EXPONENTIAL REPRODUCING KERNELS FOR SPARSE SAMPLING

Jose Antonio Uriguen
Pier Luigi Dragotti
Thierry Blu
29th June 2011

Communications & Signal Processing Group
Department of EEE
Imperial College of London

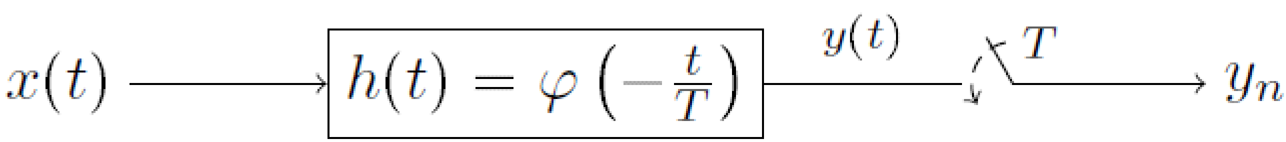
Content

Imperial College London

- Review of Finite Rate of Innovation
  - Motivation for FRI
  - Parametric signals
  - Appropriate sampling kernels
  - Sample & reconstruct a train of Diracs
- The noisy scenario
  - A subspace approach
  - Prewhitening
  - Modified E-Spline kernels

Problem statement

- **Motivation**
  - Sample a **sparse** *continuous-time* signal
  - Not necessarily band-limited → parametric / FRI signal (+) appropriate filtering



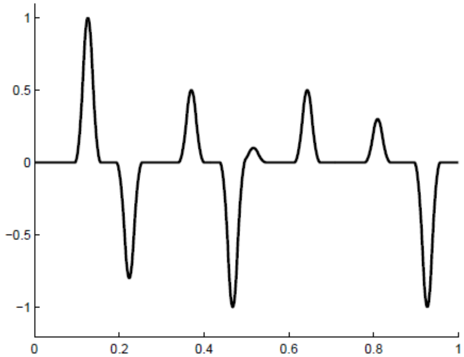
- Perfect reconstruction based on
  - Set of *N* discrete measurements
  - Taken every *T* seconds $y_n = \langle x(t), h(t - nT) \rangle$

Sparse signals to be sampled

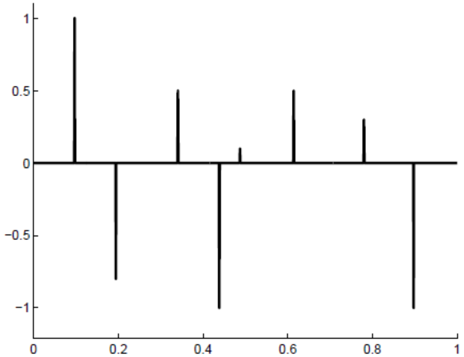
- Signals with *finite rate of innovation*
  - Finite amount of degrees of freedom
  - Parametric representation

$$x(t) = \sum_{k \in \mathbb{Z}} \sum_{r=0}^{R-1} \gamma_{k,r} g_r(t - t_k).$$

- Prototype signal



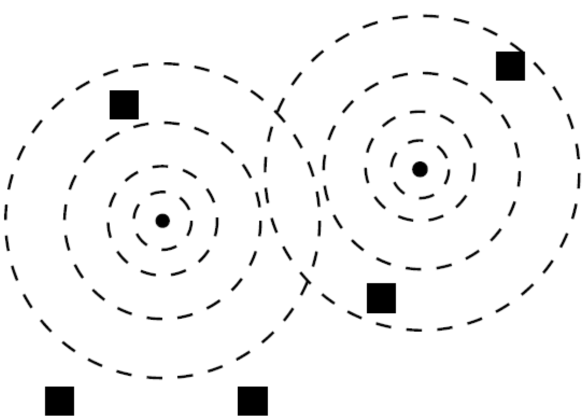
(e) Stream of Pulses

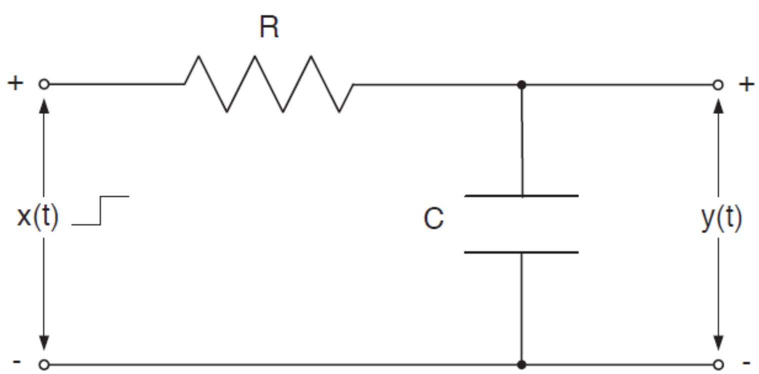


(a) Train of Diracs

Sampling kernels (i)

- Normally sampling kernels are given
  - *Natural process*: diffusion field observed by various spatially distributed sensors
  - *Acquisition device*: electric circuit, camera lens
  - Exponential reproducing kernels can model certain types of filters





Sampling kernels (ii)

- **Exponential reproducing kernels**
  - Finite support: order *P* → *P*+1
  - Reproduce exponentials

$$\sum_{n\in\mathbb{Z}} c_{m,n}\varphi(t-n) = \mathrm{e}^{\alpha_m t}$$

  - Coefficients are discrete time exponentials

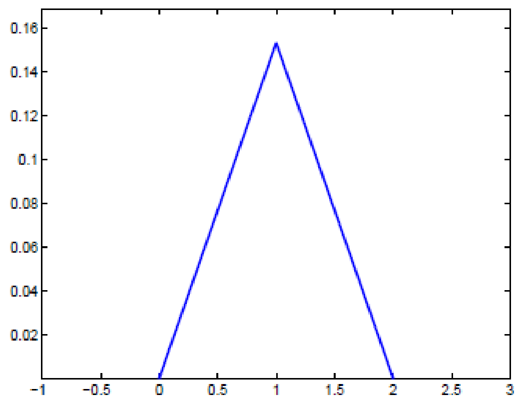
$$c_{m,n} = \mathrm{e}^{\alpha_m n} c_{m,0}$$

  - Based on E-Splines $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}_P}(t)$

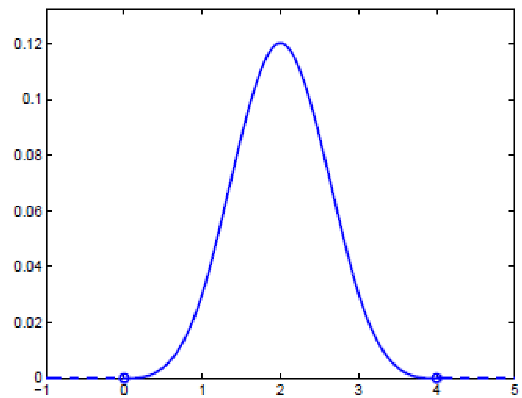
$$\hat{\beta}_{\vec{\alpha}_P}(\omega) = \prod_{m=0}^{P}\left(\frac{1-\mathrm{e}^{\alpha_m - j\omega}}{j\omega - \alpha_m}\right)$$

Sampling kernels (iii)

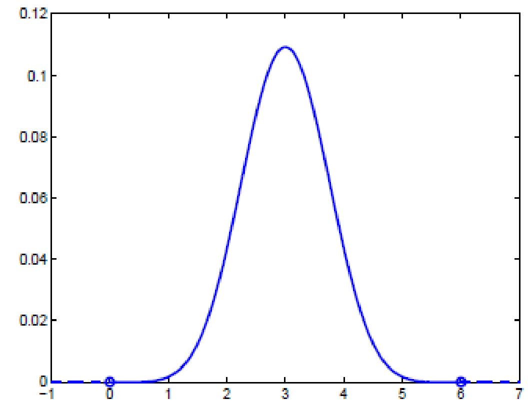
- Examples of E-Splines



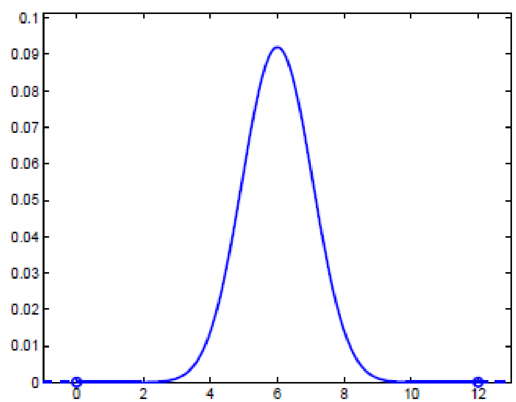
(a) $P = 1$



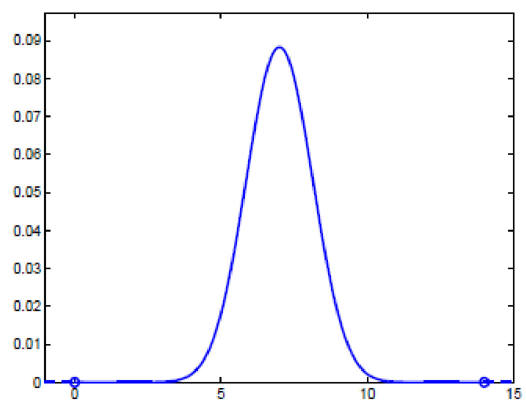
(b) $P = 3$



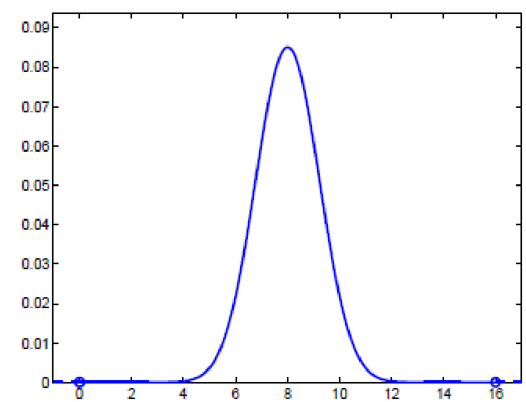
(c) $P = 5$



(d) $P = 11$



(e) $P = 13$



(f) $P = 15$

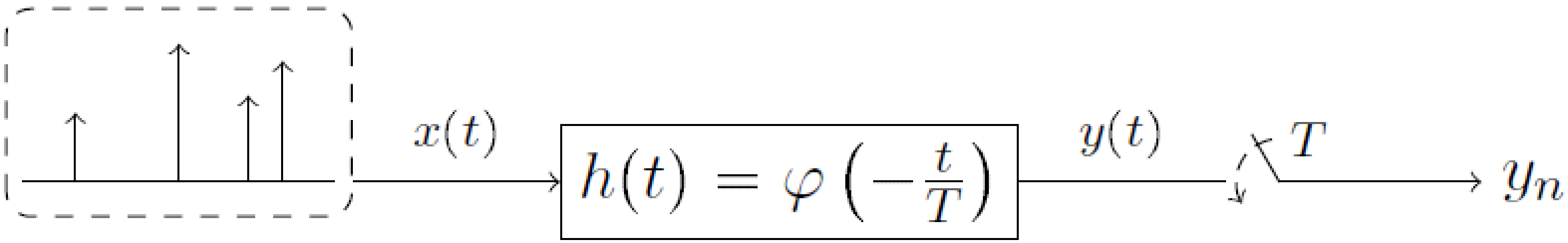
Sampling kernels (iv)

- ***Exponential reproducing kernels*** give us flexibility
  - They can be designed to accommodate certain types of given filters
    - Electric circuit, camera lens, etc
  - If we can choose the kernel, we can also optimise them
    - To handle noise effectively
    - To satisfy other requirements
- In any case, we need to find an appropriate $\gamma(t)$

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

Sample & Rec a train of Diracs (i)

- **1. Obtain the input *measurements***
  - Traditional linear sampling



  - Input characterised by (t_k, a_k) $k = 0, \ldots, K-1$
  - Set of N samples $y_n = \langle x(t), h(t - nT) \rangle$

Sample & Rec a train of Diracs (ii)

- **2.** *Modify* the samples
  - Obtain new measurements

$$s_m = \sum_{n=0}^{N-1} c_{m,n} y_n \quad m = 0, \ldots, P.$$

    - Linear transform $\mathbf{s} = \mathbf{C}\mathbf{y}$.
  - *Power sum series equivalence* $\longrightarrow$ Harmonic retrieval

$$s_m = \left\langle x(t), \sum_n c_{m,n} \varphi(t-n) \right\rangle = \int_{-\infty}^{\infty} x(t) \mathrm{e}^{\alpha_m \frac{t}{T}} \mathrm{d}t$$

$$= \sum_{k=0}^{K-1} \hat{a}_k u_k^m, \quad m = 0, 1, \ldots, P$$

$$\hat{a}_k = a_k \mathrm{e}^{\alpha_0 \frac{t_k}{T}}$$

$$u_k = \mathrm{e}^{\lambda \frac{t_k}{T}}$$

Sample & Rec a train of Diracs (iii)

- **3.** ***Retrieve*** **the input parameters (t_k, a_k)**
  - Prony's method --- Annihilating filter method

$$h_m * s_m = 0$$

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

$$\begin{pmatrix} s_L & s_{L-1} & \cdots & s_0 \\ s_{L+1} & s_L & \cdots & s_1 \\ \vdots & \vdots & \ddots & \vdots \\ s_P & s_{P-1} & \cdots & s_{P-L} \end{pmatrix}$$

    - Toeplitz matrix **S** is *rank deficient* → **h** null-space of **S**
    - Obtain u_k (t_k) from *roots* of **h**

$$\hat{h}(z) = \prod_{k=0}^{K-1} (1 - u_k z^{-1})$$

    - Find a_k using power sum series equation

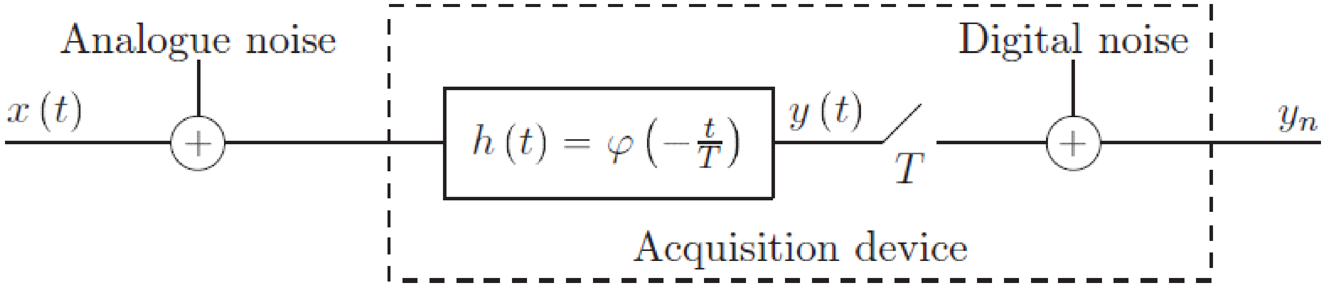
Content

Imperial College London

- Review of Finite Rate of Innovation
  - Motivation for FRI
  - Parametric signals
  - Appropriate sampling kernels
  - Sample & reconstruct a train of Diracs
- The noisy scenario
  - A subspace approach
  - Prewhitening
  - Modified E-Spline kernels

The noisy scenario

- Sampling scheme



- Consider only digital noise: AWGN(0, σ)

$$\tilde{y}_n = \langle x(t), h(t - nT) \rangle + \epsilon_n = y_n + \epsilon_n$$

- Degrades performance reconstruction algorithms

A Subspace Approach (i)

- Measurements change $\mathbf{s} = \mathbf{C}\mathbf{y} + \mathbf{C}\mathbf{e}$

$$\tilde{s}_m = \sum_{n=0}^{N-1} c_{m,n}\tilde{y}_n = s_m + \underbrace{\sum_{n=0}^{N-1} c_{m,n}\epsilon_n}_{b_m} \qquad m = 0, \ldots, P$$

- Toeplitz matrix **S** changes too

$$\tilde{\mathbf{S}} = \begin{pmatrix} s_L + b_L & s_{L-1} + b_{L-1} & \cdots & s_0 + b_0 \\ s_{L+1} + b_{L+1} & s_L + b_L & \cdots & s_1 + b_1 \\ \vdots & \vdots & \ddots & \vdots \\ s_P + b_P & s_{P-1} + b_{P-1} & \cdots & s_{P-L} + b_{P-L} \end{pmatrix}$$

  - Matrix is not rank deficient any more

$$\tilde{\mathbf{S}} = \mathbf{S} + \mathbf{B} \qquad \tilde{\mathbf{S}}\mathbf{h} \neq \mathbf{0}$$

A Subspace Approach (ii)

- Assume the term **B** in $\tilde{\mathrm{S}} = \mathrm{S} + \mathrm{B}$ is due to AWGN
  - Even though $\tilde{\mathrm{S}}\mathrm{h} \neq 0$
  - We could find **h** to minimise $\|\tilde{\mathrm{S}}\mathbf{h}\|^2$ s.t. $\|\mathbf{h}\| = 1$
- *Why?* Covariance matrix $\mathrm{B}^*\mathrm{B} = \sigma^2\mathrm{I}$
  - The noise affects equally signal and noise subspaces
- **SVD** is able to separate these subspaces
  - **h** is the vector corresponding to the noise subspace
  - Improve estimation using Cadzow

A Subspace Approach (iii)

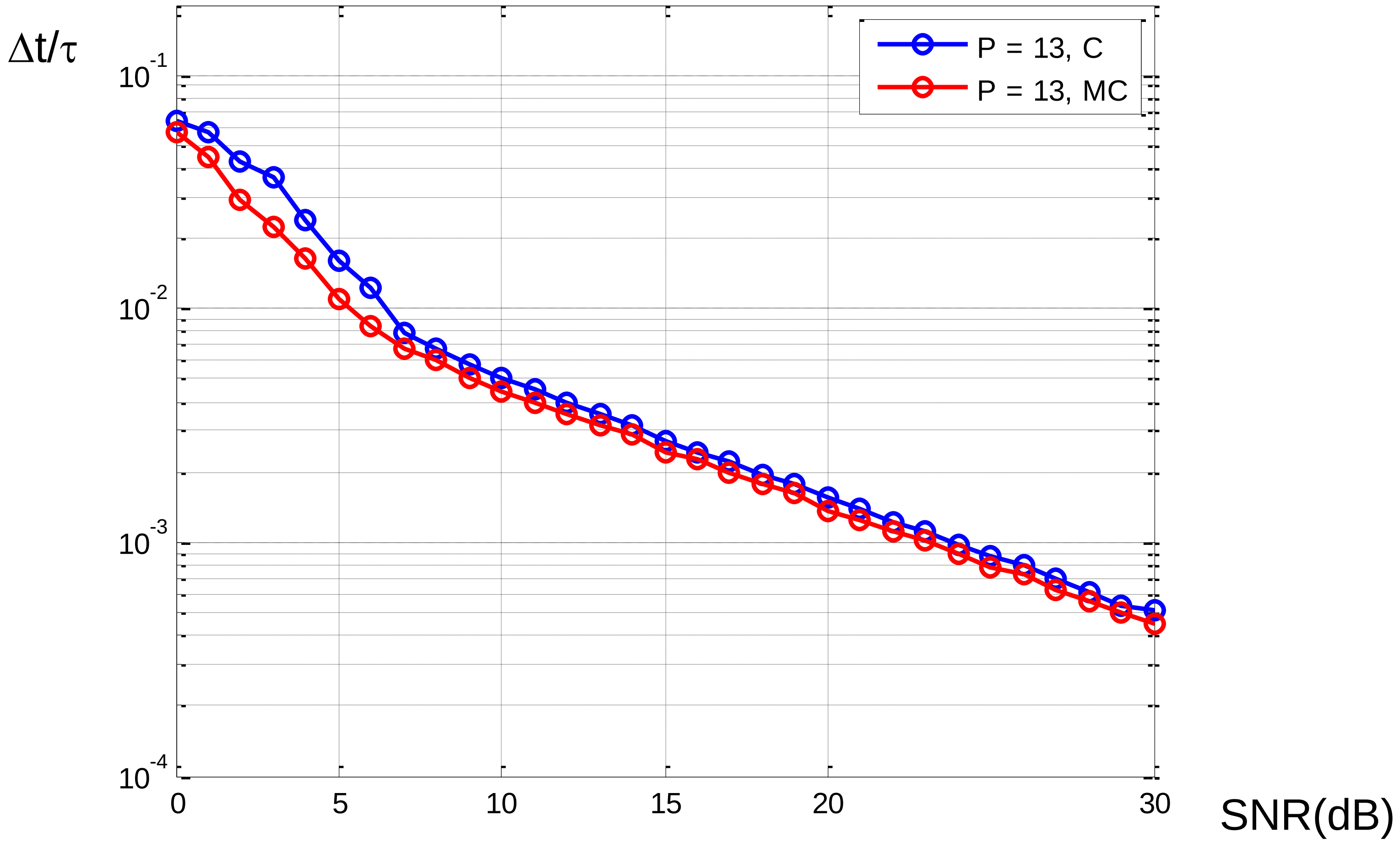
- The term **B** in $\tilde{\mathbf{S}} = \mathbf{S} + \mathbf{B}$ for exponential reproducing kernels is due to *coloured* noise
  - Now, $\mathbf{s} = \mathbf{C}\mathbf{y} + \boxed{\mathbf{C}\mathbf{e}}$ where $\mathbf{C}\mathbf{e}$ is coloured
  - We ***can't*** directly find **h** to minimise $\|\tilde{\mathbf{S}}\mathbf{h}\|^2$ s.t. $\|\mathbf{h}\| = 1$

- ***Approach***: estimate the covariance matrix of the noise $\mathbf{R} = \lambda\mathbf{B}^*\mathbf{B}$ and use Cholesky $\mathbf{R} = \mathbf{Q}^T\mathbf{Q}$
  - *pre-whiten* $\tilde{\mathbf{S}}' = \tilde{\mathbf{S}}'\mathbf{Q}^{-1}$
  - **SVD** is *now* able to separate subspaces

$$(\mathbf{B}\mathbf{Q}^{-1})^*(\mathbf{B}\mathbf{Q}^{-1}) = \sigma^2\lambda^{-1}\mathbf{I}$$

A Subspace Approach (iv)

- Simulations (K = 2 Diracs , N = 31 samples)



Modifying E-Splines (i)

- Coloured noise term → AWGN $\mathbf{s} = \mathbf{C}\mathbf{y} + \boxed{\mathbf{C}\mathbf{e}}$

$$\mathbf{C}_{(P+1)\times N} = \begin{pmatrix} c_{0,0} & c_{0,0}\mathrm{e}^{\alpha_0} & \cdots & c_{0,0}\mathrm{e}^{\alpha_0(N-1)} \\ c_{1,0} & c_{1,0}\mathrm{e}^{\alpha_1} & \cdots & c_{1,0}\mathrm{e}^{\alpha_1(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{P,0} & c_{P,0}\mathrm{e}^{\alpha_P} & \cdots & c_{P,0}\mathrm{e}^{\alpha_P(N-1)} \end{pmatrix}$$

- *Goal*: **C** to have orthonormal rows
  - Orthogonal $\alpha_m = j\omega_m = j\frac{2\pi m}{N}$
  - Orthonormal $|c_{m,0}| = 1$
    - Then, we have a DFT like transform

$$\sum_n c_{k,n} c_{l,n}^* = c_{k,0} c_{l,0}^* \sum_n \mathrm{e}^{\frac{j2\pi k}{N}} \mathrm{e}^{-\frac{j2\pi l}{N}} = |c_{k,0}|^2 \delta_{k,l}$$

Modifying E-Splines (ii)

- Orthonormality condition $|c_{m,0}| = 1$
  - For any exponential reproducing kernel we can show

$$c_{m,0}\int_{-\infty}^{\infty} \mathrm{e}^{-\alpha_m t}\varphi(t)\mathrm{d}t = 1$$

    - This means that the coefficients $c_{m,0}$ are related to the *Laplace transform* of the kernel $\varphi(t)$ at α_m.

  - Then, using $\alpha_m = j\omega_m = j\frac{2\pi m}{N}$

    We identify that $c_{m,0}$ is the inverse of the *Fourier transform* of the kernel at ω_m. **Therefore**

$$|c_{m,0}| = 1 \quad \Leftrightarrow \quad |\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m)\hat{\beta}_{\vec{\alpha}_P}(\omega_m)| = 1$$

Modifying E-Splines (iii)

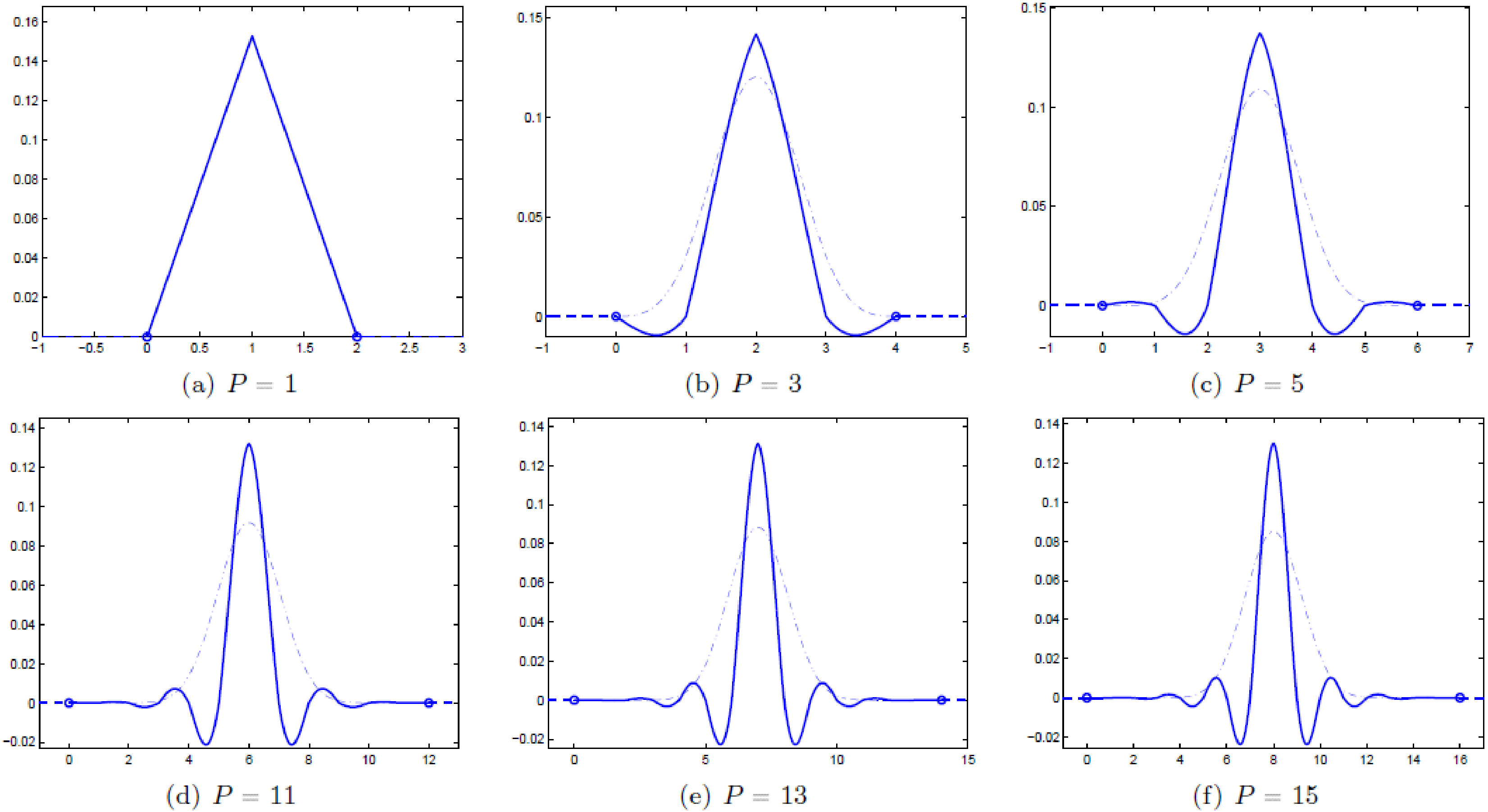
- The new condition $|\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m)\hat{\beta}_{\vec{\alpha}_P}(\omega_m)| = 1$ can be satisfied by choosing

$$\hat{\varphi}(\omega) = \hat{\beta}_{\vec{\alpha}_P}(\omega) \sum_{i=0}^{P-1} d_i (j\omega)^i$$

  - This means that we design *γ(t)* to be a polynomial that interpolates $(\omega_m, |\hat{\beta}_{\vec{\alpha}_P}(\omega_m)|^{-1})$
  - The time domain expression is a linear combination of derivatives $\varphi(t) = \sum_{i=0}^{P-1} d_i \beta_{\vec{\alpha}_P}^{(i)}(t)$ of the E-Spline
  - These functions have the characteristics of being of *maximum order* P and *minimum support* (MOMS)

Modifying E-Splines (iv)

- Kernel examples

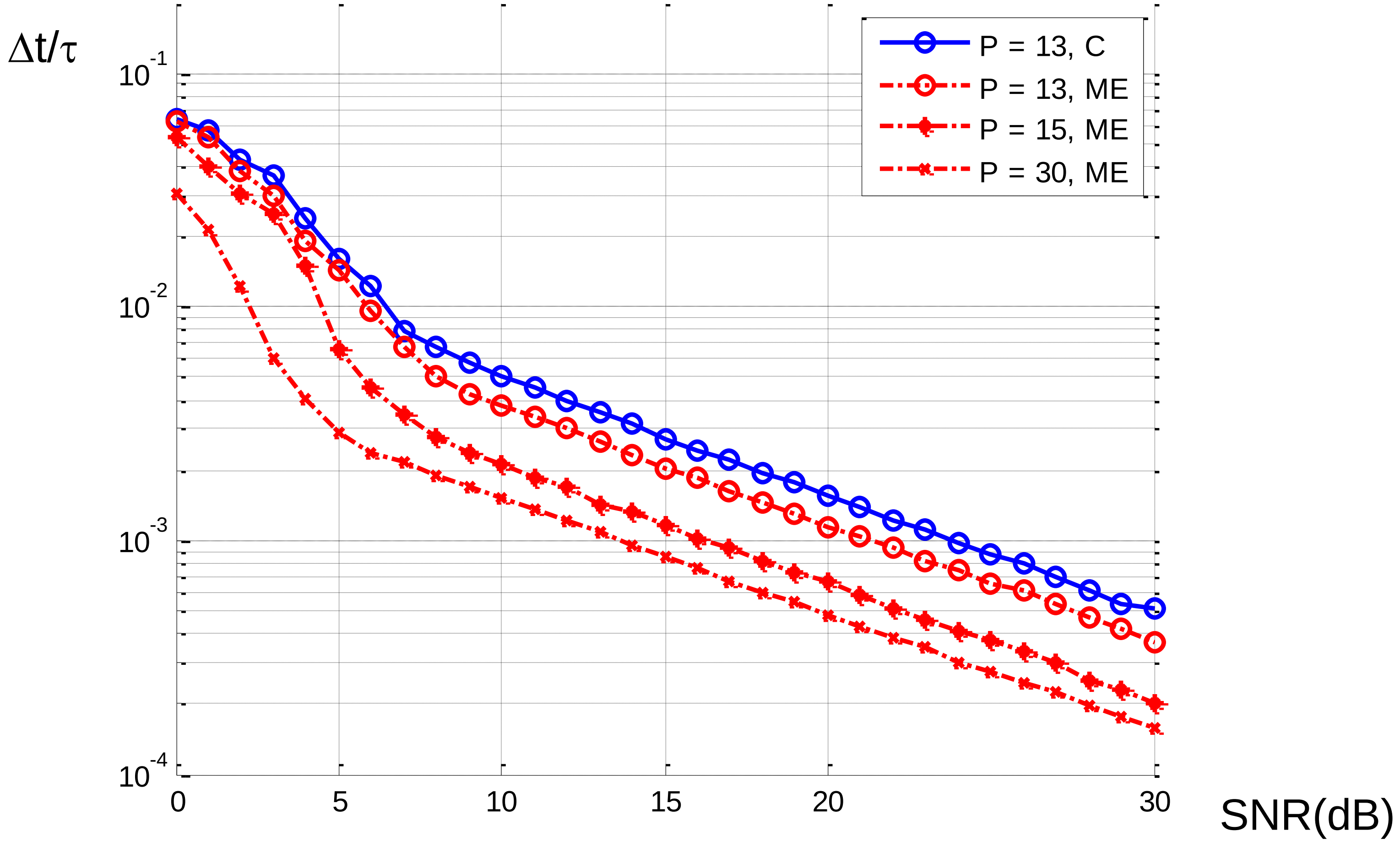


(a) $P = 1$ (b) $P = 3$ (c) $P = 5$

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

Modifying E-Splines (v)

- Simulations (K = 2 Diracs , N = 31 samples)



Conclusions

- Motivation for FRI theory
  - Sample & perfectly reconstruct continuous-time sparse signals
  - Using appropriate kernels
- Exponential reproducing kernels
  - Flexible tool to accommodate existing acquisition devices (rational FT, lens psf, …)
  - Can be modified to satisfy further conditions (MOMS)

Conclusions

- **Noisy FRI scenario**
  - *Prewhitening* to account for coloured noise
    - Standard approach
    - Doesn't perform as well as expected
  - More powerful and general approach: *Modify* kernels
    - Performance is optimal
    - Idea behind is preserve properties of noise (AWGN)
- Future work
  - How can we make default E-Splines behave optimally?

Questions

Imperial College London

- Review of Finite Rate of Innovation
  - Motivation for FRI
  - Parametric signals
  - Appropriate sampling kernels
  - Sample & reconstruct a train of Diracs
- The noisy scenario
  - A subspace approach
  - Prewhitening
  - Modified E-Spline kernels