Abstract—In this work, highly linear AlGaN/GaN laterally gated (or buried gate) high-electron-mobility transistors (HEMTs) are reported. The effect of gate dimensions on source-access resistance and the linearity of laterally gated devices are investigated experimentally in detail for the first time. Transistors with different gate dimensions and conventional planar devices are fabricated using two-step electron beam lithography (EBL). Current–voltage, source-access resistance, small-signal, and two-tone measurements are performed to evaluate the linearity of devices. Contrary to conventional planar HEMTs, the intrinsic transconductance of laterally gated devices monotonically increases with increasing gate voltage, showing a similar behavior as junction field-effect transistors (JFETs). The source-access resistance shows a polynomial increase with the drain current, which can be reduced by decreasing the filling ratio of the buried gates. Through the optimization of these two competing factors, i.e., intrinsic transconductance and the source-access resistance, flat transconductance and high linearity is achieved experimentally. The laterally gated structure shows flat transconductance and small-signal power gain over a larger span of gate voltage that is 2.5 times higher than a planar device. Moreover, 6.9-dB improvement in output intercept point (OIP3)/PDC is achieved. This approach can be used to improve the linearity of AlGaN/GaN HEMTs at the device level.

Index Terms—AlGaN/GaN high-electron-mobility transistors (HEMTs), buried gate, field-effect transistors (FETs), FinFET, FinHEMT, laterally gated HEMT, linearity, transconductance, tri-gate.

I. INTRODUCTION

GaN high-electron-mobility transistor (HEMT) devices are suitable choices for high frequency and power applications, thanks to their high mobility and high breakdown voltages [1]–[3]. However, due to the nonlinear transconductance characteristics, the gain is dependent on bias conditions, which is disadvantageous for linearity and noise performance [4], [5]. Increasing demand for low noise amplifiers with 5G technology, especially for multi–in multi-out (MIMO) amplifiers in the base stations, motivates studies to improve the linearity performance of the devices [6]. The source of the nonlinearity of transconductance is attributed to the early saturation of the electron mobility with the electric field and resultant nonlinear source-access resistance [7], [8]. In order to overcome this effect, self-aligned structures are proposed and showed superior linearity, but they suffer from low breakdown voltages that make them unsuitable for high-power applications [9]. Graded-channel HEMTs show big progress in linearity over conventional HEMTs by using the formation of three-dimensional electron gas (3DEG) [10], [11]. Multichannel structures are shown to provide good linearity performance by increasing the number of channels where the current flows and by decreasing the effect of source resistance [12]. N-polar GaN HEMTs also support very high linearity figures of merit (FOMs) even at high frequencies [13], [14].

In addition to these solutions, three-dimensional-gate (3-D-gate) HEMTs are presented, offering enhancement in linearity with acceptable device performance [15]–[17]. These devices are an implementation of fin field-effect transistor (FinFET) topology in the CMOS technology to GaN devices. In conventional planar HEMTs, the gate is above the semiconductor surface and gate bias modulates the electric field to control the channel conductivity. In normally ON devices, a negative gate bias is required to deplete the two-dimensional electron gas (2DEG) and to close the channel.
In 3-D-gate devices, the gate covers the channel from three sides. Additional horizontal depletion is achieved by buried gates in addition to a gate above nonetched regions, which provides better gate control and reduced short channel effects [18], [19]. In addition, it provides a higher ON-/OFF-current ratio [20], [21] or less subthreshold leakage [22].

3-D sketches of planar and laterally gated HEMTs are shown in Fig. 1(a) and (b), respectively. Planar devices are conventional HEMTs where the gate is placed between the source and drain perpendicular to the current direction as a solid body. It stays on top of the epitaxy and is generally coated with passivation material. The gate modulates the channel from the top and controls the conductivity of the channel. In the laterally gated structure, Fig. 1(b), the gate is positioned between the drain and source contacts like the planar device. The gate metal, however, has periodic comb-like extensions penetrating into the substrate down to the GaN layer. Because the gate metal passes through 2DEG vertically and reaches the GaN layer, 2DEG at the regions where metal exists is absent. Therefore, current passes through the 2DEG which remains between the teeth of the comb metal. The current to control the current passing from the drain to source with lateral depletion [20]. The surface is coated with passivation material like SiN and the gate metal does not touch the epitaxy from the top. This creates the difference between tri-gate and laterally gated structures. The laterally gated structure eliminates the Schottky contact on the top, so lateral depletion is the dominant factor in channel control.

The cross section of a periodic part of the structure, parallel to 2DEG, is given in Fig. 2(a). The current passing from the drain to the source and the lateral depletion of gate metals are shown. The region between the source and the gate contacts is named as the source access region and its resistance is called source-access resistance ($R_S$). It is a function of drain current and is shown in Fig. 2(a) as a series resistance in the channel. $W_C$ represents the width of the channel that remains between the combs of the gate metal. The current is passing through this region. $W_M$ represents the width of metals that penetrates into the substrate. $W_M$ and $W_C$ constitute the period of this structure. The filling ratio (FR) is defined as the ratio of $W_C$ to the period of the structure ($W_M + W_C$) in Fig. 2(a), and cross-sectional view of the device on the line AA’ of Fig. 2(a) is indicated in Fig. 2(b). The gate metal reaches the GaN substrate, and the 2DEG is eliminated at this region. Depletion of the 2DEG can be seen near the gate metal. Fig. 2(c) shows a cross section taken in line BB’ in Fig. 2(a). In this cross section, the gate metal stays on top of the passivation. This increases the distance of the gate contact to the channel in the nonetched region and reduces the effect of the vertical depletion. The cross section of one period of the structure perpendicular to the current direction is provided in Fig. 2(d), taken in CC’ line in Fig. 2(a). SiN forms a layer between the semiconductor and gate metal in the nonetched region. The teeth of the gate reach the GaN. 2DEG exists between these teeth and the horizontal depletion modulates the electron concentration, thereby controlling the current flow in 2DEG.

Referring to uniformly doped field-effect transistors (FETs) [26], the intrinsic transconductance in the saturation region, excluding the source and drain side access region resistances, can be written as (1) where $n_s$ is 2DEG charge density in cm$^{-2}$, $W_C$ is channel width, $L_G$ is gate length, $V_B$ is built-in...
potential of gate-channel junction, and $V_G$ is the gate potential. $N_x$ is the peak channel charge density in cm$^{-3}$, which is approximately equal to $n_i/\Delta d$, where $\Delta d$ is the effective width of the 2DEG charge distribution along the growth direction.

$$g_{m,i} = \frac{q \mu_n n_i W_C}{L_G} \left[ 1 - \frac{8 \varepsilon_r \varepsilon_0 (V_B - V_{GS})}{q N_x W_C} \right]$$  \hspace{1cm} (1)

$$V_{GS} = \frac{V_G - I_D \times R_S}{L_G}$$  \hspace{1cm} (2)

$$R_S = \frac{q \mu_n n_i (W_C + W_M)}{g_{m,i}}$$  \hspace{1cm} (3)

$$g_{m,ext} \approx \frac{g_{m,i} (1 - I_D \frac{dR}{dV_{GS}})}{(1 + g_{m,i} R_S)}$$  \hspace{1cm} (4)

The gate to source voltage is accepted as the difference between the biases of the gate and source contacts. However, as formulated in (2), a voltage drop occurs in the source access region due to nonzero source-access resistance ($R_S$) and the channel current. Therefore, the real gate to source voltage deviates from the applied biases, dependent on $R_S$ and drain current. The source-access resistance can be formulated as (3) and is also not constant due to the dependence of mobility ($\mu_n$) on the drain current and electric field. The changes in the gate to source voltage and $R_S$ result in extrinsic transconductance $g_{m,ext}$ as given in (4). Of the parasitic components, the extrinsic transconductance of the device is less than the intrinsic transconductance.

In 3-D-gate transistors, the channel width at gate to source region is more than effective channel width, where the current flows in the region between the teeth of the gate metal, as can be seen in Fig. 2(a). This decreases the access resistances and electric field in the access region. As a result, lower $R_S$ and higher over-drive voltage are achieved [16], [23], [27]. Not only the magnitude but also the nonlinear behavior of the source access region changes with the FR. It is shown in Fig. 2(a) that the structure is similar to the junction FETs (JFETs). In both structures, the gate is controlling the channel from two sides. Therefore, depletion mechanisms and the formulation show similarity with JFET devices.

The modified $R_S$ and the JFET-like intrinsic transconductance of the laterally gated devices can be manipulated to observe the linear external transconductance of the devices in a wider gate bias span [28]. In addition, the small values of first- and second-order derivatives of the transconductance with respect to gate voltage ($g_m' = dI_D/dV_{GS}$, $g_m'' = d^2I_D/dV_{GS}^2$) indicate better linearity performance [29], [30]. Both of these FOMs are addressed in our design. These are supported with small-signal and two-tone measurements.

Although previous works show the effects of gate parameters on device performance in tri-gate and metal-oxyde–semiconductor HEMT (MOSHEMT) devices [15]–[17], [27], [28], [31], [32], the effect of the gate dimensions on $R_S$ and the linearity of laterally gated devices are not studied experimentally in detail in earlier works.

In this work, laterally gated devices with different channel and metal widths are fabricated, the transconductance and source-access resistances are compared, and the results in terms of transconductance linearity are explained by using basic semiconductor device equations. It is shown that the linearity of the transconductance can be improved by changing the gate dimensions in laterally gated devices and there is an optimum point for the design parameters. This is the point where the balance between nonlinear $R_S$ and the JFET like intrinsic transconductance is satisfied. Small-signal gain linearity measurements show a match with transconductance and the gain. In two-tone measurements at 3–4 GHz range, 6.9-dB improvement in OIP3/$P_{DC}$ is achieved by using laterally gated HEMTs over conventional planar ones.

II. DEVICE FABRICATION

An AlGaN/GaN HEMT epitaxial structure is grown on 4H-SiC substrate using a metal–organic chemical vapor deposition (MOCVD) system. 1.2-μm thick, $1 \times 10^{17}$ cm$^{-3}$ carbon-doped GaN buffer is grown after a 20-nm AlN nucleation layer. 200-nm transition and 130-nm channel GaN layers are placed on top. Finally, 1-nm AlN spike and 20-nm thick Al$_{0.23}$Ga$_{0.76}$N barrier layer are grown, resulting in $\mu_n = 1697$ cm$^2$/V-s and $n_i = 1.04 \times 10^{13}$ cm$^{-2}$.

Mesa isolation patterns are defined by optical lithography and Cl-based 100-nm etching with inductively coupled plasma reactive ion etching (ICP-RIE). Ti/Al/Ni/Au ohmic contacts being deposited by an e-beam evaporator with 12 nm/120 nm/35 nm/65-nm thickness, respectively. Contact and sheet resistances are measured as 0.19 Ω·nm and 348 Ω/□ using TLM patterns. Until this point, the process is the same for conventional and laterally gated devices.

For laterally gated devices, 25-nm SiN is deposited by plasma-enhanced chemical vapor deposition (PECVD). Electron beam lithography (EBL) is used to define gate patterns with 100 keV. First, SiN is etched with ICP-RIE by F-based chemistry. Then, Cl-based etching of the epoxih is done to achieve a 50-nm depth to eliminate the 2DEG. Before second EBL, the sample surface is cleaned with O$_2$-plasma in an Asher system, and surface treatment is done with diluted NH$_4$OH. A second EBL step is performed for defining the gate heads. 50-nm Ni and 300-nm Au gate metals are deposited and lifted-off. Due to the lateral component of dry etch, different SiN thicknesses result in slightly different etch profiles of the buried metals. In Fig. 3, a fabricated laterally gated device is shown. The current is flowing from the drain to source contact, labeled as D and S, respectively. The gate has periodically etched regions and the gate metal is deposited on top of it. In order to see the cross section, focused ion beam (FIB) is used to cut the structure along and perpendicular to the direction of current flow. The teeth of the gate and SiN between the top contact and semiconductor surface are shown in Fig. 2(d). The conventional planar HEMTs are fabricated with the same gate lengths and EBL parameters, but the gate etch step is missing.

In order to check the effect of the top contact on the channel control, two separate samples with 25 and 50-nm SiN thicknesses are fabricated. It is observed that the lateral depletion is the dominant depletion mechanism for the reported devices, in agreement with the literature [20]. We have chosen 25 nm
thickness for further fabrications because of the aspect ratio issues in the etching process.

The effect of the device structure on transconductance behavior is investigated by using devices with different gate dimensions; all having a source to drain distance ($L_{SD}$) of 4 μm and a gate length ($L_G$) of 250 nm. All of the gates are placed at the midpoint of source–drain spacing. A gate coupled field plate (or T-gate) is used to increase the breakdown voltages and to make the devices more compatible for commercial use. The gate head length is constant and is 600 nm for all of the devices. One and two fingers conventional and laterally gated devices are fabricated with a gate width of 125 μm.

In order to observe the effect of channel width ($W_C$) on device performance, it is swept from 100 to 300 nm with 50-nm steps while all other parameters are kept constant. In addition, the metal width is investigated by fabricating devices with $W_M$ (Fig. 2) from 100 to 300 nm with 50-nm steps and constant $W_C$. Conventional planar T-gate devices with the same gate lengths are fabricated for comparison.

### III. RESULTS AND DISCUSSION

In order to observe the device performance, $I_D$–$V_{DS}$, $I_D$–$V_{GS}$, and $R_S$ measurements are taken from each device. For the measurements, Keysight B1505A is used with a 3-contact probe station. For $R_S$ measurements, a positive current of 0.5 mA/mm is applied to the gate contact while the source contact is kept at ground. The drain voltage is increased until the drain current saturates. The derivative of the gate voltage with respect to the drain current is accepted as $R_S$ [8].

In these measurements, the voltage drop in the Schottky barrier of the gate contact is assumed to be independent of the drain current. The access resistance is accepted to be the main component of $R_S$ and this assumption is tested with gate forward $I$–$V$ measurements. Same gate currents are observed from devices with the same period but different $W_M$ values, which shows the dominant resistive component of $R_S$ is the access region resistance. In these measurements, currents are normalized to the effective channel width, which is calculated as $W_{Eff} = \text{number of channels} \times W_{\text{Channel}}$ [16], [28].

#### A. Effect of Channel Width

To observe the impact of the channel width on performance, the $I_D$–$V_{DS}$ behaviors of laterally gated and planar devices are measured as shown in Fig. 4. The decrease in the ON-resistance ($R_{on}$) can be observed in the linear region. This is due to the decreased FR with respect to the planar device. In addition, the laterally gated device shows a very early knee voltage ($\sim 2.5$ V) versus the planar device ($\sim 4.5$ V), which is better to increase the output voltage swing.

As the channel width decreases, as shown in Fig. 5, the threshold voltage of the device shifts toward more positive values due to the lateral depletion mechanism. The gate controls the channel from the sides and, as the $W_C$ increases, more negative voltage is required to fully deplete the channel [27], [34]. The threshold voltage can be changed in the negative or positive direction compared to the planar device by changing the channel width of the laterally gated device. This provides the designer with a big flexibility in design without changing the epitaxy.

In order to compare the linearity of the planar and laterally gated devices, the gate voltage swing (GVS) FOM is used [28], which is the voltage range that the transconductance remains more than 80% of its maximum value. For the planar device, this is found to be 1.8 V, but for the laterally gated device with a 300-nm channel, it is calculated as more than 6.4 V, which is a $\approx 2.5$-fold increase in linearity, as shown in Fig. 6. In order to understand the reason behind this difference in
GVS between laterally gated and planar devices, the $R_S$ values are measured and normalized as shown in Fig. 7. A circuit schematic of access resistance measurement is also provided in the inset of Fig. 7. In the planar device, $R_S$ shows an early and fast increase, in accordance with previously published work [35]. In case of laterally gated device, as $W_C$ decreases, correspondingly with a decrease in the FR, the steepness of the increase in the $R_S$ is suppressed [28], [29].

This behavior can be explained in the light of the physical model explained in Section I [see (1)–(4)]. The laterally gated structure can be analyzed using JFET equations, since the gating mechanism is similar to JFET devices (Fig. 2). In JFET devices, the intrinsic transconductance increases with the gate voltage as in (1), while it is steady after the threshold in HEMT structures [32], [36].

The source-access resistance causes a potential drop from the gate to source and the gate voltage effectively becomes (2). Due to this effect, the observed (or external) transconductance becomes (3). Considering the dependence of 2DEG mobility ($\mu_n$) to the electric field, the scattering and heating effects, $R_S$ becomes a dependent variable on $I_D$ (or $V_G$) as in (4). Therefore, at high currents (when $V_G \to 1\text{V}$), a drop in transconductance is expected. The intrinsic transconductance of a laterally gated device and the increasing $R_S$ form a balance that leads to a flatter transconductance, especially at more positive gate voltages.

The small-signal measurements of 2 $\mu$m $\times$ 125 $\mu$m devices are taken on-wafer via vector network analyzer (VNA). S-parameters are measured for whole gate voltage range that the device is on, for each device. The frequency where unilateral gain becomes unity ($f_{\text{MAX}}$) of devices with different $W_C$ values is plotted in Fig. 8. Similar to Fig. 6, the planar device shows a bell-shaped characteristic with changing gate bias, and laterally gated devices show constant $f_{\text{MAX}}$ over a large range of gate bias, which shows that $f_{\text{MAX}}$ and transconductance of the device are highly correlated. As $W_C$ increases, $f_{\text{MAX}}$ shows a decreasing trend with gate bias at more positive gate voltages. This is also in line with Fig. 6, and in device with 250 nm $W_C$, the increase in source resistance and increasing trend of transconductance seems to equate at positive gate voltages that results in almost constant $f_{\text{MAX}}$ performance.

The maximum available gain (MAG) at operating frequency is also an important FOM to characterize devices for RF applications. From the same RF measurements, MAG values of the devices with respect to gate voltage at 3.6 GHz are plotted in Fig. 9. Laterally gated devices provide nearly constant MAG while the planar device has a bell-shaped characteristic. Also, with the increase of $W_C$, increasing rate of MAG values with increasing gate voltage is slowed. As it is seen in Figs. 8 and 9, $f_{\text{MAX}}$ and MAG at a certain frequency show quite similar characteristics over the varying gate-to-source voltages.
B. Effect of Metal Width

Different laterally gated devices with various metal widths and a constant channel width are fabricated and measured in order to observe the effect of the FR. Fig. 10 shows the transconductance and Fig. 11 shows the $R_S$ of these devices.

The transconductance of the laterally gated devices shows a negligible threshold shift with the metal width and are flatter than the planar device. As the metal width increases, the transconductance shows a tendency to increase at higher gate voltages. Fig. 11 shows that, as $W_M$ increases, the FR decreases and the normalized $R_S$ decreases. In this case, since $W_C$ is constant, the effect of the channel width on the transconductance can be assumed to be constant to a first order.

As the $R_S$ decreases, the extrinsic transconductance increases. The transconductance varies more (i.e., less flat) when $W_M$ is 100 and 300 nm than $W_M$ is 200 nm, which shows that there is an optimum range to make the transconductance flatter. As in $f_{\text{max}}$ versus $W_C$ characteristics, in Fig. 8, $f_{\text{max}}$ also correlates with transconductance for different $W_M$.

Leakage currents are measured at 40-V drain bias and gate voltage 5 V smaller than the threshold voltage of each device. Nearly constant drain and gate leakage currents are observed from devices with different $W_M$ values. Similar gate leakage currents are measured with planar device, which indicates that the qualities of the Schottky barriers are same in all devices.

First- and second-order derivatives of transconductance with respect to gate voltage ($g_m' = d^2 I_D/dV_{GS}^2$, $g_m'' = d^3 I_D/dV_{GS}^3$) are also used to evaluate device linearity, as shown in [29] and [30]. Smaller $g_m'$ and $g_m''$ indicate a better second-order voltage intercept point, VIP$_2$ (i.e., extrapolated gate voltage amplitudes at which the second harmonic becomes equal to the fundamental tone in the device’s drain current), third-order voltage intercept point, VIP$_3$ (i.e., extrapolated gate voltage for third-order harmonic) and IMD$_3$ (i.e., third-order intermodulation distortion), performances [29].

In Fig. 12(a) and (b), the first and second derivatives of transconductance ($g_m'$ and $g_m''$) with respect to gate bias are plotted.

The peak values are effectively dropped in laterally gated devices and they get closer to zero. In order to quantify this,
we plot the rms value of $g_m'$ for different $W_M$ values in Fig. 13. The smallest value of $g_m'RMS$ is obtained for 150-nm metal width, indicating higher IMD3 [30].

The gate voltage span increases with increasing $W_C$, but the magnitude of the transconductance is decreasing with the same change. This forms a tradeoff dependent on the application wherein the maximum transconductance and the linear voltage span range should be decided regarding the performance requirements.

Two-tone measurements are taken from 2 $\mu$m × 125 $\mu$m devices with Keysight N5244A power network analyzer (PNA-X) to evaluate the devices linearity. Center frequencies from 3 to 4 GHz with 100 MHz step size are scanned to be in line with a 5G frequency band. Two tones with 10 MHz separation are applied to the devices at each center frequency. The output spectrum of each device, the power levels at the fundamental and the harmonic frequencies, is measured to determine the third-order intermodulation distortions (IMD3) as well as input and output intercept points (IIP3 and OIP3). In order to take the changes of total channel width into account, OIP3/$P_{DC}$ ratio is used as a FOM. This correlates normalization of the OIP3 of the device with the dc power at the operating condition. Both planar and laterally gated devices are measured without search for a sweet spot for $g_m''$. Devices are biased with 100 mA/mm, 15 V $V_{DS}$ and terminated with the optimal source and load impedances for the maximum output power. The planar device results in 25.4 dBm OIP3, while laterally gated device with 250 nm $W_C$ results in 31.7 dBm at 3.5 GHz. The OIP3/$P_{DC}$ ratio of laterally gated device provides 6.3-dB improvement over planar device, which shows the enhanced linearity of the proposed design with respect to planar devices. Note that dc bias condition of both devices is the same.

The OIP3 values for 3–4 GHz range and power levels at fundamental and third harmonic frequencies for 3.5 GHz center frequency of laterally gated and planar devices are plotted in Fig. 14(a). As it is seen in Fig. 14(b), carrier to IM3 power ratio of laterally gated devices offers approximately 13 dB improvement, which corresponds to $\sim$6.5-dB OIP3 enhancement as displayed in Fig. 14(a). Additionally, the device maintains improved OIP3 performance over a wide range of operation frequencies.

IV. CONCLUSION

In this work, the working mechanism of laterally gated transistors is studied to understand the effects of channel and metal widths on the transconductance and source-access resistance in these devices. This is the first detailed experimental work in this regard to the best of our knowledge. Devices with different channel widths and FRs are fabricated in order to observe the performance changes in a controlled manner. We observed that the decrease in the FR suppresses the increase in $R_S$. There are at least two opposing mechanisms on $g_m$ that influence the $g_m'ext$ at more positive gate voltages: one is the intrinsic transconductance of the JFET-like lateral gates and the other one is the nonlinear $R_S$ of the GaN HEMTs. We found that, by optimizing the $W_C$ and $W_M$, nearly constant $g_m$ over a wider $V_{GS}$ range and more linear devices can be achieved. A linear gate voltage span of 6.8 V is achieved by properly selecting $W_C$ and $W_M$, which is 2.5 times more than planar devices and $f_{MAX}$ is shown to follow the transconductance. An improvement of 6.9 dB in OIP3/$P_{DC}$ is reached at 3.5 GHz. Laterally gated HEMTs are advantageous to use in GaN-based low noise amplifier applications, especially when the linearity is the main concern.

ACKNOWLEDGMENT

The authors would like to thank Mehmet Deniz Çalışkan, Gurur Salkım, Mahmut Can Soydan, Ahmet Toprak, Hüseyin Çakmak, and Sinan Besler for their valuable support in fabrication as well as Gizem Tendürüs for her contribution in characterization.

REFERENCES


D. S. Lee, T.-T. Liu, J. Ma, G. Kampitsis, P. Xiang, K. Cheng, and E. Matioli, “Multi-

Authorized licensed use limited to: ULAKBIM UASL - Bilkent University. Downloaded on January 28, 2022 at 12:51:57 UTC from IEEE Xplore. Restrictions apply.