Finding the Minimum Rate of Innovation in the Presence of Noise

Christopher Gilliam and Thierry Blu

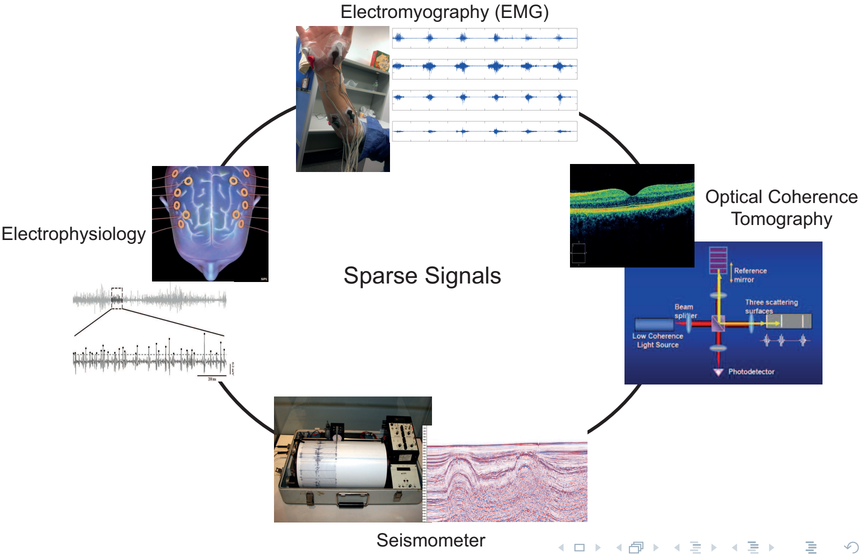
Department of Electronic Engineering
The Chinese University of Hong Kong

23rd March 2016

Outline

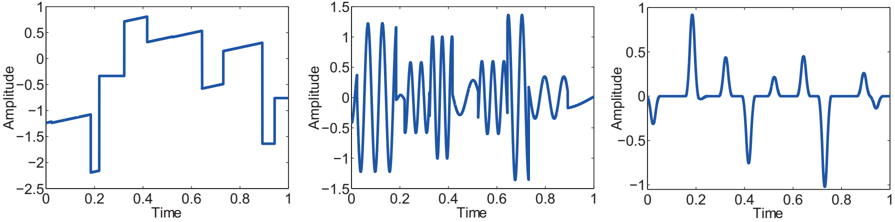
1. Signal Acquisition $\Longrightarrow$ Real World to Digital World
2. Signals with Finite Rate of Innovation (FRI)
   - General Framework
   - Reconstruction Procedure
   - Model Mismatch
3. Determining the Minimum Rate of Innovation
   - Model-Fitting Approach to FRI
   - Fitting based on a MSE Budget
   - Fast Estimation Algorithm
4. Simulation Results
5. Conclusion

Acquiring Sparse Signals



Representing Sparsity with Finite Rate of Innovation

Sparse signals $\iff$ FRI signals



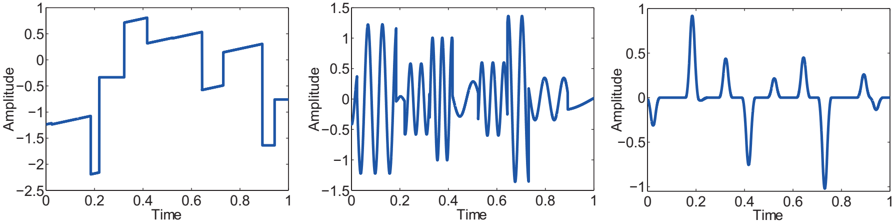
(a) Piecewise Polynomials (b) Piecewise Sinusoids (c) Stream of Pulses

Characteristics:

Finite number of sparse, unpredictable, parameters

Representing Sparsity with Finite Rate of Innovation

Sparse signals $\Longleftrightarrow$ FRI signals



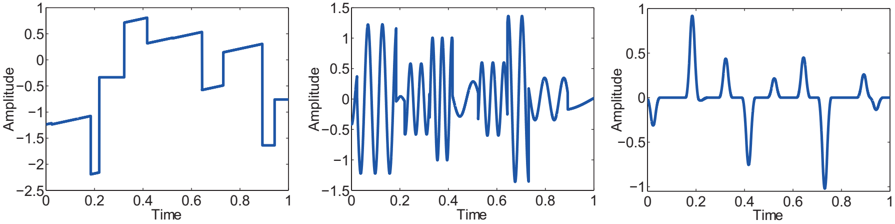
(a) Piecewise Polynomials (b) Piecewise Sinusoids (c) Stream of Pulses

Characteristics:

$$\underbrace{\text{Finite number of sparse, unpredictable, parameters}}_{\text{Innovations of the Signals}}$$

Representing Sparsity with Finite Rate of Innovation

Sparse signals $\Longleftrightarrow$ FRI signals



(a) Piecewise Polynomials (b) Piecewise Sinusoids (c) Stream of Pulses

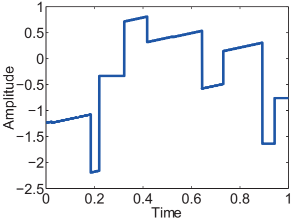
Applications in Bio-medics:

- Spike Estimation in Neurophysiological Data[1]
- Fetal heart rate detection[2]

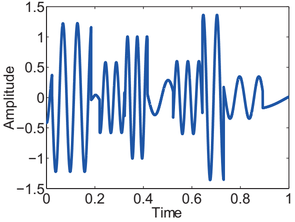
[1] J. Oñativia, R. Schultz & P.L. Dragotti, '*A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging*', J. Neural Engineering, 2013.
[2] A Nair & P. Marziliano, '*Fetal heart rate detection using VPW-FRI*', ICASSP, 2014

Representing Sparsity with Finite Rate of Innovation

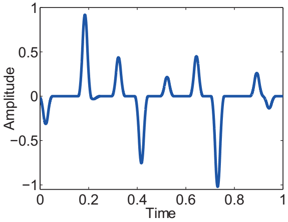
Sparse signals $\Longleftrightarrow$ FRI signals



(a) Piecewise Polynomials



(b) Piecewise Sinusoids



(c) Stream of Pulses

Two important questions...

- How to find the parameters?
- How many parameters to find? $\Longleftrightarrow$ Rate of Innovation

General FRI Framework

All FRI signals can be reduced to[1,2]....

$$\underbrace{\text{Sum of Diracs}}_{\sum_{k=1}^{K} x_k \delta(t - t_k)} \quad \underset{\text{Transform}}{\overset{\text{Linear}}{\Longleftrightarrow}} \quad \underbrace{\text{Sum of Sinusoids}}_{\sum_{k=1}^{K} x_k \mathrm{e}^{-j\omega t_k}}$$

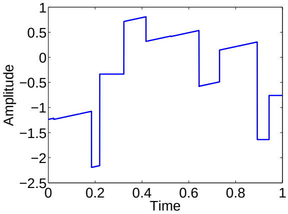
[1] P.L. Dragotti, M. Vetterli & T. Blu '*Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix*', IEEE Trans. Signal Processing, 2007.
[2] J.A. Urigüen, T. Blu & P.L. Dragotti, '*FRI sampling with arbitrary kernels*', IEEE Trans. Signal Processing, 2013.

General FRI Framework

All FRI signals can be reduced to[1,2]....

Sum of Diracs $\underbrace{}$ $\xLeftrightarrow[\text{Transform}]{\text{Linear}}$ Sum of Sinusoids

$$\sum_{k=1}^{K} x_k \delta(t - t_k) \qquad \sum_{k=1}^{K} x_k \mathrm{e}^{-j\omega t_k}$$



Finite Differences:

Discontinuities $\Longrightarrow$ Diracs

[1] P.L. Dragotti, M. Vetterli & T. Blu '*Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix*', IEEE Trans. Signal Processing, 2007.
[2] J.A. Urigüen, T. Blu & P.L. Dragotti, '*FRI sampling with arbitrary kernels*', IEEE Trans. Signal Processing, 2013.

General FRI Framework

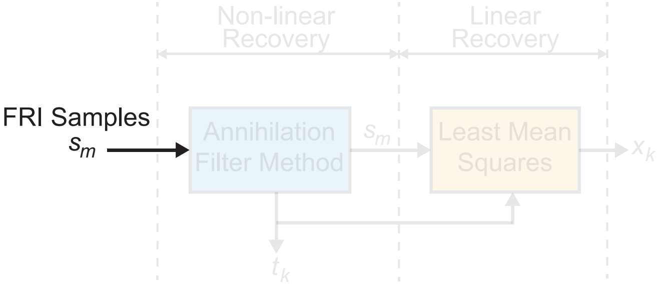
All FRI signals can be reduced to[1,2]....

$$\underbrace{\text{Sum of Diracs}}_{\sum_{k=1}^{K} x_k \delta(t - t_k)} \xLeftrightarrow[\text{Transform}]{\text{Linear}} \underbrace{\text{Sum of Sinusoids}}_{\sum_{k=1}^{K} x_k \mathrm{e}^{-j\omega t_k}}$$

↬ Estimate FRI parameters: $\begin{cases} x_k & \text{Amplitudes} \\ t_k & \text{Positions} \end{cases}$

[1] P.L. Dragotti, M. Vetterli & T. Blu '*Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix*', IEEE Trans. Signal Processing, 2007.
[2] J.A. Urigüen, T. Blu & P.L. Dragotti, '*FRI sampling with arbitrary kernels*', IEEE Trans. Signal Processing, 2013.

Reconstruction Procedure

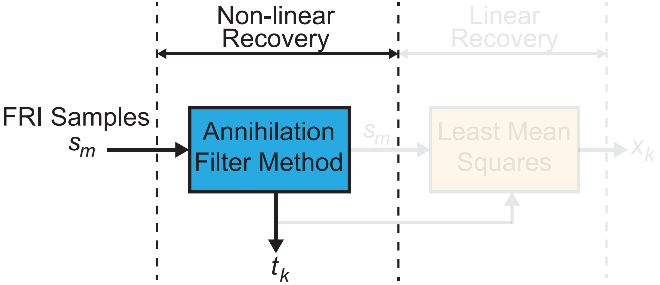


Uniformly sample:

FRI samples: $$\underbrace{s_m = \sum_{k=1}^{K} x_k \mathrm{e}^{-j2\pi m t_k}}_{\text{Sum of Sinusoids}}$$

N samples in total.

Reconstruction Procedure

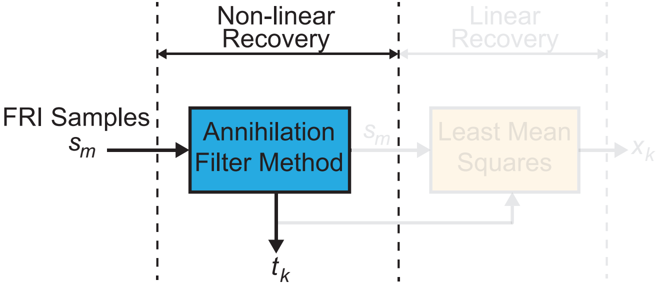


Step 1) Non-linear Recovery of t_k

Determine filter $H(z) = \sum_{k=0}^{K} h_k z^{-k}$ such that $\underbrace{h_m * s_m = 0}_{\text{Annihilation}}$

↬ Roots of filter H define the positions t_k

Reconstruction Procedure

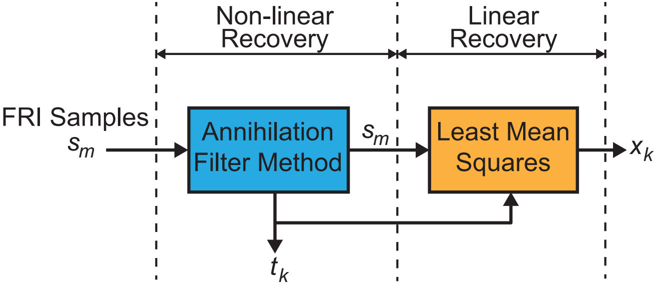


In more detail:

$$h_m * s_m = 0 \quad \Longleftrightarrow \quad \underbrace{\begin{bmatrix} s_K & s_{K-1} & \cdots & s_0 \\ s_{K+1} & s_K & \cdots & s_1 \\ \vdots & \vdots & \ddots & \vdots \\ s_{N-1} & s_{N-2} & \cdots & s_{N-K-1} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{K+1} \end{bmatrix}}_{\mathbf{h}} = 0$$

Solve Annihilation Equation: $\mathbf{A}\mathbf{h} = 0 \longrightarrow$ Requires $N = 2K$ samples

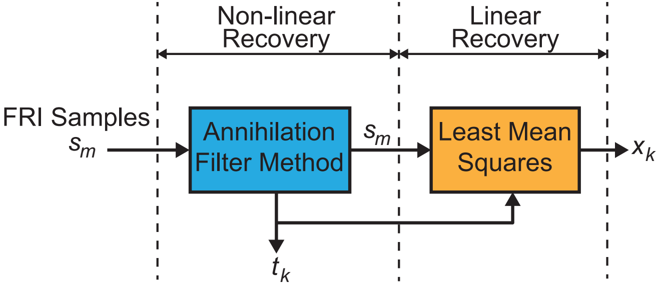
Reconstruction Procedure



Step 2) Linear Recovery of x_k

Using $t_k \implies$ Determine x_k via least mean squares

Reconstruction Procedure



Estimating the rate of innovation:

1. Over-estimate rate $\implies$ Set $K_{\text{est}} = N/2$
2. Build the annihilation matrix $\mathbf{A}$
3. Rank of $\mathbf{A}$ = Actual K

Model Mismatch

An imperfect world:

Sample Corruption: $\tilde{s}_m = s_m + \epsilon_m$ therefore $\tilde{\mathbf{A}}\mathbf{h} \neq 0$

Model Mismatch

An imperfect world:

Sample Corruption: $\tilde{s}_m = s_m + \epsilon_m$ therefore $\tilde{\mathbf{A}}\mathbf{h} \neq 0$

Many methods to estimate sinusoids in noise:

For example: Total Least Squares,
Cadzow's Method[1],
Matrix Pencil[2],
⋮

[1] J. Cadzow, '*Signal Enhancement - A composite property mapping algorithm*', IEEE Acoust., Speech and Processing, 1988
[2] Y. Hua & T.K. Sarkar, '*Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise*', IEEE Acoust., Speech and Processing, 1990

Model-Fitting approach

Noiseless Conditions:

Inverse Fourier Transform of the sum of sinusoids s_m

$\Updownarrow$

Expressed as a ratio of two Polynomials, P and H[1]

[1] C. Gilliam & T. Blu, '*Fitting instead of Annihilation: Improved recovery of noisy FRI signals*', ICASSP, 2014

Model-Fitting approach

Noiseless Conditions:

Inverse Fourier Transform of the sum of sinusoids s_m

$\Updownarrow$

Expressed as a ratio of two Polynomials, P and H[1]

In other words:

$$s_m \xleftrightarrow[\text{Transform}]{\text{Inverse Fourier}} y_n = \frac{P\left(e^{j\omega_n}\right)}{H\left(e^{j\omega_n}\right)}$$

where

- $P\left(e^{j\omega_n}\right)$ is a polynomial of order $K-1$
- $H\left(e^{j\omega_n}\right)$ is the annihilation filter, a polynomial of order K
- $\omega_n = 2\pi n/N$

[1] C. Gilliam & T. Blu, '*Fitting instead of Annihilation: Improved recovery of noisy FRI signals*', ICASSP, 2014

Model-Fitting approach

Noiseless Conditions:

Inverse Fourier Transform of the sum of sinusoids s_m

$\Updownarrow$

Expressed as a ratio of two Polynomials, P and $H^{[1]}$

In other words:

$$s_m \xleftrightarrow[\text{Transform}]{\text{Inverse Fourier}} y_n = \frac{P\left(e^{j\omega_n}\right)}{H\left(e^{j\omega_n}\right)}$$

$\looparrowright$ Samples y_n completely defined by the coefficients of the polynomials

[1] C. Gilliam & T. Blu, '*Fitting instead of Annihilation: Improved recovery of noisy FRI signals*', ICASSP, 2014

Model-Fitting approach

Noiseless Conditions:

Inverse Fourier Transform of the sum of sinusoids s_m

$\Updownarrow$

Expressed as a ratio of two Polynomials, P and H[1]

In other words:

$$s_m \xleftrightarrow[\text{Transform}]{\text{Inverse Fourier}} y_n = \frac{P\left(e^{j\omega_n}\right)}{H\left(e^{j\omega_n}\right)}$$

↬ Samples y_n completely defined by the coefficients of the polynomials

↬ Departs from the Annihilation Equation!!

[1] C. Gilliam & T. Blu, '*Fitting instead of Annihilation: Improved recovery of noisy FRI signals*', ICASSP, 2014

Model-Fitting approach

Noiseless Conditions:

Inverse Fourier Transform of the sum of sinusoids s_m

$\Updownarrow$

Expressed as a ratio of two Polynomials, P and H [1]

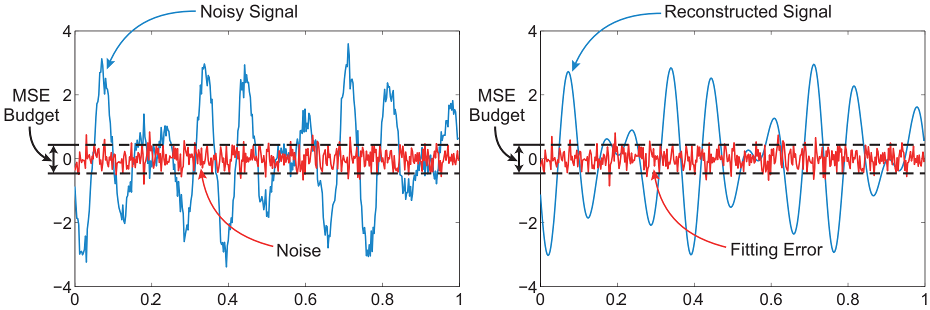
In other words:

$$s_m \xleftrightarrow[\text{Transform}]{\text{Inverse Fourier}} y_n = \frac{P\left(e^{j\omega_n}\right)}{H\left(e^{j\omega_n}\right)}$$

Model Mismatch $\Longrightarrow$ Fit ratio model to the noisy samples $\tilde{y}_n$

[1] C. Gilliam & T. Blu, '*Fitting instead of Annihilation: Improved recovery of noisy FRI signals*', ICASSP, 2014

Fitting using a MSE Budget

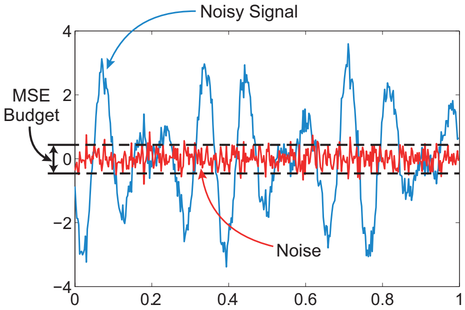


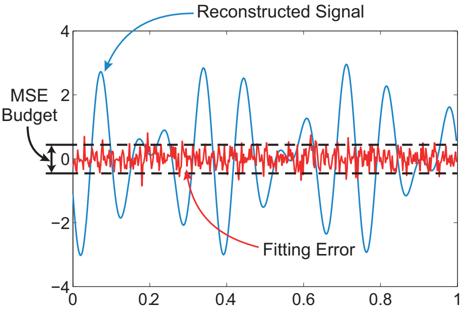
Central Concept:

Fit model to the noisy data until the MSE budget has been satisfied

$$\looparrowright \quad \|\text{Data} - \text{Model}\|_2^2 \;\leq\; \text{MSE Budget}$$

Fitting using a MSE Budget





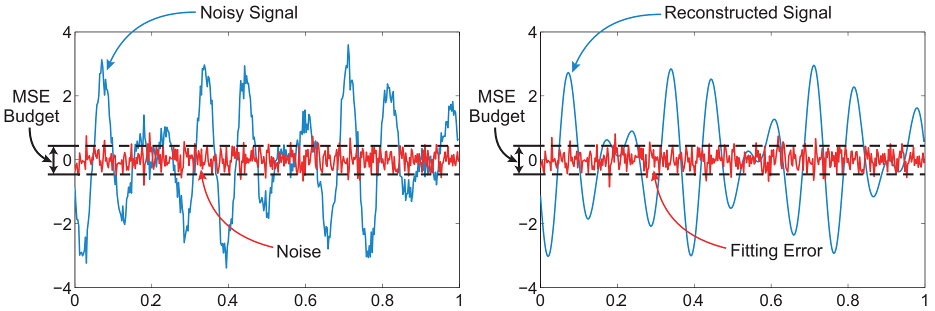
Central Concept:

Fit model to the noisy data until the MSE budget has been satisfied

$$\looparrowright \quad \|\text{Data} - \text{Model}\|_2^2 \leq \text{MSE Budget}$$

$\looparrowright$ Treat all models that satisfy the budget as equal

Fitting using a MSE Budget

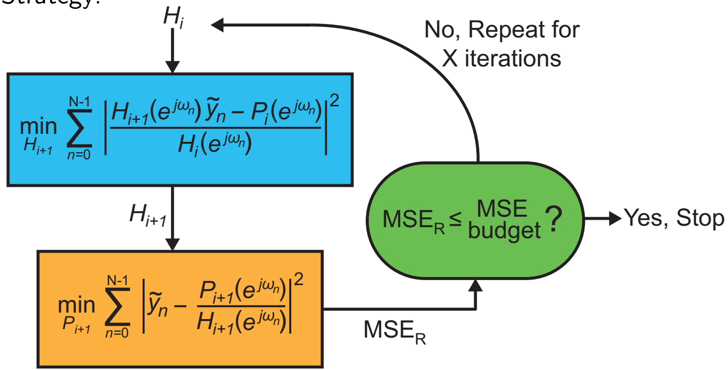


Obtaining the MSE Budget:

- ↬ Noise Level of Acquisition Device
- ↬ Estimated from Noisy Signal

Solving the Model-Fitting

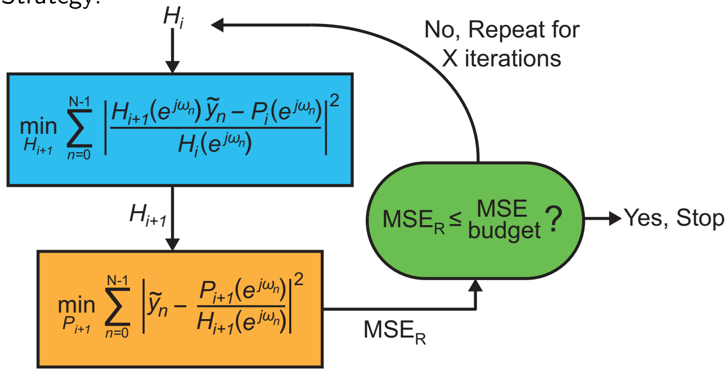
Iterative Strategy:



1. Update denominator $H_{i+1} \implies$ Solve linear system of equations

Solving the Model-Fitting

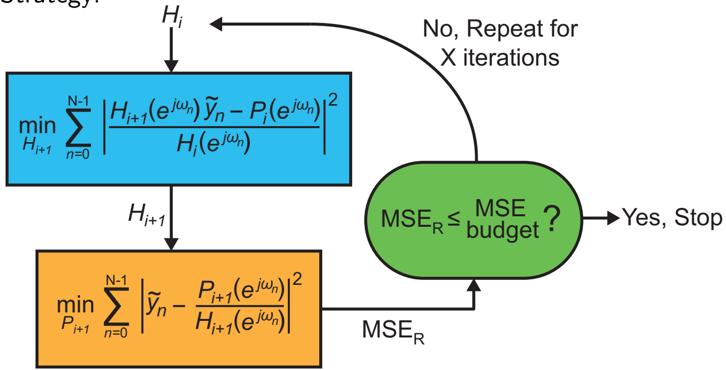
Iterative Strategy:



1. Update denominator H_{i+1} $\Longrightarrow$ Solve linear system of equations
2. Update numerator P_{i+1} $\Longrightarrow$ Obtain model update

Solving the Model-Fitting

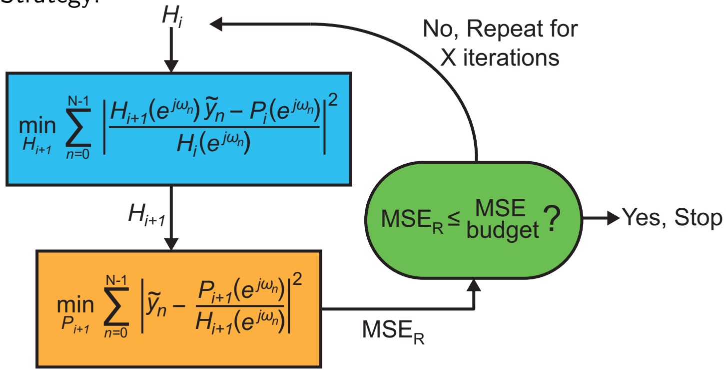
Iterative Strategy:



1. Update denominator $H_{i+1} \implies$ Solve linear system of equations
2. Update numerator $P_{i+1} \implies$ Obtain model update
3. Check model against MSE budget

Solving the Model-Fitting

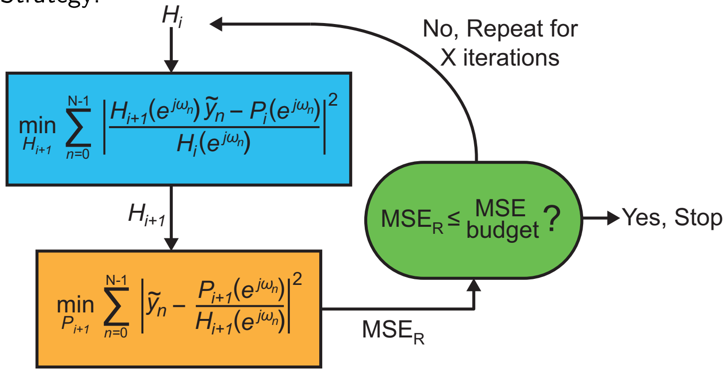
Iterative Strategy:



Based on MSE budget $\Longrightarrow$ Not based on convergence

Solving the Model-Fitting

Iterative Strategy:

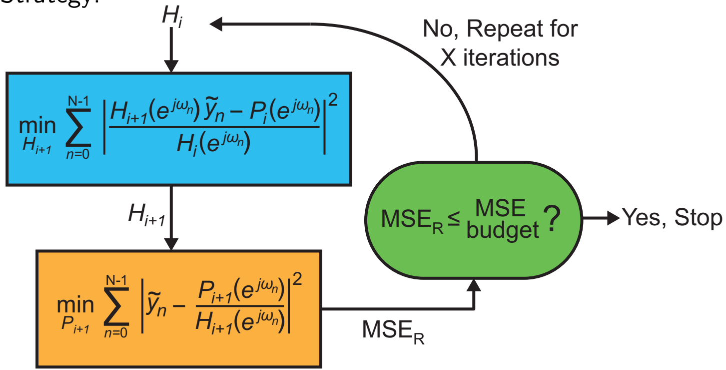


Reaching X iterations...

↬ Repeat with another random initialisation of H_0

Solving the Model-Fitting

Iterative Strategy:



Reaching X iterations...

↬ Repeat with another random initialisation of H_0

Costly?....

Solving the Model-Fitting

No $\implies$ Very few initialisations are needed

Example of Robustness

FRI signal with $K = 6$ Diracs and $N = 51$ samples:

30 random initialisations $\implies$ Algorithm failure rate $= 0.015\%$
50 random initialisations $\implies$ Algorithm failure rate $= 0.0046\%$

$\looparrowright$ 15 random initialisations required to succeed in 99.9% cases
$\looparrowright$ 3 random initialisations required to succeed in 95% cases

*Results obtained using 500,000 realisations

Solving the Model-Fitting

No $\implies$ Very few initialisations are needed

Example of Robustness

FRI signal with $K = 6$ Diracs and $N = 51$ samples:

30 random initialisations $\implies$ Algorithm failure rate $= 0.015\%$
50 random initialisations $\implies$ Algorithm failure rate $= 0.0046\%$

$\looparrowright$ 15 random initialisations required to succeed in 99.9% cases
$\looparrowright$ 3 random initialisations required to succeed in 95% cases

*Results obtained using 500,000 realisations

Finding the Rate of Innovation

Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Finding the Rate of Innovation

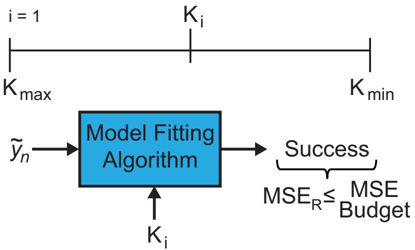
Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation



Finding the Rate of Innovation

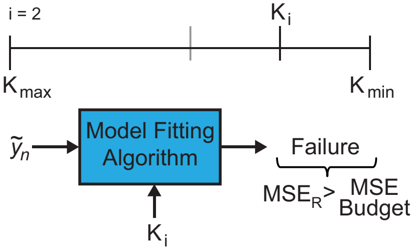
Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation



Finding the Rate of Innovation

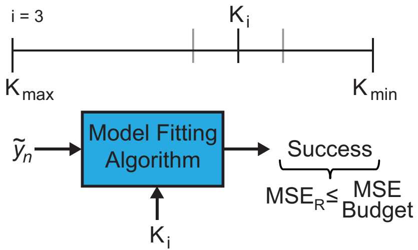
Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation



Finding the Rate of Innovation

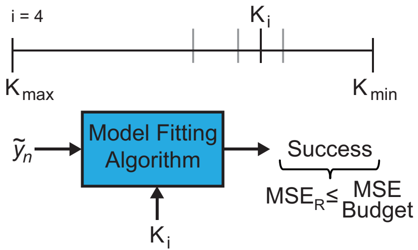
Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation



Finding the Rate of Innovation

Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation

$\looparrowright$ Possible to estimate a lower rate of innovation than the original

Finding the Rate of Innovation

Central Concept

Given a Rate of Innovation and MSE Budget:

Fitting algorithm either succeeds or fails $\Longrightarrow$ Binary outcome

Dichotomic Algorithm:

Binary search to determine the minimum K, i.e. rate of innovation

Concept of Parsimony:

The sparsest model is the most appropriate to represent the data

Simulation Results

Set-up:

- FRI signal with $K = 12$ Diracs and $N = 97$ samples
- Number of random initialisations = 50

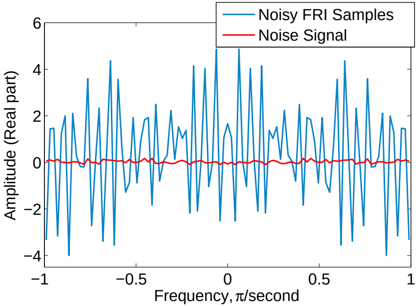
Compare against the Bayesian Information Criterion (BIC)[1]:

- Calculate the BIC for every K
- Choose K with the lowest BIC value

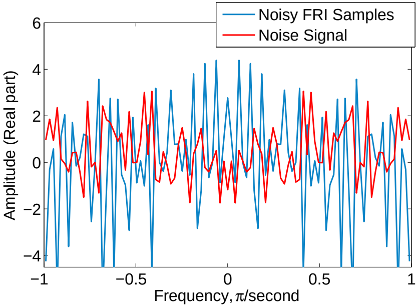
[1] P. Stoica & Y. Selen, '*Model-order selection*', IEEE Signal Processing Mag., 2004.

Simulation Results

Two sets of noisy conditions:



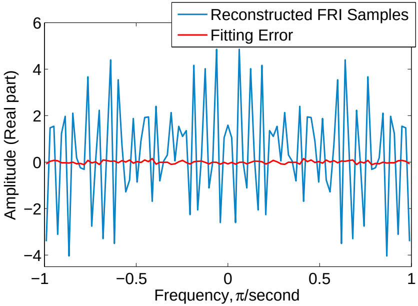
(a) Input SNR $= 30\text{dB}$



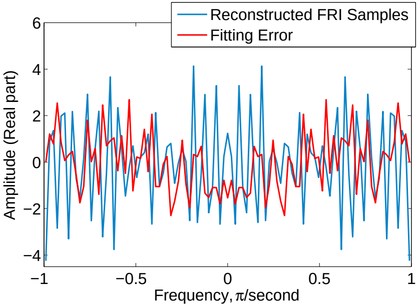
(b) Input SNR $= 5\text{dB}$

Simulation Results

Results of Dichotomic Algorithm:



(a) Input SNR $= 30$dB

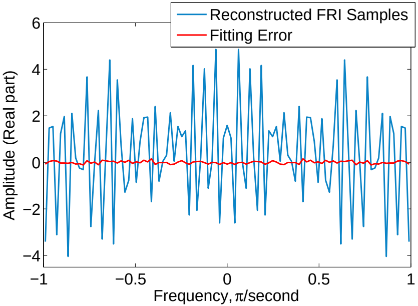


(b) Input SNR $= 5$dB

$$\text{Fit to MSE budget} \implies \underbrace{\text{Fitting Error}}_{\text{Noisy Data - Model}} \sim \text{Noise Signal}$$

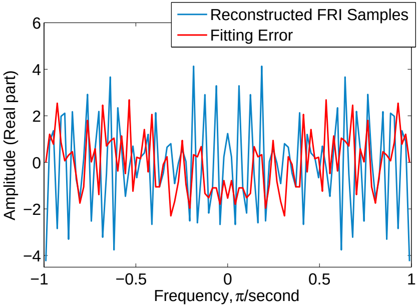
Simulation Results

Results of Dichotomic Algorithm:



(a) Input SNR $= 30$dB

↬ Output SNR = 34.2dB

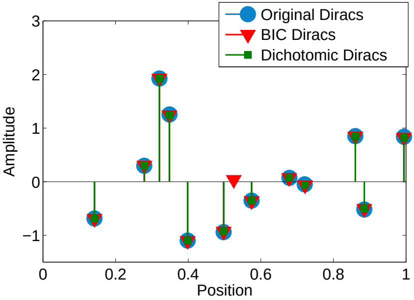


(b) Input SNR $= 5$dB

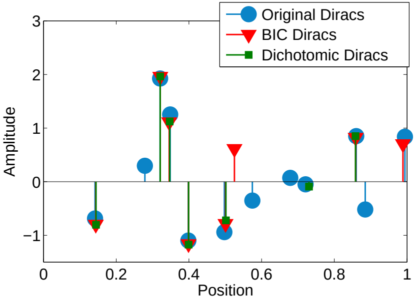
↬ Output SNR = 7.3dB

Simulation Results

Sparse parameters:



(a) Input SNR $= 30\text{dB}$



(b) Input SNR $= 5\text{dB}$

↬ BIC = 13 Diracs
↬ Dichotomic = 12 Diracs

↬ BIC = 8 Diracs
↬ Dichotomic = 7 Diracs

Conclusion

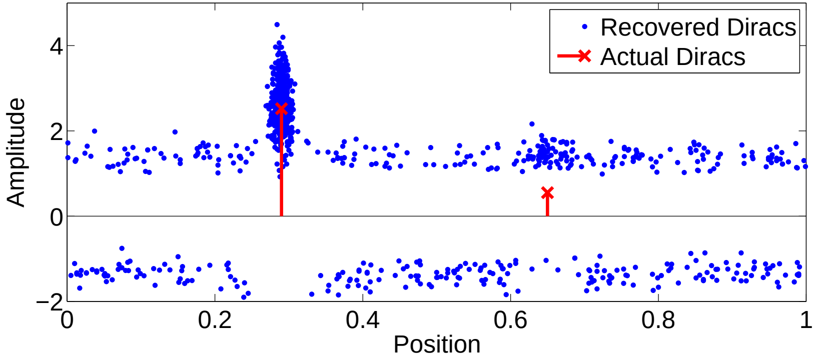
- Introduced the concept of FRI recovery based on a MSE budget
  - Model representation of the FRI samples $\Longrightarrow$ A ratio of polynomials
  - Fit model to noisy data until MSE budget is satisfied
  - Fitting algorithm is fast, accurate and robust
- Presented framework for finding rate of innovation
  - Outcome of model-fitting is binary $\Longrightarrow$ Succeeds or fails
  - Design binary search for the rate of innovation
  - Parsimony $\Longrightarrow$ Sparsest model is best
- Results demonstrating the advantages of the algorithm
  - Obtained FRI signals that met the MSE budget
  - Sparse parameters consistent with original

The End

Thank you for listening

Possibility of Losing Innovation

Recovering 2 Diracs in heavy noise, $\mathrm{SNR} = 0\mathrm{dB}$:



↬ Estimation using Maximum Likelihood over 500 realisations