Characterisation of CMOS Compatible Vertical MOSFETs with New Architectures through EKV Parameter Extraction and RF Measurement

L. Tan†, M. M. A. Hakim*, S. Connor, A. Bousquet**, W. Redman-White*, P. Ashburn*, S. Hall†**

† Department of Electrical Engineering and Electronics, University of Liverpool, * School of Electronics and Computer Science, University of Southampton, **MHS Electronics UK Ltd, Swindon UK

Abstract

Vertical MOSFETs (VMOSFETs) with channel lengths down to 100nm and reduced overlap parasitic capacitance were fabricated using 0.35µm lithography, with only one extra mask step compared to standard CMOS technology. EKV modelling produced reasonable fitting of the DC and AC characteristics for short channel devices. It is noted that achieving sufficiently long channels in vertical pillar devices is difficult and introduces challenges for accurate and scalable compact modelling. The measured peak f_T was 7.8 GHz and is significantly limited by high contact resistance and affected by un-optimised junction formation. The study comprehensively reveals structure issues that affect the RF performance. The performance inhibitors have then been optimised using process and device simulation. It is demonstrated that f_T and f_{MAX} based on the measurement and numerical simulation, can reach 30.5GHz, and 41GHz respectively.

Introduction

A major advantage of vertical MOSFETs fabricated by implantation and etch techniques is that a low cost process can be used to produce deca-nano channel devices suitable for analogue/RF circuits [1, 2]. However, high gate to drain/source capacitances limit the operating frequency range where useful current and power gain can be achieved. In previous research, a local, self-aligned oxidation process known as fillet localised oxidation (FILOX) was demonstrated to address the problem by thickening the oxide in the overlap region [2, 3]. This process serves to thicken the oxide in the gate overlap region thus reducing the overlap capacitances both at the pillar top and bottom. Recently, a novel frame-gate architecture was proposed to reduce the total gate resistance and protect the pillar sidewall from plasma damage during poly-gate spacer dry etching [4, 5]. These devices exhibited a good sub-threshold slope of 76mV/dec at a channel length of 100nm.

In this paper, further investigations of DC and RF aspects of these surround gate VMOSFET-FILOX devices with a range of process splits are reported. The first part of the work focuses on investigation of modelling issues that emerged during the compact model parameter extraction procedure.

The second part of the work focuses on measurement and simulation of the cut-off frequency f_T and evaluation of f_{MAX} for the VMOSFET-FILOX surround gate devices with both double gate contact (DGC) and frame gate contact (FGC) with FILOX thicknesses of 40 nm and 60 nm. As a result, the influences of contact resistance, FILOX thickness, bias and gate type on f_T and f_{max} are identified with underlying device physics explored.

Experiment Setup

DGC VMOSFET-FILOX devices with 40nm thick FILOX and 30 um channel width were selected for compact modelling. The EKV model was selected due to its good linkage with fundamental device physics and model simplicity which lightens the workload for parameters extraction [6, 7]. The short and long devices had nominal channel length of 100nm and 250nm respectively with body doping of $1x10^{18}cm^{-3}$. A gate oxide thickness of 2.9 nm was obtained from spectro-ellipsometry and CV measurements. The S-parameters were extracted for frequencies up to 10GHz during RF modelling and f_T characterisation. Only the open dummy structure was needed to de-embed the RF pad parasitics which were found to consist only of capacitive components. This is evident in the Smith chart of the 'open' on-wafer dummy structure where S11 and S22 parameters are confined onto a closed circle (equal resistance) in the negative half of the chart as the frequency increases and capacitance impedance reduces. The negligible magnitudes of S12 and S21 also suggest the absence of resistive and inductive components between the gate, drain to source/substrate pads. The extraction technique relates f_T to the product of the magnitude of h_{21} and the test frequency. In this way, f_T values for various gate and drain biases can be extracted conveniently. It was found an excellent agreement with the -20dB/dec method using the h_{21} Bode plot.

EKV Modelling

The threshold voltage at zero body bias, VTO and the body factor GAMMA were extracted from the fits of threshold voltage and sub-threshold slope for the I_D-V_G curve with V_{DS}=50mV and V_{BS} varied from 0V to -2.5V with -0.5V steps. Fitting of the current and transconductance measurements above the threshold

978-1-4244-3705-4/09/$25.00 ©2009 IEEE

voltage, gave the transconductance factor KP and mobility parameters E0, E1, and ETA.

Using the 100nm channel length (short) device, the series resistance related parameter RLX was first extracted by fitting both the on-current of the I_D-V_G curves with V_{DS}=50mV and V_{DS}=1.5V later on self-consistently. The SCE coefficient LETAD and the slope factor dependence for the charge-sharing parameter NCS were extracted from fitting the I_D-V_G curves with V_{DS}=50mV. The DIBL effect coefficient ETAD and its body dependent factor SIGMAD were extracted from the difference of the threshold voltage between the V_{DS}=50mV and V_{DS}=1.5V curves under a range of back biases. The drain to source leakage current characteristic at negative gate bias was left unfitted at this stage due to anomalies arise from non-uniform oxides (gate and FILOX transition in FILOX encroachment region) and the asymmetrical junction structure inherent in VMOSFETs. The final I_D-V_G and g_m fitted curves under high drain bias are shown in Fig 1 (a, b). The velocity saturation parameters UCRIT and Early voltage factor LAMBDA were then optimized from the fitting of the short channel device I_D-V_D and output resistance characteristics with V_{BS}=0V and V_{BS}= -2.5V. The later fitted curves are shown in Fig 1 (c, d).

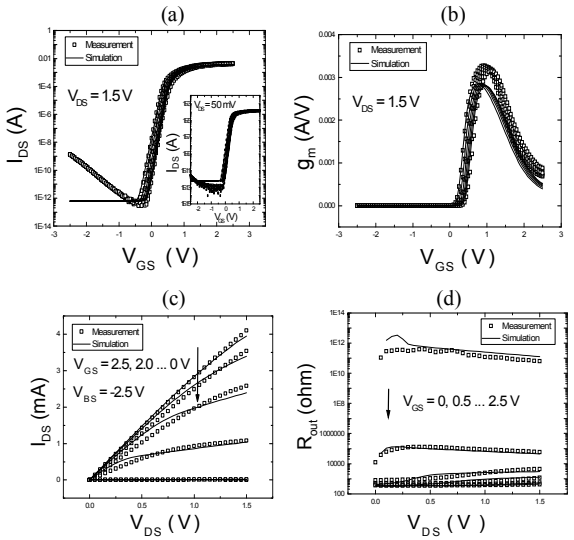

Fig. 1 I_DV_G and g_m , I_DV_D, output resistance curves of the short channel device with V_{DS}=0.5V and 1.5V in (a) (b) and V_{BS}=-2.5V in (c) (d).

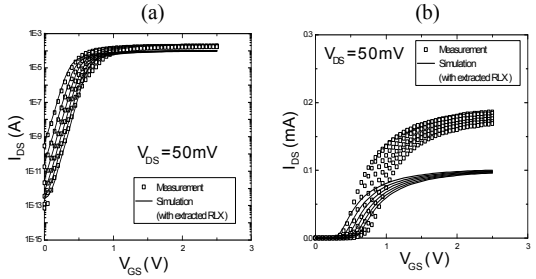

Fig. 2 Comparison of measured and simulated I_DV_G curves of the long device using the RLX extracted from the short device with linear scale in (a) log scale (b).

After RLX extraction from the short channel device, the simulated I_D-V_G curve of the long channel device with V_{DS}=50mV above VTO, now gives an anomalously lower current as shown in Fig 2. This lower current level above threshold suggests the influence of series resistance in the long channel device. This is an issue relates to the fact that the channel length in the 'long' device is not sufficiently long enough. However, increasing the channel length in the vertical pillar process is challenged by the limitation of the boron diffusion range in the vertical direction of the Si body. It is still possible to realise a longer channel by some degree in principle, by using an optimised combination of pillar height and body implantation and diffusion (RTA) condition.

The frequency responses of S-parameters for the short channel device fitted by the simulation are demonstrated in Fig 3 where V_{GS} and V_{DS} are swept from 0-2.5V and 0.05-1.5V respectively. Based on the extractions from CV plots, the junction/overlap capacitance related parameters were improved from S12, S21, S22 and relevant phase/magnitude curves. It was found that the junction and gate overlap capacitance related parameters at the drain side have much more significant impact on the above three S-parameters than at the source. It is also concluded that a smaller drain junction capacitance in VMOSFET can induce lower output impedance in the RF range at high gate biases. The input impedance was largely determined by gate-bulk capacitance and gate sheet resistance which are extracted from S11, S22 curves. However, accurate fitting in some parts were proved difficult to achieve due to small variations between the devices (same type) used for DC and S-parameters measurements.

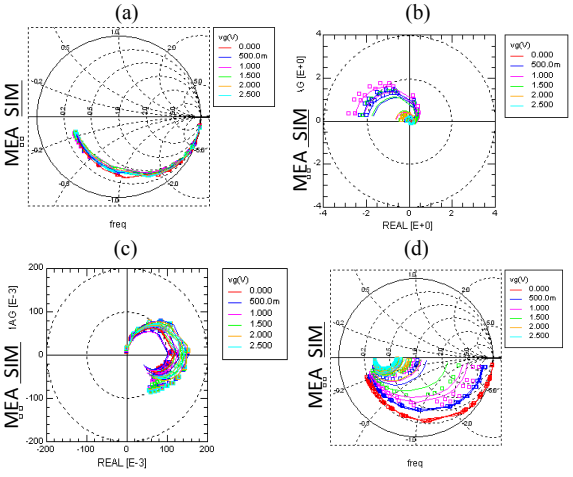

Fig. 3 Fitting of (a) S11, (b) S12, (c) S21, (d) S22 parameter curves of the short channel device with V_{GS}=0, 0.5V, 1.0V, 1.5V, 2.0V, 2.5V and V_{DS}=50mV, 0.5V, 1.0V, 1.5V (orders not shown).

f_T, f_{MAX} Characterisation, Evaluation and Analysis

The f_T values were also extracted from the numerical model which was generated by process

simulation in Silvaco in accordance with the process described in [4, 5]. The f_T values of two VMOSFET structures are shown in Fig 4. It can be observed that for a VMOSFET with FILOX=40nm and 60nm, the initial simulated f_T at V_{GS}=V_{DS}=1.5 V is 24GHz and 30.5GHz respectively. The improvement shows evidence for the contribution of FILOX encroachment on reducing the gate to drain/source capacitance. For the 40nm FILOX device, it can be seen that the measured maximum f_T is about 14.2 GHz lower than the simulation result. This degradation in RF performance is attributed to the fact that the fabricated device has higher specific contact resistance than that in the simulation. The effect of this resistance on the on-current is shown in Fig 5 where a good fit is seen for the on-currents of the simulation and measurement after adding a 5.0K Ω·μm lumped resistor to both drain and source. The presence of this high contact resistance is also in agreement with the result extracted using a test structure which has two separated contact pads. It can therefore be concluded that in the fabricated devices, the contact resistance plays a dominant role in degrading the f_T performance. Lower values of specific contact resistance could be achieved by siliciding the source/drain contacts [7]. The 6.5 times difference in I_{on} between the measurement and simulation at V_{GS}=1.5V, indicates that the intrinsic resistance of the fabricated device has an approximate value of 1590 Ω·μm. This value is similar to that predicted by the analytical model and also values extracted from a numerical model using an RF technique [8]. A lower intrinsic resistance could be obtained using a higher RTA thermal budget to diffuse the junctions beyond the region of FILOX encroachment. This will allow optimisation of the series resistances and gate oxide at the boundary area of both source and drain junctions. Without the

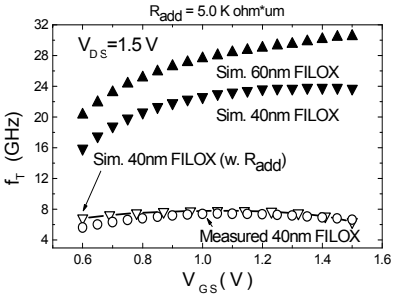

Fig. 4 f_T vs. V_{GS} for a 100nm DGC device with V_{DS}=1.5V.

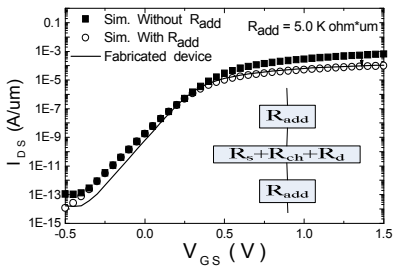

Fig. 5 I_{DS} vs. V_{GS} with V_{DS}=1.5V.

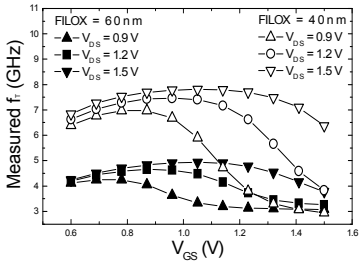

Fig. 6 Measured f_T vs. V_{GS} for 100nm DGC devices.

proposed junction resistance optimisation, with a thicker FILOX such as 60nm the f_T can be even lower than that of 40nm FILOX as shown in Fig 6 despite the reduced overlap capacitance.

Fig 7 shows that in the saturation region, f_T increases with drain bias due to the influence of enhanced saturation current arising from short channel effects (charge sharing, DIBL and channel length modulation). In the linear region, f_T exhibits less roll-off with gate bias as the g_m theoretically gains little enhancement by the gate bias. For each value of drain bias, f_T first increases with gate bias until it reaches its maximum value at a specific gate bias. The initial increase of the f_T is driven by the increasing on-current level due to the increasing gate bias; then it falls with gate bias due to the combined effect of the accelerated increasing of the total gate-to-drain/source parasitic capacitances (largely due to carrier accumulation under the oxide) and the limiting of g_m due to mobility reduction by the gate field. The comparison of the above two dominant terms and the resulting f_T as a function of gate bias are numerically simulated and illustrated in the inset of Fig 7. The peak point of f_T corresponds to the point between the minimum gate-drain/source capacitance and peak of transconductance.

The structure of the FGC devices is shown in Fig 8(a). The thin frame of polysilicon surrounds all parts of the Si-pillar sidewall and a large portion of area on top of the pillar. This serves to protect the channel during etching (a critical process step) and also reduces the gate series resistance compared to the double gate contact structure, as shown in Fig 6(b). From measurements of the resistance of the poly-silicon surround gate made on test structures, it is evident that a FGC device can have a gate resistance

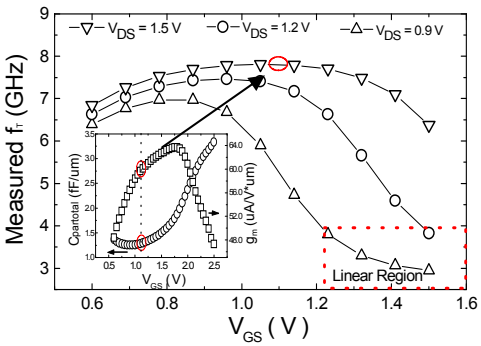

Fig. 7 Measured f_T vs. V_{GS} with V_{DS}=1.5, 1.2, 0.9V.

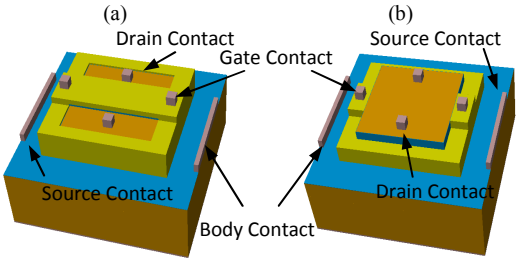

Fig. 8 3D structures of FGC and DGC devices.

as low as 340 Ω·μm while a comparable DGC device has a higher resistance, of the order of 680 Ω·μm when half the depth of the gate spacer is over-etched. The gate resistance varies depending on the degree of etching of the poly-silicon spacer. It is well known that the gate resistance R_g plays an important part in influencing the frequency response f_{max}:

$$f_{max} = \frac{f_T}{2\sqrt{g_{ds}(R_g + R_s) + 2\pi f_T R_g C_{gd}}}$$

Therefore a smaller gate resistance introduced by the FG structure is certainly advantageous for f_{max}. However, the frame gate also causes an increase to the gate to drain/source overlap area which can in turn degrade the peak f_T up to 2.2 times compared to the DG device as shown by the measurements illustrated in Fig 9 (left axis). In Fig 9 (right axis), the fmax of FGC and DGC device are evaluated using f_T source resistance R_s[8], gate-to-drain parasitic capacitance C_{gd} and output admittance g_{ds} values from the optimised/calibrated numerical model and aforementioned measured R_g. Note that the f_T difference of the two structures is taken into account with reference to the difference between measured values. For the 40nm FILOX devices, the FGC has an f_{max} (23.3 GHz at peak) which is significantly lower than the DGC device (35.5 GHz at peak) despite of the gate resistance advantage. The peak point is at the highest gate bias when g_{ds} and gate bias dependent source resistance reach to its minimum. Its corresponded gate bias is higher than that of peak f_T. For a 60nm FILOX the f_{max} reaches the maximum value of 41GHz also at the highest gate bias.

The best f_T (30.5GHz) and f_{max} (47.8GHz) evaluated have shown similar values to the advanced

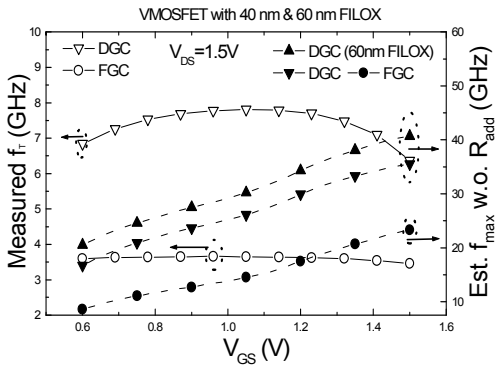

Fig. 9 Measured f_T and estimated f_{max} (no R_{add}) vs. V_{GS}, FG/DGC devices, with 40nm FILOX (if not specified), V_{DS}=1.5mV.

lateral MOSFETs at the 0.35μm technology node according to the ITRS road map. Further device simulation suggests that with the junction abruptness improved from the original 10dec/nm to the ITRS requested 2.7dec/nm [9] for a 50nm channel length, the peak f_T within gate bias of 1.5V can reach as high as 99.4GHz. This corresponds to an almost three-generation-hop in RF performance that can be ultimately achieved with VMOSFETs using a 0.35μm lithography stepper if ultra abrupt junctions are achieved by advanced annealing techniques.

Conclusion

This paper presents detailed compact modelling and RF performance characterisation of 100nm, CMOS compatible vertical MOSFETs. The problems of realising a truly 'long channel device' for model fitting in contrast to the good fitting for the short channel model and device variations for S-parameters fitting highlight challenges in the compact modelling of vertical MOSFETs. The f_T characteristic was compromised by high contact resistance which could be reduced by silicidation. The use of thicker FILOX to reduce the gate overlap parasitic capacitance also brings the risk of a significant degradation in junction series resistance and gate oxide at source end which can be mitigated by controlling the RTA conditions. The best combination of g_m and overlap capacitance determines the bias point of the peak value of f_T. In addition, the frame gate structure serves to reduce the gate resistance but the associated larger poly-silicon spacer area degrades f_T even more than the beneficial effect of the reduced gate resistance. The numerical simulation shows that the best optimised f_T and f_{max} for this technology can reach 30.5GHz and 41GHz. Finally, it is expected that with 0.35μm lithography stepper, up to three-generation-hop in f_T can be ultimately achieved with VMOSFETs if 50nm channel length is targeted using the junction abruptness required by ITRS road map.

Acknowledgement: Funded by the Engineering and Physical Science Research Council, UK.

[1] M. Jurczak, E. Josse, R. Gwoziecki, M. Paoli, T. Skotnicki, Proc. ESSDERC'98, pp. 172–175. 1998.
[2] V.D. Kunz, T.Uchino, C.H. Groot, P. Ashburn, D.C. Donaghy, E. Gili, S. Hall, Y. Wang, P. L. F. Hemment, IEEE TED, Vol. 50, Issue 6, pp. 1480-1487, Jun. 2003.
[3] V.D. Kunz, C.H. de Groot, E. Gili, T. Uchino, S. Hall, and P. Ashburn, ESSDERC'2004, pages 221-224, September 2004.
[4] M.M.A. Hakim, L. Tan, T. Uchino, O. Buiu, W. Redman-White, S. Hall and P. Ashburn, ESSDERC 2008, Aug. 2008.
[5] M.M.A. Hakim, L. Tan, O. Buiu, W. Redman-White, S. Hall and P. Ashburn, Solid State Electronics, Special Issue, 2008.
[6] S. C. Terry, J.M. Rochello, et. al, Nuclear Science Symposium Conference Record, Vol.1, page 317-321, Nov, 2002.
[7] J. kedzierski, et al, IEDM Technical Digest, P57, 2000.
[8] L. Tan, S. Hall, O. Buiu, M.M.A. Hakim, T. Uchino, P. Ashburn, W. Redman-White, ULIS 2008, March 2008.
[9] M.Y. Kwong, K. Reza, P. Griffin, et. al, IEEE TED, Vol. 49, Issue 11, pp. 1882-1890, Nov. 2002.