



# Sampling in the Age of Sparsity

Martin Vetterli EPFL & UC Berkeley

# Acknowledgements

## Sponsors and supporters

- Conference organizers, PinaMarziliano in particular!
- NSF Switzerland and ERC Europe: thank you!
- Qualcomm, Chong Lee in particular, for collaboration

## Co-authors and collaborators

- Thierry Blu, CUHK
- Yue Lu, Harvard
- PinaMarziliano, NTU
- David Gontier, ENS-EPFL
- YannBarbotin, EPFL
- Ali Hormati, EPFL
- MihailoKolundzija, EPFL
- Jury Ranieri, EPFL
- Jay Unnikrishnan, EPFL

## Discussions and Interactions

- V. Goyal, MIT
- M. Unser, EPFL

# Outline

## 1. Introduction

The world is analog but computation is digital!

## 2. Sampling: The linear case

Classic results and generalizations: Known locations

## 3. Application: Sampling physics

The plenacoustic function and sampling wave fields

Mobile sensing

## 4. Sampling based on timing

Integrate and fire

## 5. Sampling: The non-linear case

Finite rate of innovation sampling

Compressed sensing

## 6. Applications: The non-linear case

Multichannel sampling

Super-resolution imaging

## 7. Conclusions: Power game!

# The Situation

- The world is analog
- Computation is digital
- How to go between these representations?



**Ex: Audio, sensing, imaging, computer graphics, simulations etc**

# From Analog to Digital



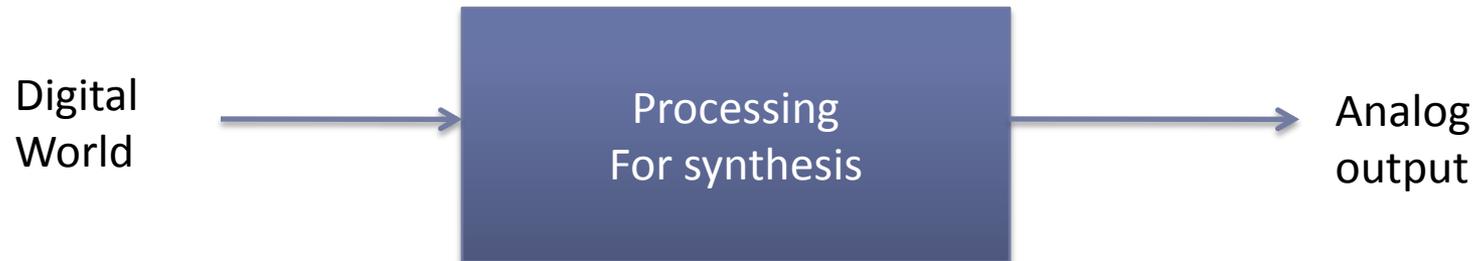
# The Situation (bis)

The world is analog, computation is digital, and

- You want an analysis



- You want a synthesis



# The Dual Situation

- Information is digital
- Communication is in the analog world
- How to go between these representations?



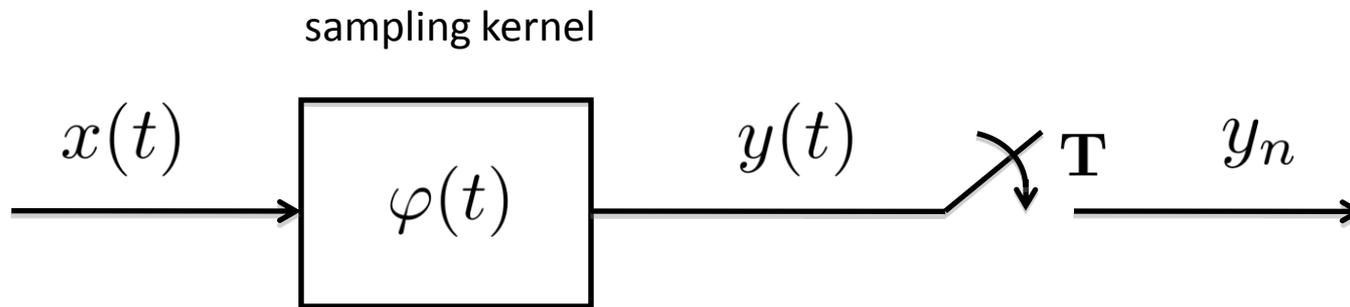
**Ex: Mobile communication, fiber cables, storage etc**

# The Question:

Given a class of objects, like a class of functions (e.g. bandlimited)

And given a sampling device, as usual to acquire the real world

- Smoothing kernel or low pass filter
- Regular, uniform sampling
- That is, the workhorse of sampling!



Obvious question: When does a minimum number of samples uniquely specify the function?

$$x(t) \iff y_n$$

# Kernel and Sampling Rate

## 1. About the observation kernel:

### Given by nature

- Diffusion equation, Green function  
Ex: sensor networks

### Given by the set up

- Designed by somebody else, it is out there  
Ex: Hubble telescope

### Given by design

- Pick the best kernel  
Ex: engineered systems, but constraints

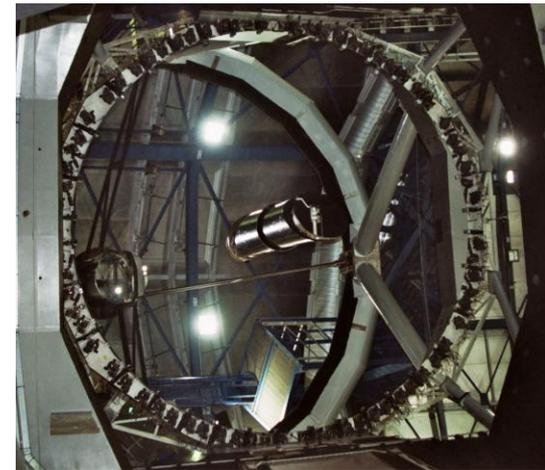
## 2. About the sampling rate:

### Given by problem

- Ex: sensor networks

### Given by design

- Usually, as low as possible (UWB)

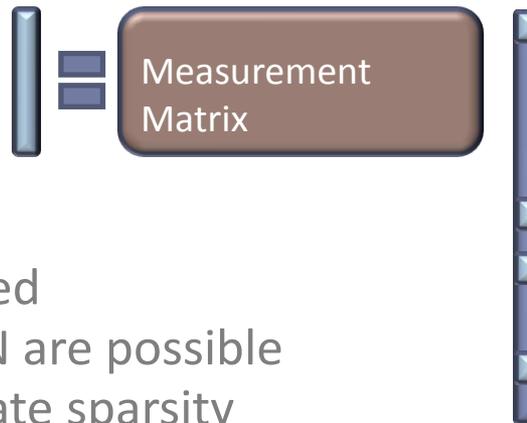


ESO, Chile

# Variation: Compressed Sensing

**Finite dimensional problem:  $K$  sparse vector in  $N$  dimensional space**

- $x$ : Input in  $\mathbb{R}^N$  but only  $K$  non zero elements
- $y$ : Output in  $\mathbb{R}^M$ , where  $M < N$
- $F$ : Frame sensing matrix  $M$  by  $N$
- Ill posed inverse problem....
- Key:  $K < M \ll N$



## Questions

- Can this be inverted
- What sizes  $K$ ,  $M$ ,  $N$  are possible
- What if approximate sparsity
- What algorithms

**Problem is non-linear in the location of non-zero entries of  $x$ !**

# Variation: Multichannel Sampling

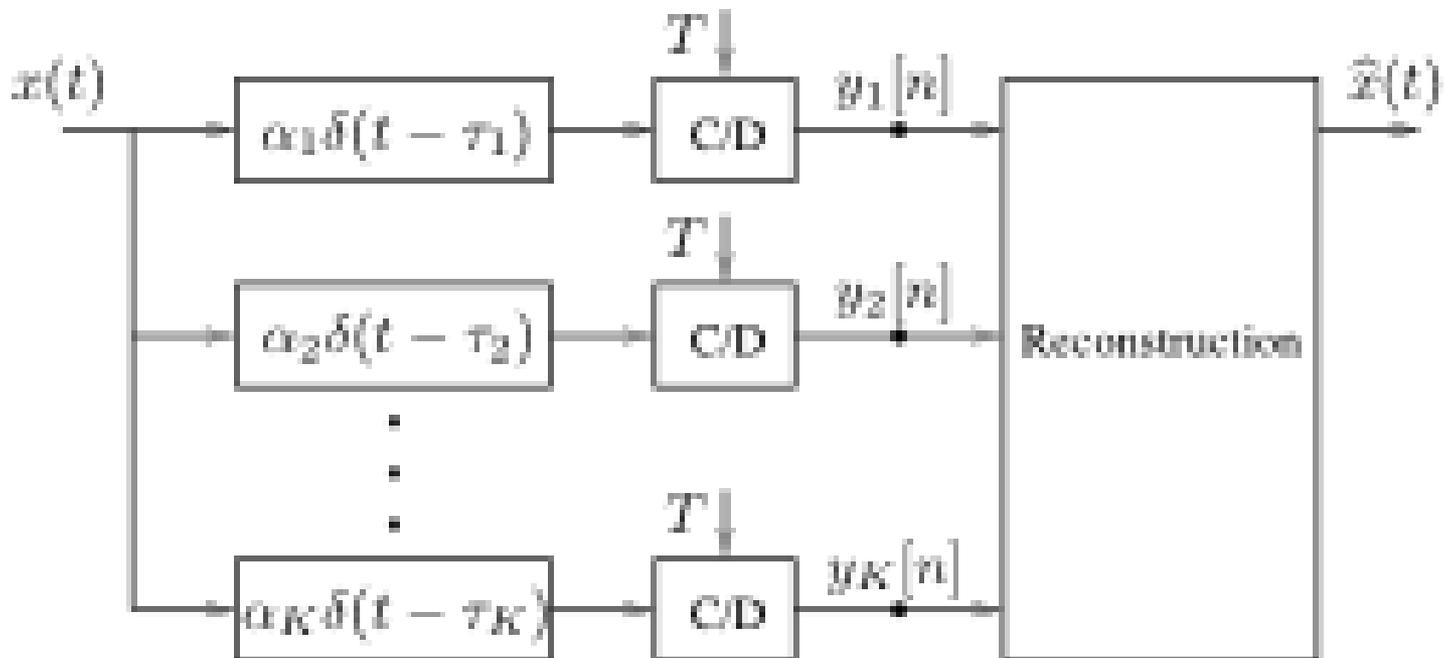
Signal is observed in  $K$  different channels

Sampling rate can be diminished by at most  $K$

Shifts, however, are unknown

Some redundancy needed to find the unknown shifts

**Problem is non-linear in the shifts!**

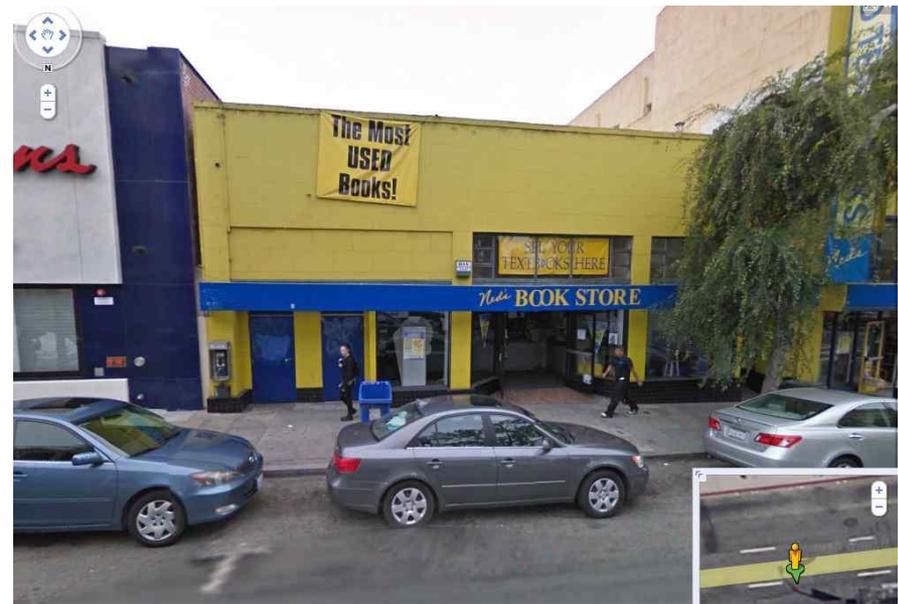


# Are these Real Problems?

## Google Street view as a popular example

How many images per second to reconstruct the real world

What resolution images to give a precise view of the world

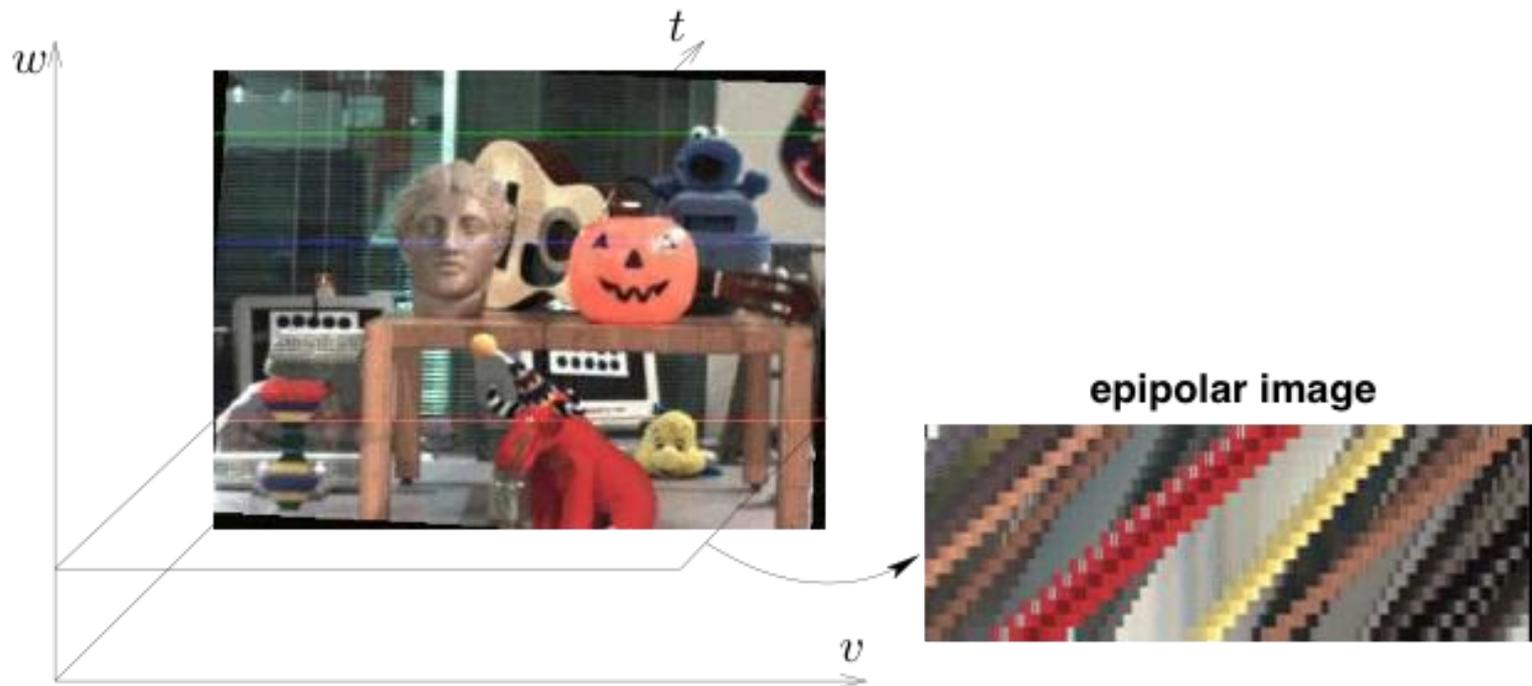


# Plenoptic Sampling

## Computer vision problem: Epipolar geometry

Points map to lines

Approximate sampling theorem



# When there are Problems....

## Rolex Learning Center at EPFL

*SANAA Architects* ( Kazuyo Sejima, Ryue Nishizawa )



# Are these Real Problems?

## Google maps as another popular example

How to register images

What resolution images to give an adequate view of the world



# Super-resolution

## Actual acquisition with a digital camera (Nikon D70)

- registration using FRI with psf and moments
- captured image: 60 images of 48 by 48 pixels
- super-resolved image 480 by 480 pixel



[Dragotti et al, 2008]

# Are these Real Problems?

## Sensor networks as another relevant example

How many sensors

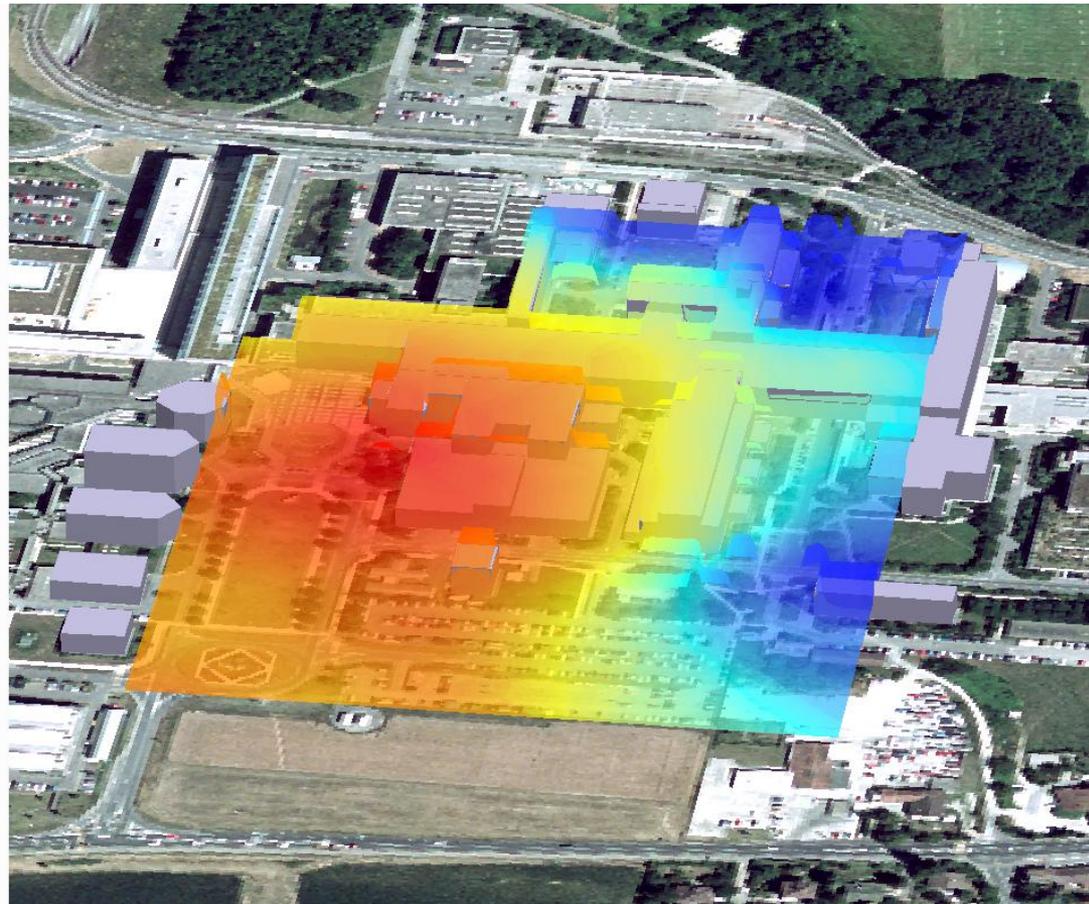
How to reconstruct

November 13th 2006  
Air Temperature Kriging  
5h00 pm local time

Air Temperature [°C]



batiments5



# Diffusion Equation and Inversion

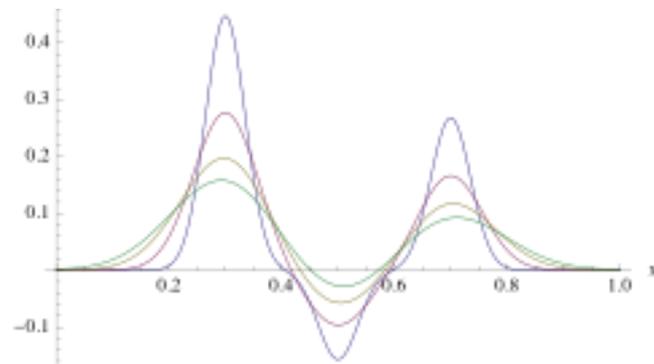
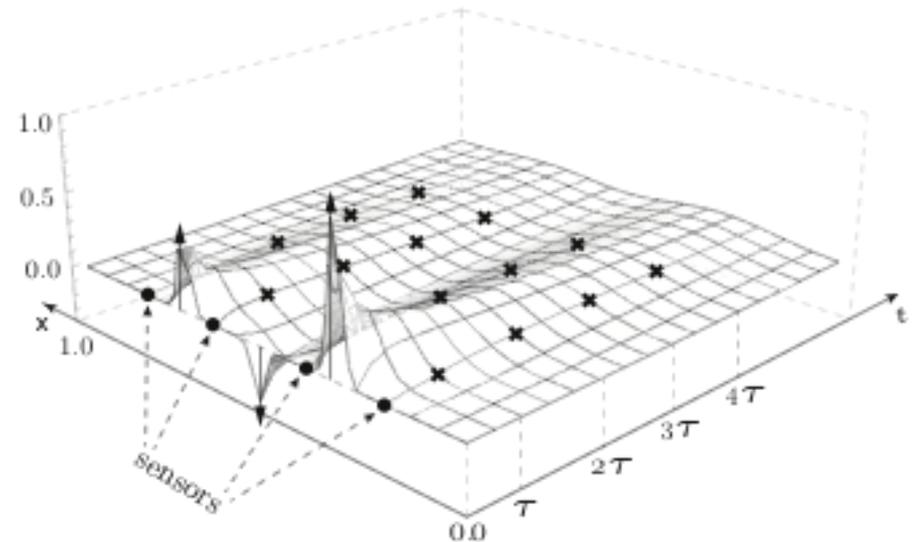
Point sources

Observation by sensors

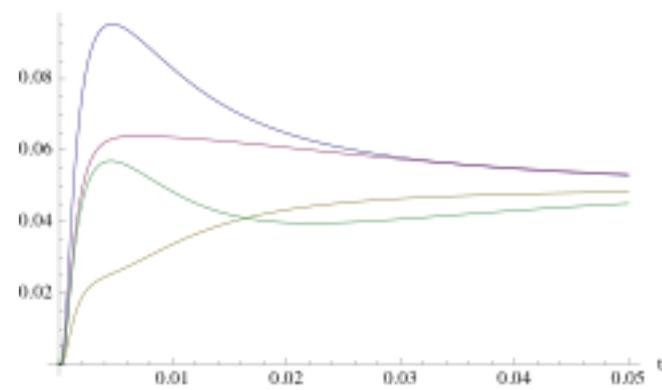
Ex: Process over  $(x,t)$

3 point sources

4 sensors



Over space

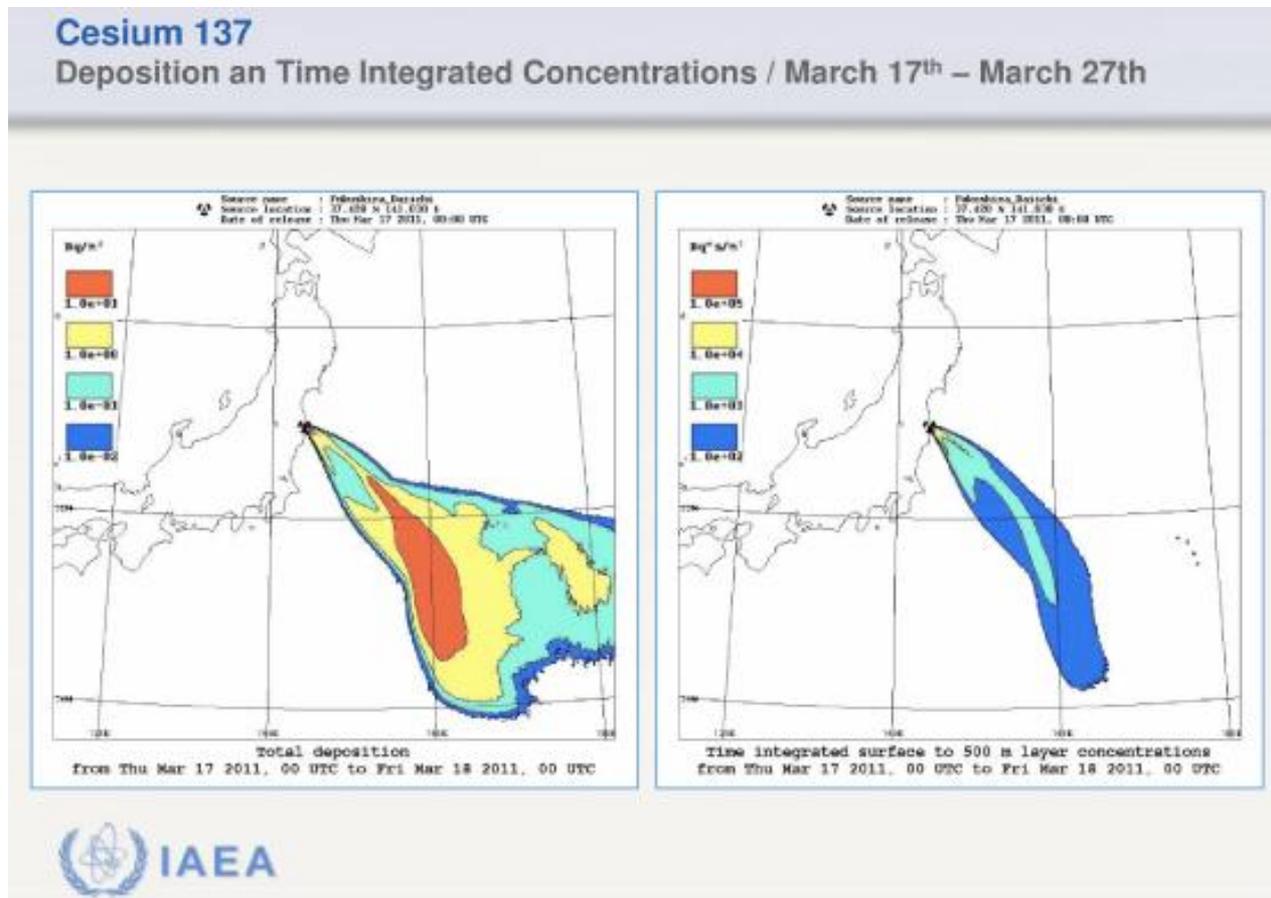


Over time

# Diffusion Equation: Real World Problem

Point source, know location, unknown time-varying intensity

Observation by multiple, possibly mobile sensors



# Outline

## 1. Introduction

## 2. Sampling: The linear case

Shannon sampling

Subspace sampling

Irregular sampling with known locations

## 3. Application: Sampling physics

## 4. Sampling based on timing

## 5. Sampling: The non-linear case

## 6. Applications: The non-linear case

## 7. Conclusions: Power game!

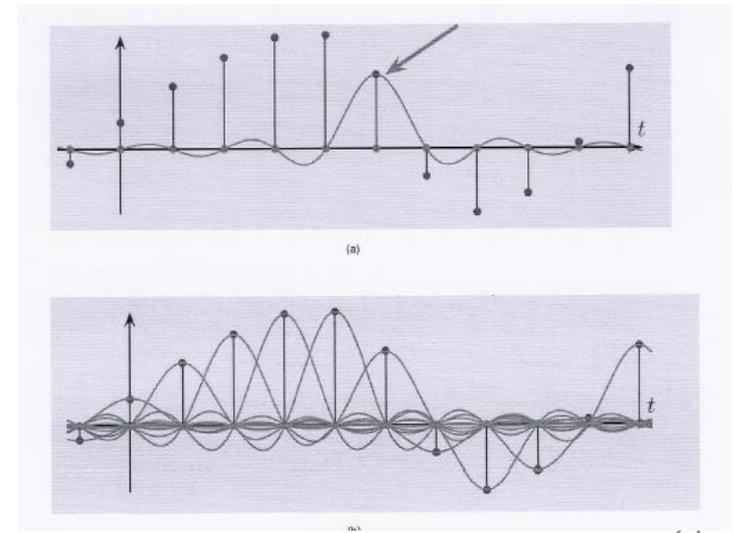
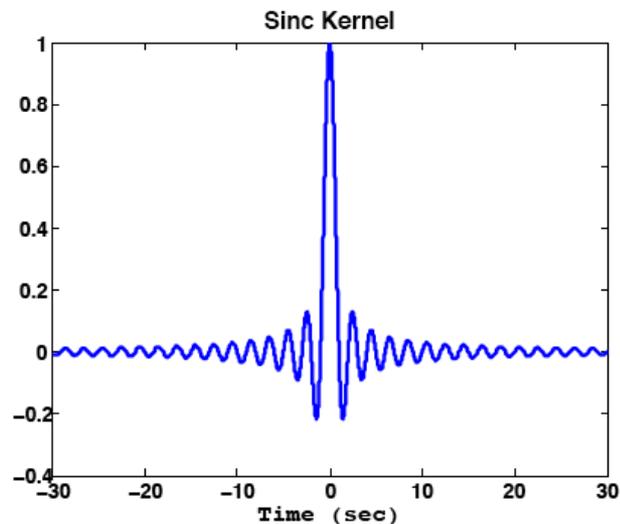
# Classic Sampling Case [WNKWRGSS, 1915-1949]

*If a function  $x(t)$  contains no frequencies higher than  $W$  cps, it is completely determined by giving its ordinates at a series of points spaced  $1/(2W)$  seconds apart. [if approx.  $T$  long,  $W$  wide,  $2TW$  numbers specify the function]*

**It is a representation theorem:**

- $\text{sinc}(t-n)$  is an orthobasis for  $\text{BL}[-\pi, \pi]$
- $x(t) \in \text{BL}[-\pi, \pi]$  can be written as

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \text{sinc}\left(\frac{t - nT_s}{T_s}\right)$$



# Classic Sampling Case (cont.)

## What if not bandlimited?

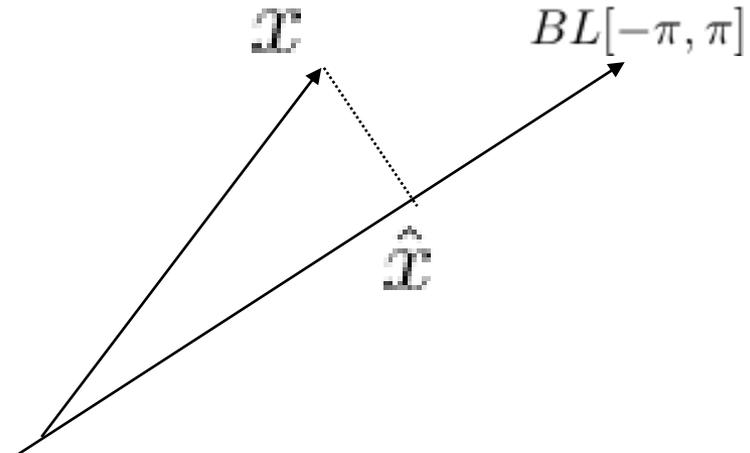
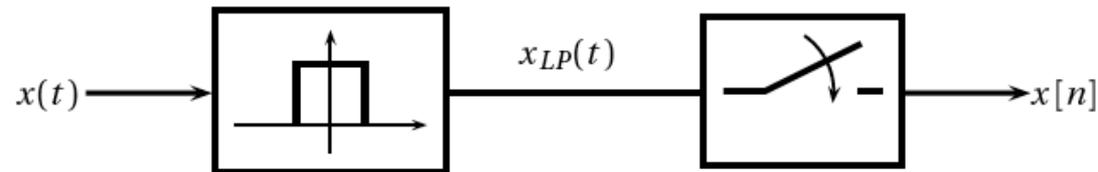
- Project onto  $BL[-\pi, \pi]$
- Use sampling theorem

## Space of bandlimited functions

- Linear subspace
- Shift-invariant

## Notes:

- Slow convergence of the sinc series
- Shannon-Bandwidth and information rate
- Bandlimited is sufficient, not necessary!  
(e.g. bandpass sampling)



# Shannon's Theorem... a Bit of History

**Whittaker 1915**

**Nyquist 1928**

**Kotelnikov 1933**

**Whittaker 1935**

**Raabe 1938**

**Gabor 1946**

**Shannon 1948**

**Someya 1949**

# Shannon's Theorem: Variations on the Subspace Theme

## Non uniform

- Kadec 1/4 theorem

## Derivative sampling

- Sample signal and derivative...  
... at half the rate

## Stochastic

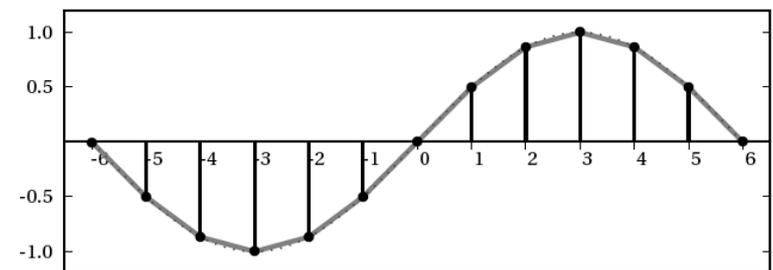
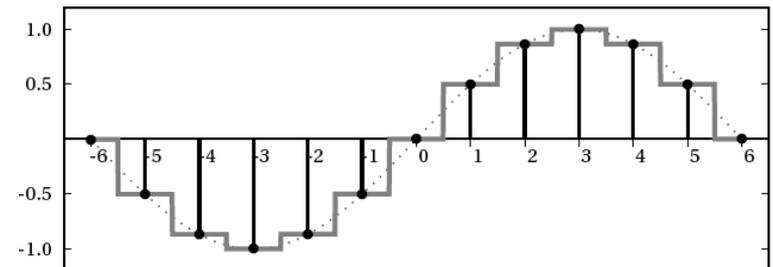
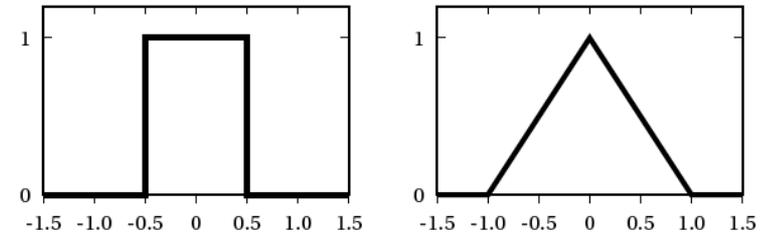
- Power spectrum is good enough

## Shift-invariant subspaces SISS

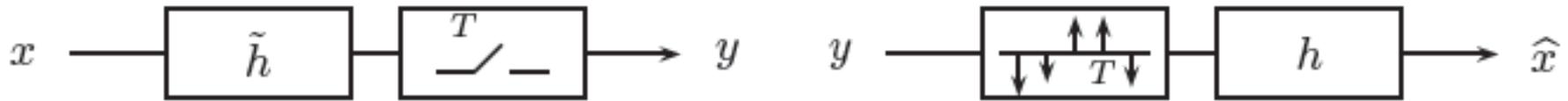
- Generalize from sinc to other shift-invariant Riesz bases (ortho. and biorthogonal)
- Structurally, it is the same thing!

## Multichannel sampling

- Known shift: easy
- Unknown shift: interesting (super-resolution imaging)



# SISS and the Biorthogonal Case



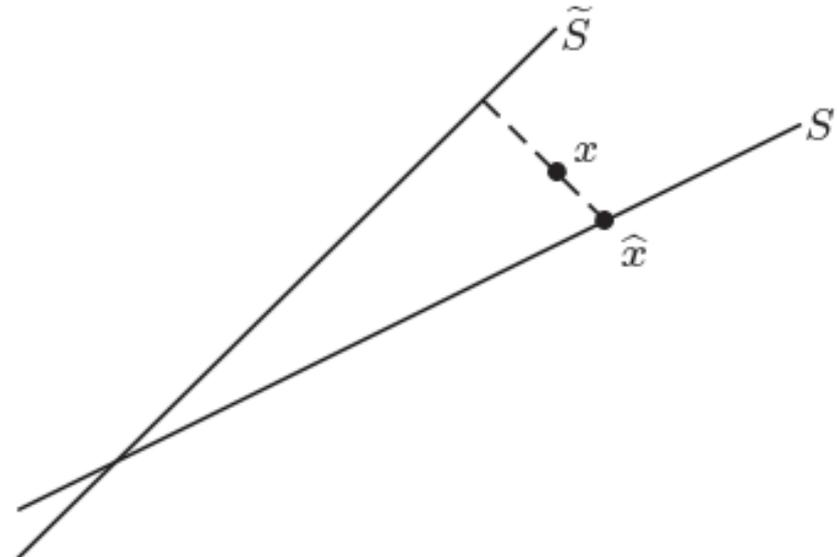
Consider shift-invariant subspace  $S$  with a generating function  $h$ ,  
e.g. B-spline and uniform shifts

Find a biorthogonal function

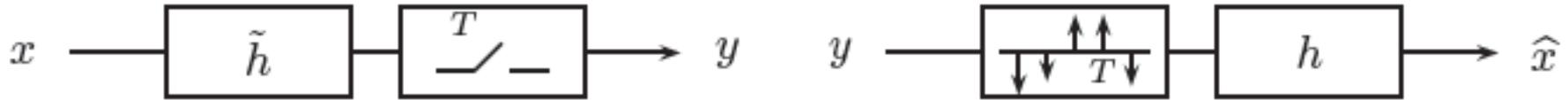
$$\langle \tilde{h}(t), h(t - n) \rangle = \delta_n$$

Oblique Projection  $P$

$$P = \Phi \tilde{\Phi}^*, \quad P^2 = P$$



# SISS and the Biorthogonal Case



Find a biorthogonal function in S

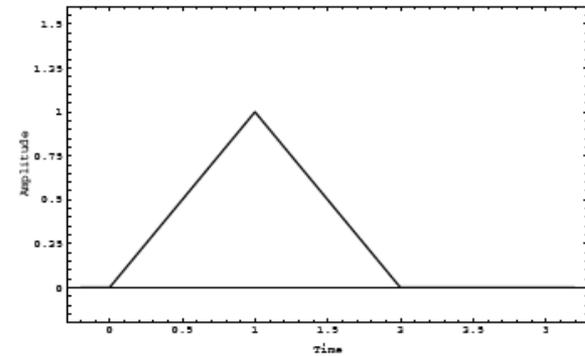
$$\check{\Phi} = \Phi(\Phi^* \Phi)^{-1}$$

Orthogonal projection

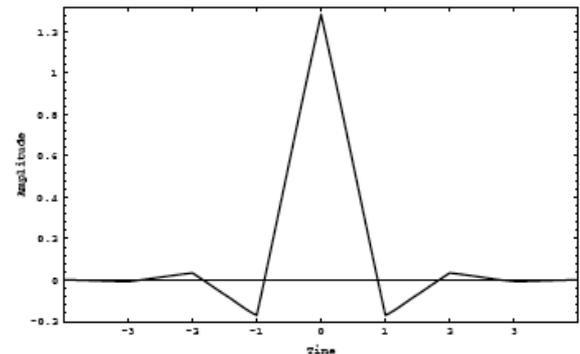
$$P^* = P$$

Many variations [Blu, Unser et al]

$\varphi(t)$



$\check{\varphi}(t)$



# The Frame View and Irregular Sampling

Consider bandlimited signals  $x(t)$ ,  $x(t) \leftrightarrow X(\omega)$ ,  $|X(\omega)| = 0 \quad |\omega| > \Omega$

If the sampling times are such that we have a frame operator  $A$ : invertible!

$$\alpha \|x\|^2 \leq \|A(x(t_n))\|^2 \leq \beta \|x\|^2$$

Thm [Grochenig et al 1992] Consider  $\delta$ -dense sequences,  $\sup(t_{i+i} - t_i) < \delta$  and piecewise constant midpoint approximation followed by LP filtering

Then

$$Ax = P \left( \sum_{i \in \mathbb{Z}} x(t_i) \chi_i \right)$$

$$\|x - Ax\| \leq \frac{\delta \Omega}{\pi} \|x\|$$

And thus, for  $\delta < \frac{\pi}{\Omega}$  one can reconstruct  $x$  from the irregular samples

Many variations [Feichtinger, Aldroubi et al]

Includes Kadec-1/4

# Outline

## 1. Introduction

## 2. Sampling: The linear case

## 3. Application: Sampling physics

Microphone and loudspeaker arrays

The plenacoustic function and a sampling theorem for wave fields

Mobile sensing

Sampling along trajectories

## 4. Sampling based on timing

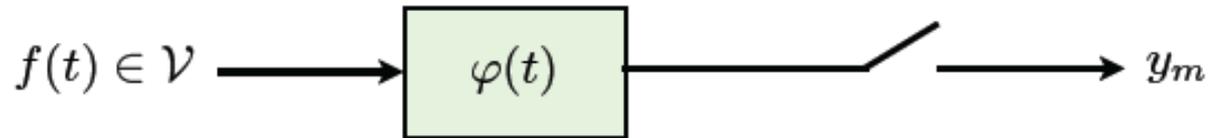
## 5. Sampling: The non-linear case

## 6. Applications: The non-linear case

## 7. Conclusion: The power game!

# Sampling Versus Sampling Physics

## Sampling:

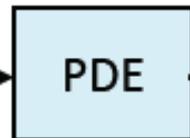


- Signal class  $\mathcal{V}$ : subspace, sparse, parametric (manifold)
- Observation kernel: by the setup, by design (e.g. random matrix in CS)

## Sampling physics:

sources:

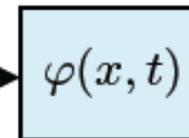
$$s(x, t) \in \mathcal{V}$$



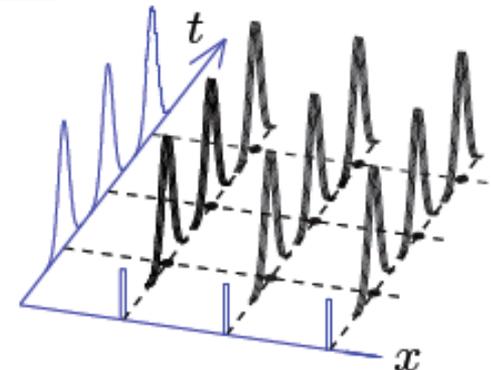
dynamic system

field:

$$f(x, t) \in \mathcal{W}$$



- Observation kernel: only temporal filtering
- **No spatial filtering** in  $\varphi(x, t)$



# Sampling and Interpolating Acoustic Fields

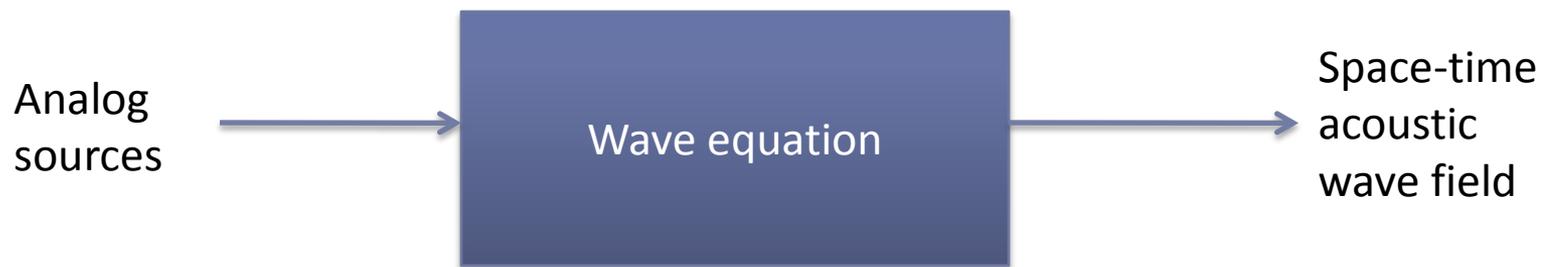
## Wave fields governed by the wave equation

- Space-time distribution
- Constrained by wave equation
- Not arbitrary, but smoothed



## What can we say about sampling/interpolation?

- Spatio-temporal Nyquist rate
- Perfect reconstruction
- Aliasing
- Space-time processing



How do we sample and interpolate?

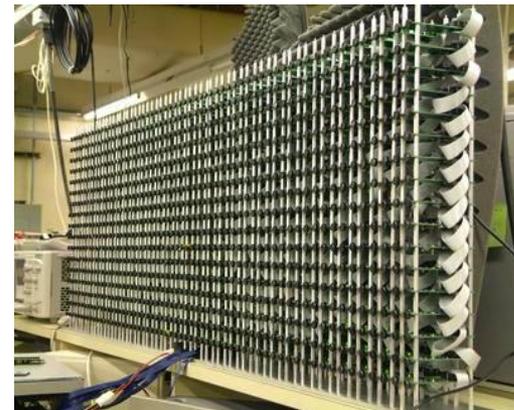
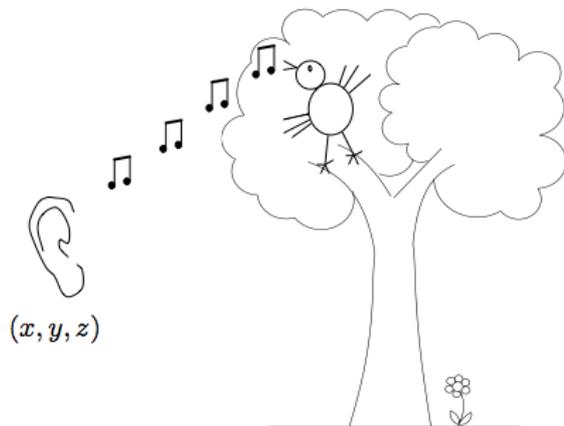
# Many Microphones and Loudspeakers

## Multiple microphones/loudspeakers

- physical world (e.g. free field, room)
- distributed signal acquisition of sound with “many” microphones
- sound rendering with many loudspeakers (wavefield synthesis)

## This is for real!

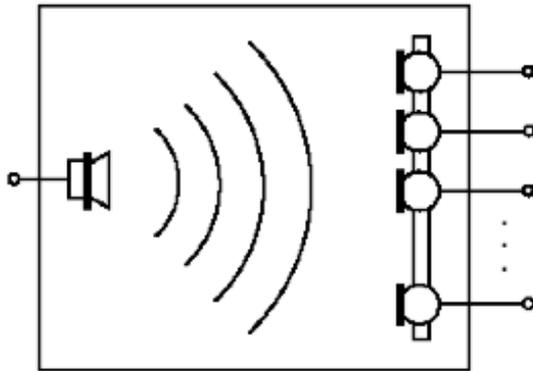
- sound recording
- special effects
- movie theaters (wavefield synthesis)
- MP3 surround etc



MIT1020 mics

# The Plenacoustic Function and its Sampling

## Setup

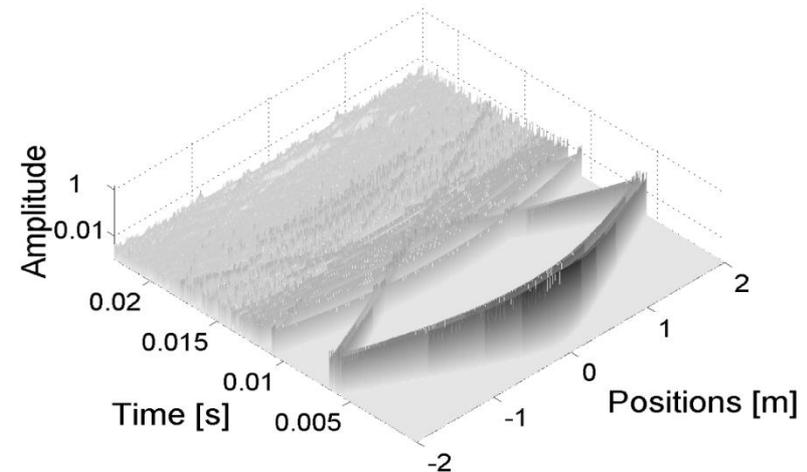
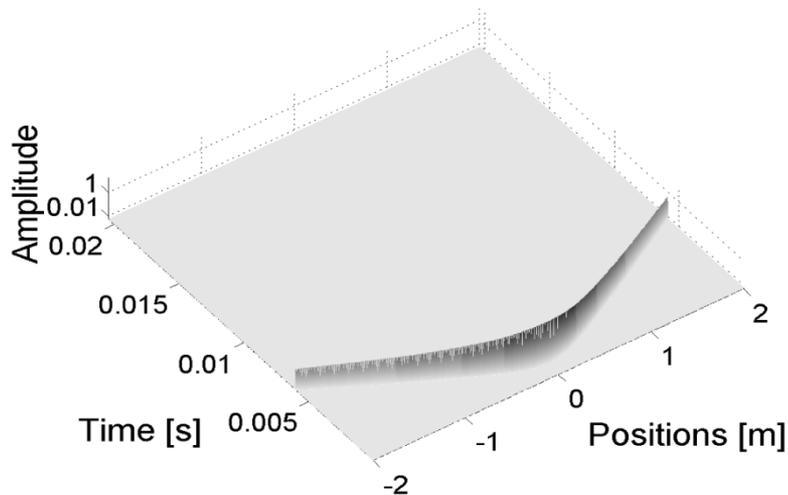


## Questions:

- Sample with “few” microphones and hear any location?
- Solve the wave equation? In general, it is much simpler to sample the plenacoustic function
- Dual question also of interest for synthesis (moving sources)
- Implication on acoustic localization problems
- Application for acoustic cancellation

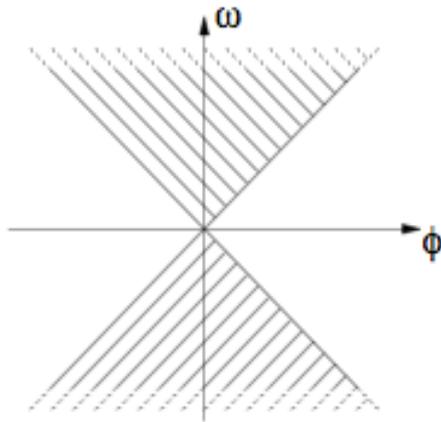
# The Plenacoustic Function and its Sampling

PAF in free field and in a room for a given point source



- We plot:  $p(x,t)$ , that is, the spatio-temporal impulse response
- The key question for sampling is:  $P(\phi, \omega)$  that is, the Fourier transform
- A precise characterization of  $P(\phi, \omega)$  for large  $\phi$  and  $\omega$  will allow sampling and reconstruction error analysis

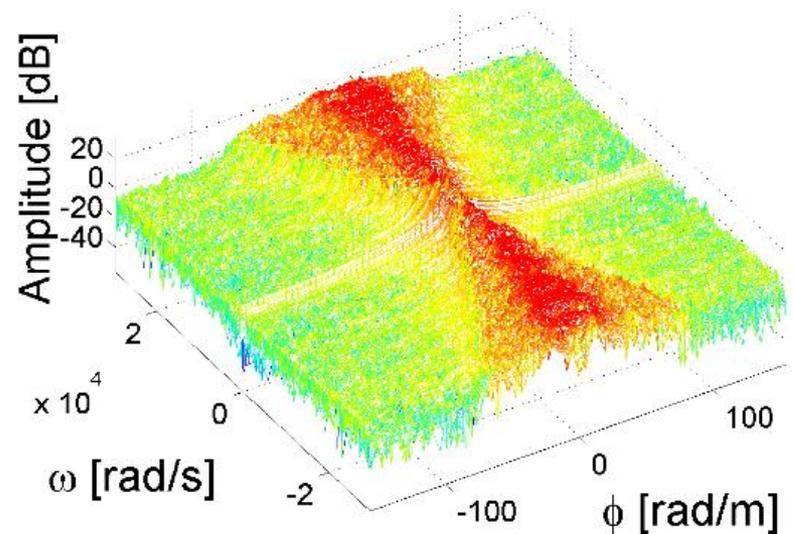
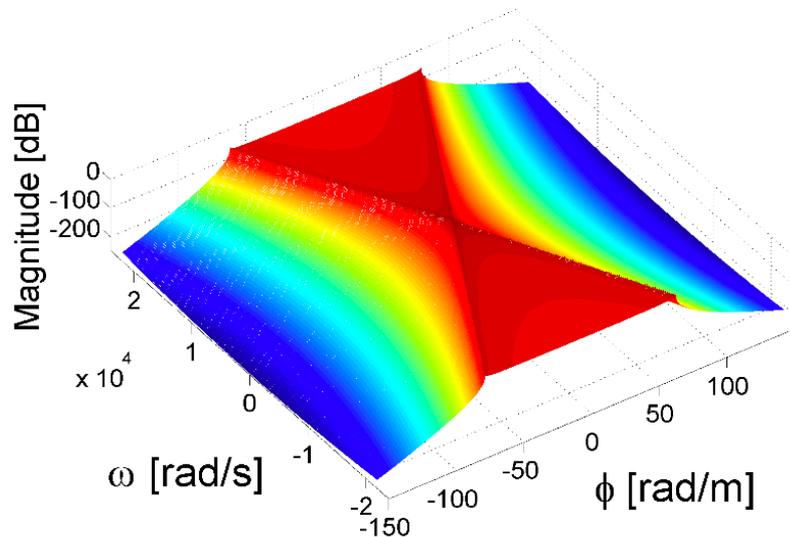
# The Plenacoustic Function in Fourier Space



slope:  $\omega = c \phi$

$\omega$ : temporal frequency

$\Phi$ : spatial frequency



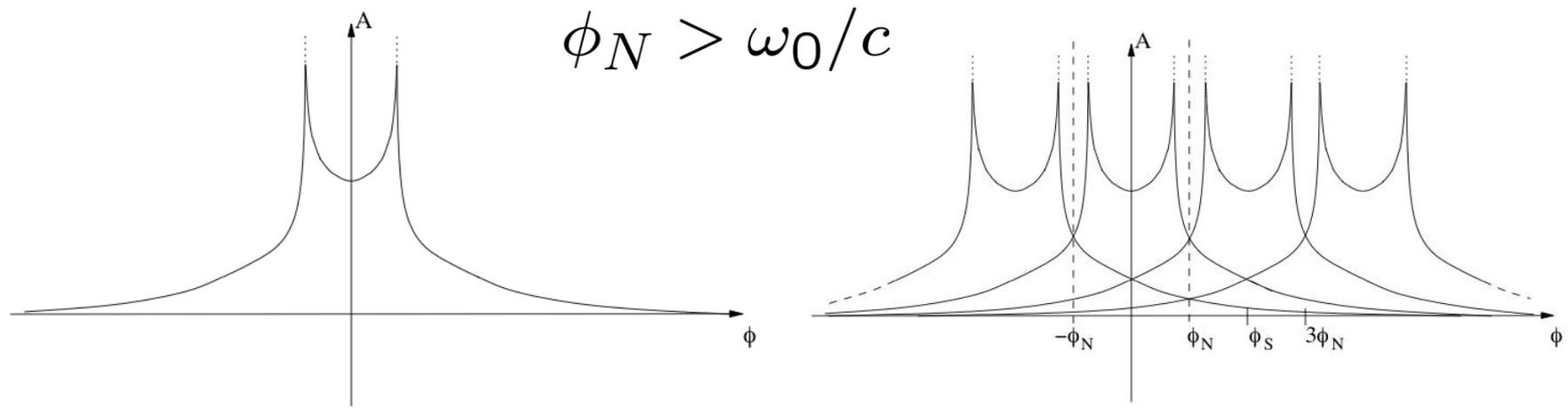
# A Sampling Theorem for Acoustic Fields

## Theorem [ASV:06]:

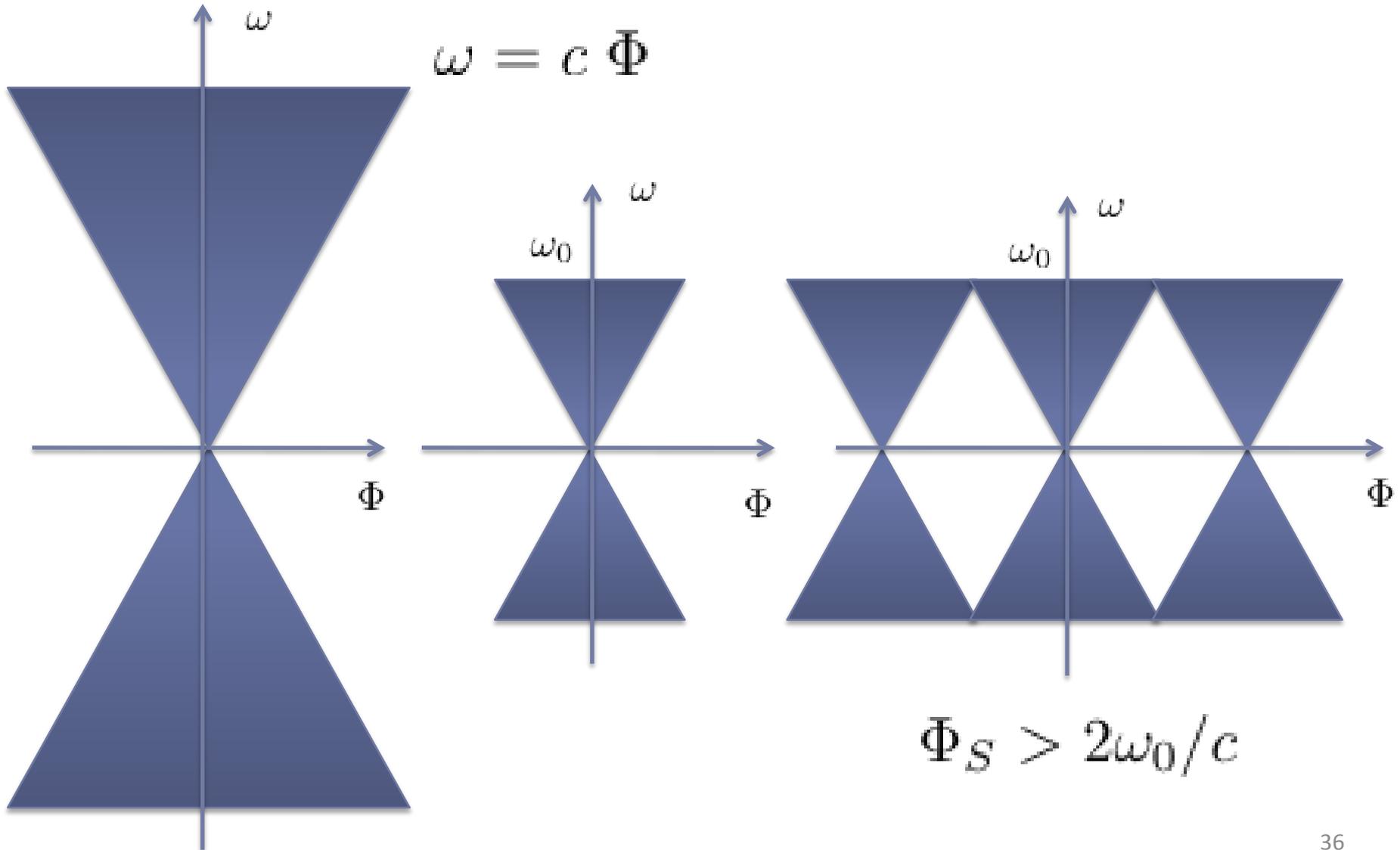
- Assume a max temporal frequency  $\omega_0$
- Pick a spatial sampling frequency  $\phi_N > \omega_0/c$
- Spatio-temporal signal interpolated from samples taken at  $(2\omega_0, 2\phi_N)$

## Argument:

- Take a cut through PAF
- Use exp. decay away from central triangle to bound aliasing
- Improvement using quincunx lattice



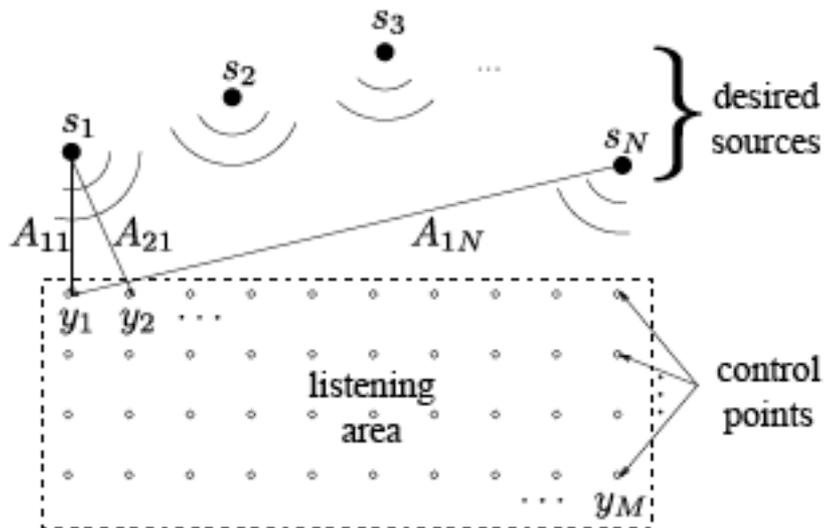
# Visual Proof ☺



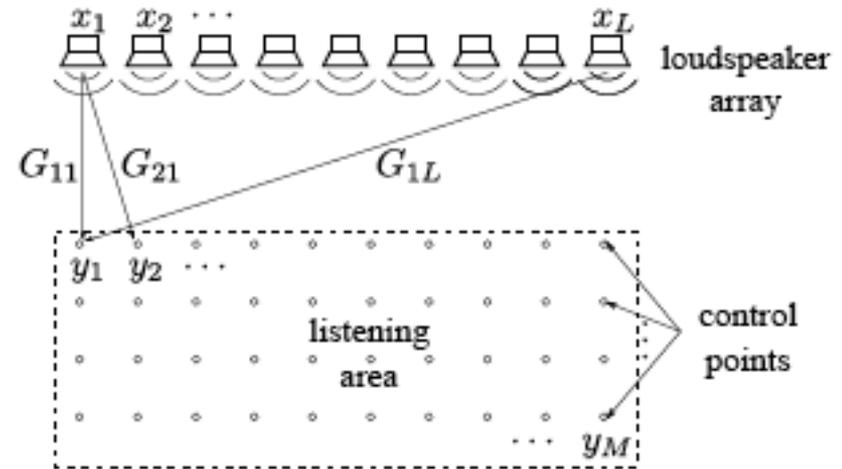
# Application: Sound Field Reproduction

Use the wave equation as an interpolation kernel!

- Control points at spatial Nyquist rate
- MIMO inversion
- The wave field is reproduced accurately in an entire region



Desired acoustic scene



Reproduced acoustic scene

# Mobile Sensing

Many signals exist in space and time but cannot be precisely monitored

## Pollution monitoring on cars/buses

- Sensors for CO<sub>2</sub>, NO<sub>2</sub>, fine particles and so on
- Known, moving locations
- Spatio-temporal distribution

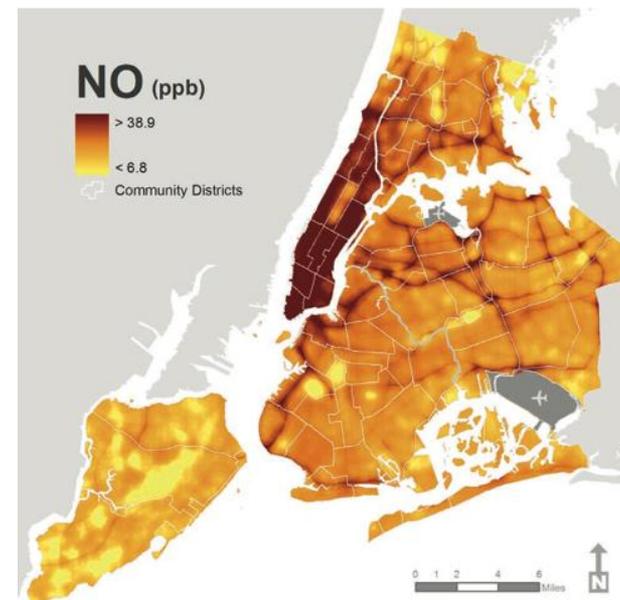
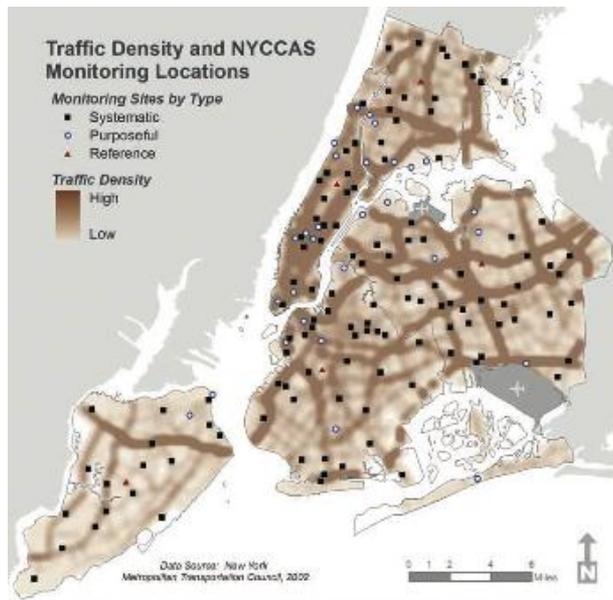
## Citizen sensing

- Collaborative sensing using mobile phones
- Known, moving locations
- Possible to gather data that is not necessarily published ...

# Pollution Monitoring Example

## NO measurement with 115 stations in NYC

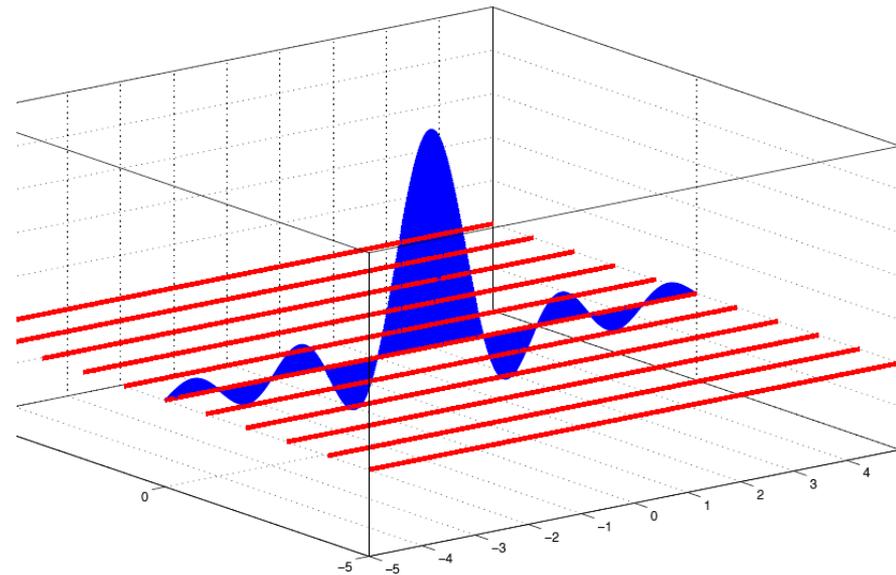
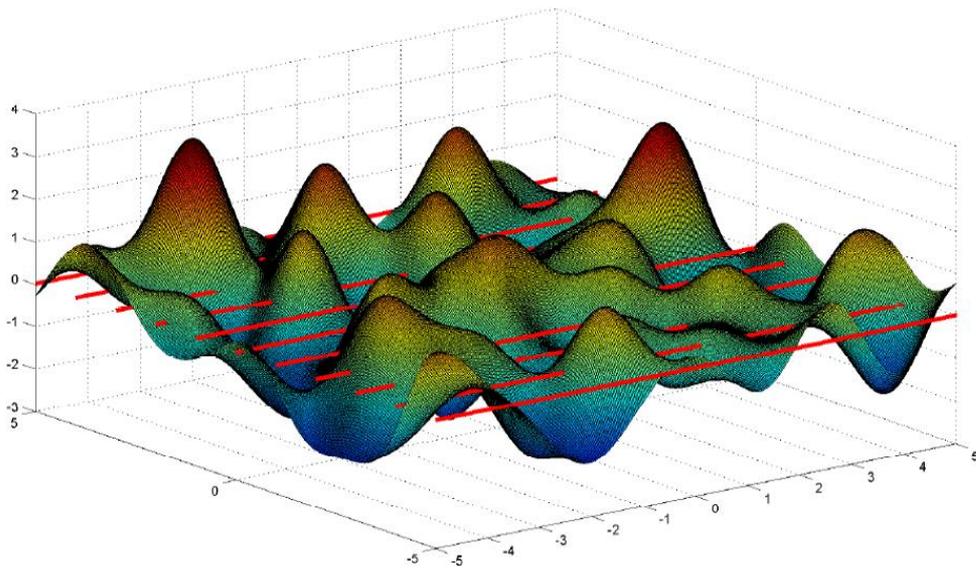
- Density of stations
- Interpolation



# Manhattan Sampling

Continuous along trajectories, discrete in between

- Original sampling question
- Interpolation/reconstruction
- Estimation



# Outline

## 1. Introduction

## 2. Sampling: The linear case

## 3. Application: Sampling physics

## 4. Sampling based on timing

Integrate and fire

Sampling theorem based on timing

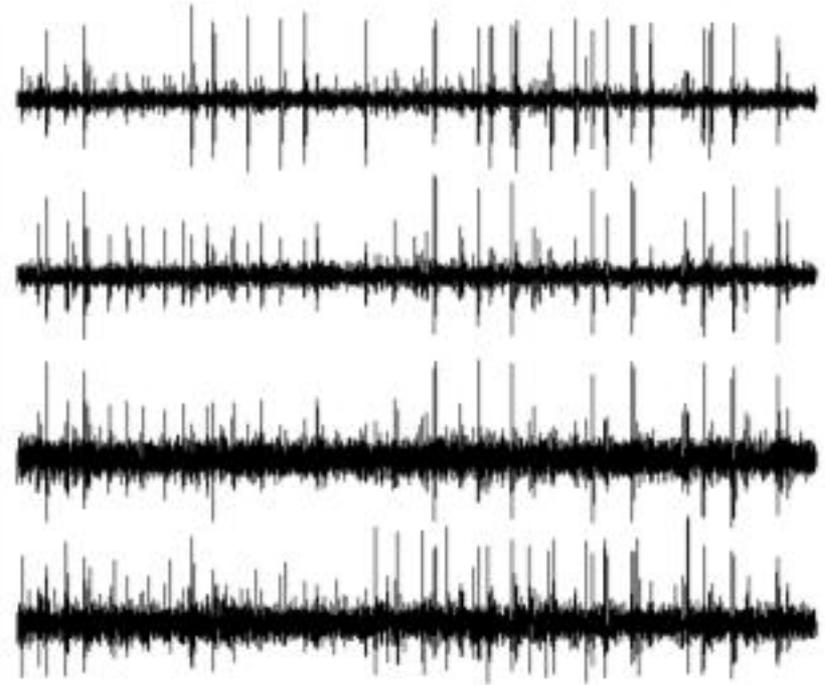
## 5. Sampling: The non-linear case

## 6. Applications: The non-linear case

## 7. Conclusions: Power game!

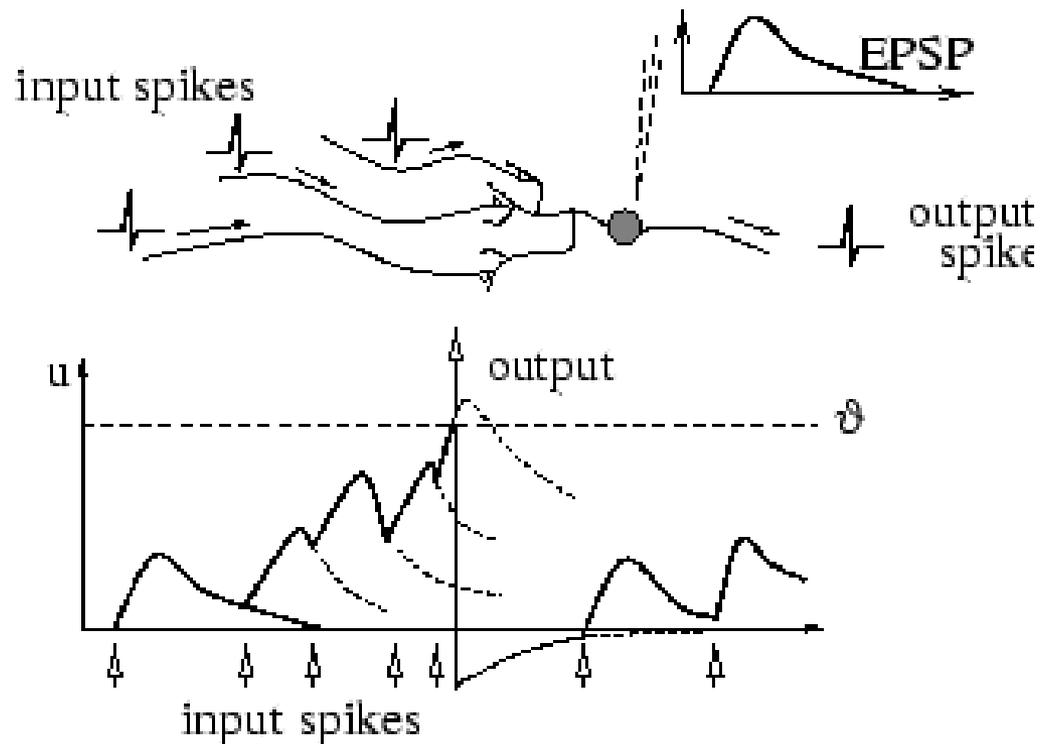
# Sampling Based on Timing

Don't try this at home!



# Sampling Based on Timing

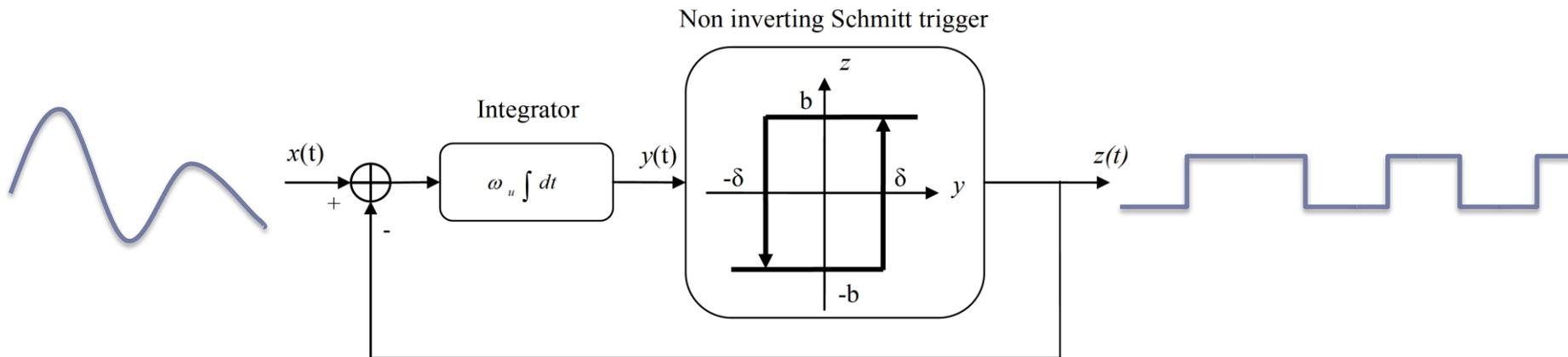
Neuron firings\*:



\*Gerstner and Kistler,  
*Spiking Neuron Models. Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002)

# Sampling Based on Timing

A possible model for neuron firing: Time encoding machine [Lazar et al]



It can be shown that for  $x$  in BL and bounded

- the set of timings are a sampling set in the sense of frames
- $x$  can be perfectly reconstructed

This can be generalized in several ways

# Outline

1. Introduction
2. Sampling: The linear case
3. Application: Sampling physics
4. Sampling based on timing
5. Sampling: The non-linear case
  - Position information
  - Finite rate of innovation
  - Sparse sampling
  - Compressed sensing
6. Applications: The non-linear case
7. Conclusions: Power game!

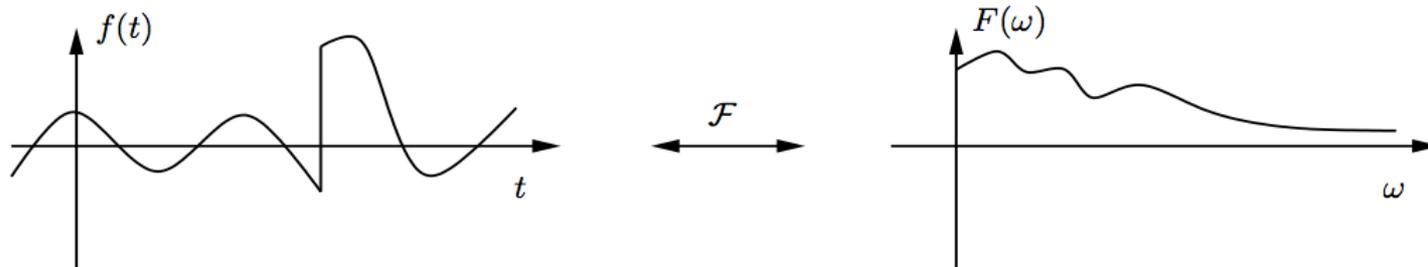
# Classic Case: Subspaces

Shannonbandlimitedcase

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \operatorname{sinc}\left(\frac{t - nT_s}{T_s}\right)$$

or  $1/T$  degrees of freedom per unit time

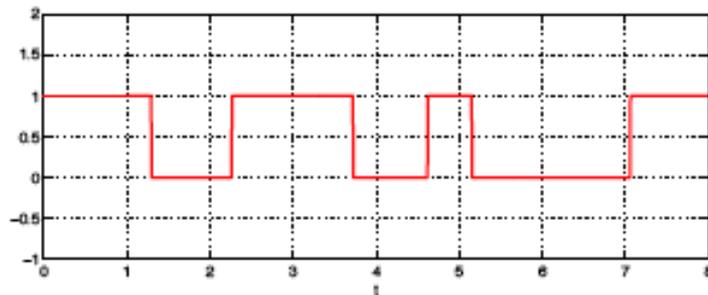
**But: a single discontinuity, and no more sampling theorem...**



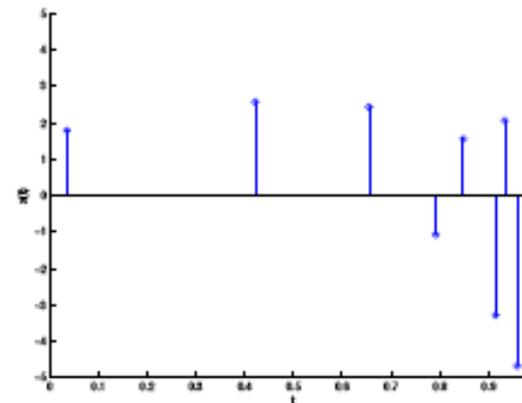
**Are there other signals with finite number of degrees of freedom per unit of time that allow exact sampling results?**

# Examples of Non-bandlimited Signals

Bilevel signals  
PPM, CDMA



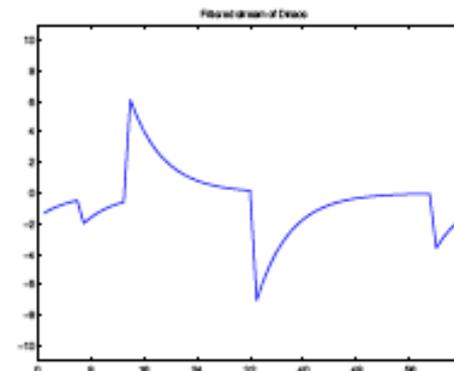
Stream of Diracs  
Poisson process



Piecewise Polynomial  
Woodcut picture



Filtered stream of Diracs  
Neural spikes  
UWB



# Classic Case and Beyond...

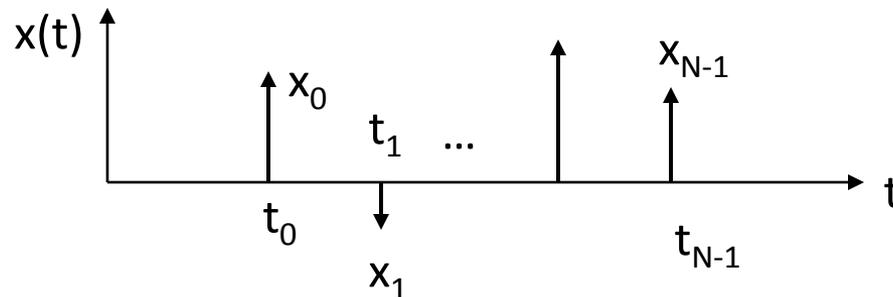
## Is there a sampling theory beyond Shannon?

- Shannon: bandlimitedness is sufficient but not necessary
- Shannon bandwidth: dimension of subspace
- Shift-invariant subspaces: Similar dimension of subspace

## Is there a sampling theory beyond subspaces?

- Finite rate of innovation: Similar to Shannon rate of information
- Non-linear set up

**Thus, develop a sampling theory for classes of non-bandlimited but sparse signals!**



Generic, continuous-time sparse signal

# Sparsity and Signals with Finite Rate of Innovation

## Sparsity:

- CT: parametric class, with degrees of freedom (e.g.  $K$  diracs, or  $2K$  degrees of freedom)
- DT:  $N$  dimensional vector  $x$ , and its  $l_0$  norm  $K = \|x\|_0$ . Then  $K/N$  indicates sparsity

## $\rho$ : Rate of innovation or degrees of freedom per unit of time

- Call  $C_T$  the number of degrees of freedom in the interval  $[-T/2, T/2]$ , then

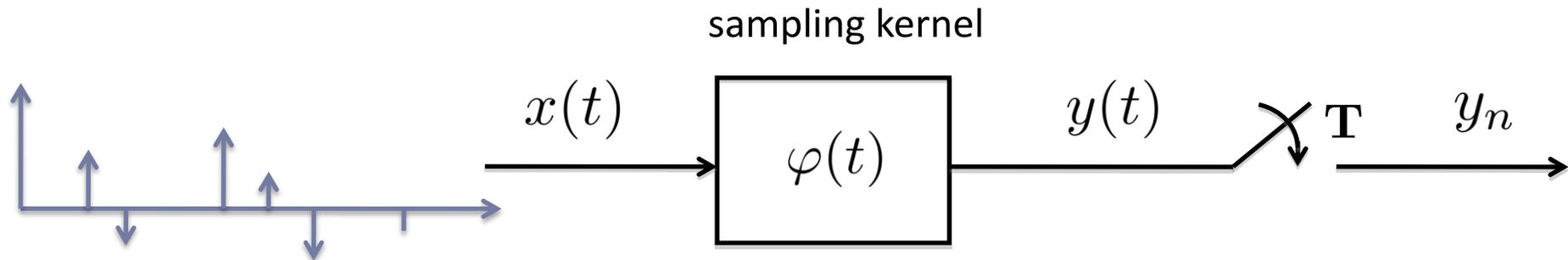
$$\rho = \lim_{T \rightarrow \infty} \frac{1}{T} C_T$$

## Note: Compressibility:

- Object expressed in an ortho-basis has fast decaying NLA  
For ex., decay of ordered wavelet coefficients in  $O(k^{-a})$ ,  $a > 1$ .

# Signals with Finite Rate of Innovation

The set up:



**For a sparse input, like a weighted sum of Diracs**

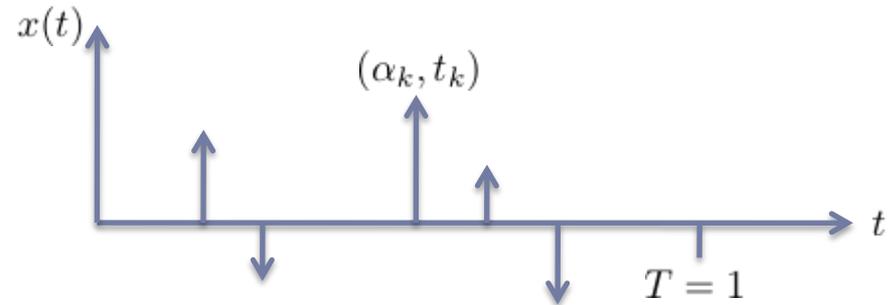
- One-to-one map  $y_n \Leftrightarrow x(t)$ ?
- Efficient algorithm?
- Stable reconstruction?
- Robustness to noise?
- Optimality of recovery?

# A simple exercise in Fourier series

## Periodic set of K Dirac pulses

- Is not bandlimited!
- Has a Fourier series  $X_m$

$$x(t) = \sum_{m=-\infty}^{\infty} X_m e^{-j2\pi mt}$$



## Fourier integral leads to

- K complex exponentials
- Exponents depends on location  $t_k$
- Weight depends on  $\alpha_k$

$$\begin{aligned} X_m &= \int_0^1 x(t) e^{-j2\pi mt} dt \\ &= \int_0^1 \left( \sum_{k=0}^{K-1} \alpha_k \delta(t - t_k) \right) e^{-j2\pi mt} dt \\ &= \sum_{k=0}^{K-1} \alpha_k e^{-j2\pi m t_k} \end{aligned}$$

## If we can identify the exponents

- Diracs can be recovered
- Weights are a linear problem (given the locations!)

# A Representation Theorem

## Theorem [VMB:02]

Given  $x(t)$ , a periodic set of  $K$  Diracs, of period  $\tau$ , weights  $\{x_k\}$  and locations  $\{t_k\}$ .

$$x(t) = \sum_{k=1}^K \sum_{k' \in \mathbb{Z}} x_k \delta(t - t_k - k'\tau)$$

Take a Dirichlet sampling kernel of bandwidth  $B$ , with  $B\tau$  an odd integer  $> 2K$

$$\varphi(t) = \sum_{k' \in \mathbb{Z}} \text{sinc}(B(t - k'\tau)) = \frac{\sin(\pi Bt)}{B\tau \sin(\pi t/\tau)}$$

Then the  $N$  samples,  $N > B\tau$ ,  $T = \tau/N$ ,

$$y_n = \sum_{k=1}^K x_k \varphi(nT - t_k)$$

are a sufficient characterization of  $x(t)$ .

# Linear versus Non-linear Problem

This is **not** a subspace problem!

Problem is **non-linear** in  $t_k$ ,  
and **linear** in  $x_k$  given  $t_k$

Given two such streams of  $K$  Diracs,  
and weights and locations  $\{x_k, t_k\}$  and  $\{x'_k, t'_k\}$ .

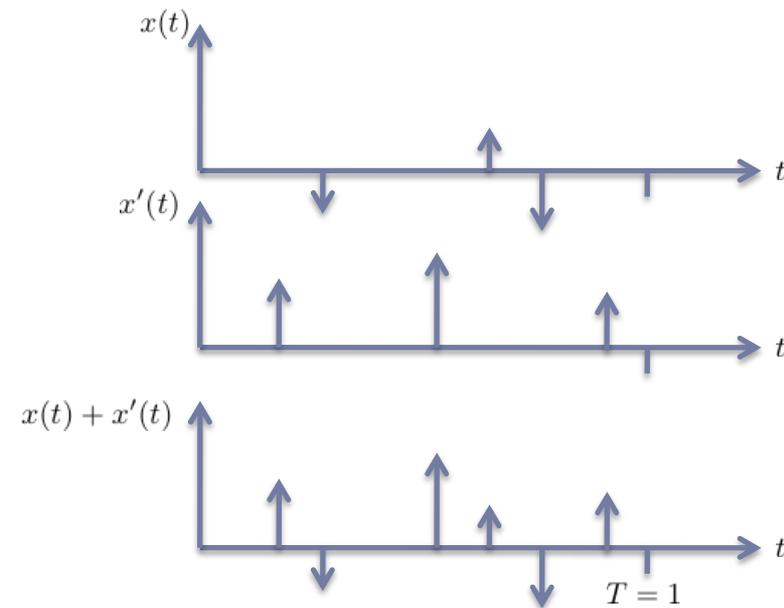
The sum is a stream with  $2K$  Diracs.

But, given a set of locations  $\{t_k\}$   
then the **problem is linear** in  $\{x_k\}$ !

**The key to the solution:**

Separability of non-linear from linear problem

**Note:** Finding locations is a key task in estimation/retrieval of sparse signals,  
but also in spectral estimation, error location in coding, in registration,  
feature detection etc



# Sketch of Proof

The signal is periodic, so consider its Fourier series

$$x(t) = \sum_{m \in \mathbb{Z}} \hat{x}_m e^{j2\pi mt/\tau}, \quad \text{where} \quad \hat{x}_m = \frac{1}{\tau} \sum_{k=1}^K x_k \underbrace{e^{-j2\pi mt_k/\tau}}_{u_k^m}.$$

1. The samples  $y_n$  are a sufficient characterization of the central  $2K+1$  Fourier series coefficients (Sampling Thm. for BL FS).
2. The Fourier series is a linear combination of  $K$  complex exponentials. These can be killed using an annihilation filter

$$H(z) = \sum_{k=0}^K h_k z^{-k} = \prod_{k=1}^K (1 - u_k z^{-1}).$$

$$h_m * \hat{x}_m = \sum_{k=0}^K h_k \hat{x}_{m-k} = 0$$

## Sketch of Proof (cont.)

3. To find the coefficients of the annihilating filter, we need to solve a convolution equation, which leads to a  $K$  by  $K$  Toeplitz system

$$\begin{bmatrix} \hat{x}_{-1} & \hat{x}_{-2} & \cdots & \hat{x}_{-K} \\ \hat{x}_0 & \hat{x}_{-1} & \cdots & \hat{x}_{-K+1} \\ \vdots & \vdots & \cdots & \vdots \\ \hat{x}_{K-2} & \hat{x}_{K-3} & \cdots & \hat{x}_{-1} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_K \end{bmatrix} = - \begin{bmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_{K-1} \end{bmatrix}.$$

4. Given the coefficients  $\{1, h_1, h_2, \dots, h_K\}$ , we get the  $\{t_k\}$ 's by factorization of

$$H(z) = \prod_{k=1}^K (1 - u_k z^{-1}). \quad u_k = e^{-j2\pi t_k / \tau}$$

5. To find the coefficients  $\{x_k\}$ , we have a linear problem, since given the  $\{t_k\}$ 's or  $\{u_k\}$ 's, the Fourier series is given by

$$\hat{x}_m = \frac{1}{\tau} \sum_{k=1}^K x_k e^{-j2\pi m t_k / \tau} = \frac{1}{\tau} \sum_{k=1}^K x_k u_k^m.$$

This is a Vandermonde linear system, proving  $2K+1$  samples are sufficient!  $\square$

# Notes on Proof

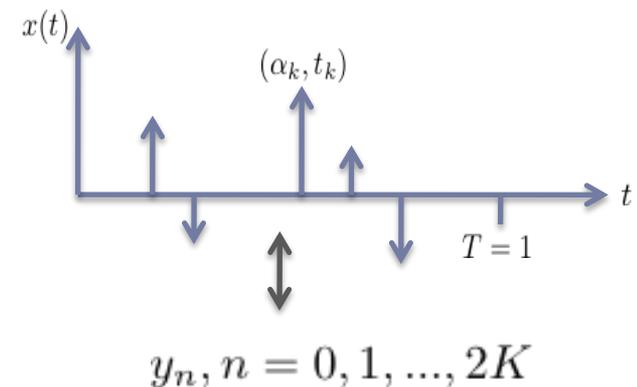
**The procedure is constructive, and leads to an algorithm:**

1. Take  $2K+1$  samples  $y_n$  from Dirichlet kernel output
2. Compute the DFT to obtain Fourier series coefficients  $-K..K$
3. Solve Toeplitz system of size  $K$  by  $K$  to get  $H(z)$
4. Find roots of  $H(z)$  by factorization, to get  $u_k$  and  $t_k$
5. Solve Vandermonde system of size  $K$  by  $K$  to get  $x_k$ .

**The complexity is:**

1. Analog to digital converter
2.  $K \log K$
3.  $K^2$
4.  $K^3$  (can be accelerated)
5.  $K^2$

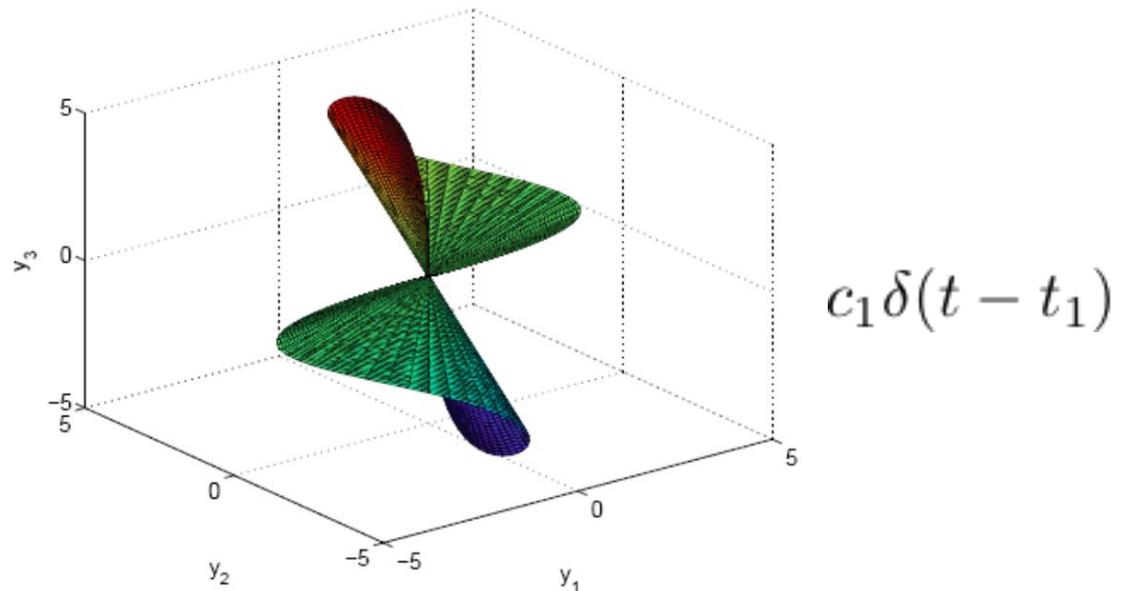
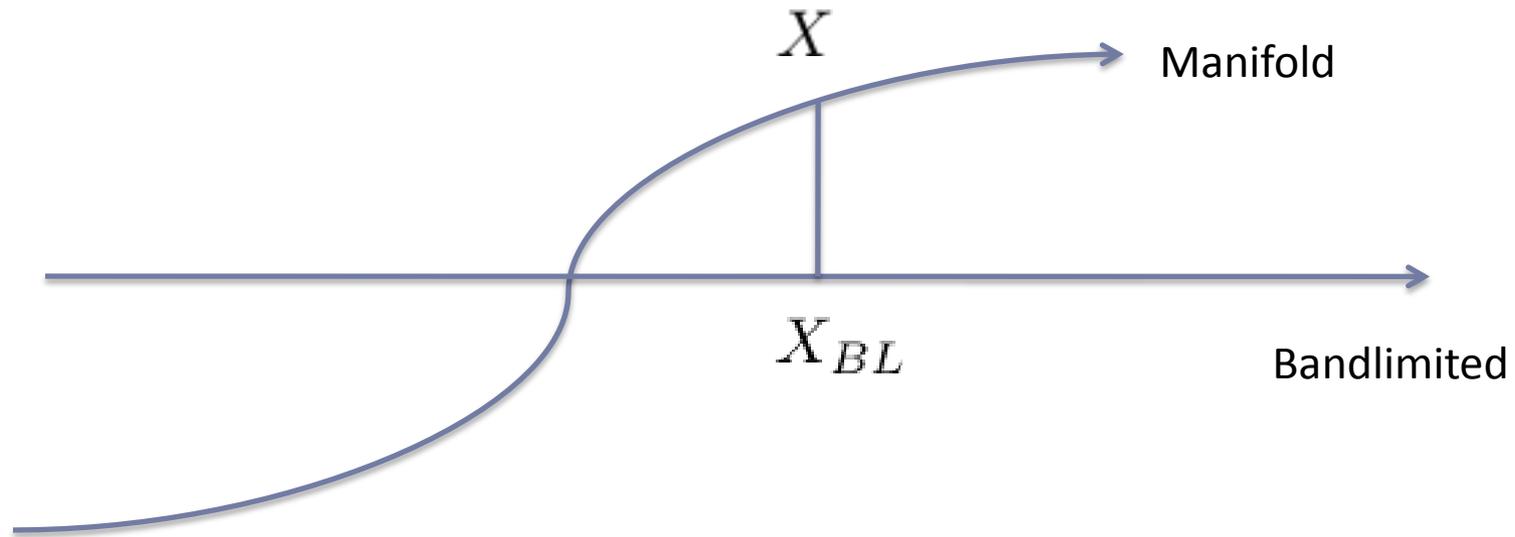
Or polynomial in  $K$ !



**Note 1:** For size  $N$  vector, with  $K$  Diracs,  $O(K^3)$  complexity, noiseless

**Note 2:** Method similar to sinusoidal retrieval in spectral estimation and

# What's Maybe Surprising....



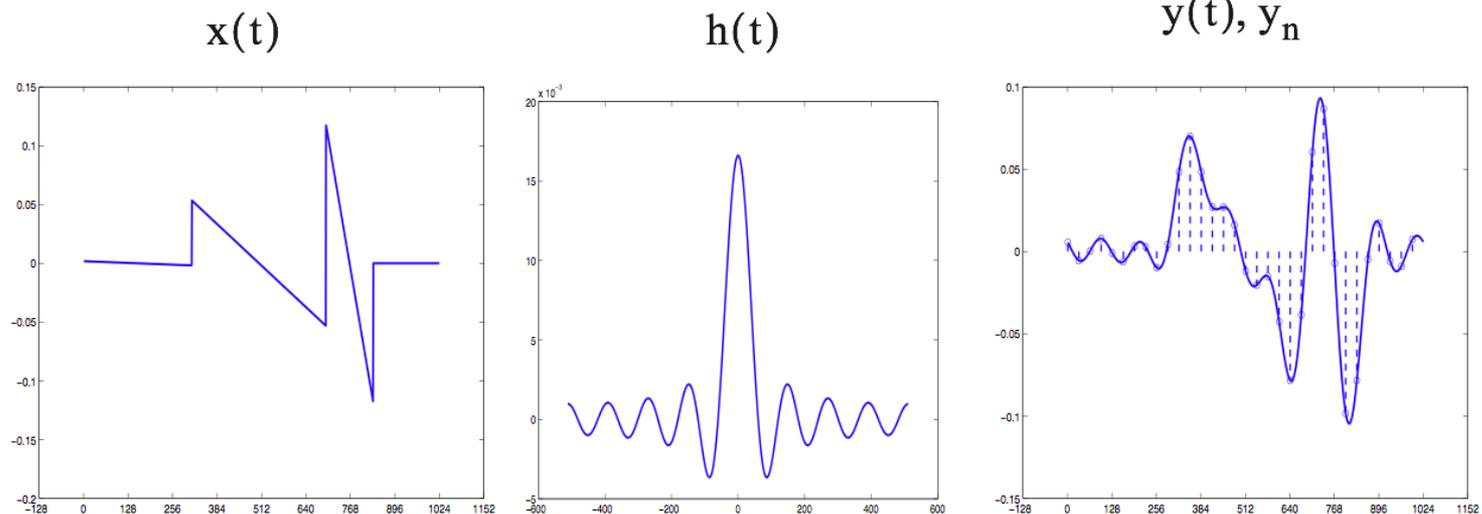
# Generalizations

For the class of periodic FRI signals which includes

- Sequences of Diracs
- Non-uniform or free knot splines
- Piecewise polynomials

There are sampling schemes with sampling at the rate of innovation with perfect recovery and polynomial complexity

Variations: finite length, 2D, local kernels etc

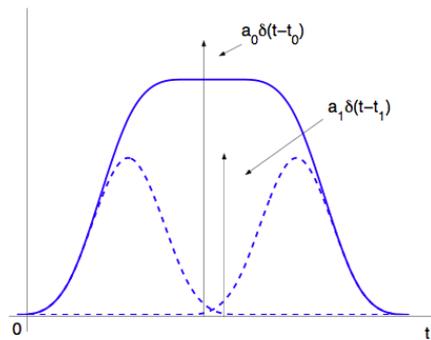


# Generalizations

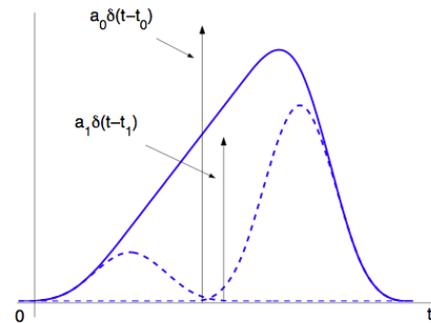
## Strang-Fix condition on sampling kernel [DVB:07]:

Local, polynomial complexity reconstruction, for diracs and piecewise polynomials

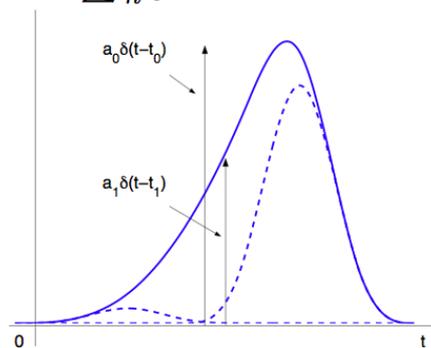
Pure discrete-time processing!



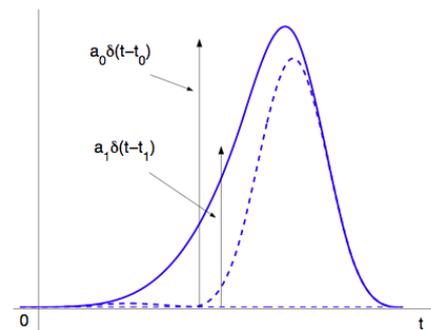
$$\sum_n y_n = a_0 + a_1$$



$$\sum_n n y_n = a_0 t_0 + a_1 t_1$$



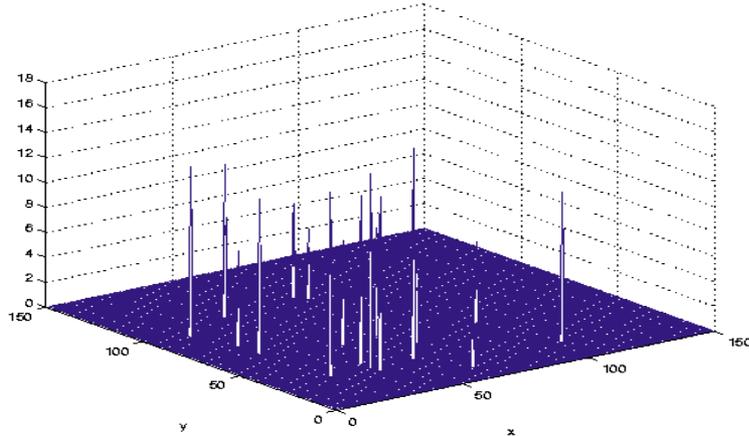
$$\sum_n n^2 y_n = a_0 t_0^2 + a_1 t_1^2$$



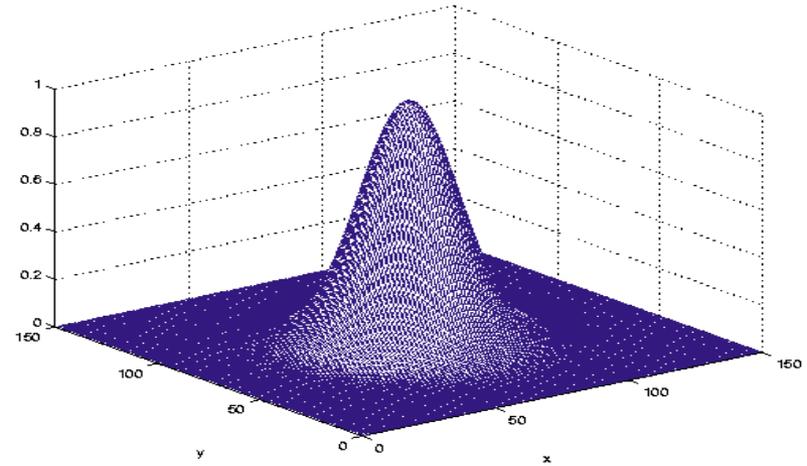
$$\sum_n n^3 y_n = a_0 t_0^3 + a_1 t_1^3$$

# Generalizations: 2D Extensions

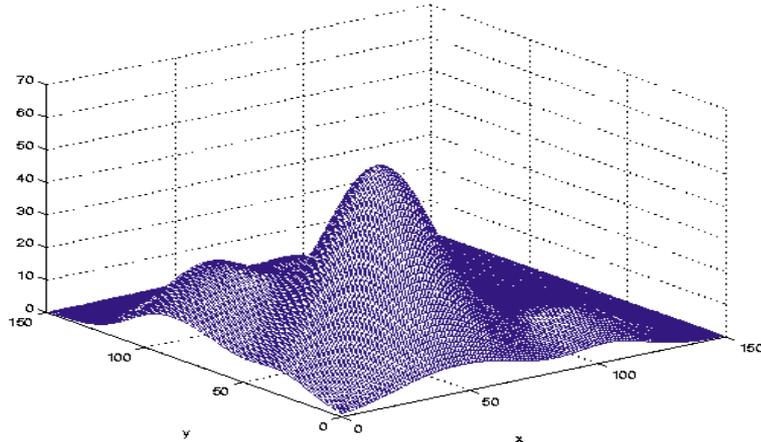
Finite set of  $M=17$  weighted Diracs



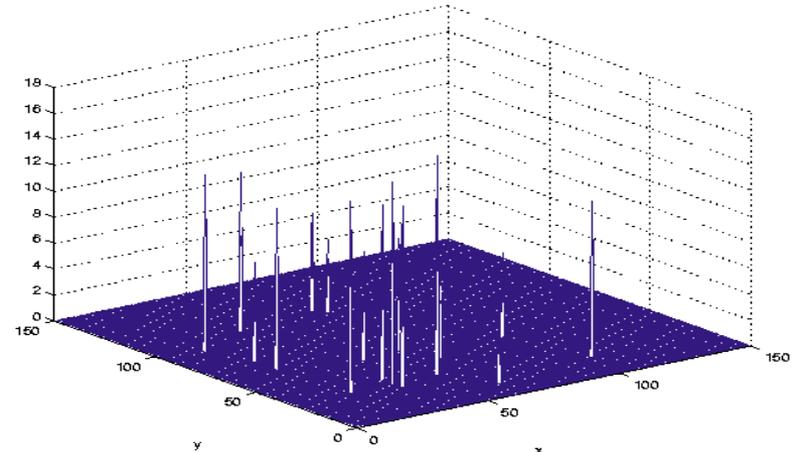
Gaussian sampling kernel



Convolution with the sampling kernel



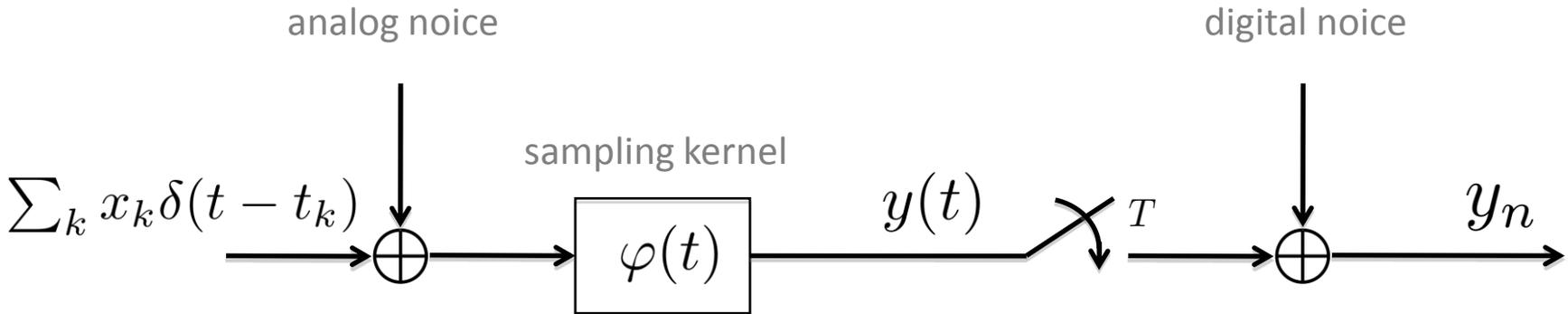
Reconstructed signal



Note: true 2D processing!

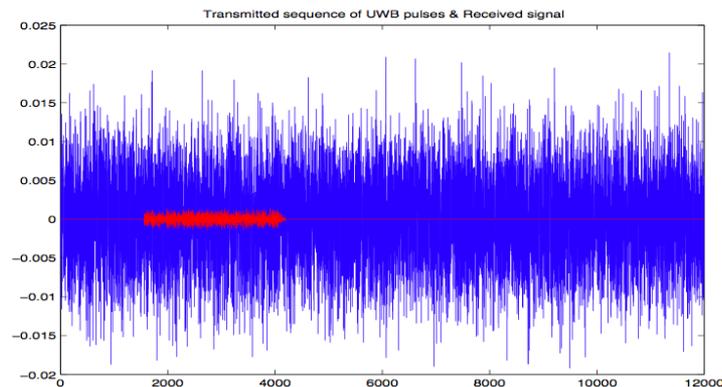
# The Real, Noisy Case...

## Acquisition in the noisy case:



where “analog” noise is before acquisition (e.g. communication noise on a channel) and digital noise is due to acquisition (ADC, etc)

**Example:** Ultrawide band (UWB) communication....



# The Real, Noisy Case...

## Total Least Squares:

Annihilation equation:  $AH = 0$  can only be approximately satisfied.

Instead:

Minimize  $\|AH\|^2$  under constraint  $\|H\|^2 = 1$

using SVD of a rectangular system

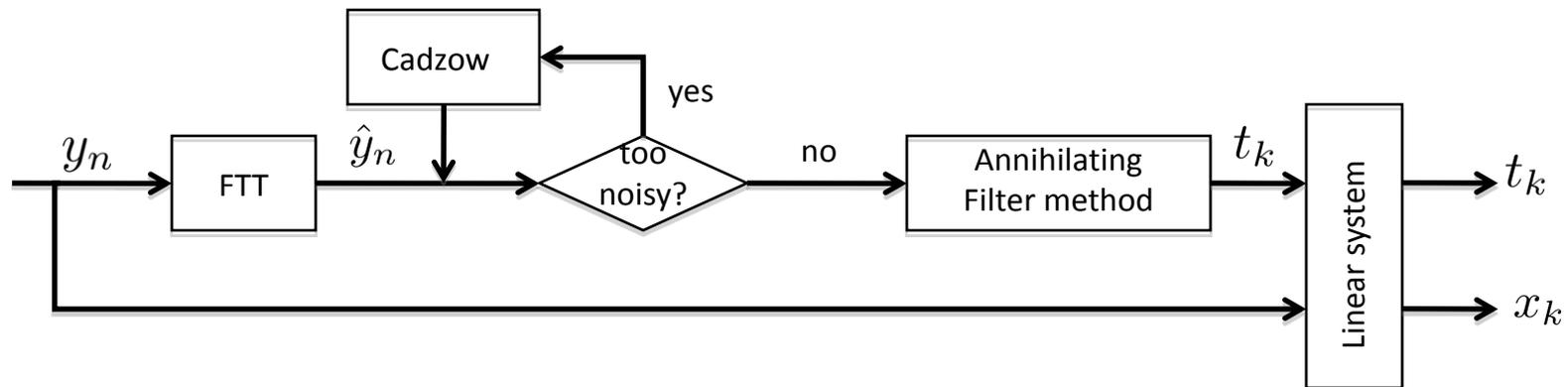
## Cadzowdenoising:

When very noisy: TLS can fail...

Use longer filter  $L > K$ , but use fact that noiseless matrix  $A$  is Toeplitz of rank  $K$

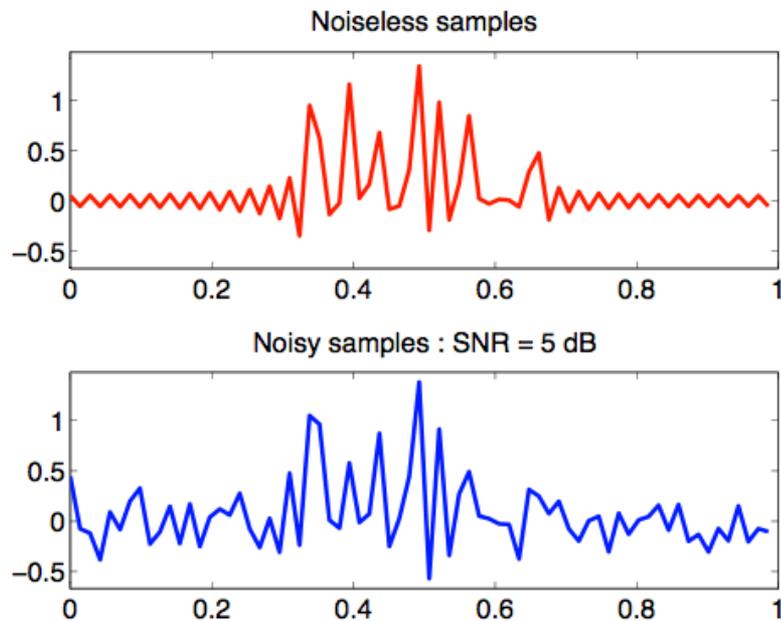
Iterate between the two conditions

## Algorithm:

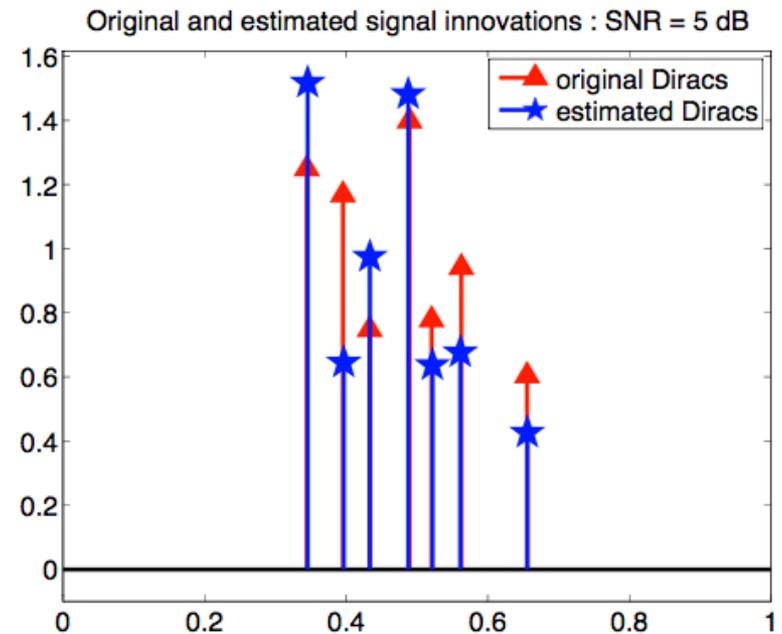


# Example

7 Diracs in 5dB SNR, 71 samples



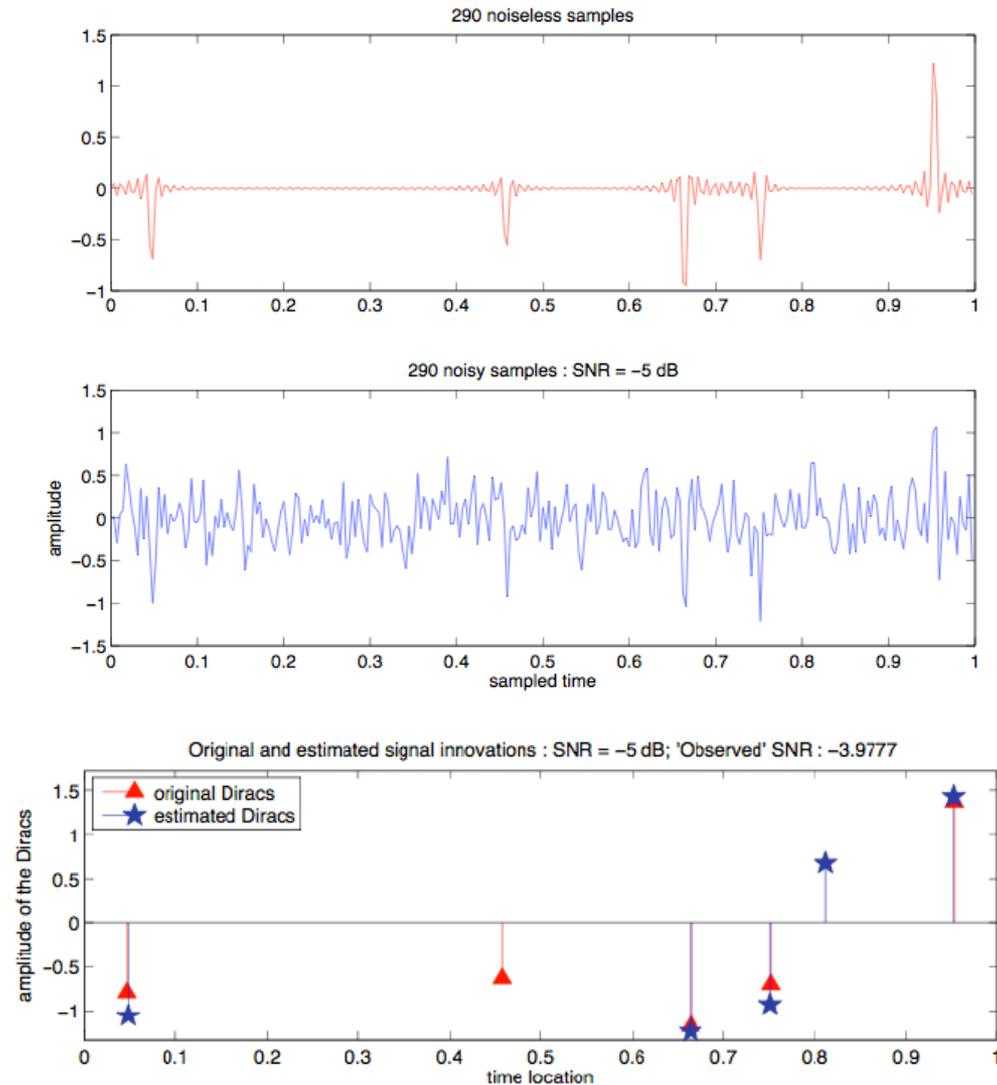
Original and noisy version



Original and retrieved Diracs

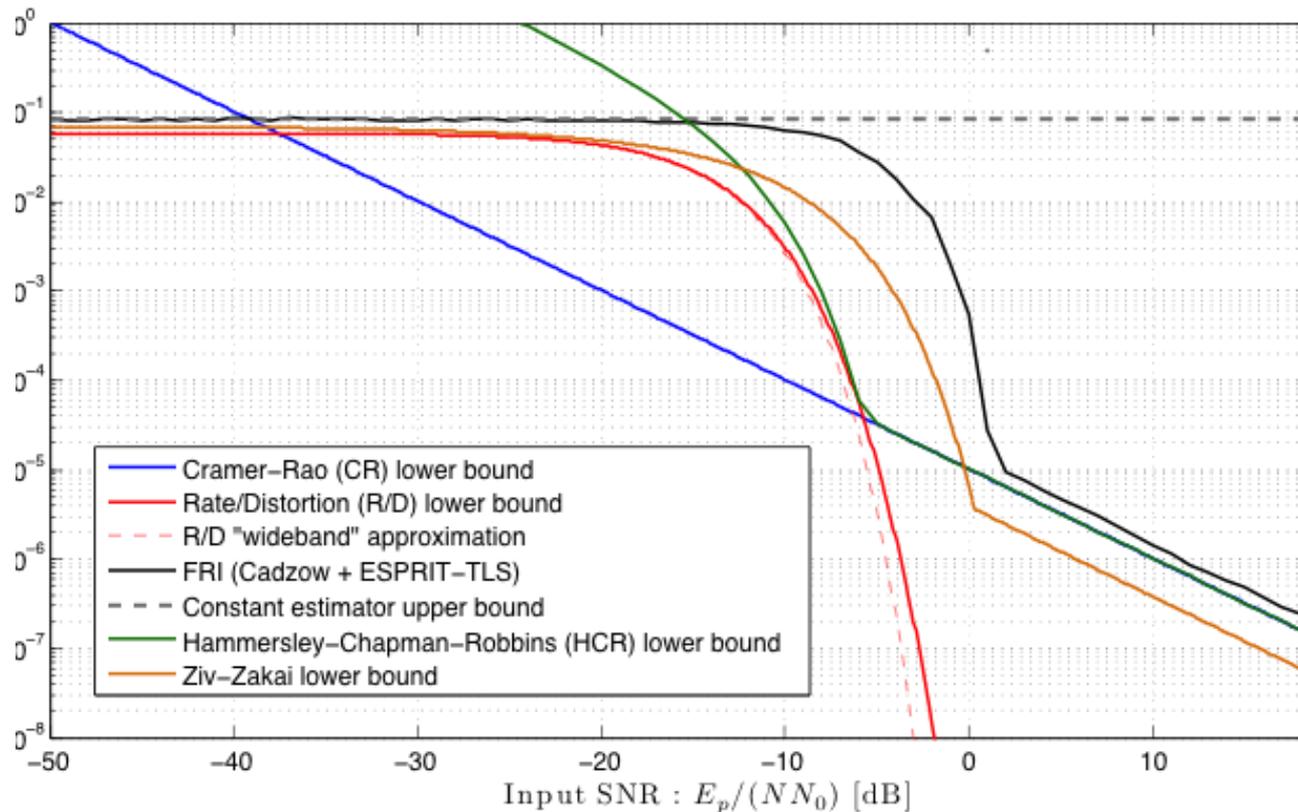
# TheReal World: -5dB Experiment

Find  $[x_k, t_k]$  from noisy samples  $[y_1, y_2, \dots, y_{290}]$



# The Bounds: There is a Phase Transition!

Single Dirac on  $[0,1]$  in noise, with various bounds versus FRI algorithm  
There is a "cut off" SNR, both in FRI and the bounds



# Compressed sensing

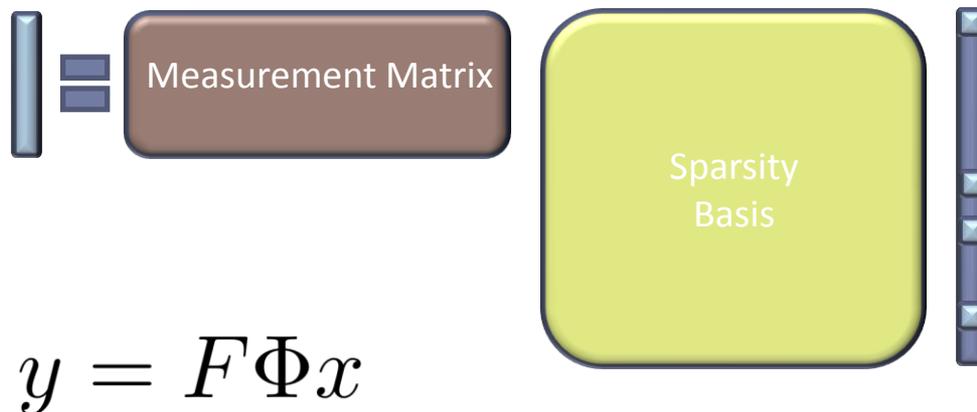
Consider a discrete-time, finite dimensional set up in  $\mathbb{R}^N$ :

**Model:**

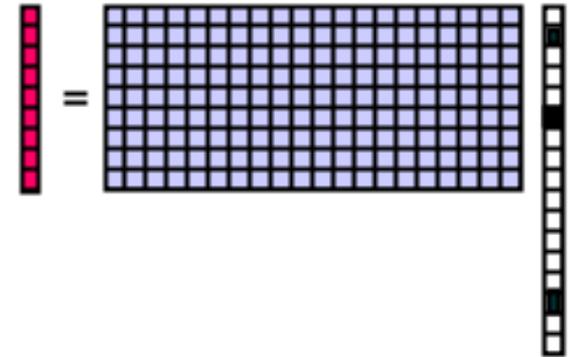
- World is discrete, finite dimensional of size  $N$
- $x \in \mathbb{R}^N$ , but  $|x|_0 = K \ll N$ , that is vector is  $K$ -sparse
- Alternatively,  $K$  sparse in a basis  $\Phi$

**Method:**

- Take  $M$  measurements, where  $K < M \ll N$
- Measurement matrix  $F$ : a fat matrix of size  $M \times N$



# Geometry of the problem



Vastly under-determined system...

Infinite number of solutions...

- $F$  is a frame matrix, of size  $M$  by  $N$ ,  $M \ll N$
- Each set of  $K$  possible entries defines  $M$  by  $K$  matrix  $F_k$
- There are  $\binom{n}{k}$  matrices  $F_k$
- Calculate projection of  $y$  onto range of  $F_k$ ,  $\hat{y}_k = F_k (F_k^* F_k)^{-1} F_k^* y$
- If  $\hat{y}_k = y$ , possible solution
- In general, choose  $k$  such that  $\min_k \|y - y_k\|$
- Note: this is hopeless in general

**Necessary conditions** (for most inputs, or prob. 1)

- $M > K$
- all  $F_k$  must have rank  $K$
- all ranges of  $F_k$  must be different

**It requires completely different attacks!**

# Example: Fourier matrix

## Vandermonde matrices satisfy the geometric conditions

- All submatrices  $F_k$  of rank  $K$
- All subspaces spanned by columns of  $F_k$  are different
- Fourier case: Discrete finite rate of innovation matrix

$$F_{2K \times n} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & W & \dots & W^{N-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & W^{2K-1} & \dots & W^{(2K-1)(N-1)} \end{pmatrix}$$

## Conditioning can be an issue

- As  $N$  grows,  $K$  subsequent columns are “almost” co-linear
- Necessary to take  $M$  measurements that grows faster than  $N$

# CS: The power of random matrices!

Flurry of activity on efficient solutions Donoho, Candes et al

Set up:  $x \in \mathbb{R}^N$ , but  $|x|_0 = K \ll N$  :  $K$  sparse,  $F$  of size  $M$  by  $N$

Measurement matrix with random entries (gaussian, bernoulli)

- Pick  $M = O(K \log N/K)$
- With high probability, this matrix is good!

Condition on matrix  $F$

- Uniform uncertainty principle or restricted isometry property  
All  $K$ -sparse vectors  $x$  satisfy an approx. norm conservation

$$(1 - \delta_K) \|x\|_2^2 \leq \|Fx\|_2^2 \leq (1 + \delta_K) \|x\|_2^2$$

Reconstruction Method:

- Solve linear program under constraint 
$$\min_{\hat{x} \in \mathbb{R}^N} \|\hat{x}\|_1 \quad y = F\hat{x}$$

**Strong result:**  $l_1$  minimization finds, with high probability, sparse solution, or  $l_0$  and  $l_1$  problems have the same solution!

# Sparse sampling and compressed sensing

## Sparse sampling of signal innovations

- + Continuous or discrete, infinite or finite dimensional
- + Lower bounds (CRB) provably optimal reconstruction
- + Close to “real” sampling, deterministic
- Not universal, designer matrices

## Compressed sensing

- + Universal and more general
- ± Probabilistic, can be complex
- Discrete, redundant
- Not continuous time

**The real game:** In the space of frame measurements matrices  $F$

- Best matrix? (constrained grassmanian manifold)
- Tradeoff for  $M$ :  $2K$ ,  $K \log N$ ,  $K \log^2 N$
- Complexity (measurement, reconstruction)

# Outline

1. Introduction
2. Sampling: The linear case
3. Application: Sampling physics
4. Sampling based on timing
5. Sampling: The non-linear case
6. Applications: The non-linear case
  - Multichannel sampling
  - Super-resolution imaging
7. Conclusions: Power game!

# Super-resolution imaging

## 1. What is super-resolution?

Registration: Non-linear... Reconstruction: Linear!

## 2. Multi-channel sampling

Unknown shifts, unknown weights

## 3. Rank condition

Correct shifts lead to a low rank solution

## 4. A new algorithm

Efficient rank minimization



# 1. Registration and Reconstruction



$\Delta x, \Delta y$



# 1. Registration and Reconstruction



$\Delta x', \Delta y'$



# 1. Registration and Reconstruction

## 1. Registration

Is a non-linear problem

Exhaustive search is possible but not computable...

Need for efficient and precise registration

Rank testing is a possibility

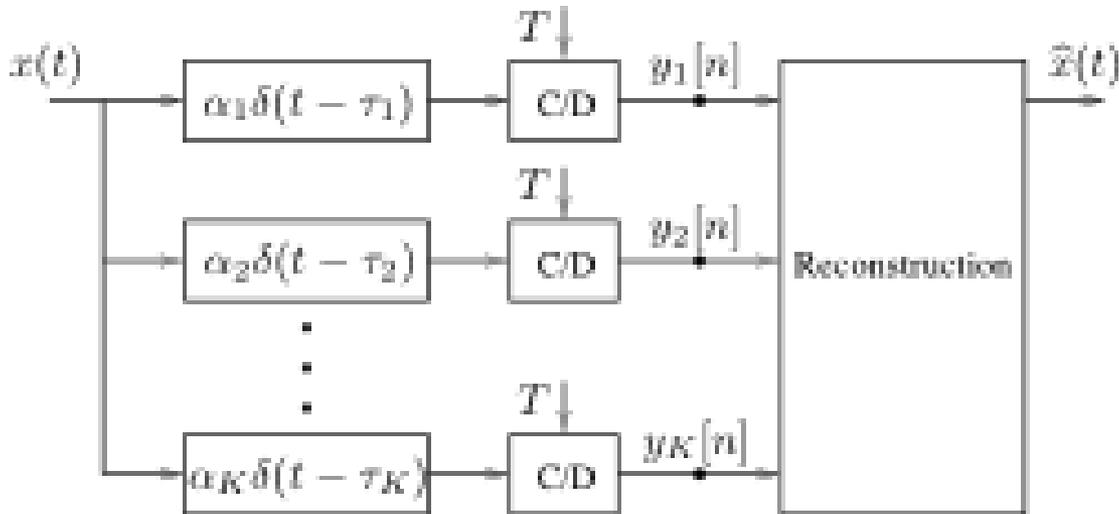
## 2. Reconstruction

Is a linear problem

Solution lives on a subspace

It amounts to solving a linear system of equations

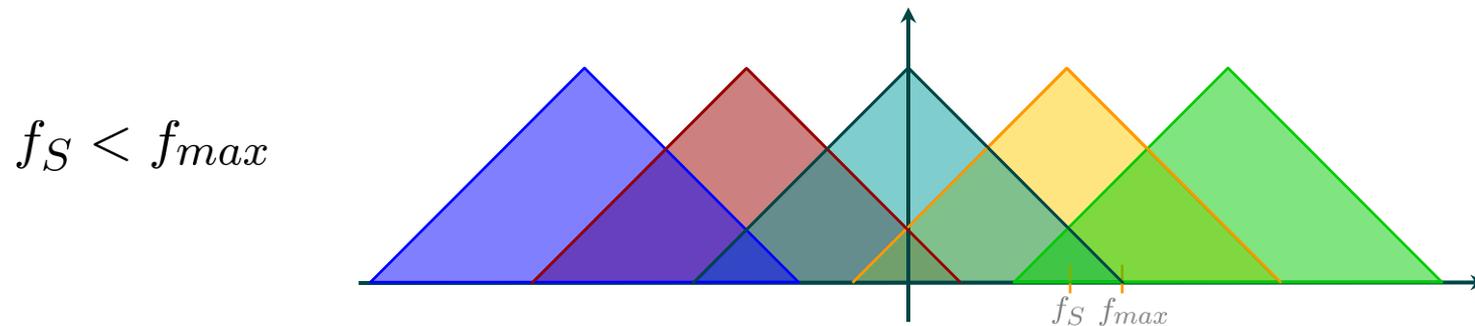
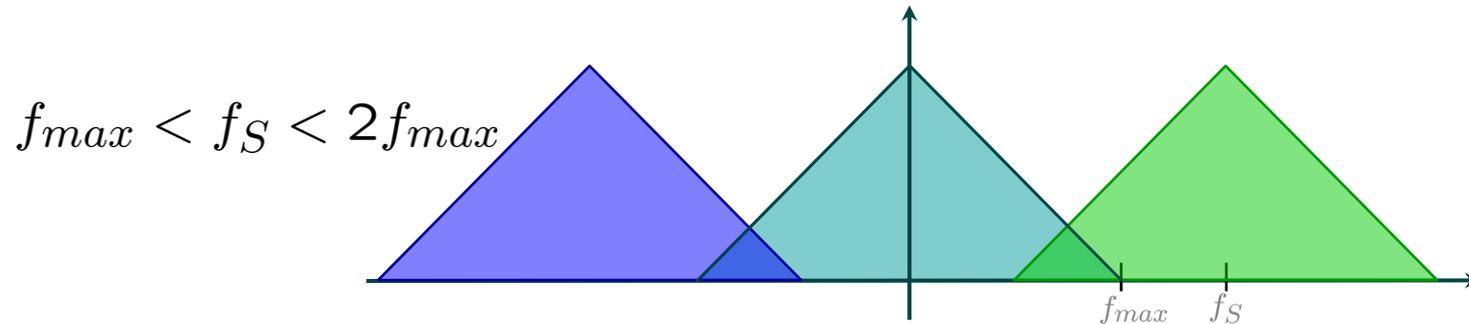
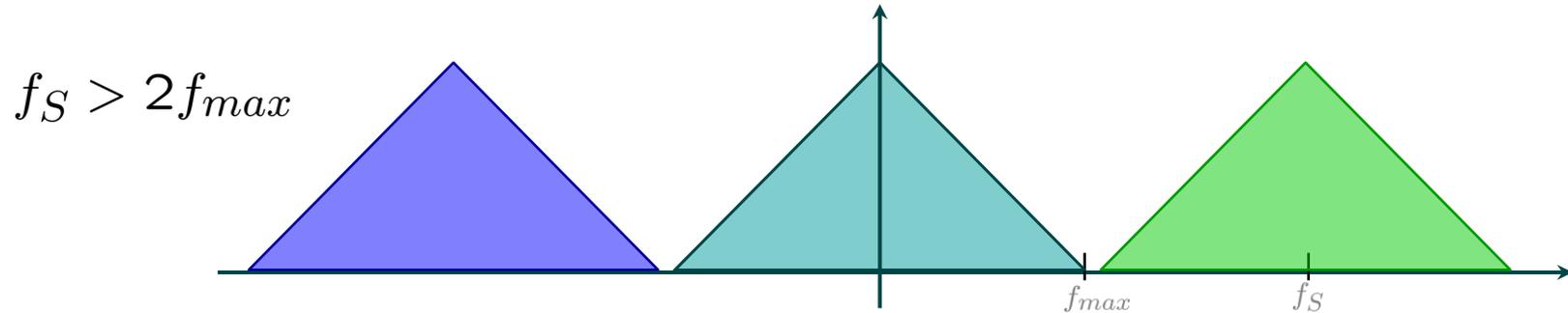
## 2. Multichannel Sampling



- Input:  $x(t)$  bandlimited to  $[-\sigma, \sigma]$
- **Unknown** gains  $\{\alpha_k\}$  and offsets  $\{\tau_k\}$
- Sub-Nyquist sampling:  $\frac{1}{T} < \frac{\sigma}{\pi}$  **aliasing!**

# Subsampling and Aliasing

Depending on sampling rate, 3 different cases



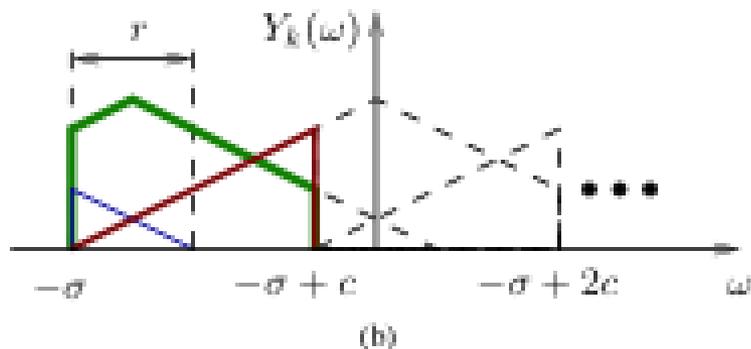
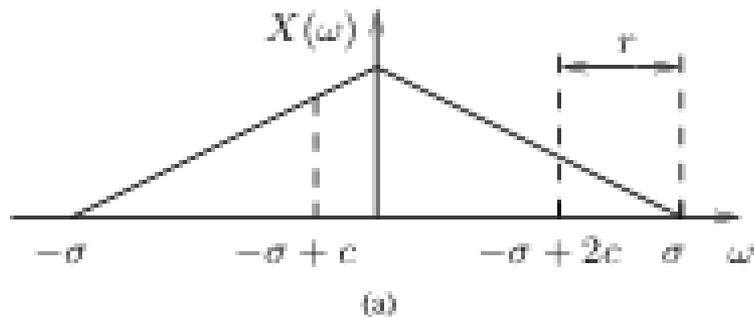
### 3. Rank Condition

Time:

$$y_k[n] = \alpha_k x(nT - \tau_k)$$

Frequency:

$$Y_k(\omega) = \frac{\alpha_k}{T} \sum_{m \in \mathbb{Z}} X(\omega + mc) e^{-j(\omega + mc) \tau_k}$$

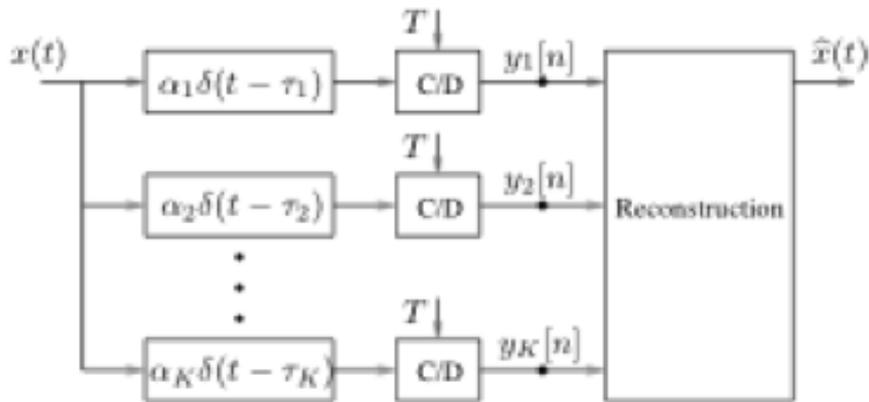


Discrete-time FT:  $Y_k(\omega)$

- Periodic

- A **finite** number of frequency segments folding on top of each other

### 3. Rank Condition (cont.)



System parameters:

unknown gains:  $\alpha \stackrel{\text{def}}{=} [\alpha_1, \dots, \alpha_K]^T$

unknown offsets:  $\tau \stackrel{\text{def}}{=} [\tau_1, \dots, \tau_K]^T$

In the Fourier domain:

channel output

$$Y(\omega) = \Lambda_{\tau}(\omega) \Lambda_{\alpha} V_{\tau} X(\omega)$$

diagonal matrix depending  
on channel gains

unknown input

diagonal matrix depending  
on channel offsets

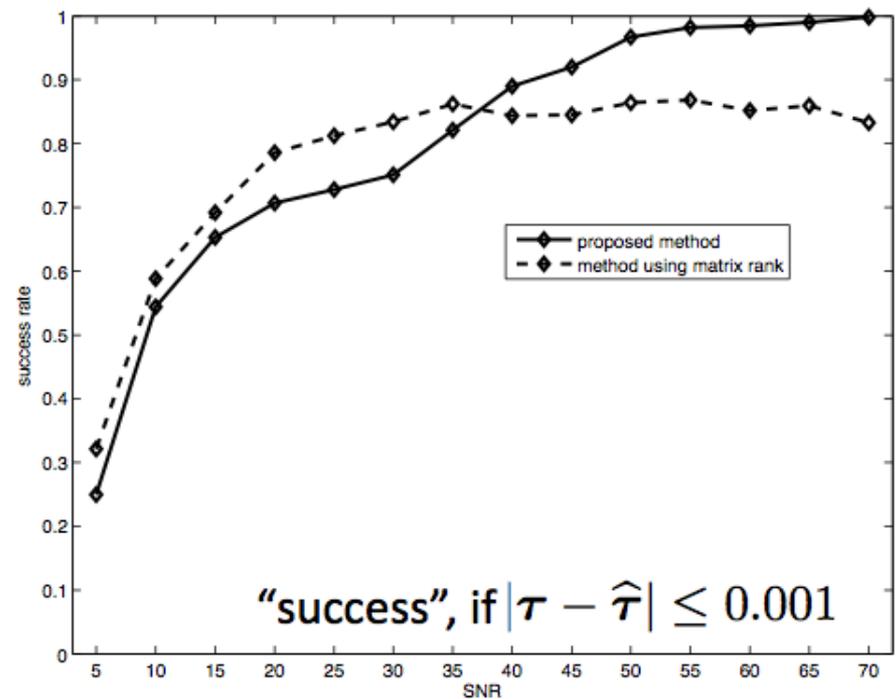
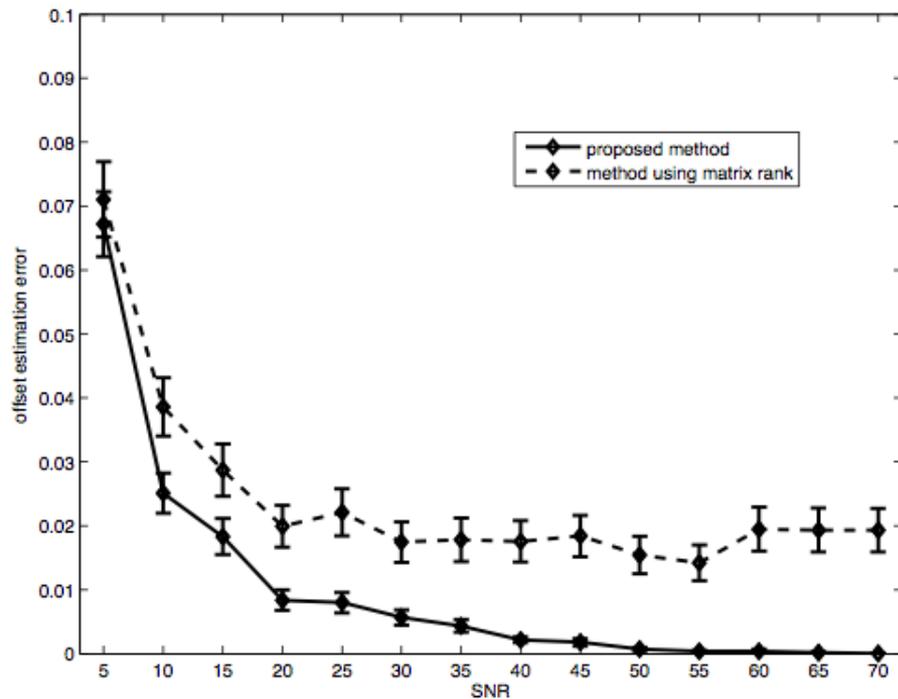
Vandermonde matrix depending  
on channel offsets

# Example

**Setup:** three channels, each channel samples at one-half of the Nyquist rate.

**Estimate:** two unknown offsets  $\tau_2, \tau_3$  (by assumption,  $\tau_1 = 0$ )

**Comparison:** [Vandewalle et al: 07] multiscale search



# Conclusions: Sampling is Alive and Well!

## Classic sampling

- It is a linear problem
- Nice applications in sampling of physics

## Sampling of sparse signals

- It is a non-linear problem
- Separation of location and value
- Different possible algorithms

## Many actual and potential applications

- Fit the model (needs some work)
- Apply the “right” algorithm

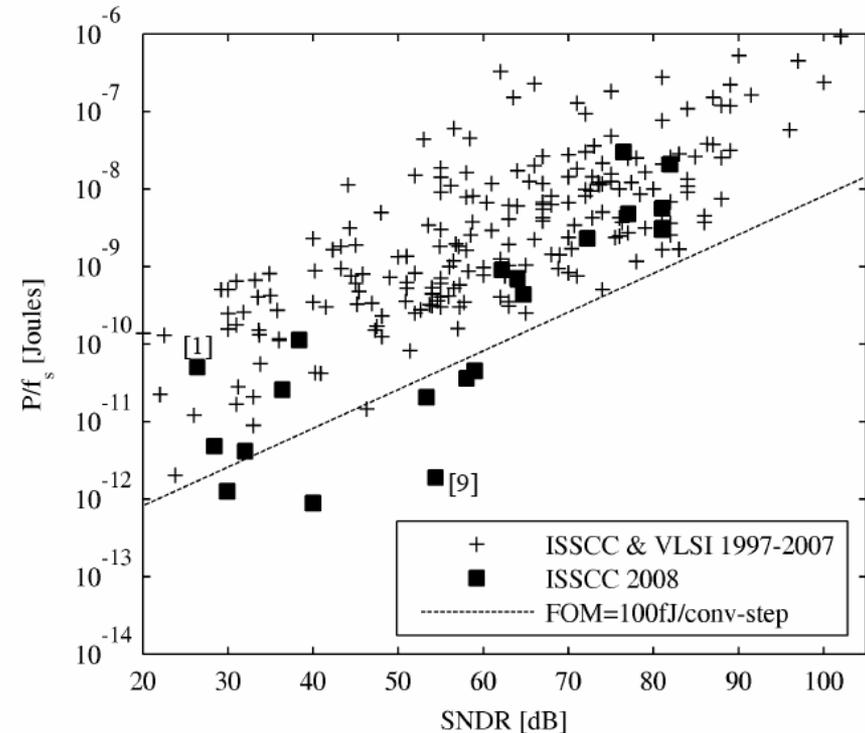
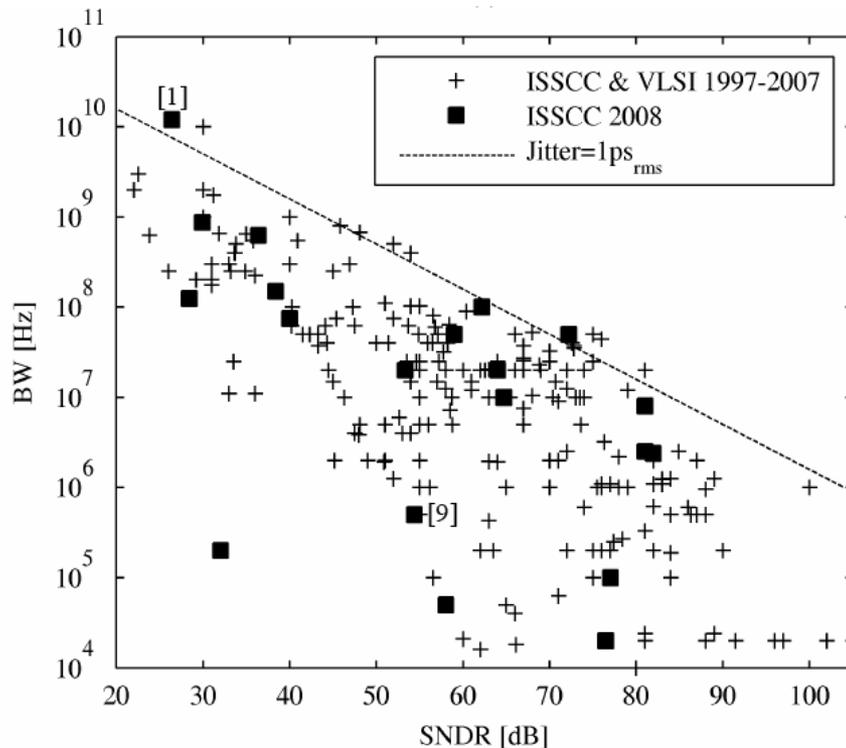
## Still a number of good questions open, from the fundamental to the algorithmic and the applications

- Non-linear problems
- Regularization of the inverse problem
- Faster algorithms
- Designer matrices

# Conclusions: Power game!

## Power or energy is the key driver

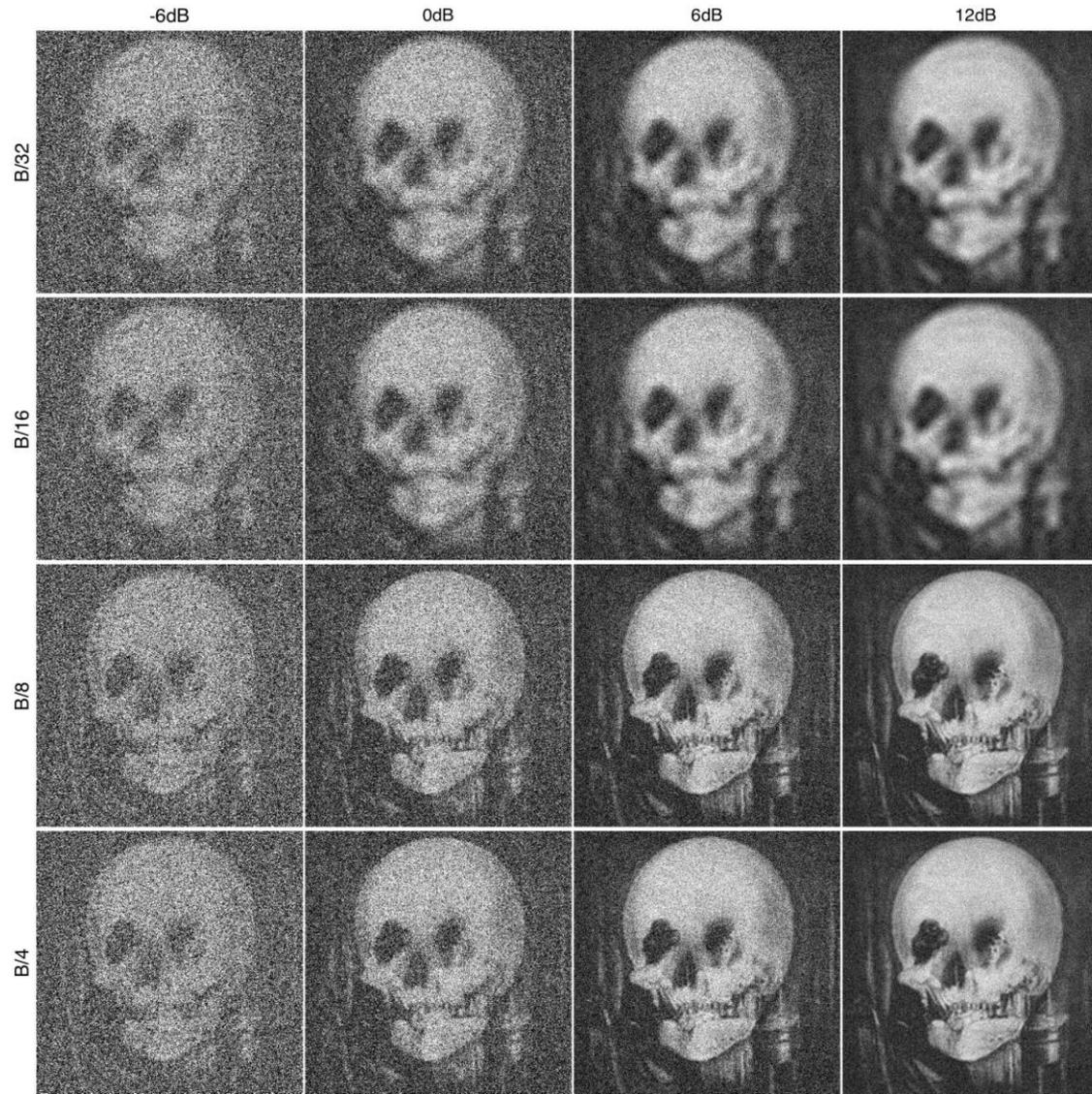
- Insufficiently understood
- Tackles limit of physics



[From B.Murmann, Stanford]

# Conclusions: The Resolution Game!

Resolution is an evasive notion.... Bandwidth and noise!



# Publications

## Special issue on Compressive Sampling:

- R.Baraniuk, E.Candes, R.Nowak and M.Vetterli (Eds.) IEEE Signal Processing Magazine, March 2008. 10 papers overviewing the theory and practice of Sparse Sampling, Compressed Sensing and Compressive Sampling (CS).

## Basic paper:

- M.Vetterli, P. Marziliano and T. Blu, “Sampling Signals with Finite Rate of Innovation,” IEEE Transactions on Signal Processing, June 2002.

## Main paper, with comprehensive review:

- T.Blu, P.L.Dragotti, M.Vetterli, P.Marziliano, and L.Coulot, “Sparse Sampling of Signal Innovations: Theory, Algorithms and Performance Bounds,” IEEE Signal Processing Magazine, Special issue on Compressive Sampling, March 2008.



[Theory, algorithms, and performance bounds]

Thierry Blu,  
Pier-Luigi Dragotti,  
Martin Vetterli,  
Pina Marziliano,  
and Lionel Coulot

Signal acquisition and reconstruction is at the heart of signal processing and sampling theorems provide the bridge between the continuous and the discrete-time worlds. The most celebrated and widely used sampling theorem is often attributed to Shannon (and many others, from Whittaker to Kotelnikov and Nyquist, to name a few) and gives a sufficient condition, namely bandlimitedness, for an exact sampling and interpolation formula. The sampling rate, at twice the maximum frequency present in the signal, is usually called the Nyquist rate. Bandlimitedness, however, is not necessary as is well known but only rarely taken advantage of [1]. In this broader, nonbandlimited view, the question is: when can we acquire a signal using a sampling kernel followed by uniform sampling and perfectly reconstruct it? The Shannon case is a particular example, where any signal from the subspace of bandlimited signals, denoted by BL, can be acquired through sampling and perfectly

# Publications

## For more details:

- P.L. Dragotti, M. Vetterli and T. Blu, “Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon Meets Strang-Fix,” IEEE Transactions on Signal Processing, May 2007.
- P. Marziliano, M. Vetterli and T. Blu, Sampling and exact reconstruction of bandlimited signals with shot noise, IEEE Transactions on Information Theory, Vol. 52, Nr. 5, pp. 2230-2233, 2006.
- I. Maravic and M. Vetterli, Sampling and Reconstruction of Signals with Finite Rate of Innovation in the Presence of Noise, IEEE Transactions on Signal Processing, Aug. 2005.
- A. Hormati, O. Roy, Y.M. Lu and M. Vetterli, Distributed Sampling of Signals Linked by Sparse Filtering: Theory and Applications, IEEE Transactions on Signal Processing, Vol. 58, Nr. 3, pp. 1095 - 1109, 2010.
- Y.Lu and M.Vetterli, Multichannel sampling with unknown gains and offsets: A fast reconstruction algorithm, Allerton 2010.
- A. Hormati, I. Jovanovic, O. Roy, and M. Vetterli, "Robust ray-based reconstruction in ultrasound tomography", SPIE Medical Imaging, 2010.

# Thankyou for your attention !



Lausanne, November 6th, 2009

Any questions ?

# The Frame View: Reconstruction

Consider  $A$ , a bounded linear operator on  $B$ , satisfying

$$\|I-A\| < 1$$

It can be inverted using the Neuman series

$$A^{-1} = \sum_{n=0}^{\infty} (I - A)^n$$

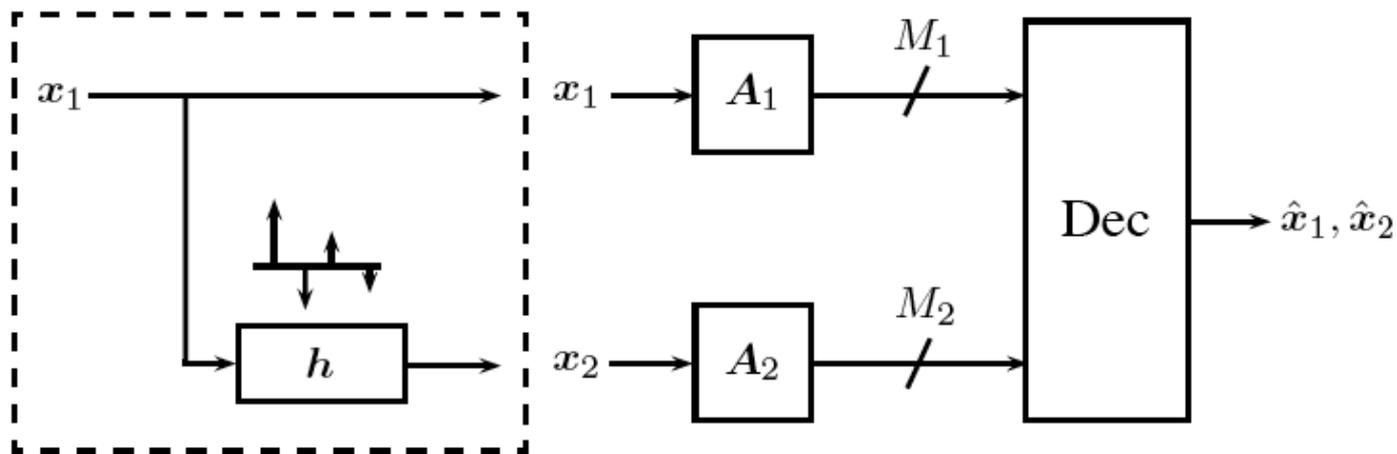
All  $x$  on  $B$  can be reconstructed through the iteration

$$\begin{aligned}\phi_0 &= Ax \\ \phi_{n+1} &= \phi_n - A\phi_n \\ x &= \sum_{n=0}^{\infty} \phi_n\end{aligned}$$

# Applications: Joint Sparsity Estimation

## Two non-sparse signals, but related by sparse convolution

- Often encountered in practice
- Distributed sensing
- Room impulse measurement



Question: What is the sampling rate region?

Similar to Slepian-Wolf problem in information theory

# Joint Sparsity Estimation

## Sampling rate region:

- Universal (all signals): no gain!
- Almost surely (undecodable signals of measure zero)

$$M_1 \geq \min\{K + r, N\}$$

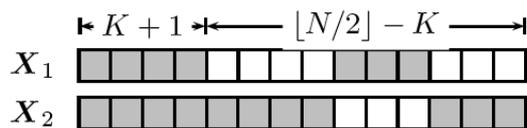
$$M_2 \geq \min\{K + r, N\}$$

$$M_1 + M_2 \geq \min\{N + K + r, 2N\}$$

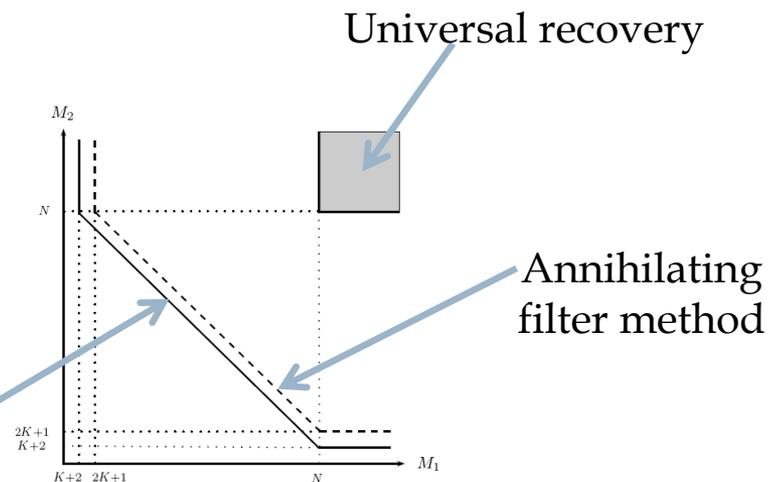
$$M_1 \geq \min\{2K + 1, N\}$$

$$M_2 \geq \min\{2K + 1, N\}$$

$$M_1 + M_2 \geq \min\{N + 2K + 1, 2N\}$$



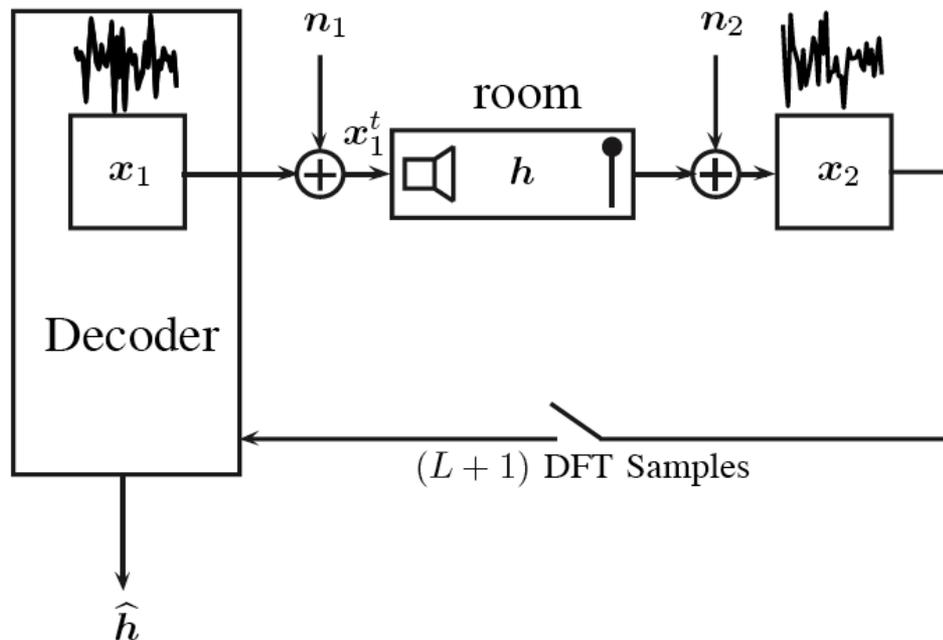
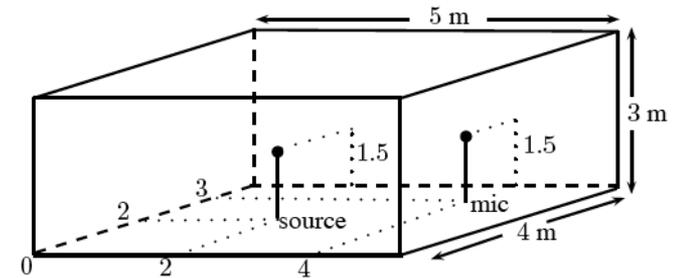
Almost sure  
recovery



# Joint Sparsity Estimation

Ranging, room impulse response, UWB:

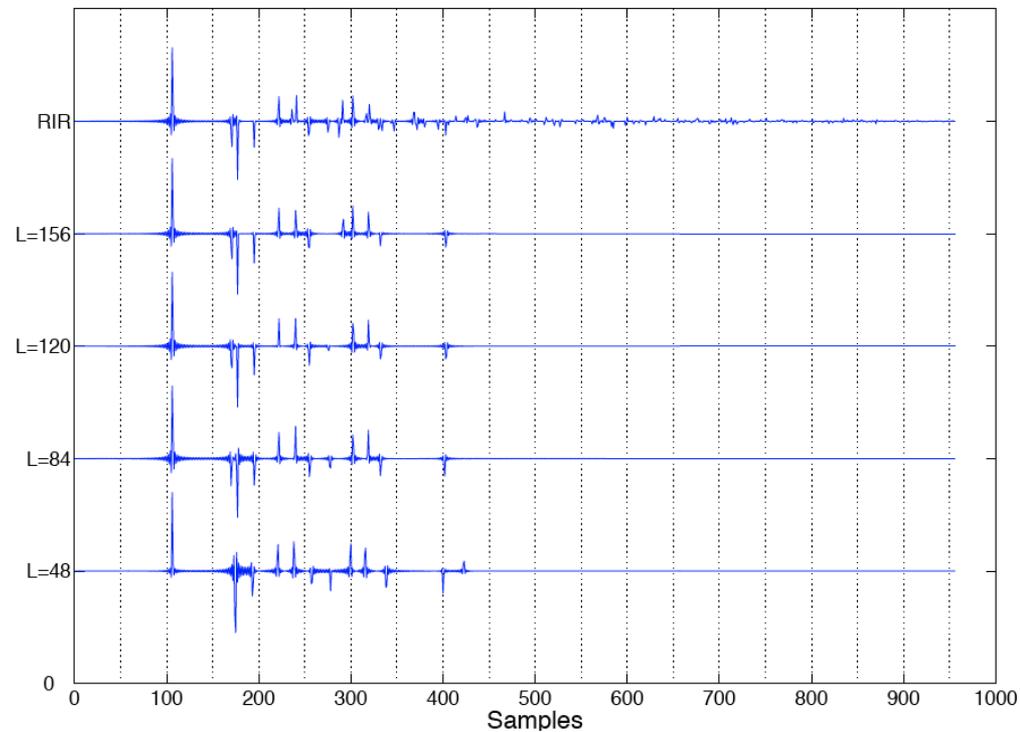
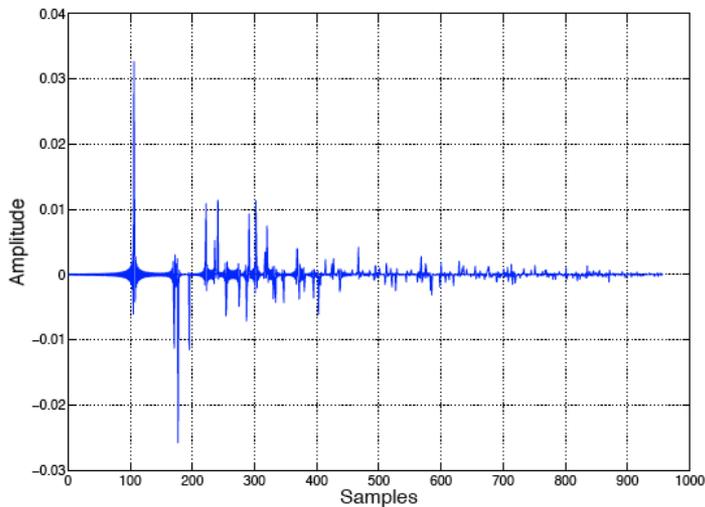
- Know signal  $x_1$ , low rate acquisition of  $x_2$



# Joint Sparsity Estimation

## Experiment

- Known signal  $x_1$ ,
- Low rate acquisition of  $x_2$ ,
- Various sparsity levels for acquisition of  $x_2$



# From Theorems to Patents to Products...

Theorem 1 -> Patent 1

{Patents 1..N} -> Patent troll? No!

Real technology transfer!

(19) **United States**  
 (12) **Patent Application Publication** (10) **Pub. No.:** US 2007/0143078 A1  
**Vetterli et al.** (43) **Pub. Date:** Jun. 21, 2007

---

(54) **SAMPLING METHOD, RECONSTRUCTION METHOD, AND DEVICE FOR SAMPLING AND/OR RECONSTRUCTING SIGNALS** (30) **Foreign Application Priority Data**  
 Mar. 26, 2001 (EP)..... 01107530.6  
 Aug. 15, 2001 (EP)..... 01119537.7

(76) Inventors: **Martin Vetterli**, Grandvaux (CH); **Pina Marziliano**, Montreal (CA); **Thierry Blu**, Lausanne (CH)

Correspondence Address:  
**Peter A. Businger, Esq.**  
 344 Valleyscent Avenue  
 Scotch Plains, NJ 07076-1170 (US)

(21) Appl. No.: **10/680,833**  
 (22) Filed: **Oct. 7, 2003**

**Related U.S. Application Data**  
 (63) Continuation of application No. PCT/EP02/03380, filed on Mar. 26, 2002.

(51) **Int. Cl.**  
**G06F 15/00** (2006.01)  
 (52) **U.S. CL.** ..... **702/189**

(57) **ABSTRACT**  
 Reconstruction method for reconstructing a first signal  $x(t)$  regularly sampled at a sub-Nyquist rate, comprising the step of retrieving from the regularly spaced sampled values  $\{y_s[n], y(nT)\}$  a set of weights  $\{c_n, c_{nT}\}$  and shifts  $\{t_n, t_n\}$  with which said first signal  $x(t)$  can be reconstructed.  
 The reconstructed signal  $x(t)$  can be represented as a sequence of known functions  $\{r(t)\}$  weighted by the weights  $\{c_n\}$  and shifted by the shifts  $\{t_n\}$ . The sampling rate is at least equal to the rate of innovation  $(\rho)$  of the first signal  $x(t)$ .

