**L12 ECE 4243/6243 November 15, 2016 F. Jain**

 **Short Channel Effect and FET Scaling Laws**

 **Background (Sections 9.1-9.6)**

9.1 Short-Channel Effects 587

9.2 Mobility Degradation 590

9.3 Subthreshold Conduction 595

9.4 Subthreshold Figure of Merit α 597

9.5 High-field Effects 599

9.6 Punch-through 601

 **Scaling Laws and NanoFET (Section 9.7 for HW 11**

9.7 Scaling Laws and Design Steps for a 0.25 μm MOSFET 606

 9.7.1 Constant Electric Field (CE) Scaling: 606

 9.7.2.Constant Voltage (CV) and Quasi Constant Voltage (QCV) Scaling: 609

 9.7.3.Generalized Scaling Laws [1984] 610

 9.7.4 Design Steps using Generalized Scaling Laws 613

 9.7.5 Nano-FET Design Example 614

 Solved NanoFET Design Set-II 619

 **FIN-FETs and NanoFET degradation Issues (Section 9.8-9.9)**

9.8 Topics on Nanodevices 624

 9.8.1 FIN-FET 625

9.9 Nanotransistors Issues 626

 9.9.1 Degradation mechanisms 627

 9.9.1.1.Poly-Si gate depletion and increased gate capacitance 627

 9.9.1.2.Thin gate insulator 628

 9.9.1.3 Injection of hot carriers 629

 9.9.1.4. Junction leakage current 630

 9.9.1.5.Oxide (dielectric) breakdown 631

 9.9.1.6.Trapped electrons 632

 9.9.1.7. Hole generation during electron tunneling 632

 9.9.2 The depletion region and carrier velocity in a MOSFET 633

 **CNT-FETs**

9.10 Carbon nanotube transistors 634

# 9.1 Short-Channel Effects

Channel Length Modulation is modifies the current out put as W/L ratio increases as the length of the channel L is reduced above VDS(sat).



At pinch-off



Beyond pinch-off

Fig .1(a) Channel length modulation as a function of drain voltage exceeding saturation or pinch off.



Λ = ?

**Derivation of Λ:**

Using Poisson Equation in the drain end of the MOSFET.



Boundary conditions

 φ = φsat at x = L - Λ

 and φ = φBI + VDS at x = L

Solution to the Poissons equation gives

φ = φsat + (φBI + VDS - φsat)(x – L + Λ)/Λ + (x – L + Λ)/2

Using , we get

 (Form I)

The drain current in situation illustrated in Fig. (b)

 

Under saturation 

 or 

For VDS > VDS (sat)



or 

 ID′ = ID(sat) =  (1)

The drain current for situation shown in Fig. (a)

  (2)

fΛ

Eqs. (1) + (2) yield

  (3)

Λ is expressed in various ways by different authors.

 FORM II: 

 FORM III: 



the Built–in voltage at drain

 FORM IV: 

Here, a = , εMax is the field at which velocity saturates to γMax.

Substituting Form II of Λ in Eq. (3).



 

This equation simplifies to

 (4)

Here,

 = Constant.

B is a proportionality constant.

Eq. 4

VDS(sat)

-VA + VDS(sat)

ID

VDS

ID′

An alternative empirical relation for ID′ is

  (5)



# 9.2 Mobility Degradation

The carrier mobility in the inversion layer of a MOSFET depends on the magnitude of electric field (both parallel and perpendicular components) in the channel. The perpendicular filed is determined by gate voltage and lateral field by the drain voltage. In reality it the relative voltage along vertical and lateral axes that determines it (Fig. 1b). It is governed by Poisson’s equation.



x

y

z

Fig. 1(b). Electrical field components in a MOSFET.

The parallel component (εz) is also referred as ‘longitudinal’, ‘lateral’, or ‘horizontal,’ and the perpendicular (εx) is identified as ‘normal’ or ‘vertical.’ The low-field carrier mobility μ0 is generally treated as constant for devices operating under linear or nonsaturation regime. μ0 depends on temperature, surface orientation [e.g. (111) or (100)], quality of Si-SiO2 interface, substrate doping, and interface charge density.

For a given device and operating temperature, the mobility degrades due to velocity saturation effects. The effective mobility is expressed as

  (6)



*fh =*Mobility degradation due to εz or horizontal component.

 The carrier velocity (electron, hole) is a function of electric field. This is true in thin channels as well as in bulk conduction. Fig. 2 shows the drift velocity as a function of εz in inversion layers.



Fig. 2 Field dependence of Carrier velocity in an inversion layer

The region between the pinch-off and drain hosts large electric field εz. This is due to the fact that voltage drop [VDS - VDS(sat)], if of sufficient magnitude, can result in εz > εc.

The mobility degradation factor fh is empirically expressed:

  (7)

  (8)

*fy=*Mobility degradation due to ⊥ or vertical field εx.

 The vertical field tends to accelerate the electrons towards the SiO2-Si interface, causing additional scattering due to the ‘roughness of interface.’ This surface scattering results in the mobility reduction. The surface mobility is a function of an average electric field .

  (9)

 εxs = surface value

 εxb = bulk field just below the inversion layer (i.e. below about 50Å)

Gauss’s law gives

  (10)

 and  (11)

Hence,  (12)

Note that QN, QB depend on VGS as well as on VDS.

In strong inversion

  (13)

δ term is usually dropped.

  (14)

ALSO

2.139 (Text)

The substrate bias is accounted by introducing a θBVSB term

  (15)

Accounting for the mobility degradation term, we get

 ID = ID] \* fν \* fh (16)

without

mobility

degradation

VTE variation

channel length modulation

Mobility degradation

Incorporating:

We get in non saturation

 the following Equation, in place of Eq. (2.108) for ID.

  (17)

In this equation, the influence of Λ (channel length modulation) is not there (as there is no modulation in the nonsaturation region).

However for VDS ≥ VDS(sat)

  (18)

  (Eq. 3)

V′TE was defined in supplementary notes (dated 2/22/88).

  (19)



shows the dependence

ΔVTE also depends on the channel width

 Unlike the channel length L, a decrease in channel width w results in the increase of the threshold voltage. The increase ΔVTE is

  (20)

[Compare Eq. (19) with Eq. (20)]

 Therefore,

  (21)

Long

channel

Eq. (19)

Eq. (20)

Note that the value of V′TE, which one should use in determining ID [from Eqs (17) + (18)], is given by Eq. (21) for narrow channel devices.

One to remember:

Decrease in L decreases VTE

Decrease in w increases VTE

# 9.3 Subthreshold Conduction

We have defined the threshold conduction as when the channel potential was reduced by



The device is in weak inversion if the channel potential is between  and . Under the condition of weak inversion the electron density is

  (22)

**Optional:**

  (23)

(Eq. 4.39)
Brews

Eq. 23 is from Brews, who assumes charge sheet approximation.

Here,





ϕfn = quasi Fermi level for electrons

  (24)

(4.18)
Brews



 In the case of weak inversion ϕs is pretty much uniform along the channel. Its value is close to ϕs(sat). ϕs(sat) is the surface potential, when VDS > VDS(sat), pinch-off and drain region.

At the source end



and at the drain end (25)



Using Eqs (23) and (25)



or  (26)

The channel current under weak inversion is

  (27)

  (28)

From Eqs (26), (27), (28)

  (29)

The saturation potential can be shown to

  (30)

End of optional

Equation 29 illustrates the following:

1. The drain current varies exponentially with ϕs(sat) or VGS due to the term  (since ϕs(sat) has a linear dependence on VGS).
2.  term drops out if ; i.e. VDS has no control over ID in the sub threshold region.
3. ID depends or reduces significantly when VBS is applied. That is, VGS and VBS reinforce each other in reducing ID.

Whereas Eq. 29 is a complicated equation (given the fact that ϕs(sat) is expressed by Eq. 30 – equally complex!), ID is many times simply expressed as



 (31)

In Eq. 31 ϕs(sat) ≈ VGS - V′TE. For V′TE, see the following discussion.

Our textbook uses

  (32)

Here, n is a parameter, which is determined from experiment.

Under sub threshold approximation

  (33)

This threshold yields [using Eqs (32) and (33)]

  (34)

(Eq. 2.159) Text

From Eq. (34) we see that ID has a finite value when VGS < VTE.

# 9.4 Sub threshold Figure Of Merit α

 Sub threshold figure of merit expresses the required reduction in gate voltage VGS to decrease the leakage drain current ID by an order of magnitude.

  (35)

Taking log10 of Eq. (34) and differentiating it



  (36)

Eq. (35) and (36) give

  (37)

 

 n ranges between 1 and 2.5.

The subthreshold figure of merit α is between 60 and 150 mV/decade at 300°K.

Subthreshold currents pose severe problems in short-channel devices. This is illustrated in Fig. 3.

****

# Figure 3 Subthreshold leakage characteristics of transistors with a variety of channel lengths 9.5 High-field Effects

 Small channel MOSFETS manifest significantly high electric field in the vicinity of the drain region, particularly when the device is operating in the saturation mode. This leads to undesirable effects including velocity saturation, hot-carrier effects, dielectric breakdown, drain sustaining breakdown, and wear-out of thin insulators.

 Hot carrier effects give rise to various current or carrier injection mechanisms described in Fig. 4. Takeda et al [1985] have classified these as follows:

1. Channel hot electron injection (CHE) – Channel electrons (near the drain end) obtain enough energy to escape to the gate. These are called ‘Lucky electrons’ and they give rise to the gate current. The process is significant at 77°K and lower temperatures.
2. Drain-Avalanche Hot Carrier (DAHC) injection

leads to time-dependent dielectric breakdown

Fowler-Norheim

1. Tunneling injection

Direct

1. Substrate hot-electron (SHE) injection
2. Secondarily Generated Hot-Electron (SGHE)
3. Oxide clearing (Trap, defect generation in gate oxide)

Device degradation:

1. Transconductance gm degradation
2. VTH variation

gm variation (more severe in PMOS):

1. Interface state generation due to hole-hole hitting
2. Interface state generation due to hole-electron hitting
3. Trapped electrons



Fig.4 Hot-Electron induced device degradation in a n-channel MOSFETs. (Illustration of various current components)

1. Isub:

Holes generated by hot electrons via the impact ionization in the vicinity of drain are collected by the substrate, thus forming thin current components. Isub causes; (a) VTH variations as the potential variations are caused, (b) latchup in CMOS, (c) overloading of chip substrate bias generators.

1. IG:

When hot electrons gain enough energy to overcome the potential barrier (qϕb) at the Si-SiO2 interface [qϕb = 3.1 eV at EOX = 0, or 2.4 eV at EOX = 106 ν/cm] to reach the gate.

  (38)

  (39)

IG → causes

VTH shift

μ degradation

generate oxide traps

Impact ionization cutoff = 

EOX = oxide field near the drain

 =

 COX = 10-3 at EOX = 0

 COX = 4×10-3 at EOX = 106 ν/cm

 λ = hot electrons mean free path (≈78Å)

C2 = Probability for an energetic electron that can be injected in the gate oxide.

Em = field at the drain end of the channel.

 Some of these electrons may be trapped in the oxide causing oxide charging.

1. Icoll:

The minority-carrier current due photogeneration mechanism. It has an additional component available in p-n junctions. (IN)

Icoll = Iphoto + IN

Photo generation: photons are generated by the process of ‘Bremsstrahlung’ when electron-hole recombine at the drain end. Photons, in turn travel quite far, and generate electron-hole pairs.

1. IN:

The source (n+) – substrate junction gets slightly forward biased as Isub becomes significant. This results in injection of minority electrons from n+ source. This occurs when:

Isub \* Rsub ≥ Vsub + 0.6

Once the electrons are injected into the substrate, they are collected by other electrodes (as leakage current). This current may be collected by the drain. This in turn may cause additional electron-hole pairs. Thus, a regenerative process gets underway.

 Icoll can discharge: storage capacitors (e.g. DRAM)

 Low-current carrying nodes

 The source to substrate injection can initiate

 This regenerative process causes snap-back breakdown. It occurs when

Multiplication factor \* αnpn ≥ 1

(parasitic)

#

# 9.6 Punch-through

 As the drain voltage in a MOSFET is increased beyond saturation VD(sat), two phenomena generally take place in long channel devices. These are:

1. oxide breakdown
2. drain junction breakdown

 In long channel MOSFETs, these breakdowns are independent of the channel length. However, as the channel length is reduced, the maximum supportable VDS drops abruptly. This behavior signifies the onset of punchthrough.

VDS(MAX)

volts

Effective Channel length

μm

0

2

4

6

10

15

5

Avalanche breakdown limited (at the drain)

Punch-

through

limited

Figure 5: Max volts vs. Channel length of MOSFET breakdown

VDS(max) → VDS drawing /μA ID.

Device parameters: tOX = 0.4 μm Channel implant dosage = 5×1011 cm2

 NA = 1015 cm-3 keV = 25 keV

 rj = 0.28 μm

 Punchthrough is associated with a flow of subsurface current which is not controlled by the gate. A two-dimensional potential study (Kolani + Kawazu) reveals that a saddle point (pot. minimum) forms below the surface under the gate.



Figure 6: Equipotential and contours of constant minority carrier density in punchthrough



Figure 7: Long channel behavior (upper) vs. short channel behavior (lower)

 Gate control does not exist in the saddle point region. In this region, field lines originating from the drain terminate on the source. Field lines originating at the gate do not pass through the saddle point region to terminate on the substrate in the punchthrough mode. This can also be viewed as drain-induced barrier lowering below the surface.

 Punchthrough voltage is: (1) Proportional to NA

 (2) Proportional to L2

 The substrate bias VSUB (VSB) increases the punchthrough voltage.

 Note that this definition of punchthrough is different from the conventional long-channel characterization of punchthrough (which defines it to be occurring when the source and drain depletion regions merge).



Figure 8: Electron velocity as a function of applied electric field

# 9.7 Scaling Laws and Design Steps for a 0.25 μm MOSFET

Various design criteria have been proposed in the literature to realize scaled-down MOSFET. These are divided into three categories.

1. Constant electric field (CE) scaling [4]\*; 1974
2. Constant voltage (CV) and Quasi constant voltage scaling (QCV) [2]; 1980
3. Generalized scaling [1]; 1984

(\* For references see the original notes dated 1/25/88)

 The primary objective of scaling techniques is to design MOSFETs, which exhibit “long-channel” behavior and manifest no (or fewer) adverse effects as the channel length and width are reduced. That is, it is a methodology to minimize the effect of mobility degradation, threshold variation, punchthrough and other high field phenomena.

###

## 9.7.1 Constant Electric Field (CE) Scaling:

Dennard and coworkers (1974) presented the first systematic approach to scaling. The central concept in this approach is to maintain the electric field constant as the device dimensions are reduced. This is achieved by transforming various parameters according to the following laws:

* 1. Linear dimensions are reduced by a factor k – L, w, TOX are divided by k (k>1) to obtain scaled-down values.

e.g. L′ = L/k, w′ = w/k, T′OX = TOX/k (40)

* 1. Voltages are reduced by k

e.g. V′DS = VDS/k, V′GS = VGS/k (41)

* 1. Substrate doping is increased by k

e.g. N′A = NA\*k (42)

 The applied voltages are reduced to maintain electric field strengths as in the case of long-channel or unscaled devices. In case of substrate doping, the rationale is to keep the depletion width x′p = xp/k.

 In an unscaled device

  (43)



Figure 9: Long channel unscaled device with body revers bias.

W’=W/K

εsi=εrε0

For the scaled-down device

  (44)

  (45)

If N′A = kNA

  (46)

Note: ϕ′s should be ϕs/k for Eq. 44 to reduce to Eq. 45.

 Example: At threshold  (47)

 =/K

  (48)

In light of Eq. 46  (49)

Therefore, Eq. 45 works if VSB>>ϕs, otherwise ϕ′s ≠ ϕs/k. In fact, ϕ′s ≈ ϕs as k ≈ 5-10.

 In the CE scheme of scaling the drain current I′D = ID/k, and threshold voltage V′TH or V′TE ≈ VTE/k (here V′TE stands for scaled-down value). The scaling of other device and circuit parameters is shown below.

TABLE I

|  |  |  |
| --- | --- | --- |
| Parameter | Unscaled | Scaled Value |
| Capacitance (total; F) | cOXwL | w′L′c′OX = cOXwL/k |
| Charge (Total, coulomb) | QwL | Q′w′L′ = QwL/k2 |
| Capacitance per unit area | cOX | c′OX = cOX\*k |
| Charge per unit area | Q | Q′ = Q |
| Power Dissipation (VI) | P | p′ = p/k2 |
| Power Density (VI/wL) | (p/wL) | (p′/w′L′) = (p/wL) |
| Delay Time (wLQ/I) | τ | τ′ = τ/k |
| Resistance | R | R′ = Rk |

Note that k>1.

In this approach the effect of velocity saturation, and drain and source resistances is avoided by maintaining the drain field & power density at the same level as the unscaled situation. On the other hand, power-delay product [p′τ′] of scaled down devices is improved by k3 as

P

τ

Figure 10: Power density vs. drain field

p′τ′ = pτ/k

 However, the power density remains unchanged.

**Limitations:**

 Chatterjee et al [1980] pointed out that the parameter tolerances do not scale linearly as is the case for voltages and dimensions in constant field scaling. In particular, the other limitations include noise margin considerations and compatibility with existing logic families (e.g. TTL). In addition, the interconnect capacitance varies logarithmically rather than linearly with scaling. Therefore, a higher capacitance (of interconnects) requires an increase in the current driving capability of a driver MOSFET to maintain acceptable level of delays. Voltage drop across interconnect resistance does not scale.

### 9.7.2 Constant Voltage (CV) and Quasi Constant Voltage (QCV) Scaling:

1. Constant Voltage Scaling (Cv)

 The scaling laws in which device dimensions (e.g. L, w, TOX) are reduced while maintaining the voltages unchanged constitute the constant voltage scaling.

 Since the voltages are not scaled, a reduction in TOX by the same factor as (w or L) will yield significantly high oxide electric fields. This may result in mobility degradation (higher fν, fh values) and oxide charging effects. Therefore, in this scheme L, w, NA are scaled the same way as in constant-field scaling. TOX is scaled by a factor less than that used for L and w. The voltages are not scaled. The primary advantage is its compatibility with existing logic families.

1. Quasi Constant Voltage (QCV) Scaling

 Unlike constant voltage scaling, in QCV the voltages are scaled down by a factor smaller than used for L and w. The oxide thickness is scaled down by the same factor as used for L and w. In this scheme, the oxide fields are pretty much of similar magnitude as in the case of CV scaling.

CV + QCV are summarized below in Table II

TABLE II

|  |  |  |
| --- | --- | --- |
| Parameter (unscaled) | Constant Voltage | Quasi Constant Voltage |
| Channel length (L) | L′ = L/k | L′ = L/k |
| Channel width (w) | w′ = w/k | w′ = w/k |
| Gate oxide thickness (TOX) | TOX′ = TOX/β | TOX′ = TOX/k |
| Substrate Doping (NA) | N′A = NAk | N′A = NAk |
| Voltages (V) | V′ = V | V′ = V/β |

Note: k>1

k > β > 1

 β>1

Chatterjee et al use λ in their paper.



 and

### 9.7.3 Generalized Scaling Laws [1984]

Baccaarani and coworkers (including Dennard) have proposed a generalized approach to scaling including the points mentioned by Chatterjee et al. Their approach is based on maintaining the shape of the electric field and potential distributions in the scaled-down devices. They do allow, similar to Chatterjee et al, an increase in electric field in certain regions (unlike Dennard’s CE scaling).

Use is made of different scaling factors for various parameters. For example,

 The scaled-down potentials/voltages ϕ′ = ϕ/k

 here, k>1

 The scaled-down dimensions (x′,y′,z′) = (x,y,x)/λ

 here, λ>1

And, the scaled-down concentrations (n′,p′,N′D,N′A) = (n,p,ND,NA)/δ

 It will be shown, following Baccarani et al that

δ = k/λ2, if potential distribution shapes are to be preserved in scaled-down devices.

**Relationships between various scaling factors:**

δ = k/λ2

 This relationship is derived by comparing the governing equations (e.g. Poisson and continuity) of unscaled (long-channel) and scaled-down devices. The idea is if these equations are similar then the shape of potential and field distributions will be preserved, and the device behavior will be long-channel (i.e. no short-channel effects).

In an unscaled device the distributions are obtained solving the following equations:

  (50)

 ∇.*J*n=0 (51)

 here  (52)

Using ϕ′, (x′,y′,z′), and (n′,p′,N′D,N′A) for a scaled-down MOSFET, we get the Poisson’s equation in the form

  (53)

 ∇.D=R=q(p-n+ND-NA)

 D = εr ε0ε

 Ɛ= -∇V

 = -∇ϕ

 let ϕ′ = ϕ/k (54)

(x′,y′,z′) = (x,y,x)/λ (55)

 and (n′,p′,N′D,N′A) = (n,p,ND,NA)/δ (56)

Eqs. (50), (54), (55), and (56) give



 or  (57)

Eqs. (53) and (57) are same if  (58)

 Eqs. (56) + (58) show

(n′,p′,N′D,N′A) = (n,p,ND,NA) λ2/k (59)

We have shown that Eqs. (50) and (53) are similar if transformations represented by Eqs. (54), (55), and (56) are followed.

The next question is if Eqs. (51) and (52) retain their shapes under variable transformation. It can be seen that these equations transform so long as the value of n (and p) is small and the variation ∇n (spatial coordinates) is negligible. This is true under subthreshold and (only partly) in saturation regimes.

Eqs. (54), (55) and (56) can now be used to obtain scaling factors for other parameters such as electric field, power dissipation, and gate delay. These are listed in Table III.

TABLE III: Generalized Scaling (1984)

|  |  |
| --- | --- |
| Parameter (unscaled) | Generalized scaling |
| Channel length (L) | L′ = L/λ |
| Channel width (w) | w′ = w/λ |
| Junction depth (xj) | xj′ = xj/λ |
| Oxide thickness (TOX) | TOX′ = TOX/λ |
| Potentials: VGS, VDS, ϕ | V′GS = VGS/kV′DS = VDS/kϕ′ = ϕ/k |
| Impurity Concentration (NA) | N′A = NA λ2/k |
| Electric Field (ε) | ε′ = ε λ/k |
| \* Total Capacitance (cwL) | c′w′L′ = (cwL)/λ |
| Power (P) | P′ = P λ/k3 |
| Current (I) \*\* | I′ = I λ/k2 (Unsaturated velocity) |

\* Capacitance per unit area C|| C′ = Cλ \*\* Room temp. unsaturated

 (F/cm2)

In addition, Baccarini et al have distinguished between scaling factors under 300°K (room temperature) and liquid nitrogen (77°K) conditions.

For example:

The scaling factor used for current with no saturation velocity effects is

 I′ = I λ/k2 , (60)

and if velocity saturation occurs

 I′ = I \* 1/k (61)

Eq. (61) is generally prescribed for 77°K operation.

The factors used in equations (60) + (61) can be obtained if

  (62)

Correction term

(not the scaling factor as used in Eq. 9)

(No velocity saturation)

Note: In the event IDS saturation due to velocity saturation (+ not pinch-off),

  (63)

 (64)

In Eq. (62) (under non velocity saturation conditions, and assuming no change in μn) only the cOX and voltage terms transform.

[λ due to c′OX = λcOX

 V′ \* V′ = V2/k2]

 Therefore, I′ = I λ/k2

However in Eq. (64) νmax remains constant and w, cOX + Voltages transform under scaling. Hence

 I′ (under velocity saturation) = I/k

 [c′OX = cOXλ, w′ = w/λ, V′ = V/k; the product yields 1/k]

## 9.7.4 Design Steps

0.25 μm MOSFET from a 1.3 μm MOSFET

Dimensional scaling factor λ = 1.3/0.25 ≅ 5.

1. Selection of an appropriate value of VTH (VTE)

This depends on the variation in VTE due to short-channel effects, VDS, operating temperature range (0-75°C), and process related tolerances (e.g. channel implant level fluctuations, oxide charges, etc.)

ΔVTE = 20% VTE

VTE ≈ 0.25 Volt.

Supply voltage VDD = 4VTE = 1.0 Volt.

In contrast, 1.3 μm MOSFET used 2.5 Volt.

Voltage scale down factor k = 2.5/1 = 2.5

1. Choosing a substrate bias VSB value

VSB = 0

1. 1/δ = λ2/k =  = 10

Once VTE, VSB, NA (background doping), and oxide thickness are selected one can determine other parameters related to ion implantations [For details, see Eq. 6 of Baccarani].

1. Finally, the effect of source and drain parasitic resistance (RS, RD), finite inversion layer thickness, and mobility degradation are incorporated to optimize the device operation.



Figure 11: Drain parasitic resistances RS and RD (not shown)

**9.7.5 NanoFET Design Example:**

(a) Outline the steps used in scaling a FET from 1.3 micron to 0.25 micron is described in Baccarani et al (1984) Table IV).

(b) Scale down a 0.25 micron FET to design the 0.025 micron (25nm) transistor following the 10-fold scaling. Given VTH = 0.015V for 25 nm process. Justify the values given in Table I.

For 0.025micron FET, use VTH = ¼ of VDD. Assume VDD of 0.4 Volt. Obtain gate oxide thickness, Si doping levels, source and drain thicknesses and doping etc.

Source and drain resistances are function of junction depth xj and source and drain length and width. Knowing the resistivity of the source and drain region, we can calculate the resistance of the source and drain regions.

 (c) Summarize in few lines CE, CV, QCV and GS scaling schemes. What do you think are the most difficult parameters to scale in the design of 25nm channel length FET?

**Solution**. NANOFET Design: (a) Outline the steps used in scaling a FET from 1.3 micron to 0.25 micron is described in Baccarani et al (1984 paper) and shown in Table IV below.

**Table IV Hint set for 25nm FET design**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **LATV****1.3micron** | **0.25 micron QMDT\* (IBM)** | **25nm SiGe** **(Home work)** | **Scaling** **factor** | **Comments** |
| Channel Length L(μm) | 1.3 | 0.25 | 0.025 | λ=10 |  |
| Gate InsulatorOxide tox (nm) | 25 | 5.0 (SiO2) | 0.5 (SiO2) | 10 | 0.5(HfO2 /SiO2)= 1.7nm |
| Junction Depth xj(nm) | 350 | 70-140 | 7-14 | 10 | Gives high Rs and Rd  |
| VTH (V)VTH 4xVTH | 0.6 | 0.25V’TH 4xV’TH V’TH =0.06 | 0.06V’TH 4xV’TH V’TH =0.015 | κ=4 | VTH=0.015V |
| VDS = VDD (V)VDD = 4 x VTH | 2.5 | 1.0 | 0.25 | κ=4 |  |
| Band Bending (V) | 1.8 | 0.8 | 0.3 | ? |  |
| Doping NA (cm-3) | 3x1015 | NA =3×1016 | N’A=7.5×1017 | NA x λ2/κ=25 |  |
| (Rs + Rd)\*IDIR Drop (mV) | NA | < 10mV | < 1mV | ?? | How to solve this |
| RC Delay  (ps) | Not appl. (NA) | 100ps | 2-5 ps?? | ?? | How to solve this |

**Solution**. NANOFET Design: (a) Outline the steps used in scaling a FET from 1.3 micron to 0.25 micron is described in Baccarani et al (1984 paper) and shown in Table I below.

**Design steps:**

1. Find the dimensional scaling factor l which is L/L’ or W/W’ or Tox/Tox’.

2. Find the V’TH for the new scaled down FET including process variations, masking registration errors, doping changes. Find V’TH= 4 V’TH. Find V’DD = 4 V’TH.

Calculate scaling factor for voltages and potentials: k = VTH/ V’TH

3. Find 

4. Compute new substrate doping N’A = NA/ = (2/k) NA

Generalized scaling is better than constant electric field (CE) scaling (Table–IV) and constant voltage (CV) and Quasi constant voltage (QCV) in as it permits different scaling for voltages/ potential Dimensions (L,W,Tox ) and doping. GS is based on preserving Poisson’s equation. GS gives relationship forNA’ andND’ with NA and ND multiplied by 1/δ.

**Scaled down FET(with prime) Original FET (L, W, Tox, NA)**

1. Dimensions

L’, W’, Tox’ $^{L}/\_{λ,} $ $^{W}/\_{λ,}^{T\_{ox}}/\_{λ}$ $ λ>1$

1. Voltages

VGS’,VDS’,VTH’ $^{ V\_{GS}}/\_{k,}^{V\_{DS}}/\_{k, }^{V\_{TH}}/\_{k}$ $k>1$

1. Doping Levels

NA’ or ND’ $^{ N\_{A}}/\_{δ,}^{N\_{D}}/\_{δ }$

Since $δ=^{k}/\_{λ^{2}}$ and (1/= 2/k, substrate doping NA’ is higher. [Pages 579-580 Notes III]

(b) Scale down a 0.25 micron FET to design the 0.025 micron (25nm) transistor following the 10-fold scaling. Given VTH = 0.015V for 25 nm process. Justify the values given in Table I.

For 0.025micron FET, use VTH = ¼ of VDD. Assume VDD of 0.4 Volt. Obtain gate oxide thickness, Si doping levels, source and drain thicknesses and doping etc.

The three design steps to obtain 0.25μm FET from 1.3μm FET is shown below

 Table V (Baccharini et al. 1984)

|  |  |  |
| --- | --- | --- |
| 1. 1.3 μm channel length L FET
 | 1. 0.25 μm FET channel length L’
 | 1. Scaling factor
 |
| 1. L = 1.3 μm
2. Tox = 25 nm
3. VDD = 2.5 V
4. VTH = 0.6 V
 | 1. L’ = 0.25 μm
2. Tox ’= 5 nm
3. VDD ’= 1 V
4. VTH ’= 0.25 V
 | 1. L’=$^{ L}/\_{λ,}$ $λ=$ $^{ 1.3}/\_{0.25} $= 5
2. VDD’=$^{ V\_{DD}}/\_{k,}$ $k=$ $^{ 2.5}/\_{1} $= 2.5
 |
| 1. Substrate doping NA
2. Scaling factor δ
 | 1. NA’ = NA$ \frac{ λ^{2}}{k}$ = NA / δ
2. δ = k / $λ^{2}$
 | 1. NA’ = NA$ \frac{25}{2.5}$ = NA \* 10
2. N’A =3×1016 cm-3
 |

These parameters are given in Reference 6 Baccharini et al. 1984.

Now our task is to scale 0.25μm to 0.025μm (25nm) FET.

Table VI

|  |  |  |
| --- | --- | --- |
| 1. L=0.25μm
 | 1. L’ =0.025μm = 25nm FET
 | Scaling factor |
| 1. L = 0.25 μm
2. Tox = 5 nm
3. VDD = 1.0 V
4. VTH = 0.25 V
5. NA =3\*1016cm-3.
 | 1. L’ = 0.025 μm
2. Tox’=5/10 =0.5 nm = 5Å
3. VDD’= 0.25V (given). VDD’= 4\* VTH’
4. k = VDD/ VDD’= 1.0/0.25 = 4
5. VTH’= 4 VTH’=4\*0.015 = 0.06V
6. k = VTH/ VTH’= 0.25/0.06 = 4
7. NA’ = NA/(1/= 2/k= 100/4=25.
8. NA’= 3\*1016 \*25 = 7.5\*1017 cm-3.
 | 1. L’/L = 10; 10 =

5Å results in direct tunneling; Tox’(HfO2) Tox’(SiO2)\*(HfO2/SiO2);HfO2~17-211. VDD’=VDD/k=1.0/0.25=4; So $k=$ 4

Voltage scaling factor k can be found from VTH or VDD scaling(1/= 2/k= 100/4=25 |

Certain issues following scaling:

1. Oxide thickness in 0.25μm FET is ≈ 5nm or 50Å. Scaled down Tox’ in 0.025μm is ≈ 5 Å, which is too small. As you have done home work 8, it will lead to direct tunneling of electrons.
2. Tunneling from source to drain takes place when channel length L’ = 25nm and lower.
3. The source and drain parasitic resistances lead to Ohmic voltage drop. This voltage driop [ID \* ( R S + RD )] should be < 1 mV.

(c) Generalized scaling (GS, Table IV) is better than constant electric field (CE, Table–II) and constant voltage (CV) and Quasi constant voltage (QCV) in Table-III as it permits different scaling for voltages/ potential, dimensions (L,W,Tox) and doping. GS is based on preserving Poisson’s equation. GS gives relationship for NA’ and ND’ with NA and ND multiplied by 1/δ.

**Constant electric (CE) field scaling:** All dimensions, voltages are divided by k (or scaling factor). Doping levels are multiplied by k.

**Quasi Constant Voltage (QCV) Scaling:** All dimensions are divided by k (or scaling factor). Doping levels are multiplied by k. The voltages are divided by (1/k)1/2 which is less than k.

**Constant Voltage (CV) Scaling:** All dimensions are divided by k (or scaling factor). Doping levels are multiplied by k. The voltages are not changed.

**Most difficult parameters are:**

(i) Gate Insulator. Solution is to increase Tox‘= 5 $A^{o}$ by using HfO2  which has a higher dielectric constant ɛHf O2 ≈ 13.7 and dielectric constant for SiO2 is 3.9.

Tox’for HfO2 as gate insulator is = $\frac{ɛ HfO2}{ɛ SiO2}$ \* Tox’(SiO2),

= $\frac{13.7}{3.9}$ \* 5 $A^{o}$ = 17.5 Å= 1.75 nm

Eg  for Hf O2  = 5.7 eV and electron affinity qHf O2  = (2.0 ± 0.25 ) eV

The other parameter is the voltage drop in R S + RD

(ii) Source and Drain resistances and voltage drops for a 25nm FET are calculated below.



L=22nm

Xj for 0.25 µm FET (Baccarani et al Table II) Ls= 22nm (contact) +11nm +11nm=44nm

 = 0.07 – 0.14 µm RS = RD  =ρn\* $\frac{Ls}{X\_{j \*W}}$

Xj’ for 0.025 µm NanoFET For $10^{20}$ n+ doping the resistivity s is

 =$\frac{ 0.07}{10}$ 🡪 $\frac{0.14 µm }{10}$ ρs  = 7 \* $10^{-4}$Ω-cm. Given W=44nm,

 = 0.007 µm 🡪0.014 µm RS = RD  =$\frac{7 \* 10^{-4}\*44 nm}{0.007 \* 10^{-4} \*44 nm}$

 = 7 nm 🡪 14 nm = 1.0 kΩ

 If the drain current is I D ≈ 1mA or 0.1 mA. Taking 0.1mA or 100 microamp, we get

 I D \* (R S + RD) ≈ 0. 2V =200mV, which is too high.

 **How can we reduce RS and RS ?**

Method # 1 Increase thickness Xj and contact thickness Tsource above the 14 nm source region.



New source resistance RS’ = Rs \*[Xj/(Xj + Tsource]  = $\frac{7 \* 10^{-7}}{154 \* 10^{-7} }$\*1.0 kΩ = 45.45Ω

The new IRs’ drop is

I \*RS’= 100 µA \* 45.45Ω = 4.5mV.This value is still higher than 1mV.

Method 2: (a) Increase W of the channel, and (b) further increase thickness of source TSource. Both are not desirable as they increase the area of FET (option a) and make processing difficult when using masks or metal contacts (option b).

Method 3. Use metal silicides for vertical extensions in place of n+ Si which have reduced resistivity.What is done is to replace Tsource contact with lower resistivity silicides or TiN or TaN. RS (with silicides) is reduced by a factor of 10. The new drop is 0.45mV.

Reduce RS  from45.4Ω to approximately 4.5 Ω.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameters | Symbols | BSIM 3.2 | BSIM4.0.0 | BSIM4.6.0 |
|  |  | NMOS | PMOS | NMOS | PMOS | NMOS | PMOS |
| Channel Length L(μm) | L | 0.5 | 0.5 | 0.18 | 0.18 | 0.05 | 0.05 |
| Gate Insulator Oxide tox (nm) | TOX | 15 | 15 | 4 | 4.2 | 1.4 | 1.4 |
| Junction Depth xj(nm) | XJ | 150 | 150 | 60 | 70 | 20 | 20 |
| VTH (V) at Vbs=0 | VTH0 | 0.6 | -0.7 | 0.3999 | -0.42 | 0.22 | -0.22 |
| 1. Channel doping NA (cm-3)
 | NCH | 2.3E+17 | 2.3E+17 | 5.95E+17 | 5.92E+17 | 3E+18 | 3E+18 |

**Capacitances:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameters |  | BSIM 3.2 | BSIM4.0.0 | BSIM4.6.0 |
|  |  | NMOS | PMOS | NMOS | PMOS | NMOS | PMOS |
| Channel Length L(μm) | L | 0.5 | 0.5 | 0.18 | 0.18 | 0.05 | 0.05 |
| Gate-Source Capacitance per meter of gate width F/m | CGSO | 3.40e-10 | 4.50e-10 | 2.786e-10 | 2.786e-10 | 6.23e-10  | 7.43e-10 |
| Gate-Drain Capacitance per meter of gate width F/m | CGDO | 3.40e-10 | 4.50e-10 | 2.786e-10 | 2.786e-10 | 6.23e-10 | 7.43e-10 |
| Gate-Substrate Capacitance per meter of gate length F/m2 | CGBO | 5.75e-10 | 5.75e-10 | 2.56e-11 | 2.56e-11 | 2.56e-11 | 2.56e-11 |
| 1. Source/drain sidewall junction capacitance per unit length F/m
 | Cjsw | 1.05e-10 | 1.19e-10 | 7.9e-10 | 1.44e-09 | 2e-10 | 2e-10 |
| 1. Bottom junction capacitance per unit area F/m2
 | Cj | 6.80e-04 | 5.28e-04 | 9.65028e-4  | 0.00138 | 0.0015 | 0.0015 |

1. *C. Hu, “Hot Electron effects in MOSFETs,” IEDM, pp. 176-179, 1983.*
2. *E. Takeda, Y. Ohji, and H. Kume, “High field effects in MOSFETs,” IEDM, pp. 60-63, 1985.*
3. *D. Kahng, Physics of MOS transistor (by Brews), 1981, Academic Press*.
4. *Ref: D. Kahng (1781)*
5. *Reference: VLSI Handbook, Einspruch (Ed.), Academic Press 1985; see chapter 37 by T. Gheewala.*
6. *Baccarani, Wordeman, and Dennard, “Generalized Scaling Theory and its Application to a ¼ Micrometer MOSFET Design,” IEEE Trans. Electron Devices, E0-31, pp. 452-462, April 1984.*
7. *Chatterjee, Hunter, Holloway, and Lin, “The Impact of Scaling
Laws on the Choice of n-Channel or p-Channel for MOS VLSI,” IEEE Electron Dev. Letters, EDL-1, pp. 220-223, October 1980.*
8. *Brews, Fichtner, Nicollian, and Sze, “Generalized Guide for MOSFET Miniaturization,” IEEE Electron Dev. Letters, EDL-1, pp. 2-4, January 1980.*
9. *Dennard, Gaenssten, Yu, Rideout, Bassous, and LeBlanc, “Design of Ion-Implanted MOSFETs with very small physical Dimensions,” IEEE J. Solid-State Circuits, SC-9, pp. 245-268, October 1974**.*

# Solved SET NANOFET Design

# FET Scaling Laws and Design of Nano-FETs-I :

Q1 NANOFET Design:

Table IV shows the parameter of a 0.025 micron (25nm) transistor following the 10-fold scaling of a 0.25 micron FET.

(a) Show that the design is sound and follows Generalized Scaling.

Given VTH = 0.015V for 25 nm process. Justify the values given in Table I.

For 0.025micron FET, use VTH = ¼ of VDD. Assume VDD of 0.4 Volt. Obtain gate oxide thickness, Si substrate doping levels, source and drain thicknesses and their doping etc.

(b) What are the critical parameters?

**Table IV 1.3 m to 0.25m and**

**0.25 m to 0.025 m (or 25nm) FET design**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **LATV****1.3micron** | **0.25 micron QMDT\* (IBM)** | **25nm SiGe** **(Home work)** | **Scaling** **factor** | **Comments** |
| Channel Length L(μm) | 1.3 | 0.25 | 0.025 | λ=10 |  |
| Gate InsulatorOxide tox (nm) | 25 | 5.0 (SiO2) | 0.5 (SiO2) | 10 | 0.5(HfO2 /SiO2)= 1.7nm |
| Junction Depth xj(nm) | 350 | 70-140 | 7-14 | 10 | Gives high Rs and Rd  |
| VTH (V)VTH 4xVTH | 0.6 | 0.25V’TH 4xV’TH V’TH =0.06 | 0.06V’TH 4xV’TH V’TH =0.015 | κ=4 | VTH=0.015V |
| VDS = VDD (V)VDD = 4 x VTH | 2.5 | 1.0 | 0.25 | κ=4 |  |
| Band Bending (V) | 1.8 | 0.8 | 0.3 | ? |  |
| Doping NA (cm-3) | 3x1015 | NA =3×1016 | N’A=7.5×1017 | NA x λ2/κ=25 |  |
| (Rs + Rd)IDIR Drop (mV) | NA | < 10mV | < 1mV | ?? | How to solve this |
| RC Delay  (ps) | Not appl. (NA) | 100ps | 2-5 ps?? | ?? | How to solve this |

Source and drain resistances are function of junction depth xj and source and drain length and width. Knowing the resistivity of the source and drain region, we can calculate the resistance of the source and drain regions.

 Q.2. NANOFET Design: Scale down a 0.025 micron FET to design the 0.00125 micron (12.5 nm) transistor following the 2-fold scaling. Given VTH = 0.010V for 12.5 nm process.

Justify your design in light of

(a) Source to drain tunneling. Find barrier height and width and compute probability of tunneling.

(b) Channel to gate tunneling.

(c) If VTH = 0.010V is not given to you, how would you proceed to compute it?

(d) What happens to current conduction if gate width is reduced to 8nm.

(e) What happens to current-voltage characteristics if both channel length and width are reduced to 8nm?

(f) Can you compute tunneling using Fowler-Nordheim equation and direct tunneling?

#  Solution Set

**Q1 NANOFET Design: (a) Show that the design is sound and follows Generalized Scaling.**

In this problem, we have given inconsistent VDD and VTH parameters. We have specified VDD=0.4V and also VTH= 0.015V. We will do design for both.

**First we assume** VDD of 0.4 Volt. Using VTH = ¼ of VDD we get VTH= 0.1V, V’TH =0.025V

Table IV shows the parameter of a 0.025 micron (25nm) transistor following the 10-fold scaling of a 0.25 micron FET.

For scaling to 0.025micron FET from 0.25 micron, the length, width and oxide thickness scaling factor is =10.

The voltage scaling k = 2.5

The doping multiplication factor is 2/k = 100/2.5 = 40=1/.

The new substrate doping NA’ = 40\*3x1016= 1.2x1018 cm-3

Obtain gate oxide thickness, 0.5nm, which is too thin and will cause direct tunnelihg. We use a high permittivity oxide such as HfO2. The new gate dielectric HfO2 is (13.7/3.9)\*0.5 nm = 3.5\*0.5 = 1.75nm. Generally, the HfO2 has higher permittivity of 21, the oxide thickens could be higher.

Source and drain thicknesses (xj) = 7-14 nm. The source and drain doping n+ = 1020 cm-3.

For Rs, Rd and voltage drop I(RS + RD), see the solution distributed in the last class (pp. 14, 15).

The band bending is related to kT\*(lnNA’/ni) = 0.47V. It is not linear.

**Table IV-A 1.3 m to 0.25m and**

**0.25 m to 0.025 m (or 25nm) FET design: Given** VDD=0.4V, VTH=0.06V

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **LATV****1.3micron** | **0.25 micron QMDT\* (IBM)** | **25nm SiGe** **(Home work)** | **Scaling** **factor** | **Comments** |
| Channel Length L(μm) | 1.3 | 0.25 | 0.025 | λ=10 |  |
| Gate InsulatorOxide tox (nm) | 25 | 5.0 (SiO2) | 0.5 (SiO2) | 10 | 0.5(HfO2 /SiO2)= 1.7nmHfO2 = 13.7SiO2=3.9 |
| Junction Depth xj(nm) | 350 | 70-140 | 7-14 | 10 | Gives high Rs and Rd  |
| VTH (V)VTH 4xVTH | 0.6 | 0.25V’TH 4xV’TH V’TH =0.06 | 0.1V’TH 4xV’TH V’TH =0.015 | κ=2.5 | VTH=0.06V |
| VDS = VDD (V)VDD = 4 x VTH | 2.5 | 1.0 | 0.4 | κ=2.5 |  |
| Band Bending (V) | 1.8 | 0.8 | 0.47 | 1.69 | Not linear, natural log 2 |
| Doping NA (cm-3) | 3x1015 | NA =3×1016 | N’A=1.2×1018 | NA x λ2/κ=40 |  |
| (Rs + Rd)ID | NA | < 10mV | < 1mV | ?? | IR Drop (mV) |
| RC Delay  (ps) | Not appl. (NA) | 100ps | 2-5 ps?? | ?? | How to solve this |

**Second** **assuming** V’TH =0.015V.

**Using the design criteria** V’TH = 4\* V’TH = 0.06V. The VDD=4 \*VTH = 0.24V~0.25V.

Table IV shows the parameter of a 0.025 micron (25nm) transistor following the 10-fold scaling of a 0.25 micron FET.

For scaling to 0.025micron FET from 0.25 micron, the length, width and oxide thickness scaling factor is =10.

The voltage scaling k = 4

The doping multiplication factor is 2/k = 100/4 = 25=1/. The new substrate doing

NA’ = 25\*3x1016= 7.5x1017 cm-3

Obtain gate oxide thickness, 0.5nm, which is too thin and will cause direct tunnelihg. We use a high permittivity oxide such as HfO2. The new gate dielectric HfO2 is (13.7/3.9)\*0.5 nm = 3.5\*0.5 = 1.75nm. Generally, the HfO2 has higher permittivity of 21, the oxide thickens could be higher.

Source and drain thicknesses (xj) = 7-14 nm. The source and drain doping n+ = 1020 cm-3.

For Rs, Rd and voltage drop I(RS + RD), see the solution distributed in the last class (pp. 14, 15).

The band bending is related to kT\*(lnNA’/ni) = 0.47V

**Table IV B V’TH =0.015V**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **LATV****1.3micron** | **0.25 micron QMDT\* (IBM)** | **25nm SiGe** **(Home work)** | **Scaling** **factor** | **Comments** |
| Channel Length L(μm) | 1.3 | 0.25 | 0.025 | λ=10 |  |
| Gate InsulatorOxide tox (nm) | 25 | 5.0 (SiO2) | 0.5 (SiO2) | 10 | 0.5(HfO2 /SiO2)= 1.7nm |
| Junction Depth xj(nm) | 350 | 70-140 | 7-14 | 10 | Gives high Rs and Rd  |
| VTH (V)VTH 4xVTH | 0.6 | 0.25V’TH 4xV’TH V’TH =0.06 | 0.06V’TH 4xV’TH V’TH =0.015 | κ=4 | VTH=0.015V |
| VDS = VDD (V)VDD = 4 x VTH | 2.5 | 1.0 | 0.25 | κ=4 |  |
| Band Bending (V) | 1.8 | 0.8V | 0.45 V | 1.77 | Not linear |
| Doping NA (cm-3) | 3x1015 | NA =3×1016 | N’A=7.5×1017 | NA x λ2/κ=25 |  |
| (Rs + Rd)IDIR Drop (mV) | NA | < 10mV | < 1mV | ?? | How to solve this |
| RC Delay  (ps) | Not appl. (NA) | 100ps | 2-5 ps?? | ?? | How to solve this |

**Q2. (b)** What are the critical parameters?

***1)*** Gate oxide thickness as it controls tunneling of electrons from inversion channel to the gate. This constitutes the gate current IG that is undesirable as it loads gate power supply VGS.

***2)*** IR drop (mV). This reduces the effective magnitude of VDD.

***3).*** RC delay  (ps). The delay depends on the product of (RS + RD) and capacitance. The capacitance is the CGD of stage #1 and CGS of stage #2. This gets little bit involved.