Announcements

• Exam 1 is on 9/29
  • It will be in WEB 236C
  • 2:20-3:45PM (10 extra minutes)
  • Closed book w/ one standard note sheet
  • 8.5”x11” front & back
  • Bring your calculator
  • Covers material through lecture 5
  • Previous years’ exam 1s are posted on the website for reference

• Reference Material
  • Razavi Chapter 18 & 19
Agenda

- MOS Fabrication Sequence
- CMOS Design Rules
- Layout Techniques
- Layout Examples
Substrate is always connected to the most negative voltage, and is shared by all N-type transistors.
MOS Fabrication Sequence

[Razavi]

(a) $p$-substrate

(b) N-Type Implant

(c) $n$-well

(d) Channel-Stop Implant

Regions where transistors are to be formed.
MOS Fabrication Sequence

[Razavi]

(c)

(p-substrate)

(f)

(p-substrate)

(g)

(p-substrate)

(h)

(p-substrate)
MOS Fabrication Sequence

[Razavi]

“Front-End”

“A silicide” step, where highly conductive metal is deposited on the gate and diffusion regions, reduces transistor terminal resistance

To prevent potential gate-source/drain shorting an “oxide spacer” is first formed before silicide deposition

“Back-End”

- A “silicide” step, where highly conductive metal is deposited on the gate and diffusion regions, reduces transistor terminal resistance
- To prevent potential gate-source/drain shorting an “oxide spacer” is first formed before silicide deposition
Contact and Metal Fabrication

[Razavi]
Transistor Geometries

- $\lambda$-based design rules allow a process and feature size-independent way of setting mask dimensions to scale
  - Due to complexity of modern processing, not used often today

- Minimum drawing feature $= \lambda$
  - Assume w.c. mask alignment $< 0.75\lambda$
  - Relative misalignment between 2 masks is $< 1.5\lambda$

\[
AGate = W \times L
\]
\[
AD, AS = W \times X
\]
\[
PS, PS = W + 2X \text{ (3 sides)}
\]

- $X$ depends on contact size
  - $5\lambda$ in this example
## Basic SCNA CMOS Layers

### Physical Layer

<table>
<thead>
<tr>
<th>Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-well</td>
</tr>
<tr>
<td>Silicon Nitride</td>
</tr>
<tr>
<td>Polysilicon Layer 1</td>
</tr>
<tr>
<td>Polysilicon Layer 2</td>
</tr>
<tr>
<td>P+ Ion Implant</td>
</tr>
<tr>
<td>N+ Ion Implant</td>
</tr>
<tr>
<td>Contact cut to n+/p_-</td>
</tr>
<tr>
<td>Metal 1</td>
</tr>
<tr>
<td>Via Oxide Cuts</td>
</tr>
<tr>
<td>Metal 2</td>
</tr>
<tr>
<td>Pad Contact (Overglass)</td>
</tr>
</tbody>
</table>

### P-channel MOSFET

- Drain
- Metal 1
- CVD Oxide
- Poly Gate
- n+/
- Gate Oxide
- p substrate
- Bulk

### N-channel MOSFET

- Source
- Metal 1
- CVD Oxide
- Poly Gate
- p substrate
- Bulk
Design Rule Basics

Minimum width and spacing

Patterning sequence for a doped n+ line.

Depletion regions due to parallel n+ lines

Contact spacing rule

Depletion regions

Mask definition

After annealing

Contact size

Side view

Geometry of a contact cut

Metal

Implanted dopants

p, Na

Contact size

Masking Design

Registration tolerance

Contact spacing rule
(a) Correct mask sizing

(b) Incorrect mask sizing

Formation of n+ regions in an n-channel MOSFET

Effect of misalignment without overhang

Gate spacing form an n+ edge

Gate overhang in MOSFET layout

(a) No overhang

(b) With misalignment

Effect of misalignment without overhang

(a) Resist pattern
(b) Isotropic etch
(c) anisotropic etch
<table>
<thead>
<tr>
<th>Mask Number</th>
<th>Mask Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NWELL</td>
</tr>
<tr>
<td>2</td>
<td>ACTIVE</td>
</tr>
<tr>
<td>3</td>
<td>POLY</td>
</tr>
<tr>
<td>4</td>
<td>SELECT</td>
</tr>
<tr>
<td>5</td>
<td>POLY CONTACT</td>
</tr>
<tr>
<td>6</td>
<td>ACTIVE CONTACT</td>
</tr>
<tr>
<td>7</td>
<td>METAL1</td>
</tr>
<tr>
<td>8</td>
<td>VIA</td>
</tr>
<tr>
<td>9</td>
<td>METAL2</td>
</tr>
<tr>
<td>10</td>
<td>PAD</td>
</tr>
<tr>
<td>11</td>
<td>POLY2</td>
</tr>
</tbody>
</table>

Difference between the drawn and physical values for channel length and the channel width
Structure of a n-channel MOSFET

Perspective view of an n-channel MOSFET

- Minimum transistor width is set by minimum diffusion width
  - 2 or 3\(\lambda\) (check with TA)
- Often, we use a slightly larger “minimum” that is equal to the contact height (4\(\lambda\) in this example)
N-channel MOSFET

(a) Cross section

(b) Circuit symbol

(c) Top view

P-channel MOSFET
Stick Diagrams

(a) Definitions

(b) MOSFET

Stick diagrams for the CMOS Inverter
The CMOS Inverter

Basic Inverter Layout

Alternate Inverter Layout
Standard Cells: VDD, VSS and output run in Parallel

CMOS NAND2 logic gate

CMOS NOR2 logic gate
Wide Analog Transistor: Analog techniques

- Unacceptable drain and source resistance
- Stray resistances in transistor structure
- Contacts short the distributed resistance of diffused areas

Most of the current will be shrunk to this side

Current is spread
Transistor orientation

• Orientation is important in analog circuits for matching purposes
Stacked Transistors

- Wide transistors need to be split
- Parallel connection of $n$ elements ($n = 4$ for this example)
- Contact space is shared among transistors
- Parasitic capacitances are reduced (important for high speed)
- Gate resistance is reduced

Note that parasitic capacitors are lesser at the drain
Matched Transistors

- Simple layouts are prone to process variations, e.g. $V_T$, $K_P$, $C_{ox}$
- Matched transistors require elaborated layout techniques
Interdigitized Layout

• Averages the process variations among transistors
• Common terminal is like a serpentine
Why Interdigitized?

- Process variations are averaged among transistors
  - KP for $M_1$: $KP_1 + KP_4 + KP_5 + KP_8$ (Avg=4.5)
  - KP for $M_2$: $KP_2 + KP_3 + KP_6 + KP_7$ (Avg=4.5)
- Technique maybe good for matching dc conditions
- Uneven total drain area between $M_1$ and $M_2$. This is undesirable for ac conditions: capacitors and other parameters may not be equal
- A more robust approach is needed (Use dummies if needed !!)
A method of achieving better capacitive matching:

Using the previous slide values, we would have $M_2$ with $KP=3.5$ and $M_1$ with $KP=5.5$

- Each transistor is split in four equal parts interleaved in two by two’s. So that for one pair of pieces of the same transistor we have currents flowing in opposite direction.

- Transistors have the same source and drain area and perimeters, but this topology is more susceptible to gradients (not common centroid)
Common Centroid Layouts

Usually routing is more complex

CENTROID
(complex layout)

M1: 8 transistors
(0,3) (0,1)
(1,2) (1,0)
(2,3) (2,1)
(3,2) (3,0)

M2: 8 transistors
(0,2) (0,0)
(1,3) (1,1)
(2,2) (2,0)
(3,3) (3,1)
Common Centroid Layouts

- Split into parallel connections of even parts
- Half of them will have the drain at the right side and half at the left
- Be careful how you route the common terminal
- Cross talk (effect of distributed capacitors ⇒ RF applications)!
• Many contacts placed close to one another reduces series resistance and make the surface of metal connection smoother than when we use only one contact; this prevents microcracks in metal;

• Splitting the transistor in a number of equal parts connected in parallel reduces the area of each transistor and so reduces further the parasitic capacitances, but accuracy might be degraded!
Integrated Resistors

- Highly resistive layers (p\textsuperscript{+}, n\textsuperscript{+}, well or polysilicon)
- \( R \) defines the resistance of a square of the layer
- Accuracy less than 30%

\[ R = \frac{\rho}{t} \ (\Omega/\text{cm}) \]

Resistivity (volumetric measure of material’s resistive characteristic)

Sheet resistance (measure of the resistance of a uniform film with arbitrary thickness \( t \))

\[ R = 2R_{\text{contact}} + \left(\frac{L}{W}\right) R \]
Diffusion resistors

Diffused resistance

Well resistance

Pinched n-well resistance
Poly Resistors

a) First polysilicon resistance

f) First polysilicon resistance with a well shielding

g) Second polysilicon resistance

h) Second polysilicon resistance with a well shielding
# TYPICAL INTEGRATED RESISTORS

\[
R = 2R_{\text{cont}} + \frac{L}{W} R_{\square}
\]

<table>
<thead>
<tr>
<th>Type of layer</th>
<th>Sheet Resistance W/0</th>
<th>Accuracy %</th>
<th>Temperature Coefficient ppm/°C</th>
<th>Voltage Coefficient ppm/V</th>
</tr>
</thead>
<tbody>
<tr>
<td>n + diff</td>
<td>30 - 50</td>
<td>20 - 40</td>
<td>200 - 1K</td>
<td>50 - 300</td>
</tr>
<tr>
<td>p + diff</td>
<td>50 - 150</td>
<td>20 - 40</td>
<td>200 - 1K</td>
<td>50 - 300</td>
</tr>
<tr>
<td>n - well</td>
<td>2K - 4K</td>
<td>15 - 30</td>
<td>5K</td>
<td>10K</td>
</tr>
<tr>
<td>p - well</td>
<td>3K - 6K</td>
<td>15 - 30</td>
<td>5K</td>
<td>10K</td>
</tr>
<tr>
<td>pinched n - well</td>
<td>6K - 10K</td>
<td>25 - 40</td>
<td>10K</td>
<td>20K</td>
</tr>
<tr>
<td>pinched p - well</td>
<td>9K - 13K</td>
<td>25 - 40</td>
<td>10K</td>
<td>20K</td>
</tr>
<tr>
<td>first poly</td>
<td>20 - 40</td>
<td>25 - 40</td>
<td>500 - 1500</td>
<td>20 - 200</td>
</tr>
<tr>
<td>second poly</td>
<td>15 - 40</td>
<td>25 - 40</td>
<td>500 - 1500</td>
<td>20 - 200</td>
</tr>
</tbody>
</table>

Special poly sheet resistance for some analog processes might be as high as 1.2 KΩ/
Large Resistors

In order to implement large resistors:

• Use long strips (large L/W)

• Use layers with high sheet resistance (bad performances – see previous table)
  • High temperature coefficient and non-linearity

Layout: rectangular “serpentine”

\[ R = \frac{L}{W} R_{\square} = \frac{L}{W} \cdot \rho \frac{1}{x_j} \]

• Estimating the resistance in the corners can be difficult
Well-Diffusion Resistor

- Example shows two long resistors for $K\Omega$ range
- Alternatively, "serpentine" shapes can be used
- Noise problems from the body
  - Substrate bias surrounding the well
  - Substrate bias between the parallel strips
Factors affecting accuracy:

Plastic packages cause a large pressure on the die (= 800 Atm.). It determines a variation of the resistivity.

For <100> material the variation is unisotropic, so the minimum is obtained if the resistance have a 45° orientation.

Temperature:

Temperature gradient on the chip may produce thermal induced mismatch.
Etching

Wet etching: isotropic (undercut effect)
HF for SiO₂; H₃PO₄ for Al
Δx for polysilicon may be 0.2 – 0.4 μm with standard deviation 0.04 – 0.08 μm.
Reactive ion etching (R.I.E.) (plasma etching associated to “bombardment”): unisotropic.
Δx for polysilicon is 0.05 μm with standard deviation 0.01 μm

Boundary:

The etching depends on the boundary conditions

• Use dummy strips
Side diffusion effect: Contribution of endings

Side Diffusion “widens” $R$
$R_{\square}$ is not constant with $W$

Impact of $R_{\text{cont}}$ depends on relative geometry
Best to always use a resistor $W$ that is at least as large as the contact

Interdigitized structure:
MOS Capacitors

- One of the most dense capacitors available (fF/um²)
- Often used to de-couple DC power supply and bias signals
- Capacitor non-linearity can be important if used in the signal path

(a) MOSFET configured as a capacitor, (b) nonlinear C/V characteristic.
Integrated Capacitors

Poly - Diffusion

Poly - Poly

Metal1 - Poly

$C = C_1 + \ldots + C_4$

Vertical Metal “Sandwich”

Lateral Metal-Oxide-Metal (MOM)

Substrate

[Razavi]

[Ho]

[Wang]
Poly1 - Poly2 Capacitor

- Area is determined by poly2
- Problems
  - undercut effects
  - nonuniform dielectric thickness
  - matching among capacitors
Factor affecting RELATIVE accuracy/matching

- Oxide damage
- Impurities
- Bias condition
- Bias history (for CVD)
- Stress
- Temperature
- Grow rate
- Poly grain size
- Etching
- Alignment

$$\left(\frac{\Delta \varepsilon_{\text{ox}}}{\varepsilon_{\text{ox}}}\right)$$

$$\left(\frac{\Delta t_{\text{ox}}}{t_{\text{ox}}}\right)$$

$$\left(\frac{\Delta L}{L}\right); \left(\frac{\Delta W}{W}\right)$$

$$\left(\frac{\Delta C}{C}\right)^2 = \left(\frac{\Delta \varepsilon_r}{\varepsilon_r}\right)^2 + \left(\frac{\Delta t_{\text{ox}}}{t_{\text{ox}}}\right)^2 + \left(\frac{\Delta L}{L}\right)^2 + \left(\frac{\Delta W}{W}\right)^2$$

Note, the absolute C may vary as high as 20% due to process variations.
Accuracy of integrated capacitors

Perimeter imperfections effect the total capacitance:

\[ C = C_A A \]

\[ A = (x-2\Delta x)(y-2\Delta y) = xy - 2x\Delta y - 2y\Delta x + 4\Delta x \Delta y \]

Assuming that \( \Delta x = \Delta y = \Delta e \)

\[ A = xy - 2\Delta e(x + y) + 4(\Delta e)^2 \]

\[ A \approx xy - 2\Delta e(x + y) \]

\[ \therefore C_e = -2\Delta e(x + y) \]

The relative error is

\[ \varepsilon = \frac{C_e}{C} = \frac{-2\Delta e(x + y)}{xy} \]

Then maximize the area and minimize the perimeter \( \rightarrow \) use squares!!!
Capacitor Matching

- If we want to match the ratio of two caps $C_1$ and $C_2$

\[
\frac{C_1}{C_2} = \frac{C_{1,\text{ideal}}(1+e_1)}{C_{2,\text{ideal}}(1+e_2)}
\]

- To minimize the error in the cap ratio, we need to have $e_1 = e_2$. This implies that the Perimeter/Area should be equal.

- Generally, we break up both caps $C_1$ and $C_2$ into square unit caps, $C_u$

\[
A_u = x_u^2
\]
Capacitor Matching

- If both $C_1$ and $C_2$ are integer multiples of $C_u$, then simply use $I_1$ unit caps for $C_1$ and $I_2$ unit caps for $C_2$

$$\frac{C_1}{C_2} = \frac{I_1C_u}{I_2C_u}$$

- If $C_1 = I_1C_u$ and $C_2 = I_2C_u + (1 + f)C_u$, where $f$ is a fraction, then we want the non-unit cap, $C_{nu} = (1 + f)C_u$, to have equal Perimeter \[\frac{\text{Area}}{\text{Area}}\] as a unit cap.

- Why is the non-unit cap $(1 + f)C_u$? Because we don't want a cap smaller than the unit cap, or bigger than $2C_u$. Overall, we want to use as many unit caps as possible, but none smaller than $C_u$. 


Capacitor Matching

• How to size $C_{nu}$?

Define $N = 1 + f = \frac{A_{nu}}{A_u} = \frac{x_{nu}y_{nu}}{x_u^2}$

For matching $\frac{P_{nu}}{A_{nu}} = \frac{P_u}{A_u}$

$\frac{P_{nu}}{P_u} = \frac{A_{nu}}{A_u} = N = \frac{2(x_{nu} + y_{nu})}{4x_u}$

$x_{nu} + y_{nu} = 2Nx_u$

From $N$ definition $\Rightarrow x_{nu} = \frac{N x_u^2}{y_{nu}}$

Plugging this into the previous expression, we can solve for $y_{nu}$

$y_{nu}^2 - 2Nx_u y_{nu} + N x_u^2 = 0$

$y_{nu} = x_u \left( N \pm \sqrt{N^2 - N} \right)$

If you want $y_{nu} > x_{nu}$ choose "+", $y_{nu} < x_{nu}$ choose "-"
Capacitor Matching

- If $C_1 = I_1 C_u + (1 + f_1)C_u$ and $C_2 = I_2 C_u + (1 + f_2)C_u$
  where $N_1 = 1 + f_1$ and $N_2 = 1 + f_2$

$$x_{nu1} = \frac{N_1 x_u^2}{y_{nu1}} \quad x_{nu2} = \frac{N_2 x_u^2}{y_{nu2}}$$

$$y_{nu1} = x_u \left( N_1 \pm \sqrt{N_1^2 - N_1} \right) \quad y_{nu2} = x_u \left( N_2 \pm \sqrt{N_2^2 - N_2} \right)$$

- Although, generally we have the flexibility to size $C_u$ to set $C_1 = I_1 C_u$
Common Centroid Capacitor Layout

- Unit capacitors are connected in parallel to form a larger capacitance.
- Typically the ratio among capacitors is what matters.
- The error in one capacitor is proportional to perimeter-area ratio.
- Use dummies for better matching (See Razavi Book, page 752).
Approximate Common Centroid Structure

\[ \begin{align*}
C_2 &= C_1 \\
C_3 &= 2C_1 \\
C_4 &= 4C_1 \\
C_5 &= 8C_1
\end{align*} \]
“Floating” Capacitors

Be aware of parasitic capacitors

**Polysilicon-Polysilicon:** Bottom plate capacitance is comparable (10-30 %) with the poly-poly capacitance

➤ **Metal1-Metal2:** More clean, but the capacitance per micrometer square is smaller. Good option for very high frequency applications (C~ 0.1-0.3 pF).

CP1, CP2” are very small (1-5 % of C1)
CP2’ is around 10-50 % of C1

CP2 is very small (1-5 % of C1)
Analog Cell Layout Key Points

- Use transistors with the same orientation

- Minimize S/D contact area by stacking transistors (to reduce parasitic capacitance to substrate)

- Respect symmetries

- Use low resistive paths when current needs to be carried (to avoid parasitic voltage drops)

- Shield critical nodes (to avoid undesired noise injection)

- Include guard rings everywhere; e.g. Substrate/well should not have regions larger than 50 um without guard protections (latchup issues)
Bipolar Transistors – Latchup

- Potential for parasitic BJTs (Vertical PNP and Lateral NPN) to form a positive feedback loop circuit
- If circuit is triggered, due to current injected into substrate, then a large current can be drawn through the circuit and cause damage
- Important to minimize substrate and well resistance with many contacts/guard rings
Stacked Layout for Analog Cells

- Stack of elements with the same width

- Transistors with even number of parts have the source (drain) on both sides of the stack

- Transistors with odd number of parts have the source on one end and the drain on the other. If matching is critical use dummies

- If different transistors share a same node they can be combined in the same stack to share the area of the same node (less parasitics)

- Use superimposed or side by side stacks to integrate the cell
• M1 and M2 must match. Layout is interdigitized
• M3 and M4 must match. M6 must be wider by 4*M3
• M7 must be 2*M5
• Layout is an interconnection of 3 stacks; 2 for NMOS and 1 for PMOS
• Capacitor made by poly-poly

Pay attention to your floor plan! It is critical for minimizing iterations: Identify the critical elements

Not the best floorplan, but OK
Following slides were provided by some of Dr. Silva’s graduate students.

Special thanks to Fabian Silva-Rivas, Venkata Gadde, Marvin Onabajo, Cho-Ying Lu, Raghavendra Kulkarni and Jusung Kim
Figure: Layout of a single stage fully differential amplifier and its CMFB circuit. 
1. I/p NMOS diff pair  2. PMOS (Interdigitated) 3. Resistors for $V_{CM}$ 4. Capacitors (Common centroid)
Figure: Layout of a second order Active RC low-pass Filter (Bi-quad)

- Resistive network
- Capacitive network (Common centroid)
- Fully differential amplifier
• 3-bit quantizer in Jazz 0.18μm CMOS technology
• S/H: sample-and-hold circuit that is used to sample the continuous-input signal
• Core: contains matched differential pairs and resistors to create accurate reference levels for the analog-to-digital conversion
• Latches: store the output bits; provide interface to digital circuitry with rail-to-rail voltage levels
- High-speed D-Flip-Flop in Jazz 0.18\textmu m CMOS technology
- Resolves a small differential input with $10\text{mV} < V_{\text{p-p}} < 150\text{mV}$ in less than 360ps
- Provides digital output (differential, rail-to-rail) clocked at 400MHz
- The sensitive input stage ($1^{st}$ differential pair) has a separate “analog” supply line to isolate it from the noise on the supply line caused by switching of digital circuitry
Design example (industrial quality): Simplest OTA
Overall amplifier: Have a look on the guard rings and additional well!
BIAS: you may be able to see the dummies, symmetry and S/D connections
From downstairs

Differential pair
Details on the P-type current mirrors
Q-value of Spiral Inductors in CMOS Process

Most of the following slides were taken from Seminar by: Park, Sang Wook

TAMU, 2003
What is $Q$?

$$Q = \frac{\omega \text{ energy stored}}{\text{average power dissipated}}$$

Simple Inductor Model: $L_S$ $R_S$

$$Q = \frac{\omega L_s}{R_s}$$

Integrated Spiral Inductor “Pi” Model
Equivalent Circuit & Calculation

Equivalent Circuit

Parameter Calculation

Q-factor vs Frequency (GHz)
Layout & Structure

- N: # of Turns
- R: Radius
- S: Space (=1.5um)
- W: Width (=14.5um)
- Metal-5 under path
- Metal-6 (thickness=2um)

Layers:
- Oxide
- Substrate
- Via-5
Layout Split 1

Shape & Radius

- **Shape**: Octagon > Square
- **Radius**: 60 > 30
Layout Split 2

PGS (Patterned Ground Shield) material

Q-factor

\[ \text{NxRxWxS} \]
\[ (4.5\times60\times15\times1.5) \]

PGS : Poly > Nwell > none > Metal1
GS (Ground Shield) type

Q-factor

\[ N \times R \times W \times S \]

(4.5x60x15x1.5)

- GS : Poly PGS > Poly(nonsal) SGS > Poly(nonsal) PGS > none
- PGS = patterned ground shield
- SGS = solid ground shield
Layout Split 4

Metal Stack

Q-factor
NxRxWxS
(4.5x60x15x1.5)

- Stack: M6 > M5/6 > M4/5/6 > M3/4/5/6
TAMU Mixed-Signal Research Chip

- A 10 Gb/s Hybrid ADC-Based Receiver w/ Embedded Analog and Per-Symbol Dynamically Enabled Digital Equalization

- 10GS/s asynchronous SAR ADC with embedded 3-tap FFE
- Digital equalizer with 4-tap FFE and 3-tap DFE
- Fabricated in GP 65nm CMOS

BER w/ 36.4dB Loss Channel

Next Time

• Table-Based \((g_m/I_D)\) Design Examples