in Figure 3.36, again illustrates the importance of including the V_A increase with V_{DS} predicted by Equation 3.74.

The g_{ds} versus V_{DS} behavior seen in Figures 3.36 and 3.38 for pMOS and nMOS devices in the example 0.18 μm CMOS process is similar to that observed for devices in the much larger example 0.5 μm process (both processes are described in Table 3.2). However, comparisons of predicted and measured g_{ds} indicate that V_E must be increased to 0.45 V for the 0.5 μm process from 0.15 V for the 0.18 μm process. This results in an effective scaling of V_{DS} where a 3 V increase in V_{DS} in the 0.5 μm process results in a similar decrease of g_{ds} as a 1 V increase in V_{DS} for the 0.18 μm process. In effect, the decrease of g_{ds} over the full supply voltage range of 5 V and 1.8 V, respectively, is similar for the two processes.

The larger value of V_E for the 0.5 μm process is consistent with a larger value of characteristic length, L_C, where V_E is proportional to L_C in the predictions of g_{ds} given in [14, 28, p. 39]. From Equation 3.68, L_C has predicted values of 27 and 82 nm for the 0.18 and 0.5 μm processes, assuming $t_{ox} = 4.1$ and 13.5 nm, and $x_j = 60$ and 165 nm. If V_E is taken as $V_E = L_C E_{CRIT}$ from [14, 28, p. 39], it then has predicted values of 0.15 and 0.33 V for nMOS devices in the 0.18 and 0.5 μm processes, assuming $E_{CRIT} = 5.6$ and 4 V/μm. Process values are from Table 3.2 except for the drain junction depth, x_j, which is estimated at one-third the minimum channel length. The predicted value of $V_E = 0.15$ V is equal to the extracted value for the 0.18 μm process, while the predicted value of $V_E = 0.33$ V is below the extracted value of 0.45 V for the 0.5 μm process. The values cannot be directly compared since the actual values of x_j are unknown, but do show the trends where L_C and, correspondingly, V_E, decrease in smaller-geometry processes. As mentioned, measurements here indicate that V_E is equal for nMOS and pMOS devices in a given process, whereas $V_E = L_C \cdot E_{CRIT}$ predicts higher values for pMOS values because of higher E_{CRIT}. This suggests that V_E is best found by fitting to measured data.

to as DIBL here and described by a decreasing value of V_T (from the unaffected, large-geometry value) with increasing V_{DS}.

As shown later in measured $V_A = I_D/g_{ds} = I_D \cdot r_{ds}$ data in Section 3.8.4.5, DIBL effects significantly increase g_{ds}, or decrease r_{ds} and V_A, for minimum channel length devices operating in weak inversion. DIBL effects must be carefully considered in all processes, especially for devices having channel lengths near the process minimum.

The change in drain current, ΔI_D, resulting from a change in V_T, ΔV_T, due to DIBL is given by

$$\Delta I_D \text{ (DIBL)} = g_m (-\Delta V_T) \tag{3.76}$$

ΔV_T acts as a negative, small-signal, gate–source voltage since its increase lowers $V_{EFF} = V_{GS} - V_T$, lowering the drain current. Dividing Equation 3.76 by ΔV_{DS} that causes ΔV_T and ΔI_D gives

$$g_{ds} \text{ (DIBL)} = \frac{\Delta I_D \text{ (DIBL)}}{\Delta V_{DS}} = \frac{g_m (-\Delta V_T)}{\Delta V_{DS}} = g_m \left(\frac{-\partial V_T}{\partial V_{DS}} \right) \tag{3.77}$$

where $\delta V_T/\delta V_{DS}$ is the change in threshold voltage with respect to V_{DS}. Since $\delta V_T/\delta V_{DS}$ is negative (V_T decreases as V_{DS} increases), g_{ds} due to DIBL is positive, consistent with the increase in drain current caused by increasing V_{DS}. The Early voltage associated with DIBL is then estimated as

$$V_A \text{ (DIBL)} = \frac{I_D}{g_{ds} \text{ (DIBL)}} = \frac{I_D}{g_m \left(\dfrac{-\partial V_T}{\partial V_{DS}} \right)} = \frac{1}{\left(\dfrac{g_m}{I_D} \right) \dfrac{-\partial V_T}{\partial V_{DS}}} = \frac{V_{gm}}{\dfrac{-\partial V_T}{\partial V_{DS}}} \tag{3.78}$$

where, in the rightmost term, g_m/I_D is alternatively expressed from the transconductance effective voltage, $V_{gm} = (g_m/I_D)^{-1}$, described earlier in Section 3.8.2.1. As mentioned for Equation 3.52 near the beginning of Section 3.8.4, excluding the value of V_{DS} in the g_{ds} or r_{ds} relationship with V_A effectively includes the value of V_{DS} in the value of V_A.

Predicting V_A and the corresponding g_{ds} or r_{ds} due to DIBL appears simple using Equation 3.78 above. However, the prediction is complex because of the difficulty in predicting the change in threshold voltage with drain–source voltage, $\delta V_T/\delta V_{DS}$. This is analogous to predicting V_A due to CLM in Equation 3.59, which also appears simple. In the case of CLM, though, the prediction is complex because of the difficulty in predicting the change in pinch-off length with drain–source voltage, $\delta l_p/\delta V_{DS}$.

One prediction of $\delta V_T/\delta V_{DS}$ is based on a pseudo, two-dimensional solution of Poisson's equation giving [40]

$$\begin{aligned} \Delta V_T \text{ (DIBL)} \approx & -\left[3(\phi_D - 2\phi_F) + V_{DS}\right] e^{-L/L_{DIBL}} \\ & - 2\sqrt{(\phi_D - 2\phi_F)(\phi_D - 2\phi_F + V_{DS})}\, e^{-L/(2L_{DIBL})} \end{aligned} \tag{3.79}$$

ΔV_T is the threshold voltage change for a given value of V_{DS}, ϕ_D (expression given in Table 3.1) is the built-in potential for the drain and channel junction, and $2\phi_F$ is the silicon surface potential taken at the threshold of surface inversion (the boundary of weak and moderate inversion [15, p. 87]); $2\phi_F$ is equal to the process parameter *PHI* that describes twice the Fermi potential. L_{DIBL} is a characteristic length for DIBL given by

$$L_{DIBL} = \lambda_{DIBL} \sqrt{\frac{\varepsilon_{Si}}{C'_{OX}} t_{dep}} = \lambda_{DIBL} \sqrt{\frac{\varepsilon_{Si}}{\varepsilon_{SiO_2}} t_{ox} t_{dep}} \approx \lambda_{DIBL} \sqrt{3 t_{ox} t_{dep}} \tag{3.80}$$

which depends on vertical dimensions and is identical in form to the characteristic length for CLM, L_C, given by Equation 3.68 except that the depletion-region width below the channel, t_{dep} (expression

given in Table 3.1), is used in place of the drain junction depth, x_j, and the fitting parameter λ_{DIBL} is included. λ_{DIBL} has values in the vicinity of unity.

Equation 3.79 predicts that ΔV_T depends negatively on V_{DS} and the square root of V_{DS}, with the square-root relationship dominating for small values of V_{DS} and channel lengths somewhat above L_{DIBL} [40]. The equation also predicts that ΔV_T is proportional to the exponential of $-L/L_{DIBL}$ and the exponential of $-L/(2L_{DIBL})$, as seen in the first term and the second, square-root term. While making the prediction of ΔV_T very dependent on the value of L_{DIBL}, this shows that the magnitude of ΔV_T decreases very rapidly as L increases. ΔV_T is roughly proportional to $1/L^2$ in the MOS model described in [28, p. 36] and proportional to $1/L^3$ in the SPICE Level 3 MOS model [36, pp. 57–58]. All models predict a rapid decrease of ΔV_T with increasing channel length.

A rapid decrease in the magnitude of ΔV_T with increasing channel length is consistent with measured data where the DIBL increase in g_{ds} and the corresponding decrease in V_A are nearly eliminated by increasing channel length modestly from the process minimum. This is shown in Figures 3.43–3.47 and 3.49 in Section 3.8.4.5 where measured V_A for $L = 0.28\,\mu\text{m}$ increases very rapidly towards normal expected values compared to reduced values for the process minimum (L_{min}) of $L = 0.18\,\mu\text{m}$ where DIBL effects are severe. The rapid decrease of ΔV_T is also consistent with increased, simulated, channel surface potential due to DIBL effects that decreases rapidly as channel length increases from the process minimum [40].

In addition to a rapid decrease in the magnitude of ΔV_T with increasing channel length, Equation 3.79 also predicts decreased ΔV_T for high channel doping concentration where t_{dep} and, correspondingly, L_{DIBL} are smaller. The equation also predicts increased ΔV_T for non-zero V_{SB} where t_{dep} and L_{DIBL} are larger. Both trends are observed in practice [15, p. 267]. The equation, however, does not predict the decrease in ΔV_T resulting from decreased drain junction depth, x_j, as this does not appear in the expression [15, p. 267]. In smaller-geometry processes, the depletion region thickness below the channel, t_{dep}, is reduced by retrograde doping where doping is increased vertically below the channel surface. Additionally, the depletion region thickness at the source and drain is reduced by halo doping where doping is increased laterally near the source and drain. This reduces DIBL effects, especially increased doping near the drain. The reduction of L_{DIBL} predicted by Equation 3.80 with decreasing t_{ox} and t_{dep} (t_{dep} decreases because of higher channel doping) and the reduction of x_j suggest that ΔV_T could remain at comparable levels as L_{min} decreases in smaller-geometry processes.

L_{DIBL}, unlike the characteristic length L_C given in Equation 3.68 for CLM predictions, is a function of the inversion level (through the silicon surface potential) and V_{SB} since these influence t_{dep}. As mentioned in Section 3.8.3.2, $t_{dep} \approx 9t_{ox}$ in weak inversion for $n = 1.33$ ($V_{SB} = 0\,\text{V}$), where $n - 1 = \eta = 0.33 \approx 3t_{ox}/t_{dep}$. For the $0.18\,\mu\text{m}$ CMOS process described in Table 3.2 with $t_{ox} = 4.1\,\text{nm}$, this gives $t_{dep} \approx 37$ nm. Using $t_{dep} \approx 37$ nm, $t_{ox} = 4.1\,\text{nm}$, and $\lambda_{DIBL} = 1$, Equation 3.80 predicts $L_{DIBL} \approx 21\,\text{nm} \approx L_{min}/8$ for $L_{min} = 0.18\,\mu\text{m}$ (this neglects the reduction associated with DL, which lowers L_{min} somewhat). When using such a small value of L_{DIBL}, Equation 3.79 predicts that the magnitude of ΔV_T decreases very, very rapidly as channel length increases from L_{min}. While ΔV_T does decrease rapidly as channel length increases from L_{min}, it is unknown if it decreases as rapidly as predicted by $L_{DIBL} \approx L_{min}/8$ using Equation 3.79. Measured data would be required to verify this.

While Equation 3.79 shows that the magnitude of ΔV_T and, correspondingly, $\delta V_T/\delta V_{DS}$ decrease exponentially with increasing channel length with a length constant related to L_{DIBL}, we will use the simple relationship mentioned earlier for some MOS models where $\delta V_T/\delta V_{DS}$ is proportional to $(1/L)$ raised to some power. This avoids the complexity of estimating L_{DIBL} while providing some ability to select fitting parameters that correspond to measured values of $\delta V_T/\delta V_{DS}$. $\delta V_T/\delta V_{DS}$ is then estimated by

$$\frac{\partial V_T}{\partial V_{DS}} \approx DVTDIBL \cdot \left(\frac{L_{min}}{L}\right)^{DVTDIBLEXP} \tag{3.81}$$

where $DVTDIBL = \delta V_T/\delta V_{DS}$ for the minimum channel length, L_{min}, in the process, and *DVTDIBLEXP* is the exponent describing the decrease of $\delta V_T/\delta V_{DS}$ with increasing channel length. Estimated *DVTDIBL* is -12 and $-8\,\text{mV/V}$, respectively, for nMOS devices in the example 0.5 and $0.18\,\mu\text{m}$ CMOS processes listed in Table 3.2. As listed, the magnitude of pMOS values is slightly higher. Estimated *DVTDIBLEXP* is three, corresponding to a $1/L^3$ decrease in $\delta V_T/\delta V_{DS}$ with increasing channel length. While there is uncertainty about the value of $\delta V_T/\delta V_{DS}$ for channel lengths above L_{min}, Equation 3.81 provides a good estimate for the minimum channel length because it uses the extracted value of $\delta V_T/\delta V_{DS}$ here. If non-zero V_{SB} is present, *DVTDIBL* should be higher because of higher t_{dep} as mentioned earlier.

Substituting $\delta V_T/\delta V_{DS}$ from Equation 3.81 into Equation 3.78 gives V_A due to DIBL as

$$V_A\ (\text{DIBL}) = \frac{1}{\left(\frac{g_m}{I_D}\right)\frac{-\partial V_T}{\partial V_{DS}}} = \frac{V_{gm}}{\frac{-\partial V_T}{\partial V_{DS}}} \approx \frac{1}{\left(\frac{g_m}{I_D}\right)\left(-DVTDIBL \cdot \left(\frac{L_{min}}{L}\right)^{DVTDIBLEXP}\right)} \tag{3.82}$$

Substituting g_m/I_D, excluding velocity saturation and VFMR effects, from Table 3.17 then gives

$$V_A\ (\text{DIBL}) = \frac{1}{\left(\frac{g_m}{I_D}\right)\frac{-\partial V_T}{\partial V_{DS}}} = \frac{V_{gm}}{\frac{-\partial V_T}{\partial V_{DS}}} \approx \frac{nU_T\left(\sqrt{IC+0.25}+0.5\right)}{-DVTDIBL \cdot \left(\frac{L_{min}}{L}\right)^{DVTDIBLEXP}} \tag{3.83}$$

This predicts that V_A due to DIBL is equal to the reciprocal of the product of g_m/I_D and $-\delta V_T/\delta V_{DS}$ or V_{gm} divided by $-\delta V_T/\delta V_{DS}$. As mentioned for Equation 3.78, $\delta V_T/\delta V_{DS}$ and, correspondingly, *DVTDIBL* are negative giving the expected positive value for V_A.

Equations 3.82 and 3.83 show that increasing L, where the magnitude of $\delta V_T/\delta V_{DS}$ decreases rapidly, favorably maximizes V_A, minimizing g_{ds}, due to DIBL. Also, operating in strong inversion ($IC > 10$) increases V_A through the reduction in g_m/I_D, where the additional reduction in g_m/I_D caused by velocity saturation and VFMR effects can be included in Equation 3.83 by the *IC* modification described in Table 3.17. Operating short-channel devices in weak inversion gives the worst case DIBL reduction in V_A and increase in $g_{ds} = I_D/V_A$ because of maximum $\delta V_T/\delta V_{DS}$ and g_m/I_D.

Figure 3.39 illustrates g_{ds} increases caused by DIBL. The figure shows drain current curves for a $W/L = 25.2\,\mu\text{m}/0.5\,\mu\text{m}$, nMOS device in the $0.5\,\mu\text{m}$ CMOS process described in Table 3.2. The curves are for operation in weak inversion or the weak-inversion side of moderate inversion where DIBL effects are severe for the minimum channel length device shown. As seen in the figure, the curves turn up exponentially with increasing V_{DS}. This is a result of weak-inversion drain current that is exponentially proportional to $V_{EFF} = V_{GS} - V_T$, where V_{EFF} increases by 12 mV per volt of increasing V_{DS}. This is caused by a 12 mV/V decrease in V_T, given by $DVTDIBL = \delta V_T/\delta V_{DS} = -12\,\text{mV/V}$ listed in Table 3.2 for the minimum channel length device in the process.

The I_D versus V_{DS} characteristics in Figure 3.39 are very steep, giving unattractively high g_{ds} for analog design. Additionally, the increasing slope (g_{ds}) at increasing V_{DS} indicates that DIBL or hot-electron effects (described in Section 3.8.4.3) are present as CLM effects result in decreasing slope and g_{ds} at increasing V_{DS} as seen earlier in Figures 3.35–3.38. Interestingly, a tangent line touching any of the curves in Figure 3.39 at high V_{DS} intercepts the V_{DS} axis at $I_D = 0\,\mu\text{A}$ for positive values of V_{DS}. This corresponds to a negative value of the Early voltage, V_A, for the definition shown in Figure 3.34 and indicates an unusally high value of g_{ds}. DIBL effects are very easy to observe in Figure 3.39 because the large V_{DS} allowed in the $0.5\,\mu\text{m}$ process results in a significant drop in V_T, resulting in a significant increase in drain current.

In Figure 3.39, the identified bias point is at $I_D = 1\,\mu\text{A}$, corresponding to $IC = 0.094$ in weak inversion. This is found from $IC = I_D/[I_0(W/L)]$ from Table 3.6 with a technology current $I_0 = 0.21\,\mu\text{A}$ from Table 3.2. The channel length of $L = 0.5\,\mu\text{m}$ includes the reduction by $DL = 0.1\,\mu\text{m}$ (also listed

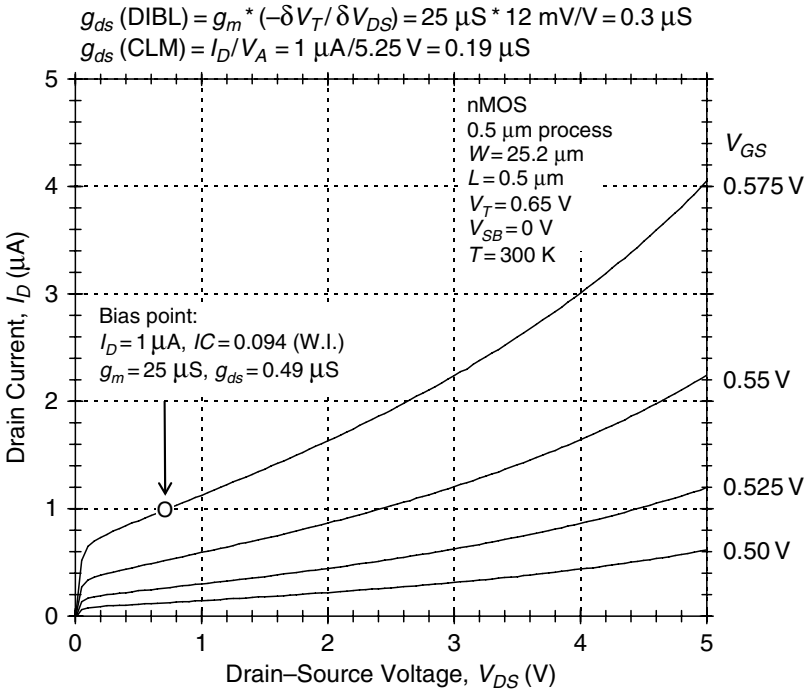

Figure 3.39 Drain current curves for a $W/L = 25.2\,\mu m/0.5\,\mu m$, nMOS device in a $0.5\,\mu m$ CMOS process illustrating increased drain current slope, g_{ds}, at high V_{DS} due to DIBL. The threshold voltage decreases by approximately 12 mV per volt with increasing V_{DS}, causing increased drain current and turn-up in the curves. Increasing channel length from the process minimum significantly reduces DIBL effects, while increasing the inversion level from the weak-inversion operation shown also reduces these effects by lowering g_m/I_D. pMOS DIBL effects are similar

in Table 3.2) from a drawn length of $0.6\,\mu m$. At the bias point, measured g_m is approximately $25\,\mu S$, corresponding to $g_m/I_D = 25\,\mu S/\mu A$ (or per volt). Measured g_{ds}, found from the slope of the tangent line at the bias point, is $0.49\,\mu S$.

For the bias point shown in Figure 3.39, evaluating Equation 3.77 for g_{ds} due to DIBL gives $g_{ds} = g_m \cdot (-\delta V_T/\delta V_{DS}) = 25\,\mu S \cdot (12\,mV/V) = 0.3\,\mu S$. The combined DIBL and CLM effects give a total $g_{ds} = 0.3 + 0.19 = 0.49\,\mu S$, where g_{ds} due to CLM is estimated at $0.19\,\mu S$ as summarized at the top of the figure. As expected by the shape of the drain current curves, DIBL clearly dominates the overall value of g_{ds} and raises it significantly from the value associated with CLM alone.

Figure 3.40 shows drain current curves for $L = 0.18$, 0.28, 0.48, and $4\,\mu m$, nMOS devices in the $0.18\,\mu m$ CMOS process described in Table 3.2. The devices have widths of $W = 3.2$, 4.8, 8, and $64\,\mu m$, respectively, giving nearly equal shape factors of $S = W/L = 16$. This permits their curves to be overlaid since their drain currents are nearly equal. The curves are for operation in weak inversion where DIBL effects are severe for the minimum channel length, $L = 0.18\,\mu m$ device. Unlike the curves shown in Figure 3.39, the $L = 0.18\,\mu m$ curve does not show an exponential turn-up of drain current with increasing V_{DS} because of smaller $\delta V_T/\delta V_{DS} = -8\,mV/V$, compared to $-12\,mV/V$, and reduced V_{DS}. Instead, nearly equal contributions of DIBL and CLM result in nearly constant drain current slope and g_{ds} as V_{DS} increases compared to the typical decrease observed in Figure 3.35 when CLM effects dominate. Figure 3.40 also shows that DIBL increases in slope and g_{ds} are significantly reduced by increasing channel length modestly from the process minimum of $L = 0.18\,\mu m$ to $L = 0.28\,\mu m$. For the $L = 0.28\,\mu m$ and longer-channel devices, CLM effects dominate, giving the typical reduction of slope and g_{ds} as V_{DS} increases. For the bias point identified at $I_D = 1\,\mu A$ for the $L = 0.18\,\mu m$ device, estimated g_{ds} due to DIBL and CLM is $0.2\,\mu S$ and $0.23\,\mu S$, respectively, for a total g_{ds} of $0.43\,\mu S$

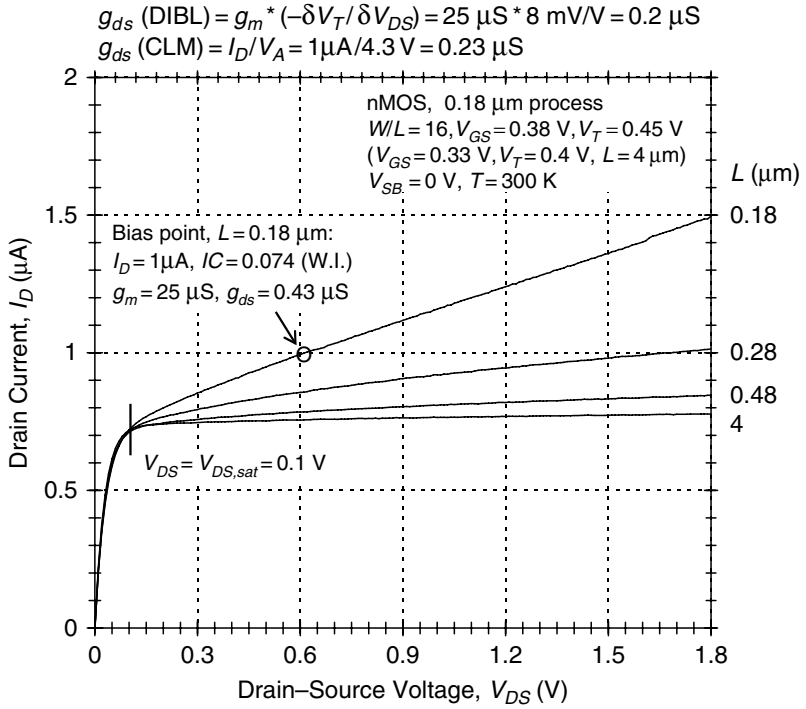

Figure 3.40 Drain current curves for $L = 0.18$, 0.28, 0.48, and 4 µm, nMOS devices in a 0.18 µm CMOS process illustrating increased drain current slope, g_{ds}, due to DIBL for the $L = 0.18$ µm device. The lower threshold-voltage reduction of approximately 8 mV per volt with increasing drain–source voltage and limited drain–source voltage reduce DIBL effects compared to those shown in Figure 3.39. Increasing channel length modestly from the process minimum of $L = 0.18$ µm to $L = 0.28$ µm significantly reduces DIBL contributions to g_{ds} while also lowering CLM contributions. pMOS DIBL effects are similar

as summarized at the top of the figure. The bias point corresponds to operation at $IC = 0.074$ in weak inversion found using the technology current $I_0 = 0.64$ µA and the channel length reduction of $DL = 0.028$ µm from Table 3.2.

While the examples shown are for minimum channel length, nMOS devices, DIBL effects are also significant and even slightly higher for minimum channel length, pMOS devices in the example processes. This can be seen by the slightly higher values of $DVTDIBL = \delta V_T/\delta V_{DS}$ (at the minimum channel length) listed in Table 3.2.

For the processes illustrated here, DIBL effects are nearly non-existent for an $L = 0.5$ µm device in the 0.18 µm process compared to the same channel length, now the minimum channel length device, in the 0.5 µm process. This is typical of processes where short-channel, charge-sharing effects are substantially reduced for channel lengths modestly above the process minimum. In contrast, velocity saturation effects on V_{EFF} (Section 3.7.2.2) and g_m/I_D (Section 3.8.2.2) are similar across processes for equal channel length devices, varying only slightly through the value of E_{CRIT}.

3.8.4.3 Due to hot-electron effects

Hot-electron or impact ionization effects cause an increase in drain current with increasing V_{DS} for nMOS devices, giving rise to another component of MOS drain conductance. When the electric field across the depletion region between the pinched-off channel and drain is sufficiently high, above E_{CRIT} associated with velocity saturation, electrons acquire kinetic energy from lattice collisions that limit