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

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

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

- 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:
  \[ f \mapsto \{ f(nT) : n \in \mathbb{Z} \} \]
  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]
Beyond Nyquist (Duke CS Workshop, Durham, Feb. 2009)
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
A *normalized* model for signals sparse in time–frequency:

- Let $W$ exceed the signal bandwidth (in Hz)
- Let $\Omega \subset \{-W/2 + 1, \ldots, -1, 0, 1, \ldots, W/2\}$ be *integer* frequencies
- For each one-second time interval, signal has the form
  \[
  f(t) = \sum_{\omega \in \Omega} a(\omega) e^{2\pi i \omega t} \quad \text{for } t \in [0, 1)
  \]
- The set $\Omega$ 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

\[ f(t) \times p_c(t) \]

\[ \int_{t-\frac{1}{R}}^{t} \]

\[ t = \frac{n}{R} \]

\[ y[n] \]

- \( p_c(t) \) alternates randomly between levels \( \pm 1 \) 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 $X(\omega)$

\[ \ast \]

pseudorandom sequence spectrum $P_c(\omega)$

\[ = \]

modulated input and integrator (low-pass filter)
Exploded View of Passband
Reconstruction from Samples

The matrix $\Phi$ summarizes the action of the random demodulator

$$\Phi = HDF : \mathbb{C}^W \rightarrow \mathbb{C}^R$$

Maps a (sparse) amplitude vector $s$ to a vector of samples $y$

Given samples $y = \Phi s$, signal reconstruction can be formulated as

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

The $\ell_0$ function counts nonzero entries of a vector
Signal Reconstruction Algorithms

Approach 1: Convex Relaxation

- Can often find sparsest amplitude vector by solving

\[ \hat{s} = \arg \min \|c\|_1 \quad \text{subject to} \quad \Phi c = y \]  

(P1)

Approach 2: Greedy Pursuit

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

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

\[ R \geq C \cdot K \cdot \log^6 W \]

Then the matrix \( \Phi \) has the restricted isometry property

\[
(1 - c) \| \mathbf{x} \|_2^2 \leq \| \Phi \mathbf{x} \|_2^2 \leq (1 + c) \| \mathbf{x} \|_2^2 \quad \text{when} \quad \| \mathbf{x} \|_0 \leq 2K
\]
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 $\ell_1$ minimization (via IRLS)

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

$$\| s - \hat{s} \| < \varepsilon_{\text{mach}}$$
$K = 5$, regression line $R = 1.69K \log(W/K + 1) + 4.51$
$W = 512$, regression line $R = 1.71K \log(W/K + 1) + 1.00$
Beyond Nyquist (Duke CS Workshop, Durham, Feb. 2009)
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 $K = 5000$
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