The influence of nitride thickness variations on the switching speed of MNOS memory transistors

Bruun, Erik

Published in:
I E E E Transactions on Electron Devices

Link to article, DOI:
10.1109/T-ED.1978.19275

Publication date:
1978

Document Version
Publisher's PDF, also known as Version of record

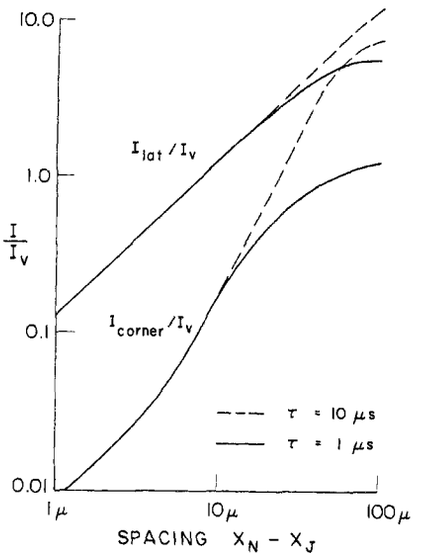

Fig. 2. Ratios of lateral current (I_{lat}) and corner current (I_{corner}) normalized with respect to vertical current (I_v), plotted versus spacing $X_N - X_j$ between p^+-n junction and isolation diffusion for lifetimes of 10 and 1 μs. The curves are for a rectangular diffused diode measuring 80 μm × 20 μm and are obtained by numerical integration of (4).

kind of order zero. The equation may also be solved simply using numerical integration, by choosing an initial guess for p at $r = X_j$ and iterating on this value until the boundary condition $I_p = 0$ at $r = X_N$ is satisfied. However, in the special case where the n^+ diffusion is within one diffusion length of the p^+-n junction, the corner current may be calculated quite simply by setting $L_p = \infty$ in (4). This yields

$$I_{corner} = q(X_N^2 - X_j^2) X_{epi} p(0)/\tau. \quad (5)$$

We recall that in the case where $X_N - X_j \gg L_p$, the result given in [2] may be used

$$I_{corner} = K_1 qp(0) X_{epi} D_p \quad (6)$$

where K_1 is of order unity and nearly independent of lifetime.

III. Results for Typical Structure

In order to illustrate the significance of the relative contributions of each of the three components (vertical, lateral, and corner) we show in Fig. 2 results obtained by numerical integration of (4) for a structure made by diffusing a rectangular p^+ region 20 μm × 80 μm to a depth of 1 μm on an epitaxial layer of thickness $X_{epi} = 5$ μm. In interpreting these results it should be noted that for all practical purposes they are almost independent of the value of X_{epi} (unless X_j becomes comparable to X_{epi}); in addition, it should be remembered that the lateral current is proportional to $2(B + L)$ and the vertical current is proportional to BL. Thus for a square diffusion 15 μm × 15 μm, the I_{lat}/I_v curve would shift upwards by about a factor of two and the I_{corner}/I_v curve would shift upwards by a factor of about eight.

For $X_N - X_j \gg L_p$, all the curves tend to a constant value as predicted by the results in [1], [2]. For $X_N - X_j \ll L_p$, the I_{lat}/I_v curve shows a linear dependence on spacing between the junction and the n^+ isolation diffusion. The linear dependence also exists for the corner current for low values of spacing (X_N close to X_j); this changes to a square dependence for $X_N \gg X_j$ as predicted by (5).

As an example of a particular case, consider a 20-μm spacing with a hole lifetime of 10 μs in the epitaxial layer. The results in Fig. 2 indicate that the lateral current is about twice the vertical current and the corner current is about 80 percent of the vertical current. This would represent a total reduction in current gain of an IIL transistor of about three (with the simplifying assumption that the gain is originally limited only by recombination in the epitaxial layer under the p^+ diffusion). Notice that for higher lifetimes (for the same spacing) the ratios remain unchanged.

IV. Conclusions

The above results enable rapid estimation of "excess" currents due to sideways diffusion of carriers from a p^+-n junction in the presence of a deep n^+ isolation diffusion. The additional components of current are of importance not only in diode design but also in gain calculations for IIL structures with n^+ isolation diffusions at a finite distance from the p diffused base. In this respect, it must be noted that the n-n^+ interface has been assumed to be a plane of zero hole current. For very small spacing, it is clear that in the limit, the lateral and corner currents discussed in this correspondence tend to zero (as seen in Fig. 1) and the nonzero contribution arising from direct injection into the n^+ region would have to be considered. The result given by [3, eq. (16)] could be used to estimate the importance of this current.

Acknowledgment

Invaluable discussions with M. Depey of SESCOSEM, Thomson-CSF, Saint-Egrève, France, are gratefully acknowledged.

References

[1] C. A. Grimbergen, "The influence of geometry on the interpretation of the current in epitaxial diodes," *Solid State Electron.*, vol. 19, pp. 1033–1037, 1976.
[2] D. J. Roulston, M. H. Elsaid, M. Lau, and L. A. K. Watt, "Corner currents in rectangular diffused p^+-n-n^+ diodes," *IEEE Trans. Electron Devices*, vol. ED-25, pp. 392–393, Mar. 1978.
[3] M. H. Elsaid, D. J. Roulston, and L. A. K. Watt, "Vertical current components in integrated injection logic," *IEEE Trans. Electron Devices*, vol. ED-24, pp. 643–647, June 1977.

The Influence of Nitride Thickness Variations on the Switching Speed of MNOS Memory Transistors

ERIK BRUUN

Abstract—**The influence of nitride thickness variations on the switching speed of MNOS memory transistors is examined. The switching time constant is calculated as a function of the nitride thickness using a model of modified Fowler-Nordheim injection. The calculated characteristics compare well with measured characteristics and show a strong dependence on the nitride thickness.**

I. Introduction

In the literature on MNOS memory transistors considerable attention has been paid to the impact of the oxide thickness on the switching speed [1], [2]. The influence of the nitride thickness has been somewhat neglected because the switching properties are normally examined with the nitride field (or the oxide field) as the externally applied parameter. However, for a practical memory circuit it is necessary to specify the gate voltage rather than the gate insulator fields. With a fixed value

Manuscript received February 15, 1978; revised March 3, 1978.
The author is with the Laboratory for Semiconductor Technology, Technical University of Denmark, DK-2800 Lyngby, Denmark.

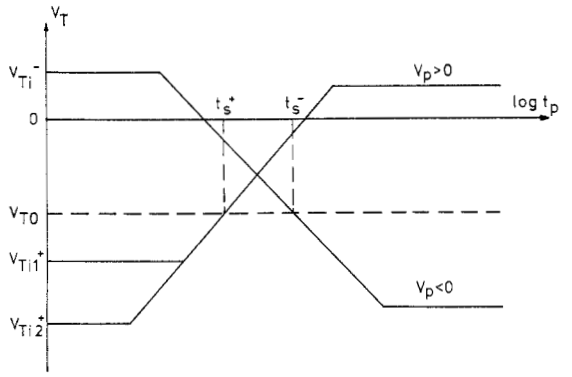

Fig. 1. Idealized switching characteristics for MNOS memory transistor. MNOST threshold voltage V_T versus programming pulse length t_p. For positive programming voltage ($V_p > 0$) the characteristics are shown for two different initial values of V_T, V_{Ti1}^+, and V_{Ti2}^+. V_{TO} is the threshold voltage corresponding to zero stored charge in the nitride. The switching time constants are t_s^+ for the positive value of V_p and t_s^- for the negative value of V_p.

of the gate voltage, variations in the nitride thickness result in variations in the oxide field which, in turn, affects the switching speed. It is, therefore, important to determine the variations in switching speed which can be expected from manufacturing tolerances on the nitride thickness.

II. Switching Theory

The switching characteristics of MNOS memory transistors are normally plotted in a semilogarithmic diagram, showing the MNOS transistor threshold voltage V_T versus the length t_p of the applied gate voltage pulse. Fig. 1 shows an idealized switching diagram.

The switching characteristics may be divided into three regions:

1) An initial region for very small values of t_p, where virtually no shift of V_T is observed.

2) A shift region where V_T shows a logarithmic dependence on t_p.

3) A saturation region for very large values of t_p.

The transition point between region 1) and region 2) is determined by the initial value of V_T prior to the application of the programming voltage V_p. In Fig. 1 the switching characteristic for positive V_p is shown for two different values of the initial threshold voltage.

The transition point between region 2) and region 3) depends on the saturation value of V_T which depends on the applied voltage V_p and the properties of the gate insulators.

The value V_{TO} of the threshold voltage which is shown in Fig. 1 is the value of V_T which corresponds to zero stored charge in the nitride. V_{TO} can be found experimentally either from the switching characteristics [1], [2] or from discharge measurements [3].

The shift region of the switching characteristics is uniquely determined by the slope of the characteristics and the intersection time t_s between $V_T = V_{TO}$ and the switching curve. The intersection time t_s is referred to as the switching time constant [2], and the slope $(dV_T)/(d \log t_p)$ of the switching characteristics is given by [2]

$$\frac{dV_T}{d \log t_p} = 2.3\, \epsilon_{ox}\, \frac{E_A}{C_G} \tag{1}$$

where ϵ_{ox} is the oxide permittivity and C_G is the total gate capacitance per unit area. E_A is called the switching field parameter [2] and is a property of the MNOST which depends on the oxide field and the oxide thickness.

The most important parameter describing the switching speed is the switching time constant t_s, and it is, therefore, of prime interest to examine the influence of nitride thickness variation on t_s.

The switching time constant t_s is a function of the oxide thickness and the oxide field. The oxide field is a function of the applied voltage and the insulator thickness, hence the influence of the nitride thickness on t_s. Based on an assumption of an approximately exponential dependence of the oxide current density $J_{ox}(E_{ox})$ on the oxide field E_{ox} it can be shown [1], [2] that t_s is given by

$$t_s = \frac{C_N \epsilon_{ox}}{C_G J_{ox}(E_{ox\,0})} \cdot \left[\frac{\partial \ln J_{ox}}{\partial |E_{ox}|} \right]^{-1}_{E_{ox\,0}} \tag{2}$$

where C_N is the nitride capacitance per unit area and $E_{ox\,0}$ is the oxide field resulting from the application of the gate voltage V_p when no charge is stored in the nitride. C_N is given by

$$C_N = \frac{\epsilon_N}{W_N}. \tag{3}$$

Similarly, the oxide capacitance per unit area is given by

$$C_{ox} = \frac{\epsilon_{ox}}{W_{ox}}. \tag{4}$$

W_N and W_{ox} are the nitride thickness and the oxide thickness, respectively, and ϵ_N is the nitride permittivity.

In a practical memory circuit, it is more convenient to relate the switching time constant to the programming voltage V_p than to the oxide field E_{ox}. $E_{ox\,0}$ is given by

$$E_{ox\,0} = \frac{1}{W_{ox}}\, V_p\, \frac{C_N}{C_{ox} + C_N}. \tag{5}$$

Now, assume that a fixed value of the programming voltage V_p is specified. The problem then is to determine the insulator thickness variations which can be tolerated in the manufacturing process in order to obtain a guaranteed switching time.

In practical memory circuits, the oxide thickness is 20–25 Å and it can be very well controlled, e.g., using a dry oxidation at a rather low temperature (600°C) for the preparation of the gate oxide. Curves showing the switching time constant versus the oxide thickness are given in [2] and will not be dealt with in the present correspondence.

The nitride thickness in practical memory circuits is normally around 600 Å. This gives a programming voltage V_p about 30 V corresponding to an oxide field about 8 MV/cm. The nitride is deposited by the silane-ammonia reaction. The variations in thickness may well be on the order of 5 percent or more, and this can be shown to have a major influence on the switching time constant.

It has been shown by several investigators [1], [2], [4] that for oxide fields above approximately 6 MV/cm a good agreement between measured switching data and calculated characteristics is obtained if the oxide current is assumed to be a modified Fowler–Nordheim injection from the silicon into the nitride. From the theoretical model of Fowler–Nordheim injection $J_{ox}(E_{ox})$ can be computed [2] and, using (2) and (5), t_s can be calculated as a function of W_N for a fixed value of V_p. Fig. 2 shows the relative variation of t_s versus W_N for a device with typical parameters.

The curves are calculated for a device with an oxide thickness $W_{ox} = 20$ Å. The nitride thickness is normalized versus $W_{N0} = 600$ Å. The programming voltage V_p is assumed to be ±30 V. The switching time constant t_s is normalized versus $t_{s0} = 5.06 \times 10^{-5}$ s for $V_p = +30$ V and $t_{s0} = 2.89 \times 10^{-4}$ s for $V_p = -30$ V. These values of t_{s0} correspond to $W_N = W_{N0} = 600$ Å.

From the figure it is seen that a relative variation of only 5 percent in W_N results in a relative variation of t_s of almost 3 times.

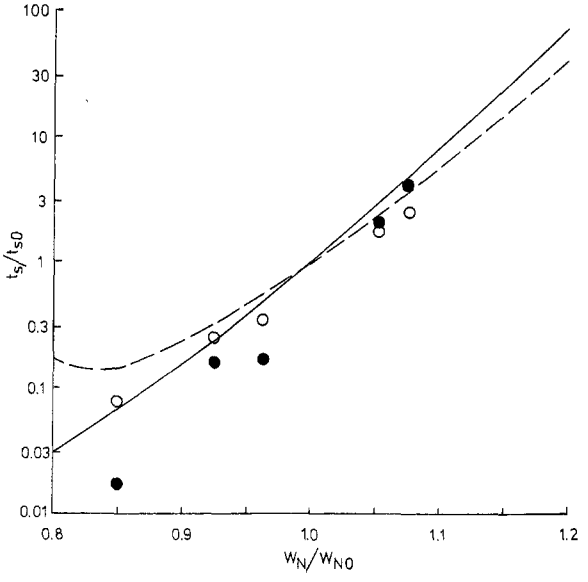

Fig. 2. Normalized switching time constant versus normalized nitride thickness for $V_p = \pm 30$ V, $W_{ox} = 20$ Å. W_N has been normalized versus $W_{N0} = 600$ Å, and t_s has been normalized versus $t_{s0} = 5.06 \times 10^{-5}$s for $V_p = +30$ V, and $t_{s0} = 2.89 \times 10^{-4}$s for $V_p = -30$ V. ——, calculated curve for $V_p = +30$ V; – – –, calculated curve for $V_p = -30$ V; •, measured results for $V_p = +30$ V; ○, measured results for $V_p = -30$ V.

As one would expect, this variation has a major effect on the threshold voltage window obtained using a specified length and amplitude of the programming pulse. From (1) the threshold window may be found

$$\Delta V_T = 2.3 \frac{\epsilon_{ox}}{C_G} \left(E_A^+ + \log \frac{t_p^+}{t_s^+} + E_A^- \log \frac{t_p^-}{t_s^-} \right) \tag{6}$$

where the superscripts + and − are used for positive V_p and negative V_p, respectively. In (6), t_s^+ and t_s^- are strong functions of W_N as shown in Fig. 2. E_A^+ and E_A^- are weak functions of W_N, and so is C_G.

As an example of the impact on ΔV_T of variations in W_N, ΔV_T is calculated versus the relative nitride thickness (W_N normalized versus $W_{N0} = 600$ Å) for $V_p = \pm 30$ V and $t_p = 10 \cdot t_{s0}$. ($t_p^+ = 5.06 \times 10^{-4}$ s, $t_p^- = 2.89 \times 10^{-3}$ s.) The resulting curve is shown in Fig. 3.

III. Experimental Results

Measurements of t_s have been performed on p-channel devices with an oxide thickness $W_{ox} = 20$ Å, measured by ellipsometer, and a nitride thickness W_N around 600 Å. The nitride thickness has, for each device, been determined from a gate capacitance measurement using

$$W_N = \epsilon_N \left(\frac{1}{C_G} - \frac{W_{ox}}{\epsilon_{ox}} \right). \tag{7}$$

A relative dielectric constant of 6.5 for the nitride has been assumed. The switching time constant has been found from switching characteristics which were measured in the following way. Before each measurement V_T was made highly positive (negative) by the application of a gate voltage of +30 V (−30 V) for 5 s. Then a negative (positive) gate voltage pulse with an amplitude of 30 V and length t_p was applied, and the threshold voltage was measured by an automatic equipment applying to the transistor a gate voltage which ensures that the transistor conducts a drain current of 3 μA. The measuring of the threshold voltage takes place within 80 μs after the termination of the gate programming pulse.

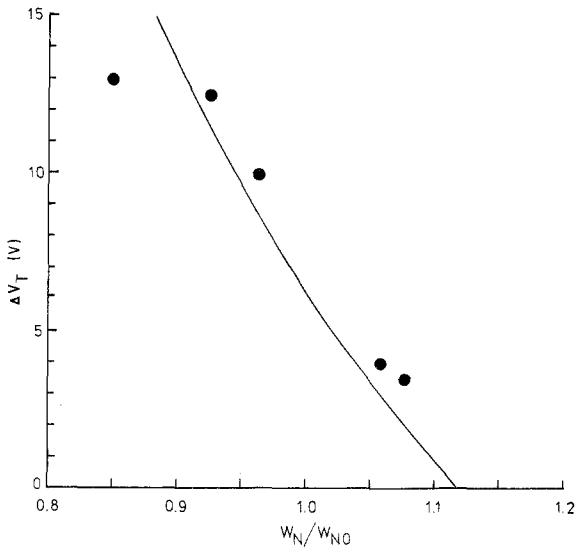

Fig. 3. Calculated and measured threshold voltage window ΔV_T versus normalized nitride thickness W_N/W_{N0}. $W_{ox} = 20$ Å, $W_{N0} = 600$ Å. Programming conditions: $V_p^+ = +30$ V, $t_p^+ = 10 \cdot t_{s0}^+ = 5.06 \times 10^{-4}$ s; $V_p^- = -30$ V, $t_p^- = 10 \cdot t_{s0} = 2.89 \times 10^{-3}$s.

The measured values of t_s have been normalized with the calculated values of t_{s0} for $V_p = \pm 30$ V, $W_N = 600$ Å, and $W_{ox} = 20$ Å. The experimental results are plotted in Fig. 2.

Although the agreement between the calculated curves and the measured results is not perfect, it is evident that the switching time constant shows the expected strong dependence on the nitride thickness. It is also seen that the dependence is stronger for positive V_p than for negative V_p as expected from the theory. The reasons for the nonperfect agreement between the calculated curves and the measured values may simply be inaccuracies in the measured values of t_s and the gate capacitance C_G. Also, in the calculation of the curves, only the modified Fowler-Nordheim injection is taken into account as this is the dominant injection mechanism at the fields involved. Other injection mechanisms may, however, be present [2].

In Fig. 3 we show the measured values of the threshold voltage window obtained with $t_p = 10 \cdot t_{s0}$. The measured results are in reasonable agreement with the calculated curve except for the result obtained for the device with $W_N = 0.85 \cdot W_{N0} = 510$ Å. For this device, a saturation of the threshold voltage takes place with the applied values of V_p and t_p, thus limiting the threshold voltage window.

The measured results clearly demonstrate the necessity of maintaining a close tolerance on the nitride thickness if critical switching requirements are to be fulfilled.

IV. Conclusion

It has been shown both theoretically and experimentally that the switching time constant of MNOS memory transistors is strongly dependent on the nitride thickness for a fixed value of the programming voltage. Curves showing the variation of t_s versus the variation of the nitride thickness are given. It is demonstrated that for a typical MNOS device a variation in the nitride thickness of only 5 percent results in a relative variation of the switching time constant of almost 3 times. Consequently, an extremely good control of the nitride deposition is required in order to obtain reproducible switching characteristics.

References

[1] K. I. Lundström and C. M. Svensson, "Properties of MNOS structures," *IEEE Trans. Electron Devices*, vol. ED-19, pp. 826-836, 1972.

[2] E. Bruun, "Switching characteristics of m.n.o.s. memory transistors," *IEE J. Solid-State and Electron Devices*, vol. 1, pp. 133–138, 1977.
[3] L. Lundkvist, C. Svensson, and B. Hansson, "Discharge of MNOS structures at elevated temperatures," *Solid-State Electron.*, vol. 19, pp. 221–227, 1976.
[4] H. E. Maes and R. J. Van Overstraeten, "Low-field behavior of MNOS devices," *J. Appl. Phys.*, vol. 47, pp. 664–666, 1976.

Influence of Surface States on the Measurement of Field-Effect Mobility

S. T. HSU

Abstract—**The effect of surface states on the measurement of field mobility has been analyzed. The result shows that this effect is important when the rate of change of the surface charge density with respect to the surface potential is larger than the mobile charge density at the conductive channel of the MOSFET.**

The conventional method of obtaining the field-effect mobility of electrons and holes in MOSFET's is by measuring the slope of the transfer current–voltage characteristics of the MOSFET's operated in the linear region [1]. The results that one obtains are found to be dependent upon the fabrication process. This process-dependent field-effect mobility is interpreted as arising from charges in the surface scattering mechanism and the redistribution of impurities in the silicon near the Si–SiO_2 interface [2].

During the course of life testing p-channel MOSFET's we found that the field-effect hole mobility measured by the conventional method is approximately linearly proportional to the absolute value of the threshold voltage, as is shown in Fig. 1. In this figure we plotted the field-effect hole mobility as a function of threshold voltage of six chips in plastic packages. Each chip contains three p-channel and three n-channel MOSFET's and other test devices. The MOSFET's were stressed with negative bias at high temperature (NBT). The typical NBT stress was done with gate voltage between −10 to −20 V, the temperature was between 150 to 200°C for a time interval of up to 1000 h. After NBT stress, the threshold voltage of p-channel MOSFET's increased. The amount of the p-channel threshold voltage increase depends on the device process, package conditions as well as the bias voltage, and time duration of NBT stress. This threshold voltage increase can be reduced by a positive bias temperature (PBT) stress.

The threshold voltage of our n-channel MOSFET's can be changed only very slightly by NBT stress. The threshold voltage increase of p-channel MOSFET's caused by NBT stress is believed due to holes injected into the oxide trap states. One would expect a stronger surface scattering mechanism in p-channel MOSFET's after NBT stress. Hence, the field-effect hole mobility should decrease instead of increase with NBT stress as was observed. We present an analysis in this correspondence showing that the conventional method measures the field-effect mobility of MOSFET's only when the charge density at the fast surface states is independent of the surface potential when the surface is inverted.

The gate voltage of a MOSFET can be written as

$$V_G = \varphi_{MS} + \varphi_S - \frac{1}{C_o}\,(Q_B + Q_{ss} + Q_{st}^{+} - Q_{st}^{-} + Q_c) \qquad (1)$$

Manuscript received February 24, 1978.
The author is with RCA Corporation, RCA Laboratories, Princeton, NJ 08540.

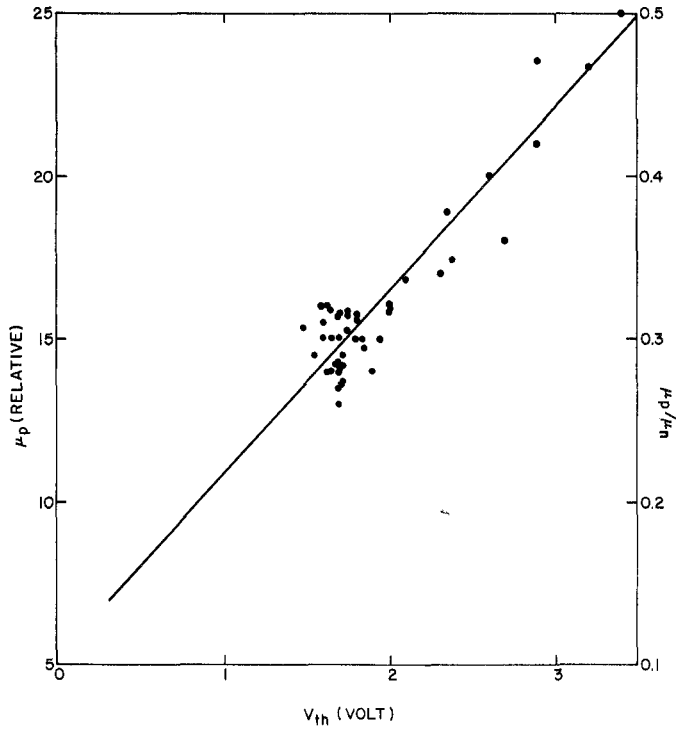

Fig. 1. Relative field-effect hole mobility at small drain current as a function threshold voltage. The threshold voltage is changed by bias-temperature stress.

where

φ_{MS}	work function difference between gate and substrate
φ_S	surface potential
Q_B	charge density per unit area in bulk Si
Q_{ss}	fixed surface state density
Q_{st}^{+}	positive charged fast surface state density
Q_{st}^{-}	negative charged fast surface state density
Q_c	mobile charge density at the conduction channel
C_o	unit area oxide capacitance.

The threshold voltage V_T, therefore, is given by

$$V_T = \varphi_{MS} + 2\varphi_F - \frac{1}{C_o}\,(Q_B + Q_{ss} + Q_{st}^{+} - Q_{st}^{-}). \qquad (2)$$

The mobile charge density at the conduction channel can be written as

$$Q_c = \pm Q_{co} \exp\left(\mp \frac{q\varphi_s}{2kT}\right) \qquad (3)$$

where the upper sign is for p-channel MOSFET's and the lower sign is for n-channel devices. The drain current I_D is given by

$$I_D = \frac{W}{L}\,\mu Q_C V_D = \frac{W}{L}\,\mu C_o\,(V_G - V_T)\,V_D. \qquad (4)$$

The field-effect mobility measured with the conventional method is

$$\mu = \frac{dI_D}{dV_G}\Big/\left(\frac{W}{L}\,C_o V_D\right). \qquad (5)$$

In this expression the threshold voltage V_T is assumed to be constant. If we examine (2) carefully we find that V_T is a function of V_G through the variation of fast surface state, Q_{st},

0018-9383/78/1100-1331$00.75