Magnetic Storage to DNA

Bridging Signal Processing from the Magnetic to the DNA Domain

Austin Striegel, Iouri Oboukhov, Niranjay Ravindran, Rick Galbraith – RW Channel Architecture
6/19/2025

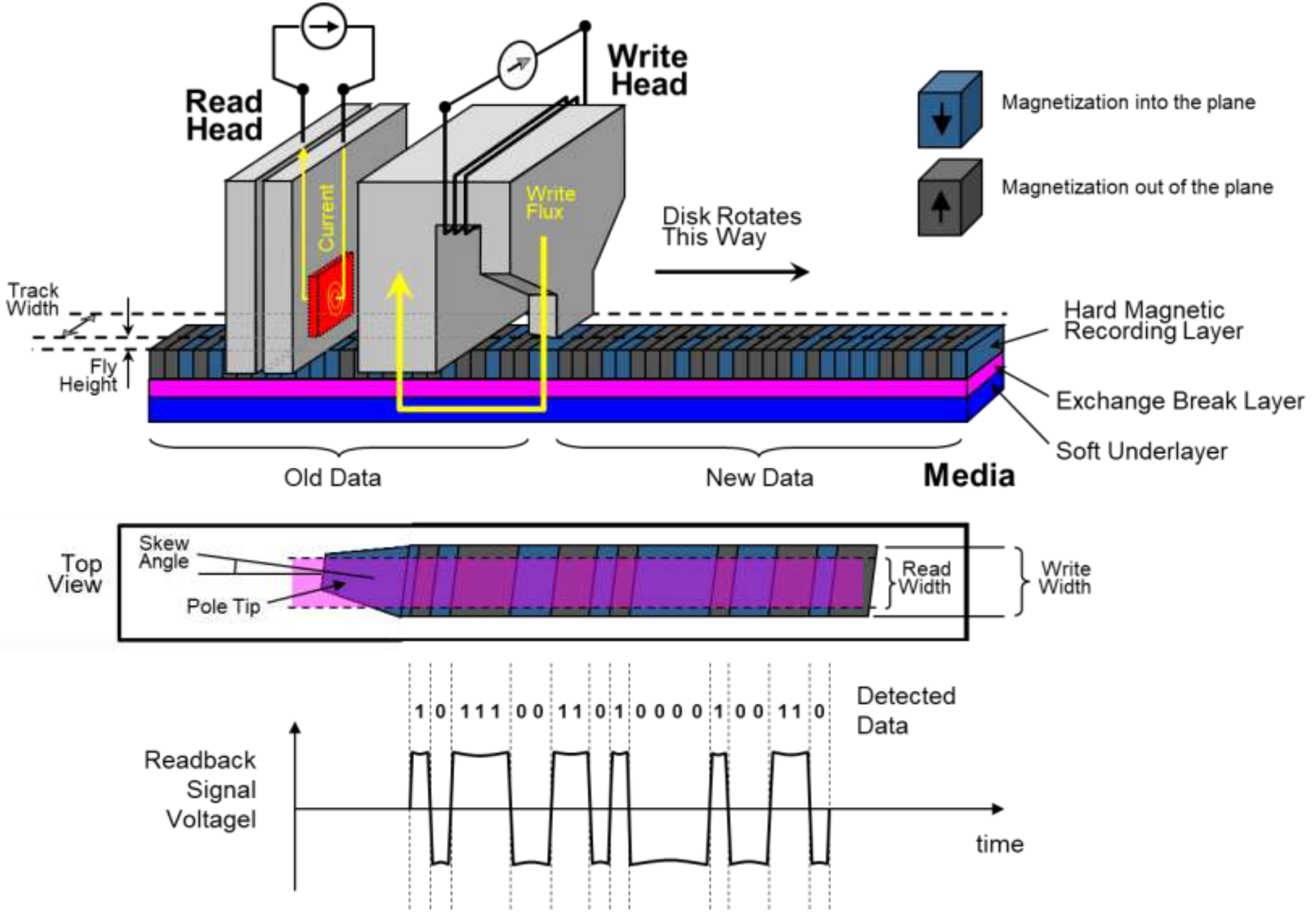
Agenda

- Overview of Hard-Disk Drive (HDD) Magnetic Storage
- Soft-Decoded Error Correction Capability
  - Hard Disk Drives
  - Example DNA Storage System
- Mutual Information (MI) & Error Correction
  - Overview
  - Application to DNA Decoding System
- Summary / Wrap-up

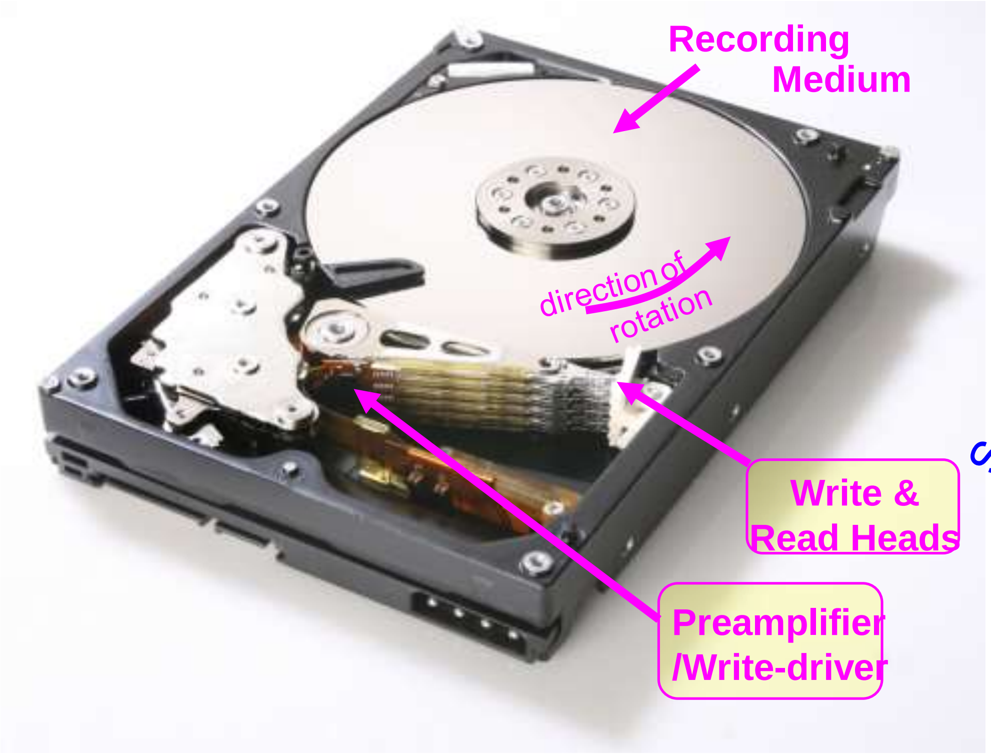
Recording Mechanism

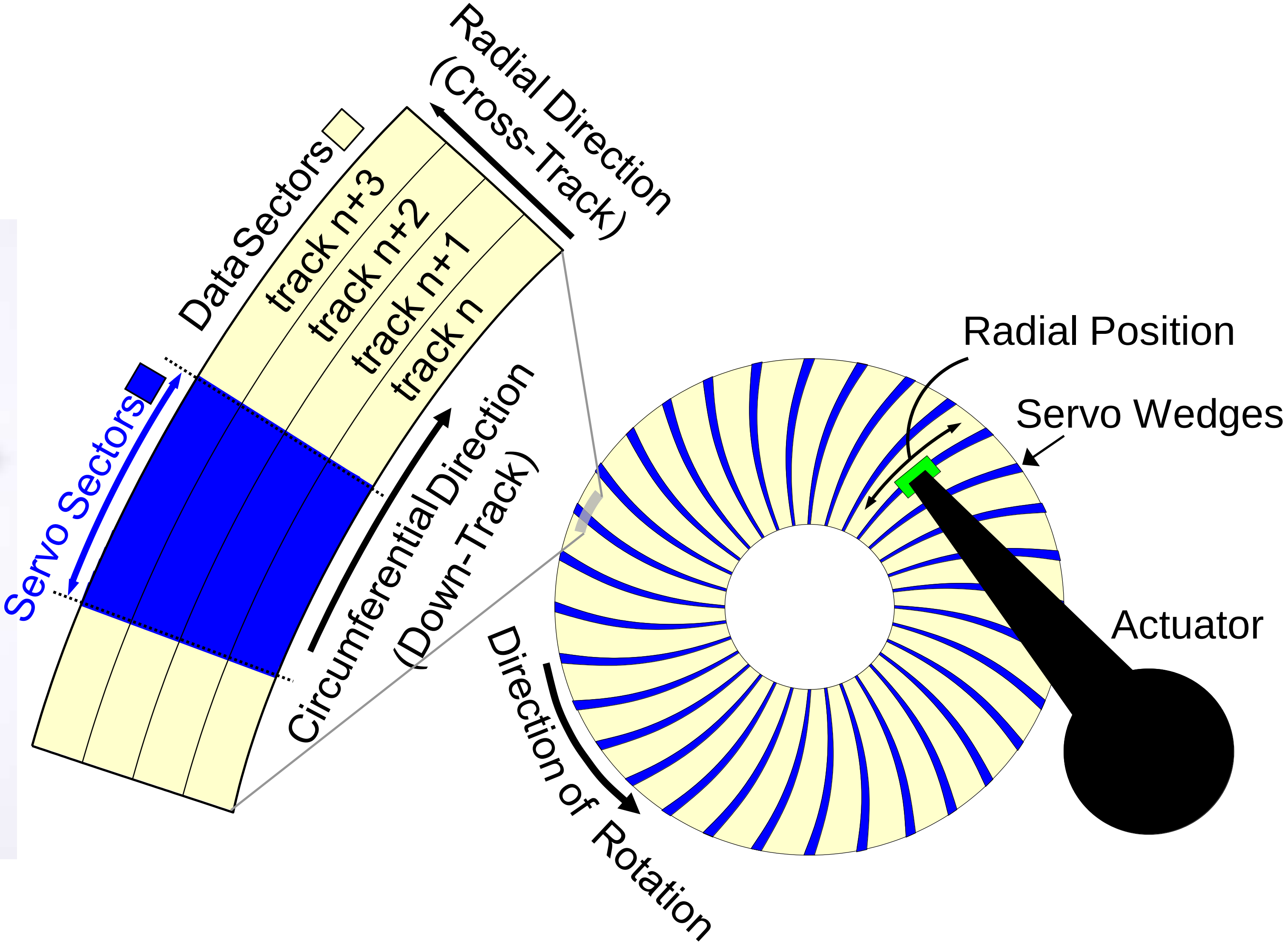
Write Head Magnetizes grains on media to align in up or down direction

Read head senses direction of magnetization



Hard Disk Drive & Drive Format



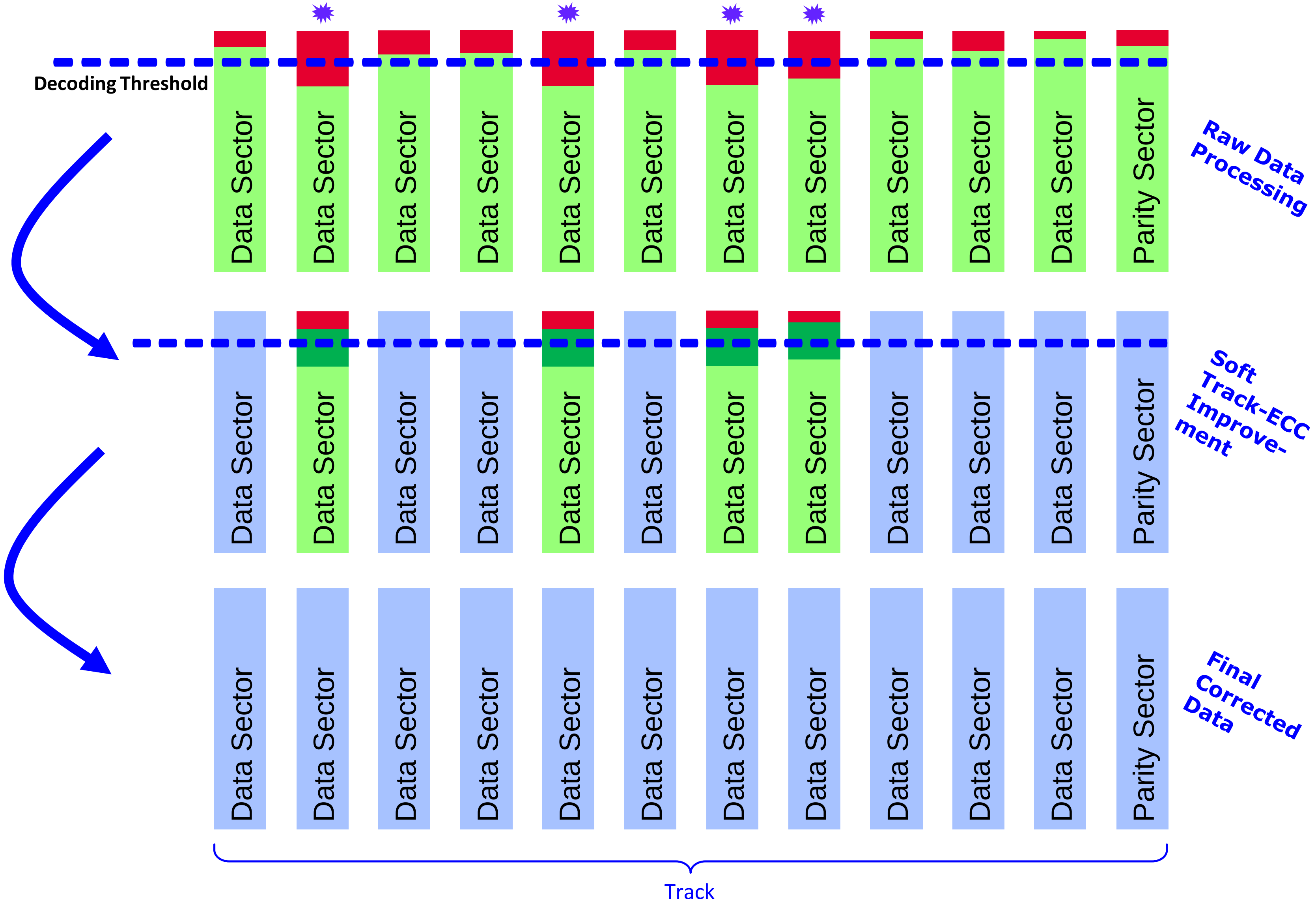


HDD Soft-Decoded Track ECC

Sectors above the decoding threshold are fully corrected by Sector level LDPC

Sectors below the decoding threshold are improved by soft-decoded track level LDPC

Sectors improved by soft-decoded track level LDPC are now decodable by Sector Level LDPC

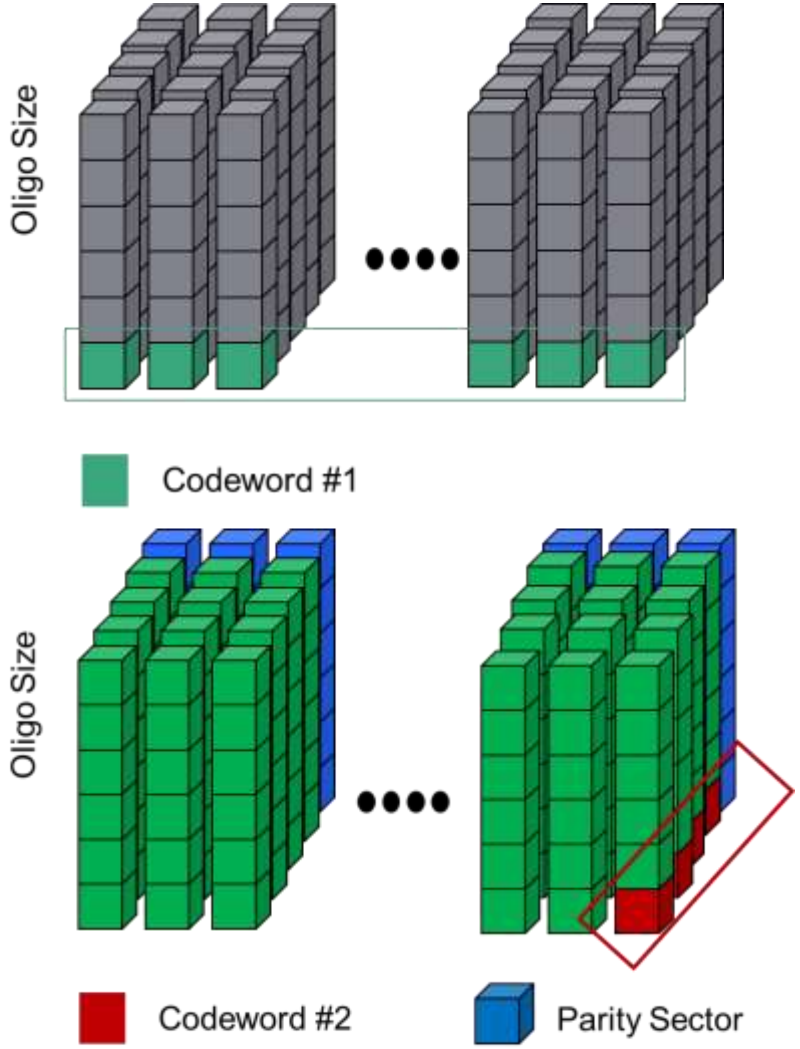


DNA Storage

Example Design Strategy - Encoding

- In a pool of oligos, codewords can be constructed of a target size

- Codewords combined into a larger block (Superblock) by Parity

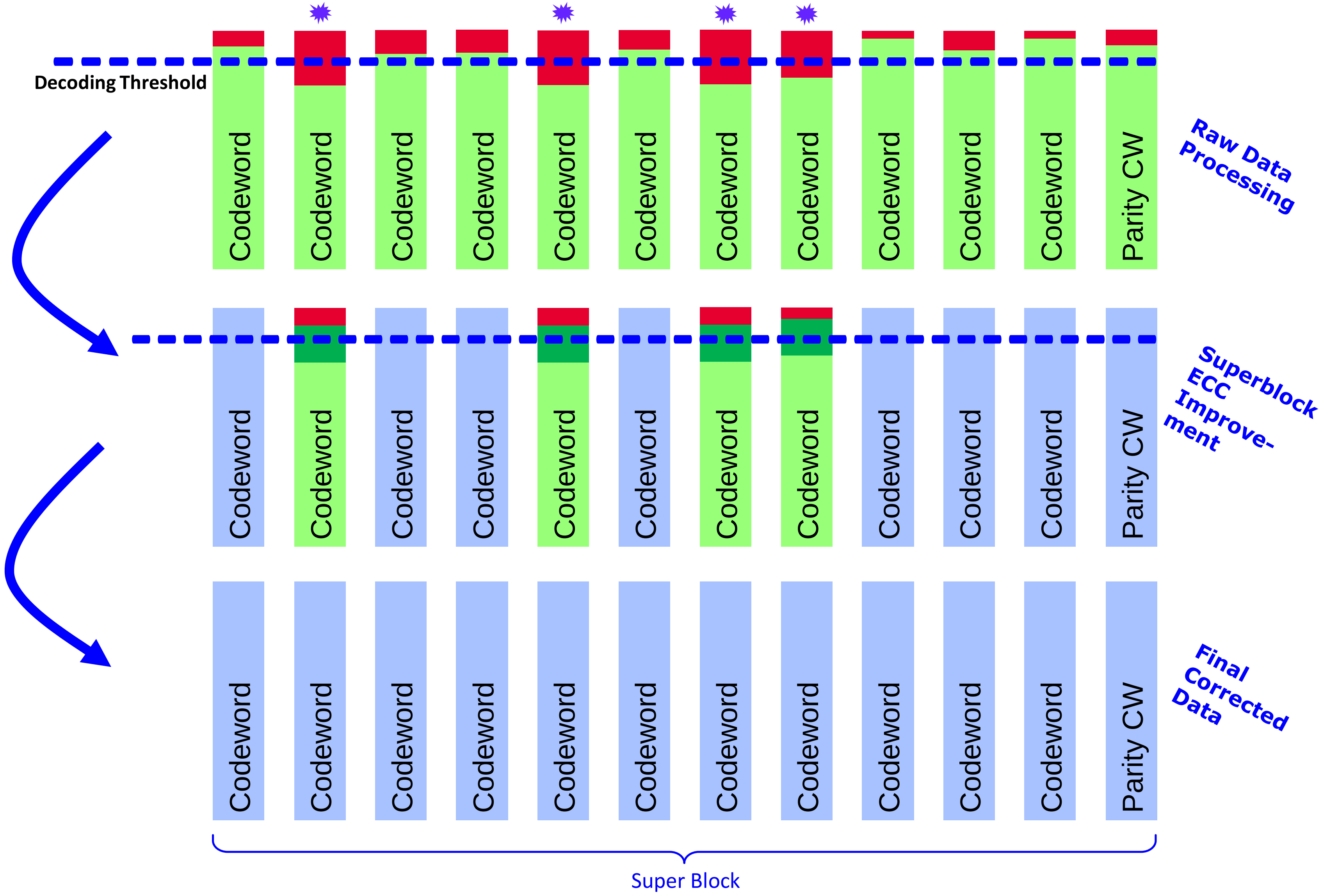


DNA Soft-Decoded Parity CW ECC

Codewords above the decoding threshold are fully corrected by CW level LDPC

Codewords below the decoding threshold are improved by soft-decoded Superblock LDPC

Codewords improved by soft-decoded Superblock LDPC are now decodable by CW LDPC

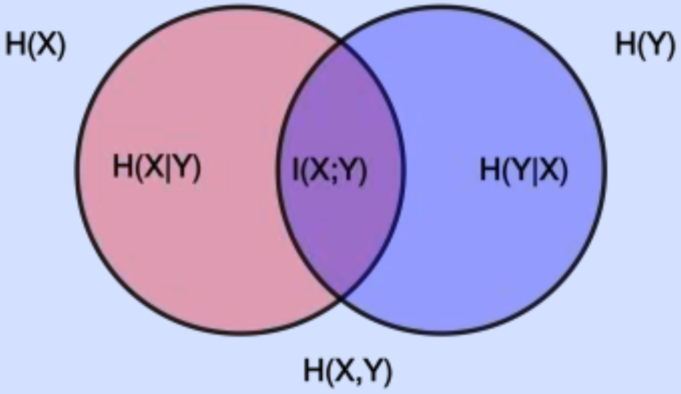


Mutual Information (MI)

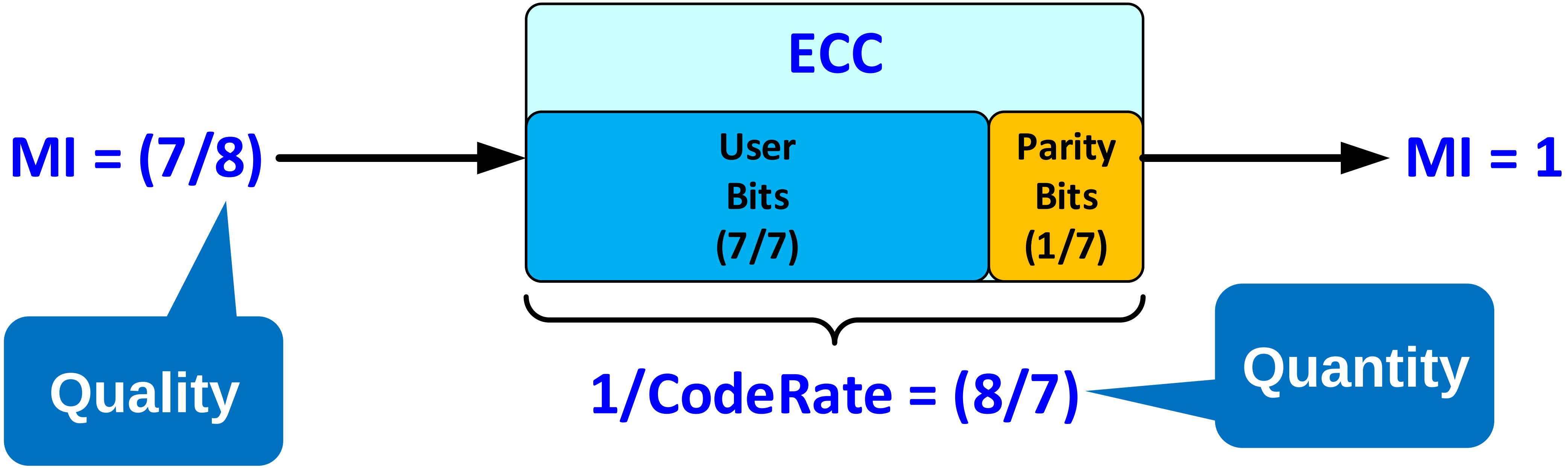
Mutual Information

- Starring role in "Correlation for the 21st Century."
- Superior to Pearson's Correlation
- "Mutual' Relationship between Written & Read
- "Information" as the resolution of uncertainty
- Information Rate in Bits or Nats or Dits

$$I(X,Y) = \sum_{X}\sum_{Y} P(x,y) log \frac{P(x,y)}{P(x)P(y)}$$



MI & Error Correction Code (ECC)



Success → MI/Coderate >= 1
Success → MI >= Coderate

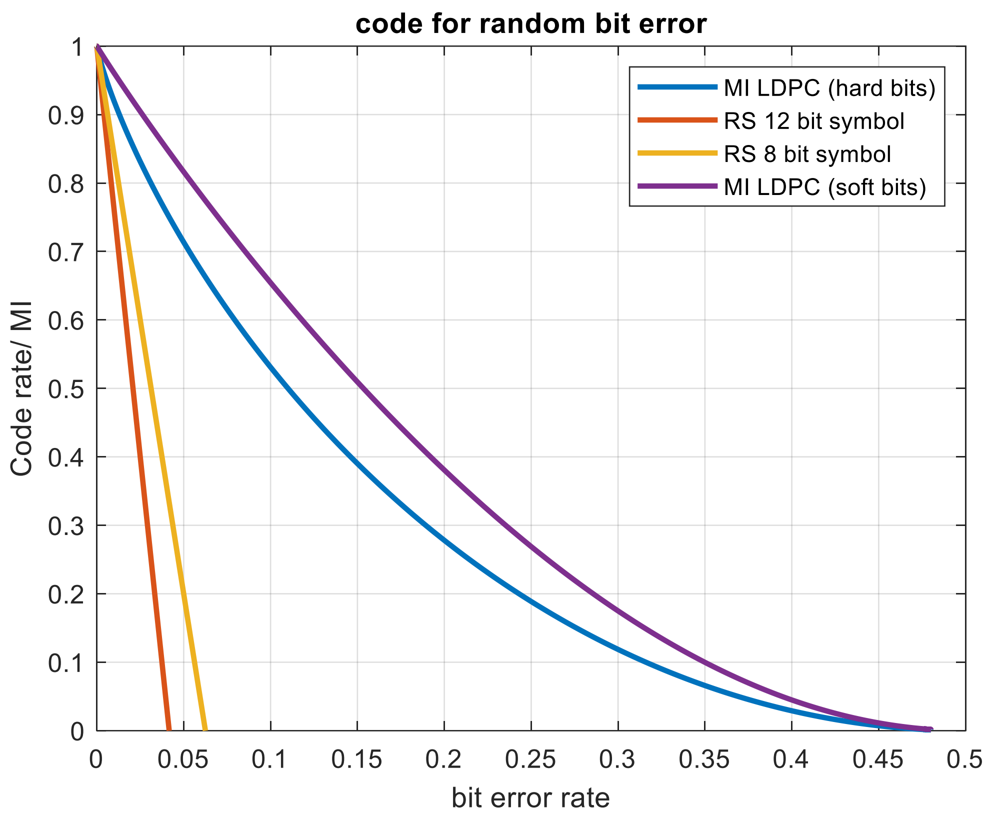
ECC Viewed as Information Passing & Recovery

Usefulness of MI in DNA Storage

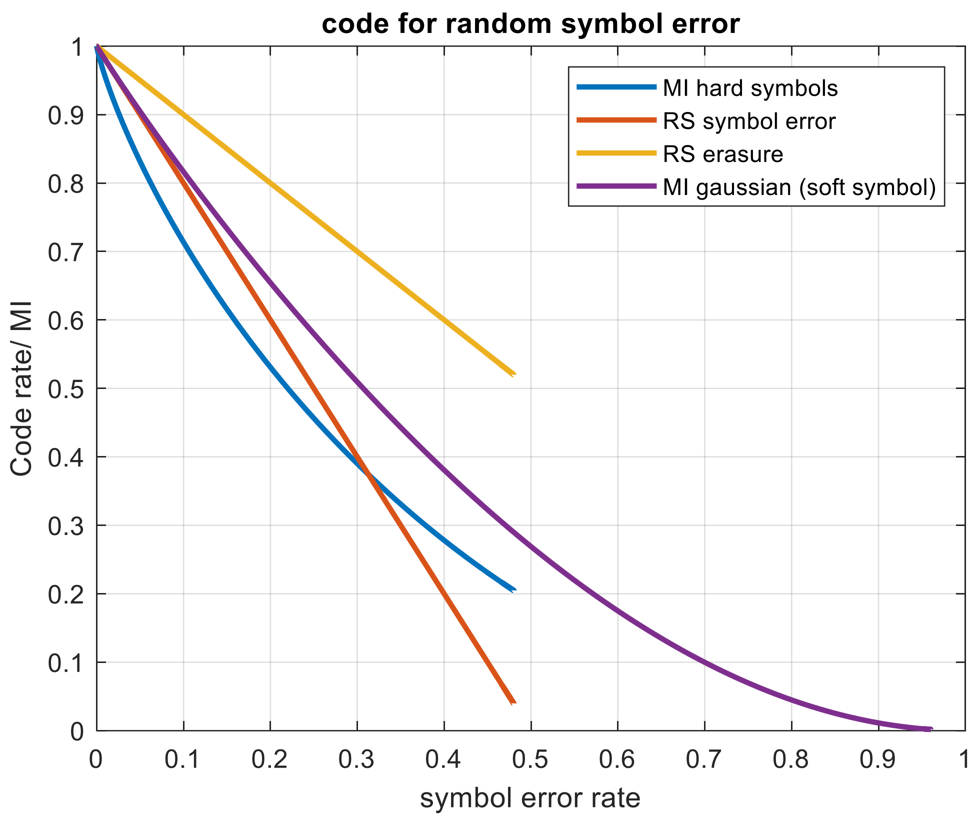
- Define a DNA Storage System
  - Mutation Errors - α
  - Erasure Errors - β
  - Code Rate – R
- Reed-Solomon Decoding (Hard)
  - $R = 1 - 2\alpha - \beta$
- LDPC (Hard)
  - $R = (1-\beta)[1 + (1-\alpha)log_2(1-\alpha) + \alpha log_2(\alpha)]$
- LDPC (Soft)
  - Assume Gaussian Probability of Correct Answer

MI Analysis – DNA

Codec efficiency for Random Bit Errors



Codec Efficiency for Symbol (8 bit) errors



Summary

- Techniques utilized by the Magnetic Storage Industry can be readily applied to potential DNA Storage Coding
  - Soft-Decoded Superblock Parity Correction
- Mutual information provides
  - Information relationship between written and read
  - Valuable metric to assess coding requirements

TM