MODELING OF ACCUMULATION MOS CAPACITORS FOR ANALOG DESIGN IN DIGITAL VLSI PROCESSES

Shanthi Pavan, Yannis Tsividis, and Krishnaswamy Nagaraj
Texas Instruments
15 Independence Boulevard
Warren, New Jersey NJ 07059

Department of Electrical Engineering
1318 S.W. Mudl Building, 500 W 120th Street
Columbia University, New York NY 10027

ABSTRACT

In this paper, we present an explicit model for the MOS capacitor in accumulation and compare it with device simulation and measurement results. The model is physically-based and tracks substrate doping concentration and temperature, and incorporates the polysilicon gate depletion effect. Harmonic distortion generated when the proposed model is used in circuit simulation agrees with that obtained by using a very closely spaced piecewise linear model generated by device simulation.

1. INTRODUCTION

There has been a growing interest in implementing mixed analog-digital circuits on a single chip. For economical reasons, the preferred technology is CMOS. In low-cost digital CMOS processes, high density poly-poly capacitors are not available. Metal-metal capacitors can be used, but their specific capacitance is very low. The focus of this work is the use of MOS capacitors in place of metal-metal capacitors. This means that the regions of interest are the "flat" parts of the C-V characteristic. A two terminal MOS structure can be operated in accumulation or inversion. It can be shown that for typical process parameters, and a given gate-body bias, accumulation capacitors tend to be more linear. For this the reader is referred to an excellent paper [1]. Hence, we do not consider inversion capacitors. Moreover, existing SPICE MOS transistor models are not accurate in accumulation since it is typically not important for transistor operation. The above two factors are the motivation behind this work.

In this paper, we present an explicit physics-based simulation model for the two-terminal MOS capacitor in accumulation. In Section 2, we derive the model from first principles and compare it with results obtained by numerical solution of the device equations. Model implementation issues are considered in Section 3. In deep submicron CMOS technologies, it is not possible to dope the gate with arbitrarily high concentrations of acceptors/donors. Then the assumption that the potential throughout the gate is constant is not valid anymore. Finite gate doping causes what is known as the "polysilicon gate depletion effect." This phenomenon is analyzed and incorporated into our model in Section 4. Section 5 compares the complete model derived in Section 4 to capacitor measurements in some typical fine line CMOS processes. Section 6 contains the conclusions of this work.

2. THE TWO TERMINAL MOS STRUCTURE

Consider the two-terminal poly-n-well MOS structure shown in Figure 1. For a given bias voltage \( V_{GB} \) applied across the device, we present relations for charge, capacitance, and surface potential. Derivations are not shown here for lack of space, the reader is referred to [2]. \( \psi_{ox} \) is the potential drop across the oxide, \( \psi_s \) is the surface potential, \( \phi_{MS} \) is the potential corresponding to the work function difference of the bulk and the gate materials, \( Q_G \) and \( Q_e \) denote the charge per unit area on the gate and in the semiconductor respectively, \( \mu_{ox} \) is the effective interface charge per unit area of the oxide, \( \varepsilon_s, N_D \) and \( \phi_t \) stand for the permittivity of silicon, donor density in the n-well and the thermal voltage respectively. \( C_{ox} \) is the capacitance of the oxide per unit area and \( \phi_P \) is the Fermi potential of the n-well. Using Kirchhoff's voltage law and charge conservation along with Poisson's equation in the y-direction, it can be shown that in accumulation, after neglecting the presence of holes in the n-well [2],

\[
Q_e = -\sqrt{2\varepsilon_s N_D} \sqrt{\phi_t - \psi_s - \phi_t} \\
V_{GB} = V_{FB} + \psi_s + \gamma \sqrt{\phi_t - \psi_s - \phi_t} \\
Q_c + Q_e + Q'_e = 0 \\
where \gamma = \frac{\sqrt{2\varepsilon_s N_D}}{C_{ox}} \\
V_{FB} = \phi_{MS} - \frac{Q_e}{C_{ox}}
\]

For a given bias \( V_{GB} \), we can solve (2) numerically for \( \psi_s \) and calculate \( Q_e \) using (1). However, if an explicit expression for the surface potential can be obtained, calculations become much more straightforward. It can be found that, for typical values of \( \gamma \) and temperature, such an explicit expression is (see Appendix A),

\[
\psi_s = 2\phi_t \left[ \frac{V_{GB} - V_{FB} + \gamma \phi_t}{V_{GB} - V_{FB} + k_2 \phi_t} \right] \log \left( 1 + \frac{V_{GB} - V_{FB}}{\gamma \phi_t} \right)
\]

where \( k_1 = 3 \) and \( k_2 = 6 \) provide a good fit to a numerical solution of (2), for a wide range of process parameters. The capacitance of the two-terminal structure per unit area \( C_{ph} \) is obtained from

\[
\frac{1}{C_{ph}} = \frac{1}{C_{ox}} + \frac{1}{C_c}
\]

where \( C_c = -\frac{\partial Q_e}{\partial \psi_s} \) is the capacitance of the accumulation layer per unit area. Differentiating (1),

\[
C_c = \gamma C_{ox} \frac{\exp(\frac{\psi_s}{\gamma}) - 1}{2 \sqrt{\phi_t - \psi_s - \phi_t}}
\]

Figure 2 compares the capacitance obtained by numerical methods with the explicit model proposed in this paper. \( C_{ox} \) was kept constant and the \( N_D \) values used were \( 10^{16} \) cm\(^{-3} \), \( 10^{17} \) cm\(^{-3} \) and \( 10^{18} \) cm\(^{-3} \). Note that this corresponds to the variation of \( \gamma \) by a factor of ten.

0-7803-5471-0/99/$10.00 © 1999 IEEE

VI-202
3. MODEL IMPLEMENTATION

In this section we examine issues that arise in implementing the model in a general purpose circuit simulator. In this work, the model was implemented in TISFICE, an in-house circuit simulation program at Texas Instruments. The model is implemented as a user defined utility, coded as a ‘C’ routine. \( C_{ox} \) is calculated using (7) for use in AC simulations. For transient simulations, \( Q' \) is needed. One needs to be careful with the way in which it is obtained. One could use (3) and (1) to obtain

\[
Q' = -Q_s - Q_o = \sqrt{2q \epsilon_s N_D} \frac{\phi_t \exp(\frac{\psi_s}{\phi_t}) - \psi_s - \phi_t - Q_o}{\phi_t} \tag{9}
\]

Alternatively

\[
Q' = C_{ox} \psi_o = C_{ox}(V_{GB} - \psi_s - \phi_{MS}) \tag{10}
\]

The two approaches would yield identical results if \( \psi_o \) was known exactly. However, our expression (6) for \( \psi_o \) is approximate. When the applied gate bias is very large, even a small error in determining the surface potential would lead to a large error in computing charge from (9). On the other hand, (10) has no such problem. Hence, this expression should be used to calculate capacitor charge.

The performance of the implemented model was compared with a very finely spaced (voltage steps of 1 mV) piecewise linear capacitor model generated from device simulation. A MOS capacitor was driven with a sinusoidal voltage source and the distortion in the capacitor current was used as a metric for comparison of the two models. Distortion depends on derivatives of the C-V characteristic, and is hence a very sensitive indicator of modeling accuracy. Since third harmonic distortion is very small at high bias voltages and/or low signal levels, simulations were run at very tight tolerances, and using a very small timestep compared to the period of the voltage source. The dynamic range of the simulations is about 100 dB.

Figure 3 and 4 compare the levels of second and third harmonic currents.

3.1. Circuit considerations and choice of polysilicon type

Until now, the results we derived for \( Q_{G} \), \( Q' \), and \( \psi_o \) were independent of the type of polysilicon used as the gate material. We now examine the two choices for the the type of polysilicon (n or p) from a circuit design viewpoint. As mentioned in the introduction to this paper, we are interested in using MOS accumulation capacitors in place of metal-metal capacitors. We would like to be operating in deep accumulation even for a very small \( V_{GB} \). Hence, the flatband voltage \( V_{FB} \) should be as negative as possible.

In modern CMOS technologies \( C_{ox} \) is very large, and the term \( C_{ox}/\phi_t \) in (5) is typically about 100 mV. If the gate is degenerately doped (either n+ or p+), \( \phi_{MS} \) is given by

\[
\phi_{MS} = \begin{cases} 
-\phi_p - 0.56V, & \text{n+ poly} \\
-\phi_p + 0.56V, & \text{p+ poly} 
\end{cases} \tag{11}
\]

From the above equation, n+ polysilicon is the one of choice. We will henceforth assume that this is the case.

4. THE POLYSILICON GATE DEPLETION EFFECT

The analysis presented in the preceding sections assumes that the “top plate” of the MOS capacitor is infinitely highly doped. In practice, the finite doping density of the n+ polysilicon causes a depletion layer in the gate, whose thickness depends on the amount of negative charge in the n-well. In the literature\(^3\)\(^4\), the effects of gate depletion are analyzed when the MOS structure is in inversion. Accumulation is not considered as this region is not important for transistor operation.

We redraw the cross section of the MOS capacitor in Figure 5, now with a depletion layer in the gate. Assume that the gate is uniformly doped n type with a donor concentration of \( N_{POLY} \). Let \( \psi_{dep} \) denote the potential drop across the gate depletion layer. Let a voltage \( V_{GB} \) be applied to the structure as shown. The potential balance equation can be written as

\[
V_{GB} = \phi_{MS} + \psi_{dep} + \psi_o + \psi_s \tag{12}
\]

Charge balance as given by (3) still holds. Using this along with \( \psi_o = Q_{G}/C_{ox} \) in (12), we get

\[
V_{GB} = \phi_{MS} + \psi_{dep} - \frac{Q_{G}}{C_{ox}} + \psi_s \tag{13}
\]

Using (5), the above equation can be rewritten as

\[
V_{GB} = V_{FB} + \psi_{dep} + \psi_s - Q_{G} \frac{1}{C_{ox}} \tag{14}
\]

\( Q' \) is only a function of \( \psi_o \) and is given by (1). Using (1) with (14)

\[
V_{GB} = V_{FB} + \psi_s + \psi_{dep} + \gamma \sqrt{\phi_t \exp(\frac{\psi_s}{\phi_t}) - \psi_s - \phi_t} \tag{15}
\]

From basic electrostatics,

\[
Q_{G} = \sqrt{2q \epsilon_s N_{POLY} \psi_{dep}} \tag{16}
\]

Using \( Q' = C_{ox} \psi_o \) alongwith (12) we have

\[
Q' = C_{ox} \psi_o = C_{ox}(V_{GB} - \phi_{MS} - \psi_s - \psi_{dep}) \tag{17}
\]

Hence,

\[
\sqrt{2q \epsilon_s N_{POLY} \psi_{dep}} = C_{ox}(V_{GB} - \phi_{MS} - \psi_s - \psi_{dep}) \tag{18}
\]

Defining \( \gamma_p = \sqrt{2q \epsilon_s N_{POLY} / \phi_t} \) enables us to rewrite (18) as

\[
\psi_{dep} + \gamma_p \sqrt{\psi_{dep}} - (V_{GB} - \phi_{MS} - \psi_s) = 0 \tag{19}
\]

Assuming \( \psi_s \) is known, the above equation can be solved to give

\[
\psi_{dep} = \left( \frac{\gamma_p}{\phi_t} \right)^2 \left[ 1 + 4 \right] \left( V_{GB} - \phi_{MS} - \psi_s \right) \frac{1}{\gamma_p} \tag{20}
\]

To get simpler analytic expressions, (20) can be further simplified by noting that for practical values of process parameters and reasonable bias voltages, \( \frac{\gamma_p}{\phi_t} \gg V_{GB} \).

Hence, (20) can be simplified to give

\[
\psi_{dep} = \frac{1}{\gamma_p} (V_{GB} - \phi_{MS} - \psi_s)^2 \tag{21}
\]

In order to compute \( \psi_s \), (20) (or the simplified form (21)) should be used for \( \psi_{dep} \) in (15). The resulting equation is implicit and even more complicated that (2). Since \( \frac{\gamma_p}{\phi_t} \gg V_{GB} \), \( \psi_{dep} \) is a very small fraction of \( V_{GB} \).
we can neglect $\psi_{dep}$, on the right hand side of (15), which can now be written as
\[ V_{GB} = V_{FB} + \psi_s + \gamma \sqrt{\frac{\phi_s}{\phi_l}} \exp\left(\frac{\psi_s}{\phi_l}\right) - \psi_s - \phi_l \] (22)
This is exactly (2). Hence, as in the case of an infinitely highly doped gate
\[ \psi_s \approx 2 \phi_l \ln \left( \frac{V_{GB} - V_{FB} + 3 \phi_l}{V_{GB} - V_{FB} + 6 \phi_l} \right) \log \left( 1 + \frac{V_{GB} - V_{FB}}{\gamma \phi_l} \right) \] (23)
$C_{pb}$ is calculated from
\[ \frac{1}{C_{pb}} = \frac{1}{C_{os}} + \frac{1}{C_c} + \frac{1}{C_{dep}} \] (24)
where $C_c$ is given by (8). $C_{dep}$ is the capacitance of the gate depletion region per unit area. Differentiating (16), we get
\[ C_{dep}' = \frac{1}{2} n_p N_{POLY} \frac{2 \psi_{dep}}{\sqrt{\psi_{dep}}} \] (25)
Using (21) in the above equation gives
\[ C_{dep}' = \frac{1}{2} \left( \frac{n_p N_{POLY}}{V_{GB} - \phi_{MS} - \psi_s} \right) \] (26)

5. MEASUREMENT RESULTS AND DISCUSSION

Figure 6 compares our model with measurement results made in a modern n-well salicided CMOS processes. The gate oxide thickness for the data of Figure 6 is about 85 A. When oxide thickness is reduced to a few tens of angstroms, it has been shown that the charge sheet model is quite accurate. The thickness of the charge sheet and quantum mechanical effects need to be taken into account [5] [6]. If these effects are not separately modeled, the value of $N_{POLY}$ may need to be changed to fit measurements.

6. CONCLUSION

In this paper, an explicit physics based simulation model was proposed for the MOS capacitor in accumulation. The results were compared to device simulation results and measurements. The model has been implemented in a general purpose circuit simulator (T1-Spice), and can be used to calculate distortion in circuits.

7. ACKNOWLEDGMENTS

The authors wish to acknowledge useful technical discussions with Keith Green and Bert Epler of Texas Instruments and Simon Willard of Silicon Systems.

APPENDIX 1

In order to find an explicit expression for $\psi_s$ in terms of $V_{GB}$, we rewrite (2) as
\[ \frac{V_{GB} - V_{FB}}{\phi_l} = \psi_s + \gamma \sqrt{\frac{\phi_s}{\phi_l}} \exp\left(\frac{\psi_s}{\phi_l}\right) - \psi_s - 1 \] (27)
Working with normalized variables
\[ z = \frac{V_{GB} - V_{FB}}{\phi_l} \] (28)
\[ x = \frac{\psi_s}{\phi_l} \] (29)
\[ a = \frac{\gamma}{\phi_l} \] (30)
we have
\[ z = x + a \sqrt{e^x - x - 1} \] (31)
We need to find $x$ in terms of $z$. Notice that the only parameter in the equation is $a$. If we could find an approximate explicit equation for $x$ for practical values of $a$, the surface potential will “track” changes in substrate doping concentration and temperature. Then, the model formulation is based entirely on the physics of the device, and is devoid of function fitting for particular values of process parameters and temperature.

As a guide towards an approximation, we consider the asymptotic behavior of (31). For small $x$, the exponential can be approximated by the first three terms of its Taylor expansion, whereas for large $x$ it becomes dominant. Thus, we obtain
\[ x \approx \begin{cases} \exp \left( \frac{z}{a} \right), & \text{small } x \\ a \exp \left( \frac{x}{2} \right), & \text{large } x \end{cases} \] (32)
Inverting these relations, we obtain:
\[ x \approx \begin{cases} \frac{a \sqrt{e^x - x - 1}}{z}, & \text{small } z \\ 2 \exp \left( \frac{z}{a} \right), & \text{large } z \end{cases} \] (33)
The second factor in these relations can be approximated by $\log(1 + z/a)$, for both small and large $x$. The first factor is typically 1 for small $z$, and becomes 2 for large $z$. This behavior can be satisfied by the function $2 \left( \frac{z + k_1}{z + k_2} \right)$, with $k_1$ and $k_2$ typically being 3 and 6, respectively. Thus the behavior in (33) can be approximated by:
\[ x \approx 2 \left( \frac{z + k_1}{z + k_2} \right) \log \left( 1 + \frac{z}{a} \right) \] (34)
Although this relation was developed from the asymptotic behavior of (31), we find that it approximates the latter well for the entire useful range of $z$ and for typical values of $a$, when $k_1 = 3$ and $k_2 = 6$ are chosen. Thus, rewriting (34) by using (28-30), we obtain (6).

REFERENCES

Figure 1. Cross section of an MOS capacitor.

Figure 2. Comparison of numerical simulation and proposed model for changes in substrate doping concentration (– numerical, ⊕ proposed model) $C_{ox} = 3.84 \, \text{fF/\mu m}^2$, $V_{FB} = 0$, $T = 300 \, \text{K}$.

Figure 3. 2nd harmonic relative to fundamental (– PWL, ⊕ proposed model).

Figure 4. 3rd harmonic relative to fundamental (– PWL, ⊕ proposed model).

Figure 5. MOS capacitor cross section indicating the polysilicon depletion region.

Figure 6. Comparison of complete model with measurements, $t_{ox} = 85 \, \text{A} (– \text{model}, \oplus \text{data})$.