Layout, Short-Channel Effects and Device Models

Tuesday, April 12th, 9:15 – 11:00

Snorre Aunet (sa@ifi.uio.no)
Nanoelectronics group
Department of Informatics
University of Oslo

April 12th:

More from 18.2 analog layout techniques,
18.3 substrate coupling
18.4 packaging
16.1 Scaling theory
16.2 Short channel effects
16.3 MOS device m
16.4 process corner
16.5 analog in a digital world (*sic!*)

Constant field scaling

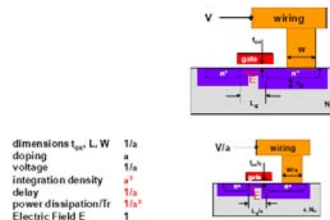


Fig.2 – The constant field scaling theory predicts an increased speed and a lower power consumption of digital MOS circuits when the critical dimensions are scaled down.

Layout of passive devices

The circuit of Fig. 18.35(a) is designed for a nominal gain of $C_1/C_2 = 8$. How should C_1 and C_2 be laid out to ensure precise definition of the gain?

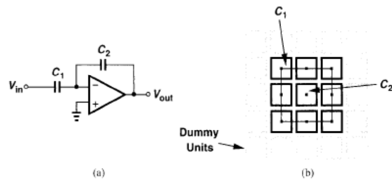


Figure 18.35

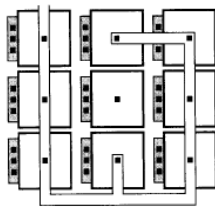


Figure 18.36 Layout of capacitors along with interconnections.

- C_2 is symmetrically surrounded by the units comprising C_1 so that vertical or horiozontal gradients are cancelled to the 1st order.
- Dummy capacitors are placed around (not seen here) , creating approximately the same environment for the units of C_1 as that seen by C_2.
- The wiring seen in Fig. 18.36 inevitably leads to some error in the C_1/C_2 ratio.

3

Capacitive coupling and reduction through differential signalling

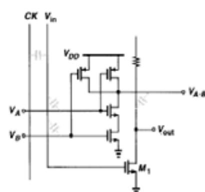


Figure 18.38 Capacitive coupling between various lines in a typical layout.

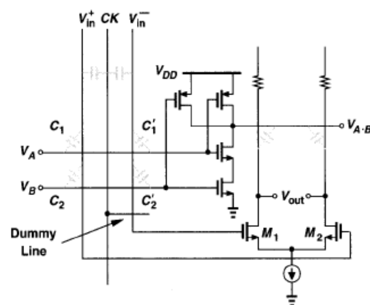


Figure 18.39 Reduction of capacitive coupling through the use of differential signaling.

- 65 nm CMOS : 9 layers of interconnect
- Capacitance may degrade speed
- Coupling of signals through capacitance between signal lines.
- Fig. 18.38: example of cross-talk between signals
- Fig. 18.39: Cancellation through differential signaling. Shielding is another technique..

4

Shielding – 2 techniques.

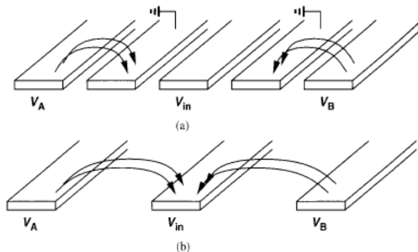


Figure 18.40 (a) Shielding sensitive signals by additional ground lines, (b) greater spacing between lines to reduce coupling.

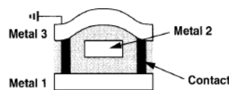


Figure 18.41 Shielding a sensitive line (metal 2) by lower and upper ground planes.

- Most of the electric field lines terminate on ground rather than the signal lines (Fig. 18.40)
- This is more effective than simply allowing more space between the signal and noisy lines.
- Fig. 18.41: The sensitive line is surrounded by a grounded shield consisting of a higher and lower metal layer and hence isolated from electric fields.

5

Wiring may add substantial thermal noise

- Vt = SQRT(4kTBR) = 0.49 nV /Hz

Example 18.5

In the layout of Fig. 18.42, a 100-μm metal 4 line is connected to a sequence of vias and contacts to reach the gate of a transistor. Calculate the thermal noise contributed by the line and the contacts.

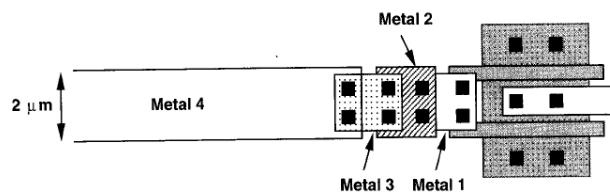


Figure 18.42

Solution

Assuming $R_\square = 40\ \text{m}\Omega/\square$ for metal 4, a via resistance of 5 Ω, and a contact resistance of 10 Ω, we have $R_{tot} = 2+2.5+2.5+2.5+5 = 14.5\ \Omega$. The thermal noise voltage is thus equal to $0.49\ \text{nV}/\sqrt{\text{Hz}}$ at room temperature.

6

Interconnects may introduce significan delay and "dispersion"

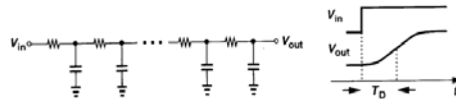


Figure 18.43 Delay and dispersion of a signal in a long line.

The distributed resistance and capacitance of long interconnects may introduce significant delay and "dispersion" in signals. Illustrated in Fig. 18.43, the delay can be approximated as

$$T_D = \frac{1}{2} R_u C_u L^2, \quad (18.15)$$

where R_u and C_u denote the resistance and capacitance per unit length, respectively, and L is the total length. For example, consider the circuit shown in Fig. 18.44, where an array

- Dispersion: increase in transition time as the signal propagates through a line (especially troublesome if it's a sampling clock signal).
- The design of power and ground busses on a chip require atention to a number of issues:
- The DC or transient voltage drop may be significant, which may affect sensitive circuits supplied by the same lines.
- Electromigration
- Two or more layers may be used in parallell, reducing the series resistance, alleviating electromigration constraints.

Pads

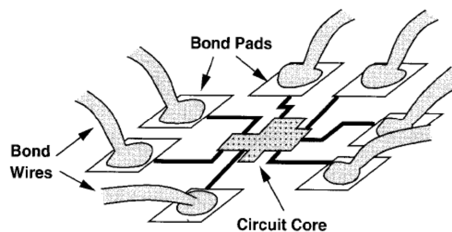
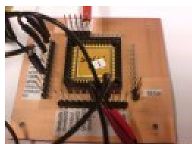


Figure 18.46 Addition of bonding pads to a chip.



- Large pads to attach bond wires
- Dimensions dictated by reliability issues.
- Bond wire diameters; 25um – 50 um
- Minimum pad size; 70um x 70um – 100um x 100 um.
- From the circuit point of view the pads should be small, to reduce area and capacitance

ESD (Electro Static Discharge) protection

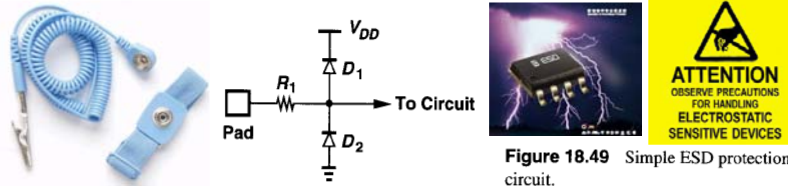


Figure 18.49 Simple ESD protection circuit.

- ESD may produce a large voltage that may damage devices on chip.
- A person touching a chip may damage it.
- ESD may occur via the air and without actual contact.
- If not properly grounded, various objects in a typical chip assembly line accumulate charge, rising to high potential levels.
- 1) MOS potential permanent damage due to ESD: gate oxide breakdown, when the field exceeds 10^7 V/cm (10 V for oxide thickness of 100 Å), typically leading to low resistance between the gate and the channel.
- 2) S/D junction diodes may melt if they carry a large current in forward or reverse bias, creating a short to the bulk.
- Both phenomena may occur for today's short-channel devices
- Fig. 18.49: Clamp external discharge to Vdd or gnd. Potential problems: degrading speed and impedance matching. Adding thermal noise via R1. Noise may be coupled through parasitics to the input. Latchup, if not properly designed.

Substrate noise due to low resistive paths between devices

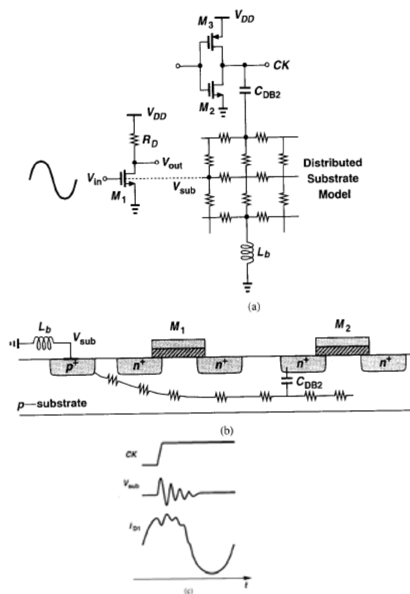


Figure 18.50 (a) Mixed-signal circuit including the effect of substrate coupling, (b) sideview of device layout, (c) signal waveforms.

- Coupling through body effect of M_1.
- Thousands of circuits might inject noise into the substrate in mixed-signal circuits, especially during clock transitions.
- Remedies:
- Increase spacing (helps)
- Differential operation
- Distribute digital signals and clocks in complementary form
- Critical operations should be performed well after clock transitions (sampling, charge transfer)
- Minimize inductance of the bond wire
- Guard rings
- Separate analog and digital ground

Guard rings to protect sensitive circuits

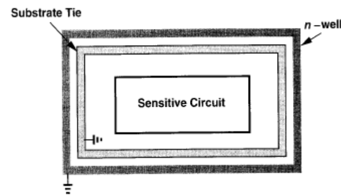


Figure 18.52 Use of guard ring to protect sensitive circuits.

- In circuits fabricated in lightly doped substrates ”guard rings” may be used to isolate sensitive sections from the substrate noise produced by other sections. (continous ring of substrate ties providing a low impedance path to ground for charge carriers produced in the substrate).

11

Tying substrate and ground together to reduce noise to transistors

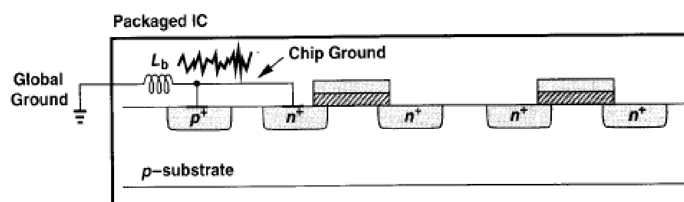


Figure 18.53 Substrate bounce.

- Since in large mixed-signal ICs it may be impossible to avoid subsrate ”bounce” with respect to the external ground, due to high transient currents drawn by the devices and the finite impedance of the bond wire connected to the substrate.
- Fig. 18.53: ground and substrate connected on the chip and brought out through a single wire.

12

Separate analog and digital ground to avoid corrupting the analog section

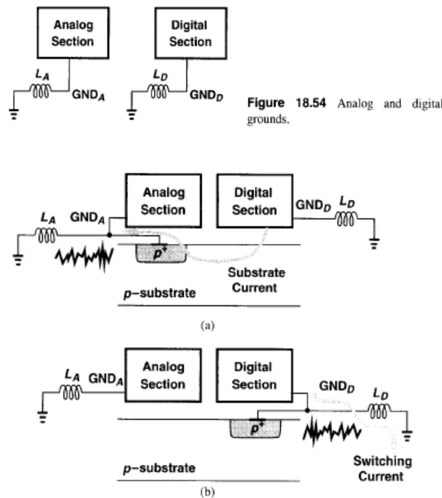


Figure 18.54 Analog and digital grounds.

Figure 18.55 Connection of substrate contact to (a) analog ground, (b) digital ground.

- Choosing the solution in Fig. 18.55 a) or b) depends on the transient currents drawn by the digital section as well as the magnitudes of L_A and L_B.
- The 1st one is mostly preferred because it ensures the analog ground and substrate potential vary in unison.

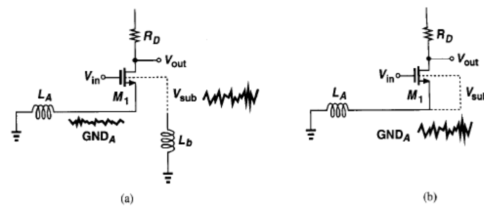
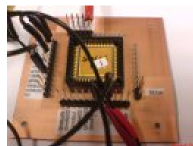


Figure 18.56 (a) Large source-bulk noise voltage due to separating substrate contact from analog ground, (b) suppression of the effect.



Electrical model of the package

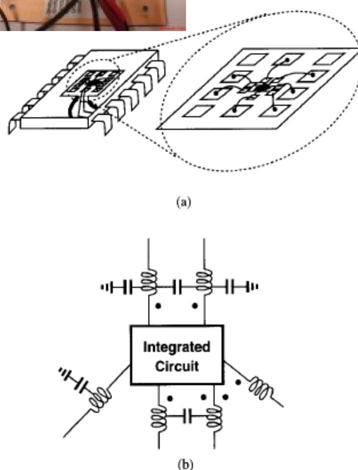
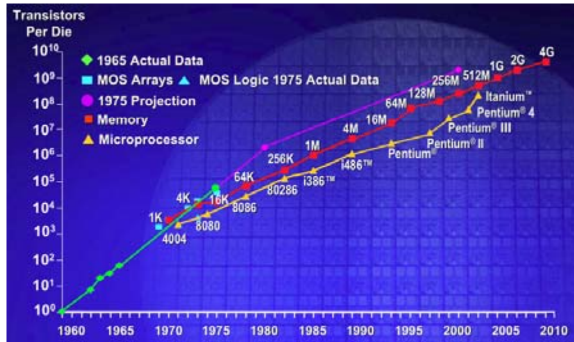


Figure 18.58 (a) Dual-in-line package, (b) electrical model of the package.

Parasitics in package and connections to the chip.
Bond-wire self inductance.
Trace self-inductance.
Trace to ground capacitance.
Trace-to-trace mutual inductance.
Trace-to-trace capacitance.
While device scaling and circuit speed have been going on, many packages have not improved significantly, due to the unscalable nature of the packages.
Simulations should include a reasonable circuit model of the package

Doubling the functionality per area every 18-24 months is nice for digital (but not for analog)



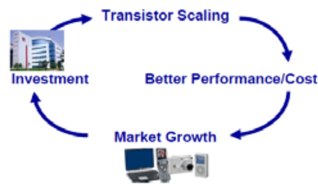


Fig. 1: The virtuous circle of the semiconductor industry

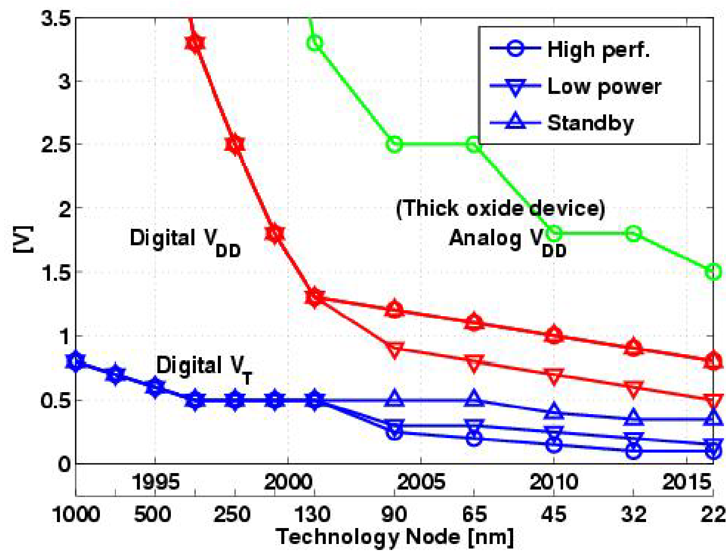
Table A Improvement Trends for ICs Enabled by Feature Scaling

TREND	EXAMPLE
Integration Level	Components/chip, Moore's Law
Cost	Cost per function
Speed	Microprocessor throughput
Power	Laptop or cell phone battery life
Compactness	Small and light-weight products
Functionality	Nonvolatile memory, imager

Supply voltages

From **Analog Circuit Design Techniques at 0.5V**

Shouri Chatterjee, Yannis Tsividis and Peter Kinget



16

Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

By Gordon E. Moore
Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

The author

Dr. Gordon E. Moore is one of the new breed of electronic engineers, schooled in the physical sciences rather than in electronics. He earned a B.S. degree in chemistry from the University of California and a Ph.D. degree in physical chemistry from the California Institute of Technology. He was one of the founders of Fairchild Semiconductor and has been director of the research and development laboratories since 1959.

Electronics, Volume 38, Number 8, April 19, 1965

The future of integrated electronics is the future of electronics itself. The advantages of integration will bring about a proliferation of electronics, pushing this science into many new areas.

Integrated circuits will lead to such wonders as home computers—or at least terminals connected to a central computer—automatic controls for automobiles, and personal portable communications equipment. The electronic wristwatch needs only a display to be feasible today.

Increasing the yield

There is no fundamental obstacle to achieving device yields of 100%. At present, packaging costs so far exceed

Heat problem

Will it be possible to remove the heat generated by tens of thousands of components in a single silicon chip?

17

Scaling and digital circuits

$$I_{D,scaled} = \frac{1}{2}\mu_n(\alpha C_{ox})\left(\frac{W/\alpha}{L/\alpha}\right)\left(\frac{V_{GS}}{\alpha} - \frac{V_{TH}}{\alpha}\right)^2 = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2\frac{1}{\alpha},$$

$$C_{ch,scaled} = \frac{W}{\alpha}\frac{L}{\alpha}(\alpha C_{ox}) = \frac{1}{\alpha}WLC_{ox}.$$

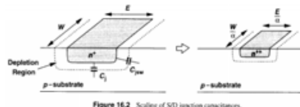


Figure 16.2 Scaling of S/D junction capacitances.

$$C_{S/D,scaled} = \frac{W}{\alpha}\frac{E}{\alpha}(\alpha C_j) + 2\left(\frac{W}{\alpha} + \frac{E}{\alpha}\right)(C_{jsw}) = [WEC_j + 2(W+E)C_{jsw}]\frac{1}{\alpha}.$$

$$T_{d,scaled} = \frac{C/\alpha}{I/\alpha}\frac{V_{DD}}{\alpha} = \left(\frac{C}{I}V_{DD}\right)\frac{1}{\alpha}.$$

- Current capability drops by $1/\alpha$ ($\alpha \approx 1/0.7$)
- Capacitances decrease by the scaling factor, $1/\alpha$
- Delay decreases by $1/\alpha$
- Number of transistors per area scales by α^2.
- The reduction of power and delay and the increase in circuit density make scaling extremely attractive for digital systems.

18

Scaling and analog circuits

$$g_{m,scaled} = \mu(\alpha C_{ox}) \frac{W/\alpha}{L/\alpha} \frac{V_{GS} - V_{TH}}{\alpha} = \mu C_{ox} \frac{W}{L}(V_{GS} - V_{TH}),$$

$$r_{O,scaled} = \frac{1}{\alpha\lambda \frac{I_D}{\alpha}} = \frac{1}{\lambda I_D}.$$

capacitance	Sqrt(kT/c)	electrons
1 fF	2 mV	25 e-
10 fF	640uV	40 e-
100 fF	200uV	125 e-
1uF	64uV	400 e-
10uF	20uV	1250 e-

- Transconductance remains constant if all of the dimensions and voltages (and currents) scale down.
- The scaling of the supply voltage has the greatest impact. If the lower end of the dynamic range is limited by thermal noise, the scaling of Vdd by 1/α reduces the dynamic range by 1/α.
- For constant thermal-noise limited dynamic range, ideal scaling of linear analog circuits require a constant power dissipation and a higher device capacitance (= αWLC_{OX})
- If the lower end of the dynamic range is determined by kT/c noise, then to maintain a constant slew rate in SC circuits, the bias current must scale up by a factor of α^2.

19

Short channel effects (L < 3 um)

- Arise due to 5 factors:
- The electric fields tend to increase because the supply voltage has not scaled proportionally.
- The built-in potential term in eq. 16.5 is neither scalable nor negligible
- The depth of the S/D junctions cannot be reduced easily
- The mobility decreases as the substrate doping increases
- The subthreshold slope is not scalable

Threshold voltage variation (Ch. 16.2.1)
Mobility degradation with Vertical field (Ch. 16.2.2)
Velocity saturation (Ch. 16.2.3)
Hot carrier effects (Ch. 16.2.4)
Output impedance variation with Drain-Source voltage (Ch. 16.2.5)

$$W_d = \sqrt{\frac{2\epsilon_{si}}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(\phi_B + V_R)}, \qquad (16.5)$$

20

Threshold Voltage Variation

Let us first consider the subthreshold behavior. For long-channel devices, the subthreshold drain current can be expressed as

$$I_D = \mu C_d \frac{W}{L} V_T^2 \left(\exp\frac{V_{GS}-V_{TH}}{\zeta V_T}\right)\left(1-\exp\frac{-V_{DS}}{V_T}\right), \quad (16.16)$$

$$S = 2.3 V_T \left(1+\frac{C_d}{C_{ox}}\right) \text{ V/dec.}$$

For example, if $C_d = 0.67 C_{ox}$, then $S = 100$ mV/dec

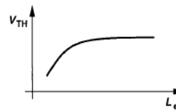
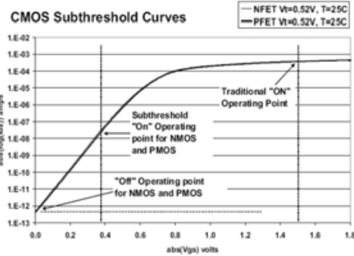
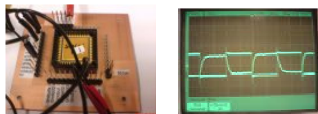


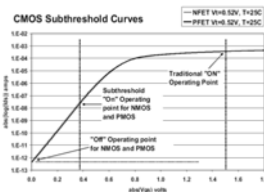
Figure 16.5 Variation of threshold with channel length.



- V_{TH} upper bound : (3/4) Vdd, lower bound dependent on subthreshold behaviour, process and temperature variations, dependency on L.
- Ex: Subthreshold slope of 80 mV/dec. Imposes a lower bound of 400 mV for V_{TH} if the off-current must be > 5 orders of magnitude.
- Temp. dependency – 1mV/°K , yielding a 50 mV change accross the commercial temp. range of 0-50°C.
- Process-induced variation in the vicinity of 50 mV raises the margin to 100 mV → difficult to reduce V_{TH} below several hundred millivolts
- Analog design; L increase for higher output impedance may increase V_{TH} by 100-200 mV

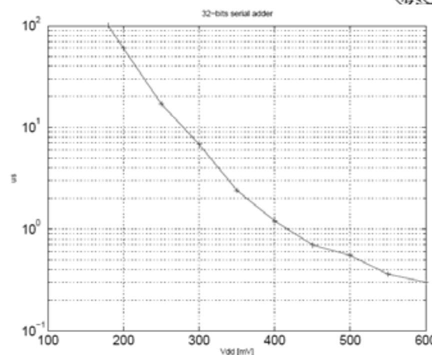
Reduced power supply voltage for the lowest current levels and reduced power consuption



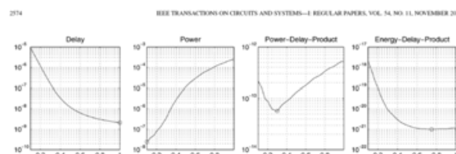


Power = Voltage x Current

Own chip prototypes working for supply voltage of 0.2 V (and below).



14. Delay, capacitively loaded 32-bit serial adder as a function of V_{dd}.



Mobility degradation with vertical field

- At large gate-source voltages, the high electric field developed between the gate and the channel confines the charge carriers to a narrower region below the oxide-silicon interface, leading to more carrier scattering and hence lower mobility.
- This lowers the current capability and transconductance of the MOSFET and deviates the I/V characteristic from the simple square-law behaviour.

Velocity saturation – potentially leading to current linearly proportional to the overdrive voltage

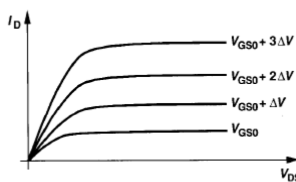
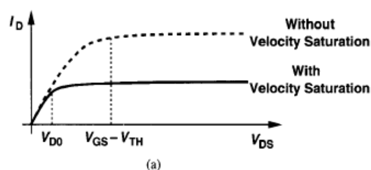


Figure 16.8 Effect of velocity saturation on drain current characteristics.

$$I_D = v_{sat} Q_d \quad (16.23)$$

$$= v_{sat} W C_{ox} (V_{GS} - V_{TH}). \quad (16.24)$$

- Carrier velocity v = μE approaches a saturated value of 10^7 cm/s, for sufficiently high fields.
- In the extreme case, eq. 2.2 may be rewritten as in equations 16.23 and 16.24.
- Fig. 16.9 a) constant current for increasing V_{DS}. Fig. 16.9b): smaller increase in $I_D(V_{GS})$ than normal.
- Short channel effect, L<1



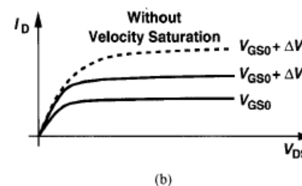


Figure 16.9 Effect of velocity saturation: (a) premature drain current saturation, (b) reduction of transconductance.

Hot carrier effects producing a finite drain-substrate current

Short-channel MOSFETs may experience high lateral electric fields if the drain-source voltage is large. While the *average* velocity of carriers saturates at high fields, the instantaneous velocity and hence the kinetic energy of the carriers continue to increase, especially as they accelerate towards the drain. These are called "hot" carriers [2].

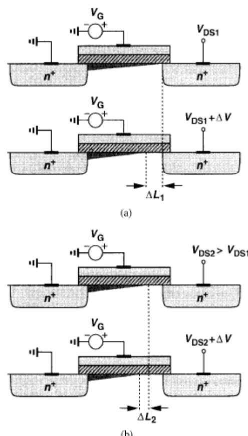
In the vicinity of the drain region, hot carriers may "hit" the silicon atoms at high speeds, thereby creating impact ionization. As a result, new electrons and holes are generated, with the electrons absorbed by the drain and the holes by the substrate. Thus, a finite drain-substrate current appears. Also, if the carriers acquire a very high energy, they may be injected into the gate oxide and even flow out the gate terminal, introducing a gate current. The substrate and gate currents are often measured to study hot carrier effects.

The scaling of technologies proceeds so as to minimize hot carrier effects. This limitation and other breakdown phenomena make the supply voltage scaling inevitable.

- Utilize similar effects in memory

25

Output impedance variation with Drain-Source Voltage



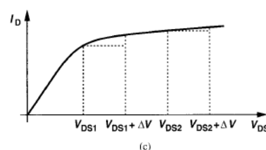


Figure 16.10 Decrement in channel length for (a) small V_{DS} and (b) large V_{DS}.

- R_o varies with V_{DS}
- As V_{DS} increases and the pinch off moves towards the source, the rate at which the depletion region around the source becomes wider decreases, resulting in a higher incremental output impedance.

26

MOS Device models

- Since the mid-1960s tremendous research has been expended on improving the accuracy of models as device dimensions scale down.

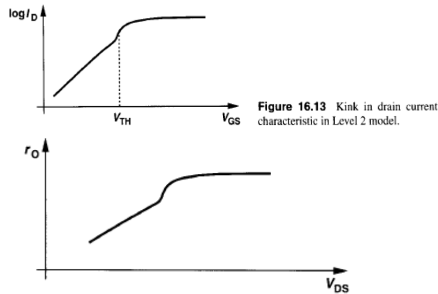


Figure 16.13 Kink in drain current characteristic in Level 2 model.

Figure 16.14 Kink in output resistance in Level 3 model.

name	special
Level 1	No subthreshold or short channel effects. Reasonably accurate for I/V down to 4µm. Poor prediction of output impedance
Level 2	Represent many higher order effects. Not good at modelling finite output impedance. Models V_{TH} variation and velocity saturation. I/V acc. To 0.7 µm
Level 3	Similar to level 2 but with many empirical constants to improve accuracy for L down to 1 µm.
BSIM series	More empirical parameters → losing touch with actual device operation. Meas. In 0.7 CMOS indicate good I/V char, but not as good for narrow, short tr.
EKV	Bulk as reference potential. Single current equation valid for subthreshold and saturation regions

27

Charge and capacitance modeling

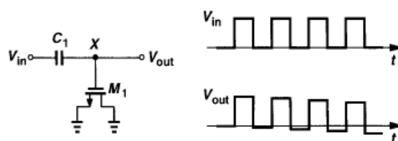


Figure 16.15 Annihilation of charge in simulation.

- Level 1 model: For example a periodic rectangular waveform applied to a voltage divider consisting of an ideal capacitor and a MOSFET experiences "droop" at the output since some charge at node X is lost every period.
- Efforts have been done to improve the charge and capacitance models, especially for analog.

28

Process corners

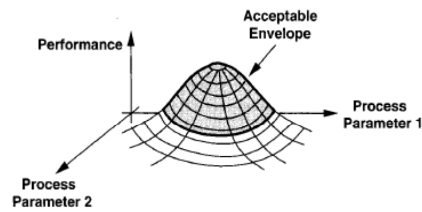


Figure 16.16 Performance envelope as a function of process parameters.

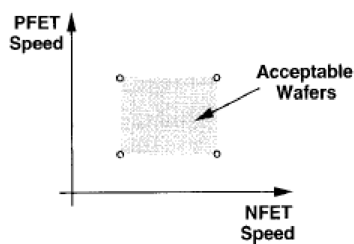


Figure 16.17 Process corners based on speed of NMOS and PMOS devices.

- Monte Carlo more and more common.

29

Analog design in a digital world

- Most CMOS technologies are designed, optimized and characterized for digital applications.
- Digital design earn from scaling, analog do not.
- Digital is automated. Analog design relies on experience, intuition, and measured data (to a much larger extent).

30

SUBSTITUTE ANALOG COMPONENTS BY DIGITAL CIRCUIT TECHNIQUES

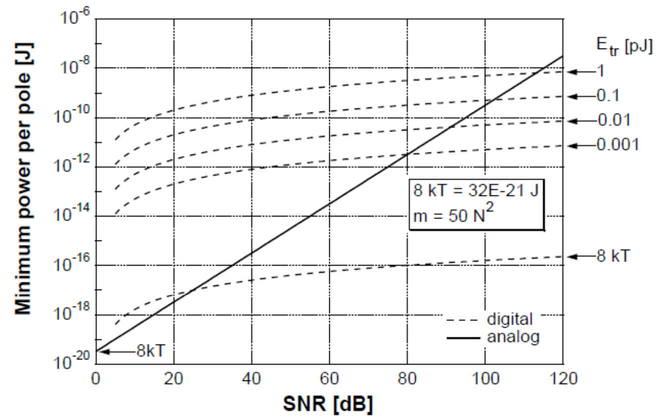


Figure 2.2 Minimum power for analog and digital circuits

CMOS Low-Power Analog Circuit Design

Christian C. Enz and Eric A. Vittoz

- No teaching at the Department next week..
- Keep an eye on the web for further plans..

32