated by the Fang–Howard variational form,[11] $\phi_z(z) = \sqrt{b^3/2}ze^{-bz/2}$, where b is given by $b = (33m_w^* e^2 n_s / 8\hbar^2\epsilon_0\epsilon_b)^{1/3}$. The ground-state energy E_0 is

$$E_0 = \left[\frac{5}{16}\left(\frac{33}{2}\right)^{2/3}\right]\left[\frac{\hbar^2}{2m_w^*}\left(\frac{e^2 n_s}{\epsilon_0\epsilon_b}\right)^2\right]^{1/3}, \quad (2)$$

where F_w is the electric field in the well, ϵ_w is the dielectric constant of the well material, and m_w^* is the electron effective mass in the well. If the electron density in the second subband is negligible compared to that in the first subband (which is the case in AlN/GaN heterojunctions with very thin AlN barriers), the electron density can be approximated by the electric quantum limit

$$n_s = \frac{m_w^*}{\pi\hbar^2}(E_F - E_0). \quad (3)$$

Multiple subband occupation should be considered at high gate overdrives when the channel electron density is large, but in this work, we restrict the treatment to the electric quantum limit. This is valid when the electron density in the channel is low and the channel is close to pinch-off; indeed the speed and transconductance of GaN HEMTs peak near this bias point. Combining Eqs. (2) and (3), we can write E_F as a function of n_s

$$E_F = \frac{\pi\hbar^2}{m_w^*}n_s + 2.025\left[\frac{\hbar^2}{2m_w^*}\left(\frac{e^2 n_s}{\epsilon_0\epsilon_b}\right)^2\right]^{1/3}. \quad (4)$$

Substituting Eq. (4) into Eq. (1), we obtain the required relation between n_s, t, and V_g,

$$\frac{e^2 t(\sigma_\pi - n_s)}{\epsilon_0\epsilon_b} = e(V_b - V_g) - \Delta E_c + \frac{\pi\hbar^2}{m_w^*}n_s + 2.025\left[\frac{\hbar^2}{2m_w^*}\left(\frac{e^2}{\epsilon_0\epsilon_b}\right)^2\right]^{1/3} n_s^{2/3}. \quad (5)$$

The equation is cubic in n_s. For a given V_g, the 2DEG density n_s can be either analytically or numerically solved as a function of the barrier thickness t. Figure 2(a) shows the calculated 2DEG density n_s as a function of the gate bias V_g. A fixed surface barrier height $eV_b = 3$ eV (Ref. 1) and AlN barrier thicknesses of $t = 2$, 3, and 4 nm are used for the calculation. $t = 1$ nm is used for the calculation in E-mode devices later. The different slopes in Fig. 2(a) are indicative of the different capacitances.

The electron volume density along the z direction is $\rho(z) = -en_s\phi_z^2 = -en_s(b^3/2)z^2 e^{-bz}$. Identifying the heterointerface as $z = 0$, the electric field $F_w(z)$ in the well $\rightarrow 0$ when $z \rightarrow \infty$. The electric field variation $F_w(z)$ in the triangular quantum well is thus

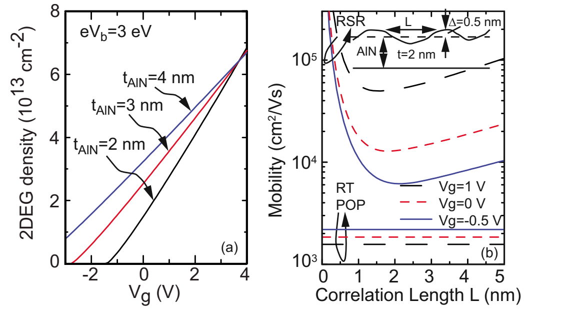

FIG. 2. (Color online) (a) The 2DEG density n_s is plotted against the gate bias V_g with different AlN barrier thicknesses. (b) Mobilities given by RSR scattering in the D-mode HEMTs with t_{AlN} of $t = 2$ nm under different gate biases of 1 V (long dashed lines), 0 V (short dashed lines), and -0.5 V (solid lines) are plotted against different correlation lengths L assuming that Gaussian correlation is applied with screening effect considered. Δ is 0.5 nm (~2 MLs). Mobilities limited by POP scattering at RT with corresponding V_g are also plotted in the dashed lines.

$$F_w(z) = \int_\infty^z \frac{\rho(z)}{\epsilon_0\epsilon_w}dz = \frac{en_s}{\epsilon_0\epsilon_w}e^{-bz}\left(\frac{1}{2}b^2z^2 + bz + 1\right). \quad (6)$$

The potential energy of electrons in the well is then

$$V_z(z) = e(V_b - V_g) + eF_b t - \Delta E_c + \int_0^z eF_w(z)dz. \quad (7)$$

The barrier thickness varies gradually in the $r = (x, y)$ plane and is modeled as $t(r) = t + \Delta t(r)$, with the average thickness $\langle t(r)\rangle = t$. A Gaussian distribution is assumed for the barrier thickness "roughness" with a correlation function given by $\langle \Delta t(r)\Delta t(r')\rangle = \Delta^2 e^{-(r-r')^2/L^2}$, where Δ is the roughness amplitude and L is the correlation length. The first-order in-plane variation of the channel potential energy $\Delta V(r,z)$, which constitutes the scattering potential for RSR scattering is

$$\Delta V(r,z) = \left(\frac{\partial V_z}{\partial t} + \frac{\partial V_z}{\partial n_s}\frac{\mathrm{d}n_s}{\mathrm{d}t} + \frac{\partial V_z}{\partial b}\frac{\mathrm{d}b}{\mathrm{d}n_s}\frac{\mathrm{d}n_s}{\mathrm{d}t}\right)\Delta t(r) \equiv -eF_{\mathrm{eff}}\Delta t(r), \quad (8)$$

where $V_z(z)$ is the channel potential derived in Eq. (7), F_{eff} is a polarization-dependent effective electric field given by $F_{\mathrm{eff}} = P_0 + P_1 e^{-bz} + P_2 e^{-bz}z + P_3 e^{-bz}z^3$, where $P_0 = [e(\sigma_\pi - n_s)]/(\epsilon_0\epsilon_b) - (\mathrm{d}n_s/\mathrm{d}t)[(et/\epsilon_0\epsilon_b) + (2e/\epsilon_0\epsilon_w b)]$, $P_2 = (\mathrm{d}n_s/\mathrm{d}t) \times (e/\epsilon_0\epsilon_w)$, $P_1 = P_2(2/b)$, $P_3 = -P_2(b^2/6)$, and dn_s/dt is obtained from Eq. (5).

Thus, the matrix element for scattering of the 2DEG electrons from state $|k\rangle \rightarrow |k'\rangle$ by the RSR potential is

$$H_{k',k} = \int \phi_{k'}^*(r,z)\Delta V(r,z)\phi_k(r,z)d^2r dz \equiv e\sqrt{\frac{\pi}{\Omega}}\Delta L \times \exp\left(-\frac{q^2L^2}{8}\right)F_{\mathrm{avg}}, \quad (9)$$

where Ω is the area in x-y plane and $F_{\mathrm{avg}} = P_0 + (P_1/8) + (3P_2/16b) + (15P_3/16b^3)$. The total scattering rate due to remote roughness scattering is given by

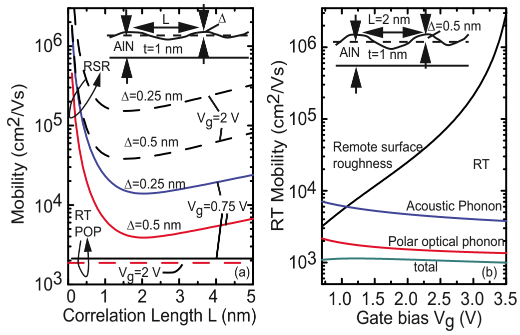

FIG. 3. (Color online) In E-mode devices, assuming that Gaussian correlation is applied with screening effect considered, mobilities given by RSR scattering are plotted (a) against different correlation lengths L under the gate bias V_g=0.75 V (solid lines) and V_g=2 V (dashed lines) with Δ=0.5 nm (~2 MLs) and Δ=0.25 nm (~1 ML). Mobilities limited by POP scattering at RT with corresponding V_g are also plotted. (b) Mobilities given by RSR scattering are plotted against the gate bias V_g with L=2 nm, Δ=0.5 nm, and t=1 nm, where the mobilities limited by POP and AP scattering at RT are also plotted, as well as the total mobility calculated with Matheissen's rule applied.

$$\frac{1}{\tau_{\mathrm{RSR}}} = \frac{\pi e^2 m_w^*}{\hbar^3}\Delta^2 L^2 F_{\mathrm{avg}}^2 e^{-Z}[I_0(Z) - I_1(Z)], \quad (10)$$

where $Z=\pi n_s L^2$, $I_0(Z)$, and $I_1(Z)$ are modified Bessel functions of the first kind and are the solutions of $Z^2(\mathrm{d}^2 f/\mathrm{d}Z^2)+Z(\mathrm{d}f/\mathrm{d}Z)-(Z^2+\nu^2)f=0$, with ν equal to 0 and 1 respectively. We note that the RSR scattering rate increases as the square of the roughness Δ, the correlation length L, and the polarization charge σ_π.

Figure 2(b) shows the calculated electron mobilities limited by RSR scattering in a D-mode AlN HEMT with a barrier thickness $t_{\mathrm{AlN}}\sim 2$ nm. At a negative gate bias of V_g =−0.5 V, the 2DEG density in the channel is $n_s \sim 9 \times 10^{12}/\mathrm{cm}^2$. The RSR scattering rate is found to be of the same order of magnitude (2–3×) as polar optical phonon (POP) scattering, which is the strongest RT scattering mechanism. This result indicates that under negative gate bias, RSR scattering can degrade the 2DEG mobilities in D-Mode HEMTs with thin AlN barriers. Thus, either a smoother surface or a thicker barrier is preferred in D-mode devices for the uniformity of 2DEG transport properties and high mobilities.

In recently reported E-mode InAlN HEMTs,[9] the AlN barrier is only t=1 nm thick under the gate. With such barrier, mobilities limited by RSR scattering under different gate biases V_g (0.75 V in solid lines and 2.0 V in dashed lines) are plotted in Fig. 3(a), together with mobilities limited by POP scattering. With an increasing gate overdrive, the channel charge density increases, leading to a smaller electric field in the AlN barrier. Thus, the potential fluctuation due to the remote surface roughness is less for the same Δ. Therefore, the 2DEG experiences less RSR scattering at high V_g. POP scattering rate increases with higher V_g due to the increased n_s. To study the effect of the surface roughness Δ on RSR scattering, Δ=0.25 nm (~1 MLs) and 0.5 nm (~2 MLs) are selected. When the surface is rough (Δ=0.5 nm, V_g =0.75 V, and L=2 nm), the mobility is ~2000 cm²/V s limited by POP scattering and ~4000 cm²/V s limited by RSR scattering. The total mobility can be degraded from 2000 to ~1300 cm²/V s due to the presence of RSR scattering. If smaller roughness is achieved, e.g., Δ=0.25 nm (~1 ML), RSR scattering can be suppressed and POP scattering is still the major scattering mechanism that limits 2DEGs mobility in the channel. This indicates that for E-mode HEMTs using gate recessing technique, a low-damage etching process is critical for achieving high channel mobilities.

Figure 3(b) shows the V_g dependence of the mobility in E-mode HEMTs with t=1 nm, Δ=2 MLs, and L=2 nm. RSR scattering is weak at high V_g. However, in the range V_g<1 V, which is close to the threshold voltage of most E-mode devices, RSR scattering is more severe than acoustic phonon (AP) scattering. At V_g=0.7 V, the mobility limited by RSR scattering is close to that limited by POP scattering, which is the dominant scattering mechanism in nitride HEMTs at RT. Due to the presence of RSR scattering, the total mobility reduces from ~2000 cm²/V s given by POP scattering to ~1000 cm²/V s, indicating a substantial degradation.

In summary, the combined effect of high polarization and ultrathin barriers can lead to strong degradation of mobility in 2DEGs in the presence of surface roughness. In highly polar heterostructures with thin barriers, this scattering is severe and comparable to optical phonon scattering at RT. The study of RSR scattering in ultrascaled E-mode GaN HEMTs indicates that a uniform barrier thickness is highly desired to avoid the degradation of 2DEG mobilities. The model also applies to 2DEG transport study in other polar heterostructures. To minimize the influence of RSR scattering in E-mode HEMTs, the electric field in the barrier F_b has to be reduced. A thin layer of dielectric between gate metal and recessed AlN barrier can effectively reduce the F_b as well as gate leakage but with a penalty of a lower transconductance. Since the highest transconductance g_m and cut-off frequency f_t of GaN HEMTs occurs near pinch-off voltage, RSR scattering will assume an increasing important role in future design of high-speed GaN HEMTs. Improved epitaxy and processing techniques that preserve a smooth surface and a uniform barrier thickness will be critical.

[1] Y. Cao and D. Jena, Appl. Phys. Lett. **90**, 182112 (2007).
[2] J. Li and T.-P. Ma, J. Appl. Phys. **62**, 4212 (1987).
[3] S. Saito, K. Torii, Y. Shimamoto, S. Tsujikawa, H. Hamamura, O. Tonomura, T. Mine, D. Hisamoto, T. Onai, J. Yugami, M. Hiratani, and S. Kimura, Appl. Phys. Lett. **84**, 1395 (2004).
[4] A. Tsukazaki, A. Ohtomo, T. Kita, Y. Ohno, H. Ohno, and M. Kawasaki, Science **315**, 1388 (2007).
[5] A. Ohtomo, D. Muller, J. Grazul, and H. Hwang, Nature (London) **419**, 378 (2002).
[6] N. Reyren, S. Thiel, A. Caviglia, L. Kourkoutis, G. Hammerl, C. Richter, C. Schneider, T. Kopp, A. Ruetschi, D. Jaccard, M. Gabay, D. Muller, J. Triscone, and J. Mannhart, Science **317**, 1196 (2007).
[7] J.-S. Moon, S. Wu, D. Wong, I. Milosavljevic, A. Conway, P. Hashimoto, M. Hu, M. Antcliffe, and M. Micovic, IEEE Electron Device Lett. **26**, 348 (2005).
[8] W. Lanford, T. Tanaka, Y. Otoki, and I. Adesida, Electron. Lett. **41**, 449 (2005).
[9] R. Wang, P. Saunier, X. Xing, C. Lian, X. Gao, S. Guo, G. Snider, P. Fay, D. Jena, and H. Xing, IEEE Electron Device Lett. **31**, 1383 (2010).
[10] B. Ridley, W. Schaff, and L. Eastman, J. Appl. Phys. **94**, 3972 (2003).
[11] J. Davies, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, Cambridge, UK, 1997).