Capacitance Degradation due to Fringing Fields in Deep Sub-Micron MOSFETs with High-K Gate Dielectrics

A. Inani, *V.R. Rao, †B. Cheng, #P. Zeitzoff, and J. C. S. Woo  
Department of Electrical Engineering, University of California,  
Los Angeles, CA 90095-1594, USA  
*Indian Institute of Technology, Bombay, India  
†APRDL, Motorola, Austin, USA  
#SEmatech Inc., Austin, USA

1. Abstract
High-K gate dielectrics have been under extensive investigation for use in sub-100nm MOSFETs to suppress gate leakage. However, thicker gate dielectrics can result in degradation of the electrical performance due to increased fringing fields from the gate to source/drain. In this paper, the capacitance degradation resulting from this effect is analyzed and a simple technique to model this effect is presented.

2. Introduction
Scaling of MOS devices into the sub-100nm regime requires a reduced oxide thickness (<3nm) to control the short channel effects. This can lead to undesirably large direct tunneling gate leakage currents [1]. To alleviate this problem, alternative gate insulators with higher dielectric permittivities are under extensive investigation. However, not only does introducing a new material in the CMOS process have its own constraints, but the electrical characteristics of the device are also degraded due to the large fringing fields resulting from a larger physical thickness of such gate dielectrics [2]. In this work, we discuss fringing field induced degradation in sub-100nm devices, e.g., undesirably large C_GS_D, which is known to worsen circuit performance, and reduced C_GSUB, which exacerbates short channel effects and lowers I_on. These effects have also been modeled using basic electromagnetic theory to gain insight into the physics behind the fringing fields.

3. Simulation
Several possible gate stack structures suggested in literature, shown schematically in Fig. 1, have been studied. Structure A uses a high-K dielectric for the gate insulator, sidewall and all other insulation. B has a high-K dielectric under the gate electrode and sidewalls while the remaining insulation is SiO_2. Structure C has a high-K dielectric only under the gate electrode. The sidewalls and all other insulation are formed using SiO_2. D has a high-K dielectric under the gate and all other insulation material has a permittivity of 1, corresponding to air. This structure is used to bring out a very important point as will be clear in later discussions. The “effective” gate oxide thickness is kept at 1.5nm in all simulations. The thickness of the gate insulator is therefore determined by:

\[ T_K = 1.5 \times (K/3.9) \text{ (nm)} \]  (1)

for a particular value of K. Furthermore, to
isolate the capacitance degradation purely due to fringing fields, the effects of polydepletion and quantum models for the inversion layer were not included in the simulations. The metal contacts to the silicided S/D region are assumed to be two λ (0.1μm) away from the gate electrode. L is 100nm unless specified otherwise and all other physical parameters have been kept the same in the simulations performed using SILVACO tools [3].

4. Results and Discussion

To quantitatively understand the effect of fringing electric fields, the increase in the total gate capacitance of the device has been studied. For this, the capacitance has been extracted from low frequency C-V simulations in the deep inversion region. The effect of increasing the K value of the dielectric (and hence the thickness of the gate insulator, $T_K$) is to increase fringing fields from the gate to the S/D electrodes. As shown previously [2], a good parameter to model this effect is the aspect ratio $T_K/L$ (gate insulator thickness to channel length ratio). In Fig. 2, we plot the ratio of the gate capacitance obtained from these simulations ($C_{tot}$) to the theoretical capacitance ($C_{thr}$=$A_G$($K\varepsilon_0$)$T_K$) versus the $T_K/L$ ratio. As shown here, the capacitance degradation in structure A is the worst and gets better as we approach structure D, where the smaller increase in the $C_{GDS}$ is compensated at the expense of a reduced $C_{GSUB}$. For example, for K=30 (~Ta$_2$O$_5$), the increase in the capacitance due to fringing fields is 25% in the worst case (Struct A) while $C_{tot}$ remains close to $C_{thr}$ in structures C and D. This is because with a high-K material between the gate and S/D electrodes, the fringing capacitance, $C_F$, is aggravated. $C_F$ is reduced as the K value of the sidewall or insulating material decreases. Fig. 3 shows typical vertical electric field in structures A, C and a conventional structure using SiO$_2$ as gate insulator. The magnitude of the (fringing) E-field over the S/D region is the largest in A followed by C and the conventional structure respectively. Fig. 4 shows the variation in the capacitance by changing the K value of the sidewall with a high-K gate insulator (K=200). The capacitance ($C_{tot}$) increases from close to $C_{thr}$ to over 2 times the theoretical value as the K value of the sidewall varies from 3.9 to 200. It is clear that a combination of a high-K material for gate insulator with a low K sidewall is very effective in suppressing the effects of
5. Modeling

The effect of fringing fields on parallel plate capacitors has been modeled in the past using the Method of Moments (MoM) [4]. We use a similar technique to understand the charge distribution in MOS devices including the effect of fringing fields. For this purpose, a MOSFET has been approximated as a parallel plate capacitor with metal plates, with a longer bottom plate to account for the presence of the S/D electrodes (insert in Fig. 5). In this work, the length of the S/D regions has been fixed at 0.25μm. With this topology, simulations have been performed with MoM using MATLAB [5] to study the worst case discussed above (Struct A). The capacitance of the system has been evaluated for a range of values of $T_K/L$. Fig. 5 shows the ratio of $C_{GSOB}/C_{thr}$ versus $T_K/L$ by varying $T_K$ (and keeping $L$ constant) and by varying $L$ (and keeping $T_K$ constant). The capacitance increase is almost 4 times when $T_K/L$ is close to 1.5. These results match very well with those obtained by 2-D device simulations. The small discrepancy is due to an over simplified representation of the MOS capacitor. Typical charge distributions over the two plates of the system for two extreme cases of $T_K/L$ are plotted in Fig. 6. This clearly shows an increased charge concentration in the S/D regions as $T_K/L$ increases. Fig. 7 shows the various components of the gate capacitance obtained from these simulations with variation in $T_K/L$. In the worst case, the gate-to-source/drain capacitance becomes almost equal to the gate-to-substrate capacitance. We model these effects as an increase in the insulator thickness and express it as an equivalent electrical oxide thickness extracted from the gate-to-substrate capacitance. Note that a decreasing $C_{GSOB}$ (i.e., higher $T_{ox}$(elec)) implies worse short channel effects such as DIBL and $V_{th}$ roll-off. Fig. 8 shows that as $T_K/L$ increases (by varying $L$), the $T_{ox}$(elec) increases rapidly even though the $T_{ox}$(eff) is kept at 1.5nm.
(Eq. 1). Also shown in Fig. 9 are the results obtained for Struct C from 2-D device simulations ($T_K/L$ varies by varying $T_K$). In this case, $T_{ox}(elec)$ has been extracted as an equivalent oxide thickness leading to the same DIBL in a conventional MOS device. Again, $T_{ox}(elec)$ increases rapidly for $T_K/L$ larger than 0.5 even though the $T_{ox}(eff)$ is kept at 1.5nm. However, as mentioned earlier, the degradation in Struct C is substantially less than in A.

6. Conclusion

In this paper, the fringing field induced degradation caused by high-K gate dielectrics in sub-100nm MOS devices is discussed quantitatively. It is found that a good means of suppressing these effects is to use a combination of dielectric values for the gate insulator and the sidewall. We have also demonstrated the use of a simple technique based on basic EM field theory as an effective means of studying fringing fields in MOS devices.

7. References


Figure 6: Charge density distribution in the parallel plate system shown in the insert of Fig. 4 for cases of (a) $K=200$ and (b) $K=10$. Simulations were done using MATLAB with the Method of Moments. Note that the z-axis scale is different in the two figures. $L=100\text{nm}$ in both cases.

Figure 7: Ratio of the two components of capacitance (gate-to-source/drain and gate-to-substrate) to the total gate capacitance ($C_{tot}$) versus the aspect ratio $T_K/L$ obtained using MATLAB simulations with MoM. $L=100\text{nm}$ in all cases.

Figure 8: Variation of the electrical oxide thickness ($T_{ox}(elec)$) obtained from 2-D device simulations of Struct C (variation in $T_K$) and MATLAB simulations of Struct A with the Method of Moments (variation in $L$).