Fabrication and Characterization of GaN/AlN Resonant Tunneling Diodes

W-D. Zhang[1], T. A. Growden[2], E. R. Brown[1], P. R. Berger[2], D. F. Storm[3], and D. J. Meyer[3]

Abstract This chapter reviews our recent efforts on growth, fabrication and characterization of GaN/AlN resonant tunneling diodes (RTDs). Working GaN/AlN RTDs were successfully demonstrated, and they could function well under the flux of very high current densities (e. g.~431 kA/cm^2) without thermal breakdown. The high-speed nature of these devices was confirmed through switching experiments, achieving a 10-90% switching time of ≈55 ps. A f_{max} calculation shows a small-signal oscillation with frequency up to 165 GHz is possible. Unlike InGaAs/AlAs RTDs, the peak to valley current ratios (PVCRs) of GaN/AlN RTDs remain ~1.5. Through computer modeling, temperature measurements, and material diagnosis, we reveal that there could be stronger inelastic scattering processes contributing to the valley current other than the coherent tunneling in the GaN/AlN RTDs. The possible inelastic mechanisms include optical phonons, interface roughness and dislocations. Thus, the growth of high quality GaN/AlN heterostructures as well as the evolution of bulk GaN substrates are critical for getting better performance devices. Finally, unipolar electroluminescence, without the presence of p-type doping, was observed in GaN/AlN RTDs. The interband tunneling process, which generates holes for the optical recombination, is likely due to the strong electric fields originated from the polarization effects native to wurtzite heterostructures.

---

[1] Departments of Physics and Electrical Engineering Wright State University, Dayton, OH 45435, USA. Emails: wzzhang@fastmail.fm; elliott.brown@wright.edu
[2] Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210, USA.
[3] U.S. Naval Research Laboratory, Washington, DC 20375, USA

Table of Contents

I. Introduction and Summary 3
II. Material Growth and Fabrication 4
  - **II.A. GaN/AlN RTDs** 4
  - **II.B. InGaAs/AlAs RTDs** 5

III. Characterization 6
  - **III.A Current-Voltage Curves** 6
    - **III.A.1 GaN RTDs at Room Temperature** 6
    - **III.A.2 GaN RTDs at Cryogenic Temperatures and high temperatures** 9
    - **III.A.3 InGaAs RTDs** 10
    - **III.A.4 High Current Density GaN RTDs** 11

IV. High-Speed Characterization 12
  - **IV.A. Switching** 12
    - **IV.A.1. Switching Methods** 12
    - **IV.A.2. Qualification with InGaAs RTDs** 14
    - **IV.A.3. GaN RTD Switching** 15
    - **IV.A.4. Effect of GaN Growth Methods on Switching Speed** 18
  - **IV.B. GaN RTD Oscillations** 19
  - **IV. C. A SPICE MODEL for GaN RTDs** 20

V. Electroluminescence 21
  - **V.A. GaN RTDs** 21
  - **V.B. InGaAs RTDs** 25
  - **V.C. Explanation for EL in GaN RTDs** 26
  - **V.D. Explanation for EL of InGaAs RTDs** 27
  - **V.E. Comparison of EL in GaN and InGaAs RTDs** 29
  - **V.F. Estimate of Quantum Efficiency in GaN RTDs** 31

VI. Conclusion 33
VII. Acknowledgements 34
VIII. References 34

I. **Introduction and Summary**

Electron tunneling through potential-energy barriers in semiconductor heterostructures is a fast process as its transit time can be 100 fs, or less. In double-barrier resonant tunneling diode (RTD) structures, such tunneling can support high-speed, negative differential resistance (NDR), as first described in the pioneering work of Tsu and Esaki [1] and [2]. The NDR generally occurs under applied bias as the lowest occupied (quasibound) energy level in the heterostructure drops in energy below the occupied states on the emitter side. This is the famous resonant-tunneling process, and its "N-type" I-V characteristic can be utilized for high frequency self-oscillation and fast switching. Among tunneling devices, the double-barrier resonant tunneling diode (DBRTD) is the most extensively studied NDR device. Waveguide-based RTD oscillations up to 712 GHz were demonstrated in the early 1990s [3], and then extended up to 1.1 THz by planar-circuit resonator techniques [4]. More recently the oscillation frequency has been extended to ~1.5 THz and beyond with efficient antenna coupling and electronic frequency tuning [5-7], providing a rejuvenated interest in resonant-tunneling devices.

This chapter summarizes our own efforts in the past few years on high-speed double-barrier RTD devices made from GaN/AlN heterostructures, and using InGaAs/AlAs RTDs as a benchmark. The former is relatively new and did not provide stable, room-temperature NDR until our recent effort was carried out [8]. The latter is much older and still serves as the gold standard in high-speed RTDs after ~30 years of development. Our major achievements are briefly listed as follows. First, we applied plasma-assisted molecular-beam epitaxy on bulk GaN substrates to grow GaN/AlN heterostructures having very high material quality with abrupt hetero-interfaces and low defect density. We then developed a fabrication procedure using state-of-the art dry etching and ohmic-contact technology to produce mesa-isolated device on the micron size scale. The resulting GaN/AlN RTDs displayed stable, reproducible NDR at room temperature with a peak-to-valley current ratio (PVCR) of ≈ 1.5 [8-11]. This was after a long period of study by other groups which showed either no NDR at room temperature, or NDR was likely associated with electron trapping, not resonant tunneling [12-24]. We then demonstrated the high-speed nature of our devices through NDR switching experiments, achieving a 10-90% peak-to-valley switching time of ≈55 ps. In parallel, our benchmark InGaAs/AlAs RTDs, produced by similar fabrication procedures, displayed a room-temperature PVCR of ≈10 and a switching time of ≈ 22 ps, both consistent with the expected performance of these devices. Its 10-90% switching time was measured to be ~22 ps [25].

In addition, we discovered bright near-UV electroluminescence (EL) in the new GaN/AlN RTD devices in spite of them being unipolar, n-doped everywhere in the structure [26]. Through spectral measurements, the near-UV emission was found to be centered near the GaN band-gap wavelength around 365 nm. Furthermore, recent noise measurements showed that the transport displayed normal shot noise except for a suppression effect associated with the resonant tunneling [27]. These results combined with detailed quantum transport computation suggested that the near-UV emission was by cross-gap radiative recombination between electrons accumulated on the emitter side of the device, and holes created in the same region by interband tunneling, which is enabled in vertical GaN heterostructures by the huge polarization fields at the GaN/AlN heterointerfaces [26].

The same mechanism for light emission in GaN-based RTDs was then observed in an $In_{0.53}Ga_{0.47}As$/AlAs DBRTD but at the $In_{0.53}Ga_{0.47}As$ band-gap wavelength around 1650 nm [28]. This discovery strongly suggests that the resonant- and interband- co-tunneling of electrons is a universal feature of unipolar DBRTDs and, remarkably, one that has never been reported in the vast literature of resonant-tunneling diodes over the past 40^{+} years.

This chapter is organized as follows: section II presents material growth and fabrication; Section III discusses dc properties of both GaN/AlN RTDs and InGaAs/AlAs RTDs; section IV discusses switching time measurements of both GaN/AlN and InGaAs/AlAs RTDs; section V discusses electroluminescence of both GaN/AlN and InGaAs/AlAs RTDs; finally, section VI is conclusion.

II. Material Growth and Fabrication

II.A. GaN/AlN RTDs

The samples were synthesized via plasma-assisted molecular-beam epitaxy (PAMBE) at 860°C on freestanding Ga-polar semi-insulating GaN substrates grown separately using hydride vapor phase epitaxy (HVPE) [Kyma Technologies, Inc.]. The substrates have low dislocation densities of approximately 10^{6} cm^{-2}. The substrate wafers were cleaned using an aggressive wet chemical etch prior to loading in the ultrahigh vacuum MBE system [29]. Once loaded into the ultra-high vacuum environment, the wafers were de-gassed for 30 minutes at 600°C and transferred into the MBE deposition chamber. The device layers were grown continuously without interrupts at constant temperature and at a constant growth rate of 3 nm/minute. The layer structure and

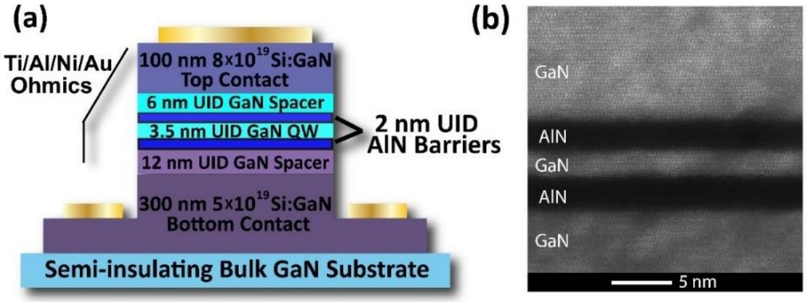

(a)
Ti/Al/Ni/Au
Ohmics
100 nm 8×10^19 Si:GaN
Top Contact
6 nm UID GaN Spacer
3.5 nm UID GaN QW
2 nm UID
AlN Barriers
12 nm UID GaN Spacer
300 nm 5×10^19 Si:GaN
Bottom Contact
Semi-insulating Bulk GaN Substrate
(b)
GaN
AlN
GaN
AlN
GaN
5 nm

Fig. 1. (a)The stack structure of GaN RTDs, (b) The TEM figure showing sharp transition between layers.

doping profile are shown in Figure 1 (a). A TEM image in Figure 1 (b) displays abrupt interfaces between the heterostructure layers.

The GaN/AlN RTD devices were fabricated using a 5-level mask set that consisted of the following steps: (i) top contact/mesa definition, (ii) bottom contact definition, (iii) device isolation, (iv) via creation, and (v) RF pad definition. Both the top and bottom ohmic contacts were a Ti/Al/Ni/Au stack. The isolation was done with a patterned PECVD-SiO_2 top layer, and the via holes were dry-etched with a CF4 plasma.

II.B. InGaAs/AlAs RTDs

The benchmark InGaAs/AlAs double-barrier RTD device under test (DUT) was grown by molecular beam epitaxy as an $In_{0.53}Ga_{0.47}As$/AlAs heterostructure on a semi-insulating InP substrate with a layer structure and doping profile as shown in Figure 2. Its active region is comprised of two unintentionally doped (UID) AlAs barriers (thickness=2.4 nm) separated by an undoped $In_{0.53}Ga_{0.47}As$ quantum-well (width = 4.4 nm) layer, such that a quasibound level E_1 occurs in the quantum well at an energy of ≈0.40 eV above the $In_{0.53}Ga_{0.47}As$ conduction band edge under zero bias. This rather high confinement energy compared to typical DBRTDs means that a large bias is required to reach the condition of NDR, especially under forward bias (positive on top contact), because of the 100-nm low-doped spacer layer on the top side that depletes and supports a large voltage drop and electric field.

The InGaAs/AlAs RTDs were fabricated following the similar procedures of GaN/AlN devices. The contact pads are a Ge/Au based stack, while the RF pads are a Ti/Au stack, different from the GaN/AlN RTDs.

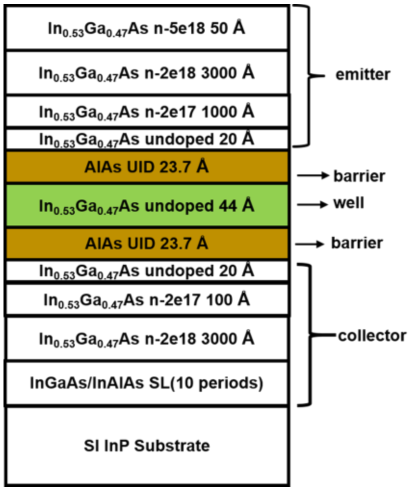

$In_{0.53}Ga_{0.47}As$ n-5e18 50 Å
$In_{0.53}Ga_{0.47}As$ n-2e18 3000 Å
$In_{0.53}Ga_{0.47}As$ n-2e17 1000 Å
$In_{0.53}Ga_{0.47}As$ undoped 20 Å
emitter
AlAs UID 23.7 Å
barrier
$In_{0.53}Ga_{0.47}As$ undoped 44 Å
well
AlAs UID 23.7 Å
barrier
$In_{0.53}Ga_{0.47}As$ undoped 20 Å
$In_{0.53}Ga_{0.47}As$ n-2e17 100 Å
$In_{0.53}Ga_{0.47}As$ n-2e18 3000 Å
InGaAs/InAlAs SL(10 periods)
collector
SI InP Substrate

Fig. 2. The stack structure of InGaAs RTDs.

III. Characterization

III.A Current-Voltage Curves

III.A.1 GaN RTDs at Room Temperature

Room temperature dc electrical characterization of the GaN RTD devices was carried out with a Keysight B1500A Parameter Analyzer [8]. Voltage-source bias was applied to the top contact via tungsten probes. To verify correct operation of a device initially, a full voltage sweep from -5V up to +6V and then back to -5V was employed (illustrated in Figure 3 (a) inset). Negative differential resistance was routinely observed around +4.1V across the sample on roughly 90% of the devices, regardless of size. The remaining 10% of the devices were usually flawed in some way as a result of process error and manifested as a short or an open. The consistency of the NDR onset voltage across devices of different sizes can be observed in Fig. 3(a). The peak current density, J_p, and peak-to-valley current ratio (PVCR) scale with device size. The I-V curves seen in Fig. 3(b) show very stable NDR over 1000 continuous sweeps, with only a very slight shift with time from Joule heating effects. This is contrary to any past reports, which generally detail a degradation of any observed NDR with subsequent sweeps [12-24]. A representative 7×10 µm device was measured and its I-V is shown in Fig. 3(b) where it was swept from +2.5V up to +5.5V and then from +5.5V down to +2.5V. This up-and-down

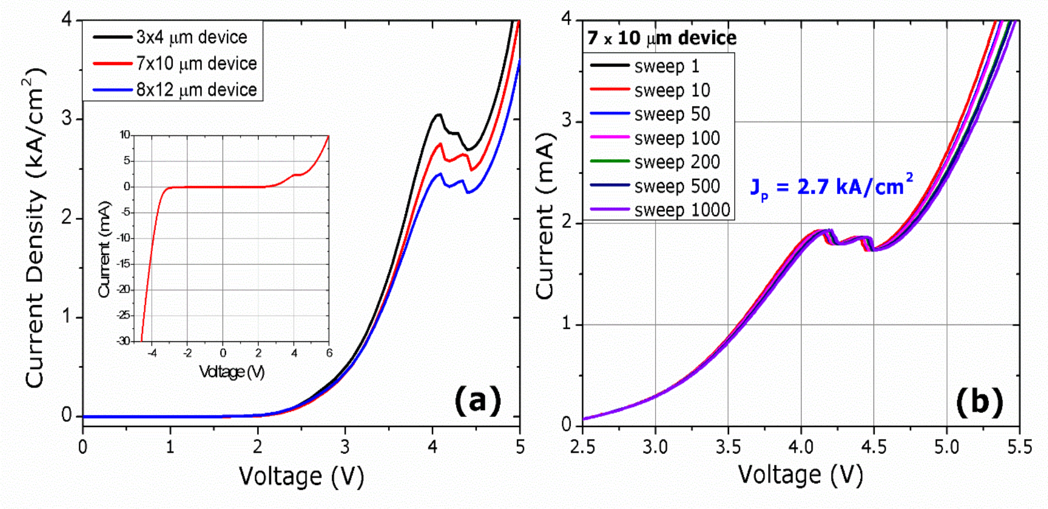


Fig. 3 (a) Current density-voltage characteristics of three different device sizes illustrating the similarities in both current density and NDR onset voltage. Inset displays a full sweep from -5V up to 6V and then from 6V back to -5V of a 7×10 µm device. (b) Current-voltage curve of a 7×10 µm device after increasing sweeps. One sweep was 2.5V-5.5V followed by 5.5V-2.5V (up-and-down sweep). Each sweep shown displays both the up and down curve in order to display the lack of hysteresis.

measurement was considered one "sweep". This approach was taken to display the lack of hysteresis which is usually observed in III-nitride tunneling devices during the downward part of the sweep. Additionally, this particular measurement illustrates the high degree of thermal stability available to III-nitride tunneling devices. The chair-like pattern in the NDR region is indicative of self-oscillation as the voltage-source bias presents negligible series resistance.

Traditionally, RTDs exhibit an antisymmetric I-V characteristic with respect to bias if the heterostructure and doping profiles are both symmetric about the center of the quantum well. As seen in Fig. 3(a) (inset displaying negative bias), our GaN/AlN RTD I-Vs are distinctly not antisymmetric partly because of asymmetry in the spacer-layer thickness and in the doping concentration in the contact layers. However, we believe a more important asymmetric effect is from the polarization fields present at the heterojunctions. The asymmetry is apparent in the band diagrams of Figures. 4 (a) and (b). The polarization charge is responsible for inducing a significant increase in the height of the first (top) barrier, as well as the creation of a thick space-charge immediately outside of it. Both of these features greatly affect the probability for tunneling, especially through the first quasibound state, E_1. To help minimize these effects, very high doping levels were utilized in the top contact layer. Additionally, an accumulation region is formed immediately following the second barrier.

Energy band diagrams for the early generation GaN/AlN RTDs calculated using numerical simulations indicate the existence of three quasibound states located within the RTD quantum well and one bound state located in the accumulation region [30]. The first state is located deep within the potential well where the barriers are much thicker, the tunneling probability resonance is very narrow, and thus it has little or no impact on the overall current contribution. The second state, E_2, has a much broader resonance and is responsible for most of the measured current and NDR observed in this study. Additionally, the bound state in the outer accumulation well, E_0,

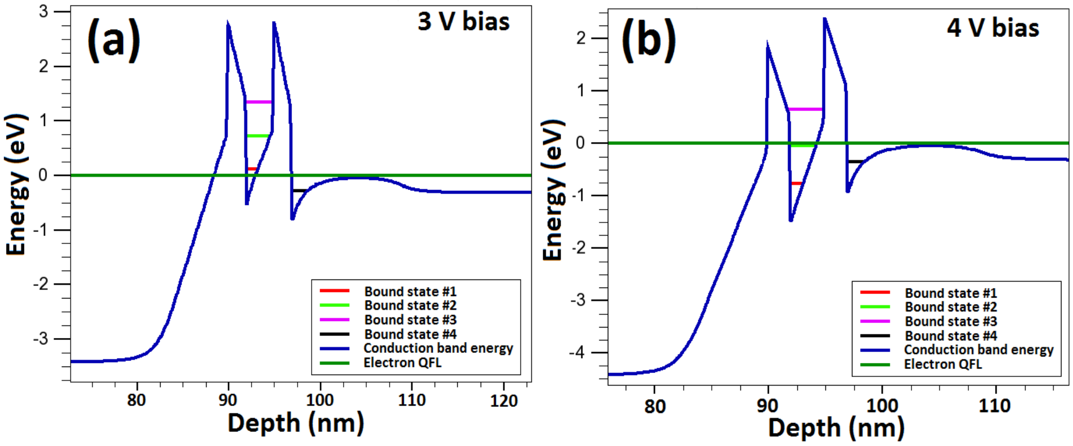


Fig. 4. (a) Energy band diagram at 3 volts. (b) Energy band diagram at 4 volts.

could be creating another, less pronounced resonance, contributing to the current significantly at the valley point and beyond. As the bias is increased past the 1st NDR region one would expect to see a 2nd NDR region resulting from the third quasibound state, E_3, this, however, is not the case. With increasing bias, as shown in Fig. 4(b), the first barrier eventually drops too low with respect to the second barrier, resulting in very poor confinement prior to the occurrence of resonant tunneling.

Despite a high degree of repeatability, further testing was carried out to verify the NDR observed in our devices was not the result of severe trapping effects. To this end, the probe station and the devices were isolated from all light sources and a baseline I-V was performed (in the same manner as described previously) and plotted in Fig. 5. After a period of time, a high power (1.15 W) white light was put directly over the device and another I-V sweep was measured. This same process was carried out for lower power (~150 mW) white, blue, green, and red-light sources. If there were any significant contribution from trapping effects, one would expect to observe a noticeable difference in the device operation, however, as seen in Fig. 5, the device remains stable and repeatable during the forward and backward voltage sweeps. The inset does show a very small difference when zoomed in on a point. The observed difference is only ~1 nA and is most likely a result of photocurrent

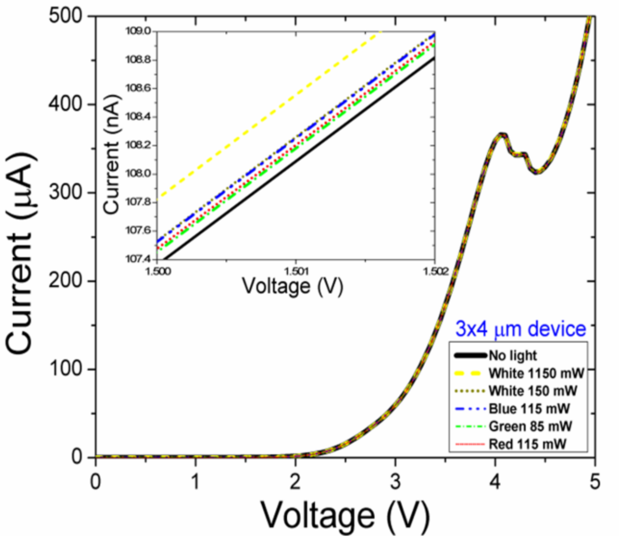


Fig. 5. I-V curve of a 3×4 μm device under different light exposures. Inset displays the current at +1.5V to illustrate that the only difference is related to a very small photocurrent.

III.A. 2 GaN RTDs at Cryogenic and High Temperatures

Low temperature measurements were carried out at 8, 50, 100, and 297 K for the device stack displayed in Fig. 6 (a) [10]. It was observed that the peak current I_p remains stable; however, the valley current I_v begins to rise at temperatures above 100 K (Fig. 6 (b)). Figure 6 (c) displays high temperature I-V measurements, from a different device, and again, I_p remained

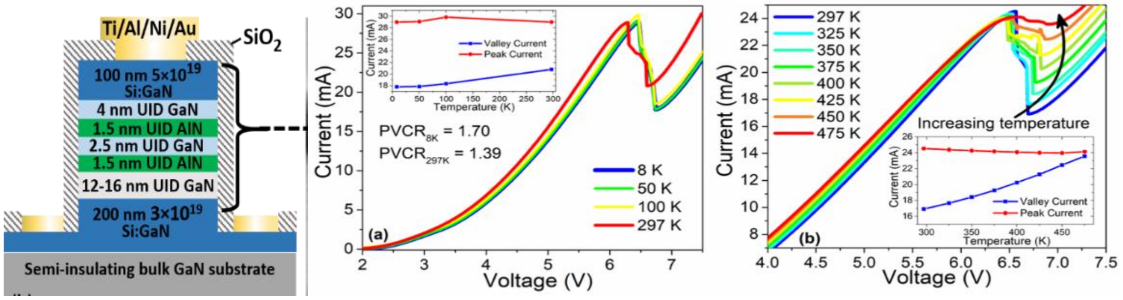
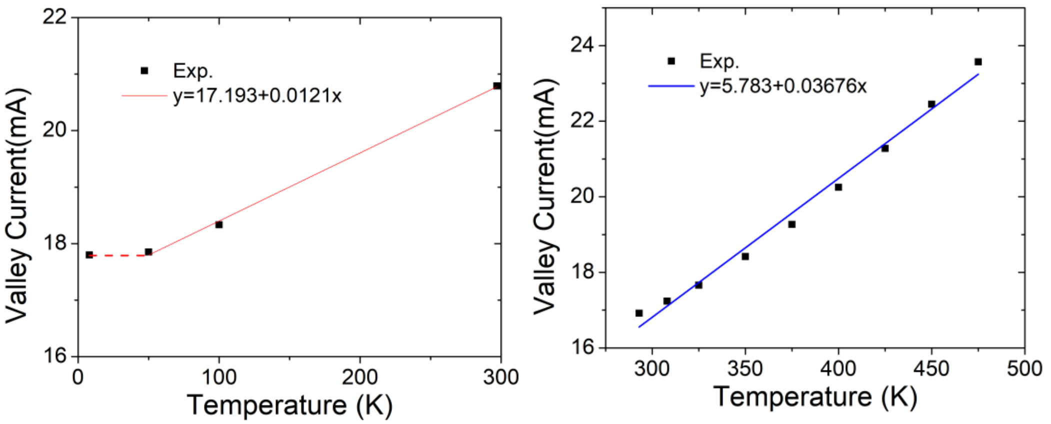


Fig. 6. (a) The device stack with the UID GaN collector-spacer region decreased to 4 nm, (b) Low temperature I-V sweeps illustrating a significant change in valley current. (c) High temperature I-V sweeps showing the disappearance of NDR as temperature increases. Insets: peak and valley current values displaying the temperature dependence of differing current mechanisms.



(a) **(b)**

Fig. 7. (a) and (b) are the fittings of valley currents vs temperatures.

stable, while I_v continued to increase. The heated probe station platen could not safely operate above 475 K, but if the I_v trend observed in Fig. 6 (c) continues, then the NDR would disappear above ≈485 K. This sets an upper limit on operational temperatures for these specific RTDs.

Based on a theoretical model on GaN/$Al_xGa_{1-x}N$ heterostructures given by Ridley [31] and an approximation that the incoherent phonon-assisted tunneling current is proportional to the inverse of scattering time constant $1/\tau$, a simple analysis of I_v at both low and high temperatures (Fig. 6) was performed, showing I_v can be written into a sum of two parts: one tends to be a constant as the temperature approaches zero and the other has a linear temperature dependence (Fig. 7). The former is likely contributed by LO phonon scattering via phonon emission, which has a very weak temperature dependence as the relevant temperature term is $\exp(-\hbar\omega_{LO}/k_BT)<<1$ because of $\hbar\omega_{LO}>>k_BT$ ($\hbar\omega_{LO}$-LO phonon energy ~87.3meV(1011K) for GaN and k_B-Boltzmann constant). Other scattering mechanisms such as interface roughness are possible candidates too. The latter likely comes from piezoelectric acoustic phonon scattering, which has a linear dependence on temperature.

Thus, a better understanding of the underlying thermal effects— primarily a mixing of LO phonon and piezoelectric acoustic phonon scattering — should help mitigate the high valley current and thus improve the PVCR.

III.A.3 InGaAs RTDs

The experimental I-V curve was measured with a Keithley 2400. As displayed in Figure 8, a local peak in current occurs at the start of the NDR region at 1.80 V, and a valley at the end of the NDR region at 2.65 V with a peak-to-valley current ratio (PVCR) of 9.2. This excellent PVCR is characteristic of InGaAs vs GaAs-based DBRTDs going back to their first demonstrations [32, 33]. The peak current density J_P is 3.5×10^4 A/cm^2 in this 9×6 μm^2 active-area device.

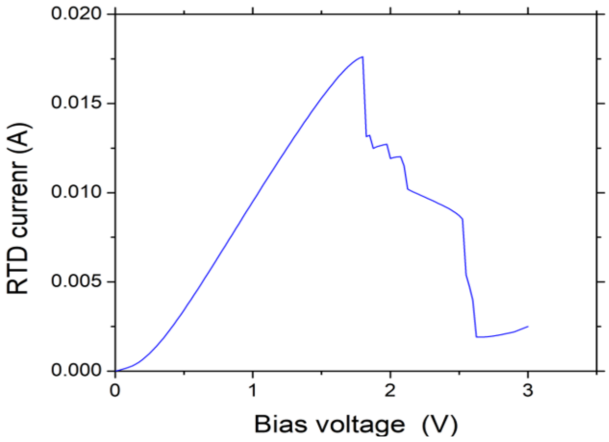


Fig. 8. I-V characteristic of an InGaAs/AlAs RTD.

III.A.4 High Current Density GaN RTDs

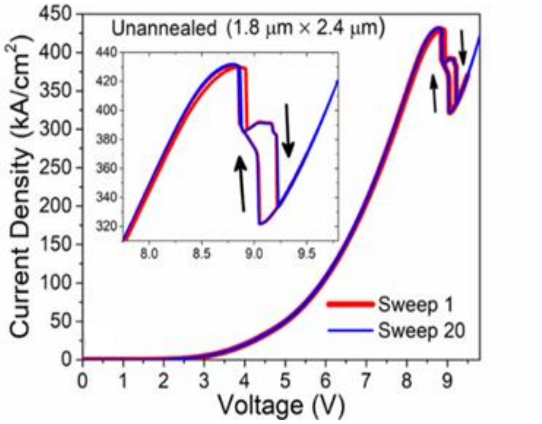


Figure 9. Bi-directional J-V curve displaying anomalous high current density from a small area device with slight hysteresis in the NDR region. Inset: magnified NDR region.

Another figure-of-merit achieved in the GaN/AlN RTDs is high peak current density, much larger than expected comparatively for a wide bandgap versus an InGaAs-based RTD. This is partly due to the excellent thermal conductivity of GaN (K~1.4W/cm-K). For example, the device displayed in Fig. 9 has an area of ~20 μm^2, thus the current density is at the level of ~140kA/cm^2. By shrinking the AlN barrier thickness to 1.5 nm (Fig. 6 (a)), a record of 431 kA/cm^2 is obtained in an RTD with a small area of ~4.3 μm^2. Despite the high current density, the I-V characteristic is still highly repeatable (Fig. 9).

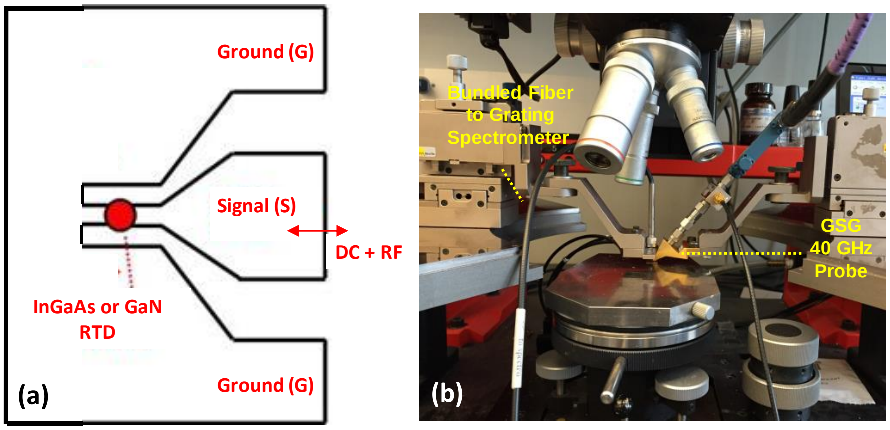

Ground (G)
Signal (S)
DC + RF
InGaAs or GaN
RTD
Ground (G)
(a)
Bundled Fiber
to Grating
Spectrometer
GSG
40 GHz
Probe
(b)

Fig. 10. (a) The RTD packed into a monolithic coplanar waveguide. (b) The probe station for RF characterization.

IV. High-Speed Characterization

IV.A. Switching

IV.A.1. Switching Methods

The measurement of switching events on a ps-time scale has a long history going back to the 1980s, and is usually done by one of two methods: [34] electrically with a fast oscilloscope (o-scope), or optically by ultrafast electro-optic sampling [35]. O-scopes are generally more *accurate* than electro-optic sampling, providing voltage or current measurements in MKSA units. But electro-optic sampling is inherently more precise at the ps time scale where modern o-scopes, even the sampling variety, are hampered by risetime limitations. Nevertheless, we have opted to carry out switching time measurements on GaN (and the benchmark InGaAs) RTDs using o-scopes primarily because of the packaging and coupling issues.

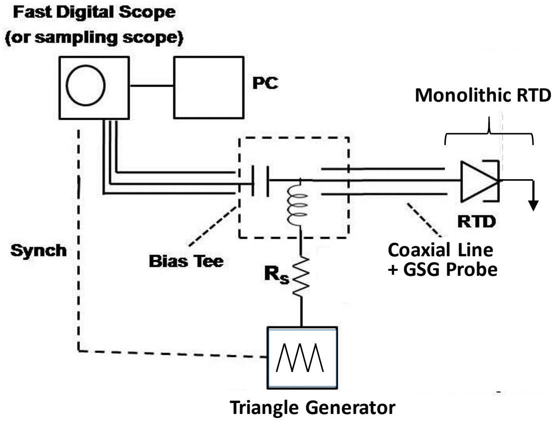

Fast Digital Scope
(or sampling scope)
PC
Monolithic RTD
Synch
Bias Tee
R_S
RTD
Coaxial Line
+ GSG Probe
Triangle Generator

Fig. 11. Experimental setup used to measure switching time.

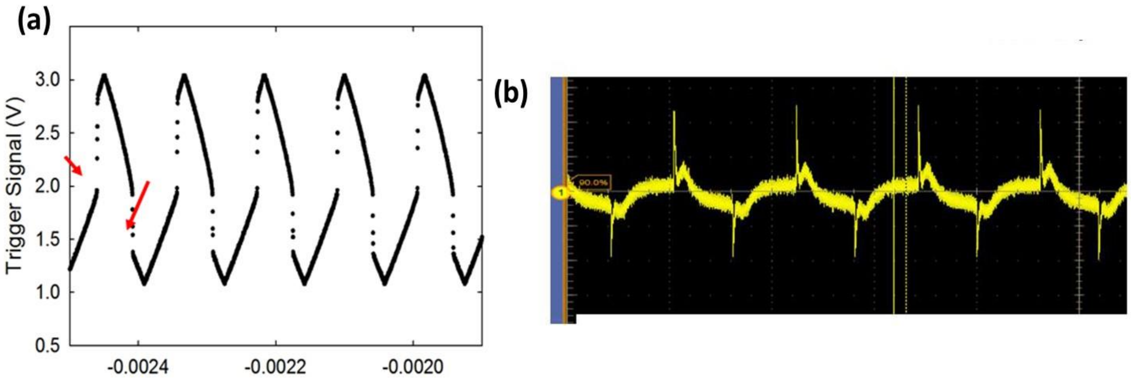


Fig. 12. (a) An example of a triangle signal (RF+DC port) on the standard o-scope. The red arrows show the switching events. The measurement was done for an InGaAs/AlAs RTD. (b) An example of a time-domain signal on the fast scope (RF port of the bias tee).

For switching time and all other high-speed characterization, we packaged the RTDs in the monolithic coplanar waveguide (CPW) structure shown in Fig. 10 (a). The device chip was mounted on the probe station shown in Fig. 10 (b) and contacted with a 200-um-pitch GSG probe operable in the frequency domain up to 40 GHz. As shown in Fig. 11, DC bias was applied with a 40-GHz bias tee, which was carefully chosen to have low series resistance (~1 Ohm) and short risetime (t ~ 5 ps). The bias tee was connected to both a triangle-wave generator and the high-speed o-scope, the first of which was a digital "real-time" variety (Tektronix MS073304DX). Since modern digital o-scopes are still limited to risetimes in the 10-100 ps range, we also employed a digital "sampling" unit (Tektronix DSA 8200). The latter has a risetime of approximately 7 ps but is more difficult to trigger.

Prior to conducting the high-speed measurement, a standard 50-Ohm o-scope was used to set the DC bias voltage and triangle amplitude correctly for balanced switching, meaning that peak-to-valley and valley-to-peak both occur within one cycle of the triangle. It is wise to keep the triangle amplitude just high enough to see the balanced switching, since much higher amplitude uses o-scope digital capacity with no additional information. The DC bias point is best adjusted to occur either in the positive differential resistance region just before the peak, or just beyond the valley points. Figure. 12 (a) shows a typical trace from a standard o-scope with

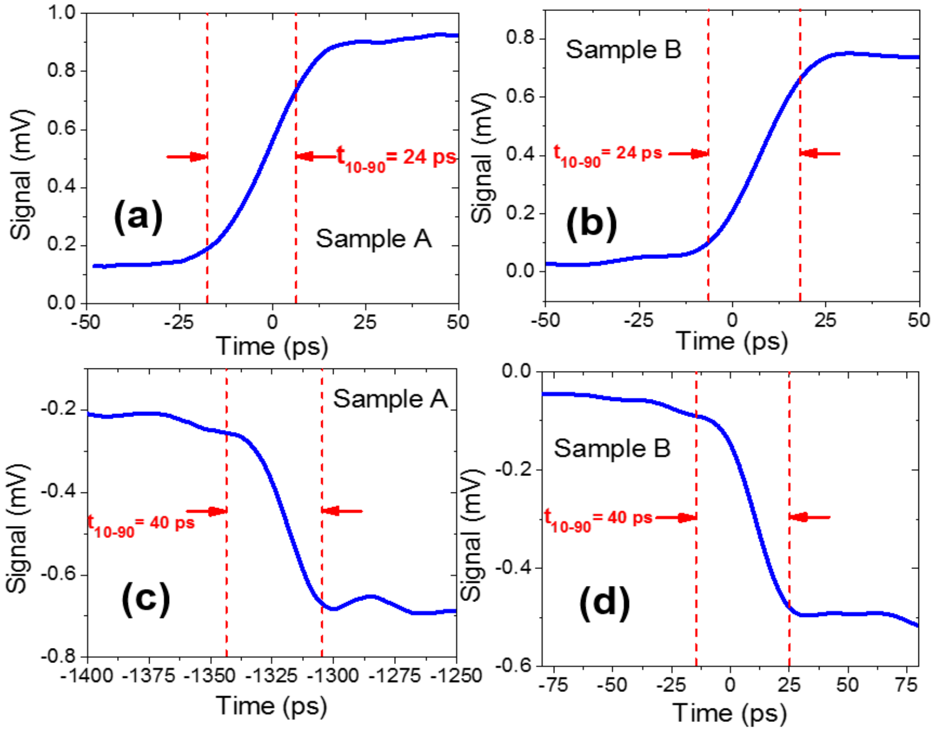


Fig. 13. (a) The high-speed switching time measurement shown for sample A was taken using a real-time Tektronix MS073304DX oscilloscope which has a 10-90 rise time of 13 ps. (b) The switching time measurement for sample B was measured in the same manner as sample A. (c) The valley-to-peak switching time data of sample A. (d) The valley-to-peak switching time data of sample B.

an 8.6 KHz triangle and after adjustment of the amplitude and DC bias are made. The replacement of o-scope by the high-speed real-time variety scope then yields the trace in Fig. 12 (b) where the switching events appear as spikes in voltage.

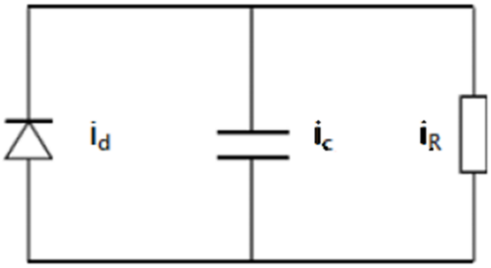
IV.A.2. Qualification with InGaAs RTDs

To qualify the switching technique on the ps time scale, we tested two InGaAs/AlAs RTDs on the same substrate, labeled A and B in Fig. 13 [25]. The resulting P-V switching times between the two samples are practically identical. Additionally, the V-P switching times are slower than the P-V times, as expected. This is caused by the fact that the valley region of RTDs is generally much broader (in voltage) than the peak region, so the available current (difference between the RTD resistive current and the load current at any given voltage) is smaller for P-V switching which makes the process slower, especially during the initial stage. However, the values measured here are not the intrinsic RTD switching times, but rather a combination of all significant rise-times in the circuit from the RTD, the bias tee, and the o-scope. The measured rise-time is then given by $t_{tot} = [(t_1)^2 + (t_2)^{2\,+}(t_3)^2]^{1/2}$. A rise-time of 13 ps is reported for the Tektronix MS073304DX oscilloscope [36], and the rise-time of our 40-GHz bias tees is ≈7 ps.

The risetime measured by the oscilloscope for the P-V switching time was determined to be approximately 24 ps for both samples A and B, as seen in Figs. 13 (a)-(d). This value was determined using the on-tool functionality of the oscilloscope as well as an analysis of the point-by-point data. After many measurements, there existed an average value of 25.3 ps and a standard deviation of 2.6 ps. This results in a coefficient of variation of approximately 10%. Given the effects of jitter in the triangle generator and scope trigger, physical noise and numerical errors (e.g., roundoff) in the oscilloscope, this is a reasonable variation. De-convolving the intrinsic RTD switching time yields a value of around 21.7 ps. This time would be much faster if the barriers were thinner, leading to a much higher speed index that tracks linearly with current density, but for the purpose of this qualification a robust InGaAs/AlAs RTD device design was chosen to enhance the reliability and measurability.

Calculation of the theoretical peak-to-valley switching time is possible if a few assumptions are made. The diode is biased at the peak tunneling voltage, V_P, through a load resistance, R_L, consistent with DC bistability. In this particular case, $R_L > \Delta V/\Delta I$, where $\Delta V/\Delta I$ is the NDR resistance, R_{NDR}. The values for the voltage space, ΔV, and current span, ΔI, can be calculated by taking the difference of the peak voltage and valley voltage ($V_P - V_V$) and the peak

Fig. 14. Circuit model used to model the RC-limited switching behavior through the NDR region.



current and the valley current ($I_P - I_V$) respectively. The diode capacitance is assumed to be constant and equal to the value at the V_P. A slight increase in the bias at this point causes a switch to the stable point (valley or beyond). The I-V curve in the NDR and valley regions is modeled as a parabolic form [37],

$$i_d = \frac{\Delta I}{\Delta V^2}(V - V_V)^2 + I_V \tag{1}$$

where i_d is the diode current as seen in the circuit model in Fig. 14. Applying Kirchoff's current law for a load resistance equal to ΔV/ΔI the 10%-90% RC switching time, t_R , can be determined by evaluating using the equation

$$t_R = \int_{V_P+0.1\Delta V}^{V_V-0.1\Delta V} \frac{CdV}{-\frac{\Delta I}{\Delta V}(V - V_V) - \frac{\Delta I}{\Delta V^2}(V - V_V)^2} \tag{2}$$

where C is the diode capacitance. Evaluation of this integral yields

$$t_R \approx 4.4\frac{\Delta V}{S} \tag{3}$$

where $S \equiv \Delta J/C'$, which is the speed index. The capacitance for the RTDs in this study were measured using a standard 4-point-probe, C-V measurement and consistently yielded values of approximately 1.8 fF/μm^2. Applying the experimental values to Eq. (3) resulted in a RC switching time, t_R, of 14.5 ps.

We also considered the effect of quasibound-state lifetime on the measured switching speed of the InGaAs/AlAs RTDs. In a previous analysis, the quasibound-state lifetimes τ_{QW} of various RTD designs with varying barrier thicknesses was calculated using a quasi-stationary-state, envelope-function method to compute the transmission probability, T(E), where E is the kinetic energy of electrons incident on the double barrier structure from the emitter side [38]. The lifetime was then calculated using the Breit-Wigner approximation, $\tau_{QW} \approx \hbar/(2\Gamma)$, where Γ is the full-width at half-maximum of the first resonant peak in T(E). By utilizing this data, it was determined that the first quasibound state lifetime for the RTDs considered in this study was equal to 5.5 ps. By taking the sum of the theoretical values for the RC switching time and the first quasibound state lifetime, a resulting intrinsic switching time for the RTD is 20 ps. When compared to the experimental value of 21.7 ps for these devices, there is only a slight difference (1.7 ps) which is well within any error due to assumptions, jitter, and oscilloscope physical noise and numerical errors. These results suggest the RC time is independent of the QW lifetime and that both contribute in an additive manner to the intrinsic switching time of an RTD.

IV.A.3. GaN RTD Switching

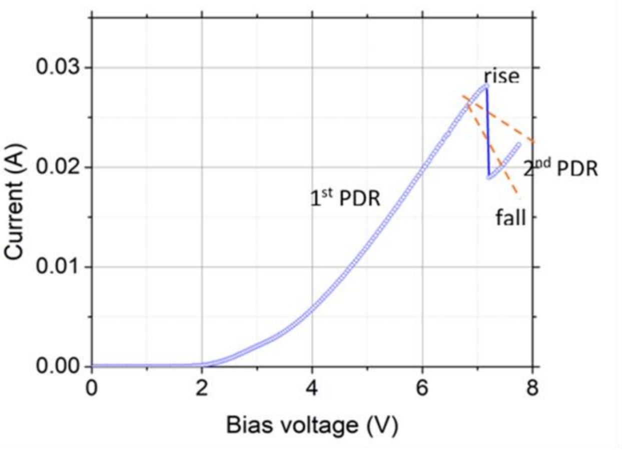
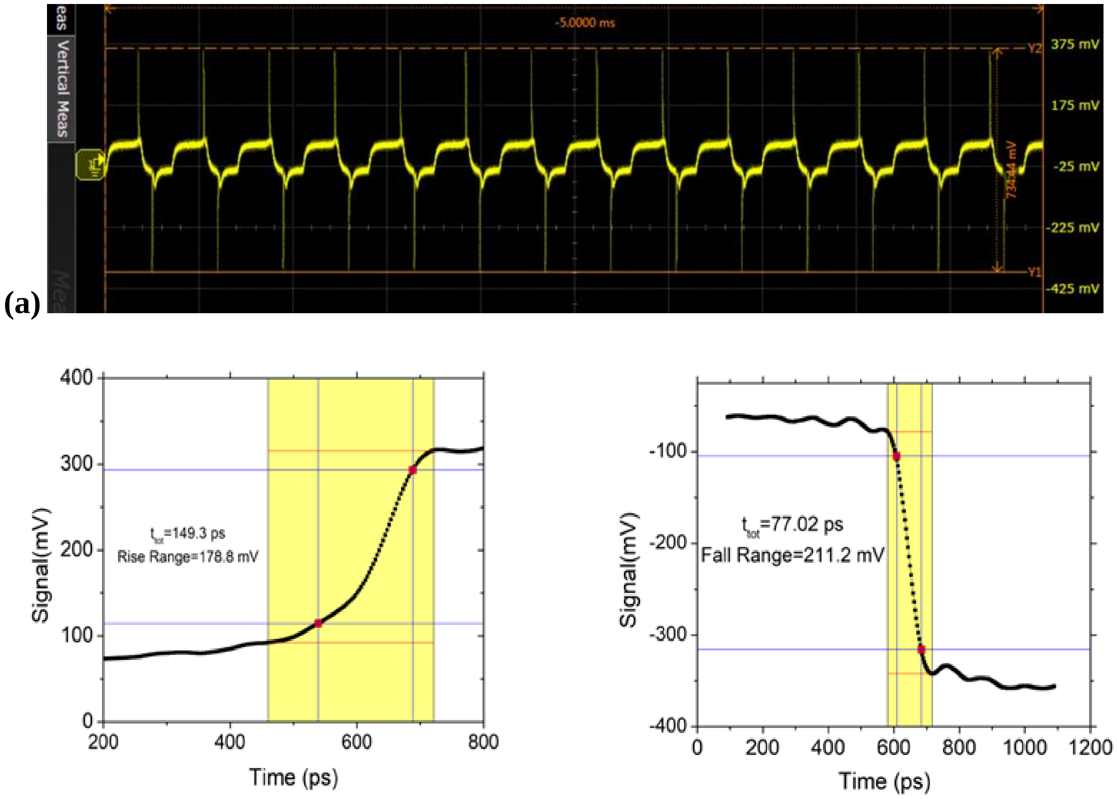


Fig. 15. The I-V characteristic of an GaN/AlN RTD.

Next we investigated the switching time constants of GaN/AlN RTDs using a similar experimental setup as above. An Anritsu bias tee was used. The DC+RF port of the bias tee was connected to the triangle signal generator while the RF port was connected to an Infiniium MSOS804A 8GHz oscilloscope. The Infiniium scope itself has an internal rise/fall time of t_2=53.8 ps [39], which along with the bias-tee risetime of ≈7 ps, needs to be deducted from the total 10-90% time t_{tot} measured by the scope.

The I-V characteristic of a GaN/AlN RTD diode under test is shown in Fig. 15. It has a peak voltage at 6.9 V, and a valley at 7.1V. Accordingly, the peak current is I_p=24.6 mA, and the valley current is I_v=17.5 mA, reaching a PVCR of ~1.41, which is much less than that of an InGaAs/AlAs RTD. In this case, a respectable PVCR is important, as small PVCRs don't trigger switching events, and according to equation (3), a large PVCR will increase the speed index, thereby shortening the switching time. The negative resistance at the NDR region is estimated R_D ~ -21 ohm. Hence any self-oscillation can be subdued simply with the 50-ohm impedance of the coaxial cable connected to the bias tee. We performed SPICE-model fitting following the procedure laid out later in Section IV. C. The full width of maximum, Γ_T, is roughly ~0.6 meV, and thus the related time constant is $\tau_T \sim \hbar/\Gamma_T$ =1.8 ps, indicating the inelastic scattering processes are significantly suppressed compared to the examples to be discussed in Sec. IV. C. This is beneficial to the quantum transport through the quasi-bound state, meaning the switching processes can be potentially very fast.

As shown in Figure 16, the 10%-90% rise time measured with the Infiniium scope was determined to be $t_{tot(rise)}$~149.3 ps, while the fall time was estimated to be $t_{tot(fall)}$~77.0 ps. Measurements on other devices from the same substrate yielded similar results. After the subtraction of the o-scope and bias-tee rise (fall) times through $\sqrt{t_{tot}^2 - t_2^2 - t_3^2}$, the actual rise time of the switching is t_r ~139.3 ps, but the fall time is only t_f ~55.1 ps. Noticeably, the fall time is ~2.53 times shorter than the rise time. This is opposite to the case in InGaAs RTDs, where the



(b) (c)

Fig. 16. (a) The switching events of a GaN RTD device. (b) The rise time measurement of the GaN/AlN RTD, (c) the fall time of the GaN/AlN RTD.

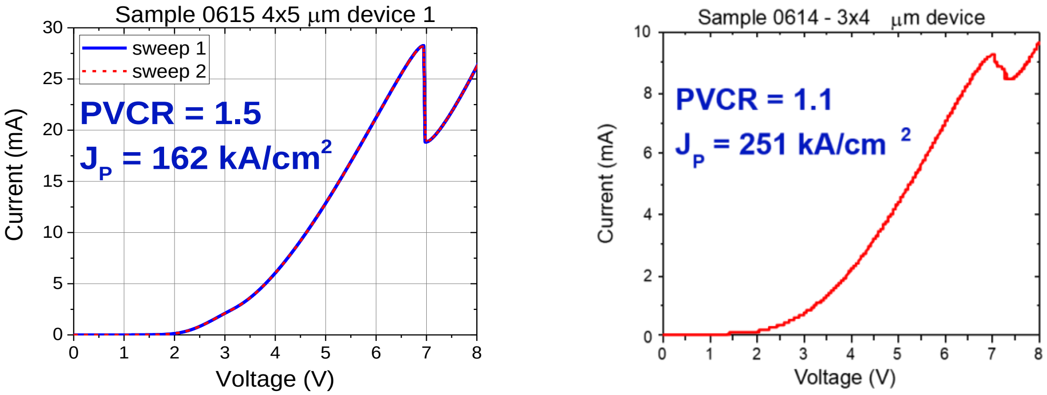
fall time is the shorter one of the two time-constants. Nevertheless, to the best of our knowledge, this is the first direct measurement of switching time constants ever done for the vertical transport in GaN/AlN/GaN heterostructure RTDs.

Then we estimated the capacitance of the GaN/AlN/GaN RTDs by evaluating the space charge vs bias voltage relationship with a Poisson solver provided by Silvaco Atlas [30]. The capacitance per unit area is roughly ~4.8 fF/μm^2. Given the device area of A~11 μm^2, the capacitance for the RTD is about C~52.8 fF. Solutions from other Poisson solvers yield similar results. According to equation (3), the peak-to-valley switching time is estimated to be t_R~4.9 ps, which is significantly less than the experimental values. A possible explanation for the discrepancy is that the calculation of t_R with Eq. (3) considers only the switching operation in the NDR region. However, switching events generally involve not only in the NDR but also the second positive differential region (PDR) as well as the first PDR. In particular, it may be the case as the voltage range of the valley region of the GaN RTD is very narrow (ΔV only ~0.2V). The inclusion of the PDRs can significantly slow down the switching action. This has been demonstrated in our large-signal-model numerical studies [40]. And, the rising and falling switching processes may come across different points of the second PDR, each seeing different load-line impedances (Fig. 15). This explains why the rising and falling switching time constants are different, and the falling is shorter because of the smaller impedance.

IV.A.4. Effect of GaN Growth Methods on Switching Speed

In addition to the DC and switching measurements, we also investigated the influence on the devices' function by changing the recipe of material growth. For example, our research reveals that varying the surface preparation of the freestanding GaN substrates prior to epitaxial growth can lead to substantially different performances of GaN/AlN RTD devices even though the heterostructure layers deposited are the same and the growth conditions are kept identical.

Devices on standard-clean substrates have higher PVCRs (~1.5), and faster switching time constants (Fig. 17 (a)); on the contrary, devices on substrates cleaned with Ga-deposition-



(a) (b)

Fig. 17. IV curves for devices on two GaN substrates prepared with different clean procedures.

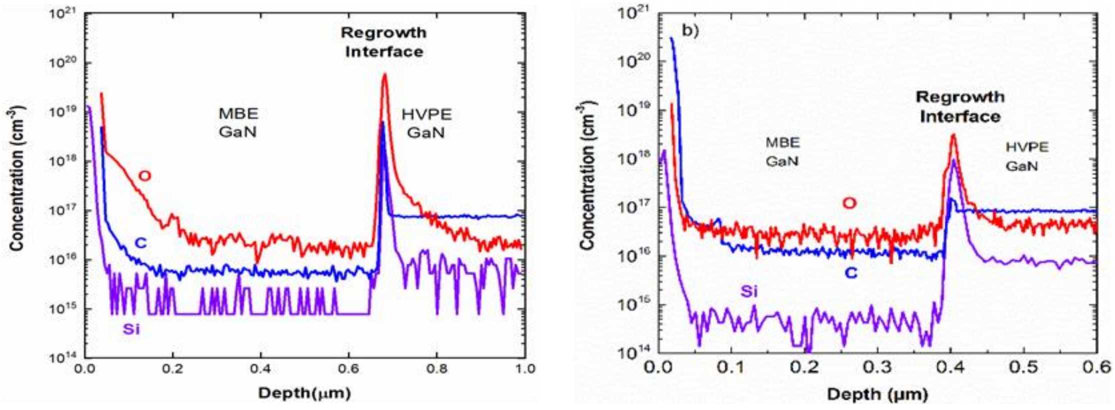
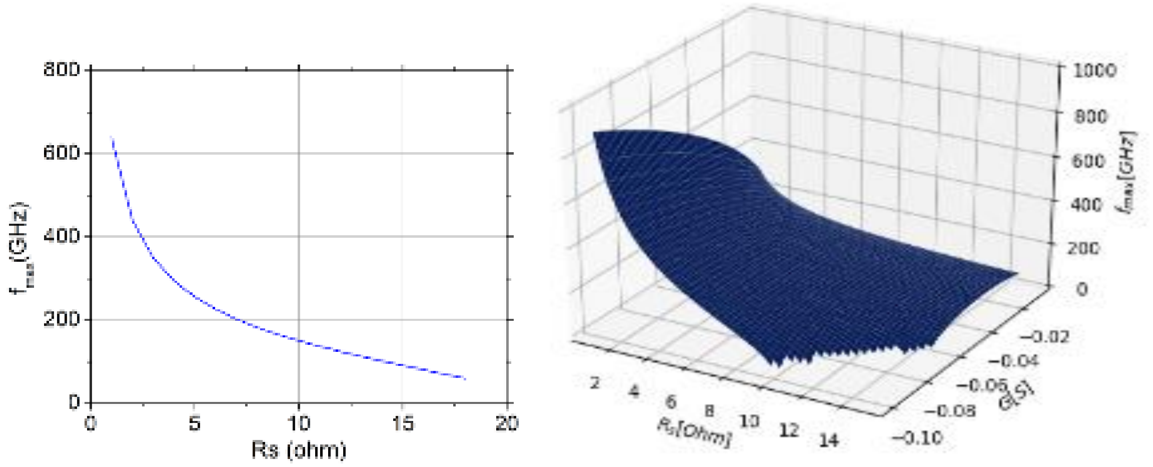


Fig. 18. SIMS on epilayers grown on two GaN substrates prepared with different clean procedures.



(a) **(b)**

Fig. 19. The estimation of f_{max}. (a) f_{max} vs. R_s, (b) f_{max} vs. R_s and G.

desorption tend to have smaller PVCRs (~1.1), and slower switching time constants, but allow for high current densities (Fig. 17 (b)). Also, the SIMS results plotted below support the subtle differences between the two heterostructures when their respective substrates were prepared using different cleaning methods (Fig. 18).

Descriptions of the surface preparation recipes can be found in [11, 41]. Clearly, continuous improvements of GaN/AlN heterostructure's quality are crucial for getting better performance devices.

IV.B. GaN RTD Oscillations

When biased into the NDR region, the GaN RTDs are capable of self-oscillation by the same mechanism that RTD oscillators demonstrated in the past going back 30^+ years. Unlike the RTD switching process, the start-up condition for such oscillations is simply that the real part of the terminal impedance (R_T) be negative, which is a small-signal condition. The frequency where R_T crosses from negative to positive is called the maximum frequency of oscillation, f_{max}. Knowing the capacitance C, the differential conductance (which is negative) G, and the device series resistance R_S, f_{max} is given by [42]

$$f_{max} = \frac{1}{2\pi C}\sqrt{-\frac{G}{R_S} - G^2} \tag{4}$$

where $G=-1/R_D$ is the conductance and R_S is the parasitic resistance. In the present case (Figs. 15, and 16), Rs is ~9.1 ohm from the contact resistance measurement $\sim 1.1\times10^{-7}$ Ohm-cm^2 and the device active area of ~11 μm^2. So f_{max} is estimated ~165 GHz. As shown in Figure 19, in principle, according to Eq. (4), a reduction of the parasitic resistance can boost the maximum frequency to an even higher range of ~600 GHz.

IV.C. A SPICE Model for GaN RTDs

Now we apply a SPICE model developed previously for InGaAs RTDs [43] to study the I-V characteristics of the GaN/AlN RTDs. The I-V curves can be described by the following equation,

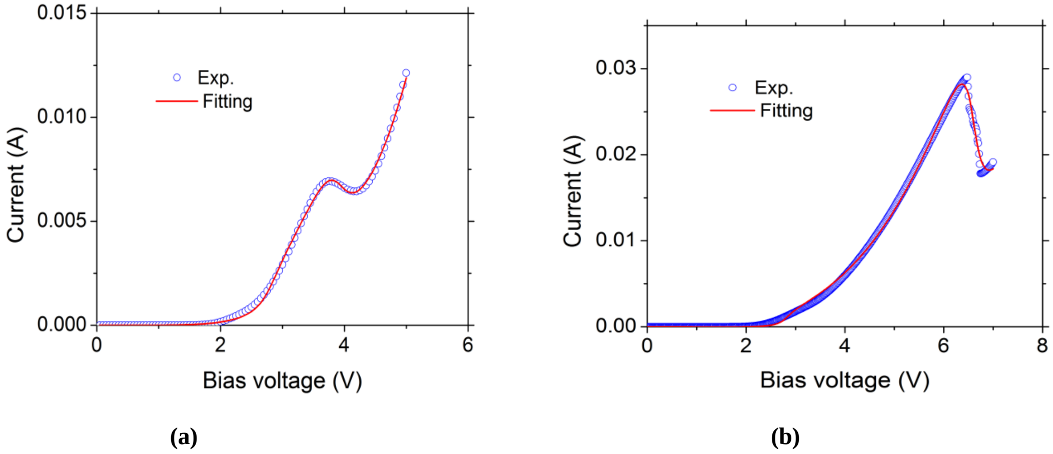
$$I = C_1 V^i \{\tan^{-1}[C_2(V - V_T)] - \tan^{-1}[C_2(V - V_N)]\} + C_3 V^j + C_4 V^k \qquad (5)$$

where C_1, C_2, C_3 and C_4 are numerical constants. V_T is defined as the threshold voltage where the second-order derivative d^2I/dV^2 has a local maximum. V_N is defined as the bias in the NDR region where the current's variation is the steepest. i, j and k are integer exponents ≥2. In the case of InGaAs/AlAs RTDs, the exponents were taken i=3, j=5 and k=3; these values are kept the same when modeling the GaN/AlN RTDs.

Equation (5) has proven to be able to generate excellent fittings on I-V curves of GaN/AlN RTDs. An example is shown in Fig. 20 (a), and the fitting parameters are $C_1=4.36\times10^{-5}$, $C_2=3.9$, $C_3=3.7\times10^{-6}$, $C_4=-3.1\times10^{-6}$, $V_T=2.8$ and $V_N=3.9$, respectively. Furthermore, the fitting enables us to extract the full width of maximum from

$$\Gamma_T \approx \frac{2E_F}{C_2(V_N - V_T)} \qquad (6)$$

Γ_T is the equal to $\Gamma_T = \Gamma_L + \Gamma_R + \Gamma_S$, where the first two terms are partial-width factors associated with the left and right barriers respectively, and the third term is a phenomenological parameter determined by inelastic scattering processes in the double-barrier structure. From the emitter's doping density ~$5\times10^{19}/\text{cm}^3$, the Fermi level is approximately E_F~0.25eV away from the conduction band. With these parameters, we estimate Γ_T ~115 meV for the diode shown in Fig.



(a) **(b)**

Fig. 20. The fitting of I-V curves with the SPICE model: (a) for a low current-density RTD, (b) for the RTD in Fig. 6 (b) at 8K.

20 (a). And the time constant is $\tau_T \sim \hbar / \Gamma_T \sim 5.7$ fs evaluated from the uncertainty principle. This is much shorter than the quantum lifetime expected from the half width of the resonant peaks that is in the order of a few ps. For example, typically $\Gamma_L + \Gamma_R$ is ~0.2 meV for an InGaAs/AlAs RTD. Figure 20 (b) shows the second example of fitting on the 8-Kelvin I-V curve in Fig. 6. The full width of maximum is reduced to $\Gamma_T \sim 22.5$ meV but remains significantly large. For comparison, the RTD device which displayed the shortest switching time constant (~55ps) has an estimated full width of maximum only $\Gamma_T \sim 0.6$ meV (Fig. 15).

Therefore, there could be stronger inelastic scattering processes contributing to electric current other than the coherent quantum tunneling in the GaN/AlN RTDs compared to the InGaAs/AlAs RTDs [42]. For the GaN/AlN RTDs, the possible inelastic scattering mechanisms include optical phonons, interface roughness and dislocations. The scatterings from the latter two can be the dominant ones at cryogenic temperatures [44]. Furthermore, they could have contributed to the plateau valley currents between ~8K and 50K shown in Fig. 6 (b).

V. Electroluminescence

V.A. GaN RTDs

Light emission was initially observed by eye through the probe-station microscope when the RTDs were biased beyond ~5 V. Strong violet light was observed coming from the mesa periphery, as displayed in Fig. 21, with fully repeatable NDR at room temperature [26]. The light emission was sufficiently bright that it could easily be measured by a commercial grating spectrometer coupled to the device through a bundled fiber probe placed in close proximity. The emission spectra of all three samples [Table I, Figs. 22(a)-(c)] exhibited a dominant peak centered near 360 nm that increases in intensity with increasing bias voltage and has full-width at

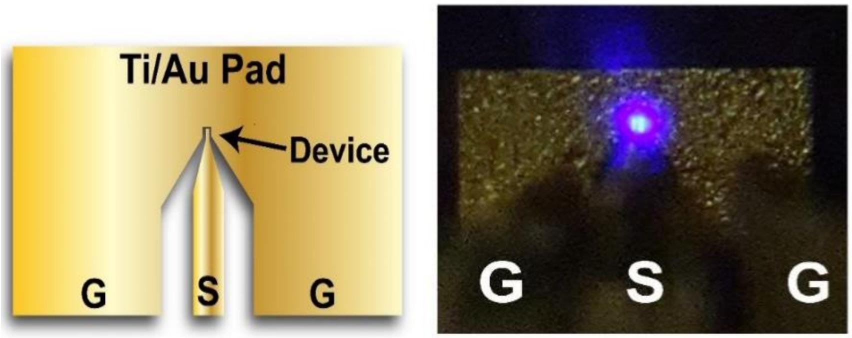


Fig. 21. A top-down image of the Ti/Au pad contact. Photograph of 7×10 μm^2 GaN RTD structure showing the three dc-coupled electrodes, and the RTD mesa device under test (DUT). Strong violet light was observed emitting from the RTD structures under bias, but this was found to be the long-wavelength tail of a much stronger near-UV emission around 360 nm.

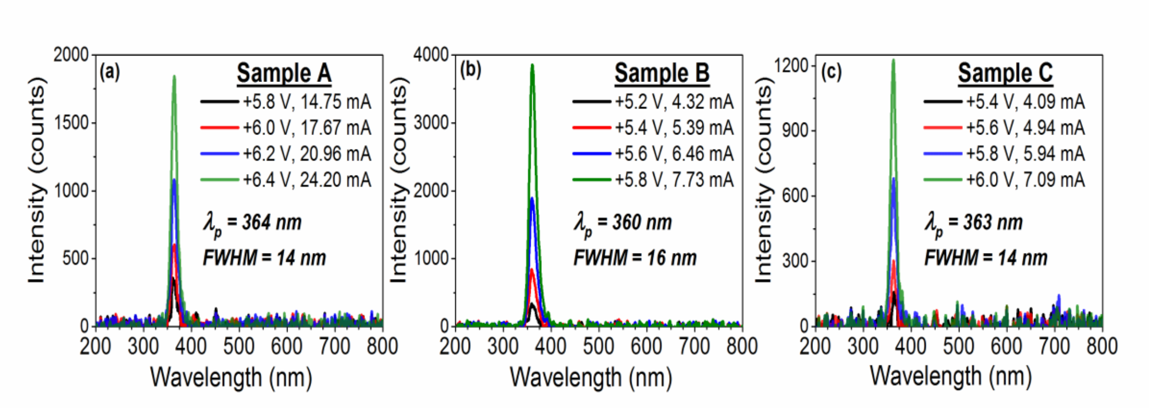


Fig. 22. The measured light spectrum emitted from representative 7×10 μm^2 devices as a function of voltage bias for (a) Sample A, (b) Sample B and (c) Sample C. The associated device current is also displayed. The increased wavelength under negative bias is a result of heavier doping in the collector contact layer, where recombination takes place (as opposed to the emitter contact layer under positive bias). All the measurements were conducted at room temperature.

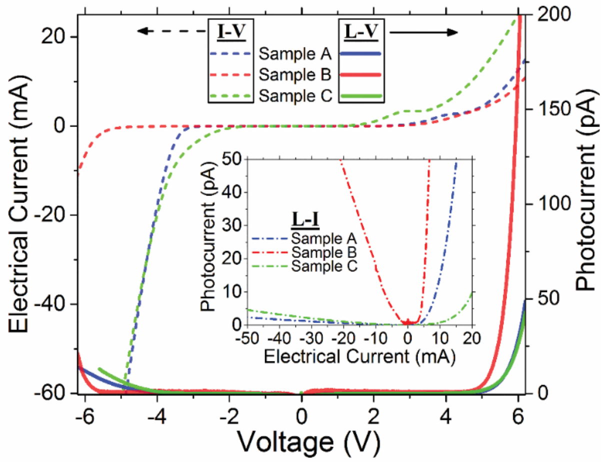


Fig. 23. The dc I-V and L-V curves from all three samples during both positive voltage bias and negative voltage bias. The inset displays the L-I curves for all three samples.

half-maximum (FWHM) values of 14 nm for Sample A, 16 nm for Sample B and 14 nm for Sample C. While under positive voltage bias, the FWHM remains ≤ 16 nm, even at the highest applied bias levels, with no significant spectral broadening. The 360-nm emission is attributed to cross-gap transitions because the wavelength corresponds closely to the 3.44 eV bandgap of GaN at room temperature. With increasing voltage bias, the emitted light remains quite optically pure in all three samples without any significant sub-bandgap emission, as reported in many other GaN light-emission results. However, the devices studied in this work often failed in the form of a short circuit as the bias was raised above a critical breakdown voltage (e.g., ≈7 V).

Table I   Device Structures and Parameters

Sample	A	B	C
Emitter Layer	300nm GaN 5E19/cm^3	300 nm GaN 5E19/cm^3	300 nm GaN 5E19/cm^3
Emitter Spacer	12-nm UID GaN	Digital AlGaN alloy (2.5 Å AlN/ 10 Å GaN)	12-nm UID GaN
Barriers	2 nm UID AlN	2 nm UID AlN	2 nm UID AlN
Quantum Well	3 nm UID GaN	3 nm UID GaN	3.5 nm UID GaN
Collector Spacer	6 nm UID GaN	6 nm UID GaN	6 nm UID GaN
Collector Layer	100 nm GaN 8E19/cm^3	100 nm GaN 8E19/cm^3	100 nm GaN 8E19/cm^3

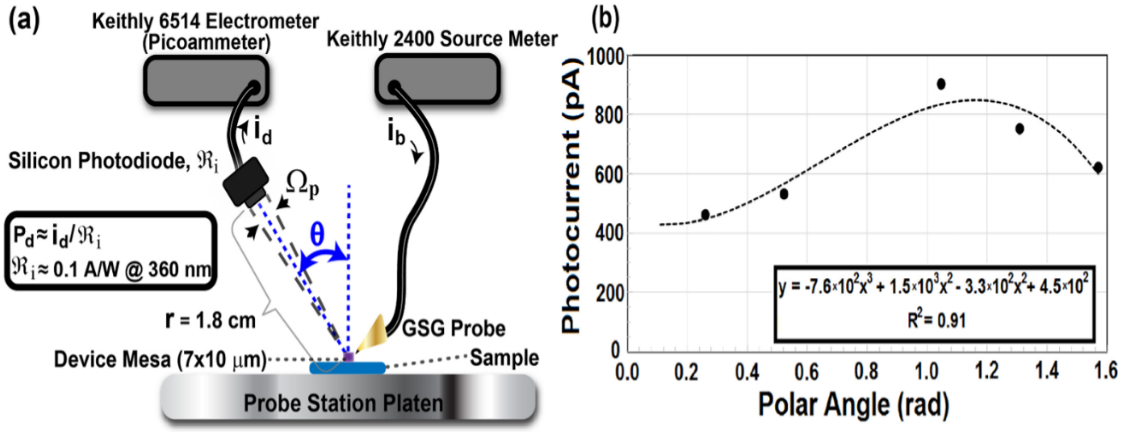


Fig. 24. (a) Experimental set-up used to measure the emitted power at various angles θ from the polar axis and into a solid-angle Ω_P defined by the silicon photodiode area and range r from the GaN unipolar-doped light emitter. (b) Photocurrent vs elevational angle obtained with the set-up in (a). The data points are shown as solid circles, and the cubic-polynomial curve-fit as a dashed line.

The room-temperature current-voltage (I-V), light intensity vs. voltage (L-V), and light intensity vs. current (L-I) characteristics for all three samples are displayed in Fig. 23. Although both the L-V and L-I curves exhibit a threshold effect, the L-I curves are distinct between the samples, whereas the L-V curves display a common threshold (~4.7 V). Above the threshold, the L-V curve displays an exponential increase of light emission vs. bias voltage.

The far-field intensity of the RT-LEDs was measured, despite the shadowing of the GSG probe pads atop the DUT. The far field was found to be significantly dependent on the elevational angle θ in Fig. 24 (a) but relatively independent of the azimuthal angle φ. This observation is consistent with symmetry considerations given that the radiating structure is a mesa with exposed sidewalls around the periphery. Optical measurements at five elevational angles for Sample B are shown in Fig. 24 (b), all at a range of 1.8 cm from the mesa. The resulting data points were fit with a cubic polynomial, and the best fit was $I(\theta) = -7.6\times10^{2}\theta^{3} + 1.5\times10^{3}\theta^{2} - 3.3\times10^{2}\theta + 450$. The total power was estimated into the upper hemisphere by a rectangle approximation $P_{tot} \approx \frac{2\pi}{\Re_I \Omega_p} \cdot \sum_{j=1}^{j_{max}} I(\theta_j) \cdot \sin\theta_j \cdot \Delta\theta$, where $\Re_I \approx 0.1$ A/W is the current responsivity of the photodiode at 360 nm, and Ω_p is the solid angle subtended by the photodiode with respect to an origin defined by the emitting diode 1.8 cm away, such that $\Omega_p \approx 0.010$ str. Setting $\Delta\theta = 1.0°$ (0.017 rad), we find $P_{tot} = 4.7\times10^{-6}$ W. The EQE, η_{ext}, was calculated with additional parameters I_B=18.8 mA and hν = 3.4 eV and led to the lower limit estimation of $\eta_{ext} \approx$ 0.0074%.

This η_{ext} value is well below the state-of-the-art values of ~50% for optimized bipolar-doped (p-n) GaN LEDs [45]. However, we emphasize that this value is conservative, considering emission into only the upper hemisphere and ignoring internal loss mechanisms such as total internal reflection. In addition, these devices are not designed to balance electron and hole currents such that the electron-hole radiative recombination reported is currently hole-limited. Even so, to the best of our knowledge, this value is higher than the values reported for any other unipolar-doped GaN emitter to date, such as the 10^{-6}% UV-value reported in reference Zimmler *et al* [46].

V.B. InGaAs RTDs

After the discovery of unipolar EL in GaN/AlN RTDs, we went forward to investigate the EL phenomena in InGaAs/AlAs RTDs [28]. Shown in Fig. 25 (a) (right vertical axis) is the photocurrent from the Ge diode as a function of RTD bias voltage (L-V curve). The photocurrent from the electrometer rises significantly above the noise floor at a bias voltage of ≈1.0 V and increases monotonically with higher voltage through the NDR region up to the valley point. Then there is a precipitous drop at the valley voltage followed by a slow increase above that. That is, the change in photocurrent in the NDR region is anticorrelated to the change in

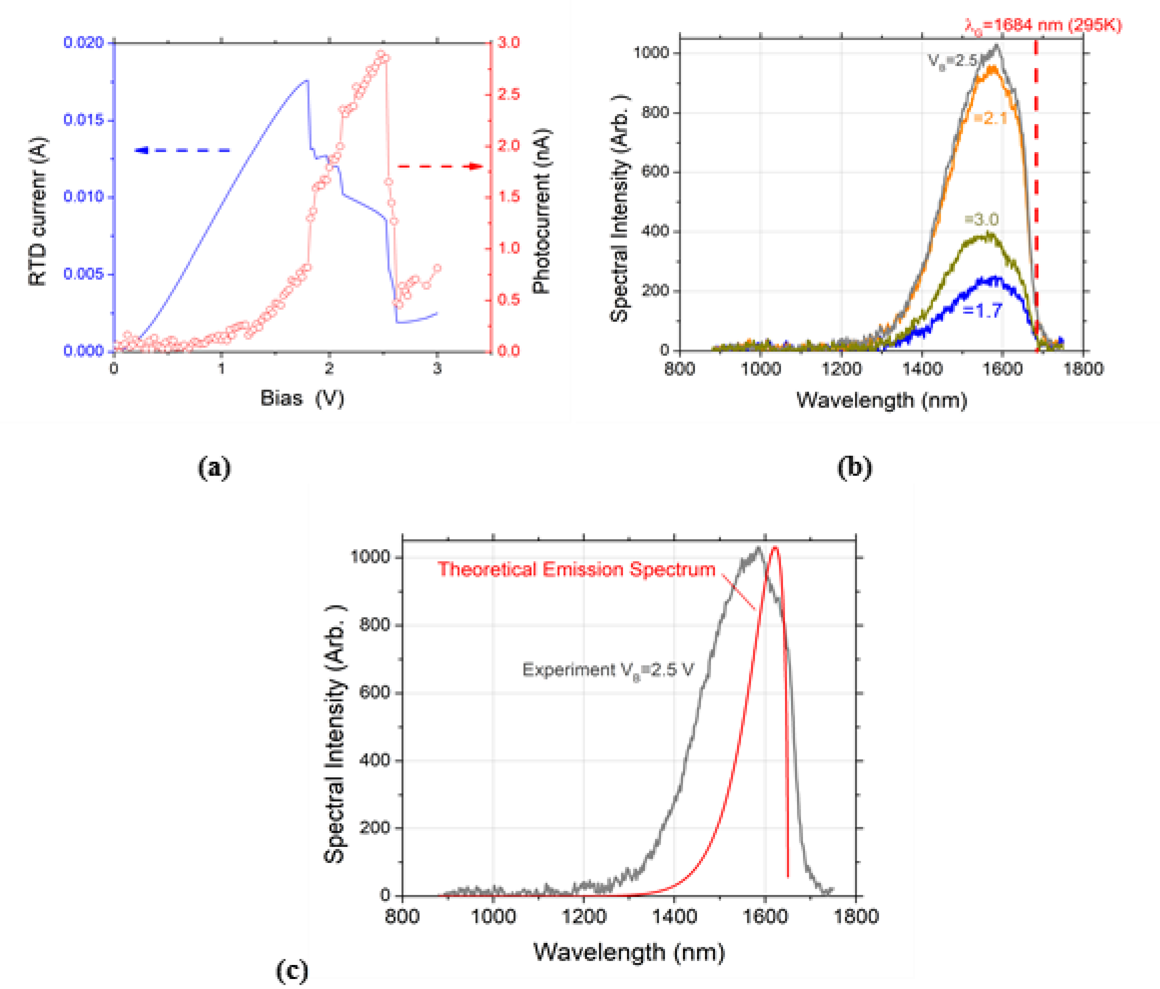


Fig. 25. (a) Room-temperature I-V curve (left vertical axis), and L-V curve measured with the NIR Ge photodiode. (b) Spectrally resolved emission at four different bias voltages. (c) Spectral emission at the most intense spectrum in (b) and the plot of theoretical electro-luminescence according to Eqn. 7.

electrical current. This behavior is similar to that observed for the near-UV photocurrent from one of the GaN/AlN DBRTDs, but for reasons that are not yet understood.

Plotted in Fig. 25 (b) are the spectral emission curves plotted vs. wavelength and parameterized by bias voltage at V_B = 1.7, 2.1, 2.5, and 3.0 V. The middle two bias points are in the NDR region, and the first and last points are just below and above it, respectively. All four curves show a peak emission λ around 1580 nm, and a long-wavelength cut-off behavior around 1650 nm. Superimposed in Fig. 25 (b) is the $In_{0.53}Ga_{0.47}As$ band-edge wavelength reference, λ = 1684 nm corresponding to the band-gap energy of 0.736 eV at 295 K, and calculated with the Varshni formula [47]. The intersection of this reference line with all four spectral curves in their steeply rising edge suggests that the observed emission is occurring at or near the $In_{0.5}Ga_{0.47}As$ band edge. However, the strongest curves in the NDR region (i. e. V_B=2.1, 2.5V) are distinctly asymmetric with short-wavelength tails that extend to 1300 nm, or less. In addition, the weaker emission curve at bias outside the NDR region (i. e. V_B=1.7, 3.0 V) also display short-wavelength tails, but appear more symmetric. The light emission with a UV-VIS fiber spectrometer was examined next. No peak feature was observed in the wavelength range of 200-800 nm. This together with the IR spectrum suggest little possibility of recombination between confined electrons and holes in their respective potential wells.

To understand the emission process better, Fig. 25 (c) shows the brightest of the emission curves plotted against the ideal spontaneous emission expression for a bulk semiconductor EL [48]:

$$S(\nu) = A\nu^2 (h\nu - E_G)^{1/2} \exp[(E_G - h\nu)/k_B T)] \qquad (7)$$

where E_G is the $In_{0.53}Ga_{0.47}As$ band gap [0.736 eV at 295 K], *A* is a frequency-independent constant, and no external cavity effects are included. The agreement is satisfactory on the low frequency (long-wavelength) end, but clearly Eqn (7) decays much faster than the experiment on the short-wavelength end.

V.C. Explanation for EL in GaN RTDs

Initial Non-Equilibrium-Green-Function (NEGF) modeling indicates that the holes necessary for the observed cross-gap emission are created by interband tunneling across the UID GaN collector spacer [Fig. 26 (a)] [30]. A large internal electric field is present as a consequence of polarization-induced charge density caused by two mechanisms: one from piezoelectric polarization because of the abrupt lattice mismatch between the c-axes of AlN (4.982 Å) and GaN (5.185 Å) and the other from the discontinuity of spontaneous polarization between AlN (-0.081 C/m^2) and GaN (-0.029 C/m^2) [49]. The induced surface charge density might reach levels of $\sigma \sim 5.5\times10^{13}/cm^2$, which leads to fields approaching 10 MV/cm in the AlN and at its interface with the GaN layers [49]. These enormous polarization-induced electric fields present in III-nitride heterostructures have been recently confirmed by direct measurement with nano-beam

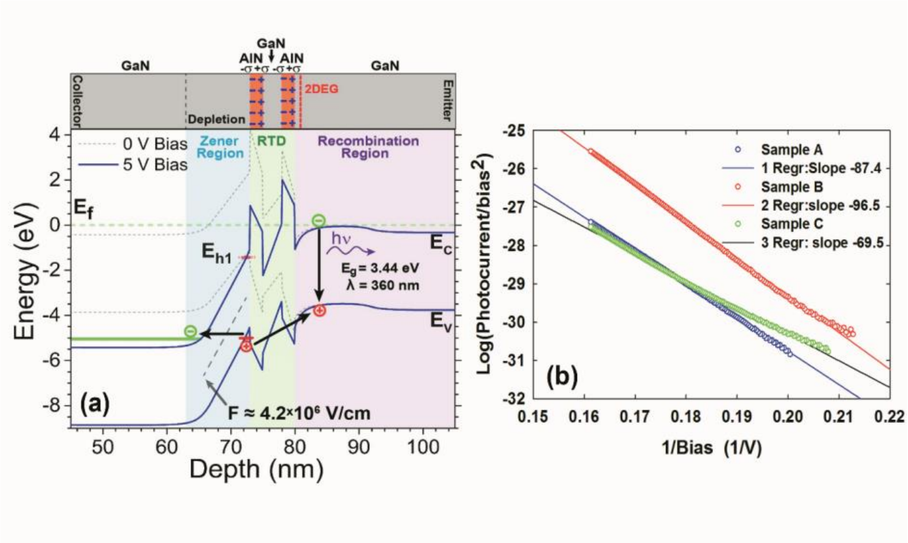
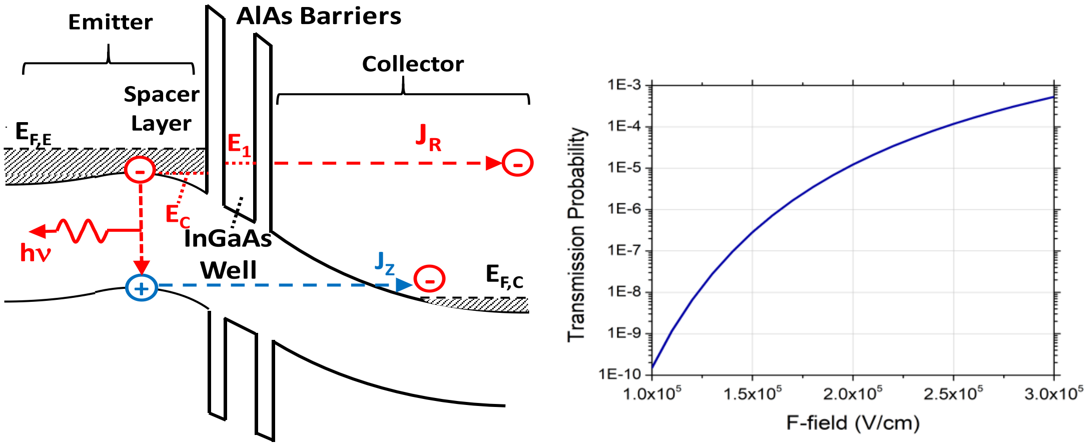


Fig. 26. (a) The band diagram of the GaN/AlN heterostructure generated with a NEGF simulation. The holes are generated in the interband region, they then tunnel through the RTD region into the emitter spacer where they recombine. The lack of observable emission from transitions between bound conduction and valence band states within the quantum well is attributed to the quantum-confined Stark effect, resulting in a small wave function overlap. (b) Forward bias fittings of emission vs. bias voltage with Kane's Zener interband tunneling model.

electron diffraction [50]. The induced field creates a depletion region within the UID collector spacer and an accumulation region in the UID emitter spacer. Under the external voltage bias, the field increases further, which makes interband tunneling possible even though the potential barrier (cross-bandgap GaN) is ~3.44 eV. For perspective, if the internal field is F=2 MV/cm, the interband hole generation density is estimated to be ~0.66/cm^3/s with Kane's model [51], whereas when it increases to F=5 MV/cm, the hole density rate increases to ~3.1×10^{20}/cm^3/s.

Once generated, the holes can migrate by tunneling (possibly by Auger recombination as well) to the emitter side of the structure where electron-hole recombination occurs. For small bias, estimations with a Bardeen Transfer Hamiltonian method indicate the hole transmission through the double-barrier structure is smaller than the electron transmission due to the larger light-hole mass ($m_{lz} \approx 1.1$ vs. $m_e \approx 0.2\ m_0$), despite a smaller valence band offset barrier ($\Delta E_{v\_GaN/AlN} \approx 0.7$ eV vs. $\Delta E_{c\_GaN/AlN} \approx 2.0$ eV [52]). However, the hole transmission increases considerably because the hole quasi-bound level moves downward as the internal field increases [Fig. 26(a)]. This observation is essential to the "bipolar tunneling" effect. Fitting of the experimental photocurrent at both bias polarities was conducted, and the results agree well with Kane's model, thus supporting interband tunneling as the primary source of hole generation [Fig. 26 (b)].

V.D. Explanation for EL of InGaAs RTDs



(a) **(b)**

Fig. 27. (a) Band bending model with positive bias below the NDR region showing simultaneous electron resonant- and interband-tunneling. (b) Interband tunneling probability according to Kane model over the range of bias fields in the present device, along with bias voltages at the boundary fields.

A simple qualitative model that explains the experimental data is shown schematically in Fig. 27(a). The spectra of Fig. 25 (b) and Fig. 25 (c) clearly indicate that the emission is most likely free-carrier cross-gap recombination occurring at or near the $In_{0.53}Ga_{0.47}As$ band edge, which requires free holes. Judging from the threshold in emission shown in Fig. 25 (a) just below 1.0 V bias, the likely generation mechanism for holes is the interband tunneling mechanism of Fig. 27(a). The band-bending in Fig. 27(a) is such that electrons can readily flow by resonant tunneling from the emitter to the collector through the quantum-well quasibound level (E_1). Furthermore, the bias is large enough that unoccupied conduction band states on the collector side line up energetically with occupied valence band states on the emitter side, so that interband tunneling can occur while conserving energy and crystal momentum. The lowest possible threshold bias for this process is approximately the $In_{0.53}Ga_{0.47}As$ bandgap of ≈0.75 eV, which is reasonably close to the experimental threshold. Note that this model is subtly different than that proposed for the GaN/AlN DBRTD where the interband tunneling can occur from the valence-band quasibound level in the quantum well to the collector side, followed by tunneling of the holes to the emitter side where radiative recombination occurs. This is because in GaN, the key factor in the interband tunneling is the huge interfacial polarization field, whereas with $In_{0.53}Ga_{0.47}As$ the key factor is the narrow bandgap.

We present a calculation of the valence-to-conduction band tunneling probability according to the classic Kane expression:

$$T = \frac{\pi^2}{9} exp\left(-\frac{\pi^2 E_G^2 \cdot m}{2hP \cdot F}\right) \qquad (8)$$

where m is the electron mass in vacuum, h is Planck's constant, P is the momentum matrix element between the valence and conduction band cell-periodic wavefunctions, which is generally defined as $E_P \equiv P^2/2m$, and F is the electric field in units of eV/cm [53].

In Fig. 27(b) we plot Eqn. (8) assuming $E_G = 0.736$ eV, $E_P = 25.3$ eV [54], and as a function of F between 1.0×10^5 and 3.0×10^5 eV/cm (this value of E_G is established by device and bandgap modeling described below). The tunneling probability increases more than 6-orders-of-magnitude over this range of bias field, and is essentially a decaying exponential dependence on the length of the band-gap barrier given roughly as $L_B \approx E_G/F$. Between the bias voltage where we first see significant light emission, $V_B \approx 0.75$ eV, and the peak voltage $V_B = 1.75$ V, we observe T increase ~50 times from 2×10^{-7} to 1×10^{-5}. While these values may at first seem small compared to the transmission probabilities for resonant tunneling, which routinely fall in the range 0.1 to 1.0, the overall interband tunneling current also depends on the "supply function" of electrons occupying the valence band on the emitter side, which is very large because of the large effective density-of-states and the high Fermi occupancy factor.

V.E. Comparison of EL in GaN and InGaAs RTDs

To further emphasize the universal nature of the co-tunneling and enhance the accuracy of the analysis, Figure 28 compares the physical characteristics of the $In_{0.53}Ga_{0.47}As$ /AlAs emitter structure studied here to a GaN/AlN structure studied previously. The band-bending plots in Figs 28 (a) and (b) were computed as self-consistent solutions to the coupled Poisson-Schrödinger equations at a bias voltage just below the respective NDR regions. The high electric field in the barrier region of the InGaAs structure, combined with its relatively narrow band gap, makes interband tunneling a significant transport mechanism. The much greater polarization-induced electric field in the GaN/AlN again makes interband tunneling plausible in spite of the much larger GaN bandgap. The essential tunneling parameters of Eqn. (8) for $In_{0.53}Ga_{0.47}As$ and GaN are listed in Table II. Of utmost importance are the electric fields at the peak voltage, F_P [from Figs. 27(a) and (b)], 2×10^5 V/cm and 5×10^6 V/cm for the $In_{0.53}Ga_{0.47}As$ and GaN RTDs, respectively. This large difference makes the factor $(E_G)^2/F$ in Eqn. (8) remarkably close at $F = F_P$: $(E_G)^2/F_P = 2.7\times10^{-8}$ and 2.3×10^{-8} for the InGaAs and GaN, respectively. The only other material-dependent factor in Eqn. (8) is P, which is only ≈25% different between the two materials and is similarly comparable amongst all the common semiconductors independent of bandgap [54].

Also included in Table II is the bandgap at the operating temperature of each device. The bandgap is calculated using the Varshni expression $E_G(T)= E_G(T=0) - \alpha T^2/(T+\beta)$ with parameters

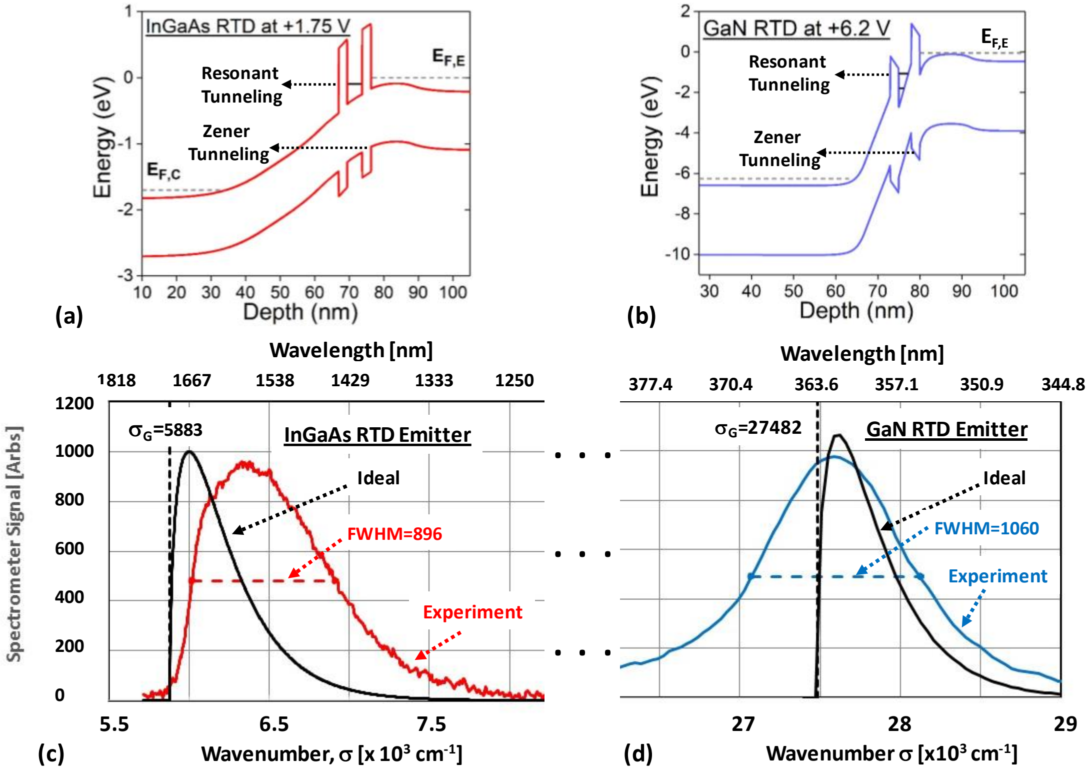


Fig. 28. (a) Band-bending at a bias voltage just below the NDR region for (a) InGaAs RTD, and (b) GaN/AlN RTD, both computed by a self-consistent Poisson-Schrödinger solver. Spectral emission curves at the same bias as in (a) and (b) for (c) the InGaAs emitter, and (d) the GaN emitter, respectively. Also plotted are the ideal electroluminescence curves according to Eqn (8).

.

given in Table II. The operating temperature is estimated by T = 295 K + ΔT with $\Delta T \approx P_0 \cdot R_{TH}$, where P_0 is the dc power dissipation, and R_{TH} is the thermal resistance, also included in Table I. As both devices were mesas having 54 μm^2 active area, the only difference in R_{TH} is the higher thermal conductivity of the GaN-based device compared to the InP-based device in the heat “spreading” region below the mesa. Heat transport to above the mesa through the contact metal is negligible in comparison.

The EL spectra for both structures are plotted in Figs. 28(c) and (d) vs wavenumber (σ [cm^{-1}]) on an identical scale, with 28(c) being the same data as in 25 (b) at 1.7 V bias. Also plotted in Figs. 28(c) and (d) are the ideal EL curves according to Eqn. (7) assuming for the $In_{0.53}Ga_{0.47}As$ device: T = 318 K and E_G = 0.729 ($\sigma_G = 5.883\times10^3$ cm^{-1}); and for the GaN device, T = 355 K and E_G = 3.410 ($\sigma_G = 27.48\times10^3$ cm^{-1}). For the InGaAs device, the experimental EL curve peaks well above (in σ) the maximum of its ideal EL spectra, so it emits the majority of its radiation above the band-edge σ_G, consistent with the “pre-well” quantization effect described above. However, for the GaN device the experimental EL curve peaks close to the ideal-spectrum maximum and has a much broader width, such that the emission above and below the band gap are roughly equal. We again attribute the blue-shifted radiation to the "pre-well" quantization effect, which is strong in GaN as well as $In_{0.53}Ga_{0.47}As$. The red-shifted radiation is not as straightforward. In our previous analysis, the red-shift was obviated by renormalization of the GaN bandgap – an effect which decreases the bandgap energy in proportion to the free carrier concentration [55-58]. However, the lack of red-shifted radiation in the InGaAs device of Fig. 28 (a), even in the presence of the high accumulated electron density in the emitter region, suggests that bandgap renormalization may not be significant. Another possibility for the red-shifted GaN radiation is shallow traps that occur at the GaN emitter layer or at the GaN/AlN interfaces. This is supported by the experimental fact that the total GaN emission spectrum is significantly broader [FWHM =1060 cm^{-1} in Fig. 28(d)] than the InGaAs spectrum [FWHM = 896 cm^{-1} in Fig. 28(c)]. However, more research is necessary to resolve this discrepancy.

Table II Parameters for RTDs

	Parameter	**$In_{0.53}Ga_{0.47}As$ on InP**	**GaN on GaN**
Physical	Mesa Area [um^2]	54	54
Electrical	I_P [mA]	17.1	23.0
	V_P [V]	1.7	6.2
	P_0 [mW]	29.1	144.9
Thermal	R_{TH} [K/W]	798	418
	ΔT [K]	23.2	60.5
Varshni	$E_G(T=0)$	0.803	3.51
	α[meV/K]	0.400	0.909
	β [K]	226	830
	E_G(295K) [eV]	0.736	3.44
Kane	$E_G(T + \Delta T)$ [eV]	0.729	3.41
	F_P [V/m]	$2.0x10^7$	$5.0x10^8$
	E_P [eV]	25.3	20.2
	$(E_G)^2/F_P$	$2.7x10^{-8}$	$2.3x10^{-8}$

V.F. Estimate of Quantum Efficiency in GaN RTDs

This section is focused on the light emission efficiency in GaN/AlN RTDs. To estimate the internal quantum efficiency (IQE), we must first determine the injection efficiency (IE) and the light extraction efficiency (LEE). However, given the uniqueness of the device layout and hole injection, certain assumptions must be made. Because light emission from these devices is hole limited, we can estimate the IE from the ratio between the measured electron and calculated hole current densities (J_p/J_n). Applying this methodology to Sample B (Table I) results in an IE value of ~1.0%. As mentioned earlier, the current RTD/LED structure was designed for stable

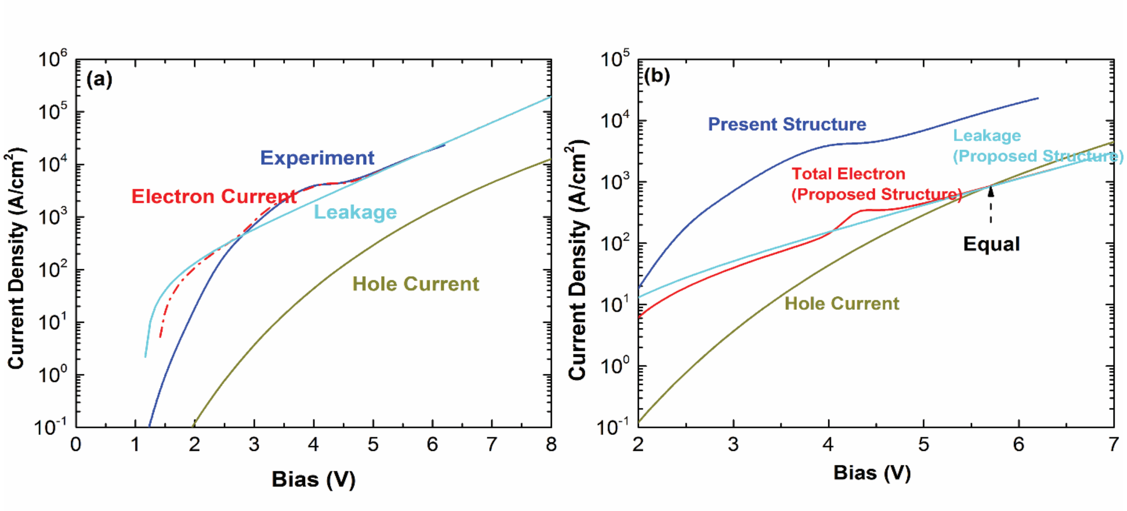


Fig. 29. (a) The fitting of the experimental electron current of sample A with the analytic model [26]. (b) The proposed approach for balancing electron and hole current densities by reducing the n-type doping concentration on the emitter side.

NDR at room temperature and is therefore non-optimal as an LED. The top surface is largely covered by a thick metal (>400 nm), resulting in ~40-60% of the surface area emitting light. Additionally, due to a large refractive index difference between GaN (n=2.6 at 360 nm) and air (n=1.0), the maximum emission efficiency dictated by the narrow escape cone (22.6°) is 3.8%. A concatenation of these effects led us to a LEE estimation of ~1.5-2.3%. Subsequently, combining the measured EQE for Sample B and the estimated IE and LEE values, we approximate the IQE to fall within a range of ~30-50% (IQE = EQE/[IE×LEE]). This range seems quite reasonable because the IQE generally reflects the quality of the crystal and these devices were grown on low dislocation bulk GaN and are also functional RTDs, which is indicative of excellent epitaxy. Improvement of the EQE would involve a more evenly balanced electron and hole current ratio.

To balance the electron and hole current, we investigated the separate current mechanisms of Sample A (Table I) using the previously discussed modeling techniques. Figure 29(a) compares the experimental J-V curve (current of Sample A in Fig. 22 divided by the device area) against our electron and hole current models. The combination of resonant tunneling and leakage current of the electrons offers a good fit to the experimental J-V clearly showing the NDR while the hole current is much smaller in comparison. Above ~5.0 V, near where the experimentally measured device displays a threshold in near-UV light emission, the hole current density begins to take off as well. Further investigation shows that the simplest way to shrink the electron-hole difference is to reduce the electron current density while holding the hole density nearly constant. A reduction in the Fermi energy (E_F) on the emitter side does exactly this, and the electron resonant-tunneling and leakage mechanisms both fall monotonically, whereas the Interband-tunneling of holes have practically no dependence on E_F at all. The n-type doping concentration outside the spacer layer on the emitter side determines E_F, and for the existing structure with $N_D = 5\times10^{19}$ cm^{-3}, $E_F = 0.25$ eV using the conduction-band parameter of GaN, and

$m_e = 0.20\ m_0$. A reduction of N_D to $5\times10^{18}\ cm^{-3}$ would drop E_F to 0.05 eV, and the resulting model J-V curves are plotted in Fig. 29 (b). The electron current drops dramatically but not the hole current such that the two currents are equal at ~5.7 V bias. This simple scaling in an easily controllable material growth parameter should significantly improve the balance between electrons and holes, thereby greatly enhancing the light emission.

VI. Conclusion

We have successfully demonstrated working GaN/AlN and InGaAs/AlAs RTDs. The GaN/AlN RTDs can operate well even under the flux of very high current densities (e. g.~431 kA/cm^2) without thermal breakdown. Their switching time constants were measured in the order of ~55 ps. A f_{max} calculation shows a small-signal oscillation with frequency up to 165 GHz is possible. Compared to the InGaAs/AlAs RTDs, the NDR region early generation GaN/AlN RTDs occurs through the second quasi-bound energy level instead of the first quasi-bound energy level. This is largely attributed to the significant electron accumulation in the interface of the emitter and the left barrier due to the strong field polarization at the GaN/AlN heterostructures. However, in recent batches of high-current-density devices, there are indications of resonant tunneling through the first quasibound state. Unlike the InGaAs/AlAs RTDs, the PVCRs of GaN/AlN RTDs remain low with maximum at only ~1.5, so far. This would substantially hinder any useful high-speed power output. A possible cause to the low PVCR is the scattering mechanisms to the valley current from LO phonon, piezoelectric acoustic phonons, surface roughness and dislocations. Some of the scatterings may be alleviated by the growth of high quality GaN/AlN heterostructures with abrupt interfaces as well as the evolution of bulk GaN substrates. Further, the bias voltage of the NDR region (>4V) is significantly higher than that of InGaAs/AlAs (~2.5V), and the PDR region is broader. This would generate a lot of wasteful DC power consumption when the device is used for RF circuits. A possible solution is to increase the width of the quantum well region lowering the quasi-bound energy levels. All these factors need be considered when designing next generation GaN/AlN RTDs for high speed, high power applications.

Unipolar electroluminescence, without the presence of p-type doping, was observed in both GaN/AlN and InGaAs/AlAs RTDs. This is strong evidence for a co-tunneling mechanism occurring in the RTDs: one is resonant tunneling and the other is interband tunneling. The two processes are relevant to electron injection and hole generation, respectively. For GaN/AlN, the interband tunneling process is most likely due to the strong electric fields originated from polarization effects native to wurtzite structures; and for InGaAs/AlAs, the interband tunneling is largely due to the narrow band gap of InGaAs semiconductor. The scaling study of the Kane model suggests the unipolar electroluminescence is not an isolated effect, it may occur to other RTDs based upon different heterostructure systems. An interesting question will be how to engineer useful devices out of this unique cotunneling mechanism and the resulted electroluminescence.

VII. Acknowledgements

All these works were performed under the sponsorship of a Multi-University-Research-Initiative (MURI), “Devices and Architectures for THz Electronics (DATE)”- managed by Dr. Paul Maki, and have been either published in the literature, dissertations, or under preparation for near-term publications. We also acknowledge the National Science Foundation (Dr. Dimitris Pavlidis) for support under Grants #1711733 & #1711738, and we thank Dr. Ravi Droopad for providing the InGaAs/AlAs RTD structures used as a benchmark for this work.

VIII. References

1) Tsu and L. Esaki, Applied Physics Letters **22**, 562 (1973).
2) L. L. Chang, L. Esaki, and R. Tsu, Applied Physics Letters **24**, 593 (1974).
3) E. R. Brown, J. R. Soderstrom, C. D. Parker, L. J. Mahoney, K. M. Molvar, and T.C. McGill, Appl. Phys. Lett. **58**, 2291 (1991).
4) M. Feiginov, C. Sydlo, O. Cojocari, and P. Meissner, Appl. Phys. Lett. **99**, 233506 (2011).
5) M. Feiginov, H. Kanaya, S. Suzuki, and M. Asada, Appl. Phys. Lett. **104**, p. 243509(1-4) (2014).
6) T. Maekawa, H. Kanaya, S. Suzuki and M. Asada, Electron. Lett., vol. 50, 17, pp. 1214-1216, 2014.
7) S. Kitagawa, S. Suzuki, and M. Asada, IEEE Electron Dev. Lett. **35**, pp. 1215-1217 (2014).
8) T. A. Growden, D.F. Storm, W. Zhang, E.R. Brown, D.J. Meyer, P. Fakhimi, and P.R. Berger, Appl. Phys. Lett. **109**, 083504 (2016).
9) T.A. Growden, Ph.D. Dissertation, The Ohio State University, 2016.
10) T. A. Growden, W-D. Zhang, E.R. Brown, D. F. Storm, K. Hansen, P. Fakhimi, D. J. Meyer, and P. R. Berger, Appl. Phys. Lett. **112**, 033508 (2018).
11) D. F. Storm, T.A. Growden, W. Zhang, E.R. Brown, N. Nepal, D.S. Katzer, M.T. Hardy, P.R. Berger, and D.J. Meyer, J. of Vac. Sci. & Tech B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena **35(2)**, 02B110 (2017).
12) A. Kikuchi, R. Bannai, K. Kishino, C. M. Lee, and J. I. Chyi, Appl. Phys. Lett. **81**, 1729 (2002).
13) K. Kishino, A. Kikuchi, Phys. Stat. Sol. (a) **190**, 23 (2002).
14) S. N. Grinyaev and A.N. Razzhuvalov, Semiconductors **37**, 450 (2003).
15) C. T. Foxon, S.V. Novikov, A.E. Belyaev, L.X. Zhao, O. Makarovsky, D.J. Walker, L. Eaves, R.I. Dykeman, S.V. Danylyuk, S.A. Vitusevich, M.J. Kappers, J.S. Barnard, and C.J. Humphreys, Phys. Stat. Sol. (c) **7**, 2389 (2003).
16) M. Hermann, E. Monroy, A. Helman, B. Baur, M. Albrecht, B. Daudin, O. Ambacher, M. Stutzmann, and M. Eickhoff, Phys. Stat. Sol. (c) **8**, 2210 (2004).
17) S. Golka, C. Pflugl, W. Schrenk, and G. Strasser, Appl. Phys. Lett. **88**, 172106 (2006).
18) C. Bayram, Z. Vashaei, and M. Razeghi, Appl. Phys. Lett. **96**, 042103 (2010).
19) Z. Vashaei, C. Bayram, and M. Razeghi, Appl. Phys. Lett.**107**, 0835053 (2010).
20) L. Yang, H. He, W. Mao, and Y. Hao, Appl. Phys. Lett. **99**, 153501 (2011).
21) T. A. Growden, S. Krishnamoorthy, D.N. Nath, A. Ramesh, S. Rajan, and P.R. Berger, *in Proceedings of Device Research Conference*, University Park, pp. 163-164 (2012).
22) D. Li, L. Tang, C. Edmunds, J. Shao, G. Gardner, M. J. Manfra, and O. Malis, Appl. Phys. Lett. **100**, 252105 (2012).
23) D. Li, J. Shao, L. Tang, C. Edmunds, G. Gardner, M.J. Manfra, and O. Malis, Semicon Sci. and Tech. **28**, 074024 (2013).
24) A. Grier, A. Valavanis, C. Edmunds, J. Shao, J.D. Cooper, G. Gardner, M.J. Manfra, O. Malis, D. Indjin, Z. Ikonic, and P. Harrison, Appl. Phys. Lett. **118**, 224308 (2015).
25) T. A. Growden, E. R. Brown, W-D. Zhang, R. Droopad, and P. R. Berger, Appl. Phys. Lett. **107**, p. 153506 (2015).

26) Tyler A. Growden,Weidong Zhang, Elliott R. Brown, David F. Storm, David J. Meyer, and Paul R. Berger, Nature Light: Science and Applications **7**, 17150 (2018).
27) E. R. Brown, W-D. Zhang, T. A. Growden, D. F. Storm, D. J. Meyer, and P. R. Berger, Noise measurements of high-speed, light-emitting GaN resonant-tunneling diodes, (2018) https://arxiv.org/pdf/1806.09270.pdf.
28) E. R. Brown, W-D. Zhang, T. A. Growden, P. R. Berger, R. Droopad, (2018), https://arxiv.org/abs/1804.07666.
29) D. F. Storm, D. A. Deen, D. S. Katzer, D. J. Meyer, S. C. Binari, T. Gougousi , T. Paskova, E. A. Preble, K. R. Evans, D. J. Smith, Journal of Crystal Growth **380**, 14-17 (2013).
30) Silvaco ATLAS. (2016) www.silvaco.com.
31) B. K. Ridley, B. E. Foutz and L. F. Eastman, Physical Review B **61**, 16862(2000).
32) T. P. E., Broekaert, W. Lee, and C. G. Fonstad, Pseudomorphic $In_{0.53}Ga_{0.47}As$/AlAs/InAs resonant tunneling diodes with peak‐ to‐ valley current ratios of 30 at room temperature. Appl. Phys. Lett. **53**, 1545 (1988).
33) T. Inata, S. Muto, Y. Nakata, S. Sasa, T. Fujii, and S. A. Hiyamizu, Japanese Journal of Applied Physics **26**, L1332-L1334 (1987).
34) E. Ozbay, D.M. Bloom, D.H. Chow, and J.N. Schulman, IEEE Electron Dev. Lett. **14**, 400 (1993).
35) J. F. Whitaker, G. A. Mourou, T.C. L. G. Sollner and W.D. Goodhue, Appl.Phys. Lett. **53,** 385(1988**).**
36) Digital and Mixed Signal Oscilloscopes, MSO/DPO70000 Series Datasheet, (Tektronix, U.S., 2015) 17.
37) E.R. Brown, C.D. Parker, T.C.L.G. Sollner, A.R. Calawa, M.J. Manfra, C.L. Chen, S.W. Pang, and K.M. Molvar, "High-speed resonant-tunneling diodes made from the $In_{0.53}Ga_{0.47}As$/AlAs system," *SPIE Proceedings on High-Speed Electronics and Device Scaling* **1288**, p. 122, 1990.
38) E. R. Brown,"High-speed resonant-tunneling diodes," *in Heterostructure and Quantum Devices*, ed. by N.G. Einspruch and W.R. Frensley, (Academic, Orlando, 1994), pp. 306-347.
39) Keysight: https://literature.cdn.keysight.com/litweb/pdf/5991-3904EN.pdf?id=2447379
40) W-D. Zhang, E. R. Brown, T. A. Growden , P. R. Berger, and R. Droopad, IEEE Trans. Electron Devices **63**, 4993-4997 (2016).
41) D. F. Storm, T. A. Growden, W-D. Zhang, D. S. Katzer, M. T. Hardy, D. J. Meyer, E. R. Brown and P. R. Berger, “RF-MBE growth of AlN/GaN/AlN resonant tunneling diodes on freestanding GaN and GaN templates,” *in Proceedings of 34th North American Molecular Beam Epitaxy Conference,* Alberta, 2019.
42) E. R. Brown, W. D. Goodhue, and T. C. G. Sollner, J. Appl. Phys. **64**, 1519 -1529(1988).
43) E. R. Brown, O. B. McMahon, L. J. Mahoney, and K. M. Molvar, Electron. Lett. **32**, p. 938-940 (1996).
44) D. Zanato, S. Gokden, N. Balkan, B. K. Ridley, and W. J. Schaff, Semicond. Sci. Technol. **19** , pp. 427-432 (2004).
45) S. P. DenBaars, D. Feezell, K. Kelchner, S. Pimputkar, C. C. Pan, C-C Yen, S. Tanaka, Y. Zhao, N. Pfaff N, R. Farrell, M. Iza, S. Keller, U. Mishra, J. S. Speck, S. Nakamura, Acta Materialia **61**, 945-995 (2013).
46) M. A. Zimmler, J. Bao , I. Shalish , W.Yi , V. Narayanamurti, and F. Capasso, Nanotechnology **18**, 395201 (2007).
47) Y. P. Varshni , Physica **34**, pp. 149-154 (1967).
48) S. M. Sze, Physics of Semiconductor Devices, 2nd Ed., (John Wiley and Sons, New York, 1981).
49) 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, Journal of Applied Physics **87**, 334-344 (2000).
50) D. Carvalho, K. Müller-Caspary, M. Schowalter, T. Grieb, T. Mehrtens, A. Rosenauer, T. Ben , R. García1, A. R. Cubero, K. Lorenz, B. Daudin B and F. M. Morales.. Scientific reports (2016). DOI: 10.1038/srep28459.
51) E. O. Kane, Journal of Physics and Chemistry of Solids **12**, 181-188 (1959).
52) G. Martin, A. Botchkarev, A. Rockett, and H. Morkoc, Applied Physics Letters **68**: 2541-2543 (1996).
53) W. Vandenberghe, B. Soree, W. Magnus, and F. Groeseneken, J. Appl. Phys. **107**, 054520 (2010).
54) I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
55) K. F. Berggren and B.E. Sernelius, Phys. Rev. B **24**, 3240 (1984).
56) E. Burstein, Phys. Rev. **93**, 632 (1954).
57) T. S. Moss, Proc. Phys. Soc. B **67**, 775 (1954).
58) M. Bouzidi, Z. Benzarti, I. Halidou, S. Soltani, Z. Chine, and B. El Jani, Mat. Sci in Semicon. Processing, **42**, 273 (2016).