Beyond Nyquist

Joel A. Tropp

Applied and Computational Mathematics
California Institute of Technology
jtropp@acm.caltech.edu

With M. Duarte, J. Laska, R. Baraniuk (Rice DSP),
D. Needell (UC-Davis), and J. Romberg (Georgia Tech)

Research supported in part by NSF, DARPA, and ONR

The Sampling Theorem

Theorem 1. *Suppose f is a continuous-time signal whose highest frequency is at most $W/2$ Hz. Then*

$$f(t) = \sum_{n \in \mathbb{Z}} f\left(\frac{n}{W}\right) \operatorname{sinc}(Wt - n).$$

where $\operatorname{sinc}(x) = \sin(\pi x)/\pi x$.

- The *Nyquist rate* W is twice the highest frequency
- The *cardinal series* represents a bandlimited signal by uniform samples taken at the Nyquist rate

Analog-to-Digital Converters (ADCs)

- An ADC consists of a *low-pass filter*, a *sampler* and a *quantizer*
- For sampling rate R, low-pass filter has cutoff $R/2$ to prevent aliasing
- Ideal sampler produces a sequence of amplitude values:

$$f \longmapsto \{f(nT) : n \in \mathbb{Z}\}$$

  where the sampling interval $T = R^{-1}$

- The quantizer maps the real sample values to a discrete set of levels
- Commonly, analog signals are acquired by sampling at the Nyquist rate and samples are processed with digital technology

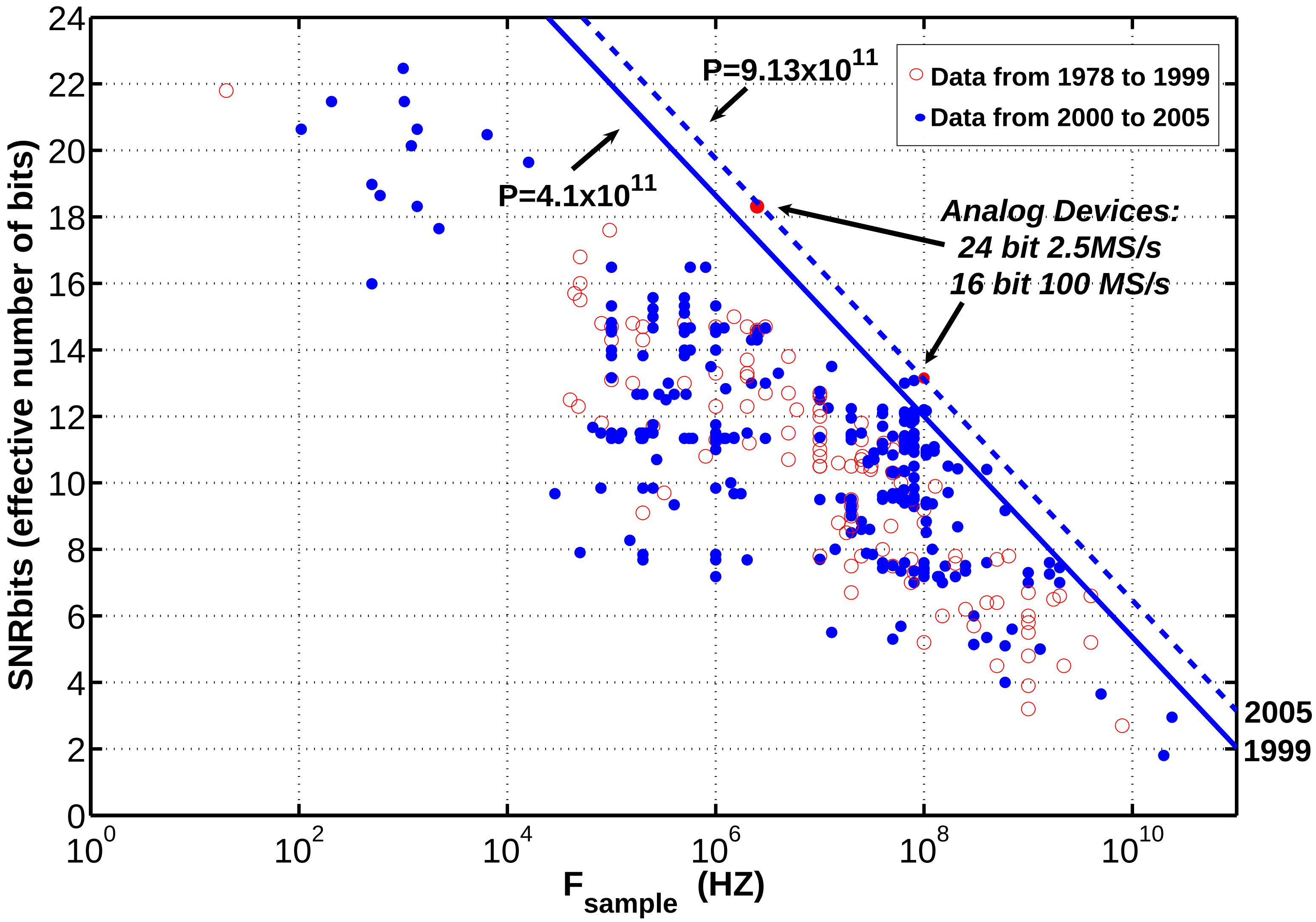
ADCs: State of the Art

- The best current technology (2005) gives
  - 18 effective bits at 2.5 MS/s (MegaSamples/sec)
  - 13 effective bits at 100 MS/s
- Performance degradation about 1 effective bit per frequency octave
- The standard performance metric is

$$P = 2^{\#\text{ effective bits}} \cdot \text{sampling frequency}$$

- At all sampling rates, one effective bit improvement every 6 years

References: [Walden 1999, 2006]



P=9.13x10^11
P=4.1x10^11
Data from 1978 to 1999
Data from 2000 to 2005
Analog Devices:
24 bit 2.5MS/s
16 bit 100 MS/s
2005
1999
SNRbits (effective number of bits)
F sample (HZ)

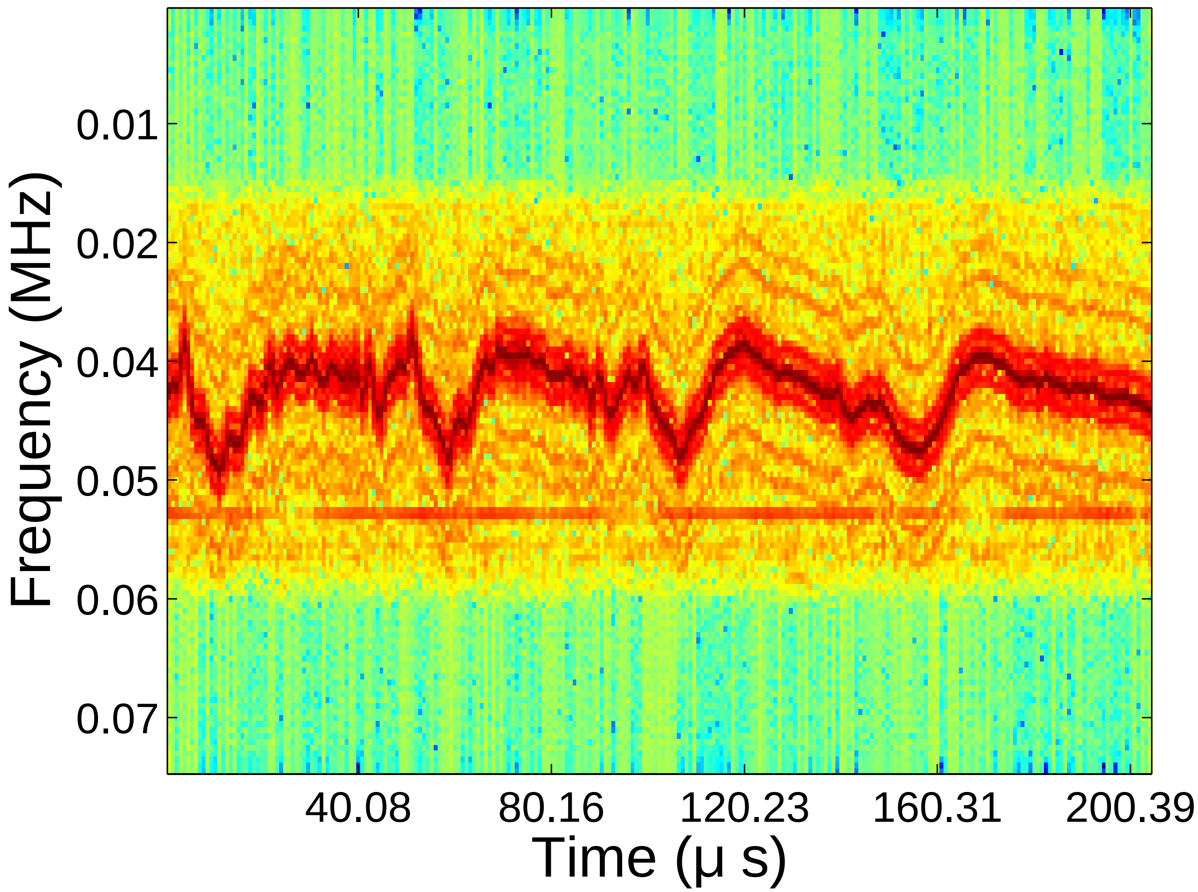
Train Wreck

- Modern applications already exceed ADC capabilities
- The Moore's Law for ADCs is too shallow to help

Conclusion:
We need fundamentally new approaches

Idea: Exploit structure...

Example: An FM Signal



Data provided by L3 Communications

Sparse, Bandlimited Signals

A *normalized* model for signals sparse in time–frequency:

- Let W exceed the signal bandwidth (in Hz)
- Let $\Omega \subset \{-W/2+1, \dots, -1, 0, 1, \dots, W/2\}$ be *integer* frequencies
- For each one-second time interval, signal has the form

$$f(t) = \sum_{\omega \in \Omega} a(\omega)\, \mathrm{e}^{2\pi \mathrm{i} \omega t} \qquad \text{for } t \in [0, 1)$$

- The set Ω of frequencies can change every second
- In each time interval, number of frequencies $|\Omega| = K \ll W$

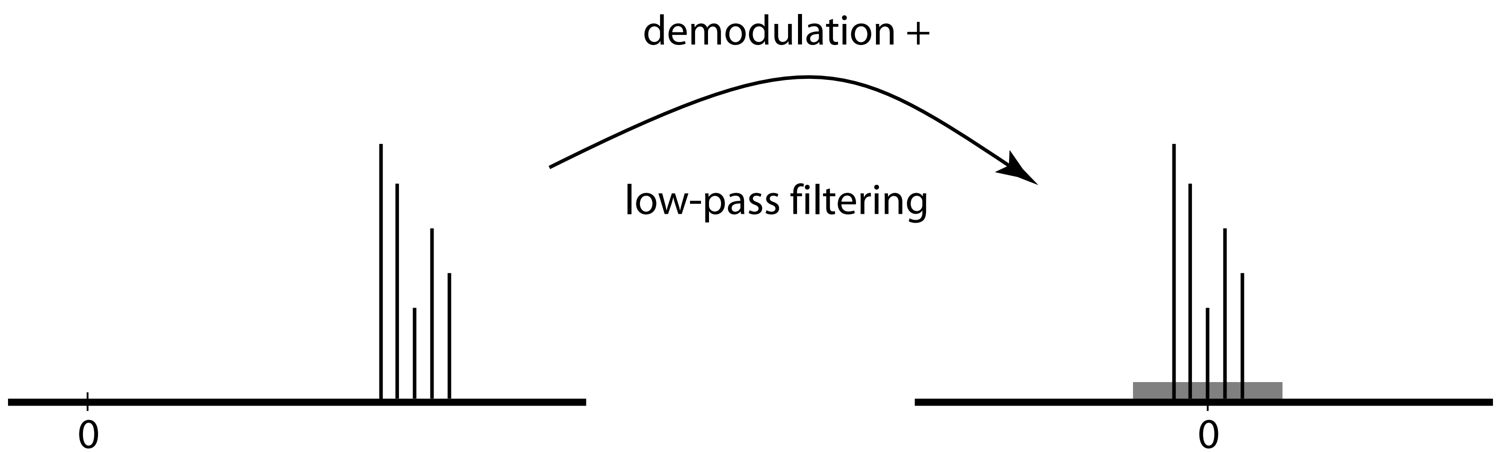
Other models: [Mishali–Eldar–T 2008, 2009]

Information and Signal Acquisition

- Signals in our model contain little information
  - In each time interval, have K frequencies and K coefficients
  - Total: About $K \log W$ bits of information
- Idea: We should be able to acquire signals with about $K \log W$ nonadaptive measurements
- Challenge: Achieve goal with current ADC hardware
- Approach: Use randomness!

Random Demodulator: Intuition

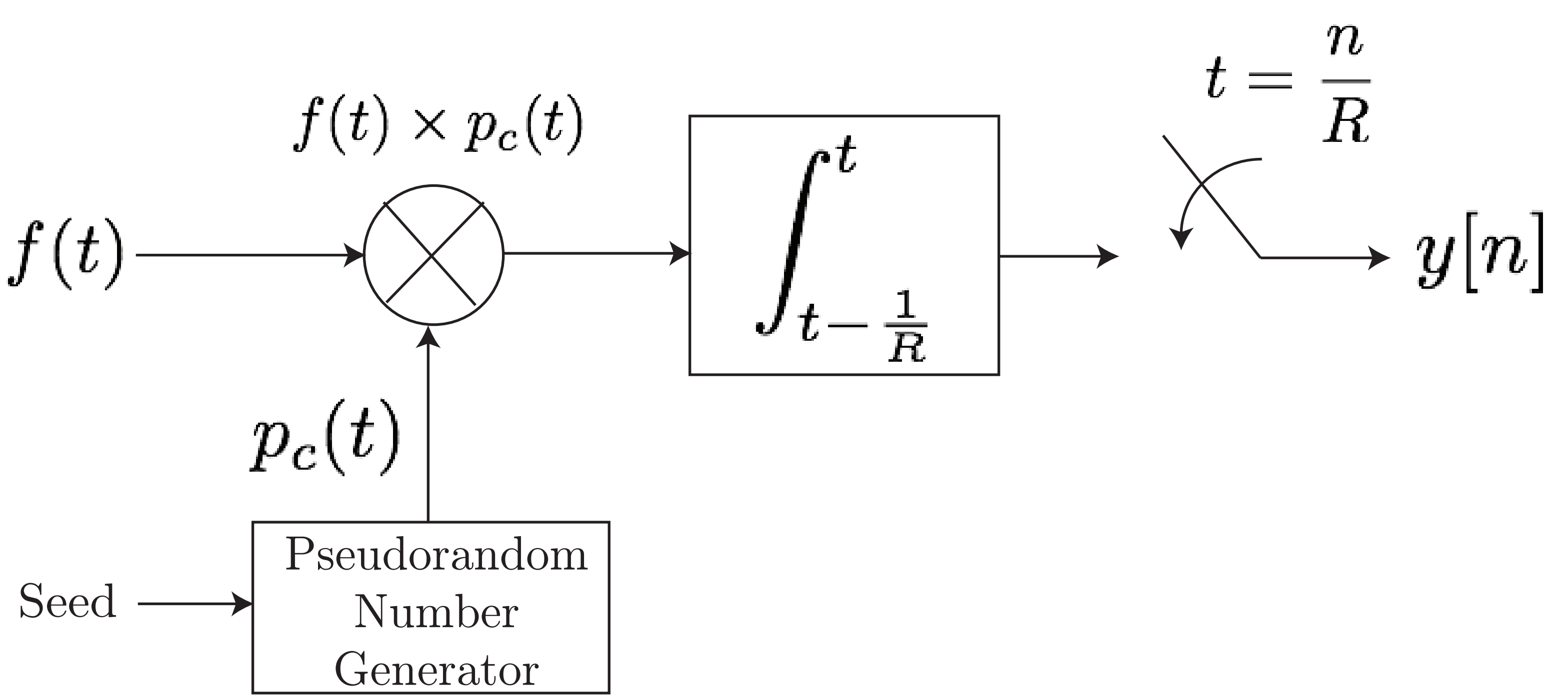
- With clustered frequencies, demodulate to baseband and low-pass filter



- Don't know locations, so demodulate *randomly* and low-pass filter

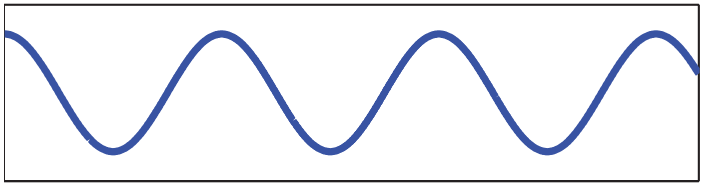
- Analogy with spread-spectrum communications methods

Random Demodulator: System Model



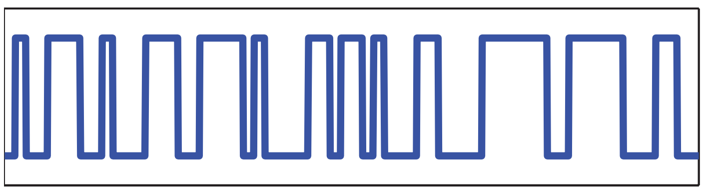
- $p_c(t)$ alternates randomly between levels ± 1 at Nyquist rate W
- Sampler runs at rate $R \ll W$

input signal $x(t)$



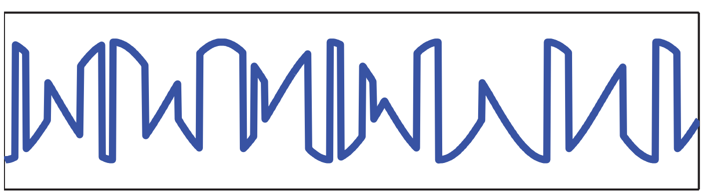
$\times$

pseudorandom sequence $p_c(t)$



$=$

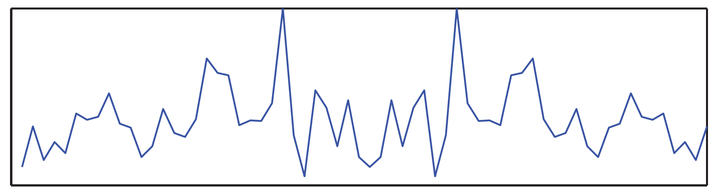
modulated input



input signal $X(\omega)$

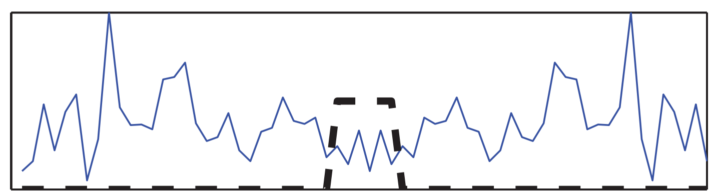
$*$

pseudorandom sequence spectrum $P_c(\omega)$

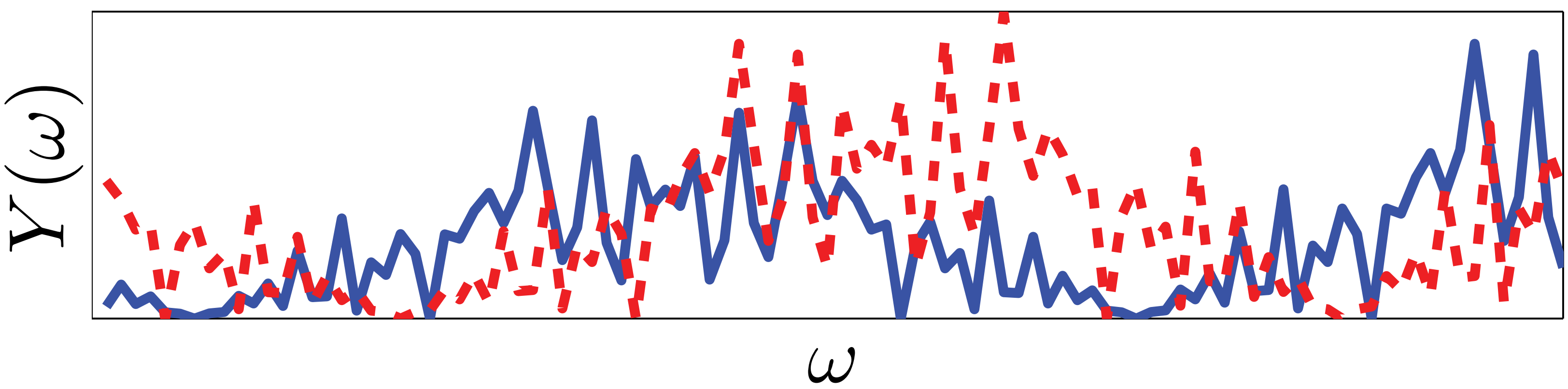


$=$

modulated input and integrator (low–pass filter)



Exploded View of Passband



Reconstruction from Samples

- The matrix $\mathbf{\Phi}$ summarizes the action of the random demodulator

$$\mathbf{\Phi} = \boldsymbol{HDF} : \mathbb{C}^W \longrightarrow \mathbb{C}^R$$

- Maps a (sparse) amplitude vector $\boldsymbol{s}$ to a vector of samples $\boldsymbol{y}$
- Given samples $\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{s}$, signal reconstruction can be formulated as

$$\widehat{\boldsymbol{s}} = \arg\min \|\boldsymbol{c}\|_0 \quad \text{subject to} \quad \mathbf{\Phi}\boldsymbol{c} = \boldsymbol{y}$$

- The ℓ_0 function counts nonzero entries of a vector

Signal Reconstruction Algorithms

Approach 1: Convex Relaxation

- Can often find sparsest amplitude vector by solving

$$\widehat{\boldsymbol{s}} = \arg\min \|\boldsymbol{c}\|_1 \quad \text{subject to} \quad \boldsymbol{\Phi}\boldsymbol{c} = \boldsymbol{y} \tag{P1}$$

Approach 2: Greedy Pursuit

- Identify a small set of significant frequencies and iteratively refine
- Examples: OMP and CoSaMP

References: [Candès et al. 2006, Donoho 2006, Tropp–Gilbert 2007, Tropp–Needell 2008]

Shifting the Burden

- These algorithms are much more computationally intensive than linear reconstruction via cardinal series
- Move the work from the analog front end to the digital back end

Moore's Law for ICs
saves us from
Moore's Law for ADCs!

Theoretical Analysis

Theorem 2. [T 2007] *Suppose the sampling rate satisfies*

$$R \geq \mathrm{C} \cdot K \cdot \log^6 W$$

Then the matrix $\mathbf{\Phi}$ *has the restricted isometry property*

$$(1 - \mathrm{c})\, \|\boldsymbol{x}\|_2^2 \leq \|\mathbf{\Phi}\boldsymbol{x}\|_2^2 \leq (1 + \mathrm{c})\, \|\boldsymbol{x}\|_2^2 \quad \textit{when} \quad \|\boldsymbol{x}\|_0 \leq 2K$$

except with probability W^{-1}.

- Abstract property supports efficient sampling and reconstruction
- Intuition: Sampling operator preserves geometry of sparse vectors

Simulations

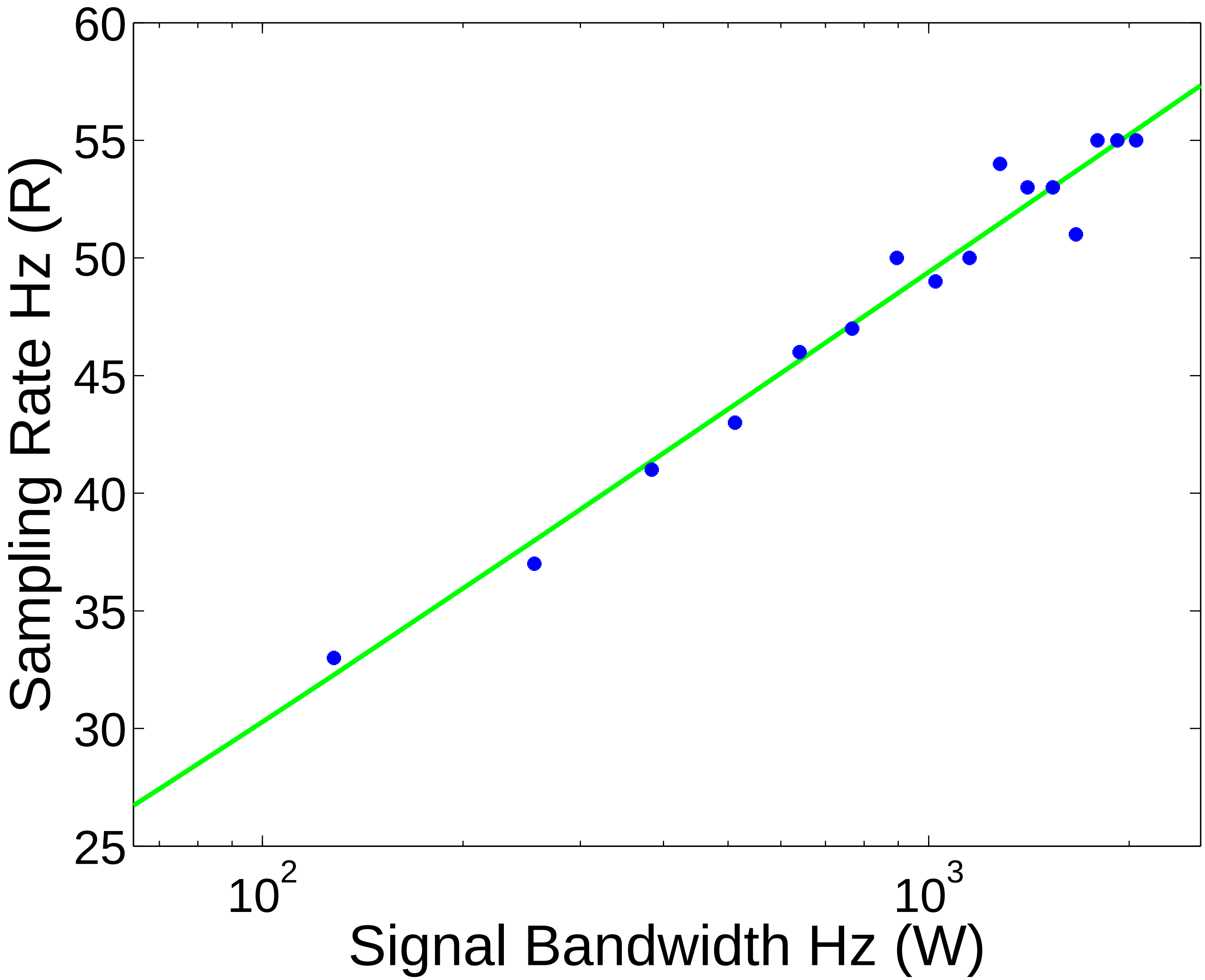
Goal: Estimate sampling rate R to achieve success probability 99%

For each of 500 trials,

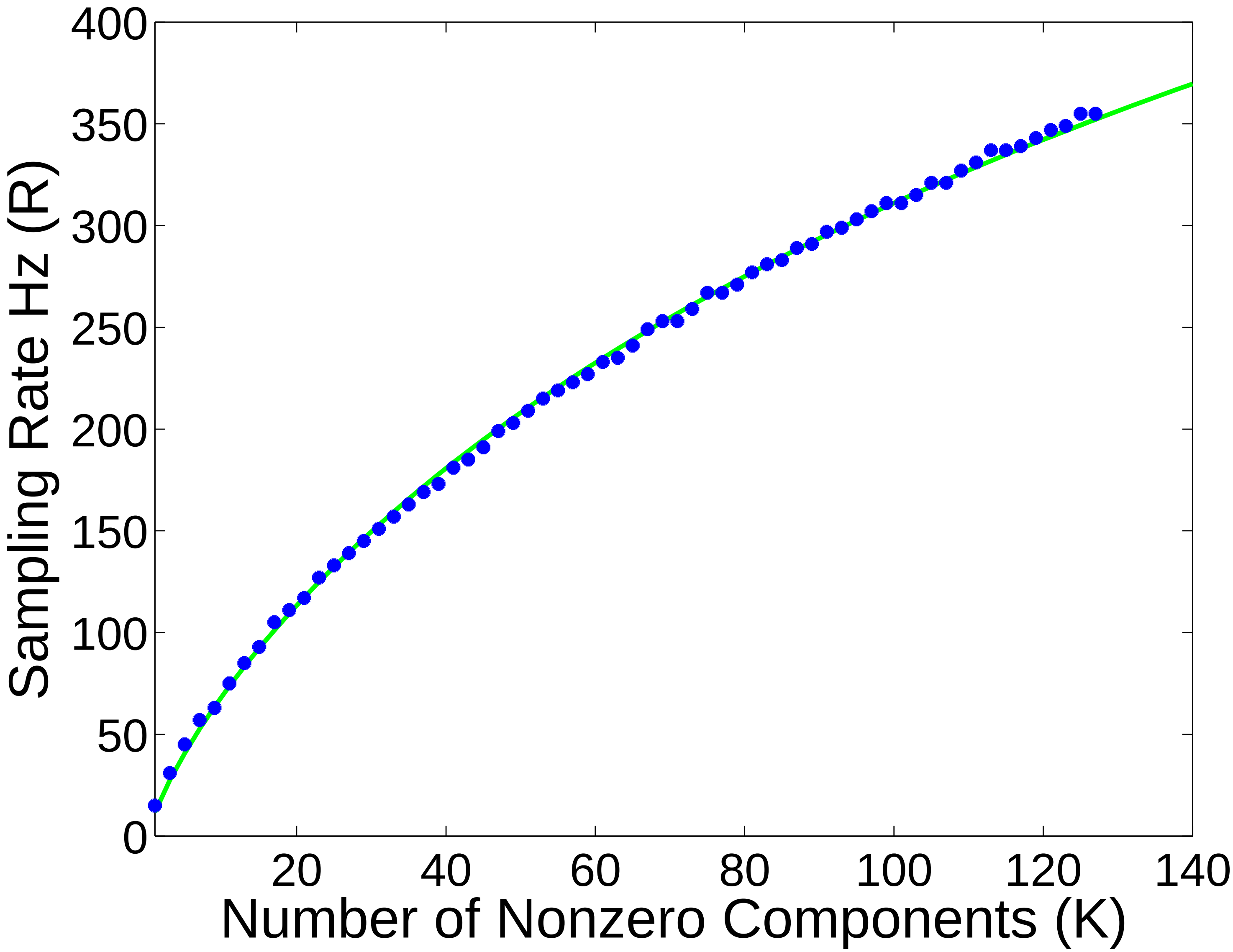
- Draw a random demodulator with dimensions $R \times W$
- Choose a random set of K frequencies
- Set their amplitudes equal to one
- Take measurements of the signal
- Recover with ℓ_1 minimization (via IRLS)

Define *success* at rate R when 99% of trials result in

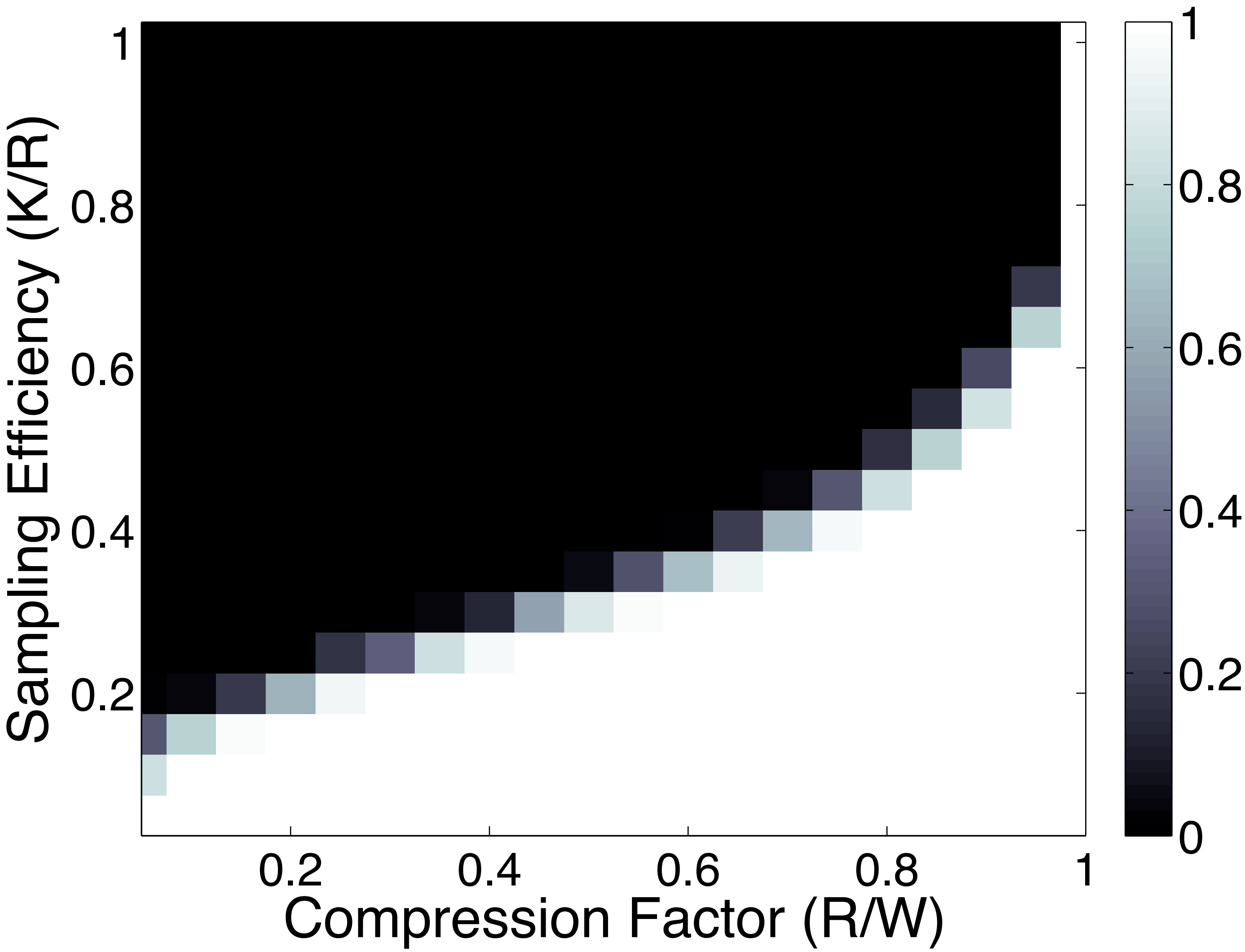
$$\|\boldsymbol{s} - \widehat{\boldsymbol{s}}\| < \varepsilon_{\mathrm{mach}}$$



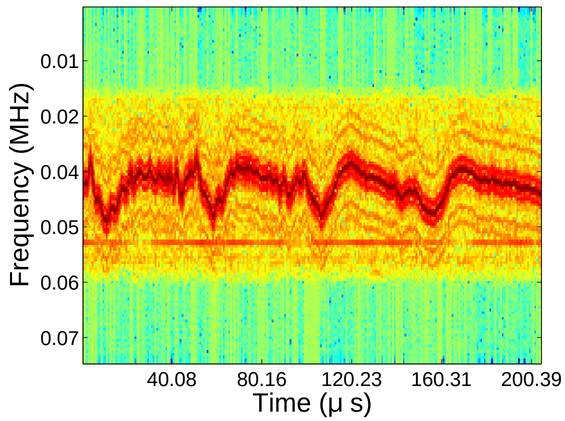
$K = 5$, regression line $R = 1.69K \log(W/K + 1) + 4.51$



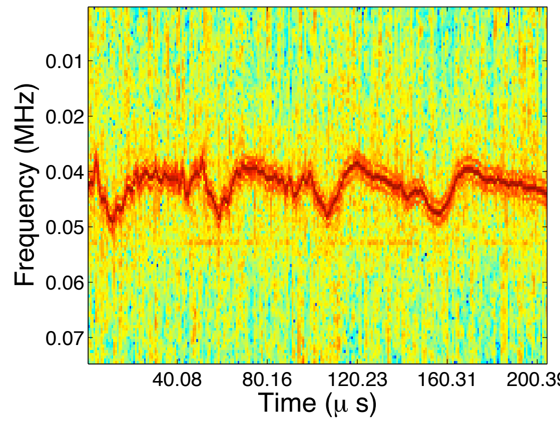
$W = 512$, regression line $R = 1.71K\log(W/K + 1) + 1.00$


Sampling Efficiency (K/R)
Compression Factor (R/W)
0.2
0.4
0.6
0.8
1
0

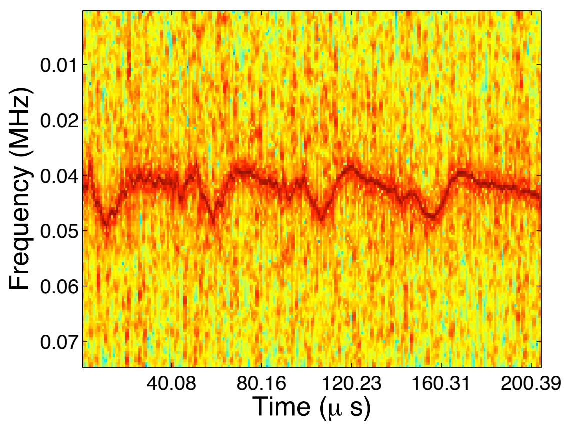
Reconstruction of FM Signal



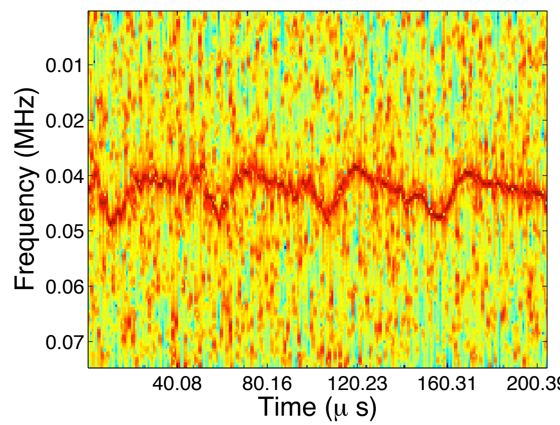
(a) Original Signal (1.25 MHz)



(b) Rand Demod (0.63 MHz)

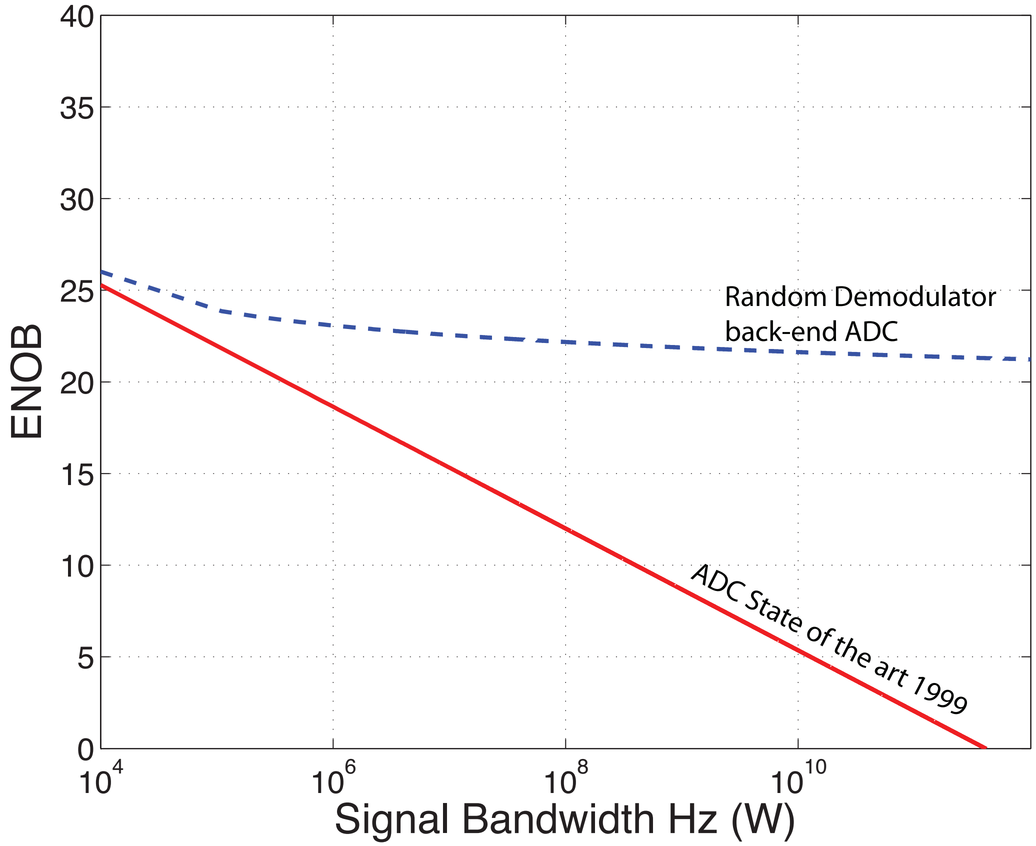


(c) Rand Demod (0.31 MHz)



(d) Rand Demod (0.16 MHz)

On Walden Pond



Fixed sparsity $K = 5000$

To learn more...

E-mail: `jtropp@acm.caltech.edu`

Web: `http://acm.caltech.edu/~jtropp`

`http://www.dsp.rice.edu/cs/`

`http://www.dsp.rice.edu/a2i/`

Papers

- Mishali, Eldar, T, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," in preparation
- T, Romberg, Rice DSP, "Beyond Nyquist: Efficient sampling of sparse, bandlimited signals," submitted 2009
- Mishali, Eldar, T, "Efficient sampling of sparse, wide band signals," IEEEI 2008
- Needell and T, "CoSaMP: Iterative signal recovery from incomplete and inaccurate measurements," ACHA 2008
- T and Gilbert, "Signal recovery from random measurements via Orthogonal Matching Pursuit," Trans. IT 2007