An Improved MOS Transistor Model with an Integrated Mobility Model

J. R. Hauser
N. C. State University, Raleigh, NC 27695, hauser@ncsu.edu

ABSTRACT

An improved compact MOS device model is derived based upon a first-order model for the dependency of MOS surface mobility on surface field and lateral drain field. The mobility model is incorporated into the basic device equations before integrating the equation along the device channel. A comparison with experimental data shows that a consistent set of physical parameters can be used to describe both long channel nMOS devices and short channel devices. The model can form the basis for improved MOS models for circuit analysis.

keywords: MOS device model, mobility model, compact model, MOS transistor

1 INTRODUCTION

Accurate analytical models of MOS device current are very important for circuit analysis and for projecting the performance of CMOS to future nanoscale device dimensions. Our understanding of MOS device physics has greatly improved over the past several decades. However, not all this improved understanding has been incorporated into MOS device models. This is especially true with respect to surface mobility in the inversion channel of an MOS device. As device dimensions have decreased over the years, the surface electric field in a typically MOS device has increased and at the same time the doping density in the channel region has increased in order to maintain a proper device threshold voltage. Both of these effects have led to a decreasing carrier mobility with a strong dependency on surface electric field.

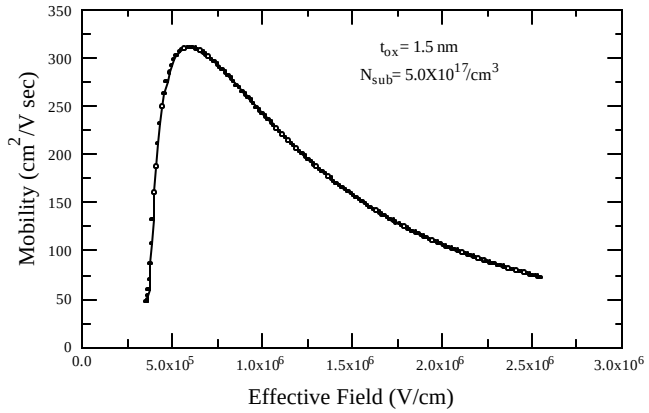

Figure 1. Typical variation of channel mobility with effective surface electric field for a 1.5 nm oxide.

Figure 1 shows a typical variation of effective surface mobility at low lateral fields with effective surface field for a nominal 1.5 nm oxide thickness. Of importance here is the falloff of the mobility with surface field. Our present understanding of device physics is that this falloff is due to the dominance of surface roughness scattering. The falloff at low fields is in turn due to impurity ion scattering either in surface interface states or from the bulk doping density. This curve is for electrons and the peak mobility is considerably reduced from that which can be achieved in bulk silicon at low doping densities. The surface scatterings are particularly important near the source end of the MOS channel as it is at such a point that the surface electric field has its maximum value. As one moves along the channel of an MOS device toward the drain end, the surface field decreases and the mobility increases from considering only surface scatterings. However, near the drain end of the channel, for reasonable drain voltages, the lateral electric field easily exceeds values appropriate for rapid optical phonon scatterings and the velocity tends to saturate at some maximum value. One thus has a picture where the MOS surface mobility is low at both ends of the surface channel and attains its largest value somewhere along the channel.

Including a position-dependent carrier mobility into the basic MOS device equation has in the past been considered intractable in an analytical model so that various device models have assumed a constant mobility in integrating the basic differential equation along the channel and then attempted to add in mobility corrections to the basic derived device equation. An example of this is the basic "quadratic" MOS current expression developed in the early 60's [1-3] which can be expressed below the current saturation region by:

$$I_d = \mu_n C_{ox}(W/L)(V_g - V_T - 0.5V_d)V_d. \qquad (1)$$

One of the most common is to modify the mobility so that it becomes a function of the form:

$$\mu_n = \frac{\mu_o}{[1+\theta_1(V_g - V_T)+\theta_2(V_g - V_T)^2][1+V_d/E_cL]} \qquad (2)$$

where the first term in square brackets in the denominator is intended to account for the known reduction in mobility due to the vertical gate oxide field and the second term in square brackets is intended to account for velocity saturation effects due to a large horizontal drain field.

In recent years modeling attention has been addressed to "inversion charge approaches" (EKV, ACM and BSIM5 models) and to "surface potential" based

models (HiSIM, SP, MM11 and PSP models)[4-14]. These models have made possible the smooth integration of the subthrehold region with the above threshold region and form the basis for modern circuit analysis programs. Much attention has also been given in these recent models to ensuring that high order derivatives of the current expressions are finite and continuous for all voltages as needed for many circuit simulations.

In these past models, mobility reduction effects due to the vertical and horizontal fields are typically taken into account by use of multiplicative factors such as Equation (2) although the exact form of the mobility reduction factors as well as the voltage dependent terms may be more complex than that given in this example equations. The mobility reduction effects are thus taken as some average effect along the channel. Using such a multiplicative correction factor for mobility works well when the mobility correction is a small correction (few percent) to the differential equation. However, this approach does not properly account for mobility effects when the mobility reduction is a first-order correction to the basic differential equation. Including these as multiplicative effects gives too low a mobility as it multiplies a mobility term that is already reduced by the gate field factor. The net result is a mobility reduction factor where the parameters can only be treated as adjustable parameters with no direct connection to underlying device physics. In a recent paper it has been shown that a first-order model of MOS mobility can be incorporated directly into an analytical model for the MOS transistor[15]. The major features of this model are reviewed in the next section

2 NEW MOS DEVICE MODEL

The basic differential equation describing voltage along the channel of an MOS transistor is:

$$I_d = m_n C_{ox} W (V_g - V_T - V) \frac{dV}{dx} \qquad (3)$$

This equation neglects depletion layer charge and is frequently referred to as the sheet charge layer model. Depletion layer effects will subsequently be considered as a second order effect. The two major physical effects that must be considered for the mobility are: (a) dependency on vertical field and (b) dependency on horizontal field. All theoretical models for the surface roughness mobility predict a dependence on the square of the vertical electric field [16]. With such a field dependency, it is possible to express the fall-off in mobility at high fields by an expression of the form (see [15]):

$$\frac{1}{m_n} = \frac{1}{m_o}\left[1 + \left(\frac{V_g - V_T}{V_{gx}} + b\right)^2\right] \qquad (4)$$

where the parameters V_{gx} and b are defined as:

$$V_{gx} = \frac{e_{si}}{hC_{ox}}\sqrt{K_{sr}/m_o}, \quad b = \frac{Q_d}{e_{si}}\sqrt{m_o/K_{sr}}\ . \qquad (5)$$

In these equations, K_{sr} is a surface roughness constant with a value of about $6X10^{14}$ Volts/sec [16], Q_d is the depletion layer charge density and h is a parameter relating the inversion layer charge to the surface field through the expression:

$$E_{eff} = (hQ_i + Q_d)/e_{si}. \qquad (6)$$

Equation (4) has been found to provide a very good fit to experimental surface mobility for MOS devices at large effective fields. Figure 2 shows a comparison of this equation to experimental data for an MOS device with an oxide thickness of 1.66 nm. Two possible fits are shown with one taking b = 0 which neglects depletion layer charge and the other optimized for the best-fit value of depletion layer charge. It is seen that for this example, the neglect of depletion layer charge for the mobility model provides a good fit to the data. Neglecting b also provides a simpler set of MOS device equation.

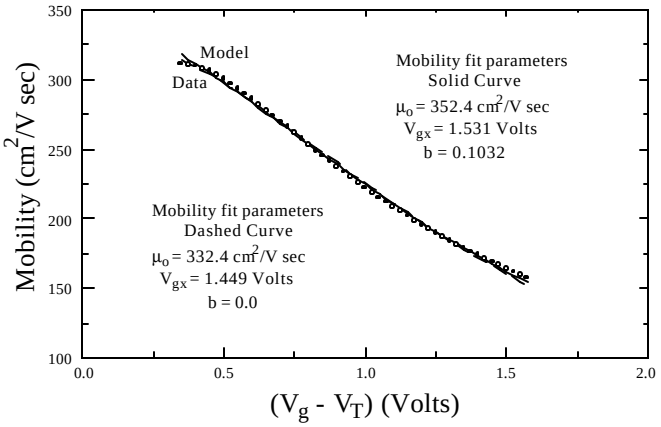

Figure 2. Comparison of mobility model with experimental data for an n-channel MOS device with oxide thickness of 1.66 nm.

As discussed in the previous paper [15], the proper manner to include the high-field limit to carrier velocity is in a reciprocal addition manner to the other mobility effects. Then a first-order mobility model is:

$$\frac{1}{m_n} = \frac{1}{m_o}\left[1 + \left(\frac{V_g - V_T - V}{V_{gx}} + b\right)^2 + \frac{1}{E_c}\left|\frac{dV}{dx}\right|\right] \qquad (7)$$

In addition to m_o, a low field them, the mobility model has two parameters associated with the fall-off at large surface fields (V_{gx} and b) and the more well known high field parameter of E_c which is related to the saturated velocity by $v_{sat} = m_o E_c$.

The resulting differential differential with this mobility model is:

$$\frac{I_d}{m_o} = \frac{[C_{ox}W(V_g - V_t - aV) - I_d/m_o E_c]}{[1+((V_g - V_t - aV)/V_{gx} + b)^2]} \frac{dV}{dx} \qquad (8)$$

The term a has been introduced into the equation for channel charge in order to provide a first-order correction term for depletion layer charge which causes the channel charge to decrease slightly faster than the linear dependence shown in Equation (7). A theoretical value for alpha can be found in many references as [4]:

$$a = 1 + \frac{0.5\sqrt{2e_{si}qN_b}}{C_{ox}\sqrt{2f_F + V_s}} \qquad (9)$$

The unique feature of the previous publication is the realization that the basic MOS device Equation (8) can be integrated explicitly with this improved mobility model.

Equation (8) is now readily integrated and the I_d terms collected together to give the final device equation in the form:

$$I_d = \frac{I_o}{2}\frac{[\ln(1+y_s^2) - \ln(1+y_d^2) - 2b[\tan^{-1}(y_s) - \tan^{-1}(y_d)]}{1+Z|\tan^{-1}(y_s) - \tan^{-1}(y_d)|}, \qquad (10)$$

where

$$I_o = m_o C_{ox}(W/aL)V_{gx}^2,$$
$$y_s = b + (V_g - V_T - aV_s)/V_{gx} = b + q_s/Q_{gx},$$
$$y_d = b + (V_g - V_T - aV_d)/V_{gx} = b + q_d/Q_{gx},$$
$$Z = m_o V_{gx}/(a v_{sat} L) = V_{gx}/V_{dx},$$
$$\text{with } V_{dx} = (a v_{sat}/m_o)L = a E_c L,\ Q_{gx} = C_{ox}V_{gx}.$$

For the case of $b = 0$ the equation simplifies to the simpler form:

$$I_d = \frac{I_o}{2}\frac{[\ln(1+x_s^2) - \ln(1+x_d^2)]}{1+Z|\tan^{-1}(x_s) - \tan^{-1}(x_d)|}, \qquad (11)$$

where

$$x_s = (V_g - V_T - aV_s)/V_{gx} = q_s/Q_{gx},$$
$$x_d = (V_g - V_T - aV_d)/V_{gx} = q_d/Q_{gx}.$$

In the denominator the term multiplying the Z factor involves the magnitude of some $\tan^{-1}()$ terms. This magnitude arises because of the magnitude term in the mobility expression of Equation (7).

Equation (10) (or (11)) provides an important starting point for a new compact MOS device model with an integrated model of the channel mobility. The I_d model is slightly more complex than previous MOS device models and includes in addition to the usual MOS parameters the additional parameters of b and V_{gx}. For the simpler model of Equation (11) only the additional parameter V_{gx} is present and this has the simple physical interpretation of being the gate voltage at which the surface mobility is reduced by a factor of 2 from its low field value due to surface roughness scattering. The new I_d expression reduces to the quadratic voltage dependence of Equation (1) for the parameter values: $b = 0, V_{gx} \gg (V_g - V_t)$, and $Z = 0.$ The model also has the required feature that the interchange of source and drain voltages produces a reversal of sign of the current. The current value as well as the derivative is also continuous for all source, gate and drain voltages.

Theoretical I_d-V_d curves as predicted by the new device model are shown in Figures 3 and 4 for the $a = 1.0$ case. The case of $b = 0$ is shown in each figure as the set of solid curves while the case of $b = 0.2$ is shown as a set of dotted curves. The assumed value of the V_{gx} parameter in both figures is 1.5 Volts which is a

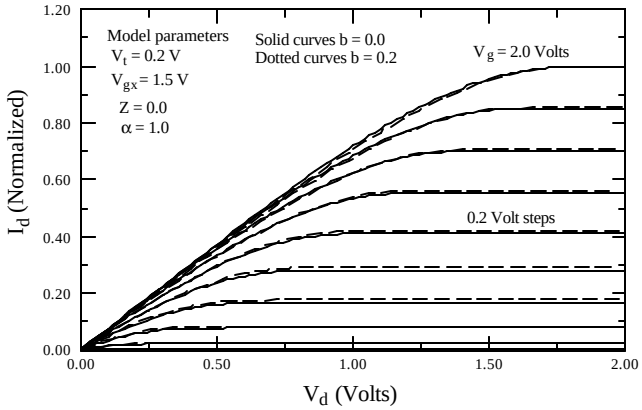

Figure 3. Theoretical I_d-V_d curves for parameters typical of a long channel device with an oxide thickness of about 1.5nm.

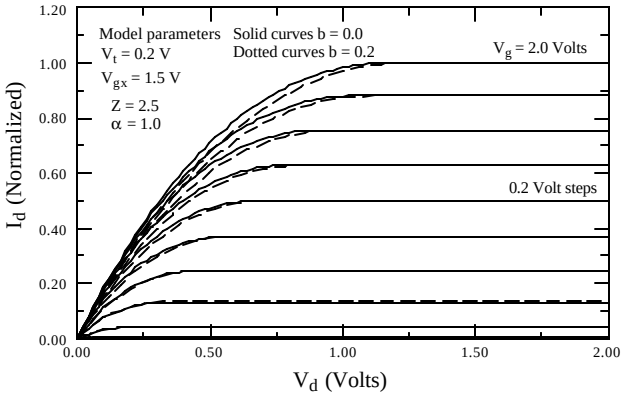

Figure 4. Theoretical I_d-V_d curves for parameters typical of a short channel device with an oxide thickness of about 1.5nm.

typical theoretical value for an oxide thickness of about 1.5nm. The long channel case corresponding to $Z = 0$ is shown in Figure 3. As in most device equations, one must find the value of drain saturation voltage or the point at which the theoretical curve has a maximum value and then extend the maximum current value into the saturation region for larger drain voltages. The drain saturation voltage is found from the device equation in the normal manner of:

$$\left.\frac{dI_d}{dV_d}\right|_{V_{dsat}} = 0, \text{ or } \left.\frac{dI_d}{dx_d}\right|_{x_{dsat}} = 0 \qquad (12)$$

where either Equation (10) or (11) is used for the model equation. For the simpler model of Equation (10) the resulting expression is:

$$x_{dsat} = \frac{Z}{2}\frac{\ln(1+x_s^2) - \ln(1+x_{dsat}^2)}{1 + Z[\tan^{-1}(x_s) - \tan^{-1}(x_{dsat})]} \qquad (13)$$

This transcendental equation can be solved by a few successive substitutions into the right hand side or by a few iterations of Newton's method. The saturation drain current is then found by using this expression in Equation (10). The corresponding expression for Equation (11) is only slightly more complex, and is not presented here.

One of the main features of the long channel characteristics in Figure 3 is the almost equal steps in saturated drain current for equal steps in gate voltage. This is typically seen in experimental data. This is in contrast to the basic Equation (1) which predicts a quadratic increase in drain current with gate voltage in the saturation region. Since the present model directly includes the voltage dependent mobility in the differential equation, this behavior is included in a first-order manner. For the long channel device there is little difference between the curves for the two values of b shown. Figure 4 shows the theoretical characteristic of a corresponding short channel device with $V_{dx} = 0.6$ Volts. This would correspond theoretically to a channel length of approximately 150nm (using $v_{sat} = 1X10^7$ cm/sec and $m_o = 250$ cm^2/Vsec). Again the predicted saturation curves are approximately equally spaced for equal gate voltage steps. The major difference is below saturation where the predicted curves are more rounded than for the long channel device. For the short channel device, values of the b parameter have a slightly larger effect on the shape of the characteristic below the saturation region. The dotted curves increase less rapidly below saturation than do the solid curves. Thus this parameter will be a significant factor primarily below saturation and at short channel lengths.

3 ADDITIONAL ENHANCEMENTS

As a model for use in circuit simulations, the equations so far developed have some undesirable features. The current model is everywhere continuous and symmetrical in the source and drain voltages. However, using an abrupt transition to the saturation region, leads to a discontinuous second derivative in the current at the saturation voltage. This is characteristic of many MOS device models and several techniques are available for smoothing the transition to the saturation region. One such technique is to replace the value of x_d in Equation (13):

$$x_d \rightarrow x_{dsat} + 0.5\left(x_d - x_{dsat} - x_{do} + \sqrt{(x_d - x_{dsat} - x_{do})^2 + 4x_{do}x_{dsat}}\right) \qquad (14)$$
$$\rightarrow \begin{cases} x_d & \text{for } x_d \gg x_{dsat} \text{ or } V_d \ll V_{dsat} \\ x_{dsat} & \text{for } x_d \ll x_{dsat} \text{ or } V_d \gg V_{dsat} \end{cases}$$

where x_{do} is some small value selected to give a smooth transition to the saturation region. This expression is similar to the smoothing function used in BSIM4 for a similar purpose [5].

The magnitude function in the denominator around the inverse tangent functions also gives a discontinuous second derivative at $x_d \rightarrow x_s$ or at $V_d \rightarrow V_s$. This can be traced back to the use of a linear field dependent term in Equation (7) to describe the high field mobility reduction. It is known that this is not a good approximation at very low fields and a more accurate expression involves a second power of the electric field at low fields. To obtain a smoothing function that eliminates this discontinuous second derivative is only a little more involved than that for the saturation voltage. Consider just

the denominator of Equation (10) or (11) which can be written as:

$$1+|\tan^{-1}(x_s)-\tan^{-1}(x_d)|$$
$$=1+|\tan^{-1}((x_s-x_d)/(1+x_s x_d))| \qquad (15)$$
$$\rightarrow 1+|(x_s-x_d)/(1+x_s x_d)| \text{ as } x_d \rightarrow x_s$$

The third line shows the limiting value as the drain voltage approaches the source voltage. An appropriate smoothing function for the denominator can then be obtained by replacing the $(x_s - x_d)$ term in the above equation by the expression:

$$(x_s-x_d) \rightarrow x_{sdz} = x_{so}\left(\sqrt{1+((x_s-x_d)/x_{so})^2}-1\right) \quad (16)$$
$$\rightarrow \begin{cases} (x_s-x_d)^2/x_{so} & \text{for } (x_s-x_d) \ll x_{so} \\ (x_s-x_d) & \text{for } (x_s-x_d) \gg x_{so} \end{cases}$$

By selecting an appropriate small value of x_{so} the replacement will have little effect on the model at large drain-to-source voltages but results in a smooth transition in all derivatives as $V_d \rightarrow V_s$ or $x_d \rightarrow x_s$. With this substitution one can eliminate the magnitude sign on the denominator term. As previously shown [15] this new parameter is equivalent to taking a quadratic dependence of mobility on electric field at low lateral fields.

These two smoothing functions eliminate the second derivative discontinuities in the new MOS model and result in a model that is symmetrical in the source and drain voltages and has no singularities in the derivatives. The smoothing functions with appropriate values, have very little effect on the fitting of the model to $I_d - V_d$ as previously shown [15].

In order to compare the model with experimental data, the model must be enhanced by some additional physical effects. The most obvious enhancements must be some means of providing a finite slope on the I_d-V_d characteristic in the current saturation region. The simplest means of accomplishing this is through a $(1+l_n V_{ds})$ term multiplying the current expression and a drain induced barrier lowering (DIBL) factor s_D. With these modifications, the new device model for $b = 0$ becomes:

$$I_d = \frac{I_o}{2}\frac{[\ln(1+x_s^2)-\ln(1+x_d^2)]}{1+Z[\tan^{-1}(x_{sdz}/(1+x_s x_d))]}[1+l_n(V_d-V_s)], \quad (17)$$

with

$$x_s = (V_g - V_T - s_D(V_d - V_s) - aV_s)/V_{gx},$$
$$x_d = (V_g - V_T - s_D(V_d - V_s) - aV_d)/V_{gx}.$$

The model can be further extended into an inversion charge approach by expressing the model in terms of surface charges q_s and q_d at the source and drain respectively. This is shown in Equations (10) and (11). This provides a pathway for extending the basic model into the subthreshold region by use of appropriate expressions for surface charge density in terms of device voltages. For the subthreshold region, the mobility should be further modified to account for a falloff in value at low surface fields as seen in Figure 1. Work is continuing on including these effects in an appropriate manner in the basic improved device model.

4 COMPARISON WITH MOS DATA

This model has been compared with experimental I_d-V_d characteristics for a series on nMOS devices of varying channel lengths. The experimental devices have a gate oxide thickness of approximately 1.66 nm extracted from C-V measurements and correcting for Quantum mechanical effects and poly depletion effects. The devices were all 10 μm in width with drawn channel lengths varying from 0.13 μm to 10 μm. Before fitting to the device equations, low field mobility was extracted from I_d-V_g measurements taken at a low drain voltage of 25mV using standard techniques [17]. The resulting mobility for the 0.13 μm device in the high field region is the data shown in Figure 2. A modified shift and ratio method was also used to extract the effective channel lengths of the devices and the source and drain resistance. This technique gave a channel length reduction factor of 0.033 μm from the drawn channel length, so the effective channel length of the shortest channel device was about 0.097 μm. The device source and drain resistance was found to be a negligible correction for the device data.

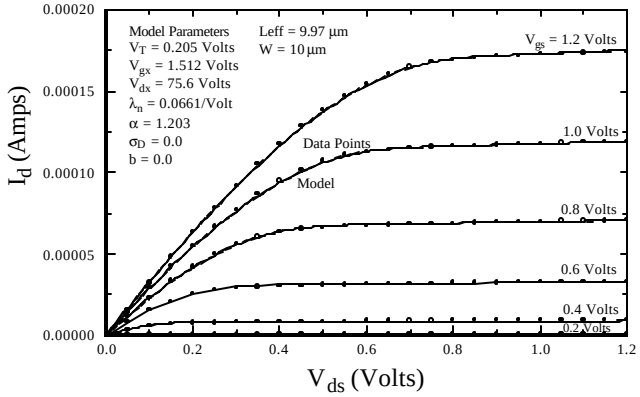

Figure 5. . Least-squares fit of the new device model to long-channel MOS device with W = L = 100 µm. Oxide thickness of the device is 1.66 nm. Solid and dotted curves overlay on figure.

Figures 5 and 6 show fits of the device model to experimental data for the long-channel and short-channel devices as described above. As seen in both figures, the model produces close to equal steps in output current for equal steps of gate voltage as is typically observed in experiments. Best fit values of the various model parameters are listed on the figures. There are actually two sets of data plotted in Figures 5 and 6, one using solid lines and one using dotted lines. The dotted set of curves is the least squares fit without the two smoothing functions discussed in the previous section and the solid lines are with the smoothing functions The two sets of fitted curves are so close to each other that the difference in the solid and dotted curves can not be seen on the scale of the figures.

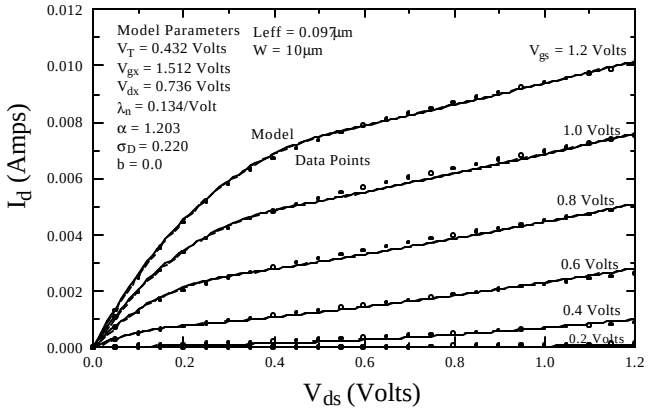

Figure 6. . Least-squares fit of the new device model to short-channel MOS device with L_{eff} = 0.097 μm, W = 100 μm. Oxide thickness of the device is 1.66 nm. Solid and dotted curves overlap on figure.

5 CONCLUSIONS

The model fit to the experimental data is very good. For the long channel device, the model provides almost an exact fit to the experimental data. For the short channel device the fit is not quite as good but the model still provides a very close fit to the experimental data. The long channel data and short channel data provide different tests of the fundamental model and model parameters. The long channel data is most sensitive to the mobility parameters m_o, V_{gx} and b plus the depletion charge parameter a . The long channel data is relative insensitive to the velocity saturation parameter as indicated by the large value of the V_{dx} parameter listed in Figure 5. Values without the smoothing functions were just slightly different from these values.

The short channel device data is sensitive not only to the surface mobility parameters, but also to the velocity saturation parameter in the basic model. Also the short channel data shows a considerable DIBL effect as evidenced by the fact that the threshold voltage is near 0.4 Volt for low drain voltages, but decreases to below 0.2 Volts for large drain voltage as can readily be seen in the I_d-V_d characteristic in Figure 6. As previously discussed, the best-fit model parameters are very close to what one would expect from a purely theoretical basis. For a comparison to the small signal parameters, the reader is referred to the previous paper [15].

This improved MOS device model with an integrated model for surface mobility can form the basis for improved compact models for circuit simulation.

REFERENCES

[1] C. T. Sah, IEEE Trans. Elecron. Dev. ED-11, 1964, pp. 324-345.

[2] J. T. Wallmark and H. Johnson, Field Effect Transistors: Physics, Technology and Applications, Prentice-Hall, 1966.

[3] H. Sichman and D. Hodges, IEEE J. Sol. St. Circ. Vol SC-3, 1968.

[4] N. Arora, MOSFET Models for VLSI Circuit Simulation, Springer-Verlag, 1993.

[5] X. Xi, K. M. Cao, H. Wan, M. Chan and C. Hu, "BSIM4.2.1 MOSFET Model," Dept. of Elec. And Comp. Sci., Univ of Calif. At Berkeley, 2001.

[6] C. Enz, K. Krummenacher, E. Vittoz., J. on Analog Int. Circuits and Signal Proc., Kluwer Academic Pub., July 1995.

[7] P. H. Woerlee, M. J. Knitel, R. van LanLangevelde, D. B. M. Klaassen, L. F. Tiemeijer, A. J. Scholten and A. T. A. Zegers-van Duijnhoven, IEEE Trans. On Elec. Dev., vol. 48, 2001.

[8] Jin He, Jane Xi, Mansun Chan, Hui Wan, Mohan Dunga, Babak Heydari, Ali M. Niknejad, Chenming Hu. ISQED, vol. 00, 2005.

[9] A.I.A. Cunha, O. C. Gouveia Filho, M. C. Schneider and C. Galop-Montoro, Proc. Of IEEE-IGCAS, Hong-Kong, June 1997.

[10] M. Suetake, K. Suematsu, H. Nagakura, M. Miura-Mattausch, H. J. Mattausch, S. Kumashiro, T. Yamaguchi, S. Odanaka, and N. Nakayama, Proc. SISPAD, 2000.

[11] M. Miura-Mattausch, H. Ueno, M. Tanaka, H. J. Mattausch, S. Kumashiro, T. Yamaguchi, K. Yamashita and N. Nakayama, Proc. IEDM 2002.

[12] G. Gildenblat, H. Wang, T.-L. Chen, Y. Guad and X. Cai, IEEE J. of Solid-State Cir., vol. 39, September 2004.

[13] R. vanLangevelde, A. J. Scholten and D. B. M. Klaassen, NL-UR 2004/00085, Koninklijke Philips Electronics, September 2004.

[14] R. vanLangevelde, A. J. Scholten, G. D. J. Smit, D. B. M. Klaassen, G. Gildenblat, X. Li, H. Wang and W. Wu, "PSP Model", NL-UR 2005/00303 Koninklijke Philips Electronics, September 2004.

[15] J. R. Hauser, IEEE Trans. On Elec. Dev., vol. 52, 2005.

[16] H. Shin, G. M. Yeric, A. F. Tasch and C. M. Maziar, Solid-State Elec., vol. 34, 1991.

[17] J. R. Hauser, IEEE Trans. On Elec. Dev., vol. 43, 1996.