Ideas from Compressive Sampling for Radar

Justin Romberg
Georgia Tech, Electrical and Computer Engineering

IPAM Workshop on Challenges in SAR
February 8, 2012
Los Angeles, California

Compressive sensing and linear algebra

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

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

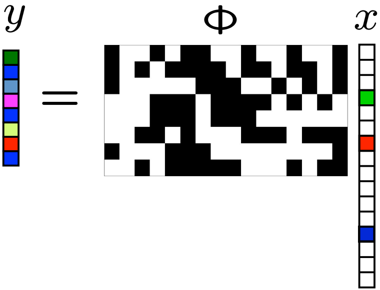
- Fewer measurements than degrees of freedom, $M \ll N$

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

- Treat acquisition as a *linear inverse problem*
- Compressive Sampling: for *sparse* x_0, we can "invert" *incoherent* Φ

Random matrices

Example: $\Phi_{i,j} \sim \pm 1$ w/ prob $1/2$, iid



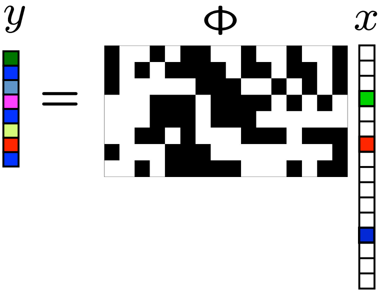
Can recover S-sparse x from

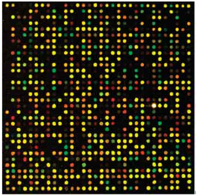
$$M \gtrsim S \cdot \log(N/S)$$

measurements, and the sensing is *universal*

Random matrices

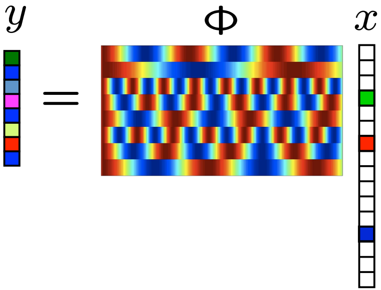
Example: $\Phi_{i,j} \sim \pm 1$ w/ prob $1/2$, iid





Random matrices

Example: Φ consists of *random rows* from a *global orthobasis*



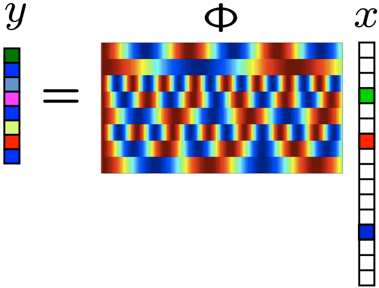
Can recover S-sparse x from

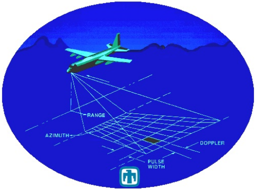
$$M \gtrsim S \cdot \log N$$

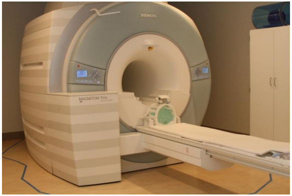
measurements

Random matrices

Example: Φ consists of *random rows* from a *global orthobasis*







Sparse recovery algorithms

Many recovery methods are based on *convex programming*
ℓ_1 minimization (or Basis Pursuit)

$$\min_x \ \|\Psi^T x\|_1 \quad \text{s.t.} \quad \Phi x = y$$

Ψ = sparsifying transform, Φ = measurement system

Convex (linear) program, can relax for robustness to noise

Sparse recovery algorithms

ℓ_1 can recover sparse signals almost anytime it is possible

- perfect recovery with no noise
- stable recovery in the presence of noise
- robust recovery when the signal is not exactly sparse

Sparse recovery algorithms

Other recovery techniques have similar theoretical properties
(their practical effectiveness varies with applications)

- greedy algorithms
- iterative thresholding
- belief propagation
- specialized decoding algorithms

Agenda

What we will talk about today:

- CS via random convolution

Agenda

What we will talk about today:

- CS via random convolution
- MIMO channel estimation using random probes

Agenda

What we will talk about today:

- CS via random convolution
- MIMO channel estimation using random probes
- Analog-to-information receivers

Agenda

What we will talk about today:

- CS via random convolution
- MIMO channel estimation using random probes
- Analog-to-information receivers
- Parametric estimation
  (PDW extraction, e.g. time-of-arrival, carrier frequency)

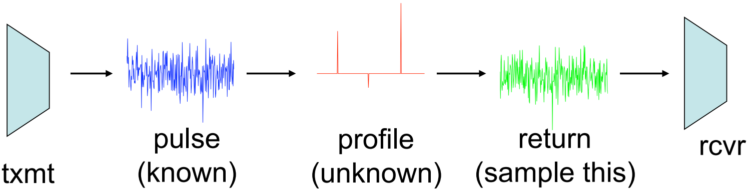
Agenda

What we will talk about today:

- CS via random convolution
- MIMO channel estimation using random probes
- Analog-to-information receivers
- Parametric estimation
  (PDW extraction, e.g. time-of-arrival, carrier frequency)
- Sampling architectures for *ensembles* of signals
  (exploit *correlation structure* rather than sparisty)

Random convolution

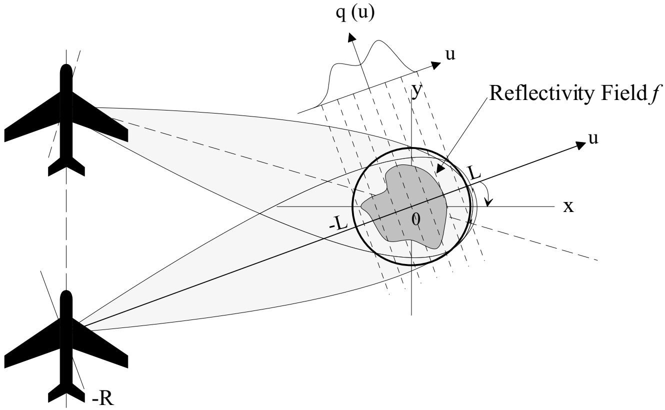
- Many *active imaging* systems measure a pulse convolved with a *reflectivity profile* (Green's function)



- Applications include:
  - radar imaging
  - sonar imaging
  - seismic exploration
  - channel estimation for communications
  - super-resolved imaging
- Using a random pulse = compressive sampling

(R, Rauhut, Tropp, Haupt, Bajwa, ...)

SAR spotlight imaging

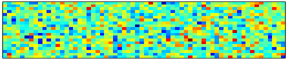


- Send out pulse $p(t)$, return is $p(t)$ convolved with range profile $q(t)$
- CS $\Rightarrow$ ADC sampling rate is determined by complexity of profile, and *not* the bandwidth of the pulses

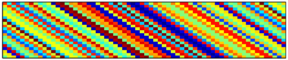
(Munson et al. '83, Cetin '01)

CS via random convolution

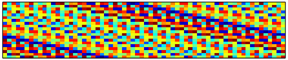
iid subgaussian



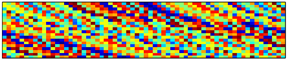
Three models for random convolution



burst sampling



uniform sampling

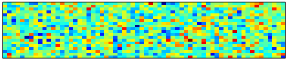


random sampling

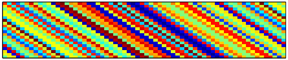
Matrices are random, but structured...

CS via random convolution

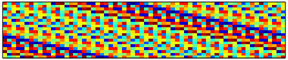
iid subgaussian



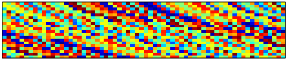
Three models for random convolution



burst sampling
(sparse in time)



uniform sampling
(sparse in time)



random sampling
(*universal*)

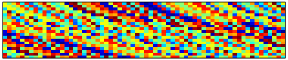
Can guarantee recovery for

$$M \gtrsim S \log^q N$$

in different situations for small q

CS via random convolution

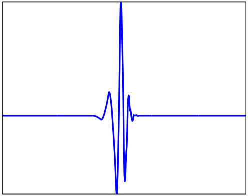
- Signal model: sparsity in *any orthobasis* Ψ
- Acquisition model:
  generate a “pulse” whose FFT is a sequence of random phases,
  convolve with signal,
  sample result at M *random* locations Ω



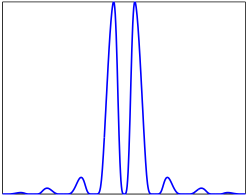
$$\Phi = R_\Omega \mathcal{F}^* \Sigma \mathcal{F}, \quad \Sigma = \mathrm{diag}(\{\sigma_\omega\})$$

Randomizing the phase = spreading in time

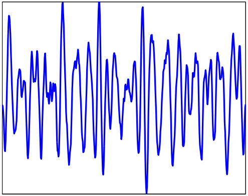
local in time



local in freq

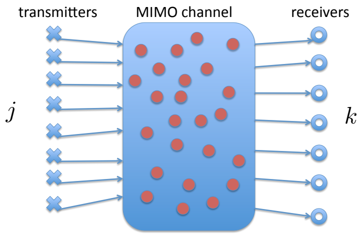


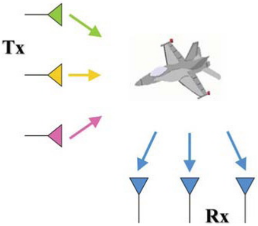
not local in M



sample here

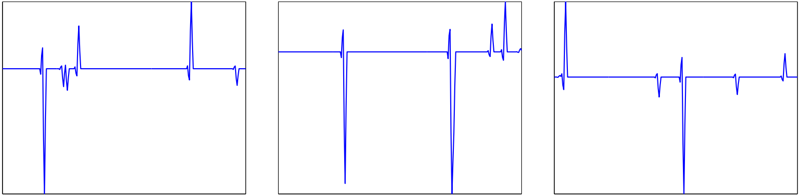
MIMO channel estimation





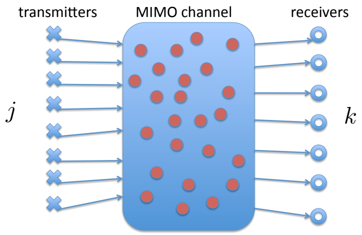
- Estimate all channel responses $h_{j,k}$ = between source j and receiver k

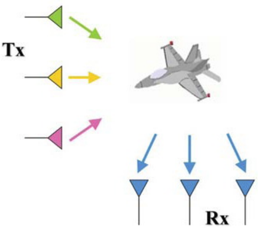
$h_{1,1}$ $h_{1,2}$ $h_{2,4}$ etc



$\ldots$

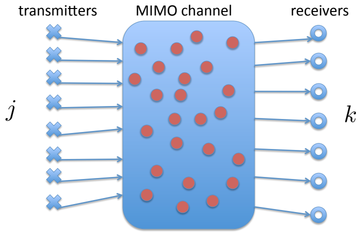
MIMO channel estimation

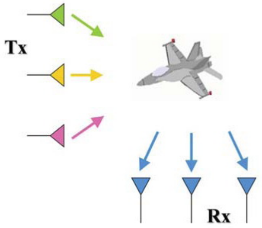




- There are many way to estimate each channel individually (single channel deconvolution)

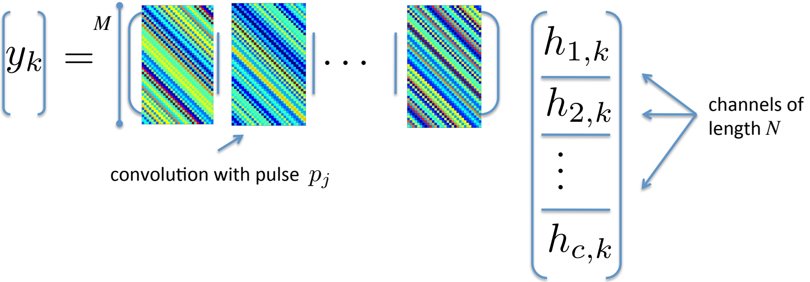
MIMO channel estimation





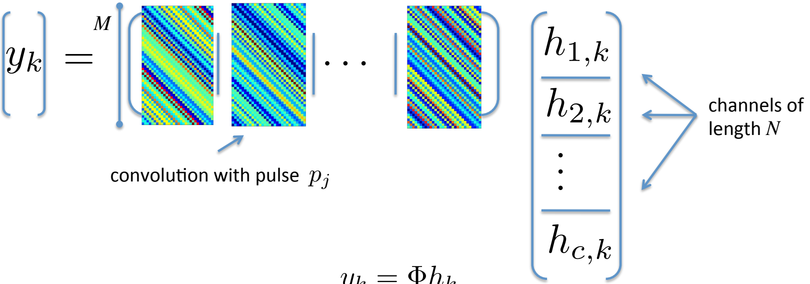
- Activating *simultaneously* with diverse source signatures allows us to separate the cross-talk
- This reduces the total amount of time we spend probing the channels

Multiple channel linear algebra



- How long does each pulse need to be to recover all of the channels? (the system is $M \times NC$, $M =$ pulse length, $C =\#$ channels)
- Of course we can do it for $M \geq NC$

Restricted isometries for multichannel systems



$$y_k = \Phi h_k$$

- With each of the pulses as iid Gaussian sequences, Φ is a "compressive sensing matrix" when

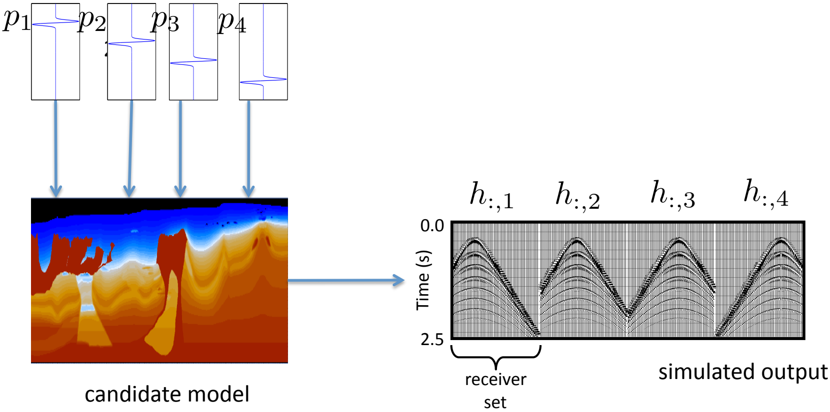
$$M \gtrsim S \cdot \log^q(NC) + N$$

for small q. (R and Neelamani '09)

- **Consequence:** we can separate the channels using short random pulses (using ℓ_1 min or other sparse recovery algorithms)

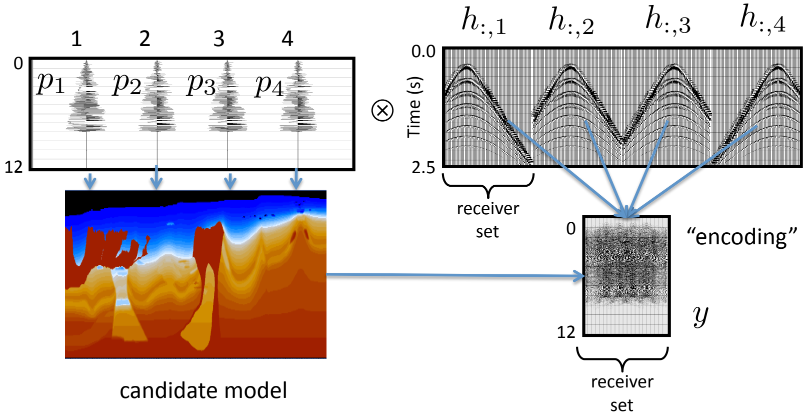
Forward modeling/simulation

- Given a candidate model of the earth, we want to estimate the channel between each source/receiver pair

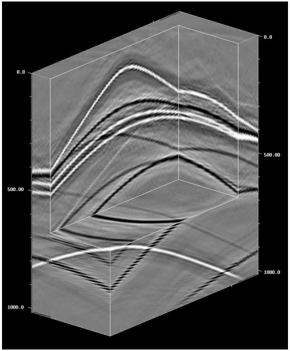


Simultaneous activation

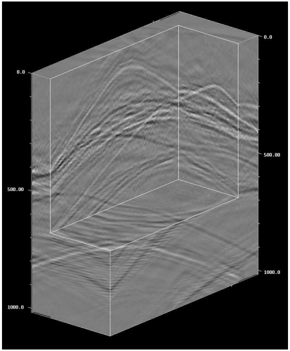
- Run a single simulation with all of the sources activated simultaneously with random waveforms
- The channel responses interfere with one another, but the randomness "codes" them in such a way that they can be separated later



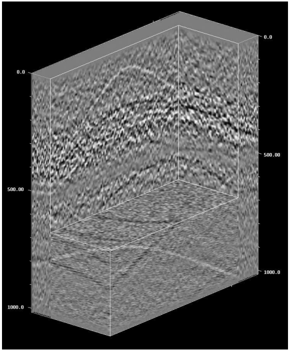
Seismic imaging simulation



(a) Estimated
(16x faster, SNR=9.6 dB).



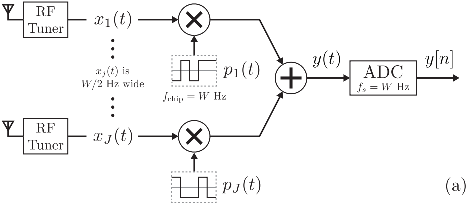
(b) Estimation error
(Figure 2b minus 5(a))

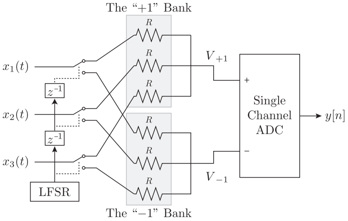


(c) Cross-correlation
estimate.

- Result produced with $16\times$ “compression” in the computations
- Can even take this example down to $32\times$

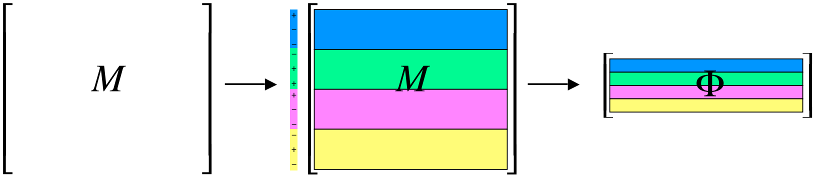
Application: Compressive Multiplexing





- Architecture proposed by Slavinsky et al '10

Randomly modulated integration

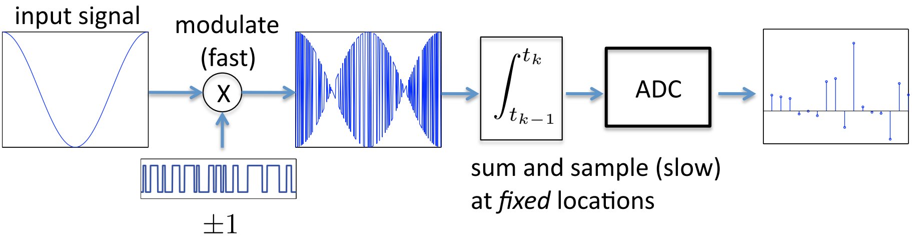


- Measurement system M with coherence μ
- Sampling process:
  Divide rows into m blocks,
  randomly *flip sign* of each row, *sum* over block
- Recovery guarantees for

$$M \gtrsim S \cdot \log^5 N$$

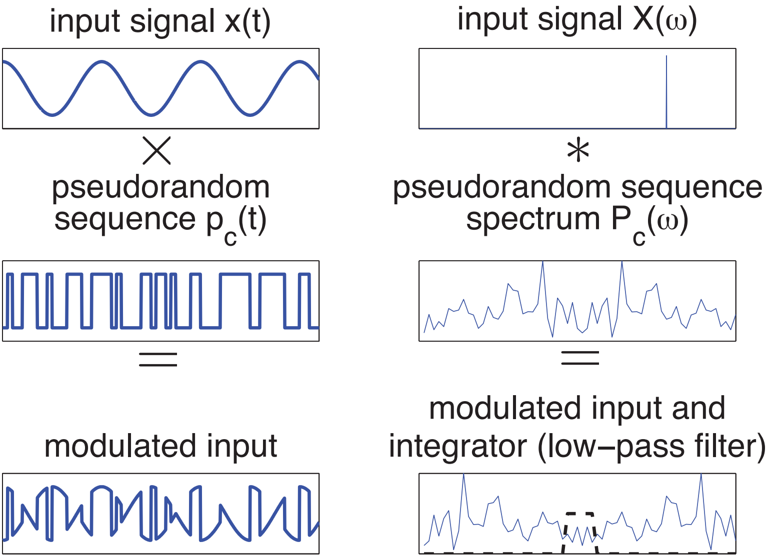
for spectrally sparse signals (Tropp, Duarte, Laska, R, Baraniuk '10)

Randomly modulated integration

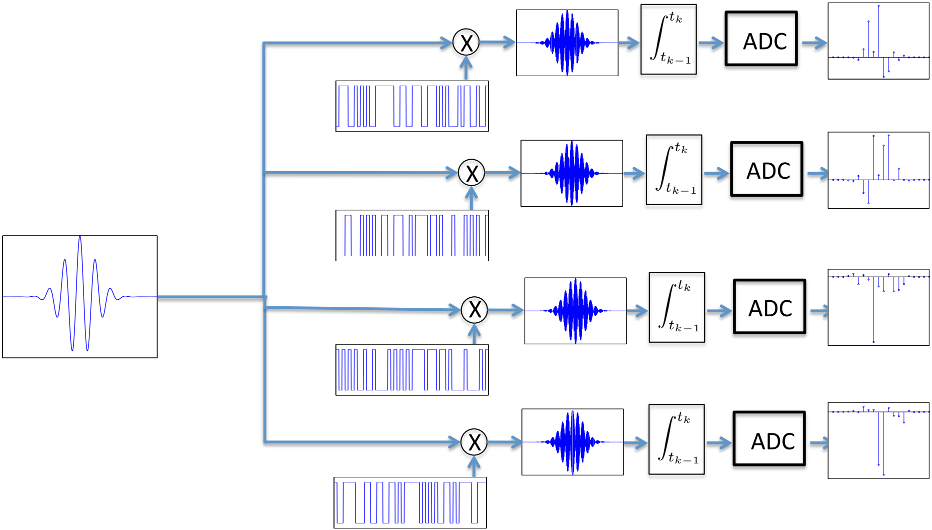


- Uses a standard "slow" ADC preceded by a "fast" binary mixing
- Mixing circuit much easier to build than a "fast" ADC
- In each sampling interval, the signal is summarized with a random sum
- Sample rate $\sim$ total *active* bandwidth

Random modulated integration in time and frequency

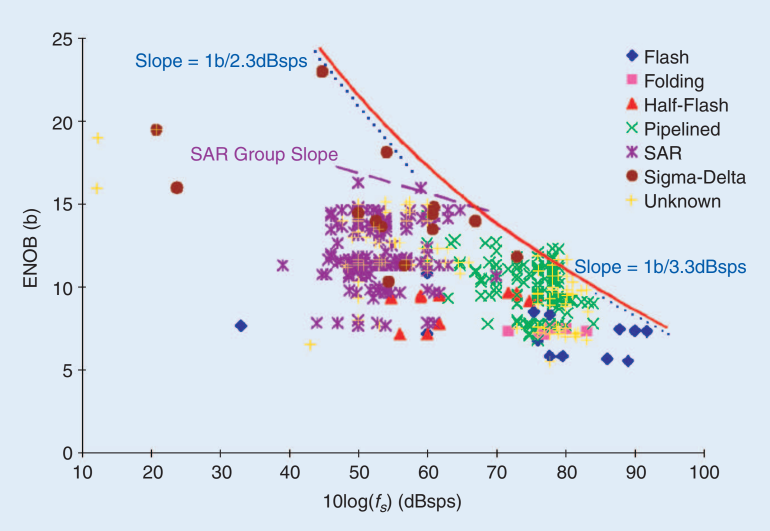


Multichannel modulated integration



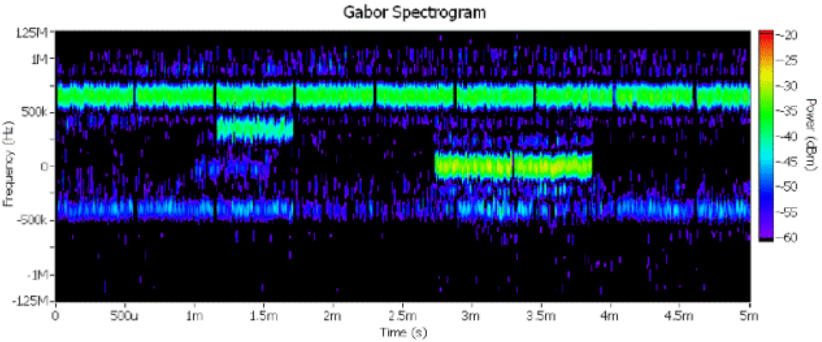
This architecture is being implemented as part of DARPA's Analog-to-Information program

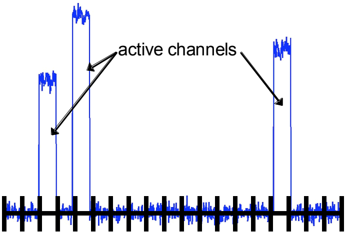
Analog-to-digital converter state-of-the-art



The bad news starts at 1 GHz (Le et al '05)

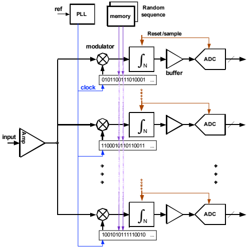
Spectrally sparse RF signals

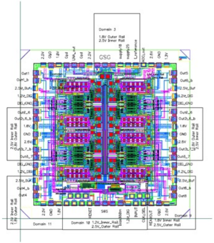




candidate channels

Randomly modulated integration receiver

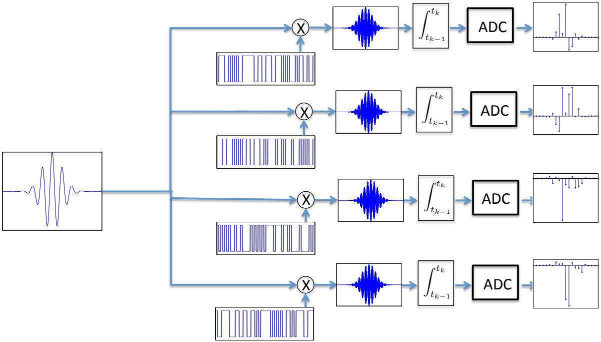




(arch. of Yoo and Emami)

- Random demodulator being built at part of DARPA A2I program
- Multiple (8) channels, operating with different mixing sequences
- Effective BW/chan = 2.5 GHz
  Sample rate/chan = 50 MHz
- Applications: radar pulse detection, communications surveillance, geolocation

Parameter estimation



Say we are interested in just extracting some key parameters from a pulse (carrier frequency, time-of-arrival, etc.)...

Instead of reconstructing the signal, we can estimate these *directly from the compressed samples*

Example formulation: frequency of a pure tone

Suppose that the input signal $x(t)$ is a pure tone at an unknown frequency, amplitude, phase

$$\begin{aligned} x(t) = A_0 \cos(2\pi f_0 t + \theta_0) &= a_1 \cos(2\pi f_0 t) + a_2 \sin(2\pi f_0 t) \\ &= U_{f_0}[a] \end{aligned}$$

Example formulation: frequency of a pure tone

If Φ is the measurement device, we want to solve

$$\hat{f} = \min_{f} \left(\min_{a} \ \|y - \Phi[U_f[a]]\|_2^2 \right)$$

Example formulation: frequency of a pure tone

If Φ is the measurement device, we want to solve

$$\hat{f} = \min_{f} \left(\min_{a} \ \|y - \Phi[U_f[a]]\|_2^2 \right)$$

Let V_f be the $M \times 2$ matrix

$$V_f = \begin{bmatrix} \langle \phi_1(t), \cos(2\pi f t) \rangle & \langle \phi_1(t), \sin(2\pi f t) \rangle \\ \langle \phi_2(t), \cos(2\pi f t) \rangle & \langle \phi_2(t), \sin(2\pi f t) \rangle \\ \vdots & \vdots \\ \langle \phi_M(t), \cos(2\pi f t) \rangle & \langle \phi_M(t), \sin(2\pi f t) \rangle \end{bmatrix}$$

Example formulation: frequency of a pure tone

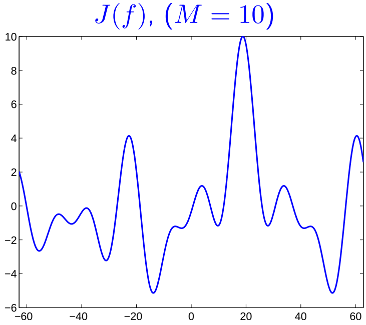
If Φ is the measurement device, we want to solve

$$\begin{aligned}\hat{f} &= \min_f \left(\min_a \ \|y - V_f a\|_2^2 \right) \\ &= \min_f \|(I - P_f)y\|_2^2 \\ &= \max_f \|P_f y\|_2^2\end{aligned}$$

where $P_f = V_f (V_f^T V_f)^{-1} V_f^T$.

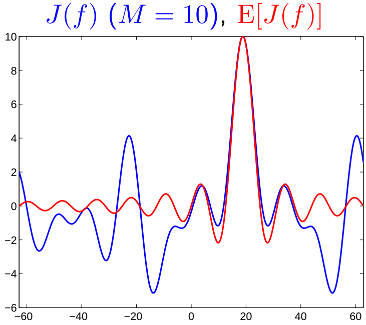
Estimating the frequency of a pure tone

We are looking for the maximum of $J(f) = P_f y$



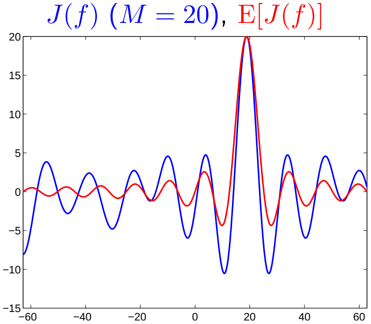
Estimating the frequency of a pure tone

We are looking for the maximum of $J(f) = P_f y$



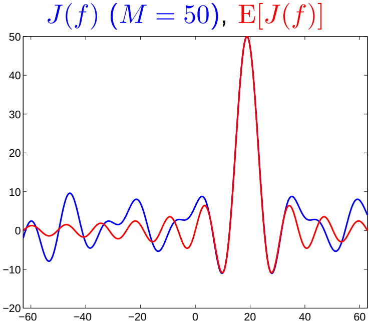
Estimating the frequency of a pure tone

We are looking for the maximum of $J(f) = P_f y$



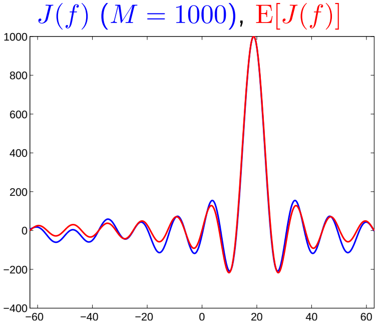
Estimating the frequency of a pure tone

We are looking for the maximum of $J(f) = P_f y$



Estimating the frequency of a pure tone

We are looking for the maximum of $J(f) = P_f y$



Frequency estimation: theory

If we take Nyquist samples and use the FFT, we can withstand noise levels of

$$\sigma^2 \;\lesssim\; \frac{A_0^2}{\log \Omega}$$

where Ω is the total bandwidth.

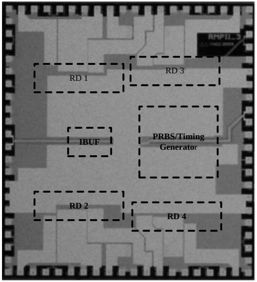
From compressed samples, we can withstand noise levels of

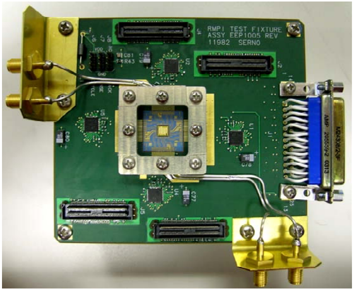
$$\sigma^2 \;\lesssim\; \frac{A_0^2}{\log \Omega} \cdot \frac{M}{\Omega}$$

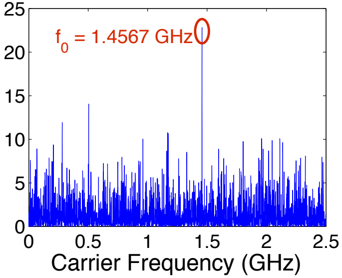
when $M \gtrsim \sqrt{\log \Omega}$.

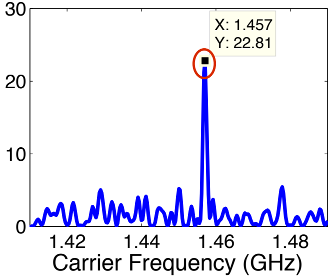
We can interpret M/Ω as the *noise penalty*

Frequency estimation on actual hardware

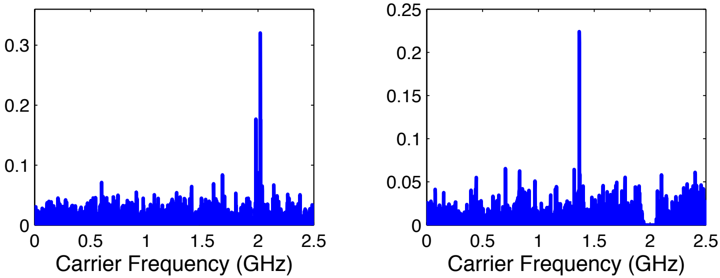








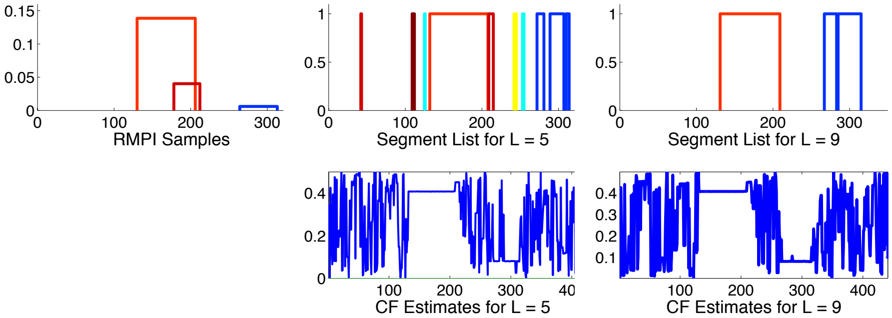
Nulling out interferers



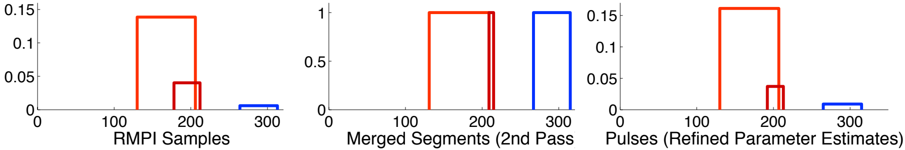
Large, dominating interferer between 1.976 and 2.026 GHz, removed a posteriori using linear algebra

Nulling out an interferer "costs" a number of measurements proportional to its bandwidth

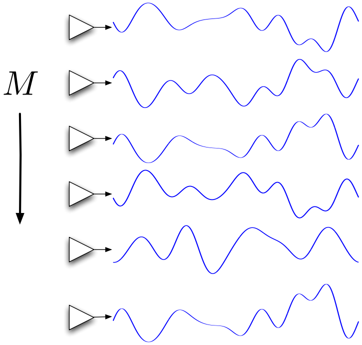
Pulse detection and segmentation



Pulse detection and segmentation

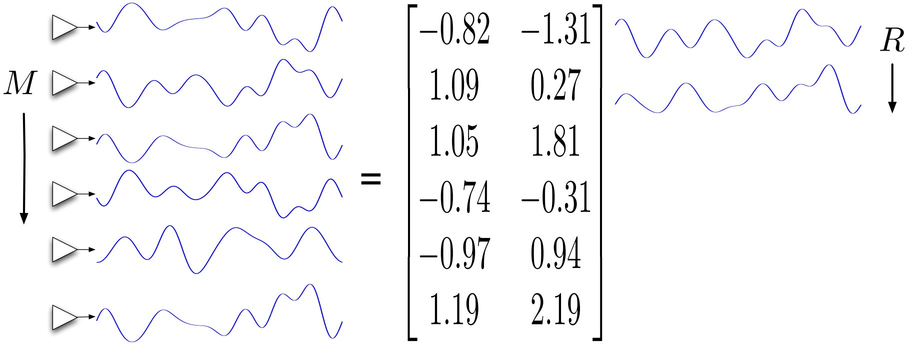


Sampling correlated signals



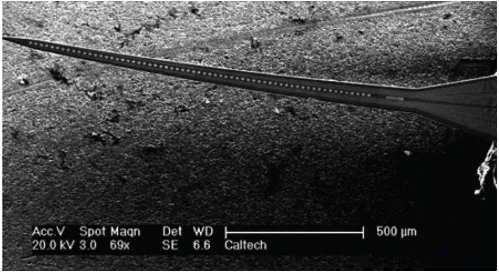
- Goal: acquire an *ensemble* of M signals
- Bandlimited to $W/2$
- "Correlated" $\rightarrow$ M signals are $\approx$ linear combinations of R signals

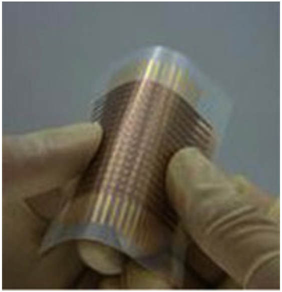
Sampling correlated signals



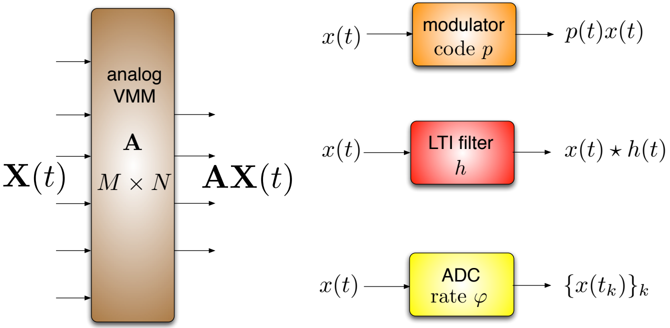
- Goal: acquire an *ensemble* of M signals
- Bandlimited to $W/2$
- "Correlated" $\rightarrow$ M signals are $\approx$ linear combinations of R signals

Sensor arrays



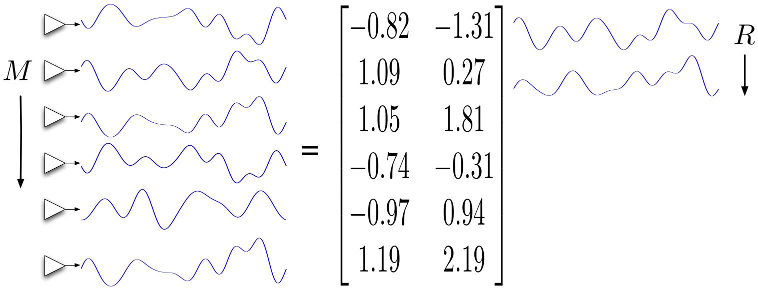


Components



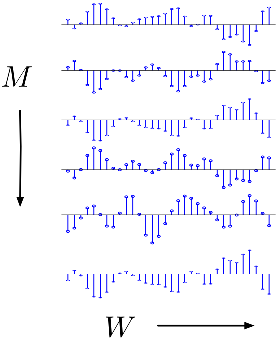
- Analog vector-matrix multiplier spreads energy across channels
- Modulators spread energy across frequency
- Filters spread energy in one channel across time
- We will use both uniform and non-uniform ADCs

Sampling correlated signals

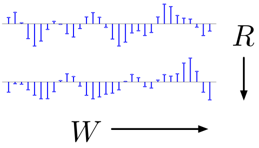


Bandlimited $\Rightarrow$ there is a natural way to discretize this problem ...

Sampling correlated signals



$$= \begin{bmatrix} -0.82 & -1.31 \\ 1.09 & 0.27 \\ 1.05 & 1.81 \\ -0.74 & -0.31 \\ -0.97 & 0.94 \\ 1.19 & 2.19 \end{bmatrix}$$



- Bandlimited $\Rightarrow$ this is just a *low-rank recovery problem*
- Sampling each channel separately takes MW total samples, we want strategies that take $\sim RW$ total samples

Low-rank matrix recovery

- Given p *linear samples* of a matrix,

$$\mathbf{y} = \mathcal{A}(\mathbf{X_0}), \quad \mathbf{y} \in \mathbb{R}^p, \quad \mathbf{X_0} \in \mathbb{R}^{M \times W}$$

  we solve

$$\min_{\mathbf{X}} \ \|\mathbf{X}\|_* \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}$$

  where $\|\mathbf{X}\|_*$ is the nuclear norm: the sum of the singular values of $\mathbf{X}$.

- "optimal" sampler $\mathcal{A}$ would (stably) recover $\mathbf{X_0}$ from $\mathbf{y}$ when

$$\begin{aligned} \#\text{samples} &\gtrsim \ R \cdot \max(M, W) \\ &\gtrsim RW \quad \text{(in our case)} \end{aligned}$$

Low-rank matrix recovery

- Given p *linear samples* of a matrix,

$$\mathbf{y} = \mathcal{A}(\mathbf{X_0}), \quad \mathbf{y} \in \mathbb{R}^p, \quad \mathbf{X_0} \in \mathbb{R}^{M \times W}$$

  we solve

$$\min_{\mathbf{X}} \ \|\mathbf{X}\|_* \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}$$

  where $\|\mathbf{X}\|_*$ is the nuclear norm: the sum of the singular values of $\mathbf{X}$.

Recht, Fazel, Parrilo, Candès, Plan, ...

- If $\mathcal{A}$ is a *random projection*, then we can take

$$\#\text{samples} \ \gtrsim \ RW$$

Low-rank matrix recovery

- Given p *linear samples* of a matrix,

$$\mathbf{y} = \mathcal{A}(\mathbf{X_0}), \quad \mathbf{y} \in \mathbb{R}^p, \quad \mathbf{X_0} \in \mathbb{R}^{M \times W}$$

  we solve

$$\min_{\mathbf{X}} \; \|\mathbf{X}\|_* \quad \text{subject to} \;\; \mathcal{A}(\mathbf{X}) = \mathbf{y}$$

  where $\|\mathbf{X}\|_*$ is the nuclear norm: the sum of the singular values of $\mathbf{X}$.

Recht, Fazel, Parrilo, Candès, Plan, ...

- If $\mathcal{A}$ is a *random projection*, then we can take

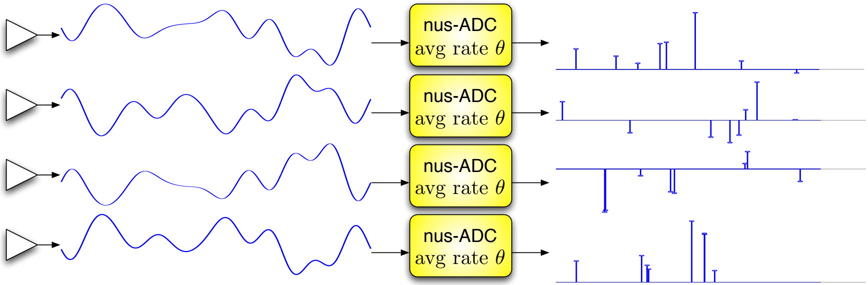
$$\#\text{samples} \;\gtrsim\; RW$$

- *Matrix completion*: $\mathcal{A}$ observes a subset of the entries; we can take

$$\#\text{samples} \;\gtrsim\; RW\log(RW)$$

  under *incoherence* assumptions

Architecture 1: One non-uniform ADC per channel



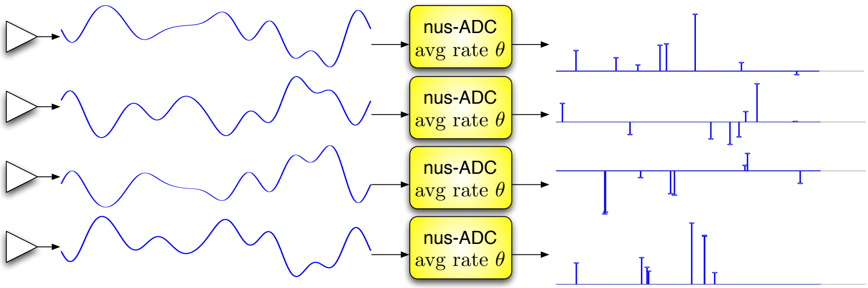
- Direct application of these results: we can recover "incoherent" ensembles when

$$\text{total samples} = M\theta \geq \mathrm{Const.} \cdot RW \cdot \log^2(W)$$

so *we can take $\theta \sim \frac{R}{M}W$ instead of W*.

- Incoherent $\Rightarrow$
signal energy is spread out evenly across time and channels

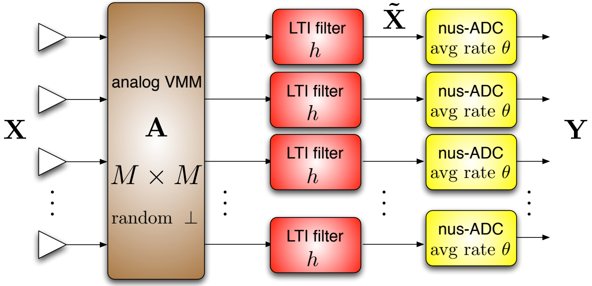
Architecture 1: One non-uniform ADC per channel



Drawbacks:

- Incoherence assumptions (not universal)
- Requires M ADCs (time-multiplexing would be delicate...)

Architecture 2: Pre-mix + prefilter + non-uniform ADCs



- Ahmed, R '11: We can recover the ensemble $\tilde{\mathbf{X}}$ when

$$\text{total samples} \;\gtrsim\; RW \log^4(W)$$

- From $\tilde{\mathbf{X}}$, we recover $\mathbf{X}$ using

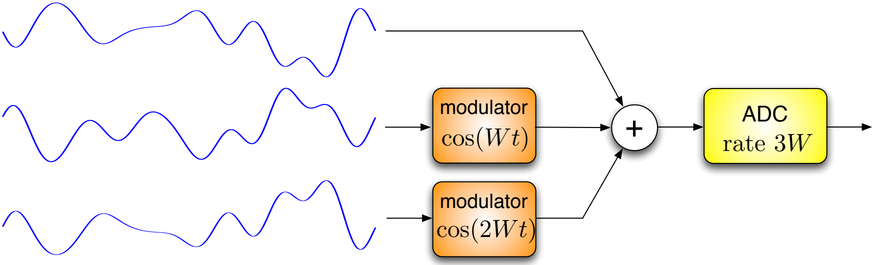
$$\mathbf{X} = \mathbf{A}^{\mathbf{T}}\tilde{\mathbf{X}}\mathbf{H}$$

- Universal, but still using an ADC for every channel...

Multiplexing onto one channel

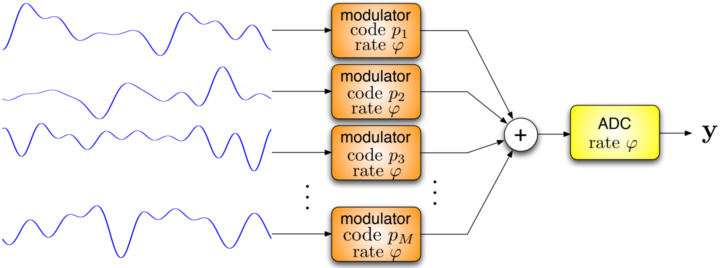
- We can always combine M channels into 1 by *multiplexing* in either time or frequency

  Frequency multiplexer:



- Replace M ADCs running at rate W with 1 ADC at rate MW

Architecture 3: modulated multiplexing

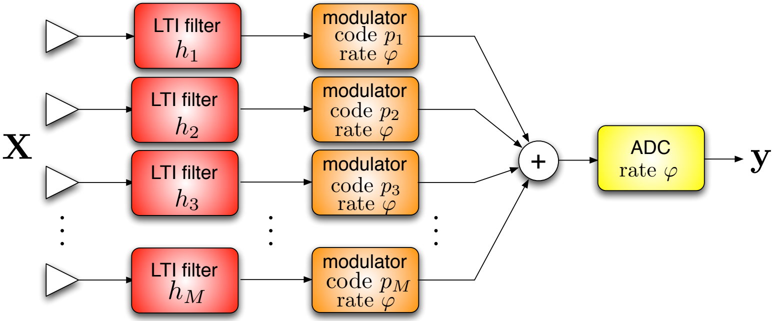


- Ahmed, R '11: If the signals are spread out uniformly in time, then the ADC and modulators can run at rate

$$\varphi \gtrsim RW \log^{3/2}(MW)$$

- This requires a (milder) "incoherence across time" assumption, and a (slightly) different recovery algorithm for the theory to work

Architecture 4: prefilter + modulated multiplexing



- We can stably recover a rank-R ensemble X when the modulators and ADC operate at rate

$$\varphi \;\gtrsim\; RW \log^4(mW)$$

- This architecture is *universal* in that it works for any low-rank correlation structure

Recap

- CS via random convolution
- MIMO channel estimation using random probes
- Analog-to-information receivers
- Parametric estimation
  (PDW extraction, e.g. time-of-arrival, carrier frequency)
- Sampling architectures for *ensembles* of signals
  (exploit *correlation structure* rather than sparisty)

jrom@ece.gatech.edu