As GaN HEMTs are scaled down to push performance into 100’s of GHz range, it is timely to investigate their performance limits. Unlike Si MOSFETs and most other III-V semiconductor based HEMTs, the electron - polar optical phonon interaction is exceptionally strong in GaN. As a result, the mean free path of hot electrons in GaN is $\lambda_{np} \sim 3.5\text{nm}$, far shorter than typical HEMT gate lengths ($L_g$). Thus while Si MOSFETs and other III-V HEMTs can approach near ballistic behavior by reduction of parasitic delays and $L_g$, the situation is starkly different for GaN HEMTs. Here, we investigate the intrinsic performance limits of GaN HEMTs by incorporating the effect of polar optical phonon backscattering into a quasi-ballistic model. Then, we include parasitic elements and quantitatively investigate the degradation in performance. The method used is semi-analytical, and will prove very helpful in designing future generations of devices. The work not only sets a roadmap for scaling to high speeds, it also offers clear physical reasons for a number of unexplained features observed in state-of-the-art GaN HEMTs.

The device structure in Fig.1(a) shows the carrier injection from the source-injection point. The carrier density at this point is controlled by the gate voltage $V_{gs}$. Fig. 1(b) depicts the k-space electron distributions at the top of the barrier at 0 K. At low carrier concentrations (near pinchoff), electrons at the injection point cannot emit optical phonons since maximum electron kinetic energy is less than optical phonon energy $\hbar\omega_o$. Even if they absorb phonons down the channel, they cannot return to the source injection point at high fields. A critical density when the Fermi energy of right-going carriers becomes equal to $\hbar\omega_o$ marks the onset of optical phonon emission. Above this density, hot electrons are scattered back into source, and the distribution function now includes left-going carriers as in Fig 1(b). This picture repeats for the second (and higher) subbands if the carrier density increases further. By assuming the Fermi level of injection electrons is exactly one phonon energy ($\hbar\omega_o$) higher than back-scattered electrons, the ensemble injection velocity $v_{inj}$ is calculated from the k-space distribution, and shown in Fig. 2(a) for 0K & 300 K. The injection velocity $v_{inj}$ increases until the optical phonon emission occurs at an injection point 2DEG density-$4\times10^{12}/\text{cm}^2$, at which point it reaches $1.6\times10^7\text{cm/s}$ at 0K, but degrades to $1.3 \times 10^7\text{cm/s}$ at 300K. This peak velocity point also leads to peak $g_m$ and $f_T$. The injection velocity curve has a small bump at the onset of injection from the second subband ($-2.5\times10^{13}/\text{cm}^3$), which is clear at 0K, but smoothened out at 300K. Fig 2(a) is the central result of phonon-limited high-field transport properties in GaN HEMTs, and all device characteristics (drain current, transconductance, and speed) of the HEMT can be derived from this result by combining it with device parasitics. The corresponding saturation current $J_d = enu_{inj}$ is shown in Fig 2(b). In the following calculations, we ignore the second subband.

Using the model, we calculate the device characteristics of a prototype 5 nm 35% AlGaN/GaN HEMT. We ignore short channel effects, but consider parasitic elements. The carrier density $n$ and gate capacitance $C_{gs}$ are calculated self-consistently using Schrodinger-Possion equations, and shown in Fig. 2(c). The carrier concentration at the injection point decreases with the source resistance $R_S$ due to the voltage drop across $R_S$. The intrinsic drain current $J_d = enu_{inj}$ and intrinsic transconductance $g_m = \partial J_d/\partial V_{gs}$ at different $V_{gs}$ are calculated and shown in Fig. 3(a). Note the sharp drop in $g_m$ beyond the peak, and the sublinear increase of the drain current with gate bias. Both these experimentally observed features in GaN HEMTs remain unexplained; our model uncovers the physics behind this behavior. The phonon model predicts a second peak in the $g_m$-vs-$V_{gs}$ curve at the onset of occupation of the second subband (not shown here).

The device characteristics and the effect of parasitic resistances (contact+access, S+D) are shown in Fig. 3(a). The sharply peaked ‘intrinsinc’ curves are smoothed by the parasitic resistances. Fig. 3(b) shows the cutoff frequency $f_T$ with $L_g=50\text{nm}$ vs $V_{gs}$. The trend mimics the $g_m$ curve. The strong degradation of the peak $f_T$ with parasitic resistances is evident. The scaling of the peak $f_T$ with $L_g$ is shown in Fig. 3(c). The intrinsic peak $f_T$ ($C_{gd}=0$) curve is the straight-line $f_TL_g^{-23}\text{GHz} \cdot \mu\text{m}$. $C_{gd}$ and $R_S$ (and $R_P$) further decrease the peak $f_T$. Some recent experimental reports of record high performance GaN HEMTs are plotted in the figure [1-4]. The performances of these devices are rather close to the limits based on the model. Our model clearly shows the intrinsic limit of $f_T$ that we can achieve at the respective gate lengths. The nature of ultrafast optical phonon emission forces a high degree of vertical scaling to speed up GaN HEMTs. In addition, the model quantitatively shows the significant role played by parasitic resistances and capacitances in limiting the $f_T$ for short gate length (<50nm) GaN HEMTs.

In summary, a phonon emission model is able to explain the DC and RF behavior of ultrafast GaN HEMTs. A major departure from earlier models is the dependence of the injection velocity on the 2DEG density and on optical phonons in GaN. It is this injection velocity that determines both DC and RF characteristics, and in combination with parasitic elements is successful in explaining experimental data. The model should prove valuable for further improvement of GaN transistors.

Figure 1: (a) Device geometry of a GaN HEMT and band-diagram under bias. (b) Bandstructure and distribution functions $f(k)$ at source injection point. The gate voltage $V_{gs}$ controls the total carrier concentration at the source injection point, but at high field, phonon emission determines the shape of $f(k)$. Wavy arrows show the optical phonon emission by hot electrons.

Figure 2: (a) Carrier ensemble saturation velocity as a function of 2DEG density based on the phonon model at two temperatures. (b) Current density as a function of 2DEG carrier concentration at the source-injection point. (c) Gate capacitance and carrier concentration versus gate voltages at 300K, evaluated by self-consistent solution of Schrodinger-Poisson equations.

Figure 3: (a) The net drain current and transconductance versus $V_{gs}$. The peak in $g_m$ occurs at the onset of optical phonon emission at the source injection point. The degradation of transconductance and saturation current with parasitic access+contact resistances are also shown. (b) The corresponding $f_t$ versus $V_{gs}$ plot follows the transconductance behavior, with the peak $f_t$ near the onset of phonon emission. Parasitic resistances degrade the peak $f_t$ severely. (c) The peak $f_t$ is plotted against the gate length $L_g$. Interestingly, the intrinsic limit on the peak $f_t$ follows the locus $f_tL_g=23\text{GHz} \cdot \mu\text{m}$. The severe degradation with parasitics is also evident, and experimental record values follow the predicted trend.