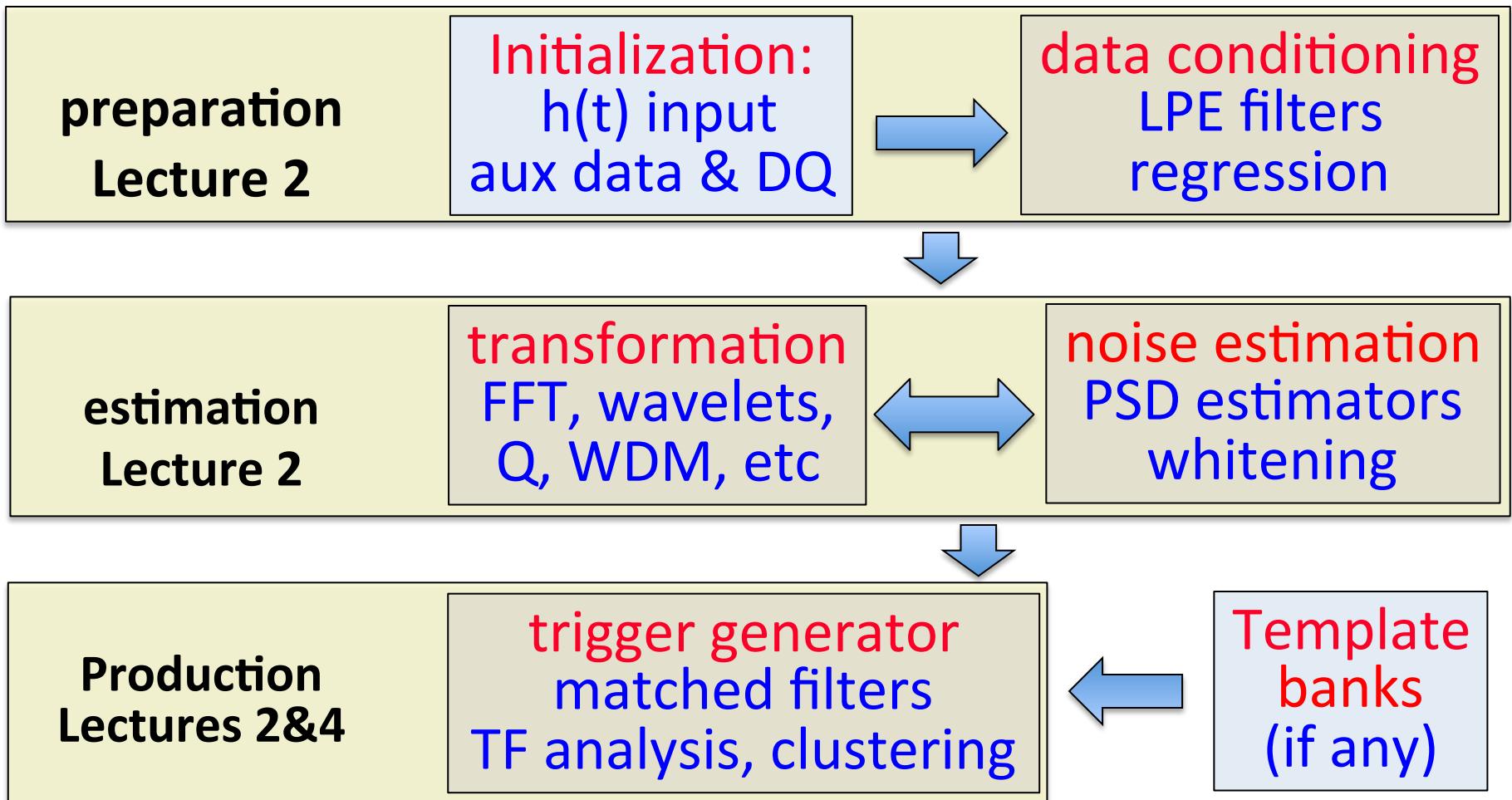




Stages of search pipeline: ETG

2



Event Trigger Generator (ETG): search preprocessor which uses basic signal processing algorithms with the goal to prepare data, produce intermediate data products (triggers), reduce complexity and processing time

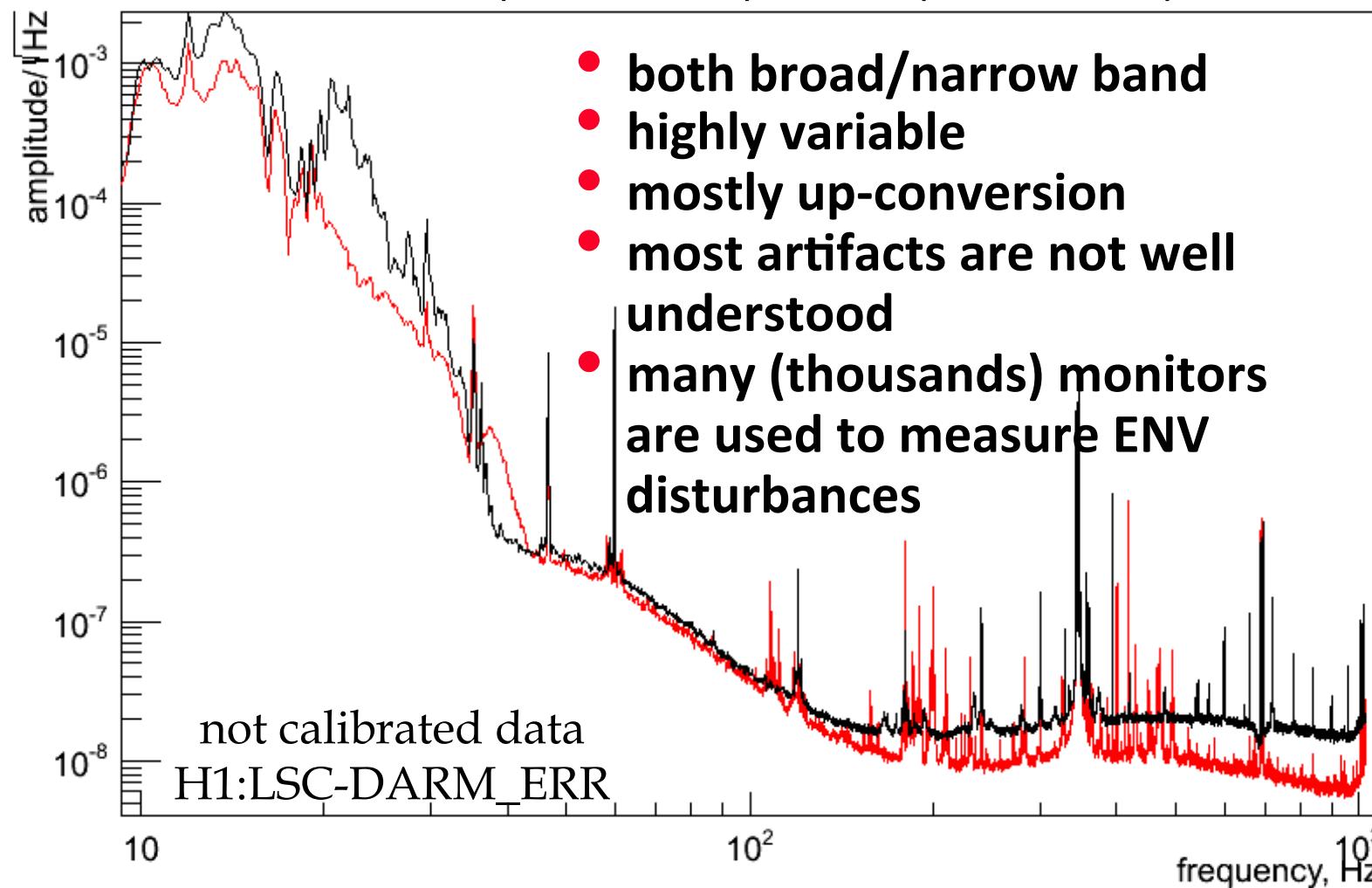




Instrumental/Environmental Noise

3

Snapshots of H1 GW channel data (15 min)
black S5(820707090), red S6(942451300)

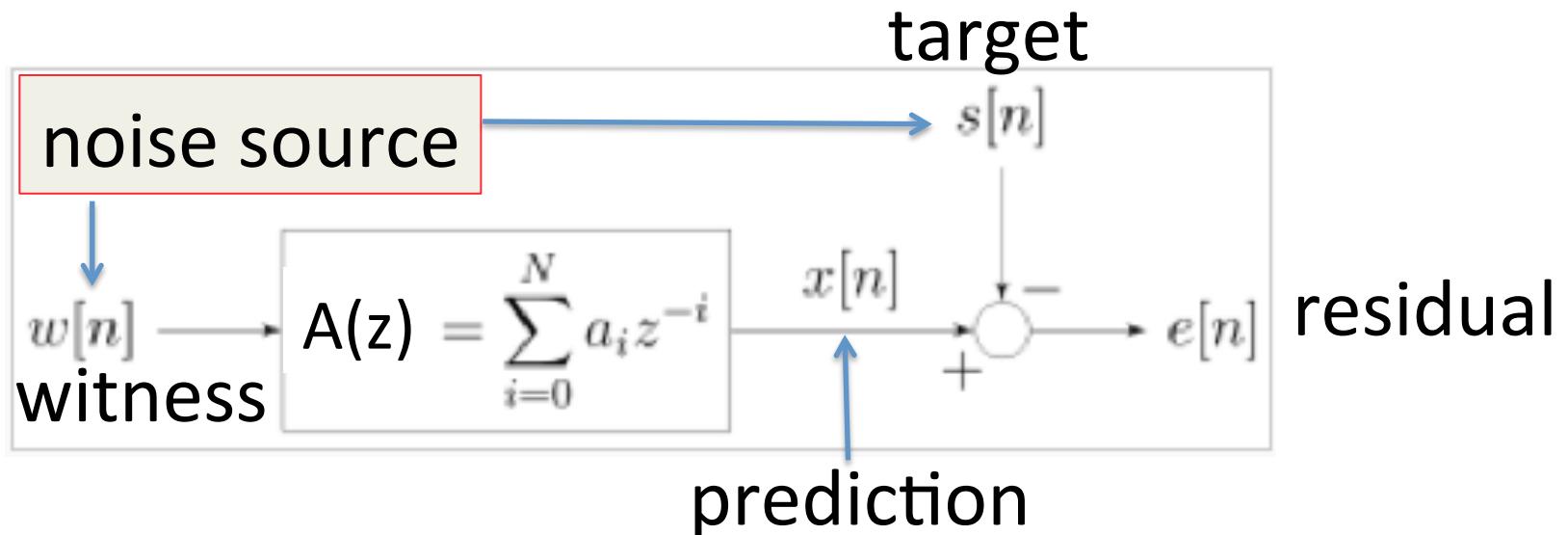


data conditioning – remove “predictable” noise to improve search sensitivity



Wiener(1949)-Kolmogorov(1941) Filter

4



- regression: target $s[n]$ can be predicted if there is a linear association with witness channel $w[n]$.
- find $A = \{a[0], \dots, a[K]\}$ by minimization of residual

$$\sum_{n=0}^{N-1} e^2[n] = \sum_{n=0}^{N-1} \left| s[n] - \sum_{k=0}^K a[k]w[i-k] \right|^2$$

- N – filter training length, K – filter length



Wiener-Hopf Equation

5

$$\begin{bmatrix} r_{ww}[0] & r_{ww}[1] & \dots & r_{ww}[K] \\ \dots & \dots & \dots & \dots \\ r_{ww}[K] & r_{ww}[K-1] & \dots & r_{ww}[0] \end{bmatrix} \begin{bmatrix} a[0] \\ \dots \\ a[K] \end{bmatrix} = \begin{bmatrix} c_{sw}[0] \\ \dots \\ c_{sw}[K] \end{bmatrix}$$

$R \cdot A = C$

- R is (Hermitian) Toeplitz matrix constructed from autocorrelation of w and C is a x-correlation function between s and w

$$r_{ww}[k] = \sum_{n=0}^{n=N} w[n]w[n-k] \quad c_{sw}[k] = \sum_{n=0}^{n=N} s[n]w[n-k]$$

- Complexity: to capture noise spectral features with the resolution df in a sampled (rate fs) LIGO signal need a filter with ~ 2 (fs/df) coefficients A. $K \sim 8000$ for $fs=4\text{kHz}$
 - solved by using efficient Levinson-Durbin algorithm [1]



Linear Predictors

6

- What if $w=s$? – predict s by using past (or future) samples

- forward prediction
(lattice prediction)
- backward prediction

$$x[n] = \sum_{k=1}^K a[k] s[n-k]$$

$$x[n-K] = \sum_{k=0}^{K-1} a[K-k] s[n-k]$$

$$a[0]=0 \rightarrow \begin{bmatrix} r_{ww}[0] & r_{ww}[1] & \dots & r_{ww}[K-1] \\ \dots & \dots & \dots & \\ r_{ww}[K-1] & r_{ww}[K-1] & \dots & r_{ww}[0] \end{bmatrix} \begin{bmatrix} a[1] \\ \dots \\ a[K] \end{bmatrix} = \begin{bmatrix} r_{sw}[1] \\ \dots \\ r_{sw}[K] \end{bmatrix}$$

- Efficient tool to remove spectral lines (power, violin, ...)
 - filter should be significantly longer than the GW transient duration $T \rightarrow K \gg T*f_s$, f_s – data sampling rate

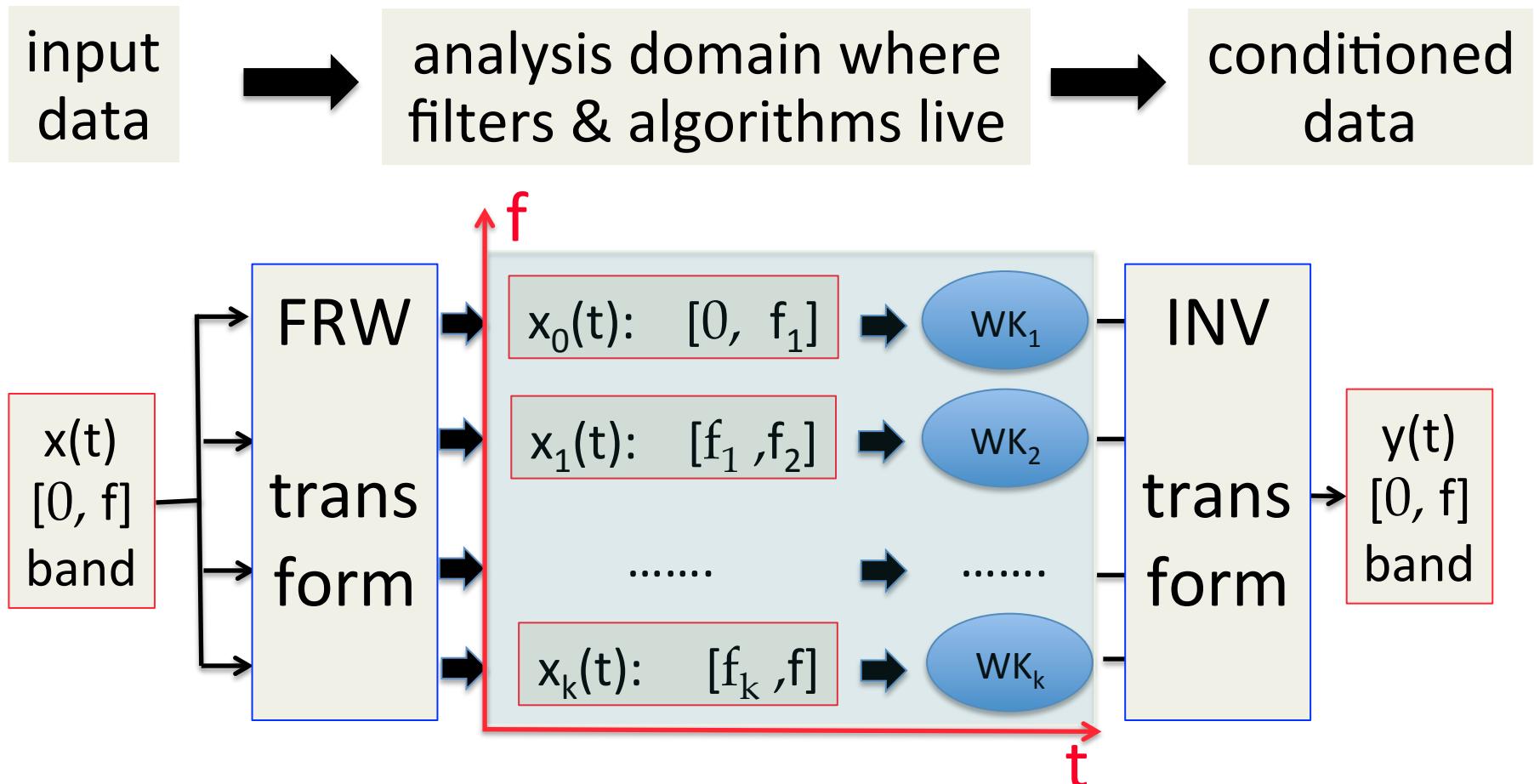


Practical Applications

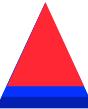
7

- **Simple case:** remove first 15 harmonics of power lines
 - re-sample data down to 1000Hz
 - select filter length: $K_o \sim 1000\text{Hz}/0.1\text{Hz} = 10000$
 - estimate auto&cross – correlation functions on selected data segment
 - solve Weiner-Hopf equations, find & apply filter
- **Problems:**
 - invert 10000×10000 matrix and find 10000 filter coefficients, while only ~ 30 parameters are relevant
 - 30 line parameters are spread among 10000 filter coefficients and masked by other interfering spectral features in 0-1000Hz band
- **Solution:** band-pass data into multiple (M) time series and construct individual WK filter for each one
 - filter length: $K_o \rightarrow K_o/M$, solve 15 WH equations of size $10000/M$
→ significantly reduce complexity and interference

Analysis Domain



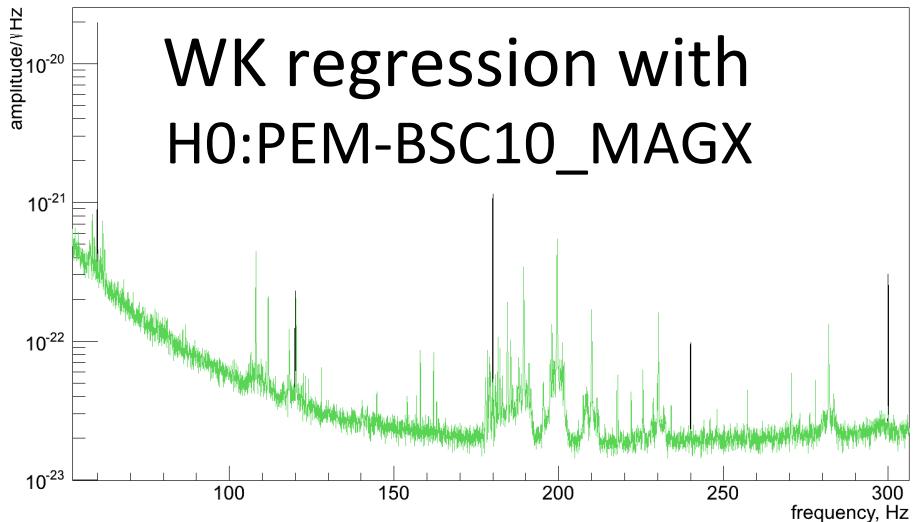
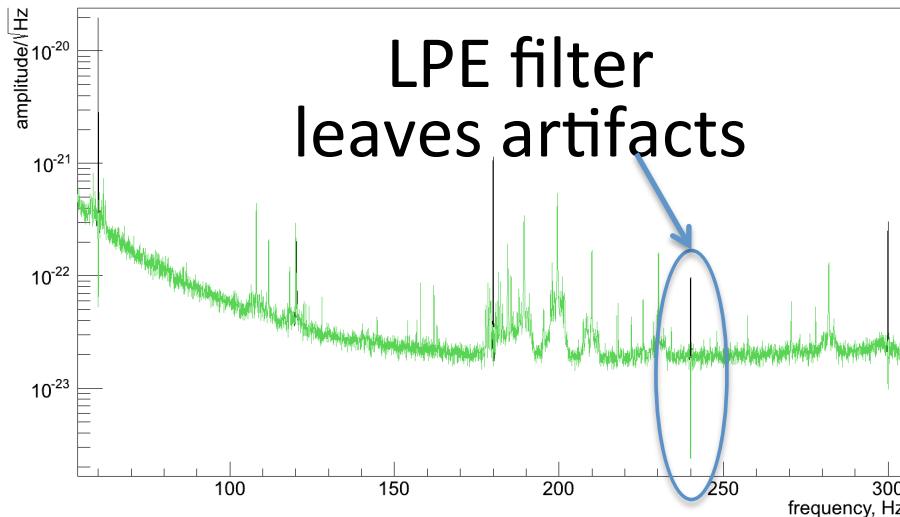
- For practical filter applications need an efficient & invertible transformation, which band-pass data into multiple time series in the analysis (TF) domain



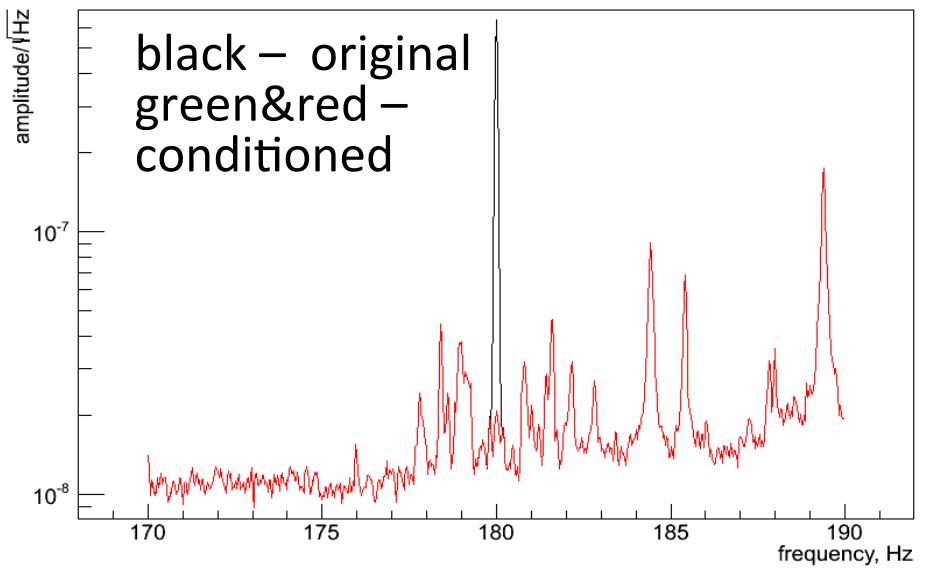
Power Lines

9

H1: 15min of S6 data



- removed by LPEF and WK regression (in WDM domain – 1Hz resolution) with power monitors or magnetometers (H0:PEM-BSC10_MAGX)
- WK can remove other cases of linear coupling, particularly broad-band
- How to identify and remove non-linear coupling?





Multiple Witness Channels

10

- $$\sum_n \left| s[n] - \sum_k a_w[k]w[n-k] - \sum_k a_u[k]u[n-k] - \sum_k a_v[k]v[n-k] \right|^2$$
- In general, $s[n]$, $w[n]$, $u[n]$, $v[n]$ and filters A are complex

$$\begin{bmatrix} R_{ww} & R_{wu} & R_{wv} \\ R_{uw} & R_{uu} & R_{uv} \\ R_{vw} & R_{vu} & R_{vv} \end{bmatrix} \begin{bmatrix} A_w \\ A_u \\ A_v \end{bmatrix} = \begin{bmatrix} c_{sw} \\ c_{su} \\ c_{sv} \end{bmatrix}$$

$$\begin{aligned} Rq_k &= \lambda_k q_k \\ Q &= \{q_1, \dots, q_K\} \\ \Lambda &= \text{diag}\{\lambda_1, \dots, \lambda_K\} \\ A &= Q^T \Lambda^{-1} Q C \end{aligned}$$

- Combine different monitors measuring noise sources at different locations, increase SNR and hence estimation of noise
- Significantly more complex problem
- Levinson-Durbin algorithm does not work - not a big loss!
 - need to do a complete eigenvalue analysis anyway (LD not stable)
 - formally WH always has a solution, but in fact it can be rand-deficient





Regulators

11

-L < k < L

$$\begin{pmatrix} a_{-L} \\ a_{-L+1} \\ \vdots \\ a_L \end{pmatrix} = O \begin{pmatrix} 1/\lambda_{-L} & 0 & \cdots & 0 \\ 0 & 1/\lambda_{-L+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & & & 1/\lambda_L \end{pmatrix} O^T \begin{pmatrix} C_{yx}(-L) \\ C_{yx}(-L+1) \\ \vdots \\ C_{yx}(L) \end{pmatrix},$$

hard

$$\begin{pmatrix} a_{-L} \\ a_{-L+1} \\ \vdots \\ a_L \end{pmatrix} = O \begin{pmatrix} 1/\lambda_{-L} & 0 & \cdots & 0 \\ 0 & 1/\lambda_{-L+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & & & 0 \end{pmatrix}_{\lambda_\tau} O^T \begin{pmatrix} C_{yx}(-L) \\ C_{yx}(-L+1) \\ \vdots \\ C_{yx}(L) \end{pmatrix},$$

soft

$$\begin{pmatrix} a_{-L} \\ a_{-L+1} \\ \vdots \\ a_L \end{pmatrix} = O \begin{pmatrix} 1/\lambda_{-L} & 0 & \cdots & \cdot & 0 \\ 0 & \cdot & \cdots & \cdot & \cdot \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & \cdot & \cdot & \cdot & 1/\lambda_l \end{pmatrix}_{\lambda_\tau} O^T \begin{pmatrix} C_{yx}(-L) \\ C_{yx}(-L+1) \\ \vdots \\ C_{yx}(L) \end{pmatrix},$$

- address rank deficiency of WH matrix
- for each filter (in a set) typically only few λ are significant
example: regression of 60Hz with 10 MAGs – only 2 first λ are relevant
- reduce filter noise, suppress irrelevant channels

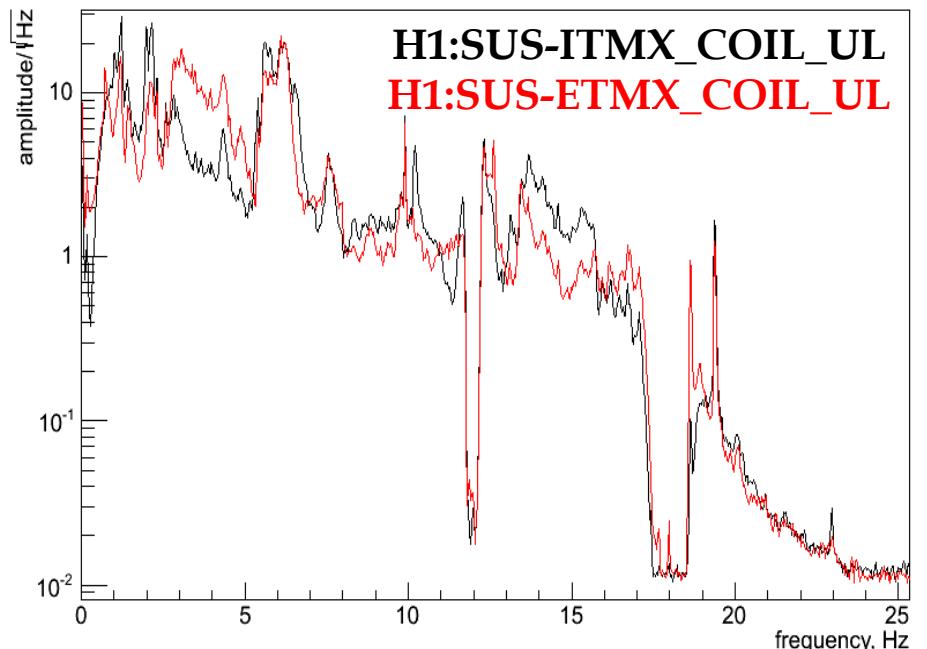
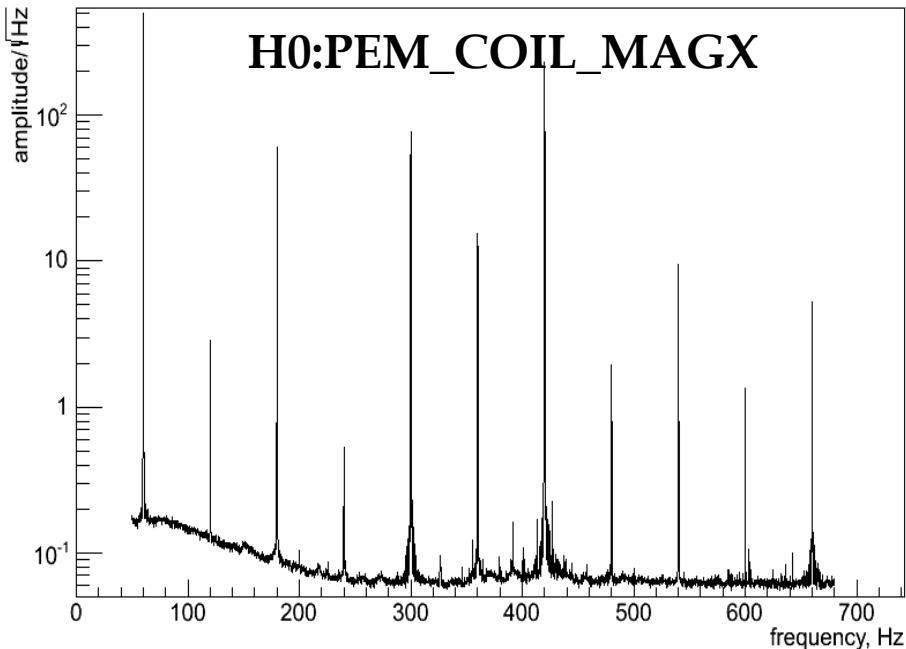




Bi-Linear coupling

12

(second term in Volterra series – see I.Pinto's talk)



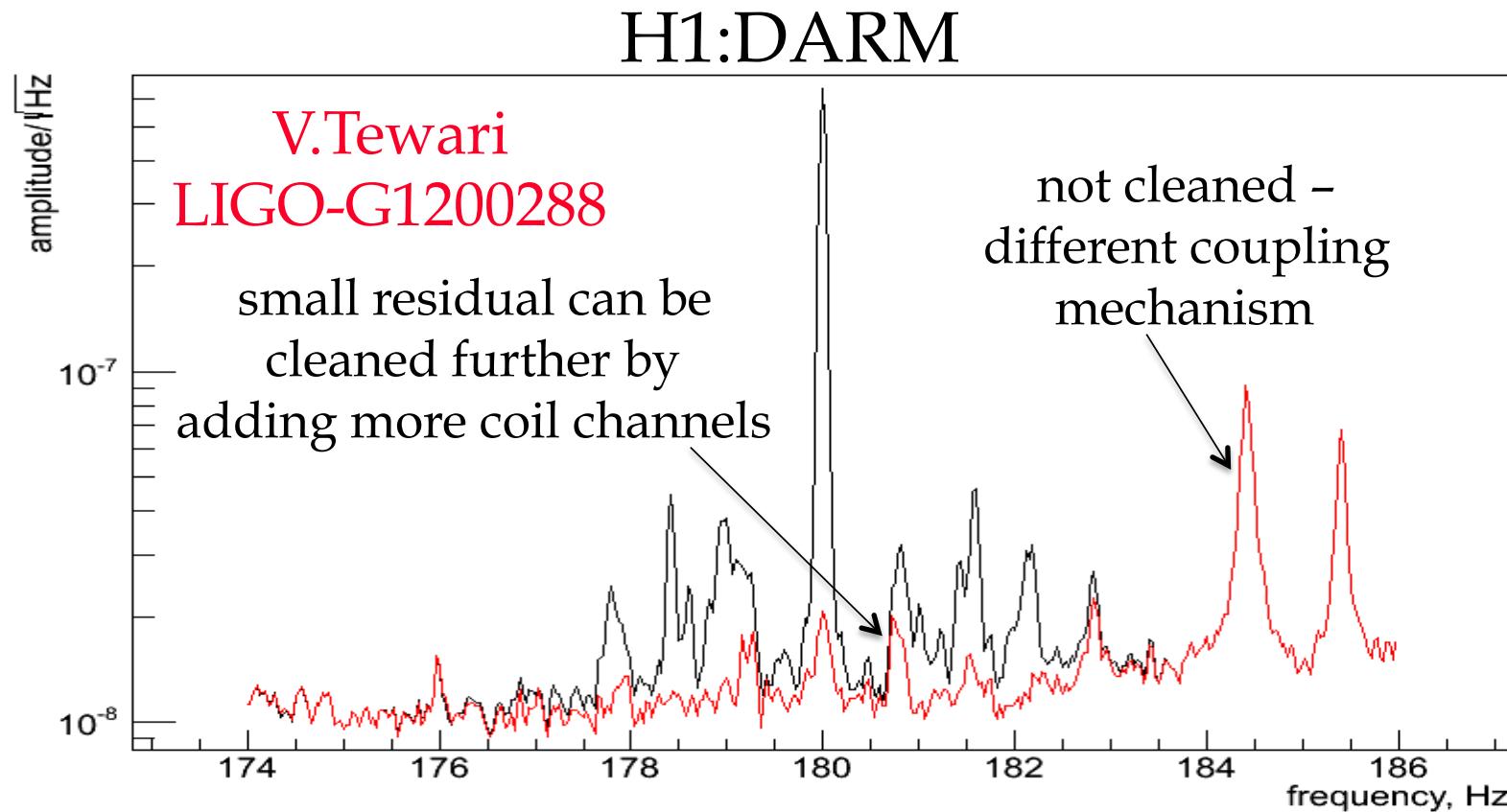
- Interaction of mirror's magnets with ambient magnetic field from power mains and low frequency coil current (seismic).
- Construct artificial witness channels as
 $\text{BICO}_{\text{XX}}_{\text{YY}}(t) = \text{H0:PEM_COIL_MAGX}(t) \times \text{H1:SUS-XX_COIL YY}(t)$
- Use coils from **ITMX, ETMX, RM, BS, MMT,...**





Regression of PL bi-coherence

13



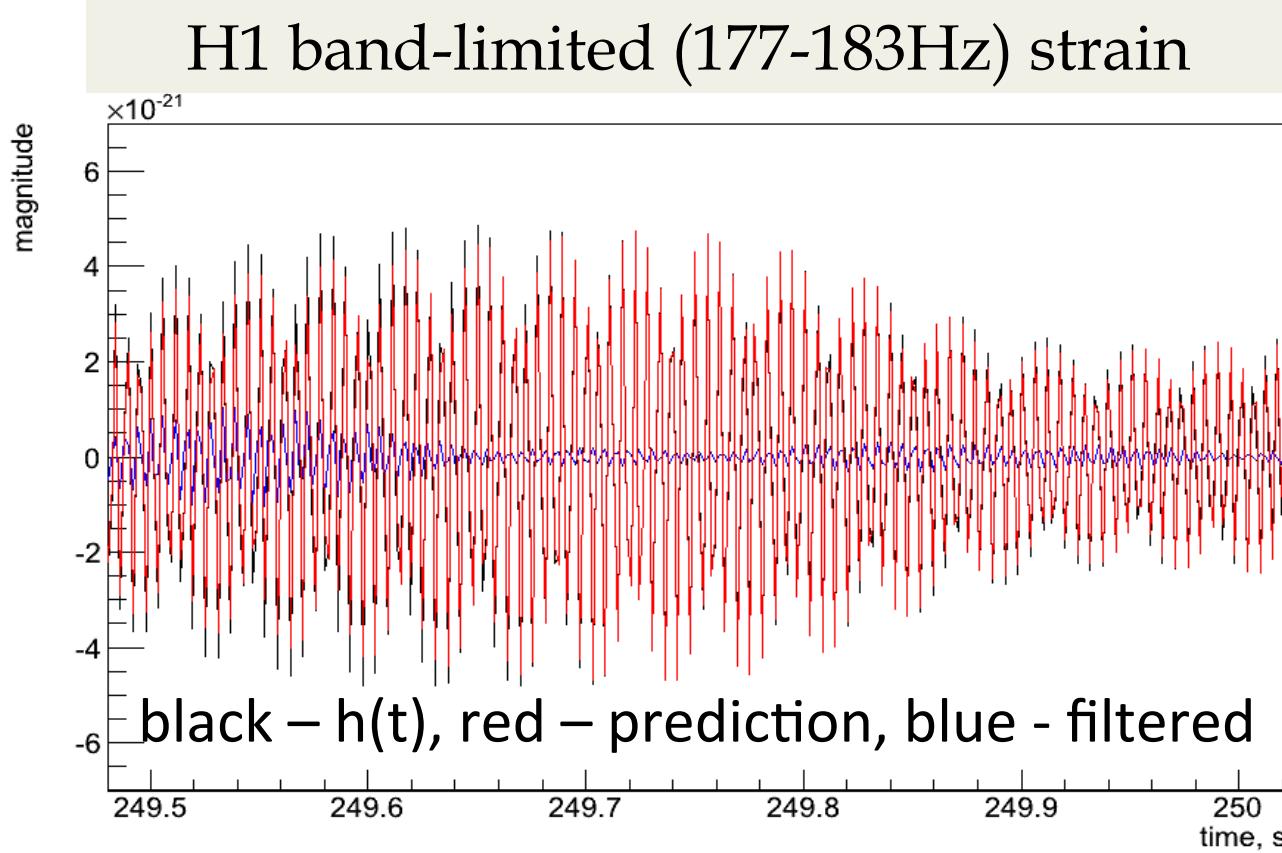
- 8 BICO(t) witnesses constructed from ITMX and ETMX coil channels and H0:PEM-BSC10_MAGX magnetometer
- first example of up-conversion noise removed from LIGO data
- useful not only in transient GW analysis, but also GW from pulsars



Approaches to Signal Processing

14

- textbook approach: $R \cdot A = C$ - “solution exists” (but often useless)
- real-world approach – always in context of a specific problem
 - identify analysis goals and select regression method
 - create analysis domain (WK filters can also be constructed in the Fourier domain)
 - obtain stable solutions (regularization)
 - create regression tool for conditioning and monitoring of GW data

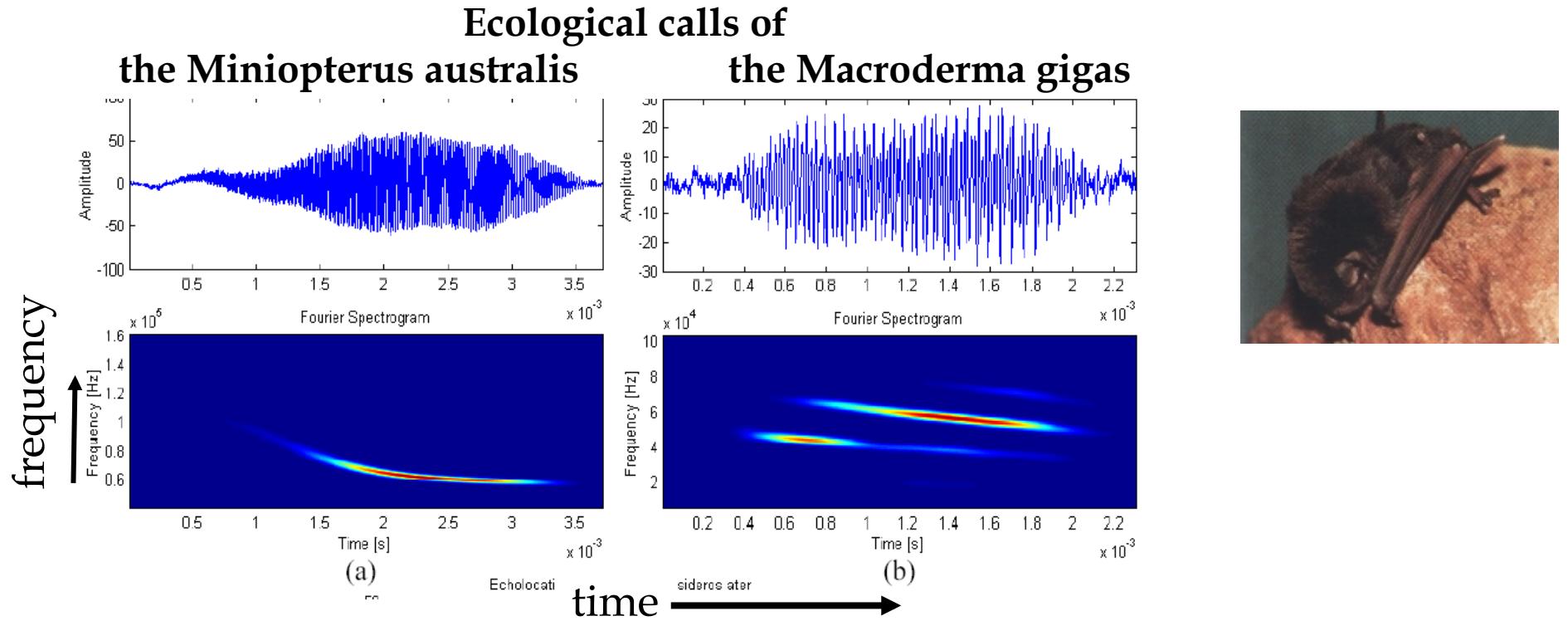




time-frequency analysis

15

simultaneous analysis of time and frequency properties of data to identify and characterize transient events (speech, music, GWs,...)



● Objective

- identify (& visualize) transients as TF patterns (detection)
- from TF pattern characterize GW “ecological calls” (reconstruction)



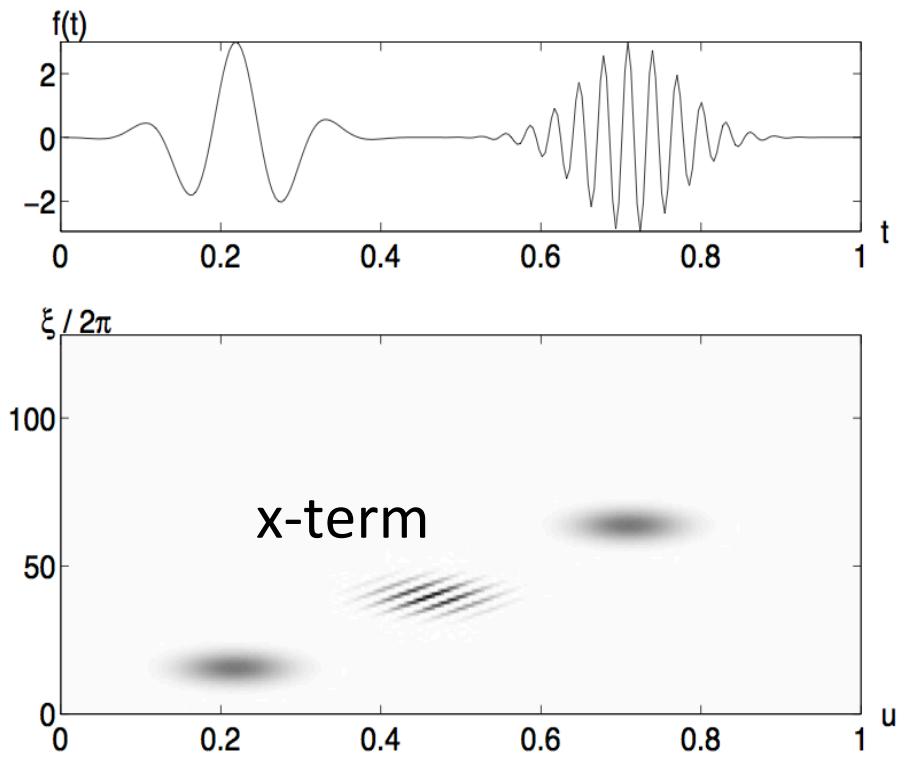
Wigner(1932)-Ville(1948) Transform

16

$$W_x(u, \xi) = \int_{-\infty}^{+\infty} f\left(u + \frac{\tau}{2}\right) f^*\left(u - \frac{\tau}{2}\right) e^{-i\tau\xi} d\tau$$

symmetric w.r.t time and frequency

- being a great theoretical tool for analysis of TF structure of data, it has very limited practical use
 - localize TF energy → x-terms created by quadratic properties of WV transform
- What good TF transform should do except to localize & display TF patterns?
 - provide convenient analysis domain for construction & application of data-processing filters & algorithms





Time-frequency atoms

17

describe data as a superposition of waveforms (atoms)
with “minimal” TF spread to capture TF details

- Gabor atoms, 1946 (windowed Fourier)

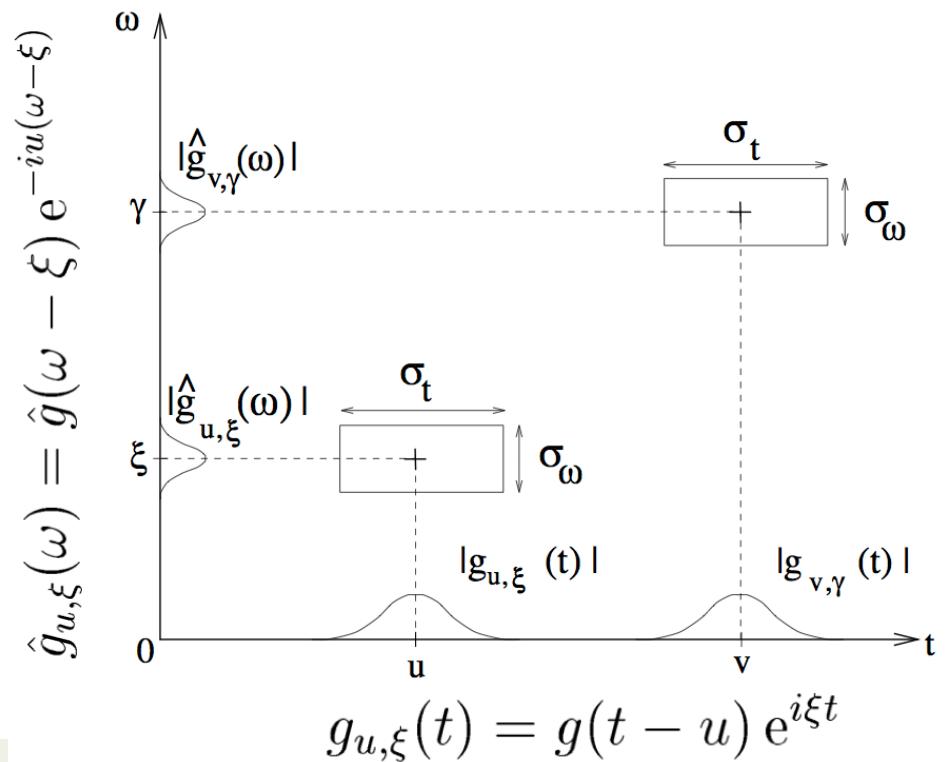
- uncertainty principle

$$\sigma_t^2 = \frac{1}{\|f\|^2} \int_{-\infty}^{+\infty} (t-u)^2 |f(t)|^2 dt$$

$$\sigma_\omega^2 = \frac{1}{2\pi\|f\|^2} \int_{-\infty}^{+\infty} (\omega - \xi)^2 |\hat{f}(\omega)|^2 d\omega$$

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4}$$

$\sigma_t \sigma_\omega = 0.5$ for Gaussian

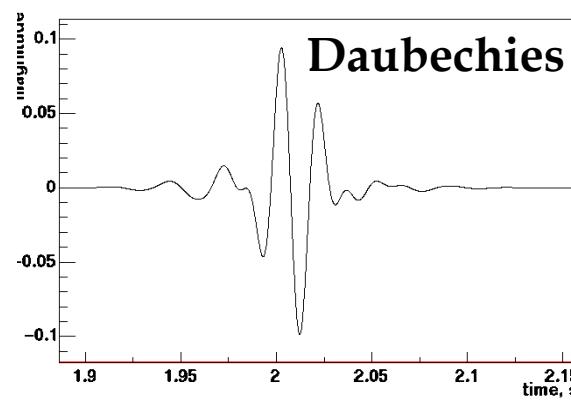
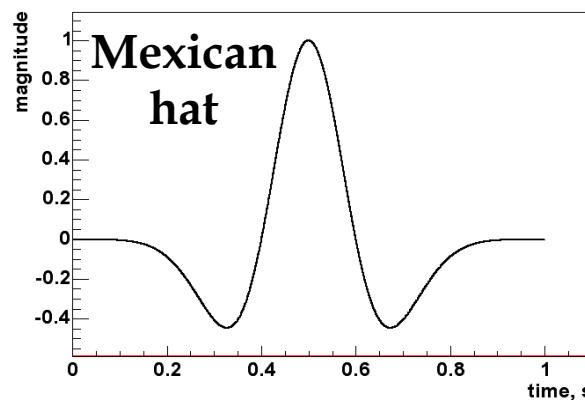
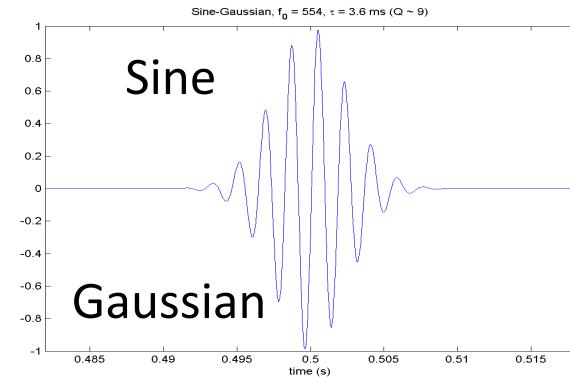




What TF atoms to select?

18

- TF atoms are natural waveforms to describe transients but
 - how to select them?
 - need sets of atoms → organize TF atoms in frames or bases
 - Linear (and preferably invertible) operator (transform) should be constructed
 - waveform shape itself could be important, but rarely a deciding factor





Frames & Bases

19

- **Transform:** Linear time-frequency transform correlates signal with a family of time-frequency atoms
 - assorted collection of TF atoms is simply a template bank which is hard to construct and use. Such banks are useless in the analysis, unless matched filters are constructed based on a source model.
 - all practical TF transforms rely on banks of TF atoms organized as frames or bases
 - ✓ possible to construct & apply FIR filters in TF domain
 - ✓ inverse exists → reconstruction of filtered signal in time domain
- **Frames:** Duffin&Schaeffer - a family of atoms ϕ_n that describes any **discrete** signal f from inner products $\langle f, \phi_n \rangle$ if for constant $A > 0$ & $B > 0$

$$A \|f\|^2 \leq \sum_{n \in \Gamma} |\langle f, \phi_n \rangle|^2 \leq B \|f\|^2$$

- **Bases:** $A = B = 1$
- **Why do we need frames & bases?** GW events are represented with several atoms that need to be processed and combined in a single event



Windowed Fourier transform

20

- Windowed Fourier waveforms $g_{u,\xi}(t) = g(t - u) e^{i\xi t}$
 - u – translation in time, ξ – translation in frequency
 - uniform tiling of the TF plane: σ_t, σ_ω const.
- Not any set $\{g_{u_n, \xi_k}\}_{(n,k) \in \mathbb{Z}^2}$ is a frame (Daubechies)

- Balian-Low theorem: If $\{g_{u_n, \xi_k}\}_{(n,k) \in \mathbb{Z}^2}$ is a Fourier frame with differentiable window than

$$\int_{-\infty}^{+\infty} t^2 |g(t)|^2 dt = +\infty \quad \text{or} \quad \int_{-\infty}^{+\infty} \omega^2 |\hat{g}(\omega)|^2 d\omega = +\infty.$$

- not possible to construct an orthonormal basis
- Good tool to display TF transients, does PSD estimation, but does not provide a convenient domain for construction of filters & application of signal processing algorithms





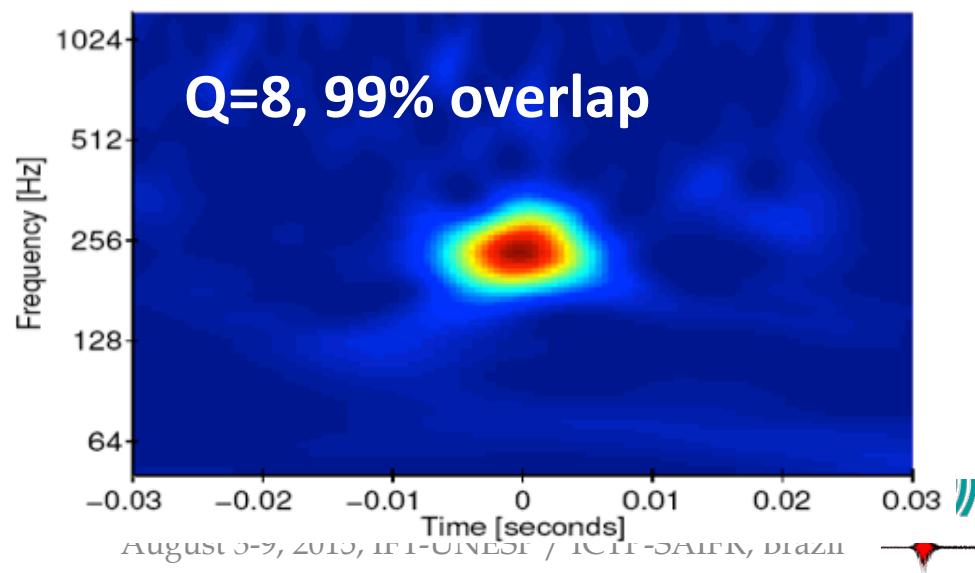
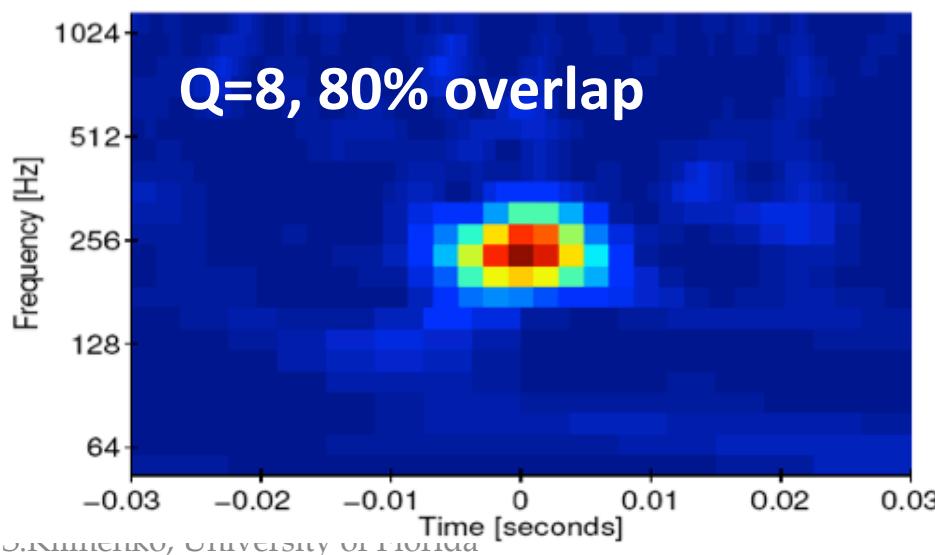
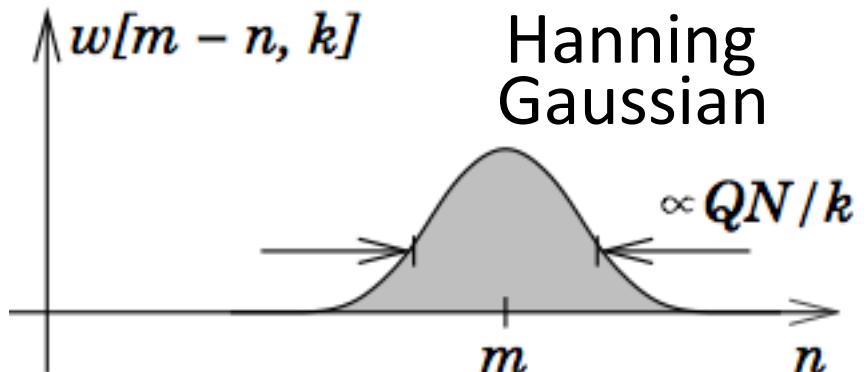
Q-transform

21

- TF tiling may not be uniform
 - project data $x[n]$ onto time-shifted windowed sinusoids, with widths inversely proportional to their center freq.

$$X[m, k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk/N} w[m - n, k]$$

- great visualization tool, use FFT (fast)
- cumbersome analysis domain – not a frame (not clear how to build filters)





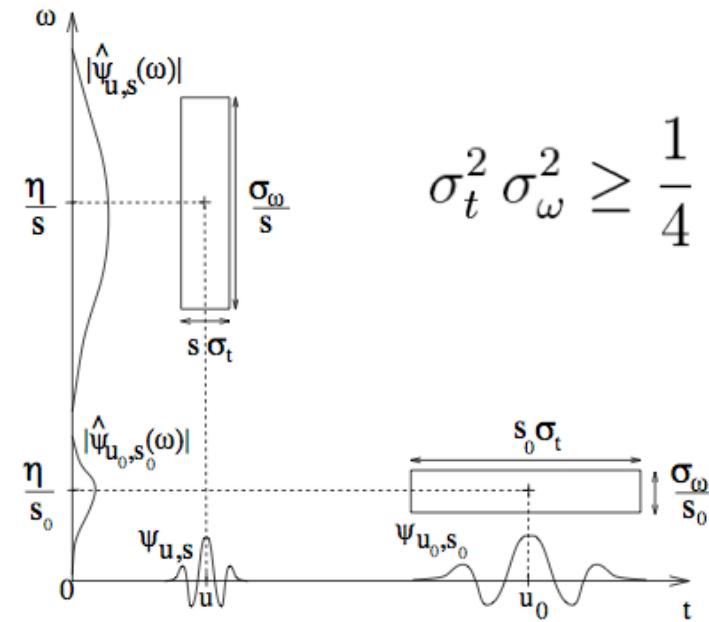
Wavelet transforms

22

- Wavelet frames [4-7] are constructed by sampling the time translation (u) and scale dilation (s) parameters of a mother wavelet ψ

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad Wf(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt.$$

- Handful of wavelet bases have been constructed: Haar, Marr, Meyer, Daubechies, Symlet, biorthogonal, ...
- Transformation is performed in iterative steps with low/high pass filters





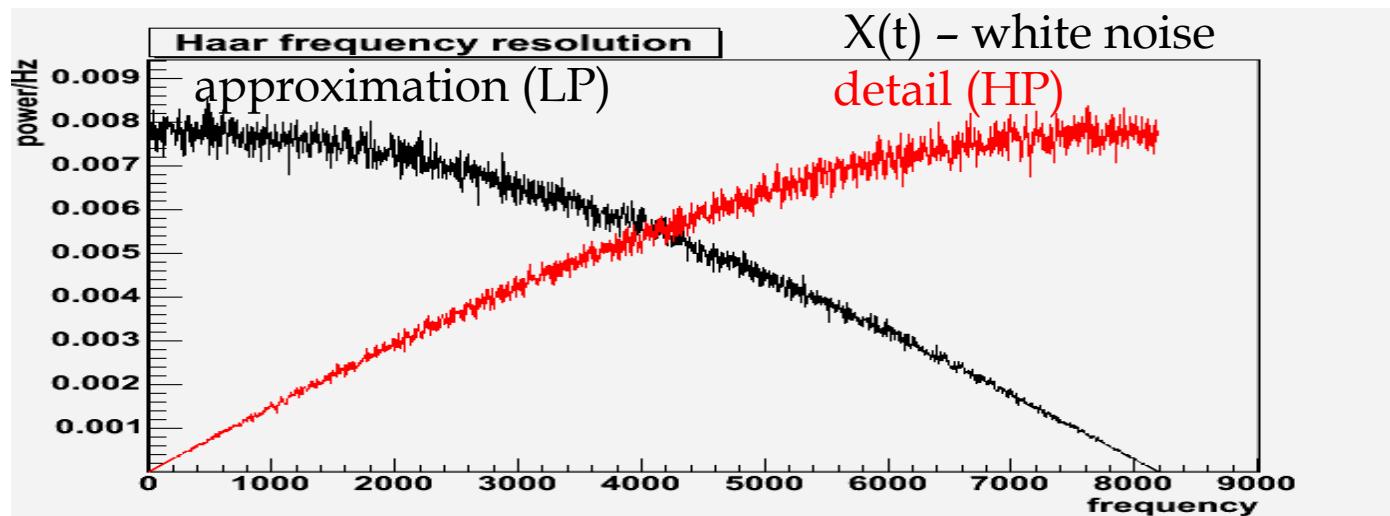
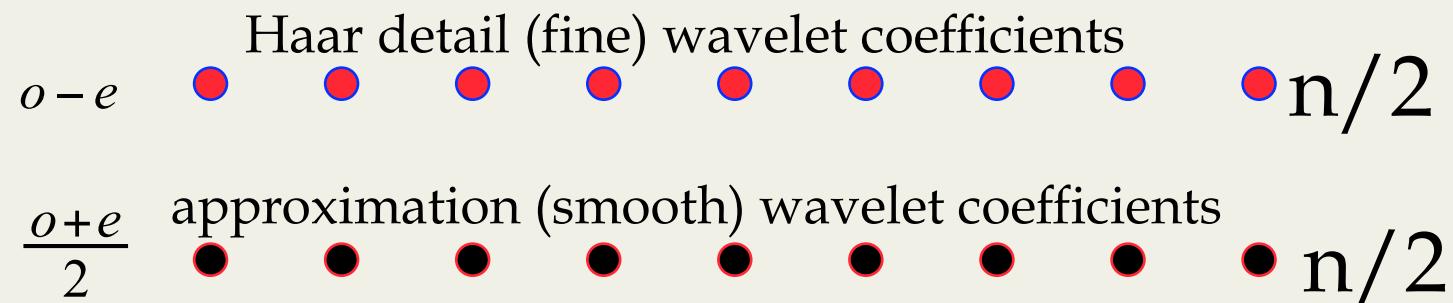
Discrete Wavelet Transforms

23

- Filter banks (S.Mallat, I.Daubechie)
- Lifting (W.Swelden) - biorthogonal wavelets,..

slow compared to FFT

$X(t)$ ● ● ● ● ● ● ● ● ● ● ● ● ● ● n





Popular Wavelet Families

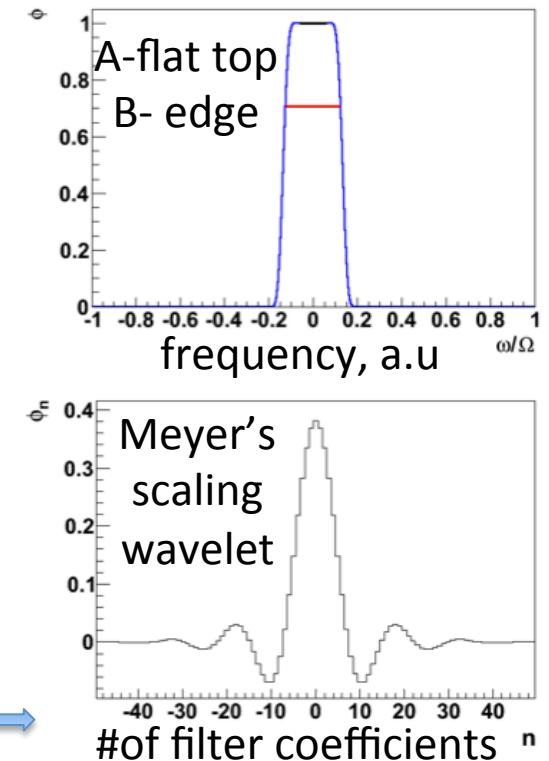
24

- Meyer [4-6] – Shannon-like wavelet constructed in F domain

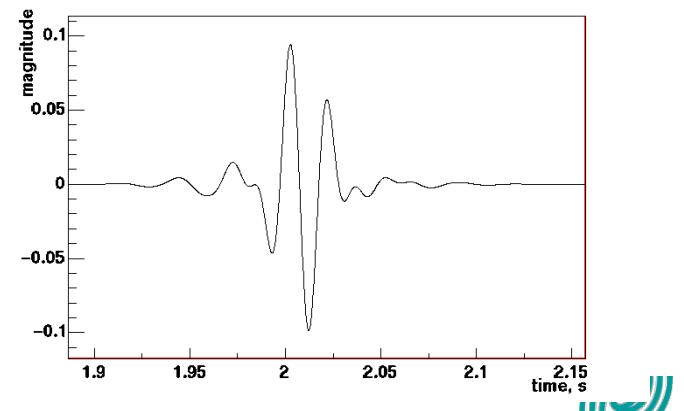
$$\tilde{\phi}(\omega) \propto \text{const}, |\omega| < A, \quad \nu_n(x) = \frac{b(x; n, n)}{b(1; n, n)}$$

$$\tilde{\phi}(\omega) \propto \cos\left[\nu_n\left(\frac{|\omega|-A}{B}\right)\frac{\pi}{2}\right], A < |\omega| < A + B$$

- $\nu_n(x)$ – Meyers edge function constructed from incomplete Beta function $b(x; n, m)$
- Meyer's filter should be truncated (no compact support)



- Daubechies [5,7] – first wavelets with compact support: $\phi(t)=0, t>|T|$
 - asymmetric wavelet function – frequency dependent LP/HP phase delay
 - Symlets – more symmetric Daubechies wavelets

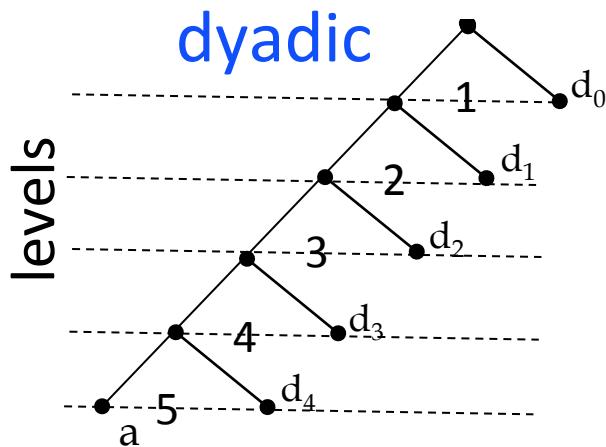




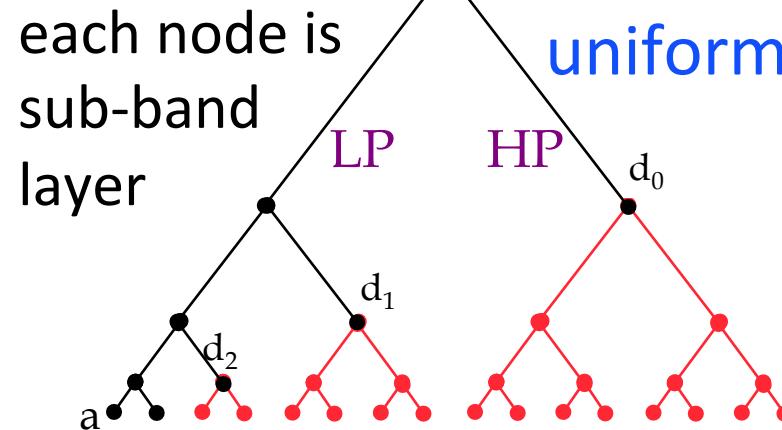
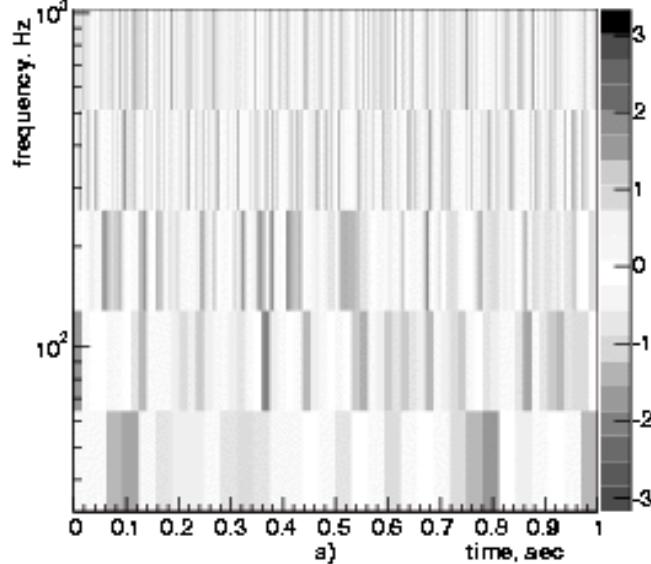
Wavelet tiling of TF plane

25

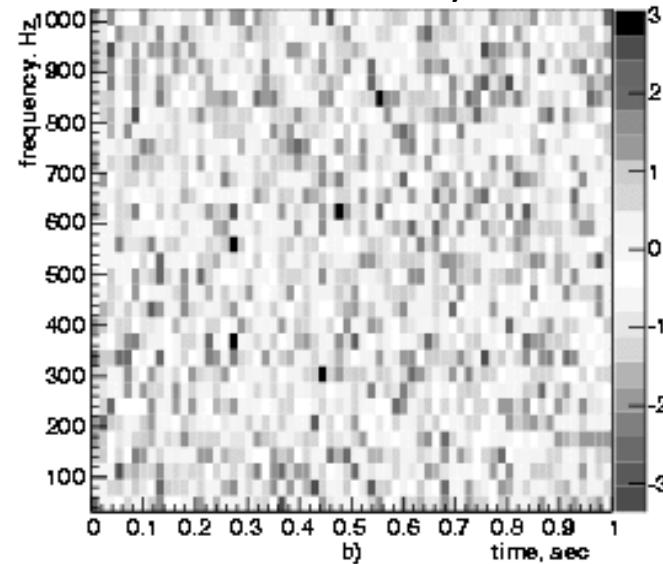
critically sampled – same # of samples before and after



a. wavelet decomposition tree



b. wavelet binary tree



Great data resampling tool based on interpolating properties of wavelets

TF data is displayed with pixels of minimum size $\Delta f \times \Delta t = 0.5$
distinguish from $\sigma_t \times \sigma_\omega$!!!



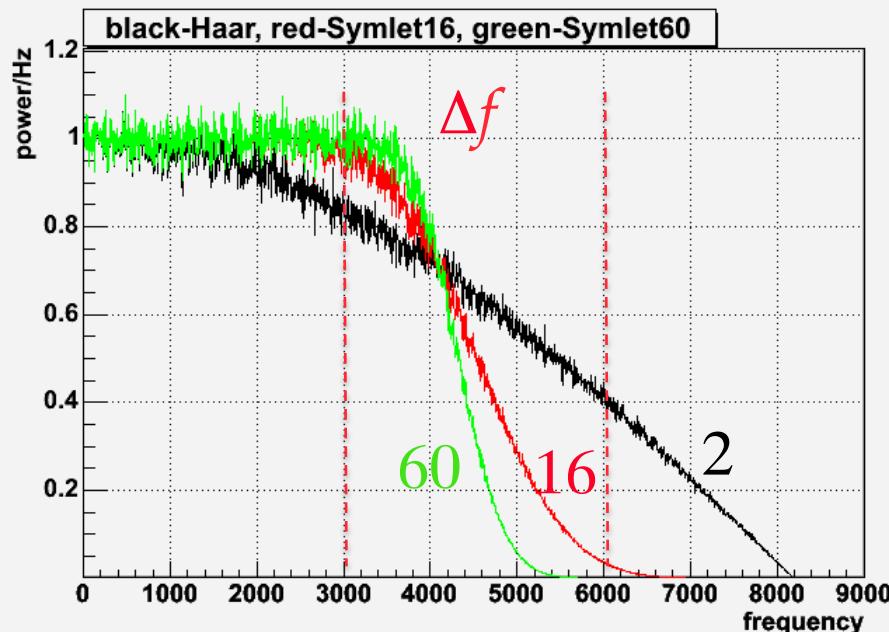


Spectral Leakage

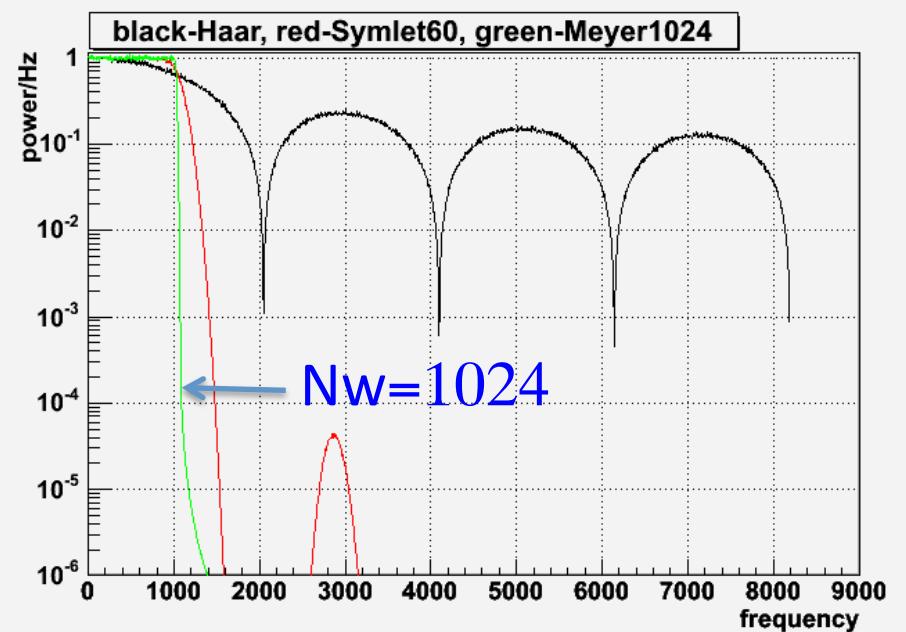
26

- x-talk between different frequency bands
 - limits wavelet performance as a band-pass filter (Haar – worst leakage)

1-step decomposition (2 layers)
leakage from 0 layer to first



3-step decomposition (8 layers)
leakage from 0 layer to 1-7 layers



- LIGO/Virgo data has a huge dynamic range
 - need long wavelet filters to achieve good intrinsic frequency resolution ($\Delta f = f_s/Nw$) and get acceptable spectral leakage
→ significantly slow down the DWT and complicates the analysis





Windowed Fourier meets Wavelets

27

- **Wilson** [8] circumvented the Balian-Low theorem by introducing a general alternative to Gabor frames where each frequency band has its own window function.
- **Daubechies** [9] showed that one can build an orthonormal Wilson basis by using just one (mother) window function $\phi(t)$ and construct the basis functions in F domain as:

$$\tilde{g}_{n0}(\omega) = e^{-in\omega T} \tilde{\phi}(\omega) \quad \tilde{g}_{nm}(\omega) = \frac{1}{\sqrt{2}} e^{-in\omega T/2} \tilde{\psi}_{nm}(\omega)$$
$$\tilde{\psi}_{nm}(\omega) = C_{m+n}^* \tilde{\phi}(\omega + m\Delta\Omega) + C_{m+n} \tilde{\phi}(\omega - m\Delta\Omega) , \quad m > 0$$
$$\Delta\Omega = 2\pi/T , \quad C_{2k} = 1 , \quad C_{2k+1} = i .$$

- **Meyer** [6] constructed wavelet function with compact support in F domain used in Meyer's wavelet
 - satisfies Daubechies admissibility condition

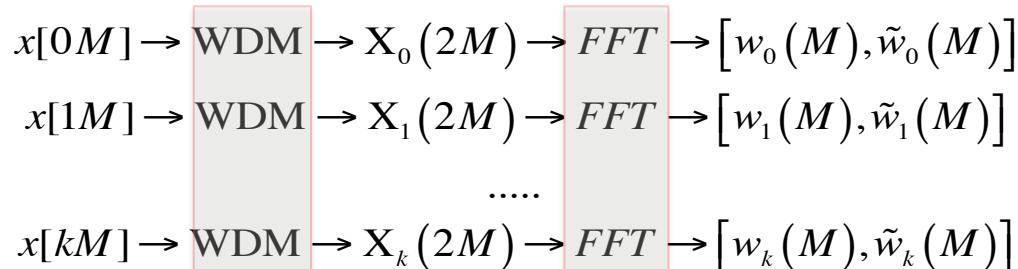
$$\sum_{l \in \mathbb{Z}} \tilde{\phi}(\omega + 2\pi l) \tilde{\phi}(\omega + 2\pi l + 4\pi m) = (2\pi)^{-1} \delta_{m0}$$



Fast WDM transform

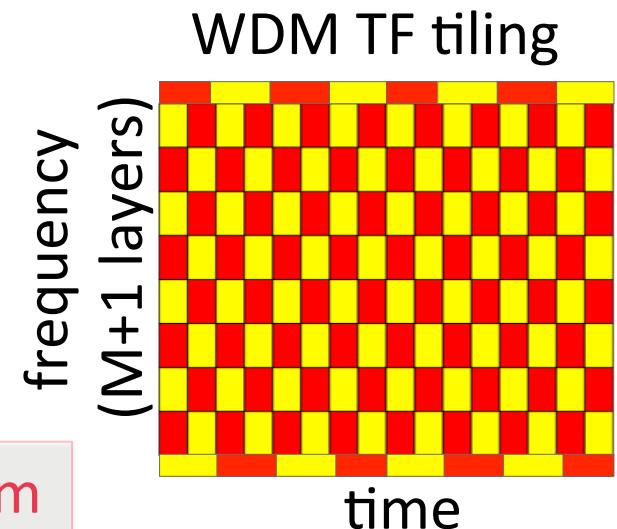
28

- Necula, Klimenko, Mitselmakher [10] developed fast WDM transform applied in 2 steps: WDM (apply filter) & FFT



- FWDM transforms x into two orthonormal bases which form a dual frame

- w - zero phase
- \tilde{w} - 90 degrees phase (quadrature) **dual stream**
- used for construction of Hermitian WK filters, reconstruction of GW polarizations and sparse TF analysis



$$C_{\text{WDM}} = 2N \left[\log_2 2M + \frac{L(K, M, P)}{M} \right]$$

- Complexity: N -# data points, M - # freq.layers, L -filter length, $L/M \sim 10$ (defined by typical WDM precision of $-\log_{10}(P)=5$ (float) and by $K=fs/2B$)

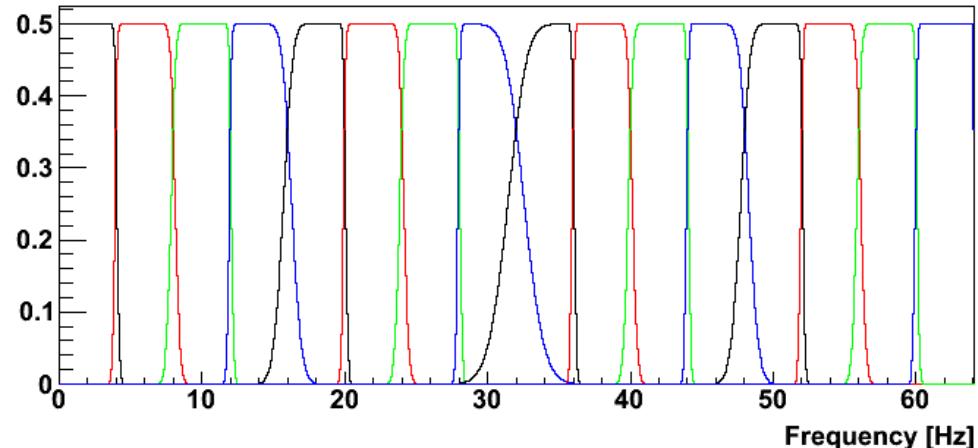
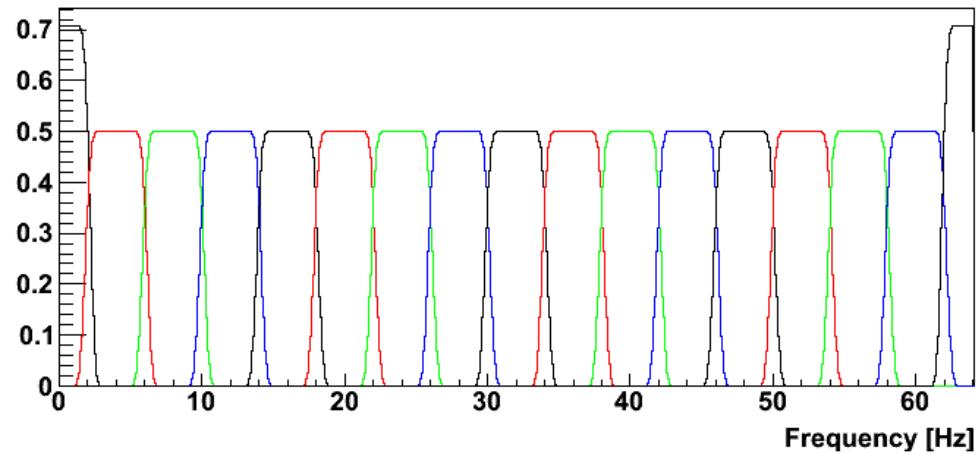
$$C_{\text{FFT}} = 2N [\log_2 M + 1]$$



Meyer's and WDM band-pass filters

29

- Wavelets and WDM split input time series into a set of down-sampled and band-passed time series → perfect for construction and application of FIR (WK&TD) filters
- WDM frequency layers
 - “uniform” resolution in the whole frequency band
 - great alternative to WFT
 - Semi-analytical time-delay (TD) filters
- Meyer's freq. layers
 - resolution depends on decomposition level
 - hard to build TD filters
 - hard to achieve segmentation of the frequency band as for WFT



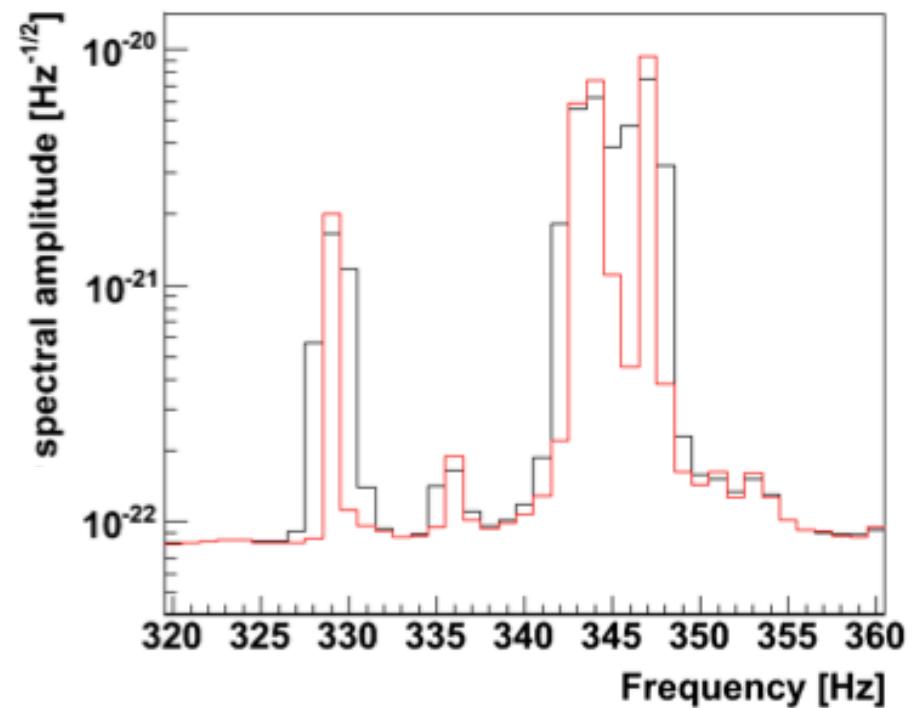
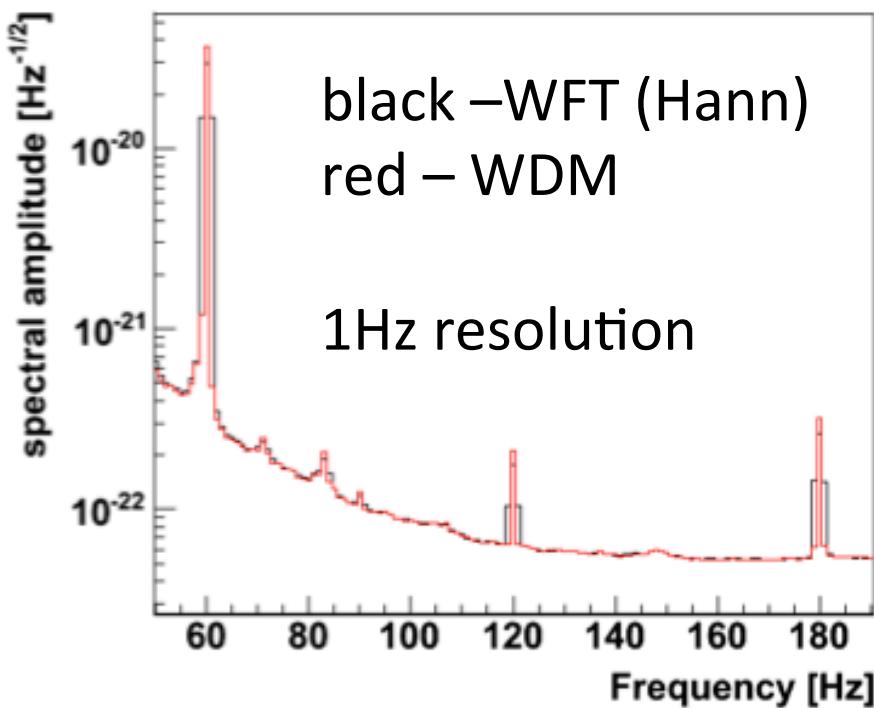


PSD Estimators

30

- along with WFT, wavelets and WDM are excellent PSD estimators: double-sided $S(\omega)$ for each frequency layer k is estimated from the variance of the wavelet (WDM) coefficients in each frequency layers. Estimated variance is used to whiten data.

$$S(\omega_k) = \sigma^2(\omega_k) / f_s$$



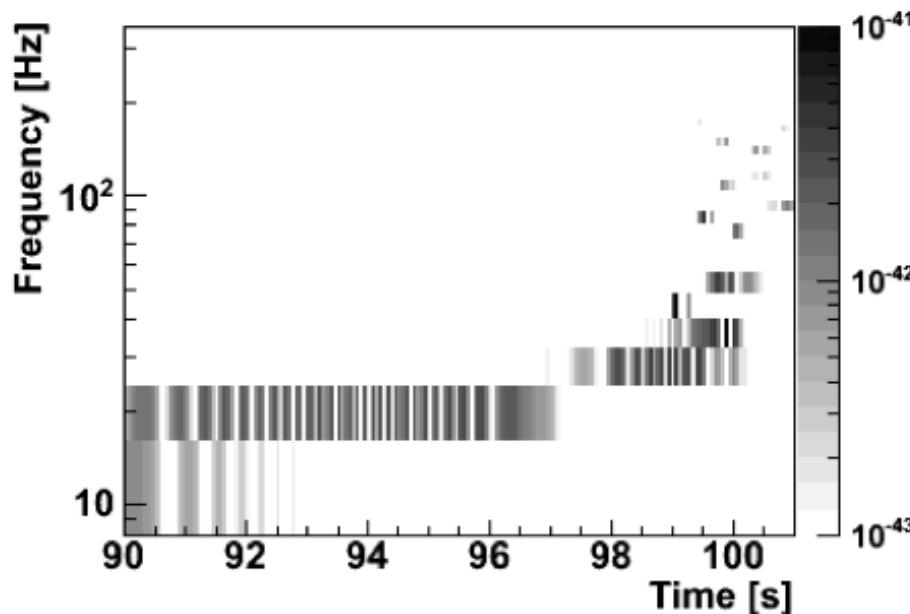


TF localization

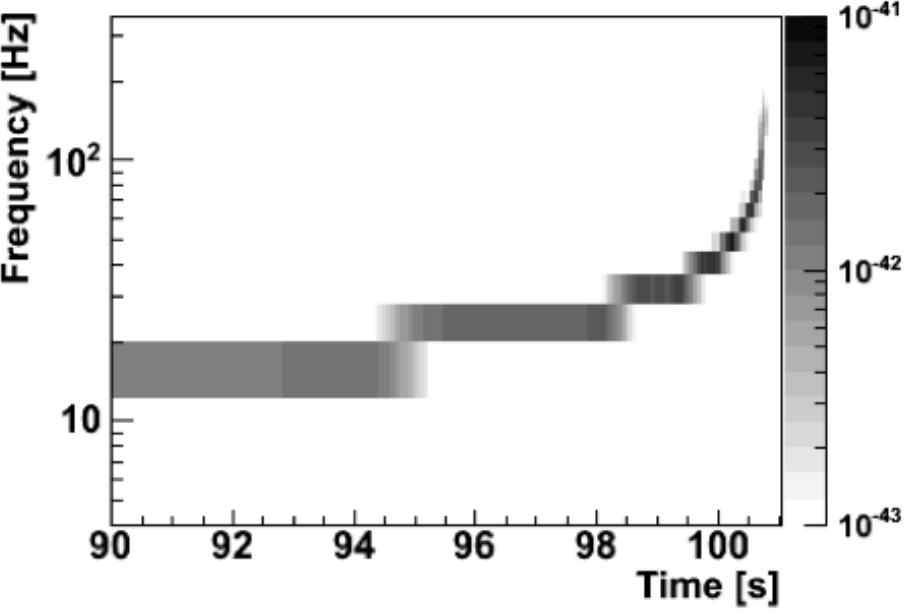
31

- Determined by
 - spread (σ_t, σ_ω) of atomic waveforms
 - spectral leakage (important for data with large dynamic range)
 - symmetry of atomic waveforms (*const* phase delay)

TF power of GW signal from 10 + 10 Mo binary black holes



$\sigma_t \sigma_\omega = 1.85$
leakage can be low
asymmetric waveforms



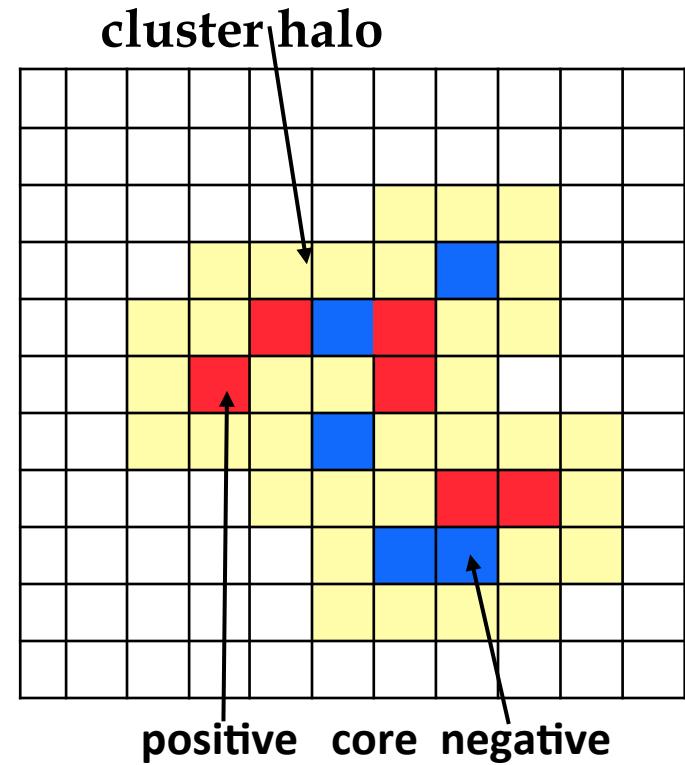
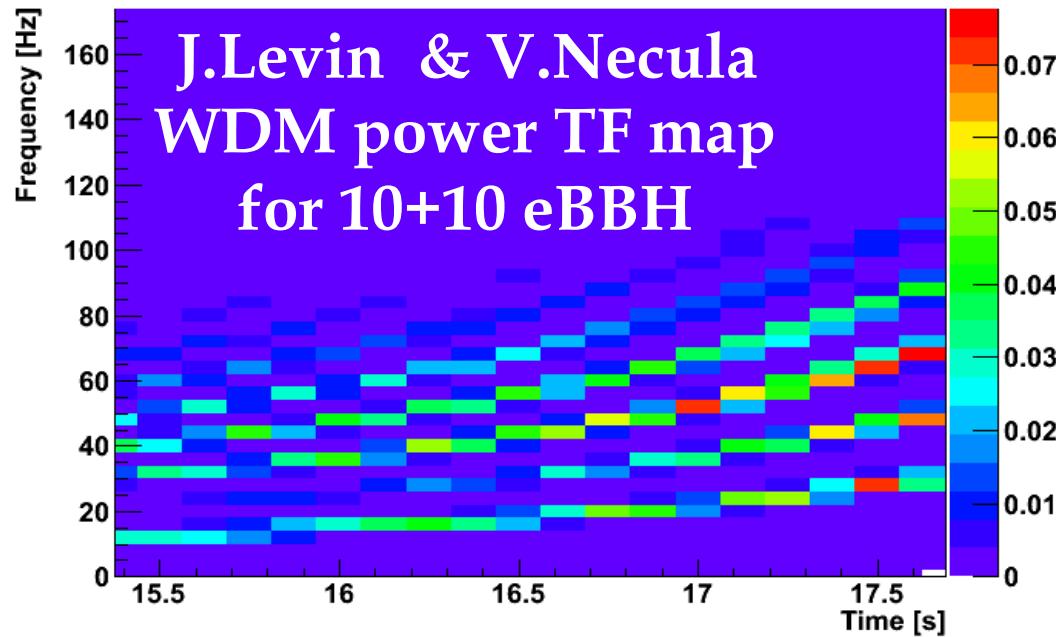
$\sigma_t \sigma_\omega = 0.62$
superior leakage control
symmetric waveforms



Cluster Analysis

32

cluster → TF area with high occupancy of hot pixels (defines an event)



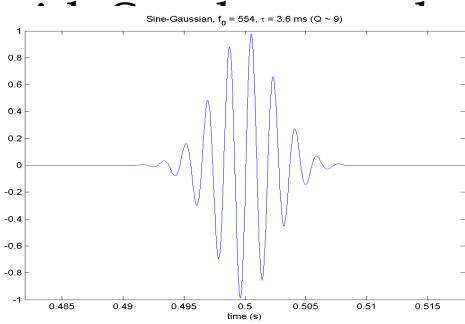
- identification – apply pattern recognition & clustering algorithms
- de-noising – construct and apply TF filters: WK, ...
- characterization – measure cluster parameters: size, t, f, Δt , Δf , ...
- reconstruction – apply Likelihood & other methods, extract signal parameters



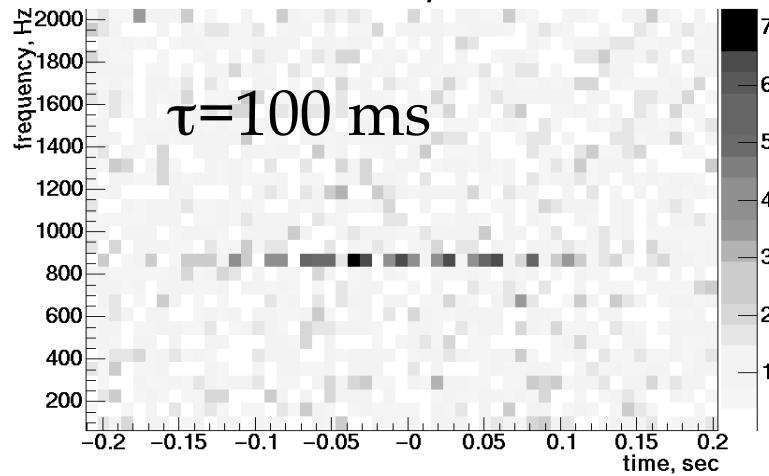
Optimal Resolution

33

representation of SG
850Hz

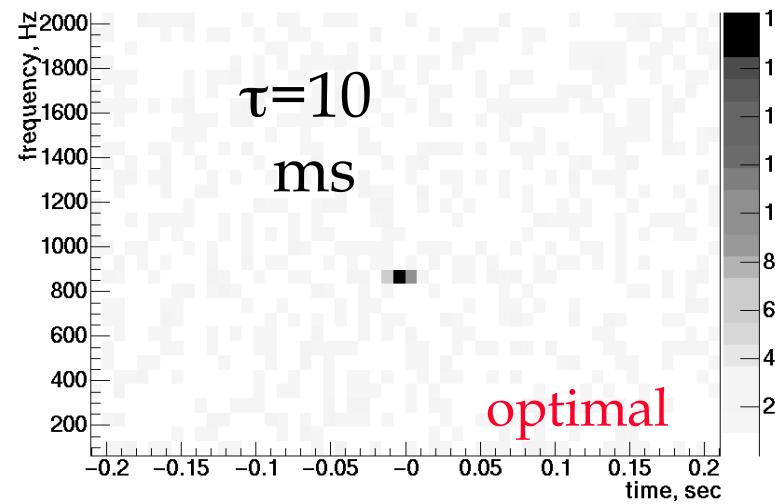


TF resolution: 1/128 sec X 64 Hz

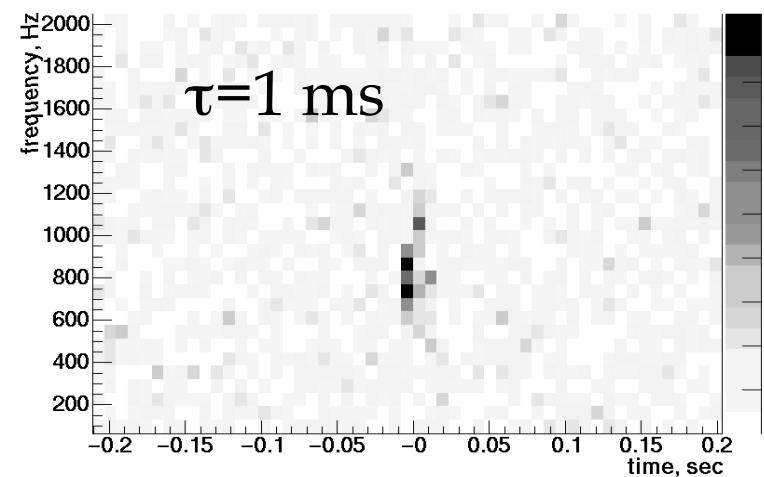


$\tau=10$
ms

optimal



$\tau=1$ ms



- How to select optimal resolution when transient duration and bandwidth are not known? → produce several TF maps and select the “optimal” one.

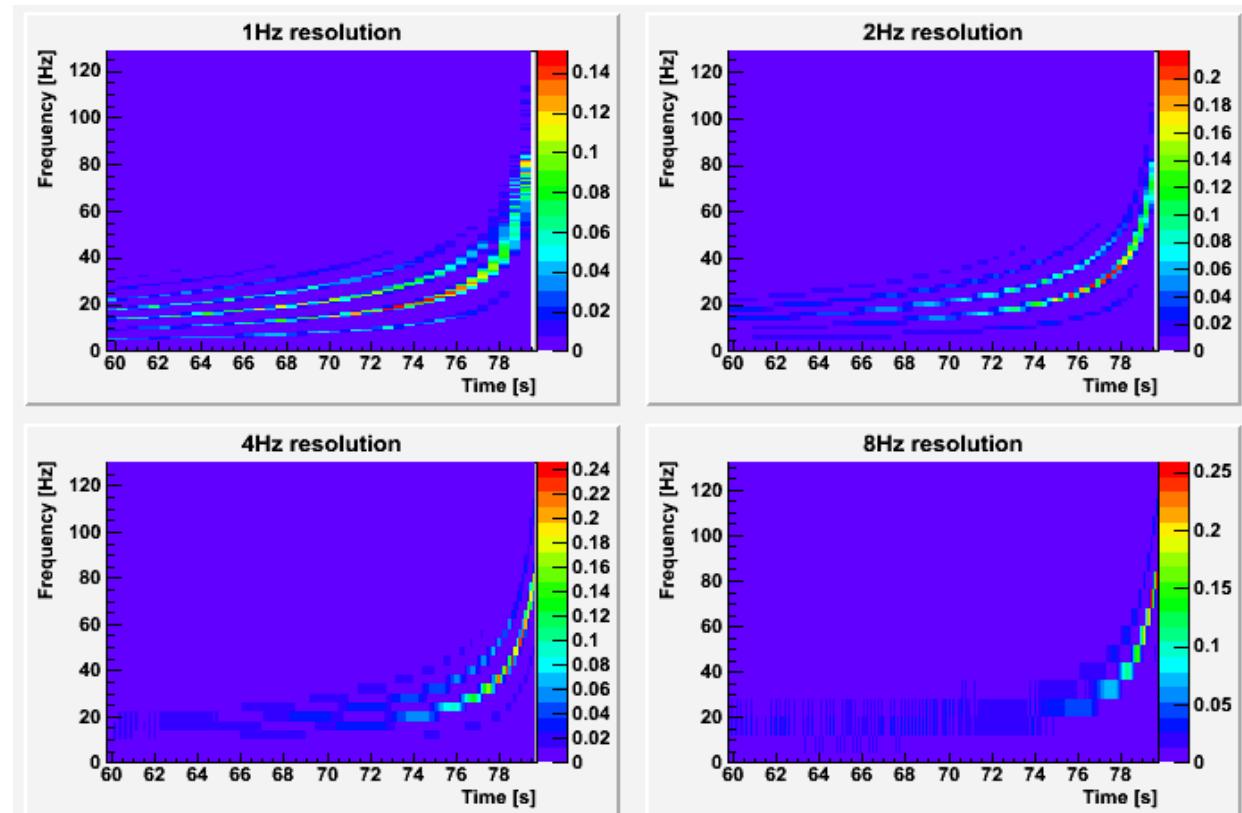


WDM frames

34

- Problem: real GW signals may not have “optimal” TF domain. For example, GWs from binary sources: inspiral requires good frequency resolution and merger requires good time resolution

10+10 Mo
binary black
holes with
eccentric orbit
at 4 different
WDM resolutions



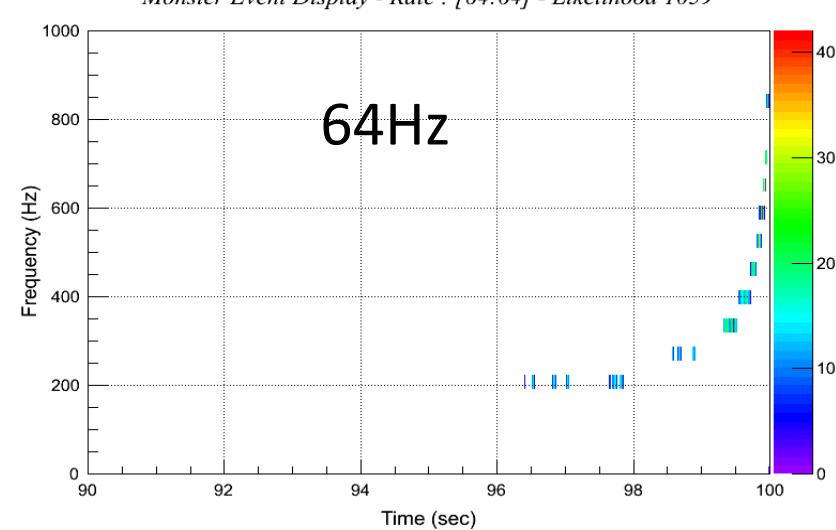
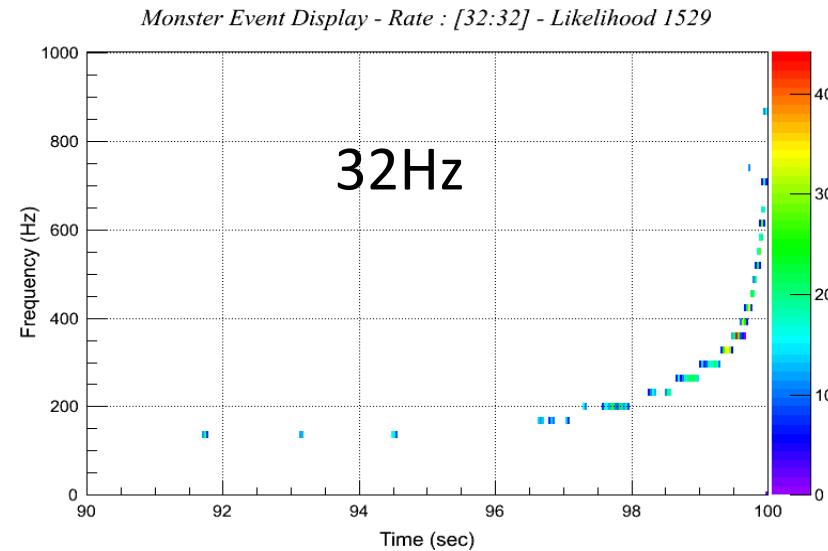
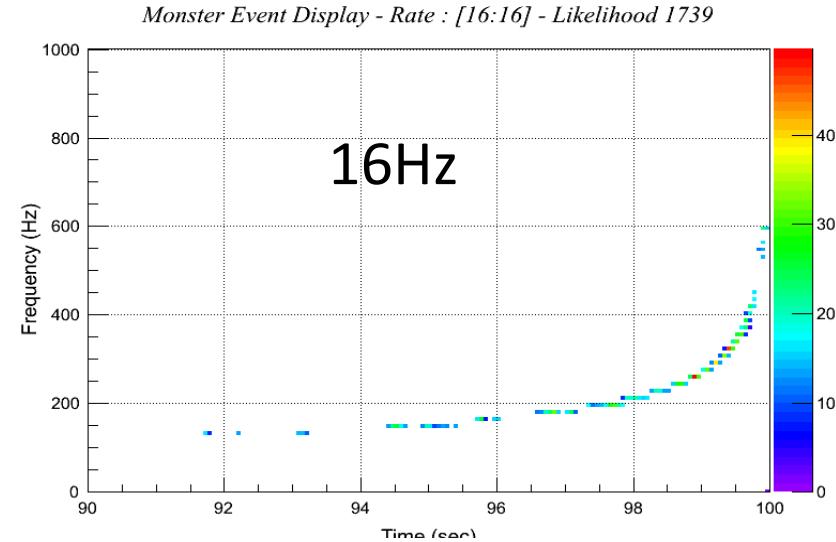
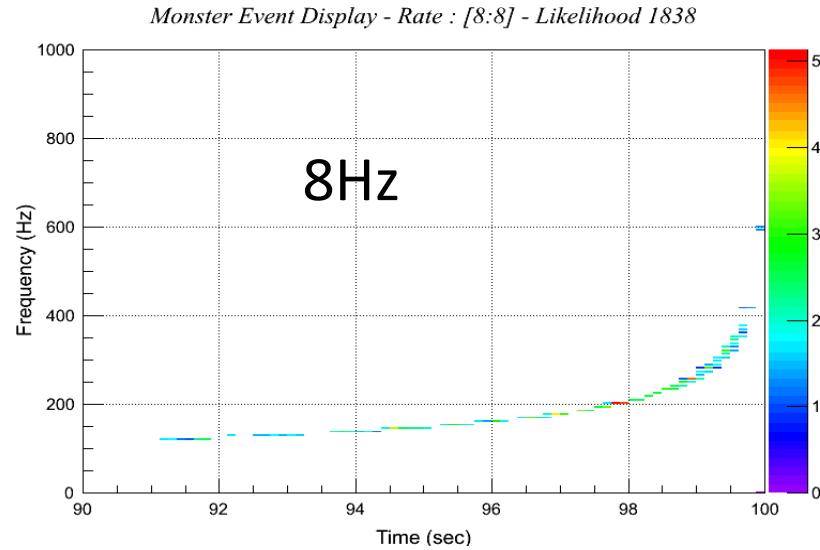
- Several WDM TF maps (bases) can be combined together to form a frame



NS-NS reconstructed aLIGO/aVirgo network

35

- NS-NS cluster with pixel amplitude $> 1.4\sigma_n$ at 4 different WDM resolutions

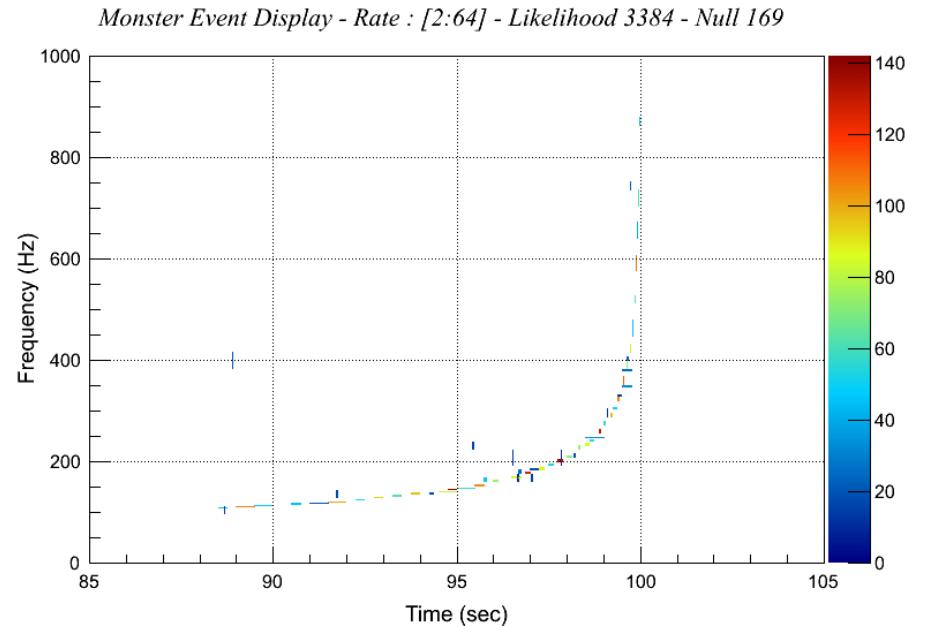
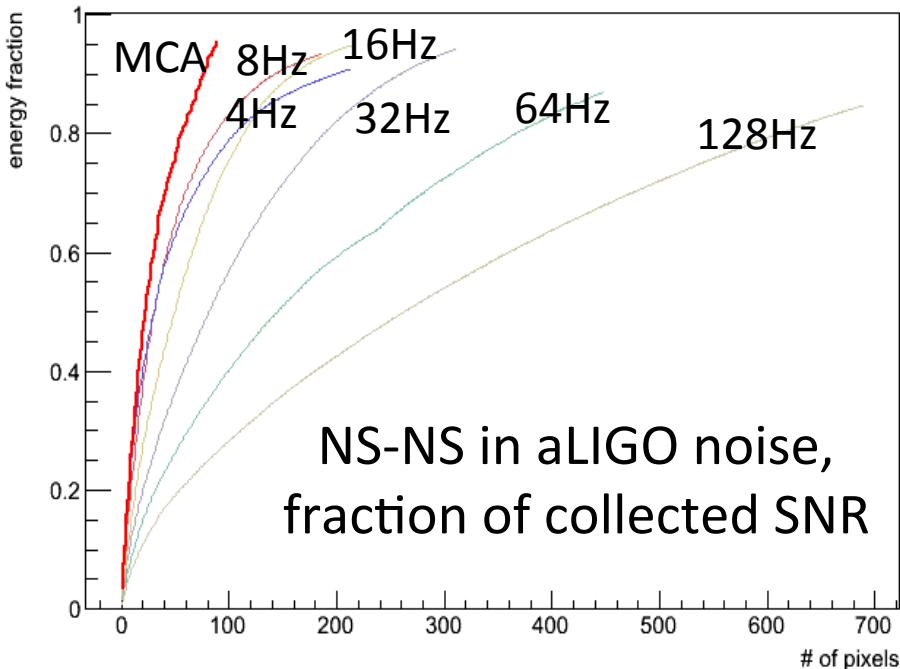




Sparse TF representation

36

- Several WDM TF maps (bases) can be combined together to form a frame



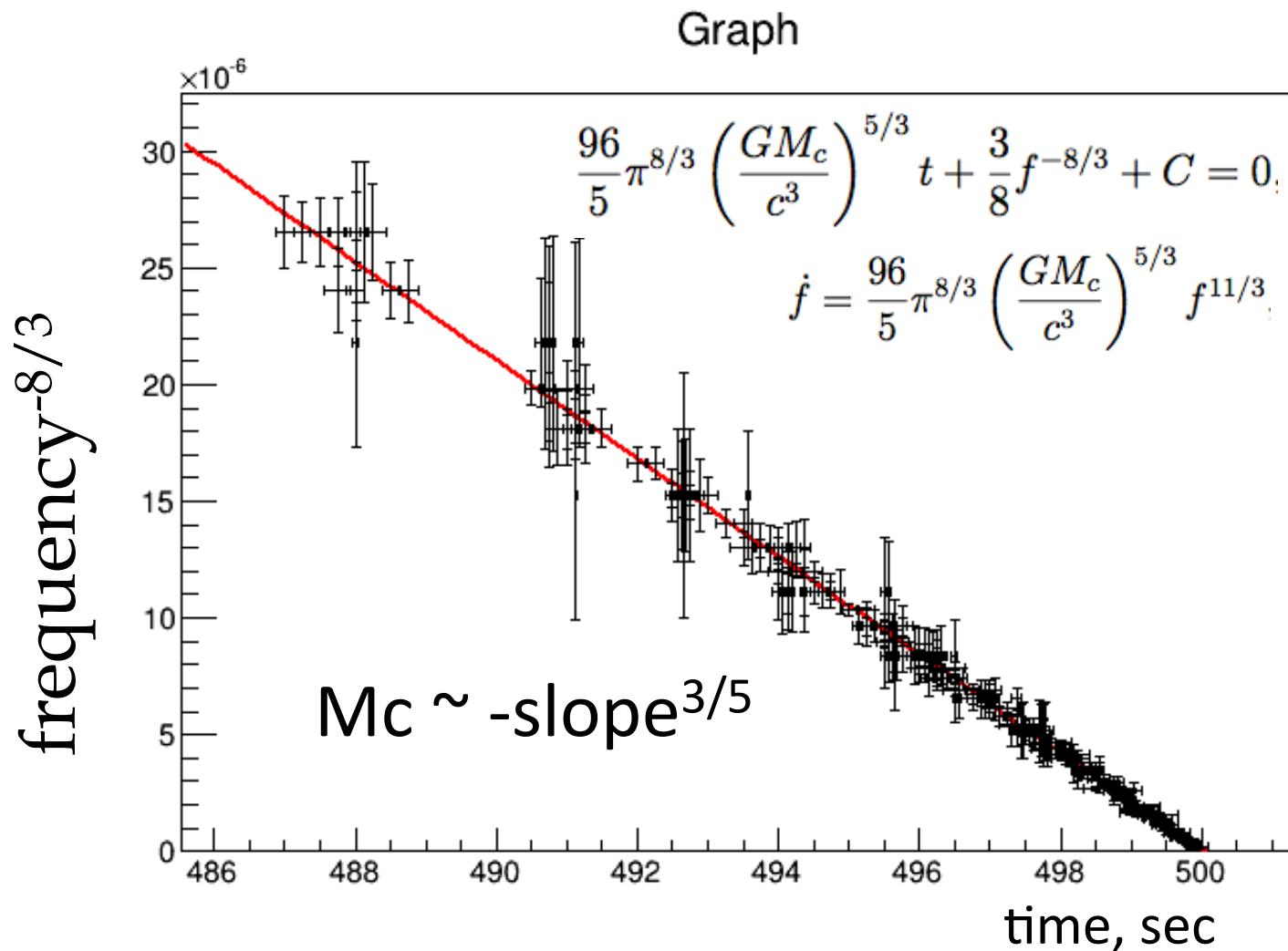
- More sparse representation (wrt individual TF resolutions) can be obtained by extraction of principle components of the signal as a superposition of pixels from different resolutions





NS-NS in aLIGO noise

37



Necula,
Tewari
Klimenko

- Mc can be estimated from TF data without detail knowledge of the waveforms.





Periodic table of TF transforms

38

| property | WFT | Q | Haar | Meyer | Daub | WDM |
|---------------------|------|------|------|-------|------|------|
| visualization | good | good | bad | fair | bad | fair |
| TF localization | good | good | bad | fair | bad | good |
| frame/base | some | no | yes | yes | yes | yes |
| compact support | no | no | yes | no | yes | no |
| invertible | some | no | yes | yes | yes | yes |
| energy conservation | no | no | yes | yes | yes | yes |
| critical sampling | no | no | yes | yes | yes | yes |
| PSD estimation | good | no | bad | fair | bad | good |
| data resampling | no | no | yes | yes | yes | no |
| performance | fast | fair | fast | slow | slow | fast |

- What transform to use depends on practical application
 - Q - visualization & detector characterization
 - Meyer, bi-orthogonal wavelets - resampling
 - WDM - data conditioning, TF filters & network analysis





Reading Material

- [0] L.Wainstein and V.Zubakov, Extraction of signals from noise, ISBN 0-486-62625-3
Statistical theory of optimal linear filtering
- [1] P. Delsarte and Y. Genin. On the splitting of classical algorithms in linear prediction theory. IEEE, ASSP-35(5), 1987.
- [2] Klimenko S, Yakushin I, Mercer A and Mitselmakher G 2008 Class. Quantum Grav. 25 114029 - application of LPR in burst analysis by cWB & Ω
- [3] RSI, 83, 024501 (2012) – active noise cancellation in suspended interferometers
- [4] S.Mallat, A wavelet tour of signal processing, ISBN 978-0123743701
Probably the best overview of TF analysis
- [5] B.Vidakovic, Statistical modeling by wavelets, 1999, ISBN 0-471-29365-2
Wavelets and their applications
- [6] Meyer Y 1992 Wavelets and Operators (Cambridge: Cambridge University Press)
- [7] Daubechies I 1992 Ten Lectures on Wavelets (Philadelphia, PA: SIAM)
- [8] Wilson K G, preprint, Cornell University
- [9] Daubechies I, Jaffard S and Journé J L 1991 J. Math. Anal. 22 pp 554-673
- [10] V Necula *et al* 2012 J. Phys.: Conf. Ser. **363** 012032
Construction and properties of fast Wilson-Daubechies-Meyer transform