Proceedings of the 3rd International Conference on Materials Chemistry and Environmental Engineering (CONF-MCEE 2023), Part II
DOI: 10.54254/2755-2721/7/20230423

Overcoming sub-threshold slope degradation in sub 10 nm technologies

Peiyin Xu

[1]School of Microelectronics, Fudan University, Shanghai, 200433, China

*pyxu20@fudan.edu.cn

Abstract. Sub-threshold slope degradation is one of the difficulties preventing microchip scaling. This paper mainly concentrates on sub-threshold region of MOSFET and introduces the definition of sub-threshold swing (inverse of sub-threshold slope). The reason for this work is also told. Basing on ideal assumptions, the work considers 8 influence factors of sub-threshold. After comparison, high-k material, lower doping density, higher bias voltage and more ideal surface are good ways to solve the problem. Changing work function is conditional. The result is meaningful for further understanding of sub-threshold region and further scaling-down.

Keywords: MOSFET; sub-threshold region; sub-threshold slope; scaling-down.

1. Introduction: sub 10 nm technology

With the rapid development of IC design and fabrication, the scale of transistors goes down exponentially for years. According to Moore's law, the number of transistors on a microchip doubles about every 18 months, which also improves the performance of the processors and reduces power consumption. Nowadays, microchip technology has already reached 10 nm (gate length in MOSFETs). In the near future, sub 10 nm technology will still be the most cutting-edge technology.

Sub 10 nm technology has been widely known so far. However, there exist plenty of difficulties in sub 10 nm technologies. Sub-threshold slope degradation is one of the tough problems among sub 10 nm technologies. This paper is mainly for those who have already learned the basic knowledge of transistors, especially MOSFET, to get deeper understanding of sub-threshold slope degradation and the solution. In terms of IC industry, overcoming sub-threshold slope degradation is rather meaningful for further scaling-down.

2. Sub-threshold characteristic of MOSFET

2.1. Sub-threshold phenomenon

When MOSFET is working, the voltage of gate and source (V_{GS})is higher than threshold voltage of MOSFET (V_T). However, if the gate bias is slightly lower than V_T, the semiconductor surface is in weak inversion, and the current is called sub-threshold current.

Figure 1 shows $V_g - I_{ds}$ characteristic of MOSFET, I_{ds} is expressed both in linear plot and semi log plot.

In linear plot, I_{ds} is relatively small, near 0 when in sub-threshold region. To make it clear, we apply

DOI: 10.54254/2755-2721/7/20230423

semi log plot to amplify the figure. In semi log plot, the characteristic is linear, which means I_{ds} rises exponentially as V_g rises.

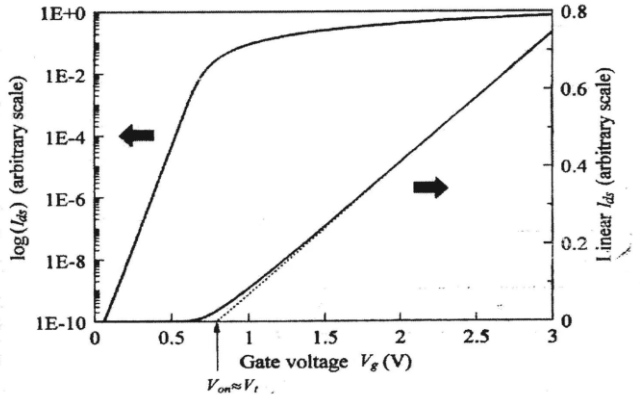

Figure 1. $V_g - I_{ds}$ characteristic of MOSFET.

2.2. Carriers' distribution & fiffusion current of sub-threshold region

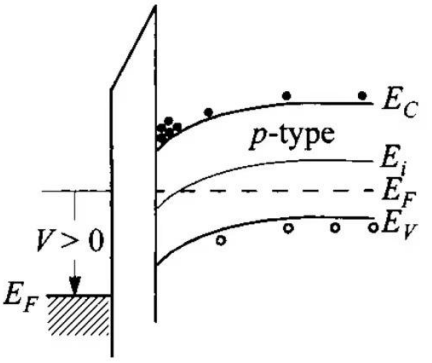

Figure 2. Energy band of MIS capacitor (p-type semiconductor).

Figure 2 is the energy band of MIS capacitor, when the gate voltage V_S satisfies $V_B < V_S < 2V_B$, CB (Conduction Band) and VB (Valence Band) will bend so that E_i on the surface is lower than E_F, which leads to n-type surface. Because E_F is close toE_i, the density of electrons is low.

In weak inversion, the density of carriers is relatively low. A simple model of drift current density is

$$j_{drift} = nq\mu E \tag{1}$$

The density n is so small that drift current is ignorable. But the concentration gradient cannot be ignored. Thus, In sub-threshold region, diffusion current matters.

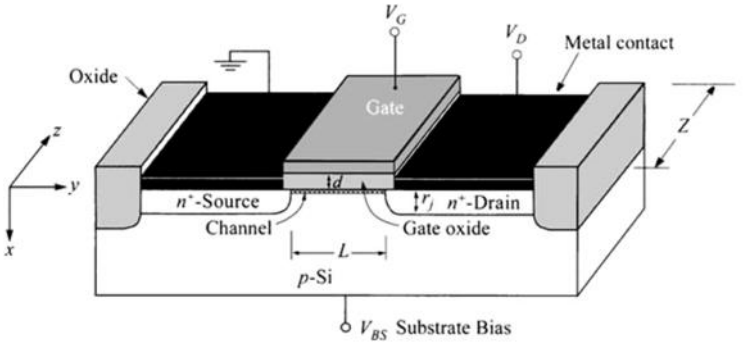

Figure 3. Structure of MOSEFET.

DOI: 10.54254/2755-2721/7/20230423

Figure 3 shows the structure of MOSFET, we define the orientation from source to drain as y. The diffusion current of y-orientation is given by

$$I_y = qD_nA\frac{dn}{dy} \tag{2}$$

For N channel MOSFET. A is the sectional area of current flow and D_n is diffusion coefficient. Because the current is continuous from source to drain.

$$\frac{\partial I_y}{dy} = 0 \tag{3}$$

I_y is constant.

To simplify, we can assume $y-n$ characteristic is linear, and now I_y is given by

$$I_y = qD_nA\frac{n(L)-n(0)}{L} \tag{4}$$

A is the sectional area of current flow

$$A = W \cdot d_{ch} \tag{5}$$

and W is known,

Assume inversion electron distribution is ideal in x-orientation and set y fixed.

$$n(x,y) = n(0,y)\exp\left({}^{-x}/_{d_{ch}}\right) \tag{6}$$

d_{ch} can be written as

$$d_{ch} = \frac{kT}{qE_s} \tag{7}$$

, in which E_s is the electric field of surface.

According to Gauss's law. E_s is known as

$$E_s = \frac{-Q_B}{\varepsilon_s} = \frac{1}{\varepsilon_s}qN_A \cdot \sqrt{\frac{2\varepsilon_s V_s}{qN_A}} = \sqrt{\frac{2qN_AV_s}{\varepsilon_s}} \tag{8}$$

Put formula (8) into (7) and (5), and A is now known.

According to Boltzmann distribution, the density of electrons is

$$n(0) = n_p^0\exp\left(\frac{qV_s}{kT}\right) \tag{9}$$

$$n(L) = n_p^0\exp\left(\frac{q(V_s-V_{DS})}{kT}\right) \tag{10}$$

Then put formula (5)(9)(10) into (4). According to definition, I_{DS} is reverse of I_y ,so I_{DS} can be written as

$$I_{DS} \equiv -I_y = \frac{W\mu_n}{L}\left(\frac{kT}{q}\right)^2 q\left(\frac{\varepsilon_s}{2qN_AV_s}\right)^{\frac{1}{2}}\frac{{n_i}^2}{N_A}\exp\left(\frac{qV_s}{kT}\right)\left[1-\exp\left(\frac{-qV_{DS}}{kT}\right)\right] \tag{11}$$

In order to get the function $I_{DS} = f(V_{GS})$, it is needed to get the relation between V_{GS} and V_S.

The gate-source voltage is the sum of semiconductor surface voltage and oxide voltage

$$V_{GS} = V_s + V_{ox} \tag{12}$$

The surface electric field is

$$E_s = \sqrt{\frac{2qN_AV_s}{\varepsilon_s}} \tag{13}$$

Electric flux is continuous from oxide to surface:

DOI: 10.54254/2755-2721/7/20230423

$$\varepsilon_{ox}E_{ox} = \varepsilon_s E_s \tag{14}$$

The oxide voltage can be written as:

$$V_{ox} = t_{ox}E_{ox} \tag{15}$$

Simulate equation (12)(13)(14)(15), the relation between V_{GS} and V_s is given by

$$V_s = V_{GS} - \frac{B}{C_{ox}}\left[\left(1 + 2\frac{C_{ox}}{B}V_{GS}\right)^{\frac{1}{2}} - 1\right] \tag{16}$$

$$\text{and } B = \frac{\varepsilon_s}{\varepsilon_{ox}} q N_A t_{ox}$$

Simulate equation (11) and (16), and the relation between V_{GS} and I_{DS} is gained.

2.3. Sub-threshold slope

Sub-threshold slope is an important parameter to measure how sharply MOSFET is turned off by the gate voltage. Sub-threshold swing, inverse of sub-threshold slope, is widely used and defined as

$$S = \frac{dV_{GS}}{d(log I_{DS})} \tag{17}$$

S shows the gate voltage change needed to induce a drain current change of one order of magnitude.

3. Why is it important to overcome sub-threshold slope degradation in sub 10 nm technologies

3.1. Application of sub-threshold region: MOSFET as a switch

In sub-threshold region, the current drops sharply with gate bias while gate bias is relatively low. The characteristic is particularly fit for low-voltage, low-power applications such as switch in digital logic and memory applications [1].

For switches, sub-threshold slope should be sharper in order to thoroughly turn on and off. Because of the demand for lower consumption, the high level of voltage on transistors is usually from 1V to 5V. So sub-threshold swing S should be low (sub-threshold slope should be high) in order to get a sharp change of current based on little change of voltage.

Table 1. Different voltage level in digital applications.

	V_{CC}	V_{ih}	V_{il}	V_{oh}	V_{ol}
TTL	5.0V	2.0V	0.8V	2.4V	0.5V
LVTTL	3.3V	2.0V	0.8V	2.4V	0.4V
LVTTL	2.5V	1.7V	0.7V	2.0V	0.2V
LVTTL	1.8V	1.17V	0.63V	1.35V	0.45V
CMOS	5.0V	3.5V	1.5V	4.45V	0.5V
LVCMOS	3.3V	2.0V	0.8V	2.4V	0.4V
LVCMOS	2.5V	1.7V	0.7V	2.0V	0.4V
LVCMOS	1.8V	1.17V	0.63V	1.35V	0.45V

Table 1 shows some digital applications and their respect voltage levels. Some units are based on bipolar transistors and the other are based on MOSFETs. For the command of low power consumption, all the voltage levels are relatively low. Set CMOS for example, the smallest voltage gap between switch "on" and "off" is about 0.5V. If the sub-threshold swing S is too high, the current change will be small, which leads to the mix of "on" and "off".

3.2. Sub-threshold slope degradation in sub 10 nm technologies

Formula (11) shows the relation between I_{DS} and V_s

$$I_{DS} \equiv -I_y = \frac{W\mu_n}{L}\left(\frac{kT}{q}\right)^2 q\left(\frac{\varepsilon_s}{2qN_A V_s}\right)^{\frac{1}{2}} \frac{{n_i}^2}{N_A} \exp\left(\frac{qV_s}{kT}\right)\left[1 - \exp\left(\frac{-qV_{DS}}{kT}\right)\right] \tag{11}$$

DOI: 10.54254/2755-2721/7/20230423

$$\text{Moreover, } V_s = V_{GS} - \frac{B}{C_{ox}}\left[\left(1 + 2\frac{C_{ox}}{B}V_{GS}\right)^{\frac{1}{2}} - 1\right] \tag{16}$$

Fix the gate-source voltage V_{GS}, and the only variate in (11) is L. Consider the influence of L and ignore the influence of $\sqrt{\frac{\varepsilon_S}{2qN_A V_s}}$, I_{DS} is given by

$$I_{DS} \propto \frac{1}{L}\exp\left(\frac{qV_s}{kT}\right) \tag{18}$$

Rewrite the definition of S,

$$S = \frac{dV_{GS}}{d(logI_{DS})} = ln10 \cdot \frac{dV_{GS}}{d(lnI_{DS})} = ln10 \cdot \frac{dV_{GS}}{-\frac{1}{L}dL + \frac{q \cdot dV_S}{kT}} \tag{19}$$

In this formula, it is clear that smaller L will cause higher S.

Figure 4 shows $V_{gs} - I_d$ characteristic with different L simulated by DESSIS, and the model fits the simulation result.

It should be noticed that the present work deals with subthreshold characteristics only. Therefore, there is a large discrepancy between calculated results and simulation results above the subthreshold region [2].

Generally, smaller scale comes with higher S, leading to worse switching performance. That is because the decrease of gate length weakens the control of channel by gate, and the switching performance will be worse. That's why overcoming sub-threshold slope degradation is rather important in sub 10 nm technologies.

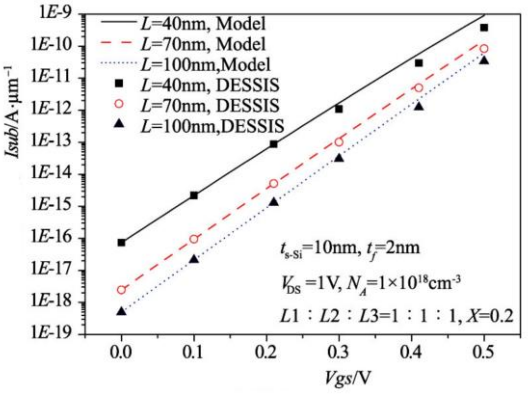

Figure 4. $V_{gs} - I_d$ Characteristic with different L.

4. How to improve sub-threshold slope

4.1. Different influence factors

In order to overcome sub-threshold slope degradation, the first thing is to figure out influence factors of S. There will be quantitative analysis and qualitative analysis.

(1) According to the definition of S, it is convenient to rewrite the formula (17)

$$S = \frac{dV_{GS}}{d(logI_{DS})} = ln10 \cdot \frac{dV_{GS}}{d(lnI_{DS})} = \frac{kT}{q}ln10 \cdot \frac{dV_{GS}}{dV_s} \tag{20}$$

Derivative of gate-source voltage is given by

$$dV_{GS} = dV_s + dV_{ox} = dV_s + \frac{C_d + C_{N_{ss}}}{C_{ox}}dV_s \tag{21}$$

C_d is the depletion capacitor of semiconductor, C_{ox} is the oxide capacitor and $C_{N_{ss}}$ is the capacitor of surface state.

DOI: 10.54254/2755-2721/7/20230423

Finally sub-threshold swing S is given by

$$S = \frac{kT}{q} ln10 \cdot \frac{dV_{GS}}{dV_S} == \frac{kT}{q} ln10 \cdot (1 + \frac{C_d + C_{N_{ss}}}{C_{ox}}) \quad (22)$$

From formula (22), S is controlled by 4 factors.

For T,temperature: Most of MOSFET applications work under constant temperature (usually about room temperature). Although S has direct ratio with T, it is not a feasible way to enhance MOSFET performance.

For C_{ox},oxide capacitor: As C_{ox} is denominator, increasing oxide capacitor will decrease S, making the better performance. In a simple model, oxide capacitor can be regarded as plane-parallel capacitor. However, the distance between two plates cannot change arbitrarily. We can only apply different materials. In this case, high-k material is useful to increase the capacitor, and consequently enhance the performance.

For C_d,depletion capacitor: There is positive correlation between S and C_d. Depletion capacitor can also be regarded as plane-parallel capacitor. Moreover, the length of depletion depends on doping density and substrate bias. If doping density is lower, the length of depletion will be longer, leading to lower C_d and lower S. Substrate bias effect also influences depletion. If bias effect voltage is higher, the depletion region will be narrower, leading to lower C_d and lower S.

For $C_{N_{ss}}$,surface state capacitor: Surface state capacitor depends on the density of surface. If the surface is more ideal, the density of surface state $C_{N_{ss}}$ will go down, then causing lower capacitor and lower S.

(2) Aside from T, C_{ox} , C_d and $C_{N_{ss}}$, there are also some factors missed in the formula (22). Because the demonstration of (22) is based on ideal assumption. There are other factors like L, V_{GS} and φ_M(work function of metal).

For L, gate length: We have already discussed it in part **3.2**. The conclusion is that L is counteractive.

For V_{GS}, gate-source voltage: V_{GS} is independent variable in $V_{GS} - I_D$ characteristic of MOSFET. According to the definition of S. It is clear that lower V_{GS} can result in lower S. This characteristic tells us that low voltage not only benefits power consumption, but also leads to better sub-threshold performance. However, the voltage mostly depends on the device itself, so V_{GS} is usually fixed.

For φ_M, work function of metal: If the work function difference between metal and semiconductor is greater, the flat band voltage of control gate will be higher, which leads to the sharp decrease of I_D [3]. So the greater work function difference leads to lower S and better performance.

Figure 5 shows the result of different work function. However, this model is based on the assumption of sub-threshold region or weak inversion, more exactly. When V_{GS} is high that surface is in strong inversion, the model doesn't work anymore.

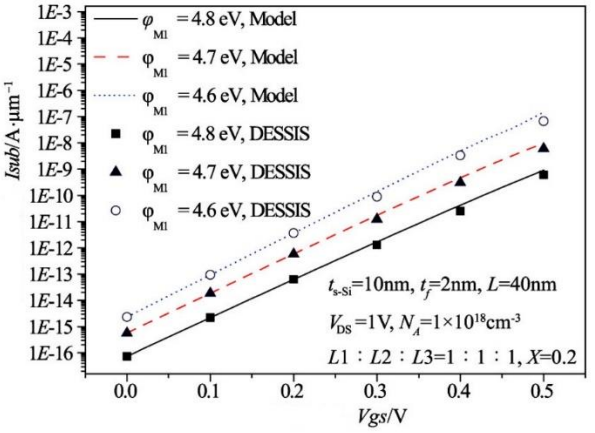

Figure 5. Vgs-Id Characteristic with different work function.

DOI: 10.54254/2755-2721/7/20230423

4.2. Compare different ways

After demonstration and analysis, we can firstly ignore some factors that cannot change: T, L and V_{GS}.

There are 5 ways left, number them ① to ⑤.

① High-k material: Apply high-k insulator like HfO_2 to take the place of SiO_2.

But some high-k material is not fit for MOSFET because some of them lead to more defects on the surface. Overall consideration is important for different occasions.

② Lower doping density: It seems to be a simple way to change the doping density for better sub-threshold performance. But doping density is rather important in MOSFET, causing great change of other parameters like threshold voltage. It also takes overall consideration.

③ Higher bias voltage: Higher bias voltage leads to better sub-threshold slope. Meanwhile, bias voltage will change threshold voltage. So, designers should find the balance between sub-threshold and threshold region.

④ More ideal surface: Surface state mainly depends on the material and fabrication. But material also influences ①, there should be a balance.

⑤ Greater work function difference: The model is conditional, only fits weak inversion.

5. Conclusion

This paper concentrates on sub-threshold region and sub-threshold slope in sub 10 nm technologies, starting from the definition of sub-threshold region and sub-threshold slope, then demonstrate the model and find out influence factors of sub-threshold slope.

After demonstration and analysis, T, L and V_{GS} cannot be changed arbitrarily.

High-k material, Lower doping density, Higher bias voltage and more ideal surface are good ways. But overall consideration is important because the parameters also influence other performance. Changing work function of metal is conditional, only for low voltage (weak inversion).

Reference

[1] S.M.Sze, Kwok K.Ng. (2007) Physics of Semiconductor Devices. A John Wiley&Sons, Inc., Hoboken,New Jersey.

[2] Gopal Rawat,Sanjay Kumar,Ekta Goel,Mirgender Kumar,Sarvesh Dubey,S.Jit.Analytical modeling of subthreshold current and subthreshold swing of Gaussiandoped strained-Si-on-insulator MOSFETs[J].Journal of Semiconductors,2014,35(08):56-63.

[3] XIN Yan-hui，YUAN He-cai，XIN Yang. Analytical Models of Subthreshold Current and Subthreshold Slope for Symmetrical Triple-Material Double-Gate s-Si MOSFETs[J]. ACTA ELECTRONICA SINICA, 2018,46(11):2768-2772.