Mechanism to Achieve PMOS and NMOS Band Edge Work Function using Low Temperature Tuning Process for Low Power Application


Abstract: Band edge work function metal gates of ~4.1 eV and ~5.0 eV for N- and P-MOSFETs, respectively, were achieved at low EOT (~0.9 nm) and low leakage (~<0.1 A/cm² at V₉=g=V₉h=0.9 V). These data result from mechanistic understanding of the low temperature process flow using WF tuning techniques with simple TiN metal and various tuning species. P-type WF was obtained by suppressing V₉h roll-off and N-type WF was achieved by controlling various tuning species in simple TiN.

Introduction

The need for high-k/metal gate stacks in CMOS [1,2] has led to intensive study on electrode work function tuning processes and the associated mechanisms. In the past, high temperature approaches using metal doping into high-k (La, Sc, Y for NMOS [3] and Al, Ti for PMOS [4,5]) and dual channel (Si for NMOS and SiGe for PMOS [6]) were explored. Achieving a high V₉h (low Vₑ) metal gate for future, more aggressively scaled stacks (EOT<0.9 nm) continues to be a challenge due to the V₉h roll-off issue. Low temperature approaches have been employed to eliminate thermal budget concerns [1] and also as a mechanistic tool to understand V₉h roll-off associated with high temperature process [7]. This report for the first time presents a PMOS mechanism capable to suppress V₉h roll-off and achieve band edge PMOS work function even at EOT ~9 Å at ~<0.1 A/cm² at V₉=g=V₉h=0.9 V through comprehensive study on V₉h roll-off. This study also reports a low temperature WF tuning of the same electrode for n-type WF using a doping process. A study on device failure mechanism for n-type metal gate is also discussed.

Experimental

Terraced oxide capacitors (with the SiO₂ thicknesses in the 1-4 nm range) were fabricated using Hf-based high-k (of 2 to 3 nm physical thickness) followed by the TiN films capped with doping layers. The polySi, film was then deposited and the patterned gate stack was annealed at ~1000°C, then followed by the final forming gas anneal. The TiN work function tuning process is applied after S/D activation.

Results and Discussion

P-type WF. Obtaining a P-type WF value is a challenging task due to the well-known V₉h roll-off phenomenon, which becomes more pronounced with higher electrode WF [8]. Based on the earlier proposed understanding of the V₉h roll-off mechanism we implemented several options for its suppression. The proposed model explains the V₉h roll-off phenomenon as caused by a build-up of positive charges attributed to the oxygen vacancies at the SiO₂/Si interface, as shown in Fig. 1(a) [8]. These vacancies are formed by the metal/high-k stack induced oxygen out-diffusion from SiO₂ driven by high-temperature processes. According to the model, the V₉h roll-off can be reduced by using a thinner high-k film, although this is a rather limited option. Alternative approaches are to block oxygen out-diffusion from the interface by inserting an inert barrier layer between the SiO₂ and high-k (fig. 1(b)), substituting O for F in the SiO₂ (fig. 1(c)) and adding deuterium (D) in the SiO₂ (fig. 1(d)). The stronger Si-F bonds are expected to reduce the roll-off [9] and less deuterium dissociation (WF) by capping of TiN with tuning species (TS). TiN WF is in a range of 4.1~4.6 eV depending on TiN thickness and following annealing temperature. Low WF is most likely due to TS element incorporation. The obtained high 5.0 eV WF is confirmed by the direct barrier height measurements (fig. 6) [10]. Fig. 7 shows excellent C-V and gate leakage characteristics of the optimized MOSFET V₉h=0.15 V at EOT=0.93nm and leakage current ~<0.1 A/cm² at V₉=g=V₉h=0.9 V. Lower gate leakage current is attributed to higher barrier height associated with higher WF. These results are well suited for sub 28 nm node low power application.

N-type WF. Fig. 8 shows tunability of the TiN work function (WF) by capping of TiN with tuning species (TS). TiN WF is in a range of 4.1–4.6 eV depending on TiN thickness and following annealing temperature. Low WF is most likely due to TS element accumulation at the HK/TiN interface (fig. 9). The TiN WF change due to TS doping is also confirmed by the direct measurements of the barrier height change using the J-V technique [10] as seen in Fig. 10. Low temperature tuning process requires different dopants rather than lanthanide metal elements (La, Gd, Sc) that were usually utilized in high temperature tuning approaches because of the very different process window. However, TS doping may also cause device failure when the annealing is performed at the above critical temperature. The J-V characteristics of the terraced oxide MOSCs with the TiN/TS gate stack annealed at 480°C show initial failure regardless of the bottom SiO₂ thickness, Fig. 11 (inset TEM images in fig. 11). Backside SIMS data show that annealing at 480°C results in a significant TS diffusion through the TiN, HK and even thick SiO₂ films (fig. 12). TS clusters are also observed in the active area after Si is removed for the backside SIMS measurements (fig. 13). Further analyses suggest the TS-induced device failure could be due to a conductive filaments formation in the gate stack by the localized TS diffusion at a critical temperature. Fig. 14 shows excellent C-V and gate leakage characteristics of the optimized NMOS showing -0.9 V V₉h at 0.93nm EOT and ~<0.1 A/cm² at V₉=g=V₉h=0.9 V.

Conclusion

To obtain high WF, we employed a low work function metal, which is characterized by minimal V₉h roll-off, combined with low temperature tuning process to increase the effective WF value while maintaining minimal V₉h roll-off. This approach has resulted in a targeted WF value of 5.0 eV without EOT penalty. Low WF was achieved by using TiN/TS and a low temperature process to control TS doping. Annealing above a certain critical temperature is shown to result in device failure due to a metal-induced conductive filament formation. This CMOS-compatible tuning process can also be applied to both dielectric-last/metal-
last and dielectric-first/metal-last integration schemes.


Fig. 1 (a) Schematic $V_{th}$ roll-off model based on oxygen deficiency in IL. Based on the model, several approaches are suggested to oxygen vacancy (b) SiN$_x$ layer between high-k and IL for blocking oxygen or oxygen vacancy transport, (c) F ion implantation for oxygen vacancy passivation, (d) pre-incorporation of D into IL, and (e) low temperature WF tuning of no or negligible $V_{th}$ roll-off metal gate.

Fig. 4 Low temperature tuning process in both gate first and gate last process boost WF but it contributes to EOT increase and does not help suppress $V_{th}$ roll-off.

Fig. 9 Backside SIMS of gate stack with low WF TiN metal gate (split #01 in Fig. 1). Analyses were made after backside Si is completely removed. TS is diffused through TiN and accumulated at HK/TiN interface causing WF shift to 4.1eV.

Fig. 10 Barrier height measurement confirms WF change of TiN (split #01 and #05 in Fig. 8).

Fig. 11 J-V characteristics of TiN/TS MOSCs with (a) thin and (b) thick bottom SiO$_2$ that got annealing at 480°C.

Fig. 12 Backside SIMS profile of 4.0nm-IL/HK/TiN/TS MOSCs that got annealing at 400°C and 480°C.

Fig. 13 Microscopic images of backside area that got SIMS sputtering. Defects are observed on the sample got annealing at 480°C.