Threshold Voltage Model for Deep-Submicrometer MOSFET's

Zhi-Hong Liu, Member, IEEE, Chenming Hu, Fellow, IEEE, Jian-Hui Huang, Tung-Yi Chan, Member, IEEE, Min-Chic Jeng, Ping-K. Ko, Member, IEEE, and Y. C. Cheng, Senior Member, IEEE

Abstract—The threshold voltage, $V_{th}$, of lightly doped drain (LDD) and non-LDD MOSFET's with effective channel lengths down to the deep-submicrometer range has been investigated. Experimental data show that in the very-short-channel length range, the previously reported exponential dependence on channel length and the linear dependence on drain voltage no longer hold true. We use a simple quasi-two-dimensional model, taking into account the effects of gate oxide thickness, source/drain junction depth, and channel doping, to describe the accelerated $V_{th}$ roll-off and the nonlinear drain voltage dependence. Relative to non-LDD devices, LDD devices have a smaller dependence of $V_{th}$ on channel length due to their lower drain-substrate junction built-in potentials. LDD devices also show less $V_{th}$ dependence on drain voltage because the LDD region reduces the effective drain voltage. Based on consideration of the short-channel effects, it is shown that the minimum acceptable length is equal to $(0.0035 - 0.005) X_{ch}^{1/3} T_{OX}$ micrometer ($X_{ch}$ in μm, $T_{OX}$ in Å).

I. INTRODUCTION

The reduction in threshold voltage with decreasing channel length and increasing drain-source voltage is widely used as an indicator of the short-channel effect for evaluating technologies. Two general approaches have previously been used to model this phenomenon. Two-dimensional numerical simulation is one of them [1]. The alternative is to develop a two-dimensional analytical solution, using the charge sharing approach [2], [3] or simplifying Poisson's equation in the depletion region [4]--[6]. Between the two analytical approaches, the two-dimensional analytical solutions of Poisson's equation in the depletion region reveal an expression for the threshold voltage shift that is exponentially dependent on effective channel length and agrees better with experimental results than do the charge sharing models which predict a $1/L_{eff}$-dependent threshold voltage shift. Furthermore, in the range of submicrometer channel length, the charge-sharing model assumptions of constant surface potential and no divergence of electric field lines in the gate oxide are invalid for high drain and substrate biases. Therefore, the charge-sharing model is unable to model the drain-induced barrier lowering (DIBL) or the body effects [7]. On the other hand, when deriving the analytical solutions for the two-dimensional Poisson's equation in the depletion region, various approximations have been made for the boundary conditions. Consequently, model parameters lack physical meanings. In addition, the simple exponential models also fail to predict the accelerated $V_{th}$ reduction at very short-channel lengths.

In this work, a short-channel threshold voltage model for both lightly doped drain (LDD) and non-LDD MOSFET's is derived using a quasi-two-dimensional approach, similar to those used for modeling the substrate current and other hot-electron phenomena in MOSFET's [8]--[10]. With this model, the accelerated $V_{th}$ reduction observed in the very-short-channel range can be accurately predicted. It is shown that in addition to its hot-carrier immunity, the LDD device has a small $V_{th}$ roll-off with decreasing channel length and a lower $V_{th}$ sensitivity to $V_{GS}$. Since this model is based on solutions of the quasi-two-dimensional Poisson's equation in the depletion region, no error is introduced by partitioning the channel, yet the solution of the two-dimensional problem is simplified. Because of its simple functional form and computational efficiency, this quasi-two-dimensional $V_{th}$ model is suitable for the guidelines of technology design and can be used in circuit simulation.

II. MODEL

By applying Gauss's law to a rectangular box (Gaussian box) of height $X_{dep}$ and length $Δy$ in the channel depletion region (Fig. 1) and neglecting mobile carrier charge, the following equation can be derived [10]:

$$X_{dep} \frac{dE_{S}(y)}{dy} + \frac{V_{GS} - V_{FB} - V_{th}(y)}{T_{OX}} = qN_{SUB} X_{dep}$$

(1)

where $E_{S}(y)$ is the lateral surface electric field, $V_{th}(y)$ is the channel potential at the Si-SiO$_2$ interface, $V_{GS}$ is the...
gate-source voltage, \( V_{FB} \) is the flatband voltage, \( N_{SUB} \) is channel doping, \( T_{OX} \) is the gate-oxide thickness, and \( \varepsilon_{Si} \) and \( \varepsilon_{SiO2} \) are the permittivity of SiO2 and Si, respectively. The depletion layer thickness, \( X_{dep} \), is equal to 
\[
\sqrt{2 \varepsilon_{Si} (\phi_s - \phi_f)/qN_{SUB}},
\]
where \( V_{FB} \) is the substrate bias, \( \phi_s = 2\phi_f \) is the surface potential at the threshold of surface inversion, and \( \eta \) is a fitting parameter [10].

The first term on the left hand side of (1) is equal to the net electric flux entering the Gaussian box along the y direction. The second term represents the electric flux entering the top surface of the Gaussian box. There is no electric flux passing through the bottom of the Gaussian box.

The solution to (1) under the boundary conditions of \( V_s(0) = V_{bi} \) and \( V_s(L) = V_{DS} + V_{bi} \) (the substrate potential is taken as ground) is

\[
V_s(y) = V_{sl} + (V_{bi} + V_{DS} - V_{sl}) \frac{\sinh \left(\frac{y}{L}\right)}{\sinh \left(\frac{L}{l}\right)} + \left(\frac{V_{bi} - V_{sl}}{\sinh \left(\frac{L}{l}\right)}\right)
\]

In (2), \( V_{sl} = V_{GS} - V_{th} + \phi \), represents the long-channel surface potential, and \( V_{th} = V_{FB} + qN_{SUB} X_{dep} T_{OX}/\varepsilon_{SiO2} \) + \( \phi \) represents the long-channel threshold voltage. \( V_{bi} \) is the built-in potential between the source-substrate and drain-substrate junctions, and \( l \) is the characteristic length defined as

\[
l = \frac{T_{OX} X_{dep}}{\varepsilon_{Si} \varepsilon_{SiO2} \eta}.
\]

The channel surface potential expressed by (2) can be thought of as the long-channel surface potential modified by the S/D finger field. Note that \( X_{dep} \) is assumed to be a constant when solving (1). In reality, \( X_{dep} \) is a function of the drain voltage and the channel length [1], [6]. Keeping in mind that the effects of the variation of the lateral field in the depletion layer under the channel are incorporated through the fitting parameter \( \eta \) [8], [9], one may treat the term \( X_{dep}/\eta \) in (3) as an average of the depletion layer thickness along the channel. Although \( \eta \) (hence \( l \)) may also be a function of the drain voltage, it is a second-order effect as will be seen later. This quasi-two-dimensional approximation simplifies the solution of (1), yet retains accuracy. Therefore, \( \eta \) is treated as a constant for a given technology in the following discussion unless otherwise specified.

At a given \( V_{GS}, V_{FB}, \) and \( V_{DS} \), the channel potential distribution calculated using the new model is plotted in Fig. 2(a) for devices with different channel lengths. In contrast to the constant channel potential assumed by the charge sharing model, the new model predicts a large variation in potential along the channel for devices with short-channel length even when the drain voltage is low. This result has been verified by 2D numerical simulation [1], [11]. The channel potential has a minimum at \( y_0 \) which can be found by solving the equation \( dV_s(y)/dy = 0 \). The minimum value of channel potential will increase, i.e., the potential barrier for electron flow from source to drain will decrease, with decreasing channel length and increasing drain voltage. Location \( y_0 \) and minimum potential \( V_{smin} \) can be obtained numerically by solving

\[
V_{smin} = V_s(y_0)
\]

\[
\left. \frac{dV_s}{dy} \right|_{y=y_0} = 0.
\]

However, when \( V_{DS} \ll V_{th} - V_{sl}, y_0 \) may be approximated as \( L/2 \) (see Fig. 2(a)). \( V_{smin} \) can therefore be obtained analytically from (4a)

\[
V_{smin} = V_{sl} + \left[2(V_{bi} - V_{sl}) + V_{DS}\right] \frac{\sinh \left(\frac{L}{2l}\right)}{\sinh \left(\frac{L}{l}\right)}.
\]

Recall that \( V_{sl} \) is a function of the gate voltage. To determine the threshold voltage, we will assume that \( V_{smin} \) is equal to

\[
2\phi_f - 2 \frac{kT}{q} \ln \left(\frac{N_{SUB}}{n_i}\right)
\]

when \( V_s = V_{th} \), which is approximately true with an error of \( kT/q \) [12]. Thus defined as the gate voltage which causes \( V_{smin} \) to equal \( 2\phi_f \), \( V_{th} \) can be solved as

\[
V_{th} = V_{th0} - \frac{\left[2(V_{bi} - \phi_f) + V_{DS}\right]}{2 \cosh \left(\frac{L}{2l}\right) - 2} = V_{th0} - \Delta V_{th}.
\]

When \( l \ll L, 1/[2 \cosh \left(\frac{L}{2l}\right) - 2] \) can be approximated as follows:

\[
\frac{1}{2 \cosh \left(\frac{L}{2l}\right) - 2} \approx \frac{1}{2e^{L/2l} - 2} = e^{-L/2l} \left(1 + 2e^{-L/2l}\right).
\]
In this case, the threshold voltage shift $\Delta V_{th}$ can be expressed as

$$\Delta V_{th}(L) = [2(V_{th} - \Phi_d) + V_{DS}] e^{-L/2L} + 2 e^{-L/l}. \quad (7)$$

It is worth noting that the right-hand side of (7) reduces to a single exponential term similar to the sample exponential models proposed in [4]–[6] when $L > \eta L$. For most technologies, the value of $\eta L$ is about 0.1–0.15 $\mu$m, hence the simple models are valid only when channel length exceeds 0.5–0.8 $\mu$m. As $L$ decreases, (7) predicts an accelerated $V_{th}$ reduction due to the existence of the second exponential term. It also should be mentioned that the more exact (6) correctly predicts the physical fact that $\Delta V_{th} \to \infty$ as $L \to 0$, i.e., the transistor cannot be turned off when channel length approaches zero. According to (7), a higher channel doping level, a lower $S/D$ doping level, or a thinner $T_{ox}$ will help to suppress the $V_{th}$ roll-off.

In general, when $V_{DS}$ is not small, $y_0$ will no longer equal $L/2$ as shown in Fig. 2(b) and (c). Therefore, (5)–(7) are not valid for large $V_{DS}$. When $L >> l$, (2) can be approximated as:

$$V_s(y) = V_{sl} + (V_{th} + V_{DS} - V_{sl}) e^{-y/L/L} + (V_{th} - V_{sl}) e^{-y/l} + (V_{th} + V_{DS} - V_{sl}) e^{-L/l}. \quad (8)$$

Similarly, $y_0$ can be found by equating the derivative of (8) to zero

$$y_0 = \frac{L}{2} - \frac{1}{2} \ln \left( \frac{V_{th} - V_{sl} + V_{DS}}{V_{th} - V_{sl}} \right). \quad (9)$$

Calculated results using (9) are plotted in Fig. 2(b) and (c), together with the results of the numerical solution of (4). Equation (9) is a good approximation of $y_0$, especially for the important case of small $L$. Although larger error will be seen at large $L$, this does not seriously affect the accuracy of $V_{min}$. Since $V_3$ is only a weak function of variable $y$ in the vicinity of $y_0$ (see Fig. 2(a)). Using (9), $V_{min}$ can be found from (2) and (4a)

$$V_{min} = V_{sl} - (V_{th} + V_{DS} - V_{sl}) e^{-L/L} + 2 \sqrt{(V_{th} - V_{sl} + V_{DS})(V_{th} - V_{sl}) e^{-L/L}}. \quad (10)$$
Fig. 3. The calculated $V_\text{th}$ shifts versus channel length at $V_{DS} = 0.05$ V. Note that the simple analytical solutions, (7) and (11), agree well with the numerical solution. The device parameters are the same as those in Fig. 2. Note that when $L > 5l$, all the curves have the same slope of $1/(2l \ln 10)$. (b) Comparison between simple analytical solutions and numerical solution of threshold voltage versus drain voltage. The solution based on the assumption $y_0 = L/2$, i.e., (7), overestimates the variation in threshold voltage.

Just as $\Delta V_\text{th}$ was calculated for the low $V_{DS}$ case, the more general case may be similarly derived from (10).

$$\Delta V_\text{th} = \frac{2(V_{bi} - \phi_0) + [V_{DS} + (V_{bi} - \phi_0)](1 - e^{-L/l}) + 2\sqrt{(V_{bi} - \phi_0)^2 + (V_{bi} - \phi_0)(V_{bi} - \phi_0) + V_{DS}}(e^{L/l} - 1)}{4 \sinh^2 \frac{L}{2l}}.$$  

For $L >> l$

$$\Delta V_\text{th} = \frac{3(V_{bi} - \phi_0) + V_{DS} + 2(V_{bi} - \phi_0)\sqrt{1 + [V_{DS}/(V_{bi} - \phi_0)]e^{L/l}}}{e^{L/l}}$$

$$= [3(V_{bi} - \phi_0) + V_{DS}]e^{-L/l} + 2\sqrt{(V_{bi} - \phi_0)(V_{bi} - \phi_0) + V_{DS}}e^{-L/l}.$$  

Equation (11b) reduces (7) for large $L/l$ and small $V_{DS}$ as expected. However, (11) predicts a weaker $V_{DS}$ sensitivity than (7) when $V_{DS}$ is high. $\Delta V_\text{th}$ is not proportional to $V_{DS}$ but has a functional form of $AV_{DS} + B\sqrt{V_{DS}}$. To first order, $A$ and $B$ are dependent only on device parameters.

The above analysis ignored possible voltage drop inside the drain diffusion. It is valid for both non-LDD and LDD devices as long as $V_{DS}$ is small. For an LDD device, $V_{bi}$ is the built-in potential of the n+/p junction. When $V_{DS}$ is large ($V_{DS} > 1$ V), the voltage drop in the drain region should be subtracted from the $V_{DS}$+ $V_{bi}$ terms in (7) and (11) for LDD devices. We will discuss this further in the following section.

Fig. 3 shows the calculated results using (7) and (11). The numerical solution of (4) is also shown in the same figure for comparison. When $L >> l$ and $V_{DS} < V_{bi} - \phi_0$, (7) gives a reasonable estimate of $V_{th}$ shift (Fig. 3(a)). Note that when $L > 5l$, the data can be approximated by a straight line with slope of $1/(2l \ln 10)$. Obviously, (7) overestimates $\Delta V_{th}$ high $V_{DS}$ since the assumption of $y_0 = L/2$ is not valid. However, (11) still accurately predicts the value of $\Delta V_{th}$ (Fig. 3(b)). To evaluate the accuracy of this model, results from the charge sharing model [3], 2D device simulation using MINIMOS [11], and the new model are compared in Fig. 4. The figure shows that while the charge sharing model underestimates the $V_{th}$ roll-off, the new quasi-2D model yields results comparable to those from 2D numerical simulation. In addition, the new model has a simple functional form.

III. EXPERIMENTAL

The devices used in this study are non-LDD and LDD nMOSFET’s with a phosphorus-doped polysilicon gate. The non-LDD MOSFET’s were fabricated using a photoresist ashing technique [13]. The LDD-MOSFET’s have an oxide spacer structure with an n” region length of 0.15-0.2 μm. The n” region was formed by P+ implant with a dose of $5 \times 10^{12}$ cm$^{-2}$. The gate oxides were grown in dry O$_2$ at 950°C with the thicknesses ranging from 55 to 400 Å. The dose and energy for both the channel (B+) and the S/D (As+) implants were adjusted for individual gate oxide thicknesses.

The effective channel length $l_{eff}$ was determined by the $C-V$ technique [14], which is suitable for both LDD and conventional devices. Two methods were used to extract the threshold voltage. One is to define the threshold voltage as the gate voltage at $I_{DS} = 10^{-7} \times W/L$ amperes with
Fig. 4. A comparison of the $V_{th}$ calculated using the charge sharing model, the two-dimensional numerical simulation (MINIMOS), and our model. The device parameters used are the same as those in Fig. 3.

$V_{DS} = 0.05$ V. Another method is to make an extrapolation from the $I_{DS}$ versus $V_{GS}$ curve, starting from the point of maximum $dI_{DS}/dV_{GS}$ to $I_{DS} = 0$, and define the $x$ intercept as the threshold voltage. Both methods produced nearly identical $\Delta V_{th}$ values. However, since no round-up error is introduced by calculating the derivative of $I_{DS}$ versus $V_{GS}$, the first method is more reproducible when $\Delta V_{th}$ is small. The shift of $V_{th}$ at high $V_{DS}$ was measured from the parallel shift of the log ($I_{DS}$) versus $V_{GS}$ in the subthreshold region. Typically, when $V_{DS} > 1$ V, the $V_{th}$ shift is defined as the shift of the log ($I_{DS}$) versus $V_{GS}$ curves at the $I_{DS}$ three decades lower than the drain current at which low drain voltage $V_{th}$ is defined.

IV. RESULTS AND DISCUSSION

A. Non-LDD Devices

The typical behavior of $V_{th}$ versus $L_{eff}$ i.e., $V_{th}$ roll-off, is shown in Fig. 5, where both experimental data and theoretical results from the present model are plotted for devices from several technologies. As can be seen, the proposed model correctly predicts the effects of $T_{OX}$ and $N_{SUB}$ on $V_{th}$. In addition, the effects of bias conditions on $V_{th}$ can be evaluated using the new model. The effects of drain bias will be discussed in detail in a later section.

As mentioned in Section II, the simple one-term exponential expression relating $\Delta V_{th}$ to $L_{eff}$ underestimates the reduction in $V_{th}$ with decreasing $L_{eff}$. Instead, the two-term exponential expression models this behavior better. This is illustrated in Fig. 6, where log ($\Delta V_{th}$) is plotted as a function of $L_{eff}$. As predicted by the new model, the experimental data show a steeper slope than and deviate from the straight line (dashed as dashed lines) predicted by the simpler exponential models at very-short-channel length ($L_{eff} < 5f$). This change in slope for the plot of log ($\Delta V_{th}$) versus $L_{eff}$ is also observable from the data of previous studies [4] although it was not pointed out. When $L_{eff} >> t$, the $V_{th}$ shift versus $L_{eff}$ for different substrate biases can be approximated by straight lines with an intersection of about $2(V_{th} - \phi_p)$, in agreement with (7). The slope of these straight lines is $1/21 \ln 10$, which is a function of the substrate bias since $t$ is proportional to the square root of $X_{dep}$ as indicated by (3).

Fig. 5. Experimental and calculated threshold voltage versus effective channel length for non-LDD MOSFET’s from different technologies, i.e., Device A: $T_{OX} = 55$ A, $N_{SUB} = 3.6 \times 10^{19}$ cm$^{-3}$, $X_i = 0.25$ pm, $t = 0.04$ pm; Device B: $T_{OX} = 86$ A, $N_{SUB} = 1.5 \times 10^{17}$ cm$^{-3}$, $X_i = 0.2$ pm, $t = 0.05$ pm; and Device C: $T_{OX} = 156$ A, $N_{SUB} = 4 \times 10^{16}$ cm$^{-3}$, $X_i = 0.2$ pm, $t = 0.09$ pm.

Fig. 6. Threshold-voltage shift versus effective channel length at $V_{GS} = 0.05$ V and different $V_{th}$ for non-LDD device. The solid lines are calculated results and the dashed lines are the lines best fitting the experimental data of $L_{eff} > 5f$. Note all the dashed lines intersect at the same point of $2(V_{th} - \phi_p) + V_{GS}$.

B. LDD Devices

Fig. 7 shows $V_{th}$ as a function of $L_{eff}$ for LDD devices operating at different bias conditions. It can be seen that the quasi-2D model still correctly predicts the $V_{th}$ behavior if proper refinements are made. For example, the value of $V_{th}$ should be the built-in potential of the n- substrate junction instead of the n+ substrate junction; the voltage drop in the n- region at higher $V_{DS}$ may not be negligible; and the bias for determination of $L_{eff}$ must be carefully chosen so that the n- region is not depleted when measuring the gate-channel capacitance [14].

To compare the sensitivity of $V_{th}$ to $L_{eff}$ for both LDD and non-LDD devices, $\Delta V_{th}$ for LDD and non-LDD devices from the same technology and with identical $T_{OX}$, $X_i$, and $N_{SUB}$ are plotted in Fig. 8 as a function of $L_{eff}$. In Fig. 8, the two intercepts differ by the difference in $V_{th}$ as explained earlier. As predicted by (7), the LDD device shows less $V_{th}$ reduction than the non-LDD device because the value of $V_{th}$ for the LDD device is smaller than that of the conventional device, although their $t$ values are comparable. Consequently, under the constraint of a tolerable $V_{th}$ roll-off, the shortest acceptable channel length of LDD devices will be shorter than that of non-LDD de-
Fig. 7. Typical threshold voltage behavior for LDD device.

Fig. 8. Comparison of $\Delta V_{th}$ versus $L_{eff}$ for LDD and non-LDD devices at $V_{DS} = 0.05$ V. The solid lines are calculated results and the dashed lines are the results from best fitting the experimental data for $L_{eff} > 5$. The $\Delta V_{th}$ of LDD devices is smaller by a factor of $[V_{th} \text{(non-LDD)} - \phi_s]/[V_{th} \text{(LDD)} - \phi_s]$.

The devices by a factor of

$$2 \ln \frac{V_{th} \text{(non-LDD)} - \phi_s}{V_{th} \text{(LDD)} - \phi_s}.$$  

It is worth pointing out that the accelerated $V_{th}$ roll-off is still visible as indicated in Fig. 8.

By applying an appropriate reverse bias to the substrate and S/D junctions of the LDD device, the junction potential can be made equal to the $V_{th}$ of a conventional device. $V_{th}$ of the LDD device with this special bias shows a $V_{th}$ reduction comparable to that of a normally unbiased non-LDD device. This is indicated by the data in Fig. 9 where a reverse bias of 0.13 V is applied to the S-substrate and D-substrate junctions of an LDD device.

C. The Effects of Drain Voltage

The present model and those from 2D analyses show that source/drain charge sharing and DIBL result from the same mechanism, namely, the channel potential lowering, although source/drain charge sharing usually refers to the $\Delta V_{th}$ measured at low drain voltage while DIBL usually refers to the $\Delta V_{th}$ induced by large drain voltage.

Fig. 10 presents $V_{th}$ of the non-LDD device as a function of $V_{DS}$. In agreement with previous studies, $V_{th}$ decreases as $V_{DS}$ increases. The shorter the channel length, the more severe the decrease of $V_{th}$. As accurately predicted by the present model, when $L_{eff}$ decreases, $V_{th}$ is no longer linearly dependent on $V_{DS}$ as predicted by common DIBL models [5], [15]. This nonlinear $V_{DS}$ dependency for very-short-channel devices is predicted by (11). According to (11), $V_{th}$ approaches a linear function of $V_{DS}$ at large $V_{DS}$. However, at low $V_{DS}$, $V_{th}$ of short-channel devices is approximately reduced to a square-root function of $V_{DS}$.

For comparison, results for LDD and non-LDD devices fabricated in the same technology are shown in Fig. 11. Relative to their non-LDD counterparts, LDD devices show suppressed DIBL effects as indicated by the smaller slope of $V_{th}$ versus $V_{DS}$. We attribute this to the drain voltage drop in the n-region. To correctly predict the DIBL effects in LDD devices, this voltage drop must be taken into account. Fig. 12 shows the surface field $E_S$ and the surface potential $V_S$ along the channel in an LDD device.
where \( E_j \) is the channel field at the metallurgical junction and \( L_{n^-} \) is the length of the \( n^- \) region. Let \( V_{\text{D eff}} \) equal \( V_{DS} \) minus the voltage drop in the LDD region. \( E_j \) can be approximated as

\[
E_j = \frac{V_{\text{D eff}} + V_{n^-} - V_{\text{min}}}{l}
\] (12)

i.e., the area under the approximately exponential \( E_S(y) \) curve is equal to \( E_j \times l \). Since for \( V_{th} \) measurements, \( V_{n^-} - V_{\text{min}} \) (≈ 0.1 V) is much lower than \( V_{\text{D eff}} \), one has

\[
V_D = V_{\text{D eff}} + V_{n^-} = E_j (l + L_{n^-}) \approx \frac{V_{\text{D eff}}}{l} (l + L_{n^-})
\] (13)

where \( V_{n^-} \) is the voltage drop in the \( n^- \) region and \( L_{n^-} \) is the length of the \( n^- \) region. Based on (13), \( V_{\text{D eff}} \) can be solved

\[
V_{\text{D eff}} = \frac{V_D}{1 + \frac{L_{n^-}}{l}}
\] (14)

Note that for a practical LDD structure, \( E_S \) is not constant in the LDD region as shown in Fig. 12, and the voltage drop is not \( E_j L_{n^-} \) but is proportional to \( \alpha E_j L_{n^-} \), where \( \alpha \) is a fitting parameter between 0 and 1, (14) can be rewritten as

\[
V_{\text{D eff}} = \frac{V_D}{1 + \frac{\alpha L_{n^-}}{l}}
\] (15)

Therefore, (6)–(11) can be modified for use in LDD devices by substituting \( V_{\text{D eff}} \) for \( V_{DS} \). Results with above correction are presented in Fig. 11. For a given technology, \( \alpha \) can be uniquely determined from the LDD doping profile [10] and hence does not depend on \( L_{eff} \). For most technologies, the empirical value of \( \alpha \) is between 0.3 to 0.7. Since value of \( l \) is generally 0.1–0.15 \( \mu \)m and \( L_{n^-} \) is 0.1–0.3 \( \mu \)m, the effective drain voltage for the LDD device is reduced by a factor of 1–3. Consequently, the DIBL effects can be greatly suppressed. We see that in addition to their hot-carrier immunity, LDD structures also promise to reduce \( V_{th} \) sensitivity to \( L_{eff} \) and \( V_{DS} \).

D. The Determination of Characteristic Length \( l \)

In this section, we discuss the determination of characteristic length \( l \) as it affects the accuracy of our quasi-2D \( V_{th} \) model. Although \( l \) calculated from (3) has the correct order of magnitude and function form, exact values of \( l \) need to be extracted from actual devices because of the unknown parameter \( \eta \). The extraction of \( l \) can be done by fitting the experimental data of \( \log (\Delta V_{th}) \) versus \( L_{eff} \) in the region of \( L_{eff} > 5l \). Based on (7) and Fig. 3(a), the slope of the fitted straight lines is equal to \( 1/(2l \ln 10) \). Extracted \( l \)'s versus the depletion layer thickness \( X_{dep} \) for several technologies are shown in Fig. 13. The different \( X_{dep} \)'s in this figure for a given technology correspond to different substrate biases. These straight lines with similar slopes of 2/3 suggest that \( l \) is proportional to \( X_{dep}^{1/3} \). Note that \( l \) is not proportional to \( X_{dep}^{1/2} \), and increasing \( V_{DS} \) somewhat decreases the slope of \( l \) versus \( X_{dep} \). This can be interpreted as saying \( \eta \) is also a function of \( X_{dep} \) and \( V_{DS} \). According to Fig. 13, the effects of \( V_{DS} \) on \( l \) or \( \eta \) should be of the second order. It has been found that for a given technology, a unique \( l \) (or \( \eta \)) extracted by the technology described above can be used for a wide range of \( L_{eff} \) and \( V_{DS} \) values (\( L_{eff} = 0.2 \text{–} 5 \mu \text{m}, \ V_{DS} = 0.05 \text{–} 3.5 \text{ V} \)).

According to (7), at \( L = 5l \), \( \Delta V_{th} \) is about 0.03 V. Assuming a typical subthreshold swing of 100 mV/decade, this \( \Delta V_{th} \) will increase the subthreshold current by two times. Therefore, the minimum acceptable channel length \( L_{min} \) [16] should be about 5l. According to Brews et al. [16], \( L_{min} \) follows an empirical expression of

\[
L_{min} = 0.41 (X_j T_{OX} X_{dep}^{1/3})^{1/3}
\] (16)

where \( L_{min}, X_j, X_{dep} \) are in micrometers and \( T_{OX} \) is in angstroms. Comparing with (16) and interpreting (16) as 5l, \( l \) should also be a cubic root function of \( X_j T_{OX} X_{dep} \). Using the same devices, both \( L_{min} \) from (16) and \( l \) from Fig. 13 are plotted in Fig. 14 against \( X_j T_{OX} X_{dep} \). By best fitting the
Fig. 13. Measured characteristic length $l$ as depletion layer thickness, i.e., substrate bias, is varied, for devices from different technologies. The solid diamond markers are the data collected at $V_{DD} = 3$ V and all the others are measured at $V_{DD} = 0.05$ V.

Fig. 14. The reported minimum channel length at which devices still show minimal short-channel effects is $L_{min}$, and measured characteristic length $l$, are plotted against $X_{OX}X_{dep}$, showing that the two are proportional to each other.

data of $l$, one has

$$l = 0.1(X_{OX}X_{dep})^{1/3}.$$  \hspace{1cm} (17)

This further confirms our previous estimation, i.e., $l \approx L_{min}/5$. In addition, the effects of $X_{OX}$ on $V_{th}$ are also incorporated into the model through (17). For an nMOSFET with an $n^+$ poly gate, it can be shown that in order to maintain $V_{th} = 0.7$ V, i.e.,

$$V_{th} = 0.7 = V_{FB} + \phi_{s} + \frac{T_{OX}}{\epsilon_{OX}} \frac{2\epsilon_{Si}\phi_{s}}{X_{dep}}$$

it is necessary that

$$X_{dep} = \frac{T_{OX}}{\epsilon_{OX}} \left( V_{th} - V_{FB} - \phi_{s} \right) \approx \frac{2\epsilon_{Si}}{\epsilon_{OX}} T_{OX}.$$  \hspace{1cm} (18)

By substituting (18) into (17), $l$ can be rewritten as

$$l = 0.0007 X_{OX}^{1/3} T_{OX}$$

where $X_{i}$ and $l$ are in micrometers and $T_{OX}$ is in angstroms. On the other hand, for an nMOSFET with a $p^+$ poly gate,

$$X_{dep} = \frac{4\epsilon_{Si}}{\epsilon_{OX}} T_{OX}$$

(assuming $V_{th} = 1.2$ V). Therefore, one has $l \approx 0.001X_{OX}^{1/3} T_{OX}$. This helps to explain why buried-channel devices, either nMOSFET's or pMOSFET's, generally have worse short-channel effects.

V. SUMMARY

The proposed model predicts the effects of $V_{DS}$, $L_{eff}$, $T_{OX}$, $N_{SUB}$, $X_{i}$, and body bias on threshold voltage. Previously reported simple exponential $\Delta V_{th} \sim L_{eff}$ model fails at $L_{eff} < 0.8$ $\mu$m, while the new model performs satisfactorily until $L_{eff}$ is as small as 0.1 $\mu$m. The new model also shows that $V_{th}$ is no longer a linear function of $V_{GS}$. Instead, a square root function of $V_{GS}$ can better model this nonlinear $V_{DS}$ dependence of $V_{th}$ for short-channel device. A simple scaling rule is that the channel length must be at least five times the characteristic length $l$, which in turn is proportional to $X_{OX}^{1/3} T_{OX}$ or $(X_{i}T_{OX}X_{dep})^{1/3}$.

LDD devices are found to have threshold voltages less sensitive to both the charge sharing and the DIBL effects. This finding suggests that in addition to their ability to suppress hot-carrier effects, LDD devices are also promising in terms of $V_{th}$ stability due to short-channel effects. The proposed new model also works well for LDD devices after properly taking into account the smaller $V_{th}$ and the voltage drop in the LDD region.

REFERENCES


Zhi Hong Liu (S’89–M’91) was born in Beijing, China. He received the B.Sc. and Master’s degrees from the South China University of Technology, Guangzhou, China, in 1978 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Hong Kong, Hong Kong, in 1990. In 1990, he joined the Department of Electrical Engineering and Computer Science, University of California, Berkeley as a research engineer, where he is now a research specialist. From 1978 to 1986 he was engaged in teaching and research of semiconductor material physics and device physics in the Department of Physics at the South China University of Technology. From 1986 to 1987, he was a visiting scholar in the Department of Electrical and Electronic Engineering at University of Hong Kong. His major research interests are in the areas of characteristics and reliability of MOSFET’s, thermally nitrided thin oxides, the physics, and modeling of the deep-submicrometer MOSFET’s. He has published over 30 technical papers and received a Best Paper Award at the 1991 International Electron Devices Meeting.

Chenming Hu (F’90) received the B.S. Degree from the National Taiwan University and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1970 and 1973, respectively. From 1973 to 1976 he was an Assistant Professor at the Massachusetts Institute of Technology, Cambridge. In 1976 he joined the University of California, Berkeley, as Professor of Electrical Engineering and Computer Sciences. He is Director of the Joint Services Electronics Program at Berkeley. While on industrial leave from the University during 1980 and 1981, he was Manager of nonvolatile memory development of National Semiconductor. Since 1973 he has served as a Consultant to the electronics industry. He has also been an advisor to many government and educational institutions. His present research areas include VLSI devices including silicon-on-insulator devices, hot-electron effects, thin dielectrics, electromigration, circuit reliability simulation, and nonvolatile semiconductor memories. He has also conducted research on electrooptics, solar cells, and power electronics. He has been awarded several patents on semiconductor devices and technology. He has authored or co-authored three books and 300 research papers. He has delivered twenty keynote addresses and invited papers at scientific conferences, and has received five best paper awards. He is an Honorary Professor of Beijing University and of Tsinghua University, China and of the Chinese Academy of Science. Dr. Hu was an Associate Editor of IEEE TRANSACTIONS ON ELECTRON DEVICES from 1986 to 1988, and Vice Chairman of IEEE Electron Devices Society, Santa Clara Valley Chapter, 1980–1982. He was appointed the first National Science Council Invited Chair Lecturer, Republic of China, from 1988 to 1991. He received the 1991 Design News Excellence in design Award for leading the development of an IC Reliability Simulator, BERT. He has been listed in Who is Who in America, American Men and Women of Science, Who is Who in Science and Engineering, and Who’s Who in American Education.

Jian-hui Huang received the B.S. degree in mechanical engineering from the University of Science and Technology of China, China, in 1985, the M.S. degree in physics from University of California in Berkeley in 1989, and the M.S. degree in electrical engineering from University of California at Berkeley in 1992. Currently, he is working toward the Ph.D. degree in the Department of Electrical Engineering and Computer Science, University of California at Berkeley. His research involves deep-submicrometer MOSFET modeling for VLSI circuit simulation, SOI MOSFET modeling, device physics, and hot-electron effects.

Tung-Yi Chan (S’85–M’87) was born in Taiwan, Republic of China. He received the B.S. degree in electrical engineering from the National Taiwan University in 1981, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1985 and 1987, respectively.

From 1987 to 1990, he was with Intel Corporation, where he was involved in the studies of deep-submicrometer CMOS and SOI devices, MOSFET reliability, high-reliability gate dielectric, low-temperature device operation, and the applications of CAD tools in technology development. In February 1990, he joined Cypress Semiconductor Corporation, and is currently in charge of the technology development for high-speed SRAM.

Min-Chieh Jeng was born in Taiwan, ROC. He received the B.S. degree from National Taiwan University in 1980, the M.S. degree from the University of Maryland, College Park, in 1983, and the Ph.D. degree in 1989 from the University of California, Berkeley, in electrical engineering. The subject of his master’s thesis was the ultrafast optoelectronic switches and semiconductor lasers. His research for the doctoral dissertation was on the design, characterization, and modeling of deep-submicrometer MOSFET’s. Since 1989, he has been with the Analog Devices, Cadence Design Systems, where he has been working on the development of semiconductor device models for circuit simulators. His other interests include submicrometer MOSFET performance and reliability studies, device characterization, and parameter extraction systems.

Ping K. Ko (S’78–M’82) received the B.S. degree in physics with special honors from Hong Kong University in 1974, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1978 and 1982, respectively.

In 1982 and 1983, he was a member of Technical at Bell Labs, Holmdel, NJ, and was responsible for developing high-speed MOS technologies for communication circuits. He joined the Berkeley faculty in 1984, where he is now Professor and Vice Chairman of the Department of Electrical Engineering and Computer Sciences. He is also the director of the Berkeley Microfabrication Laboratory. His present research interests include high-speed VLSI technologies and devices, device modeling for circuit simulation, CAD tools for IC, and electronic neural network. He has authored and co-authored one book and over 100 research papers.

Dr. Ko has served on program committees of the International VLSI Technology Symposium and the International Electron Device Meeting. He was Associate Editor of IEEE TRANSACTIONS ON ELECTRON DEVICES from 1988 to 1990.
Y. C. Cheng (M'78-SM'91) received the B.Sc degree in physics and mathematics from the University of Hong Kong in 1963 and the Ph.D. degree in theoretical physics from the University of British Columbia, Vancouver, B.C., Canada, in 1967.

From 1963 to 1978, he held various appointments at a number of universities in Canada and at Bell-Northern and Xerox Research Laboratories. During this period, he had undertaken research work on silicon devices and had co-invented and developed the "HCl-oxidation" technique for the production of clean oxides for integrated-circuit applications. In 1977, he was appointed as an Adjunct Professor at the University of Waterloo, Waterloo, Ont., Canada. He was a Professor in the Department of Electrical and Electronic Engineering and Dean of the Engineering Faculty, University of Hong Kong. In 1986, he was appointed as a Visiting Professor in the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. He is also an Advisory Professor of the Physics Department, the South China University of Technology. He has authored or co-authored over 100 technical papers and holds three U.S. and Canadian patents on semiconductor technology. Presently, he is the Director of the City Polytechnic of Hong Kong.