ECEN474: (Analog) VLSI Circuit Design
Fall 2011

Lecture 6: Layout Techniques

Sebastian Hoyos

Analog & Mixed-Signal Center

Texas A&M University

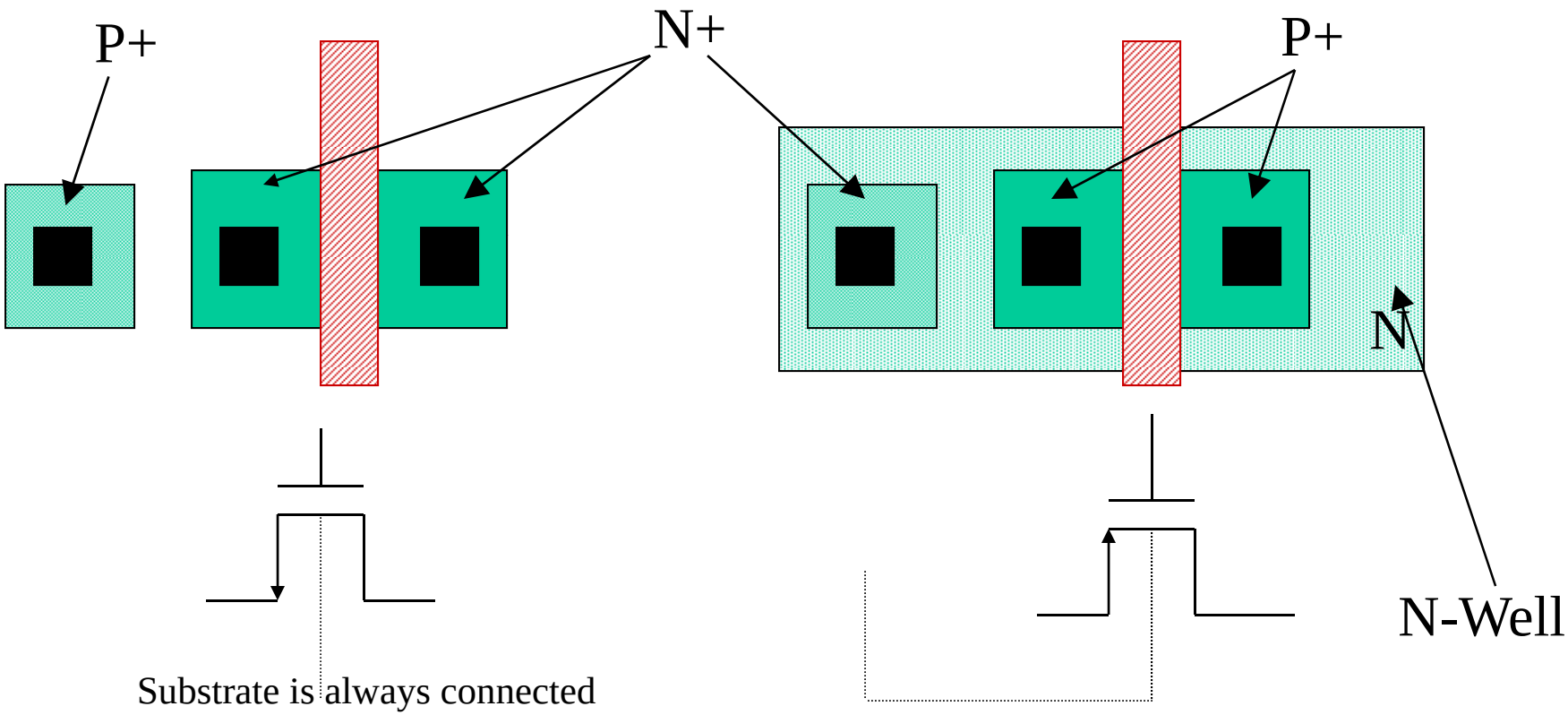
Announcements

- Lab 2 next week
  - Has a prelab
  - Upgrading Cadence version
- Reading
  - Razavi's CMOS chapter 17 & 18
- HW1 was assigned today and is due in a week

Agenda

- MOS Fabrication Sequence
- CMOS Design Rules
- Layout Techniques
- Layout Examples

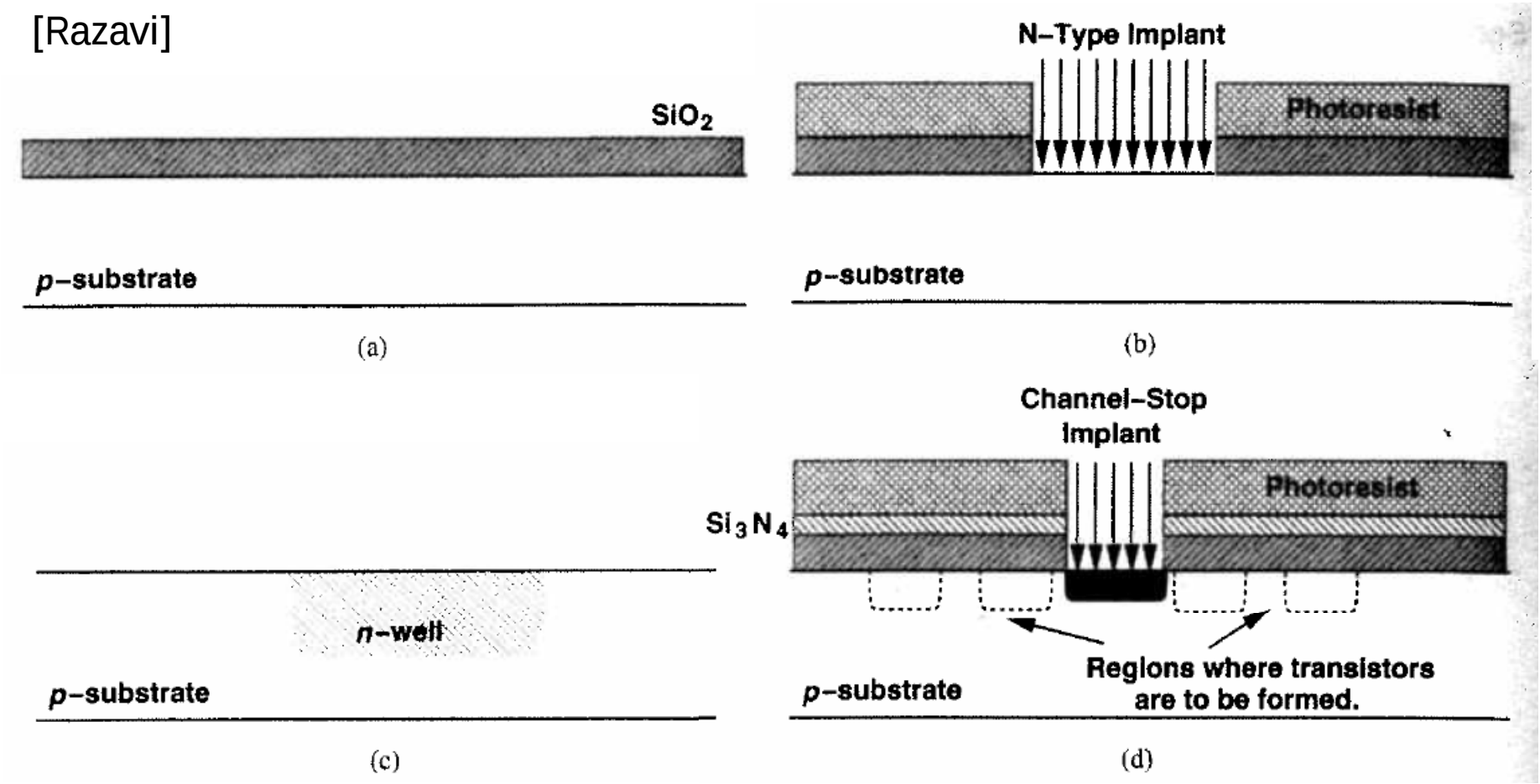
Fundamentals on Layout Techniques: N-Well CMOS Technologies



Substrate is always connected to the most negative voltage, and is shared by all N-type transistors

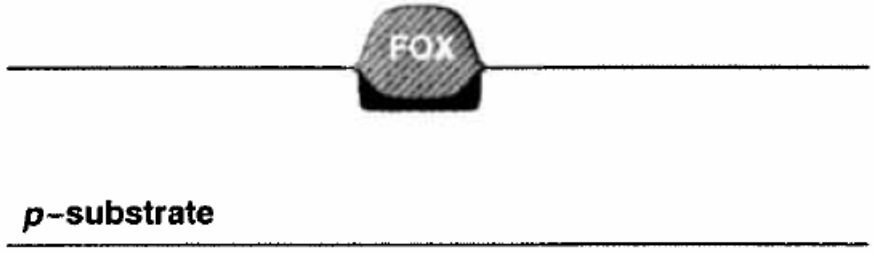
MOS Fabrication Sequence

[Razavi]

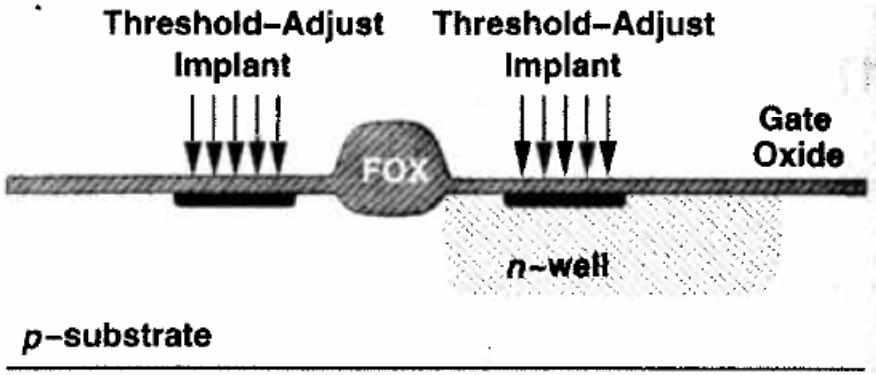


MOS Fabrication Sequence

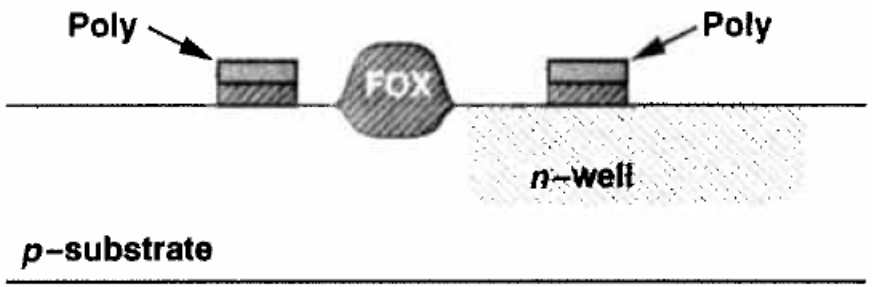
[Razavi]



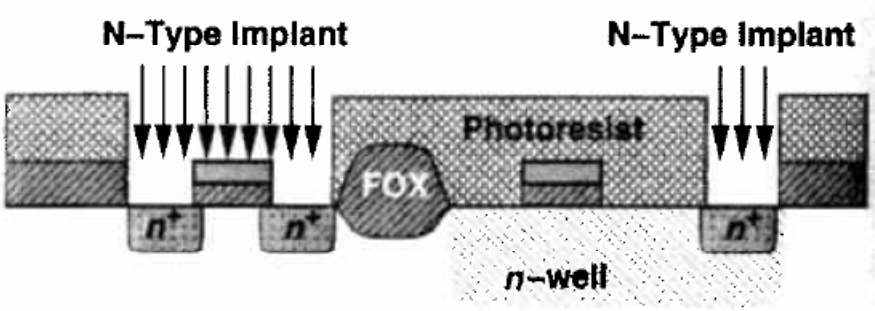
(e)



(f)



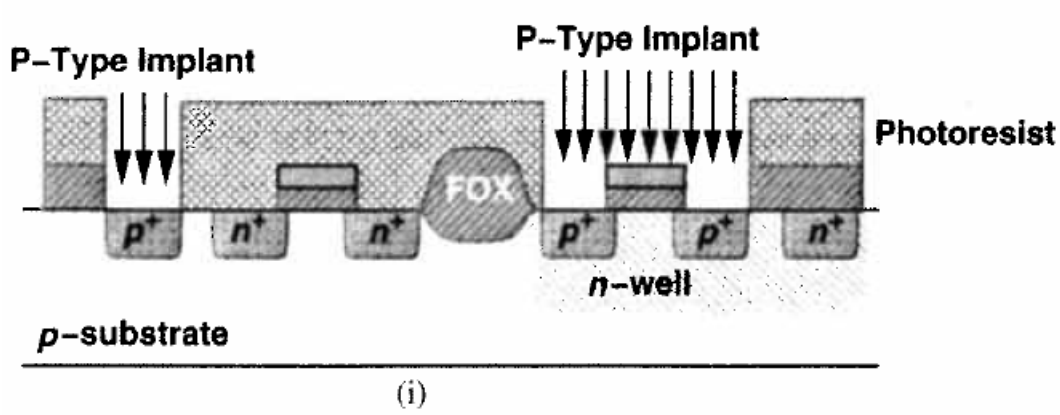
(g)



(h)

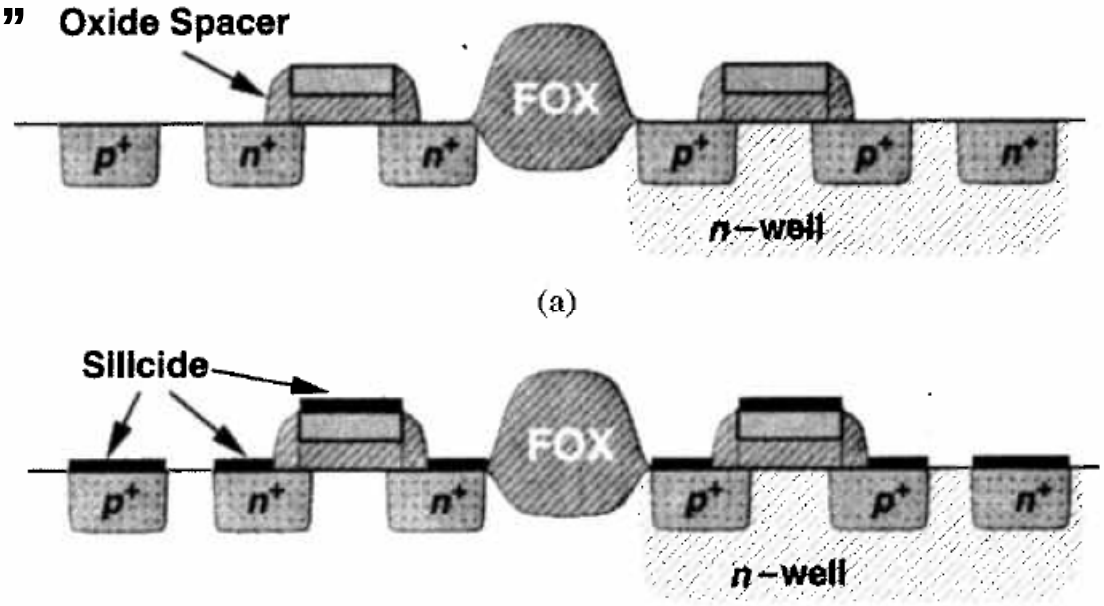
MOS Fabrication Sequence

[Razavi]



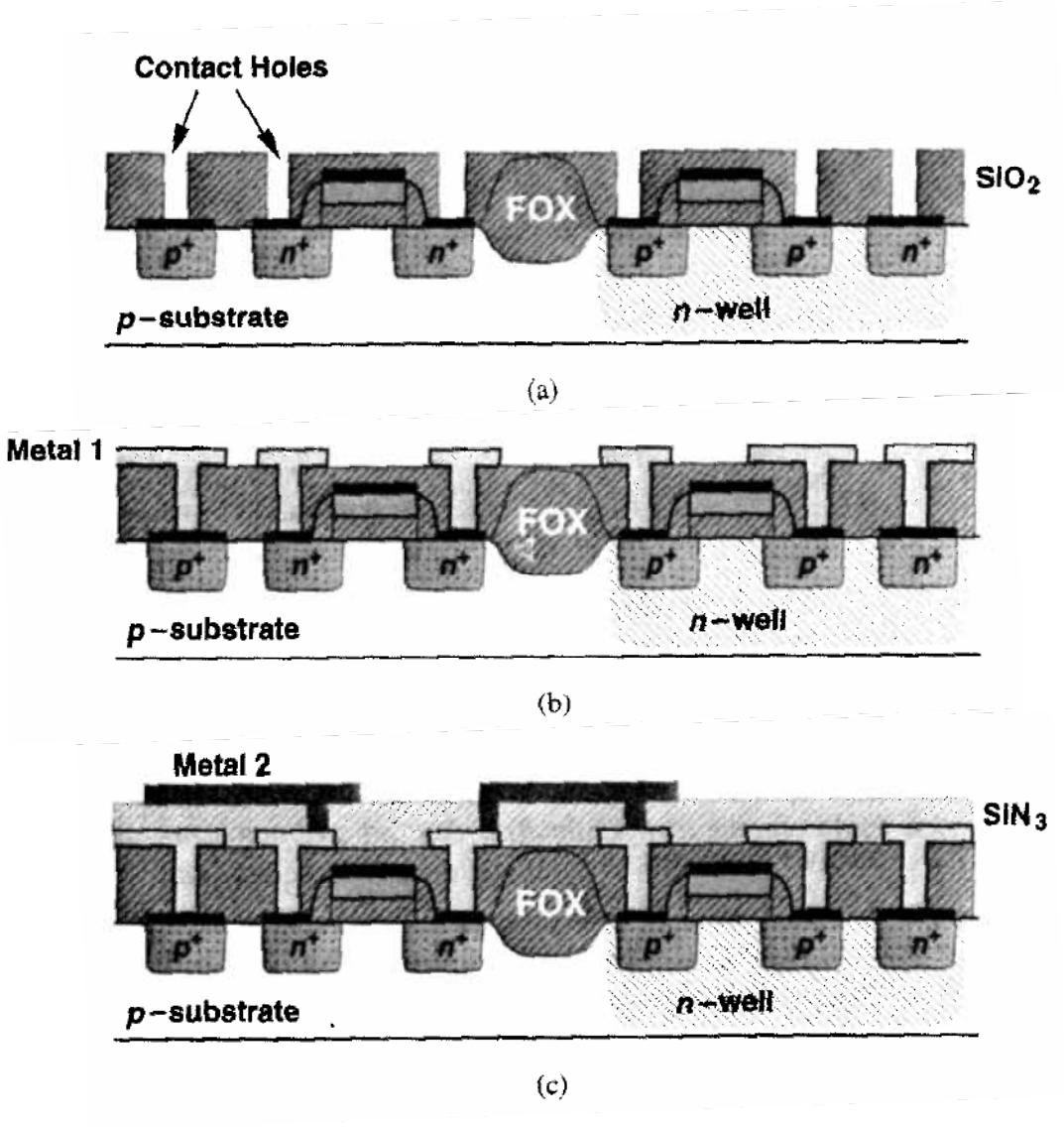
"Front-End"

"Back-End"



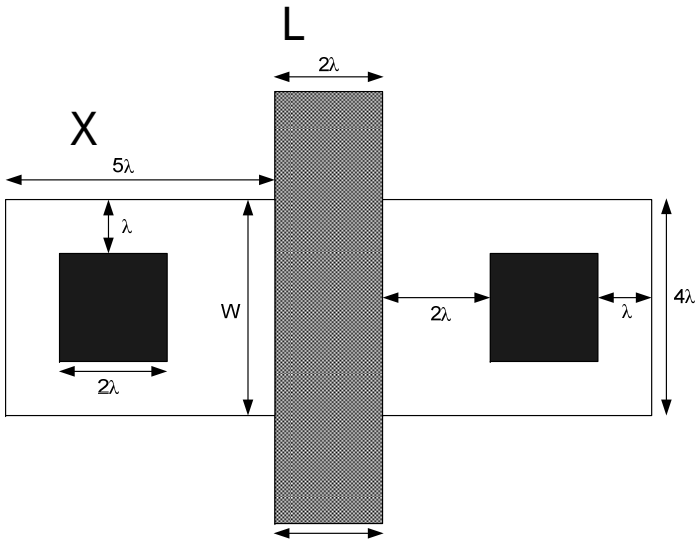
Contact and Metal Fabrication

[Razavi]



Transistor Geometries

- λ-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 = λ
  - Assume w.c. mask alignment <0.75λ
  - Relative misalignment between 2 masks is <1.5λ

$$AGate = W * L$$

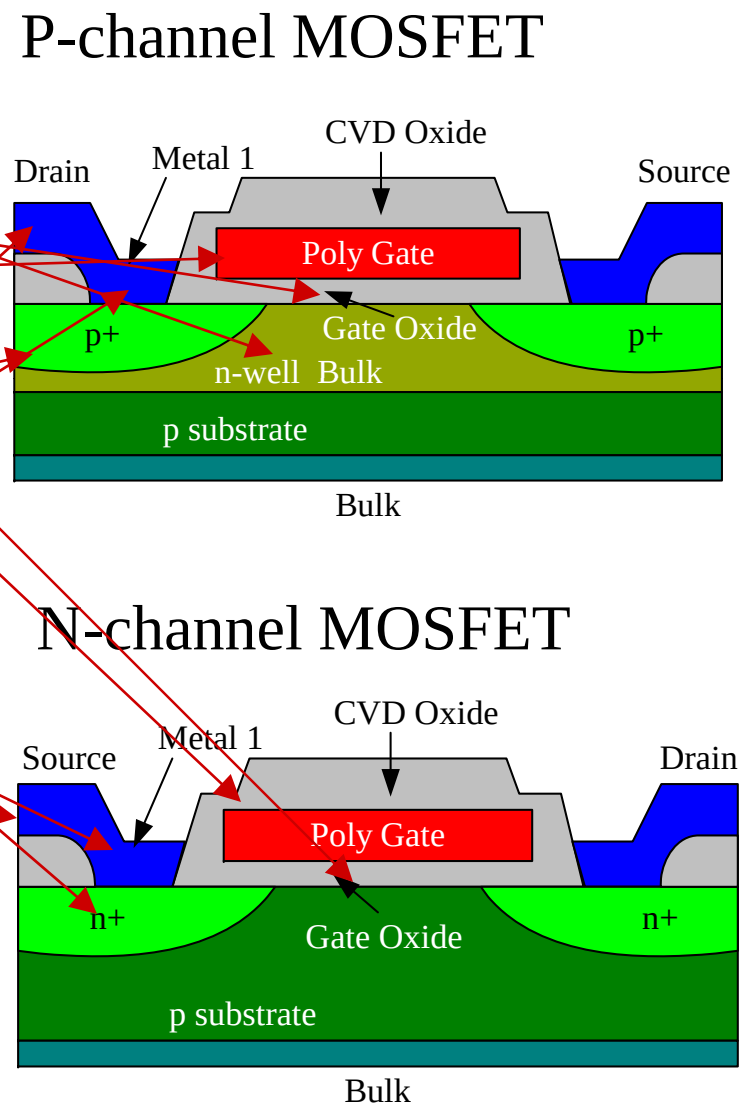
$$AD, AS = W * X$$

$$PS, PS = W + 2X \textbf{ (3 sides)}$$

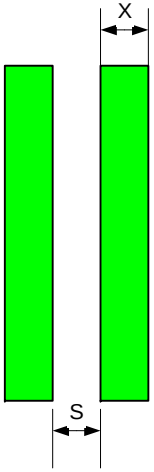
- X depends on contact size
  - 5λ in this example

BASIC SCNA CMOS LAYERS

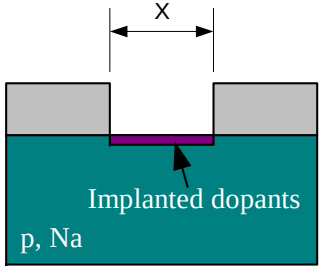
Physical Layer
N-well
Silicon Nitride
Polysilicon Layer 1
Polysilicon Layer 2
P+ Ion Implant
N+ Ion Implant
Contact cut to n+/p_
Metal 1
Via Oxide Cuts
Metal 2
Pad Contact (Overglass)



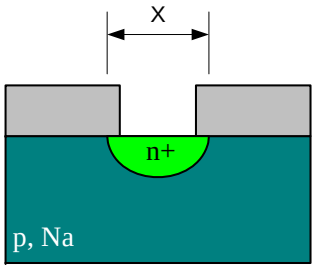
Design Rule Basics



Minimum width and spacing

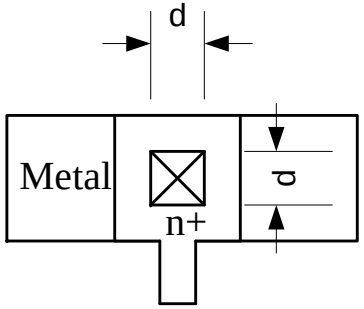


(a) Mask definition

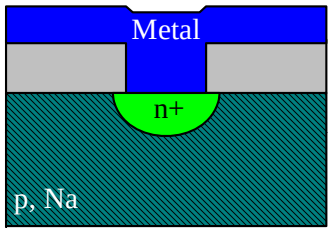


(b) After annealing

Patterning sequence for a doped n+ line.

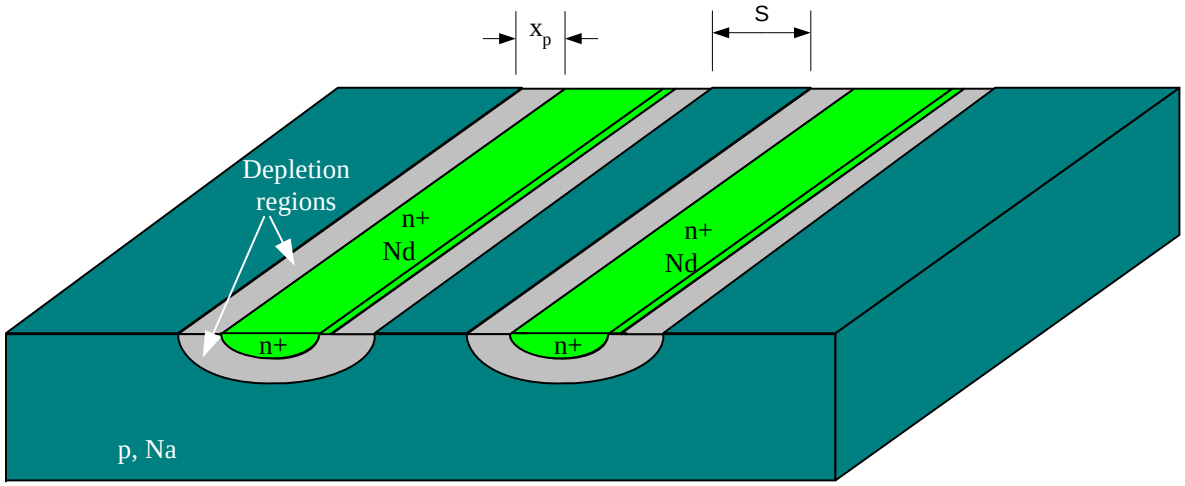


(a) Contact size

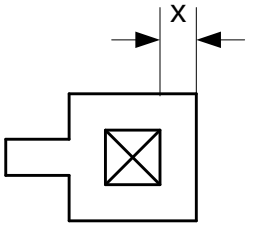


(b) Side view

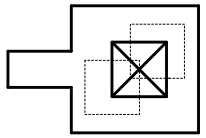
Geometry of a contact cut



Depletion regions due to parallel n+ lines

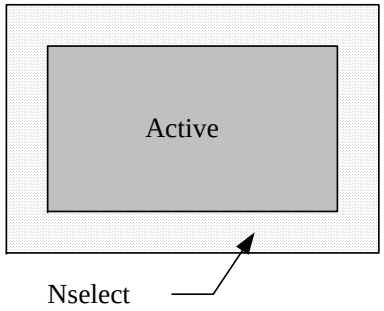


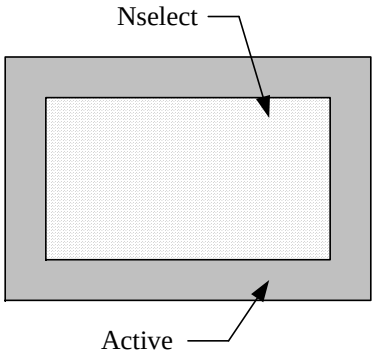
(a) Masking Design

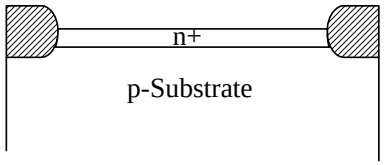


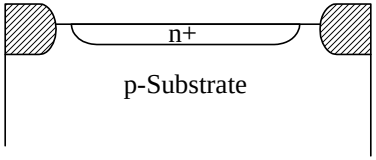
(b) Registration tolerance

Contact spacing rule



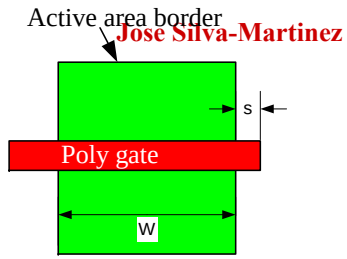




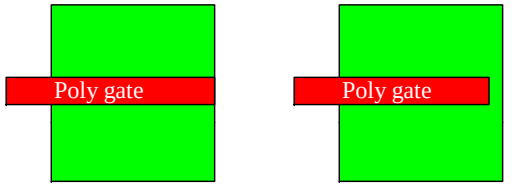


(a) Correct mask sizing (b) Incorrect mask sizing

Formation of n+ regions in an n-channel MOSFET

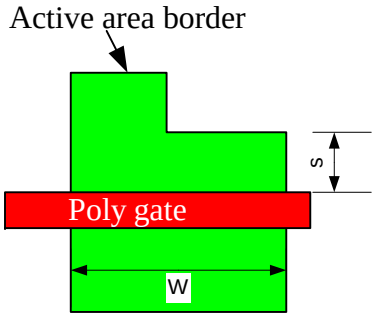


Gate overhang in MOSFET layout

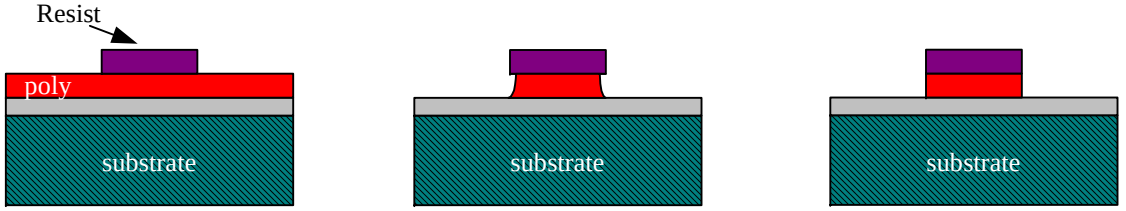


(a) No overhang (b) With misalignment

Effect of misalignment without overhang



Gate spacing form an n+ edge

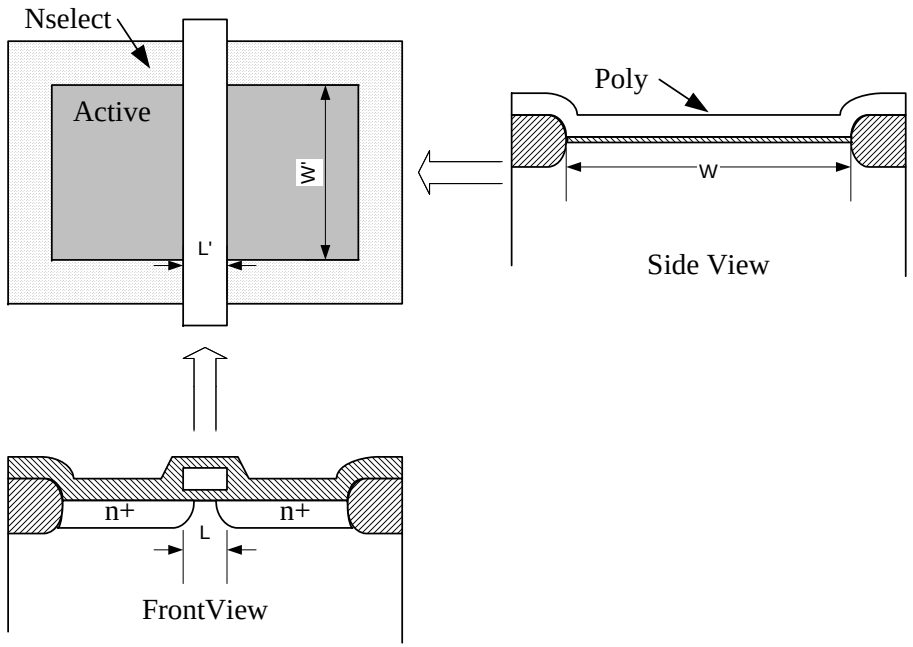


(a) Resist pattern (b) Isotropic etch (c) anisotropic etch

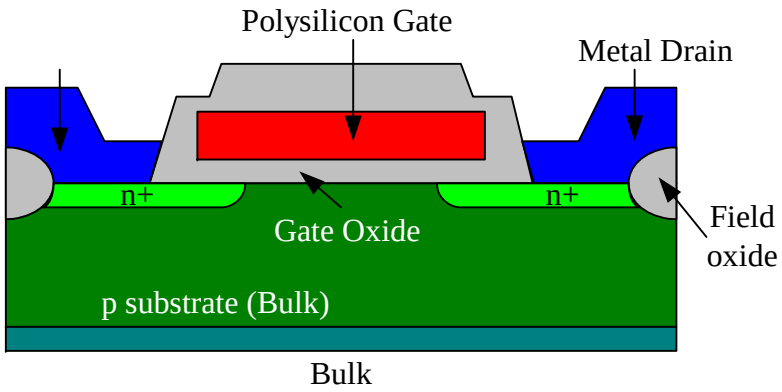
Effect of misalignment without overhang

Mask Number	Mask Layer
1	NWELL
2	ACTIVE
3	POLY
4	SELECT
5	POLY CONTACT
6	ACTIVE CONTACT
7	METAL1
8	VIA
9	METAL2
10	PAD
11	POLY2

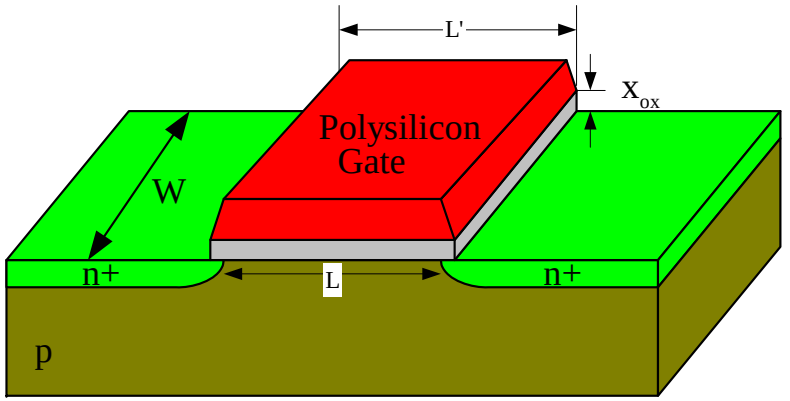
Design Rule Layers



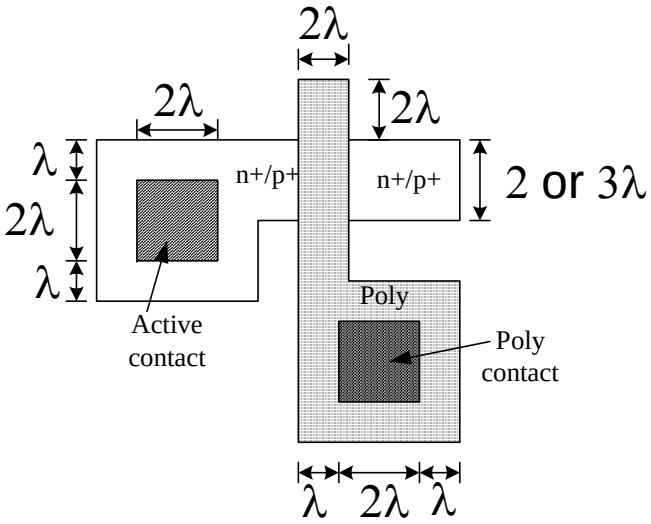
Difference between the drawn and physical values for channel length and the channel width



Structure of a n-channel MOSFET



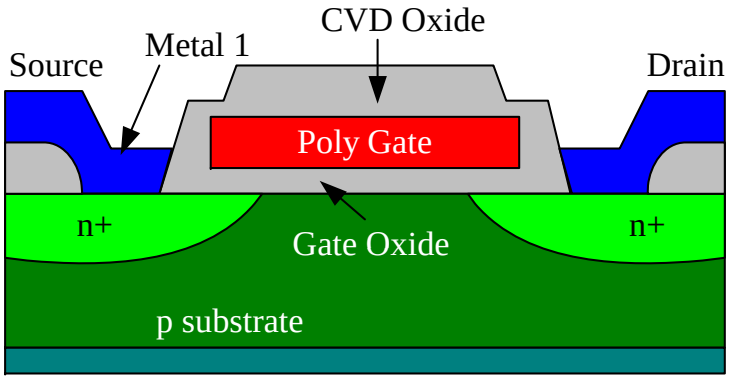
Perspective view of an n-channel MOSFET



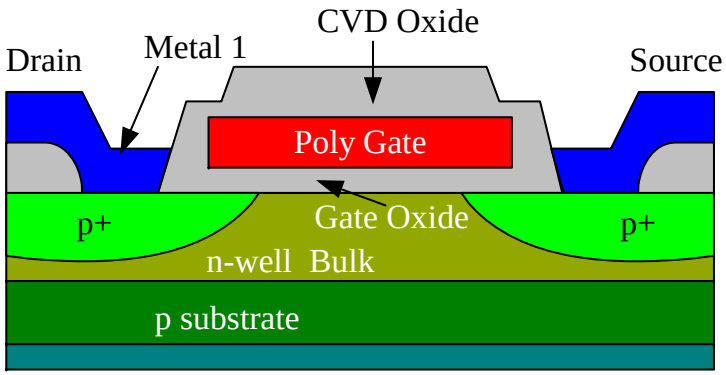
Example of Layout Rules

- Minimum transistor width is set by minimum diffusion width
  - 2 or 3λ (check with TA)
- Often, we use a use a slightly larger "minimum" that is equal to the contact height (4λ in this example)

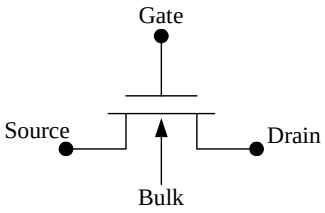
N-channel MOSFET

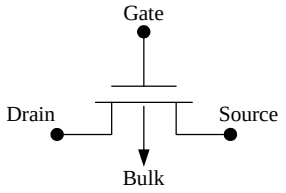


P-channel MOSFET

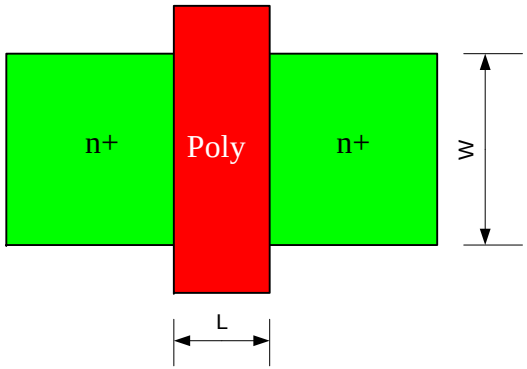


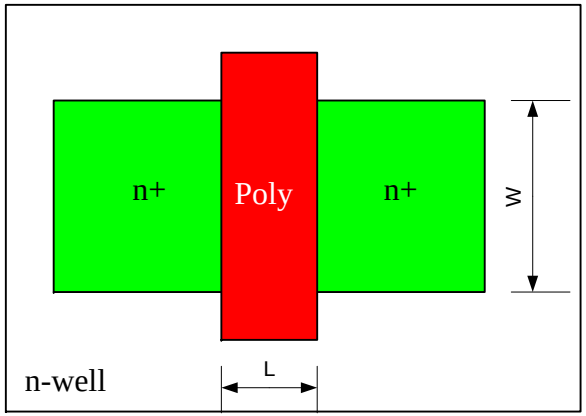
(a) Cross section





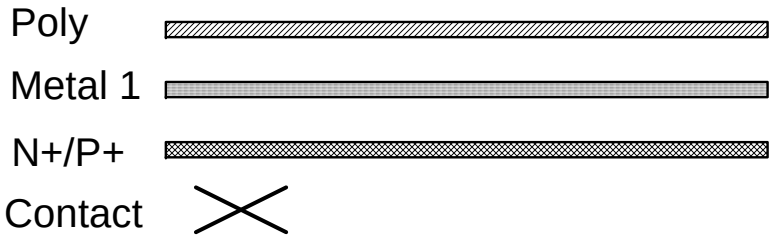
(b) Circuit symbol



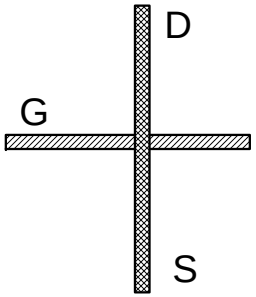


(c) Top view

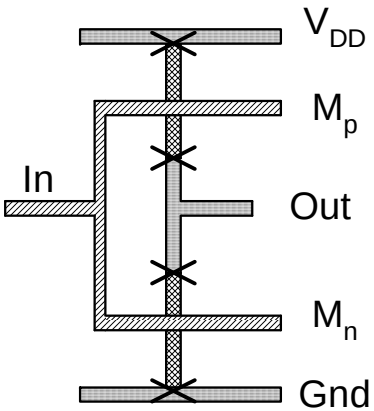
Stick Diagrams

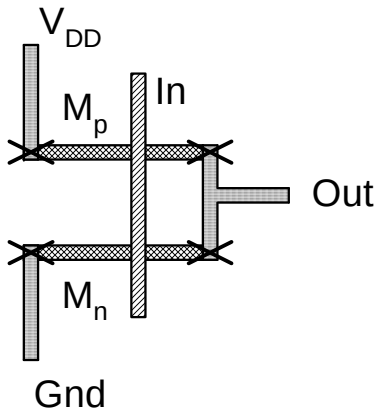


(a) Definitions



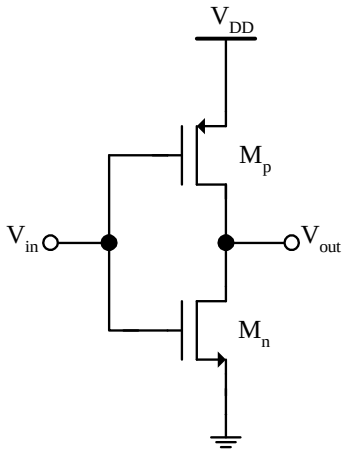
(b) MOSFET

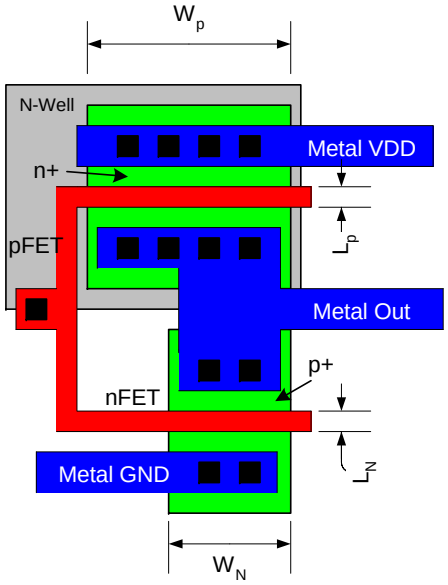




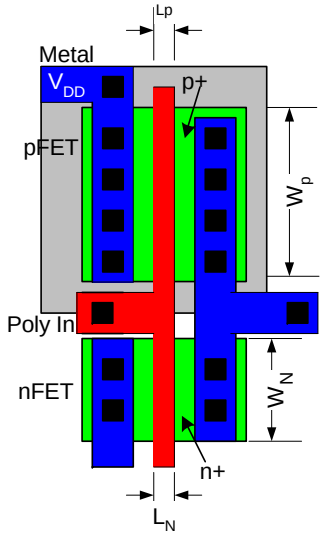
Stick diagrams for the CMOS Inverter

The CMOS Inverter



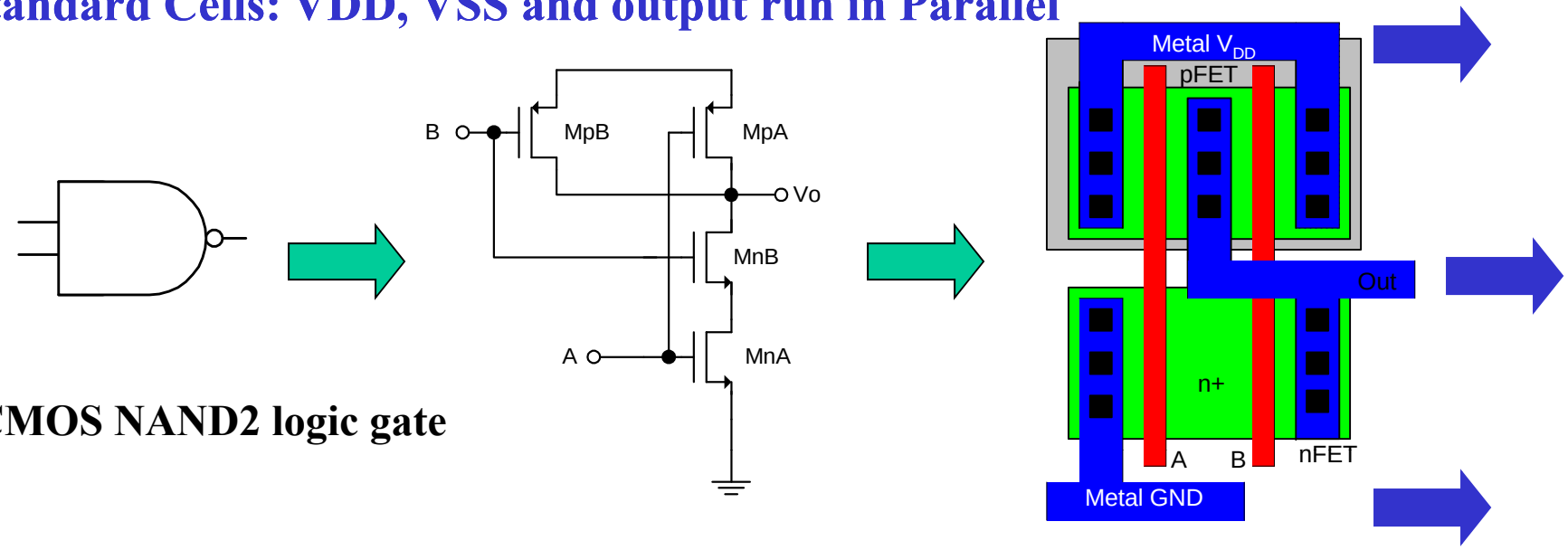


Basic Inverter Layout

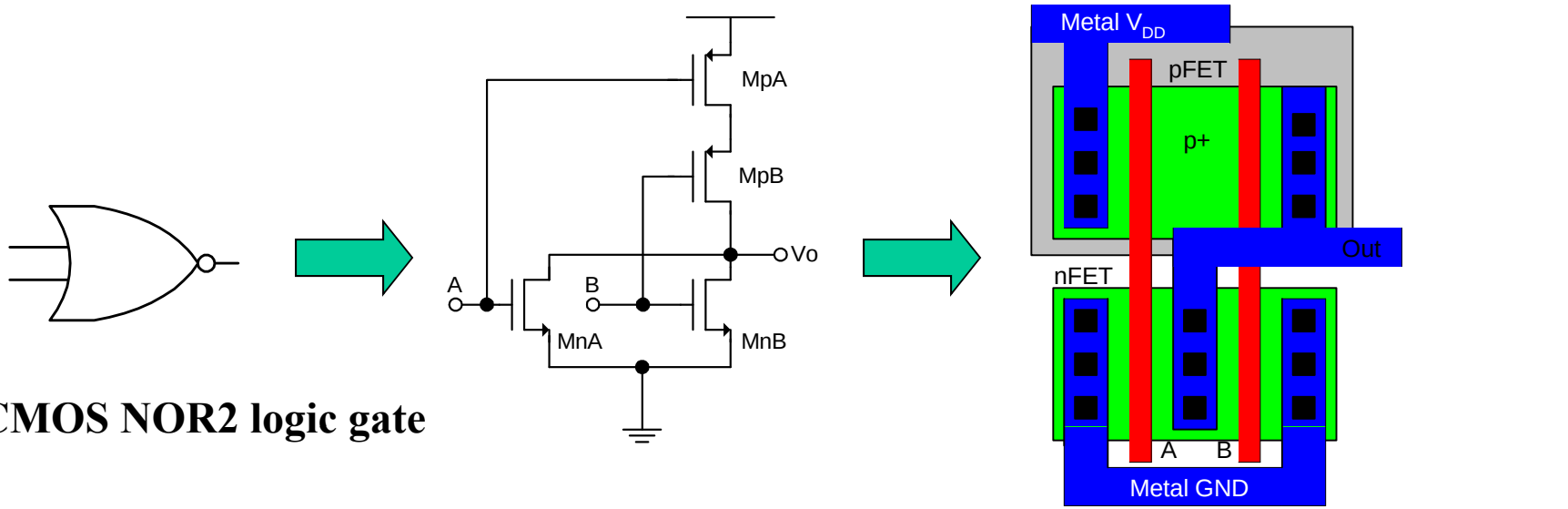


Alternate Inverter Layout

Standard Cells: VDD, VSS and output run in Parallel



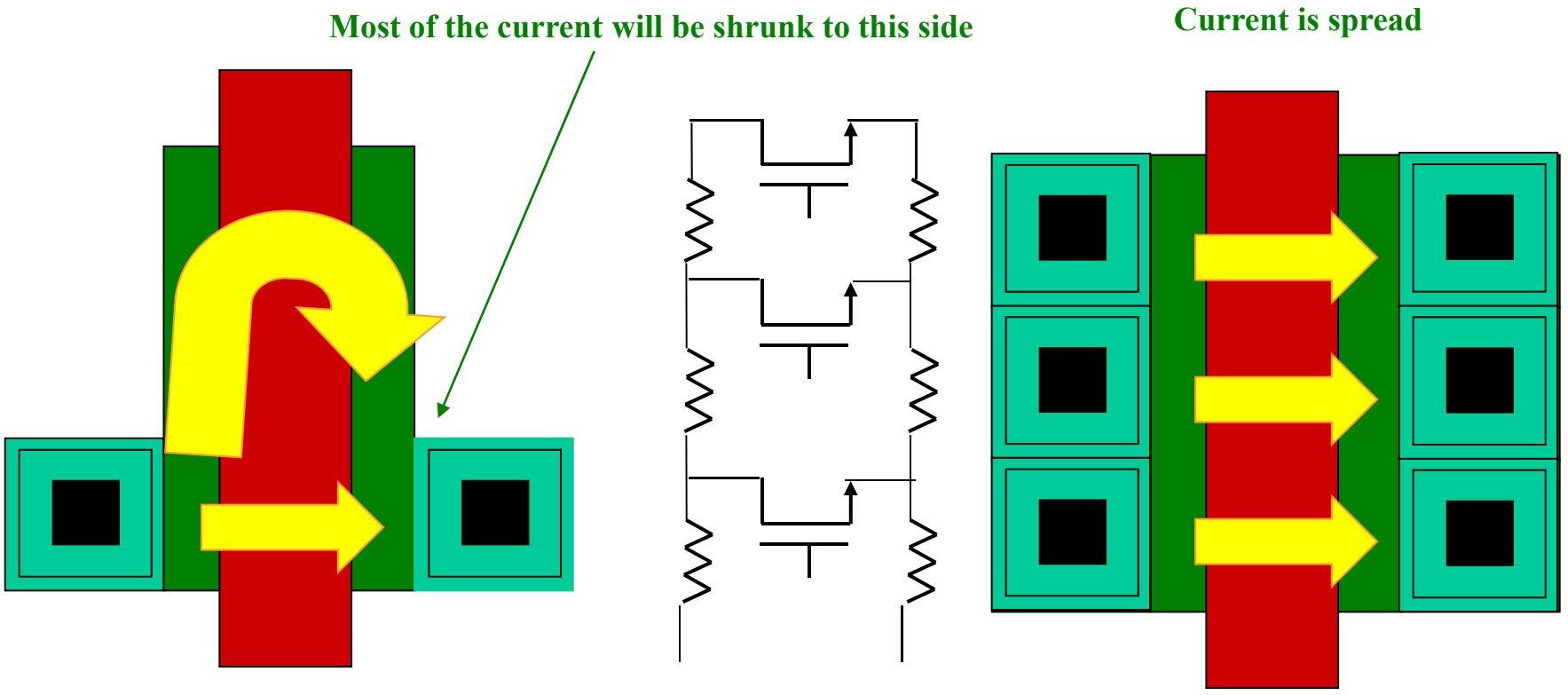
CMOS NAND2 logic gate



CMOS NOR2 logic gate

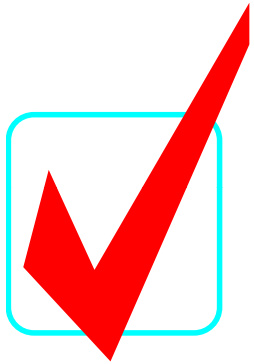
Wide Analog Transistor: Analog techniques

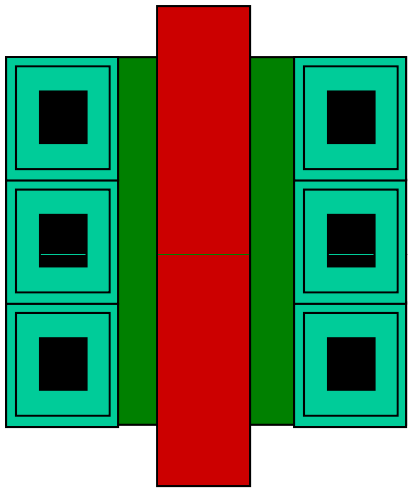
- Unacceptable drain and source resistance
- Stray resistances in transistor structure
- Contacts short the distributed resistance of diffused areas

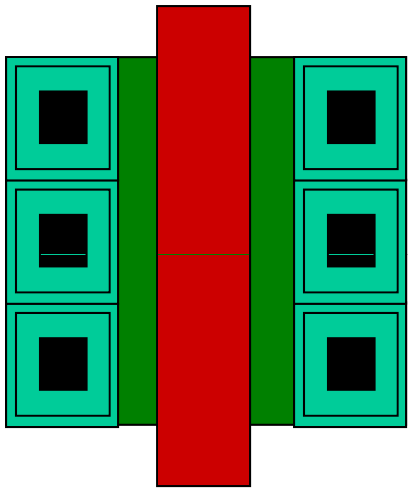


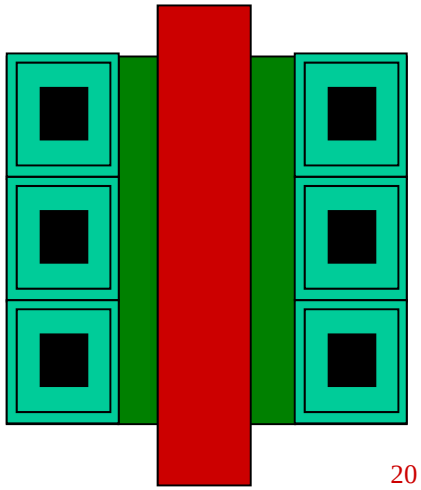
Transistor orientation

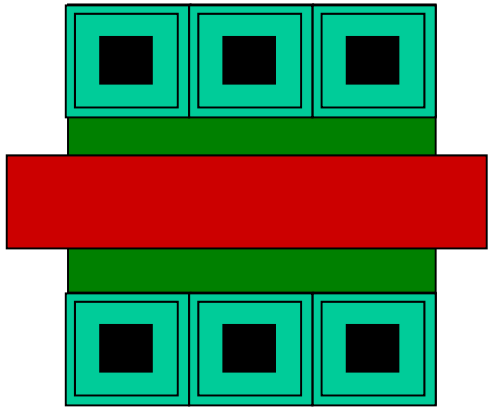
- Orientation is important in analog circuits for matching purposes





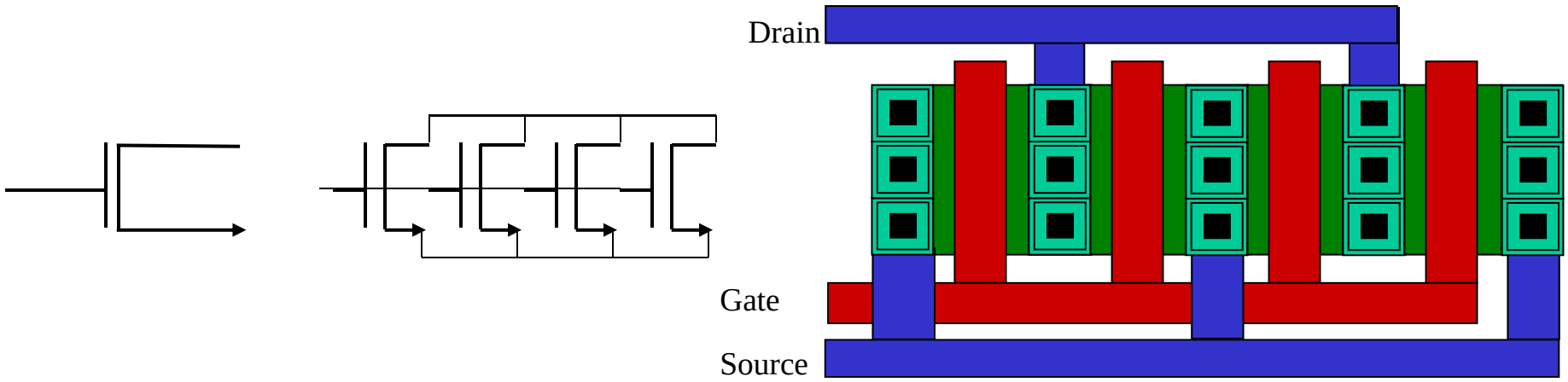






Stacked Transistors

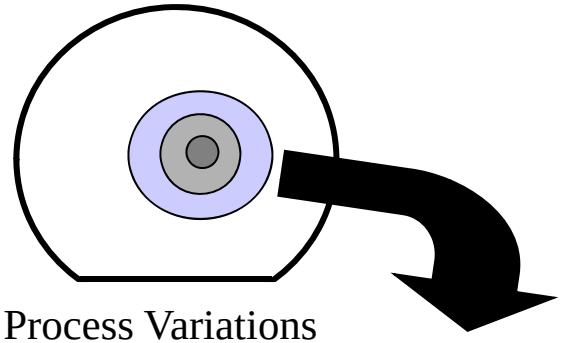
- Wide transistors need to be split
- Parallel connection of *n* elements ($n = 4$ for this example)
- Contact space is shared among transistors
- Parasitic capacitances are reduced (important for high speed)



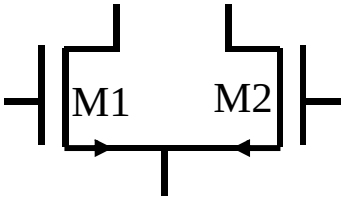
Note that parasitic capacitors are lesser at the drain

Matched Transistors

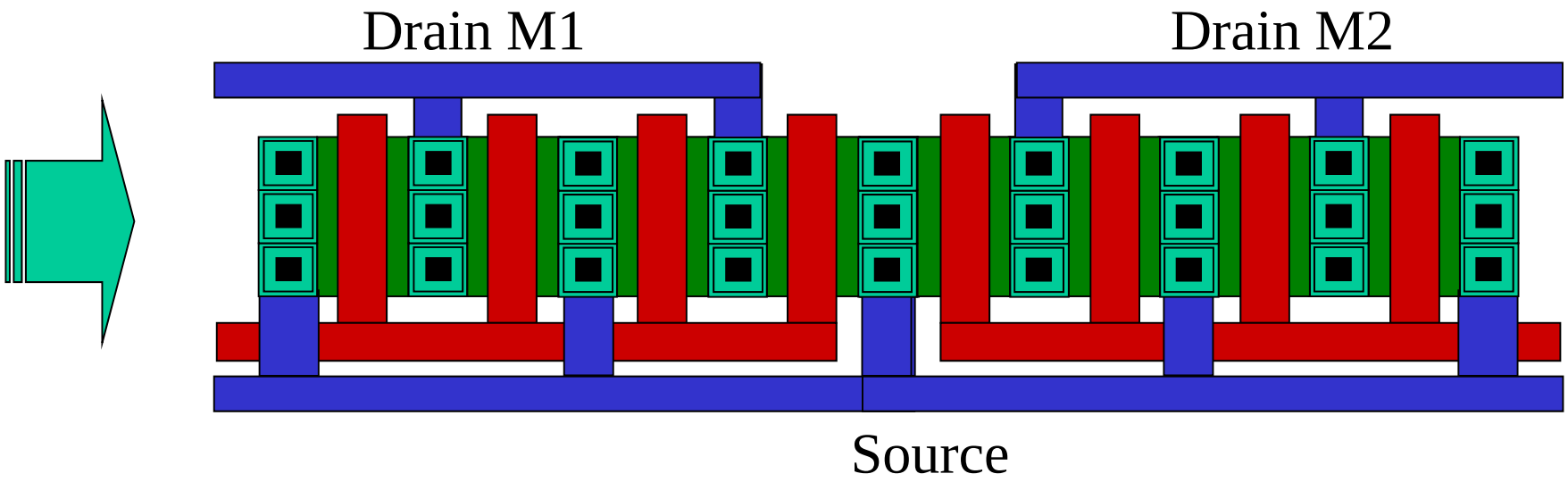
- Simple layouts are prone to process variations, e.g. V_T, KP, C_{ox}
- Matched transistors require elaborated layout techniques



Process Variations

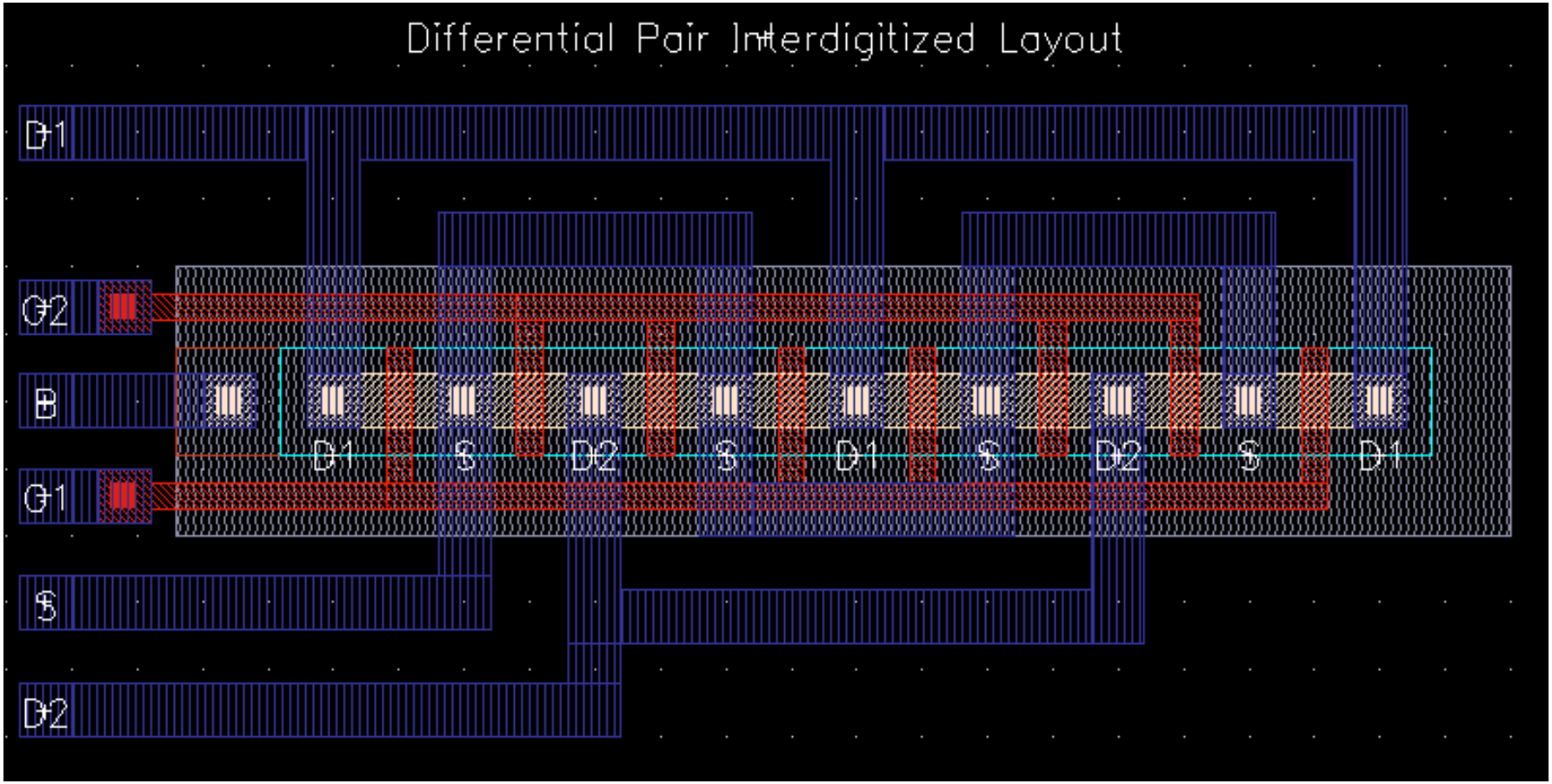


Differential pair requiring “matched transistors”

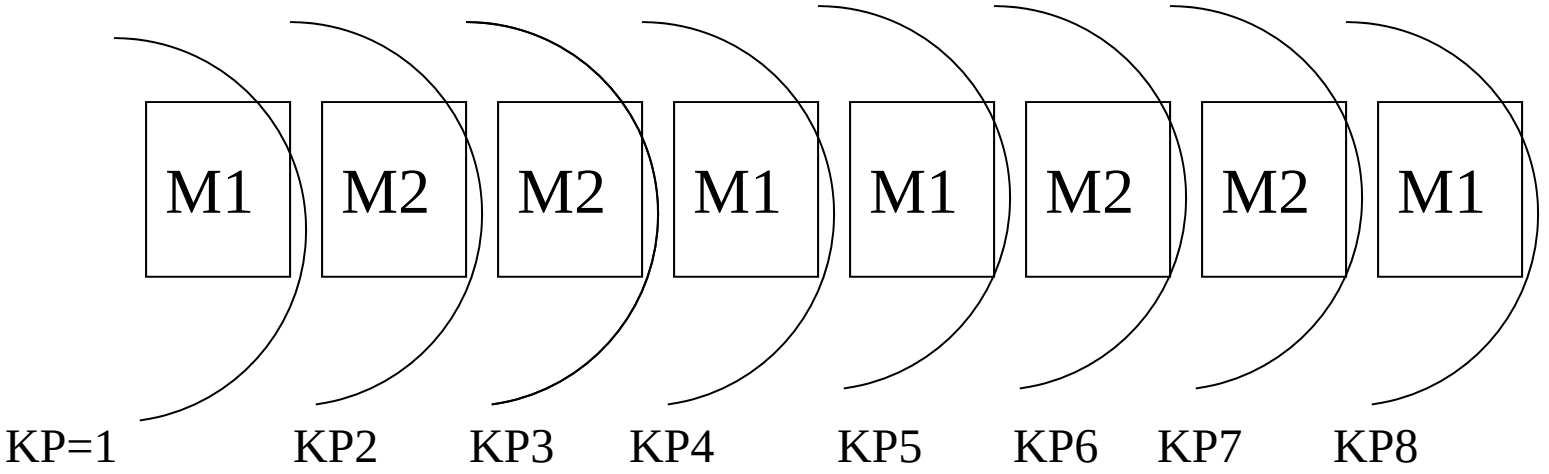


Interdigitized Layout

- Averages the process variations among transistors
- Common terminal is like a serpentine

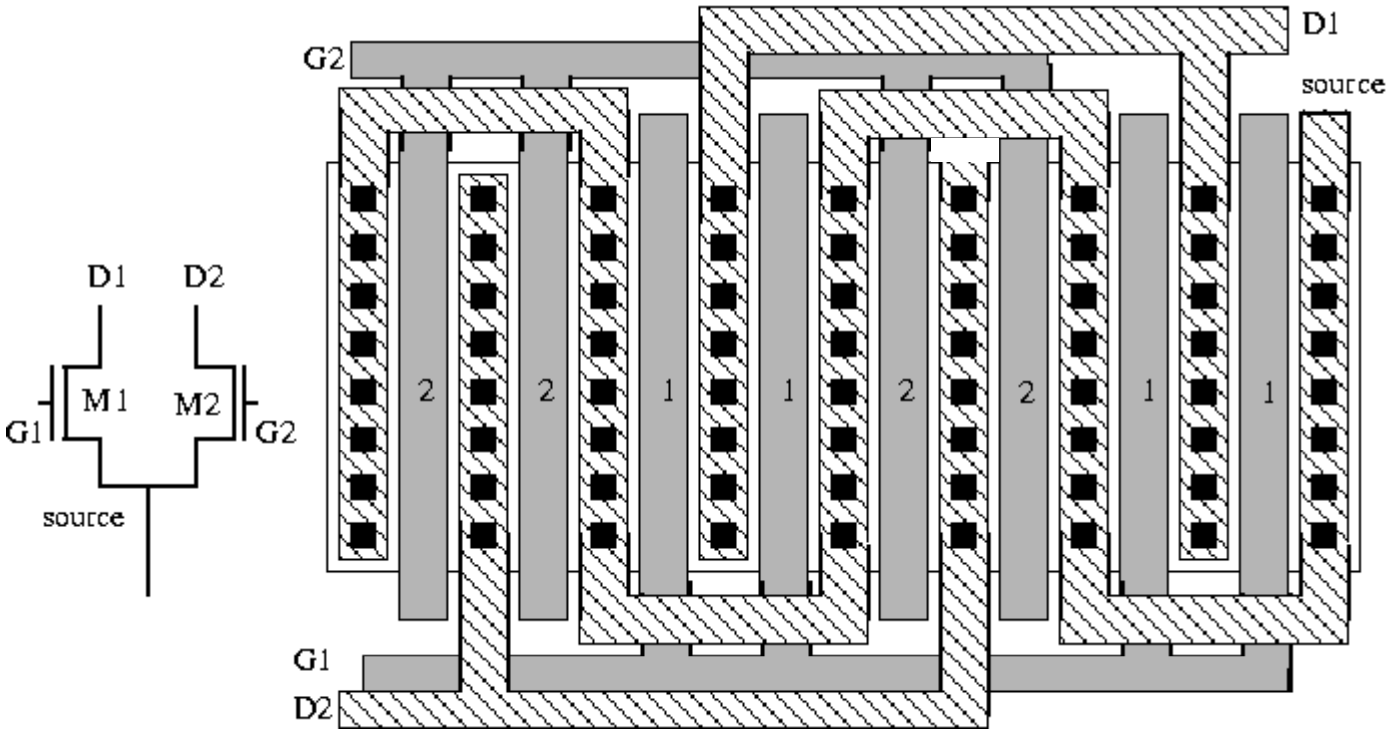


Why Interdigitized?



- Process variations are averaged among transistors
  KPs for M1: KP1+KP4+KP5+KP8 M2: KP2+KP3+KP6+KP7
- Technique maybe good for matching dc conditions
- Uneven total drain area between M1 and M2. 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 good matching is shown in the following figure :



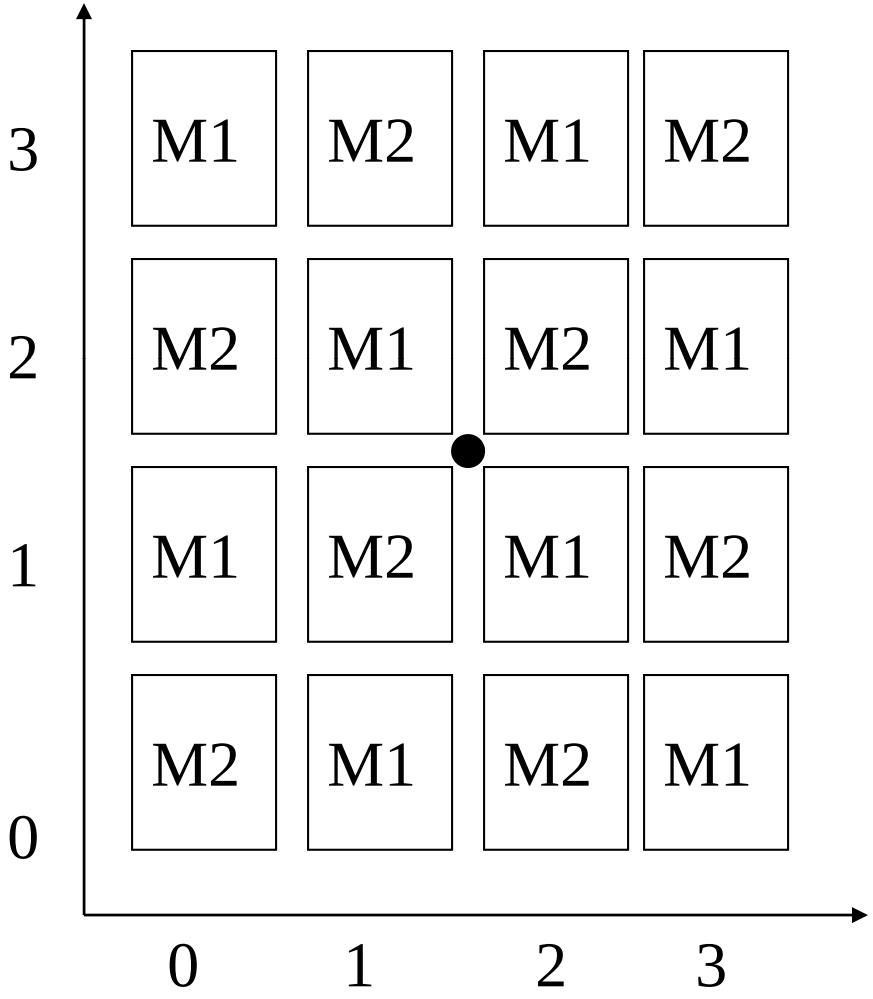
- Each transistor is split in four equal parts interleaved in two by two's.

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



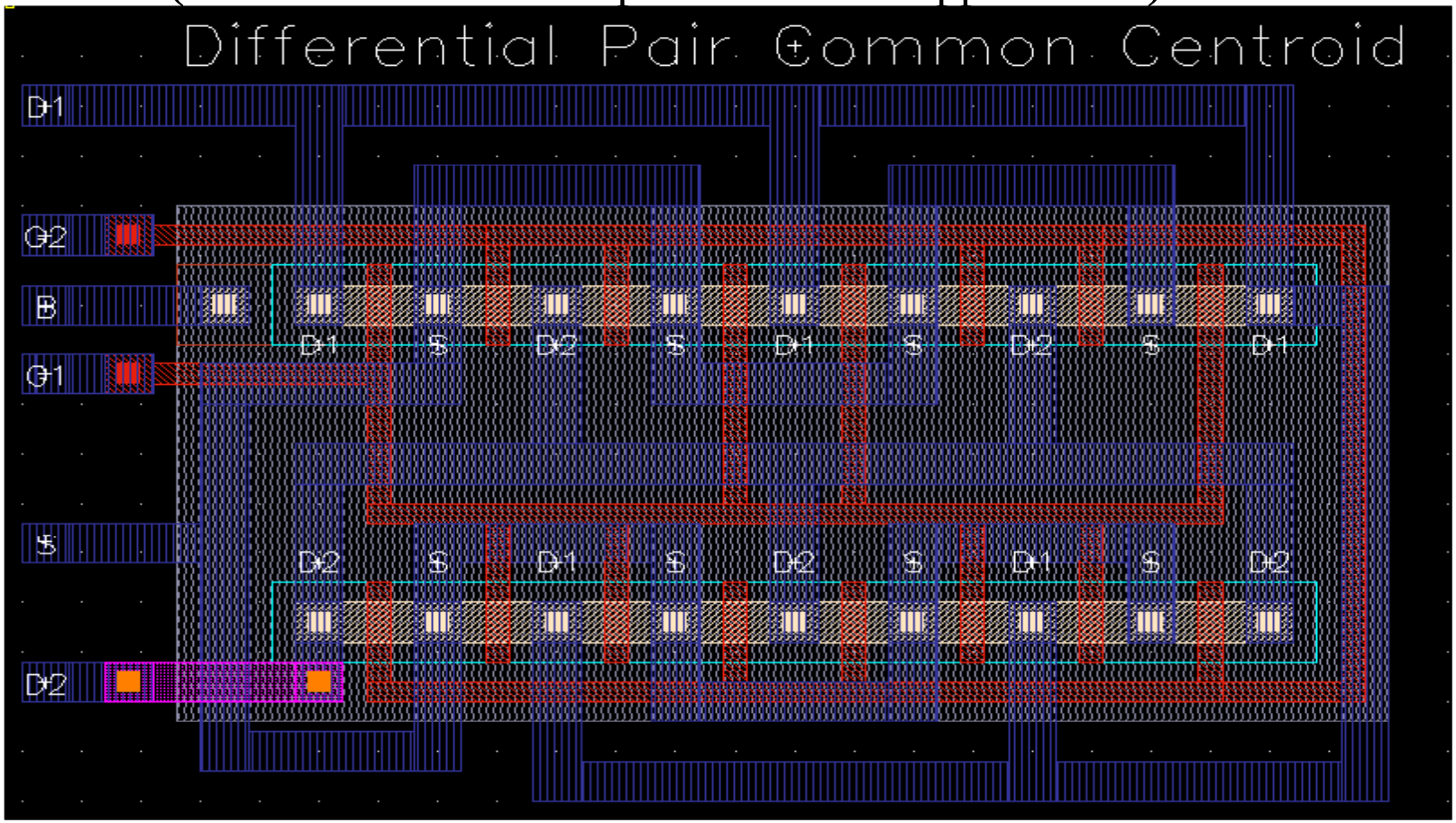
CENTROID
(complex layout)

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)

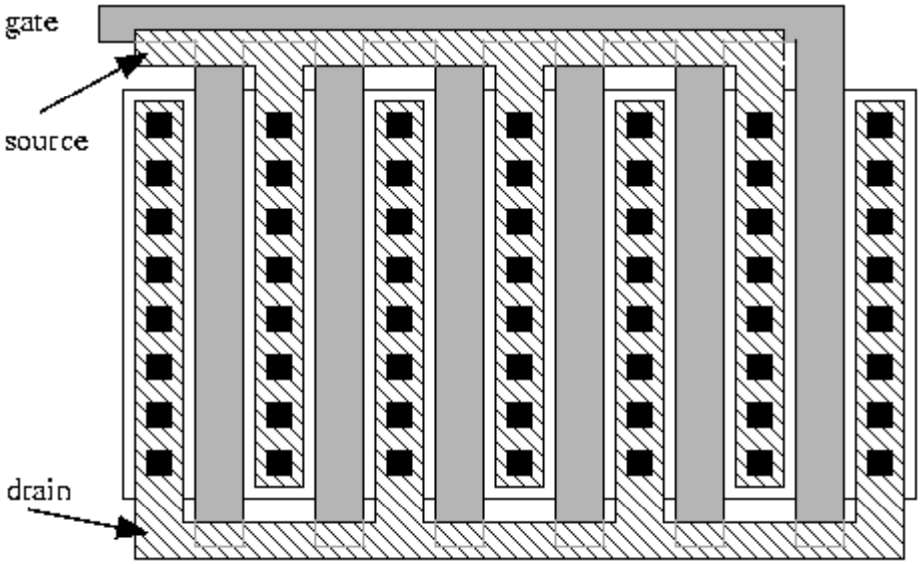
Common Centroid Layouts

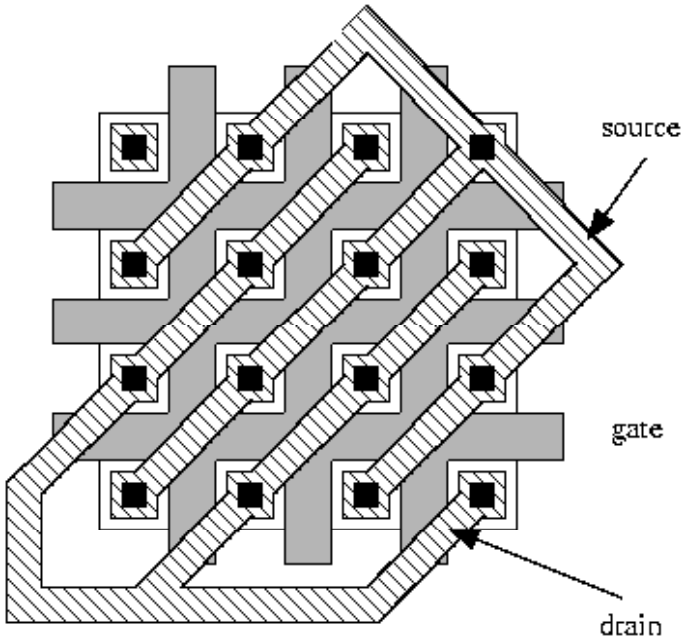
- Split into parallel connections of even parts
- Half of them will have the drain at the right side and half at the left
- Be careful how you route the common terminal
- Cross talk (effect of distributed capacitors ➔ RF applications)!



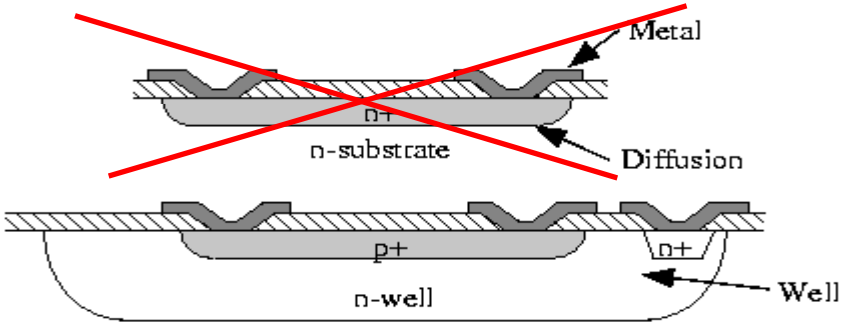
•Many contacts placed close to one another reduces series resistance and make the surface of metal connection smoother than when we use only one contact; this prevents microcraks in metal;

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



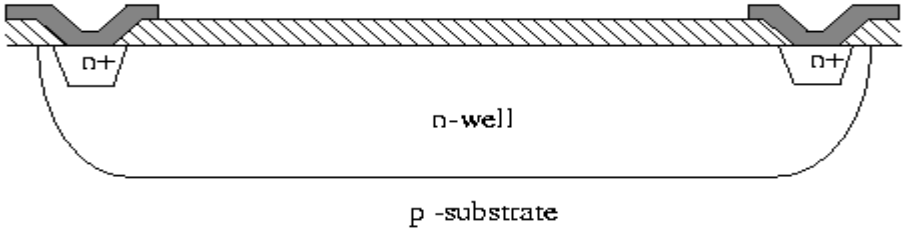


Diffusion resistors

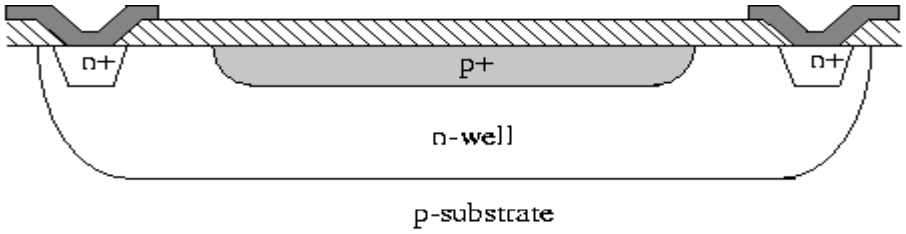


Diffused resistance

Diffused resistance



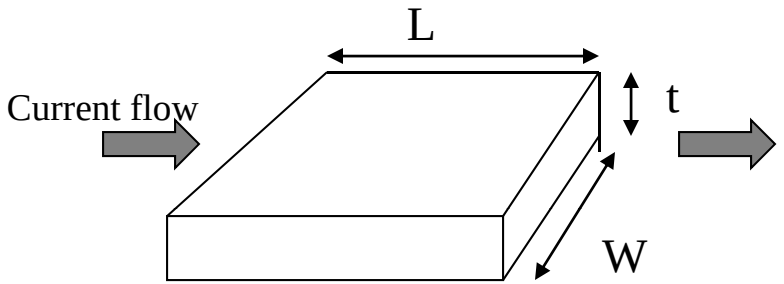
well resistance



Pinched n-well resistance

Integrated Resistors

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

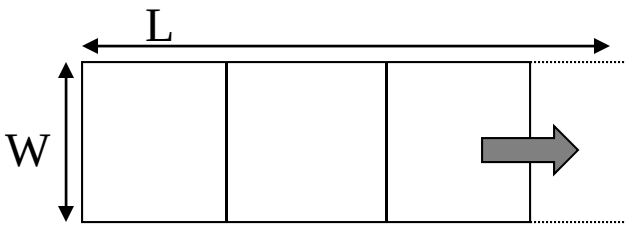


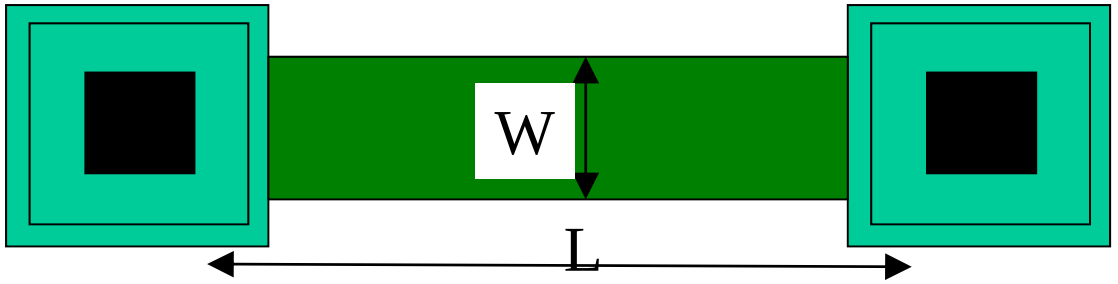
Resistivity (volumetric measure of material's resistive characteristic)

$$\rho\ (\Omega\text{-cm})$$

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

$$R = \rho/t\ \ (\Omega/)$$

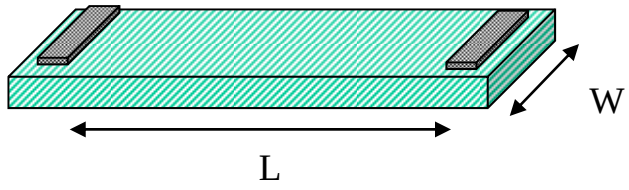




$$R = 2R_{contact} + (L/W)\ R$$

TYPICAL INTEGRATED RESISTORS

$$R = 2R_{cont} + \frac{L}{W} R_{\square}$$



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

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

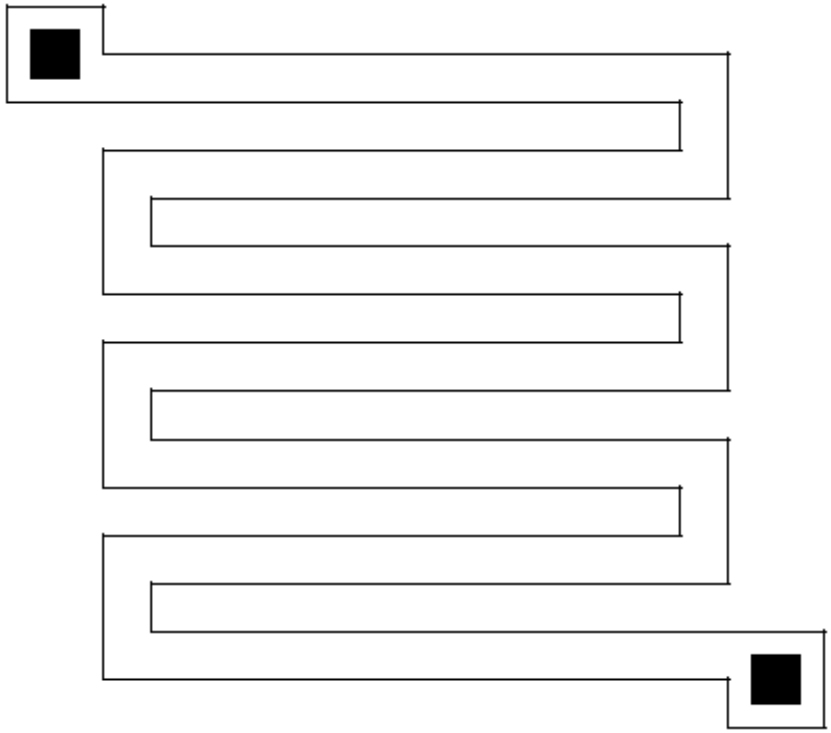
Large Resistors

In order to implement large resistors :

- Use of long strips (large L/W)
- Use of layers with high sheet resistance (bad performances)

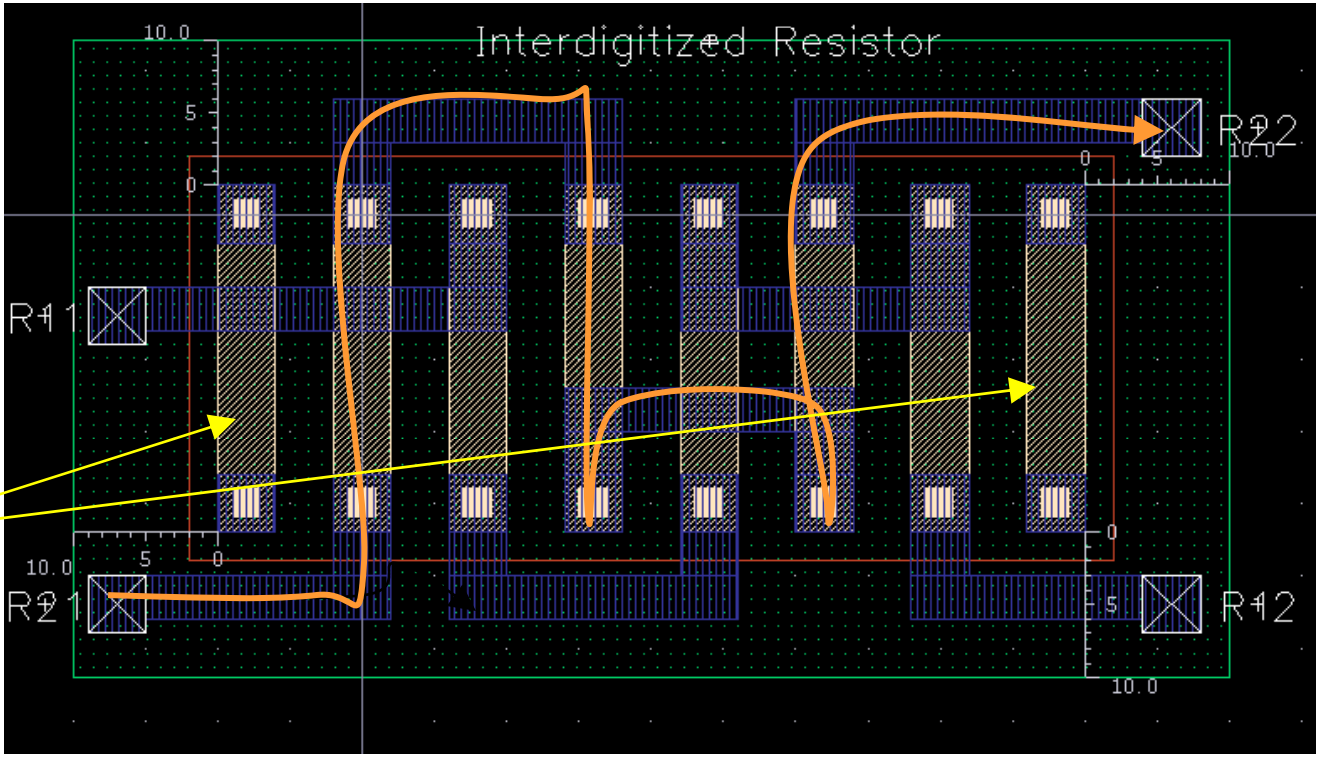
Layout : rectangular “serpentine”

$$R = \frac{L}{W} R_{\square} = \frac{L}{W} \cdot \frac{\bar{\rho}}{x_j}$$



Well-Diffusion Resistor

- Example shows two long resistors for KΩ range
- Alternatively, “serpentine” shapes can be used
- Noise problems from the body
  - Substrate bias surrounding the well
  - Substrate bias between the parallel strips



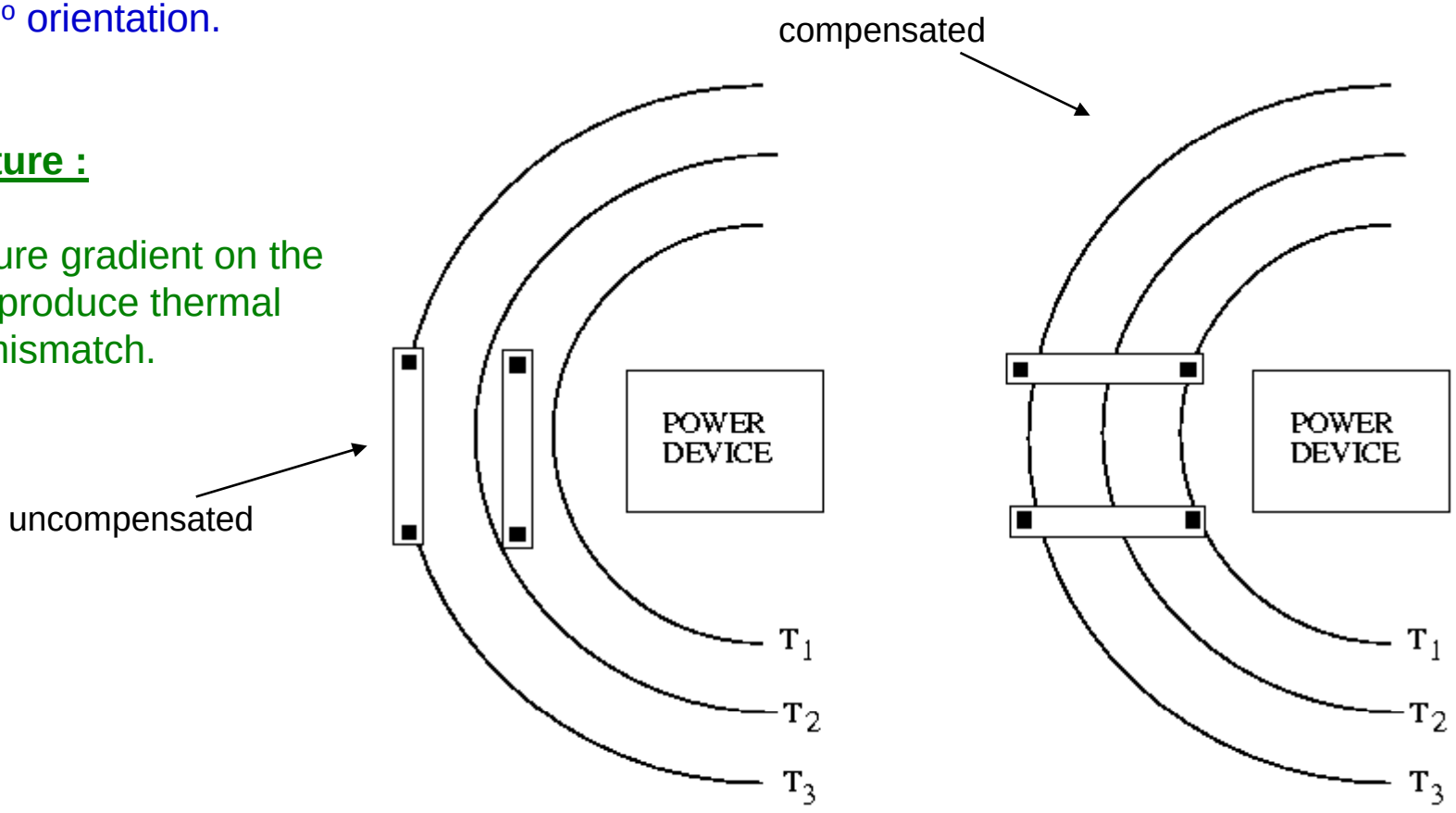
Factors affecting accuracy :

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

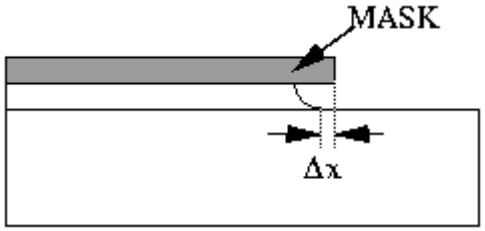
For <100> material the variation is unisotropic, so the minimum is obtained if the resistance have a 45^{o} orientation.

Temperature :

Temperature gradient on the chip may produce thermal induced mismatch.



Etching



Wet etching : isotropic (undercut effect)

HF for SiO_2 ; H_3PO_4 for Al

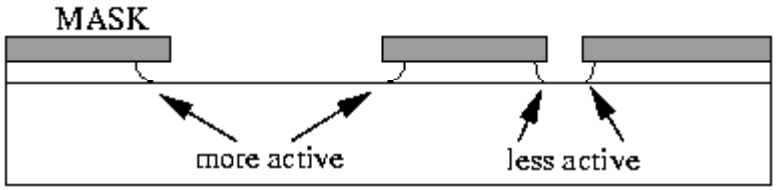
Δx for polysilicon may be 0.2 – 0.4 μm with standard deviation 0.04 – 0.08 μm.

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

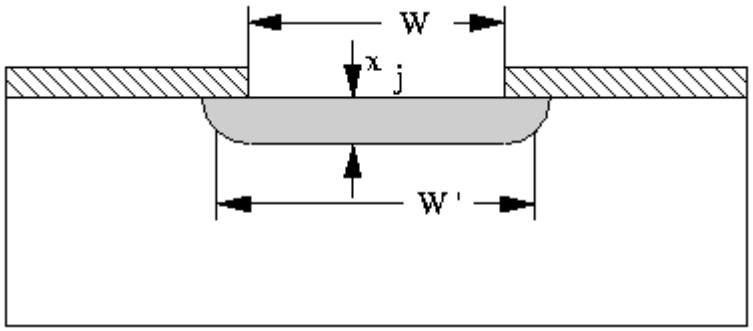
Boundary :

The etching depends on the boundary conditions

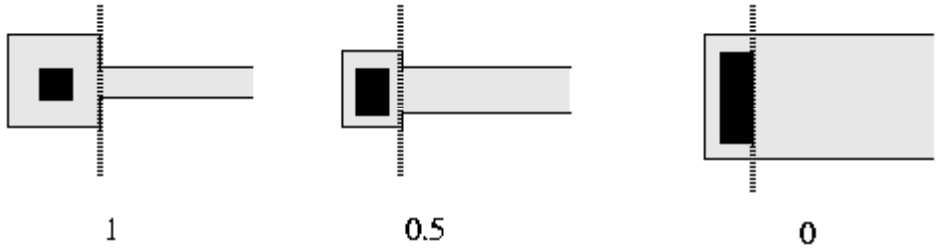


- **Use dummy strips**

Side diffusion effect : Contribution of endings

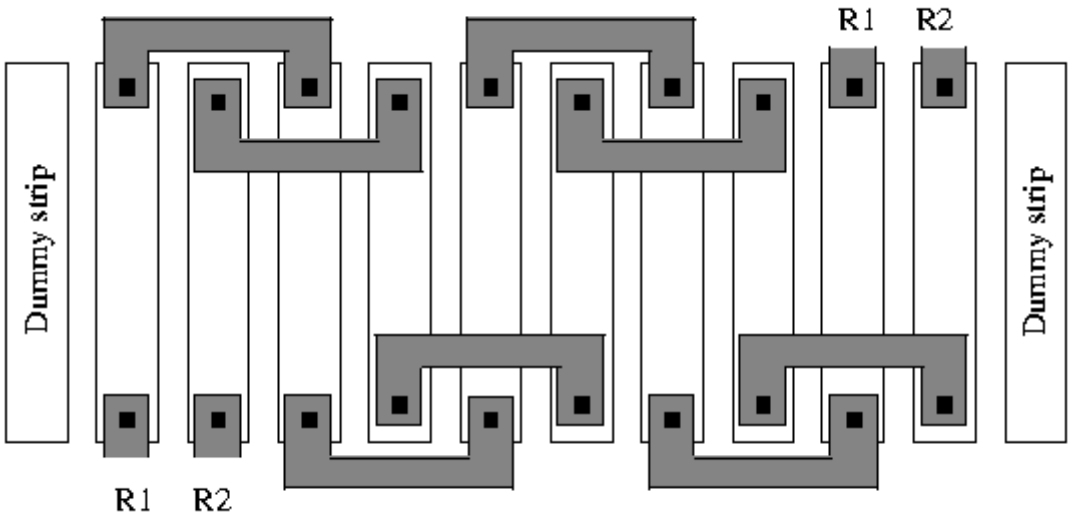


Side Diffusion "widens" R

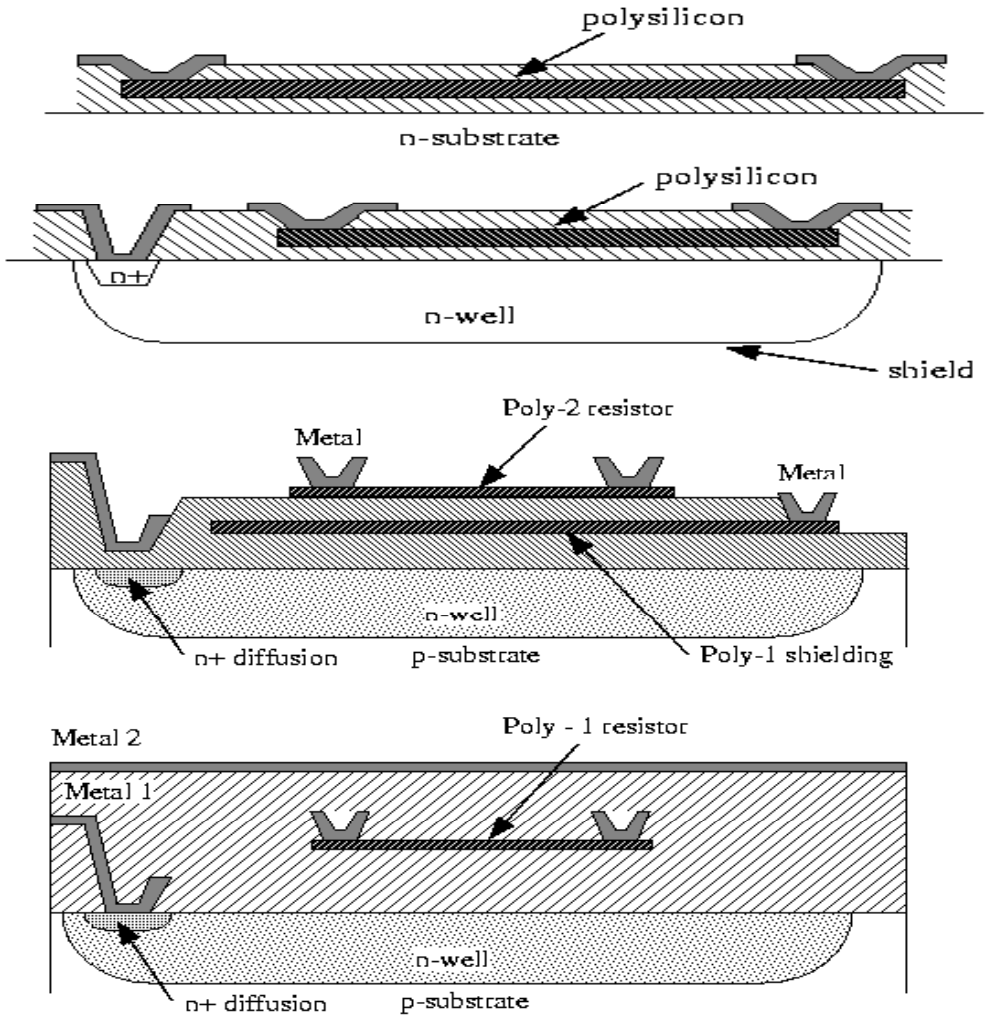


Impact of R_{cont} depends on relative geometry

Interdigitized structure :



Poly Resistors



a) First polysilicon resistance

b) First polysilicon resistance with a well shielding

Second polysilicon resistance

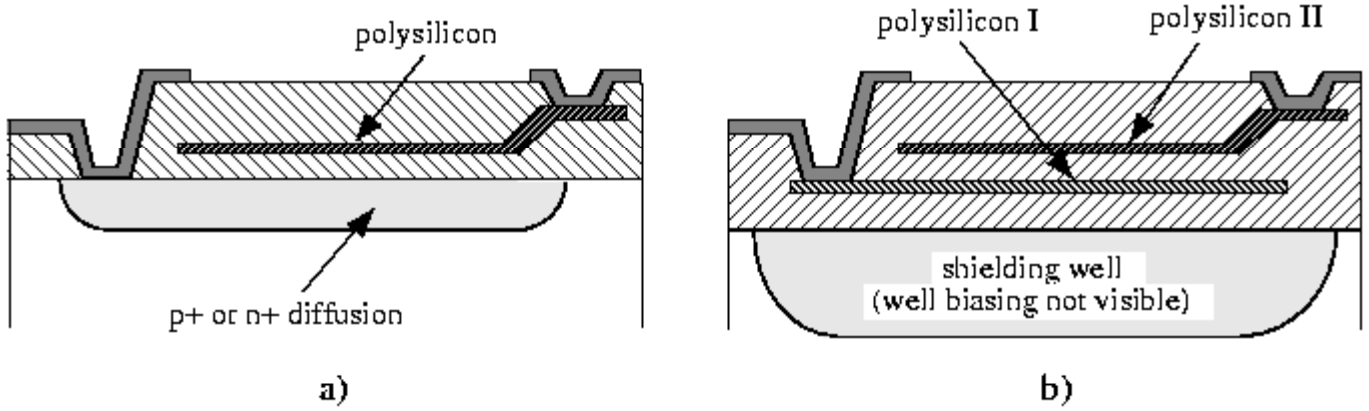
Second polysilicon resistance with a well shielding

Typical Resistance Process Data 0.8 μm process

	Sheet Resistance (Ω/□)	Width Variation (μm) (measured-drawn)	Contact Resistance (Ω)
N+Actv	52.2	-0.66	66.8
P+Actv	75.6	-0.73	37.5
Poly	36.3	-0.10	30.6
Poly 2	25.5	0.31	20.7
Mtl 1	0.05	0.56	0.05
Mtl 2	0.03	-0.06	
N-Well	1513		

Gate oxide thickness 316 angstroms

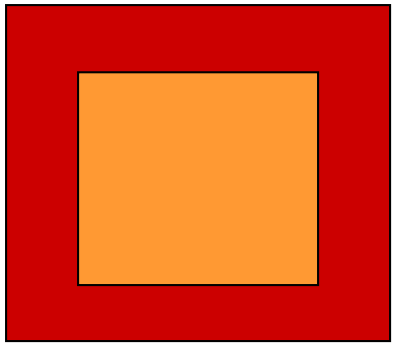
TYPES OF INTEGRATED CAPACITORS



Electrodes : metal; polysilicon; diffusion

Insulator : silicon oxide; polysilicon oxide; CVD oxide

$$C = \frac{\varepsilon_{ox}}{t_{ox}} WL$$



TOP VIEW

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

Factor affecting accuracy

- Oxide damage
- Impurities
- Bias condition
- Bias history (for CVD)
- Stress
- Temperature

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

- Grow rate
- Poly grain size

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

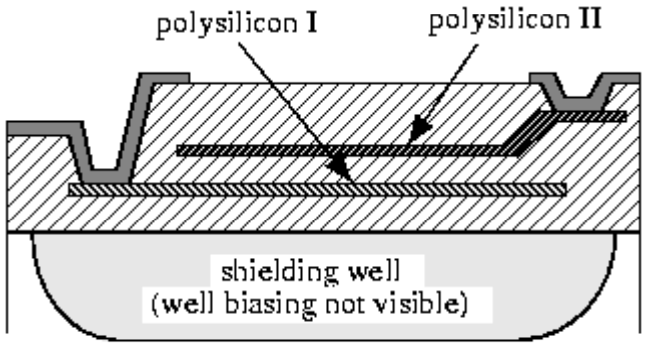
- Etching
- Alignment

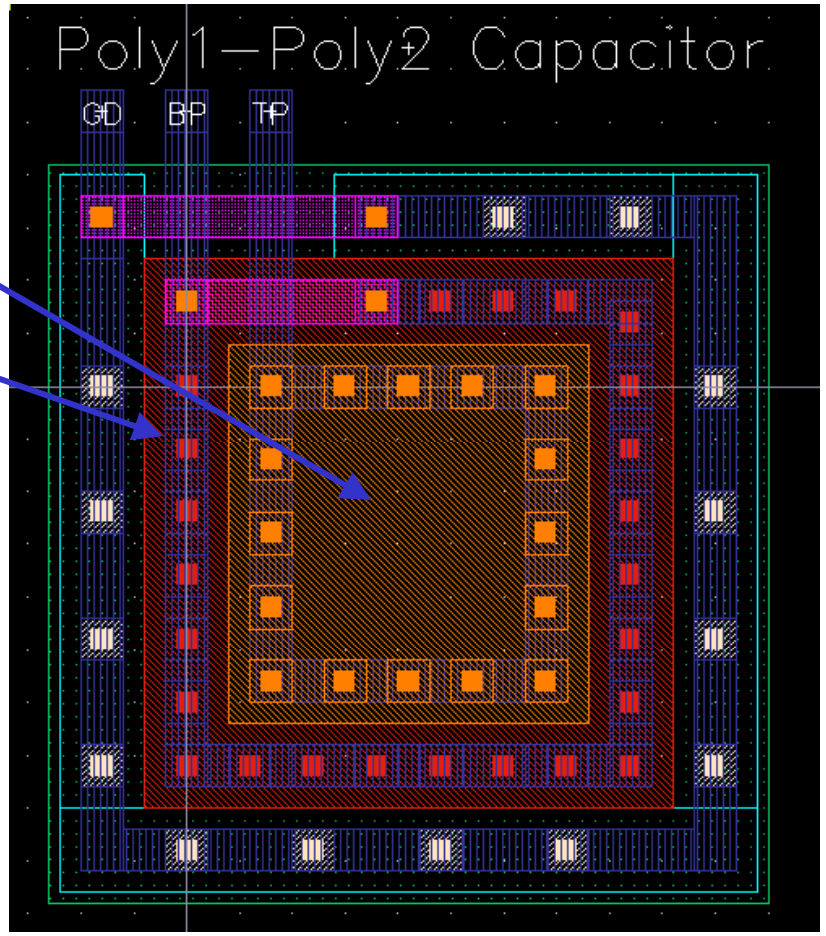
$$\left(\frac{\Delta L}{L}\right);\left(\frac{\Delta W}{W}\right)$$

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

$$\frac{\Delta C}{C}\approx 1-0.1\%$$

Poly1 - Poly2 Capacitor





- Area is determined by poly2
- Problems
  - undercut effects
  - nonuniform dielectric thickness
  - matching among capacitors
  - Minimize the rings (inductors)

Accuracy of integrated capacitors

Perimeter effects led 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 $\Delta x = \Delta y = \Delta e$

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

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

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

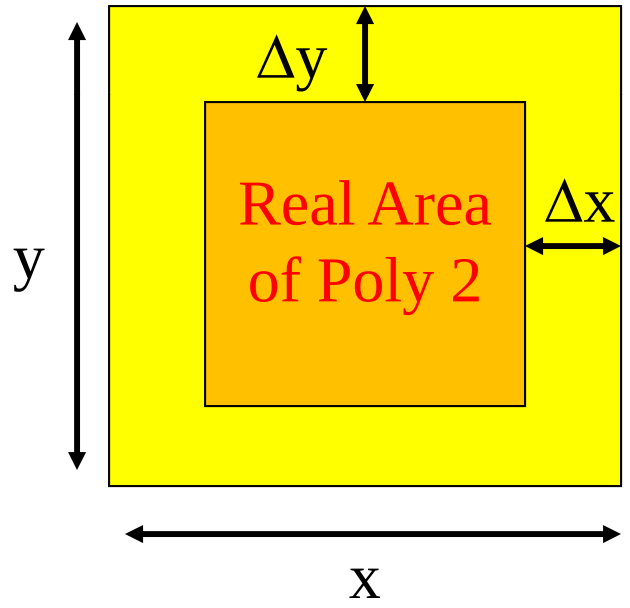
The relative error is

$\varepsilon = C_e/C$

$= -2\Delta e(x + y) / xy$

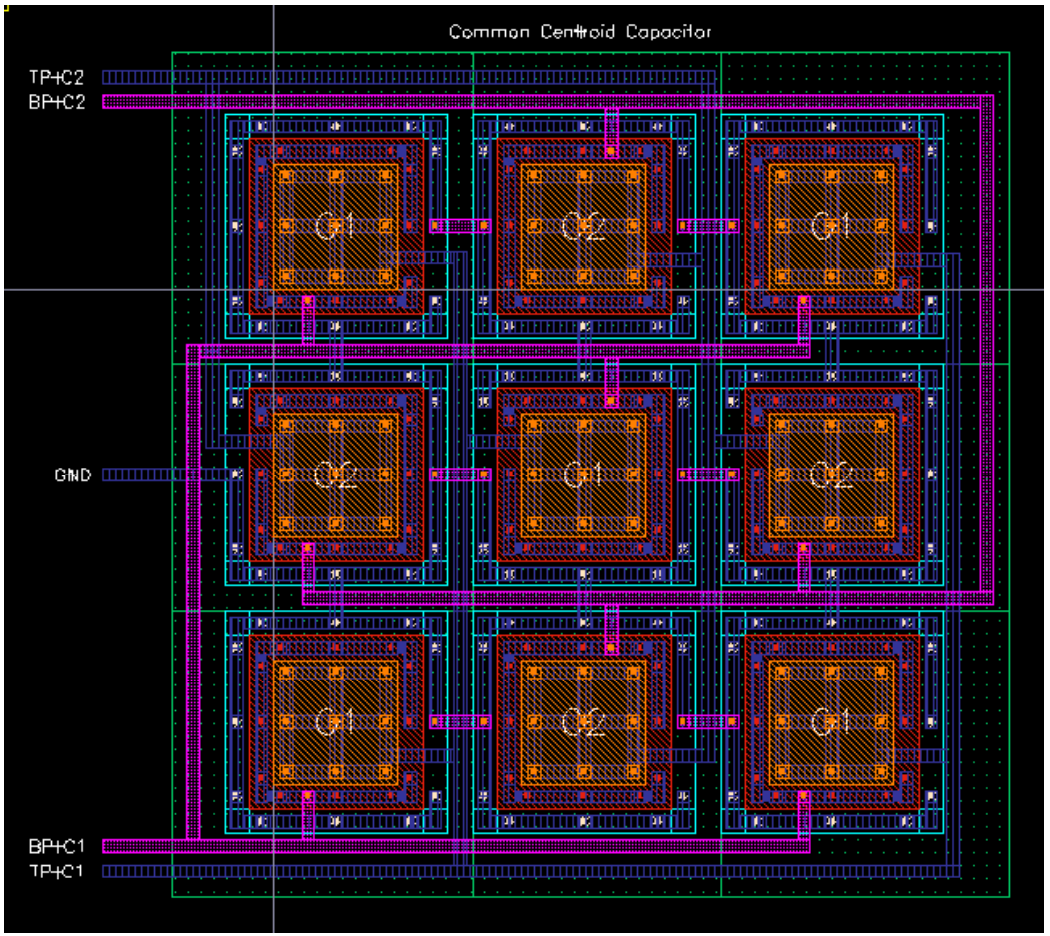
Then maximize the area and minimize the perimeter ➔ use squares!!!

C_A = capacitance per unit area

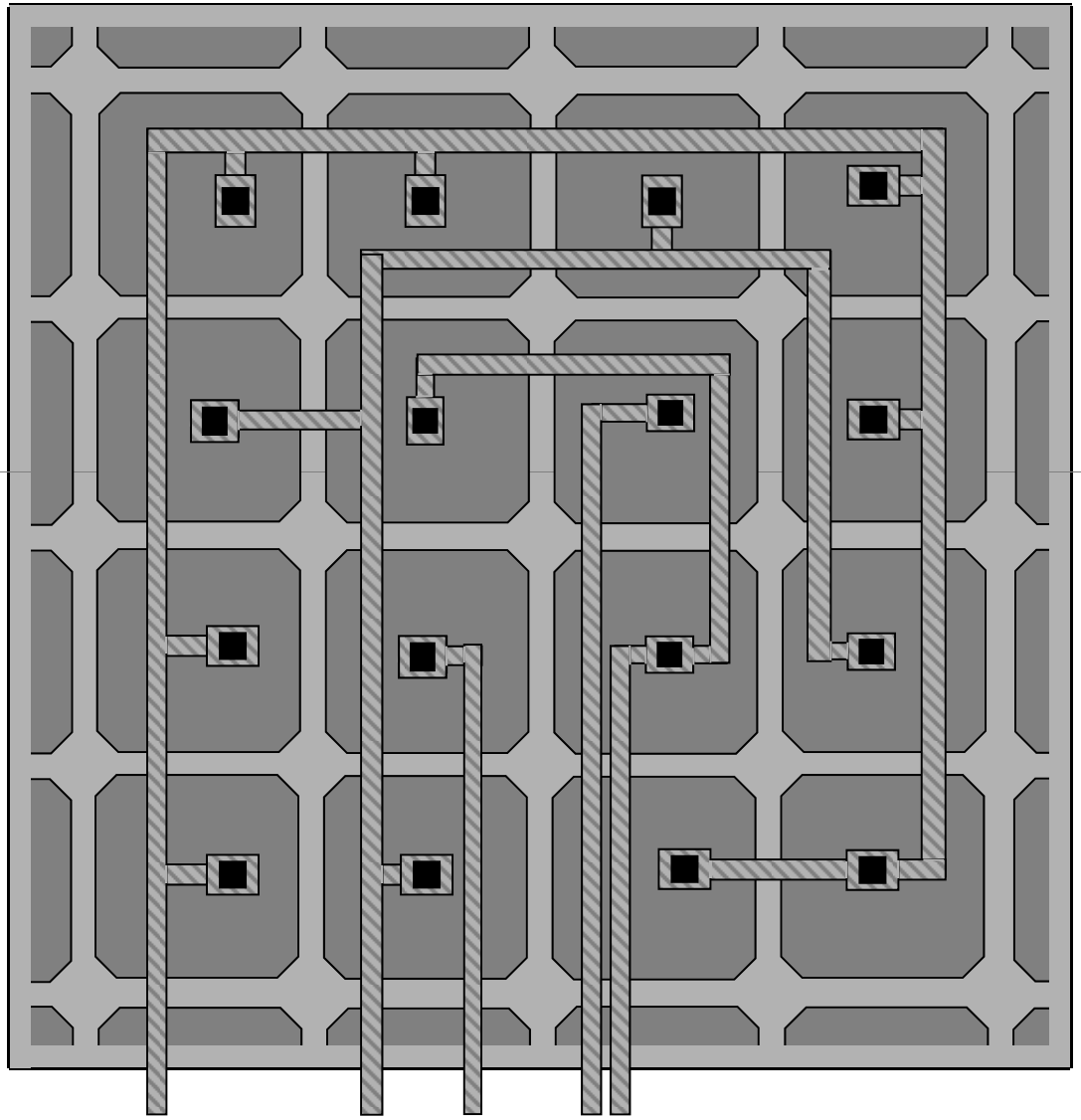


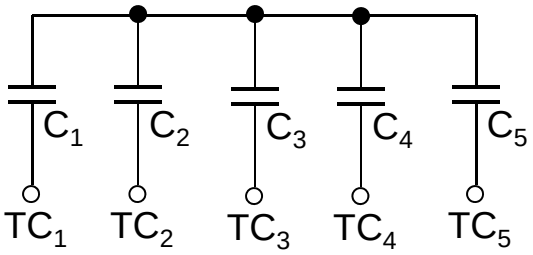
Common Centroid Capacitor Layout

- 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 Johns & Martin Book, page 112)



Common centroid structures



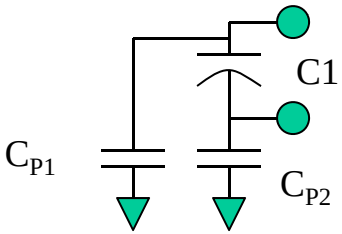


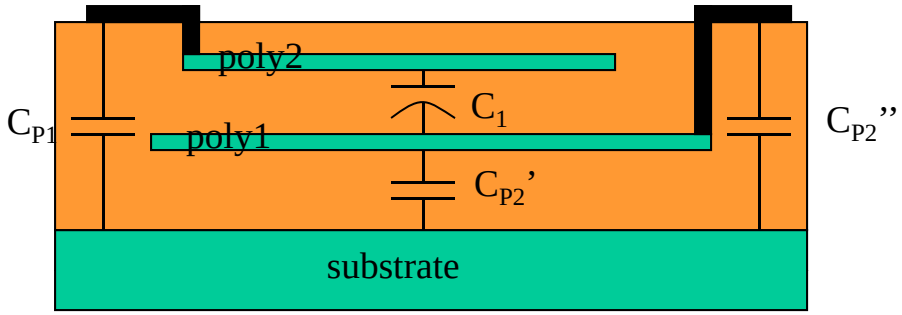
$C_2 = C_1$
$C_3 = 2C_1$
$C_4 = 4C_1$
$C_5 = 8C_1$

"Floating" Capacitors

Be aware of parasitic capacitors

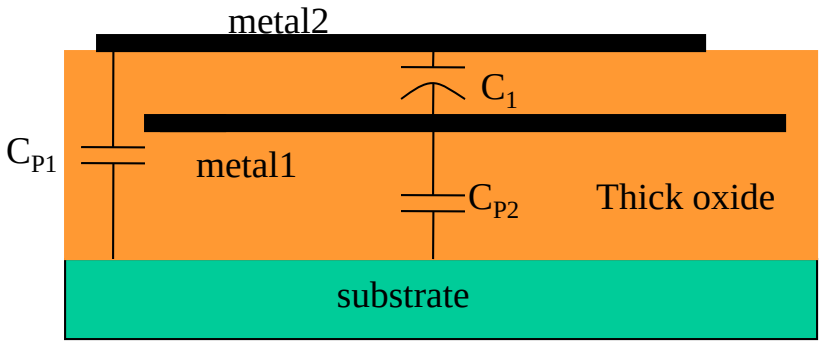
Polysilicon-Polysilicon: Bottom plate capacitance is comparable (10-30 %) with the poly-poly capacitance





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

➔**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).



CP2 is very small (1-5 % of C1)

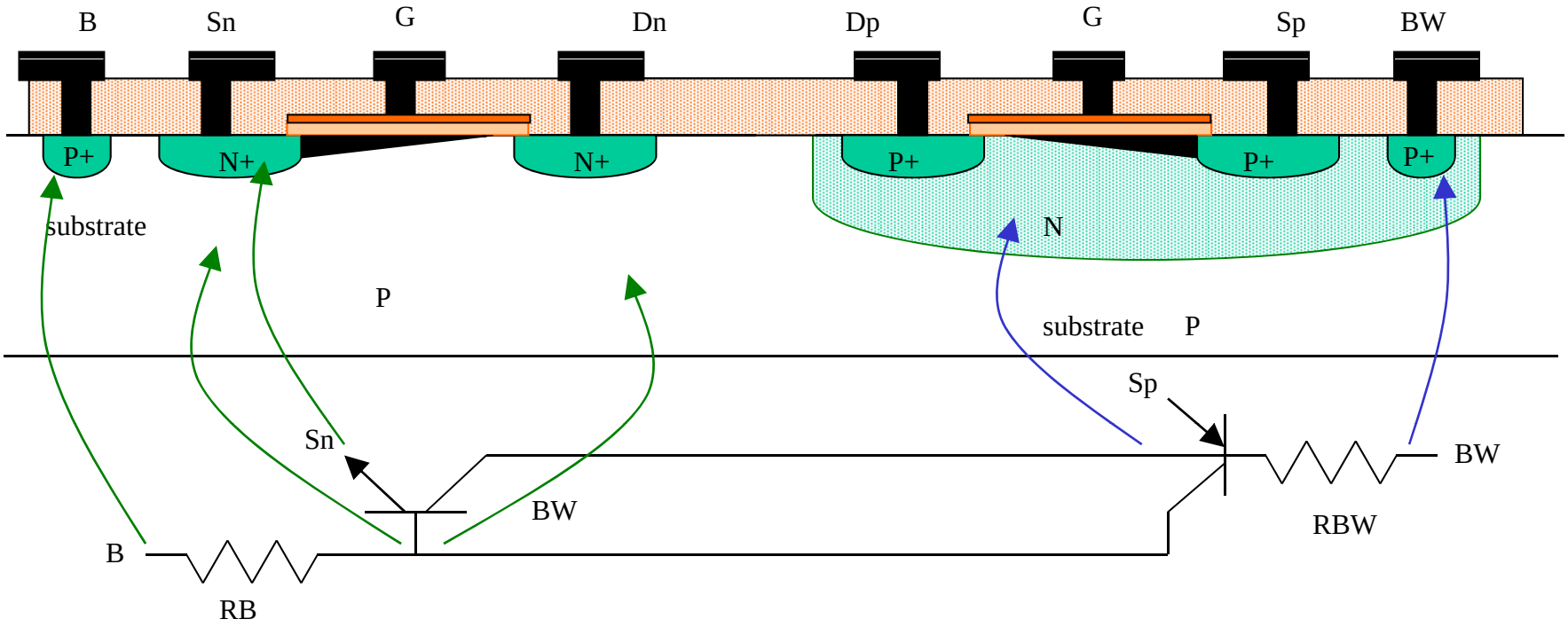
Typical Capacitance Process Data (See MOSIS webside for the AMI 0.6 CMOS process)

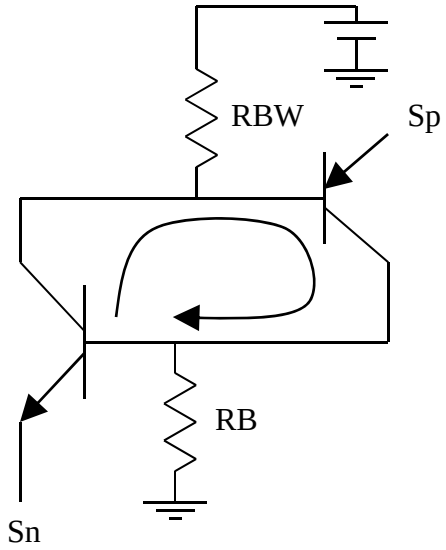
Capacitance	N+Actv	P+Actv	Poly	Poly 2	Mtl 1	Mtl 2	UNITS
Area (substrate)	292	290	35		20	13	aF/μm^2
Area (N+active)			1091	684	49	26	aF/μm^2
Area (P+active)			1072	677			aF/μm^2
Area (poly)				599	45	23	aF/μm^2
Area (poly2)			**900**		45		aF/μm^2
Area (metal1)						42	aF/μm^2
Fringe (substrate)	80	170			36	25	aF/μm
Fringe (poly)					59	39	aF/μm

a=10^{-18}, f=10^{-15}, p=10^{-12}, n=10^{-9} μ=10^{-6}, m=10^{-3}

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





POSITIVE FEEDBACK!

System may lock!

Reduce as much as possible RB and RBW: Place guard contacts everywhere

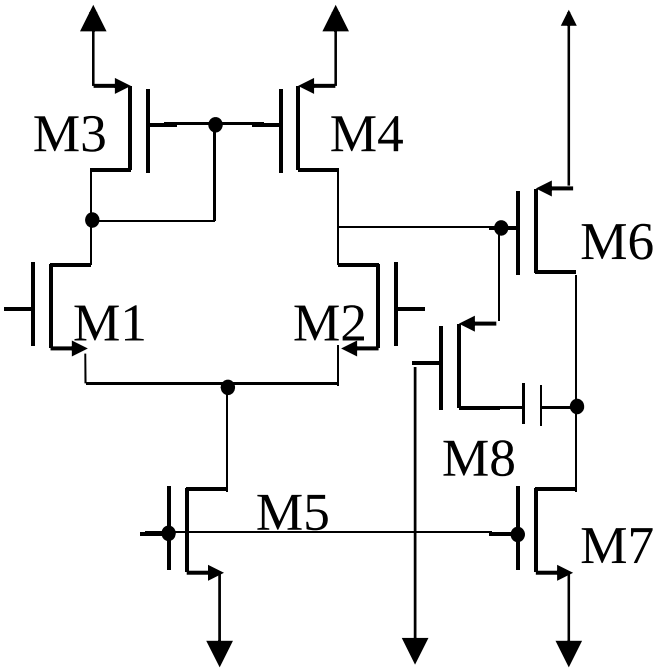
Be sure Base-Emitter voltages are such that the BJTs are off!

If possible, B and BW must be connected to the most negative and positive voltages, respectively!

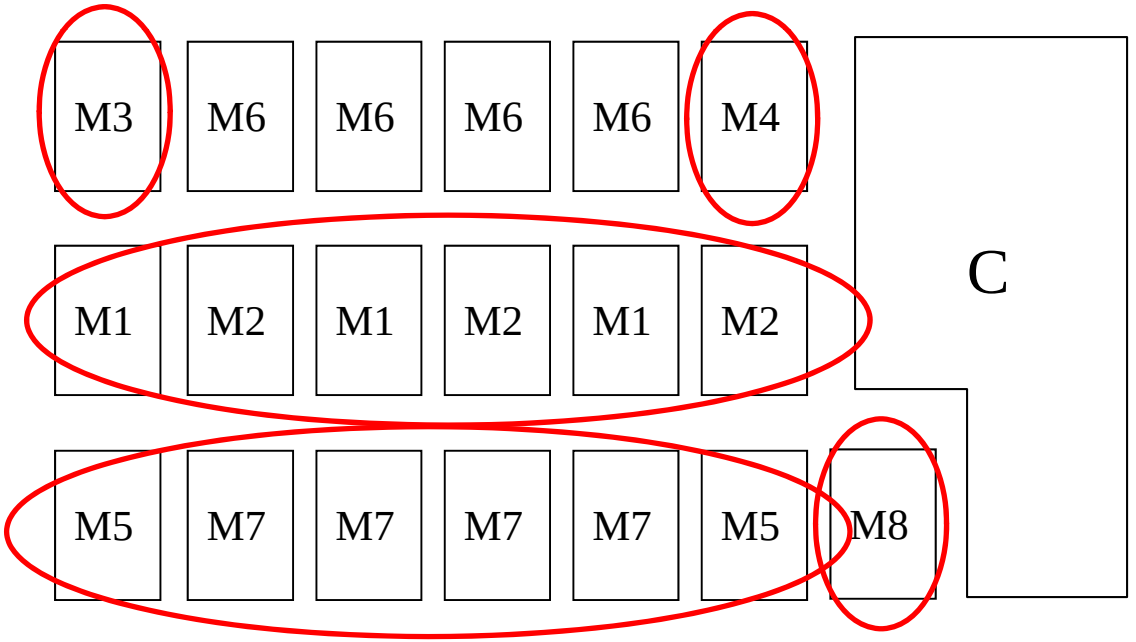
Analog Cell Layout

- 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)

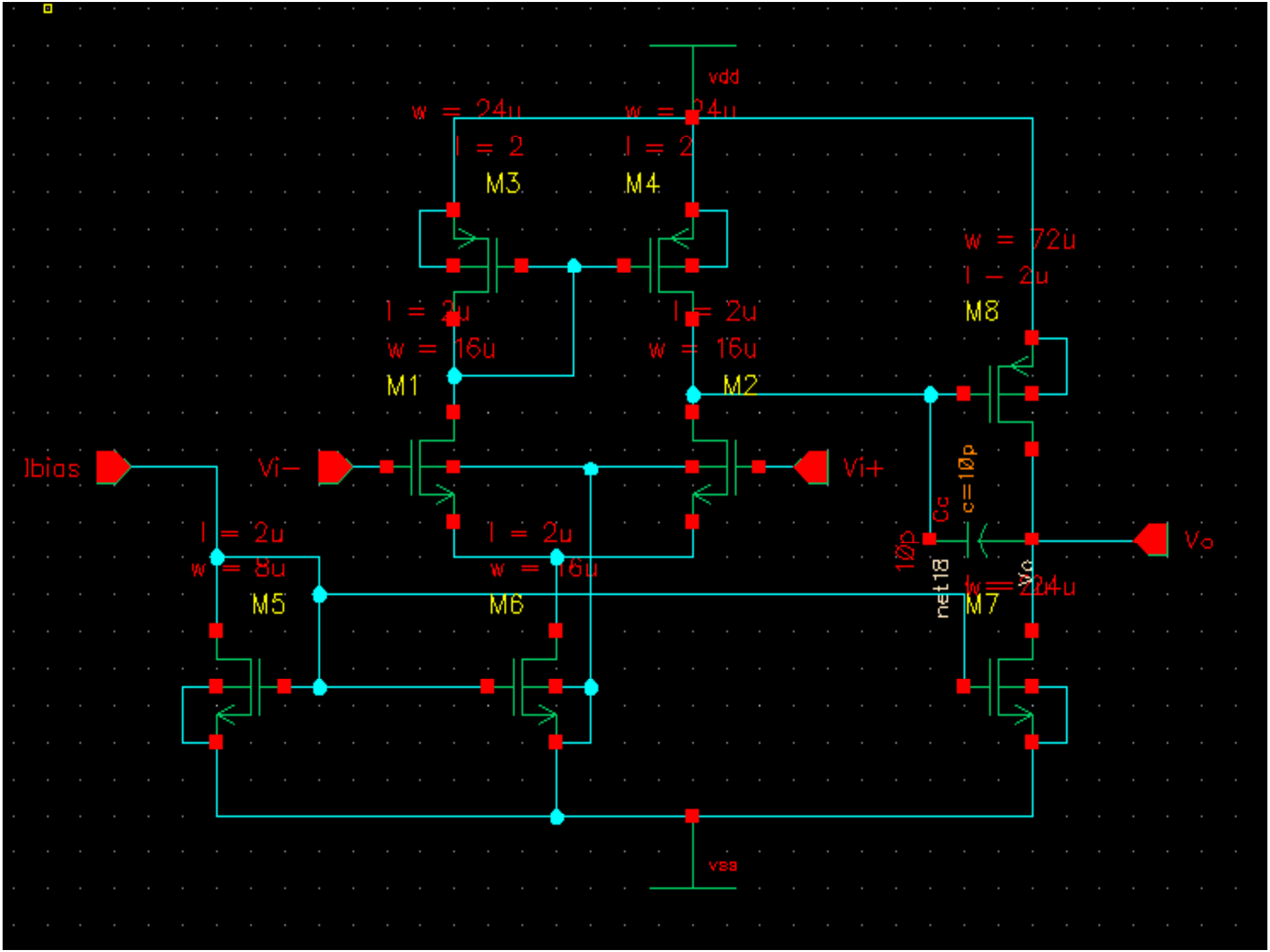
- M1 and M2 must match. Layout is interdigitized
- M3 and M4 must match. M6 must be wider by 4*M3
- M7 must be 2*M5
- Layout is an interconnection of 3 stacks; 2 for NMOS and 1 for PMOS
- Capacitor made by poly-poly

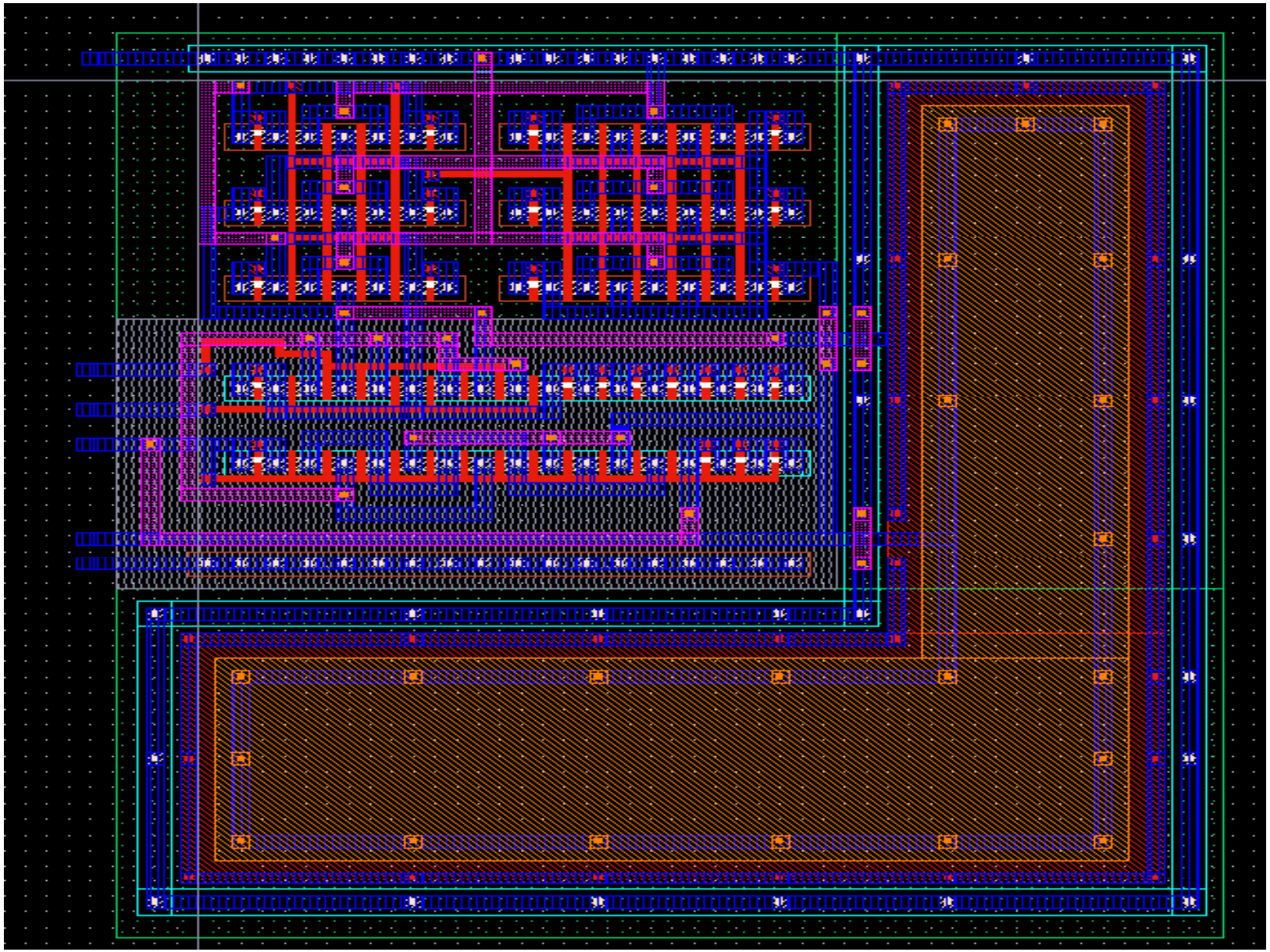


Not the best floorplan



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


vdd
w = 24u
l = 2
M3
w = 24u
l = 2
M4
w = 72u
l = 2u
M8
l = 2u
w = 16u
M1
l = 2u
w = 16u
M2
Ibias
Vi-
Vi+
Cc
c=10p
10p
Vo
l = 2u
w = 8u
M5
l = 2u
w = 16u
M6
net18
w = 24u
M7
vss



Layout (of something we should not do) example (cap related)

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

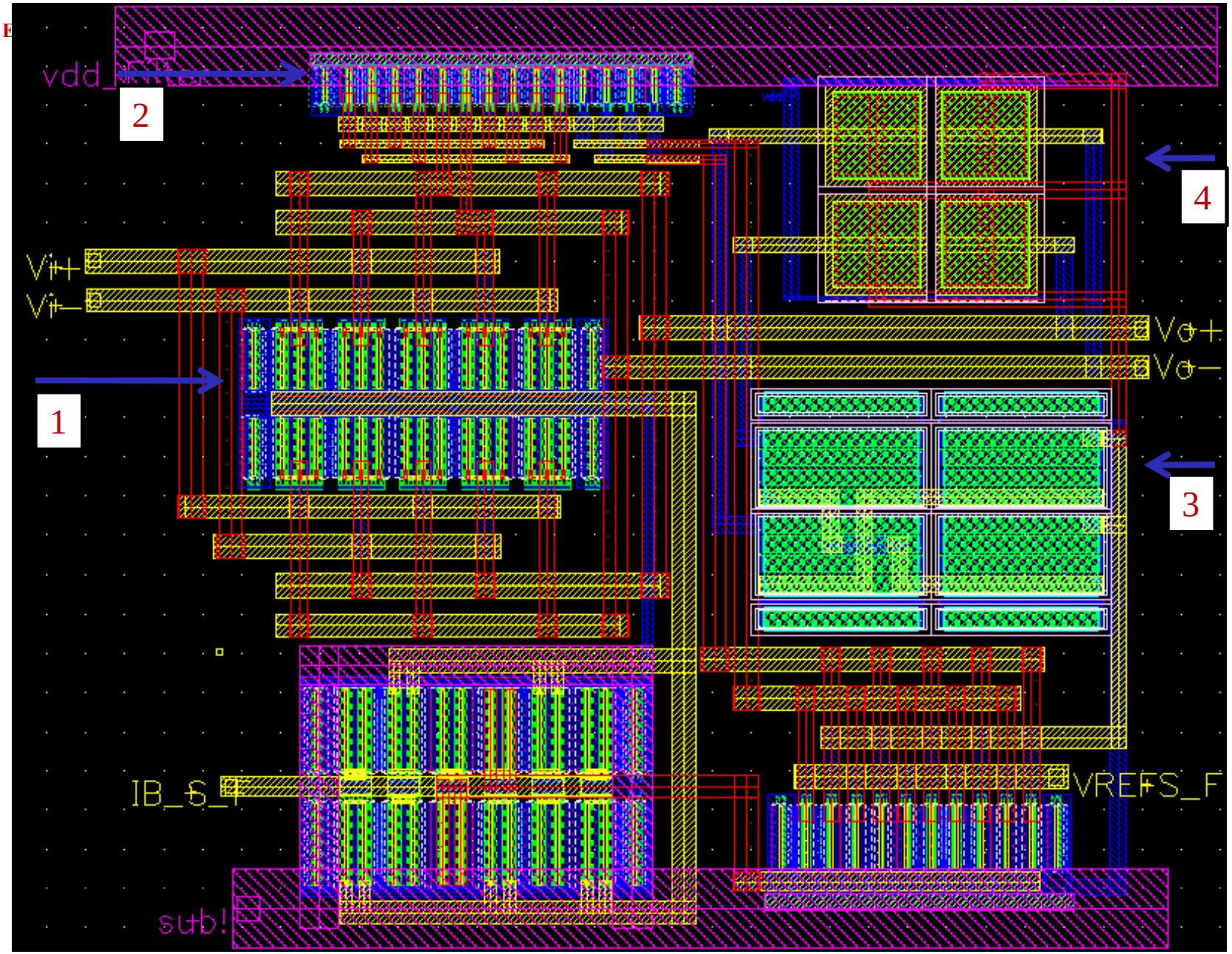


Figure: Layout of a single stage fully differential amplifier and its CMFB circuit.

1. I/p NMOS diff pair 2. PMOS (Interdigitated) 3. Resistors for V_{CM} 4.Capacitors (Common centroid)

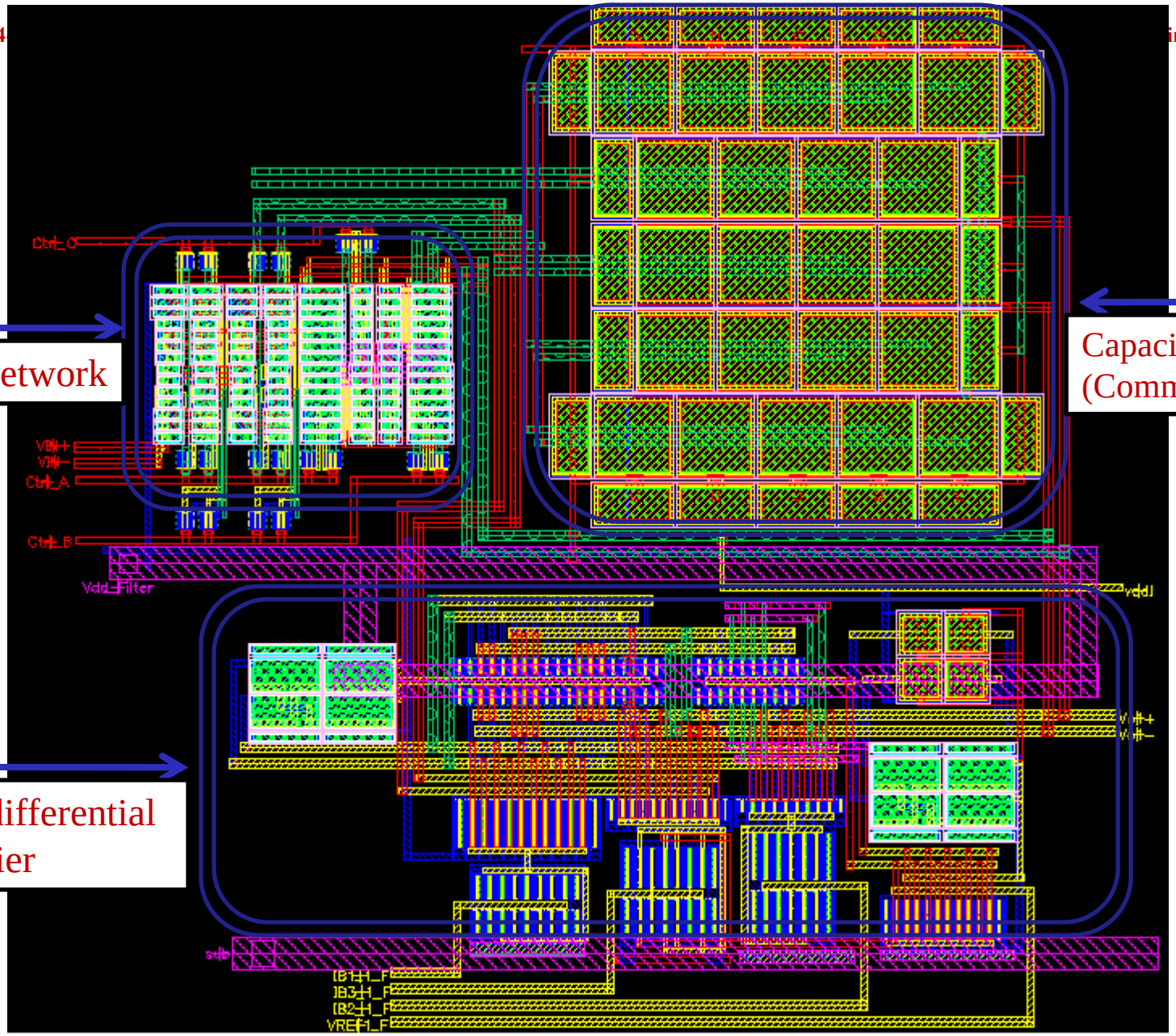
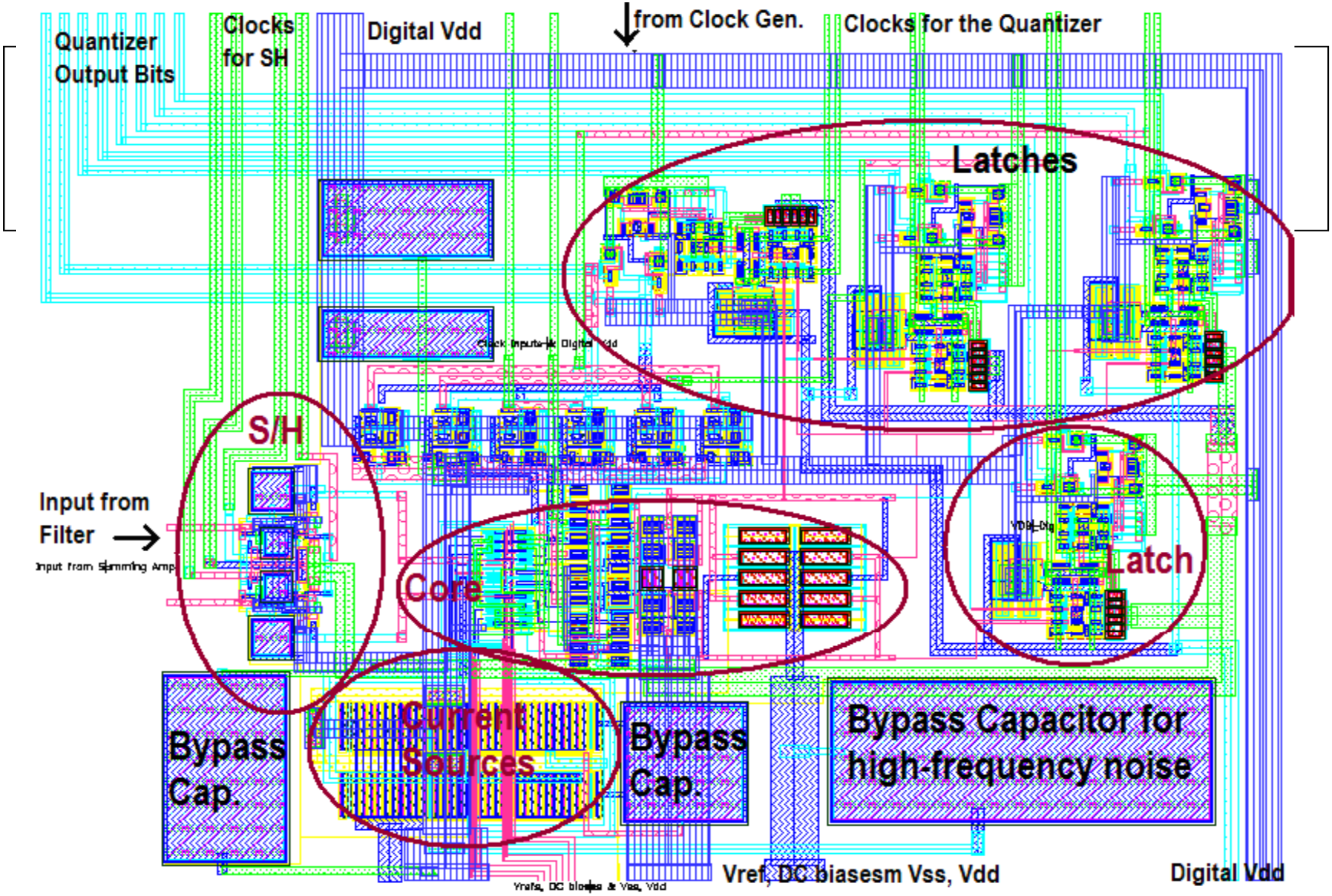
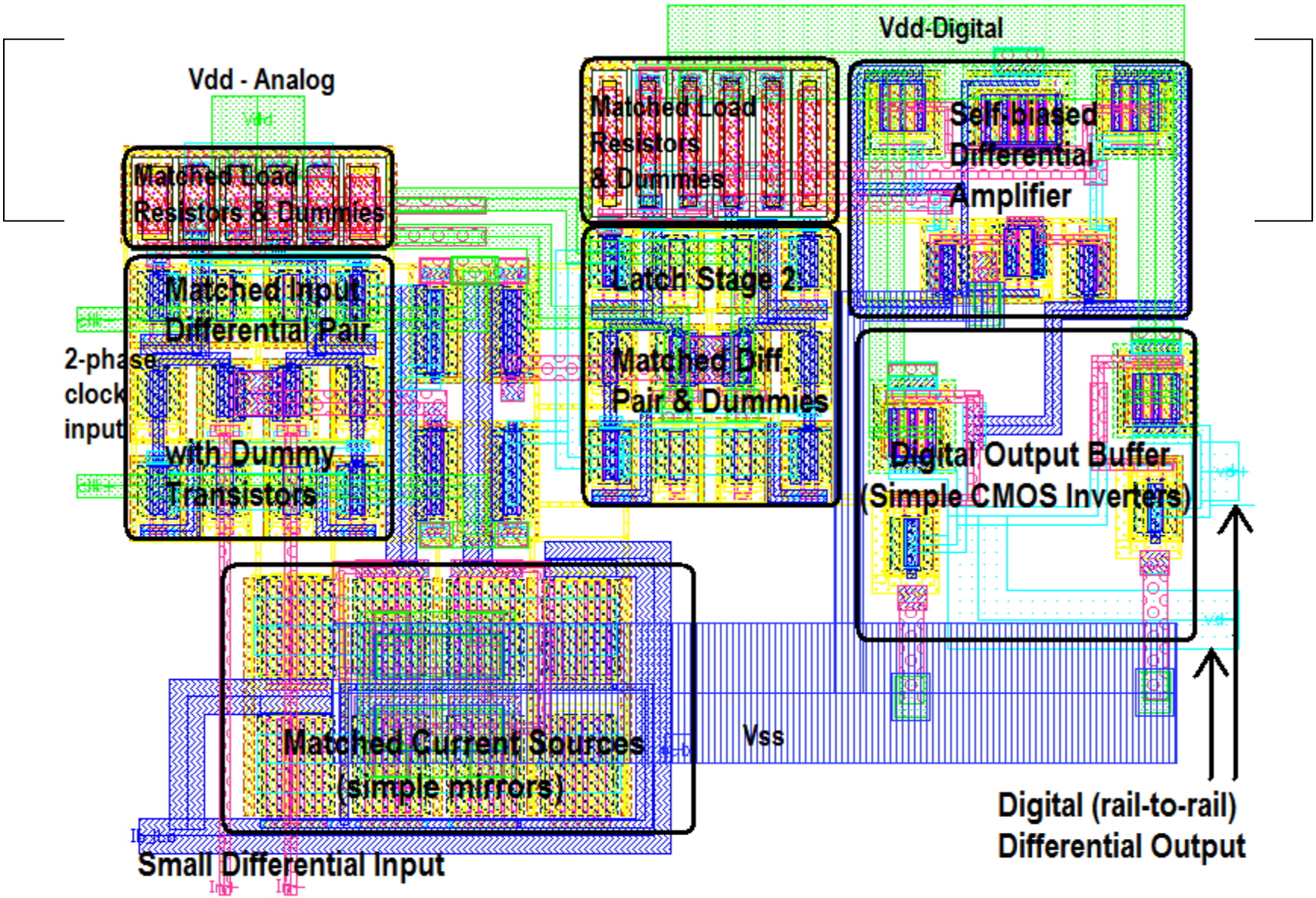


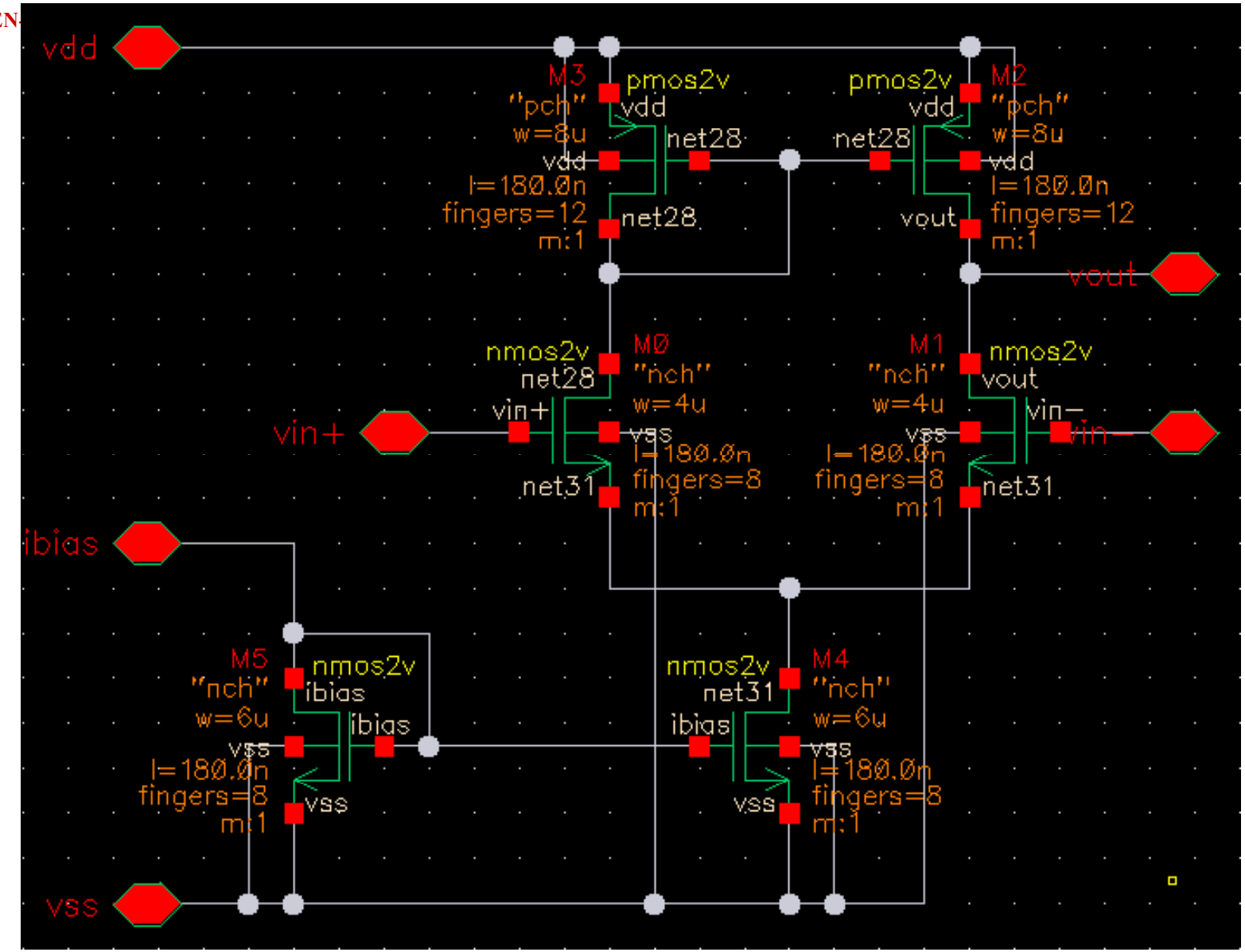
Figure: Layout of a second order Active RC low-pass Filter (Bi-quad)



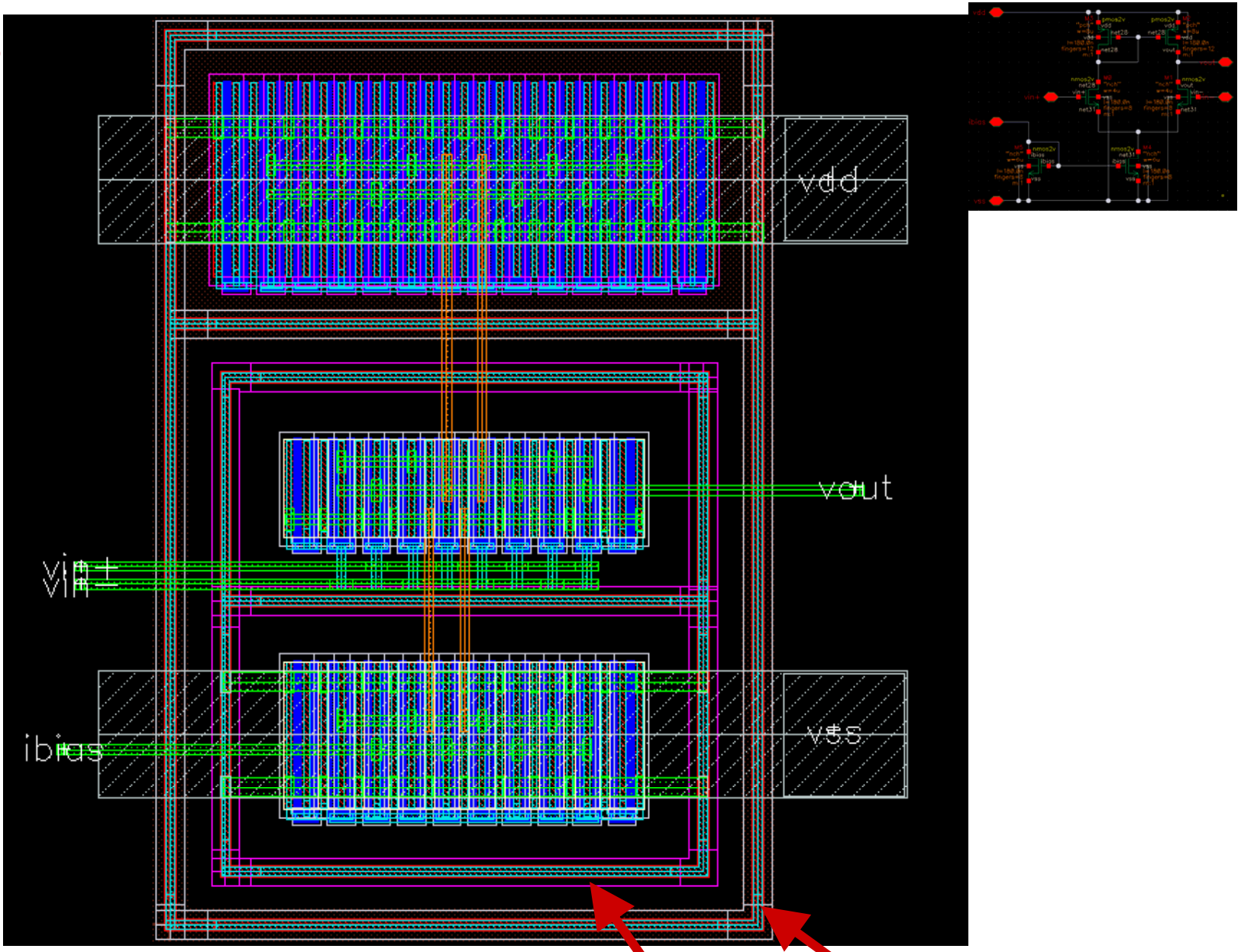
- **3-bit quantizer in Jazz 0.18µm CMOS technology**
- S/H: sample-and-hold circuit that is used to sample the continuous-input signal
- Core: contains matched differential pairs and resistors to create accurate reference levels for the analog-to-digital conversion
- Latches: store the output bits; provide interface to digital circuitry with rail-to-rail voltage levels



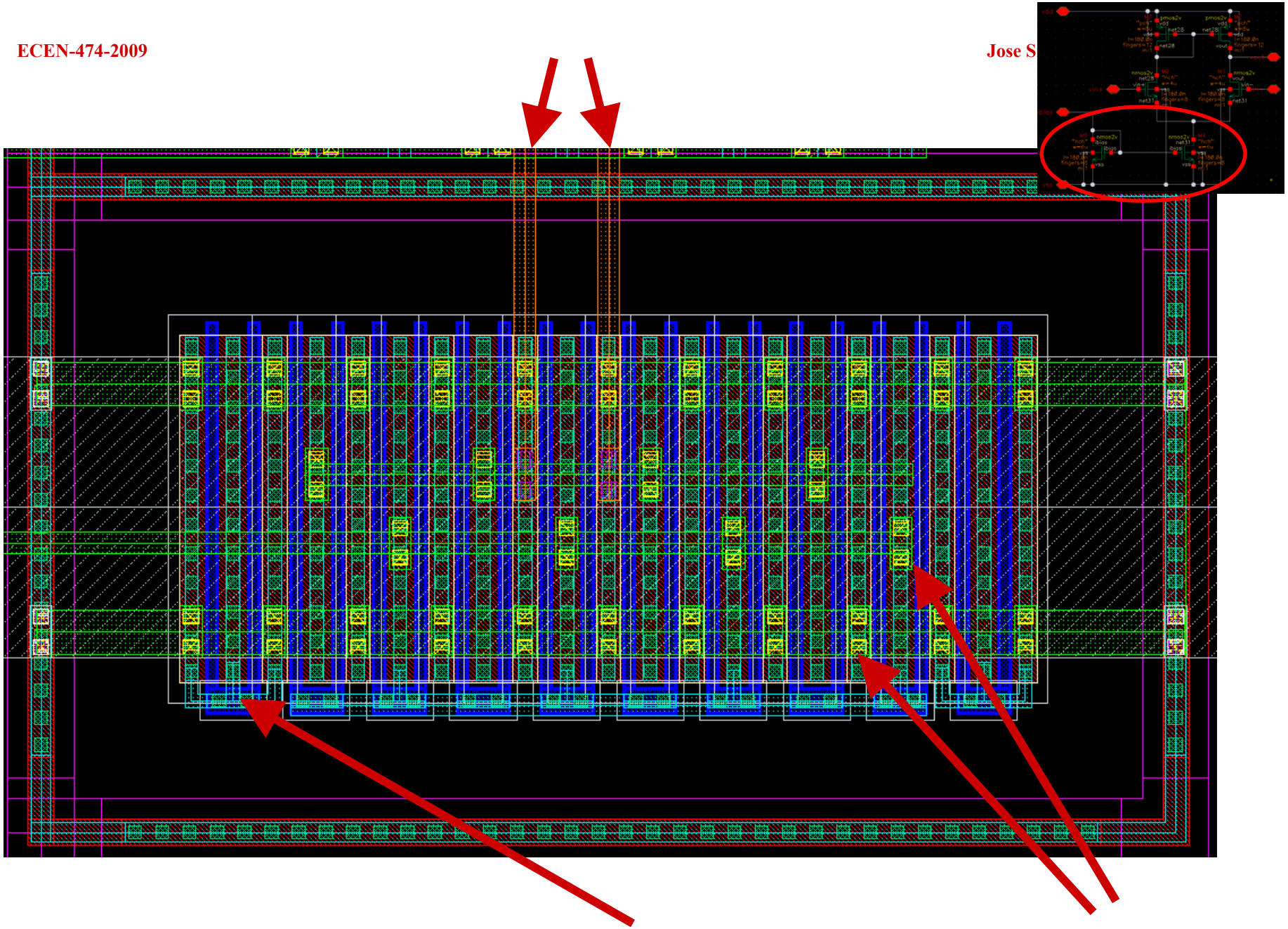
- **High-speed D-Flip-Flop in Jazz 0.18µm CMOS technology**
- Resolves a small differential input with $10mV < V_{p-p} < 150mV$ in less than 360ps
- Provides digital output (differential, rail-to-rail) clocked at 400MHz
- The sensitive input stage (1st differential pair) has a separate “analog” supply line to isolate it from the noise on the supply line caused by switching of digital circuitry



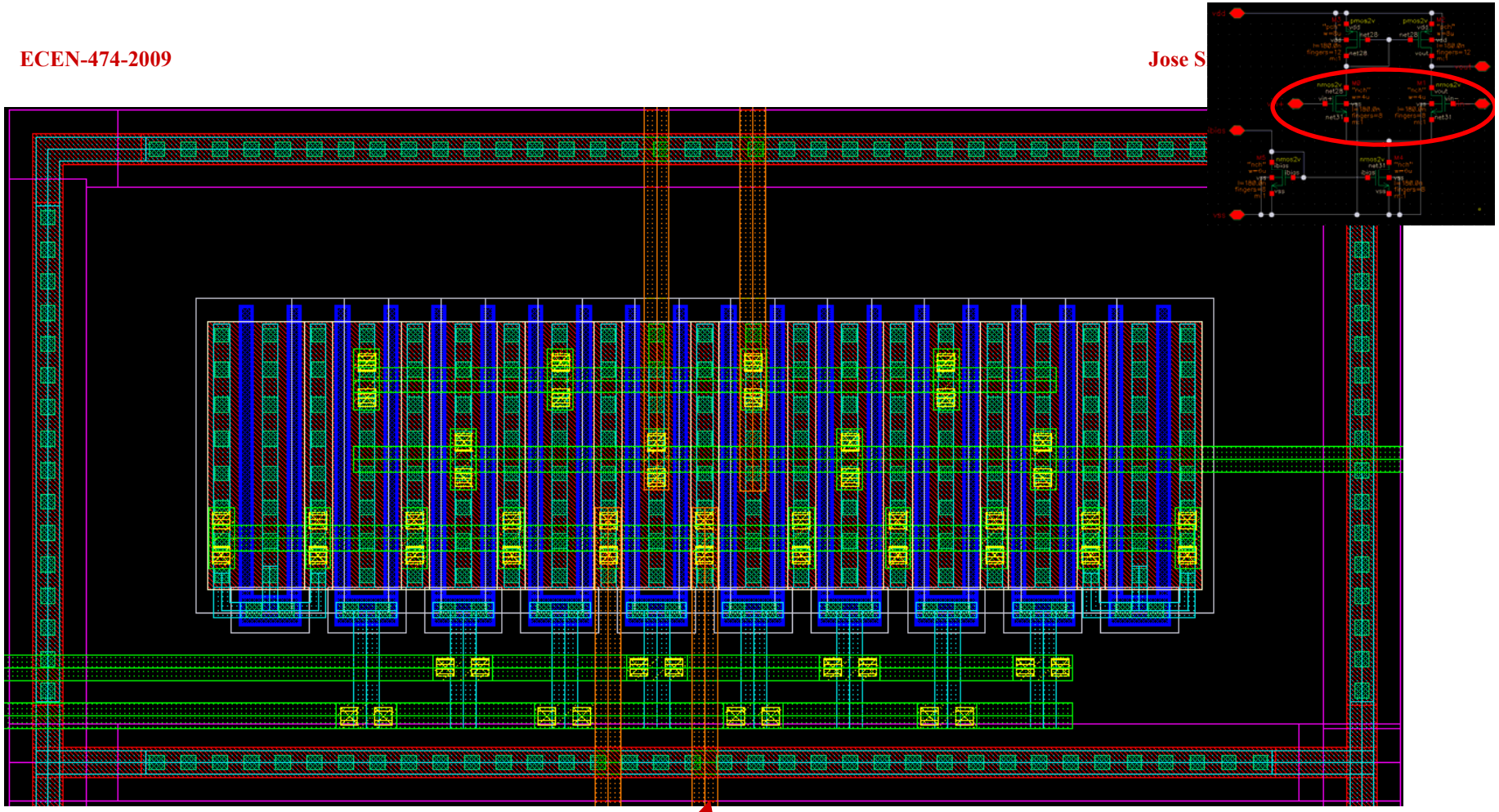
Design example (industrial quality): Simplest OTA



Overall amplifier: Have a look on the guard rings and additional well!

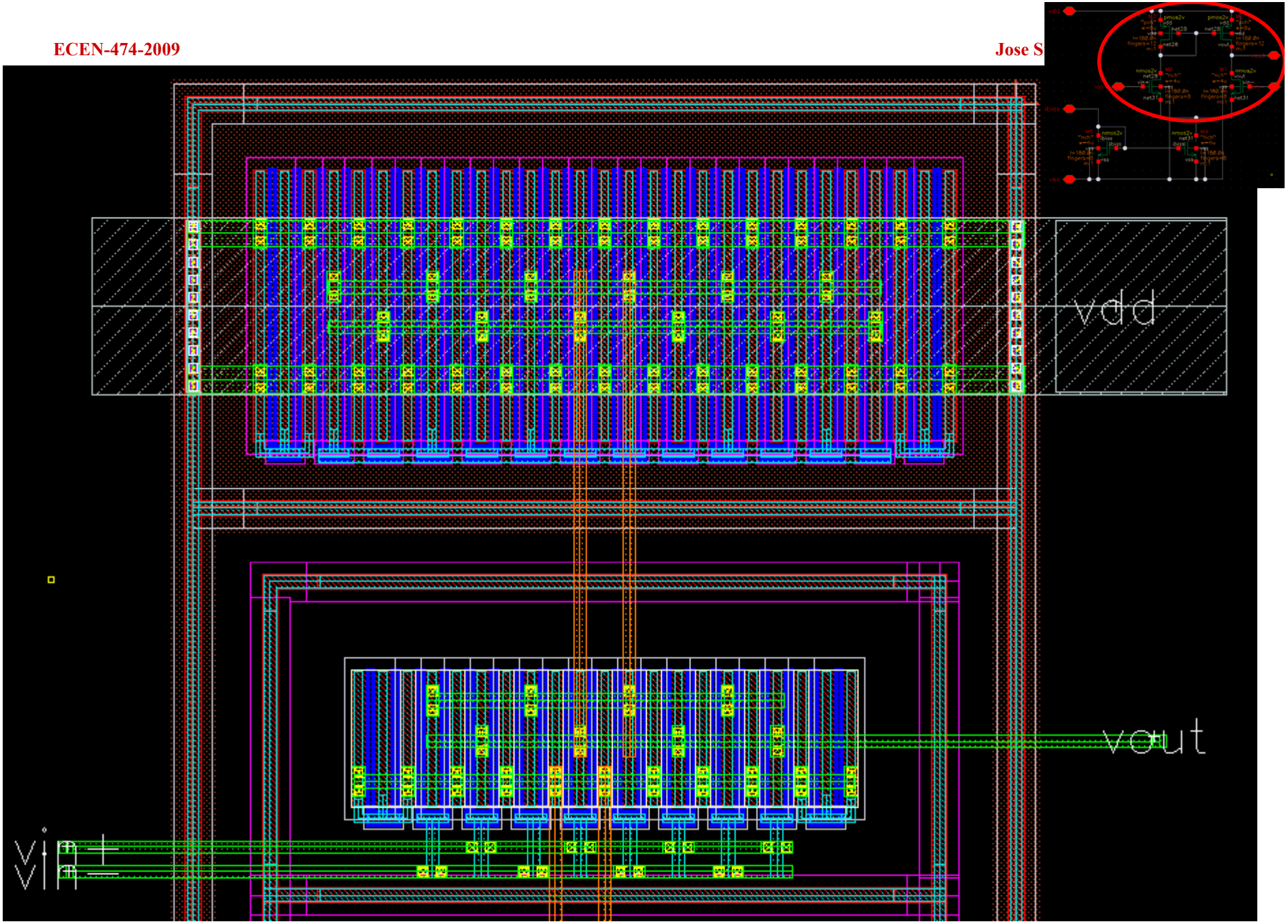


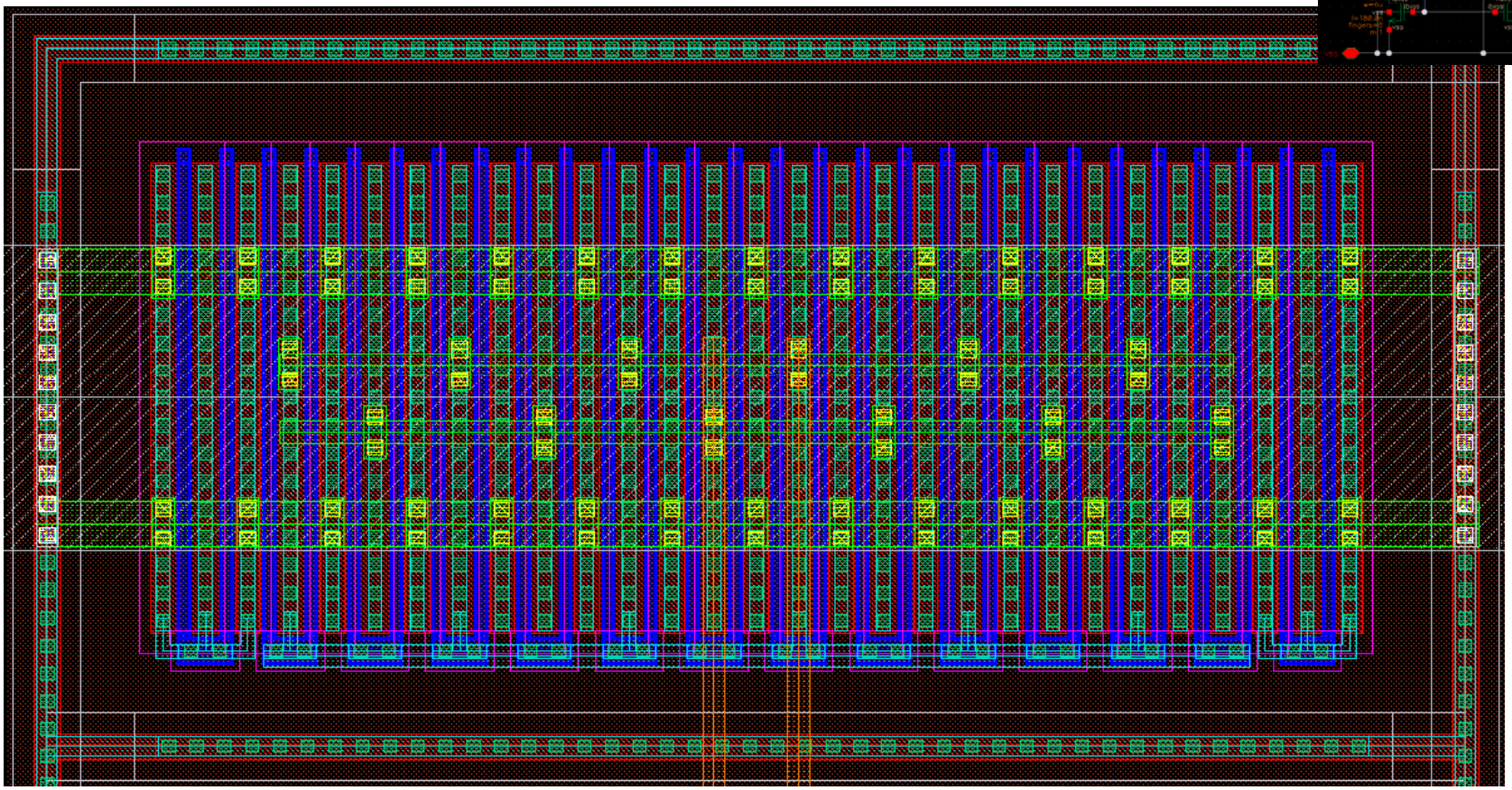
BIAS: you may be able to see the dummies, symmetry and S/D connections

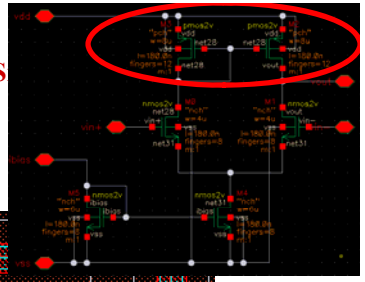


From downstairs

Differential pair


vdd
vout
vin





Details on the P-type current mirrors

Q-value of Spiral Inductors in CMOS Process

Most of the following slides were taken from

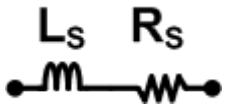
Seminar by: Park, Sang Wook

TAMU, 2003

What is Q?

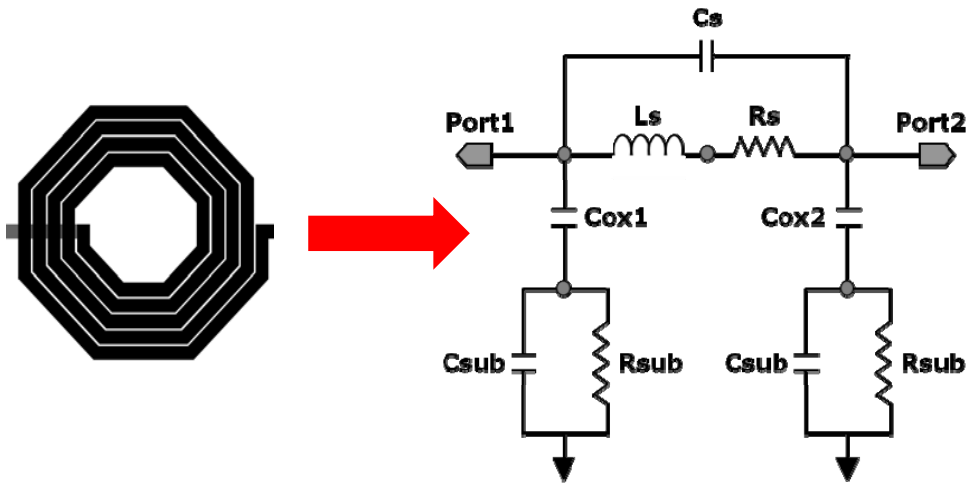
$$Q \equiv \omega \frac{\text{energy stored}}{\text{average power dissipated}}$$

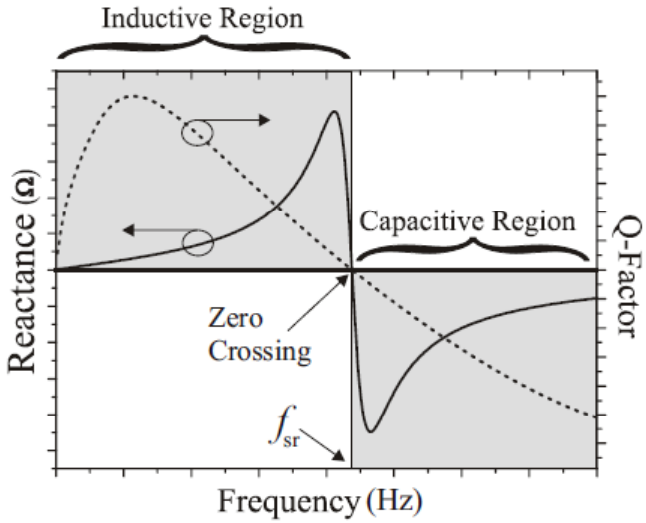
Simple Inductor Model:



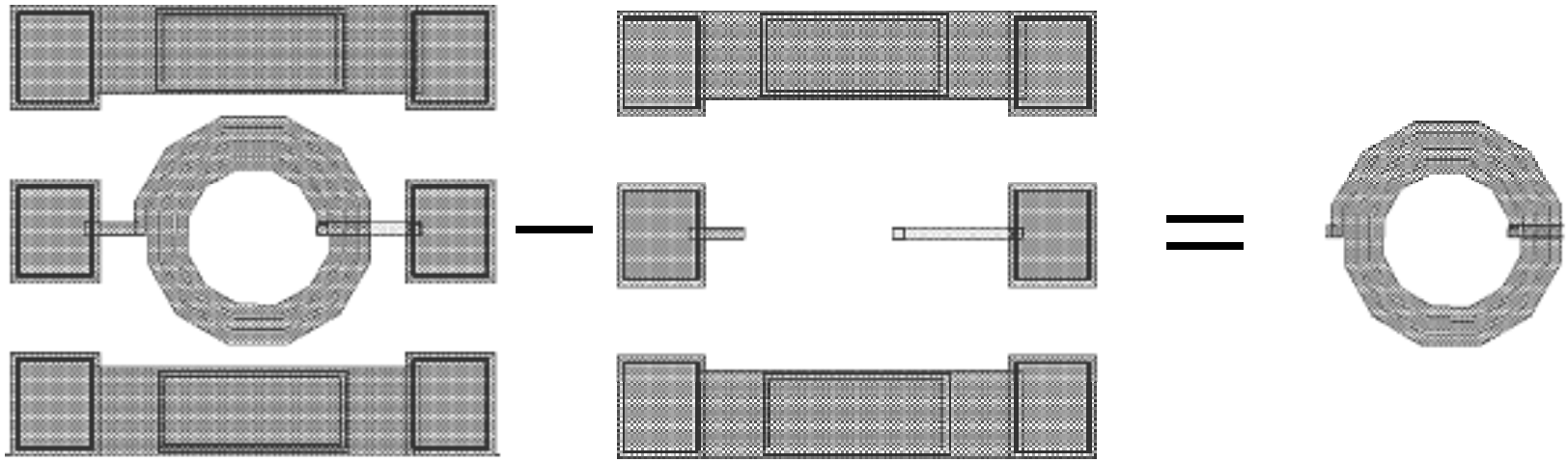
$$Q = \frac{\omega L_s}{R_s}$$

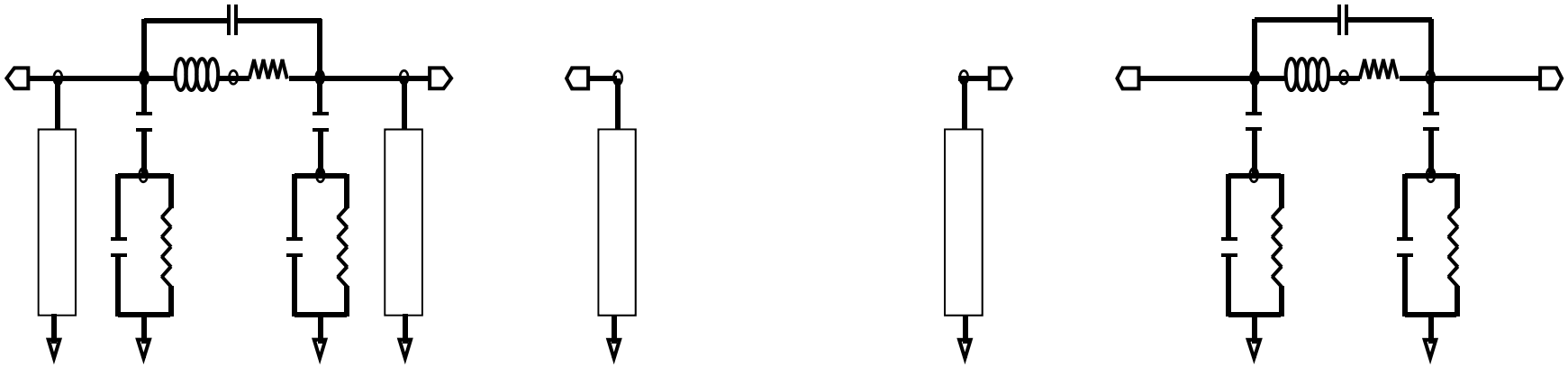
Integrated Spiral Inductor "Pi" Model



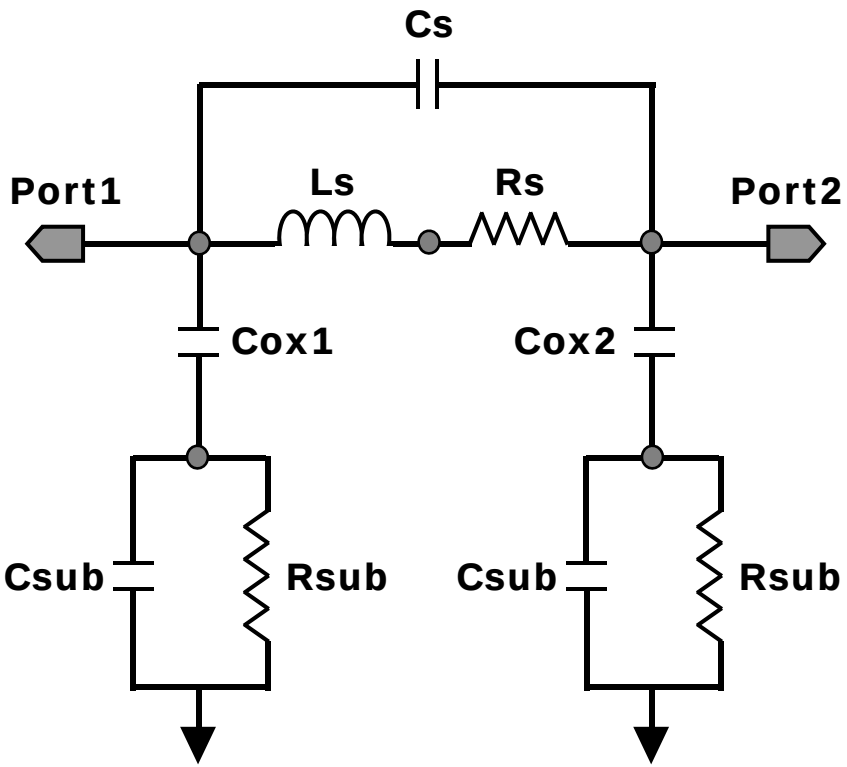


De-Embedding

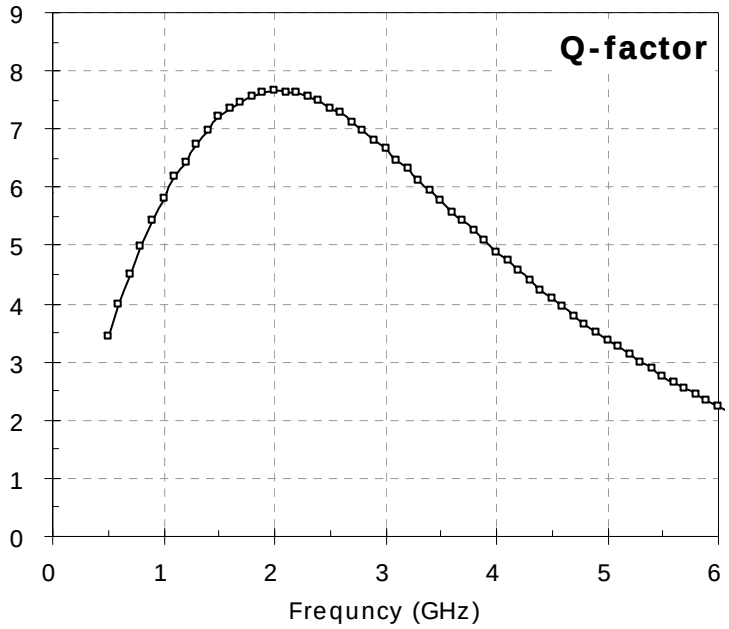




Equivalent Circuit & Calculation

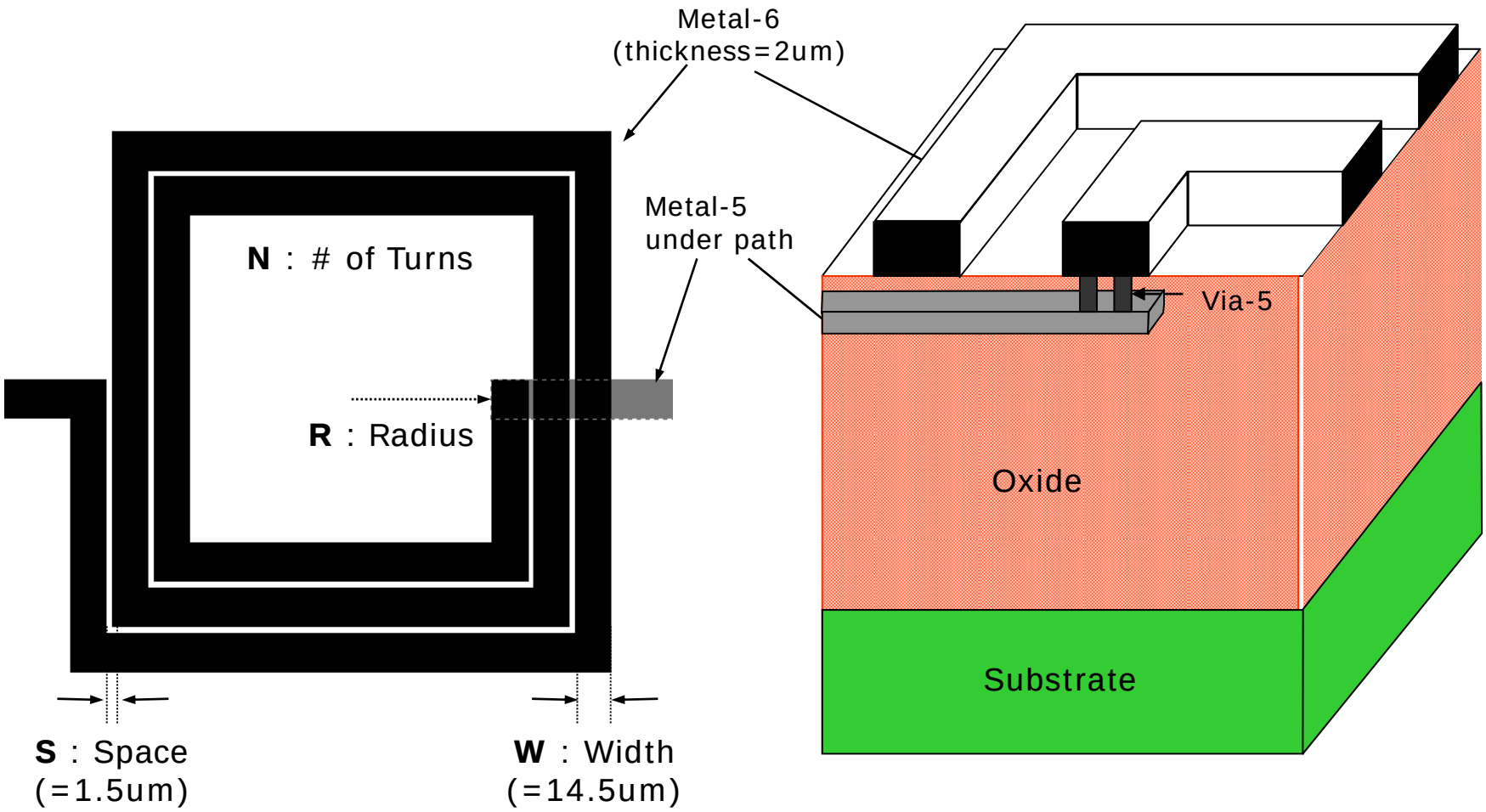


Equivalent Circuit



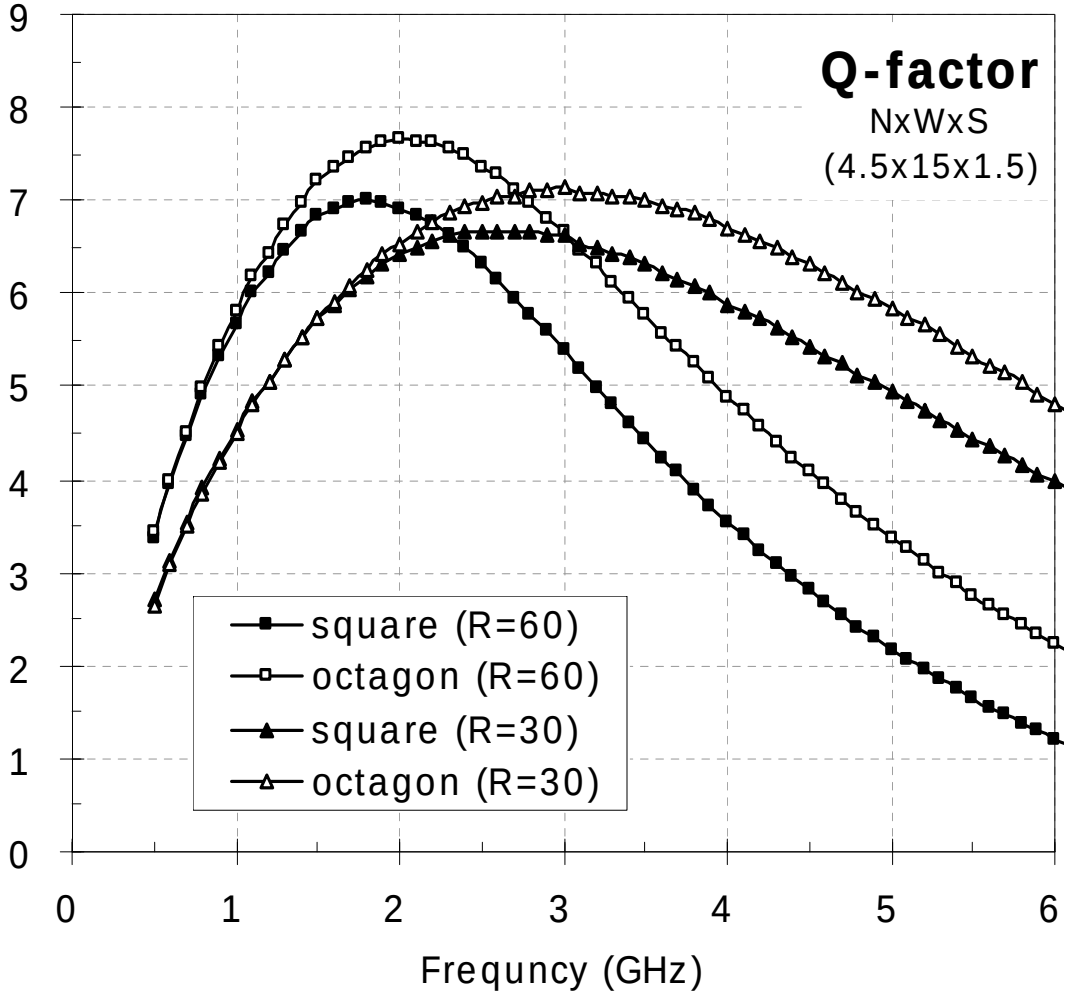
Parameter Calculation

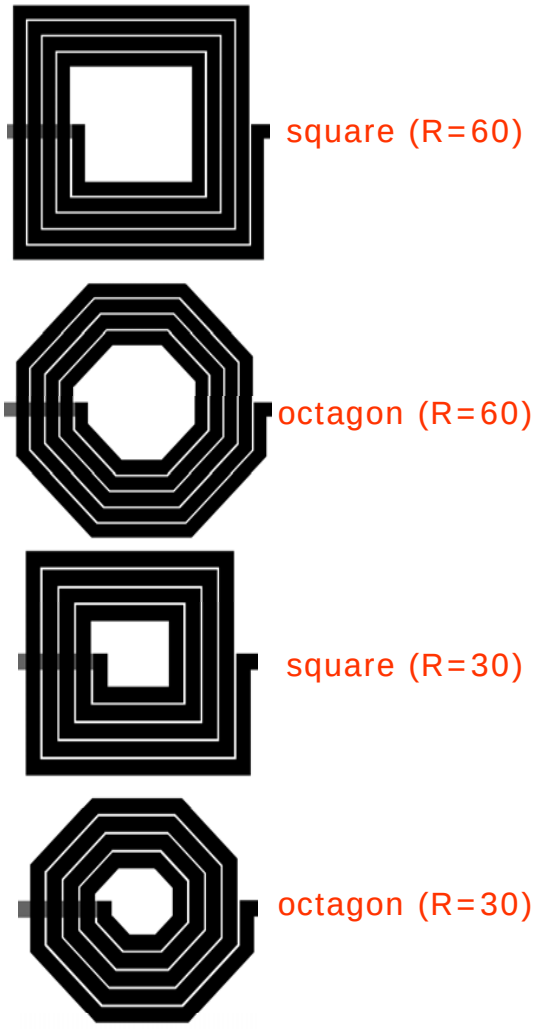
Layout & Structure



Layout Split 1

Shape & Radius

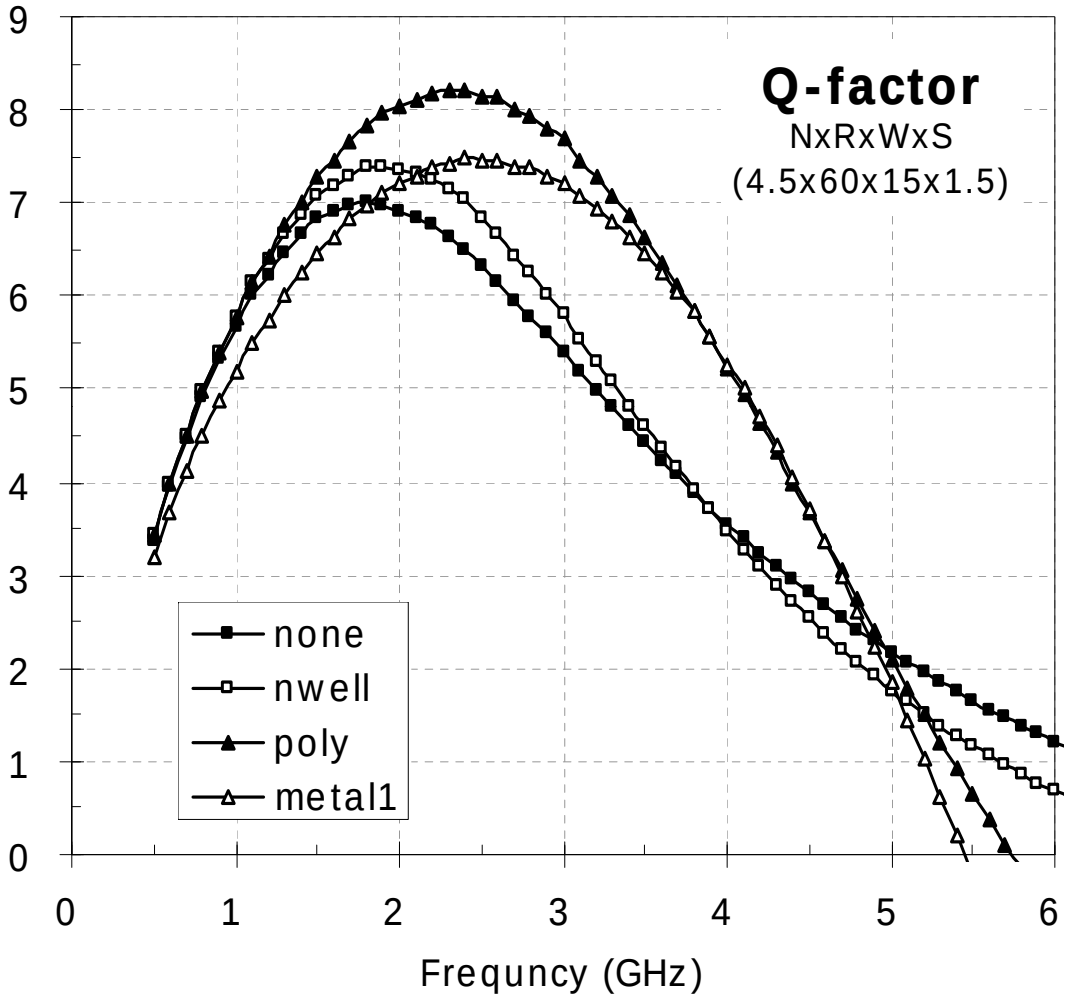


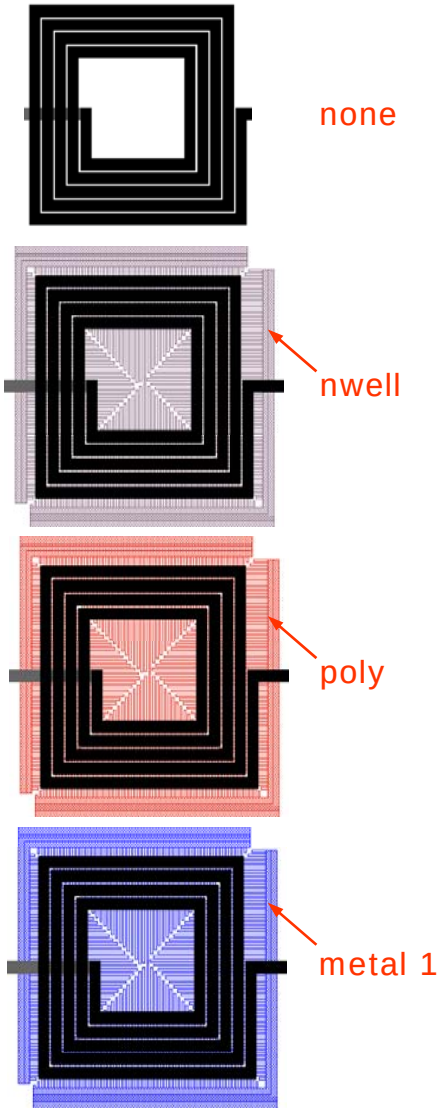


- **Shape : Octagon > Square**
- **Radius : 60 > 30**

Layout Split 2

PGS (Patterned Ground Shield) material

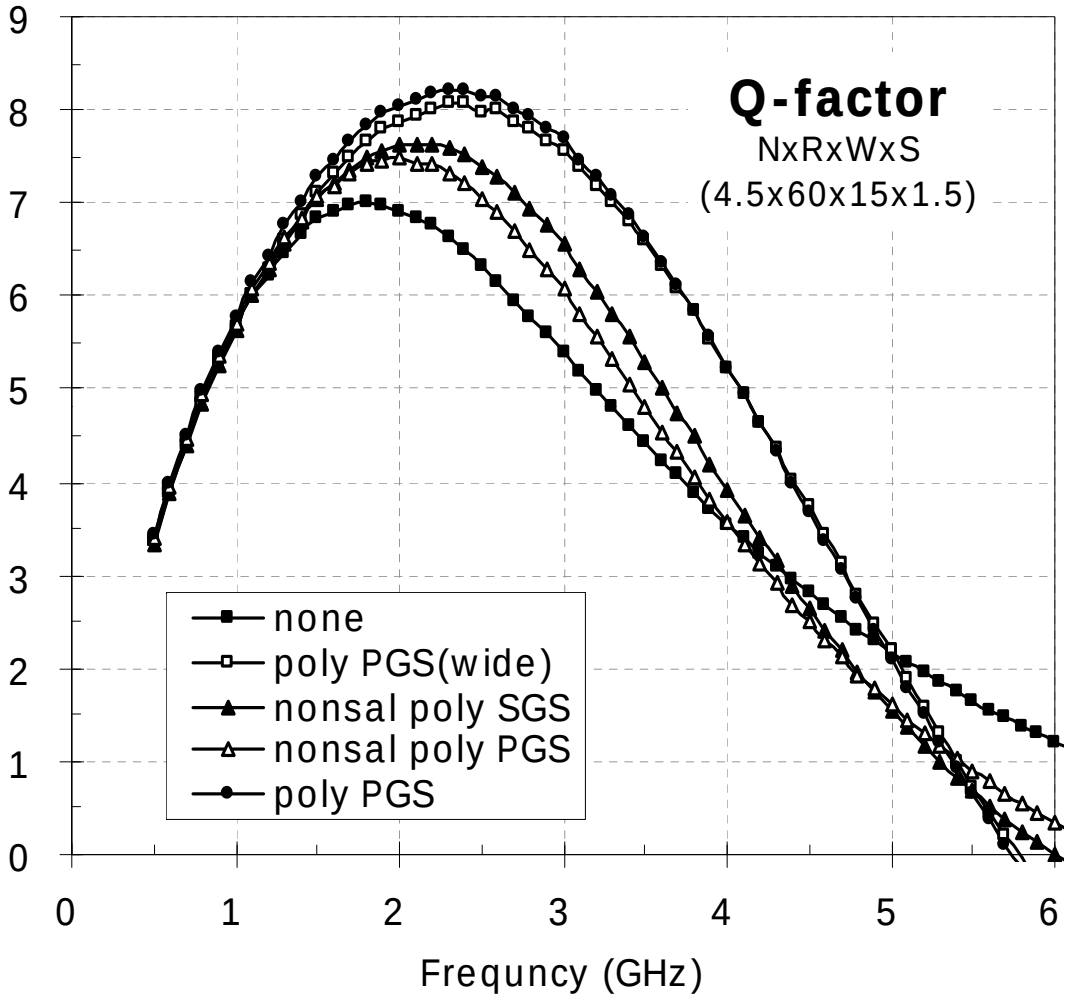


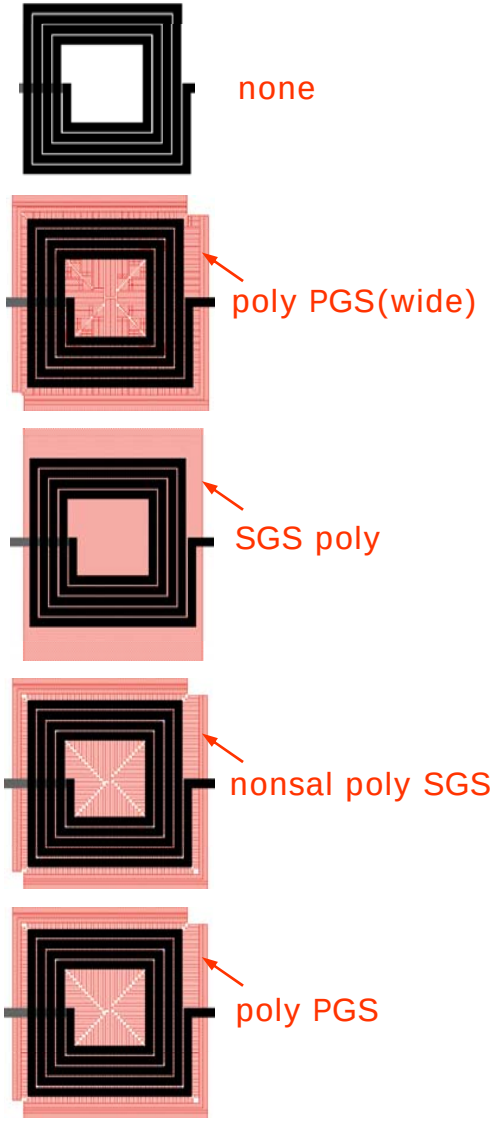


PGS : Poly > Nwell > none > Metal1

Layout Split 3

GS (Ground Shield) type

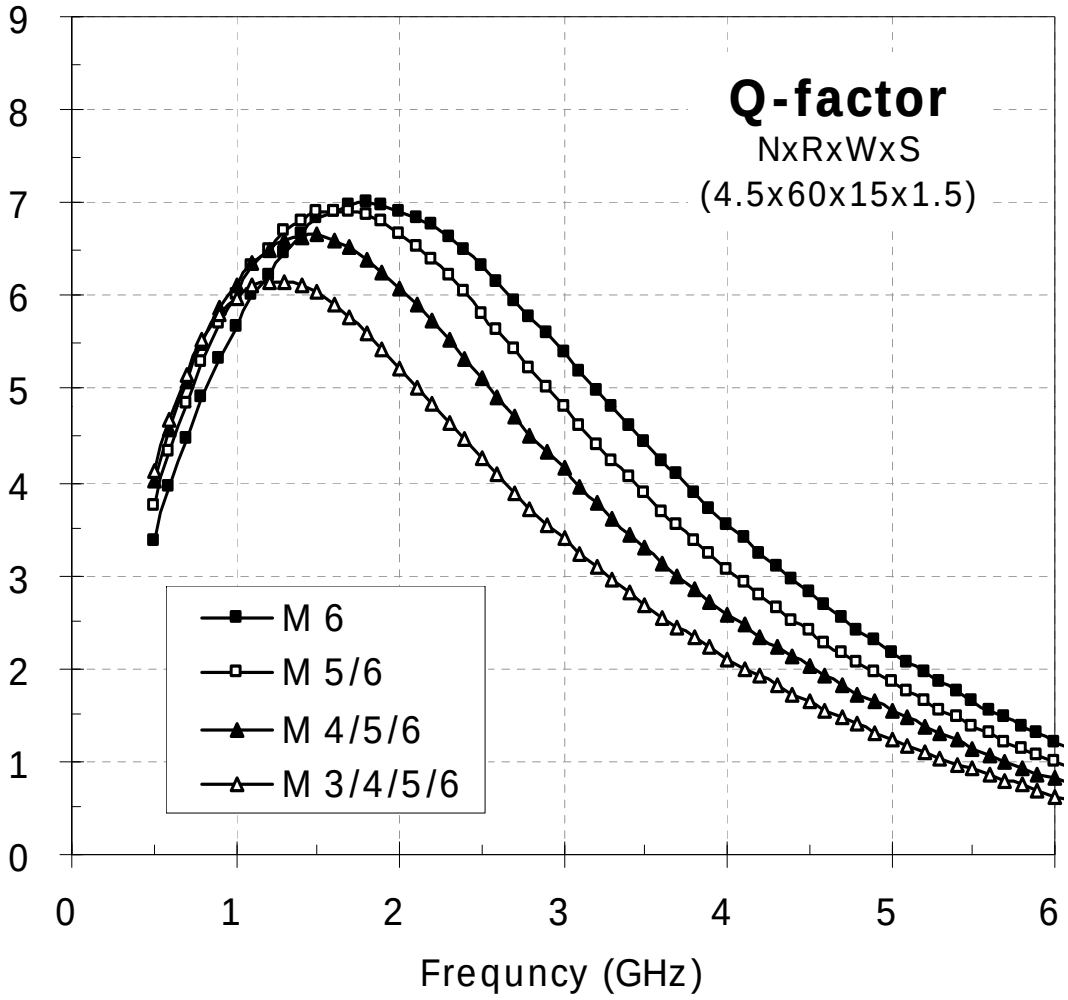


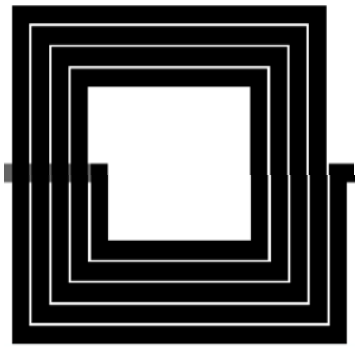


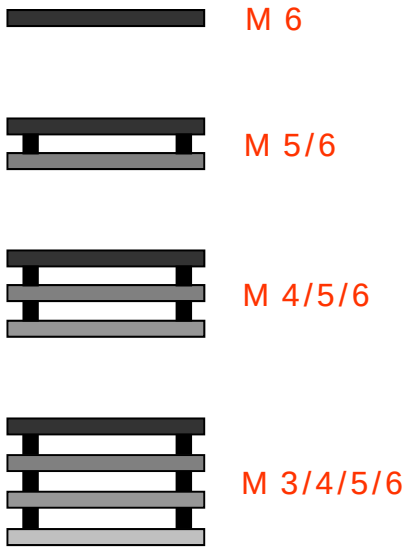
- **GS : Poly PGS > Poly(nonsal) SGS > Poly(nonsal) PGS > none**

Layout Split 4

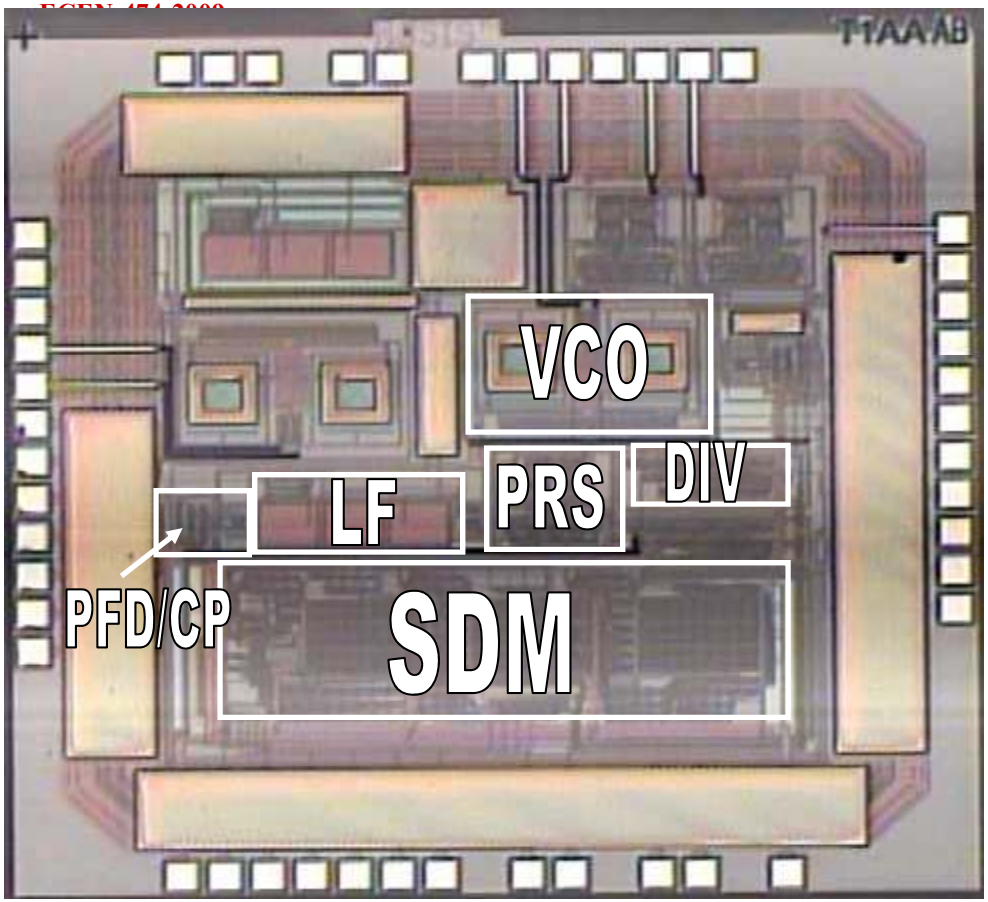
Metal Stack







- **Stack : M6 > M5/6 > M4/5/6 > M3/4/5/6**



Chip was fabricated in 0.35um CMOS through MOSIS.

Total area 2mm×2mm.

It includes the monolithic PLL, standalone prescaler, loop filter and VCO, etc.

The chip was packaged in 48-pin TPFQ.

Best student paper award: Radio Frequency Intl Conference 2003

IEEE-JSSC-June 2003

Keliu Shu[1], Edgar Sánchez-Sinencio[1], Jose Silva-Martinez[1], and Sherif H. K. Embabi[2]

[1]**Texas A&M University**

[2]**Texas Instruments**

Next Time

- Table-Based (g_m/I_D) Design Examples