Compressive Sensing

Part II: Sensing Matrices and Real-World Sensors

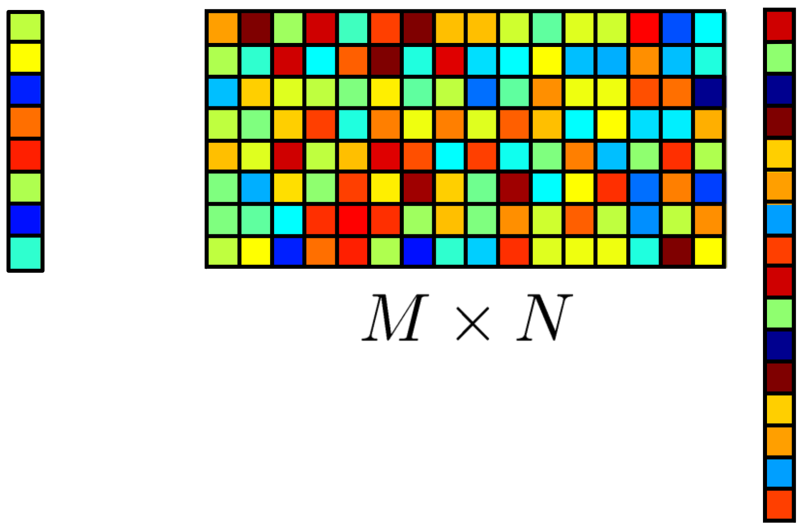
Mark A. Davenport

Stanford University
Department of Statistics

Compressive Sensing

Replace samples with general *linear measurements*

$$y = \Phi x$$

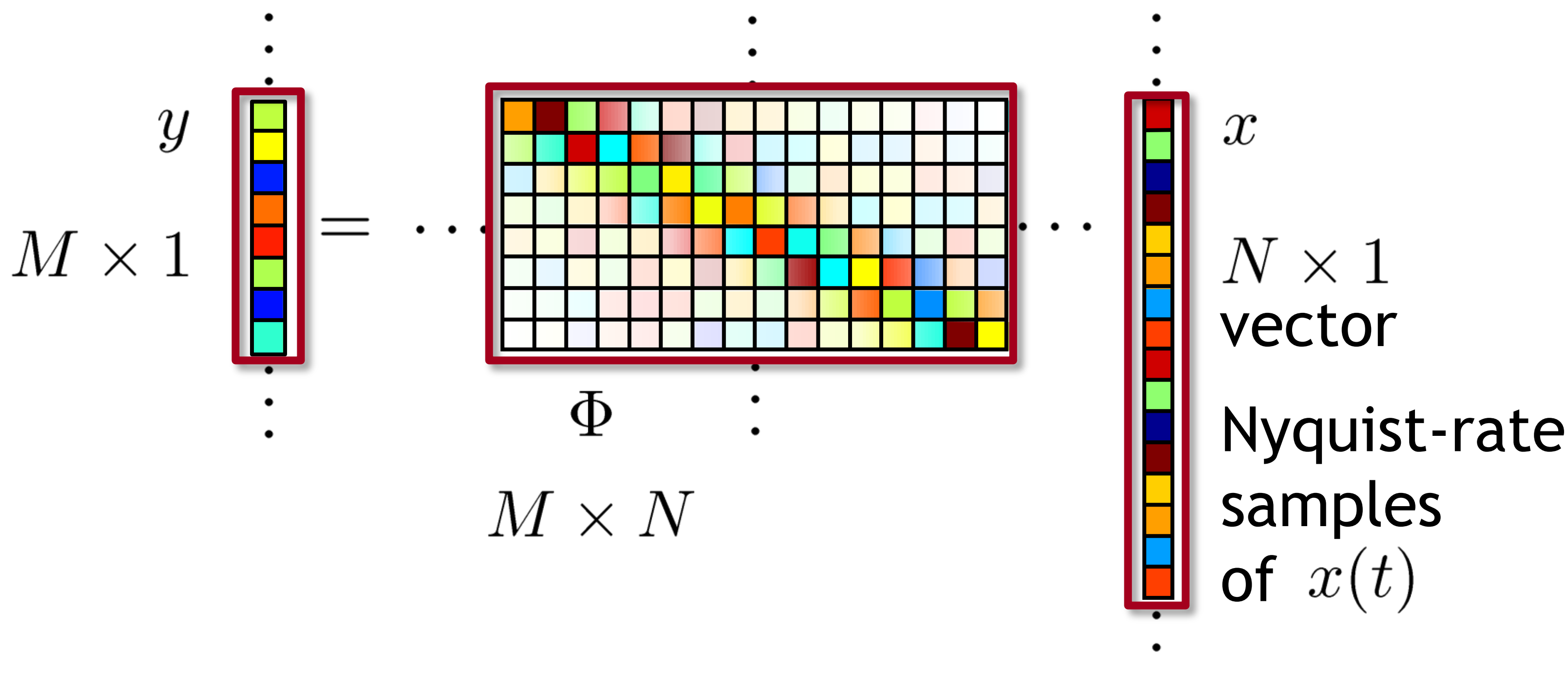


[Donoho; Candès, Romberg, and Tao - 2004]

Analog Sensing *is* Matrix Multiplication

If $x(t)$ is bandlimited,

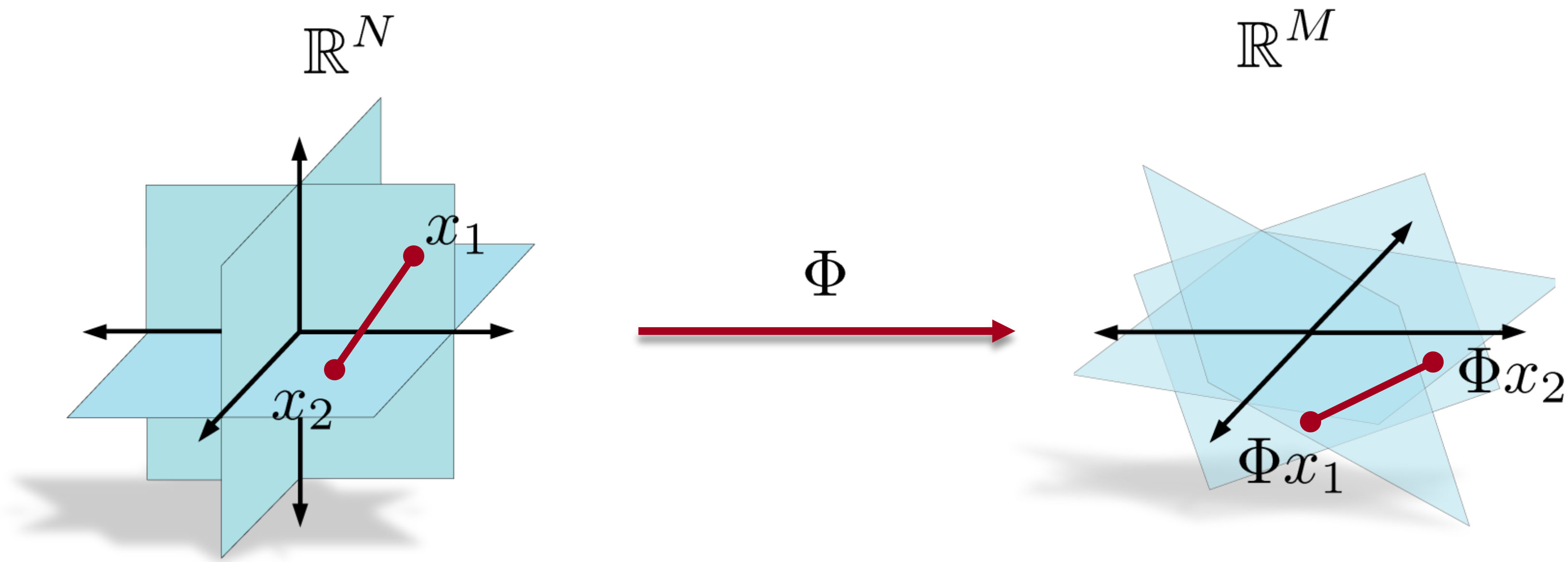
$$y[m] = \langle \phi_m(t), x(t) \rangle = \sum_{n=-\infty}^{\infty} x[n] \, \langle \phi_m(t), \mathrm{sinc}(t/T_s - n) \rangle$$



Sensing Matrix Design

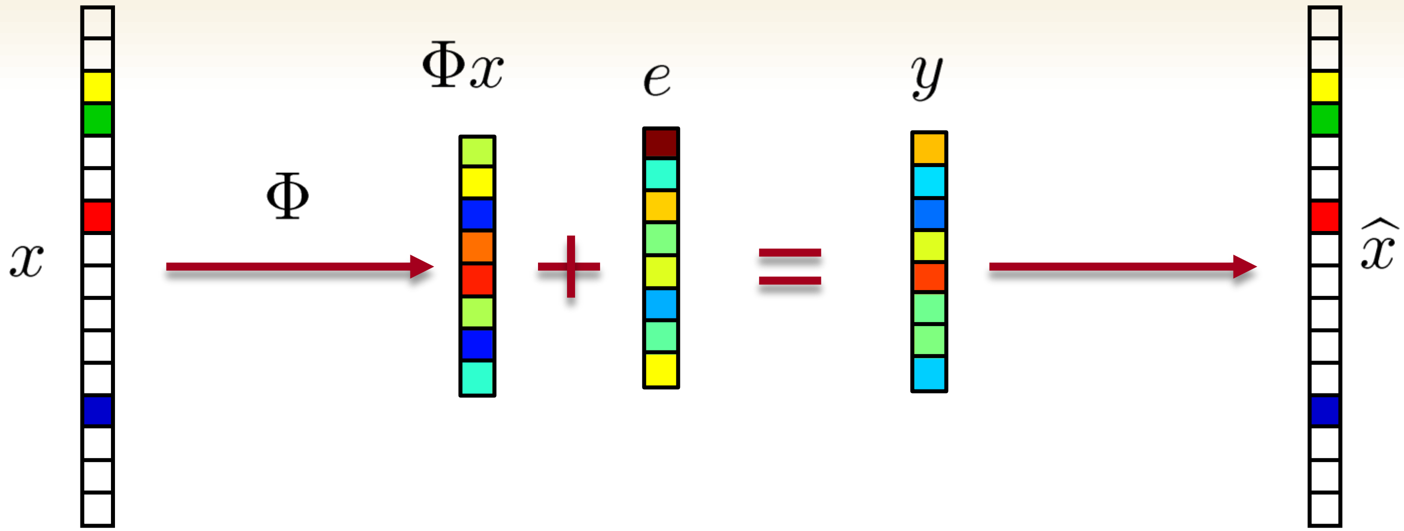
Restricted Isometry Property (RIP)

$$1 - \delta \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq 1 + \delta \qquad \|x_1\|_0, \|x_2\|_0 \leq S$$



$$1 - \delta \leq \frac{\|\Phi x\|_2^2}{\|x\|_2^2} \leq 1 + \delta \qquad \|x\|_0 \leq 2S$$

RIP and Stability



If we want to guarantee that

$$\|x - \widehat{x}\|_2 \leq C\|e\|_2$$

then we must have

$$\boxed{\frac{1}{C} \leq \frac{\|\Phi x\|_2^2}{\|x\|_2^2} \qquad \|x\|_0 \leq 2S}$$

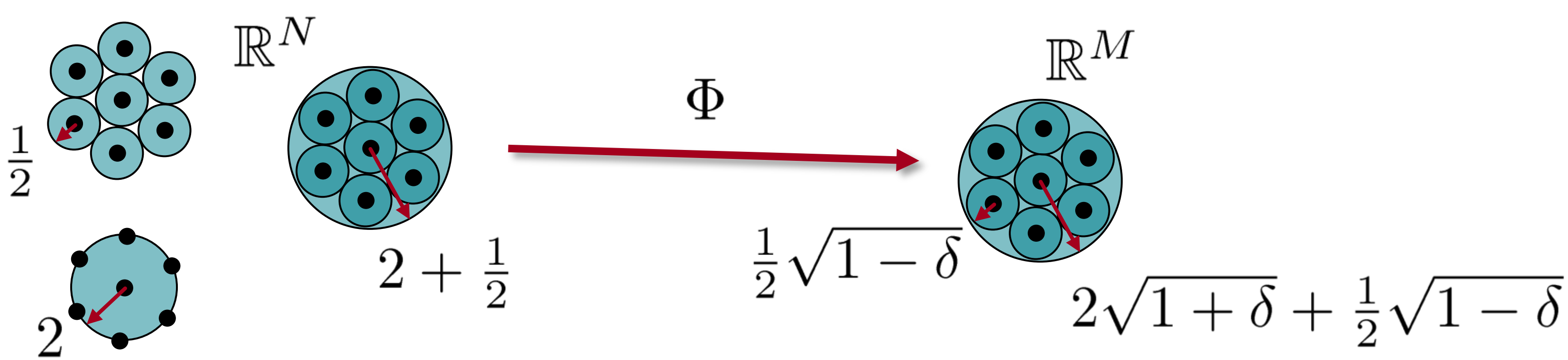
How Many Measurements?

If Φ satisfies the RIP with constant δ, then

$$M > C_{S,\delta} S \log\left(N/S\right)$$

Sketch of proof: Construct a set $\mathcal{X}$ such that

- for any $x \in \mathcal{X}$, $\|x\|_0 = S$
- $|\mathcal{X}| \approx (N/S)^S$
- for any pair $x, y \in \mathcal{X}$, $1 \leq \|x - y\|_2 \leq 2$



Sub-Gaussian Distributions

- Sub-Gaussian: $\mathbb{E}\left(e^{Xt}\right) \leq e^{c^2 t^2/2}$
  - Gaussian
  - Bernoulli/Rademacher (± 1)
  - any bounded distribution

- Strictly sub-Gaussian: $\mathbb{E}\left(e^{Xt}\right) \leq e^{\sigma^2 t^2/2}$

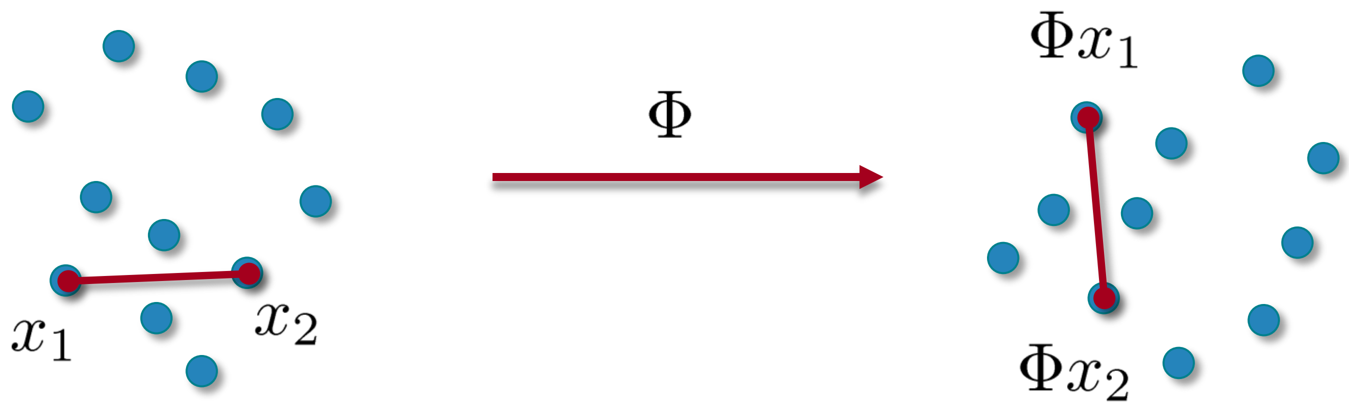
- For any x, if the entries of Φ are sub-Gaussian, then there exist α and β such that w.h.p.

$$\alpha \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq \beta \|x\|_2^2$$

Strictly sub-Gaussian $\Rightarrow$ $\alpha = 1 - \delta\,, \quad \beta = 1 + \delta$

Johnson-Lindenstrauss Lemma

- Stable projection of a discrete set of P points



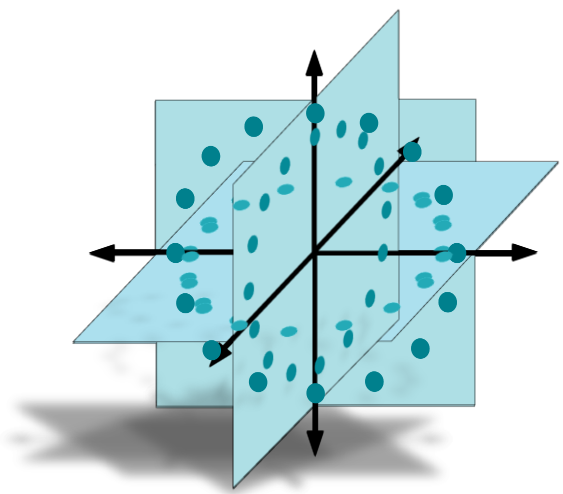
- Pick Φ at *random* using a *sub-Gaussian* distribution

- For any fixed x, $\|\Phi x\|_2$ concentrates around $\|x\|_2$ with (exponentially) high probability

- We preserve the length of all $O(P^2)$ difference vectors simultaneously if $M = O(\log P^2) = O(\log P)$.

JL Lemma Meets RIP

$$1 - \delta \leq \frac{\|\Phi x\|_2^2}{\|x\|_2^2} \leq 1 + \delta \qquad \|x\|_0 \leq 2S$$

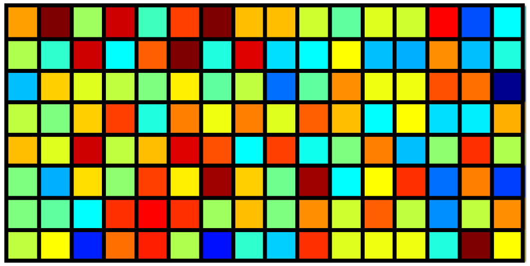


$$P = O\left((N/S)^S\right) \quad \Rightarrow \quad M = O(S \log(N/S))$$

[Baraniuk, D, DeVore, and Wakin –2008]

RIP Matrix: Option 1

- Choose a ***random matrix***
  - fill out the entries of Φ with i.i.d. samples from a sub-Gaussian distribution
  - project onto a “random subspace”

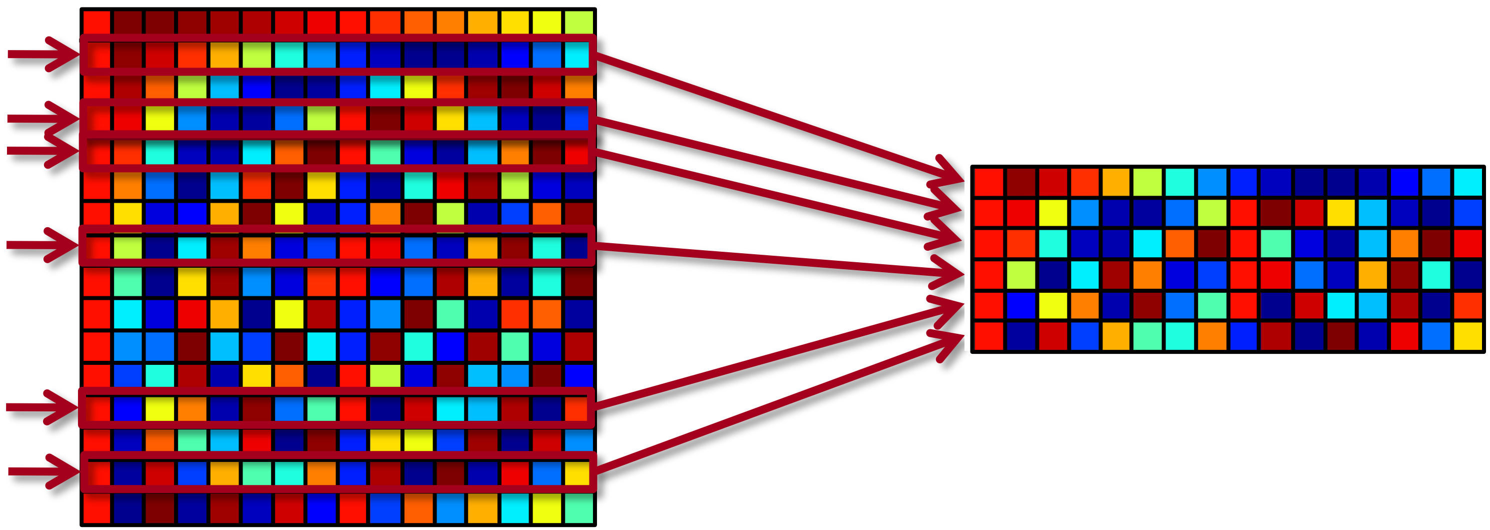


$$M = O(S \log(N/S)) \ll N$$

[Baraniuk, D, DeVore, and Wakin –2008]

RIP Matrix: Option 2

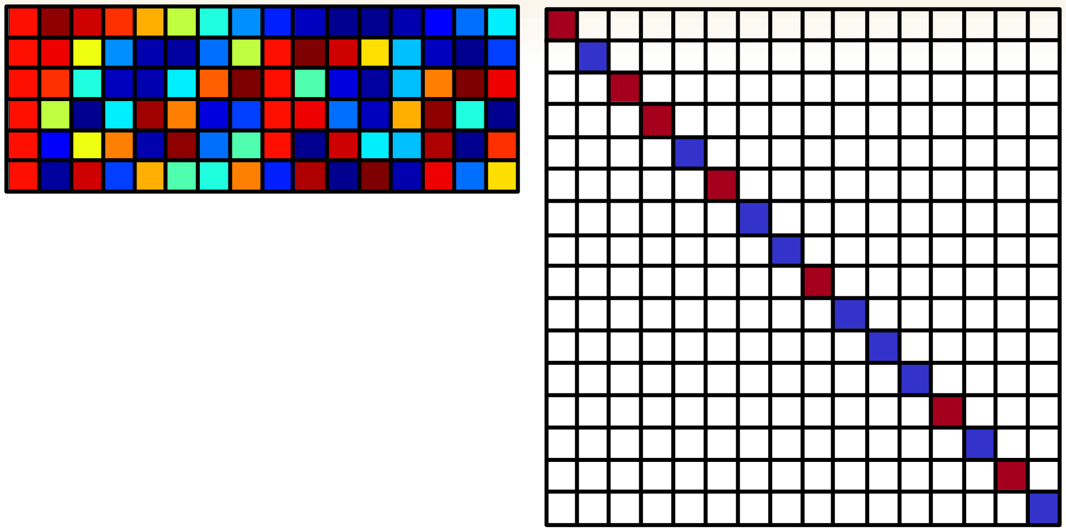
- Random Fourier submatrix



$$M = O(S \log^p (N/S)) \ll N$$

[Candès and Tao - 2006]

"Fast JL Transform"



- By first multiplying by random signs, a random Fourier submatrix can be used for efficient JL embeddings

- If you multiply the columns of *any* RIP matrix by random signs, you get a JL embedding!

[Ailon and Chazelle – 2007; Krahmer and Ward - 2010]

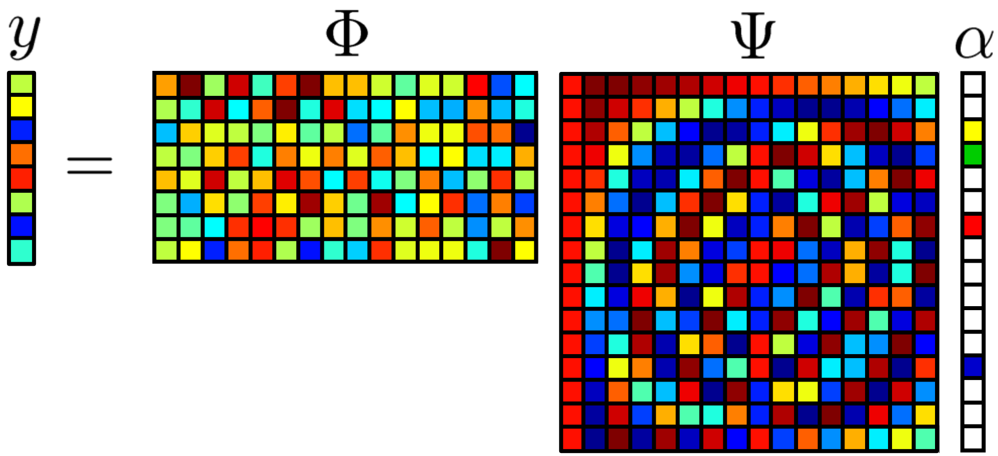
Hallmarks of Random Measurements

Stable

With high probability, Φ will preserve information, be robust to noise

Universal

Φ will work with ***any*** fixed orthonormal basis (w.h.p.)

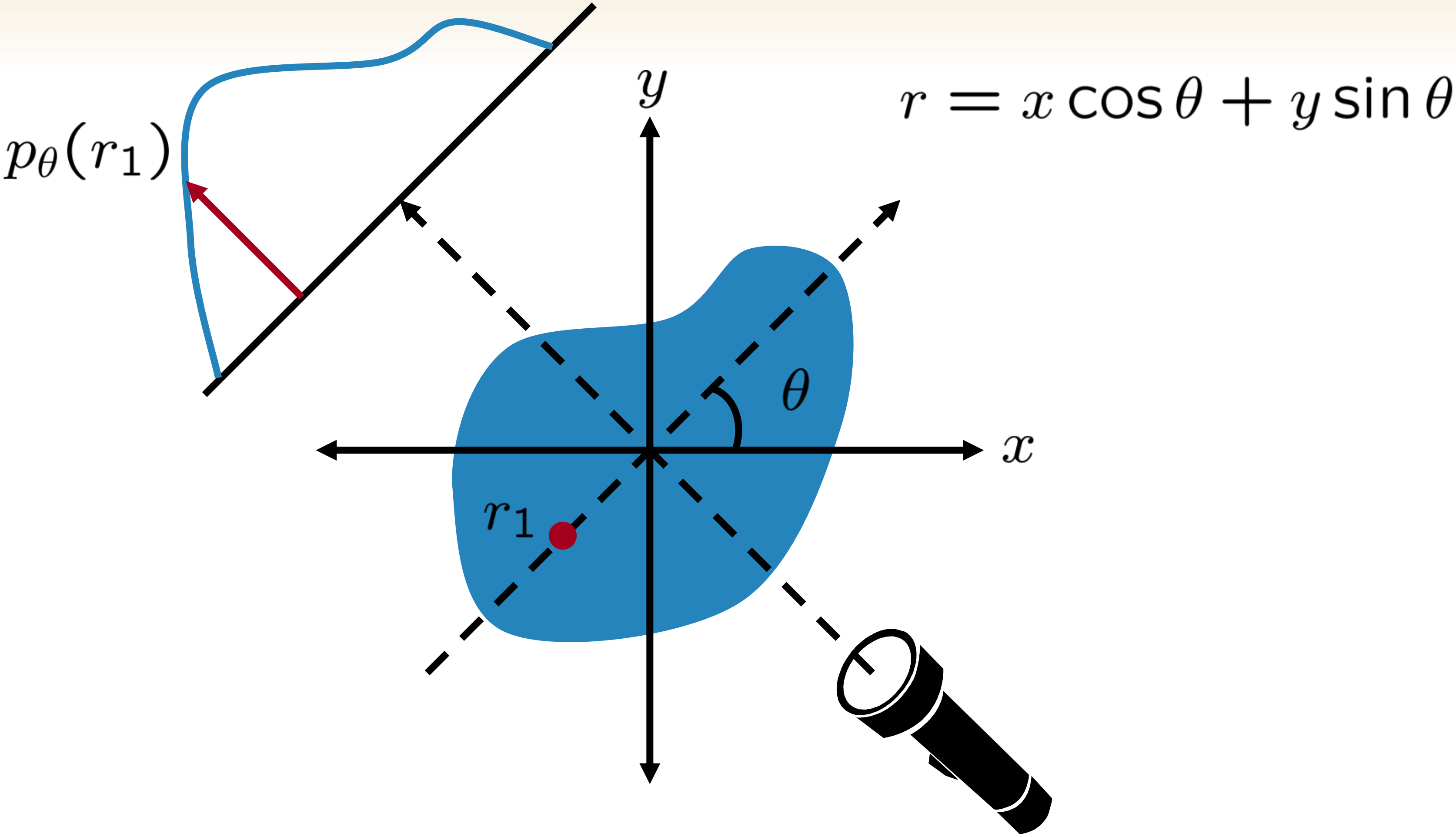


Democratic

Each measurement has "equal weight"

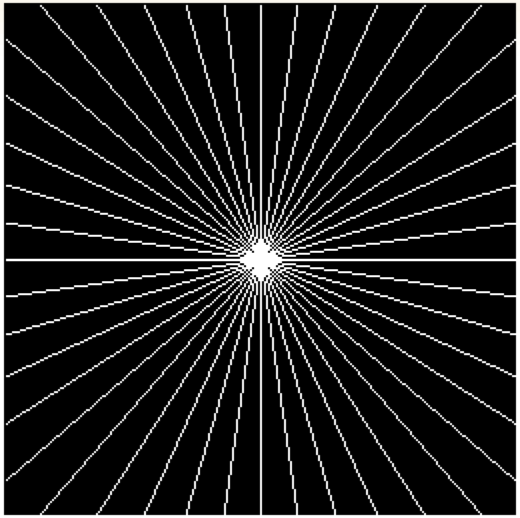
Compressive Sensors in Practice

Tomography in the Abstract



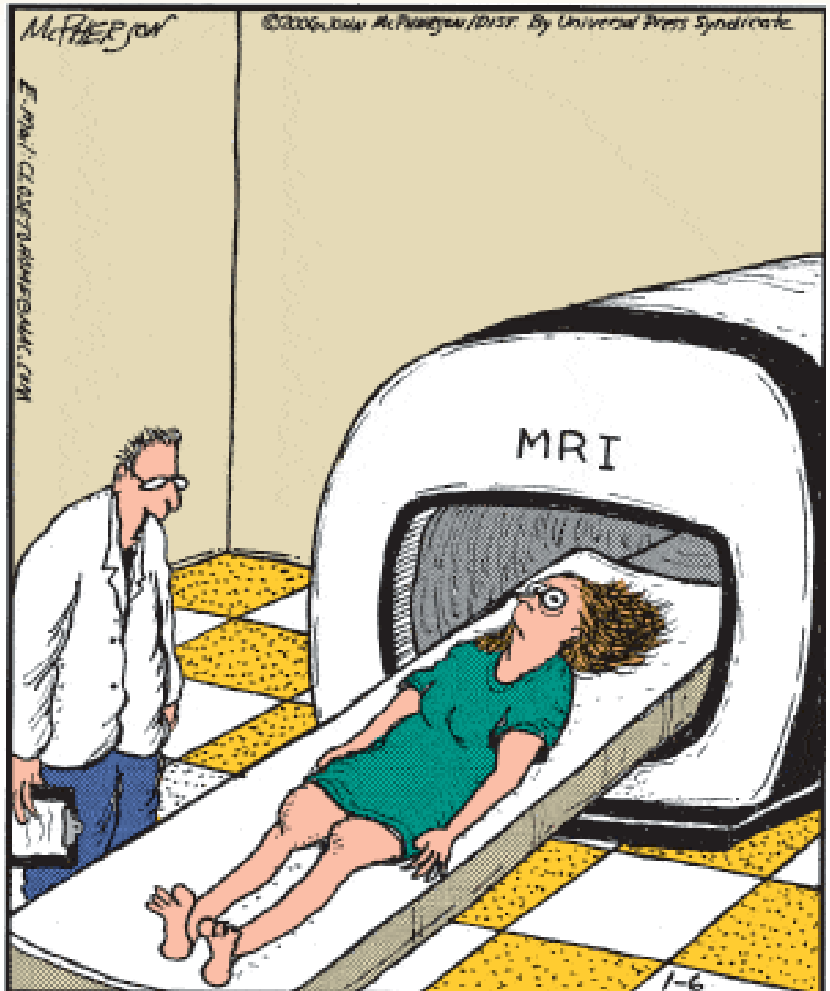
$$p_\theta(r) = \int\int f(x,y)\delta(x\cos\theta + y\sin\theta - r)dxdy$$

Fourier-Domain Interpretation



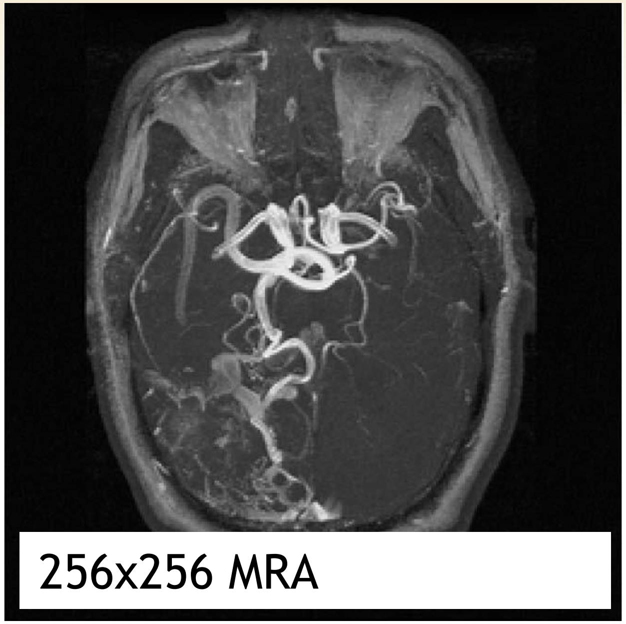
- Each projection gives us a "slice" of the 2D Fourier transform of the original image
- Similar ideas in MRI
- Traditional solution: Collect lots (and lots) of slices

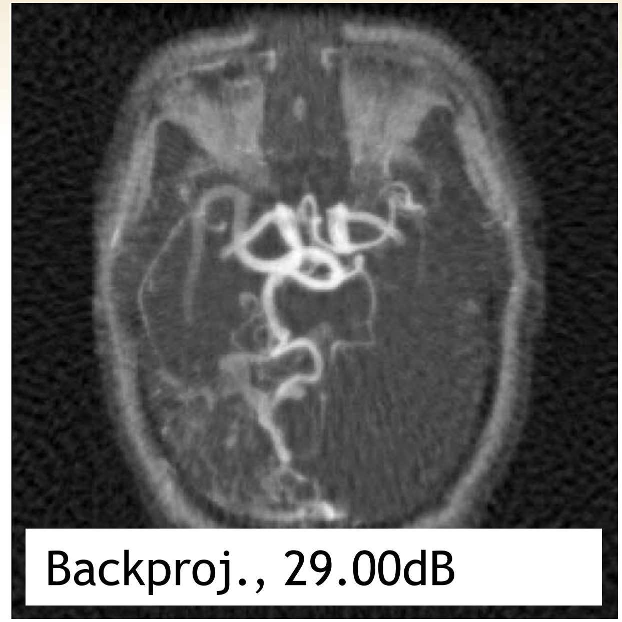
Why CS?

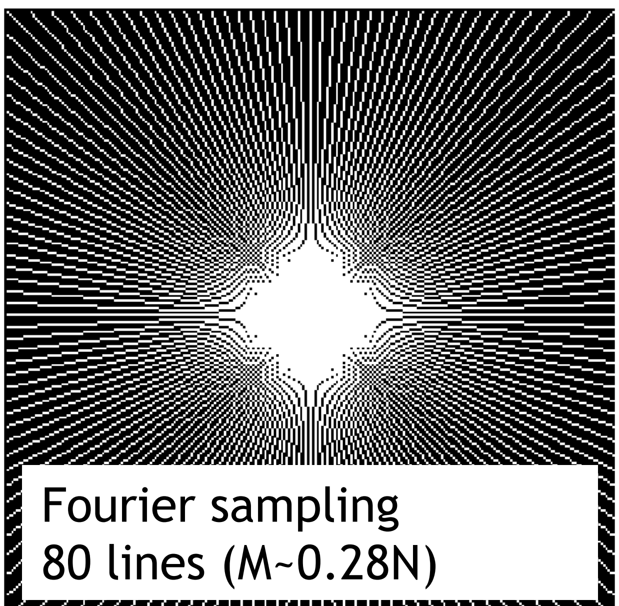


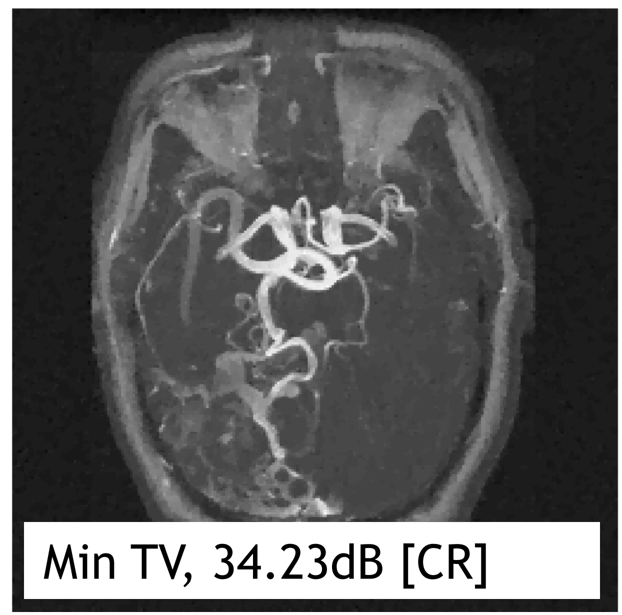
"OK, Mrs. Dunn. We'll slide you in there, scan your brain, and see if we can find out why you've been having these spells of claustrophobia."

CS for MRI Reconstruction

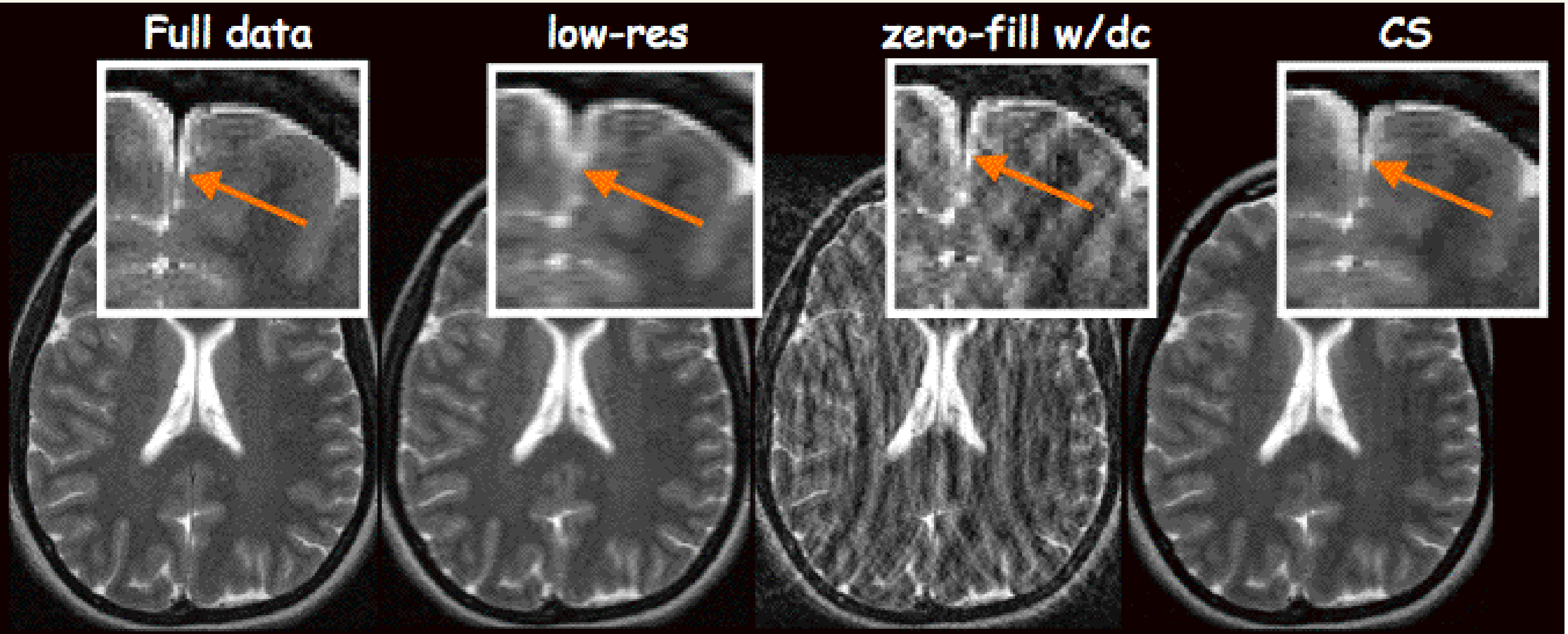




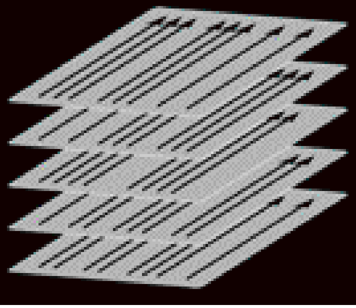




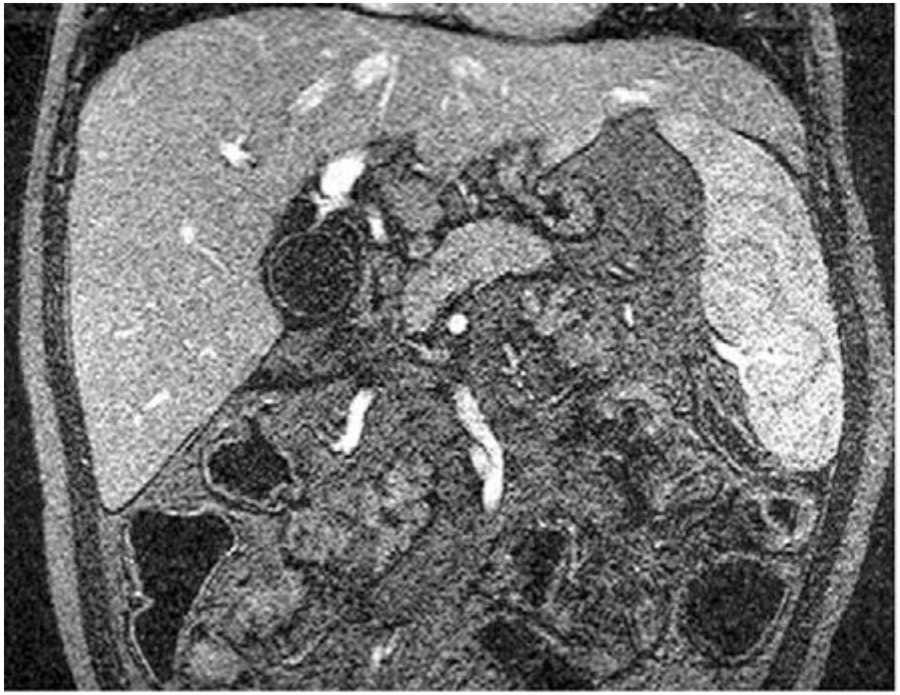
Multi-Slice Brain Imaging



- Scan reduction: x2.4
- Transform: wavelet



Pediatric MRI



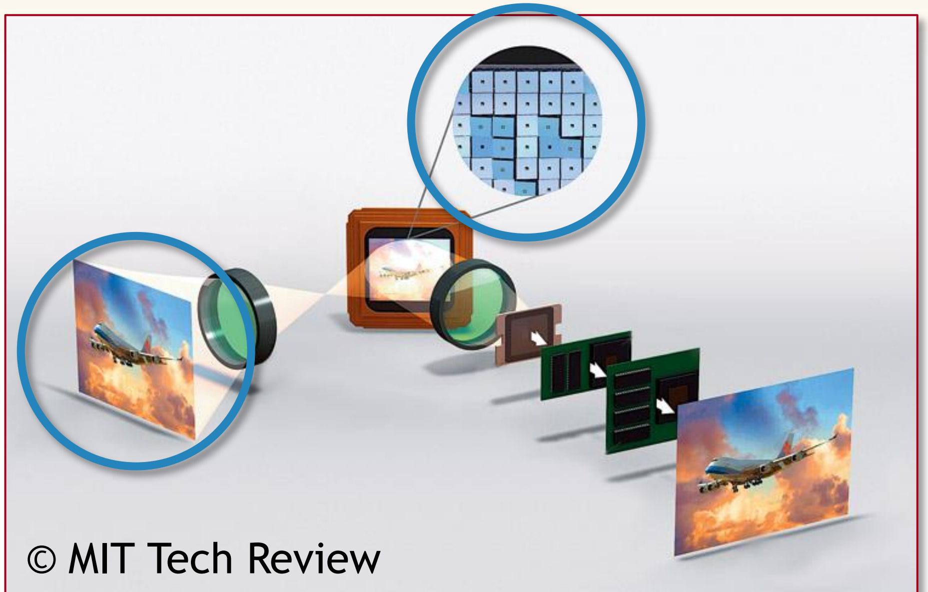
Traditional MRI

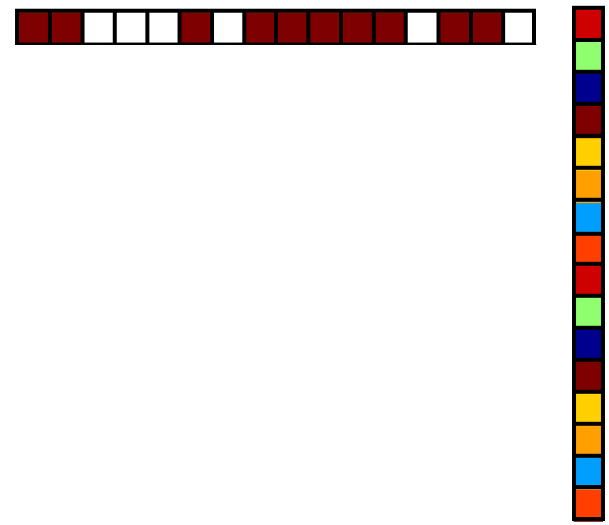
CS MRI

4-8 x faster!

[Vasanawala, Alley, Hargreaves, Barth, Pauly, and Lustig - 2010]

"Single-Pixel Camera"



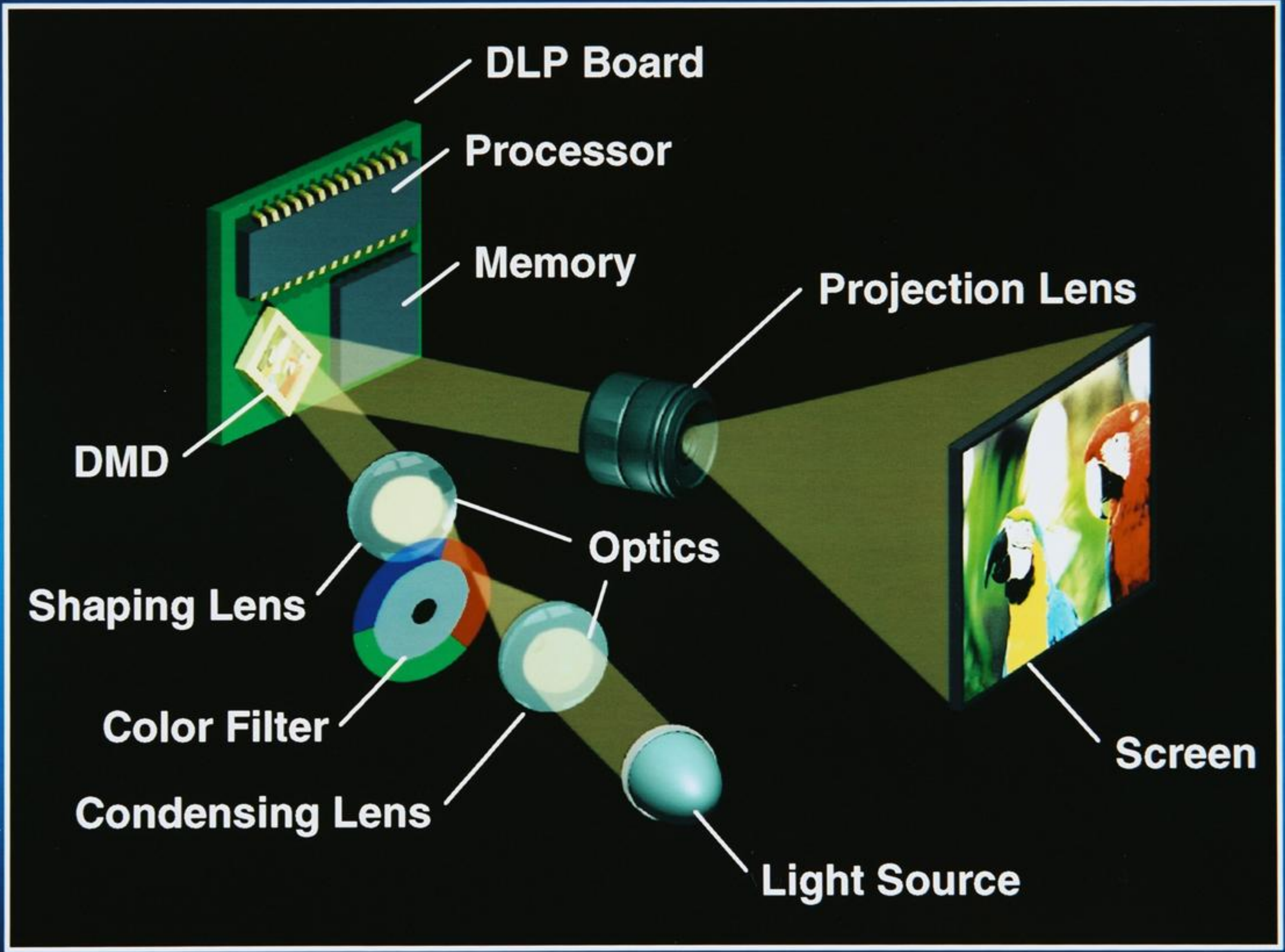


$$y[m] = \sum_{n \in I_m} x[n]$$

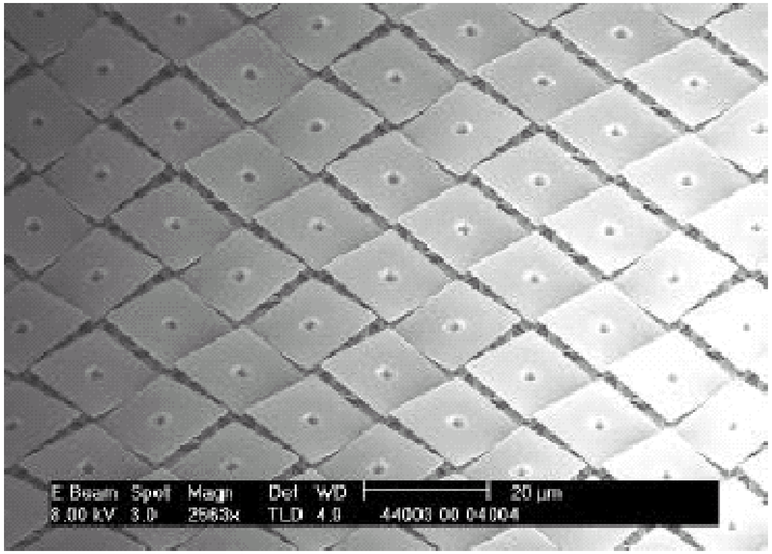
$$x[n] = \int\int_{\text{pixel } n} x(t_1, t_2)\, dt_1\, dt_2$$

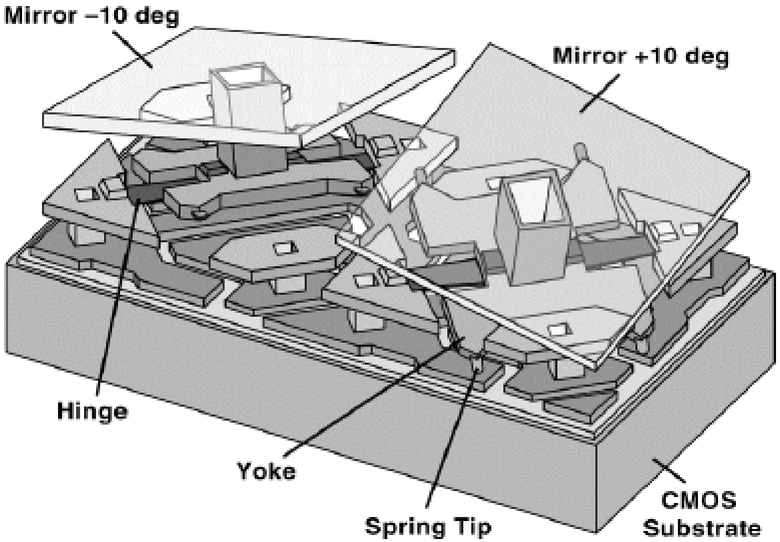
[Duarte, D, Takhar, Laska, Sun, Kelly, and Baraniuk - 2008]

1 Chip DLP™ Projection

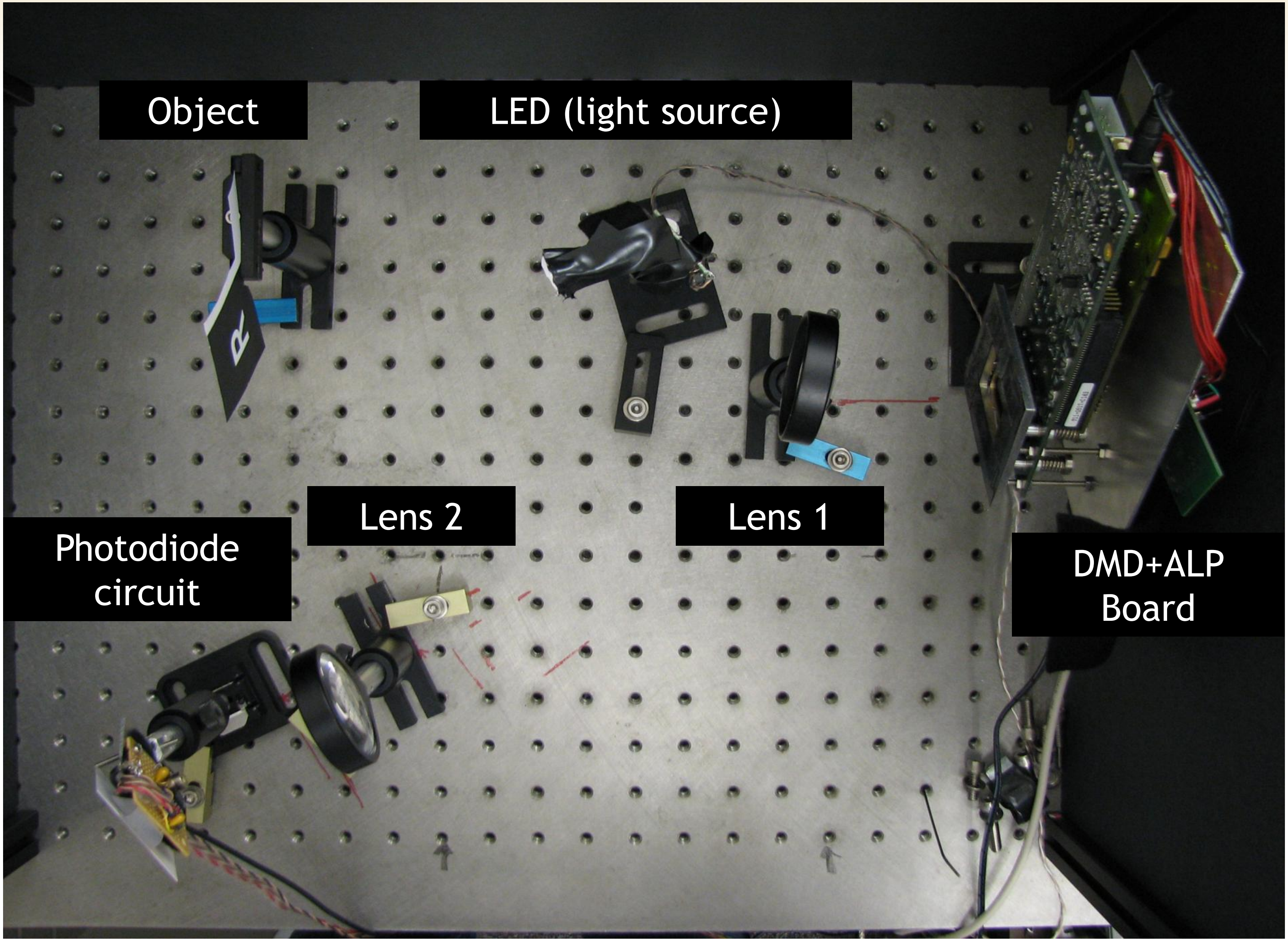


TI Digital Micromirror Device

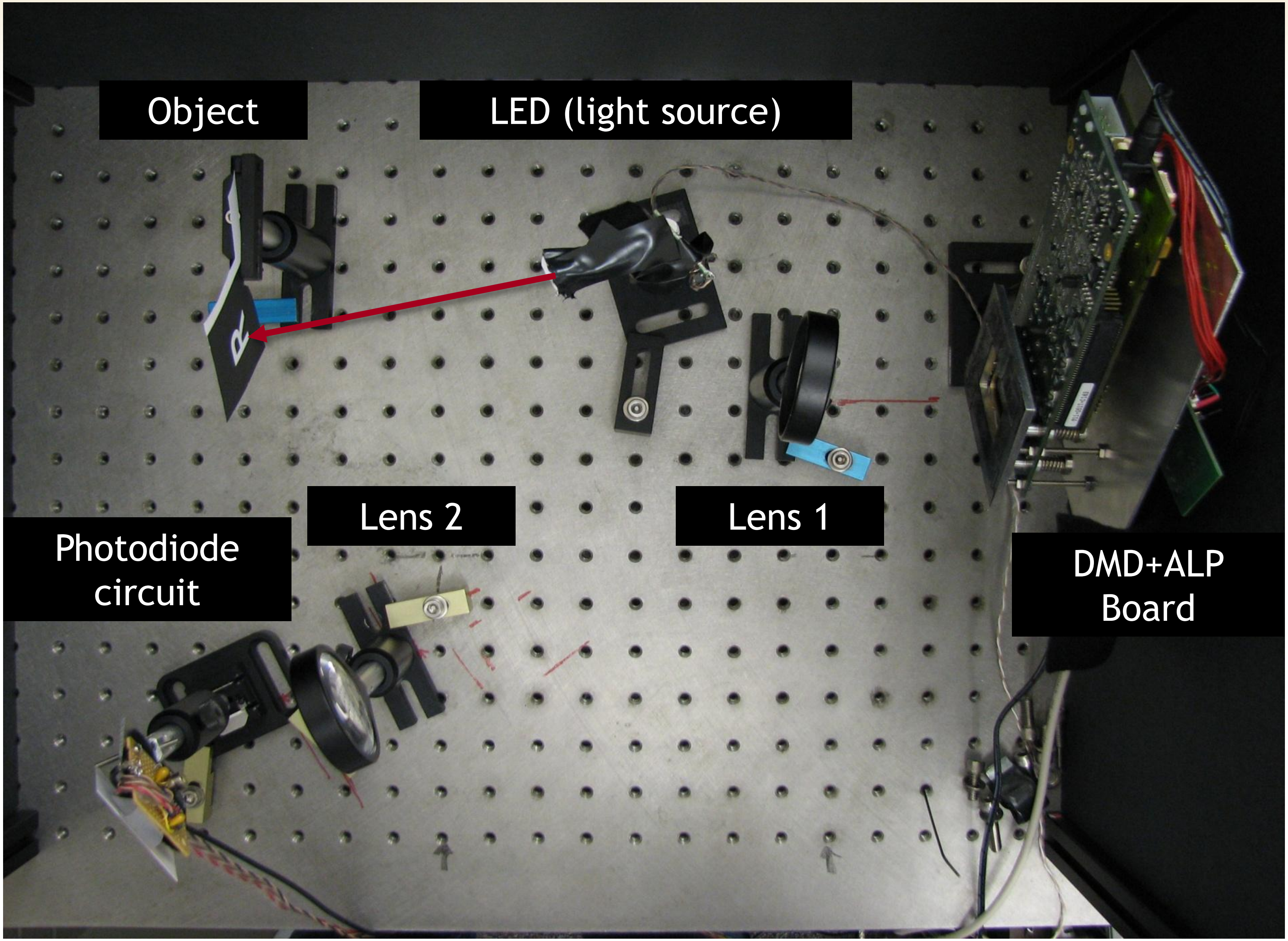




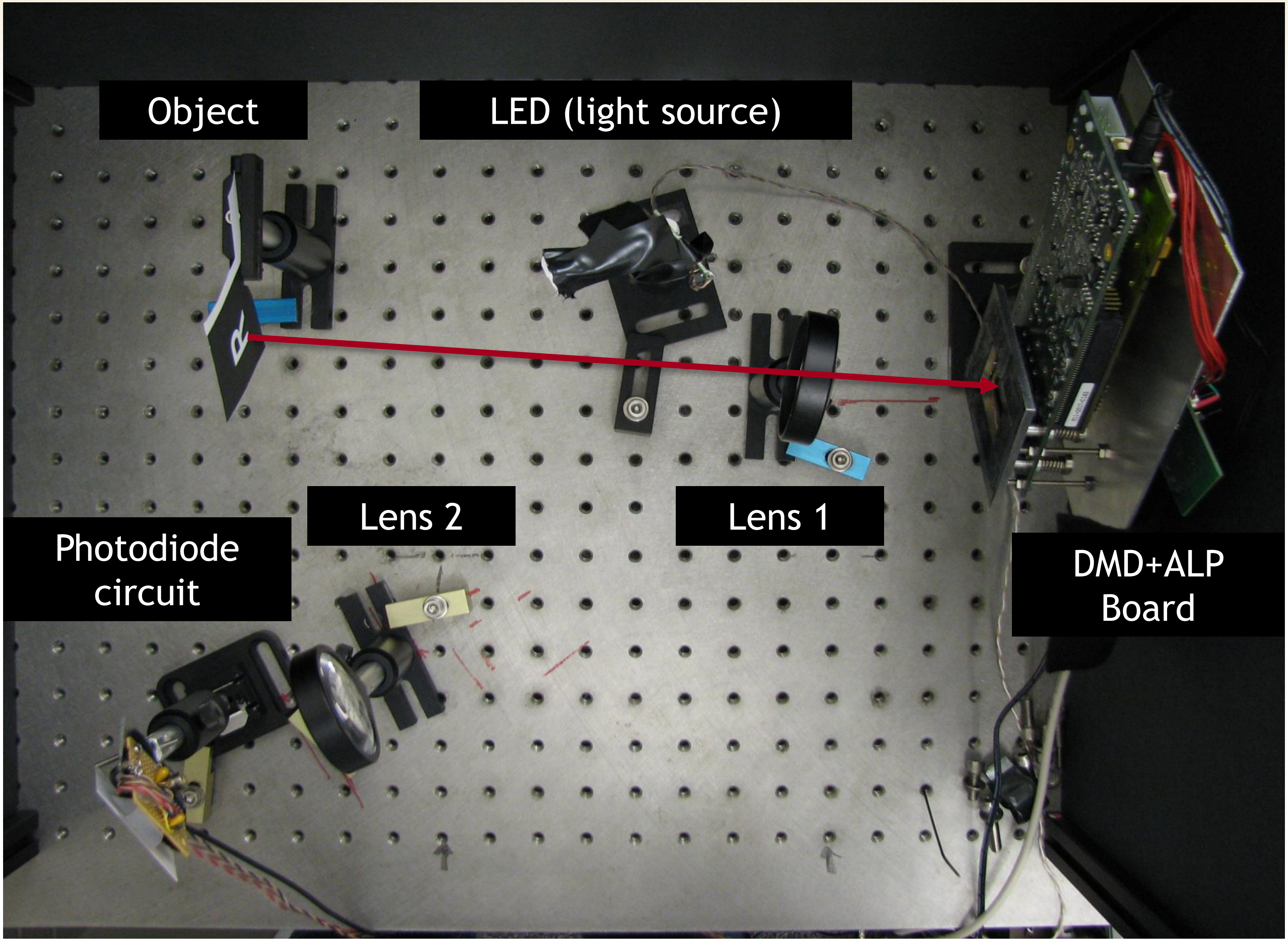
“Single-Pixel” Camera



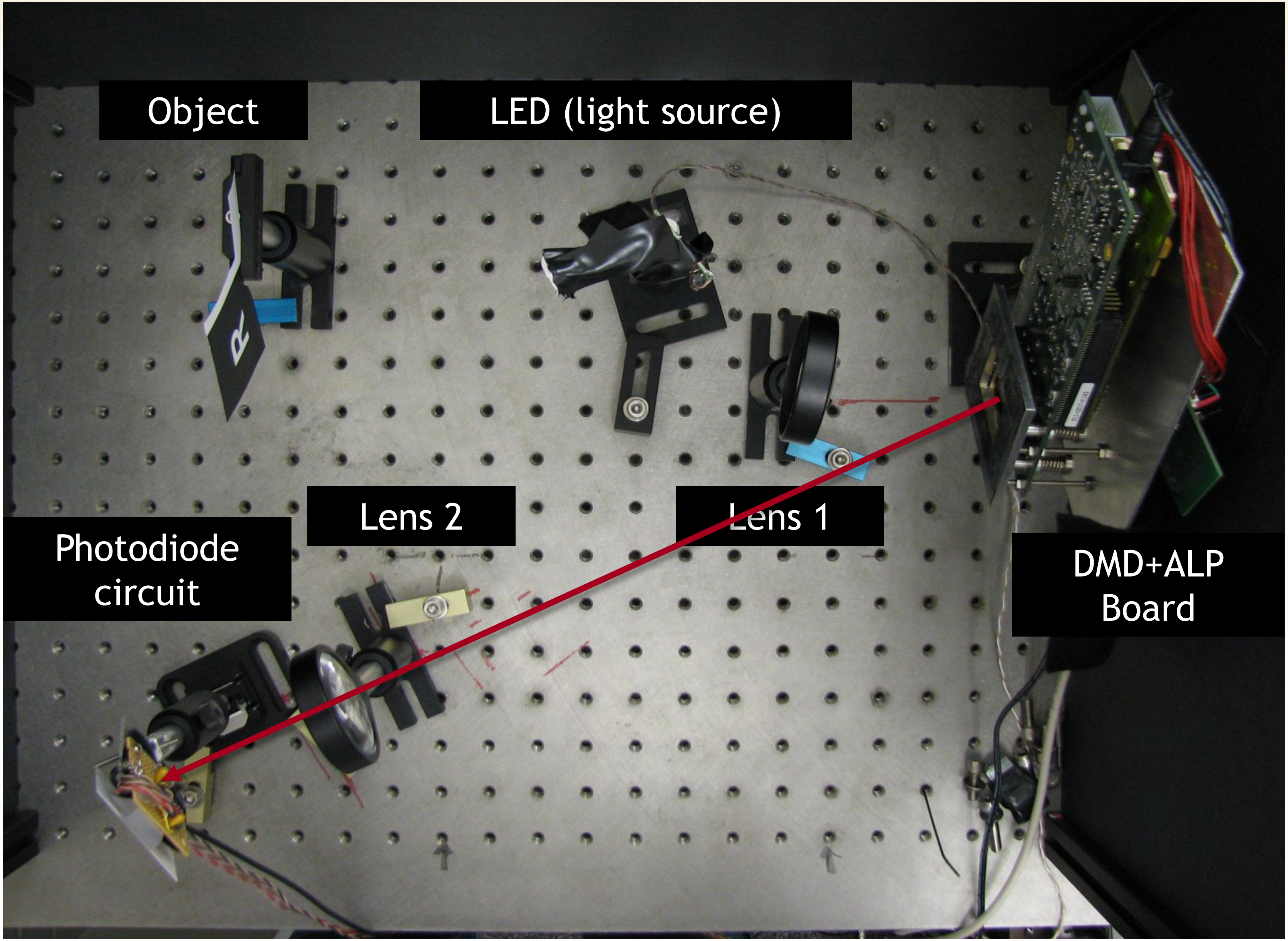
“Single-Pixel” Camera



“Single-Pixel” Camera



“Single-Pixel” Camera



Slashdot

News for Nerds. Stuff that matters.

oops, crash, seven million years bad luck !?!

I can’t wait to take this on my next vacation

This is me skydiving

.

This is me swimming with dolphins

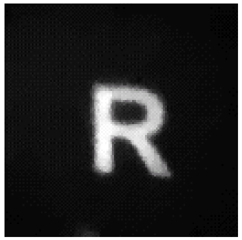
.

This is me at the grand canyon

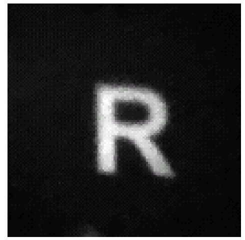
.

First Images

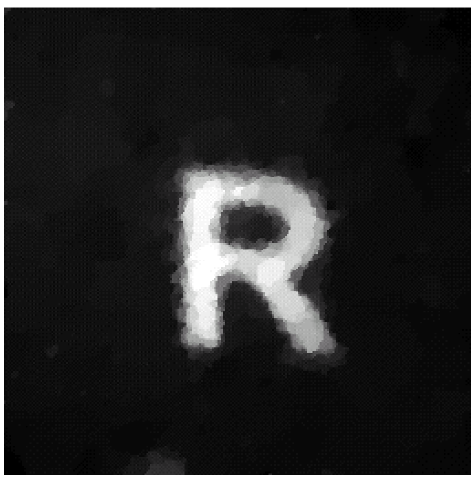
Original



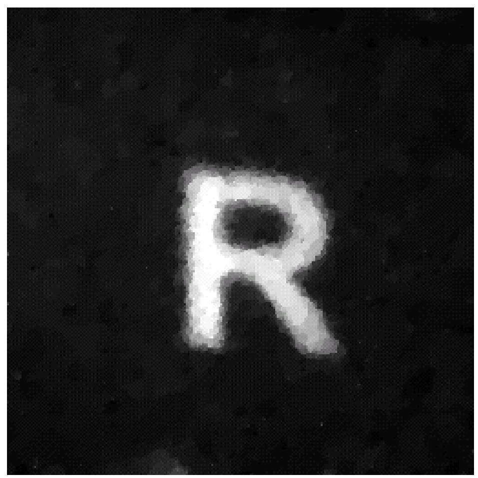
16384 Pixels
1600 Measurements
(10%)



16384 Pixels
3300 Measurements
(20%)



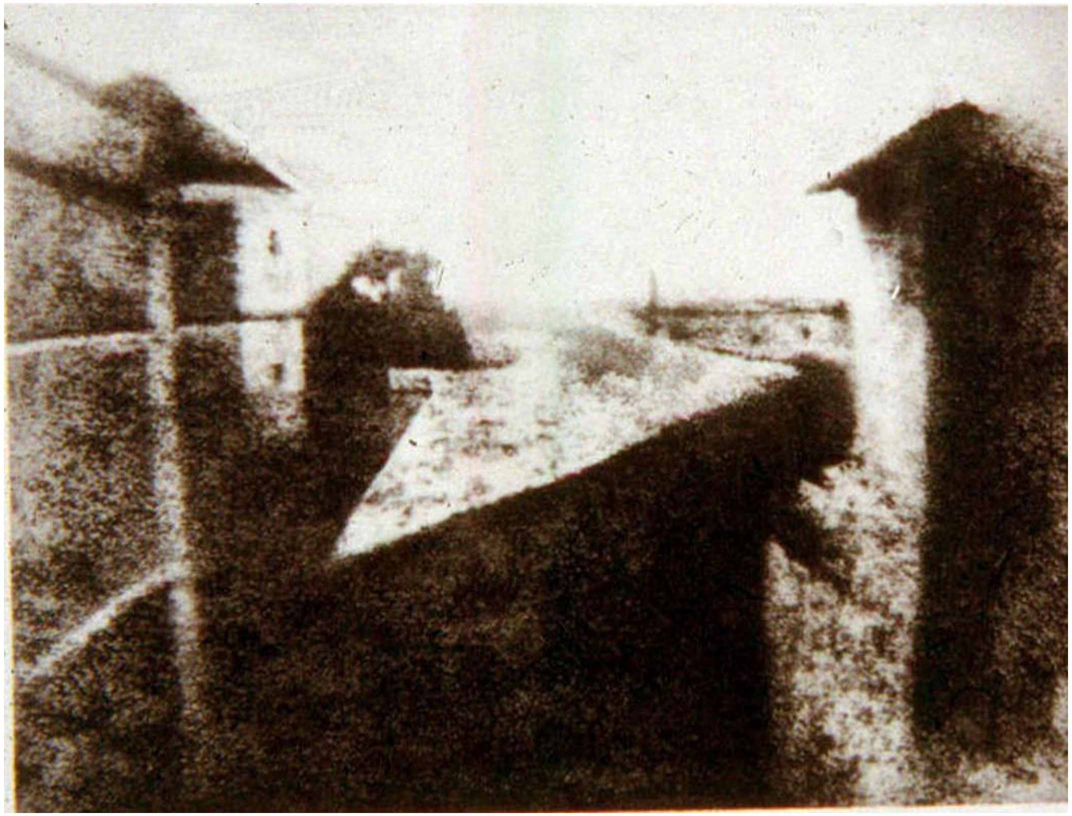
65536 Pixels
1300 Measurements
(2%)



65536 Pixels
3300 Measurements
(5%)

World’s First Photograph

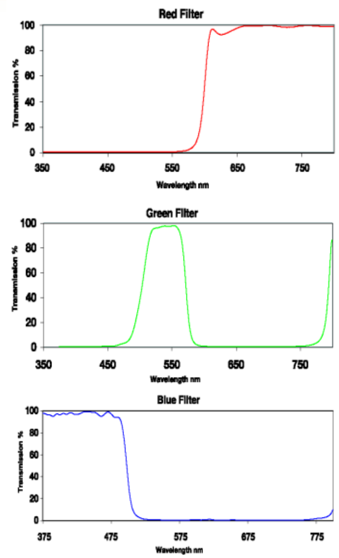
- 1826, Joseph Niepce
- Farm buildings and sky
- 8 hour exposure

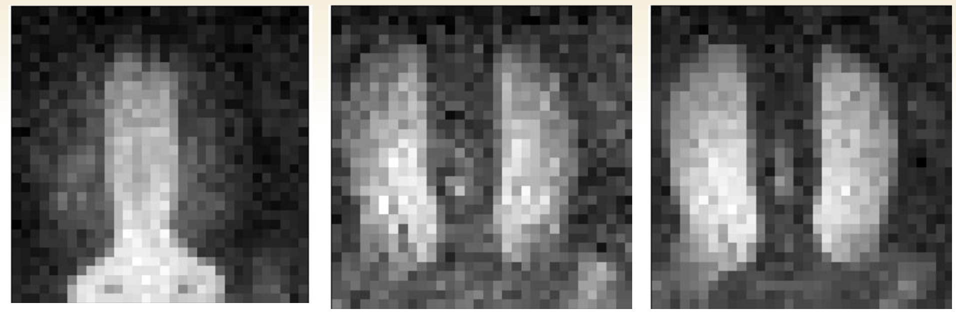


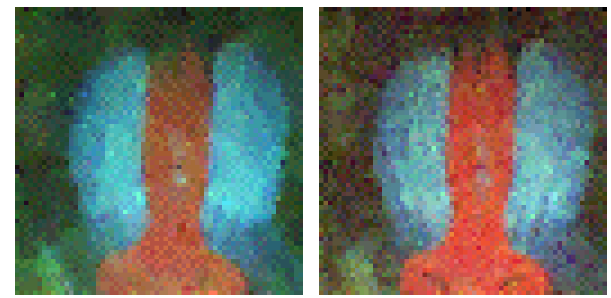
Color Imaging

Merging RGB channels

Color Filter Wheel





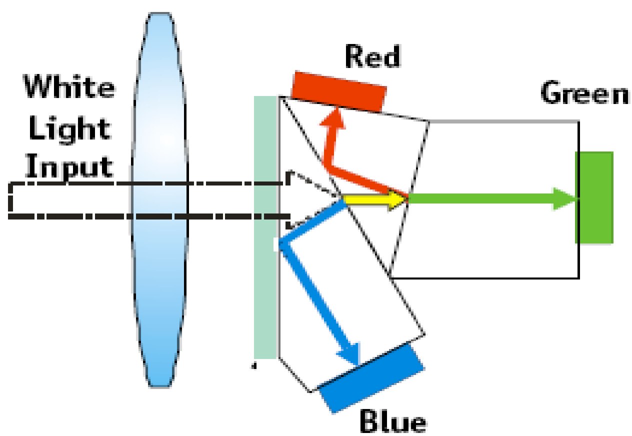


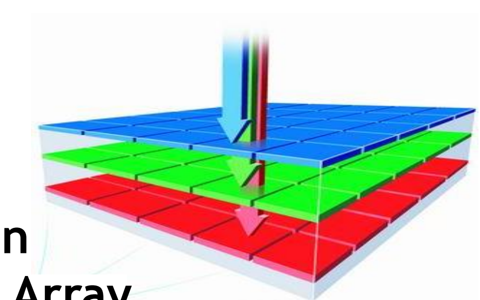
4096 Pixels
800 (20%)
Measurements

4096 Pixels
1600 (40%)
Measurements

Two strategies:

1. Prism assembly
2. Layered sensors
   (ala Foveon)

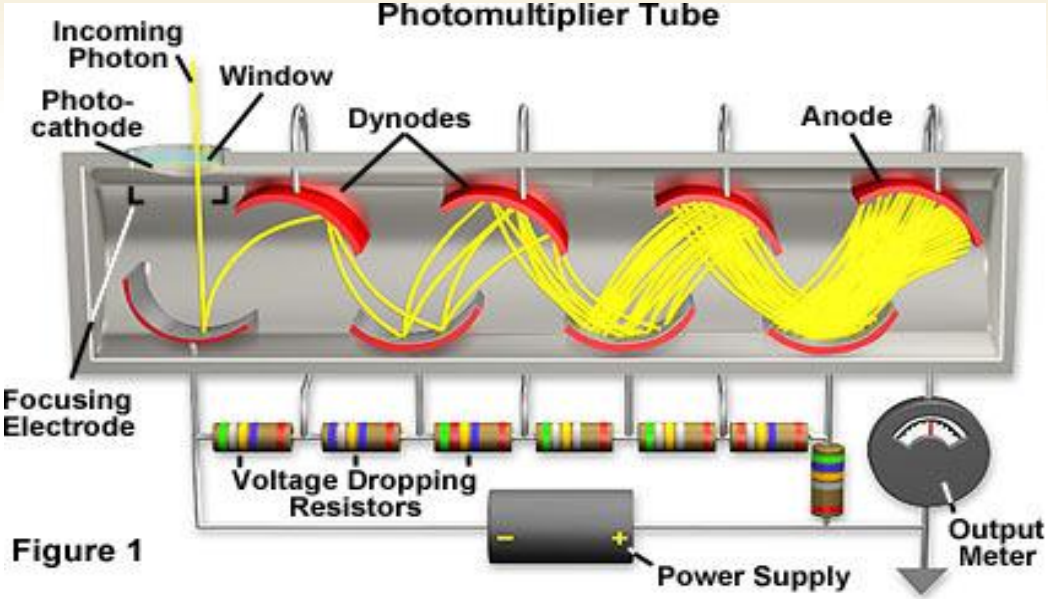


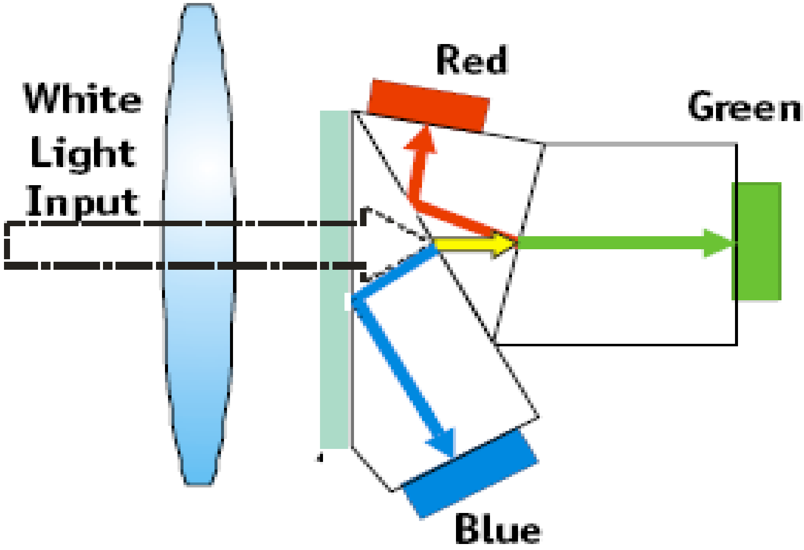


Foveon Image Array

“Single-Pixel” Camera

Low-Light Imaging with PMT

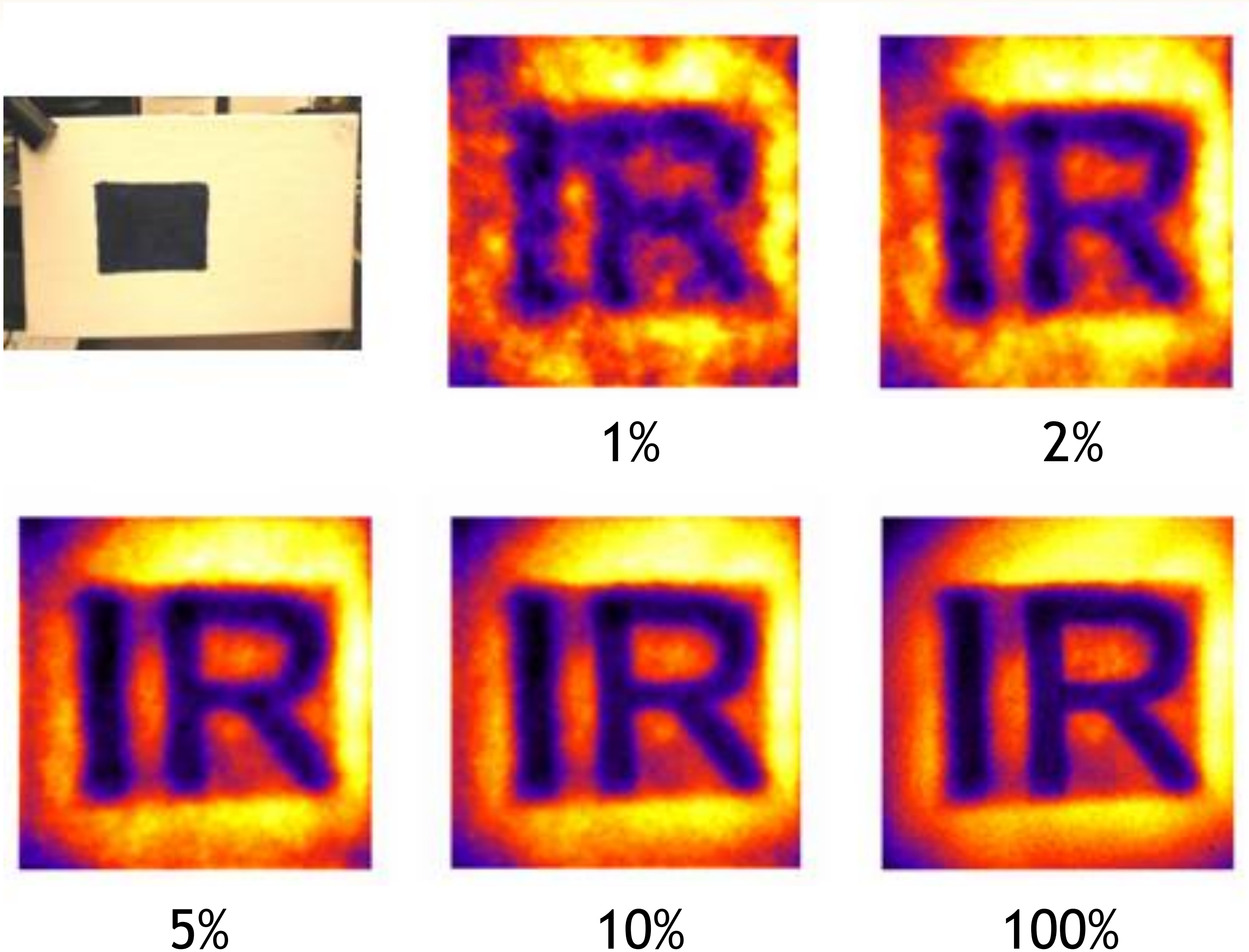




True color low-light imaging:

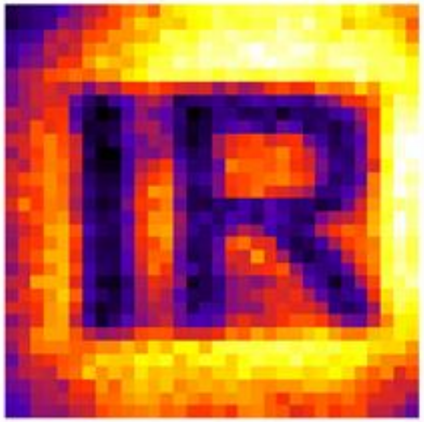
256 x 256 image with 10:1 compression

IR Imaging

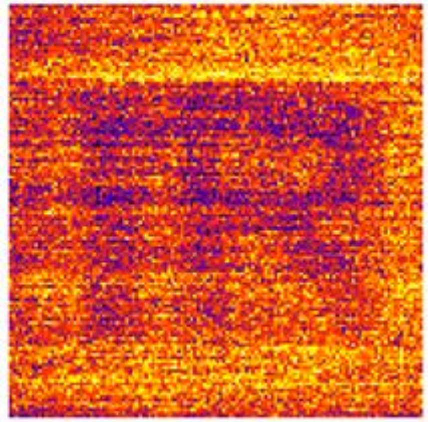


IR Imaging

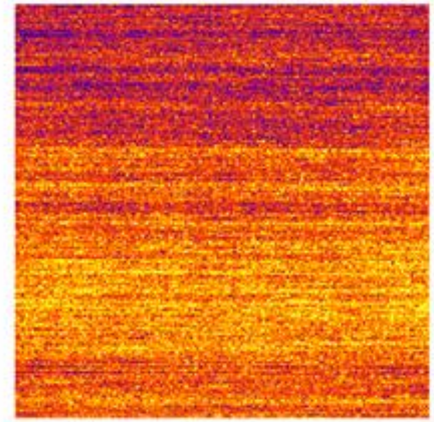
Raster scans: Light from only one pixel



32×32



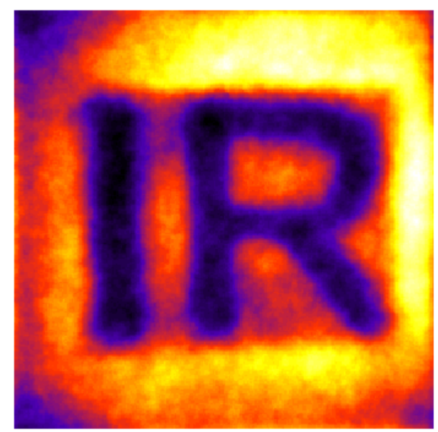
128×128



256×256

Compressive sensing:

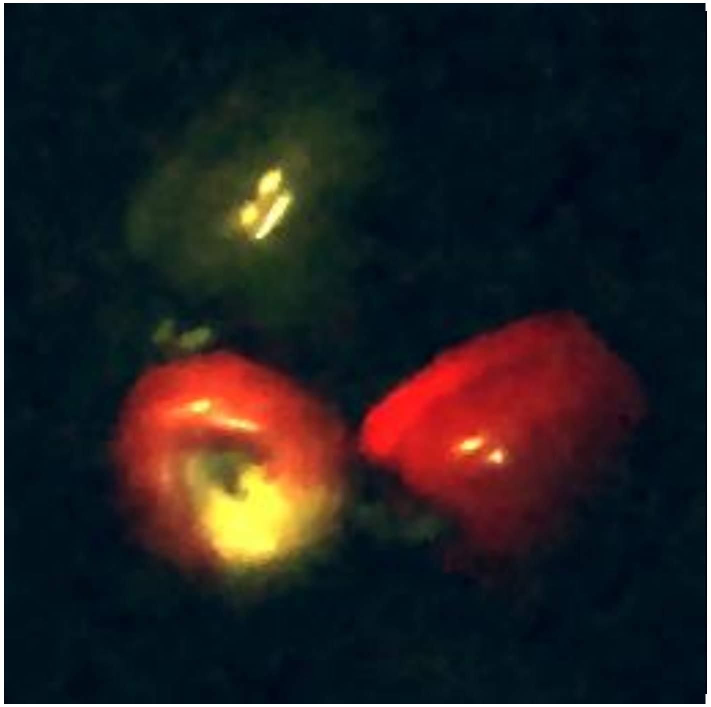
Light from half the pixels



256×256

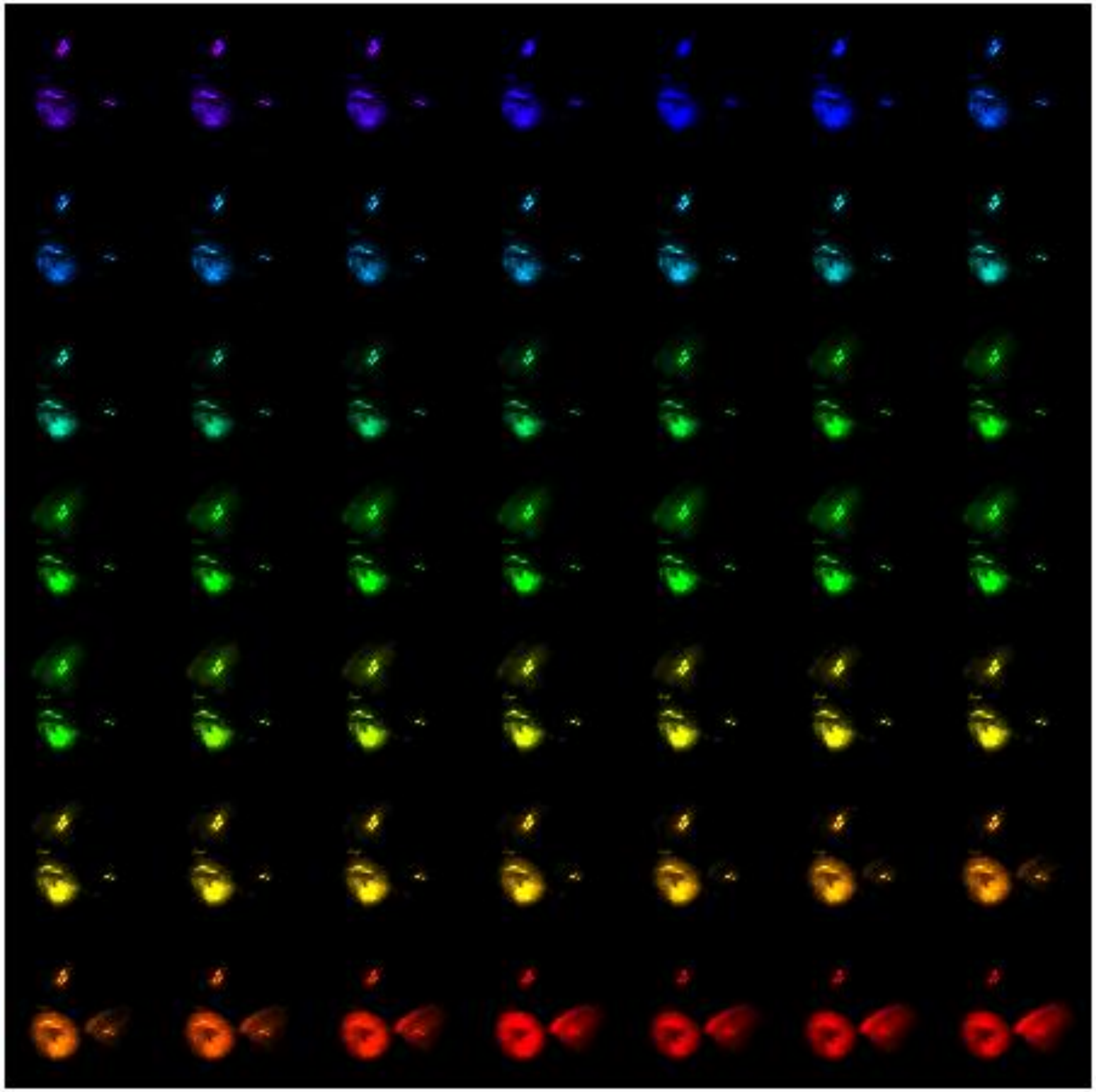
Hyperspectral Imaging

Sum of all bands

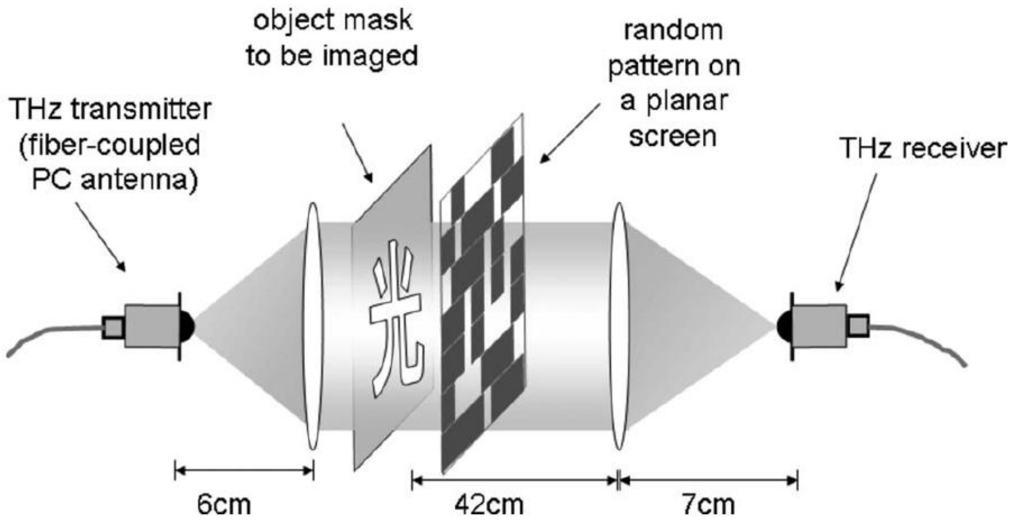


Real target

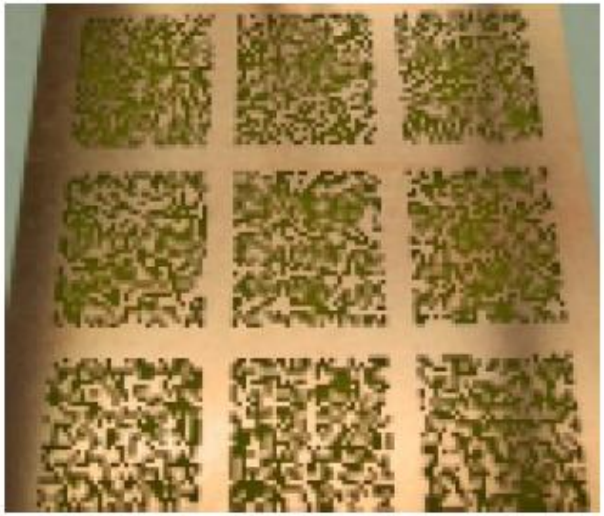
Hyperspectral Imaging

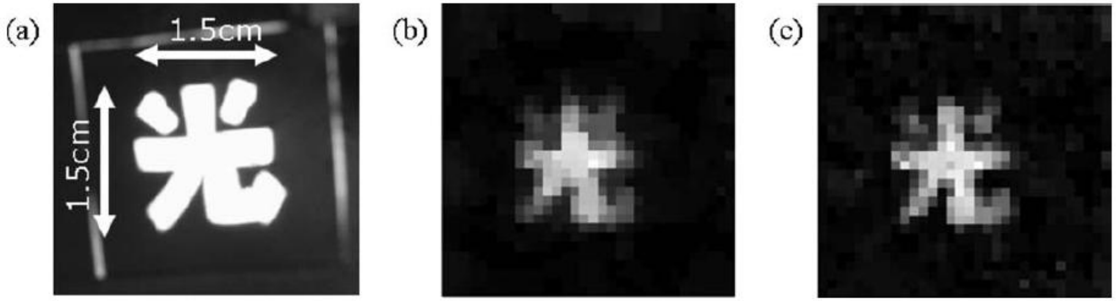


THz Imaging



32 x 32 PCB masks





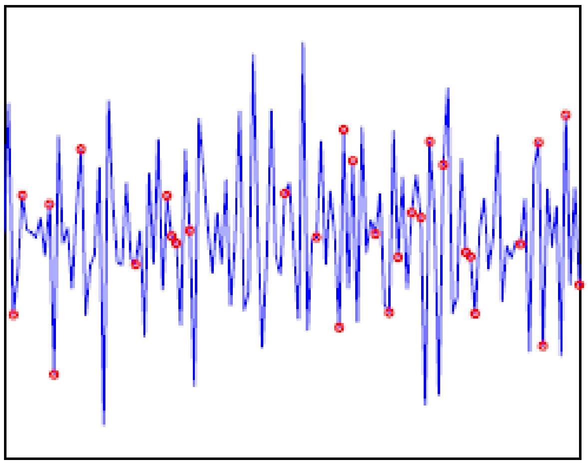
Object mask

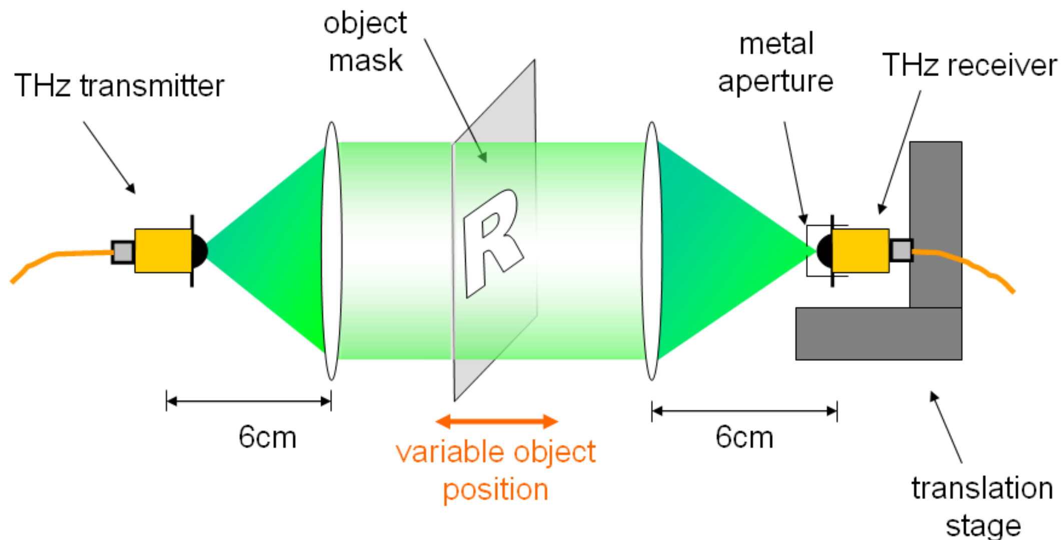
300 measurements

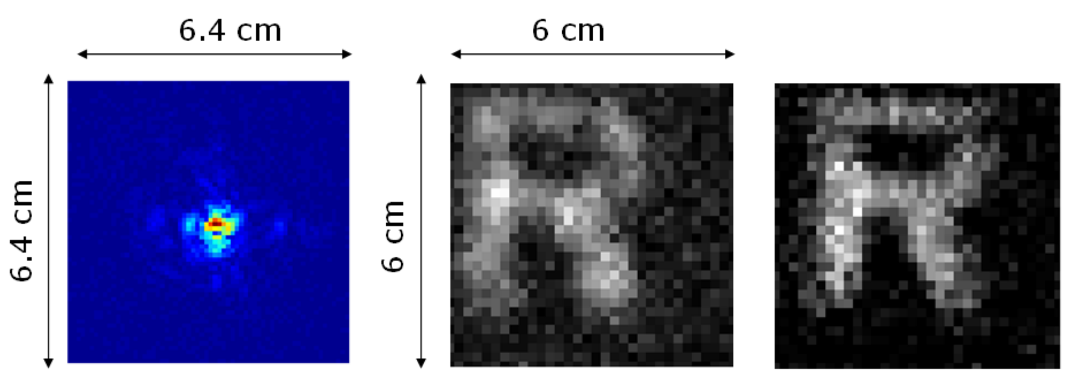
600 measurements

[Mittleman Group, Rice University]

THz Imaging: Sampling in Fourier







Fourier Transform
of object
(Magnitude-only)

CPR Reconstruction
(4096 measurements)

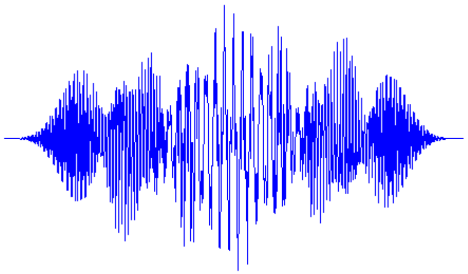
CSPR Reconstruction
(1000 measurements)

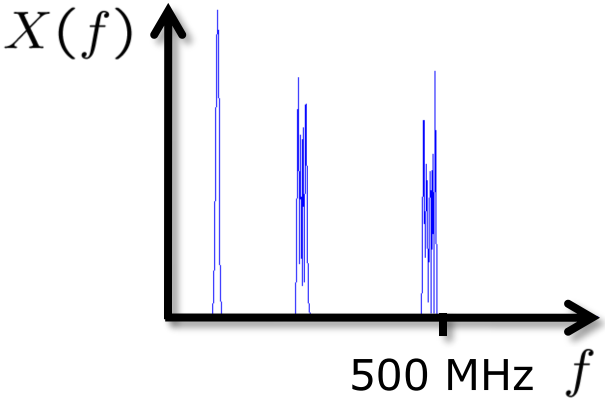
[Mittleman Group, Rice University]

Compressive ADCs

DARPA "Analog-to-information" program:

Build high-rate ADC for signals with sparse spectra

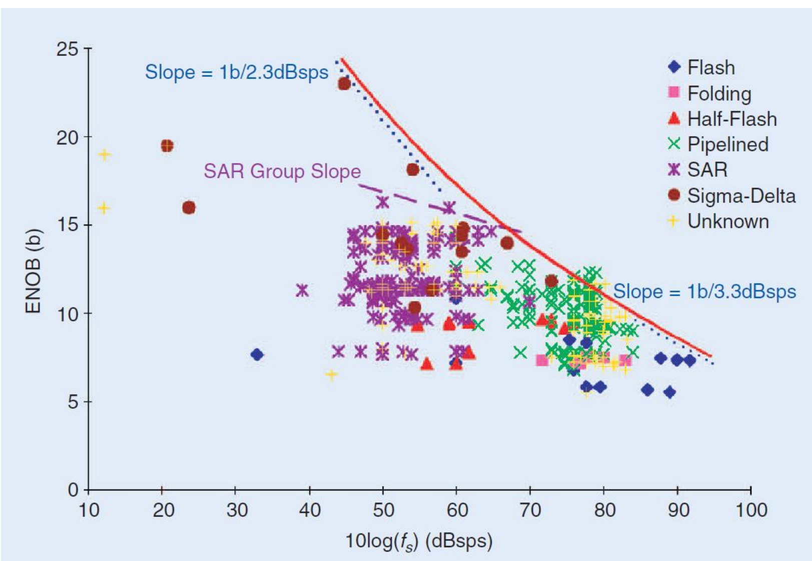


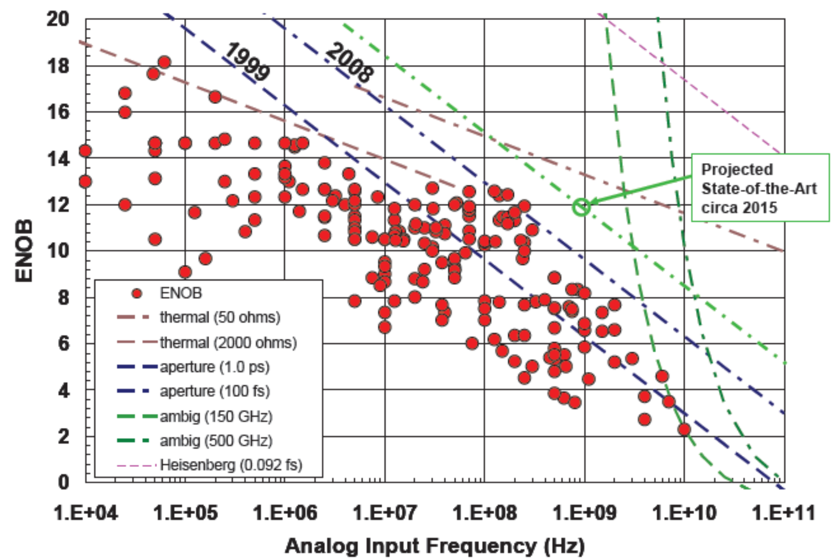


Compressive ADCs

DARPA "Analog-to-information" program:

Build high-rate ADC for signals with sparse spectra





[Le – 2005; Walden – 2008]

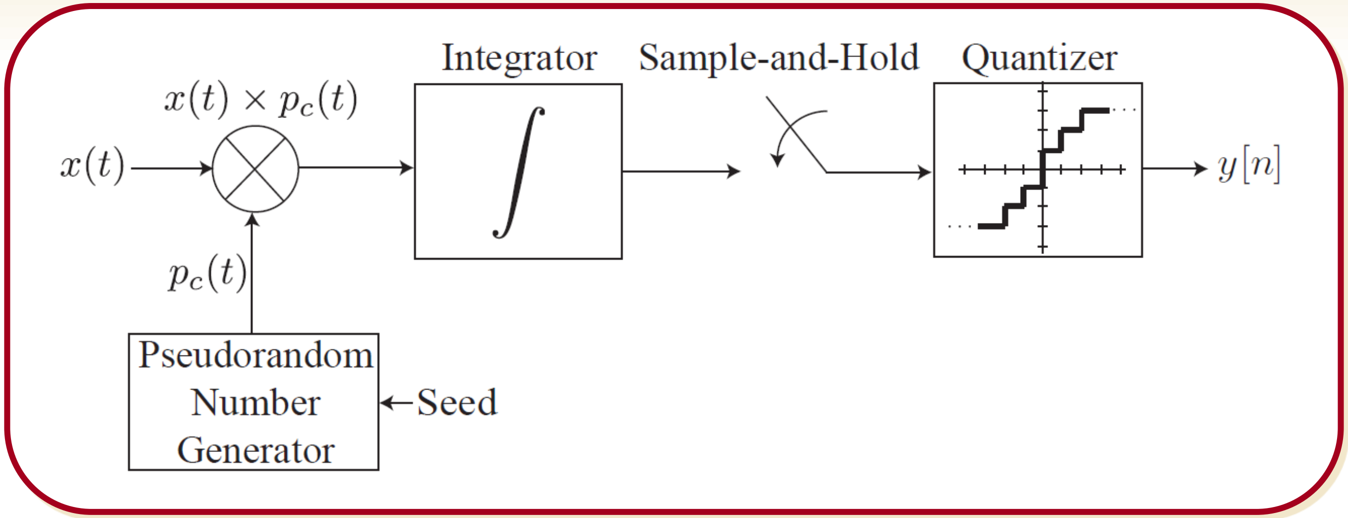
Analog-to-*Information* Conversion

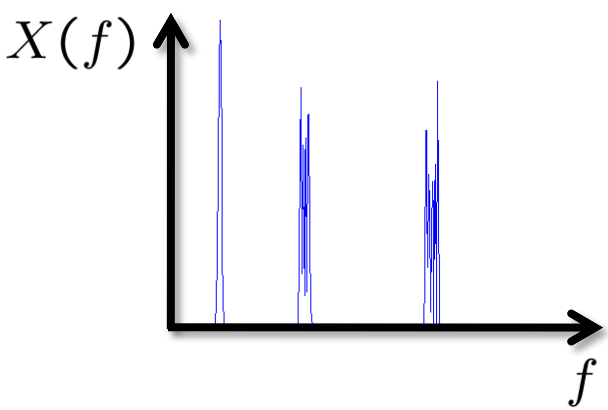
- Many applications – particularly in RF – have hit an ADC performance ***brick wall***
  - limited bandwidth (# Hz)
  - limited dynamic range (# bits)
  - deluge of bits to process downstream

- "Moore's Law" for ADC's: doubling in performance only every 6 years

- Inspiration from CS:
  - "analog-to-information" conversion

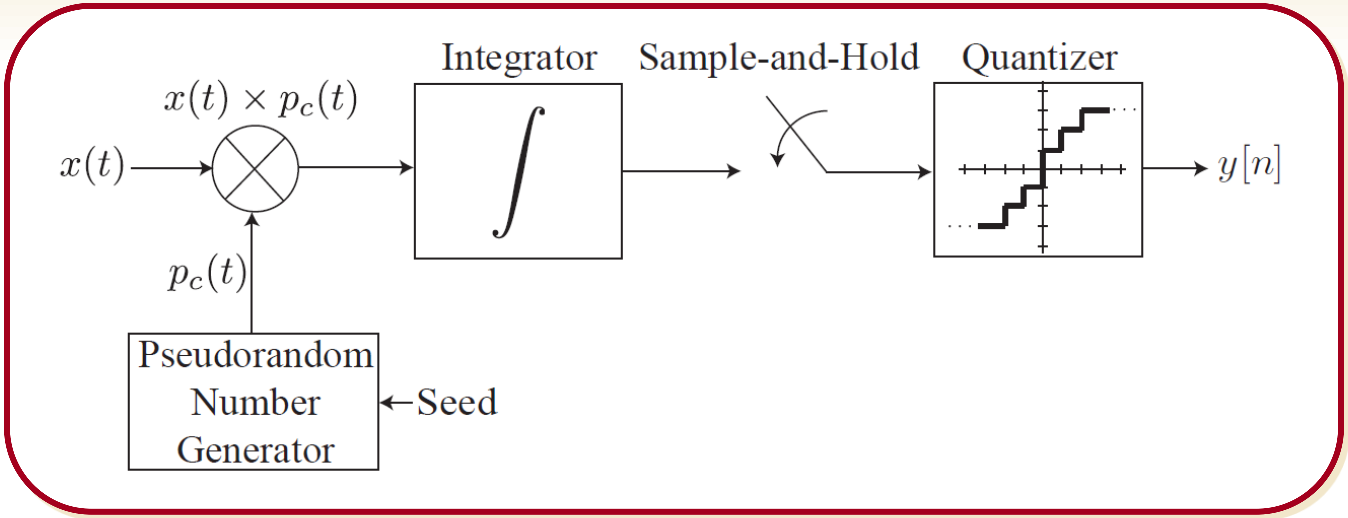
Random Demodulator

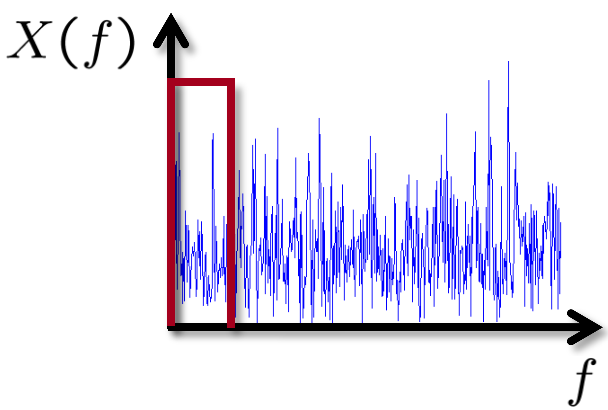




[Tropp, Laska, Duarte, Romberg, and Baraniuk – 2010]

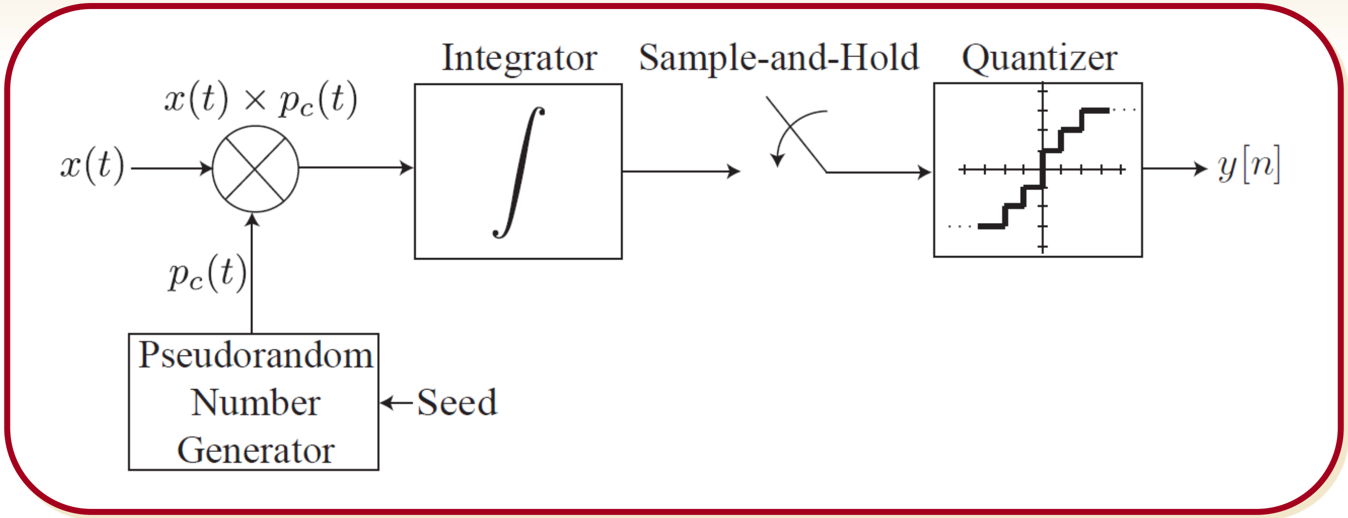
Random Demodulator

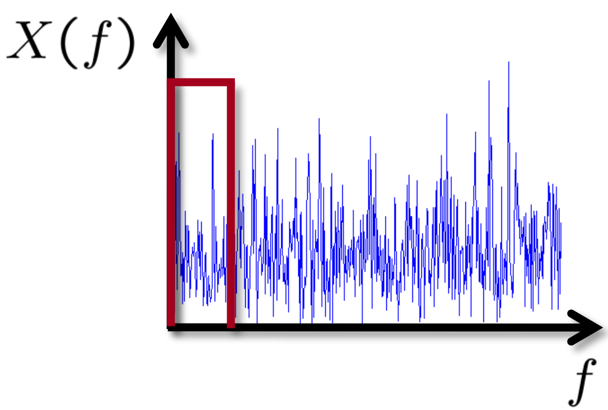


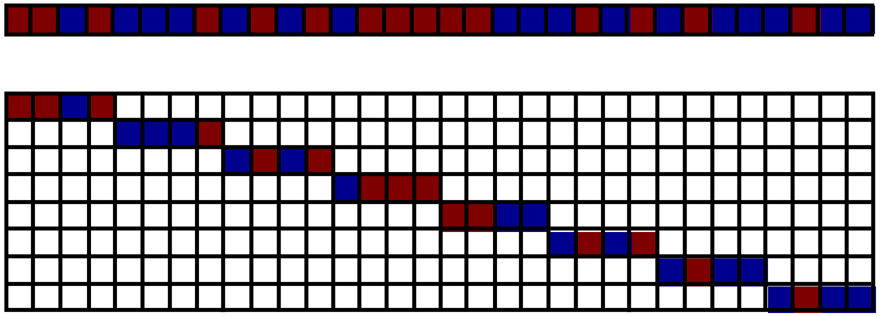


[Tropp, Laska, Duarte, Romberg, and Baraniuk – 2010]

Random Demodulator

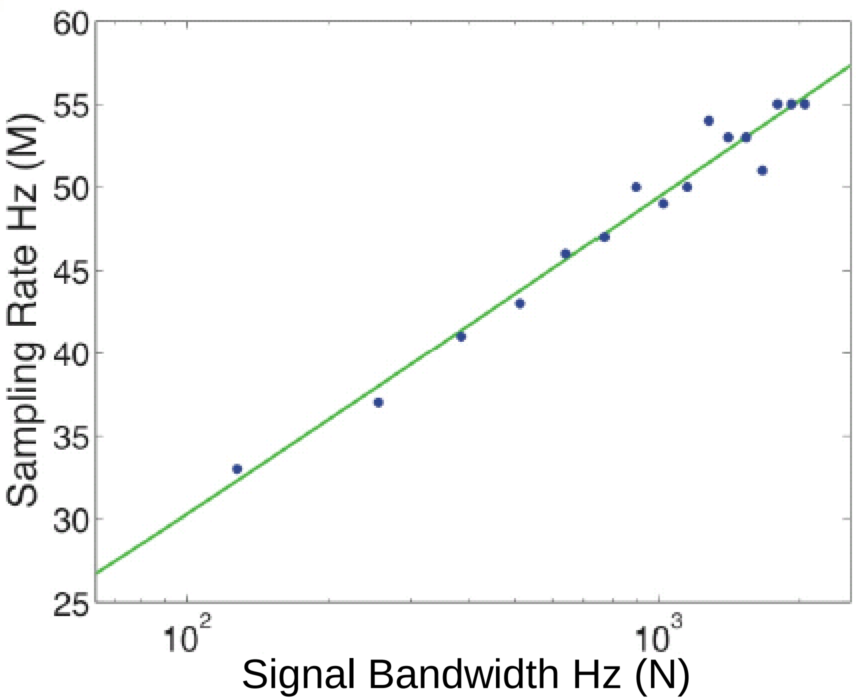


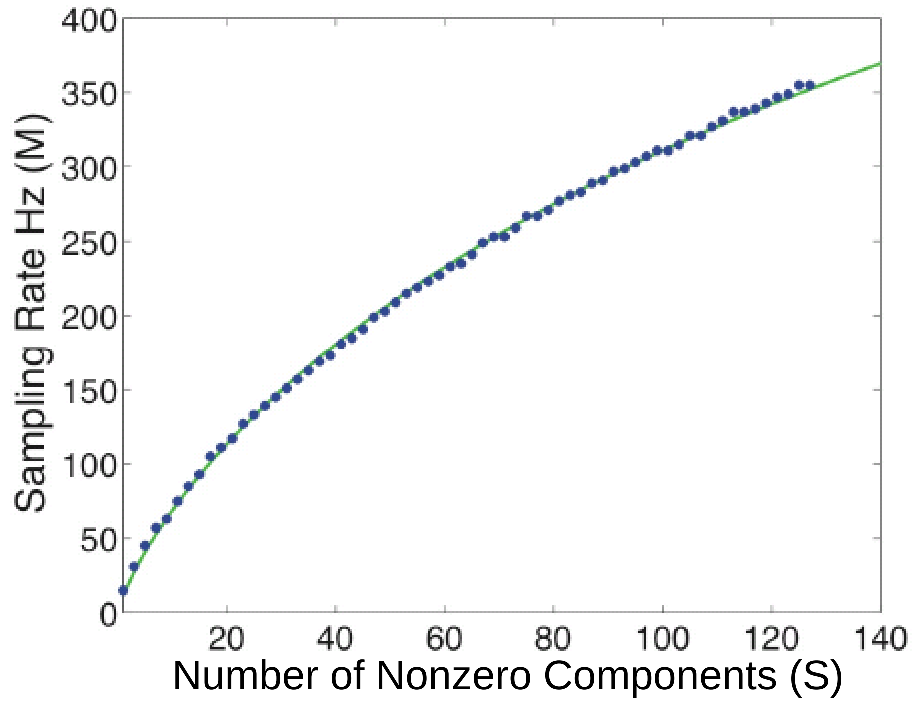




[Tropp, Laska, Duarte, Romberg, and Baraniuk – 2010]

Empirical Results





$$1.69S\log(N/S+1)+4.51$$

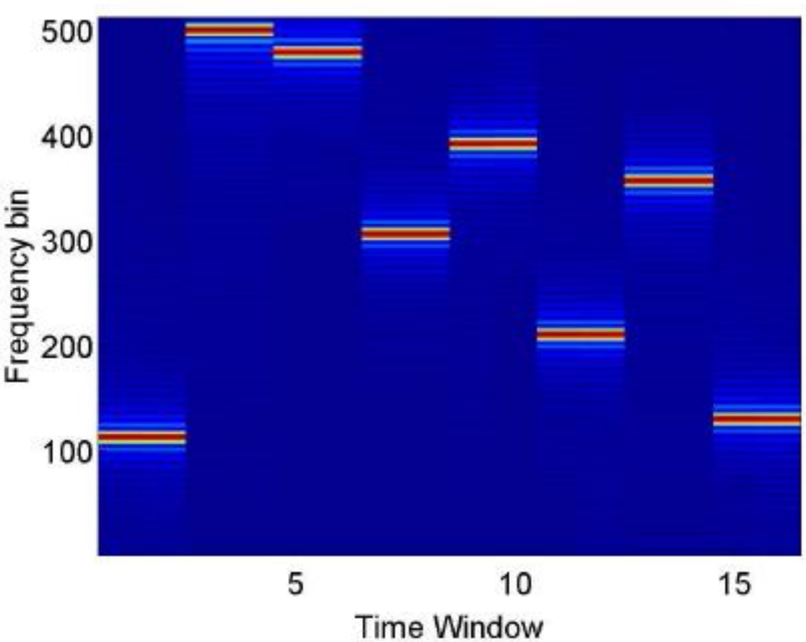
$$1.71S\log(N/S+1)+1$$

$$M \approx 1.7S\log(N/S+1)$$

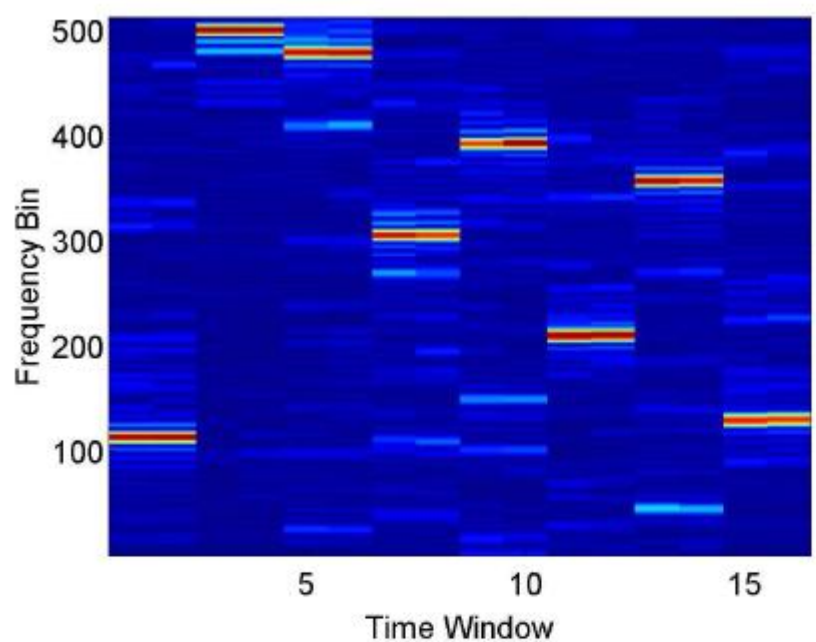
[Tropp, Laska, Duarte, Romberg, and Baraniuk – 2010]

Example: Frequency Hopper

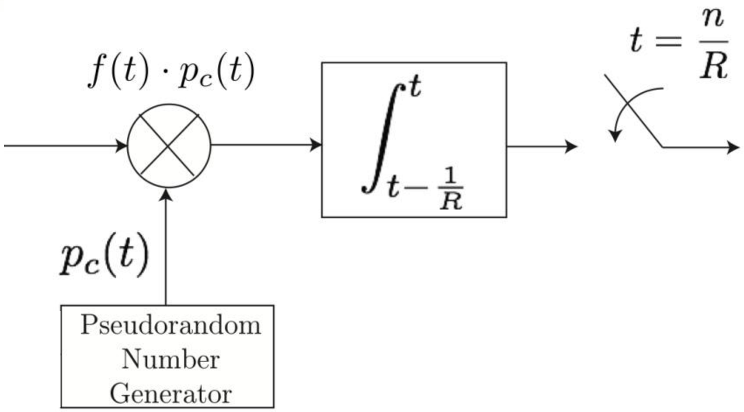
Nyquist rate sampling



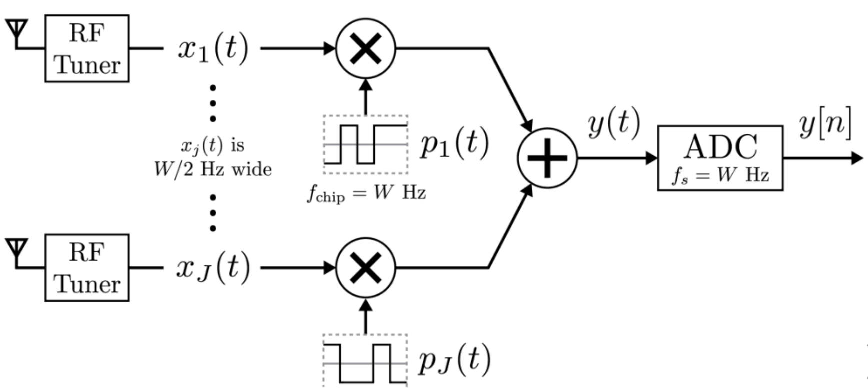
20x sub-Nyquist sampling



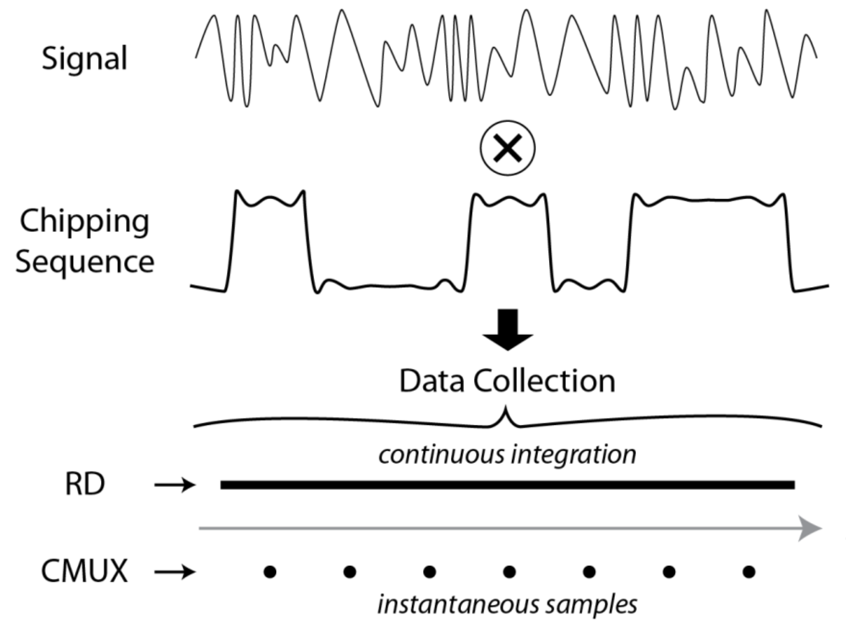
Compressive Multiplexor



Random Demodulator



Compressive Multiplexor

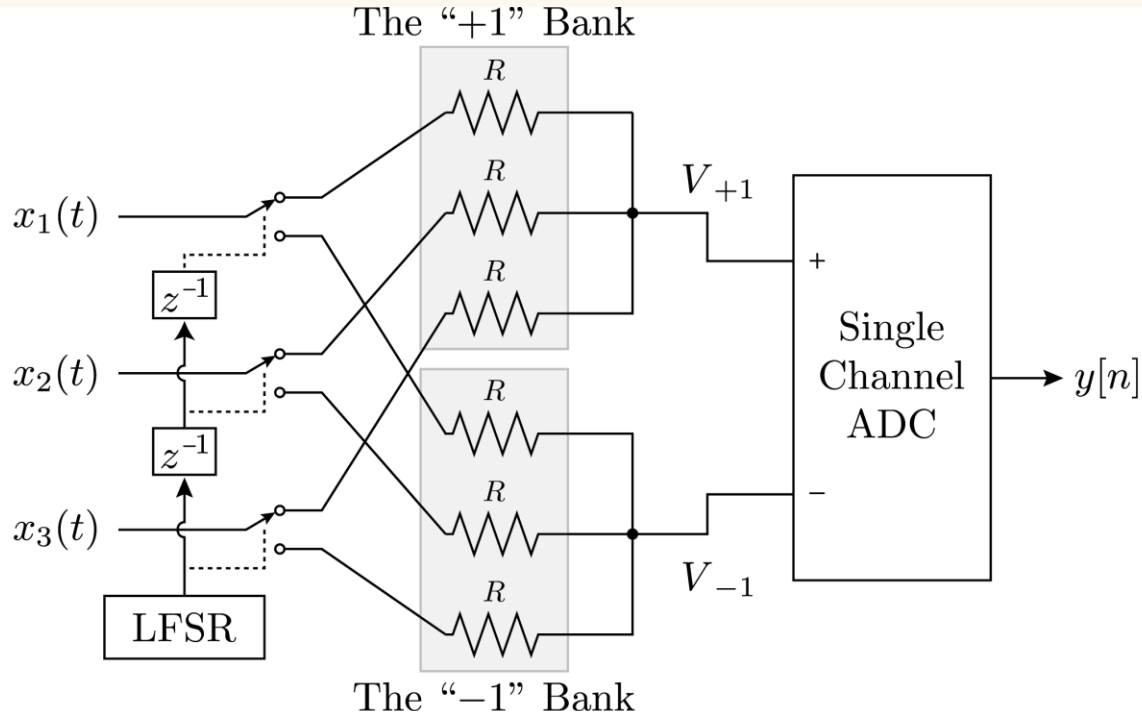


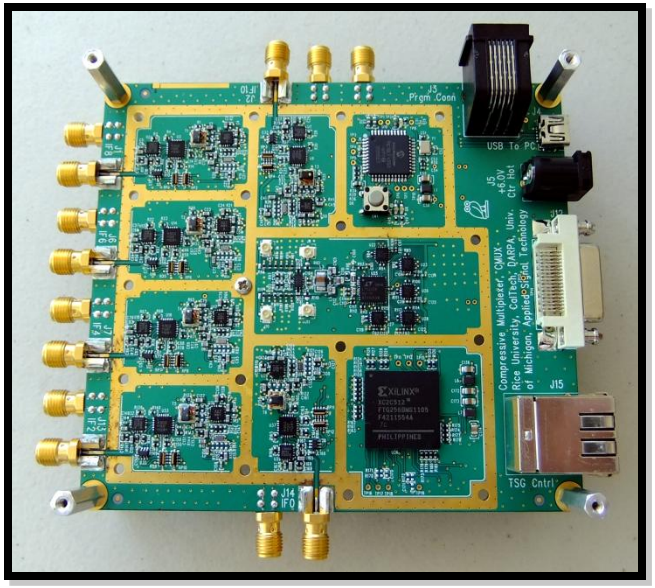
Compressive multiplexor is *more digital*

[Slavinsky, Laska, D, and Baraniuk – 2011]

Compressive Multiplexor in Hardware

- Boils down to:
  - 1 LFSR
  - J switches
  - $2J$ resistors
  - 2 op amps
  - 1 low-rate ADC





1.1mm x 1.1mm ASIC on its way!

[Slavinsky, Laska, D, and Baraniuk – 2011]

Compressive ADCs: Challenges Ahead

- Calibration!
  - you must know Φ to recover (or do anything else)
  - big challenge for all approaches
  - can often be mitigated by certain design choices

- Algorithms
  - recovery algorithms are much faster than a few years ago, but still can't operate in real time on GHz bandwidths
  - is recovery always necessary?

- Applications
  - noise can be a problem
  - good signal models are key

Compressive Sensors Wrap-up

- CS is built on a theory of *random measurements*
  - Gaussian, Bernoulli, random Fourier, fast JLT
  - stable, universal, democratic

- Randomness can often be built into real-world sensors
  - tomography
  - cameras
  - compressive ADCs
  - microscopes, sensor networks, DNA microarrays, radar, …

- OK, we can build these devices. What are they actually good for? When are they appropriate?