REGRESSION METHOD

Regression Basics

Wiener-Kolgomorov filter (*)

- Estimating which features of h(t) can be predicted by its correlation with a set of witness channels

- Simple case: considering one auxiliary channel x(t):
  - s: prediction
  - L: filter length

$$s_i = \left(\sum_{j=-L}^{L} a_j x_{i+j} \right)$$

a: predicting filter
i, j : time indexes

- Least square minimization of the residuals:
  - N: filter training length

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} \left[h_i - \left(\sum_{j=-L}^{L} a_j x_{i+j} \right) \right]^2$$

* [Rev. Sci. Instrum. **83**, 024501 (2012)]

Solution

- The minimization of residual $\left(\frac{\delta e^2}{\delta a_k} = 0 \right)$ leads to:

$$\sum_{j=-L}^{L} a_j \left(\sum_{i=1}^{N} x_{i+j} x_{i+k} \right) = \left(\sum_{i=1}^{N} x_{i+k} h_i \right)$$

- In a matrix form:

$$R_{jk}^{xx} a = C^{hx}$$

$$R_{jk}^{xx} = \sum_{i=1}^{N} x_{i+j} x_{i+k} \qquad a = \begin{Bmatrix} a_{-L} \\ \cdots \\ a_0 \\ \cdots \\ a_L \end{Bmatrix} \qquad C_k^{hx} = \sum_{i=1}^{N} x_{i+k} h_i$$

Multiple witness channels

- Enhance regression **but** add noise to prediction

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} \left[h_i - \left(\sum_{j=-L}^{L} a_j x_{i+j} \right) - \left(\sum_{j=-L}^{L} b_j y_{i+j} \right) - \ldots \right]^2$$

- Minimization on each filter component (a_k, b_k, …) leads to:

$$\begin{pmatrix} R^{xx} & R^{yx} & \ldots \\ R^{xy} & R^{yy} & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \\ \ldots \end{pmatrix} = \begin{pmatrix} C^{hx} \\ C^{hy} \\ \ldots \end{pmatrix} \quad \begin{cases} R_{xy}^{jk} = \sum_{i=1}^{N} x_{i+j} y_{i+k} \\ C_{hx}^{k} = \sum_{i=1}^{N} h_i x_{i+k} \end{cases}$$

- Which is a similar form of the one-channel case, with M*(2L+1) components: $R\mathbf{a} = C$

Multi-linear correlation

- Some spectral features are described by correlation with a carrier line and low-frequencies
  - Side-bands formation

$$\sin(2\pi f_0 t)\sin(2\pi f_1 t) \propto \sin[2\pi(f_0 + f_1)t] + \sin[2\pi(f_0 - f_1)t]$$

- Cross-correlation matrix R can be constructed using also the multiplication of different witness channels:
  - describe up-conversion of low frequency signals

$$x_{sideband} = x_{carrier} * x_{lowfreq}$$

Regulators

- R matrix can be written considering related eigen-values λ ($\lambda_1 > \lambda_2 > \ldots > \lambda_N$) and eigenvectors P

$$P^{-1}RP = \Lambda$$

$$a = P\Lambda^{-1}P^{-1}C$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_N \end{pmatrix}$$

Typically, few eigenvalues are significant

- Regulators: Impose a threshold on eigen-values and select only the bigger eigen-values

$$\Lambda_r^{-1} = \left(\begin{array}{ccccccc} 1/\lambda_1 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & 1/\lambda_2 & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & 1/\lambda_{th} & \ldots & \ldots & \ldots \\ \hline \ldots & \ldots & \ldots & \ldots & \lambda' & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \lambda' \end{array}\right)$$

Different values of λ' can be considered

- Hard: $\lambda' = 0$
- Soft: $\lambda' = 1/\lambda_{th}$
- Mild: $\lambda' = 1/\lambda_1$

Regression in cWB

- Calculation of filter in small frequency bands
  - Better description of noise features in each band
  - Reduce computational complexity

- Introduction of 90 degree phase data
  - More complete description of noise features
  - Double filter length

$$S_k = \sum_j (A_j X_{j+k}) = \sum_j (a_j + i\tilde{a}_j)(x_{j+k} + i\tilde{x}_{j+k})$$

x -> 0 phase
$\tilde{x}$ -> 90 phase

- Prediction can be constructed from only 0 or 90 degree phase, or both

Regression parameters

- Filter length L -> ***setFilter(L)***

- Regulators -> ***solve(th, nE, c)***
  - Eigen-values thresholds
    - Select $\lambda_k > th$
    - Select $th < nE$
  - Regulators
    - c='h' -> hard: $\lambda' = 0$
    - c='m' -> mild: $\lambda' = l_1$
    - c='s' -> soft: $\lambda' = l_{th}$

$$\Lambda_r^{-1} = \left(\begin{array}{ccccccc} 1/\lambda_1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & 1/\lambda_2 & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1/\lambda_{th} & \cdots & \cdots & \cdots \\ \hline \cdots & \cdots & \cdots & \cdots & \lambda' & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \lambda' \end{array}\right)$$

Ranking

$$Rank(x) = \frac{1/N \sum_{i=1}^{N} s_i}{1/N \sum_{i=1}^{N} x_i}$$

Rank(x) gives the effective contribution of channel x to the prediction

- It can be used to evaluate the useful channel

- It is possible to set a threshold on rank on the most contributing channels
  - ***apply(th,c)***
    - th: threshold on the channel rank
    - c: calculate prediction from Inverse Wavelet transform starting from only 0, only 90, both phases

LPE filter

- When there are no witness channel describing spectral features, we can use target channel as "fake witness"

- In this case matrix R and vector C are slightly modified (R', C'):

$$R'_{ij} = \begin{cases} 0 & i = j \\ R_{ij} & i \neq j \end{cases}$$

$$C' = \begin{Bmatrix} C_{-L} \\ \dots \\ 0 \\ \dots \\ C_L \end{Bmatrix}$$