Chapter 5
The Real Transistor

The basic E.K.V. model considered in the previous chapter is not suited for real transistors for it makes use of the "gradual channel" approximation, like the C.S.M. Non-uniform doping, mobility degradation, short channel effects, etc. are ignored. Advanced models like BSIM and PSP, which are primarily circuit simulation tools, take care of these but don't offer the degree of flexibility that is desirable.

We show in this chapter that as long as the source and drain voltages with respect to the substrate remain constant, DC currents, g_m/I_D and g_d/I_D ratios of real transistors, even sub-micron devices, can be reconstructed by means of the basic E.K.V model. Once V_S or V_D is modified, the parameters must be updated. The model remains unchanged however.

5.1 Short Channel Effects

Figure 5.1 shows the constituents of a short channel transistor. The region under the thin oxide in the middle embodies the active region. The rest makes up passive parts. The source and drain consist of narrow n- implanted regions, which run on larger n + diffused regions, themselves connected to the contact regions through silicide layers.

With long channel devices, the proportion of fixed charge below the inversion layer that is controlled by the gate is always much larger than that controlled by the source and drain. Consequently, as the gate length decreases the threshold voltage remains practically constant for the gate-controlled charge varies almost like the gate length. This isn't true with short channel devices. The depleted charge controlled by the gate decreases faster than the gate length owing to the consistent contributions of the source and drain. As a result, the threshold voltage begins to roll-off. Dedicated ion implantations help postponing the effect but generally at the expense of a slight increase of the threshold voltage just before roll-off, an effect called the ***reverse short channel effect***.

Short channel roll-off is not the only issue. The drain voltage influences also the threshold voltage. As the drain voltage increases, the drain takes over a larger share of the depleted region previously controlled by the gate. The threshold voltage

P.G.A. Jespers, *The g_m/I_D Methodology, A Sizing Tool for Low-voltage Analog CMOS Circuits*, Analog Circuits and Signal Processing, DOI 10.1007/978-0-387-47101-3_5,

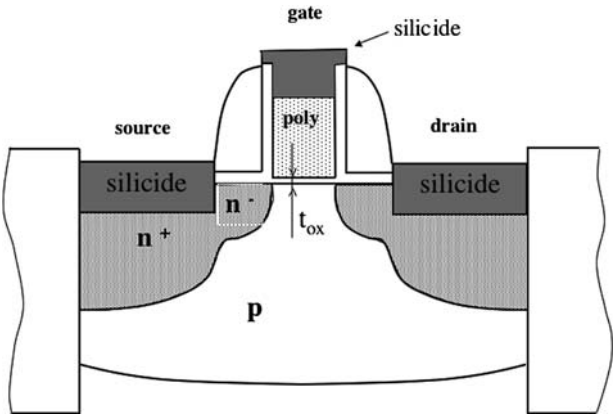

Fig. 5.1 A typical short channel N-MOS transistor

decreases thus, an effect designated currently by the name ***Drain Induced Barrier Lowering*** (**D.I.B.L**.). Contrarily to roll-off, D.I.B.L. is bias dependent and therefore a source of non-linear distortion for it affects V_{To} dynamically.

Other effects plague short channel devices, like mobility degradation caused by high electrical fields. With short channel devices, the electrical field along the channel increases rapidly not only for gate lengths are getting smaller but also because supply voltages often don't scale down at the same rate. The increasing mobility degradation produces a loss of current capability that can be partly compensated by the introduction of lowly doped stripes bridging the channel to the drain region as illustrated in Fig. 5.1. These lessen the impact of the longitudinal electrical field somehow but don't restore the original transconductances.

The model considered in the previous chapter doesn't take any of these effects in consideration. However, real $I_D(V_{GS})$ characteristics look very similar to characteristics predicted by the Charge Sheet and E.K.V.–A.C.M. models. A quasi-exponential region and a more or less quadratic strong inversion region are clearly identifiable. Does this mean that it is possible to reconstruct $I_D(V_G)$ characteristics with the compact model of Chapter 4 even with short channel devices? As long as the source-to-substrate and the drain-to-substrate voltages don't change, the answer may be yes. The spatial distributions of the charge in the inversion layer and in the depleted region underneath may explain this. The thickness of the inversion layer is small compared to the gate length (even down to 100 nm gate lengths). To push a little, what happens in the inversion layer boils down to a 1D problem while in the depleted region beneath, things are different. As 'long channel' conditions prevail more or less in the inversion layer, the spatial distribution of the electric field in the depleted region conforms to a 2D problem owing to the large source and drain contributions. This may be the reason why $I_D(V_G)$ characteristics of submicron transistors can be modeled reasonably well with the compact model as long as the source and drain voltages don't change, which implies that all the parameters must be updated as soon as one of these is modified.

5.2 Checking the Validity of the Compact Model when its Parameters vary with the Source and Drain Voltages

The sample displayed in Fig. 5.2[1] represents drain currents of a 10 µm wide N-channel MOS transistor whose drain-to-source voltage is stepped from 0.2 to 1.2 V, considering two source voltages V_S (0 and 0.8 V) and two gate lengths (0.1 and 1 µm). Although distinct, all curves are similar.

The question is: can we reconstruct each of these characteristics dependably by means of the compact model of the previous chapter taking advantage of parameters that depend on the source and drain voltages? To answer the question, we must compare drain currents predicted by the model to real $I_D(V_{GS})$ characteristics. An identification algorithm is needed up therefore.

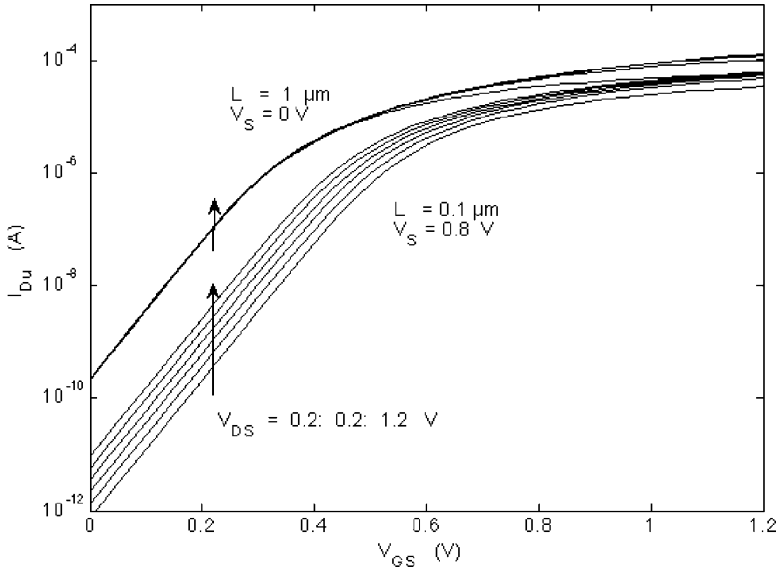

Fig. 5.2 Drain currents of an N-channel unary transistor (W/L = 1) considering two gate length 0.1 and 1 µm, various drain-to-source and back-bias voltages. The device belongs to a low-power, low-voltage 90 nm technology developed by IMEC (Courtesy of IMEC)

[1] The currents shown in this figure are reconstructed drain currents obtained by means of the PSP compact MOSFET model. The parameters were extracted from measurements carried out on real physical transistors (courtesy of IMEC). The assumption that reconstructed currents agree fairly well with the physical currents is accepted implicitly. The PSP compact MOSFET model is a product of Philips Semiconductors and Penn State University (now respectively NXP and Arizona State University) (PSP 2006).

5.2.1 E.K.V Parameter Identification (MATLAB IdentifDemo.m)

The identification algorithm[2] makes use of the E.K.V equations introduced in the previous chapter. These are divided in two groups, general equations:

$$V_P - V_S = U_T\,(2\,(q_F - 1) + \log\,(q_F)) \tag{5.1}$$

$$V_P - V_D = U_T\,(2\,(q_R - 1) + \log\,(q_R)) \tag{5.2}$$

$$i = q_F^2 + q_F - q_R^2 - q_R \tag{5.3}$$

and equations involving the parameters n, V_{T0} and I_S:

$$V_P = \frac{V_G - V_{TO}}{n} \tag{5.4}$$

$$I_D = i I_S \tag{5.5}$$

Before we review the acquisition algorithm, three preliminary remarks ought to be made. The first concerns the transistor configuration for data acquisition. The algorithm described in the previous chapter cannot be used for the acquisition method makes use of the common-gate configuration, violating thus the conditions formulated above regarding constant source and drain voltages. The parameters must be extracted from $I_D(V_{GS})$ characteristics exclusively.

The second remark concerns the reference terminal when carrying out measurements. The reference terminal is generally the source of the transistor while the substrate is back-biased. This requires to rewrite the equations above accordingly. The pinch-off voltage of Eq. 5.1 becomes V_{PS}, the left part of Eq. 5.2 is replaced by $V_{PS} - V_{DS}$ while the expression below is substituted to Eq. 5.4. Notice that the threshold voltage V_{To} (with a lower case zero) below is defined also with respect to the source.

$$V_{PS} = \frac{V_{GS} + V_{To}}{n} \tag{5.6}$$

The third remark concerns geometry. All the 'experimental' drain currents are divided by W/L prior to identification. We consider only ***unary drain currents*** $I_{Du}(V_{GS})$ and ***unary specific currents***.

Let us consider now the acquisition algorithm. The two parameters that are identified first are the slope factor and the threshold voltage. Both emanate from the derivative of the $\log(I_D(V_{GS}))$ characteristics, in other words from g_m/I_D. The slope factor is derived from the maximum of g_m/I_D as usual, while the threshold

[2] The identification algorithm can be found in the 0start directory under IdentN.m and IdentP.m. The algorithm makes use of the 'semi-empirical' N- and P-channel data listed under n90.mat and p90.mat. The compact model parameters outputted by the identification algorithm are stored under ParamN.mat and ParamP.mat,. These are turned into ***global variables*** when running Glob.m (see also Annex 1).

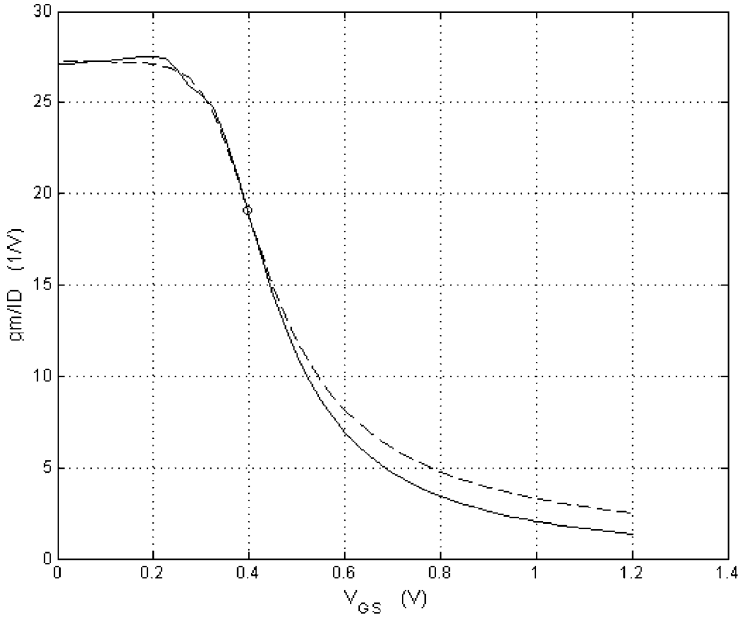

Fig. 5.3 Reconstructed and original g_m/I_D's (respectively dashed and plain lines) of a grounded source N-channel transistor whose L is equal to 100 nm and V_{DS} equal to 0.6 V. The circle corresponds to R equal to 0.7. The MATLAB IdentifDemo.m file illustrates the identification mechanism when the variable M on top of the data list is made equal to one[3]

voltage is the result of a fitting procedure illustrated by Fig. 5.3. The idea is to search the V_{To} that forces the g_m/I_D predicted by the E.K.V model to pass through a predefined point $(g_m/I_D)_o$ of the 'experimental' g_m/I_D curve supposed to lie in moderate inversion, say 80% to 50% below the maximum of g_m/I_D. The acquisition starts with the evaluation of the normalized mobile charge density q_{Fo}, which is derived from the ratio $(g_m/I_D)_o$ over the maximum g_m/I_D, called R, and Eq. 4.34:

$$R = n\,U_T \left(\frac{g_m}{I_D}\right)_o = \frac{1}{1+q_{Fo}} \tag{5.7}$$

Knowing q_{Fo}, we evaluate the pinch-off voltage V_{PSo} by means of Eq. 5.1:

$$V_{PSo} = U_T\,(2\,(q_{Fo}-1) + \log\,(q_{Fo})) \tag{5.8}$$

This allows extracting the threshold voltage from the expression below derived from Eq. 5.6, where V_{GSo} represents the gate-to-source voltage at the selected coincidence point:

$$V_{To} = -nV_{PSo} + V_{GSo} \tag{5.9}$$

[3] The MATLAB file IdentifDemo.m illustrates dynamically the evolution of Fig. 5.3 when the drain voltage symbolized by a vertical landmark is swept from 0 to 1.2 V.

The extraction method is fairly reliable for changes of R by 5–10 cent do not affect V_{To} by more than 1–2 mV.

Now that n and V_{To} are known, we identify the unary specific current. The evaluation is straightforward. All what is needed therefore indeed is to divide the drain current I_{Duo} (the drain current at the point considered for the evaluation of V_{To}) by the normalized drain current i_o, which is know since q_{Fo} has been assessed already.

In a nutshell, the slope factor is extracted from the subthreshold drain current characteristic, the threshold voltage from the progressive bending of the drain current in moderate inversion and the specific current from the drain currents coincidence.

One may argue that Eq. 5.7 supposes that the transistor be saturated, which may not be the case. To take care of non-saturation, the expression below, demonstrated further under Eq. 5.20, must substituted to Eq. 5.7:

$$R = \frac{1}{1 + q_{Fo} + q_{Ro}} \tag{5.10}$$

The introduction of q_{Ro} requires however having at one's disposal an additional expression linking q_{Ro} to V_{DS}. To get this equation, we subtract Eq. 5.2 from Eq. 5.1:

$$V_{DS} = U_T \left(2 \left(q_{Fo} - q_{Ro} \right) + \log \left(\frac{q_{Fo}}{q_{Ro}} \right) \right) \tag{5.11}$$

Equations 5.10 and 5.11 form a system of non-linear implicit equations that can be solved by means of MATLAB interpolation instructions. All what is needed therefore is to generate a ***logspace*** vector q_R encompassing all possible reverse normalized mobile charge densities and extract the corresponding forward q_F's from Eq. 5.10. One makes then use of Eq. 5.11 to find the concomitant drain to source voltage vector U_{DS}. The q_{Fo} to be put in Eq. 5.8 is found by running the MATLAB interpolation instruction below making use of the U_{DS} and q_F vectors. Notice that the problem requires to be solved only once.

$$q_{Fo} = \textbf{\textit{interp1}} \left(U_{DS}, q_F, V_{DS}, \text{'cubic'} \right) \tag{5.12}$$

The reconstructed g_m/I_D curve (represented by the dashed line in Fig. 5.3) calls for a few comments. The 'experimental' and reconstructed curves coincide of course at the reference point. Differences appear else. In weak inversion, the model operates like a filter wiping out a number of local disparities that may be inherent to the 'experimental' data or reflect physical effects like side currents or a shift of the drain current in volume (with P-channel transistors namely). The fact that these differences are smeared out does not represent a problem per see but raises the question as how to define the maximum of g_m/I_D and, more specifically, what is the impact of small variations of the slope factor n on the final threshold voltage V_{To}? The answer is little, for small variations of n are synonymous of small variations of R and

the threshold voltage does not depend much on R. The departure in strong inversion between original and reconstructed g_m/I_D curves is more serious. This difference is discussed in the next section.

A final remark concerns the source voltage of the reverse transistor. Since it is not the same as that of the original transistor, results may be questionable. Yet, examples show that Eq. 5.12 yields generally more consistent results than Eq. 5.7.

5.2.2 How to Introduce Mobility Degradation?

After g_m/I_D, let us reconstruct the drain current. The result is shown in Fig. 5.4. In weak and moderate inversion the 'experimental' and model-driven curves coincide practically, but diverge substantially in strong inversion. The reason is that we didn't take mobility degradation into consideration. In the acquisition method described in the previous section, the mobility factor μ that appears in the specific current expression of Eq. 4.19 is evaluated at the reference point, in other words in moderate inversion. It is supposed not to vary. Mobility degradation can be modeled however by making μ a function of the electrical field like in Chapter 2.

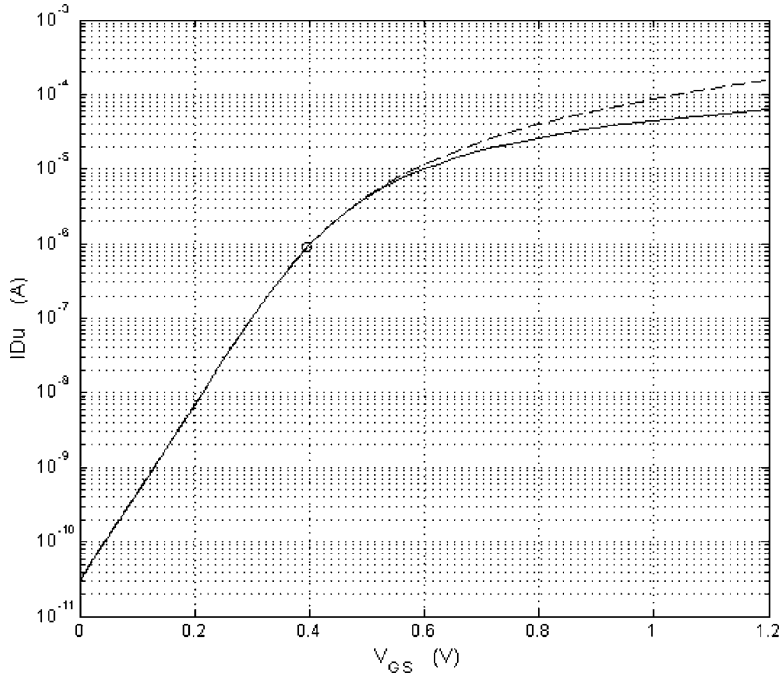

Fig. 5.4 Representation of the real drain current (*plain lines*) and model-reconstructed current (*dashed lines*) making use of the three parameters identified so far. The circle refers to the point selected for the threshold identification (see curve M = 2 of the MATLAB IdentifDemo.m file)

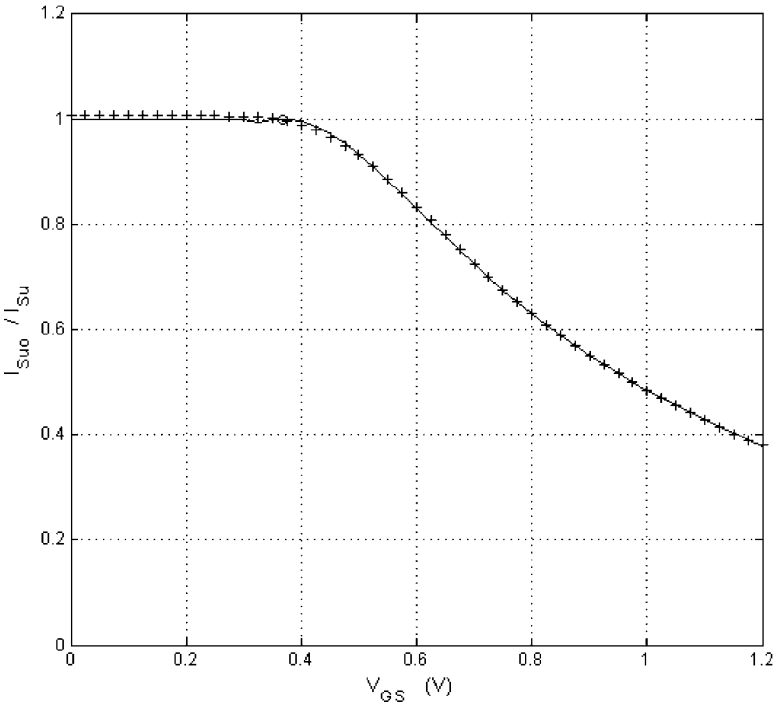

Fig. 5.5 The impact of mobility degradation in strong inversion is illustrated by the roll-off of the plain line curve representing the ratio of the constant weak inversion specific current I_{Suo} over the effective specific current I_{Su}. Crosses represent the approximation based on the theta polynomial $\theta(i)$. The gate length is 100 nm, the source is grounded and the drain voltage equal to 0.6 V like in the two previous figures (see curve M = 3 of the MATLAB IdentifDemo.m file)

The continuous curve of Fig. 5.5 shows the ratio of the 'semi-empirical' over model-reconstructed current represented by means of a dashed line in Fig. 5.4. Two regions are clearly identifiable. Left, the ratio is more or less constant and equal to one for the reconstructed and 'experimental' drain currents coincide practically. Above 0.4 V, mobility degradation is taking over steadily. One can model the trend by turning the unary specific current into a variable. We can define I_{Su} for instance as the product of the constant specific current I_{Suo} of the previous section times a function that rolls-off progressively in strong inversion. In weak inversion, the specific current I_{Su} boils down to the constant ***weak inversion unary specific current*** I_{Suo}. Else, mobility degradation is acknowledged by dividing μ by a polynomial function (like in Eq. 4.56). Generally, the polynomial is expanded versus V_{GS} and V_{DS}. We take a different approach. We expand the polynomial versus the normalized drain current instead of the gate and drain voltages. The idea is to coalesce the effects of the gate, drain and source voltages by means of the sole normalized drain current. The plot illustrated by crosses in Fig. 5.5, which makes use of a fourth order polynomial fit $\theta(i)$, shows that this is feasible. Eventually a third or even second order polynomials can be put to use.

5.2.3 Drain Current Reconstruction

In the previous section, we described the acquisition of the parameters, n, V_{To}, I_{Suo} and the fitting polynomial $theta(i)$.[4] Figure 5.6 compares drain currents predicted by the model (dots) to the data shown in Fig. 5.2. The two match reasonably well. Relative errors are of the order of 1–2% in moderate and strong inversion. In weak inversion, they attain 4–8% because the model doesn't take into consideration the gradual decline of the g_m/I_D ratio mentioned before. The errors with P-channel transistors are generally larger reaching eventually 10–15% over 8 decades drain current. The reason is probably due to the different nature of the inversion layer, which may be deeper in the substrate than with N-channel transistors.

A 3D representation comparing experimental and reconstructed drain currents versus the drain and gate voltages is represented in Fig. 5.7.

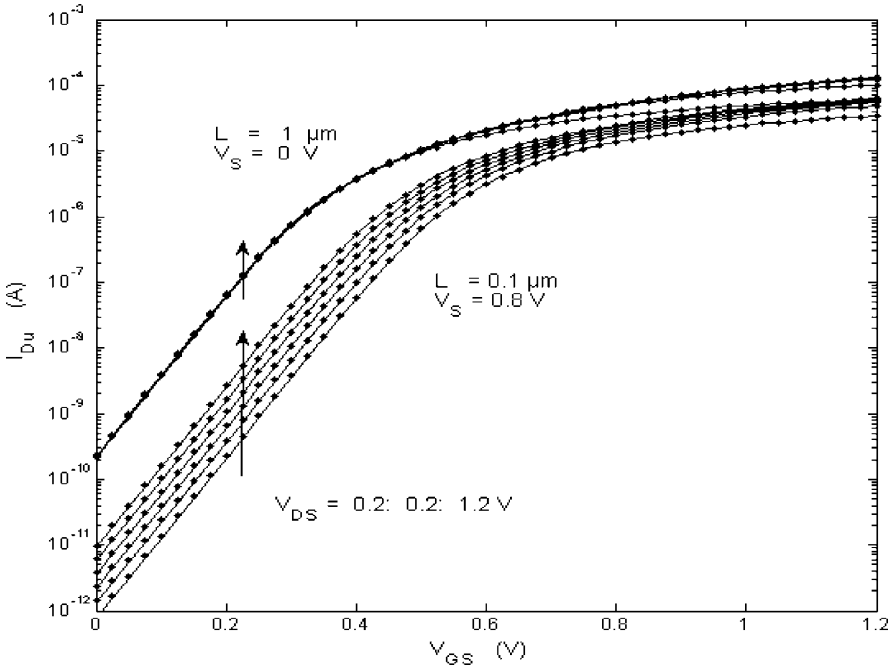

Fig. 5.6 This figure compares reconstructed drain currents (*dots*) to the currents of Fig. 5.2 (*plain lines*) where from the E.K.V. parameters were evaluated by means of the identification algorithm described in the previous section

[4] The MATLAB IdentifN.m and IdentifP.m files implementing the acquisition algorithm can be found in the Glob directory together with the 'semi-empirical' data where from the compact model parameters are extracted. It is possible to retrieve the extraction algorithm with other 'experimental' data when desired. To get familiar with the data organization, please consult Annex 1.

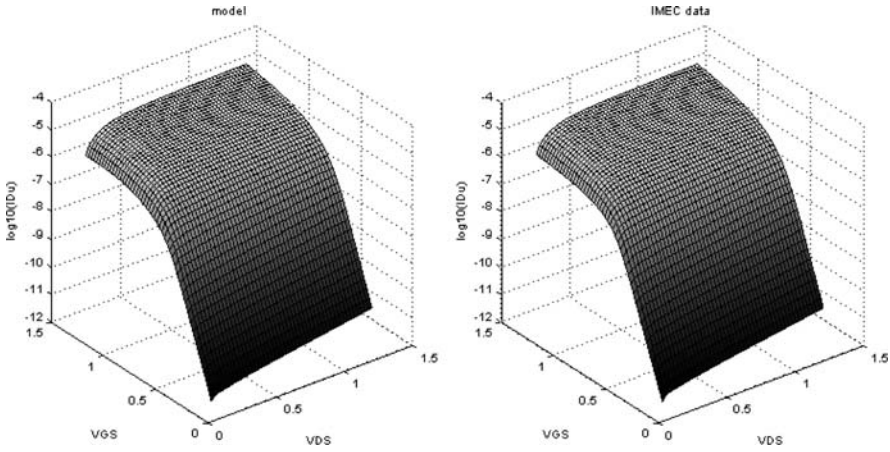

Fig. 5.7 3D representations of the drain current of the 100 nm N-channel grounded source transistor. The reconstructed model is shown *left*, the original *right*. The gate voltage varies from 0 to 1.2 V and the drain from 0.15 to 1.2V

5.3 Compact Model Parameters Versus Bias and Gate Length

The fact that the model reproduces characteristics of short channel devices with few parameters opens interesting prospects. Measurements carried out on large numbers of 'identical' transistors pave the road towards sensitivity analyses by assessing mismatches affecting the slope factor, threshold voltage and specific current. The impact of the temperature can be transposed likewise in terms of parameters sensitivities (see Annex 3). Small modifications of the terminal voltages can be expressed in terms of parameter modifications moreover, which allow evaluating small signal parameters. In a nutshell, the possibility to scrutinize the dependence of the parameters on the gate length and bias conditions opens interesting investigation fields. A few examples are reviewed hereafter.

5.3.1 The Influence of the Gate Length on the Model Parameters

The gate length brings to the fore a number of well-known effects, such as threshold voltage roll-off, reverse short channel effect, D.I.B.L. and C.M.L.

The plot of Fig. 5.8 illustrates the impact of the gate length on the slope factors of N- and P-channel transistors. The slope factors tend to increase when the gate length is shrinking. The effect is more pronounced for P- than for N-channel devices owing to their distinct structure. The drain voltage has very little effect on the slope factor.

The curves of Fig. 5.9 illustrate the influence of the gate length on V_{To}. The threshold voltage of long channel devices does not depend practically on the gate length nor the drain voltage, whether N or P channel transistors are considered.

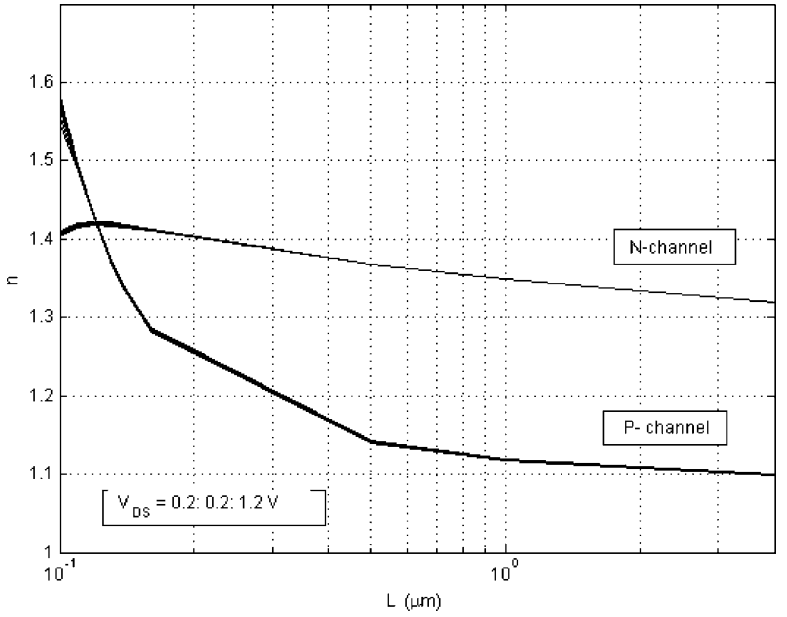

Fig. 5.8 Plot of the subthreshold slope n versus the gate length of grounded source N and P channel transistors for six equally spaced drain voltages (MATLAB SlopeFact1.m)

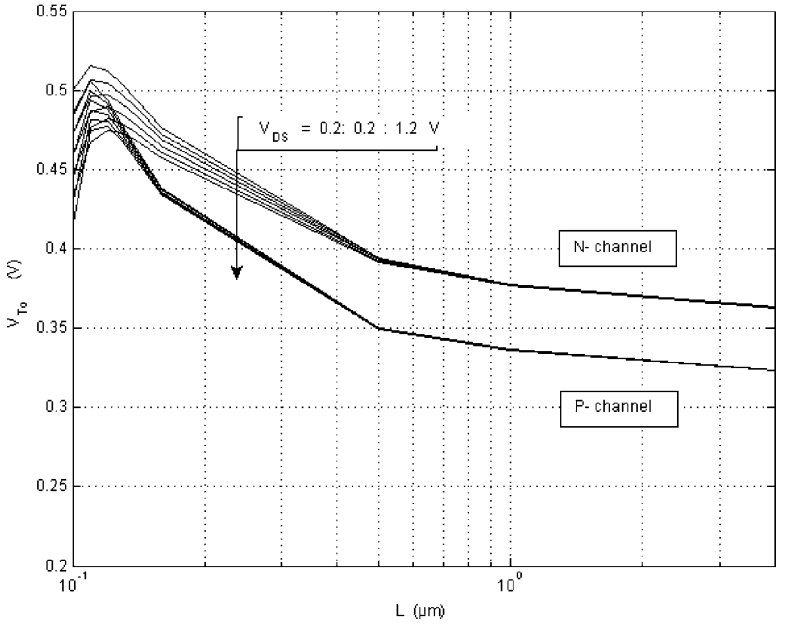

Fig. 5.9 Plot of the threshold voltage V_{To} versus the gate length of grounded source N and P channel transistors considering six equally spaced drain voltages comprised between 0.2 and 1.2 V (MATLAB ThresVolt1.m)

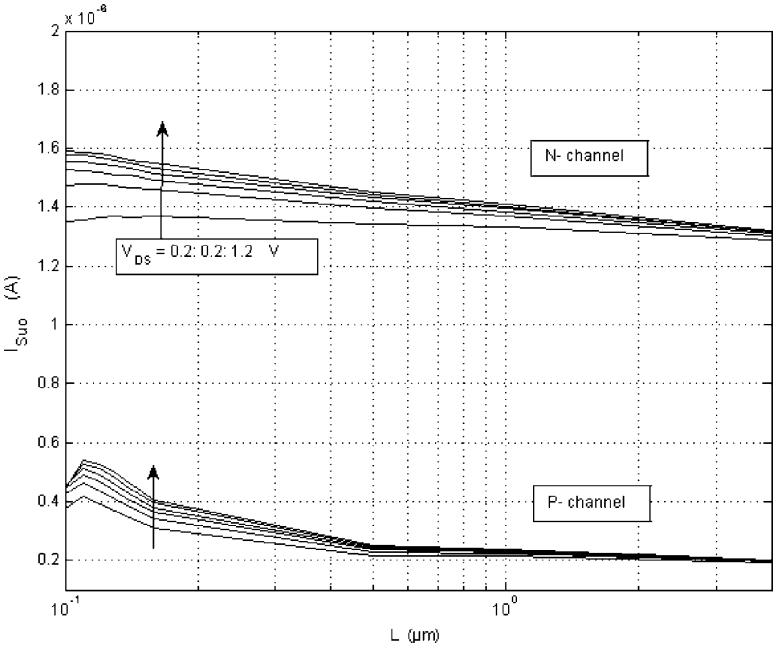

Fig. 5.10 Plot of the weak inversion specific current I_{suo} versus the gate length of the grounded source N and P channel transistors considering six equally spaced drain voltages comprised between 0.2 and 1.2 V (MATLAB SpecCur1.m)

Below 1 µm, the threshold voltage starts to increase progressively until a rapid roll-off occurs at short gate lengths. The global increase, called the ***reverse short channel effect***, reflects the actions taken during fabrication in order to postpone roll-off. The rise contrasts sharply with the abrupt ***roll-off*** due to the source and the drain depleted regions taking over a larger share of the gate-controlled depleted region. It shows that one is getting close to the minimum achievable gate length.

The data displayed in Fig. 5.10 illustrate the influence of the gate length on I_{Suo}. Though the W/L ratio of unary transistors is constant and equal to one, unary specific currents increase slightly when the drain voltage increases. The widening depleted region near the drain is shortening indeed the effective gate length. As a result, I_{Suo} tends to increase. The effect is commonly designated by the acronym C.L.M for ***Channel Length Modulation***.

5.3.2 The Influence of Bias Conditions on the Parameters

The next figures illustrate the influence of the drain-to-source and source-to-substrate voltages. The impact of the drain-to-source voltage on the slope factor n is relatively small and can be ignored as shown already in Fig. 5.8. The influence

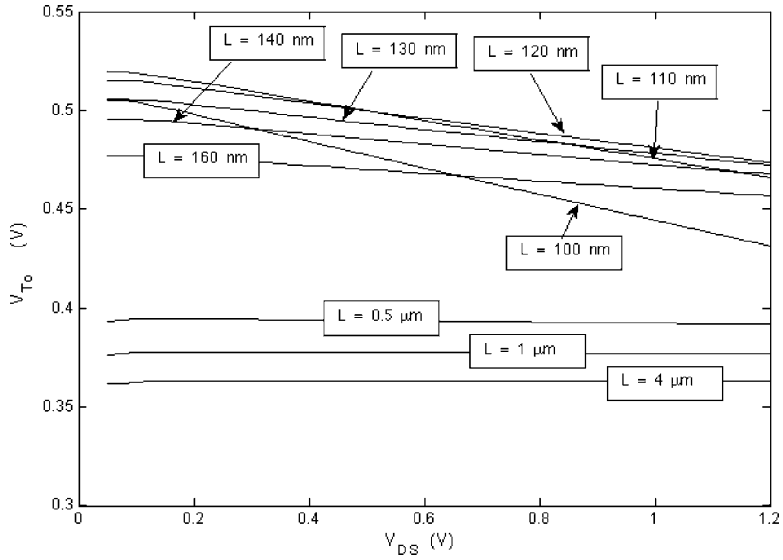

Fig. 5.11 The threshold voltage of the N-channel transistor exhibits almost a linear dependence on the drain-to-source voltage over a wide range, whatsoever the gate length (V_S is equal to zero) (MATLAB ThresVolt2.m)

of V_{DS} on the threshold voltage is more central as shown in Fig. 5.11. As V_{DS} increases, the drain takes over a larger share of the control exercised by the gate, especially when the gate length is small. This lowers the potential barrier carriers must overcome to reach the drain. As a result, the threshold voltage decreases. The effect is designated by the acronym D.I.B.L ***for Drain Induced Barrier Lowering***. It affects strongly the derivative $\mathrm{d}V_{To}/\mathrm{d}V_{DS}$, called the sensitivity factor S_{VTo}, that characterizes the quasi-linear evolution of V_{To}. S_{VTo} is of the order of $-0.13\,\mathrm{mV/V}$ with the $4\,\mu\mathrm{m}$ transistor but reaches $-66\,\mathrm{mV/V}$ and $-85\,\mathrm{mV/V}$ with the 100 nm transistor considering back-bias voltages respectively equal to 0 and 0.8 V. While negligible when L is larger than $2\,\mu\mathrm{m}$, D.I.B.L plays a major role with submicron transistors.

A number of other effects are visible also in the same figure. The global shift upwards with shorter gate lengths illustrates the reverse short channel effect mentioned in connection with Fig. 5.9. Threshold voltages grow until the trend changes once roll-off starts to take place below 130 nm. The P-channel transistor is a little less sensitive to D.I.B.L. Its S_{VTo} is equal to $-34\,\mathrm{mV/V}$ for 100 nm transistors and vanishes faster than with N-channel transistors.

Figure 5.12 displays the influence of back-bias on the threshold voltage of the 100 nm N- and P-channel transistors considering several drain-to-source voltages. The systematic increase of the threshold voltage is an illustration of the well-known ***body effect***. It is more pronounced for the N- than for the P-channel transistors.

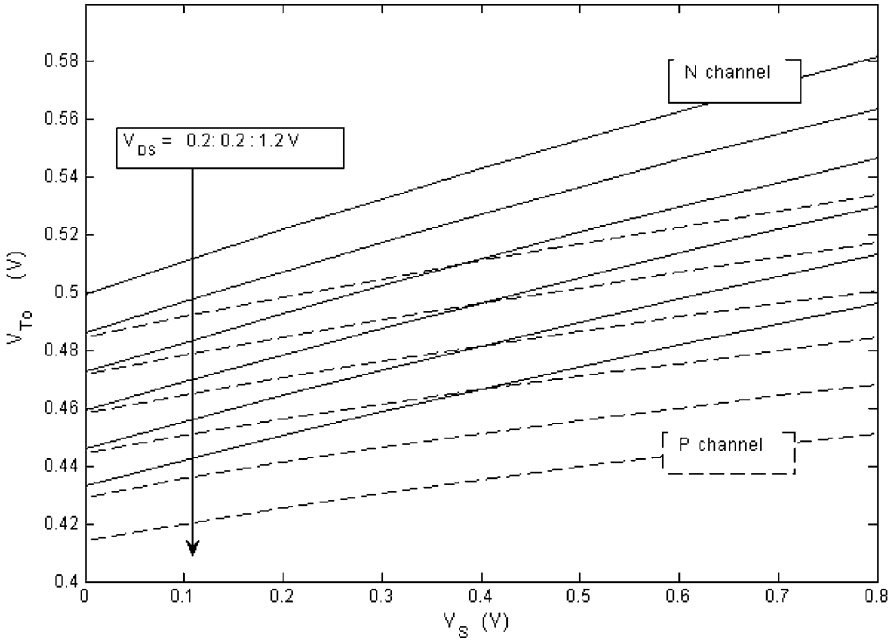

Fig. 5.12 Plot of the threshold voltage V_{To} versus the source-to-substrate voltage V_S considering N- and P-channel transistors with a gate length of 100 nm and six equally spaced drain-to-source voltages (MATLAB ThresVolt3.m)

The impact V_{DS} and V_{GS} have on the specific current I_{Su} is trickier but exemplifies several interesting effects. The overall decline of I_{Su} that is clearly visible in the 3D representation of Fig. 5.13 when the gate-to-source voltage trespasses 0.4 V reflects the growing mobility degradation caused by the electrical field. Below 0.4 V, V_{GS} has little effect but the impact of the drain-to-source voltage is subtler. When V_{DS} decreases, the longitudinal field lessens so that the specific current should be increasing instead of decreasing. Mobility degradation is not the only item to consider however for the specific current depends also on C.L.M. When the drain voltage decreases, the channel length increases slightly owing to the lessening W/L ratio. Mobility and C.L.M impact the specific current in opposite directions thus. The first tends to decrease, the second to increase I_{Su}. According to Fig. 5.13, C.M.L overwhelms mobility degradation in weak and moderate inversion. In strong inversion, the explanation is a bit trickier and requires separating more clearly the impact of D.I.B.L and C.M.L. Fig. 5.14 proposes an interpretation.

The figure represents the unary specific current I_{Su} divided by I_{Suo}, in other words the reciprocal of '*theta*' function. Dividing the specific current by I_{Suo} eludes C.L.M. The fact that the ratio remains practically equal to one in weak and partly in moderate inversion whichever V_{DS} supports the idea. When the gate-to-source voltage trespasses 0.4–0.5 V and mobility degradation starts to grow, we observe a smooth lift up in the non-saturated region. In this region, the longitudinal electrical field

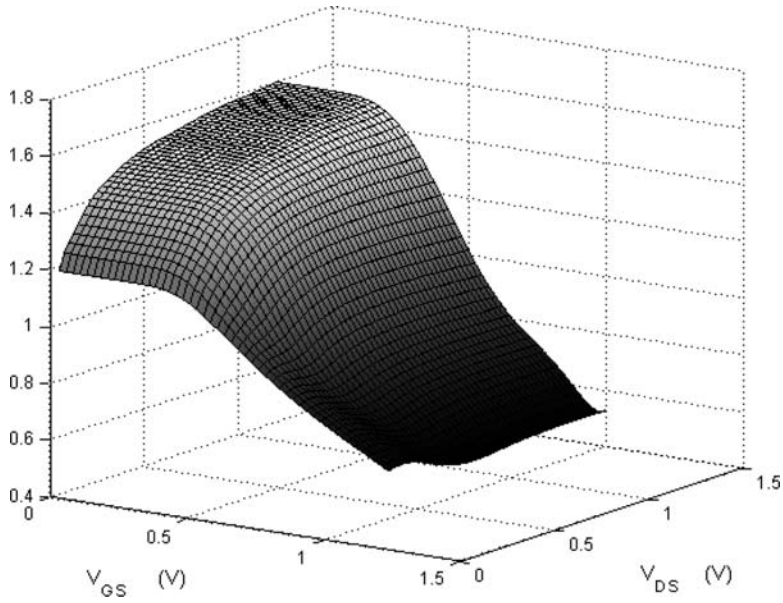

Fig. 5.13 Illustration of the dependence of I_{Su} on the gate and drain voltages for the grounded-source N-channel 100 nm transistor (MATLAB SpecCur2.m)

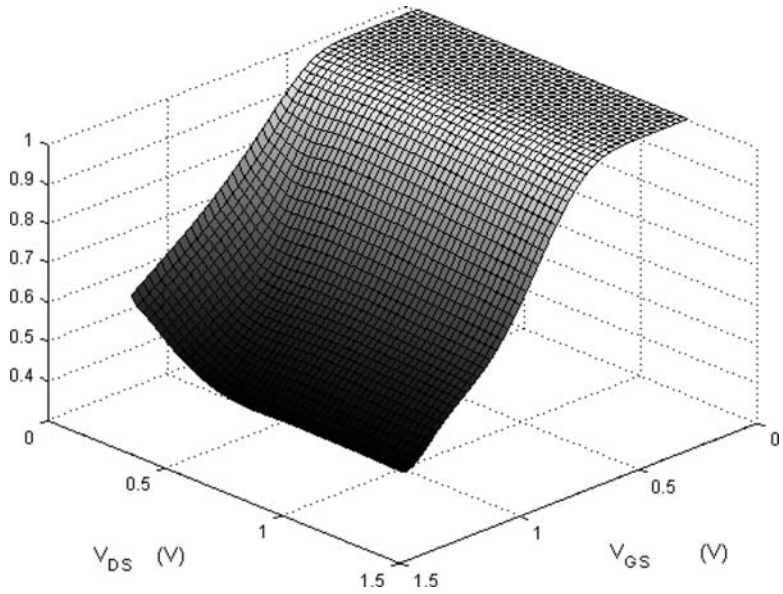

Fig. 5.14 3D representation of the reciprocal of the theta polynomial considering the 100 nm N-channel transistor with zero back-bias (MATLAB SpecCur2.m)

is lessening, mobility degradation decreases thus. Ultimately, when V_{DS} is equal to zero, only the vertical electrical field remains. It looks like if the '*theta*' function discriminates the contributions of the vertical and longitudinal electrical fields. The interpretation of Fig. 5.13 is more intricate for mobility degradation and C.L.M combine their effects with non-saturation.

5.4 Reconstructing $I_D(V_{DS})$ Characteristic

The crucial role played by bias dependent parameters is clearly illustrated when we reconstruct $I_D(V_{DS})$ characteristics. We proceed like in the previous chapter. The specific current is multiplied by the normalized drain current, which requires to know the normalized mobile charge densities q_F and q_R, themselves derived from the applied voltages and the pinch-off voltage. But, contrarily to what happens in the Charge Sheet model where the drain current remains practically constant in saturation, in the compact model the current varies for all the parameters vary with V_{DS}. With short channel devices, the forward mobile carrier density increases with V_{DS} for the threshold voltage decreases owing to D.I.B.L, whereas in long channel devices the gate length decreases owing to C.L.M.

Figures 5.15 and 5.16 compare reconstructed drain currents (represented by means of crosses) to original (continuous) drain currents considering the same

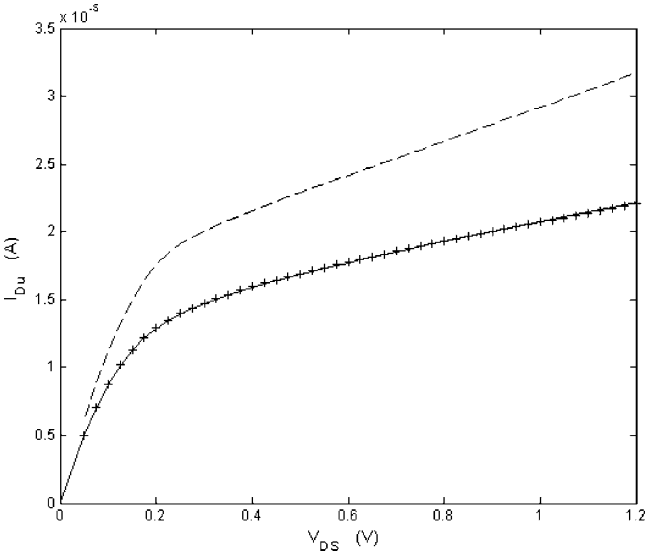

Fig. 5.15 Comparison of the drain currents of the 100 nm grounded source N- channel transistor when V_{GS} is equal to 0.70 V. The *plain line curve* represents the 'experimental' data, crosses illustrate the predicted drain current. The *dashed lines* relate to the model when the mobility is supposed to be invariant (MATLAB fig515.m)

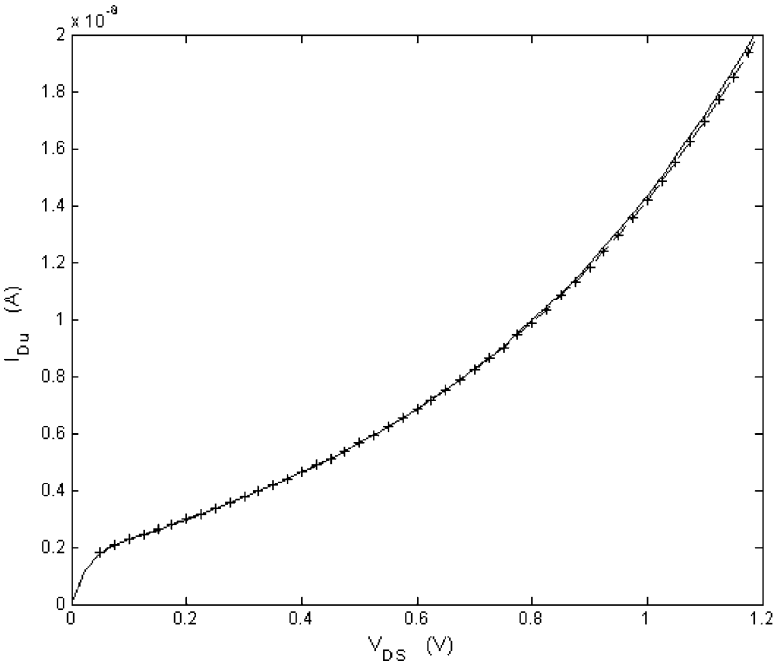

Fig. 5.16 Comparison of the drain currents of the 100 nm grounded source N-channel transistor when V_{GS} is equal to 0.20 V (weak inversion). The *plain line curve* represents the 'experimental' data, crosses the predicted drain current. There is no difference between the *dashed* characteristic (constant mobility) and the current predicted by the model (MATLAB fig515.m)

100 nm N-channel grounded source transistor as above. Two distinct gate voltages, are contemplated, respectively 0.70 and 0.20 V. With the first, the transistor is in strong inversion, with the second it is in weak inversion. Notice that the dashed curve in the first figure represents the drain current without mobility degradation. The curve lies definitely above the actual drain current while in the second figure the curves coincide for mobility does not take place.

Predicted and 'experimental' drain currents differ by less than a few per-cent. The large dissimilarity between the two figures calls for an explanation. In strong inversion, the saturated drain current increases steadily whereas in weak inversion the current displays a quasi-exponential behavior. Avalanche breakdown is not the reason of course for the drain voltage is too low. The explanation is related to the impact of the threshold voltage on the pinch-off voltage. The mechanism is illustrated by means of Fig. 5.17, which takes advantage of the graphical construction introduced in Chapter 3. Left, we consider a large gate voltage so that strong inversion prevails. Right, the opposite holds true. Hatched areas represent the drain currents divided by beta as explained in Chapter 3. Since the drain voltages and gate lengths are identical in the two figures, the impact of the drain voltage on the threshold voltages is the same. Increasing V_{DS} shifts V_{To} downwards as illustrated by

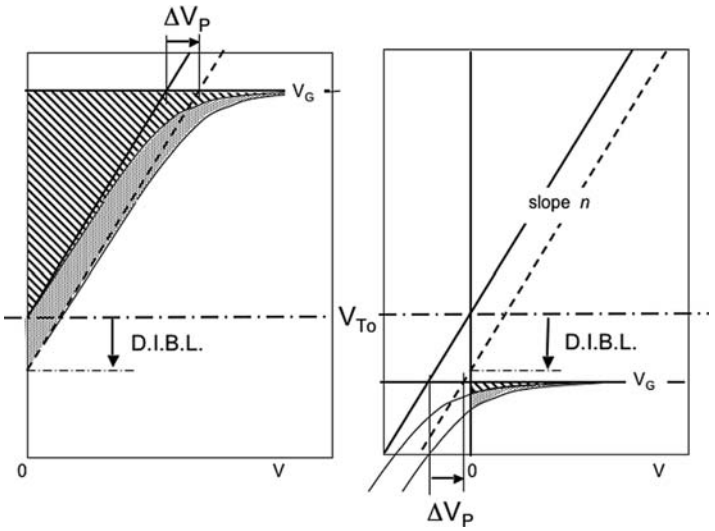

Fig. 5.17 The impact of DIBL on the drain current in strong (*left*) and weak (*right*) inversion is illustrated by means of the graphical construction presented in Chapter 3. The current increases almost linearly *left*, exponentially *right*

the two thick equal lengths arrows visible in both figures. Grey areas represent the concomitant increases of the drain currents. In strong inversion, the current grows almost linearly. In weak inversion, though the current is small, the relative increase is much larger because currents encompass a region where V_T varies exponentially.

5.5 Evaluation of g_x/I_D Ratios

The g_m/I_D and g_d/I_D ratios require to evaluate the derivatives of $\log(I_{Du})$ with respect to V_{GS} and V_{DS}. Both derivatives can be derived from the general expression:

$$\frac{g_x}{I_D} = \frac{d}{dV_x}\log\left(I_{Du}\right) = \frac{d}{dV_x}\log\left(i\right) + \frac{d}{dV_x}\log\left(I_{Su}\right) \tag{5.13}$$

where:

$$\log\left(I_{Su}\right) = \log\left(I_{Suo}\right) - \log\left(\theta\left(i\right)\right) \tag{5.14}$$

Thus:

$$\frac{g_x}{I_D} = \frac{1}{i}\frac{di}{dV_x} - \frac{1}{\theta}\frac{d\theta}{dV_x} + \frac{1}{I_{Suo}}\frac{dI_{Suo}}{dV_x} \tag{5.15}$$

which can be rewritten also as follows:

$$\frac{g_x}{I_D} = \left(1 - \frac{i}{\theta}\frac{d\theta\left(i\right)}{di}\right)\frac{1}{i}\frac{di}{dV_x} + \frac{1}{I_{Suo}}\frac{dI_{Suo}}{dV_x} \tag{5.16}$$

For the evaluation of the differential of the log of the normalized drain current, we take advantage of Eqs. 4.8 and 4.21–4.23 (remember V_P, V_S and V_D are defined with respect to the substrate):

$$\frac{1}{i}\frac{di}{dV_x} = \frac{1}{U_T}\left[\frac{1}{1+q_F+q_R}\frac{dV_P}{dV_x} - \frac{q_F}{i}\frac{dV_S}{dV_x} + \frac{q_R}{i}\frac{dV_D}{dV_x}\right] \tag{5.17}$$

5.5.1 The g_m/I_D Ratio

Equations 5.16 and 5.17 boil down to the expression below in the common-source configuration, since I_{Suo} doesn't depend on V_{GS}:

$$\frac{g_m}{I_D} = \frac{1}{U_T}\frac{1}{1+q_F+q_R}\left(1 - \frac{i}{\theta}\frac{d\theta(i)}{i}\right)\frac{dV_P}{dV_G} \tag{5.18}$$

which, can be rewritten as follows according to Eq. 4.14:

$$\frac{g_m}{I_D} = \frac{1}{nU_T}\frac{1}{1+q_F+q_R}\left(1 - \frac{i}{\theta}\frac{d\theta(i)}{di}\right) \tag{5.19}$$

In weak and in moderate inversion, the g_m/I_D ratio can be further simplified for mobility degradation must not be considered. This gives birth to the expression put to use by the acquisition algorithm:

$$\frac{g_m}{I_D} = \frac{1}{nU_T}\frac{1}{1+q_F+q_R} \tag{5.20}$$

In strong inversion, the evaluation of the derivative inside the parenthesis can be implemented by means of the MATLAB ***polyval*** and ***polyder*** instructions:

$$\frac{i}{\theta}\frac{d\theta(i)}{di} \Rightarrow \frac{polyval([polyder(PSu)\ 0], i)}{polyval(PSu, i)} \tag{5.21}$$

The denominator makes use of the polynomial counterpart of the 'theta' function to return θ (the polynomial *PSu* is derived from the ***global variable*** PolyN or PolyP -). The ***polyder*** instruction in the numerator takes care of the derivative of *PSu* with respect to i. The zero after the ***polyder*** instruction increments the order of the derivative to multiply the result by the normalized drain current i.

Figure 5.18 compares predicted to 'exact' g_m/I_D's considering the 100 nm N-channel transistor (the 'exact' data are obtained by taking the numerical derivative of the log of the 'semi-empirical' drain current). The difference between dashed and crossed lines in strong inversion illustrates the impact of mobility degradation predicted by Eq. 5.21. In weak and moderate inversion, g_m/I_D is not affected legitimating the assumptions made in the acquisition algorithm.

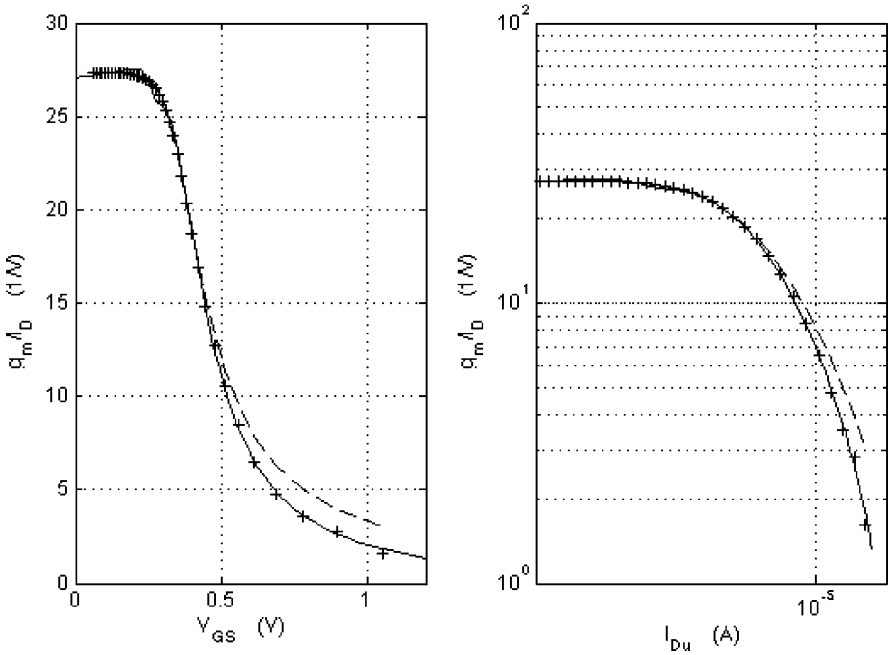

Fig. 5.18 Exact (*plain lines*) and model (*crosses*) g_m/I_D ratios versus the gate-to-source voltage V_{GS} (*left*) and unary drain current (*right*) considering the grounded source 100 nm N-channel transistor. The drain-to-source voltage is equal to 0.6 V (MATLAB fig518.m)

The three next figures illustrate the influence of the gate length, the drain-to-source voltage and back-bias on the semi-empirical and model-driven g_m/I_D ratios. In Fig. 5.19, the gate length is expanded from 0.1 to 4 μm. The smaller g_m/I_D of the 0.1 μm transistor in weak inversion reflects the larger slope factor illustrated by Fig. 5.8 that is characteristic of short channel devices. Similarly, the larger gate voltages required by the short channel device in moderate and strong inversion result from the reverse short channel effect mentioned under Fig. 5.9.

In the right figure, the impact of the gate length on the mobility degradation is clearly visible. The decay of g_m/I_D is much faster with the short channel device. Both g_m/I_D's are compared to the asymptotic construction put to use in Fig. 4.15 for the ideal transistor. The large difference with respect to the ideal transistor, even with long channel devices, is a clear warning not to infer specific currents from measurements based on the intersection of the strong and weak inversion asymptotes.

The two plots of Fig. 5.20 illustrate the influence of the drain voltage. The little impact V_{DS} has on the slope factor is corroborated by the almost similar g_m/I_D ratios in weak inversion. When the transistor is not saturated ($V_{DS} = 0.1$ V), g_m/I_D collapses very rapidly.

Figure 5.21 shows the influence of back-bias on the g_m/I_D ratio. The left side illustrates the anticipated threshold voltage and slope factor increase associated with

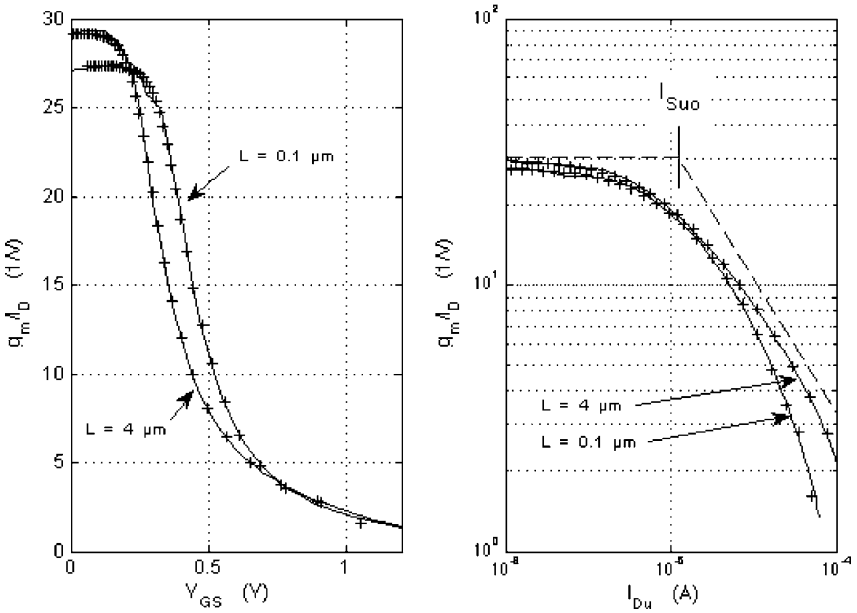

Fig. 5.19 Exact (*plain lines*) and compact model (*crosses*) g_m/I_D ratios versus gate-to-source voltage V_{GS} (*left*) and unary drain current (*right*) considering 0.1 and 4.0 µm gate lengths. The source is grounded and the drain-to-source voltage equal to 0.6 V (MATLAB fig519.m)

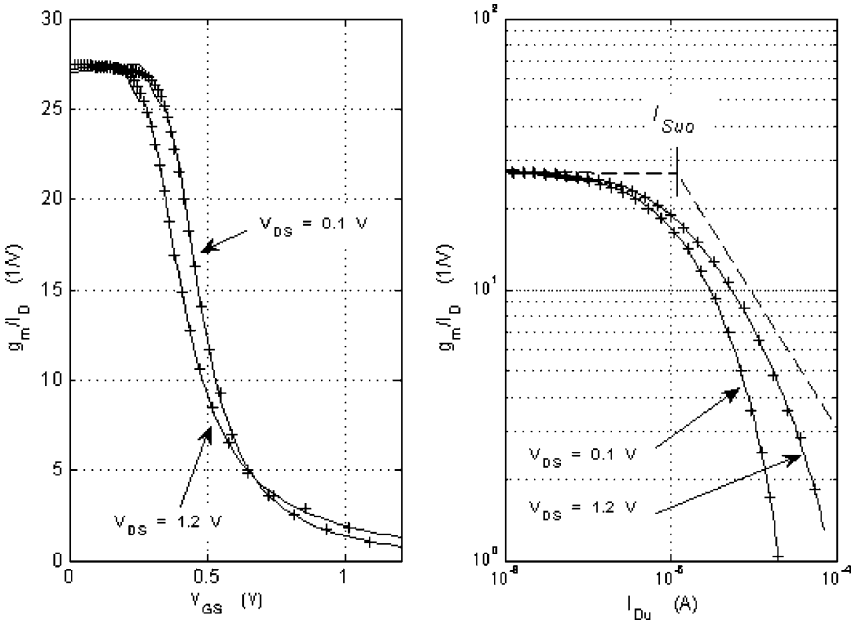

Fig. 5.20 Exact (*plain lines*) and compact model (*crosses*) g_m/I_D ratios versus the gate-to-source voltage V_{GS} (*left*) and unary drain current (*right*) considering a non-saturated ($V_{DS} = 0.1$V) and a saturated transistor ($V_{DS} = 1.2$ V). The source is grounded and the gate length equal to 100 nm (MATLAB fig520.m)

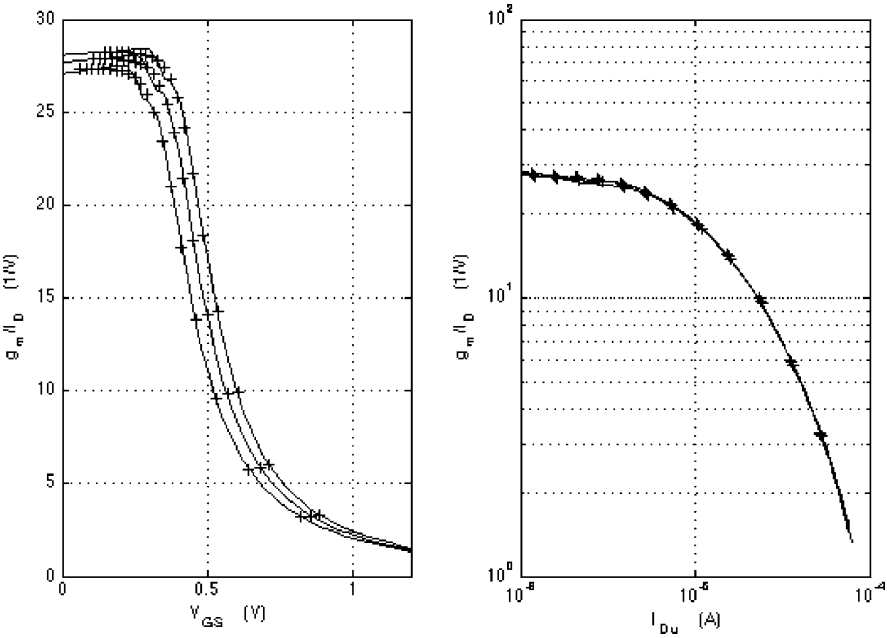

Fig. 5.21 Exact (*plain lines*) and compact model (*crosses*) g_m/I_D ratios versus the gate-to-source voltage V_{GS} (*left*) and unary drain current (*right*) considering three source voltages equal to 0, 0.4 and 0.8 V (*left to right*). The gate length is equal to 100 nm and the drain-to-source voltage 0.6 V (MATLAB fig521.m)

the growing back-bias voltage. In the right side, the curves merge practically in strong inversion (MATLAB fig521.m).

Figure 5.22 shows a magnified view of the g_m/I_D of the 100 nm N-channel transistor in weak and moderate inversion for V_{DS} equal to 0.6 V. The figure illustrates the 'filtering' effect of the compact model mentioned earlier. The model ignores the small dip near 2.7 V, which is probably due to side current.

5.5.2 The g_d/I_D Ratio

The drain conductance over drain current ratio g_d/I_D derived from Eqs. 5.16 and 5.17 boils down to the expression below where the influence of the drain voltage on the slope factor n has been neglected for it is small compared to the influence of V_{To}. When the transistor is saturated, the impact of the drain voltage is reflected by the sensitivity factor S_{VTo} and by the derivative of the log of the specific current.

$$\frac{g_d}{I_D} = \frac{1}{nU_T}\left(1 - \frac{i}{\theta}\frac{d\theta}{di}\right)\left(\frac{1}{1+q_F+q_R}\,|S_{VTo}| + \frac{n\,q_R}{i}\right) + \frac{d}{dV_D}\log\left(I_{Suo}\right) \tag{5.22}$$

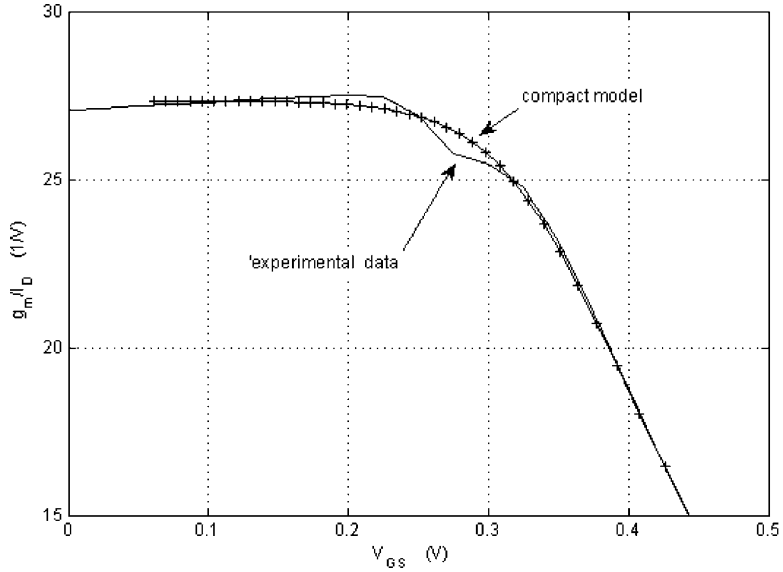

Fig. 5.22 Second order effects are ignored by the compact model (MATLAB fig522.m)

The merit of this expression is that it separates clearly the contributions of mobility degradation (the first parenthesis), D.I.B.L. (the first term in the second parenthesis), de-saturation (the second term in the second parenthesis) and C.L.M. (the last term). The impact every item has on the reciprocal of g_d/I_D, can be assessed separately thus. The point is illustrated in Figs. 5.23 and 5.24, which represent the Early voltages[5] of the N-channel transistors. In the first, V_{GS} is equal to 0.3 V (moderate inversion), in the second 0.6 V (strong inversion). Both figures report results obtained with two gate lengths: 0.1 µm left and 1 µm right. Curve (1) represents the Early voltage without D.I.B.L and C.L.M terms. As soon as the transistor enters saturation, the output conductance gets very small, almost zero, making the Early voltage very large. The transistor becomes practically an ideal current source like in the C.S.M. When second order effects are introduced, the picture changes drastically. Curve (2) shows the influence of D.I.B.L without C.L.M, whereas curve (3) combines the two. With the 1 µm transistor, the Early voltage in saturation is fixed essentially by C.L.M. The impact of D.I.B.L. is almost negligible

[5] The Early voltage is defined generally as the voltage where the tangent to the $I_D(V_{DS})$ characteristic crosses the horizontal axis. The Early voltage considered here is the difference between the aforementioned crossing point and the actual drain-to-source voltage. This makes I_D/V_A identical to g_d. When the Early voltage is large, the two definitions coincide more or less, but this doesn't hold true with short channel devices. In weak inversion, the zero crossing may be located even to the right of the origin owing to the exponential characteristic of the drain current like in Fig. 5.16. The Early voltage would be negative with the first definition.

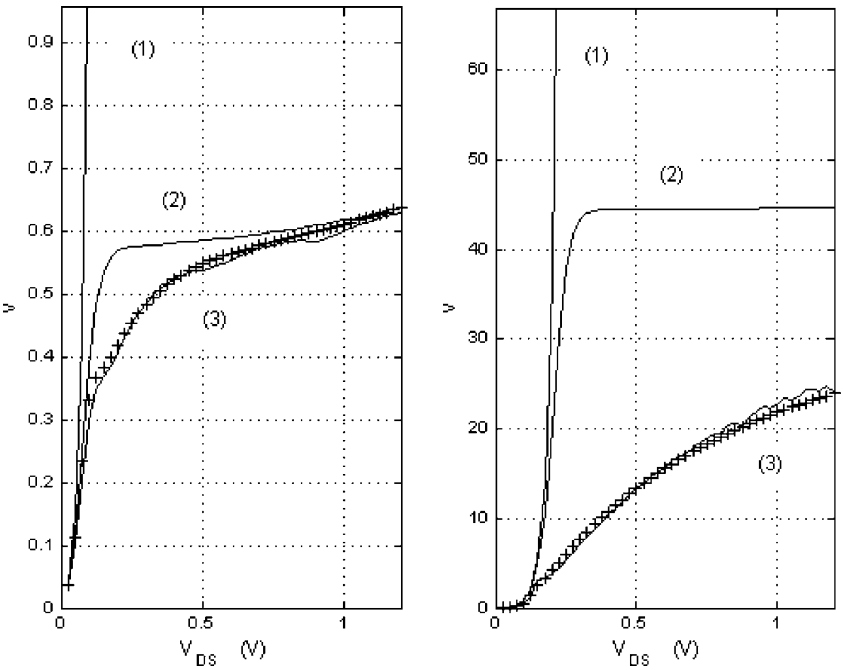

Fig. 5.23 Cumulated contributions to the Early voltage predicted by Eq. 5.23, considering a grounded 100 nm N-channel transistor (*left*) and a 1 μm (*right*). Crosses illustrate the actual semi-empirical Early voltage. The gate-to-source voltage is equal to 0.3 V (MATLAB gdID.m)

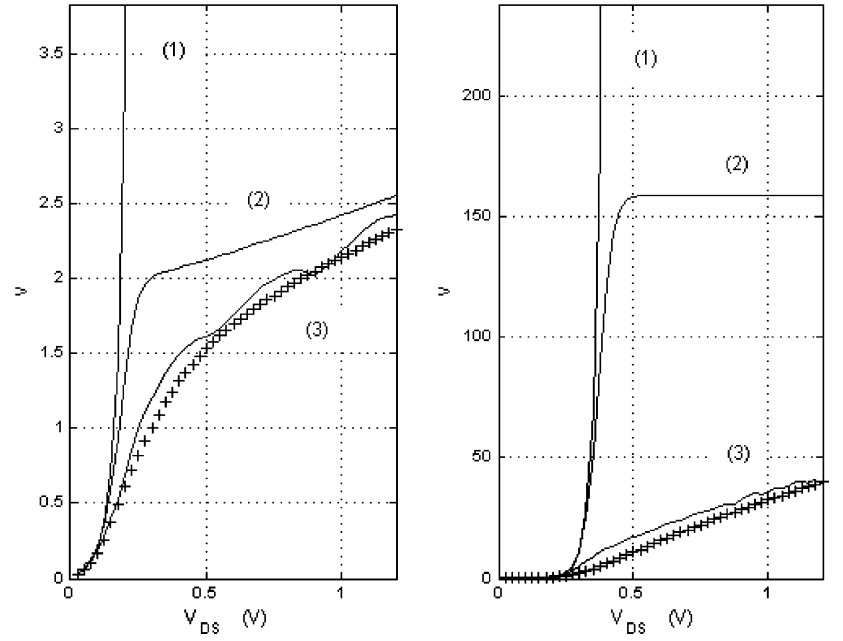

Fig. 5.24 Cumulated contributions to the Early voltage predicted by Eq. 5.23, considering a grounded 100 nm N-channel transistor (*left*) and a 1 μm (*right*). Crosses illustrate the actual semi-empirical Early voltage. The gate-to-source voltage is equal to 0.6 V (MATLAB gdID.m)

in strong inversion, small but not negligible in moderate inversion. With the 0.1 μm transistor, the opposite holds true. D.I.B.L overwhelms C.L.M.

The line consisting of crosses in the two figures illustrates the Early voltage predicted by the semi-empirical model. The 'semi-empirical' and model-driven Early voltages are more similar in moderate than in strong inversion. The difference increases with longer gate lengths owing to the increasing noise that comes with the drastic reduction of the derivatives of I_{Suo}. A more accurate approach is considered in the next chapter that does not require the derivatives of V_{To} and I_{Suo}.

5.6 Conclusions

Drain currents predicted by the model of Chapter 4 are very similar to real drain currents. Can one extend the compact model to real transistors? The answer is yes prided some conditions are fulfilled. The model reproduces reasonably well real $I_D(V_{GS})$ characteristic even those of short-channel devices as long as the source and drain voltages are kept constant. Not only drain currents, but also g_m/I_D and g_d/I_D ratios can be reconstructed with acceptable accuracy. As soon as the drain or source voltage are modified, all parameters must be updated.

A parameter extraction algorithm is set up evaluating the slope factor, the threshold voltage, the specific current and a polynomial fit taking care of mobility degradation. The result brings about a number of interesting observations highlighting the impact of short channel effects on the parameters of the compact model, namely D.I.B.L and C.L.M.

The simplicity of the model lays down the grounds for analytical expressions. These allow performing sizing without the need to explore blindly wide ranges of drain currents. The idea is to control MOS transistors by means of variables like the normalized drain current or the forward mobile charge density q_F. A first example is considered is the next chapter concerning the sizing the real I.G.S. The method takes advantage of few parameters instead of complex advanced models with large numbers of parameters.