Defying Nyquist in Analog to Digital Conversion

Yonina Eldar

Department of Electrical Engineering
Technion – Israel Institute of Technology
Visiting Professor at Stanford, USA

http://www.ee.technion.ac.il/people/YoninaEldar  yonina@ee.technion.ac.il

In collaboration with my students at the Technion
Digital Revolution

“The change from analog mechanical and electronic technology to digital technology that has taken place since c. 1980 and continues to the present day.”

- Cell phone subscribers: 4 billion (67% of world population)
- Digital cameras: 77% of American households now own at least one
- Internet users: 1.8 billion (26.6% of world population)
- PC users: 1 billion (16.8% of world population)
Sampling: “Analog Girl in a Digital World…” Judy Gorman 99

Analog world

- Music
- Radar
- Image...

Digital world

- Signal processing
- Image denoising
- Analysis...

Very high sampling rates: hardware excessive solutions

High DSP rates (Interpolation)

Sampling A2D

\[ x(t) \rightarrow \tilde{x}(t) \rightarrow c[n] \]

ADCs, the front end of every digital application, remain a major bottleneck
Today’s Paradigm

- Analog designers and circuit experts design samplers at Nyquist rate or higher
- DSP/machine learning experts process the data
- Typical first step: Throw away (or combine in a “smart” way) much of the data …
- Logic: Exploit structure prevalent in most applications to reduce DSP processing rates

Can we use the structure to reduce sampling rate + first DSP rate (data transfer, bus …) as well?
Key Idea

Exploit structure to improve data processing performance:

- Reduce storage/reduce sampling rates
- Reduce processing rates
- Reduce power consumption
- Increase resolution
- Improve denoising/deblurring capabilities
- Improved classification/source separation

Goal:

- Survey the main principles involved in exploiting “sparse” structure
- Provide a variety of different applications and benefits
Talk Outline

- Motivation
- Classes of structured analog signals
- Xampling: Compression + sampling
- Sub-Nyquist solutions
  - Multiband communication: Cognitive radio
  - Time delay estimation: Ultrasound, radar, multipath medium identification
- Applications to digital processing
Shannon-Nyquist Sampling

Signal Model

Bandlimited signals

$-f_{\text{max}} \quad f_{\text{max}}$

Minimal Rate

Shannon-Nyquist: $f_{\text{sample}} \geq 2f_{\text{max}}$

Analog + Digital Implementation

ADC

$x(t)$

$y[n]$

Digital Signal Processor

d[n]

sinc$(\cdot)$ Interpolation

DAC

$\hat{x}(t)$

Previous work extends theory to arbitrary subspaces

Many beautiful results, and many contributors (Unser, Aldroubi, Vaidyanathan, Blu, Jerri, Vetterli, Grochenig, DeVore, Daubechies, Christensen, Eldar, Dvorkind …)

Recovery from nonlinear samples as well (Dvorkind, Matusiak and Eldar 2008)

More information:

Structured Analog Models

Multiband communication:

- Can be viewed as $f_{\text{max}}$–bandlimited (subspace)
- But sampling at rate $\geq 2f_{\text{max}}$ is a waste of resources
- For wideband applications Nyquist sampling may be infeasible

Question:
How do we treat structured (non-subspace) models efficiently?
Cognitive Radio

- Cognitive radio mobiles utilize unused spectrum ``holes’’
- Spectral map is unknown a-priori, leading to a multiband model

Federal Communications Commission (FCC) frequency allocation
Structured Analog Models

Medium identification:

\[ g(t) \]

1

0

t

Channel

\[ a_1 \]
\[ a_2 \]
\[ a_3 \]

\[ t_1 \]
\[ t_2 \]
\[ t_3 \]

t

Similar problem arises in radar, UWB communications, timing recovery problems …

- Digital match filter or super-resolution ideas (MUSIC etc.) (Brukstein, Kailath, Jouradin, Saarnisaari …)
- But requires sampling at the Nyquist rate of \( g(t) \)
- The pulse shape is known – No need to waste sampling resources!

Unknown delays – non-subspace

Question (same):
How do we treat structured (non-subspace) models efficiently?
Ultrasound

- High digital processing rates
- Large power consumption

(Collaboration with General Electric Israel)

Echoes result from scattering in the tissue
The image is formed by identifying the scatterers
To increase SNR the reflections are viewed by an antenna array.

SNR is improved through beamforming by introducing appropriate time shifts to the received signals.

- Requires high sampling rates and large data processing rates.
- One image trace requires 128 samplers @ 20M, beamforming to 150 points, a total of $6.3 \times 10^6$ sums/frame.
Resolution (1): Radar

- **Principle:**
  - A known pulse is transmitted
  - Reflections from targets are received
  - Target’s ranges and velocities are identified

- **Challenges:**
  - Targets can lie on an arbitrary grid
  - Process of digitizing
    - Loss of resolution in range-velocity domain
  - Wideband radar requires high rate sampling and processing which also results in long processing time
Resolution (2): Subwavelength Imaging

(Collaboration with the groups of Segev and Cohen)

Diffraction limit: Even a perfect optical imaging system has a resolution limit determined by the wavelength $\lambda$

- The smallest observable detail is larger than $\sim \lambda/2$
- This results in image smearing

Nano-holes as seen in electronic microscope

Sketch of an optical microscope: the physics of EM waves acts as an ideal low-pass filter

Blurred image seen in optical microscope
Imaging via “Sparse” Modeling

- Radar:

- Subwavelength Coherent Diffractive Imaging:

Recovery of sub-wavelength images from highly truncated Fourier power spectrum

Bajwa et al., '11

Szameit et al., Nature Photonics, ‘12
Proposed Framework

- Instead of a single subspace modeling use union of subspaces framework

- Adopt a new design methodology – Xampling
  - Compression+Sampling = Xampling
  - X prefix for compression, e.g. DivX

- Results in simple hardware and low computational cost on the DSP

Union + Xampling = Practical Low Rate Sampling
Talk Outline

- Motivation
- Classes of structured analog signals
- Xampling: Compression + sampling
- Sub-Nyquist solutions
  - Multiband communication: Cognitive radio
  - Time delay estimation: Ultrasound, radar, multipath medium identification
- Applications to digital processing
Union of Subspaces

Model: $\mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda$

$x(t) \in \mathcal{A}_{\lambda^*} \rightarrow \lambda^* \text{ is unknown a-priori}$
Each $\mathcal{A}_\lambda$ has low dimension

Examples:

Multiband communication

Union over possible band positions $f_i \in [0, f_{\text{max}}]$
Union of Subspaces

Model: \[ \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda \]

\[ x(t) \in \mathcal{A}_{\lambda^*} \rightarrow \lambda^* \text{ is unknown a-priori} \]

Each \( \mathcal{A}_\lambda \) has low dimension

Standard approach: Look at \textit{sum} of all subspaces

\[ \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda \]

\[ \mathcal{U} = \bigoplus_{\lambda \in \Lambda} \mathcal{A}_\lambda \]

Signal bandlimited to \( f_{\text{max}} \) → \textbf{High rate}
Union of Subspaces

Model: \[ \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda \]

\( x(t) \in \mathcal{A}_{\lambda^*} \rightarrow \lambda^* \) is unknown a-priori
Each \( \mathcal{A}_\lambda \) has low dimension

Examples:

Time-delay estimation

Union over possible path delays \( t_i \in [0, \tau] \)
Union of Subspaces

Model: \[ U = \bigcup_{\lambda \in \Lambda} A_\lambda \]

\[ x(t) \in A_{\lambda^*} \rightarrow \lambda^* \text{ is unknown a-priori} \]

Each \( A_\lambda \) has low dimension

- Allows to keep low dimension in the problem model
- Low dimension translates to low sampling rate

(Lu and Do 08, Eldar and Mishali 09)
Motivation

Classes of structured analog signals

**Xampling: Compression + sampling**

Sub-Nyquist solutions
- Multiband communication: Cognitive radio
- Time delay estimation: Ultrasound, radar, multipath medium identification

Applications to digital processing
Naïve attempt: direct sampling at low rate
Most samples do not contain information!!

Most bands do not have energy – which band should be sampled?

\[ B \cdot f_{NYQ} \sim 10 \text{’s GHz} \]
Intuitive Solution: Pre-Processing

- Smear pulse before sampling
- Each samples contains energy
- Resolve ambiguity in the digital domain

- Alias all energy to baseband
- Can sample at low rate
- Resolve ambiguity in the digital domain
Xampling: Main Idea

- Create several streams of data
- Each stream is sampled at a low rate
  (overall rate much smaller than the Nyquist rate)
- Each stream contains a combination from different subspaces

Hardware design ideas

- Identify subspaces involved
- Recover using standard sampling results

DSP algorithms
For linear methods:
- Subspace techniques developed in the context of array processing (such as MUSIC, ESPRIT etc.)
- Compressed sensing
(Deborah and Noam’s talks this afternoon)

For nonlinear sampling:
- Specialized iterative algorithms (Tomer’s talk this afternoon)
Compressed Sensing

\[ \mathbf{y} = \mathbf{\Phi} \mathbf{x} \]

Short
\[ \approx 2K \text{ meas.} \]

Main ideas:
- Sparse input vector with unknown support
- Sensing by sufficiently incoherent matrix (semi-random)
- Polynomial-time recovery algorithms

\[ m \times n, \quad m \ll n \]

Long
\[ K\text{-sparse} \]

(Candès, Romberg, Tao 2006)
(Donoho 2006)
Compressed Sensing

\[ y = \Phi \times x \]

Short
\[ \approx 2K \text{ meas.} \]

Xampling:
- Sparsity of \( x \) represents that only a few subspaces participate
- The matrix \( \Phi \) represents the aliasing of the hardware
- Support detection is equivalent to subspace detection
Compressed Sensing and Hardware

- Explosion of work on compressed sensing in many digital applications
- Many papers describing models for CS of analog signals
- None of these models have made it into hardware
- CS is a digital theory – treats vectors not analog inputs

<table>
<thead>
<tr>
<th>Input</th>
<th>Standard CS</th>
<th>Analog CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sparsity</td>
<td>vector $x$</td>
<td>analog signal $x(t)$</td>
</tr>
<tr>
<td>Measurement</td>
<td>few nonzero values</td>
<td>?</td>
</tr>
<tr>
<td>Recovery</td>
<td>random matrix</td>
<td>real hardware</td>
</tr>
<tr>
<td></td>
<td>convex optimization</td>
<td>need to recover analog input</td>
</tr>
<tr>
<td></td>
<td>greedy methods</td>
<td></td>
</tr>
</tbody>
</table>

We use CS only after sampling and only to detect the subspace. Enables real hardware and low processing rates.
Xampling Hardware

- $p_i(t)$ - periodic functions
- $p_i(t) = \sum a_i n e^{-j \frac{2\pi}{T_p} n t}$ sums of exponentials
- The filter $H(f)$ allows for additional freedom in shaping the tones
- The channels can be collapsed to a single channel
Talk Outline

- Motivation
- Classes of structured analog signals
  - Xampling: Compression + sampling
  - Sub-Nyquist solutions
    - Multiband communication
      - Time delay estimation: Ultrasound, radar, multipath medium identification
  - Applications to digital processing
Signal Model

(Mishali and Eldar, 2007-2009)

1. Each band has an uncountable number of non-zero elements
2. Band locations lie on the continuum
3. Band locations are unknown in advance

\[ \mathcal{M} = \{ x(t) \mid \text{no more than } N \text{ bands, max width } B, \text{ bandlimited to } [-\frac{1}{2}f_{NYQ}, +\frac{1}{2}f_{NYQ}) \} \]
Rate Requirement

Theorem (Single multiband subspace)

Let $R$ be a sampling set for $\mathcal{B}_\mathcal{F} = \{ x(t) \in L^2(\mathbb{R}) \mid \text{supp } X(f) \subseteq \mathcal{F} \}$. Then,

$$D^-(R) \geq \lambda = |\mathcal{F}|$$

(Landau 1967)

Average sampling rate

Theorem (Union of multiband subspaces)

Let $R$ be a sampling set for $\mathcal{N}_\lambda = \bigcup_{|\mathcal{F}| \leq \lambda} \mathcal{B}_\mathcal{F}$. Then,

$$D^-(R) \geq \min \{ 2\lambda, f_{NYQ} \}$$

(Mishali and Eldar 2007)

1. The minimal rate is doubled.
2. For $x(t) \in \mathcal{M}$, the rate requirement is $2NB$ samples/sec (on average).
The Modulated Wideband Converter

$T_p$-periodic $p_i(t)$ gives the desired aliasing effect and many more...
Recovery From Xamples

Using ideas of compressed sensing

Modifications to allow for real time computations and noise robustness

Cleverly combine data across samples to improve support detection

Details in Deborah’s talk this afternoon
A 2.4 GHz Prototype

(Mishali, Eldar, Dounaevsky, and Shoshan, 2010)

- 2.3 GHz Nyquist-rate, 120 MHz occupancy
- 280 MHz sampling rate
- Wideband receiver mode:
  - 49 dB dynamic range
  - SNDR > 30 dB over all input range
- ADC mode:
  - 1.2 volt peak-to-peak full-scale
  - 42 dB SNDR = 6.7 ENOB
- Off-the-shelf devices, ~5k$, standard PCB production
Sub-Nyquist Demonstration

Carrier frequencies are chosen to create overlayed aliasing at baseband

FM @ 631.2 MHz  +  AM @ 807.8 MHz  +  Sine @ 981.9 MHz  →  MWC prototype  →  aliasing around 6.171 MHz

Reconstruction (CTF)

FM @ 631.2 MHz  10 kHz  1.5 MHz  AM @ 807.8 MHz  100 kHz

Mishali et al., 10
Online Demonstrations

- GUI package of the MWC

![GUI package of the MWC]

- Video recording of sub-Nyquist sampling + carrier recovery in lab

![Video recording of sub-Nyquist sampling + carrier recovery in lab]
Demos – Supported By NI

Demo this afternoon by Rolf and Idan
Talk Outline

- Brief overview of standard sampling
- Classes of structured analog signals
- Xampling: Compression + sampling
- Sub-Nyquist solutions
  - Multiband communication
  - Time delay estimation: Ultrasound, radar, multipath medium identification
- Applications to digital processing
Streams of Pulses

- $H(f)$ is replaced by an integrator
- Can equivalently be implemented as a single channel with $T = T_p/m$
- Application to radar, ultrasound and general localization problems such as GPS

$$s(t) = \sum_n b_n e^{j \frac{2\pi}{T_p} n t} \text{rect}(t/T_p)$$
Output corresponds to aliased version of Gabor coefficients

Recovery by solving 2-step CS problem $Y = AB^T$

1. Solve $Y = AC$ with $C = ZB^T$ ⇒ Since $Z$ is row-sparse $C$ is row-sparse
2. Solve CS problem $C^T = BZ$ where $Z$ is row sparse

(Matusiak and Eldar, 11)
Noise Robustness

- MSE of the delays estimation, versus integrators approach (Kusuma & Goyal)

\[ L = 2 \text{ pulses, 5 samples} \]

\[ L = 10 \text{ pulses, 21 samples} \]

The proposed scheme is stable even for high rates of innovation!
Application: Multipath Medium Identification

Medium identification (collaboration with National Instruments):
- Recovery of the time delays
- Recovery of time-variant gain coefficients

The proposed method can recover the channel parameters from sub-Nyquist samples

\[ x_T(t) = \sum_{n \in \mathbb{Z}} h(t - nT) \]

\[ x_R(t) = \sum_{l \in \mathbb{Z}} a_l h(t - t_l) \]

LTV channel \( L \) propagation paths

(Gedalyahu and Eldar 09-10)
New paradigm for wireless communications: Joint effort with Prof. Andrea Goldsmith from Stanford (Transformative Science Grant)

- Main bottleneck today in wireless are ADCs
- Multiuser detection, which enables many users to share joint resources, is not implemented because of high rates – channels are interference limited
- SDR and Cognitive radio are limited by ADCs
- Capacity tools are limited to Nyquist-rate channels

- New multiuser receiver that substantially reduces hardware requirements
- Capacity expressions for sampling-rate limited channels
- Applications to LTE standards (with Prof. Murmann and Ericsson)
Application: Radar

- Each target is defined by:
  - Range – delay
  - Velocity – doppler

- Targets can be identified with **infinite resolution** as long as the time-bandwidth product satisfies $TW \geq 2\pi(K + 1)^2$

- Previous results required infinite time-bandwidth product

(Bajwa, Gedalyahu and Eldar, 10)
A scheme which enables reconstruction of a two dimensional image, from samples obtained at a rate 10-15 times below Nyquist.

The resulting image depicts strong perturbations in the tissue.

Obtained by beamforming in the compressed domain.

More details in Noam’s talk.

Xampling in Ultrasound Imaging

Wagner, Eldar, and Friedman, ’11

- Standard Imaging: 1662 real-valued samples, per sensor per image line.
- Xampled beamforming: 200 real-valued samples, per sensor per image line (assume L=25 reflectors per line).
Structure in Digital Problems

- Union of subspaces structure can be exploited in many digital models
- Subspace models lead to block sparse recovery
- Block sparsity: algorithms and recovery guarantees in noisy environments (Eldar and Mishali 09, Eldar et. al. 10, Ben-Haim and Eldar 10)
- Hierarchical models with structure on the subspace level and within the subspaces (Sprechmann, Ramirez, Sapiro and Eldar, 10)
Source Separation Cont.

**Texture Separation:**

<table>
<thead>
<tr>
<th>2.80 0.42</th>
<th>2.06 0.00</th>
<th>0.33 0.25</th>
<th>0.96 0.01</th>
<th>1.02 1.00</th>
<th>2.26 0.32</th>
<th>4.37 1.39</th>
<th>0.09 0.98</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.36 0.00</td>
<td>3.69 0.00</td>
<td>2.67 0.02</td>
<td>3.69 0.00</td>
<td>2.30 0.00</td>
<td>3.69 0.07</td>
<td>0.97 1.97</td>
<td>2.25 0.09</td>
</tr>
<tr>
<td>110 214</td>
<td>107 182</td>
<td>107 182</td>
<td>107 182</td>
<td>107 182</td>
<td>107 182</td>
<td>107 182</td>
<td>107 182</td>
</tr>
<tr>
<td>117 69</td>
<td>76 141</td>
<td>68 120</td>
<td>81 120</td>
<td>86 120</td>
<td>81 120</td>
<td>86 120</td>
<td>86 120</td>
</tr>
<tr>
<td>2.67 0.02</td>
<td>1.74 0.00</td>
<td>2.42 0.00</td>
<td>3.39 0.16</td>
<td>2.85 0.35</td>
<td>2.85 0.35</td>
<td>2.85 0.35</td>
<td>2.85 0.35</td>
</tr>
<tr>
<td>17 18</td>
<td>69 57</td>
<td>59 57</td>
<td>69 57</td>
<td>69 57</td>
<td>69 57</td>
<td>69 57</td>
<td>69 57</td>
</tr>
<tr>
<td>18 069</td>
<td>102 102</td>
<td>102 102</td>
<td>102 102</td>
<td>102 102</td>
<td>102 102</td>
<td>102 102</td>
<td>102 102</td>
</tr>
<tr>
<td>85 174</td>
<td>132 132</td>
<td>83 183</td>
<td>83 183</td>
<td>145 145</td>
<td>128 128</td>
<td>128 128</td>
<td>128 128</td>
</tr>
<tr>
<td>102 102</td>
<td>141 141</td>
<td>81 120</td>
<td>81 120</td>
<td>148 148</td>
<td>148 148</td>
<td>148 148</td>
<td>148 148</td>
</tr>
<tr>
<td>245 162</td>
<td>178 178</td>
<td>95 95</td>
<td>95 95</td>
<td>178 178</td>
<td>178 178</td>
<td>178 178</td>
<td>178 178</td>
</tr>
<tr>
<td>7 43</td>
<td>27 3</td>
<td>68 105</td>
<td>95 19</td>
<td>21 93</td>
<td>102 10</td>
<td>26 89</td>
<td>85 10</td>
</tr>
<tr>
<td>102 10</td>
<td>107 107</td>
<td>124 124</td>
<td>107 107</td>
<td>124 124</td>
<td>107 107</td>
<td>124 124</td>
<td>107 107</td>
</tr>
<tr>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
<td>3.31 1.00</td>
</tr>
<tr>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
<td>3.04 0.24</td>
</tr>
<tr>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
<td>1.90 0.18</td>
</tr>
</tbody>
</table>
Prior work has focused primarily on learning a single subspace (Vidal et. al., Ma et. al., Elhamifar …)

We developed methods for multiple subspace learning from training data (Rosenblum, Zelnik-Manor and Eldar, 10)

Subspace learning from reduced-dimensional measured data: Blind compressed sensing (Gleichman and Eldar 10)

Current work: Extending these ideas to more practical scenarios (Carin, Silva, Chen, Sapiro)
Interpolation by learning the basis from the corrupted image
Conclusions

- Compressed sampling and processing of many signals
- Wideband sub-Nyquist samplers in hardware
- Union of subspaces: broad and flexible model
- Practical and efficient hardware
- Many applications and many research opportunities: extensions to other analog and digital problems, robustness, hardware ...

Exploiting structure can lead to a new sampling paradigm which combines analog + digital

More details in:
Thank you