

ECE 6473  
Lecture 8  
Date: 11/09/2015

Shaloo Rakheja  
Assistant Professor  
Electrical and Computer Engineering, NYU

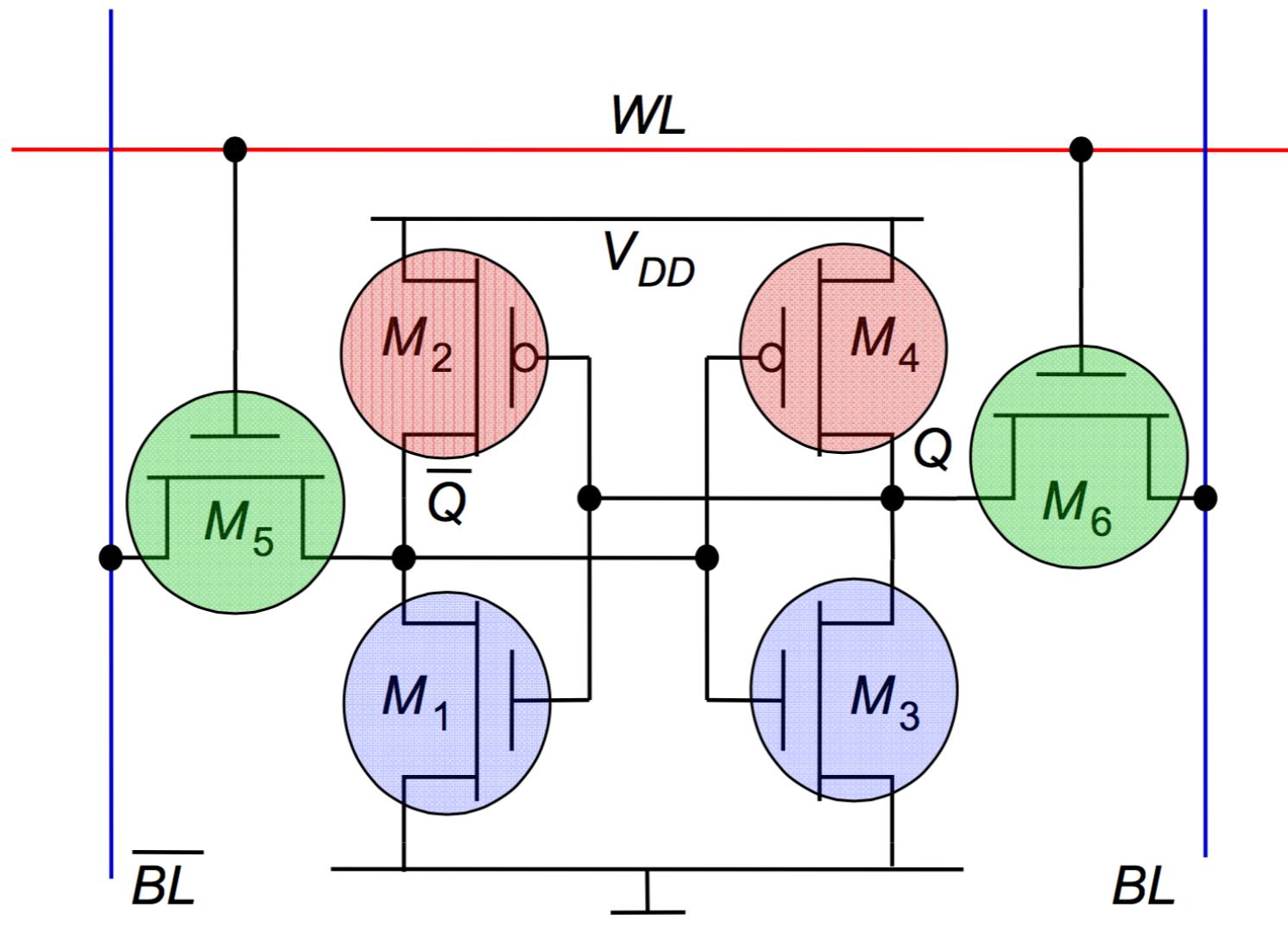
# Reading

- Power point and hand-written lecture notes posted on  
[newclasses.nyu.edu](http://newclasses.nyu.edu)

# Homework # 5

- **Decoder design with in groups of FOUR. Submission on 11/10/2015. If you miss the deadline you will forfeit the 15% credit for timely submission.**

# 6T CMOS SRAM cell



● pull-up or load devices ( $M_2, M_4$ )

● pull-down devices ( $M_1, M_3$ )

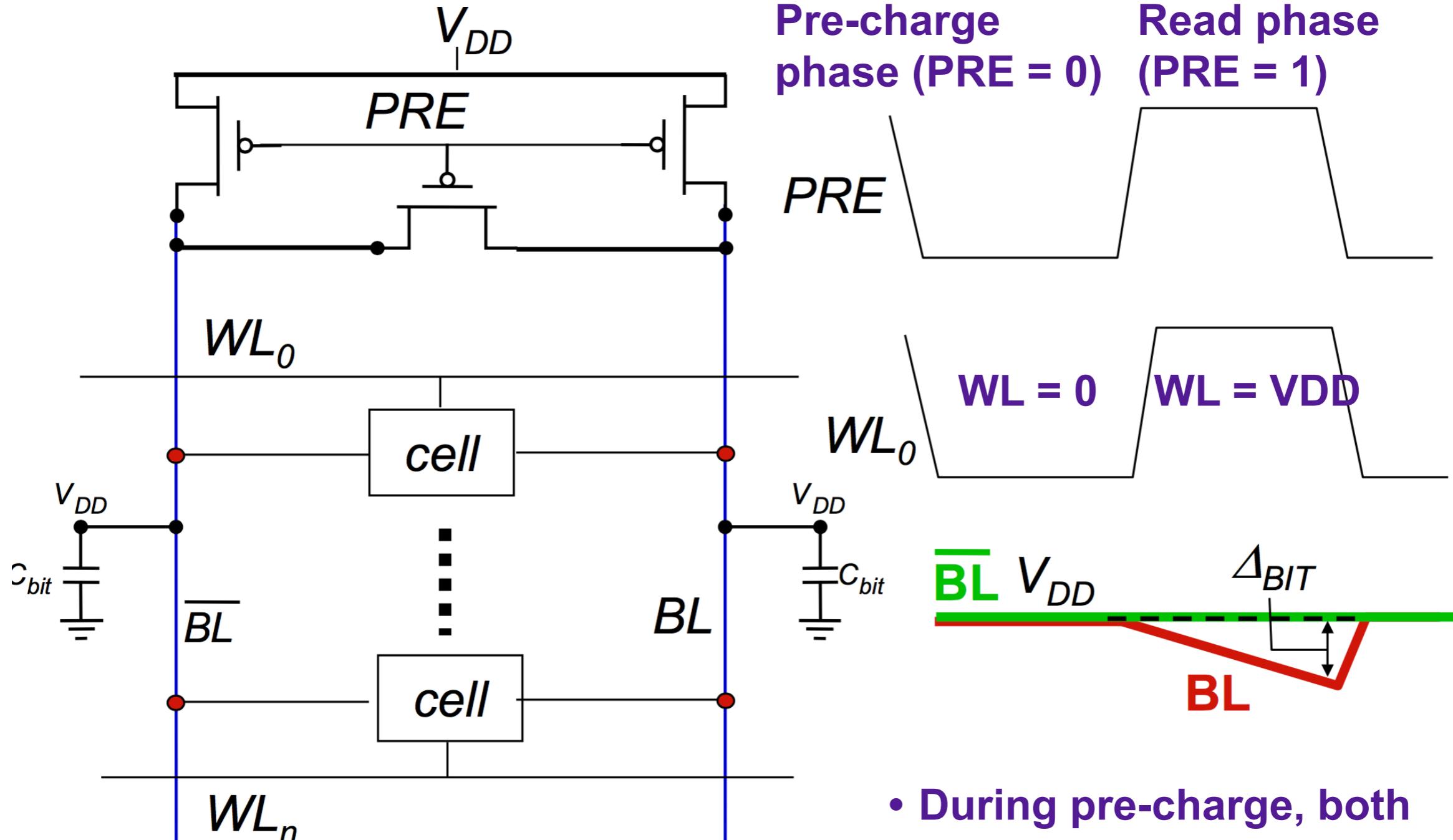
● access or pass-gate devices ( $M_5, M_6$ )

□ Data Holding

□ Read

□ Write

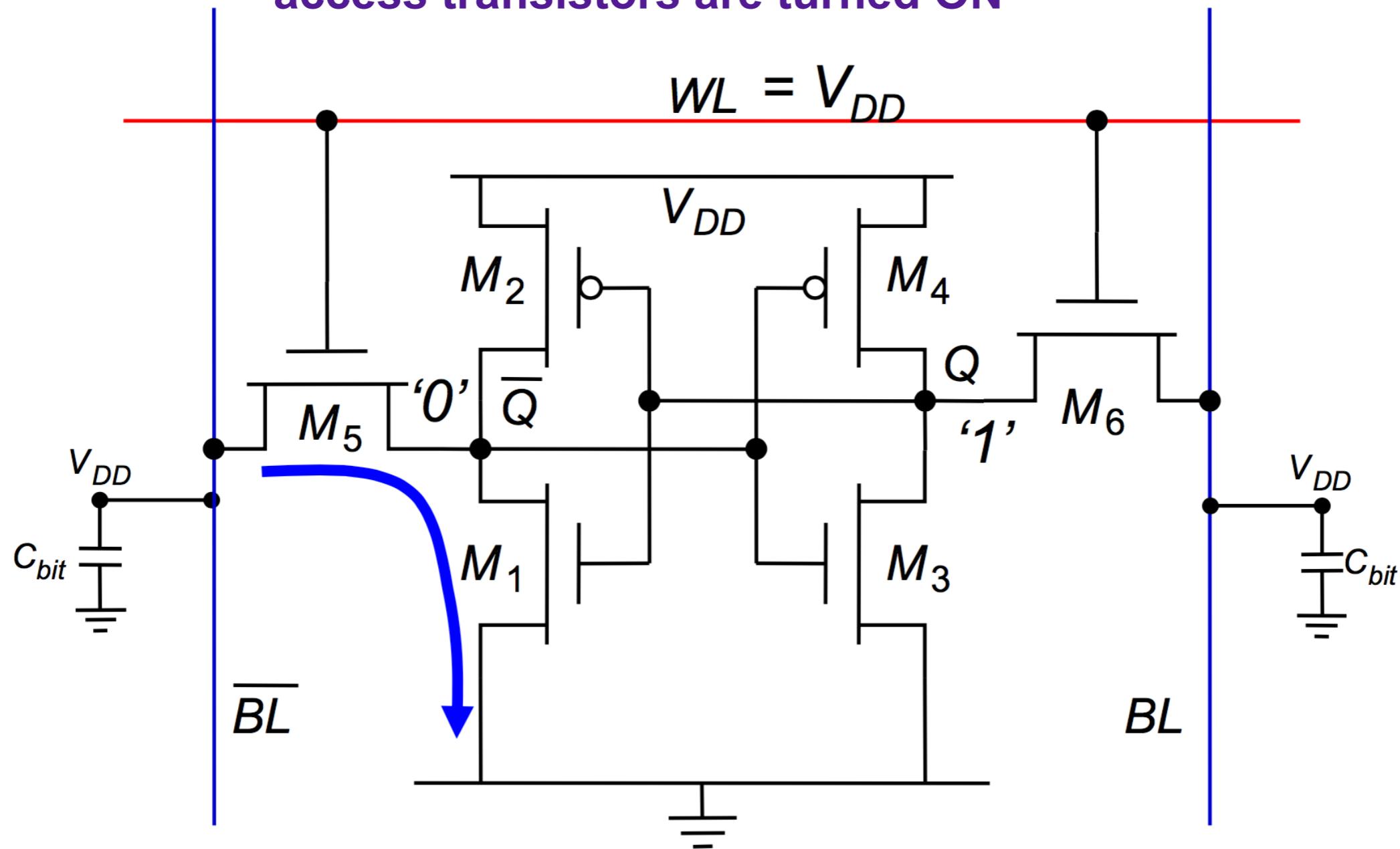
# READ analysis



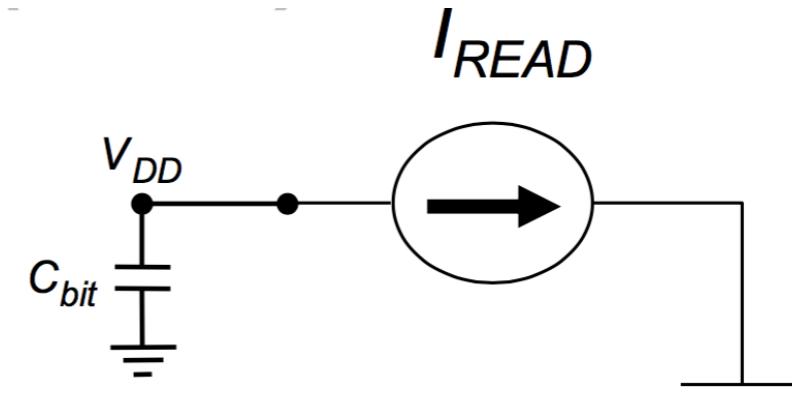
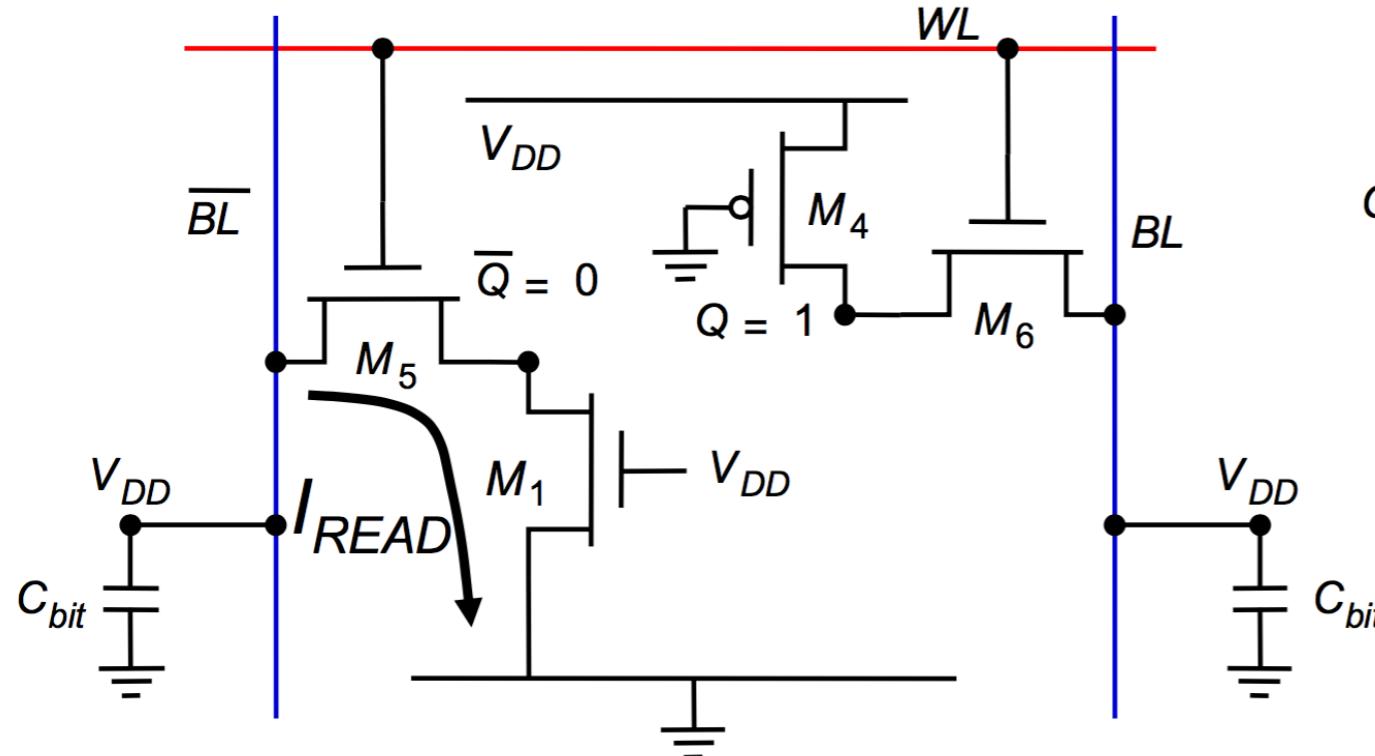
- During pre-charge, both bit line and its complement are charged to VDD.

# READ analysis

Read phase ( $\text{PRE} = 1, \text{WL} = 1$ )  $\rightarrow$   
access transistors are turned ON



# READ analysis



$$C_{bit} \frac{dV_{bl}}{dt} + I_{READ} = 0$$

$$C_{bit} \frac{\Delta_{BIT}}{I_{READ}} = T_{access}$$

**I<sub>READ</sub>** = read current

**T<sub>access</sub>** = Access Time

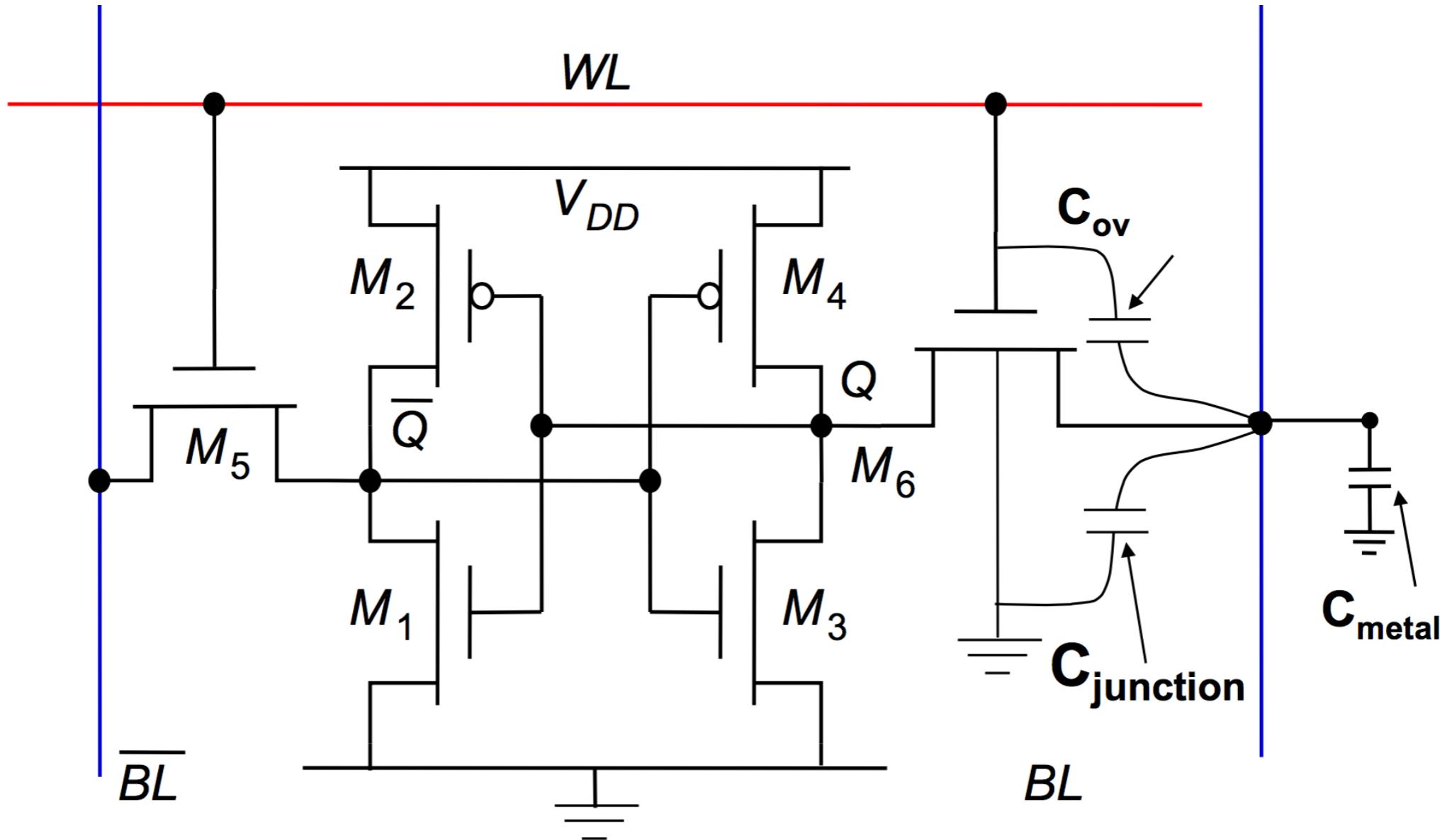
**M<sub>1</sub>** in linear region

**Δ<sub>BIT</sub>** = bit - differential

Assume, bitline voltage drops slowly

=> **M<sub>5</sub>** is in saturation

# Bitline capacitances



$$C_{bit\_percell} = (C_{junction} + C_{ov} + C_{metal})$$

# Bitline capacitances

- Junction capacitance of access transistors
- Gate-to-drain overlap capacitances
- Metal capacitance of the bit-line

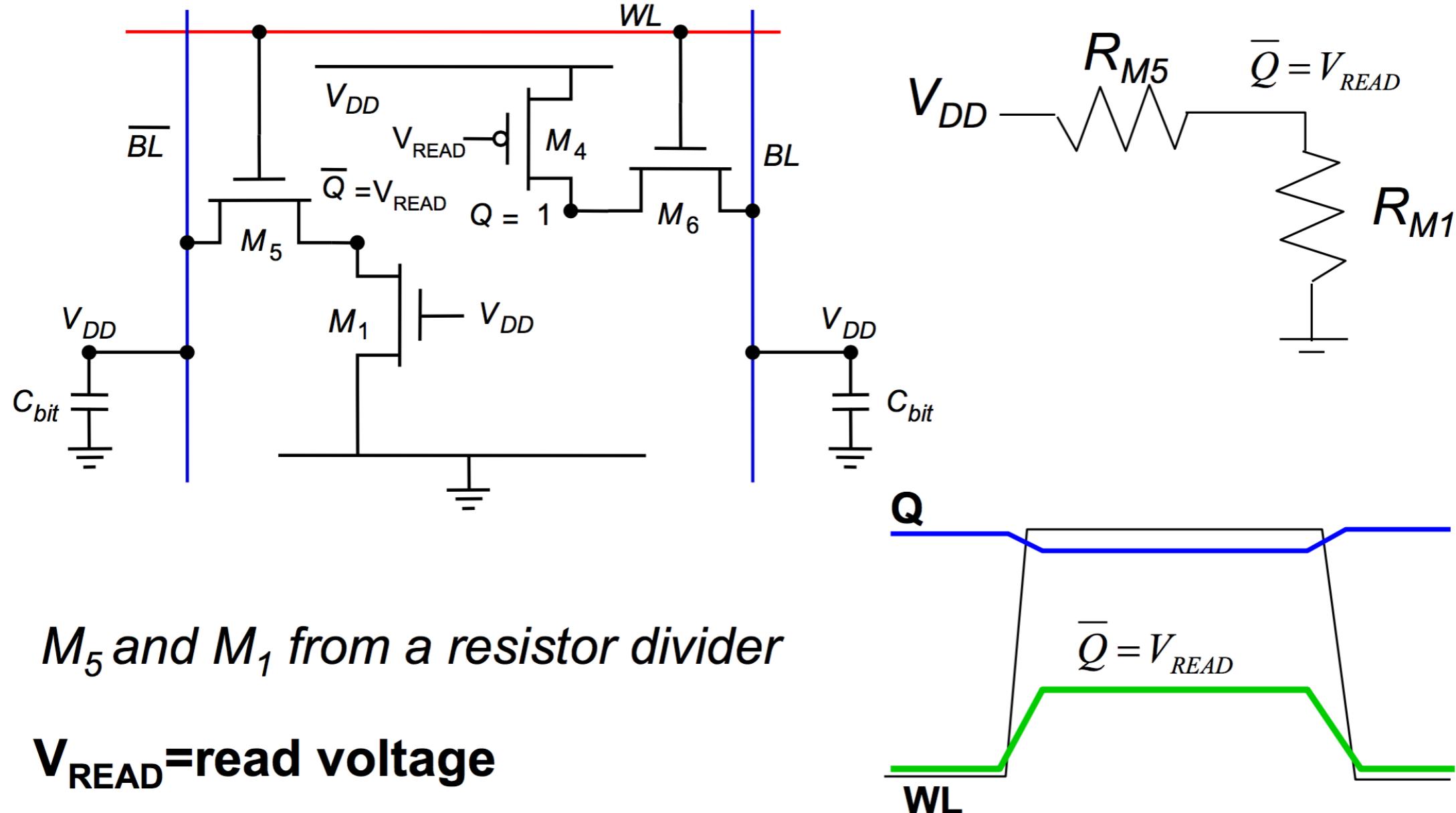
$$C_{bit} = N_{row} \left( C_{junction} + C_{ov} + C_{metal} \right)$$

$C_{junction}$  = junction cap of each access device

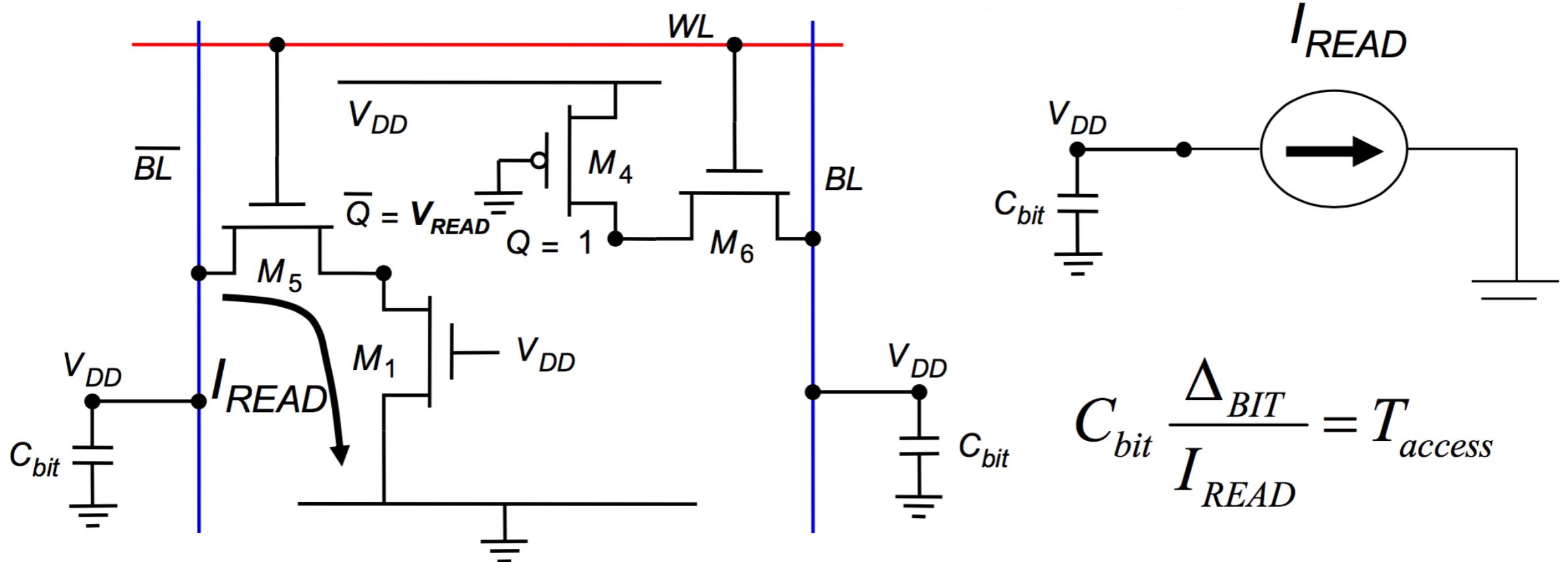
$C_{ov}$  = gate – to – drain overlap cap of each access device

$C_{metal}$  = metal cap of per cell

# READ analysis



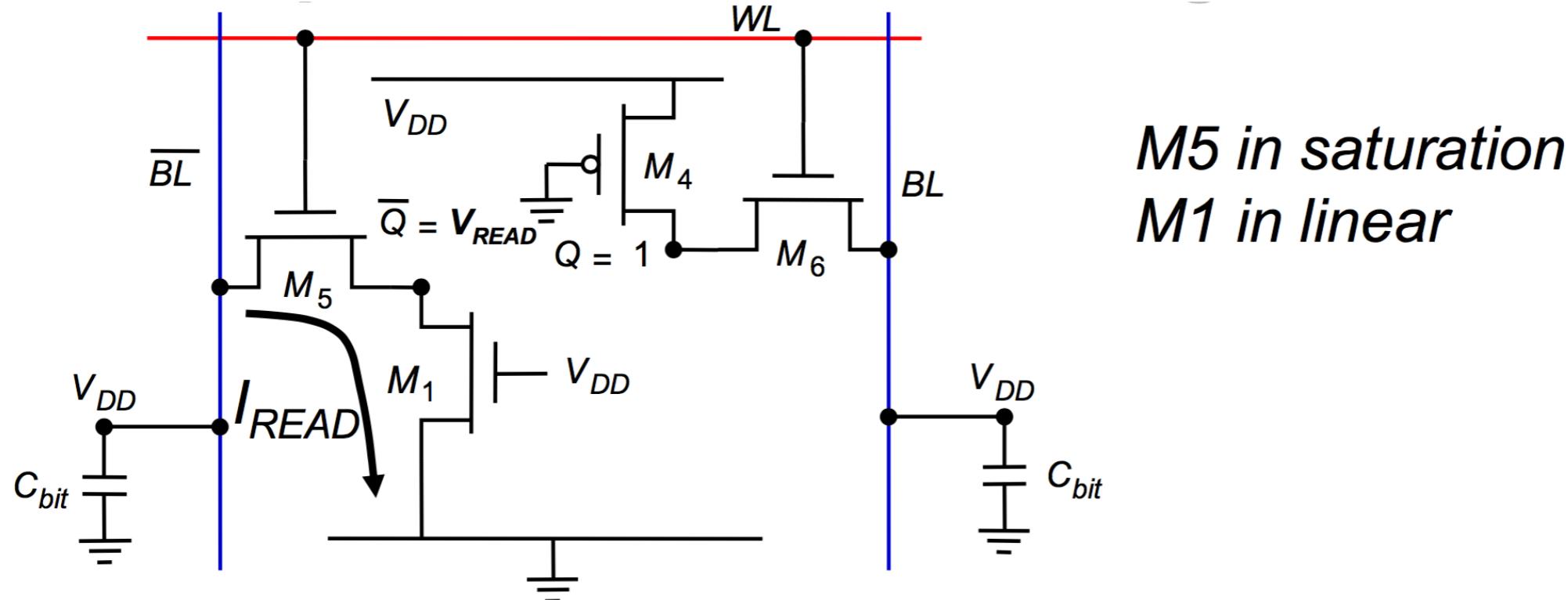
# READ analysis



$$I_{READ} = \mu_n C_{ox} \frac{W_{M5}}{2L_{M5}} (V_{DD} - V_{READ} - V_{th})^2 = 0.5 \beta_{M5} (V_{DD} - V_{READ} - V_{th})^2$$

$$T_{ACCESS} = C_{bit} \frac{\Delta_{BIT}}{0.5 \beta_{M5} (V_{DD} - V_{READ} - V_{th})^2}$$

# Calculation of read voltage

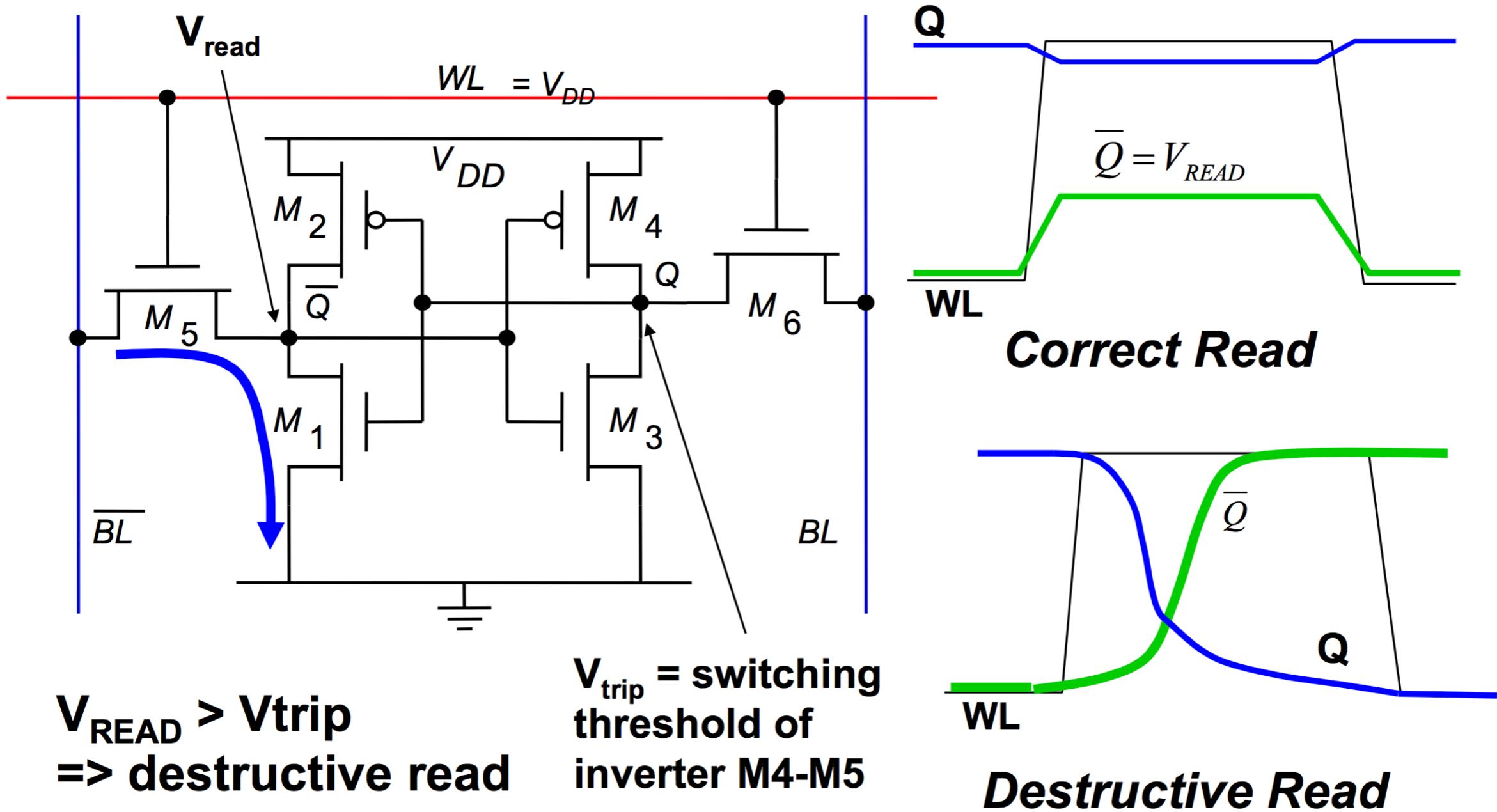


$$0.5\beta_{M5} (V_{DD} - V_{READ} - V_{th})^2 = \beta_{M1} [(V_{DD} - V_{th}) - 0.5V_{READ}] V_{READ}$$

$$\left( 1 + \underbrace{\beta_{M1}/\beta_{M5}}_{\beta_{ratio-pd-ax}} \right) V_{READ}^2 - 2(1 + \beta_{M1}/\beta_{M5})(V_{DD} - V_{th})V_{READ} + (V_{DD} - V_{th})^2 = 0$$

$$V_{READ} = (V_{DD} - V_{th}) \left( 1 - \sqrt{1 - 1/(1 + \beta_{ratio-pd-ax})} \right)$$

# Destructive read



# Noise margin for read

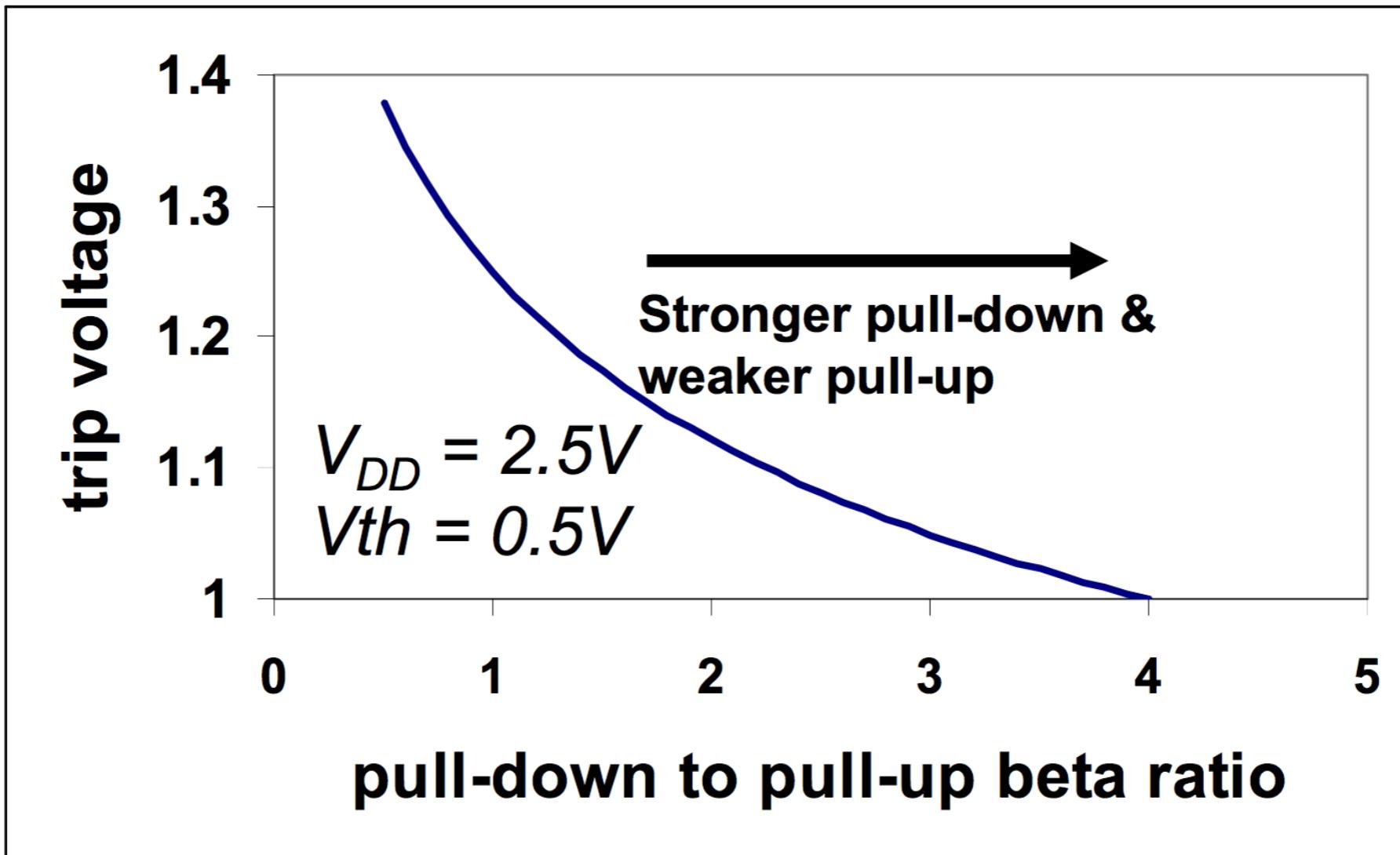
**A simple definition (can be used for first-order design)**

$$\text{Read Margin} = V_{trip} - V_{read}$$

$$V_{trip} = \frac{V_{DD} - |V_{thp}| + \sqrt{\beta_{M3}/\beta_{M4}} V_{thn}}{1 + \sqrt{\beta_{M3}/\beta_{M4}}} \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{trip} \\ \text{stronger pull-up} \Rightarrow \text{higher } V_{trip} \end{array} \right\}$$
$$= \frac{V_{DD} - |V_{thp}| + \sqrt{\beta_{ratio-pd-pup}} V_{thn}}{1 + \sqrt{\beta_{ratio-pd-pup}}} \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{trip} \\ \text{stronger pull-up} \Rightarrow \text{higher } V_{trip} \end{array} \right\}$$

$$V_{READ} = (V_{DD} - V_{thn}) \left( 1 - \sqrt{1 - \frac{1}{1 + \beta_{ratio-pd-ax}}} \right) \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{read} \\ \text{stronger access} \Rightarrow \text{higher } V_{read} \end{array} \right\}$$

# Trip voltage



- A stronger pull-up device increases  $V_{trip}$ .
- A weaker pull-down device increases  $V_{trip}$ .

# Read access speed

For lower read access time:

**Higher  $I_{read}$**

- ⇒ stronger access & stronger pull-down
- ⇒ Pull-up device does not contribute to read speed
- ⇒ Pull-up can be made weaker (i.e. of smaller width) which will also reduce area

# Read noise margin

**For better read margin:**

**Smaller  $V_{read}$  => Higher  $\beta_{ratio-pd-ax}$**

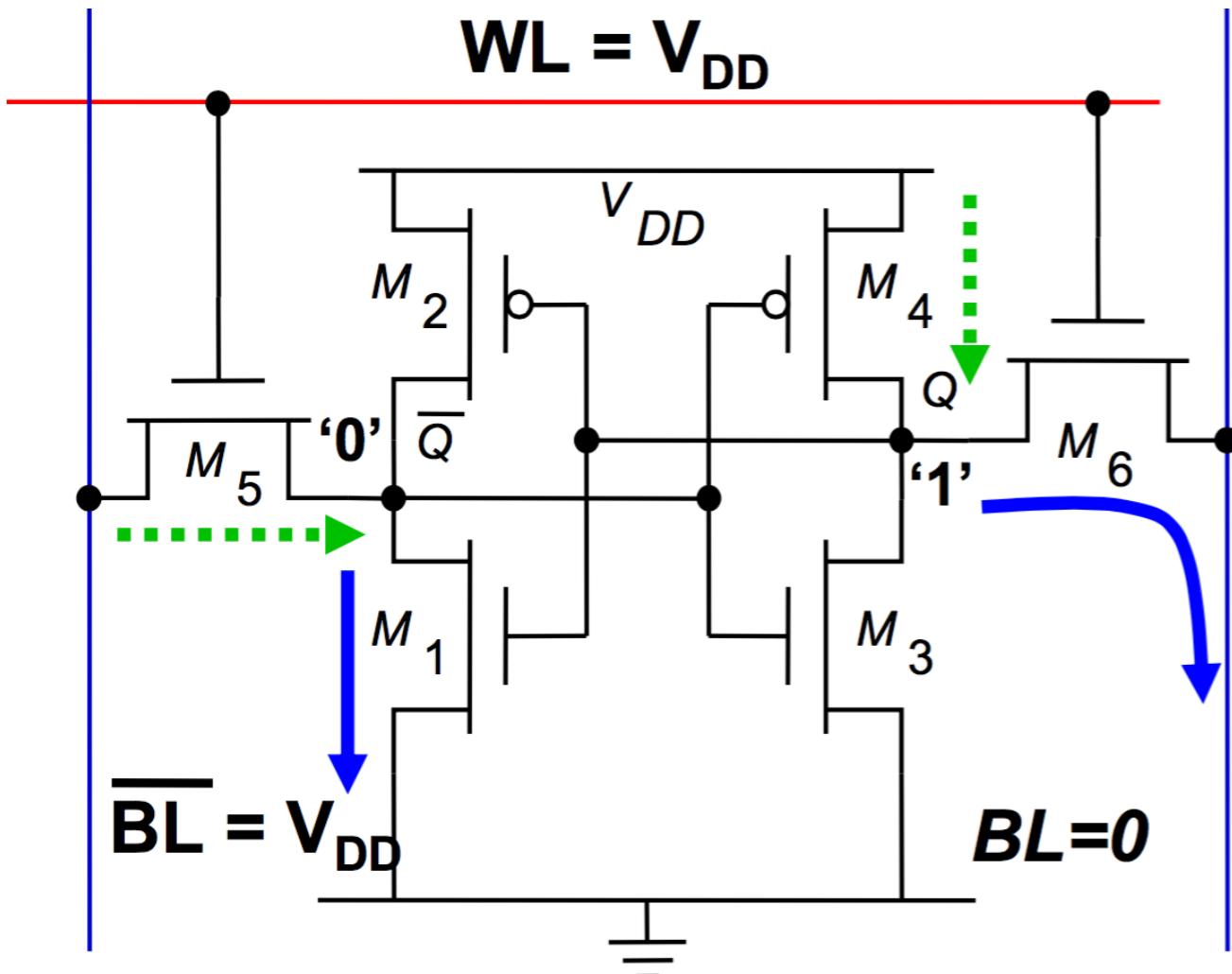
**=> weaker access & stronger pull-down**

**Higher  $V_{trip}$  => Lower  $\beta_{ratio-pd-pup}$**

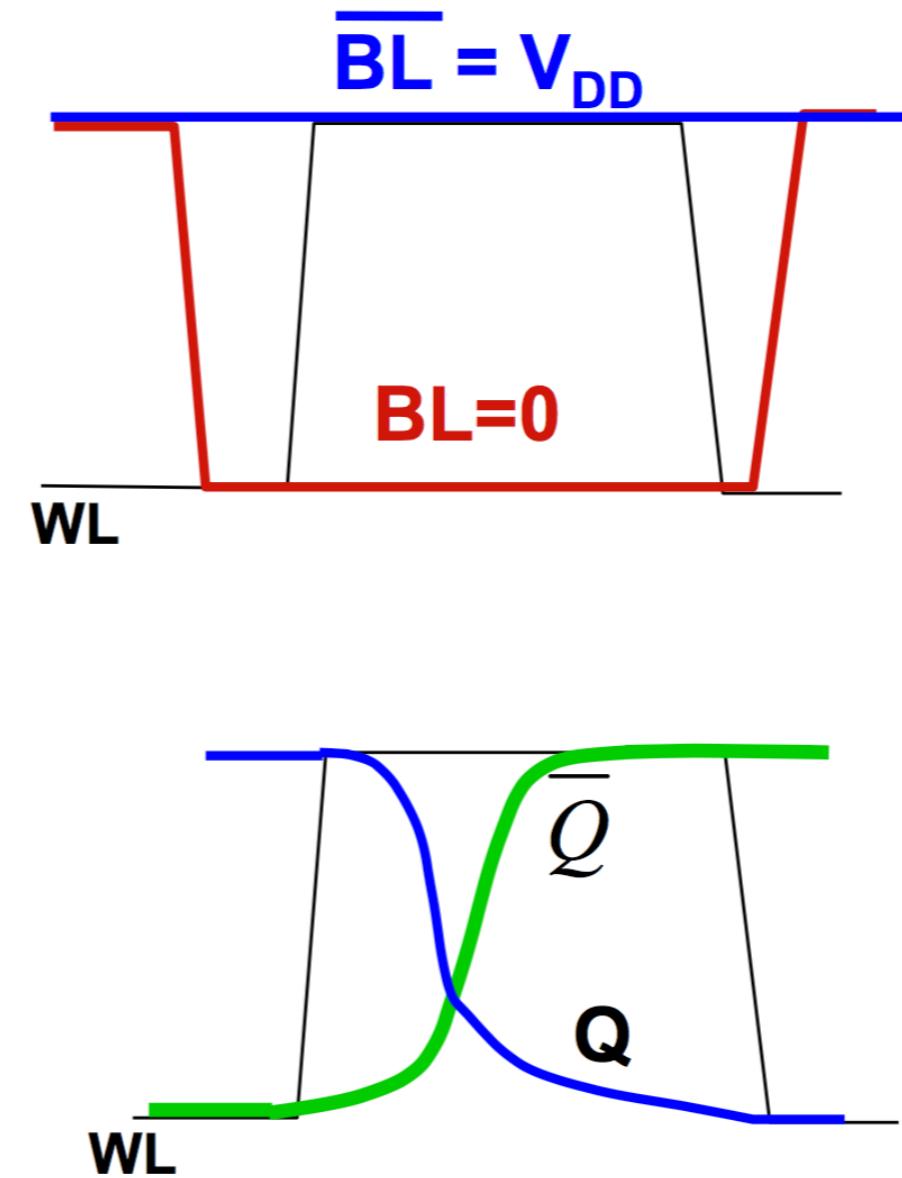
**=> weaker pull-down & stronger pull-up**

- **Read margin requirement for access device contradicts with read speed requirement**
- **Read voltage requirement for pull-down device contradicts trip point requirement**
  - Read voltage is more sensitive to pull-down strength compared to trip voltage
  - Generally a larger pull-down helps read margin but it cannot be too large

# Write operation



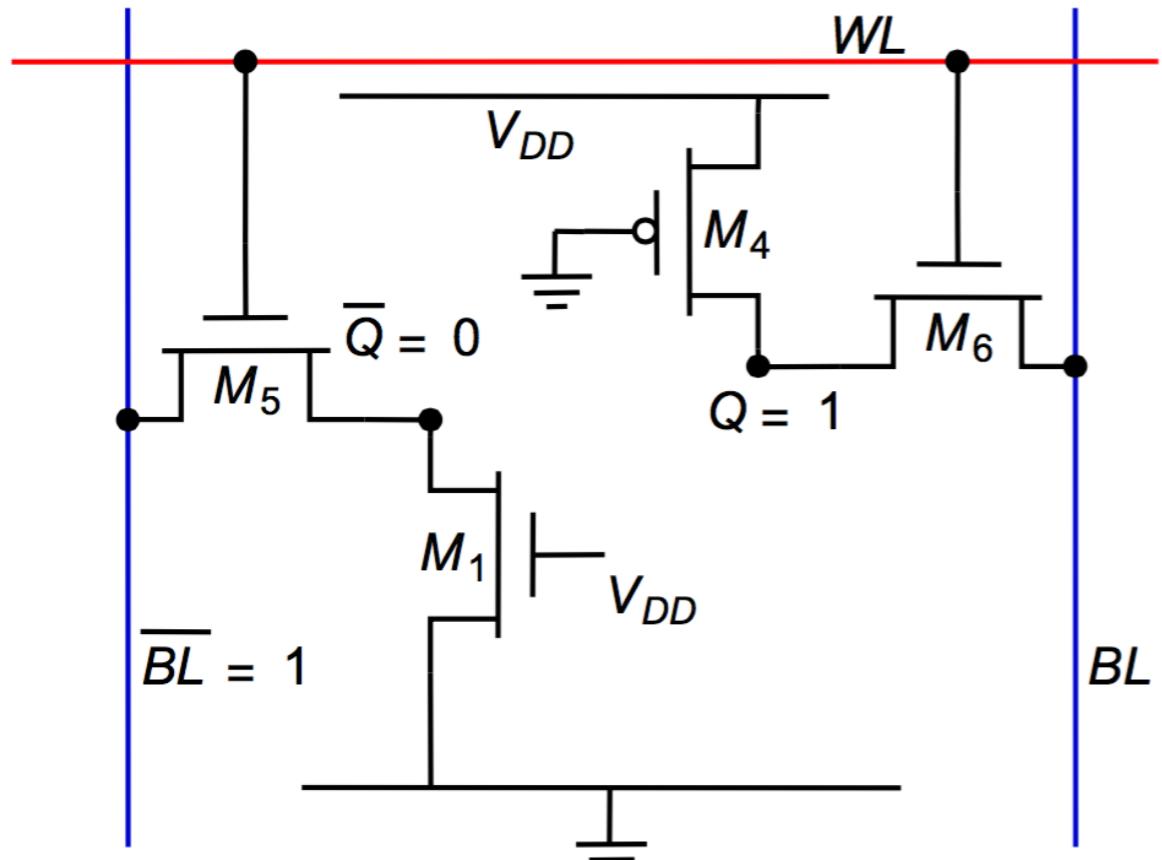
.....→ charging current  
→ dis-charging current



# Write operation: key

- The major writing process is discharging the node storing ‘1’ to ‘0’.
- During this discharging process (initially) the “ON” pull-up PMOS fights with the “ON” access device.
- We need to make sure the node storing ‘1’ can be reduced below the switching threshold of the other inverter.

# Write operation analysis



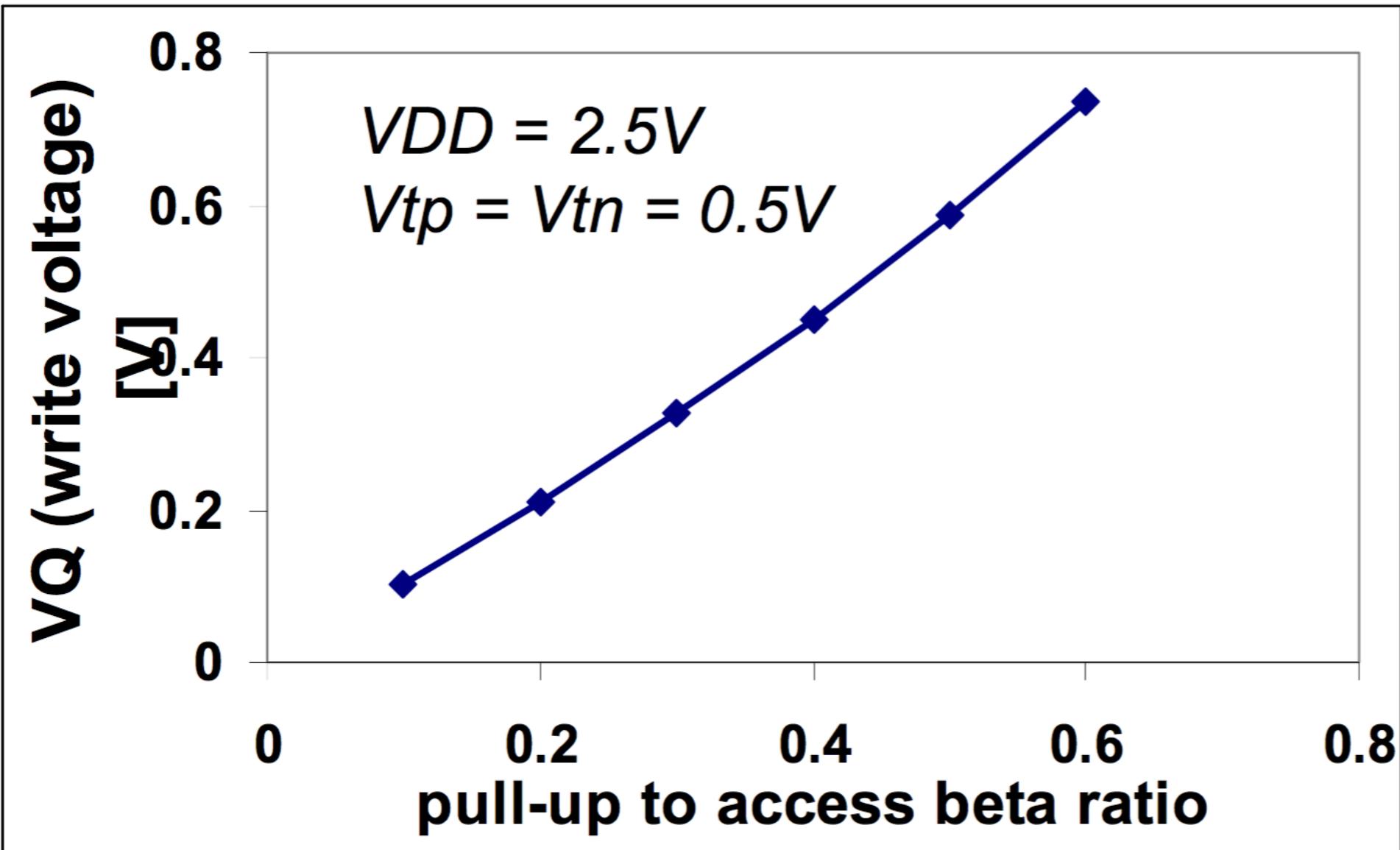
We want to reduce Q to a very low value  $V_Q$   
 $\Rightarrow$  Assume  $V_Q$  is low enough such that  $M_4$  goes to saturation and  $M_6$  goes to linear

$$0.5\beta_{M4}(V_{DD} - V_{th})^2 = \beta_{M6}[(V_{DD} - V_{th}) - 0.5V_Q]V_Q$$

$$\Rightarrow V_Q^2 - 2(V_{DD} - V_{th})V_Q + \underbrace{\beta_{M4}/\beta_{M6}}_{\beta_{ratio-pup-ax}}(V_{DD} - V_{th})^2 = 0$$

$$\Rightarrow V_Q = (V_{DD} - V_{th}) \left( 1 - \sqrt{1 - \beta_{ratio-pup-ax}} \right)$$

# Write operation: cell sizing



- A stronger pull-up PMOS degrades writability
- A weaker access device degrades writability

# SRAM sizing conflicts

	Read speed	Read noise margin	Writability
Pull-down	Strong	Strong	Less important
Access	Strong	Weak	Strong
Pull-up	Don't care	Strong	Weak

- Careful sizing of each device is necessary to meet all the requirement for cell design
- A good starting point :  $W_p:W_{access}:W_{pull-down} = 1:1.5:2$

# SRAM design parameters

## □ Read Margin

- Read and trip voltage
- Static Noise Margin

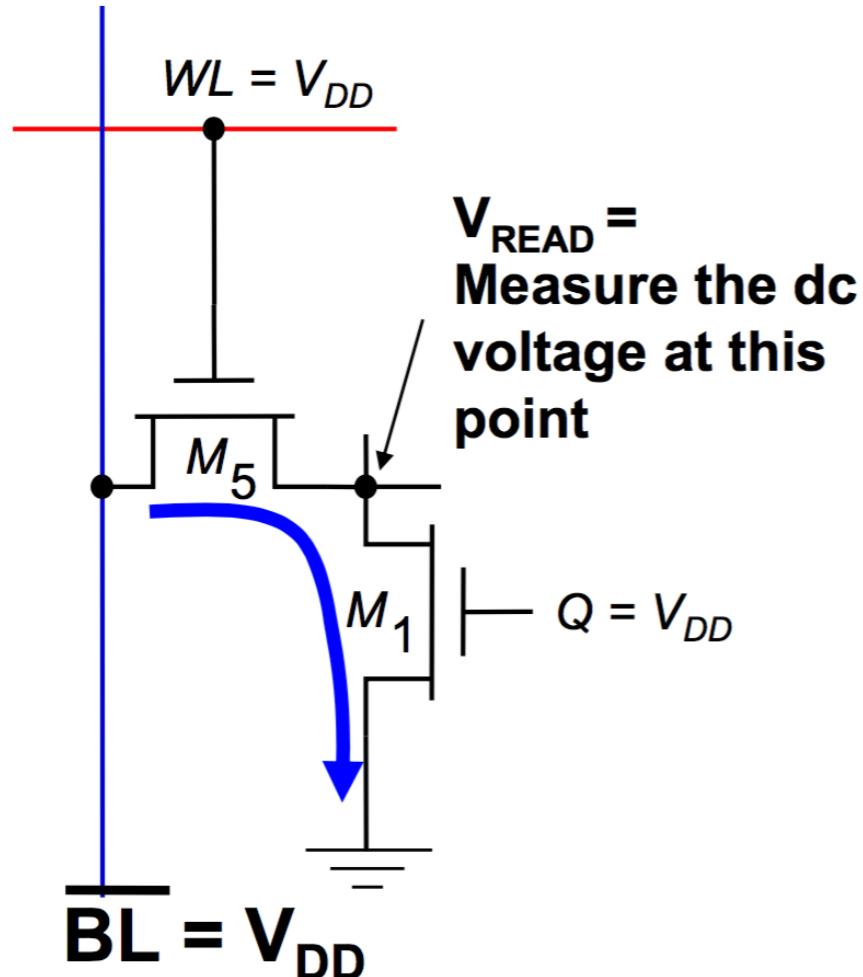
## □ Speed

- Read current
- Access time
- Bit-differential

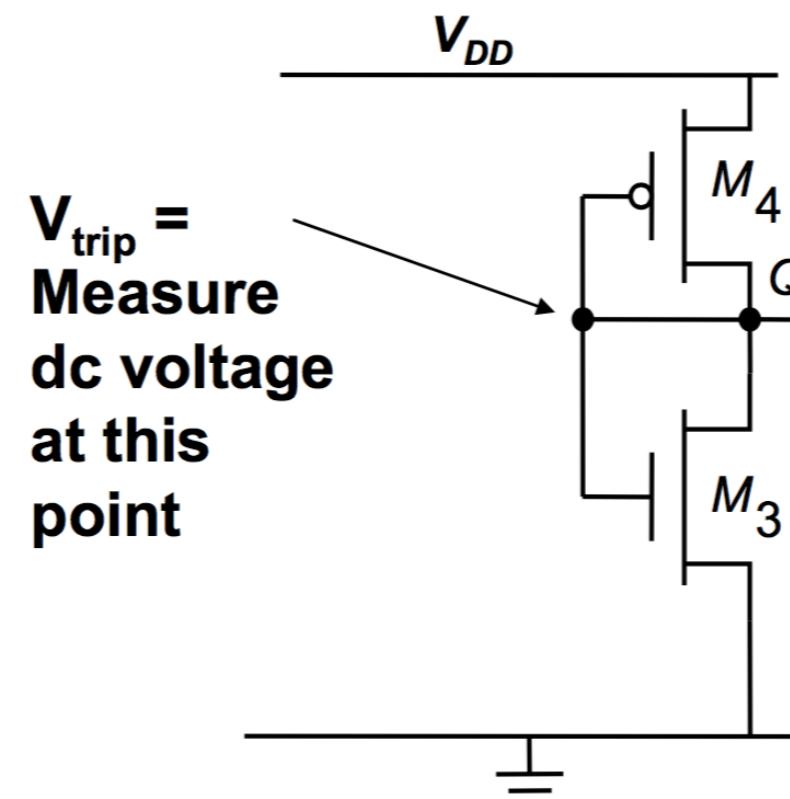
## □ Write

- Write Margin
- Write time

# Cell design parameters: read margin

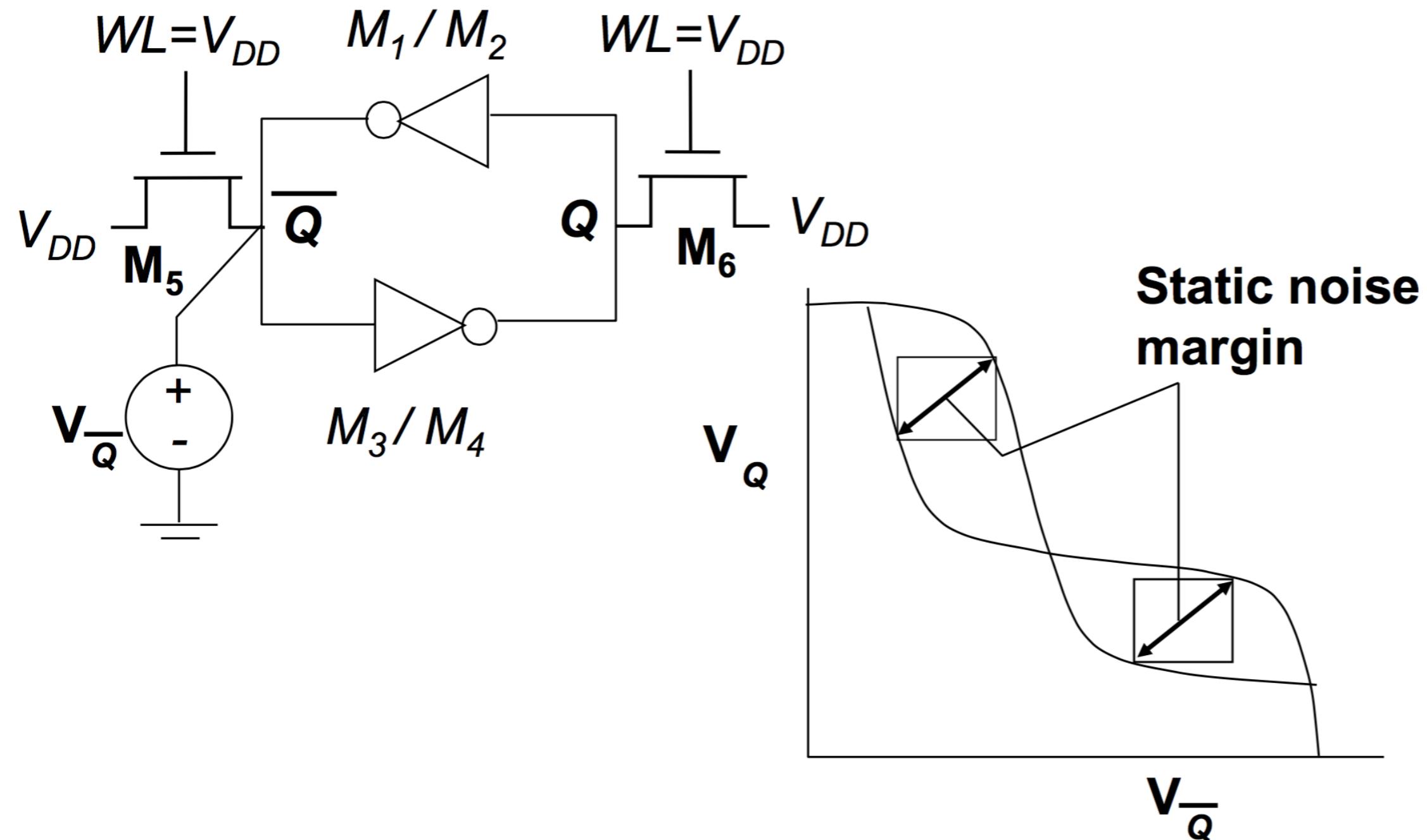


Read voltage estimation

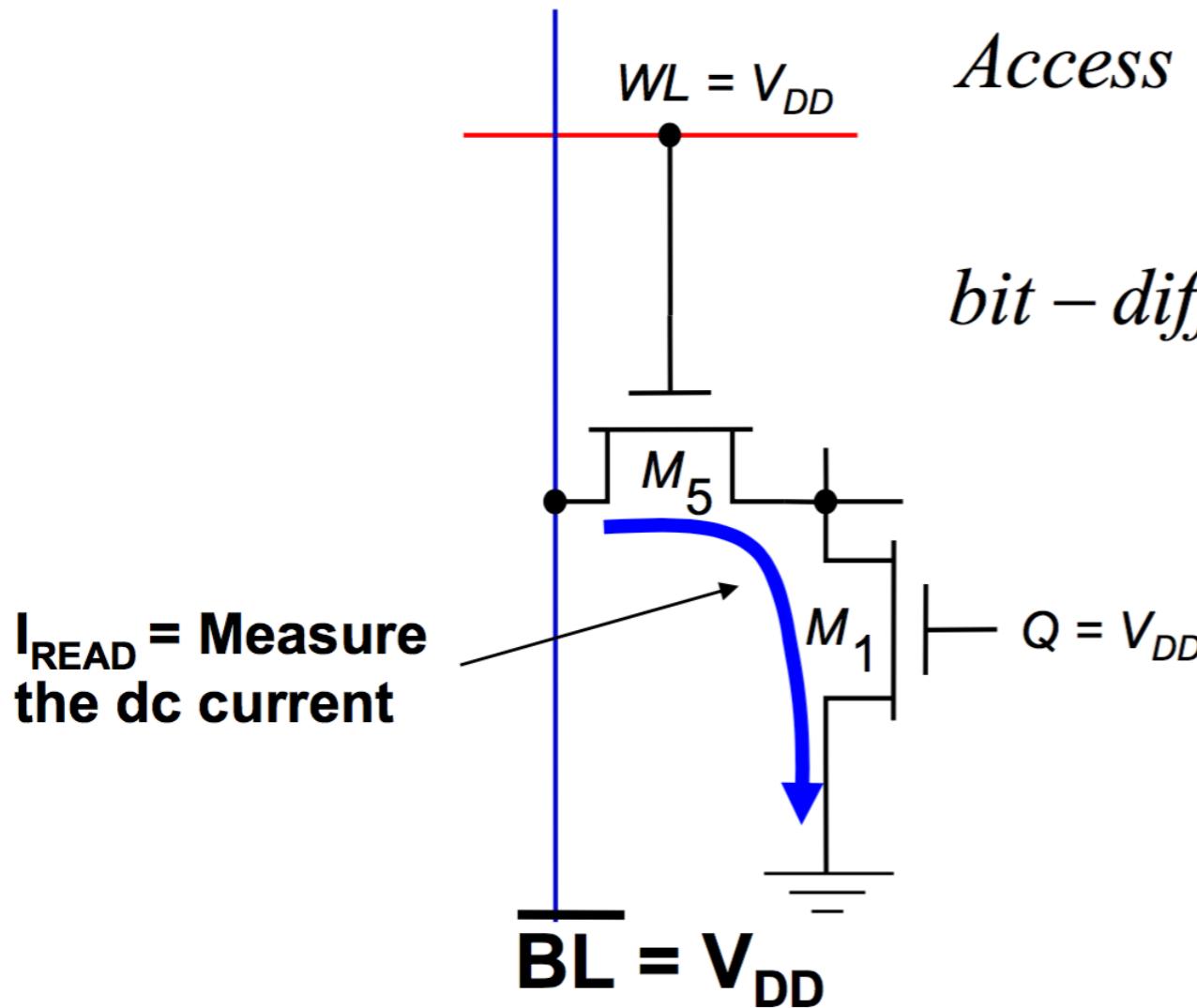


Trip voltage estimation

# A more accurate measure of read noise margin



# Read current

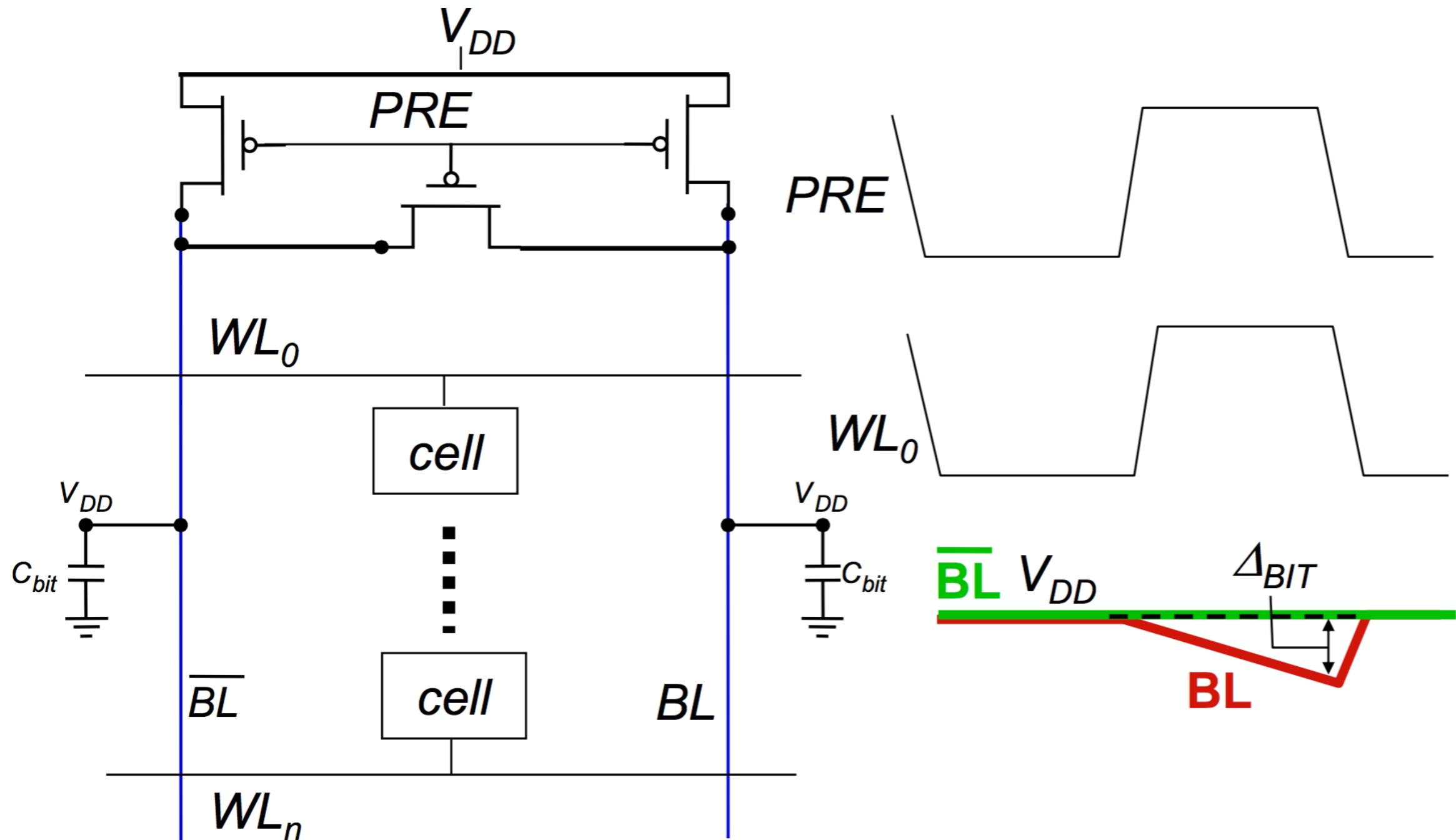


$$\text{Access Time} = T_{access} = C_{bit} \frac{\Delta_{BIT}}{I_{READ}}$$

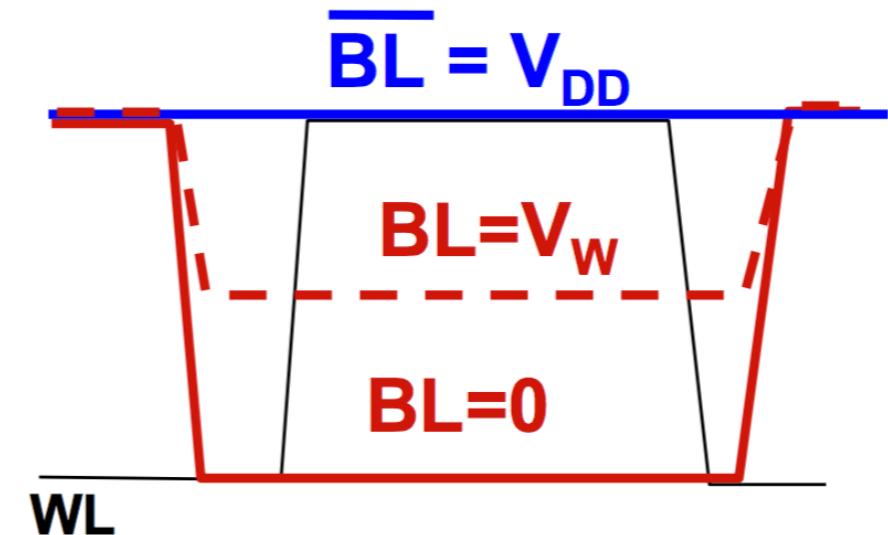
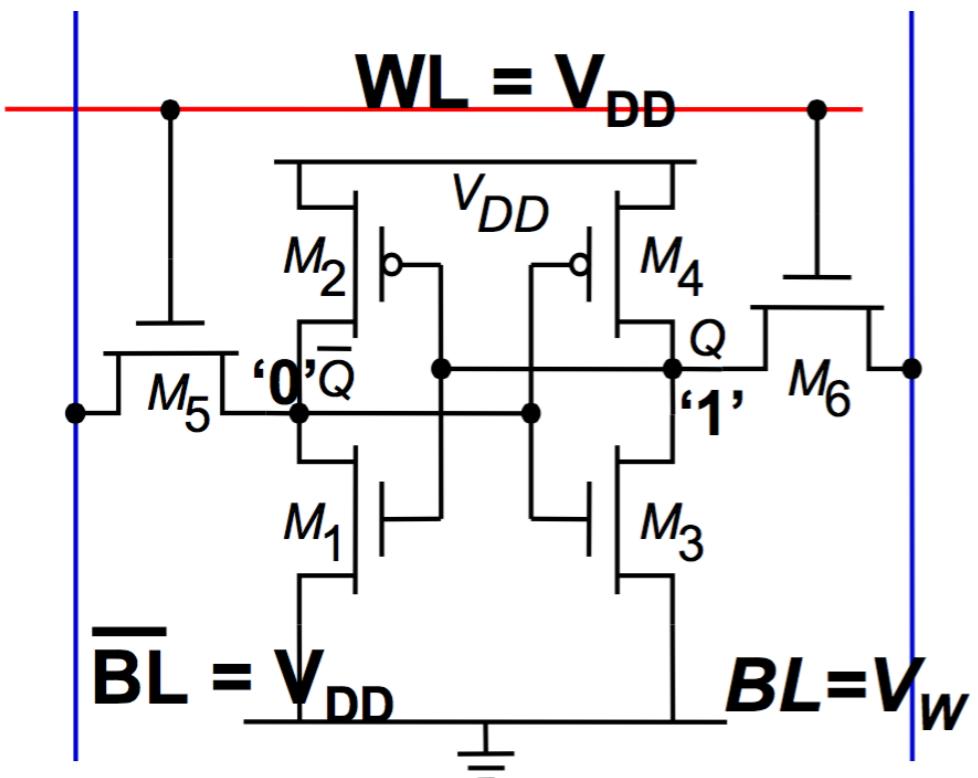
$$\text{bit-differential} = \Delta_{BIT} = \frac{I_{READ} T_{access}}{C_{bit}}$$

**Access time or bit differential estimation**

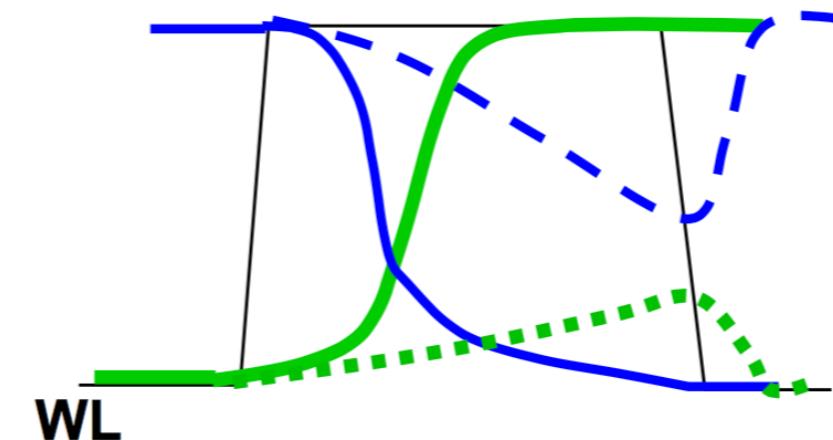
# Bit differential



# Write margin



**Write Margin** = Increase  $V_W$  from 0 to  $V_{DD}$  and find out the maximum value of  $V_W$  till which you can write to the cell



A larger value of write margin ensures a more reliable write operation

# Peripheral circuits

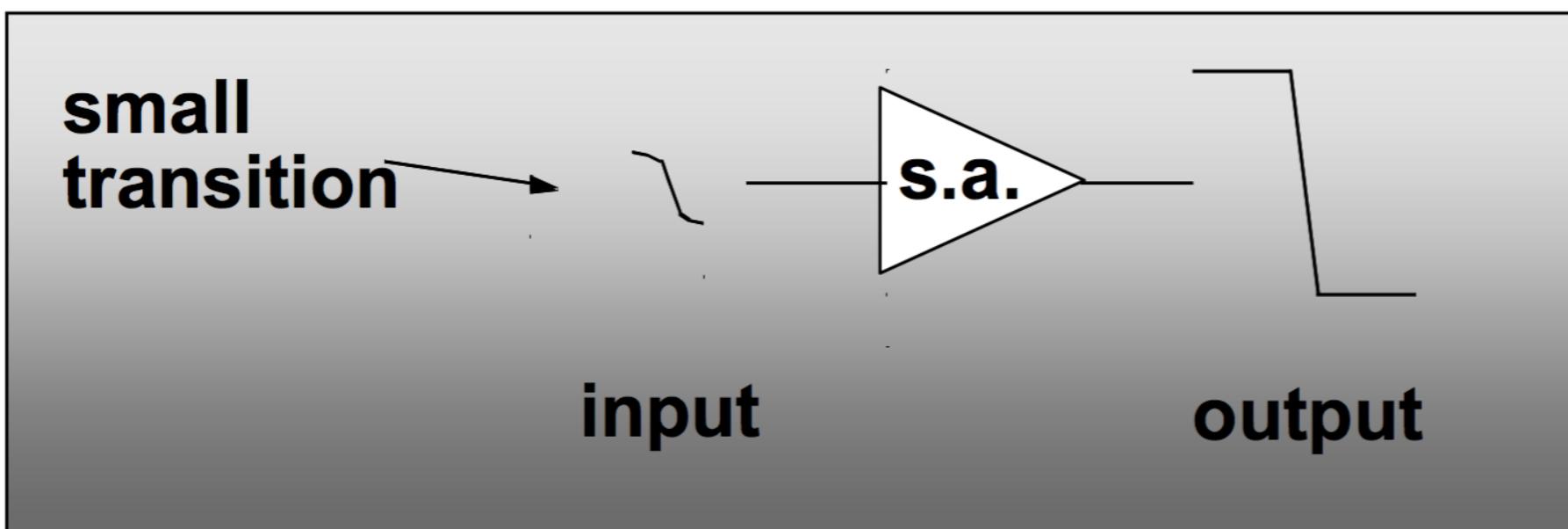
- Decoders
- Column multiplexers
- Read circuits
  - Pre-charge circuits
  - Sense amplifiers
  - Timing requirements
- Write circuits

# Sense amplifier

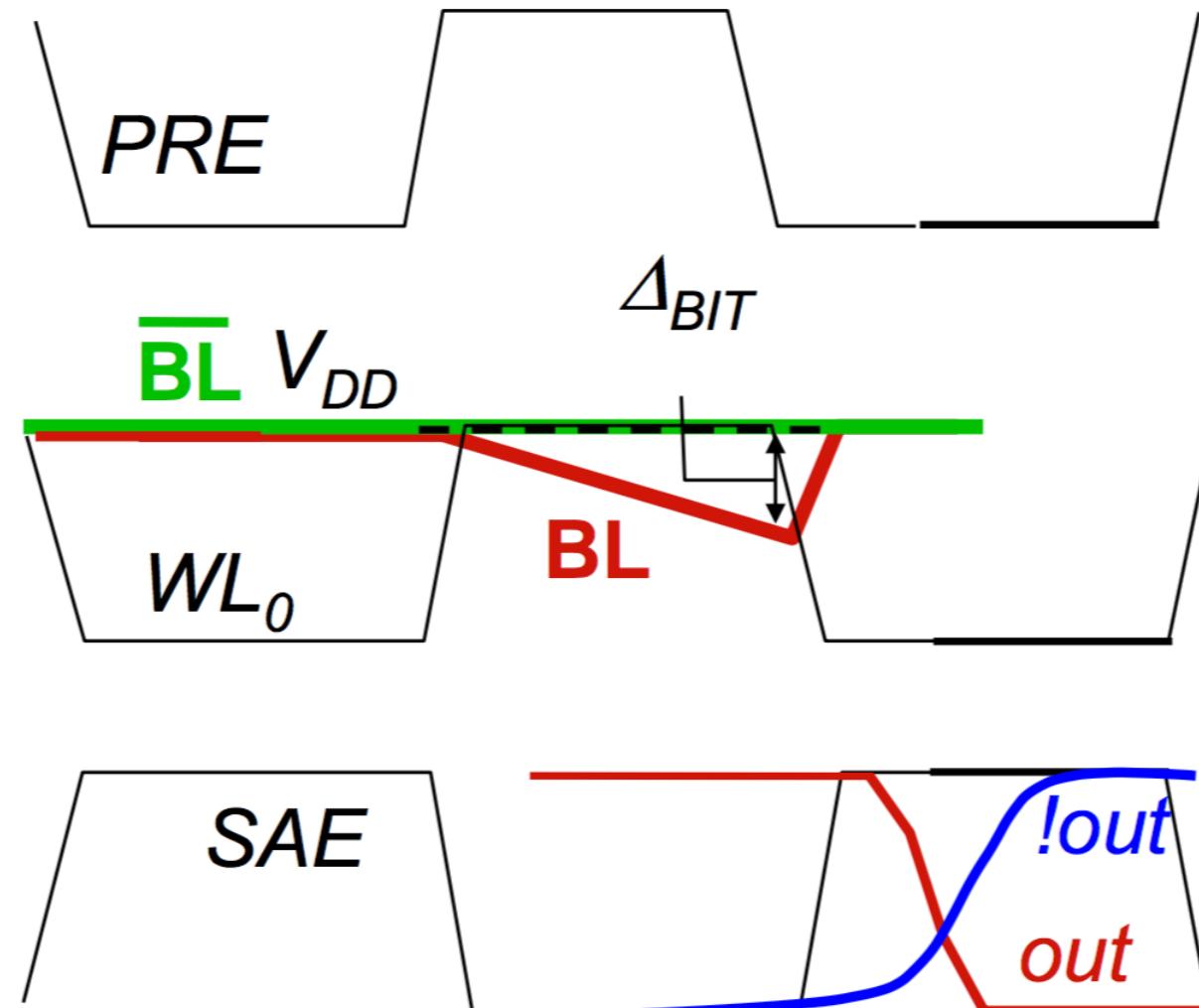
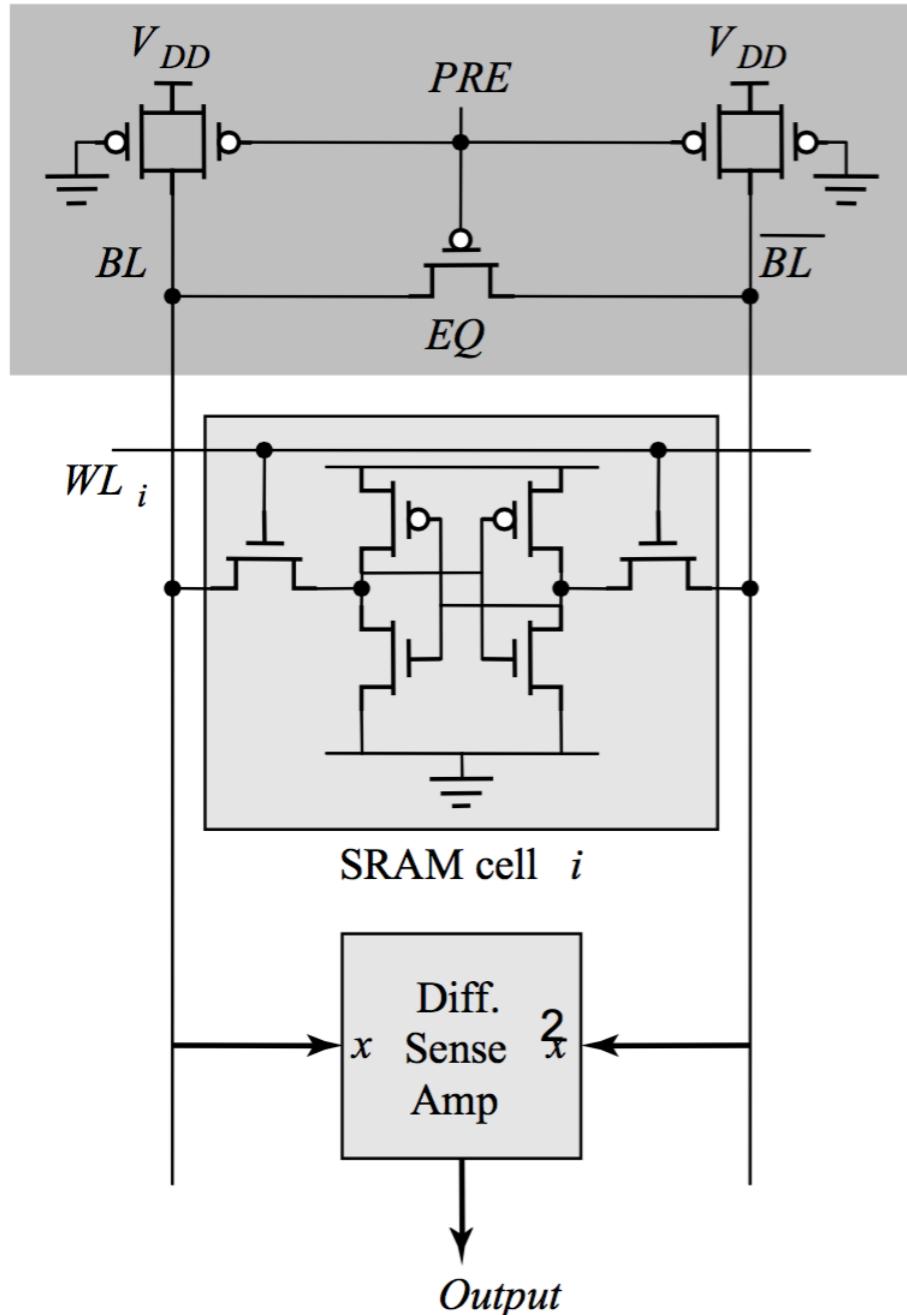
$$t_p = \frac{C \cdot \Delta V}{I_{av}}$$

make  $\Delta V$  as small as possible

# Idea: Use Sense Amplifier

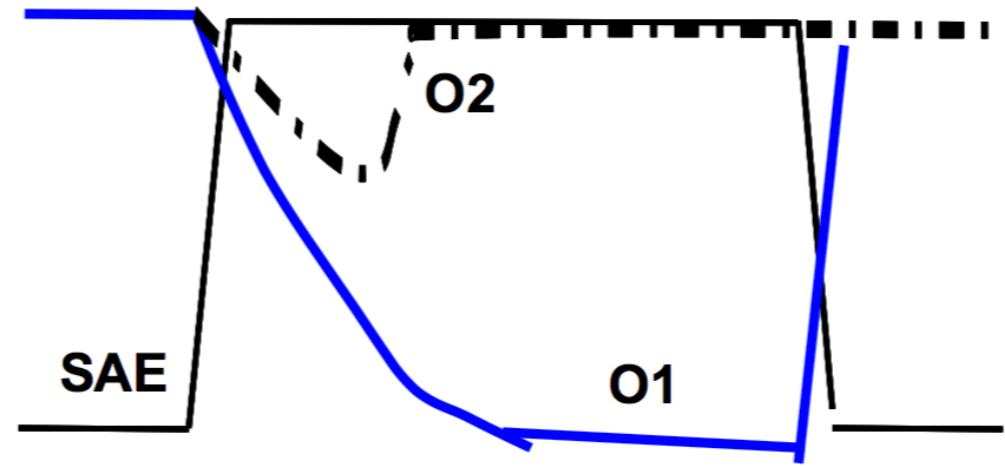
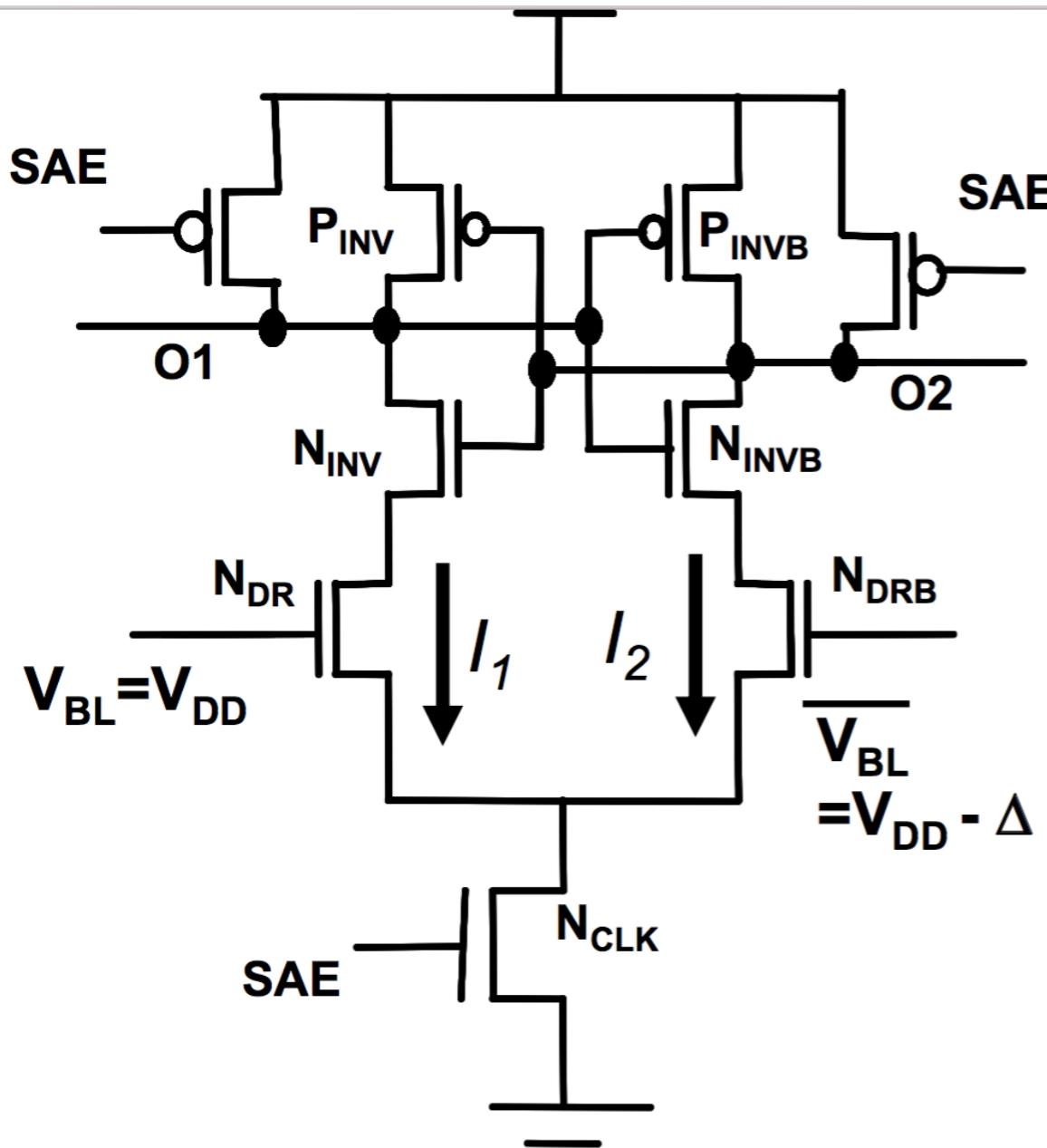


# Sense amplifier



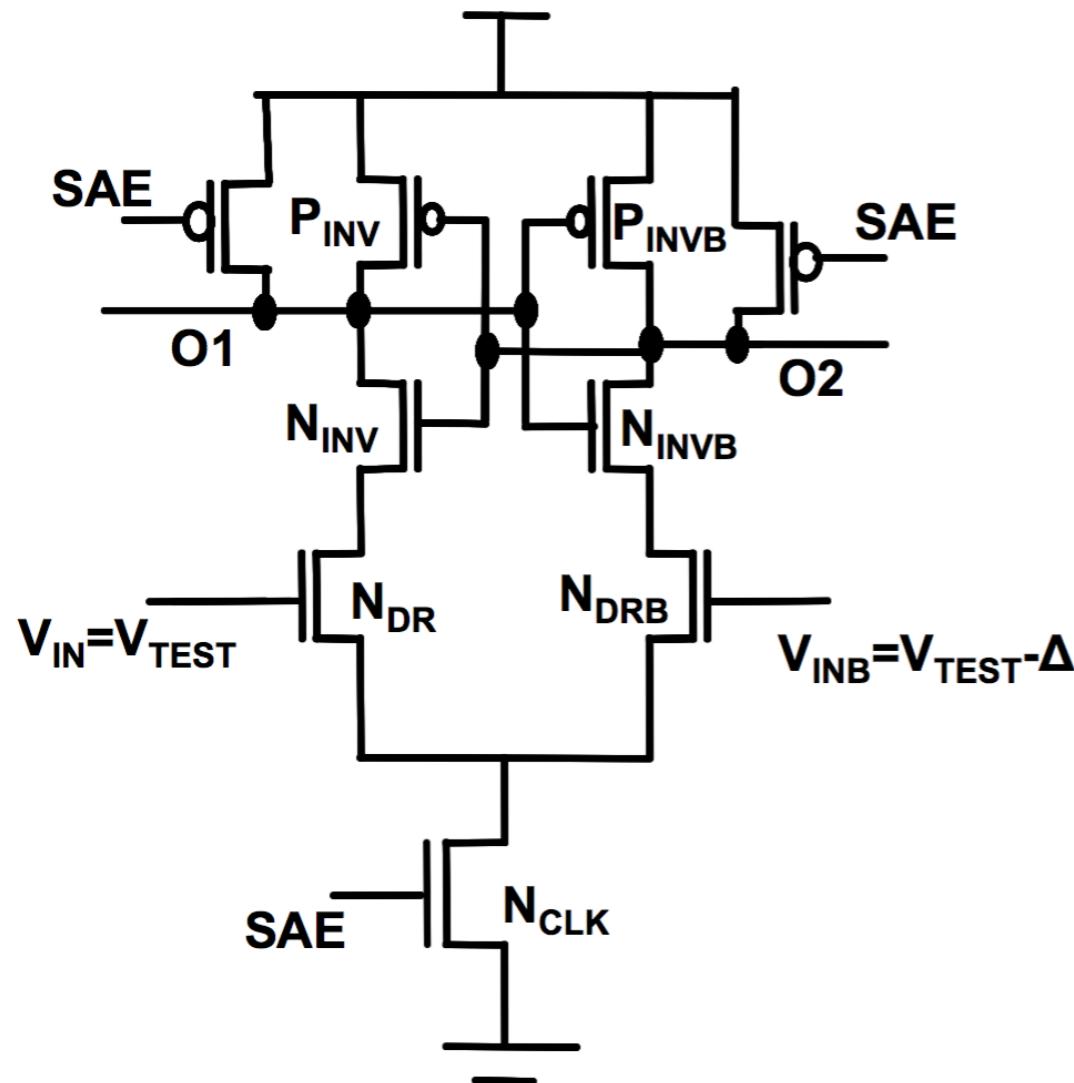
SRAM sensing scheme

# Current latch based sense amp.

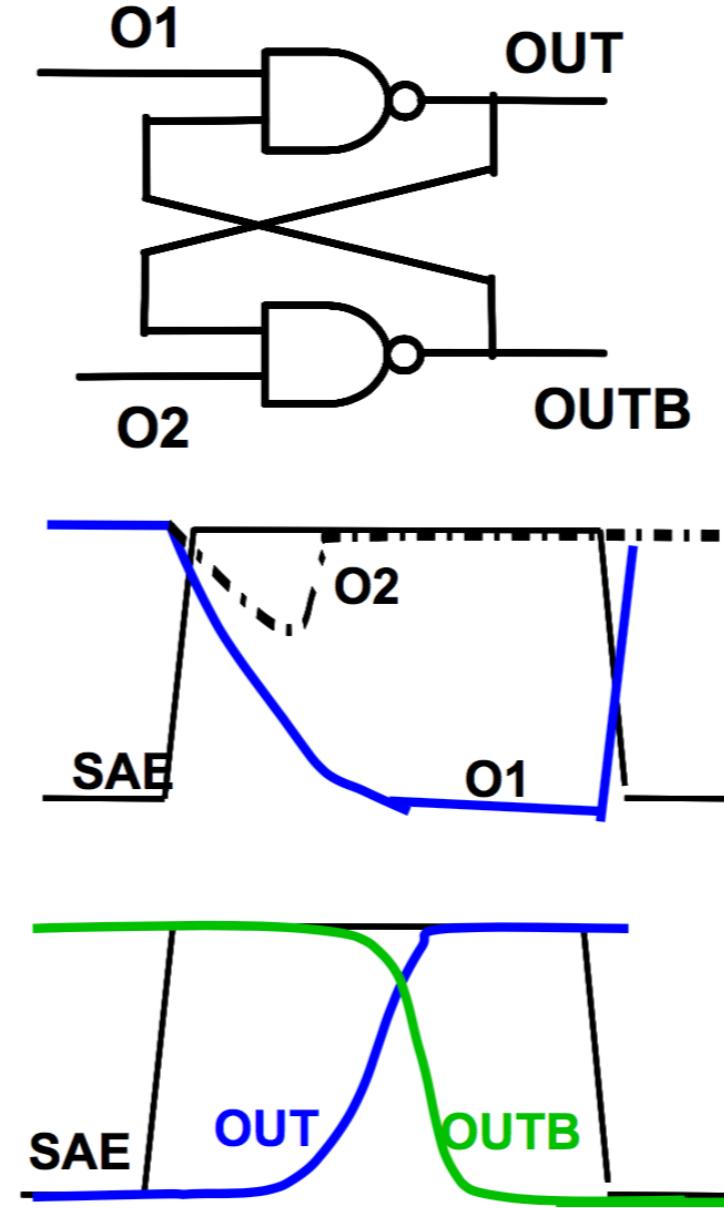


1. **SAE = low => O1 and O2 are precharged to  $V_{DD}$**
2. **SAE = high => both O1 and O2 starts to discharge**
3. **If  $V_{BL} < V_{BL} - \Delta \Rightarrow I_1 > I_2 \Rightarrow O1$  discharges at a faster rate**
4. **When the voltage diff. between O1 and O2 is high enough the cross-coupled inverter amplifies it**

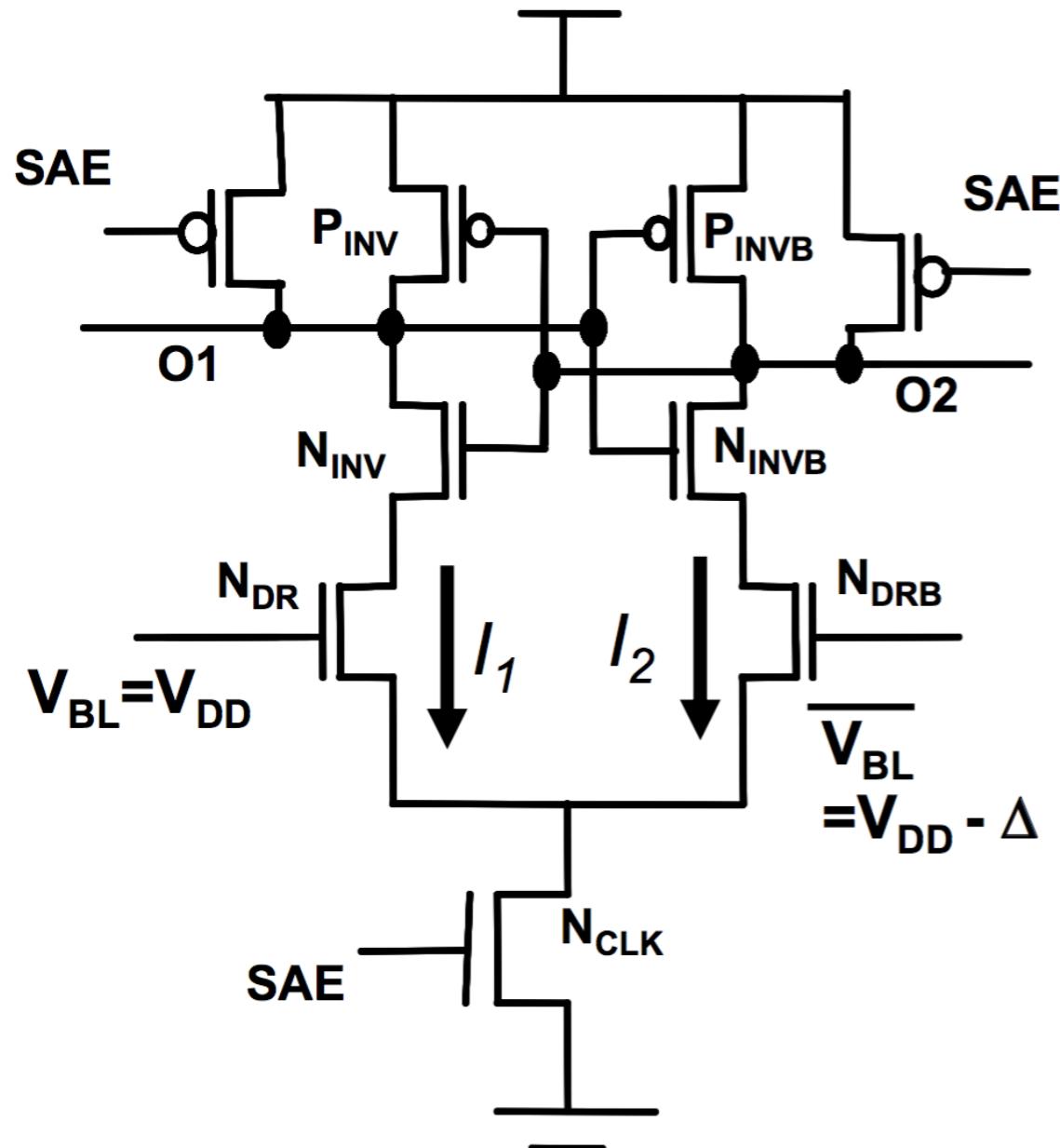
# Current latch based sense amp.



The output latch converts the dynamic O1 & O2 to static signals OUT and OUTB



# Offset voltage



**Assume width of  $N_{DRB}$  is larger than the width of  $N_{DR}$**

$$I_1 \propto W_{DR} (V_{DD} - V_s - V_{th})$$

$$I_2 \propto W_{DRB} (V_{DD} - V_s - \Delta - V_{th})$$

if  $W_{DRB} = W_{DR} + w$

even if  $\Delta > 0$ ,

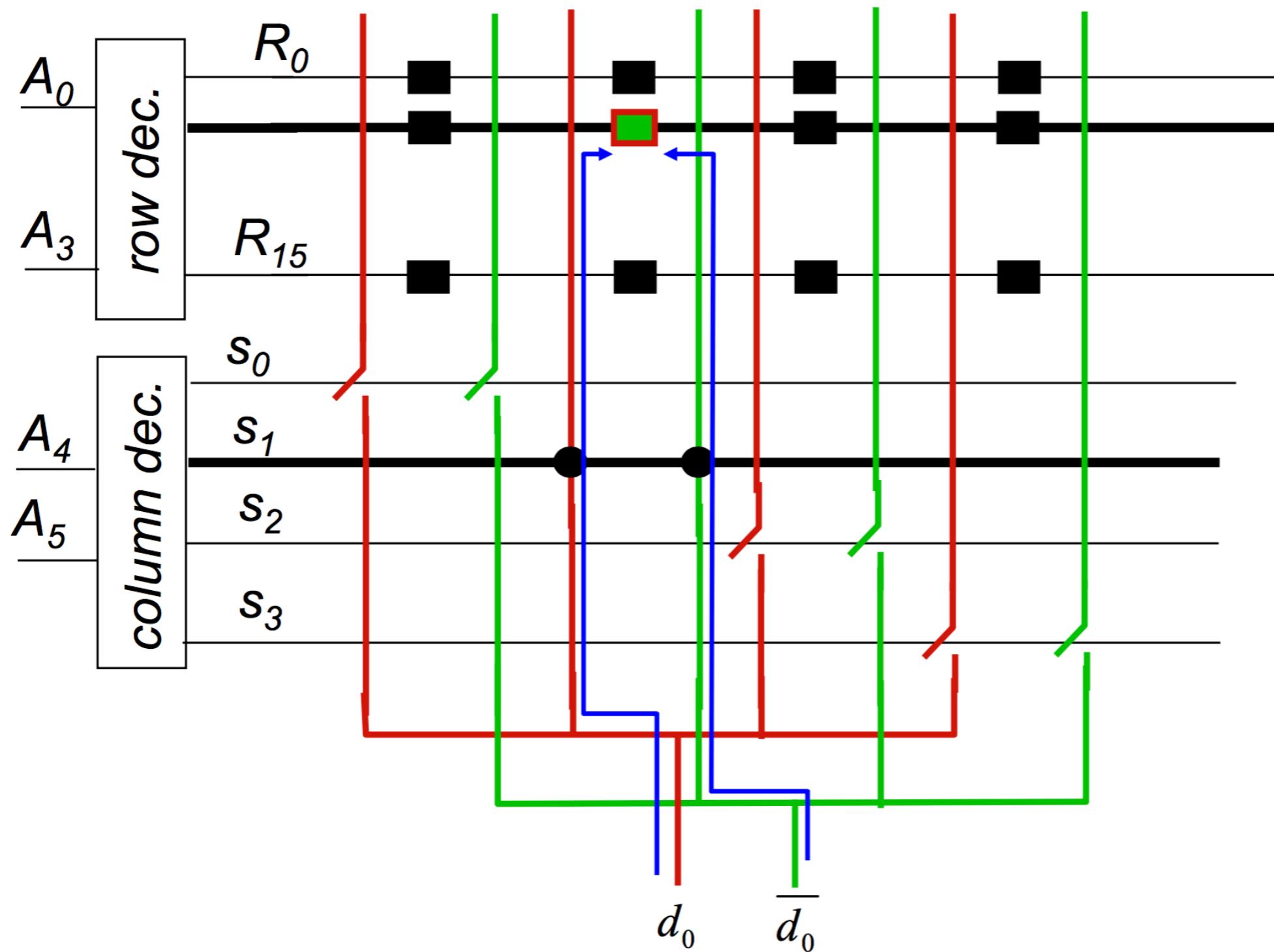
it is possible  $I_2 > I_1$

To have  $I_1 = I_2$  under  $W$  mismatch  
we need to have a minimum  $\Delta$

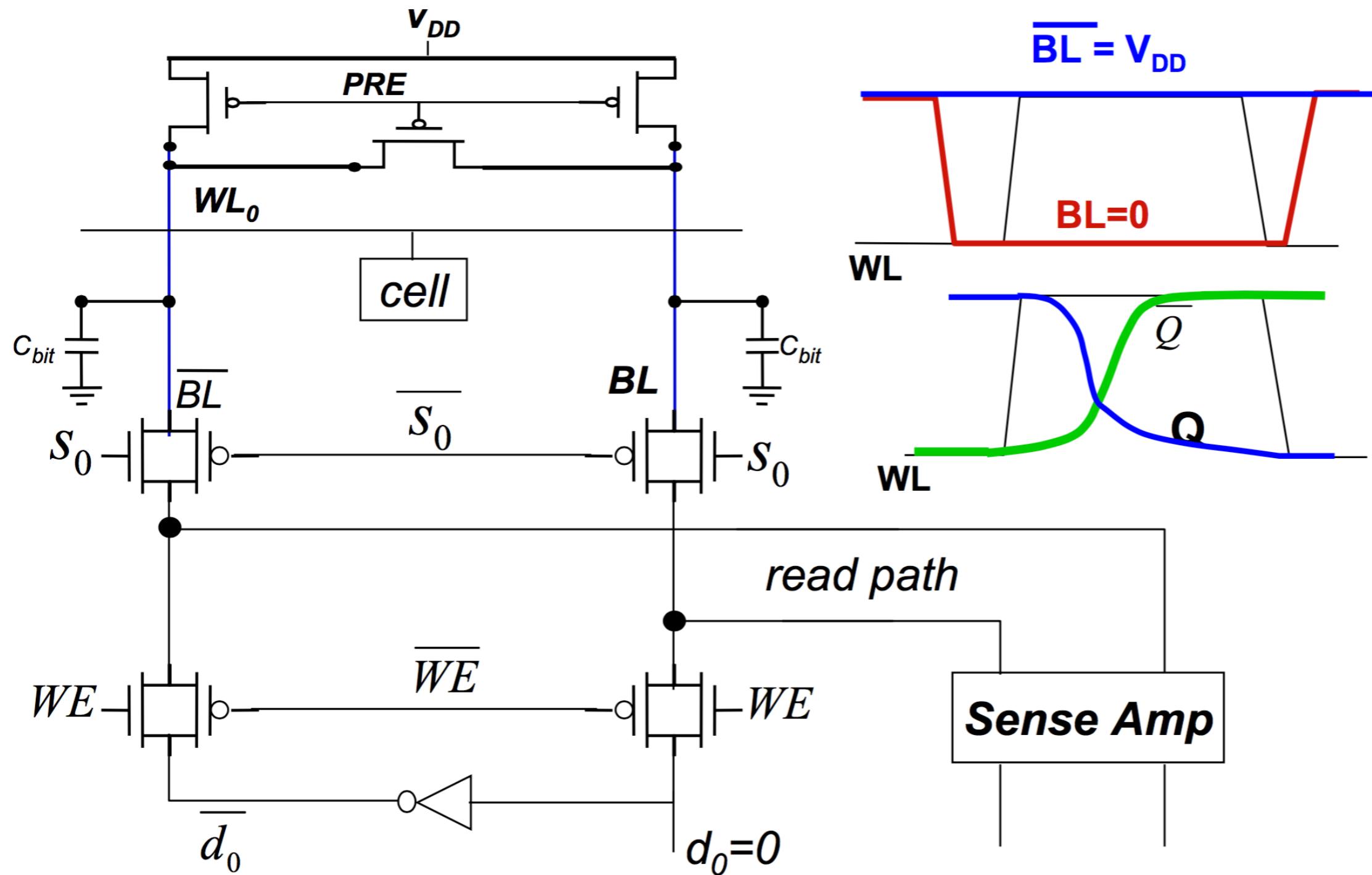
**Minimum input voltage difference required for correct sensing is defined as the offset voltage**

# Write operation

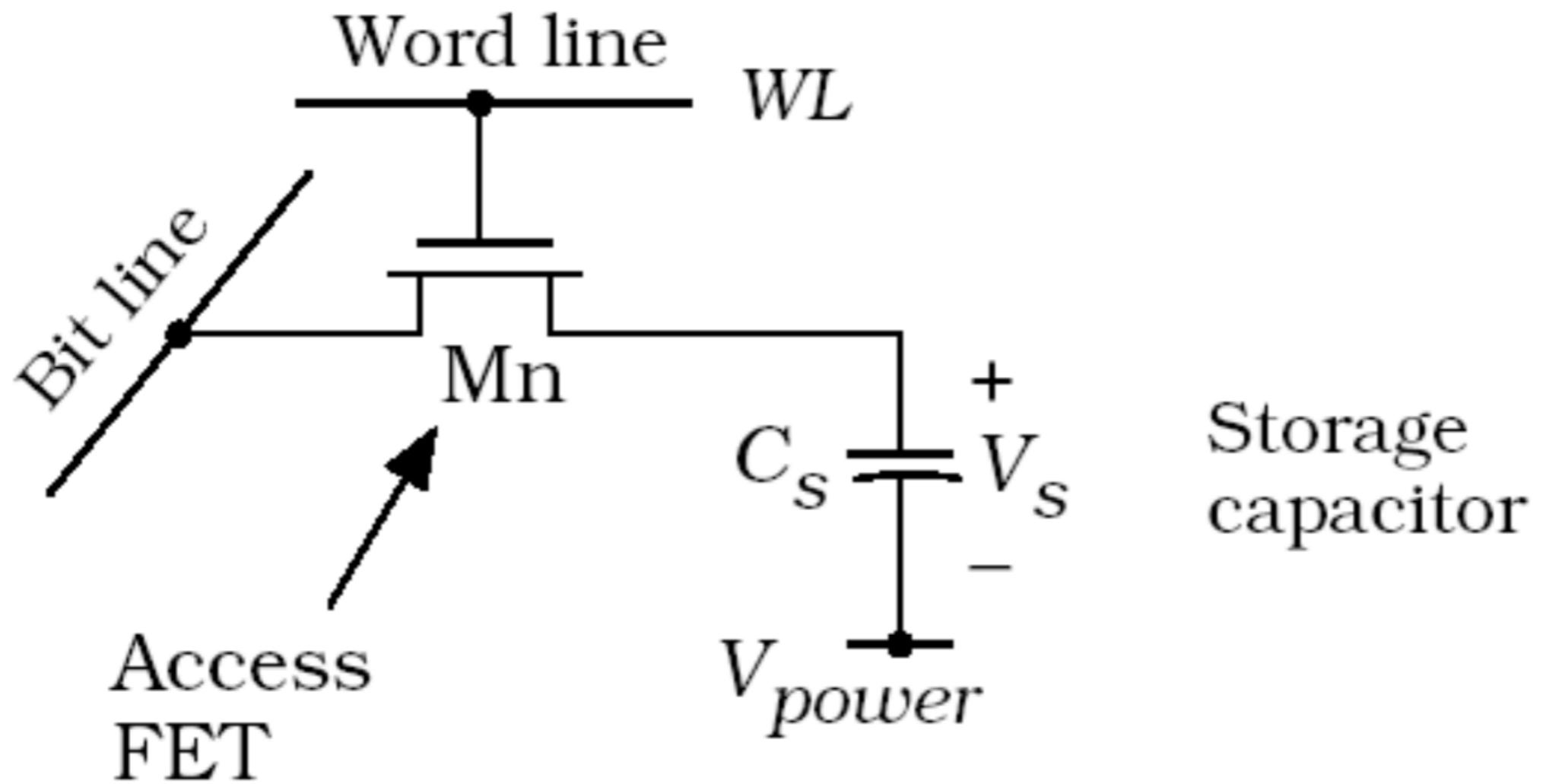
$$BL = Red \quad \overline{BL} = Green$$



# Write circuits



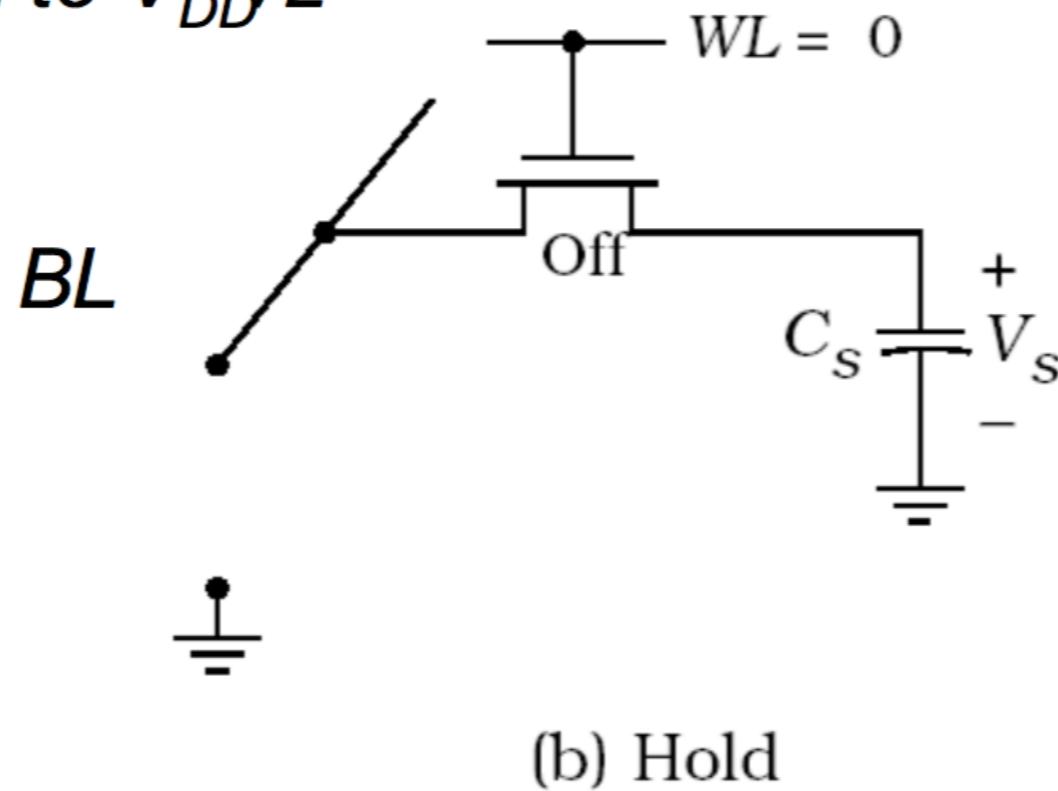
# 1T DRAM



**Only one transistor: Very high density**

# DRAM hold operation

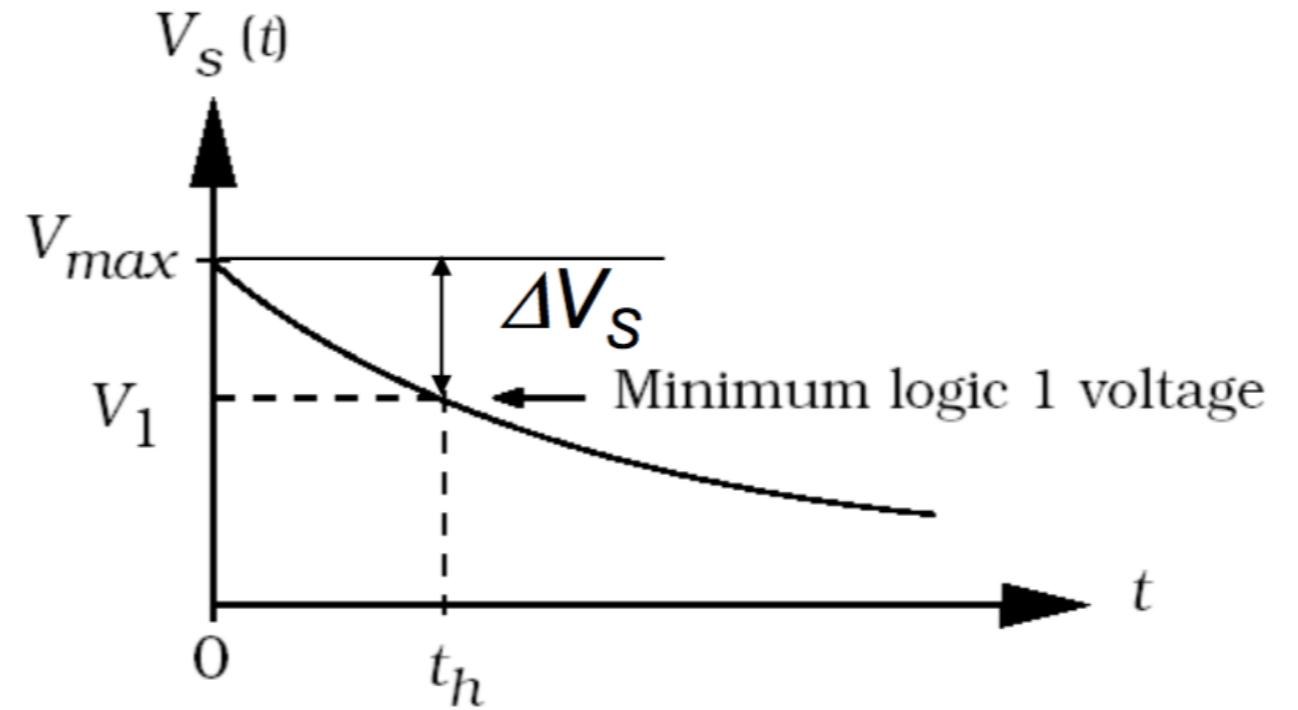
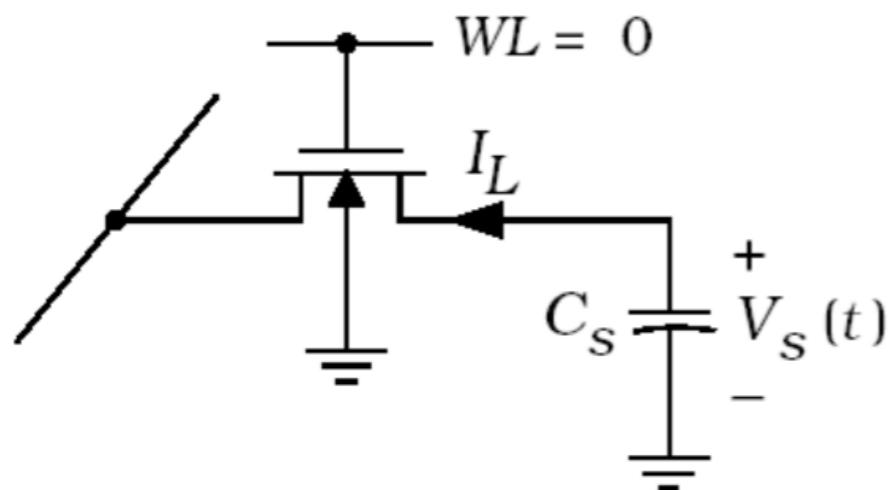
*BL* is precharged to  $V_{DD}/2$



*Logic (1): Charge stored in the capacitor and  $V_s = Q_s C_s$*

*Logic (0): No-charge stored in the capacitor and  $V_s = 0$*

# DRAM charge leakage

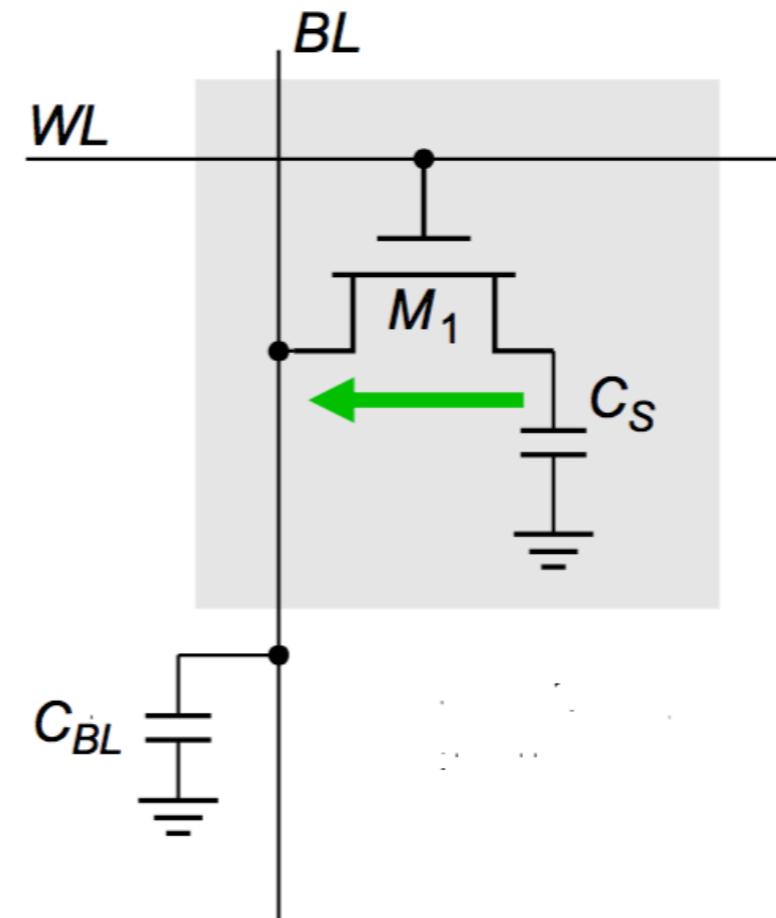
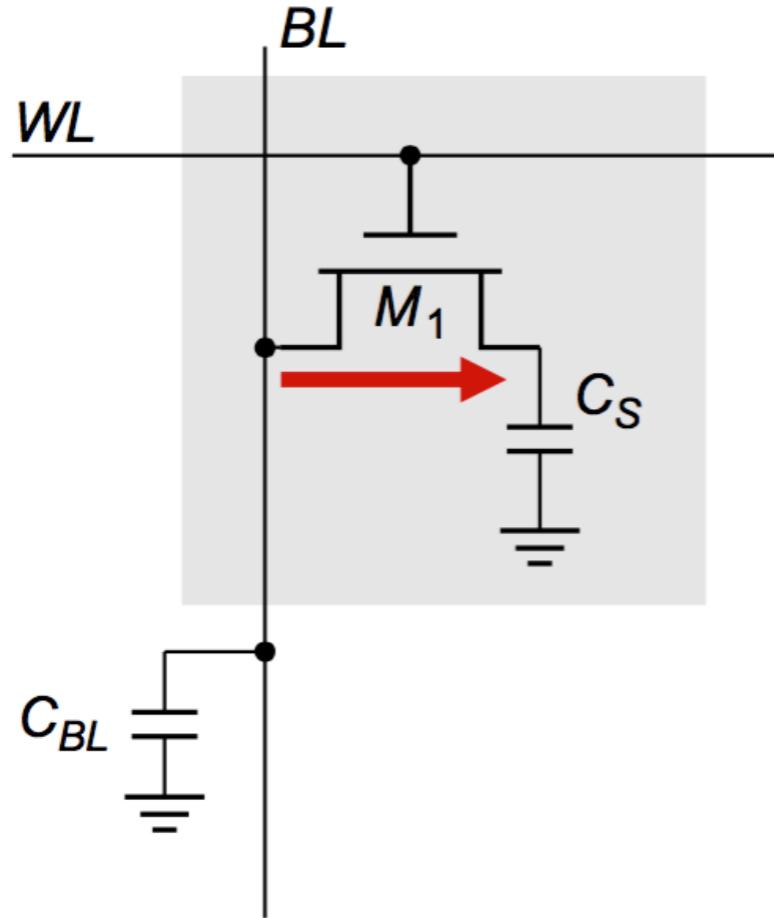


**Assuming a constant leakage current =  $I_L$**

**Retention time or hold time =  $t_h = (C_S/I_L)\Delta V_S$**

**The memory need to be refreshed periodically while storing the data**

# 1T DRAM cell



**Write '1':  $V_{BL} = V_{DD}$**

**$C_s$  is charged through  $M_1$**

$$V_s = (V_{WL} - V_{thM1}) = V_{DD} - V_{th}$$

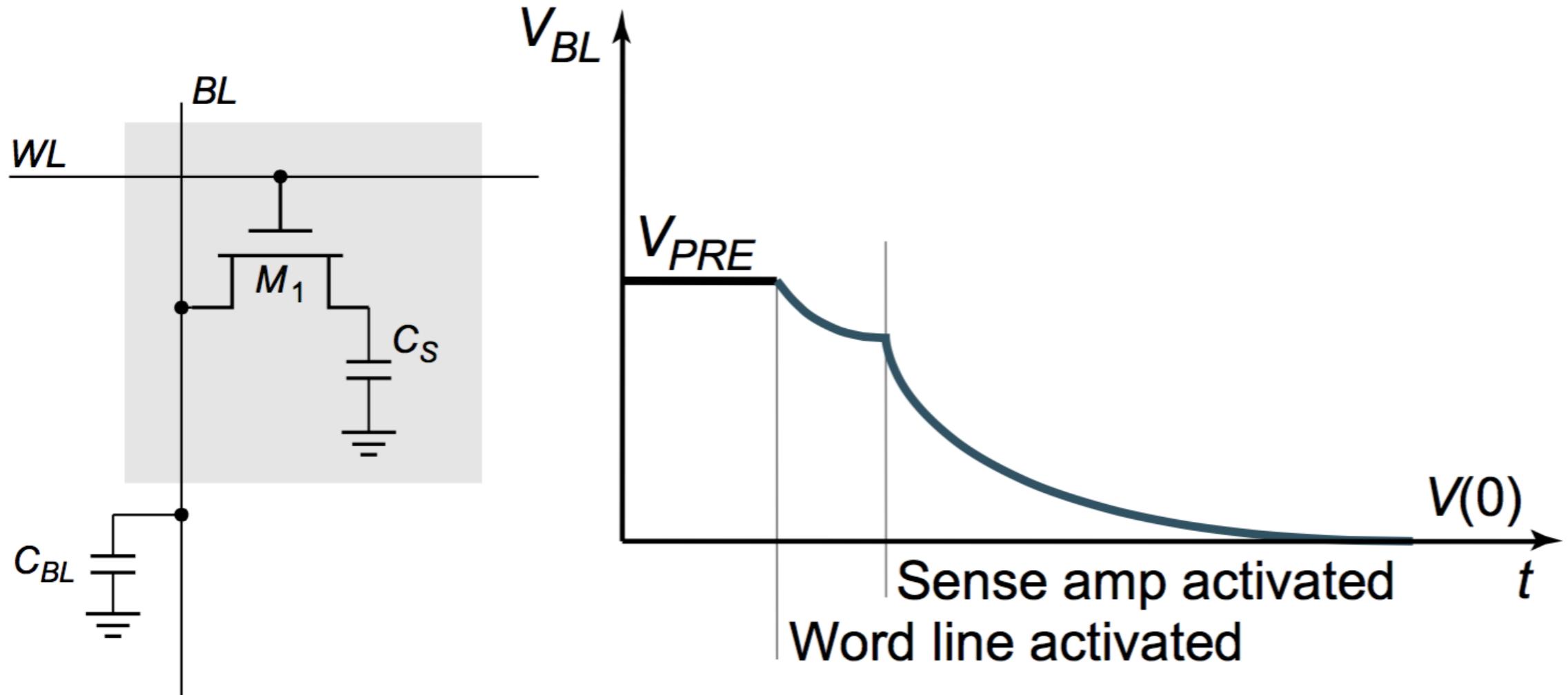
**Boosting the WL can help**

**Write '0':  $V_{BL} = 0$**

**$C_s$  is dis-charged through  $M_1$**

$$V_s = 0$$

# DRAM read operation



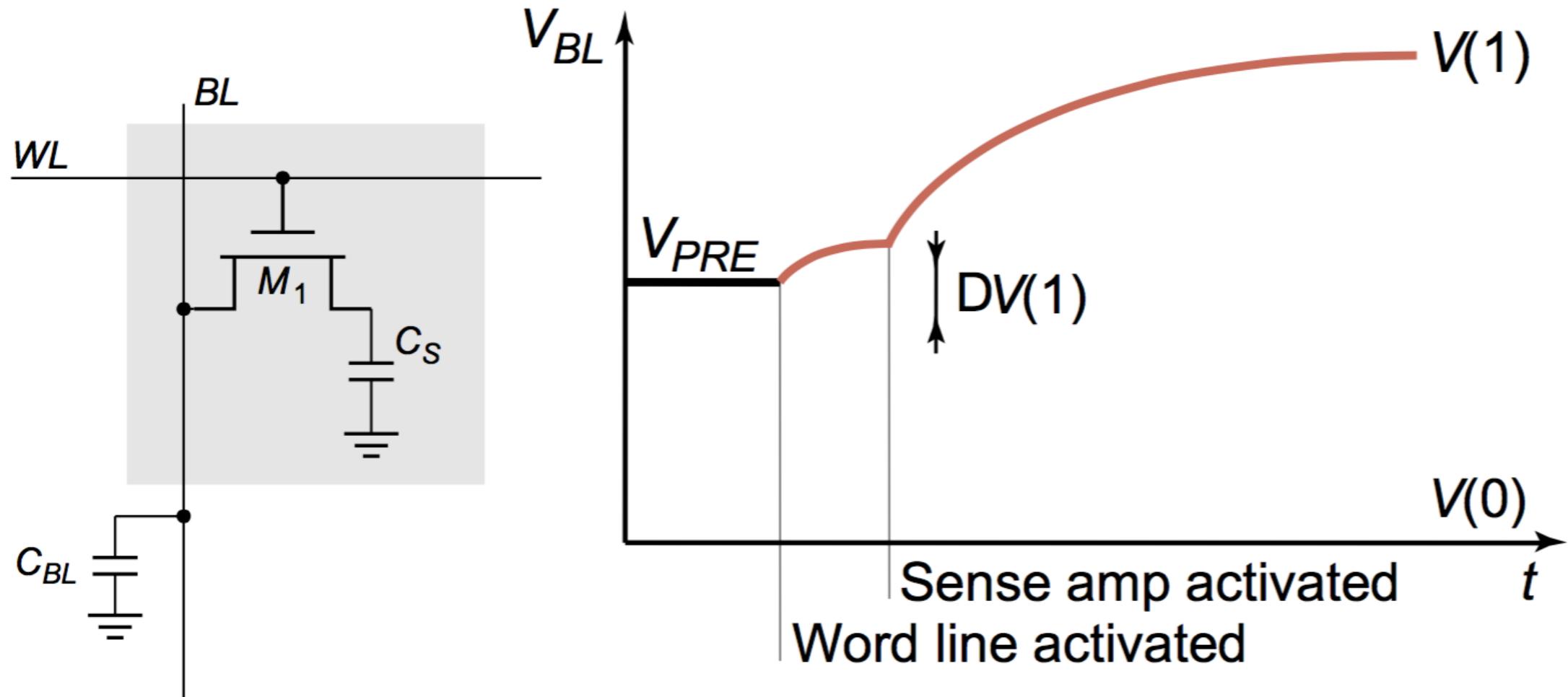
**Read '0':**

$$V_s = V_{BIT} = 0 < V_{BL} = V_{PRE}$$

Charge from  $C_{BL}$  goes to  $C_s$

**Bit-lines are precharged to  $V_{PRE}$**   
**Often  $V_{PRE} = V_{DD}/2$**

# DRAM read operation



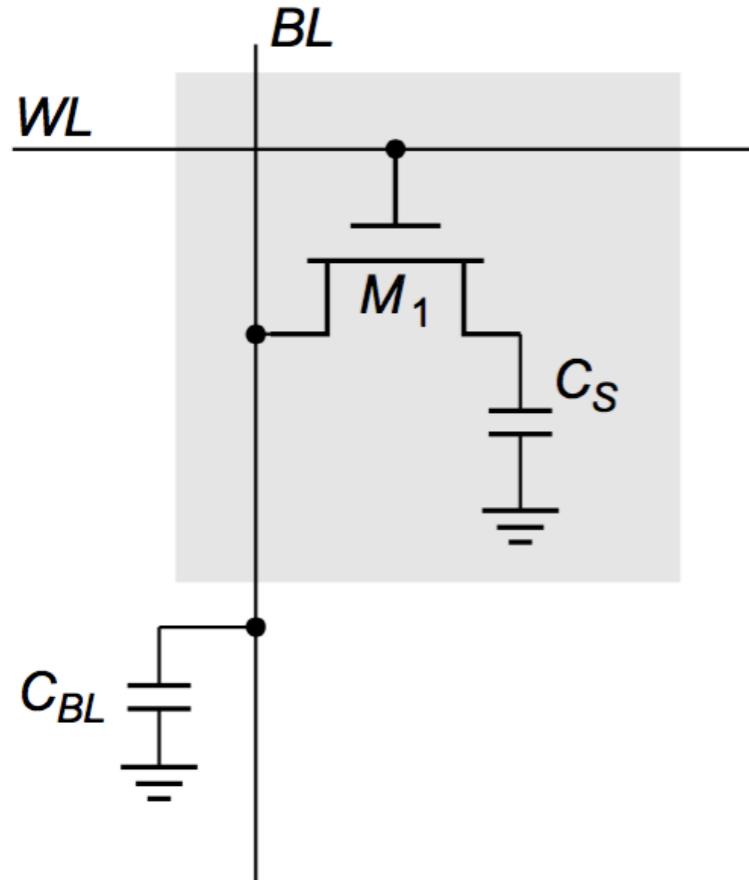
**Read '1':**

$$V_s = V_{BIT} = V_{DD} > V_{BL} = V_{PRE}$$

Charge from  $C_s$  goes to  $C_{BL}$

**Bit-lines are precharged to  $V_{PRE}$**   
Often  $V_{PRE} = V_{DD}/2$

# DRAM read operation



**charge before read operation**

$$Q = C_S V_{BIT} + C_{BL} V_{PRE}$$

**charge after read operation**

$$Q = C_S V_{BL} + C_{BL} V_{BL}$$

**Voltage change at bit line**

$$= \Delta V = V_{BL} - V_{PRE}$$

**Bit-lines are precharged to  $V_{PRE}$**   
 $V_{PRE} = V_{DD}/2$ , Cell voltage =  $V_S = V_{BIT}$

$$\Delta V = V_{BL} - V_{PRE}$$

$$C_S V_{BIT} + C_{BL} V_{PRE} = V_{BL} (C_S + C_{BL})$$

$$\Rightarrow V_{BL} = \frac{C_S V_{BIT} + C_{BL} V_{PRE}}{C_S + C_{BL}}$$

$$\Rightarrow \Delta V = \frac{C_S V_{BIT} + C_{BL} V_{PRE}}{C_S + C_{BL}} - V_{PRE}$$

$$\Rightarrow \Delta V = (V_{BIT} - V_{PRE}) \frac{C_S}{C_S + C_{BL}}$$

# DRAM read operation

*Read '0':  $V_{BIT} = 0 \Rightarrow \Delta V = -0.5V_{DD} C_s / (C_s + C_{BL})$*

*Read '1':  $V_{BIT} = V_{WL} - V_{th} = V_{DD}$  (assuming boosted WL)*

$\Rightarrow \Delta V = 0.5V_{DD} C_s / (C_s + C_{BL})$

$\frac{C_s}{C_s + C_{BL}}$  = charge transfer ratio ~ 1-10%

For 5% charge transfer ratio,  $V(1) = V_{DD}$ ,  $V(0) = 0$ ,  $V_{DD} = 2.5V$

$$\Delta V = 0.5 * 0.05 * 2.5V = 62.5mV$$

**Sense amplifier is a must for every bitline  
for correct functionality**

# What happens to cell voltage ?

$$\Delta V_{CELL} = V_{BL} - V_{BIT}$$

$$C_S V_{BIT} + C_{BL} V_{PRE} = V_{BL} (C_S + C_{BL}) \Rightarrow V_{BL} = \frac{C_S V_{BIT} + C_{BL} V_{PRE}}{C_S + C_{BL}}$$

$$\Rightarrow \Delta V_{CELL} = \frac{C_S V_{BIT} + C_{BL} V_{PRE}}{C_S + C_{BL}} - V_{BIT} = (V_{PRE} - V_{BIT}) \frac{C_{BL}}{C_S + C_{BL}}$$

**Cell voltage for bit storing ‘1’ reduces from  $V_{BIT} = V_{WL} - V_{th}$**

**Cell voltage for bit storing ‘0’ increases from  $V_{BIT} = 0$**

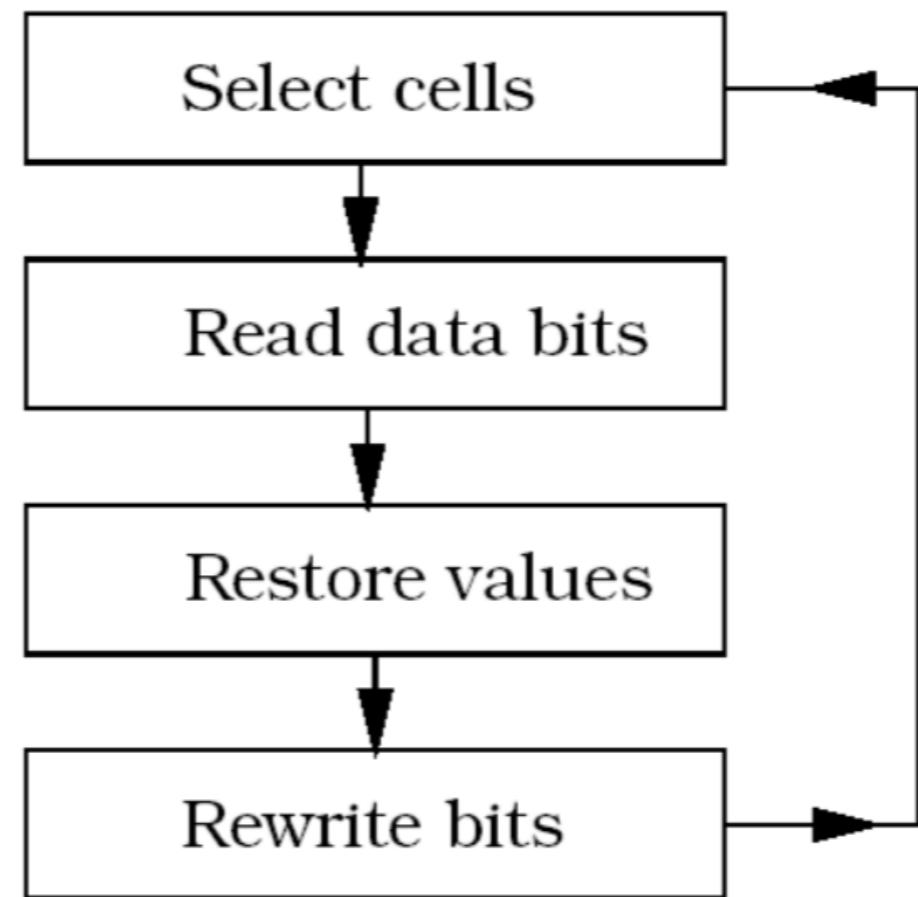
**Read operation destroy the cell data  
=> destructive read**

# What happens to cell voltage ?

**After reading original value has to be restored.**

**Normally, after reading the sense-amplifier output is applied back to the BL and WL is kept on.**

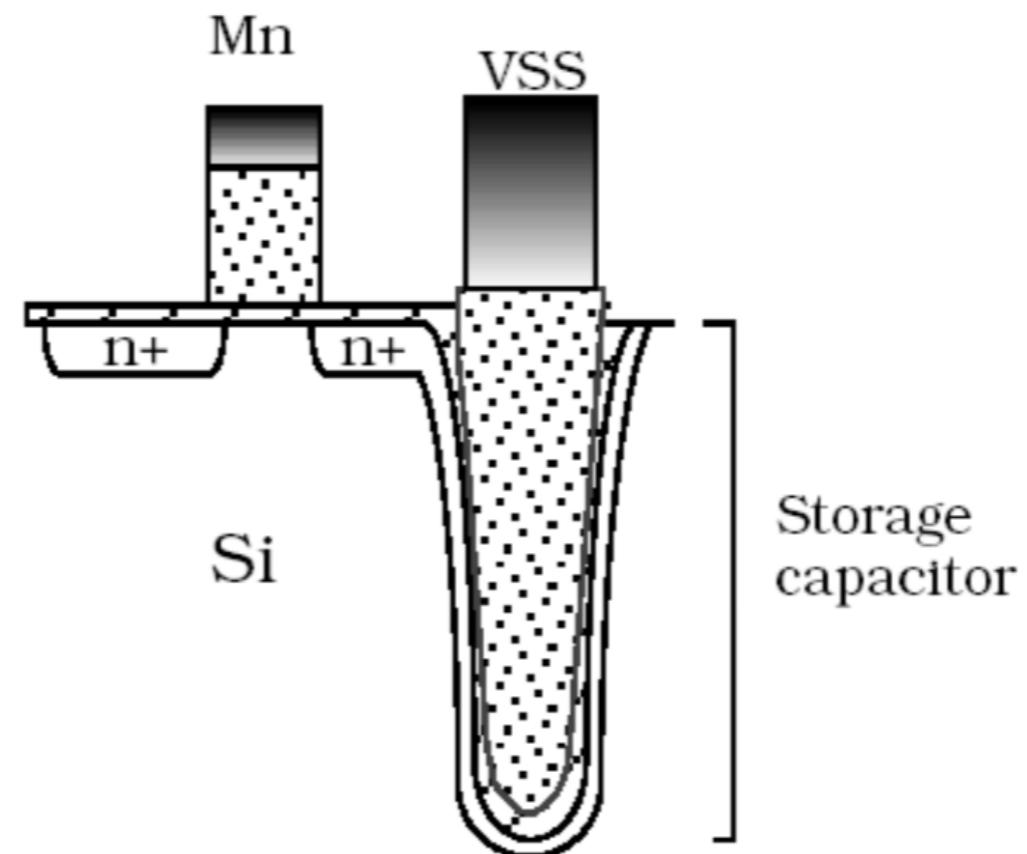
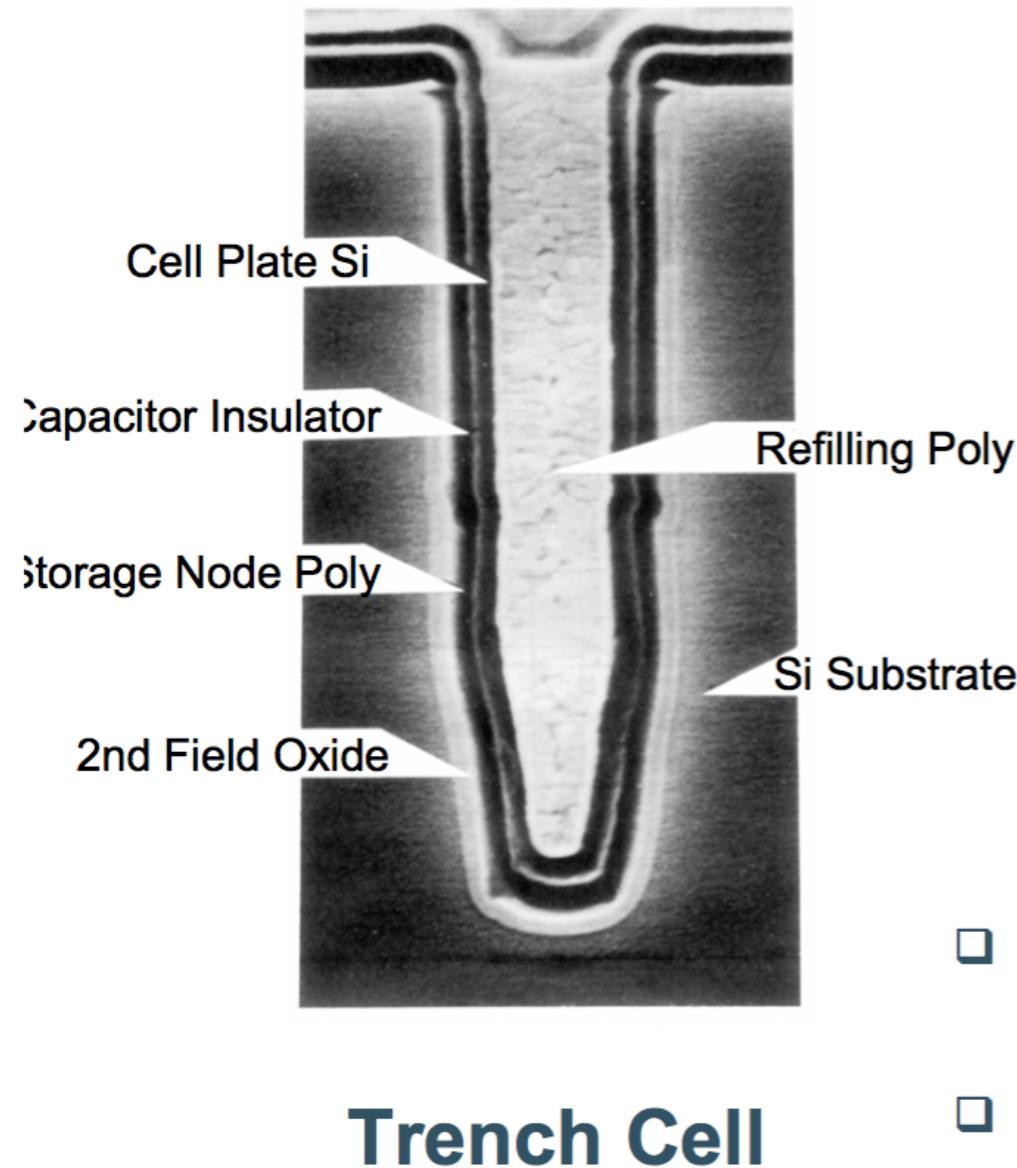
**=> Read and restore**



# DRAM key observations

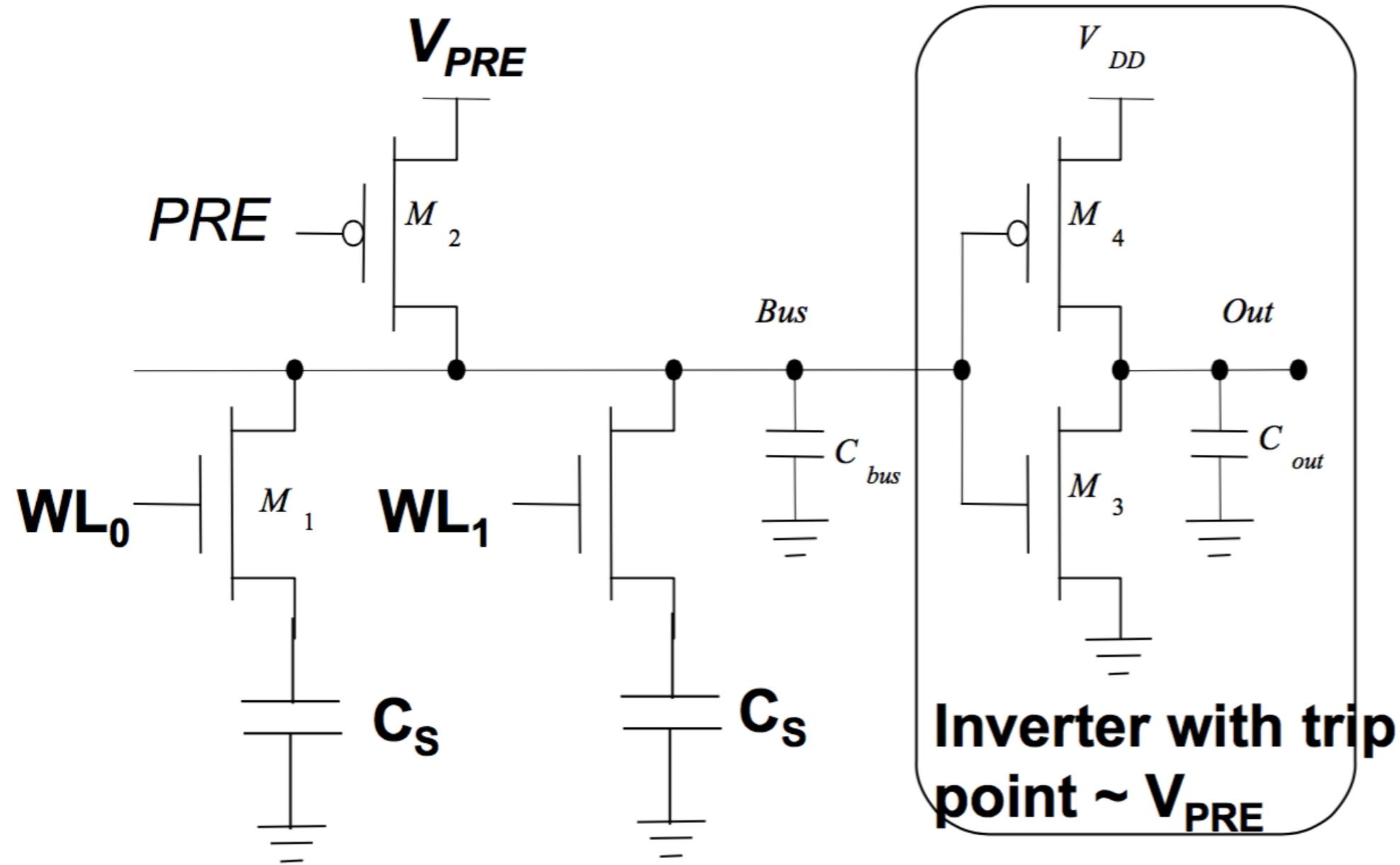
- 1T DRAM requires a sense amplifier for each bit line, due to charge redistribution read-out.
- DRAM memory cells are single ended in contrast to SRAM cells.
- The read-out of the 1T DRAM cell is destructive; read and refresh operations are necessary for correct operation.
- 1T DRAM cell requires presence of an extra capacitance that must be explicitly included in the design.
- When writing a “1” into a DRAM cell, a threshold voltage is lost. This charge loss can be circumvented by bootstrapping the word lines to a higher value than  $V_{DD}$

# DRAM capacitor: trench capacitor



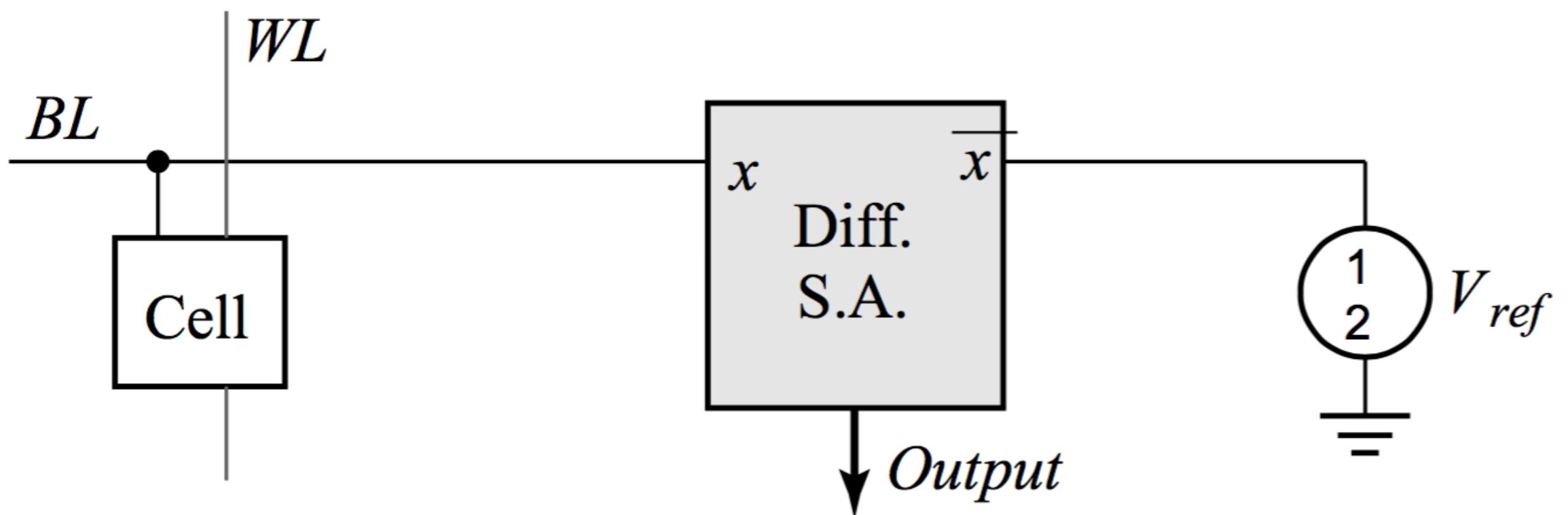
- **Vertical trench (~5μm) is deep etched into the substrate**
- **The side-wall and the bottom of the trench act as the capacitor**

# DRAM sensing: single-ended

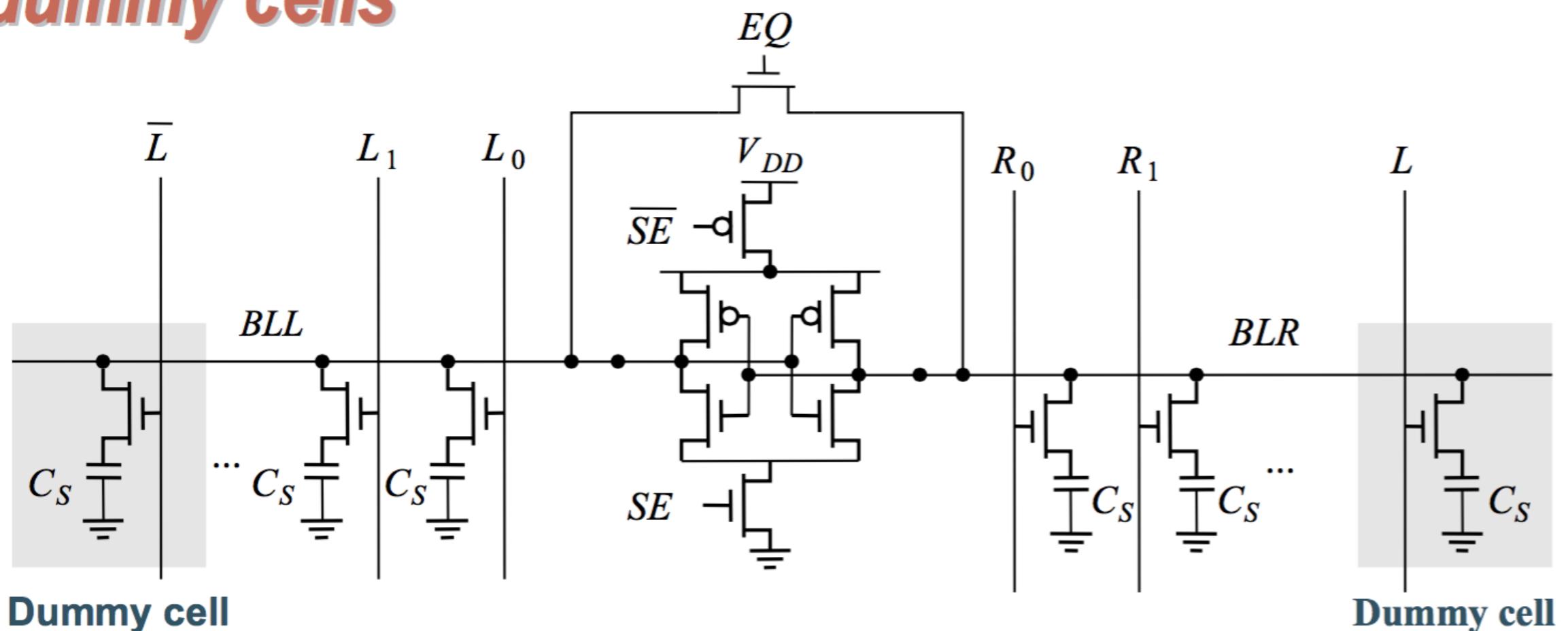


**Susceptible to noise at the bitline**

# Single-to-differential conversion



# **Open bitline architecture with dummy cells**



- Pre-charge and equalize the two bitlines to  $V_{DD}/2$
- Raise  $L$  and  $\bar{L}$  to ensure both dummy cells are charged to  $V_{DD}/2$ .
- Select a cell from one half and the dummy cell from the other half  
=> differential sensing

# **Physical design**

## CMOS fabrication

## Layout and design rules