3 / Capacitance-Voltage Characteristics

3.3 EXACT CHARGE ANALYSIS

The delta-depletion characteristics, as typified by Fig. 3.4, are a rather crude representation of reality. The first-order theory does a credible job for gate voltages comfortably within a given biasing region, but fails badly in the neighborhood of the transition points going from accumulation to depletion and from depletion to inversion. A more accurate modeling of the observed characteristics is often required in practical applications and is established by working with the exact charge distribution inside the MOS-capacitor. The exact charge analysis for the low frequency capacitance is quite tractable and is reviewed herein. The high frequency analysis, however, is very complex and involved. For this reason the high frequency results are simply quoted and appropriately combined with the low frequency results.

Let us consider an ideal MOS-capacitor with a low frequency ac signal applied to the gate of the device. When the ac gate voltage, v_g', is added to the dc gate voltage, V_G', the charge on the MOS-C gate is of course modified to $Q_G + q_g$, where Q_G and q_g are the dc gate charge per unit area and ac gate charge per unit area, respectively. Provided the device can follow the ac change in gate potential quasi-statically, the assumed case at low operational frequencies, one can state $Q_G(V_G') + q_g$ equals $Q_G(V_G' + v_g')$ or

$$q_g = Q_G(V_G' + v_g') - Q_G(V_G') = \Delta Q_G \tag{3.14}$$

Since, quite generally,

$$C = A_G \frac{q_g}{v_g} \tag{3.15}$$

we have, in the low frequency limit for the ideal structure,

$$C = A_G \frac{q_g}{v_g'} = A_G \frac{\Delta Q_G}{\Delta V_G'} \rightarrow A_G \frac{dQ_G}{dV_G'} \tag{3.16}$$

Equation (3.16) states that the low frequency capacitance can be determined by simply differentiating the dc expression for the gate charge. Instead of working with Q_G directly, it is more convenient to note that in an ideal structure the charge on the gate must balance the charge inside the semiconductor, or $Q_G = -Q_S$, where Q_S is the total semiconductor charge per unit area of the gate. We can therefore write

$$C = -A_G \frac{dQ_S}{dV_G'} = -A_G \frac{dQ_S}{dU_S} \frac{dU_S}{dV_G'} \tag{3.17}$$

The latter form of Eq. (3.17) suggests a mode of attack for completing the analysis. We already know, restating Eq. (2.42),

$$V_G' = \frac{kT}{q}\left[U_S + \hat{U}_S \frac{x_o'}{L_D} F(U_S, U_F)\right] \tag{3.18}$$

In addition, applying Gauss' law, we find that

$$Q_S = -K_S \varepsilon_0 \mathscr{E}_S = -\hat{U}_S \frac{kT}{q} \frac{K_S \varepsilon_0}{L_D} F(U_S, U_F) \tag{3.19}$$

Thus, performing the required differentiations, substituting into Eq. (3.17), and reorganizing the result, we conclude

$$C = \frac{C_O}{1 + W_{eff}/x_o'} \tag{3.20}$$

$$W_{eff} = \hat{U}_S L_D \left[\frac{2F(U_S, U_F)}{e^{U_F}(1 - e^{-U_S}) + e^{-U_F}(e^{U_S} - 1)/(1 + \Delta)} \right] \tag{3.21}$$

where

$$\Delta = 0 \qquad \text{in the low frequency limit} \tag{3.22}$$

and, in the high-frequency limit, for a p-type device,

$$\Delta = \begin{cases} 0 & \text{acc} \quad (U_S < 0,\ U_F > 0) \qquad (3.23a) \\ \dfrac{(e^{U_S} - U_S - 1)/F(U_S, U_F)}{\displaystyle\int_0^{U_S} \frac{e^{U_F}(1 - e^{-U})(e^U - U - 1)}{2F^3(U, U_F)}\, dU} & \begin{matrix}\text{depl, inv} \\ (U_S > 0,\ U_F > 0)\end{matrix} \qquad (3.23b) \end{cases}$$

Unlike the delta-depletion result, C cannot be expressed explicitly as a function of V_G' in the exact charge formulation. Both variables, however, have been related to U_S and it is possible to compute numerically the capacitance expected from the structure for a given applied gate voltage using Eqs. (3.20) through (3.23) in conjunction with Eq. (3.18). The usual and most efficient computational procedure is to calculate C and the corresponding V_G' for a set of assumed U_S values. Typically, a sufficient set of (C, V_G') points to construct the C–V_G' characteristic will be generated if U_S is stepped by whole-number units $(-5, -4, \cdots)$ over the normal operating range of U_S values ($U_F - 21 \lesssim U_S \lesssim U_F + 21$ at room temperature). It should be noted that care must be exercised if $U_S = 0$ is included as one of the computational points. At $U_S = 0$ the Eq. (3.21) expression for W_{eff} becomes indeterminate (0/0) and, as is readily established, must be replaced by $W_{eff} = \sqrt{2}\, L_D/[\exp(U_F) + \exp(-U_F)]^{1/2}$. Also, the quoted high frequency results hold only for p-type devices. It is nevertheless possible to obtain an n-type characteristic by simply running the calculations for an equivalently doped p-type device and then changing the sign of all computed V_G' values. This procedure works because of the voltage symmetry between ideal n- and p-type devices.

EXACT-CHARGE C-V THEORY SUPPLEMENT

BASIC RELATIONSHIPS

$$\underset{\text{TOTAL}}{v_G'} = \underset{\text{D.C.}}{V_G'} + \underset{\text{A.C.}}{v_g'} \qquad \underset{\text{TOTAL}}{q_G} = \underset{\text{D.C.}}{Q_G} + \underset{\text{A.C.}}{q_g} \qquad \underset{\text{TOTAL}}{\mathcal{E}_S} = \underset{\text{D.C.}}{\mathcal{E}_{S\phi}} + \underset{\text{A.C.}}{\mathcal{E}_s}$$

$$v_G' = \frac{kT}{q} u_S + \chi_0' \mathcal{E}_S \qquad V_G' = \frac{kT}{q} u_S + \chi_0' \mathcal{E}_{S\phi} \qquad v_g' = \frac{kT}{q} u_s + \chi_0' \mathcal{E}_s$$

$$q_G = -q_S = K_S \epsilon_0 \mathcal{E}_S \qquad Q_G = -Q_S = K_S \epsilon_0 \mathcal{E}_{S\phi} \qquad q_g = -q_s = K_S \epsilon_0 \mathcal{E}_s$$

$$\underset{\text{TOTAL}}{\mathcal{E}_S} = \hat{u}_S \frac{kT}{q} \frac{F(u_S, u_N, u_P)}{L_D} \qquad \underset{\text{D.C.}}{\mathcal{E}_{S\phi}} = \hat{U}_S \frac{kT}{q} \frac{F(U_S, U_F)}{L_D} \qquad \underset{\text{A.C.}}{\mathcal{E}_s}$$

$$F(u_S, u_N, u_P) = \left[e^{u_P}(e^{-u_S} + u_S - 1) + e^{-u_N}(e^{u_S} - u_S - 1) \right]^{\frac{1}{2}}$$

$$F(U_S, U_F) = \left[e^{U_F}(e^{-U_S} + U_S - 1) + e^{-U_F}(e^{U_S} - U_S - 1) \right]^{\frac{1}{2}}$$

$$C = \frac{q_g}{v_g'} = \frac{-q_s}{\frac{kT}{q} u_s + \chi_0' \mathcal{E}_s} = \frac{-q_s}{\frac{kT}{q} u_s - q_s / C_0} = \frac{C_0 \left(\frac{-q_S}{kT/q\, u_s} \right)}{C_0 + \left(\frac{-q_S}{\frac{kT}{q} u_s} \right)}$$

or

$$C = \frac{C_0 C_s}{C_0 + C_s} \qquad \text{where } C_s \equiv -\frac{q_s}{\frac{kT}{q} u_s} = \text{SEMICONDUCTOR CAPACITANCE}$$

NOTE: all q's and C's are per unit area of the gate.

General C_s Relationship

We note

$$-q_s = K_s \epsilon_o \mathcal{E}_s = K_s \epsilon_o (\mathcal{E}_s - \mathcal{E}_{s\phi}) = \hat{U}_s \frac{kT}{q} \frac{K_s \epsilon_o}{L_D} \left[F(U_s, U_N, U_p) - F(U_s, U_F) \right]$$

where

$$u_s = U_s + u_s \; ; \; u_N = U_F + u_m \; ; \; u_p = U_F + u_p$$

Expanding $F(u_s, u_N, u_p)$ in a Taylor series expansion about the d.c. operating point, and only keeping first order u_s, u_m, and u_p terms in the expansion, one obtains

$$F(u_s, u_N, u_p) = F(U_s, U_F) + \left.\frac{\partial F}{\partial u_s}\right|_{U_s, U_F} u_s + \left.\frac{\partial F}{\partial u_N}\right|_{U_s, U_F} u_m + \left.\frac{\partial F}{\partial u_p}\right|_{U_s, U_F} u_p$$

$$= F(U_s, U_F) + \frac{e^{U_F}(1 - e^{-U_s}) + e^{-U_F}(e^{U_s} - 1)}{2F(U_s, U_F)} u_s$$

$$- \frac{e^{-U_F}(e^{U_s} - U_s - 1)}{2F(U_s, U_F)} u_m + \frac{e^{U_F}(e^{-U_s} + U_s - 1)}{2F(U_s, U_F)} u_p$$

Substituting the expansion into the $-q_s$ relationship, and then substituting the new $-q_s$ relationship into the C_s expression, we conclude

$$C_s = \hat{U}_s \frac{K_s \epsilon_o}{L_D} \left[\frac{e^{U_F}(1 - e^{-U_s}) + e^{-U_F}(e^{U_s} - 1)}{2F(U_s, U_F)} - \frac{e^{-U_F}(e^{U_s} - U_s - 1)}{2F(U_s, U_F)} \frac{u_m}{u_s} + \frac{e^{U_F}(e^{-U_s} + U_s - 1)}{2F(U_s, U_F)} \frac{u_p}{u_s} \right]$$

Low-Frequency Capacitance

In the low-frequency limit, the MOS-C follows the a.c. gate signal quasistatically, meaning $u_m = 0$ and $u_p = 0$. Introducing the symbols $C_{s\downarrow}$ and $C_\downarrow$ to be the low-frequency semiconductor capacitance and MOS-C capacitance (both per unit area), we conclude

$$C_{s\downarrow} = \hat{U}_s \frac{K_s \varepsilon_0}{L_D} \frac{e^{U_F}(1-e^{-U_s}) + e^{-U_F}(e^{U_s}-1)}{2F(U_s, U_F)}$$

and

$$C_\downarrow = \frac{C_0 C_{s\downarrow}}{C_0 + C_{s\downarrow}} = \frac{C_0}{1 + \frac{C_0}{C_{s\downarrow}}} = \frac{C_0}{1 + \frac{\hat{U}_s L_D}{K_0'}\left[\frac{2F(U_s, U_F)}{e^{U_F}(1-e^{-U_s}) + e^{-U_F}(e^{U_s}-1)}\right]}$$

same as Eqs. (3.20)/(3.21)/(3.22) reseelt on p. T6-2.

High-Frequency Capacitance

In the high-frequency limit, the majority carrier can still follow the applied a.c. signal. Thus, $u_m = 0$ if the semiconductor is n-type and $u_p = 0$ if the semiconductor is p-type. Moreover, if the MOS-C is accumulation or weakly depletion biased, only the majority carriers affect the observed capacitance; i.e., the minority carrier quasi-Fermi level term in the general C_s expression can be ignored. Therefore, under accumulation and weak depletion biasing $C_{s\uparrow} = C_{s\downarrow}$ and $C_\uparrow = C_\downarrow$, where $C_{s\uparrow}$ and $C_\uparrow$ are the high-frequency semiconductor and MOS-C capacitance per unit area, respectively. On the other hand, for depletion biases where $|U_s| > |U_F|$ and for inversion biases, the minority carrier term in the general C_s expression cannot be neglected. To complete the analysis, we are therefore required to obtain explicit expressions for u_m/u_s and u_p/u_s.

The relationship between u_n or u_p and u_s is obtained from the requirement that the total number of minority carriers in an inversion layer at a given d.c. bias does not change during a high-frequency capacitance measurement. Mathematically, to bring the semiconductor to be p-type, we require

$$\int_0^{\infty} (n - n_b)\,dx \Big|_{D.C.+A.C.} = \int_0^{\infty} (n - n_b)\,dx \Big|_{D.C.} \qquad \ldots n_b \text{ is the electron concentration in the semiconductor bulk}$$

Noting $n = n_i e^{\frac{u - u_N}{U - U_F}} \cdots$ D.C. + A.C.

$n = n_i e^{U - U_F} \cdots$ D.C.

changing variables from x to u or U, and cancelling the same quantities appearing in both integrals, we obtain

$$\int_0^{u_S} \frac{e^{u - u_N} - e^{-u_N}}{F(u, u_N, U_F)}\,du = \int_0^{U_S} \frac{e^{U - U_F} - e^{-U_F}}{F(U, U_F)}\,dU \qquad \ldots u_F = U_F \text{ for a p-type semi}$$

Setting $u_S = U_S + u_s$ and $u_N = U_F + u_M$, expanding the left-hand integral in terms of the a.c. quantities u_s and u_M, and keeping only up to first order terms in u_s and u_M, one ultimately deduces

$$\frac{u_M}{u_s} = \frac{1}{\left(\frac{F(U_S, U_F)}{e^{U_S} - 1}\right) \int_0^{U_S} \left[\frac{(e^U - 1)}{F(U, U_F)}\left(1 - \frac{e^{-U_F}(e^U - U - 1)}{2F^2(U, U_F)}\right)\right] dU}$$

Finally, substituting the above u_M/u_s relationship into the general C_s expression, remembering $u_p = 0$ for a p-type semiconductor, and manipulating the result into as simple a form as possible, we conclude

$$C_f = \frac{C_0}{1 + \frac{\hat{Q}_s L_D}{x_0'}\left[\frac{2F(U_S,U_F)}{e^{U_F}(1-e^{-U_S}) + e^{-U_F}(e^{U_S}-1)/(1+\Delta)}\right]}$$

where

$$\Delta = \begin{cases} 0 \quad \ldots\ U_S < U_F \\ \dfrac{(e^{U_S} - U_S - 1)/F(U_S,U_F)}{\int_0^{U_S} \frac{e^{U_F}(1-e^{-U})(e^{U}-U-1)}{2F^3(U,U_F)}\,dU} \quad \ldots\ U_S > U_F \end{cases}$$

A Comparison of Quantum-Mechanical Capacitance–Voltage Simulators

C. A. Richter, A. R. Hefner, and E. M. Vogel

Abstract—**We have systematically compared the results of an extensive ensemble of the most advanced available quantum-mechanical capacitance–voltage (*C–V*) simulation and analysis packages for a range of metal-oxide-semiconductor device parameters. While all have similar trends accounting for polysilicon depletion and quantum-mechanical confinement, quantitatively, there is a difference of up to 20% in the calculated accumulation capacitance for devices with ultrathin gate dielectrics. This discrepancy leads to large inaccuracies in the values of dielectric thickness extracted from capacitance measurements and illustrates the importance of consistency during *C–V* analysis and the need to fully report how such analysis is done.**

Index Terms—**Capacitance, effective oxide thickness, gate dielectric, inversion quantization, MOS devices, polysilicon depletion.**

I. Introduction

IN RECENT years, the continuing decrease of the gate dielectric thickness in conventional silicon MOS devices has made it increasingly difficult to predict capacitance–voltage (*C–V*) curves accurately. Thus, predictive TCAD is problematic and likely to be incorrect. Researchers have developed sophisticated methods to simulate *C–V* curves [1]–[8], [10], [11] in an effort to overcome this difficulty. However, there has been little work comparing the outputs from such simulators to check their agreement. In an effort to determine the variability between such simulators, we have acquired an ensemble of five advanced quantum-mechanical (QM) simulation and analysis packages and compared their results for a matrix of *C–V* curves produced for a range of device parameters.

Currently, one of the most prevalent uses of these QM *C–V* simulators is to determine the equivalent oxide thickness (EOT) of MOS devices made from alternate (high-k) dielectrics proposed to replace SiO_2 as the gate dielectric in future MOS technologies. Finding a replacement gate dielectric is necessary in order for traditional CMOS scaling to continue, and EOT is a primary criteria used to assess which materials are technologically most promising. Thus, it is critical that the interrelationship of QM CV simulators be well characterized to properly compare EOT extracted from experimental *C–V* curves.

Traditionally, the thickness, d, of the dielectric in an MOS capacitor was easily found by $d = \varepsilon A/C_{ox}$, where ε is the dielectric's permittivity, A is the device area, and C_{ox} is the device capacitance (in accumulation). However, this simple relationship does not hold for thin oxides, and extracting a accurate physical thickness from *C–V* curves is increasingly difficult for oxide thicknesses at and below 2 nm. Two primary effects, illustrated in Fig. 1, must be considered: quantum mechanical confinement, and the finite voltage drop across polysilicon gates (poly-Si depletion). To avoid the confusion often associated with *C–V*-derived film thickness, the following terminology will be used. EOT is the equivalent thickness of SiO_2 that would produce the same *C–V* curve as that obtained from the alternate dielectric system. The capacitive effective thickness (CET(V)) is simply the thickness that is derived directly from the relationship $CET(V) = (\varepsilon_0 \varepsilon_{SiO_2} A)/C(V)$, where ε_0 is the permittivity of free space, ε_{SiO_2} is the relative permittivity of SiO_2, and C(V) is the capacitance at bias voltage V.

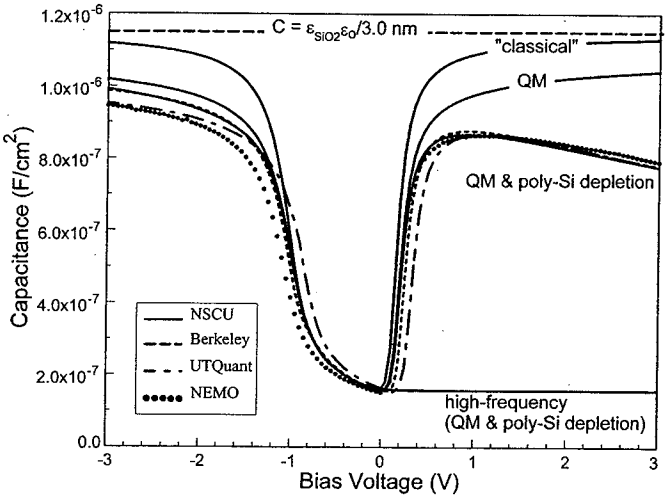

Fig. 1. Simulated *C–V* curves accounting for both QM confinement and poly-Si depletion. Simulated parameters are $d_{ox} = 3.0$ nm (2.987 nm for NEMO [10]), $N_d = 3 \times 10^{17}$ cm^{-3}, and $N_{poly} = 5 \times 10^{19}$ cm^{-3}. A classical *C–V* with no QM confinement or poly-Si depletion and a QM *C–V* which accounts for QM confinement only are also shown for illustrative purposes.

II. Simulators

An ensemble of five of the most advanced, one-dimensional (1-D) , quantum-mechanical *C–V* software packages was used in this comparison:

1) Quantum mechanical *C–V* simulator developed by the Device Group at UC Berkeley (Berkeley) [5];
2) Nanotechnology Engineering Modeling Program (NEMO) [2], [11];
3) CVC, a program developed by Hauser [1], [10] at NCSU (NCSU);

Manuscript received June 13, 2000. This work was funded in part by the NIST Office of Microelectronics Programs. The review of this letter was arranged by Editor K. De Meyer.

The authors are with Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8121 USA (e-mail: Curt.Richter@NIST.Gov).

Publisher Item Identifier S 0741-3106(01)00686-3.

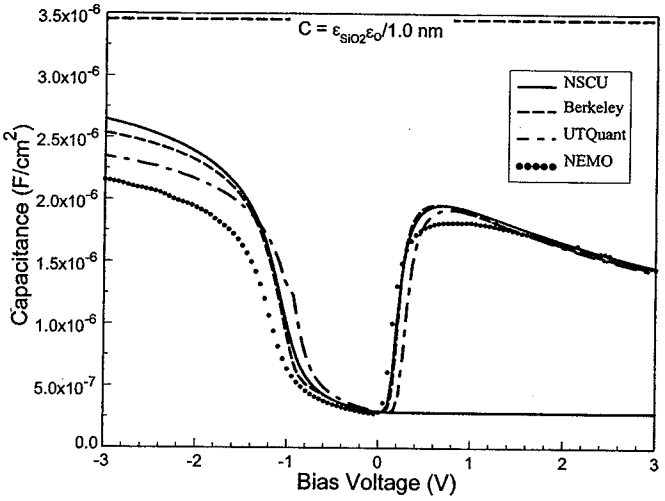

Fig. 2. Simulated *C–V* curves accounting for both QM confinement and poly-Si depletion. Simulated parameters are $d_{ox} = 1.0$ nm (1.086 nm for NEMO), $N_d = 1 \times 10^{18}$ cm^{-3}, and $N_{poly} = 1 \times 10^{20}$ cm^{-3}.

4) UTQuant developed at UT-Austin [4];
5) IBM's Tqm_v6 [3].

Three programs are purely simulation packages; Berkeley, NEMO, and UTQuant. NCSU is both a simulation and analysis package, while IBM's Tqm_v6 is only a *C–V* analysis program. The simulators all assume a source of minority carriers in the Si-substrate, and therefore produce ideal quasistatic *C–V* curves, or curves obtained in transistor measurements.

In simulating *C–V* curves, Berkeley calculates the electron/hole distributions in both inversion and accumulation derived by solving the Schrödinger and Poisson equations self-consistently with the Fermi–Dirac distribution [5]. NEMO, the most physically comprehensive of the ensemble, is a nonequilibrium Green's function simulator [2], [11]. In it, fundamental physics—such as multiband scattering, inelastic scattering, and interface roughness—can be simulated. NCSU is based upon a model containing first order physics approximations that can be rapidly calculated [1], [10]. UTQuant is another self-consistent, QM Poisson solver [4]. IBM's Tqm_v6 is a fast QM *C–V* analysis program that is based upon polynomial interpolation from the results of IBM's extensive QM simulations [3].

An n-channel MOS capacitor was chosen as the test structure to compare the various simulators. The MOS capacitor consists of a p-type silicon substrate, ideal SiO_2 gate dielectric, and n-type poly-Si as the gate material. This idealized structure was chosen in order to best compare the results of the simulators themselves. The parameters of the capacitor were varied to create a matrix of *C–V* curves: oxide thickness, $d_{ox} = 1.0$ nm, 2.0 nm, 3.0 nm, and 10.0 nm; silicon substrate doping, $N_d = 1 \times 10^{15}$ cm^{-3}, 1×10^{17} cm^{-3}, 3×10^{17} cm^{-3}, and 1×10^{18} cm^{-3}; and poly-Si doping, $N_{poly} = 1 \times 10^{19}$ cm^{-3}, 5×10^{19} cm^{-3}, and 1×10^{20} cm^{-3}. An effort was made to insure that comparable values were used for other parameters in the simulations such as work functions because their default values sometimes varied between simulators. Our simulations were done from −5 V to 5 V in nominally 50 mV steps [1]. It should be noted that this voltage range can lead to artificially large electric fields in the thinnest devices simulated.

TABLE I
PARAMETERS EXTRACTED BY USING Tqm_v6. SIMULATED PARAMETERS ARE: TOP, $d_{ox} = 1.0$ nm, $N_d = 1 \times 10^{18}$cm^{-3}, and $N_{poly} = 1 \times 10^{20}$cm^{-3}; BOTTOM, $d_{ox} = 3.0$ nm, $N_d = 3 \times 10^{17}$cm^{-3}, and $N_{poly} = 5 \times 10^{19}$cm^{-3}. CET IS DETERMINED AT −2.5 V

Simulator	CET (nm)	EOT (nm)	V_{fb} (V)	N_d (cm^{-3})
Berkeley	1.408	0.878	-1.141	1.12E18
NCSU	1.357	0.833	-1.133	1.12E18
NEMO (1.086 nm)	1.663	1.108	-1.281	1.12E18
UTQuant	1.520	0.979	-0.948	1.20E18

Simulator	CET (nm)	EOT (nm)	V_{fb} (V)	N_d (cm^{-3})
Berkeley	3.562	2.849	-1.087	3.44E17
NCSU	3.552	2.840	-1.076	3.44E17
NEMO (2.987 nm)	3.737	3.012	-1.153	3.19E17
UTQuant	3.695	2.973	-0.920	3.44E17

III. RESULTS AND DISCUSSION

Figs. 1 and 2 show typical *C–V* simulations for $d_{ox} = 3.0$ nm and for $d_{ox} = 1.0$ nm, respectively (with the exception of the thickness used in NEMO, which is nominally the same [2]). These figures illustrate that the simulators are in qualitative agreement; i.e., the overall shape of the *C–V* curves is the same for these four different simulators. More specifically, the capacitance value at the minimum is the same for the set. This agreement indicates that the simulators agree on how substrate doping affects the shape of the *C–V* curves. The flatband and threshold values of the *C–V* curves are also in good agreement with the exception that the default parameters used for UTQuant lead to a slight shift with respect to the others. This agreement between the simulators, where some are based on very different fundamentals (such as NSCU and NEMO) and others that have nominally the same basis (i.e., UTQuant and Berkeley), instills confidence in the overall capabilities of all these simulators.

However, there are systematic trends and important differences between the various simulators. The largest disparity is in the values of the accumulation region capacitance. The NCSU and Berkeley simulators tend to have the largest accumulation capacitance for a given set of parameters, while NEMO and UTQuant have a lower accumulation capacitance. This disagreement leads to *C–V* curves that appear to be from devices with different oxide thicknesses. On the other hand, there is surprisingly little difference in the simulated *C–V* curves in the inversion region. It should be noted that there is slightly greater variability (not shown) in the inversion capacitance when a metal gate is simulated (i.e., no poly-Si depletion).

[1]The Berkeley simulator simulates data in steps that are uniform in Si-surface potential, and are therefore not uniform in gate bias. A comparable number of data points (200) were simulated

[2]In NEMO, a minimum mesh of the silicon lattice constant (0.271 547 nm) is most physically realistic and gives consistent results. Therefore, the thicknesses simulated in NEMO are near, but not exactly, the same as the nominal thickness values

In order to quantitatively determine the relationships among the various simulators, we have used Tqm_v6, IBM's QM *C–V* analysis program, to extract a reduced set of parameters (or assessment criteria) for each simulated *C–V* curve. These extracted values are then used to compare the results for the simulators. This analysis also compares the simulators with the calculations utilized by Tqm_v6. Table I shows the extracted EOT, CET, and N_d for two different sets of parameters; one set at $d_{ox} = 3.0$ nm and the other set at $d_{ox} = 1.0$ nm. There is little variation in the extracted substrate doping values, as expected. In addition, the extracted flatband values are also in moderate agreement, as expected from the results shown in Figs. 1 and 2. The simulators are not in such strong agreement for EOT. There is a maximum difference [3] of 0.185 nm between the simulations for the $d_{ox} = 3.0$ nm parameter set, and a maximum difference of 0.189 nm for $d_{ox} = 1.0$ nm. This illustrates that the offset between simulators is not scaling with the thickness of the SiO_2, and thus becomes a larger problem for thinner gate dielectrics. The apparent thickness variability is a significant proportion of the total film thickness (up to 20%). This observed thickness variability is larger than the EOT thickness control value ($<\pm 0.08$ nm) required for the year 2001 by the 1999 ITRS [9]. We are currently investigating the possible physical assumptions leading to differences between the various simulators such as: the use of approximations for quantum effects versus a full solution of the Schrödinger equation, wave function boundary conditions at the Si/SiO_2 interface, type of carrier statistics, and models for handling highly doped poly-Si. In addition to a comprehensive understanding of the physical assumptions, a detailed knowledge of the computational methods used to solve the underlying physical equations is necessary to fully understand why this suite of programs is not in exact agreement in accumulation.

Although a significant discrepancy has been identified, it is difficult experimentally to determine which simulator is most "accurate." Typically, physical thickness measurements have minimum uncertainties on the order of 0.2 nm with regards to accuracy. In addition, the active doping concentration and profiles in the poly-Si gate affect the derived thickness and therefore also must be determined accurately. Thus, while there are systematic differences among the simulators they are in agreement to levels that are experimentally verifiable. Choosing the "most correct" one remains extremely challenging.

IV. Conclusions

While this extensive comparison of *C–V* simulator/analysis packages increases the confidence in the individual packages in this ensemble, important systematic differences in the resulting *C–V* curves have been observed. The most noticeable of these variations occurs in the accumulation capacitance region and leads to variations in extracted EOT on the order of 0.2 nm for total SiO_2 film thicknesses in the range 1.0 nm to 3.0 nm. This thickness variation is a significant proportion of the total film thickness—up to 20% of the total film thickness. The demonstrated discrepancy illustrates that when reporting experimentally derived electrical thickness results, it is important to describe fully how these values were obtained. The same experimental curve can lead to different extracted EOT values depending upon which QM software is used for the analysis. Therefore, it is essential that EOT results be presented consistently and with sufficient detail to allow the technical community to reliably compare *C–V* results.

[3]The maximum difference is determined from comparing $d_{ox}^{\text{simulated}}$-EOT(Tqm_v6) for the various simulators.

Acknowledgment

The authors would like to acknowledge R. Lake (Raytheon), J. Hauser (NCSU), D. Buchanan, E. Gusev (IBM), A. Tasch (UT Austin), the Device Group at UC Berkeley, S. Satterfeld, D. Blackburn, and J. Ehrstein (NIST) for their help during this comparison.

References

[1] J. R. Hauser and K. Ahmed, "Characterization of ultra-thin oxides using electrical C-V and I–V measurements," in *Characterization and Metrology for ULSI Technology*, Seiler *et al.*, Eds. Woodbury, NY: AIP, 1998, pp. 235–239.

[2] R. Lake *et al.*, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, pp. 7845–7869, 1997.

[3] S.-H. Lo, D. A. Buchanan, and Y. Taur, "Modeling and characterization of quantization, polysilicon depletion and direct tunneling effects in MOSFET's with ultrathin oxides," *IBM J. Res. Develop.*, vol. 43, pp. 327–337, 1999.

[4] W.-K. Shih *et al.*, *UTQUANT 2.0 User's Guide*. Austin, TX: Univ. Texas Press, Oct. 1997.

[5] Berkeley Device Group.. [Online].. Available: www.device.eecs.berkeley.edu/qmcv/html

[6] K. S. Krisch, J. D. Bude, and L. Manchanda, "Gate capacitance attenuation in MOS devices with thin gate dielectrics," *IEEE Electron Device Lett.*, vol. 17, pp. 521–524, 1996.

[7] C. Bowen *et al.*, "Physical oxide thickness extraction using quantum mechanical simulation," in *IEDM Tech. Dig.*, 1997, pp. 869–872.

[8] S. V. Walstra and C.-T. Sah, "Thin oxide thickness extrapolation from capacitance-voltage measurements," *IEEE Trans. Electron Devices*, vol. 44, pp. 1136–1142, 1997.

[9] Semiconductor Industry Association, *International Technology Roadmap for Semiconductors: 1999*. Austin, TX: SEMATECH, 1999.

[10] W. K. Henson *et al.*, "Estimating oxide thickness of tunnel oxides down to 1.4 nm using conventional capacitance-voltage measurements on MOS capacitors," *IEEE Electron Device Lett.*, vol. 20, pp. 179–181, 1999.

[11] D. K. Blanks *et al.*, "NEMO: General release of a new comprehensive quantum device simulator," in *Proc. IEEE 24th Int. Symp. Compound Semiconductors*, 1998, pp. 639–642.