Announcements

• Homework 1 will be assigned this week

Outline

• Module 2
  * MOS transistor I-V and C-V models

Module 2 Goals

• Models that traverse design hierarchy
• Start with transistor models
• Gate delay models
• Use models to time the design
• Modeling variability

* Based on 251A, approach
  * Increase accuracy, when needed

Device Models

• Transistor models
  * I-V characteristics
  * C-V characteristics

• Interconnect models
  * R, C, L
  * Covered in EE240A

Transistor Modeling

• Different levels:
  * Hand analysis
  * Computer-aided analysis (e.g., Matlab)
  * Switch-level simulation (some flavors of “fast Spice”)
  * Circuit simulation (Hspice)

* These levels have different requirements in complexity, accuracy and speed of computation
• We are primarily interested in delay and energy modeling, rather than current modeling
• But we have to start from the currents…
Transistor Modeling

- DC
  - Accurate I-V equations
  - Well behaved conductance for convergence (not necessarily accurate)

- Transient
  - Accurate I-V and Q-V equations
  - Accurate first derivatives for convergence
  - Conductance, as in DC

- Physical vs. empirical

from BSIM group

Goal for Today

- Develop velocity-saturated model for \( I_{on} \) and apply it to sizing and delay calculation

  - Similar approach as in 251A, just use an analytical model

Transistor I-V Modeling

- BSIM
  - Superthreshold and subthreshold models
  - Need smoothing between two regions

- EKV/PSP
  - One continuous model based on channel surface potential

MOS I-V (BSIM)

Start with the basics:

\[
I_{DS} = WC_{ox}(V_{GS} - V_{Th} - V_{C}(x)) \mu E
\]

Transistor saturates when \( V_{GD} = V_{Th} \), - the channel pinches off at drain's side.

MOS Currents (32nm CMOS with \( L >> 1 \mu m \))

Simulated 32nm Transistor
Simulation vs. Model

Major discrepancies:
- shape
- saturation points
- output resistances

Model

<table>
<thead>
<tr>
<th>Device</th>
<th>0.00E+00</th>
<th>1.00E-04</th>
<th>2.00E-04</th>
<th>3.00E-04</th>
<th>4.00E-04</th>
<th>5.00E-04</th>
<th>6.00E-04</th>
<th>7.00E-04</th>
<th>8.00E-04</th>
</tr>
</thead>
</table>

2.C Velocity Saturation

$E_c = \frac{V_{sat}}{\mu_{sat}}$

$v_{sat} = 10^5 \text{ m/s}$

Constant velocity

Constant mobility (slope = $\mu$)

$E_c = 1.5$

Modeling Velocity Saturation

- Fit the velocity-dependence curve

$x_n (\text{m/s})$

$E_c/2 = \frac{v_n}{\mu}$

NMOS: $n = 2$
PMOS: $n = 1$

Approximation $n \to \infty$

1) $v = \mu_{sat}E, E < E_c$

$I_{ds} = \mu C_{ox} \frac{W}{L} \left( V_{gs} - V_{th} \right) \left( V_{ds} - \frac{V_{ds}^2}{2} \right)$

2) $v = v_{sat}, E > E_c$

$I_{ds} = \mu C_{ox} \frac{W}{L} \left( V_{gs} - V_{th} \right) \left( V_{ds} - \frac{V_{ds}^2}{2} \right)$

$V_{dsat} = ?$

Can be reduced to Rabaey DIC model by $V_{dsat} = \text{const}$

2.D Short-Channel MOS On-Current

MOS Models

From Rabaey, 2nd ed.
Unified MOS Model

* Model presented is compact and suitable for hand analysis.
* Still have to keep in mind the main approximation: that $V_{DSat}$ is constant.
  When is it going to cause largest errors?
  * When does $E$ scale? – Transistor stacks.
* But the model still works fairly well.
  * Except for stacks

Approximation $n = 1$, piecewise

* $n = 1$ is solvable, piecewise closely approximates

$$
V = \begin{cases} 
\frac{\mu C}{1 + E/E_s} & E < E_0 = \frac{2V_{sat}}{\mu L} \\
V_{sat} & E > E_0
\end{cases}
$$

Sodini, Ko, Moll, TED’84
Toh, Ko, Meyer, IJSSC’88
BSIM model

Drain Current

* We can find the drain current by integrating

$$I_{DS} = WC_{ox}(V_{GS} - V_{th} - \frac{V_{out}^2}{2})$$

In saturation:

$$I_{DSat} = \frac{\mu C}{1 + (V_{sat}/E_s)} L (V_{GS} - V_{th} \frac{V_{out}^2}{2})$$

Drain Current in Velocity Saturation

* Solving for $V_{DSat}$

$$V_{DSat} = \frac{(V_{GS} - V_{th}) E_s L}{(V_{GS} - V_{th}) + E_s L}$$

> And saturation current

$$I_{DSat} = \frac{\mu C}{2} (V_{GS} - V_{th}) E_s L$$

Velocity Saturation

* $E_s$ is $V_{GS}$ dependent
  * Can calculate $V_{DSat}$ ($V_{in} \sim 0.4V$ in 32nm)

<table>
<thead>
<tr>
<th>$V_{GS}$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{Sat}$</td>
<td>0.05</td>
<td>0.11</td>
<td>0.19</td>
<td>0.25</td>
<td>0.33</td>
<td></td>
</tr>
</tbody>
</table>

* For $V_{GS} - V_{th} << E_s$, $V_{DSat}$ is close to $V_{GS} - V_{th}$
* For large $V_{GS}$ $V_{DSat}$ bends upwards toward $E_s$
* Therefore $E_s$ can be sometimes approximated with a constant term

Application of Models: NAND Gate

* 2-input NAND gate

Sizing for equal transistors:
* $P/N$ ratio ($l$-ratio) of 1 in < 22nm, 1.6 > 22nm
* Upsizing stacks by a factor proportional to the stack height
Transistor Stacks

- With transistor stacks, $V_{DS}$, $V_{GS}$ reduce.
- Unified model assumes $V_{DSat}$ = const.
- For a stack of two, appears that both have exactly double $R_{dev}$ of an inverter with the same width.
- Therefore, doubling the size of each should make the pull down $R$ equivalent to an inverter.

Velocity Saturation

- As $(V_{GS} - V_{Th})/ECL$ changes, the depth of saturation changes.

$$I_{DSat} = \frac{W \mu_{ef} C_{ox} ECL (V_{GS} - V_{Th})^2}{2 (V_{GS} - V_{Th}) + ECL}$$

- For $V_{DS} = 1.0V$, $ECL$ is ~0.75V.
- With double length, $ECL$ is 1.5V (in this model).
- Stacked transistors are less saturated.
- $V_{GS} = 0.6V$, $I_{DSat} = 2/3$ of inverter $I_{DSat}$ (64%).
- Therefore NAND2 should have pull-down sized 1.5X.
- Check any library NAND2’s.

Examples

Note about FinFETs

- Widths are quantized.