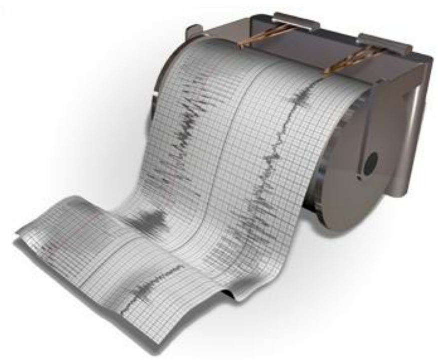
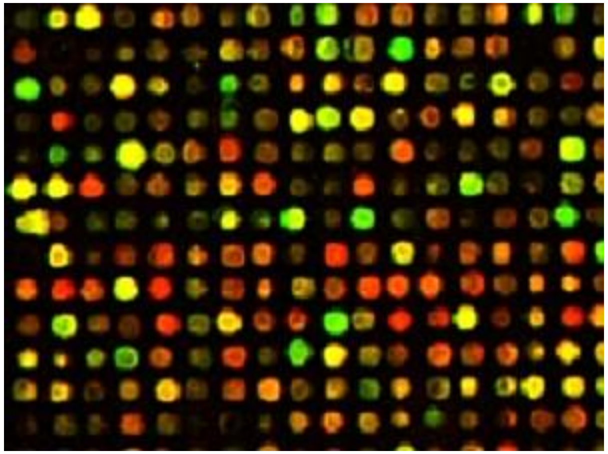
Compressive Measurements for Signal Acquisition and Processing

Mark Davenport

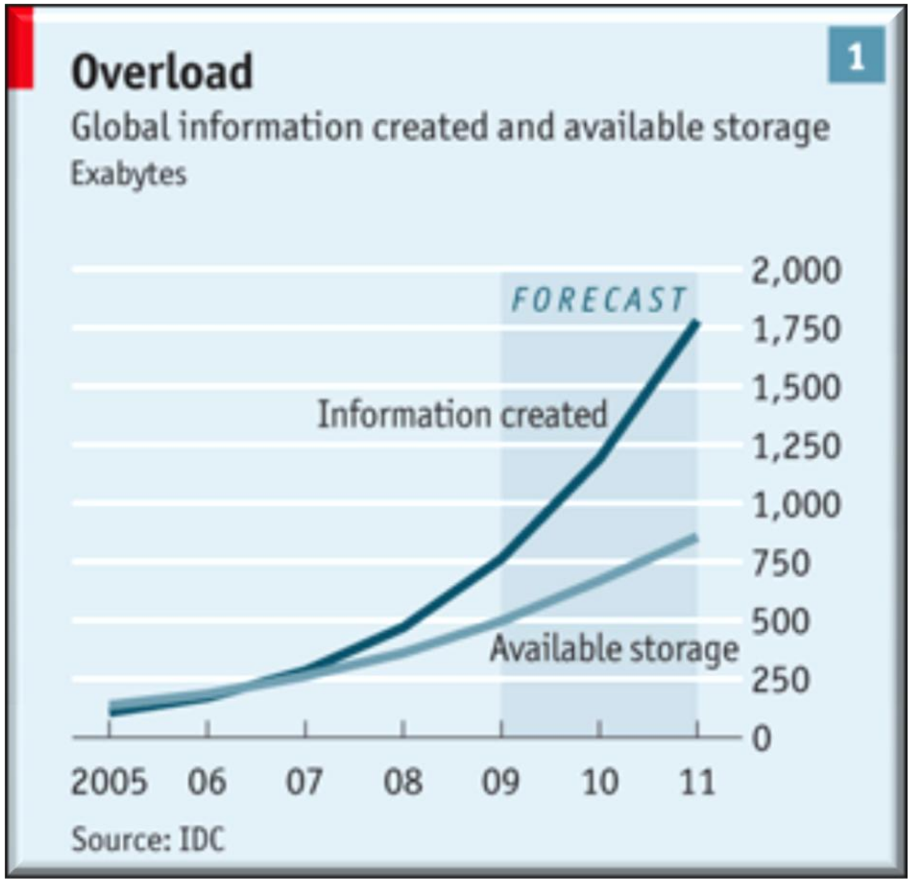
Rice University
ECE Department

Sensor Explosion





Data Deluge



By 2011, ½ of digital universe will have no home

[The Economist – March 2010]

Dimensionality Overload

~~How can we get as much data as possible into the digital domain?~~

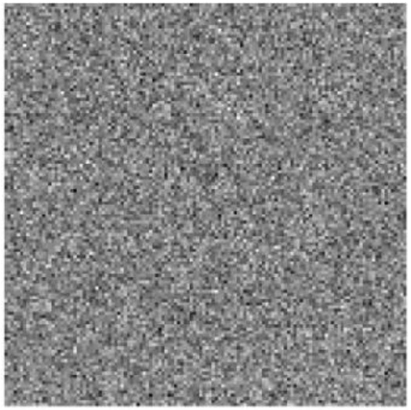
~~How can we extract as much information as possible from a limited amount of data?~~

How can we avoid having to acquire so much data?

How can we extract any information at all from a massive amount of high-dimensional data?

Dimensionality Reduction

Data is rarely intrinsically high-dimensional



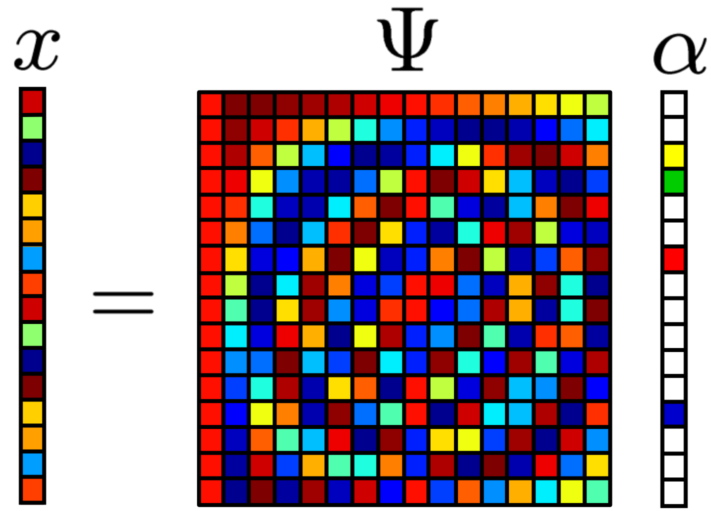
Signals often obey ***low-dimensional models***

- sparsity
- manifolds
- low-rank matrices

The intrinsic dimension K can be much less than the ambient dimension N, which enables ***dimensionality reduction***

Sparsity

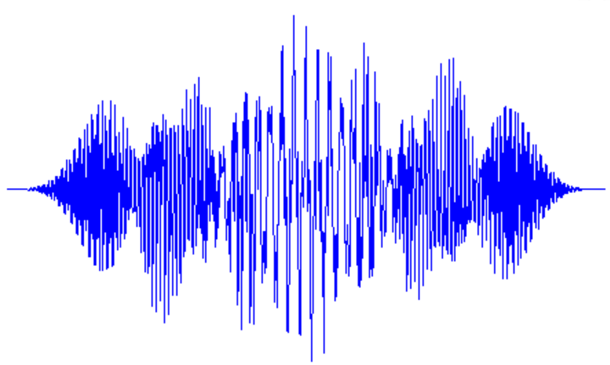
$$x = \sum_{j=1}^{N} \alpha_j \psi_j$$
$$= \Psi \alpha$$

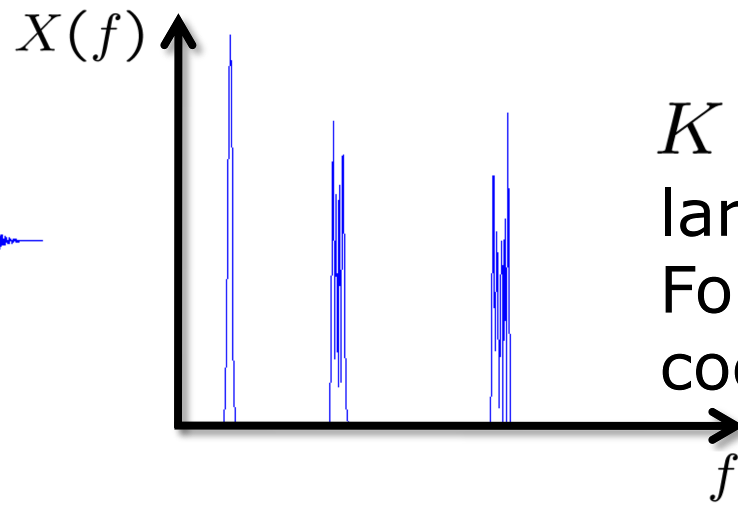


K nonzero entries

$\|\alpha\|_0 = K$

N samples

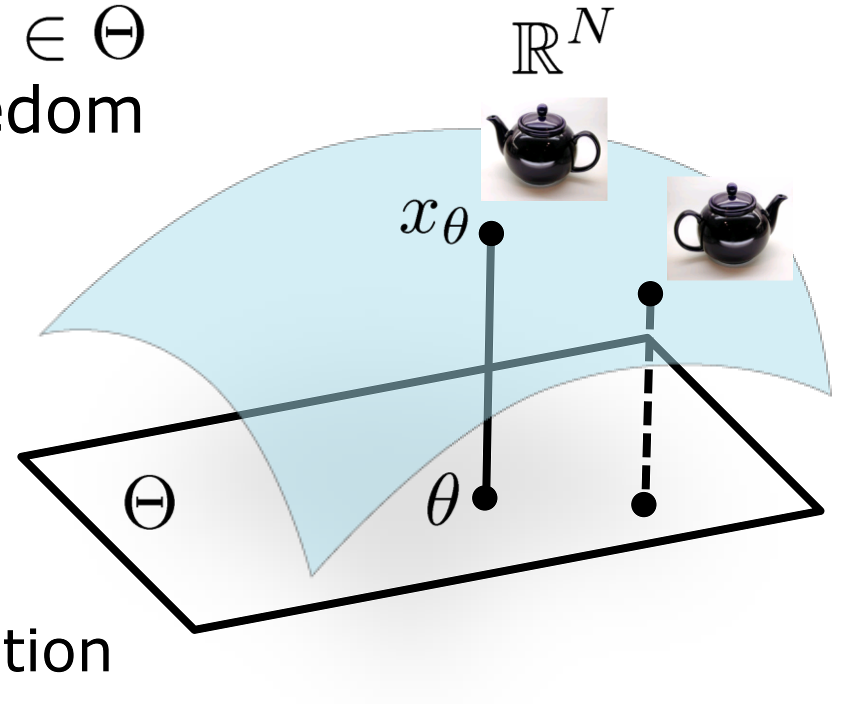


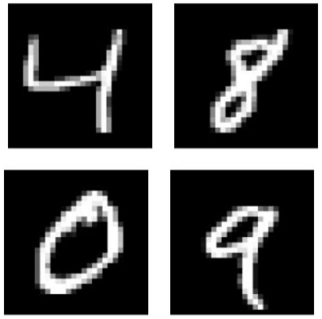


$K \ll N$ large Fourier coefficients

Manifolds

- K-dimensional *parameter* $\theta \in \Theta$ captures the degrees of freedom of signal
- Signal class forms a K-dimensional *manifold*
  - rotations, translations
  - robot configuration spaces
  - signal with unknown translation
  - sinusoid of unknown frequency
  - faces
  - handwritten digits
  - speech
  - …

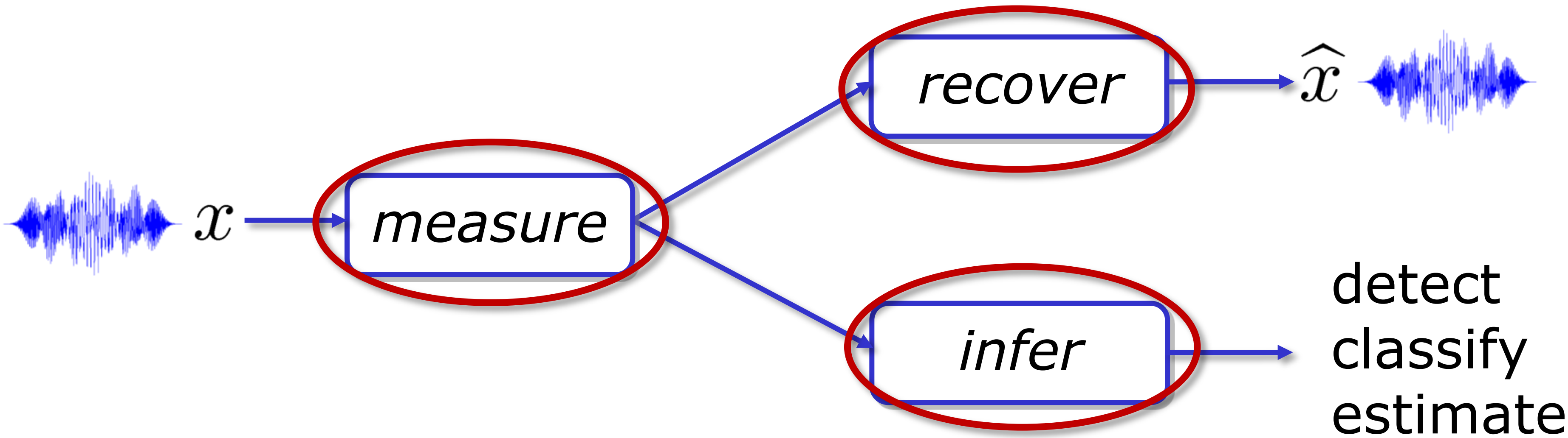




Compressive Signal Processing

How can we exploit low-dimensional models in the design of signal processing algorithms?

We would like to operate at the *intrinsic dimension* at all stages of the DSP pipeline



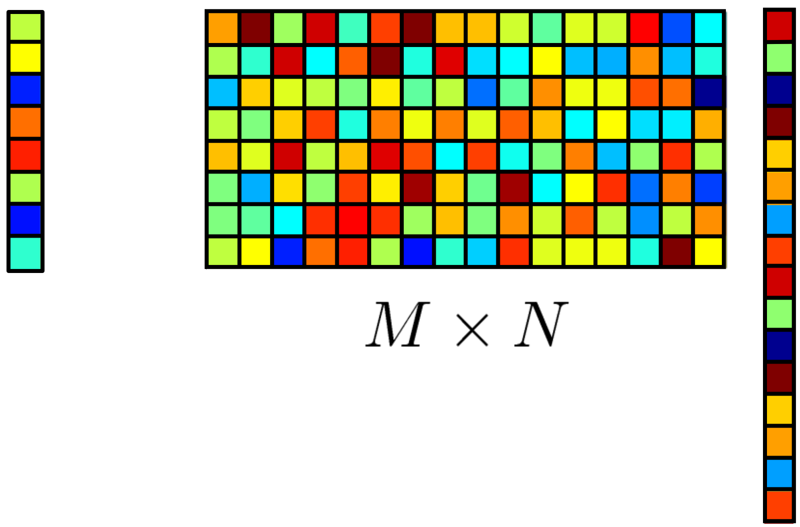
Compressive Measurements

Compressive Measurements

Compressive sensing [Donoho; Candes, Romberg, Tao – 2004]

Replace samples with general *linear measurements*

$$y = \Phi x$$



$M \times 1$ measurements

$M \times N$

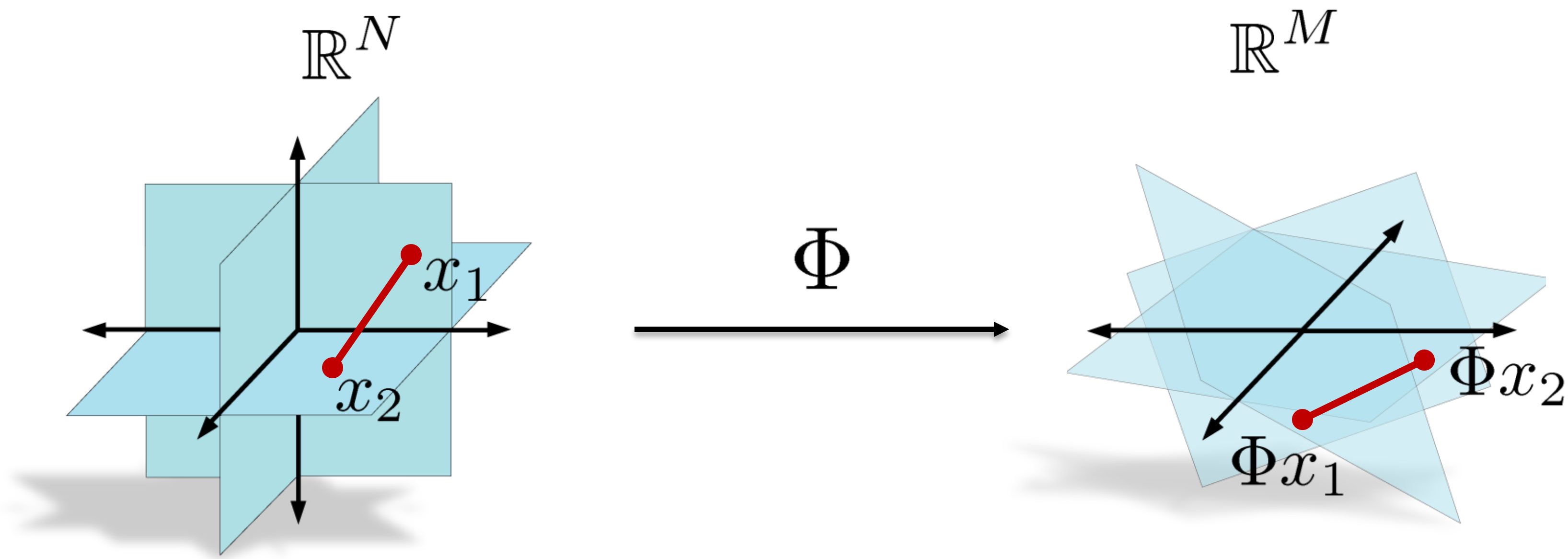
$N \times 1$ *sampled* signal

K-sparse

$$\boxed{M = O(K \log(N/K))}$$

Restricted Isometry Property (RIP)

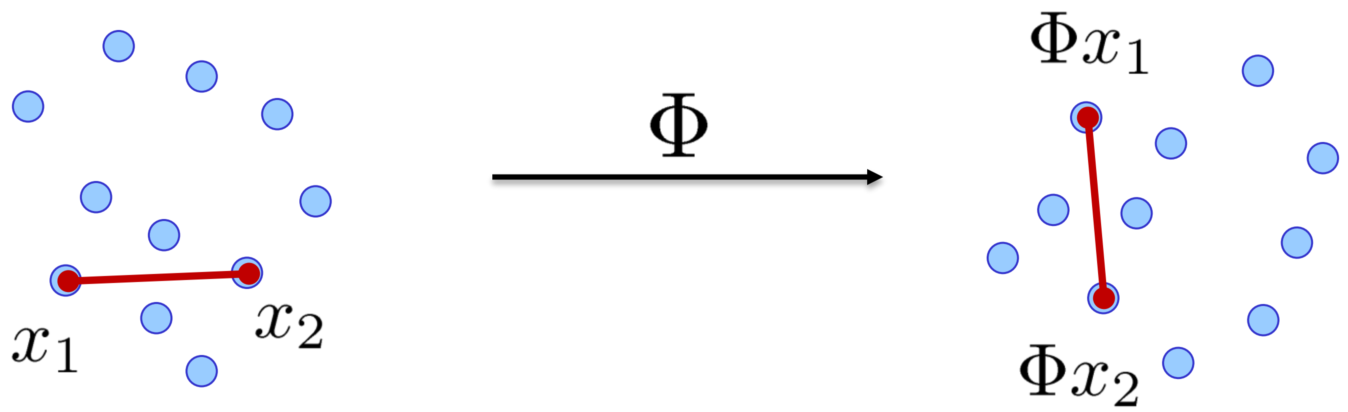
$$1 - \delta \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq 1 + \delta \qquad \|x_1\|_0, \|x_2\|_0 \leq K$$



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

Johnson-Lindenstrauss Lemma

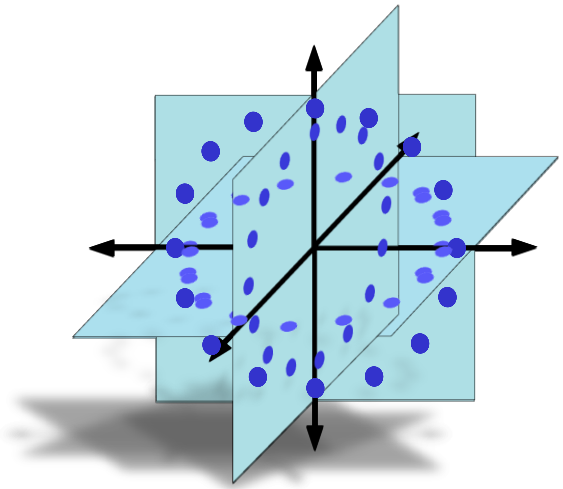
- Stable projection of a discrete set of P points



- 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

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



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

[Baraniuk, M.D., DeVore, Wakin – Const. Approx. 2008]

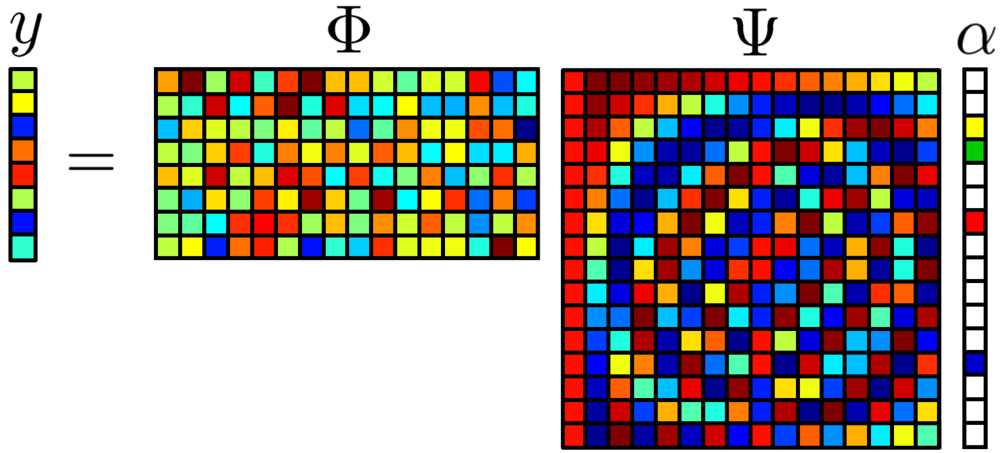
Hallmarks of Random Measurements

Stable

Φ will preserve information, be robust to noise

Universal

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

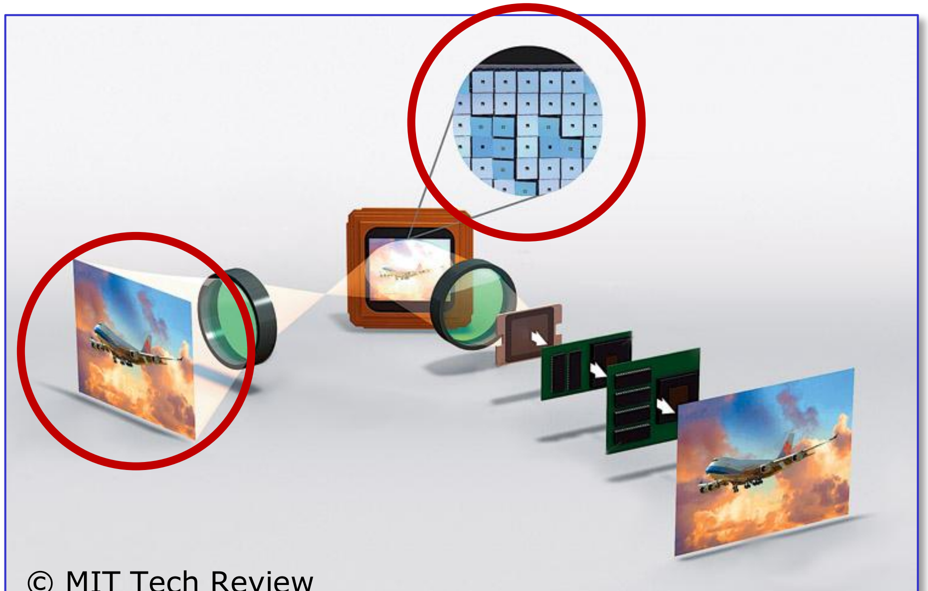


Democratic

Each measurement has “equal weight”

Compressive Measurements: Imaging

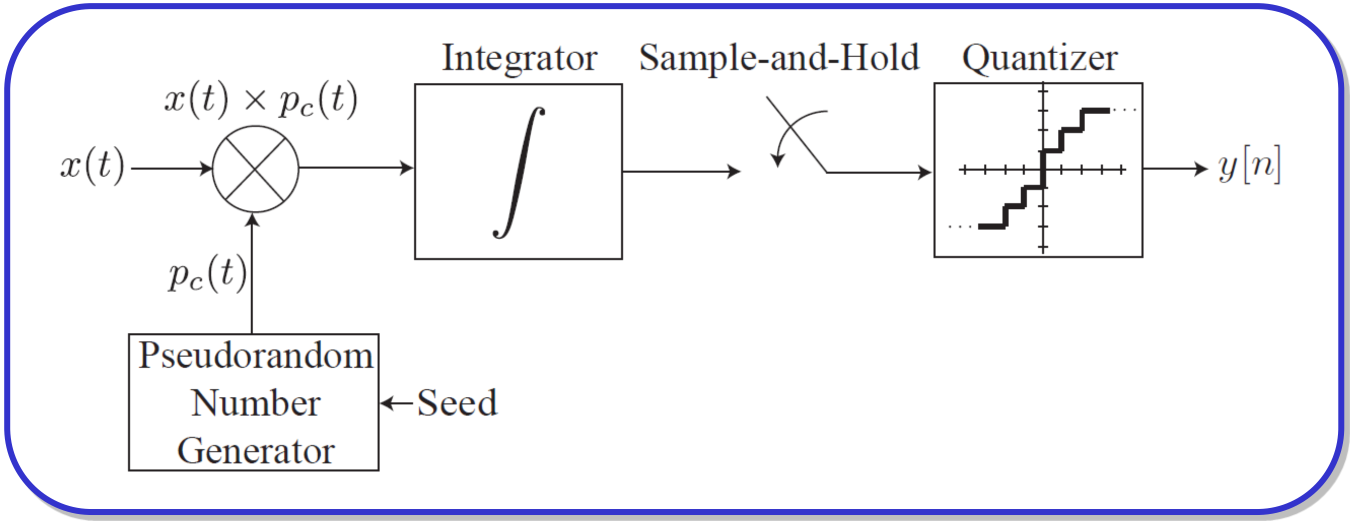
Rice "single-pixel camera"

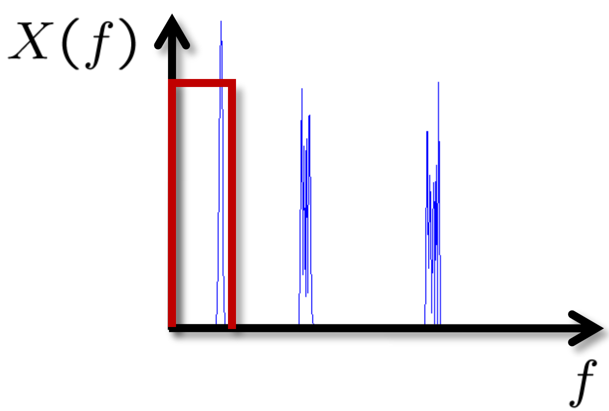


[Duarte, M.D., Takhar, Laska, Sun, Kelly, Baraniuk – Sig. Proc. Mag. 2008]

Compressive Measurements: ADCs

"Random demodulator"

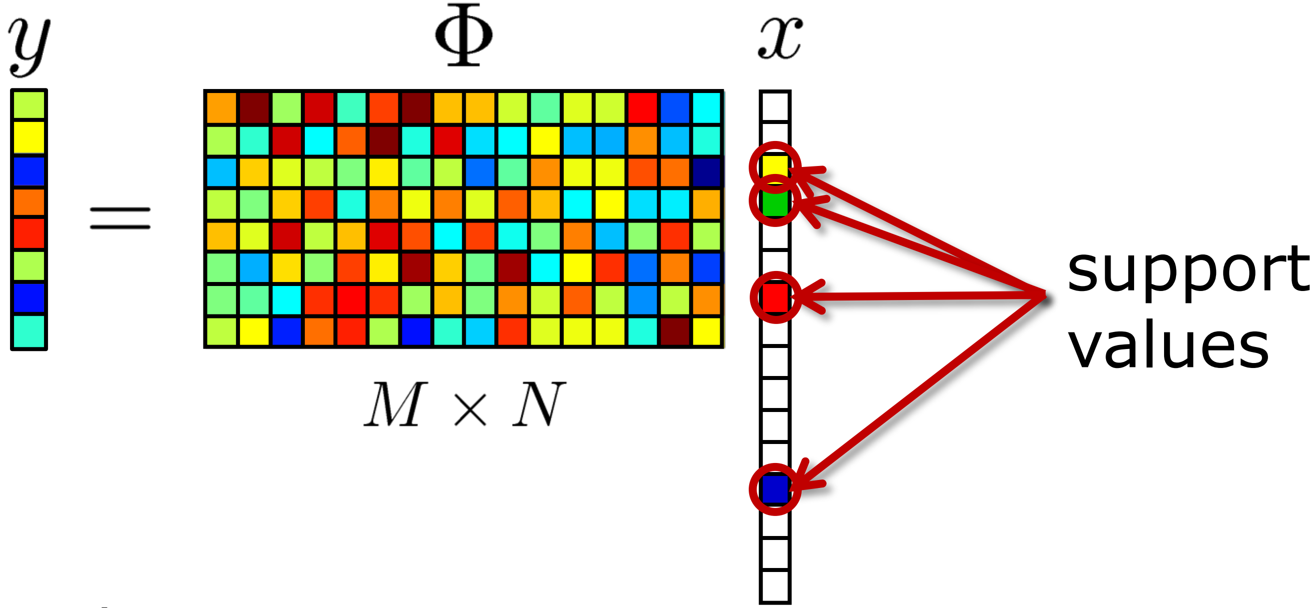




[Tropp, Laska, Duarte, Romberg, Baraniuk – Trans IT 2010]

Signal Acquisition and Recovery

Sparse Signal Recovery

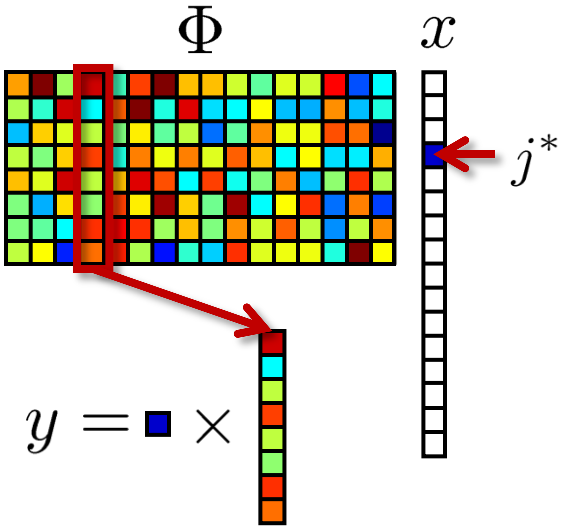


- Optimization / ℓ_1-minimization
- Greedy algorithms
  - matching pursuit
  - orthogonal matching pursuit (OMP)
  - regularized OMP
  - CoSaMP, Subspace Pursuit, IHT, …

Orthogonal Matching Pursuit

OMP selects one index at a time

Iteration 1:



$$j^* = \arg\max_j |\langle y, \Phi_j \rangle|$$

If Φ satisfies the RIP of order $\|u \pm v\|_0$, then

$$|\langle \Phi u, \Phi v \rangle - \langle u, v \rangle| \leq \delta \|u\|_2 \|v\|_2$$

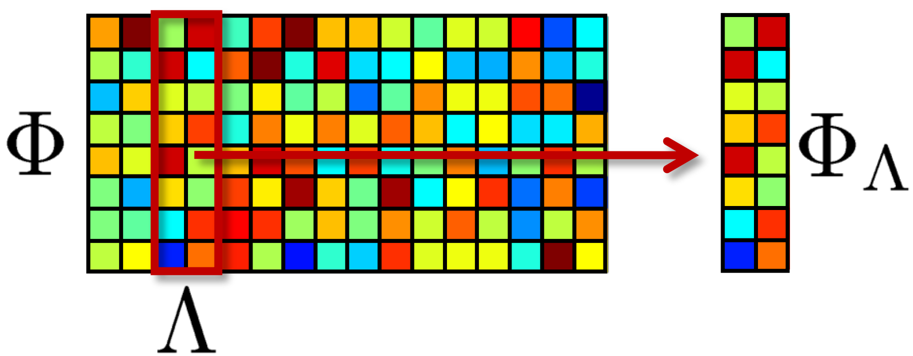
Set $u = x$ and $v = e_j$

$$|\langle y, \Phi_j \rangle - x_j| \leq \delta \|x\|_2$$

Orthogonal Matching Pursuit

Subsequent Iterations:

$$\boxed{j^* = \arg\max_j |\langle Py, P\Phi_j \rangle|}$$



$$P = I - \underbrace{\Phi_\Lambda \Phi_\Lambda^\dagger}_{\text{Projection onto } \mathcal{R}(\Phi_\Lambda)}$$

$$P\Phi_\Lambda = 0 \quad \Rightarrow \quad P\Phi x = P\Phi x_{\Lambda^c}$$

Interference Cancellation

Lemma

If Φ satisfies the RIP of order K, then

$$\left(1 - \frac{\delta}{1-\delta}\right)\|x\|_2^2 \le \|P\Phi x\|_2^2 \le (1+\delta)\|x\|_2^2$$

for all x such that $\|x\|_0 \le K - |\Lambda|$ and $\operatorname{supp}(x) \cap \Lambda = \emptyset$.

$$\Rightarrow \quad |\langle Py, P\Phi_j\rangle - x_j| \le \frac{\delta}{1-\delta}\|x_{\Lambda^c}\|_2$$

[M.D., Boufounos, Wakin, Baraniuk – J. Selected Topics in Sig. Proc. 2010]

Orthogonal Matching Pursuit

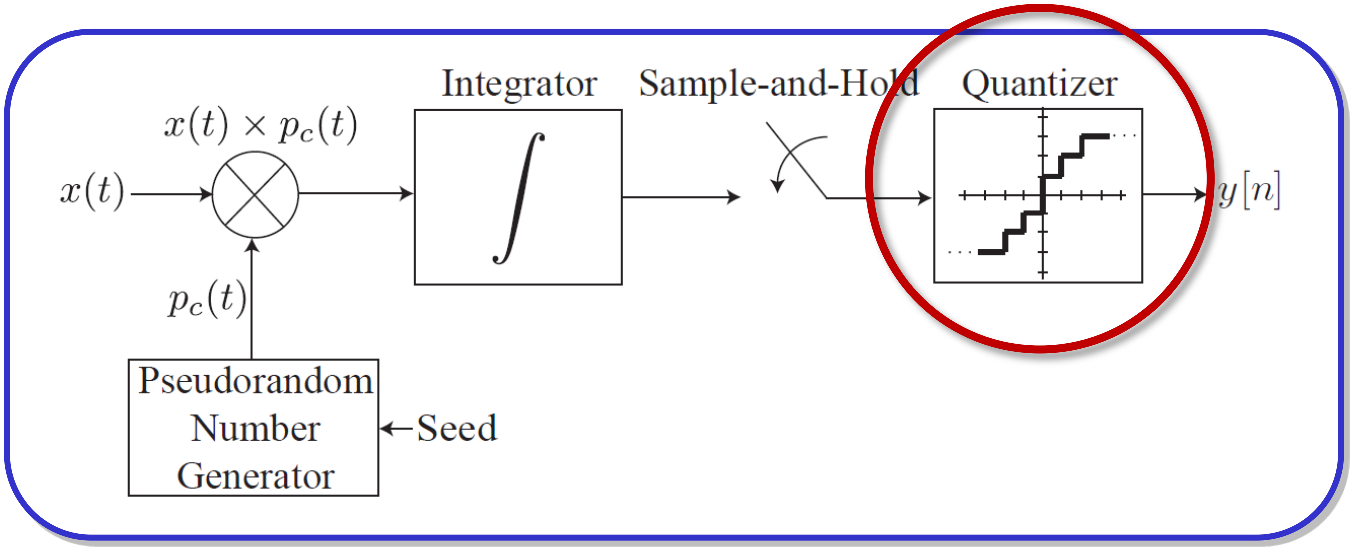
Theorem
Suppose x is K-sparse and $y = \Phi x$.
If Φ satisfies the RIP of order $K+1$ with constant $\delta < \frac{1}{3\sqrt{K}}$, then the j^* identified at each iteration will be a nonzero entry of x.

➡ Exact recovery after K iterations.

Argument provides simplified proofs for other orthogonal greedy algorithms (e.g. ROMP) that are robust to noise

[M.D., Wakin – Trans. IT 2010]

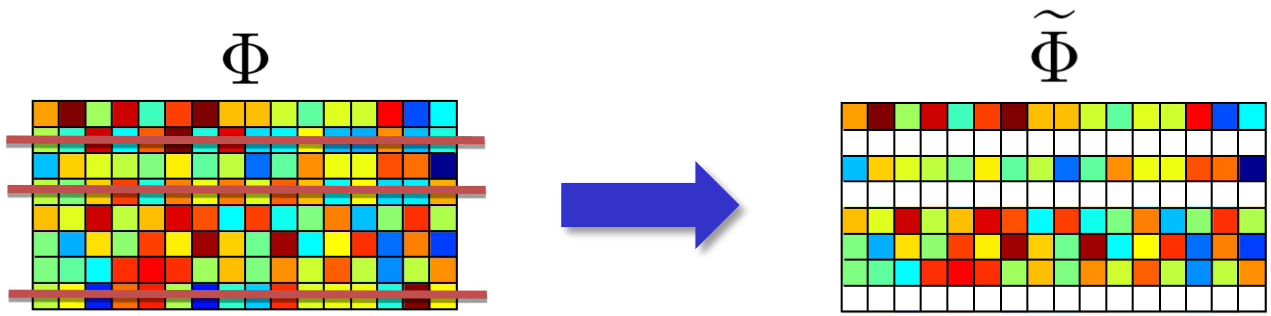
Signal Recovery with Quantization



- Most algorithms are designed for *bounded errors*
- Finite-range quantization leads to *saturation* and *unbounded errors*
- Being able to handle saturated measurements is ***critical*** in any real-world system

Saturation Strategies

- **Rejection:** Ignore saturated measurements



- **Consistency:** Retain saturated measurements. Use them only as inequality constraints on the recovered signal

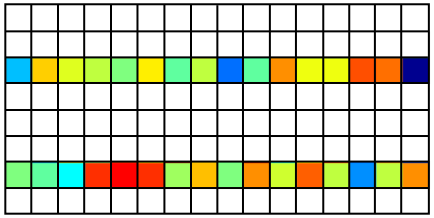
- If the rejection approach works, the consistency approach should automatically do better

Rejection and Democracy

- The RIP is ***not sufficient*** for the rejection approach

- Example: $\Phi = I$
  - perfect isometry
  - *every* measurement must be kept

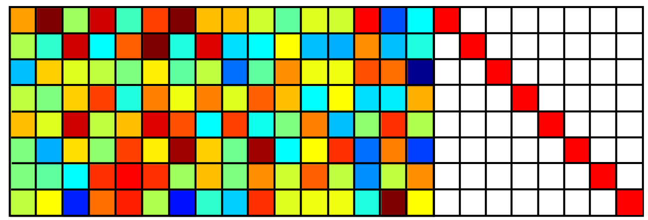
- We would like to be able to say that ***any*** submatrix of Φ with sufficiently many rows will still satisfy the RIP



- Strong, ***adversarial*** form of "democracy"

Sketch of Proof

- Step 1: Concatenate the identity to Φ



Theorem:

If Φ is a sub-Gaussian matrix with

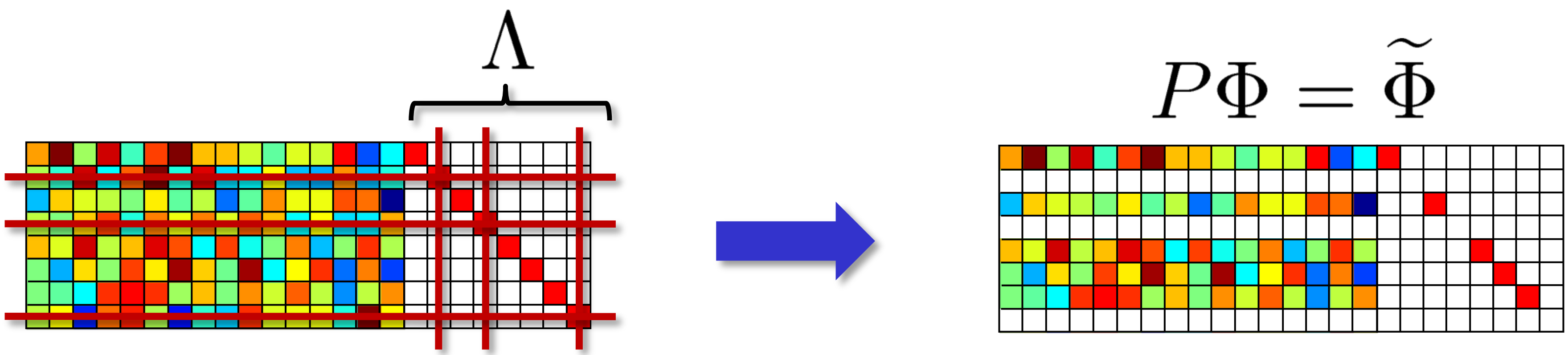
$$M = O\left(K \log\left(\frac{N}{K}\right)\right)$$

then $[\Phi \;\; I]$ satisfies the RIP of order K with probability at least $1 - 3e^{-CM}$.

[M.D., Laska, Boufounos, Baraniuk – Tech. Rep. 2009]

Sketch of Proof

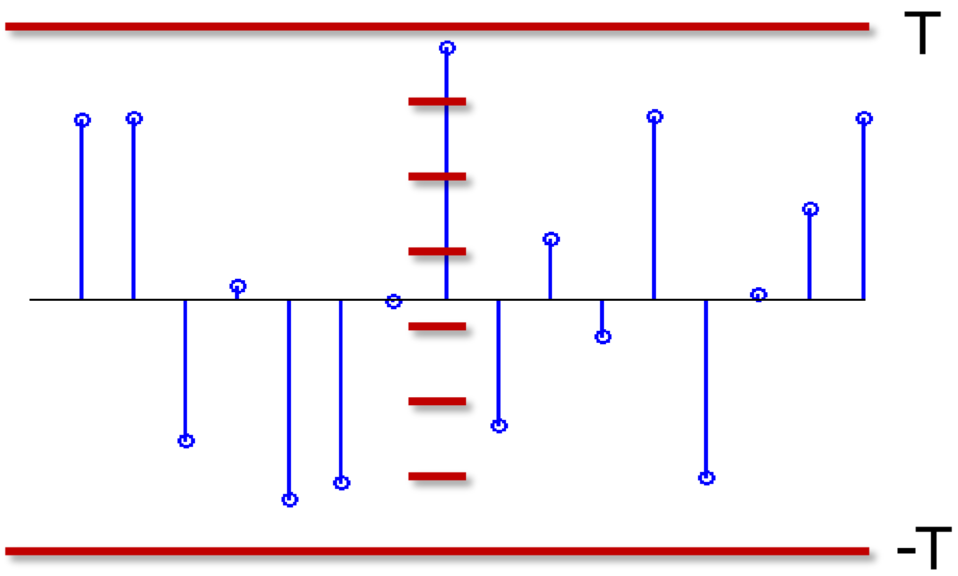
- Step 2: Combine with the “interference cancellation” lemma



- The fact that $[\Phi \ \ I]$ satisfies the RIP implies that if we take D extra measurements, then we can delete $O(D)$ arbitrary rows of Φ and retain the RIP

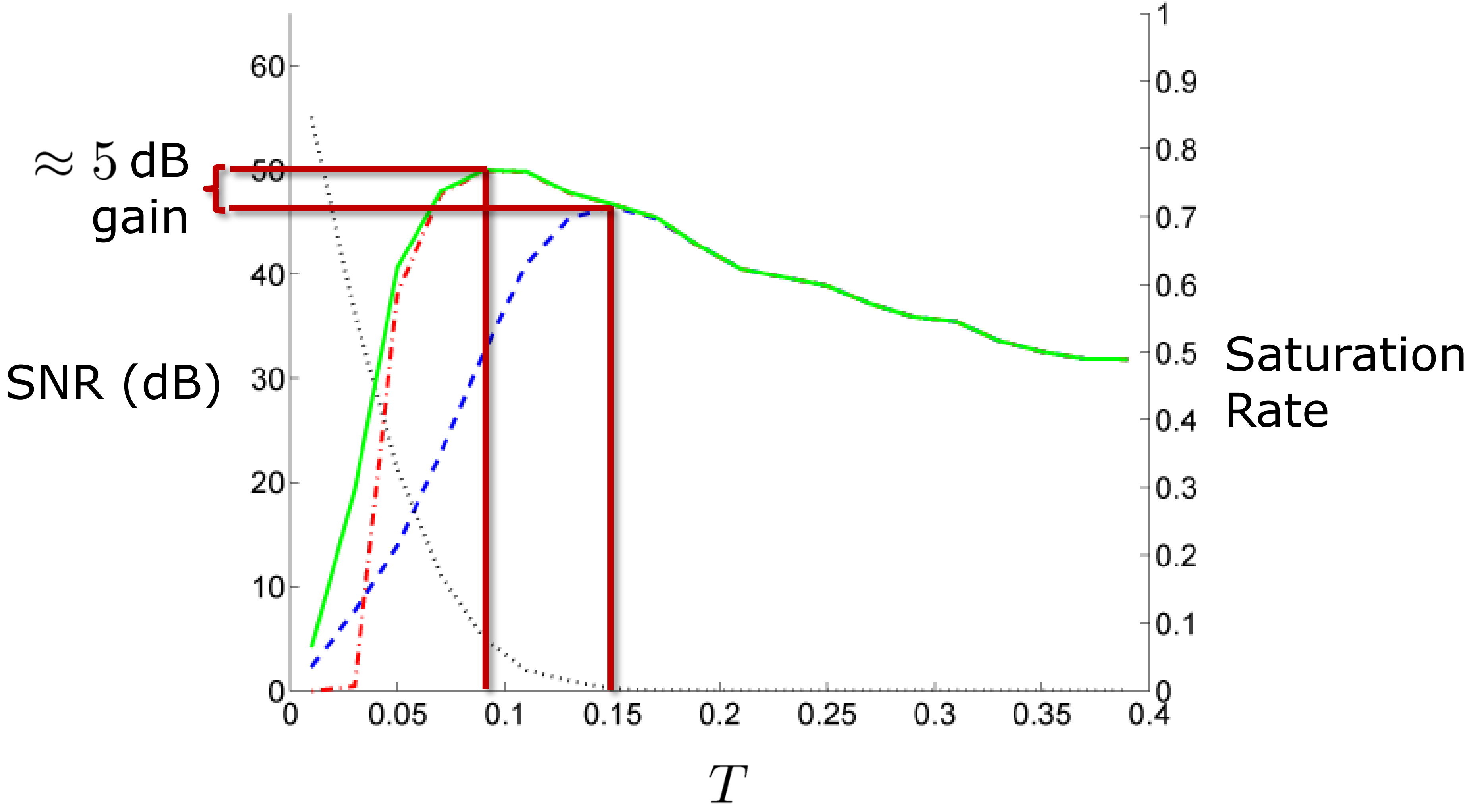
[M.D., Laska, Boufounos, Baraniuk – Tech. Rep. 2009]

Rejection In Practice



$$\mathrm{SNR} = 10\log_{10}\left(\frac{\|x\|_2^2}{\|\widehat{x}-x\|_2^2}\right)$$

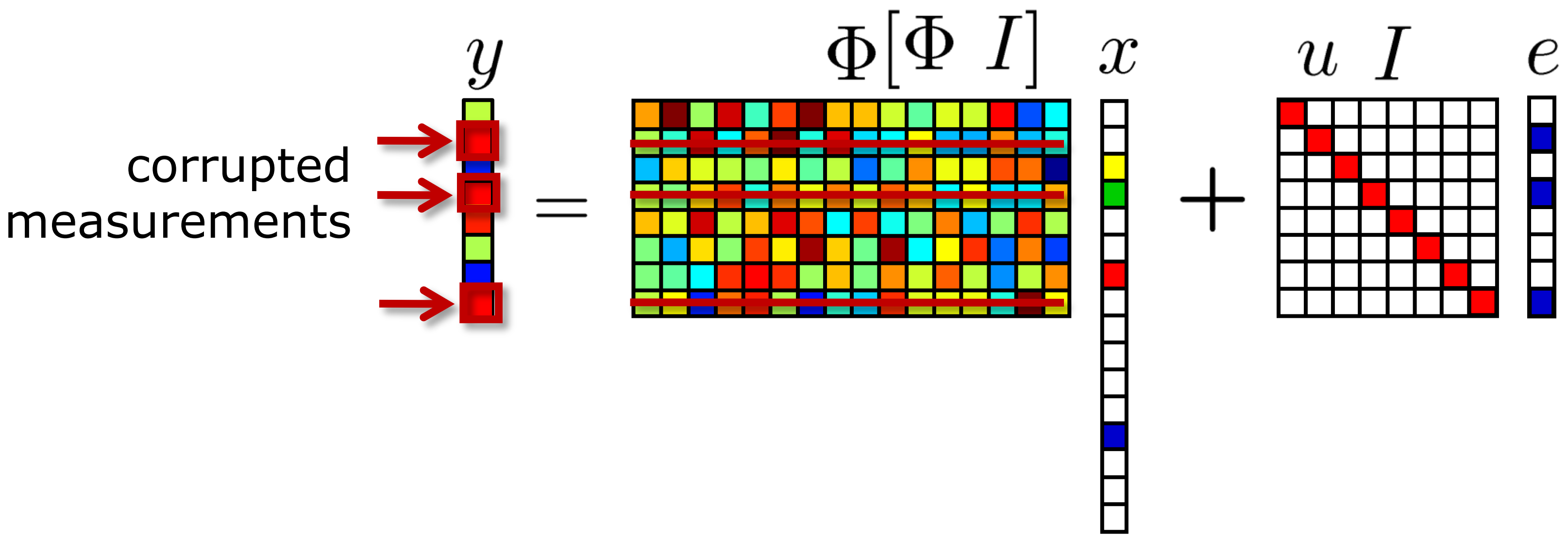
Benefits of Saturation?



[Laska, Boufounos, D, and Baraniuk, 2009]

Recovery in Structured Noise

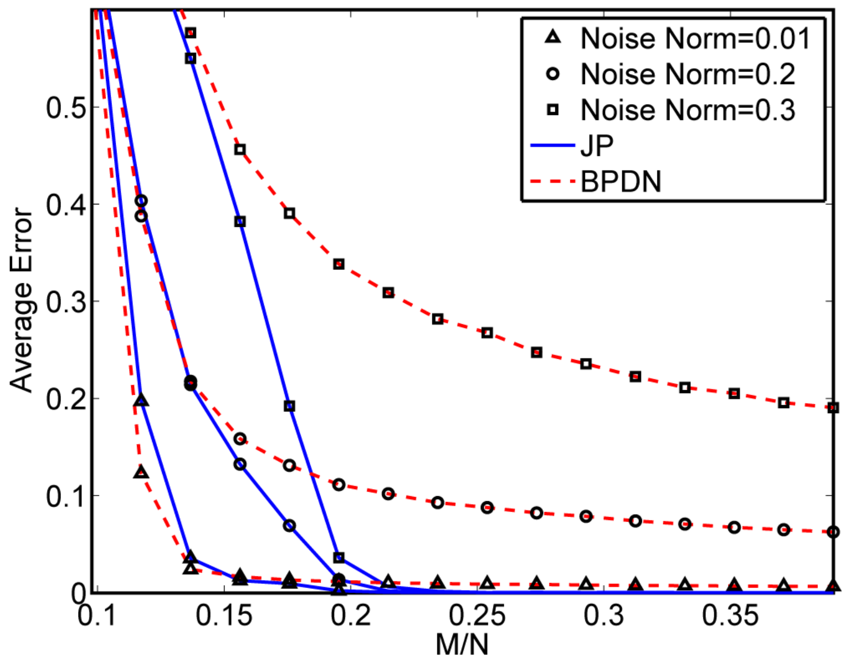
What about structured measurement noise?



Justice Pursuit

- Since $[\Phi \ I]$ satisfies the RIP, we can apply standard sparse recovery algorithms to recover u
- Analogous to joint source/channel coding for sparse signals with erasure channel

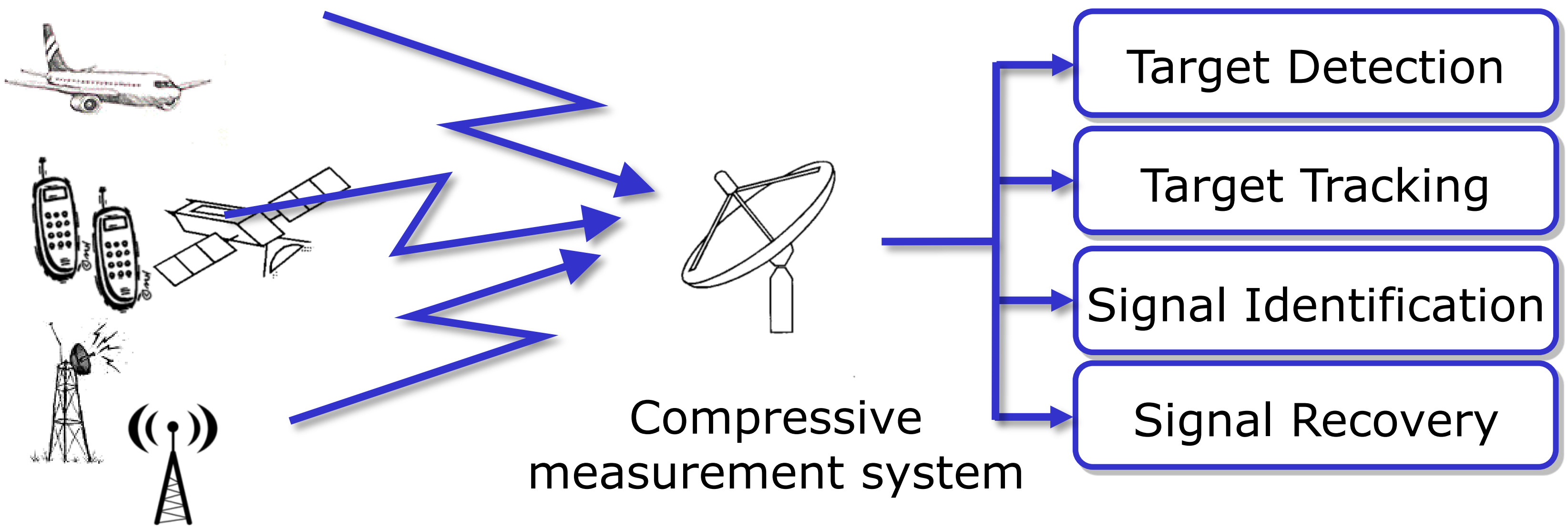
Fixed $\kappa = 10$



Compressive Signal Processing

Compressive Signal Processing

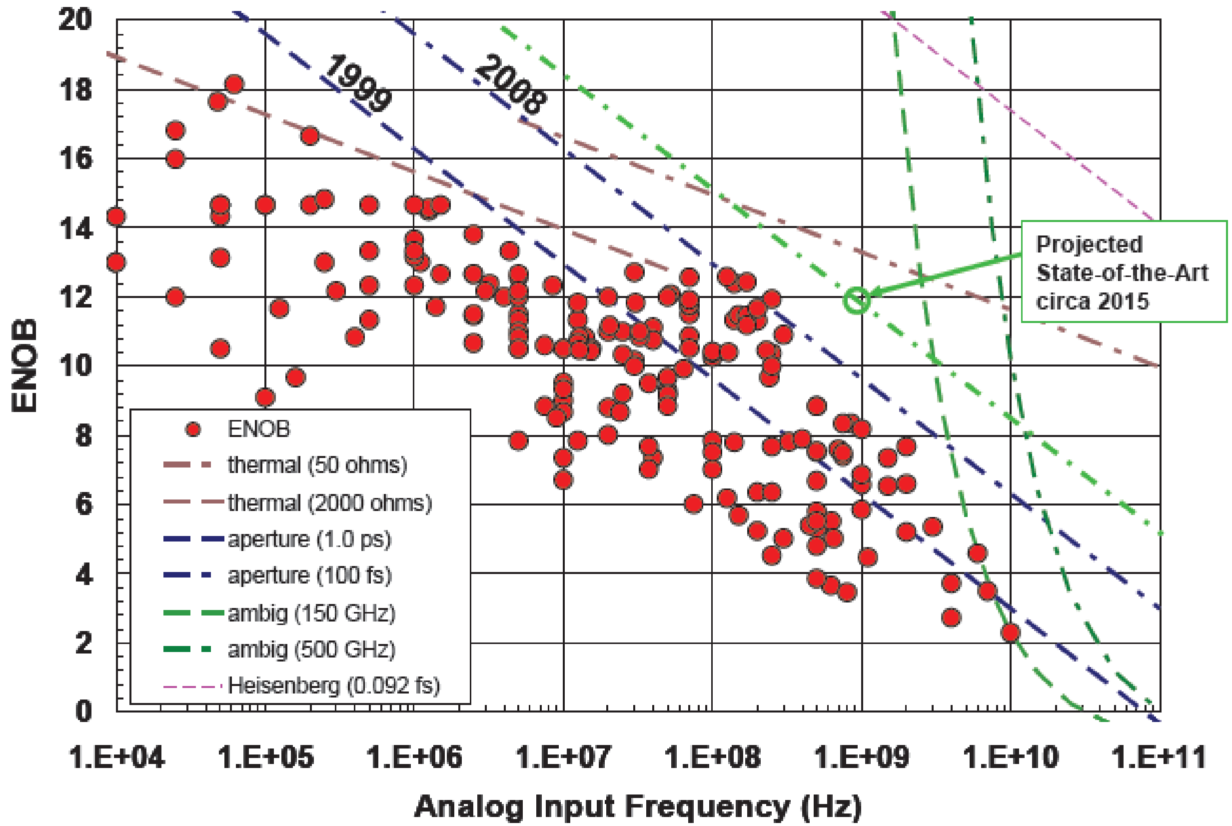
Random measurements are ***information scalable***



When and how can we directly solve signal processing problems directly from compressive measurements?

Compressive ADCs

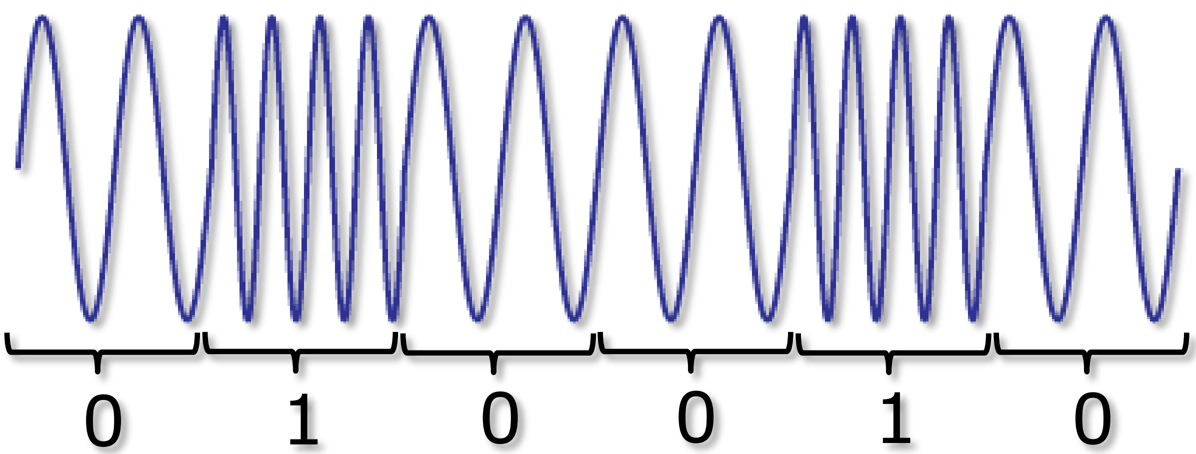
DARPA "Analog-to-information" program:
Build high-rate ADC for signals with sparse spectra



From: R.H. Walden, "Analog to Digital Converters and Associated IC Technologies," 2008

Example: FM Signals

- Can we directly recover a *baseband voice signal* without recovering the modulated waveform?
- Suppose we have compressive measurements of a digital communication signal (FSK modulated)

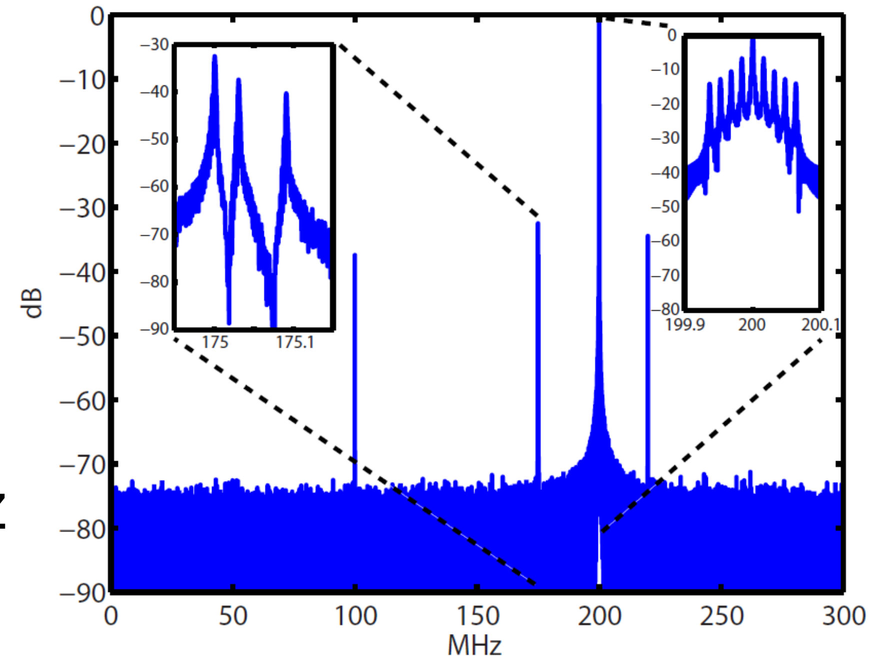


- Can we directly recover the encoded *bitstream* without first recovering the measured waveform?

Compressive Radio Receivers

Example Scenario

- 300 MHz bandwidth
- 5 FM signals (12 kHz)
- TV station interference
- Acquire compressive measurements at 30 MHz (20 x undersampled)



We must simultaneously solve several problems

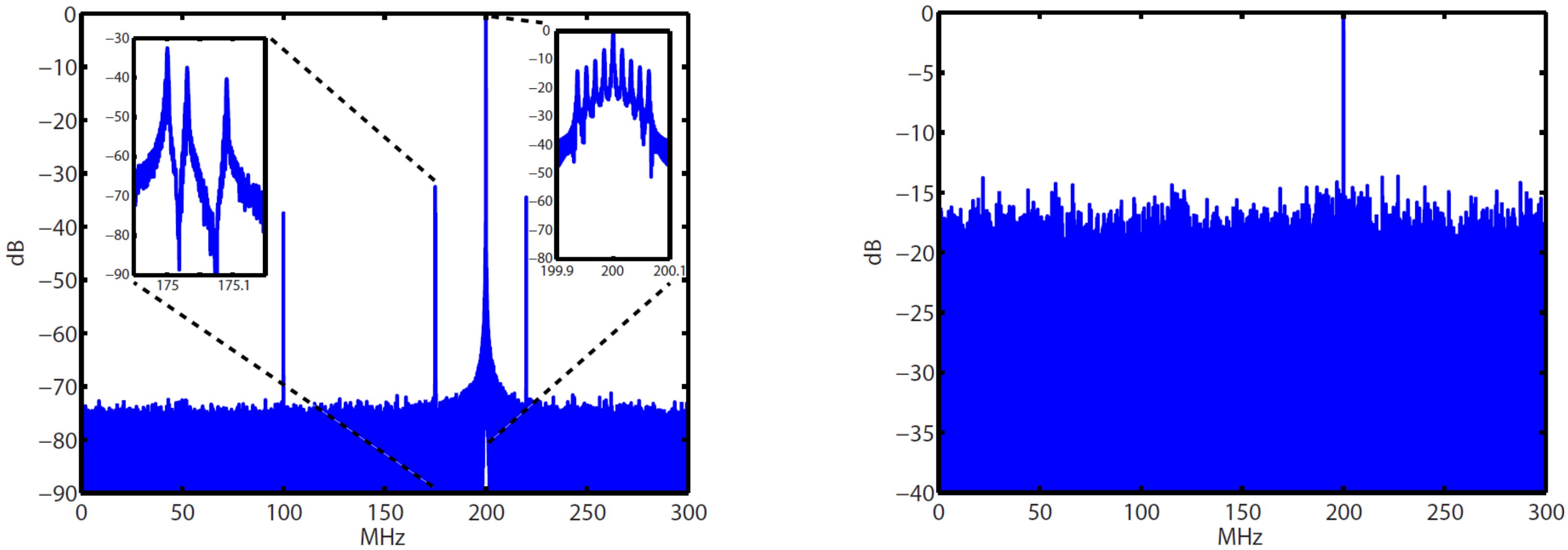
y → detect signal energy → cancel known interferers → filter signals of interest → demod → baseband signals

Energy Detection

We need to identify where in frequency the important signals are located

Correlate measurements with projected tones

$$\widehat{F}(k) = |\langle \Phi \cos(2\pi f_k t), y \rangle|$$



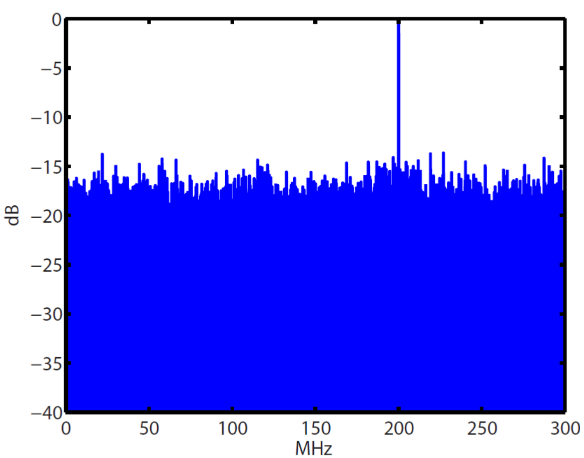
[M.D., Schnelle, Slavinsky, Baraniuk, Wakin, Boufounos – In Prep. 2010]

Filtering

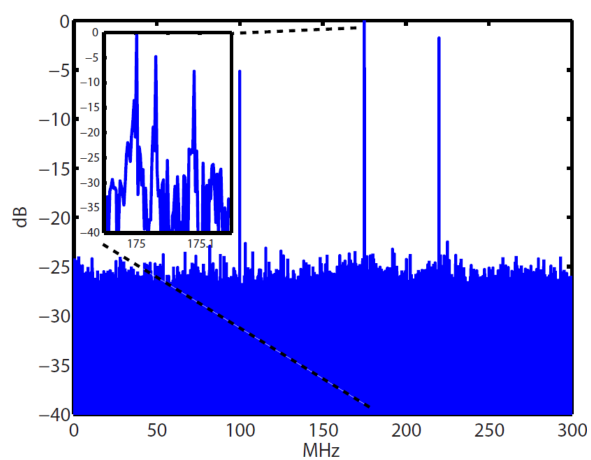
If we have multiple signals, must be able to filter to isolate and cancel interference

$$P = I - \Phi S (\Phi S)^{\dagger}$$

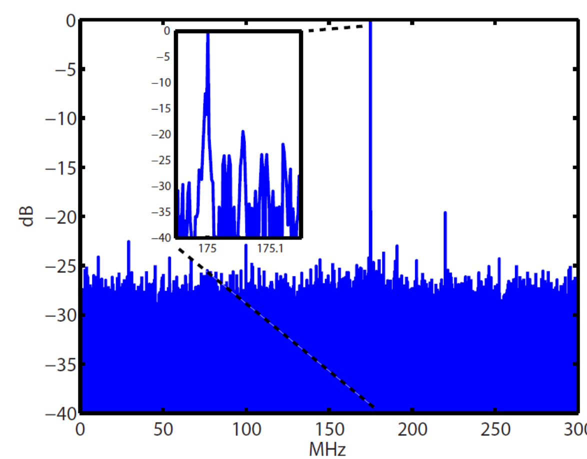
S : Discrete prolate spheroidal sequences



original



after interference cancellation

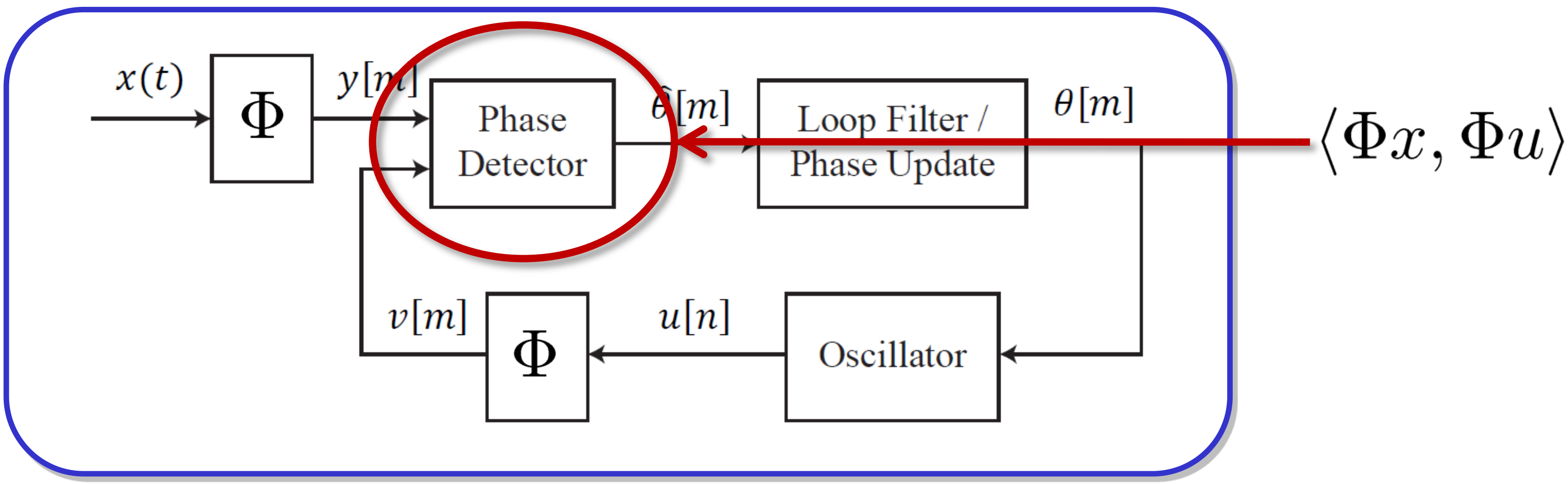


after isolation filtering

[M.D., Schnelle, Slavinsky, Baraniuk, Wakin, Boufounos – In Prep. 2010]

Unsynchronized Demodulation

We can use a phase-locked-loop (PLL) to track deviations in frequency by directly operating on compressive measurements



We can directly demodulate signals from compressive measurements *without recovery*

[M.D., Schnelle, Slavinsky, Baraniuk, Wakin, Boufounos – In Prep. 2010]

CSP – Summary

- **Compressive signal processing**
  - integrates sensing, compression, processing
  - exploits signal sparsity/compressibility
  - enables new sensing modalities, architectures, systems
  - exploits randomness at many levels

- Why CSP works: preserves information in signals with concise geometric structure

  sparse signals | manifolds | low-dimensional models

- **Information scalability** for compressive inference
  - compressive measurements ~ sufficient statistics
  - much less computation required than for recovery

Looking Forward

Some Open Problems

- Links with **information theory**
  - ex: random projection design via codes
  - ex: new decoding algorithms (BP, etc.)
  - ex: democracy and multiple description coding
- Links with **machine learning**
  - ex: Johnson-Lindenstrauss, manifold embedding, RIP
- **Processing/inference** on random measurements
- **Multi-signal** CSP
  - sensor networks, localization, multi-modal data…
- **New sensors**
  - single-pixel gas sensor, hyperspectral cameras
  - scientific imaging (astronomy, microscopy)
  - genomic data, DNA microarrays

Beyond Sparsity

- Learned dictionaries, structured sparsity
- Manifold models
  - connections to “finite rate of innovation”
- Low-rank matrix models
- Models for non-numerical data
  - graphical models

Acquisition

- how to design Φ
- practical devices
- adaptivity

Recovery

- practical algorithms
- robust
- stable

Processing

- classification
- estimation
- learning

Acknowledgements

- Rich Baraniuk
- Ron DeVore
- Piotr Indyk
- Mark Embree
- Kevin Kelly
- Petros Boufounos
- Marco Duarte
- John Treichler
- Mike Wakin
- Chinmay Hegde
- Jason Laska
- Stephen Schnelle

"I not only use all the brains I have, but all I can borrow."

-Woodrow Wilson

More Information

http://dsp.rice.edu/~md

md@rice.edu