Ultrafast Hot Carrier Dynamics in GaN and Its Impact on the Efficiency Droop
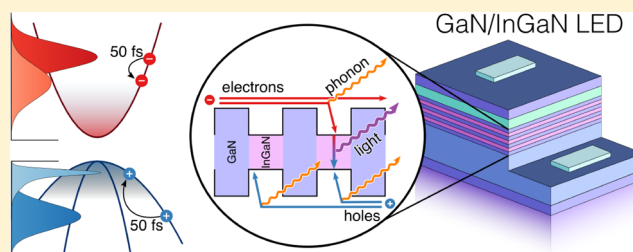
Vatsal A. Jhalani, Jin-Jian Zhou, and Marco Bernardi*

Department of Applied Physics and Materials Science, Steele Laboratory, California Institute of Technology, Pasadena, California 91125, United States

Ⓢ *Supporting Information*

ABSTRACT: GaN is a key material for lighting technology. Yet, the carrier transport and ultrafast dynamics that are central in GaN light-emitting devices are not completely understood. We present first-principles calculations of carrier dynamics in GaN, focusing on electron−phonon (e-ph) scattering and the cooling and nanoscale dynamics of hot carriers. We find that e-ph scattering is significantly faster for holes compared to electrons and that for hot carriers with an initial 0.5−1 eV excess energy, holes take a significantly shorter time (∼0.1 ps) to relax to the band edge compared to electrons, which take ∼1 ps. The asymmetry in the hot carrier dynamics is shown to originate from the valence band degeneracy, the heavier effective mass of holes compared to electrons, and the details of the coupling to different phonon modes in the valence and conduction bands. We show that the slow cooling of hot electrons and their long ballistic mean free paths (over 3 nm at room temperature) are a possible cause of efficiency droop in GaN light-emitting diodes. Taken together, our work sheds light on the ultrafast dynamics of hot carriers in GaN and the nanoscale origin of efficiency droop.

KEYWORDS: *Gallium nitride, light-emitting diodes, ultrafast dynamics, electron−phonon scattering*

Wurtzite GaN has emerged as a promising material for solid-state lighting[1] and power electronics;[2,3] its potential technological benefits are driving intense research in industry and academia. However, material properties essential for device performance and energy efficiency, such as carrier transport and recombination, are not completely understood in GaN and remain the subject of debate. Carrier transport and ultrafast dynamics are regulated by scattering with phonons, carriers, and impurities.[4] In particular, the electron−phonon (e-ph) interaction[5,6] plays a dominant role in transport at room temperature in relatively pure materials. It further regulates the energy loss (or "cooling") of excited carriers injected at heterojunctions, a scenario of relevance in GaN light-emitting diodes. The excited (so-called "hot") carriers rapidly lose their excess energy with respect to the band edges, dissipating heat by phonon emission through e-ph coupling. Hot carriers (HCs) are also central in degradation and current leakage in GaN transistors for power electronics[7,8] and set the operational basis for hot electron transistors.[9]

Microscopic understanding of carrier dynamics is challenging in GaN since experimental results are modulated by defects and interfaces, and are typically interpreted with empirical models.[10−16] For example, the hot electron cooling times measured by different groups range over 2 orders of magnitude,[10,12−17] and reports of hot hole dynamics are scarce.[11] In addition, the efficiency decline in GaN light-emitting diodes (LEDs) at high current, a process known as efficiency droop,[18] has been intensely investigated but its origin

remains unclear. First-principles calculations focused on Auger recombination[19,20] as a possible cause, though other mechanisms that have been proposed,[18] including HC effects and electron leakage, have seen less extensive theoretical treatment compared to Auger, leaving simplified models to guide device design.

We recently developed first-principles calculations of carrier dynamics[6] that can obtain carrier mobility,[21,22] ultrafast dynamics,[23−26] HC relaxation times,[23,24] and ballistic mean free paths,[23,25] which are in excellent agreement with experiment. These approaches are free of empirical parameters and use the structure of the material as the only input. In particular, we recently developed a method[22] to accurately compute the e-ph relaxation times (RTs), namely the average time between e-ph collisions, in polar materials, as is needed for GaN. These approaches are extended in this work to investigate HC dynamics in GaN from first principles.

Here, we compute the e-ph RTs over a wide energy range, and study the cooling of HCs by numerically solving the electron Boltzmann transport equation (BTE). Both the RTs and the simulated ultrafast dynamics reveal a large asymmetry between the hot electron and hole dynamics, with hot holes relaxing to the band edges significantly faster than hot electrons. The origin of this asymmetry, the role of different
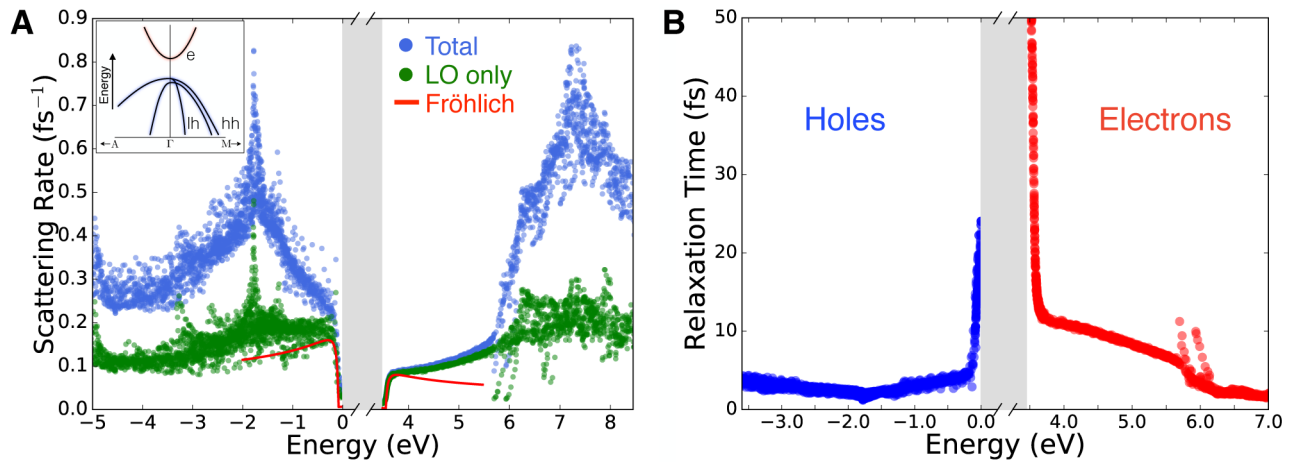
Figure 1. First-principles e-ph scattering. (A) e-ph scattering rates in GaN at 300 K, for electrons and holes with energies within 5 eV of the band edges. Both the LO contribution and the total rate due to all phonon modes are shown. The red lines are the empirical Fröhlich scattering rates (see eq 1). A schematic of the bandstructure of GaN near the band edges is shown in the inset. (B) The e-ph relaxation times, defined as the inverse of the total scattering rates in (A). The zero of the energy axis is the valence band maximum, and the band gap is shown as a shaded area.

phonon modes and the limitations and failure of phenomenological models are analyzed in detail. We additionally find significantly longer mean free paths (MFPs) for electrons compared to holes, with implications for GaN devices. We show that the slow cooling rate of hot electrons can lead to inefficient light emission at high current, thus demonstrating that the nanoscale dynamics of HCs play a key role in LED efficiency droop.

Electron−Phonon Scattering. In polar materials like GaN, empirical models typically assume that polar optical phonons, and in particular the longitudinal optical (LO) mode, dominate carrier scattering due to their long-range interactions with carriers. The empirical Fröhlich model[27] for the LO mode predicts an e-ph coupling matrix element $g_F(q) = [(e^2\hbar\omega_0)/(2q)]\epsilon_{ph}^{-1}$, where $\mathbf{q}$ is the phonon wavevector (and q is its magnitude), $\hbar\omega_0$ is the LO phonon energy, and $\epsilon_{ph}^{-1} = \epsilon_\infty^{-1} - \epsilon_0^{-1}$ is the phonon contribution to the dielectric screening, with ϵ_∞ and ϵ_0 as the high- and low-frequency dielectric constants, respectively. The intravalley e-ph scattering rate $\Gamma_{\mathbf{k}}$ (where $\mathbf{k}$ is the electron crystal momentum) due to the empirical Fröhlich coupling g_F can be obtained analytically for carriers in a spherical parabolic band with effective mass m^* and energy $E_{\mathbf{k}} = \hbar^2 k^2 / 2m^*$[5,28]

$$\Gamma_{\mathbf{k}} = \tau_0^{-1}\left(\frac{E_{\mathbf{k}}}{\hbar\omega_0}\right)^{-1/2}\left[N_0\sinh^{-1}\left(\frac{E_{\mathbf{k}}}{\hbar\omega_0}\right)^{1/2} + (N_0+1)\sinh^{-1}\left(\frac{E_{\mathbf{k}}}{\hbar\omega_0}-1\right)^{1/2}\right] \tag{1}$$

where N_0 is the Bose−Einstein occupation factor for LO phonons and $\tau_0^{-1} = \epsilon_p^{-1}\left[(e^2\omega_0)/(2\pi\hbar)\right]\cdot[m^*/(2\hbar\omega_0)]^{1/2}$ is the inverse Fröhlich time. In this work, the rate in eq 1 is compared with our first-principles results.

Following an approach we recently developed,[22] we combine density functional theory (DFT),[29] density functional perturbation theory (DFPT),[30] and ab initio Fröhlich[31] calculations to compute the short- and long-range contributions to the e-ph coupling matrix elements, which are then interpolated on fine Brillouin zone (BZ) grids to converge the e-ph scattering rates $\Gamma_{n\mathbf{k}}$ for each electronic band n and crystal momentum $\mathbf{k}$ (see Methods). For the polar LO mode scattering, our ab initio Fröhlich calculations[31] differ in important ways from the

empirical Fröhlich model as they include Born effective charges and anisotropic dielectric tensors, both computed with DFPT, and account for the electronic bandstructure and phonon dispersions (see Methods). Here and in the following, the carrier excess energies are defined as the energy above the conduction band minimum (CBM) for the electrons, and the energy below the valence band maximum (VBM) for the holes.

The scattering rates and their inverse, the e-ph RTs $\tau_{n\mathbf{k}} = \Gamma_{n\mathbf{k}}^{-1}$, contain rich microscopic information on the carrier dynamics. The bandstructure near the band edges, given for reference in the inset of Figure 1A, crucially determines the e-ph scattering rates. In GaN, the conduction band exhibits a single parabolic valley at Γ, and the valence band edge consists of a light-hole and two heavy-hole bands with anisotropic effective masses and degeneracy along the Γ−A direction. Figure 1A shows our computed e-ph scattering rates of electrons and holes with energies within 5 eV of the band edges. Both the total scattering rate contributed by all phonons and the LO mode contribution alone are shown; the empirical Fröhlich rate of eq 1, computed using parameters from the literature,[32,33] is also given for comparison. The full mode-resolved scattering rates are shown in Figure S1 of the Supporting Information. For both electrons and holes, the scattering rate in Figure 1A is very small within an LO phonon energy ($\hbar\omega_0 \approx 100$ meV) of the band edges, because the dominant process in this energy range is LO phonon absorption. The scattering rate increases rapidly above $\hbar\omega_0$ up to ∼150 meV excess energy due to an increase in the phase space for LO phonon emission. The trend at low energy in the conduction band shows that LO phonon emission dominates intravalley scattering in the conduction valley at Γ, roughly up to 2 eV above the CBM. At higher energy in the conduction band, the total and LO mode electron scattering rates differ substantially due to interband scattering mediated by all phonon modes in roughly equal measure (see Figure S1).

In the valence band, due to the presence of multiple bands at Γ, both intravalley and intervalley small-$\mathbf{q}$ transitions are possible within ∼1 eV of the VBM, thus resulting in a higher LO scattering rate (by roughly a factor of 2) than in the conduction band within 1 eV of the CBM. The LO contribution becomes roughly constant at hole excess energies greater than ∼0.2 eV. Different from the conduction band

where LO phonon emission is the only active process, the total scattering rate keeps increasing for holes with excess energy above 0.2 eV due to intravalley and intervalley scattering contributed in roughly equal measure by all acoustic and optical phonon modes (see Figure S1). Within 2 eV of the band edges, the DOS in the valence band is greater than in the conduction band due to the presence of multiple valence bands and to the higher effective masses of holes compared to electrons. These effects result in scattering from phonon modes other than the LO (which are referred to here as "nonpolar" phonons) becoming important at lower excess energies in the valence band due to a greater phase space for large-$\mathbf{q}$ scattering. Overall, the stronger LO polar and nonpolar contributions in the valence band result in significantly higher scattering rates for holes compared to electrons within 2 eV of the band edges. This asymmetry has important consequences for carrier dynamics.

The empirical Fröhlich model shows two major limitations in reproducing the first-principles trends (see Figure 1A). The empirical and first-principles rates exhibit opposite trends at excess energies greater than ∼150 meV. For energies where the LO phonon emission rate is roughly constant, the discrepancy of the empirical Fröhlich rate is as large as 30−50% for both electrons and holes. Note that first-principles calculations include all phonon modes on the same footing at all energies, while empirical e-ph calculations would account for modes other than the LO through mode-specific empirical deformation potentials.[28] Traditionally employed empirical models are thus inadequate to compute the e-ph scattering rates due to all phonon modes over a wide energy range, as is done here, and the first-principles approach is necessary.

As a consequence of the scattering rate asymmetry, the computed e-ph RTs (see Figure 1B) of holes are overall significantly shorter than the RTs of electrons. Within 2 eV of the band edges, the electron RTs range between 10−50 fs, while the hole RTs are of order 3−20 fs. The electron RT above the threshold for LO phonon emission in the conduction band is ∼12 fs, which is in very good agreement with the LO phonon emission time of 16 fs recently measured for electrons by Suntrup et al.[17] The detailed energy dependence of the RTs and scattering rates reported here is valuable for GaN device design.

Origin of the Carrier Relaxation Asymmetry. We address the question of whether the asymmetry between the electron and hole scattering rates found here is a mere consequence of the heavier effective mass of holes compared to electrons in GaN. In doing so, we develop an intuition for the origin of this asymmetry by analyzing separately the polar and nonpolar e-ph scattering contributions. As noted above, the two sources of e-ph scattering are the long-range interaction from the LO polar mode and the short-range interactions from all other nonpolar phonons. Because e-ph processes are determined by the e-ph coupling strength and the phase space available for scattering (see Methods, eq 3), the nonpolar scattering rate $\Gamma^{(NP)}$ approximately follows the same energy trend as the electronic DOS, $D(E)$, multiplied by an average e-ph coupling strength $\langle g^2 \rangle$, so that $\Gamma^{(NP)}(E) \propto \langle g^2 \rangle D(E)$. On the other hand, due to the long-range electrostatic nature of the polar interaction, the strength of LO coupling is insensitive to the specific electronic states involved in the scattering process. The LO mode polar coupling behaves as $|g^{LO}(q)|^2 \sim 1/q^2$ at small phonon wavevector q, resulting in much stronger LO scattering for small-$\mathbf{q}$ transitions. Because of this particular

phonon wavevector dependence, the scattering rate $\Gamma^{(P)}$ due to the polar LO mode approaches the band edge with a constant trend in energy, as opposed to being proportional to the DOS as in the nonpolar case. In the empirical Fröhlich formula (see eq 1), the polar scattering rate behaves as a simple function of the effective mass, $\Gamma^{(P)} \propto (m^*)^{1/2}$.

To quantify the scattering rate asymmetry, we compute in Figure 2 the ratio of the BZ averaged scattering rate of holes
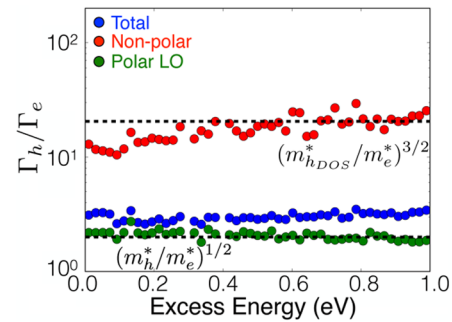

Figure 2. Origin of the scattering rate asymmetry. Ratio of the Brillouin zone averaged e-ph scattering rates of holes (Γ_h) to those of electrons (Γ_e) as a function of carrier excess energy. The zeros of the excess energy are the conduction and valence band edges for electrons and holes, respectively. The data points are computed using rates due to polar LO phonons (green), nonpolar phonons (red), and all phonon modes (blue). The dashed lines indicate the ratios $(m_h^*/m_e^*)^{1/2}$ and $(m_{h_{DOS}}^*/m_e^*)^{3/2}$ discussed in the text.

(Γ_h) to that of electrons (Γ_e). The ratio Γ_h/Γ_e is shown for the average total, polar and nonpolar scattering rates as a function of excess energy. For the LO polar contribution, the ratio closely matches the result expected based on the empirical Fröhlich model (see eq 1), namely, $\Gamma_h^{(P)}/\Gamma_e^{(P)} \approx (m_h^*/m_e^*)^{1/2}$, where m_h^* is the experimental hole effective mass.[32] The agreement between the empirical and ab initio ratios of LO mode scattering rates indicates that the heavier hole effective mass is the main source of the LO polar scattering asymmetry, and that the empirical and ab initio treatments roughly factor out in the polar scattering rate ratio when a single m_h^* value is employed as a proxy of the multiple hole bands.

Because the nonpolar scattering rate is proportional to the DOS, the ratio between the ab initio nonpolar scattering rates is compared in Figure 2 with a heuristic DOS ratio. The DOS of a parabolic band is proportional to $(m^*)^{3/2}$,[34] so we expect that the ratio of the nonpolar scattering rates at low energy is approximately $\Gamma_h^{(NP)}/\Gamma_e^{(NP)} \approx (m_{h_{DOS}}^*/m_e^*)^{3/2}$, where $m_{h_{DOS}}^*$ is the hole DOS effective mass.[33] The ab initio and DOS-based empirical nonpolar ratios are in reasonable agreement. However, the inaccuracy of approximating multiple valence bands with a single DOS effective mass for holes, combined with the stronger average e-ph coupling strength for electrons (see Figure S2 and Section A in the Supporting Information), both push the ratio below the heuristic prediction, with an energy-dependent discrepancy. Therefore, the ratio between the nonpolar scattering rates cannot be accurately estimated without detailed knowledge of the bandstructure and e-ph coupling strengths.

Finally, since the total scattering rate is determined by the sum of the polar and nonpolar contributions, the ratio between the total scattering rates of holes and electrons cannot be estimated by a simple heuristic model based on the effective
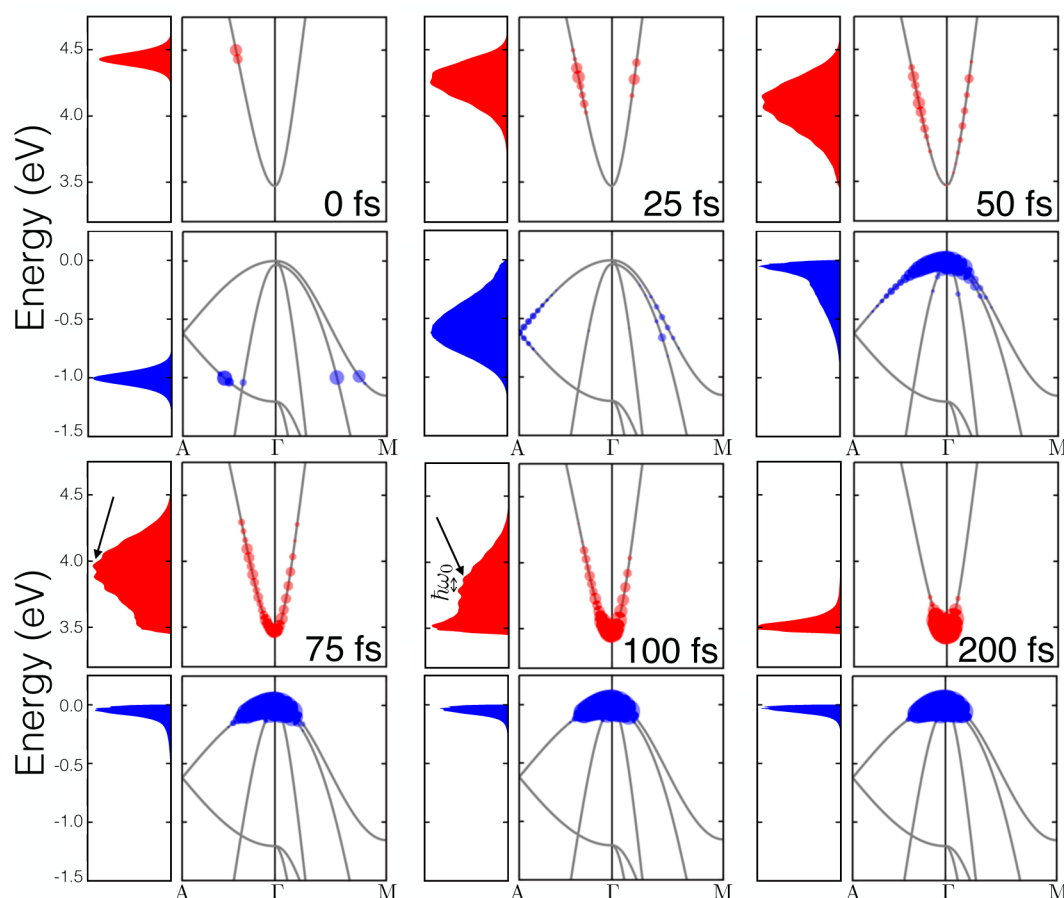
Figure 3. Ultrafast dynamics of hot carriers in GaN. Hot electrons (red) and holes (blue) are injected with a 1 eV excess energy with respect to the band edges, as modeled by creating initial Gaussian carrier distributions with a small energy width. The occupations $f_{n\mathbf{k}}$ for electrons and $1 - f_{n\mathbf{k}}$ for holes are shown at various times in the square panels along the A−Γ−M line of the Brillouin zone with the point size proportional to the logarithm of the occupations. Left of each occupation panel we plot the average carrier concentrations (in arbitrary units) as a function of energy. The oscillations in the electron concentration are indicated with an arrow in the 75 and 100 fs frames, along with their spacing of $\hbar\omega_0 \approx 100$ meV.

masses. The average ab initio ratio between the total scattering rates found here is $\Gamma_h/\Gamma_e \approx 3$ within 1 eV of the band edges and is bracketed by the nonpolar and polar ratios. Detailed first-principles calculations, as employed here, are necessary to quantify this asymmetry in the scattering rates.

Real-Time Hot Carrier Dynamics. Simulating the HC dynamics in real time further highlights the different behavior of electrons and holes in GaN. To this end, we carry out numerical simulations of the dynamics of hot electron and hole populations injected in GaN with a range of initial excess energies. To represent the injected carriers, we employ narrow Gaussian distributions centered at the initial excess energy and solve the electron BTE[5] in real time fully ab initio,[6] using first-principles e-ph matrix elements, bandstructures, and phonon dispersions (see Methods). The carrier occupations are time-stepped using a fourth-order Runge−Kutta algorithm, while the phonon occupations are kept at 300 K, so that hot phonon effects are neglected.

Figure 3 shows the time evolution of the electron and hole populations after injection with a 1 eV initial excess energy. At each time step, we analyze both the carrier occupations ($f_{n\mathbf{k}}$ for electrons and $1 - f_{n\mathbf{k}}$ for holes) along the A−Γ−M line of the BZ and the carrier concentrations as a function of energy, $\bar{f}(E)$, obtained by integrating the occupations at each energy over the BZ (see Methods). Following HC generation, the electron and hole distributions broaden in energy and approach a Fermi−

Dirac-like shape as they shift toward lower excess energies, eventually reaching the band edges. Both electrons and holes ultimately thermalize to a 300 K Fermi−Dirac distribution in equilibrium with the phonons, thus reaching the correct long-time limit for our simulations.

We find that while holes reach the band edges and cool in ~80 fs, electrons are still far from equilibrium at the same time. Electron cooling is roughly five times slower, with electrons relaxing to the band edges in ~200 fs after injection and fully thermalizing to the 300 K equilibrium distribution in 400 fs. Interestingly, the dominant LO phonon emission for electrons results in distribution oscillations in energy spaced apart by $\hbar\omega_0 \approx 100$ meV. These wiggles in the distributions are seen most clearly for electrons in the 75 and 100 fs panels in Figure 3. The same simulation carried out for an initial carrier excess energy of 0.5 eV (see Figure S3 in Supporting Information) shows even more pronounced oscillations in the electron population. The asymmetry found here in the time scale for hot electron and hole cooling has not been reported previously and is distinct from the transport asymmetry[18] due to the different effective masses and mobilities of electrons and holes.

Comparing our real-time dynamics with experiments is not straightforward and should be done keeping in mind that our results pertain to the low carrier concentration regime. Several ultrafast pump−probe experiments have been carried out in GaN but due to the high HC concentrations reached in some
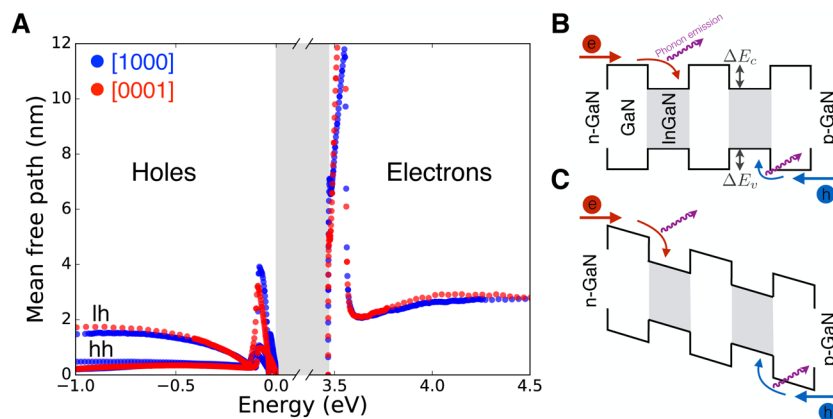
Figure 4. Role of hot carriers in efficiency droop. (A) Mean free paths, computed using calculated band velocities and e-ph relaxation times, along two crystal directions in GaN. The c-axis direction is labeled as [0001] and the in-plane direction as [1000]. For holes, the lower mean free path branch is due to the heavy-hole (hh) band, and the upper branch to the light-hole (lh) bands. The zero of the energy axis is the valence band maximum, and the band gap is shown with a shaded area. (B,C) Schematic band diagram of the active region of a InGaN/GaN LED. Shown are the stacked quantum wells, both without and with a driving electric field in (B,C), respectively. The valence and conduction band offsets, ΔE_v and ΔE_c, are also indicated.

of these measurements, hot phonon effects become relevant as the emitted phonons re-excite the carriers, thus slowing down HC cooling compared to the low carrier concentration regime studied here. For example, Ye et al.[10] measured hot electron cooling in n-doped GaN. They concluded that LO emission is the dominant mechanism for hot electron cooling, consistent with our findings. They extracted a 0.2 ps LO emission time from their measurements, and observed that this value is significantly longer than the ∼10 fs empirical Fröhlich RT; this discrepancy is attributed in their work to hot phonon effects resulting from the high carrier density. In a subsequent work, the same authors measured the cooling of hot holes in p-doped GaN.[11] They found that the hot hole data is difficult to fit with an LO phonon emission model because of the complexity of the valence band. We argue that the reason for this discrepancy, as shown in our work, is the significant scattering by phonon modes other than the LO in the valence band, even at low excess energy. While differences in the electron and hole dynamics are not discussed by Ye et al.,[11] note that in their work the details of the early time evolution cannot be seen due to the 100 fs duration of the pump pulse, which sets the time resolution of the measurement. Our data provides evidence for an asymmetry between the electron and hole cooling times and mechanisms, which should be observable in new experiments with ∼10 fs resolution.

Role of Hot Carriers in Efficiency Droop. We analyze the role of HCs in GaN LED devices. GaN and III-nitride LEDs are bipolar devices that emit light as a result of the recombination of electrons and holes in quantum well (QW) active layers.[1,35] A key figure of merit in LEDs is the external quantum efficiency (EQE), defined as the ratio of generated photons per electron–hole pair introduced at the contacts. One of the main barriers toward commercialization of GaN LEDs is efficiency droop, a decay of the EQE at high driving current. Almost all GaN LEDs exhibit a maximum EQE at low current densities, with the EQE rapidly decreasing above the optimal current.[18]

In a typical GaN LED design,[1] electrons and holes are injected at the opposite ends of an active region that contains multiple InGaN QWs and has the bandstructure shown schematically in Figure 4B. The multi-QW active region typically consists of up to 10 periods of an ∼3 nm thick

In$_x$Ga$_{1-x}$N QW followed by an undoped ∼10 nm GaN spacer,[18] a structure designed to maximize carrier confinement and recombination in the light-emitting InGaN QWs. Because of the band offsets at the GaN/InGaN interface, electrons and holes are hot when injected into the QW, and need to cool to the band edges to emit light efficiently. The excess energy of the HCs injected in the QW depends on the band offsets ΔE_c for electrons and ΔE_v for holes (see Figure 4B), which are typically of order 0.5 eV for InGaN/GaN devices.

Carriers injected into the active region scatter with phonons, and after several phonon emission events lose enough energy to be confined in the QW by the band offset potential, following which the carriers cool to the QW band edges and recombine to emit light. The e-ph mean free paths (MFPs) L_{nk} characterize the average length traveled between phonon emission events by carriers moving in the direction $\hat{\mathbf{k}}$ in band n (here, the top valence bands for holes and bottom conduction band for electrons). The MFPs are obtained here as $L_{nk} = v_{nk}\tau_{nk}$, using Wannier-interpolated band velocities v_{nk}[36,37] and the e-ph relaxation times τ_{nk} given in Figure 1B. Figure 4A shows the computed e-ph MFPs for carriers moving along the c-axis [0001] and in-plane [1000] directions in GaN, for energies within 1 eV of the band edges, which are typical for carrier injection in GaN LEDs. The slow hot electron cooling via LO emission results in relatively long electron MFPs ranging from 10 nm near the CBM to 3 nm between 0.25−1 eV above the CBM. For comparison, the hole MFPs are much shorter, ∼3 nm close to the VBM and only 0.5 nm for the heavy-hole and 1.5 nm for the light-hole bands between 0.25−1 eV below the VBM. Note that these MFPs are lower bounds that pertain to the low carrier concentration regime, and that hot phonon effects can increase these values at high carrier concentration, especially for electrons where the only active e-ph scattering mechanism is LO phonon emission.

Because the typical QW thickness is ∼3 nm, the longer electron MFPs indicate that electrons need to cross on average several QWs before cooling to the band edge, while holes can cool effectively within a single QW. A fact relevant for droop is that as the current is increased, the driving electric field increases and further tilts the band edges in the active region, resulting approximately in the tilted staircase potential shown in

Figure 4C. For typical LEDs operated at droop regime currents, using the devices in ref 38 as an example, the driving voltage is roughly 5 V. Assuming that ~3 V fall over a roughly 150 nm thick undoped active region, the resulting active region electric field is of order 2×10^5 V/cm. For electrons losing an LO phonon energy of ~100 meV as they travel a 3 nm MFP length, the cooling rate is the order of 100 meV over 3 nm, and thus roughly 3×10^5 eV/cm. Droop thus occurs in a regime where the band tilting and the electron cooling rate are comparable, so that electrons cannot cool effectively against the tilted band edges. In this scenario, electrons that do not cool in the QW are swept by the electric field into the next QW, until a fraction of electrons leaks out of the active region without recombining radiatively, leading to a decreased EQE at high current as observed in efficiency droop. Our analysis highlights that HC cooling, and in particular the slow LO phonon emission rate of hot electrons occurring over lengths comparable with the QW thickness, can be an important factor determining efficiency droop.

The mitigation of droop seen experimentally by increasing the QW thickness[39] and number of QWs in the active region[38] is consistent with our findings. In our model of hot electron leakage, increasing the QW thickness or the total active region thickness will improve HC cooling and decrease the probability of electrons overflying the QWs, thus increasing the EQE. Interestingly, several reports attribute the droop mitigation primarily to reduced Auger recombination rather than HC effects. Although Auger processes may play a role in droop, experimental evidence for Auger recombination in GaN LEDs is still scarce.[18,40] Auger processes, if present with a high enough rate, would need to coexist with the e-ph scattering processes studied here, which are relevant at all carrier concentrations. Most notably, it has been found that the mechanism responsible for droop increases in strength at lower temperatures.[41] This behavior is in contrast with Auger scattering, because the dominant phonon-assisted Auger processes[20] would lead to stronger droop at higher temperatures. Note instead that e-ph scattering increases with increasing temperatures, so that lower temperatures are associated with longer hot carrier MFPs. Our model thus predicts that lower temperatures lead to longer electron MFPs and thus increased electron leakage and efficiency droop, consistent with experiment. The temperature dependence clearly suggests that e-ph scattering plays a key role in the efficiency droop.

Several authors have discussed HC effects and QW overfly as a potential source of droop,[18,42] showing that the use of electron blocking layers to stop carrier leakage and staircase injectors to lower carrier excess energy can drastically mitigate droop. Our first-principles calculations of e-ph scattering, HC cooling, and MFPs provide a quantitative basis for understanding such HC effects. We conclude that the different cooling rates of hot electrons and holes found here, together with the longer electron MFPs compared to holes, can play an important role in GaN LEDs efficiency droop and deserve further experimental investigation.

In summary, we apply first-principles calculations to shed light on the microscopic HC dynamics in GaN, providing details beyond the reach of previously employed theoretical and experimental methods. Our key findings include a significant difference in hot hole and electron scattering and cooling, the origin of this asymmetry, and the non-negligible role of scattering mechanisms besides the polar LO mode for holes. A model is presented to explain droop as a consequence of slow hot electron cooling. Our results can be employed to include HC dynamics in GaN device scale modeling, which typically neglects HC effects. Future work will extend our calculations to InGaN and AlGaN. Taken together, our computational approach advances the design of GaN devices and enables the engineering of novel lighting materials with microscopic insight.

Methods. *Computational Methods.* We carry out ab initio calculations on GaN in the wurtzite structure with relaxed lattice parameters of $a = 3.17$ Å and $c = 5.16$ Å. The ground-state electronic structure is computed using the Quantum Espresso code[43] within the local density approximation (LDA)[44] of DFT. We employ a plane-wave kinetic energy cutoff of 80 Ry and scalar-relativistic norm-conserving pseudopotentials[45] for both Ga and N. The pseudopotential of Ga includes a nonlinear core correction[46] to account for the effect of the shallow 3d core states. The ground-state charge density is obtained using a $12 \times 12 \times 12$ **k**-point grid, following which a nonself-consistent calculation is employed to obtain the Kohn−Sham eigenvalues and wave functions on an $8 \times 8 \times 8$ **k**-point grid. We construct maximally localized Wannier functions (WFs)[47] using the Wannier90 code.[48] The Kohn−Sham wave functions are first projected onto 4 sp^3 orbitals on each Ga and N atom for a total of 16 wannierized bands. The WF spread is then minimized, and the relevant energy windows[47] are adjusted until the interpolated bandstructure can smoothly reproduce the LDA result within ~10 meV throughout the BZ. The LDA eigenvalues are then corrected using the GW[49] quasiparticle energies of GaN computed in ref 50, by extracting an average energy dependent self-energy that is employed to apply a scissor shift of the conduction band minimum and a linear energy stretch of the valence and conduction band energies. Using the GW-corrected eigenvalues together with the WFs, interpolated GW quasiparticle energies with quality comparable to those in ref 50 are obtained and employed in all calculations described below. The WF projected DOS in Figure S2 of the Supporting Information is computed by projecting the Kohn−Sham wave functions $|\psi_{n\mathbf{k}}\rangle$ onto the WFs localized on the Ga and N atoms, respectively, using $\sum_{n\mathbf{k}, \alpha \in \{Ga, N\}} \langle \psi_{n\mathbf{k}} | \alpha \rangle \langle \alpha | \psi_{n\mathbf{k}} \rangle \delta(E - E_{n\mathbf{k}})$.

Electron−Phonon Scattering Calculations. We use density functional perturbation theory (DFPT)[30] to compute lattice dynamical properties and e-ph matrix elements $g_{nn'\nu}(\mathbf{k},\mathbf{q})$[6] on coarse $8 \times 8 \times 8$ **k**-point and $4 \times 4 \times 4$ **q**-point grids in the BZ. Here and in the following, the e-ph matrix elements[6] describe an electron in Bloch state $|n\mathbf{k}\rangle$ with quasiparticle energy $E_{n\mathbf{k}}$ that scatters into the state $|n'\mathbf{k}+\mathbf{q}\rangle$ with quasiparticle energy $E_{n'\mathbf{k}+\mathbf{q}}$ due to a phonon with branch index ν, wavevector $\mathbf{q}$, and frequency $\omega_{\nu\mathbf{q}}$. The electron and phonon energies and the e-ph matrix elements are then interpolated on significantly finer grids using WFs.[51] Wannier interpolation of the matrix elements relies on the spatial localization of the e-ph interaction. In polar materials like GaN, however, the interaction of an electron with an LO phonon diverges like $\sim 1/q$ as predicted by the Fröhlich model.[27,31] This singularity for $q \to 0$ results from the long-range field generated by the LO mode and is incompatible with Wannier interpolation. To overcome this issue, the e-ph matrix elements are separated into short- (S) and long-range (L) parts,[31] $g_{nn'\nu}(\mathbf{k},\mathbf{q}) = g_{nn'\nu}^{S}(\mathbf{k},\mathbf{q}) + g_{nn'\nu}^{L}(\mathbf{k},\mathbf{q})$, where g^S is computed by Wannier interpolation and g^L is evaluated using an analytical ab initio Fröhlich e-ph vertex based on the Vogl model[52]

$$g^{L}_{nn'\nu}(\mathbf{k}, \mathbf{q}) = i\frac{4\pi}{\Omega}\frac{e^2}{4\pi\epsilon_0}\sum_{\kappa}\left(\frac{\hbar}{2\omega_{\nu\mathbf{q}}NM_{\kappa}}\right)^{1/2}\times\sum_{\mathbf{G}\neq-\mathbf{q}}\frac{(\mathbf{q}+\mathbf{G})\cdot Z_{\kappa}^{*}\cdot\mathbf{e}_{\kappa\nu}(\mathbf{q})}{(\mathbf{q}+\mathbf{G})\cdot\epsilon^{\infty}\cdot(\mathbf{q}+\mathbf{G})}$$

$$\times\langle\Psi_{n'\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{q}+\mathbf{G})\cdot(\mathbf{r}-\tau_{\kappa})}|\Psi_{n\mathbf{k}}\rangle \quad (2)$$

where $\mathbf{G}$ is a reciprocal lattice vector, Ω is the volume of the unit cell, N is the number of points in the $\mathbf{q}$-grid, ϵ^{∞} is the high-frequency permittivity, M_{κ}, τ_{κ}, and Z_{κ}^{*} are the mass, position, and Born effective charge tensor of atom κ in the unit cell, respectively, and $\mathbf{e}_{\kappa\nu}(\mathbf{q})$ is a vibrational eigenmode normalized in the unit cell. Note that we neglected a quadrupole moment contribution to polar phonon scattering due to the piezoelectric interaction;[22,52] we have verified that this approximation has a negligible effect on the hot carrier dynamics.

The e-ph scattering rate $\Gamma_{n\mathbf{k}}$ for an electronic state with band n and crystal momentum $\mathbf{k}$ is obtained from the imaginary part of the lowest-order e-ph self-energy, using[5,6]

$$\Gamma_{n\mathbf{k}} = \frac{2\pi}{\hbar}\sum_{n'\mathbf{q}\nu}|g_{nn'\nu}(\mathbf{k}, \mathbf{q})|^2$$

$$\times\left[(N_{\nu\mathbf{q}}+1-f_{n'\mathbf{k}+\mathbf{q}})\delta(E_{n\mathbf{k}}-\hbar\omega_{\nu\mathbf{q}}-E_{n'\mathbf{k}+\mathbf{q}})\right.$$

$$\left.+(N_{\nu\mathbf{q}}+f_{n'\mathbf{k}+\mathbf{q}})\delta(E_{n\mathbf{k}}+\hbar\omega_{\nu\mathbf{q}}-E_{n'\mathbf{k}+\mathbf{q}})\right] \quad (3)$$

where $E_{n\mathbf{k}}$ and $\hbar\omega_{\nu\mathbf{q}}$ are the electron quasiparticle and phonon energies, respectively, and $f_{n\mathbf{k}}$ and $N_{\nu\mathbf{q}}$ the corresponding equilibrium occupations at 300 K. The scattering rate due to a given phonon mode is obtained by restricting the sum in eq 3 to the corresponding phonon branch index ν. Using an approach we recently developed,[22] we compute and converge the scattering rate in eq 3 with an in-house developed code that can carry out, among other tasks, e-ph calculations analogous to the EPW code.[53] We split $|g|^2$ in eq 3 into the long-range part $|g^L|^2$ and the remainder ($|g|^2 - |g^L|^2$) and separately compute and converge $\Gamma_{n\mathbf{k}}$ for these two contributions, which are then added up to obtain the total scattering rate. This approach leads to a dramatic speed-up compared to converging eq 3 with $|g|^2 = |g^S + g^L|^2$ directly.[22] The scattering rate is computed for $\mathbf{k}$ points on a fine $100\times100\times100$ grid in the BZ, using Gaussian broadening with a small parameter of 8 meV to approximate the δ function in eq 3. For each $\mathbf{k}$ point, we converge the long-range contribution with 10^6 random $\mathbf{q}$ points sampled from a Cauchy distribution,[22] and the remainder contribution with 10^5 random $\mathbf{q}$ points from a uniform distribution. Convergence of all quantities is carefully verified.

Carrier Dynamics Simulations. We simulate HC cooling due to e-ph scattering in bulk GaN in the absence of external fields. The initial HC distribution is modeled with a narrow Gaussian centered at the initial HC energy. Separate calculations are carried out for electrons and holes and for the two values of initial excess energy studied here (0.5 and 1 eV). The time evolution of the carrier distributions is obtained by solving the BTE[5,6]

$$\frac{\partial f_{n\mathbf{k}}(t)}{\partial t} = -\frac{2\pi}{\hbar}\sum_{n'\mathbf{q}\nu}|g_{nn'\nu}(\mathbf{k}, \mathbf{q})|^2[\delta(E_{n\mathbf{k}}-\hbar\omega_{\nu\mathbf{q}}-E_{n'\mathbf{k}+\mathbf{q}})F_{em}(t)$$

$$+\delta(E_{n\mathbf{k}}+\hbar\omega_{\nu\mathbf{q}}-E_{n'\mathbf{k}+\mathbf{q}})F_{abs}(t)] \quad (4)$$

where $f_{n\mathbf{k}}(t)$ is the time-dependent electron distribution, the e-ph matrix elements $g_{nn'\nu}(\mathbf{k},\mathbf{q})$ are computed in the ground state, and the phonon emission (F_{em}) and absorption (F_{abs}) terms are computed at each time step as[6]

$$F_{abs} = f_{n\mathbf{k}}(1-f_{n'\mathbf{k}+\mathbf{q}})N_{\nu\mathbf{q}} - f_{n'\mathbf{k}+\mathbf{q}}(1-f_{n\mathbf{k}})(N_{\nu\mathbf{q}}+1)$$

$$F_{em} = f_{n\mathbf{k}}(1-f_{n'\mathbf{k}+\mathbf{q}})(N_{\nu\mathbf{q}}+1) - f_{n'\mathbf{k}+\mathbf{q}}(1-f_{n\mathbf{k}})N_{\nu\mathbf{q}} \quad (5)$$

Phonon–phonon scattering and the change in phonon occupations are neglected in our simulations, where we fix $N_{\nu\mathbf{q}}$ to the equilibrium phonon occupations at 300 K. This approximation is justified in the low carrier concentration limit, in which hot phonon effects are negligible. We solve eq 4 numerically using the fourth-order Runge–Kutta method with a time step of 1 fs, using uniform fine grids of up to $100\times100\times100$ $\mathbf{k}$- and $\mathbf{q}$-points in the BZ. The BZ-averaged energy-dependent carrier populations $\bar{f}(E,t)$ at energy E used in Figure 3 are obtained as $\bar{f}(E, t) = \sum_{n\mathbf{k}}f_{n\mathbf{k}}(t)\,\delta(\epsilon_{n\mathbf{k}}-E)$ via tetrahedron integration. We developed an efficient scheme to speed up the solution of the BTE to be detailed elsewhere, which combines MPI and OpenMP parallelizations to repeatedly compute the e-ph matrix elements and converge the scattering integral at each time step.

■ ASSOCIATED CONTENT

ⓈSupporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.nano-lett.7b02212.

> (Section A) Origin of the stronger average e-ph coupling strength for electrons compared to holes; (Figure S1) Phonon mode-resolved electron–phonon scattering rates; (Figure S2) Origin of the stronger average electron–phonon coupling strength for electrons compared to holes; (Figure S3) Simulated hot carrier dynamics for 0.5 eV initial excess energy (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bmarco@caltech.edu. Phone: +1 (0)626 3952515.

ORCID Ⓘ

Vatsal A. Jhalani: 0000-0003-0866-0858

Author Contributions

M.B. conceived and designed the research. V.A.J. and J.-J.Z. developed the computational codes and carried out the calculations. All authors wrote the manuscript. V.A.J. and J.-J.Z. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Pimputkar, S.; Speck, J. S.; DenBaars, S. P.; Nakamura, S. *Nat. Photonics* **2009**, *3*, 180−182.

(2) Dimitrijev, S.; Han, J.; Moghadam, H. A.; Aminbeidokhti, A. *MRS Bull.* **2015**, *40*, 399−405.

(3) Okumura, H. *MRS Bull.* **2015**, *40*, 439−444.

(4) Ziman, J. M. *Electrons and Phonons: The Theory of Transport Phenomena in Solids*; Oxford University Press, 1960.

(5) Mahan, G. D. *Condensed Matter in a Nutshell*; Princeton University Press, 2010.

(6) Bernardi, M. *Eur. Phys. J. B* **2016**, *89*, 239.

(7) Meneghesso, G.; Verzellesi, G.; Danesin, F.; Rampazzo, F.; Zanon, F.; Tazzoli, A.; Meneghini, M.; Zanoni, E. *IEEE Trans. Device Mater. Reliab.* **2008**, *8*, 332−343.

(8) Chowdhury, S.; Swenson, B. L.; Wong, M. H.; Mishra, U. K. *Semicond. Sci. Technol.* **2013**, *28*, 074014.

(9) Dasgupta, S.; Nidhi; Raman, A.; Speck, J. S.; Mishra, U. K. *IEEE Electron Device Lett.* **2011**, *32*, 1212−1214.

(10) Ye, H.; Wicks, G. W.; Fauchet, P. M. *Appl. Phys. Lett.* **1999**, *74*, 711.

(11) Ye, H.; Wicks, G. W.; Fauchet, P. M. *Appl. Phys. Lett.* **2000**, *77*, 1185−1187.

(12) Stanton, N. M.; Kent, A. J.; Akimov, A. V.; Hawker, P.; Cheng, T. S.; Foxon, C. T. *J. Appl. Phys.* **2001**, *89*, 973−979.

(13) Balkan, N.; Arikan, M. C.; Gokden, S.; Tilak, V.; Schaff, B.; Shealy, R. J. *J. Phys.: Condens. Matter* **2002**, *14*, 3457.

(14) Wang, K.; Simon, J.; Goel, N.; Jena, D. *Appl. Phys. Lett.* **2006**, *88*, 022103.

(15) Wu, S.; Geiser, P.; Jun, J.; Karpinski, J.; Wang, D.; Sobolewski, R. *J. Appl. Phys.* **2007**, *101*, 043701.

(16) Tripathy, S. K.; Xu, G.; Mu, X.; Ding, Y. J.; Wang, K.; Cao, Y.; Jena, D.; Khurgin, J. B. *Appl. Phys. Lett.* **2008**, *92*, 013513.

(17) Suntrup, D. J.; Gupta, G.; Li, H.; Keller, S.; Mishra, U. K. *Appl. Phys. Lett.* **2014**, *105*, 263506.

(18) Verzellesi, G.; Saguatti, D.; Meneghini, M.; Bertazzi, F.; Goano, M.; Meneghesso, G.; Zanoni, E. *J. Appl. Phys.* **2013**, *114*, 071101.

(19) Kioupakis, E.; Rinke, P.; Delaney, K. T.; Van de Walle, C. G. *Appl. Phys. Lett.* **2011**, *98*, 161107.

(20) Kioupakis, E.; Steiauf, D.; Rinke, P.; Delaney, K. T.; Van de Walle, C. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *92*, 035207.

(21) Mustafa, J. I.; Bernardi, M.; Neaton, J. B.; Louie, S. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *94*, 155105.

(22) Zhou, J.-J.; Bernardi, M. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *94*, 201201.

(23) Bernardi, M.; Vigil-Fowler, D.; Lischner, J.; Neaton, J. B.; Louie, S. G. *Phys. Rev. Lett.* **2014**, *112*, 257402.

(24) Bernardi, M.; Vigil-Fowler, D.; Ong, C. S.; Neaton, J. B.; Louie, S. G. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 5291−5296.

(25) Bernardi, M.; Mustafa, J.; Neaton, J. B.; Louie, S. G. *Nat. Commun.* **2015**, *6*, 7044.

(26) Palummo, M.; Bernardi, M.; Grossman, J. C. *Nano Lett.* **2015**, *15*, 2794−2800.

(27) Fröhlich, H. *Adv. Phys.* **1954**, *3*, 325−361.

(28) Ridley, B. K. *Quantum Processes in Semiconductors*; Oxford University Press, 2013.

(29) Martin, R. M. *Electronic Structrure: Basic Theory and Practial Methods*; Cambridge University Press, 2008.

(30) Baroni, S.; de Gironcoli, S.; Dal Corso, A.; Giannozzi, P. *Rev. Mod. Phys.* **2001**, *73*, 515−562.

(31) Verdi, C.; Giustino, F. *Phys. Rev. Lett.* **2015**, *115*, 176401.

(32) Pankove, J. *RCA Rev.* **1975**, *36*, 163−176.

(33) Bougrov, V.; Levinshtein, M.; Rumyantsev, S.; Zubrilov, A. In *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe*; Levinshtein, M. E., Rumyantsev, S. L., Shur, M. S., Eds.; John Wiley & Sons: New York, 2001; pp Chapter 1, pp 1−30.

(34) Ashcroft, N.; Mermin, N. *Solid State Physics*; Holt, Rinehart, and Winston, 1976.

(35) Hardy, M. T.; Feezell, D. F.; DenBaars, S. P.; Nakamura, S. *Mater. Today* **2011**, *14*, 408−415.

(36) Yates, J. R.; Wang, X.; Vanderbilt, D.; Souza, I. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *75*, 195121.

(37) Mostofi, A. A.; Yates, J. R.; Pizzi, G.; Lee, Y.-S.; Souza, I.; Vanderbilt, D.; Marzari, N. *Comput. Phys. Commun.* **2014**, *185*, 2309−2310.

(38) Tanaka, S.; Zhao, Y.; Koslow, I.; Pan, C.-C.; Chen, H.-T.; Sonoda, J.; DenBaars, S.; Nakamura, S. *Electron. Lett.* **2011**, *47*, 335−336.

(39) Gardner, N. F.; Müller, G. O.; Shen, Y. C.; Chen, G.; Watanabe, S.; Götz, W.; Krames, M. R. *Appl. Phys. Lett.* **2007**, *91*, 243506.

(40) Zhang, J.; Zhang, D.; Ye, D.; Xu, C.; Huang, M. In *Optoelectronics - Materials and Devices*; Pyshkin, S. L., Ballato, J., Eds.; Intech, 2015.

(41) Hader, J.; Moloney, J. V.; Koch, S. W. *Appl. Phys. Lett.* **2011**, *99*, 181127.

(42) Ni, X.; Li, X.; Lee, J.; Liu, S.; Avrutin, V.; Özgür, Ü.; Morkoç, H.; Matulionis, A. *J. Appl. Phys.* **2010**, *108*, 033112.

(43) Giannozzi, P.; et al. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(44) Perdew, J. P.; Zunger, A. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1981**, *23*, 5048.

(45) Troullier, N.; Martins, J. L. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *43*, 1993.

(46) Louie, S. G.; Froyen, S.; Cohen, M. L. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1982**, *26*, 1738−1742.

(47) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. *Rev. Mod. Phys.* **2012**, *84*, 1419−1475.

(48) Mostofi, A. A.; Yates, J. R.; Lee, Y.-S.; Souza, I.; Vanderbilt, D.; Marzari, N. *Comput. Phys. Commun.* **2008**, *178*, 685−699.

(49) Hybertsen, M. S.; Louie, S. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5390−5413.

(50) Rubio, A.; Corkill, J. L.; Cohen, M. L.; Shirley, E. L.; Louie, S. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *48*, 11810−11816.

(51) Giustino, F.; Cohen, M. L.; Louie, S. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *76*, 165108.

(52) Vogl, P. *Phys. Rev. B* **1976**, *13*, 694−704.

(53) Noffsinger, J.; Giustino, F.; Malone, B. D.; Park, C.-H.; Louie, S. G.; Cohen, M. L. *Comput. Phys. Commun.* **2010**, *181*, 2140−2148.