2005 Viterbi Conference

Applications of the Viterbi Algorithm in Data Storage Technology

Paul H. Siegel

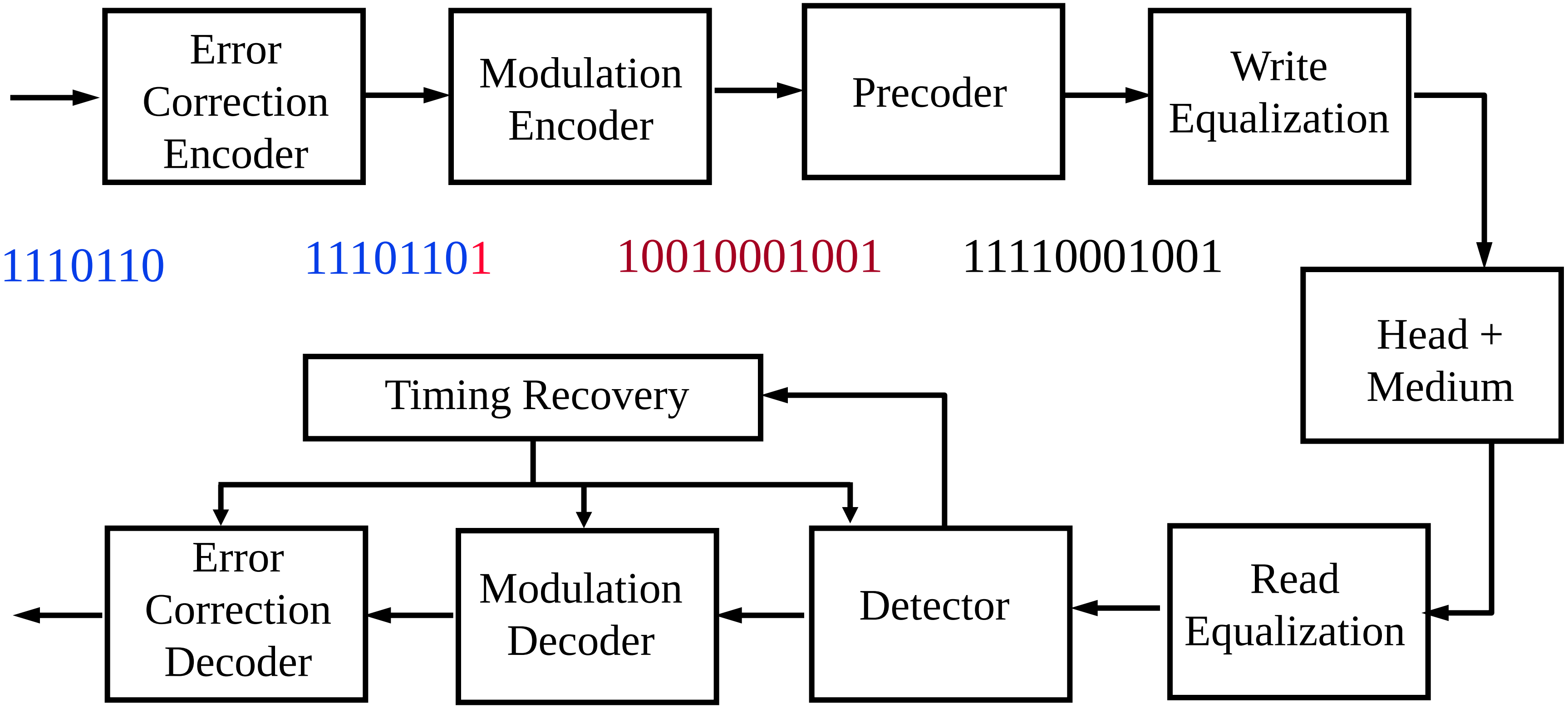
Director, CMRR
Electrical and Computer Engineering
University of California, San Diego

Center for Magnetic Recording Research

Outline

- Data storage trends
- Recording channel technology
  - PRML
  - Coded PRML
  - Turbo equalization
- Channel capacity
- Concluding remarks

Digital Recording Channel



Error Correction Encoder

Modulation Encoder

Precoder

Write Equalization

Head + Medium

1110110

11101101

10010001001

11110001001

Timing Recovery

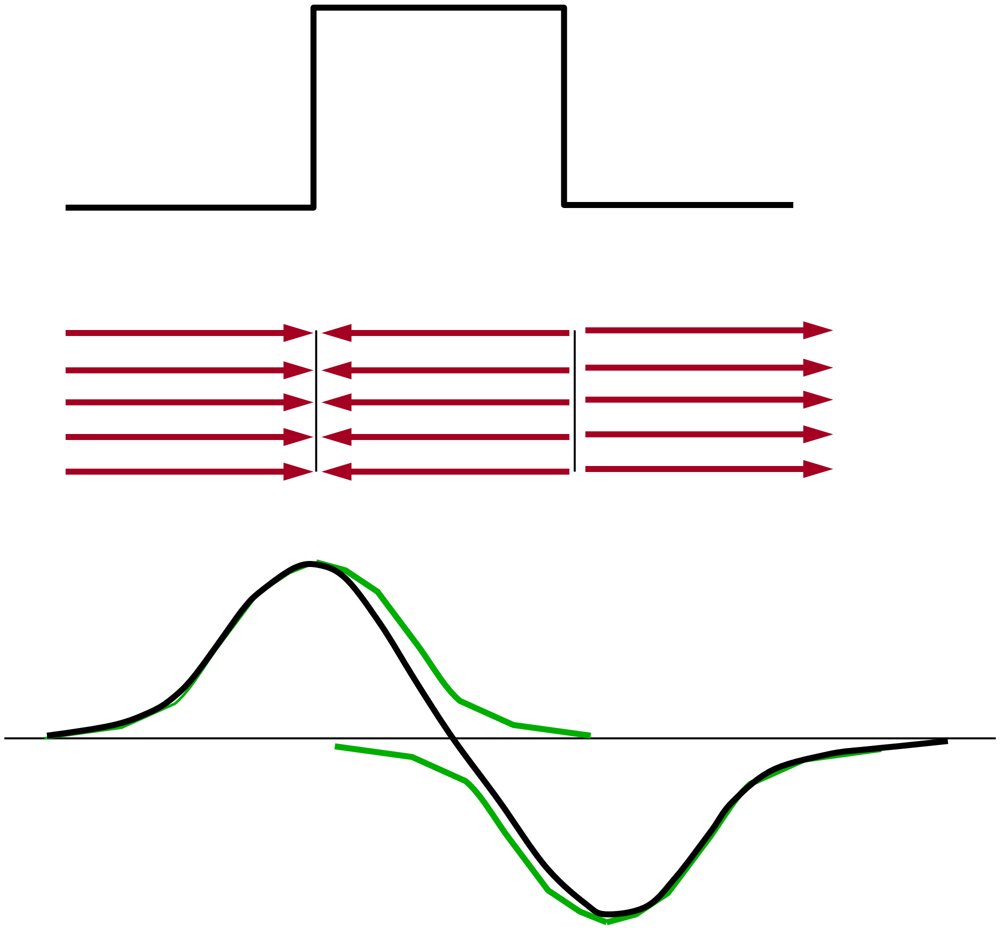
Error Correction Decoder

Modulation Decoder

Detector

Read Equalization

Magnetic Recording Process



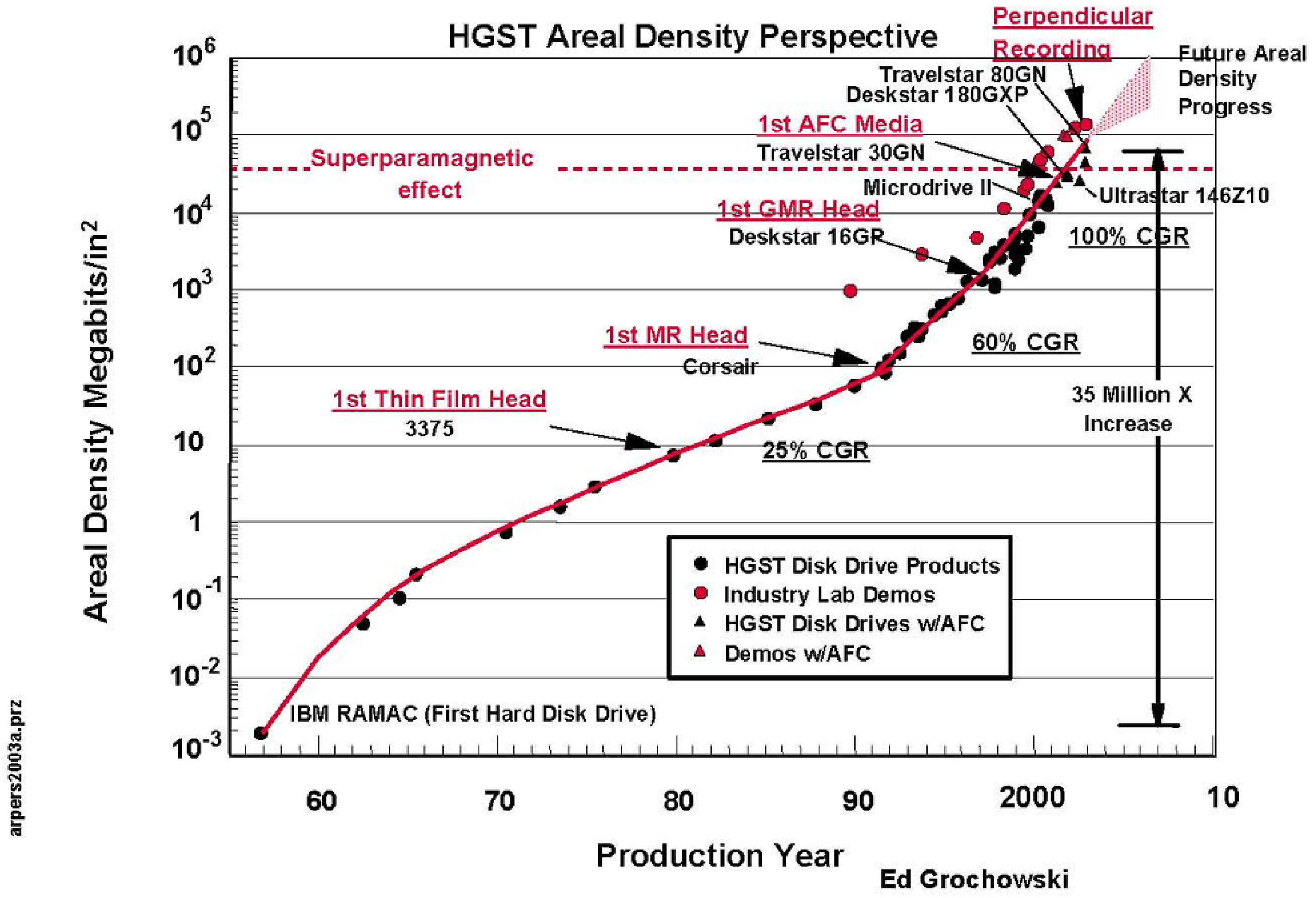
Input signal

Magnetized
Medium

Readback
Signal

Areal Density Progress

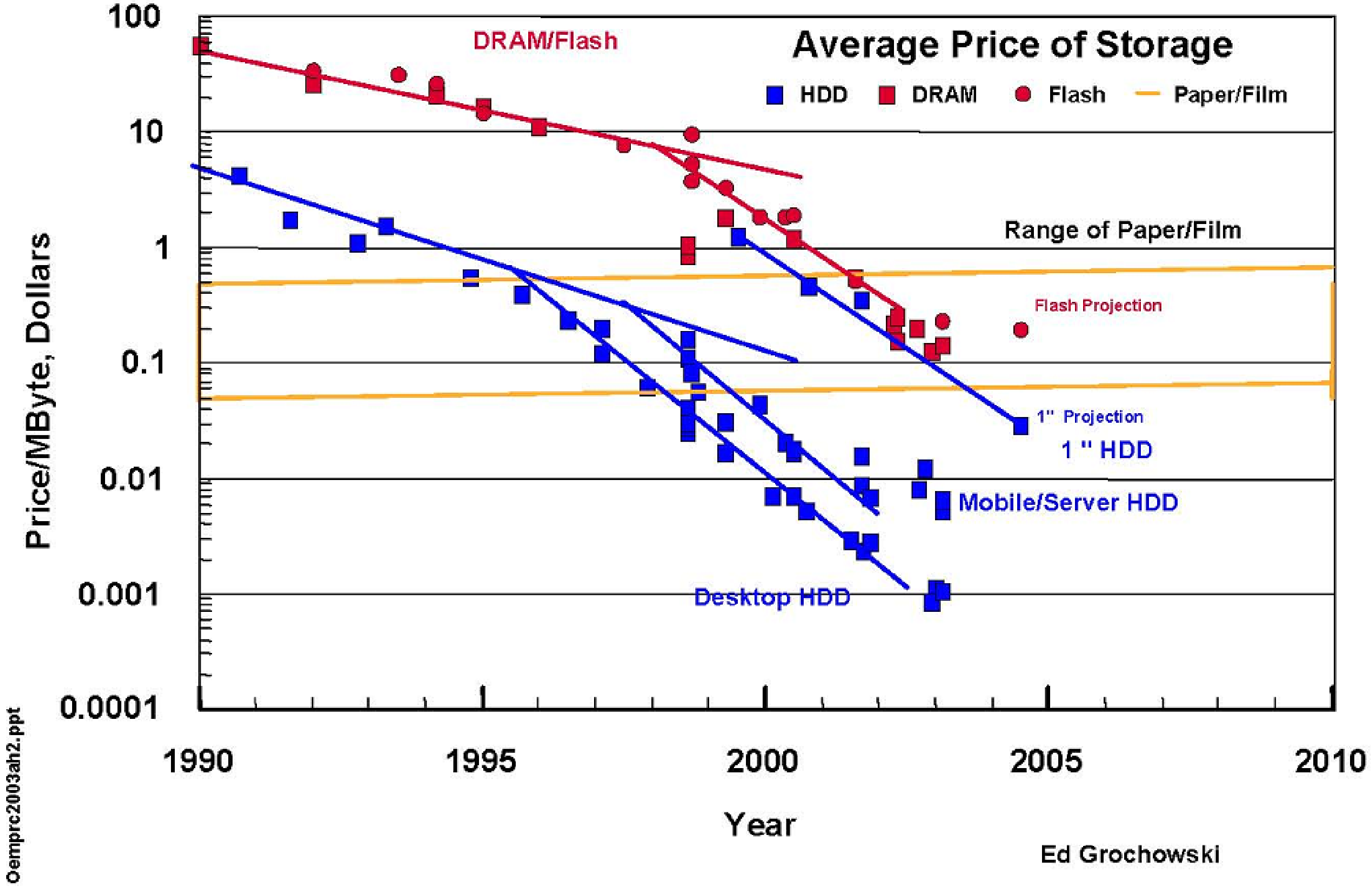
HITACHI
Inspire the Next



2005 Viterbi Conference

Average Price of Storage

HITACHI
Inspire the Next



A Disk Drive (and VA) in Every Pocket

10,000 songs

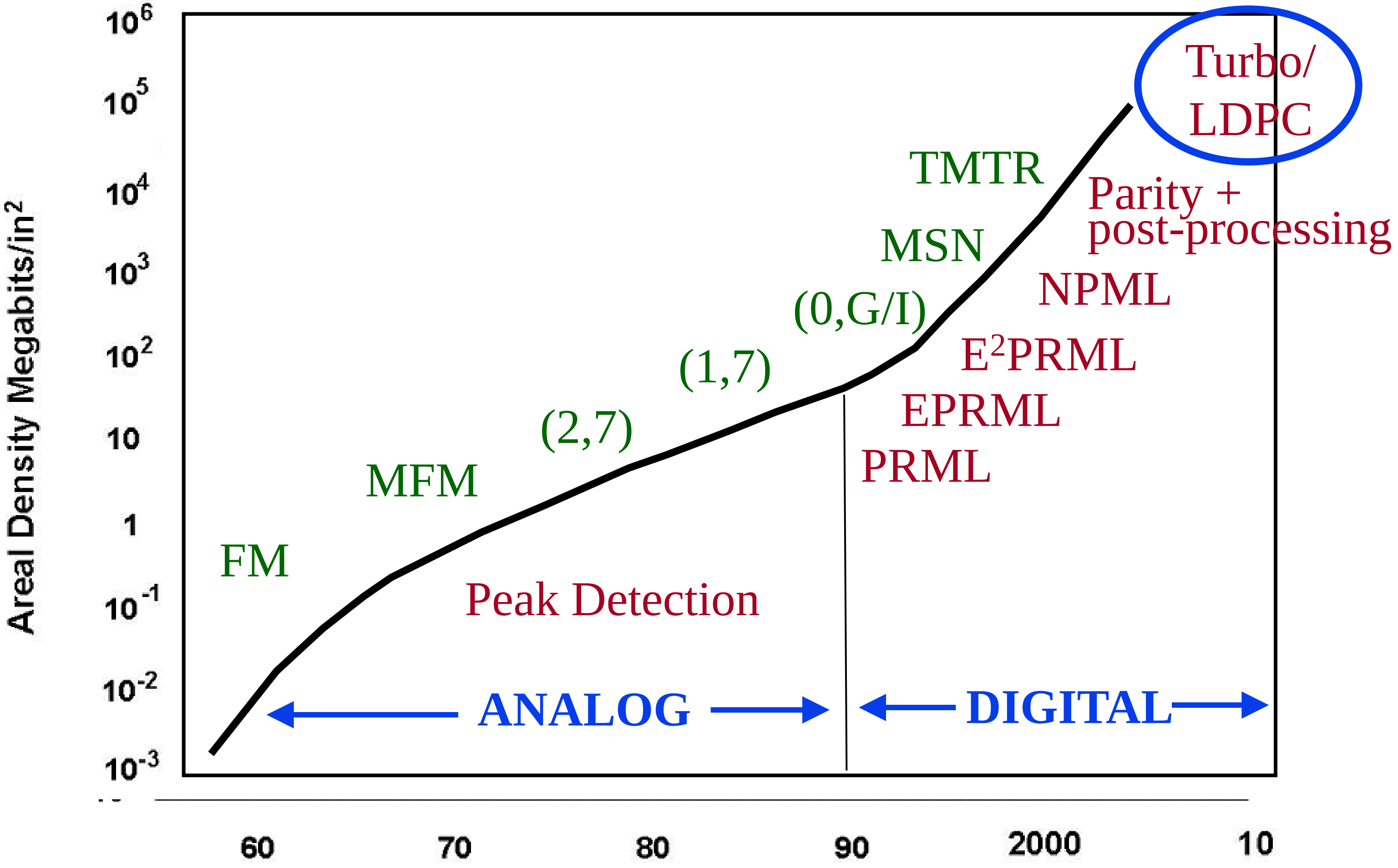
with album covers

Toshiba 1.8" drive

40.0 Gigabytes

(80GB on the way!)

Signal Processing and Coding Innovation



Key References and Their Impact…

[1] A.J. Viterbi, “Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm,” *IEEE Transactions on Information Theory*, vol. IT-13, no. 2, pp. 260-269, April 1967.

[2] A.J. Viterbi, “Convolutional Codes and Their Performance in Communication Systems,” *IEEE Transactions on Communications Technology*, vol. COM-19, no. 5, pp. 751-772, October 1971.

[3] A.J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York, NY: McGraw-Hill, Inc., 1979, Ch. 4.9, pp. 272-284.

[4] A.J. Viterbi, “An Intuitive Justification and a Simplified Implementation of the MAP Decoder for Convolutional Codes,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 260-264, February 1998.

PRML …

[1] "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm"

- **Since the introduction of PRML technology in 1990, the VA has been the standard detection method in disk drives.**

Coded PRML …

[2] "Convolutional Codes and Their Performance in Communication Systems"

[3] *Principles of Digital Communication and Coding*

- **Since the mid-1990's, error event characterization of partial-response channels has been used to bound performance and to design constrained modulation codes that detect and/or forbid dominant error events.**

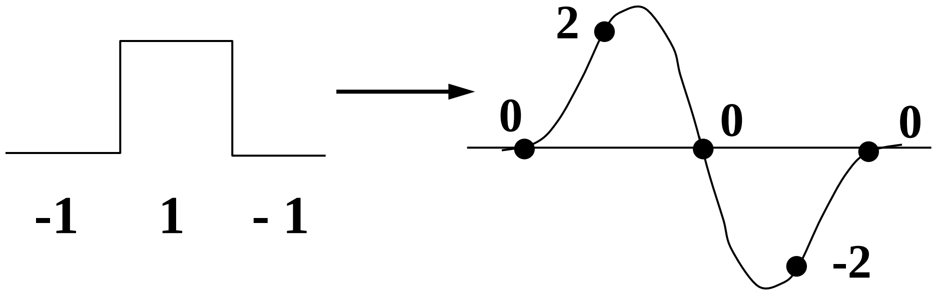
Turbo Equalization and Channel Capacity

[4] “An Intuitive Justification and a Simplified Implementation of the MAP Decoder for Convolutional Codes”

- **“Turbo-equalized” recording channels (proposed) use a modified “dual-max” algorithm for detection and a difference-metric LDPC decoder.**
- **Sharp estimates of the recording channel capacity are calculated using a “generalized VA.”**

What is PRML?

- “PR” = Partial Response [Class-4] Equalization



$$y_n = x_n - x_{n-2}$$

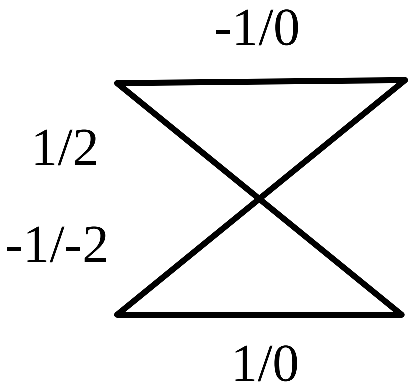
$$h(D) = (1 - D^2)$$
$$= (1 - D)(1 + D)$$

- “ML” = Maximum Likelihood Sequence Detection (VA)

“Dicode” trellis for even/odd interleaves

$$\boldsymbol{y_n = x_n - x_{n\text{-}1}}$$

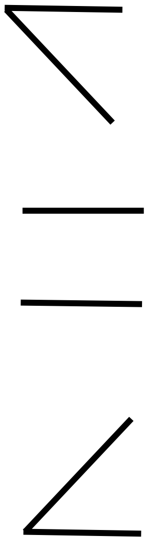
$$\boldsymbol{h(D)=1\text{-}D}$$



*The acronym “PRML” was coined by Andre Milewski, of IBM LaGaude.

Difference Metric VA for Dicode

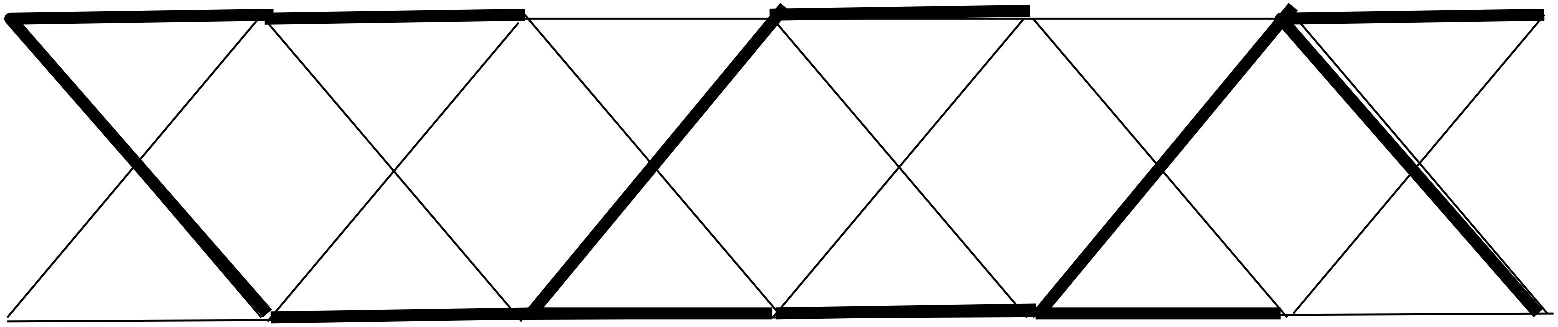
$$DM_n = \begin{cases} r_n - 1 & if \quad DM_{n-2} - r_n \leq -1 & \Rightarrow \\ DM_{n-2} & if \quad \text{-}1 < DM_{n-2} - r_n < 1 & \Rightarrow \\ r_n + 1 & if \quad 1 \leq DM_{n-2} - r_n & \Rightarrow \end{cases}$$



- Used in first commercial disk drive with PRML: IBM 681 (1990)

Difference Metric VA for Dicode

r	2.6	1	0.2	1.5	-1	1.3
DM	1.6	1.6	1.2	1.2	0	0.3



$\hat{y}$	2	0	0	0	-2	?

Beyond PRML

- Extended PRML - "E^NPRML"

$$h(D) = (1 - D)(1 + D)^{N+1}, \; N \geq 1$$

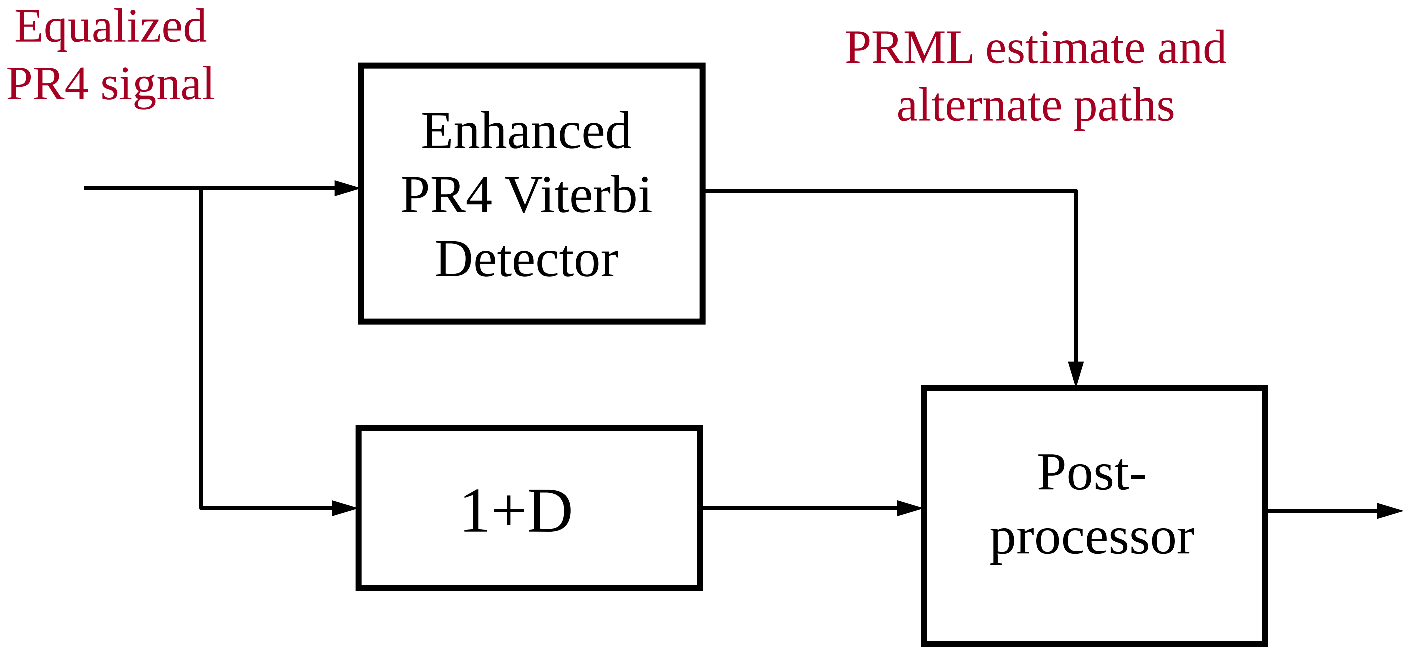
  - – Viterbi detector has 2^{N+2} states.
  - – EPR4 and E^2PR4 have been widely used in commercial drives.

- Noise-predictive PRML (a.k.a. Generalized PRML)

$$h(D) = (1 \underset{\substack{\uparrow \\ \text{PR4}}}{-} D^2)(1 + p_1 D \underset{\substack{\uparrow \\ \text{Noise-whitening filter}}}{+} p_2 D^2)$$

Post-Processor EPRML Detector

- "Turbo-PRML" (1993)


Equalized PR4 signal

Enhanced PR4 Viterbi Detector

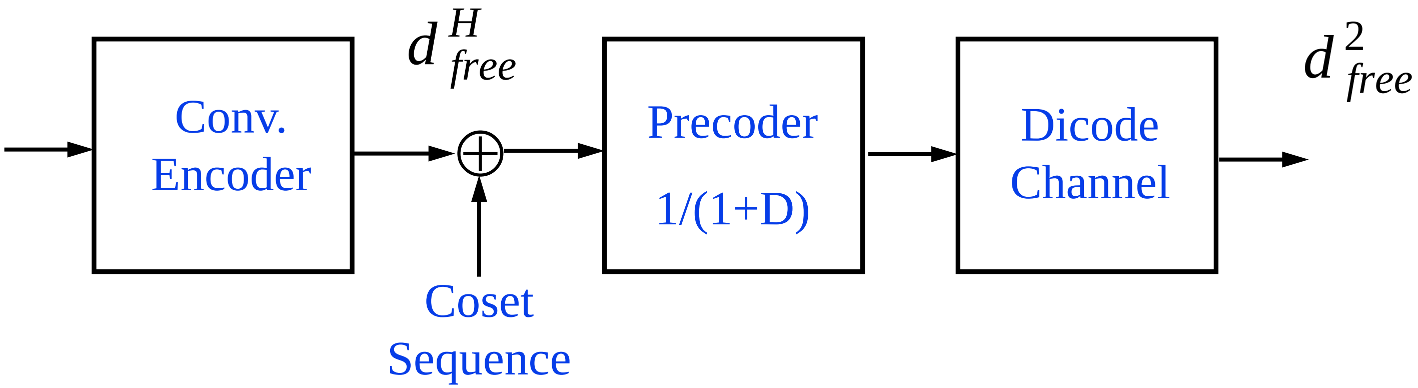
PRML estimate and alternate paths

1+D

Post-processor

Trellis-coded PRML

- Convolutional code with channel precoder
- Combined convolutional code and channel trellis detector



$$d^2_{free} \geq \begin{cases} d^H_{free} & if \quad d^H_{free} \text{ is even} \\ d^H_{free}+1 & if \quad d^H_{free} \text{ is odd} \end{cases}$$

Distance-Enhancing Constrained Codes

- Characterize PR channel error-events using error-state diagram analysis. (See [2], [3].)

- Determine modulation constraints that reduce and/or forbid dominant error events, and design code.

- Incorporate channel and code constraints into detector trellis, or use reduced-state trellis and a post-processor.

Error Event Analysis – E²PR4

- E²PR4: $h(D) = (1-D)(1+D)^3$

  $d^2_{free} = 6$

- Input "error" events: $e(D) = x_1(D) - x_2(D)$

$d^2(e)$	$e = x_1 - x_2$
6	**+ — +**
8	**+ — + 0 0 + — +** **+ — + — (+ —)** **+ — + — (+ —) +**

Distance-Enhancing Codes

- Matched-Spectral-Null (MSN) codes
  - ➢DC-null and order-K Nyquist null on E^2PR4:

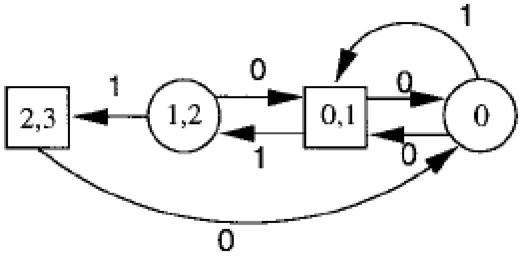
$$d_{free}^{2} \geq 2(K+3)$$

- Maximum-Transition-Run MTR(*j*,*k*) codes
  - ➢Limit number of consecutive 1's to *j* (*k*) on even (odd) phase
  - ➢For E^2PR4, the MTR(2,3) constraint yields:

$$d_{free}^{2} = 10$$

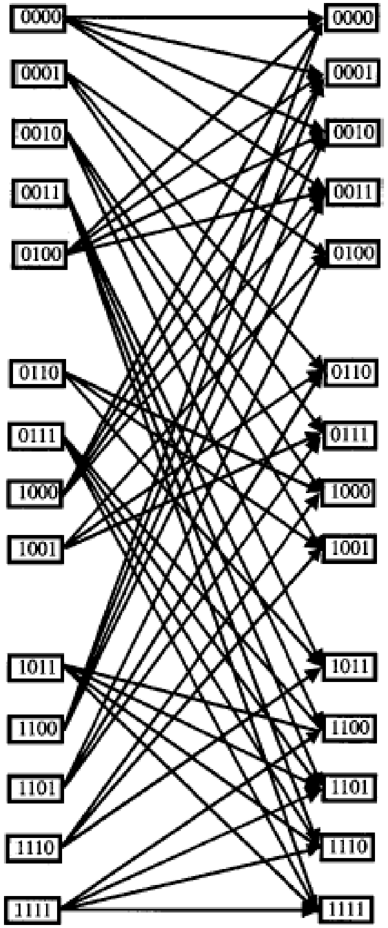
- Parity-check codes
  - ➢Detect variety of error events

Combined Code-Channel Trellis



G

MTR(2,3) constraint graph
(NRZI format)

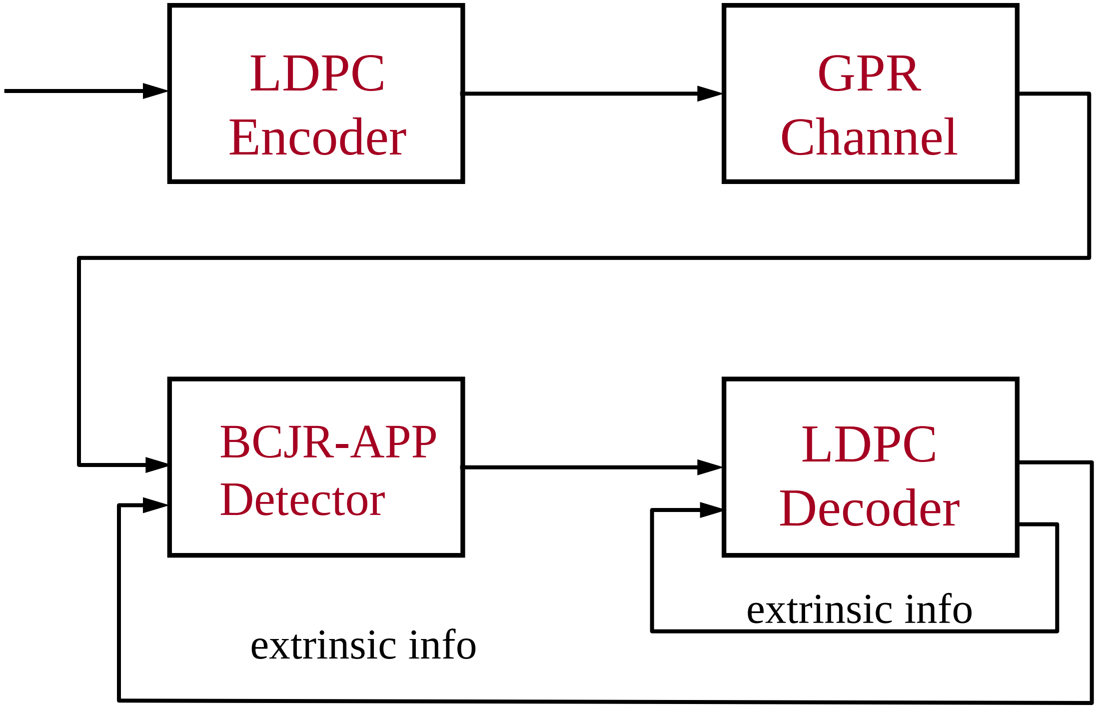


Combined MTR(2,3)
and E^2PR4 trellis
(NRZ format)

State-of-the-Art Channel

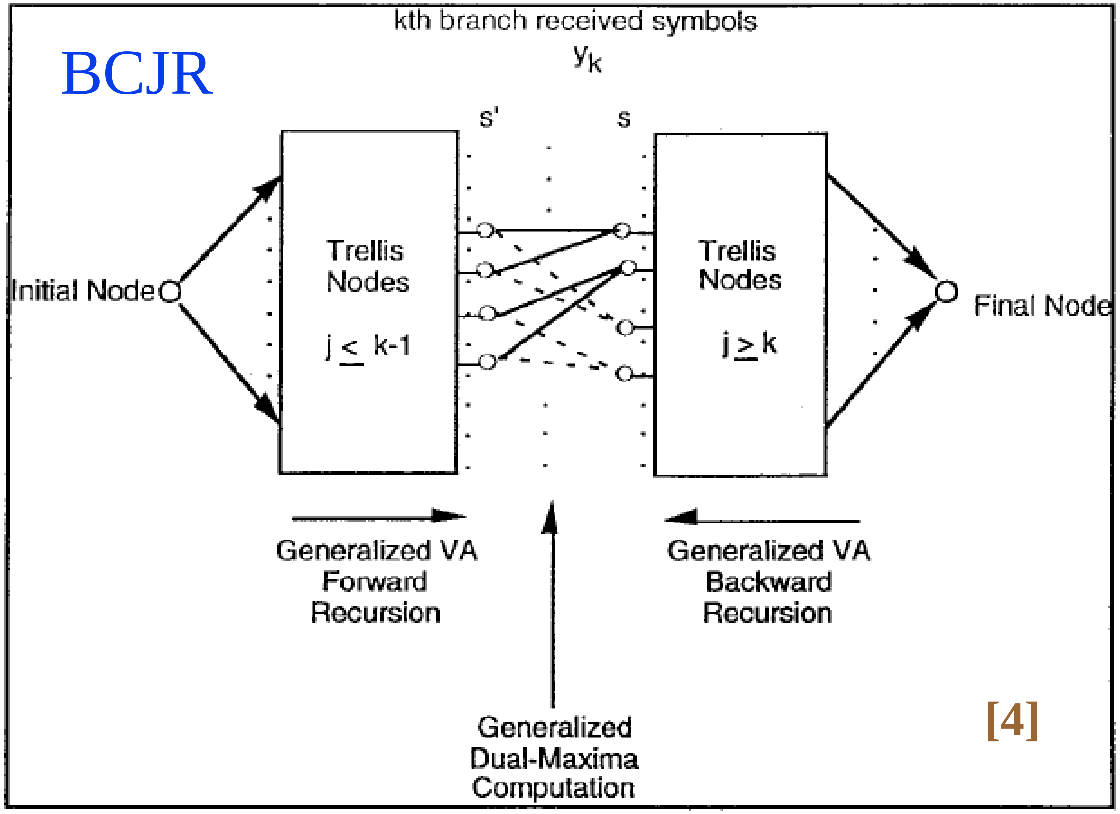
- Rate-96/104 dual-parity code with MTR(3,3) constraints
  - Eliminates all error events of type: + – + – , + – + – + , + – + –
  - Eliminates half of events of type: + – +
  - Detects error events of type: +, + –, + – +, and +00+
- 16-state NPML detector with dual-parity post-processing
  - Gain of 0.75dB over rate-48/49, no parity, at P_e(sector)=10^{-6}

Turbo Equalization



- Length-4376 LDPC code
- Gain ~4 dB over uncoded NPML at P_e(symbol)=10^{-5}
- Gap to capacity ~1.5dB

Simplified BCJR: Dual-Max Detector



$$L_n = \max_{s',s:\,x=1}\{A_{n-1}(s') + B_n(s) + c_n(s',s)\} - \max_{s',s:\,x=-1}\{A_{n-1}(s') + B_n(s) + c_n(s',s)\}$$

Capacity of Magnetic Recording Channels

- Binary input, linear ISI, additive, i.i.d. Gaussian noise

$$y[i] = \sum_{k=0}^{n-1} h[k]\,x[i-k] + n[i]$$

- Capacity C

$$\begin{aligned} C &= \max_{P(X)} I(X;Y) \\ &= \max_{P(X)} H(Y) - H(Y \mid X) \\ &= \max_{P(X)} H(Y) - \frac{1}{2}\log(\pi e N_0) \end{aligned}$$

For a given $P(X)$, we want to compute $H(Y)$

Computing Entropy Rates

- Shannon-McMillan-Breimann theorem implies

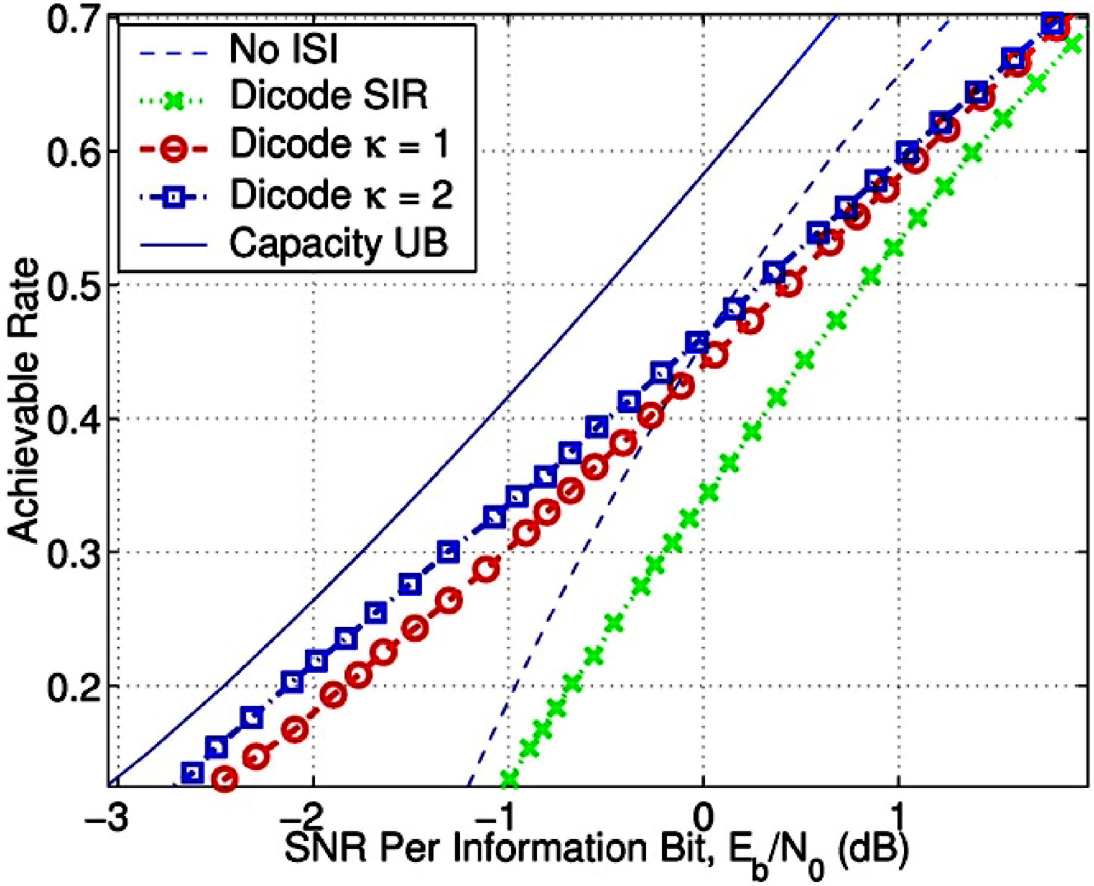
$$-\frac{1}{n} log\ p\left(y_1^n\right) \xrightarrow[a.s.]{} H(Y)$$

as $n \to \infty$, where y_1^n is a single long sample realization of the channel output process.

- The probability $p(y_1{}^n)$ can be computed using the forward recursion of the BCJR - APP algorithm.

- In the log domain, this forward recursion can be interpreted as a "generalized Viterbi algorithm." (See [4].)

Capacity Bounds for Dicode h(D)=1-D



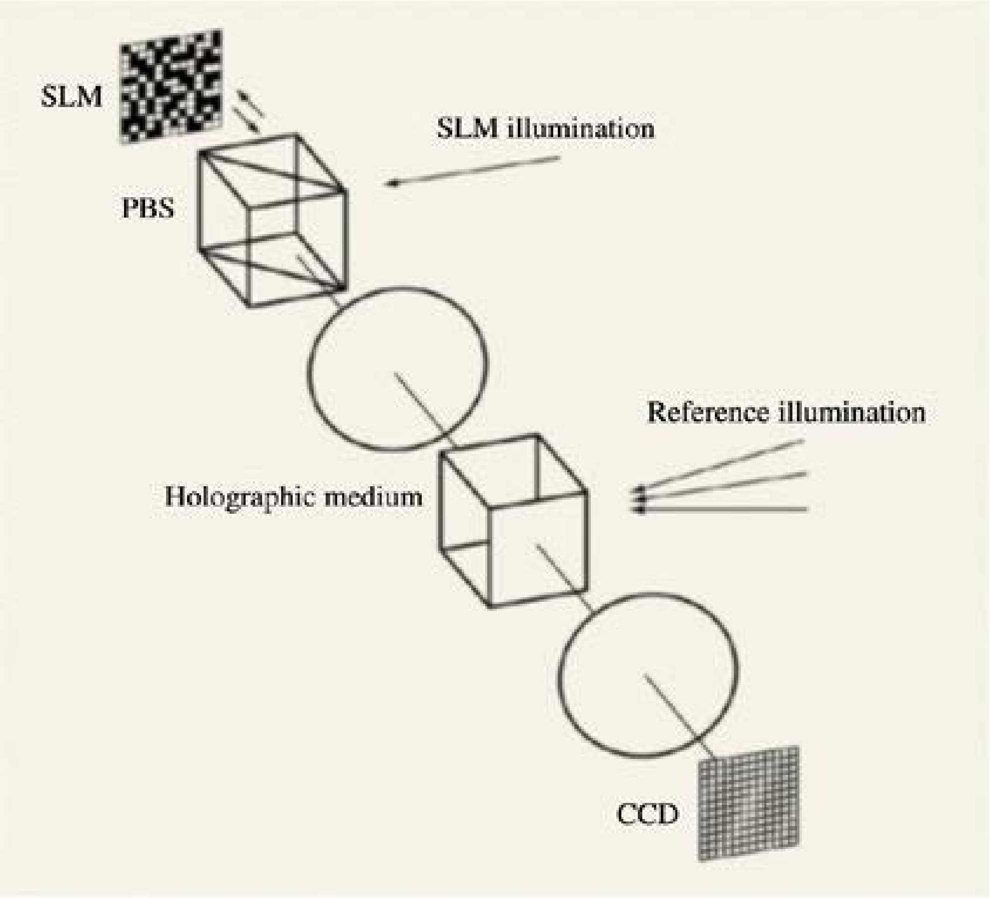
Concluding Remarks

- The Viterbi Algorithm and related ML performance evaluation techniques have been vital to the advancement of data storage technology – magnetic and optical - since 1990.

- The “Viterbi architecture” for APP computation has influenced the development and evaluation of capacity-approaching coding schemes for digital recording applications.

- Future storage technologies offer interesting challenges in detection and decoding…

Holographic Recording

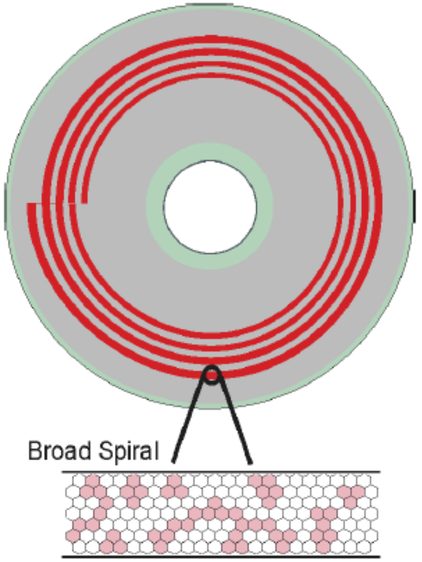


2-D Intersymbol Interference

$$h = \begin{bmatrix} 0.05 & 0.21 & 0.05 \\ 0.20 & 0.91 & 0.19 \\ 0.01 & 0.10 & 0.02 \end{bmatrix}$$

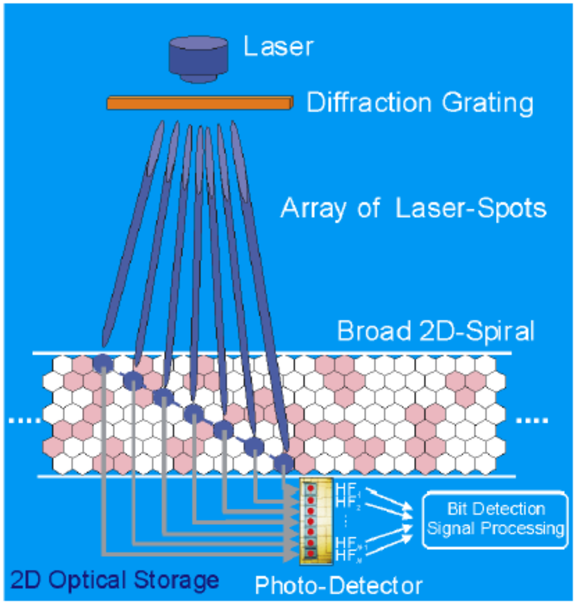
Two-Dimensional Optical Storage (TwoDOS)

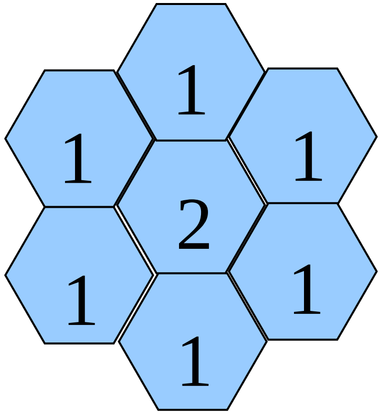
Disc



(spiral contains 11 bit-rows)

Readout





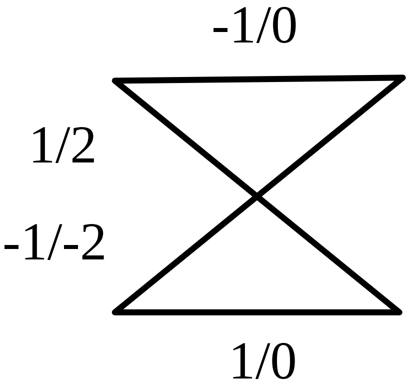
2-D Impulse response

Courtesy of Wim Coene, Philips Research

And, finally…

- **Congratulations – and many thanks – Andy!!**

  on the occasion of your milestone birthday, and for your many landmark contributions to science, technology, and engineering education.



PRML References

- H. Kobayashi and D.T. Tang, “Application of partial-response channel coding to magnetic recording systems,” *IBM J. Res. Develop*., vol. 14, pp. 368-375, July 1970.
- H. Kobayashi, “Application of probabilistic decoding to digital magnetic recording systems,” *IBM J. Res. Develop*., vol. 15, pp. 65-74, Jan. 1971.
- H. Kobayashi, “Correlative level coding and maximum-likelihood decoding,” *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 586-594, Sept. 1971.
- G.D. Forney, Jr., “Maximum likelihood sequence detection in the presence of intersymbol interference,” *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363-378, May 1972.
- R.D.Cideciyan, et al., "A PRML System for Digital Magnetic Recording," *IEEE J. Select. Areas Commun*., vol. 10, no. 1, pp. 38 – 56, Jan. 1992.

EPRML References

- H.K. Thapar and A.M. Patel, “A class of partial response systems for increasing storage density in magnetic recording,” *IEEE Trans. Magn.*, pp. 3666-3678, Sept. 1987.
- G. Fettweis, R. Karabed, P. H. Siegel, and H. K. Thapar, “Reduced-complexity Viterbi detector architectures for partial response signaling,” in *Proc. 1995 Global Telecommun. Conf. (Globecom’95)*, Singapore, pp. 559–563.
- R.Wood, “Turbo-PRML: A compromise EPRML detector,” *IEEE Trans. Magn.*, vol. 29, pp. 4018–4020, Nov. 1993.
- K. K. Fitzpatrick, “A reduced complexity EPR4 post-processor,” *IEEE Trans. Magn.*, vol. 34, pp. 135–140, Jan. 1998.
- J. D. Coker, E. Eleftheriou, R. L. Galbraith, and W. Hirt, “Noise-predictive maximum likelihood (NPML) detection,” *IEEE Trans. Magn.*, pt. 1, vol. 34, pp. 110–117, Jan. 1998.

Coded PRML References

- J. K. Wolf and G. Ungerboeck, ``Trellis coding for partial-response channels," *IEEE Trans. Commun*., vol. COM-34, no. 8, pp. 765-773, Aug. 1986.
- R. Karabed and P. Siegel, “Matched spectral-null codes for partial response channels,” *IEEE Trans. Inform. Theory*, vol. 37, no. 3, pp. 818–855, May 1991.
- J. Moon and B. Brickner, “Maximum transition run codes for data storage systems,” *IEEE Trans. Magn*., vol. 32, pp. 3992–3994, Sept. 1996.
- W. Bliss, “An 8/9 rate time-varying trellis code for high density magnetic recording,” *IEEE Trans. Magn*., vol. 33, pp. 2746–2748, Sept. 1997.
- S.A. Altekar, M. Berggren, B.E. Moision, P.H. Siegel, J.K. Wolf, “Error event characterization on partial-response channels”, *IEEE Trans. Inform. Theory*, vol. 45, no. 1 , pp. 241 –247, Jan. 1999.

Coded PRML References (cont.)

- R. Karabed, P.H. Siegel, and E. Soljanin, ``Constrained coding for binary channels with high intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1777-1797, Sept. 1999.
- T. Conway, “A new target response with parity coding for high density magnetic recording channels,” *IEEE Trans. Magn*., vol. 34, no. 4, pp. 2382–2486, July 1998.
- Cideciyan R.D., Coker, J.D., Eleftheriou, E., and Galbraith, R.L.: “Noise predictive maximum likelihood detection combined with parity-based post-processing”, *IEEE Trans. Magn*., vol. 37, no. 2, pp. 714-720, March 2001.
- R.D. Cideciyan, E. Eleftheriou, B.H. Marcus, and D. S. Modha, “Maximum transition run codes for generalized partial response channels,” *IEEE J. Select. Areas Commun*., vol. 19, no. 4, pp. 619-634, April 2001.
- R.D. Cideciyan and E. Eleftheriou, “Codes satisfying maximum transition run and parity-check constraints, *Proc. IEEE Int. Conf. Commun*., vol. 27, no. 1, June 2004, pp. 635 – 639.

Turbo Equalization References

- L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Sep 1974.

- W. Ryan, "Performance of high rate turbo codes on a PR4-equalized magnetic recording channel," *Proc. 1998 Int. Conf. Commun.*, vol. 2, June 1998, pp. 947-951.

- T. Souvignier, A. Friedmann, M. Oberg, P. H. Siegel, R. E. Swanson, and J. K. Wolf, “Turbo decoding for PR4: parallel versus serial concatenation,” *Proc. IEEE ICC’99*, Vancouver, Canada, June 1999, pp. 1638–1642.

- B. M. Kurkoski, P. H. Siegel, J. K. Wolf, “Joint Message-Passing Decoding of LDPC Codes and Partial-Response Channels,” *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1410-1422, June 2002.

Capacity Calculation References

- D. Arnold and H.-A. Loeliger, “On the information rate of binary-input channels with memory,” *Proc. IEEE ICC 2001*, (Helsinki, Finland), June 2001, pp. 2692–2695.
- H. D. Pfister, J. B. Soriaga, and P. H. Siegel, “On the achievable information rates of finite state ISI channels,” *Proc. IEEE GLOBECOM 2001*, (San Antonio, Texas), Nov. 2001, pp. 2992–2996.
- A. Kavcic, “On the capacity of Markov sources over noisy channels,” *Proc. IEEE GLOBECOM 2001*, (San Antonio, Texas), Nov. 2001, pp. 2997–3001.
- P. Vontobel and D. M. Arnold, “An upper bound on the capacity of channels with memory and constraint input,” *Proc. IEEE Inform. Theory Workshop*, (Cairns, Australia), Sept. 2001.
- S. Yang and A. Kavcic, “Capacity of Partial Response Channels,” *Handbook on Coding and Signal Processing for Recording Systems*, CRC Press 2004, Ch. 13.