An Analytical Solution to a Double-Gate MOSFET with Undoped Body

Yuan Taur, *Fellow, IEEE*

***Abstract*—A one-dimensional (1-D) analytical solution is derived for an undoped (or lightly-doped) double-gate MOSFET by incorporating only the mobile charge term in Poisson's equation. The solution gives closed forms of band bending and volume inversion as a function of silicon thickness and gate voltage. A threshold criterion is derived which serves to quantify the gate work function requirements for a double-gate CMOS.**

***Index Terms*—MOSFET's, semiconductor devices, transistors.**

AS CMOS scaling is approaching the limit imposed by gate oxide tunneling, double-gate MOSFET is becoming an intense subject of VLSI research because in theory it can be scaled to the shortest channel length possible for a given gate oxide thickness [1]. Among the advantages advocated for double-gate MOSFET's are: ideal 60 mV/decade subthreshold slope, volume inversion [2], setting of threshold voltage by the gate work function thus avoiding dopants and associated number fluctuation effects, etc. In this paper, a one-dimensional (1-D) Poisson's equation is solved for an undoped (or lightly-doped) double-gate MOSFET to derive analytical expressions for band bending and inversion charge density. A threshold criterion is established for quantifying the gate work function requirements in a double-gate CMOS.

Fig. 1 shows the schematic band diagrams of a symmetric double-gate MOSFET. Same voltage is applied to the two gates having the same work function. At zero gate voltage, the position of the silicon bands is largely determined by the gate work function. This is because as long as the thin silicon is lightly-doped and the depletion charge is negligible, the bands remain essentially flat throughout the thickness of the film. Since there is no contact to the silicon body, the energy levels are referenced to the electron quasi-Fermi level or the conduction band of the n^+ source–drain (not shown), represented by the long dotted line in Fig. 1. As the gate voltage increases toward the threshold voltage in Fig. 1(b), mobile charge or electron density becomes appreciable when the conduction band of the silicon body moves to near the conduction band of the source-drain.

By defining the coordinates and potential as in Fig. 1, one can write Poisson's equation for the silicon region with only the mobile charge (electron) density as

$$\frac{d^2\psi}{dx^2} = \frac{q}{\varepsilon_{\rm si}} n_i e^{q\psi/kT} \tag{1}$$

Manuscript received October 27, 1999. The review of this letter was arranged by Editor E. Sangiorgi.

The author is with IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: taur@watson.ibm.com).

Publisher Item Identifier S 0741-3106(00)03721-6.

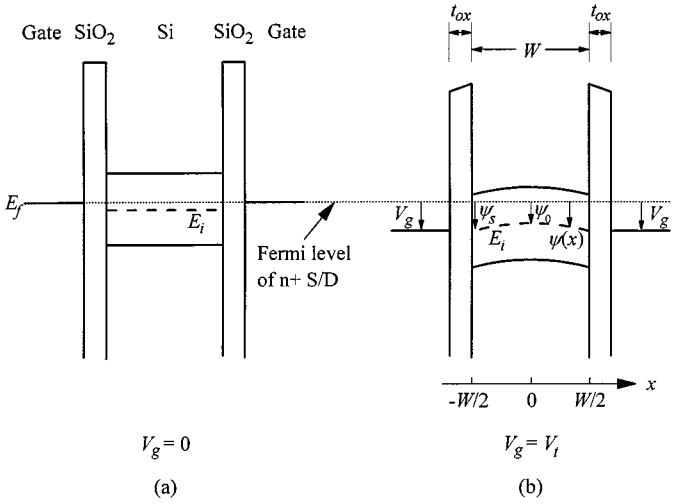

Fig. 1. Schematic band diagrams of a symmetric, undoped double-gate nMOSFET. At zero gate voltage (a), the silicon bands are flat for the gate work function (slightly toward n^+ than the midgap work function) shown in the example. Near the threshold voltage (b), the conduction band of the silicon body at the surface is bent to near the conduction band of the n^+ source-drain (long dotted line).

where q is the electronic charge,$\varepsilon_{\rm si}$ is the permittivity of silicon, and n_i is the intrinsic carrier density. Here we consider an nMOSFET with $q\psi/kT \gg 1$ so that the hole density is negligible.

Integrating (1) once with the symmetry boundary condition $d\psi/dx|_{x=0} = 0$, one obtains

$$\frac{d\psi}{dx} = \sqrt{\frac{2kTn_i}{\varepsilon_{\rm si}}(e^{q\psi/kT} - e^{q\psi_0/kT})} \tag{2}$$

for $0 \le x \le W/2$. Here $\psi_0 \equiv \psi(x=0)$ is the potential at the center of the silicon film, to be solved later as a function of V_g. Integrating again, one obtains the potential as a function of x:

$$\frac{q(\psi-\psi_0)}{2kT} = -\ln\left[\cos\left(\sqrt{\frac{q^2n_i}{2\varepsilon_{\rm si}kT}}e^{q\psi_0/2kT}x\right)\right]. \tag{3}$$

Note that the right-hand side (RHS) of (3) is always positive.

The surface potential $\psi_s \equiv \psi(x=W/2)$ is then given by

$$\frac{q(\psi_s-\psi_0)}{2kT} = -\ln\left[\cos\left(\sqrt{\frac{q^2n_i}{2\varepsilon_{\rm si}kT}}e^{q\psi_0/2kT}(W/2)\right)\right]. \tag{4}$$

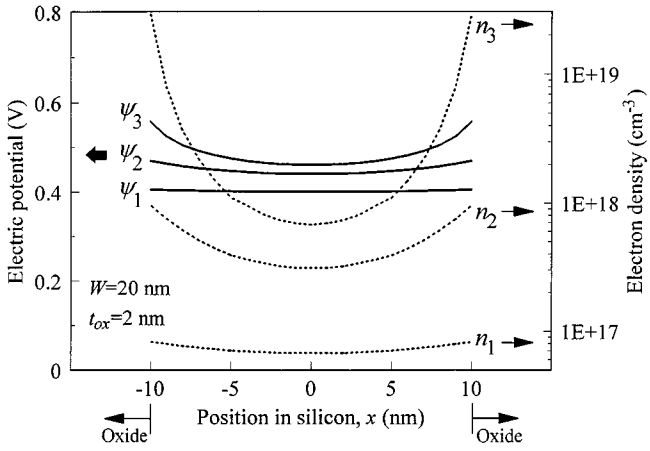

Fig. 2. Potential ψ and electron volume density $n = n_i \exp(q\psi/kT)$ as a function of position in the silicon film for three different values of ψ_0. The corresponding gate voltages are: $V_{g1} - \Delta\phi_i = 0.412$ V for ψ_1 and n_1, $V_{g2} - \Delta\phi_i = 0.573$ V for ψ_2 and n_2, and $V_{g3} - \Delta\phi_i = 0.845$ V for ψ_3 and n_3.

ψ_s is also related to V_g and $t_{\rm ox}$ through the boundary condition at the Si-SiO_2 interface:

$$\varepsilon_{\rm ox}\frac{V_g - \Delta\phi_i - \psi_s}{t_{\rm ox}} = \varepsilon_{\rm si}\left.\frac{d\psi}{dx}\right|_{x=W/2} = \sqrt{2\varepsilon_{\rm si}kTn_i(e^{q\psi_s/kT} - e^{q\psi_0/kT})}. \quad (5)$$

Here $\Delta\phi_i$ is the work function difference between the gate electrode and intrinsic silicon. In other words, $\Delta\phi_i = 0$ for midgap work function gate, $-E_g/2q$ for n^+ poly, and $+E_g/2q$ for p^+ poly, etc. Given V_g, equations (4) and (5) are coupled equations that can be solved for ψ_s and ψ_0.

Numerical examples of the solution to the above equations are shown for $W = 20$ nm (or 5 nm in some cases), $t_{\rm ox} = 2$ nm, and $\Delta\phi_i = 0$ in Figs. 2–4. Fig. 2 plots the potential $\psi(x)$ and the electron density $n = n_i \exp(q\psi/kT)$ for three different values of ψ_0 or V_g. Fig. 3 plots the solutions ψ_s and ψ_0 versus V_g. Fig. 4 shows the sheet density of mobile charge, $Q_i = 2\varepsilon_{\rm si}(d\psi/dx)_{x=W/2}$ (the factor of two arises from the two surfaces), versus V_g. There are two distinct regions of operation in the double-gate MOSFET, just like in a conventional bulk MOSFET.

Below the threshold voltage, the mobile charge density is low and $\psi_s \approx \psi_0 \approx V_g - \Delta\phi_i$. In other words, the bands move as a whole as both ψ_s and ψ_0 closely follow V_g (for $V_g < 0.4$ V in Fig. 3) and volume (weak) inversion takes place. This is illustrated by the case of ψ_1 and n_1 in Fig. 2 where they are essentially flat within the silicon film. The total inversion charge density increases exponentially with V_g at an inverse slope of 60 mV/decade, as shown in Fig. 4.

When the conduction band of the undoped body is bent close to the conduction band of the n^+ source-drain, however, mobile charge terms on the RHS of (4) and (5) are no longer negligible. As the mobile charge near the silicon surfaces screens the gate field from the center of the silicon film, ψ_s and ψ_0 become de-coupled, i.e., there is no further volume inversion. This is illustrated by the case of ψ_3 and n_3 in Fig. 2, and also by the behavior of ψ_s and ψ_0 in Fig. 3 for $V_g > 0.5$ V. In fact, since the angle of the cosine function in (4) cannot exceed $\pi/2$, ψ_0 is pinned to an upper bound of $(kT/q)\ln(2\pi^2\varepsilon_{\rm si}kT/q^2n_iW^2)$, which gives the saturation values of ψ_0 in Fig. 3. ψ_s, on the other hand, continues to increase slowly as governed by (5) with the $e^{q\psi_0/kT}$ term in the square root neglected:

$$V_g - \Delta\phi_i = \psi_s + \frac{Q_i/2}{C_{\rm ox}} \approx \psi_s + \frac{\sqrt{2\varepsilon_{\rm si}kTn_i}}{C_{\rm ox}}e^{q\psi_s/2kT}. \quad (6)$$

It is worth noting here that the above-threshold behavior described by (6) is independent of silicon thickness. This is also reflected in the similarity between the two curves of different W in Fig. 4.

When V_g is substantially above the threshold voltage, the second or the inversion charge term on the RHS of (6) becomes dominant over the first term. The sheet density of mobile charge, Q_i, then increases nearly linearly with V_g (Fig. 4) with a slope $\approx 2C_{\rm ox}$, i.e., $Q_i = 2C_{\rm ox}(V_g - \Delta\phi_i - \psi_s)$. Here ψ_s is a weak function of V_g (Fig. 3) that can be solved iteratively from (6). With a second iteration, the extrapolated threshold voltage, $V_t \equiv \Delta\phi_i + \psi_s$, can be expressed as

$$V_t = \Delta\phi_i + \frac{2kT}{q}\ln\left[\frac{C_{\rm ox}(\Delta V_{\rm gt})}{\sqrt{2\varepsilon_{\rm si}kTn_i}}\right] = \Delta\phi_i + \frac{E_g}{2q} + \frac{2kT}{q}\ln\left[\frac{C_{\rm ox}(\Delta V_{\rm gt})}{\sqrt{2\varepsilon_{\rm si}kT\sqrt{N_cN_v}}}\right]. \quad (7)$$

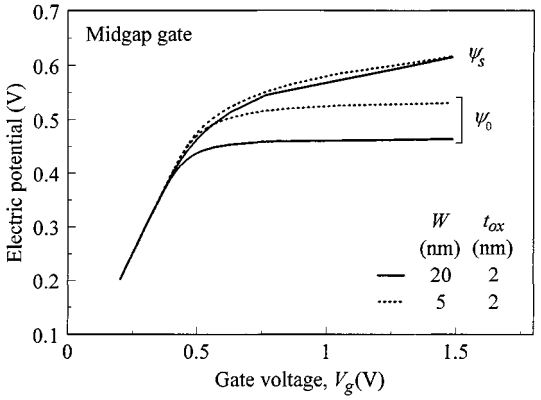

Fig. 3. Solutions ψ_s and ψ_0 from the coupled (4) and (5) versus V_g for two sets of values of W and $t_{\rm ox}$. $\Delta\phi_i = 0$ is assumed.

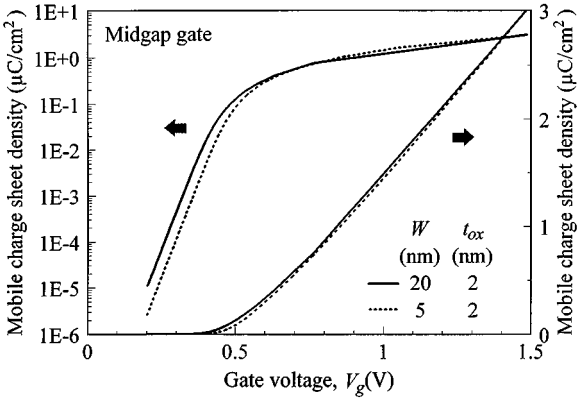

Fig. 4. Sheet density of mobile charge, $Q_i = 2\varepsilon_{\rm si}(d\psi/dx)_{x=W/2}$, in both logarithmic (left) and linear (right) scales versus gate voltage for two sets of values of W and $t_{\rm ox}$. $\Delta\phi_i = 0$ is assumed.

Here $\Delta V_{\rm gt} \equiv V_g - \Delta\phi_i - \psi_s$ is the gate overdrive for the region of interest, and the last step used the expression $n_i = (N_c N_v)^{1/2} \exp(-E_g/2kT)$, where E_g is the silicon bandgap and N_c, N_v are the effective densities of states [3]. The value of V_t is insensitive to the exact choice of $\Delta V_{\rm gt}$; a factor of two difference only results in a threshold change of 0.036 V. If we take $\Delta V_{\rm gt} \approx (10\text{–}20)kT/q$, a practical choice applicable to the cases in Fig. 4, the last term on the RHS of (7) is small (<0.05 V) and the extrapolated threshold voltage for a midgap work function gate ($\Delta\phi_i = 0$) is about 0.55 V as seen in Fig. 4. The same estimate also suggests that V_t is insensitive to temperature and oxide thickness.

For future CMOS logic circuits with a scaled-down power supply voltage [4], the above V_t is too high. Dual n^+ and p^+ polysilicon gates, on the other hand, render the V_t too low: slightly negative for nMOSFET's and slightly positive for pMOSFET's. One can always put the dopants into a double-gate CMOS to obtain the desired threshold voltages [5]. But that, of course, gives up the ideal subthreshold slope and lets the discrete dopant fluctuations back into the picture.

One key physical effect neglected in the above analysis is the inversion layer quantum effect, which tends to shift the peaks of electron concentration away from the oxide interface and toward the center of the silicon film. How far the electron peak is from the surface depends on the magnitude of the surface electric field, i.e., charge density inside the silicon. For silicon surface fields in the range of 5×10^5 V/cm, such as the case of ψ_3 and n_3 in Fig. 2, recent numerical quantum calculations [6] showed that the electron peaks are about 1 nm from the interface (no difference from that in a single-gate bulk device) and remain separated for the case of $W = 5$ nm. This indicates that the bimodal picture of electron distribution beyond threshold as depicted by the classical model is basically valid. Only the detailed peak-to-valley ratio differs between the two models. For silicon film thicknesses much thinner than 5 nm, however, the quantum shift of V_t due to the finite ground-state electron energy, which is inversely proportional to the square of the silicon thickness, becomes significant. Under those circumstances, V_t is higher than that of the classical model and is difficult to control because of its high sensitivity to the silicon thickness [1].

In conclusion, a 1-D Poisson's equation with only the mobile charge term is solved analytically for a symmetric, undoped double-gate MOSFET. It is shown that "volume inversion" only occurs in the subthreshold region, insensitive to the silicon film thickness above 5 nm. Although Fermi–Dirac distribution is not used and quantum effect is neglected, the basic conclusions remain the same. A key technological challenge for double-gate CMOS is to find two gate materials with proper work functions for the desired n and p-threshold voltages, respectively.

REFERENCES

[1] D. Frank, S. Laux, and M. Fischetti, "Monte Carlo simulation of a 30 nm dual-gate MOSFET: How far can silicon go?," *1992 IEDM Tech. Dig.*, p. 553.

[2] F. Balestra and S. Cristoloveanu *et al.*, "Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance," *IEEE Electron Device Lett.*, vol. EDL-8, p. 410, 1987.

[3] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[4] Y. Taur *et al.*, "CMOS scaling into the nanometer regime," *Proc. IEEE*, vol. 85, pp. 486–504, 1997.

[5] H.-S. P. Wong, D. J. Frank, and P. M. Solomon, "Device design considerations for double-gate, ground-plane, and single-gate ultra-thin SOI MOSFET's at the 25 nm gate length generation," *IEDM Tech. Dig.*, 1998.

[6] M. Shoji and S. Horiguchi, "Electronic structures and phononlimited electron mobility of double-gate silicon-on-insulator silicon inversion layers," *J. Appl. Phys.*, vol. 85, pp. 2722–2731, 1999.