Faster than Nyquist, Slower than Tropp

The Constrained Random Demodulator

Andrew Harms   Waheed Bajwa   Robert Calderbank

From Nyquist Sampling to Parameter Estimation

1928 

Shannon-Nyquist Sampling: Perfect reconstruction of any signal with bandwidth W from $2W$ samples

1974 **Rife and Boorstyn**: Information in the signal determines sampling rate

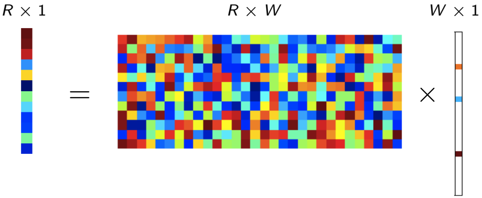
- Sub-Nyquist sampling followed by parameter estimation to identify the frequency, amplitude, and phase offset of an unknown tone

2006 

Compressed Sensing: Systematic exploration of sparsity as a prior

- Chirp Sampling – Applebaum et al.
- Xampling – Eldar et al.
- Random Demodulator – Tropp et al.

Compressed Sensing - More Variables than Equations



$(R << W)$

- **Candés and Tao** Measurement matrix Φ acts as a near isometry on S-sparse vectors x

$$(1 - \delta_S)\|\mathrm{x}\|^2 \leq \|\Phi \mathrm{x}\|^2 \leq (1 + \delta_S)\|\mathrm{x}\|^2$$

- **Theorem** (Signal Reconstruction): If $\mathrm{y} = \Phi \mathrm{x} + w$, $\|w\|_2 \leq \eta$ and Φ is an $R \times W$ matrix satisfying RIP($2S$,δ_{2S}) with $\delta_{2S} \leq \sqrt{2} - 1$, then we can find $\hat{\mathrm{x}}$ from y such that $\|\hat{\mathrm{x}} - \mathrm{x}\|_2 \leq C \max\left\{\eta, \frac{1}{\sqrt{S}}\|\mathrm{x} - \mathrm{x}_S\|_1\right\}$ where x_S is a best S-sparse approximation to x.

The French Revolution in Signal Processing

1795

Prony's Curve Fitting:

- $f(x) \approx C_1 e^{\gamma_1 x} + C_2 e^{\gamma_2 x} + \cdots + C_n e^{\gamma_n x}$
- To solve for the unknowns C_i and γ_i, we need $f(x)$ specified at $N \geq 2n$ points.

1949

Shannon Sampling: If a signal is bandlimited to $1/2T$, then it has $1/T$ degrees of freedom per unit time – *sufficient* sampling

2002

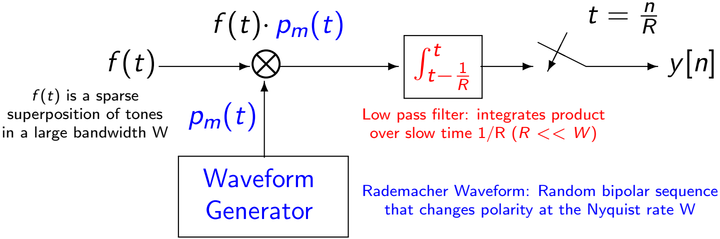
Vetterli – Finite Rate of Innovation: If a signal has C_T degrees of freedom in $[-T/2, T/2]$, then sample at the "Rate of Innovation" – *necessary* sampling

$$\rho = \lim_{T \to \infty} \frac{1}{T} C_T$$

The Random Demodulator

2010:  et al $\quad \Phi = HDF$

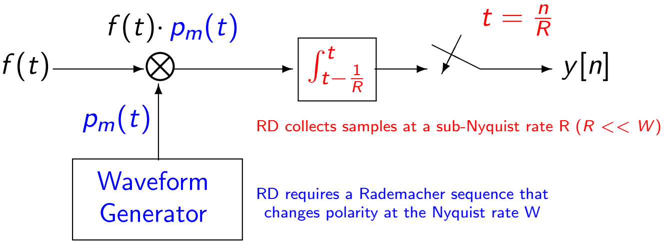
$$D = \text{diag}(\text{Rademacher})$$



$$H = \begin{bmatrix} 1 & 1 & & & & \\ & & 1 & 1 & & \\ & & & & 1 & 1 \end{bmatrix} \quad (R \times W)$$

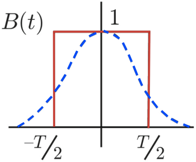
Slower than Nyquist?

- Doubling the sampling rate of a traditional ADC causes a 1 bit reduction in resolution
- Capacitors used to build traditional ADC circuits take time to switch between charged and uncharged states, forcing designers to limit either the sampling rate or the resolution

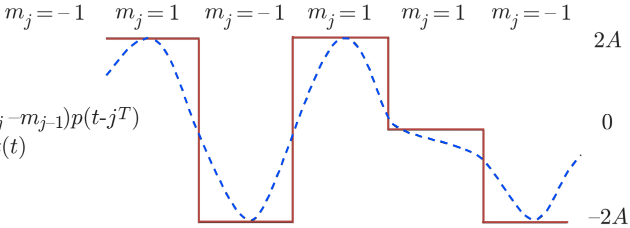


- Generating a waveform that changes polarity at a high rate might be just as hard as building an ADC

Reading Data on Magnetic Media

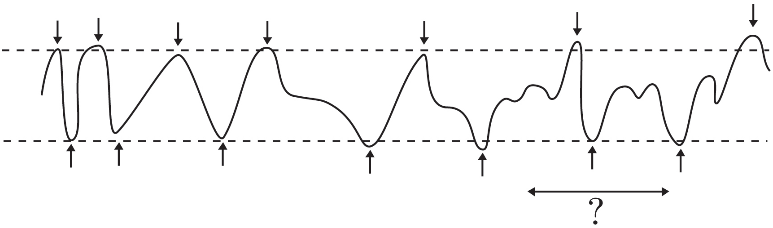


$p(t)$: response to isolated transition
$B(t)$: idealization of $p(t)$



$$y(t) = \sum A(m_j - m_{j-1})p(t-jT) + z(t)$$

Higher Recording Density Means Lower Fidelity



Transitions are indicated by peaks in the waveform that should alternate in sign.

The ability of the readback voltage to support symbol by symbol detection depends on the minimum spacing between transitions. If two transitions are too close the peaks are reduced in amplitude and shifted.

Challenge: How to write more data while separating transitions so they do not interfere.

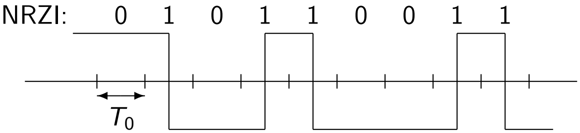
Run Length Limited Codes

NRZI data: 0 – no transition; 1 – transition

(d, k) constrained binary sequences

$$
\begin{aligned}
d &= \text{minimum number of 0's between adjacent 1's} \\
k &= \text{maximum number of 0's between adjacent 1's}
\end{aligned}
$$

$$
\begin{aligned}
T_{\min} &= (d+1)\,T_0 \quad \text{– maintains separation between transitions} \\
T_{\max} &= (k+1)\,T_0 \quad \text{– aids timing recovery}
\end{aligned}
$$



- rate loss and increased correlation from converting arbitrary binary sequences to (d, k) sequences
- rate gain and increased transition time from spacing data bits more finely than transitions

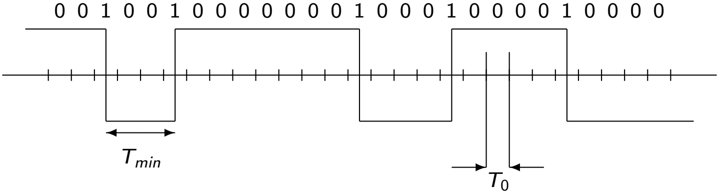
Comparison of Linear Information Densities

Variable Length Fixed Rate (2,7) Code: Rate $\frac{1}{2}$ code found in IBM 3370, 3375 and 3380 disk drives.

Data	Message
10	0100
11	1000
000	000100
010	100100

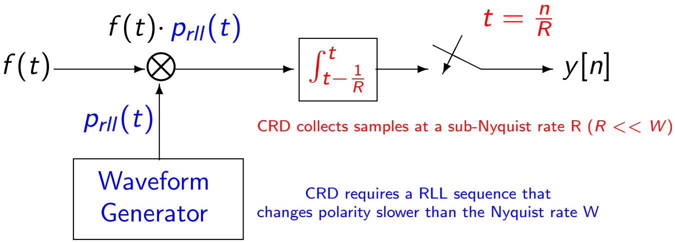
Data	Message
011	001000
0010	00100100
0011	00001000

(2,7) NRZI: Factor of 3 rate gain: $-T_{min} = 3T_0$



Value: 50% gain in a real density over uncoded NRZI

Why *Constrain* the Random Demodulator?



- The RD waveform switches at the Nyquist rate W – the faster the rate, the less ideal is the modulating waveform
- The CRD waveform switches at a rate lower than W – RLL sequences measure time instances at a finer scale than the minimum transition width of the waveform

What is the impact of correlations?

$$\Phi = HDF\colon\ \varphi_{r\omega} = \sum_{\frac{W}{R}(r-1)}^{\frac{W}{R}r} \varepsilon_j f_{j\omega}$$

If Φ is a near isometry on S-sparse vectors

$$\sup_{\Omega, |\Omega| \leq S} ||(\Phi^*\Phi - \mathrm{I})|_{\Omega\times\Omega}|| \approx 0$$

How good is an "average" system?

$$\mathbb{E}[\Phi^*\Phi] = \mathrm{I} + \Delta$$

$$\Rightarrow \sup_{\Omega, |\Omega| \leq S} ||\Delta|_{\Omega\times\Omega}|| \approx 0$$

- $\Delta_{\alpha\omega} = \sum_{j\neq k} \mathbb{E}[\varepsilon_j \varepsilon_k] \eta_{jk} f_{j\alpha}^* f_{k\omega}$, $\eta_{jk} = \langle h_j, h_k \rangle$

Enter the Addressable Spectrum

The spectral norm of Δ

$$(\Delta^*\Delta)_{\alpha\omega} = \sum_j \exp\left(-\frac{2\pi\imath}{W}(\omega-\alpha)j\right)\hat{F}^*(j,\alpha)\hat{F}(j,\omega)$$

- 'Windowed' Spectrum: $\hat{F}(j,\omega) = \sum_{m\neq 0} \eta_{j(j+m)} R_\varepsilon(m) f_{m\omega}$

As $\frac{W}{R}\uparrow$,

$$(\Delta^*\Delta)_{\alpha\omega} \approx \delta_{\alpha\omega}(F_\varepsilon(\alpha)-1)(F_\varepsilon(\omega)-1)$$

- Spectrum: $F_\varepsilon(\omega) = \sum_m R_\varepsilon(m)\exp(-\frac{2\pi\imath}{W}m\omega)$

$$||\Delta|_{\Omega\times\Omega}|| \approx \sup_{\omega\in\Omega}|F_\varepsilon(\omega)-1|$$

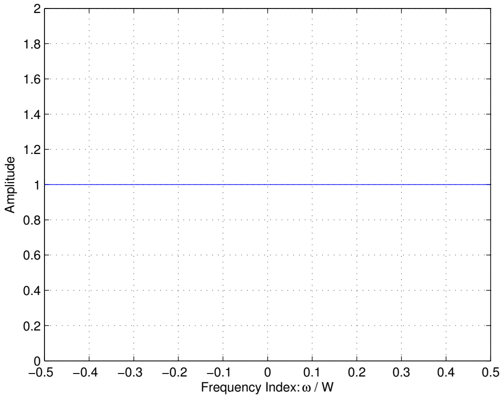
Random Demodulator – Addressable Spectrum

- Flat spectrum provides uniform illumination for all sparse signals:

$$|F_\varepsilon(\omega) - 1| = 0 \ \forall \omega$$

$$\Rightarrow \Delta = 0$$

$F_\varepsilon(\omega)$



Repetition Code – Addressable Spectrum

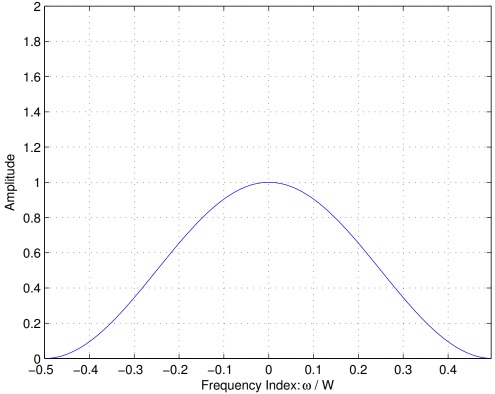
Rate-$\frac{1}{2}$ code: Rademacher sequence with repeated entries

$F_\varepsilon(\omega)$

- Spectrum rolls off to zero at high frequency:

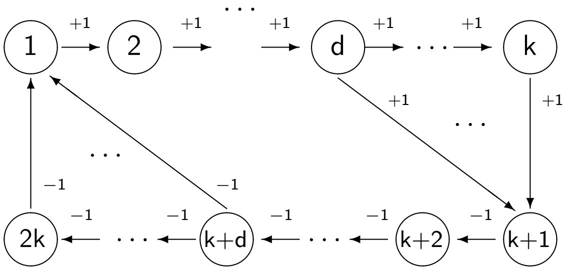
  $$|F_\varepsilon(W/2) - 1| = 1$$

- High frequency signals not well illuminated



General RLL sequences – RLL-Coded Random Data

- Generated from a Markov chain



- Autocorrelation – transition matrix and output symbols

$$R_\varepsilon(m) = \mathrm{a}^T \mathrm{P}^m \mathrm{b}$$

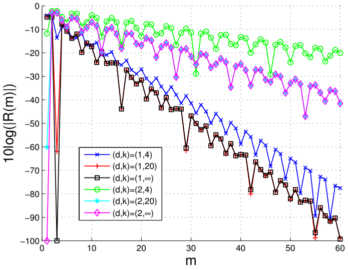
- Geometric decay – rate of second largest eigenvalue of P

$$|\mathrm{a}^T \mathrm{P}^m \mathrm{b}| \leq \lambda^m$$

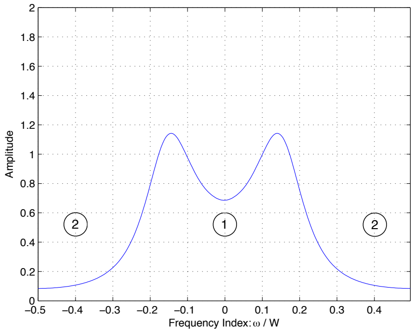
General RLL Sequences – Addressable Spectrum

Geometric decay of correlation

$$|R_\varepsilon(m)| \le \lambda^m$$

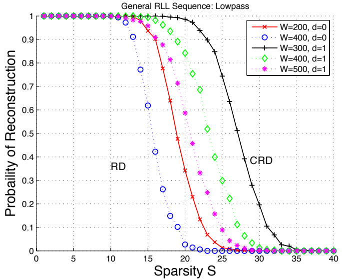


Low frequency signals are better illuminated

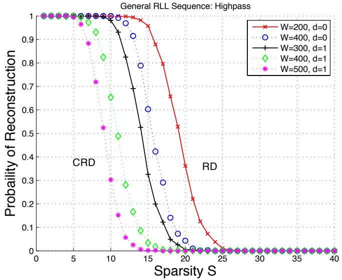


General RLL Sequences – Reconstruction

1: Reconstruct low frequency signals better than RD



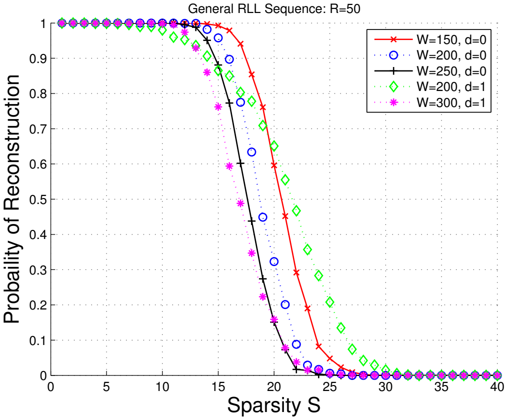
2: Reconstruct high frequency signals worse than RD



Modulating sequence chosen to illuminate different regions of the input spectrum

Bandwidth Advantage

- Reconstruct unrestricted sparse signals
  - Compare performance to the RD (d=0)
  - Increase bandwidth with a small sacrifice in sparsity
  - (d=0,W=250) vs (d=1,W=300): 13% sparsity penalty for a 20% bandwidth gain



RIP for the Constrained Random Demodulator

Theorem (RIPCoRD)

Let Φ_{CRD} be a $R \times W$ CRD matrix using a modulating sequence with maximum dependence distance ℓ and $|||\Delta||| < \delta$ for a fixed $\delta \in (0, 1)$. Suppose that R satisfies

$$R \geq \ell^3 (\delta - |||\Delta|||)^{-2} C \cdot S \log^6(W)$$

where C is a fixed constant. Then with probability $O(1 - W^{-1})$ the CRD matrix Φ_{CRD} satisfies RIP(S,δ_S) with $\delta_S \leq \delta$.

Note: $|||A||| = \sup_{|\Omega| \leq S} \|A|_{\Omega \times \Omega}\|$

Conclusions

- Bandwidth increase given a fundamental limit on waveform fidelity due to the charging/discharging rate of a capacitor
- Tunability – Based on the spectrum, choose the demodulating sequence to illuminate different input signals