Sub-Nyquist Sampling without Sparsity and Phase Retrieval

Yonina Eldar

Department of Electrical Engineering
Technion – Israel Institute of Technology

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

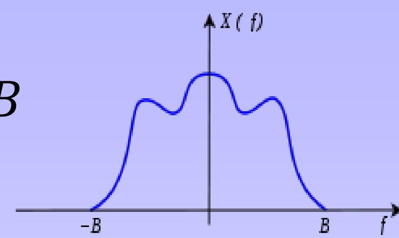
Signal Acquisition Modeling and Processing Lab

London Workshop on Sparse Signal Processing

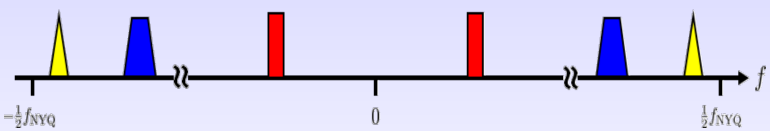
Traditional Sampling

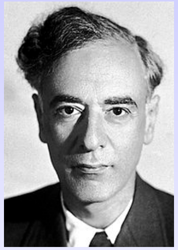
Sampling rate required in order to recover $x(t)$ from its samples

- Shannon-Nyquist theorem:
  - Bandlimited signal with bandwidth $2B$
  - Minimal sampling rate: $f_{Nyq} = 2B$



- Landau rate:
  - Multiband signal with known support of measure Λ
  - Minimal sampling rate: Λ



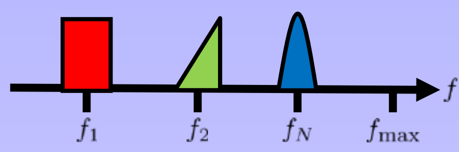


- Extension to arbitrary subspaces:
  - Signal in a subspace with dimension D requires sampling at rate D
  - Shift-invariant subspaces $\sum_{n\in\mathbb{Z}} a[n]h(t-nT)$ require sampling at rate $1/T$

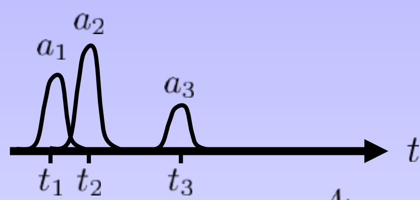
Sampling of Structured Signals

Exploit analog structure to reduce sampling rate

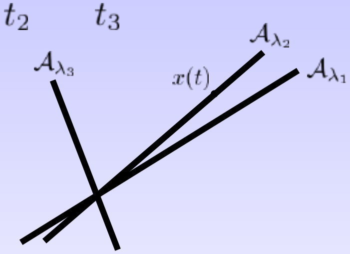
- Multiband signal with unknown support of measure Λ
  Minimal sampling rate: 2Λ **(Mishali and Eldar '09)**



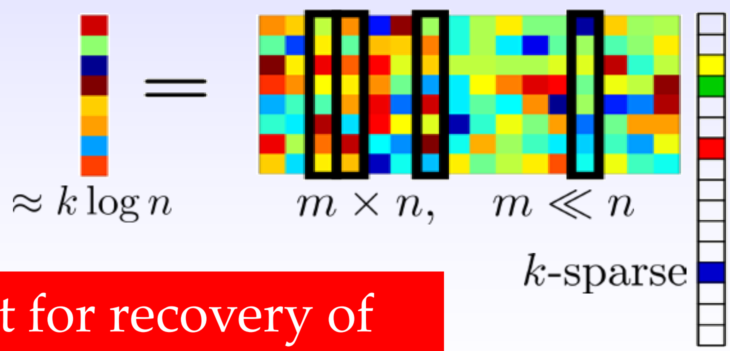
- Stream of k pulses (finite rate of innovation)
  Minimal sampling rate: 2k **(Vetterli, Dragotti et. al '02)**



- Union of subspaces **(Lu and Do '08, Mishali and Eldar '09)**



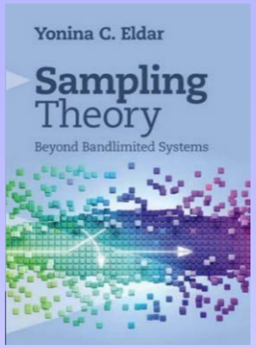
- Sparse vectors
  (Candes, Romberg, Tau '06, Donoho '06)

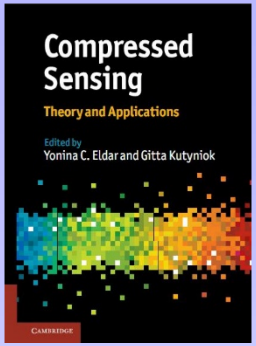


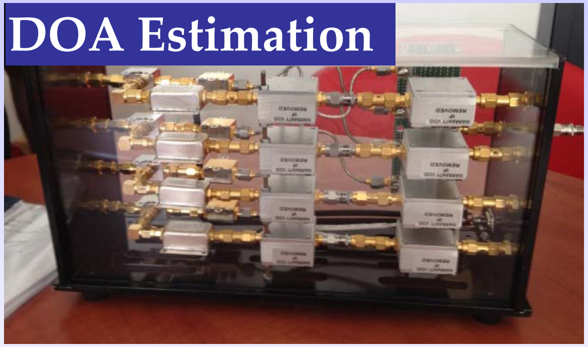
Sampling rate below Nyquist for recovery of $x(t)$ by exploiting structure

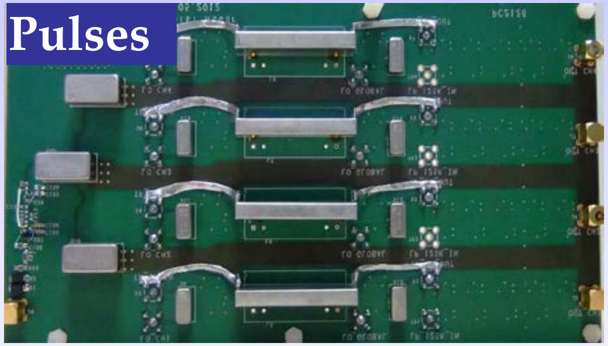
Sub-Nyquist Sampling

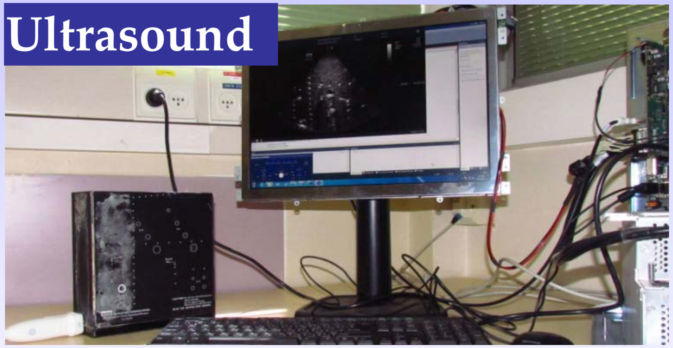
- Many examples in which we can reduce sampling rate by exploiting structure
- Xampling: practical sub-Nyquist methods which allow low-rate sampling and low-rate processing in diverse applications

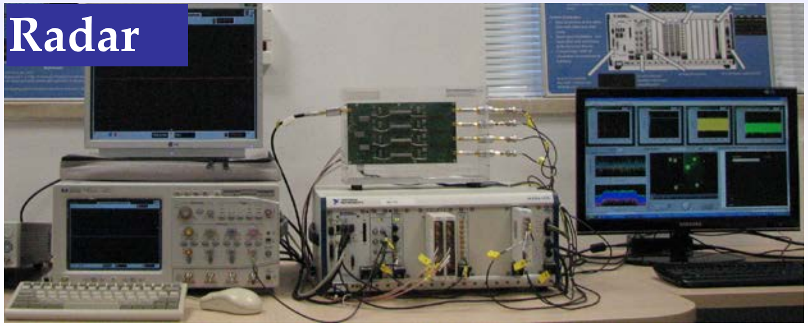


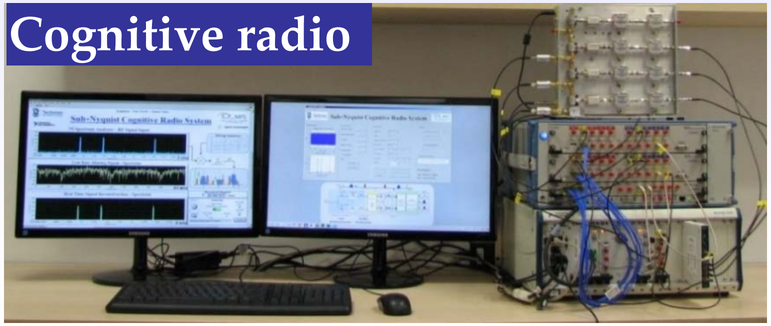



DOA Estimation

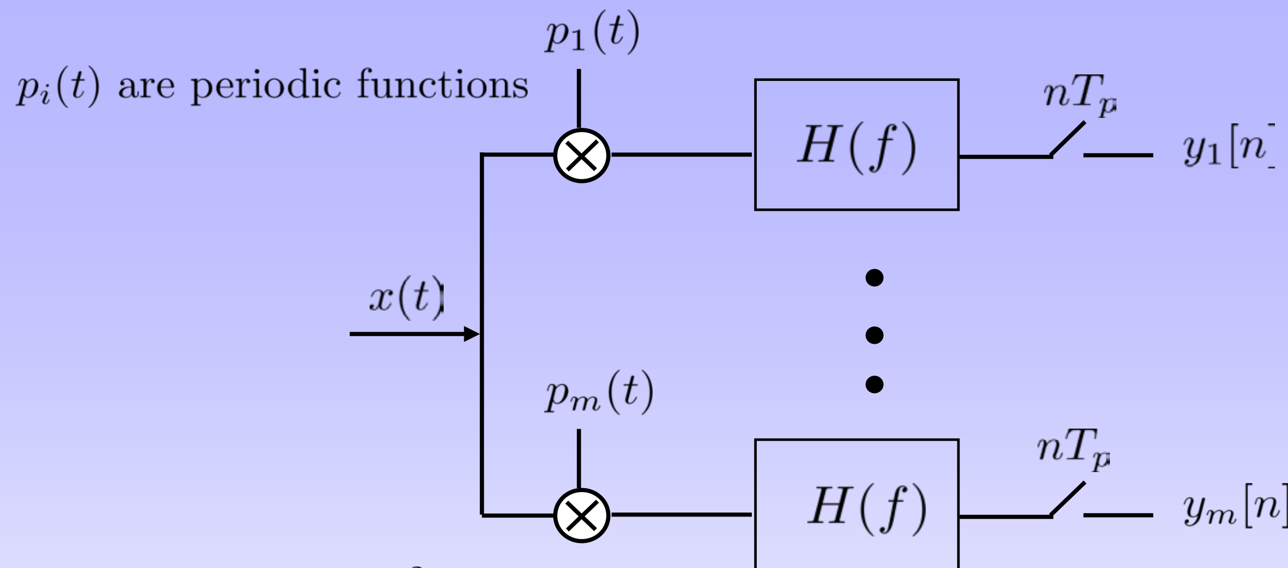

Pulses

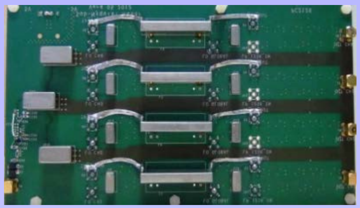

Ultrasound


Radar


Cognitive radio

Xampling Hardware





- $p_i(t) = \sum a_{in} e^{-j\frac{2\pi}{T_p}nt}$ sums of exponentials
- The filter *H(f)* shapes the tones and reduces bandwidth
- The channels can be collapsed to a single channel

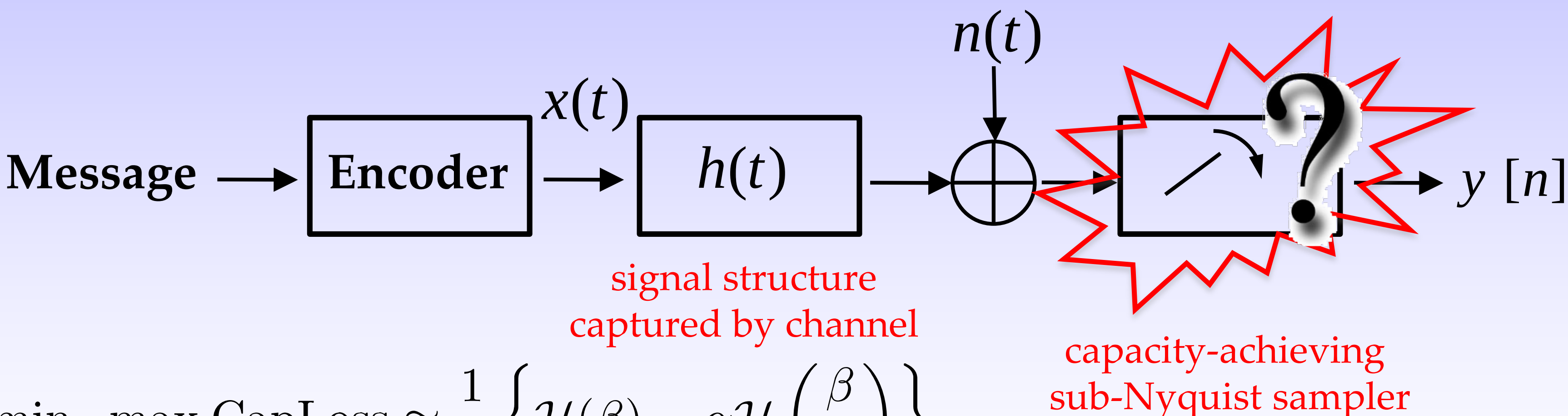
x(t) →  → Acquisition → Compressed sensing and processing → recovery

Analog preprocessing

Low rate (bandwidth)

Optimality of Xampling Hardware

- Achieves the Cramer-Rao bound for analog recovery given a sub-Nyquist sampling rate **(Ben-Haim, Michaeli, and Eldar 12)**
- Minimizes the worst-case capacity loss for a wide class of signal models **(Chen, Eldar and Goldsmith 13)**
- Capacity provides further justification for the use of random tones



$$\min_{\text{sampler}} \max_{\text{states}} \mathrm{CapLoss} \approx \frac{1}{2}\left\{\mathcal{H}(\beta) - \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right)\right\}$$

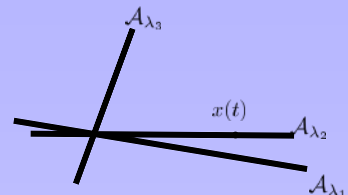
$$\mathcal{H}(\beta) = -\beta\log\beta - (1-\beta)\log(1-\beta)$$

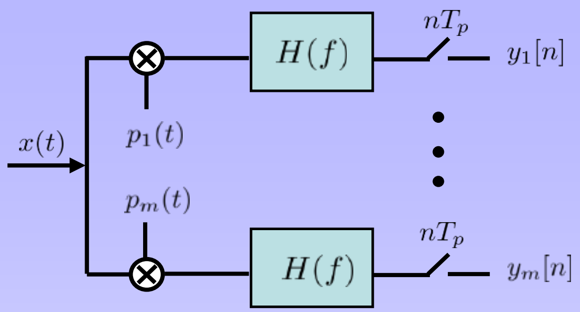
binary entropy function

- α: undersampling factor
- β: sparsity ratio

Sampling of Structured Signals

- Nonlinear prior





- Careful design of measurement scheme
- Typically non-linear recovery methods
- Often nonlinear processing needs to be accounted for (such as beamforming, quantization etc.)

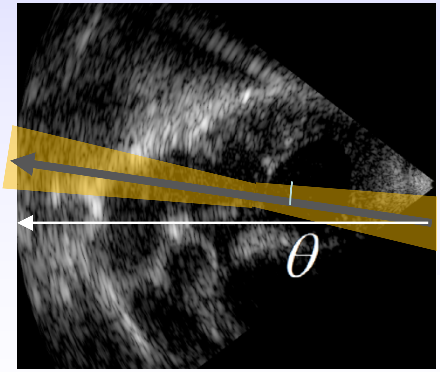
Extensions:

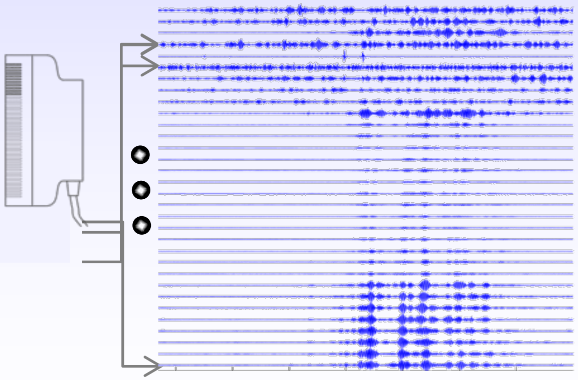
- Is structure necessary for sub-Nyquist sampling?
- Can careful measurement design and optimization-based recovery methods help in other nonlinear problems?

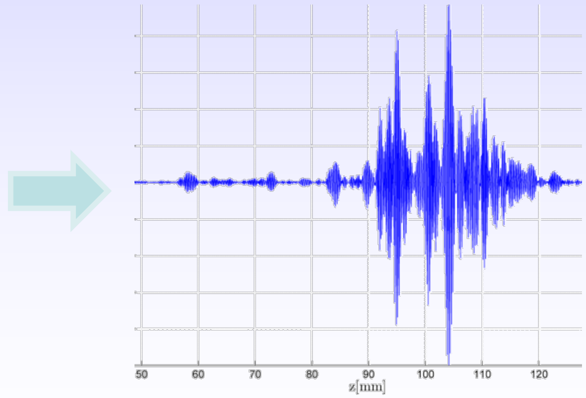
Sub-Nyquist Without Structure

Can we reduce sampling rates when the signals do not have structure?

- Goal: Recovery of some function of the signal
  - Signal statistics: Power spectrum estimation **with Geert Leus and Deborah Cohen**
  - Quantized version of the signal **with Andrea Goldsmith and Alon Kipnis**
  - Sampling of a set of signals that are used for beamforming where the beamformed signal has structure **with Tanya Chernyakova**





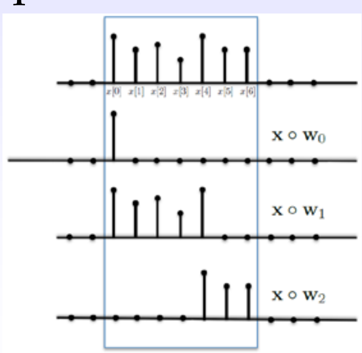


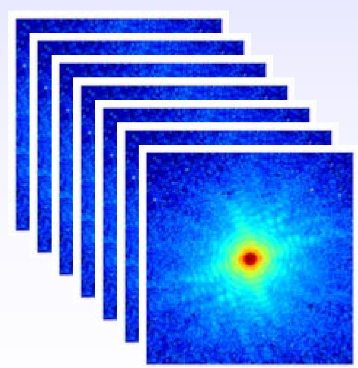
Measurement Design for Phase Retrieval

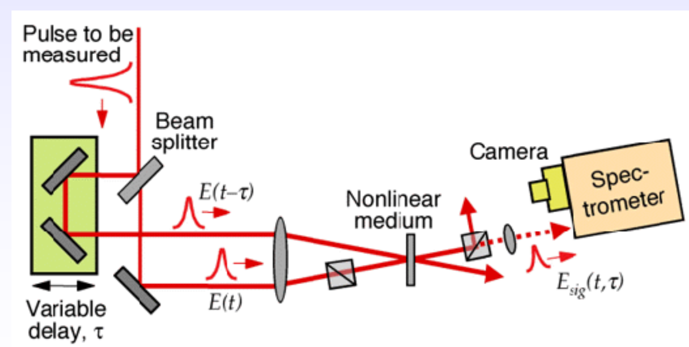
Can we design measurement schemes to enable phase retrieval from Fourier measurements?

- Goal: Recover signals from their Fourier magnitude
  - Known to be impossible for 1D problems
  - No known stable methods for 2D problems
  - Recent methods rely on random measurements rather than Fourier $|\langle a_k, x\rangle|^2$
  - Proper design of deterministic Fourier measurements together with optimization methods allows for recovery even in 1D problems!

$x[n] \rightarrow$ Fourier + Absolute value $\rightarrow |X[k]|^2$







Talk Outline

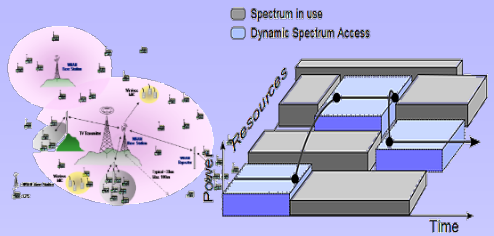
- Power spectrum estimation from sub-Nyquist samples
- Rate-distortion theory of sampled signals
  - Unify sampling theory and rate distortion theory
  - Optimal distortion at sub-Nyquist rates
- Sub-Nyquist beamforming in ultrasound
  - Beamforming at sub-Nyquist rates
  - Wireless ultrasound
- Phase retrieval from Fourier measurements

Part 1:

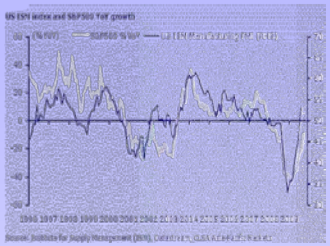
Xampling Without Structure

Power Spectrum Reconstruction

- Sometimes reconstructing the covariance rather than the signal itself is enough:
  - Support detection
  - Statistical analysis
  - Parameter estimation (e.g. DOA)



Cognitive Radios


Financial time Series analysis

What is the minimal sampling rate to estimate the signal covariance?

- Assumption: Wide-sense stationary ergodic signal
- If all we want to estimate is the covariance then we can substantially reduce the sampling rate even without structure!

Covariance Estimation

Cohen, Eldar and Leus 15

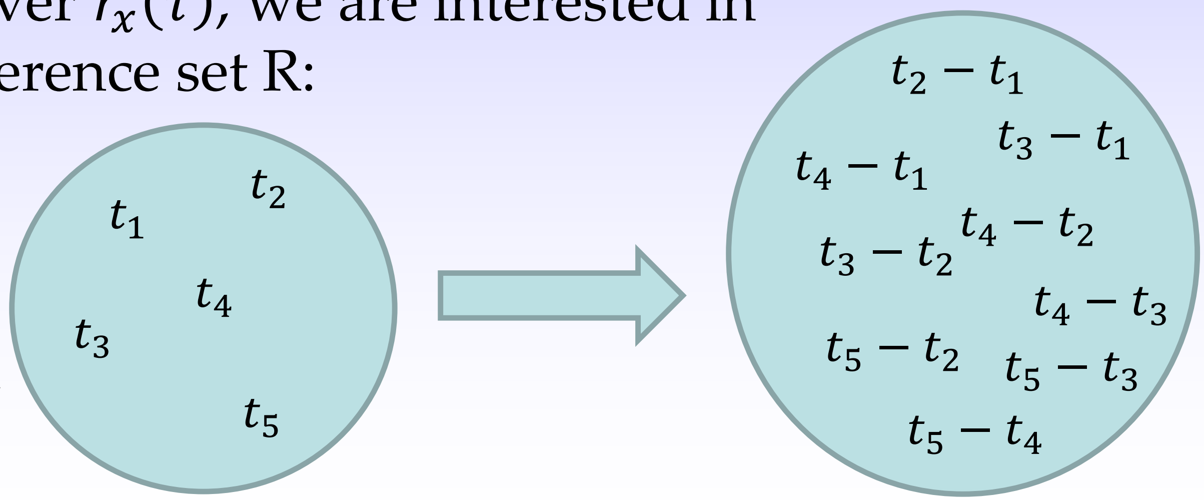
- Let $x(t)$ be a wide-sense stationary ergodic signal
- We sample $x(t)$ with a stable sampling set at times $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$
- We want to estimate $r_x(\tau) = \mathbb{E}[x(t)\mathrm{x}(t-\tau)]$

What is the minimal sampling rate to recover $r_x(\tau)$?

- Sub-Nyquist sampling is possible!

Intuition:

- The covariance $r_x(\tau)$ is a function of the time lags $\tau = t_i - t_j$
- To recover $r_x(\tau)$, we are interested in the difference set R:



Sampling set
$\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$

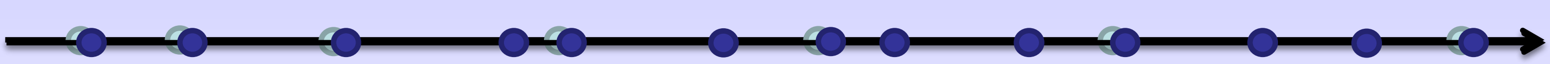
Difference set
$R = \{t_i - t_j\}_{i,j\in\mathbb{Z}}$
$t_i > t_j$

Difference Set Density

It is possible to create sampling sets with Beurling density 0 for which the difference set has Beurling density ∞!

- There should be enough distinct differences so that the size of the difference set goes like the square of the size of the sampling set
- The density of the set should go to 0 slower than the square root

⇨ the density of the square (difference set) goes to ∞



$\tilde{R}$

Theorem

Let $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$, be a sampling set with lower Beurling density $D^-(\tilde{R}) = 0$, so that the set of differences between two sets of size p and q is of the order of pq. Let $R = \{t_i - t_j\}, \forall t_i > t_j \in \tilde{R}$ be the associated difference set. If $\lim_{r\to\infty} \frac{d_{\tilde{R}}(r)}{\sqrt{r}} = \infty$, then, $D^-(R) = \infty$

Universal Minimal Sampling Rate

Under the previous conditions on the sampling set, we can reconstruct $r_x(\tau)$ from $\{x(t_i)\}_{i\in\mathbb{Z}}$

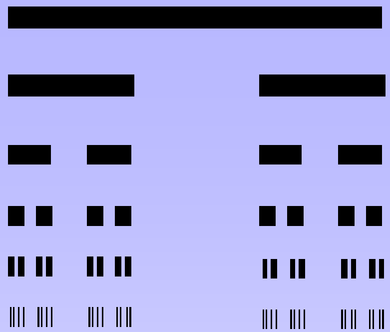
We can reconstruct the covariance from signal samples with density 0!

Theorem

Let $x(t)$ be a wide-sense stationary ergodic signal. Let $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$, be a sampling set with lower Beurling density $D^{-}(\tilde{R}) = 0$, so that $\lim_{r\to\infty} \frac{d_{\tilde{R}}(r)}{\sqrt{r}} = \infty$ and the set of differences between two sets of size p and q is of the order of pq. Then, $r_x(\tau)$ can be perfectly recovered from the samples $x(t_i), i \in \mathbb{Z}$

Sampling Sets Examples

- Cantor ternary set: repeatedly delete the open middle third of a set of line segments, starting with the interval [0,1]
  - Sampling set: $D^{-}\left(\tilde{R}_C\right) \longrightarrow 0$
  - Difference set: $D^{-}(R_C) \longrightarrow \infty$ (both conditions hold)

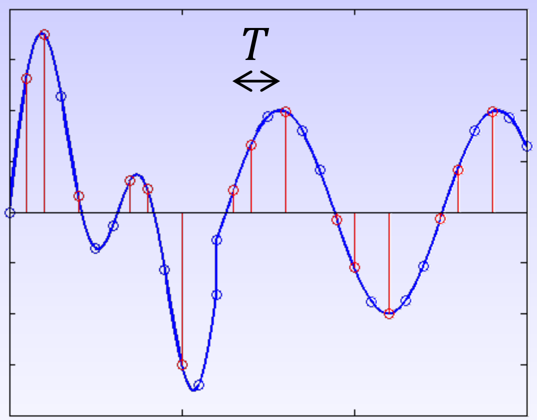


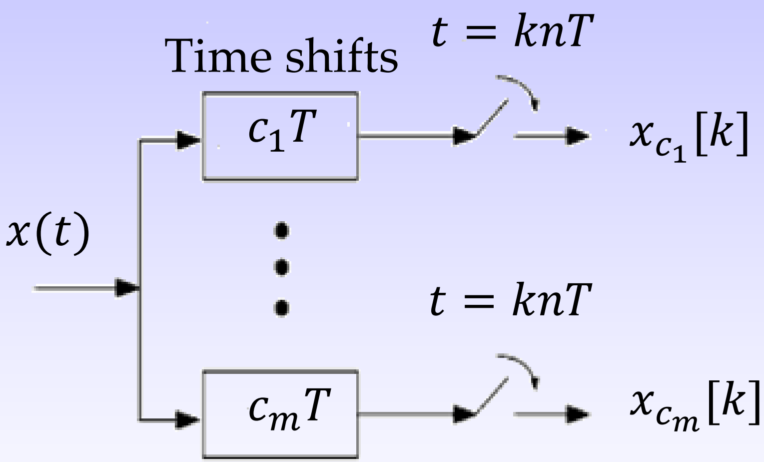
- Uniform sampling: let $\tilde{R}_U = \{kT\}_{k\in\mathbb{Z}}$ be a uniform sampling set spaced by T. It holds that $R_U = \tilde{R}_U$. If $T \longrightarrow \infty$, then
  - Sampling set: $D^{-}\left(\tilde{R}_U\right) \longrightarrow 0$
  - Difference set: $D^{-}(R_U) \longrightarrow 0$ (not enough distinct differences)

Can we analyze practical sampling sets with positive Beurling density?

Multicoset Sampling

- Practical sampling set with finite rate
  - Divide the Nyquist grid into blocks of n consecutive samples (cosets)
  - Keep m samples from each block
  - Sampling set: $D^{-}(\tilde{R}) = \frac{m}{nT}$





What is the minimal sampling rate for perfect covariance recovery from multicoset samples with n cosets?

Multicoset – Bandlimited Signal

Theorem

Let $x(t)$ be bandlimted with bandwidth $1/T$. The minimal rate for perfect recovery of $r_x(\tau)$ when using multicoset sampling with n channels is given by

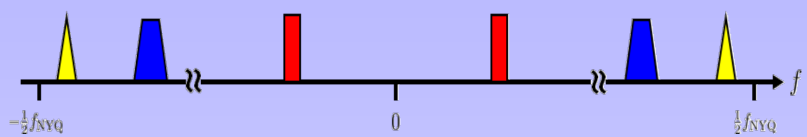
$$\frac{m}{nT} \geq \frac{1+\sqrt{4n-3}}{2nT} \sim \frac{1}{\sqrt{n}T}$$

Signal recovery: $m \geq n$
Covariance recovery: $m \gtrsim \sqrt{n}$

- Achieved when the differences between two distinct cosets are unique, namely $c_i - c_j \neq c_k - c_l, \forall i \neq k, j \neq l$
- Known as the Golomb ruler
- Sparse ruler special case when sampling on the Nyquist grid

Multicoset – Sparse Signal

- Let $x(t)$ be sparse with unknown support with occupancy $\varepsilon < {}^{1}/_{2}$



- Minimal sampling rate for signal recovery: ${}^{2\varepsilon}/_{T}$ **(Mishali and Eldar '09)**

Theorem

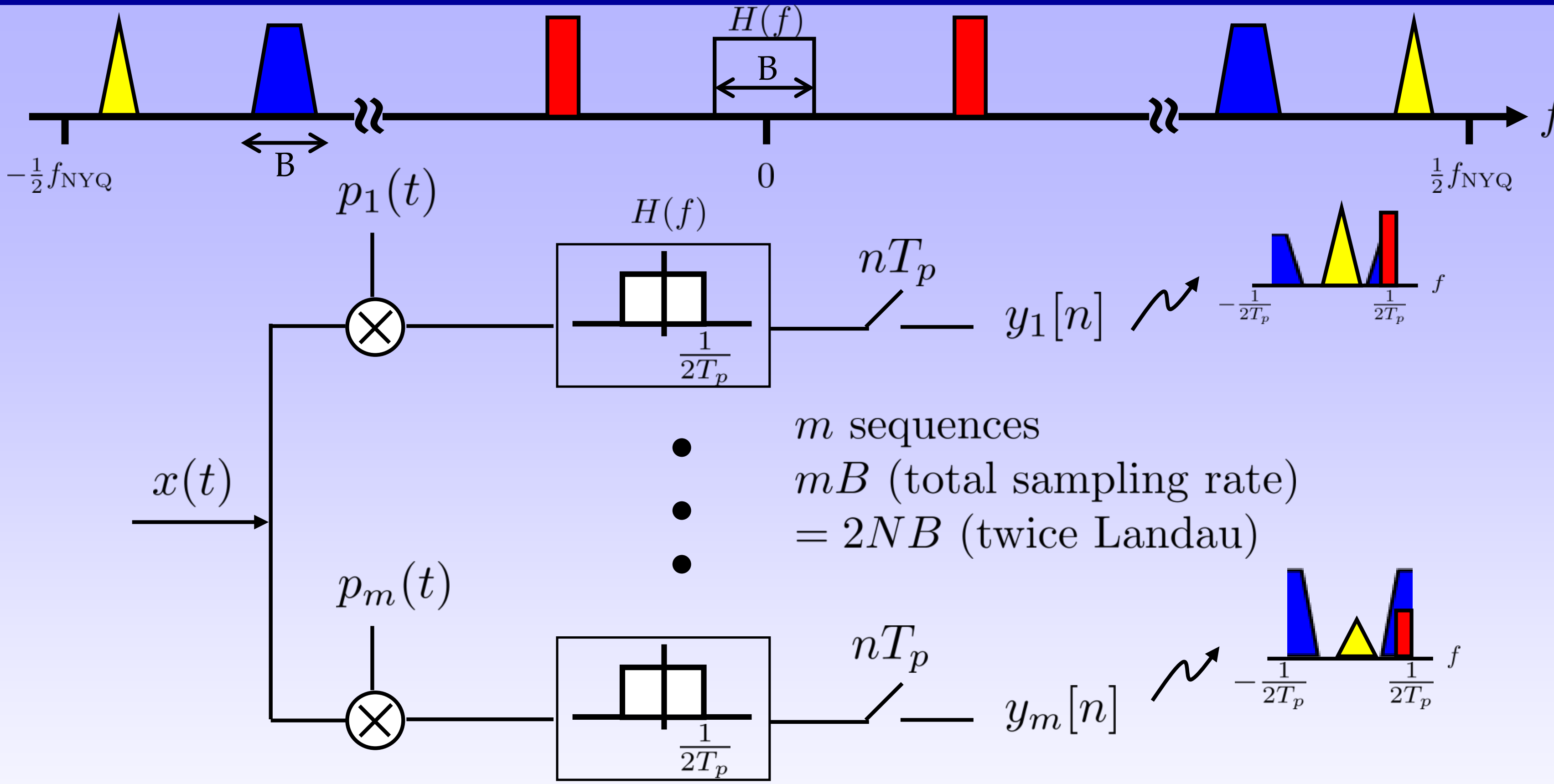
Let $x(t)$ be sparse with unknown support with occupancy $\epsilon < 1/2$. The minimal rate for perfect recovery of $r_x(\tau)$ when using multicoset sampling with n channels is given by

$$\frac{m}{nT} \geq \frac{1+\sqrt{8\epsilon n - 3}}{2nT} \approx \frac{\sqrt{2\epsilon}}{\sqrt{n}T}$$

Signal recovery: $m \geq 2\varepsilon n$
Covariance recovery: $m \gtrsim \sqrt{2\varepsilon n}$

The Modulated Wideband Converter

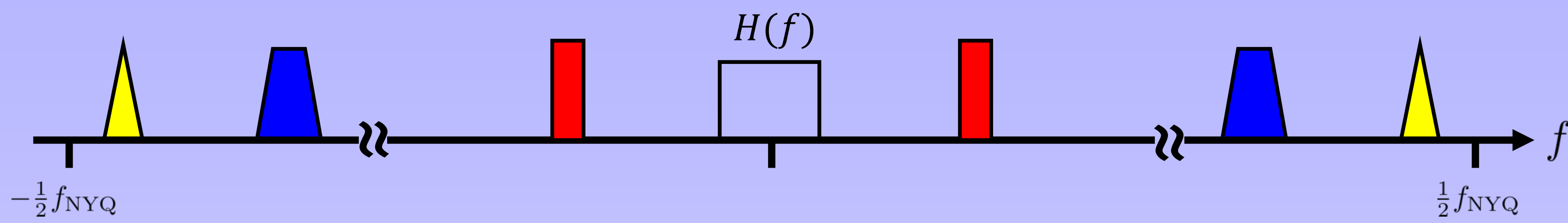
Mishali and Eldar, 11

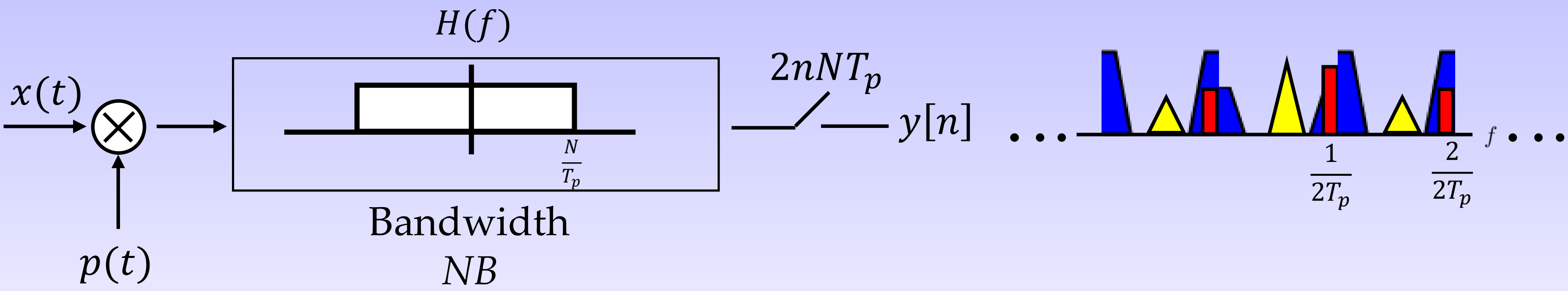


T_p–periodic $p_i(t)$ gives the desired aliasing effect

Single Channel Realization

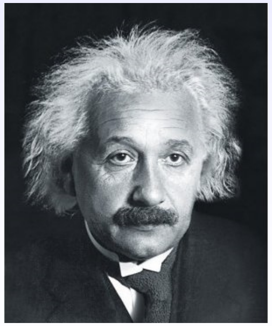
Mishali and Eldar, 11





- The MWC does not require multiple channels
- Does not need accurate delays
- Does not suffer from analog bandwidth issues

Application: Cognitive Radio

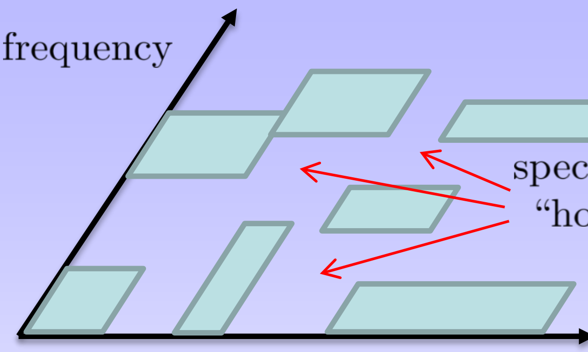


"In theory, theory and practice are the same.
In practice, they are not."

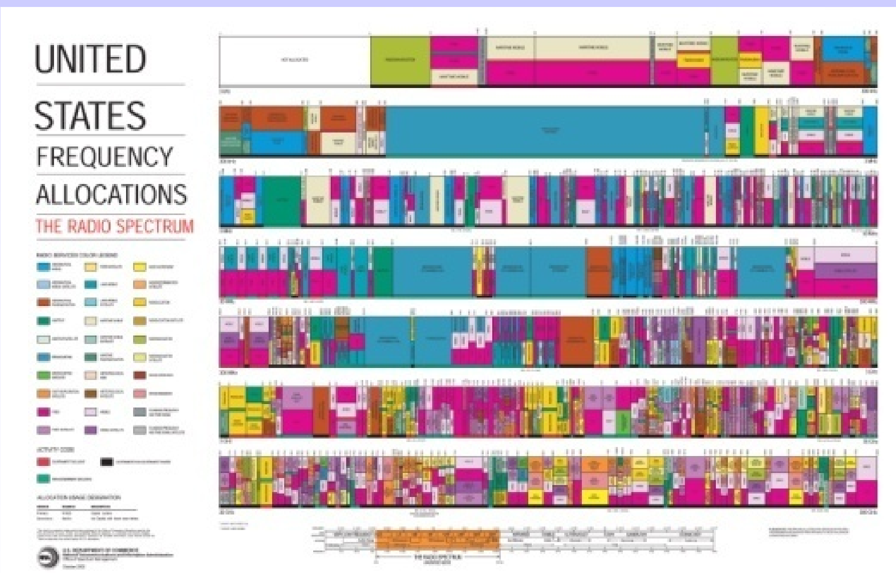
Albert Einstein

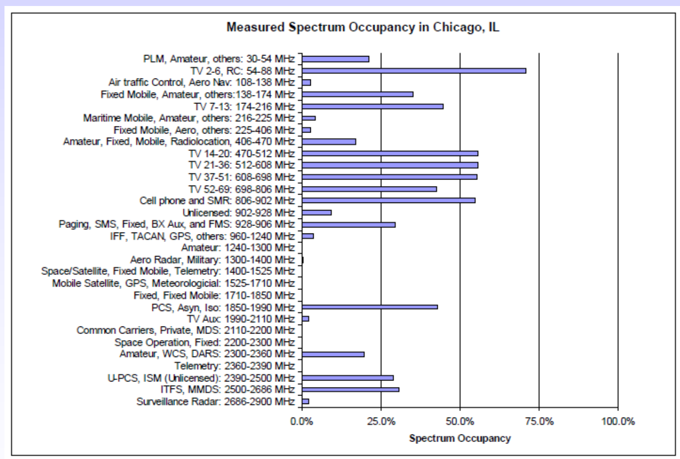
Cognitive Radio

- Cognitive radio mobiles utilize unused spectrum ``holes''
- Need to identify the signal support at low rates



Federal Communications Commission (FCC) frequency allocation





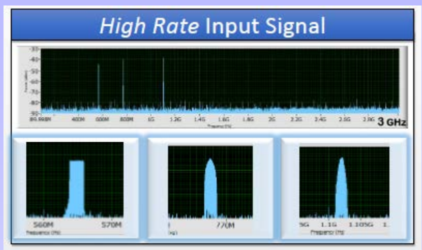
Shared Spectrum Company (SSC) – 16-18 Nov 2005

Licensed spectrum highly underused: E.g. TV white space, guard bands and more

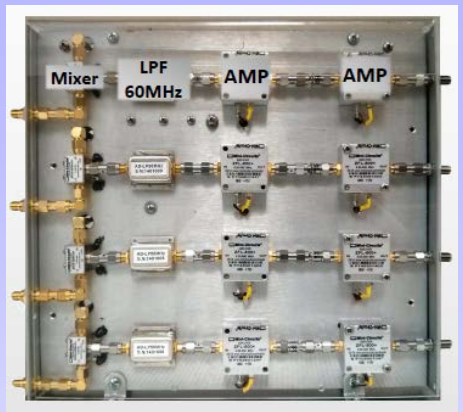
Nyquist: 6 GHz
Sampling Rate: 360MHz

Mishali, Eldar, Dounaevsky, and Shoshan, 2010
Cohen et. al. 2014

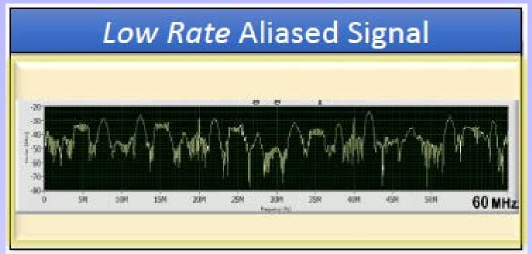
6% of Nyquist rate!



$x(t)$



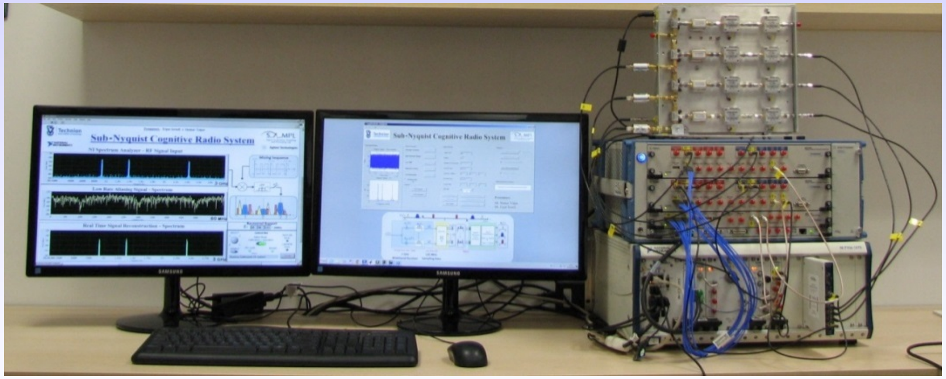
MWC analog front-end



$y_i[n]$

Parameters:

- Nyquist rate: 6 GHz
- Xampling rate: 360 MHz
  (6% of Nyquist rate)



Performance:

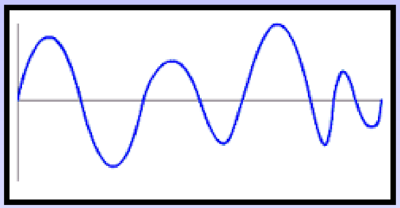
- Wideband receiver mode: 49 dB dynamic range, SNDR > 30 dB
- ADC mode: 1.2v peak-to-peak full-scale, 42 dB SNDR = 6.7 ENOB


Technion
Israel Institute of Technology
Headed by Yonina Eldar
S_A_MPL
Signal Acquisition Modeling and Processing Lab
Sub-Nyquist
Cognitive Radio System

Reducing Rate with Quantization

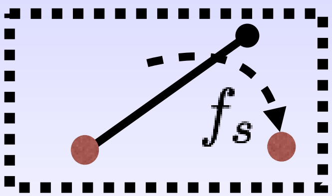
Kipnis, Goldsmith and Eldar 15

- Until now we ignored quantization
- Quantization introduces inevitable distortion to the signal
- Since the recovered signal will be distorted due to quantization do we still need to sample at the Nyquist rate?

01001001001
010010…

$x(t)$ →



→ $y[n]$ → quantizer → $\hat{y}[n]$

$f_s = \#\ \text{samples/sec}$

$R = f_s \log_2(\#\text{levels})$ bit/sec

Sampling Theory

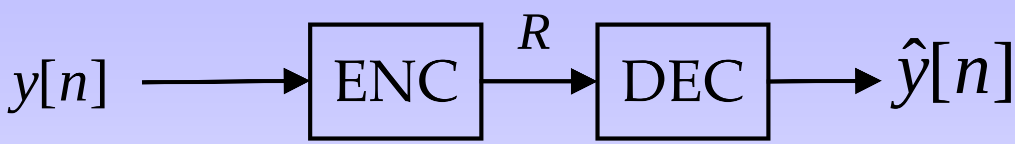
Source Coding [Shannon]

Goal: Unify sampling and rate distortion theory

Unification of Rate-Distortion and Sampling Theory

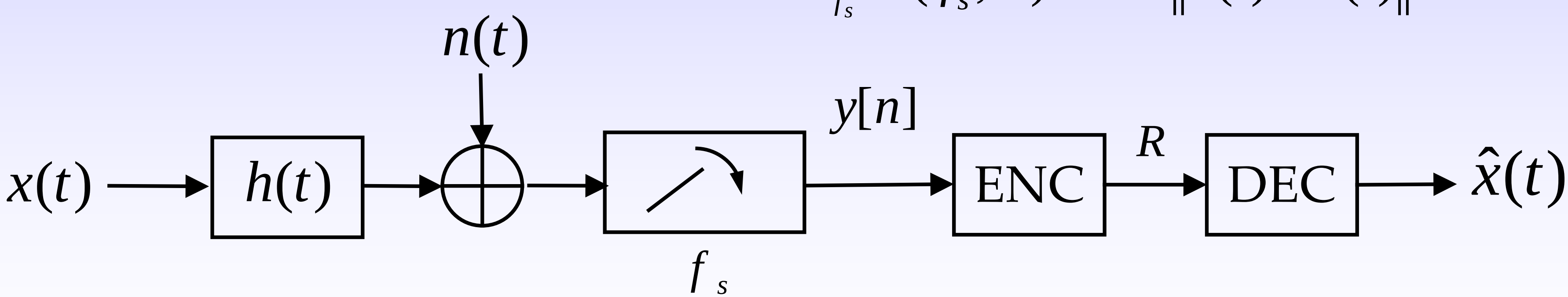
- Standard source coding:

  For a given discrete-time process y[n] and a given bit rate R what is the minimal achievable distortion $D(R) = \inf \left\| y[n] - \hat{y}[n] \right\|^2$



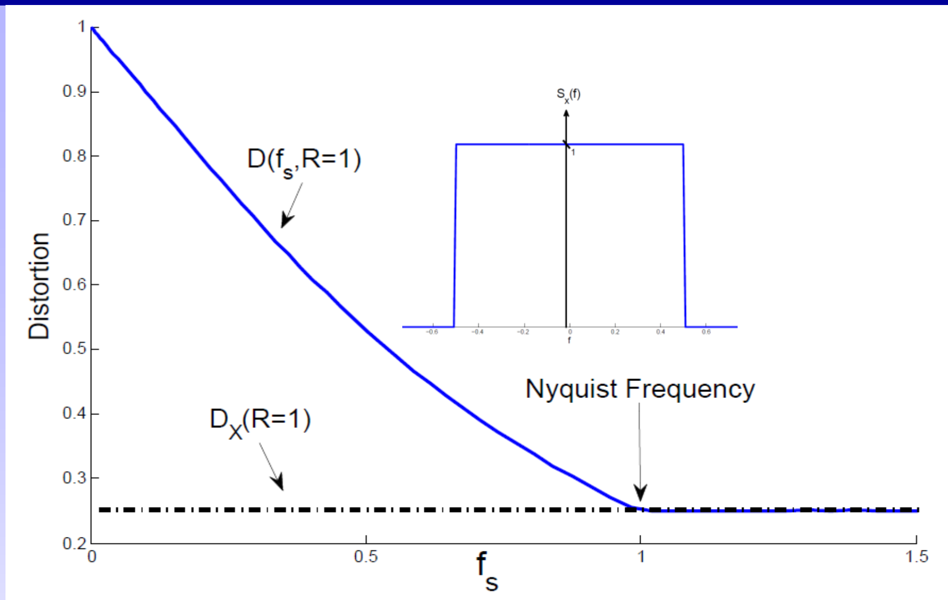
- Our question:

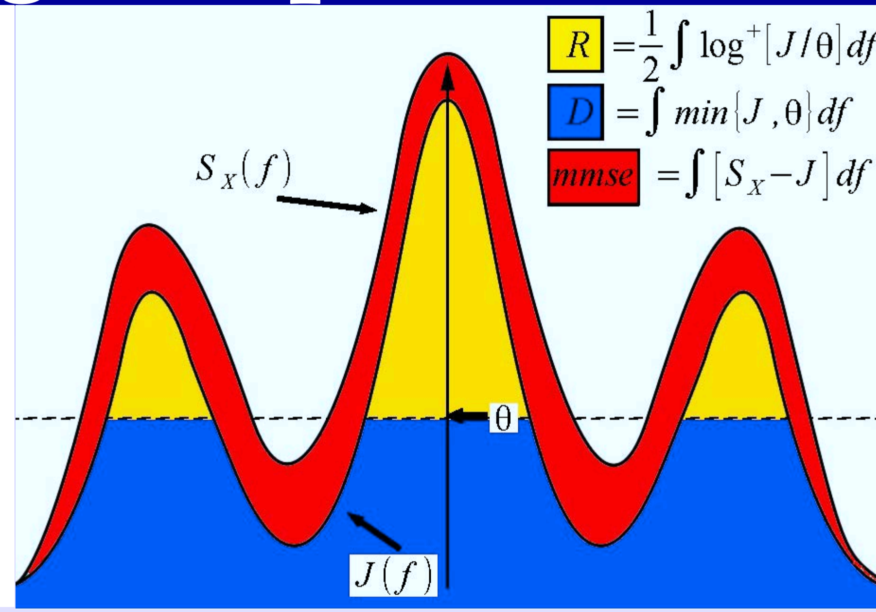
  For a given continuous-time process *x(t)* and a given bit rate R what is the minimal distortion $\inf_{f_s} D(f_s, R) = \inf \left\| x(t) - \hat{x}(t) \right\|^2$



- What sampling rate is needed to achieve the optimal distortion?

Quantizing the Samples: Source Coding Perspective





Preserve signal components above "noise floor" q , dictated by R

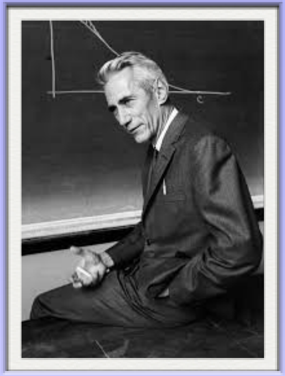
Distortion corresponds to mmse error + signal components below noise floor

Theorem (Kipnis, Goldsmith, Eldar, Weissman 2014)

$$R(f_s,\theta) = \frac{1}{2}\int_{-\frac{fs}{2}}^{\frac{fs}{2}} \log^+\left[\tilde{S}_{X|Y}(f)/\theta\right]df$$

$$D(f_s,\theta) = mmse_{X|Y}(f_s) + \int_{-\frac{fs}{2}}^{\frac{fs}{2}} \min\{\tilde{S}_{X|Y}(f),\theta\}df$$

Optimal Sampling Rate



Shannon [1948]:

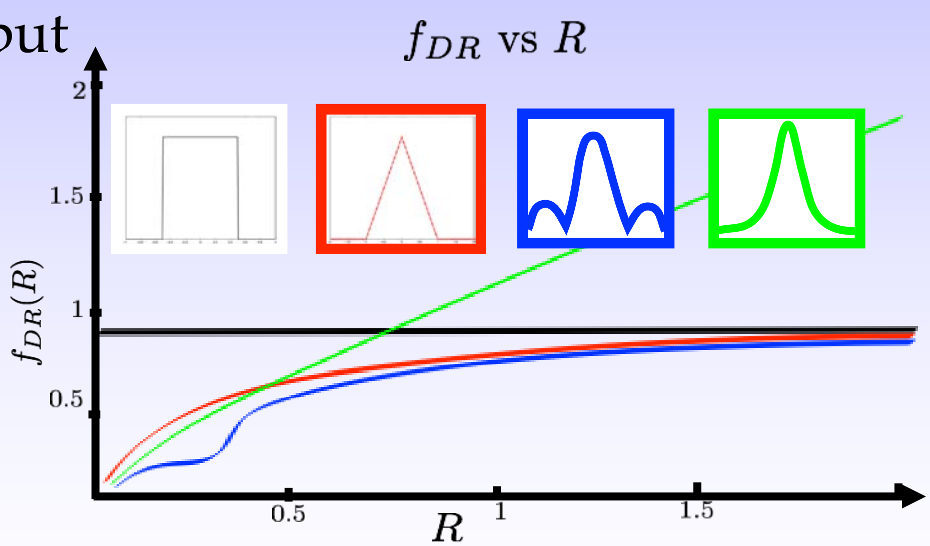
"we are not interested in exact transmission when we have a continuous source, but only in transmission to within a given tolerance"

- Can we achieve D(R) by sampling below f_{Nyq}?

- Yes! For any non-flat PSD of the input

$$D(R, f_s) = D(R) \text{ for}$$

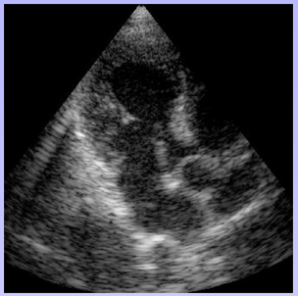
$$f_s \geq f_{DR}(R)!$$



No optimality loss when sampling at sub-Nyquist (without input structure)!

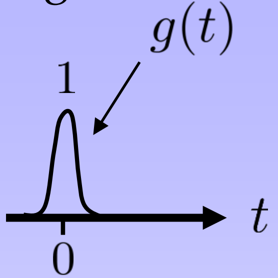
Ultrasound

- Relatively simple, radiation free imaging



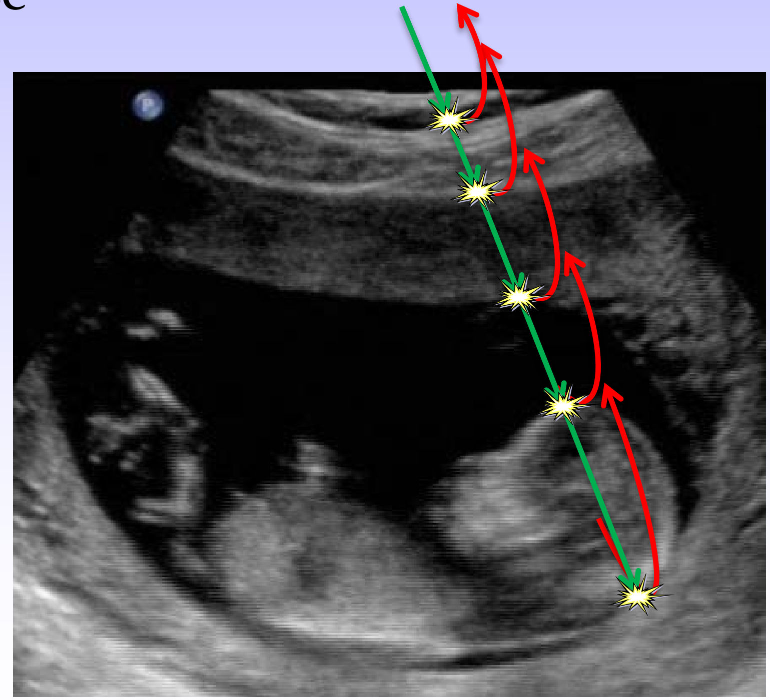
Cardiac sonography

Obstetric sonography

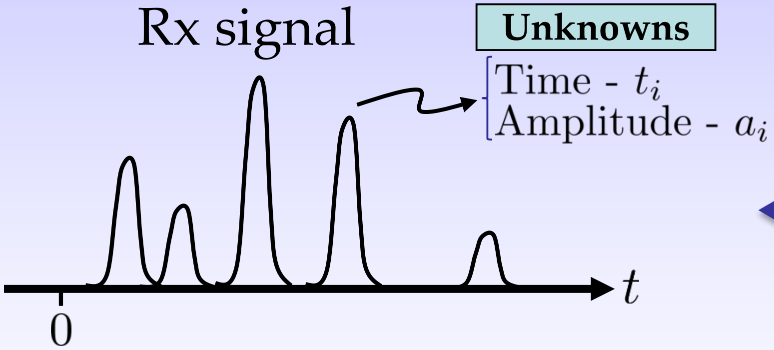


Tx pulse

Ultrasonic probe



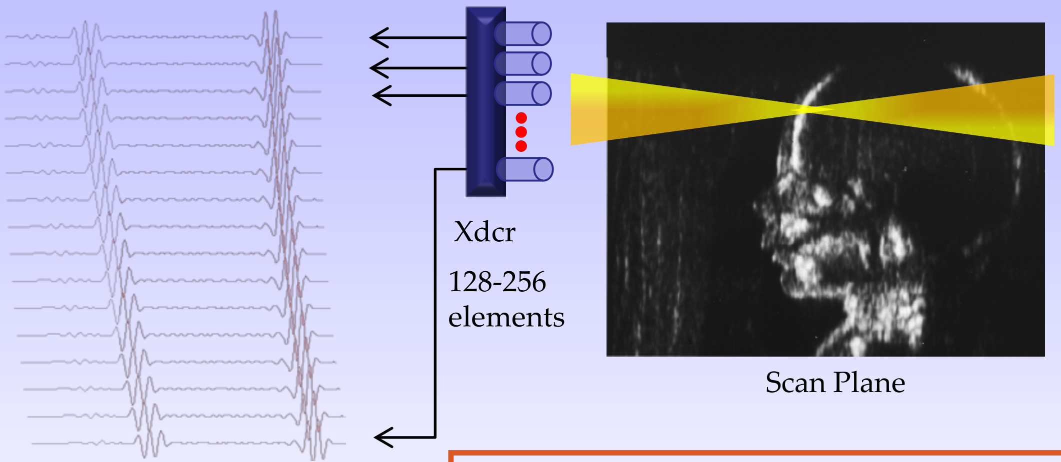
Rx signal



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

Processing Rates

- To increase SNR and resolution an antenna array is used
- SNR and resolution are improved through beamforming by introducing appropriate time shifts to the received signals



Focusing the received beam by applying nonlinear delays

$$\Phi(t;\theta)=\frac{1}{M}\sum_{m=1}^{M}\varphi_m\left(t-\frac{1}{2}\left(t-\sqrt{t^2-4(\delta_m/c)t\sin\theta+4(\delta_m/c)^2}\right)\right)$$

- Requires high sampling rates and large data processing rates
- One image trace requires 128 samplers @ 20M, beamforming to 150 points, a total of $6.3\text{x}10^6$ sums/frame

Challenges

- Can we reduce analog sampling rates?
- Can we perform nonlinear beamforming on the sub-Nyquist samples without interpolating back to the high Nyquist-rate grid digitally?

Compressed Beamforming

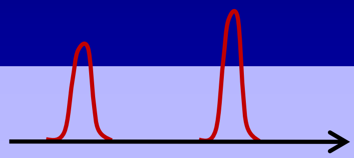
Goal: reduce ultrasound machine size at same resolution

Enable 3D imaging

Increase frame rate

Enable remote wireless ultrasound

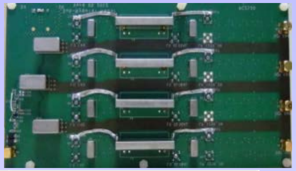
Streams of Pulses



Gedalyahu, Tur, Eldar 10, Tur, Freidman, Eldar 10

- L pulses can be entirely recovered from only $2L$ Fourier coefficients – finite-rate-of-innovation framework by Vetterli, Marziliano, Blu, Dragotti
- Efficient hardware:

$$x(t) \longrightarrow \boxed{s^*(-t)} \text{—/ } \longrightarrow \boxed{FFT} \longrightarrow c[k]$$

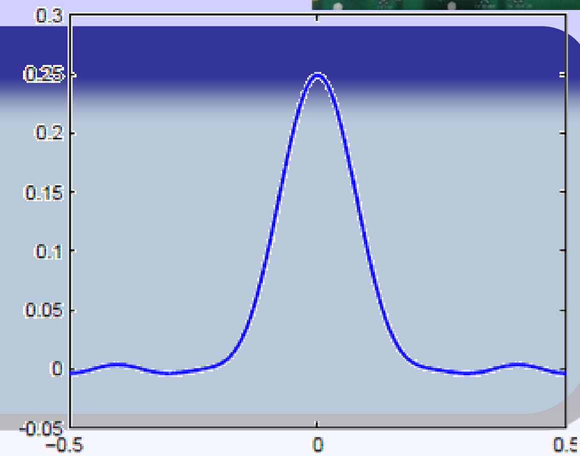


Theorem (Tur, Eldar and Friedman 11)

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

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

then $c[k]$ are the desired Fourier coefficients



Here $\mathcal{K}$ are the desired set of Fourier coefficients

- Sum-of-Sincs filter with compact support $S(\omega) = \frac{\tau}{\sqrt{2\pi}} \sum_{k \in \mathcal{K}} b_k \text{sinc}\left(\frac{\omega}{2\pi/\tau} - k\right)$

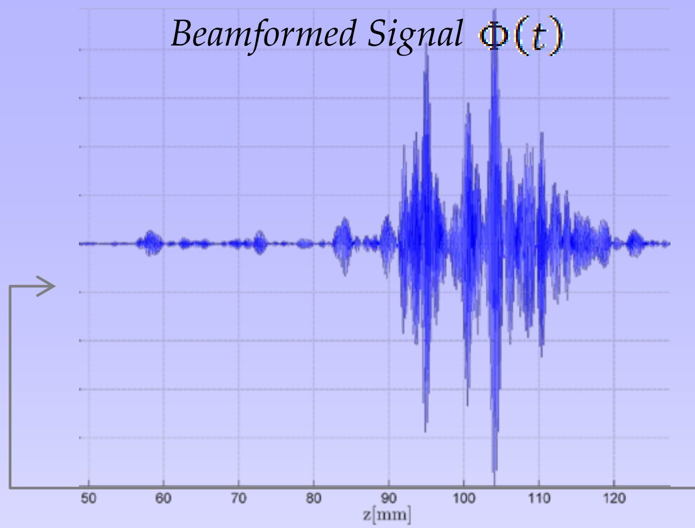
Conventional Beamforming

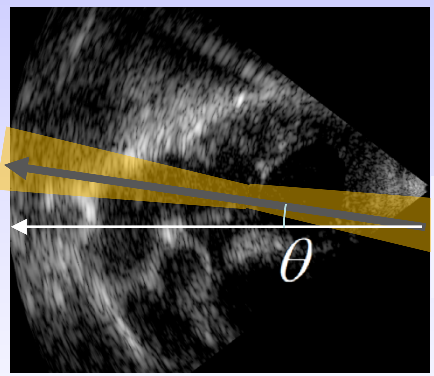
Non-linear scaling of the received signals

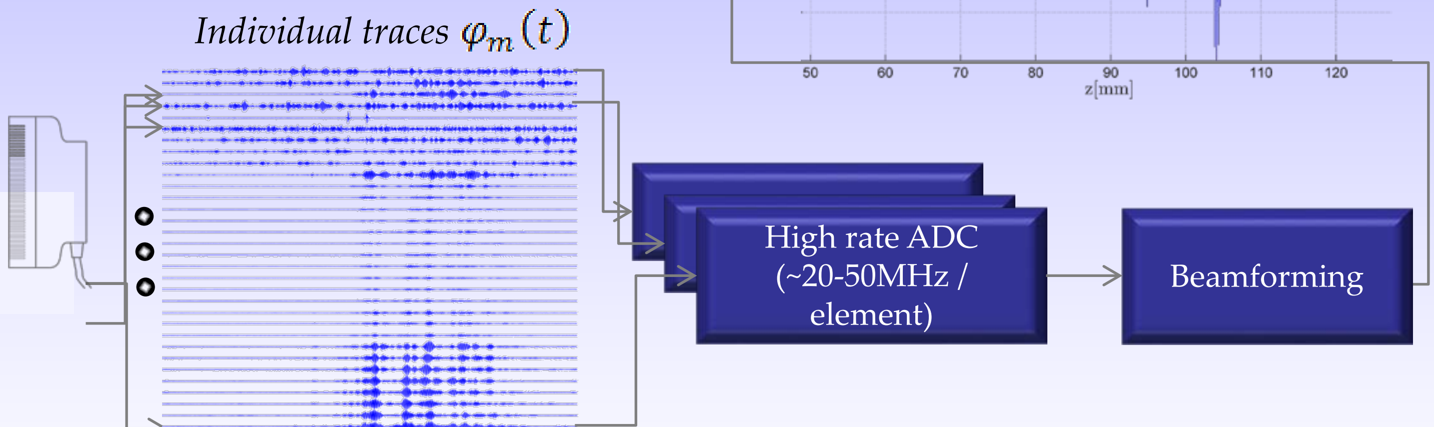
$$\Phi(t;\theta) = \frac{1}{M}\sum_{m=1}^{M} \varphi_m\left(\frac{1}{2}\left(t + \sqrt{t^2 - 4\gamma_m t\sin\theta + 4\gamma_m^2}\right)\right)$$

γ_m - distance from m'th element to origin, normalized by c.

Performed digitally after sampling at sufficiently high rate







- **Focusing along a certain axis – reflections originating from off-axis are attenuated (destructive interference pattern)**
- **SNR is improved**

Difficulty in Low Rate Sampling

- Each individual trace is buried in noise and has no structure
- Structure exists only after beamforming which improves resolution/SNR
- How can we perform beamforming on low rate data? How can we obtain small time shifts without interpolation?

- **Compressed beamforming: Enables beamforming from low rate samples**
- **Key idea: Perform beamforming in frequency**

$$c_k = \frac{1}{M}\sum_{m=1}^{M}\sum_{n}\varphi_m[n]Q_{k,m;\theta}[k-n]$$

Fourier coefficient of BMF signal

Fourier coefficient of signal at element m

Logic:

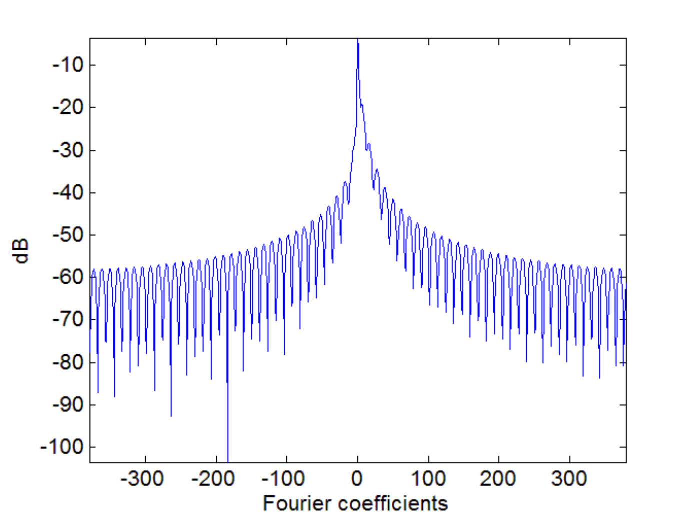
1. BMF signal is a stream of pulses => can be recovered from a small number of c_k
2. Small number of c_k requires only a small number of $\varphi_m[n]$

Low rate sampling of $\varphi_m(t)$!

$Q_{k,m;\theta}[n]$ are the Fourier coefficients of $q_{k,m}(t;\theta)$

$$q_{k,m}(t;\theta) = 1_{[\delta_m/c,\,\tau_m(T;\theta))}(t)\left(1+\frac{(\delta_m/c)^2\cos^2\theta}{(t-(\delta_m/c)\sin\theta)^2}\right)\times$$

$$\exp\left\{i\frac{2\pi}{T}k\frac{\delta_m/c-t\sin\theta}{(t-(\delta_m/c)\sin\theta)}\delta_m/c\right\}$$



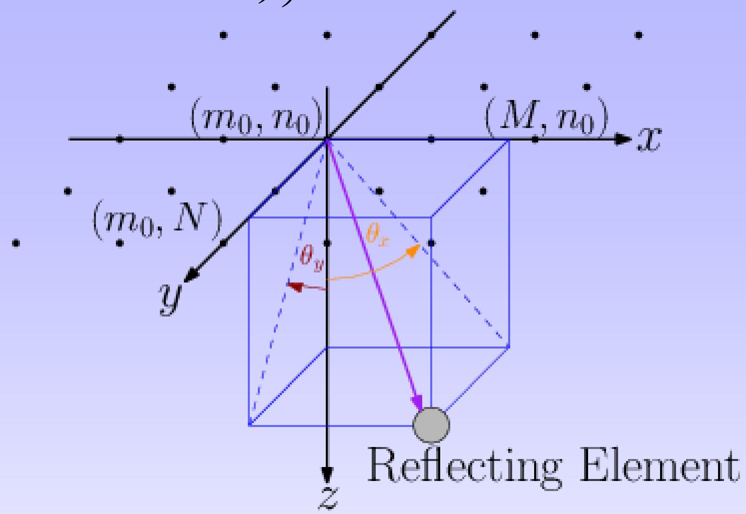
Volumetric Ultrasound Imaging

$$\Phi\left(t;\theta_x,\theta_y\right)=\frac{1}{N_{\mathrm{RX}}}\sum_{(m,n)}\varphi_{m,n}\left(t-\frac{1}{2}\left(t-\sqrt{t^2-4t\left(\gamma_m x_\theta+\gamma_n y_\theta\right)+4\left|\gamma_{m,n}\right|^2}\right)\right)$$

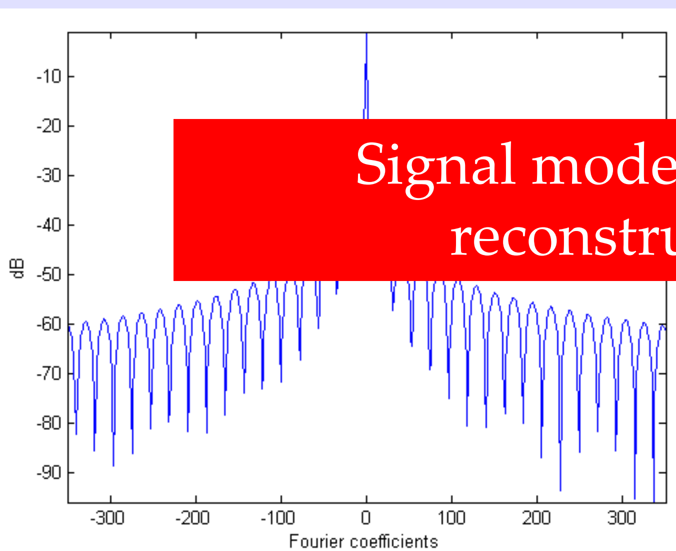
$$c[k]\approx\frac{1}{N_{\mathrm{RX}}}\sum_{(m,n)}\sum_{l=-L_1}^{L_2}c_{m,n}[k-l]\,Q_{k,m,n;\theta_x,\theta_y}[l]$$

Fourier coefficient of BMF signal

Fourier coefficient of signal detected at element (m,n)



20 elements of $\left\{Q_{k,m,n;\theta_x,\theta_y}[l]\right\}$ contain more than 95% of the energy



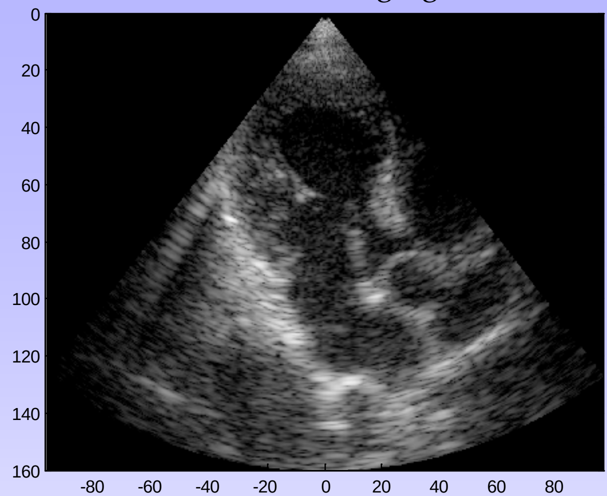
$$q_{k,m,n}\left(t;\theta_x,\theta_y\right)=I_{\left[\left|\gamma_{m,n}\right|,\tau_{m,n}\left(T;\theta_x,\theta_y\right)\right]}(t)\times$$

Signal model still holds, allowing the same reconstruction technique to be used

$$\exp\left\{-i\frac{2\pi}{T}k\cdot\frac{t\left(\gamma_m x_\theta+\gamma_n y_\theta\right)-\left|\gamma_{m,n}\right|^2}{t-\left(\gamma_m x_\theta+\gamma_n y_\theta\right)}\right\}$$

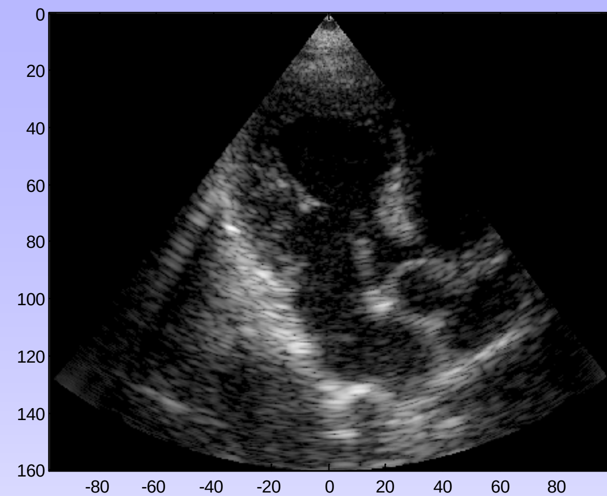
Ultrasound Results

Standard Imaging



3328 real-valued samples, per sensor per image line

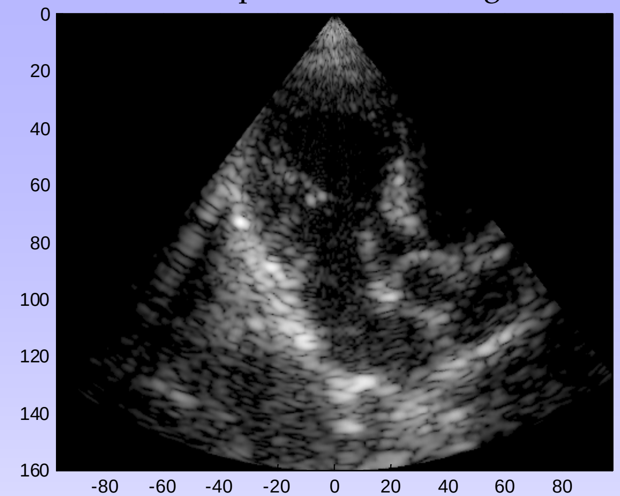
Xampled beamforming



360 complex-valued samples, per sensor per image line

~1/10 of the Nyquist rate

Xampled beamforming



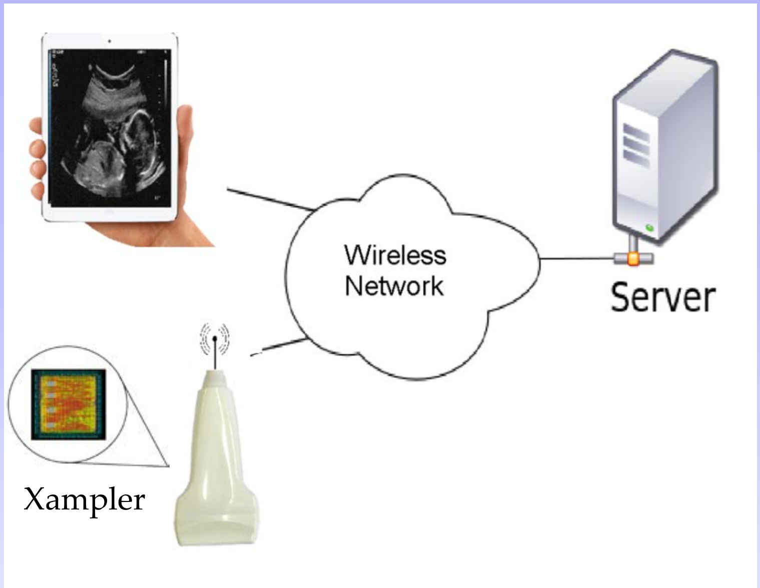
100 complex-valued samples, per sensor per image line

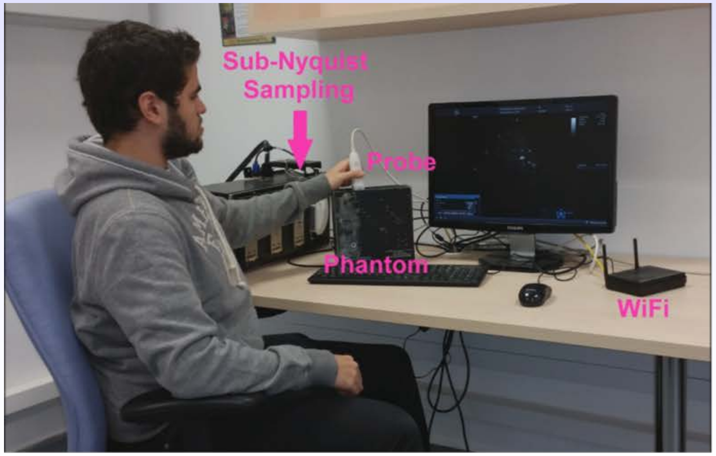
~1/32 of the Nyquist rate

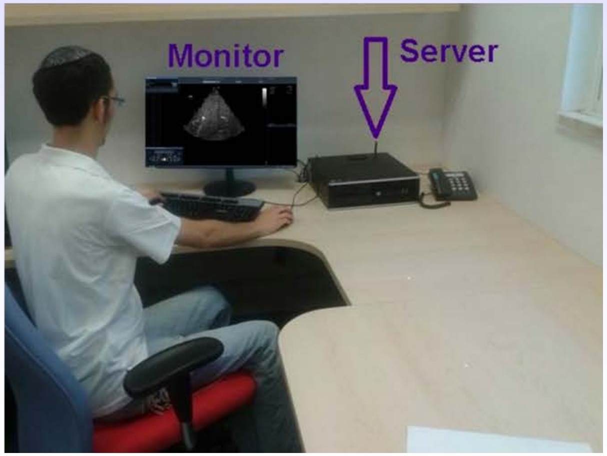
- We obtain a 32-fold reduction in sample rate and 1/16-fold reduction in processing rate
- All digital processing is low rate as well
- Almost same quality as full rate image

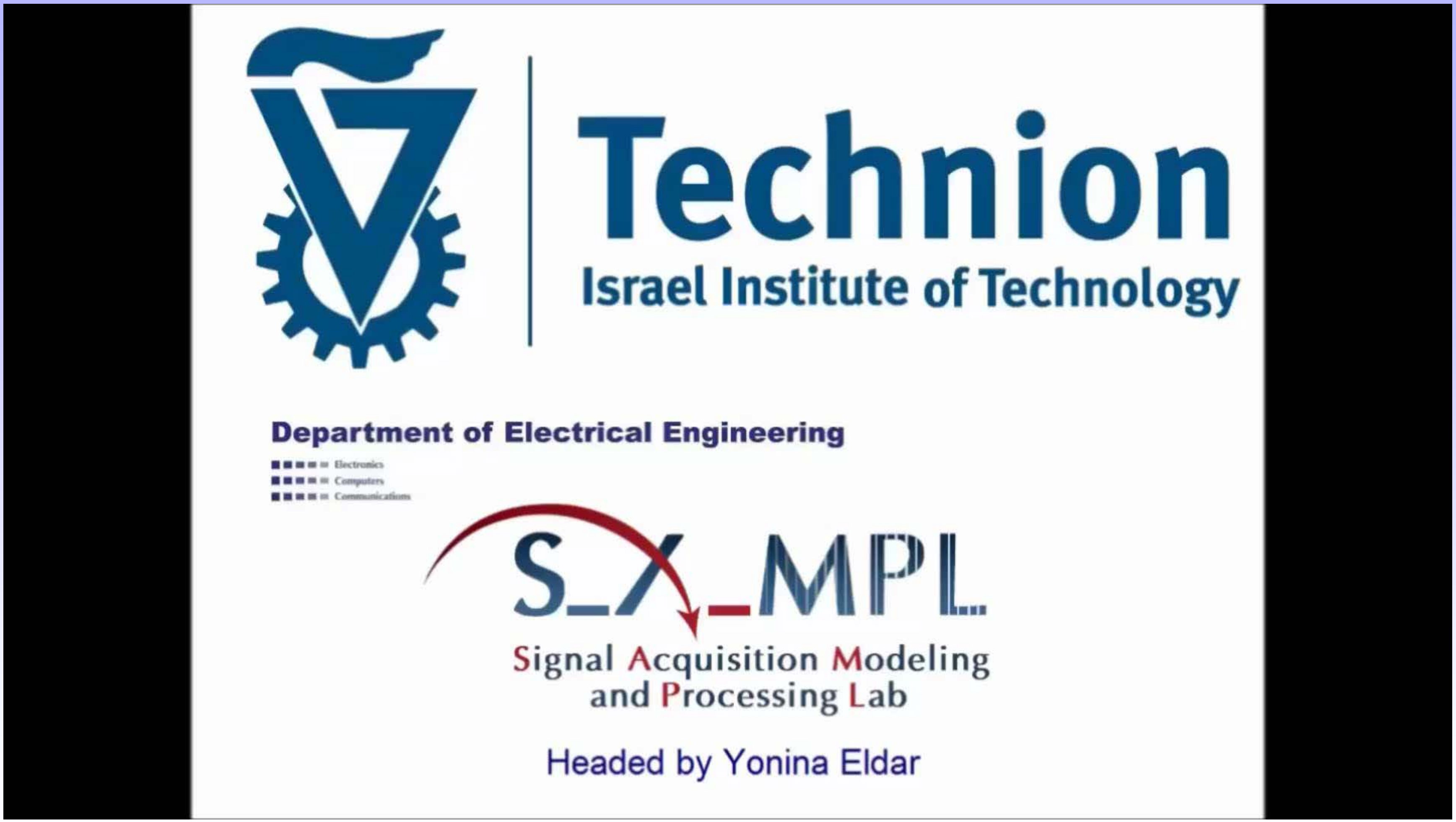
Wireless Ultrasound Imaging

- A wireless probe performs Xampling and transmits the low rate data to a server for processing
- Frequency Domain Beamforming and image reconstruction is performed by the server
- The image is sent for display on a monitor





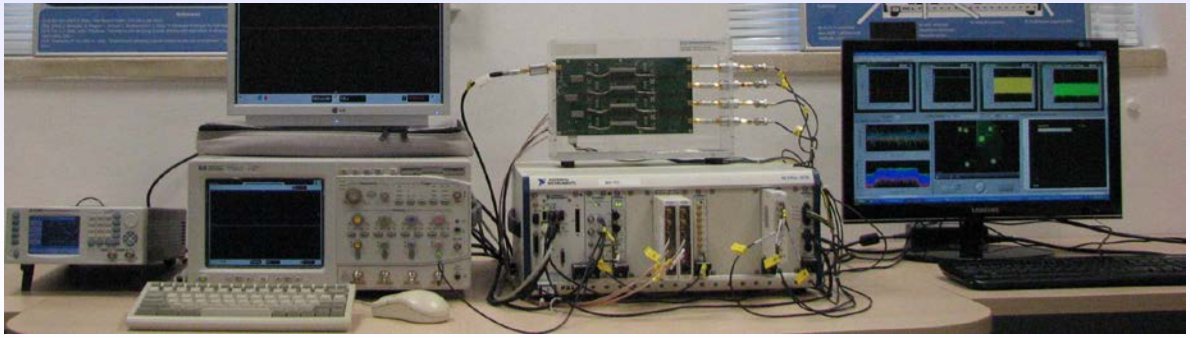



Technion
Israel Institute of Technology
Department of Electrical Engineering
Electronics
Computers
Communications
S_X_MPL
Signal Acquisition Modeling
and Processing Lab
Headed by Yonina Eldar

Pulse-Doppler Radar

Bar-Ilan and Eldar, 13

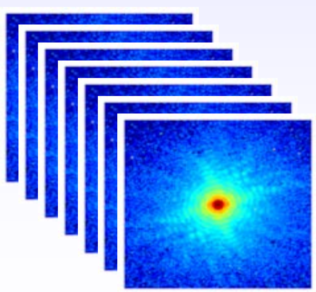
- Same beamforming idea can be used in radar in order to obtain high resolution radar from low rate samples
- Our radar prototype is robust to noise and clutter
- Doppler Focusing (beamforming in frequency):
  - Optimal SNR scaling
  - CS size does not increase with number of pulses
  - No restrictions on the transmitter
  - Clutter rejection and the ability to handle large dynamic range

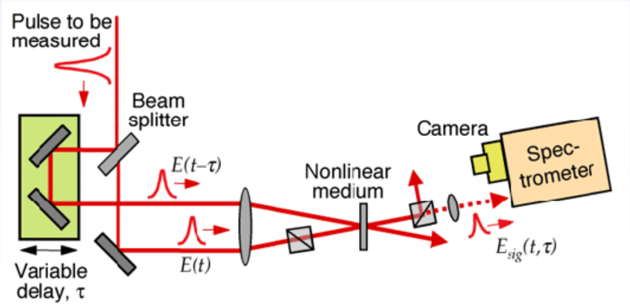


Part 2:

Measurement design for phase retrieval

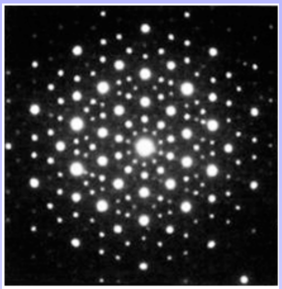
Enabling phase retrieval from Fourier measurements using practical devices!



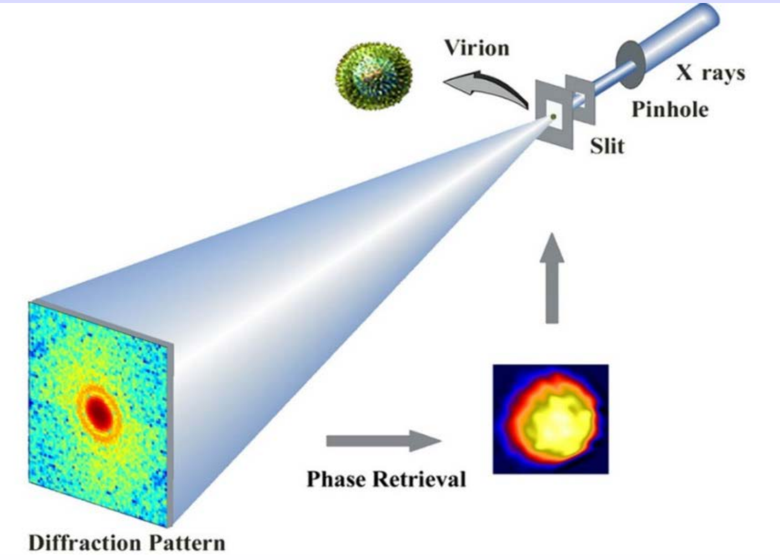


Phase Retrieval:
Recover a signal from its Fourier magnitude

$$x[n] \rightarrow \boxed{\text{Fourier + Absolute value}} \rightarrow y[k] = |X[k]|^2$$



- Arises in many fields: crystallography (Patterson 35), astronomy (Fienup 82), optical imaging (Millane 90), and more
- Given an optical image illuminated by coherent light, in the far field we obtain the image's Fourier transform
- Optical devices measure the photon flux, which is proportional to the magnitude
- Phase retrieval can allow direct recovery of the image



Theory of Phase Retrieval

Difficult to analyze theoretically when recovery is possible

- No uniqueness in 1D problems **(Hofstetter 64)**
- Uniqueness in 2D if oversampled by factor 2 **(Hayes 82)**
- No guarantee on stability
- No known algorithms to achieve unique solution

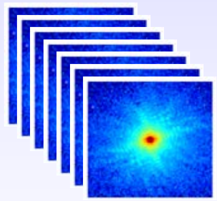
Recovery from Fourier Magnitude Measurements is Difficult!

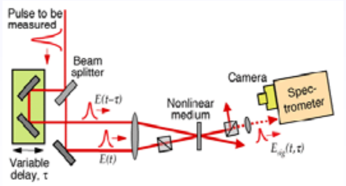
Progress on Phase Retrieval

- Assume random measurements to develop theory **(Candes et. al, Rauhut et. al, Gross et. al, Li et. al, Eldar et. al, Netrapalli et. al, Fannjiang et. al …)**

- Introduce prior to stabilize solution
  - Support restriction **(Fienup 82)**
  - Sparsity **(Moravec et. al 07, Eldar et. al 11, Vetterli et. al 11, Shechtman et. al 11)**
  - GESPAR: Greedy sparse phase retrieval **(Shechtman, Beck and Eldar 14)**

- Add redundancy to Fourier measurements
  - Impulse addition and least-squares recovery **(Huang et. al 15)**
  - Short-time Fourier transform **(Nawab et. al 83, Eldar et. al 15, Jaganathan et. al 15)**
  - Masks **(Candes et. al 13, Bandeira et. al 13, Jaganathan et. al 15)**





Analysis of Phase Retrieval

Analysis of Random Measurements:

$$y_i = |\langle a_i, x \rangle|^2 + w_i \leftarrow \text{noise} \quad x \in R^N$$

random vector

- $4N - 2$ measurements needed for uniqueness

(Balan, Casazza, Edidin 06, Bandira et. al 13)

Stable Phase Retrieval (Eldar and Mendelson 14):

$O(N)$ measurements needed for stability

$O(k\log(N/k))$ measurements needed for stability with sparse input

Solving $\sum_{i=1}^{M} \left(y_i - |\langle a_i, x \rangle|^2\right)^p \quad 1 < p \leq 2$ provides stable solution

How to solve objective function?

Recovery via Semidefinite Relaxation

Candes, Eldar, Strohmer ,Voroninski 12

- $|\langle a_k, x\rangle|^2 = \mathrm{Tr}(A_k X)$ with $A_k = a_k {a_k}^T$, $X = xx^T$
- Phase retrieval can be written as

  minimize rank (X)

  subject to A(X) = b, X ≥ 0

- SDP relaxation: replace rank(X) by Tr(X) or by logdet($X + \varepsilon I$) and apply reweighting
- PhaseCut: semidefinite relaxation based on MAXCUT **(Waldspurger et. al 12)**

Advantages / Disadvantages

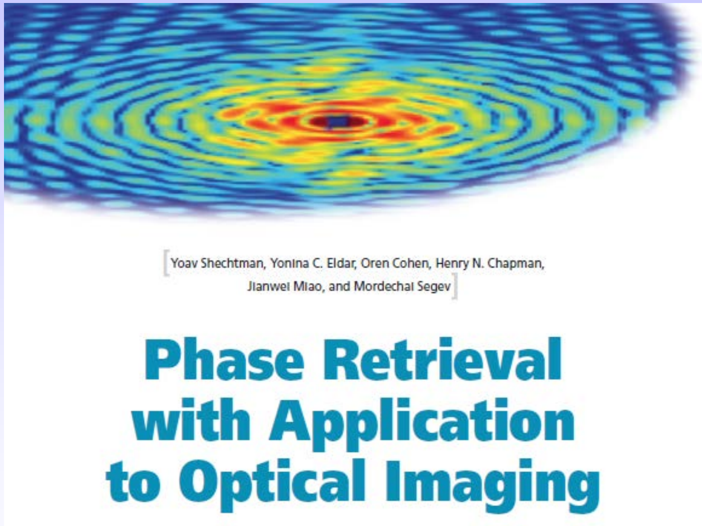
- Yields the true vector whp for $O(N)$ Gaussian meas. **(Candes et al. 12)**
- Recovers sparse vectors whp for $O(k^2\log(N))$ Gaussian meas. **(Candes et al. 12)**
- Computationally demanding
- Difficult to generalize to other nonlinear problems

Provable Efficient Algorithms for Phase Retrieval

- Wirtinger flow: Gradient descent on $|\langle a_i, x\rangle|^2$ (Candes et. al 14,15)
- Amplitude flow: Gradient descent on $|\langle a_i, x\rangle|$ (Wang, Giannakis and Eldar 16)
- All recovery results for random measurements (or random masks)

Recent overview:

Y. Shechtman, Y. C. Eldar, O. Cohen,
H. N. Chapman, J. Miao, and M. Segev,
"Phase retrieval with application to optical imaging,"
SP magazine 2015



Moving to practice: Provable recovery from Fourier measurements?

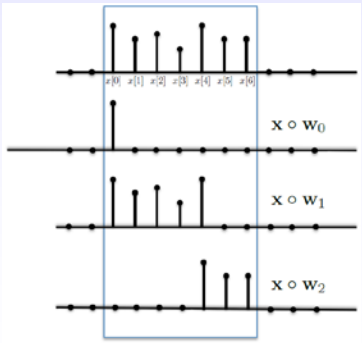
Design Measurements + Optimization Methods

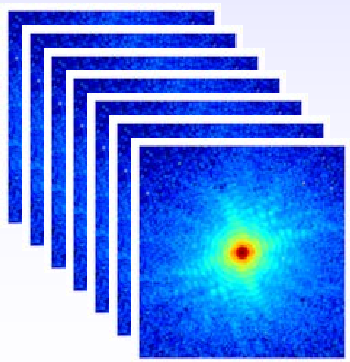
Lessons learned from sub-Nyquist sampling:

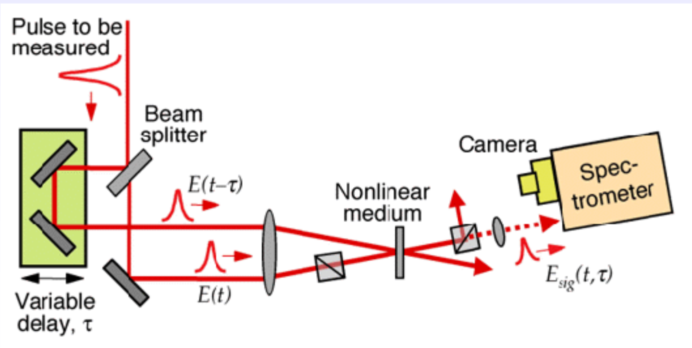
- Measurement design is crucial!
- Combine with modern optimization tools for recovery

Fourier measurements with a twist:

- Impulse addition and least-squares recovery
- Short-time Fourier transform
- Small number of fixed masks







Least-Squares Phase Retrieval

Huang, Eldar and Sidiropoulos 15

- We have seen that SDP relaxation can recover the true signal for sufficiently many random Gaussian measurements
- We can show that in fact SDP relaxation for Fourier phase retrieval is tight!

Theorem (Huang, Eldar and Sidiropoulos 15)

Consider LS recovery $\sum_{i=1}^{M}(y_i - |\langle f_i, x\rangle|^2)^2$ from Fourier measurements. Then:

1. The SDP relaxation $\min_{X\succeq 0}\sum_{i=1}^{M}\left(y_i - \mathrm{Tr}(f_i f_i^* X)\right)^2$ is tight for any M
2. We can always find a rank-one solution from the SDP solution X

- Create the correlation sequence $r_k = \mathrm{sum}(\mathrm{diag}(X, k))$
- Any choice of x such that $r_k = \mathrm{sum}(\mathrm{diag}(xx^*, k))$ is optimal

Spectral Factorization

Theorem

Minimum phase factor can be found by solving

$$\max_{X \succeq 0} X(1,1) \quad \text{s.t. } r_k = \text{sum}(\text{diag}(X,k))$$

The solution is always rank one

Minimum phase solution:

Let $X(z) = \sum_{n=0}^{N-1} x[n] z^{-n}$ be the z-transform of $x[n]$

$x[n]$ is minimum phase if the zeros of $X(z)$ are all inside the unit circle

Minimum phase solution can always be found in polynomial time

Summary: LS Phase Retrieval

Huang, Eldar and Sidiropoulos 15

- By solving two SDPs we can always solve the LS phase retrieval problem from Fourier measurements $\min_x \sum_{i=1}^{M} (y_i - |\langle f_i, x\rangle|^2)^2$
- The solution can be found by implementing:

1. $\widehat{X} = \min_{X \succeq 0} \sum_{i=1}^{M} \left(y_i - \mathrm{Tr}(f_i f_i^* X)\right)^2$
2. $r_k = \mathrm{sum}(\mathrm{diag}(\widehat{X}, k))$
3. $\max_{X \succeq 0} X(1,1)$ s.t. $r_k = \mathrm{sum}(\mathrm{diag}(X, k))$ gives rank-one optimal solution

- The minimum phase solution is optimal namely minimizes the LS error
- However, solution may not be equal to the true x since there are no uniqueness guarantees in 1D phase retrieval

Convert any signal into a minimum phase signal and then measure it!

Impulse Addition

Huang, Eldar and Sidiropoulos 15

- Any signal can be made minimum phase by adding an impulse at zero

Theorem (Huang, Eldar and Sidiropoulos 15)

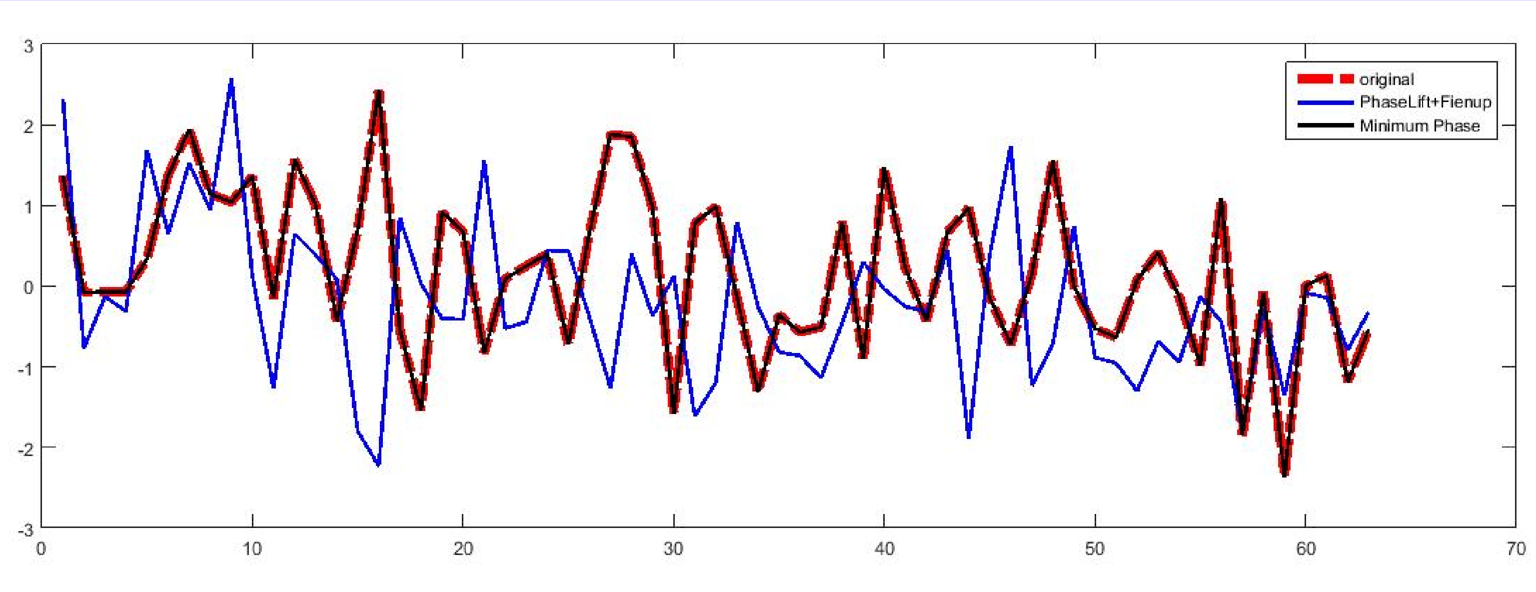
An arbitrary complex signal is minimum phase if $|x_0| \geq \|x\|_1$

- Add an impulse at zero
- Take Fourier magnitude measurements
- Recover the minimum phase signal
- Subtract the impulse

Robust recovery of any 1D complex signal from Fourier measurements using SDP!

Simulation: Exact Recovery

- Compare with Fourier measurements with recovery using PhaseLift initialized by Fienup
- Both produced zero fitting error but only our approach led to recovery of the true signal



Recovery From the STFT Magnitude

$$X_g(m,k) = \left| \sum_{n=0}^{N-1} x[n] g[mL-n] e^{-j2\pi kn/N} \right|^2$$

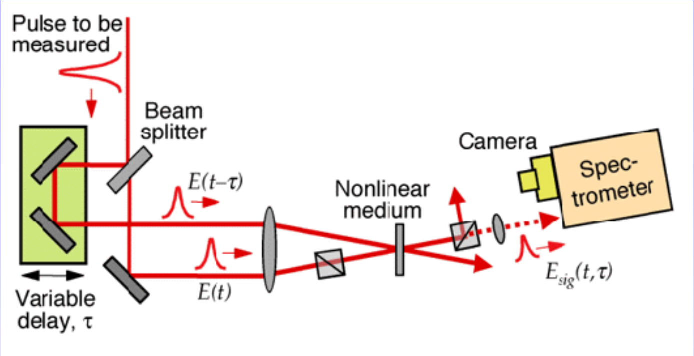
L – step size
N – signal length
W – window length

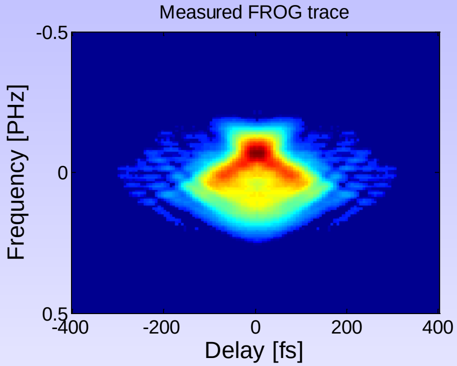
- Easy to implement in optical settings
  - FROG – measurements of short pulses **(Trebino and Kane 91)**
  - Ptychography – measurement of optical images **(Hoppe 69)**
- Also encountered in speech/audio processing **(Griffin and Lim 84, Nawab et. al 83)**
- Almost all signals can be recovered as long as there is overlap between the segments
- Almost all signals can be recovered using semidefinite relaxation
- Simple least-squares recovery for many choices of windows

Frequency-Resolved Optical Gating (FROG)

Trebino and Kane 91

- Method for measuring ultrashort laser pulses
- The pulse gates itself in a nonlinear medium and is then spectrally resolved





- In XFROG a reference pulse is used for gating leading to STFT-magnitude measurements:

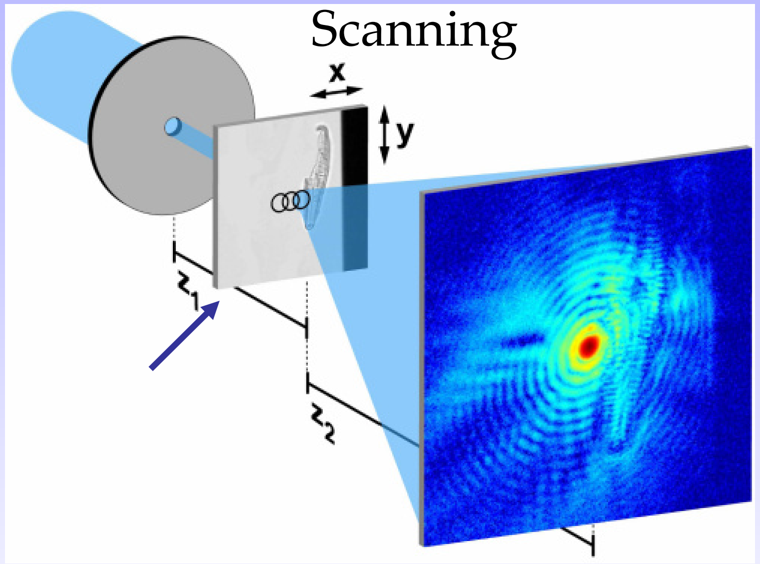
$$X_g(m,k) = \left| \sum_{n=0}^{N-1} x[n] g[mL-n] e^{-j2\pi kn/N} \right|^2$$

L – step size
N – signal length
W – window length

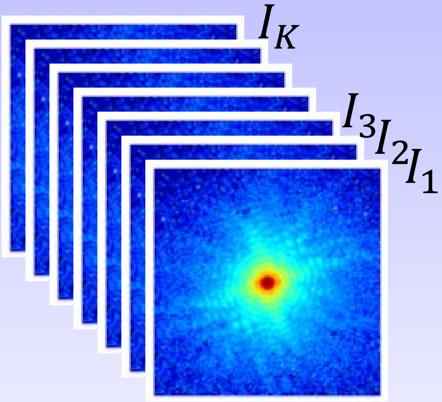
Ptychography

Hoppe 69

Plane wave



For all positions (X_i, Y_j) record diffraction pattern I_k



- Method for optical imaging with X-rays
- Records multiple diffraction patterns as a function of sample positions
- Mathematically this is equivalent to recording the STFT

Theoretical Guarantees

- Uniqueness condition for *L=1* and all signals:

Theorem (Eldar, Sidorenko, Mixon et. al 15)

The STFT magnitude with *L=1* uniquely determines any *x[n]* that is everywhere nonzero (up to a global phase factor) if:

1. The length-N DTFT of $|g[n]|^2$ is nonzero
2. $N \geqslant 2W - 1$
3. N and W-1 are coprime

- Uniqueness condition for general overlap and *almost* all signals:

Theorem (Jaganathan, Eldar and Hassibi 15)

The STFT magnitude uniquely determines *almost* any *x[n]* that is everywhere nonzero (up to a global phase factor) if:

1. The window *g[n]* is nonzero
2. $L < W \leqslant N/2$ (segments overlap)

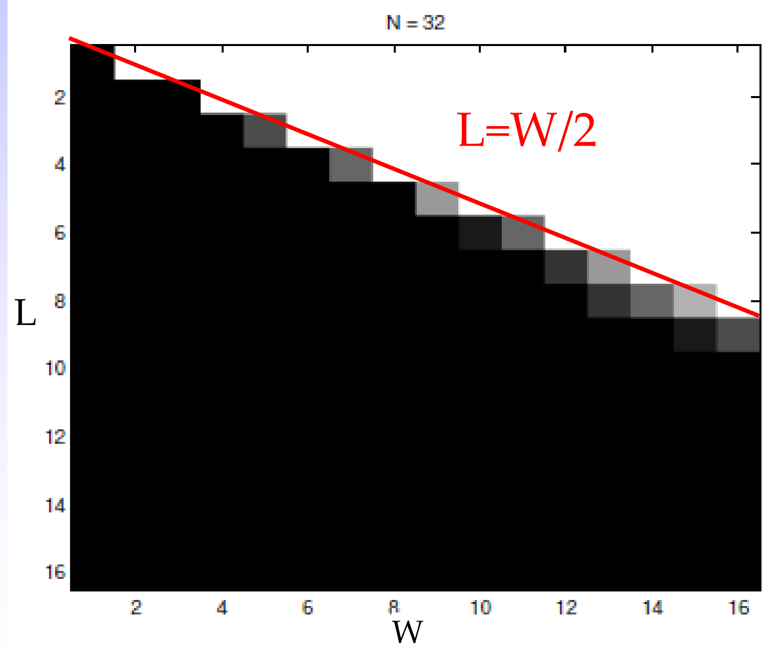
Strong uniqueness for 1D signals and Fourier measurements

Recovery From STFT via SDP

Theorem (Jaganathan, Eldar and Hassibi 15)

SDP relaxation uniquely recovers any *x[n]* that is everywhere nonzero from the STFT magnitude with *L=1* (up to a global phase factor) if $2 \leqslant W \leqslant N/2$

- In practice SDP relaxation seems to works as long as $L \leqslant W/2$ (at least 50% overlap)
- Proof in progress …
- Strong phase transition at *L=W/2*



Probability of Success for N=32 for different L and W

- Can prove the result assuming the first W/2 values of *x[n]* are known

Nonconvex Recovery From STFT Magnitude

Bendory and Eldar 16

- We consider the data in a transformed domain (1D DFT with respect to the frequency variable)

$$\hat{X}_g(m,\ell) = \sum_{n=0}^{N-1} x[n]x^*[n+\ell]g[mL-n]g[mL-n-\ell] = x^* H_{m,\ell} x$$

where $H_{m,\ell}$ is the (non-Hermitian) measurement matrix

- We suggest using gradient descent to minimize the non-convex loss

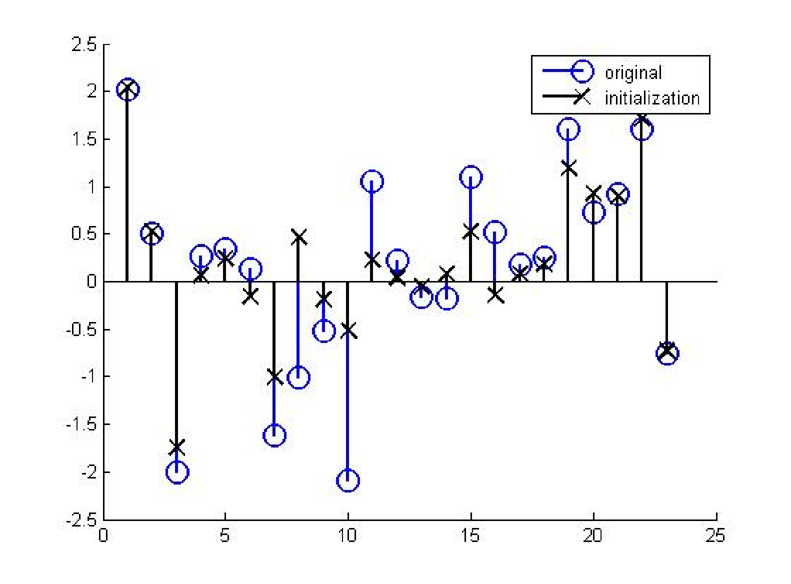
$$f(z) = \sum_{m,\ell} \left(\hat{X}_g(m,\ell) - z^* H_{m,\ell} z \right)^2$$

- Initialization by the principle eigenvector of a matrix, constructed as the solution of a least-squares problem
- Under appropriate conditions, initialization is close to the true solution

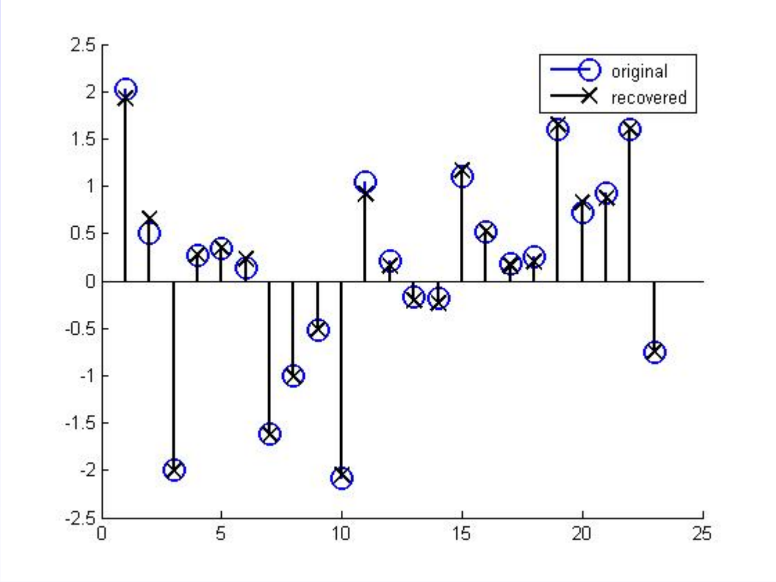
Nonconvex Recovery From STFT Magnitude

Simple example (N=23, W=7, L=1, SNR=20db)

Initialization



Recovery



Conclusions

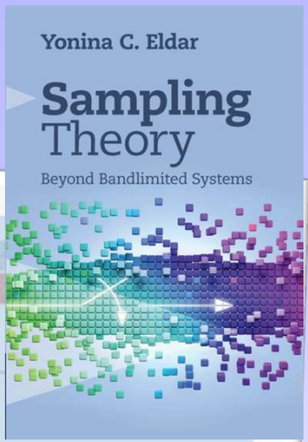
- Compressed sampling and processing of many analog signals even without structure
- Wideband sub-Nyquist samplers in hardware
- Many new applications like wireless ultrasound
- Merging information theory and sampling theory
- Importance of measurement design in phase retrieval

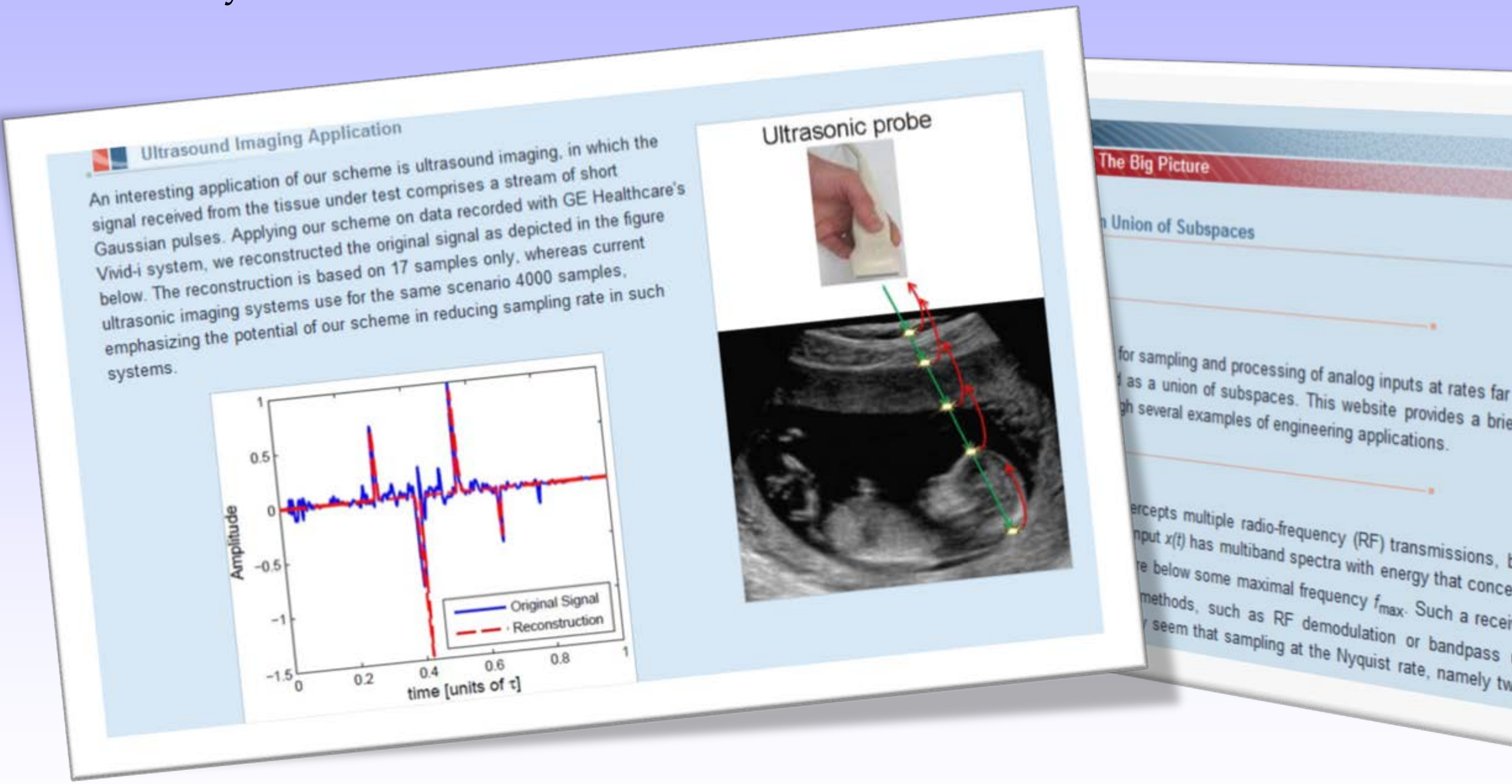
Exploiting processing task and careful design of measurements can lead to new sampling and processing techniques

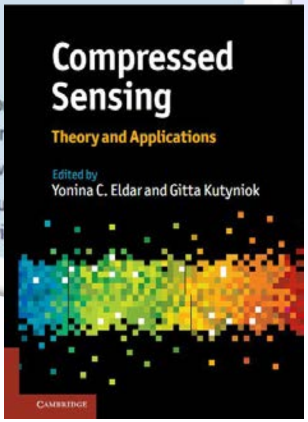
Xampling Website

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

Y. C. Eldar, “Sampling Theory: Beyond Bandlimited Systems", Cambridge University Press, 2015







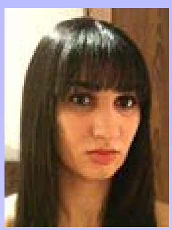
Y. C. Eldar and G. Kutyniok, "Compressed Sensing: Theory and Applications", Cambridge University Press, 2012

SAMPL Lab Website

www.sampl.technion.ac.il

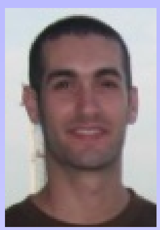
Acknowledgements

- Graduate students (Past and Present)

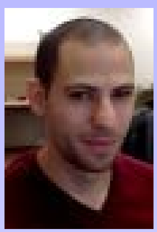


Deborah Cohen

Tanya Chernyakova



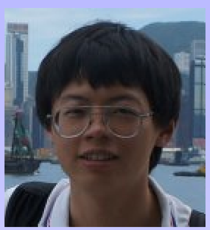
Shahar Tsiper



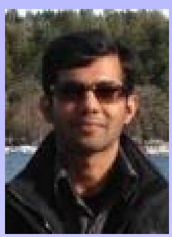
Alon Kipnis

Tamir Ben-Dori

Pavel Sidorenko



Kejun Huang



Kishore Jaganathan

- Engineers: Eli Shoshan, Yair Keller, Robert Ifraimov, Alon Eilam, Idan Shmuel
- Collaborators
- Funding: ERC, ISF, BSF, SRC, Intel
- Industry: GE, NI, Agilent

Thank you

If you found this interesting …

Looking for a post-doc!