Lecture 1.

CMOS PROCESS

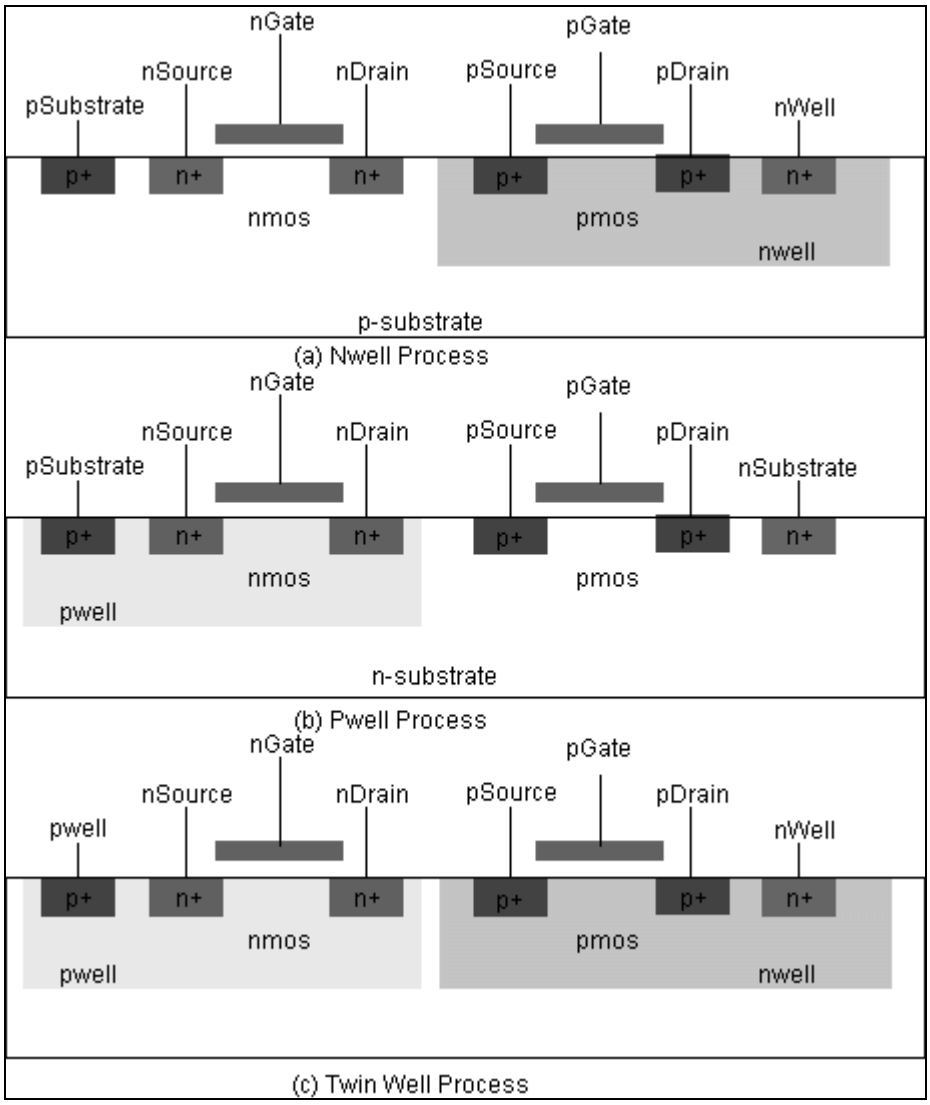

Figure 1. Three types of CMOS processing: (a) nwell, (b) pwell, and (c) twin nwell

In complimentary MOS (CMOS) technology, both PMOS and NMOS devices are used. Since the PMOS and NMOS devices require substrate material of opposite type of doping, at least two different CMOS technologies occur. The cross section of an n-well CMOS technology is shown in Fig. 1(a). The PMOS transistor is located in a deep, lowly doped n-well that serves as its bulk. The NMOS, on the contrary, is located directly on the p-substrate material. The opposite is true for p-well CMOS technology (see Fig. 1(b)). In a twin-well process (see Fig. 1(c).) both transistors are located in separate wells.

Native transistors are transistors that lie directly in the substrate whereas well transistors are transistors that lie in wells. In an n-well CMOS technology, the wells are n-type. The native transistors have n-type sources and drains, and the well transistors have p-type sources and drains (see Fig. 1(a)). The channels formed by the native transistors in the p-type substrate will be n-channels.

All MOSFET chips are extremely prone to damage by static electricity. The current through the transistors is controlled by an insulated gate. Even a few tens of volts can blow out the gate. A short walk across the room can build up kilovolts of static potential. There are electrostatic discharge (ESD) protection structures on the chip, but often this will not be enough. There are two simple precautions that can definitely keep the chip safe.

- When the chip is not powered up in a socket, keep it stuck into a piece of black conductive foam. This will short the susceptible inputs to both the power supply and ground pins.
- Always ground yourself to chassis ground before picking up or touching a chip. This will discharge the static charge.

An important difference between p-well and n-well CMOS technologies is the doping levels of the substrate and well. Typical doping levels for the substrate material are approximately 2×10^{14} to 10^{15} cm^{-3} . Since the wells are realized by means of diffusion, they are doped at a higher level than the substrate itself. Typical doping levels of the wells are about 10^{16} cm^{-3}. As a result, the bulk doping level of an nMOST in a p-well CMOS technology is much higher than in an n-well CMOS technology. Typically this ratio is a factor of 10 to 50. These two values of bulk doping levels will give different values of transistor parameters.

To clarify the meaning of the terms substrate, bulk, and well. The substrate is always the material just underneath the gate. For n-well CMOS technology, the p-substrate is the substrate for the NMOS; on the other hand, the n-well is also the substrate for the PMOS. The term bulk (B) is used instead of substrate to avoid confusion with the use of S to denote source. The opposite is true for p-well CMOS technology (see Fig. 1).

For n-well CMOS process, the bulk of the PMOS is the n-well. It is isolated from the substrate and thus can be connected to the source. On the other hand, the bulk of the NMOS is the substrate itself and thus the bulk of the NMOS can't be connected to the source. If you do, all the sources of the different NMOS transistors will be connected to each other. The opposite is true for p-well CMOS technology (see Fig. 1).

The NMOS and PMOS double-metal, double-poly processes are each analogous to a five level printed-circuit board, consisting of five levels of conducting materials, each layer electrically isolated from the layers immediately above and below by silicon dioxide.

Figure 2. CMOS layers.

These layers are color coded according to the different photolithographic masks needed to manufacture the devices.

Oxide thickness for the 2 μ, double poly, double metal n-well CMOS process (Orbit CN2, Low Noise Analog) are as follows:

Oxide thickness (angstroms)	Min	Typ	Max
Poly 1 gate oxide	370	400	430
Poly 2 oxide	470	500	530
Field oxide (poly 1 &2 to sub)	5500	6000	6500

Metal 1 to poly 1 & 2	8000	8500	9000
Metal 1 to sub	13500	14500	15500
Metal 1 to n+/p+ diff	8500	9000	9500
Metal 2 to metal 1	6000	6500	7500
Poly 1 to poly 2	650	750	850

Metal can cross polysilicon or diffused areas with no functional effect other than to produce a parasitic capacitance. Polysilicon crossing a diffused area creates a transistor; wherever a poly cross an nplus region an n-channel transistor is formed, and a p-channel transistor is formed wherever poly crosses a pplus region.

SCMOS Rules:

1. Metal 2 can only connect to metal1.
2. Metal3 can only connect to metal2.
3. Nplus and Pplus can not be connected directly. Must connect through metal1.
4. To prevent latch up tub ties (well/substrate ohmic contacts) must be placed every one to two transistors.
5. Metal 1 is used for power VDD, and VSS; it usually runs horizontally.
6. Metal2 is used for signal (I/O).
7. Via is used to connect metal1 and metal2 only
8. Contact is used for connecting poly to metal1, substrate to poly, and substrate to metal1.

ohmic contact means low resistance connection. In nwell/n-substrate region, connection is done at n+(nplus) diffusion area; while in pwell/p-substrate region, connection is done at p+ (pplus) diffusion area.

Layer	GDS	CIF	CIF Synonym	Rule Section	Notes
N_WELL	42	CWN		1	
ACTIVE	43	CAA		2	
POLY	46	CPG		3	
N_PLUS_SELECT	45	CSN		4	
P_PLUS_SELECT	44	CSP		4	
POLY2	56	CP2	CEL	11, 12, 13	Optional
HI_RES_IMPLANT	34	CHR		27	Optional
CONTACT	25	CCC	CCG	5, 6, 13	
POLY_CONTACT	47	CCP		5	Can be replaced by CONTACT
ACTIVE_CONTACT	48	CCA		6	Can be replaced by CONTACT
POLY2_CONTACT	55	CCE		13	Can be replaced by CONTACT.
METAL1	49	CM1	CMF	7	
VIA	50	CV1	CVA	8	
METAL2	51	CM2	CMS	9	
VIA2	61	CV2	CVS	14	
METAL3	62	CM3	CMT	15	
GLASS	52	COG		10	
PADS	26	XP			Non-fab layer used to highlight pads
Comments	--	CX			Comments

It is sometimes easier to use Polysilicon (Poly) as a routing layer in double metal process. This prevents using the formation of vias, which complicate the mask generation and end up using more area. Polysilicon is essentially formed by doping Si heavily with donor impurities (since electrons have higher mobility, lower resistivity is achieved for the same impurity concentration as compared to acceptor doping).

NMOS Layout

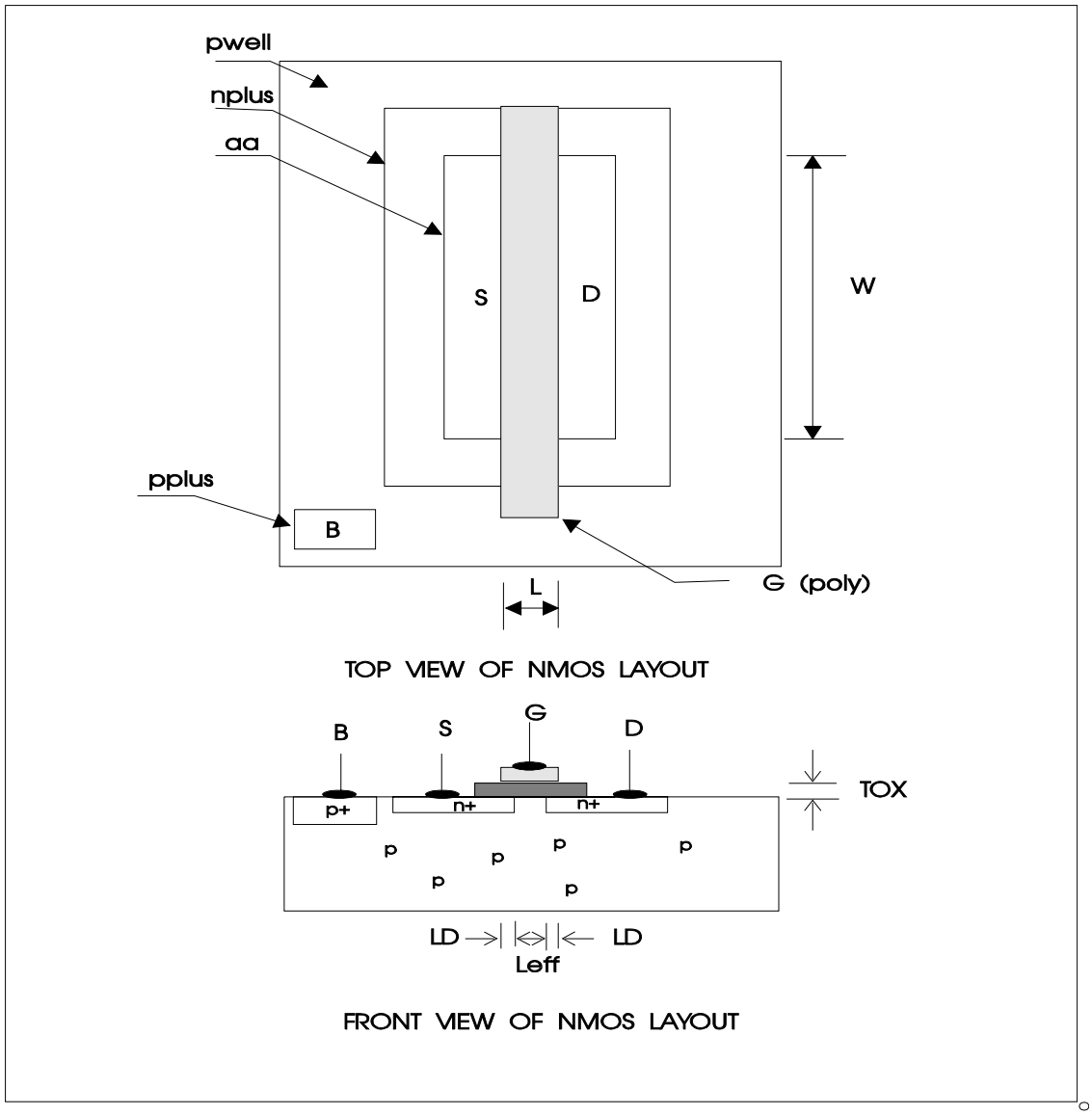

Figure 3. Nmos layout

Figure 3 shows the NMOS layout. The S (source) and D (drain) are indistinguishable. What determine the S and D is how they are connected to the external supply. D is connected a higher potential than that of the S. Normally D is connected to VDD, and S is connected to VSS. The B (bulk) is the material underneath the gate, in the NMOS case, the P substrate material. To provide a good ohmic connection to B, a pplus doping is introduced in the pwell, where connection is to be applied.

Interconnection Between Layers are provided by the following contacts:

NTAP	Metal1 to Nwell low ohmic contact (nplus and aa on nwell)
PTAP	Metal1 to Pwell low ohmic contact (pplus and aa on pwell)
M1_P	Metal1 to Pplus contact
M1_N	Metal1 to Nplus contact
M1_POLY	Metal1 to Poly 1 contact
M1_ELEC	Metal1 to Poly 2 (or Electrode) contact
M2_M1	Metal2 to Metal1 contact
M3_M2	Metal3 to Metal 2 contact

Standard CMOS Inverter Layout Steps:

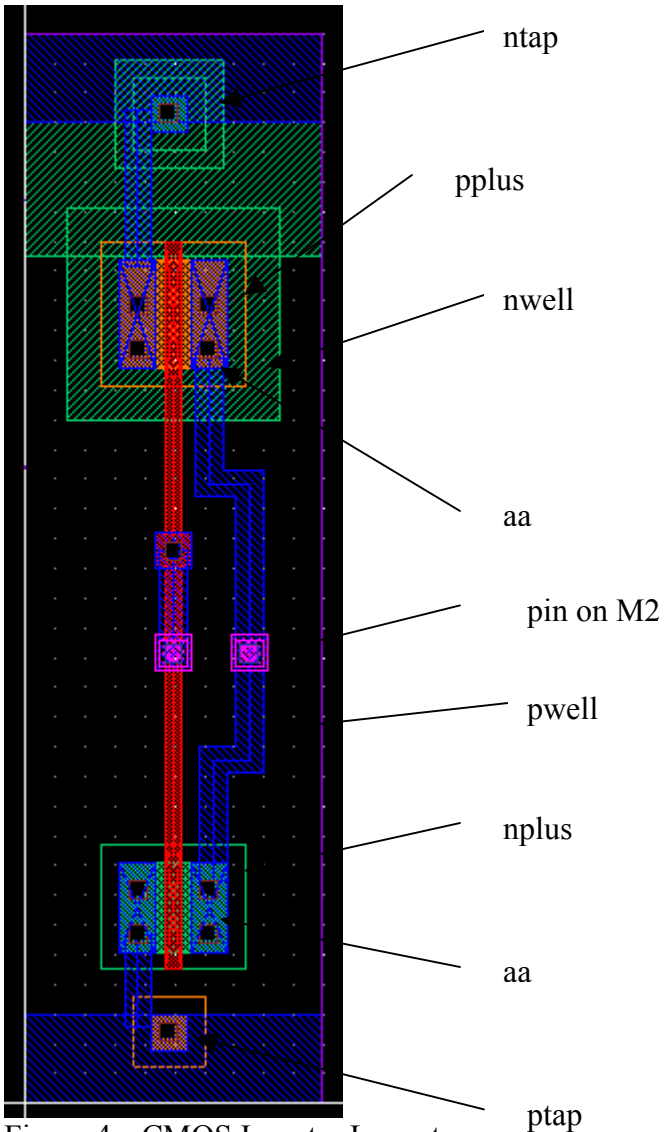

Figure 4a. CMOS Inverter Layout

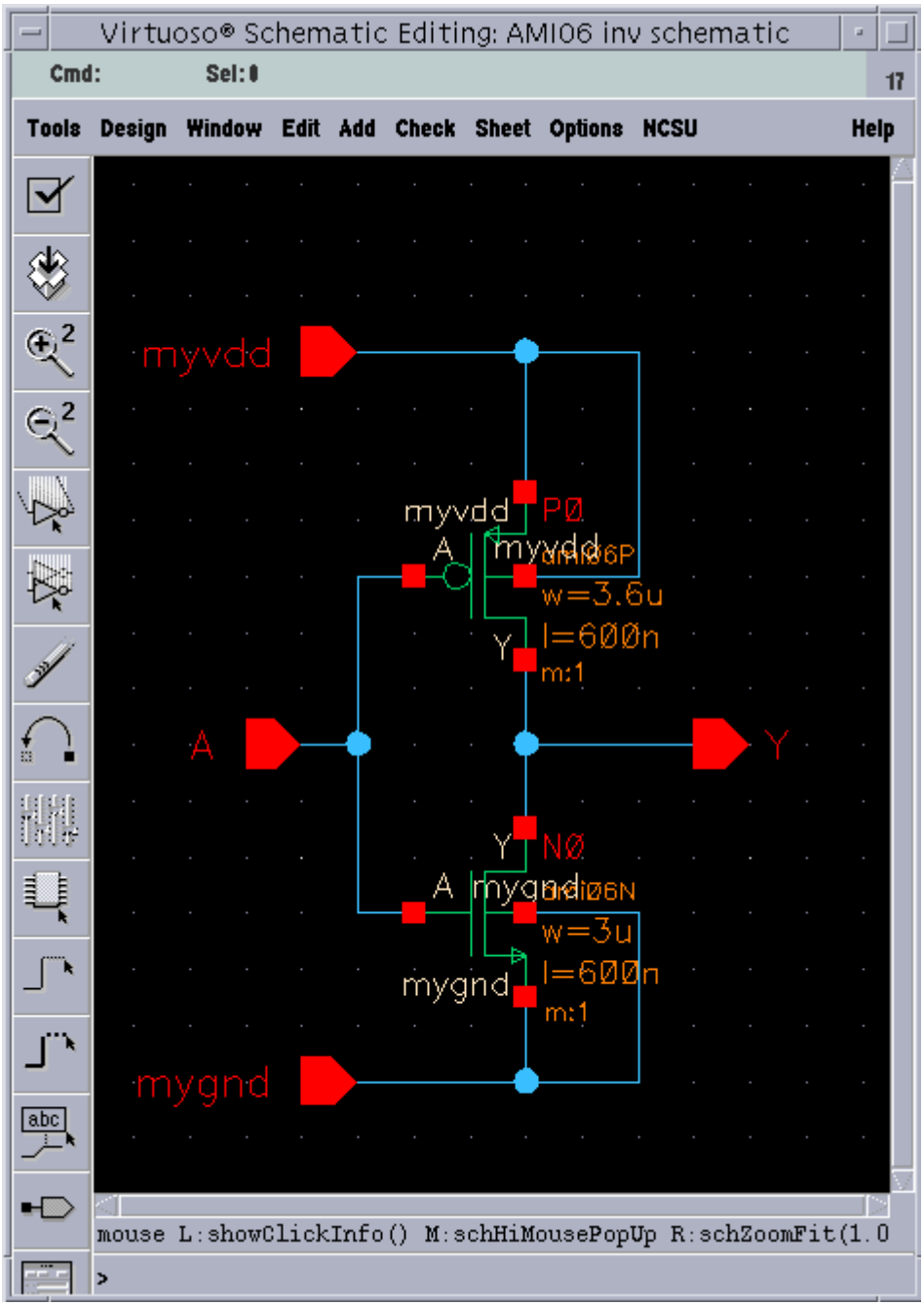

Figure 4b. CMOS Inverter Schematic.

Standard cell is designed so that each cell has a standard height. The requirements for automatic layout is that when two standard cells abut the VDD and VSS power busses must also abut.

NMOS transistor

1. Add pwell shape of appropriate size; not drawn in nwell process the entire p-substrate is considered the pwell.
2. Add nplus shapes within pwell; this defines the n-channel transistor area.
3. Add aa(active area) within nplus, this define the effective nmos transistor area.
4. Add poly across nplus to create NMOS transistor.
5. Add contact to provide connection between nplus and metal1, poly and metal1.

PMOS transistor

1. Add nwell shape of appropriate size.
2. Add pplus shapes within nwell; this defines the p-channel transistor area.
3. Add aa(active area) within pplus, this define the effective pmos transistor area.
4. Add poly across pplus to create PMOS transistor.
5. Add contact to provide connection between pplus and metal1, poly and metal1.

NCSU (North Carolina State Univ) has provided skill functions that automatically generate:

1. Nmos4, Pmos4 transistor layout that accept the transistor size W and L as parameters.
2. Ptap has been provided to provide low ohmic contact to pwell or p-substrate. It consists of pplus diffusion on pwell or p-substrate, aa and metal1 contact.
3. Ntap has been provided to provide low ohmic contact to nwell. It consists of nplus diffusion on nwell, aa and metal1 contact.

CMOS Inverter Layout

1. Draw the power rails for standard cell by executing the command "**pr**". This is a skill function provided by WSU(Wayne State Univ) to facilitate the standard cell generation.
2. Draw the nmos4 and pmos4 transistor of specified W and L within the power rails. Draw pmos4 on top of nmos4, with the poly-gate lign-up. The two gates will be connected together.
3. Connect the two-gate poly together using polysilicon.
4. The two-gate is also to be connected to the input pin. I/O pins are on metal2 layer. To achieve this a common layer is required. Metal1 can connect to poly using M_POLY contact, and metal1 can connect to metal2 using M2_M1 contact. Metal1 is then used to connect them together.
5. Connect the two-drain together using metal1. The two-drain is also connected to the output pin by M2_M1 contact.

6. Place an ntap between nwell and vdd. This is the low ohmic contact of the bulk (nwell) to vdd.
7. Place a ptap between pwell(p-substrate) and gnd. This is the low ohmic contact of the bulk (pwell/p-substrate) to gnd.
8. Connect the source of nmos4 to gnd using metal1.
9. Connect the source of pmos4 to vdd using metal1.

DRC (design rule check) each time you made a connection or place a new component to make sure that you did not have any design rule violation.

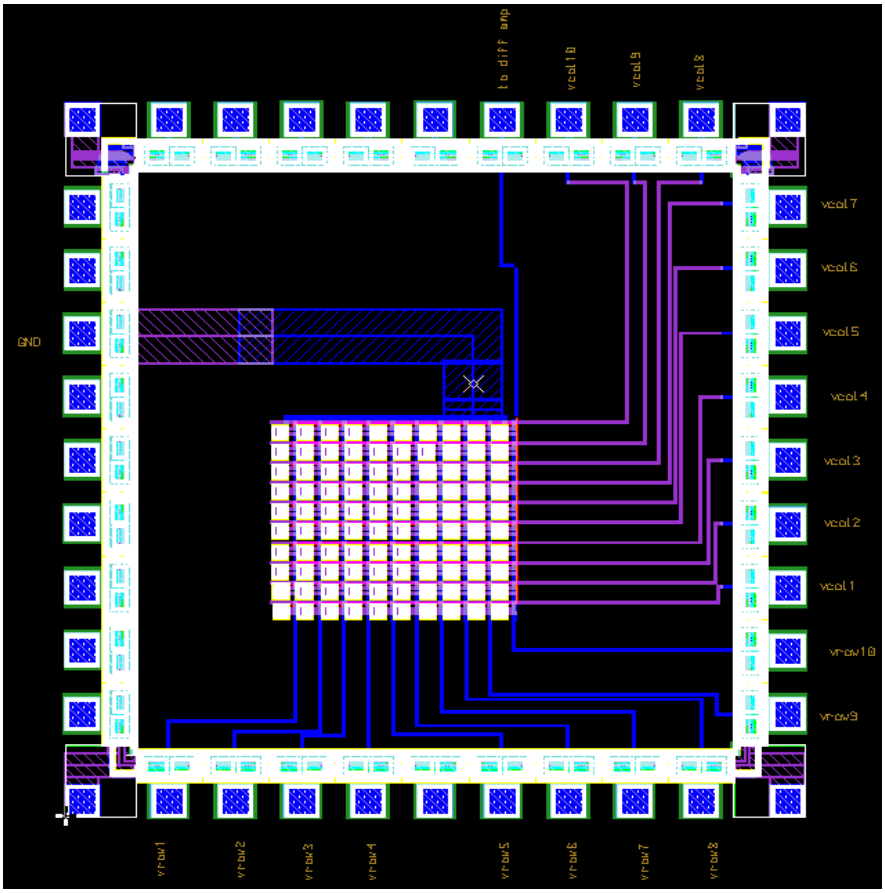

Figure 5. Complete layout with Pad Frame

Low Frequency NMOS Model

The three operating region of nmos transistor are given below:

$$I_{DS} = 0 \qquad ; V_{GS} < V_T,\ V_{DS} \geq 0 \text{ (cutoff)}$$

$$= \beta(V_{GS} - V_T - \frac{V_{DS}}{2})V_{DS} \qquad ; V_{GS} > V_T, 0 < V_{DS} < V_{GS} - V_T \text{ (ohmic)}$$

$$= (\beta/2)(V_{GS} - V_T)^2(1 + \lambda V_{DS}) ; V_{GS} > V_T, V_{DS} > V_{GS} - V_T \text{ (saturation)}$$

where:

$$\beta = K(W/L)$$

$$V_T = V_{TO} + \gamma(\sqrt{\phi - V_{BS}} - \sqrt{\phi})$$

The corresponding SPICE parameters are given below:

$K = KP \qquad A/V^2$

$V_{TO} = VTO \quad V$

$\gamma = GAMMA \ V^{1/2}$

$\phi = PHI \qquad V$

$\lambda = LAMBDA \ 1/V$

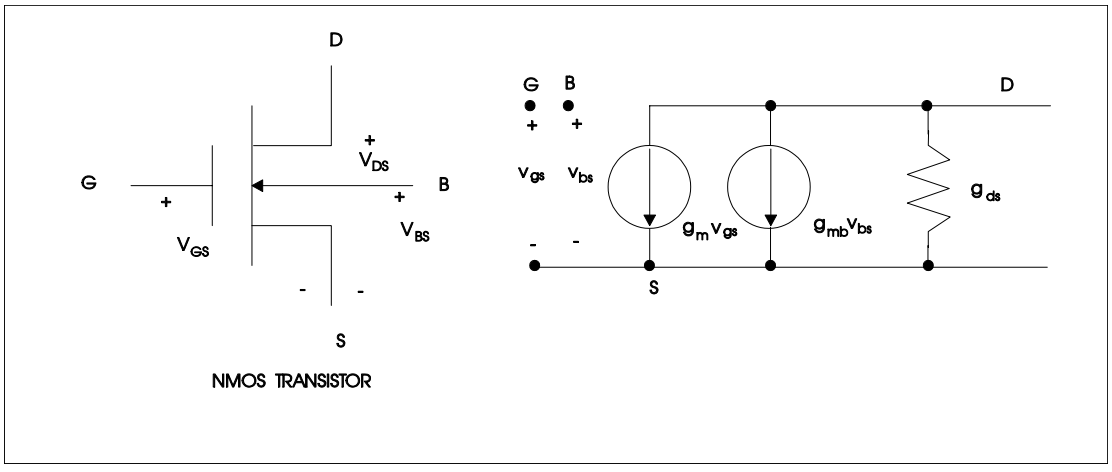

Figure 5. Low frequency small signal equivalent circuit of nmos transistor.

The low frequency model parameters are derived assumming that the transistor is biased at the saturation mode of operation.

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} = \beta(V_{GS} - V_T)(1 + \lambda V_{DS})$$
$$= \sqrt{2\beta I_{DQ}(1 + \lambda V_{DS})}$$
$$\approx \sqrt{2\beta I_{DQ}}$$

$$g_{mb} = \frac{\partial I_{DS}}{\partial V_{BS}} = (\frac{\partial I_{DS}}{\partial V_T})(\frac{\partial V_T}{\partial V_{BS}})$$
$$= -(\frac{\partial I_{DS}}{\partial V_{GS}})(\frac{\partial V_T}{\partial V_{BS}}) = -g_m(\frac{\partial V_T}{\partial V_{BS}});\ \text{since}\ (\frac{\partial I_{DS}}{\partial V_{GS}}) = -(\frac{\partial I_{DS}}{\partial V_T})$$

where:

$$\frac{\partial V_T}{\partial V_{BS}} = \frac{\partial}{\partial V_{BS}}[V_{TO} + \gamma(\sqrt{\phi - V_{BS}} - \sqrt{\phi})]$$
$$= \frac{\partial}{\partial V_{BS}}[\gamma(\sqrt{\phi - V_{BS}})] = -\frac{\gamma}{2\sqrt{\phi - V_{BS}}}$$

Therefore,

$$g_{mb} = \frac{\gamma}{2\sqrt{\phi - V_{BS}}} g_m$$

$$g_{ds} = \frac{\partial I_{DS}}{\partial V_{DS}} = \lambda(\beta/2)(V_{GS} - V_T)^2$$
$$= \frac{\lambda I_{DQ}}{1 + \lambda V_{DS}} = \lambda I_{DQ};\ \text{since}\ \lambda \approx 0$$

Implementation of transistor with large W/L ratio

It will be shown that transistor with large W/L ratio can be constructed with n transistors connected in parallel with ratio smaller by a factor of n, (W/L)/n. This principle will allow us to construct any standard cell for any desired W/L ratio. This is achieved by dividing original W/L by a suitable factor of n such that the resulting height satisfies the height constraint of standard cell.

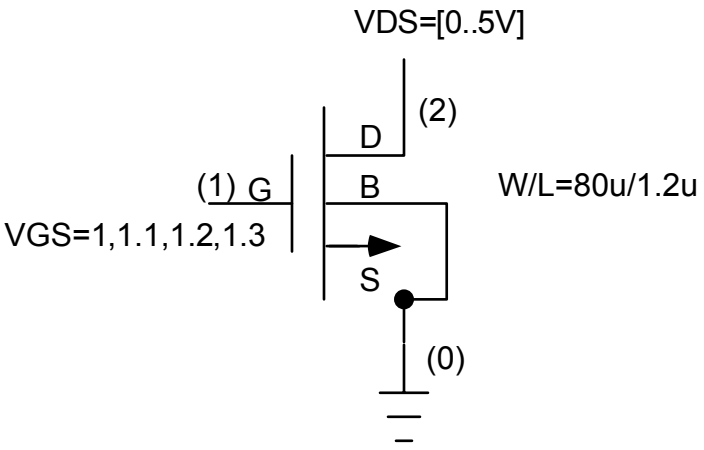

Figure 6. Schematic diagram of single nmos transistor with large W/L ratio.

The nmos transfer characteristic is simulated with Pspice. The Pspice listing is given in Listing 1 and simulation results in Figure 7. The layout of single transistor is shown in Figure 8.

Listing 1

```
* Filename="large1.cir"
* Pspice file for NMOS
* SIGNAL TRACER

VDS 2 0 DC 2.50VOLT
VGS 1 0 DC 0.0VOLT
M1 2 1 0 0 MN  W=80U L=1.2U

.MODEL MN NMOS VTO=1.0 KP=40U
+ GAMMA=1.0 LAMBDA=0.02 PHI=0.6
+ TOX=0.05U LD=0.5U CJ=5E-4 CJSW=10E-10
+ U0=550 MJ=0.5 MJSW=0.5 CGSO=0.4E-9 CGDO=0.4E-9

* Analysis
.STEP VGS 1V 1.3V 0.1V
.DC VDS 0.0 5.0 0.05
.PROBE
.END
```

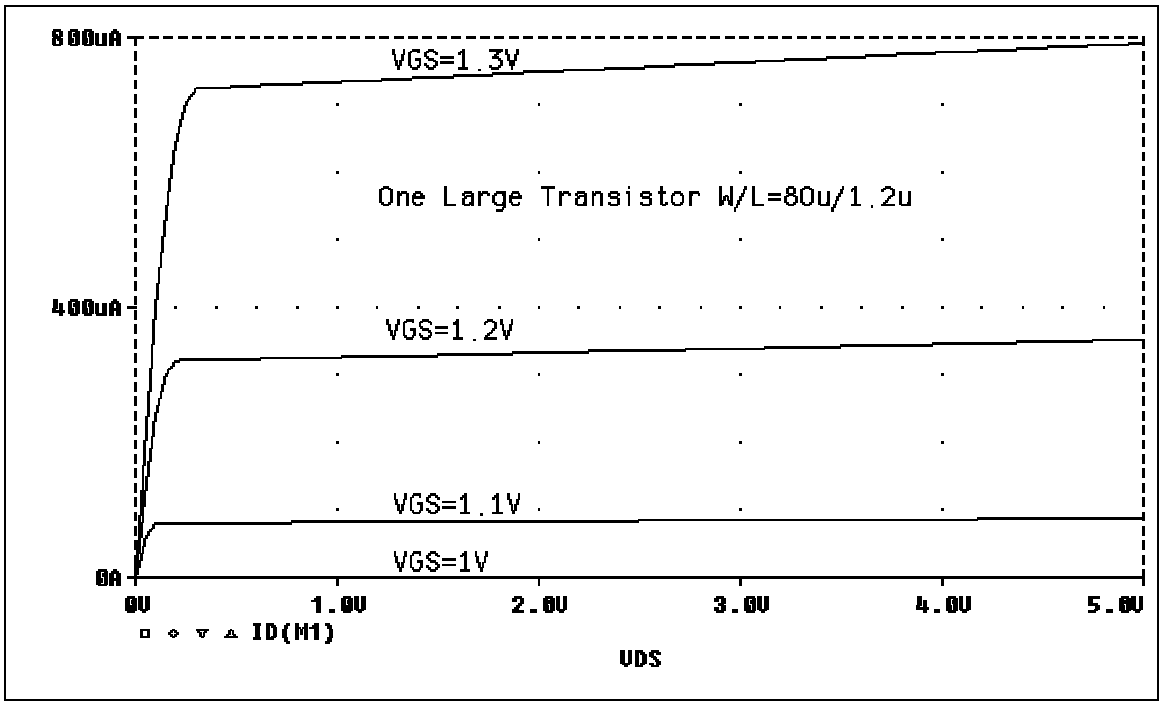

Figure 7. The transfer characteristic of nmos with W/L=80/1.2

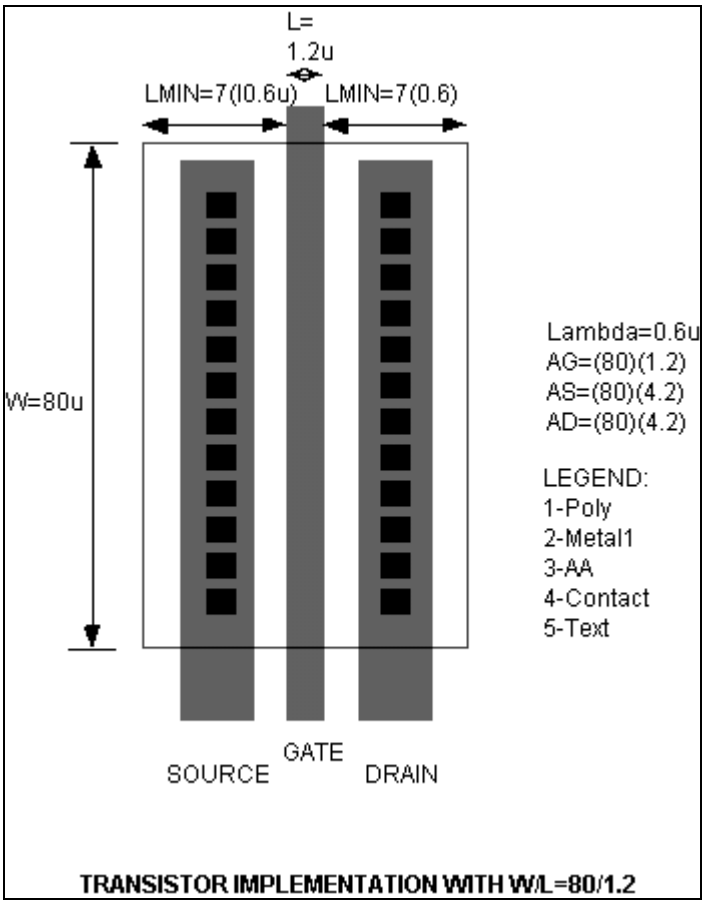

Figure 8. Layout of nmos transistor with W/L=80/1.2.

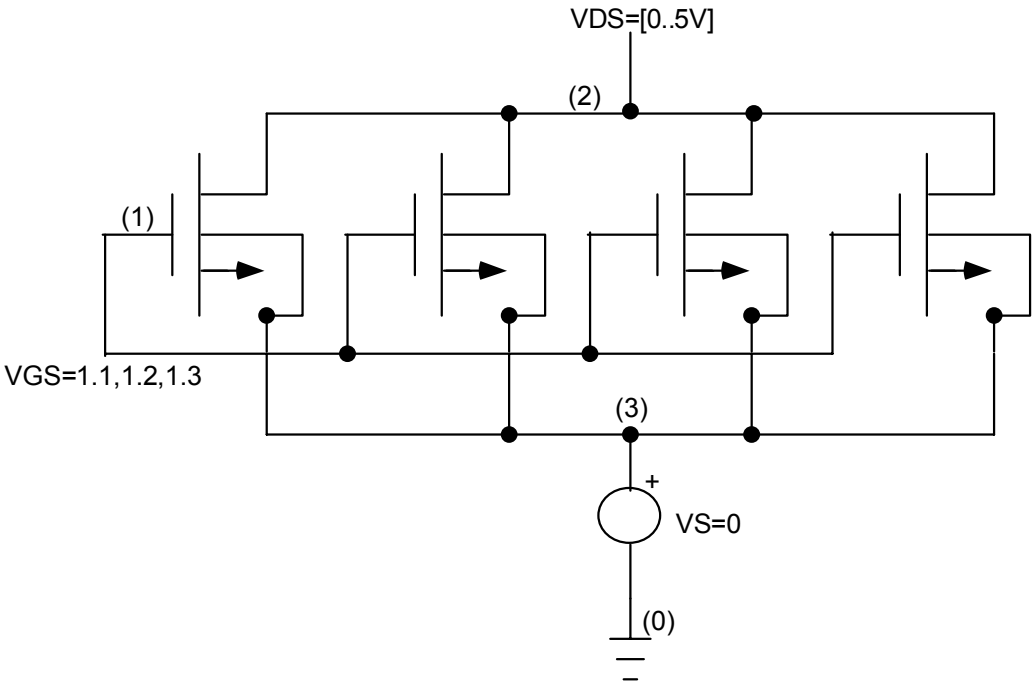

Figure 9. Schematic diagram of 4 parllel connected nmos transistors with W/L=20/1.2 each.

The four parallel connected nmos transistors transfer characteristic is simulated with Pspice. The Pspice listing is given in Listing 2 and simulation results in Figure 10. The parallel connected transistors layout is shown in Figure 11. Comparing Figure 7 and 10 shows that they have identical transfer characteristics. Hence, the two configurations are equivalent.

Listing 2

```
* Filename="largepar.cir"
* Pspice file for NMOS
* SIGNAL TRACER

VDS 2 0 DC 2.50VOLT
VGS 1 0 DC 0.0VOLT
VS  3 0 DC 0V
M1 2 1 3 3 MN  W=20U L=1.2U
M2 2 1 3 3 MN  W=20U L=1.2U
M3 2 1 3 3 MN  W=20U L=1.2U
M4 2 1 3 3 MN  W=20U L=1.2U

.MODEL MN NMOS VTO=1.0 KP=40U
+ GAMMA=1.0 LAMBDA=0.02 PHI=0.6
+ TOX=0.05U LD=0.5U CJ=5E-4 CJSW=10E-10
+ U0=550 MJ=0.5 MJSW=0.5 CGSO=0.4E-9 CGDO=0.4E-9
* Analysis
```

```
.STEP VGS 1V 1.3V 0.1V
.DC VDS 0.0 5.0 0.05
.PROBE
.END
```

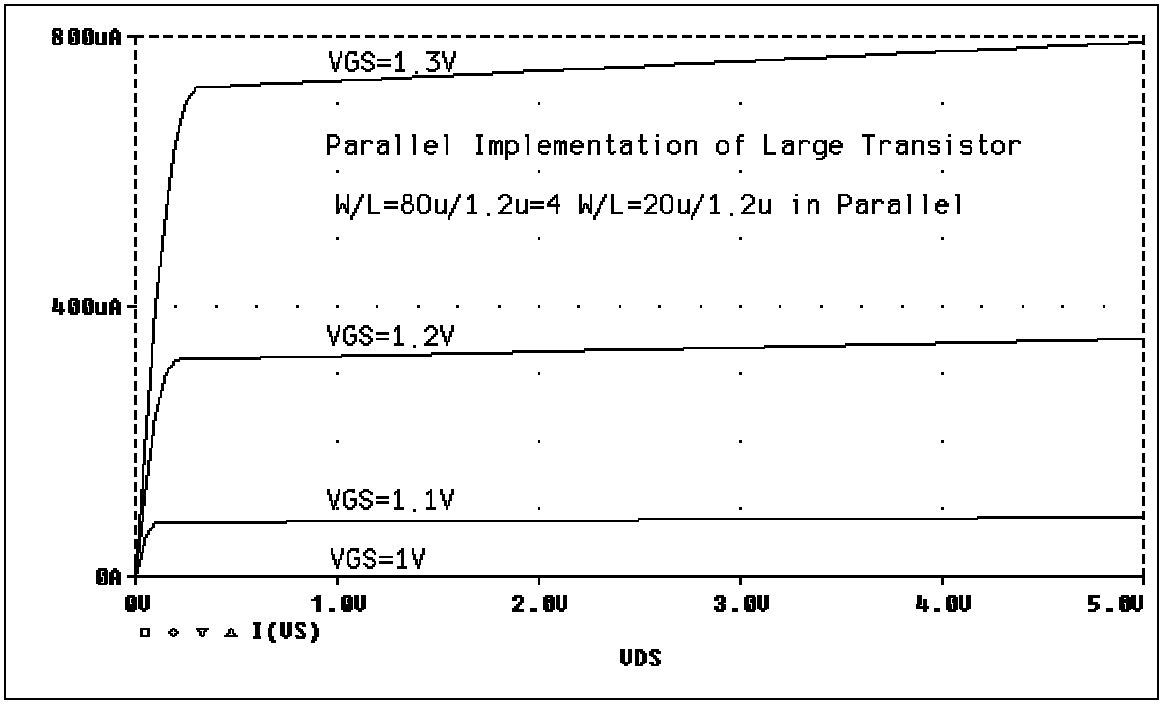

Figure 10. The transfer characteristic of 4 parallel connected transistors with W/L=20/1.2 each.

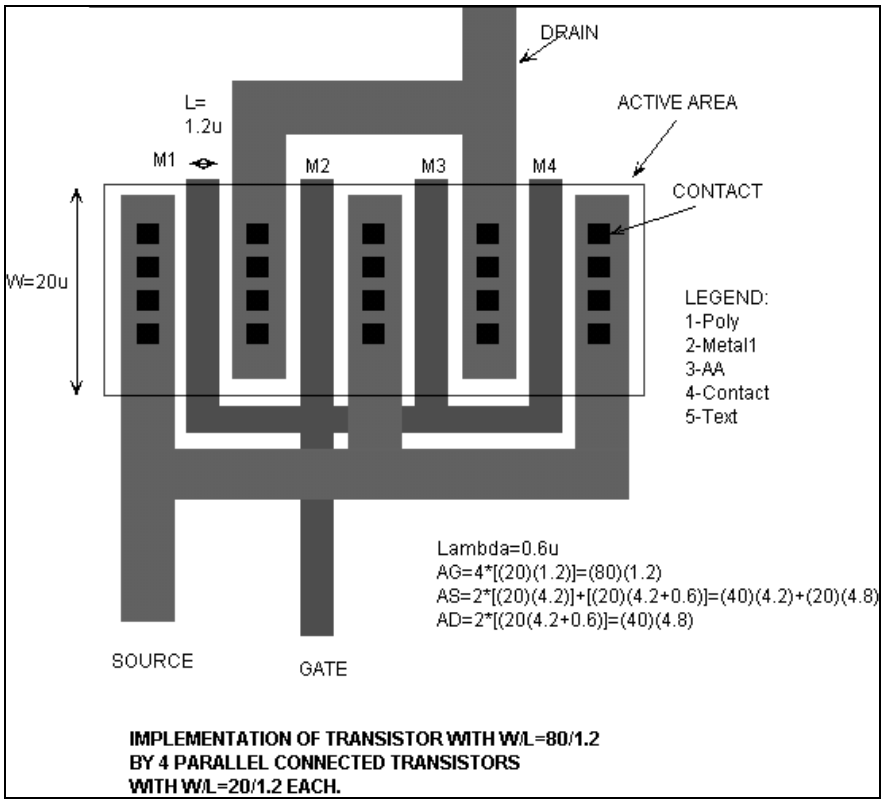

Figure 11. Layout of 4 nmos transistors connected in parallel.

PSPICE

Analysis

Analysis Requests	Pspice Command
Operating point	.OP
DC sweep	.DC source_name start_value stop_value step_value
AC frequency response	.AC DEC poinrs_per_decade freq_start freq_stop .AC OCT points_per_octave freq_start freq_stop .AC LIN total_points freq_start freq_stop
Transient response	.TRAN time_step time_stop [no_print_time max_step_size] [UIC] .IC V(node1)=value V(node2)=value …..

Pspice Output Variables:

V(node) - voltage at any node
V(node1,node2) - the voltage difference between two nodes
I(Vname) - current through a voltage source

AC output variables can also specify :

Vr, Ir : real part
Vi, Ii : imaginary part
Vm, Im : magnitude
Vp, Ip : phase
Vdb, Idb : decibels
20*log(vm(node2)/vm(node1)) : voltage gain in decibel between node2 and node1

Scale-factor Suffix

Power of ten Suffix	Metrix Prefix	Multiplying Factor
T	Tera	10^{+12}
G	Giga	10+9
Meg	Mega	10+6
K	Kilo	10+3
M	Milli	10-3
U	Micro	10-6
N	Nano	10-9
P	Pico	10-12
F	Femto	10-15

Element dimension suffix

Spice Suffix	Units
V	Volts
A	Amps
Hz	Hertz
Ohm	Ohm
H	Henry
F	Farad
Degree	Degree

Basic Element Type

First Letter Representation	Element
B	MESFET
C	Capacitor
D	Diode
E	Voltage-controlled voltage source (VCVS)
F	Current-controlled current source (CCCS)
G	Voltage-controlled current source (VCCS)
H	Current-controlled voltage source (CCVS)
I	Independent current source
J	JFET
K	Coupled inductors
L	Inductor
M	MOSFET
Q	Bipolar transistor (BJT)
R	Resistor
V	Independent voltage source

Variables Generated by Pspice
V(node)
V(node1, node2)
V(element_name)
Vx(trans_name)
I(element_name)
Ix(trans_name)

where: x can be any one of the following transistor terminals:
BJT (Q): C(collector) B (base) E (emitter) S (substrate)

FET (B,J,M): D(drain) G (gate) S (source) B (bulk, substrate)

MOSFET Model Description

Mname drain gate source bulk NMOS_model_name L=value W=value
.MODEL NMOS_model_name NMOS (parameter_name=value …)

Mname drain gate source bulk PMOS_model_name L=value W=value
.MODEL PMOS_model_name PMOS (parameter_name=value …)

Independent Source Representation

Independent Source Description	Type of Analysis
Vname n+ n- DC value Iname n+ n- DC value	All Types
Vname n+ n- AC magnitude phase_degrees Iname n+ n- AC magnitude phase_degrees	AC Frequency Response
Vname n+ n- SIN(Vo Va freq td damp) Iname n+ n- SIN(Io Ia freq td damp)	Transient $V = V_o + V_a e^{-damp(t-t_d)} \sin[(2\pi freq(t-t_d)] \quad t \geq t_d$
Vname n+ n- PULSE(V1 V2 td tr tf PW T) Iname n+ n- PULSE(I1 I2 td tr tf PW T)	Transient
Vname n+ n- PWL (t1,v1 t2,v2 … tn,vn) Iname n+ n- PWL (t1,i1 t2,i2 … tn,in)	Transient

Filename=”spicedsk.doc”

Downloadwing procedures for pspice_dos / pspice_win

In your home directory issue the following commands:

```
1   ftp> open emitsun1              -- login to emitsun1 via ftp
2   ftp> cd ../ee595                -- go to ee595 directory
3   ftp> lcd a:\                    -- go to pc local a: directory
4   ftp> bin                        -- set to bin transfer
5   ftp> prompt                     -- to turn off interactive prompt
6   ftp> ls                         -- the following two directories must be listed.
        pspice_dos
        pspice_win
7   ftp> cd pspice_dos              -- go to PSPICE DOS version – need 2 disks
8   ftp> ls
        disk1                       -- must show two directories
        disk2
9   ftp> cd disk1                   -- go to disk1 directory
```

put a blank formatted floppy in you’re a drive

10 ftp> mget *

This will copy disk1 to your floppy. Repeat 9 and 10 for the remaining disks.

To copy the window version change step 7 and 8 as follows

7 ftp>cd pspice_win -- go to PSPICE WINDOW version – need 4 disks
8 ftp> ls
disk1 -- must show four directories
disk2
disk3
disk4