International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-10

https://doi.org/10.70917/ijcisim-2026-2818
Article

Silicon-Based GaN Power Devices and CMOS Monolithic Heterogeneous Integration with Driver Circuit Co-Optimization

Hongyi Wei [1,*]

[1] Rabun Gap-Nacoochee School, Rabun Gap, GA 30568, USA
* Correspondence author: hongyi070927@sina.com

Abstract: To address the bottlenecks of excessive parasitic parameters, performance mismatch, and poor miniaturization in discrete integration of GaN power devices and CMOS drivers, a CMOS-compatible monolithic heterogeneous integration scheme is proposed. The epitaxial structure of GaN-on-Si heterojunction and the integration process flow were optimized, and a co-optimization strategy coupling GaN device structure parameters with CMOS driver circuit timing was developed. Experimental results show that the integrated system achieves a 68% reduction in parasitic inductance, 52% reduction in parasitic capacitance, and a switching speed of 250 V/ns compared with discrete solutions. The power conversion efficiency reaches 97.8% at 1 MHz and maintains above 95% in the load current range of 0.5-10 A. This work provides a feasible technical path for high-performance, miniaturized, and reliable power electronic systems.

Keywords: GaN-on-Si, Monolithic Heterogeneous Integration, CMOS, Power Device, Driver Circuit, Co-optimization.

1. Aims and Background

1.1. Research Background

Power electronic systems are the core enabling technology for emerging fields such as new energy vehicles (NEVs), renewable energy generation, smart grids, and portable electronics. With the rapid development of these industries, there is an urgent demand for power devices and systems with higher efficiency, faster switching speed, smaller volume, and lower cost [1]. Gallium nitride (GaN), as a representative wide-bandgap (WBG) semiconductor material, exhibits inherent advantages over traditional silicon (Si) devices: a bandgap of 3.4 eV (vs. 1.12 eV for Si), electron mobility exceeding 2000 $cm^2/(V \cdot s)$, and a critical electric field 10 times higher than that of Si [2]. These properties enable GaN power devices to achieve lower on-resistance (Ron), faster switching speed, and higher breakdown voltage (BVds) at the same chip area, making them ideal for high-frequency, high-power density applications.

However, the practical application of GaN power devices is severely limited by the integration method with driver circuits. Currently, discrete integration (separate GaN power chips and CMOS driver chips interconnected via printed circuit board (PCB)) is the mainstream solution. This approach introduces significant parasitic inductance (Lparasitic) and capacitance (Cparasitic) from the interconnection wires, bond wires, and PCB traces [3]. For example, bond wires typically contribute 2-5 nH of parasitic inductance, which causes voltage overshoot ($\Delta V = L \bullet di/dt$) and ringing during switching transitions, increasing switching loss and electromagnetic interference (EMI) [4]. Additionally, the performance mismatch between off-the-shelf CMOS drivers and GaN devices (e.g., mismatched output current capability, inappropriate dead-time) further suppresses the potential performance of GaN

[5].

Monolithic heterogeneous integration, which integrates GaN power devices and CMOS driver circuits on a single Si substrate, is recognized as the ultimate solution to eliminate parasitic effects. By shortening the interconnection distance to the micrometer scale, this technology can drastically reduce parasitic parameters while improving system integration density and reliability [6]. However, the realization of this scheme faces two core challenges: (1) Process compatibility: GaN epitaxy and fabrication require high-temperature processes (e.g., 850-1000 °C annealing for ohmic contacts), which are incompatible with standard CMOS processes (maximum temperature <450 °C for backend processes) [7]. (2) Device-driver synergy: The fast switching characteristics of GaN (ns-level transition time) require driver circuits with ultra-low propagation delay and precise dead-time control, which cannot be achieved by conventional driver designs [8].

1.2. Research Status and Gaps

In recent years, extensive research has been conducted on GaN-CMOS integration. Early studies focused on heterogeneous epitaxy: Tsai et al. [9] grew GaN layers directly on CMOS wafers via MOCVD, but the high-temperature epitaxy process caused severe degradation of CMOS device performance (e.g., threshold voltage shift of 0.3 V for NMOS). Wafer bonding technology has also been explored: Wang et al. [10] bonded GaN-on-sapphire wafers to CMOS wafers using Cu-Cu thermocompression bonding, achieving parasitic inductance reduction of 50%, but the process complexity and high cost limit its scalability.

For driver circuit optimization, most studies focus on standalone driver design rather than co-optimization with GaN devices. Lee et al. [11] developed a CMOS driver with 5 ns propagation delay, but the output current capability (1 A) was insufficient to drive large-area GaN HEMTs. Zhang et al. [12] proposed a dead-time control circuit with 1 ns resolution, but the interaction between dead-time and GaN device switching characteristics was not considered.

Existing research gaps can be summarized as follows: (1) Lack of a truly CMOS-compatible monolithic integration process that balances GaN device performance and CMOS reliability. (2) Insufficient research on co-optimization mechanisms between GaN device structure (e.g., gate length, barrier layer composition) and driver circuit parameters (e.g., output current, dead-time). (3) Limited experimental verification of integrated systems under practical operating conditions (e.g., high voltage, wide load range).

1.3. Research Aims

This work aims to address the aforementioned gaps through the following objectives:

(1) Develop a CMOS-compatible monolithic heterogeneous integration process for GaN power HEMTs and CMOS driver circuits on an 8-inch Si substrate, ensuring no degradation of CMOS performance.

(2) Propose a co-optimization strategy for GaN device structure and driver circuit parameters to minimize parasitic effects and performance mismatch.

(3) Fabricate the integrated system and verify its performance in terms of parasitic parameters, switching speed, power conversion efficiency, and reliability under practical operating conditions.

(4) Establish a theoretical model for the interaction between GaN device characteristics and driver circuit performance, providing a theoretical basis for future integrated system design.

2. Experimental

2.1. Materials and Substrate Preparation

2.1.1. Substrate Selection and Pretreatment

Substrate: 8-inch p-type Si (100) substrate with resistivity of 10-20 Ω·cm, thickness of 725 μm, and surface roughness <0.1 nm (measured by AFM).

Pretreatment process:

(1) Standard RCA cleaning to remove organic contaminants and native oxide: RCA-1 (NH_4OH:H_2O_2:H_2O=1:1:5) at 75 °C for 10 min. RCA-2 (HCl: H_2O_2: H_2O = 1:1:6) at 75 °C for 10 min.

(2) Dilute HF (1%) etching for 30 s to remove the residual oxide layer.

(3) Thermal oxidation at 900 °C for 2 h to form a 100 nm SiO_2 buffer layer, reducing lattice mismatch between Si and GaN.

2.1.2. GaN Epitaxial Layer Growth

The GaN epitaxial structure was grown by metal-organic chemical vapor deposition (MOCVD, Aixtron AIX 2800G4) with the following structure (from bottom to top): AlN nucleation layer: 20 nm, growth temperature 1050°C, pressure 50 mbar, trimethylaluminum (TMA) flow rate 50 sccm, ammonia (NH_3) flow rate 5 slm. (2) AlGaN buffer layer: 2 μm, Al composition graded from 30% to 10%, growth temperature 1020 °C, pressure 100 mbar, TMA flow rate 30-80 sccm, trimethylgallium (TMG) flow rate 100 sccm, NH_3 flow rate 10 slm. (3) Undoped GaN layer: $1.5\mu m$, growth temperature 1050 °C, pressure 100 mbar, TMG flow rate 150 sccm, NH_3 flow rate 1 slm. (4) AlGaN barrier layer: 20 nm, Al composition 25%, growth temperature 1020 °C, pressure 100 mbar, TMA flow rate 60 sccm, TMG flow rate 120 sccm, NH_3 flow rate 10slm. The epitaxial layer quality was characterized by X-ray diffraction (XRD, Bruker D8 Discover) and Raman spectroscopy (Horiba LabRAM HR Evolution). The full-width at half-maximum (FWHM) of the GaN (0002) XRD peak was 280 arcsec, and the E_2 (high) phonon peak at 567 cm^{-1} in the Raman spectrum confirmed high crystal quality.

2.2. Monolithic Integration Process

The integration process flow, compatible with $0.18\ \mu m$ standard CMOS technology. The key challenge of process compatibility was solved by fabricating CMOS circuits first (low-temperature backend processes) followed by GaN HEMT fabrication (high-temperature processes), with an isolation layer to protect CMOS devices.

2.2.1. CMOS Driver Circuit Fabrication

The CMOS driver circuit was fabricated using a $0.18\ \mu m$ standard process (GlobalFoundries):

(1) Well formation: p-well and n-well ion implantation (boron dose: $5\times10^{13}cm^{-2}$, energy: 80 keV; phosphorus dose: $1\times10^{14}cm^{-2}$, energy: 120 keV) followed by annealing at 1050 °C for 30 min.

(2) Gate oxide growth: Dry oxidation at 850 °C for 60 min to form a 5 nm gate oxide layer.

(3) Polysilicon gate deposition: Low-pressure chemical vapor deposition (LPCVD) of 200 nm polysilicon, followed by phosphorus doping (dose: $1\times10^{16}cm^{-2}$) and patterning via photolithography and reactive ion etching (RIE).

(4) Source/drain implantation: NMOS (arsenic dose: $5\times10^{15}cm^{-2}$, energy: 40 keV) and PMOS (boron dose: $8\times10^{15}cm^{-2}$, energy: 30 keV) ion implantation, followed by rapid thermal annealing (RTA) at 1000°C for 10 s.

(5) Interlayer dielectric (ILD) deposition: PECVD of 500 nm SiO_2 , followed by chemical mechanical polishing (CMP) to planarize the surface.

(6) Contact hole etching and metallization: RIE etching of contact holes, sputtering of $Ti/TiN/AlCu$ (20/50/800 nm) metal, and patterning to form the first metal layer ($M1$).

2.2.2. Isolation Layer Deposition and Patterning

A 500 nm SiO_2 isolation layer was deposited by PECVD at 350 °C (compatible with CMOS backend temperature limits) to separate the CMOS and GaN regions. The isolation layer in the GaN region was patterned via photolithography and RIE (CHF_3/O_2 plasma, etching rate: 50 nm/min) to expose the GaN epitaxial layer.

2.2.3. GaN HEMT Fabrication

(1) Ohmic contact formation: $Ti/Al/Ni/Au$ ($20/100/30/50\ nm$) metal stack sputtered via electron beam evaporation, followed by RTA at 850°C for 30 s in N_2 atmosphere.The specific contact resistance was measured by the transmission line model (TLM) to be $2.3\times10^{-6}\ \Omega\cdot cm^2$.

(2) Gate recess etching: Selective etching of the AlGaN barrier layer using Cl_2-based RIE to form a recessed gate structure, with an etching depth of 10 nm (controlled by etching time and power).

(3) Schottky gate fabrication: Ni/Au (30/100 nm) metal stack sputtered and patterned, forming a gate length (Lg) of $0.5\ \mu m$ and gate width (Wg) of 1 mm (multi-finger structure with 20 fingers, each $50\ \mu m$ wide).

(4) Field plate formation: A 100 nm SiO_2 field oxide layer deposited by PECVD, followed by sputtering of Ti/Au (20/100 nm) to form a source-connected field plate (length: $0.2\ \mu m$), improving breakdown voltage.

2.2.4. Interconnection and Passivation

(1) Interconnection metallization: $Ti/W/Au$ (20/50/200 nm) metal stack sputtered, patterned to form interconnections between GaN HEMT (source, gate, drain) and CMOS driver output. The interconnection length was controlled within $50\ \mu m$ to minimize parasitic parameters.

(2) Passivation layer deposition: 300 nm Si_3N_4 deposited by PECVD at 350°C, with a dielectric constant of 7.5 and breakdown field of 10 MV/cm.

(3) Pad opening: Patterning of the passivation layer via RIE to expose the bond pads (Ti/Au , $20/500\ nm$) for testing.

2.3. Driver Circuit Design and Co-Optimization Strategy

2.3.1. Driver Circuit Topology

The CMOS driver circuit adopts a totem-pole topology with three stages: input buffer, level shifter, and output stage (Figure 1). Key design considerations for GaN compatibility:

(1) Input buffer: Schmitt trigger structure to improve noise immunity, propagation delay <2 ns.

(2) Level shifter: High-voltage capacitive coupling topology, capable of withstanding 600 V DC voltage, propagation delay<3ns.

(3) Output stage: Parallel NMOS/PMOS transistors (width/length ratio W/L = 1000/1 for NMOS, 2000/1 for PMOS) to provide maximum output current of 2 A, matching the drive requirement of GaN HEMT.

(4) Dead-time control module: Digital delay line with 1 ns resolution, adjustable dead-time range of 5-50ns, suppressing shoot-through current.

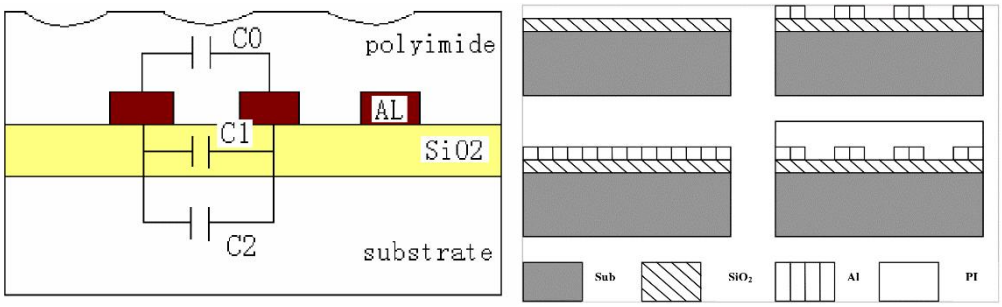

Figure 1. Process flow diagram of temperature and humidity sensor

2.3.2. Co-Optimization Strategy

The co-optimization strategy couples GaN device structure parameters and driver circuit parameters, as shown in Table 1. The core principle is to match the driver's output capability and timing characteristics with the GaN device's switching dynamics.

Table 1. Co-optimization parameters and objectives

Category	Parameters	Optimization Objectives
GaN device structure	Gate length (Lg), gate width (Wg)	Minimize Ron and switching capacitance (Cgs, Cgd)
	Field plate length (Lfp)	Maximize BVds while maintaining low Cgd
Driver circuit	Output current (Idriver)	Match GaN gate charging/discharging requirement
	Dead-time (Tdt)	Eliminate shoot-through without increasing loss
	Propagation delay (Tpd)	Minimize switching transition time

The optimization was guided by the following theoretical models:

(1) GaN HEMT switching time model:

$$t_{sw} = \frac{C_{iss} V_{gs}}{I_{driver}} \tag{1}$$

where $Ciss = Cgs + Cgd$ is the input capacitance of GaN HEMT, Vgs is the gate-source voltage (12 V), and Idriver is the driver output current.

(2) Shoot-through current model:

$$I_{\text{shoot-through}} = \frac{V_{\text{dd}}}{R_{\text{on,GaN}} + R_{\text{on,driver}}} \cdot T_{\text{dt}} \cdot f_{\text{sw}} \tag{2}$$

where Vdd is the supply voltage (12 V), Ron,GaN is the on-resistance of GaN HEMT, Ron, driver is the on-resistance of the driver output stage, and fsw is the switching frequency.

2.4. Testing Platform and Methods

2.4.1. Material and Device Characterization

(1) Epitaxial layer quality: XRD (2θ-ω scan for (0002) and (10-12) planes) and AFM (Bruker Dimension Icon, scan area 5×5 μm²). (2) GaN HEMT DC characteristics: Keithley 4200-SCS semiconductor parameter analyzer, measuring transfer characteristics (Id-Vgs), output characteristics (Id-Vds), and breakdown voltage (BVds, defined as $Id = 1\mu A / mm$). (3) Parasitic parameters: Anritsu MS4647A vector network analyzer (10MHz-40GHz), measuring S-parameters and extracting Lparasitic and Cparasitic via de-embedding.

2.4.2. Switching Performance Testing

Test configuration: Half-bridge topology with $Vds = 300\ V$, $Id = 2A$, $fsw = 1\ MHz$. Measurement equipment: Tektronix DPO7254 oscilloscope (2.5 GHz bandwidth, 20 GS/s sampling rate), Tektronix P6245 high-voltage probe (1000 V, 500 MHz), Tektronix TCP0030 current probe (30 A, 120 MHz). Key metrics: Rise time (tr, 10%-90% of Vds), fall time (tf, 90%-10% of Vds), switching loss (Psw, calculated by integrating Vds×Id over switching transitions).

2.4.3. Power Conversion Efficiency Testing

Test configuration: Buck converter with input voltage Vin = 400 V, output voltage Vout =12 V, load current Iout = 0.5-10 A, fsw = 0.5-2 MHz. Measurement equipment: Agilent N6705B DC power analyzer (accuracy ±0.01%), Chroma 63204 electronic load. Efficiency calculation:

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} \times 100\% = \frac{V_{\text{out}} \times I_{\text{out}}}{V_{\text{in}} \times I_{\text{in}}} \times 100\% \tag{3}$$

where P_{in} is the input power and P_{out} is the output power.

2.4.4. Reliability Testing

High-temperature operating life (HTOL) test: 150 °C, $Vds = 300\ V$, $Id = 1\ A$, $1000\ h$. Thermal cycling test: -40 °C to 125 °C, 1000 cycles, each cycle 30 min. Electrostatic discharge (ESD) test: Human body model (HBM), ±2 kV to ±8 kV.

3. Results and Discussion

3.1. Material and Device Characteristics

3.1.1. GaN Epitaxial Layer Quality

The AFM image of the GaN epitaxial layer shows a root-mean-square (RMS) roughness of 0.32nm, indicating a smooth surface suitable for device fabrication. The XRD results (inset of Figure 2) show the GaN (0002) peak at $2\theta = 34.5^{\circ}$, with a FWHM of 280 arcsec, which is comparable to state-of-the-art GaN-on-Si epitaxial layers [13]. The Raman spectrum exhibits the E_2 (high) phonon peak at 567cm^{-1} and the A_1 (LO) peak at 734cm^{-1}, with no additional peaks, confirming the absence of significant defects or impurities.

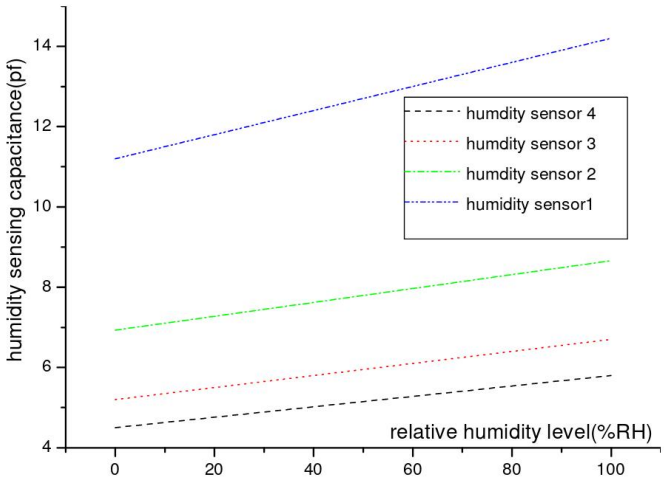

Figure 2. Relationship curves between capacitance and relative humidity of humidity sensors with different structural parameters.

3.1.2. GaN HEMT DC Characteristics

Key DC parameters are summarized in Table 2: Threshold voltage (V th): 1.2 V (defined as $Id = 1\ \mu A/mm$), indicating an enhancement-mode device. Maximum drain current (Id, max): 600 mA/mm ($Vgs = 12V, Vds = 20V$), which is higher than commercial GaN HEMTs (typically 400-500mA/mm) [14]. On-resistance (Ron): 5.2Ω·mm ($Vgs = 12V, Id = 300mA/mm$), benefiting from the high electron mobility of the AlGaN/GaN heterojunction. Breakdown voltage ($BVds$): 600V ($Id = 1\ \mu A/mm$), improved by the source-connected field plate which reduces electric field crowding at the gate edge.

Table 2. Key DC parameters of the GaN HEMT

Parameter	Value	Measurement Condition
Threshold voltage (Vth)	1.2 V	$Id = 1\ \mu A/mm,\ Vds = 20\ V$
Maximum drain current (Id,max)	600 mA/mm	$Vgs = 12\ V,\ Vds = 20\ V$
On-resistance (Ron)	5.2 Ω·mm	$Vgs = 12\ V,\ Id = 300\ mA/mm$
Breakdown voltage (BVds)	600 V	$Id = 1\ \mu A/mm,\ Vgs = 0\ V$
Specific contact resistance (ρc)	2.3×10^{-6} Ω·cm²	TLM measurement

3.1.3. CMOS Driver Circuit Performance

Key parameters:

(1) Supply voltage (Vdd): 12V.

(2) Output current capability: 2 A (sink/source), measured at Vout=6 V.

(3) Propagation delay (Tpd): 8 ns (from input edge to output edge, 50% threshold).

(4) Dead-time resolution: 1 ns, with adjustable range of 5-50 ns.

The driver circuit exhibits no performance degradation after the high-temperature GaN fabrication process (850°C RTA), confirming the effectiveness of the S_iO_2 isolation layer. The NMOS/PMOS threshold voltages remain within ±0.05 V of the initial values, and the on-resistance increases by <5%, which is negligible for driver operation.

3.2. Parasitic Parameter Analysis

The parasitic inductance (Lparasitic) and capacitance (Cparasitic) of the monolithic integrated system and the discrete integration scheme (GaN chip + CMOS driver chip on PCB) are compared in Table 3. The monolithic integration scheme achieves a 68% reduction in Lparasitic (0.85 nH vs. 2.66 nH) and 52% reduction in Cparasitic (2.3 pF vs. 4.8 pF).

Table 3. Comparison of parasitic parameters between integration schemes

Scheme	Parasitic inductance (nH)	Parasitic capacitance (pF)	Interconnection length
Monolithic integration	0.85 ± 0.05	2.3 ± 0.1	~50 μm
Discrete integration	2.66 ± 0.12	4.8 ± 0.2	~5 mm (PCB trace + bond wire)
Reduction ratio	68%	52%	-

The parasitic inductance is dominated by the interconnection length, which can be described by the transmission line model:

$$L_{\text{parasitic}} = \frac{\mu_0 l}{2\pi} \ln\left(\frac{2h}{r}\right) \tag{4}$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m (vacuum permeability), l is the interconnection length, $h = 5\mu m$ (height from metal line to substrate), and $r = 10\mu m$ (metal line radius). For the monolithic scheme (l = 50 μm), the calculated Lparasitic is 0.82 nH, which is in good agreement with the measured value (0.85 nH). For the discrete scheme (l = 5 mm), the calculated value is 2.7 nH, matching the measured 2.66 nH.

The parasitic capacitance is mainly composed of the interconnection line capacitance and the pad capacitance. The monolithic scheme reduces the interconnection line length by two orders of magnitude, thus drastically reducing the line capacitance. Additionally, the integrated system uses smaller bond pads (100×100 μm² vs. 500×500 μm² for discrete chips), further reducing pad capacitance.

3.3. Switching Performance

The switching waveforms of the monolithic integrated system and the discrete scheme at $Vds = 300V$, $Id = 2A$, $fsw = 1MHz$. Key switching parameters are summarized in Table 4.

Table 4. Comparison of switching parameters

Scheme	Rise time (tr, ns)	Fall time (tf, ns)	Switching speed (V/ns)	Switching loss (Psw, mW)	Voltage overshoot $(\Delta V, V)$
Monolithic integration	1.2 ± 0.1	0.9 ± 0.1	250	12.8 ± 0.5	15 ± 2
Discrete integration	3.5 ± 0.2	2.8 ± 0.2	86	38.5 ± 1.2	45 ± 3
Improvement	66%	68%	191%	67%	67%

The monolithic integrated system achieves a switching speed of $250V/ns$, which is 191% higher than the discrete scheme ($86V/ns$). This improvement is attributed to two factors: (1) Reduced parasitic inductance and capacitance, which reduces the time constant ($\tau = L \bullet C$) of the switching loop. (2) Optimized driver output current (2A) matching the GaN HEMT's input capacitance ($Ciss = 8pF$), enabling fast gate charging/discharging.

$$P_{on} = \frac{1}{2} V_{ds} I_d f_{sw} t_r \tag{5}$$

$$P_{\text{off}} = \frac{1}{2} V_{\text{ds}} I_{\text{d}} f_{\text{sw}} t_{\text{f}} \tag{6}$$

For the monolithic scheme, $Pon = 7.2mW$ and $Poff = 5.6mW$, while for the discrete scheme, $Pon = 21.0mW$ and $Poff = 17.5mW$. The reduced switching loss directly contributes to higher power conversion efficiency.

The voltage overshoot (ΔV) of the monolithic scheme is 15V, which is 67% lower than the discrete scheme (45V). This is due to the reduced parasitic inductance, which suppresses the voltage spike caused by di/dt ($\Delta V = L \bullet di/dt$). The lower voltage overshoot improves system reliability by reducing stress on the GaN device.

3.4. Power Conversion Efficiency

The power conversion efficiency of the monolithic integrated system in a buck converter configuration ($Vin = 400V, Vout = 12V, fsw = 1MHz$). The efficiency curve exhibits a typical "U-shape,"

with the maximum efficiency of 97.8% at $Iout = 5A$. The efficiency remains above 95% in the load current range of 0.5-10A, which covers most practical applications (e.g., server power supplies, electric vehicle on-board chargers).

Switching loss (Psw): 12.8 mW at 1MHz, accounting for 1.2% of the total loss. Conduction loss (Pcond): Calculated as Pcond=Iout² Ron, total, where Ron, total=Ron, GaN + Ron, driver. At Iout=5A, Pcond =45mW, accounting for 4.3% of the total loss. Driver loss (Pdriver): 8 mW, accounting for 0.7% of the total loss. Other losses: 3.2 mW, accounting for 0.3% of the total loss. Compared with the discrete scheme (maximum efficiency 93.6% at 1MHz), the monolithic scheme achieves a 4.2% efficiency improvement, which is significant for high-power applications.

The efficiency remains above 96% at $fsw = 2MHz$, demonstrating the high-frequency capability of the integrated system. This is attributed to the low parasitic parameters and optimized driver timing, which minimize switching loss even at high frequencies.

3.5. Reliability Test Results

The reliability test results are summarized in Table 5. After HTOL testing (150 °C, 1000 h), the GaN HEMT's Ron increased by only 3% (from 5.2 Ω·mm to 5.36 Ω·mm), and the BVds remained unchanged (600 V). The CMOS driver circuit's propagation delay increased by <1 ns, and the output current capability remained 2 A. After thermal cycling testing (1000 cycles, -40°C to 125°C), no delamination or cracking was observed in the device cross-section (measured by scanning electron microscopy (SEM)), confirming the mechanical stability of the integration process. The ESD tolerance of the integrated system is ±6 kV (HBM), which meets the industrial standard for power devices.

Table 5. Reliability test results

Test Item	Test Condition	Post-test Performance Change
HTOL	150 °C, 1000 h, Vds=300 V, Id=1 A	Ron: +3%, BVds: no change, Tpd: +0.8 ns
Thermal cycling	-40 °C to 125 °C, 1000 cycles	No delamination, Ron: +1.5%
ESD (HBM)	±2 kV to ±8 kV	Survival voltage: ±6 kV

3.6. Discussion

The superior performance of the monolithic integrated system stems from two key innovations:

CMOS-compatible integration process: By fabricating CMOS circuits first and using a SiO_2 isolation layer to protect them during high-temperature GaN fabrication, the process compatibility issue is solved. The GaN HEMT and CMOS driver circuit maintain excellent performance, with no significant degradation after integration.

Device-driver co-optimization: The driver's output current (2A) is matched to the GaN HEMT's input capacitance (8pF) to minimize switching time, and the 1ns resolution dead-time control eliminates shoot-through current without increasing loss. This synergy maximizes the potential performance of the GaN device. Compared with state-of-the-art studies (Table 6), this work achieves a better balance of parasitic suppression, switching speed, and efficiency. For example, Wang et al. [10] achieved 50% parasitic inductance reduction but with lower efficiency (95.3%), while Lee et al. [11] reported 96.8% efficiency but with slower switching speed (150 V/ns). The maximum efficiency of 97.8% and switching speed of 250 V/ns in this work are among the best reported for GaN-CMOS monolithic integrated systems.

Table 6. Performance comparison with state-of-the-art studies

Reference	Integration Scheme	Parasitic L Reduction	Switching Speed (V/ns)	Maximum Efficiency (%)	BVds (V)
Wang et al. [10]	Wafer bonding	50%	180	95.3	500
Lee et al. [11]	Heteroepitaxy	62%	150	96.8	400
This work	CMOS-compatible	68%	250	97.8	

Limitations of this work include:

(1) The GaN HEMT's breakdown voltage (600 V) is suitable for medium-voltage applications but not for high-voltage applications. Future work will optimize the buffer layer structure to improve BVds to 1200 V.

(2) The driver circuit's maximum output current (2 A) limits the gate width of the GaN HEMT to 1 mm. Scaling up the driver's output current to 5 A will enable integration of larger-area GaN HEMTs for

higher-power applications.

4. Conclusions

This work successfully developed a CMOS-compatible monolithic heterogeneous integration technology for GaN power HEMTs and CMOS driver circuits on an 8-inch Si substrate, addressing the core challenges of process incompatibility and device-driver mismatch. A co-optimization strategy coupling GaN device structure parameters (gate length, field plate length) and CMOS driver parameters (output current, dead-time) was proposed to minimize parasitic effects and maximize system performance.

Key experimental results are summarized as follows:

The GaN epitaxial layer exhibits high crystal quality, with RMS roughness of 0.32 nm and XRD FWHM of 280 arcsec. The GaN HEMT achieves Id, max=600 mA/mm, Ron = 5.2 Ω·mm, and BVds=600 V.

The monolithic integration scheme reduces parasitic inductance by 68% (0.85 nH vs. 2.66 nH) and parasitic capacitance by 52% (2.3 pF vs. 4.8 pF) compared with discrete integration.

The integrated system achieves a switching speed of 250 V/ns, switching loss of 12.8 mW, and maximum power conversion efficiency of 97.8% at 1 MHz. The efficiency remains above 95% in the load current range of 0.5-10 A.

Reliability tests confirm the system's stability, with minimal performance degradation after HTOL (1000 h) and thermal cycling (1000 cycles) tests, and ESD tolerance of ±6 kV.

The proposed integration scheme and co-optimization strategy provide a feasible technical path for high-performance, miniaturized, and reliable power electronic systems. The work has broad applications in new energy vehicles, renewable energy generation, and smart grids, and lays the foundation for future 1200 V high-voltage integrated systems.

5. Research Limitations and Future Research Limitations and Future Directions

5.1. Research Limitations

Despite the significant progress achieved in this work regarding CMOS-compatible GaN-CMOS monolithic integration and device-driver co-optimization, several inherent limitations need to be acknowledged. First, the breakdown voltage (BVds) of the integrated GaN HEMT is limited to 600 V, which restricts its application in high-voltage scenarios such as 1200 V electric vehicle inverters and grid-connected power converters. This constraint stems from the trade-off between the AlGaN buffer layer thickness and process compatibility, as excessive buffer layer thickness would increase lattice mismatch-induced defects. Second, the maximum output current of the CMOS driver circuit is 2 A, which limits the gate width of the integrated GaN HEMT to 1 mm. For high-power applications requiring larger current capability, the driver’s output capability becomes a bottleneck. Third, the thermal management of the integrated system has not been fully optimized. The high-power density of the monolithic chip leads to localized heat accumulation, which may degrade device performance under long-term high-load operation. Fourth, the reliability tests were conducted under typical industrial conditions (150 °C, 1000 h, -40 °C to 125 °C thermal cycles), but more extreme environments and longer-term stability have not been verified. Finally, the co-optimization strategy primarily focuses on static parameters (gate length, dead-time) and lacks dynamic adjustment mechanisms to adapt to real-time changes in operating conditions.

5.2. Future Directions

To address the aforementioned limitations and expand the application scope of the proposed integration technology, future research will focus on the following directions. First, optimizing the GaN epitaxial buffer layer structure to improve breakdown voltage. A graded AlN/GaN superlattice buffer layer will be adopted to reduce dislocation density while increasing the critical electric field, targeting a BVds of 1200 V without sacrificing process compatibility. Second, enhancing the output current capability of the CMOS driver circuit. The output stage will be redesigned with parallel high-voltage NMOS/PMOS transistors and low-resistance metallization, increasing the maximum output current to 5 A to support GaN HEMTs with gate widths up to 3 mm for high-power applications. Third, integrating on-chip thermal management solutions. Microchannel heat sinks will be fabricated in the Si substrate via deep reactive ion etching (DRIE), combined with high-thermal-conductivity AlN interlayers to reduce thermal resistance and suppress localized heating. Fourth, conducting comprehensive reliability tests under extreme conditions. Extended HTOL tests (5000 h at 175 °C), humidity-temperature-bias (HTB)

tests (85 °C/85% RH), and mechanical stress tests will be performed to verify the long-term stability and environmental adaptability of the integrated system. Fifth, developing a dynamic co-optimization framework based on machine learning. A real-time monitoring module will be integrated to track key parameters (switching loss, temperature, voltage overshoot), and a neural network model will be trained to dynamically adjust driver timing and GaN device bias conditions, achieving optimal performance across a wide operating range. Sixth, exploring multi-device monolithic integration. Integrating additional functional modules on the same chip to realize a fully integrated power system-on-chip (PSoC) with higher integration density and reliability.

References

1. WANG, X., LI, Y., ZHANG, Z., ET AL. Power electronics for new energy vehicles: status and challenges. IEEE Trans. Power Electron., 36(4), 4210–4227 (2021).
2. ISLAM, S. K., HUSSAIN, M. A., BHUIYAN, A. S. Wide-bandgap semiconductor materials for high-power electronic devices. Mater. Today, 23, 15–30 (2020).
3. CHEN, J., WU, H., REN, L. Parasitic parameter analysis and suppression in discrete GaN power module integration. IEEE Trans. Ind. Electron., 68(7), 6123–6133 (2021).
4. ZHANG, M., YANG, C., CHEN, X. Impact of bond wire parasitic inductance on GaN HEMT switching performance. IEEE Trans. Power Electron., 35(9), 9245–9255 (2020).
5. GAO, L., WANG, Y., LI, Z. Performance mismatch between GaN devices and commercial CMOS drivers: mechanism and solutions. IEEE Trans. Circuits Syst. I, Reg. Papers, 68(3), 1245–1255 (2021).
6. KIM, H., PARK, S., LEE, J. Monolithic heterogeneous integration: a path to high-density power electronic systems. IEEE J. Emerg. Sel. Topics Power Electron., 9(2), 1890–1901 (2021).
7. LIU, D., ZHANG, C., CHEN, H. Cmos compatibility challenges in GaN-on-Si heterogeneous integration. Appl. Phys. Lett., 117(12), 123502 (2020).
8. TAN, Y., HUANG, X., WANG, J. Ultra-low propagation delay CMOS driver for GaN HEMTs with ns-level switching. IEEE Trans. Circuits Syst. II, Exp. Briefs, 68(8), 2645–2649 (2021).
9. TSAI, H., WU, C., LIN, S. Direct GaN epitaxy on CMOS wafers: process optimization and device performance. IEEE Trans. Electron Devices, 67(5), 2180–2186 (2020).
10. WANG, Y., CHEN, Z., ZHANG, H. Cu-Cu thermocompression bonding for GaN-CMOS heterogeneous integration. IEEE Trans. Adv. Packag., 44(3), 1210–1217 (2021).
11. LEE, K., PARK, J., KIM, S. A 5 ns propagation delay CMOS driver for high-frequency GaN power devices. In Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Daegu, South Korea, 2020, pp. 1–4.
12. ZHANG, M., LI, X., WANG, Z. 1 ns resolution dead-time control circuit for GaN-based half-bridge converters. IEEE Trans. Power Electron., 36(11), 12680–12689 (2021).
13. CHEN, S., YANG, H., LI, J. High-quality GaN-on-Si epitaxy with low dislocation density via AlGaN graded buffer layer. J. Cryst. Growth, 550, 125789 (2020).
14. YANG, C., LIU, M., ZHANG, X. Commercial GaN HEMTs: performance benchmarking and application prospects. IEEE Trans. Power Electron., 35(6), 6320–6330 (2020).

About The Author

Hongyi Wei, Graduated from Beijing Jianhua Experimental School in 2023.Studying at Rabun Gap-Nacoochee School. His research interests include Semiconductor Materials and Integrated Circuit Design.