Sub-30 nm InAs Quantum-Well MOSFETs with Self-aligned Metal Contacts and Sub-1 nm EOT HfO_2 Insulator

J. Lin, D. A. Antoniadis, and J. A. del Alamo

Microsystems Technology Laboratories, Massachusetts Institute of Technology, 60 Vassar St., Rm 39-619, Cambridge, MA, 02139 USA. E-mail: linjq@mit.edu. Phone: 1-617-253-0714. Fax: 1-617-324-5341.

Abstract Sub-30 nm III-V planar Quantum-Well (QW) n-type MOSFETs are fabricated through a self-aligned CMOS compatible front-end process. Good performance and short-channel effect mitigation are obtained through the use of a QW-channel that incorporates a thin pure InAs subchannel and extremely scaled HfO_2 gate dielectric on a very thin InP barrier (total barrier EOT<1 nm). The devices also feature self-aligned metal contacts that are 20-30 nm away from the edge of the gate. At L_g=30 nm, transconductance of 1420 μS/μm and subthreshold swing of 114 mV/dec at 0.5 V are obtained. 22-nm gate-length devices are demonstrated through this process. Long-channel devices exhibit nearly ideal subthreshold swing of 69 mV/dec, and high channel mobility of 4650 $cm^2V^{-1}s^{-1}$ at $N_s=4x10^{12}$ cm^{-2}.

Introduction

InAs and InGaAs are considered promising candidates for channel material in future CMOS applications **[1,2]**. Impressive device prototypes have recently been demonstrated **[3-8]**. However, a III-V MOSFET structure that combines high performance, ability to harmoniously scale down to sub-30 nm gate length dimensions and CMOS-type manufacturability is still to be realized. In this work, we demonstrate a novel QW-MOSFET that addresses these challenges. For this, we use an extremely-scaled HfO_2 gate insulator. The fabrication closely follows CMOS requirements, particularly self-alignment of the refractory metal gate and metal contacts, very low thermal budget, gate-last process that uses RIE extensively and an entirely lift-off free process in the frontend. MOSFETs with gate length dimensions in the 20-30 nm range and outstanding electrical characteristics are obtained.

Device Fabrication

The process flow is shown in **Fig. 1**, a device cross-section schematic is illustrated in **Fig. 2** and TEM pictures of finished devices are shown in **Fig. 3**. The starting heterostructure is grown by MBE and it features a 2 nm InAs sub-channel clad by 3 and 5 nm of $In_{0.7}Ga_{0.3}As$ and a novel n^+-InP/n^+-InAlAs/i-InP ledge design to reduce access resistance.

Low-ρ Mo is first sputtered as metal contact (R_{sh}=5 Ω/□) followed by CVD SiO_2 deposition. After mesa isolation, the gate pattern is defined by E-beam lithography. The SiO_2 and Mo layers are then patterned by RIE. A thermal annealing step at 350°C in N_2 is performed to remove RIE damage **[9]**. The top n^+ InGaAs cap is wet-etched in a well-controlled manner that yields an undercut < 20 nm, as shown in **Fig. 3 (a)**. The top n^+ InP spacer is etched through highly-directional Ar etch, stopping at n^+ InAlAs. Subsequently, the Mo at the S/D regions is pulled back by isotropic RIE **[10]**. We use a self-limiting digital etch technique (plasma oxidation + diluted H_2SO_4) to remove the n^+-InAlAs and i-InP in a well-controlled manner. One cycle of this etch removes about 0.9 nm of InAlAs or InP. Through careful calibration, we can obtain about 1 nm InP barrier remaining above the channel.

Immediately after the last digital etch cycle, the sample is loaded into ALD for gate dielectric deposition. 2 nm HfO_2 is used as gate dielectric. From the dielectric constant extracted by a separate calibration experiment, the EOT for the HfO_2 alone is ~ 0.4 to 0.5 nm. The total EOT including the InP barrier is ~0.8 nm. Evaporated Mo is used as gate metal and patterned by RIE. The device is finished by pad formation. This is the only lift-off step in the backend of the process.

The final gate length is defined by the recess opening in the SiO_2. This can reach below 20 nm. The ALD gate dielectric is conformal and passivates the access region between the edges of the gate and the n^+ InGaAs cap. This distance is L_{side}~ 20-30 nm. HR-TEM shows the high quality device interface in **Fig. 3 (c)**. The highest temperature step in the entire process is the 350°C damage annealing. There is no high temperature step after the ALD film is deposited.

Results and discussion

The output characteristic of a device with L_g= 30 nm are shown in **Fig. 4**. Its R_{on} is 475 Ω.μm. R_{sd}~455 Ω.μm is obtained by L_g-extrapolation method. In our self-alignment design, ohmic metal Mo with R_{sh}=5 Ω/□ is placed in very close proximity to the gate edge. This represents an order of magnitude reduction in the access region R_{sh} with respect to n^+-InP or n^+-InGaAs, as used in other schemes where the metal contacts are still far apart **[4,5]**. Previous work also shows that Mo offers a very low contact resistance to n^+ InGaAs **[11]**. Our relatively high R_{sd} mainly arises from the un-capped L_{side} regions. This can be solved through an improved n^+-InP S/D ledge design.

The subthreshold characteristics of this device are shown in **Fig. 5**. Steep subthreshold swings (103 mV/dec at 50 mV, 114 mV/dec at 0.5 V) over a wide range of V_{gs} values are obtained. DIBL is 236 mV/V. We believe this relatively high value is associated with the heterostructure buffer and is not related to the fabrication process. Despite using a pure HfO_2 gate dielectric of 2 nm, the maximum I_g over the voltage range of **Fig. 5** is below $1x10^{-9}$ A/μm. The transfer and g_m characteristics of the same L_g= 30 nm device are shown in **Fig. 6**. A peak g_m of 1420 μS/μm is obtained at V_{ds}=0.5 V. Very small hysteresis (< 10 mV) in I_d is shown. A maximum $g_{m,pk}$=1530 μS/μm is obtained at L_g=60 nm.

A fully operational L_g= 22 nm MOSFET has been obtained (XTEM in Fig. 3b). The output and transfer characteristics are shown in **Figs. 7** and **8**. At V_{ds} = 0.5 V, the device delivers a transconductance of 1050 μS/μm. **Fig. 9** shows device subthreshold characteristics from L_g=150 nm to 22 nm at 0.5 V.

Fig. 10 shows S at 0.5 V *vs.* physical gate length L_g for our transistors, recent III-V MOSFETs, and InGaAs HEMTs **[2-8]** down to 30 nm. Our ultra-scaled barrier MOSFETs exhibit a subthreshold swing that is superior to any other planar III-V MOSFET and that matches the best Tri-gate III-V devices **[3]**. In addition, our devices feature outstanding transport properties. As shown in **Fig. 11** among transistors with $L_g \leq 60$ nm, for a given S, the transconductance is superior to any other planar or Tri-gate III-V MOSFET. **Fig. 12** shows I_{on} for an I_{off}=100 nA/μm at V_{dd}=0.5 V. This FOM combines short-channel effects and drive current. For L_g>70 nm, our devices are among the best reported to date. At L_g=60 nm, our device is superior to a Tri-gate III-V MOSFET **[3]**.

A further manifestation of the scalability of this technology is shown in **Fig. 13** that graphs saturation V_t roll-off. Good scalability is demonstrated down to 30 nm with 2 nm HfO_2 gate dielectric. For comparison, a device run made with a 2 nm thick Al_2O_3 gate dielectric shows relatively poor V_t roll-off as a result of poor electrostatic control.

S at 0.5 V *vs.* I_d for L_g= 150 nm devices is plotted for three different gate dielectrics in **Fig. 14**. For the sample with pure Al_2O_3, S is high as a consequence of its higher EOT. With pure HfO_2, S stays almost constant across 3 orders of magnitude in I_d (10^{-8} to 10^{-5} A/μm). With Al_2O_3 at the interface and HfO_2 on top, the minimum value of S is comparable to that with pure HfO_2, but as I_d decreases it rises at higher values of I_d just as with pure Al_2O_3. In both samples with Al_2O_3 at the InP interface, the rise of S happens while $I_g \ll I_d$. This suggests that the shape of S is due to the different distribution of D_{it} across the InP bandgap. With HfO_2, a lower and more uniform D_{it} distribution is present close to midgap. Closer to the conduction band edge, Al_2O_3 has a slightly lower D_{it}. These results are consistent with recent findings **[12]**.

As part of our process development, we fabricated devices on a different heterostructure with a 15 nm thick $In_{0.53}Ga_{0.47}As$ channel and an Al_2O_3/HfO_2 (0.4/2 nm) composite gate dielectric. The InP barrier is scaled down by a time-controlled low-energy Ar dry etch **[10]**. The final InP thickness is 1 nm. Long-channel MOSFETs (L_g=300 μm) with nearly ideal S= 69 mV/dec at V_{ds}=50 mV were obtained (**Fig. 15**). This result is among the best ever reported for III-V MOSFETs. This is seen in **Fig. 16** that graphs S *vs.* dielectric EOT for several long-channel III-V FETs at low V_{dd}. This graph illustrates the importance of D_{it} and dielectric scaling. Our gate-last process, in spite of its multiple RIE steps delivers good interfacial quality with a very thin dielectric.

Fig. 17 shows split C-V measurements of long-channel MOSFETs with HfO_2 and Al_2O_3 gate dielectrics. A noticeable D_{it} bump is observed in the case of Al_2O_3 in weak inversion at low frequencies, indicating a significant amount of slow traps close to mid-gap. This is absent when using a pure HfO_2 dielectric. **Fig. 18** shows electron mobilities extracted from 20-μm long MOSFETs. Mobility of 4650 $cm^2V^{-1}s^{-1}$ at N_s of $4x10^{12}$ cm^{-2} is obtained. This is the one of the highest mobility values at this N_s published in InGaAs MOSFETs to date **[4, 13]**.

Gate current densities (J_g) for devices with different dielectrics on the two heterostructures are shown in **Fig. 19**. Due to a larger CB discontinuity, Al_2O_3 results in reduced J_g for the same physical thickness. A 2 nm HfO_2 gate dielectric delivers J_g <1 A/cm^2 across the entire 0.5-V forward V_g span.

Conclusion

In this work, we have demonstrated the shortest functional III-V MOSFETs to date. Our InAs QW-MOSFETs feature an ultra-scaled pure HfO_2 gate dielectric and are fabricated through a self-aligned CMOS compatible process with L_{side}=20-30 nm. Fully operational devices as short as L_g=22 nm have been achieved. Outstanding performance and short-channel effects in devices down to L_g=30 nm has been demonstrated. A long-channel subthreshold swing of 69 mV/dec and an electron channel mobility of 4650 $cm^2V^{-1}s^{-1}$ at high N_s=$4x10^{12}$ cm^{-2} have also been demonstrated.

Acknowledgements: This work is sponsored by FCRP-MSD Center and Intel Corp. Device fabrication was carried out at the Microsystems Technology Laboratories of MIT. MBE growth is provided by Intelliepi.

References:
[1] J. A. del Alamo, Nature, 479 (2011), p.317.
[2] D.-H. Kim *et. al*., IEDM 2008, p.719.
[3] M. Radosavljevic *et. al.,* IEDM 2011, p.765.
[4] M. Egard *et. al.,* IEDM 2011, p. 304.
[5] Y. Yonai *et. al.,* IEDM 2011, p.307.
[6] R. Terao *et. al.,* Applied Phys. Exp.5, 054201, (2011).
[7] S. H. Kim *et. al.,* VLSI 2012, p. 177.
[8] T. W. Kim *et. al.,* VLSI 2012, p.179.
[9] J. Lin *et. al.,* Applied Phys. Exp.5, 064002, (2012).
[10] N. Waldron *et. al*., T-ED, 56 (2010) p.297.
[11] T. W. Kim *et. al.,* IEDM 2010, p.696.
[12] R. Galatage *et al*., Techcon 2012.
[13] F. Xue *et. al.,* Applied Phys. Lett. 98, 082106 (2011).

- Sputtered Mo contact
- CVD SiO_2 hard mask
- Mesa isolation
- Gate lithography
- Gate recess: SiO_2/Mo/Cap etch
- Digital etching of InAlAs/InP barrier
- H_2SO_4 cleaning
- ALD gate dielectric deposition
- Mo gate evaporation
- Gate head photo and pattern
- Pad formation

Fig. 1 Process flow of buried-channel self-aligned gate-last QW-MOSFET.

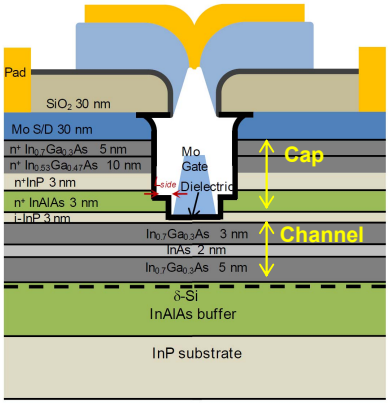

Fig. 2 Cross sectional schematic of InAs QW-MOSFET with ultra-scaled HfO_2/InP composite barrier.

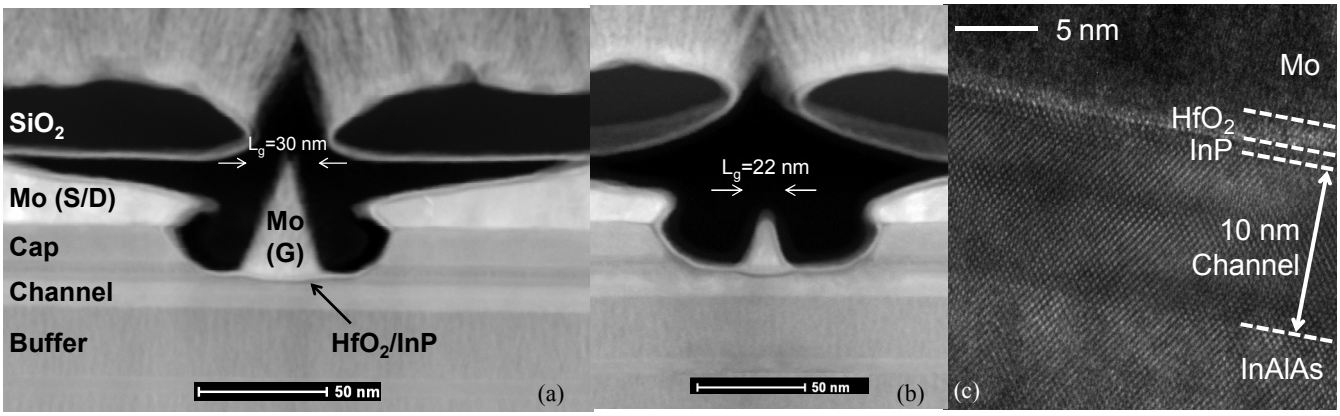

Fig. 3 XTEM of QW-MOSFET with (a) L_g= 30 nm, and (b) L_g= 22 nm. (c) HR image near the intrinsic gate region for L_g= 30 nm MOSFET.

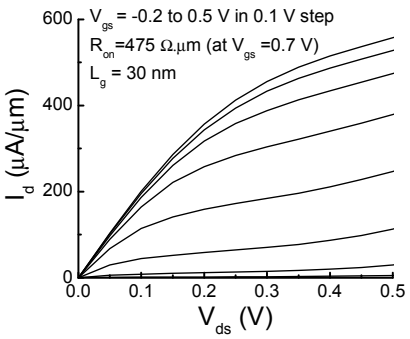

Fig. 4 Output characteristics of QW-MOSFET with L_g= 30 nm.

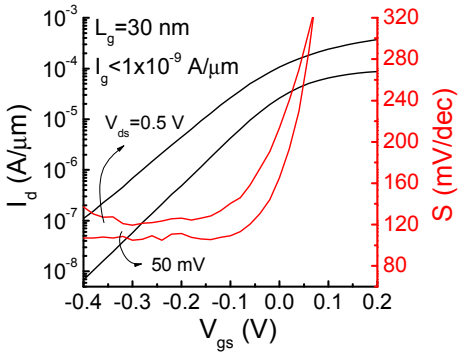

Fig. 5 Subthreshold characteristics and subthreshold swing of L_g = 30 nm QW-MOSFET.

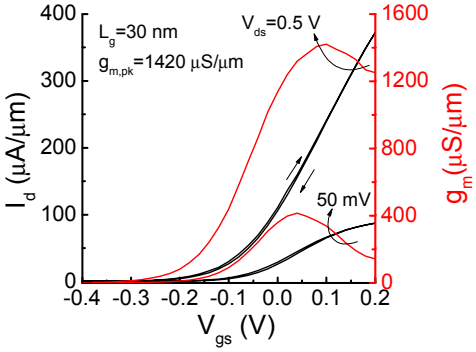

Fig. 6 Transfer and g_m characteristics of L_g= 30 nm QW-MOSFET.

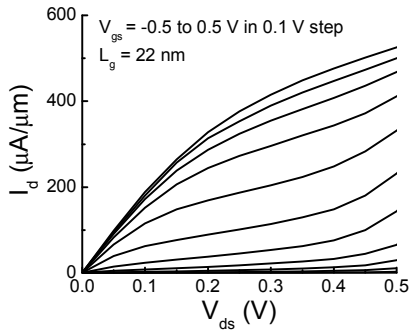

Fig. 7 Output characteristics of QW-MOSFET with L_g= 22 nm.

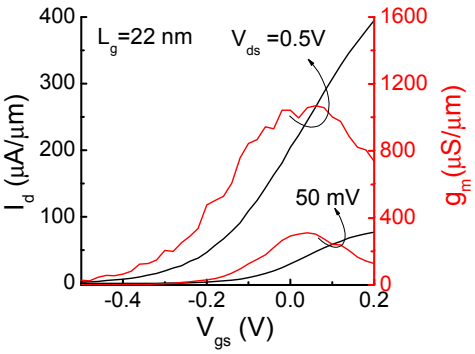

Fig. 8 Transfer and g_m characteristics of L_g= 22 nm QW-MOSFET.

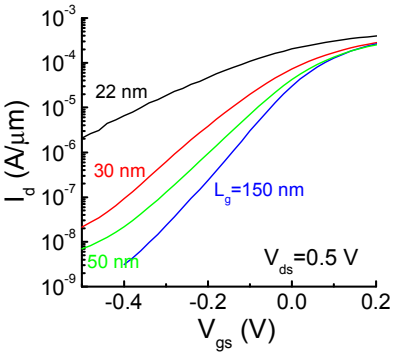

Fig. 9 Subthreshold characteristics of QW-MOSFETs from 150 nm down to 22 nm.

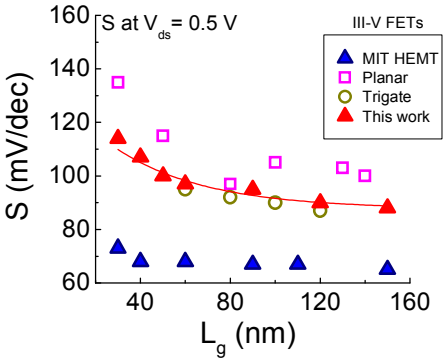

Fig. 10 S at 0.5 V for III-V MOSFETs and HEMTs.

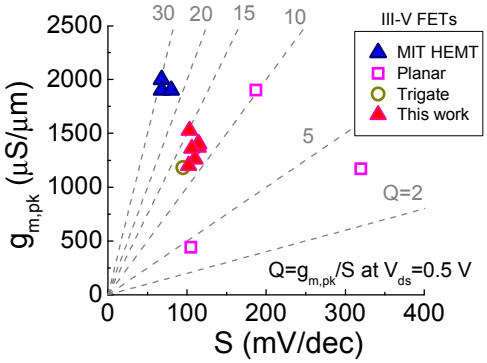

Fig. 11 Peak transconductance *vs.* S for III-V MOSFETs and HEMTs with $L_g \leq 60$ nm. $Q=g_{m,pk}/S$ [Unit: $\mu S.\mu m^{-1}.dec.mV^{-1}$].

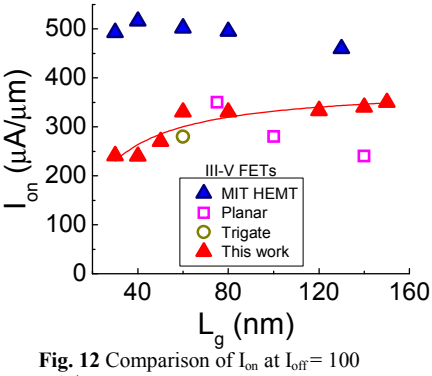

Fig. 12 Comparison of I_{on} at I_{off} = 100 nA/μm and V_{dd} = 0.5 V for III-V MOSFETs and HEMTs.

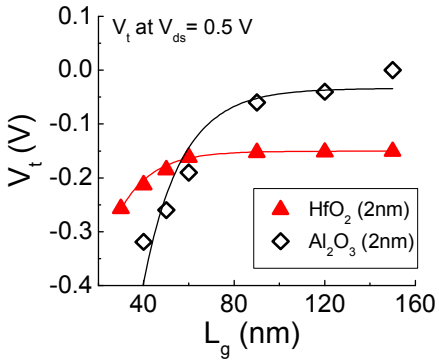

Fig. 13 $V_{t\text{-}sat}$ *vs.* L_g. Very small V_t roll-off is observed.

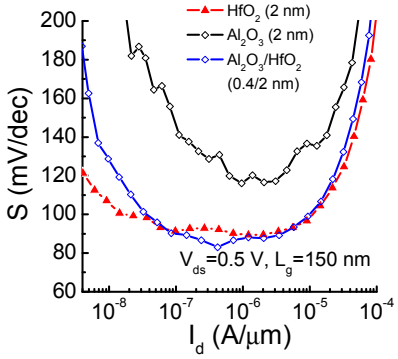

Fig. 14 S *vs.* I_d for three representative devices with different gate dielectrics (in all cases $I_g<10^{-9}$ A/μm).

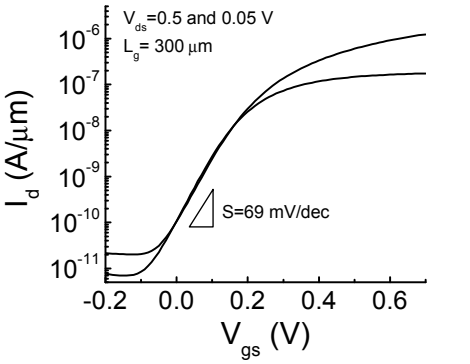

Fig. 15 Subthreshold characteristics of a long-channel $In_{0.53}Ga_{0.47}As$ FET with Al_2O_3 (0.4 nm)/HfO_2 (2 nm) gate dielectric. The InP barrier is 1 nm thick.

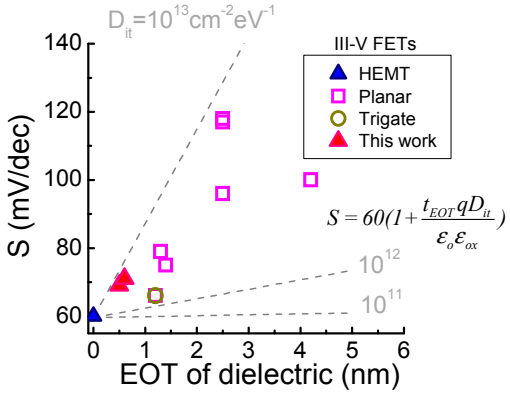

Fig. 16 Low-voltage subthreshold swing *vs.* dielectric EOT of several long-channel III-V MOSFETs and HEMTs.

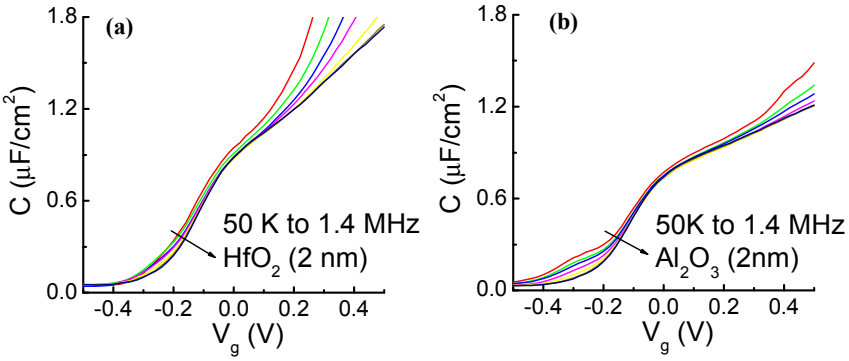

Fig. 17 Capacitance *vs.* V_g characteristics for devices with 2 nm of (a) HfO_2 (b) Al_2O_3 by split C-V measurement. I_g starts affecting the measurement at $V_g \sim 0.3$ V, depending on dielectric used and frequency.

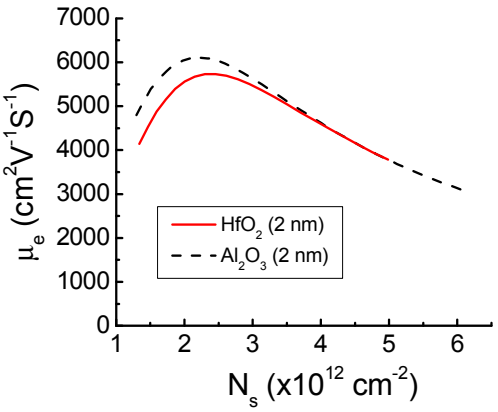

Fig.18 Channel mobility extracted by split C-V method without D_{it} correction.

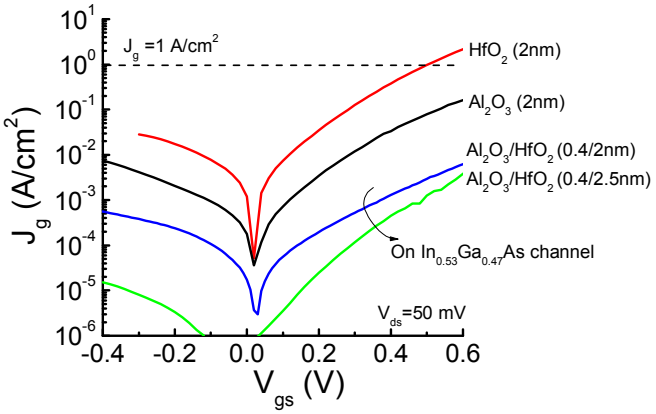

Fig.19 Gate leakage of QW-MOSFETs with different gate dielectrics.