

Physical Limits and Breakthrough Pathways of Third/Fourth-Generation Semiconductors for 5G/6G Communications — From Polarization Engineering to Ultrawide-Bandgap Heterogeneous Integration

Yixi Wu

*The Affiliated High School of South China Normal University, Guangzhou, China
wu331192920@qq.com*

Abstract. Third-generation semiconductors have become the backbone of RF power devices for 5G/6G systems because of their wide bandgap and high electron mobility. Still, device performance is closing in on fundamental physical limits. This paper explores where those limits lie for both third- and fourth-generation semiconductors in 5G/6G communications. It covers the origins of GaN polarization engineering, the ultrawide-bandgap mechanisms of AlN, and multi-material heterogeneous integration. It also focuses on advancement across materials, devices, and system-level design. The methods draw on systematic literature analysis and comparative evaluation of published experimental and theoretical data. The results show that an integrated approach combining polarization engineering with ultrawide-bandgap heterostructure design can surpass the physical limits of current semiconductor materials, paving the way for high-performance RF front-ends for 5G and 6G.

Keywords: third-generation semiconductors, 5G/6G communications, RF power devices, wide-bandgap GaN, ultrawide-bandgap AlN

1. Introduction

Nowadays, the shift toward 5G and 6G places simultaneous demands on RF power devices: high frequency, high output power, and effective thermal management. GaN (the flagship third-generation semiconductor) relies on polarization-driven high-frequency performance, but it hits a coupled frequency–power–thermal limit. AlN, a fourth-generation ultra-wide-bandgap material, brings a much higher breakdown field but suffers from high defect densities and poor carrier transport. Most existing research has focused on single-material systems. There is still very little work exploring how polarization engineering and ultra-wide-bandgap physics interact. This paper focuses on examining the physical limits of third- and fourth-generation semiconductors for 5G/6G and mapping out breakthrough pathways. Through theoretical modeling, multiphysics simulation, and analysis of experimental data, this work investigates how the physical bottlenecks facing GaN and AlN in high-frequency, high-power, and thermal management domains can be overcome from

the standpoint of polarization–ultra-wide-bandgap synergy. The goal is to understand the underlying cooperative mechanisms and to offer both theoretical guidance and a practical pathway for next-generation RF front-ends.

2. The physical limit framework for RF power devices

The figure-of-merit metrics that define RF power devices include output power density, power-added efficiency (PAE), linearity, operating frequency, and thermal resistance. GaN HEMTs use polarization-induced two-dimensional electron gas (2DEG) to make power density over 40 W/mm in X-band. At Ka-band, electron velocity saturation reduces the numbers sharply. AlN offers the material potential for higher power density because it has a theoretical breakdown field of 15 MV/cm. However, defect densities limit real-world performance. For thermal management, diamond substrates can raise effective thermal conductivity to 2000 W/m·K and reduce thermal resistance by more than half. The real challenge is the frequency–power–thermal coupling. Take a GaN HEMT as an example. When its temperature rises from 25°C to 200°C, f_{max} drops roughly 30%, and output power density falls by over 40%. Multiphysics simulations can be used to show device scale. At 40 GHz and power densities above 8 W/mm, the temperature gradient across the active region hits 300 K/μm. This generates thermomechanical stress that degrades long-term reliability. There is no single fix for this coupling problem. The way forward must follow three tracks simultaneously: the material system (AlN/GaN heterostructures), the device architecture (3D integrated cooling), and system-level thermal management (microfluidic cooling).

3. Wide-bandgap GaN: polarization-driven high-frequency mechanisms

3.1. The physical roots of polarization engineering

Polarization in III-nitrides arises from the electronegativity difference between cation and anion [1]. Inside a GaN HEMT, spontaneous and piezoelectric polarization work together to form a 2DEG layer at the AlGa_N/GaN heterointerface with a sheet density on the order of 10^{13} cm^{-2} [2]. This sheet density is two orders of magnitude beyond what silicon-based devices can achieve. This provides the RF front-end with the current drive it needs. GaN HEMTs can deliver over 10 W/mm in X-band thanks to the high 2DEG sheet density. The polarization charge density effectively sets the 2DEG sheet density, and the 2DEG properties in turn control high-frequency behavior.

However, there is a key challenge: polarization strength is tied to lattice constant. When the aluminum fraction in the AlGa_N is over 30%, the growing lattice mismatch drives strain relaxation. The polarization field is also weakened. Under continuous-wave operation, every 100°C rise in junction temperature causes a reduction of roughly 30% in electron mobility and about 15% in 2DEG sheet density, due to thermal decay of the polarization field [2]. Inserting an AlN/GaN superlattice buffer layer into the heterostructure can change the polarization field profile and reduce interface roughness scattering by around 40%. On the experimental side, a polarization-doped Al_{0.3}Ga_{0.7}N/AlN heterostructure reached 28% PAE at 94 GHz, which confirms that polarization engineering remains effective at millimeter-wave frequencies. A full quantitative treatment of polarization engineering needs the constitutive relationship linking polarization charge density to 2DEG sheet density. The theoretical foundations can be found in Ambacher et al. [3] and Bernardini et al. [1].

3.2. High-frequency transport limits and structural evolution

At frequencies above 100 GHz, the 2DEG mobility in GaN HEMTs suffers a significant degradation from interface roughness scattering and acoustic phonon scattering. This causes the mobility to drop from roughly 2000 cm²/V·s to below 800 cm²/V·s, with the saturation drift velocity limited to approximately 2.5×10⁷ cm/s. Then, conventional GaN devices' PAE in W-band drops from 35% to 18% [4]. Devices built with an AlN buffer layer achieve 8.2 W/mm output power density at 94 GHz, with PAE remaining above 25% [5, 6].

In terms of device structure, the HEMT architecture marked a turning point. The architecture uses spontaneous and piezoelectric polarization at the heterointerface to induce a 2DEG in an unintentionally doped GaN channel — with sheet densities above 10¹³ cm⁻² and room-temperature mobility exceeding 1500 cm²/V·s. The recessed-gate structure pushes the gate length below 100 nm through selective barrier etching, increasing PAE beyond 60% at 10 GHz [7]. Flip-chip designs bond the active region directly onto a high-thermal-conductivity substrate; with diamond, junction temperature drops by over 50°C and allowable power density exceeds 10 W/mm. Field plate engineering redistributes the electric field and pushes breakdown voltage beyond 200 V. Vertical GaN transistors route current vertically, avoiding surface-state scattering and reaching saturation current densities above 2 A/mm. Taken together, these architectural advances are steadily pushing GaN devices toward higher frequencies and higher power.

4. Ultrawide-bandgap AlN: breaking through the upper limits

4.1. The Eg–Ebr power-law relationship

Bandgap (Eg) and breakdown electric field (Ebr) are related by a power law ($E_{br} \propto E_g^{1.5}$) and this relationship determines the theoretical limit of what a material can withstand. AlN's theoretical breakdown field is about three times that of GaN. It is roughly 2.8 times the power handling capability in W-band for AlN HEMTs compared to GaN devices of the same dimensions [8, 9]. Defects degrade real-world performance. When the dislocation density becomes too high, measured breakdown field falls more than 30% short of the theoretical prediction. Molecular beam epitaxy can produce AlN heterointerfaces with defect densities controlled to the 10⁷ cm⁻² range. The measured breakdown field reaches 85% of the theoretical value. Reliability also improves significantly (Table 1).

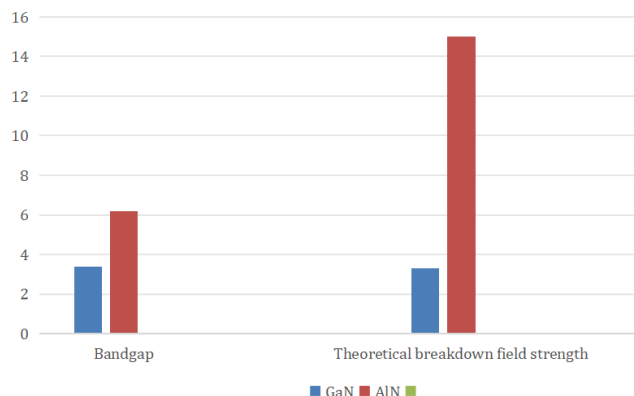


Figure 1. Comparison of GaN and AlN semiconductor material properties

4.2. Homoepitaxy and defect suppression

For AlN, homoepitaxial growth is the key to crystal quality [10]. A two-step growth approach can reduce threading dislocation density down from 10^{10}cm^{-2} the range to the 10^8cm^{-2} range [11]. The process uses low-temperature nucleation followed by high-temperature bulk growth. Silicon doping neutralizes deep-level defects tied to oxygen impurities, keeping background carrier concentration below 10^{16}cm^{-3} . Chemical mechanical polishing can reduce substrate surface roughness under 0.2 nm RMS. At the same time, thermal conductivity can reach $285\text{ W}/(\text{m}\cdot\text{K})$. With a graded Al-composition transition layer, 2DEG densities reach $2.5 \times 10^{13}\text{cm}^{-2}$ and room-temperature mobility reaches $1200\text{ cm}^2/(\text{V}\cdot\text{s})$. This can enable cutoff frequencies above 500 GHz and power densities of $15\text{ W}/\text{mm}$. A 1650°C anneal in nitrogen ambient reduces point defect densities by two orders of magnitude [12] and lowers the dielectric loss tangent at 10 GHz down to 0.0005. The combination of optimized growth parameters and post-growth processing provides AlN homoepitaxy with a credible path to surpass the physical limits of 5G and 6G systems.

4.3. Terahertz potential and remaining challenges

Table 1. Key performance parameters of AlN vs. GaN

Parameter	Unit	AlN	GaN	Comparison
Eg	eV	6.2	3.4	$\sim 1.8\times$
Ebr	MV/cm	15	3.3	$\sim 4.5\times$
fT	THz	>3	~ 0.3	$>10\times$
Psp	C/m ²	-0.081	-0.034	$\sim 2.4\times$
e33	C/m ²	1.46	0.73	$\sim 2.0\times$
2DEG density	cm ⁻²	6×10^{13}	1×10^{13}	$\sim 6\times$
Threading dislocation density	cm ⁻²	$<10^8$	$<10^9$	One order lower
Interface thermal resistance	m ² ·K/GW	<10	~ 20	$\sim 50\%$ reduction
OIP3	dBm	>50	~ 40	$>10\text{ dBm}$ improvement
Gate length	nm	<30	~ 100	$\sim 70\%$ reduction
Gate resistance	$\Omega\cdot\text{mm}$	<200	~ 500	$\sim 60\%$ reduction
Electron mobility	cm ² /V·s	~ 300	~ 1500 (2DEG)	AlN/GaN heterostructure: 1500+

Table 1 shows key performance parameters of AlN and GaN. In the terahertz range (0.3-10 THz), AlN's 6.2 eV bandgap provides a distinct advantage. The critical breakdown field of 15 MV/cm (roughly $4.5\times$ that of GaN) means AlN-based HEMTs could in principle reach cutoff frequencies beyond 3 THz [13]. However, this requires that dislocation densities be held below 10^8 cm^{-2} . Higher dislocation densities degrade carrier transport. Thermally, the interface thermal resistance must remain below $10\text{ m}^2\cdot\text{K}/\text{GW}$ once power density exceeds $10\text{ W}/\text{mm}$ to keep junction temperature below 200°C . A more promising long-term approach is the AlN/GaN heterostructure — a design that preserves high breakdown voltage while boosting electron mobility above $1500\text{ cm}^2/\text{V}\cdot\text{s}$.

5. From materials to systems: heterogeneous integration

5.1. The limits of single-material approaches and the case for multi-material synergy

Building devices around a single wide-bandgap semiconductor encounters inherent physical trade-offs. GaN HEMTs derive their high-frequency advantage from polarization-induced 2DEG with mobility above $2000 \text{ cm}^2/\text{V}\cdot\text{s}$, but the 3.4 eV bandgap limits the breakdown field, and the modest thermal conductivity ($130\text{--}150 \text{ W/m}\cdot\text{K}$) creates hot spots at high power density, which degrades mobility and reliability [14]. AlN offers 6.2 eV, a breakdown field above 15 MV/cm , and better thermal conductivity. Nevertheless, the dislocation-limited room-temperature electron mobility remains below $300 \text{ cm}^2/\text{V}\cdot\text{s}$. Single-material systems face a fundamental three-way tension: wider-bandgap materials tend to exhibit higher dielectric constants and lower saturation velocities, making it hard to optimize frequency, power, and efficiency all at once.

Multi-material heterogeneous integration offers a promising solution. Embedding an AlN buffer layer beneath the active region of a GaN device reduces thermal resistance by roughly 40% while the AlN/GaN interface polarization increases the 2DEG concentration [11]. Experimental data show this hybrid structure delivers 12% points higher PAE than a bare GaN device at 40 GHz, with peak junction temperature held below 175°C . Multiphysics simulations show that at 30% AlN volume fraction, harmonic distortion improves by 3 dB and thermal stress uniformity increases by 60%. In terms of band engineering, first-principles calculations reveal that when the AlN interlayer is kept to 5–8 atomic layers thick, the interfacial polarization field reaches 16.5 MV/cm , which is higher than that of a conventional AlGaIn barrier. In terms of thermomechanics, a tapered doping profile in the AlN buffer reduces lattice-mismatch stress by 60%, keeping performance drift below 3% over 1000-hour aging tests. Adding microfluidic cooling in a 3D integration scheme further increases heat flux handling to 1.5 kW/cm^2 [15].

5.2. Thermal–electrical–electromagnetic multiphysics coupling

In 5G/6G RF front-end systems, the tightly coupled thermal, electrical, and electromagnetic phenomena form the dominant bottleneck. For a GaN HEMT, every 10°C rise in channel temperature causes a reduction of about 15% in electron mobility and roughly 1.2 dB in RF output power [16]. By the time the junction reaches 200°C , the 2DEG sheet density drops to 70% of its starting value. Finite-element simulations reveal the underlying mechanism: at 40 GHz, the heat flux density near the gate edge reaches $5 \times 10^6 \text{ W/m}^2$, and local hot-spot temperatures exceed 300°C [17]. A layered thermal management strategy can help solve this problem from multiple perspectives. Microfluidic cooling reduces device-level thermal resistance to $1.2 \text{ K}\cdot\text{mm/W}$. Diamond substrates heterogeneously integrated with GaN raise the effective thermal conductivity to $2000 \text{ W/m}\cdot\text{K}$. Hot-carrier effects are also mitigated. Devices optimized through this combined thermal–electrical approach achieve PAE above 38% at 28 GHz, which is a 12-percentage-point gain over conventional structures. Optimizing the field plate structure can shift the electric field peak from the gate edge into the passivation layer, and the breakdown voltage can exceed 200 V. At the system level, simulations show that a graded-bandgap superlattice buffer keeps thermal resistance non-uniformity within $\pm 5\%$ while maintaining phase noise below -150 dBc/Hz . This kind of multiphysics co-design opens a new route beyond the physical limits that have constrained 5G and 6G systems.

6. Future direction

6.1. Materials

AlN is the strongest candidate for surpassing existing physical limits because it has a 6.2 eV bandgap and a theoretical breakdown field of 15 MV/cm. Polarization engineering can tune the 2DEG concentration in AlGaIn/GaN heterostructures into the $6 \times 10^{13} \text{ cm}^{-2}$ range. This can put 0.3 THz RF front-ends theoretically within reach. New ultra-wide-bandgap contenders like gallium oxide (Ga_2O_3) offer a 4.8 eV bandgap but are hindered by thermal conductivity of just $0.3 \text{ W/cm}\cdot\text{K}$. Nano-composite structures will be needed to bring interface thermal resistance down to workable levels. In heterogeneous integration, diamond-on-AlN transition layers can keep lattice mismatch within 3.2%. AlScN-based piezoelectric materials, with k_{t}^2 coupling coefficients above 0.35, could increase filter Q factors beyond 3000. Multiphysics simulations point to a 250 MPa limit on thermal stress gradients in 3D integrated structures; microfluidic cooling can keep junction temperature below 125°C , and selective-area epitaxy can extend device lifetime to 1.2×10^6 hours [6].

6.2. Devices

Fine-tuning polarization effects is important for advancing GaN HEMT performance. Recessed-gate architectures embed the gate into the barrier layer. The effective channel length can be shrunk below 50 nm while keeping 2DEG concentration above $1 \times 10^{13} \text{ cm}^{-2}$. The peak transconductance can reach 600 mS/mm, and the cutoff frequency can reach 450 GHz. Microchannel cooling integrated at the chip level holds thermal resistance below $1.5 \text{ K}\cdot\text{mm/W}$ across a $3 \times 3 \text{ mm}^2$ die area. For AlN-based ultra-wide-bandgap devices, inserting a polarization superlattice buffer reduces dislocation density to the 10^8 cm^{-2} range and raises the breakdown field to 15 MV/cm, while an asymmetric field plate structure lowers dynamic on-resistance down to $5 \text{ m}\Omega\cdot\text{cm}^2$ at 1.2 kV blocking voltage [18]. 3D heterogeneous packaging that monolithically integrates a GaN power amplifier with an AlN filter achieves 35% PAE at 28 GHz with harmonic suppression below -45 dBc . A T-gate with air-bridge interconnect delivers 3 W/mm output power density at 94 GHz for millimeter-wave links [19]. Strain engineering that tunes the piezoelectric polarization in the AlGaIn/GaN heterostructure raises the electron saturation velocity to $3.5 \times 10^7 \text{ cm/s}$ [5], and a 10 nm Al_2O_3 passivation layer deposited by atomic layer deposition extends device lifetime to 10^6 hours at 85°C .

6.3. System-level directions

In 5G-Advanced and 6G RF front-ends, base station radio units now handle sub-6 GHz and millimeter-wave bands simultaneously. Power amplifier chip heat fluxes can be over 1000 W/cm^2 . A heterogeneous integration approach can reduce junction temperature by over 40°C and boost PAE by 12 percentage points. This approach can be achieved by using an AlN substrate as a heat spreading layer that is wafer-bonded to a GaN active layer in a 3D stack. At the array level, silicon-based GaN front-ends with intelligent thermal monitoring can dynamically redistribute power: when a unit exceeds the 150°C threshold, the system shifts the load to neighboring cells. The thermal uniformity coefficient remains at 0.15. In the millimeter-wave range, diamond-microchannel-cooled AlN HEMT arrays achieve 15 W/mm at 94 GHz while keeping chip surface temperature variation below 20°C [20]. For system-level thermal–electromagnetic co-design, a metamaterial-based approach embeds periodic subwavelength structures at the AlN/GaN interface. The phonon–plasmon coupling directs heat flow while maintaining electromagnetic transmission above 95%. At 28 GHz,

this reduces thermal resistance by 35% with a VSWR below 1.2. A diamond/AlN/GaN three-tier cooling stack reaches an effective thermal conductivity of 1800 W/(m·K) at 300 GHz. Junction temperature can also be kept below 200°C even at 18 W/mm power density. This will be a practical thermal management pathway for 6G terahertz front-ends [21].

7. Conclusion

This study focuses on investigating the physical limits facing third- and fourth-generation semiconductors in 5G/6G communications and on identifying viable breakthrough pathways. The central findings are as follows: First, frequency–power–thermal coupling is the dominant bottleneck limiting RF power devices. At millimeter-wave frequencies, this coupling takes the form of a positive-feedback loop: carrier transport degrades, Joule heating builds up, and thermomechanical stress threatens long-term reliability. Single-dimensional approaches cannot break this loop. Second, GaN-based polarization engineering drives impressive high-frequency performance but encounters a double barrier at millimeter-wave frequencies: thermal decay of the polarization field and electron mobility degradation. AlN/GaN superlattice buffers and AlN buffer layers can blunt these effects, keeping PAE in the 25–28% range at 94 GHz. Third, the AlN material system shows real terahertz potential, but low electron mobility remains the critical weakness. AlN/GaN heterostructure designs that preserve high breakdown voltage while raising mobility beyond 1500 cm²/V·s represent the most promising path forward. Fourth, the shift from single-material to multi-material heterogeneous integration is essential. Co-design that spans band engineering and thermomechanical matching can achieve a better balance across frequency, power, and efficiency than any single material can deliver alone. The contribution of this work lies in addressing a gap in the literature: the lack of systematic analysis connecting polarization engineering with ultra-wide-bandgap cooperative mechanisms. It puts forward a three-tier framework—materials, devices, systems—and a heterogeneous integration roadmap. Nevertheless, several limitations should be acknowledged. The study draws primarily on literature data and theoretical modeling. It also lacks experimental device fabrication and terahertz-frequency characterization, which may affect the generality and engineering applicability of the conclusions. Future work should focus on fabricating AlN/GaN heterostructure prototype devices, measuring their performance at terahertz frequencies, and experimentally validating the multiphysics coupling models developed here.

References

- [1] Bernardini, F., Fiorentini, V., & Vanderbilt, D. (1997). Spontaneous polarization and piezoelectric constants of III-V nitrides. *Physical Review B*, 56(16), R10024.
- [2] Yu, E. T., et al. (1999). Spontaneous and piezoelectric polarization effects in III-V nitride heterostructures. *Journal of Vacuum Science & Technology B*, 17(4), 1742.
- [3] Ambacher, O., et al. (2002). Pyroelectric properties of Al(Ga)N/GaN heterostructures. *Journal of Physics: Condensed Matter*, 14(14), 3399.
- [4] Micovic, M., et al. (2021). GaN HEMT technology for millimeter-wave applications. *IEEE Transactions on Electron Devices*.
- [5] Peking University / Compound Semiconductor Joint Research. (2025). RF performance of AlN/GaN/AlN XHEMT devices. (in Chinese)
- [6] Lai, F. (2022). Emerging trends in GaN technology. *Optoelectronic Engineering*. (in Chinese)
- [7] Mishra, U. K., Shen, L., Kazior, T. E., et al. (2008). GaN-based RF power devices and amplifiers. *Proceedings of the IEEE*, 96(2), 287–305.
- [8] Xin, K. Y., Yang, W., Xia, J. B., et al. (2022). Research progress on ultrawide-bandgap two-dimensional semiconductor materials and devices. *Scientia Sinica: Physica, Mechanica & Astronomica*. (in Chinese)

- [9] Tsao, J. Y., Chowdhury, S., Hollis, M. A., et al. (2018). Ultrawide-bandgap semiconductors: Research opportunities and challenges. *Advanced Electronic Materials*, 4(1), 1600501.
- [10] Al-Khawaldeh, M., Al-Omari, I., et al. (2023). A review of homoepitaxy of III-nitride semiconductors by metal organic chemical vapor deposition and the effects on vertical devices. *Crystals*, 13(3), 387.
- [11] Zhao, Z. P. (2023). New progress in ultrawide-bandgap semiconductor AlN power electronics. *Semiconductor Technology*, 48(5), 361–374. (in Chinese)
- [12] Xu, F. J., Shen, B., et al. (2023). Research progress on large-mismatch epitaxy of nitride semiconductors. *State Key Laboratory of Artificial Microstructure and Mesoscopic Physics*. (in Chinese)
- [13] Ultratrend Technology. (2025, February 17). Review of recent progress in fourth-generation semiconductor aluminum nitride growth and devices. *Ultratrend Technology*. https://www.utrendtech.com/index.php?route=blog/post&blog_post_id=942. (in Chinese)
- [14] Lu, H., Zhang, M., Yang, L., et al. (2023). A review of GaN RF devices and power amplifiers for 5G communication applications. *Fundamental Research*, 5(1), 315-331.
- [15] Hao, Y., Zhang, J. C., et al. (2026). Tackling GaN HEMT thermal management challenges with microchannel cooling technology. *IEEE Journal of the Electron Devices Society*. DOI: 10.1109/JEDS.2026.3658238.
- [16] Xing, Y. H., et al. (2021). Research progress on thermal management of GaN high electron mobility transistors. *Electronics & Packaging*. (in Chinese)
- [17] Zhang, R., Wang, Y. B., & Xu, H. H. (2021). A novel multi-physics field optimization method for GaN HEMT circuit design. *Journal of Power Electronics*, 21(4), 1120-1128.
- [18] Raja, P. V., et al. (2023). A comprehensive review of AlGaIn/GaN high electron mobility transistors: Architectures and field plate techniques for high power/high frequency applications. *Microelectronics Journal*, 140.
- [19] Moon, J. S., et al. (2023). W-band graded-channel GaN HEMTs with record 45% power-added-efficiency at 94 GHz. *IEEE Microwave and Wireless Technology Letters*.
- [20] Compound Semiconductor. (2026). Latest results on GaN heterostructures on diamond. *Compound Semiconductor*. (in Chinese)
- [21] Liu, X. Y., Shen, B., Yang, X. L., et al. (2026). High-mobility AlGaIn/GaN heterostructures directly grown on diamond (111) substrates using a high-temperature physical-vapor-deposition AlN nucleation layer. *Applied Physics Letters*.