Flash Memory SUMMIT

densbits TECHNOLOGIES

Enabling 3D NAND Devices of the Future

The Necessity for a Memory Modem™ in 3D Memories

Hanan Weingarten

DensBits Technologies

Outline

- 2D Vs 3D NAND
- 3D NAND – How Does It Work
  - p-BiCS example
- 3D NAND (unique) Considerations
  - Inter-Cell / Block Interference
  - TLC on 3D NAND
  - Scaling
- Summary

FlashMemory SUMMIT

2D Vs 3D NAND (1)

densbits TECHNOLOGIES

- 2D NAND limitations below 1xnm:
  - Lithographic limitations:
    - Additional scaling requires quad-patterning, Ultra-violet
  - Reliability limitations (retention, read and program disturbs):
    - Channel isolation deterioration
    - Stored charge with small number of electrons (a few dozens)
    - Increased cell to cell interference
  - Performance
    - Longer programming times
    - Longer read times

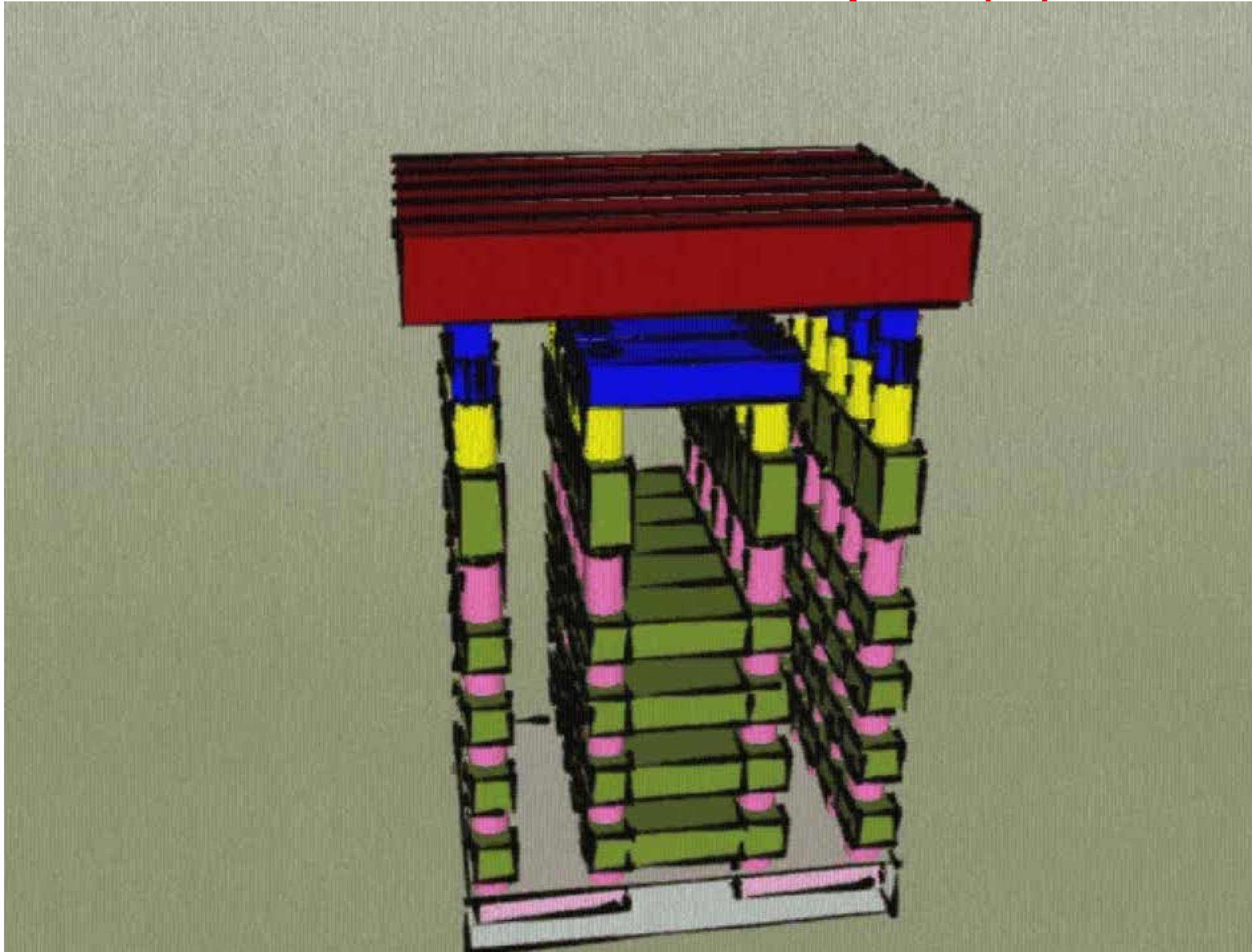
2D Vs 3D NAND (2)

- 3D NAND:
  - Solves many of the issues of 2D NAND
  - Instead of scaling x-y dimensions, scale-up
  - Scale back lithography:
    - Feature size > 3Xnm
    - No double / quad patterning
    - Improved reliability
  - Higher performance (Samsung)
  - Scaling:
    - Through additional number of layers

Flash Memory SUMMIT

3D NAND – How Does It Work: P-BiCS Example (1)

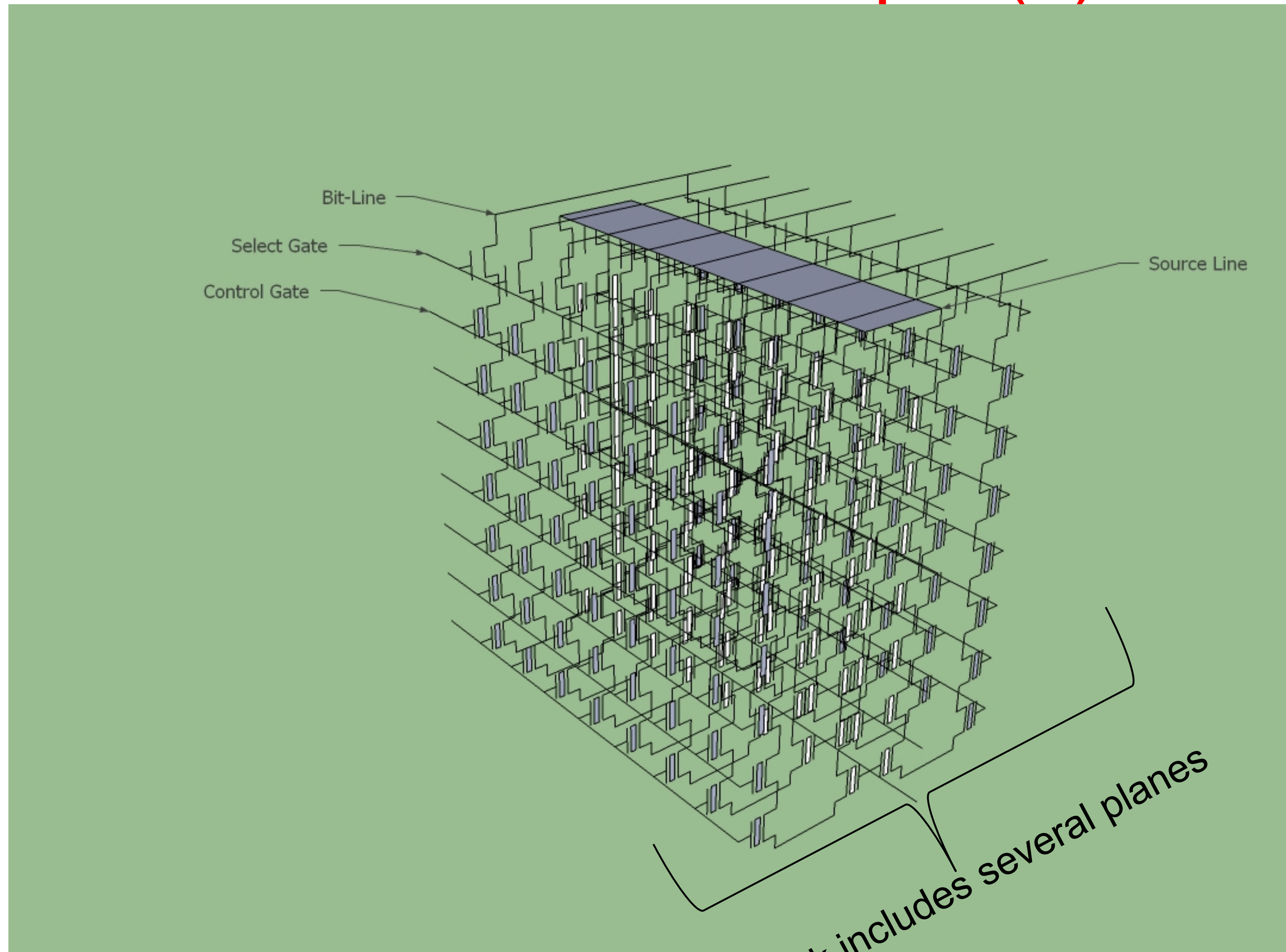
densbits TECHNOLOGIES



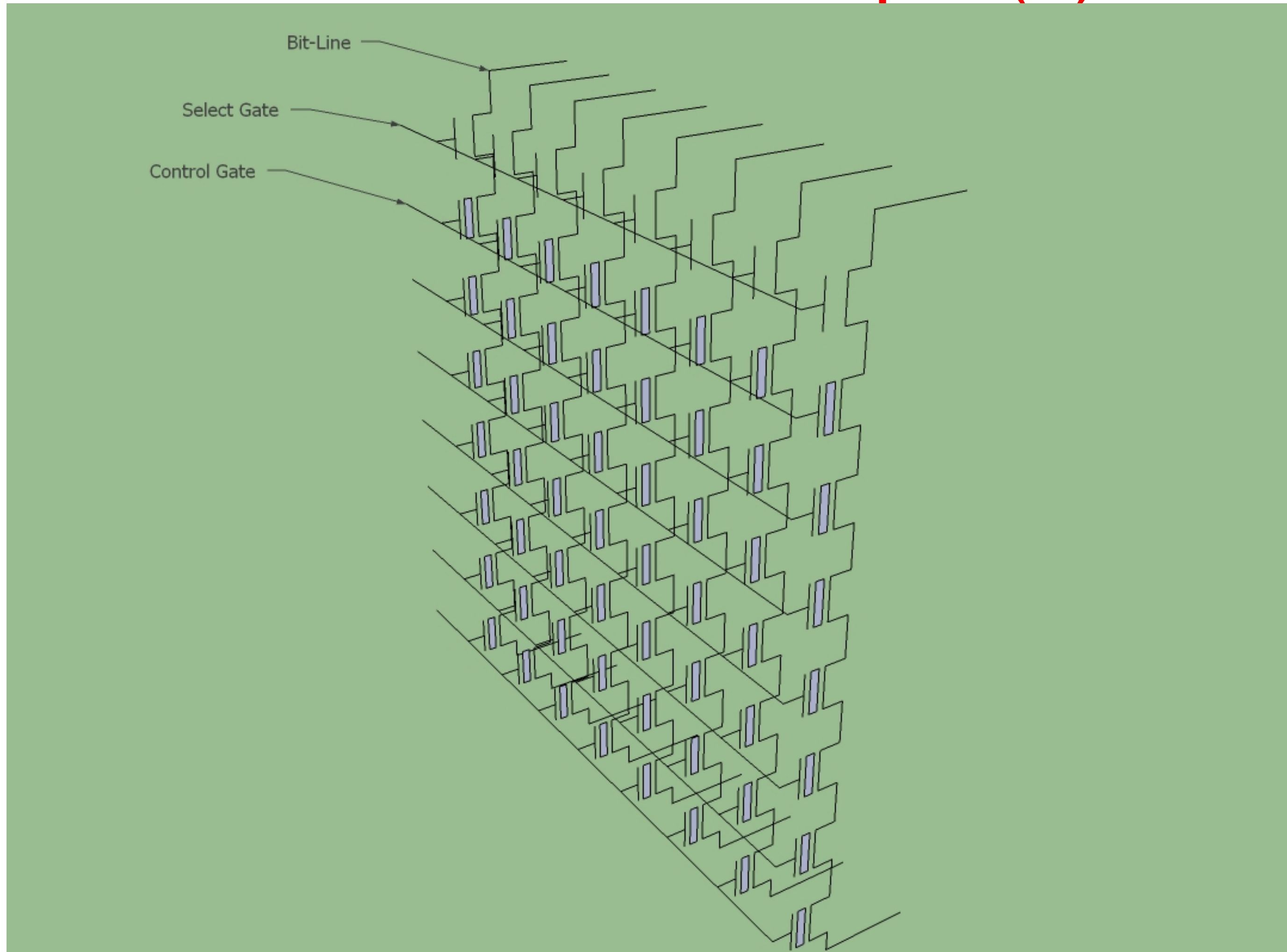
Flash Memory SUMMIT

3D NAND – How Does It Work: P-BiCS Example (2)

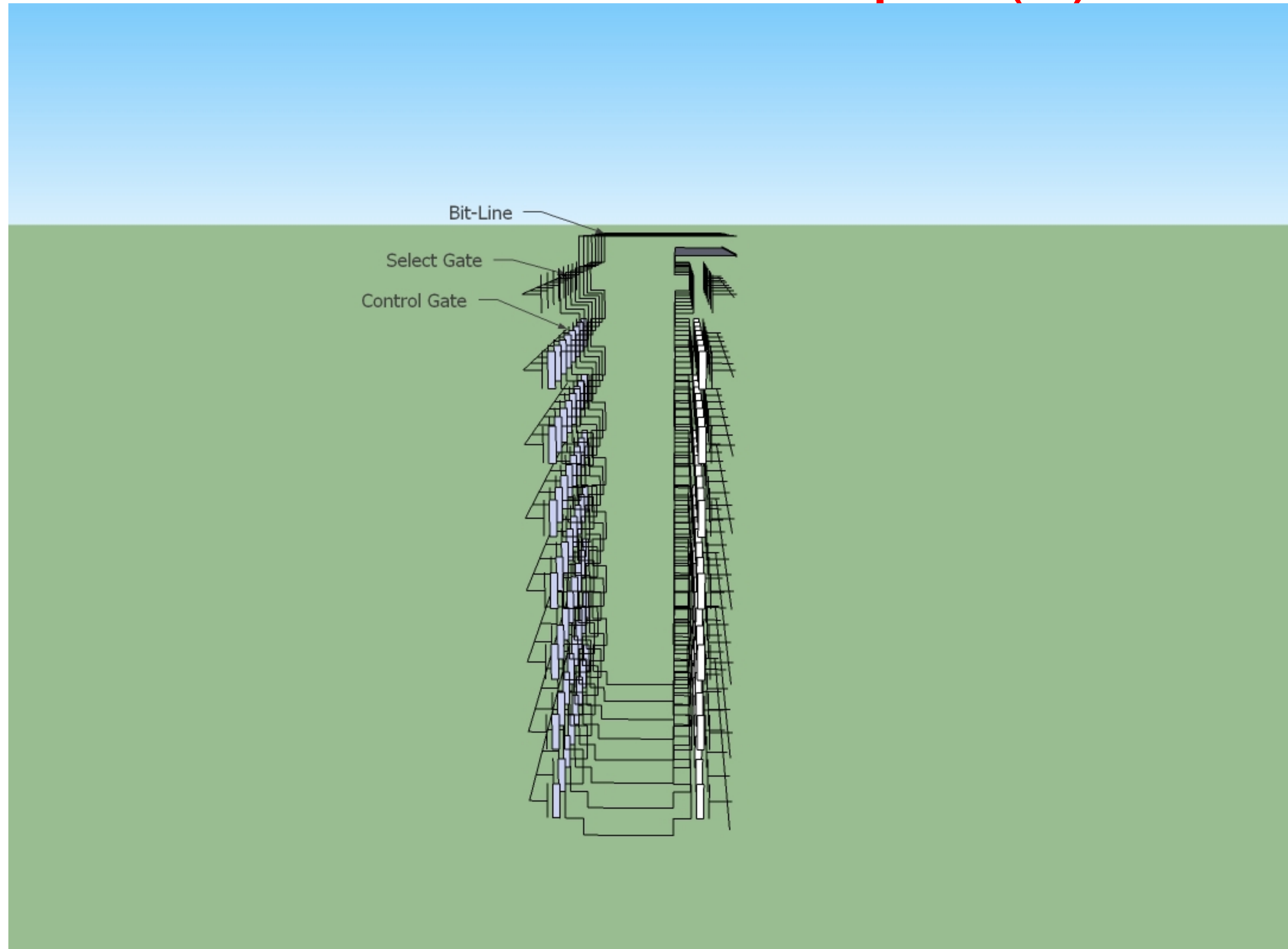
densbits TECHNOLOGIES



3D NAND – How Does It Work: P-BiCS Example (3)

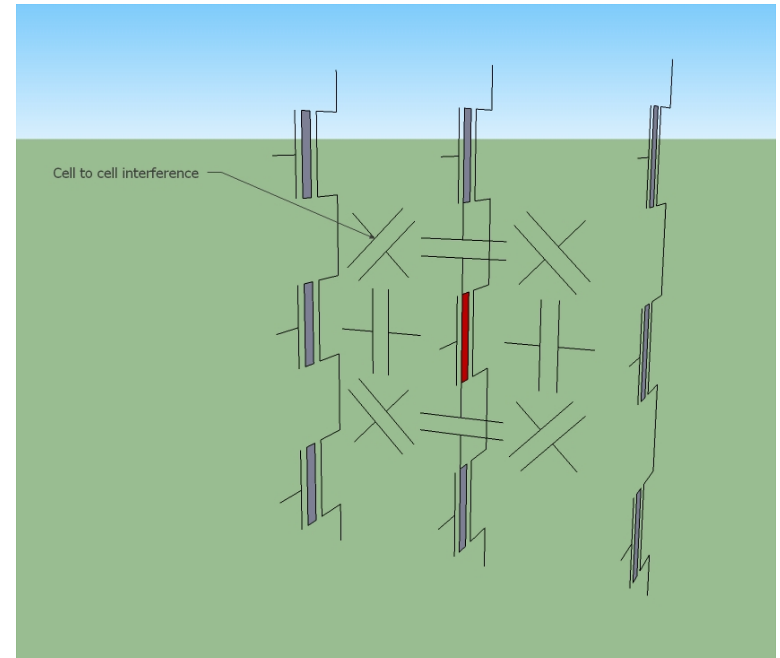


3D NAND – How Does It Work: P-BiCS Example (4)



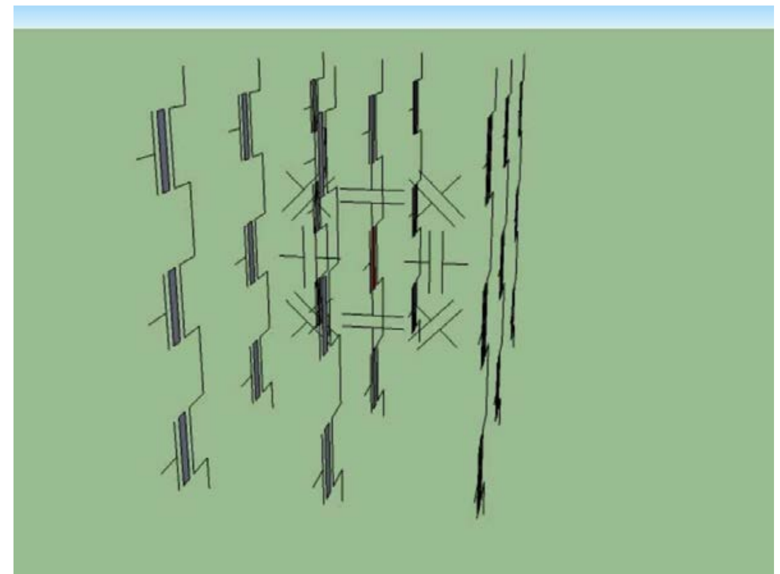
Inter-cell Interference (1)

- 2-D / Planar NAND Interference:
  - Row to Row: x2
  - Bitline to Bitline: x2
  - Diagonal (Bitline – Row): x4
  - Total: x8



Inter-cell Interference (2)

- 3-D NAND Interference:
  - Row to row: x2
  - Bitline to bitline: x2
  - Diagonal (Bitline – row): x4
  - Plane to plane: x2
  - Plane to row: x4
  - Plane to bitline: x4
  - Diagonal: (Plane – bitline – row): x8
  - Total: x26

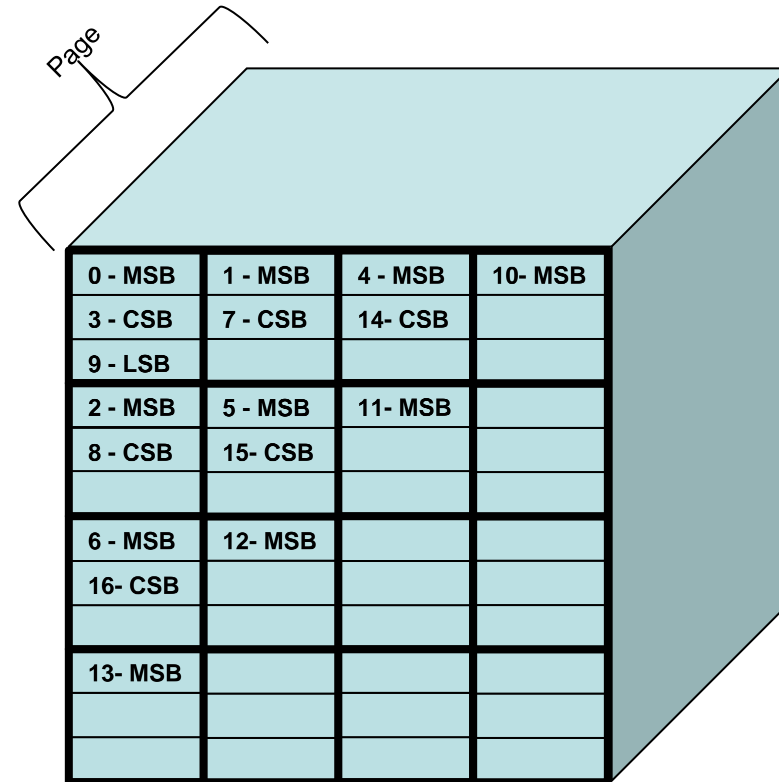


Inter-cell Interference (3)

- Decoupling during read operation:
  - A Memory Modem™ is used to improve read reliability and perform decoupling:
    - Divide interfering cells according to severity
    - Decode attempt with increased number of decoupling reads, according to severity

Inter-cell Interference (4)

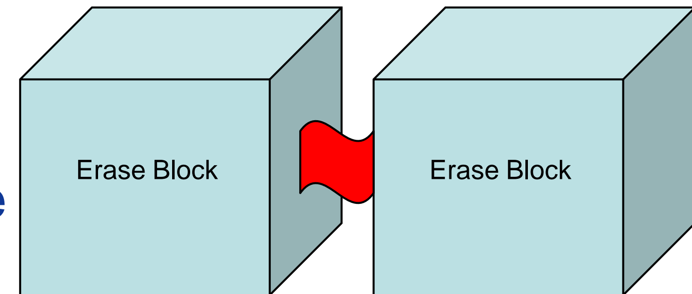
- Limiting interference through page program ordering:
  - A two dimensional programming order
  - More rows are not fully programmed (more pages at risk) at any one time
  - May be an issue during ungraceful power-down
    - **A Memory Modem™ is required to efficiently handle such cases with minimal impact on performance**



TLC Program Ordering Example

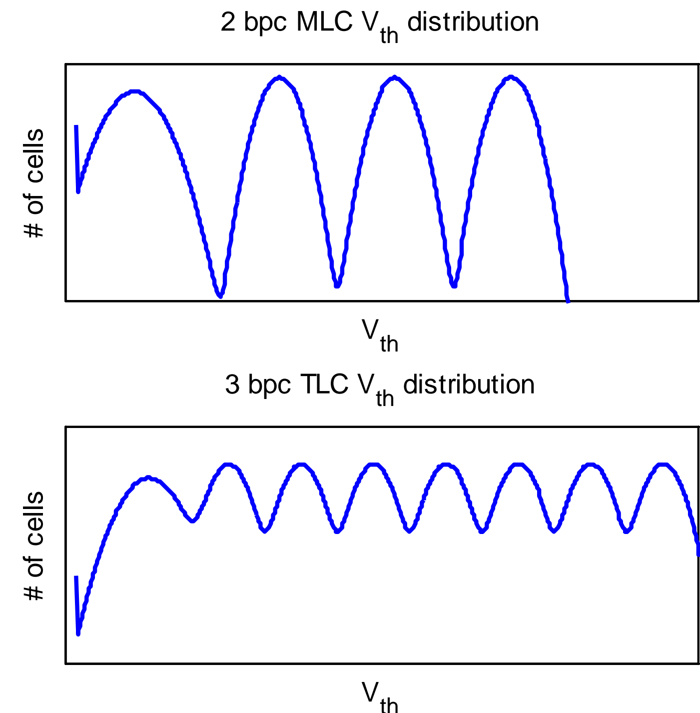
Inter-Block Interference

- With scaling block to block may become an issue:
  - Current management algorithms assume independence, allowing erasing / programming blocks independently of others
  - Solution may include:
    - Management modifications
    - Stronger correction on edge plane
    - Special allocation on edge plane



TLC on 3D

- Scaling back lithography improves MLC reliability
- TLC deteriorates reliability and requires a Memory Modem™
- With Memory Modem™ TLC performance may improve without impacting target reliability
- TLC is key for scaling in 3D (next slides)



Scaling (1)

- Scaling is required for reducing manufacturing costs
- Scaling can be performed in several ways:
  - Scaling through additional layers:
    - Samsung 1st gen 24, 2nd gen 32, …
    - Limits:
      - Diminishing returns with number of layers
      - Tougher requirements from manufacturing tools
      - Lower reliability: read disturbs, programming disturbs

Scaling (2)

- Scaling through lithography:
  - Current 3D manufacturing is planned with no double patterning >38nm
- Scaling through more levels:
  - MLC -> TLC: cost effective
  - Requires Memory Modem™ technology (ECC, DSP, low level memory management):
    - **Increasing number of bits -> Impact on reliability -> Need state of art decoder and signal processing**

Memory Modem™

- ECC:
  - Superior hard and soft decoding
  - High throughput
  - Low power / gate-count
  - Highly configurable
- DSP:
  - Blind, No overheads
  - Negligible impact on performance
- Low level memory management:
  - Configurable and optimal data allocation
  - Variable data allocation

Summary

- As scaling continues in 3D, reliability will be affected
- TLC is as cost effective in 3D as it is in 2D (probably even more)
- Memory Modem™ technology is required for 3D NAND scaling