

# 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](mailto:yonina@ee.technion.ac.il)



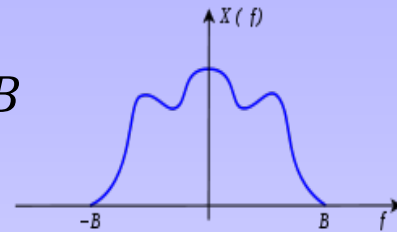
**Second International Matheon Conference on  
Compressed Sensing and its Applications**

# Traditional Sampling

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

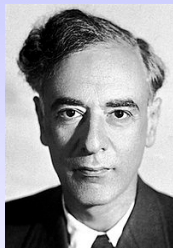
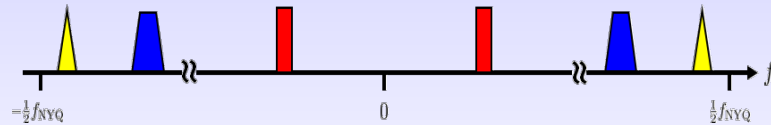
- Shannon-Nyquist theorem:

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



- Landau rate:

- Multiband signal with known support of measure  $\Lambda$
- Minimal sampling rate:  $\Lambda$

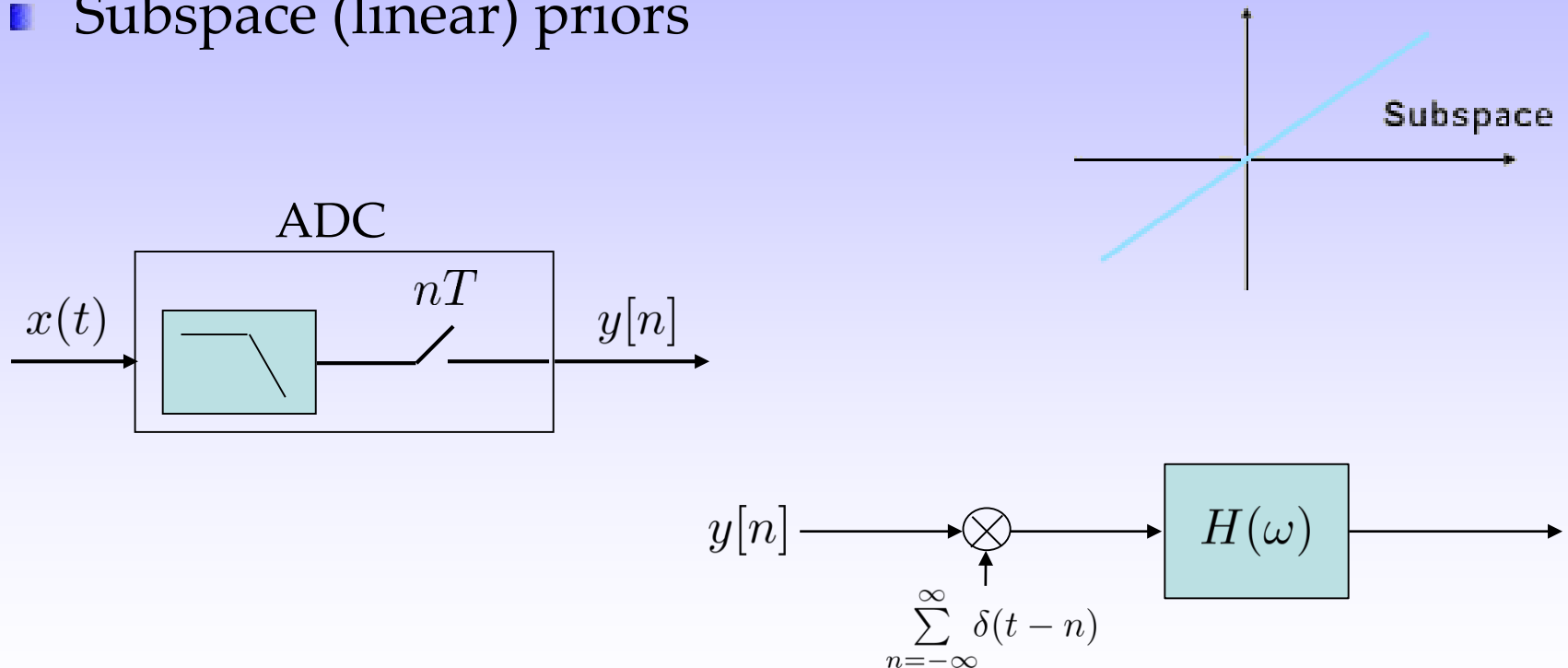


- Extension to arbitrary subspaces:

- 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  $1/T$

# Traditional Sampling

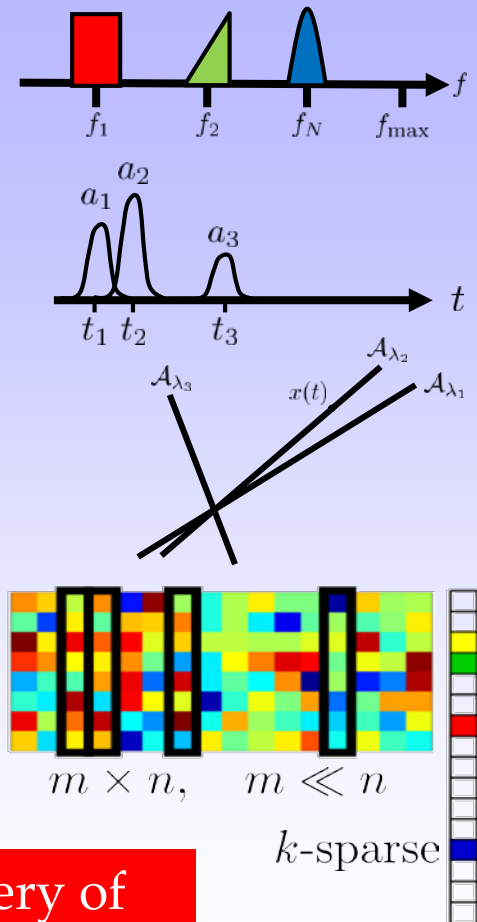
- Typically simple linear sampling schemes
- Linear recovery methods
- Subspace (linear) priors



# Sampling of Structured Signals

Exploit analog structure to reduce sampling rate

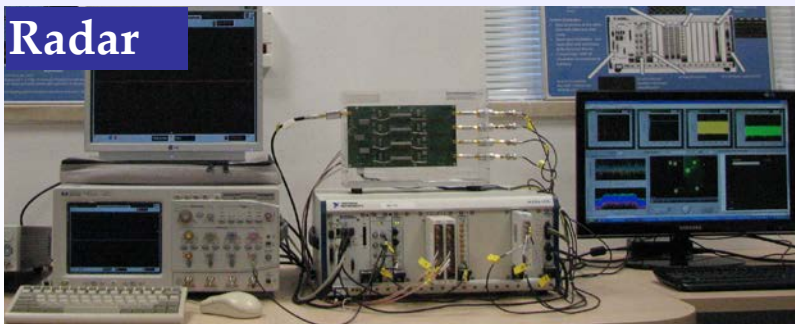
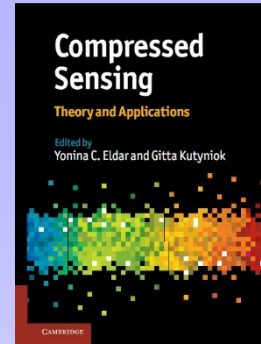
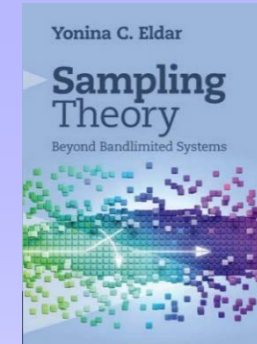
- Multiband signal with unknown support of measure  $\Lambda$   
Minimal sampling rate:  $2\Lambda$  (Mishali and Eldar '09)
- Stream of  $k$  pulses (finite rate of innovation)  
Minimal sampling rate:  $2k$  (Vetterli et. al '02)
- Union of subspaces (Lu and Do '08, Mishali and Eldar '09)
- Sparse vectors  
(Candes, Romberg, Tau '06, Donoho '06)



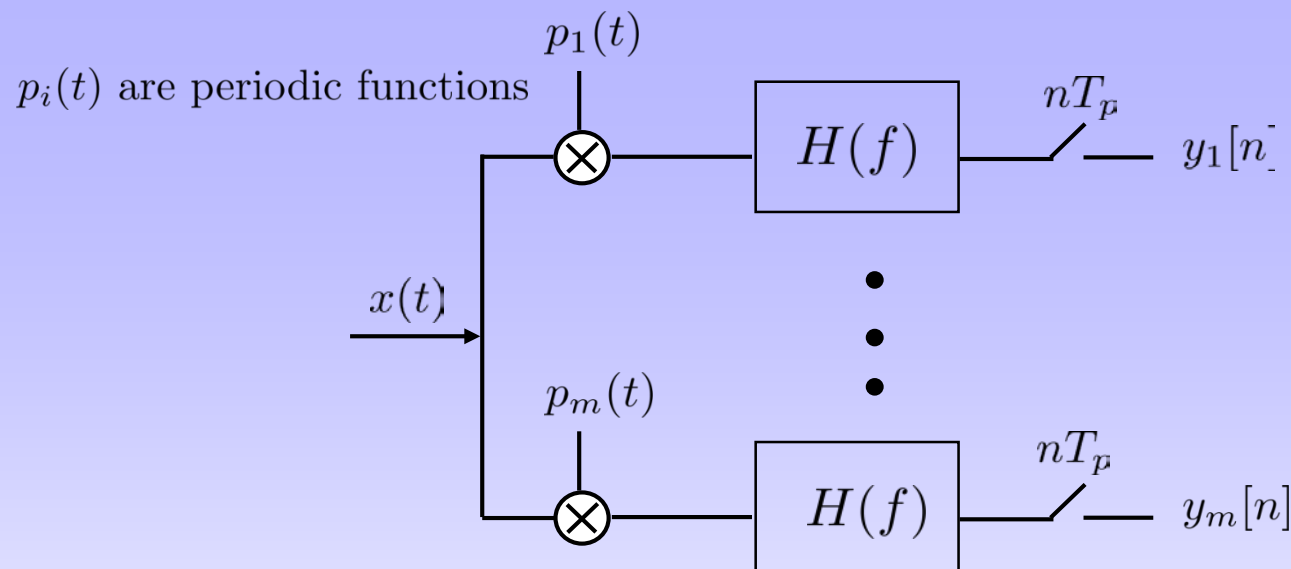
Sampling rate below Nyquist for recovery of  $x(t)$  by exploiting structure

# Sub-Nyquist Sampling

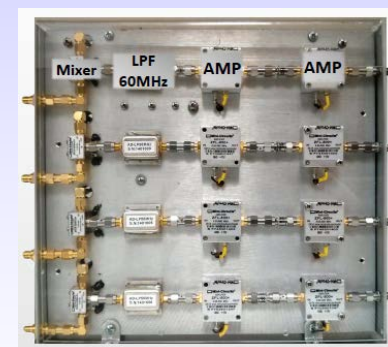
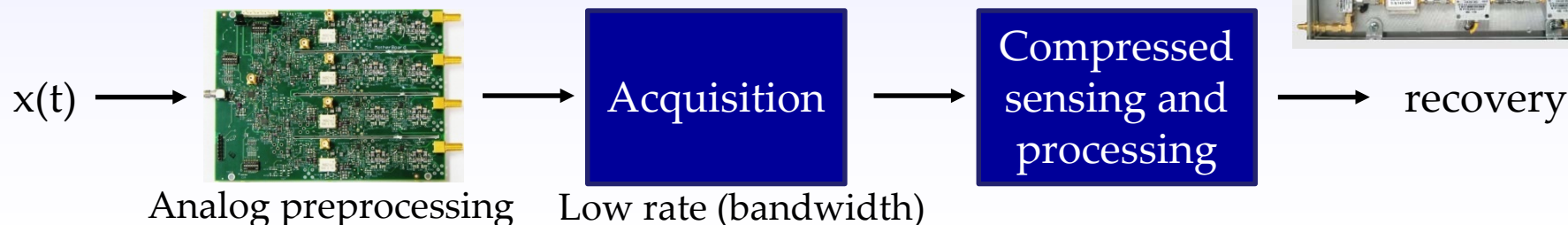
- Many examples in which we can reduce sampling rate by exploiting structure
- Xampling: practical sub-Nyquist methods which allow low-rate sampling and low-rate processing in diverse applications



# Xampling Hardware



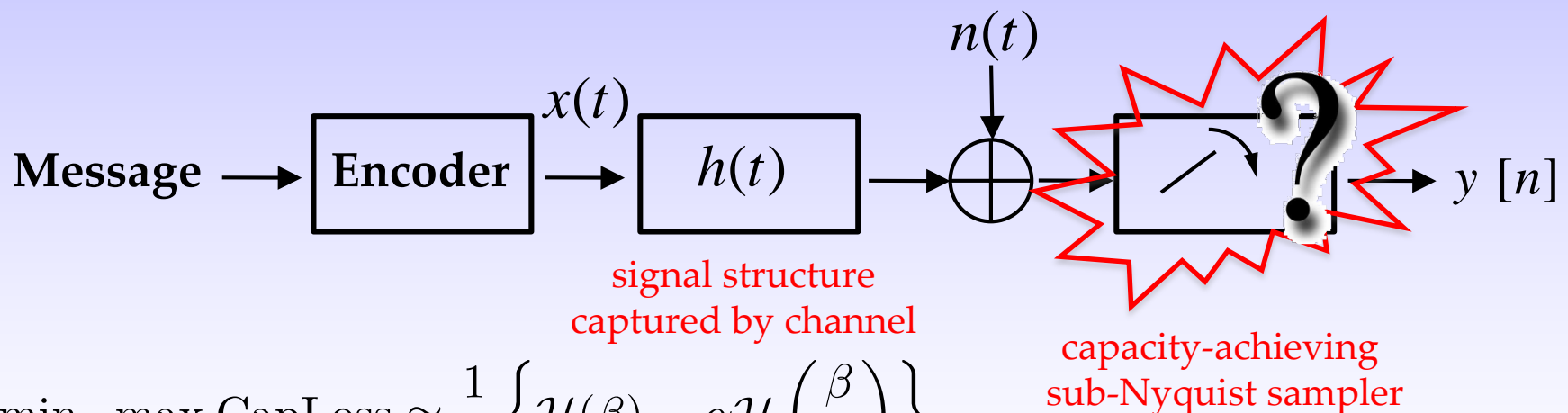
- $p_i(t) = \sum a_{in} e^{-j \frac{2\pi}{T_p} nt}$  sums of exponentials
- The filter  $H(f)$  shapes the tones and reduces bandwidth
- The channels can be collapsed to a single channel





# 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



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

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

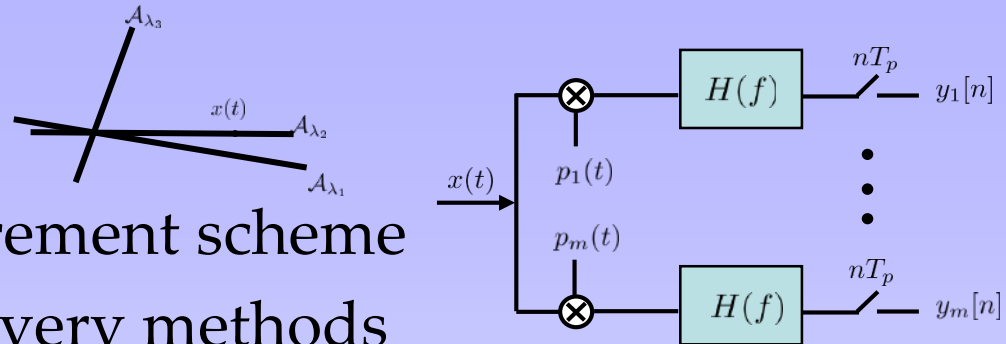
binary entropy function

■  $\alpha$ : undersampling factor

■  $\beta$ : 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:

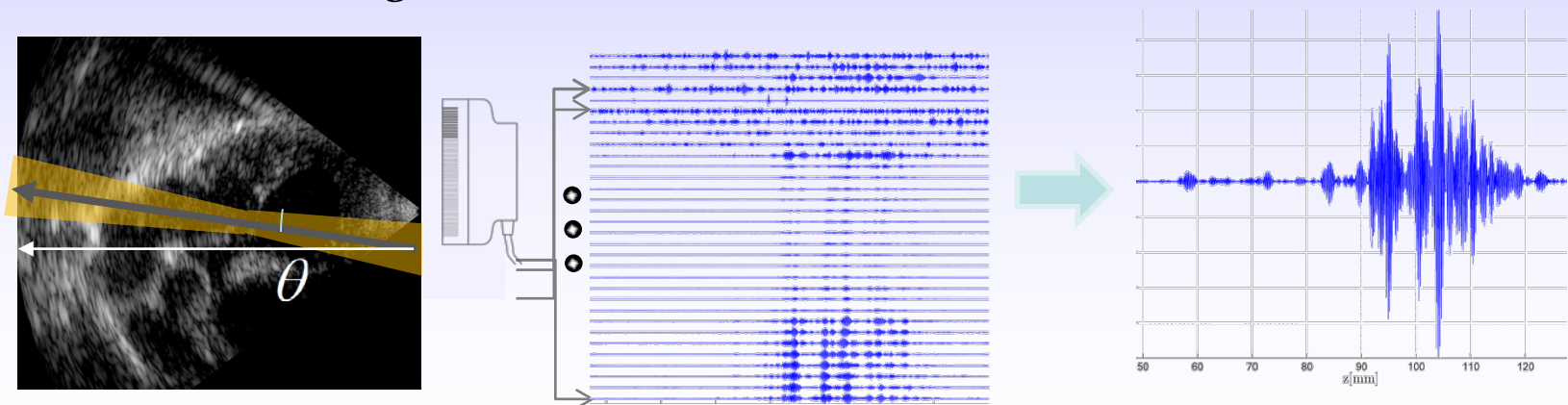
- Is structure necessary for sub-Nyquist sampling?
- Can careful measurement design and optimization-based recovery methods help in other nonlinear problems?



# Sub-Nyquist Without Structure

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

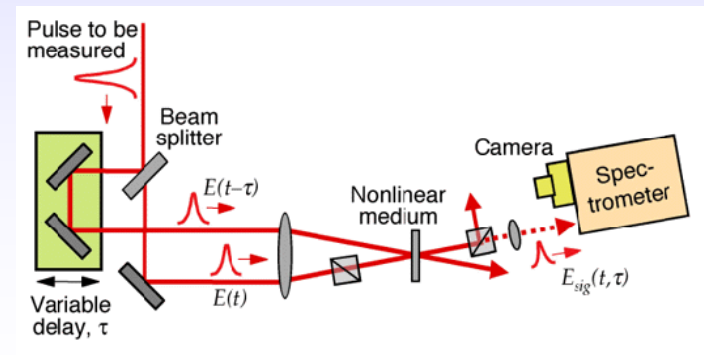
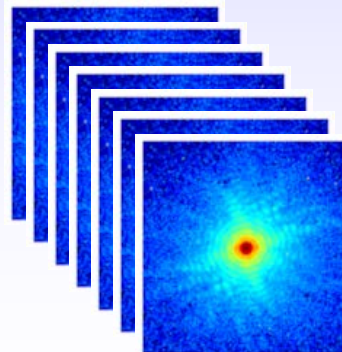
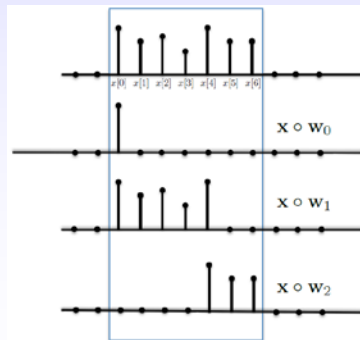
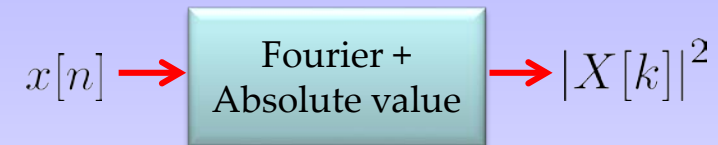
- Goal: Recovery of some function of the signal
  - Signal statistics: Power spectrum estimation *with Geert Leus and Deborah Cohen*
  - Quantized version of the signal *with Andrea Goldsmith and Alon Kipnis*
  - Sampling of a set of signals that are used for beamforming where the beamformed signal has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



# 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
  - Applications to optics



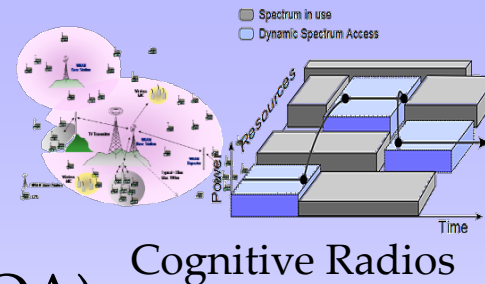
**Part 1:**

# **Xampling Without Structure**

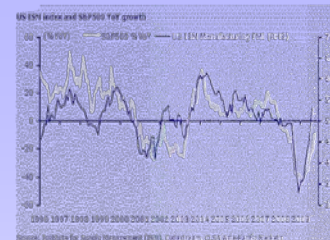
# Power Spectrum Reconstruction

- 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

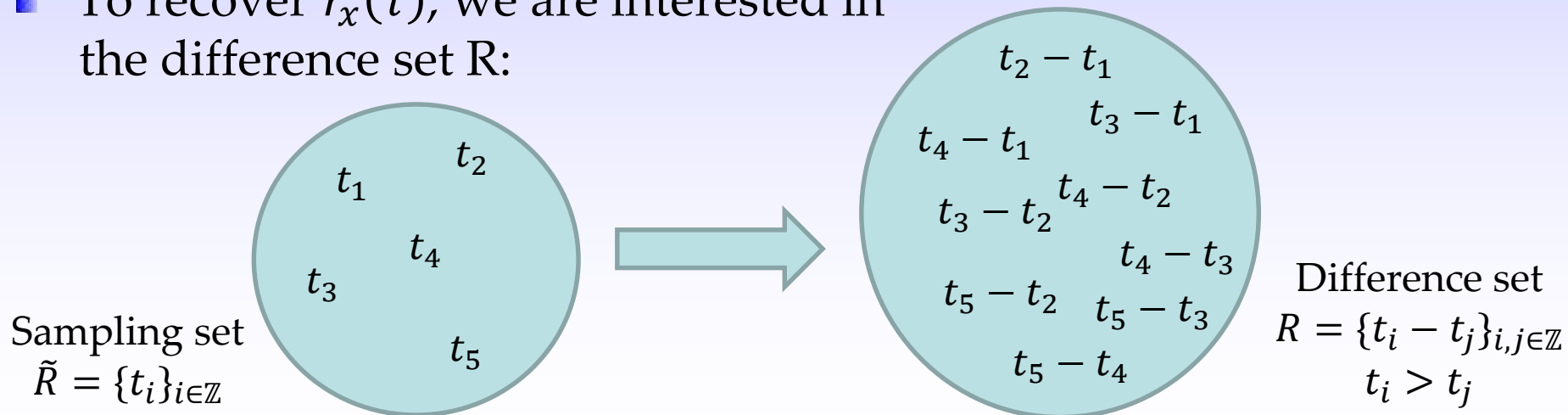
- Let  $x(t)$  be a wide-sense stationary ergodic signal
- We sample  $x(t)$  with a stable sampling set at times  $\tilde{R} = \{t_i\}_{i \in \mathbb{Z}}$
- We want to estimate  $r_x(\tau) = \mathbb{E}[x(t)x(t - \tau)]$

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:





# Previous Work

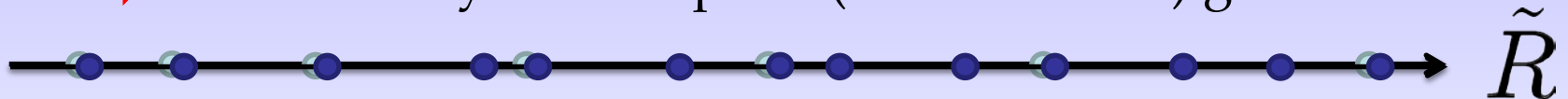
- Random sampling:
  - Masry '76: The power spectrum from **Poisson** samples can be consistently estimated for all positive average sampling rates
- Deterministic sampling:
  - Vaidyanathan '11: estimates the covariance from a **co-prime** pair of sparse sampler with arbitrarily low sampling rate
  - Tarczynski '07 , Davies '11, Leus '12: estimates the covariance from **multicoset** samples with arbitrarily low sampling rate

Under what conditions on a sampling set can we reconstruct the covariance?

# Difference Set Density

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

- 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  
 $\Rightarrow$  the density of the square (difference set) goes to  $\infty$



## 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  $\lim_{r \rightarrow \infty} \frac{d_{\tilde{R}}(r)}{\sqrt{r}} = \infty$ , then,  $D^-(R) = \infty$

# Universal Minimal Sampling Rate


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

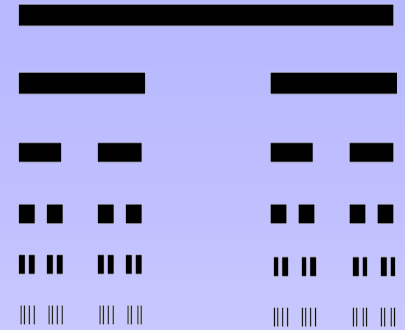
We can reconstruct the covariance from signal samples with density 0!

## 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  $\lim_{r \rightarrow \infty} \frac{d_{\tilde{R}}(r)}{\sqrt{r}} = \infty$  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:  $D^-(\tilde{R}_C) \rightarrow 0$
  - Difference set:  $D^-(R_C) \rightarrow \infty$  (both conditions hold)
- 



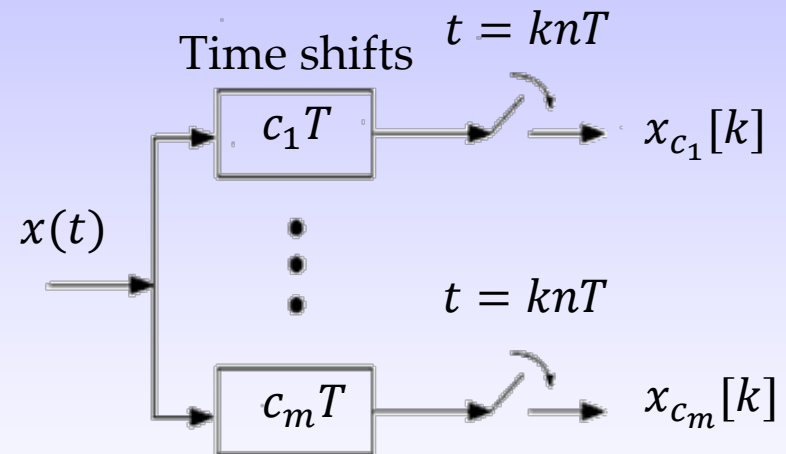
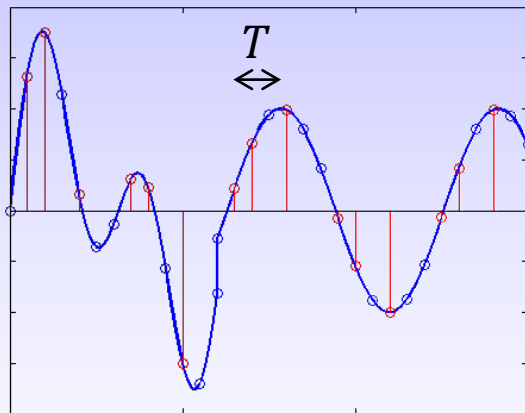
- Uniform sampling: let  $\tilde{R}_U = \{kT\}_{k \in \mathbb{Z}}$  be a uniform sampling set spaced by  $T$ . It holds that  $R_U = \tilde{R}_U$ . If  $T \rightarrow \infty$ , then
  - Sampling set:  $D^-(\tilde{R}_U) \rightarrow 0$
  - Difference set:  $D^-(R_U) \rightarrow 0$  (not enough distinct 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 bandlimited with bandwidth  $1/T$ . The minimal rate for perfect recovery of  $r_x(\tau)$  when using multicoset sampling with  $n$  channels is given by

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

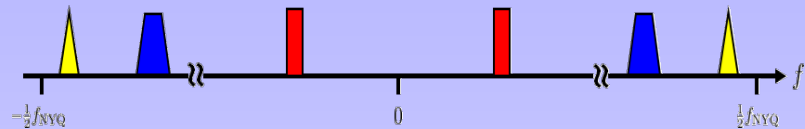
Signal recovery:  $m \geq n$   
Covariance recovery:  $m \gtrsim \sqrt{n}$

- 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

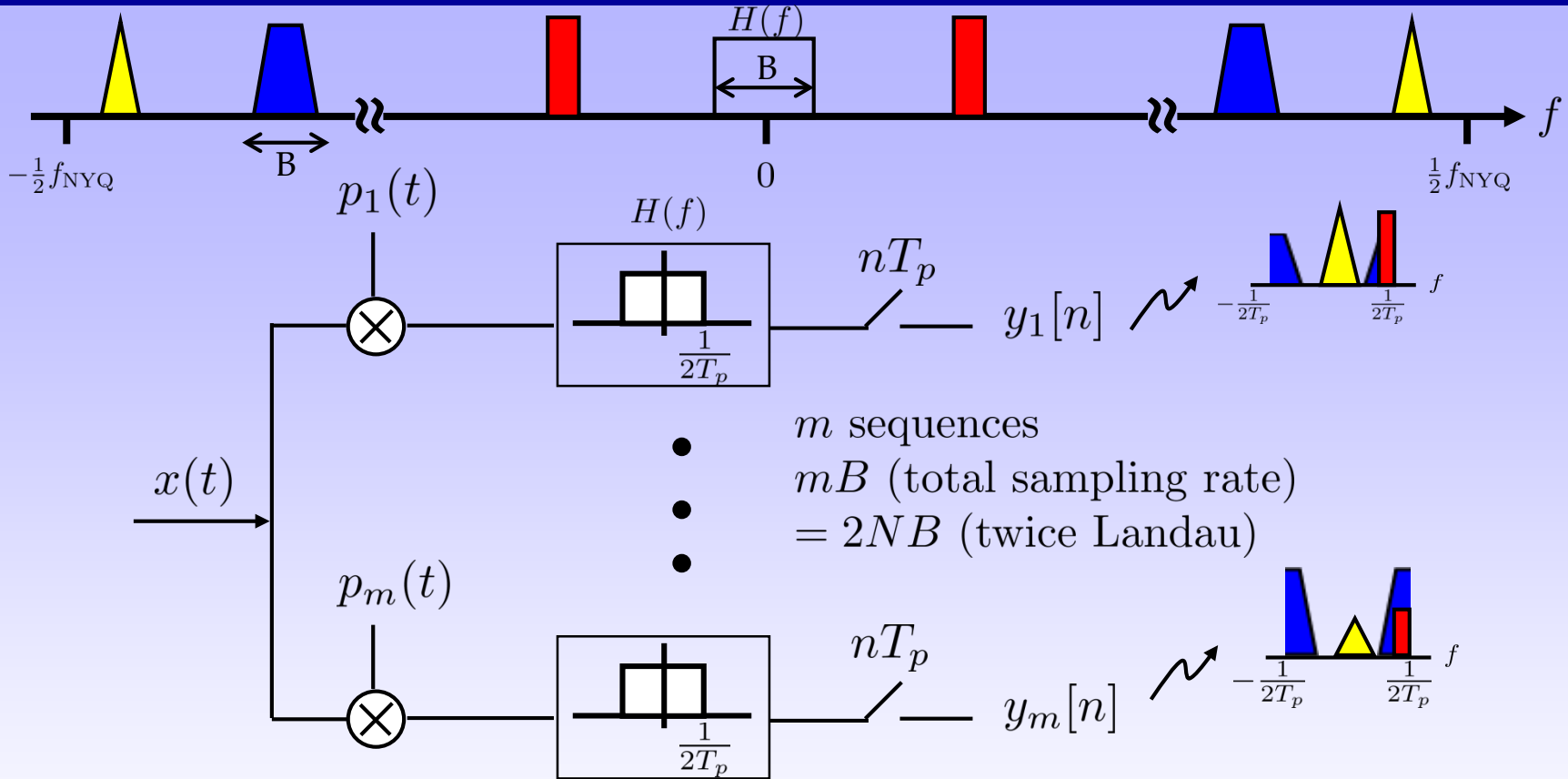
Let  $x(t)$  be sparse with unknown support with occupancy  $\varepsilon < 1/2$ . The minimal rate for perfect recovery of  $r_x(\tau)$  when using multicoset sampling with  $n$  channels is given by

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

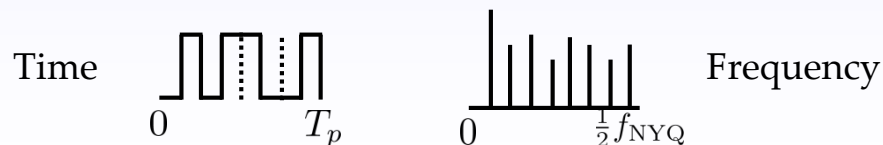
Signal recovery:  $m \geq 2\varepsilon n$   
 Covariance recovery:  $m \gtrsim \sqrt{2\varepsilon n}$

# The Modulated Wideband Converter

Mishali and Eldar, 11

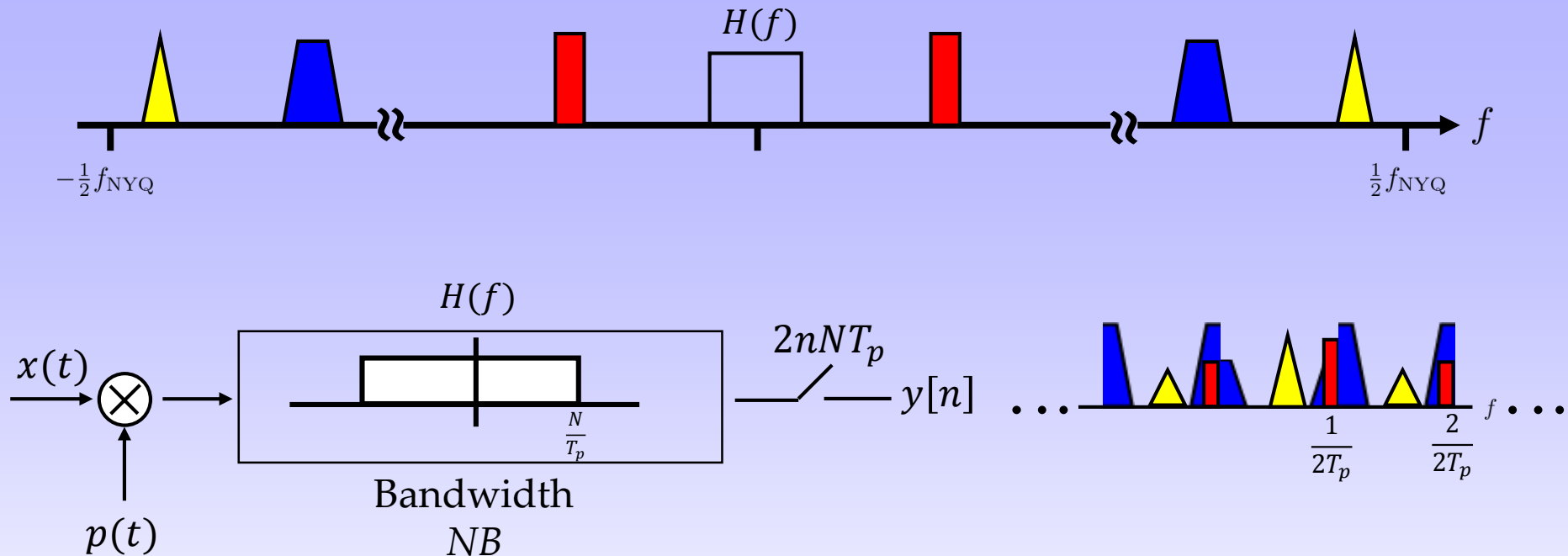


$T_p$ -periodic  $p_i(t)$  gives the desired aliasing effect



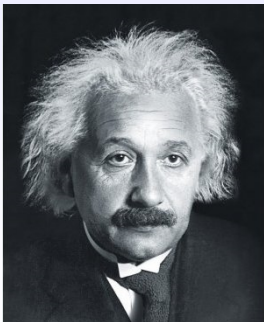
# Single Channel Realization

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



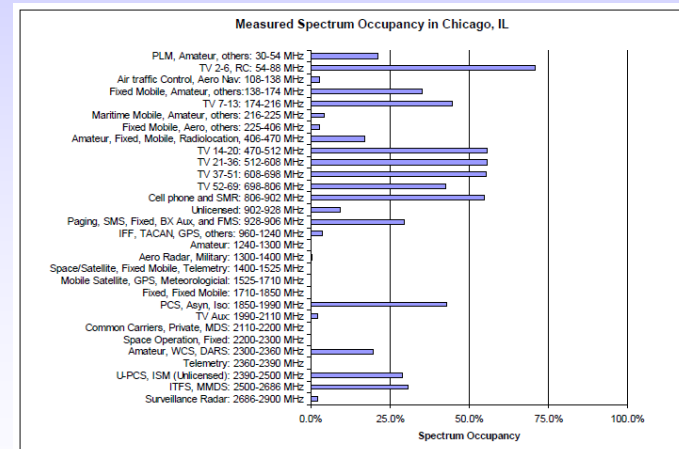
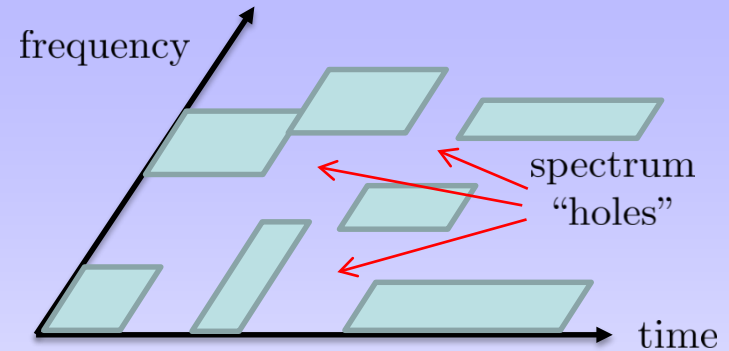
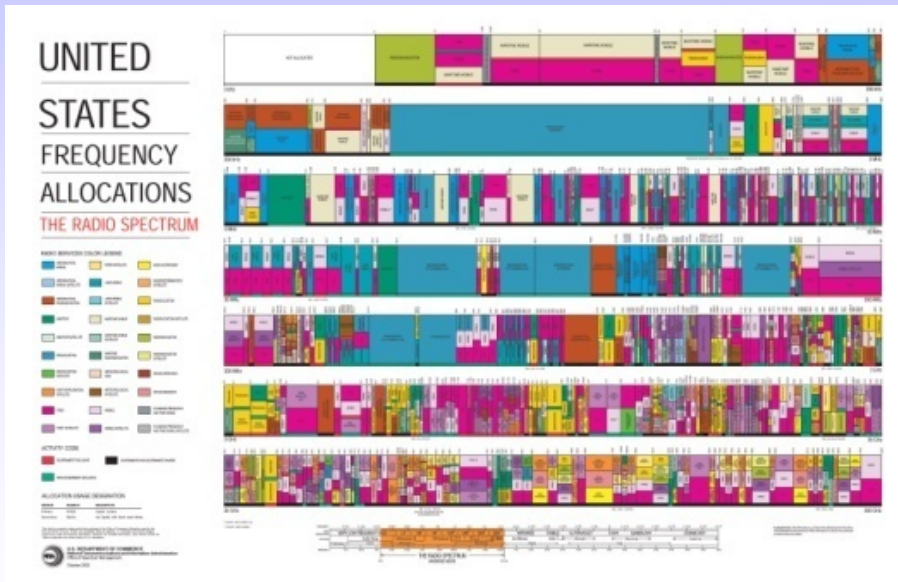
“In theory, theory and practice are the same.  
In practice, they are not.”

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

# Power Spectrum Recovery

Cohen and Eldar, 2013

- Sampling using the MWC we arrive at the equations  $y(f) = Ax(f)$
- Instead of exploiting sparsity in  $x(f)$  we estimate the covariance

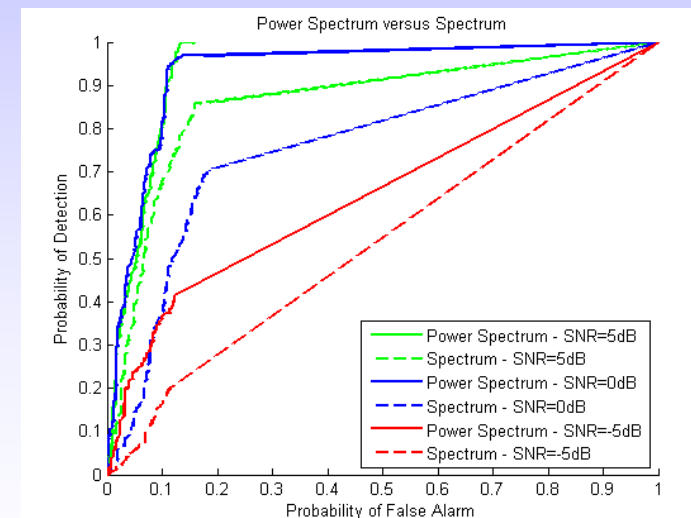
$$R_y(f) = AR_x(f)A^*$$

- From the power spectrum properties

$$R_x(f) = E[x(f)x^*(f)] = \begin{bmatrix} \text{yellow} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \text{cyan} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \text{red} & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & \text{red} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \text{cyan} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \text{yellow} \end{bmatrix}$$

- Can be used to reduce sampling rate by  $\frac{1}{2}$  and to improve robustness

3 signals 80 MHz each  
Nyquist rate 10 Ghz  
Sampling rate 1.04 Ghz



Power spectrum detection outperforms signal detection

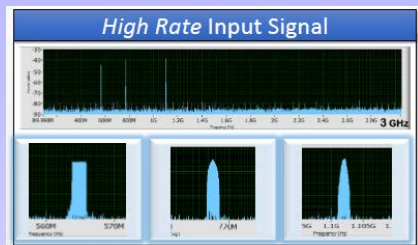


# Nyquist: 6 GHz

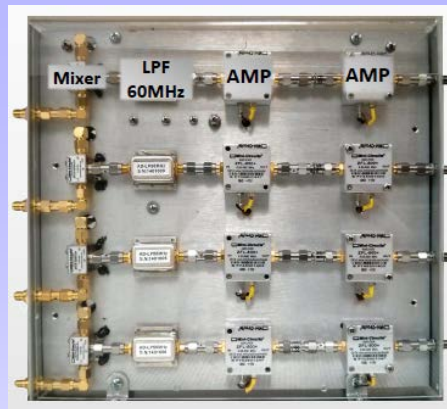
## Sampling Rate: 360MHz

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

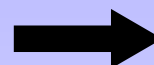
6% of Nyquist rate!



$x(t)$



MWC analog front-end



$y_i[n]$

### Parameters:

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

### Performance:

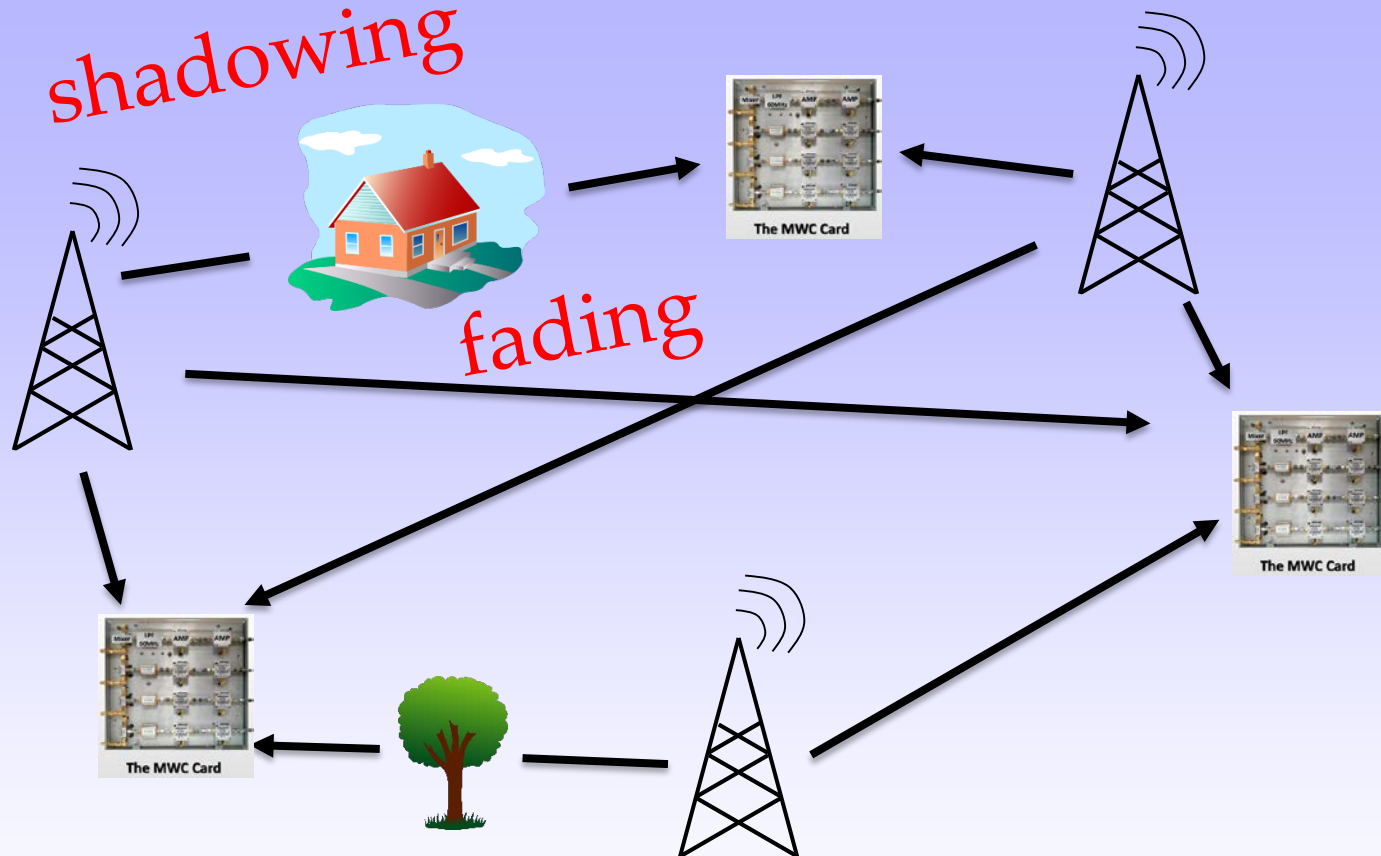
- Wideband receiver mode: 49 dB dynamic range, SNDR > 30 dB
- ADC mode: 1.2v peak-to-peak full-scale, 42 dB SNDR = 6.7 ENOB



# Sub-Nyquist Cognitive Radio System



# Collaborative Spectrum Sensing



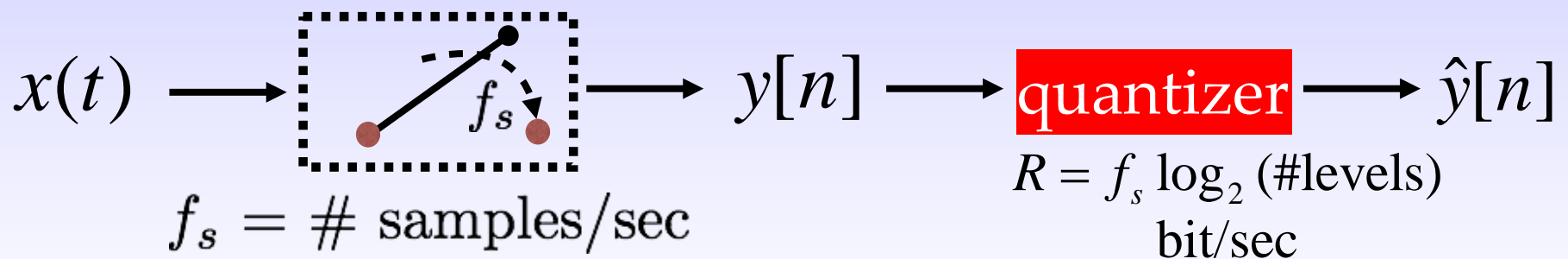
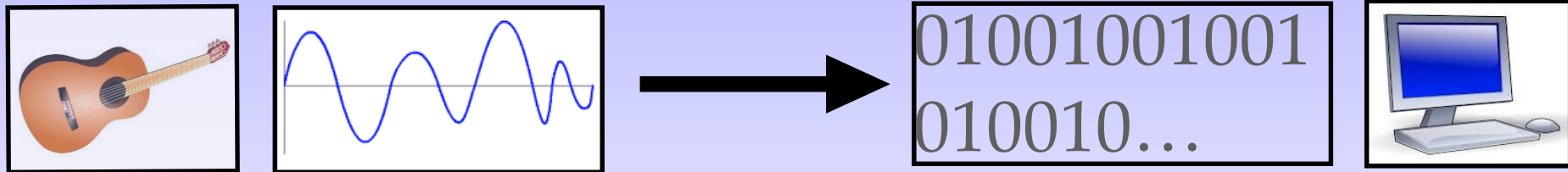
Collaborative spectrum sensing can overcome fading and shadowing issues



# 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?



Sampling Theory

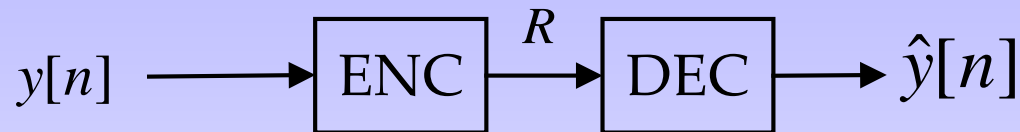
Source Coding [Shannon]

Goal: Unify sampling and rate distortion theory

# Unification of Rate-Distortion and Sampling Theory

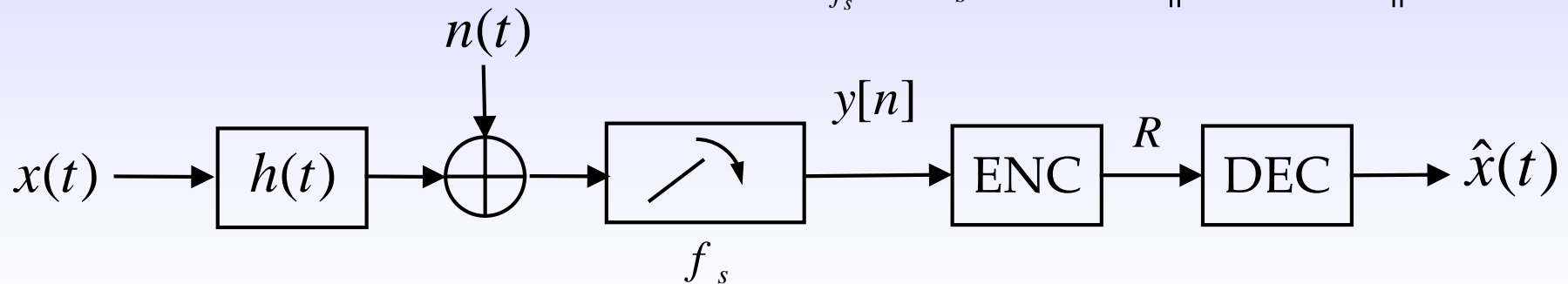
- Standard source coding:

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



- Our question:

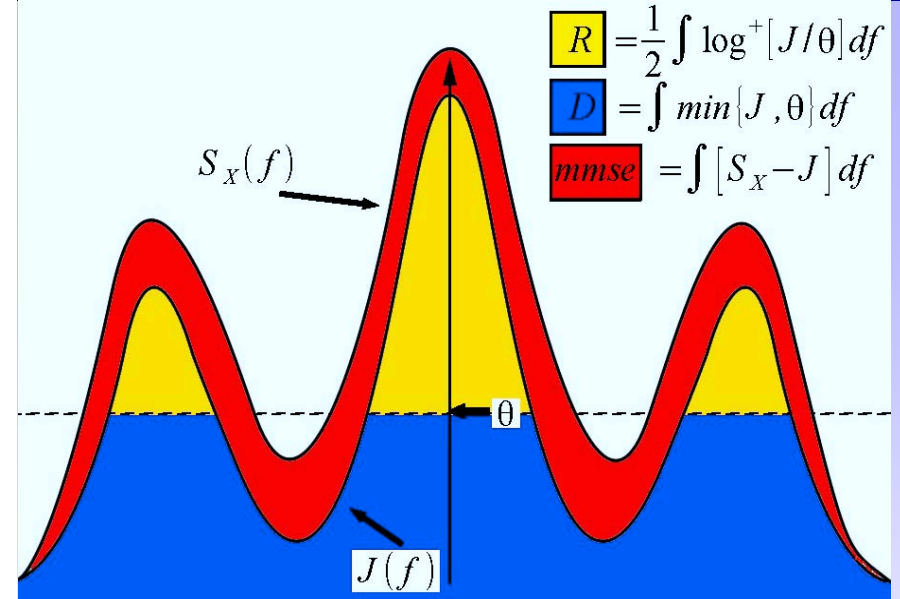
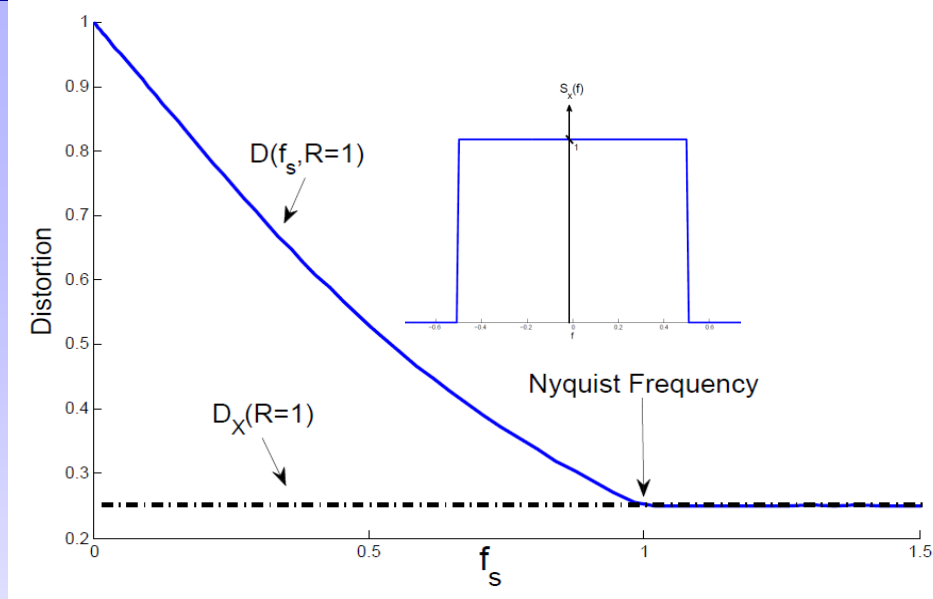
For a given continuous-time process  $x(t)$  and a given bit rate  $R$   
what is the minimal distortion  $\inf_{f_s} D(f_s, R) = \inf \|x(t) - \hat{x}(t)\|^2$



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



# Quantizing the Samples: Source Coding Perspective



Preserve signal components above “noise floor”  $\theta$ , dictated by  $R$   
 Distortion corresponds to mmse error + signal components below noise floor

**Theorem (Kipnis, Goldsmith, Eldar, Weissman 2014)**

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

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

# Optimal Sampling Rate

Shannon [1948]:

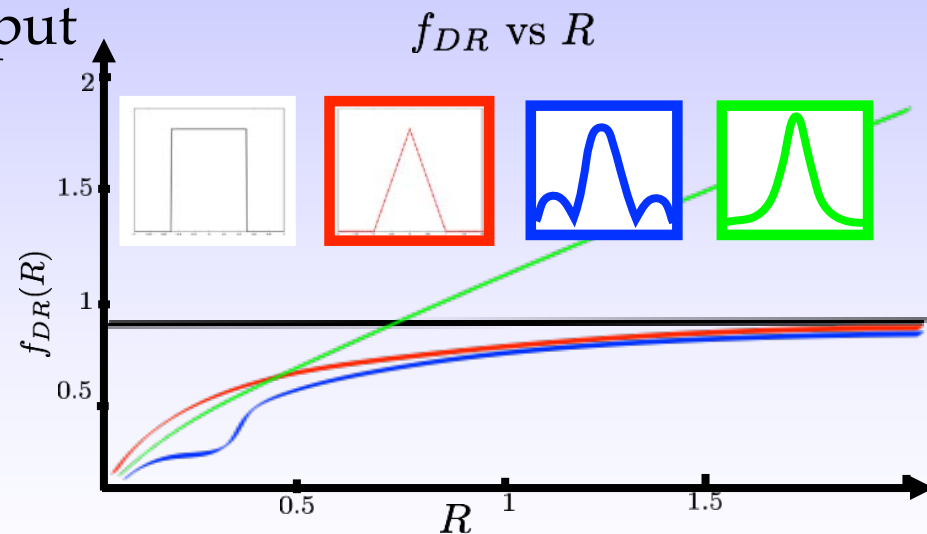
“we are not interested in exact transmission when we have a continuous source, but only in transmission to within a given tolerance”

■ Can we achieve  $D(R)$  by sampling below  $f_{\text{Nyq}}$ ?

■ 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

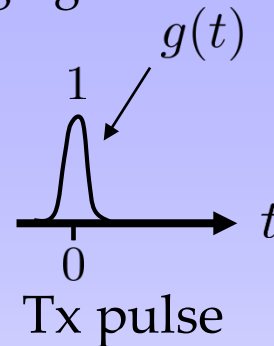
- Relatively simple, radiation free imaging



Cardiac sonography



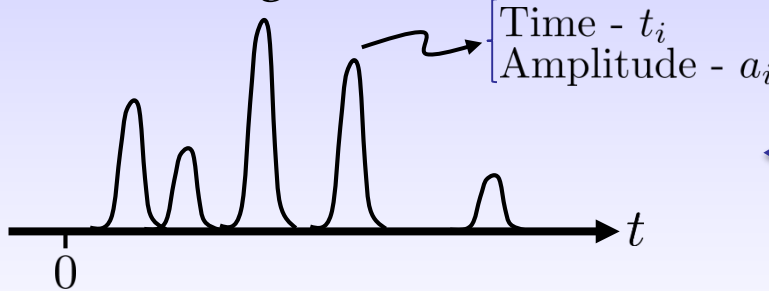
Obstetric sonography



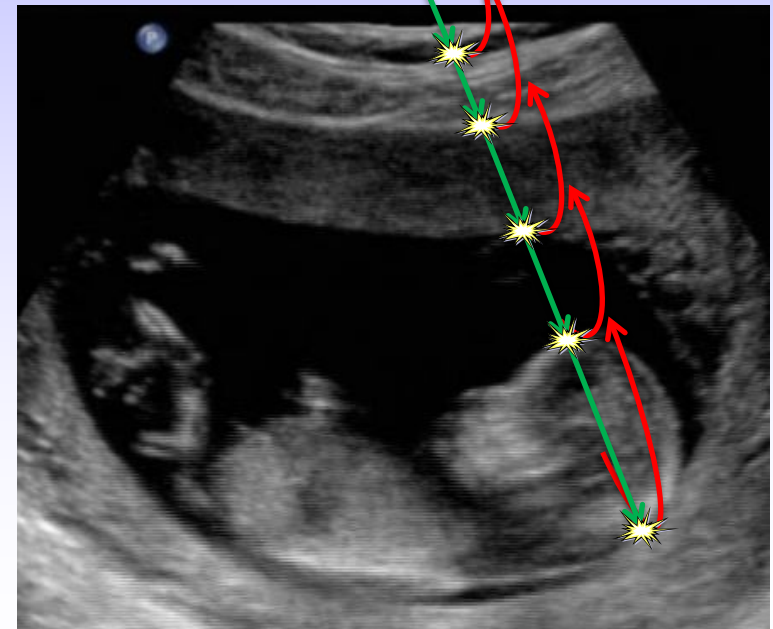
Ultrasonic probe



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



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

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

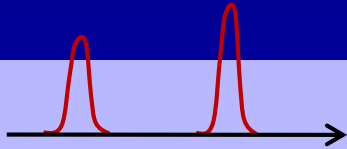
Enable 3D imaging

Increase frame rate

Enable remote wireless ultrasound

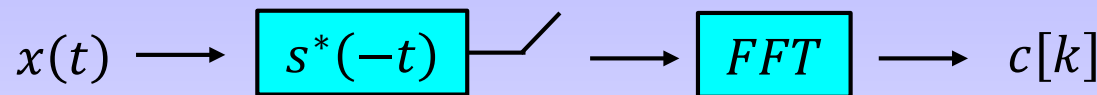


# Streams of Pulses



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:

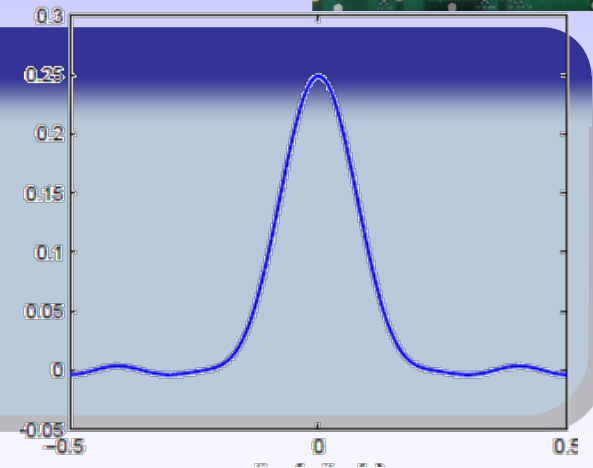


## Theorem (Tur, Eldar and Friedman 11)

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

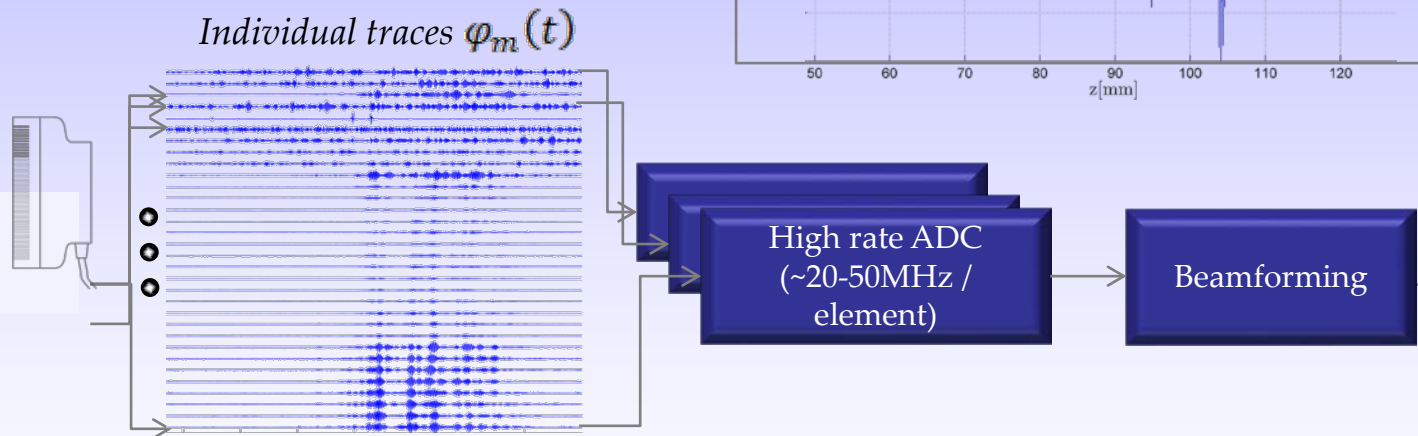
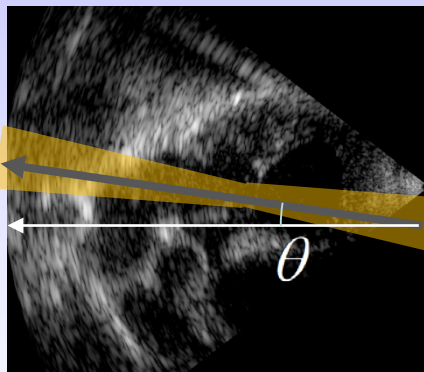
# Conventional Beamforming

Non-linear scaling of the received signals

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

$\gamma_m$  - 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

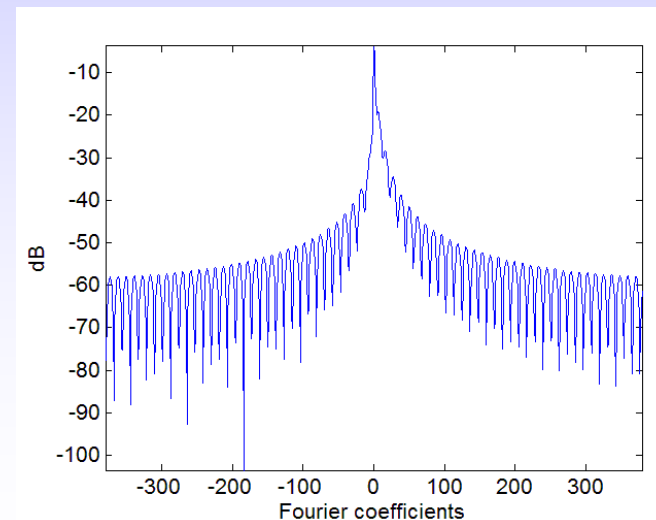
↖  
Fourier coefficient of  
signal at element  $m$

Logic:

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

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

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





# Volumetric Ultrasound Imaging

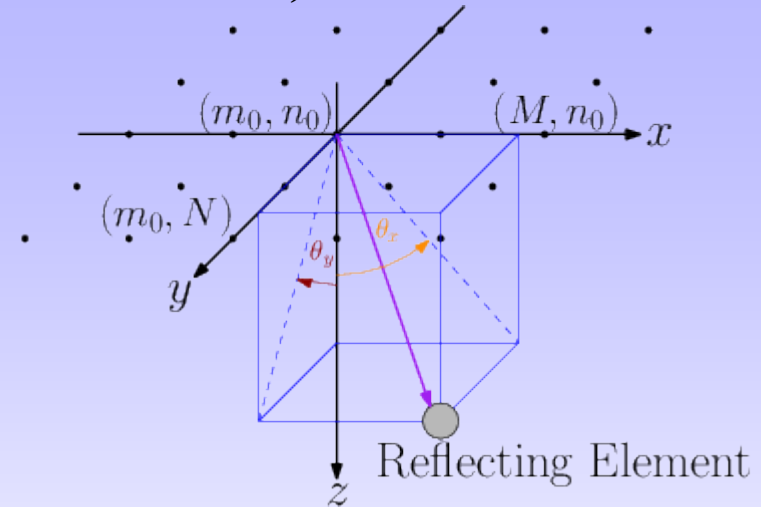
$$\Phi(t; \theta_x, \theta_y) = \frac{1}{N_{\text{RX}}} \sum_{(m,n)} \varphi_{m,n} \left( t - \frac{1}{2} \left( t - \sqrt{t^2 - 4t(\gamma_m x_\theta + \gamma_n y_\theta) + 4|\gamma_{m,n}|^2} \right) \right)$$

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

Fourier coefficient  
of BMF signal

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

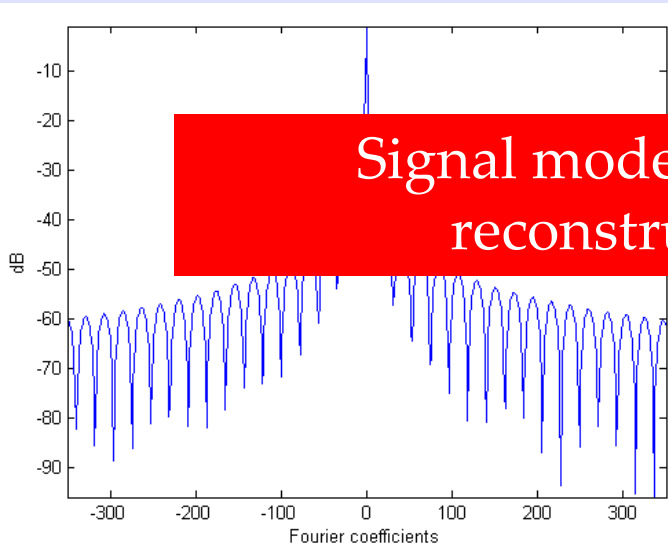
20 elements of  $\{Q_{k,m,n;\theta_x,\theta_y}[l]\}$   
contain more than 95% of the energy



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

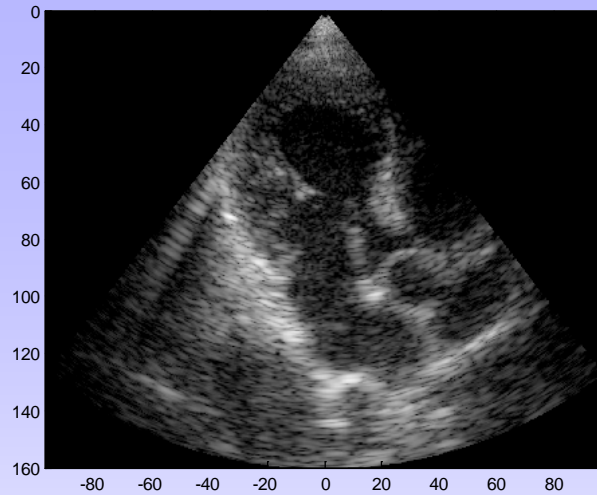
Signal model still holds, allowing the same  
reconstruction technique to be used

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



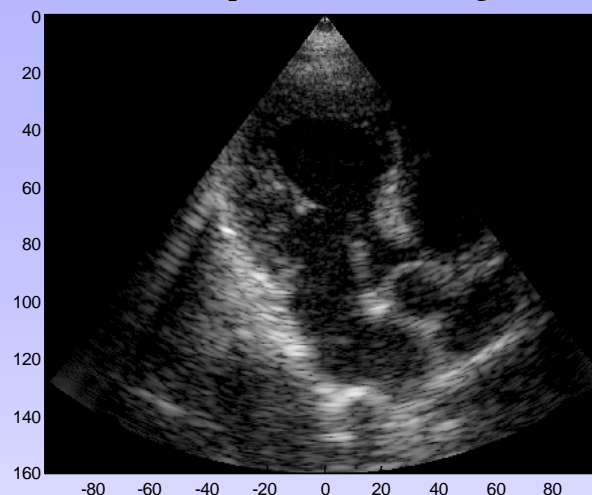
# Ultrasound Results

Standard Imaging



3328 real-valued samples, per sensor  
per image line

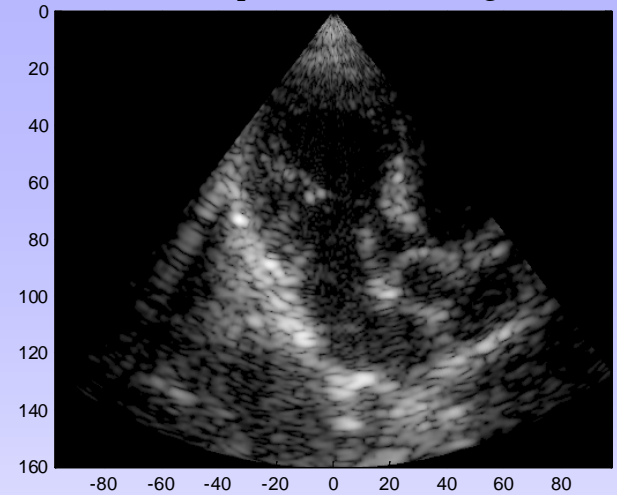
Xampled beamforming



360 complex-valued samples, per sensor per  
image line

~1/10 of the Nyquist rate

Xampled beamforming



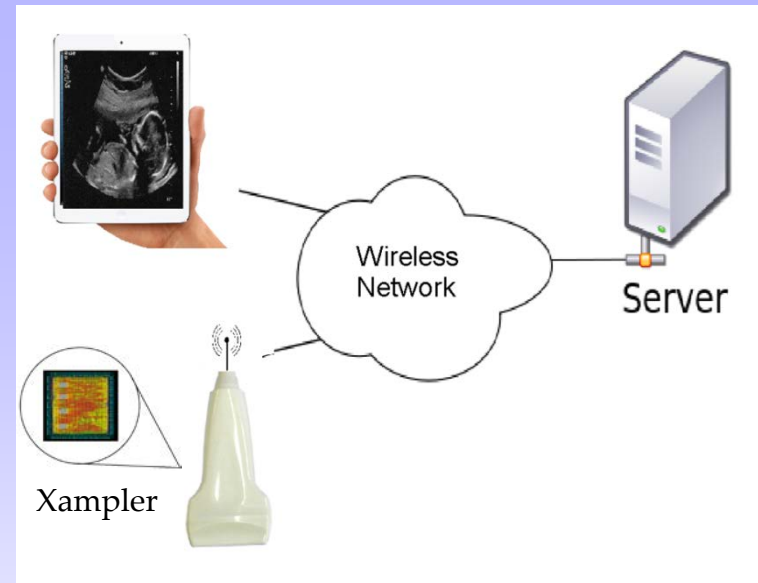
100 complex-valued samples, per sensor per  
image line

~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



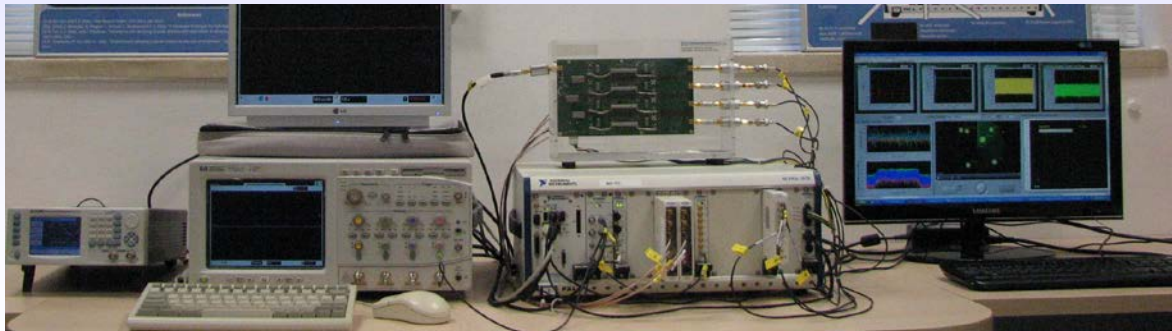
Signal Acquisition Modeling  
and Processing Lab

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





# Technion

Israel Institute of Technology

## Department of Electrical Engineering

■ ■ ■ ■ ■ Electronics  
■ ■ ■ ■ ■ Computers  
■ ■ ■ ■ ■ Communications



Headed by Yonina Eldar

# Conclusions

- Sub-Nyquist sampling of arbitrary signals by exploiting processing task
- Can accommodate nonlinearities in both measurements and processing task with appropriate design of measurement system
- Robust cognitive radio at sub-Nyquist rates
- Sub-Nyquist sampling of almost any signal in the presence of quantization
- Compressed beamforming: ultrasound and radar at very low rates even in the presence of low SNR, smearing and clutter

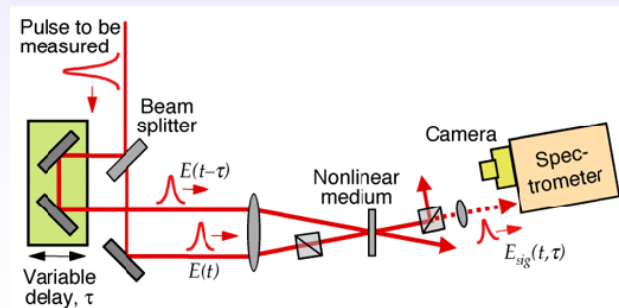
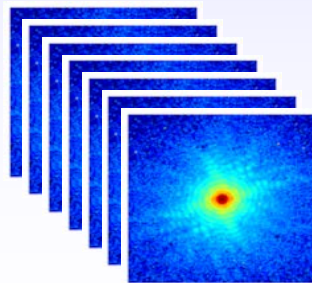
Can substantially reduce rates and accommodate nonlinearities by careful design!



## Part 2:

# Measurement design for phase retrieval

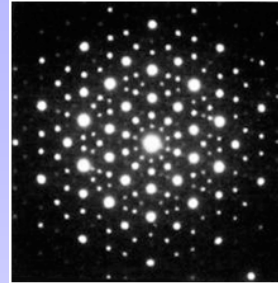
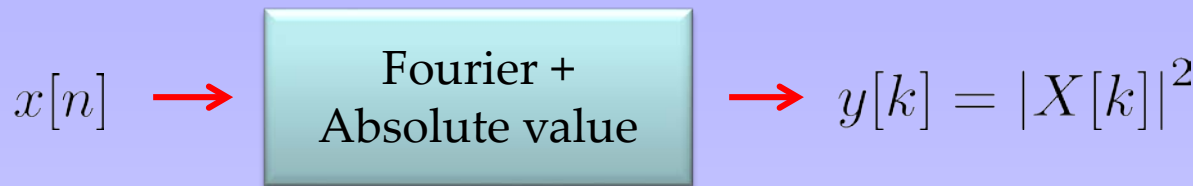
Enabling phase retrieval from Fourier measurements  
using practical devices!



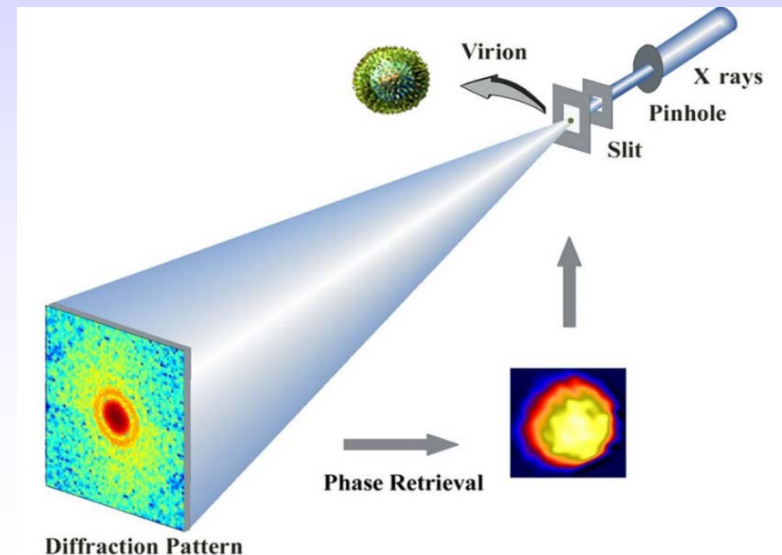


# Phase Retrieval:

## Recover a signal from its Fourier magnitude



- 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
  - Random masks (Candes et. al 13, Bandeira et. al 13)
  - 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)
  - Small number of fixed masks (Jaganathan et. al 15)

} Today

# Analysis of Phase Retrieval

## Analysis of Random Measurements:

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

↑  
random vector

- $4N - 2$  measurements needed for uniqueness

(Balan, Casazza, Edidin 06, Bandeira 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

- $|\langle a_k, x \rangle|^2 = \text{Tr}(A_k X)$  with  $A_k = a_k a_k^T$ ,  $X = x x^T$
- Phase retrieval can be written as  
minimize rank (X)  
subject to  $A(X) = b$ ,  $X \geq 0$
- SDP relaxation: replace rank(X) by  $\text{Tr}(X)$  or by  $\log \det(X + \varepsilon I)$  and apply reweighting
- PhaseCut: semidefinite relaxation based on MAXCUT (Waldspurger et. al 12)

## Advantages / Disadvantages

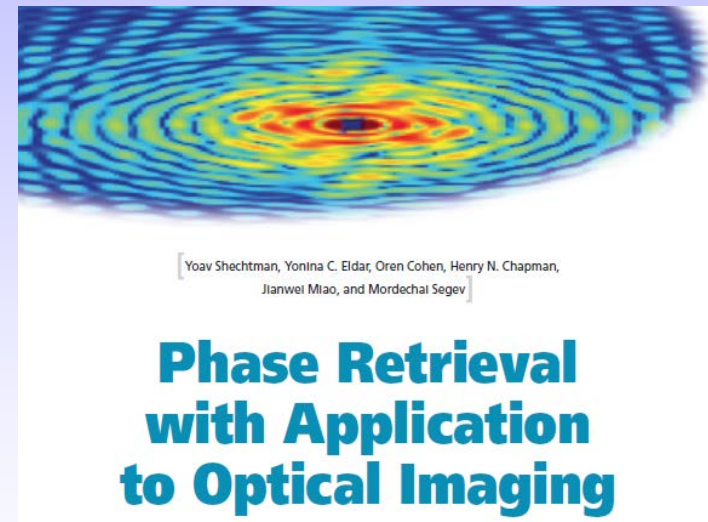
- Yields the true vector whp for  $O(N)$  Gaussian meas. (Candes et al. 12)
- Recovers sparse vectors whp for  $O(k^2 \log(N))$  Gaussian meas. (Candes et al. 12)
- Computationally demanding
- Difficult to generalize to other nonlinear problems

# Provable Efficient Algorithms for Phase Retrieval

- Wirtinger flow achieves the bound  $O(N)$  (Candes et. al 14,15)
- Wirtinger flow+thresholding requires  $O(k^2 \log(N))$  measurements (Li et. al 15)
- 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?

# Nonlinear Sparse Recovery

Yang, Wang, Liu, Eldar and Zhang 15

- We can generalize the theory and algorithms to more general nonlinear models  $y = f(Ax) + w$  where  $x$  is  $k$ -sparse and has length  $n$
- If  $f$  is an invertible function, and  $A$  satisfies RIP-like properties, then  $x$  can be recovered from  $O(k \log n)$  measurements

- Any stationary point of

$$\min_x \frac{1}{m} \sum_{i=1}^m (y_i - f(a_i^T x))^2 + \lambda \|x\|_1$$

with  $\lambda = C\sigma\sqrt{\log n/m}$  will recover the true  $x$  with vanishing error

- Such a point can be found by using a Gradient-decent like algorithm combined with soft thresholding

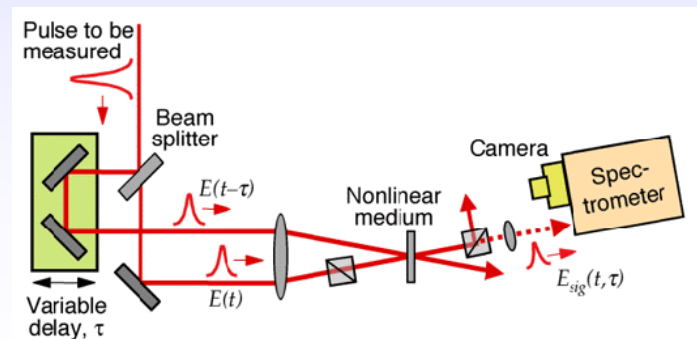
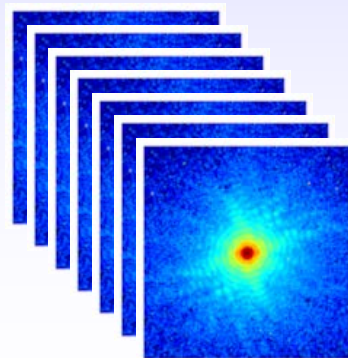
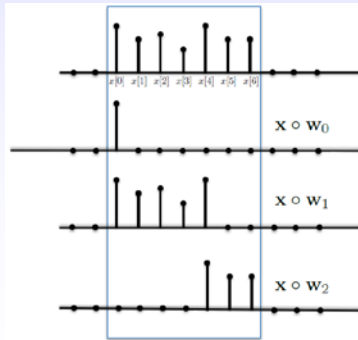
# Design Measurements + Optimization Methods

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

- Create the correlation sequence  $r_k = \text{sum}(\text{diag}(X, k))$
- Any spectral factorization is an optimal LS solution

Optimal LS solution can always be found in polynomial time

# Spectral Factorization

- Let  $R(z)$  be the  $z$ -transform of a correlation sequence  $r_k$
- $R(z)$  can be factored as

$$R(z) = S(z)S^*(1/z^*) = a \prod_{n=1}^{N-1} (1 - c_n z^{-1})(1 - c_n^* z)$$

with roots  $c_n, 1/c_n^*$

- Any  $X(z)$  such that  $R(z) = X(z)X^*(1/z^*)$  is a spectral factor of  $R(z)$
- If all roots of  $X(z)$  are inside the unit circle then  $X(z)$  is minimum phase

## 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 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.  $\hat{X} = \min_{X \succeq 0} \sum_{i=1}^M (y_i - \text{Tr}(f_i f_i^* X))^2$
  2.  $r_k = \text{sum}(\text{diag}(\hat{X}, k))$
  3.  $\max_{X \succeq 0} X(1, 1)$  s.t.  $r_k = \text{sum}(\text{diag}(X, k))$  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)

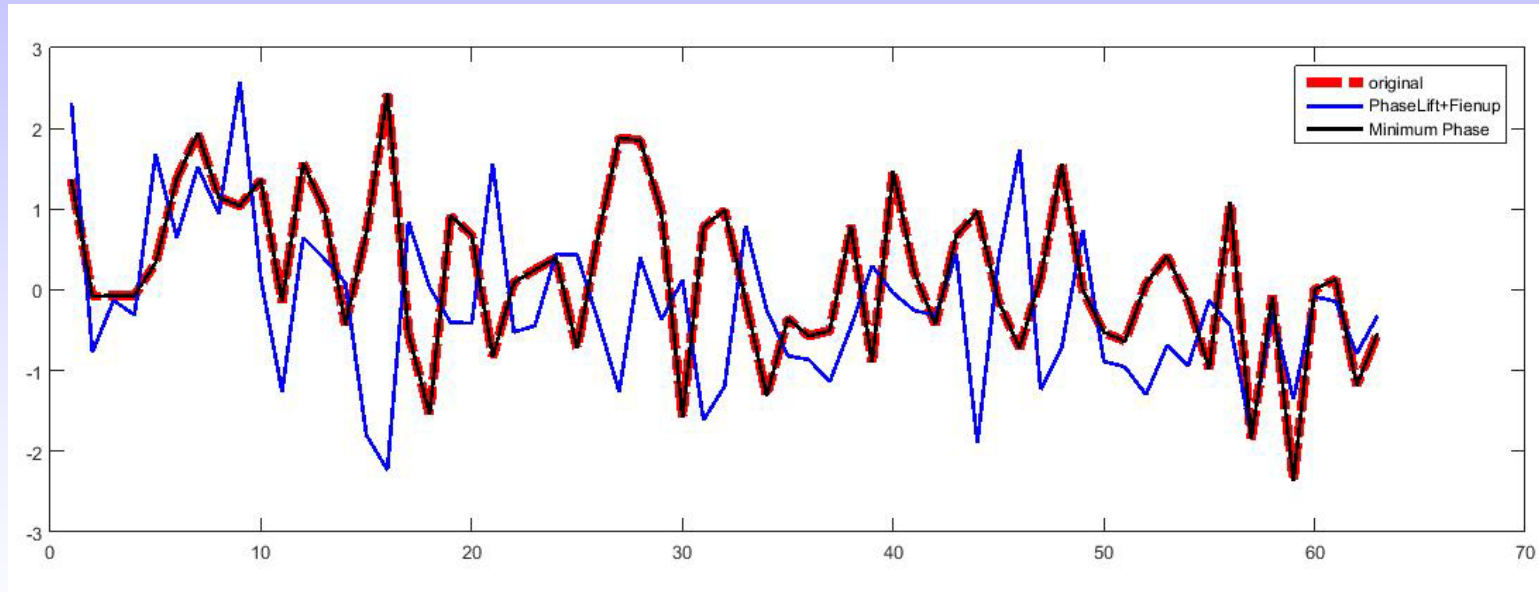
An arbitrary complex signal is minimum phase if  $|x_0| \geq \|x\|_1$

- 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

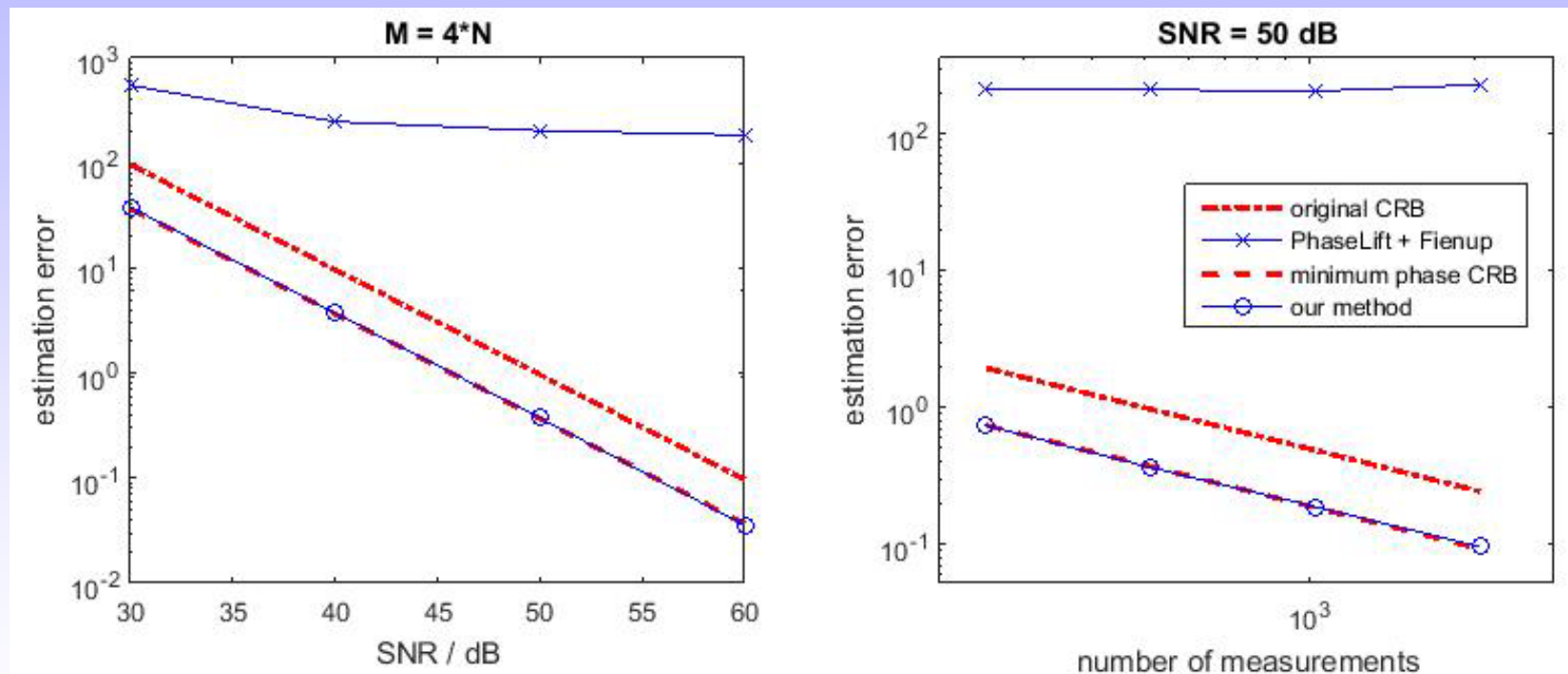


# Simulation: Gaussian Noise

Signal length:  $N = 128$

Left:  $M = 4N$ , SNR increases from 30dB to 60dB

Right: SNR=50dB,  $M$  increases from  $2N$  to  $16N$



We achieve the Cramer-Rao bound in all cases

# Recovery From the STFT Magnitude

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

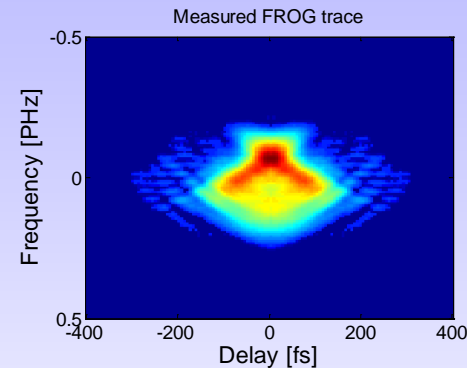
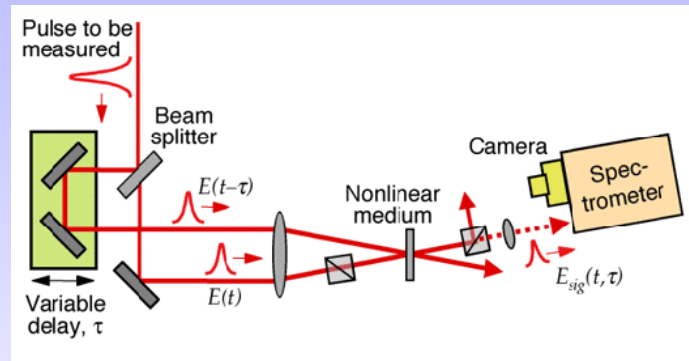
L – step size  
N – signal length  
W – window length

- 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:

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

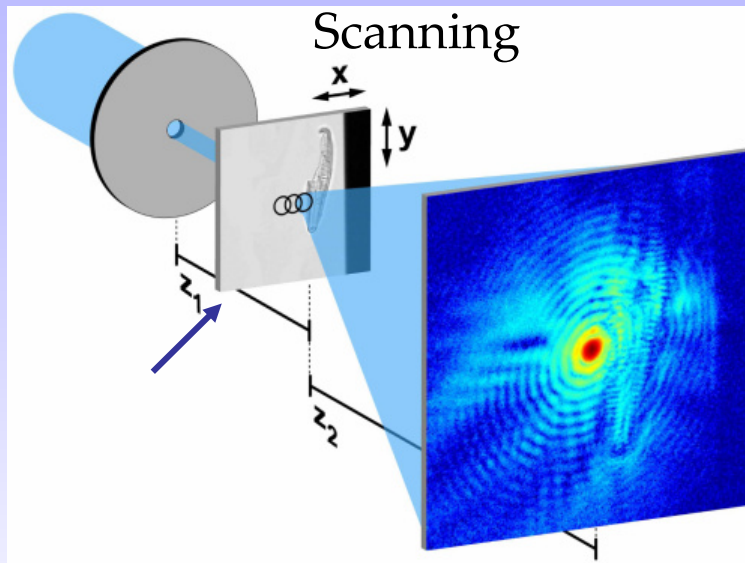
L – step size  
 N – signal length  
 W – window length



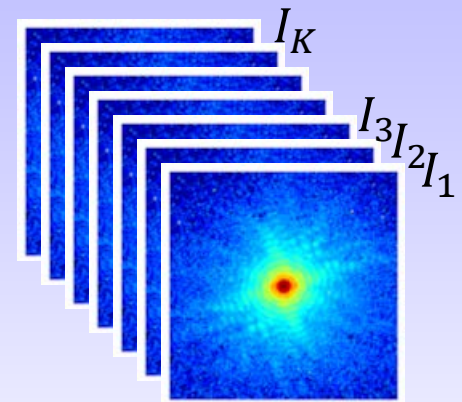
# Ptychography

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

# Current and New Results

## Existing results:

- Fienup-type algorithms – iterate between constraints in time and constraints in frequency
  - Griffin-Lim (Griffin and Lim 84)
  - XFROG (Kane 08)
- No general conditions for unique solution
- No general proofs of recovery

## Our contribution:

- Uniqueness results from STFT measurements
- Guaranteed recovery using semidefinite relaxation
- Extension of results to masked measurements

# 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.  $N \geq 2W - 1$
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.  $L < W \leq N/2$  (segments overlap)

Strong uniqueness for 1D signals and Fourier measurements

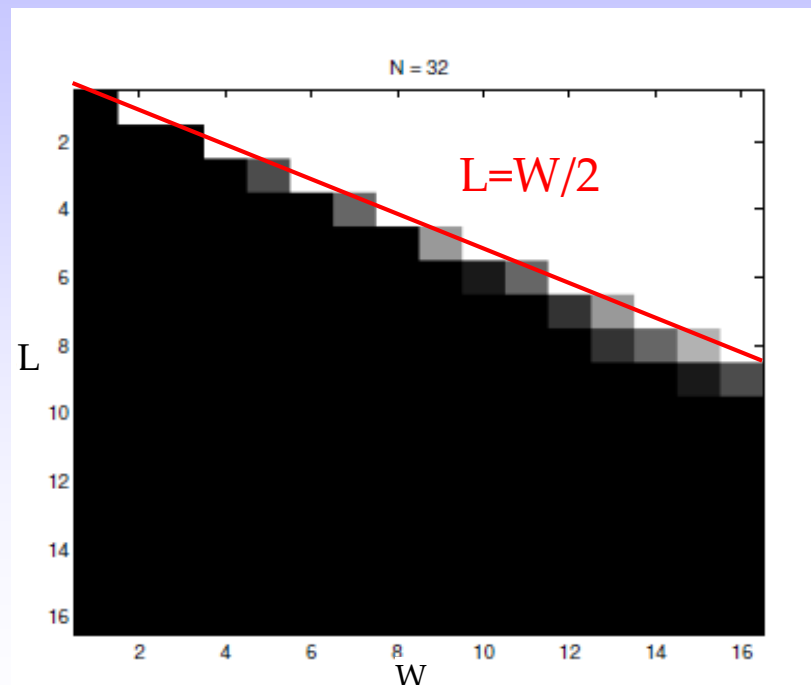
# Recovery From STFT via SDP

## 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  $2 \leq W \leq N/2$

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

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



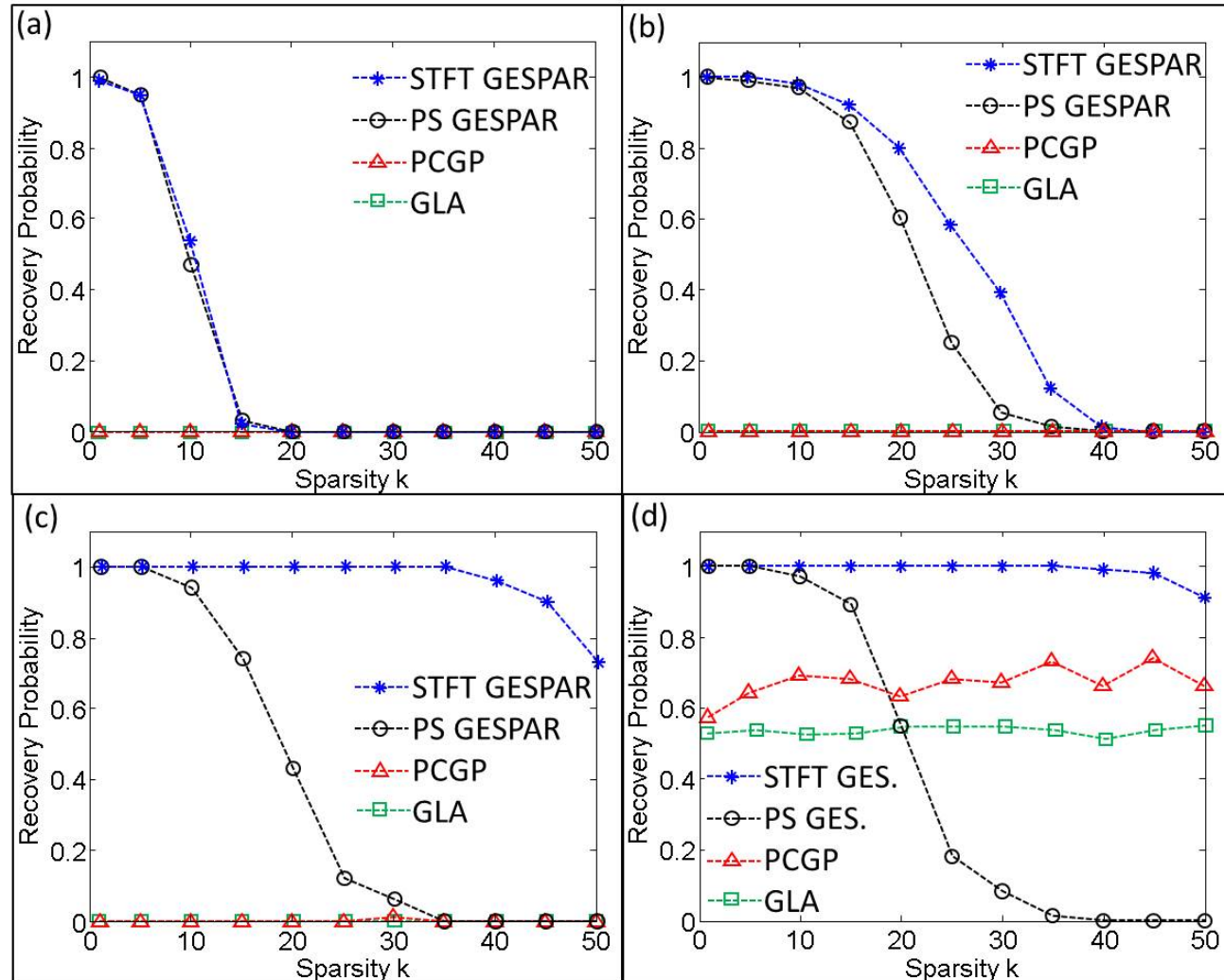
- Can prove the result assuming the first  $W/2$  values of  $x[n]$  are known

# GESPAR for STFT Recovery

Eldar, Sidorenko et. al 15

$W=16$ ,  $N=64$  varying  $L$

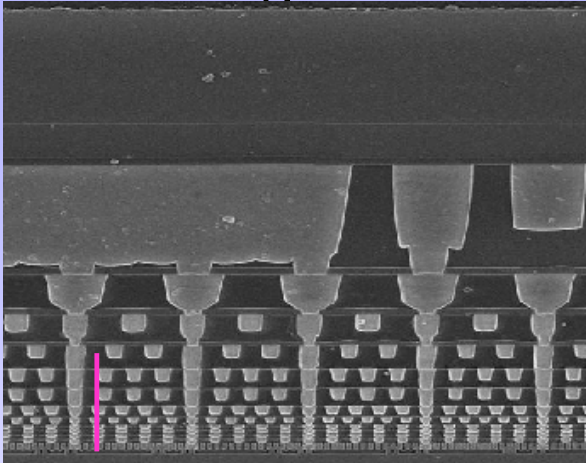
$L=2,4,8,16$  (no overlap)



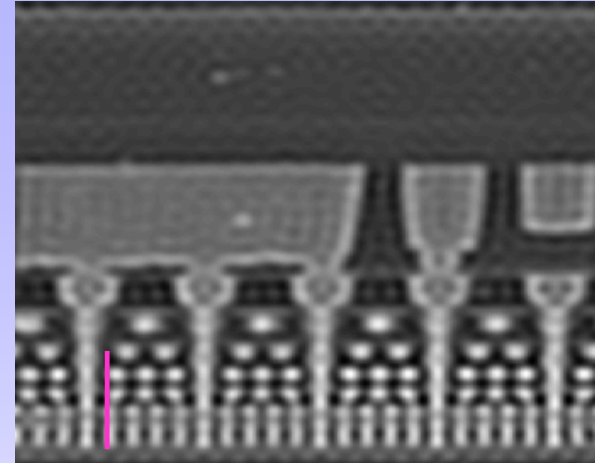
PS GESPAR uses oversampled FT with the same number of measurements as STFT

# Recovery from Lowpass Data

Original

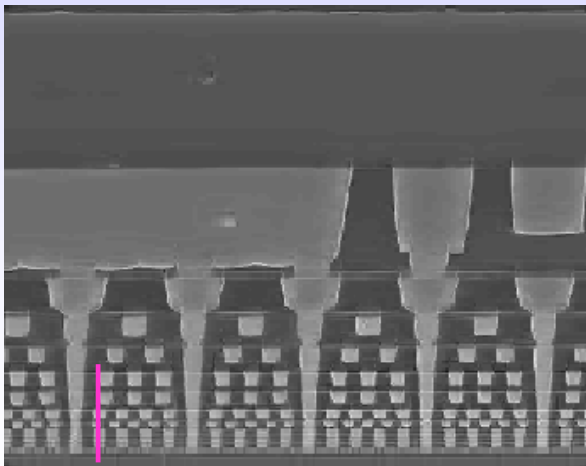


Blurred

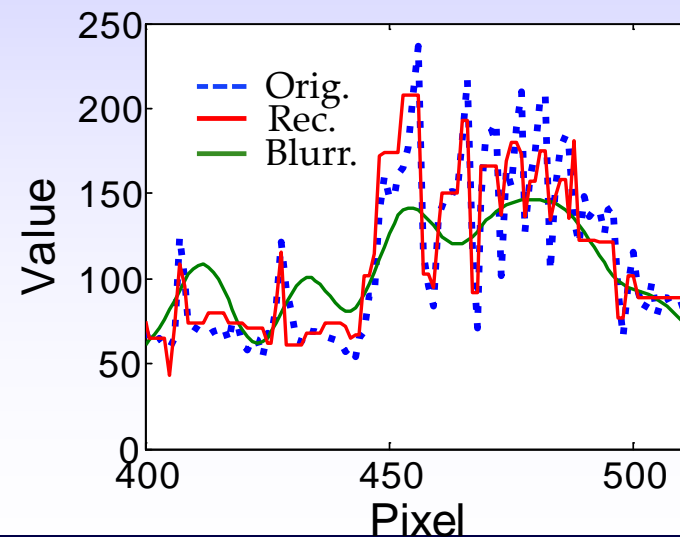


Sidorenko et al

Reconstruction



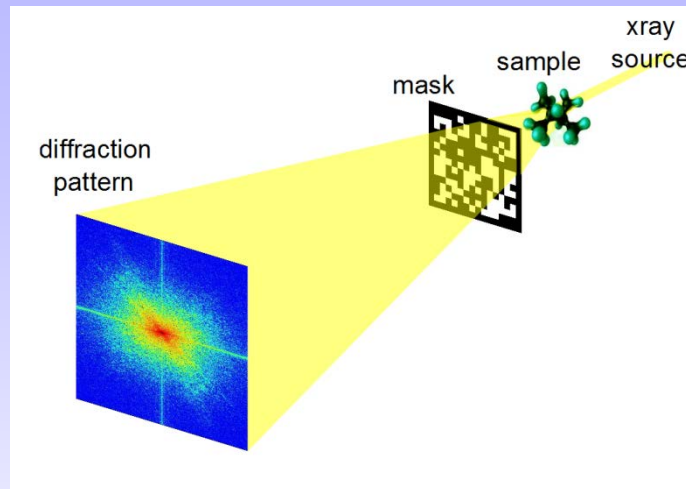
Line out



# Coded Diffraction Patterns

Candes, Li and Soltanolkotabi 13

- Modulate the signals with masks before measuring  $x[k]d[k]$



- Assume an admissible mask  $d[k]$  for which  $Ed = 0$ ,  $Ed^2 = 0$ ,  $E|d|^4 = 2E|d|^2$
- Using  $O(\log^4 N)$  masks the signal can be recovered
- Gross et. al improved to  $O(\log^2 N)$
- Difficulty: random masks, large constants

# Efficient Masks

- 2 deterministic masks each with  $2N$  measurements suffice!

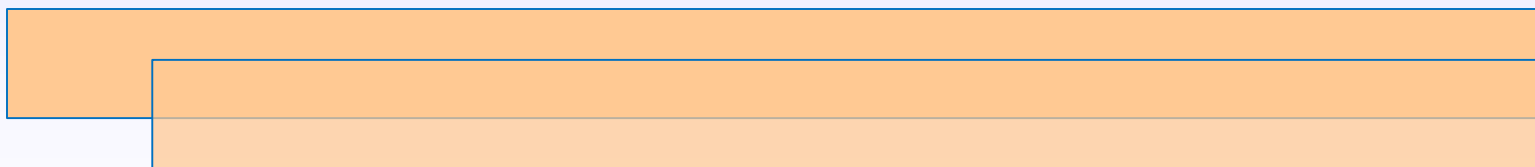
## Theorem (Jaganathan, Eldar and Hassibi 15)

The STFT magnitude of length  $2N$  taken with two distinct masks  $d_1[n], d_2[n]$  uniquely determines almost any nonvanishing  $x[n]$  if

1. for each  $n$ ,  $d_1[n] \neq 0$  or  $d_2[n] \neq 0$
2. there is at least one value  $m$  for which  $d_1[m], d_2[m] \neq 0$

- To recover all signals  $x[n]$  we can use specific masks which also guarantee stable recovery via SDP

$$d_1[n] = 1, 0 \leq n \leq N-1 \quad d_2[n] = \begin{cases} 0 & n = 0 \\ 1 & 1 \leq n \leq N-1 \end{cases}$$



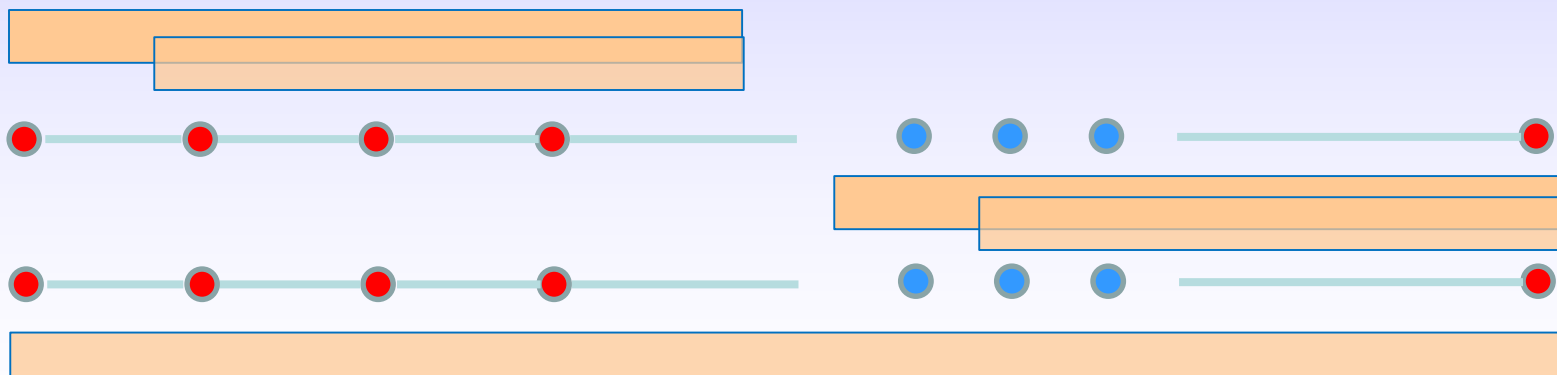


# Efficient Masks

## Theorem (Jaganathan, Eldar and Hassibi 15)

Any  $x[n]$  with  $x[0] \neq 0$  can be recovered in a robust fashion from the two previous masks using a semidefinite relaxation

- Idea can be extended to allow for 5 masks, and each one measured with only  $N$  measurements (previous: 2 masks,  $2N$  measurements each)



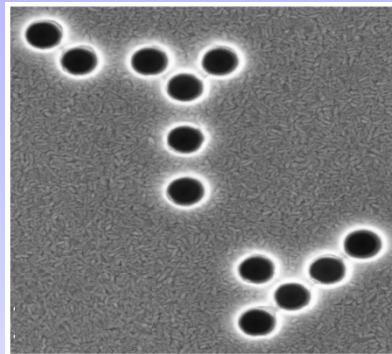
# Sparsity Based Subwavelength CDI

Szameit et al., Nature Materials, 12

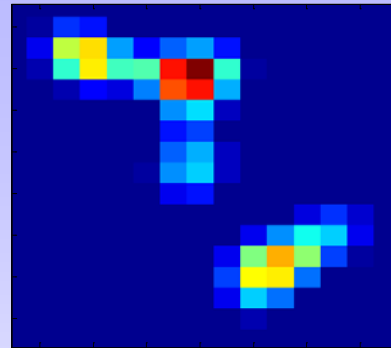
Circles are  
100 nm  
diameter

Wavelength  
532 nm

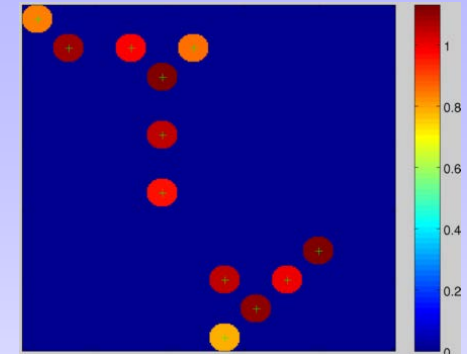
SEM image



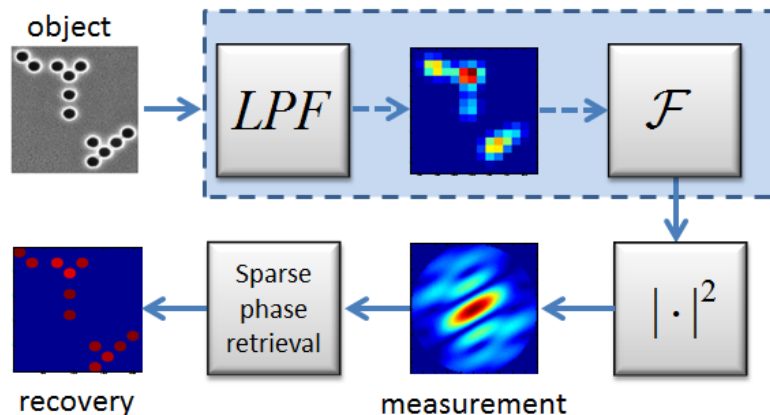
Blurred image



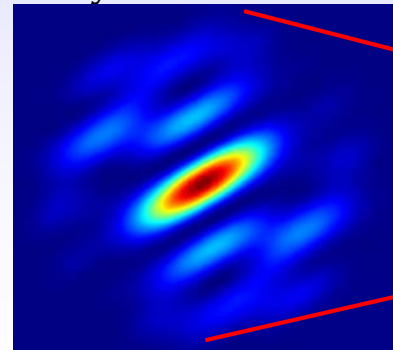
Sparse recovery



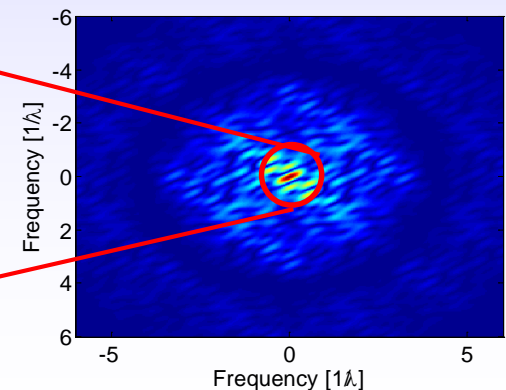
Free space propagation



Diffraction-limited  
(low frequency)  
intensity measurements



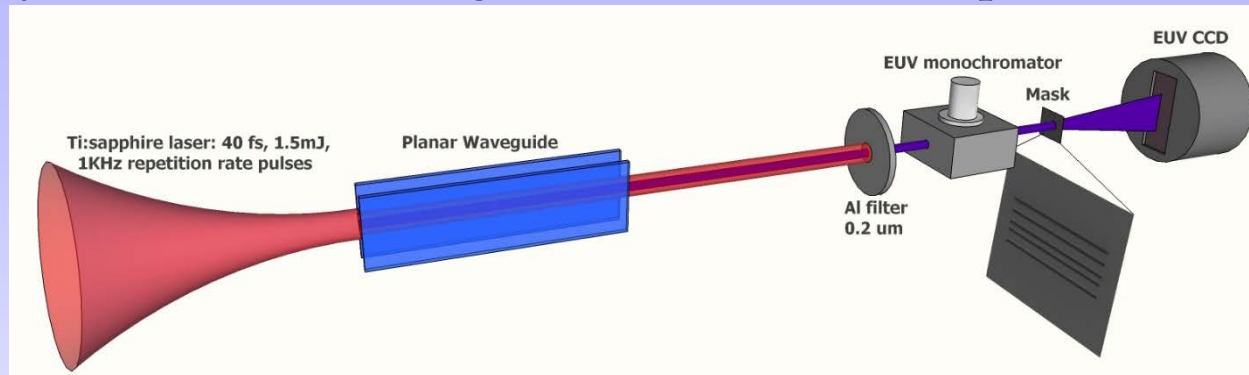
Model  
Fourier transform



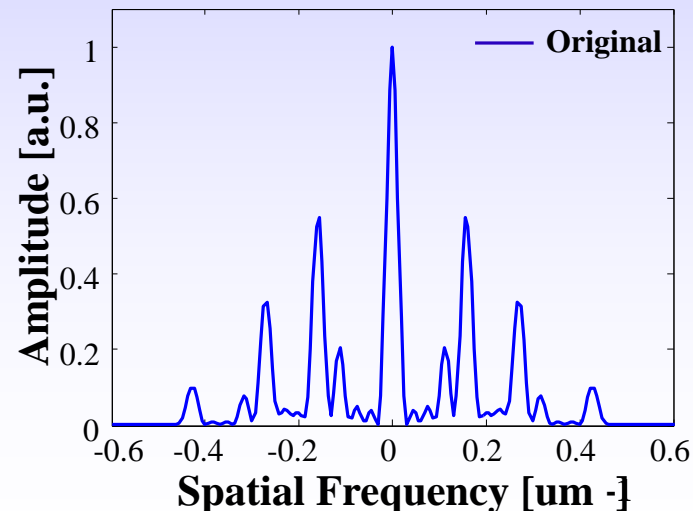
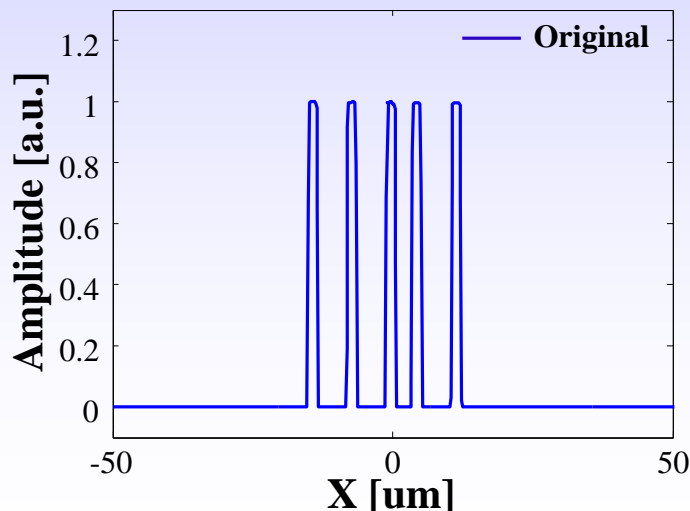
# Experimental 1D CDI

Sidorenko et al., Nature Comm, 15

- 1D phase retrieval is ill posed, admitting multiple solutions
- Sparsity can remove ambiguities due to 1D + super-resolution



Experimental Setup

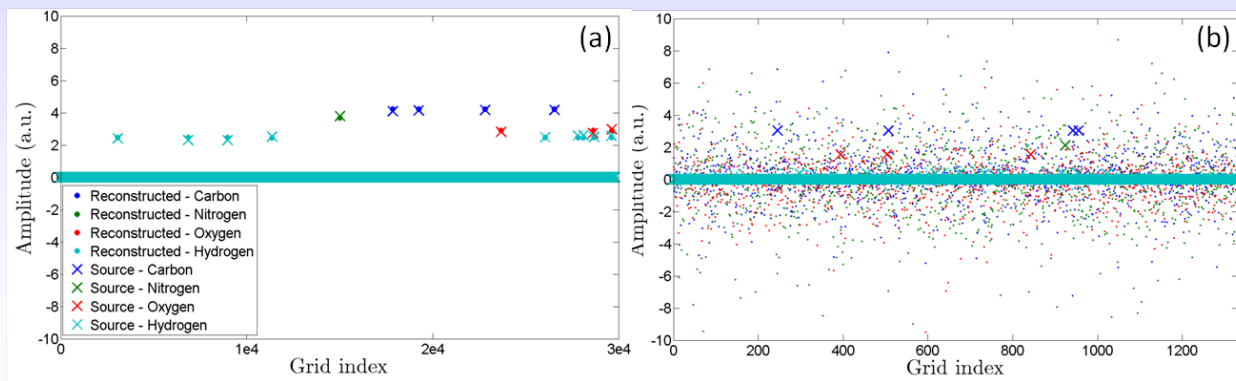
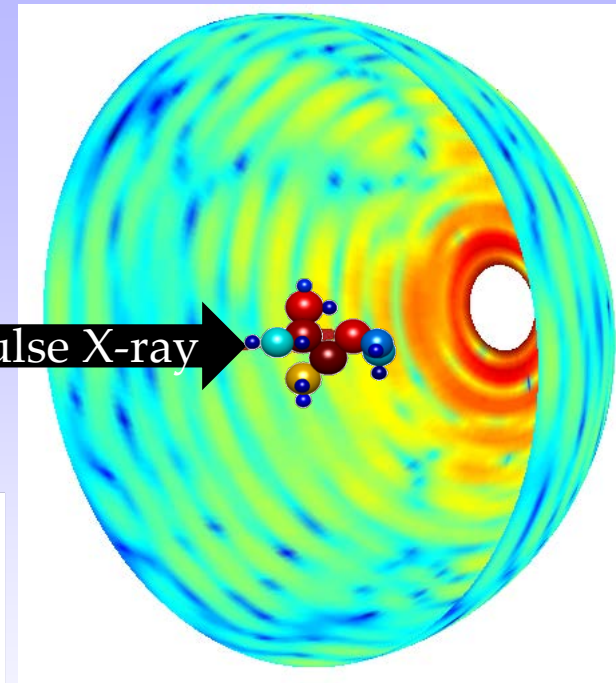


# Sparsity Based Ankylography

K.S. Raines et al. Nature 463, 214 ,(2010).  
Mutzafi et. al., (2013).

## Concept:

- A short x-ray pulse is scattered from a 3D molecule combined of known elements. The 3D scattered diffraction pattern is then sampled in a single shot
- Recover a 3D molecule using 2D sample



# Conclusions

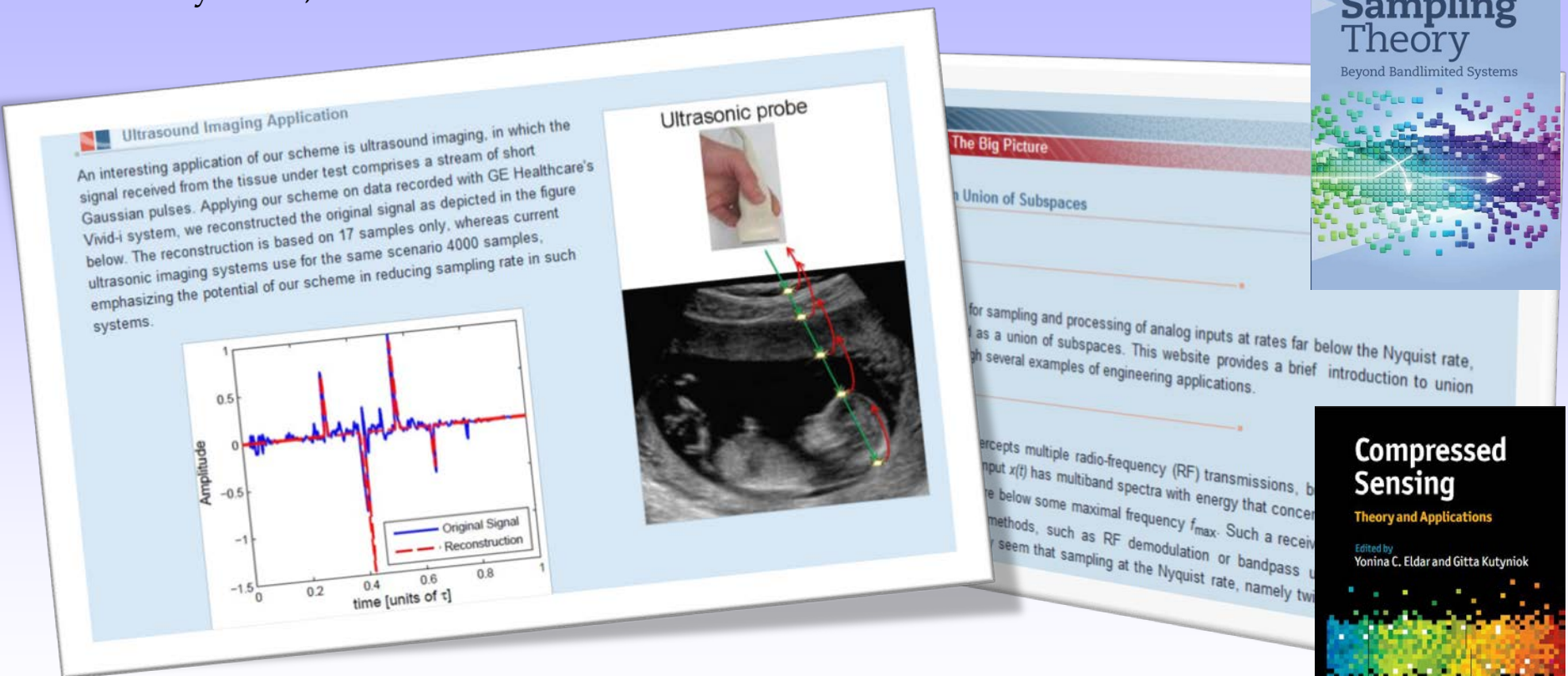
- Compressed sampling and processing of many analog signals even without structure
- Wideband sub-Nyquist samplers in hardware
- Many new applications like wireless ultrasound
- Merging information theory and sampling theory
- Importance of measurement design in phase retrieval

**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](http://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](http://www.sampl.technion.ac.il)

**SAMPL**  
Signal Acquisition Modeling  
and Processing Lab

Overview People Projects Sub-Nyquist Prototypes Software Publications Events News & Announcements

**People** **Technical Staff**

**Eli Shushan**  
Systems Engineering Consultant  
E-mail: [icon] [email]  
Phone: 4794  
Office: 307-0, Fishbach Bldg.

**Robert Ifraimov**  
Engineer  
E-mail: [icon] [email]  
Office: 4784 - Fishbach Bldg.

**Alon Eilam**  
System Engineer  
E-mail: [icon] [email]  
Phone: 1955  
Office: 306-1, Fishbach Bldg.

**Aviad Arobas**  
Computer assistant  
E-mail: [icon] [email]  
Office: 304-2, Fishbach Bldg.

**Project Proposals**

**Dictionary Reduction in Magnetic Resonance Fingerprinting**

**Lab Team**

Our team is constantly exploring innovative sampling methods that combine theory and practice

**Project Details:**

**Increased sampling capacity in optical communication**  
Project ID: 2705  
Abstract: In optical communication, signals are transmitted at a very high rate (here, we consider 20G) but only one bit per interval. A simple sampler can then be used that only needs one bit per interval. The question is whether we can use MWC ideas in order to allow for more bits. The goal of this project is to determine whether we can achieve gain in term of the information bits using the MWC.  
Description: File not available  
Supervisor(s): **Debby Cohen**  
Student(s): Omri Lev, Tal Wiener

**Cyclostationary detection from sub-Nyquist samples for Cognitive Radio - demo**  
Project ID: 2474  
Abstract: The goal of this project is to demonstrate cyclostationary detection from sub-Nyquist samples in a cognitive radio scenario. The demo will consist of an analog part, the MWC, that mixes and samples the analog input signal at a low rate. Then, the cyclic spectrum is reconstructed from the sub-Nyquist samples. Last, the carrier frequencies and symbol rates are estimated from the cyclic spectrum (using the existing algorithm developed in a previous project).  
Description: **PDR**  
Supervisor(s): **Debby Cohen, Shahar Tsiper**  
Student(s): Noga Shaul, Ulad Pollick

The extensive use of computers in every aspect of modern life vastly increases the need for digital signal processing and signal sampling. Sampling research is becoming more applications-oriented: discoveries in this field quickly find their way into a wide array of products ranging from cell phones and medical imaging equipment to defense technologies.

The **SAMPL Lab** focuses on sampling, modeling and processing of continuous-time and discrete-time signals. Traditional systems treat the sampling and processing stages separately and require sampling at the well-known Nyquist rate. In contrast, **SAMPL** research focuses on new design paradigms in which sampling and processing are designed jointly in order to exploit signal properties already in the sampling stage. This approach has the potential to drastically reduce the sampling and processing rates well below the Nyquist rate, typically considered as the ultimate limit for analog to digital conversion. The proposed prototypes can reduce power consumption, hardware size and complexity, and enable efficient wideband sensing.

The laboratory facilitates the transition from pure theoretical research to the development, design and implementation of prototype systems. **SAMPL Lab** integrates these new ideas directly into technology by close collaboration with leading industrial companies. Systems developed in the lab demonstrate the research results and technology advances in a wide range of applications ranging from biomaging such as 3D ultrasound and MRI through communications, laser optics, cognitive radio and radar systems.

**SAMPL Lab Details**

Contact Us: [SAMPL.Lab.Team@technion.ac.il](mailto:SAMPL.Lab.Team@technion.ac.il)

**Lab Demos**

**Open Positions**

**more project proposals**

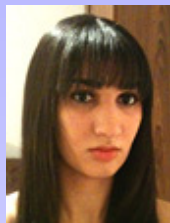
**more pictures**

**more info**

**more demo movies**

# Acknowledgements

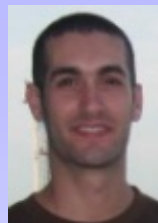
- Graduate students (Past and Present)



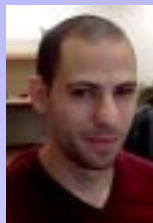
Deborah  
Cohen



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!*