Chapter 2: Threshold Voltage Model

2.1 Long-Channel Model With Uniform Doping

Accurate modeling of threshold voltage V_{th} is important for precise description of device electrical characteristics. V_{th} for long and wide MOSFETs with uniform substrate doping is given by

(2.1.1)

$$V_{th} = VFB + \Phi_s + g\sqrt{\Phi_s - V_{bs}} = VTH0 + g\left(\sqrt{\Phi_s - V_{bs}} - \sqrt{\Phi_s}\right)$$

where *VFB* is the flat band voltage, *VTH0* is the threshold voltage of the long channel device at zero substrate bias, and γ is the body bias coefficient given by

(2.1.2)

$$g = \frac{\sqrt{2qe_{si}N_{substrate}}}{C_{oxe}}$$

where $N_{substrate}$ is the uniform substrate doping concentration.

Equation (2.1.1) assumes that the channel doping is constant and the channel length and width are large enough. Modifications have to be made when the substrate doping concentration is not constant and/or when the channel is short, or narrow.

2.2 Non-Uniform Vertical Doping

The substrate doping profile is not uniform in the vertical direction and therefore γ in (2.1.2) is a function of both the depth from the interface and the substrate bias. If $N_{substrate}$ is defined to be the doping concentration (*NDEP*) at X_{dep0} (the depletion edge at V_{bs} = 0), V_{th} for non-uniform vertical doping is

(2.2.1)

$$V_{th} = V_{th,NDEP} + \frac{qD_0}{C_{oxe}} + K1_{NDEP}\left(\sqrt{j_s - V_{bs} - \frac{qD_1}{e_{si}}} - \sqrt{j_s - V_{bs}}\right)$$

where $K1_{NDEP}$ is the body-bias coefficient for $N_{substrate}$ = *NDEP*,

(2.2.2)

$$V_{th,NDEP} = VTH0 + K1_{NDEP}\left(\sqrt{j_s - V_{bs}} - \sqrt{j_s}\right)$$

with a definition of

(2.2.3)

$$j_s = 0.4 + \frac{k_B T}{q}\ln\left(\frac{NDEP}{n_i}\right)$$

where n_i is the intrinsic carrier concentration in the channel region. The zero-th and 1st moments of the vertical doping profile in (2.2.1) are given by (2.2.4) and (2.2.5), respectively, as

(2.2.4)

$$D_0 = D_{00} + D_{01} = \int_0^{X_{dep0}} \left(N(x) - NDEP\right)dx + \int_{X_{dep0}}^{X_{dep}} \left(N(x) - NDEP\right)dx$$

(2.2.5)

$$D_1 = D_{10} + D_{11} = \int_0^{X_{dep0}} (N(x) - NDEP)xdx + \int_{X_{dep0}}^{X_{dep}} (N(x) - NDEP)xdx$$

By assuming the doping profile is a steep retrograde, it can be shown that D_{01} is approximately equal to $-C_{01}V_{bs}$ and that D_{10} dominates D_{11}; C_{01} represents the profile of the retrograde. Combining (2.2.1) through (2.2.5), we obtain

(2.2.6)

$$V_{th} = VTH0 + K1\left(\sqrt{\Phi_s - V_{bs}} - \sqrt{\Phi_s}\right) - K2 \cdot V_{bs}$$

where $K2 = qC_{01} / C_{oxe}$, and the surface potential is defined as

(2.2.7)

$$\Phi_s = 0.4 + \frac{k_B T}{q} \ln\left(\frac{NDEP}{n_i}\right) + PHIN$$

where

$$PHIN = -qD_{10}/e_{si}$$

$VTH0$, $K1$, $K2$, and $PHIN$ are implemented as model parameters for model flexibility. Appendix A lists the model selectors and parameters.

Detail information on the doping profile is often available for predictive modeling. Like BSIM3v3, BSIM4 allows $K1$ and $K2$ to be calculated based on such details as *NSUB*, *XT*, *VBX*, *VBM*, etc. (with the same meanings as in BSIM3v3):

$$K1 = \gamma_2 - 2K2\sqrt{\Phi_s - VBM} \tag{2.2.8}$$

$$K2 = \frac{(\gamma_1 - \gamma_2)\left(\sqrt{\Phi_s - VBX} - \sqrt{\Phi_s}\right)}{2\sqrt{\Phi_s}\left(\sqrt{\Phi_s - VBM} - \sqrt{\Phi_s}\right) + VBM} \tag{2.2.9}$$

where γ_1 and γ_2 are the body bias coefficients when the substrate doping concentration are equal to *NDEP* and *NSUB*, respectively:

$$\gamma_1 = \frac{\sqrt{2q\epsilon_{si} NDEP}}{C_{oxe}} \tag{2.2.10}$$

$$\gamma_2 = \frac{\sqrt{2q\epsilon_{si} NSUB}}{C_{oxe}} \tag{2.2.11}$$

VBX is the body bias when the depletion width is equal to *XT*, and is determined by

$$\frac{qNDEP \cdot XT^2}{2\epsilon_{si}} = \Phi_s - VBX \tag{2.2.12}$$

2.3 Non-Uniform Lateral Doping: Pocket (Halo) Implant

In this case, the doping concentration near the source/drain junctions is higher than that in the middle of the channel. Therefore, as channel length becomes shorter, a V_{th} roll-up will usually result since the effective channel doping concentration gets higher, which changes the body bias effect as well. To consider these effects, V_{th} is written as

$$V_{th} = VTH0 + K1\left(\sqrt{\Phi_s - V_{bs}} - \sqrt{\Phi_s}\right)\cdot\sqrt{1+\frac{LPEB}{L_{eff}}} - K2\cdot V_{bs} + K1\left(\sqrt{1+\frac{LPE0}{L_{eff}}}-1\right)\sqrt{\Phi_s} \tag{2.3.1}$$

In addition, pocket implant can cause significant drain-induced threshold shift (DITS) in long-channel devices [3]:

$$\Delta V_{th}(DITS) = -nv_t\cdot\ln\left(\frac{\left(1-e^{-V_{ds}/v_t}\right)\cdot L_{eff}}{L_{eff}+DVTP0\cdot\left(1+e^{-DVTP1\cdot V_{ds}}\right)}\right) \tag{2.3.2}$$

For V_{ds} of interest, the above equation is simplified and implemented as

$$\Delta V_{th}(DITS) = -nv_t\cdot\ln\left(\frac{L_{eff}}{L_{eff}+DVTP0\cdot\left(1+e^{-DVTP1\cdot V_{ds}}\right)}\right) \tag{2.3.3}$$

2.4 Short-Channel and DIBL Effects

As channel length becomes shorter, V_{th} shows a greater dependence on channel length (SCE: short-channel effect) and drain bias (DIBL: drain-induced barrier lowering). V_{th} dependence on the body bias becomes weaker as channel length becomes shorter, because the body bias has weaker control of the depletion region. Based on the quasi 2D solution of the Poisson equation, V_{th} change due to SCE and DIBL is modeled [4]

$$\Delta V_{th}(SCE, DIBL) = -q_{th}(L_{eff}) \cdot \left[2(V_{bi} - \Phi_s) + V_{ds}\right] \tag{2.4.1}$$

where V_{bi}, known as the built-in voltage of the source/drain junctions, is given by

$$V_{bi} = \frac{k_B T}{q} \ln\left(\frac{NDEP \cdot NSD}{n_i^2}\right) \tag{2.4.2}$$

where *NSD* is the doping concentration of source/drain diffusions. The short-channel effect coefficient $\theta_{th}(L_{eff})$ in (2.4.1) has a strong dependence on the channel length given by

$$q_{th}(L_{eff}) = \frac{0.5}{\cosh\left(\frac{L_{eff}}{l_t}\right) - 1} \tag{2.4.3}$$

l_t is referred to as the characteristic length and is given by

(2.4.4)

$$l_t = \sqrt{\frac{e_{si} \cdot TOXE \cdot X_{dep}}{EPSROX \cdot h}}$$

with the depletion width X_{dep} equal to

(2.4.5)

$$X_{dep} = \sqrt{\frac{2e_{si}\left(\Phi_s - V_{bs}\right)}{qNDEP}}$$

X_{dep} is larger near the drain due to the drain voltage. X_{dep} / η represents the average depletion width along the channel.

Note that in BSIM3v3 and [4], $\theta_{th}(L_{eff})$ is approximated with the form of

(2.4.6)

$$q_{th}\left(L_{eff}\right) = \exp\left(-\frac{L_{eff}}{2l_t}\right) + 2\exp\left(-\frac{L_{eff}}{l_t}\right)$$

which results in a phantom second V_{th} roll-up when L_{eff} becomes very small (e.g. L_{eff} < *LMIN*). In BSIM4, the function form of (2.4.3) is implemented with no approximation.

To increase the model flexibility for different technologies, several parameters such as *DVT*0, *DVT*1, *DVT*2, *DSUB*, *ETA*0, and *ETAB* are introduced, and SCE and DIBL are modeled separately.

To model SCE, we use

(2.4.7)

$$\theta_{th}(\mathrm{SCE}) = \frac{0.5 \cdot DVT0}{\cosh\left(DVT1 \cdot \frac{L_{eff}}{l_t}\right) - 1}$$

(2.4.8)

$$\Delta V_{th}(\mathrm{SCE}) = -\theta_{th}(\mathrm{SCE}) \cdot (V_{bi} - \Phi_s)$$

with l_t changed to

(2.4.9)

$$l_t = \sqrt{\frac{\varepsilon_{si} \cdot TOXE \cdot X_{dep}}{EPSROX}} \cdot (1 + DVT2 \cdot V_{bs})$$

To model DIBL, we use

(2.4.10)

$$\theta_{th}(\mathrm{DIBL}) = \frac{0.5}{\cosh\left(DSUB \cdot \frac{L_{eff}}{l_{t0}}\right) - 1}$$

(2.4.11)

$$\Delta V_{th}(\mathrm{DIBL}) = -\theta_{th}(\mathrm{DIBL}) \cdot (ETA0 + ETAB \cdot V_{bs}) \cdot V_{ds}$$

and l_{t0} is calculated by

(2.4.12)

$$l_{t0} = \sqrt{\frac{\varepsilon_{si} \cdot TOXE \cdot X_{dep0}}{EPSROX}}$$

with

(2.4.13)

$$X_{dep0} = \sqrt{\frac{2\varepsilon_{si}\Phi_s}{qNDEP}}$$

*DVT*1 is basically equal to $1/(\eta)^{1/2}$. *DVT*2 and *ETAB* account for substrate bias effects on SCE and DIBL, respectively.

2.5 Narrow-Width Effect

The actual depletion region in the channel is always larger than what is usually assumed under the one-dimensional analysis due to the existence of fringing fields. This effect becomes very substantial as the channel width decreases and the depletion region underneath the fringing field becomes comparable to the "classical" depletion layer formed from the vertical field. The net result is an increase in V_{th}. This increase can be modeled as

(2.5.1)

$$\frac{\pi qNDEP \cdot X_{dep,\max}{}^2}{2C_{oxe}W_{eff}} = 3\pi \frac{TOXE}{W_{eff}}\Phi_s$$

This formulation includes but is not limited to the inverse of channel width due to the fact that the overall narrow width effect is dependent on process (i.e. isolation technology). V_{th} change is given by

(2.5.2)

$$\Delta V_{th}(Narrow\_width1) = (K3 + K3B \cdot V_{bs})\frac{TOXE}{W_{eff}' + W0}\Phi_s$$

In addition, we must consider the narrow width effect for small channel lengths. To do this we introduce the following

(2.5.3)

$$\Delta V_{th}\left(Narrow\_width2\right) = -\frac{0.5\cdot DVT0W}{\cosh\left(DVT1W\cdot\frac{L_{eff}W_{eff}'}{l_{tw}}\right)-1}\cdot\left(V_{bi}-\Phi_s\right)$$

with l_{tw} given by

(2.5.4)

$$l_{tw} = \sqrt{\frac{e_{si}\cdot TOXE\cdot X_{dep}}{EPSROX}}\cdot\left(1+DVT2W\cdot V_{bs}\right)$$

The complete V_{th} model implemented in SPICE is

(2.5.5)

$$\begin{aligned}V_{th} = {} & VTH0+\left(K_{1ox}\cdot\sqrt{\Phi_s-V_{bseff}}-K1\cdot\sqrt{\Phi_s}\right)\sqrt{1+\frac{LPEB}{L_{eff}}}-K_{2ox}V_{bseff}\\ & +K_{1ox}\left(\sqrt{1+\frac{LPE0}{L_{eff}}}-1\right)\sqrt{\Phi_s}+\left(K3+K3B\cdot V_{bseff}\right)\frac{TOXE}{W_{eff}'+W0}\Phi_s\\ & -0.5\cdot\left[\frac{DVT0W}{\cosh\left(DVT1W\frac{L_{eff}W_{eff}'}{l_{tw}}\right)-1}+\frac{DVT0}{\cosh\left(DVT1\frac{L_{eff}}{l_t}\right)-1}\right]\left(V_{bi}-\Phi_s\right)\\ & -\frac{0.5}{\cosh\left(DSUB\frac{L_{eff}}{l_{t0}}\right)-1}\left(ETA0+ETAB\cdot V_{bseff}\right)\cdot V_{ds}\end{aligned}$$

where $TOXE$ dependence is introduced in model parameters $K1$ and $K2$ to improve the scalibility of V_{th} model over $TOXE$ as

$$K_{1ox} = K1 \cdot \frac{TOXE}{TOXM} \tag{2.5.6}$$

and

$$K_{2ox} = K2 \cdot \frac{TOXE}{TOXM} \tag{2.5.7}$$

Note that all V_{bs} terms are substituted with a V_{bseff} expression as shown in (2.5.8). This is needed in order to set an upper bound for the body bias during simulations since unreasonable values can occur during SPICE iterations if this expression is not introduced.

$$V_{bseff} = V_{bc} + 0.5 \cdot \left[\left(V_{bs} - V_{bc} - d_1 \right) + \sqrt{\left(V_{bs} - V_{bc} - d_1 \right)^2 - 4 d_1 \cdot V_{bc}} \right] \tag{2.5.8}$$

where $\delta_1 = 0.001\text{V}$, and V_{bc} is the maximum allowable V_{bs} and found from $dV_{th}/dV_{bs} = 0$ to be

$$V_{bc} = 0.9 \left(\Phi_s - \frac{K1^2}{4K2^2} \right) \tag{2.5.9}$$