Physically-Based Approach to Deep-Submicron MOSFET Compact Model Parameter Extraction

Siau Ben Chiah*, Xing Zhou*, Khee Yong Lim†, Alex See†, and Lap Chan†

*School of Electrical & Electronic Engineering, Nanyang Technological University
Nanyang Avenue, Singapore 639798, exzhou@ntu.edu.sg
†Chartered Semiconductor Manufacturing Ltd, 60 Woodlands Industrial Park D, St. 2, Singapore 738406

ABSTRACT

This paper demonstrates a physically-based approach to parameter extraction of the compact \( I_{ds} \) model we have developed for deep-submicron technology development. A two-iteration parameter-extraction scheme is described, which improves the previous one-iteration approach. Parameter calibration for the \( V_t \) model is revisited. Comparison of parabolic and linear body-bias dependency with new calibration sequence for the \( V_t \) model is carried out, which shows higher accuracy in \( V_t \) modeling for the new parabolic interpolation. Optimization for the halo pile-up centroid, LDD lateral diffusion as well as saturation velocity is carried out to improve the overall \( V_t \) and \( I_{ds} \) modeling. This has been verified with the experimental data from a 0.18-\( \mu \)m CMOS technology wafer.

Keywords: Deep-submicron MOSFET, compact model, parameter extraction, nonlinear regression, optimization.

1 INTRODUCTION

A unified compact \( I_{ds} \) model [1] from subthreshold to saturation region, which conforms to the well-known long-channel characteristics and encapsulates all major short-channel effects (SCEs) and reverse short-channel effects (RSCEs) in deep-submicron MOSFETs, has been formulated through physics-based effective transformation. Simple one-iteration parameter-extraction method has been introduced, which covers full range of channel lengths (without “binning”) and bias conditions. In the one-iteration approach, there are several assumptions made. First, simple equation is used before extraction of non-calibrated parameters at each step. This is based on the assumption that the effect of non-calibrated parameters can be ignored at the condition when the parameter in the current step is being extracted. Secondly, fixing the parameter value after extraction is based on the assumption that the calibrated effect will not be affected by subsequent extraction. Finally, if a fitting parameter is found to have different values at different bias conditions, a linear bias dependence of that parameter is assumed and extracted at two extreme bias conditions.

These assumptions have been validated by the application of the model to the 0.25-\( \mu \)m technology data with reasonable accuracy [1]. However, when the same approach is applied to the 0.18-\( \mu \)m technology data, error due to these assumptions becomes fairly large.

In this paper, the approach to physically-based parameter extraction is revisited. The threshold-voltage parameters are calibrated with parabolic-\( V_{th} \) dependency for increased accuracy. A two-iteration extraction scheme is followed, in which the first iteration result is used as the initial guess. Together with local optimization of a few physical parameters (such as LDD lateral diffusion \( \sigma \), the pile-up charge centroid \( l_p \) as well as saturation velocity \( v_{sat} \)), we show that the new extraction approach can provide more physical and accurate parameters for the compact model.

2 PARAMETER EXTRACTION

The \( V_t \) model conforms to the well-known long-channel model with effective quantities for SCEs and RSCEs:

\[
V_t = V_{FB} + \phi_s + \gamma_{eff} \sqrt{\phi_{s0} - V_{bs}} \tag{1}
\]

\[
\gamma = \sqrt{2q\varepsilon N_{eff}} \frac{1}{C_{ox}} \tag{2}
\]

\[
\phi_s = \phi_{s0} - \Delta \phi_s = 2\phi_F + \Delta - \Delta \phi_s \tag{3}
\]

in which \( \Delta \) is introduced to model long-channel drain-induced barrier lowering (DIBL) effect with an assumed linear dependency on \( V_{ds} (\Delta = \Delta V_{ds}) \).

Charge-sharing effect is modeled by \( \lambda \):

\[
\gamma_{eff} = \gamma - \frac{\lambda}{L_{eff}} \frac{2\varepsilon}{C_{ox}} \left( \sqrt{\phi_s - V_{bs}} + \frac{\delta V_{ds}}{\sqrt{\phi_s - V_{bs}}} \right) \tag{4}
\]

Barrier-lowering effect is modeled by \( \alpha \):

\[
\Delta \phi_s = \frac{1}{\cosh(L_{eff}/2\alpha)} \left( V_{bs} - \phi_{s0} \right) \cosh \left( \frac{z}{2} \right) \frac{\phi_{V_{ds}}}{2} \sinh \left( \frac{L_{eff} - z}{2\alpha} \right) \sinh \left( \frac{L_{eff} + z}{2\alpha} \right) \tag{5}
\]

\[
l_{\alpha} = \alpha (\phi_{s0} - V_{bs})^{0.25} \tag{6}
\]

where the short-channel DIBL effect is modeled by \( \delta \) and \( \varphi \) in \( \gamma_{eff} \) and \( \Delta \phi_s \) at high \( V_{ds} \). RSCE is modeled by \( \kappa \) and \( \beta \) [2]:

\[
\Sigma_{eff} = \kappa \phi_{V_{ds}} \sinh \left( \frac{L_{eff} - z}{2\alpha} \right) \sinh \left( \frac{L_{eff} + z}{2\alpha} \right) \tag{7}
\]

\[
l_{\beta} = \beta (\phi_{s0} - V_{bs})^{0.5} \tag{8}
\]
\[ N_{eff} = \frac{\sqrt{\pi k N_{ch}}}{L_{eff}^4 \beta} \left[ \text{erf} \left( \frac{L_{eff} - l_\mu}{l_\beta} \right) + \text{erf} \left( \frac{l_\mu}{l_\beta} \right) \right] + N_{ch} \] (7)

\[ I_{\beta} = \beta (\phi_{0} - V_{hs})^{0.25} \] (8)

where the pile-up charge centroid \( l_\mu \) is taken as an optimization parameter together with the LDD lateral diffusion parameter \( \sigma \) in the effective channel length:

\[ L_{eff} = L - 2\sigma x_j = L - \Delta_{CD} - 2\sigma x_j \] (9)

where \( \Delta_{CD} \) is the critical-dimension correction and \( x_j \) is the LDD junction depth.

The \( V_i \) model has 3 long-channel fitting parameters, \( V_{FB}, N_{ch}, \) and \( \Delta_{L} \), which are extracted from the long-channel \( V_i - V_{hs} \) data using simple equations (without SCE parameters) at low \( V_{ds} \) (extracting \( V_{FB} \) and \( N_{ch} \)) and high \( V_{ds} \) (extracting \( \Delta_{L} \) with \( V_{FB} \) and \( N_{ch} \) fixed). The short-channel \( V_i \) model has 6 fitting parameters, \( \lambda, \alpha, \kappa, \beta, \delta, \) and \( \varphi \), which are extracted for a given technology from the measured \( V_i - L \) data, where the measured \( V_i \) is based on constant-current definition [3] in which only 6xN point \((I, V)\) data plus one \( I_{ds} - V_{gs} \) curve are required (where \( N \) is the number of different gate-length devices).

The \( V_i \) model includes all major SCEs and RSCEs that can be appropriately modeled by fitting \( \lambda, \alpha, \kappa, \beta, \delta, \) and \( \varphi \) to a given \( V_i - L \) data at any biases for the given \( \sigma \) and \( l_\mu \) values (see Section 3 for the \( \sigma - l_\mu \) optimization). It has been observed that charge-sharing and barrier-lowering parameters \( (\lambda \text{ and } \alpha) \) are relatively independent of \( V_{hs} \) and, thus, extracted at high \( |V_{hs}| \) when the effects are largest. In the previous one-iteration approach, DIBL effect \( (\delta \text{ and } \varphi) \) is first ignored when \( \lambda, \alpha, \kappa, \) and \( \beta \) are being extracted at low \( V_{ds} = V_{do} = 0.1 \) V, then \( \delta \) and \( \varphi \) are calibrated at high \( V_{ds} = V_{ds} = 1.98 \) V with the extracted \( \lambda, \alpha, \kappa, \) and \( \beta \) values fixed. \( \kappa, \beta, \delta, \) and \( \varphi \) are assumed to be linearly dependent on \( V_{hs} \) and are calibrated at low \( V_{hs} = V_{hs} = 0 \) and high \( V_{hs} = V_{hs} = -1.8 \) V. While these assumptions have been valid for the 0.25-\( \mu \)m data, it is found that they are no longer accurate for the 0.18-\( \mu \)m technology data.

In the new two-iteration approach, we take the extracted \( \delta \) and \( \varphi \) from the first iteration as the initial guess values and re-calibrate \( \lambda, \alpha, \kappa, \) and \( \beta \) with the full \( V_i \) model where the minor DIBL effect at \( V_{ds} \) is accounted for. Also, we assume that \( \kappa, \beta, \delta, \) and \( \varphi \) have parabolic \( V_{hs} \) dependence:

\[ \begin{align*}
\kappa &= \kappa_0 + \kappa_1 V_{hs} + \kappa_2 V_{hs}^2 \\
\beta &= \beta_0 + \beta_1 V_{hs} + \beta_2 V_{hs}^2 \\
\delta &= \delta_0 + \delta_1 V_{hs} + \delta_2 V_{hs}^2 \\
\varphi &= \varphi_0 + \varphi_1 V_{hs} + \varphi_2 V_{hs}^2
\end{align*} \] (10)

They are calibrated at \( V_{hs} \), \( V_{ds} \), and \( V_{bs} = V_{bs} = -0.9 \) V to account for different amount of RSCE and DIBL effects at different \( V_{hs} \) conditions, while \( \lambda, \alpha \) are still assumed to be \( V_{bs} \) and \( V_{ds} \) independent.

Excellent fitting of the parabolic-\( V_{bs} \) dependent \( V_i \) model at the extreme \( V_{ds} \) and \( V_{bs} \) conditions is demonstrated in Fig. 1. The improvement of the parabolic-\( V_{bs} \) model over the linear-\( V_{bs} \) model is shown in Fig. 2 with the excellent prediction of the measured \( V_i = L \) at \( V_{bs} = -0.45 \) and \(-1.35 \) V, in which none of the data has been used in parameter extraction.

Similarly for the \( I_{ds} \) model, the ideal long-channel MOSFET is described by the simple equation:

\[ I_{ds0} = \mu_0 C_{\alpha V} \left[ \frac{V_{gs} - V_{t}}{V_{gs} - V_{t}} \right] \] (11)

\[ A_h = 1 + \xi \frac{\gamma}{2\sqrt{\phi_{0} - V_{hs}}} \] (12)

\[ \mu_{eff} = \mu_0 \frac{\xi}{1 + \Delta_{I_{ds}}} \left[ \frac{E_{sat} L_{eff}}{I_{eff}} \right] \] (13)

where the parameters such as vertical-field mobility \( \mu_0 \) and bulk-charge factor \( A_h \) are well defined. For the short-channel \( I_{ds} \) model, Eq. (11) is modified to include SCEs such as series resistance \( (R_{sd}) \), channel-length modulation (CLM), and subthreshold modeling.

In the one-iteration approach, we start from long-channel equation to extract \( \mu_0 \) parameters at low \( V_{ds} \), assuming that the minor bulk-charge effect can be ignored by setting \( \xi = 1 \). This is followed by calibrating \( A_h \) using the long-channel \( I_{ds} \) data at high gate voltage \( V_{gs} = V_{gs} = 1.796 \) V (with fixed \( \mu_0 \)). \( I_{ds} \) in Eqs. (11) and (13) is changed to \( I_{ds0} \) using the BSIM smoothing function [4] with long-channel saturation voltage:

\[ V_{dss0} = \frac{E_{sat} L_{eff}}{V_{gs} - V_{t}} + \frac{A_h E_{sat} L_{eff}}{V_{gs} - V_{t}} \] (14)

Then, Eq. (11) is modified to include \( R_{sd} \) in the short-channel \( I_{ds} \) model with a consistent solution for \( V_{dss0} \):
where $V_{\text{eff}}$ is an effective Early voltage with one fitting parameter ($\xi$) extracted with an optimization loop for the saturation velocity ($v_{sat}$). Finally, $V_{gs} - V_t$ in the “simple” equations is replaced by an effective gate override ($V_{gs}$) [5] for subthreshold modeling with one fitting parameter ($V_{off}$) for the full range of geometry and biases.

In the one-iteration approach, it is assumed that for each calibration step, the newly extracted parameter will not affect those that have been calibrated. Fig. 3 illustrates this assumption by comparing the $I_{ds} - V_{ds}$ characteristics using the simple $I_{ds}$ model (11) without CLM to the full compact $I_{ds}$ model (16) with CLM for long- and short-channel devices. Both simple and full $I_{ds}$ models simulate accurately for the long-channel $I_{ds} - V_{ds}$ behavior, while the full $I_{ds}$ model can predict short-channel CLM effect without affecting what has been characterized at the long-channel. Similar idea is shown in Fig. 4 during the subthreshold calibration, which does not affect the calibrated behavior in the strong inversion.

In the new two-iteration approach, the above scheme is still followed, in which the extracted parameter set is used as the initial guess for the full model equations in the second iteration. Comparison between one- and two-iteration approaches is shown in Fig. 5 for the root-mean-square (RMS) error in $I_{ds} - V_{gs}$ over different $V_{th}$ biases at high and low $V_{ds}$.

### 3 OPTIMIZATION

The philosophy of our model parameter extraction is based on separation of fitting and physical parameters, where process-dependent fitting parameters are extracted using nonlinear regression method [6] at the “given” physical parameter values, which are optimized through optimization loops for minimum RMS error in the target being fitted. This is to avoid extraction of unphysical values of these physical parameters if they were used as regression parameters. $V_t$ model parameters are extracted together with $\sigma$ and $I_{ds}$ optimization loops. RMS errors are calculated with the full $V_t$ model over all gate lengths at four combinations of $V_{th}$ and $V_{ds}$ conditions ($V_{th}$, $V_{bias}$, $V_{ds}$). Total RMS errors are computed and plotted against different $\sigma$ and $I_{ds}$ values as shown in Fig. 6. Minimum RMS error in $I_{ds}$ over all gate lengths falls on $\sigma = 0.63$ and $I_{ds} = -40$ nm, which is chosen to be the optimum values for $L_{off}$ and $N_{off}$.

The $I_{ds}$ CLM parameter is extracted together with the $v_{sat}$ optimization loop. RMS errors are computed in $v_{sat}$ optimization over $I_{ds} - V_{ds}$ at high $V_{gs}$ and low $V_{th}$ for $L = 0.16$ µm during CLM calibration. A minimum RMS error is observed at $v_{sat} = 6\times 10^6$ cm/s, as shown in Fig. 7, which also shows further improvement in RMS error with two-iteration for $I_{ds}$ model extraction during $v_{sat}$ optimization.

### 4 CONCLUSION

In conclusion, the proposed new $V_t$ calibration sequence allows either linear-$V_{th}$ or parabolic-$V_{th}$ dependent model and one- or two-iteration approach being incorporated during the $V_t$ model parameter extraction, which can be selected based on available process data to trade off accuracy and complexity.

Physical parameter extraction is performed in a more systematic way, in which the parameters are extracted through optimization loops of physical parameters based on minimum RMS error criteria. The two-iteration extraction approach effectively corrects the errors due to the assumptions made in the one-iteration approach, which provides higher accuracy in deep-submicron device modeling.

The potential impact of our improved calibration approach on $V_t$ and $I_{ds}$ lies in more accurate and physical device parameter extraction, which would lead to more meaningful statistical modeling in circuit simulation and deep-submicron technology development.

### REFERENCES

Figure 2: Measured (symbols) and modeled (lines) $V_t - L$ characteristics for the linear-$V_{bs}$ (dotted) and parabolic-$V_{bs}$ (solid) models at the indicated bias conditions.

Figure 3: Measured (symbols) and modeled (lines) $I_{ds} - V_{ds}$ characteristics before (cross-dotted) and after (solid) CLM calibration for $L = 10\, \mu m$ and $L = 0.16\, \mu m$ (inset) devices.

Figure 4: Measured (symbols) and modeled (lines) $I_{ds} - V_{gs}$ characteristics before (cross-dotted) and after (solid) subthreshold calibration for the $L = 0.16\, \mu m$ device.

Figure 5: RMS error in $I_{ds}$ over all $V_{gs}$ for one-iteration (dotted lines) and two-iteration (solid lines) at $V_{dd} = 0.1\, V$ (circles, left axis) and $V_{dd} = 1.98\, V$ (triangles, right axis).

Figure 6: RMS error in $V_t$ over all $L$ for $\sigma - l_{\mu}$ optimization. The optimal condition $\sigma = 0.63$ and $l_{\mu} = -40\, nm$ is chosen.

Figure 7: RMS error in $v_{sat}$ optimization over all $I_{ds} - V_{ds}$ for one-iteration (dotted lines) and two-iteration (solid lines) during CLM calibration (at $V_{gs} = 1.796 V$, $V_{bs} = 0$, and $L = 0.16\, \mu m$). The optimal $v_{sat} = 6\times10^6\, cm/s$ is chosen.