Department of Electrical Engineering
Electronics
Computers
Communications

Xampling

From Theory to Hardware of Sub-Nyquist Sampling

Moshe Mishali **Yonina Eldar**

Department of Electrical Engineering
Technion – Israel Institute of Technology

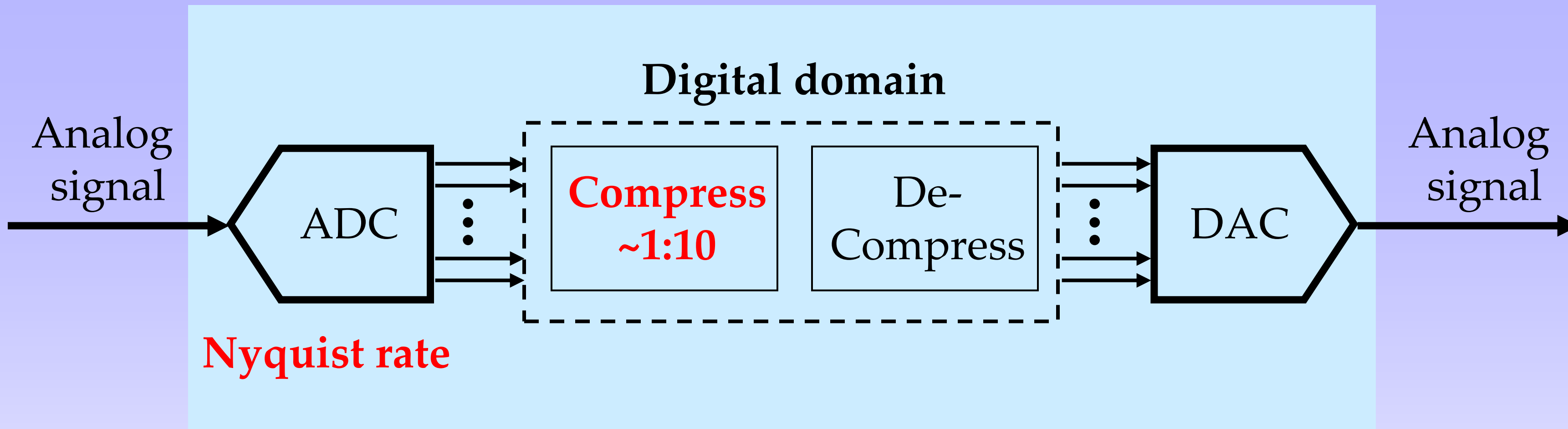
http://www.technion.ac.il/~moshiko
http://www.ee.technion.ac.il/people/YoninaEldar/

moshiko@tx.technion.ac.il
yonina@ee.technion.ac.il

ICASSP Tutorial
May 23rd, 2011

"Analog Girl in a Digital World…"

Judy Gorman `99



Voice recorder

Camera

Medical imaging

"Can we not just directly measure the part that will not end up being thrown away ?"

Donoho, '06

Key Idea

Exploit structure to improve data processing performance:

- Reduce storage/reduce sampling rates
- Reduce processing rates
- Increase imaging resolution
- Reduce power, size, cost…

Goal:

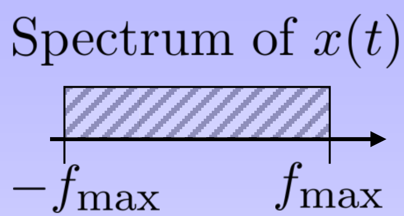
- Survey sampling strategies that exploit signal structure to reduce rate
- Present a unified framework for sub-Nyquist sampling
- Provide a variety of different applications and benefits

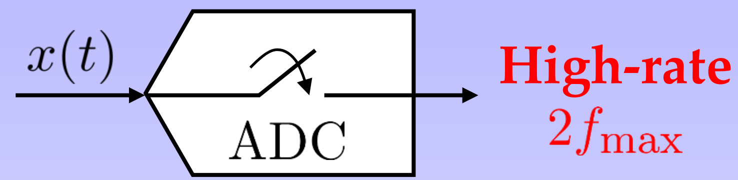
Outline

- Part 1: Introduction
- Part 2: Sub-Nyquist in a subspace
  - Generalized sampling framework
  - Examples
- Part 3: Union of subspaces
  - Model, analog and discrete applications
  - Short intro to compressed sensing
- Part 4: Xampling, Sub-Nyquist in a union
  - Functional framework
  - Modulated wideband conversion
  - Sparse shift-invariant sampling
  - Finite-rate/sequences of innovation methods
  - Random demodulation
- Part 5: From theory to hardware
  - Practical design metrics
  - Circuit challenges

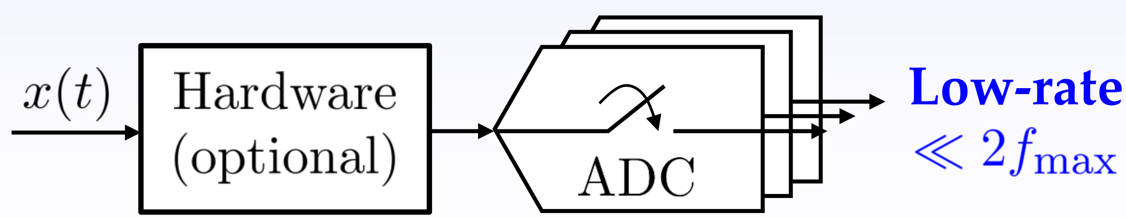
Outline Schematically

- Definition: Nyquist-rate system
  A single ADC device outputs a stream of numbers at rate $2f_{\max}$





- Sub-Nyquist system
  - One or more ADC devices
  - Each ADC device runs at a rate below $2f_{\max}$
  - With / without analog preprocessing
- Overall rate $< 2f_{\max}$ **← Our main focus**



Tutorial Goal

To be as interactive as possible!

- Feel free to ask questions
- Raise ideas
- Slow us down if things are too fast …

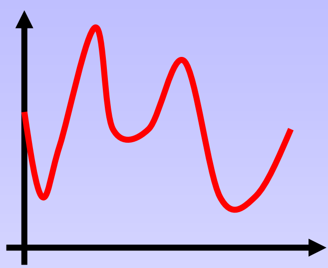
Hope you learn and enjoy!

– Part 1 –
Introduction

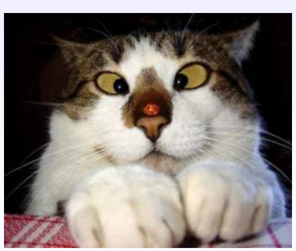
→ Outline

Sampling: "Analog Girl in a Digital World…" Judy Gorman 99

Analog world



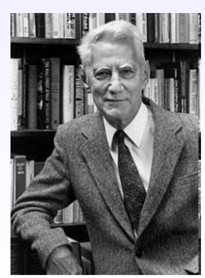
- Music
- Radar
- Image…

$x(t)$ → **Sampling A2D** → $c[n]$

$\tilde{x}(t)$ ← **Reconstruction D2A** ←

(Interpolation)

Digital world

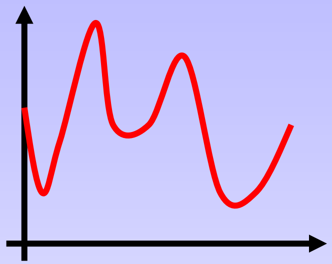
- Signal processing
- Image denoising
- Analysis…

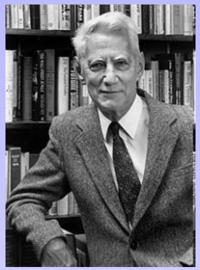


Sampling: "Analog Girl in a Digital World…" Judy Gorman 99

Analog world

Digital world



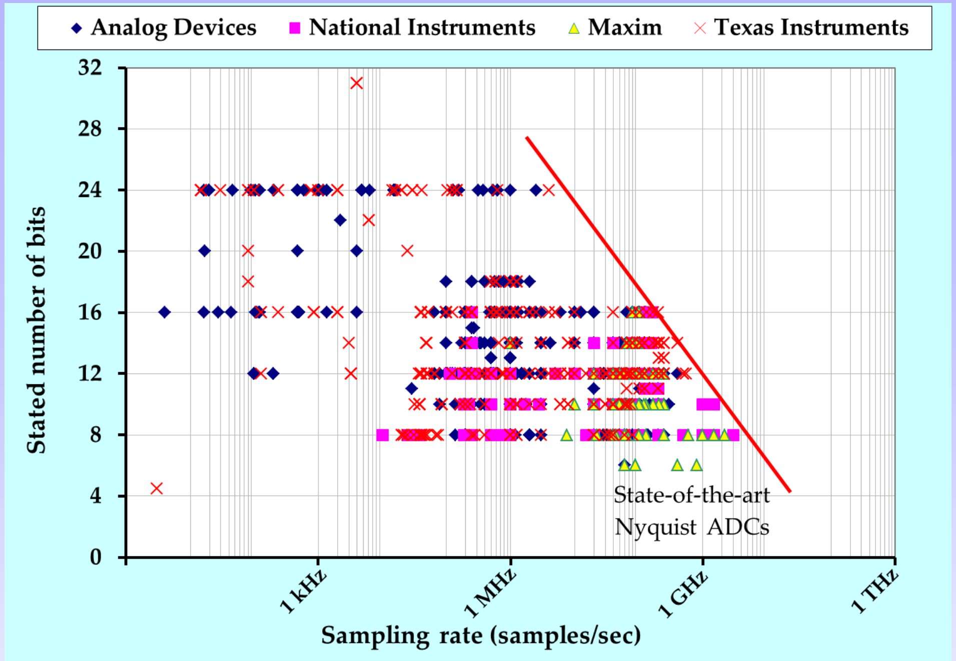


- Music
- Radar
- Image…

- Very high sampling rates: hardware excessive solutions
- High DSP rates

- Signal processing
- Image denoising
- Analysis…

ADC Market



- State-of-the-art ADCs generate uniform samples at the input's Nyquist rate
- Continuous effort to:
    - increase sampling rate (Giga-samples/sec)
    - increase front-end bandwidth
    - increase (effective) number of bits

Working in digital becomes difficult

Nyquist Rate Sampling

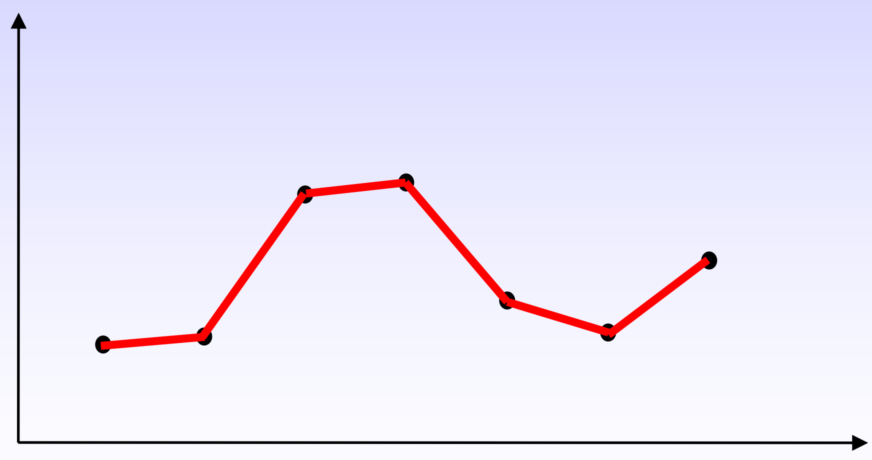
- Standard processing techniques require sampling at the Nyquist rate = twice the highest frequency
- Narrow pulse, wide sensing range = high Nyquist rate
- Results in hardware excessive solutions and high DSP rates
- Too difficult to process, store and transmit

Main Idea:
Exploit structure to reduce sampling and processing rates

The Key – Structure

- Sampling reduces ``dimenions''
- Must have some prior on $x(t)$

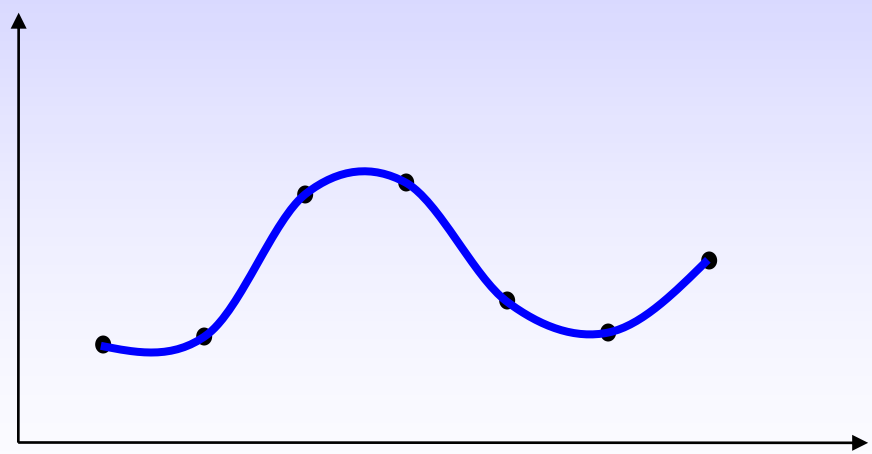
$x(t)$ piece-wise linear



The Key – Structure

- Sampling reduces ``dimenions''
- Must have some prior on $x(t)$

$x(t)$ bandlimited



Prior (= Signal Model) Necessary for Recovery

The Key – Structure

- Sampling reduces ``dimenions''
- Must have some prior on $x(t)$

- Model too narrow (e.g. pure sine) → not widely applicable
- Model too wide (e.g. bandlimited) → no rate reduction

Key: Treat signal models that are sufficiently wide and structured at the same time

Structure Types

- In this tutorial we treat 2 main structures:

Subspace

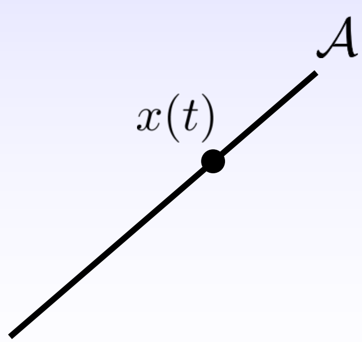
$x(t) \in \mathcal{A}$

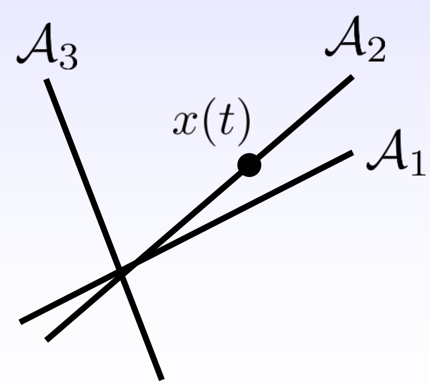
Union of subspaces

$x(t) \in \underbrace{\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots}_{\mathcal{U}}$

- **Linear:** $x, y \in \mathcal{A} \rightarrow \alpha x + \beta y \in \mathcal{A}$
- Generalized sampling theory

- **Nonlinear:** $x + y \notin \mathcal{U}$ (typically)
- **Xampling** (functional framework)





Structure Types

- In this tutorial we treat 2 main structures:

Subspace

$$x(t) \in \mathcal{A}$$

Union of subspaces

$$x(t) \in \underbrace{\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots}_{\mathcal{U}}$$

- **Linear:** $x, y \in \mathcal{A} \rightarrow \alpha x + \beta y \in \mathcal{A}$
- Generalized sampling theory

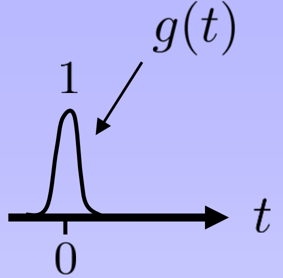
- **Nonlinear:** $x + y \notin \mathcal{U}$ (typically)
- **Xampling** (functional framework)

- Subspace modeling is used in many practical applications
- BUT, can result in unnecessary-high sampling and processing rates
- Union modeling paves the way to innovative sampling methods, at rates as low as the actual information rate

Ultrasound

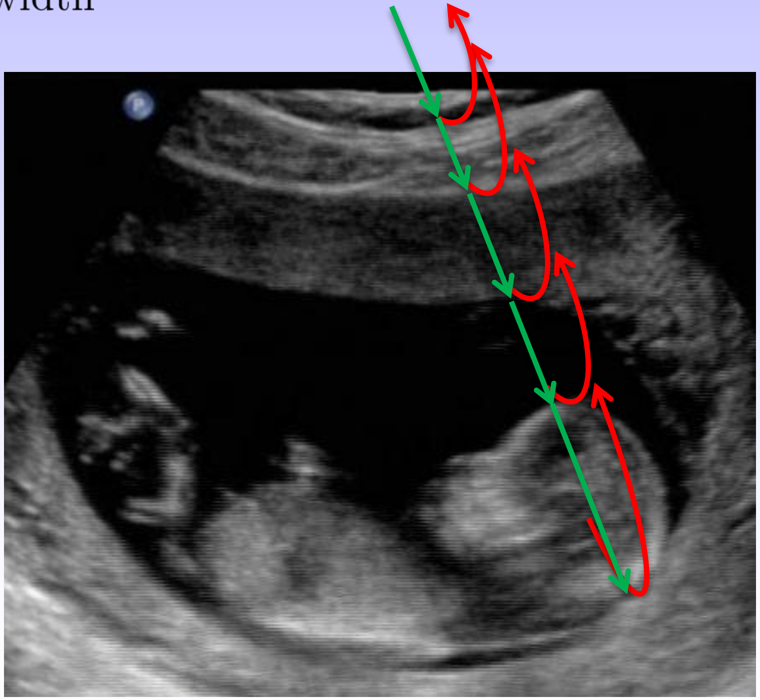
- High sampling rates
- High digital processing rates

Tx pulse



10 MHz bandwidth

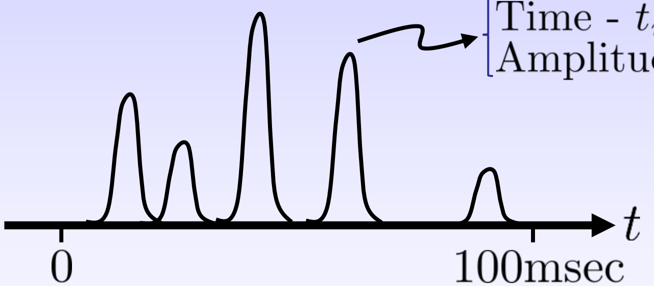
Ultrasonic probe



Rx signal

Unknowns

Time - t_i
Amplitude - a_i

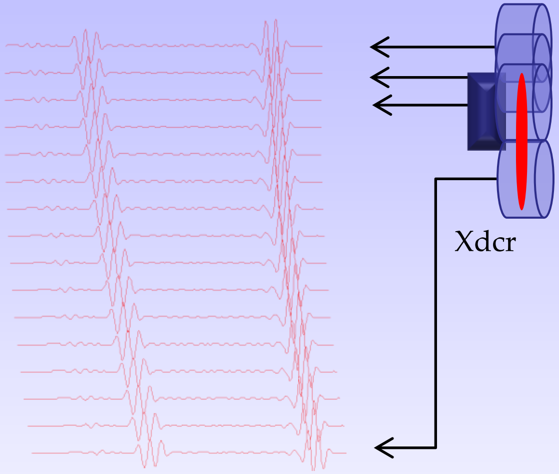


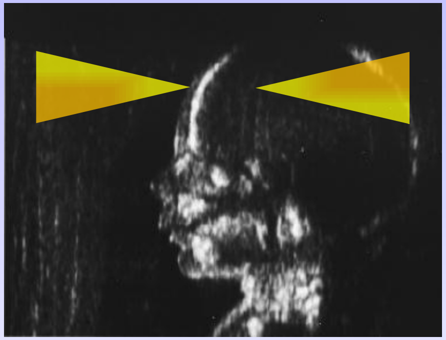
L echoes

- Echoes result from scattering in the tissue
- The image is formed by identifying the scatterers

Processing Rates

- 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



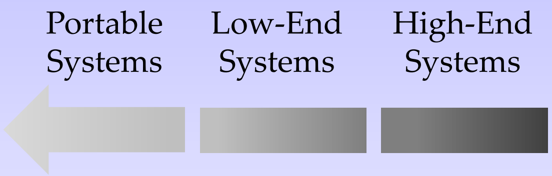


Scan Plane

- Requires high sampling rates and large data processing rates
- **Subspace:** One image trace requires 128 samplers @ 20M, beamforming to 150 points, a total of 6.3×10^6 sums/frame
- **Union:** can reduce sampling rate by orders of magnitude

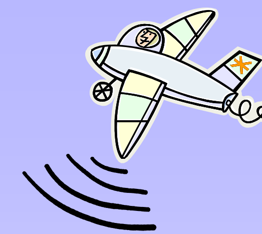
Processing Rates

Goal: reduce ultrasound machines to a size of a laptop at same resolution



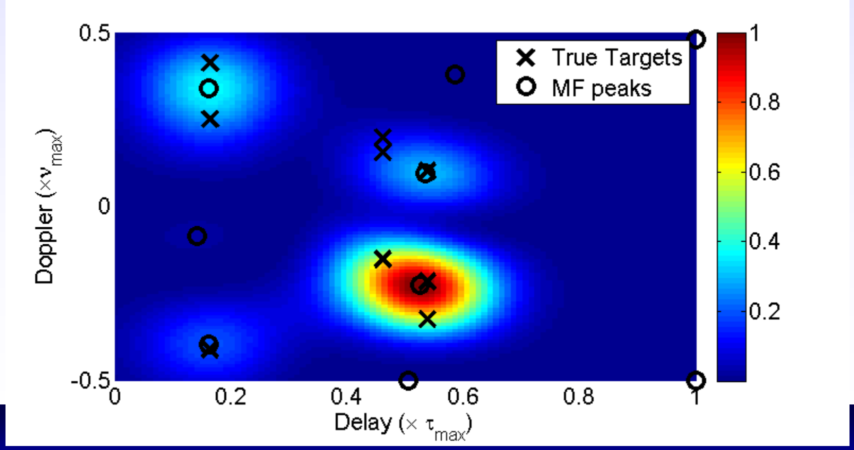
Resolution (1): Radar

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



- Challenge:
  - All processing is done digitally
  - Targets can lie on an arbitrary grid
  - Process of digitizing
    → loss of resolution in range-velocity domain

- **Subspace methods:**



Resolution (2): Subwavelength Imaging

Diffraction limit: Even a perfect optical imaging system has a resolution limit determined by the wavelength λ

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

Input

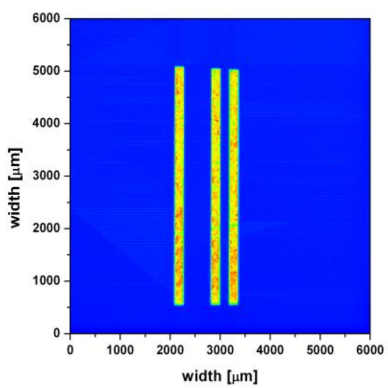
Output

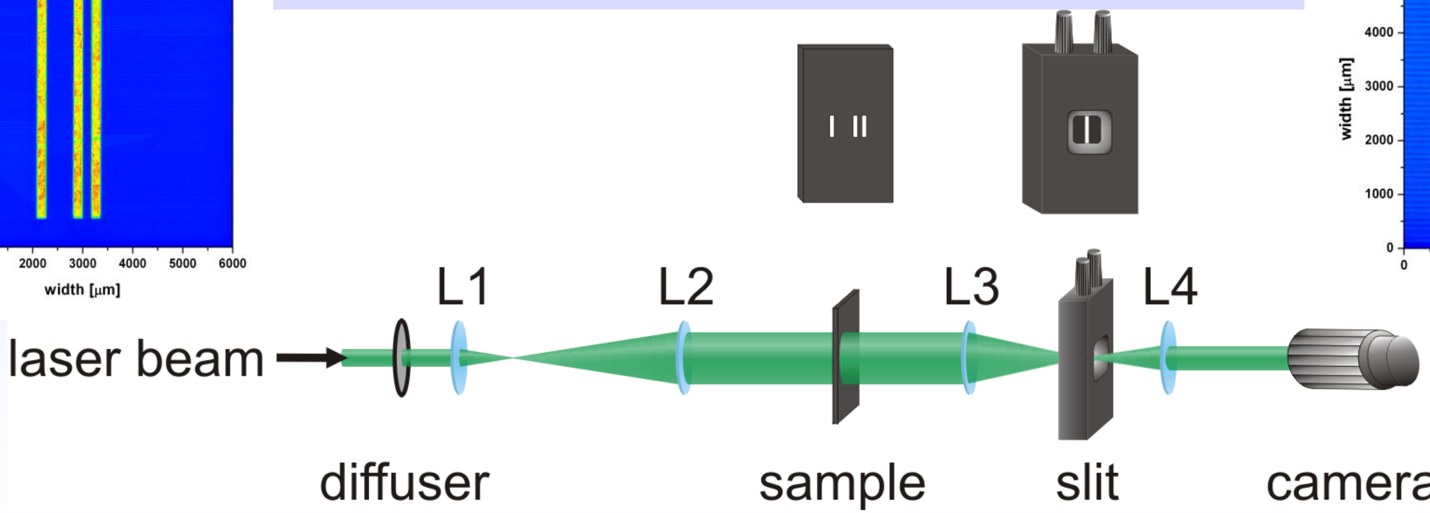
Resolution (2): Subwavelength Imaging

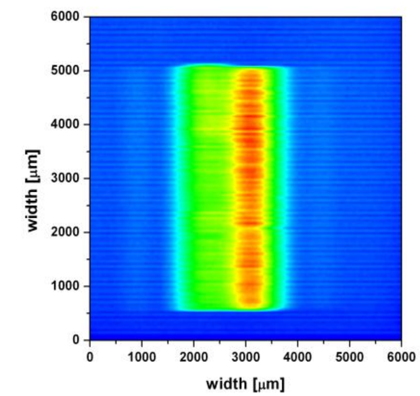
Diffraction limit: Even a perfect optical imaging system has a resolution limit determined by the wavelength λ

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

- **Subspace methods:**

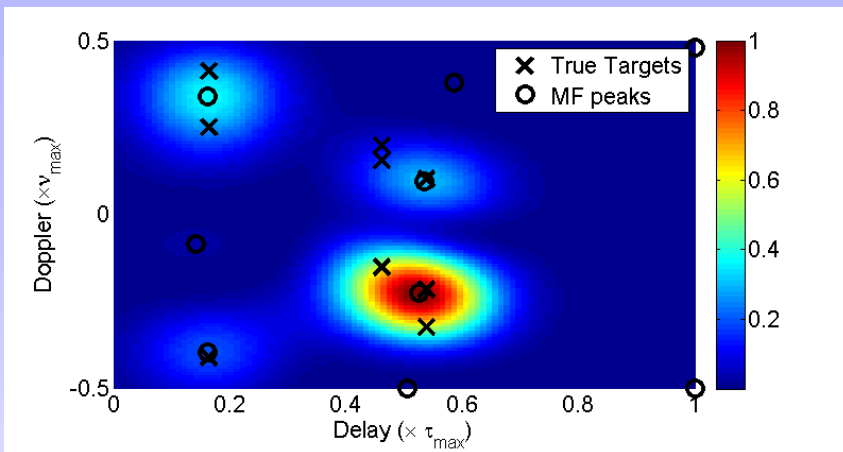


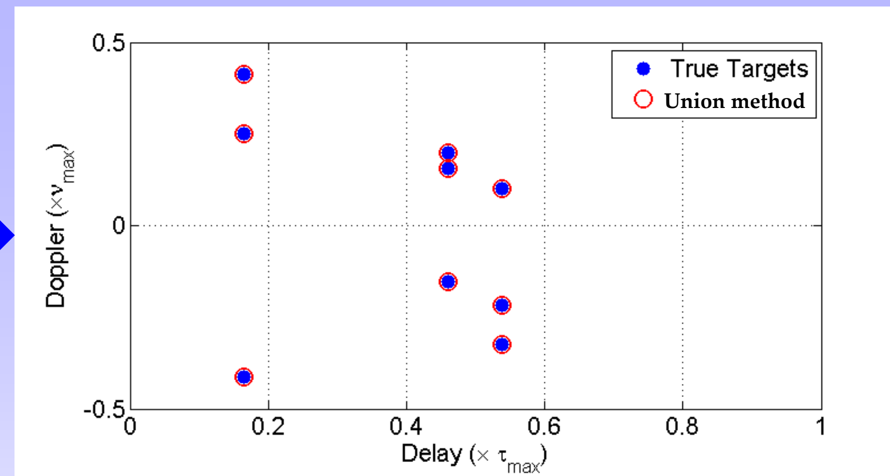




Imaging via Union Modeling

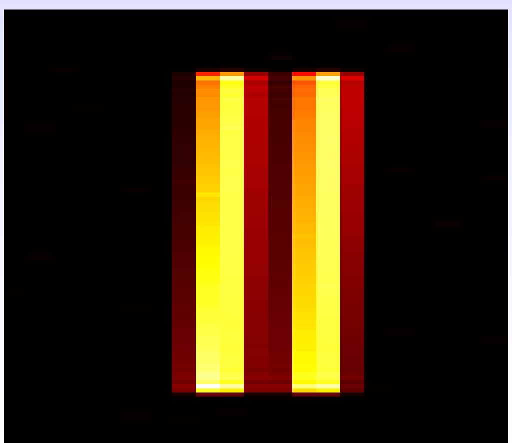
- Radar:

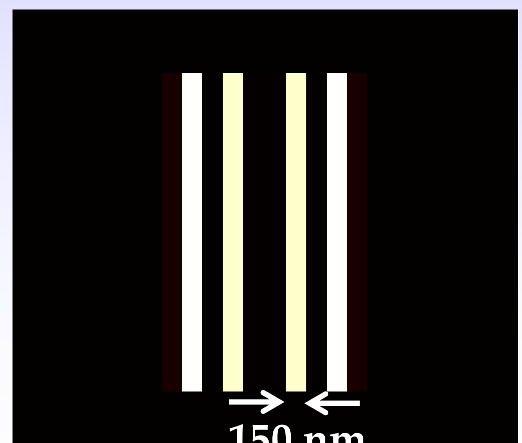




Bajwa *et al.*, '11

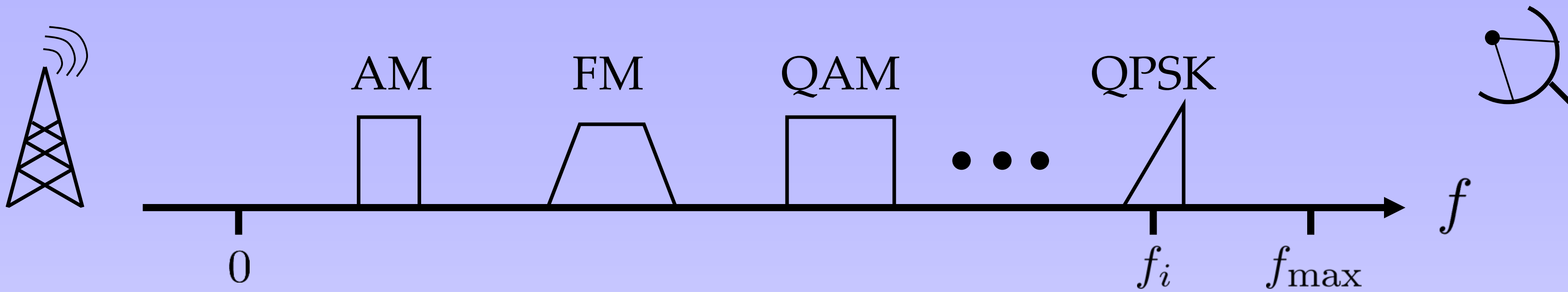
- Subwavelength:





Gazit *et al.*, '11

Wideband Communication



- **Subspace methods:**
  - RF demodulation
  - Undersampling
  - and more…

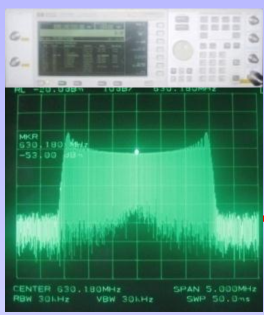
f_i are known

- Unknown f_i, *e.g.* cognitive radio. Should we sample at $2f_{\max}$?

- **Union modeling:**
  - Can sample at the actual information bandwidth, even though f_i are unknown
  - Can process at low rate (no need to reconstruct Nyquist-rate samples)

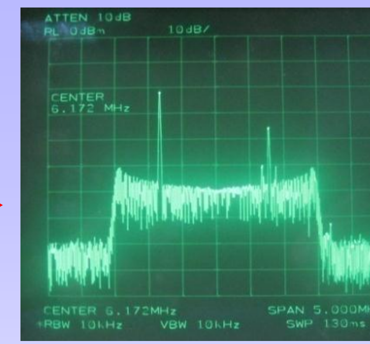
Sub-Nyquist Demonstration

Carrier frequencies are chosen to create overlayed aliasing at baseband



+

+



FM @ 631.2 MHz

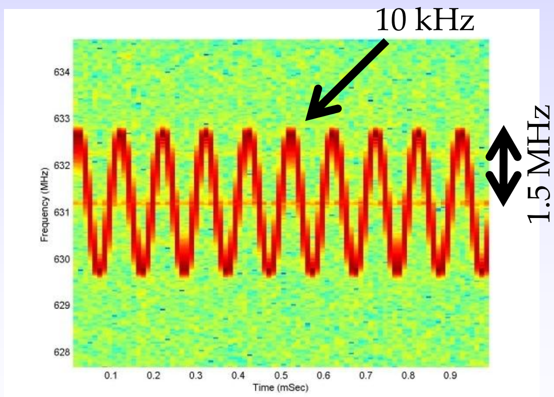
AM @ 807.8 MHz

Sine @ 981.9 MHz

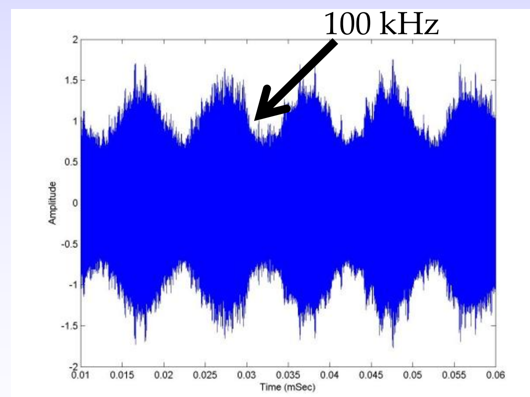
MWC prototype

aliasing around 6.171 MHz

Reconstruction
(CTF)



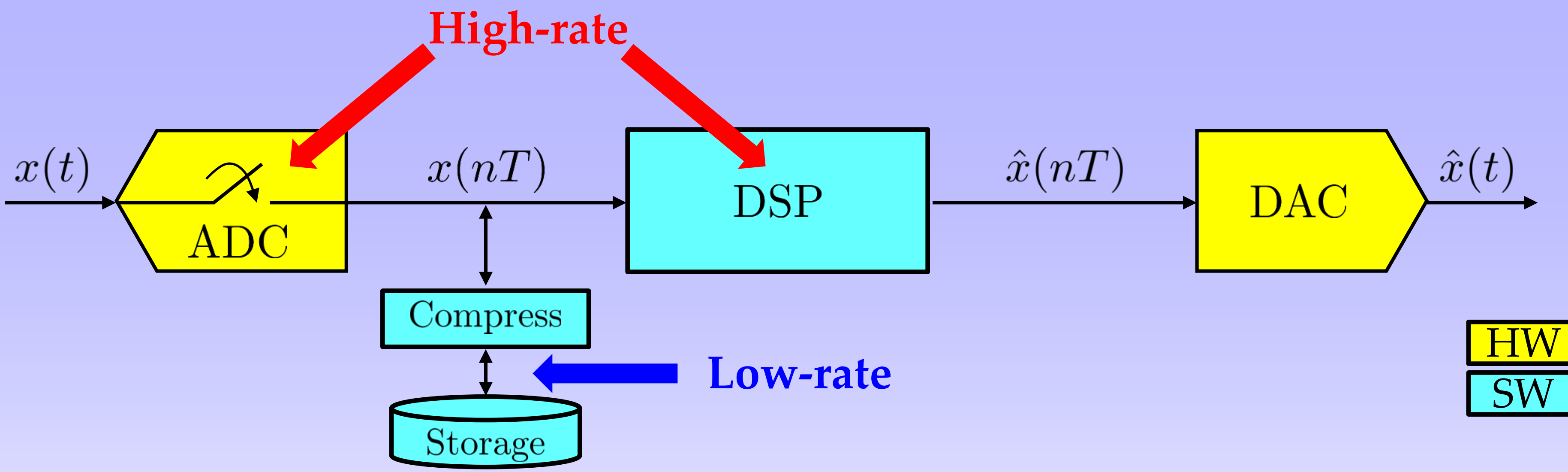
FM @ 631.2 MHz



AM @ 807.8 MHz

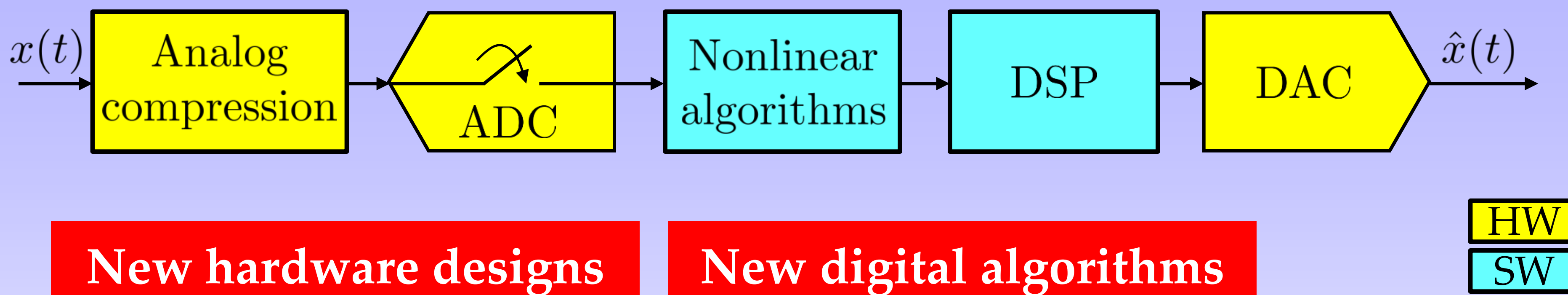
Mishali *et al.*, '10

Xampling



- Main idea:
  - Move compression before ADC
  - Use nonlinear algorithms to interface with standard DSP and signal reconstruction

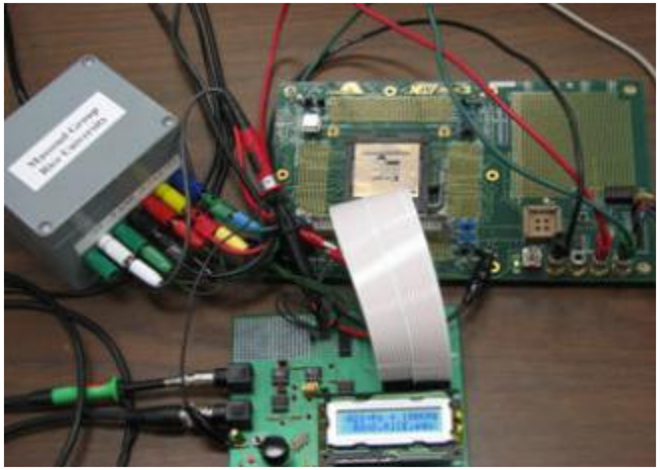
Xampling



New hardware designs **New digital algorithms**

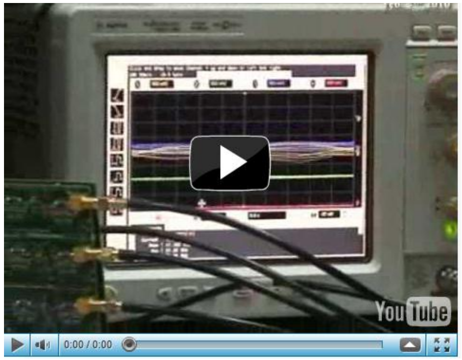
- Main idea:
  - Move compression before ADC
  - Use nonlinear algorithms to interface with standard DSP and signal reconstruction
- Follow a set of design principles → step from theory to hardware

From Theory to Hardware



- 800 kHz Nyquist-rate
- 100 kHz sampling rate

Ragheb *et al.*, 08



- 2.4 GHz Nyquist-rate, 120 MHz occupancy
- 280 MHz sampling rate
- 49 dB dynamic range
- SNDR > 30 dB over input range

Mishali *et al.*, 10

Candès *et al.*, 08 Emami *et al.*, 08 RICE 1-pixel camera DARPA A2I Project

- See many more contributors in compressive sensing hardware
- Tutorial briefly covers circuit challenges in sub-Nyquist systems

Sub-Nyquist technology becomes feasible !

Can gain significant advantages in practical applications

– Part 2 –
Sub-Nyquist in a Subspace

→ Outline

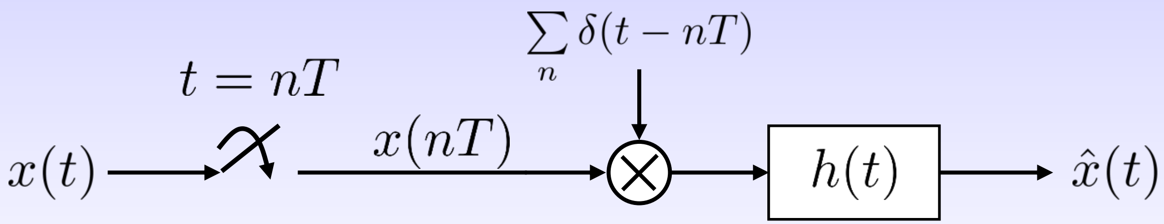
Shannon-Nyquist Sampling

Theorem [Bandlimited Sampling]

If a function $x(t)$ contains no frequencies higher than W cycles-per-second, it is completely determined by giving its ordinates at a series of points spaced $1/2W$ seconds apart

$$x(t) = \sum_n x\left(\frac{n}{2W}\right) \operatorname{sinc}(2Wt - n), \qquad \operatorname{sinc}(\alpha) = \frac{\sin(\pi\alpha)}{\pi\alpha}$$

Shannon, '49



- **Model:** W-Bandlimited signals
- **Sampling:** Pointwise at rate $1/T \geq 2W$
- **Reconstruction:** Interpolation by $h(t) = \operatorname{sinc}(2Wt)$

Avoiding High-Rate ADC

- Use several samplers:
  - Papoulis' theorem
  - Time-interleaved ADC (special case)

$\Big\}$ Overall rate = Nyquist

- Exploit signal structure (subspace):
  - Pulse streams
  - Multiband sampling

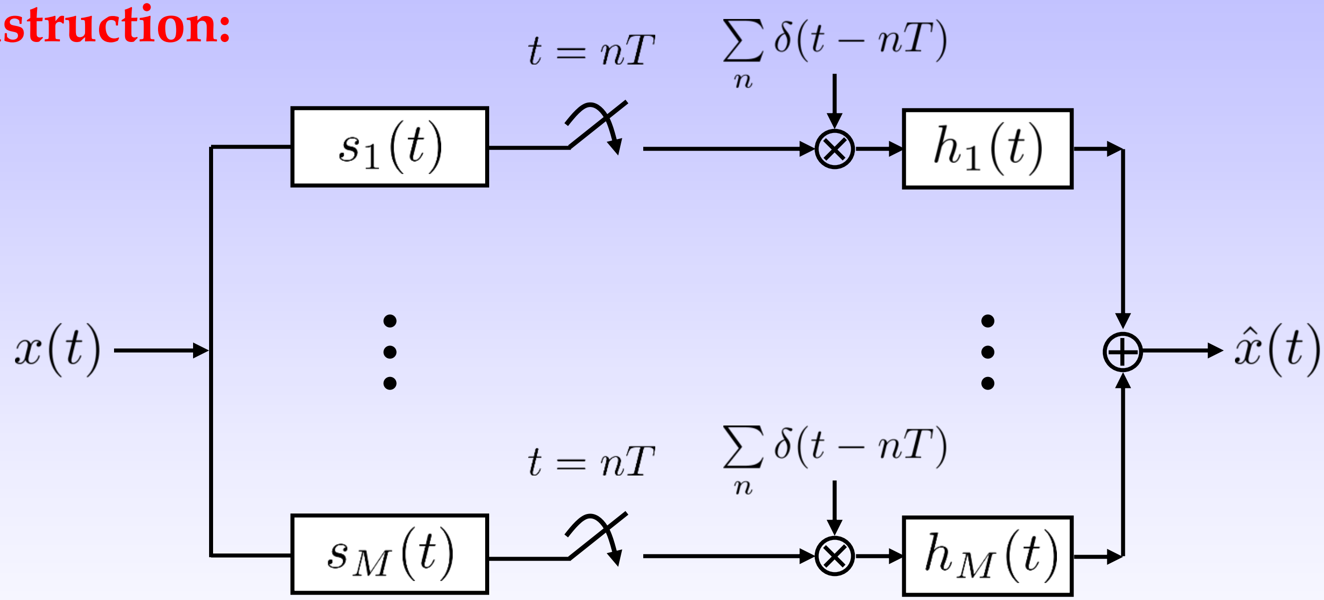
$\Big\}$ Can approach information rate

Papoulis' Theorem

- **Model:** W-bandlimited (same)
- **Sampling:** M branches sampled at $1/M$ the Nyquist rate, $\frac{1}{T} \geq \frac{2W}{M}$
  Flexible constraints on $s_i(t), h_j(t)$

- **Reconstruction:**

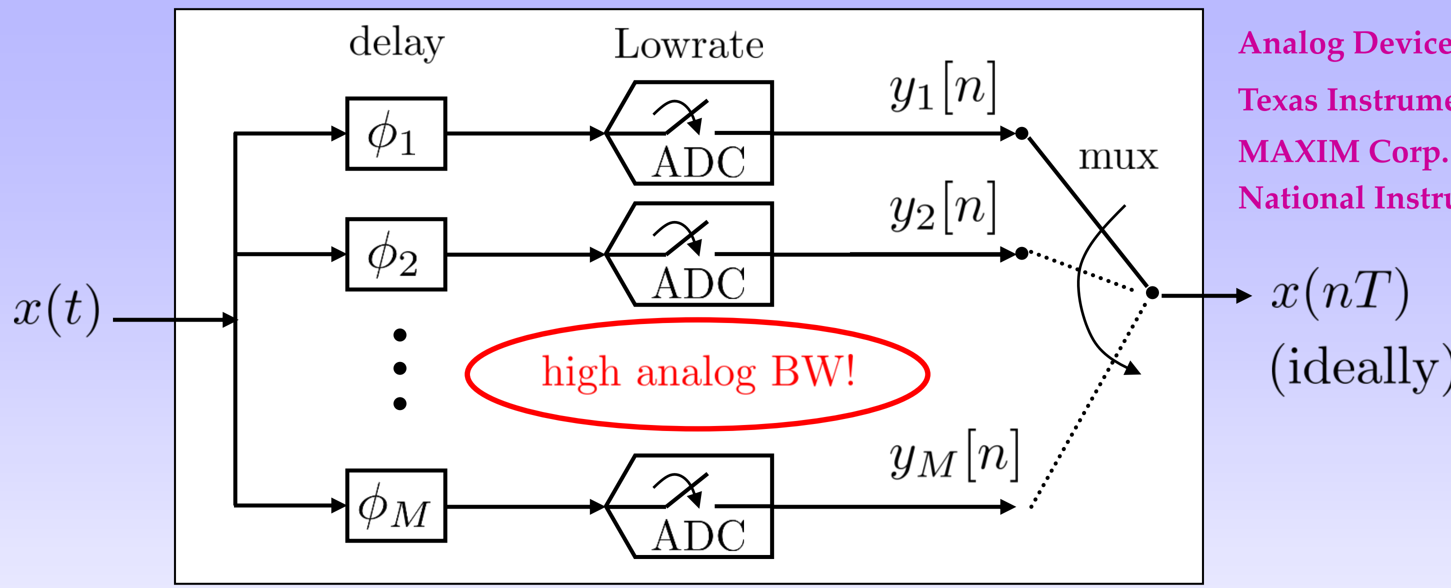
Papoulis, '77



- Overall rate is $2W$ (same)

Time-Interleaved ADCs

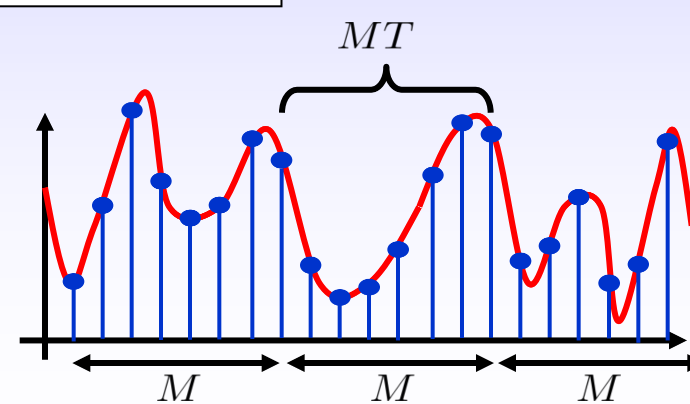
A high-rate ADC comprised of a bank of lowrate devices



Analog Devices Corp.
Texas Instruments Corp.
MAXIM Corp.
National InstrumentsCorp.

- Each branch (coset) undersamples at $1/M$ of the Nyquist-rate
- Widely-researched

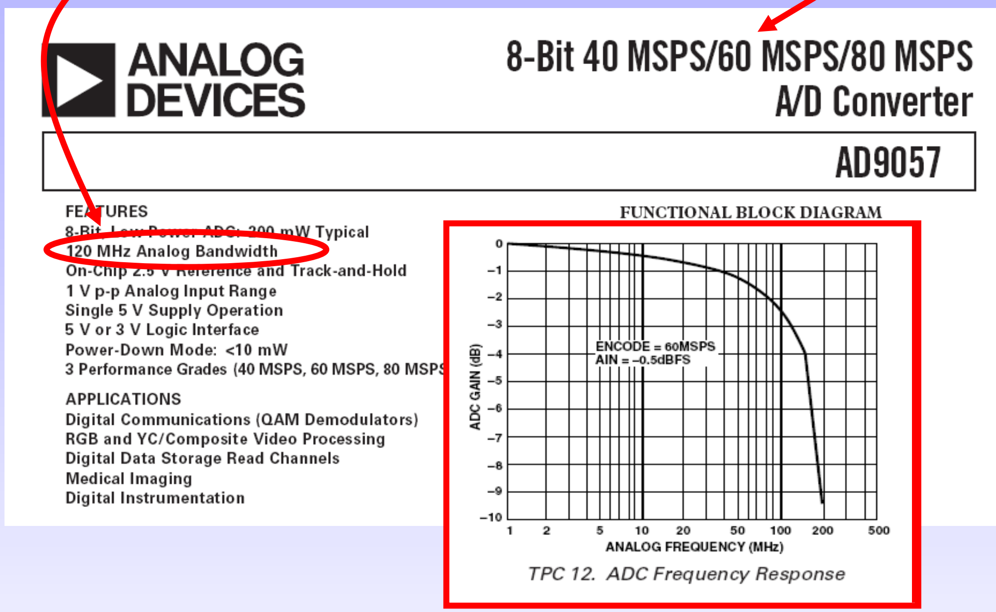
Yen, '56
Eldar and Oppenheim, '00
Johansson and Lowenborg, '02
Levy and Hurst, '04
…and more



Practical ADC Devices

Analog bandwidth limitation b

Sampling rate r


ANALOG DEVICES

8-Bit 40 MSPS/60 MSPS/80 MSPS A/D Converter

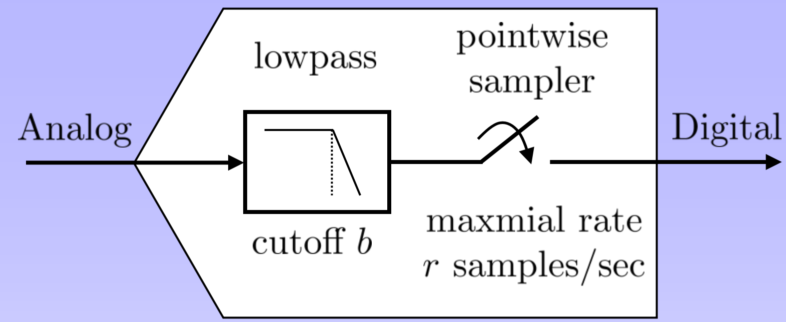
AD9057

FEATURES
8-Bit, Low Power ADC: 200 mW Typical
120 MHz Analog Bandwidth
On-Chip 2.5 V Reference and Track-and-Hold
1 V p-p Analog Input Range
Single 5 V Supply Operation
5 V or 3 V Logic Interface
Power-Down Mode: <10 mW
3 Performance Grades (40 MSPS, 60 MSPS, 80 MSPS)

APPLICATIONS
Digital Communications (QAM Demodulators)
RGB and YC/Composite Video Processing
Digital Data Storage Read Channels
Medical Imaging
Digital Instrumentation

FUNCTIONAL BLOCK DIAGRAM

TPC 12. ADC Frequency Response



In time-interleaved architectures:

- The overall rate is Nyquist
- Each branch needs front-end with Nyquist bandwidth (**will be important later**)
- Accurate time delay are required ϕ_i

Black and Hodges, '80
Jenq, '90
Elbornsson *et al.*, '05
Divi and Wornell, '09
Murmann *et al.*, '09
Goodman *et al.*, '09
…and more

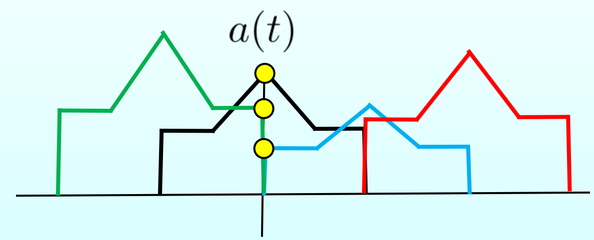
Generalized Sampling in a Subspace

- **Model:** Shift-invariant (SI) subspace of possible inputs

$$\mathcal{A} = \left\{ x(t) = \sum_n d[n] a(t - nT), \quad d[n] \in \ell_2(\mathbb{R}) \right\}$$

$$a_n(t) = \mathrm{sinc}(2Wt - n)$$

$$\mathcal{A} = W\text{-bandlimited}$$



$\mathcal{A}$ is NOT bandlimited

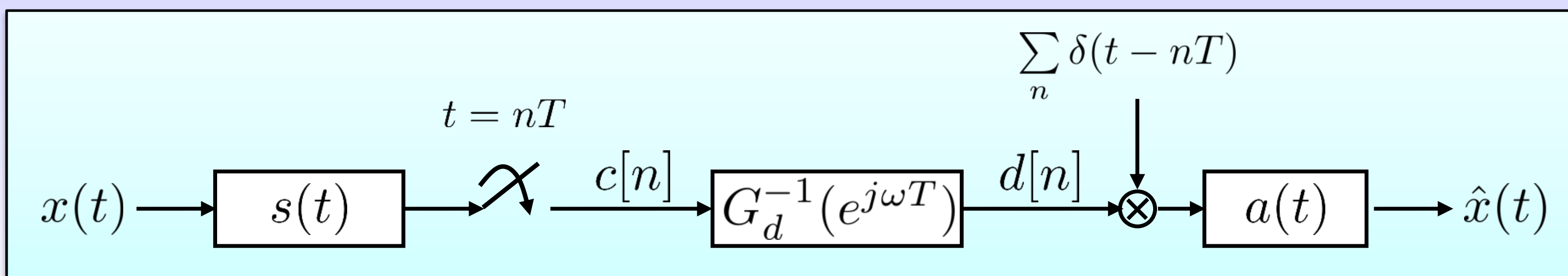
- Practical ! *e.g.*, splines, pulse amplitude modulation (PAM), and more…
- **Sampling:** Inner products, $c[n] = \langle x(t), s_n(t) \rangle$
  - $s_n(t) = \delta(t - nT) \longrightarrow$ pointwise sampling $c[n] = x(nT)$
  - $s_n(t) = s(t - nT) \longrightarrow$ $x(t) \longrightarrow \boxed{s(t)} \xrightarrow{t = nT} c[n]$

Reconstruction from Generalized Samples

- Shift-invariant case
  - **Model:** $x(t) = \sum_n d[n]a(t-nT)$ $\Rightarrow$ $X(\omega) = D(e^{j\omega T})A(\omega)$
  - **Sampling:** $c[n] = \langle x(t), s(t-nT)\rangle$

$$c(e^{j\omega T}) = \sum_k X(\omega+2\pi k)S^*(\omega+2\pi k) = D(e^{j\omega T})G_d(e^{j\omega T})$$

  - **Recovery:** Filter by $G_d^{-1}(e^{j\omega T})$ to obtain $d[n]$, then interpolate $\hat{x}(t)$

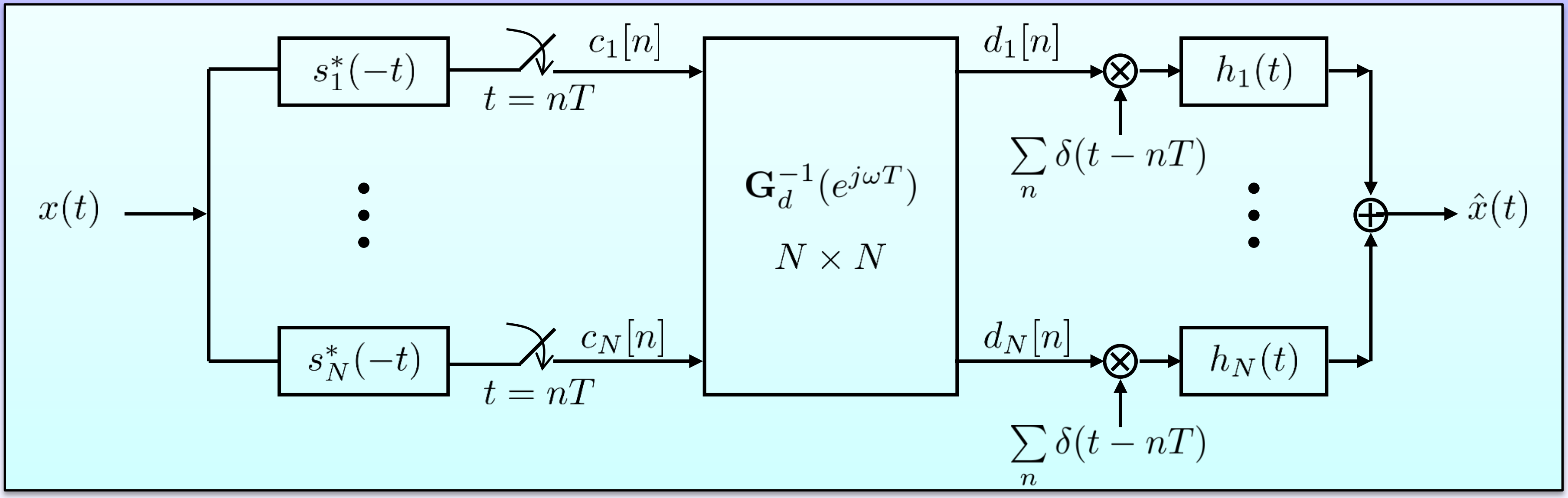


- Sampling rate is $\frac{1}{T}$ rather than the Nyquist rate of $x(t)$
- Approach does not depend on $f_{\max}$

Aldroubi and Unser, '94
Christensen and Eldar, '05

Multiple Shift-Invariant Generators

- **Model:** $$x(t) = \sum_{l=1}^{N} \sum_{n} d_l[n] a_l(t - nT)$$
- **Sampling / Reconstruction:**



$$\left[\mathbf{G}_d(e^{j\omega T})\right]_{i\ell} = \frac{1}{T} \sum_{k \in \mathbb{Z}} S_i^* \left(\frac{\omega}{T} - \frac{2\pi}{T}k\right) H_\ell \left(\frac{\omega}{T} - \frac{2\pi}{T}k\right)$$

- Sampling rate is $\frac{N}{T}$ → independent of $f_{\max}$

de Boor, DeVore and Ron , '94
Christensen and Eldar, '05

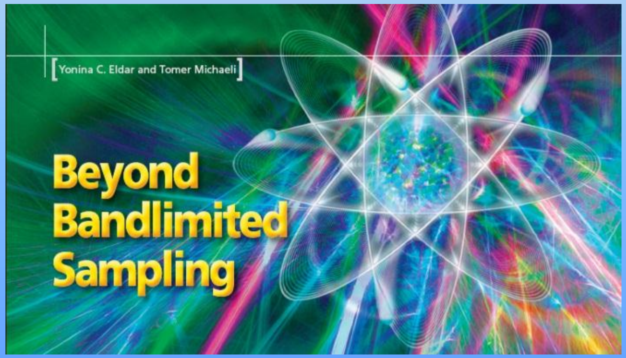
Multiple Shift-Invariant Generators

- **Model:** $$x(t) = \sum_{l=1}^{N} \sum_{n} d_l[n] a_l(t - nT)$$

- Previous work extends theory to arbitrary subspaces
- Many beautiful results, and many contributors

(Unser, Aldroubi, Vaidyanathan,Blu, Jerri, Vetterli, Grochenig, DeVore,Christensen,Schoenberg,Eldar …)

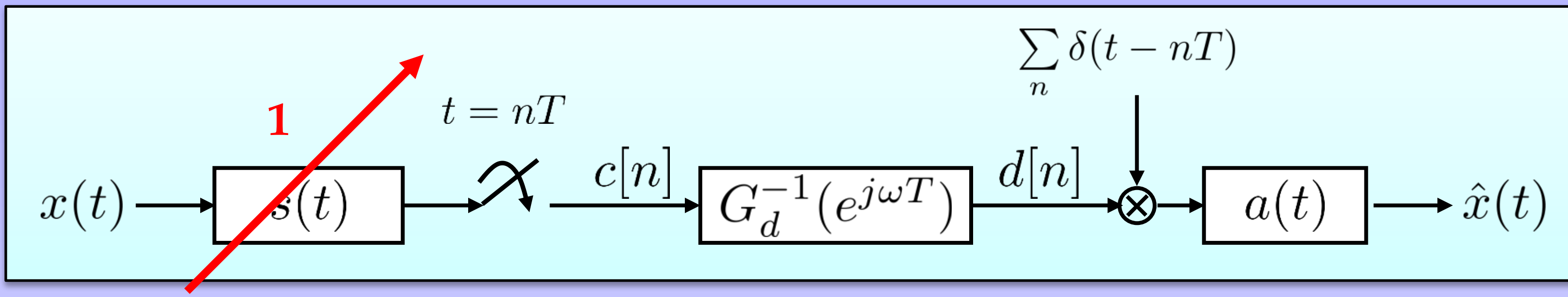
More information:
Y. C. Eldar and T. Michaeli,
"Beyond Bandlimited Sampling,"
IEEE Signal Proc. Magazine,
26(3): 48-68, May 2009

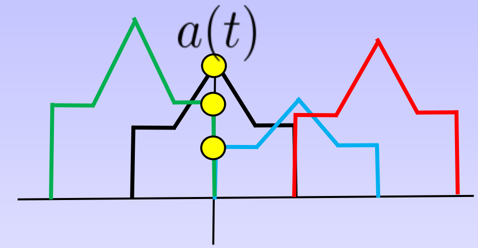


- Sampling rate is $\frac{N}{T}$ → independent of $f_{\max}$

de Boor, DeVore and Ron , '94
Christensen and Eldar, '05

Toy-Example (1)





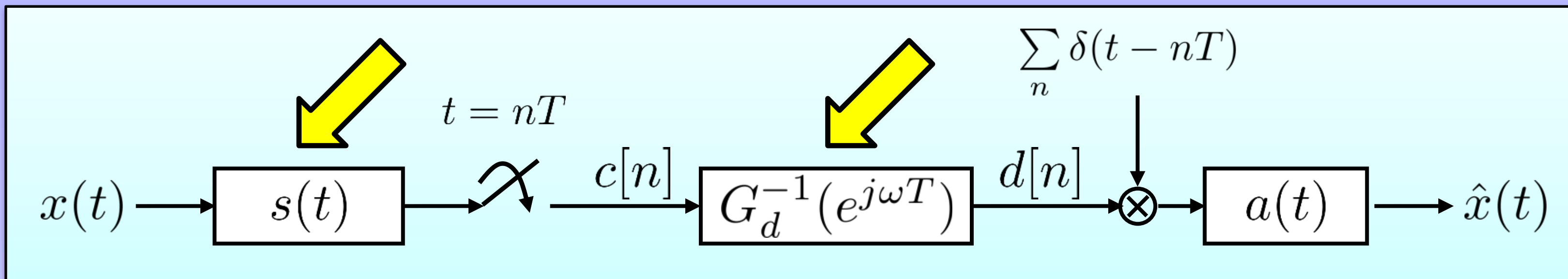
- **Model:** $x(t) = \sum d[n]a(t-nT)$
- **Sampling:** choose $s(t) = \delta(t)$
  3 adjacent shifts contributes to each sample
- **Recovery:** exploit known shape $a(t)$

Subspace prior

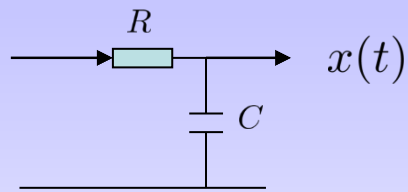
$$G_d^{-1}(e^{j\omega T}) = \frac{1}{\sum_k A(\omega - 2\pi k/T)}$$

- **Rate:** $\frac{1}{T}$
- $f_{\max}$ can be very high, since $x(t)$ is not bandlimited

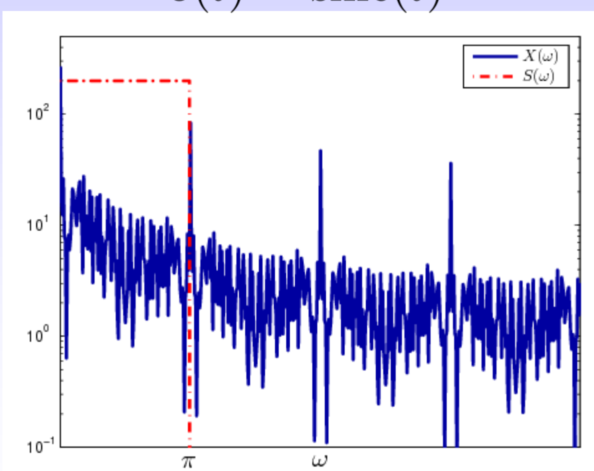
Toy-Example (2)



- **Model:** $a(t) = \frac{1}{\tau} e^{-t/\tau} u(t)$



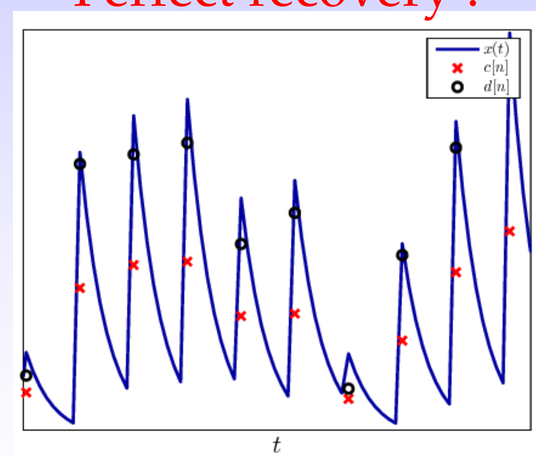
$s(t) = \mathrm{sinc}(t)$



Rate: $\frac{1}{T}$

$f_{\max}$ is high…

Perfect recovery !

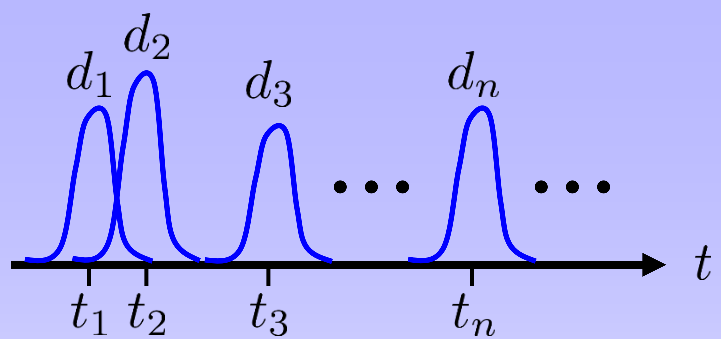


Lowpass data can contain all relevant information !

Pulse-streams (known locations)

- **Model:** fixed delays t_n , unknown d_n

$$x(t) = \sum_n d_n h(t - t_n)$$



- **Sampling:** design $s_n(t) = h(t - t_n)$ and sample $c[n] = \langle x(t), s_n(t) \rangle$
  t_n and $h(t)$ are known

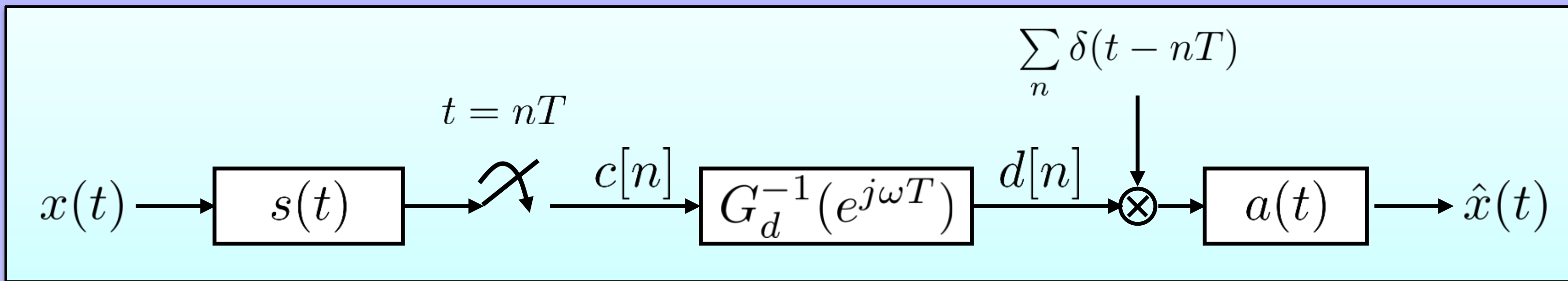
- **Recovery:** $\{d_n\}, \{c[n]\}$ satisfy a linear system, with coefficients depending on t_n and $h(t)$

$$c[n] = d_n \|h(t)\|^2 \quad \text{(for the easiest case with no overlaps)}$$

- **Rate:** $\text{information rate} = \#\text{pulses/second}$

- $f_{\max}$ is high, since $x(t)$ is not bandlimited

Generalized Sampling in Practice



So far:

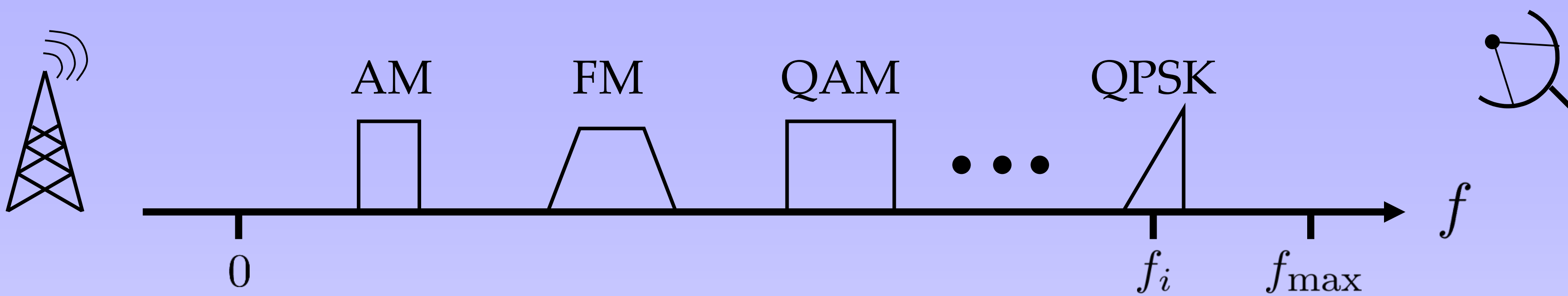
- Toy-examples: perfect recovery of nonbandlimited inputs ! ($\mathcal{A}$ =SI)
- Pulse streams, $\mathcal{A} =$ known pulse shape and fixed delays

A common denominator

Design assumption	Sampling & processing rates	
$f_{\max}$-bandlimited	High	
exact knowledge $x(t) \in \mathcal{A}$	Approach minimal	

- **Next slides:** Multiband signals, $\mathcal{A} =$ known carrier frequencies

Multiband (known carriers)



- **Model:** narrowband transmissions in wideband range, modulated on carrier frequencies $f_i \leq f_{\max}$

- **Sampling:**
  - RF demodulation
  - Undersampling
  - Nonuniform strategies

  Utilize knowledge $x(t) \in \mathcal{A}$

- Sampling and processing at rate $f_{\max}$ are often impractical

Landau's Theorem

- States the minimal sampling rate for any (pointwise) sampling strategy that utilizes frequency support knowledge

Theorem (known spectral support)

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

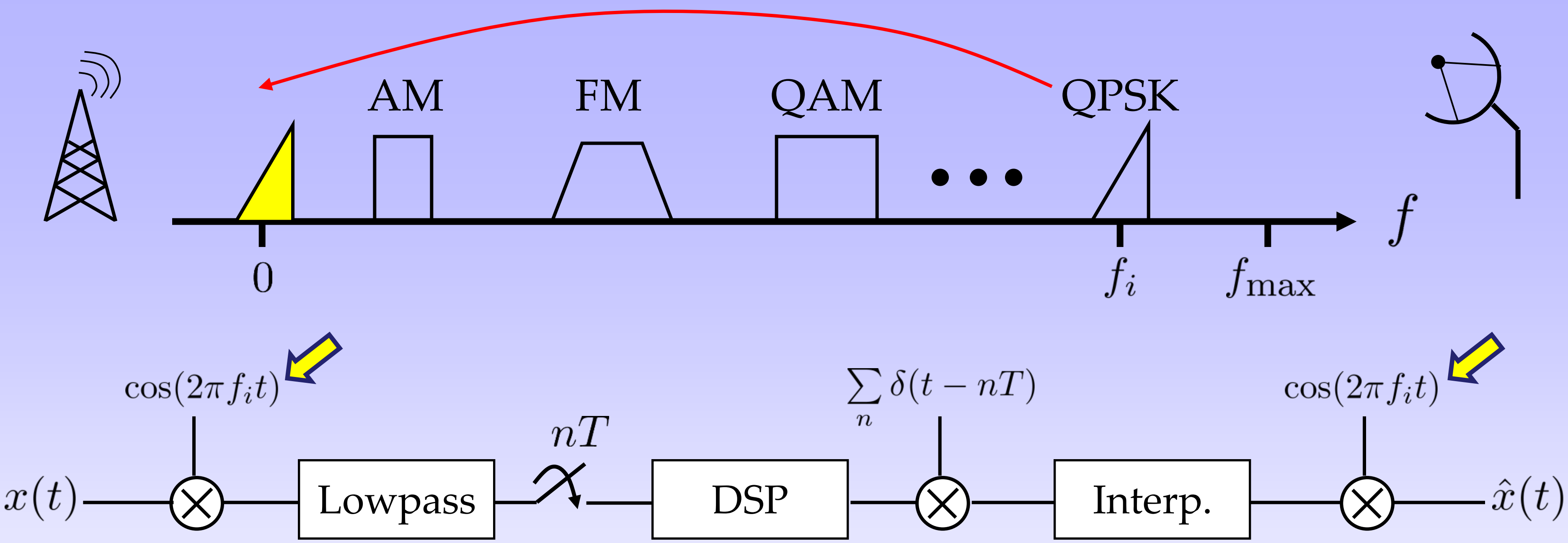
$$\underbrace{D^-(R)}_{\text{Average sampling rate}} \geq \operatorname{meas}(\mathcal{F})$$

Landau, '67

- N bands, individual widths $\leq B$, requires at least NB samples/sec

- Note: → bandpass with single-side width B requires $2B$ samples/sec
  → k transmissions result in $N = 2k$ bands (conjugate symmetry)

RF Demodulation

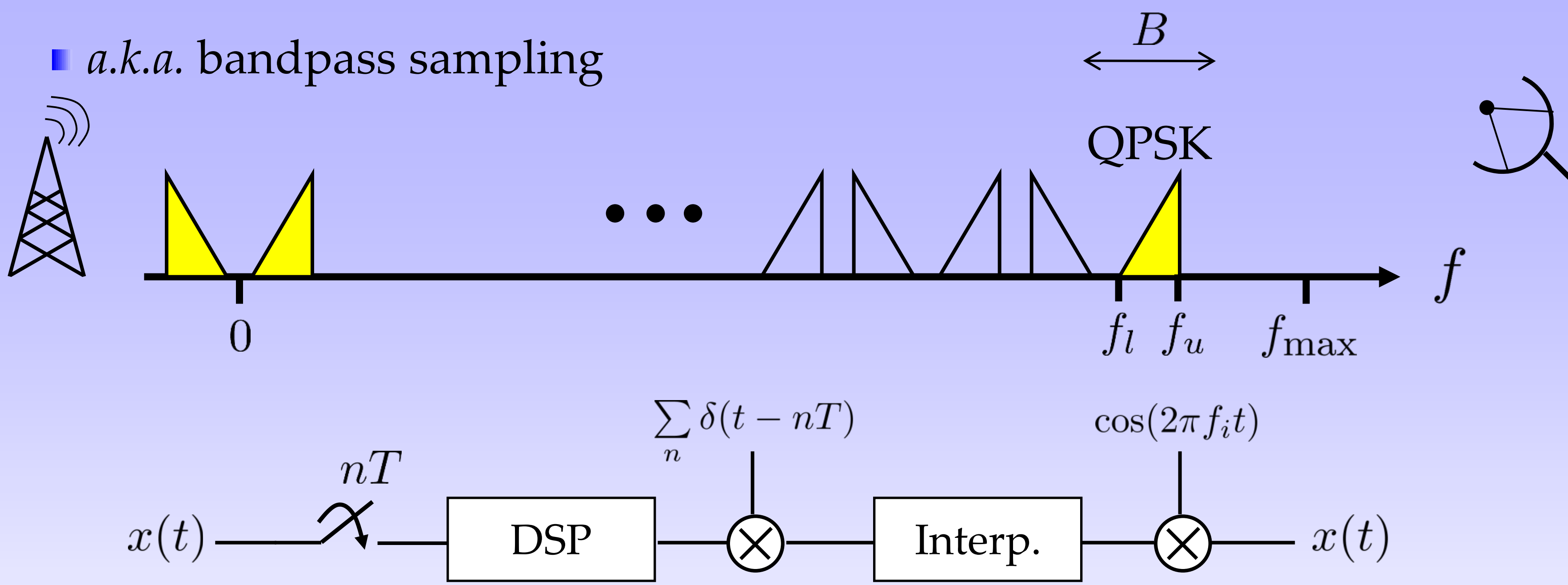


- f_i value is used in sampling and reconstruction
- Analog preprocessing with RF devices (1 branch/transmission)
- **Minimal rate:** NB
- Zero-IF, low-IF topologies

Crols and Steyaert, '98

Undersampling

- *a.k.a.* bandpass sampling



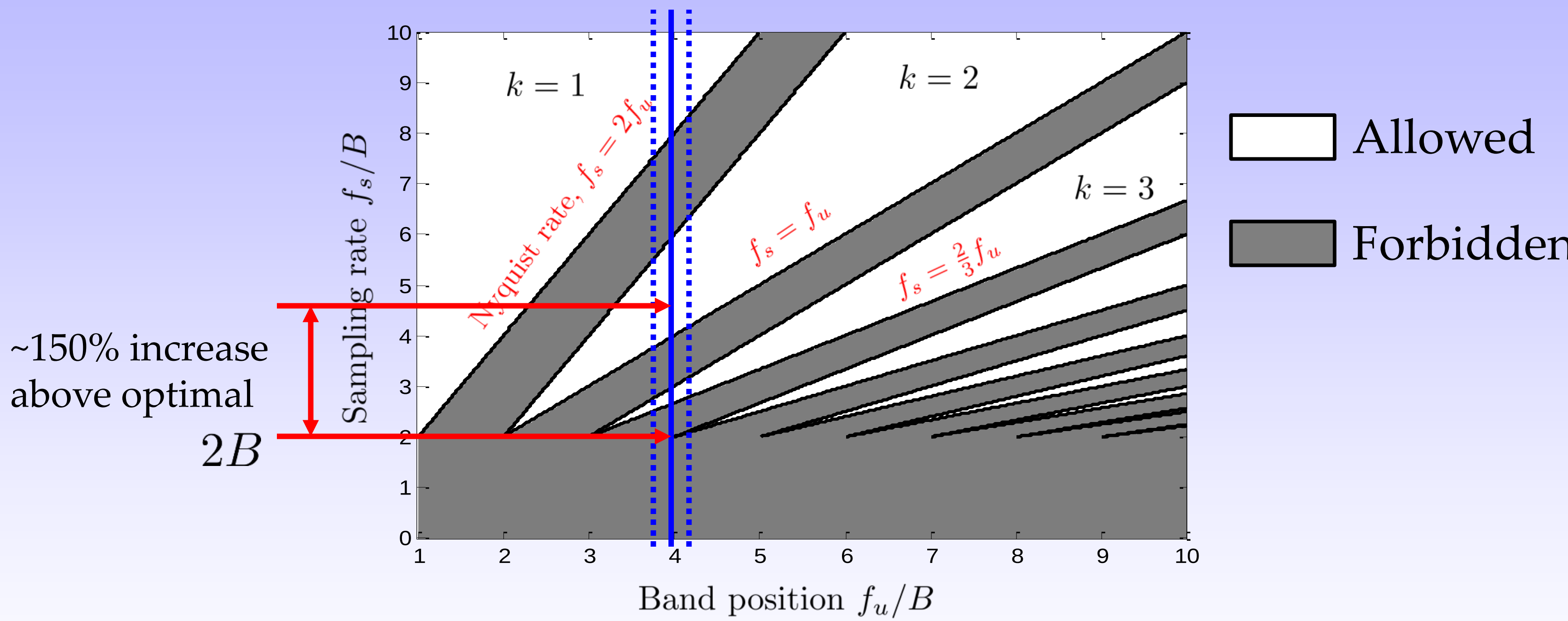
- **Sampling:** Select rate to satisfy ``alias free condition''
- **Reconstruction:** Same as in RF demodulation
- No analog preprocessing

Allowed Undersampling Rates

- Sampling rate must be chosen in accordance to band location:

$$\frac{2f_u}{k} \le f_s \le \frac{2f_l}{k-1}$$

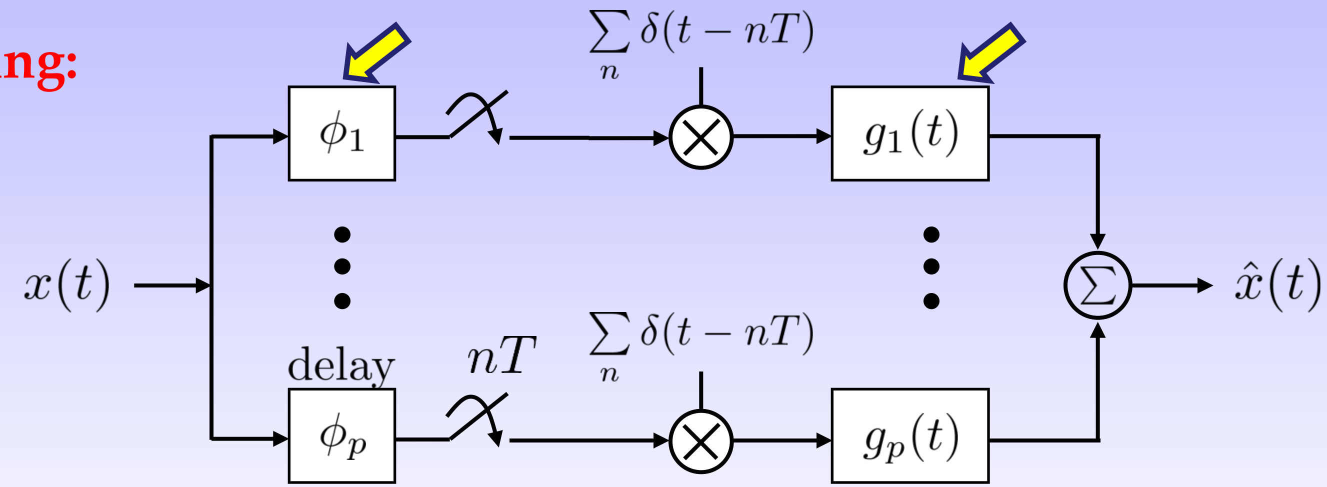
Vaughan *et al.*, '91



- Robustness to model mismatch requires significant rate increase
- Multiband alias-free conditions are complicated and generally do not result in significant rate reduction

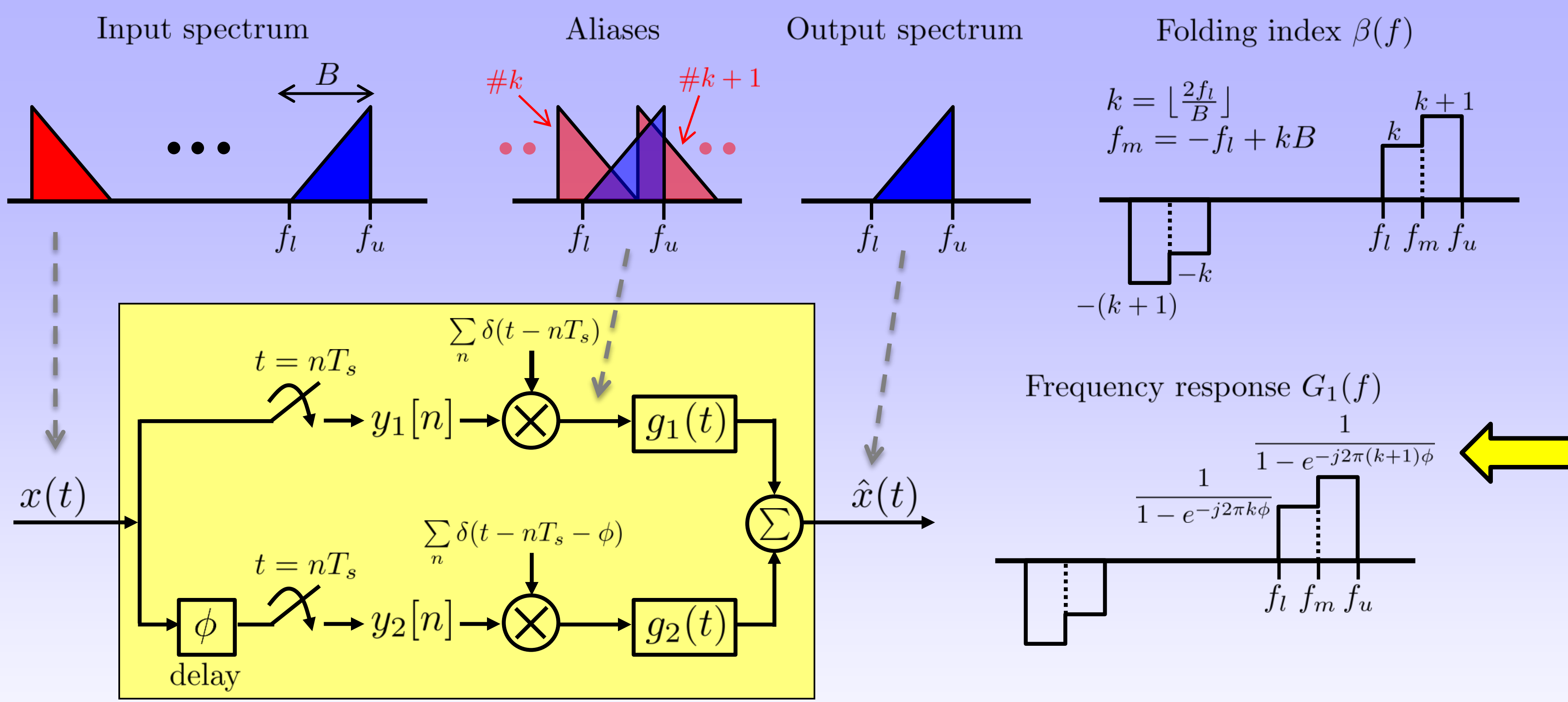
Periodic Nonuniform Sampling

- Advantages:
  - No analog preprocessing
  - No ``alias-free'' conditions, work for multiband
  - Approach minimal rate NB
- **Sampling:**



- In general, a p'th-order PNS can resolve up to p aliases:
  - Bandpass sampling at average rate $2B$ — Kohlenberg, '53
  - Multiband sampling at rate approaching minimal — Lin and Vaidyanathan, '98

Reconstruction from 2nd order PNS



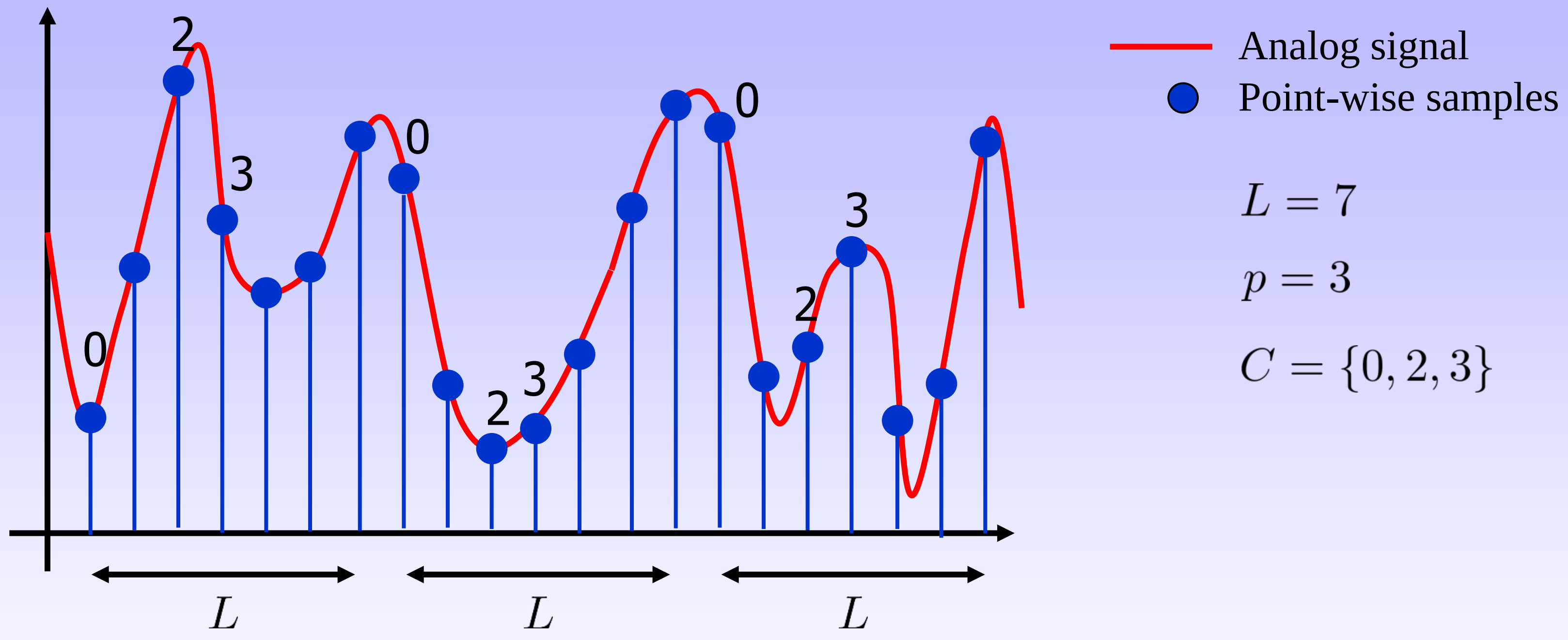
- Delays result in different linear combinations of the bands

$$\left.\begin{aligned} T_s\,Y_1(f) &= X(f) + X(f - \beta(f)B) \\ T_s\,Y_2(f) &= X(f) + X(f - \beta(f)B)e^{-j2\pi\beta(f)\phi B} \end{aligned}\right\}$$

Choose ϕ such that $e^{-j2\pi\beta(f)\phi B} \neq 1$

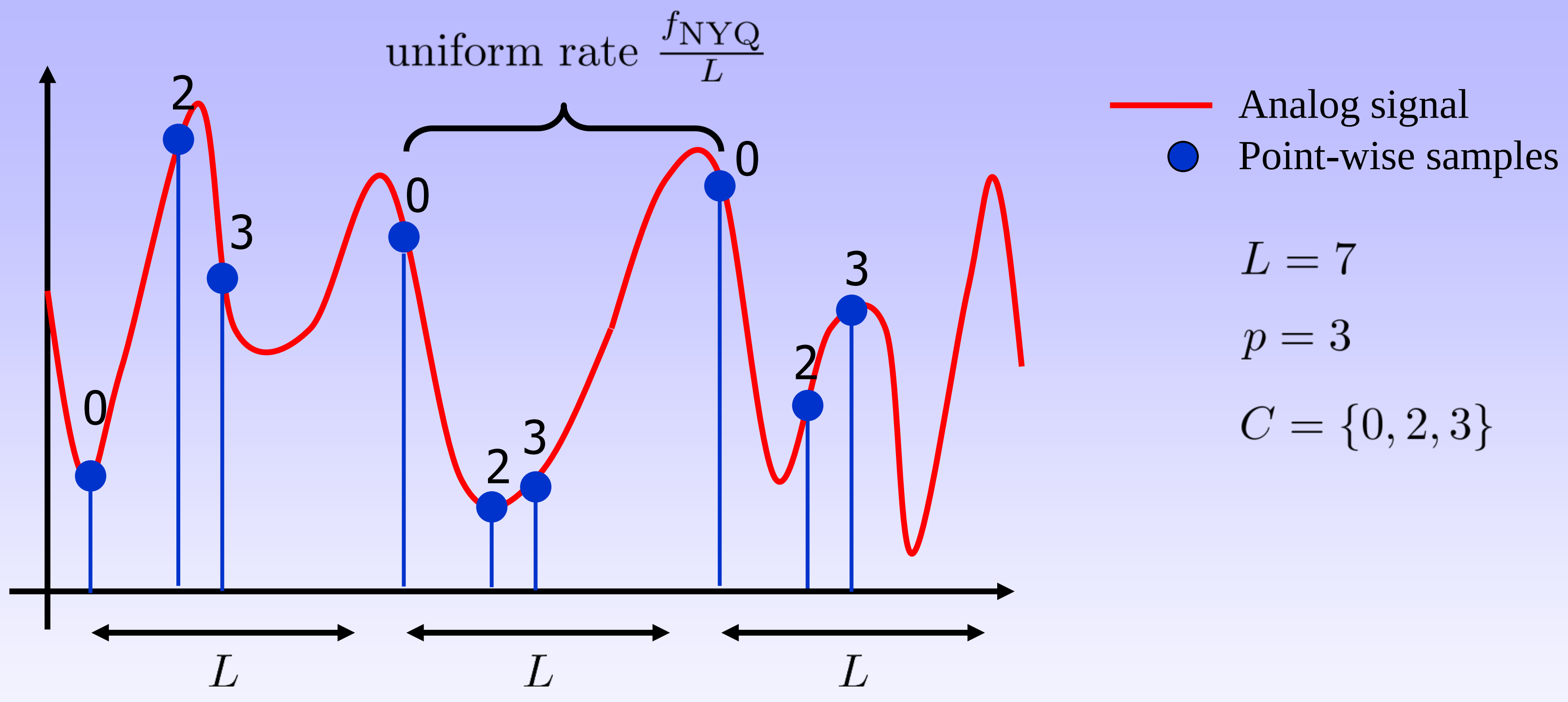
Multi-Coset Sampling

- PNS with delays $\{\phi_i\}$ on the Nyquist grid



Multi-Coset Sampling

- PNS with delays $\{\phi_i\}$ on the Nyquist grid



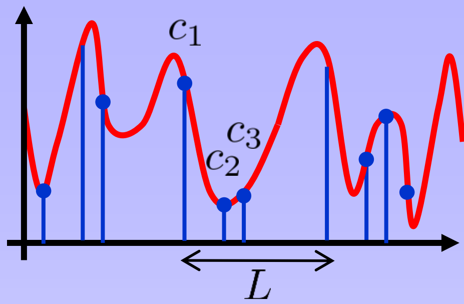
$L = 7$

$p = 3$

$C = \{0, 2, 3\}$

Multi-Coset Sampling

- PNS with delays $\{\phi_i\}$ on the Nyquist grid



- Semi-blind approaches:
  - Choose $\{\phi_i\}$ universally (or at random)
  - Design reconstruction filters $g_1(t), \ldots, g_p(t)$

Herley *et. al.*, '99
Bresler *et al.*, '00

- ``Blind'' recovery:

Bresler *et al.*, '96,'98

$$\min_{|\mathcal{K}|=q} \operatorname{trace}(P_{\mathcal{K}}\mathbf{R}) \qquad \mathbf{R} = \text{measurements covariance}$$

- Positions are implicitly assumed:
  - $q = q(x(t))$ depends on band positions
  - Recovery fails if incorrect value is used for q
  - Result requires random signal model, and holds *almost surely*

Completely blind = Unknown carriers = not a subspace model !

Short Summary

- Subspace models
  - Linear, easy to treat mathematically
  - Not necessarily bandlimited

- Generalized sampling theory
  - Treat arbitrary subspace models
  - Many classic approaches can be derived from theory
  - Rate is proportional to actual information rate rather than Nyquist

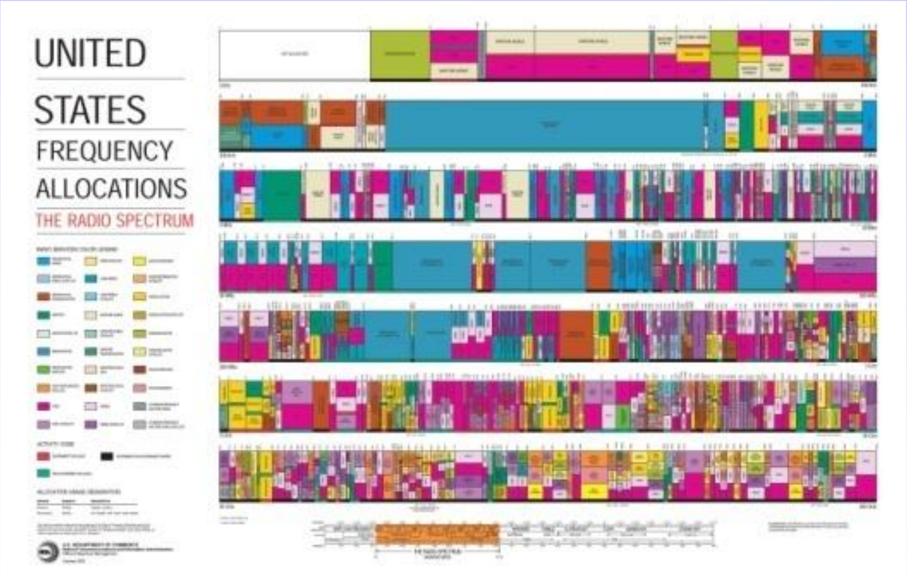
But, what if…

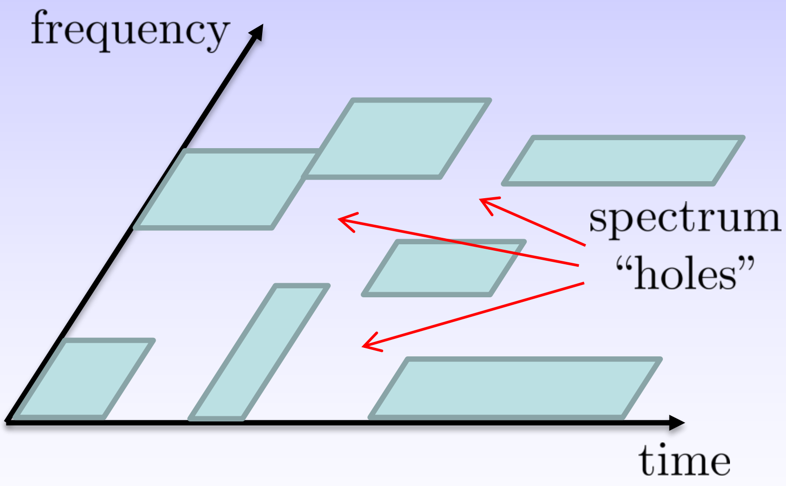
- the input model is not linear ?
  (for example, when carrier frequencies or times of arrivals are unknown)

- Answer: the rest of this tutorial

Nonlinear Models – Motivation

- Encountered in practical applications:
  - Cognitive radio mobiles utilize unused spectrum ``holes'', spectral map is unknown a-priori

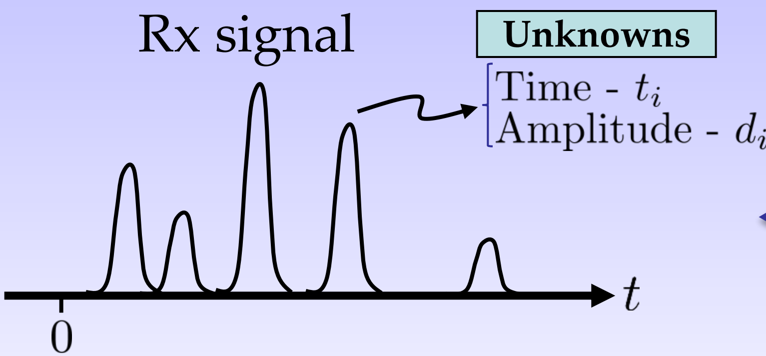


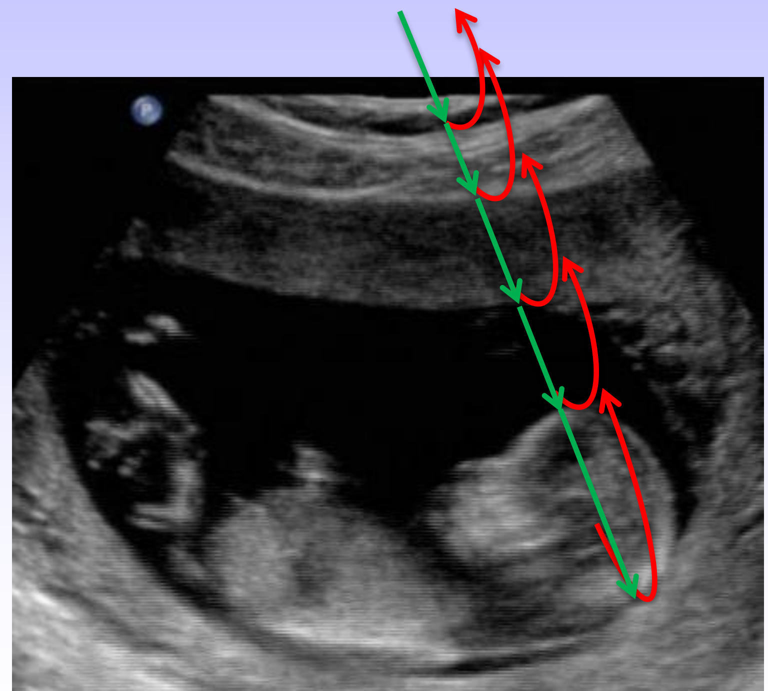


Nonlinear Models – Motivation

- Encountered in practical applications:
  - Cognitive radio mobiles utilize unused spectru es'', spectral map is unknown a-priori
  - Ultrasound, reflections are intercepted at unkn ays

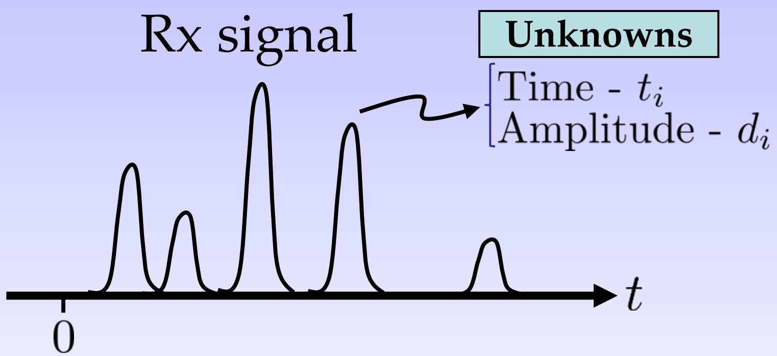
Ultrasonic probe

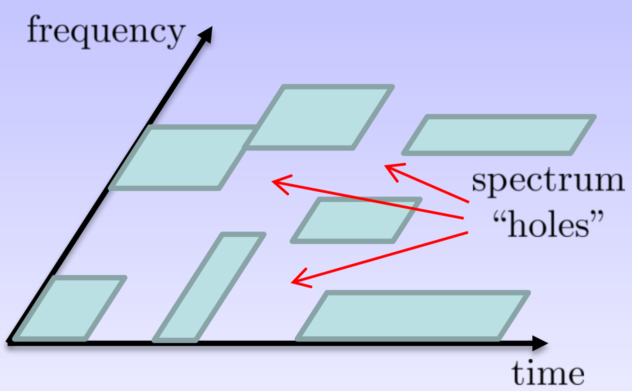




Nonlinear Models – Motivation

- Encountered in practical applications:
  - Cognitive radio mobiles utilize unused spectrum ``holes'', spectral map is unknown a-priori
  - Ultrasound, reflections are intercepted at unknown delays





- Do not fit subspace modeling … we can always sample at rate $2f_{\max}$

- Questions:
  - Better modeling? Subspace up to some uncertainty ?
  - Can we sample and process at rates below $2f_{\max}$ with proper modeling?

– Part 3 –
Union of Subspaces

→ Outline

Model

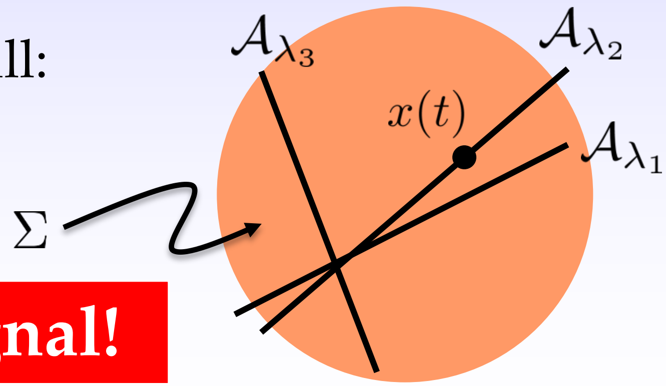
- Signal belongs to one out of (possibly infinitely-)many subspaces

$$x(t) \in \mathcal{U} \qquad \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda$$

Lu and Do, '08
Eldar and Mishali, '09

- Each λ corresponds to a different subspace $\mathcal{A}_\lambda$
- $x(t)$ belongs to $\mathcal{A}_{\lambda^*}$, for some $\lambda^* \in \Lambda$ → But, λ^* is unknown a-priori
- $\mathcal{U}$ is a nonlinear model: $x, y \in \mathcal{U} \xrightarrow{\text{typically}} x + y \notin \mathcal{U}$
- A union is generally a true subset of its affine hull:

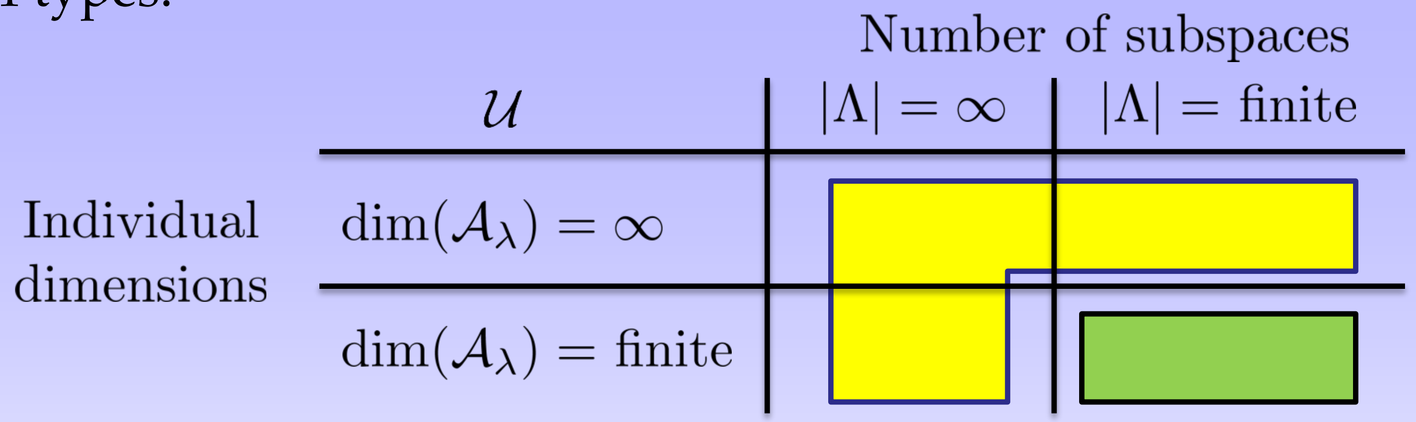
$$\mathcal{U} \subsetneq \Sigma = \{x + y \,|\, x, y \in \mathcal{U}\}$$



The union tells us more about the signal!

Union Types

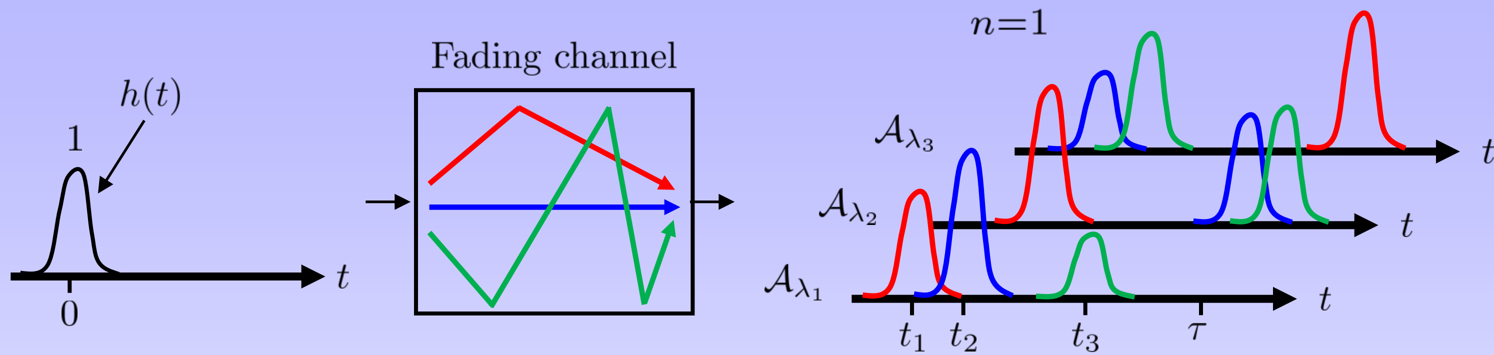
- 4 types:



- Legend:
  - = General analog union models
    Infiniteness enters in either $\dim(\mathcal{A}_\lambda)$ or $|\Lambda|$
  - = Discrete models, *e.g.*, sparse trigonometric polynomials
    $p(t) = \sum_{n=1}^{N} c_n e^{jnt}$, with only k nonzero coefficients
    continuous-time signals with finite parameterization

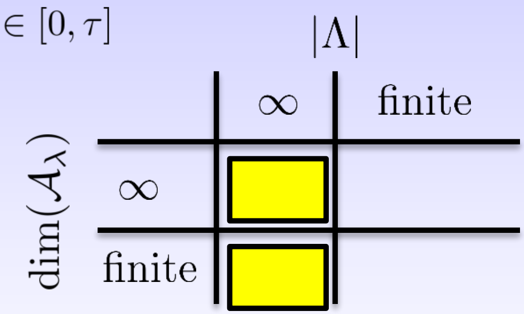
Examples: Analog Unions (1)

- Pulses with unknown time delays $x(t) = \sum_{n=1}^{L} d_n h(t - t_n)$



Union over possible path delays $t_i \in [0, \tau]$

- Dimensions:
  - $t_i \in [0, \tau], \ \lambda = \{t_i\}$
  - $\mathcal{A}_\lambda = [d_1, \ldots, d_L]^T \ \rightarrow \ \dim(\mathcal{A}_\lambda) = L$



- A special case of a broader model: finite rate of innovation (FRI)
  Here, innovation rate = $2L/\tau$

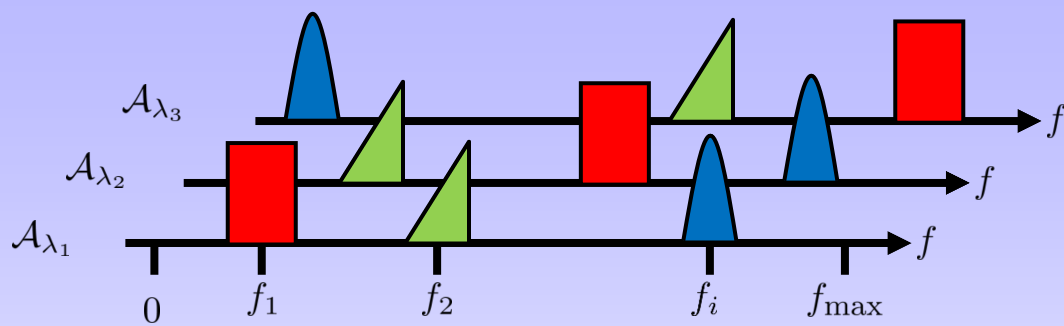
Vetterli *et al.*, '02-'11

- Sequences of innovation model has both dimensions infinite

Gedalyahu and Eldar, '09-'11

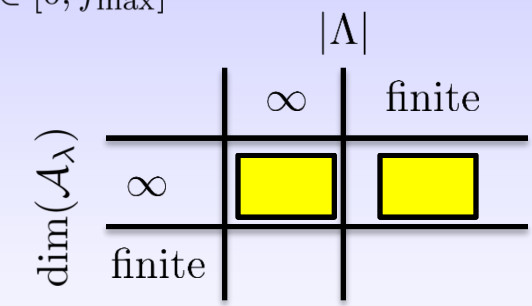
Examples: Analog Unions (2)

- Multiband with unknown carrier frequencies $\lambda = \{f_i\}$



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

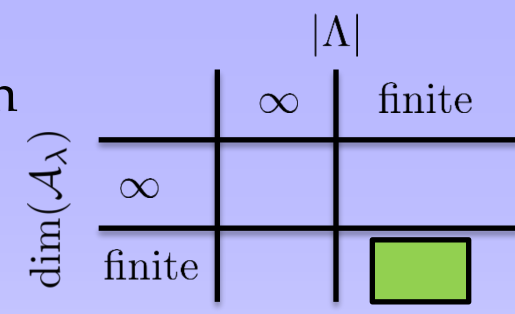
- Dimensions:
  - $f_i \in [0, f_{\max}]$
  - $\mathcal{A}_\lambda$ is a bandpass signal



- Another viewpoint with $|\Lambda|$ =finite and $\dim(\mathcal{A}_\lambda) = \infty$ is described later on
(efficient hardware and software implementation)

Mishali and Eldar '07-'11

Examples: Discrete Unions

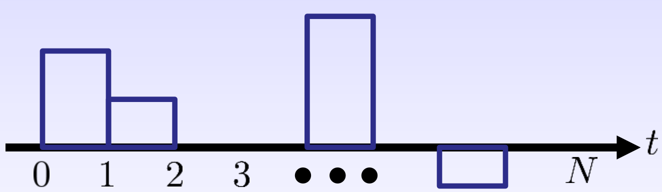


- Signal model has underlying finite parameterization

- Continuous-time examples:
  - Sparse trigonometric polynomials
    $p(t) = \sum_{n=1}^{N} c_n e^{jnt}$, with only k nonzero coefficients

Kunis and Rauhut, '08
Tropp *et al.*, '09

  - Sparse piece-wise constant with integer knots

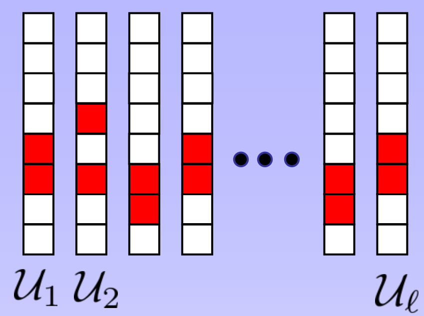


- Discrete-time examples:
  - Compressed sensing
  - Block sparsity, tree-sparse models

Donoho, Candès-Romberg-Tao, '06
Baraniuk *et al.*, Eldar *et al.*, '09-'11

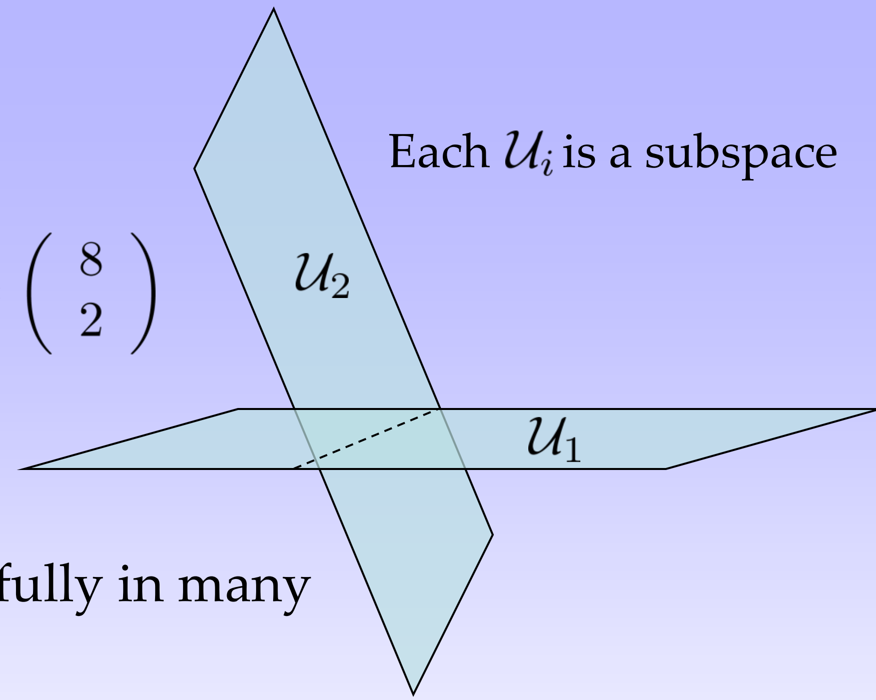
Compressed Sensing = Union

2 - sparse



$$\ell = \binom{8}{2}$$

Each $\mathcal{U}_i$ is a subspace



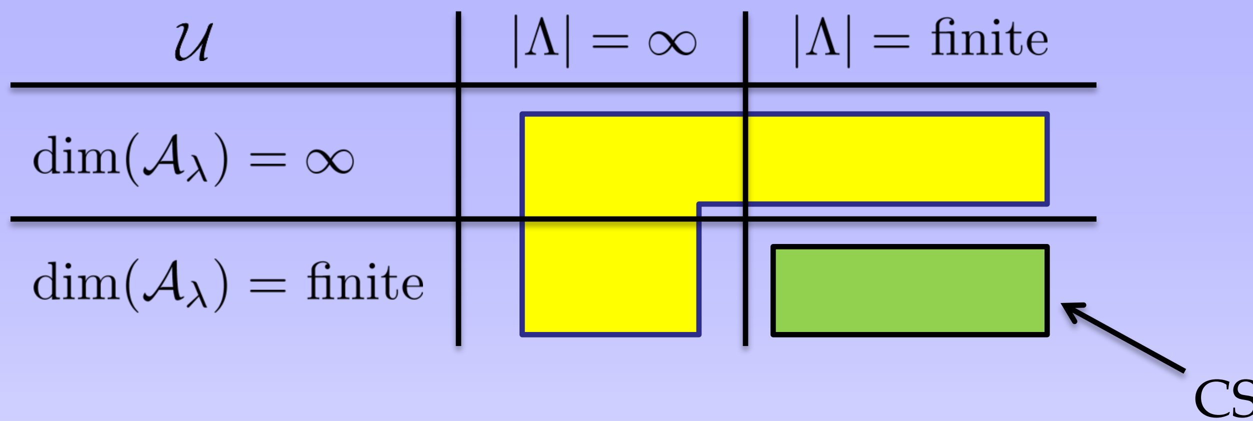
Sparsity models have been used successfully in many applications such as:

- Denoising and deblurring
- Tracking and classification
- Compressed sensing

Donoho, Johnstone, Mallat, Sapiro, Ma, Vidal, Starck, …

Candès, Romberg, and Tao '06
Donoho '06

Compressed Sensing



- For sub-Nyquist sampling, our focus is on infinite unions

- We will start with compressed sensing (CS)
  - easier to explain
  - methods for infinite unions also rely on CS algorithms

- Following a short intro on CS → Xampling and analog systems

Short Intro

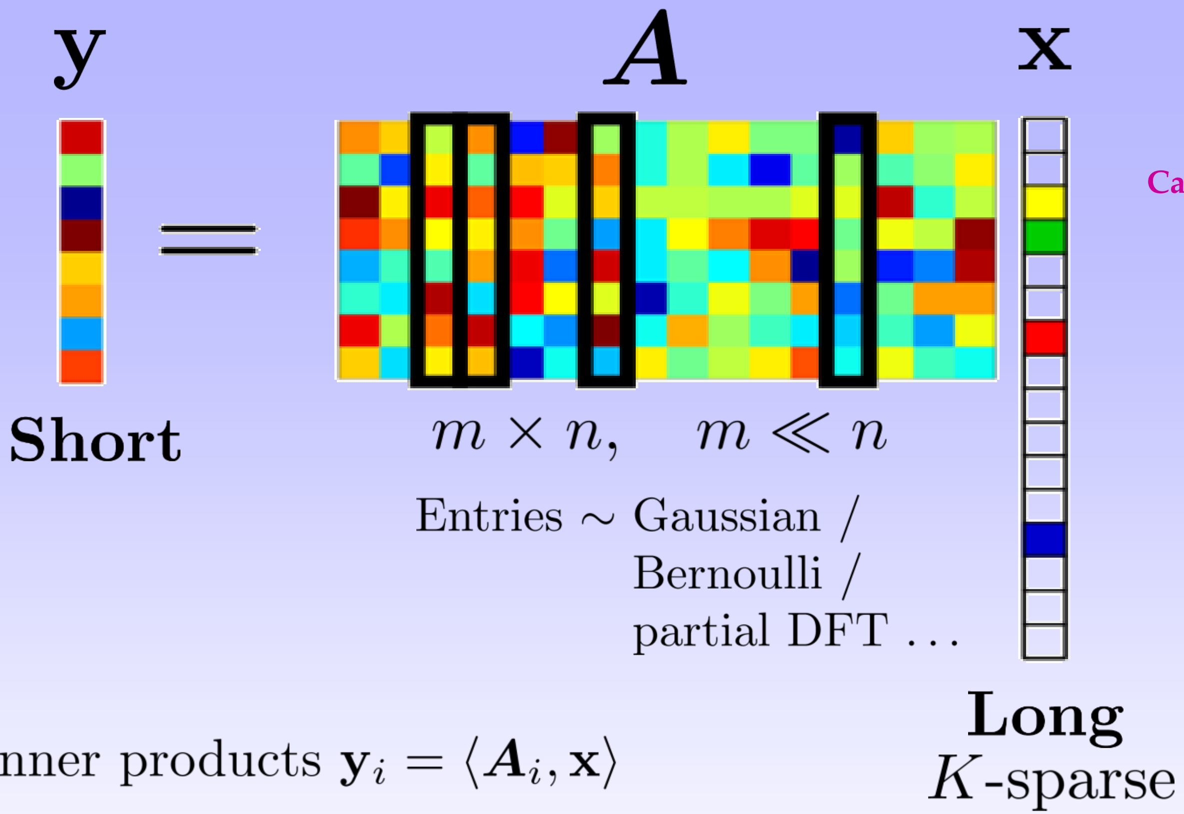
"Can we not just **directly measure** the part that will not end up being thrown away ?"

Donoho, '06

Original 2500 KB
100%

Compressed 148 KB
6%

In a Nutshell…



Donoho, '06
Candès-Romberg-Tao, '06

Main ideas:

- Sensing = inner products $\mathbf{y}_i = \langle \boldsymbol{A}_i, \mathbf{x} \rangle$
- Random projections
- K non-zero values requires at least $2K$ measurements
- Recovery: brute-force, convex optimization, greedy algorithms

Concept

Goal: Identify the bucket with fake coins.

N

  • • •   

Nyquist: → Weigh a coin from each bucket → N numbers → Compression → Bucket # / 1 number

Compressed Sensing: → Weigh a **linear** combination of coins from **all** buckets → Bucket # / 1 number

Uniqueness of Sparse Representations

- How many samples are needed to ensure uniqueness?
- Suppose there are two K-sparse vectors x_1 and x_2 satisfying

$$y = Ax_1 = Ax_2$$

- Then $A(x_1 - x_2) = 0$
- In the worst case $z = x_1 - x_2$ is $2K$ sparse
- Require that there is no z with $2K$ non-zero elements in $\mathcal{N}(A)$
- Every $2K$ columns of $A_{m \times n}$ must be linearly independent $\Rightarrow m \geq 2k$

Problem: Condition hard to verify

Coherence

Donoho *et al.*, '01
Tropp, '04

- The coherence of A is defined by (assuming normalized columns)

$$\mu = \max_{i \neq j} | < a_i, a_j > |$$

- When $n \gg m, \quad \frac{1}{\sqrt{m}} \leq \mu \leq 1$

- Uniqueness of *y*=*Ax* can be expressed in terms of μ as

$$k < \frac{1}{2}(1 + \frac{1}{\mu})$$

- Under same condition we will see that efficient recovery is possible as well

Restricted Isometry Property (RIP)

Candès and Tao, '05

- When noise is present uniqueness cannot be guaranteed
- Would like to ensure stability
- Can be guaranteed using RIP
- A has RIP of order δ if

$$(1-\delta)\|x\|^2 \leq \|Ax\|^2 \leq (1+\delta)\|x\|^2$$

  for any k–sparse vector x
- In this case A is an approximate isometry
- If A has unit-columns and coherence μ then it has the RIP with

$$\delta = k\mu$$

Recovery of Sparse Vectors

- Reconstruction: Find the sparsest and consistent x

(Requires $m = 2K$) $\min_x \|x\|_0$ s.t. $y = Ax$ NP-Hard !!

Alternative recovery algorithms (Polynomial-time):

- Basis pursuit $\min_x \|x\|_1$ s.t. $y = Ax$ (Requires $m = O(K \log(N/K))$)

  Convex and tractable — Donoho, '06; Candès *et al.*, '06

  RIP-$\delta_{2K} < \sqrt{2} - 1 \rightarrow$ exact recovery — Candès, '08

  or coherence guarantee $K < \frac{1}{2}\left(1 + \frac{1}{\mu}\right)$ — Donoho and Elad, '03

- Greedy algorithms

  OMP, FOCUSS, etc.
  OMP coherence guarantee $K < \frac{1}{2}\left(1 + \frac{1}{\mu}\right)$ — Tropp, Elad, Cotter *et* al., Chen *et* al., and many others…

Greedy Methods: Matching Pursuit

- Essential algorithm:

Mallat and Zhang, '93

  1) Choose the first "active" column (maximally correlated with y)

$$\arg\max_i \langle \mathbf{A}_i, \mathbf{y} \rangle \qquad\qquad S = \mathrm{supp}(\hat{\mathbf{x}}) \leftarrow i$$

  2) Subtract off to form a residual

$$\mathbf{y}' = \mathbf{y} - \sum_{i \in S} \langle \mathbf{A}_i, \mathbf{y} \rangle$$

  3) Repeat with $\mathbf{y}'$

- Very fast for small scale problems
- Not as accurate/robust for large signals in the presence of noise

Orthogonal MP:

Pati *et al.*, '93

- Improve residual computation

$$\mathbf{y}' = (\mathbf{I} - \mathcal{P}_S)\mathbf{y} = \mathbf{y} - \mathbf{A}\mathbf{A}_S^{\dagger}\mathbf{y}$$

Recovery In the Presence of Noise

$$y = Ax + w$$

- ℓ_1-relaxation techniques (convex optimization problems)
  - Basis pursuit denoising (BPDN) / Lasso:

$$\min_x \|x\|_1 \quad \text{s.t.} \quad \|y - Ax\|_2^2 \leq \eta \quad \text{or} \quad \min_x \|x\|_1 + \lambda \|y - Ax\|_2^2$$

Tibshirani '96
Chen *et* al., '98

  - Dantzig selector: $\min_x \|x\|_1 \quad \text{s.t.} \quad \|A^T(y - Ax)\|_\infty^2 \leq \eta$

Candès and Tao, '07

- Greedy approaches: stop when data error is on the order of the noise

Recovery Gurantees

$$y = Ax + w$$

Common settings:

- Random sensing matrix *A*, random noise $w \sim N(0, \sigma^2 I)$
  - RIP (and similar properties) can be approximated w.h.p.
  - RIP-based guarantees for Dantzig selector and BPDN: $\|x - \hat{x}\|_2^2 \leq C_0 K \sigma^2 \log N$ assuming RIP

Candès and Tao, '07
Bicket *et al., 09*

- Deterministic *A* and *x*, random $w \sim N(0, \sigma^2 I)$
  - RIP typically unknown, coherence must be used
  - Coherence-based results for BPDN, OMP, thresholding: $\|x - \hat{x}\|_2^2 \leq C_0 K \sigma^2 \log N$ assuming low μ

Ben-Haim, Eldar and Elad, '10

- Deterministic "adversarial" noise w: $\|w\|_2^2 \leq \epsilon^2$
  - Guarantees on order of $\|x - \hat{x}\|_2^2 \sim \epsilon^2$

Donoho *et* al., '06

The Sensing Matrix A

- Random IID matrices ensure recovery with high probability for sub-Gaussian distributions (Gaussian, Rademacher , Bernoulli, bounded RVs …) when $m = O(K \log(N/K))$

Donoho, '06

- Random partial Fourier matrices (or more generally unitary matrices) also ensure recovery with a slightly higher number of measurements

Candès *et al.*, '06

- Some structured matrices work as well such as a Vandermonde matrix

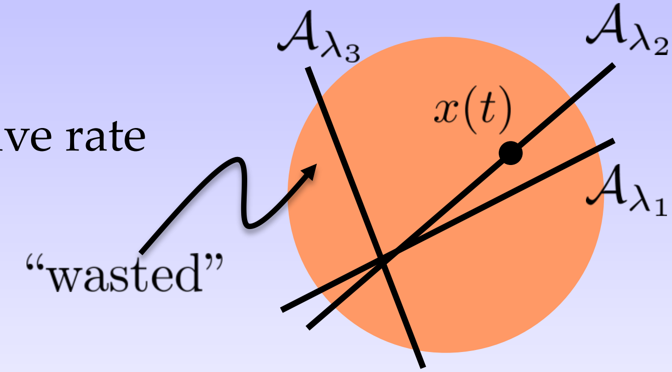
Tutorials on Compressed Sensing:

- R. G. Baraniuk, "Compressive sensing," IEEE Signal Processing Mag., 24(4), 118–124, July 2007.
- E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," IEEE Sig. Proc. Mag., 25(3), 21–30, Mar. 2008.
- M. Duarte and Y. C. Eldar, "Structured Compressed Sensing: From Theory to Applications," *TSP*.
- Y. C. Eldar and G. Kutyniok, "Compressed Sensing: Theory and Applications," Cambridge Press.

Sub-Nyquist in a Union

$$x(t) \in \mathcal{U} \qquad \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda$$

- Imposing subspace model $x(t) \in \Sigma$ is inefficient, $f_{\max}$ **problems**
  - High-sampling rate
  - Analog bandwidth issues
  - Load on the digital processing due to the excessive rate



Sub-Nyquist in a Union

$$x(t) \in \mathcal{U} \qquad \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda$$

- Imposing subspace model $x(t) \in \Sigma$ is inefficient, $f_{\max}$ **problems**

- Generalized sampling theory for unions?

  Still developing…

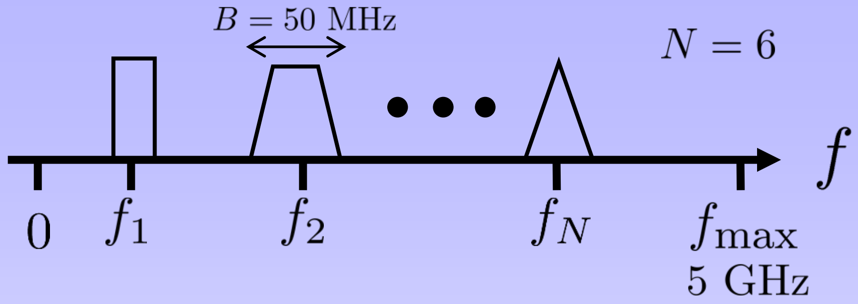
- Apply CS on discretized analog models?

…at the price of **model sensitivity, high computational loads, and loss of resolution**

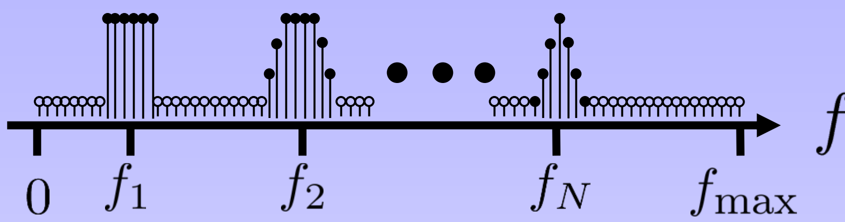
Rule of thumb: 1 MHz Nyquist = CS with 1 Million unknowns !

Multiband: Discretization ?

- Instead of **analog multiband**:



- Work with **discrete multi-tone**:



Advantages:

- Model size:

$$\mathbf{\Phi} = N \times \frac{f_{\max}}{B} \quad \approx 40 \times 200$$

Proportional to **actual** bandwidth

Problems:

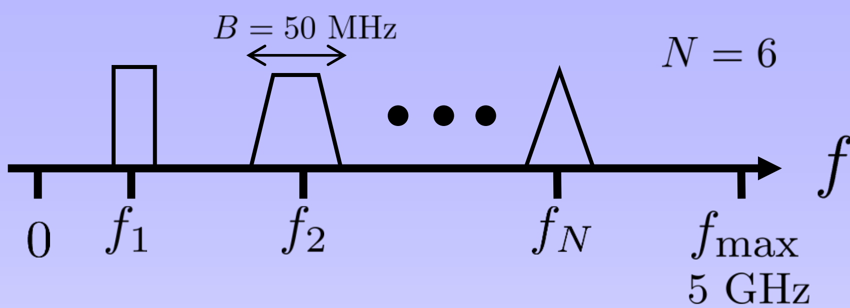
$$\mathbf{\Phi} \quad \approx 10^7 \times 10^{10}$$

Proportional to **Nyquist** rate

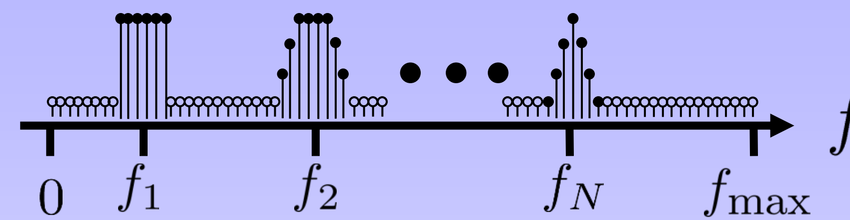
Mishali, Eldar and Elron, '10

Multiband: Discretization ?

- Instead of **analog multiband**:



- Work with **discrete multi-tone**:



Advantages:

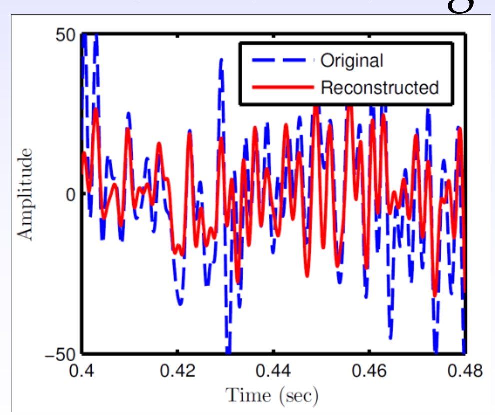
- Model size: $\mathbf{\Phi} \approx 40 \times 200$
- Sensitivity:

Negligible
(for a slight rate increase)

Problems:

$\mathbf{\Phi} \approx 10^7 \times 10^{10}$ **huge-scale**

Cannot avoid grid mismatch



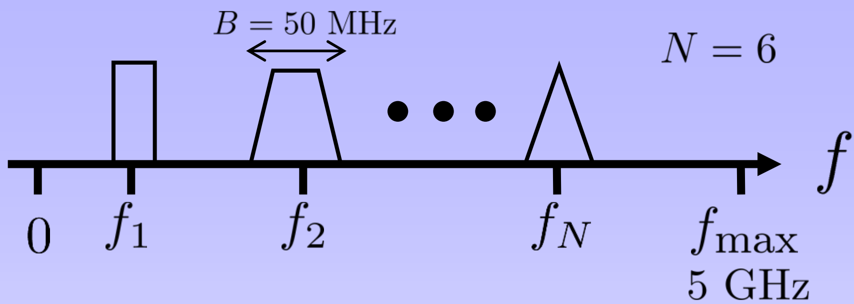
0.005% grid mismatch

$$\frac{\|f(t) - \hat{f}(t)\|^2}{\|f(t)\|^2} = 37\%$$

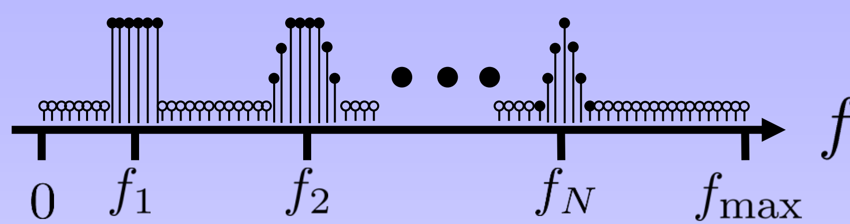
Mishali, Eldar and Elron, '10

Multiband: Discretization ?

- Instead of **analog multiband**:



- Work with **discrete multi-tone**:



Advantages:

Problems:

- Model size: $\boldsymbol{\Phi} \approx 40 \times 200$

$\boldsymbol{\Phi} \approx 10^7 \times 10^{10}$

- Sensitivity: Negligible

System "grid" must **match the unknown** signal tones grid

- Computational load (100 MHz processer):

≈ 200

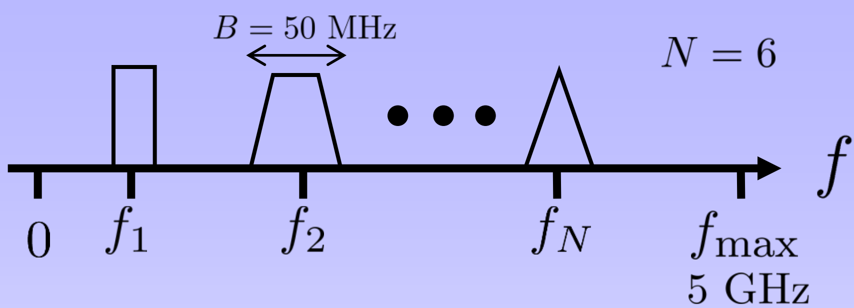
Realtime processing

$\approx 10^9$ MIPS

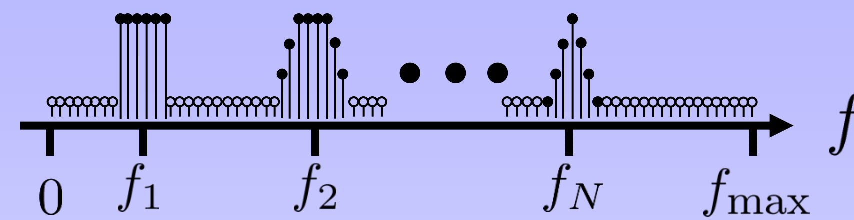
Mishali, Eldar and Elron, '10

Multiband: Discretization ?

- Instead of **analog multiband**:



- Work with **discrete multi-tone**:



Advantages:

- Model size: $\mathbf{\Phi} \approx 40 \times 200$
- Sensitivity: Negligible
- Computational load (100 MHz processer):

≈ 200

Realtime processing

Problems:

$\mathbf{\Phi} \approx 10^7 \times 10^{10}$

~~Analog Discretization ?~~

ust r

gnal id

$\approx 10^9$ MIPS

Mishali, Eldar and Elron, '10

Discrete CS Radar

- A discrete version of the channel is being estimated

Bajwa, Gedalyahu and Eldar, '11

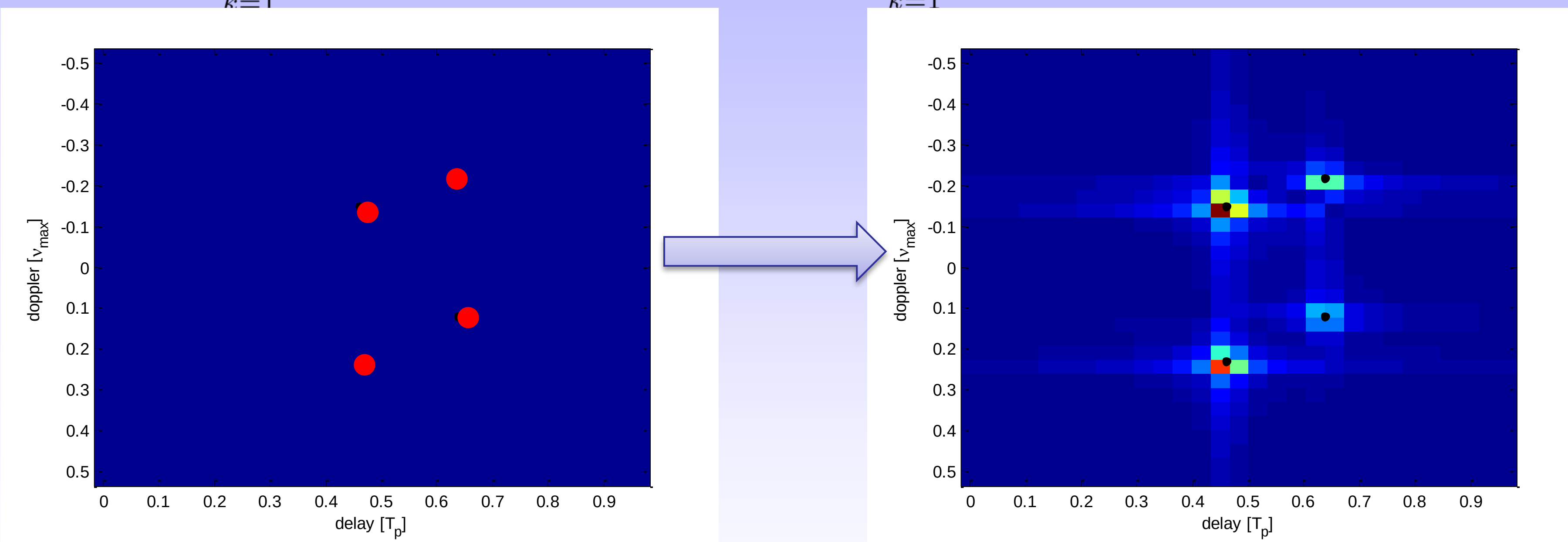
- Leakage effect → fake targets

Real channel

$$C(\tau,\nu)=\sum_{k=1}^{K}\alpha_k\delta(\tau-\tau_k)\delta(\nu-\nu_k)$$

Discretized channel

$$C(\ell,m)=\sum_{k=1}^{K}\alpha_k e^{j\pi(m-\mathcal{T}\nu_k)}\mathrm{sinc}(m-\mathcal{T}\nu_k)\mathrm{sinc}(\ell-\mathcal{W}\tau_k)$$



- Limited resolution to $1/\mathcal{W},\ 1/\mathcal{T}$
- Sampling process in hardware is unclear
- Digital processing is complex and expensive

ADCs: Why Not Standard CS?

- CS is for finite dimensional models *($y=Ax$)*
- Loss in resolution when discretizing
- Sensitivity to grid, analog bandwidth issues
- Is not able to exploit structure in analog signals
- Results in large computation on the digital side
- Samples do not typically interface with standard processing methods

More elaborate signal models needed that exploit structure to reduce sampling and processing rates

Sub-Nyquist in a Union

$$x(t) \in \mathcal{U} \qquad \mathcal{U} = \bigcup_{\lambda \in \Lambda} \mathcal{A}_\lambda$$

- Imposing subspace model $x(t) \in \Sigma$ is inefficient, $f_{\max}$ **problems**

- Generalized sampling theory for unions?

  Still developing…

- Apply CS on discretized analog models?

  Discretization issues…

Must combine ideas from Sampling theory and CS recovery algorithms

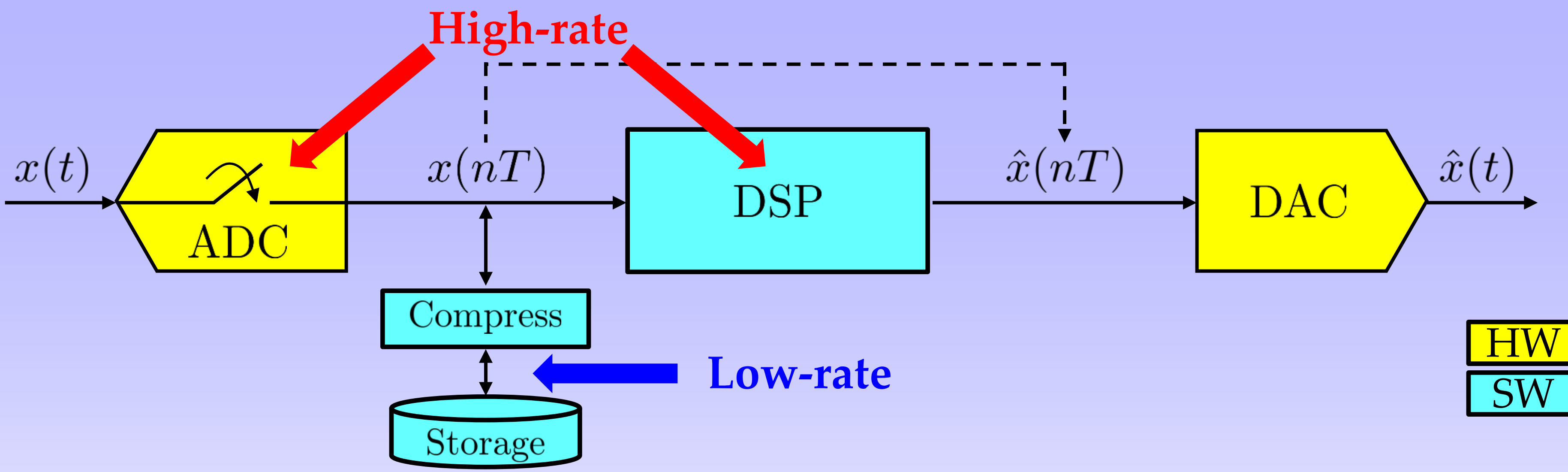
– Part 4 –
Xampling

Functional approach to sub-Nyquist in a Union

- CS+Sampling = Xampling
- X prefix for compression, e.g. DivX

→ Outline

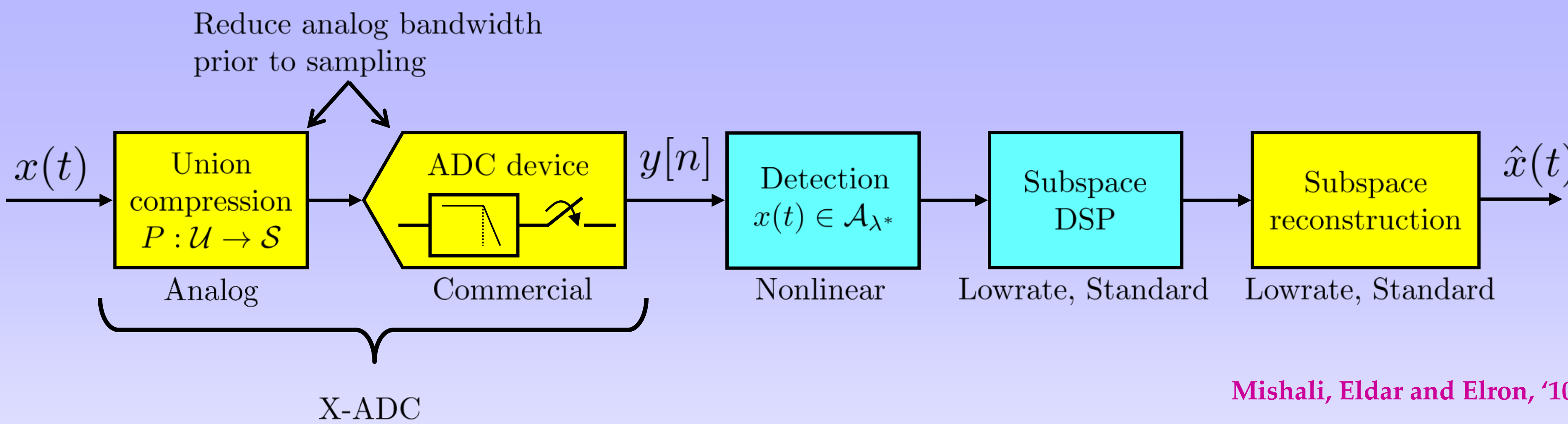
Standard DSP Systems



- Sampling and processing at high rates = Nyquist of $x(t)$
- After compression, data has low rate
- Standard DSP software expects Nyquist-rate samples
  rely on invariant properties $x(t) \leftrightarrow x(nT)$
  (enables digital filtering / digital estimation for example)

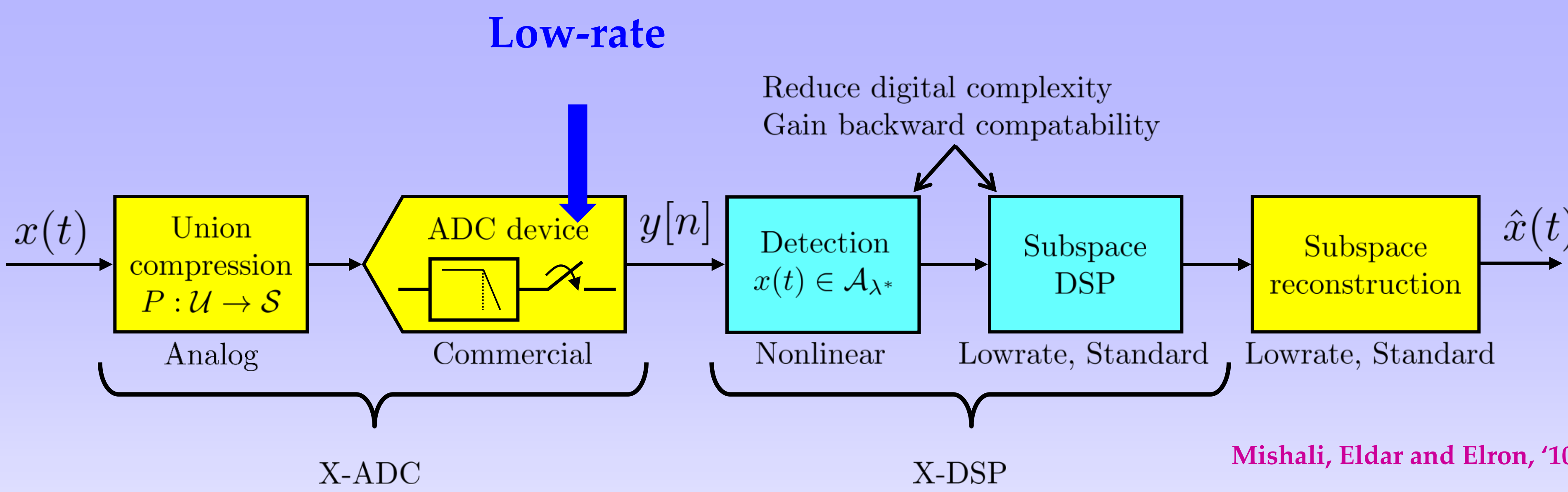
Move compression to hardware before ADC !

Xampling – Architecture



Mishali, Eldar and Elron, '10

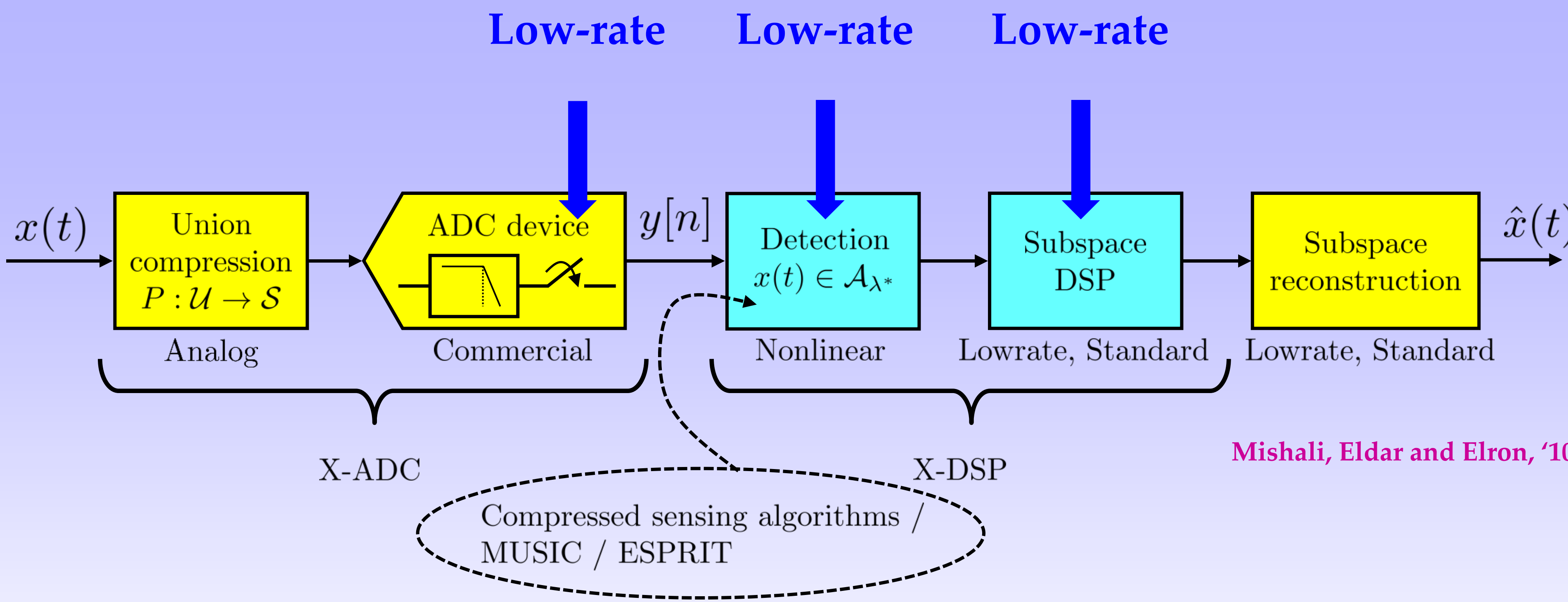
Xampling – Architecture



Mishali, Eldar and Elron, '10

- Functional architecture: Both sampling and processing at low rate
- $y[n] \neq x(nT)$ → Detection block outputs lowrate data that DSP can handle
- Built bottom-up: based on practical and pragmatic considerations

Xampling – Architecture



Mishali, Eldar and Elron, '10

- Functional architecture: Both sampling and processing at low rate
- $y[n] \neq x(nT)$ → Detection block outputs lowrate data that DSP can handle
- Built bottom-up: based on practical and pragmatic considerations

Xampling: Main Idea

Principle #1 (X-ADC):

- 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

Union
compression
$P : \mathcal{U} \to \mathcal{S}$

Analog

New hardware design ideas

Principle #2 (X-DSP):

- Identify subspaces involved (*e.g.*, using CS)
- Recover using standard sampling results

Detection
$x(t) \in \mathcal{A}_{\lambda^*}$

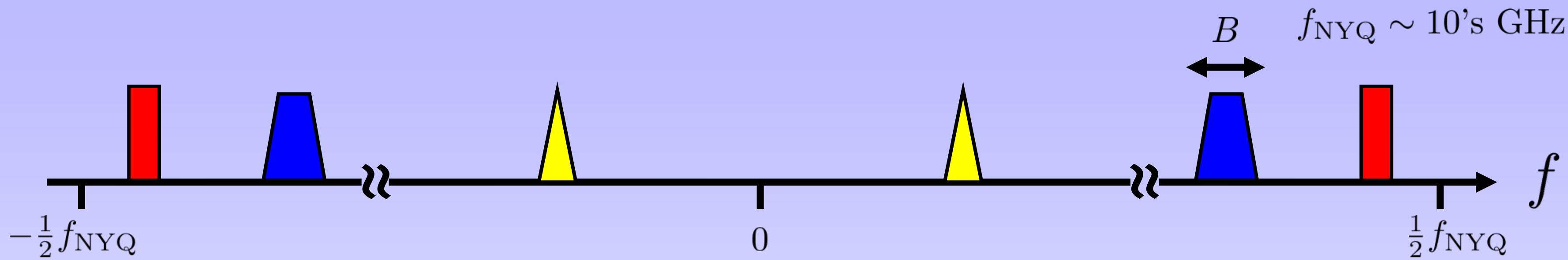
Nonlinear

New DSP algorithms

Xampling Systems

- Modulated wideband converter — Mishali and Eldar, '07-'09
- Periodic nonuniform sampling (fully-blind) — Mishali and Eldar, '07-'09
- Sparse shift-invariant framework — Eldar, '09
- Finite rate of innovation sampling — Vetterli *et al.*, '02-'07; Dragotti *et al.*, '02-'07; Gedalyahu, Tur and Eldar, '10-'11
- Random demodulation — Tropp *et al.*, '09

Multiband Union



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

$\mathcal{M} = \{\, x(t) \,|\,$ no more than N bands, max width B, bandlimited to $[-\frac{1}{2} f_{\mathrm{NYQ}}, +\frac{1}{2} f_{\mathrm{NYQ}})\}$

Mishali and Eldar, '07

Optimal Blind Sampling Rate

Theorem (known spectral support)

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

$$\underbrace{D^-(R)}_{\text{Average sampling rate}} \geq c = \operatorname{meas}(\mathcal{F})$$

Landau, '67

Theorem (unknown spectral support)

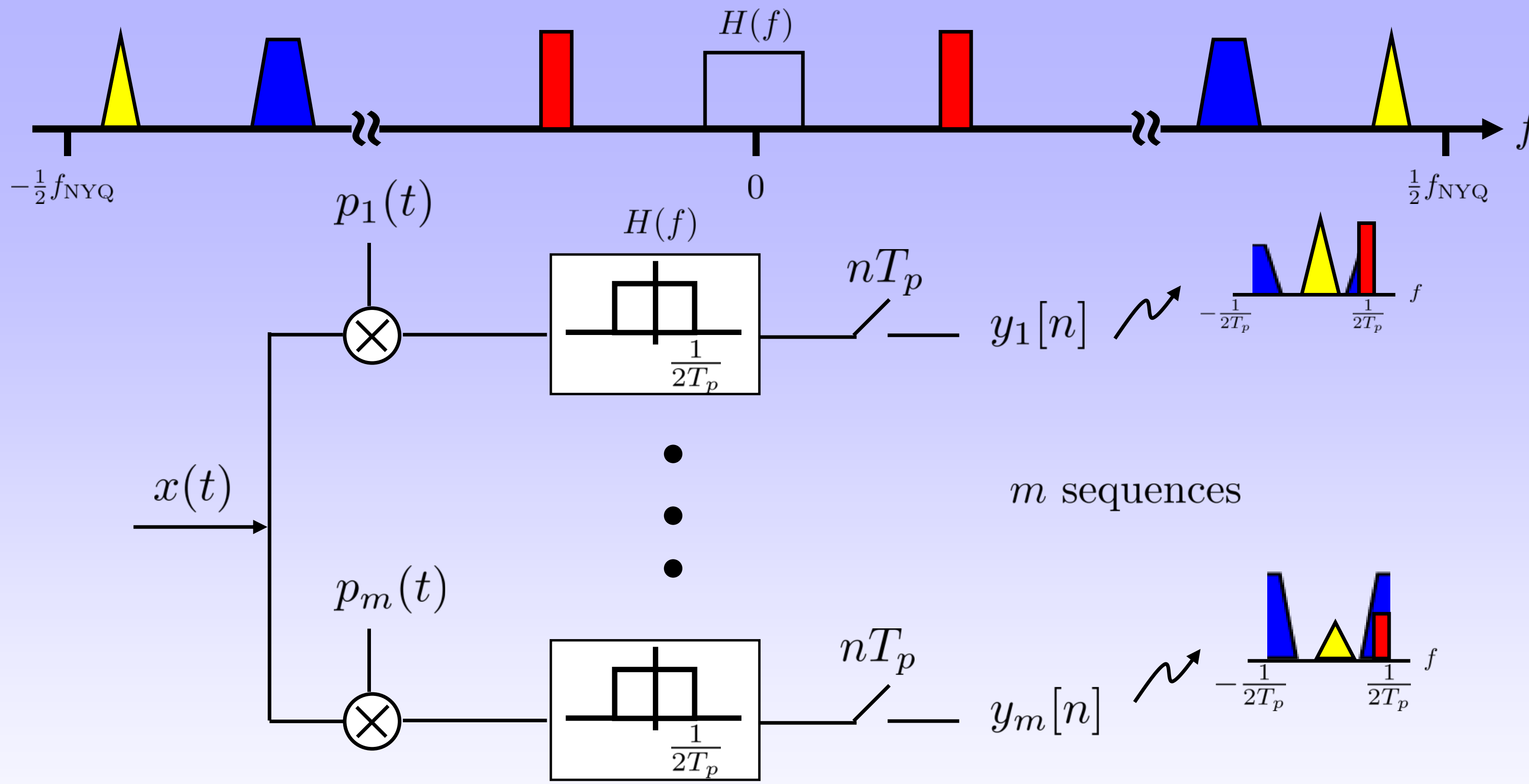
Let R be a sampling set for $\mathcal{N}_c = \{\mathcal{B}_{\mathcal{F}} \;:\; \operatorname{meas}(\mathcal{F}) \leq c\}$.
Then,

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

Mishali and Eldar, '07

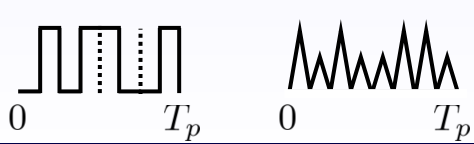
1. The minimal rate is doubled
2. N bands, individual widths $\leq B$, requires at least $2NB$ samples/sec

The Modulated Wideband Converter



T_p−periodic $p_i(t)$ gives the desired aliasing effect

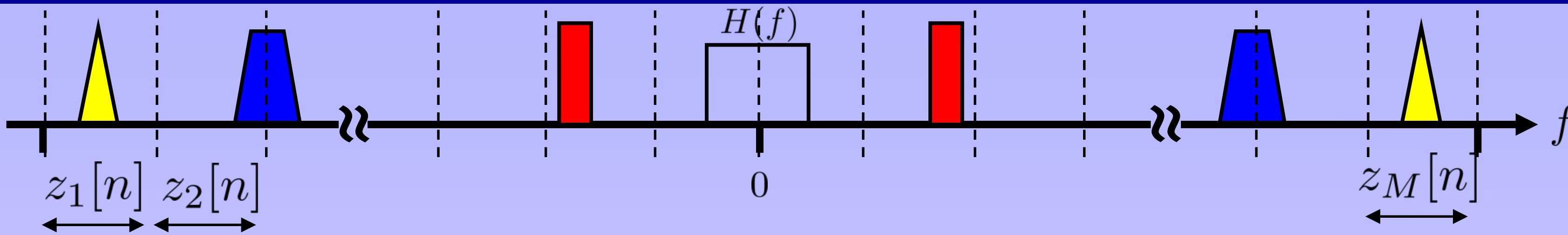
$$p_i(t) = \sum_{l=-\infty}^{\infty} c_{il} e^{j\frac{2\pi}{T_p}lt}$$



and many more…

Mishali and Eldar, '09

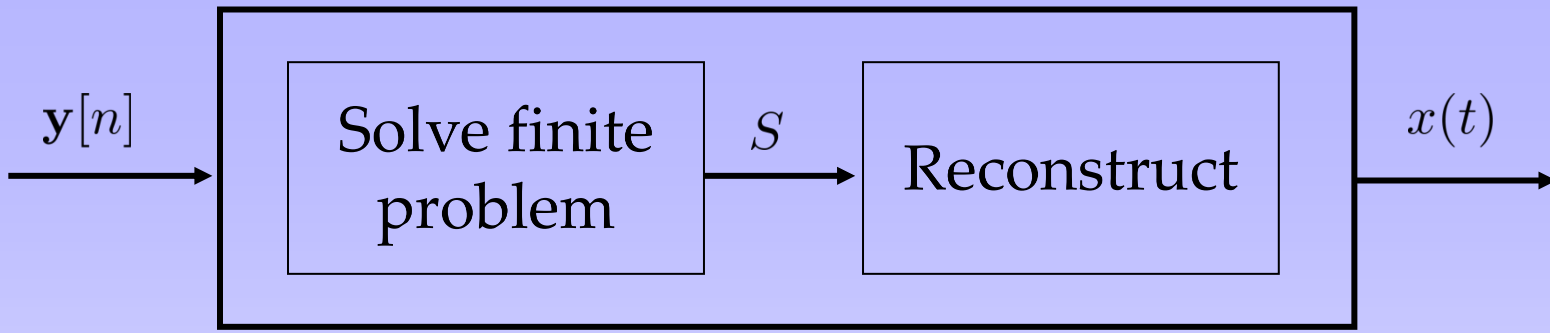
Recovery From Xamples



$$\begin{pmatrix} y_1[n] \\ \vdots \\ y_m[n] \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{il} = c_{il} \\ \\ m \times M \end{pmatrix} \begin{pmatrix} z_1[n] \\ z_2[n] \\ \vdots \\ \\ z_M[n] \end{pmatrix}$$

- Cannot invert a fat matrix!
- Spectrum sparsity: Most of the $z_i[n]$ are identically zero
- For each n we have a small size CS problem
- Problem: CS algorithms for each n → many computations

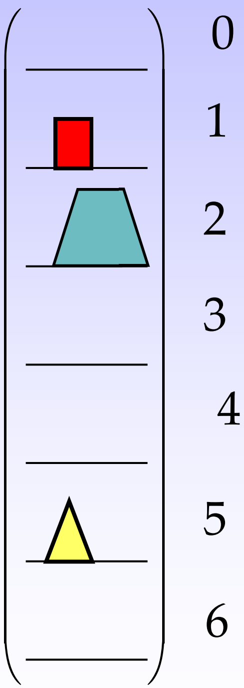
Reconstruction Approach



S = non-zero rows

CTF
(Support recovery)

$\mathbf{z}[n] =$



$S = \{1, 2, 5\}$

Continuous $\quad\quad$ **Finite**

$$\mathbf{y}[n] = \mathbf{A}\mathbf{z}[n], \quad n \in Z \quad \longrightarrow \quad \mathbf{V} = \mathbf{A}\mathbf{U}$$

The matrix V is any basis for the span of $\mathbf{y}[n]$

Underlying Theory

$$\boxed{\mathbf{y}(\lambda) = \mathbf{A}\mathbf{z}(\lambda), \quad \lambda \in \Gamma}$$

$$\mathbf{y}(\Gamma) \longrightarrow \boxed{\begin{array}{c}\text{Construct a frame} \\ \mathbf{V} \text{ for } \mathbf{y}(\Gamma)\end{array}} \xrightarrow{\mathbf{V}} \boxed{\begin{array}{c}\text{Solve MMV} \\ \mathbf{V} = \mathbf{A}\mathbf{U}\end{array}} \xrightarrow{\mathrm{supp}(\bar{\mathbf{U}})}$$

Theorem [Exact Support Recovery, CTF]

Let $\bar{\mathbf{z}}(\Gamma)$ be a k-sparse solution set. If

$$\sigma(\mathbf{A}) \geq 2k - (\mathrm{rank}(\mathbf{y}(\Gamma)) - 1)$$

then $\mathrm{supp}(\bar{\mathbf{z}}(\Gamma)) = \mathrm{supp}(\bar{\mathbf{U}})$.

Mishali and Eldar, ‘08

CTF = Continuous to Finite

Insight into CTF

$$\mathbf{y}[n] = \mathbf{A}\mathbf{z}[n]$$

Run CS recovery for each time-instance n — Poly.-time / $\mathbf{y}[n]$ — nonlinear

Computationally heavy

---

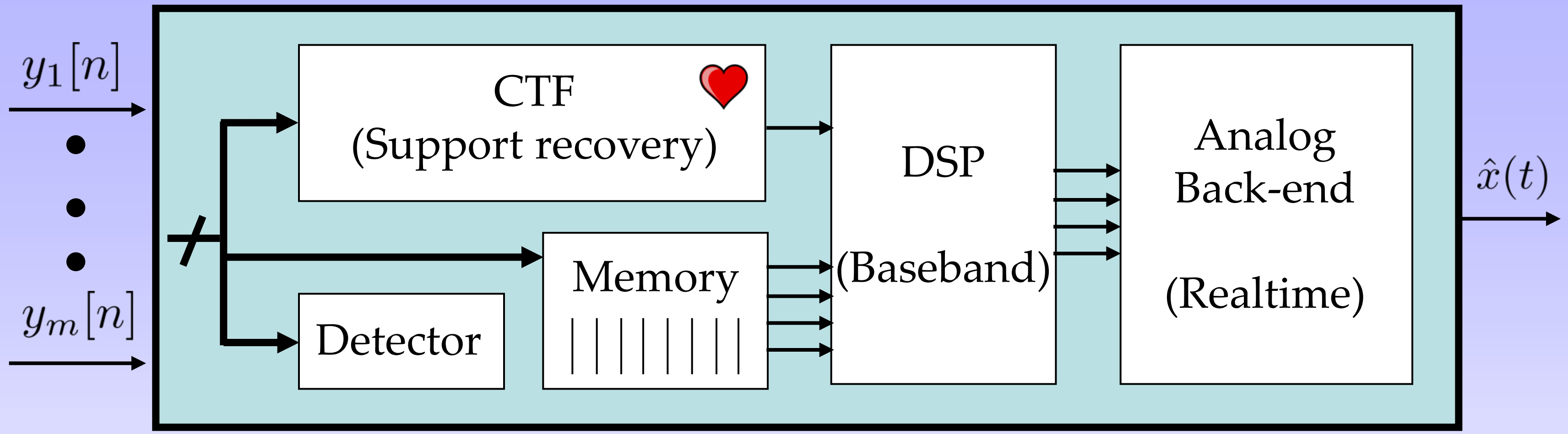
1. Construct frame $\mathbf{V}$ — $\mathcal{O}(k)$ snapshots — easy
2. Solve CS system $\mathbf{V} = \mathbf{A}\mathbf{U}$ — Poly.-time once — nonlinear
3. Apply $\mathbf{A}_S^{\dagger}$ on $\mathbf{y}[n]$ for each time-instance n — 1 matrix-vector mult. / $\mathbf{y}[n]$ — linear

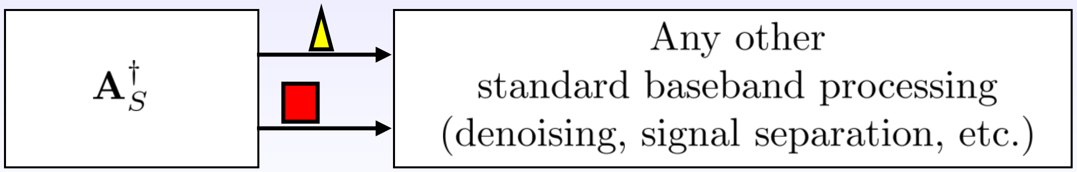
Computationally light

Reconstruction

Mishali and Eldar, '07-'10

High-level architecture



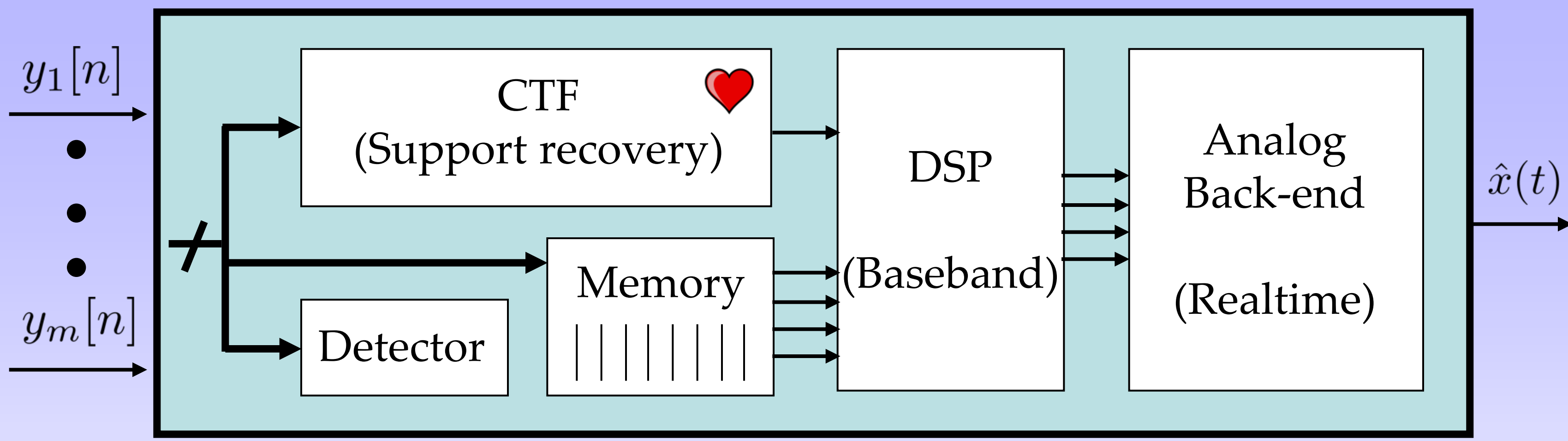


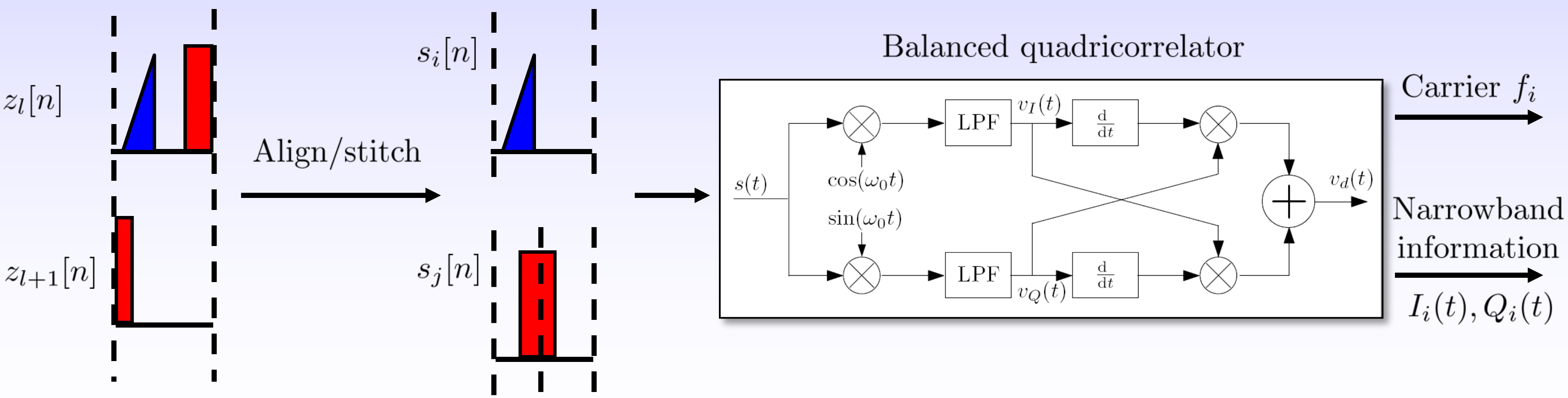
Recover any desired spectrum slice at baseband

Reconstruction

Mishali and Eldar, '07-'10

High-level architecture

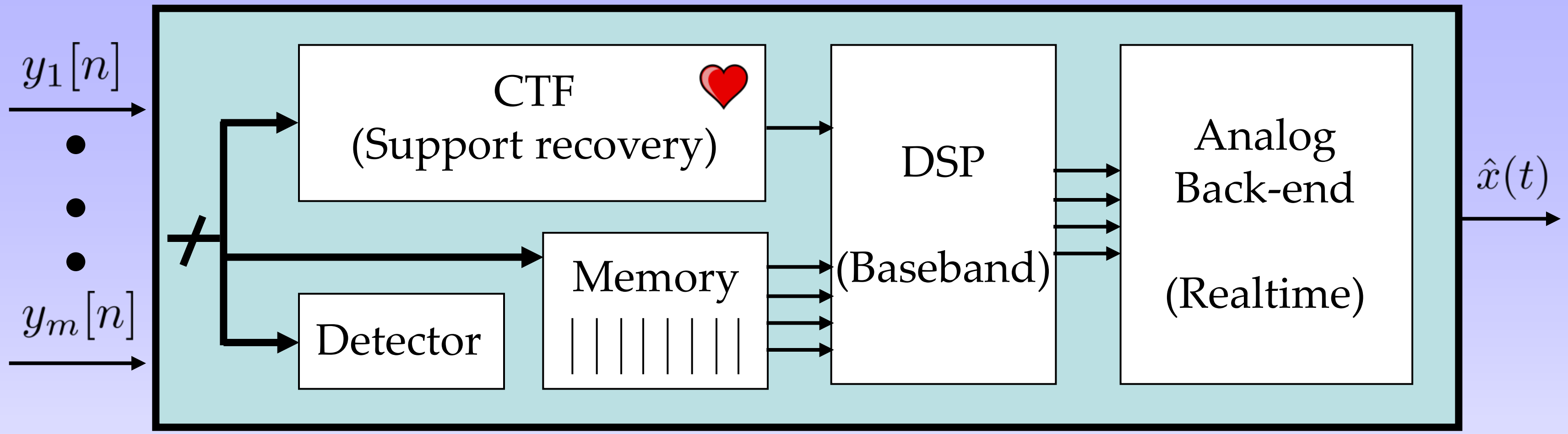




Reconstruction

Mishali and Eldar, '07-'10

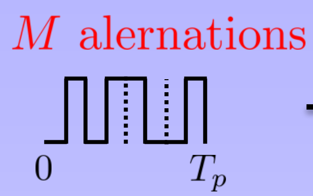
High-level architecture



Can reconstruct:

- The original analog input exactly $\hat{x}(t) = x(t)$ (without noise)
- Improve SNR for noisy inputs, due to rejection of out-of-band noise
- Any band of interest, modulated on any desired carrier frequency

Sign-Flipping Periodic Waveforms

- $p_i(t) =$ 

$\longrightarrow \mathbf{A} = \mathbf{SF} \begin{cases} \mathbf{S} = \text{rectangular (signs)} \\ \mathbf{F} = \text{square (DFT)} \end{cases}$

Theorem [Expected-RIP for MWC]

Periodic mixing with sign patterns gives $\mathbf{A}$ with ExRIP probability

$$p \geq 1 - \frac{(1-C_k)\rho_M\left(1+\boldsymbol{\alpha}(\mathbf{S}) - 2\boldsymbol{\beta}(\mathbf{S})\right) - (B_k - C_k)\rho_M\left(\boldsymbol{\gamma}(\mathbf{S}) - \boldsymbol{\beta}(S)\right) + C_k M \boldsymbol{\beta}(\mathbf{S}) - 1}{\delta_k^2}$$

TABLE II: ExRIP guarantees for different sign patterns

Family	Dimensions m	M	$2K$	Quality ×100 $\boldsymbol{\alpha}(\mathrm{S})$	$\boldsymbol{\beta}(\mathrm{S})$	$\boldsymbol{\gamma}(\mathrm{S})$	ExRIP prob. p Normal	Uniform
Maximal	80	511	24	1.438	0.196	0.408	0.932	0.931
Gold	80	511	24	1.255	0.198	0.199	0.939	0.939
Hadamard	80	512	24	1.250	1.094	1.238	0.000	0.000
Random1	80	511	24	1.439	0.198	0.202	0.927	0.927
Kasami	16	255	12	6.667	0.392	0.294	0.689	0.675
Random2	40	195	24	3.025	0.526	0.537	0.856	0.858

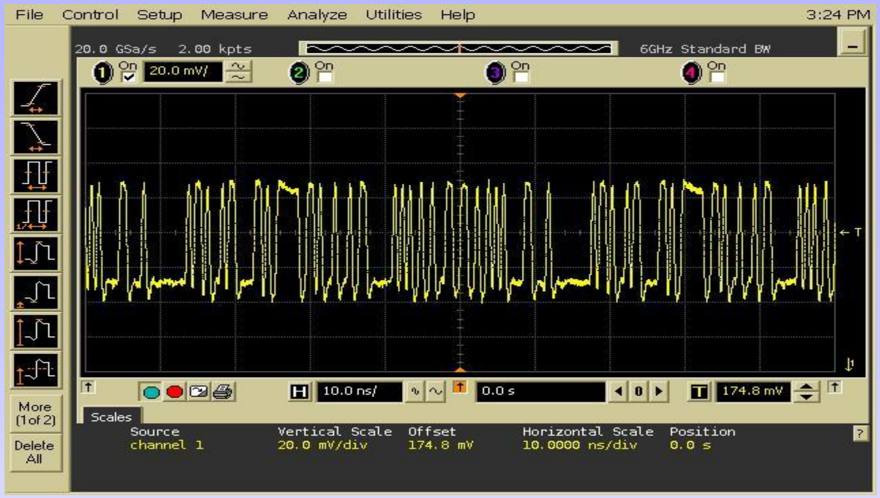
$\boldsymbol{\alpha}(\mathbf{S}) =$ correlations energy

$\boldsymbol{\beta}(\mathbf{S}) =$ auto/cross-correlations

$\boldsymbol{\gamma}(\mathbf{S}) =$ reverse-correlations

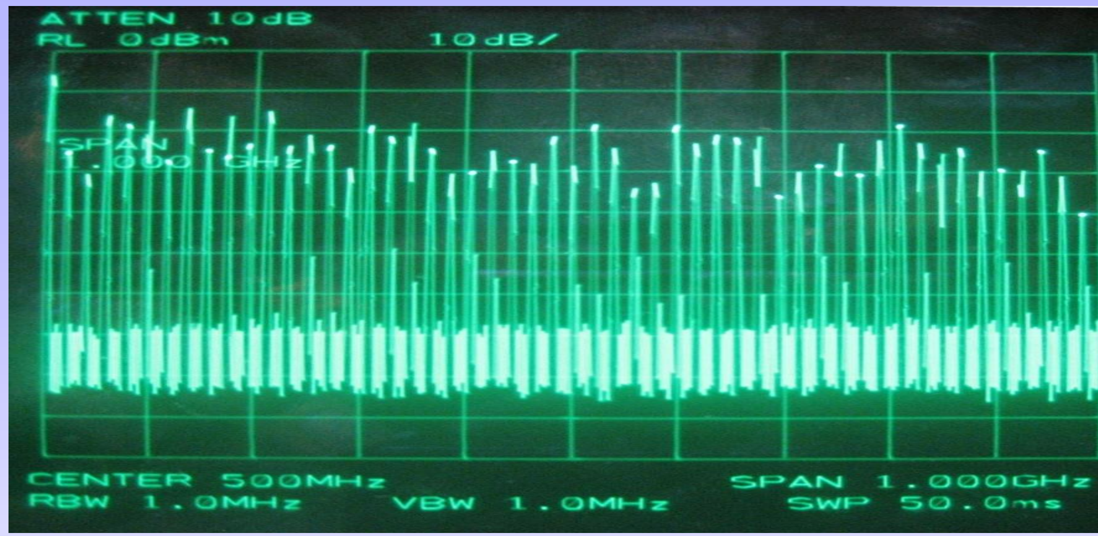
Mishali and Eldar, '09

Time Appearance of Mixing Waveforms



- Bad news: can't design nice sign patterns at GHz rates

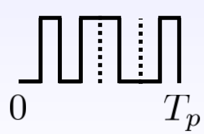
Time Appearance of Mixing Waveforms

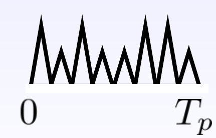


- Bad news: can't design nice sign patterns at GHz rates

- Good news: only the periodicity matters !

$$p_i(t) = \sum_{l=-\infty}^{\infty} c_{il} e^{j\frac{2\pi}{T_p} lt}$$

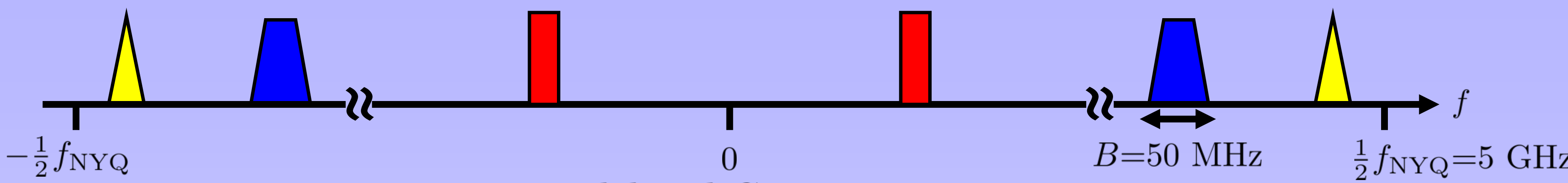




and many more...

- Competing approaches (pure CS) struggle with time appearance

Simulation



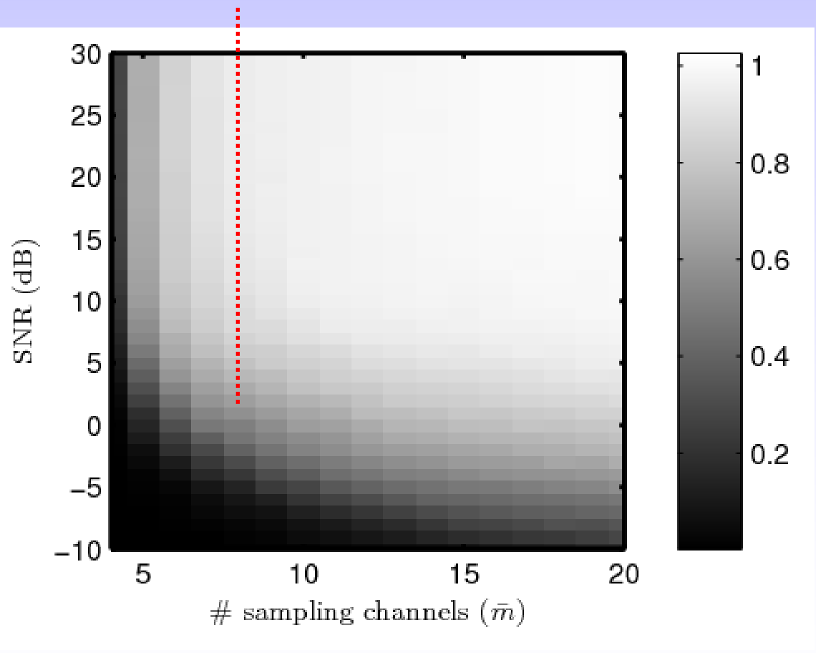
+ wideband Gaussian noise

Theory: P.R. requires 1.2 GHz (=$4NB$ with SBR4)

In practice: 99% recovery (out of 500 trials)
7 channels × 250 MHz each
=1.8 GHz (S-OMP algorithm)

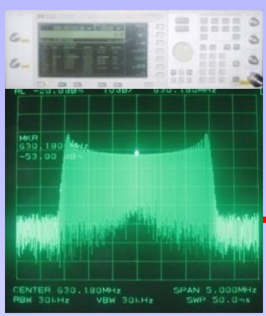
CTF observes the input for 2 μsecs only !

Can further reduce the system to
4 channels × 450 MHz (CTF with 10 μsecs)
1 channel × 1.8 GHz (CTF with 40 μsecs)



Sub-Nyquist Demonstration

Carrier frequencies are chosen to create overlayed aliasing at baseband

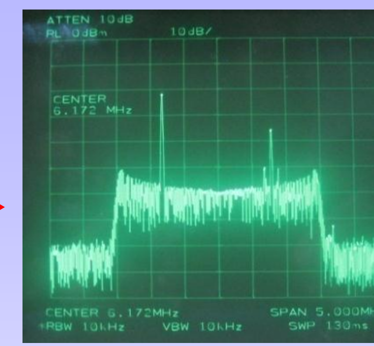


FM @ 631.2 MHz

AM @ 807.8 MHz

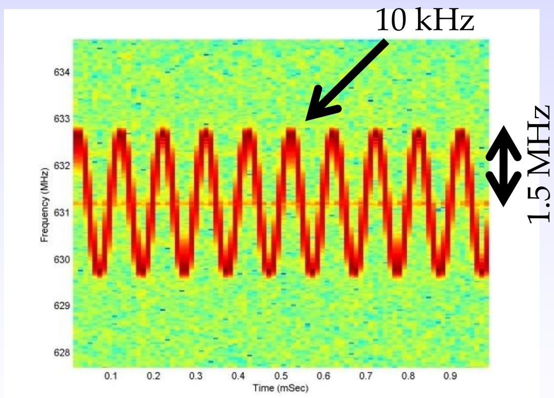
Sine @ 981.9 MHz

MWC prototype

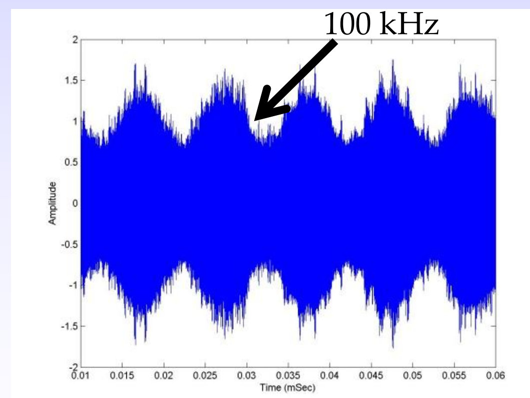


aliasing around 6.171 MHz

Reconstruction
(CTF)



FM @ 631.2 MHz



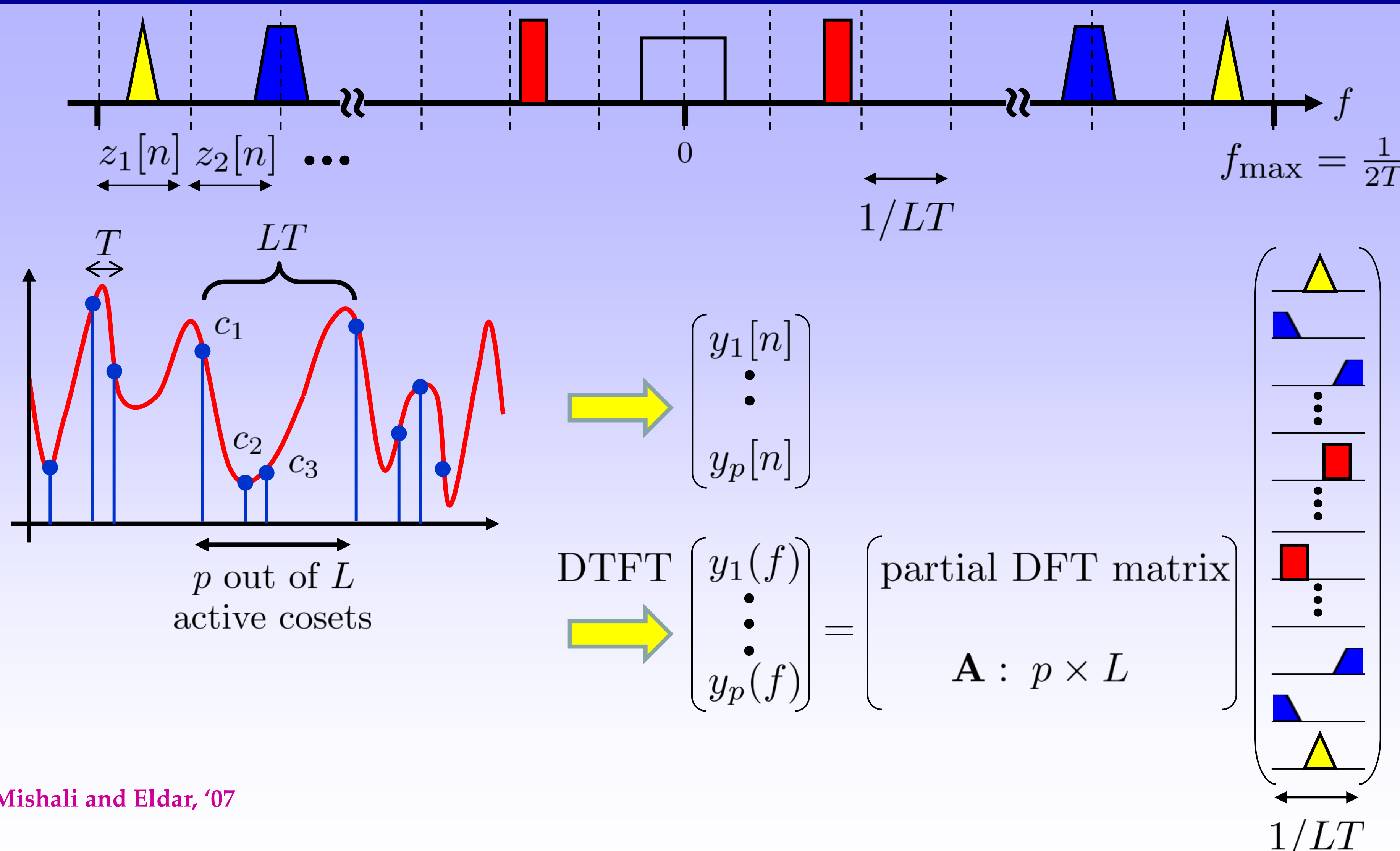
AM @ 807.8 MHz

Mishali et al., '10

Xampling Systems

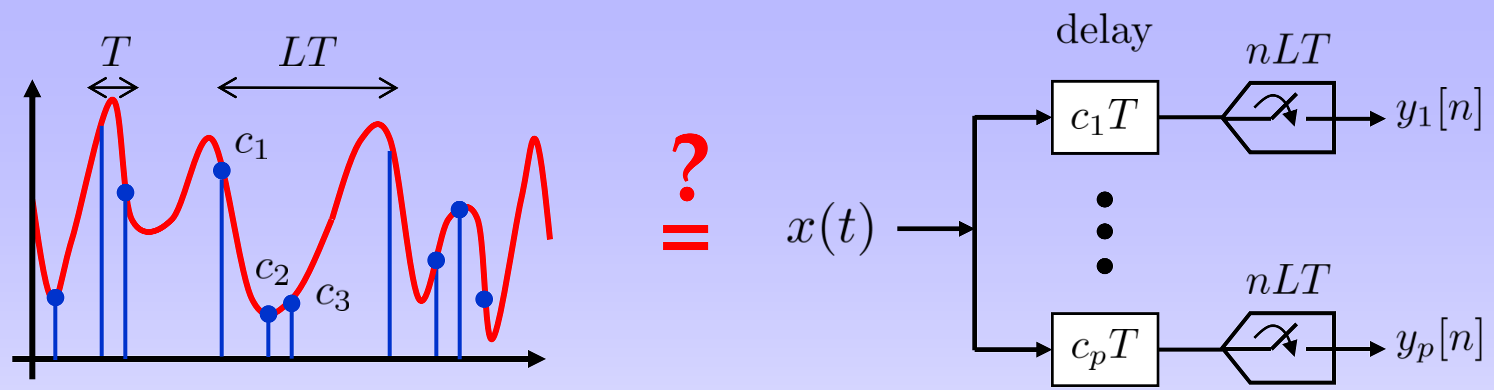
- Modulated wideband converter — Mishali and Eldar, '07-'09
- Periodic nonuniform sampling (fully-blind) — Mishali and Eldar, '07-'09
- Sparse shift-invariant framework — Eldar, '09
- Finite rate of innovation sampling — Vetterli *et al.*, '02-'07; Dragotti *et al.*, '02-'07; Gedalyahu, Tur and Eldar, '10-'11
- Random demodulation — Tropp *et al.*, '09

Fully-Blind PNS Approach



Mishali and Eldar, '07

Can Avoid RF Front-end ?

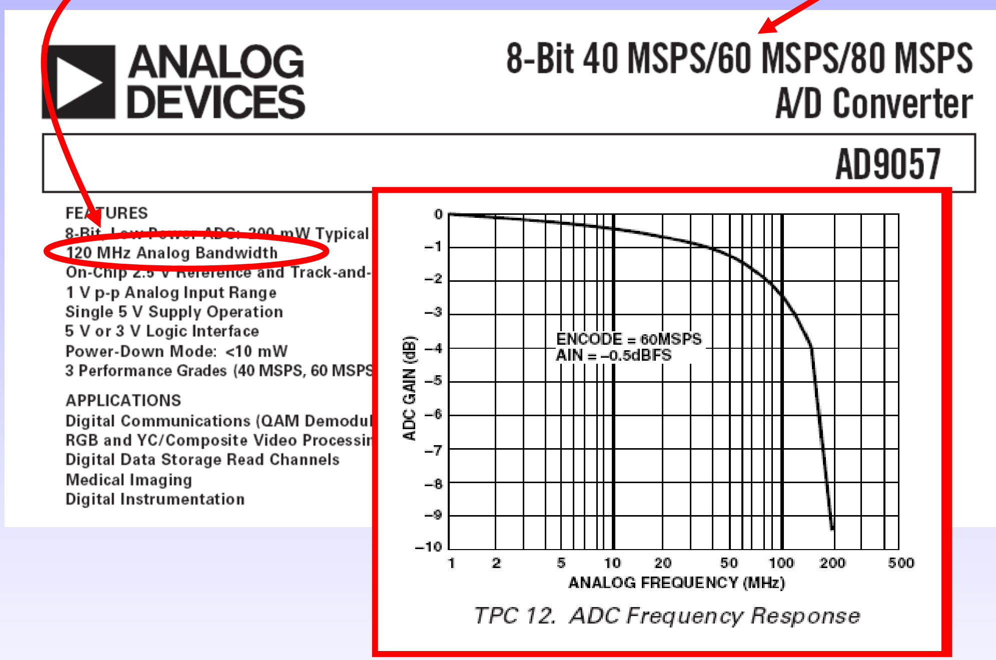


- YES ! If the input bandwidth is not too high…

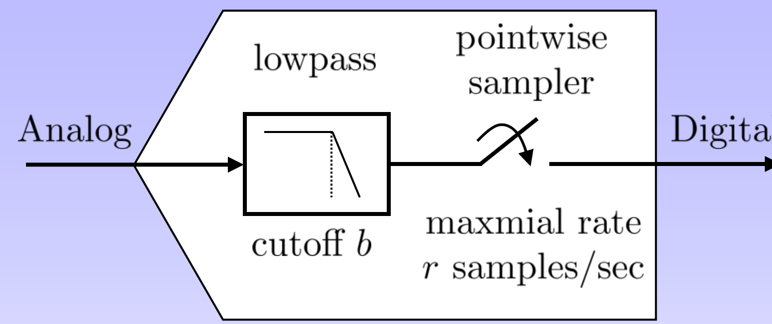
Practical ADC Devices

Analog bandwidth limitation b

Sampling rate r



TPC 12. ADC Frequency Response



In non-uniform sampling:

- Both T/H and mux operate at the Nyquist rate
- Digital processing and recovery requires interpolation to the high Nyquist grid
- Accurate time-delays ϕ_i are needed

Xampling Systems

- Modulated wideband converter — Mishali and Eldar, '07-'09
- Periodic nonuniform sampling (fully-blind) — Mishali and Eldar, '07-'09
- Sparse shift-invariant framework — Eldar, '09
- Finite rate of innovation sampling — Vetterli *et al.*, '02-'07; Dragotti *et al.*, '02-'07; Gedalyahu, Tur and Eldar, '10-'11
- Random demodulation — Tropp *et al.*, '09

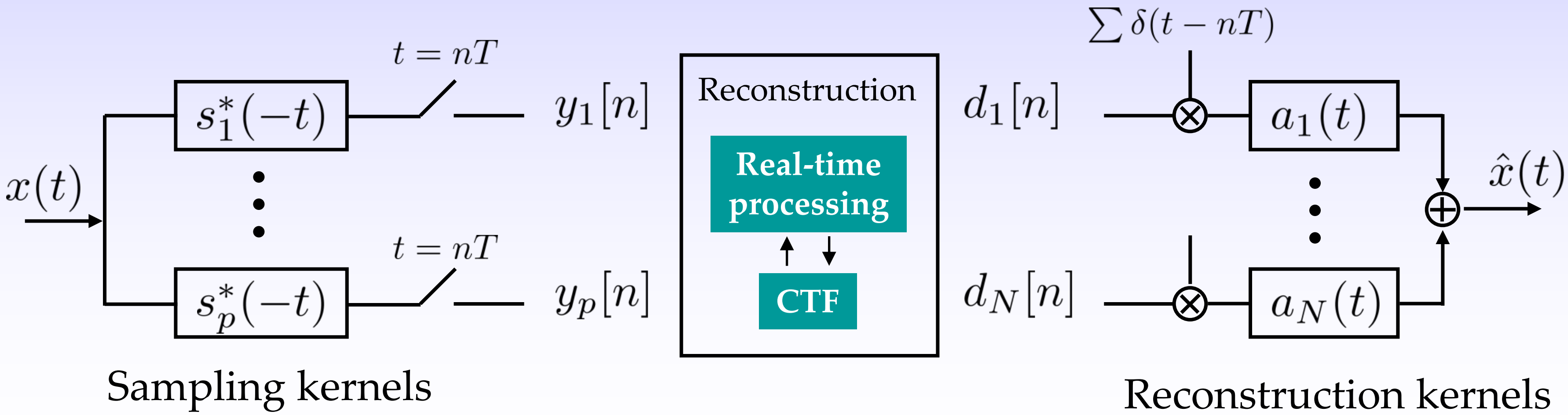
Sparse Shift-Invariant Framework

Eldar, '09

Sampling signals from a structured union of shift-invariant spaces (SI)

$$x(t) = \sum_{|l|=k} \sum_{n=-\infty}^{\infty} d_l[n] a_l(t-n)$$

There is no prior knowledge on the exact $|l| = k$ indices in the sum



Sparse Shift-Invariant Framework

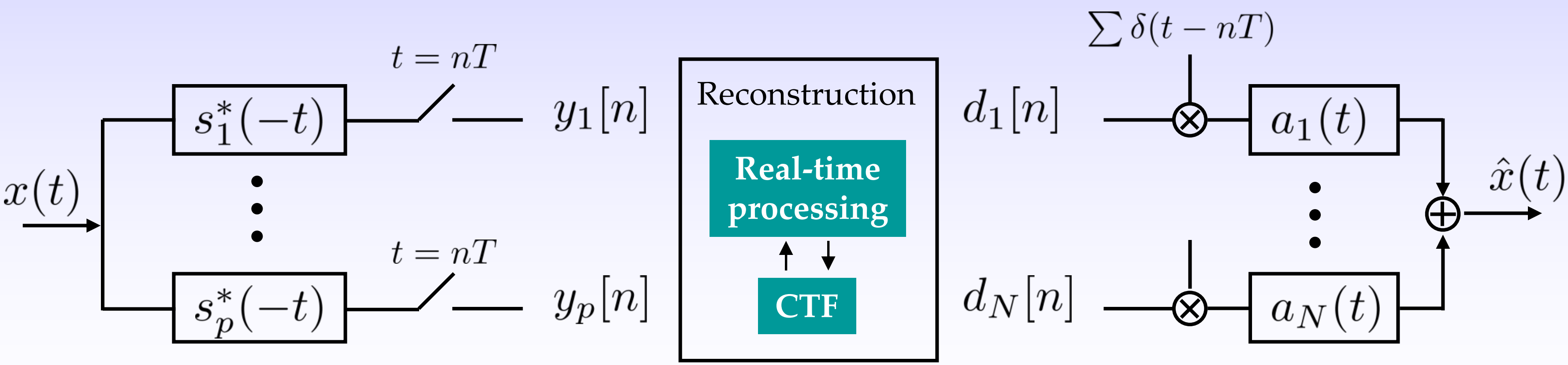
Eldar, '09

Sampling signals from a structured union of shift-invariant spaces (SI)

Theorem

$$\hat{x}(t) = x(t), \quad \text{if} \quad \mathbf{s}(\omega) = \mathbf{A}^*\underbrace{\mathbf{M}_{HA}^{-*}(e^{j\omega T})\mathbf{h}(\omega)}_{\text{Sampling theory}}$$

"Good" CS matrix



Sampling kernels

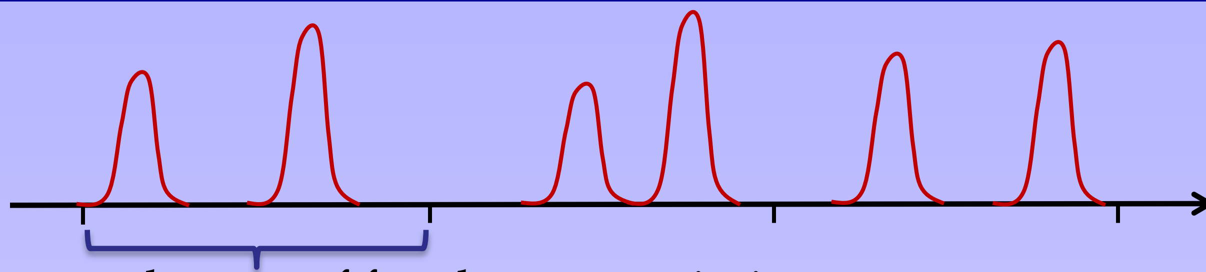
Reconstruction kernels

Xampling Systems

- Modulated wideband converter — Mishali and Eldar, '07-'09
- Periodic nonuniform sampling (fully-blind) — Mishali and Eldar, '07-'09
- Sparse shift-invariant framework — Eldar, '09
- Finite rate of innovation sampling — Vetterli *et al.*, '02-'07; Dragotti *et al.*, '02-'07; Gedalyahu, Tur and Eldar, '10-'11
- Random demodulation — Tropp *et al.*, '09

Pulse Streams

$$x(t) = \sum_{l \in \mathbb{Z}} a_l h(t - t_l)$$



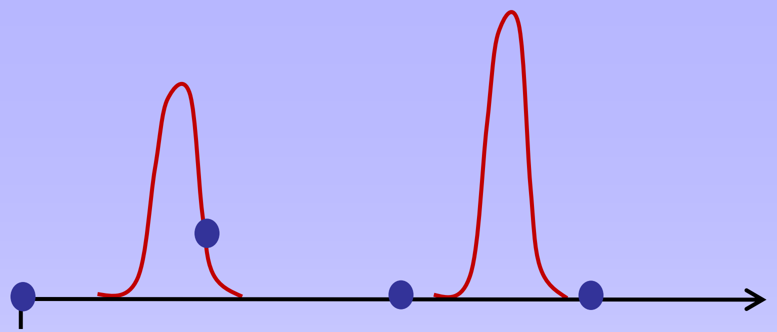
$2L$ degrees of freedom per unit time

- Delays and amplitudes are unknown

- Applications:
  - Communication
  - Radar
  - Bioimaging
  - Neuronal signals
- Special case of Finite Rate of Innovation (FRI) signals — **Vetterli *et al.*, '02**
- Minimal sampling rate – the rate of innovation: $\rho = \frac{2L}{T}$

Analog Sampling Stage

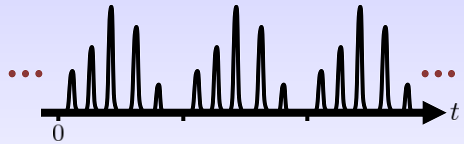
- Naïve attempt: direct sampling at low rate
- Most samples do not contain information!!



Sampling rate reduction requires proper design of the analog front-end

Special cases:

- Periodic pulse streams



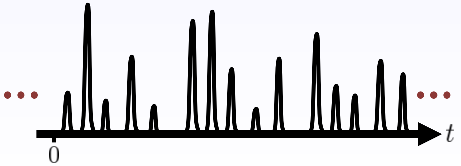
Vetterli *et al.*, '02-'05

- Finite

Dragotti *et al.*, '07-'10

Tur *et al.*, '10-'11

- Infinite pulse streams

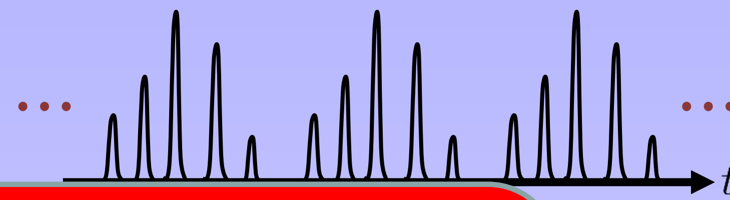


Gedalyahu *et al.*, '09

Periodic Pulse Streams

- Periodic FRI signal model:

$$x(t) = \sum \sum^{L} a_l h(t - t_\ell - k\tau), \quad t_\ell \in [0, \tau)$$



et al., '02-'05

The

- Si

Once the Fourier coefficients are known,
Standard solutions exist.
Challenge: <u>How can we obtain the coefficients?</u>

$l=1$

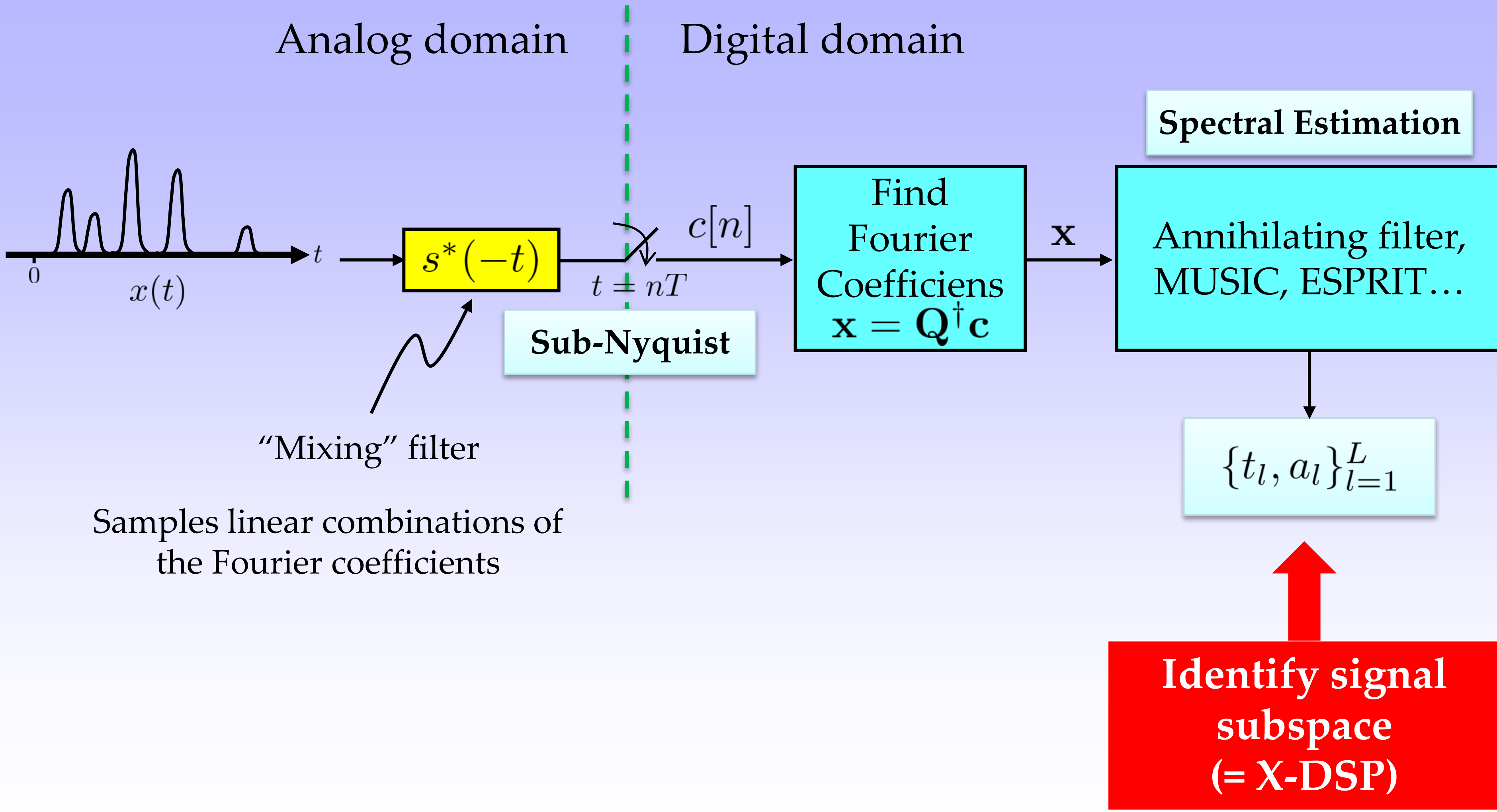
- Spectral estimation: sum of complex exponentials problem
- Solved using *2L* measurements
  - Methods: annihilating filter, MUSIC, ESPRIT

Schmidt, '86

Roy and Kailath, '89

Stoica and Moses, '97

General Approach



Find Fourier Coefficients

- Fourier series of a periodic input:

$$c[n] \longrightarrow \boxed{\begin{matrix}\text{Find}\\ \text{Fourier}\\ \text{Coefficiens}\\ \mathbf{x} = \mathbf{Q}^{\dagger}\mathbf{c}\end{matrix}} \xrightarrow{\ \mathbf{x}\ }$$

$$x(t) = \sum_{\ell=1}^{L} a_l h(t - t_\ell) \longrightarrow X[k] = H\left(\frac{2\pi k}{T}\right) \sum_{l=1}^{L} a_l e^{-j2\pi k t_l / T}$$

$$\mathbf{x} = [\cdots X[k] \cdots]^T \quad \textbf{Unknown}$$

- Sensing with lowpass:

$$c[n] = \langle s(t - nt), x(t) \rangle = \sum_k X[k] \int_{-\infty}^{\infty} e^{j2\pi kT/\tau} s^*(t - nT)\mathrm{dt}$$

$$= \sum_k X[k] e^{j2\pi knT/\tau} \underbrace{S^*\left(\frac{2\pi k}{\tau}\right)} = \sum_{k=-L}^{L} X[k] \underbrace{e^{j2\pi knT/\tau}}_{\mathbf{V}} \underbrace{S^*\left(\frac{2\pi k}{\tau}\right)}_{\text{diagonal } \mathbf{S}} \longrightarrow \mathbf{c} = \underbrace{\mathbf{VS}}_{\mathbf{Q}}\,\mathbf{x}$$

$S^*(\omega) = \mathrm{CTFT}\{\mathrm{s(t)}\}$

lowpass $\rightarrow \neq 0, -L \leq k \leq L$

$\mathbf{c} = [\cdots c[n] \cdots]^T$

Known measurements

Annihilating ``Filter''

- Goal: design a digital filter $A[k]$ with $z-$transform:

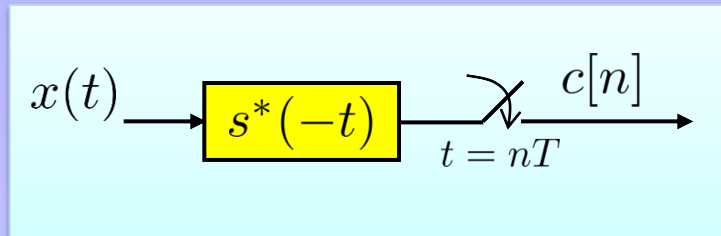
$$\mathbf{x} \longrightarrow \boxed{\{t_l, a_l\}_{l=1}^{L}}$$

$$A(z) = \sum_{k=0}^{L} A[k] z^{-k} = A[0] \prod_{l=1}^{L} \left(1 - e^{-j2\pi t_\ell/\tau} z^{-1}\right)$$

- $A[k]$ has zeros at the ``frequencies'' t_ℓ $\longrightarrow$ annihilates $X[k]$
- Filter coefficients can be computed from the measurements:

$$A[k] * X[k] = 0 \longrightarrow \begin{bmatrix} X[0] & X[-1] & \cdots & X[-L] \\ X[1] & X[0] & \cdots & X[-(L-1)] \\ \vdots & \vdots & \ddots & \vdots \\ X[L] & X[L-1] & \cdots & X[0] \end{bmatrix} \begin{pmatrix} A[0] \\ A[1] \\ \vdots \\ A[L] \end{pmatrix} = \mathbf{0}$$

X-ADC: Filter Choice

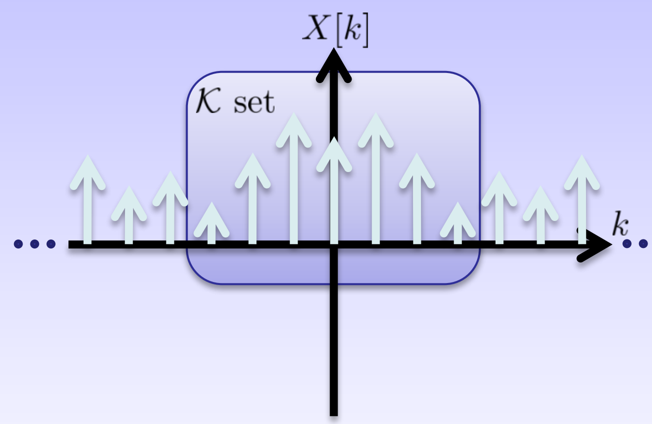


Theorem [Sufficient Condition]

If the filter $s^*(-t)$ satisfies :

$$S^*(\omega) = \begin{cases} 0 & \omega = 2\pi k/\tau, k \notin \mathcal{K} \\ \text{nonzero} & \omega = 2\pi k/\tau, k \in \mathcal{K} \\ \text{arbitrary} & \text{otherwise,} \end{cases}$$

and $N \geq |\mathcal{K}|$, then the vector $\mathbf{x}$ can be obtained from the samples $c[n]$, $n = 1 \ldots N$.



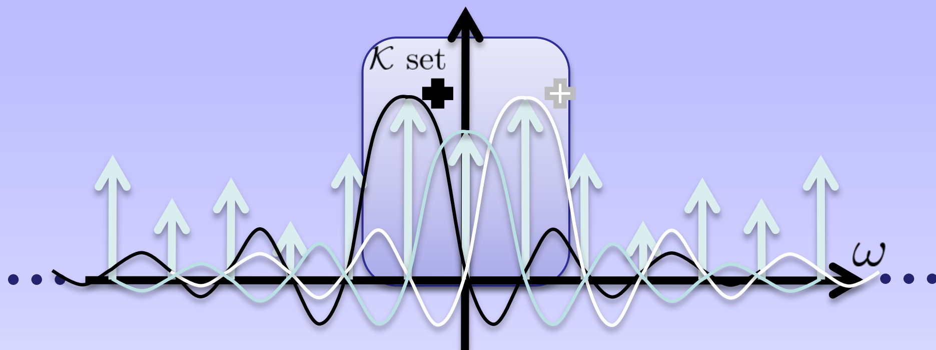
Tur, Eldar and Friedman, '11

Special Cases

- Low pass filter
- Sum of sincs (SoS) in the frequency domain

Vetterli *et al.*, '02

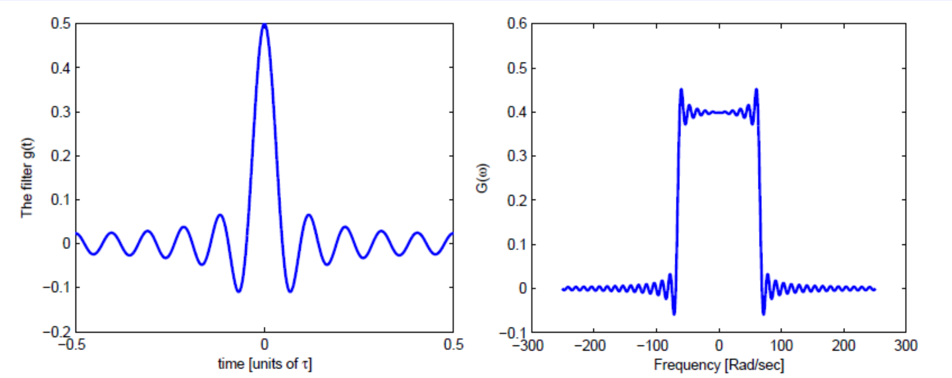
Tur, Eldar and Friedman, '11



$$\frac{\tau}{\sqrt{2\pi}} \sum_{k \in \mathcal{K}} b_k \mathrm{sinc}\left(\frac{\omega}{2\pi/\tau} - k\right)$$

Compact support!

- In the time domain $g(t) = \mathrm{rect}\left(\frac{t}{\tau}\right) \sum_{k \in \mathcal{K}} b_k e^{j2\pi kt/\tau}$
- For $b_k = 1$: $g(t) = \mathrm{rect}\left(\frac{t}{\tau}\right) D_p(2\pi t/\tau)$, $D_p(t)$ is the Dirichlet kernel

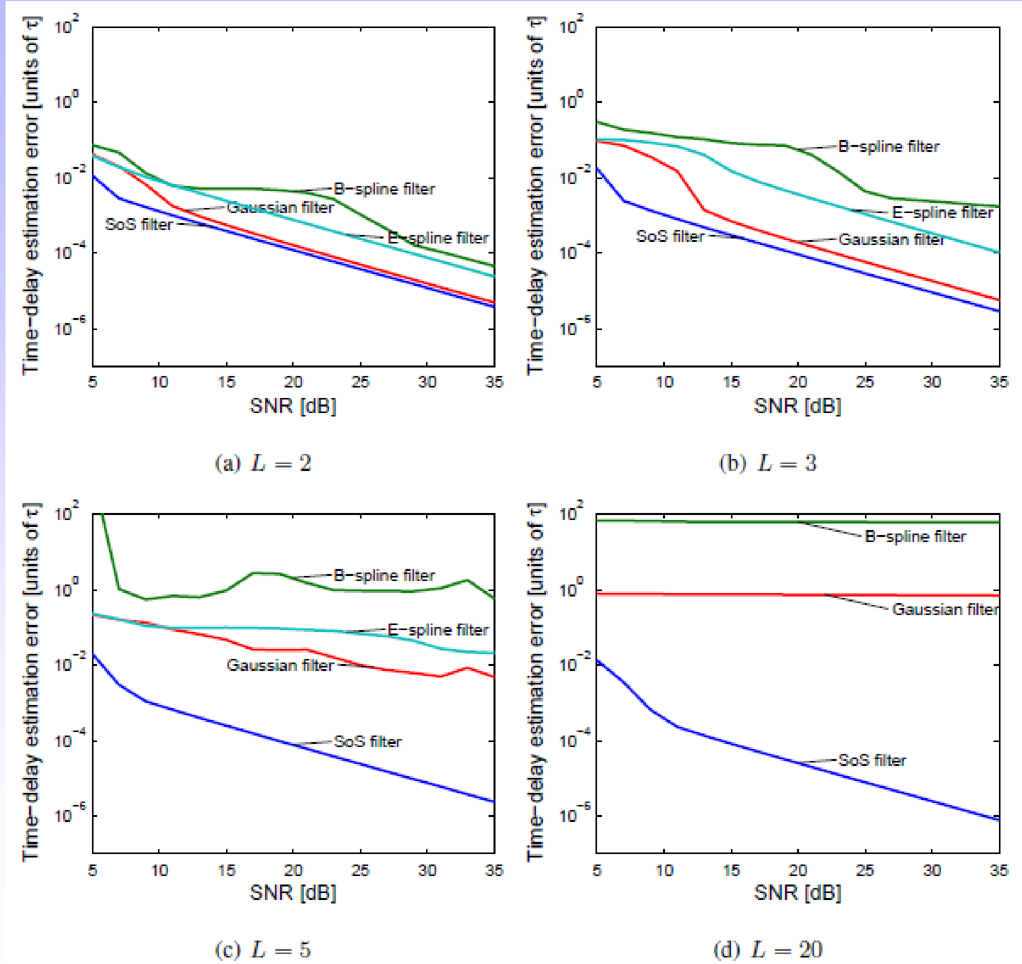


(a) Time domain (b) Frequency domain

Finite Pulse Streams

- SoS filter can be used for finite streams due to its finite support!
- Not true for LPF or other filters with long support

Far more robust than Spline based methods – works even for high *L*!



(a) $L = 2$ (b) $L = 3$ (c) $L = 5$ (d) $L = 20$

Multichannel Scheme

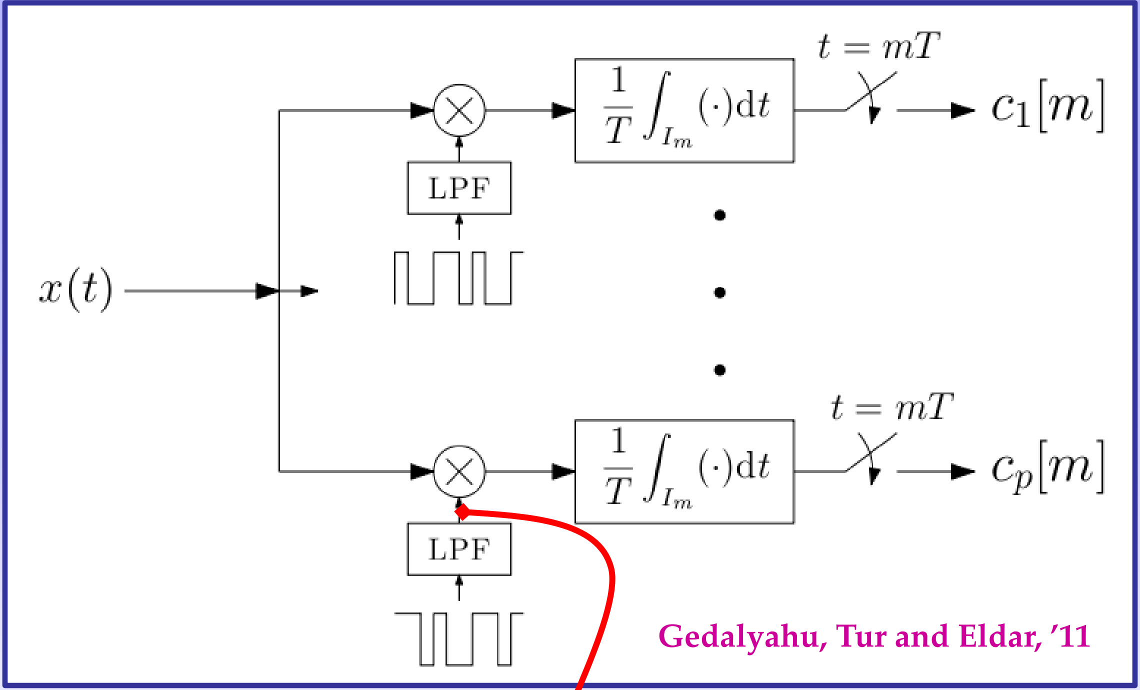
Proposed scheme:

- Mix & integrate
- Take linear combinations from which Fourier coeff. can be obtained

$$\mathbf{c} = \mathbf{S}\mathbf{x}$$

Samples

Fourier coeff. vector

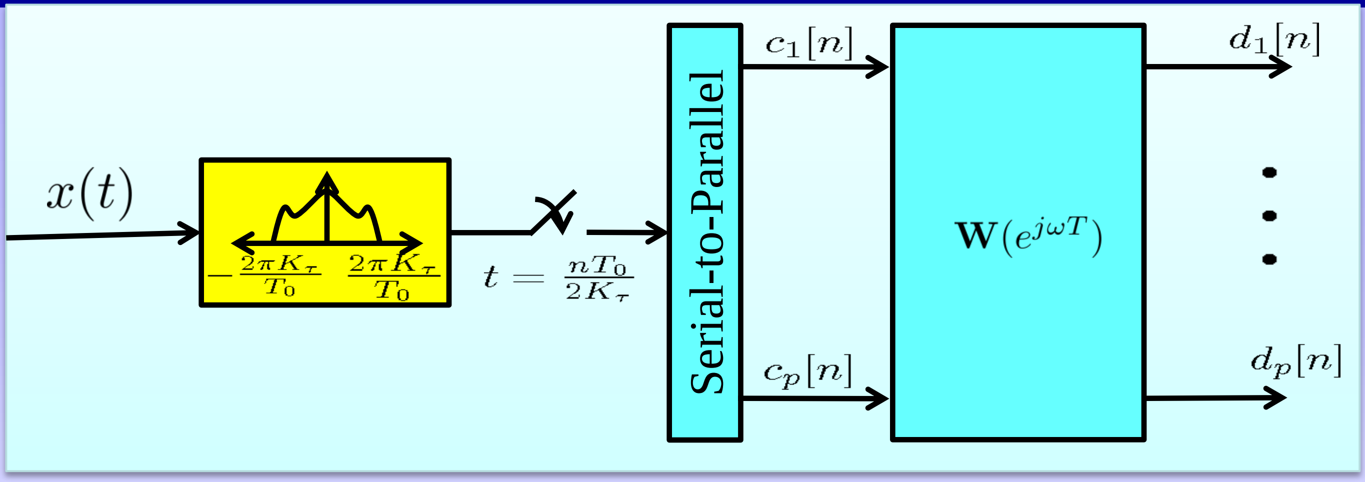


$$= \sum_k s_{i\ell} e^{-j2\frac{\pi}{T}kt}$$

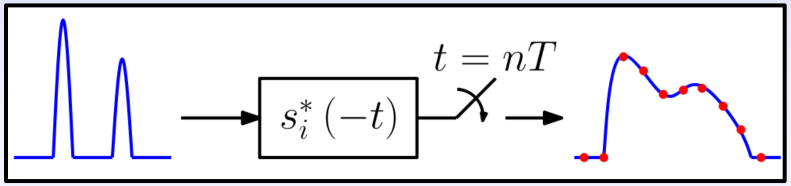
- Supports general pulse shapes (time limited)
- Operates at the rate of innovation
- Stable in the presence of noise
- Practical implementation based on the MWC
- Single pulse generator can be used

$$\mathbf{S} = [s_{i\ell}]$$

Filter Bank Approach



- The analog sampling filter "smoothens" the input signal : Gedalyahu and Eldar, '09
  - Allows sampling of short-length pulses at low rate
  - **CS interpretation:** each sample is a linear combination of the signal's values.



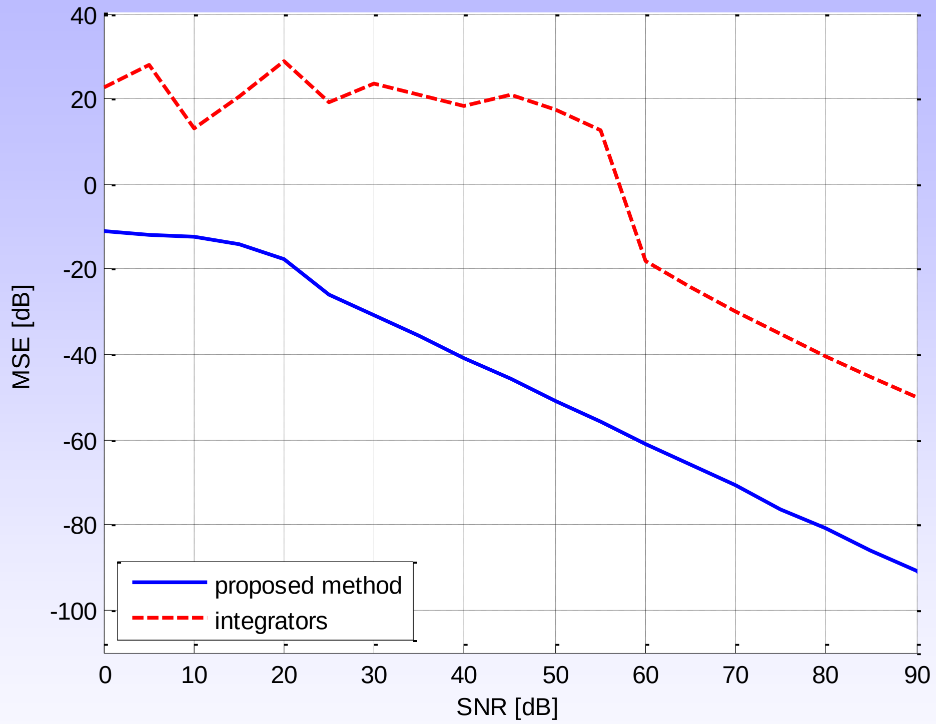
- The digital correction filter-bank:
  - Removes the pulse and sampling kernel effects
  - Samples at its output satisfy: $\boxed{\mathbf{d}[n] = \mathbf{V}(\tau_i)\mathbf{a}[n]}$ $\mathbf{V}(\tau_i)$ is Vandermonde
  - The delays can be recovered using ESPRIT as long as $W \geq 2\pi K_\tau / T_0$

Noise Robustness

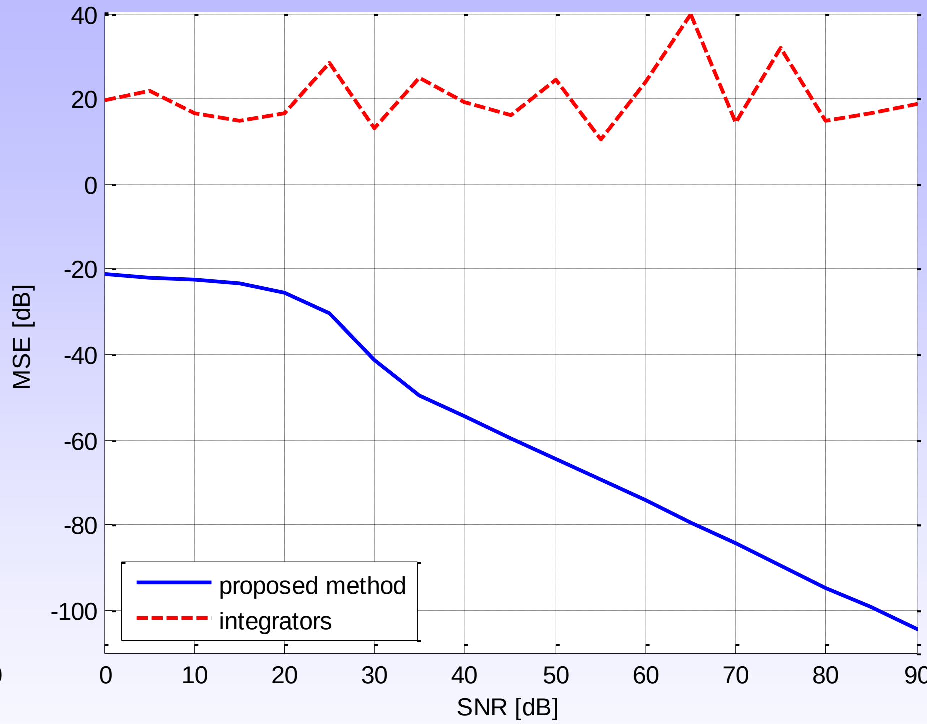
- MSE of the delays estimation, versus integrators approach

Kusuma and Goyal, '06

L=2 pulses, 5 samples



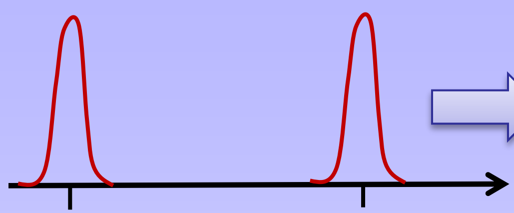
L=10 pulses, 21 samples



The proposed scheme is stable even for high rates of innovation!

Application: Multipath Medium Identification

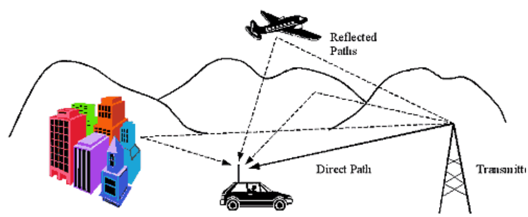
Gedalyahu and Eldar, '09-'10

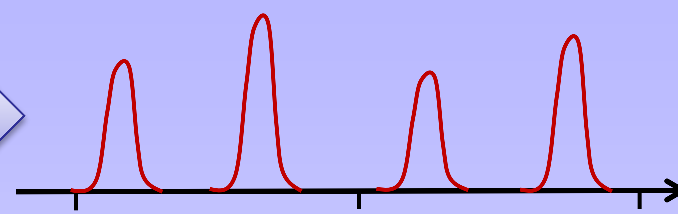


$$x_T(t) = \sum_{n\in\mathbb{Z}} h(t - nT)$$

LTV channel

L propagation paths





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

L pulses per period T

- Medium identification:
  - Recovery of the time delays
  - Recovery of time-variant gain coefficients

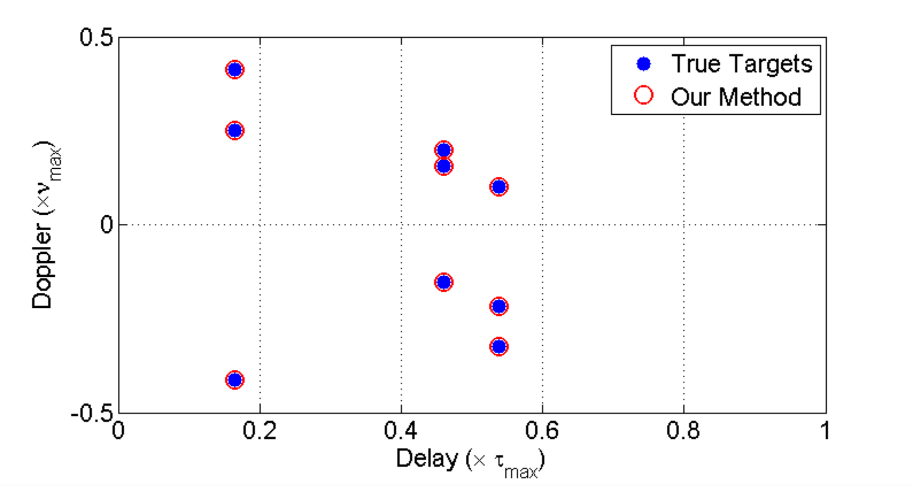
The proposed method can recover the channel parameters from sub-Nyquist samples

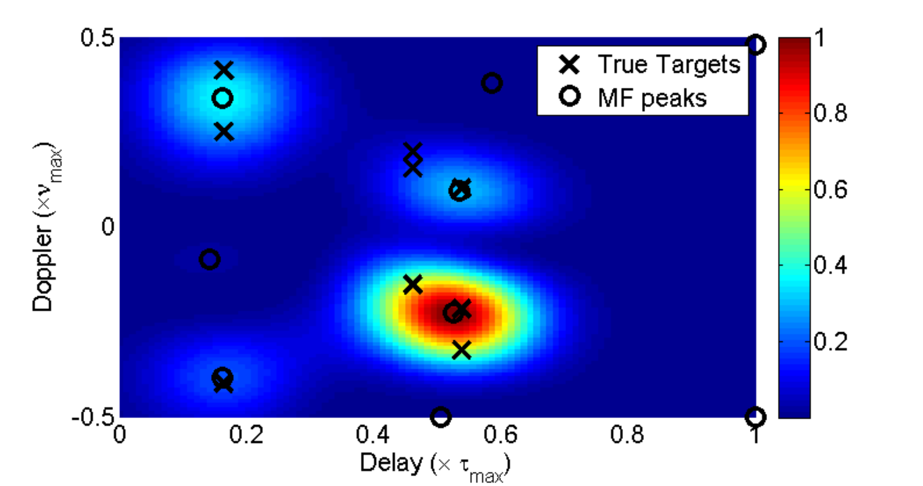
Application: Radar

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

Bajwa, Gedalyahu and Eldar, '11

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





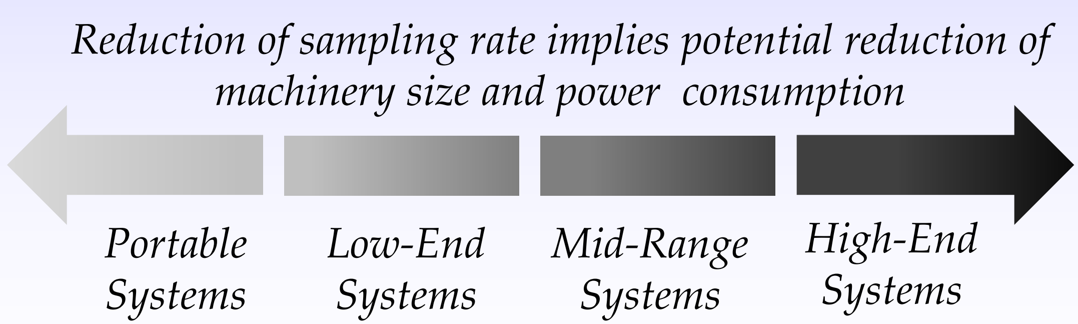
Xampling in Ultrasound Imaging

Motivation

Wagner, Eldar, Feuer, Danin and Friedman, '11

- Generate a two-dimensional focused ultrasound image while reducing the sampling rate in each active element far below the Nyquist rate

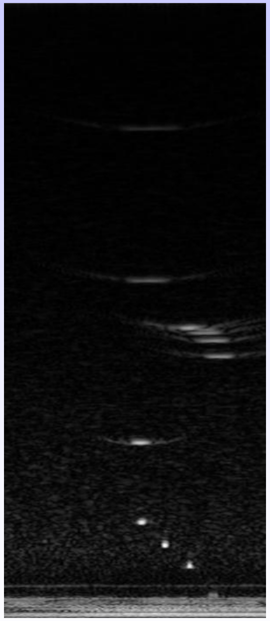
- Sample rate reduction leads to significant reduction of data size, and implies potential reduction of machinery size and power consumption, while maintaining image quality



Xampling in Ultrasound Imaging

Main Results

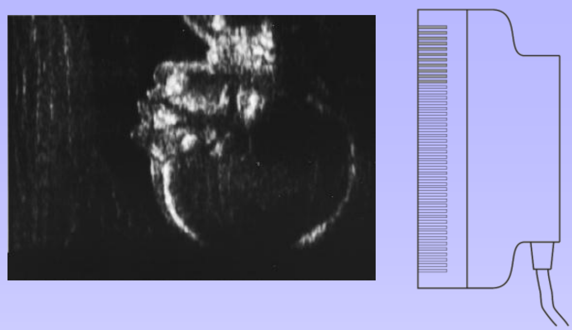
- A scheme which enables reconstruction of a two dimensional image, from samples obtained at a rate 10-15 times below Nyquist
- The resulting image depicts strong perturbations in the tissue

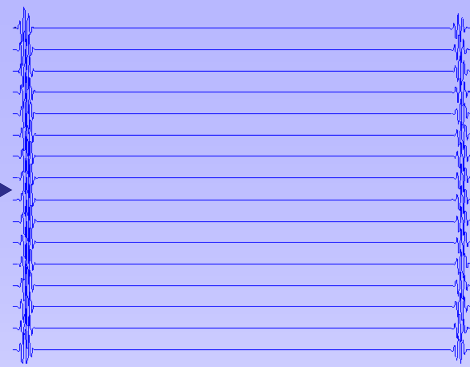


Standard B-mode image, generated from samples obtained at Nyquist Rate

Xampled B-mode image, generated from samples obtained at 0.17 Nyquist Rate

Ultrasound Imaging & FRI





Digital Beamformer

An ultrasonic pulse is transmitted into a tissue; echoes are scattered and reflected by density and propagation-velocity perturbations

$\varphi_m(t)$ - Individual traces of reflected echoes, received by the multi-element transducer array

Beamformed Signal

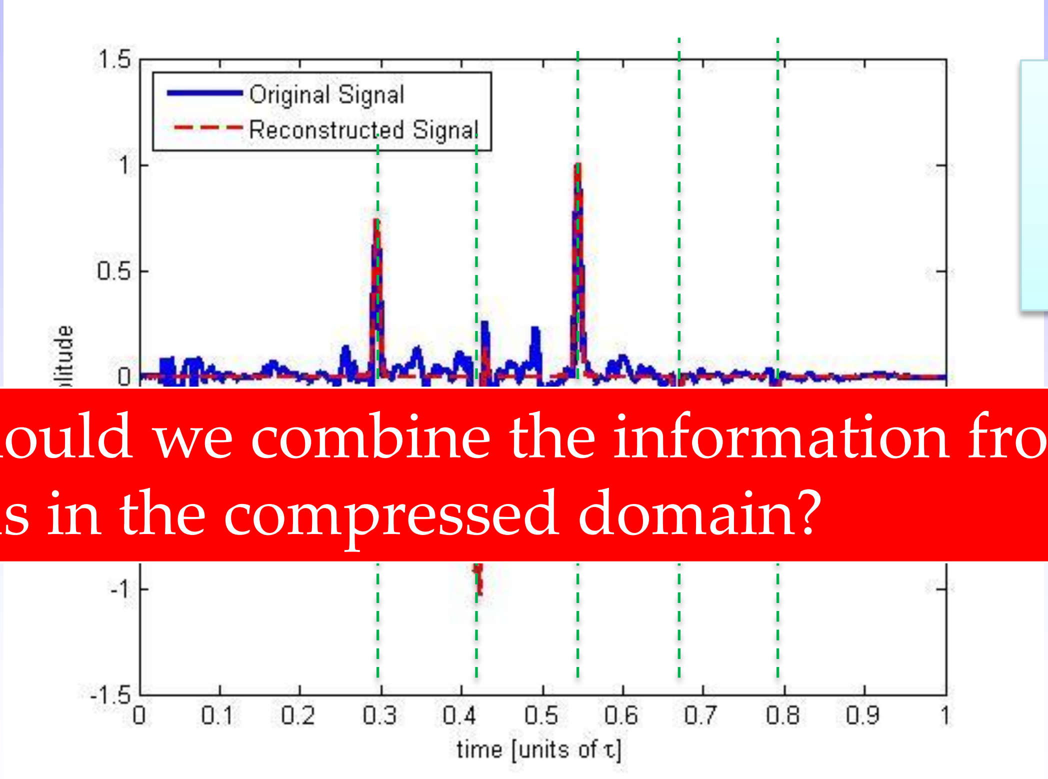
Meets the finite FRI model: $$x(t) = \sum_{l=1}^{L} a_l h(t - t_l)$$

Unknowns

Time - t_i
Amplitude - a_i

Ultrasound Experiment

- Real data acquired by GE Healthcare's Vivid-i imaging system
- Method applied on noisy signal
- Excellent reconstruction from sub-Nyquist samples
- Poor SNR motivates integration of the data from multiple receivers

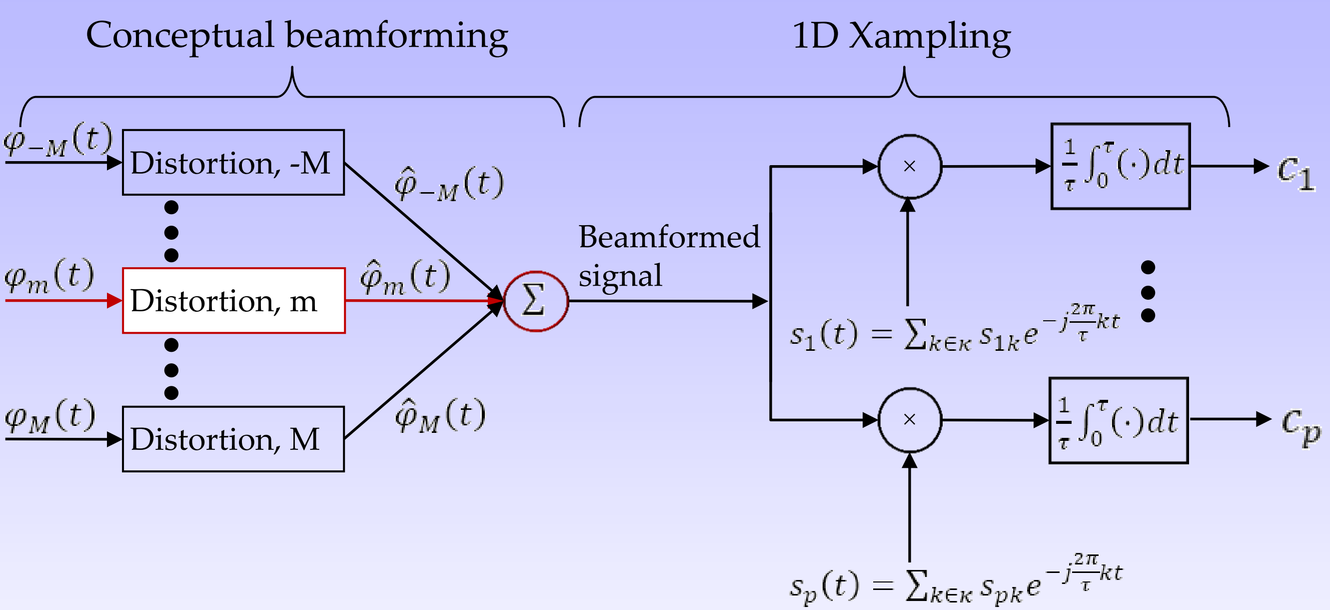


Phantom comprising 5 equally spaced scatterers

How should we combine the information from all channels in the compressed domain?

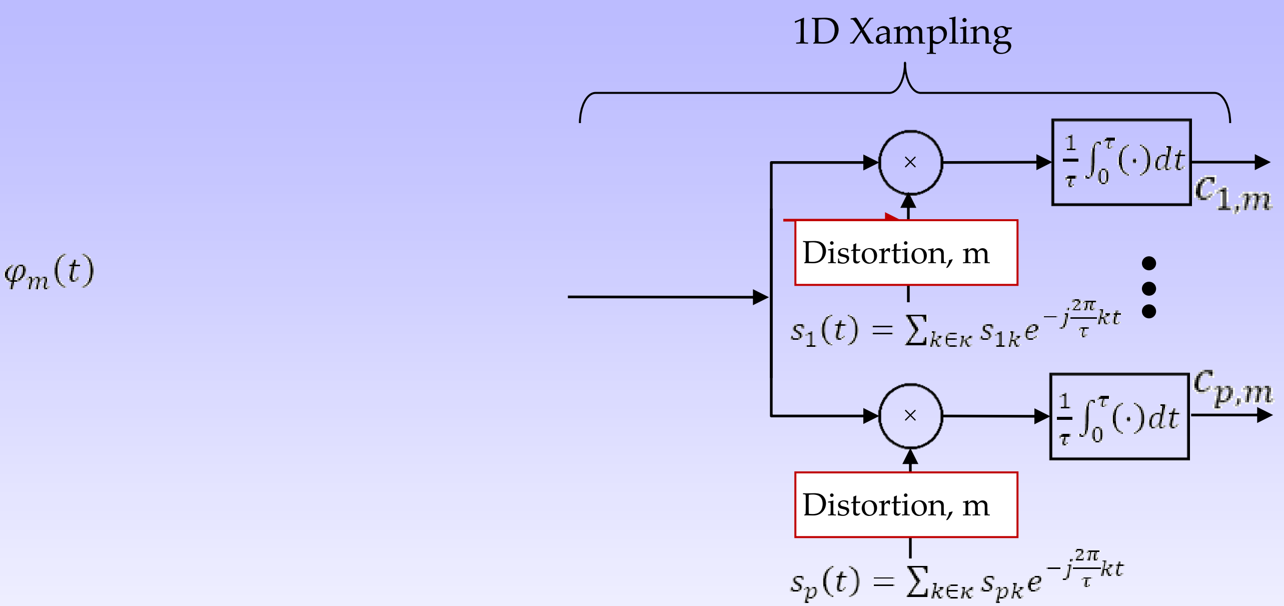
Beamforming in the Compressed Domain

Wagner, Eldar, Feuer, Danin and Friedman, '11



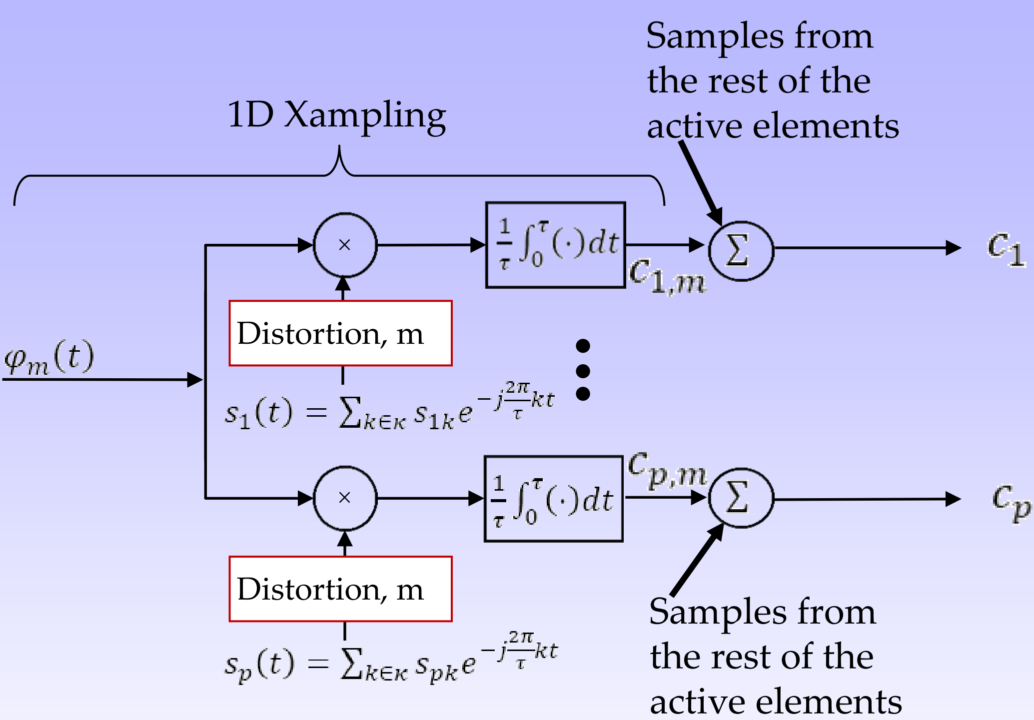
Beamforming in the Compressed Domain

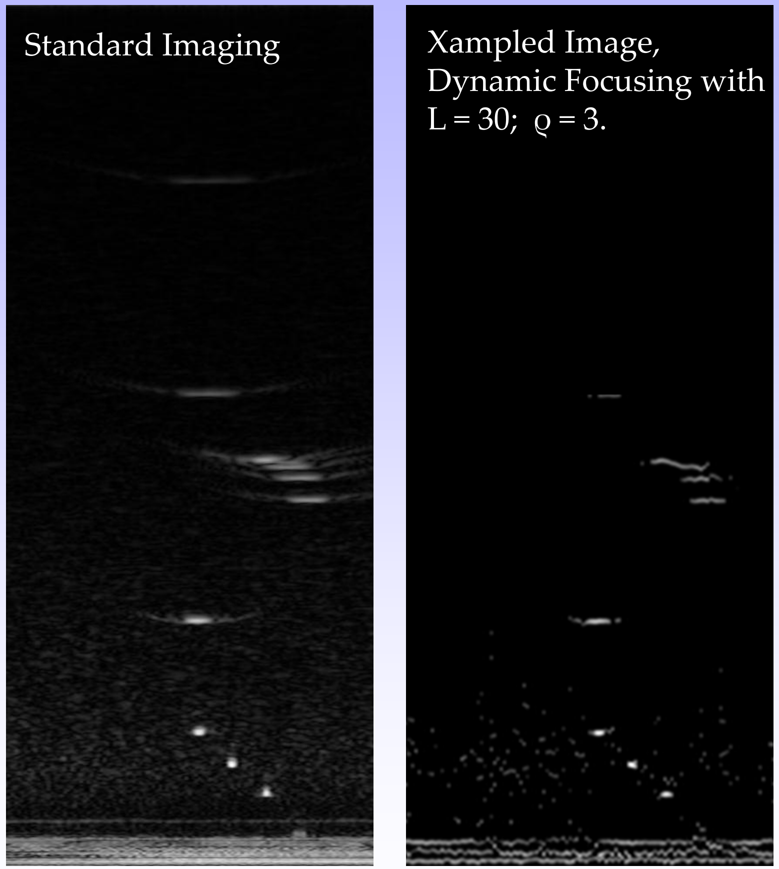
Wagner, Eldar, Feuer, Danin and Friedman, '11



Beamforming in the Compressed Domain

Wagner, Eldar, Feuer, Danin and Friedman, '11





RF ultrasound data provided by Dr. Omer Oralkan and Prof. Pierre Khuri-Yakub of the E. L. Ginzton Laboratory at Stanford University.

Xampling Systems

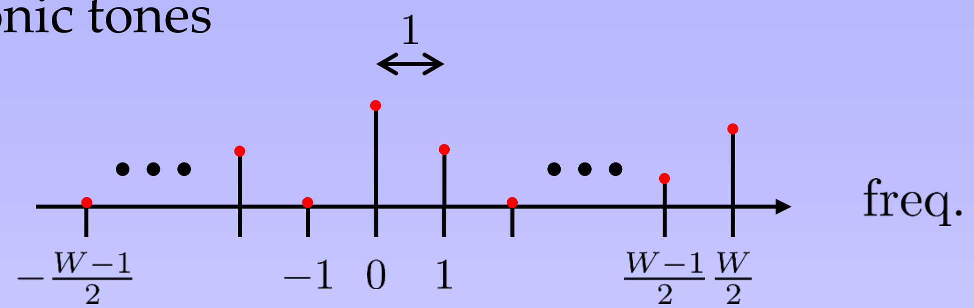
- Modulated wideband converter — Mishali and Eldar, '07-'09
- Periodic nonuniform sampling (fully-blind) — Mishali and Eldar, '07-'09
- Sparse shift-invariant framework — Eldar, '09
- Finite rate of innovation sampling — Vetterli *et al.*, '02-'07; Dragotti *et al.*, '02-'07; Gedalyahu, Tur and Eldar, '10-'11
- Random demodulation — Tropp *et al.*, '09

Random Demodulation

- **Model:** sparse sum of harmonic tones

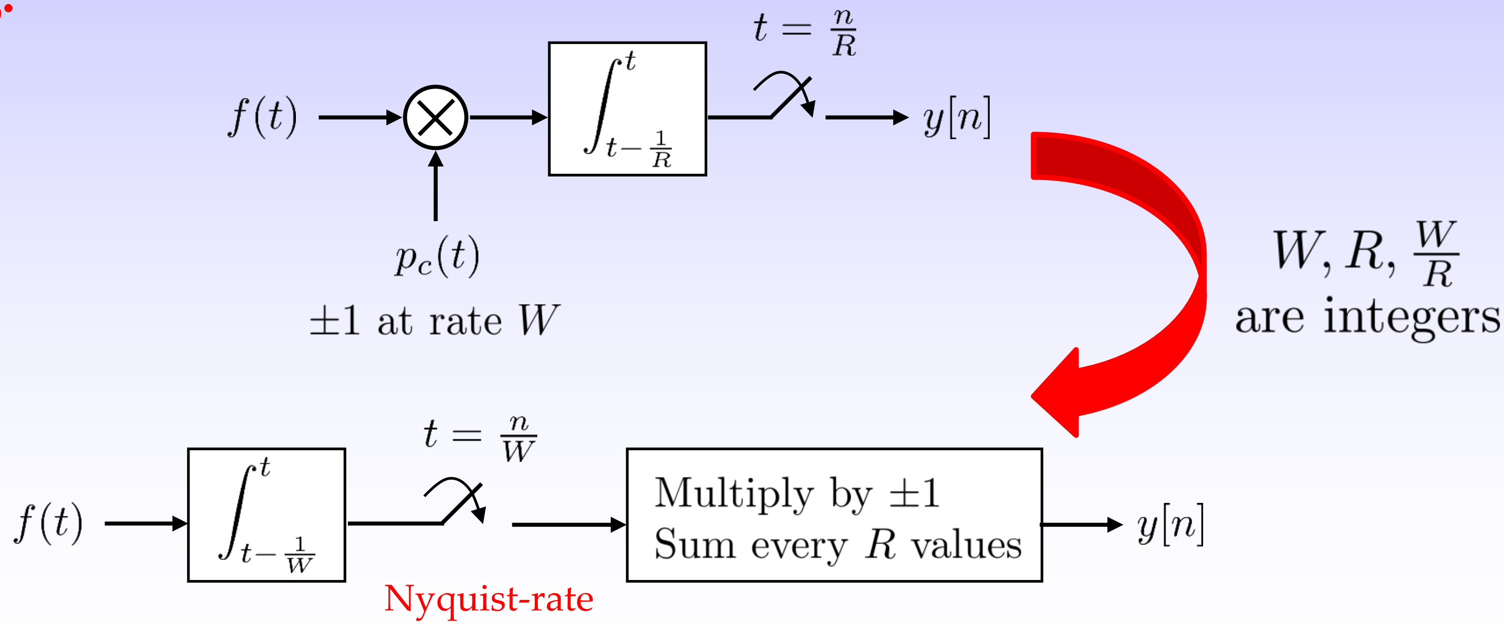
$$f(t) = \sum_{\omega \in \Omega} a_\omega e^{j2\pi\omega t}$$

K active tones, $|\Omega| \leq K$



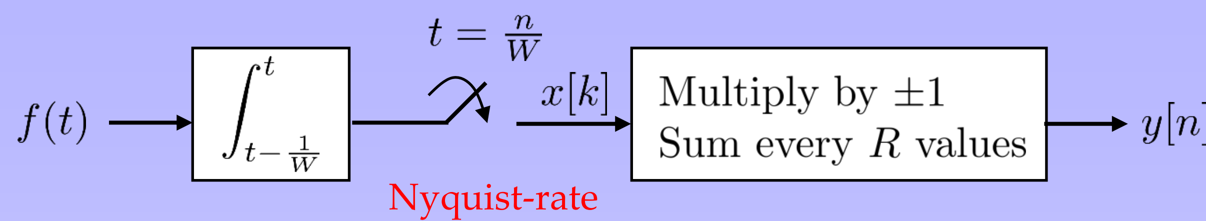
Tropp *et al.*, '09

- **Sampling:**



Random Demodulation

- **Reconstruction:**



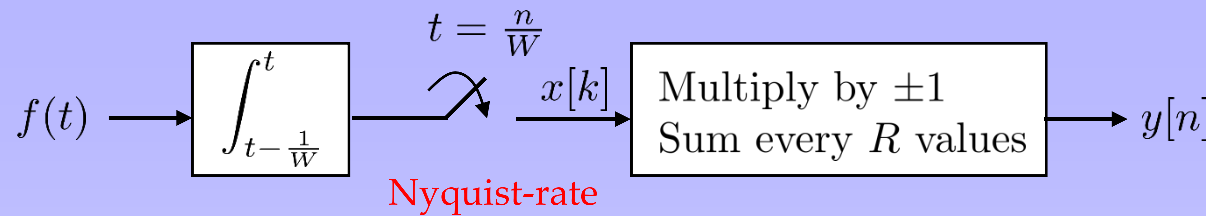
- Integers W, R, $\frac{W}{R}$

$$y[n] = \underbrace{\begin{bmatrix} 1\cdots1 & & & \\ & 1\cdots1 & \overset{\frac{W}{R}}{\longleftrightarrow} & \\ & & \ddots & \\ & & & 1\cdots1\end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix}\pm1 & & \\ & \ddots & \\ & & \pm1\end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} x[1]\\ \vdots \\ x[W]\end{pmatrix}}_{\mathbf{x}}$$

Tropp *et al.*, '09

Random Demodulation

- **Reconstruction:**



- Integers $W, R, \frac{W}{R}$ + multitone input $(a'_\omega = c_\omega a_\omega)$:

$$y[n] = \underbrace{\begin{bmatrix} 1\cdots 1 & & & \\ & 1\cdots 1 & \xleftrightarrow{\frac{W}{R}} & \\ & & \ddots & \\ & & & 1\cdots 1 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} \pm 1 & & \\ & \ddots & \\ & & \pm 1 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \text{DFT} \\ \text{matrix} \\ W \times W \end{bmatrix}}_{\mathbf{F}} \underbrace{\begin{pmatrix} \vdots \\ a'_\omega \\ \vdots \end{pmatrix}}_{\mathbf{a}}$$

k-sparse

- Use CS solvers to recover $\mathbf{a}$, then reconstruct $f(t)$

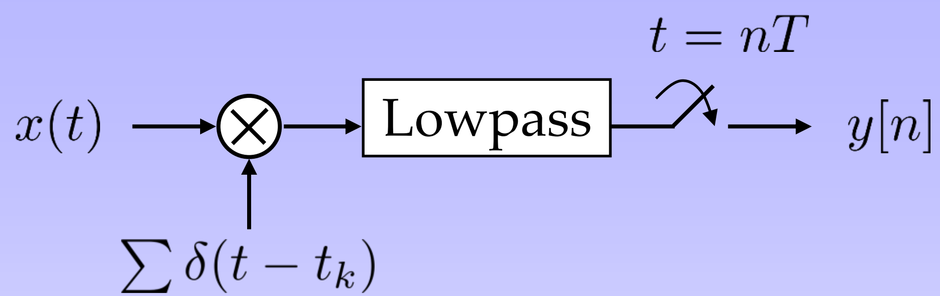
- Numerical simulations: 32 kHz AM signal recovered from sampling at 10% Nyquist rate

Tropp *et al.*, '09

- Similar to MWC? Next part describes the differences…

Nyquist Folding

- Rate reduction using nonlinear sampling effects:



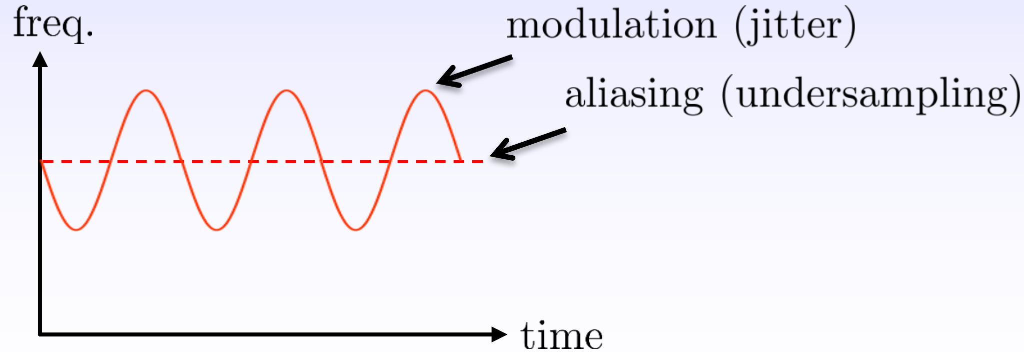
Fudge *et al.*, '08

- $t_k = nT$ $\rightarrow$ undersampling at rate $\frac{1}{T}$
- t_k jitter around nT $\rightarrow$ undersampling + frequency-dependent modulation

Input

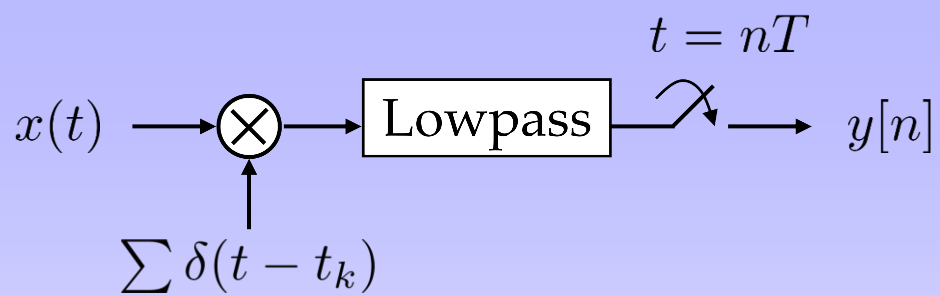
$\sin(\omega_1 t)$

Output



Nyquist Folding

- Rate reduction using nonlinear sampling effects:



Fudge *et al.*, '08

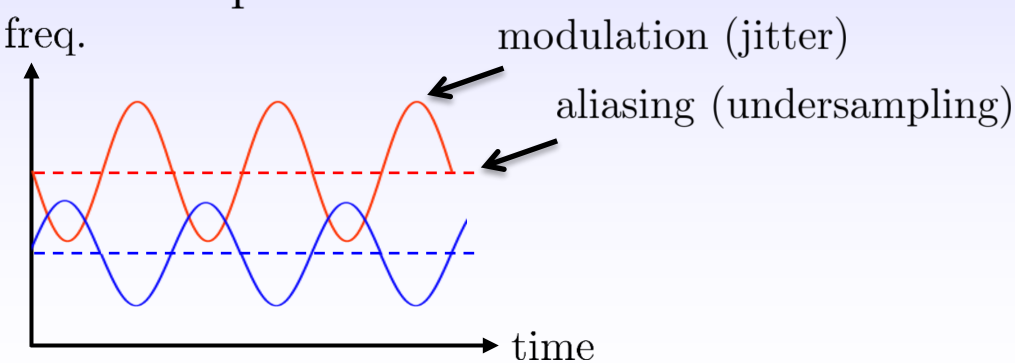
- $t_k = nT$ $\longrightarrow$ undersampling at rate $\frac{1}{T}$
- t_k jitter around nT $\longrightarrow$ undersampling + frequency-dependent modulation

Input

$\sin(\omega_1 t)$

$\sin(\omega_2 t)$

Output



Summary: Xampling Systems

Model	Union dim. $\Lambda, \mathcal{A}_\lambda$	Strategy	X-ADC	X-DSP
Multiband	finite ∞	MWC Mishali-Eldar 09	Periodic mixing	CTF
		PNS Mishali-Eldar 08	time shifts	CTF
		Nyquist-folding Fudge et al. 08	Jittered undersampling	
Sparse shift-invariant	finite ∞	Eldar 08	Filter-bank	CTF
FRI (time-delays)	∞ finite	Periodic Vetterli et al. 02-05	Lowpass	Annihilating filter
		One-shot Dragotti et al. 07	Splines	Moments factoring
		Periodic/one-shot Gedlyahu-Tur-Eldar 09-10	Sum-of-Sincs filtering	Annihilating filter
Sequences of innovation	∞ ∞	Gadlyahu-Eldar 09	Lowpass or periodic mixing + integration	MUSIC or ESPRIT
Harmonic tones	finite finite	RD Tropp et al. 09	Sign flipping + integration	CS

``Xampling: Signal Acquisition and Processing in Union of Subspaces'', Mishali, Eldar and Elron, *TSP* '11

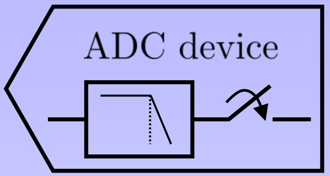
– Part 5 –
From Theory to Hardware

→ Outline

Theory vs. Practice

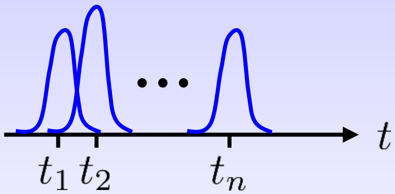
- Practical considerations affect the choice of a sampling solution

- Example 1: Multiband sampling (known carriers f_i)



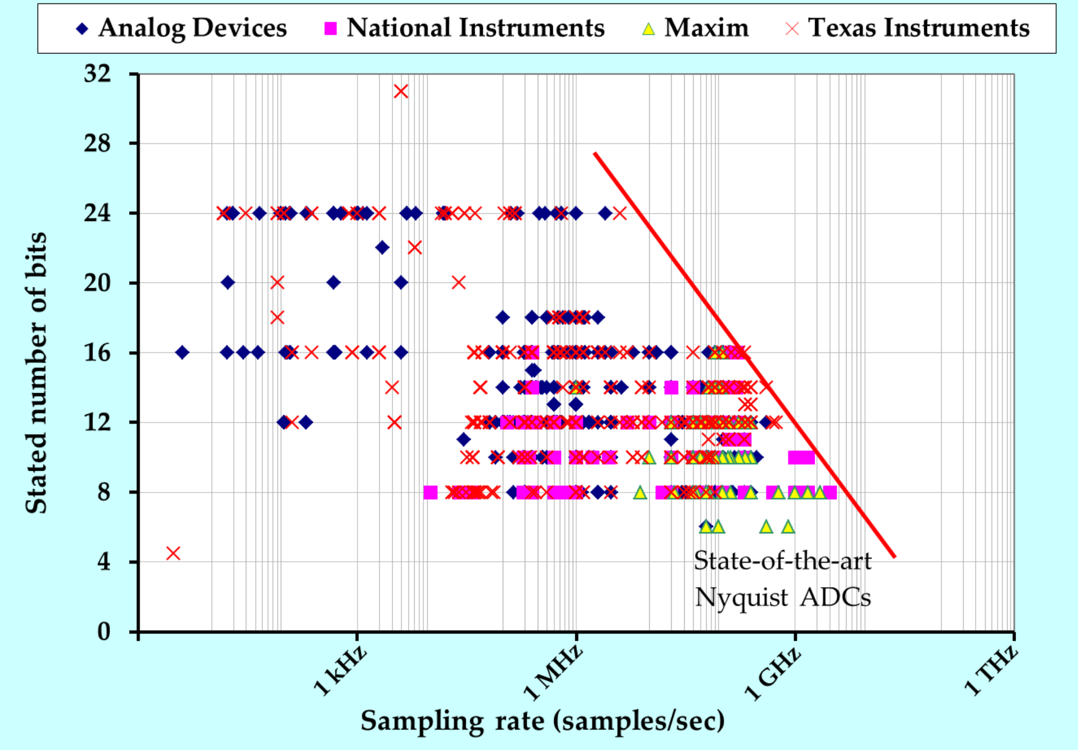
	RF demodulation	Nonuniform methods
Minimal analog preprocessing		✓
ADC with low analog bandwidth	✓	

- Example 1: Pulse streams (known delays t_n)



	$s_n(t) = h(t - t_n)$	Digital match filter
Low sampling rate	✓	
Robustness to model mismatch		✓

ADC Market



- State-of-the-art ADCs generate Nyquist samples
- Today's challenges:
  - Increase sampling rate (Giga-samples/sec)
  - Increase front-end bandwidth
  - Increase (effective) number of bits

Sub-Nyquist: Practical Challenges

- Goal: Shift $f_{\max}$ challenge away from ADC technology

- No free lunches ! Signal has frequencies until $f_{\max}$
- Nyquist will enter elsewhere into system design

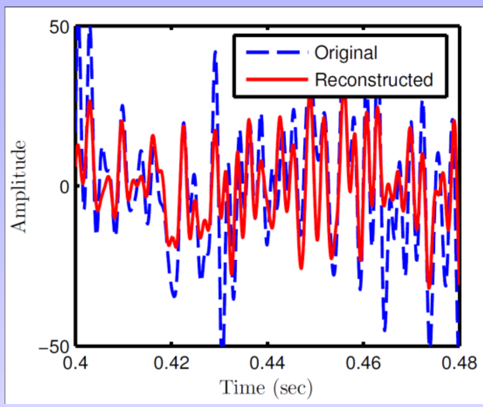
- Practical design metrics:
  - robustness to model mismatches
  - flexible hardware design
  - light computational loads
  - imaging: high resolution
  - noise performance
  - power, area, size, cost, …

Focus of this part

- Next slides:
  - Study practical metrics of example sub-Nyquist systems (RD/MWC)
  - Glance into sub-Nyquist circuit challenges
  - Sub-Nyquist imaging: analog vs. discrete CS

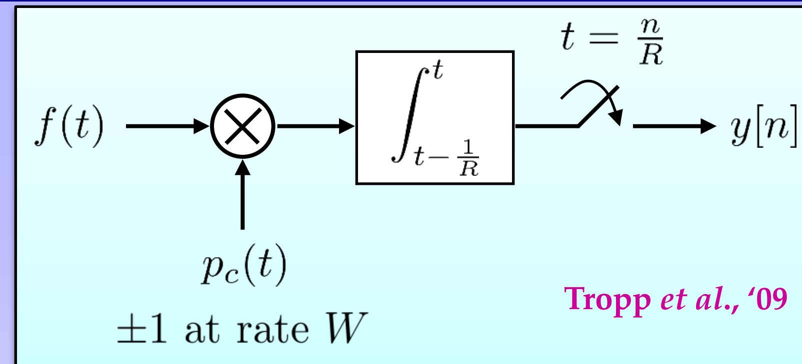
Random Demodulator

- **Robustness:**



0.005% grid mismatch

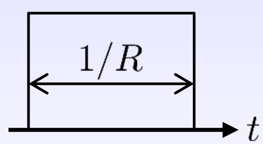
$$\frac{\|f(t)-\hat{f}(t)\|^2}{\|f(t)\|^2} = 37\%$$



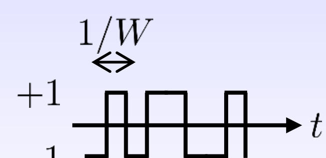
$\rightarrow W, R$ must be integer multiples of tones grid spacing

- **Required hardware accuracy** (so that $\mathbf{y} = \mathbf{HDFa}$):

Accurate integrator:



$p_c(t) =$



``Nice'' time-domain appearance

- **Computational load:** $W = 1\text{MHz} \rightarrow$ CS on 1 million unknowns

- **Reported hardware:** $W = 800$ kHz, $R = 100$ kHz
  DSP processor 160 MHz

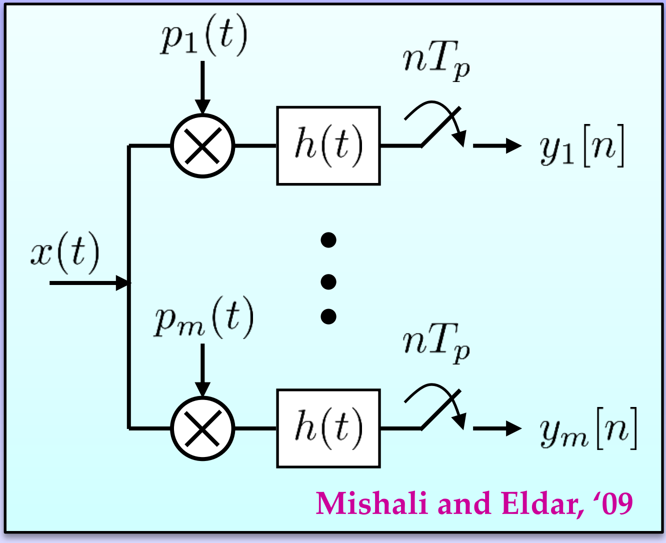
Ragheb *et al.*, '08
Yu *et al.*, '10

Modulated Wideband Converter

- **Robustness:**

$$m \geq 2N, \quad 1/T_p \geq B \quad \text{(basic setup)}$$

Inequalities allow model mismatches



- **Required hardware accuracy:**

$$\left.\begin{array}{ll} p_i(t) & = \text{periodic waveforms} \\ h(t) & = \text{lowpass} \end{array}\right\}$$ ``Nice'' freq.-domain appearance

Nonideal lowpass response can be compensated digitally — Chen *et* al., '10

- **Computational load:** $f_{\text{NYQ}} = 5 \text{ GHz}, \; N = 6, \; B = 50 \text{ MHz}$

CS system size: 40×200

linear real-time reconstruction

- **Reported hardware:** $f_{\text{NYQ}} = 2.2 \text{ GHz}$, sampling rate 280 MHz

10msec recovery (on PC-MATLAB)

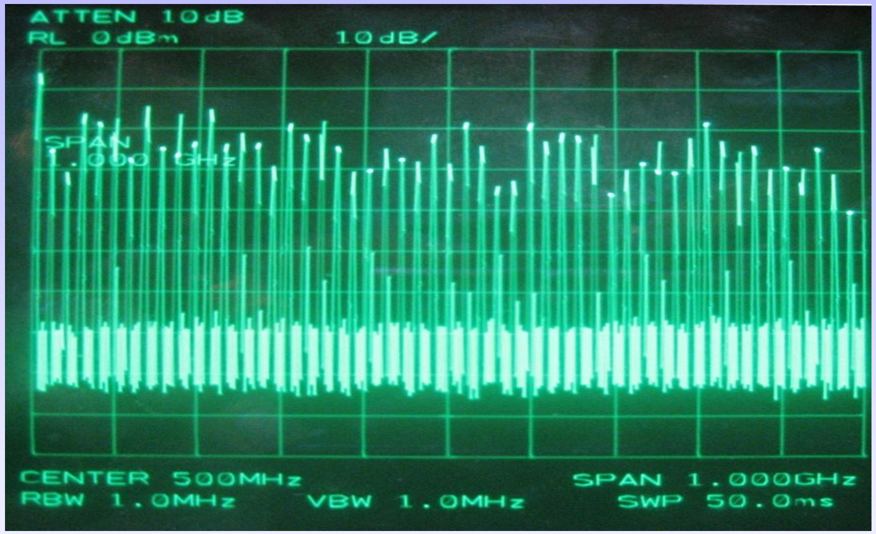
Mishali *et* al., '11

Hardware Accuracy

- Sign alternating functions at 2 GHz rate

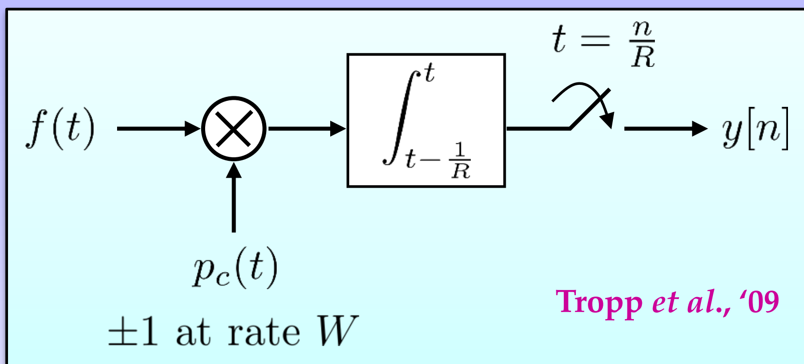
Time appearance

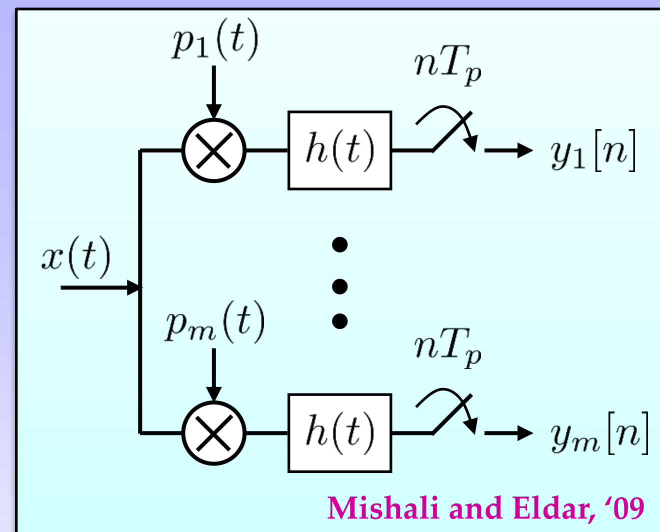
Frequency appearance



Comparison

- Visually-similar systems – major differences in practical metrics





- No free lunches… Nyquist enters in:
  - Time-domain accuracy
  - Computational loads
  - Freq.-domain accuracy (handled by RF front-end)
- Similar conclusions in other applications?

CS Radar

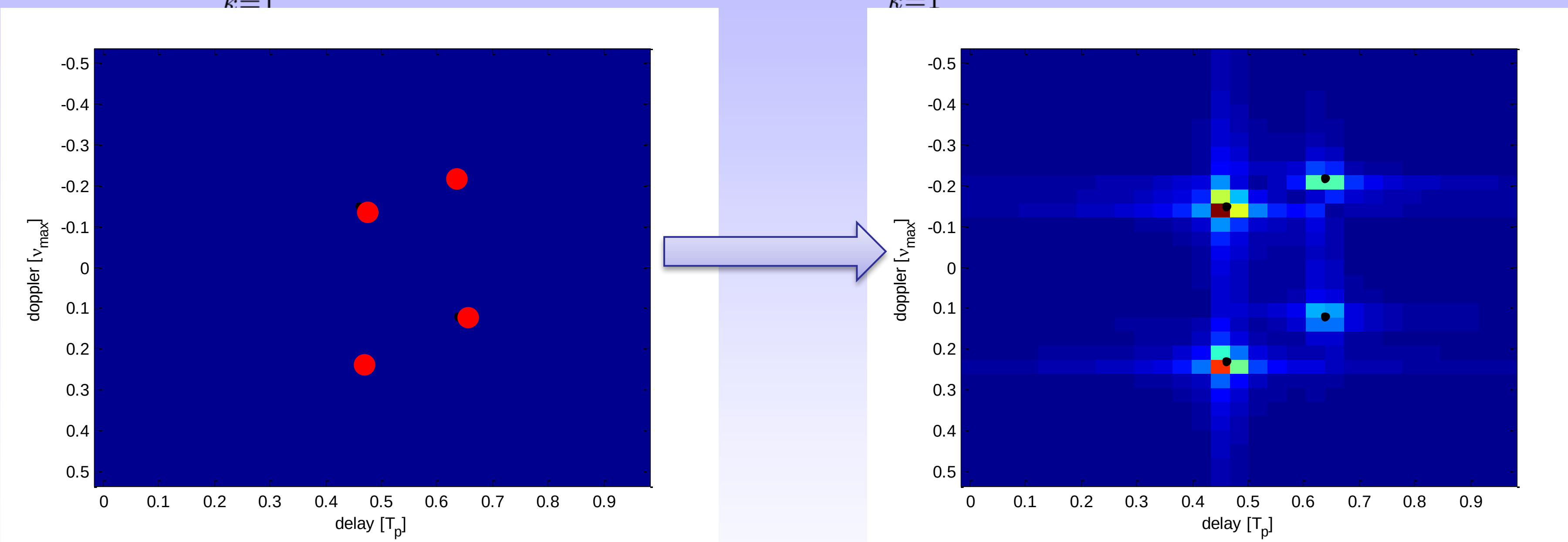
- A discrete version of the channel is being estimated
- Leakage effect → fake targets

Real channel

$$C(\tau,\nu)=\sum_{k=1}^{K}\alpha_k\delta(\tau-\tau_k)\delta(\nu-\nu_k)$$

Discretized channel

$$C(\ell,m)=\sum_{k=1}^{K}\alpha_k e^{j\pi(m-\mathcal{T}\nu_k)}\mathrm{sinc}(m-\mathcal{T}\nu_k)\mathrm{sinc}(\ell-\mathcal{W}\tau_k)$$



- Limited resolution to $1/\mathcal{W},\ 1/\mathcal{T}$
- Sampling process in hardware is unclear
- Digital processing is complex and expensive

ADCs: Why Not Standard CS?

- CS is for finite dimensional models *($y=Ax$)*
- Loss in resolution when discretizing
- Sensitivity to grid, analog bandwidth issues
- Is not able to exploit structure in analog signals
- Results in large computation on the digital side
- Samples do not typically interface with standard processing methods

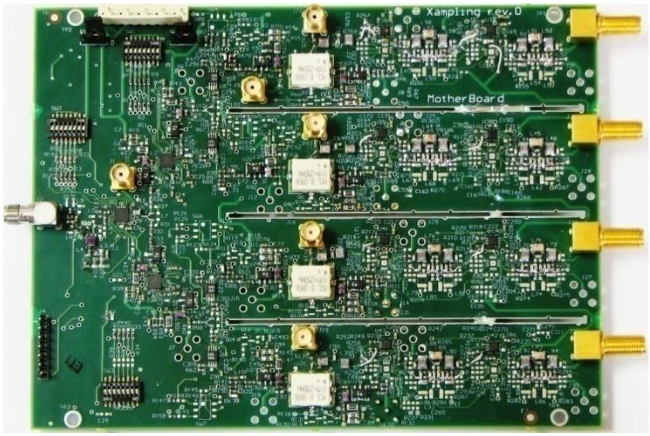
More details in: M. Mishali, Y. C. Eldar, and A. Elron, "Xampling: Signal acquisition and processing in union of subspaces"

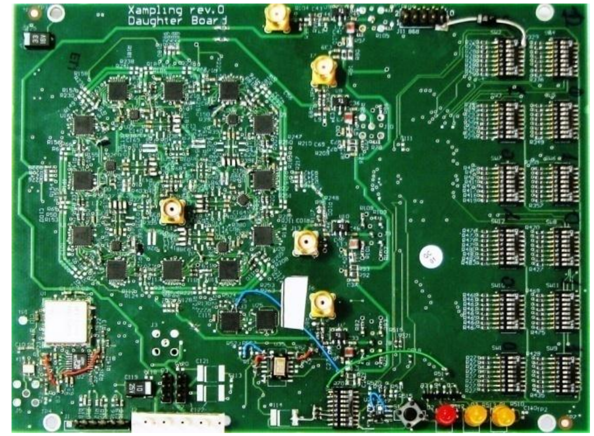
Besides union models and Xampling there are many more challenges !

Stepping CS to Practice

- Address wideband noise and dynamic range:
  - Since x is noisy: $y=A(x+e)+w$, e=wideband noise
  - MWC/PNS: Nyquist-bandwidth noise is aliased
  - RD: noise is folded from all possible tone locations
  - Large interference will swamp ADC
- Integrate into existing systems
  - Minimal (preferably no) modification to hardware
  - *e.g.*, reprogramming firmware, rewiring, etc.
  - Deal with large analog BW and wide dynamic range
- Prove cost-effective
  - Rate is only one factor ! Digital complexity is not less important
  - Improve effective number of bits / Xample
- **Next slides:** quick glance at circuit challenges + applications

A 2.4 GHz Prototype

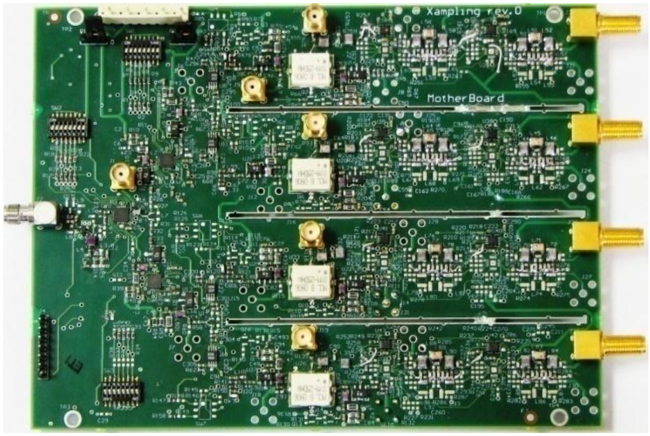




- 2.3 GHz Nyquist-rate, 120 MHz occupancy
- 280 MHz sampling rate
- Wideband receiver mode:
  - 49 dB dynamic range
  - SNDR > 30 dB over all input range
- ADC mode:
  - 1.2 volt peak-to-peak full-scale
  - 42 dB SNDR = 6.7 ENOB
- Off-the-shelf devices, ~5k$, standard PCB production

Mishali and Eldar, '08-10

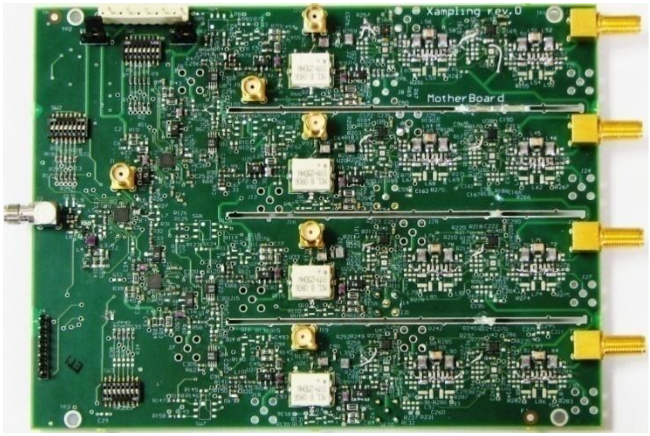
Circuit Design (2)



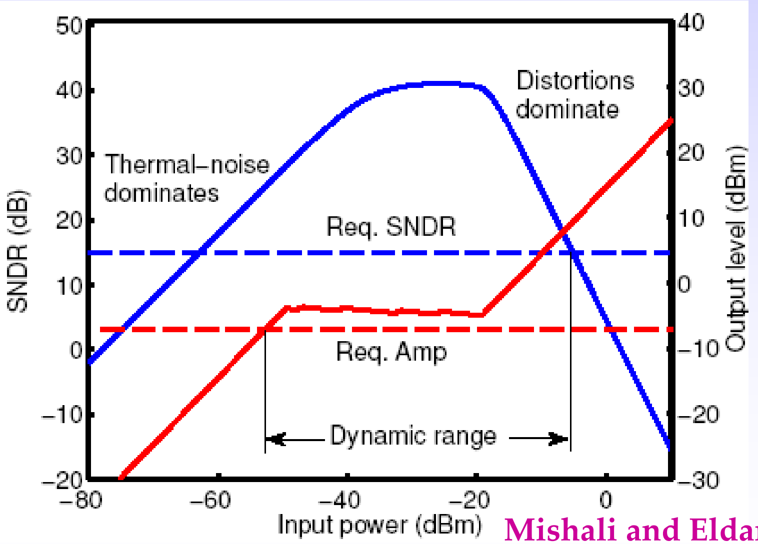
- Analog board
  - m=4 channels
  - 1:4 Split + mixing + filtering
  - Filter cutoff 33 MHz
  - Sampling rate 70 MHz per channel (scope)
- Digital board: sign alternating sequences
  - 2.075 GHz VCO
  - Discrete ECL shift-register
  - M=108 bits
  - 4 Outputs (taps of the register)

Mishali and Eldar, '08-10

Circuit Design (3)

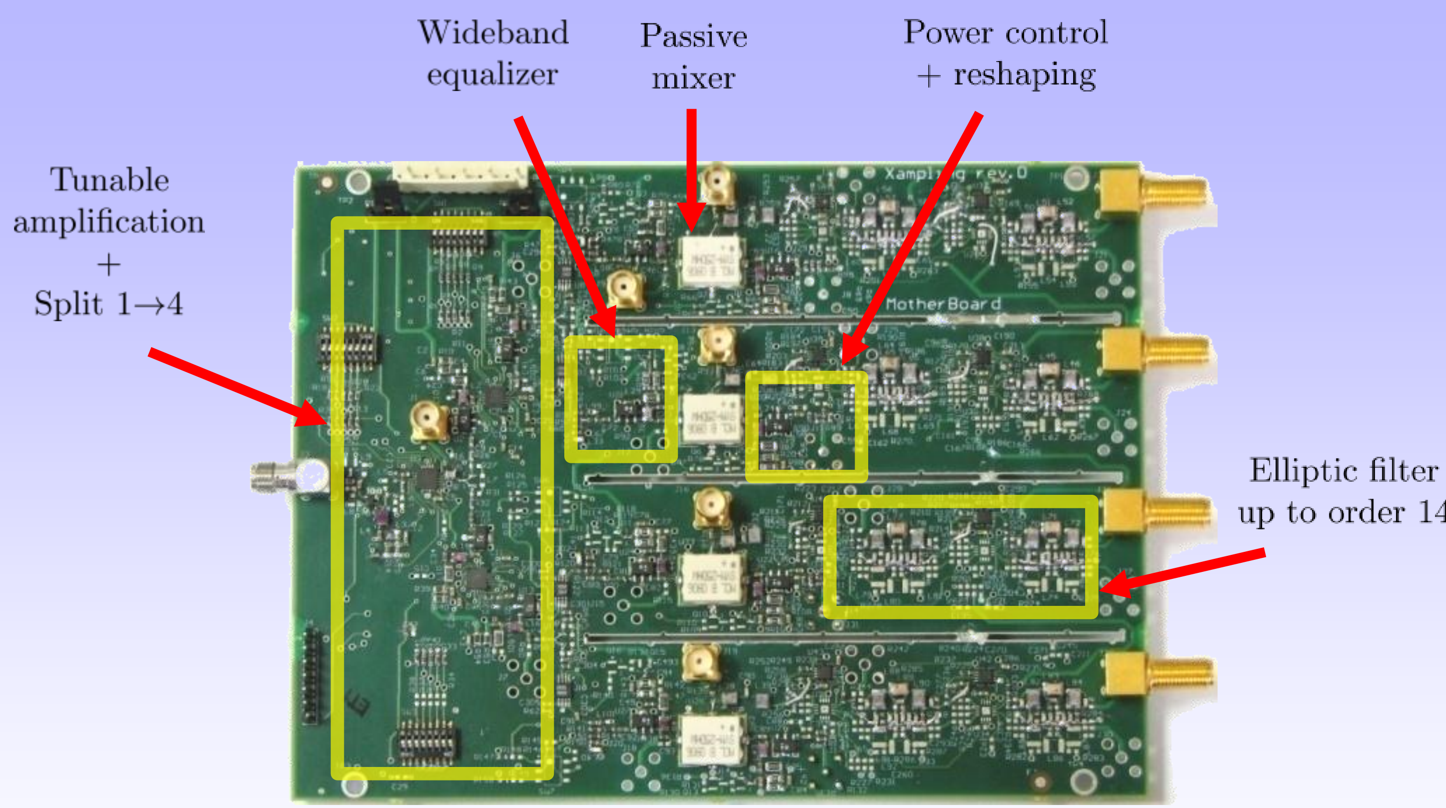


- Wideband receiver mode:
  - Gain control on the input
  - Design specifications:
    Power out > -7 dBm
    SNDR > 30 dB
    over all input range
  - Gives 49 dB dynamic range



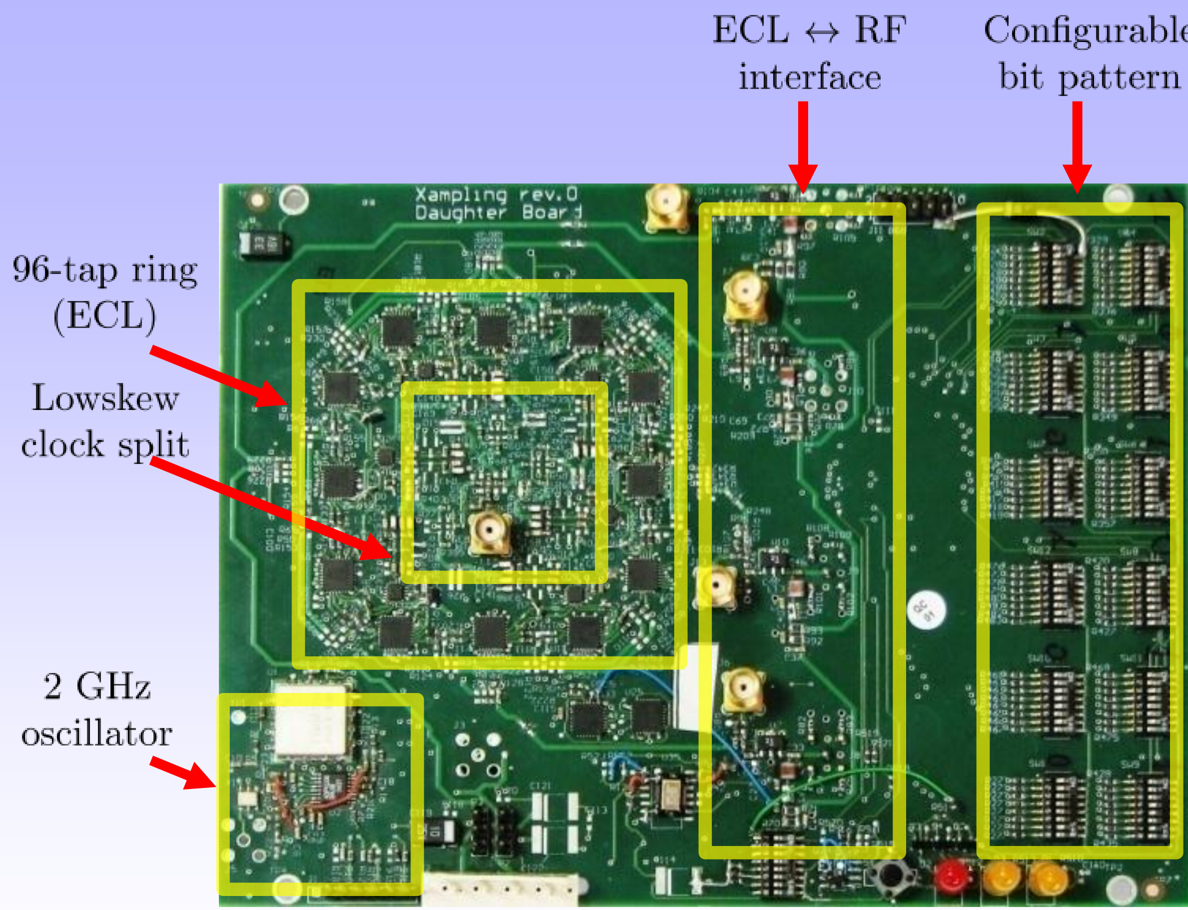
Mishali and Eldar, '08-10

Analog Design



Mishali et al., '10

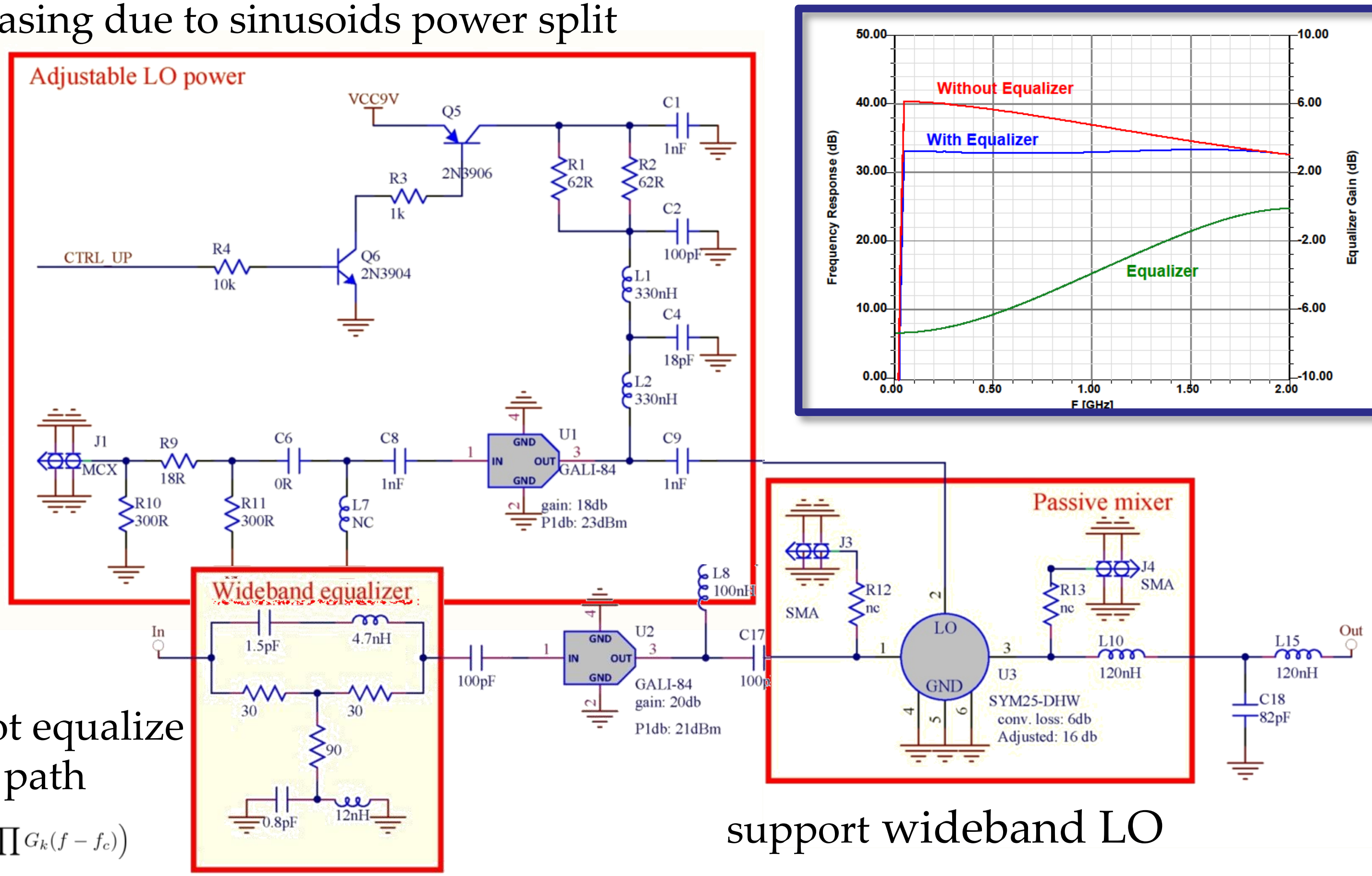
Digital Design



Mishali et al., '10

Mixing with Periodic Functions

Fine biasing due to sinusoids power split



Cannot equalize entire path

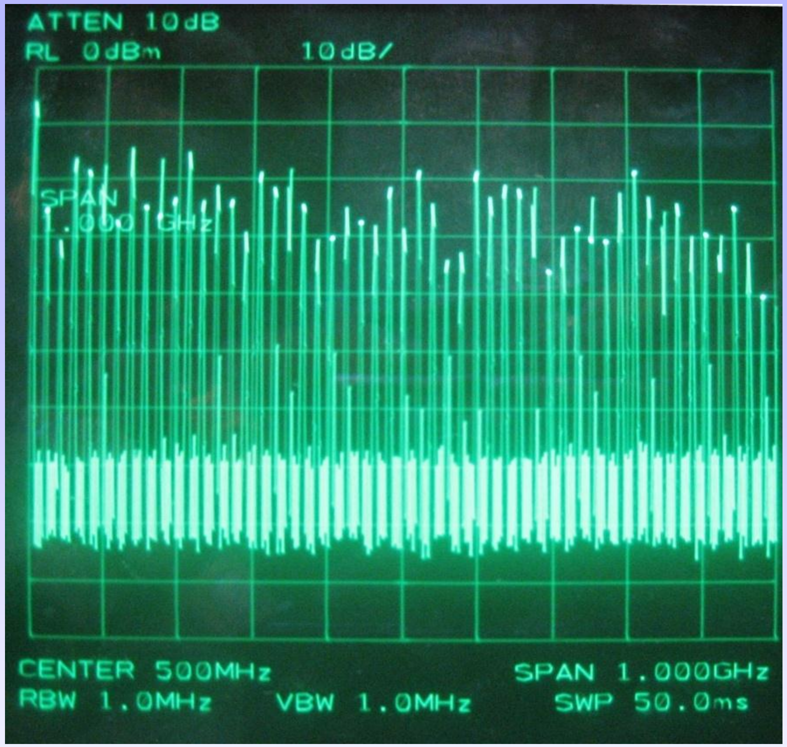
$\left(\prod H_i(f) \prod G_k(f - f_c)\right)$

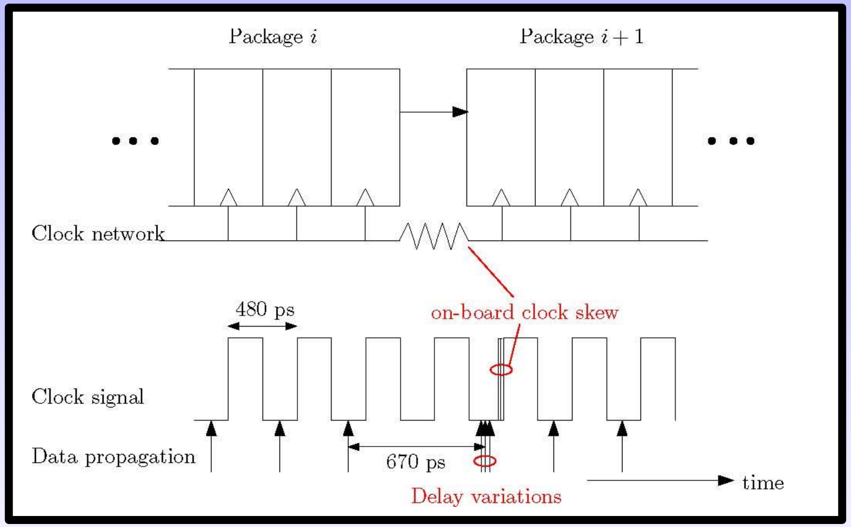
support wideband LO

Datasheet specifications are for single LO mixing (conversion loss, IP3, required power) !

Mishali et al., '10

Highly-Transient Periodic Waveforms



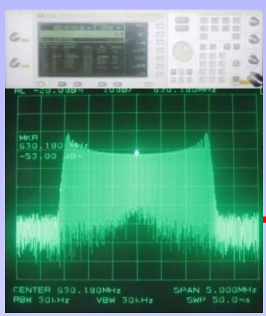


- We selected the sign pattern which gives about the same harmonic levels
- Tap locations: 5$^{\text{th}}$ bit in every consecutive 24 bits (layout considerations only)

Mishali et al., '10

Sub-Nyquist Demonstration

Carrier frequencies are chosen to create overlayed aliasing at baseband



FM @ 631.2 MHz

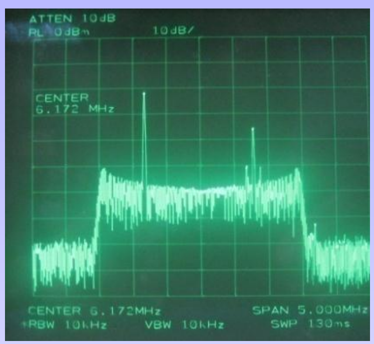
+

AM @ 807.8 MHz

+

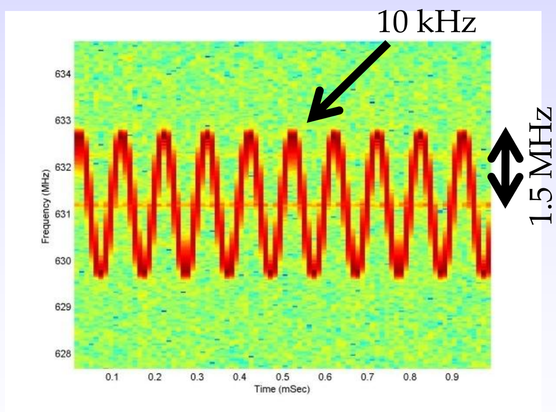
Sine @ 981.9 MHz

MWC prototype

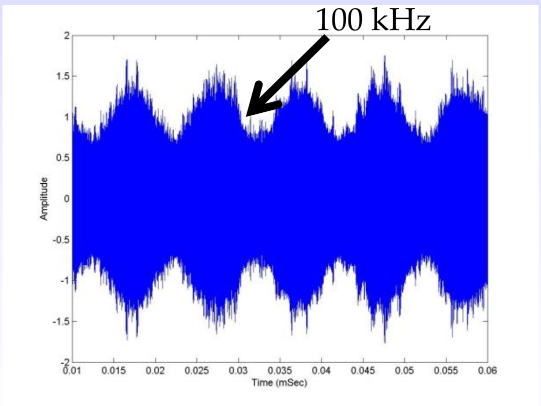


aliasing around 6.171 MHz

Reconstruction
(CTF)



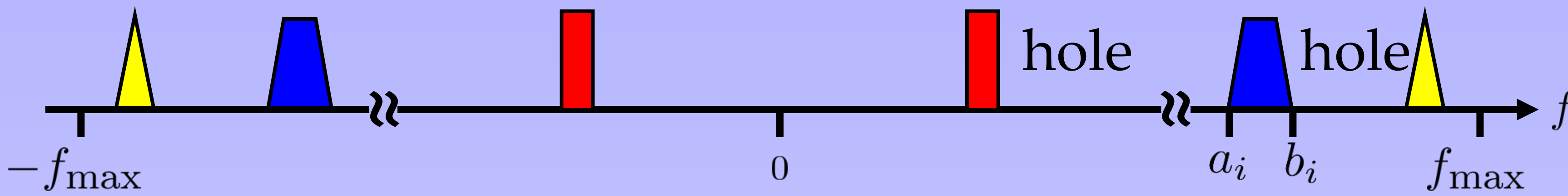
FM @ 631.2 MHz



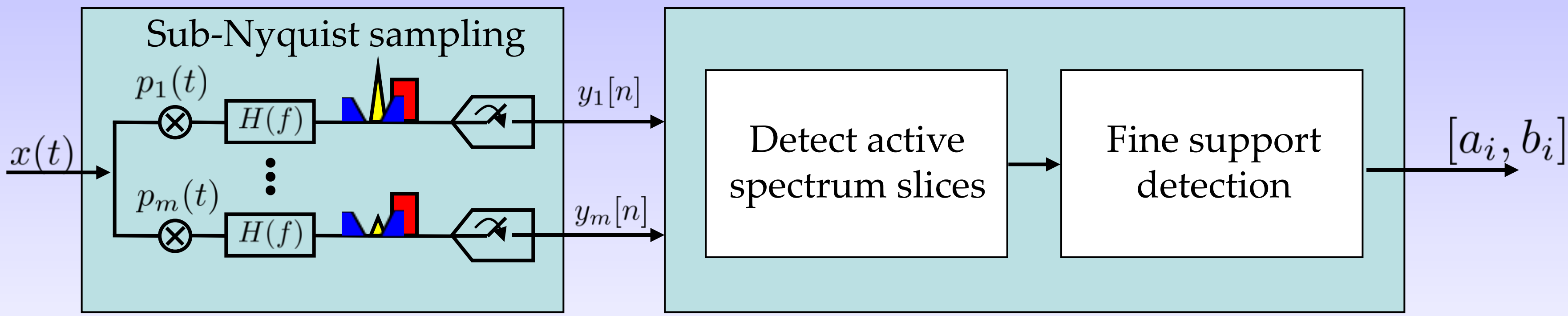
AM @ 807.8 MHz

Mishali et al., '10

Application: Cognitive Radio



Xampling for Spectrum Sensing



- For example:

$m = 4$ channels, sampling rate $= 70$ MHz/channel

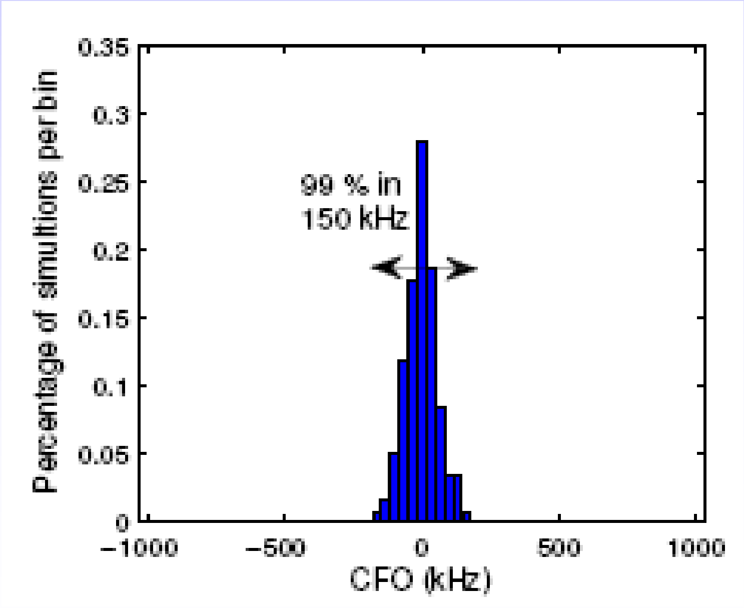
Covers 2 GHz spectrum bandwidth

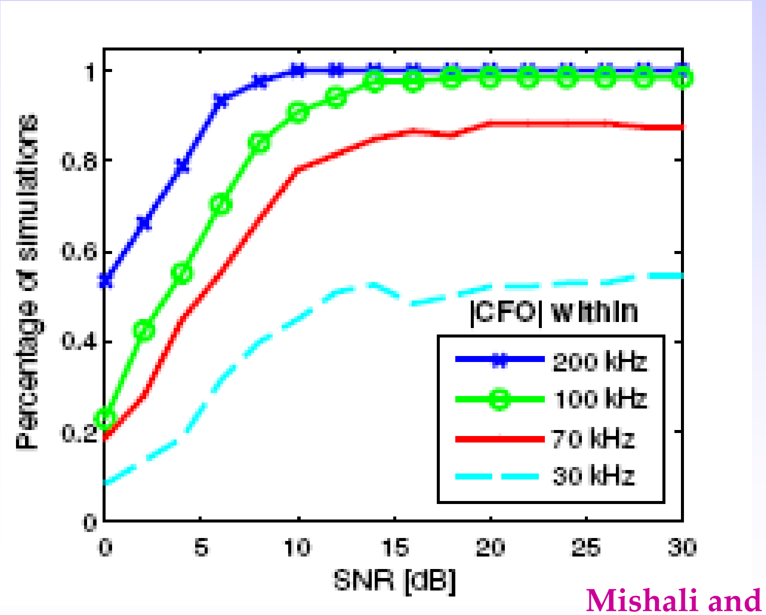
Holes detection up to tens of kHz resolution

Mishali and Eldar, '11

Simulations

- 3 QPSK transmissions, Symbol rate = 30 MHz, $f_{\max} = 5 \text{ GHz}$
- Quality measure, CFO = Carrier frequency offset
- Satisfies IEEE 802.11 40ppm specifications of standard transmissions around 3.75 GHz

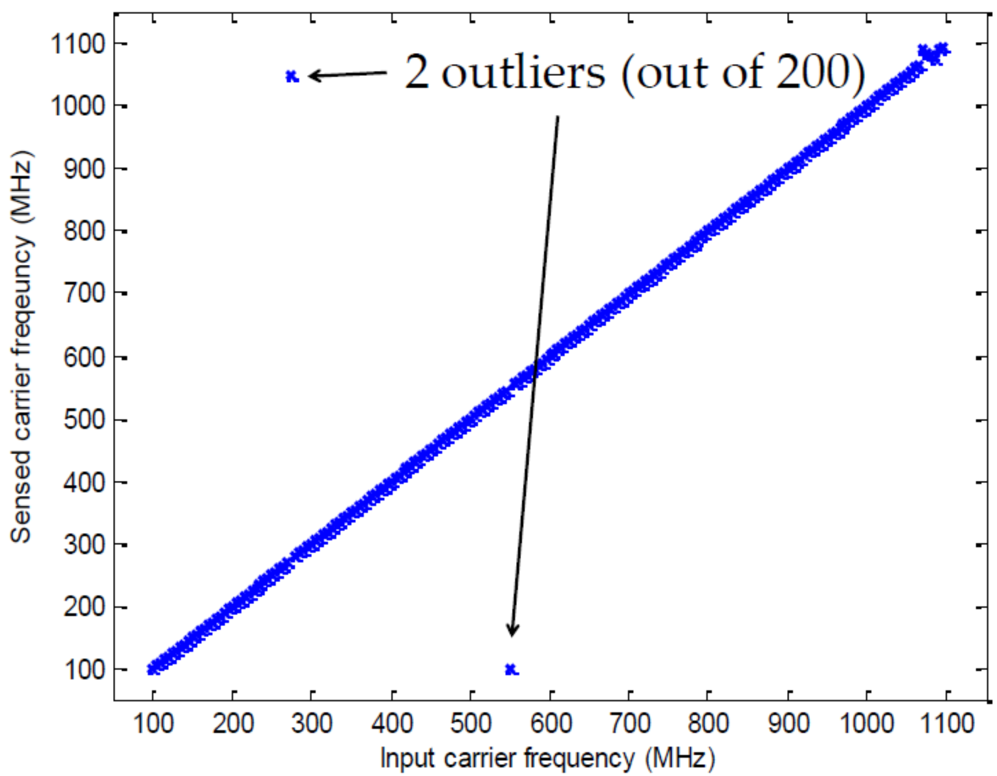




Mishali and Eldar, '11

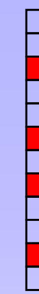
Experiments

Spectrum sensing + carrier recovery of a single sinusoid transmission



Mishali and Eldar, '11

Take-Home Message

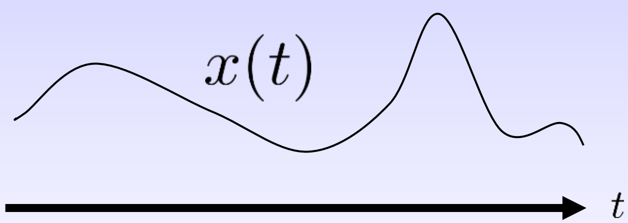


Compressed sensing uses finite models

Xampling works for analog signals

Compression

Sampling



Must combine ideas from Sampling theory and algorithms from CS

- CS+Sampling = Xampling
- X prefix for compression, e.g. DivX

Summary: Next Big Challenge

- Develop cost-effective CS **hardware** solutions
- Address wideband noise and dynamic range
- Integrate into existing hardware solutions
- Innovate at the circuit level: wideband input and large dynamic range
- Design provable hardware
  - at lab
  - on-board
  - on-chip
- Become a mature technology !

Conclusions
Q & A

→ Outline

Conclusions

- Union of subspaces: broad and flexible model
- Can lead to simple and efficient algorithms
- Includes analog signal models
- Sub-Nyquist sampler in hardware
- Compressed sensing of many classes of analog signals
- Many research opportunities: extensions, robustness, hardware, mathematical …

Compressed sensing can be extended practically to the infinite analog domain!

Opinion

- Burst of innovative publications
- Theory is still developing, yet the basic principles are understood

- Next frontier: Hardware implementations
- Become a mature technology !

More details in:

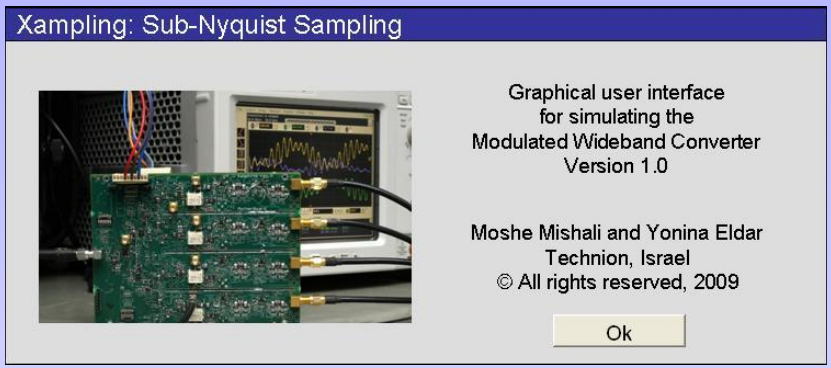
- M. Mishali and Y. C. Eldar, "Sub-Nyquist Sampling: Bridging Theory and Practice,'' *Sig. Proc. Mag.*
- M. Duarte and Y. C. Eldar, "Structured Compressed Sensing: From Theory to Applications," *TSP*.
- M. Mishali and Y. C. Eldar, "Xampling: Compressed Sensing of Analog Signals,'' in book, *Cambridge press.*

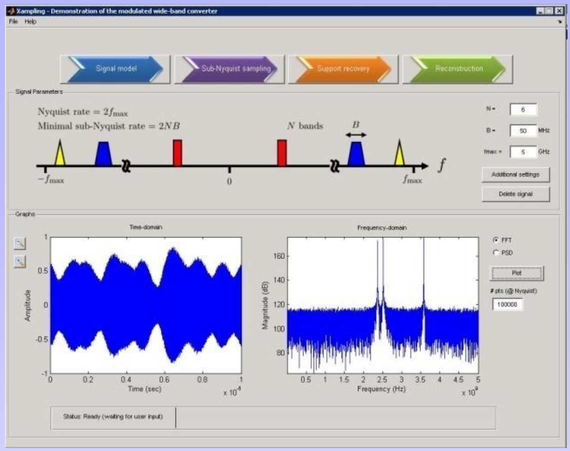
References +
Online Documentations

→ Outline

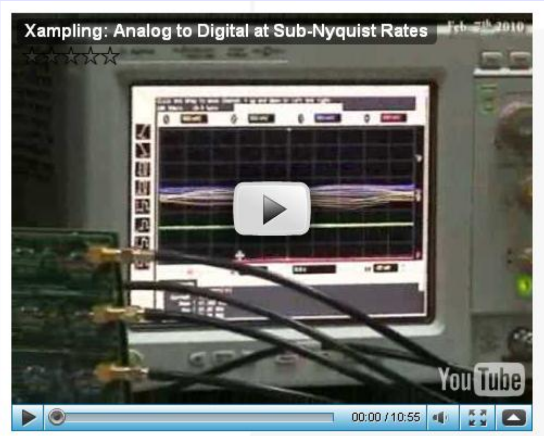
Online Demonstrations

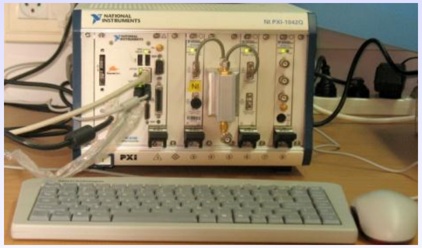
- GUI package of the MWC





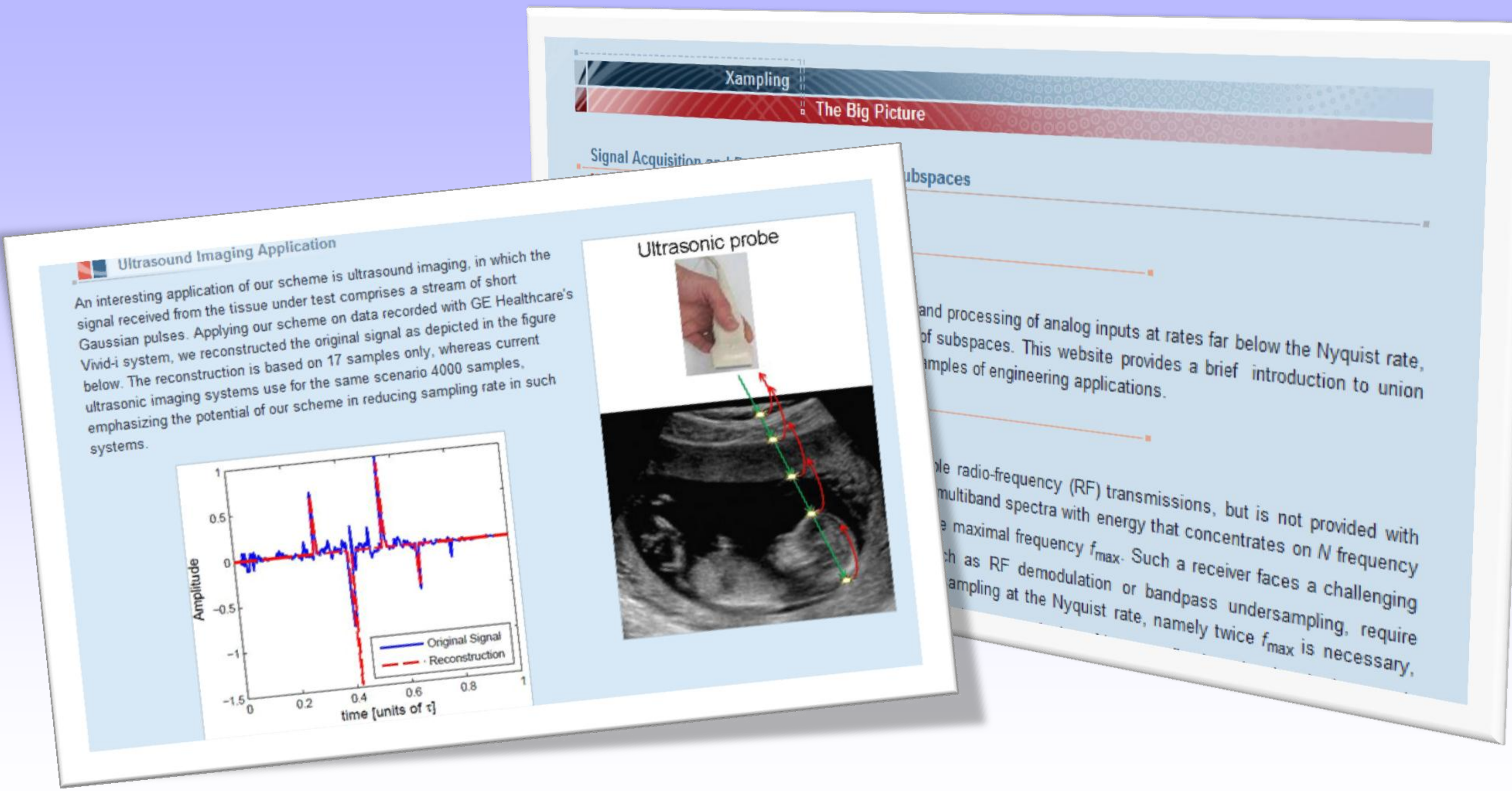
- Video recording of sub-Nyquist sampling + carrier recovery in lab





Xampling Website

webee.technion.ac.il/people/YoninaEldar/xampling_top.html



Acknowledgements

Students:

Moshe Mishali

Kfir Gedalyahu

Ronen Tur

Noam Wagner

Collaborators:

- General Electric Israel – Zvi Friedman
- National Instruments Corp. – Ahsan Aziz, Sam Shearman, Eran Castiel

Sponsors:

- Newcom Network of Excellence
- Israel Science Foundation
- Magneton

Thank you!

We'll be happy to hear your comments, ideas for future work etc:

moshiko@tx.technion.ac.il yonina@ee.technion.ac.il

References

Tutorial:

- M. Mishali and Y. C. Eldar, "Sub-Nyquist Sampling: Bridging Theory and Practice,'' *IEEE Sig. Proc. Mag.*
- M. Duarte and Y. C. Eldar, "Structured Compressed Sensing: From Theory to Applications," *TSP*
- M. Mishali and Y. C. Eldar, "Xampling: Compressed Sensing of Analog Signals,'' in book, *Cambridge press*
- Y. C. Eldar and G. Kutyniok, "Compressed Sensing: Theory and Applications," *Cambridge Press*

Other Tutorials and Summaries:

- R. G. Baraniuk, "Compressive sensing," *IEEE Sig. Proc. Mag.*, vol. 24, no. 4, pp. 118–120, 124, July 2007
- E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Mag.*, vol. 25, pp. 21–30, Mar. 2008
- J. Uriguen, Y. C. Eldar, P. L. Dragotti and Z. Ben-Haim, "Sampling at the Rate of Innovation: Theory and Applications,'' in book, *Cambridge press*

References

Generalized Sampling Theory:

- A. J. Jerry, "The Shannon sampling theorem-Its various extensions and applications: A tutorial review," *Proc. Of the IEEE*, vol. 65, no. 11, pp. 1565-1596, Nov. 1977
- A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. Optimiz.*, vol. 15, pp. 1-21, Feb. 1994
- M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices, "*IEEE Trans. Signal Process*., vol. 42, no. 11, pp. 2915-2925, Nov. 1994
- C. de Boor, R. DeVore and A. Ron, "The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$," *J. Funct. Anal*, vol. 119, no. 1, pp. 37-78, 1994
- A. Aldroubi, "Oblique projections in atomic spaces," *Proc. Amer. Math. Soc.* , vol. 124, no. 7, pp. 2051-2060, 1996
- M. Unser, "Sampling – 50 years after Shannon," *IEEE Proc.*, vol. 88, pp. 569-587, Apr. 2000
- P. P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after Nyquist," *IEEE Trans. Circuit Syst. I*, vol. 48, no. 9, pp. 1094-1109, Sep. 2001
- Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp.
- 48–68, May 2009.
- T. Michaeli and Y. C. Eldar, "Optimization Techniques in Modern Sampling Theory, " in book *Cambridge Univ. Press*, ch., pp. 266–314, 2010

References

Subspace Sampling:

- I. Djokovic and P. P. Vaidyanathan, “Generalized sampling theorem in multiresolution subspaces,” *IEEE Trans. Signal Process.* , vol. 45, pp. 583-599, Mar. 1997
- M. Unser, “Splines: A perfect fit for signal and image processing,” *IEEE Signal Process. Mag.*, pp. 22-38, Nov. 1999
- Y. C. Eldar and A. V. Oppenheim, “Filter bank reconstruction of bandlimited signals from nonuniform and generalized samples,” *IEEE Trans. Signal Processing*, vol. 48, no. 10, pp. 2864-2875, 2000
- A. Aldroubi and K. Gröchenig, “Non-uniform sampling and reconstruction in shift-invariant spaces,” *SIAM Review*, vol. 43, pp. 585-620, 2001
- Y. C. Eldar, “Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors,” *J. Fourier Analys. Appl.*, vol. 1, no. 9, pp. 77-96, Jan. 2003
- O. Christensen and Y. C. Eldar, “Oblique dual frames and shift-invariant spaces,” *Applied and Computational Harmonic Analysis*, vol. 17, no. 1, pp. 48-68, Jul. 2004
- O. Christensen and Y. C. Eldar, “Generalized shift-invariant systems and frames for subspaces,” *J. Fourier Analys. Appl.*, vol. 11, pp. 299-313, 2005
- Y. C. Eldar and T. Werther, “General framework for consistent sampling in Hilbert spaces,” *International Journal of Wavelets, Multiresolution, and Information Processing*, vol. 3, no. 3, pp. 347-359, Sep. 2005
- Y. C. Eldar and O. Christensen, “Caracterization of Oblique Dual Frame Pairs,” *J. Applied Signal Processing*, vol. 2006, Article ID 92674, pp. 1-11
- T. G. Dvorkind, Y. C. Eldar and E. Matusiak, “Nonlinear and non-ideal sampling: Theory and methods,” *IEEE Trans. Signal Processing*, vol. 56, no. 12, pp. 5874-5890, Dec. 2008

References

Multiband subspaces:

- H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, pp. 37-52, Feb. 1967
- A. Kohlenberg, "Exact interpolation of band-limited functions," *J. Appl. Phys.*, pp. 1432–1435, Dec. 1953
- R. G. Vaughan, N. L. Scott, and D. R. White, "The theory of bandpass sampling," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 1973–1984, Sep. 1991
- Y.-P. Lin and P. P. Vaidyanathan, "Periodically nonuniform sampling of bandpass signals," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 3, pp. 340-351, Mar. 1998
- C. Herley and P. W. Wong, "Minimum rate sampling and reconstruction of signals with arbitrary frequency support," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1555-1564, Jul. 1999
- R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2173–2183, Sep. 2000

References

Union of Subspaces:

- Y. M. Lu and M. N. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2334–2345, 2008
- Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Info. Theory*, vol. 55, no. 11, pp. 5302–5316, 2009
- T. Blumensath and M. E. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1872–1882, Apr. 2009

Xampling Framework:

- Y. C. Eldar, "Compressed sensing of analog signals in shift-invariant spaces", *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 2986-2997, August 2009
- Y. C. Eldar, "Uncertainty relations for analog signals," *IEEE Trans. Inform. Theory*, vol. 55, no. 12, pp. 5742 -5757, Dec. 2009
- M. Mishali, Y. C. Eldar, and A. Elron, "Xampling: Signal acquisition and processing in union of subspaces," [Online] arXiv.org 0911.0519, Oct. 2009
- M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits, Devices & Systems*, vol. 5, no. 1, pp. 8–20, Jan. 2011

References

Modulated Wideband Converter / Fully-blind Multi-Coset:

- M. Mishali and Y. C. Eldar, “Blind multiband signal reconstruction: Compressed sensing for analog signals,” *IEEE Trans. Signal Processing*, vol. 57, pp. 993–1009, Mar. 2009
- M. Mishali and Y. C. Eldar, “From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals,” *IEEE Journal of Selected Topics on Signal Processing*, vol. 4, pp. 375-391, April 2010
- M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, “Xampling: Analog to digital at sub-nyquist rates,” *IET Circuits, Devices and Systems*, vol. 5, no. 1, pp. 8–20, Jan. 2011
- M. Mishali and Y. C. Eldar, “Reduce and boost: Recovering arbitrary sets of jointly sparse vectors,” *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008
- M. Mishali and Y. C. Eldar, “Wideband spectrum sensing at sub-Nyquist rates,” to appear in *IEEE Signal Process. Mag.*; [Online] arXiv.org 1009.1305, Sep. 2010

Random Demodulator:

- J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, “Beyond Nyquist: Efficient sampling of sparse bandlimited signals,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010
- Z. Yu, S. Hoyos, and B. M. Sadler, “Mixed-signal parallel compressed sensing and reception for cognitive radio,” in *ICASSP*, 2008, pp. 3861–3864
- T. Ragheb, J. N. Laska, H. Nejati, S. Kirolos, R. G. Baraniuk, and Y. Massoud, “A prototype hardware for random demodulation based compressive analog-to-digital conversion,” in *Circuits and Systems*, 2008. MWSCAS 2008. 51st Midwest Symposium on, 2008, pp. 37–40
- Z. Yu, X. Chen, S. Hoyos, B. M. Sadler, J. Gong, and C. Qian, “Mixed-signal parallel compressive spectrum sensing for cognitive radios,” *International Journal of Digital Multimedia Broadcasting*, 2010

References

Pulse streams:

- M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, 2002
- P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang Fix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007
- J. Kusuma and V. Goyal, "Multichannel sampling of parametric signals with a successive approximation property," *IEEE Int. Conf. Image Procesing (ICIP)*, pp. 1265–1268, Oct. 2006
- C. Seelamantula and M. Unser, "A generalized sampling method for finite-rate-of-innovation-signal reconstruction," *IEEE Signal Process. Lett.*, vol. 15, pp. 813–816, 2008
- E. Matusiak and Y. C. Eldar, "Sub-Nyquist sampling of short pulses: Theory," [Online] arXiv.org 1010.3132
- Z. Ben-Haim, T. Michaeli, and Y. C. Eldar, "Performance bounds and design criteria for estimating finite rate of innovation signals," [Online] arXiv.org 1009.2221, Sep. 2010
- L. Baboulaz and P. L. Dragotti, "Exact feature extraction using finite rate of innovation principles with an
- application to image super-resolution," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 281–298, Feb. 2009
- K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3017–3031, June 2010
- R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011
- K. Gedalyahu, R. Tur, and Y. C. Eldar, "Multichannel sampling of pulse streams at the rate of innovation," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1491–1504, Apr. 2011
- W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, "Identification of underspread linear systems with application to super-resolution radar," to appear in *IEEE Trans. Signal Processing*

References

Compressed sensing (#1):

- D. L. Donoho, “Compressed sensing,” *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sep. 2006
- E. J. Candès, J. K. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Info. Theory*, vol. 52, no. 2, pp. 489–509, 2006
- E. J. Candès and T. Tao, “Near optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Info. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006
- J. A. Tropp, “Greed is good: Algorithmic results for sparse approximation,” *IEEE Trans. Info. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004
- D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2008
- J. A. Tropp, A. C. Gilbert, and M. J. Strauss, “Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit,” *Signal Processing*, vol. 86, pp. 572–588, Apr. 2006
- J. A. Tropp, “Algorithms for simultaneous sparse approximation. Part II: Convex relaxation,” *Signal Processing*, vol. 86, Apr. 2006
- D. L. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via `1 minimization,” *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003
- Z. Ben-Haim and Y. C. Eldar, "The Cramér–Rao bound for estimating a sparse parameter vector," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3384–3389, June 2010
- I. F. Gorodnitsky and B. D. Rao, “Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm,” *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 600–616, Mar. 1997

References

Compressed sensing (#2):

- Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Processing*, vol. 58, no. 10, pp. 5030–5043, Oct. 2010
- S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2477–2488, July 2005
- J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006
- M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008
- S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993
- Y. Pati, R. Rezaifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1993
- M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," Apr. 2010, *Preprint*
- D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," *Tech. Rep. TREE-0612*, Rice University, Department of Electrical and Computer Engineering, Houston, TX, Nov. 2006

Thank you