New Drain Current Model for Nano-Meter MOS Transistors On-Chip Threshold Voltage Test

Jinbo Wan and Hans G. Kerkhoff
Testable Design and Test of Integrated Systems Group (TDT), CTIT, University of Twente, Netherlands
Email: J.Wan@utwente.nl, H.G.Kerkhoff@utwente.nl

Abstract—Traditional reliability tests use complicated equipment, like probe stations and semiconductor parameter analyzers, to measure changes in transistors’ threshold voltages, which are both expensive and time consuming. This paper provides an idea to test the threshold voltage with existing low-to-moderate accuracy ADCs and DACs inside SoCs. To avoid the low-accuracy limitation of measurement results, a new MOS model for the nano-meter MOS transistor drain current is proposed. This model only uses six parameters and is valid for all regimes, being the sub-threshold/weak-inversion, moderate-inversion, strong-inversion and linear regime. Measurement results from 90nm transistors and simulation results from 65nm BSIM4.6 models are used to validate the new model. Finally, an on-chip threshold voltage measurement for reliability purpose is proposed and long-time stress measurement for 90nm PMOS transistors are shown.

I. INTRODUCTION

Recently, with metal-oxide-semiconductor (MOS) transistors’ reliability becoming increasingly important, threshold-voltage measurement gains additional emphasis. This is probably because the threshold-voltage is the most sensitive parameter in MOS transistors in terms of reliability. Several dominating ageing effects, like negative/positive bias temperature instability (NBTI/PBTI), hot carrier injection (HCI), time-dependent dielectric breakdown (TDDB) and channel hot-carrier (CHC), are mainly characterized by threshold-voltage shifts. All these ageing effects tend to increase the absolute value of the threshold-voltage with stress time. Therefore, measuring MOS threshold voltage repeatedly during a long time period is interesting to reliability researchers.

Traditionally, threshold-voltages are measured by complicated equipment like probe stations and semiconductor parameter analyzers. It increases the cost and limits the time period which reliability researchers use to observe the threshold-voltage shift. On the other side, billions of system-on-chips (SoCs) are produced every year and used in almost all kinds of application areas. Most of them include both analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) in moderate accuracy (10 to 12 bits) and countless MOS transistors which can be used as device-under-tests (DUTs). If there is a way to measure the MOS threshold-voltage through these on-chip ADCs and DACs, ageing behaviour of these SoCs can not only be monitored continuously, but also be measured in rich detail. Besides reliability, threshold-voltage information in SoCs is also of interest to process variation study in nano-meter complementary metal-oxide-semiconductor (CMOS) technologies.

In this paper, a new MOS model for the nano-meter MOS transistor drain current is proposed in Section II, which is continuously valid from the sub-threshold regime till the strong-inversion regime. The new model reduces both complexity and high-accuracy requirements of threshold-voltage measurements, which make it suitable for extracting the MOS threshold-voltage via on-chip general-purpose ADCs and DACs. Section III provides evidence based on 90nm measurement results and 65nm Berkeley short-channel IGFET model (BSIM) simulation results. In Section IV, the idea of on-chip MOS threshold-voltage measurement is described. A long time stress test together with ageing results will be shown in Section V. Conclusions will be drawn in Section VI.

II. A NEW MOS DRAIN CURRENT MODEL

Since the threshold-voltage will change with body bias, it is worthwhile to define that the threshold-voltage studied in this paper employs a zero body bias threshold-voltage ($V_{\text{th0}}$). It is the threshold-voltage which is measured at zero source to substrate bias. There are numerous methods published in literatures to extract the value of $V_{\text{th0}}$ [1], and various extractor circuits have also been proposed [2] to automatically measure $V_{\text{th0}}$. They normally measure drain currents at various gate (or both gate and drain) voltages and extract the $V_{\text{th0}}$ from these measurement results by compact models. Based on the operation regime of the DUT transistor, those methods and circuits can be divided into two groups: measuring in the sub-threshold regime and measuring in the strong-inversion regime. Measuring in the sub-threshold regime can benefit from the $V_{\text{th0}}$ definition and requires less relative accuracy (< 1 in $10^3$) when compared to measuring in the strong-inversion regime. However, the drain current in the sub-threshold regime is very small, typically in the $nA$ range. JEDEC standard [3] mentions that the measurement system must be able to measure 100$pA$ with a resolution of 1$pA$ or better, which places a tough requirement on equipment resolution and offset calibration. So it is not suitable for on-chip $V_{\text{th0}}$ measurements. Measuring in the strong inversion regime takes advantage of the large drain current and low offset influence. However, the definition of the $V_{\text{th0}}$ is in the sub-threshold regime and to extrapolate the drain current from the strong-inversion regime down to the sub-threshold regime is difficult even by compact models. It depends very much on model accuracy. Industry standard compact models like BSIM 4.6 model [4], PSP model [5] and EKV model
all have the problem to satisfy this work. The BSIM 4.6 model is known to be inaccurate in the sub-threshold regime and more than three hundred model parameters make the \( V_{th0} \) extraction difficult. The PSP model is claimed to be accurate in both the strong-inversion regime and the sub-threshold regime [5]. However, besides hundreds of model parameters, the PSP model requires extremely accurate drain current measurements (relative accuracy < 1 in \( 10^3 \)) [7] because it actually uses the information from derivation of the measured drain current. Such extremely accurate measurements are difficult to obtain by on-chip test circuits. The EKV model is good in the sub-threshold regime, but it loses accuracy in the strong-inversion regime especially for nano-meter MOS transistors. So in order to extract \( V_{th0} \) on-chip based on the drain current measured in the strong-inversion regime, a new MOS model for the nano-meter MOS drain current, which is valid for all MOS operation regimes, is the key to embedded the \( V_{th0} \) test inside SoCs.

Our new MOS model for nano-meter MOS drain current is expressed in (1) to (5).

\[
I_d = F_{sd} - F_{ds} = \beta \cdot \ln^2 \left[ 1 + \exp \left( \frac{V_s - V_i}{2V_T} \right) \right] \cdot \left( 1 + \frac{V_s - V_i}{V_A} \right)
\tag{1}
\]

\[
F_{sd} = \frac{\beta \cdot \ln^2 \left[ 1 + \exp \left( \frac{V_s - V_i}{2V_T} \right) \right]}{\left[ 1 + x \cdot V_s \cdot \ln \left[ 1 + \exp \left( \frac{V_s - V_i}{2V_T} \right) \right] \right]^{\frac{1}{2}}}
\tag{2}
\]

\[
F_{ds} = \frac{\beta \cdot \ln^2 \left[ 1 + \exp \left( \frac{V_s - V_i}{2V_T} \right) \right]}{\left[ 1 + x V_s \cdot \ln \left[ 1 + \exp \left( \frac{V_s - V_i}{2V_T} \right) \right] \right]^{\frac{1}{2}}}
\tag{3}
\]

\[
V_p = V_o - V_{th0}
\tag{4}
\]

\[
V_T = \frac{kT}{q}
\tag{5}
\]

The MOS terminal voltages all refer to the substrate and expressed as \( V_d \), \( V_o \) and \( V_s \). In (1), the model employs a similar idea from EKV model, which is forward current \( F_{sd} \) and backward current \( F_{ds} \). However, unlike the EKV model which the forward current is independent of \( V_d \) and backward current is independent of \( V_s \), the new model forward current (2) and backward current (3) are modulated by both \( V_d \) and \( V_s \). \( V_T \) is the thermal voltage, which is around 26mV at room temperature. Parameters \( k, T \) and \( q \) are the Boltzmann’s constant, temperature in Kelvin and the electrical charge of an electron respectively. \( V_p \) is the pinch-off voltage and \( n \) is the slope factor, which is linked to the weak-inversion slope. \( V_{th0} \) is the zero body bias threshold-voltage which is interesting for this paper. In (2) and (3), \( \beta \) is a fitting parameter for mobility, width-to-length ratio and gate oxide capacitance per unit area, like \( \mu C_{ox} W/L \) in the classic square-law model. Parameters \( x \) and \( \alpha \) take into account the velocity-saturation effect. Drain-induced barrier lowering (DIBL) and the short channel effect are taken into account by parameter \( V_A \).

Explaining the derivation of the model is complex. The model might be better understood using several extreme cases. Assume \( V_s \) is zero and the transistor is operating in the sub-threshold regime with \( V_d \) is very small to reduce DIBL effect (for example 0.05V). This bias condition is normally used for \( V_{th0} \) measurement in the sub-threshold regime. In this situation, \( V_p \) is negative. The sub-threshold condition can then be expressed as (6) to (8).

\[
x \cdot \frac{F_{d,or,s} \cdot \ln \left[ 1 + \exp \left( \frac{V_p - V_{s,or,d}}{2V_T} \right) \right]}{\frac{V_A - V_d}{V_A}} < 1
\tag{6}
\]

\[
\frac{\exp \left( \frac{V_p - V_{s,or,d}}{2V_T} \right)}{\frac{V_A - V_d}{V_A}} < 1
\tag{7}
\]

\[
\left[ \frac{V_A - V_d}{V_A} \right] < 1
\tag{8}
\]

The drain current in sub-threshold can then be simplified to (9), which is the known expression for sub-threshold drain current.

\[
I_d = \beta \cdot \exp \left( \frac{V_p - V_s}{2V_T} \right) - \beta \cdot \exp \left( \frac{V_p - V_d}{V_T} \right)
\tag{9}
\]

In the moderate inversion regime, the velocity saturation effect can be ignored, which means (7) is still valid. If \( V_d \) is larger than the pinch-off voltage \( V_{po} \), the transistor will work in the saturation regime. The reverse current (7) can be ignored in this case. The left side of (6) will be much larger than 1 and hence dominate. The drain current in this situation can be expressed as shown in (10), which is the classic square-law equation.

\[
I_d = \beta \cdot \frac{(V_p - V_s)^2}{2V_T^2} \cdot \left( 1 + \frac{V_d - V_s}{V_A} \right)
\tag{10}
\]

If \( V_d \) is smaller than pinch-off voltage \( V_{po} \), the transistor will work in the linear regime. The reverse current (7) can not be ignored but equation (8) will keep valid. The drain current in this situation can be expressed in (11), which is the classic linear-regime equation.

\[
I_d = \beta \cdot \left( \frac{V_p - V_s}{2V_T} \right)^2 - \beta \cdot \frac{(V_p - V_d)}{2V_T^2}
\tag{11}
\]

In the strong-inversion regime, the velocity saturation effect will dominate. The reverse current can be ignored if \( V_d \) is larger than the pinch-off voltage \( V_{po} \). Both left sides of (6) and (7) will be much larger than 1 and hence dominate. The drain current can be simplified to equation (12) and is linearly increasing with \( V_d \) instead of following a square-law, which is well known for the velocity-saturation influence. The \( V_d^{-1/\alpha} \) originally comes from the fact that the lateral electric field accelerates the velocity saturation occurrence. Now it influences the drain-to-source conductance in a way that the drain-to-source conductance will become lower with increasing of drain-to-source voltage. This change of drain-to-source conductance in the strong-inversion regime is already known.

\[
I_d = \frac{\beta}{2V_T \cdot x^{1/\alpha}} \cdot V_p \cdot \frac{1 + \frac{V_d}{V_A}}{V_A} \cdot V_d^{-1/\alpha}
\tag{12}
\]

For other situations which cannot be simplified, the drain current needs to be calculated by (1) to (5). The entire model only uses 6 parameters, \( n, \beta, x, \alpha, V_A \) and \( V_{th0} \). They are fixed and independent of \( V_d, V_p \) and \( V_s \) values. Using the measured drain currents in the strong-inversion regime, all 6 model parameters can be extracted by minimum the error between the model result and the measurement result.
In order to validate the new MOS model, measurement results from 90nm NMOS transistors are shown in Fig. 1 to 3. Moreover, to enable further study on accuracy, simulation results using TSMC 65nm BSIM4.6 model are provided in Fig. 4 and 5. During 90nm measurements, the body and source contacts of the NMOS transistor are connected to ground. Gate and drain are swept from 0 to 1.2V with 0.05V steps. Fig. 1 shows the drain current changing with drain voltage at different gate voltage level. The new MOS model well fits the measurement.

To further validate the model, the first-order derivation of the drain current, drain conductance \( G_{ds} \) of both measurement and model are shown in Fig. 2. The new model fits the measurement well in most regimes, except the part of the transfer from the linear regime to the strong-inversion regime occurs (at around \( V_d = 0.8V \)). Measurement results in Fig. 1 are used to extract six model parameters. At the same time, a more accurate measurement is done in the sub-threshold regime to check whether the model with parameters extracted from the strong-inversion regime is accurate for the sub-threshold regime. For the sub-threshold measurements, the drain current is measured at least 10 times more accurately than the measurement in Fig. 1. Results show that the model with parameters extracted in the strong-inversion fit measurements in the sub-threshold regime well as shown in Fig. 3. Here it is noted that the noise floor of sub-threshold current measurement is around 1nA.

In reliability measurements, the threshold-voltage shift is more important than the actual threshold-voltage. To validate the new model in reliability measurements, a simulation has been carried out by comparing this model-extracted \( V_{th0} \) with BSIM4.6 model results. The BSIM model parameters come from the TSMC 65nm Process Design Kit (PDK). The BSIM model is simulated in Cadence Spectre with a shift in its \( V_{th0} \) parameter of \( \pm 50mV \). The simulated drain current together with gate and drain voltage are used to extract the new model parameters. Fig. 4 compares the extracted \( V_{th0} \) between the BSIM4.6 model and this model. The x-axis is threshold-voltage shifts from \(-50mV\) to \(+50mV\). The y-axis is the \( V_{th0} \) by both the BSIM4.6 model and this model. They are very close and the error is within 1mV.

In order to realize an on-chip \( V_{th0} \) test in a SoC without adding extra cost, the drain current needs to be measured by already-existing general purpose ADCs in SoCs. These...
general purpose ADCs have normally a 10 to 14 bits accuracy. The drain current measurements can not be as accurate as a semiconductor parameter analyzer. So it is important to check the sensitivity of \(V_{\text{th}0}\), which is extracted via the new model, to the drain current accuracy. Suppose the drain current (whose maximum is below 400\(\mu\)A) is measured by an ADC with a full-scale of 1\(m\)V. By changing the accuracy of the ADC from 7bit to 14bit, the \(V_{\text{th}0}\) extracted by the new model is plotted in Fig. 5. The actual \(V_{\text{th}0}, 443\text{mV},\) is shown as a red solid line for reference. From Fig. 5, if less than 1\(m\)V \(V_{\text{th}0}\) error is required, the ADC is required to be at least 10bit and better to be 12bit.

IV. On-Chip \(V_{\text{th}0}\) Test for Reliability Purpose

The idea to measure the \(V_{\text{th}0}\) on-chip is shown in Fig. 6. Two matched MOS transistors, P1, P2 and one load resistor \(R\), are added into a SoC. The transistor P1 will be stressed, while transistor P2 will be unstressed and works as a reference fresh device. A general purpose ADC and DAC in the SoC are applied to sweep each transistor’s gate voltage and measure the voltage of the load resistor at the same time. The measurement results are processed by the central processing unit (CPU) inside the SoC with the help of the new MOS model. The detailed procedure is described in Fig. 6 and starts from the up-left sub-figure. The gate of the transistor P1 is connected to the DAC to sweep the voltage from the ground to the power supply. The transistor P2 is turned off by connecting its gate to the supply (PMOS). The voltage at the load resistor \(R\) is measured by the ADC at each DAC sweeping step. Now the drain current \(V_\text{d} \) of the P1 can be calculated by the difference of the supply voltage \(V_{\text{dd}}\) and the ADC measured voltage \(V_{\text{ADC}}\). The drain current \(I_\text{d}\) can be calculated as the ADC measured voltage \(V_{\text{ADC}}\) divided by the load resistance \(R\). These can be expressed in (13) and (14).

\[
\begin{align*}
I_\text{d} &= \frac{V_{\text{ADC}}}{R} = F_{\text{sd}} - F_{\text{ds}} \quad (13) \\
V_{\text{ADC}} &= R \cdot F_{\text{sd}} - R \cdot F_{\text{ds}} \quad (14)
\end{align*}
\]

Based on the \(F_{\text{sd}}\) and \(F_{\text{ds}}\) expression in (2) and (3), it is interesting to note that the load resistance \(R\) can be absorbed into the parameter \(\beta\) as one model fitting parameter. So the accurate value of \(R\) is not required to be known in advanced. The six model parameters (including \(V_{\text{th}0}\)) can then be extracted from (14). Now the \(V_{\text{th}0}\) of the P1 is obtained. The same procedure can be applied to the transistor P2 in upper right sub-figure in Fig. 6. Now the P1 is turned off and the gate of P2 is connected to the DAC to be swept. The ADC measures the voltage at the same point. The \(V_{\text{th}0}\) of the P2 can be obtained the same as for the P1. After \(V_{\text{th}0}\) of both P1 and P2 are measured, the transistor P1 is connected to the stress voltage (for example, to the ground in Fig. 6). The transistor P2 is turned off to remove any stress and keeps fresh. The stress will be periodically stopped and measurements will be done for threshold-voltages of the P1 and P2. So Fig. 6 shows a loop for its three sub-figures. \(V_{\text{th}0}\) values together with the age time information can be stored in local memory and monitored by users as a life status indication of the SoC. This information can also be sent via internet to the product company as a database for products reliability studies.

One problem will be the temperature because it will change the \(V_{\text{th}0}\) as well, which sometimes has a stronger influence than ageing. In order to decouple the temperature influence from ageing-induced \(V_{\text{th}0}\) shift, normally an accurate temperature sensor would need to be incorporated. However, in our opinion the temperature effect can be disconnected without a temperature sensor. This is achieved by using a well matched MOS transistor pair as shown in Fig. 6. Ageing mechanisms in nanometer MOS transistors which change \(V_{\text{th}0}\) like NBTI, PBTI, HCI, CHC and TDDB, are actually changing the surface-state charges in the silicon-dioxide. This can be proven from the definition of the threshold-voltage (15). Surface-state charges are \(Q_{ss}\) in (15). It is divided by the gate-channel capacitance \(C_{\text{ox}}\) and then added to the threshold-voltage. No temperature is involved when calculating the threshold-voltage change from the surface-state charges change. As a result, the \(V_{\text{th}0}\) shift by ageing can be simply measured by the difference in \(V_{\text{th}0}\) between the reference transistor and the stressed transistor. They are at the same temperature and matched well. Hence the difference in \(V_{\text{th}0}\) should come from ageing effects (and initial mismatch).

\[
V_{\text{th}0} = V_{FB} + \frac{Q_{ss}}{C_{\text{ox}}} + 2\phi_b + \gamma N\sqrt{2\phi_b} \quad (15)
\]
Another problem is the ageing of the ADC and DAC. Compared to a DAC, the accuracy of an ADC is more important because the DAC voltage can be measured by the ADC. However, there is not much research on ageing behaviour of an ADC. Publication on this topic [8] claims that the integral nonlinearity (INL) and differential nonlinearity (DNL) of a ADC could increase with ageing, or the offset will increase with ageing. An offset increasing can be calibrated by simple open and short structures and is thus less of a problem. Increasing of INL and DNL is a real problem. It will negatively affect the ADC accuracy. However, the new MOS model can tolerate less accurate measurement results as shown in Section III. If the ADC accuracy is not lower than 10 bits, the extracted $V_{th0}$ will not be influenced. If the ADC accuracy is lower than 10 bits, the accuracy of the extracted $V_{th0}$ will be reduced and the exact effect of the INL and DNL increasing to the on-chip $V_{th0}$ test is under research.

V. MEASUREMENT RESULTS

A long-time stress test has been carried out. The aim of the test is to study the ageing behaviour of 90nm MOS transistors, and at the same time, validate the $V_{th0}$ extraction method by the new MOS model. The 90nm test chip is put into an oven and is heated up to 127°C. The test chip contains 32 PMOS DUT transistors and those transistors are supplied with 1.15V during the stress test. The ProChek System from Ridgetop Group Inc. is used for making the measurements [9]. It contains voltage controlled resources that allow to sweep the gate and drain voltage of each PMOS transistor as well as measurement resources that allow to measure drain and gate currents concurrently and simultaneously. The stress test continues for 12 weeks, and will take around 2000 hours. After each week (167 hours), the stress is stopped and measurements are carried out to check the life status of DUTs. Till the writing of this paper, the 2000 hours test is still not finished.

In order to reduce the recovery effect, measurements for all DUTs are finished within 1% of the stress time. Measurements include two kind of tests. The first kind is measuring the DUT drain current in moderate accuracy ($0.1\mu$A) with $V_g$ and $V_d$ sweeping from 1.15V down to 0V. The sweeping step is set be 10mV to gather as many results as possible. The second kind of test is measuring the DUT drain current more accurately (around 1nA) in the sub-threshold regime with $V_d$ fixed at 50mV and $V_g$ sweeping from 1.15V down to 0V. For the same DUT transistor, these two kinds of tests are carried out at the same time (with only 20ms time difference, 0.000003% of the stress time) to avoid recovery effect, and make sure that the $V_{th0}$ extracted from both methods should come from the same physical value.

The $V_{th0}$ is extracted in three ways. First, all results of the drain current measured in moderate accuracy are used to fit the new MOS model and extract the $V_{th0}$. The $V_{th0}$ extracted in this way is named as “$V_{th0}$ by model” and can be used to validate the model accuracy. Second, select one pair of $I_d$ and $V_d$ for each set of $V_g$ sweeping measured in moderate accuracy. The selected sub-group results are used to fit the new MOS model and obtain the $V_{th0}$. The selection of $I_d$ and $V_d$ is based on $I_d = R_{load} \cdot V_d$, in which $R_{load}$ is the load resistor as shown in Fig. 6. By this method, we imitate the measurement results of on-chip $V_{th0}$ test method which is proposed in Section IV, and can validate the on-chip $V_{th0}$ test method with various of $R_{load}$ values. The $V_{th0}$ extracted in this way is named as “$V_{th0}$ by on-chip method”. Finally, from the measured higher accuracy drain current in the sub-threshold regime, the real $V_{th0}$ is calculated and named as “$V_{th0}$ sub-threshold measurement”.

The $V_{th0}$ of 32 PMOS transistors before stress are plotted in Fig. 7. Three kind of $V_{th0}$ values, “$V_{th0}$ sub-threshold measurement”, “$V_{th0}$ by model” and “$V_{th0}$ by on-chip method” are shown in the figure with different colours and symbols. From Fig. 7, the “$V_{th0}$ by model” and “$V_{th0}$ by on-chip method” agree well, because they use the same model. However, the “$V_{th0}$ sub-threshold measurement” will need much less time as it only need to sweep the $V_g$ (around 115 times faster in this case), and can be implemented on-chip. The error between “$V_{th0}$ sub-threshold measurement” and the other two $V_{th0}$ obtained from the new model changes from DUT to DUT. The mean error is within 1mV. The standard deviation of the error is around 3mV. The maximum error of these 32 DUTs is around 8mV.

The change of threshold-voltage, $\Delta V_{th0}$ due to ageing is more interesting to reliability studies. Fig. 8 shows the $\Delta V_{th0}$ by three methods of the same group DUTs after 1 week (167 hours) stress at 127°C. The $\Delta V_{th0}$ by model” and “$\Delta V_{th0}$ by on-chip method” agree well again. The maximum error
The authors like to acknowledge the help of Hans Manhaeve from Ridgetop Europe in this project. Tijl Lammetink and Michiel Schomakers from MASEr Engineering B.V. are credited for assisting in the HTOL stress tests.

REFERENCES


A close-form model for the MOS drain current is provided to facilitate MOS threshold-voltage measurements in nano-meter CMOS technologies. The model is valid for all regimes, like the sub-threshold/weak-inversion, moderate-inversion, strong-inversion and the linear regime. Measurements have carried out in 90nm. Results show that the model well agrees with measurements. Threshold-voltage extraction using this new model can tolerate less accurate measurement results in the strong-inversion regime, which can be obtained by already existing general-purpose ADCs inside SoCs. It will significantly reduce both complexity and cost of reliability measurements. A low overhead on-chip test scheme for ageing-induced threshold-voltage shift measurement is proposed. Long time stress tests show that the on-chip $V_{th0}$ test can characterize threshold-voltage shift with 3mV accuracy.

VI. CONCLUSIONS

The gray lines are $\Delta V_{th0}$, measurement results are double checked of both the strong-inversion drain current and the sub-threshold drain current for DUT No.23. There is an indeed an increasing in the drain current which corresponds to a decrease of $V_{th0}$. The reason could be large random behaviours of ageing in nano-meter MOS transistors which has been indicated in other publications [2], [10].

Fig. 9 and Fig. 10 show the $\Delta V_{th0}$ changing with time. The gray lines are $\Delta V_{th0}$ of each DUT. The average $\Delta V_{th0}$ of these 32 DUTs from “$V_{th0}$ by model” and “$V_{th0}$ by on-chip method” are shown in red and blue lines respectively. The average $\Delta V_{th0}$ from “$V_{th0}$ by on-chip method” agree well with the one from “$V_{th0}$ by model” method in the first, second and fourth weeks. However, “$V_{th0}$ by on-chip method” fail to extract $V_{th0}$ for 5 DUTs in the third week. After changing the load resistor $R_{load}$ value, the problem disappeared. It could be the reason that the “$V_{th0}$ by on-chip method” has a limitation on the range of the $R_{load}$ value. The $R_{load}$ is found to be better to keep PMOS working as much in saturation region as possible while at the same time, can provide as large $V_{th}$ swing as possible. It is similar to the requirement of the output stage of an OpAmp.