Architectures for Compressive Sampling

Justin Romberg
School of Electrical and Computer Engineering
Georgia Institute of Technology

Collaborators: *Emmanuel Candès, Joel Tropp, Richard Baraniuk, Marco Duarte, Jason Laska, Paul Voss, Paul Hasler, Ryan Robucci, Ken Chiu*

Compressive Sampling Linear Algebra

- High resolution (unknown) n-point signal x_0
- Small number of measurements

$$y_k = \langle x_0, \phi_k \rangle, \quad k = 1, \ldots, m \qquad \text{or} \qquad y = \Phi x_0$$

ϕ_k = "test function"

- Fewer measurements than degrees of freedom, $m \ll n$

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} \quad & \Phi & \quad \end{bmatrix} \begin{bmatrix} x_0 \\ \\ \\ \end{bmatrix}$$

- Compressive Sampling: for *sparse* x_0, we can "invert" certain Φ

Sparse Recovery

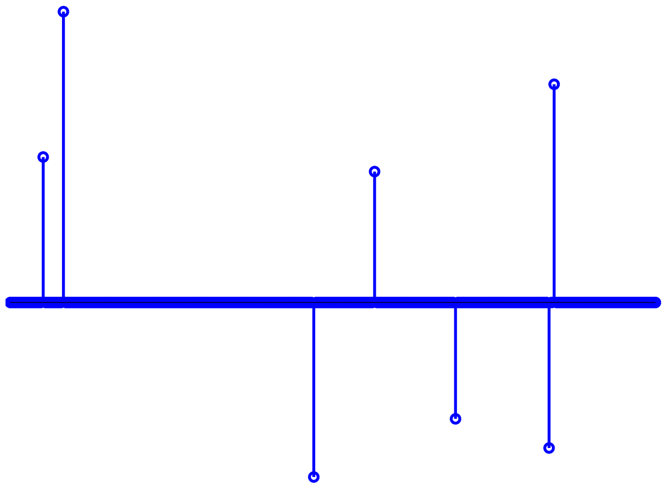
- Model: signal/image $\boldsymbol{x_0}$ is sparse in the $\boldsymbol{\Psi}$ domain
  (example: $\boldsymbol{x_0}$ is an image, $\boldsymbol{\Psi}$ is a wavelet transform)
- Acquisition: measure $\boldsymbol{y = \Phi x_0}$
- Recovery: solve

$$\min_{\boldsymbol{x}} \ \|\boldsymbol{\Psi}^{\boldsymbol{T}}\boldsymbol{x}\|_{\ell_1} \quad \text{subject to} \quad \boldsymbol{\Phi x = y}$$

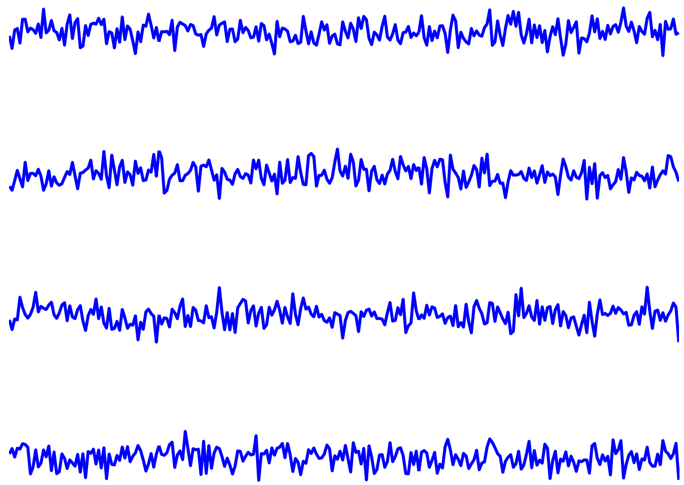
  Finds the sparsest signal which explains the measurements
- For which $\boldsymbol{\Phi}$ does this “work”?

Sensing Sparse Coefficients

S-sparse vector



incoherent measurements



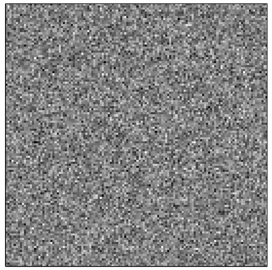
- Signal is local, measurements are global
- When the test functions are just iid random sequences, we can recover perfectly from (CT,D '06)

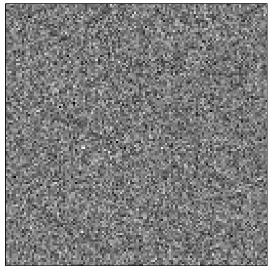
$$m \gtrsim S \cdot \log n \quad \text{measurements}$$

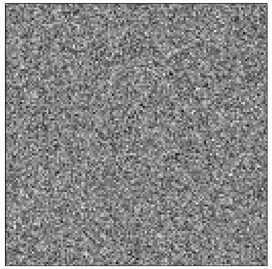
- In practice, it seems that

$$m \approx 5S \quad \text{measurements are sufficient}$$

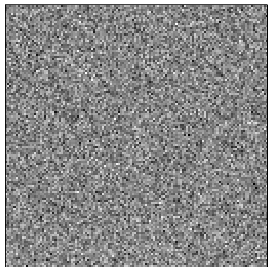
- Random sensing is a *universal* acquisition scheme

$y_1 = \Big\langle$  ,  $\Big\rangle$

$y_2 = \Big\langle$  ,  $\Big\rangle$

$y_3 = \Big\langle$  ,  $\Big\rangle$

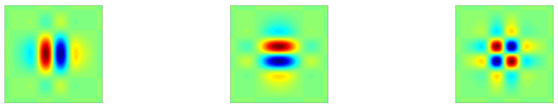
$\vdots$

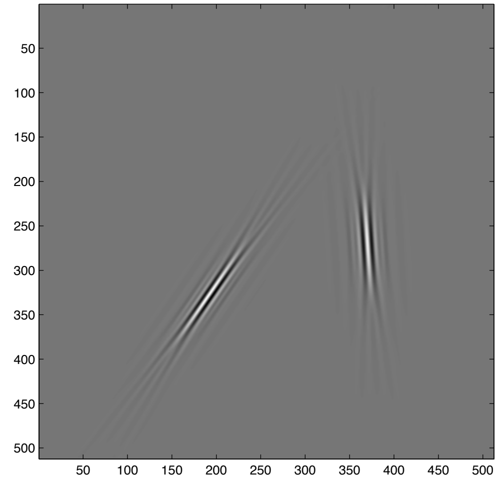
$y_m = \Big\langle$  ,  $\Big\rangle$

Representation vs. Measurements

- Image structure: *local, coherent*

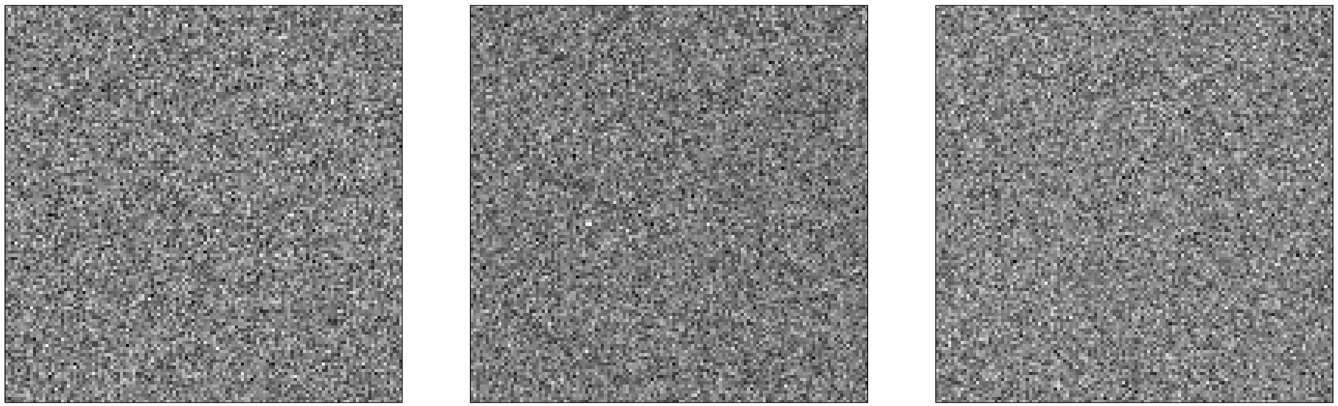
  Good basis functions:





- Measurements: *global, incoherent*

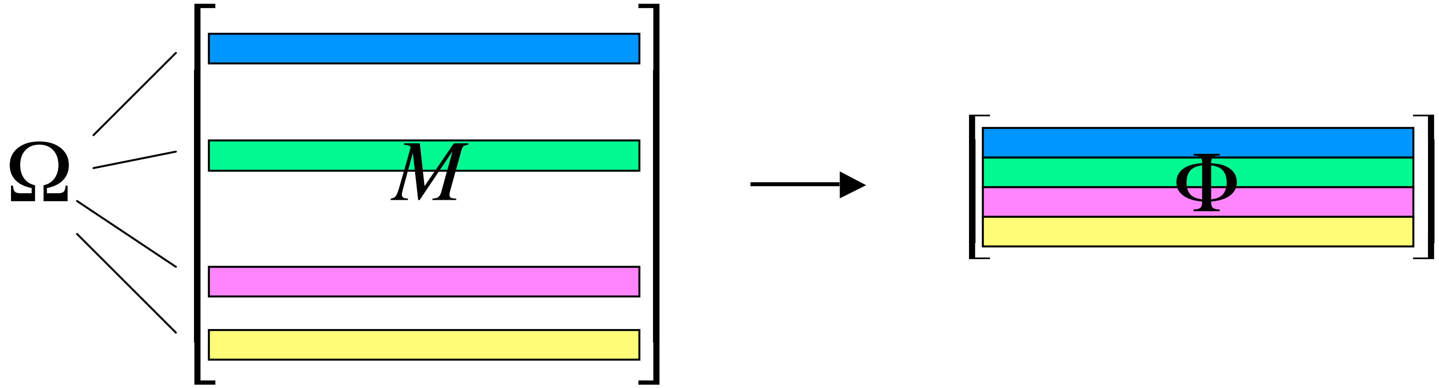
  Good test functions:



Problems with Random Measurements

- How do I compute with them?
  - Recovery algorithm will invariably require applying $\mathbf{\Phi}$ multiple times
- How do I take them physically?

Structured Recovery



- Sparsity basis Ψ (orthonormal)
- Measurement basis M (orthonormal)
- Ω = random subset of *sample locations*

  $y \;=\; M_\Omega x_0$

- Recover solving

$$\min_x \; \|\Psi^T x\|_{\ell_1} \quad \text{subject to} \quad \Phi x = y, \qquad \Phi = M_\Omega$$

- Perfect recovery for

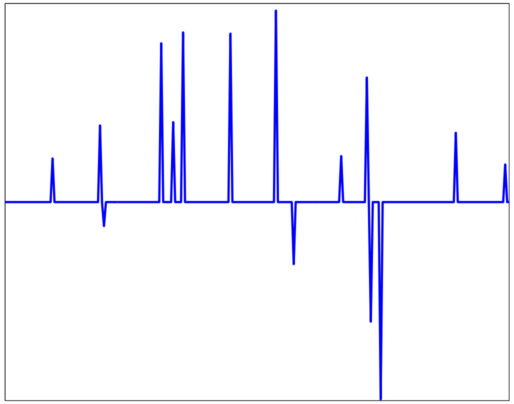
$$m \;\gtrsim\; \mu^2 \cdot S \cdot \log n$$

depends on *coherence* $1 \leq \mu^2 \leq n$ between M and Ψ (CR '07)

Examples of Incoherence

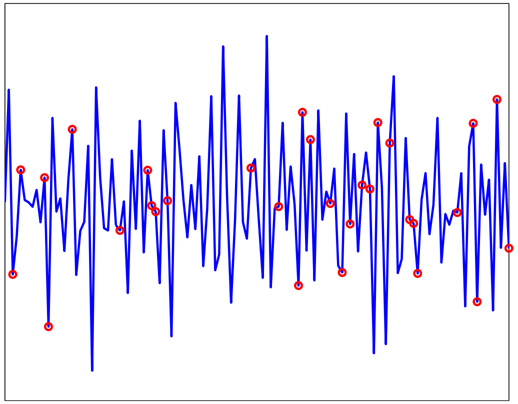
- Signal is sparse in time domain, sampled in Fourier domain

Time domain $x(t)$



S nonzero components

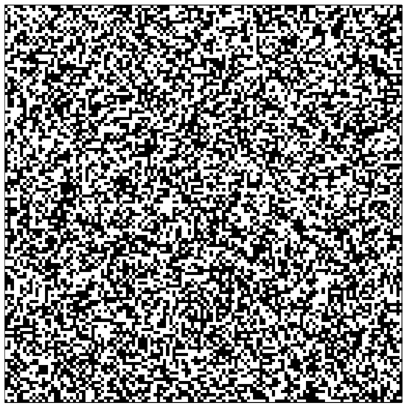
Frequency domain $\hat{x}(\omega)$



Measure m samples

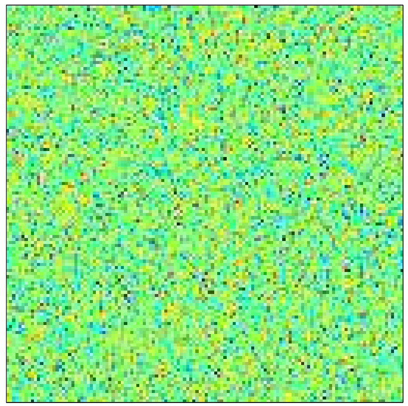
- Signal is sparse in wavelet domain, measured with noiselets (Coifman et. al)

example noiselet

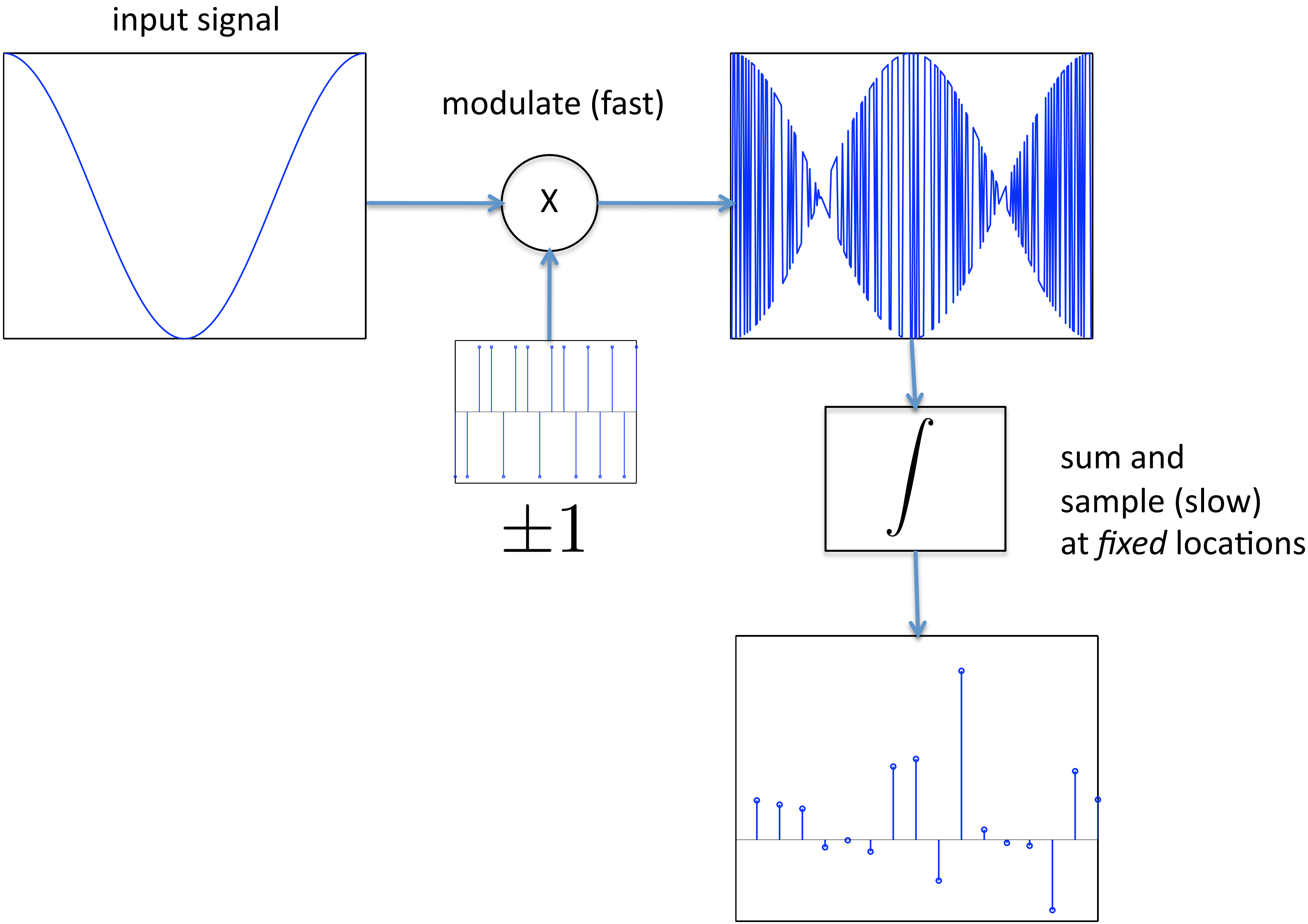


wavelet domain

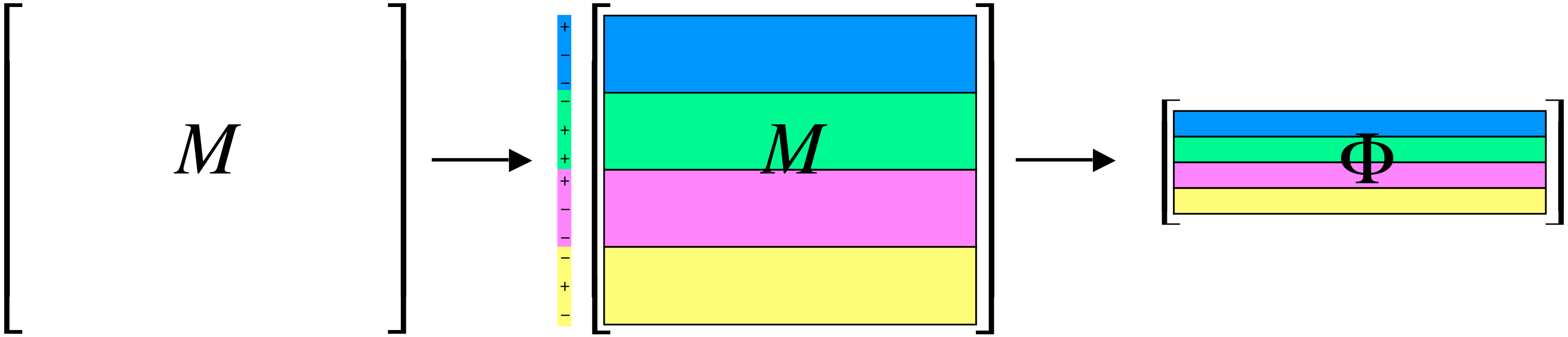
noiselet domain



Another way to downsample



Randomly Modulated Summation



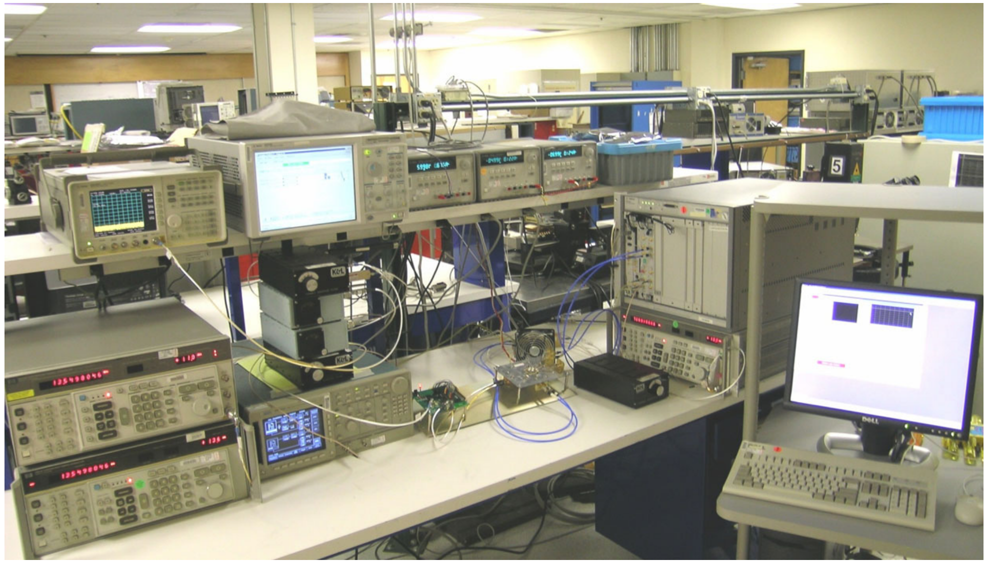
- Instead of choosing small set of random samples,
  "Downsample" by changing phases, breaking into chunks, and summing
- Measurement system M with coherence μ
- To form Φ:
  divide rows into m blocks, randomly *flip sign* of each row, *sum* over block
- S sparse x_0 can be recovered from

$$m \gtrsim \mu^2 \cdot S \cdot \log^2 n$$

measurements (TDLRB '08)

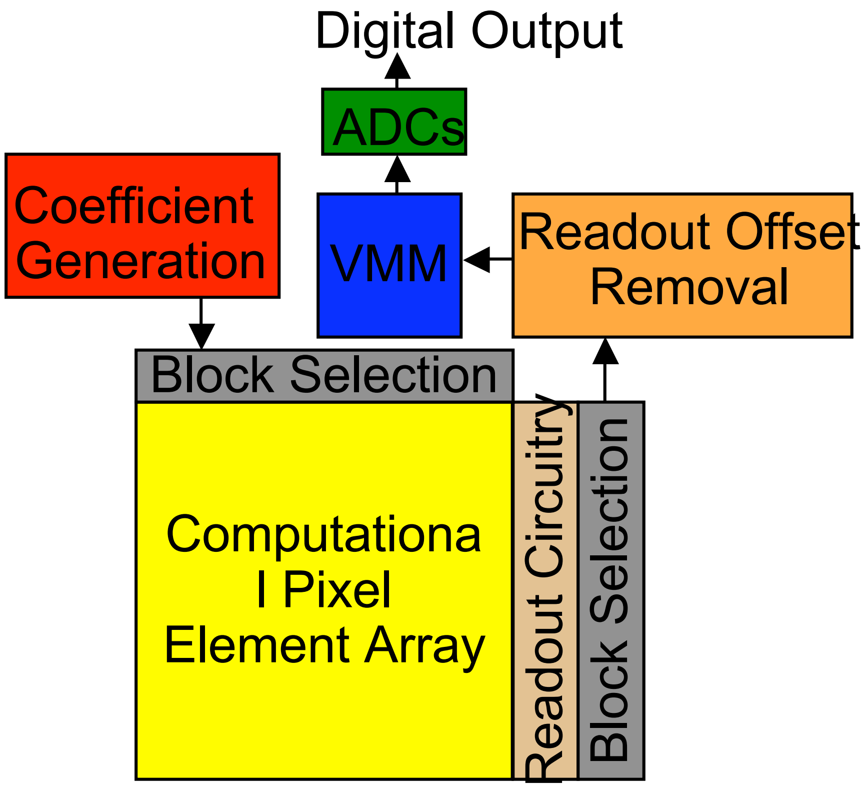
DARPA: “Analog to Information”

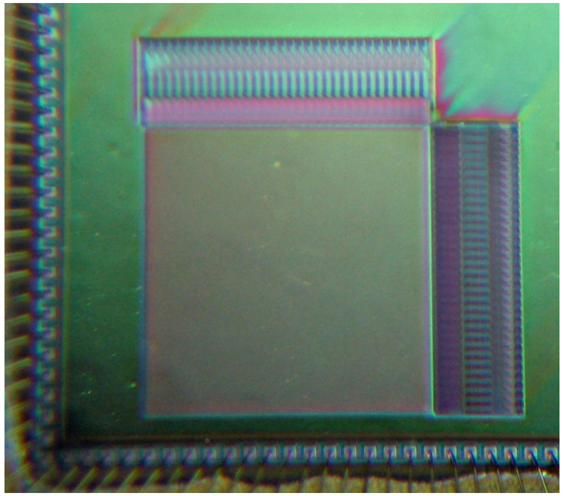
- Goal: reconstruct spectrally sparse signals with incredibly high freqs
  ADCs cannot run fast enough for Nyquist
- CS Architecture I: random non-uniform sampling
  Take standard ADC, clock it non-uniformly with “slow” average rate
- CS Architecture II: randomly modulated summation
  Modulate incoming pulse, integrate (high-speed but simple circuit), sample uniformly at slow rate
- Hardware implementation in progress ...

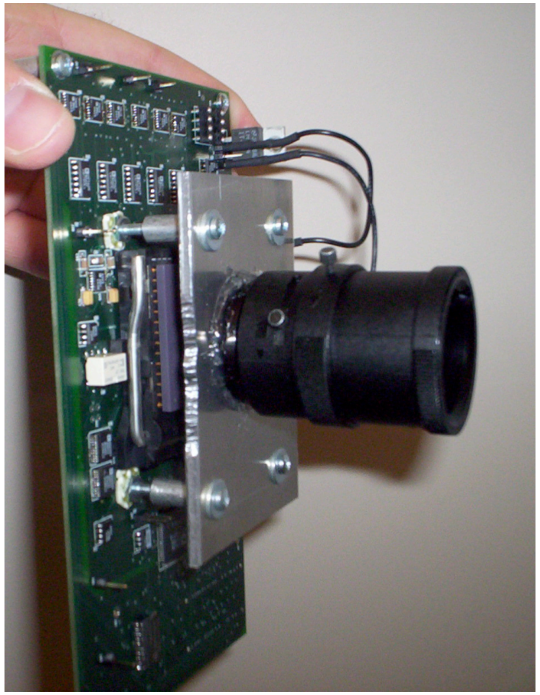


Georgia Tech Analog Imager

- Bottleneck in imager arrays is *data readout*
- Instead of quantizing pixel values, take noiselet inner products *in analog*
- Potential for tremendous (factor of $\approx \mathbf{10^4}$) power savings



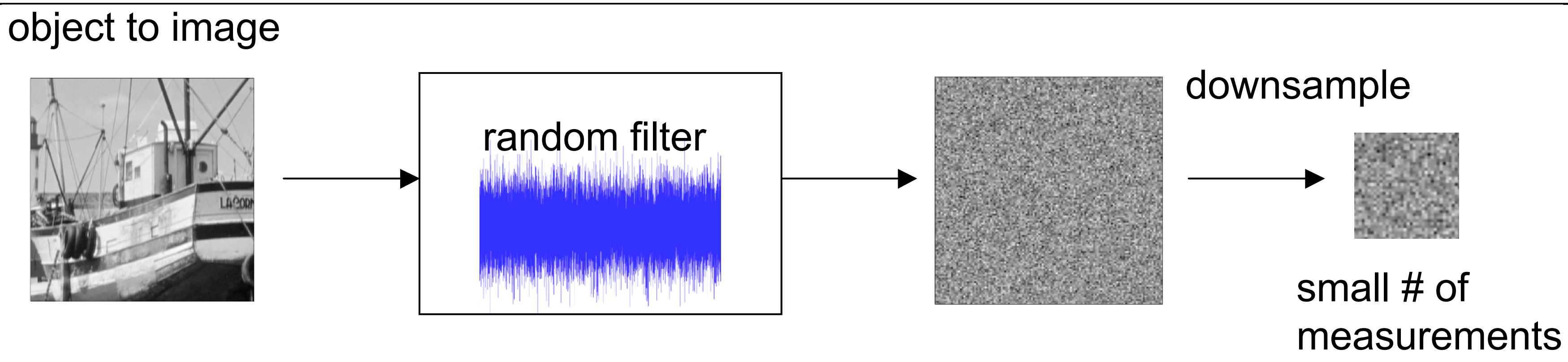




Universality

- Sampling domain M must be very different than sparsity domain $\mathbf{\Psi}$
- Are there *universal* measurement schemes that are structured for fast computations?
- Yes, but we need to add just a little more randomness...

Random Convolution



- Create a *random* orthonormal system in three easy steps:
  Take FFT, randomly change the phases, Inverse FFT

- Measurement matrix is diagonal in Fourier domain

$$M = \mathcal{F}^* \Sigma \mathcal{F}, \quad \Sigma = \mathrm{diag}(\{\sigma_\omega\}),$$

each σ_ω has unit magnitude, random phase

then randomly subsample the rows

- Equivalent to *convolving* with a random pulse, then *subsampling in time*

Intuition for Random Convolution

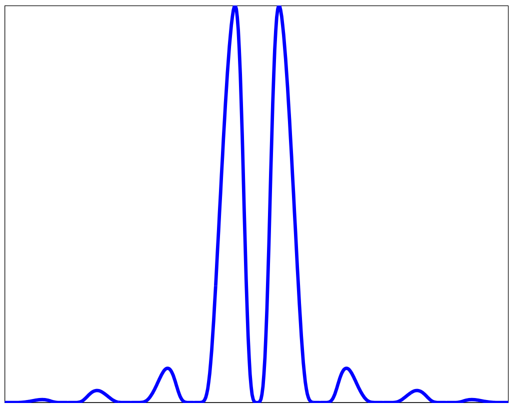
- With (extremely) high probability, measurements will be incoherent with a fixed orthosystem $\boldsymbol{\Psi}$,

$$\mathcal{F}^*\Sigma\mathcal{F} \quad \text{looks like noise in the } \boldsymbol{\Psi} \text{ domain}$$

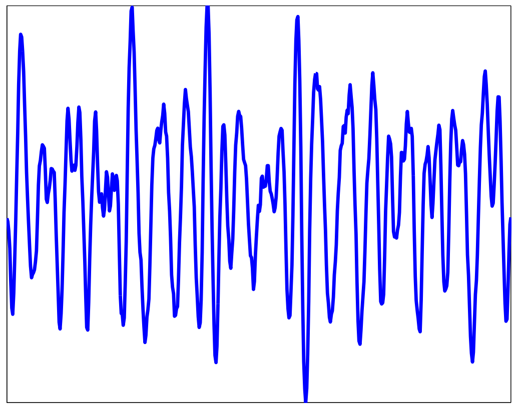
- Applying M is *fast* (two FFTs)
- Example: Wavelets

local in time

local in freq



not local in M



sample here

Theory for Random Convolution

- Fix representation Ψ, generate $M = \mathcal{F}^*\Sigma\mathcal{F}$ (*random orthobasis*)
- Coherence between Ψ and M will be $\mu^2 \sim \log n$
  $\Rightarrow$ extra $\log$ factors in the number of samples required
- Refining our notion of coherence slightly eliminates these
- Perfect recovery with random non-uniform sampling from

$$m \gtrsim S \cdot \log n \quad \text{samples,}$$

and with randomly modulate summation, from (R '08)

$$m \gtrsim S \cdot \log^2 n \quad \text{samples}$$

Why is random convolution + subsampling universal?

$$\begin{bmatrix} & & \\ & \boldsymbol{F} & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} \boldsymbol{\sigma_1} & & & \\ & \boldsymbol{\sigma_2} & & \\ & & \ddots & \\ & & & \boldsymbol{\sigma_n} \end{bmatrix} \left[\begin{array}{c|c|c|c} \hat{\psi}_{\boldsymbol{1}}(\boldsymbol{\omega}) & \hat{\psi}_{\boldsymbol{2}}(\boldsymbol{\omega}) & \cdots & \hat{\psi}_{\boldsymbol{n}}(\boldsymbol{\omega}) \\ & & & \\ & & & \end{array}\right]$$

- One entry of M:

$$\begin{aligned} \boldsymbol{M_{t,s}} &= \sum_{\boldsymbol{\omega}} e^{\boldsymbol{j2\pi\omega t}} \boldsymbol{\sigma_\omega} \hat{\psi}_{\boldsymbol{s}}(\boldsymbol{\omega}) \\ &= \sum_{\boldsymbol{\omega}} \boldsymbol{\sigma'_\omega} \hat{\psi}_{\boldsymbol{s}}(\boldsymbol{\omega}) \end{aligned}$$

- Size of each entry will be concentrated around $\|\hat{\psi}_{\boldsymbol{s}}(\boldsymbol{\omega})\|_{\boldsymbol{2}} = \boldsymbol{1}$
  does not depend on the "shape" of $\hat{\psi}_{\boldsymbol{s}}(\boldsymbol{\omega})$

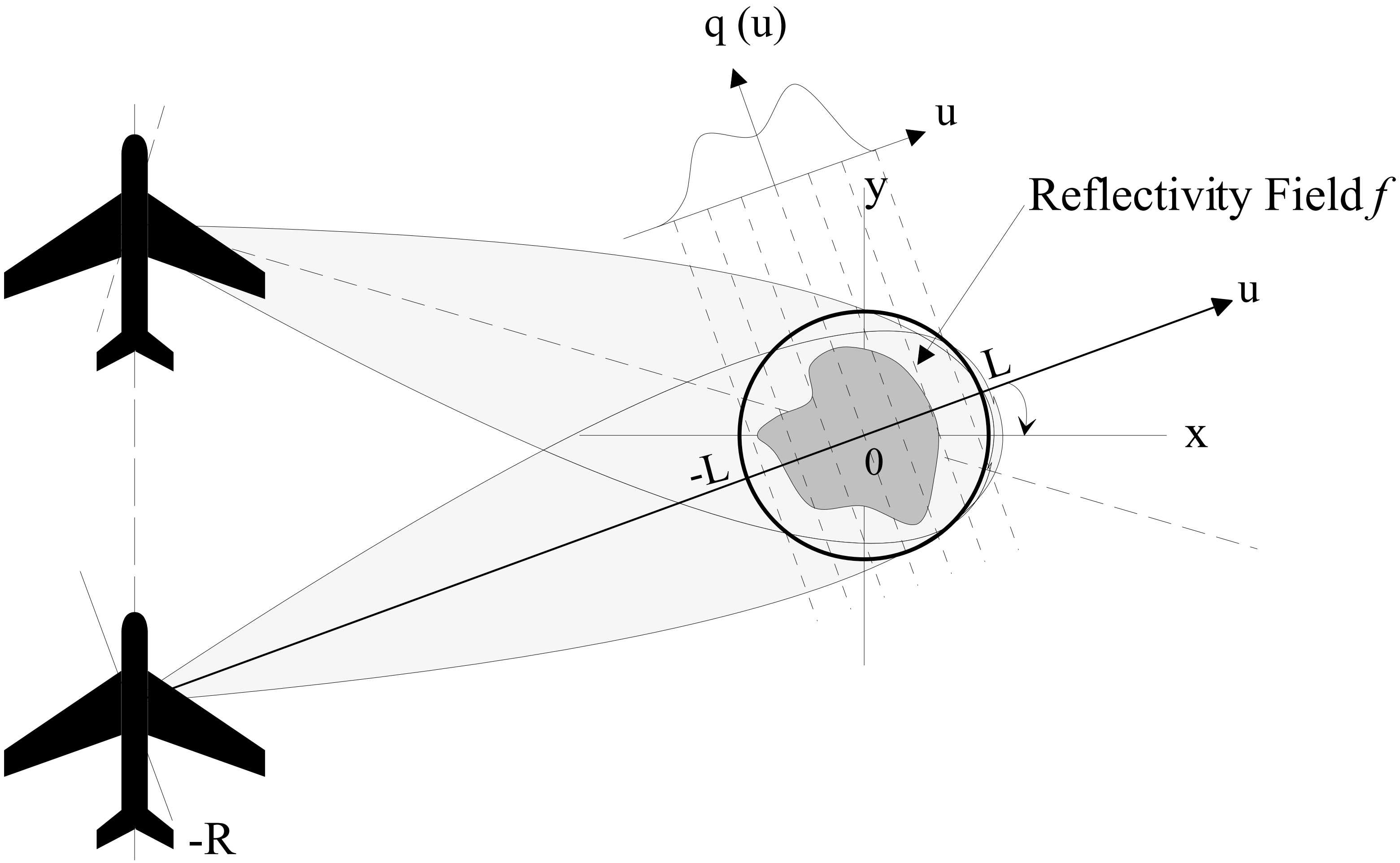
Compare to Fast Johnson-Lindenstrauss Transform

- Ailon and Chazelle, 2006
- Problem:
  k points $x_1, \ldots, x_k$ in $\mathbb{R}^n$, project onto $\mathbb{R}^m$ using Φ ($m \times n$ matrix)
  Want $\|\Phi(x_i - x_j)\|_2 \approx \|x_i - x_j\|_2$ for $m \sim \log k$, and Φ to be "fast"
- JL problem is closely related to CS (Baraniuk et al. '07)
- Their solution: take $\Phi = PHD$
  $D = \mathrm{diag}(\{\epsilon_i\})$ (makes input signs random)
  $H =$ Hadamard transform (Fourier on $\mathbb{Z}_2$)
  $P = m \times n$ subsampling matrix,
  each row has m random entries at random locations
- This Φ would be tremendous, except it is not clear how to implement it by taking $O(m)$ physical measurements
  (P has m^2 entries in it)

Random Convolution

- Convolving with a random pulse then subsampling is an efficient, *universal* acquisition strategy
- Structure allows for fast computations
  Applying $M = \mathcal{F}^*\Sigma\mathcal{F}$ is *fast* (two FFTs)
- Convolution can actually be done physically:
  Two examples:
  - Radar imaging (hi-freq wideband pulse, low-freq sampling)
  - Fourier optics (hi-res diffraction grating, low-res sensor array)

SAR Spotlight Imaging

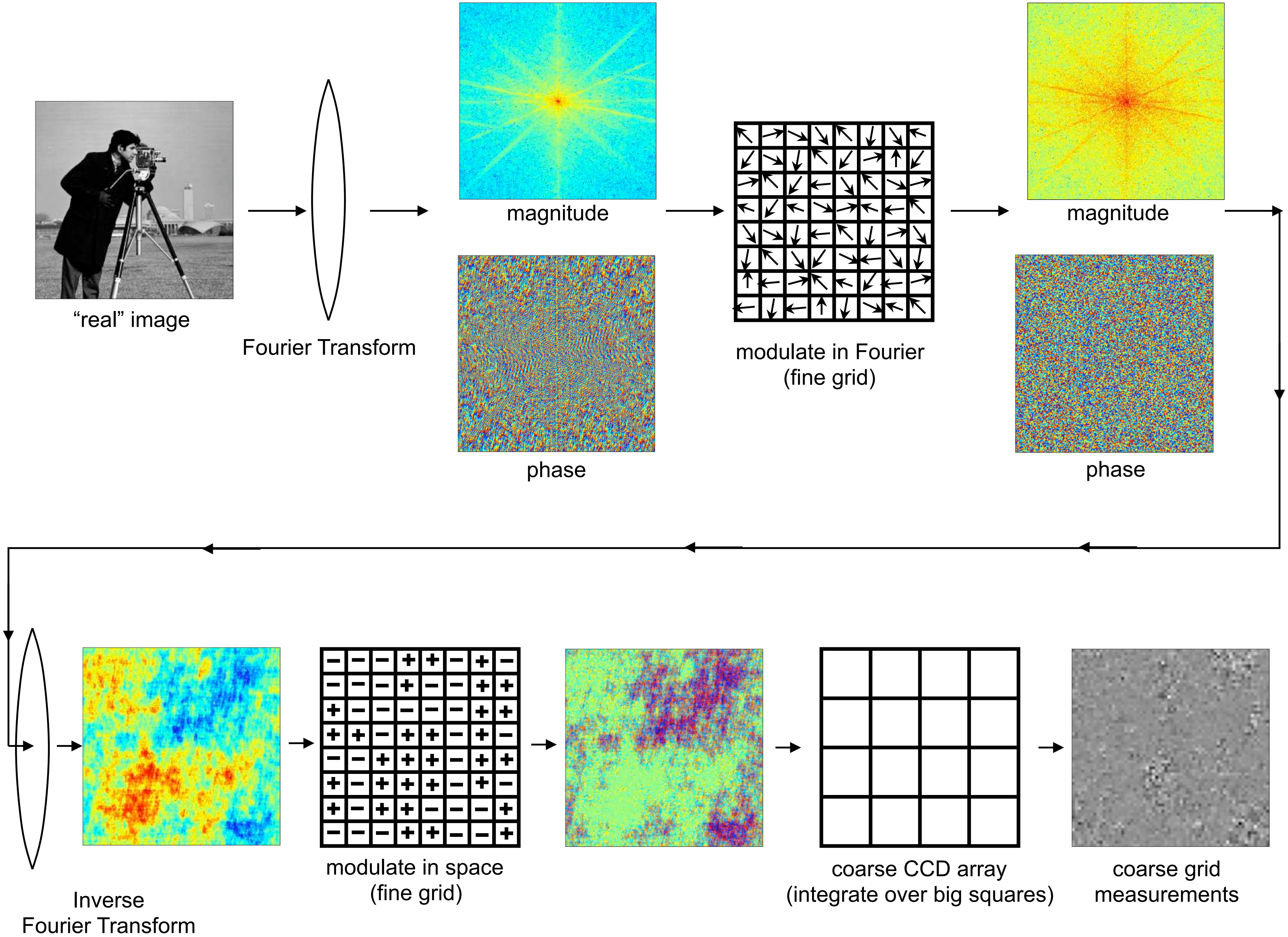


- Send out pulse $p(t)$, return signal is $p(t)$ convolved with range profile $q(t)$
- CS $\Rightarrow$ ADC sampling rate is determined by complexity of range profile, and *not* the bandwidth of the pulses

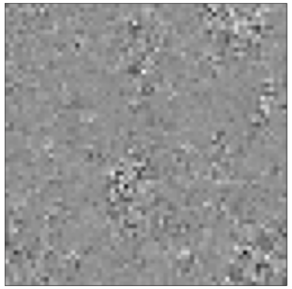
(figure from M. Cetin)

Fourier Optics

- Take Fourier transform of input image with a lens
- Apply random amplitude/phase modulation in the *Fourier domain* with a *spatial light modulator* (=random convolution in space)
- Inverse Fourier transform with another lens
- Large pixels: average consecutive rows of M
- Problem: averaging destroys incoherence ("low pass filter")
- Solution: randomly modulate the summation


"real" image
Fourier Transform
magnitude
phase
modulate in Fourier
(fine grid)
magnitude
phase
Inverse
Fourier Transform
modulate in space
(fine grid)
coarse CCD array
(integrate over big squares)
coarse grid
measurements

coarse grid
measurements



min ℓ_1

compare to standard:

pixelated image
(coarse grid)

recovered image
(fine grid)

Summary

- Random measurements:
  S-sparse recovery from $m \gtrsim S \cdot \log n$ measurements
- Structured measurements:
  S-sparse recovery from $m \gtrsim \mu^2 \cdot S \cdot \log n$ measurements
- Random convolution:
  S-sparse recovery from $m \gtrsim S \cdot \log n$ samples
  - structured yet “incoherent”
  - makes *universal*, large-scale recovery possible
- Immediately suggests architectures for CS imaging
  - Radar
  - Fourier optics