Sampling Theory and Practice: 50 Ways to Sample your Signal

Martin Vetterli  EPFL & UC Berkeley
Acknowledgements

Sponsors and supporters
– NSF Switzerland: thank you!
– Conference organizers, Akihiko Sugiyama in particular!
– My japanese friends, in particular Prof. Masaaki Ikehara 😊

Co-authors and collaborators
– Thierry Blu, CUHK
– Yue Lu, Harvard
– Amina Chebira, EPFL
– Ali Hormati, EPFL
– Yann Barbotin, EPFL
– Jury Rarienri, 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

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, simulations etc
The situation (bis)

The world is analog, computation is digital, and

- You want an analysis

![Diagram](image)

- You want a synthesis

![Diagram](image)
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 lowpass filter
  – Regular, uniform sampling
  – That is, the workhorse of sampling!

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

\[ x(t) \iff y_n \]
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)
A 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: \( K < M << N \)

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!

![Diagram of multichannel sampling system]
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

---

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

Air Temperature [°C]
- High: 8.00
- Low: 6.00

**bâtiments5**
Diffusion equation and inversion

Point sources
Observation by sensors
Ex: Process over \((x,t)\)
  3 point sources
  4 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:

- $\text{sinc}(t-n)$ 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
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 (super-resolution 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 versus sampling physics

Sampling:

\[ f(t) \in \mathcal{V} \rightarrow \varphi(t) \rightarrow y_m \]

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

Sampling physics:

sources: \[ s(x, t) \in \mathcal{V} \rightarrow \text{PDE} \rightarrow f(x, t) \in \mathcal{W} \rightarrow \varphi(x, t) \rightarrow y_{m,n} \]

- 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?

Analog sources → Wave equation → Space-time acoustic wave field
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
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

- 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

\[ \omega = c \Phi \]

\( \omega \): 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

\[ \phi_N > \omega_0 / c \]
Visual proof 😊

\[ \omega = c \Phi \]

\[ \Phi_S > \frac{2\omega_0}{c} \]
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 bandlimited case

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

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
- Piecewise Polynomial: Woodcut picture
- Stream of Diracs: Poisson process
- 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

\[ \rho : \text{Rate of innovation or degrees of freedom per unit of time} \]

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

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

**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 \( X_m \)

\[
x(t) = \sum_{m \in \mathbb{Z}} X_m e^{-j2\pi mt}
\]

Fourier integral leads to
- \( K \) complex exponentials
- Exponents depends on location
- Weight depends on \( \alpha_k \)

If we can identify the exponents
- Diracs can be recovered
- Weights are a linear problem (given the locations)
A Representation Theorem [VMB:02]

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

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

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

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

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

\[
H(z) = \prod_{k=1}^{K} (1 - u_k z^{-1}). \quad u_k = e^{-j 2\pi t_k / \tau}
\]

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

\[
\hat{x}_m = \frac{1}{\tau} \sum_{k=1}^{K} x_k e^{-j 2\pi m t_k / \tau} = \frac{1}{\tau} \sum_{k=1}^{K} x_k u_k^m.
\]

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:**

analog noise

\[ \sum_k x_k \delta(t - t_k) \]

digital noise

sampling kernel

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

![Algorithm Diagram]

\( y_n \) \( \hat{y}_n \) \( H \) \( t_k \) \( x_k \)
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 $\mathbb{R}^N$:

**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 $M \times N$

$$y = F\Phi x$$
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 $\binom{M}{K}$ possible entries defines M by K matrix $F_k$
  - There are $\binom{N}{K}$ matrices $F_k$
  - Calculate projection of $y$ onto range of $F_k$
    $$\hat{y}_k = F_k (F_k^* F_k)^{-1} F_k^* y$$
    $y_k = y$, possible solution
  - In general, choose k such that
  - 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

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

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, KlogN, Klog^2N
• 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)

\[ M_1 \geq \min\{K + r, N\} \]
\[ M_2 \geq \min\{K + r, N\} \]
\[ M_1 + M_2 \geq \min\{N + K + r, 2N\} \]
\[ M_1 \geq \min\{2K + 1, N\} \]
\[ M_2 \geq \min\{2K + 1, N\} \]
\[ M_1 + M_2 \geq \min\{N + 2K + 1, 2N\} \]
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

\[ \Delta x, \Delta y \]
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

- $f_S > 2f_{\text{max}}$

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

- $f_S < f_{\text{max}}$
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:

- Periodic
- A **finite** number of frequency segments folding on top of each other
3. Rank Condition (cont.)

System parameters:
- unknown gains: \[ \alpha = [\alpha_1, \ldots, \alpha_K]^T \]
- unknown offsets: \[ \tau = [\tau_1, \ldots, \tau_K]^T \]

In the Fourier domain:
- channel output
- diagonal matrix depending on channel gains
- diagonal matrix depending on channel offsets
- Vandermonde matrix depending on channel offsets

\[
Y(\omega) = \Lambda_\tau(\omega) \Lambda_\alpha V_\tau X(\omega)
\]
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
From Theorems to Patents to Products...

Theorem 1 -> Patent 1

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

Real technology transfer!
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
Publications

Special issue on Compressive Sampling:

Basic paper:

Main paper, with comprehensive review:
Publications

For more details:


Thank you for your attention!

Any questions?