

ECE 559 Lecture 12: Memories

SUMEET KUMAR GUPTA

ELMORE ASSOCIATE PROFESSOR OF ECE

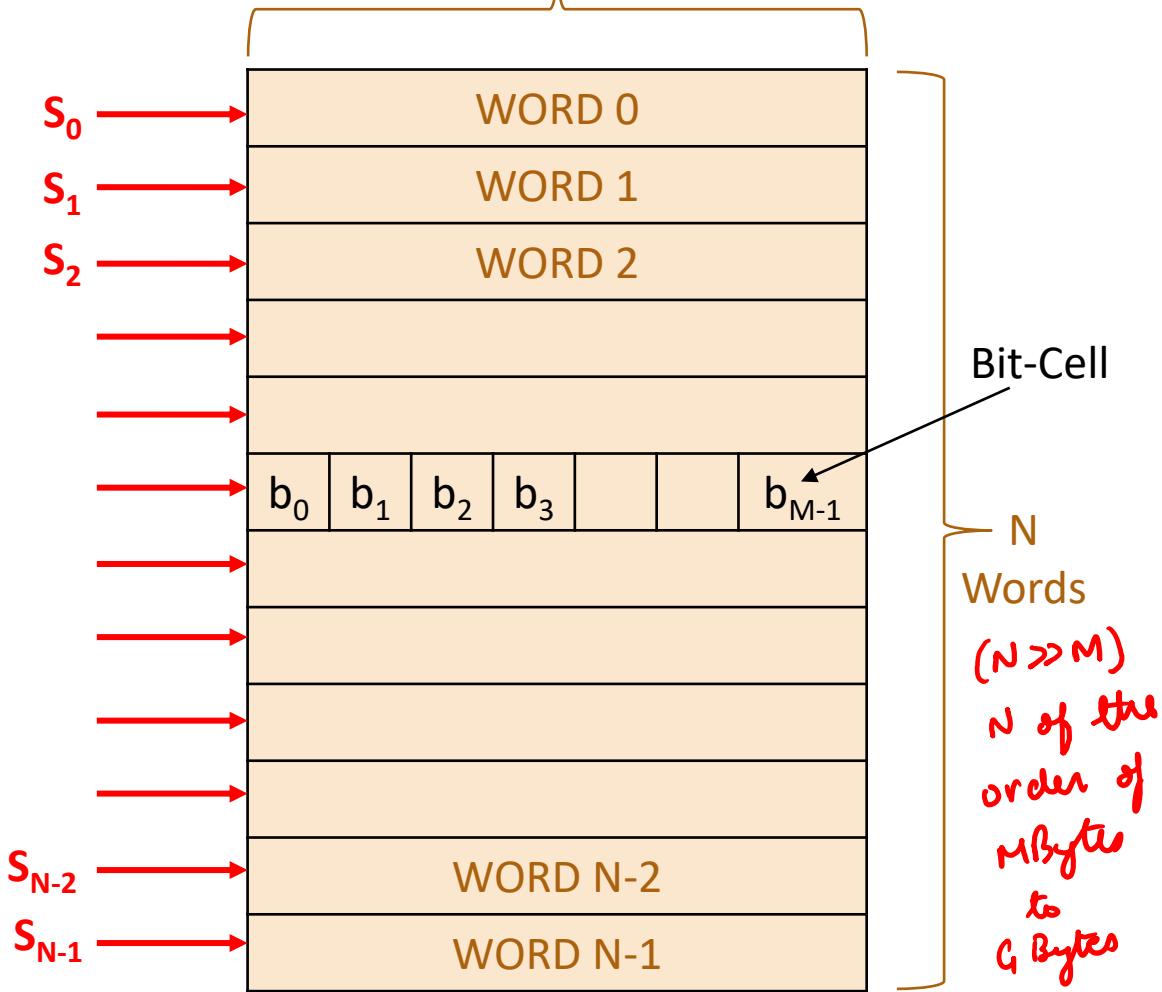
MSEE 218, GUPTASK@PURDUE.EDU

MEMORY CLASSIFICATION

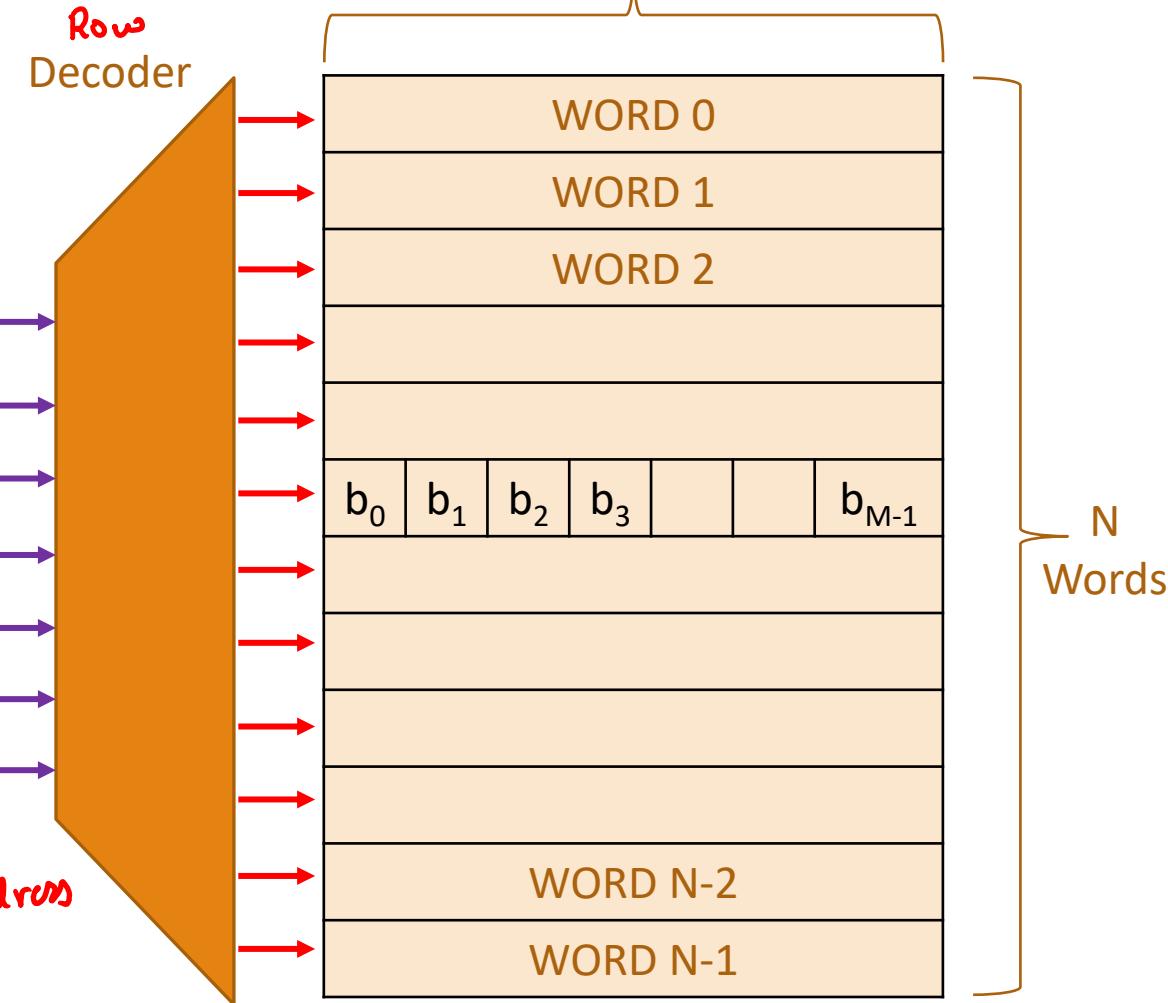
Read-Write Memory		Non-Volatile Read-Write Memory	Read-Only Memory (ROMs)
Random Access	Non-Random Access	EPROM E ² PROM FLASH ← ReRAMs PCM Spin-based Ferroelectrics	Mask-Programmed Programmable (PROM) ! One-time Programmable (OTP)
SRAM DRAM	FIFO LIFO Shift Register CAM		

MEMORY ARCHITECTURE

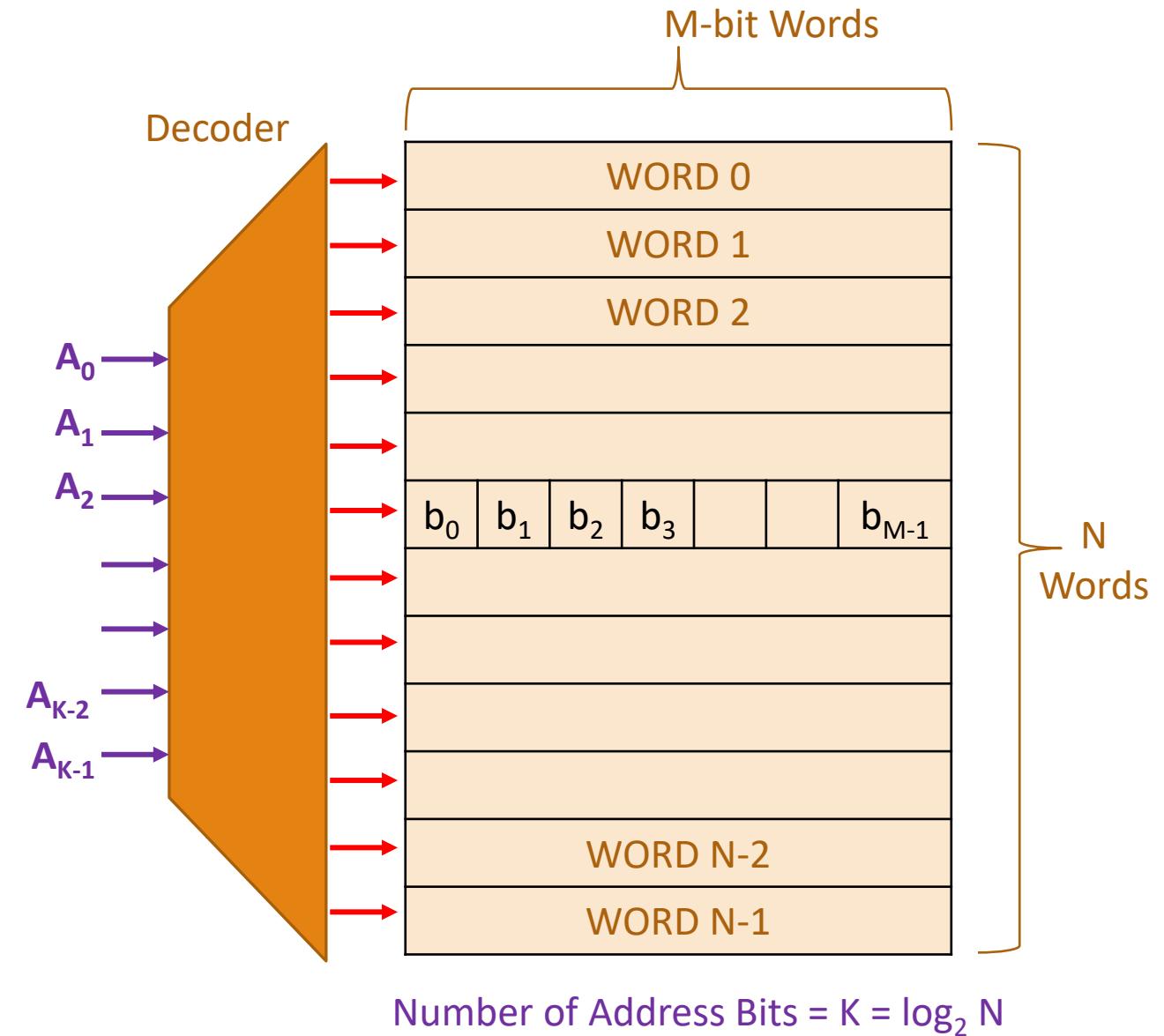
M-bit Words (e.g. $m=64$)



$A_3 A_2 A_1 A_0 = 0010$
 $S_2 = 1$, Rest of $S_0, S_1, S_3 \dots S_{15} = 0$

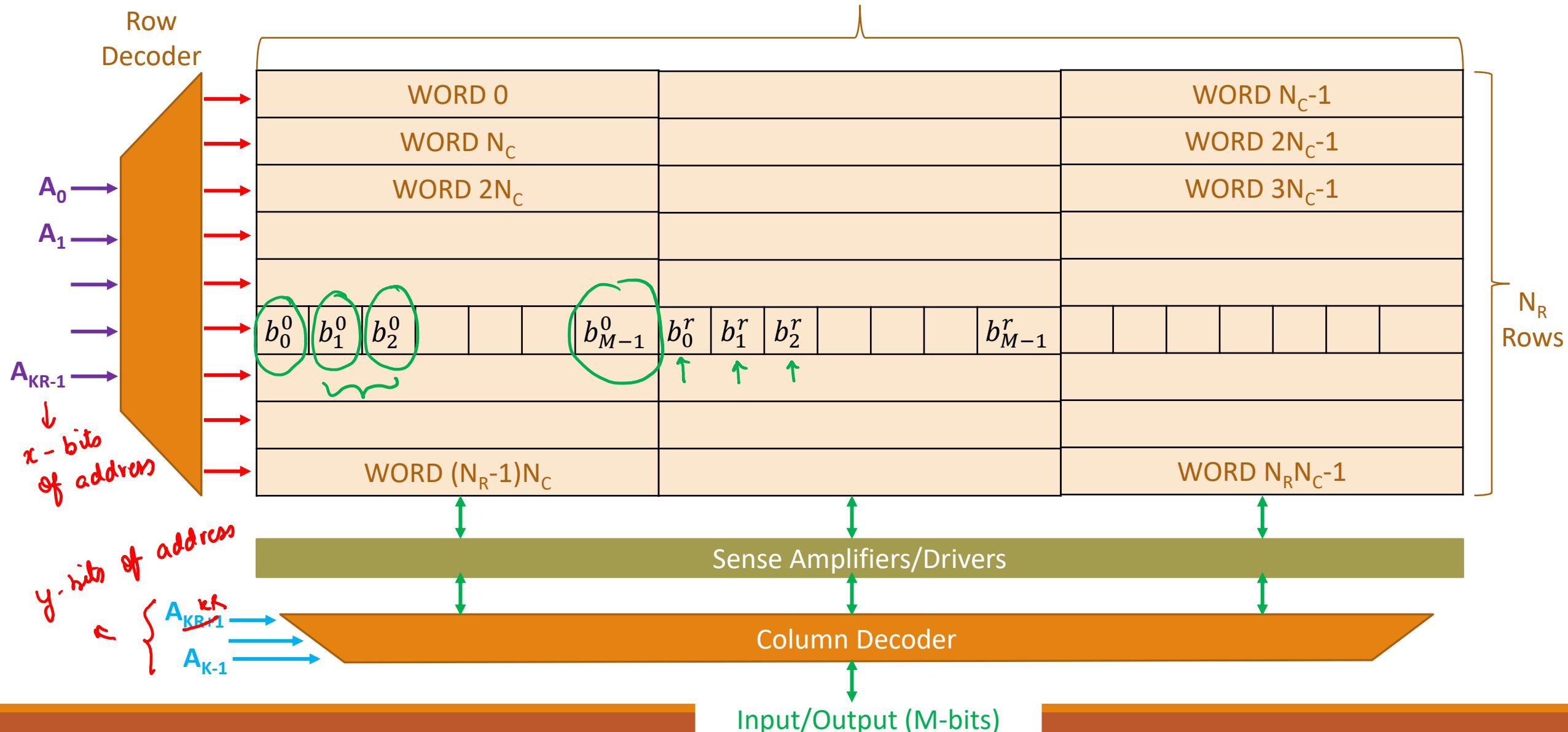


MEMORY ARCHITECTURE: SKEWED ASPECT RATIO

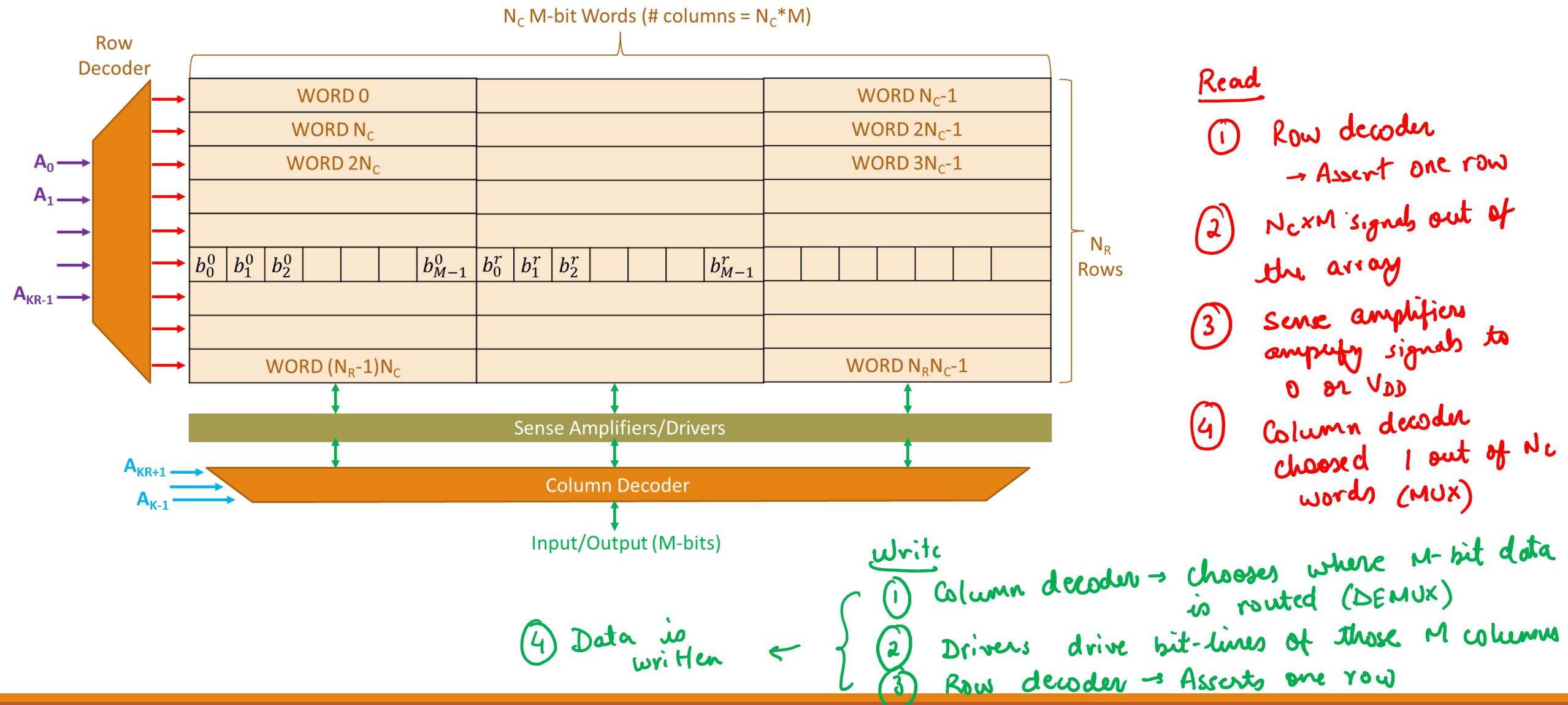


MEMORY ARCHITECTURE: BALANCED ASPECT RATIO

$$N = N_C N_R$$

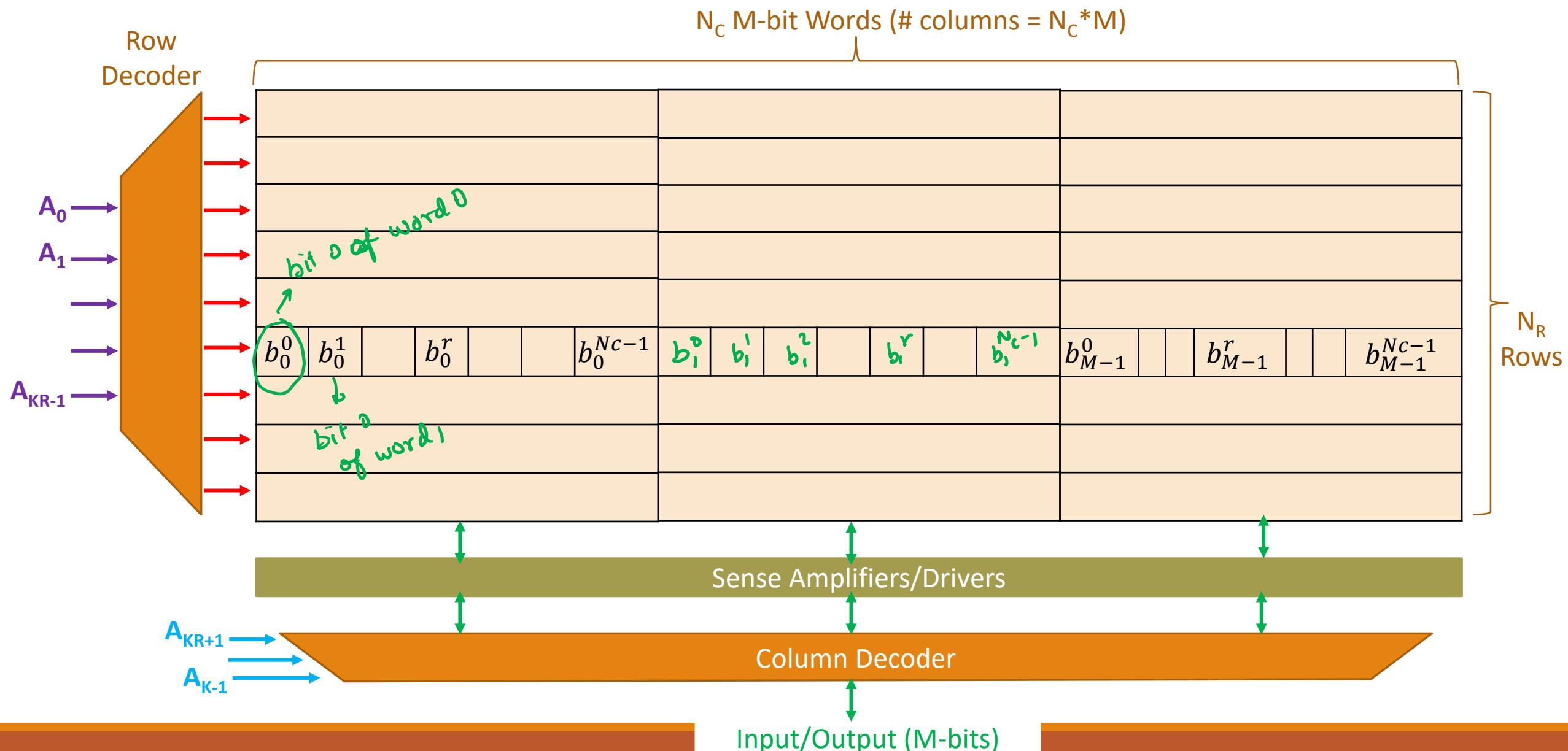


MEMORY ARCHITECTURE: BALANCED ASPECT RATIO

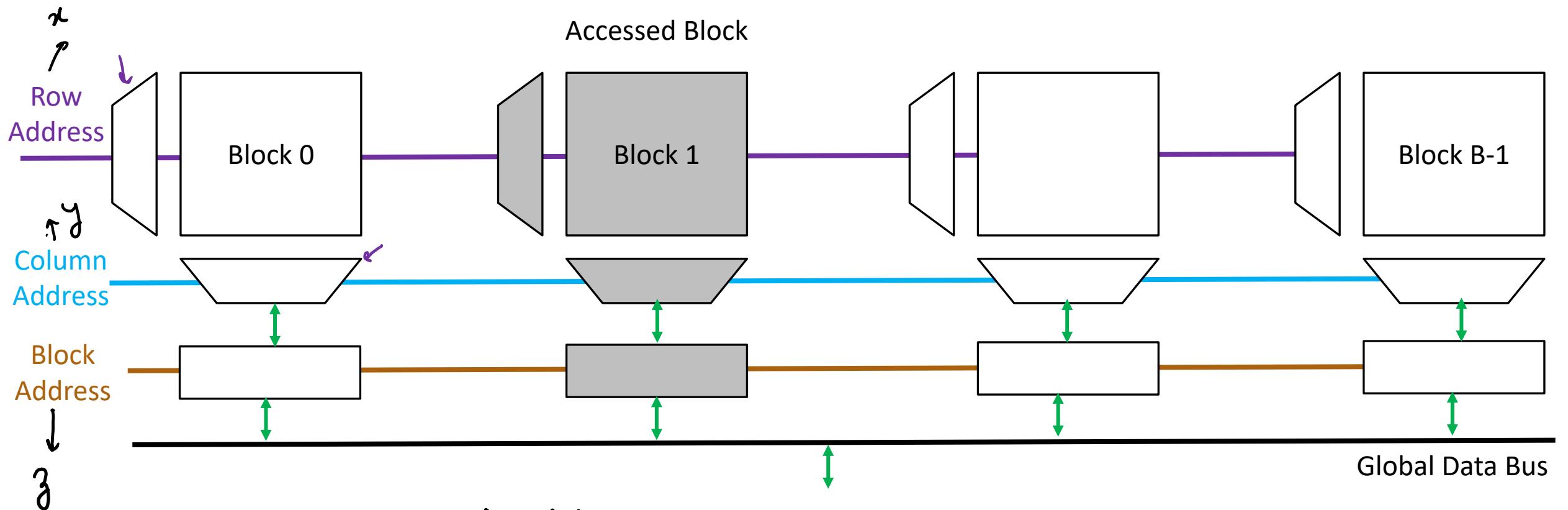


BIT-INTERLEAVED ARCHITECTURE

→ More immune to soft errors

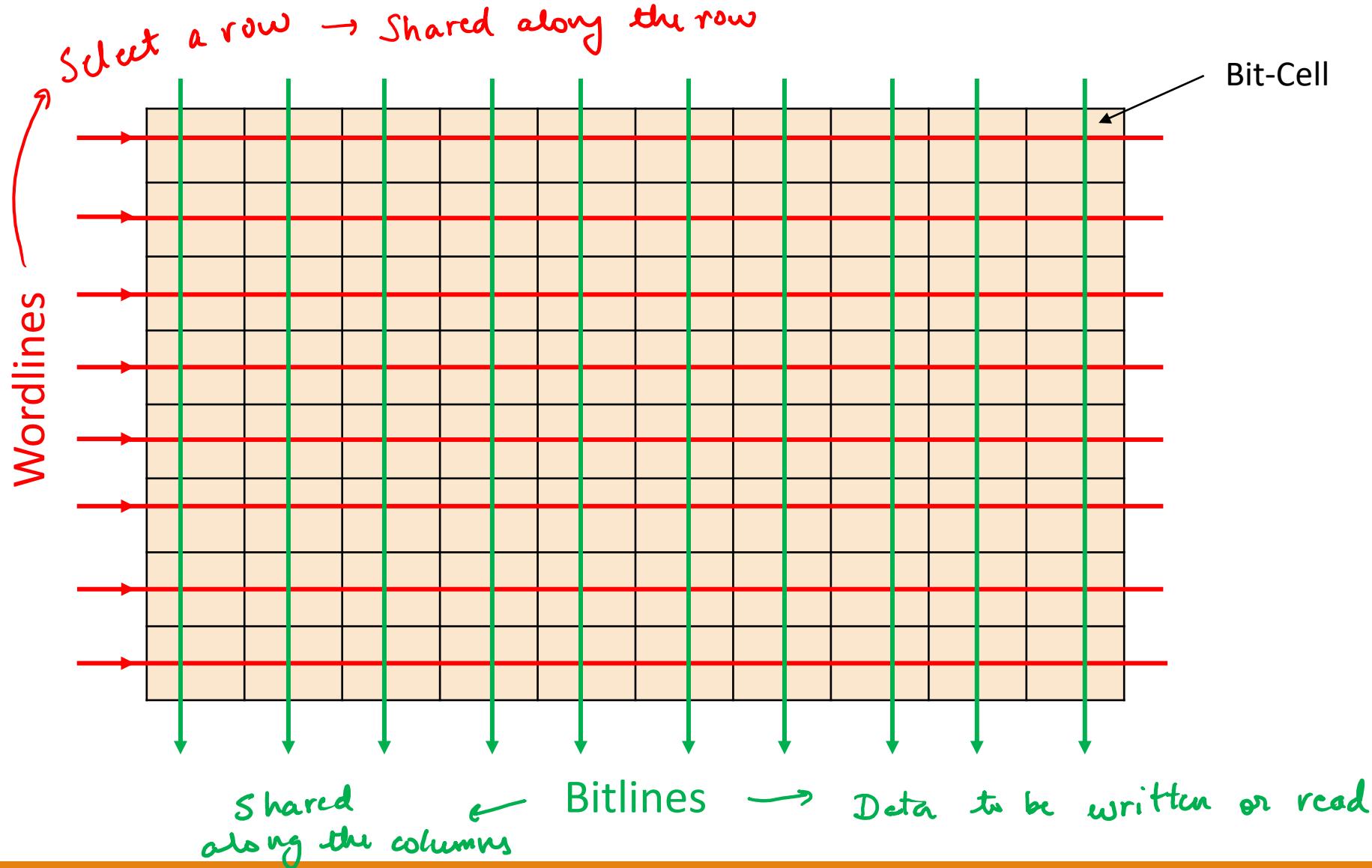


HIERARCHICAL MEMORY ARCHITECTURE



- Reduces wire delay
- Reducing standby power (Operate unaccessed blocks at low V_{DD})
 - ↓ Save leakage energy

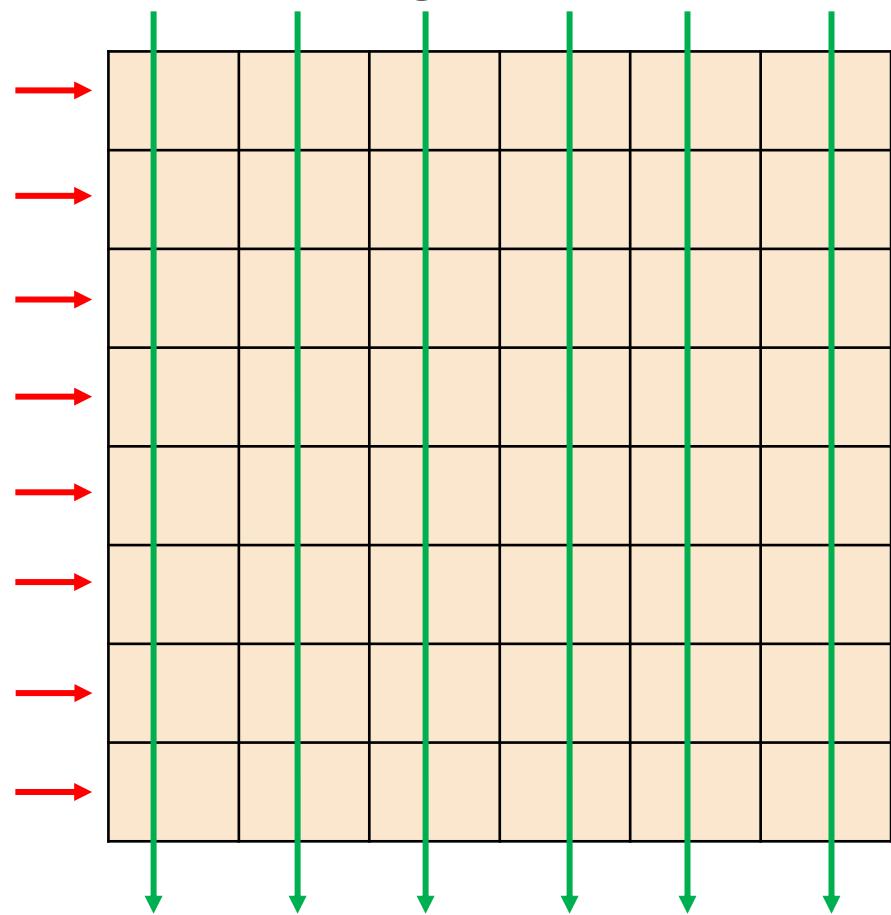
BIT-LINES AND WORD-LINES



DIFFERENTIAL AND SINGLE-ENDED MEMORIES

Single-Ended

Wordlines

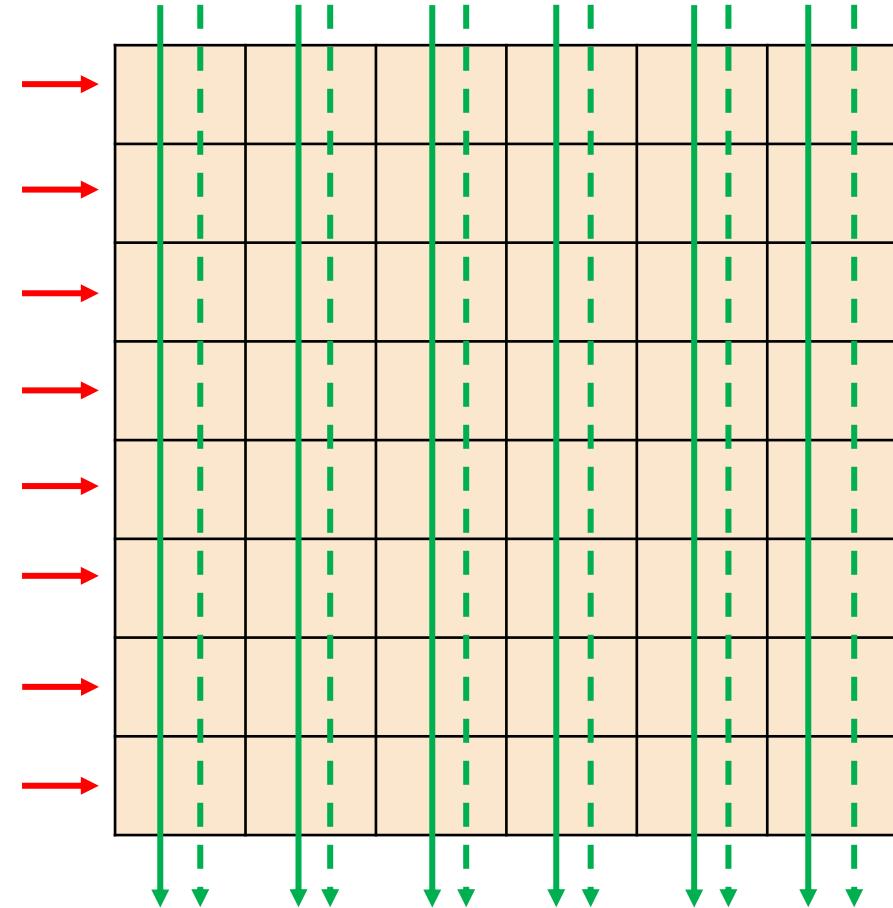


BL only

→ Area-efficient

Differential

Wordlines



BL and BLB

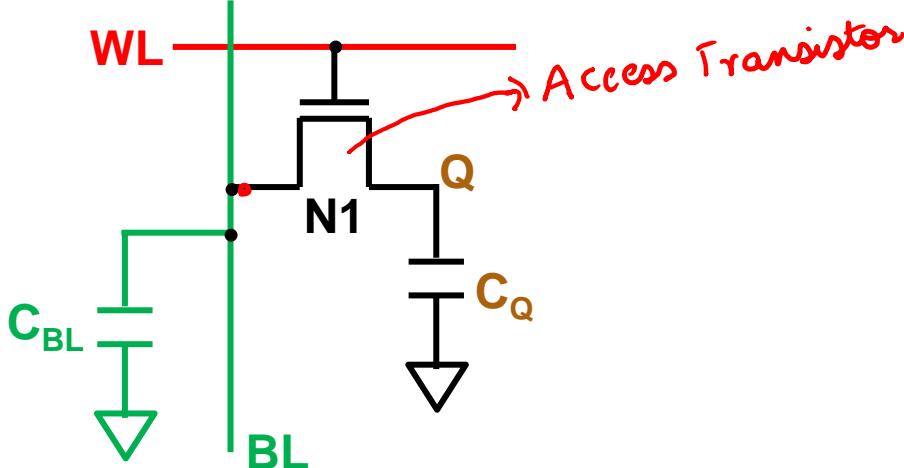
→ Better noise immunity
→ Faster

DYNAMIC RAMS (DRAMS)

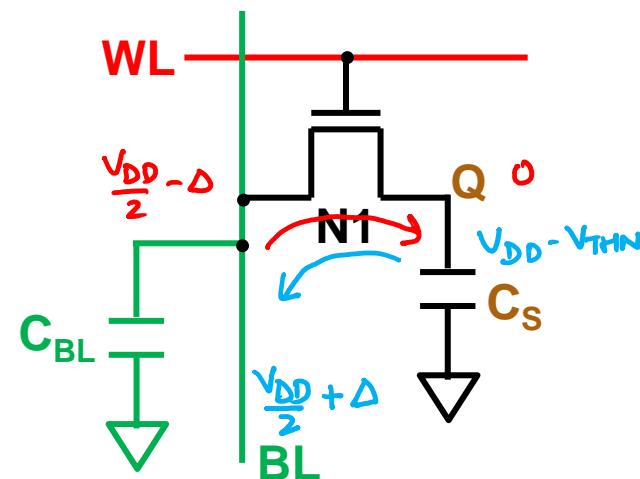
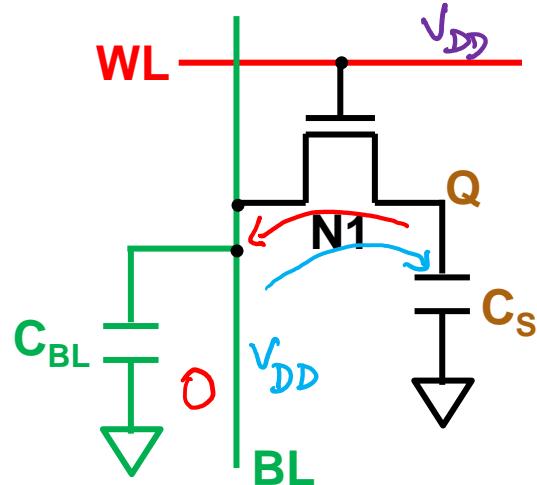
1-TRANSISTOR DRAM

→ 1T-1C

Single-ended



High Density



Write Operation

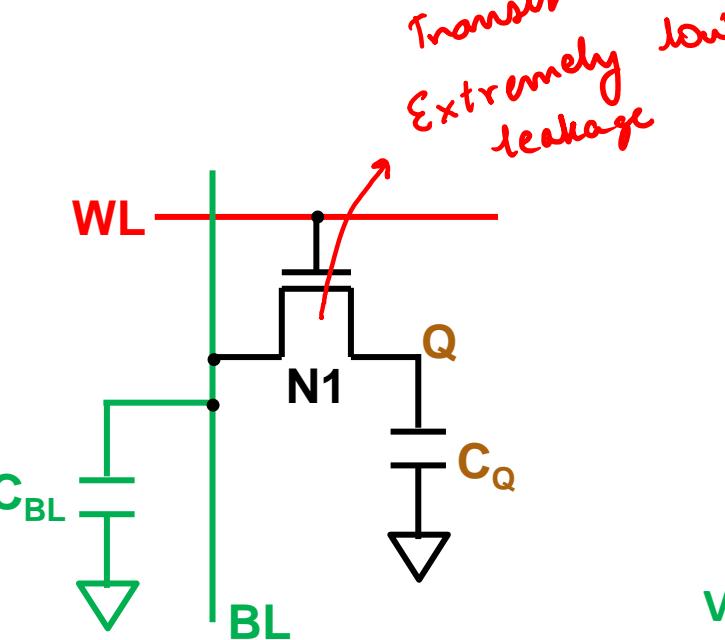
- BL driven to V_{DD} (for write 1) or 0 (for write 0)
- WL asserted
- C_S is charged to $V_{DD} - V_{THN}$ or discharged to 0

Read Operation

- BL precharged to $V_{DD}/2$
- WL asserted ($N1$ turned ON)
- BL charges if C_S stores '1'
- BL discharges if C_S stores '0'

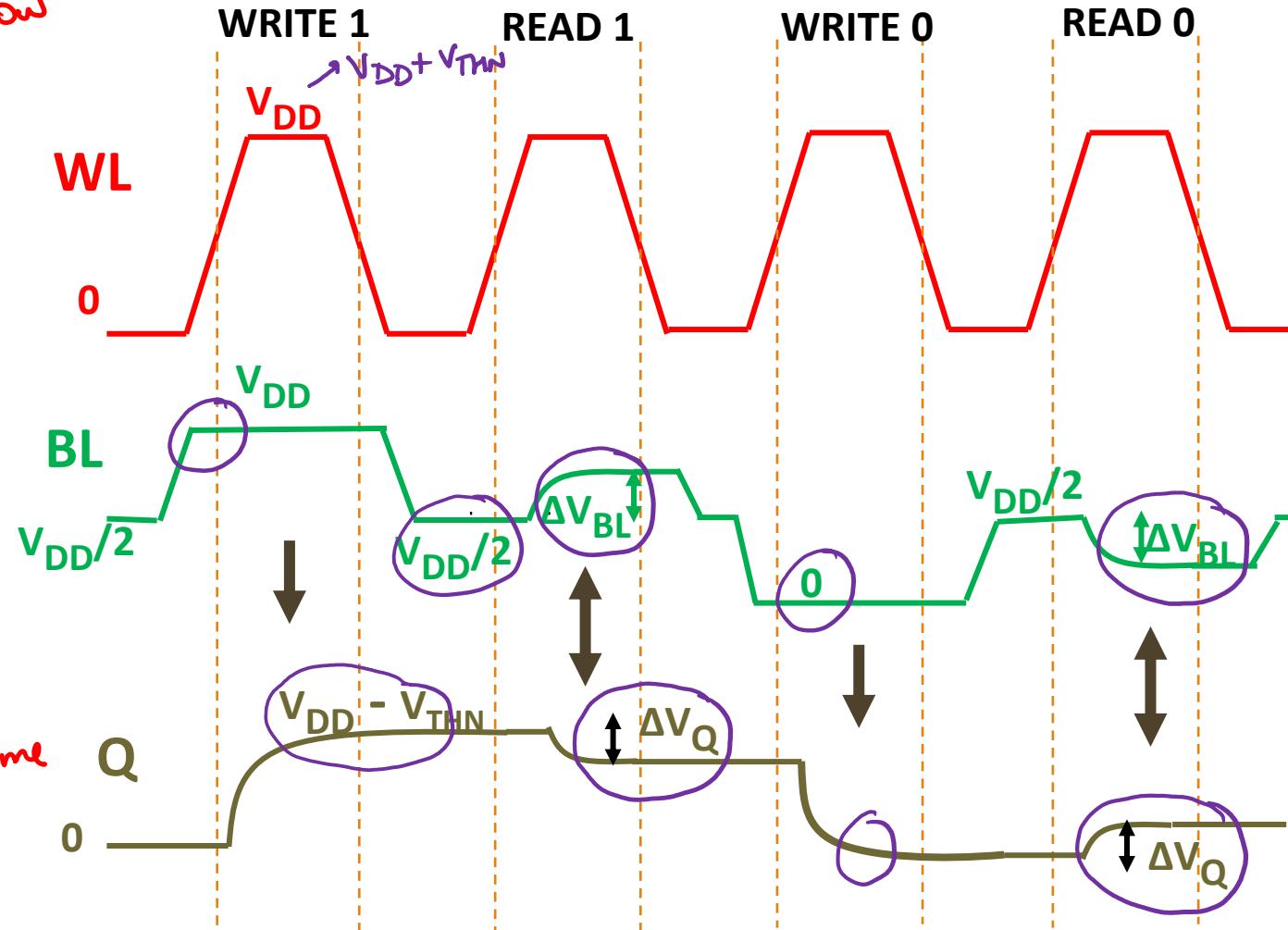
Floating

1-TRANSISTOR DRAM

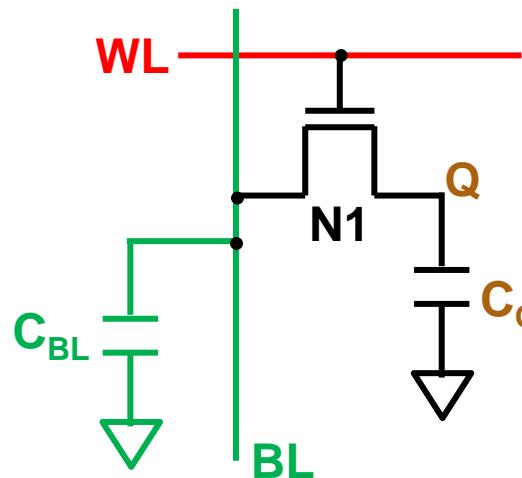


- Destructive Read
- Needs Refresh \leftarrow Refresh time $\sim ms$
- Incomplete swing at Q. (WL overdrive solves this issue)

$$V_{WL} = V_{DD} + V_{THN}$$



CHARGE SHARING DURING READ



Conservation of Charge

$$Q_{INIT} = C_{BL} V_{PRE} + C_Q V_Q$$

$$Q_{FINAL} = (C_{BL} + C_Q) V_{FINAL}$$

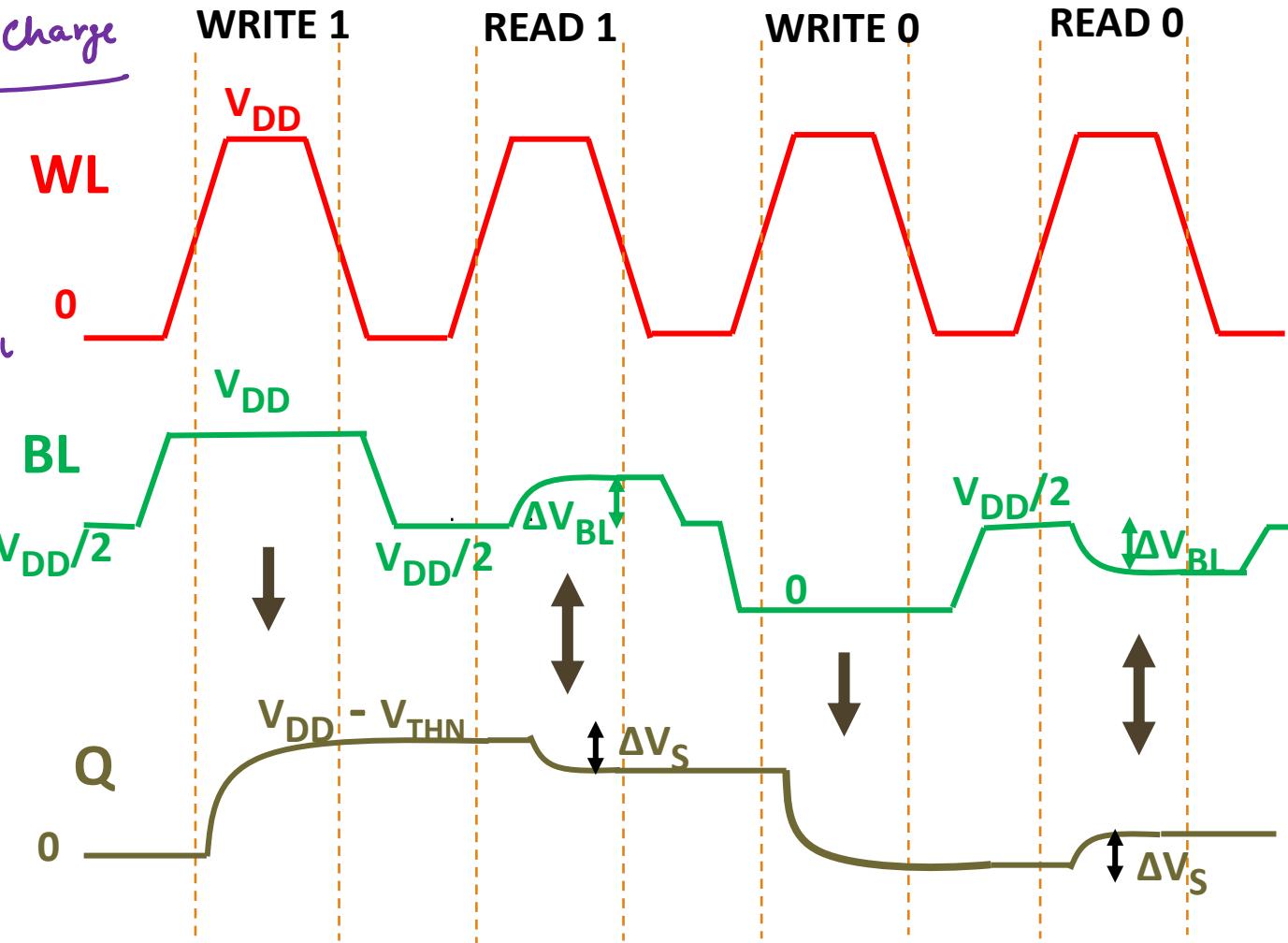
$$Q_{INIT} = Q_{FINAL}$$

$$C_{BL} V_{PRE} + C_Q V_Q = (C_{BL} + C_Q) V_{FINAL}$$

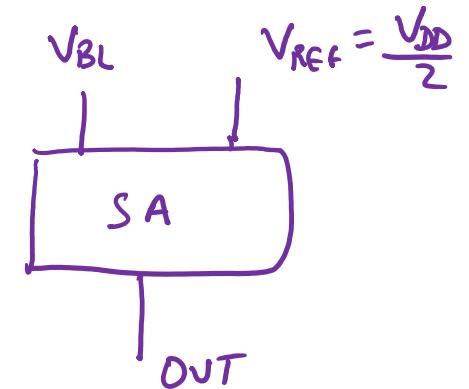
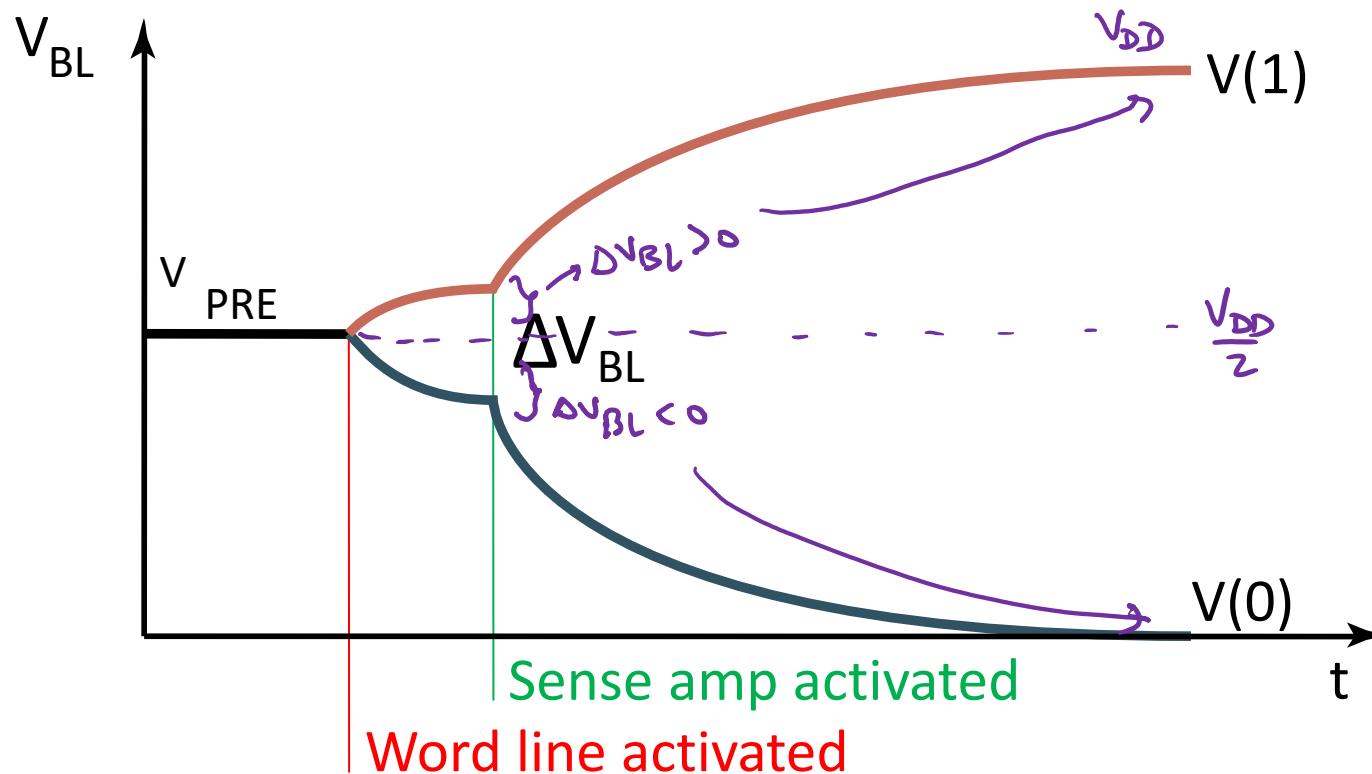
$$V_{FINAL} = \frac{C_{BL} V_{PRE} + C_Q V_Q}{C_{BL} + C_Q}$$

$$\Delta V_S = V_{FINAL} - V_Q = \frac{C_{BL}}{C_Q + C_{BL}} (V_{PRE} - V_Q)$$

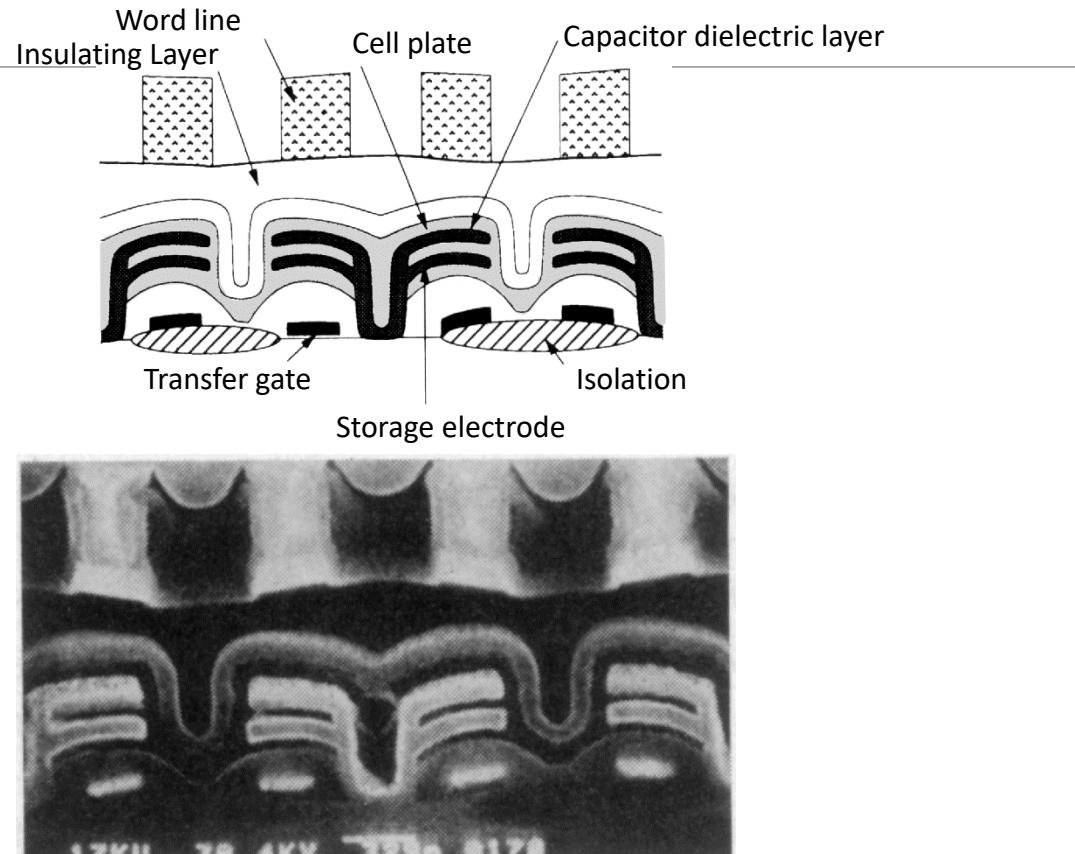
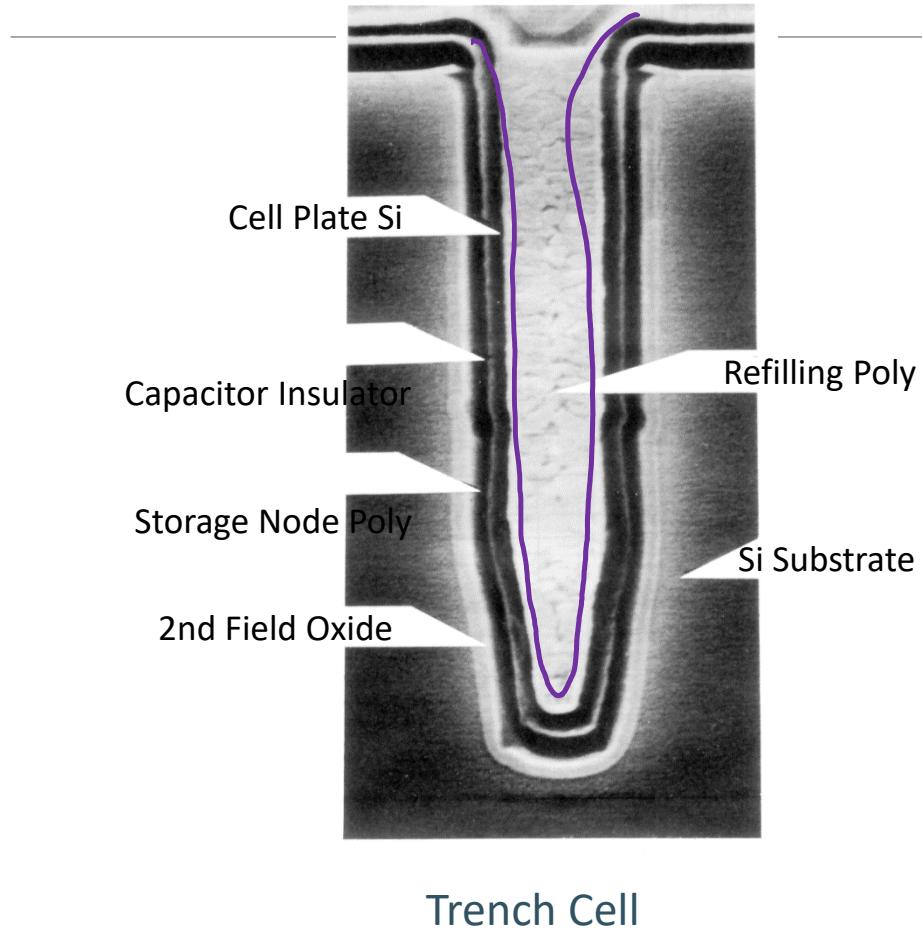
$$\Delta V_{BL} = V_{FINAL} - V_{PRE} = [C_Q / (C_Q + C_{BL})] (V_Q - V_{PRE})$$



SENSE AMPLIFIER OPERATION



ADVANCED 1T DRAM CELLS

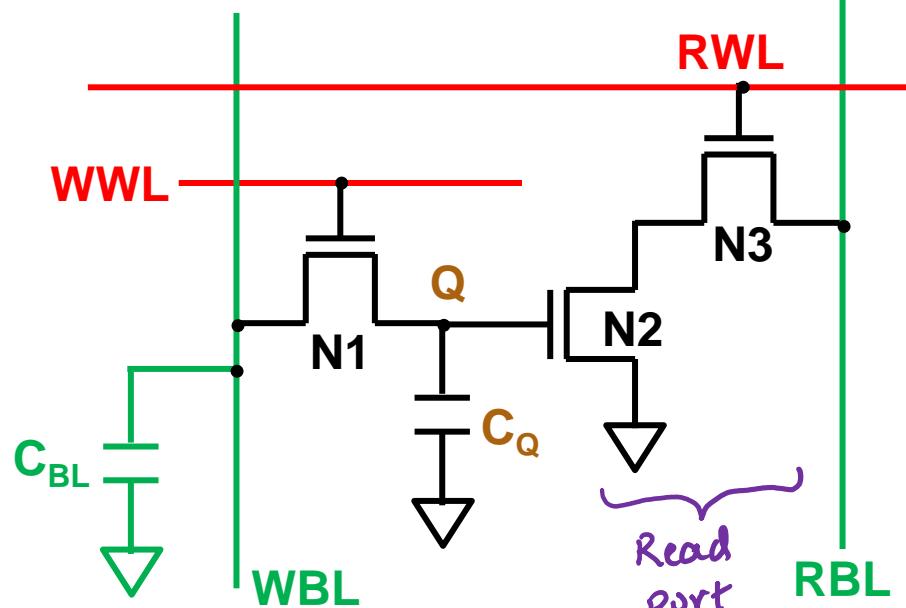


Single-ended
design

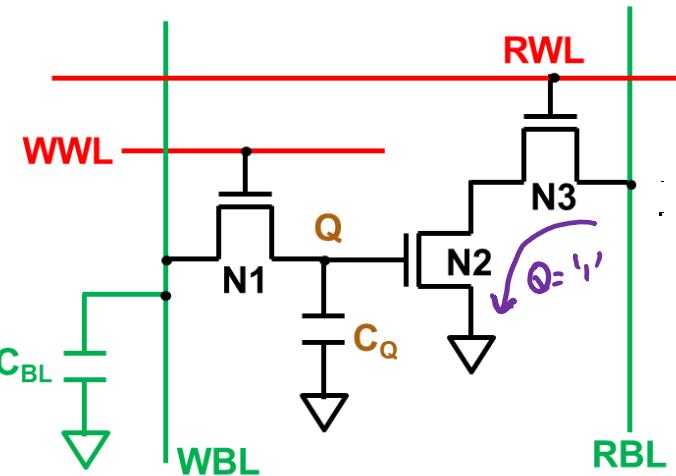
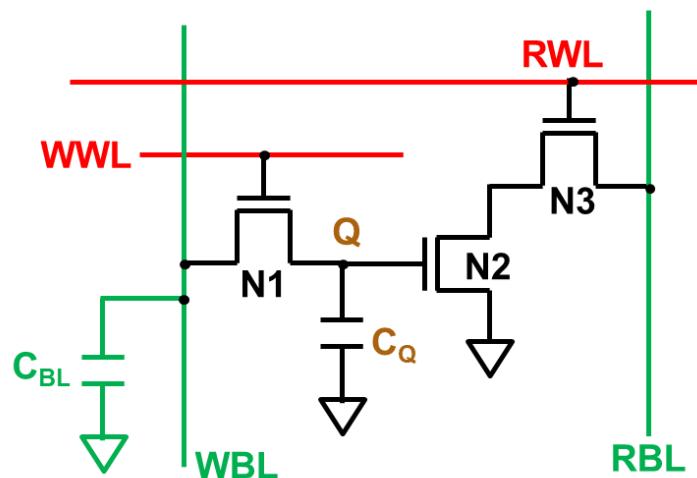
2T
(Advanced
design)

3-TRANSISTOR DRAM

→ 3T-DRAM (Gain cell)



Separate read-write paths



Write Operation

- WBL driven to V_{DD} (for write 1) or 0 (for write 0)
- WWL asserted (M1 turned ON) → $RWL @ 0$
- C_S is charged to $V_{DD} - V_{THN}$ or discharged to 0

Non-destructive
↑

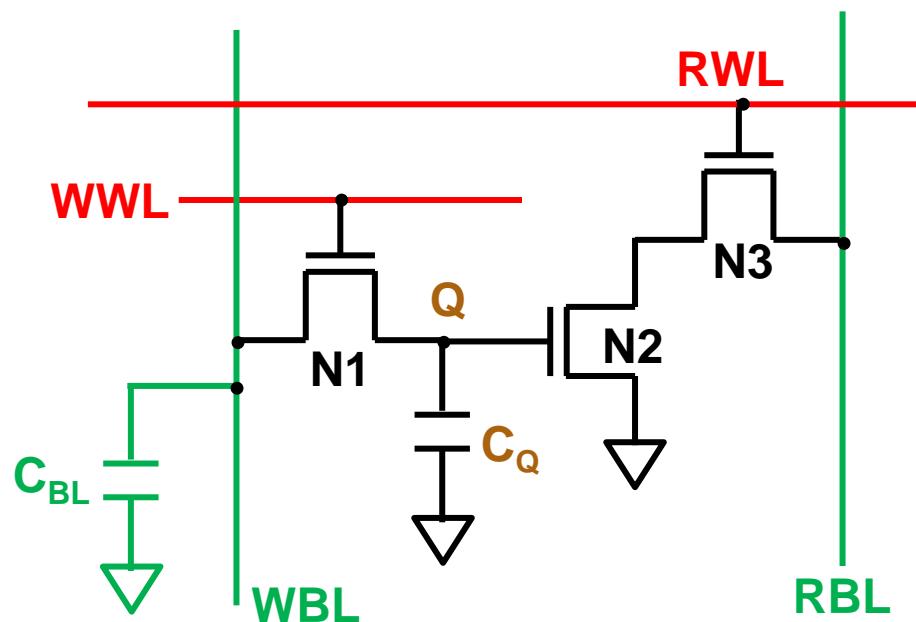
Read Operation

- RBL precharged to V_{DD}
- RWL asserted (M3 turned ON)
- RBL remains at V_{DD} if C_Q stores '0' (N_2 is OFF)
- RBL discharges if C_Q stores '1' (N_2 is ON)

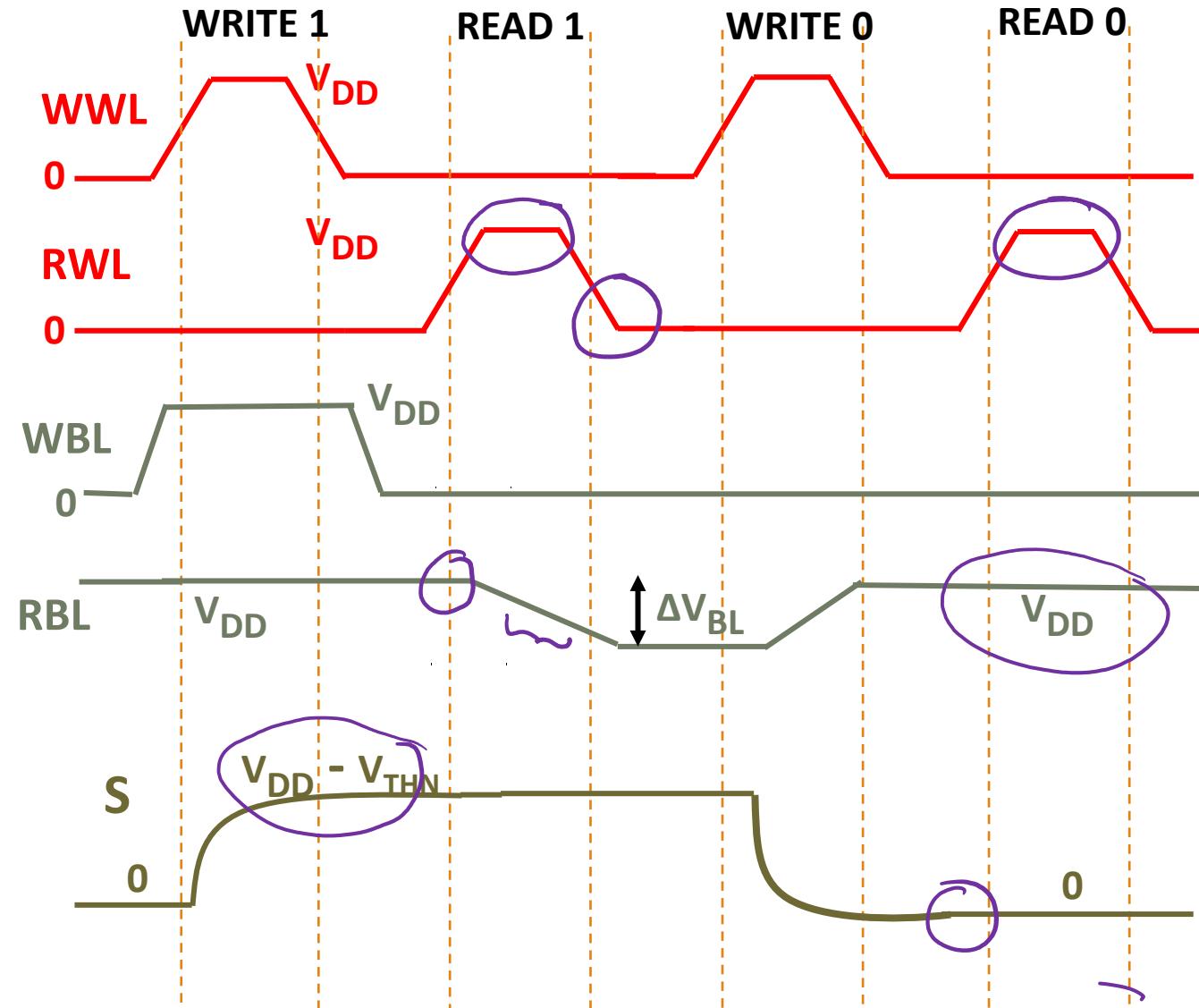
$$t_{refresh} = \frac{C_Q V_Q}{I_{LEAK}}$$

3-TRANSISTOR DRAM

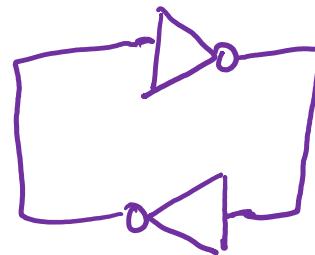
→ lower density than 1T-1C



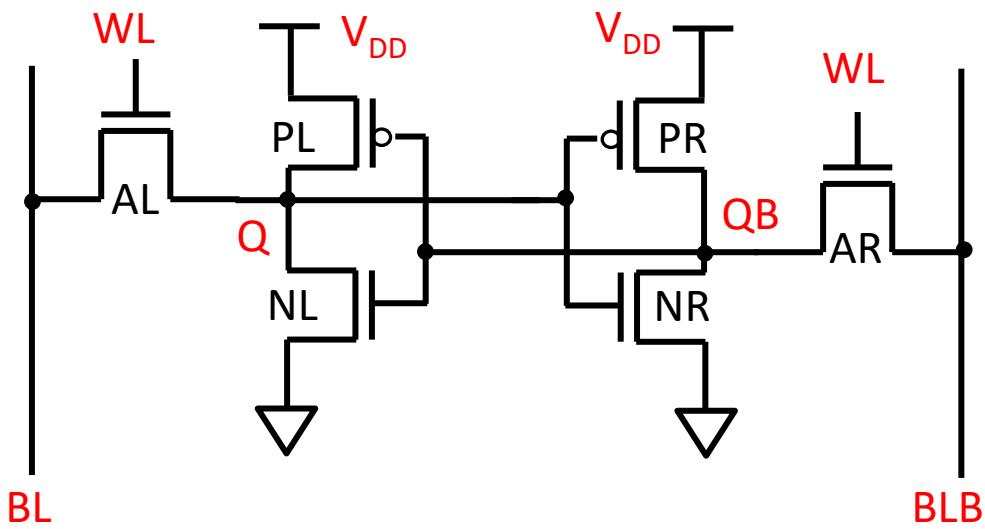
- Non-Destructive Read
- No write-back
- Incomplete swing at S. (WL overdrive solves this issue)



STATIC RAM (SRAM)

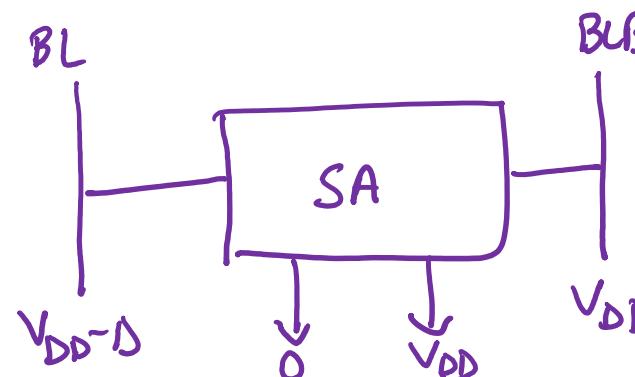
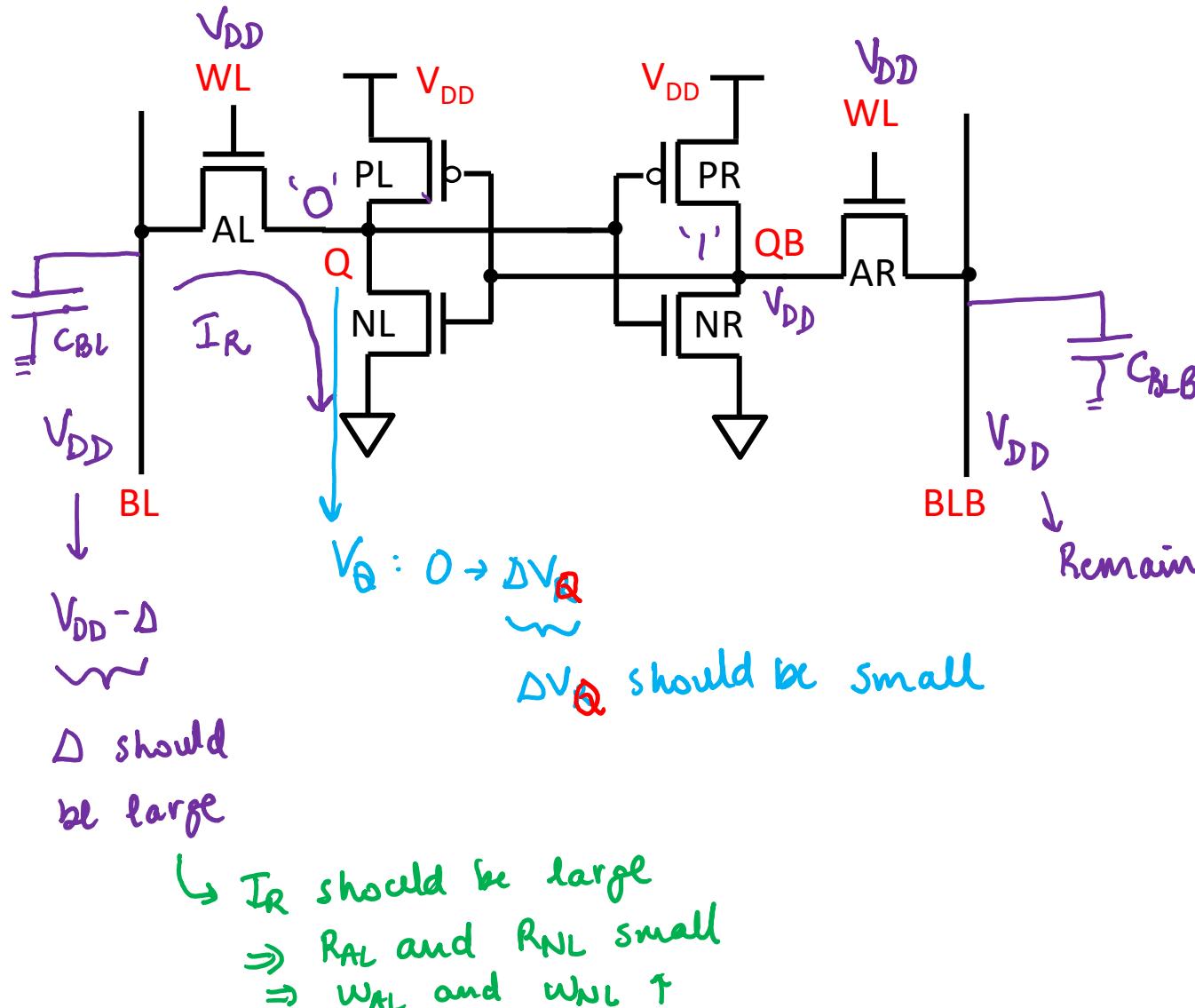


6-TRANSISTOR SRAM

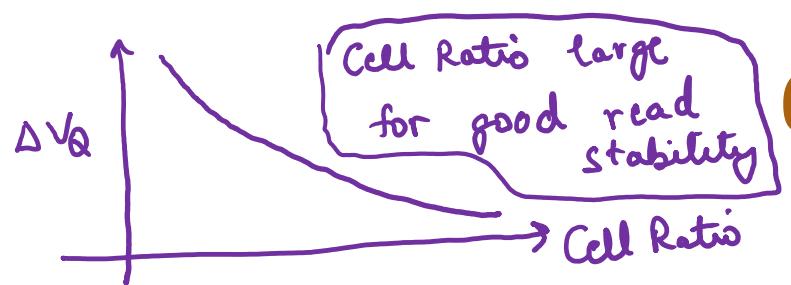


- NL-PL and NR-PR form the cross-coupled inverters ←
- AL and AR are the access transistors
- Differential cell: Has two complementary bit-lines
 - Write: Data and its complement applied on the bit-lines
 - Read: Data and its complement read on the bit-lines
- Symmetric: Must be sized so that the left and right half are the same
 - $W_{NR} = W_{NL}$ ← Pull down transistors
 - $W_{PR} = W_{PL}$ ← Pull up transistors
 - $W_{AR} = W_{AL}$ ← Access transistors

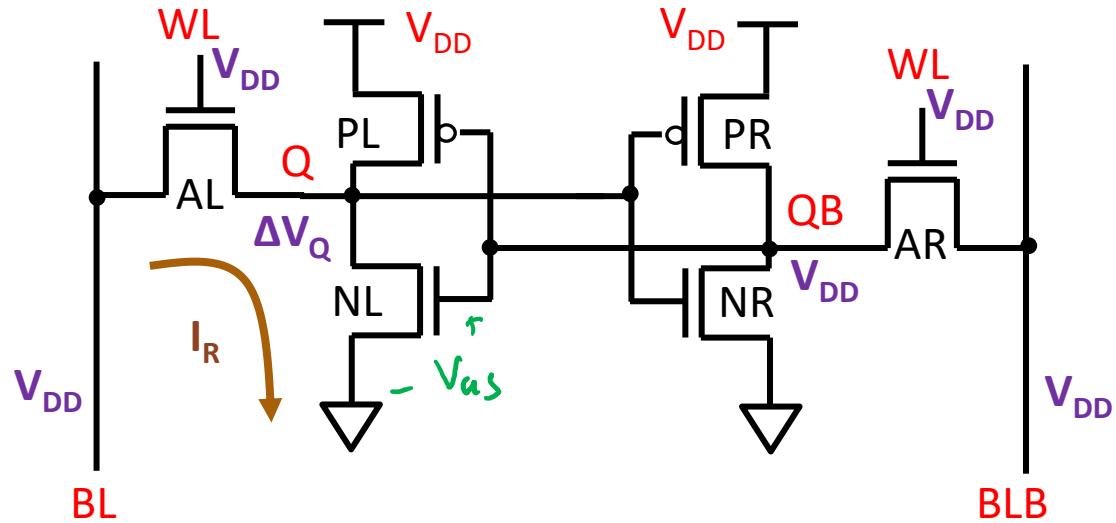
6-TRANSISTOR SRAM: READ



- Pre-charge BL and BLB to V_{DD}
- Assert WL
- Voltage differential (Δ) developed on BL and BLB
 - $Q='0'/QB='1'$: BL discharges and BLB remains at V_{DD}
 - $Q='1'/QB='0'$: BLB discharges and BL remains at V_{DD}
- Δ amplifier by sense amplifier



6-TRANSISTOR SRAM: CELL RATIO



Assuming $V_{BL} \sim V_{DD}$ (ignoring Δ)

$$\Delta V_Q = \frac{R_{NL}}{R_{NL} + R_{AL}} V_{DD} \rightarrow R_{NL} < R_{AL}$$

$\Rightarrow \frac{R_{NL}}{R_{AL}}$ should be small

For avoiding read disturb,

$$\rightarrow \Delta V_Q < V_M, PR-NR \leftarrow$$

Assume: $V_{BL} \sim V_{DD}$

$$I_R = I_{AL} = I_{NL}$$

NL: $V_{GS} = V_{DD}, V_{DS} = \Delta V_Q \rightarrow$ Assume linear
 $(\because \Delta V_Q \text{ expected to be small})$

AL: $V_{GS} = V_{DD} - \Delta V_Q, V_{DS} = V_{DD} - \Delta V_Q : \text{Saturation}$

Assume pinch-off saturation, $\lambda_N = 0$

$$I_{AL} = \frac{k_N}{2} \left(\frac{W}{L} \right)_{AL} (V_{GS} - V_{THN})^2$$

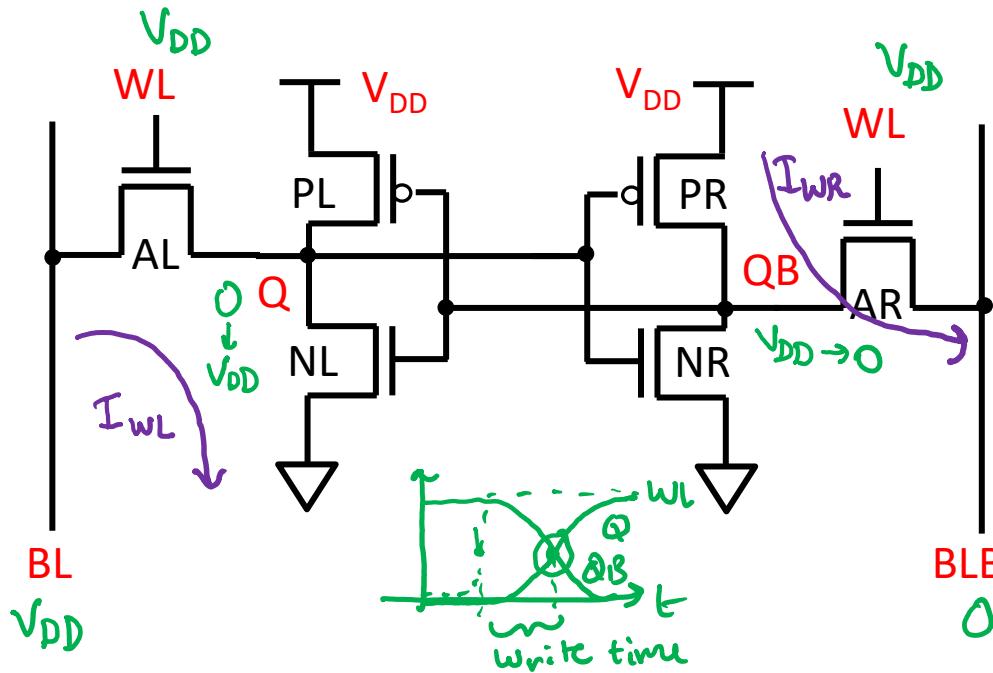
$$I_{NL} = k_N \left(\frac{W}{L} \right)_{NL} \left[(V_{GS} - V_{THN}) V_{DSN} - \frac{V_{DSN}^2}{2} \right]$$

$$\frac{k_N}{2} \left(\frac{W}{L} \right)_{AL} (V_{GS} - V_{THN})^2 = k_N \left(\frac{W}{L} \right)_{NL} \left[(V_{GS} - V_{THN}) V_{DSN} - \frac{V_{DSN}^2}{2} \right]$$

$$\frac{1}{2} \left(\frac{W}{L} \right)_{AL} (V_{DD} - \Delta V_Q - V_{THN})^2 = \left(\frac{W}{L} \right)_{NL} \left[(V_{DD} - V_{THN}) \Delta V_Q - \frac{\Delta V_Q^2}{2} \right]$$

$\Delta V_Q \sim f_1(\text{Cell Ratio})$ or $\text{cell ratio} \sim f_2(\Delta V_Q)$

6-TRANSISTOR SRAM: WRITE



- Drive BL to Data and BLB to Data'
 - Assert WL
 - Q and QB switch to the desired value

$$V_{QB} = \frac{R_{AR}}{R_{AR} + R_{pR}} V_{DD} \quad \rightarrow \quad R_{AR} < R_{pR}$$

$$W_A > W_p$$

For successful write, $v_{AB} < \sqrt{M_1 \cdot PL - NL \cdot AL}$

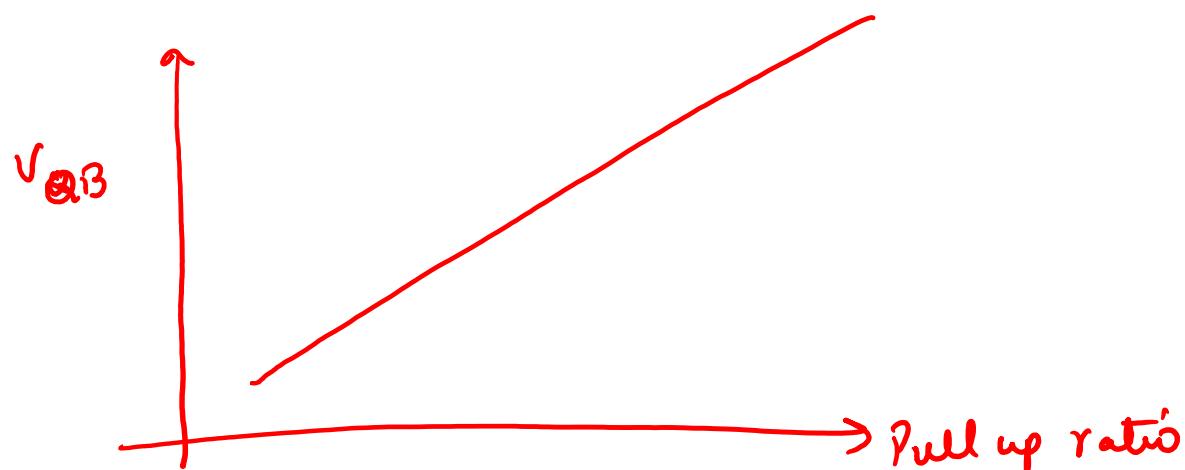
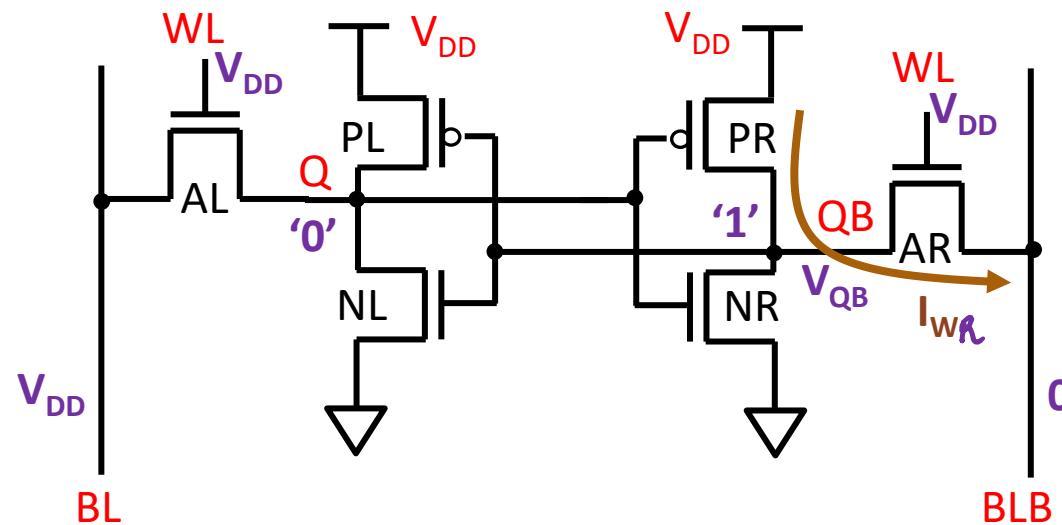
- Q: $0 \rightarrow V_{DD}$
 - NR turns ON, PR turns OFF
 - QB : $\Delta V_{QB} \rightarrow 0$

write time: Time it takes to switch Q and QB

① I_{WL} is similar (but not same) as $\underline{I_R}$
 W_{NL} & W_{RL} is dictated by read stability
 I_{UR} alone is not going to write

② Initially, I_{WR} is the main driving force for write

6-TRANSISTOR SRAM: PULL-UP RATIO



$$\text{Pull up ratio} = \frac{(W/L)_P}{(W/L)_A}$$

Exercise: Use MOSFET equations to obtain $V_{Q_B} \sim f_i$ (pull up ratio)

From the perspective of read stability and writeability,

$V_{M,P-N}$ should be large
 $\rightarrow W_P \uparrow, W_N \downarrow$

But this is in conflict with cell ratio and pull up ratio needs

\rightarrow Cell Ratio & pull-up ratio more dominant over V_M needs

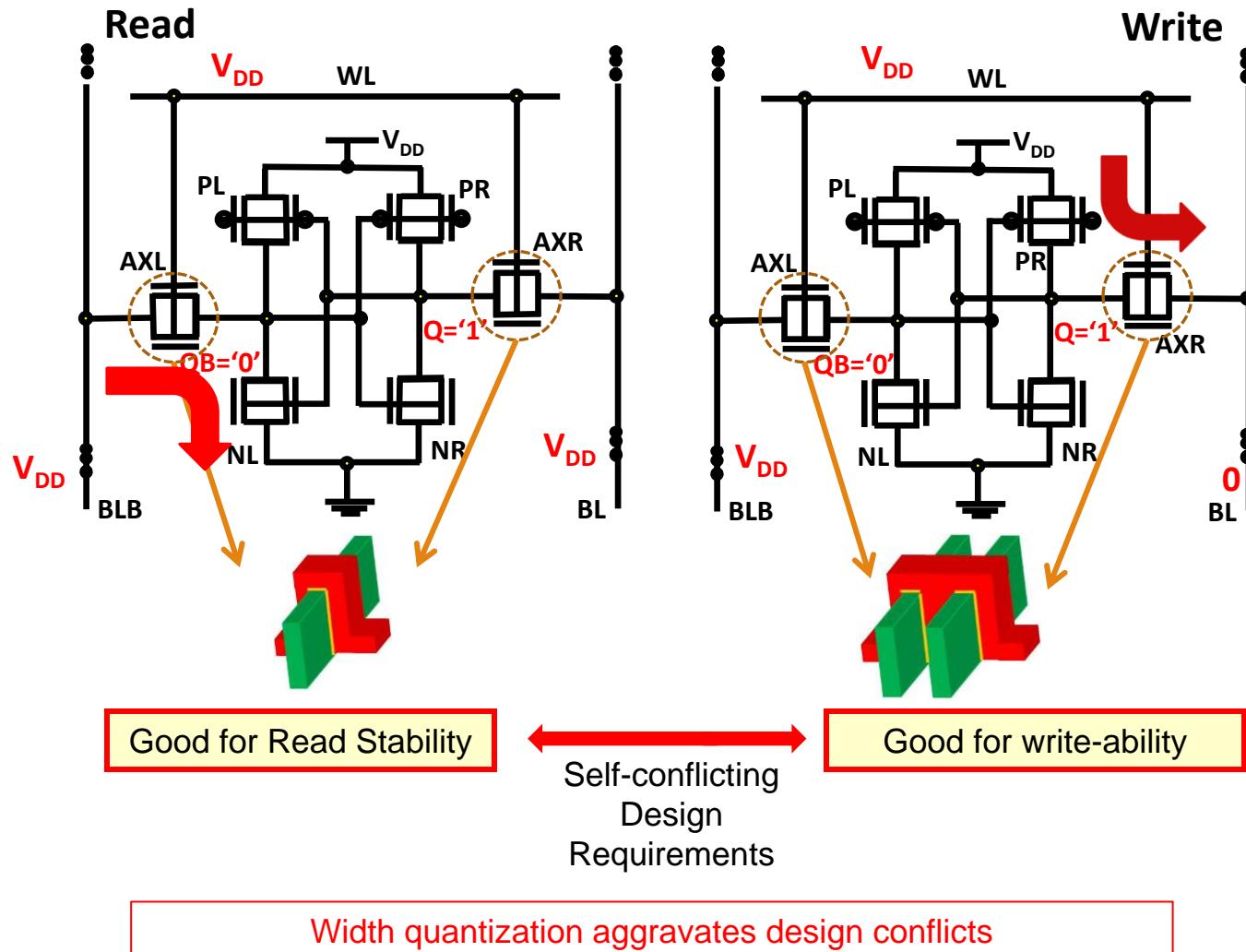
METRICS

$$W_p : W_A : W_N = 1 : 1.5 : 2$$

\uparrow
 W_{min}

- Hold stability → Balanced pull-up & pull-down — $V_M = V_{DD}/2$
 - Read stability → $W_N \uparrow, W_A \downarrow, W_p \uparrow$
 - Write-ability/write speed → $W_A \uparrow, W_p \downarrow, W_N \downarrow$
 - Access time/read speed → $W_N \uparrow, W_A \uparrow$
 - Cell leakage → $W_N \downarrow, W_p \downarrow, W_A \downarrow$
 - Cell area → $W_N \downarrow, W_p \downarrow, W_A \downarrow$
 - Read/write power
- As close to min-sized as possible
- ① BL/BLB charging/Discharging → $C_{BL} V_{DD} \Delta V_{swing}$
- ② WL switching → $C_{WL} V_{DD} \Delta V_{WL}$
- ③ Cell energy/power
- ④ Peripheral circuits

FINFET-BASED SRAMS



Bulk MOSfets

$$W_p : W_A : W_N = 1 : 1.5 : 2$$

FinFets

$$N_p : N_A : N_N = 1 : 1 : 2$$

Good for WM

$$= 1 : 2 : 2$$

$$= 1 : 1 : 1$$

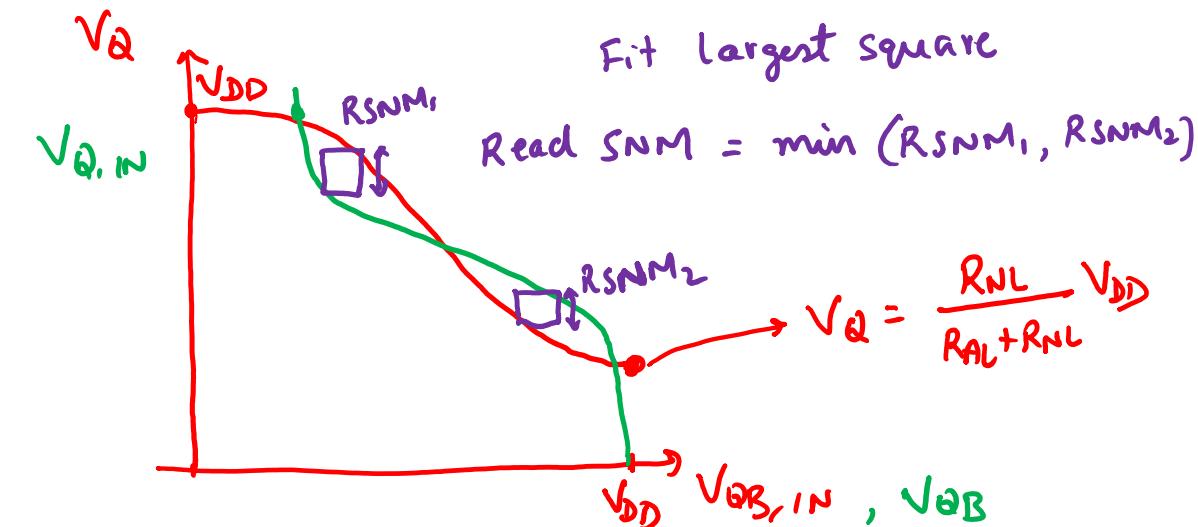
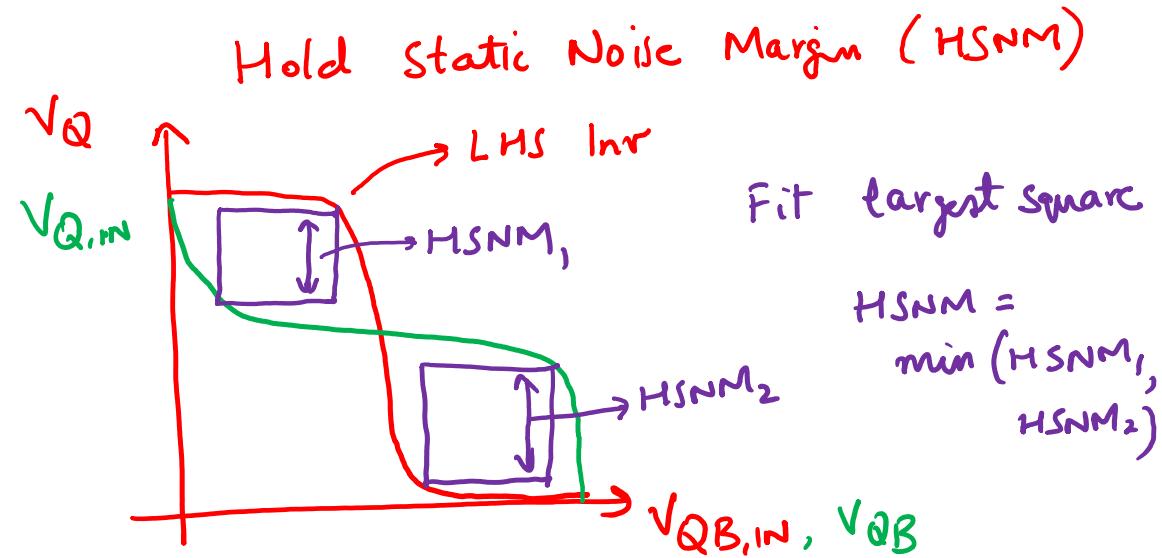
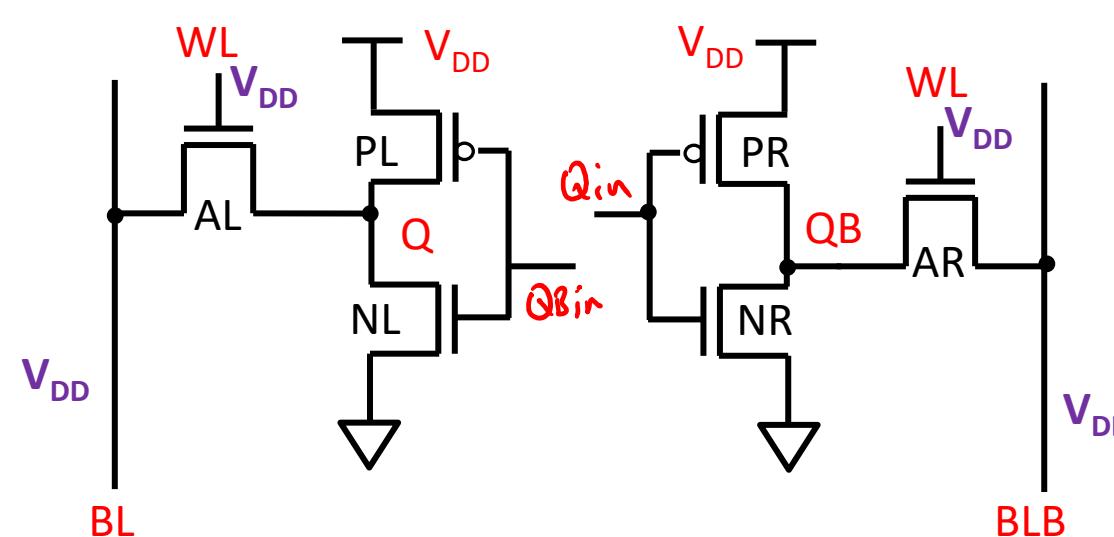
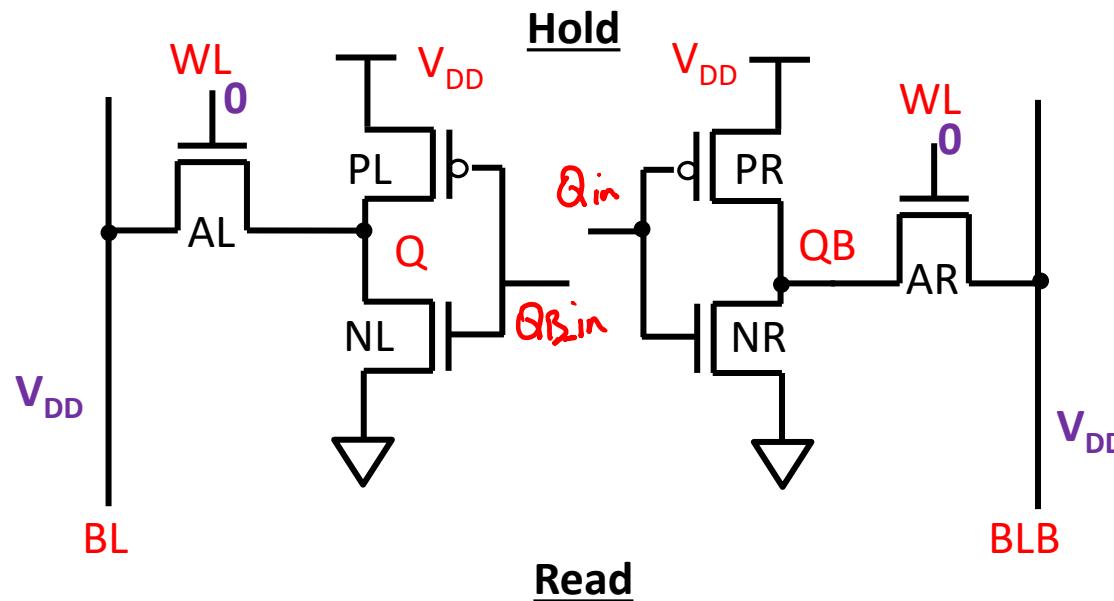
Good for RSNM

HD

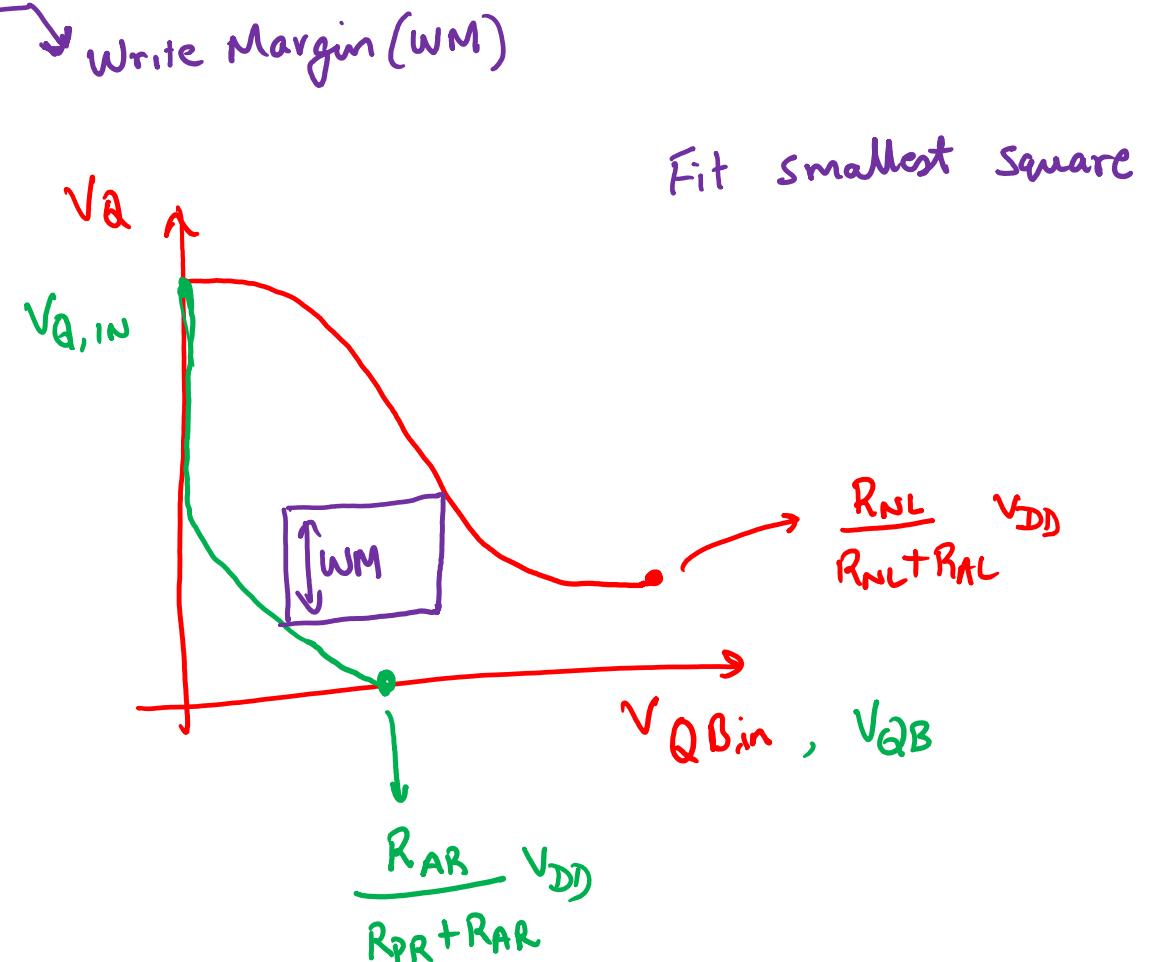
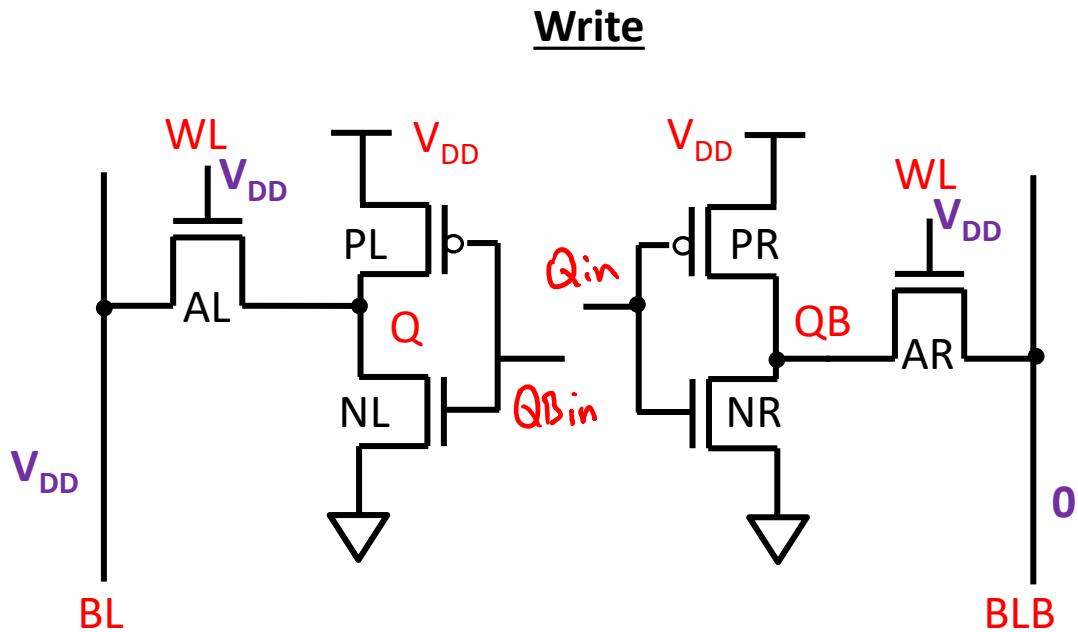
+

Read/write Assist Techniques

READ/HOLD STABILITY ANALYSIS WITH BUTTERFLY CURVES



DC WRITE-ABILITY ANALYSIS WITH BUTTERFLY CURVES



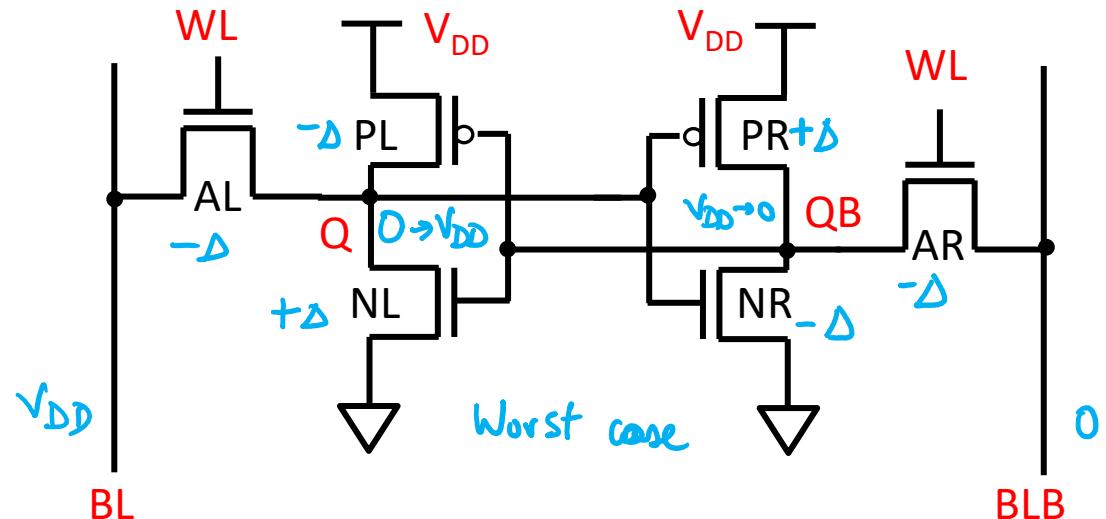
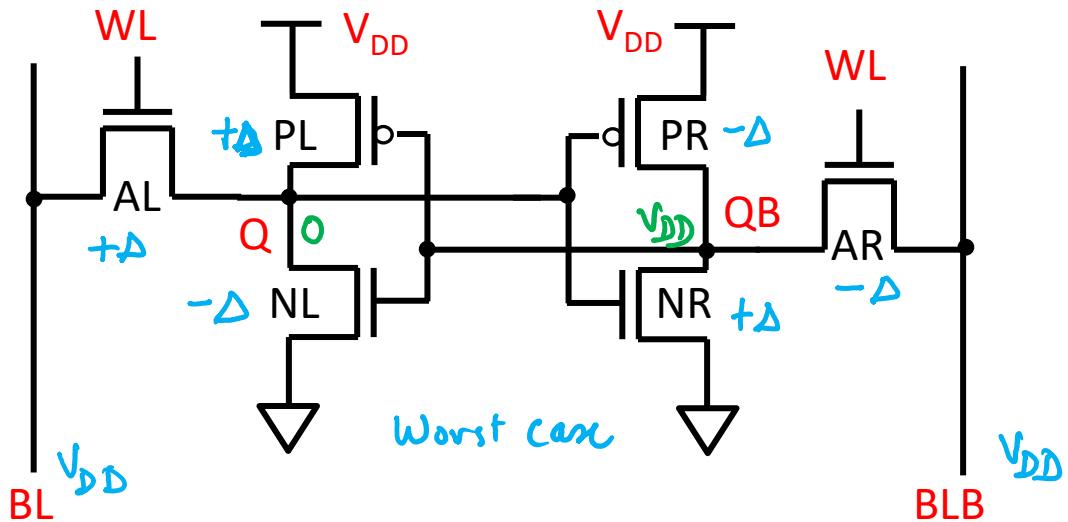
6-TRANSISTOR SRAM: VARIATIONS

→ Min-sized transistors (prone to random variations)

Design Conflicts

Write

Read/Hold



① Global Variations : Process corners (T_{Np} , $F_N S_p$, $S_N F_p$, $f_N F_p$, $S_N S_p$)

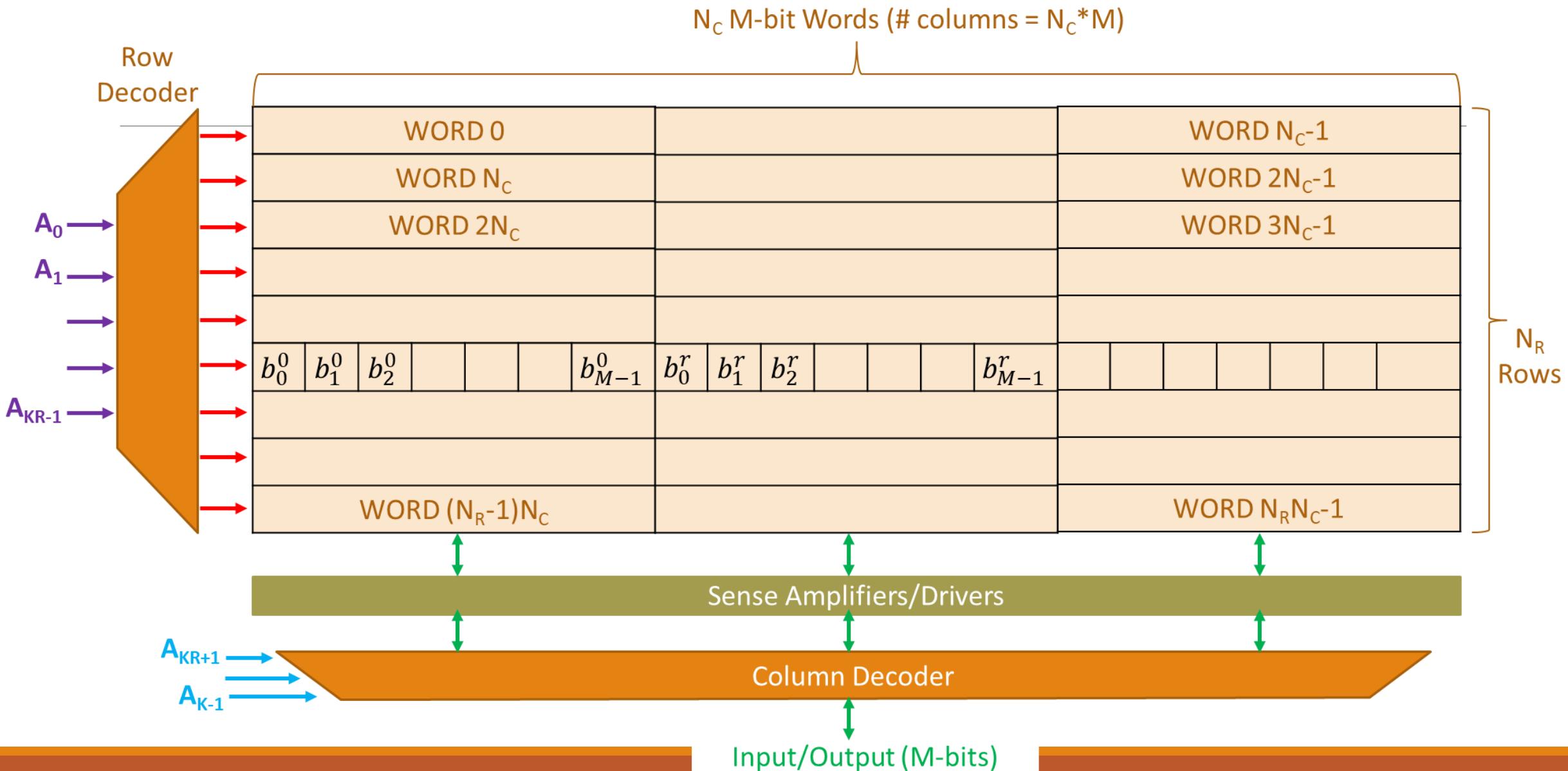
→ Read / Hold : Worst case $\rightarrow F_N S_p$

→ Write : Worst case $\rightarrow S_N F_p$

② Random Variations : $+ \Delta \rightarrow$ Transistor becomes stronger $\Rightarrow (V_{THN}/V_{THP}) \downarrow$
 $- \Delta \rightarrow$ Transistor becomes weaker $\Rightarrow (V_{THN}/V_{THP}) \uparrow$

PERIPHERAL CIRCUITS

RECAP: MEMORY ARCHITECTURE



ROW DECODERS

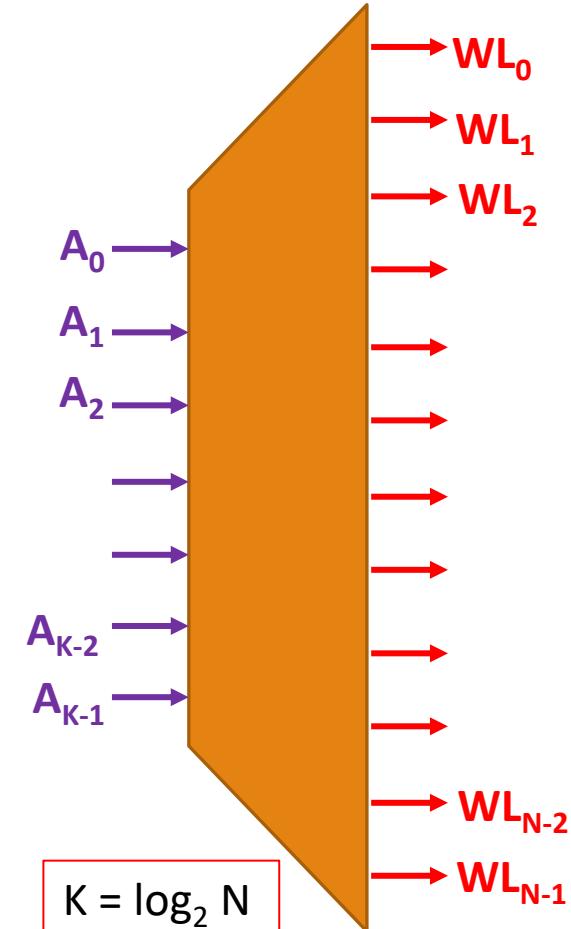
(N)AND Decoder (Example for an 8-to-256-bit decoder)

- $WL_0 = A_7'.A_6'.A_5'.A_4'.A_3'.A_2'.A_1'.A_0' \leftarrow$
- $WL_1 = A_7'.A_6'.A_5'.A_4'.A_3'.A_2'.A_1'.A_0$
- $WL_{127} = A_7'.A_6.A_5.A_4.A_3.A_2.A_1.A_0 \leftarrow$
- $WL_{255} = A_7.A_6.A_5.A_4.A_3.A_2.A_1.A_0$

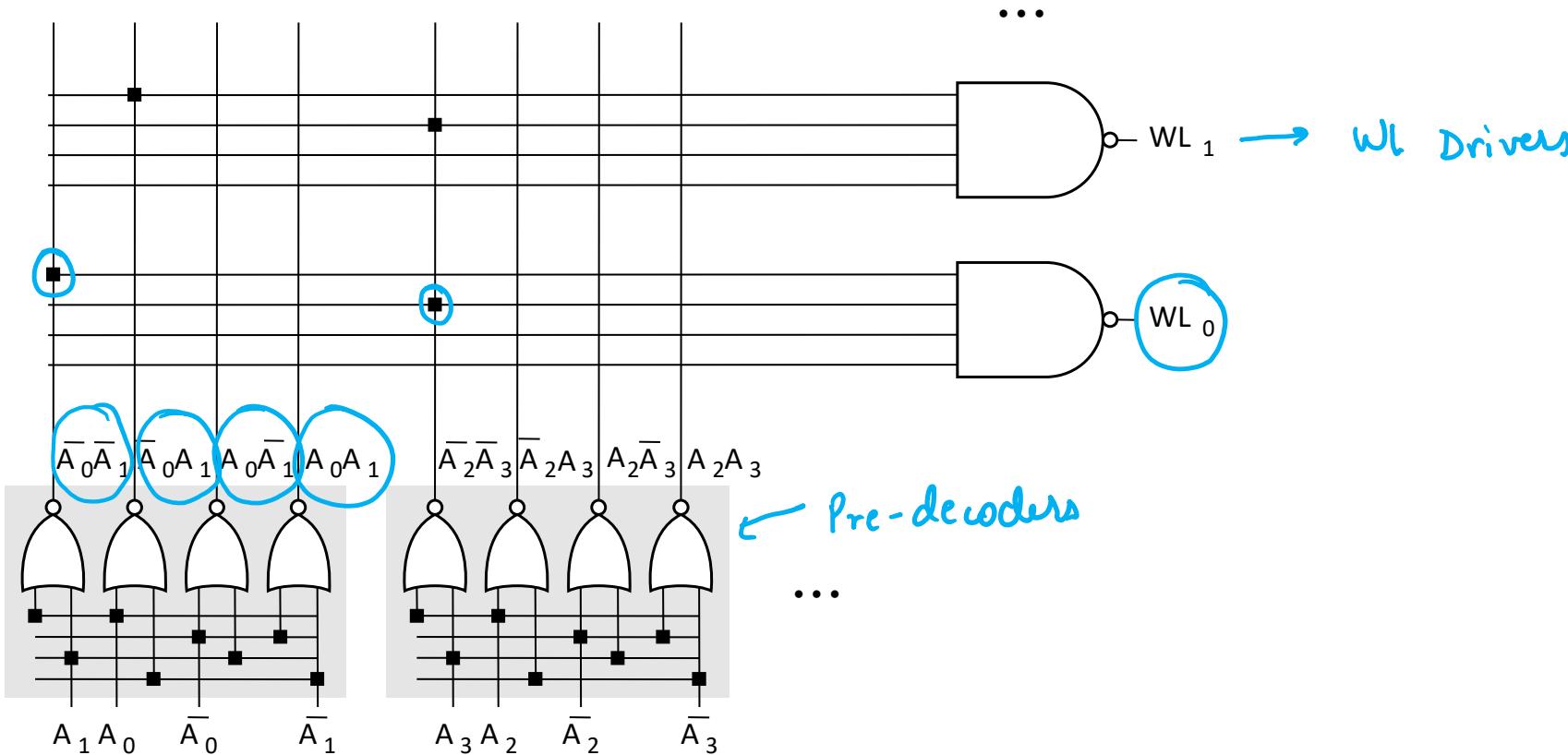
↓ De-Morgan's Theorem

(N)OR Decoder (Example for an 8-to-256-bit decoder)

- $WL_0 = (A_7 + A_6 + A_5 + A_4 + A_3 + A_2 + A_1 + A_0)' \leftarrow$
- $WL_1 = (A_7 + A_6 + A_5 + A_4 + A_3 + A_2 + A_1 + A_0)'$
- $WL_{127} = (A_7 + A_6' + A_5' + A_4' + A_3' + A_2' + A_1' + A_0')' \leftarrow$
- $WL_{255} = (A_7' + A_6' + A_5' + A_4' + A_3' + A_2' + A_1' + A_0')'$



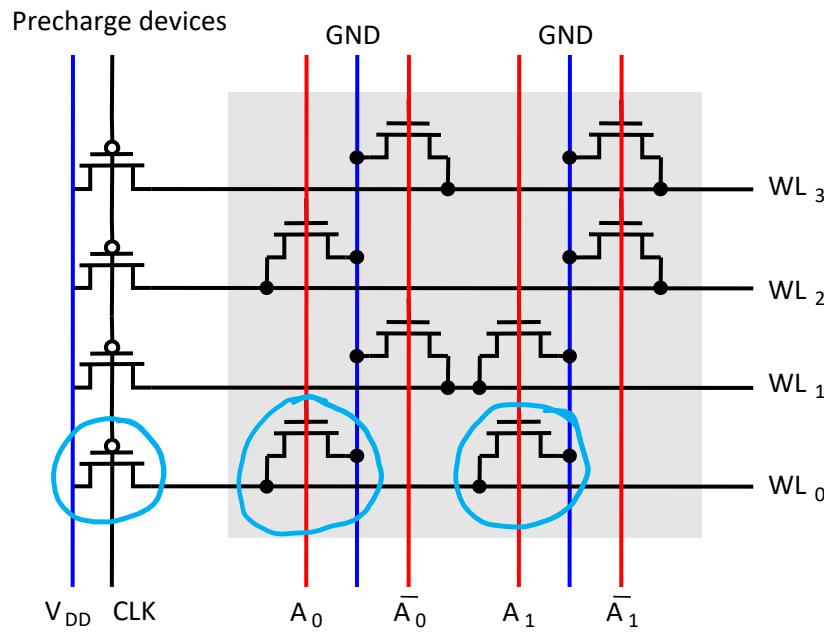
HIERARCHICAL DECODERS



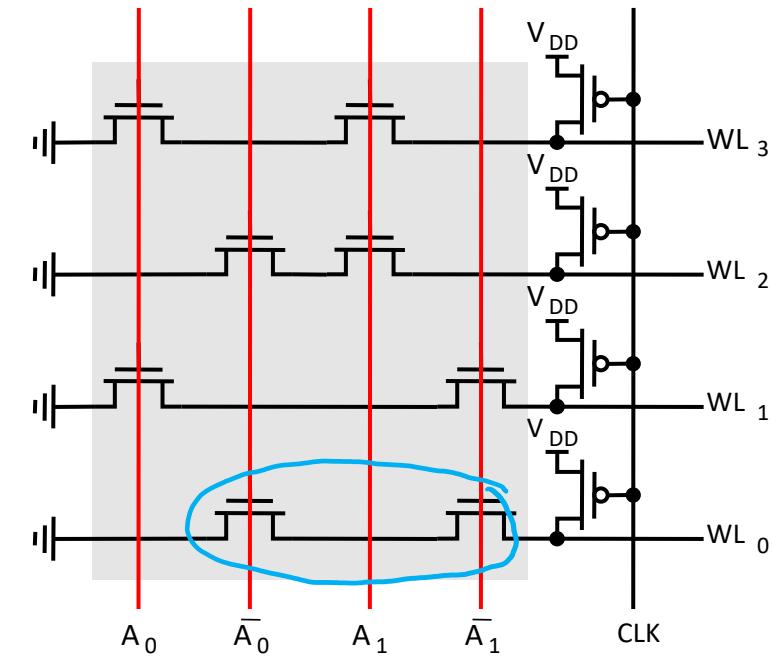
NAND Decoders with 2-input pre-decoders

DYNAMIC DECODERS

Exercise: why is there no footer transistor?

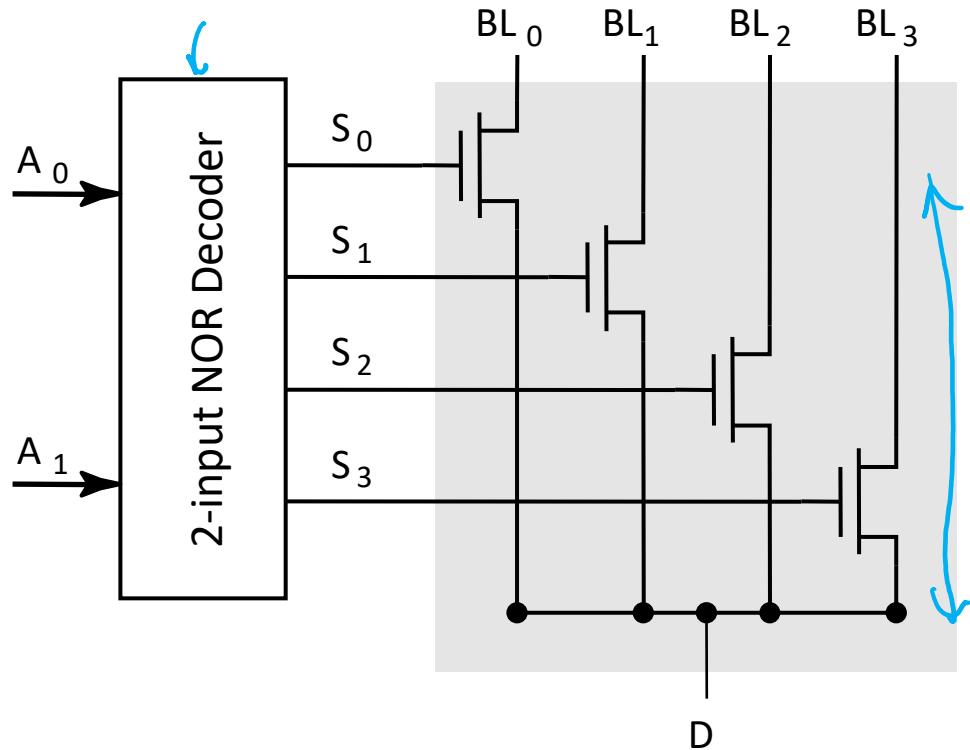


2-input NOR decoder



2-input NAND decoder

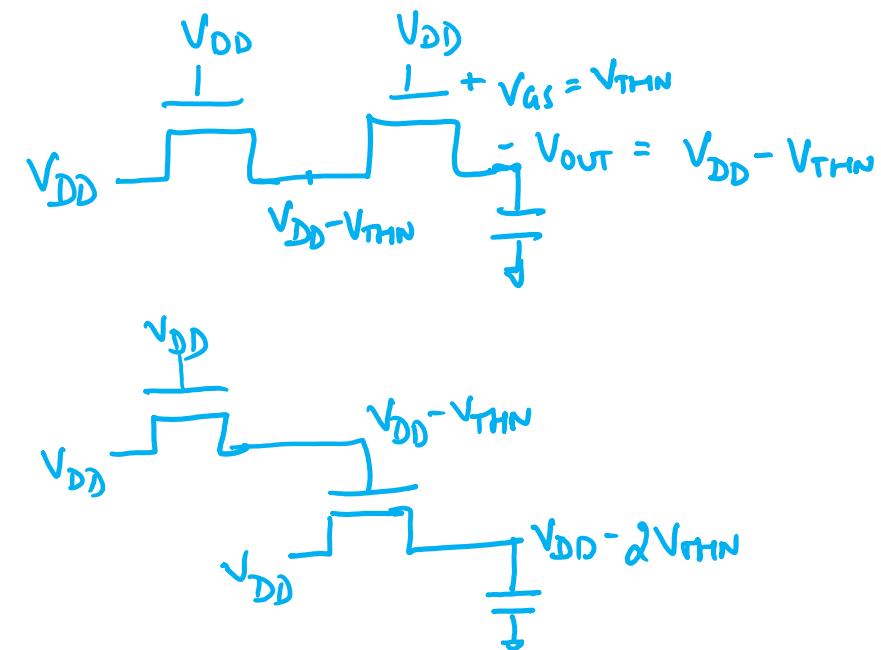
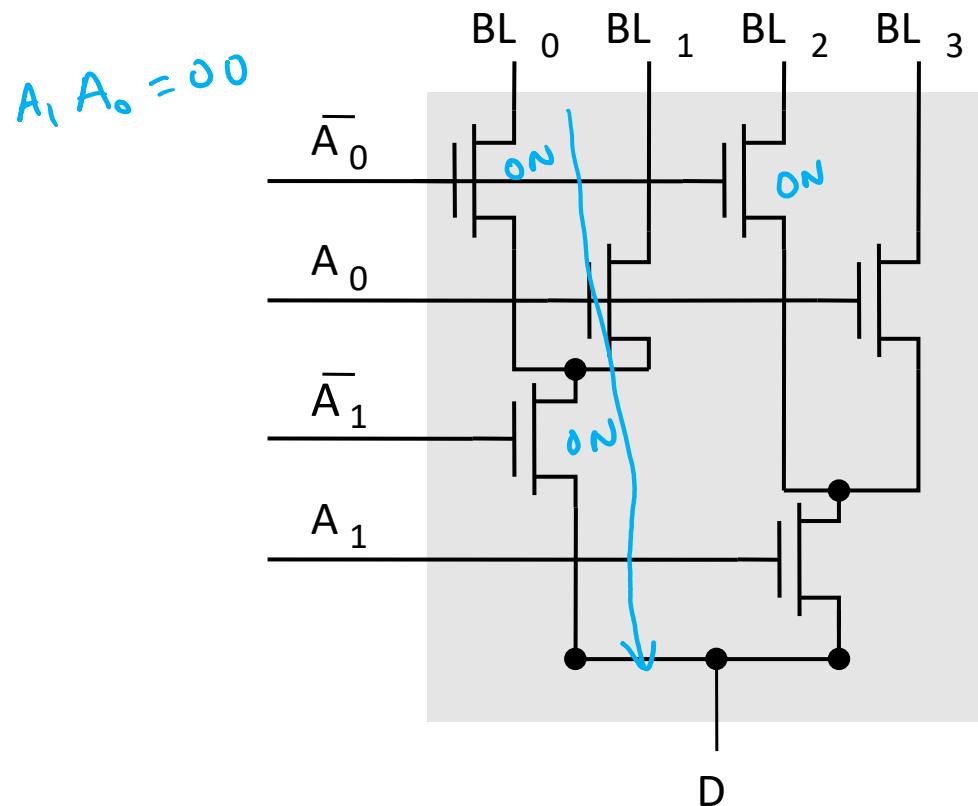
PASS-TRANSISTOR BASED COLUMN DECODER



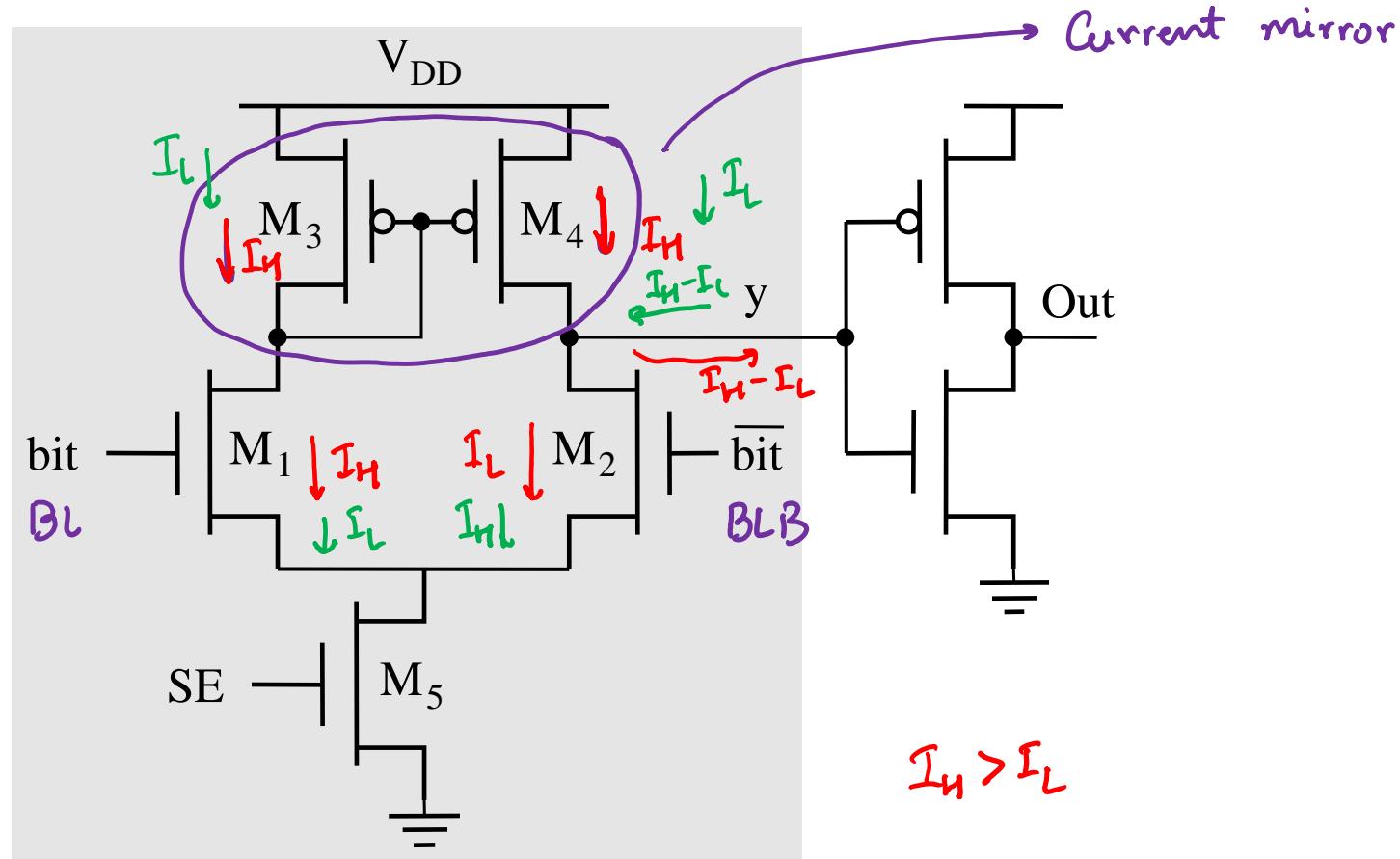
$D \rightarrow BL$: write
 $BL \rightarrow D$: read

Only one pass transistor is ON at a time

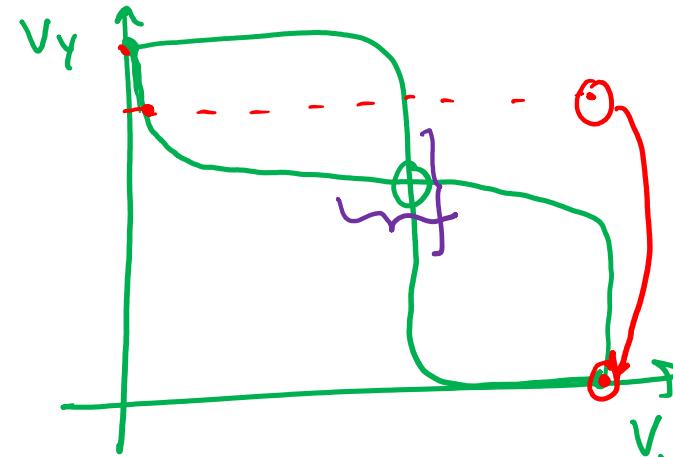
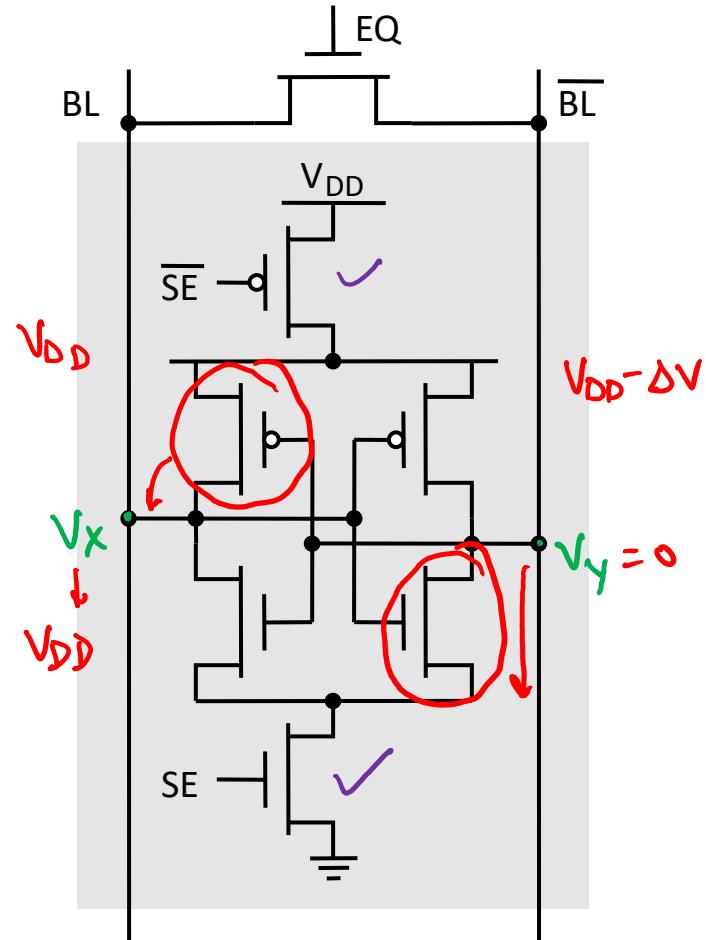
TREE BASED COLUMN DECODER



DIFFERENTIAL SENSE AMPLIFIER

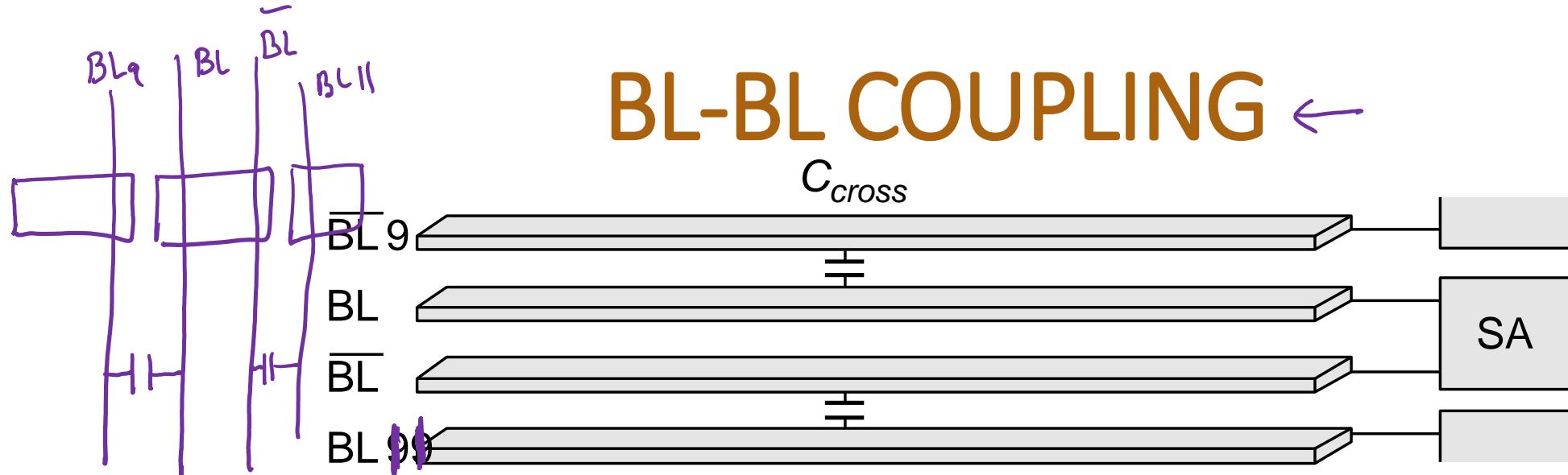


LATCH-BASED SENSE AMPLIFIER AMPLIFIER

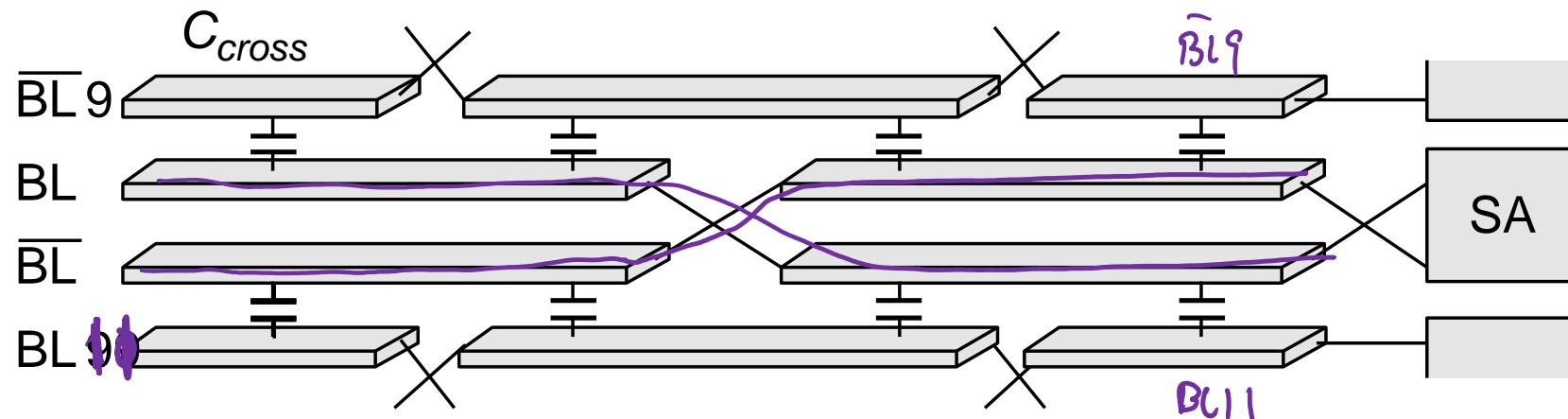


OTHER COMMON DESIGN ASPECTS

BL-BL COUPLING ←

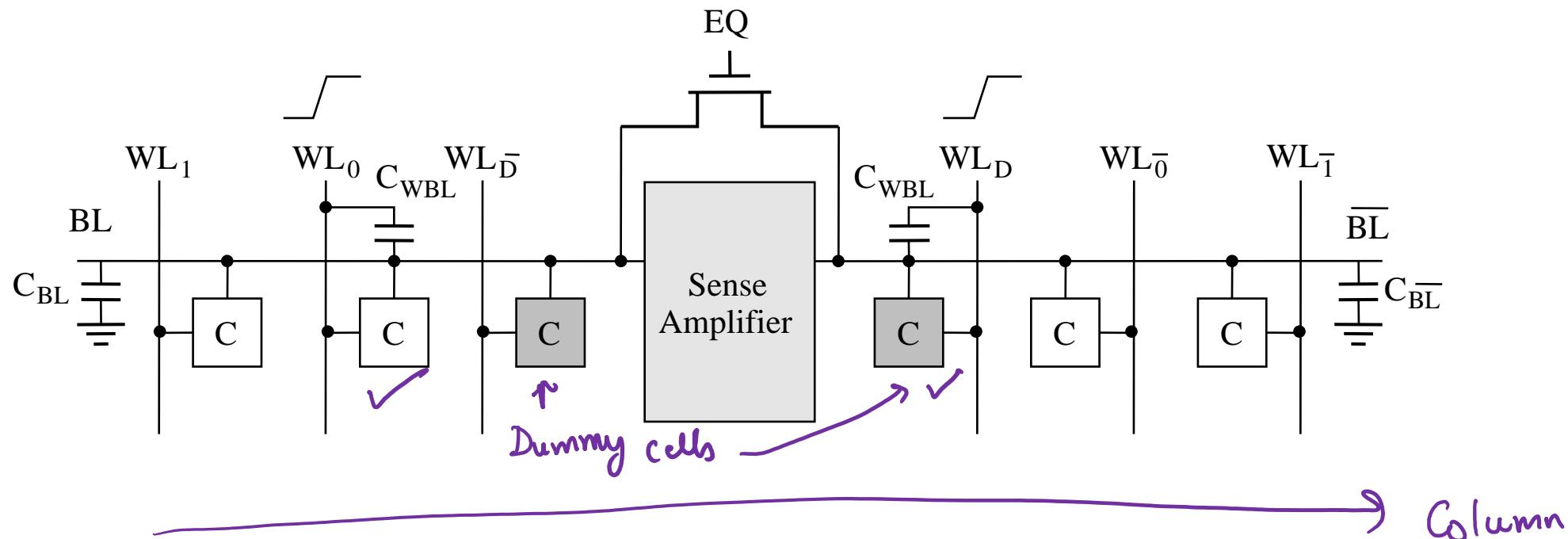


(a) Straightforward bit-line routing

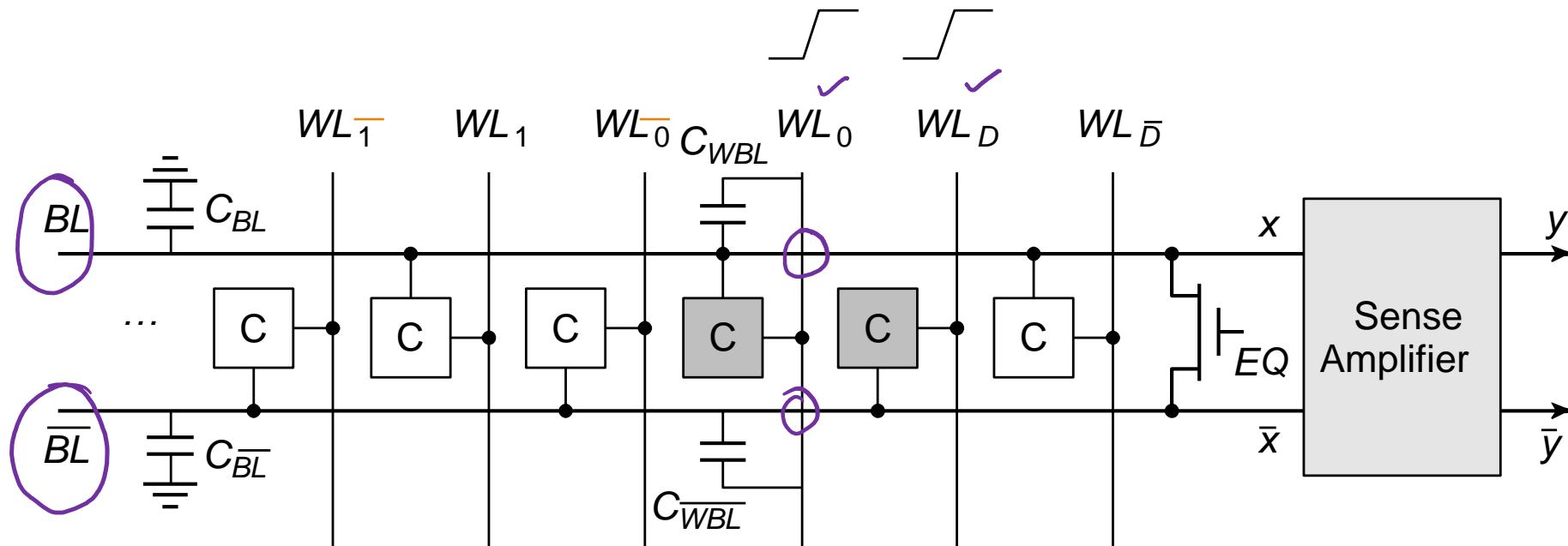


(b) Transposed bit-line architecture

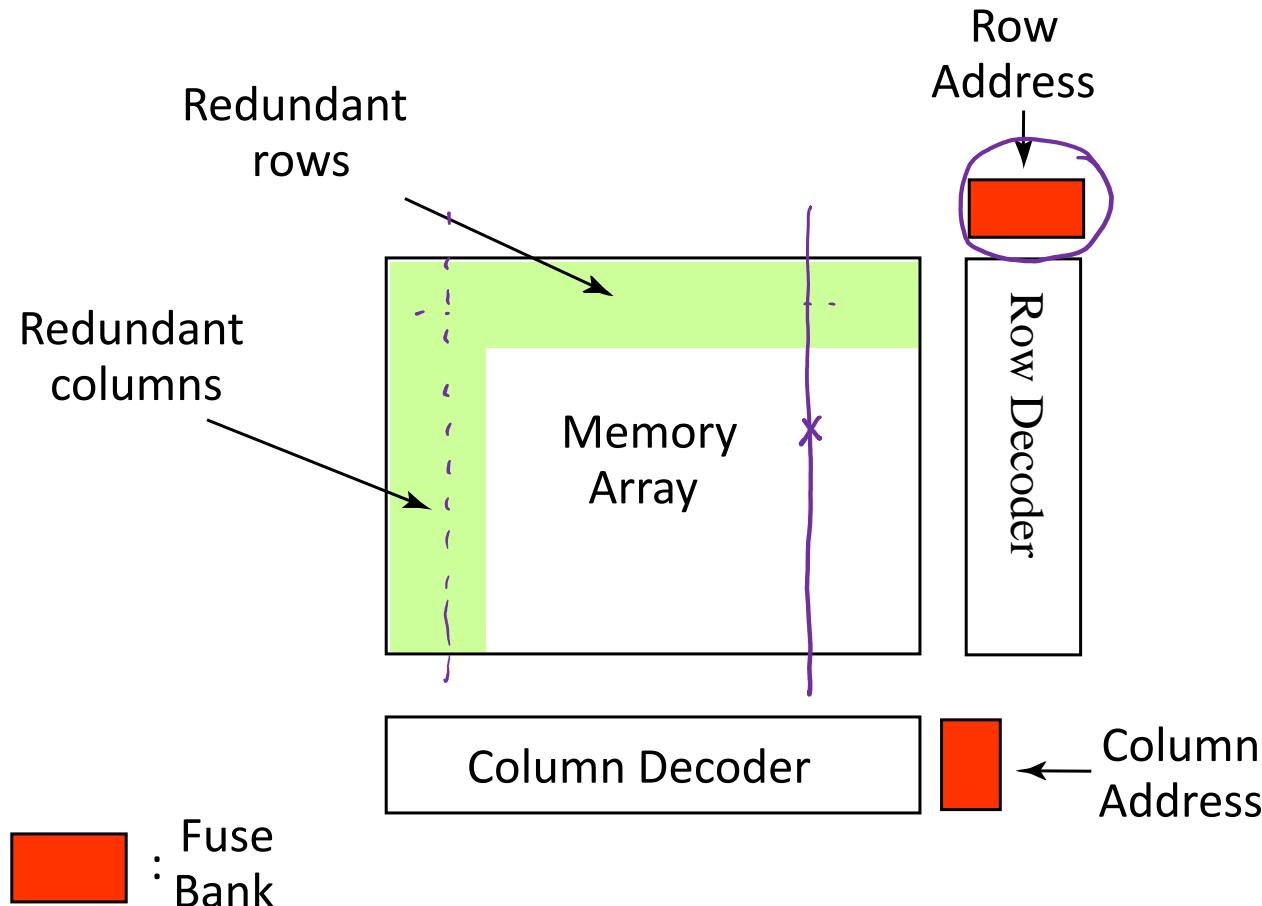
OPEN BIT-LINE ARCHITECTURE – WL-BL COUPLING



FOLDED-BITLINE ARCHITECTURE



REDUNDANCY AND ERROR DETECTION/CORRECTION



Example: Hamming Codes

$\underbrace{P_1 P_2}_{\text{Parity}} \underbrace{B_3 P_4}_{\text{with}} B_5 B_6 B_7$

e.g. B3 Wrong

$$P_1 \oplus B_3 \oplus B_5 \oplus B_7 = 0$$

$$P_2 \oplus B_3 \oplus B_6 \oplus B_7 = 0$$

$$P_4 \oplus B_5 \oplus B_6 \oplus B_7 = 0$$

1

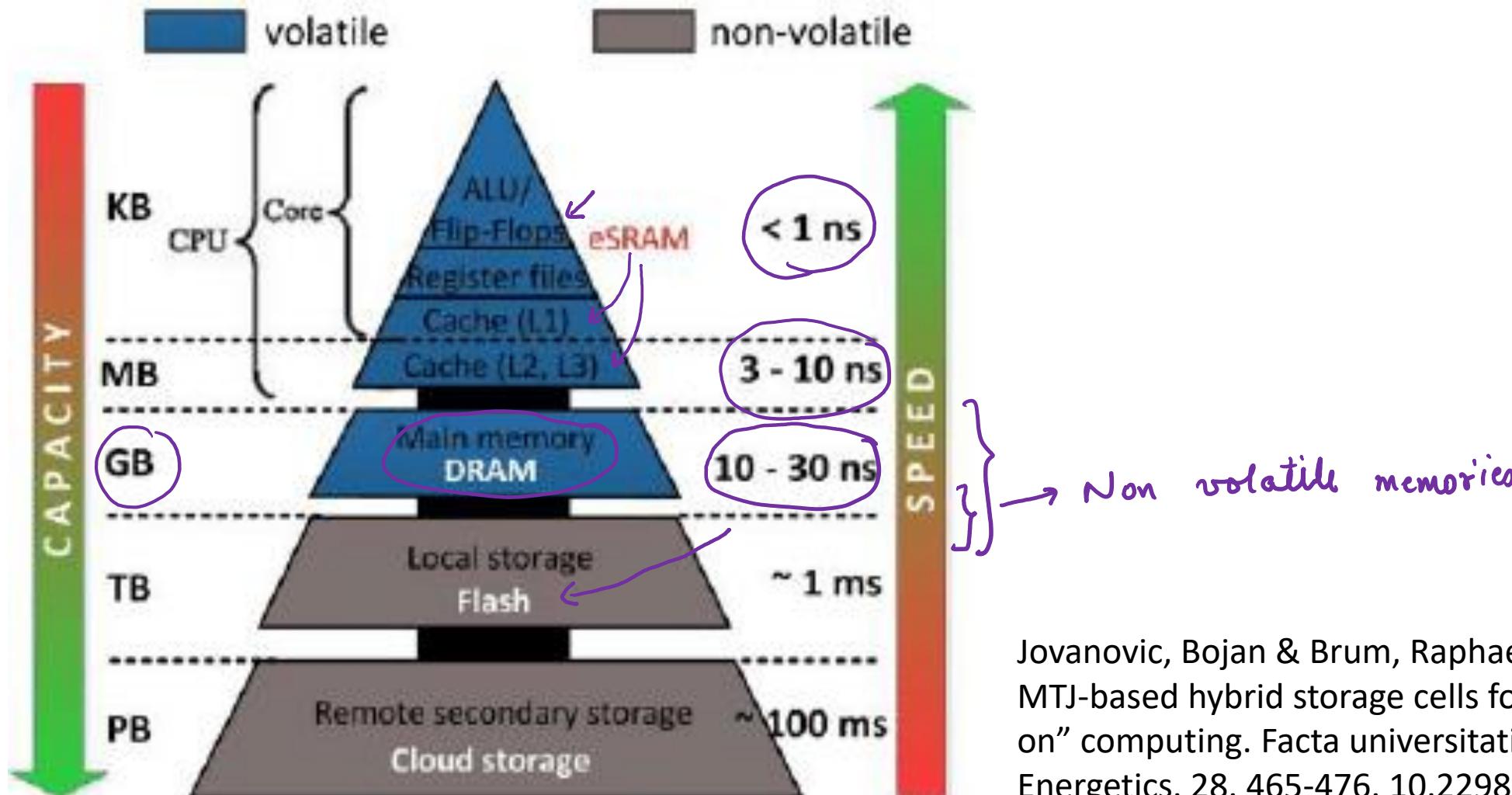
1

0

= 3

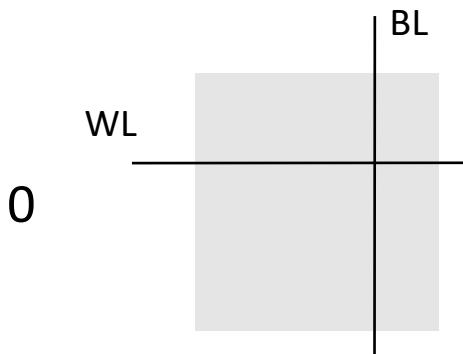
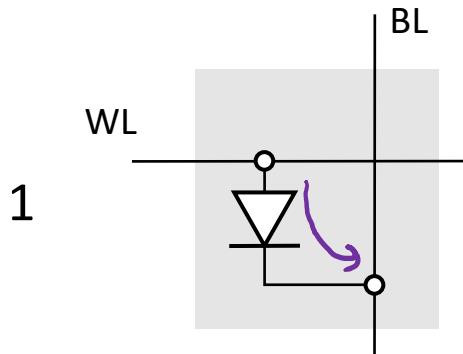
NON-VOLATILE MEMORIES

MEMORY HIERARCHY

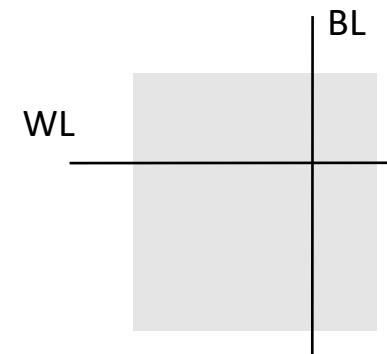
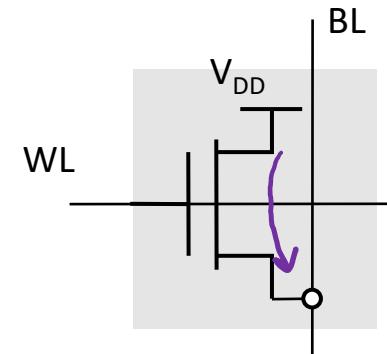


Jovanovic, Bojan & Brum, Raphael & Torres, Lionel. (2015). MTJ-based hybrid storage cells for “normally-off and instant-on” computing. Facta universitatis - series: Electronics and Energetics. 28. 465-476. 10.2298/FUEE1503465J.

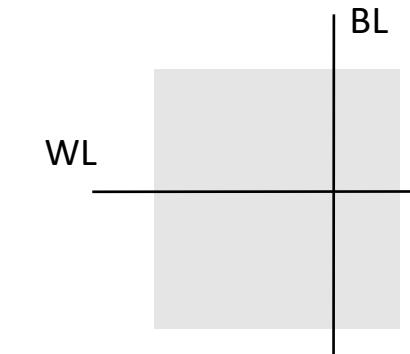
READ-ONLY MEMORY CELLS



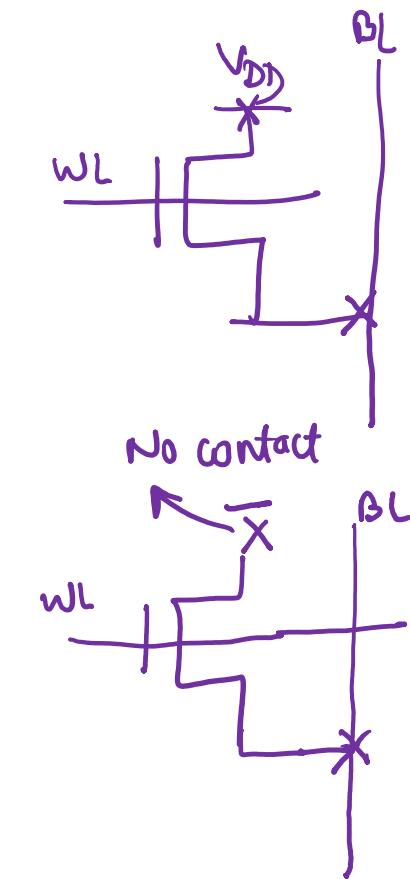
Diode ROM



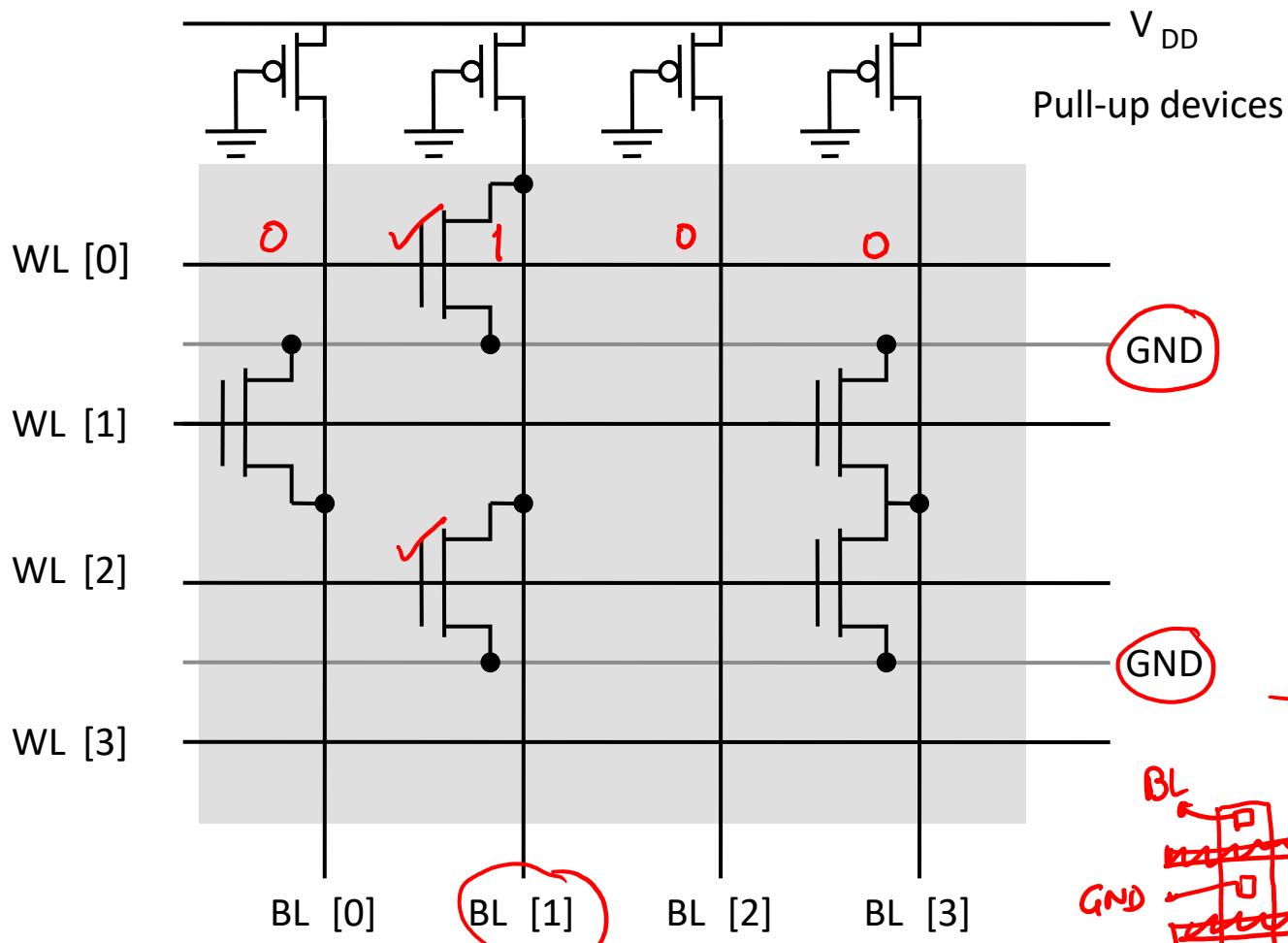
MOS ROM 1



MOS ROM 2



MOS NOR ROM



NAND
↓
NMOS in series
(Offers higher integration density than NOR)

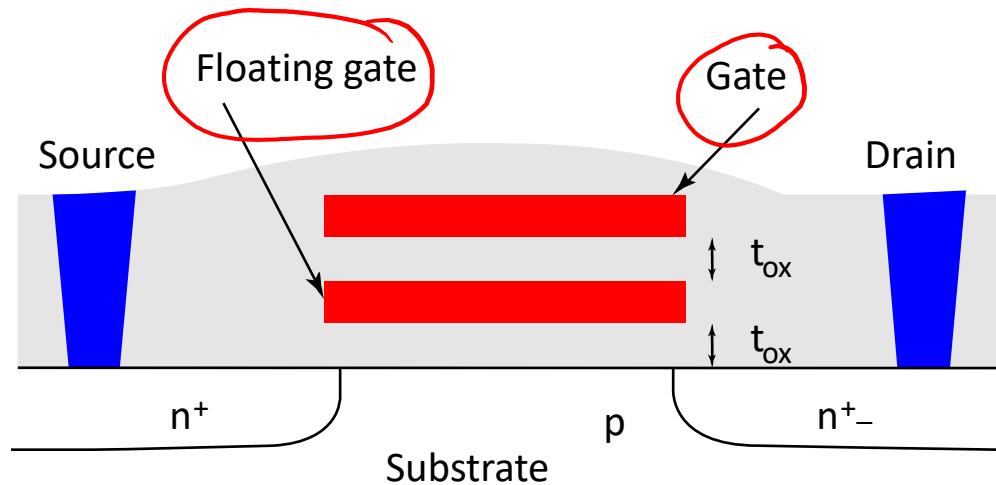
NOR

NAND

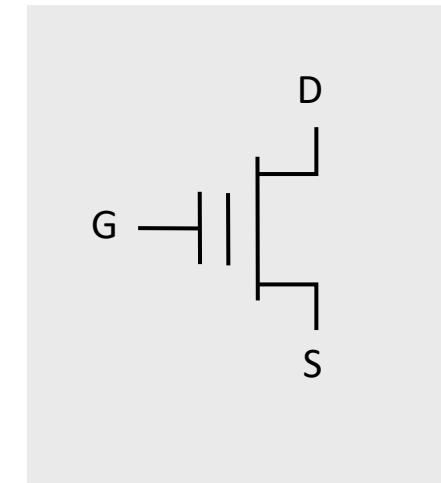
$\Delta H_1 > \Delta H_2$

$\Delta H_2 < \Delta H_1$

NON-VOLATILE MEMORIES: THE FLOATING-GATE TRANSISTOR

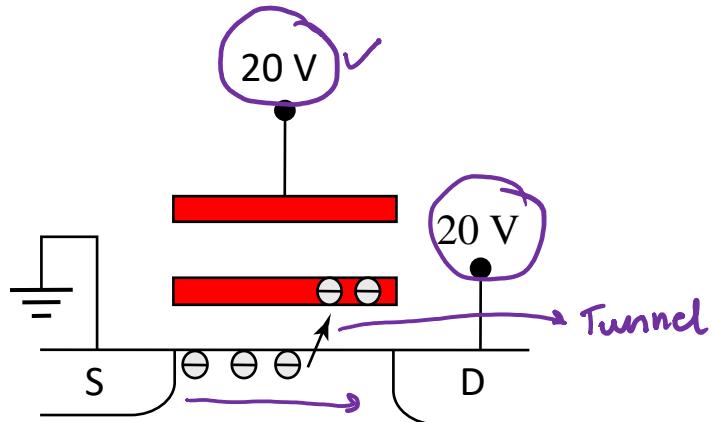


Device cross-section

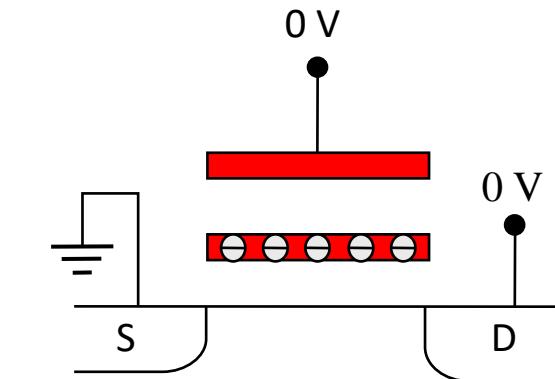


Schematic symbol

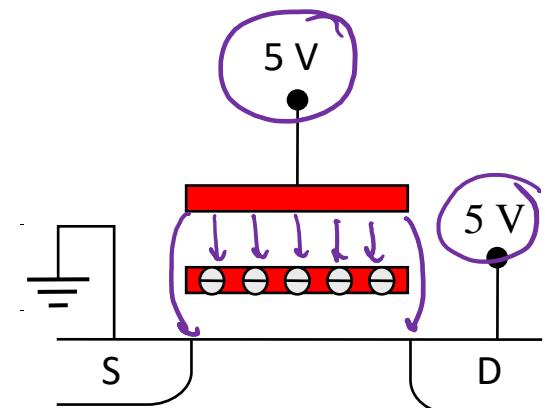
FLOATING-GATE TRANSISTOR PROGRAMMING



Avalanche injection

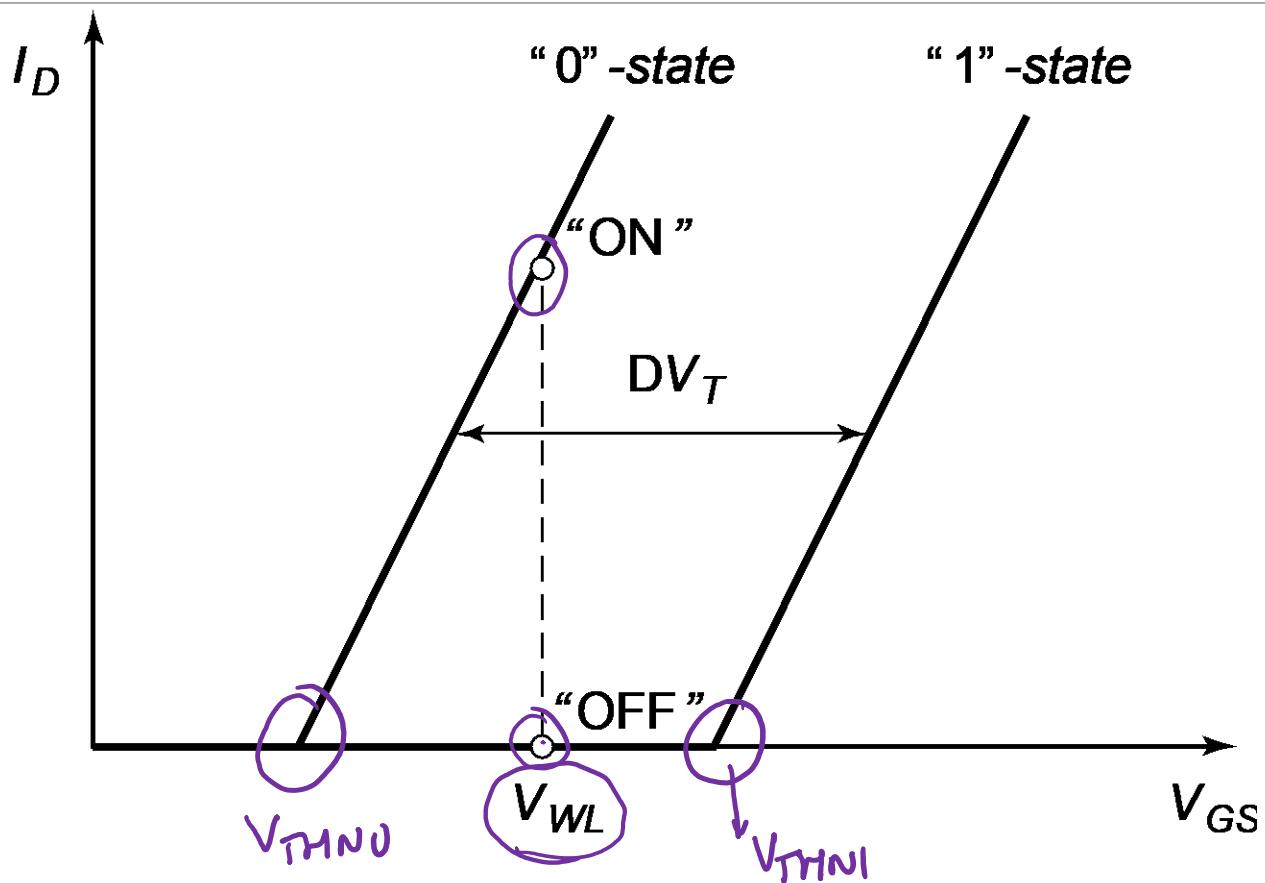


Removing programming voltage leaves charge trapped

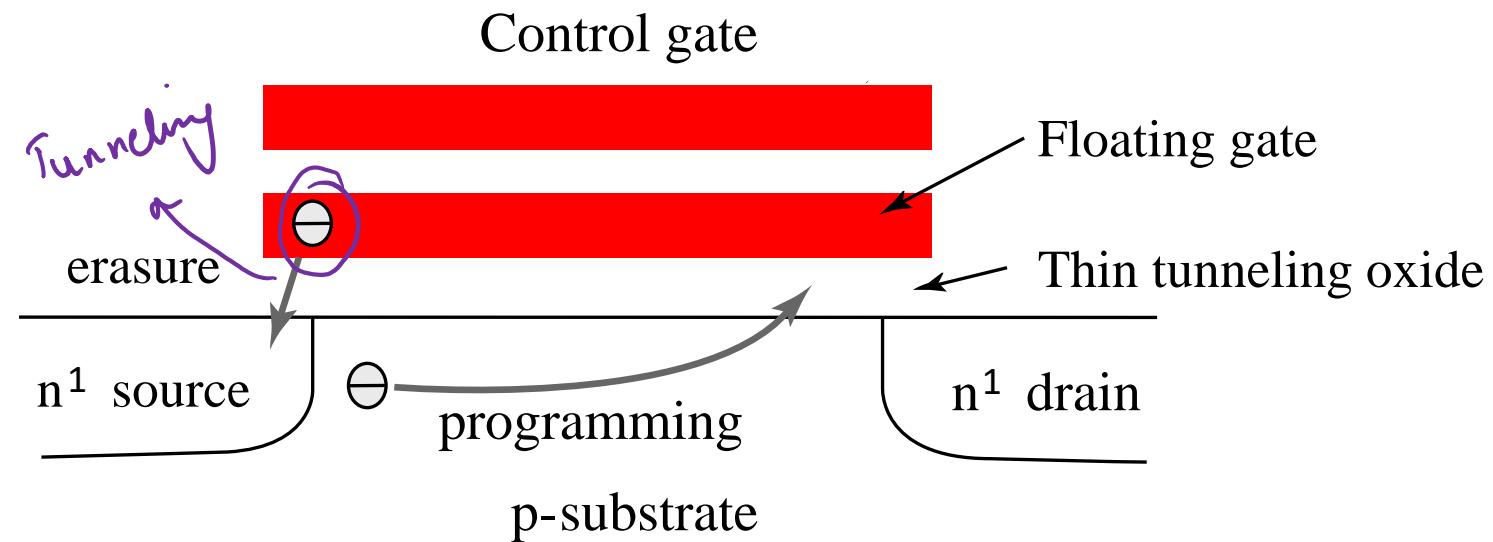


Programming results in
higher V_{THN}

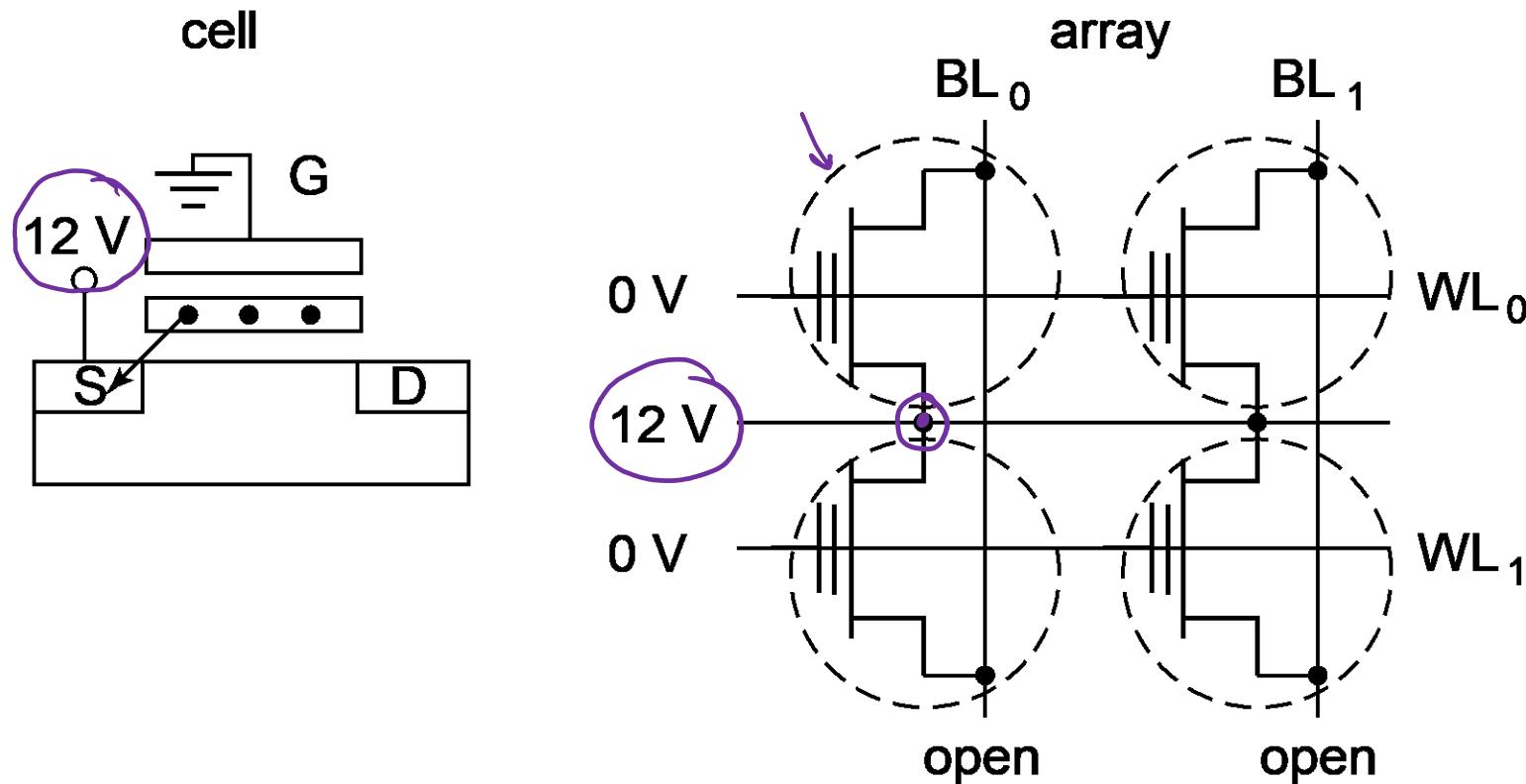
A “PROGRAMMABLE-THRESHOLD” TRANSISTOR



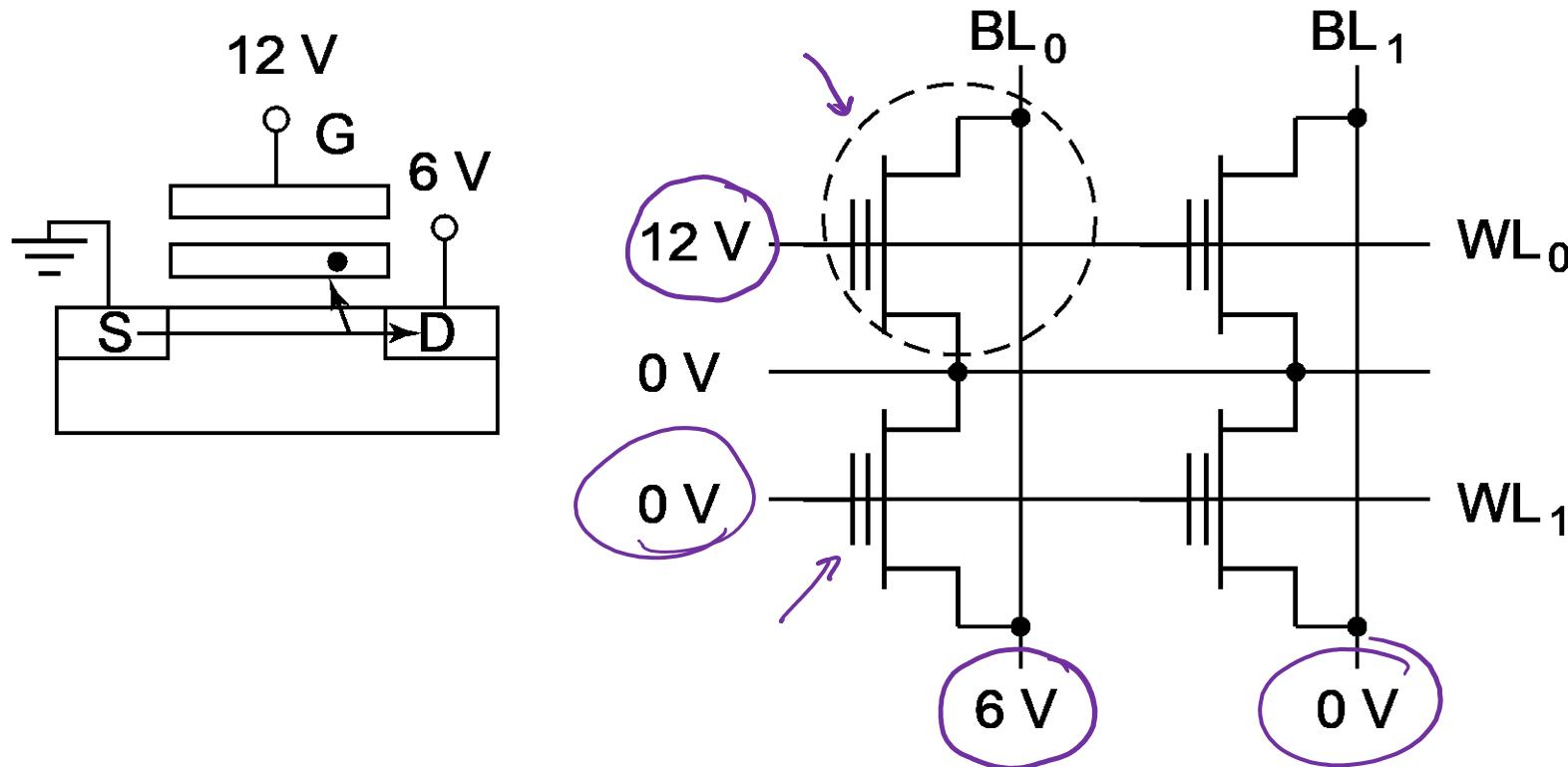
FLASH EEPROM



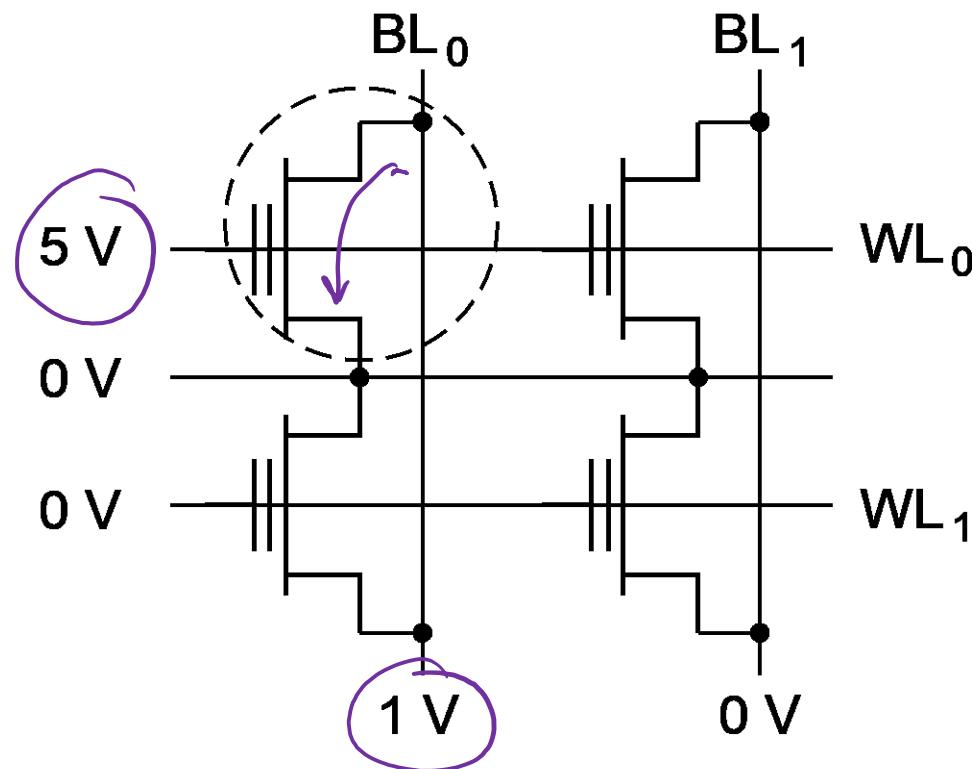
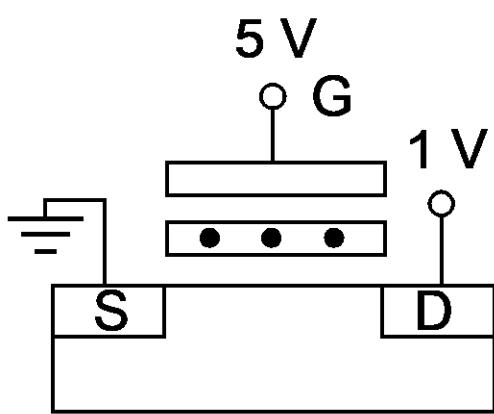
BASIC OPERATIONS IN A NOR FLASH MEMORY: ERASE



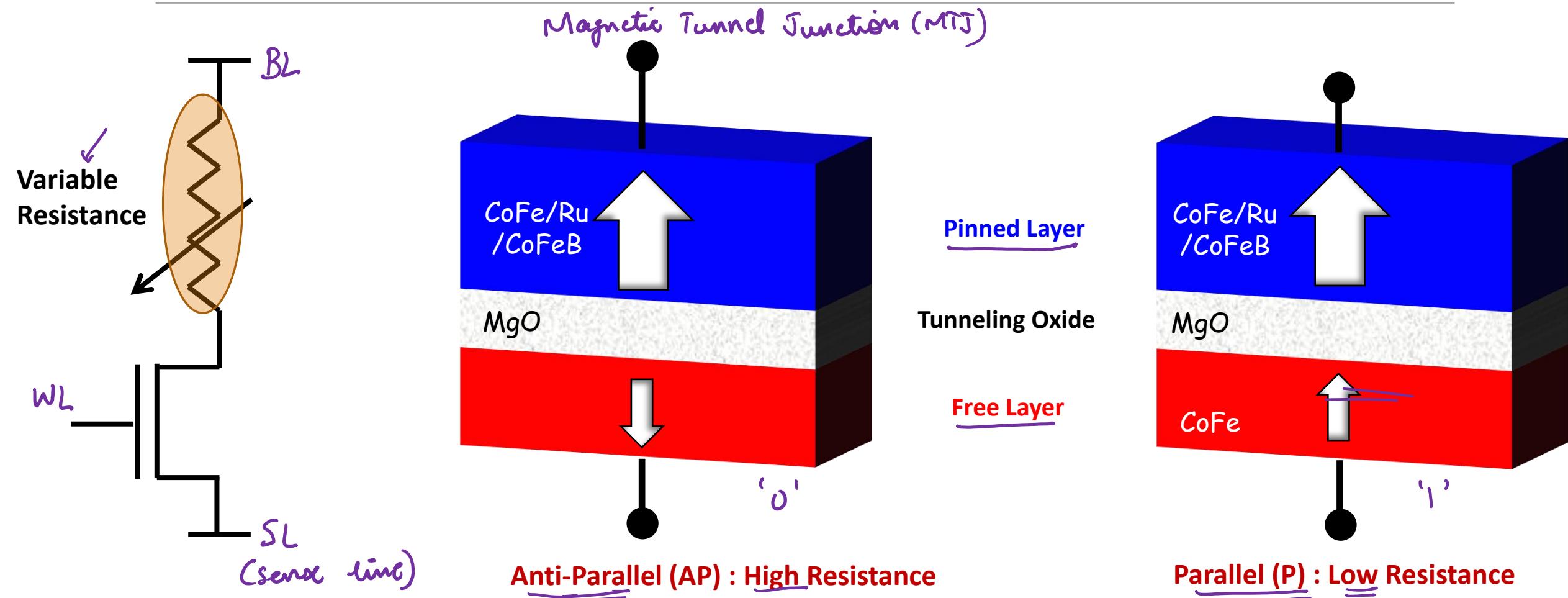
BASIC OPERATIONS IN A NOR FLASH MEMORY: WRITE



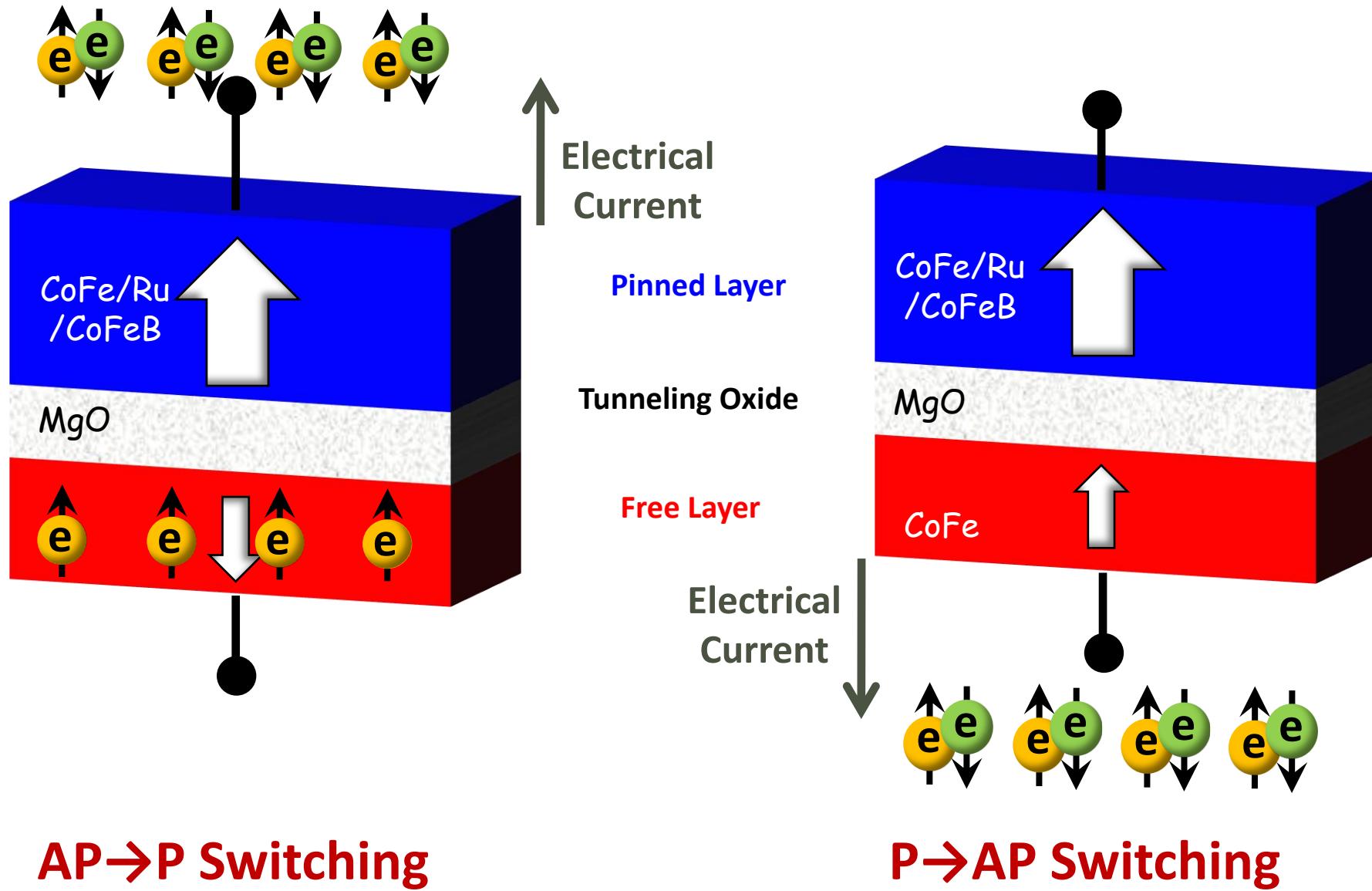
BASIC OPERATIONS IN A NOR FLASH MEMORY: READ



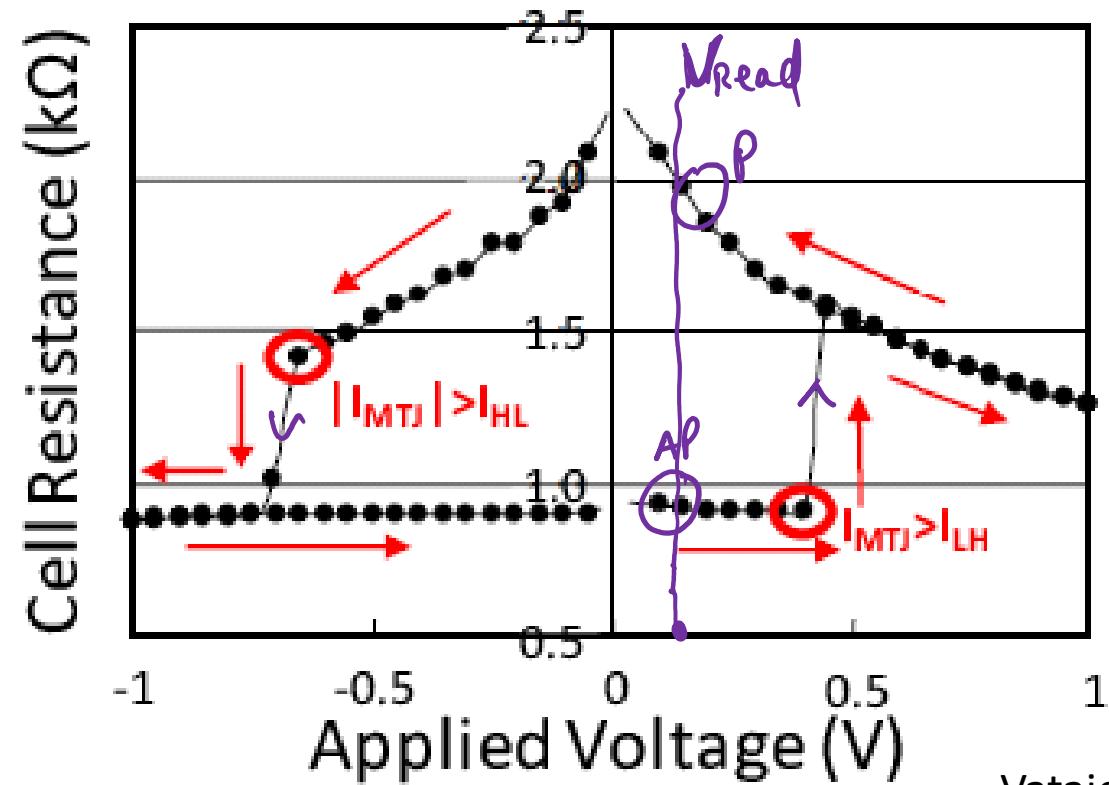
SPIN-BASED MEMORIES: SPIN TRANSFER TORQUE MAGNETIC RAMS (STT MRAMS)



SPIN-BASED MEMORIES: STT MRAMs

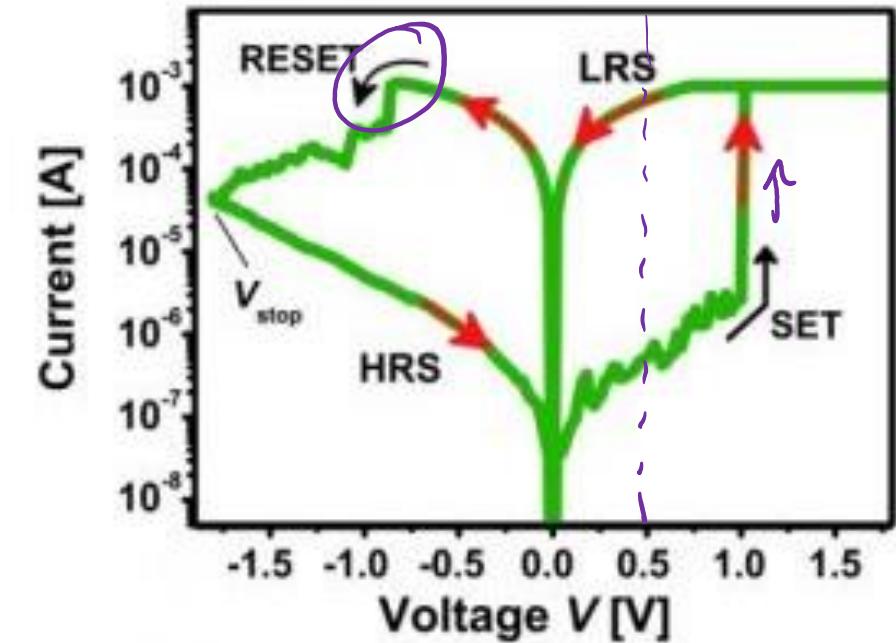
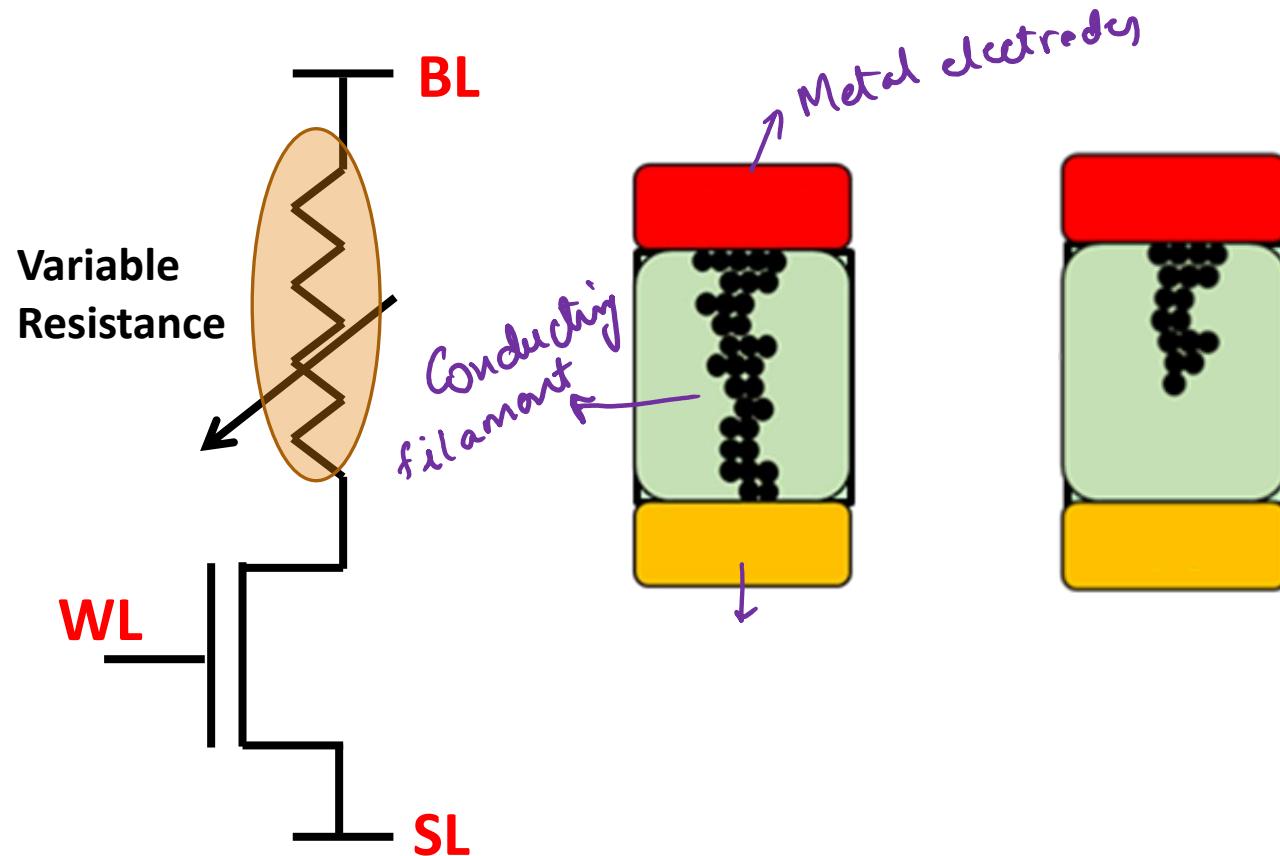


SPIN-BASED MEMORIES: STT MRAMS



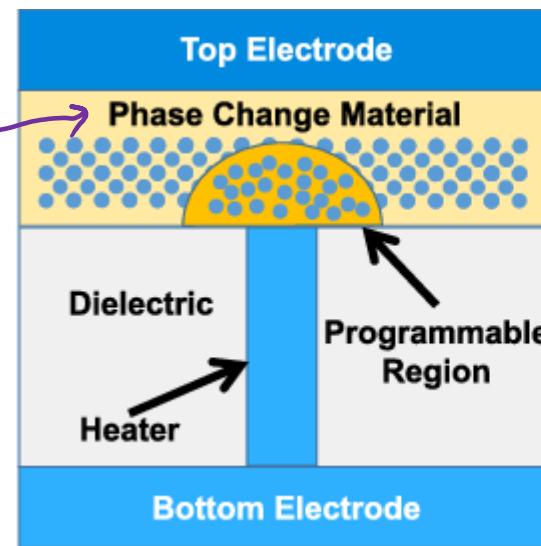
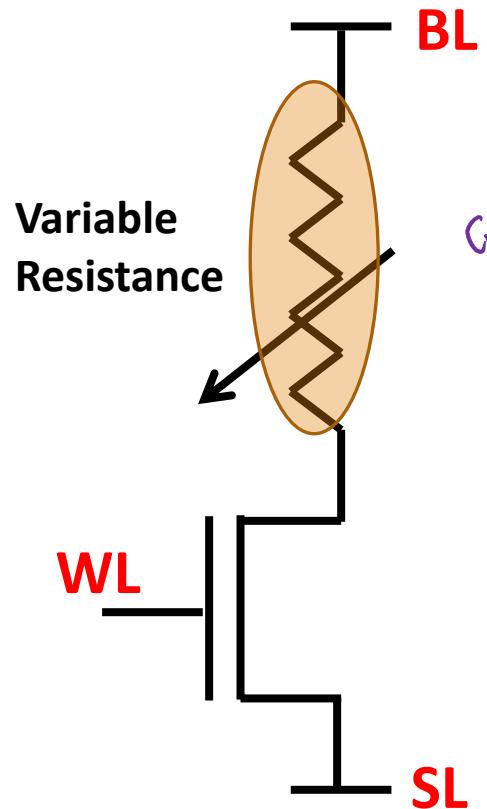
Vatajelu et al, IDT 2014

RESISTIVE MEMORIES (RERAMS)

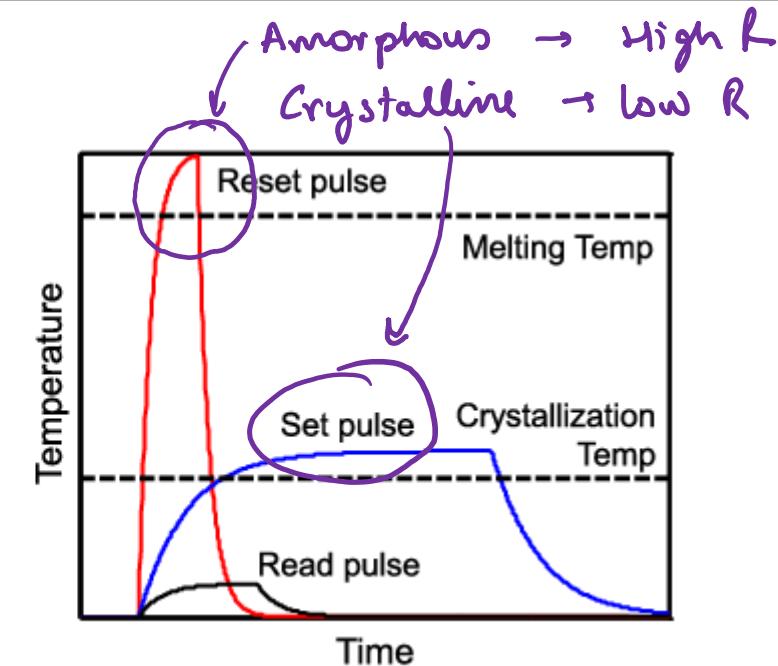


Kim et al, Scientific Reports 2014

PHASE CHANGE MEMORIES

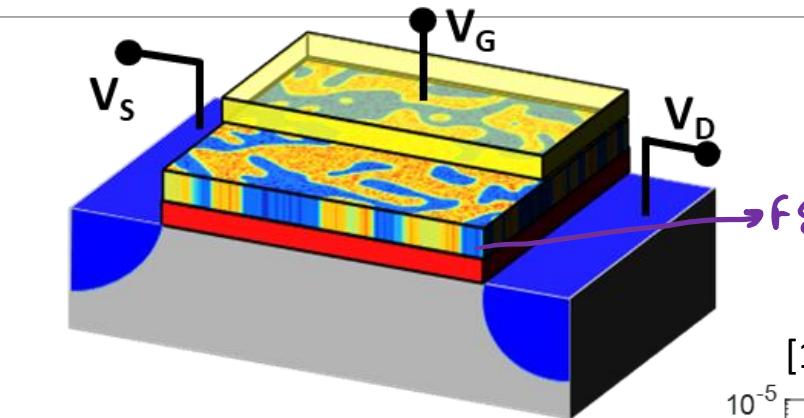
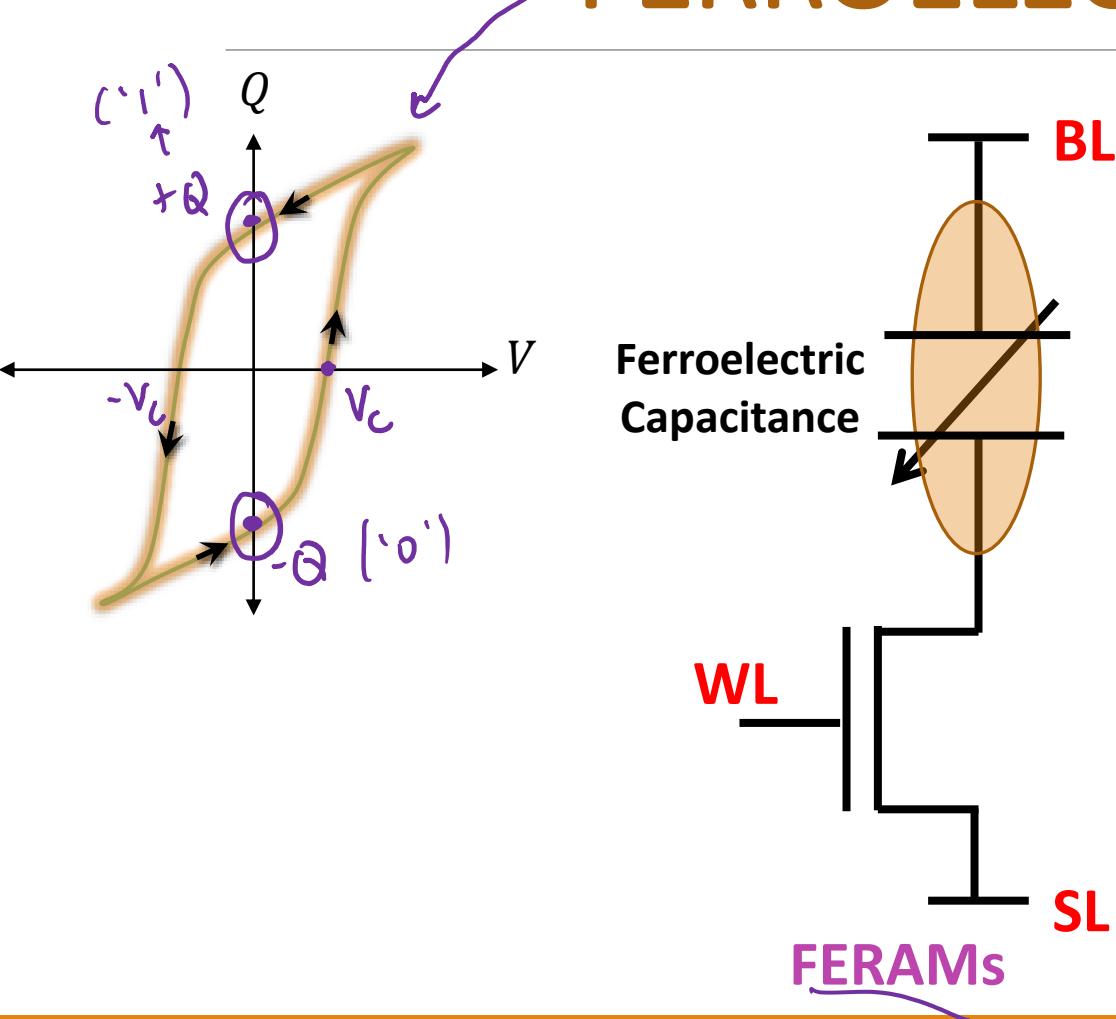


GST

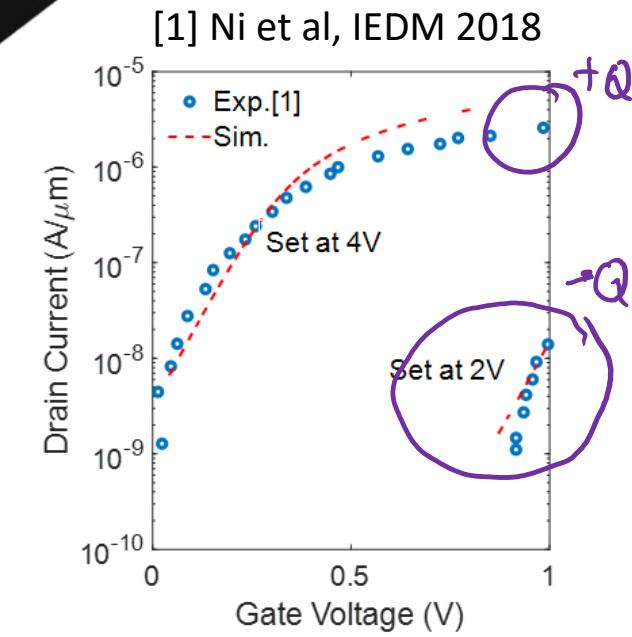
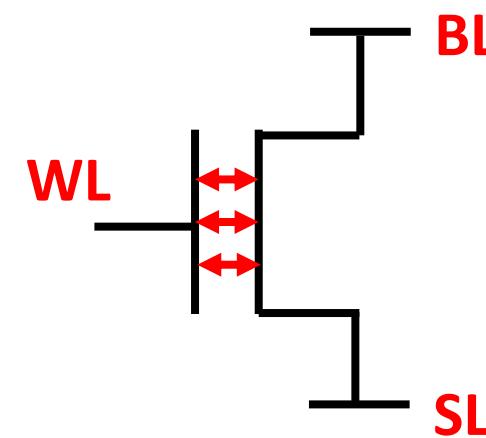


Fong et al, IEEE Transactions on Electron Devices, 2016

FERROELECTRIC MEMORIES



FEFETs



Suggested Reading

Chapter 12

- **12.1**
- **12.2.1**
- **12.2.2**
- **12.2.3**
- **12.3.1**

Questions/Comments - ??