Extended Abstracts of the 2003 International Conference on Solid State Devices and Materials, Tokyo, 2003, pp. 420-421

P2-5

A New Stable Extraction of Threshold Voltage Using Regularization Method

Woo Young Choi, Byung Yong Choi, Dong-Soo Woo, Myeong Won Lee, Jong Duk Lee, and Byung-Gook Park

Inter-university Semiconductor Research Center and School of Electrical Engineering, Seoul National University
San 56-1, Shinlim-dong, Kwanak-gu, Seoul 151-742, Korea
Phone: +82-2-880-7279 Fax: +82-2-882-4658 E-mail: claritas@korea.com

1. Introduction

Threshold voltage (V_{TH}) is one of the most important parameters for MOSFET characterization. Many V_{TH} extraction methods have been proposed to obtain physically-meaningful V_{TH} from the I_{DS} versus V_{GS} characteristics, such as split C-V, linear extrapolation (LE) and transconductance change (TC) method. Among them, only the TC method can yield a result that approaches the classically-defined V_{TH} [1]. Moreover, it is useful for nanoscale MOSFETs because it eliminates the effects of the interface state, the mobility degradation and the parasitic resistance [2]. Especially in the case of double-gate MOSFETs, the TC method reflects their volume inversion behavior well [3]. However, since in this method the second derivative of I_{DS} is required, it tends to be very noisy [4]. Numerical differentiation is known to be unstable in that small perturbations of the function to be differentiated may lead to large errors in the computed derivative. In simulation or measurement, most of errors come from round-off and truncation. There is always a trade-off: as nodes are set to be denser, data reflect the rapid variation better while differentiation of the data results in more noise [5]. Fig. 1 shows the importance of node interval selection. If we measure g_{m2} with 0.05 percent accuracy, noise and detail loss will depend on node interval. When ΔV_{GS} is set to be 0.01 V, it is difficult to find g_{m2} peak due to noise. If we increase ΔV_{GS} up to 0.1 V to remove noise, data between nodes will be omitted and then g_{m2} peak becomes flat. In both cases, we cannot detect the exact V_{TH}. Thus, we need to optimize the interval to get less noisy g_{m2} profiles with minimum detail loss. In this paper, we propose a stable extraction method for the V_{TH} defined by the TC. Adopting it, we can determine V_{GS} where the surface potential is within kT/q of $\phi_s = 2\phi_f + V_{SB}$. The knowledge of that makes it possible to calculate ϕ_s at any other point on the I_{DS} versus V_{GS} curve, which will provide a big help for device analysis.

2. Device Design

In the TC method, the V_{TH} is defined as the V_{GS} at which the derivative of the low drain voltage transconductance dg_m/dV_{GS} (= g_{m2}) is maximum. Therefore, smooth g_{m2} profiles without noise lead to the exact V_{TH}. First, the optimal node interval for g_m will be derived. When the I_{DS} is simulated or measured, some errors result from round-off and truncation. We take them into account by introducing an absolute error δ. The exact drain current becomes $I_{DS} = I_{DSm} + \delta$, where I_{DSm} is extracted drain current. Considering δ, we can get Eq. *(1)*. From *(1)*, Eq. *(2)* can be derived. The relative error is derived as depicted in Eq. *(3)*. It may look problematic because δ is unknown. However, in practice, we can get an error bound for δ, that is, a number β such that $|\delta| \leq \beta$, where β represents a characteristic sum of errors. Eq. *(3)* becomes Eq. *(4)*. If we want less than one-percent error in g_m, the condition is derived as shown in Eq. *(5)*. For g_m is approximated to $\Delta I_{DS} / \Delta V_{GS}$, the optimal interval for g_m is defined as in Eq. *(6)*. Referring to results above, the optimal node interval for g_{m2} is obtained. Assuming $g_m = g_{mm} + \varepsilon$, the case is the same. g_m is the true value of dI_{DS} / dV_{GS}, g_{mm} is the dI_{DS} / dV_{GS} from Eq. *(1)* to *(6)* and ε is the calculation error. Note that g_{mm} means g_m in Eq. *(1)* to *(6)*. We obtained the optimal node interval for g_{m2} as shown in Eq. *(7)* where γ means an error bound for ε, which requires $|\delta| \leq \gamma$. From the condition of one-percent error above, it is found that γ is equal to 0.01 times g_m. Thus, Eq. *(7)* can be rewritten into Eq. *(8)*. To profile g_{m2} with minimum loss of details, conditions *(6)* and *(8)* should be satisfied at the same time. Finally, the optimal node interval for accurate g_{m2} becomes Eq. *(9)*. By this criterion, noise-free g_{m2} is obtained within one percent error.

The condition *(9)* is applied to the simulation by estimating next node interval from g_m and g_{m2} derived referring to previous three data nodes. Fig. 2 shows the flow chart for optimal node extraction. When initial values of I_{DS}, g_m and g_{m2} are given, the optimal interval for next V_{GS} is calculated according to Eq. *(9)*. After obtaining the drain current value of the given V_{GS}, we perform the second order differentiation to extract noise-free g_{m2} without detail loss. Then, the values of I_{DS}, g_m and g_{m2} are fed back into the first stage. The algorithm leads to g_{m2} profiles satisfying one-percent noise criterion as shown in Fig. 3. The simulation was done to a 1.5 ㎛ nMOSFET at 0.1 V drain bias. Considering the noisy profiles in Fig. 1, the improvement is prominent. $V_{TH}(P)$, $V_{TH}(LE)$ and $V_{TH}(TC)$ represent the V_{TH} of the classical definition ($\phi_s = 2\phi_f + V_{SB}$), the LE and the TC method, respectively. $V_{TH}(LE)$ is selected because it is *defacto* industry standard. $V_{TH}(TC)$ extracted by our algorithm approaches $V_{TH}(P)$. To generalize our algorithm, we compared three V_{TH}*'s* in various channel lengths. In long channel, $V_{TH}(TC)$ predicts $V_{TH}(P)$ within kT/q as shown in Fig. 4. The same is the case in short channel, as depicted in Fig. 5. Moreover, $V_{TH}(TC)$ reflects more precise V_{TH} roll-off than $V_{TH}(LE)$. To make this behavior clear, correlation coefficients are calculated. The correlation coefficient R_{xy} for x and y is defined in Eq. *(10)*. Table 2 shows better correlation between $V_{TH}(P)$ and $V_{TH}(TC)$. We also performed the fast Fourier transform to confirm that our algorithm extracts noiseless g_{m2} profiles without detail loss. Fig. 6 proves that it eliminates noise components without loss of g_{m2}. Noise in high frequency range are suppressed as intervals are optimized, while data in low frequency range remain the same. If intervals exceed optimum, the peak in low frequency will be lowered due to data loss.

3. Conclusions

Consequently, it is found that our algorithm extracts noiseless g_{m2} profiles without detail loss and that $V_{TH}(TC)$ derived from the g_{m2} profiles predicts $V_{TH}(P)$ within kT/q. This V_{TH} extraction reflects the V_{TH} roll-off of nanoscale MOSFETs accurately and makes it possible to calculate the ϕ_s at any other point on the I_{DS} versus V_{GS} curve. Finally, the algorithm will endow us with powerful methodology in device modeling and characterization.

Acknowledgements

This work was supported by the BK21 program, by the National Research Laboratory Project of Ministry of Science and Technology and by the Collaborative Project for Excellence in Basic System IC Technology.

References

[1] H.-S. Wong *et. al*, *Solid-State Electronics*, vol.30, no.9, pp.953-968, 1987.

[2] R. V. Booth *et. al*, *IEEE Transactions on. Electron Devices*, vol. ED-34, no. 12, pp.2501-2509, Dec. 1987.

[3] P. Francis *et. al*, *IEEE Transactions on Electron Devices*, vol. ED-41, no. 5, pp.715-720, May 1994.

[4] N. Arora, MOSFET models for VLSI circuit simulation, Wien: Springer-Verlag, 1993.

[5] C. F. Gerald *et. al*, Applied Numerical Analysis, fourth edition, Massachusetts: Addison-Wesley, 1989.

Table 1. Equations for optimal node interval derivation.

$g_m + dg_m = dI_{DSm} / dV_{GS} = d(I_{DS} - \delta) / dV_{GS} = dI_{DS} / dV_{GS} - d\delta / dV_{GS}$.	(1)
$dg_m = - d\delta / dV_{GS}$.	(2)
$dg_m / g_m = - (d\delta / dV_{GS}) / (dI_{DS} / dV_{GS}) = -d\delta / dI_{DS} \approx \Delta\delta / \Delta I_{DS}$.	(3)
$dg_m / g_m \approx \Delta\delta / \Delta I_{DS} = 2\beta / \Delta I_{DS}$.	(4)
$\Delta I_{DS} \geq 200\beta$	(5)
$\Delta V_{GS} = \Delta I_{DS} / g_m \geq 200\beta / g_m$	(6)
$\Delta V_{GS} = \Delta g_m / g_{m2} \geq 200\gamma / g_{m2}$	(7)
$\Delta V_{GS} \geq 2g_m / g_{m2}$	(8)
$\boxed{\Delta V_{GS} \geq \max (200\beta / g_m, 2g_m / g_{m2})}$	(9)
$R_{xy} = cov(x,y)/\sigma_x\sigma_y$	(10)

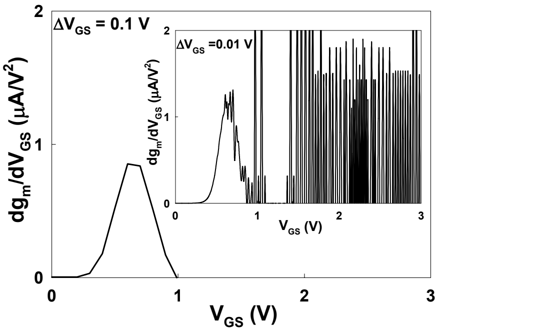

Fig. 1. Comparison of g_{m2} when node interval is 0.1 V. Inset figure shows 0.01 V case.

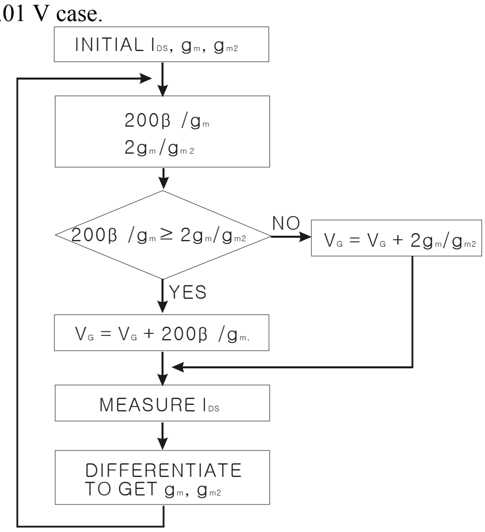

Fig. 2. Algorithm for optimized node interval derivation. Referring to previous three data points, ΔV_{GS} is determined.

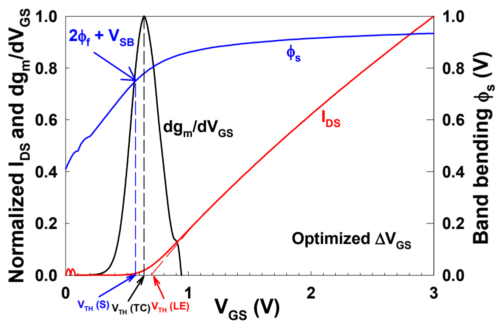

Fig. 3. Optimized g_{m2} profiles calculated from the algorithm for optimized node interval. (1.5 ㎛ nMOSFET)

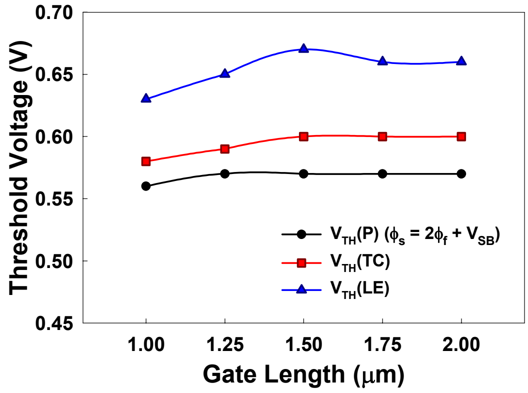

Fig. 4. Comparison of threshold voltage in long channel MOSFETs. $V_{TH}(TC)$ shows better correlation to $V_{TH}(P)$ than $V_{TH}(LE)$.

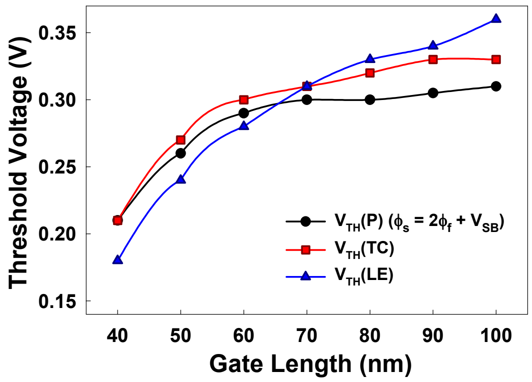

Fig. 5. Comparison of each threshold voltage in short channel MOSFETs. $V_{TH}(TC)$ shows better correlation to $V_{TH}(P)$ than $V_{TH}(LE)$. In addition, $V_{TH}(TC)$ reflects more precise threshold voltage roll-off characteristics than $V_{TH}(LE)$.

Table 2. Correlation coefficients calculated from the pairs of the threshold voltages in Fig. 4 and Fig. 5.

	Correlation coefficient between $V_{TH}(P)$ and $V_{TH}(TC)$	Correlation coefficient between $V_{TH}(P)$ and $V_{TH}(LE)$
Long channel MOSFETs	0.953	0.884
Short channel MOSFETs	0.994	0.945

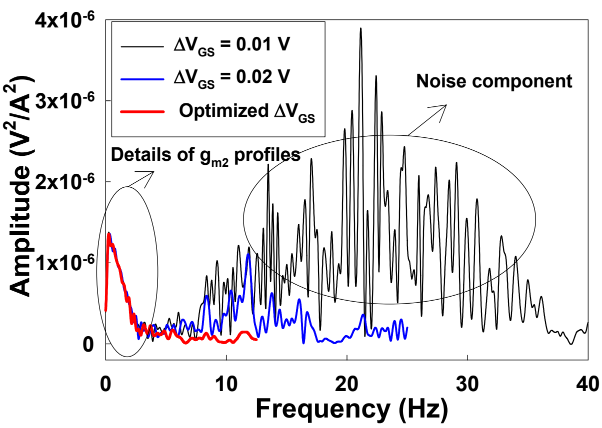

Fig. 6. Fast Fourier transform of g_{m2} profiles with variation of node intervals. Noise components in high frequency range are suppressed as node intervals get optimized, while detailed data in low frequency range remain the same.