Proc. Int. Semiconductor Device Research Symposium (ISDRS'99), pp.397-400, Charlottesville, VA, December 1-3, 1999

EXTENDED CHARGES MODELING FOR DEEP SUBMICRON CMOS

Matthias Bucher[1], Jean-Michel Sallese[1], Christophe Lallement[2], Wladyslaw Grabinski[1], Christian C. Enz[1,3], François Krummenacher[1]

[1)] *Swiss Federal Institute of Technology (EPFL), Electronics Laboratory, CH-1015 Lausanne, Switzerland*
[2)] *ERM-PHASE / Ecole Nationale Supérieure de Physique de Strasbourg, Blvd. S. Brant, F-67400 Illkirch, France*
[3)] *Swiss Center for Electronics and Microtechniques (CSEM), CH-2007 Neuchâtel, Switzerland*

Abstract - **The simulation of deep submicron CMOS circuits operating at high-frequency requires adequate models representing the dynamic behavior of the transistors. Charges in the device are affected by carrier quantization and polydepletion in the gate. Velocity saturation is one of the short-channel effects that further affect the charges. An analytical MOS transistor (MOST) model for circuit simulation which includes the above effects is described. The charges/transcapacitances expressions show good qualitative behavior at all inversion levels from weak to moderate and strong inversion and have correct asymptotic behavior. The model agrees well with characteristics obtained from 2D device simulation and measured on deep submicron CMOS technology.**

1. Introduction

Modeling of charges and transcapacitances for deep submicron CMOS technologies is an increasingly challenging task. The use of thin gate oxides, high levels of channel dopant concentration, and polysilicon gates leads to increased influence of quantum effects (QM) (e.g. [1][2]) and polysilicon gate depletion (PD) (e.g. [3]). All device characteristics are affected, in particular the charges/transcapacitances. Short-channel devices are further affected by the two-dimensional nature of fields, leading to velocity saturation among other short-channel effects. Currently available analytical MOST models have some difficulty to represent such effects in a qualitatively correct way. An alternative approach using numerical MOST models may offer increased accuracy but has to deal with the difficulty of numerical iteration.

The modeling approach proposed here is based on an accurate evaluation of charges present in the MOS structure at all inversion levels, following the 'EKV' approach [4]-[6] extended to a charge-based model [7]-[11]. The charges model is extended to account for velocity saturation and channel length modulation. First-order models are proposed to account for QM and PD effects. The compact MOST model obtained satisfies the most important criteria for MOST models, among which are correct asymptotic behavior, smooth conductances and transcapacitances in all operating regions. The model further uses a small number of parameters closely related to the underlying physics.

The increasingly important influence of parasitic elements such as fringing and overlap capacitances and junction space charge regions (e.g. [13]) will be addressed elsewhere.

2. Charge-based modeling

The basic charge-sheet model for the drain current and the charges in the MOST will be established as simple functions of physical parameters and terminal voltages. The drain current of the MOST can be expressed according to [4]-[6],

$$I_D = I_S \cdot (i_f - i_r) \qquad I_S \cong 2nU_T^2\mu C'_{ox}\frac{W}{L}$$
$$i_{f(r)} = F\left(\frac{V_P - V_{S(D)}}{U_T}\right) \tag{1}$$
$$V_P \cong \frac{V_G - V_{TO}}{n} \qquad n \cong 1 + \frac{\gamma}{2\sqrt{\Psi_0 + V_P}}$$

where I_S is the *specific current*, used as a normalization factor for the drain current, depending on the *slope factor* n, the mobility μ, and the effective device dimensions W and L. The slope factor n is in turn expressed as a function of the *pinch-off voltage* V_P depending on the gate voltage V_G. The *forward (reverse) normalized currents* $i_{f(r)}$ are the fundamental variables in this modeling approach [4]. They are symmetric in terms of source(drain) voltages $V_P - V_{S(D)}$. Note that all voltages are referred to the substrate; U_T is used to normalize voltages. The function F links the normalized variables i and υ; its asymptotes are $F(\upsilon) = (\upsilon/2)^2$ in strong inversion and $F(\upsilon) = \exp(\upsilon)$ in weak inversion [4]. Parameters in (1) are defined as,

$$V_{TO} = V_{FB} + \Psi_0 + \gamma\sqrt{\Psi_0} \qquad \Psi_0 \cong 2U_T\ln\left(\frac{N_{sub}}{n_i}\right)$$
$$\gamma = \frac{\sqrt{2q\varepsilon_{si}N_{sub}}}{C'_{ox}} \qquad C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \qquad U_T = \frac{kT}{q} \tag{2}$$

where undefined symbols have conventional meaning.

The current transport equation $I = \mu W \cdot [-Q_i' \cdot d\Psi_s/dx + U_T \cdot dQ_i'/dx]$ includes drift and diffusion in terms of inversion charge density Q_i' and surface potential Ψ_s at each coordinate x along the channel. Introducing the normalized channel current $i = I/I_S$, the channel potential $\upsilon = V/U_T$, and linearizing the inversion charge $dQ_i' \approx n \cdot C'_{ox} \cdot d\Psi_s$ [7], Q_i' can be expressed as a function of i [8][10],

$$Q_i' = -2nU_TC'_{ox}\left(\sqrt{\frac{1}{4} + i} - \frac{1}{2}\right) \tag{3}$$

Using the relationship between channel conductance and inversion charge, $dI/dV = \mu Q_i' W/L$ [4], a differential equation is obtained, linking the normalized quantities i and υ,

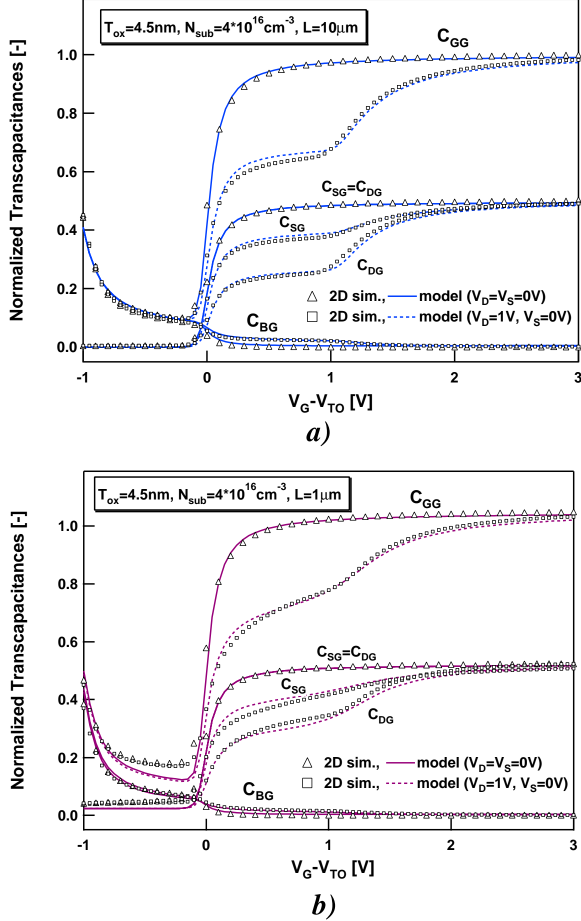

Fig. 1: *Transcapacitances, normalized to* WLC'_{ox} *, versus* $V_G - V_{TO}$*, at low and high drain voltage, for an n-channel device, a) long-channel, b) intermediate channel length. 2D device simulation (markers); model (dashed, lines).*

$$\frac{di}{d\upsilon} = \frac{1}{2} - \sqrt{\frac{1}{4} + i} = i \cdot G(i)\,. \tag{4}$$

In (4), $G(i)$ is the normalized transconductance-to-current ratio [4][10], corresponding to $G(i_f) = g_{ms} U_T / I_D$ in saturation [4], where $g_{ms} \equiv -\partial I_D / \partial V_S$. The function $G(i_f)$ shows an excellent agreement with a numerical solution of the Poisson and Gauss equations as well as with measurements at all inversion levels over several technologies [10][11]. Integrating (4) and determining the integration constant from strong inversion conditions yields the relationship $\upsilon_P - \upsilon = F^{-1}(i)$ [8][10],

$$\upsilon_P - \upsilon = 2\left(\sqrt{\frac{1}{4} + i} - \frac{1}{2}\right) + \ln\left(\sqrt{\frac{1}{4} + i} - \frac{1}{2}\right) \tag{5}$$

An approximate analytical expression can be used to invert (5) without deteriorating the accuracy of the model [9].

The drain current (1) can now be expressed in terms of the terminal voltages, at all inversion levels from weak to strong inversion and from non-saturation to saturation. It is further adapted to include field-dependent mobility μ [10]. A velocity saturated region forms near the drain when carriers reach their saturated velocity [12]. The channel is divided in two parts, a first one of length $L - \Delta L$, where the gradual channel approximation (GCA) used so far does hold, and the velocity saturated part of length ΔL. At the transition point, an effective drain voltage V_{Dsat} is reached, replacing V_D in (1) resulting in the corresponding reverse current i_r' [9].

The total charges in the device can be obtained by integrating the charge densities along the channel, leading to the charges model [7]-[11]. The effect of velocity saturation on the charges will also be considered here by integrating the charges according to the two regions, assuming a constant inversion charge density in the velocity-saturated part (e.g. [14]). The inversion charge is integrated, using $dx = -(L - \Delta L) \cdot di / (i_f - i_r')$,

$$Q_I = \frac{1}{LC'_{ox}} \cdot \left[\int_0^{L-\Delta L} Q_i' dx + Q_i'|_{(x = L - \Delta L)} \cdot \int_{L-\Delta L}^{L} dx \right]$$

$$= -2nU_T \cdot \left[\frac{L - \Delta L}{L} \cdot \left(\frac{2}{3} \cdot \frac{\chi_f^2 + \chi_f \chi_r' + \chi_r'^2}{\chi_f + \chi_r'} - \frac{1}{2} \right) + \frac{\Delta L}{L} \cdot \left(\chi_r' - \frac{1}{2} \right) \right] \tag{6}$$

where $\chi_{f(r')} = \sqrt{1/4 + i_{f(r')}}$ is an auxiliary variable.

To obtain the drain and source charges, the inversion charge is partitioned linearly [15] among source and drain when integrating the inversion charge density;

$$Q_D = \frac{1}{LC'_{ox}} \cdot \left[\int_0^{L-\Delta L} \frac{x}{L} Q_i' dx + Q_i'|_{(x = L - \Delta L)} \cdot \int_{L-\Delta L}^{L} \frac{x}{L} dx \right]$$

$$= -nU_T \left[\frac{L - \Delta L}{L} \left(\frac{3\chi_r'^3 + 6\chi_r'^2 \chi_f + 4\chi_r' \chi_f^2 + 2\chi_f^3}{(\chi_f + \chi_r')^2} - \frac{1}{2} \right) + \frac{\Delta L}{L} \cdot \left(2 - \frac{\Delta L}{L} \right) \cdot \left(\chi_r' - \frac{1}{2} \right) \right] \tag{7}$$

$$Q_S = Q_I - Q_D \tag{8}$$

The depletion charge expression is obtained through its linearization around $Q_i' = 0$ [4],

$$Q_B \cong -\gamma \sqrt{\Psi_0 + V_P} - \frac{n-1}{n} Q_I\,, \tag{9}$$

and the gate charge balances all other charges,

$$Q_G = -Q_B - Q_I - Q_{ox}\,, \tag{10}$$

where Q_{ox} is a fixed oxide charge per unit area. Note that all charges have been normalized to WLC'_{ox}.

All node charges have now been obtained through integration along the channel. Their derivatives yield the (trans-)capacitances $C_{XY} \equiv \delta(\partial Q_X / \partial V_Y)$, where $\delta = 1$ if $X = Y$ and $\delta = -1$ otherwise. In Fig. 1 some of the model's transcapacitances are compared to 2D device simulation, both for a long-channel device and a device with intermediate channel length. Correct

asymptotic behavior and good quantitative agreement in all operating regions can be observed.

The model expressions have been used as stated above, combined with charge-sharing for short-channel [9]; a slightly corrected expression for the slope factor for improved asymptotic behavior of the transcapacitances has been used. A single parameter set is used for all simulated characteristics, and the model parameters match closely those underlying the 2D simulation. Note that the model at this point requires only the five long-channel parameters μ, C'_{ox}, V_{TO}, γ and Ψ_0 (or equivalently t_{ox}, V_{FB}, and N_{sub}), and three short-channel parameters related to velocity saturation/channel length modulation and charge-sharing, to calculate the drain current, charges and transcapacitances in all operating regions and from long-channel to short-channel.

The shorter-channel device shows velocity saturation effects in strong inversion; only bias-independent fringing capacitances have been accounted for in the model. In depletion, the influence of the bias-dependent overlap and space-charge capacitances (not included in the model) can already clearly be seen. These capacitances terms become increasingly important with shorter channel lengths and will be addressed elsewhere.

3. Quantum effects and polysilicon depletion

In the following, first-order expressions for QM and PD effects will be introduced into the basic charge-sheet model. A field-dependent expression for the band-gap widening due to quantization of electron energy levels is obtained in [1] and used e.g. in [2] in the context of a surface-potential based model,

$$\Delta E_g = \frac{13}{9}\cdot\kappa\cdot\left[\frac{\varepsilon_{si}}{4kT}\right]^{1/3}\cdot E_{\perp}^{2/3} \qquad (11)$$

where the effective vertical field is defined as $E_{\perp} = |Q'_B + \eta Q'_i|/\varepsilon_{si}$ [16]; and $\kappa = 4.1\cdot 10^{-10}\ eVm$. ΔE_g corresponds to the energy shift (with respect to the Si conduction band minima) of the lower 2D state of the quantum well near the SiO_2/Si interface. This effect is considered here as an energy shift of the conduction band minima with respect to the bulk Fermi level, and the MOST will still be treated within the charge-sheet approach. Under this assumption, ΔE_g leads to a modified intrinsic carrier concentration [1], $n_i^{qm} = n_i\cdot\exp(-\Delta E_g/2kT)$, which can be considered as a shift in the conduction band bending corresponding no longer to Ψ_0 but to $\Psi_0^{qm} = \Psi_0 + \Delta\Psi_0$, where $\Delta\Psi_0 = \Delta E_g/q$. In terms of threshold voltage change, the following expression is obtained,

$$\Delta V_{TO}^{qm} = \Delta\Psi_0 + \gamma\cdot\left(\sqrt{\Psi_0^{qm}} - \sqrt{\Psi_0}\right) \approx \sigma_{qm}\cdot\Delta\Psi_0\cdot n_0 \qquad (12)$$

where $n_0 = 1 + \gamma/(2\sqrt{\Psi_0})$ results from a first-order development of the term $\sqrt{\Psi_0^{qm}}$, and $0 \le \sigma_{qm} \le 1$ is a model parameter allowing to adjust the QM effect. The charge-sheet model will now use the corrected threshold voltage $V_{TO}^{qm} = V_{FB} + \Psi_0 + \gamma\sqrt{\Psi_0} + \Delta V_{TO}^{qm}$. The model is first calculated using the bias-independent equivalent of $\Delta\Psi_0$ evaluated at $E_{\perp}^0 = \gamma C'_{ox}\sqrt{\Psi_0}/\varepsilon_{si}$, leading to

$$\Delta V_{TO}^{qm,\,0} \approx \sigma_{qm}\cdot n_0\cdot\frac{13}{9}\cdot\kappa\cdot\left[\frac{\varepsilon_{si}}{4kT}\right]^{1/3}\cdot\left[E_{\perp}^0\right]^{2/3}. \qquad (13)$$

Following the first evaluation, the model is recalculated using the bias-dependent ΔV_{TO}^{qm}.

Successive approximations have been made until here, however justified considering that the slope factor n remains a slowly varying function of the gate voltage. As will be seen, the accuracy of the approximations is acceptable. Expression (13) may serve to evaluate the scaling trend in threshold voltage in terms of oxide thickness t_{ox} and substrate doping concentration N_{sub}.

The QM effect leads to a change in threshold voltage as discussed. Mobility is reduced in strong inver-

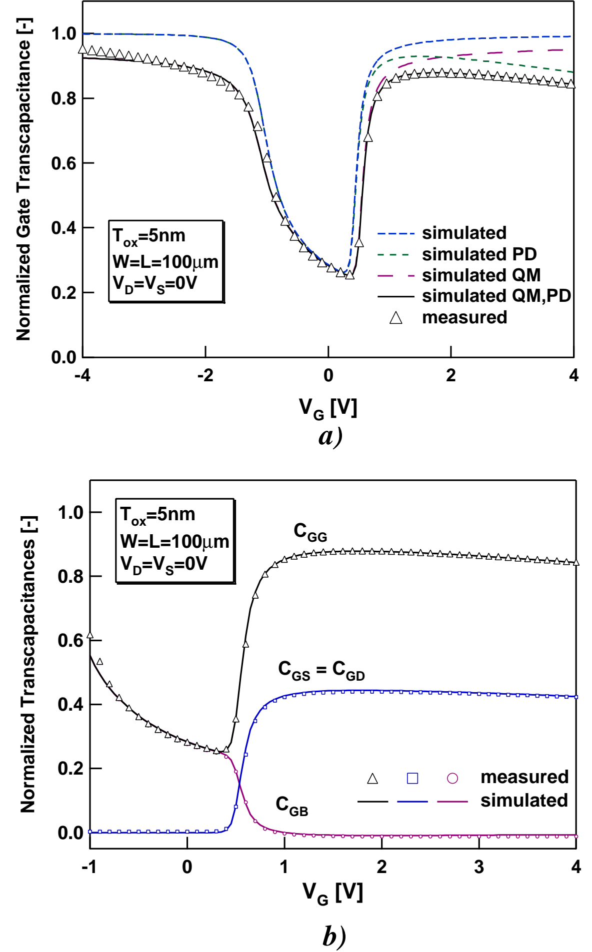

Fig. 2: *Transcapacitances normalized to WLC'_{ox}, measured (markers) from a large-area n-channel device of a $0.25\mu m$ CMOS technology, versus V_G at $V_D = V_S = 0V$. a) model corresponding to the basic charge model, without contribution of QM and PD effects individually (dashed) and cumulated (line). b) model (lines) with both QM and PD effects.*

sion, and the weak inversion slope of drain current is degraded. Further the total gate capacitance is severely reduced, both in inversion and accumulation.

In dual polysilicon gate technologies, the gates are implanted simultaneously with the source/drain regions (n-channel with n^+ poly, p-channel with p^+ poly). Insufficient active doping concentration in the poly gate leads to the formation of depletion layer causing a voltage drop in the gate, when the device is in inversion. Degradation of device characteristics becomes more severe when using thinner oxides.

The polydepletion effect will be modeled supposing uniform doping concentration in the gate and complete depletion [3]. The voltage drop in the gate can be readily expressed in terms of the gate charge,

$$\Delta V_G = \frac{q \cdot N_{poly}}{2\varepsilon_{si}} x_d^2 = \frac{q \cdot N_{poly}}{2\varepsilon_{si}} \left[\frac{C'_{ox} \cdot Q_G}{q \cdot N_{poly}}\right]^2 = \frac{Q_G^2}{\gamma_{poly}^2} \quad (14)$$

where x_d is the depletion layer width depending on the gate charge, and $\gamma_{poly} = \sqrt{2q\varepsilon_{si}N_{poly}}/C'_{ox}$ is the 'gate factor' depending on the polysilicon doping concentration N_{poly}. The effective gate voltage is then expressed as $V_{Geff} = V_G - \Delta V_G$. The gate charge Q_G (10) is first evaluated using the initial estimate $\Delta V_G{}^0 = (\gamma^2/\gamma_{poly}^2)\Psi_0$. The term $\Delta V_G{}^0$ corresponds to the threshold voltage shift due to the PD effect (showing that PD is essentially dependent on the ratio of substrate to poly doping concentrations). Then the full model accounting for the bias-dependency of ΔV_G is calculated. To account for the approximation made when evaluating the gate charge, γ_{poly} (or equivalently N_{poly}) is considered as a fitting parameter.

The total gate capacitance of an n-channel device of a standard $0.25\mu m$ CMOS technology from accumulation to strong inversion is shown in Fig. 2 a). The simulated curves correspond to the charge sheet model, including QM and PD effects individually, and the final result including both contributions. The value of the gate capacitance corresponds to the physical oxide thickness. In accumulation, the slight discrepancies observed stem from the use of a rather simple charge expression, which does not compromise the validity of the present approach. In Fig. 2 b), other transcapacitances for the same device show excellent agreement between measurement and simulation from depletion to strong inversion.

The model developed is consistently based on the evaluation of charges within the channel and the gate of the MOS structure. The simple models for both PD and QM effects are approximate, and are therefore of limited validity. Nevertheless they allow to model the bias-dependency of the effects quite well, and make use of the correct physical device parameters. The full model can be used at all inversion levels with present deep submicron CMOS technology.

4. Conclusions

Polydepletion and quantization effects have an increasingly strong impact on all device characteristics of deep submicron CMOS devices with thin gate oxides and high substrate doping. An analytical compact MOST model, based on a charge-sheet modeling approach, including velocity saturation, polydepletion and quantum effects, has been presented. The model is valid at all inversion levels from weak through moderate and to strong inversion, and shows correct asymptotic behavior and smooth and qualitatively correct transition regions. Good agreement with 2D device simulation as well as with measured characteristics from deep submicron CMOS technology has been shown. The model is efficient in terms of computation and uses a small set of physical parameters, making it adequate for deep submicron CMOS circuit simulation. Model extensions, in particular for fringing capacitances, are the object of further work.

Acknowledgment

Helpful discussions with Peter Bendix and Daniel Foty are gratefully acknowledged. This work was partially performed under CRAFT ESPRIT project (EP 25710) of the EU and supported by the Swiss Federal Office for Education and Sciences (OFES project 97.0384-1).

References

[1] M. Van Dort, P. Woerlee, A. Walker, *Solid State Electronics*, Vol. 37, N° 3, pp. 411-414, 1994.

[2] R. Rios, N. Arora, C.-L. Huang, N. Khalil, J. Faricelli, L. Gruber, *IEDM Tech. Digest,* 1995, pp. 937-940.

[3] C.-L. Huang, N. D. Arora, *IEEE Trans. ED*, Vol. 40, N° 12, pp. 2330-2337, December, 1993.

[4] C. C. Enz, F. Krummenacher, E. A. Vittoz, *Analog Int. Circ. Signal Proc. J.*, Vol. 8, pp. 83-114, 1995.

[5] H. Oguey, S. Cserveny, pp. 113-116, Bull. SEV/VSE, February 1982.

[6] M. Bucher, C. Lallement, C. Enz, F. Krummenacher, *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 703-706 , May, 1996.

[7] A. I. A. Cunha, M. C. Schneider, C. Galup-Montoro, *Solid State Electronics*, Vol. 38, N° 11, pp. 1945-1952, 1995.

[8] A. I. A. Cunha, O. C. Gouveia-Filho, M. C. Schneider, C. Galup-Montoro, *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 1608-1611, June 1997.

[9] M. Bucher, C. Lallement, C. Enz, F. Théodoloz, F. Krummenacher, Technical Report, Electronic Laboratories, EPFL, June 1997. <http://legwww.epfl.ch/ekv/>

[10] M. Bucher, C. Lallement, C. Enz, F. Théodoloz, F. Krummenacher, *Int. Semicond. Device Research Symp.*, pp. 615-618, December 1997.

[11] A. I. A. Cunha, M. C. Schneider, C. Galup-Montoro, *IEEE Journ. Solid-State Circuits*, Vol. 33, N° 10, pp. 1510-1519, 1998.

[12] P. K. Ko, Approaches to Scaling, Advanced MOS Device Physics, Ed. N. G. Einspruch, G. Gildenblat, pp. 1-37, Academic Press, 1989.

[13] P. Klein, *IEEE Trans. ED*, Vol. 44, N° 9, September, 1997.

[14] B. Iñiguez, E. G. Moreno, *Analog Int. Circ. Signal Proc. J.*, Vol. 13, N° 3, pp. 241-259, 1997.

[15] D. E. Ward, Technical Report G201-11, Integrated Circuits Laboratory, Stanford University, June 1981.

[16] S. C. Sun, J. C. Plummer, *IEEE Trans. ED*, Vol. 27, N° 7, pp. 1497-1508, 1980.