ECEN474/704: (Analog) VLSI Circuit Design Spring 2018

Lecture 4: MOS Transistor Modeling

Sam Palermo
Analog & Mixed-Signal Center
Texas A&M University

Agenda

- MOS Transistor Modeling
  - MOS Spice Models
  - MOS High-Order Effects
- Lab 1 begins Jan 31
- Current Reading
  - Razavi Chapters 2 & 17

Why Do We Need MOS Spice Models?

- Analog circuits are sensitive to detailed transistor behavior
  - Bias conditions set operation mode, gain, bandwidth, …
  - Can't simply use logical modeling methods, as in digital design flows

- Spice simulations allow us to predict the performance of complex analog circuits with models that capture high-order device operation
  - Much easier and cheaper than actually fabricating the circuit and performing physical measurements

MOS Level 1 Model

- Closely follows derived "Square-Law" Model

$$I_{DS} = \frac{W}{L-2L_D}\mu_n C_{OX}\left(V_{GS} - V_{Tn} - 0.5V_{DS}\right)V_{DS}\left(1+\lambda V_{DS}\right) \qquad \text{(Triode)}$$

$$I_{DS} = \frac{1}{2}\mu_n C_{OX}\frac{W}{L-2L_D}\left(V_{GS} - V_{Tn}\right)^2\left(1+\lambda V_{DS}\right) \qquad \text{(Saturation)}$$

$$V_T = V_{T0} + \gamma\left[\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F}\right]$$

- Note, extra $1+\lambda V_{DS}$ term in triode equation is to have continuity between triode and saturation regions
- Reasonably accurate I/V characteristics for devices with $L \geq 4\mu m$, but models output resistance poorly
- Neglects subthreshold conduction and many high-order effects in shorter-channel length devices

MOS Level 1 Model Parameters

[Razavi]

Table 2.1 Level 1 SPICE Models for NMOS and PMOS Devices.

NMOS Model			
LEVEL = 1	VTO = 0.7	GAMMA = 0.45	PHI = 0.9
NSUB = 9e+14	LD = 0.08e−6	UO = 350	LAMBDA = 0.1
TOX = 9e−9	PB = 0.9	CJ = 0.56e−3	CJSW = 0.35e−11
MJ = 0.45	MJSW = 0.2	CGDO = 0.4e−9	JS = 1.0e−8
PMOS Model			
LEVEL = 1	VTO = −0.8	GAMMA = 0.4	PHI = 0.8
NSUB = 5e+14	LD = 0.09e−6	UO = 100	LAMBDA = 0.2
TOX = 9e−9	PB = 0.9	CJ = 0.94e−3	CJSW = 0.32e−11
MJ = 0.5	MJSW = 0.3	CGDO = 0.3e−9	JS = 0.5e−8

VTO: threshold voltage with zero V_{SB} (unit: V)
GAMMA: body effect coefficient (unit: $V^{1/2}$)
PHI: $2\Phi_F$ (unit: V)
TOX: gate oxide thickness (unit: m)
NSUB: substrate doping (unit: cm^{-3})
LD: source/drain side diffusion (unit: m)
UO: channel mobility (unit: $cm^2/V/s$)
LAMBDA: channel-length modulation coefficient (unit: V^{-1})
CJ: source/drain bottom-plate junction capacitance per unit area (unit: F/m^2)
CJSW: source/drain sidewall junction capacitance per unit length (unit: F/m)
PB: source/drain junction built-in potential (unit: V)
MJ: exponent in CJ equation (unitless)
MJSW: exponent in CJSW equation (unitless)
CGDO: gate-drain overlap capacitance per unit width (unit: F/m)
CGSO: gate-source overlap capacitance per unit width (unit: F/m)
JS: source/drain leakage current per unit area (unit: A/m^2)

MOS Level 2 Model

- Improves upon the Level 1 model by modeling
  - V_T variation along the channel
  - $\lambda(V_{DS})$
    - Output conductance increases as V_{DS} increases
  - Mobility degradation due to vertical field and velocity saturation
  - Subthreshold Behavior
  - V_T dependencies on transistor W & L
- Contains 5-10 more parameters than level 1 model
- Reasonably accurate I/V characteristics for devices with $L \geq 0.7\mu m$, but still poorly models output resistance and transition point between saturation and triode

MOS Level 3 Model

- Similar in complexity to the level 2 model, but computationally more efficient

- Adds Drain-Induced Barrier Lowering effect
  - V_{DS} can lower effective V_T

- Different model for mobility degradation

Velocity Saturation

- Square-law model assumes carrier drift velocity is proportional to lateral E-field

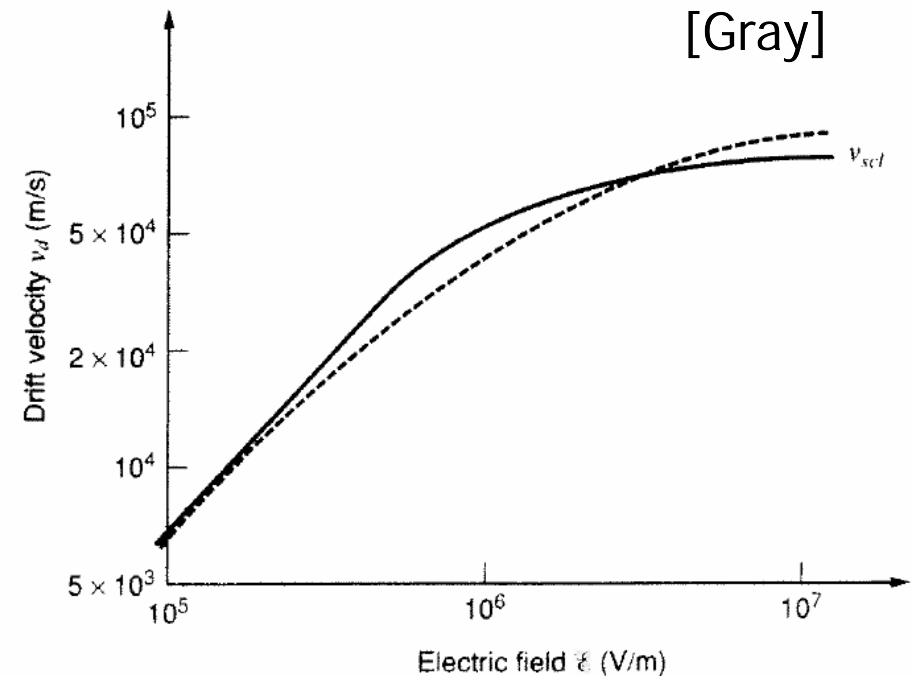
$$v_d = \mu E$$

- However, near critical electric field E_C, carrier velocity in silicon saturates due to scattering

$$v_{scl} \approx 10^5 \ m/s$$

- Causes reduction in ID relative to square-law model

$$I_D \propto \left(V_{GS} - V_T\right)$$ (fully velocity saturated)



$$v_d\left(E\right) \cong \frac{\mu E}{1 + \frac{E}{E_c}} \cong \mu E_c = v_{scl} \quad \textbf{for} \quad E >> E_c$$

$$= \frac{v_{scl}}{2} \quad \textbf{for} \quad E = E_c$$

Mobility Degradation

- Carrier mobility is degraded by lateral E-field induced velocity saturation AND vertical electric field strength

- Vertical field attracts carriers closer to silicon surface where surface imperfections impede movement

Level 2 Model

$$\mu = \mu_0 \left(\frac{\varepsilon_{si}}{C_{ox}} \frac{U_{CRIT}}{V_{GS} - V_T - U_{TRA} V_{DS}} \right)^{U_{EXP}}$$

$$U_{CRIT} = \textbf{Gate - Channel Critical Field}$$

$$U_{TRA} = \textbf{Fitting Parameter (0 - 0.5)}$$

$$U_{EXP} = \textbf{Exponent (~ 0.15)}$$

Level 3 Model

$$\mu = \frac{\mu_{eff}}{1 + \frac{\mu_{eff} V_{DS}}{v_{max} L}}$$

$$\mu_{eff} = \frac{\mu_0}{1 + \theta (V_{GS} - V_T)}$$

$$v_{max} = \textbf{Max Carrier Velocity}$$

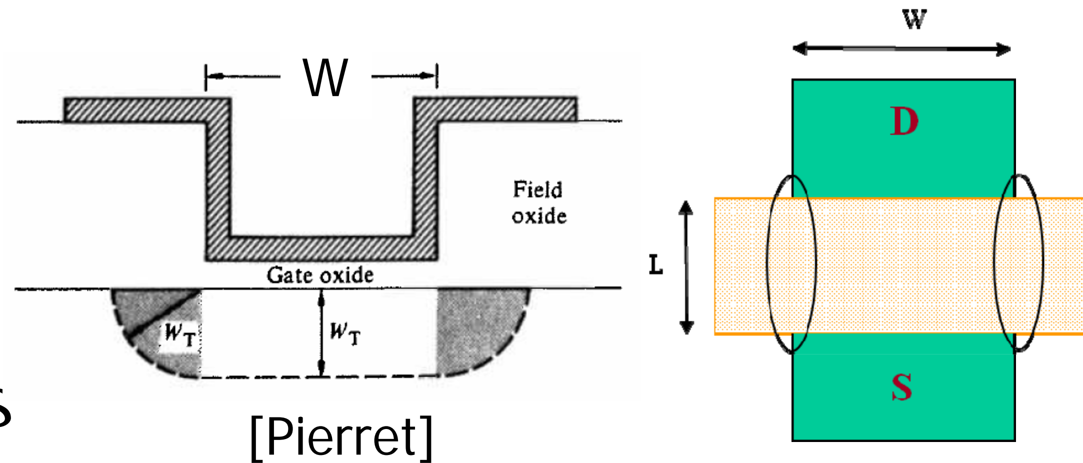
$$\theta = \textbf{Mobility Modulation Parameter (0.1 - 0.4 V}^{-1}\textbf{)}$$

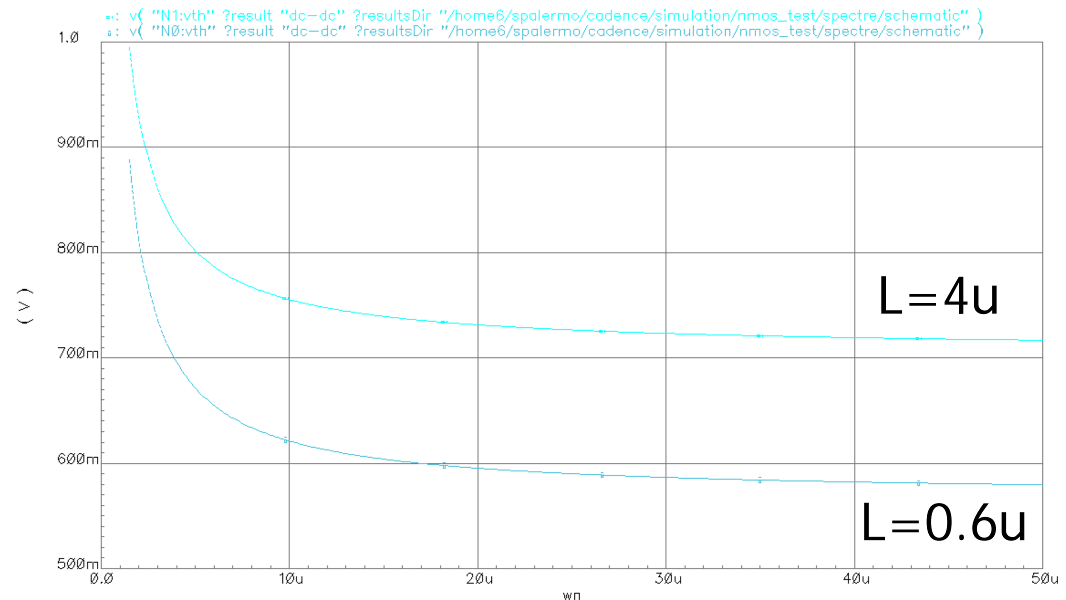
V_T Dependency on W

- Gate-controlled depletion region extends in part outside the gate width
- V_T monotonically increases with decreasing channel width

$$V_T = V_{Twide} + \Delta V_T$$

$$\Delta V_T = \frac{qN_A W_T}{C_{ox}} \frac{\pi W_T}{2W}$$

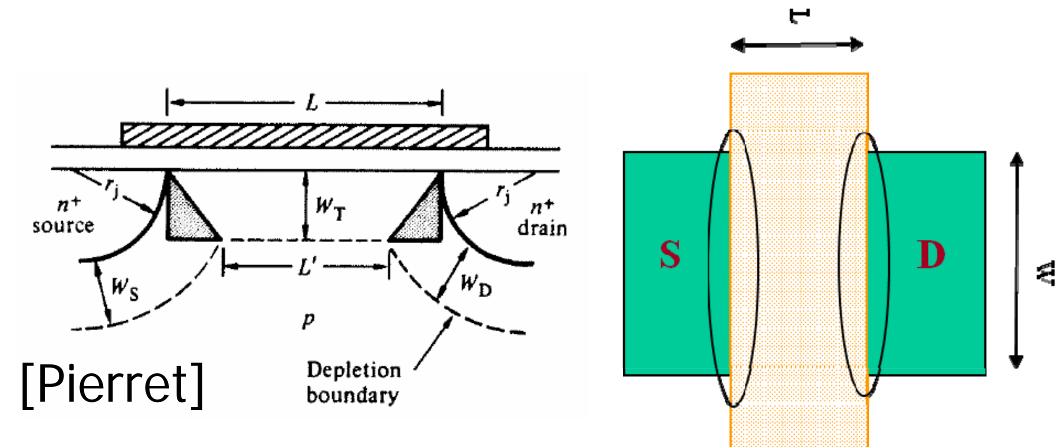


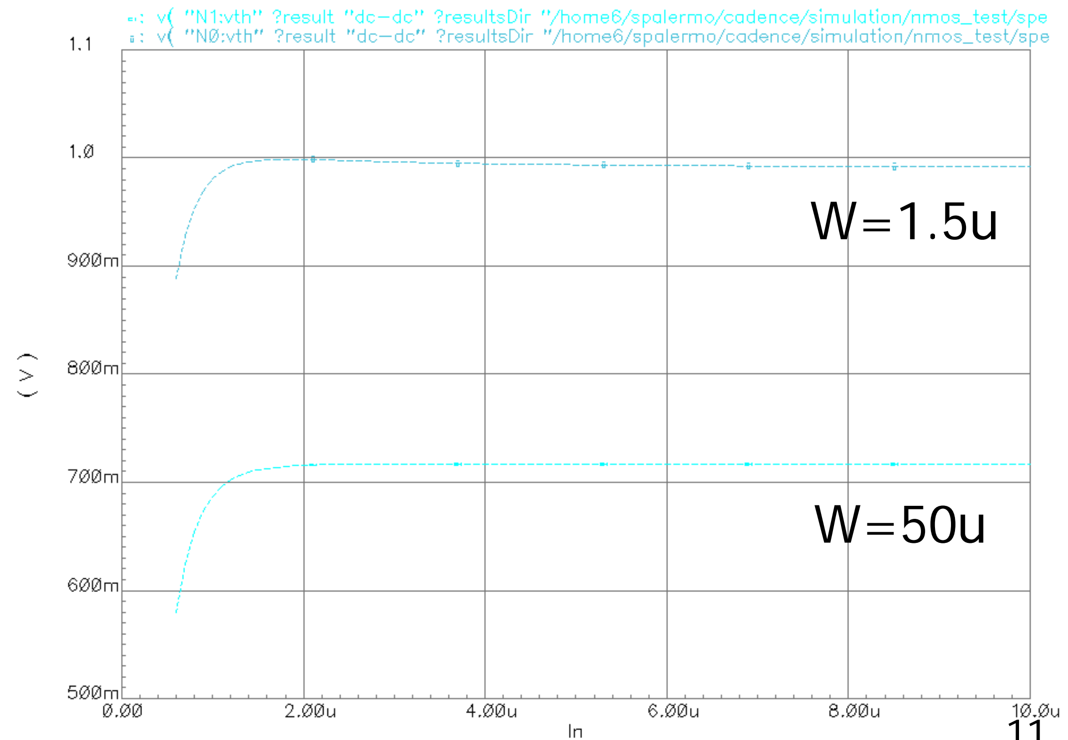


V_T Dependency on L

- Source and drain assist in forming the depletion region under the gate

- With simple model, V_T monotonically decreases with decreasing channel length

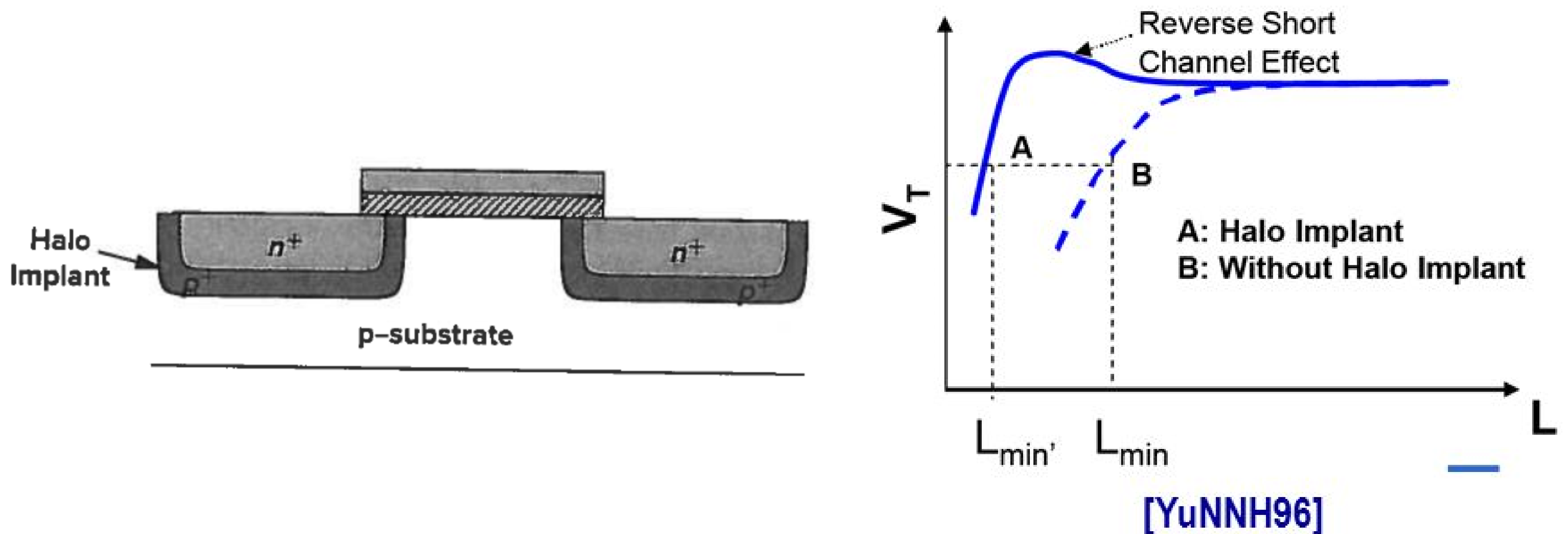




$$V_T = V_{Tlong} + \Delta V_T$$

$$\Delta V_T = -\frac{qN_A W_T}{C_{ox}} \frac{r_j}{L}\left(\sqrt{1+\frac{2W_T}{r_j}} - 1\right)$$

V_T Dependency on L: Reverse Short Channel Effect

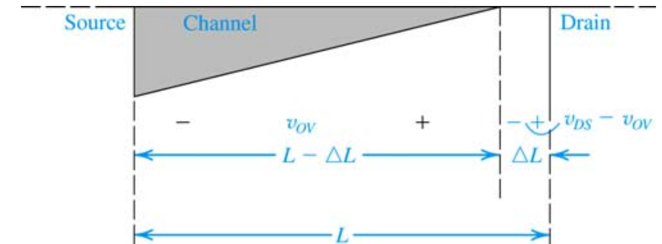


- Modern device structures use a "halo" implant of heavy doping that surrounds the source and drain junctions
- This locally increases the threshold in these regions, causing an increase in the overall threshold as L is decreased
- At very small Ls, the effect on the previous slide may dominate

SPICE LEVEL3 (NON-LINEAR OUTPUT RESISTANCE)

- Recall in the simple transistor model

$$\Delta L = \lambda' V_{DS}$$



In the L3 model, a KAPPA parameter is introduced for the computation of the channel length modulation:

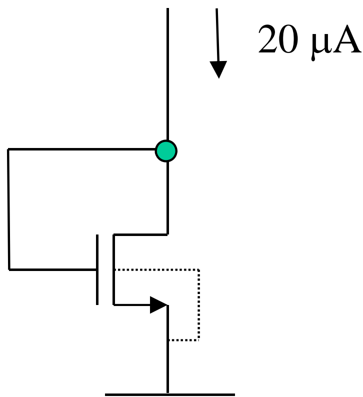
COMPUTATION OF ΔL IS MORE COMPLEX BUT MORE PRECISE:

$$\Delta L = \sqrt{\left(\frac{E_P X_D{}^2}{2}\right)^2 + K X_D{}^2 \left(V_{DS} - V_{dsat}\right)} - \frac{E_P X_D{}^2}{2}$$

Notice that ΔL is funtion of:

- **X_D (technology parameter)**
- **V_{GS}-V_T (Saturation voltage) Design parameter.**
- **Output resistance is function of L, I_D and V_{DSAT}!! R~ $\Delta L/I_D$**

Design Example: Constant current

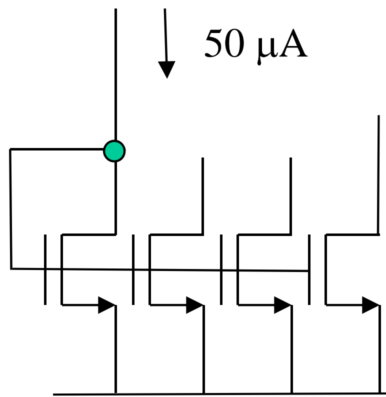



```
m1   vd1   vd1    0    0   l=0.8u  w=10u  ad=20p  ps=30u  pd=30u
m2   vd2   vd2    0    0   l=1.6u  w=20u  ad=40p  ps=40u  pd=40u
m3   vd3   vd3    0    0   l=2.4u  w=30u  ad=60p  ps=50u  pd=50u
m4   vd4   vd4    0    0   l=3.2u  w=40u  ad=60p  ps=50u  pd=50u
m5   vd5   vd5    0    0   l=4u    w=50u  ad=60p  ps=50u  pd=50u
```

element	0:m1	0:m2	0:m3	0:m4	0:m5
id	20.0000u	20.0000u	20.0000u	20.0000u	20.0000u
vgs	886.9381m	969.8944m	995.7306m	1.0083	1.0158
vds	886.9381m	969.8944m	995.7306m	1.0083	1.0158
vth	666.2501m	736.9148m	758.7246m	769.3286m	775.5978m
vdsat	154.8212m	165.3264m	168.8038m	170.5361m	171.5733m
beta	1.0881m	1.0045m	979.6687u	967.7676u	960.7872u
gam eff	571.2190m	652.2787m	677.2968m	689.4607m	696.6521m
gm	179.7789u	171.7391u	169.1462u	167.8681u	167.1073u
gds	4.6916u	2.0549u	1.3150u	966.8425n	764.4227n
gmb	47.5488u	57.1768u	59.8797u	61.1497u	61.8872u
cdtot	14.6665f	24.6840f	34.7712f	37.7241f	40.7319f
cgtot	12.3872f	40.7364f	85.1087f	145.5042f	221.9229f
cstot	23.4292f	57.2574f	106.5620f	162.0429f	233.0002f
cbtot	26.4024f	43.6549f	61.4591f	63.3113f	65.7005f
cgs	9.1292f	33.6574f	73.6620f	129.1429f	200.1002f
cgd	2.9110f	5.8540f	8.8294f	11.8370f	14.8771f

- While W/L is constant, V_T increases as transistor dimensions increase
- Check simulation operating point information to aid in the design procedure

Design Example: Constant voltages



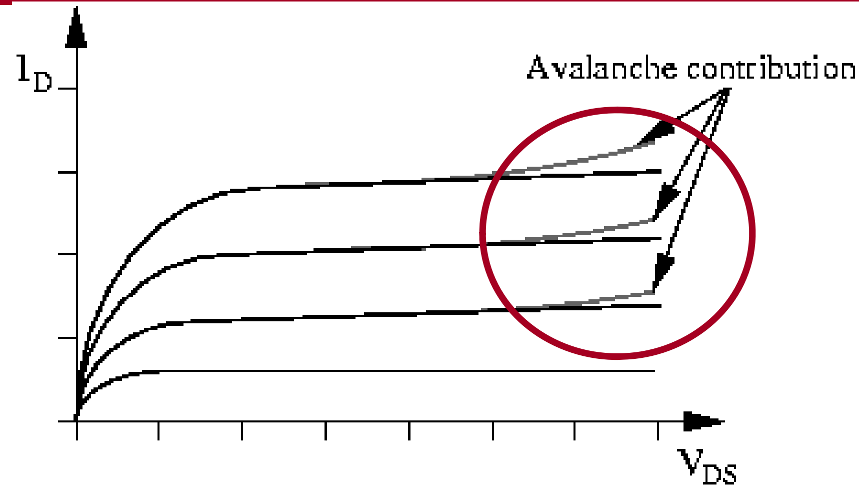
Array of transistors; only 4 transistors are shown.

```
i1   vdd   vd1    dc   50u
m1   vdd1  vd1    0    0    nmos  l=0.8u  w=10u  ad=20p  ps=30u  pd=30u
m2   vdd   vd1    0    0    nmos  l=0.8u  w=10u  ad=20p  ps=30u  pd=30u
m3   vdd   vd1    0    0    nmos  l=1.6u  w=20u  ad=40p  ps=40u  pd=40u
m4   vdd   vd1    0    0    nmos  l=2.4u  w=30u  ad=60p  ps=50u  pd=50u
m5   vdd   vd1    0    0    nmos  l=3.2u  w=40u  ad=60p  ps=50u  pd=50u
m6   vdd   vd1    0    0    nmos  l=4u    w=50u  ad=60p  ps=50u  pd=50u
```

element	0:m1	0:m2	0:m3	0:m4	0:m5	0:m6
id	50.0000u	54.0916u	29.4063u	23.7290u	21.2627u	19.8905u
vgs	1.0131	1.0131	1.0131	1.0131	1.0131	1.0131
vth	664.089m	656.629m	733.0935m	756.3653m	767.6259m	774.2667m
vdsat	234.1784m	239.2634m	197.9108m	182.8535m	175.2647m	170.7093m
beta	1.0921m	1.1291m	1.0195m	988.7500u	974.2790u	965.8602u
gam eff	568.7409m	560.1827m	647.8953m	674.5905m	687.5075m	695.1252m
gm	281.3687u	298.2144u	210.1394u	185.3381u	173.8114u	167.1665u
gds	8.4442u	7.4984u	2.1990u	1.2244u	838.6260n	635.1918n
gmb	71.9153u	75.4782u	69.0048u	65.2081u	63.1445u	61.8658u
cdtot	14.4417f	13.7280f	23.3207f	32.9599f	36.0176f	39.1218f
cgtot	12.3095f	12.3127f	40.6138f	85.0154f	145.5222f	222.1355f
cstot	23.4292f	23.4292f	57.2574f	106.5620f	162.0429f	233.0002f
cbtot	26.0967f	25.3740f	42.1100f	59.4235f	61.3917f	63.9436f
cgs	9.1292f	9.1292f	33.6574f	73.6620f	129.1429f	200.1002f
cgd	2.9126f	2.9187f	5.8836f	8.8949f	11.9526f	15.0568f

- Ideally, all transistors should have the same current!
- For best current mirror performance, use same W and L and scale the finger number

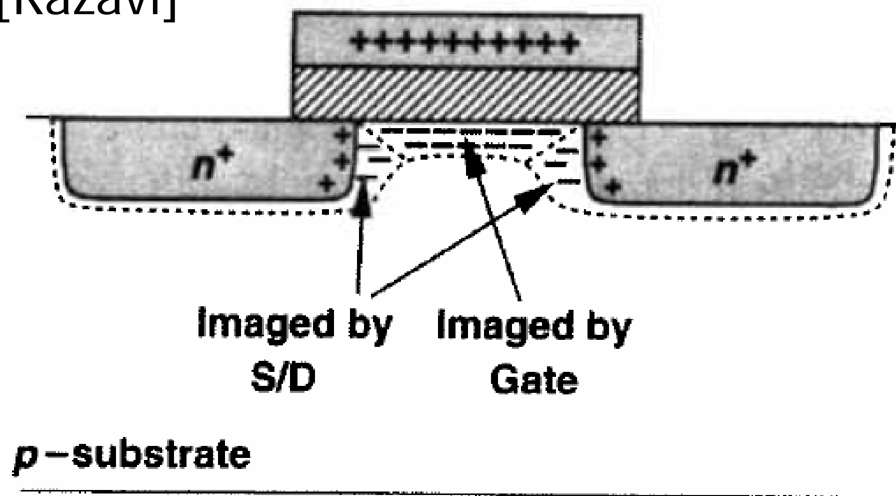
- Avalanche: drain current I_D and a substrate current I_B

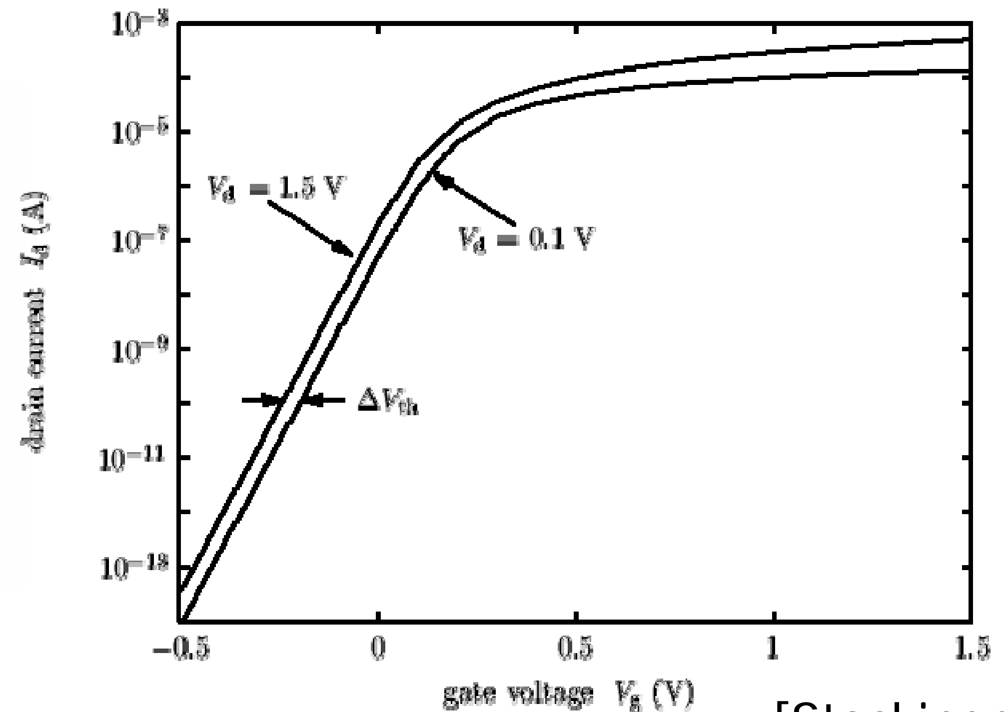


- The substrate current may contribute to latch-up
- The device noise increases
- The output impedance decreases
- Carriers can be trapped on the oxide and the V_{Th} changes (hot electron effect)

Drain Induced Barrier Lowering

[Razavi]





[Stockinger]

- Drain potential controls channel charge also
- Higher V_{DS} reduces barrier to the flow of charge, resulting in a net reduction in the threshold voltage

BSIM Model

- Berkeley Short-Channel IGFET Model (BSIM)
  - Industry standard model for modern devices
  - BSIM3v3 is model for this course
- Typically 40-100+ parameters
  - Advanced software and expertise needed to perform extraction

Example 0.6μm Technology Model (NMOS)

```
*N8BN SPICE BSIM3 VERSION 3.1 (HSPICE Level 49) PARAMETERS
* level 11 for Cadence Spectre
* DATE: Jan 25/99
* LOT: n8bn                WAF: 03
* Temperature_parameters=Default
.MODEL ami06N NMOS ( LEVEL=11 &
VERSION=3.1 &
TNOM=27 &
TOX=1.41E-8 &
XJ=1.5E-7 &
NCH=1.7E17 &
VTH0=0.7086 &
K1=0.8354582 &
K2=-0.088431 &
K3=41.4403818 &
K3B=-14 &
W0=6.480766E-7 &
NLX=1E-10 &
DVT0W=0 &
DVT1W=5.3E6 &
DVT2W=-0.032 &
DVT0=3.6139113 &
DVT1=0.3795745 &
DVT2=-0.1399976 &
U0=533.6953445 &
UA=7.558023E-10 &
UB=1.181167E-18 &
UC=2.582756E-11 &
VSAT=1.300981E5 &
A0=0.5292985 &
AGS=0.1463715 &
B0=1.283336E-6 &
B1=1.408099E-6 &
KETA=-0.0173166 &
A1=0 &
A2=1 &
RDSW=2.268366E3 &
PRWG=-1E-3 &
PRWB=6.320549E-5 &
WR=1 &
WINT=2.043512E-7 &
LINT=3.034496E-8 &
XL=0 &
XW=0 &
DWG=-1.446149E-8 &
DWB=2.077539E-8 &
VOFF=-0.1137226 &
NFACTOR=1.2880596 &
CIT=0 &
CDSC=1.506004E-4 &
CDSCD=0 &
CDSCB=0 &
ETA0=3.815372E-4 &
ETAB=-1.029178E-3 &
DSUB=2.173055E-4 &
PCLM=0.6171774 &
PDIBLC1=0.185986 &
PDIBLC2=3.473187E-3 &
PDIBLCB=-1E-3 &
DROUT=0.4037723 &
PSCBE1=5.998012E9 &
PSCBE2=3.788068E-8 &
PVAG=0.012927 &
DELTA=0.01 &
MOBMOD=1 &
PRT=0 &
UTE=-1.5 &
KT1=-0.11 &
KT1L=0 &
KT2=0.022 &
UA1=4.31E-9 &
UB1=-7.61E-18 &
UC1=-5.6E-11 &
AT=3.3E4 &
WL=0 &
WLN=1 &
WW=0 &
WWN=1 &
WWL=0 &
LL=0 &
LLN=1 &
LW=0 &
LWN=1 &
LWL=0 &
CAPMOD=2 &
XPART=0.4 &
CGDO=1.99E-10 &
CGSO=1.99E-10 &
CGBO=0 &
CJ=4.233802E-4 &
PB=0.9899238 &
MJ=0.4495859 &
CJSW=3.825632E-10 &
PBSW=0.1082556 &
MJSW=0.1083618 &
PVTH0=0.0212852 &
PRDSW=-16.1546703 &
PK2=0.0253069 &
WKETA=0.0188633 &
LKETA=0.0204965       )
```

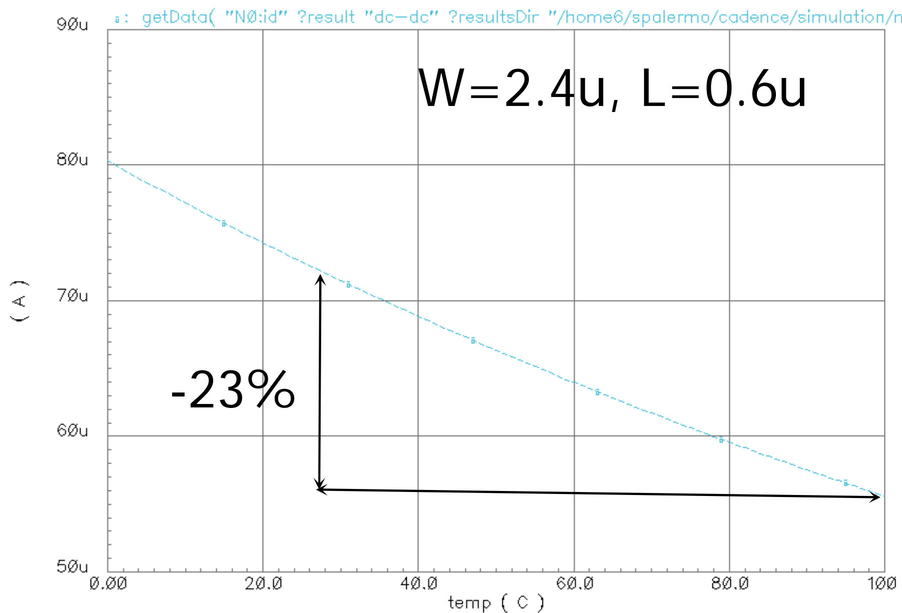
Temperature Dependence

- Transistor mobility and threshold voltage are dependent on temperature
  - Mobility $\propto T^{-3/2}$ due to increased scattering
  - Threshold voltage decreases with temperature due to reduced bandgap energy

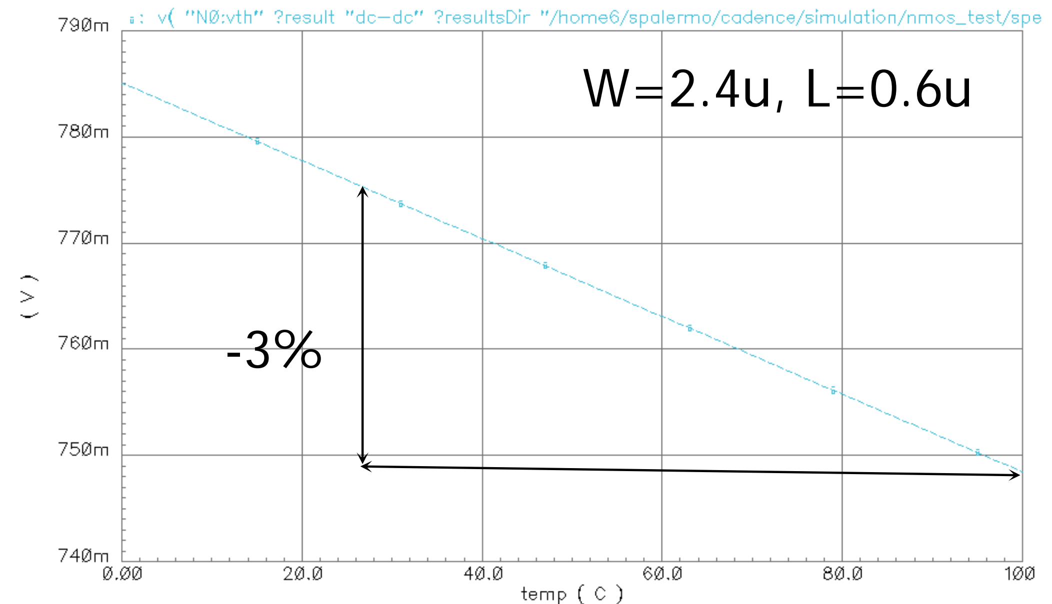
$$\mu = \mu_0 \left(\frac{300}{T} \right)^{3/2}$$

$$E_g = 1.16 - \frac{7.02 \times 10^{-4} T^2}{T + 1108}$$

I_D vs Temperature

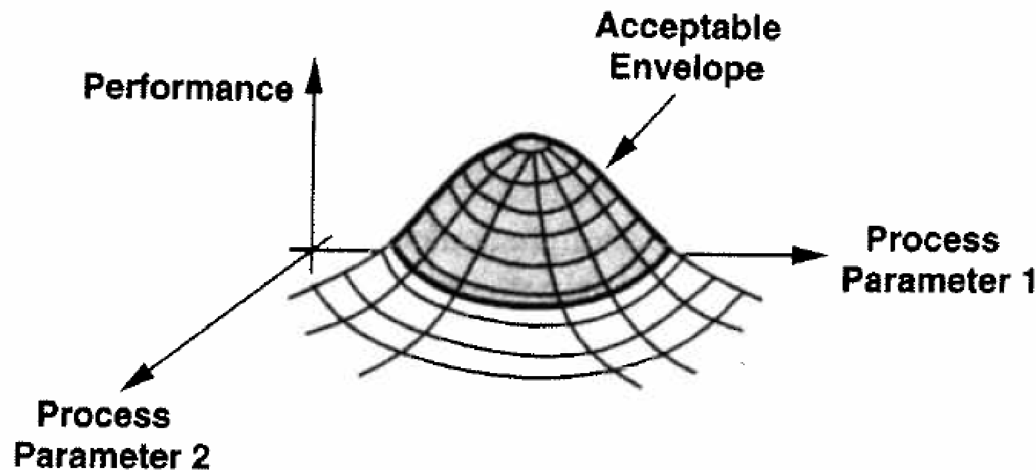


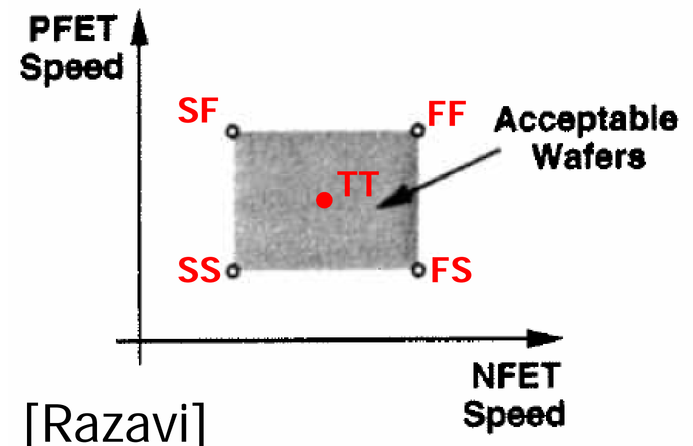
V_T vs Temperature



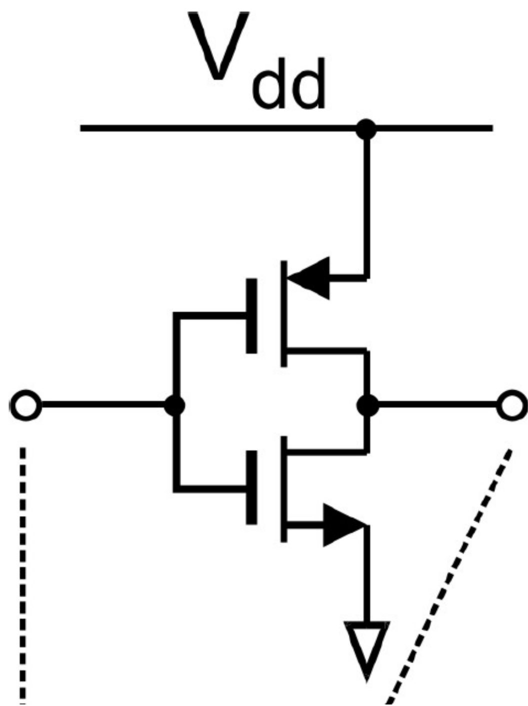
Process Corners

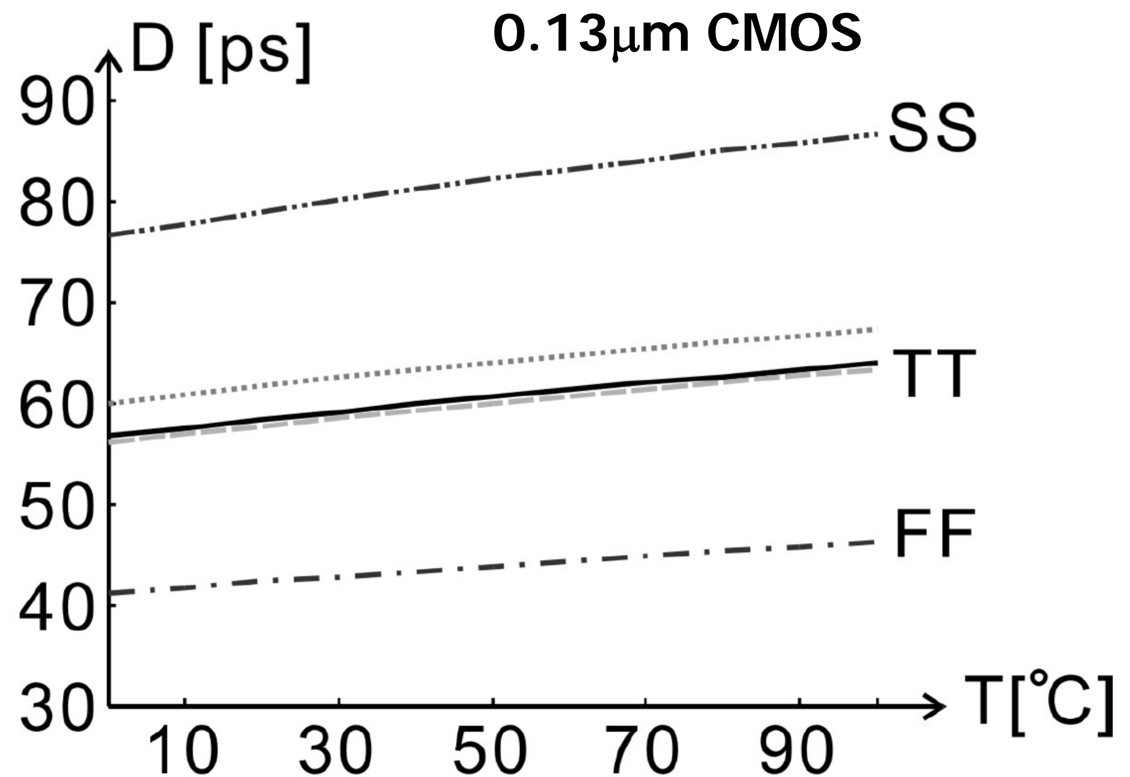
- Substantial process variations can exist from wafer to wafer and lot to lot
- Device characteristics are guaranteed to lie in a performance envelope
- To guarantee circuit yield, designers simulate over the "corners" of this envelope
- Example: Slow Corner
  - Thicker oxide (high V_T, low C_{ox}), low μ, high $R_{\square}$





Inverter Delay Variation with Process & Temperature





[Woo ISSCC 2009]

- CMOS inverter delay varies close to ±40% over process and temperature
- Also need to consider variations in supply voltage (PVT)

Next Time

- Layout Techniques