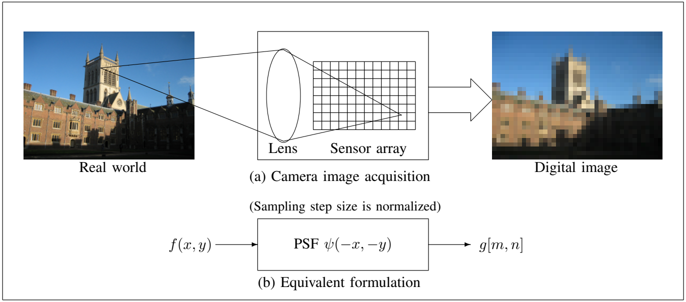
Parametric Sparse Sampling: Theory and Applications

Pier Luigi Dragotti

June 19, 2014[1]

[1] This research is supported by European Research Council ERC, project 277800 (RecoSamp)

Sparsity and Sampling: Is This Relevant?



- The lens blurs the image.
- The image is sampled ('pixelized') by the CCD array.
- You want sharper and higher resolution images given the available pixels

Motivation: Image Resolution Enhancement

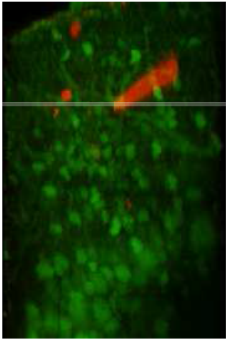
pixels

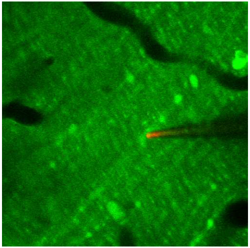
interpolation

enhancement with sparsity priors

Motivation: Application in Neuroscience

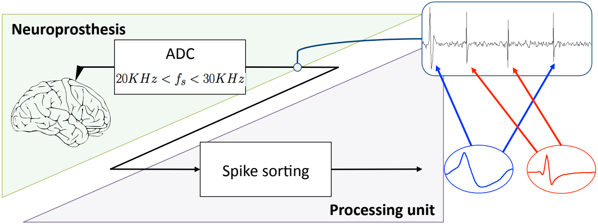
Time resolution enhancement and calcium transient detection in multi-photon calcium imaging.





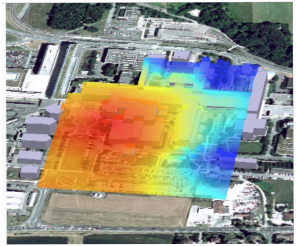
Motivation: Brain Machine Interface

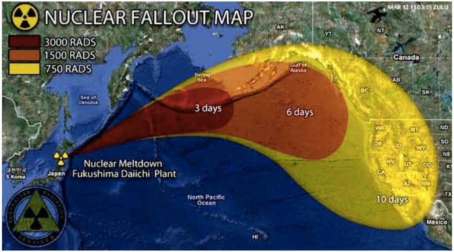
Applications in Neuroscience: Spike Sorting at sub-Nyquist rates



Wireless brain-machine interface place extreme limits on sampling.

Motivation: Sensor Networks





- Can we localise diffusion sources and estimate their activation time using sensor networks?
- Application:
  1. Check whether your government is lying ;-)
  2. Monitor dispersion in factories producing bio-chemicals

Problem Statement

What do all these problems have in common?

- The source in normally continuous in time and/or space (discretising it might not be an effective strategy)
- Measurements are discrete (e.g., pixels in a camera, sensors measurements)
- The observation process involves deterministic smoothing functions normally known a priori (e.g., point spread function in a camera, the diffusion kernel for diffusion fields)

Problem Statement

What do all these problems have in common?

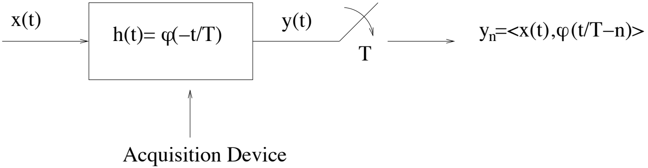
- The source in normally continuous in time and/or space (discretising it might not be an effective strategy)
- Measurements are discrete (e.g., pixels in a camera, sensors measurements)
- The observation process involves deterministic smoothing functions normally known a priori (e.g., point spread function in a camera, the diffusion kernel for diffusion fields)

Our Approach

- From the samples, using the knowledge of the observation process, estimate proper integral measurements of the source (e.g., estimate the Fourier transform at specific frequencies)
- Given the integral measurements (e.g., partial Fourier transform), solve the inverse problem using sparsity priors

Problem Statement

You are given a class of functions. You have a sampling device. Given the measurements $y_n = \langle x(t), \varphi(t/T - n)\rangle$, you want to reconstruct $x(t)$.



Natural questions:

- When is there a one-to-one mapping between $x(t)$ and y_n?
- What signals can be sampled and what kernels $\varphi(t)$ can be used?
- What reconstruction algorithm?

Signals with Finite Rate of Innovation

Consider a signal of the form:

$$x(t) = \sum_{k \in \mathbb{Z}} \gamma_k g(t - t_k). \qquad (1)$$

- Given $g(t)$, the signal is completely specified by γ_k and t_k.
- **Key intuition**: if the number of samples is larger than the number of parameters then reconstruction is possible
- This is an 'analogue' sparsity model

Imperial College London

Signals with Finite Rate of Innovation

Consider a signal of the form:

$$x(t) = \sum_{k \in \mathbb{Z}} \gamma_k g(t - t_k). \tag{2}$$

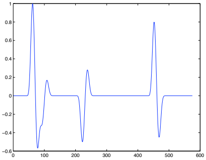
The rate of innovation of $x(t)$ is then defined as

$$\rho = \lim_{\tau \to \infty} \frac{1}{\tau} C_x \left(-\frac{\tau}{2}, \frac{\tau}{2} \right), \tag{3}$$

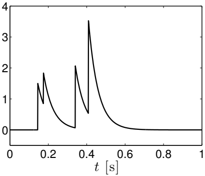
where $C_x(-\tau/2, \tau/2)$ is a function counting the number of free parameters in the interval τ.

Definition [VetterliMB:02] A signal with a **finite rate of innovation** is a signal whose parametric representation is given in (2) and with a finite ρ as defined in (3).

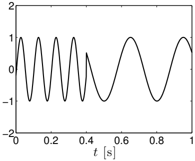
Examples of Signals with Finite Rate of Innovation



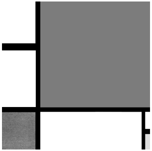
Filtered Streams of Diracs



Decaying Exponentials

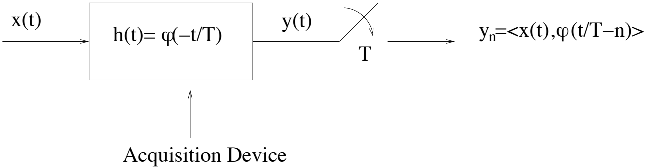


Piecewise Sinusoidal Signals



Mondrian paintings ;-)

Sampling Kernels



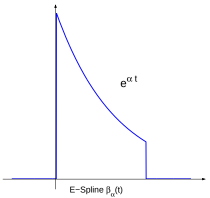
- Given by nature
  - Diffusion equation, Green function. Ex: sensor networks.
- Given by the set-up
  - Designed by somebody else. Ex: Hubble telescope, digital cameras.
- Given by design
  - Pick the best kernel. Ex: engineered systems.

Sampling Kernels

Any kernel $\varphi(t)$ that can reproduce exponentials:

$$\sum_n c_{m,n}\varphi(t-n) = e^{\alpha_m t}, \qquad \alpha_m = \alpha_0 + m\lambda \text{ and } m = 0, 1, ..., L.$$

This includes any composite kernel of the form $\gamma(t) * \beta_{\vec{\alpha}}(t)$ where $\beta_{\vec{\alpha}}(t) = \beta_{\alpha_0}(t) * \beta_{\alpha_1}(t) * ... * \beta_{\alpha_L}(t)$ and $\beta_{\alpha_i}(t)$ is an Exponential Spline of first order [UnserB:05].

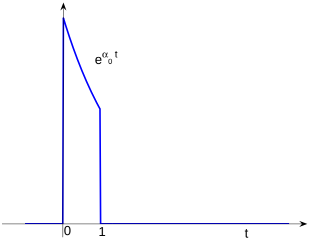


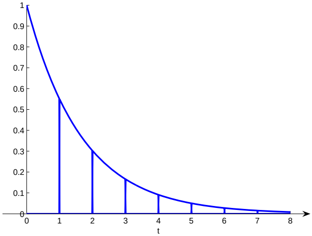
$$\beta_\alpha(t) \Leftrightarrow \hat{\beta}(\omega) = \frac{1-e^{\alpha - j\omega}}{j\omega - \alpha}$$

Notice:

- α can be complex.
- E-Spline is of compact support.
- E-Spline reduces to the classical polynomial spline when $\alpha = 0$.

Exponential Reproducing Kernels



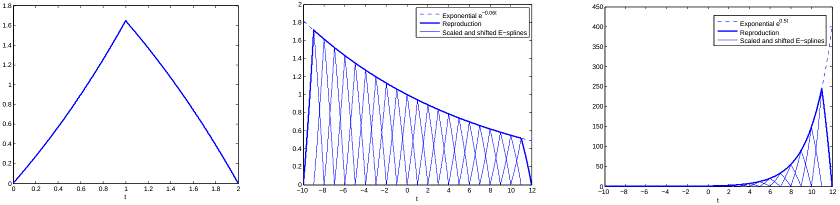


The E-spline of first order $\beta_{\alpha_0}(t)$ reproduces the exponential $e^{\alpha_0 t}$:

$$\sum_n c_{0,n}\beta_{\alpha_0}(t-n) = e^{\alpha_0 t}.$$

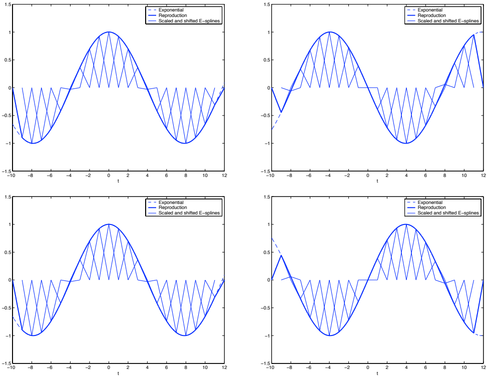
In this case $c_{0,n} = e^{\alpha_0 n}$. In general, $c_{m,n} = c_{m,0}e^{\alpha_m n}$.

Exponential Reproducing Kernels



Here the E-spline is of second order and reproduces the exponential $e^{\alpha_0 t}$, $e^{\alpha_1 t}$: with $\alpha_0 = -0.06$ and $\alpha_1 = 0.5$.

Exponential Reproducing Kernels



Here $\vec{\alpha} = (-j\omega_0, j\omega_0)$ and $\omega_0 = 0.2$. $\sum_n c_{n,m}\beta_{\vec{\alpha}}(t-n) = e^{jm\omega_0} \qquad m = -1, 1.$

Notice: $\beta_{\vec{\alpha}}(t)$ is a real function, but the coefficients $c_{m,n}$ are complex.

Generalised Strang-Fix Conditions

A function $\varphi(t)$ can reproduce the exponential:

$$e^{j\omega_m t} = \sum_n c_{m,n}\varphi(t-n)$$

if and only if

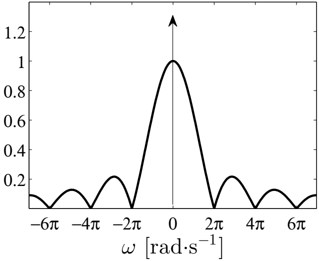
$$\hat{\varphi}(j\omega_m) \neq 0 \text{ and } \hat{\varphi}(j\omega_m + j2\pi l) = 0 \qquad l \in \mathbb{Z}\setminus\{0\}$$

where $\hat{\varphi}(\cdot)$ is the Fourier transform of $\varphi(t)$.

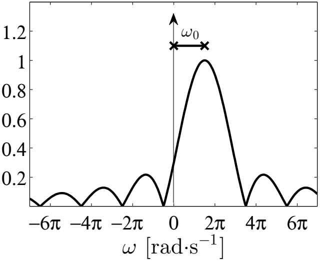
Also note that $c_{m,n} = c_{m,0}e^{j\omega_m n}$ with $c_{m,0} = \hat{\varphi}(j\omega_m)^{-1}$.

Generalised Strang-Fix Conditions

- Strang-Fix conditions are not restrictive
- Any low-pass filter approximately satisfies them.



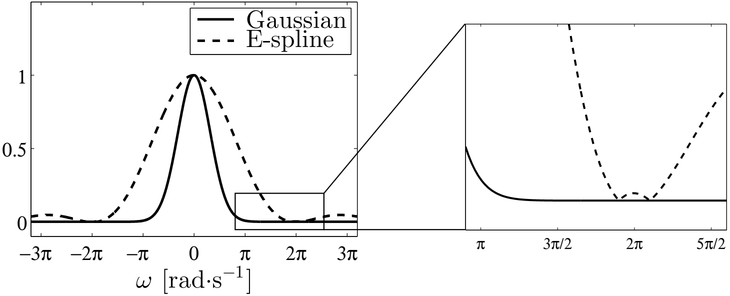
(a) $|\hat{\beta}_\alpha(\omega)|$ with $\alpha = 0$



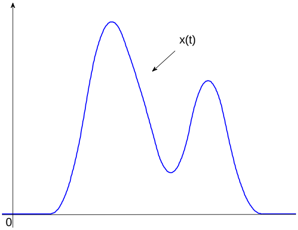
(b) $|\hat{\beta}_\alpha(\omega)|$ with $\alpha = i\omega_0$

Approximate Strang-Fix

- Strang-Fix conditions are not restrictive
- Any low-pass filter approximately satisfies them.



From Samples to Integral Measurements



- Consider any $x(t)$ with $t \in [0, N)$ and sampling period $T = 1$.
- The sampling kernel $\varphi(t)$ satisfies

$$\sum_n c_{m,n}\varphi(t-n) = e^{j\omega_m t} \qquad m = 1, ..., L,$$

- We want to retrieve $x(t)$, from the samples $y_n = \langle x(t), \varphi(t-n)\rangle$, $n = 0, 1, ..., N-1$.

From Samples to Integral Measurements

We have that

$$\begin{aligned} s_m &= \sum_{n=0}^{N-1} c_{m,n} y_n \\ &= \langle x(t), \sum_{n=0}^{N-1} c_{m,n} \varphi(t-n) \rangle \\ &= \int_{-\infty}^{\infty} x(t) e^{j\omega_m t} dt, \qquad m = 1, ..., L. \end{aligned}$$

- Note that s_m is the Fourier transform of $x(t)$ evaluated at $j\omega_m$.

From Samples to Signals

- Consider FRI signals which are completely specified by a finite number of free parameters
- For classes of parametrically sparse signals there is a one-to-one mapping between samples and signal:

$$x(t) \Leftrightarrow \hat{x}(j\omega_m) \qquad m = 1, 2, ..., L$$

- The number d of degrees of freedom of the signal must satisfy $d \leq L$

Sampling Streams of Diracs

- Assume $x(t)$ is a stream of K Diracs on the interval of size N: $x(t) = \sum_{k=0}^{K-1} x_k \delta(t - t_k)$, $t_k \in [0, N)$.
- We restrict $j\omega_m = j\omega_0 + jm\lambda \quad m = 1, ..., L$ and $L \geq 2K$.
- We have N samples: $y_n = \langle x(t), \varphi(t - n) \rangle$, $n = 0, 1, ...N - 1$:
- We obtain

$$
\begin{aligned}
s_m &= \sum_{n=0}^{N-1} c_{m,n} y_n \\
&= \int_{-\infty}^{\infty} x(t) e^{j\omega_m t} dt, \\
&= \sum_{k=0}^{K-1} x_k e^{j\omega_m t_k} \\
&= \sum_{k=0}^{K-1} \hat{x}_k e^{j\lambda m t_k} = \sum_{k=0}^{K-1} \hat{x}_k u_k^m, \qquad m = 1, ..., L.
\end{aligned}
$$

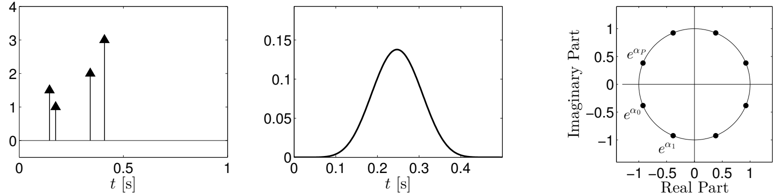
Prony's Method

- The quantity

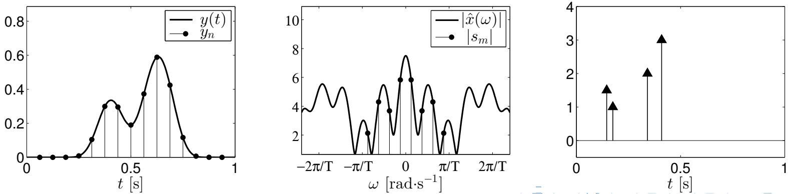
$$s_m = \sum_{k=0}^{K-1} \hat{x}_k u_k^m, \qquad m = 1, ..., L$$

  is a sum of exponentials.
- Retrieving the locations u_k and the amplitudes $\hat{x}_k$ from $\{s_m\}_{m=1}^{L}$ is a classical problem in spectral estimation and was first solved by Gaspard de Prony in 1795.
- Given the pairs $\{u_k, \hat{x}_k\}$, then $t_k = (\ln u_k)/\lambda$ and $x_k = \hat{x}_k / e^{\alpha_0 t_k}$.

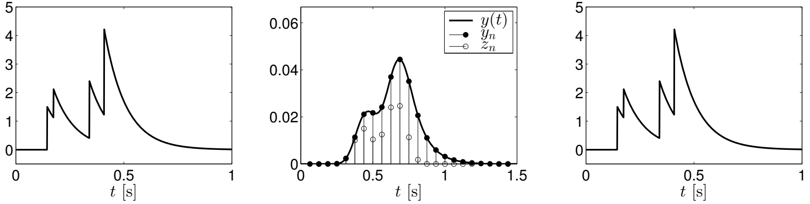
Sampling Streams of Diracs: Numerical Example



(a) Input signal, $x(t)$ **(b)** Sampling kernel, $h(t)$ **(c)** e^{α_m} reproduced by $h(t)$

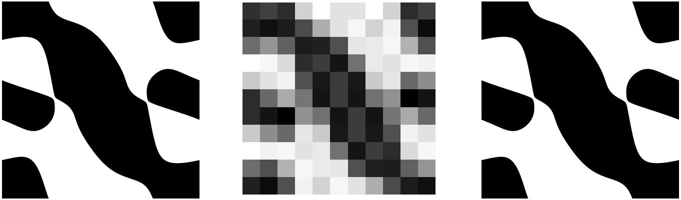


Stream of Decaying Exponentials



(a) Input signal, $x(t)$ **(b)** Filtered and sampled signal **(c)** Reconstructed signal

Sampling 2-D domains



The curve is implicitly defined through the equation [PanBluDragotti:11,14]:

$$f(x,y) = \sum_{k=1}^{K} \sum_{i=1}^{I} b_{k,i} e^{-j2\pi xk/M} e^{-j2\pi yi/N} = 0.$$

The coefficients $b_{k,i}$ are the only free parameters in the model.

Sampling 2-D domains

samples

interpolation

inter+ curve constraint

Application: Image Super-Resolution [BaboulazD:09]

Acquisition with Nikon D70

(a)Original (2014 $\times$ 3040)

(b) ROI (128 $\times$ 128)

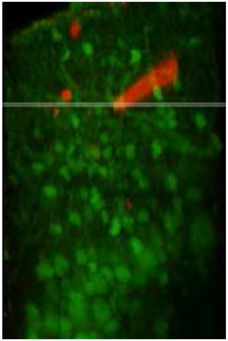
(b) Super-res (1024 $\times$ 1024)

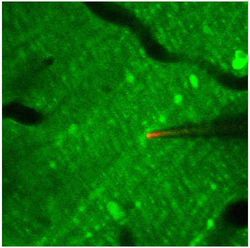
Application: Image Super-Resolution

(a)Original (48×48)

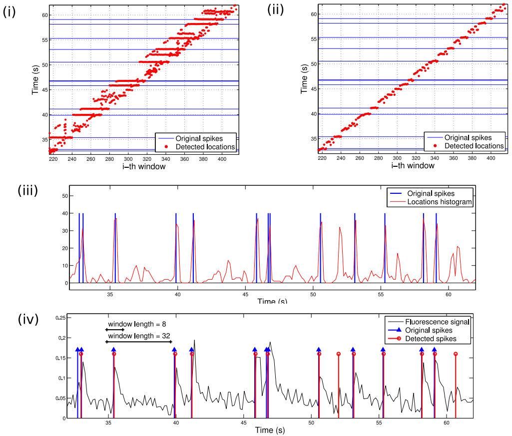
(b) Super-res (480×480)

Neural Activity Detection [OnativiaSD:13]

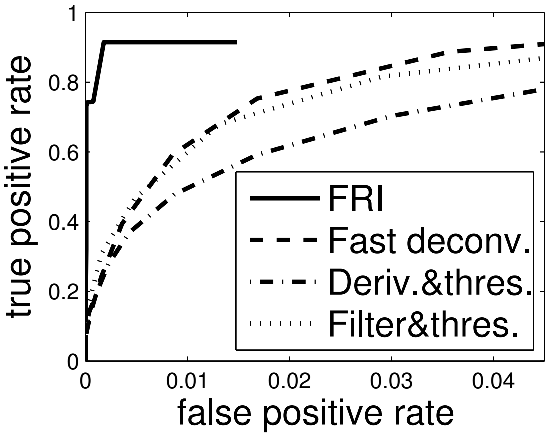




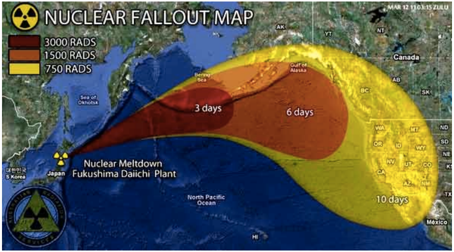
Calcium Transient Detection



Calcium Transient Detection



Localisation of Diffusion Sources using Sensor Networks [Murray-BruceD:14]

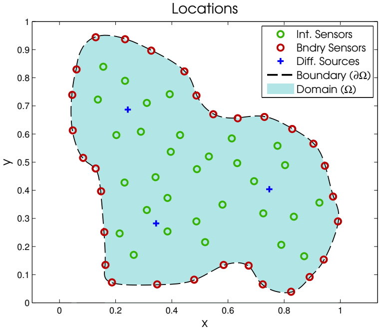


- The diffusion equation models the dispersion of chemical plumes, smoke from forest fires, radioactive materials
- The phenomenon is sampled in space and time using a sensor network.
- Sources often localised in space. Can we retrieve their location and the time of activation?

Localisation of Diffusion Sources using Sensor Networks

Good news:

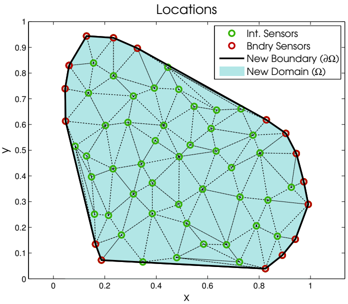
- When sources are localised in space and time, the field inversion is equivalent to a sparse sampling problem
- Proper linear combinations of sensors measurements in time and space leads to a Prony-type problem



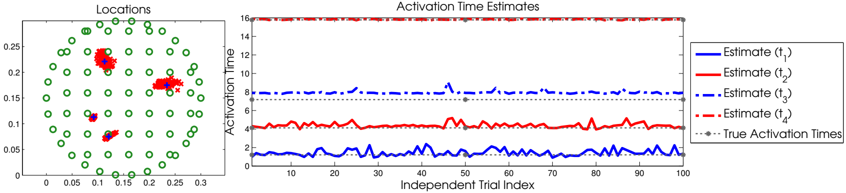
Localisation of Diffusion Sources using Sensor Networks

Good news:

- When sources are localised in space and time, the field inversion is equivalent to a sparse sampling problem
- Proper linear combinations of sensors measurements in time and space leads to a Prony-type problem



Localisation of Diffusion Sources: Numerical Results



(b) 100 independent trials using noisy sensor measurement samples (SNR=15dB).

Conclusions

Sampling signals using sparsity models:

- New framework that allows the sampling and reconstruction of continuous-time non-bandlimited signals.
- Use the knowledge of the acquisition process to map discrete measurements to specific integral measurements
- Use sparsity priors to reconstruct the original signal

Outlook:

- Promising applications in neuroscience, sensor networks, super-resolution imaging
- No golden hammer. Same framework but you need to fit the right model and carve the right solution for your problem: continuous/discrete, fast/ complex, redundant/ not-redundant

Still many open questions from theory to practice!

References

On sampling

- J. Uriguen, T. Blu, and P.L. Dragotti ‘FRI Sampling with Arbitrary Kernels’, IEEE Trans. on Signal Processing, November 2013
- 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, vol. 25(2), pp. 31-40, March 2008
- P.L. Dragotti, M. Vetterli and T. Blu, ‘Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon meets Strang-Fix’, IEEE Trans. on Signal Processing, vol.55 (5), pp.1741-1757, May 2007.
- J.Berent and P.L. Dragotti, and T. Blu,‘Sampling Piecewise Sinusoidal Signals with Finite Rate of Innovation Methods,’ IEEE Transactions on Signal Processing, Vol. 58(2),pp. 613-625, February 2010.
- J. Uriguen, P.L. Dragotti and T. Blu,‘On the Exponential Reproducing Kernels for Sampling Signals with Finite Rate of Innovation’ in Proc. of Sampling Theory and Application Conference, Singapore, May 2011.
- H. Pan, T. Blu, and P.L. Dragotti, ‘Sampling Curves with Finite Rate of Innovation’ IEEE Trans. on Signal Processing, January 2014.

References (cont'd)

On Image Super-Resolution

- L. Baboulaz and P.L. Dragotti, 'Exact Feature Extraction using Finite Rate of Innovation Principles with an Application to Image Super-Resolution', IEEE Trans. on Image Processing, vol.18(2), pp. 281-298, February 2009.

On Calcium Transient Detection

- Jon Onativia, Simon R. Schultz, and Pier Luigi Dragotti, A Finite Rate of Innovation algorithm for fast and accurate spike detection from two-photon calcium imaging, Journal of Neural Engineering, August 2013 .

On Diffusion Fields and Sensor Networks

- John Murray-Bruce and Pier Luigi Dragotti, Spatio-Temporal Sampling and Reconstruction of Diffusion Fields induced by Point Sources, Proc. of IEEE Conf. ICASSP, Florence (It), May 2014 .

Overview of Prony's Method

Assume: $y_n = \sum_{k=0}^{K-1} \alpha_k u_k^m$ and consider the polynomial:

$$P(x) = \prod_{k=1}^{K}(x - u_k) = x^K + h_1 x^{K-1} + h_2 x^{K-2} + \ldots + h_{K-1} x + h_K.$$

It is easy to verify that

$$y_{n+K} + h_1 y_{n+K-1} + h_2 y_{n+K-2} + \ldots + h_K y_n = \sum_{1 \le k \le K} \alpha_k u_k^n P(u_k) = 0.$$

In matrix-vector form for indices n such that $\ell \le n < \ell + K$, we get

$$\begin{bmatrix} y_{\ell+K} & y_{\ell+K-1} & \cdots & y_\ell \\ y_{\ell+K+1} & y_{\ell+K} & \cdots & y_{\ell+1} \\ \vdots & \ddots & \ddots & \vdots \\ y_{\ell+2K-2} & \ddots & \ddots & \vdots \\ y_{\ell+2K-1} & y_{\ell+2K-2} & \cdots & y_{\ell+K-1} \end{bmatrix} \begin{bmatrix} 1 \\ h_1 \\ h_2 \\ \vdots \\ h_K \end{bmatrix} = \mathbf{T}_{K,\ell} \mathbf{h} = \mathbf{0}$$

Overview of Prony's Method

The vector of polynomial coefficients $\mathbf{h} = [1, h_1, ..., h_K]^T$ is in the null space of $\mathbf{T}_{K,\ell}$. Moreover, $\mathbf{T}_{K,\ell}$ has size $K \times (K+1)$ and has full row rank when the u_k's are distinct. Therefore $\mathbf{h}$ is unique.

□

Prony's method summary:

1. Given the input y_n, build the Toeplitz matrix $\mathbf{T}_{K,\ell}$ and solve for $\mathbf{h}$. This can be achieved by taking the SVD of $\mathbf{T}_{K,\ell}$.
2. Find the roots of $P(x) = 1 + \sum_{n=1}^{K} h_k x^{K-k}$. These roots are exactly the exponentials $\{u_k\}_{k=0}^{K-1}$.
3. Given the $\{u_k\}_{k=0}^{K-1}$, find the corresponding amplitudes $\{\alpha_k\}_{k=0}^{K-1}$ by solving K linear equations.