Performance Analysis of Sub-22nm Metal Gate Fully Depleted SOI MOSFET with High-K Gate Dielectric

Niranjan Jyoti Borah\textsuperscript{a}, Gagnesh Kumar\textsuperscript{b}

\textsuperscript{a} M.Tech. VLSI Design Automation and Techniques, National Institute of Technology Hamirpur
\textsuperscript{b} Assistant Professor, National Institute of Technology Hamirpur
\textsuperscript{a} borah_niranjan86@yahoo.cm, \textsuperscript{b} gagnesh@nith.ac.in

Abstract: With the growing technology, the density of transistors on chips has been periodically doubling, as predicted by Moore’s Law. In order to achieve this, the device dimensions have been continuously scaled down. Leakage currents need to be controlled which come into effect due to continued scaling of gate lengths, gate oxides and threshold voltages as per the principles of scaling. The leakage currents contribute significantly to power dissipation, which is a matter of concern especially in small sized devices. This paper briefly discusses the use of high-k dielectric in combination with metal gates for deep sub-micron technology. How to have an effective control over the channel while minimizing device leakages is also been discussed.

Keywords: High-k gate dielectric, Fully-depleted SOI, metal gate

1. INTRODUCTION

In order to achieve higher packaging density coupled with high speed and lower power dissipation, CMOS devices have been scaled for more than three decades according to the Moore’s law. As the channel length is scaled down, threshold voltage ($V_{th}$) and supply voltage ($V_{dd}$) are also scaled to maintain large overdrive currents and reduce high field effects respectively. Along with the horizontal dimensions, vertical dimensions, for example, gate insulator thickness ($T_{ox}$) and junction depth ($X_{j}$) are also scaled down. But as the devices are scaled towards deep sub-micrometre regime, various leakage mechanisms or short-channel effects predominate which increase the static power dissipation per device per unit area. Since the advent of MOS devices over 40 years ago, SiO$_2$ has served as the transistor gate insulator of choice. There are many advantages of using SiO$_2$ as gate insulator, such as: i) Thermodynamically stable on Si ii) High quality SiO$_2$-Si interface iii) Interface state density, $10^{10}$ states/cm$^2$ iv) Excellent insulator. Gate insulator thickness was scaled at ~0.7x per generation up to the 130nm node, but scaling slowed at the 90nm and 65 nodes as SiO$_2$ ran out of atoms and gate leakage power limited further scaling [1]. Thinning of SiO$_2$ results in: i) leaky gate oxide ii) Poly-Si gate depletion [2] iii) Boron penetration [3]. The thickness of SiO$_2$ layer required in 45nm technology is about 1.2nm which is 4 atomic layers deep. At this gate thickness quantum mechanical tunneling of electron takes place. This tunneling current increases exponentially with decrease in oxide thickness. Further in case of thin oxide, implant dose of low energy is used for polysilicon gate. Polysilicon near the gate oxide is lightly doped which results in depletion region at the polysilicon and gate oxide interface, whose thickness becomes very much comparable to oxide thickness. This results in increase in threshold voltage, decrease in drain current and also a part of the applied gate voltage falls across this depletion region. Additionally boron penetration from a P+ poly gate electrode through the thin gate oxide into the silicon substrate can decrease the threshold voltage.

2. THE HIGH-K OXIDE SOLUTION

To reduce leakage, a thicker oxide layer is required. But this means less control over channel charge. A gate insulator is required that is thick enough to keep electrons from tunneling through it and yet permeable enough to let the gate’s electric field into the channel, so that it could turn on the transistor. In other words, the material had to be physically thick but electrically thin. To continue the scaling trend we need to increase gate-channel capacitance which given by,

$$C_{g} = \frac{\varepsilon_{ox}}{t_{ox}} \quad (1)$$

Where, $\varepsilon_{ox}$ is dielectric constant of oxide, $t_{ox}$ is the thickness of oxide. Since scaling of SiO$_2$ further is not possible so the alternate solution is to use gate oxide of another material having higher dielectric constant. By
using high-K dielectric both $\varepsilon_{ox}$ and $t_{ox}$ can be increased while maintaining the gate-channel capacitance [4] which can be shown as,

$$C_g = \frac{K_{\text{SiO}_2}\varepsilon_0}{t_{\text{SiO}_2}} = \frac{K_D\varepsilon_0}{t_D} = \frac{K_{\text{SiO}_2}\varepsilon_0}{EOT} \tag{2}$$

where $K_{\text{SiO}_2}, K_D$ are permittivity of SiO$_2$ and high-k oxide material respectively, EOT is the equivalent oxide thickness.

3. CHOICE OF HIGH-K OXIDE

High-K oxide should satisfy the following properties [4]: i) High dielectric constant and barrier height ii) Thermodynamic stability iii) Interface quality iv) Gate compatibility v) Process compatibility vi) Fixed oxide charge. Dielectric constant ($K$) and band gap (Eg) of a given material generally exhibit an inverse relationship. However HfO$_2$ and ZrO$_2$ offer higher value of $K$ and Eg. Moreover HfO$_2$ and ZrO$_2$ are stable in direct contact with Si up to high temperature. Most of high-K materials having interface state density, Dit~10$^{11}$-$10^{12}$ states/cm$^2$ have been reported which is near about to interface state density for SiO$_2$, Dit~2×10$^{10}$ states/cm$^2$. Over or under constrained interface will lead to high interface defect density. The choice of high-k gate oxide should be such that it is compatible with gate electrode and also easily be deposited without any formation of interfacial layer.

4. CHALLENGES FACED IN REPLACING SiO$_2$ WITH HIGH-K DIELECTRIC

Transistors built with poly-Si/high-K dielectric stack suffer from two major drawbacks. First, high-K dielectric and poly-Si are incompatible due to the Fermi level pinning at the poly-Si/high-K interface which results to high threshold voltage and low drive current [5]. When gate metals are deposited on SiO$_2$, each metal has n-metal and p-metal workfunctions, however on high dielectric oxide both gate electrodes show mid-metal work function. That is the workfunction of metal gets changed from its value in vacuum. The most likely cause of Fermi level pinning is due to interface defects formed at the poly-Si/high-K interface as illustrated in Figure 1. Second, poly-Si/high-K transistors exhibit severely degraded channel carrier mobility due to the coupling of low energy surface optical phonon modes arising from the polarization of the high-K dielectric to the inversion channel charge carriers [6]. The dielectric layer is made up of metal-oxygen bonds, which also give rise to low energy optical phonons which can be modelled as oscillating dipoles. This is the very aspect that gives the high-K dielectric such a high dielectric constant. These oscillating dipoles couple strongly with channel electrons when the gate plasma oscillations and the phonons in the high-K dielectric are in resonance. This resonance occurs when the gate carrier density is ~$1\times10^{18}$ cm$^{-3}$, as is the case with standard doped poly-Si gate in depletion, and the corresponding gate plasmon energy falls within the dominant LO (longitudinal optical) and TO (transverse optical) energy modes of the high-K dielectric. This resonance condition leads to significant degradation of channel carrier mobility.

![Figure 1. Defect formation at the poly Si and high-K interface is most likely the cause of the Fermi level pinning which causes high threshold voltages in MOSFET (M=Zr or Hf) [5]](image)
5. THE METAL GATE SOLUTION

The drawbacks of using poly-Si/high-K dielectric stack can be removed by using metal/high-K dielectric stack. When metal gate is used instead of poly-Si there is less amount of defects at the surface solving the fermi-level pinning to a certain extent. When the density of electrons in the gate electrode is increased significantly the resonance condition which leads to significant degradation of channel carrier mobility is not satisfied [7]. This weakens the carrier phonon coupling and leads to a recovery of the surface phonon-limited mobility. This alleviates the mobility degradation problem. One way to do that would be to switch from a polysilicon gate to a metal one, where the free carrier density exceeds 1×10^{20} cm^{-3}. The other advantage of using metal gate is that effectively there is no formation of depletion layer and its resistance is less compared to poly-Si.

The metal gates with correct work functions facilitates the right transistor threshold voltages, screen out the mobility degradation problem and enables high-performance of high-k/metal gate transistors with low gate dielectric leakage. For conventional planar CMOS applications on bulk silicon, an n+ metal work function is required for the NMOS transistor while a p+ metal work function is needed for the PMOS transistor, in order to satisfy the correct transistor threshold voltages. For fully-depleted single and multiple SOI MOSFETs, the use of midgap metals is very attractive to enable high performance CMOS transistors.

6. DEVICE STRUCTURE

To study the effect of use of metal gate in combination with high-K dielectric a schematic cross-sectional view of the SOI MOSFET is simulated using 2-D Sentaurus TCAD device simulator, is shown in the Figure 2. Gate length of the device that had been concentrated is 22nm. Gate electrode thickness, gate oxide thickness, silicon film thickness and BOX thickness are 10nm, 3.2nm, 4nm and 20nm respectively. Light channel doping concentration of 1×10^{17} cm^{-3} is given to avoid degradation of carrier mobility. The silicon film thickness has been chosen less than one fourth of the channel length to reduce short channel effect. The doping concentration of S/D region is kept at 1×10^{19} cm^{-3}. Simulation has been done taking an n channel MOSFET with gate dielectric of different dielectric value and then with different work function of metal gate. The gate dielectric used for the simulation are SiO_2 (K=3.9), Al_2O_3 (K=9), Y_2O_3 (K=15) and HfO_2 (K=25).

![Figure 2. Illustration of simulated structure of FD-SOI NMOSFET](image)
7. RESULTS AND DISCUSSIONS

Different performance parameter of FD-SOI MOSFET has been analyzed for gate dielectric of different dielectric value and then for gate electrode of different work function.

7.1 Leakage Current (OFF Current)

From Figure 3 it is seen that as dielectric value of gate insulator is increased the tunnelling current from gate to channel and vice-versa decreases resulting in lower leakage current. From Figure 4 it is also clear that leakage current decreases with increase the metal gate work function. This is because as we increase the work function of metal gate, threshold voltage increases and the leakage current of SOI MOSFET decreases exponentially with threshold voltage.

![Figure 3. OFF-state current for different gate dielectric](image1)

![Figure 4. OFF-state current for different metal gate work function](image2)

7.2 On Current to OFF Current Ratio (ION/IOFF)

ON current to OFF current ratio is a vital performance parameter for SOI MOSFET. Higher I\textsubscript{ON}/I\textsubscript{OFF} ratio provides higher switching speed. From Figure 5 it is seen that as we increase the dielectric constant of gate oxide on state current increases and off state current decreases so I\textsubscript{ON}/I\textsubscript{OFF} increases. It is also seen from the Figure 6 that I\textsubscript{ON}/I\textsubscript{OFF} ratio increases with the increase in metal gate work function.
7.3 Short Channel Effects

Figure 7 shows the subthreshold swing characteristic for the 22nm SOI device with different dielectric value of gate oxide. Small value of subthreshold slope means a small change in the input bias can modulate the output current considerably. The smaller the value of subthreshold slope the better the transistor is as a switch. As the dielectric value is increased, the effect of gate’s electric field into the channel increases which in turn increases the gate controllability and hence subthreshold swing is improved. Figure 8 also shows the improvement of subthreshold slope with the increase in work function of the metal gate.

Figure 9 and Figure 10 shows that DIBL is reduced by increase in dielectric value of gate oxide as well as by increase in metal gate work function.

Figure 7. Subthreshold slope variation for different gate dielectric
Threshold Voltage Variation

From simulation results of Figure 11 shows that the threshold voltage of SOI MOSFET increases linearly with the increase in metal gate work function because higher Vgs is required to invert the channel. So by providing the different work function of metal gate, we can set the appropriate threshold voltage of SOI MOSFET.

In order to maintain Ioff very low, it is necessary to increase the metal work function, but doing so increases the threshold voltage. So a compromise has to be done between these two parameters. Figure 12 shows that with around 4.6eV metal gate work function optimum value of threshold voltage and leakage current can be achieved.
8. CONCLUSION

From simulated results it can be concluded that metal gate electrode in combination with high-k gate dielectric are required for enabling continued equivalent gate oxide thickness scaling, and hence high performance, and for controlling gate oxide leakage. Work function engineering is an excellent way for choosing correct metal gate work function as threshold voltage shows a linear relationship with metal gate work function. For thin body CMOS devices the range of gate work is 4.4eV to 5eV that can be easily done with metallic gates. From simulated result it can be said that at 4.60eV metal gate work function SOI MOSFET device shows better performance than any other device. It can be concluded that replacing the conventional poly-Si gate and SiO$_2$ gate dielectric with metal gate electrode in combination with high-k gate dielectric will led us to continue the scaling trend as predicted by Moore’s law.

REFERENCES


