Compact Model for Manufacturing Design and Fluctuation Study

Khee Yong Lim*, Xing Zhou*

*Chartered Semiconductor Manufacturing Ltd., 60 Woodlands Ind. Park D, Street 2, Singapore 738406. Tel: 65-3946251. Fax: 65-3946505. Email: limkheey@charteredsemi.com

*School of Electrical & Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798. Email: exzhou@ntu.edu.sg

ABSTRACT

In this article, a physics based compact model [1, 2] has been used as a tool for manufacturing process variability study. Three critical end-of-line (EOL) measured electrical testing (ET) parameters, namely $V_{th}$, $I_{on}$ and $I_{off}$, of a 0.25$\mu$m CMOS technology together with three in-line device parameters ($t_{ox}$, $L_{eff}$, $N_s$) are used for the study. Based on the compact model, the sensitivity of each EOL parameter with respect to individual in-line device parameter can be easily computed. To systematically examine the correlation of final electrical parameters with the fluctuation of in-line device parameter directly related to process, a mathematical model that expresses ET parameter variance in terms of process-related parameter variance [3] has been employed. By combining the above mathematical model and sensitivity rate calculated from calibrated compact model together with computed normalized variance from actual experimental data, the device parameter fluctuation can be derived from measured ET fluctuation.

Keywords: Compact model, semiconductor process variability, fluctuation and sensitivity study, in-line device parameters and end-of-line ET parameters.

1 INTRODUCTION

Process fluctuations, which are causing final device parameters variation, are inevitable in practice. These effects of semiconductor manufacturing variability can be estimated by simple worst-case analysis, or by more sophisticated statistical techniques. Fig. 1 shows the variation of EOL experimental threshold voltage ($V_{th}$), turn-on current ($I_{on}$), and turn-off current ($I_{off}$) of a wafer with seventeen sites of measurement. These variability can be due to a number of factors, such as etch rate variation between center and edge of the wafer, temperature contour variation across the wafer and so on. It has always been a major challenge for semiconductor engineer to minimize their lot-to-lot, wafer-to-wafer and site-to-site variability and improve the overall process capability and stability indices ($C_p$ and $C_{pk}$).

A set of compact model for threshold voltage [1] and drain current [2] of MOSFET has been best fitted to nominal experimental data as shown in the line in Fig. 1. The model is single piece and continuous from subthreshold regime ($I_{on}$) to saturation regime ($I_{off}$). The model is in analytical compact form with 23 process-dependent fitting parameters. The model is made up from some in-line device parameters, such as gate oxide thickness ($t_{ox}$), effective channel doping ($N_e$), effective channel length ($L_{off}$), LDD junction depth ($X_j$), which have a direct correlation to the particular process module. Therefore, correlation between EOL ET parameters and process fluctuation can be found from the well-fitted compact model. The next section describes the role of the compact model [1, 2] for manufacturing design and fluctuation study.

2 MODEL FOR MANUFACTURING FLUCTUATION STUDY

A well-calibrated physical model is a useful tool for determining the possible cause for resultant device variation. The characterized model as shown in Fig. 1 is employed to generate $V_{th}$, $I_{on}$, $I_{off}$ variation plots in term of channel length variation ($\Delta L$), gate oxide thickness variation ($\Delta t_{ox}$), and channel doping variation ($\Delta N_e$). Fig. 2 shows nine different set of plot with a ±10% variation of $\Delta L$, $t_{ox}$ and $N_e$ from their nominal values on the final electrical parameter $V_{th}$, $I_{on}$, $I_{off}$ generated from the calibrated compact model. By comparing the characteristic of Fig. 2 with the actual variation data as shown in Fig. 1, the possible reason for the observed fluctuation can be easily attributed to channel length variation. This is because most of the observed variations in Fig. 1 are only in the short channel region, which is observed differently when $t_{ox}$ and $N_e$ are varied. Based on this piece of information, device engineer may want to work with either lithography or poly etch engineer for improving device variability by optimizing the corresponding process step.

Fig. 2 serves as a preliminary guidance for the possible cause for EOL ET variability. Indeed with analytical model, the variation of $V_{th}$, $I_{on}$ and $I_{off}$ with respect to $\Delta L$, $t_{ox}$ and $N_e$ is quantifiable. To qualitatively examine the sensitivity of final electrical parameters on the fluctuation of parameters directly related to process, a mathematical model that expresses device parameter variance in terms of process-related parameter variances [3] is employed. For example, the model applied to the dependence of $V_{th}$ variance on in-line parameter variances has the following form:
Fig. 3 shows the sensitivity rate of threshold voltage with respect to \( \Delta L \), \( t_{ox} \), and \( N_s \) for long-, medium, and short-channel devices generated from the calibrated compact model [1,2]. Similarly, sensitivity plots for \( I_{on} \) and \( I_{off} \) with respect to \( \Delta L \), \( t_{ox} \), and \( N_s \) are plotted in Fig. 4 and Fig. 5, respectively. It is observed that for long-channel devices, the parameter that should be concerned is the saturation current. As for short-channel devices, the parameter that has large percentage of fluctuation is the leakage current. Turn-on current is most sensitive to oxide thickness fluctuation, whereas deep-submicron turn-off current is most sensitive to channel length fluctuation. Hence, these sensitivity plots serve as a tool for assisting device and process engineer to focus on the right track for minimizing electrical parameter variations.

From the seventeen sites of the measured data as shown in Fig. 1, one can calculate normalized variances for \( V_{th} \), \( I_{on} \) and \( I_{off} \) of individual device. With these variances, the variances of important parameters, \((\sigma \delta_{\Delta L})^2\), \((\sigma \delta_{N_s})^2\) and \((\sigma \delta_{I_{on}})^2\), can be computed from (1), (2) and (3) provided all the other process-related parameter variances can be ignored. The long-channel example below presents the concept of determining process fluctuations from electrical parameter measurement fluctuations. From Figs. 3 to 5, the sensitivity rate with respect to channel length is very small for long-channel devices. Thus, by substituting the actual measured variances and calculated sensitivity values into (1) and (2), we get

\[
(\sigma \delta_{V_{th}})^2 = \left( \frac{\partial \delta_{V_{th}}}{\partial V_{th}} \right)_L \sigma \delta_L + \left( \frac{\partial \delta_{V_{th}}}{\partial V_{th}} \right)_{N_s} \sigma \delta_{N_s} + \left( \frac{\partial \delta_{V_{th}}}{\partial I_{on}} \right)_{I_{off}} \sigma \delta_{I_{on}} + \cdots
\]

(1)

All the quantities in (1) are normalized, which is indicated by the symbol \( \delta \), that is \( \delta V_{th} = (V_{th} - V_{th})/\bar{V}_{th} \). \((\sigma \delta V_{th})^2\) and \( \partial \delta V_{th}/\partial \Delta L \) are the normalized variance and the sensitivity rate of threshold voltage with respect to channel length, respectively. All the partial derivatives are determined at the points corresponding to specification (nominal) value. Analogous equations for other parameters’ variance, such as \( I_{on} \) and \( I_{off} \), can be written. Since all the equations are normalized, thus all the parameters \((\delta V_{th} \ V_{th} \ L)\) with different units can be compared directly.

\[
(\sigma \delta_{V_{th}})^2 = \left( \frac{\partial \delta_{V_{th}}}{\partial \Delta L} \right)_L \sigma \delta_{\Delta L} + \left( \frac{\partial \delta_{V_{th}}}{\partial V_{th}} \right)_{N_s} \sigma \delta_{N_s} + \left( \frac{\partial \delta_{V_{th}}}{\partial I_{on}} \right)_{I_{off}} \sigma \delta_{I_{on}} + \cdots
\]

(2)

Fig. 3 shows the sensitivity rate of threshold voltage with respect to \( \Delta L \), \( t_{ox} \) and \( N_s \) for long-, medium, and short-channel devices generated from the calibrated compact model [1,2]. Similarly, sensitivity plots for \( I_{on} \) and \( I_{off} \) with respect to \( \Delta L \), \( t_{ox} \) and \( N_s \) are plotted in Fig. 4 and Fig. 5, respectively. It is observed that for long-channel devices, the parameter that should be concerned is the saturation current. As for short-channel devices, the parameter that has large percentage of fluctuation is the leakage current. Turn-on current is most sensitive to oxide thickness fluctuation, whereas deep-submicron turn-off current is most sensitive to channel length fluctuation. Hence, these sensitivity plots serve as a tool for assisting device and process engineer to focus on the right track for minimizing electrical parameter variations.

From the seventeen sites of the measured data as shown in Fig. 1, one can calculate normalized variances for \( V_{th} \), \( I_{on} \) and \( I_{off} \) of individual device. With these variances, the variances of important parameters, \((\sigma \delta_{\Delta L})^2\), \((\sigma \delta_{N_s})^2\) and \((\sigma \delta_{I_{on}})^2\), can be computed from (1), (2) and (3) provided all the other process-related parameter variances can be ignored. The long-channel example below presents the concept of determining process fluctuations from electrical parameter measurement fluctuations. From Figs. 3 to 5, the sensitivity rate with respect to channel length is very small for long-channel devices. Thus, by substituting the actual measured variances and calculated sensitivity values into (1) and (2), we get

\[
(\sigma \delta_{V_{th}})^2 = \left( \frac{\partial \delta_{V_{th}}}{\partial \Delta L} \right)_L \sigma \delta_{\Delta L} + \left( \frac{\partial \delta_{V_{th}}}{\partial V_{th}} \right)_{N_s} \sigma \delta_{N_s} + \left( \frac{\partial \delta_{V_{th}}}{\partial I_{on}} \right)_{I_{off}} \sigma \delta_{I_{on}} + \cdots
\]

(4)

1.7 \times 10^{-4} = 0.89(\sigma \delta_{V_{th}})^2 + 0.31(\sigma \delta_{N_s})^2 + R_1

(5)

5.5 \times 10^{-5} = 1.75(\sigma \delta_{V_{th}})^2 + 0.16(\sigma \delta_{N_s})^2 + R_2

where \( R_1 \) and \( R_2 \) are the residue variance of \( V_{th} \) and \( I_{on} \), respectively, due to other process-related parameters. Since \( R_1 \) and \( R_2 \) are known, the following inequality equations can be formulated based on (4) and (5).

\[
0.89(\sigma \delta_{V_{th}})^2 + 0.31(\sigma \delta_{N_s})^2 \leq 1.7 \times 10^{-4}
\]

(6)

\[
1.75(\sigma \delta_{V_{th}})^2 + 0.16(\sigma \delta_{N_s})^2 \leq 5.5 \times 10^{-5}
\]

(7)

\[
(\sigma \delta_{V_{th}})^2 \geq 0
\]

(8)

\[
(\sigma \delta_{N_s})^2 \geq 0
\]

(9)

The above inequalities can be easily solved with the help of Fig. 6. \((\sigma \delta_{\Delta L})^2\) and \((\sigma \delta_{N_s})^2\) lie in the overlap region A. The maximum normalized variance of \( t_{ox} \) is 1.9 \times 10^{-5}, thus the maximum standard deviation is \( \sqrt{1.9 \times 10^{-5} t_{ox}} = 0.26 \) Å for the CMOS process. Similarly, the maximum normalized variance of \( N_s \) is 5.5 \times 10^{-5}, thus the maximum standard deviation is \( \sqrt{5.5 \times 10^{-5} N_s} = 2.38 \times 10^{15} \) cm\(^{-3}\).

One of the very interesting possible usages of physics based compact model is its predictability to the next generation MOS device design. A fully characterized compact model is much faster and more accurate to predict short-channel device characteristic than any TCAD approach. It can be utilized to assist semiconductor engineers to reduce wafer split-run. Figs. 7, 8 and 9 show process design windows for the specifications: \( I_{on} \geq 0.001A \), \( 0.5V \leq V_{th} \leq 0.6V \), \( I_{off} \leq 5 \times 10^{-5}A \). The shaded region is where all the specifications are met. Different applications require different specifications. A reliable compact model
can easily generate process design database that meets the specifications.

3 CONCLUSION

A fully characterized compact model has been employed to quantify and correlate the end-of-line critical ET parameters with some in-line process-related device parameters. The approach has illustrated the possibility of applying physics-based compact model in the field of manufacturing design and fluctuation study, which is different from conventional application for circuit simulation.

REFERENCES


Figure 1: Fitted model (lines) as compared to seventeen sites of experimental data (symbols) for $V_{th}$, $I_{on}$ and $I_{off}$.

Figure 2: $V_{th}$, $I_{on}$, $I_{off}$ variations due to small fluctuation of channel length, oxide thickness and channel doping. Solid lines: -10% of nominal value; Dotted lines: +10% of nominal value.
Figure 3: Histogram showing normalized $V_{th}$ sensitivity rate to $\Delta L$, $t_{ox}$ and $N_s$.

Figure 4: Histogram showing normalized $I_{on}$ sensitivity rate to $\Delta L$, $t_{ox}$ and $N_s$.

Figure 5: Histogram showing normalized $I_{off}$ sensitivity rate to $\Delta L$, $t_{ox}$ and $N_s$.

Figure 6: Normalized variance of $t_{ox}$ vs. normalized variance of $N_s$ for $L_g=10\mu m$. Overlap region A is where specifications are satisfied.

Figure 7: Process design window for gate oxide thickness vs. channel length.

Figure 8: Process design window for channel doping vs. channel length.

Figure 9: Process design window for gate oxide thickness vs. channel doping.