Capacity and Beyond

Erozan M. Kurtas

Seagate
We turn on ideas

Acknowledgement

M. Fatih Erden

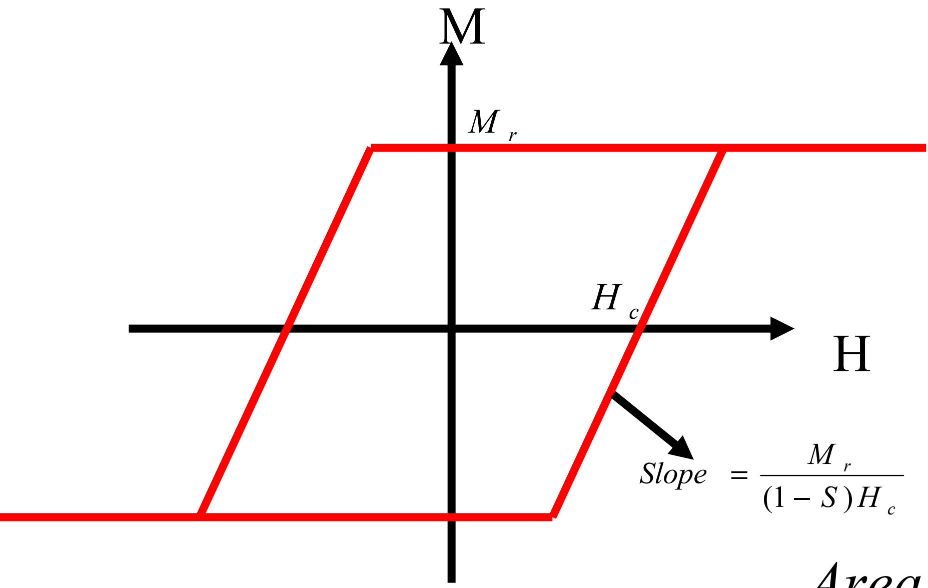
Sami Iren

Dieter Arnold

Raman Venkataramani

Inci Ozgunes

Conventional system

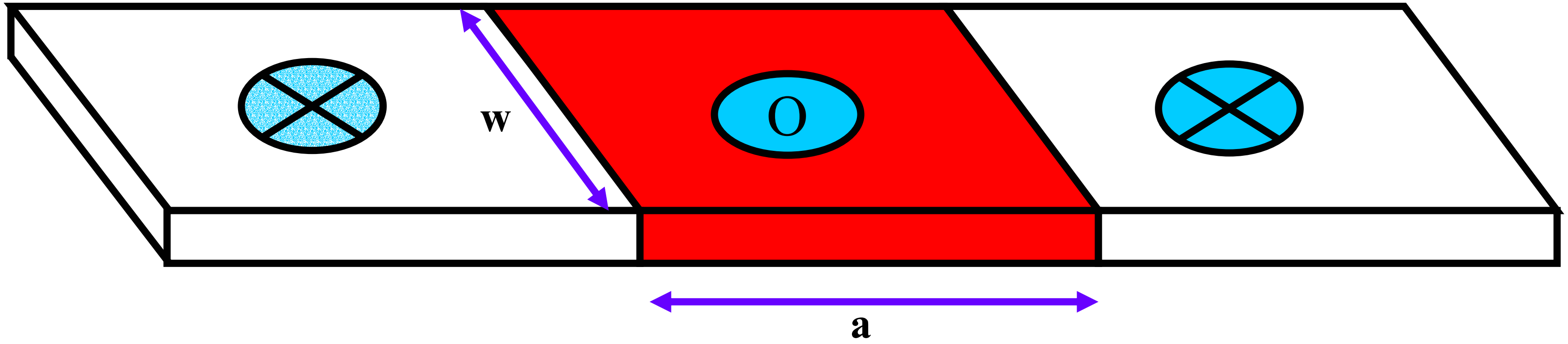


$Slope = \frac{M_r}{(1-S)H_c}$

Store bit by

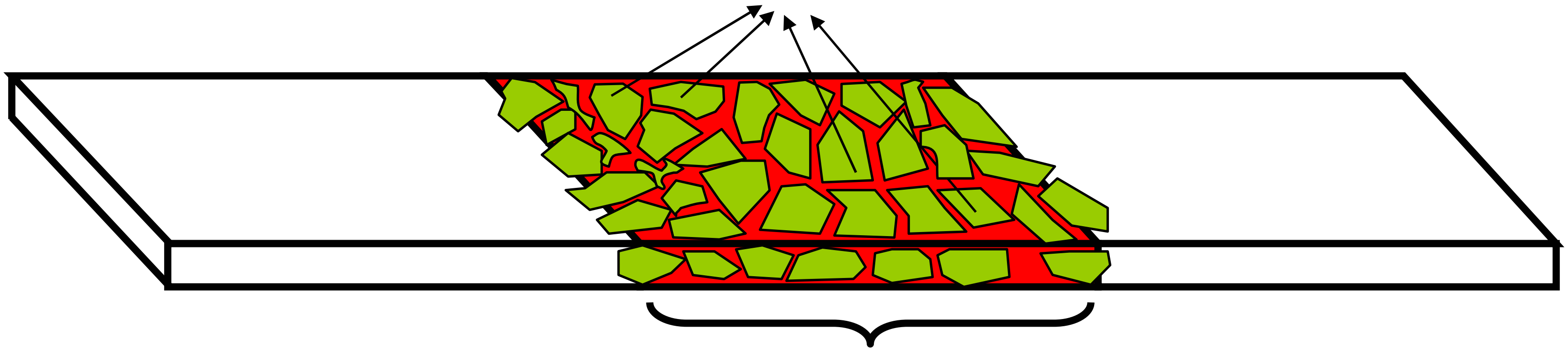
- Applying external H > Hc (magnetize up)
- Applying external H < -Hc (magnetize down)

$$AreaDensity[bits/inch^2] = \frac{1[bit]}{a[inch]w[inch]}$$



In reality there are tiny grains

Tiny grains with finite volume V and anisotropy coefficient Ku



Minimum grain number fixed to preserve SNR

Maximum value limited by maximum writer field

Super paramagnetic limit : $$\frac{K_u V}{k_B T} \geq M$$

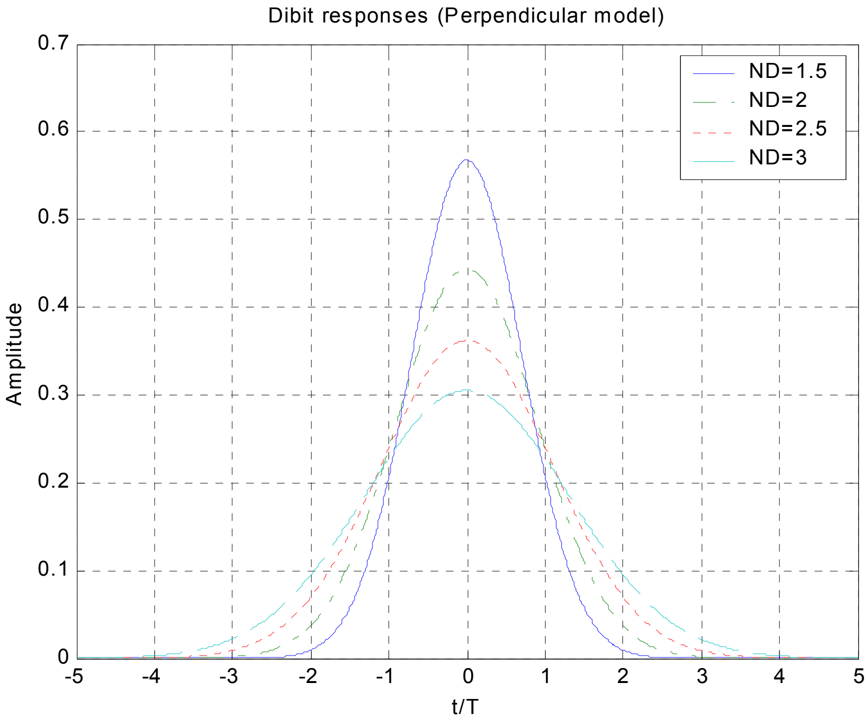
Maximum Attainable Areal Density is Limited

Boltzmann Constant

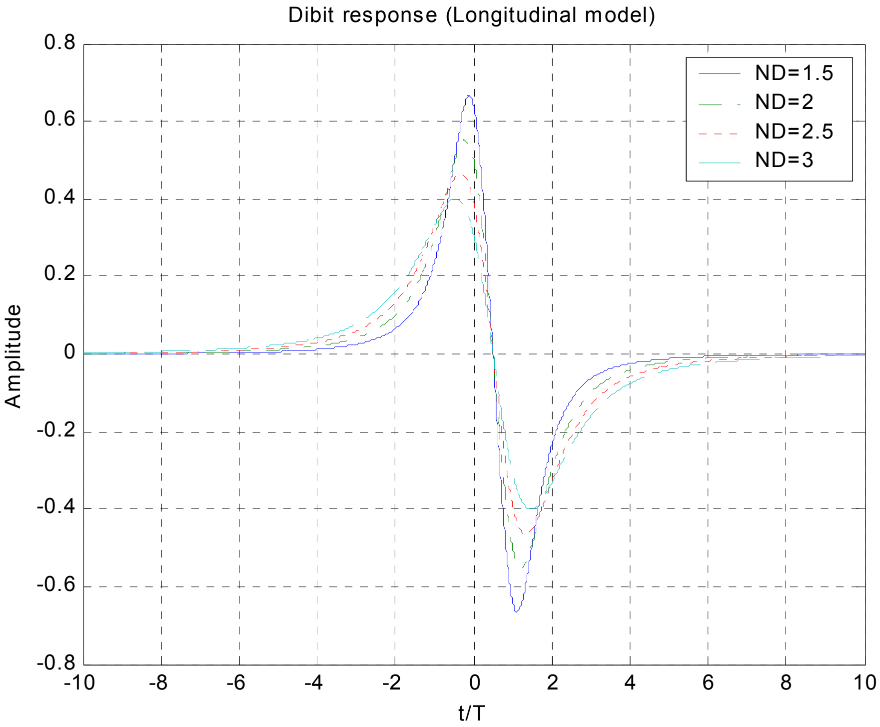
Temperature

Large number, like 60

Dibit Responses Compared

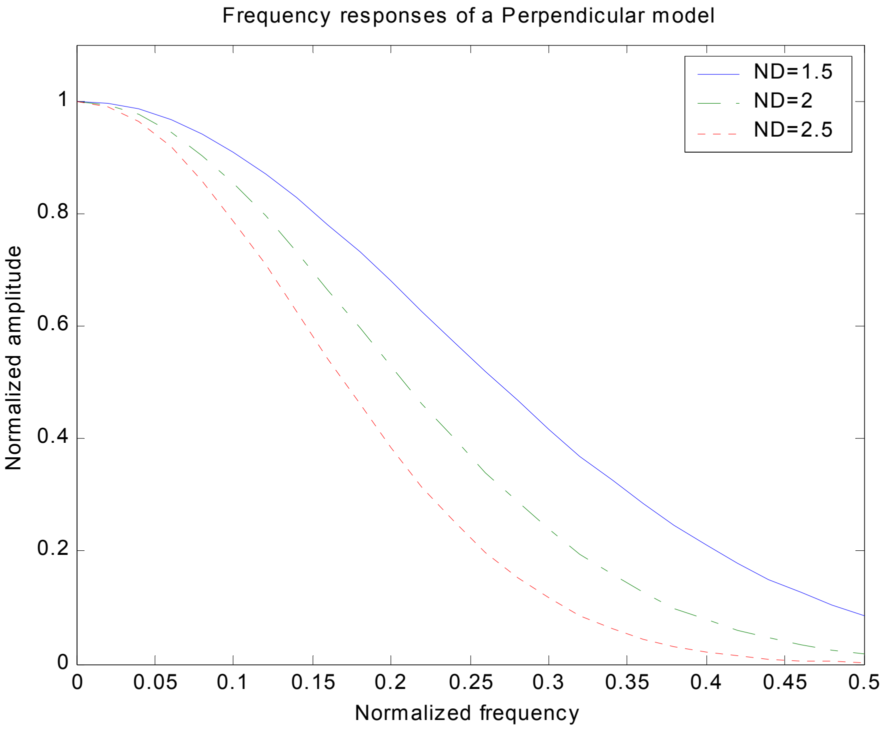


Perpendicular model

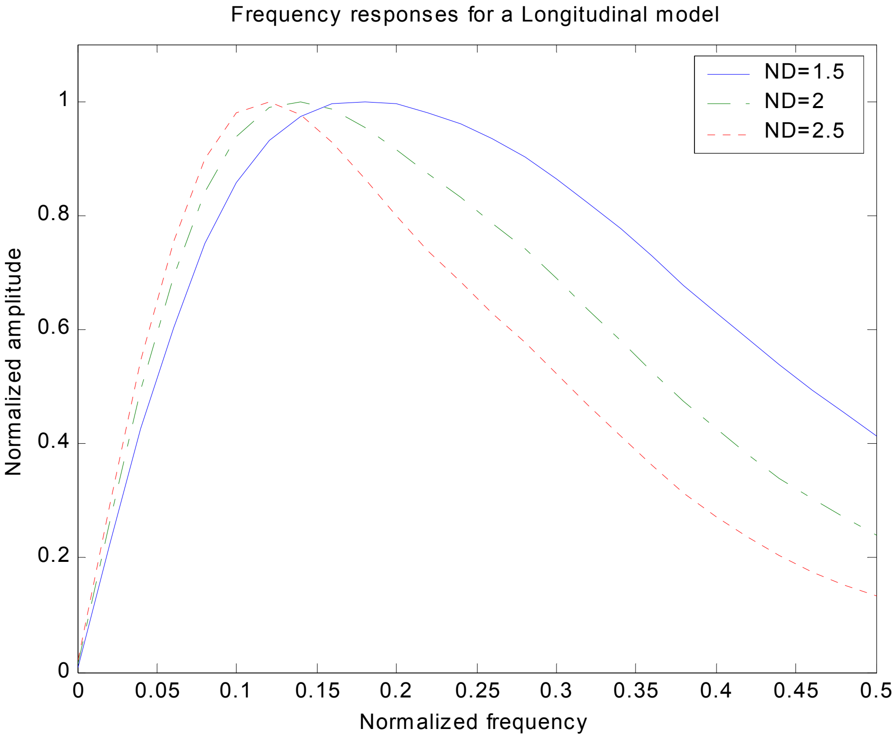


Longitudinal model

Frequency Responses of Dibit responses



Perpendicular model



Longitudinal model

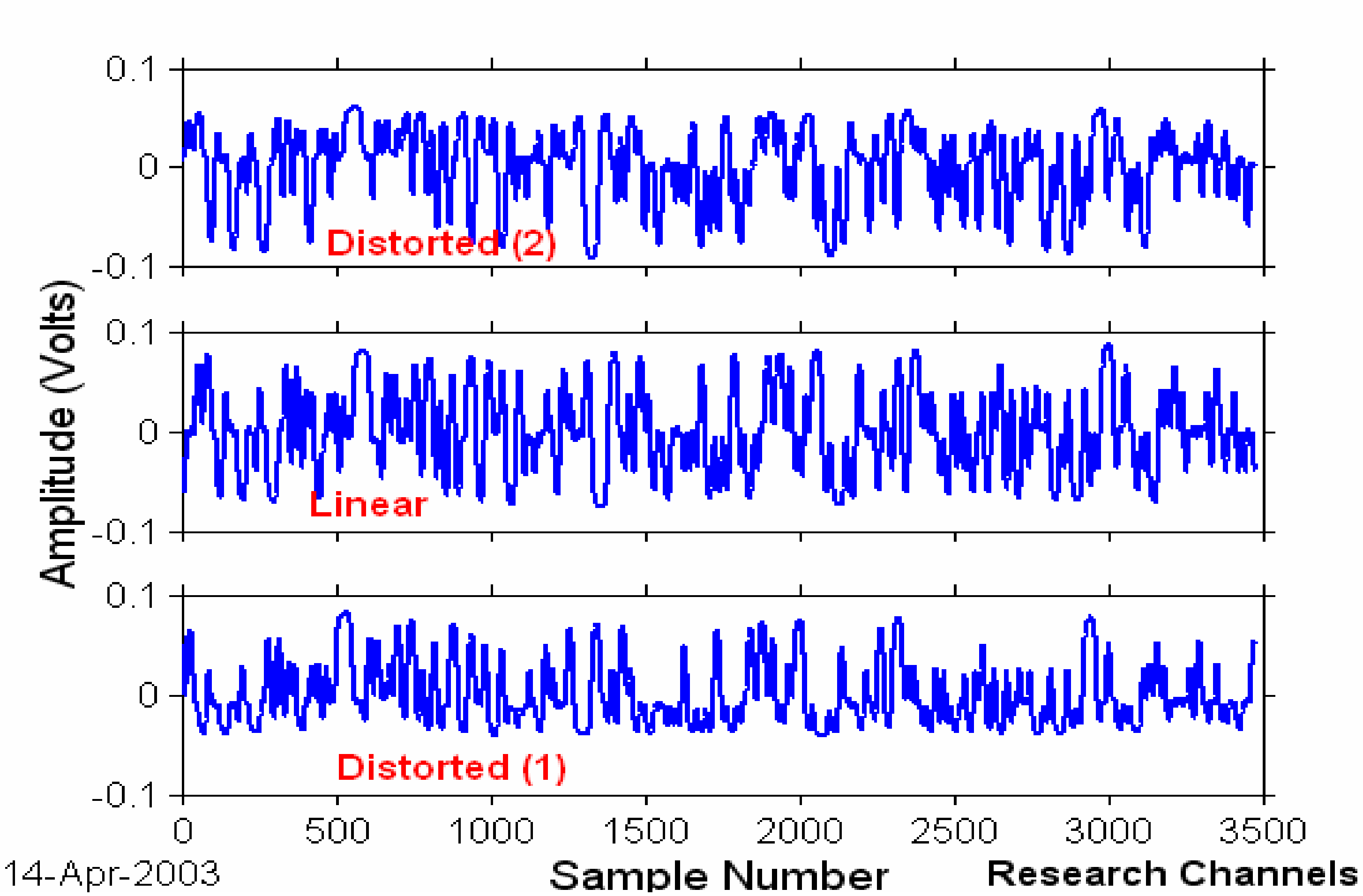
Reality is Quite Different

- Signal suffers from nonlinearities
  - NLTS
  - MR Asymmetry
  - Base Line Wanders
  - TAs
  - Other distortions

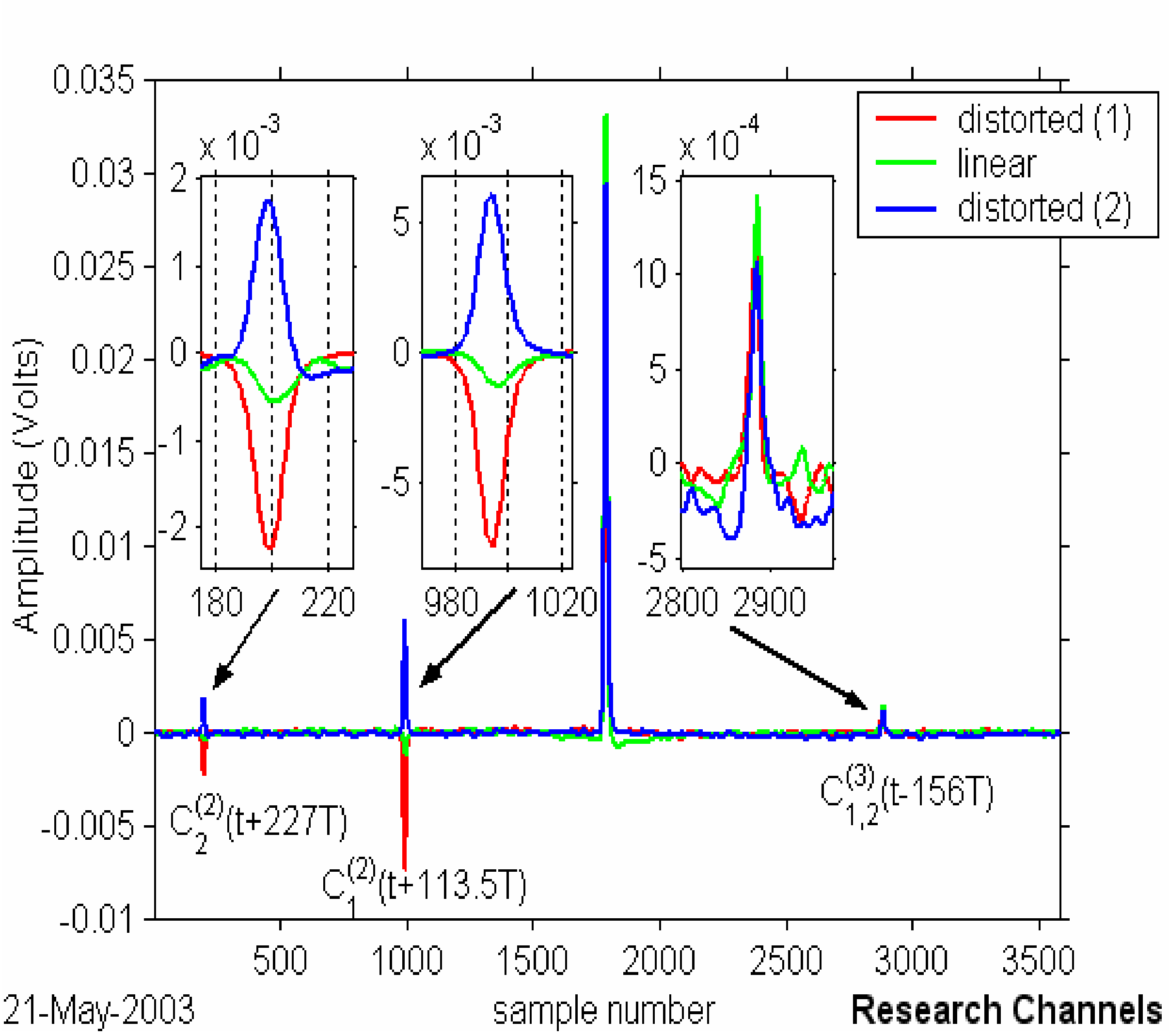
- Noise is Non-Gaussian

- Noise is Signal Dependent

Noise Free Real Signals with Nonlinearities



Extracted Dibit Response



extracted dibit

- provides information about systems linear response.
- is a convenient means for identifying nonlinearities present in system that show up as echoes around the main pulse.

Volterra Model of a Readback Signal

First Order Kernel

Linear Response

Second Order Non-linear Response

$$y(t)=\sum_k a_k C^{(1)}(t-kT)$$

Second Order Kernels

$$+\sum_k a_k a_{k-1} C_1^{(2)}(t-kT)+\sum_k a_k a_{k-2} C_2^{(2)}(t-kT)+\cdots$$

Third Order Non-linear Response

$$+\sum_k a_k a_{k-1} a_{k-2} C_{1,2}^{(3)}(t-kT)+\sum_k a_k a_{k-1} a_{k-3} C_{1,3}^{(3)}(t-kT)+\cdots$$

Memory Length

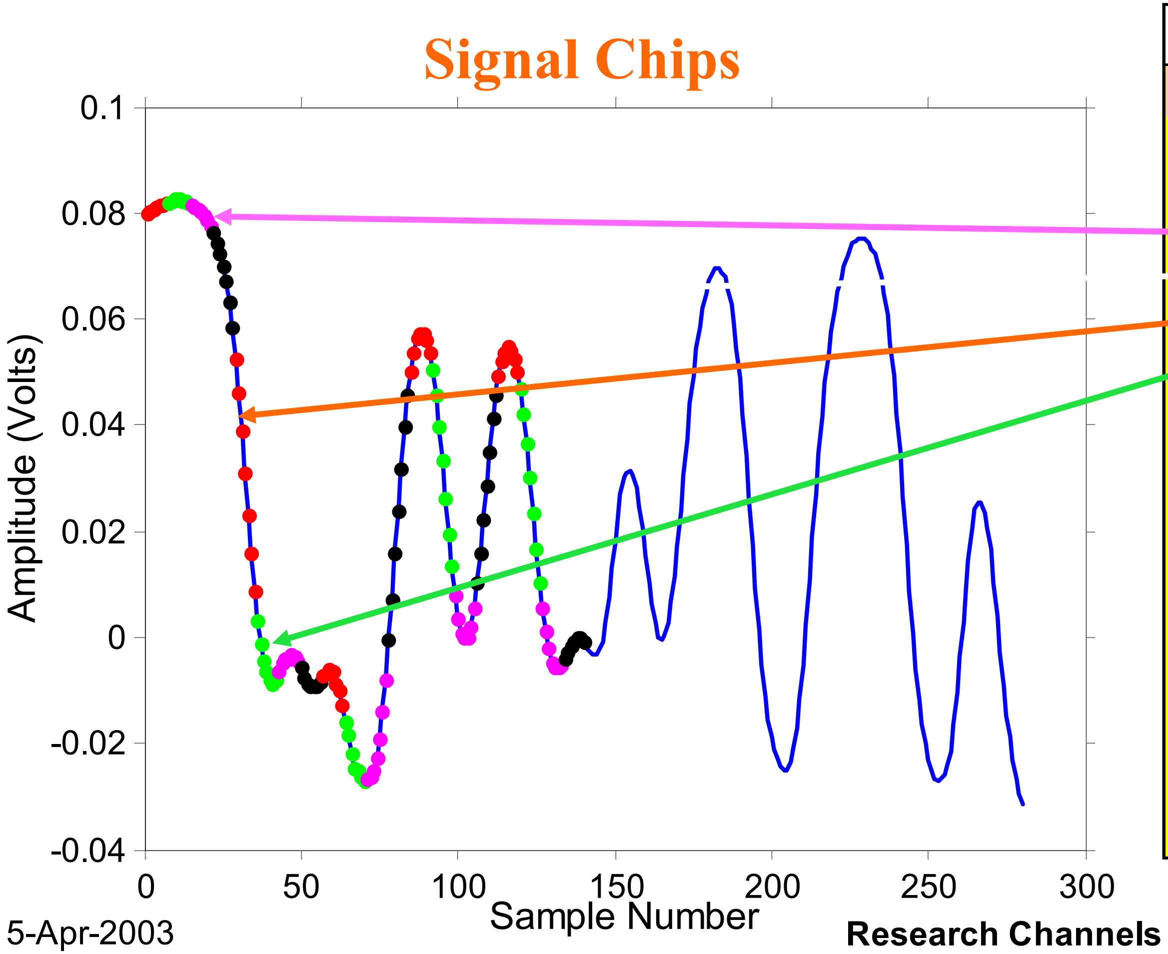
Third Order Kernels

VM can be characterized by 2^{L-1} kernels, $C^{(l)}(t), \quad l=2,\cdots L$

For magnetic recording, only a few of the kernels are significant.

A rule of thumb: L ~ the extent of the dipulse (in units of bit interval, T)

VM Kernels can be conveniently identified from measured PRBS signals

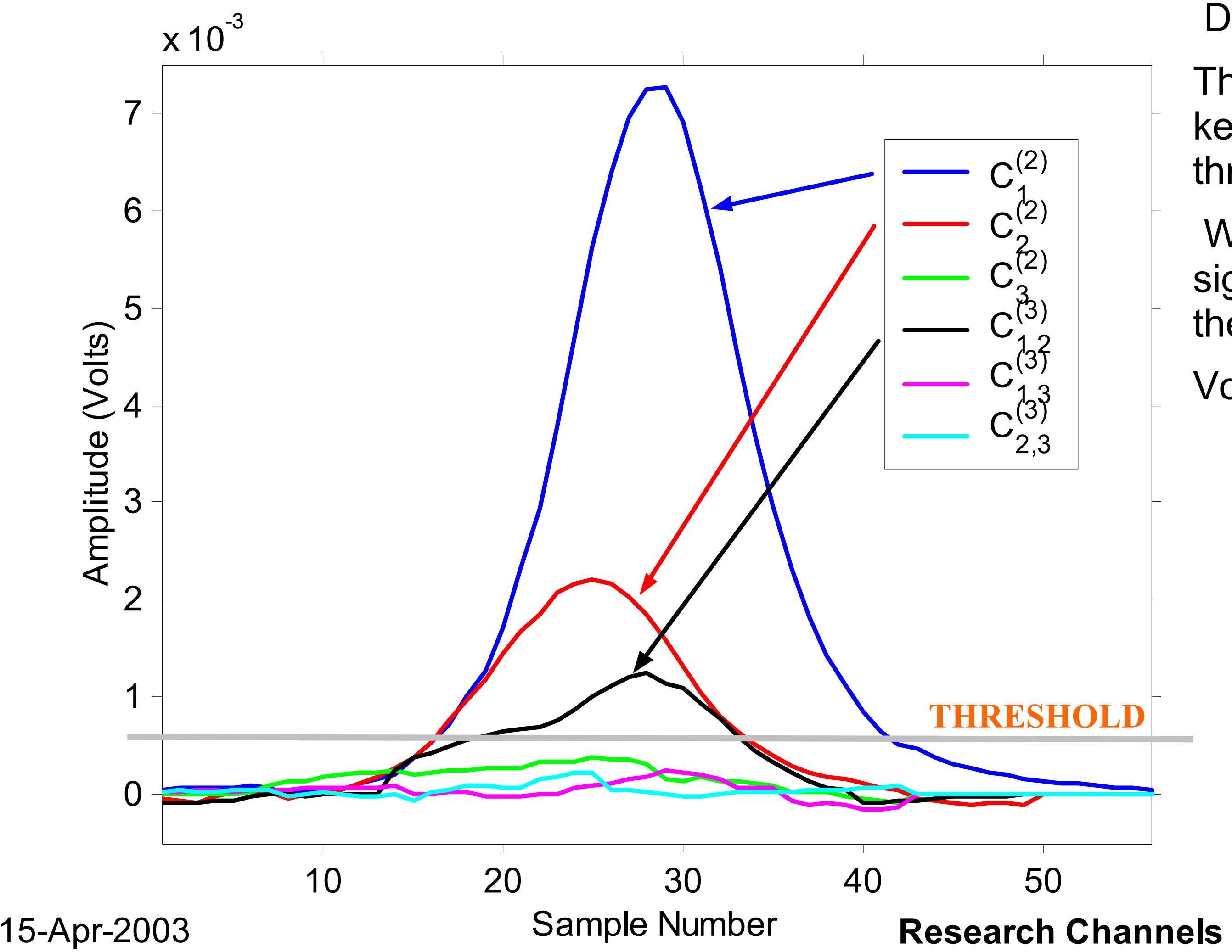


All Patterns of Length L=9

	a_{-4}	a_{-3}	a_{-2}	a_{-1}	a_0	a_1	a_2	a_3	a_4
1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	-1
3	1	1	1	1	1	1	1	-1	1
4	1	1	1	1	1	1	-1	1	-1
5	1	1	1	1	1	-1	1	-1	1
6	1	1	1	1	-1	1	-1	1	-1
7	1	1	1	-1	1	-1	1	-1	-1
•			•			•			•
•			•			•			•
•			•			•			•
506	-1	1	1	-1	1	-1	1	1	1
507	1	1	-1	1	-1	1	1	1	1
508	1	-1	1	-1	1	1	1	1	1
509	-1	1	-1	1	1	1	1	1	1
510	1	-1	1	1	1	1	1	1	1
511	-1	1	1	1	1	1	1	1	1
512	-1	-1	-1	-1	-1	-1	-1	-1	-1

No signal chip

Identification of Kernels



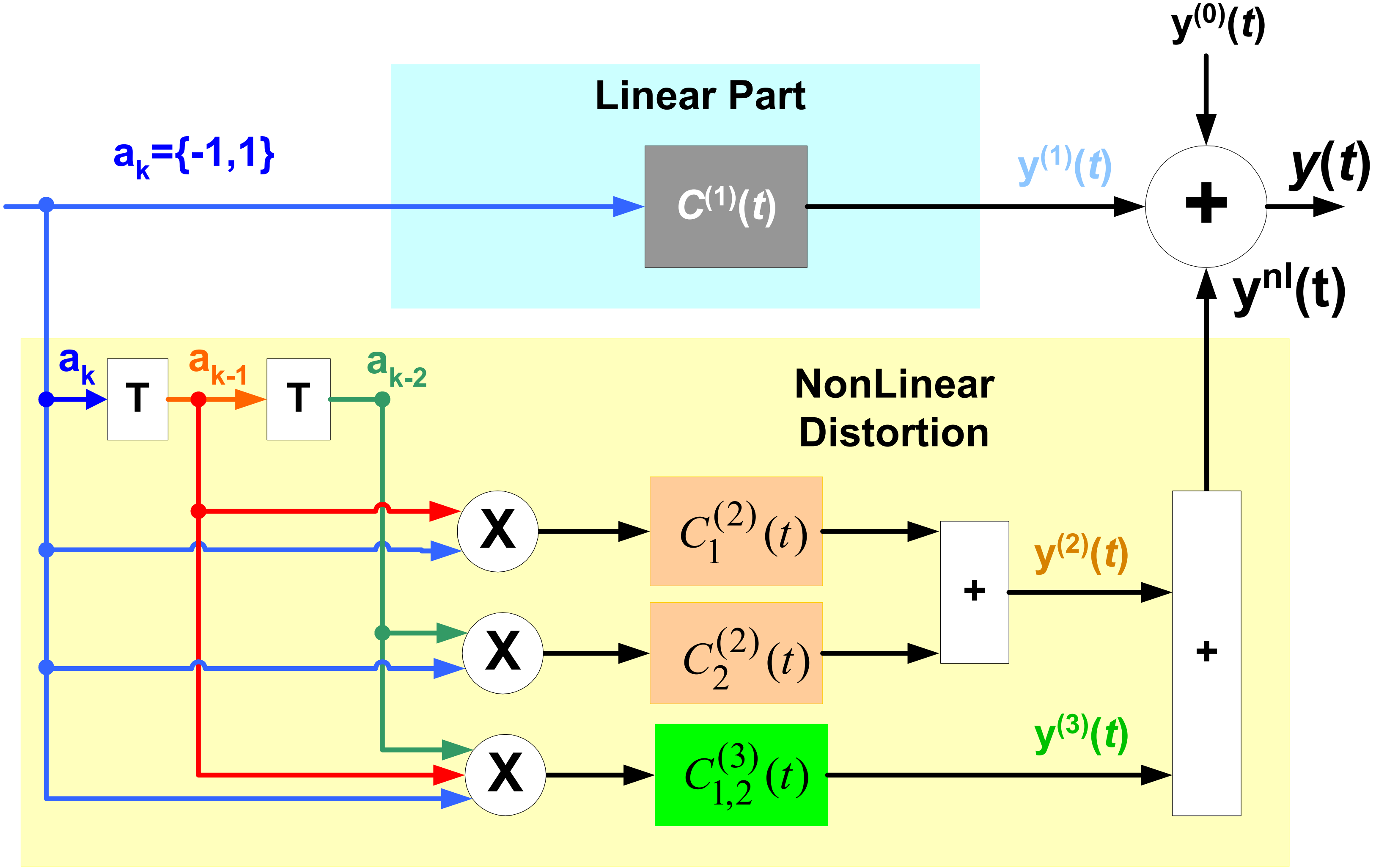
Distorted (1)

There are only three kernels above the threshold

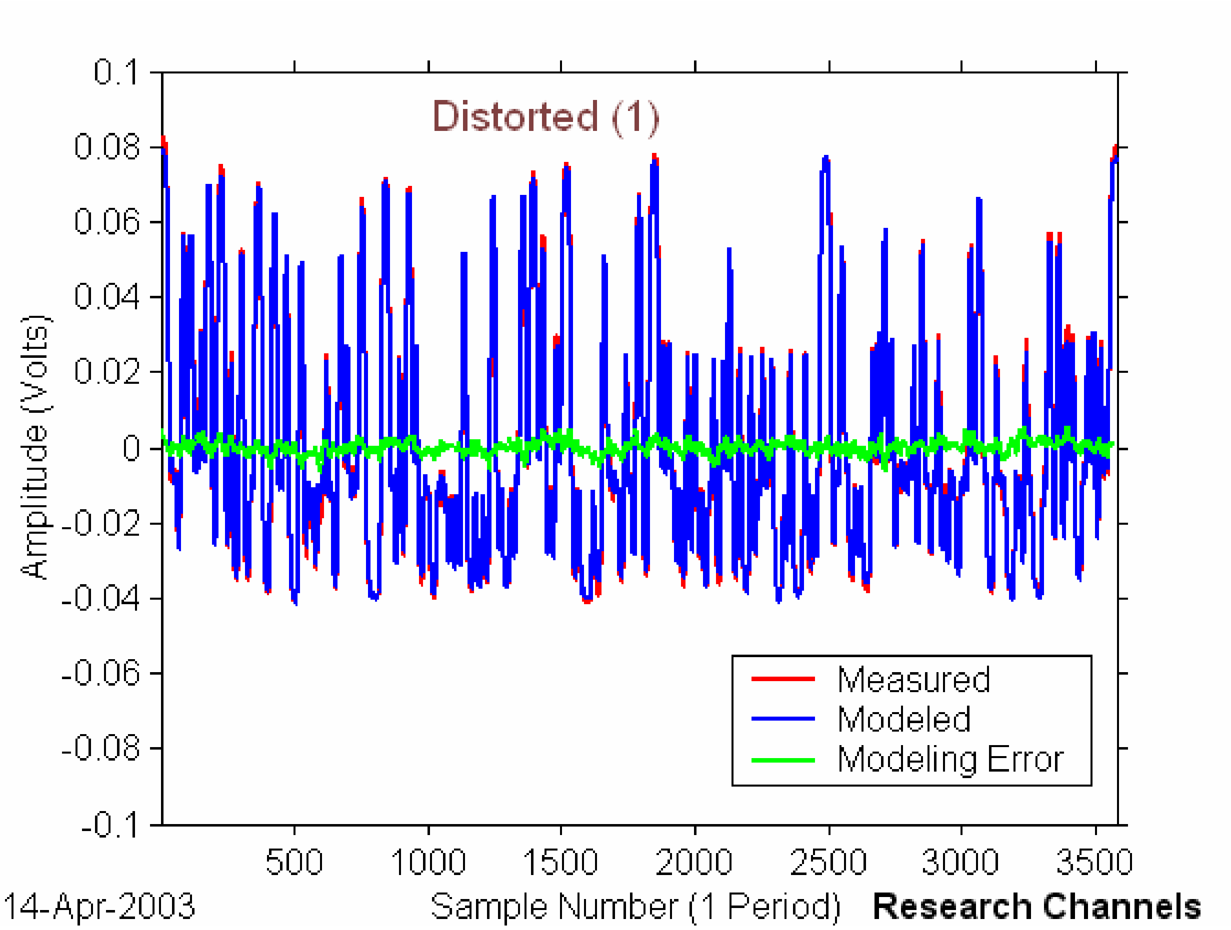
We declare them as significant and include in the Reduced Complexity

Volterra Model.

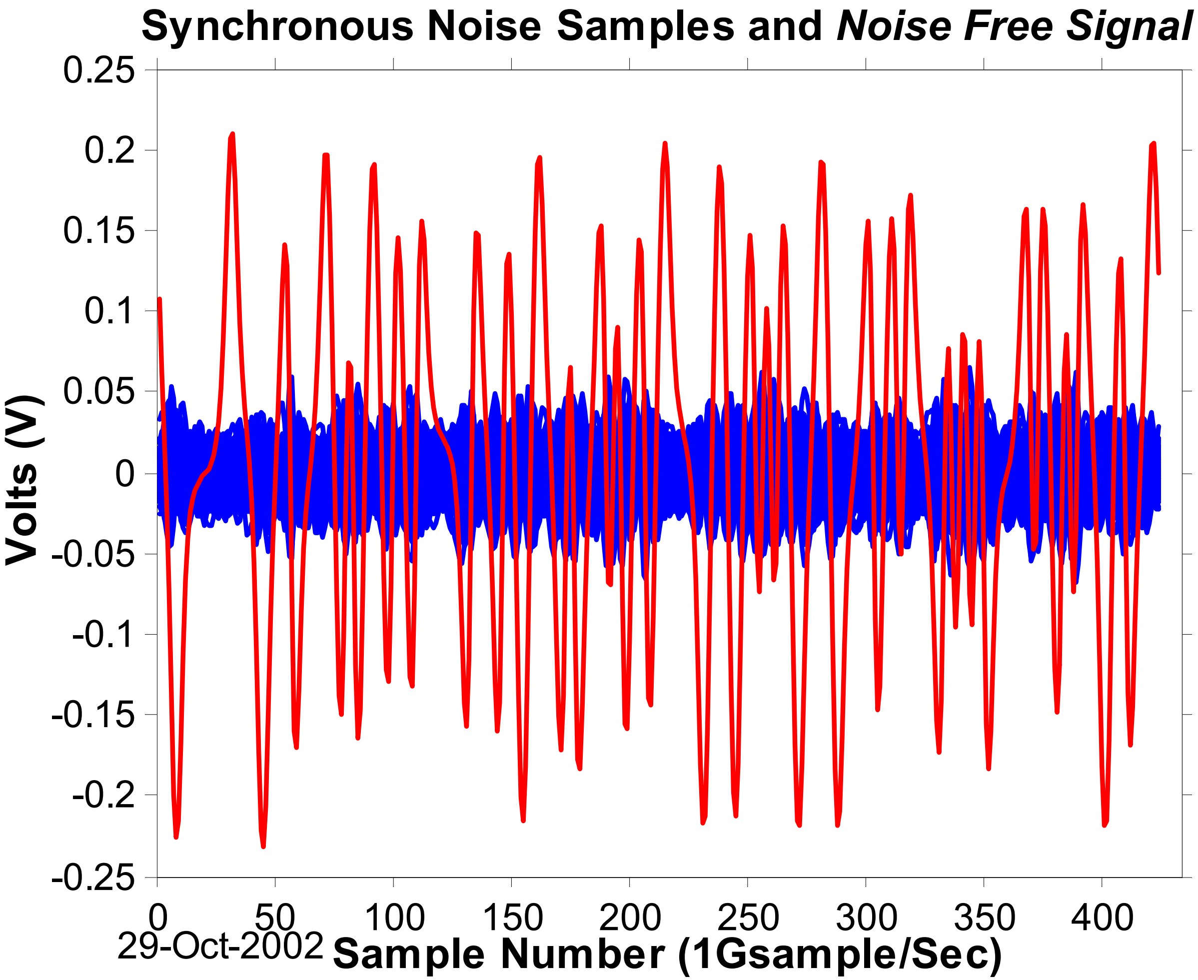
Volterra Model Block Diagram



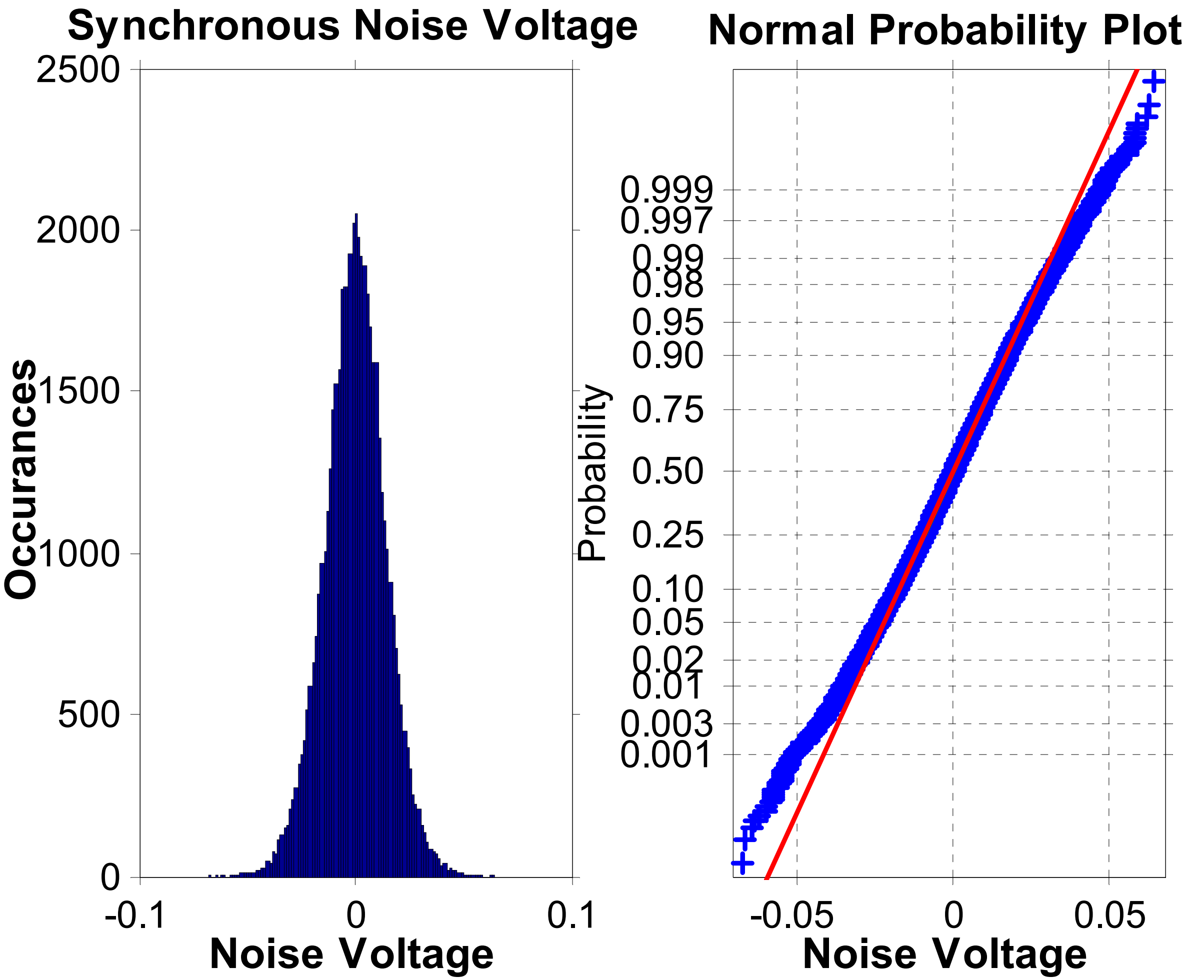
How Good is VM?



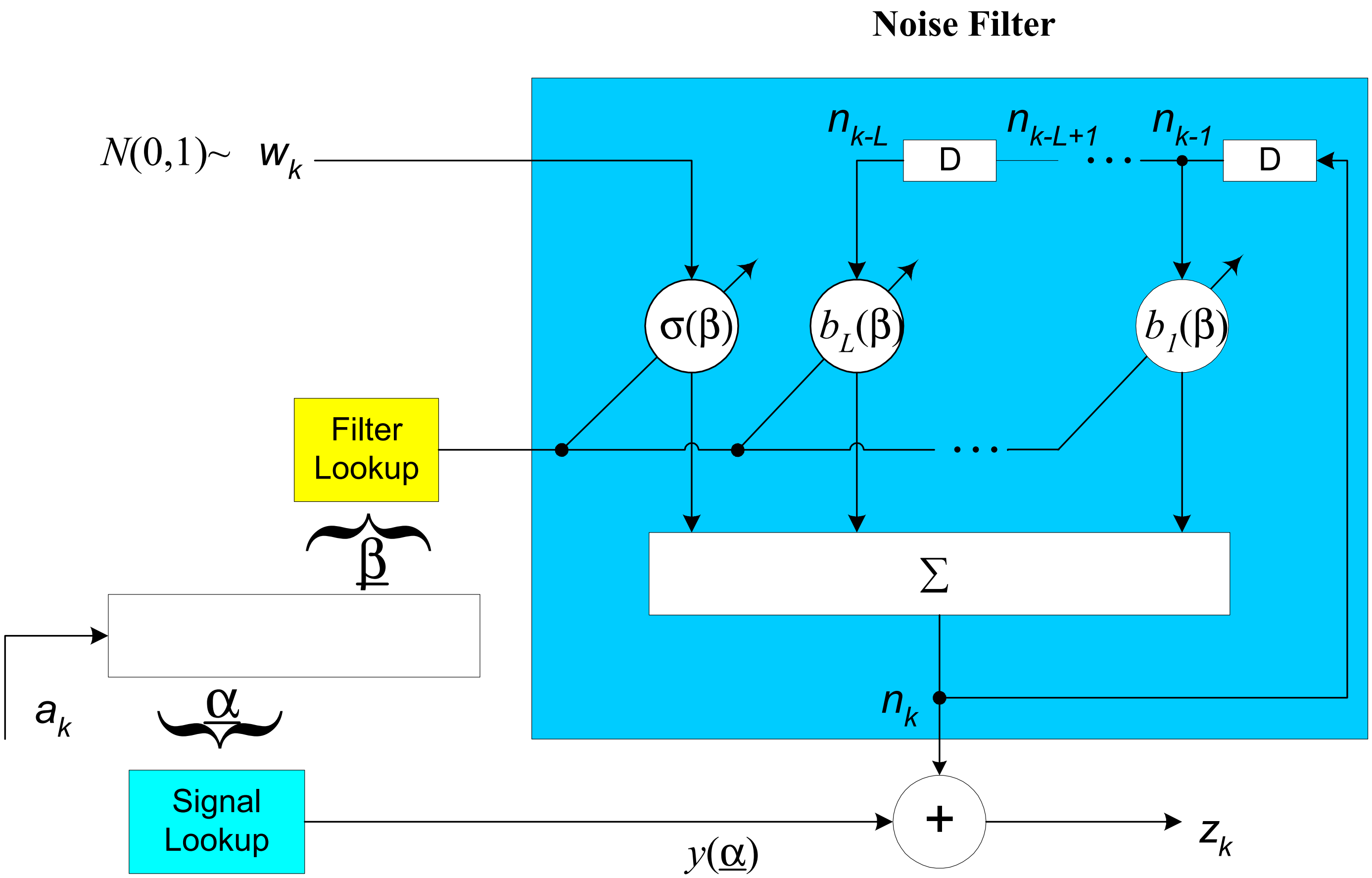
Longitudinal Media Noise and Signal



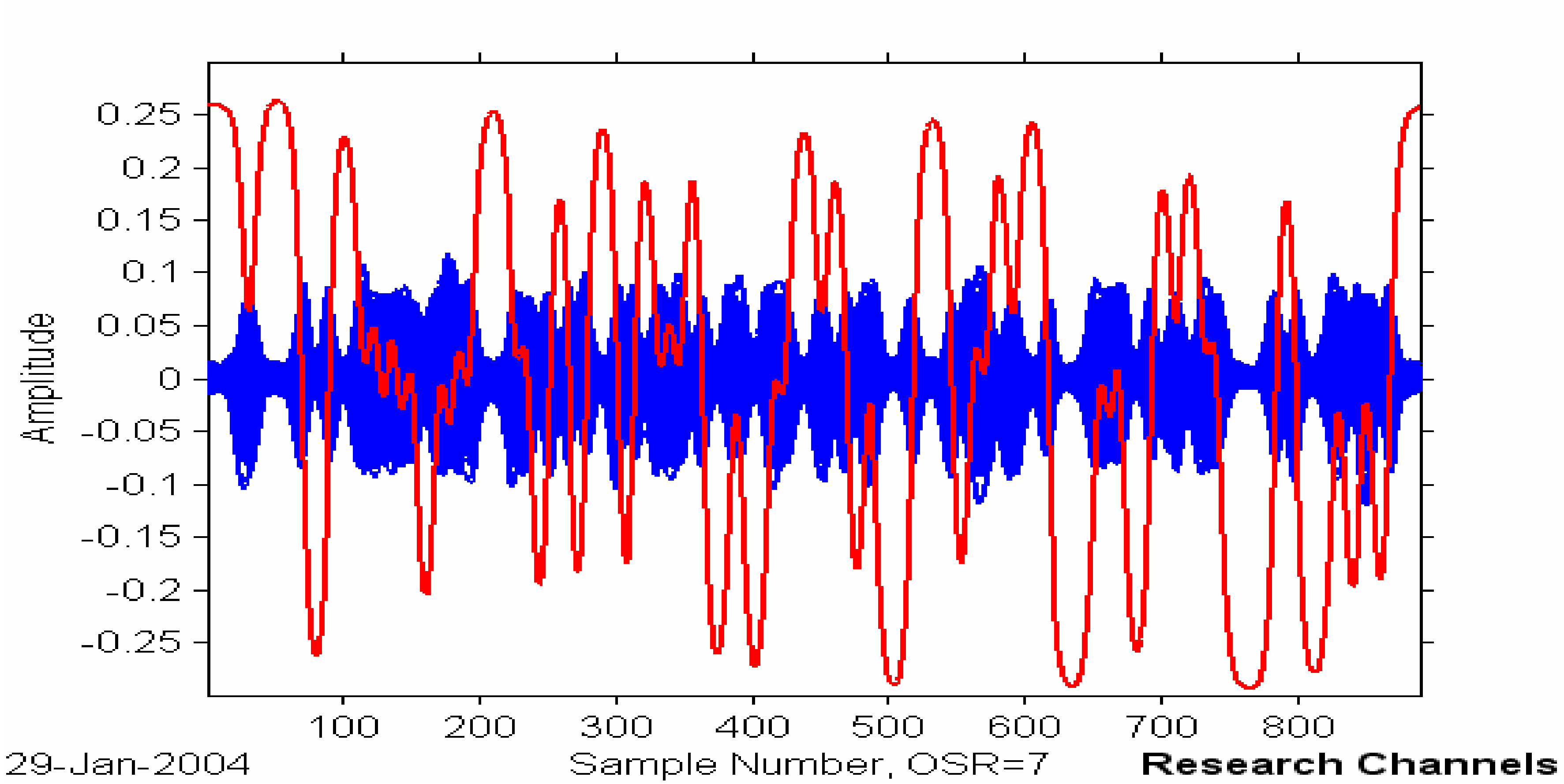
Longitudinal: Media Noise Voltage



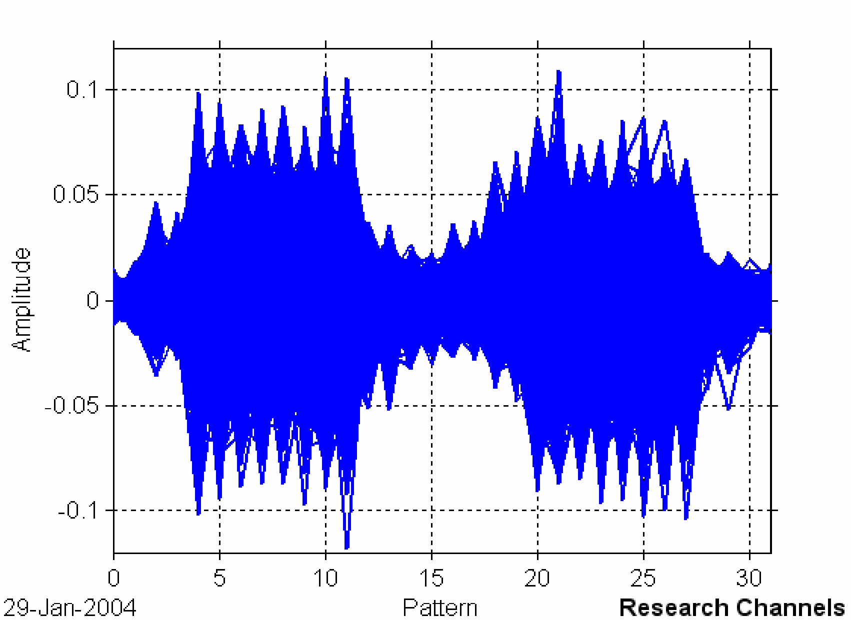
Data Dependent AR Model (AR)

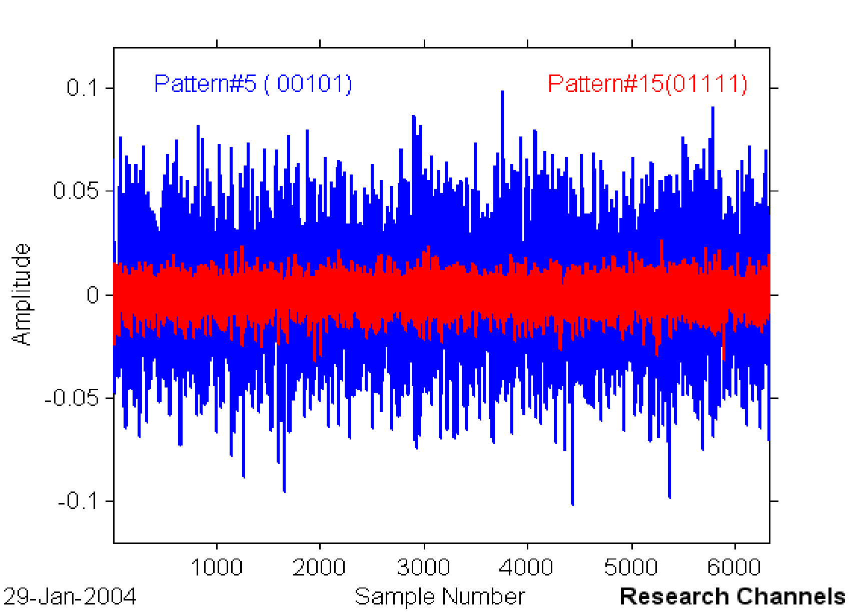


Perpendicular Signal and Media Noise

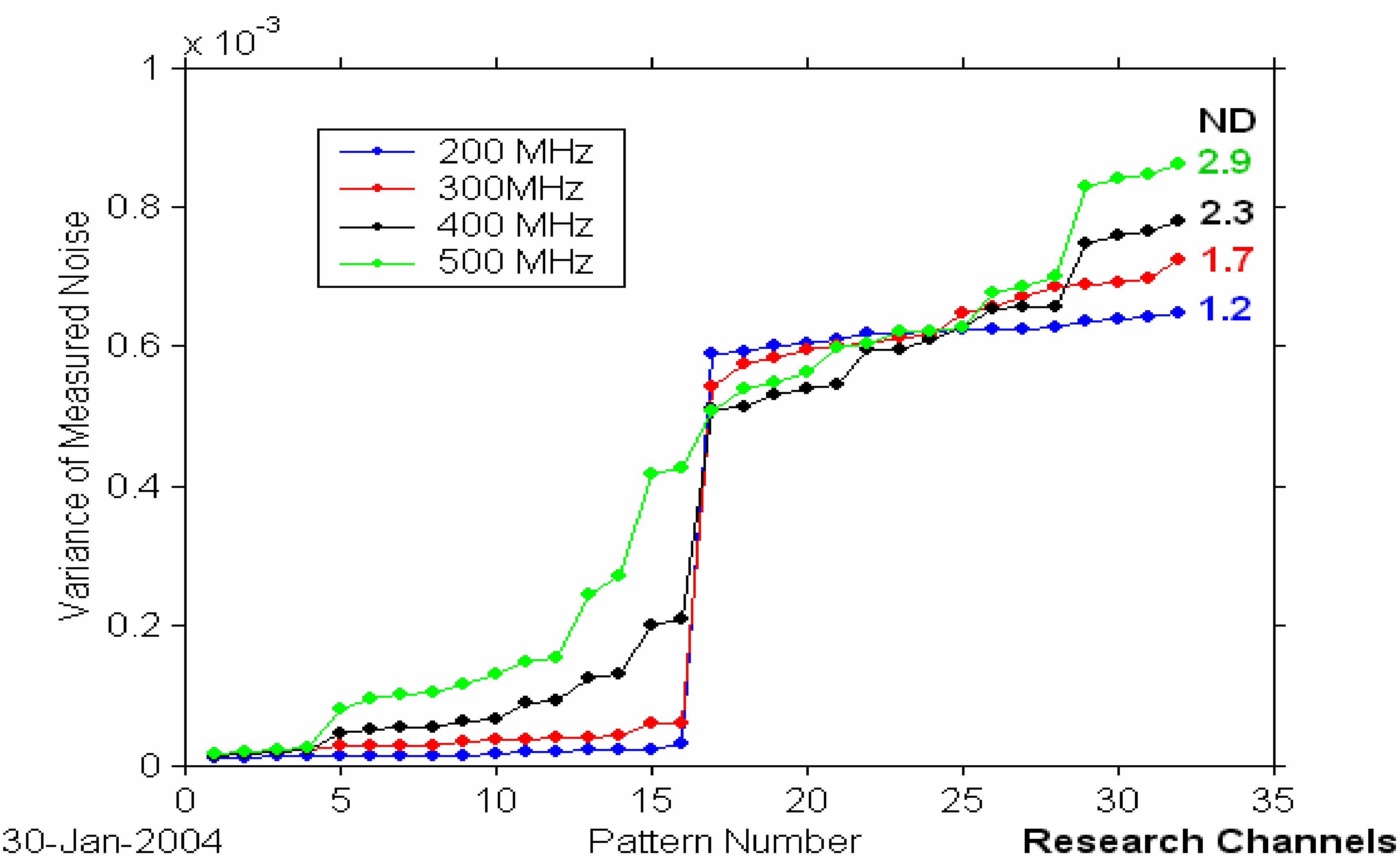


Media Noise 400 MHz

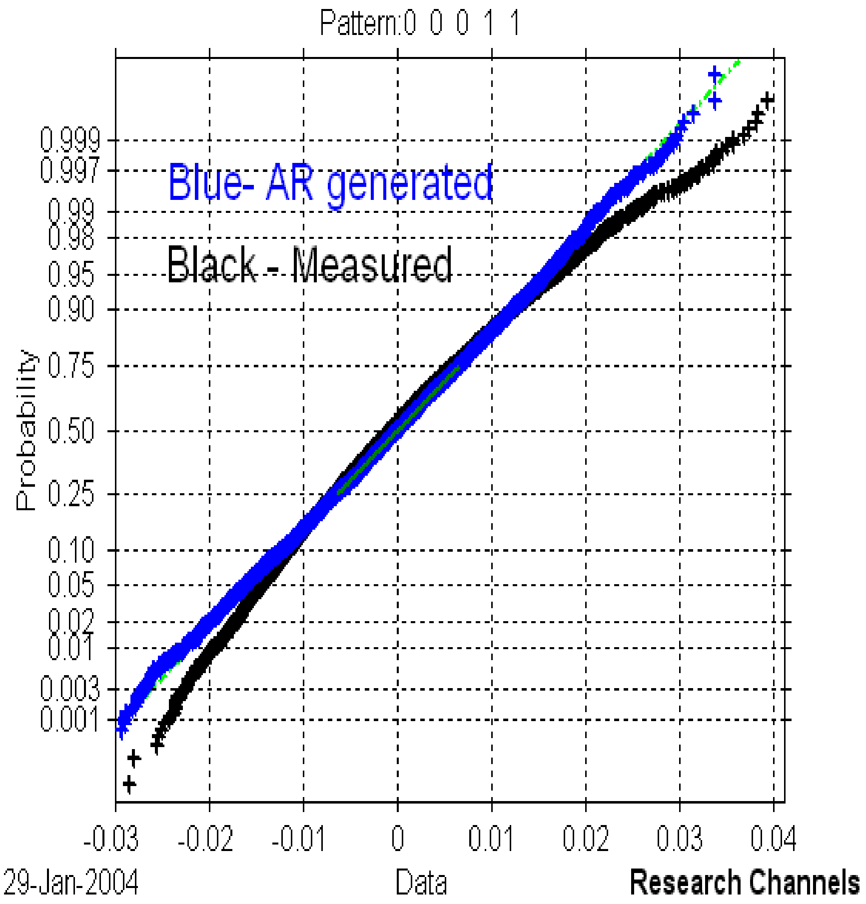




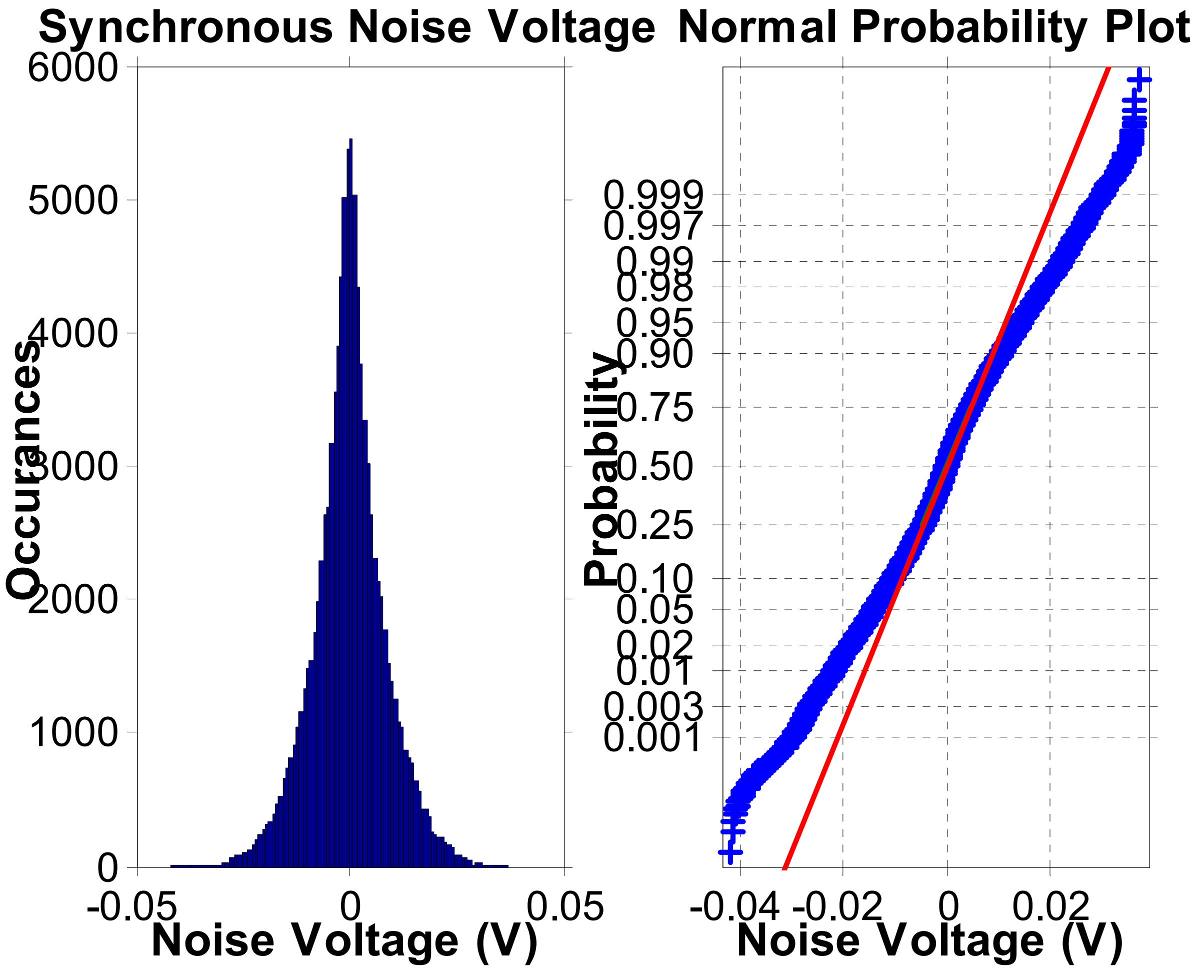
As ND Increases Noise Variance vs. Pattern



AR noise generation
- not a great fit

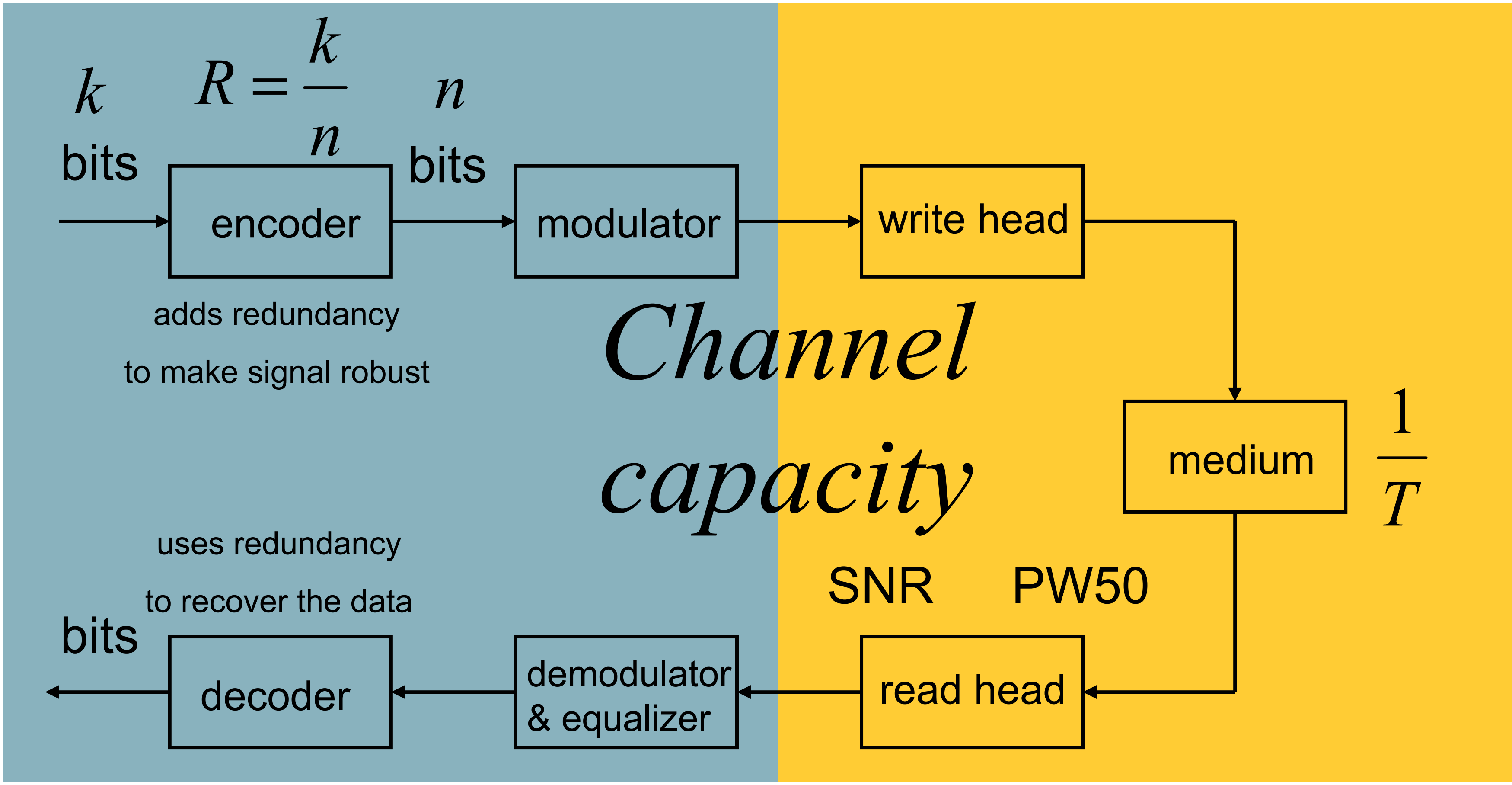


Perpendicular: Media Noise Voltage



- Noise distribution does not fit Gaussian!
- Noise distribution looks more like a laplacian.

Generic Storage Channel



k bits

$R = \frac{k}{n}$

n bits

encoder

modulator

write head

adds redundancy

to make signal robust

Channel capacity

medium

$\frac{1}{T}$

uses redundancy

to recover the data

SNR

PW50

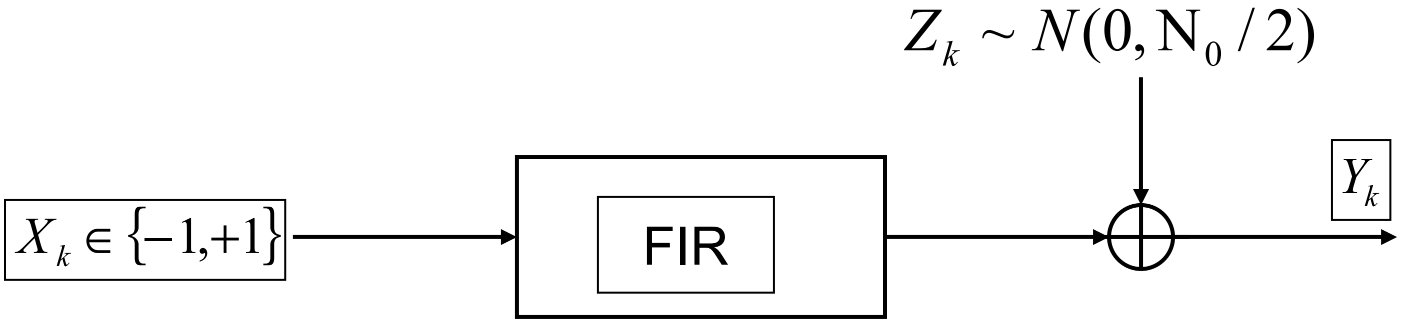
bits

decoder

demodulator & equalizer

read head

Brief Survey of the Literature



Bounds

Shamai at al. 1991, McLaughlin and Neuhoff 1993, and many others

– Direct Computation

Hirt 1988

– Markov-Chain Monte-Carlo Method

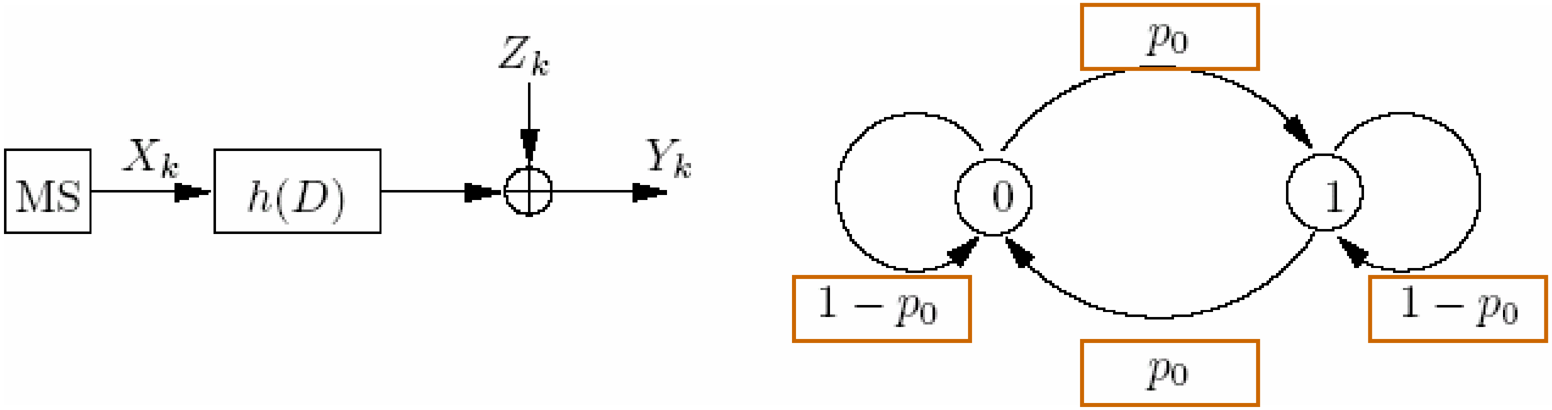
Arnold and Loeliger, Pfister et al., Sharma and Singh, Vontobel, all in 2001

– Shamai-Laroia conjecture

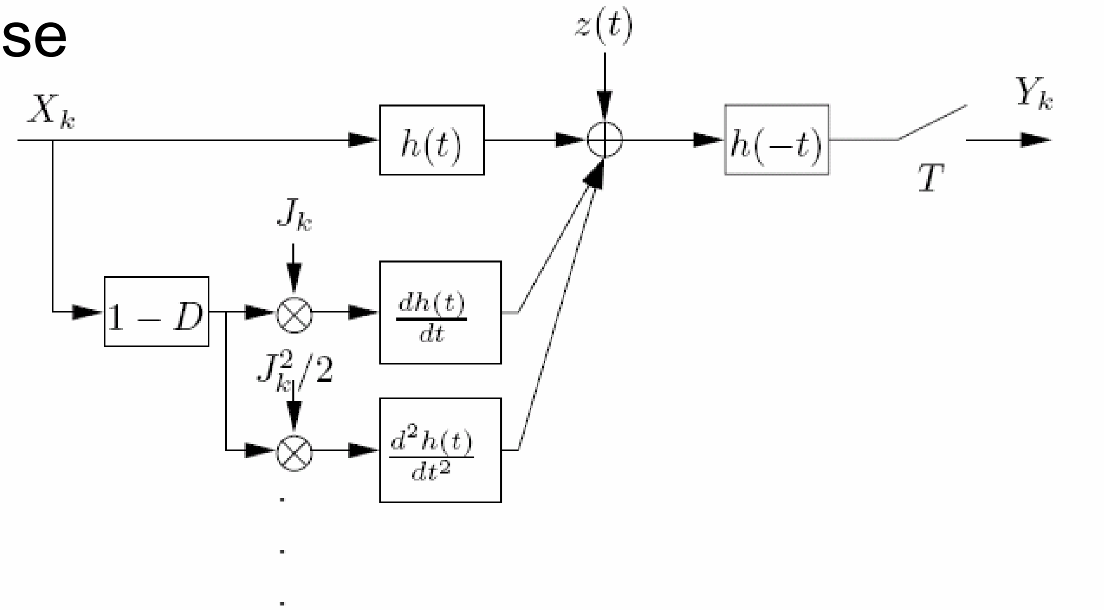
Shamai-Laroia 1996, Dholakia et al. 2000, Arnold and Eleftheriou 2002

Brief Survey of the Literature (cont.)

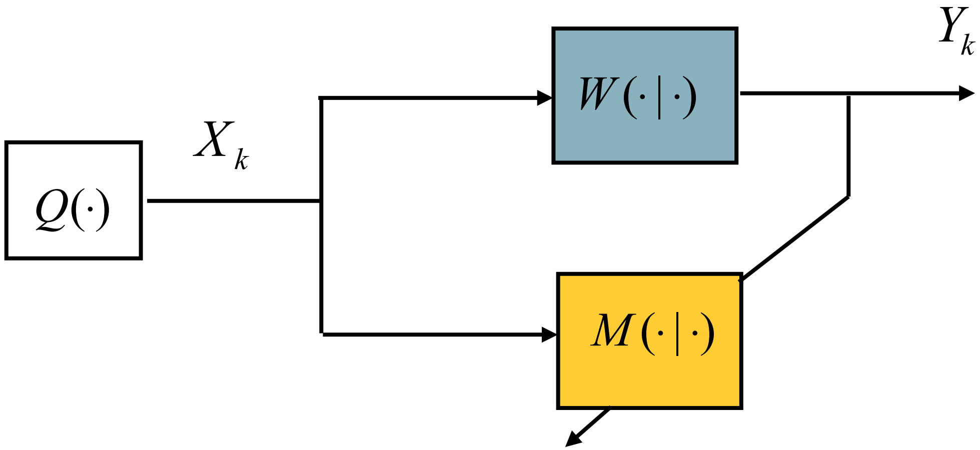
Kavcic 2001:



Zhang, Duman, and Kurtas 2002: Modeling signal-dependent noise



Set-up for Mismatch Lower Bounds



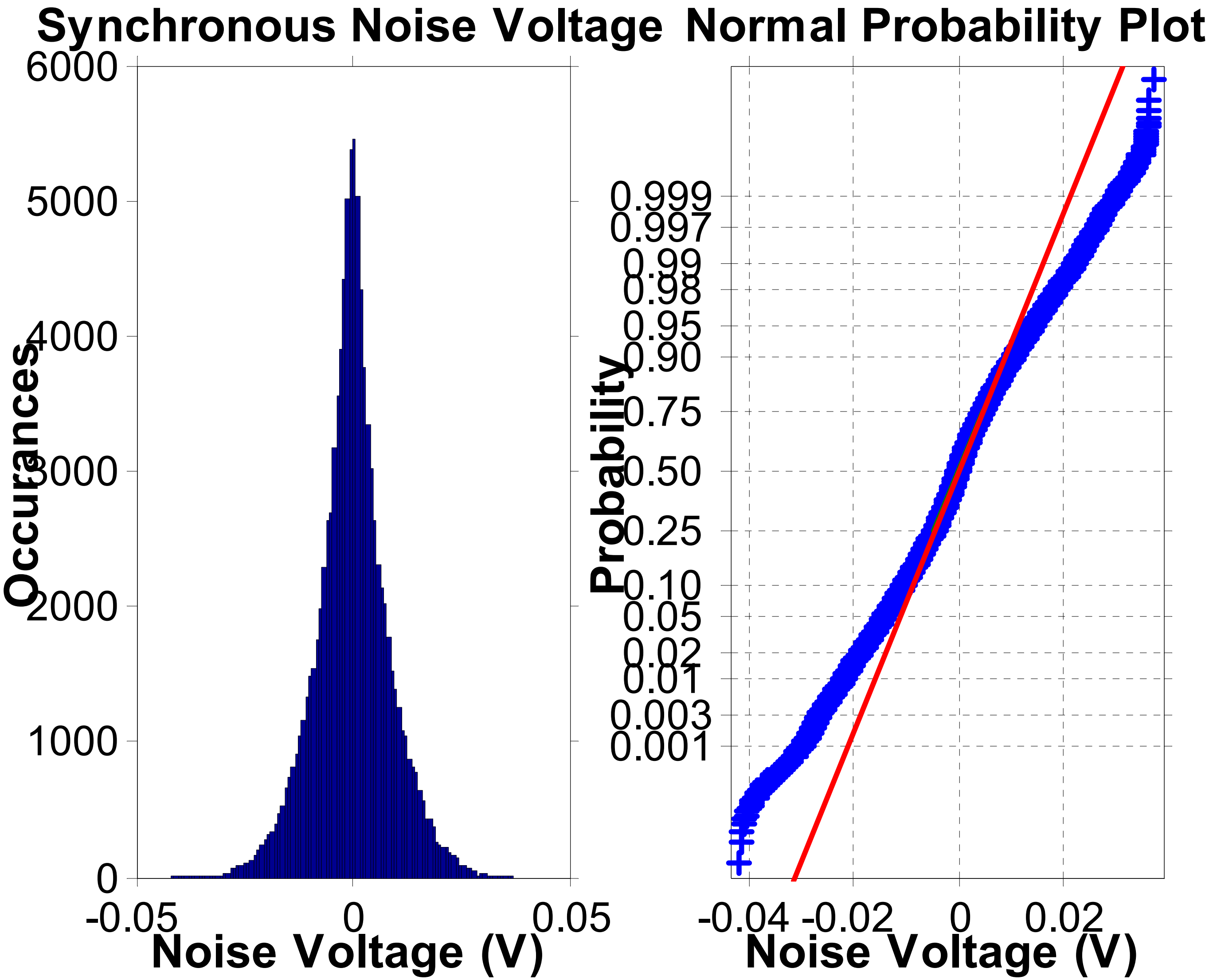
Mismatch lower bound:

$$I(Q,M) \equiv E_{QW}\left[\log(QM)(Y)\right] - E_{QW}\left[\log M(Y \mid X)\right]$$

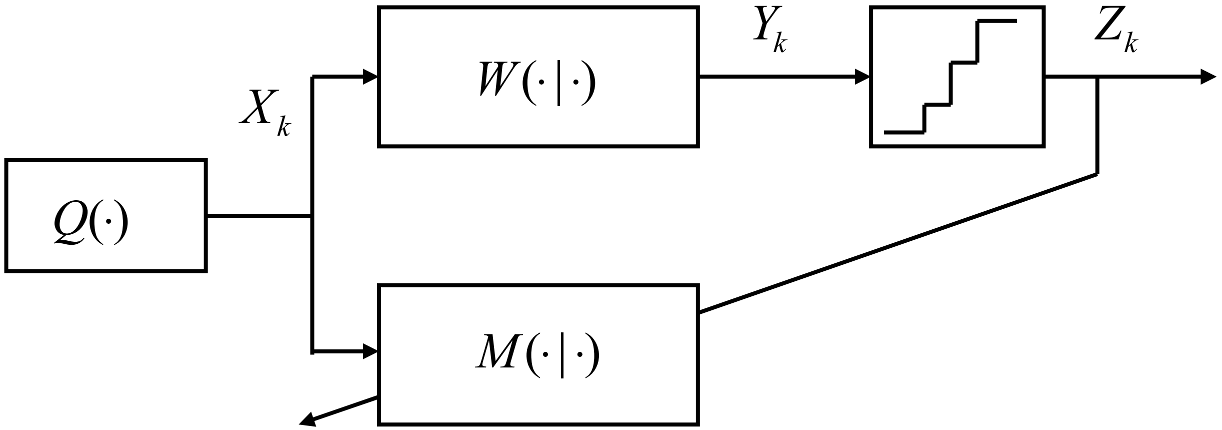
$$I(Q,W) \geq I(Q,M)$$

A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched Decoding Revisited: General Alphabets, Channels with Memory, and the Wide-Band Limit, " IEEE Trans. on Inform. Theory, pp. 2315-2328, Nov. 2000.

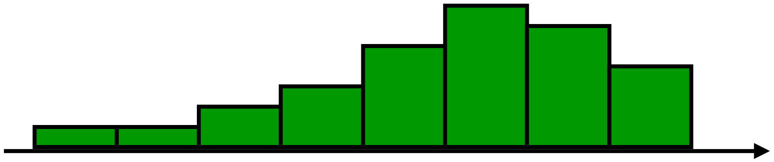
How to model the noise?



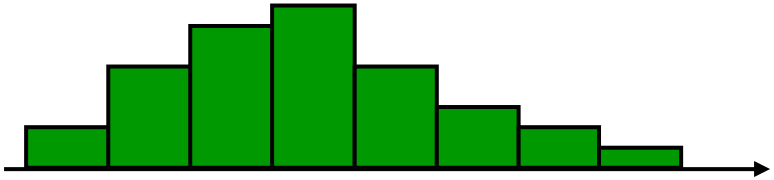
Just use histograms



$(0,0)$



$(0,1)$



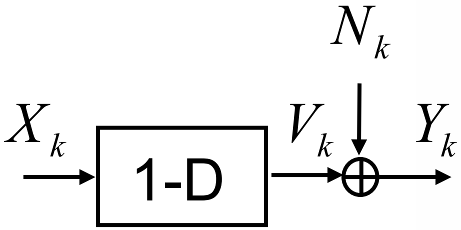
Validation on the ideal (1-D)-Channel

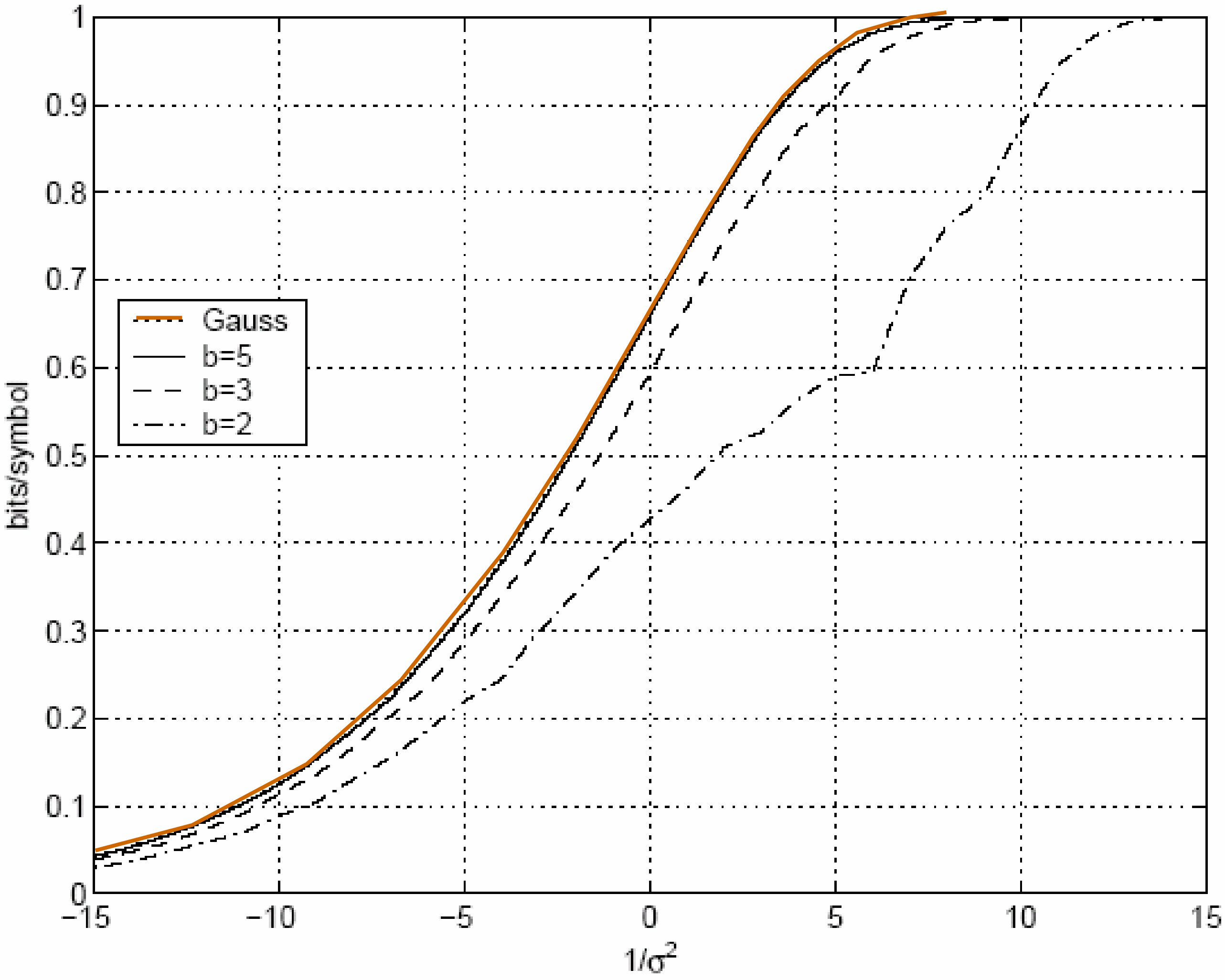
(1-D) channel

AWGN

L=1

b=1,3,5



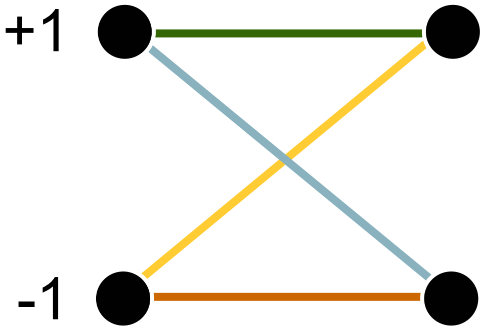


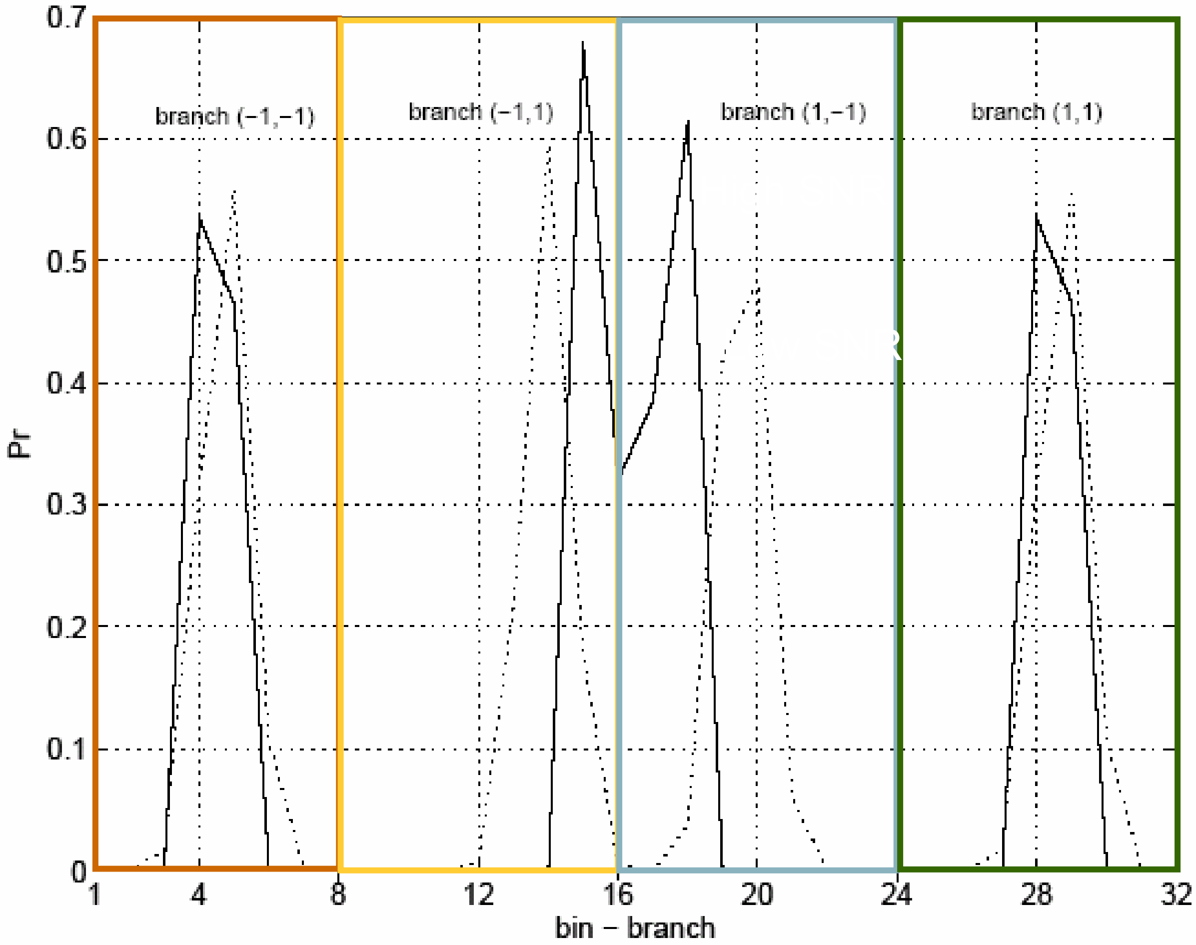
(1-D)-Channel – Histograms

(dotted for low SNR, solid for high)

L=1

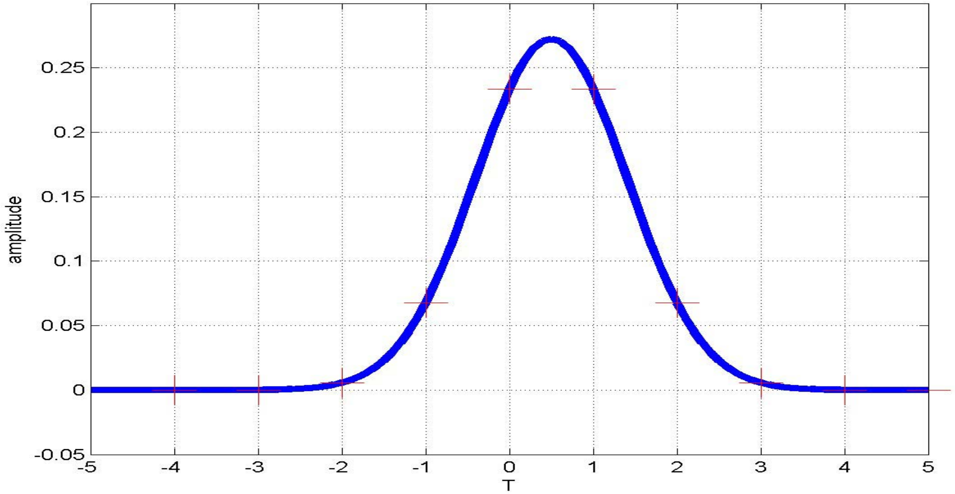
b=3





Model the channel as a GPR or FSM

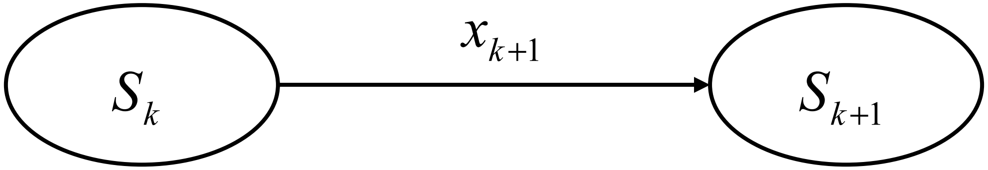
1. Capturing the influence of the neighbor bits by means of a state



Dc=2.0

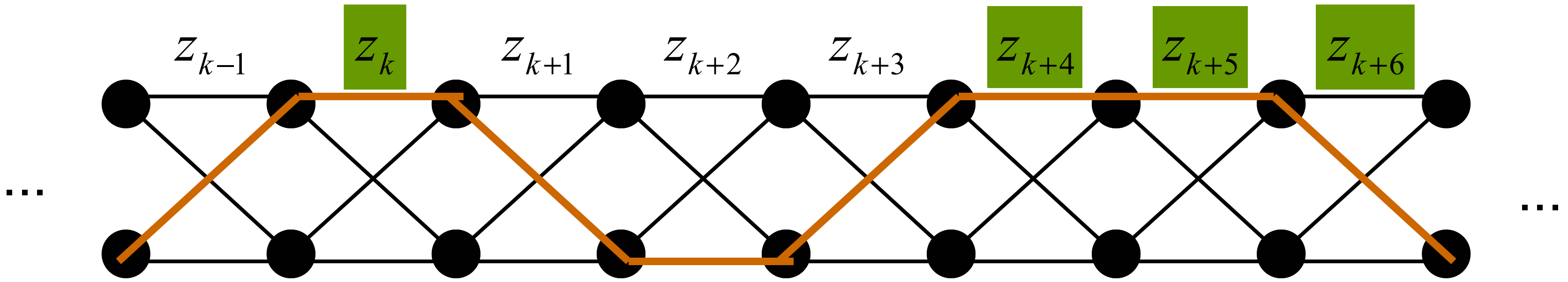
$$s_k \equiv (x_{k-m}, x_{k-(m-1)}, \ldots, x_k, x_{k+1}, \ldots, x_{k+m})$$

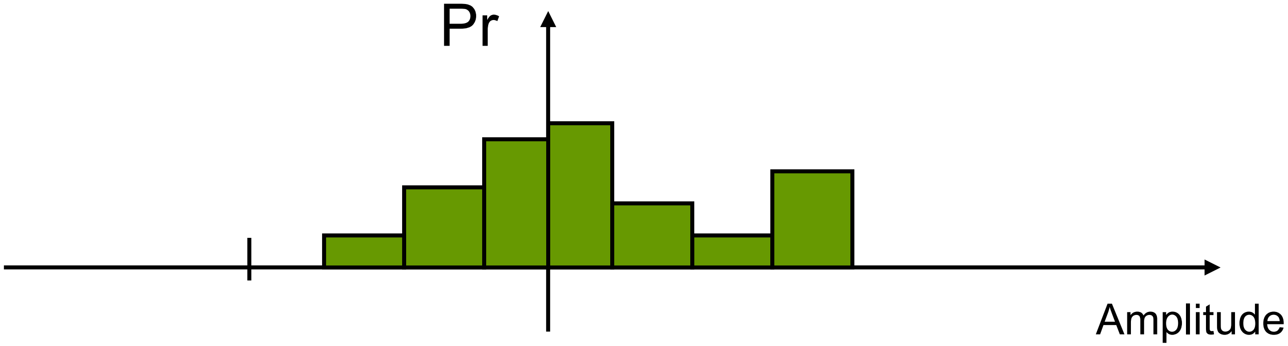
2. Sorting over time by means of a trellis of size $|S| = 2^{2m+1} = 2^L$



Use Quantized Histograms per Branch for Noise

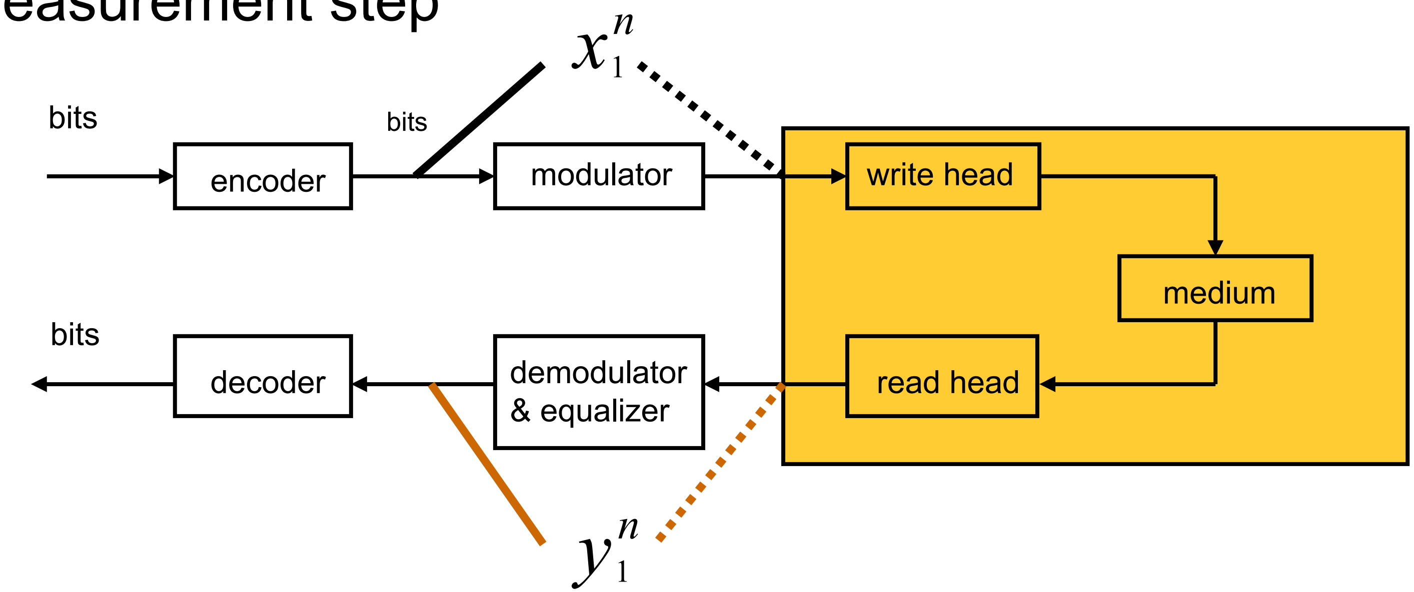
3. Representation of the noise pdfs per branch:





Big Picture

1. Measurement step



2. Quantization step

$$y_1^n \overset{Q(b)}{\mapsto} z_1^n$$

3. Computation step

$$C \geq I_{\mathrm{LB}} \leftarrow \frac{1}{n} \log \frac{M(Z_1^n \mid X_1^n)}{(QM)(Z_1^n)}$$

How complex is it?

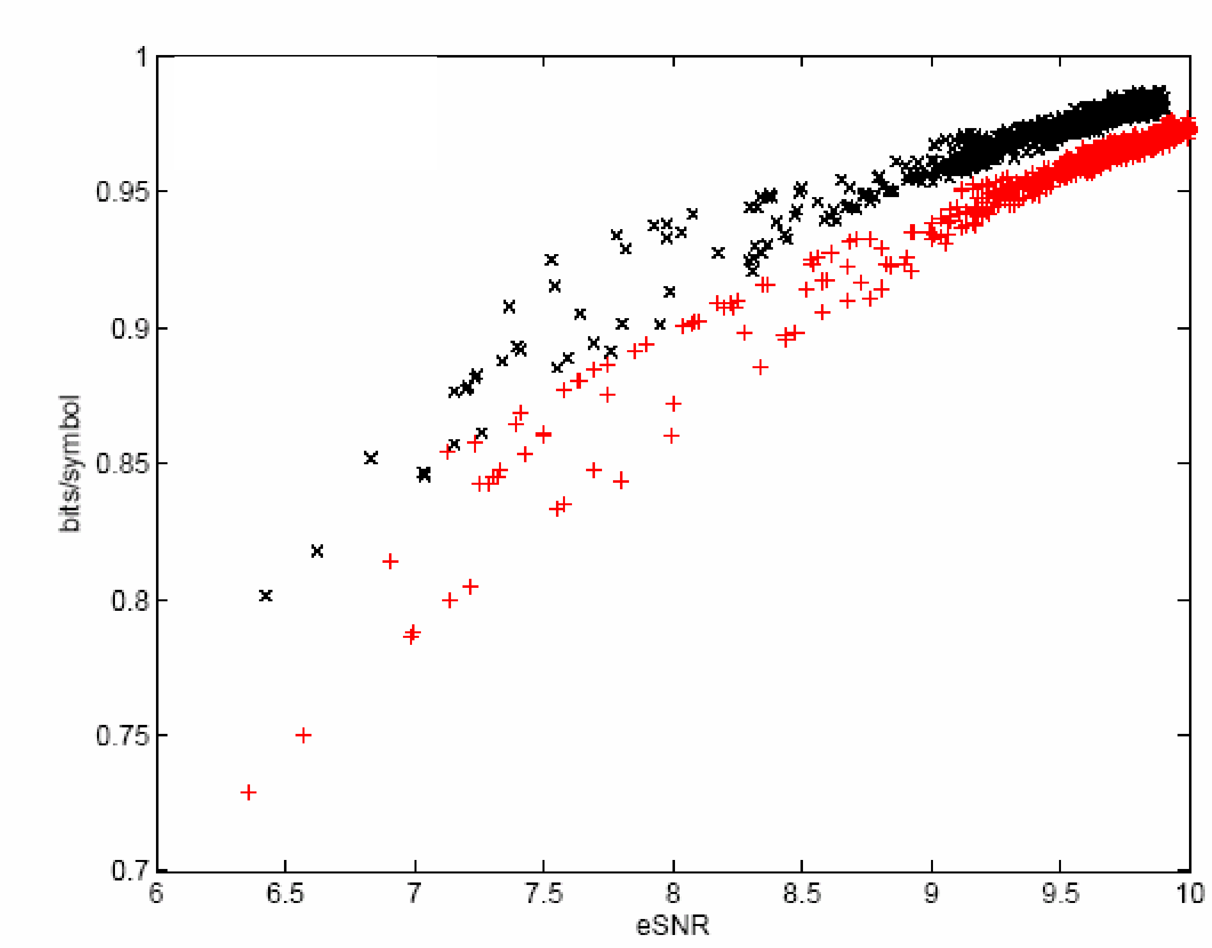
Memory	$2 \cdot n$
# of comparisons per trellis section	b
# of multiplications per branch	2

Example: $n = 10^6, |\mathrm{S}| = 32, b = 6$ < 20 s

On a P4, 2.5 GHz, 512 MB

Orders of magnitude faster than performance evaluation of Turbo-Codes.

Results with Real Waveforms

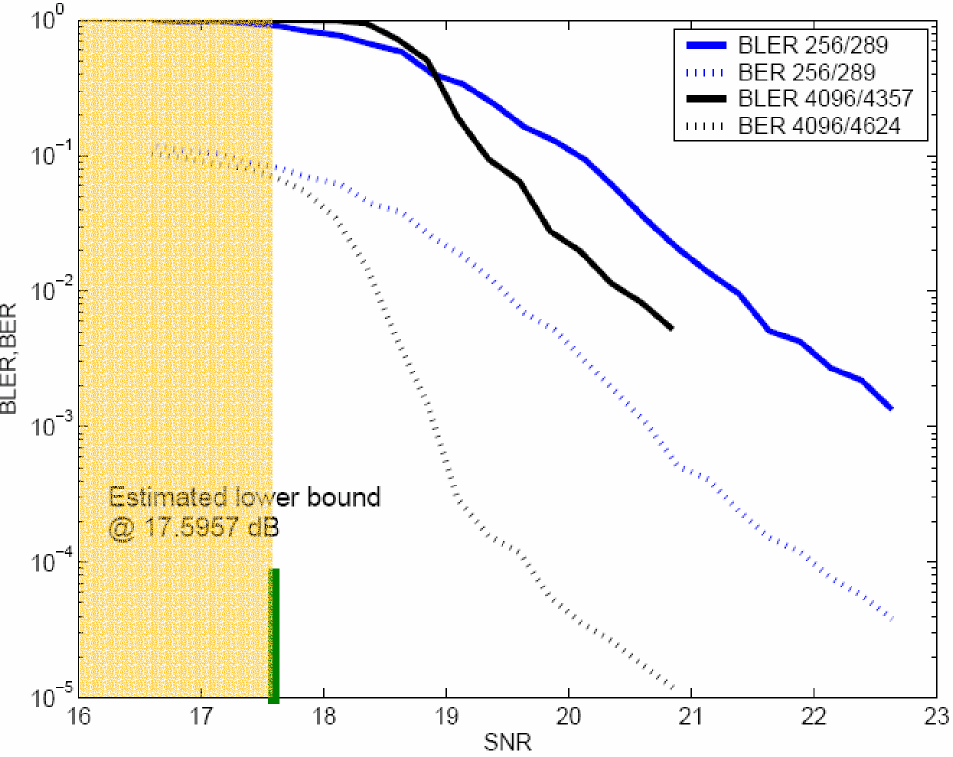


Waveform: perpendicular n=44276

Model: L=4, b=6

Compare with Turbo Codes

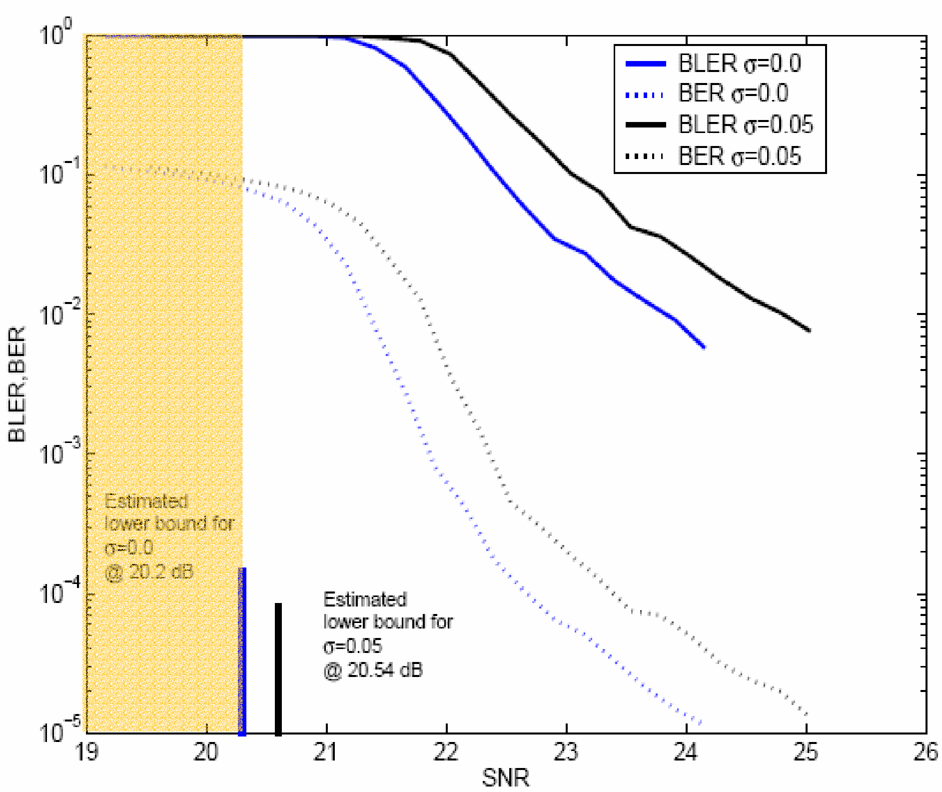
Dc=2.0



[k=256,n=289]

[k=4096,n=4624]

Dc=2.4

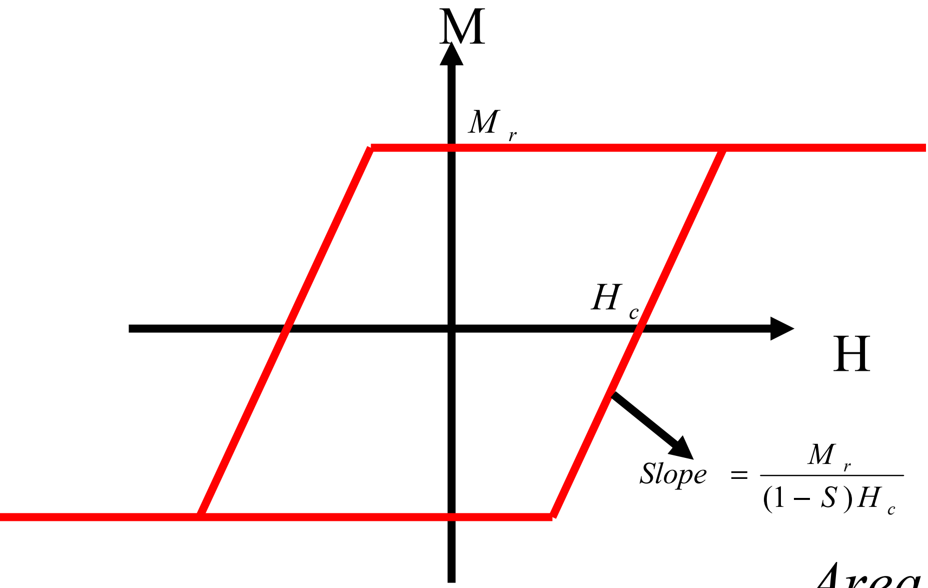


$\sigma = 0.0$

$\sigma = 0.05$

Heat Assisted Magnetic Recording

Conventional system

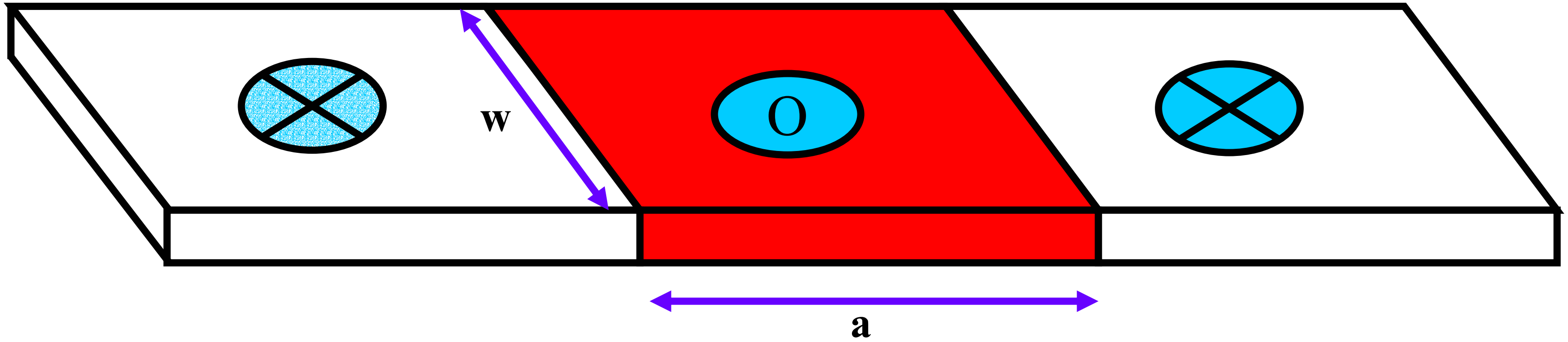


Store bit by

- Applying external H > Hc (magnetize up)

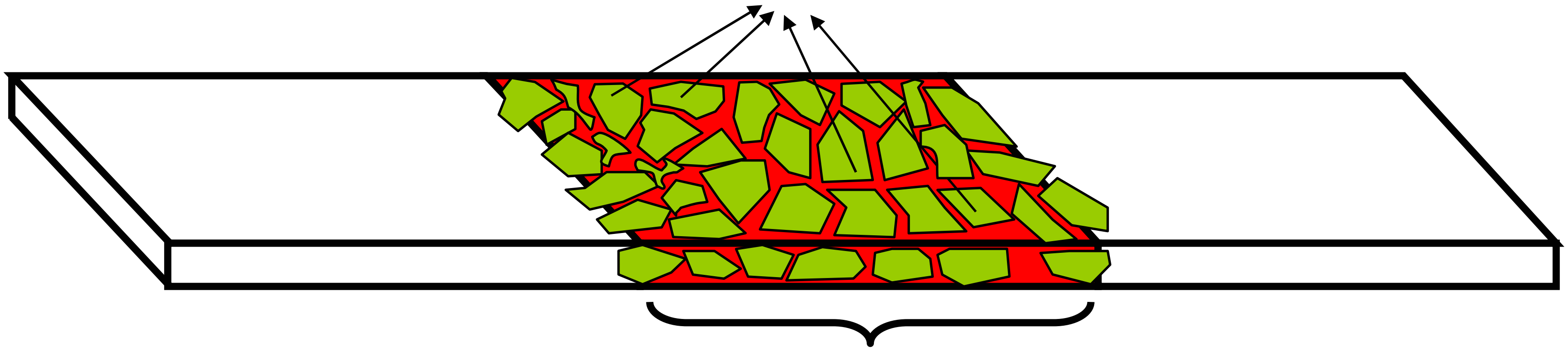
- Applying external H < -Hc (magnetize down)

$$AreaDensity[bits/inch^2] = \frac{1[bit]}{a[inch]w[inch]}$$



In reality there are tiny grains

Tiny grains with finite volume V and anisotropy coefficient Ku



Minimum grain number fixed to preserve SNR

Maximum value limited by maximum writer field

Super paramagnetic limit :

$$\frac{K_u V}{k_B T} \geq M$$

Boltzmann Constant

Temperature

Large number, like 60

Maximum Attainable Areal Density is Limited

A solution : HAMR

Heat Assisted Magnetic Recording

CENTRAL CONCEPT: Use interplay of temperature and field gradients to perform high density thermomagnetic recording on <u>very</u> high coercivity (thermally stable) media.

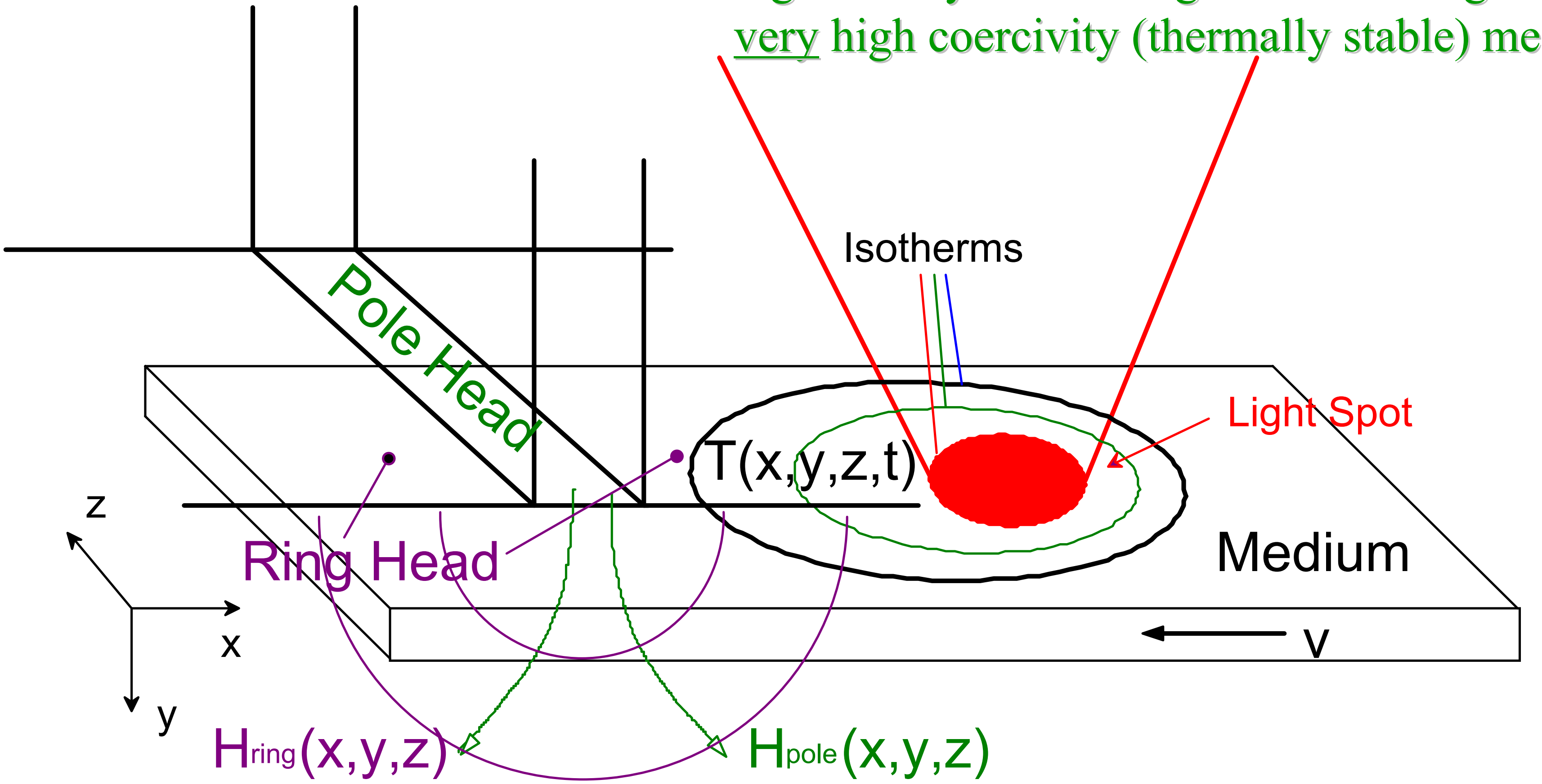
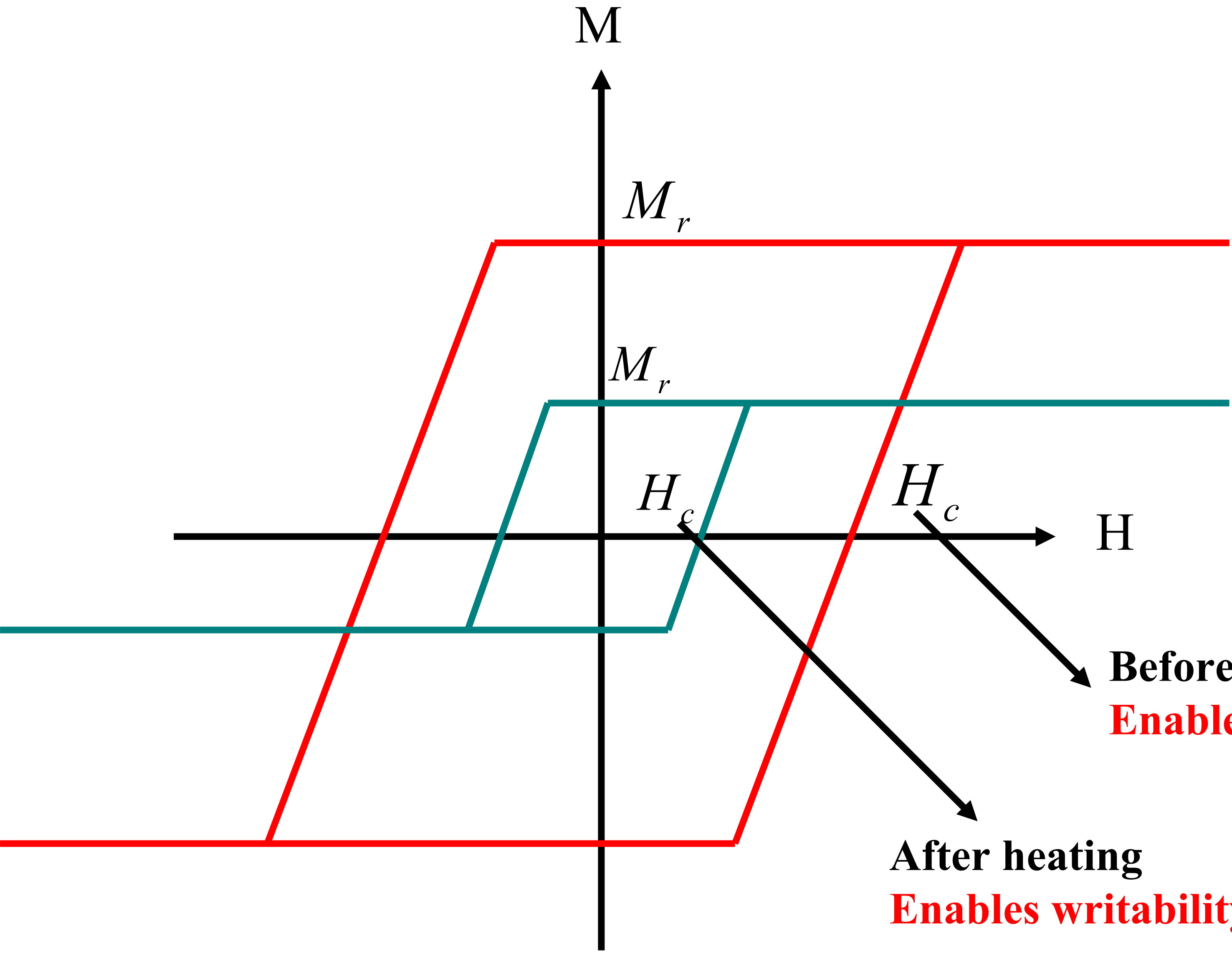


Image courtesy of T. McDaniel, Seagate Technology

Idea behind HAMR



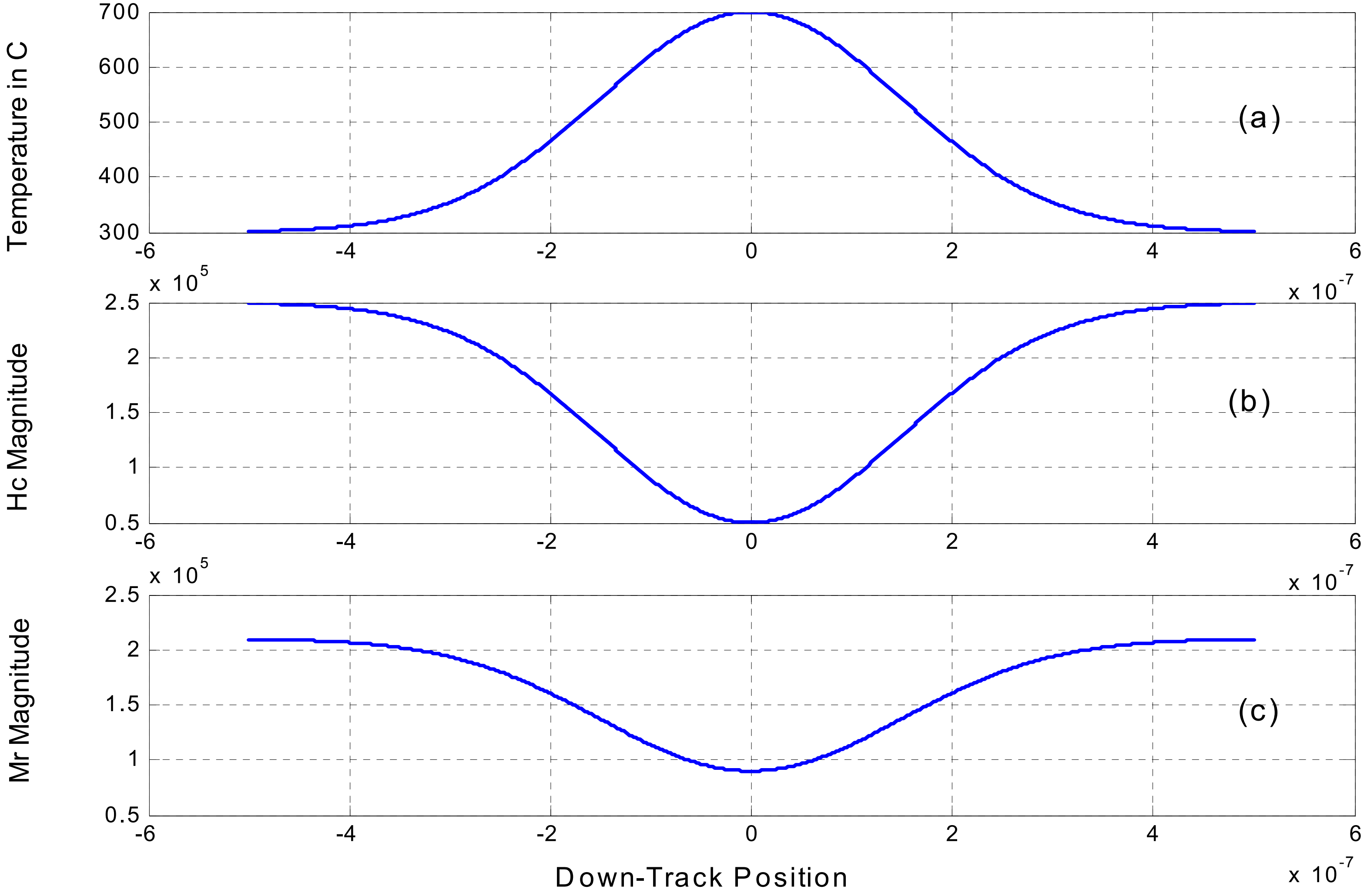
Before heading

Enables thermal stability

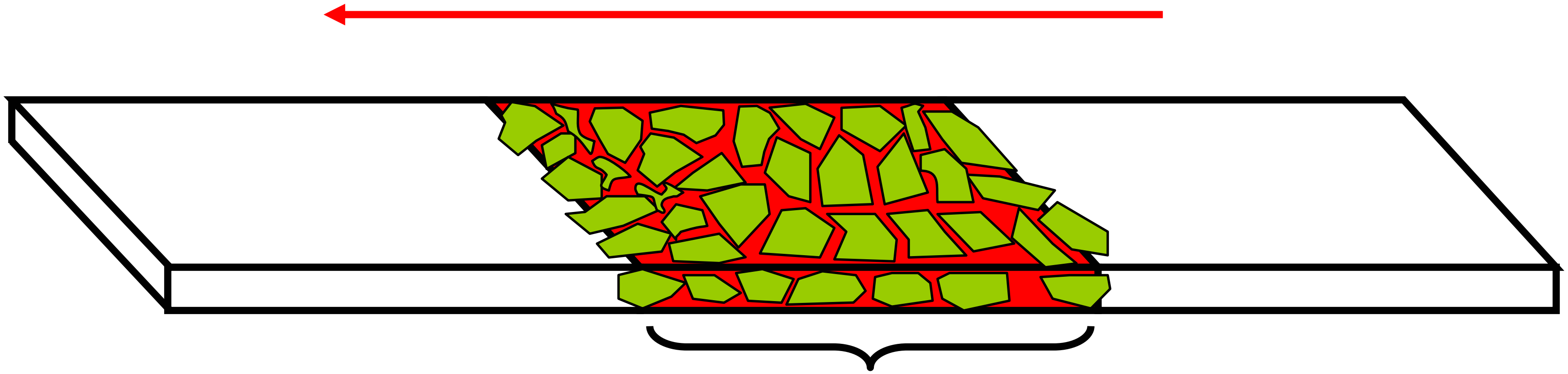
After heating

Enables writability

An illustrative example



Why should we have any issue?



Assume speed to be 25m/s

25 nm equivalent to 1ns

-Increase the heat of the medium by 300K – 400K
-Write the bit of information
-Cool the medium

!!! Complete everything in 1ns !!!

Major issues in HAMR

- Magnetic issues
  - Availability of desirable Ku(T) for particles
  - Controllability of the spatial orientation of Ku(T)

  to minimize the temperature difference

- Thermal issues
  - Lubricant and overcoat stability
  - Air bearing flying stability
  - Media-Head air bearing surface smoothness

- Optical issues -- Confining light in a very small spot. Some methods
  - Solid Immersion Lenses (SIL)
  - Apertures
  - Antennas
  - Waveguides

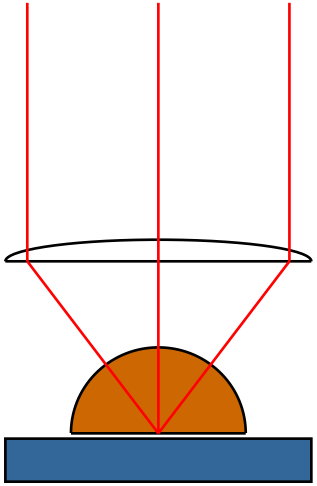
Each have their own advantages and disadvantages (*)

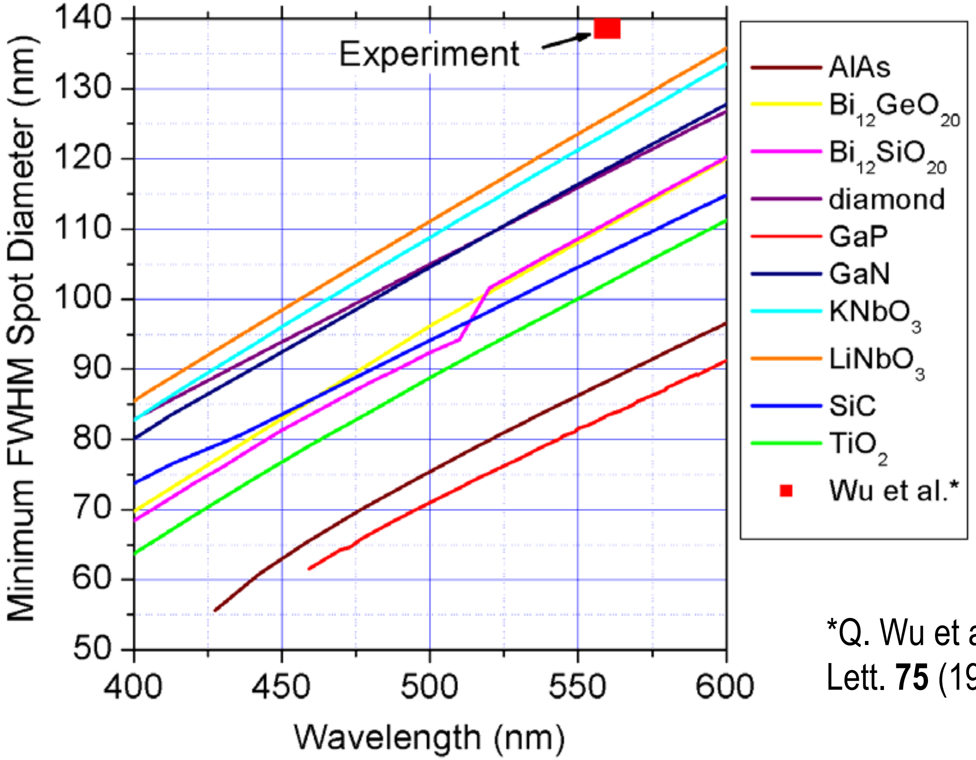
(*) Challener et al,
Jpn. J. Appl. Phys. 42, (2003) 981

Solid Immersion Lenses (SIL)

Theoretical spot size

$$\text{FWHM diameter:} \quad d \geq \frac{0.51 \cdot \lambda}{n}$$





*Q. Wu et al., Appl. Phys. Lett. **75** (1999) 4064.

Image courtesy of T. McDaniel, Seagate Technology

Circular Aperture in Ideal Conductor

$$T=\left(\frac{64\pi^2}{27}\right)\cdot\left(\frac{d}{\lambda}\right)^4 \equiv \frac{\left.{}^{transmitted\ power}\!/_{area}\right]_{hole}}{\left.{}^{incident\ power}\!/_{area}\right]_{plane\ wave}}$$

Hans Bethe,
Phys. Rev.
(1944) 163.

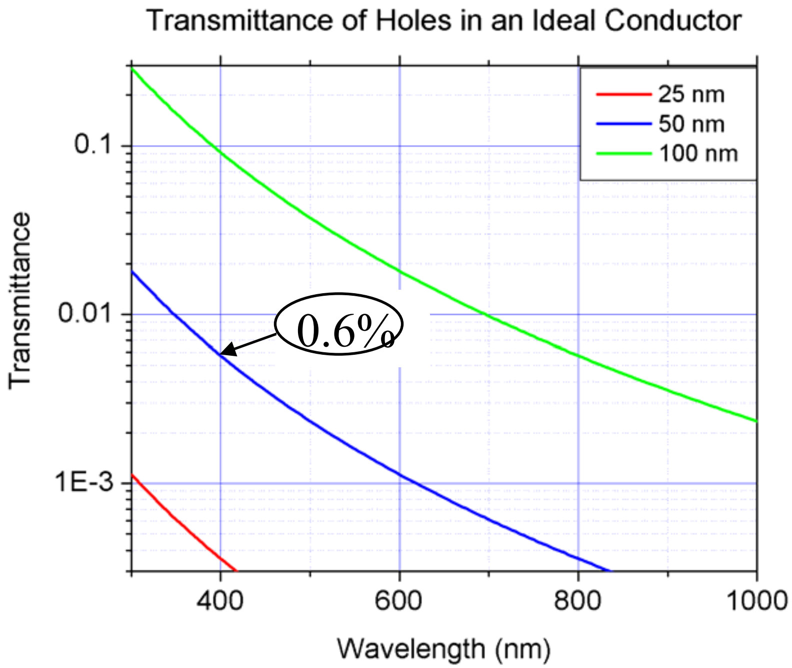
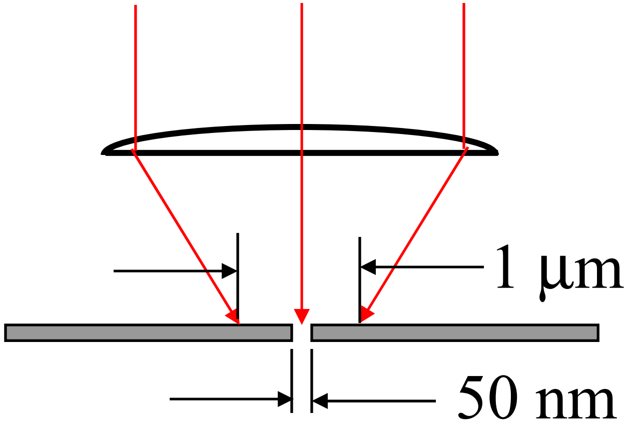


Image courtesy of T. McDaniel, Seagate Technology



$$T\cdot\frac{A_{hole}}{A_{spot}}=1.5\times10^{-5}$$

How can we get a useful channel model?

$$M = F_{loop}(H_a)$$

$$H_a = H_h + H_d$$

$$H_x = \frac{H_0}{\pi}\left[\tan^{-1}\left(\frac{x+g/2}{y}\right) - \tan^{-1}\left(\frac{x-g/2}{y}\right)\right]$$

$$H_y = -\frac{H_0}{2\pi}\ln\frac{(x+g/2)^2+y^2}{(x-g/2)^2+y^2}$$

Karlqvist Head Field Approximation

$$H_d = \frac{\partial M}{\partial x} * H_x^{step}(x)$$

Function of M

Very difficult to solve loop equation

Approximation – Thermal Williams Comstock Model

Williams Comstock equation

$$\frac{\partial M}{\partial x} = \frac{\partial M}{\partial H}\left[\frac{\partial H_h}{\partial x} + \frac{\partial H_d}{\partial x}\right]$$

Thermal Williams Comstock equation

Temperature profile

$$\frac{\partial M}{\partial x} = \frac{\partial M}{\partial H}\left[\frac{\partial H_h}{\partial x} + \frac{\partial H_d}{\partial x} - \frac{H_h + H_d}{H_c}\frac{\partial H_c}{\partial T}\frac{\partial T}{\partial x}\right]$$

Temperature dependent

Using the equations

Find transition location x_0 which satisfies

$$H_h(x_0) + H_d(x_0) = H_c(x_0)$$

Find a-parameter as a function of x_0

$$a = Function(x_0, \ldots, \ldots, \ldots)$$

Isolated transition response

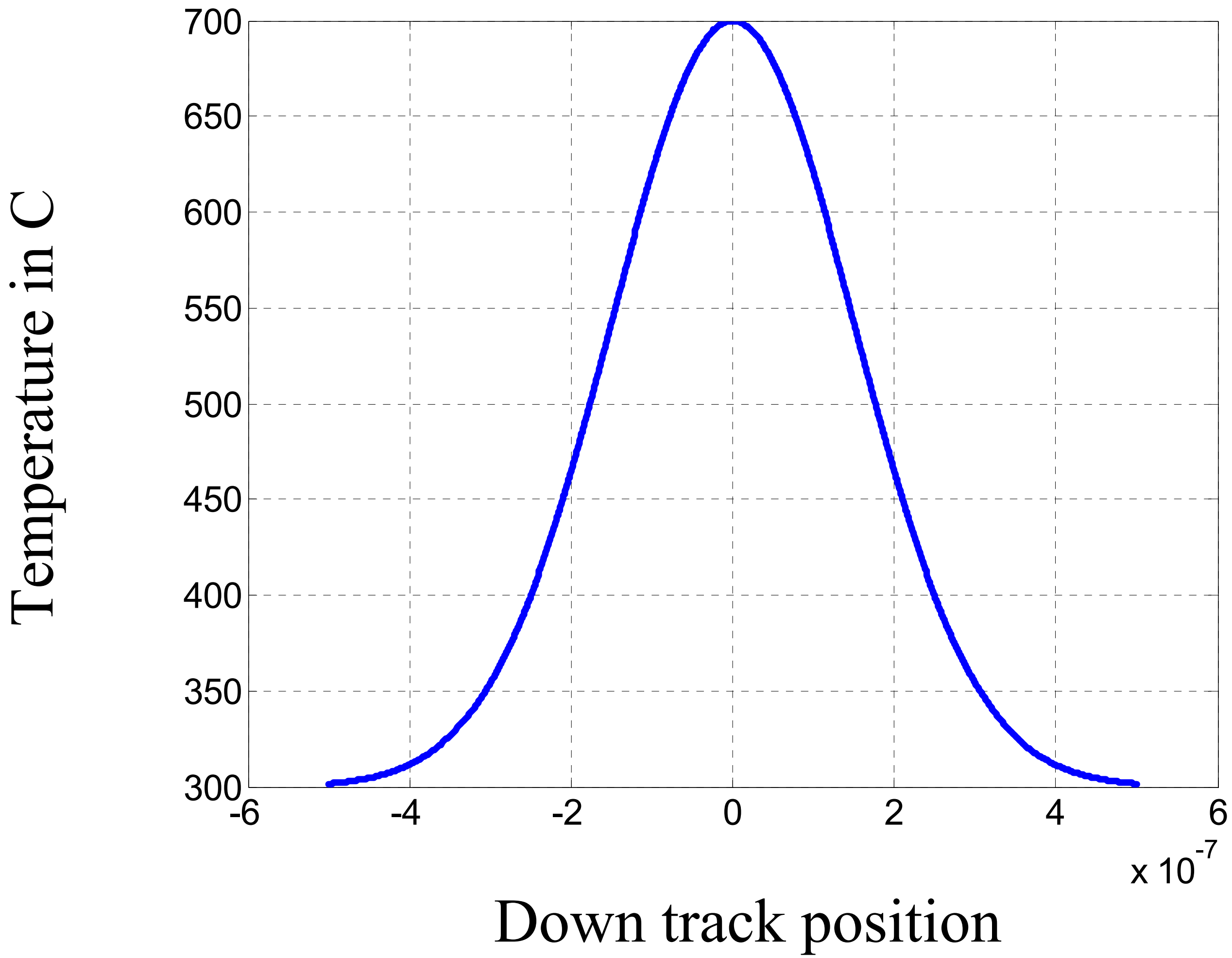
Readback voltage from isolated transition

$$V_{GMR}(x) = CM_r(x_0)\delta\left(\tan^{-1}\frac{x+g/2}{a(x_0)+d} - \tan^{-1}\frac{x-g/2}{a(x_0)+d}\right)$$

Temperature dependent

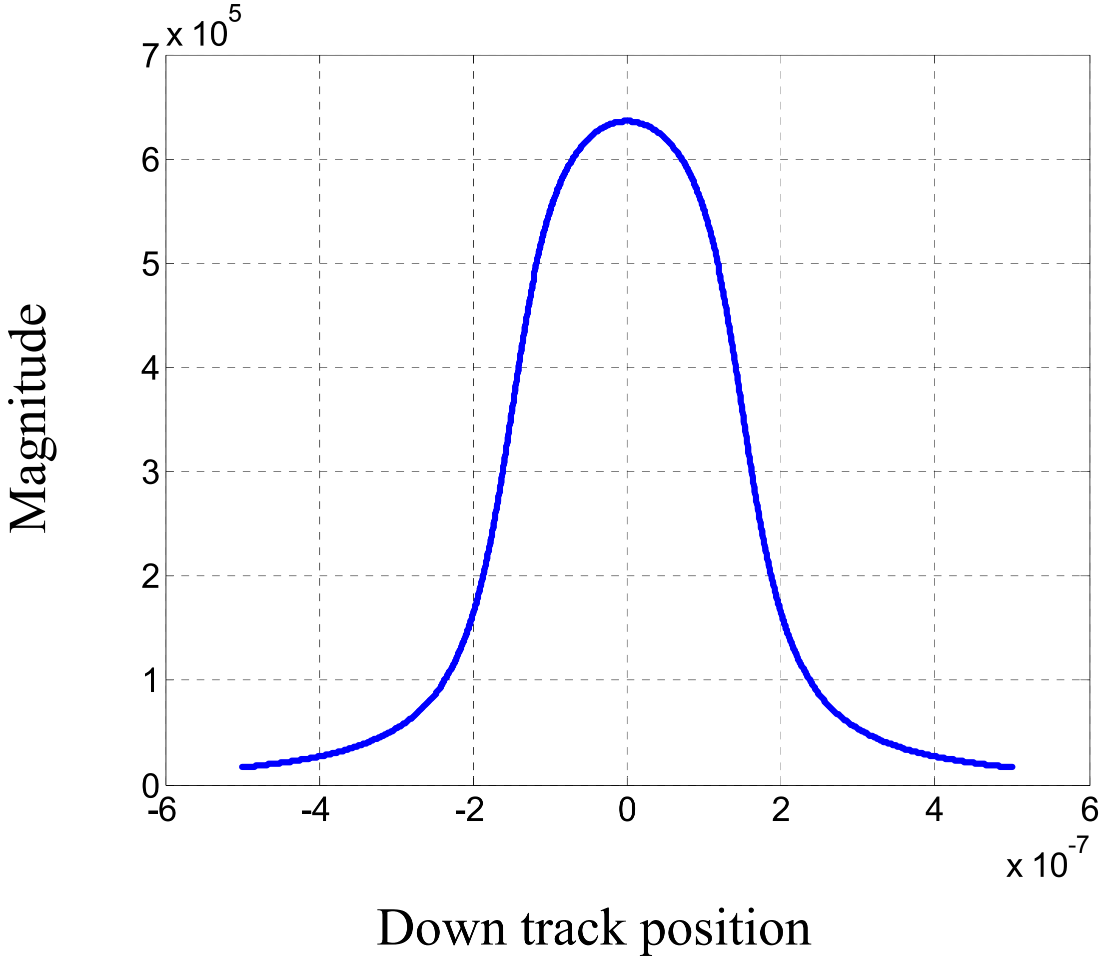
Requires iterations to find a_parameter and x_0
for given temperature profile

Temperature profile as a function of position



Peak Temp = 700
Temp sigma = 150 nm

Longitudinal Component of Karlqvist Head Field as a function of position



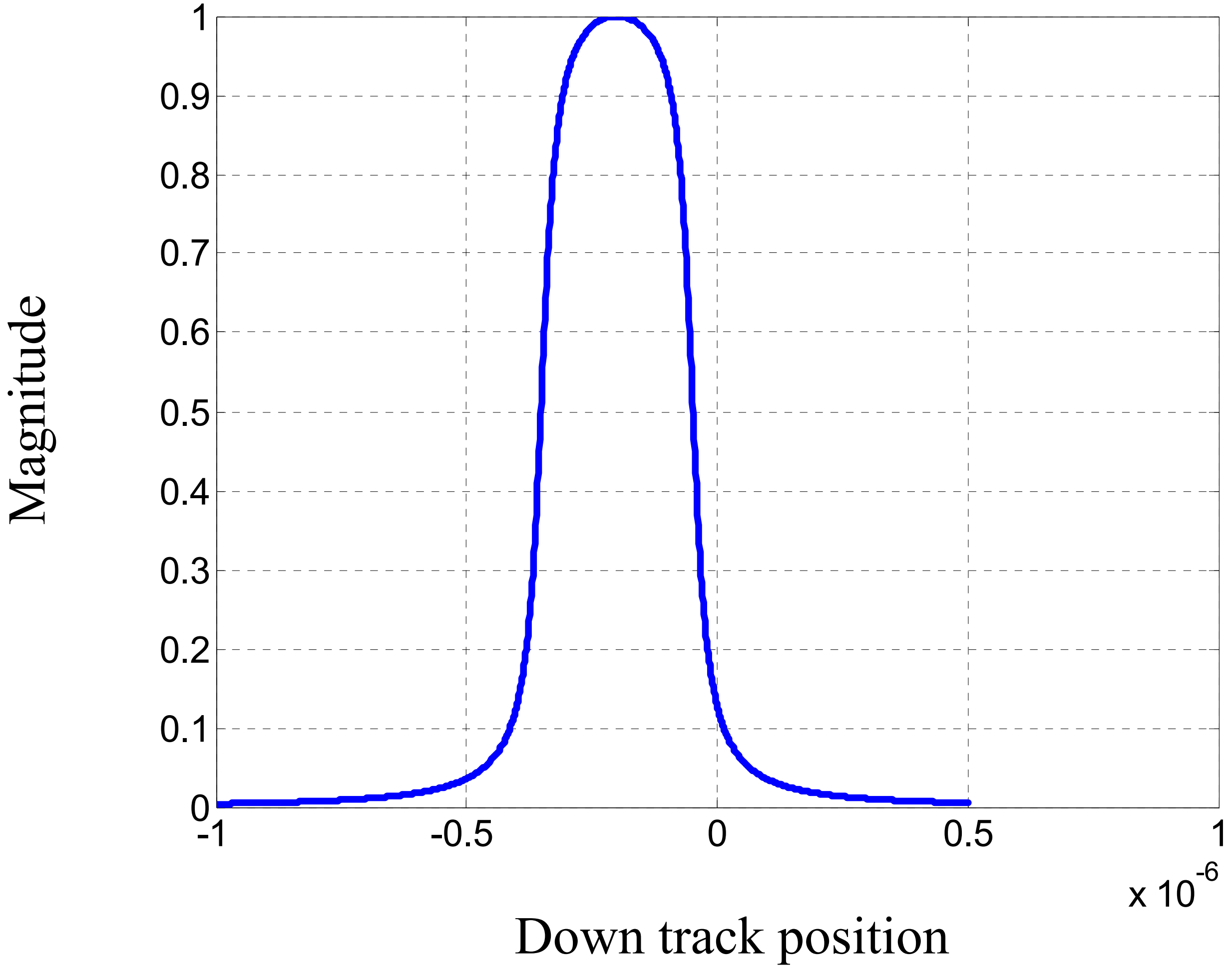
$$H_x = \frac{H_0}{\pi}\left[\tan^{-1}\left(\frac{x+g/2}{y}\right) - \tan^{-1}\left(\frac{x-g/2}{y}\right)\right]$$

Ho = 800000 A/m

g = 300nm

y = 50 nm

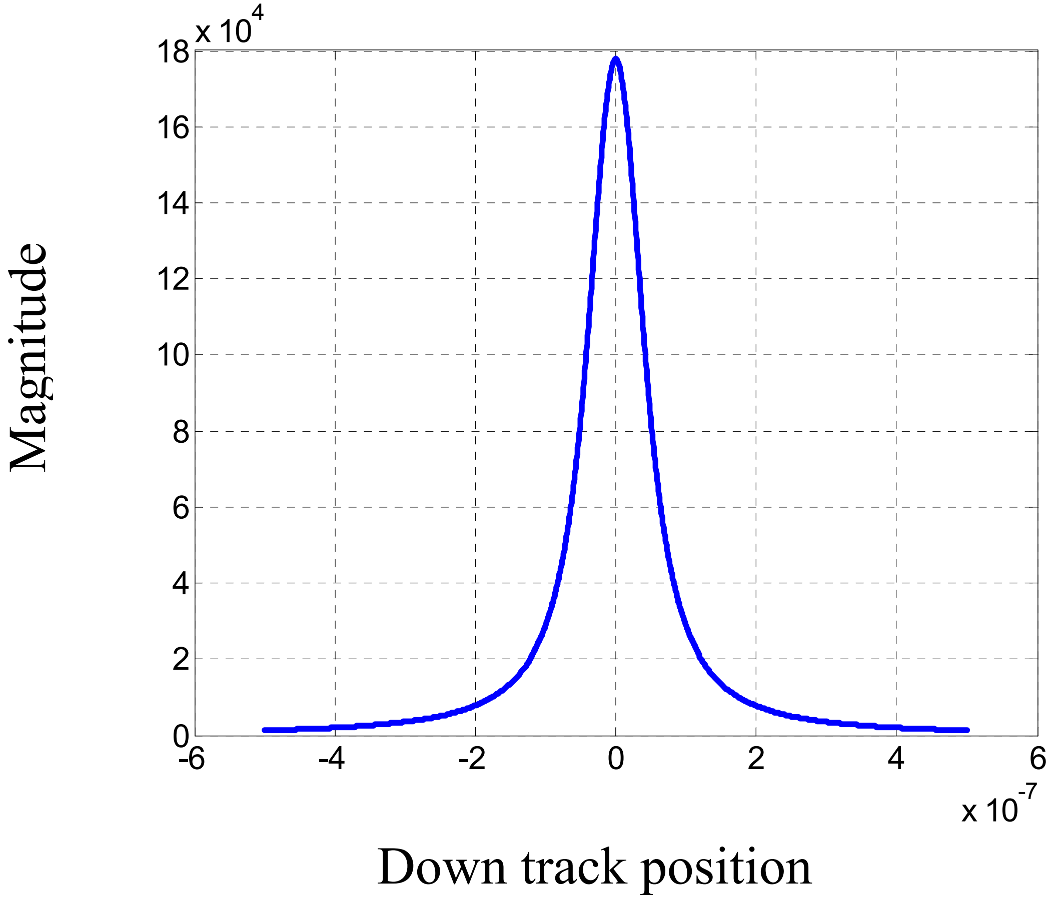
Normalized isolated transition response as a function of position



a_parameter = 22.8nm

x_0 = -200nm

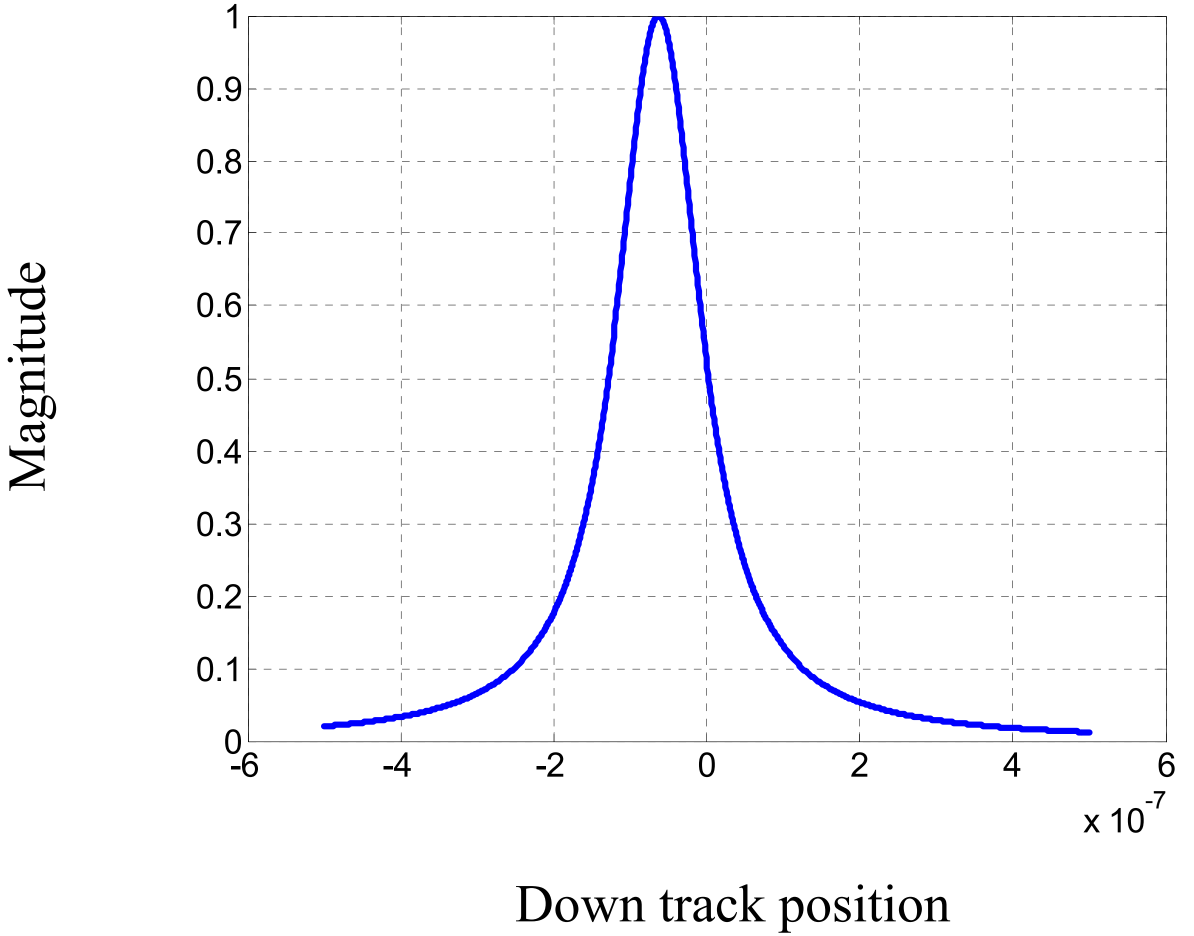
Another Head Field as a function of position



$$H_x = \frac{H_0}{\pi}\left[\tan^{-1}\left(\frac{x+g/2}{y}\right) - \tan^{-1}\left(\frac{x-g/2}{y}\right)\right]$$

Ho = 500000 A/m
g = 50 nm
y = 40 nm

Normalized isolated transition response as a function of position



A_parameter = 20.6nm

X_0 = -63nm

System should be considered as a whole

- Joint optimization of
  - Tribological system
  - Medium magnetics
  - Near field optical system

  Necessary to solve the issues

- Success in optimization will determine the attainable areal density

Object Based Storage Devices

Storage Architectures Today

Goal: Scalability, Security, Data Sharing

Direct Attached Storage (*DAS*)

I/O Application

Storage System

Storage Device

(Blocks)

- **Pros:** Simple, Physically secure
- **Cons:** Not scalable, no data sharing, limited capacity sharing

Network Attached Storage (*NAS*)

I/O Application

Network

Storage System

Storage Device

(Files)

- **Pros:** Secure data and capacity sharing
- **Cons:** Limited scalability

Storage Area Network (*SAN*)

I/O Application

Storage System

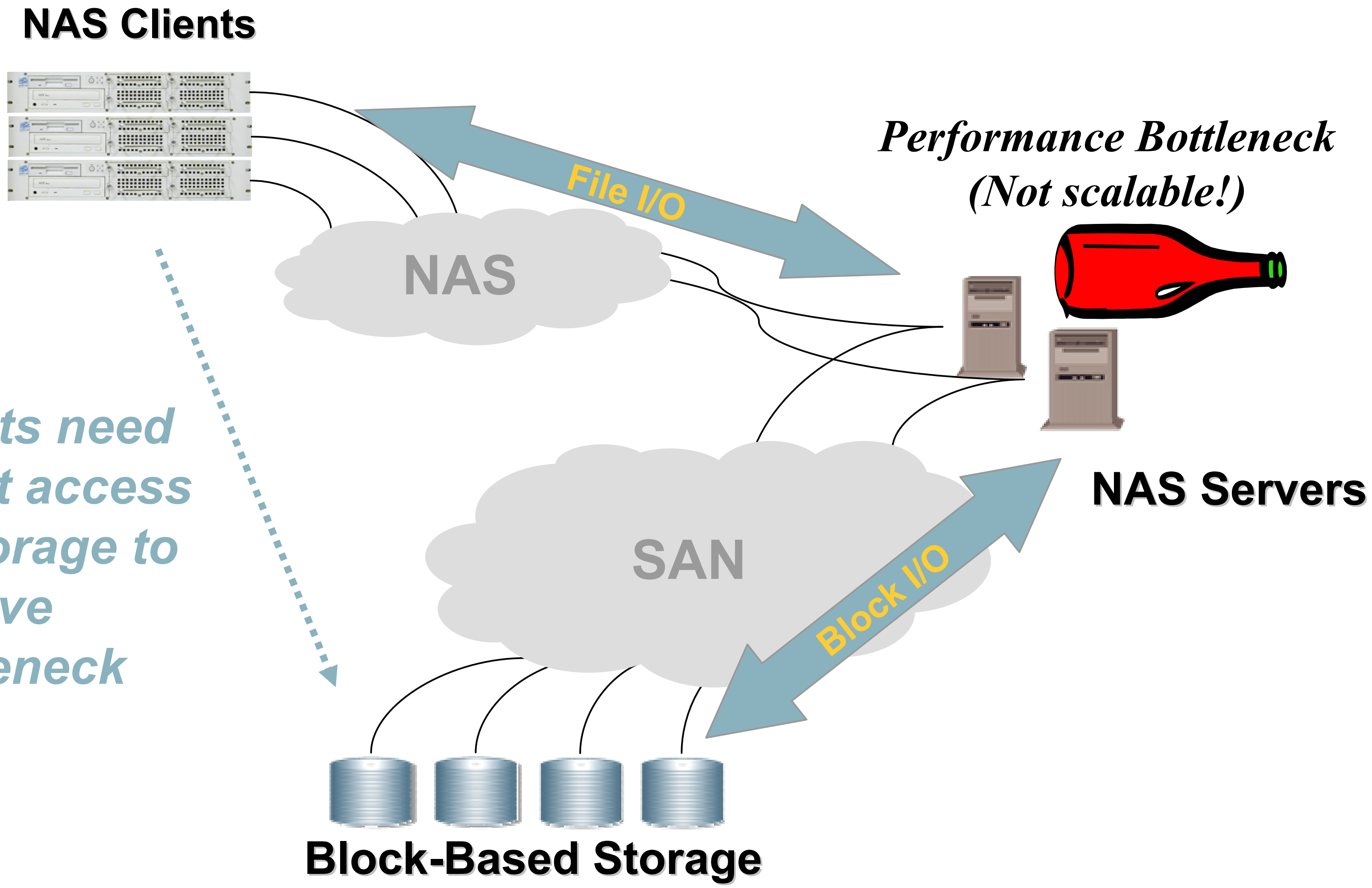
Network

Storage Device

(Blocks)

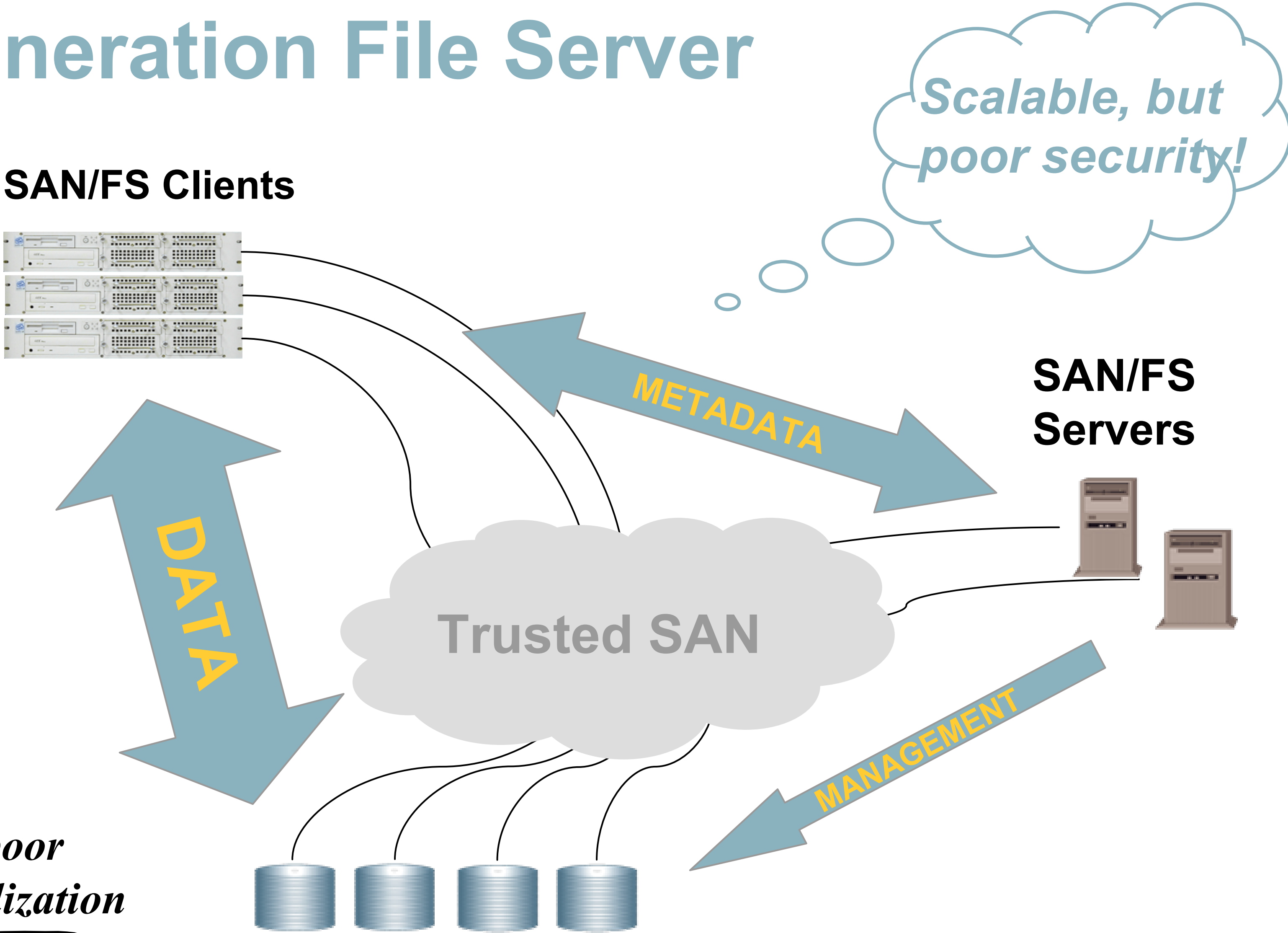
- **Pros:** Scalable capacity sharing
- **Cons:** No data sharing

Today's File Server



Clients need direct access to storage to remove bottleneck

2nd Generation File Server



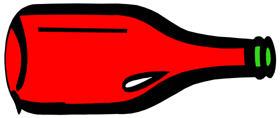
Scalable, but poor security!

SAN/FS Clients

SAN/FS Servers

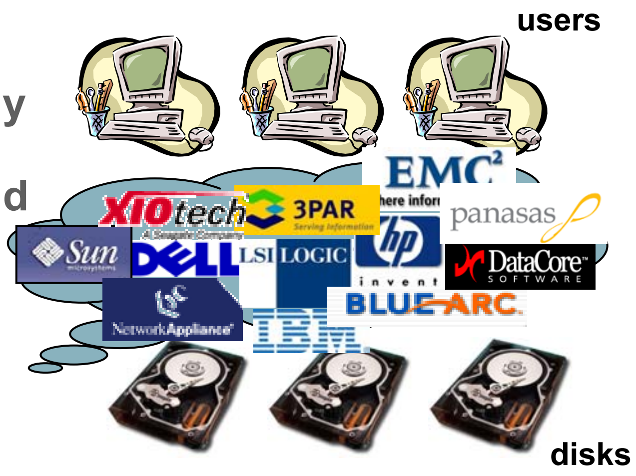
Block-Based Storage

-Security is poor
-High virtualization overhead

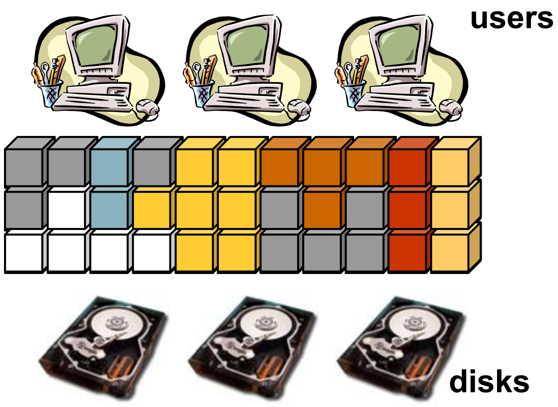


The Problem – expand endpoint function

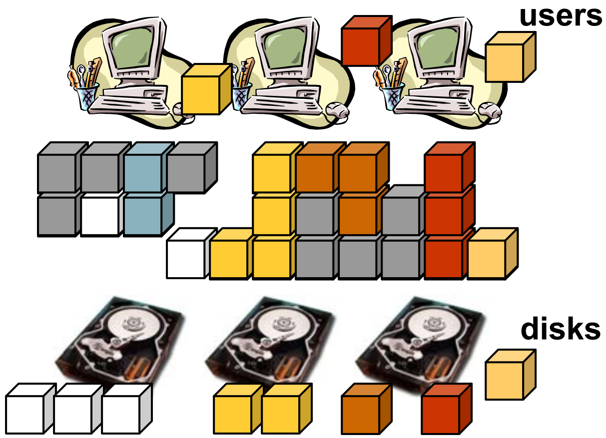
today – many layers of hardware and software between users and disks



understand what is really going on, disambiguate functions



move appropriate functions to the endpoints



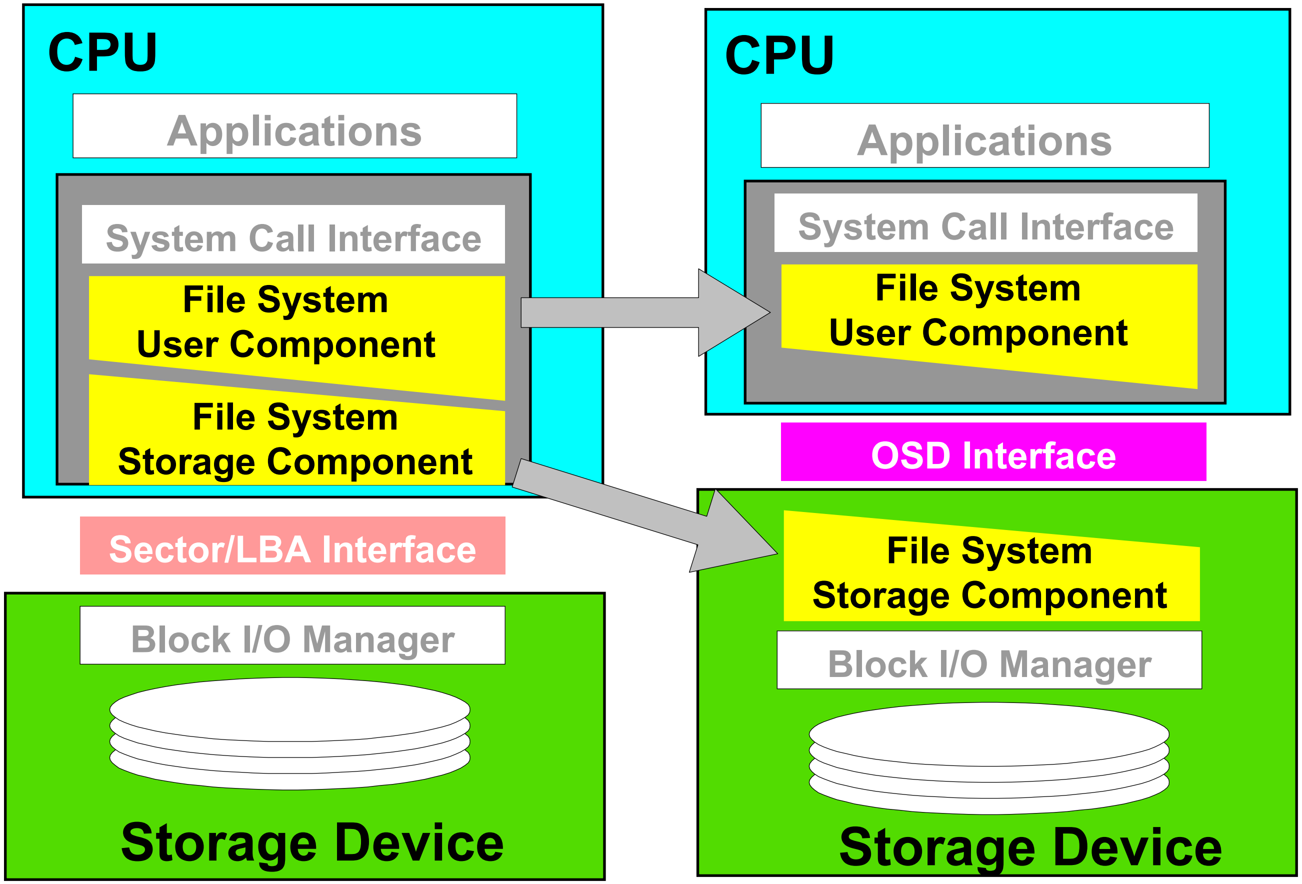
Endpoints (users at top, disks at bottom)

- not very sophisticated today
- disks don’t understand users/apps
- users/apps don’t understand disks
- need many intermediaries to translate

Intermediary functions

- some add value (e.g. data sharing)
- others simply cover limitations (e.g. reliability via RAID)

OSD Interface



OSD Functions

Basic Protocol

- READ
- WRITE

Very Basic

- CREATE
- REMOVE

Space Mgmt

- GET ATTR
- SET ATTR

Attributes

- timestamps
- vendor-specific
  - shared, opaque

Specialized

- APPEND – write w/o offset
- CREATE & WRITE – save msg
- FLUSH OBJ – force to media
- LIST – recovery of objects

Security

- Authorization – on each request
- Integrity – for args & data
- SET KEY
- SET MASTER KEY

shared secrets

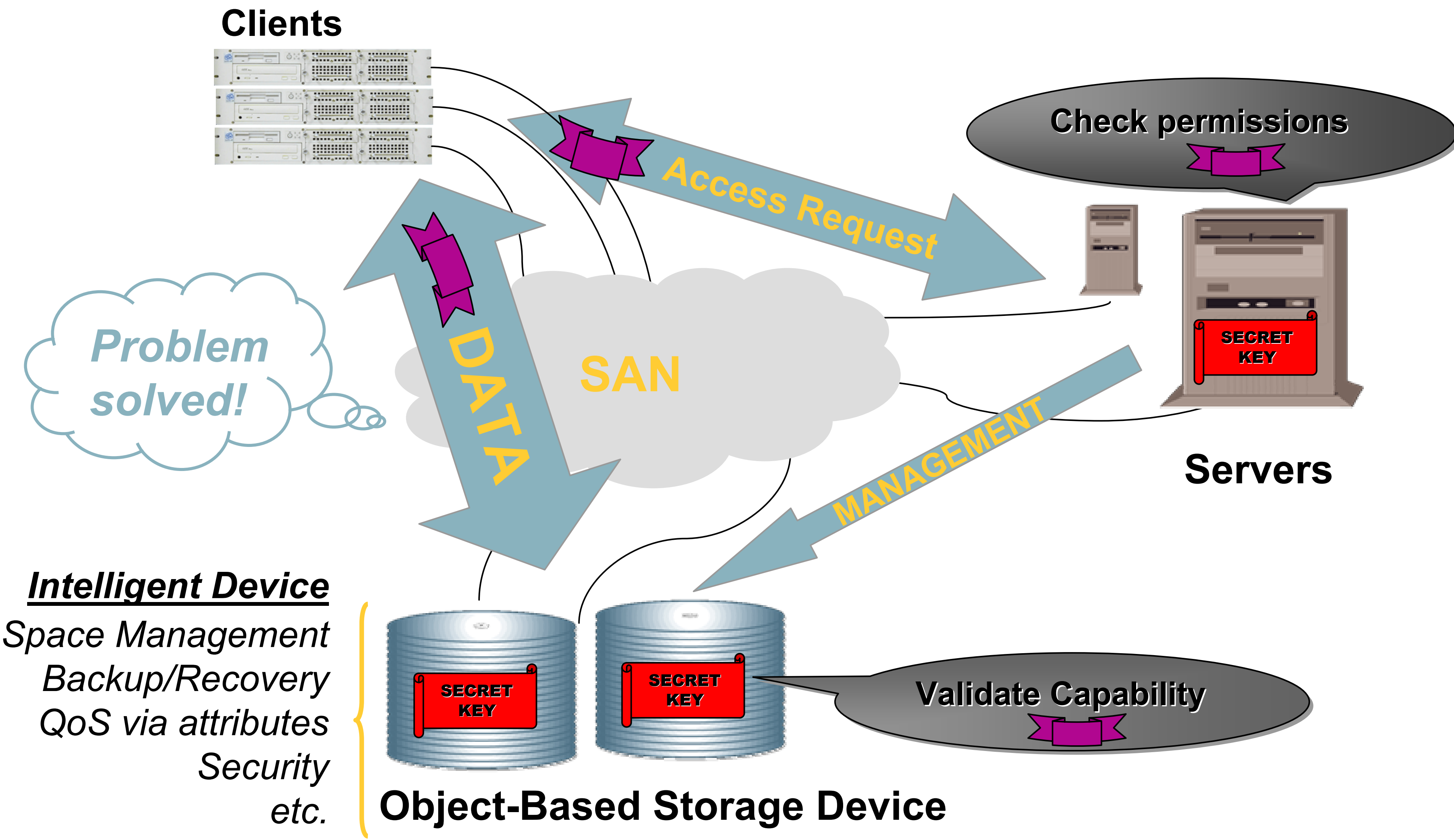
Groups

- CREATE COLLECTION
- REMOVE COLLECTION
- LIST COLLECTION

Management

- FORMAT OSD
- CREATE PARTITION
- REMOVE PARTITION

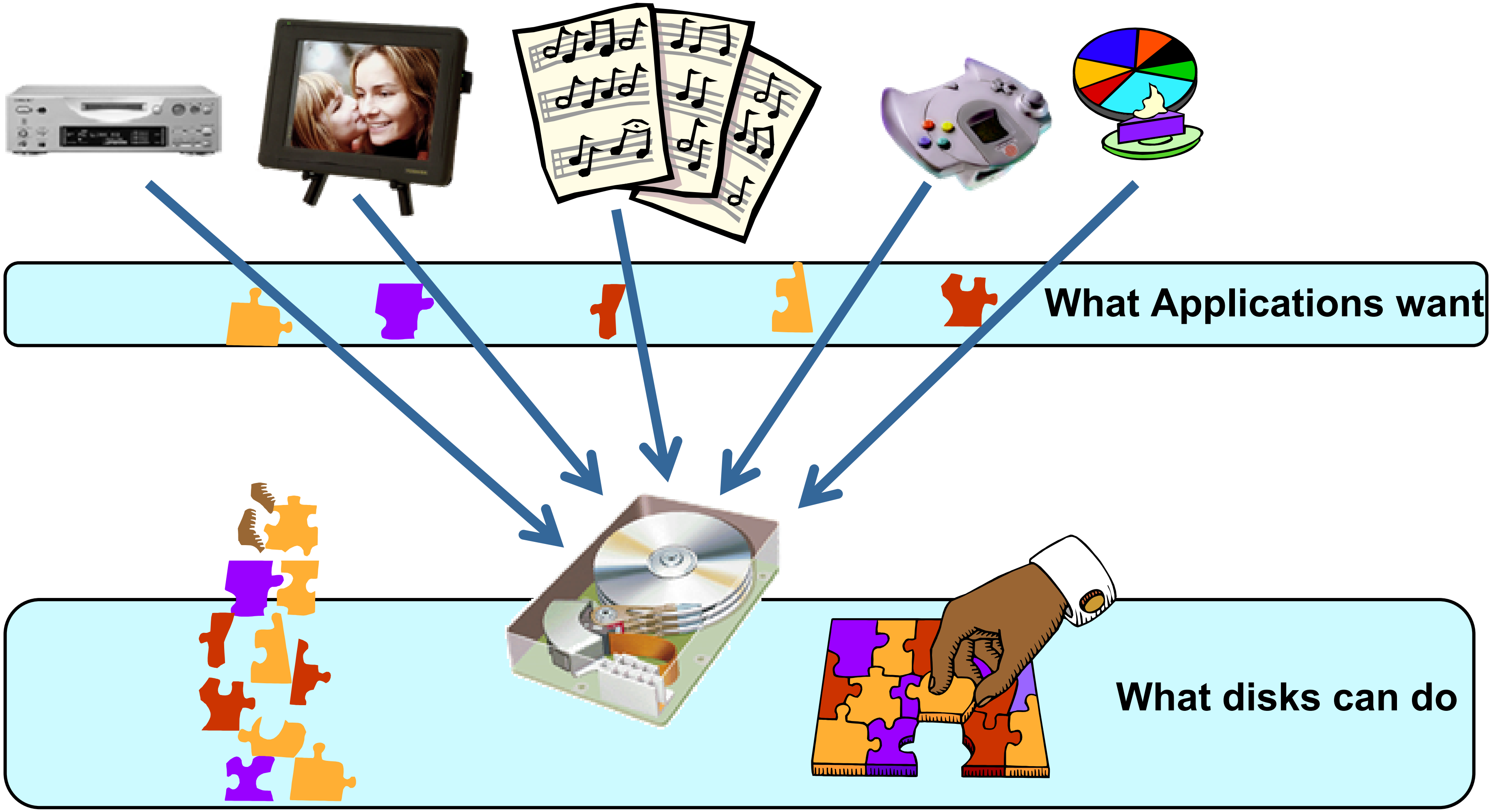
File Server with OSD



Intelligent Device

Space Management
Backup/Recovery
QoS via attributes
Security
etc.

Quality of Service on the Disk



What Applications want

What disks can do

Traditional

OSD

Example OSD Drive Uses

–Enterprise

- Data sharing
- Scalability
- Security
- Improved reliability
- Self-managed, self-configured drives

– Desktops/Notebooks

- Automatic de-fragmentation
- Object semantics
- Free space management
- QoS

–Consumer Electronics

- Video streams
  - Video can tolerate occasional bit loss
  - Drive takes advantage of this to improve performance (e.g., skipping ECC on a read)

XIOtech
A Seagate Company
DELL
EMC²
where information lives
Network Appliance
Microsoft
IBM
webtv
hp
invent

XBOX
ECHOSTAR

Summary

- –OSD enables scalable and secure data sharing
  - Highly desirable in the enterprise market
- –Pushes intelligence down to disks
  - Self-aware, self-managed, self-configured drives
  - Better communication between applications and drives (QoS)
  - Increased system performance
  - Disk level computations (searches, etc.)
- –It is REAL and HERE