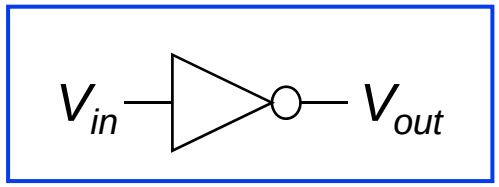
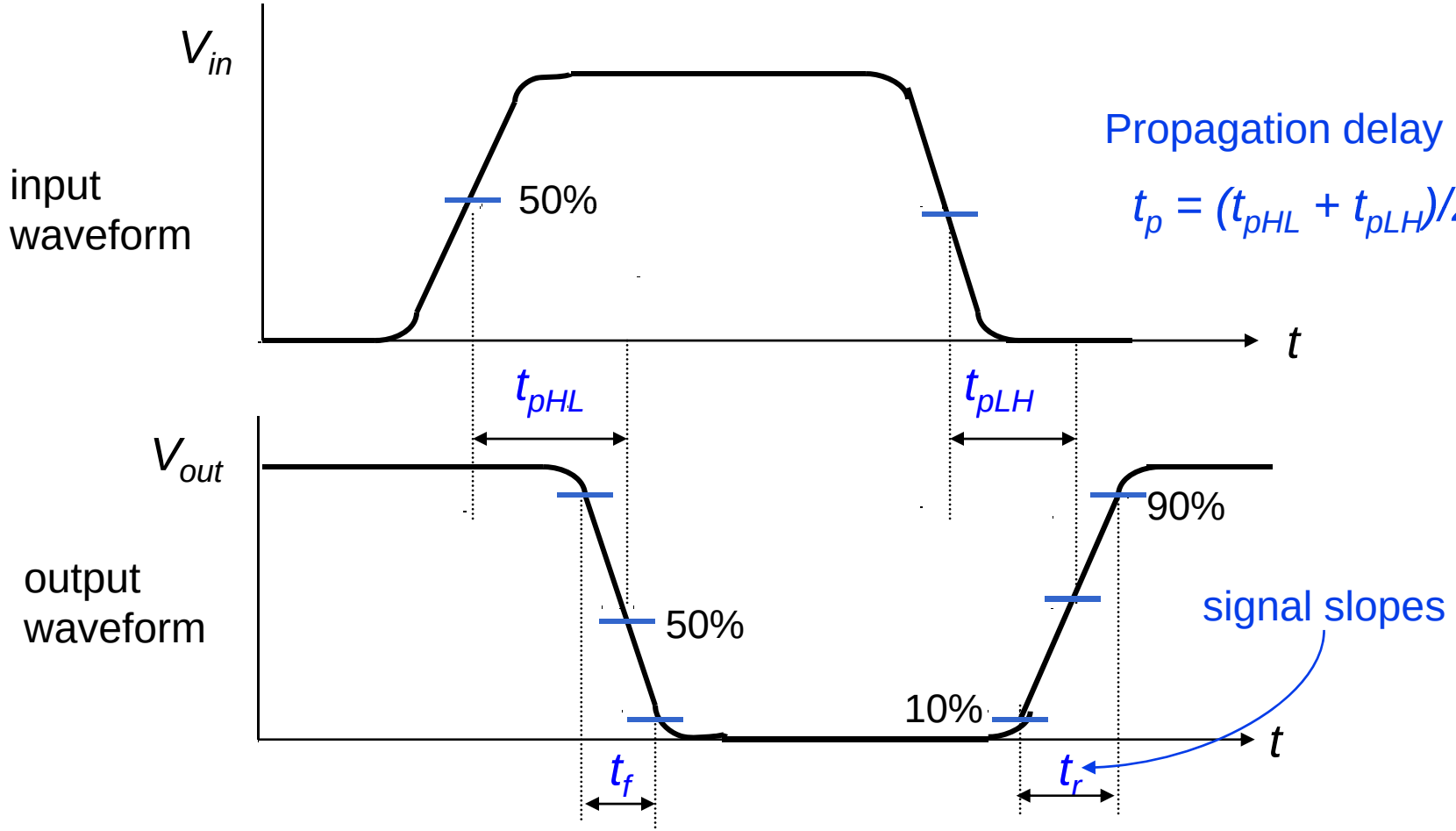
VLSI Design

MOS & Wire Capacitances

[Adapted from Rabaey's *Digital Integrated Circuits*, ©2002, J. Rabaey et al.]

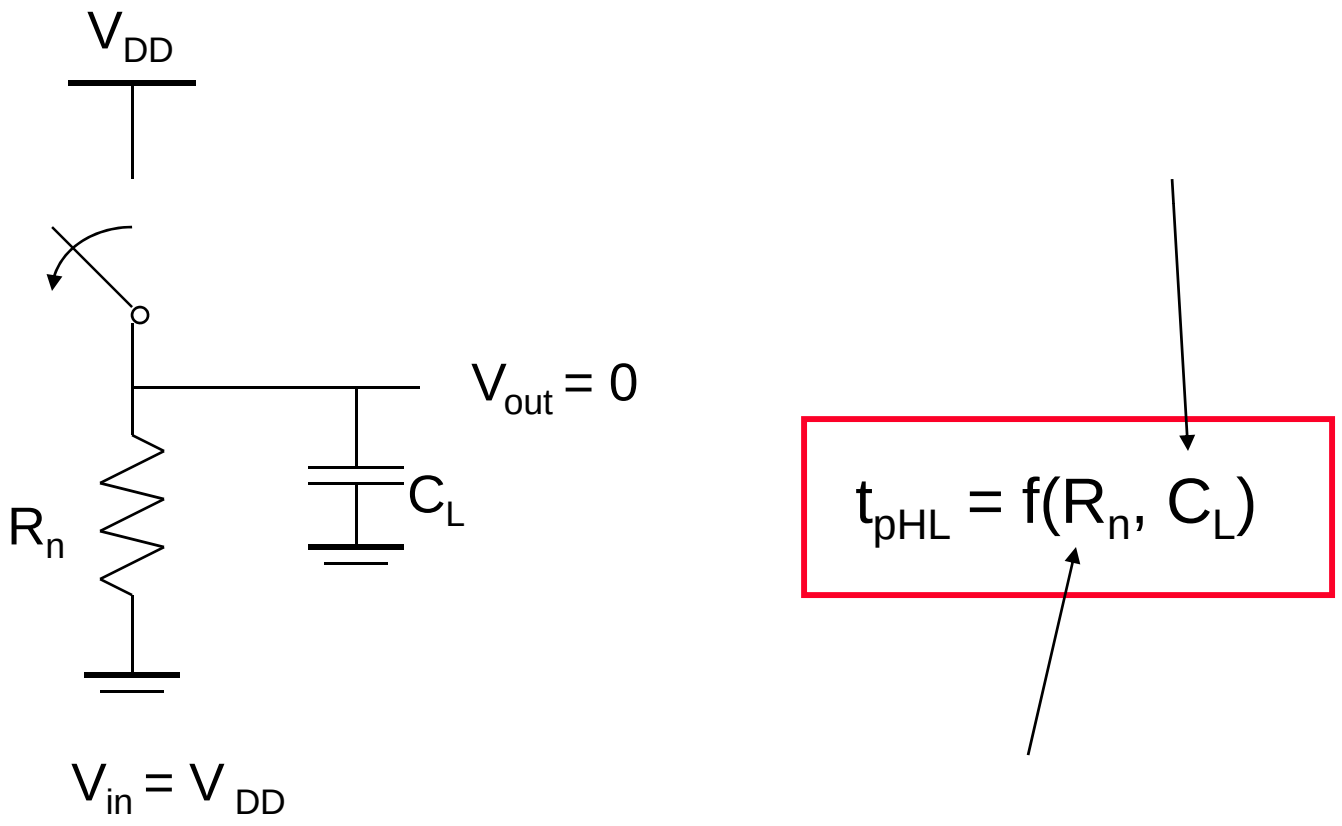
Review: Delay Definitions



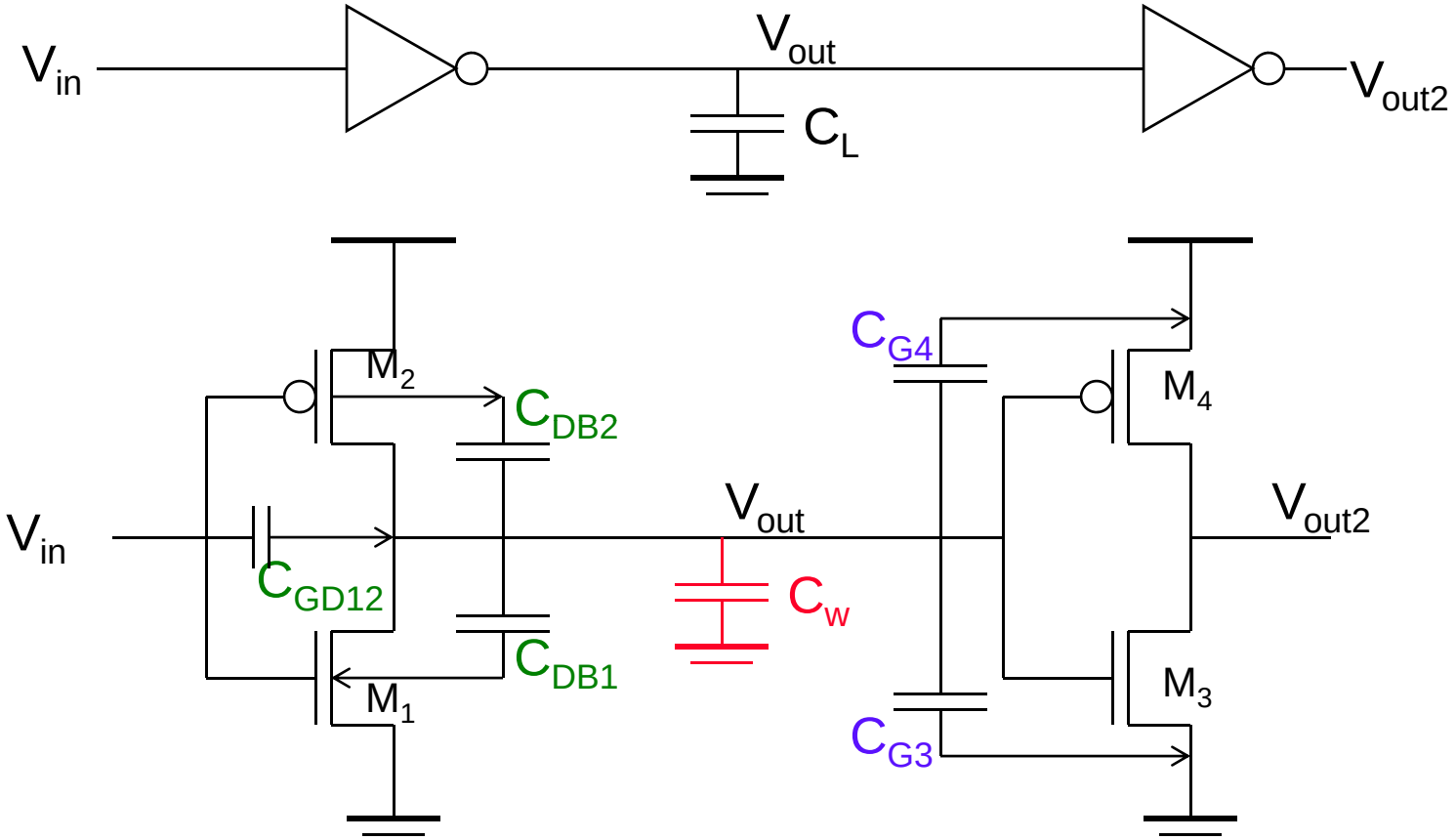


CMOS Inverter: Dynamic

- Transient, or dynamic, response determines the maximum speed at which a device can be operated.



Sources of Capacitance



intrinsic MOS transistor capacitances

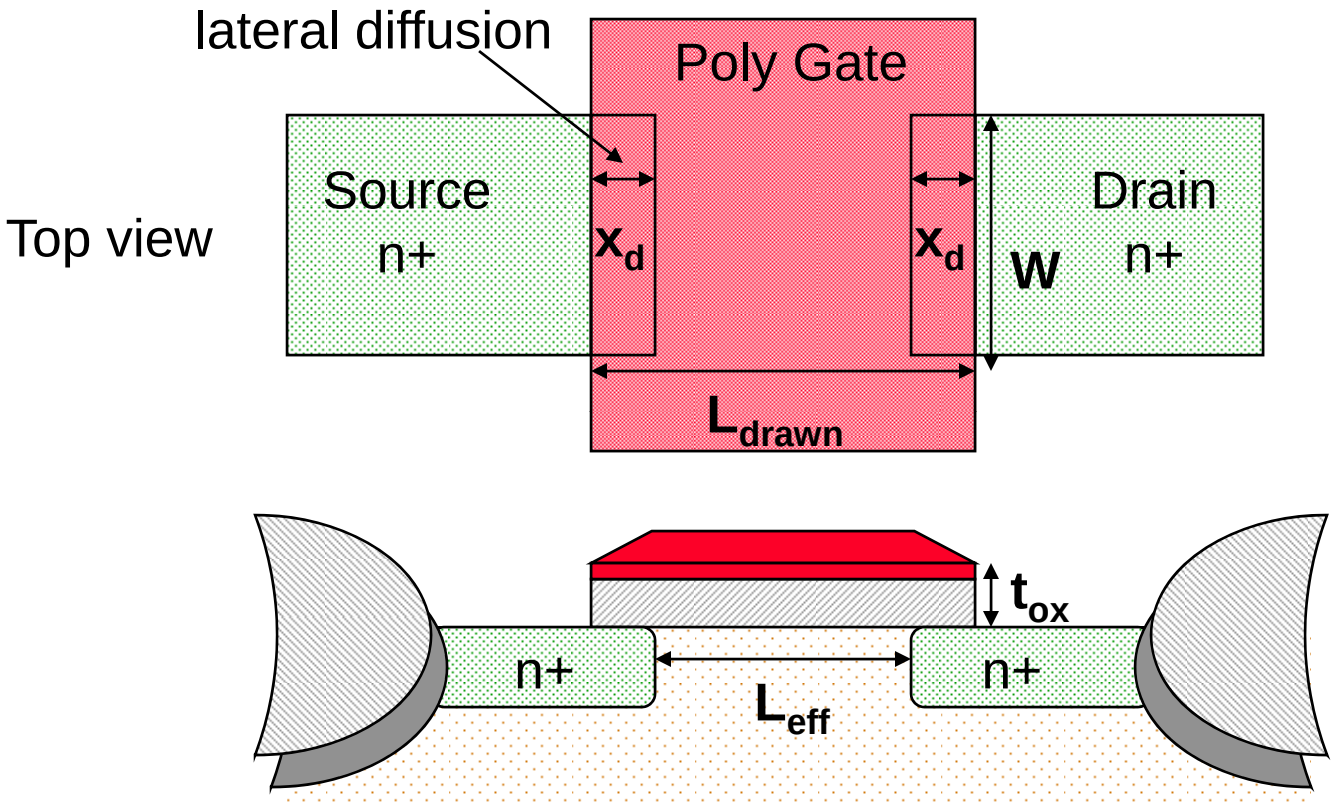
extrinsic MOS transistor (fanout) capacitances

wiring (interconnect) capacitance

MOS Intrinsic Capacitances

- Structure capacitances
- Channel capacitances
- Depletion regions of the reverse-biased *pn*-junctions of the drain and source

MOS Structure Capacitances

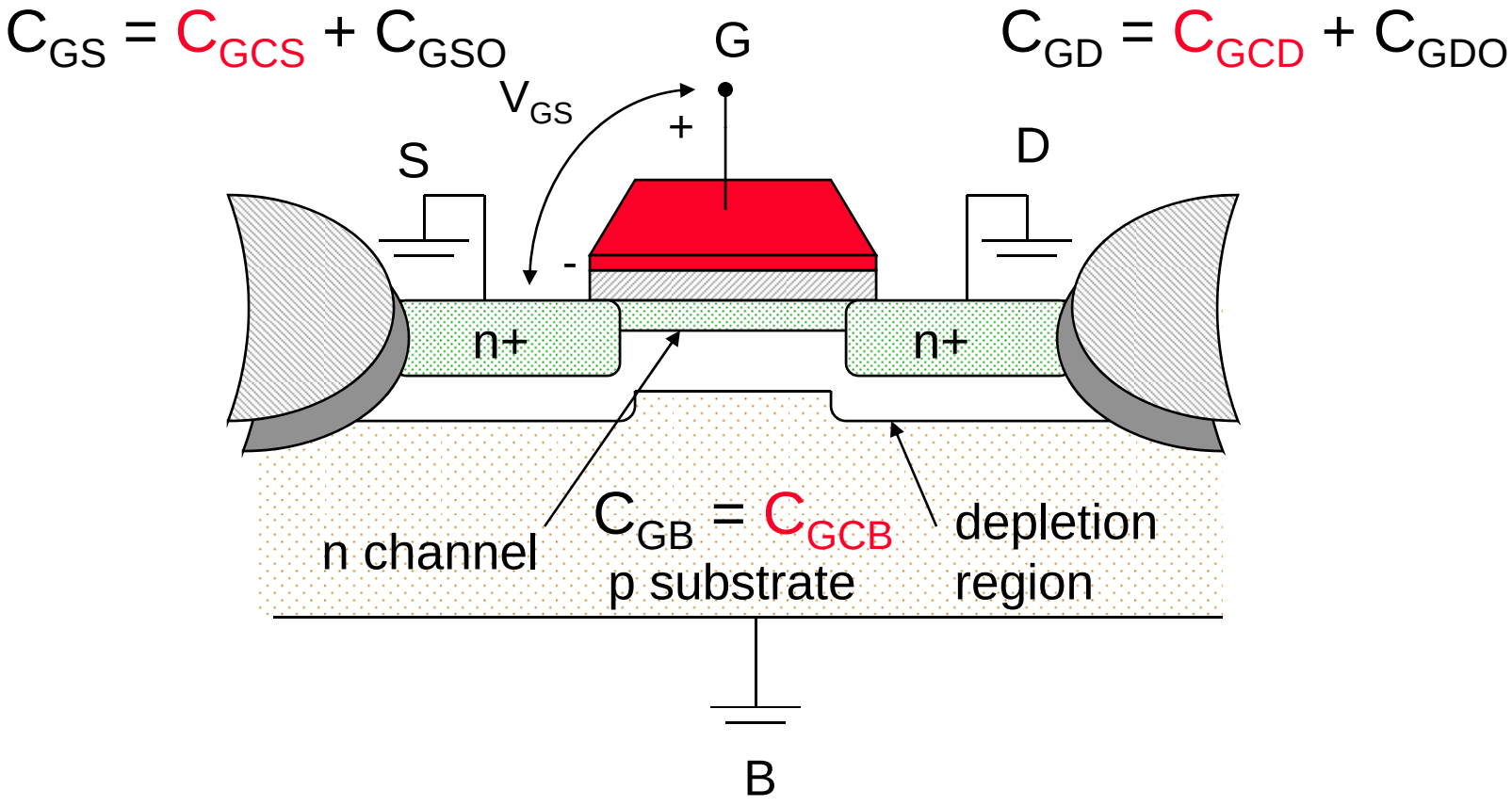


Overlap capacitance (linear)

$$C_{GSO} = C_{GDO} = C_{ox}\, x_d\, W = C_o\, W$$

MOS Channel Capacitances

- The gate-to-channel capacitance depends upon the operating region and the terminal voltages



Review: Summary of MOS Operating Regions

- Cutoff (really subthreshold) $V_{GS} \leq V_T$
  - Exponential in V_{GS} with linear V_{DS} dependence

$$I_D = I_S\, e^{(qV_{GS}/nkT)} \left(1 - e^{-(qV_{DS}/kT)}\right) (1 - \lambda V_{DS}) \quad \text{where} \quad n \geq 1$$

- Strong Inversion $V_{GS} > V_T$
  - Linear (Resistive) $V_{DS} < V_{DSAT} = V_{GS} - V_T$

$$I_D = k' W/L \left[(V_{GS} - V_T)V_{DS} - V_{DS}^2/2\right] (1+\lambda V_{DS})\, \kappa(V_{DS})$$

  - Saturated (Constant Current) $V_{DS} \geq V_{DSAT} = V_{GS} - V_T$

$$I_{DSat} = k' W/L \left[(V_{GS} - V_T)V_{DSAT} - V_{DSAT}^2/2\right] (1+\lambda V_{DS})\, \kappa(V_{DSAT})$$

	V_{T0}(V)	γ($V^{0.5}$)	V_{DSAT}(V)	k'(A/V^2)	λ(V^{-1})
NMOS	0.43	0.4	0.63	115×10^{-6}	0.06
PMOS	-0.4	-0.4	-1	-30×10^{-6}	-0.1

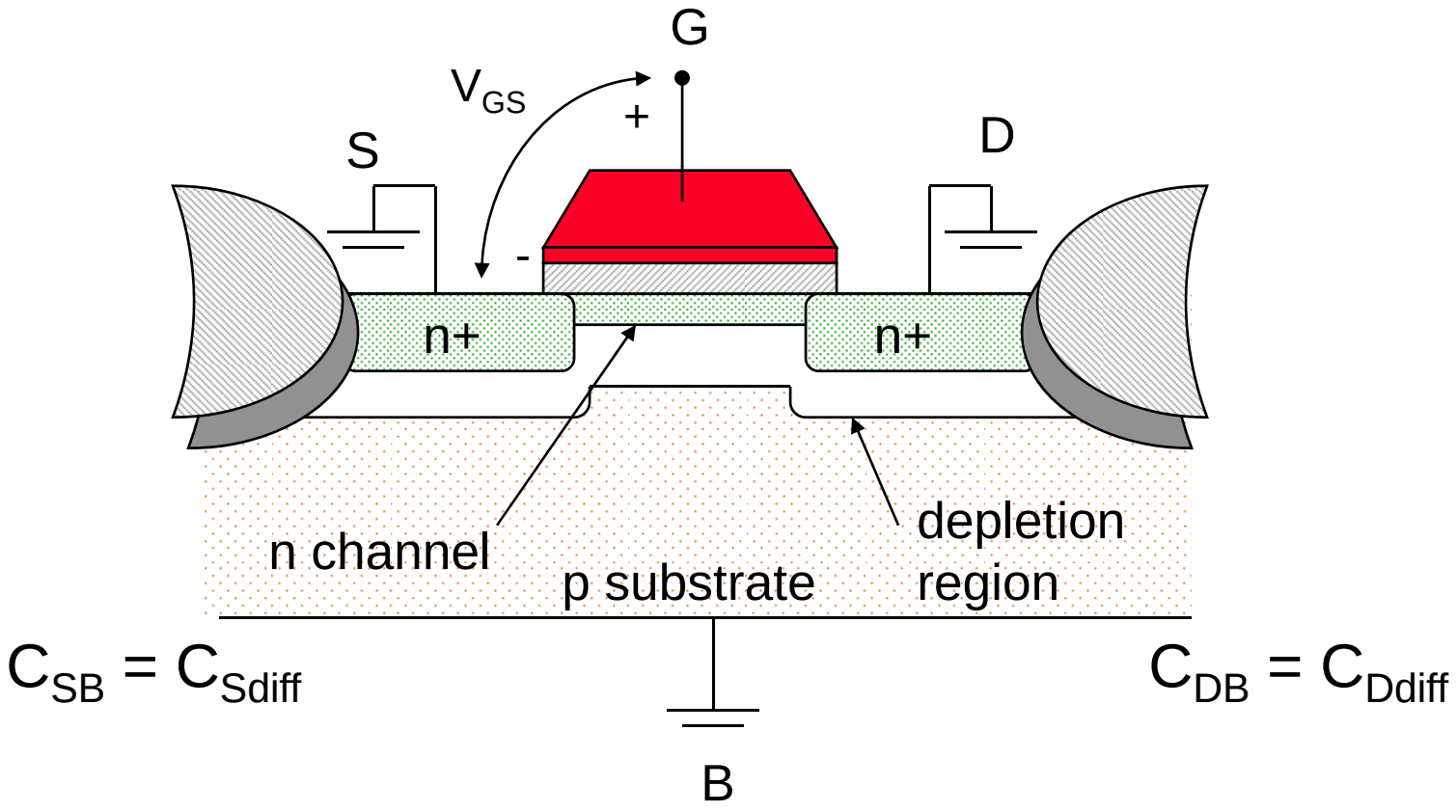
Average Distribution of Channel Capacitance

Operation Region	C_{GCB}	C_{GCS}	C_{GCD}	C_{GC}	C_G
Cutoff	$C_{ox}WL$	0	0	$C_{ox}WL$	$C_{ox}WL + 2C_oW$
Resistive	0	$C_{ox}WL/2$	$C_{ox}WL/2$	$C_{ox}WL$	$C_{ox}WL + 2C_oW$
Saturation	0	$(2/3)C_{ox}WL$	0	$(2/3)C_{ox}WL$	$(2/3)C_{ox}WL + 2C_oW$

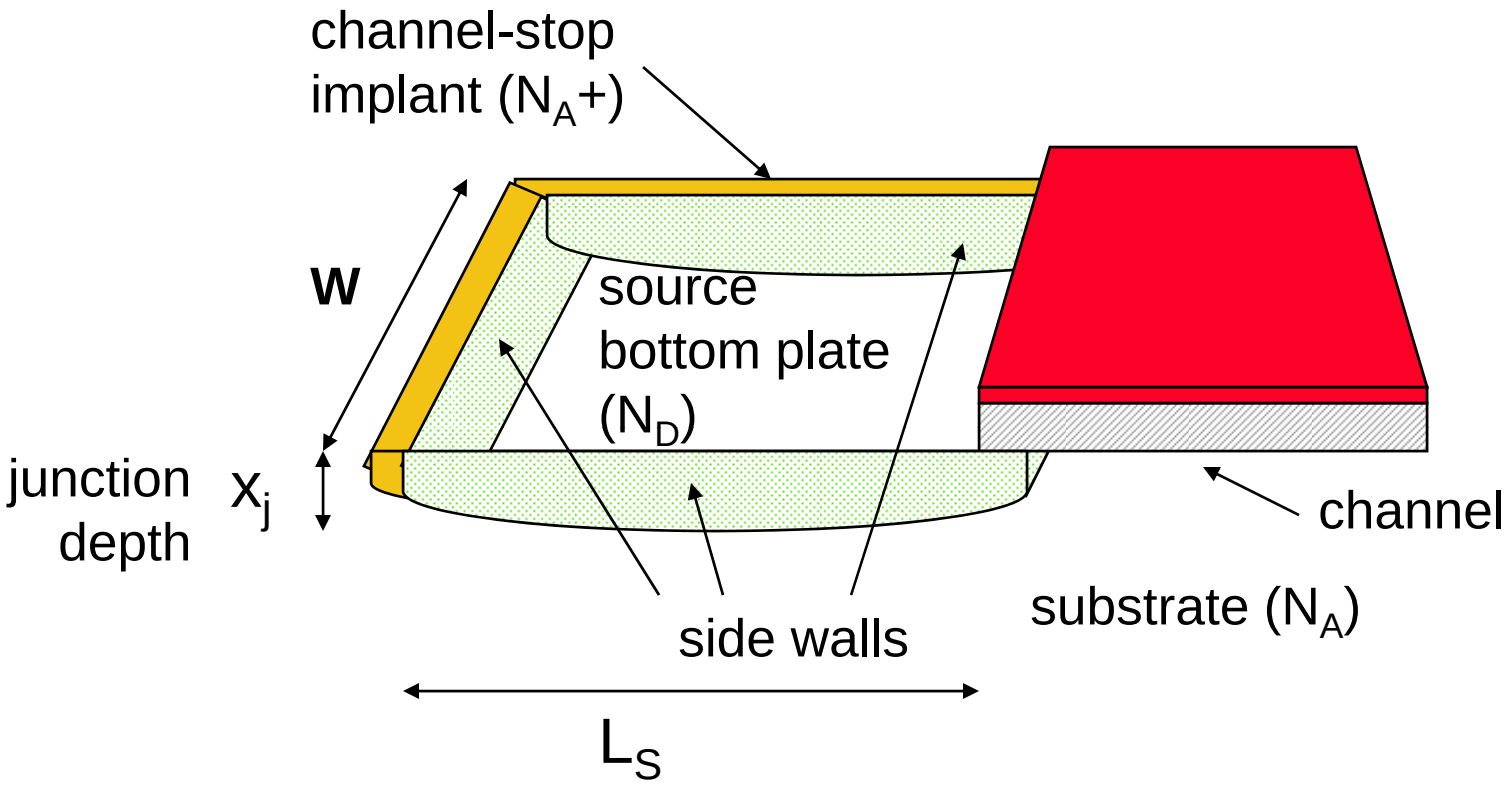
- Channel capacitance components are nonlinear and vary with operating voltage
- Most important regions are cutoff and saturation since that is where the device spends most of its time

MOS Diffusion Capacitances

- The junction (or diffusion) capacitance is from the reverse-biased source-body and drain-body pn-junctions.



Source Junction View

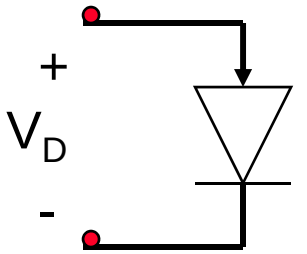


$$C_{diff} = C_{bp} + C_{sw} = C_j\, AREA + C_{jsw}\, PERIMETER$$

$$= C_j\, L_S\, W + C_{jsw}\,(2L_S + W)$$

Review: Reverse Bias Diode

- All diodes in MOS digital circuits are reverse biased; the dynamic response of the diode is determined by depletion-region charge or junction capacitance



$$C_j = C_{j0}/((1 - V_D)/\phi_0)^m$$

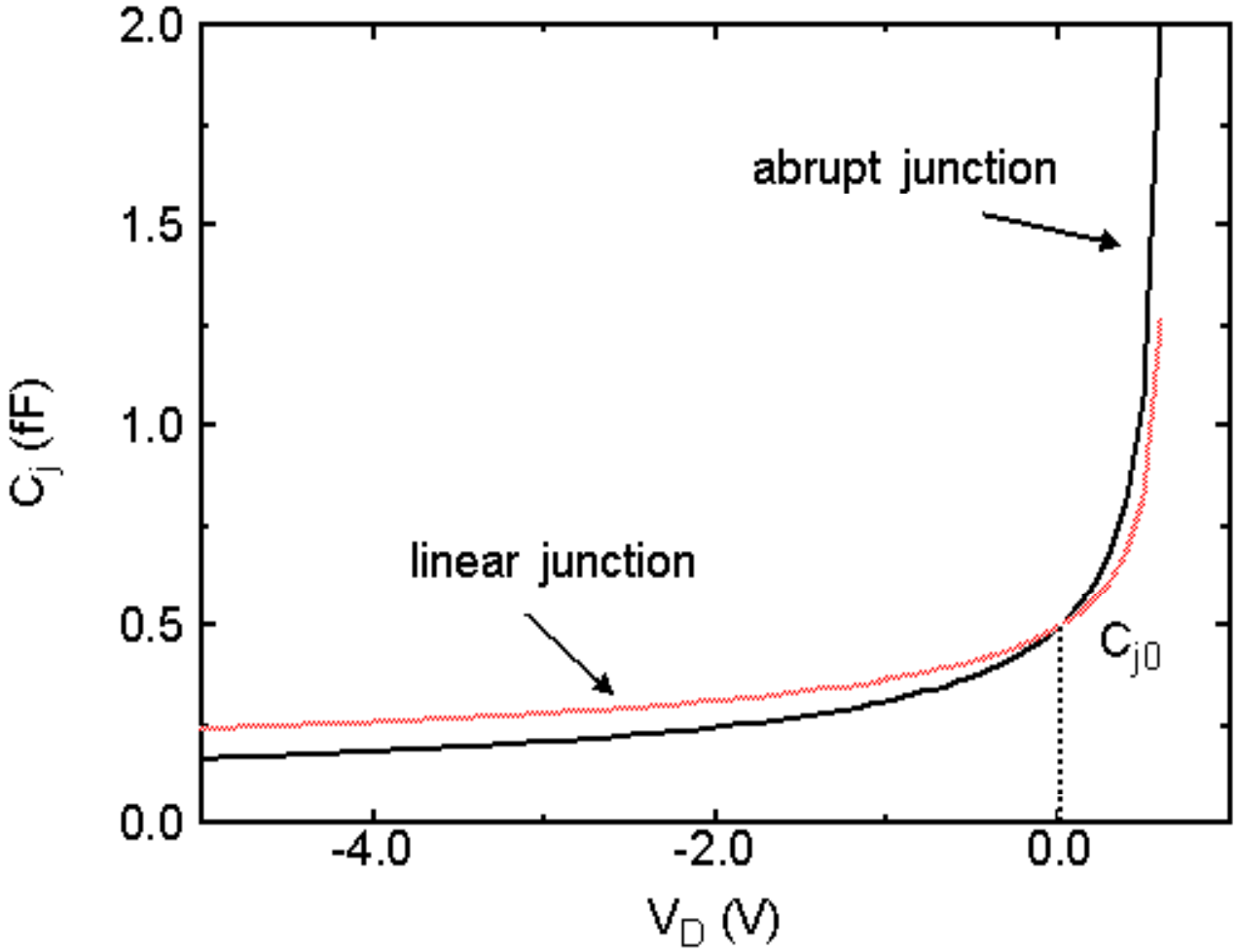
where C_{j0} is the capacitance under zero-bias conditions (a function of physical parameters), ϕ_0 is the built-in potential (a function of physical parameters and temperature)

and m is the grading coefficient

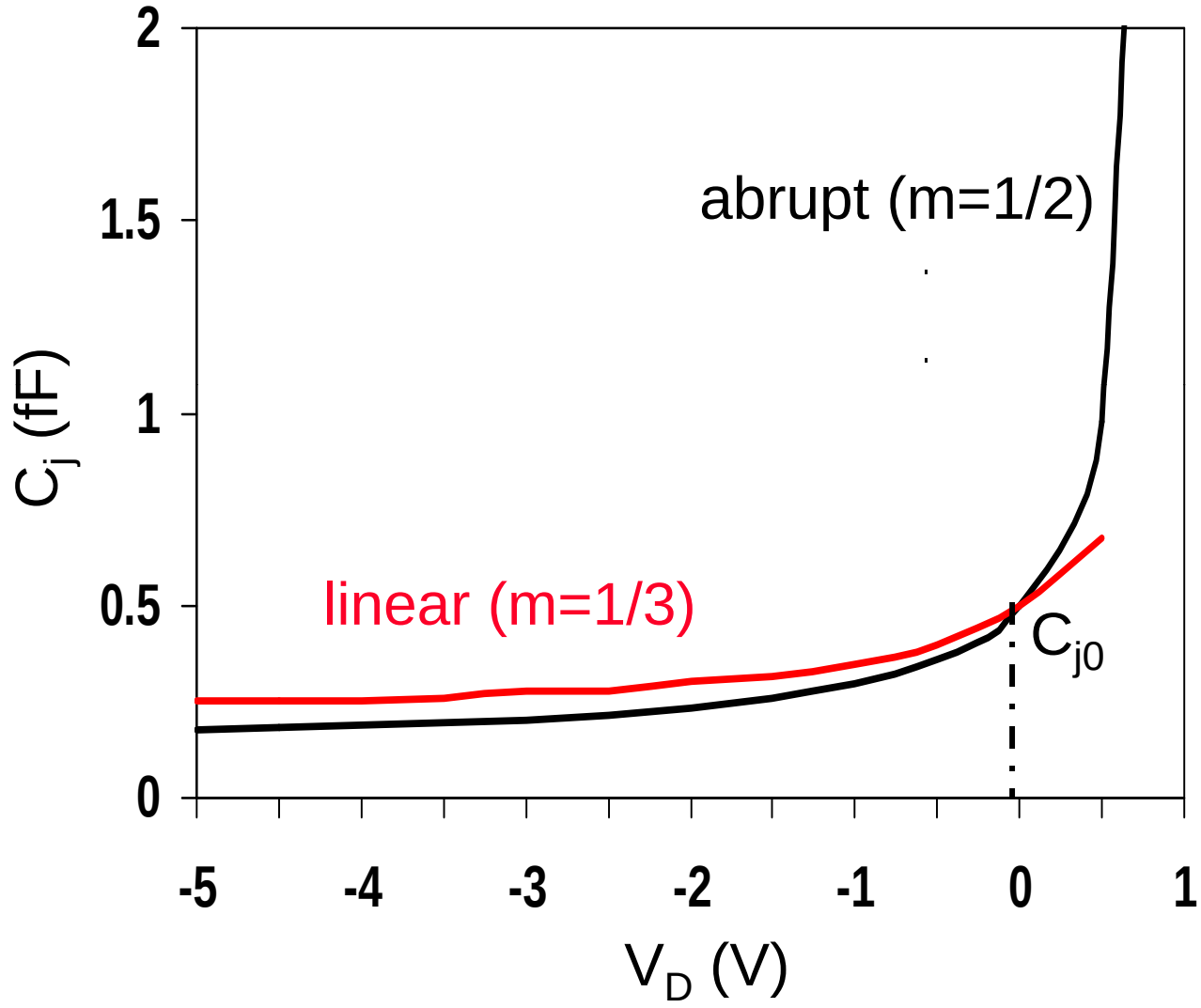
- m = ½ for an abrupt junction (transition from n to p-material is instantaneous)
- m = 1/3 for a linear (or graded) junction (transition is gradual)

- Nonlinear dependence (that decreases with increasing reverse bias)

Junction Capacitance



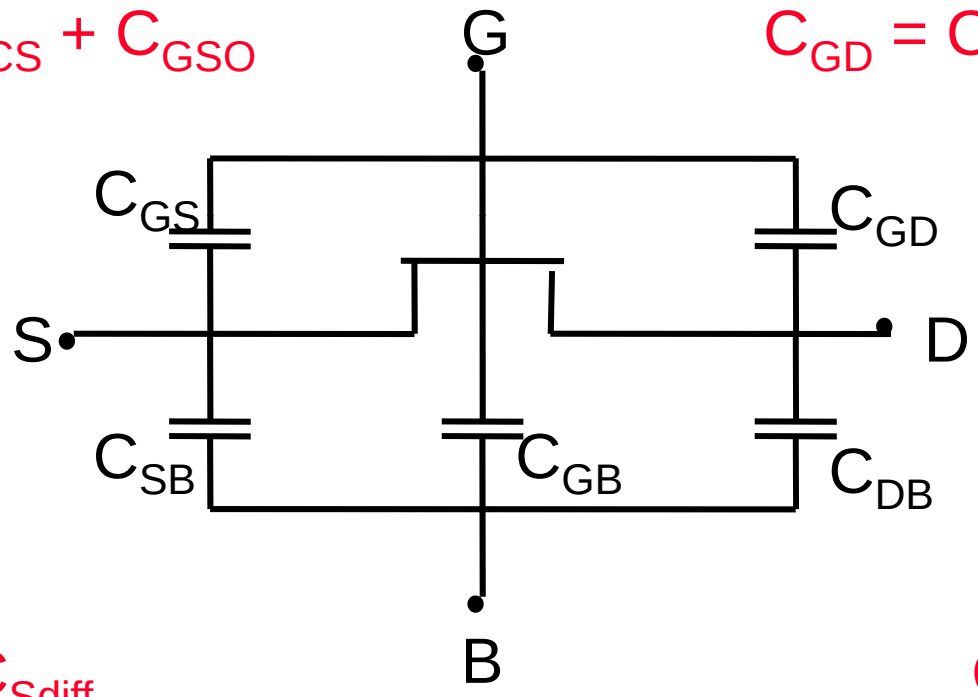
Reverse-Bias Diode Junction Capacitance



MOS Capacitance Model

$C_{GS} = C_{GCS} + C_{GSO}$

$C_{GD} = C_{GCD} + C_{GDO}$



$C_{SB} = C_{Sdiff}$

$C_{DB} = C_{Ddiff}$

$C_{GB} = C_{GCB}$

Transistor Capacitance Values for 0.25μ

Example: For an NMOS with L = 0.24 μm, W = 0.36 μm, $L_D = L_S = 0.625$ μm

$C_{GSO} = C_{GDO} = C_{ox}\, x_d\, W = C_o\, W =$

$C_{GC} = C_{ox}\, WL =$

so $C_{gate\_cap} = C_{ox}WL + 2C_oW =$

$C_{bp} = C_j\, L_S\, W =$

$C_{sw} = C_{jsw}\, (2L_S + W) =$

so $C_{diffusion\_cap} =$

	C_{ox} (fF/μm²)	C_o (fF/μm)	C_j (fF/μm²)	m_j	ϕ_b (V)	C_{jsw} (fF/μm)	m_{jsw}	ϕ_{bsw} (V)
NMOS	6	0.31	2	0.5	0.9	0.28	0.44	0.9
PMOS	6	0.27	1.9	0.48	0.9	0.22	0.32	0.9

Transistor Capacitance Values for 0.25μ

Example: For an NMOS with L = 0.24 μm, W = 0.36 μm, $L_D = L_S = 0.625$ μm

$C_{GSO} = C_{GDO} = C_{ox}\, x_d\, W = C_o\, W = 0.11$ fF

$C_{GC} = C_{ox}\, WL = 0.52$ fF

so $C_{gate\_cap} = C_{ox}WL + 2C_oW = 0.74$ fF

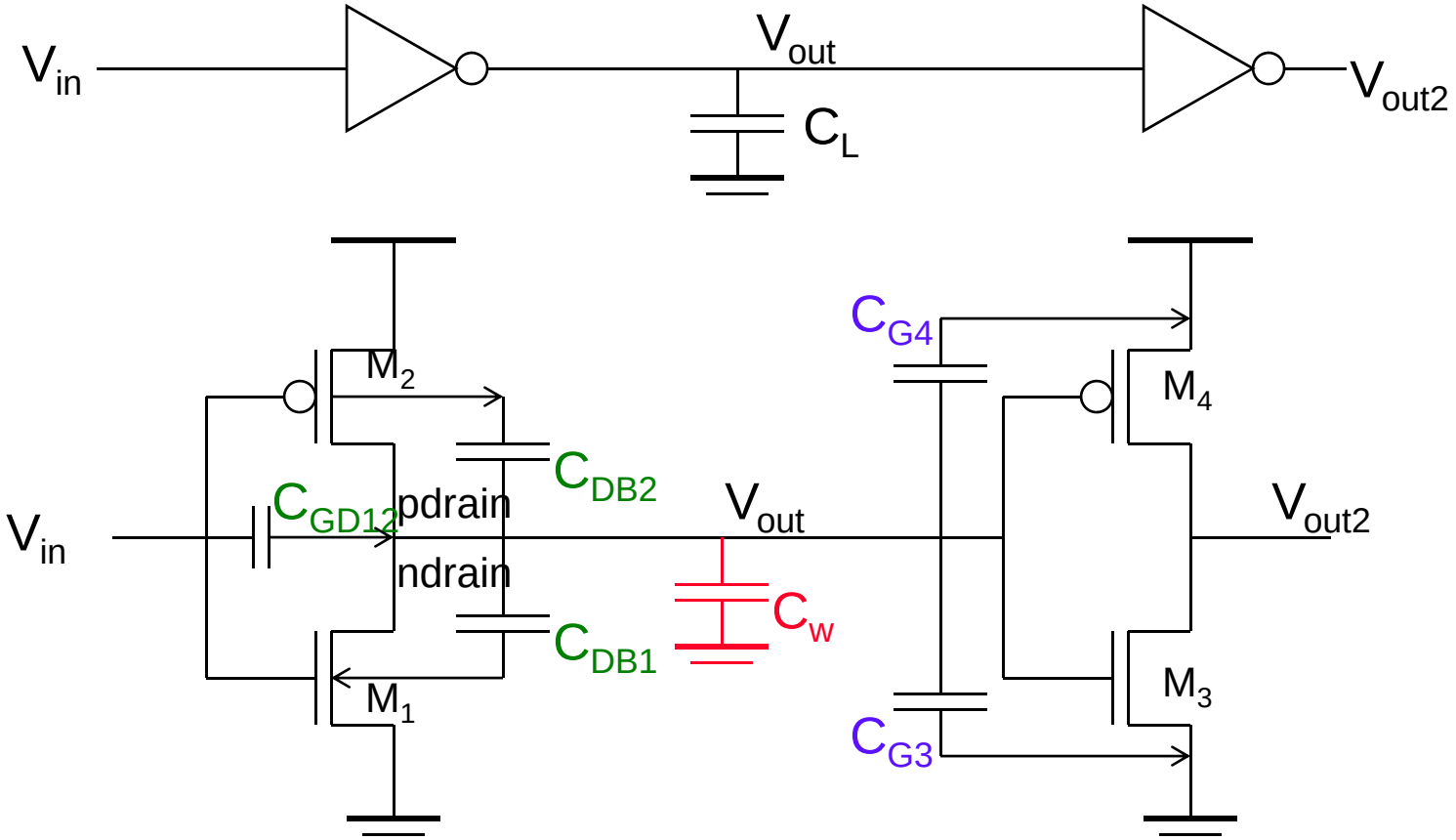
$C_{bp} = C_j\, L_S\, W = 0.45$ fF

$C_{sw} = C_{jsw}\,(2L_S + W) = 0.45$ fF

so $C_{diffusion\_cap} = 0.90$ fF

	C_{ox} (fF/μm²)	C_o (fF/μm)	C_j (fF/μm²)	m_j	ϕ_b (V)	C_{jsw} (fF/μm)	m_{jsw}	ϕ_{bsw} (V)
NMOS	6	0.31	2	0.5	0.9	0.28	0.44	0.9
PMOS	6	0.27	1.9	0.48	0.9	0.22	0.32	0.9

Review: Sources of Capacitance



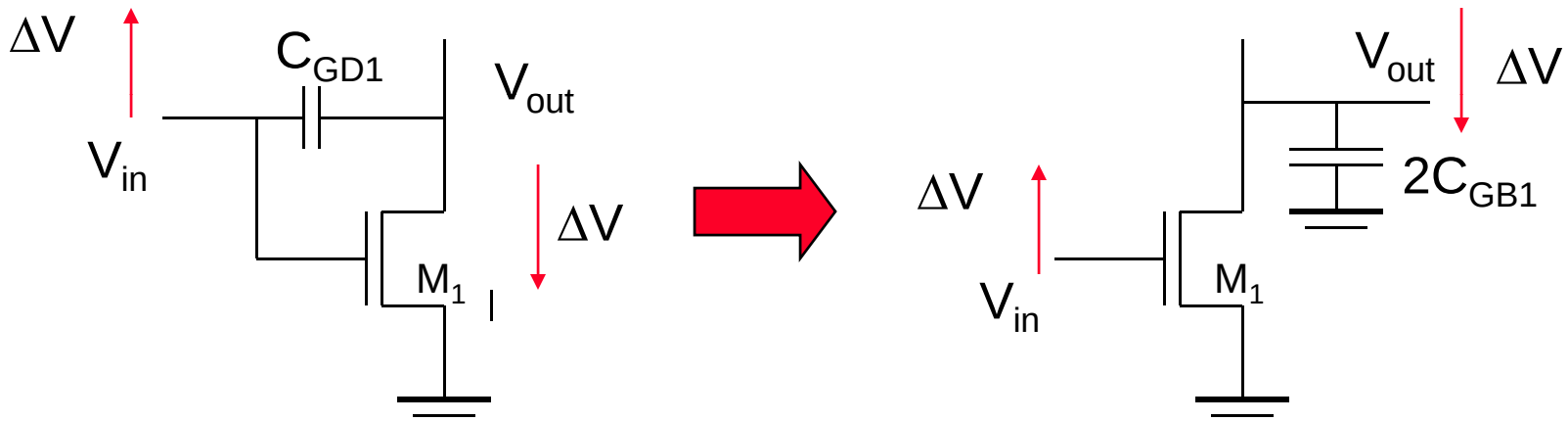
intrinsic MOS transistor capacitances

extrinsic MOS transistor (fanout) capacitances

wiring (interconnect) capacitance

Gate-Drain Capacitance: The Miller Effect

- M1 and M2 are either in cut-off or in saturation.
- The floating gate-drain capacitor is replaced by a capacitance-to-ground (gate-bulk capacitor).



- A capacitor experiencing identical but opposite voltage swings at both its terminals can be replaced by a capacitor to ground whose value is two times the original value

Drain-Bulk Capacitance: K_{eq}'s (for 2.5 μm)

- We can simplify the diffusion capacitance calculations *even* further by using a K_{eq} to relate the linearized capacitor to the value of the junction capacitance under zero-bias

$$C_{eq} = K_{eq}\, C_{j0}$$

	high-to-low		low-to-high	
	K_{eqbp}	K_{eqsw}	K_{eqbp}	K_{eqsw}
NMOS	0.57	0.61	0.79	0.81
PMOS	0.79	0.86	0.59	0.7

Extrinsic (Fan-Out) Capacitance

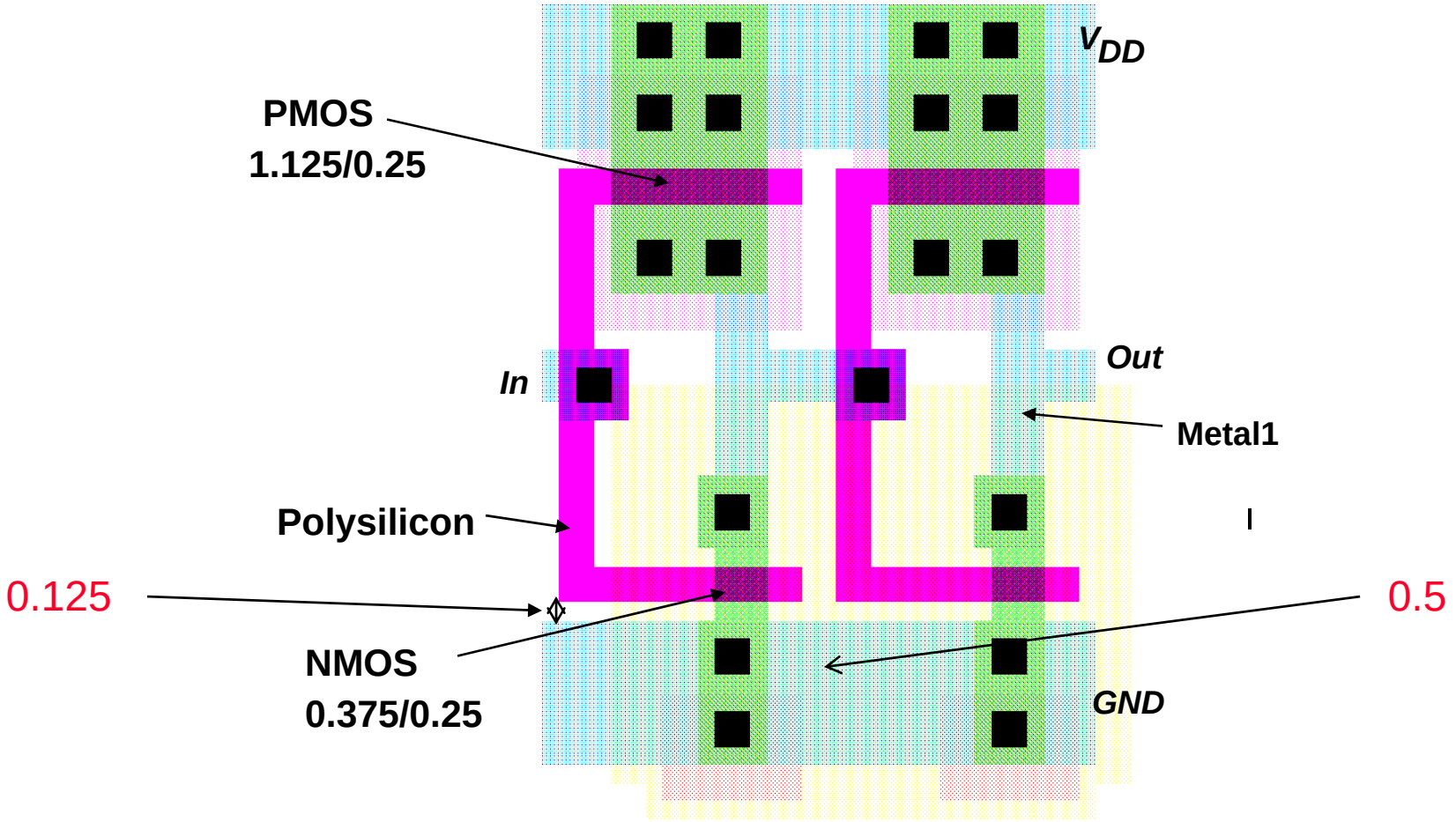
- The extrinsic, or fan-out, capacitance is the total gate capacitance of the loading gates M3 and M4.

$$C_{fan\text{-}out} = C_{gate}\,(NMOS) + C_{gate}\,(PMOS)$$

$$= (C_{GSOn} + C_{GDOn} + W_n L_n C_{ox}) + (C_{GSOp} + C_{GDOp} + W_p L_p C_{ox})$$

- Simplification of the actual situation
  - Assumes all the components of C_{gate} are between V_{out} and GND (or V_{DD})
  - Assumes the channel capacitances of the loading gates are constant

Layout of Two Chained Inverters



	W/L	AD (μm²)	PD (μm)	AS (μm²)	PS (μm)
NMOS	0.375/0.25	0.3	1.875	0.3	1.875
PMOS	1.125/0.25	0.7	2.375	0.7	2.375

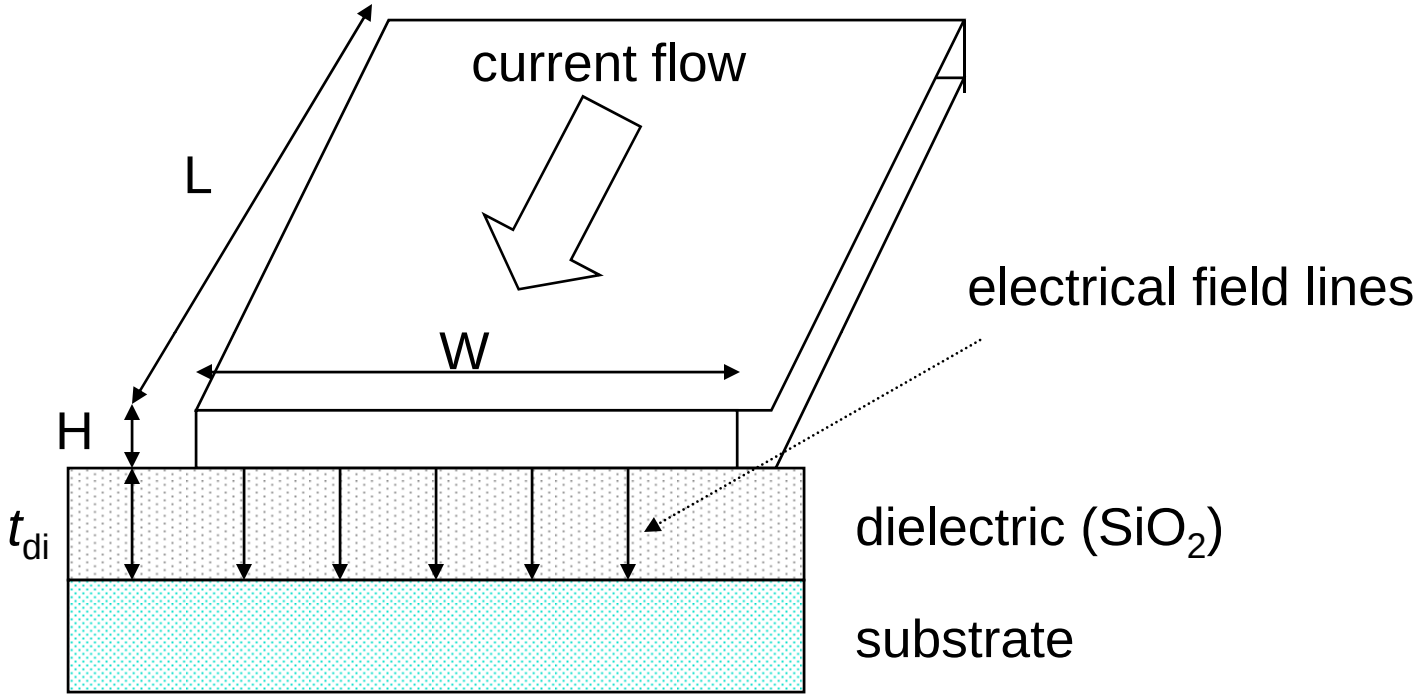
Components of C_L (0.25 μm)

C Term	Expression	Value (fF) H→L	Value (fF) L→H
C_{GD1}	$2\,C_{on}\,W_n$	0.23	0.23
C_{GD2}	$2\,C_{op}\,W_p$	0.61	0.61
C_{DB1}	$K_{eqbpn}AD_nC_j + K_{eqswn}PD_nC_{jsw}$	0.66	0.90
C_{DB2}	$K_{eqbpp}AD_pC_j + K_{eqswp}PD_pC_{jsw}$	1.5	1.15
C_{G3}	$(2\,C_{on})W_n + C_{ox}W_nL_n$	0.76	0.76
C_{G4}	$(2\,C_{op})W_p + C_{ox}W_pL_p$	2.28	2.28
C_w	from extraction	0.12	0.12
C_L	$\sum$	6.1	6.0

Wiring Capacitance

- The wiring capacitance depends upon the length and width of the connecting wires and is a function of the fan-out from the driving gate and the number of fan-out gates.
- Wiring capacitance is growing in importance with the scaling of technology.

Parallel Plate Wiring Capacitance



permittivity constant (SiO_2= 3.9)

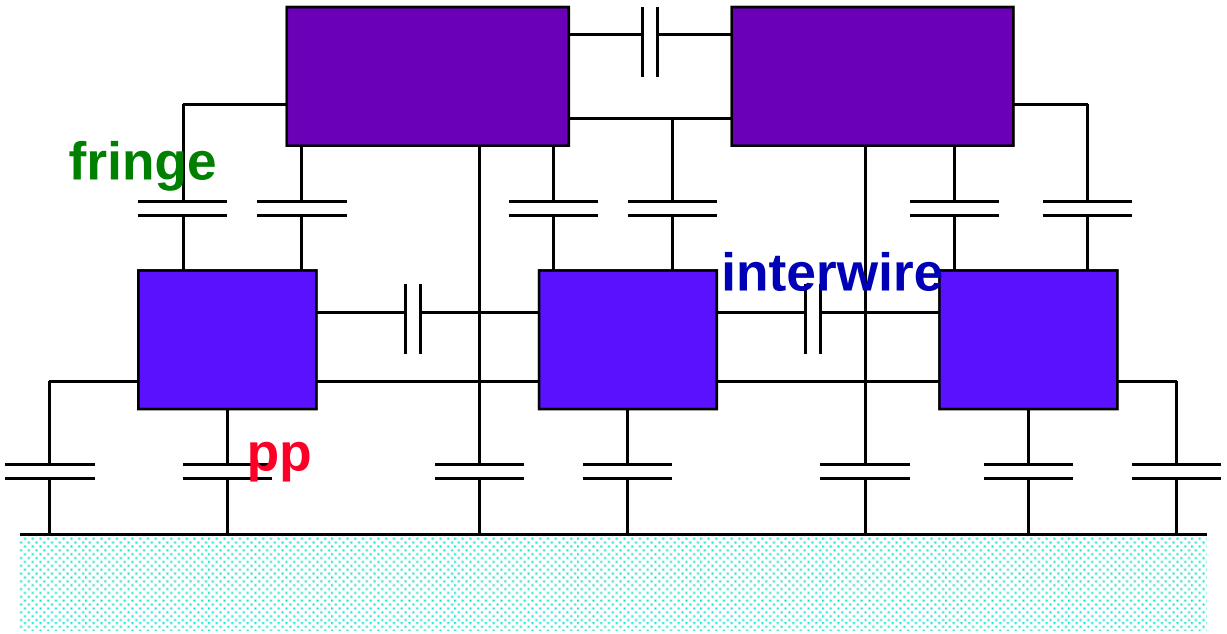
$$C_{pp} = (\varepsilon_{di}/t_{di})\ WL$$

Permittivity Values of Some Dielectrics

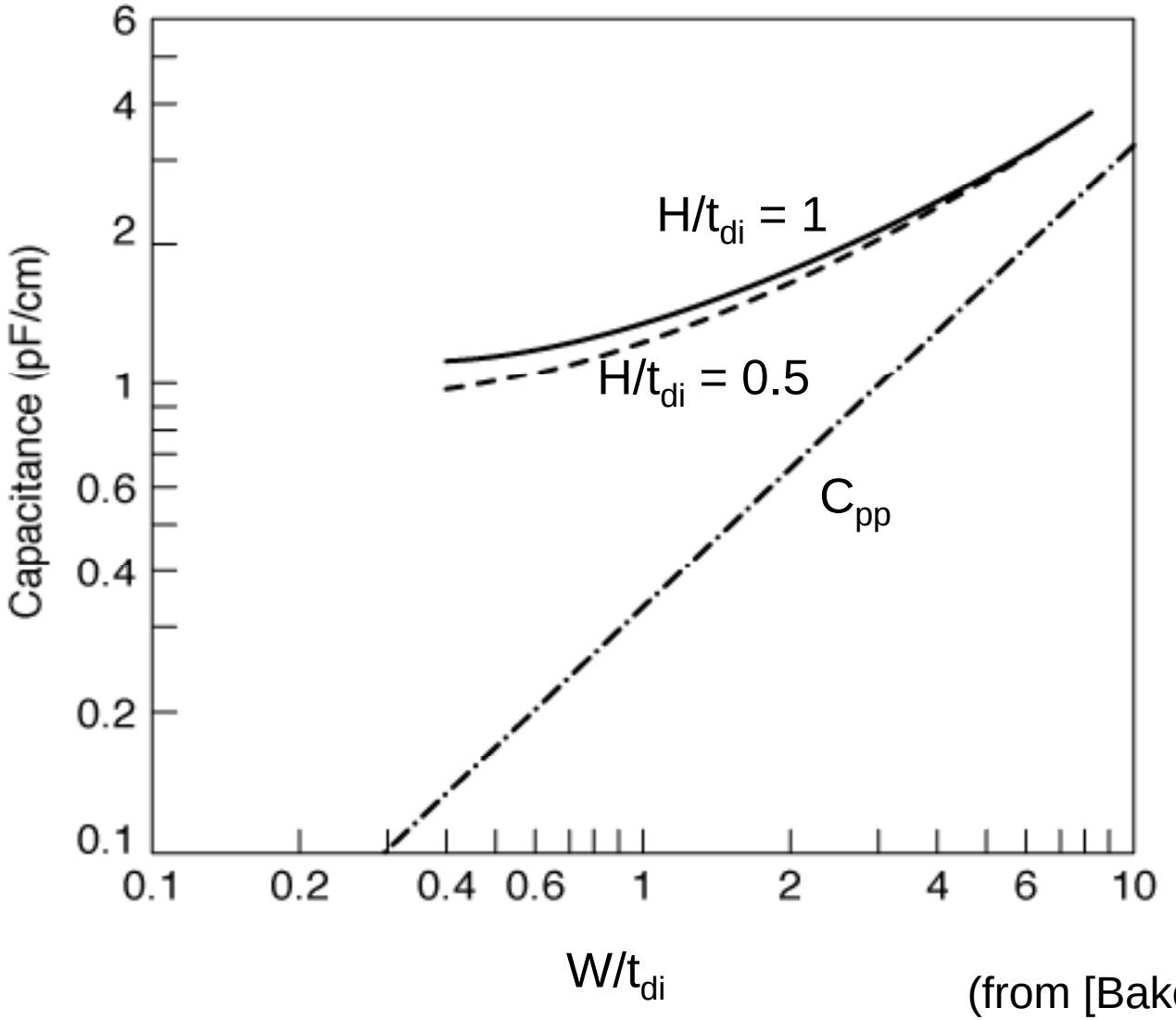
Material	ε_{di}
Free space	1
Teflon AF	2.1
Aromatic thermosets (SiLK)	2.6 – 2.8
Polyimides (organic)	3.1 – 3.4
Fluorosilicate glass (FSG)	3.2 – 4.0
Silicon dioxide	3.9 – 4.5
Glass epoxy (PCBs)	5
Silicon nitride	7.5
Alumina (package)	9.5
Silicon	11.7

Sources of Interwire Capacitance

$$C_{wire} = C_{pp} + C_{fringe} + C_{interwire}$$
$$= (\varepsilon_{di}/t_{di})WL + (2\pi\varepsilon_{di})/\log(t_{di}/H) + (\varepsilon_{di}/t_{di})HL$$

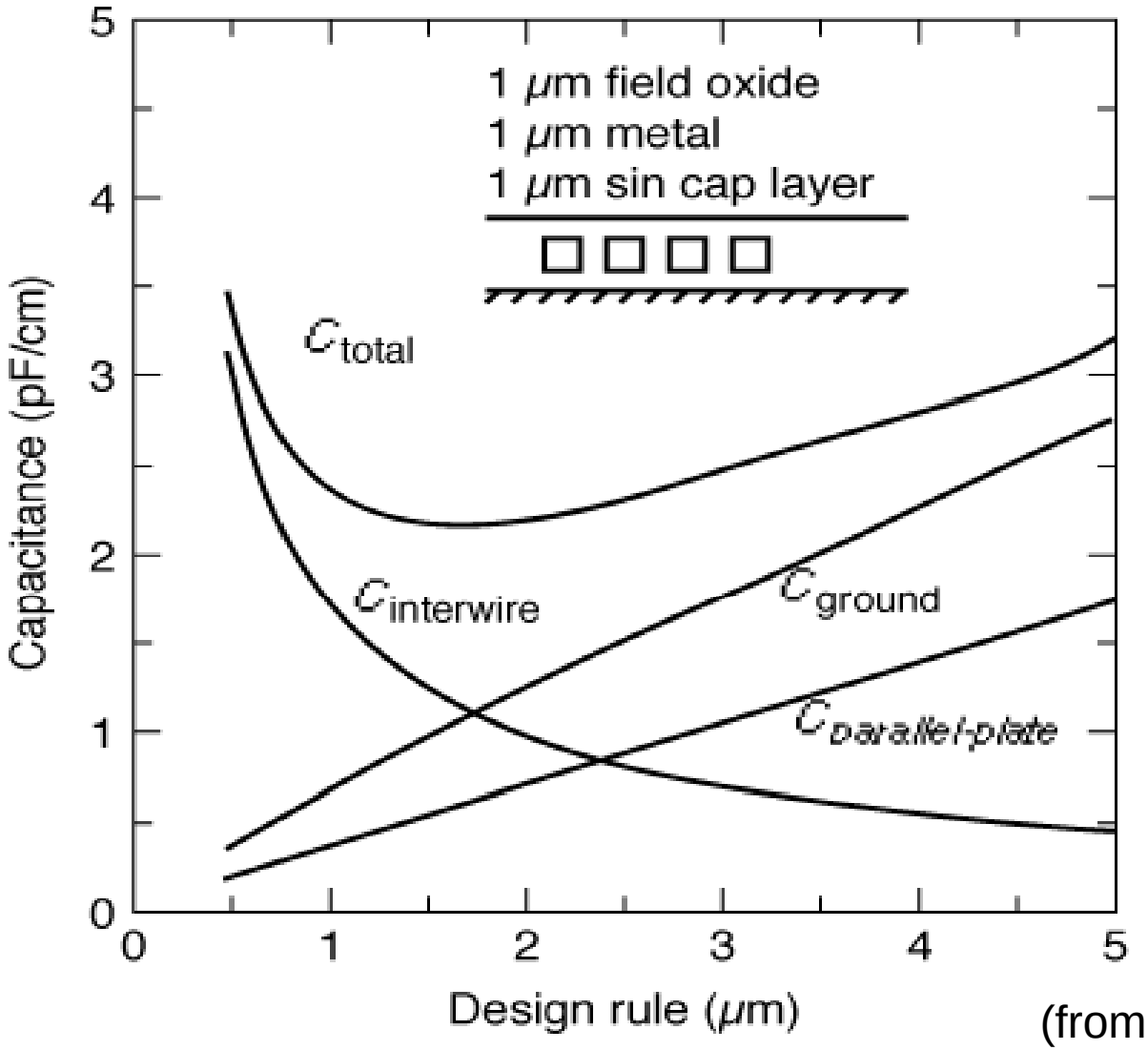


Impact of Fringe Capacitance



(from [Bakoglu89])

Impact of Interwire Capacitance



(from [Bakoglu89])

Insights

- For W/H < 1.5, the fringe component dominates the parallel-plate component. Fringing capacitance can increase the overall capacitance by a factor of 10 or more.
- When W < 1.75H interwire capacitance starts to dominate
- Interwire capacitance is more pronounced for wires in the higher interconnect layers (further from the substrate)
- Rules of thumb
  - Never run wires in diffusion
  - Use poly only for short runs
  - Shorter wires – lower R and C
  - Thinner wires – lower C but higher R
- Wire delay nearly proportional to L^2

Wiring Capacitances

	Field	Active	Poly	Al1	Al2	Al3	Al4
Poly	88						
	54				pp in aF/μm^2		
Al1	30	41	57		fringe in aF/μm		
	40	47	54				
Al2	13	15	17	36			
	25	27	29	45			
Al3	8.9	9.4	10	15	41		
	18	19	20	27	49		
Al4	6.5	6.8	7	8.9	15	35	
	14	15	15	18	27	45	
Al5	5.2	5.4	5.4	6.6	9.1	14	38
	12	12	12	14	19	27	52

	Poly	Al1	Al2	Al3	Al4	Al5
Interwire Cap	40	95	85	85	85	115

per unit wire length in aF/μm for minimally-spaced wires

Dealing with Capacitance

- Low capacitance (low-k) dielectrics (insulators) such as polymide or even air instead of SiO_2
  - family of materials that are low-k dielectrics
  - must also be suitable thermally and mechanically and
  - compatible with (copper) interconnect
- Copper interconnect allows wires to be thinner without increasing their resistance, thereby decreasing interwire capacitance
- SOI (silicon on insulator) to reduce junction capacitance

Next Time: Dealing with Resistance

- MOS structure resistance - R_{on}
- Wiring resistance
- Contact resistance

Next Lecture and Reminders

- Next lecture
  - MOS resistance
    - Reading assignment – Rabaey, et al, 4.3.2, 4.4.1-4.4.4