Imperial College London

Sampling and Reconstruction driven by Sparsity Models: Theory and Applications

Pier Luigi Dragotti

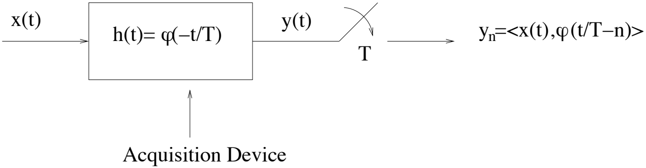
April 18, 2011

Outline

- Problem Statement
- Signals with Finite Rate of Innovation
- Sampling Kernels: E-splines and B-splines
- Sparse Sampling: the Basic Set-up and Extensions
- The Noisy Scenario
- Applications
  - Compression
  - Image Super-resolution
- Conclusions and Outlook

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 $x(t)$.

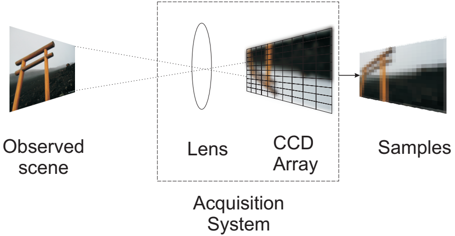


Natural questions:

- 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?

Imperial College London

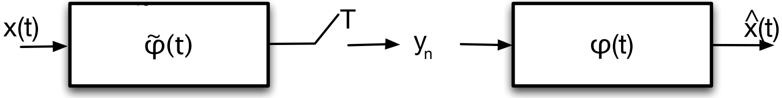
Problem Statement



- The low-quality lens blurs the images.
- The images are under-sampled by the low resolution CCD array.
- You need a good post-processing algorithm to undo the blurring and upsample the images.

Classical Sampling Formulation

- Sampling of $x(t)$ is equivalent to projecting $x(t)$ into the shift-invariant subspace $V = \text{span}\{\varphi(t/T - n)\}_{n\in\mathbb{Z}}$.
- If $x(t) \in V$, perfect reconstruction is possible.
- Reconstruction process is linear: $\hat{x}(t) = \sum_n y_n \varphi(t/T - n)$.
- For bandlimited signals $\varphi(t) = \text{sinc}(t)$.



Signals with Finite Rate of Innovation

- The signal $x(t) = \sum_n y_n \varphi(t/T - n)$ is exactly specified by one parameter y_n every T seconds, $x(t)$ has a finite number $\rho = 1/T$ of degrees of freedom per unit of time.
- In the classical formulation, innovation is uniform. How about signals where the rate of innovation is finite but non-uniform? E.g.
  - Piecewise sinusoidal signals (Frequency Hopping modulation)
  - Pulse position modulation (UWB)
  - Edges in images

Signals with Finite Rate of Innovation

Consider a signal of the form:

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

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

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

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 (1) and with a finite ρ as defined in (2).

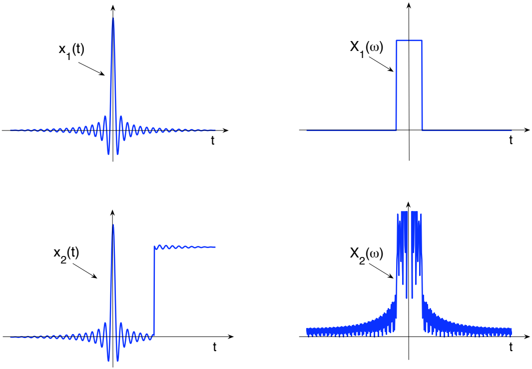
Signals with Finite Rate of Innovation

FRI signals include:

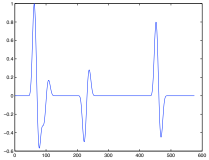
- Bandlimited signals and signals belonging to shift-invariant subspaces.
- K-sparse discrete signals (like in Compressed Sensing).
- Signals with point-like innovation, (point source phenomena), piecewise sinusoidal signals (OFDM, FH), filtered Diracs (UWB, Neuronal signals).

Signals with Finite Rate of Innovation

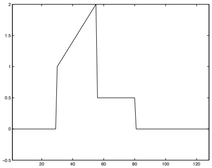
Here, $x_1(t)$ and $x_2(t)$ have the same rate of innovation. However, one discontinuity and no sampling theorems ;-)



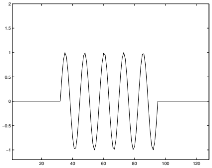
Examples of Signals with Finite Rate of Innovation



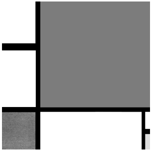
Filtered Streams of Diracs



Piecewise Polynomial Signals

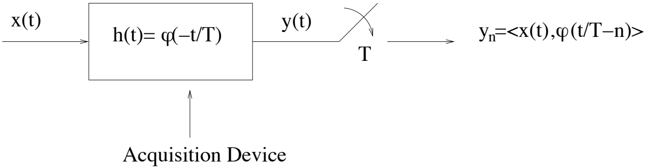


Piecewise Sinusoidal Signals



Mondrian paintings ;-)

The Sampling Kernel



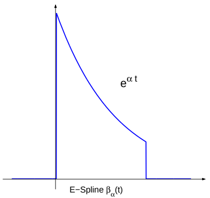
- Given by nature
  - Diffusion equation, Green function. Ex: sensor networks.
- Given by the set-up
  - Designed by somebody else. Ex: Hubble telescope, digital cameras.
- Given by design
  - Pick the best kernel. Ex: engineered systems.

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].

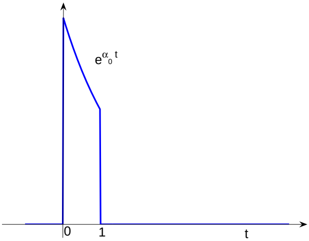


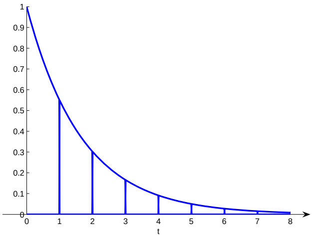
$$\beta_\alpha(t) \Leftrightarrow \hat{\beta}(\omega) = \frac{1 - e^{\alpha - j\omega}}{j\omega - \alpha}$$

Notice:

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

Kernels Reproducing Exponentials





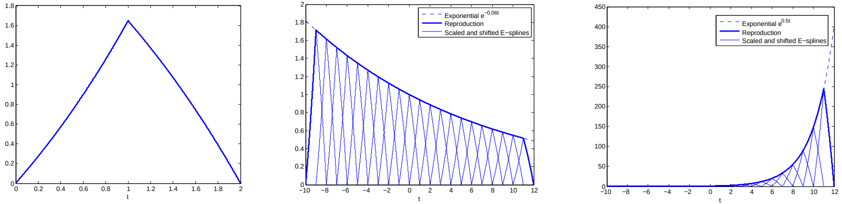
The E-spline of first order $\beta_{\alpha_0}(t)$ reproduces the exponential $e^{\alpha_0 t}$:

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

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

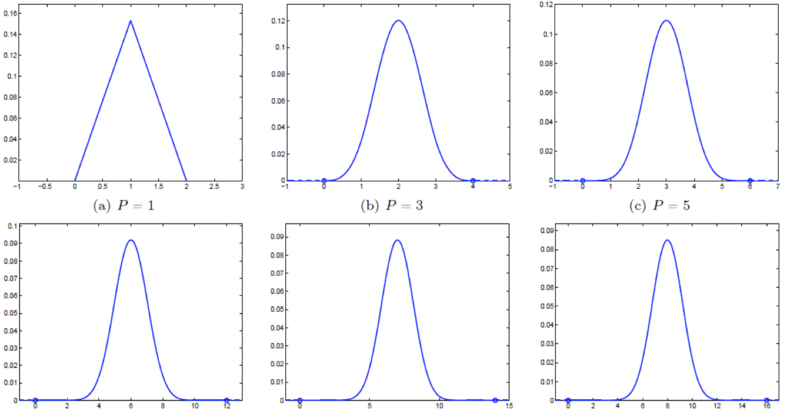
Imperial College London

Kernels Reproducing Exponentials



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 $\alpha_1 = 0.5$.

Examples of E-Splines Kernels

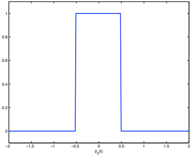


(a) $P = 1$ (b) $P = 3$ (c) $P = 5$

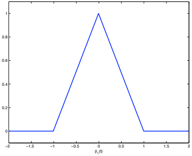
(d) $P = 11$ (e) $P = 13$ (f) $P = 15$

E-Splines and B-splines

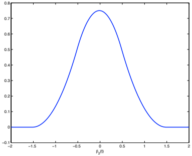
When $\alpha_m = 0$, $m = 0, 1, ..., L$. The E-spline reduce to the classical B-spline and is then able to reproduce polynomials up to degree L. Notice that any scaling function in wavelet theory is given by $\gamma(t) * \beta_L(t)$ and is therefore included in this definition.



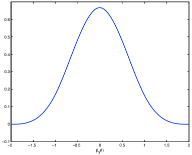
$\beta_0(t)$



$\beta_1(t)$



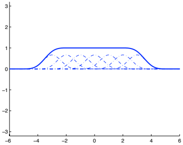
$\beta_2(t)$



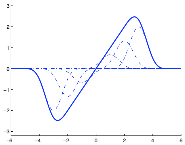
$\beta_3(t)$

E-Splines and B-splines

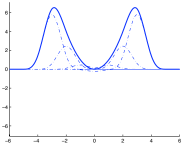
The E-spline reduces to the classical cubic B-spline when $\alpha_m = 0$, $m = 0, 1, ..., L$ and $L = 3$. In this case it can reproduce polynomials up to degree $L = 3$.



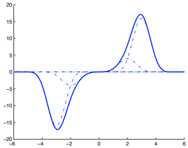
$c_{0,n} = (1, 1, 1, 1, 1, 1, 1)$



$c_{1,n} = (-3, -2, -1, 0, 1, 2, 3)$



$c_{2,n} \sim (8.7, 3.7, 0.7, -0.333, 0.7, 3.7, 8.7)$



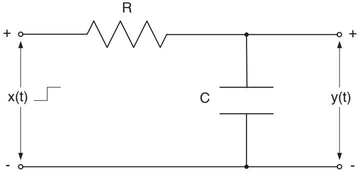
$c_{3,n} \sim (-24, -6, -0.001, 0, 0.001, 6, 24)$

Kernel Reproducing Exponential

Any functions with rational Fourier transform:

$$\hat{\varphi}(\omega) = \frac{\prod_i (j\omega - b_i)}{\prod_m (j\omega - a_m)} \qquad m = 0, 1, ..., L.$$

is a *generalized* E-splines. This includes practical devices as common as an RC circuit:



Imperial College London

Sparse Sampling: Basic Set-up

- Assume the sampling period $T = 1$.
- Consider any $x(t)$ with $t \in [0, N)$.
- Assume the sampling kernel $\varphi(t)$ is any function that can reproduce exponentials of the form

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

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

Sparse Sampling: Basic Set-up

We have that

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

- s_m is the bilateral Laplace transform of $x(t)$ evaluated at α_m.
- When $\alpha_m = j\omega_m$ then $s_m = \hat{x}(\omega_m)$ where $\hat{x}(\omega)$ is the Fourier transform of $x(t)$.
- When $\alpha_m = 0$, the s_m's are the polynomial moments of $x(t)$.

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 $\alpha_m = \alpha_0 + m\lambda \quad m = 0, 1, ..., L$ and $L \geq 2K - 1$.
- We have N samples: $y_n = \langle x(t), \varphi(t - n) \rangle, \; n = 0, 1, ...N - 1$:
- We obtain

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

Imperial College London

The Annihilating Filter Method

- The quantity

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

is a sum of exponentials.

- We can retrieve the locations u_k and the amplitudes $\hat{x}_k$ with the annihilating filter method (also known as Prony's method since it was discovered by Gaspard de Prony in 1795).
- Given the pairs $\{u_k, \hat{x}_k\}$, then $t_k = (\ln u_k)/\lambda$ and $x_k = \hat{x}_k / e^{\alpha_0 t_k}$.

The Annihilating Filter Method

1. Call h_m the filter with z-transform $H(z) = \sum_{i=0}^{K} h_i z^{-i} = \prod_{k=0}^{K-1}(1 - u_k z^{-1})$. We have that

$$h_m * s_m = \sum_{i=0}^{K} h_i s_{m-i} = \sum_{i=0}^{K}\sum_{k=0}^{K-1} \hat{x}_k h_i u_k^{m-i} = \sum_{k=0}^{K-1} \hat{x}_k u_k^m \underbrace{\sum_{i=0}^{K} h_i u_k^{-i}}_{0} = 0.$$

This filter is thus called the annihilating filter. In matrix/vector form, we have that $\mathbf{S}H = 0$ and using the fact that $h_0 = 1$, we obtain

$$\begin{bmatrix} s_{K-1} & s_{K-2} & \cdots & s_0 \\ s_K & s_{K-1} & \cdots & s_1 \\ \vdots & \vdots & \ddots & \vdots \\ s_{L-1} & s_{L-2} & \cdots & s_{L-K} \end{bmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_K \end{pmatrix} = - \begin{pmatrix} s_K \\ s_{K+1} \\ \vdots \\ s_L \end{pmatrix}.$$

Solve the above system to find the coefficients of the annihilating filter.

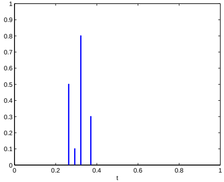
The Annihilating Filter Method

2. Given the coefficients $\{1, h_1, h_2, ..., h_k\}$, we get the locations u_k by finding the roots of $H(z)$.
3. Solve the first K equations in $s_m = \sum_{k=0}^{K-1} \hat{x}_k u_k^m$ to find the amplitudes $\hat{x}_k$. In matrix/vector form

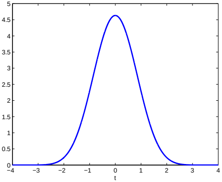
$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ u_0 & u_1 & \cdots & u_{K-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_0^{K-1} & u_1^{K-1} & \cdots & u_{K-1}^{K-1} \end{bmatrix} \begin{pmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_{K-1} \end{pmatrix} = \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_{K-1} \end{pmatrix}. \quad (3)$$

Classic Vandermonde system. Unique solution for distinct u_ks.

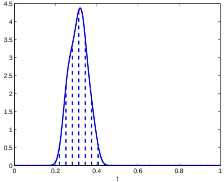
Sampling Streams of Diracs: Numerical Example



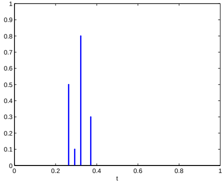
(a) Original Signal



(b) Sampling Kernel ($\beta_7(t)$)

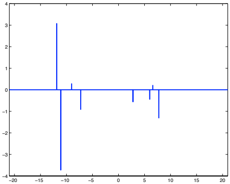


(c) Samples

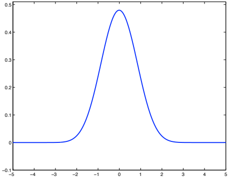


(d) Reconstructed Signal

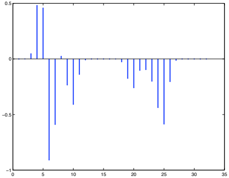
Sampling Streams of Diracs: Sequential Reconstruction



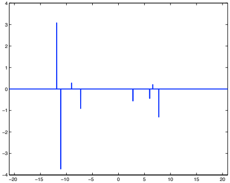
(a) Original Signal



(b) Sampling Kernel ($\beta_7(t)$)



(c) Samples



(d) Reconstructed Signal

Imperial College London

Note on the proof

Linear vs Non-linear

- Problem is **Non-linear** in t_k, but **linear** in x_k given t_k
- The key to the solution is the separability of the non-linear from the linear problem using the annihilating filter.

The proof is based on a constructive algorithm:

1. Given the N samples y_n, compute the moments s_m using the exponential reproduction formula. In matrix vector form $S = \mathbf{C}Y$.
2. Solve a $K \times K$ Toeplitz system to find $H(z)$
3. Find the roots of $H(z)$
4. Solve a $K \times K$ Vandermonde system to find the a_k

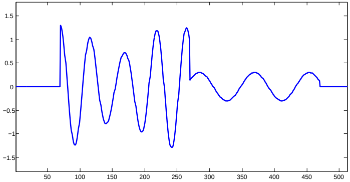
Complexity

1. $O(KN)$
2. $O(K^2)$
3. $O(K^3)$
4. $O(K^2)$

Thus, the algorithm complexity is polynomial with the signal innovation.

Sparse Sampling: Extensions

Using variations of the annihilating filter methods other signals can be sampled such as filtered streams of Diracs, multi-dimensional signals and piecewise sinusoidal signals.



Sampling Piecewise Sinusoidal Signals [BerentDB:10]

We consider signals of the type:

$$x(t) = \sum_{d=1}^{D} \sum_{n=1}^{N} A_{d,n} \cos(\omega_{d,n} t + \varphi_{d,n}) \xi_d(t),$$

where

$\xi_d(t) = u(t - t_d) - u(t - t_{d+1})$ and $-\infty < t_1 < ... < t_d < ... < t_{D+1} < \infty$.

Why is it difficult to sample them?

- Piecewise sinusoidal signals contain innovation in both spectral and temporal domains.
- They are not bandlimited.
- They are not sparse in time nor in a basis or a frame.

Sampling Piecewise Sinusoidal Signals

From the samples we can obtain the Laplace transform of $x(t)$ at $\alpha_m = \alpha_0 + m\lambda,\ m = 0, 1, ..., L$:

$$s_m = \sum_{d=1}^{D} \sum_{n=1}^{2N} \bar{A}_{d,n} \frac{\left[e^{t_{d+1}(j\omega_{d,n}+\alpha_m)} - e^{t_d(j\omega_{d,n}+\alpha_m)}\right]}{(j\omega_{d,n} + \alpha_m)},$$

where $\bar{A}_{d,n} = A_{d,n} e^{j\varphi_{d,n}}$. We define the polynomial

$$Q(\alpha_m) = \prod_{d=1}^{D} \prod_{n=1}^{2N} (j\omega_{d,n} + \alpha_m) = \sum_{j=0}^{J} r_j \alpha_m^j.$$

Sampling Piecewise Sinusoidal Signals

Multiplying both side of the equation by $Q(\alpha_m)$ we obtain:

$$Q(\alpha_m)s_m = \sum_{d=1}^{D}\sum_{n=1}^{2N}\bar{A}_{d,n}P(\alpha_m)[e^{t_{d+1}(j\omega_{d,n}+\alpha_m)} - e^{t_d(j\omega_{d,n}+\alpha_m)}], \quad (4)$$

where $P(\alpha_m)$ is a polynomial. Since $\alpha_m = \alpha_0 + \lambda m$ the right-hand side of (4) can be annihilated:

$$Q(\alpha_m)s_m * h_m = 0.$$

Sampling Piecewise Sinusoidal Signals

In matrix/vector form (assuming $h_0 = 1$), we have:

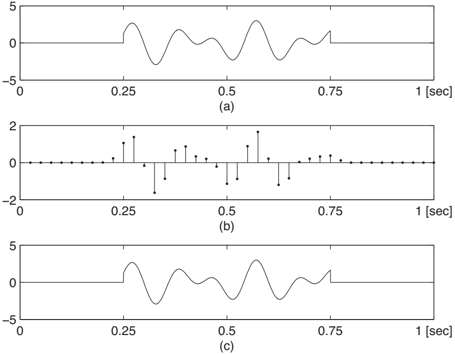
$$\begin{bmatrix} s_K & \cdots & \alpha_K^J s_K & \cdots & s_0 & \cdots & \alpha_0^J s_0 \\ s_{K+1} & \cdots & \alpha_{K+1}^J s_{K+1} & \cdots & s_0 & \cdots & \alpha_1^J s_1 \\ \vdots & \vdots & \ddots & \vdots & & & \\ s_L & \cdots & \alpha_L^J s_L & \cdots & s_0 & \cdots & \alpha_{(L-K)}^J s_{(L-K)} \end{bmatrix} \begin{pmatrix} r_0 \\ \vdots \\ r_J \\ h_1 r_0 \\ h_1 r_1 \\ \vdots \\ h_K r_J \\ \vdots h_K r_K \end{pmatrix} = 0.$$

Sampling Piecewise Sinusoidal Signals

- From the coefficients r_j, $j = 0, 1, ...J$, we obtain $Q(\alpha_m)$.
- The roots of the filter $H(z)$ and of the polynomial $Q(\alpha_m)$ give the locations of the switching points and the frequencies of the sine waves respectively.
- To solve the system we need $L \geq 4D^3N^2 + 4D^2N^2 + 4D^2N + 6DN$.

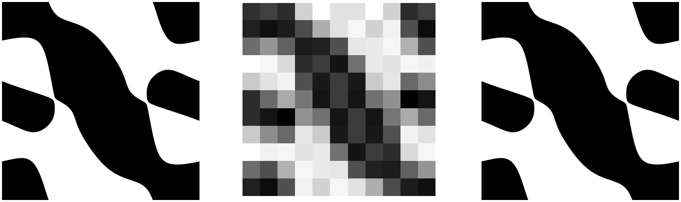
Imperial College London

Numerical Example

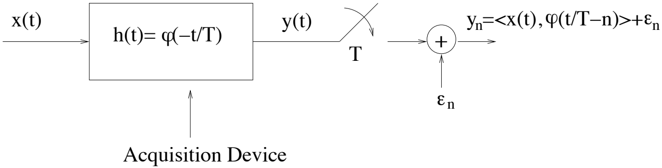


Imperial College London

Sampling 2-D domains



Robust Sparse Sampling



- The measurements are noisy
- The noise is additive and i.i.d. Gaussian

Robust Sparse Sampling

In the presence of noise, the annihilation equation

$$\mathbf{S}H = 0$$

is only approximately satisfied.
Minimize: $\|\mathbf{S}H\|_2$ under the constraint $\|H\|_2 = 1$.
This is achieved by performing an SVD of $\mathbf{S}$:

$$\mathbf{S} = \mathbf{U}\lambda\mathbf{V}^{\mathbf{T}}.$$

Then H is the last column of $\mathbf{V}$.
Notice: this is similar to Pisarenko's method in spectral estimation.

Robust Sparse Sampling: Cadzow's algorithm

For small SNR use Cadzow's method to denoise $\mathbf{S}$ before applying TLS. The basic intuition behind this method is that, in the noiseless case, $\mathbf{S}$ is rank deficient (rank K) and Toeplitz, while in the noisy case $\mathbf{S}$ is full rank.

Algorithm:

- SVD of $\mathbf{S} = \mathbf{U}\lambda\mathbf{V}^{\mathbf{T}}$.
- Keep the K largest diagonal coefficients of λ and set the others to zero.
- Reconstruct $\mathbf{S}' = \mathbf{U}\lambda'\mathbf{V}^{\mathbf{T}}$.
- This matrix is not Toeplitz, make it so by averaging along the diagonals.
- Iterate.

Robust Sparse Sampling: Best Kernel

The exponential reproducing kernel has the following form

$$\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}}(t).$$

How should we choose $\gamma(t)$ and α_m, $m = 0, 1, ..., L$ so as to minimize the effect of noise?
In the noiseless case:

$$S = \mathbf{C}Y.$$

When additive noise is present

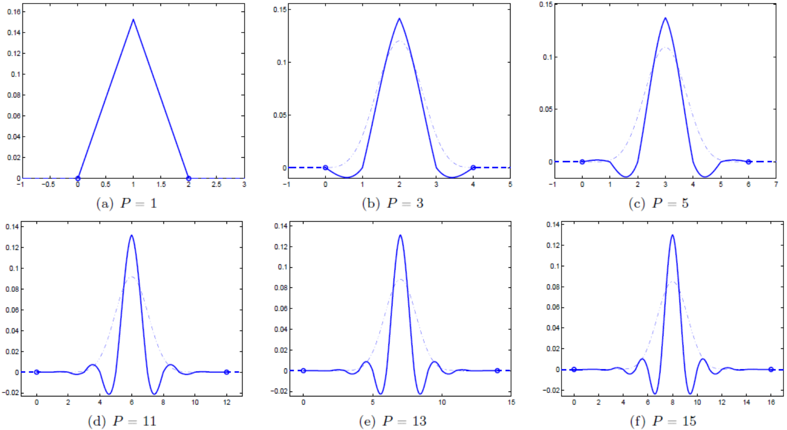
$$\hat{S} = \mathbf{C}Y + \mathbf{C}\epsilon.$$

Here $\mathbf{C}$ is the $L \times N$ matrix of the exponential reproducing coefficients $c_{m,n}$.

Robust Sparse Sampling: Best Kernel (cont'd)

- When ϵ_n is i.i.d., we want the rows of **C** to be orthonormal.
- Since $c_{m,n} = c_{m,0}e^{\alpha_m n}$, orthogonality is achieved by choosing $\alpha_m = j2\pi m/N$.
- Orthonormality requires $|c_{m,0}| = 1$, this is achieved by imposing $|\hat{\gamma}(2\pi m/N)\hat{\beta}_{\vec{\alpha}}(2\pi m/N)| = 1,\ m = 0, 1, ..., L$.
- We choose $\gamma(t)$ to be polynomial in the frequency domain: $\hat{\gamma}(\omega) = \sum_{i=0}^{L-1} d_i \omega^i$. Thus the coefficients d_i are chosen so that the polynomial $\hat{\gamma}(\omega)$ interpolates the points $(j2\pi m/N, \hat{\beta}_{\vec{\alpha}}(2\pi m/N)^{-1})$.

Examples of Best E-Splines

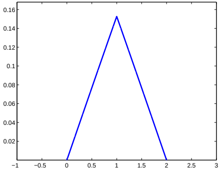


(a) $P = 1$ (b) $P = 3$ (c) $P = 5$

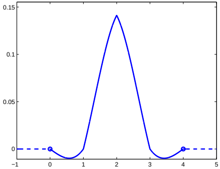
(d) $P = 11$ (e) $P = 13$ (f) $P = 15$

Imperial College London

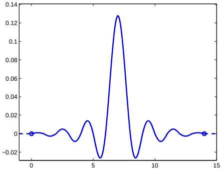
Examples of Best Kernels



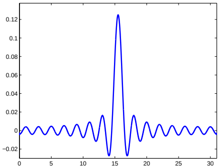
(a) $L = 2$



(b) $L = 4$

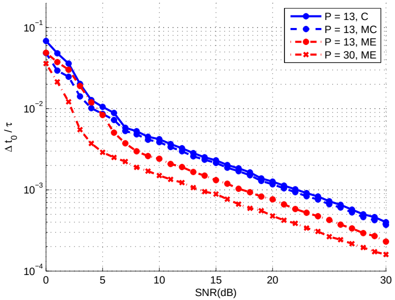


(c) $L = 14$



(d) $L = 31$

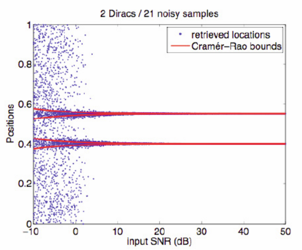
Performance of different Kernels

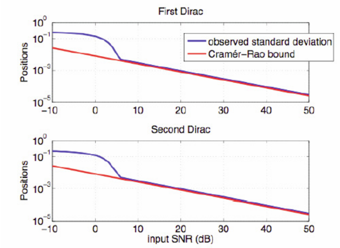


Here, $K = 2$ and we measure the error in the retrieval of the location of the Diracs.

Imperial College
London

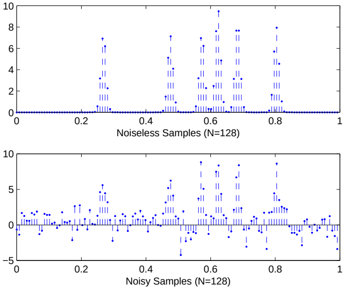
Robust Sparse Sampling

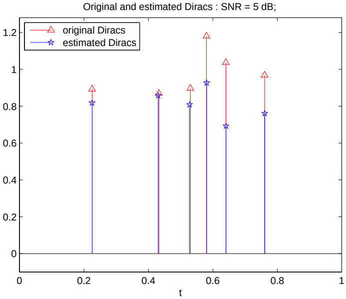




- Samples are corrupted by additive noise.
- This is a parametric estimation problem.
- Unbiased algorithms have a covariance matrix lower bounded by CRB.
- The proposed algorithm reaches CRB down to SNR of 5dB.

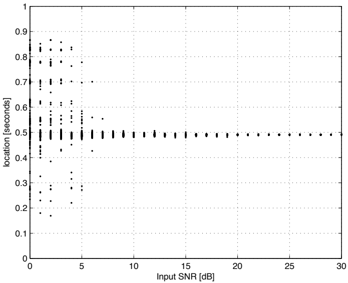
Robust Sparse Sampling

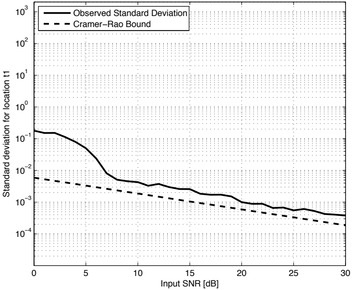




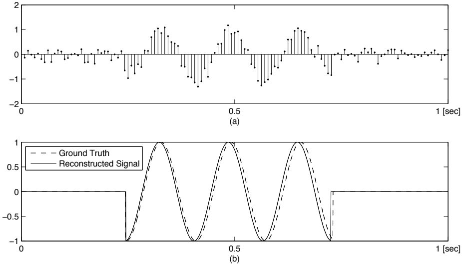
Robust Sparse Sampling

Piecewise sinusoidal signal





Robust Sparse Sampling



SNR= 8dB, N=128.

Imperial College London

Comparison with Compressed Sensing

Both use a *sparsity prior* and a non-linear reconstruction.

Sampling of Signal with Finite Rate of Innovation

+ Continuous or discrete, infinite or finite dimensional

+ Retrieval of the support of x separate from the retrieval of the coefficients.

+ Close to “real” sampling, deterministic

− Not universal, designer matrices

Compressed sensing

+ Universal

± Probabilistic, can be complex

− Discrete, redundant

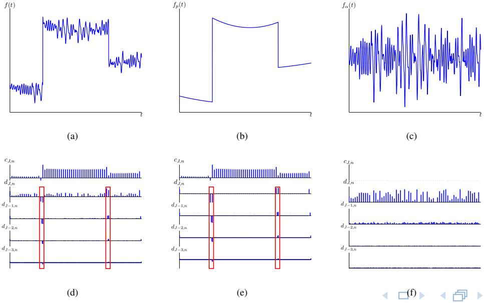
Compression

- FRI Signals can be sparsely sampled. Can they also be compressed? What happens when the samples are quantized?
- Traditional Compression is based on complex encoders and simple decoders.
- New sampling theories are characterized by a linear acquisition but non-linear reconstruction.

Compression

Signals are piecewise smooth, with α-Lipschitz regular pieces. Traditional compression algorithms use the wavelet transform and compress only the large wavelet coefficients. They achieve the optimal $D(R)$ performance:

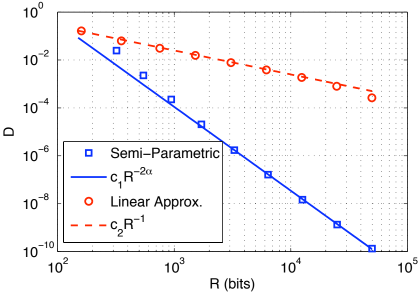
$$D(R) \sim R^{-2\alpha}$$

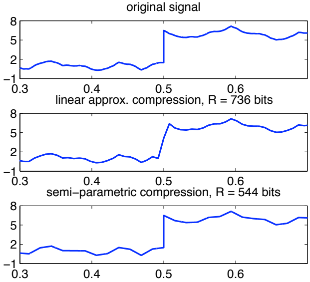


Performance Analysis

- The proposed algorithm compresses and transmits only the low-pass coefficients of the wavelet transform (linear approximation-based encoding), but uses FRI techniques to estimate the discontinuities in the signal from the low-pass coefficients (non-linear decoding).
- Any piecewise smooth signals can be decomposed into a piecewise polynomial and a globally smooth signal.
- The low-pass coefficients are a sufficient representation of the piecewise polynomial signal, but quantization and the smooth signal act as noise and this reduces the reconstruction fidelity.
- We treat both contributions as additive noise and evaluate the CR-bounds for this estimation problem. The quantization noise depends on the bit-rate R. This leads to a connection between CR-Bounds and rate-distortion analysis and leads to this performance bound [ChaisinthopD:11]:
$$D_{FRI}(R) \sim R^{-2\alpha}.$$

Simulation Results





Imperial College London

Application: Image Super-Resolution

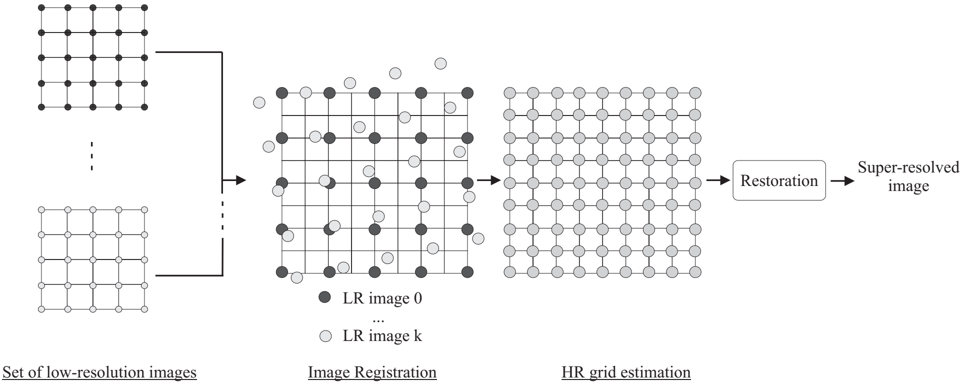
Super-Resolution is a multichannel sampling problem with unknown shifts. Use moments to retrieve the shifts or the geometric transformation between images.

(a)Original (512×512) (b) Low-res. (64×64) (c) Super-res (PSNR=24.2dB)

- Forty low-resolution and shifted versions of the original.
- The disparity between images has a finite rate of innovation and can be retrieved.
- Accurate registration is achieved by retrieving the continuous moments of the 'Tiger' from the samples.

Application: Image Super-Resolution

Image super-resolution basic building blocks



Application: Image Super-Resolution

- For each blurred image $I(x,y)$:
  - A pixel $P_{m,n}$ in the blurred image is given by

$$P_{m,n} = \langle I(x,y), \varphi(x/T - n, y/T - m) \rangle,$$

    where $\varphi(t)$ represents the point spread function of the lens.
  - We assume $\varphi(t)$ is a spline that can reproduce polynomials:

$$\sum_n \sum_m c_{m,n}^{(l,j)} \varphi(x-n, y-m) = x^l y^j \qquad l = 0,1,...,N; j = 0,1,...,N.$$

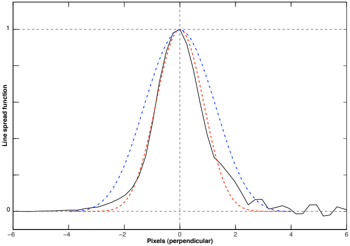
  - We retrieve the exact moments of $I(x,y)$ from $P_{m,n}$:

$$\tau_{l,j} = \sum_n \sum_m c_{m,n}^{(l,j)} P_{m,n} = \int \int I(x,y) x^l y^j \, dxdy.$$

- Given the moments from two or more images, we estimate the geometrical transformation and register them. Notice that moments of up to order three along the x and y coordinates allows the estimation of an affine transformation.

Application: Image Super-Resolution

(a)Original (2014 $\times$ 3039)



(b) Point Spread function

Application: Image Super-Resolution

(a)Original (128×128)

(b) Super-res (1024×1024)

Application: Image Super-Resolution

(a)Original (48 $\times$ 48)

(b) Super-res (480 $\times$ 480)

Conclusions

Sampling signals at their rate of innovation:

- New framework that allows the sampling and reconstruction of signals at a rate smaller than Nyquist rate.
- Robust and fast algorithms with complexity proportional to the number of degrees of freedom.
- Provable optimality (i.e. CRB achieved over wide SNR ranges).
- Intriguing connections with multi-resolution analysis, Fourier theory and analogue circuit theory.

But also

- There is no such thing as a free lunch. Core application is difficult.
- Still many open questions from theory to practice.

Imperial College London

References

On sampling

- 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' in Proc. of Sampling Theory and Application Conference, Singapore, May 2011.

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 compression

- V. Chaisinthop and P.L. Dragotti, ' Centralized and Distributed Semi-Parametric Compression of Piecewise Smooth Functions' Semi-Parametric Compression of Piecewise-Smooth Functions', accepted for publication in the IEEE Trans. on Signal Processing, January 2011.