BOAT WAKE © Pete Turner

# Sampling Theory and Practice: 50 Ways to Sample your Signal

Martin Vetterli EPFL & UC Berkeley

# Acknowledgements

**Sponsors**

– NSF Switzerland: thank you!
– Conference organizers, Thierry Blu in particular!

**Co-authors and collaborators**

– Thierry Blu, CUHK
– Yue Lu, Harvard
– Amina Chebira, EPFL
– Ali Hormati, EPFL
– Yann Barbotin, EPFL
– J.Rarienri, EPFL

**Discussions and Interactions**

– V. Goyal, MIT
– A.Fletcher, UCB
– M. Unser, EPFL

# Outline

**1. Introduction**

The world is analog, computation is digital

**2. Sampling: The linear case**

Classic results and generalization: Known locations

**3. Application: Sampling physics**

The plenacoustic function and sampling wave fields

**4. Sampling: The non-linear case**

Finite rate of innovation sampling

Compressed sensing

**5. Applications: The non-linear case**

Diffusion equation

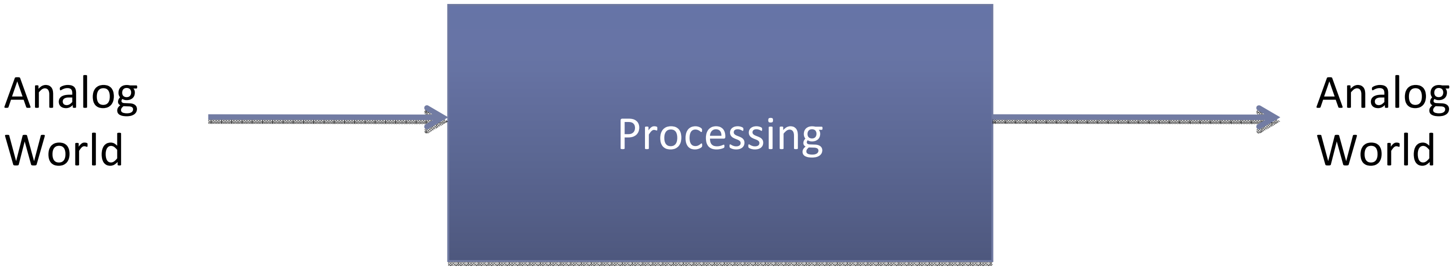
Multichannel sampling

Super-resolution imaging

**6. Conclusions**

# The situation

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



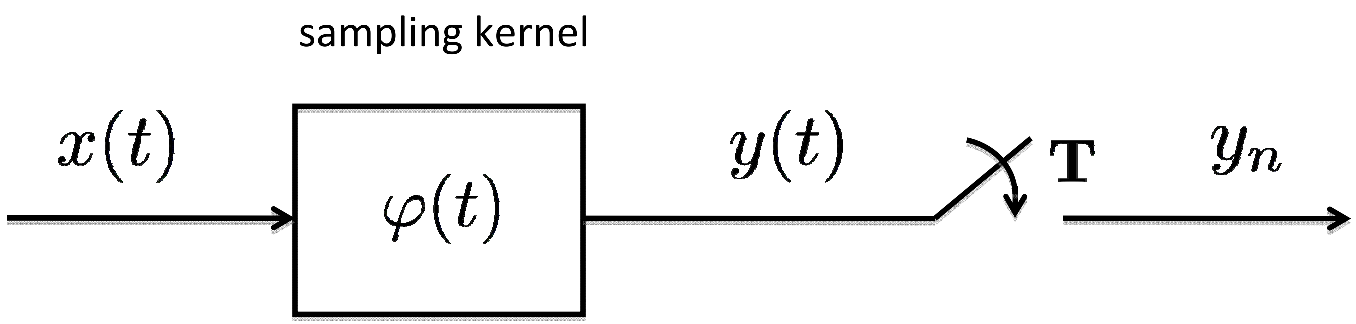

**Ex: Audio, sensing, imaging, computer graphics, simulation etc**

# The Question:

**You are given a class of objects, like a class of functions (bandlimited)**

**You have a sampling device, as usual to acquire the real world**

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




**Obvious question:**

**When does a minimum number of samples uniquely specify the function?**

$$x(t) \Longleftrightarrow y_n$$

# Kernel and sampling rate

**About the observation kernel:**

**Given by nature**

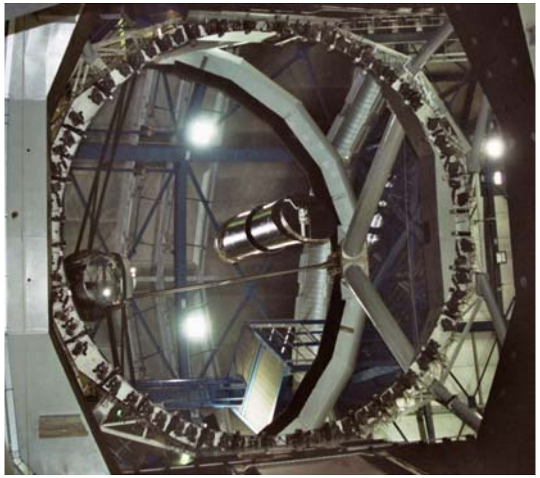
– Diffusion equation, Green function
  Ex: sensor networks

**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



**About the sampling rate:**

**Given by problem**

– Ex: sensor networks

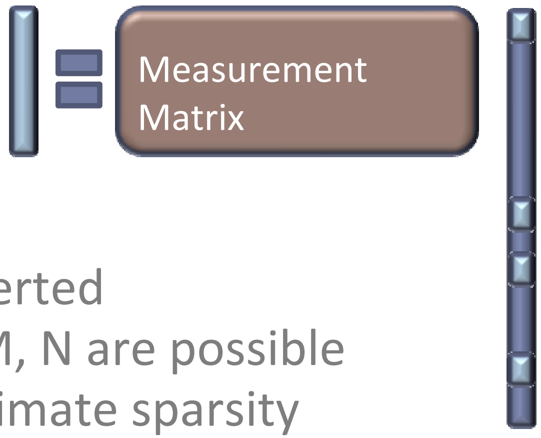
**Given by design**

– Usually, as low as possible
  Ex: digital UWB

# A Variation: Compressed Sensing

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

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




**Questions**

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

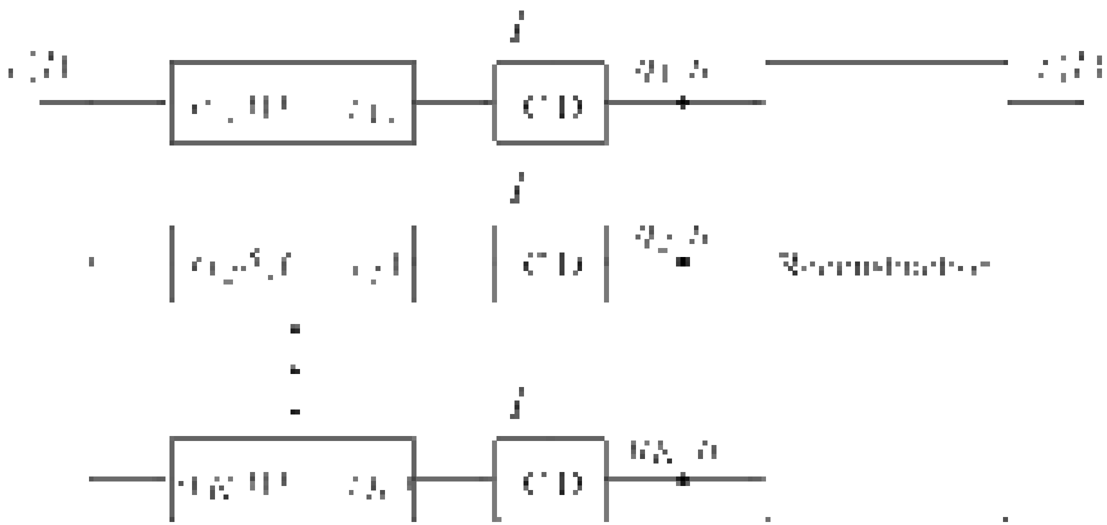
**Problem is non-linear in the location of non-zero entries of x!**

# 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!**

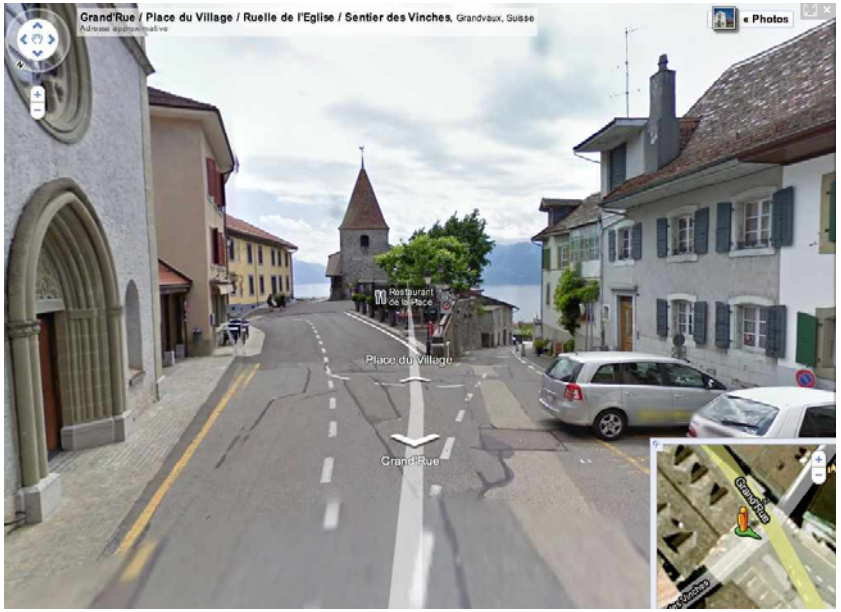


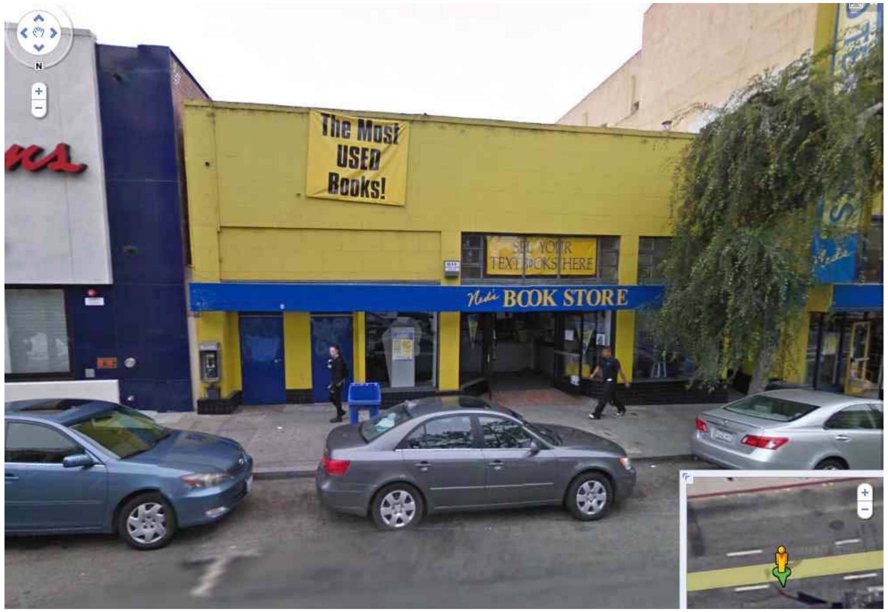
# Are these real problems? (1)

**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

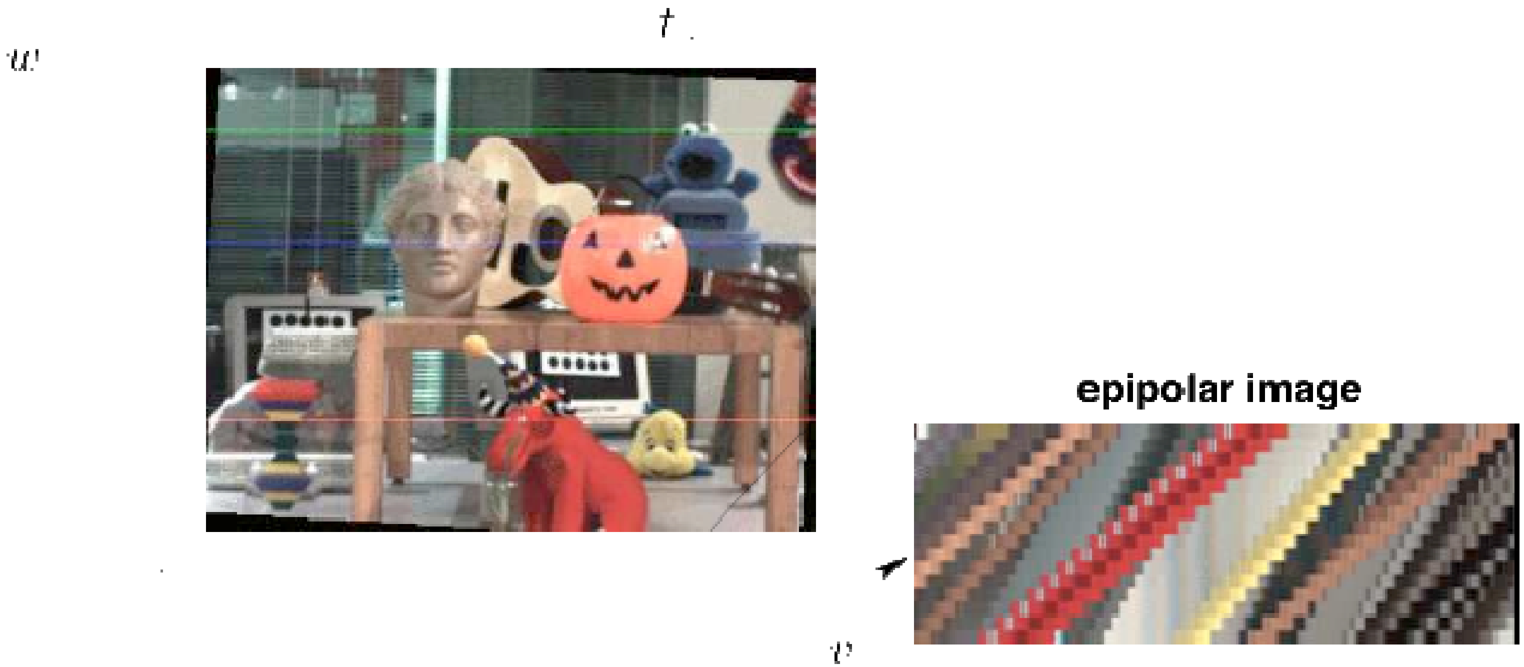


# Plenoptic sampling

**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

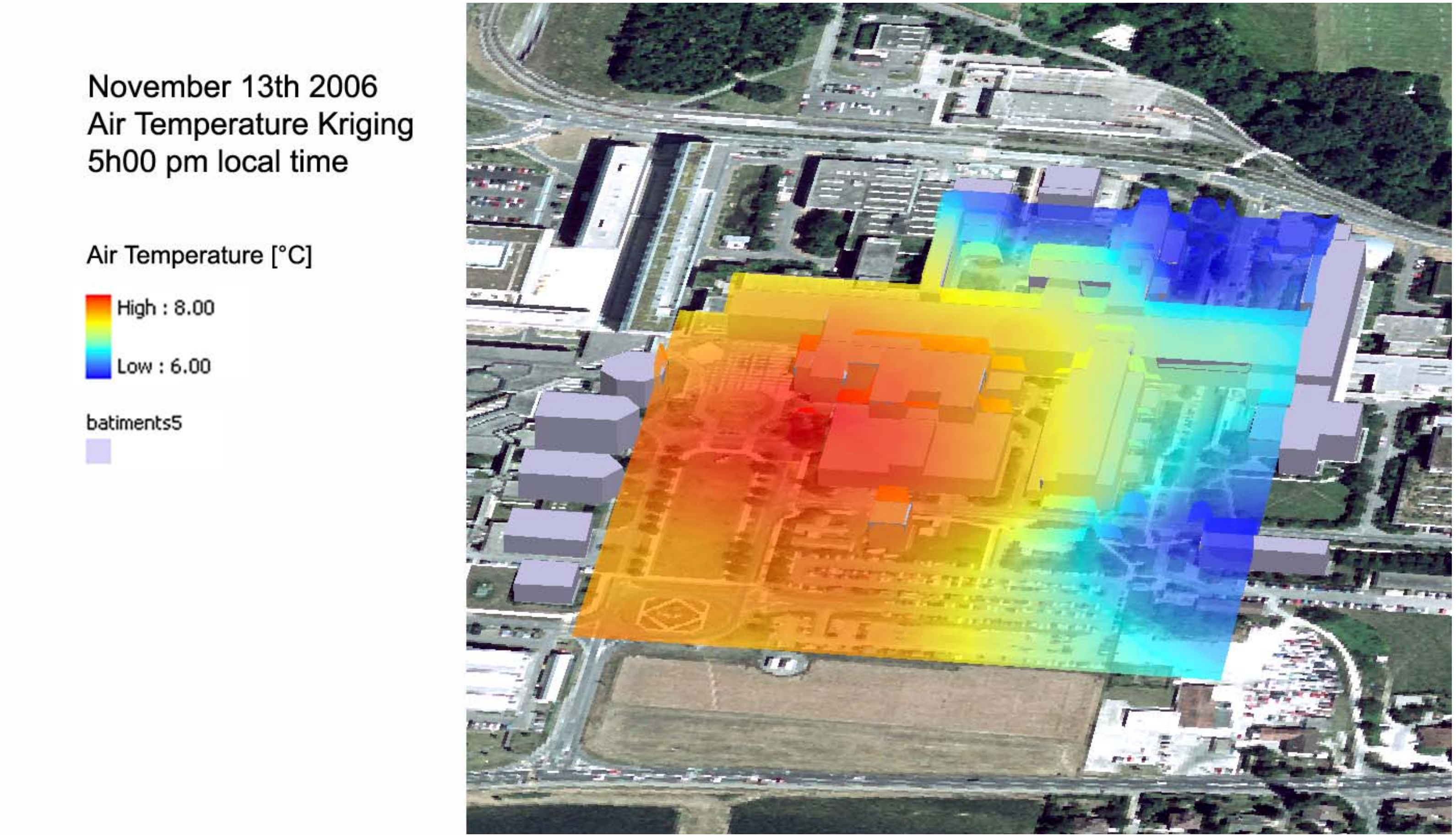
[Dragotti et al, 2008]

# Are these real problems?

**Sensor networks as another relevant example**

How many sensors

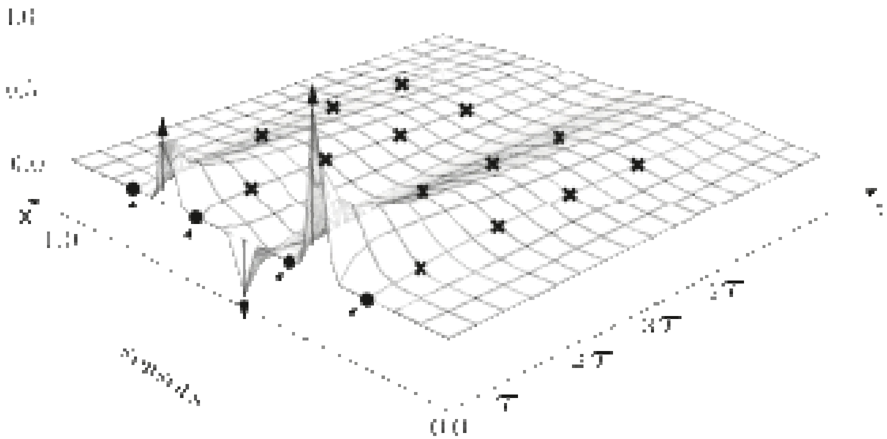
How to reconstruct



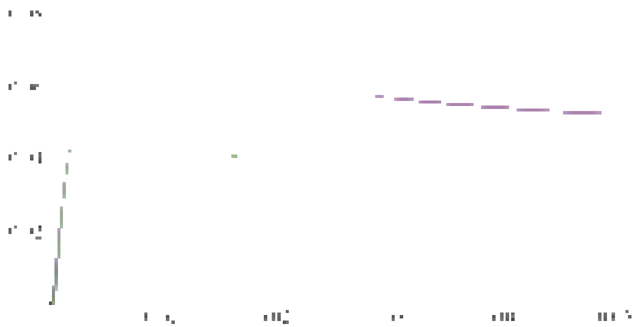
# Diffusion equation and inversion

**Point sources**

**Observation by sensors**



Over space



Over time

# Outline

**1. Introduction**

**2. Sampling: The linear case**

Shannon sampling

Subspace sampling

Irregular sampling with known locations

**3. Application: Sampling physics**

**4. Sampling: The non-linear case**

**5. Applications: The non-linear case**

**6. Conclusions**

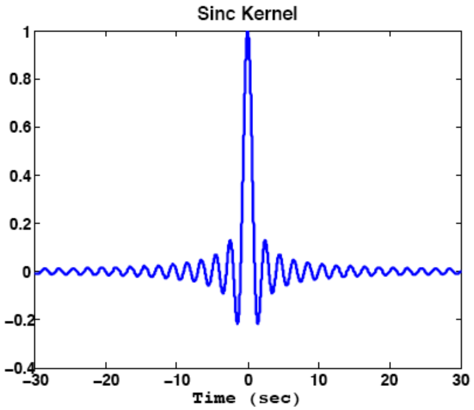
# Classic Sampling Case [WNKWRGSS, 1915-1949]

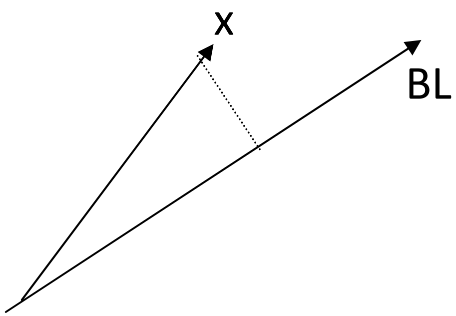
*If a function x(t) contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced 1/(2W) seconds apart. [if approx. T long, W wide, 2TW numbers specify the function]*

**It is a representation theorem:**

- sinc(t-n) is an orthogonal basis for BL[-π,π]
- x(t) ∈ BL[-π,π] can be written as

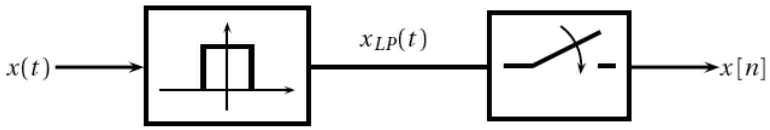
$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \mathrm{sinc}\left(\frac{t - nT_s}{T_s}\right)$$




**Note:**

- … slow convergence of the series
- Shannon-BW, BL sufficient, not necessary



# Shannon's Theorem… a bit of History

Whittaker 1915

Nyquist 1928

Kotelnikov 1933

Whittaker  1935

Raabe 1938

Gabor 1946

Shannon 1948

Isao 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**

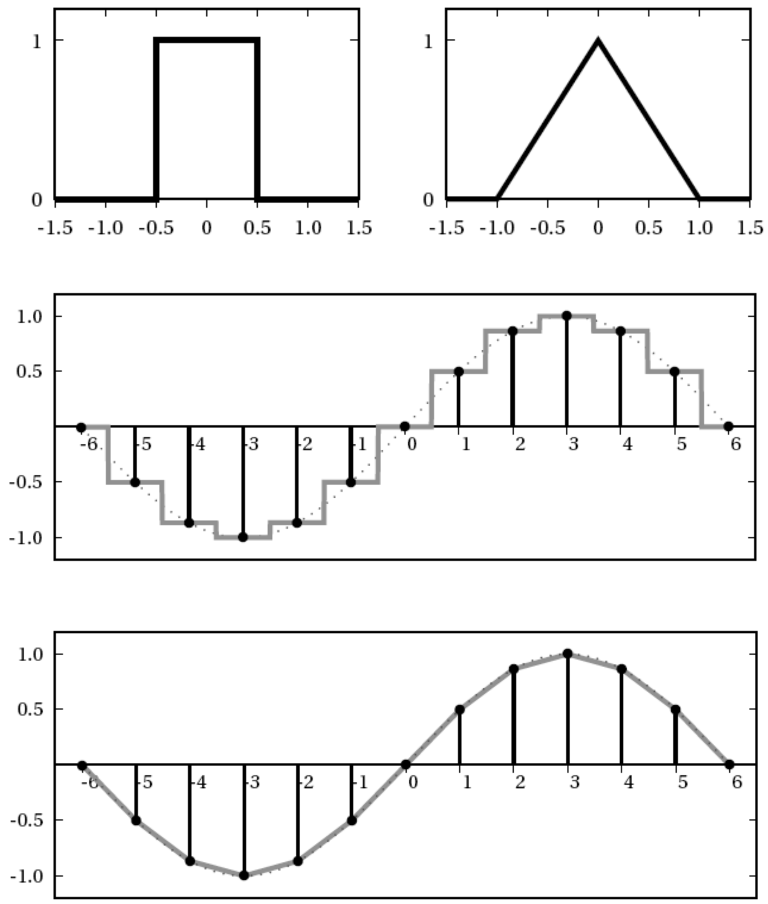
- Power spectrum is good enough

**Shift-invariant subspaces**

- 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 (superresolution imaging)



# Outline

1. **Introduction**
2. **Sampling: The linear case**
3. **Application: Sampling physics**
   - Microphone and loudspeaker arrays
   - The plenacoustic function
   - A sampling theorem for wave fields
4. **Sampling: The non-linear case**
5. **Applications: The non-linear case**
6. **Conclusions**

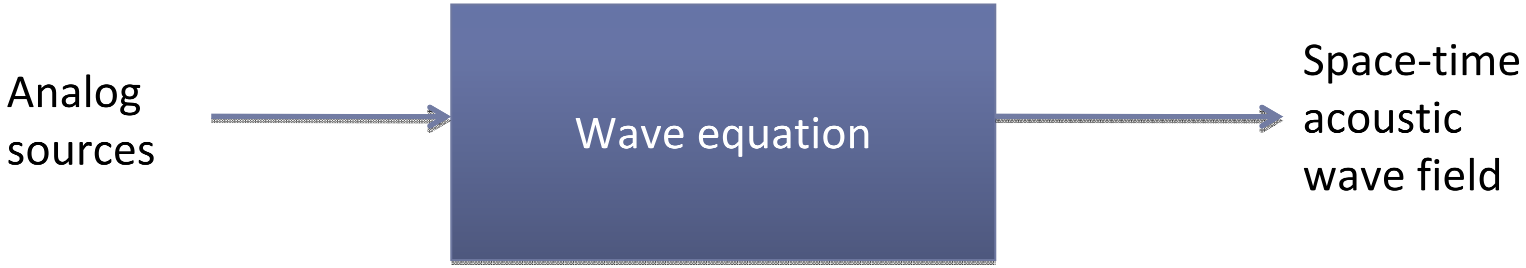
# Sampling and interpolating acoustic fields

**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?**

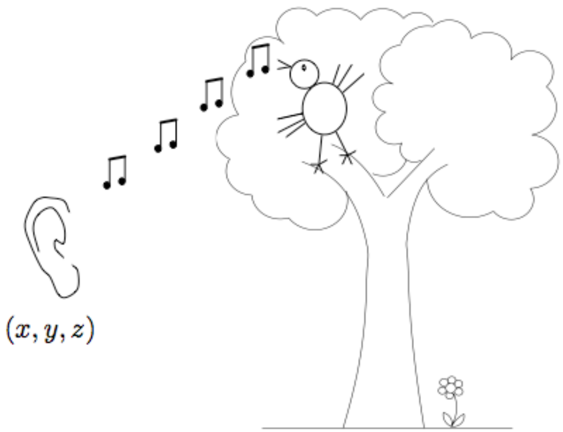
# Many microphones and loudspeakers

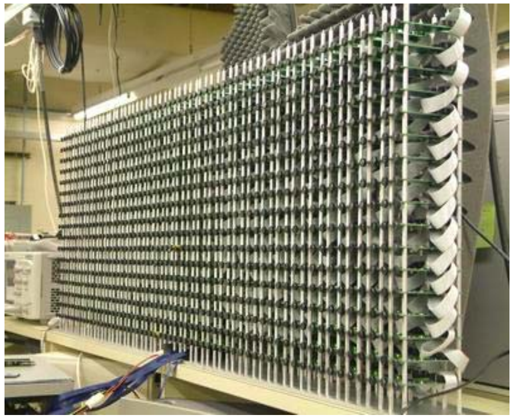
**Multiple microphones/loudspeakers**

- 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

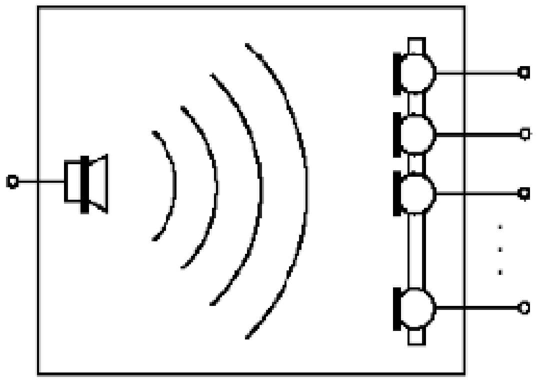


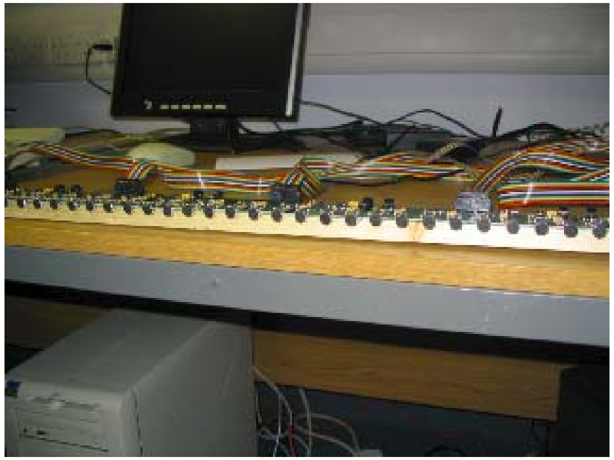



MIT1020 mics

# The plenacoustic function and its sampling

**Setup**



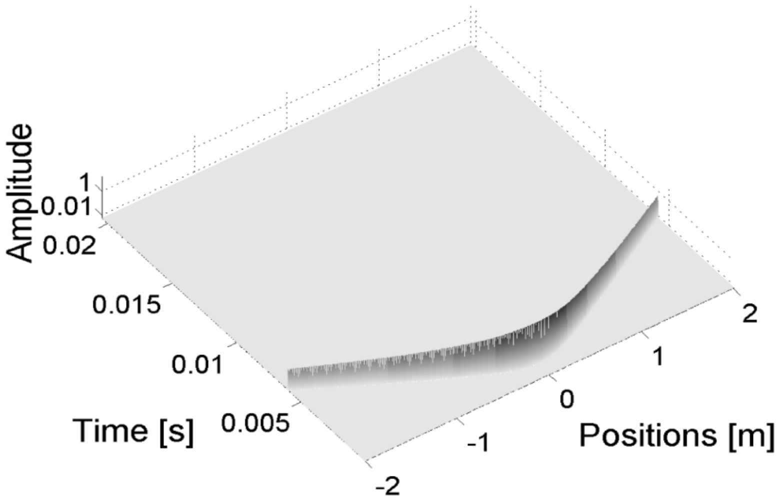


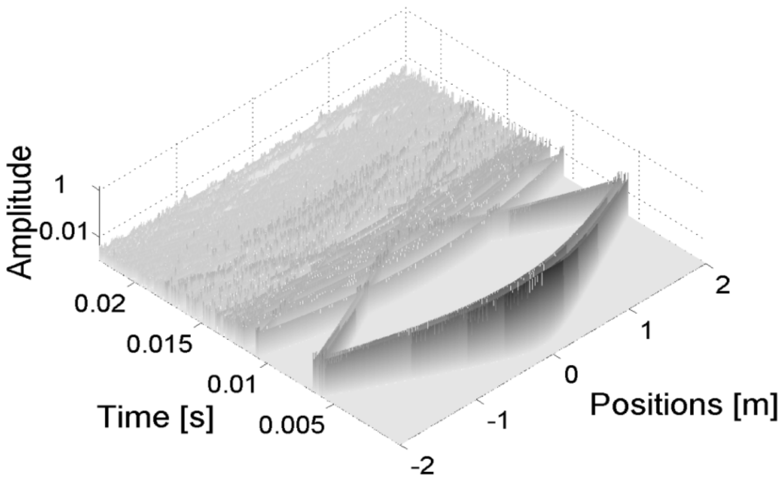
**Questions:**

- 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

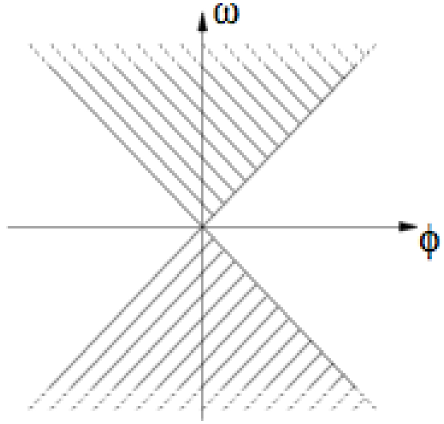
**PAF in free field and in a room for a given point source**




- 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

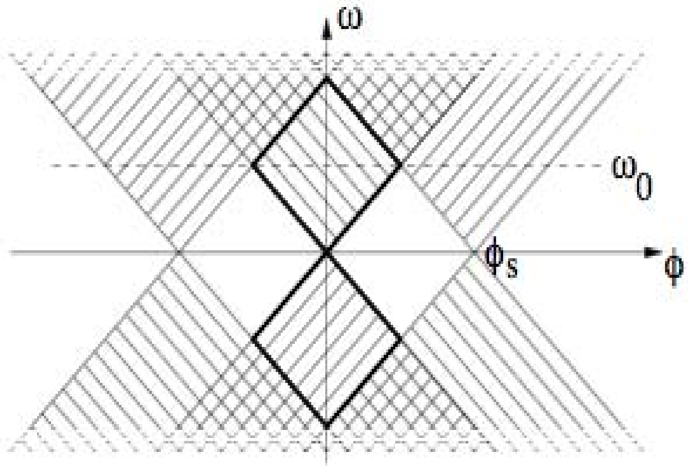



slope: $\omega = c\ \phi$

$\omega$: temporal frequency

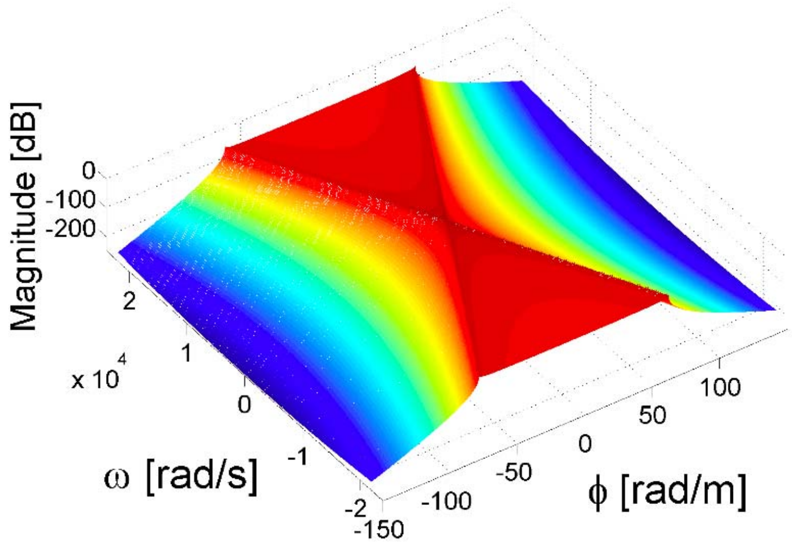
$\Phi$: spatial frequency

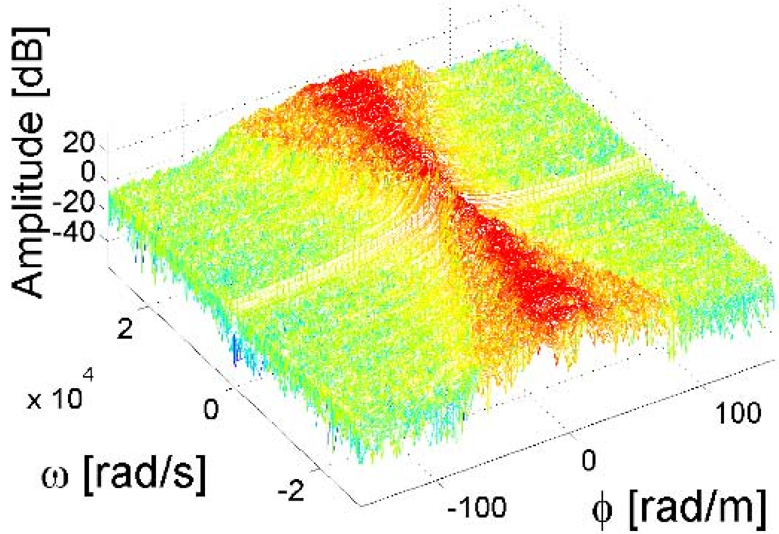
**Sampled Version:**




Thus: Spatio-temporal soundfield can be reconstructed up to $\omega_0$

# Simulated and measured PAF




Almost bandlimited!

Measurement includes noise and temperature fluctuations

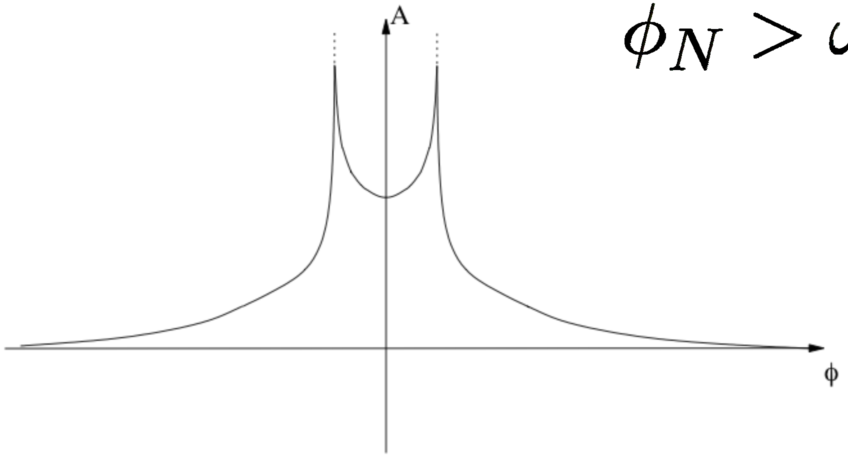
# A Sampling Theorem for Acoustic Fields

**Theorem [ASV:06]:**

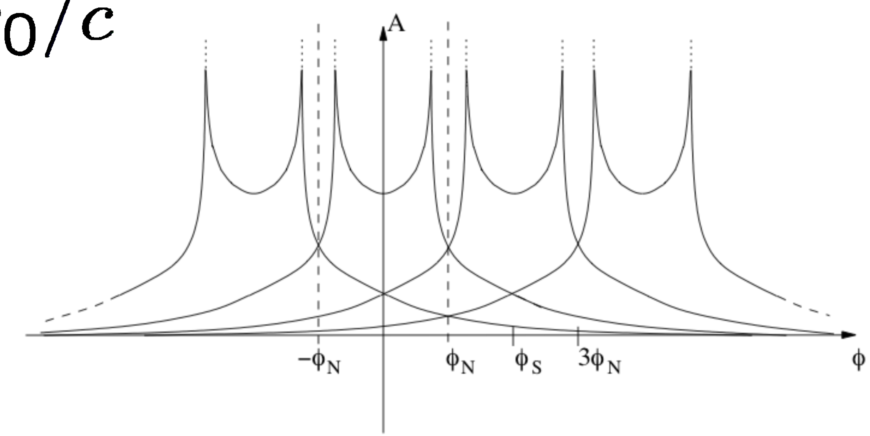
- Assume a max temporal frequency $\omega_0$
- Pick a spatial sampling frequency $\phi_N > \omega_0/c$
- Spatio-temporal signal interpolated from samples taken at $(2\omega_0, 2\phi_N)$

**Argument:**

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




$$\phi_N > \omega_0/c$$



# Outline

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

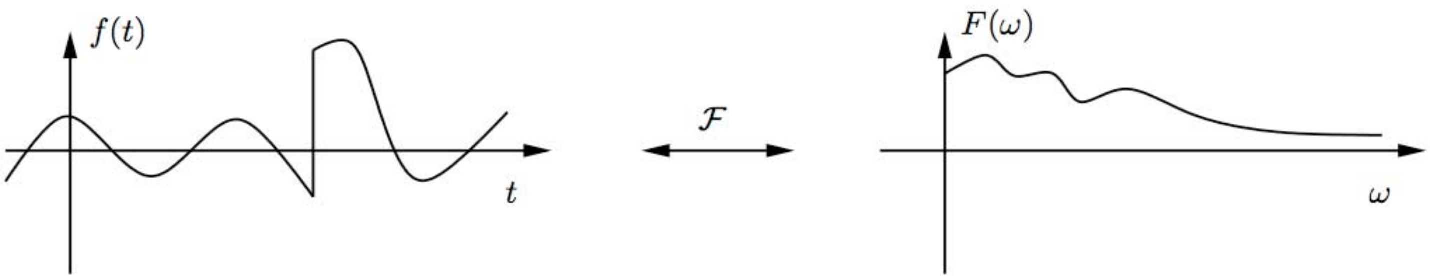
# Classic Case: Subspaces

**Shannon BL case**

$$x(t) = \sum_{n \in Z} x(nT)\mathrm{sinc}(t/T - n)$$

or 1/T degrees of freedom per unit time

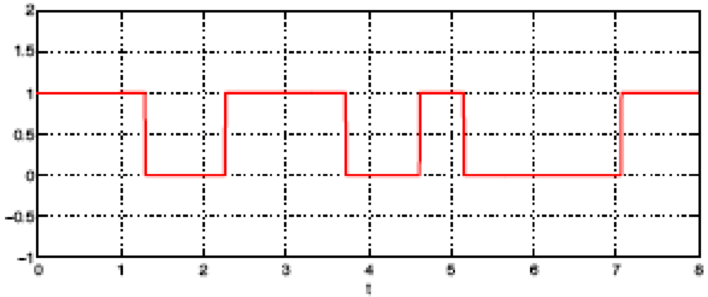
**But: a single discontinuity, and no more sampling theorem…**



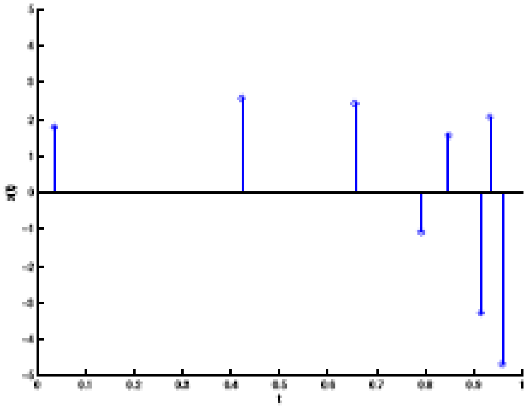

**Are there other signals with finite number of degrees of freedom per unit of time that allow exact sampling results?**

# Examples of non-bandlimited signals

Bilevel signals
PPM,CDMA

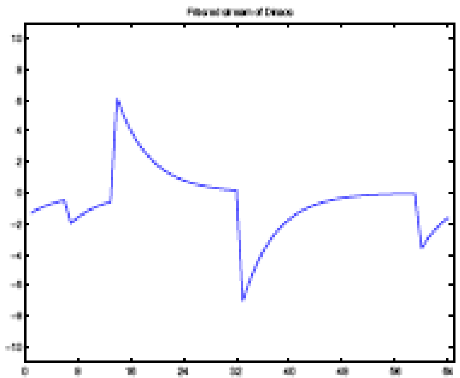



Stream of Diracs
Poisson process



Piecewise Polynomial
Woodcut picture

Filtered stream of Diracs
Neural spikes
UWB



# Classic Case and Beyond…

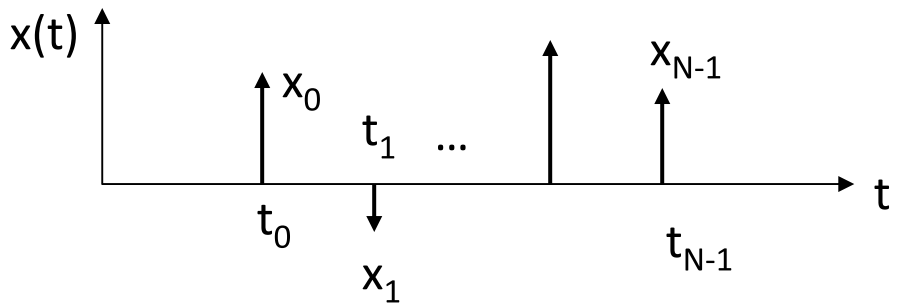
**Is there a sampling theory beyond Shannon?**

- 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

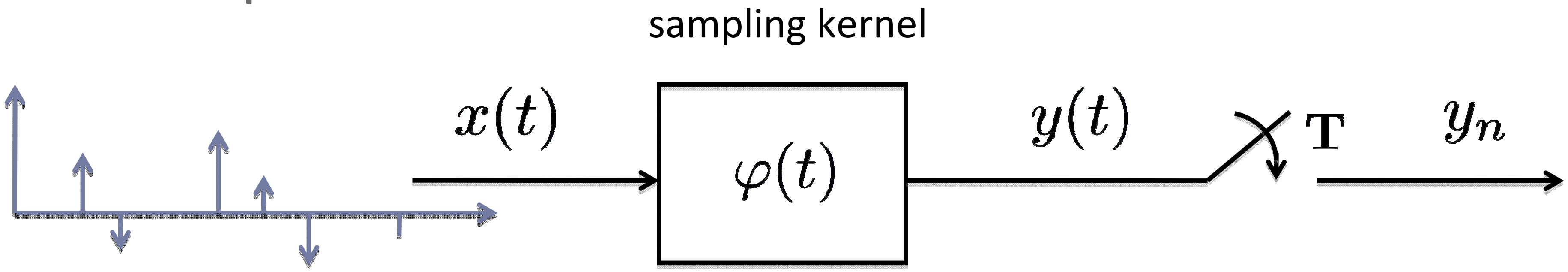
$$\rho = \lim_{T\to\infty} \frac{1}{T} C_T$$

**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 $X_m$

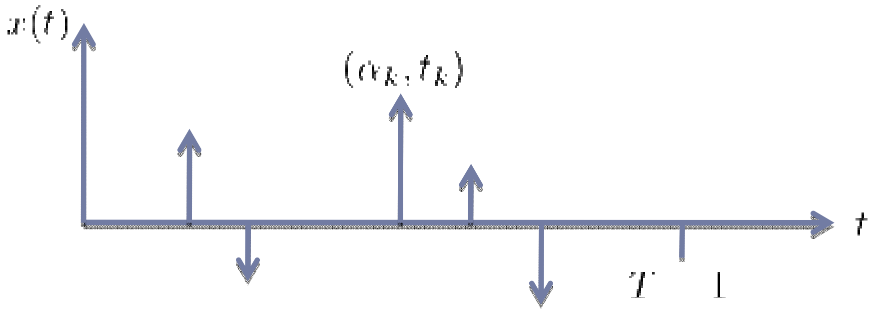
$$x(t) = \sum_{m \in Z} X_m \, e^{-j2\pi mt}$$

**Fourier integral leads to**

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

**If we can identify the exponents**

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




$$X_m = \int_0^1 x(t)\, e^{-j2\pi mt} dt$$

$$= \int_0^1 \left(\sum_{k=0}^{K-1} \alpha_k \delta(t - t_k)\right) e^{-j2\pi mt} dt$$

$$= \sum_{k=0}^{K-1} \alpha_k e^{-j2\pi m t_k}$$

# A Representation Theorem [VMB:02]

**Theorem**

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

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

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

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

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).

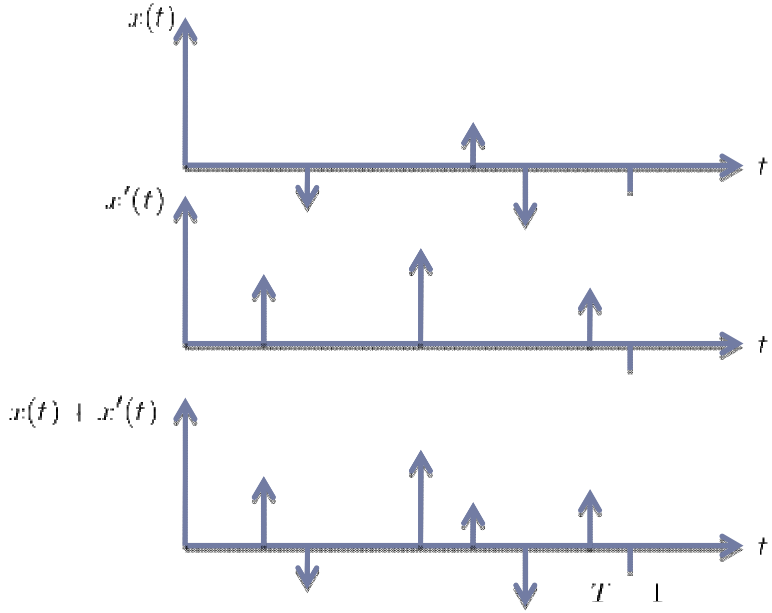
# Linear versus non-linear problem

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

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

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

$$\begin{bmatrix} \hat{x}_{-1} & \hat{x}_{-2} & \cdots & \hat{x}_{-K} \\ \hat{x}_{0} & \hat{x}_{-1} & \cdots & \hat{x}_{-K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}_{K-2} & \hat{x}_{K-3} & \cdots & \hat{x}_{-1} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_K \end{bmatrix} = - \begin{bmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_{K-1} \end{bmatrix}.$$

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

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

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! □

# Notes on Proof

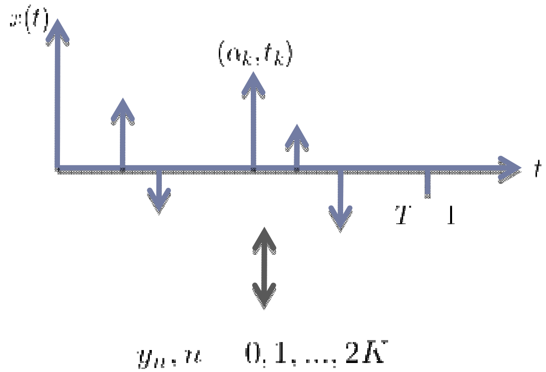
**The procedure is constructive, and leads to an algorithm:**

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

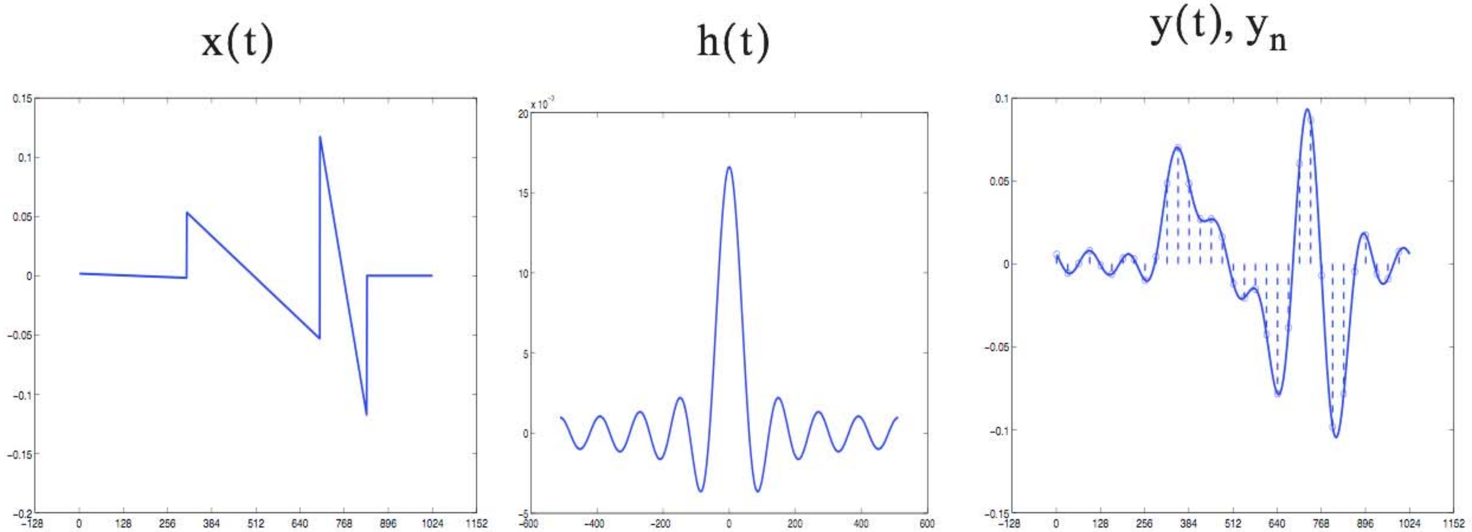
# Generalizations [VMB:02]

**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**

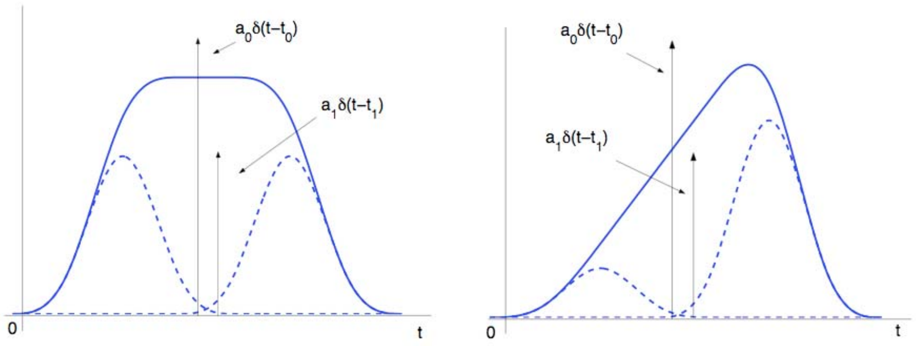


# Generalizations [DVB:07]

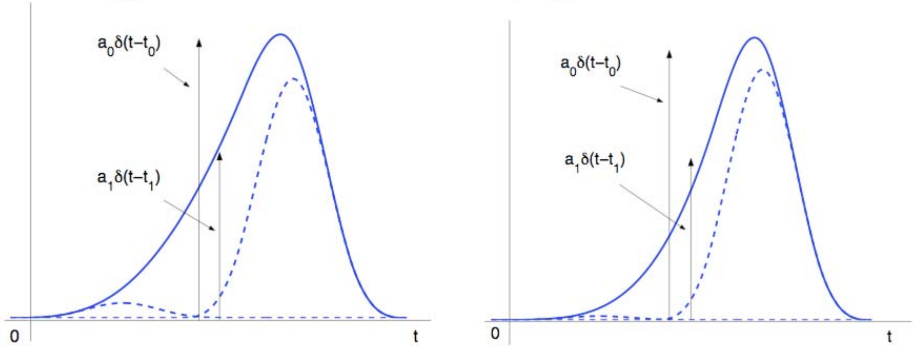
**Strang-Fix condition on sampling kernel:**

Local, polynomial complexity reconstruction, for diracs and piecewise polynomials

Pure discrete-time processing!

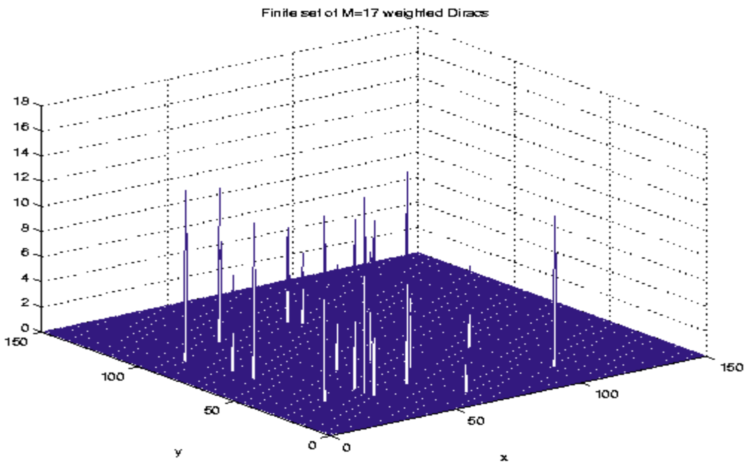


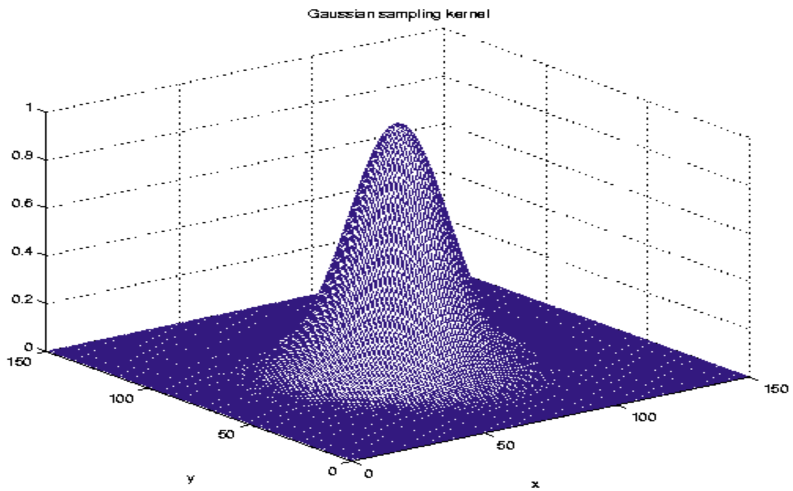

$\sum_n y_n = a_0 + a_1$ $\sum_n n y_n = a_0 t_0 + a_1 t_1$

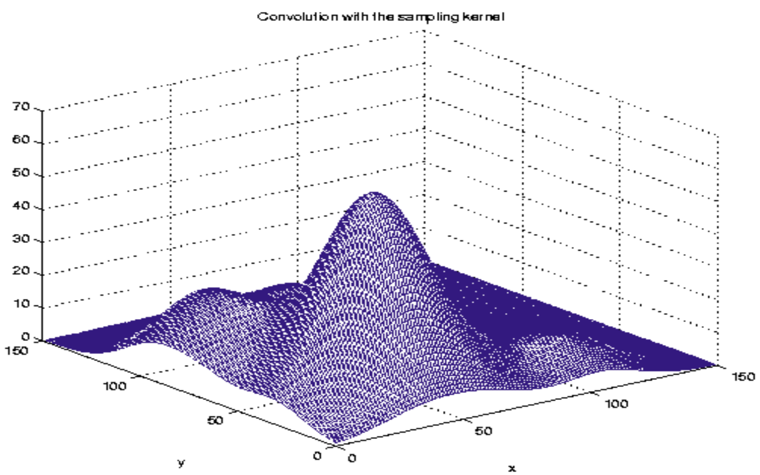


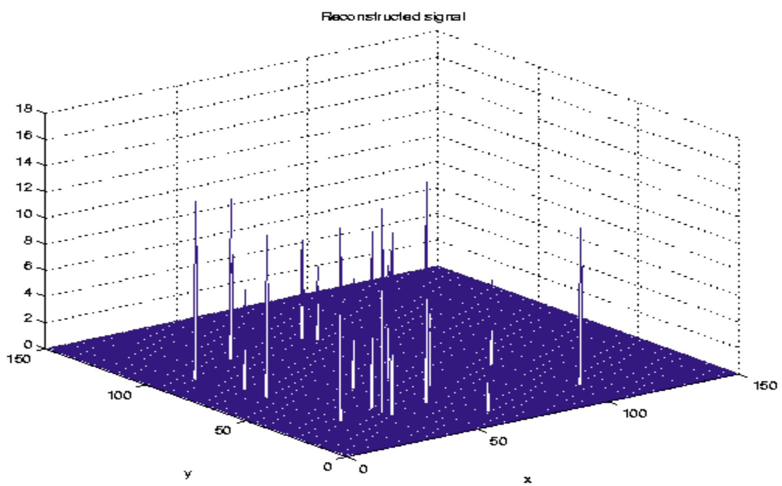

$\sum_n n^2 y_n = a_0 t_0^2 + a_1 t_1^2$ $\sum_n n^3 y_n = a_0 t_0^3 + a_1 t_1^3$

# Generalizations 2D extensions [MV:06]



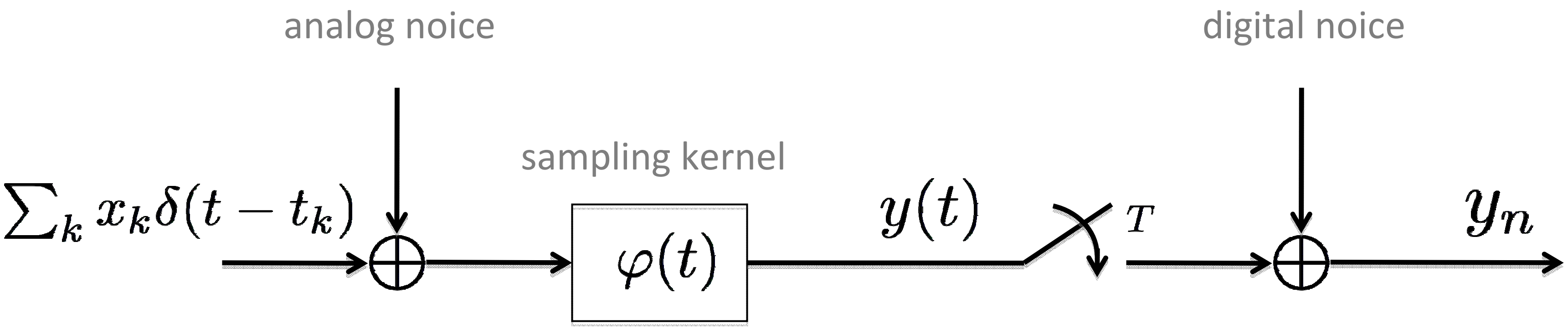




Note: true 2D processing!

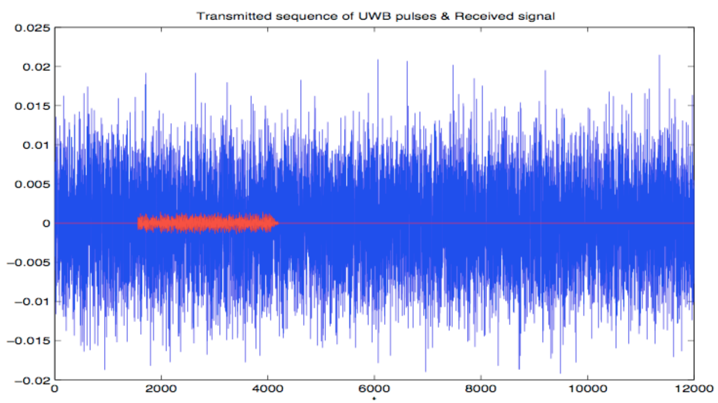
# The real, noisy case…

**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: $$\text{Minimize}\|AH\|^2 \quad \text{under constraint } \|H\|^2 = 1$$

using SVD of a rectangular system

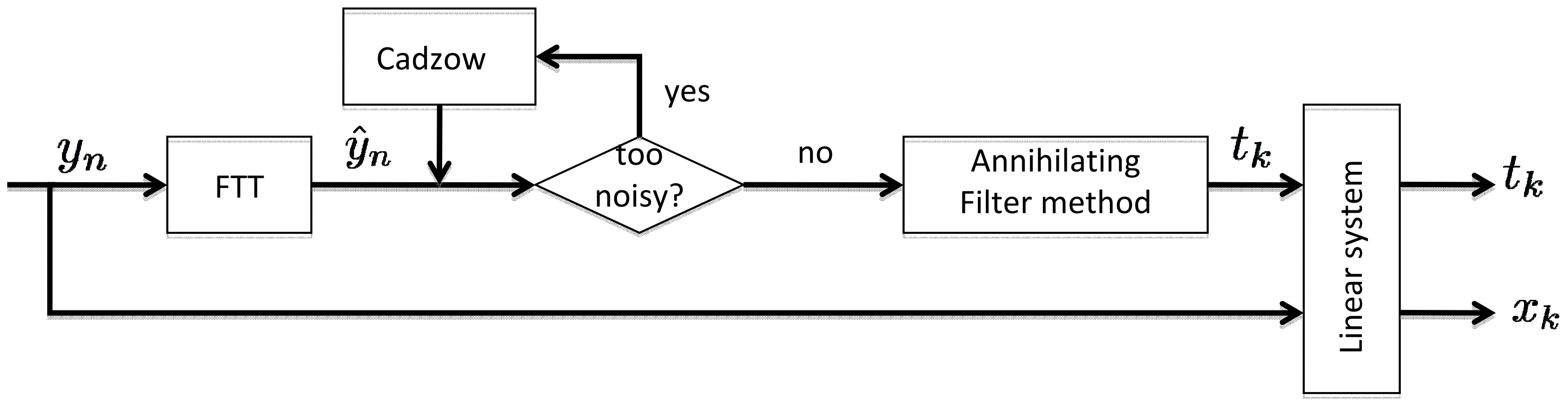
**Cadzow denoising:**

When very noisy: TLS can fail…

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

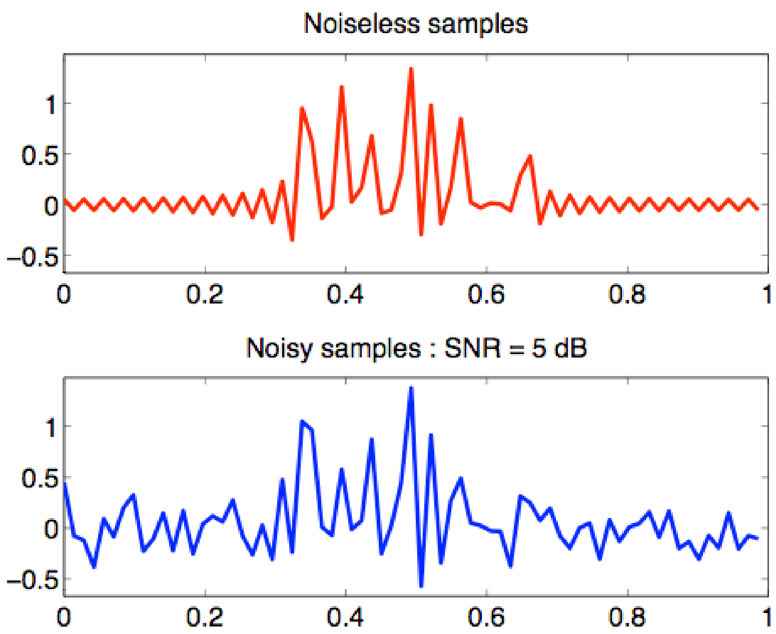
Iterate between the two conditions

**Algorithm:**

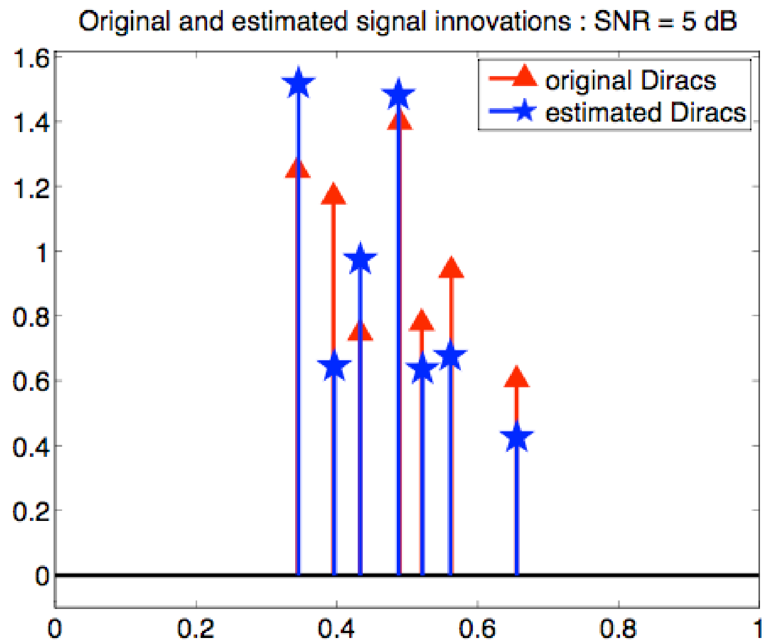


# Example

7 Diracs in 5dB SNR, 71 samples



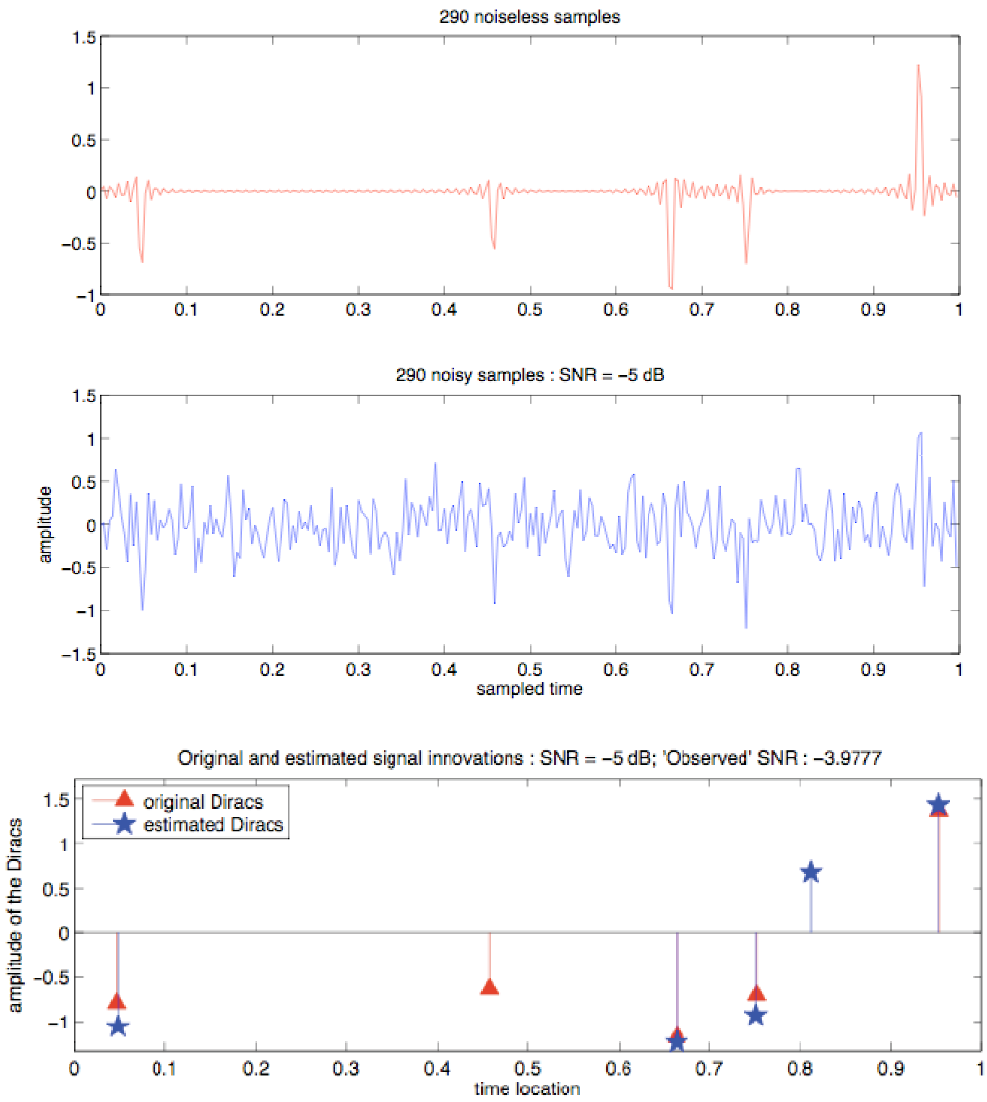

**Original and noisy version**




**Original and retrieved Diracs**

# The real world: -5dB Experiment

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



# Compressed sensing

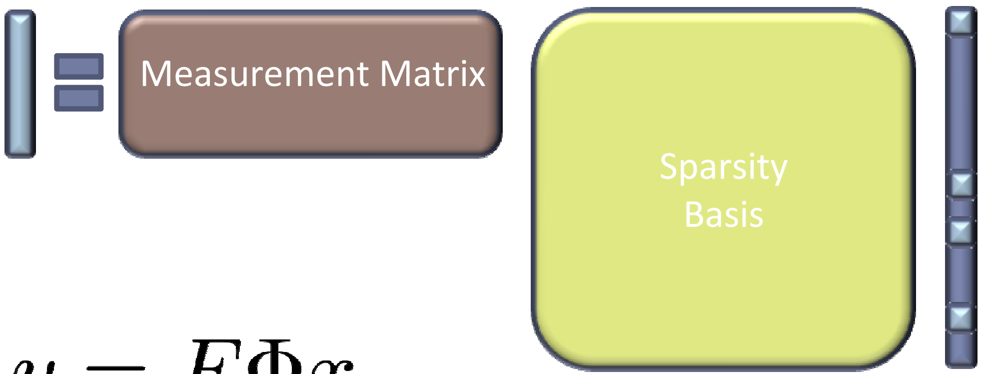
**Consider a discrete-time, finite dimensional set up in $R^N$:**

**Model:**

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

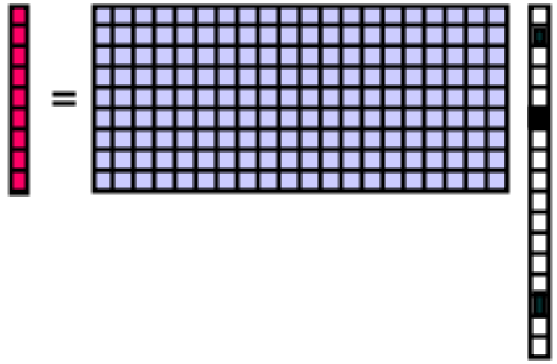
**Method:**

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




$$y = F\Phi x$$

# Geometry of the problem



**Vastly under-determined system…**

**Infinite number of solutions…**

- F is a frame matrix, of size M by N, M << N
- Each set of K possible entries defines M by K mat
- 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

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

**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:** $x \in R^N$, but $|x|_0 = K << N$ : 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

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

**Reconstruction Method:**

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

**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

**The real game:**

In the space of frame measurements matrices F

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

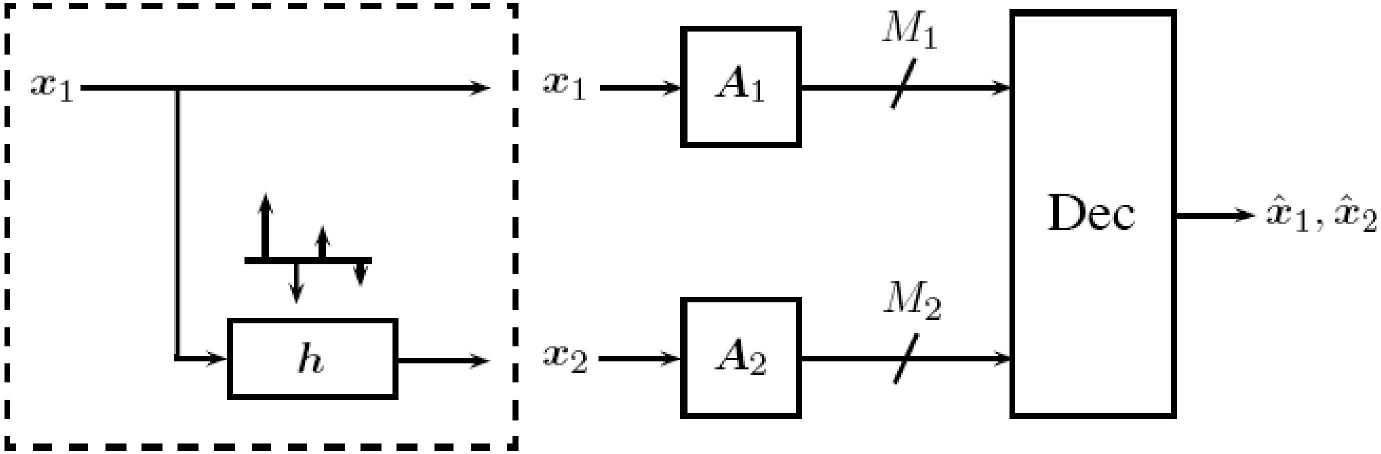
# Outline

1. **Introduction**
2. **Sampling: The linear case**
3. **Application: Sampling physics**
4. **Sampling: The non-linear case**
5. **Applications: The non-linear case**
   - Joint sparsity estimation in distributed settings
   - Multichannel sampling
   - Super-resolution imaging
   - Other applications
6. **Conclusions**

# 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**

# Applications: Joint Sparsity Estimation

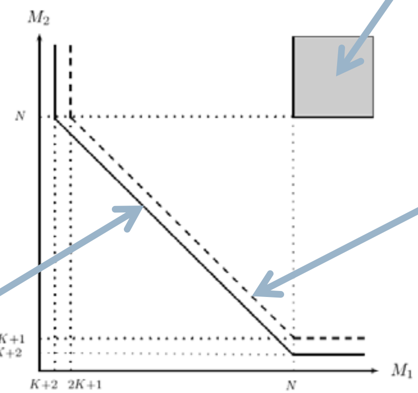
**Sampling rate region:**

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

$$\begin{aligned} M_1 &\geq \min\{K+r, N\} \\ M_2 &\geq \min\{K+r, N\} \\ M_1 + M_2 &\geq \min\{N+K+r, 2N\} \end{aligned}$$

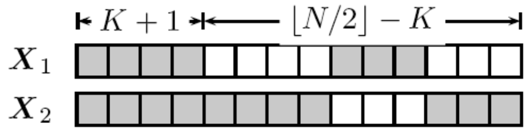
$$\begin{aligned} M_1 &\geq \min\{2K+1, N\} \\ M_2 &\geq \min\{2K+1, N\} \\ M_1 + M_2 &\geq \min\{N+2K+1, 2N\} \end{aligned}$$

Universal recovery




Annihilating filter method

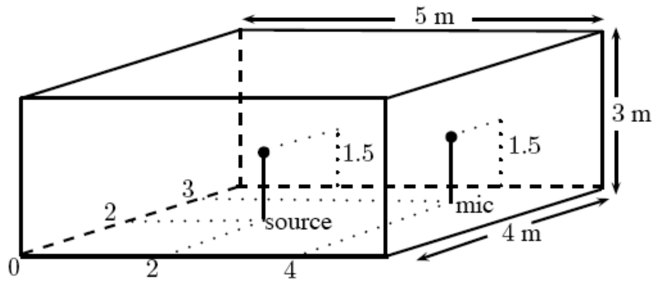
Almost sure recovery

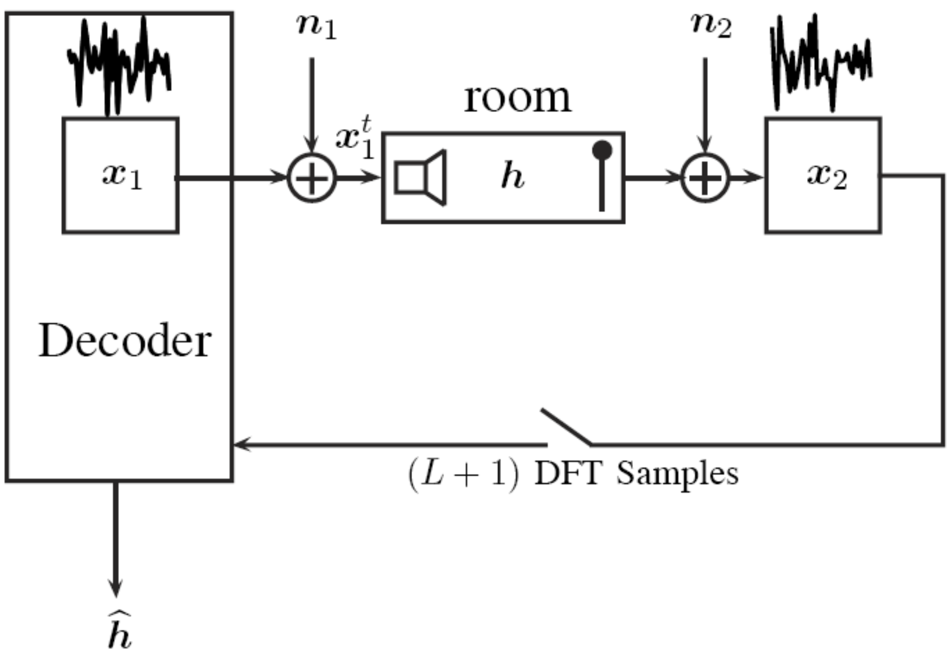


# Applications: Joint Sparsity Estimation

**Ranging, room impulse response, UWB:**

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

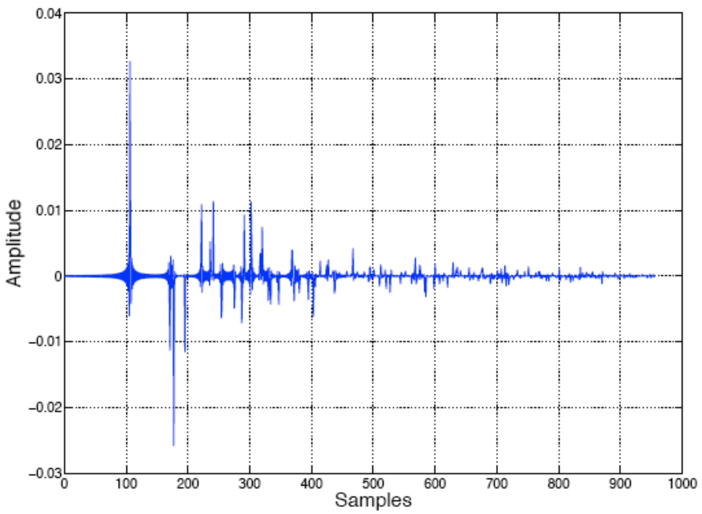


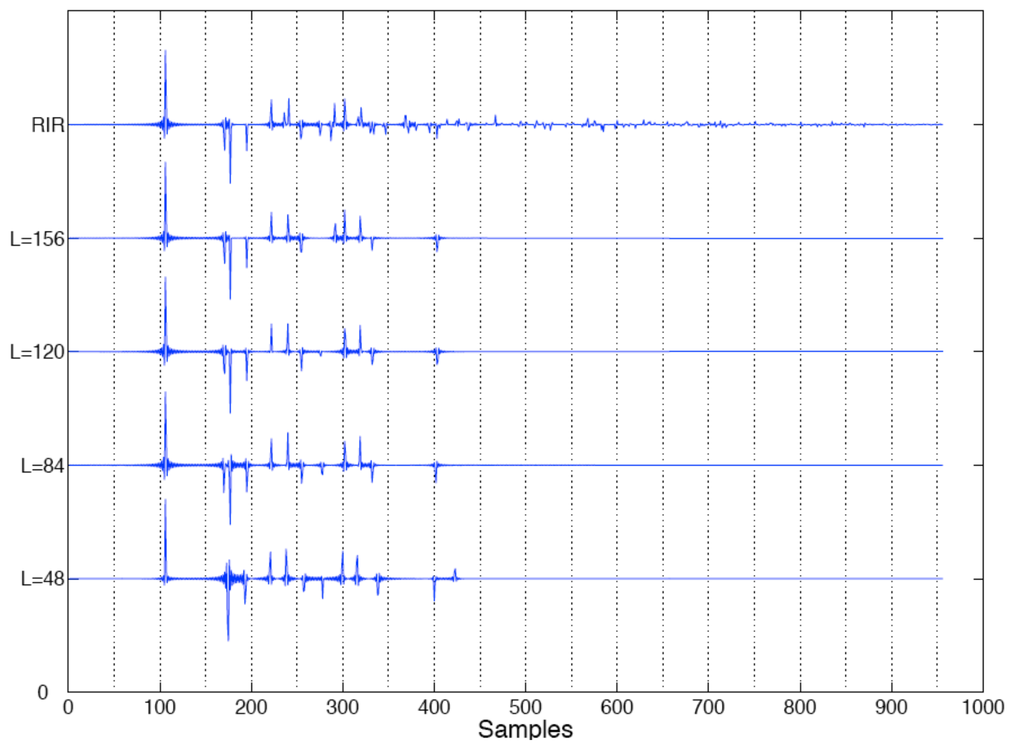



# Applications: Joint Sparsity Estimation

**Experiment**

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



# 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


# 1. Registration and Reconstruction

$\Delta x', \Delta y'$

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

# 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

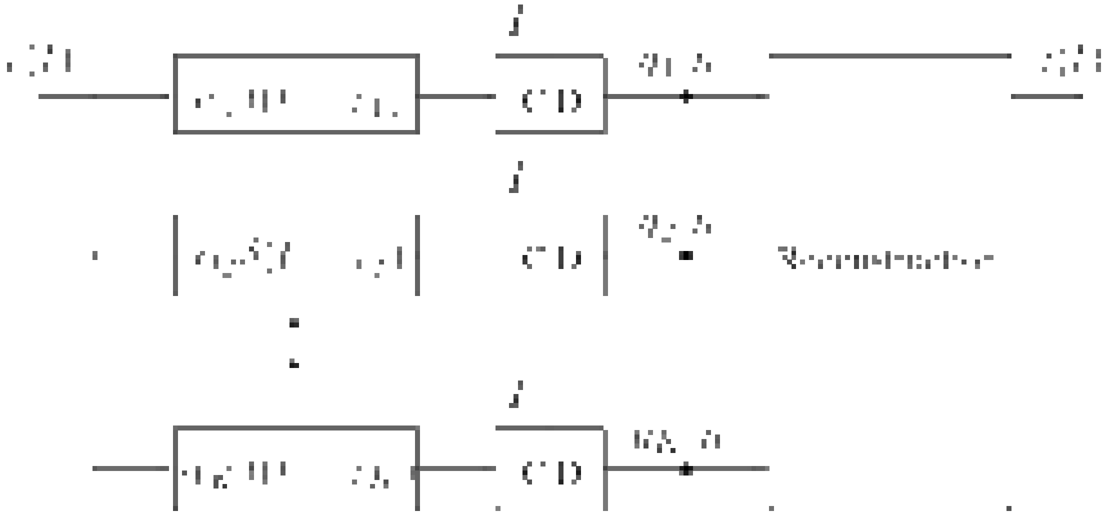
**2. Reconstruction**

Is a linear problem

Solution lives on a subspace

It amounts to solving a linear system of equations

# 2. Multichannel Sampling

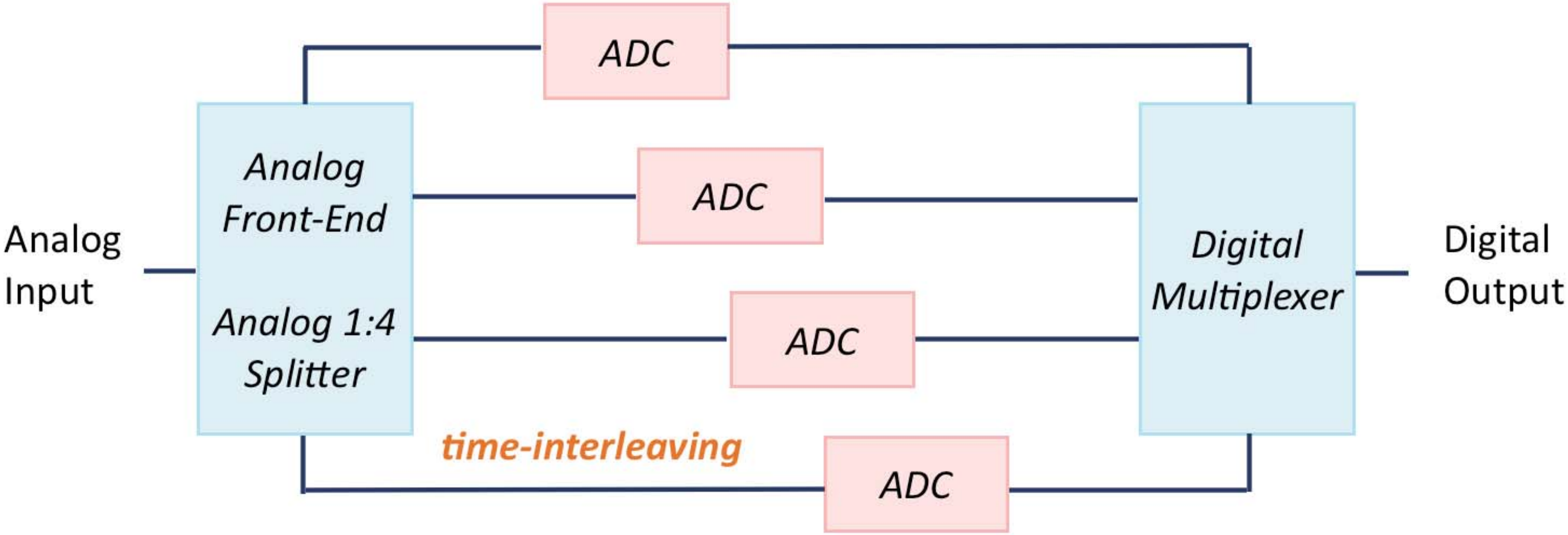


- Input: $x(t)$ bandlimited to $[-\sigma, \sigma]$

- ***Unknown*** gains $\{\alpha_k\}$ and offsets $\{\tau_k\}$

- Sub-Nyquist sampling: $\frac{1}{T} < \frac{\sigma}{\pi}$ ***aliasing!***

Ex: Emulate a single high-rate ADC by an array of low-rate ADCs

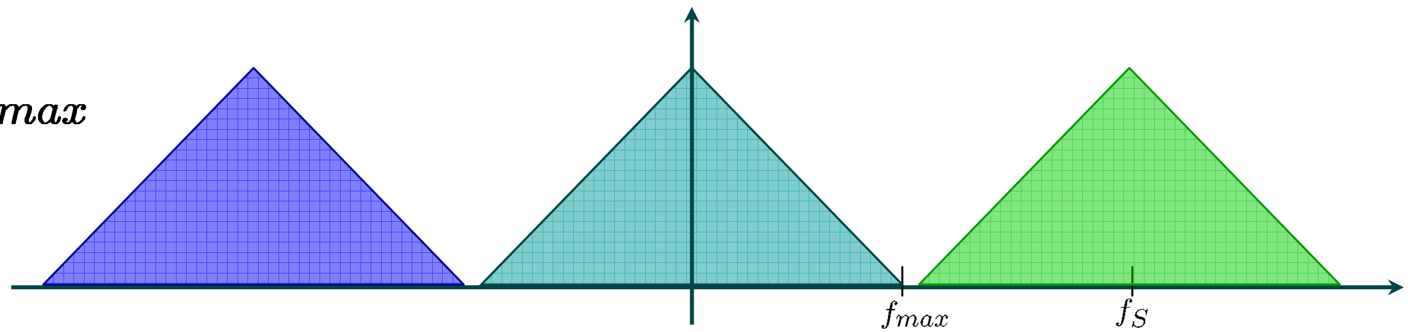



- Ideal setup: uniform channel gains and $\tau_k = \frac{k-1}{K}T$
- In practice: nonuniform gains and ***timing skews*** $\tau_k = \frac{k-1}{K}T + \delta_k$

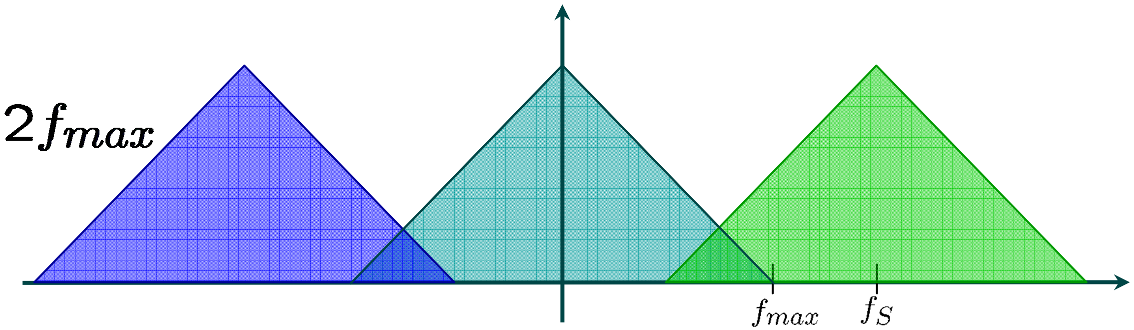
# Subsampling and Aliasing

Depending on sampling rate, 3 different cases

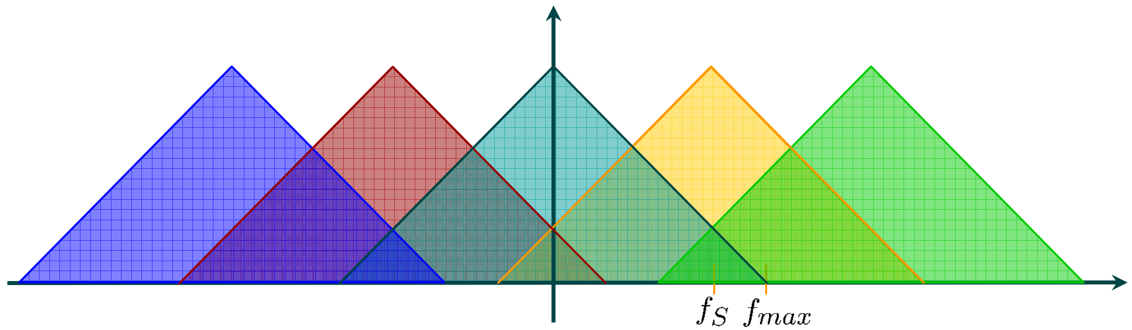
- $f_S > 2f_{max}$




- $f_{max} < f_S < 2f_{max}$




- $f_S < f_{max}$



# Related work: Lots!

***Known gains and offsets:***

- Periodic nonuniform sampling:
  - ***[Yen: 56]***
  - [Sankur & Gerhardt: 73]
  - [Seidner & Feder: 00]
    stability and noise amplification
  - [Bresler et al: 03]
    filter bank reconstruction
  - many others ...
- Generalized sampling expansions
  - ***[Papoulis: 77]***
  - [Unser & Zerubia: 98]

***Linear***

***Unknown gains and offsets:***

- Time-Interleaved ADC corrections:
  - ***[Black & Hodges: 80]***
  - [Fu et al: 98]
  - [Jamal et al: 04]
  - [Huang & Levy: 07]
  - [Lim et al: 09]
  - [McNeill et al: 09]
- Image registration and superresolution
  - [Baker & Kanade: 00]
  - [Elad, Milanfar et al: 04]
  - [Vandewalle et al: 07]
  - many others

***Nonlinear!***

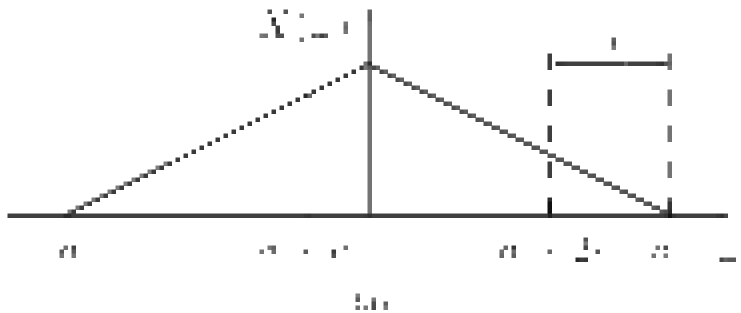
# 3. Rank Condition

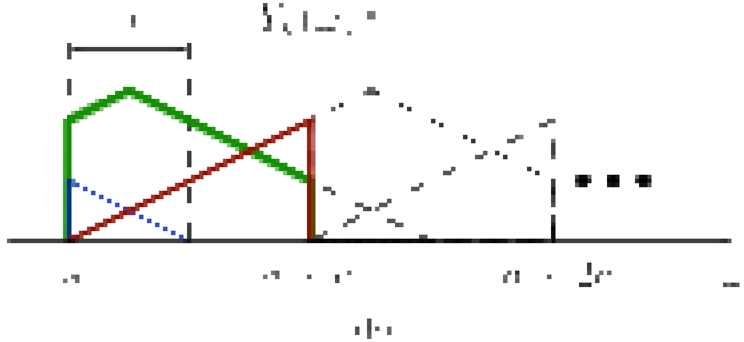
***Time:***

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

***Frequency:***

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



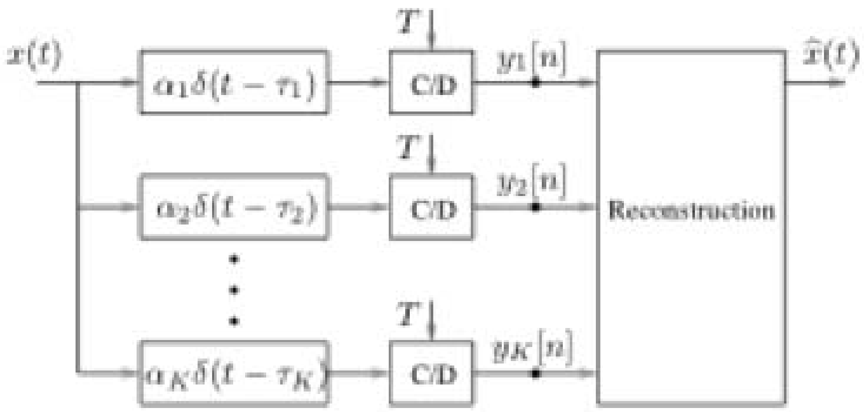


Discrete-time FT: $Y_k(\omega)$

- Periodic

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

# 3. Rank Condition (cont.)




System parameters:

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

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

In the Fourier domain:

$$\boldsymbol{Y}(\omega) = \boldsymbol{\Lambda}_{\boldsymbol{\tau}}(\omega)\, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}}\, \boldsymbol{V}_{\boldsymbol{\tau}}\, \boldsymbol{X}(\omega)$$

- channel output: $\boldsymbol{Y}(\omega)$
- diagonal matrix depending on channel offsets: $\boldsymbol{\Lambda}_{\boldsymbol{\tau}}(\omega)$
- diagonal matrix depending on channel gains: $\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$
- Vandermonde matrix depending on channel offsets: $\boldsymbol{V}_{\boldsymbol{\tau}}$
- unknown input: $\boldsymbol{X}(\omega)$

# Sampling and reconstruction is possible

With ***unknown*** channel gains and offsets:

Ambiguities: $\{x(t), \alpha_k, \tau_k\}$ vs $\left\{\alpha x(t-\tau), \frac{\alpha_k}{\alpha}, \tau_k + \tau\right\}$

Set: $\alpha_1 = 1,\ \tau_1 = 0$

***Proposition:*** Almost all input $x(t)$ can be uniquely determined by its channel samples if and only if the channel sampling rate satisfies

$$\frac{1}{T} > \frac{\sigma}{K\pi}.$$

***Oversampling*** to resolve parameter uncertainties

# Rank condition

The forward model:

$$\boldsymbol{Y}(\omega) = \boldsymbol{\Lambda}_{\boldsymbol{\tau}}(\omega)\,\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}\,\boldsymbol{V}_{\boldsymbol{\tau}}\,\boldsymbol{X}(\omega)$$

move $\boldsymbol{\Lambda}_{\boldsymbol{\tau}}(\omega)$ to the left

$$\boldsymbol{\Lambda}_{\boldsymbol{\tau}}^{*}(\omega)\boldsymbol{Y}(\omega) = \boldsymbol{\Lambda}_{\boldsymbol{\alpha}}\,\boldsymbol{V}_{\boldsymbol{\tau}}\,\boldsymbol{X}(\omega) \quad \Longrightarrow \quad \boldsymbol{\Lambda}_{\boldsymbol{\tau}}^{*}(\omega)\boldsymbol{Y}(\omega) \in \mathcal{R}\left(\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}\boldsymbol{V}_{\boldsymbol{\tau}}\right)$$

***Key observations:***

1. $\mathcal{R}\left(\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}\boldsymbol{V}_{\boldsymbol{\tau}}\right)$ has ***co-dimension*** one

2. $\boldsymbol{D}_{\boldsymbol{\tau}} \stackrel{\text{def}}{=} \left[\boldsymbol{\Lambda}_{\boldsymbol{\tau}}^{*}(\omega_1)\boldsymbol{Y}(\omega_1), \boldsymbol{\Lambda}_{\boldsymbol{\tau}}^{*}(\omega_2)\boldsymbol{Y}(\omega_2), \ldots, \boldsymbol{\Lambda}_{\boldsymbol{\tau}}^{*}(\omega_N)\boldsymbol{Y}(\omega_N)\right]$ is ***rank-deficient***

# Linearization

Example: $K = 3$ (three channels)

***Theorem:*** If the data matrix $\boldsymbol{D}_{\boldsymbol{\tau}}$ is rank deficient, then

$$\begin{aligned}(Y_{1,n+2}Y_{2,n+1}Y_{3,n})\,u - (Y_{1,n+2}Y_{2,n}Y_{3,n+1})\,v \\ - (Y_{1,n+1}Y_{2,n+2}Y_{3,n})\,u^2 + (Y_{1,n}Y_{2,n+2}Y_{3,n+1})\,u^2 v \\ + (Y_{1,n+1}Y_{2,n}Y_{3,n+2})\,v^2 - (Y_{1,n}Y_{2,n+1}Y_{3,n+2})\,uv^2 = 0,\end{aligned}$$

where $Y_{k,n} \stackrel{\text{def}}{=} Y_k(n\triangle\omega)$, $u \stackrel{\text{def}}{=} e^{j\triangle\omega\tau_1}$, $v \stackrel{\text{def}}{=} e^{j\triangle\omega\tau_2}$.

***Key observations:***

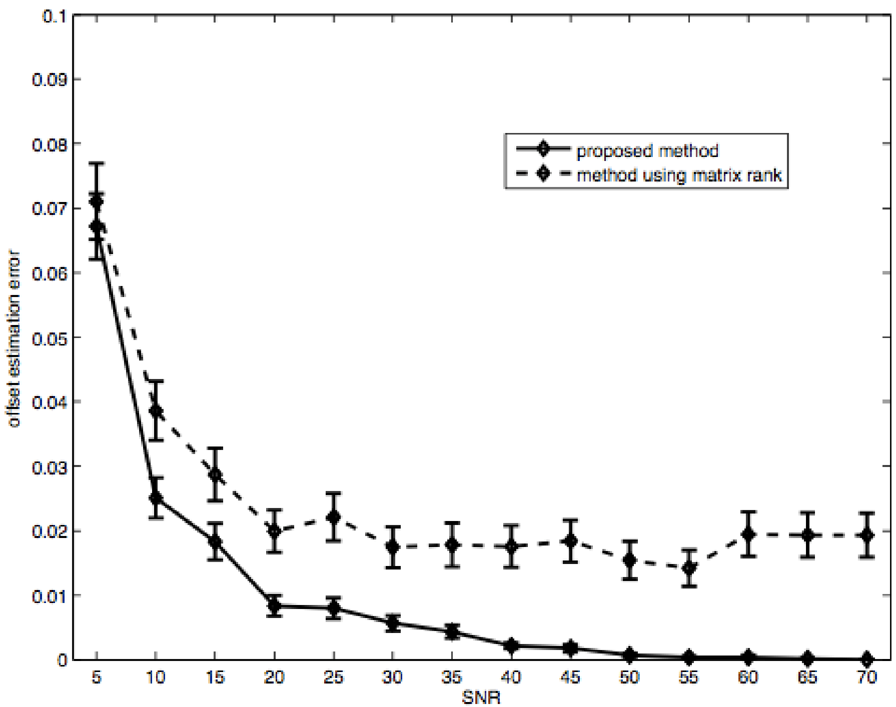
1. A system of multivariate polynomial equations (of $u, v$). ***Hard to solve***!
2. ***Linearization***: treat $u, v, u^2, u^2v, v^2, uv^2$ as if they were independent

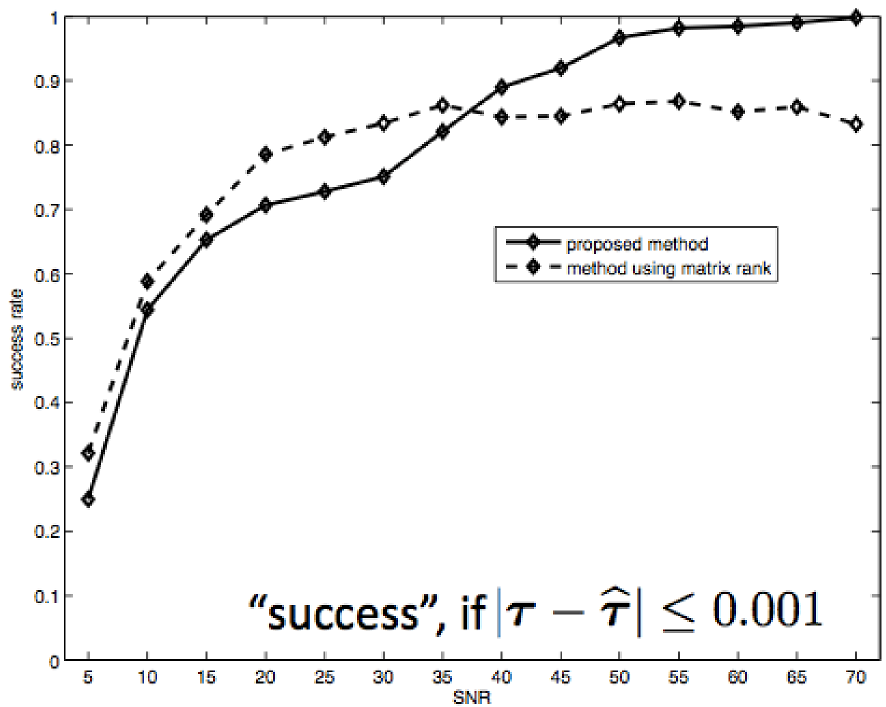
# Example

***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

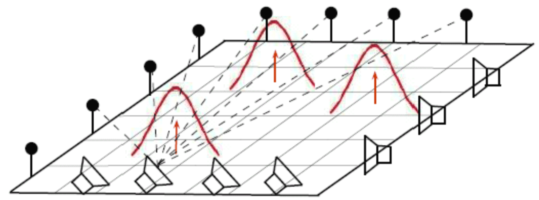


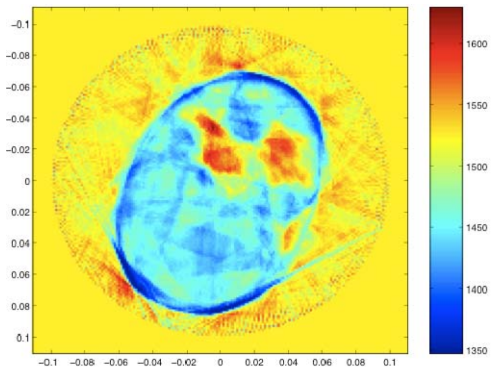
# Other examples: Inverse problems regularized by sparsity

**Breast cancer detection**

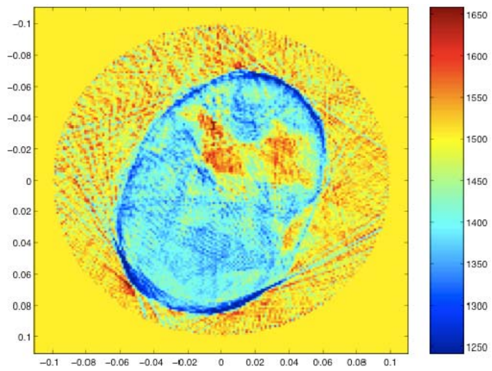
- Ultrasound transmission tomography
- Non-linear inverse problem
- Sparsity prior
- Initial results promising



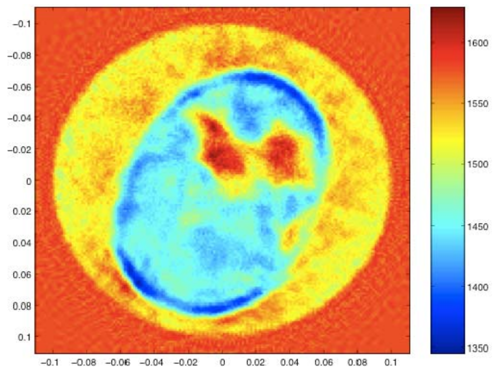
$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \quad \frac{1}{2}\|\boldsymbol{y} - \Phi \boldsymbol{x}\|_2^2 + \alpha\|\Psi \boldsymbol{x}\|_1 + \beta\,\mathrm{TV}(\boldsymbol{x})$$



Conjugate Gradient
30000 measurements



Conjugate Gradient
10000 measurements



Compressed Sensing
10000 measurements

# Conclusions: Sampling is Alive and Well!

**Sampling of sparse signals**

- Sharp theorems
- Robust algorithms
- Provable optimality over wide SNR ranges

**Many actual and potential applications**

- Fit the model (needs some work)
- Apply the ''right'' algorithm
- Catch the essence!

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

- Non-linear problems (F is part of the problem)
- Regularization of the inverse problem
- Faster algorithms
- Designer matrices

# Publications

**Special issue on Compressive Sampling:**

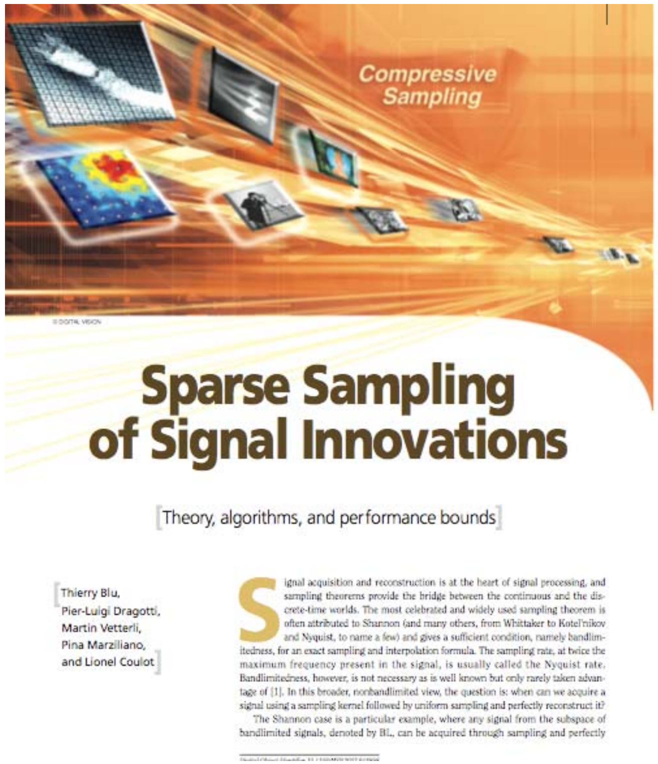
- R.Baraniuk, E.Candes, R.Nowak and M.Vetterli (Eds.)IEEE Signal Processing Magazine, March 2008. 10 papers overviewing the theory and practice of Sparse Sampling, Compressed Sensing and Compressive Sampling (CS).

**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.



# Publications

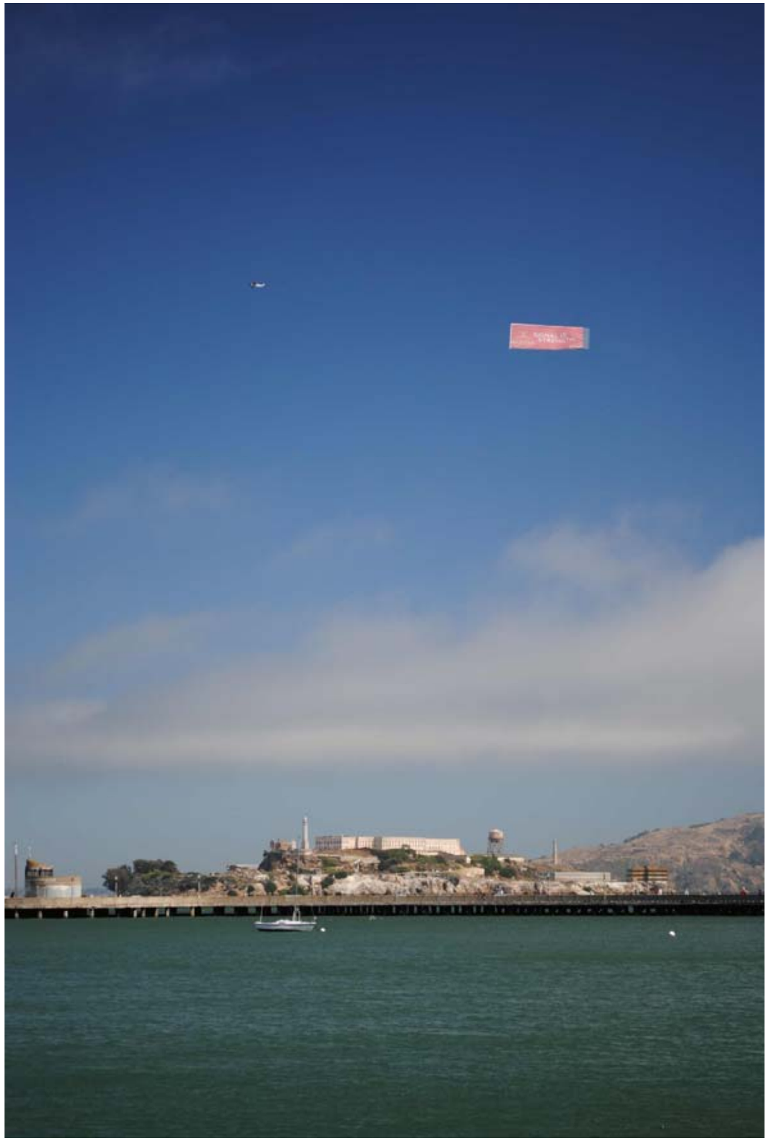
**For more details:**

– 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.

# Thank you for your attention !

Lausanne, November 6th, 2009

Any questions ?




SIGNAL IS
STRENGTH.


SIGNAL IS
STRENGTH.