An Engineering Model for Short-Channel MOS Devices

KAI-YAP TOH, MEMBER, IEEE, PING-KEUNG KO, AND ROBERT G. MEYER, FELLOW, IEEE

Abstract — An engineering model for short-channel MOS devices which includes the effect of carrier drift velocity saturation is described. Based on a piecewise carrier drift velocity model, simplified expressions for the dc drain current $I_D$, the small-signal transconductance $g_m$, and the output conductance $g_D$, in the saturation region are derived. For a given gate voltage, the expressions depend only on the threshold voltage $V_T$ and the dimensions of the device whose desired values are normally known.

I. INTRODUCTION

Much has been published on the various short-channel effects in MOS devices from the device physics point of view, but there is no analytical dc model of short-channel MOS devices which circuit designers can use for hand calculations, such as in estimating the dc bias conditions and ac parameters. Common practice is to extend the use of the long-channel square-law model for the short-channel devices, but this approach gives considerable discrepancies in both the dc and ac solutions between the calculated and simulated results and can lead to incorrect conclusions in regard to circuit performance. On the other hand, elaborate models provided in circuit simulators involve a large number of model parameters whose values must be determined or derived from device measurements, often with the help of automated parameter extraction tools. It is difficult to develop an intuitive understanding of the device electrical behavior from such a list of parameters and it is not suitable for hand calculations. A simple engineering model for short-channel MOS devices that relates the terminal voltages to the drain current, much like the well-known square-law behaviors of the short-channel MOS device from the circuit designer's perspective. It will not only be useful for circuit design and analysis, it will also be useful to device designers who need to relate device and process parameters to circuit parameters.

Based on a piecewise carrier drift velocity model, simple closed-form $I-V$ relationships between the terminal voltages and drain current $I_D$, and the transconductance $g_m$ are derived, with $V_{GS}-V_T$ as the independent variable. $I_D$ and $g_m$ essentially depend on only the device width, the channel length, the device threshold voltage $V_T$, the gate oxide thickness $t_{ox}$, and the source/drain junction depth $x_j$, whose desired values are normally known. While $I_D$ and $g_m$ are sufficient for most digital MOS circuit calculations, the device $g_D$ in the saturation region is often required in analog MOS circuit calculations. Accurate modeling of $g_D$ is difficult, even in elaborate computer models. However, based on engineering approximations, an expression for estimating $g_D$ is derived and is found to correlate fairly well with experimental data.

The dc model presented here is a modification and extension of that discussed in [1]. In this paper, a different carrier mobility model [3] is incorporated and the electrical channel length $L_e$ is used when the device is biased into saturation. This leads to a more accurate prediction of the device $g_m$. The mobility model can be used to estimate the effective mobility $\mu_{eff}$ of a device under any gate bias without taking device measurements. The use of $L_e$ leads to a slightly more complicated calculation procedure but is necessary to extend the validity of the model down to 1-\mu effective channel length $L_{eff}$. This is because the drain depletion width $x_e$ takes up a significant portion of the $L_{eff}$, particularly for devices with channel length <1 \mu and operated at high $V_{GS}$. To make the model simple for hand calculations and yet reasonably accurate, it is necessary to introduce empirical constants to replace complicated terms in the expressions that have only second-order effects on the accuracy of the model. Devices with $L_{eff}$ from 3 \mu down to 1 \mu and with different $L_{eff}$ and $x_j$ have been characterized and found to be in good agreement with the model.

II. CARRIER DRIFT VELOCITY AND $\mu_{eff}$ MODELS

The main reason that the electrical characteristics of short-channel devices deviate from those of the long-channel devices is the dependence of the carrier drift velocity $v$...
on the longitudinal channel field $E$. For long-channel devices, $v = \mu_{\text{eff}} E$, where the effective carrier mobility $\mu_{\text{eff}}$ is assumed to be a constant. For short-channel devices, $\mu_{\text{eff}}$ is no longer a constant and is a function of the transverse field $E_t$ in the inversion layer. Increasing the transverse field will reduce the value of $\mu_{\text{eff}}$. Also, $v$ is no longer directly proportional to $E$ due to high field effects. Increasing $E$ will reduce $v$. A detailed examination of the relationship between $v$ and $E$ is therefore in order.

Several expressions used for $v$ have been discussed in [1]. The carrier drift velocity model depicted in (1) is commonly used in computer modeling:

$$v_1 = \frac{\mu_{\text{eff}} E}{1 + \left(\frac{E}{E_0}\right)^2}$$

where

$$E_0 \equiv \frac{v_{\text{sat}}}{\mu_{\text{eff}}}.$$  

Unfortunately, the use of (1) does not lead to analytical expressions for the $I-V$ characteristics of the short-channel devices. However, simple analytical expressions can be obtained if the carrier drift velocity is approximated by the piecewise model depicted in (2):

$$v = \begin{cases} \frac{\mu_{\text{eff}} E}{1 + \frac{E}{E_c}}, & \text{for } E \leq E_c \\ v_{\text{sat}}, & \text{for } E > E_c \end{cases}$$

where

$$E_c = \frac{2v_{\text{sat}}}{\mu_{\text{eff}}}.$$  

The value of $\mu_{\text{eff}}$ is crucial to the accuracy of the device model. As reported in [2] and [3], $\mu_{\text{eff}}$ can be estimated as follows:

$$\frac{1}{\mu_{\text{eff}}} = \frac{1}{\mu_1} + \frac{1}{\mu_2}$$

where

$$\mu_2 = \mu_1 \left(\frac{E_t}{E_{t,\text{eff}}}\right)^n.$$  

$\mu_1$, $E_t$, and $n$ are empirical constants listed in Table I. $E_{t,\text{eff}}$ is the effective transverse field at the surface and may be approximated as follows (see Appendix A):

$$E_{t,\text{eff}} = \frac{V_{GS} - V_T}{6t_{ox}} + \frac{V_T + V_a}{3t_{ox}}$$

where, semi-empirically, $V_a = 0.5$ V for typical n$^+$ poly-silicon gate devices.

### TABLE I

<table>
<thead>
<tr>
<th>Carrier</th>
<th>$\mu_1$ (cm$^2$/V·s)</th>
<th>$E_{t,\text{eff}}$ (MV/cm)</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron</td>
<td>670</td>
<td>0.67</td>
<td>1.6</td>
</tr>
<tr>
<td>hole</td>
<td>290</td>
<td>0.35</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Calculated values of $\mu_{\text{eff}}$ versus $E_{t,\text{eff}}$ for both electron and hole carriers are shown in Fig. 1. Variations of $\mu_{\text{eff}}$ versus $V_{GS} - V_T$ for typical values of $t_{ox}$ are also shown in Fig. 2(a) and (b) for the NMOS and PMOS devices, respectively. These curves may be used to estimate the value of $\mu_{\text{eff}}$ needed in the dc drain current expression to be discussed next.

### III. LARGE-SIGNAL BEHAVIOR OF SHORT-CHANNEL MOS DEVICES

The engineering model to be described here uses the piecewise carrier drift velocity model. It can be shown [1] that the drain currents $I_D$ in the ohmic and saturation...
regions are given as follows:

\[ I_D = \begin{cases} 
\frac{\mu_{eff} C_{ox} W}{L_{eff}} \frac{1}{1 + E_c L_{eff}} \left( V_{GS} - V_T - \frac{1}{2} V_{DS} \right) V_{DS}, \\
\text{for } V_{DS} \leq V_{Dsat} \\
v_{sat} C_{ox} W (V_{GS} - V_T - V_{Dsat}) = I_{Dsat}, \\
\text{for } V_{DS} > V_{Dsat}
\end{cases} \]  

(5)

where

- \( C_{ox} \) is the gate capacitance per unit area,
- \( W \) is the device width,
- \( L_{eff} \) is the effective channel-length between the source and the drain,
- \( V_{DS} \) is the drain-to-source voltage,
- \( V_{GS} \) is the gate-to-source voltage,
- \( I_{Dsat} \) is the drain current in the saturation region,
- \( V_{Dsat} \) is the saturation drain voltage, defined here as the drain voltage at which the carrier drift velocity saturates.

By equating the \( I_D \) expressions in (5) at \( V_{DS} = V_{Dsat} \), it can be shown that

\[ V_{Dsat} = (1 - K)(V_{GS} - V_T) \]  

(6)

where

\[ K = \frac{1}{1 + E_c L_{eff} / (V_{GS} - V_T)} \]  

and

\[ L_e = L_{eff} - X_d. \]  

(7)

In the above equations, \( L_e \) is the device electrical channel length and \( X_d \) is the depletion width into the channel from the drain when \( V_{DS} > V_{Dsat} \). \( X_d \) may be calculated as follows [4]:

\[ X_d = \frac{1}{A} \ln \left[ \frac{A(V_{DS} - V_{Dsat}) + E_d}{E_c} \right] \]  

where

\[ E_d = \sqrt{1 + \left[ A(V_{DS} - V_{Dsat}) / E_c \right]^2} \]  

(8)

in which 1.5 is a semi-empirical constant. \( I_{Dsat} \) can also be expressed in terms of \( K \) and \( V_{GS} - V_T \) by substituting (6) into (5), which leads to

\[ I_{Dsat} = K v_{sat} C_{ox} W (V_{GS} - V_T). \]  

(9)

This is a useful expression for \( I_{Dsat} \) in which all the short-channel effects are modeled by the factor \( K \). As will be shown later, \( K \) may be regarded as a constant under certain conditions.

Calculated values of \( X_d \) at \( V_{GS} - V_T = 2 \, \text{V} \) for both the NMOS and PMOS devices listed in Table II are plotted in Fig. 3(a) and (b), respectively (the \( \mu_{eff} \) model described in Section II has been used and \( v_{sat} = 7 \times 10^6 \, \text{cm/s} \) is assumed for both the NMOS and PMOS devices). Devices listed in Table II are fabricated with essentially the same P-well CMOS process except that the gate oxide thicknesses, implant doses, and source/drain drive-in times are different. The value of \( K \) in (7) may be calculated with the value of \( X_d \) estimated from Fig. 3 or with one round of
The measured and calculated device $I-V$ characteristics for $V_{DS}$ from 1 to 5 V at steps of 1 V. (a) An NMOS device. (b) A PMOS device.

Fig. 5. Calculated (a) $K$ and (b) $K_m$ versus $V_{GS} - V_T$ for the NMOS devices listed in Table II.

and

$$g_m \rightarrow \frac{v_{sat} C_0 W}{1 + (E_{n,eff}/E_i)^n}.$$  

These are the ultimate relationships for very short-channel MOS devices. However, for practical devices with finite channel length and operating under a practical bias condition, $K < K_m < 1$. As an example, the calculated variations of $K$ and $K_m$ with $V_{GS} - V_T$ for the NMOS devices listed in Table II are plotted in Fig. 5(a) and (b) respectively. It should be pointed out that $K$ and $K_m$ can have different values when a different value of $v_{sat}$ is assumed, as is evident from (10a). However, other considerations, such as the accuracy of $I_D$ and $g_m$, lead to the conclusion that $v_{sat} = 7 \times 10^6$ cm/s is the best choice for this engineering model. For the 1-µm NMOS device, the variations of $K$ and $K_m$ with respect to $V_{GS} - V_T$ may be roughly divided into three regions:

1) $V_{GS} - V_T < 0.3$ V: a linear region where the “square-law” long-channel model approximately holds;
2) $0.3 V < V_{GS} - V_T < 2.0$ V: a nonlinear region, where $K$ and $K_m$ should be calculated from (7) and (10); and
3) $V_{GS} - V_T > 2.0$ V: a flat region where $K = 0.65$ and $K_m = 0.88$.

In most digital MOS circuits, during turn-on and turn-off transients, $V_{GS} - V_T$ is typically greater than 2 V and therefore $K_m$ and $K$ may be treated as constants. In analog MOS circuits, $V_{GS} - V_T$ is typically within the nonlinear region and $K_m$ and $K$ should be calculated from the formulas. For the 3-µm device, the linear region is extended to approximately 0.7 V. The variations consist of essentially only the linear and nonlinear regions for $V_{GS} - V_T \leq 5.0$ V. It is apparent that even at 3-µm channel
length, the electrical characteristics do not follow the long-channel behavior.

To compare the accuracy of the model, calculated and measured $g_m$ are shown in Fig. 6(a) and (b), respectively, for the devices listed in Table II. As shown, the agreement is better than ±6 percent over most of the gate voltage range for channel lengths from 3 µm down to 1 µm.

V. SECOND-ORDER EFFECTS IN $I_D$

The primary effect on $I_D$ in the saturation region is due to the increase of $X_d$ with $V_{DS}$. Second-order effects that affect $I_D$ are not included in the dc model depicted in (5). This is because the deviations are small in comparison with $I_D$ itself. However, the deviation may be significant in comparison with $\Delta I_D$ at high $V_{DS}$ and constant $V_{GS}$ and thus will add to the value of $g_{ds}$. It is therefore necessary to include these effects in the formulation of $g_{ds}$. Two second-order effects will be discussed, namely the drain-induced barrier lowering (DIBL) effect on $V_T$ and the substrate current ($I_{SUB}$) effect on $I_D$.

For short-channel MOS devices, $V_T$ may depend on $V_{DS}$, especially at high values of $V_{DS}$ due to the DIBL effect [5] (this can be observed as a horizontal shift in the subthreshold conduction characteristics in a plot of log $I_D$ versus $V_{GS}$ curves for different $V_{DS}$). When the DIBL effect is small, $V_T$ may be approximated by

$$V_T = V_{T0} - \eta V_{DS}.$$  (12)

$V_{T0}$ is the device threshold voltage at $V_{DS} = 0$. As $V_{DS}$ increases, $V_T$ will decrease and $V_{GS} - V_T$ will increase, causing $I_D$ to increase.

$I_{SUB}$ is normally much smaller than $I_{DSat}$. Typically, $I_{SUB}/I_D < 10^{-2}$ at low $V_{GS} - V_T$ and high $V_{DS}$, which is the worst case. A more significant effect of $I_{SUB}$ on $I_D$ is the fact that $I_{SUB}$ tends to reduce the substrate bias, which reduces $V_T$, and in turn increases $I_D$. $g_{ds}$ can therefore increase significantly even at moderately low $I_{SUB}$. Unfortunately, there is no analytical form which relates this effect of $I_{SUB}$ to $I_D$ as yet. In addition, lightly doped drain (LDD) structures and retrograded substrate or well doping profiles reduce the effects of hot-electron current and $I_{SUB}$-making $g_{ds}$ dependent on the device structure. Consequently, it is difficult to estimate $g_{ds}$ of a short-channel MOS device accurately at large $V_{DS}$. Nevertheless, an approximate expression for $g_{ds}$ can be derived which will be discussed next.

VI. DEVICE OUTPUT CONDUCTANCE $g_{ds}$

When the effect of $I_{SUB}$ is omitted, it can be shown from (5) that the device conductance in saturation, denoted by $g_{ds0}$, can be approximated by (see Appendix C)

$$g_{ds0} = \frac{\alpha_0}{g_m} \left( \frac{V_{DS} - V_{DSat}}{V_{DSat}} \right)^2 + \frac{1}{g_m} + \eta g_m.$$  (13)

For a well-designed device, the first term is the dominant factor and is due to $X_d$. The second term is due to the effect of DIBL and is expected to be proportional to $g_m$ as $\Delta I_D = g_m \Delta (V_{GS} - V_T) = g_m \Delta V_T$ at constant $V_{GS}$. A large $\eta$ will increase $g_{ds0}$ and is obviously not desirable. As the channel length is reduced, $g_m$ is increased and it is therefore important to reduce the DIBL effect or the value of $\eta$ in order to maintain a reasonable device voltage gain ($g_m/g_{ds}$).

Equation (13) normally would give a good estimate of $g_{ds}$ biased at low $V_{DS} - V_{DSat}$. At high $V_{DS} - V_{DSat}$, the substrate current effect should be included when better accuracy is needed. However, the calculation procedures are more involved and will be discussed in Appendix C. The final result is

$$g_{ds} = g_{ds0} + \frac{\beta I_{SUB}}{(V_{DS} - V_{DSat})}. \quad (14)$$

From device measurement data, $\beta$ varies from 2.0 to 3.0 for conventional devices (without the LDD structure). It is clear from the above expression that both $\eta$ and $I_{SUB}$ should be minimized to reduce $g_{ds}$ (or to increase $g_m$) and to maximize voltage gain. Ignoring the effect of $I_{SUB}$, (10) and (13) predict that the maximum achievable dc gain of a 1-µm channel-length device biased well into saturation is about 29 when $\eta = 0$ and drops to about 18 when $\eta = 0.02$, a reduction of almost 40 percent.

Fig. 7 shows the computed and measured $g_{ds}$ of a 2-µm NMOS device (device B) versus $V_{DS}$ for $V_{GS}$ around and greater than $V_{DSat}$. The calculated data are based on $\eta = -0.02$ and $\beta = 2$. The agreement is within ±25 percent.
Fig. 7. Measured and calculated $g_m$ versus $V_{GS}$ at $V_T = 2$ V for an NMOS device.

VII. APPLICABILITY TO LONG-CHANNEL REGIME

The closed-form expressions derived are a consequence of the simplified carrier drift velocity model depicted in (2). There is no restriction imposed on the device channel length $L$, assuring that the dc model described is applicable to devices with long channel length as well. For long-channel devices

$$\frac{V_{GS} - V_T}{E_c L_e} \ll 1$$

which leads to

$$K = \frac{(V_{GS} - V_T)}{E_c L_e}.$$ 

Substituting the above and (2) into (5) and (6) will confirm the square-law $I-V$ relationship. The regime for the so-called "long-channel" length can now be quantified by setting $(V_{GS} - V_T)/E_c L_e \ll 0.1$. For example, for long-channel NMOS devices, $\mu_{eff} = 700$ cm$^2$/V·s, and at $V_{GS} - V_T = 1$ V, $L_e$ should be greater than 5 μm. It is therefore not surprising to observe that the 3-μm device (discussed in Section IV) does not exhibit long-channel device behavior.

VIII. PRACTICAL EXAMPLES

To demonstrate the use of the model, the peak discharge and charging current $I_L$ of the CMOS inverter as shown in Fig. 8 are to be calculated. A different CMOS technology is deliberately chosen to check the accuracy of the proposed model. Device parameters are extracted from vendor supplied SPICE2 level-2 MOS models. For the NMOS device, $t_{ox} = 25$ nm, $V_T = 0.8$ V, $L_{eff} = 1.05$ μm, and $x_j = 0.3$ μm, and for the PMOS device, $t_{ox} = 25$ nm, $V_T = -0.8$ V, $L_{eff} = 1.05$ μm, and $x_j = 0.4$ μm.

Case 1: Peak Discharge Current

Assume $V_{DS} = 3.4$ V, $V_{in} = 3.4$ V, $V_{out} = 3.4$ V, the NMOS is in saturation, and the PMOS is in the off state. Consider the NMOS device, $V_{GS} - V_T = 3.4 - 0.8 = 2.6$ V. From Fig. 2(a), $\mu_{eff} = 510$ cm$^2$/V·s for $t_{ox} = 25$ nm. From Fig. 3(a), $X_d = 0.25$ μm, approximating it in between device B and device C at $V_{DS} = 3.4$ V. From (7), $K = 0.542$. Using these as initial values in (8) and (7) gives $X_d = 0.316$ μm and $K = 0.563$. The peak discharging current per unit width can now be calculated from (9):

$$I_L/W = 0.563 \times 7 \times 10^{-6} \times \frac{3.9 \times 8.85 \times 10^{-14}}{250 \times 10^{-10}} \times 2.6 = 142 \mu A/\mu m.$$

Case 2: Peak Charging Current

$V_{in} = 0$ V, $V_{out} = 0$ V, the NMOS is in the off state, and the PMOS is in saturation. Consider the PMOS device, $|V_{GS} - V_T| = 3.4 - 0.8 = 2.6$ V. From Fig. 2(b), estimate that $\mu_{eff} = 150$ cm$^2$/V·s for $t_{ox} = 25$ nm. From Fig. 3(b), $X_d = 0.17$ μm and from (7), $K = 0.24$. Using these as initial values in (8) and (7) gives $X_d = 0.13$ μm and $K = 0.232$. Using (9) again, the peak charging current $I_L/W = 58 \mu A/\mu m$. The corresponding values from BSIM [6] simulations are 141 and 58 μA/μm for the discharge and charging currents, respectively.

As another example, the short-channel model described can be used to evaluate the first-order trends in the electrical characteristics of submicrometer-channel devices. As illustrations, Figs. 9 and 10 show the trends in $g_m$ versus $L_{eff}$, and $K$, $K_m$ versus $L_{eff}$, respectively, for NMOS devices with $L_{eff} \leq 2$ μm and for $t_{ox}$ of 20 and 10 nm. As expected, $g_m$ increases with decreasing $L_{eff}$, and $K$ and $K_m$ approaches 1 as $L_e$ approaches 0. It should be noted that $g_m$ does not double in value as $L_{eff}$ is reduced from 1 to 0.5 μm while keeping $t_{ox}$ constant.
Fig. 10. Variations of \( K \) and \( K_m \) versus \( L_{eff} \) for NMOS devices with \( t_{ox} = 20 \text{ nm} \) (solid line) and 10 nm (dotted line).

IX. CONCLUSIONS

The above discussions demonstrated that \( I_D \) and \( g_m \) of short-channel devices can be estimated for a given gate voltage and a given set of \( V_T, t_{ox}, \) and \( x \), whose typical values are usually known for a given process. \( g_{ds} \) can be estimated with the additional knowledge of \( \eta \) and \( \beta \). Often, an educated guess in the value of \( \eta \) and \( \beta \) is sufficient. The model described is also useful for evaluating the impact of a particular short-channel technology on circuit performance before any device is available. Despite the substantial deviations from the long-channel device behavior, several simple equations are sufficient to predict the \( I_D, g_m, \) and \( g_{ds} \) of conventional short-channel MOS devices to within reasonable accuracy. For conventional MOS devices, the accuracy of the model for \( I_D \) and \( g_m \) is typically better than \( \pm 6 \) percent, and that for \( g_{ds} \), which is sensitive to the device structure and the substrate doping profile, is typically better than \( \pm 25 \) percent with properly selected values of \( \eta \) and \( \beta \). These accuracies are certainly within the acceptable range for hand-calculation purposes.

APPENDIX A
DERIVATION OF \( E_{r,eff} \)

From \[7\], \( E_{r,eff} \) may be approximated by

\[
E_{r,eff} = \frac{1}{\epsilon_{st}} \left( \frac{Q_{INV}}{2} + Q_B \right)
\]

where \( Q_{INV} \) is the inversion charge and \( Q_B \) is the bulk charge. But

\[
Q_{INV} = C_{ox}(V_{GS} - V_T)
\]

and

\[
V_T = -V_s + \frac{Q_B}{C_{ox}} + Q_B = C_{ox}(V_T + V_s).
\]

Thus

\[
E_{r,eff} = \frac{\epsilon_{ox}}{\epsilon_{st}} \left( \frac{V_{GS} - V_T}{2} + (V_T + V_s) \right)
\]

\[
= \frac{V_{GS} - V_T}{3t_{ox}} + \frac{V_T + V_s}{3\epsilon_{ox}}
\]

as

\[
C_{ox} = \frac{\epsilon_{ox}}{\epsilon_{st}}
\]

and

\[
\frac{\epsilon_{ox}}{\epsilon_{st}} = \frac{1}{3}
\]

APPENDIX B
DERIVATION OF \( g_m \)

From (5) and in saturation

\[
g_m = \frac{\partial I_D}{\partial V_{GS}} = v_{sat}C_{ox}W \left( 1 - \frac{\partial V_{DSat}}{\partial V_{GS}} \right).
\]

(B1)

From (6)

\[
\frac{\partial V_{DSat}}{\partial V_{GS}} = (1 - K) \left( V_{GS} - V_T \right) \frac{\partial K}{\partial V_{GS}}.
\]

(B2)

But

\[
\frac{\partial K}{\partial V_{GS}} = (1 - K)^2 \left( \frac{1}{E_c L_c} - \frac{a}{E_{c} V_{GS}} + \frac{a}{L_c} \frac{\partial X_d}{\partial V_{GS}} \right).
\]

(B3)

From (2)–(4)

\[
\frac{\partial E_c}{\partial V_{GS}} = -\frac{E_c}{E_d} \frac{\partial E_{r,eff}}{\partial V_{GS}} = \frac{nE_d B_s}{6t_{ox}E_{r,eff}}
\]

where

\[
B_s = \left( \frac{E_{r,eff}/E_c}{1 + (E_{r,eff}/E_c)^n} \right)^n
\]

(B4)

and from (8)

\[
\frac{\partial X_d}{\partial V_{GS}} = -\frac{1}{E_d} \left( \frac{\partial V_{DSat}}{\partial V_{GS}} + \frac{V_{DS} - V_{DSat}}{E_c} \frac{\partial E_c}{\partial V_{GS}} \right).
\]

(B5)

Substituting (B3)–(B5) into (B2) gives

\[
\frac{\partial V_{DSat}}{\partial V_{GS}} = \frac{(1 - K)(1 - K)K}{1 - K^2 \left( E_c/E_d \right)^n} \left[ \frac{V_{GS} - V_T}{E_c} \frac{\partial E_c}{\partial V_{GS}} \left( 1 + \frac{V_{DS} - V_{DSat}}{E_d L_c} \right) \right].
\]

(B6)
The denominator in (B6) is due to the weak interaction between \( X_0 \) and \( V_{dsat} \). Since typically \( K < 1 \) and \( E_c/E_d < 1 \) so that \( K^2 E_c/E_d \ll 1 \), the value of the denominator is close to 1 to the first order of approximation. Its effect is to increase \( \partial V_{dsat}/\partial V_{GS} \) and hence decrease \( g_m \) at high field which is significant only for \( L_{eff} \ll 1 \mu m \) and at high \( V_{GS} - V_T \). For hand-calculation purposes, the following simplifications are made:

\[
g_m = v_{sat} C_{ox} W \left( 1 - \frac{\partial V_{dsat}}{\partial V_{GS}} \right)
\]

\[
= v_{sat} C_{ox} W \left[ 1 - (1 - K) + K(1 - K)(1 - B_m) \right]
\]

\[
= v_{sat} C_{ox} W \left[ K(2 - K) - B_m(1 - K) \right]
\]

where

\[
B_m = \frac{V_{GS} - V_T}{E_c} \frac{\partial E_c}{\partial V_{GS}} \left( 1 + \frac{V_{ds} - V_{dsat}}{E_d L_c} \right) = pB_x.
\]

Empirically, \( p = 1.2 \) for NMOS devices and 1.0 for PMOS devices.

### APPENDIX C

**Derivation of \( g_{ds} \)**

Again, from (5) and (12)

\[
\frac{\partial g_s}{\partial V_{DS}} = v_{sat} C_{ox} W \left( \eta - \frac{\partial V_{dsat}}{\partial V_{DS}} \right).
\]

(C1)

Similar to the derivation procedures in Appendix B, it can be shown that

\[
\frac{\partial V_{dsat}}{\partial V_{DS}} = \frac{\eta(1 - K) - (1 - K) K}{1 - K^2 \left( \frac{E_c}{E_d} \right)}.
\]

(C2)

Thus from (C1) and again neglecting the denominator in (C2)

\[
g_{ds} = v_{sat} C_{ox} W \left( \eta - \frac{\partial V_{dsat}}{\partial V_{DS}} \right)
\]

\[
= v_{sat} C_{ox} W K^2 \left( \frac{E_c}{E_d} \right) + \eta g_m.
\]

Substrate current \( I_{SUB} \) may be approximated by [8]

\[
I_{SUB} = \frac{A_f}{B_f} \left( V_{DS} - V_{dsat} \right) I_{dsat} e^{-B_f/(A_f(V_{GS} - V_{th}))}
\]

where \( A_f \) and \( B_f \) are related to the impact ionization coefficient of carriers \( \alpha \) as follows:

\[
\alpha = A_f e^{-B_f/E_c}.
\]

### REFERENCES


Kai-Yap Toh (M’76–S’80–M’87) was born in Singapore in 1950. He received the B.E. (with Honors) degree from Monash University, Australia, in 1974, the M.Sc. degree from the University of Pennsylvania in 1982, and the Ph.D. degree from the University of California, Berkeley, in 1986, all in electrical engineering.

He served in the Singapore Armed Forces from 1974 to 1976. He was a civilian R&D Engineer with the Defense Ministry of Singapore from 1976 to 1980 and was the head of an R&D Department in 1980. He was a Resident Consultant with the AT&T Bell Laboratories at Holmdel in 1985 working on high-frequency analog MOS circuits. He joined the IBM Research Division at the Thomas J. Watson Research Center, Yorktown Heights, NY, in December 1986 and is working on high-speed BJT circuits. His technical interests are in the design of high-speed analog and digital IC’s, IC technologies, and the characterization of IC devices.

Ping-Keung Ko received the Ph.D. degree in electrical engineering in 1982 from the University of California at Berkeley.

He spent 1982 and 1983 at Bell Laboratories, Holmdel, NJ, working on high-speed MOS technologies. In 1984 he joined the University of California at Berkeley where he is an Associate Professor of Electrical Engineering and Computer Sciences. His current research areas include high-speed VLSI technologies and devices, and MOS device modeling for circuit simulation.

He has authored or co-authored one book and over 70 research journal articles.

Robert G. Meyer (S’64–M’68–SM’74–F’81) was born in Melbourne, Australia, on July 21, 1942. He received the B.E., M.Eng. Sci., and Ph.D. degrees in electrical engineering from the University of Melbourne in 1963, 1965, and 1968, respectively.

In 1968 he was employed as an Assistant Lecturer in Electrical Engineering at the University of Melbourne. Since September 1968 he has been employed in the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, where he is now a Professor. His current research interests are in integrated-circuit design and device fabrication. He has acted as a consultant on electronic circuit design for numerous companies in the electronics industry. He is co-author of the book *Analysis and Design of Analog Integrated Circuits* (Wiley, 1984) and editor of the book *Integrated Circuit Operational Amplifiers* (IEEE Press, 1978).

Dr. Meyer is a past President of the Solid-State Circuits Council of the IEEE and is a former Associate Editor of the IEEE Journal of Solid-State Circuits and of the IEEE Transactions on Circuits and Systems.