RPN Oxynitride Gate Dielectrics for 90 nm Low Power CMOS Applications

A. Veloso[1], M. Jurczak[1], F. N. Cubaynes[2], R. Rooyackers[1], S. Mertens[1], A. Rothschild[1],
M. Schaekers[1], H. N. Al-Shareef[3], R. W. Murto[3], C. J. J. Dachs[2], G. Badenes[1]
[1]IMEC, Kapeldreef 75, 3001 Leuven, Belgium
[2]Philips Research Leuven, Leuven, Belgium; [3] International Sematech, Austin-TX, USA

Abstract

This paper investigates the use of RPN-based oxynitride gate dielectrics for 90 nm Low Power (LP) CMOS applications. Several recipes have been developed to optimise the gate dielectric for targeted EOT, high mobility and improved EOT uniformity. Compared to conventional furnace oxynitride, significant gate leakage reduction has been found in devices with plasma nitrided oxides. This enabled reaching the spec for the I_{OFF} current of 20 pA/μm and improve the I_{ON}-I_{OFF} trade-off. The I_{ON} current obtained at 1.2 V for NMOS and PMOS devices is 427 μA/μm (at I_{OFF} =16 pA/μm) and 170 μA/μm (at I_{OFF} =16 pA/μm), respectively. The obtained results are among the best values reported in the literature.

1. Introduction

One of the major constrains in the scaling of devices dedicated to low power applications is the leakage current. According to the ITRS roadmap [1], the off-state spec for 90 nm CMOS technology targets 300 pA/μm and 1 pA/μm for low operating power (LOP) and low standby power (LSTP) applications, respectively. These transistor leakage specifications put a very stringent constraint on the gate dielectric leakage, which needs to be kept at an even lower level. As a consequence, gate dielectric thickness scaling is limited by leakage constrains, and a careful optimisation of dielectric properties is necessary so that the minimum possible leakage level for a given thickness is achieved. These targets are expected to be met with electrical equivalent oxide thickness (EOT) below 2.2 nm. Since high-k dielectrics having potential to ameliorate the gate leakage problem are still in the development phase, there is strong motivation to extent the use of oxynitride until its ultimate limits. Recently, new techniques of radical plasma nitridation, that significantly reduce the gate leakage as compared to conventional oxynitride [2], have been presented [3,4]. Oxynitrides fabricated with these techniques are considered to be the best candidates for 90 nm CMOS integration or below, at least until high-k dielectrics reach their maturity for gate applications. In this paper, we investigate the suitability of the oxynitrides formed by plasma nitridation for 90 nm low power CMOS devices. Our low power devices target the internal spec for I_{OFF} below 20 pA/μm, which is much more aggressive than that defined in the ITRS roadmap for LOP devices. In section 2, two different oxidation processes followed by plasma nitridation are investigated and compared with conventional furnace oxidation in NO. Section 3 demonstrates the performance of devices with best recipes for the gate oxides.

2. Gate dielectrics optimisation

2.1 Experiment

For the initial gate dielectric study series, NMOS long channel transistors (with the furnace and plasma nitrided gate oxides) were fabricated on 200 mm diameter silicon wafers. The reference wafer received 2 nm furnace post-annealed NO oxide [5]. The rest of the wafers were used to investigate the best recipe for the plasma nitrided (RPN) oxide. RPN gate dielectric growth was done using two alternative processes, each consisting of two steps:

1 - Growth of a pure thin oxide film using Rapid Thermal Oxidation (RTO, at 950-1000 ^oC with a total flux of 5 slm O_2) or In-Situ Steam Generation oxidation (ISSG, at 850-950 ^oC, with a total flux of 5-8 slm O_2+H_2+N_2).

2 - Nitridation (20-80%) using a remote plasma (RPN) at 550 ^oC and ~ 3 Torr. The nitridation time varied from 180 to 240s.

The plasma process used N_2 gas with helium to control the amount of atomic nitrogen present in the reaction chamber. A higher percentage of helium leads to a reduction of the probability of recombination of nitrogen radicals and hence to the incorporation of more nitrogen into the gate dielectric.

After gate dielectric growth, 150 nm thick polysilicon were deposited. After gate patterning, As extensions were implanted with 5 keV energy and 80 nm spacers were formed. For HDD, As and P were implanted at 40 keV and 10 keV, respectively. Phosphorous co-implantation and subsequent RTA anneal at 1050 °C for 10s provided high dopant activation in the gate [6].

Standard Titanium capped Cobalt silicidation completed the front end processing.

Electrical measurements were made by probing the silicide on poly/active and backside Aluminum.

2.2 Electrical results

EOT and flat-band voltage were extracted from the CV curves measured at 100×100 μm^2 square capacitors using a simulation and analysis package [7], which is based upon a model containing first order physics approximations. The model takes into account quantum mechanical effects and poly depletion. As shown in fig. 1, perfect fit between model and experimental data has been obtained.

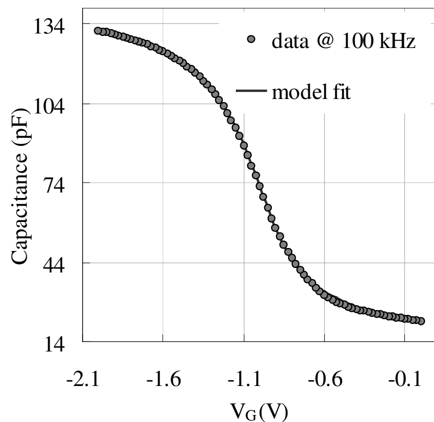

Figure 1. High frequency capacitance-voltage measurement of RTO+RPN gate dielectric. Model fit is used to estimate the EOT.

The trade-off between the gate leakage current density (J_G) and the EOT measured for reference furnace nitrided oxide and various RPN type oxynitrides is shown in fig. 2. Stronger oxide nitridation, possible with plasma nitridation, enables 3-fold reduction in the gate leakage as compared to the reference furnace oxide. Figure 2 shows that only RPN recipes targeting EOT higher than 2 nm meet the requirements for the gate leakage allowable for low power devices. Within this range, no significant difference between ISSG and RTO based oxides has been observed. It was, however, found that nitridation conditions have strong influence on the gate oxide properties. For EOT ≥ 2.0 nm, stronger nitridation, i.e. higher He concentration (65-80%), leads to an increase in EOT, whereas for EOT < 2.0 nm, an opposite trend is present.

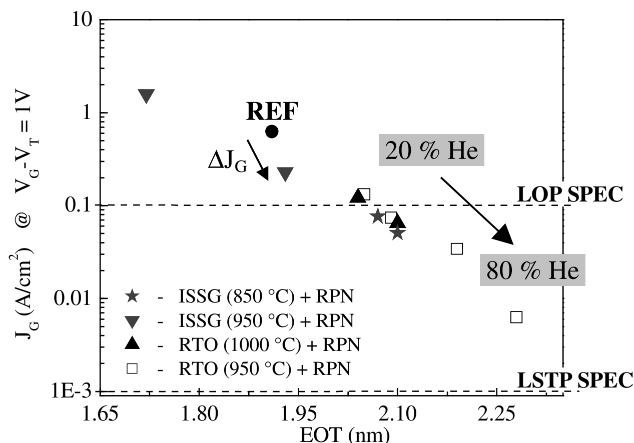

Figure 2. Gate leakage current density (J_G) vs. EOT for different gate dielectrics.

Figure 3 illustrates electron mobility extracted from the maximum transconductance (Gm) measured for various oxynitride recipes, as a function of EOT. For the same EOT, mobility measured in plasma nitrided oxide is slightly higher. It increases with EOT regardless of the content of incorporated nitrogen. For 2.3 nm ≥EOT ≥ 2.0 nm, the maximum mobility value ($\mu_{eff,max}$) is essentially constant, around 230 cm^2/V.cm, independent of the EOT value and the oxidation process (RTO vs. ISSG) used. It is also ~ 10% higher than the maximum value obtained for the reference furnace oxide.

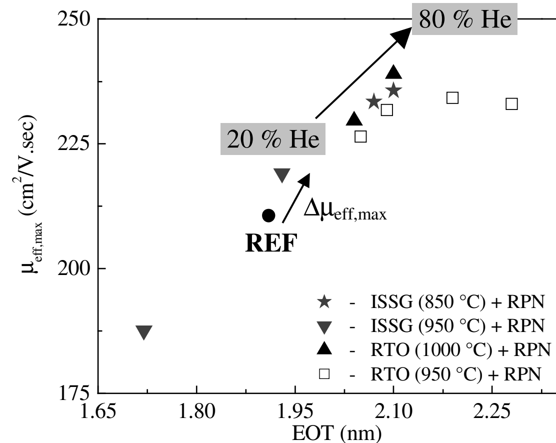

Figure 3. EOT dependence of the electron mobility (μ_{eff}=Gm/(V_{DS}×C_{ox}×W/L) for different gate oxide recipes.

The nitrogen content in the oxynitrides has strong impact on the transistor threshold voltage. Figure 4 illustrates the threshold voltage of long channel NMOS transistors with different gate dielectrics as a function of the extracted EOT. When compared to the reference furnace NO oxide, V_T for 20% He-RPN oxynitride is approximately 40 mV lower. This indicates that RPN oxynitrides exhibit more positive charges at the interface than the reference furnace NO oxide. For EOT ≥ 2.0 nm, a slight increase of the long channel threshold voltage is observed when increasing the He concentration (20% → 65-80%) in the RPN step. Figure 4 also shows that for a given EOT value (2.1 nm), an RTO based oxide leads to a slightly higher (~ 2.2 %) long channel V_T value than a ISSG based oxide.

Figure 5 shows the EOT uniformity for different gate oxide recipes across the wafer. Overall, compared to the reference furnace NO oxide with its 0.7% EOT standard deviation, RPN-based gate dielectrics exhibit much higher EOT non-uniformity. For all RPN splits, the

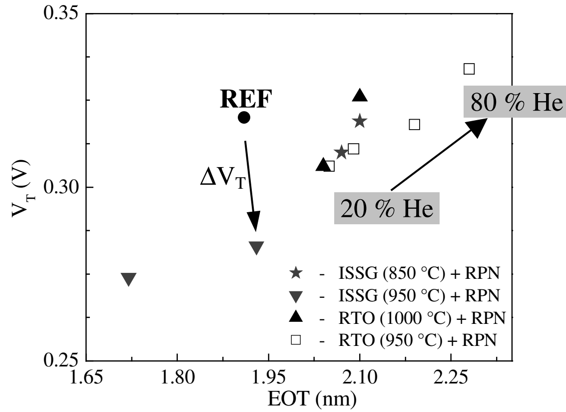

Figure. 4. EOT dependence of the long channel threshold voltage (V_T) for different gate oxide recipes.

highest value of EOT has been measured in the centre of the wafer. EOT reduces along the radial of the wafer. In the case of strong nitridation (66% He), EOT changes by 1-2 Å across the wafer. On the contrary, uniformity of RPN oxides with 20% He is much better. In this case, EOT changes by 2% within-wafer. This EOT non-uniformity is attributed to the radial distribution of nitrogen across the wafer. RTO based oxides experience more degraded film uniformity upon a more aggressive RPN step. A possible explanation for this could be related with oxide differences at intrinsic level, since the main difference between the two oxides, ISSG and RTO, is that the latter has no H_2 present during oxidation. In fact, it has been previously reported [8] that ISSG may lead to a higher quality oxide due to the hydrogen incorporated in the film during oxidation, which is believed to passivate some of the dangling bonds and provide a possible smoother interface between the silicon/silicon dioxide interface.

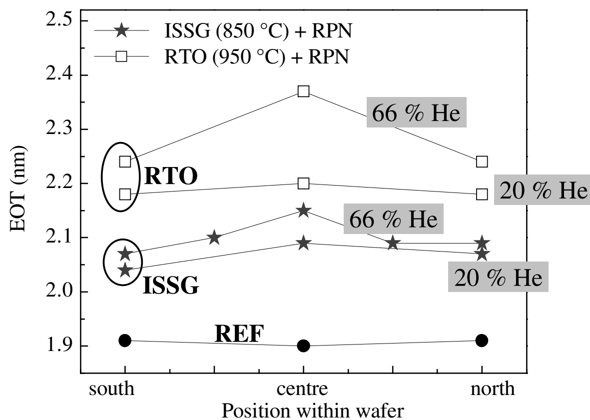

Figure 5. Comparison of the gate oxide uniformity for different oxide process recipes.

3. Transistor performance

Further evaluation, focused on the RPN-based oxide recipes targeting EOT of 2 nm and providing the best trade-off between off-state and drive currents, was done. NMOS and PMOS transistors were fabricated using 90 nm CMOS process. The process features STI, 150 nm polysilicon, extensions implanted with As (5 keV) and B (1 keV) for NMOS and PMOS, respectively, BF2 (NMOS) and As (PMOS) halos, 80 nm nitride spacers. HDD junctions were formed with 3 keV B for PMOS and 40 keV arsenic combined with 10 keV phosphorus implant for NMOS. After sharp spike anneal at 1100 °C, Co silicide was formed with a Ti capping layer.

Figure 6 shows the gate leakage in NMOS and PMOS devices in the off-state measured as a function of the gate length. An almost constant magnitude of the gate current (independent of the gate length) indicates that the gate to junction current path dominated the gate leakage in the transistor off-state [9]. With the RPN-based oxide, the gate leakage is one order of magnitude smaller when compared to the reference furnace NO oxide, for the same EOT. These results are in agreement with the capacitor results presented in fig. 2. In PMOS, gate leakage is lower due to weaker tunnelling of holes and possibly different junction to gate overlap. The small difference between RTO and ISSG type of oxides results from a small variation in EOT between them.

The intrinsic performance of NMOS and PMOS devices with 2 nm RPN-based oxides measured at the bias of 1 V and 1.2 V is plotted in fig. 7. 2 nm RPN oxide enables reducing the gate to junction current, so that the LOP spec for off-state current can be reached for both NMOS and PMOS devices. No significant differences are observed between the performance of RTO and ISSG based oxide devices. One must

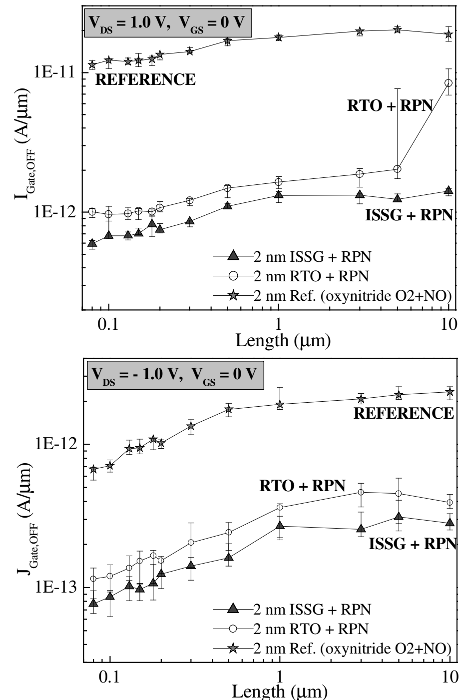

Figure 6. Gate leakage current density versus as drawn gate length for NMOS (top) and PMOS (bottom) devices.

nevertheless keep in mind that they may appear at thinner EOTs, where differences between the two oxides will become more critical in what concerns the device performance. For $|V_{DD}|$=1.2 V, I_{ON} current is 427 µA/µm (170 µA/µm) at I_{OFF} equals to 16 pA/µm (16 pA/µm) for NMOS and PMOS devices, respectively. The I_{OFF} current is determined by source to drain leakage and not by the gate leakage. The obtained results are among the best values reported in the literature [e.g., 10,11].

Figure 8 presents the saturated threshold voltage (V_T) of NMOS and PMOS transistors as a function of the gate length (Lg) as drawn. The offset between drawn and physical gate length is around 20 nm. Regarding the control of the short channel effects, there is no significant difference between the two RPN oxynitrides. For both oxides, the short channel effects in NMOS and PMOS devices are well controlled down to 80 nm. RTO and ISSG based oxides present similar immunity to boron penetration.

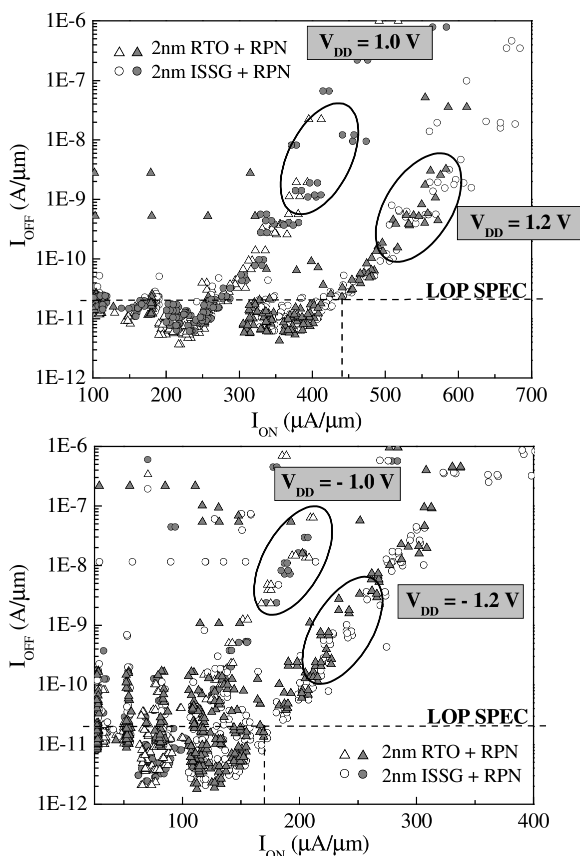

Figure 7. ITP curves for NMOS (top) and PMOS (bottom) transistors.

4. Conclusions

The suitability of the RPN-based oxynitride for low power applications was investigated. RPN process requires careful optimisation to reduce the gate leakage current without increasing EOT and gate oxide thickness non-uniformity.

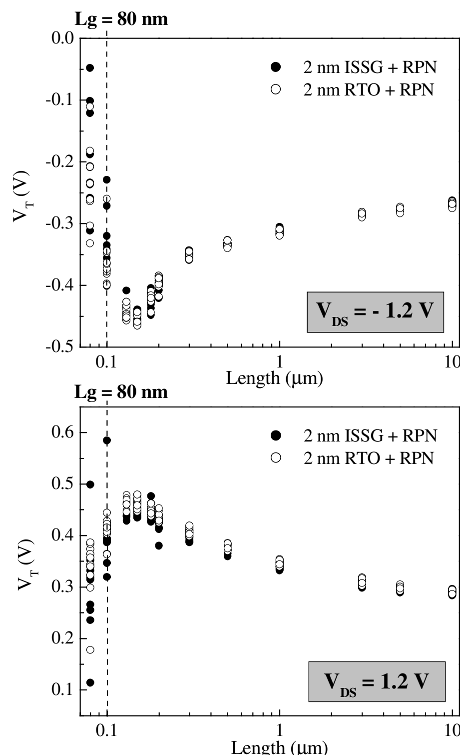

Figure 8. Saturated threshold voltage versus as drawn gate length for PMOS (top) and NMOS (bottom) transistors. A drawn gate length of 100 nm corresponds to a physical length of 80 nm.

ISSG and RTO based oxides appear to be equivalent in terms of gate leakage, mobility and intrinsic transistor performance. The differences might however consist in oxide reliability, device lifetime and oxide scalability, not being a subject of this work.

RPN oxynitrides appear to be suitable for 90 nm LOP CMOS device applications, where gate leakage is of major concern. With two 2 nm RPN-based oxides, very good NMOS and PMOS transistor performance was obtained.

5. References

[1] ITRS Road Map, 2001 Edition.
[2] M.Fujiwara *et al.*, VLSI'1999
[3] M.Togo *et al.*, VLSI'2000, pp.116-117 (2000).
[4] Y.Tasuda, *et al.*, VLSI'2001, pp.83-84 (2001).
[5] S. Kubicek *et al.*, ESSDERC, pp. 264-267 (2000).
[6] E.Augendre *et al.*, ESSDERC'2001, pp.115-118 (2001).
[7] J. R. Hauser *et al.*, Characterization and Metrology for VLSI Tech., pp. 235-239 (1998).
[8] Al-Shareef *et al.*, Appl. Phys. Lett., 78(11),pp.3875 (2001).
[9] W. K. Henson *et al.*, IEEE T-ED, 47(7), pp.1393 (2000).
[10] T. Schiml *et al.*, VLSI, pp. 101-102 (2001).
[11] S. Parihar *et al.*, IEDM, pp. 249-252 (2001).

6. Acknowledgements

This work was partially funded by the European Union under the IST-1999-11599 HUNT project.