Coherent WaveBurst
a joint LSC-Virgo project

LSC LIGO Scientific Collaboration

VIRGO

Burst search using a coherent algorithm applied with a network of detectors

M. Drago, AEI Hannover

Albert-Einstein-Institut
Hannover

Burst Search

- Bursts are waves of short duration: less than 1 s
  - Time of arrival is not (always) predictable
  - Waveform is not modeled enough and there is a great variety of possibility
  - We must be ready to the unexpected!!!

- Necessity of a general method to give a reliable confidence of a possible candidate:
  - Characterize detector noise: background study
  - Establish detection efficiency of an algorithm on test waveform (software injections)
  - Maximize detection efficiency and glitches rejection

Network of detectors

- The use of more detectors helps to give a more reliability on detection
- Data from different detectors can be combined in different way:
  - Incoherent pipelines: each detector has its trigger list, time-frequency coincidences are made
  - Coherent pipelines: data are combined as they belong to a single detector (likelihood)
  - Semi-coherent: intermediate approach

Hypothesis test

Signal presence	
Yes	No

- At each time the signal could be present or not

Hypothesis test

		Signal presence	
		Yes	No
Decision rule	Yes		
	No		

- At each time the signal could be present or not
- At each time we can decide that the signal is present or not (decision rule)

Hypothesis test

		Signal presence	
		Yes	No
Decision rule	Yes	True Alarm	False Alarm
	No	False Dismissal	True Dismissal

- At each time the signal could be present or not
- At each time we can decide that the signal is present or not (decision rule)
- 4 situations: two right and other wrong

Hypothesis test

		Signal presence	
		Yes	No
Decision rule	Yes	True Alarm	False Alarm
	No	False Dismissal	True Dismissal

- At each time the signal could be present or not
- At each time we can decide that the signal is present or not (decision rule)
- 4 situations: two right and other wrong
- Neyman-Pearson criterion: best decision rule gives greater True Alarm Rate at the same False Alarm Rate

Likelihood Analysis

- Likelihood Ratio

$$L = \frac{p(x \mid h)}{p(x \mid 0)}$$

- Matched filter for bursts
  - Noise model: Gaussian Noise

$$p(x \mid 0) \propto \exp[-x^2 / \sigma^2]$$

  - Signal model:

$$p(x \mid 0) \propto \exp[-(x - \xi)^2 / \sigma^2]$$

Detector Noise Variance

Detector Response

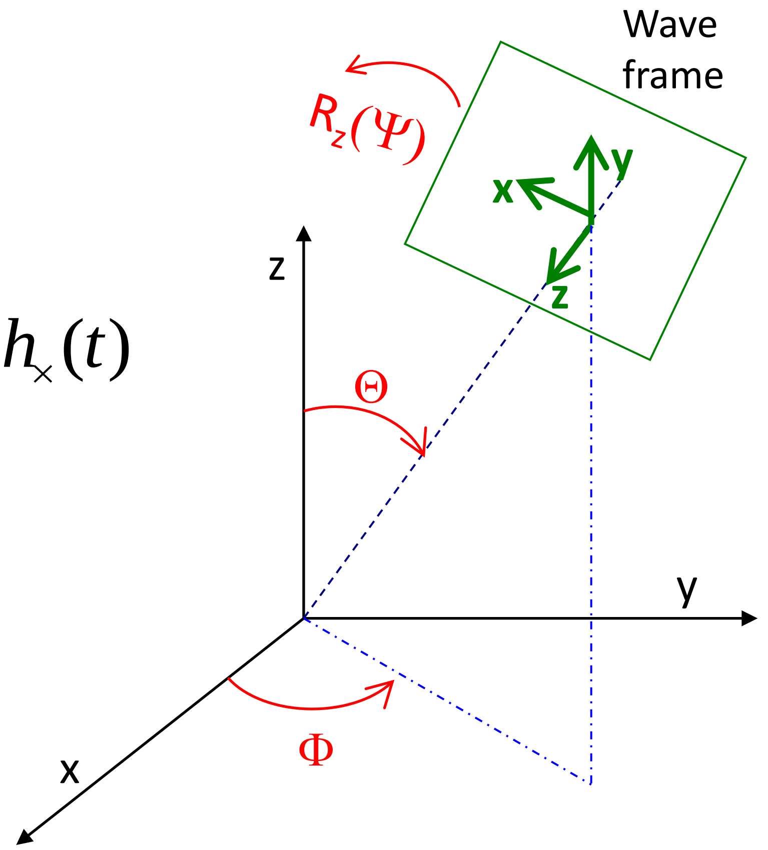
- Find best solution of h_+, h_x for maximum of L

Detector response and Antenna Patterns

- Detector response in the TT gauge can be written as:

$$\xi(t) = F_+(\Theta,\Phi,\Psi)h_+(t) + F_\times(\Theta,\Phi,\Psi)h_\times(t)$$

- Where F_+ and F_x depend on the arms orientation respect to the wave propagation and the wave polarization



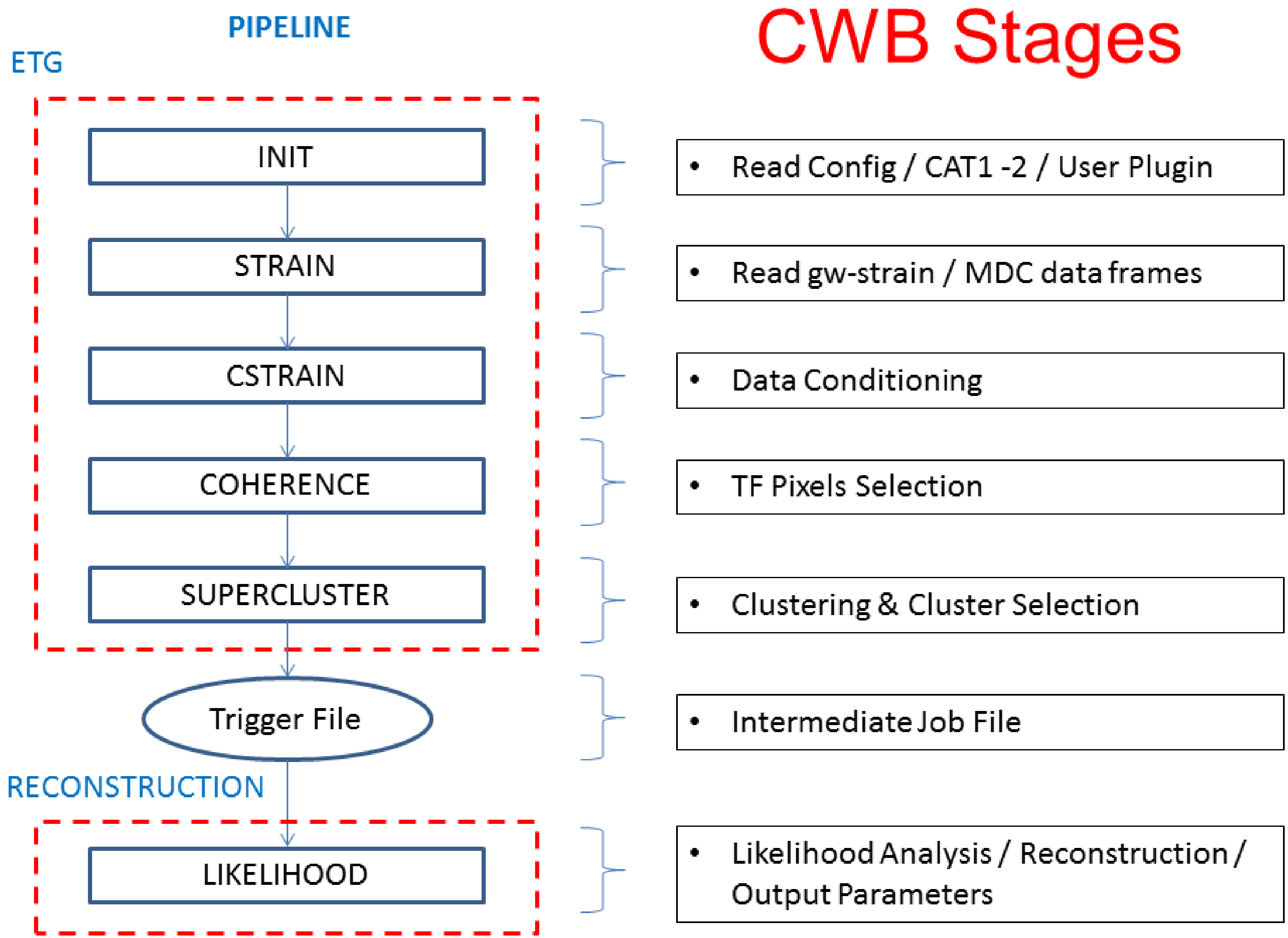
$$F_+(\Theta,\Phi,\Psi) = \frac{1}{2}(1+\cos^2\Theta)\cos 2\Phi \cos 2\Psi - \cos\Theta \sin 2\Phi \sin 2\Psi$$

$$F_\times(\Theta,\Phi,\Psi) = \frac{1}{2}(1+\cos^2\Theta)\cos 2\Phi \sin 2\Psi - \cos\Theta \sin 2\Phi \cos 2\Psi$$

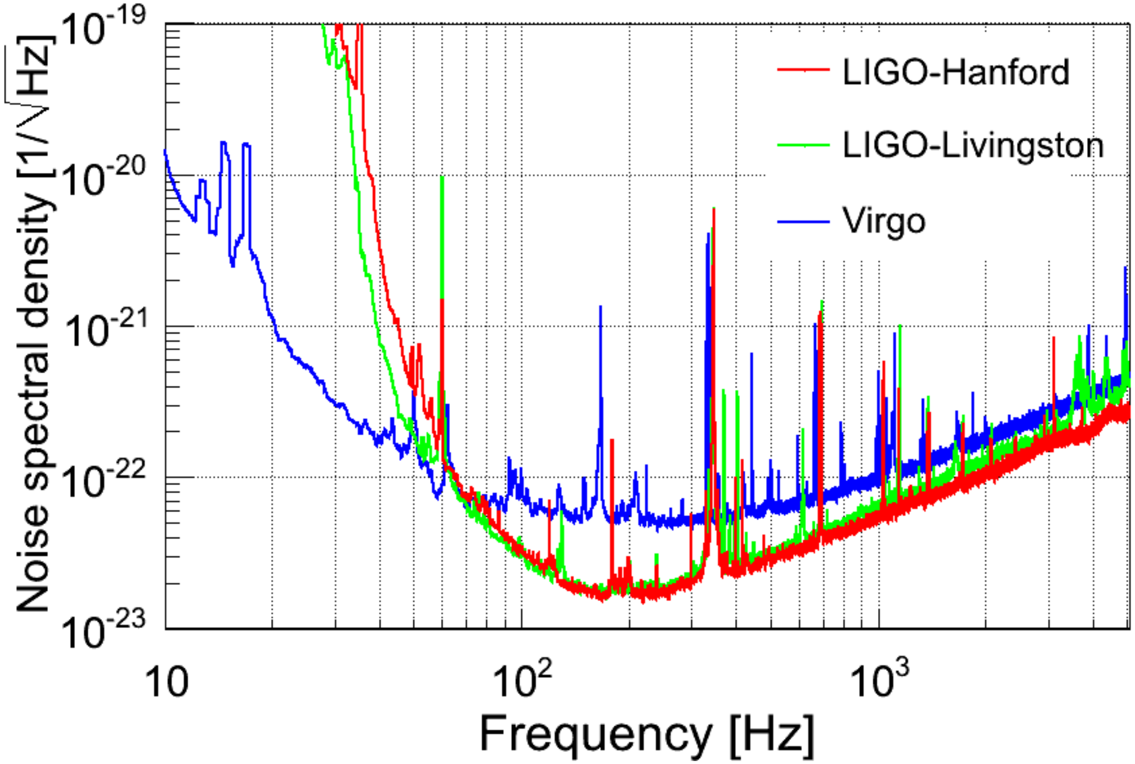
Coherent Waveburst (cWB)

- Coherent Waveburst is an algorithm of Burst search developed at LSC
- Interesting features:
  - Excess power algorithm: minimal assumption on target signal
  - Characterization of signal both in time and frequency (Wavelet)
  - Coherent analysis (Likelihood approach)
  - Reconstruction of waveforms and source coordinates
- Waveburst is applied in two steps:
  - Production: production events list
  - Post production: candidate selection

Flowchart


PIPELINE
ETG
INIT
STRAIN
CSTRAIN
COHERENCE
SUPERCLUSTER
Trigger File
RECONSTRUCTION
LIKELIHOOD
CWB Stages
Read Config / CAT1 -2 / User Plugin
Read gw-strain / MDC data frames
Data Conditioning
TF Pixels Selection
Clustering & Cluster Selection
Intermediate Job File
Likelihood Analysis / Reconstruction / Output Parameters

Data conditioning



Typical spectral feature for detector noise, one day of last joint run (2010)

1. Persistent lines
2. Not uniform in frequency

Excess power: put a threshold in energy

•Persistent lines -> Linear Predictor Filter (Regression)

•Non uniform noise in frequency -> Whitening

Time-Frequency Analysis

- Need to identify the time of arrival with great precision
- Need to identify frequency
- How we can do it at the same time?

- Time-Frequency analysis is a good solution
- In the same time we obtain the frequencies and times of the signal

Wavelet Transform

- Characterization of signal with use of template waveforms

  Ψ_0: mother wavelet

$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient

Wavelet Transform

- Characterization of signal with use of template waveforms

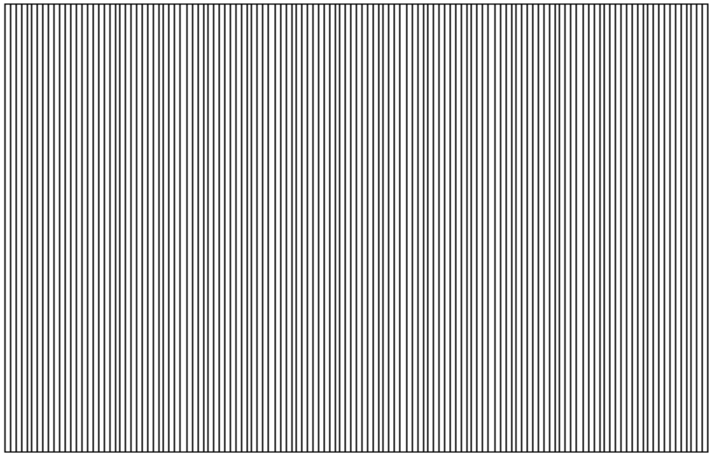
  Ψ_0: mother wavelet

$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient

•
0

Decomposition level



frequency

time

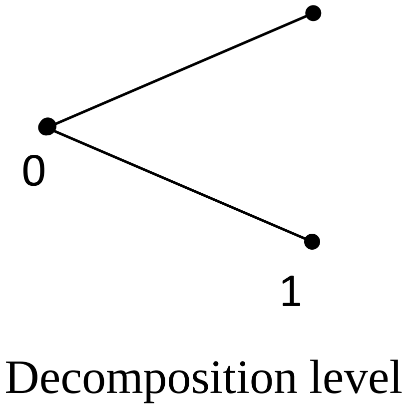
Wavelet Transform

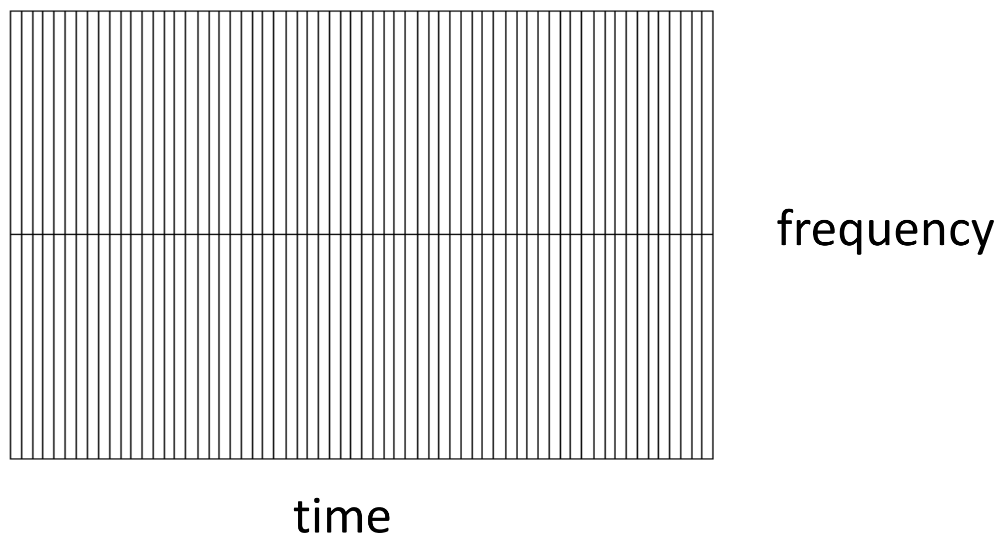
- Characterization of signal with use of template waveforms

  Ψ_0: mother wavelet

$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient





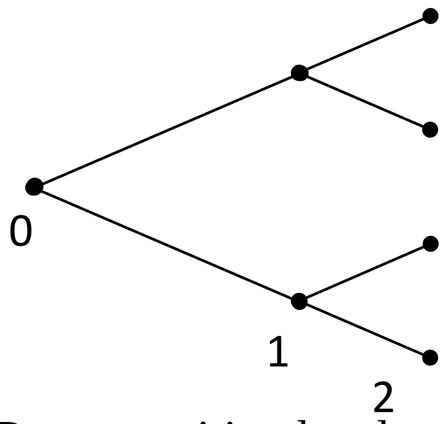
Wavelet Transform

- Characterization of signal with use of template waveforms

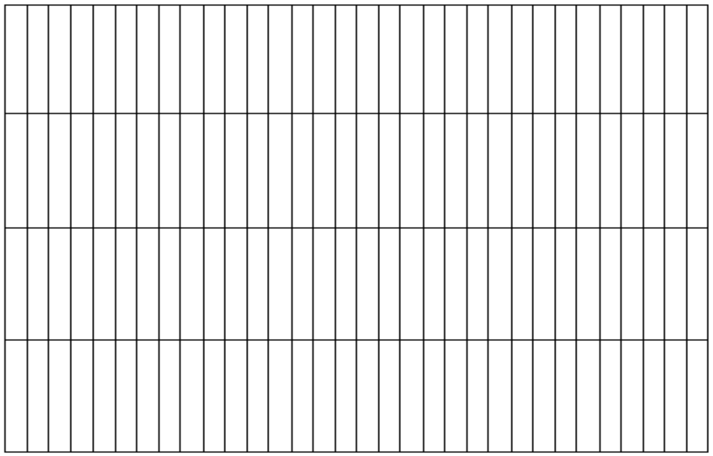
  Ψ_0: mother wavelet

$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient



Decomposition level



frequency

time

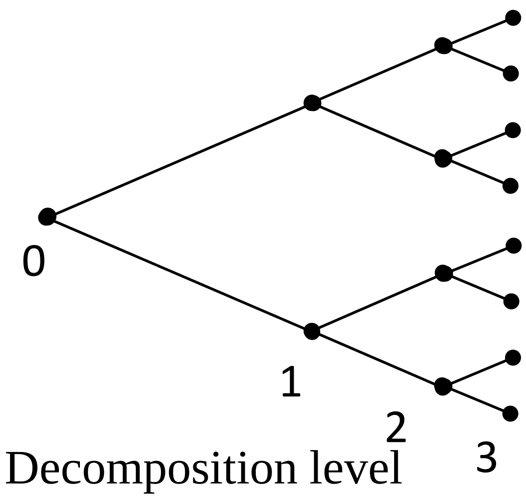
Wavelet Transform

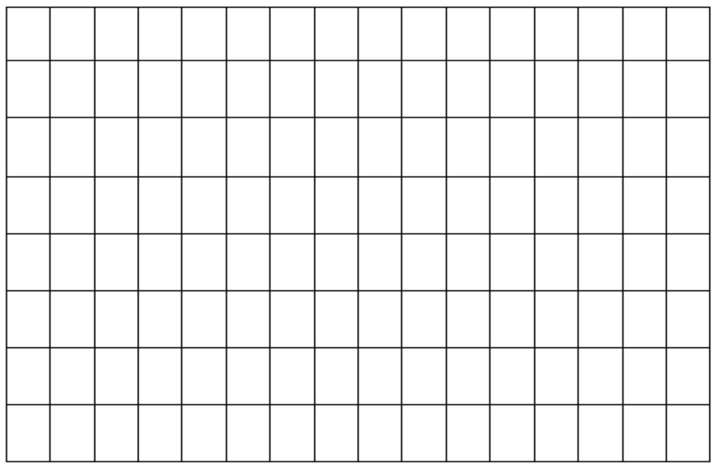
- Characterization of signal with use of template waveforms

  Ψ_0: mother wavelet

$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient





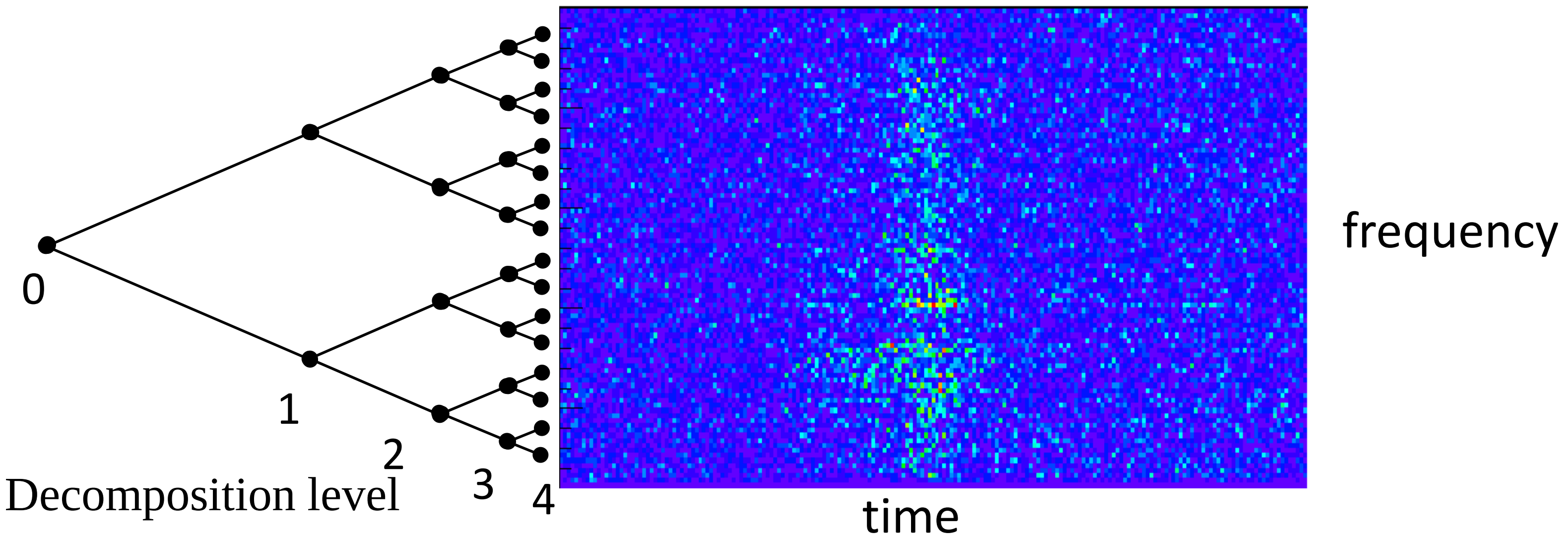
Wavelet Transform

- Characterization of signal with use of template waveforms

  Ψ_0: mother wavelet

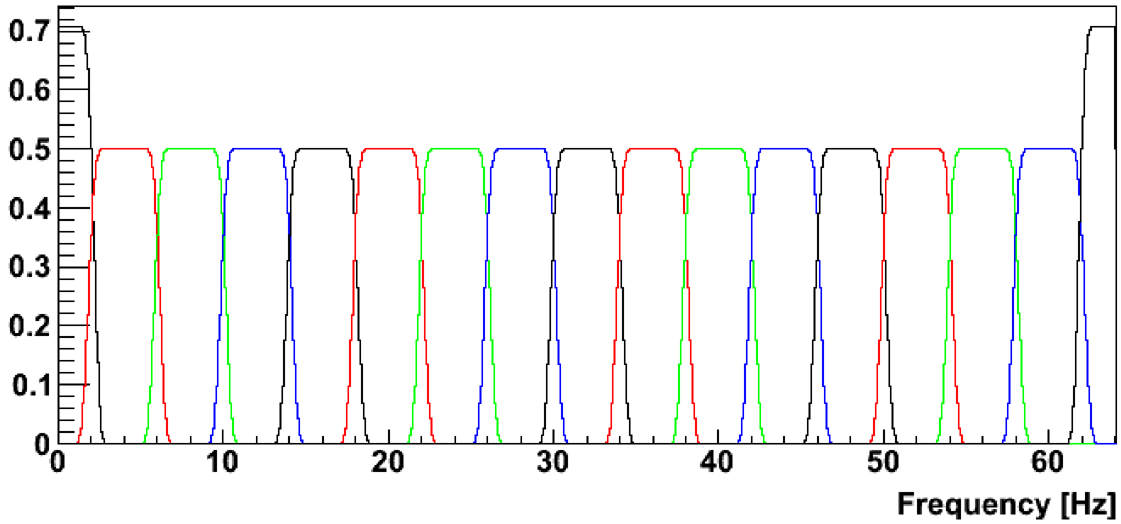
$$\Psi_{jk} = 2^{j/2}\Psi_0\left(2^j t - k\right)$$

- Natural basis of bursts
  - low spectral leakage
  - Good TF localization of transient



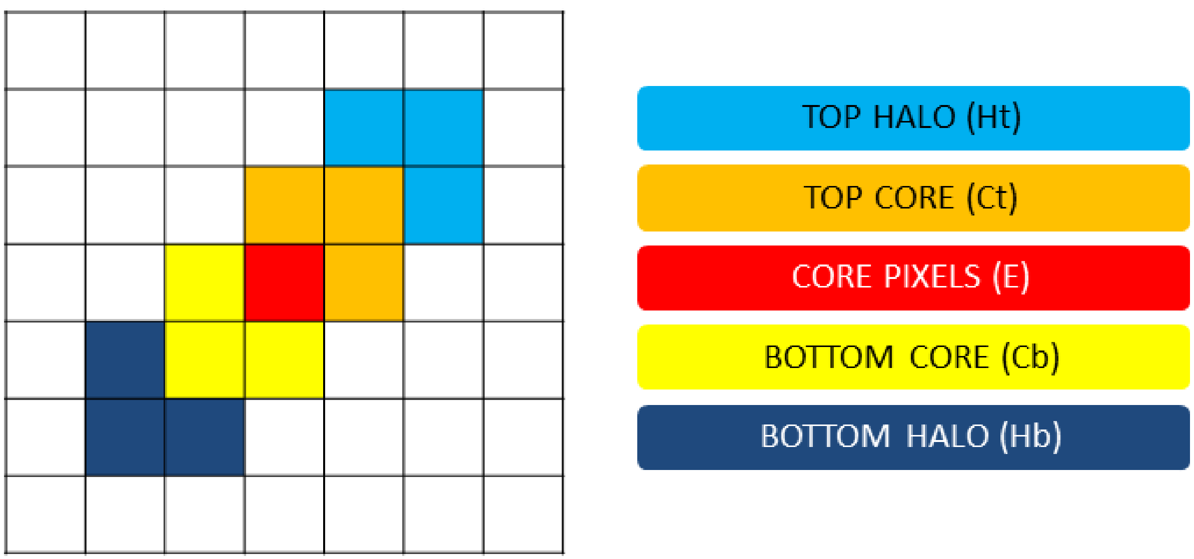
WDM

- The wavelet function used by cWB is the Wilson-Daubechier-Meyer (J.Phys.Conf.Ser. 363 (2012) 012032)
  - Transform the data into two orthonormal bases which form a dual frame
    - 0° phase
    - 90° phase
  - Uniform spectral leakage along all the frequencies



Pixel selection

- A fraction of the most energetic pixel are considered: core pixels
- For each core pixel cWB selects the neighbour in TF plane: halo
- Core pixels are considered if the energy in the core and/or in the halo is above a certain threshold



$$E_{core} > E_0$$

OR

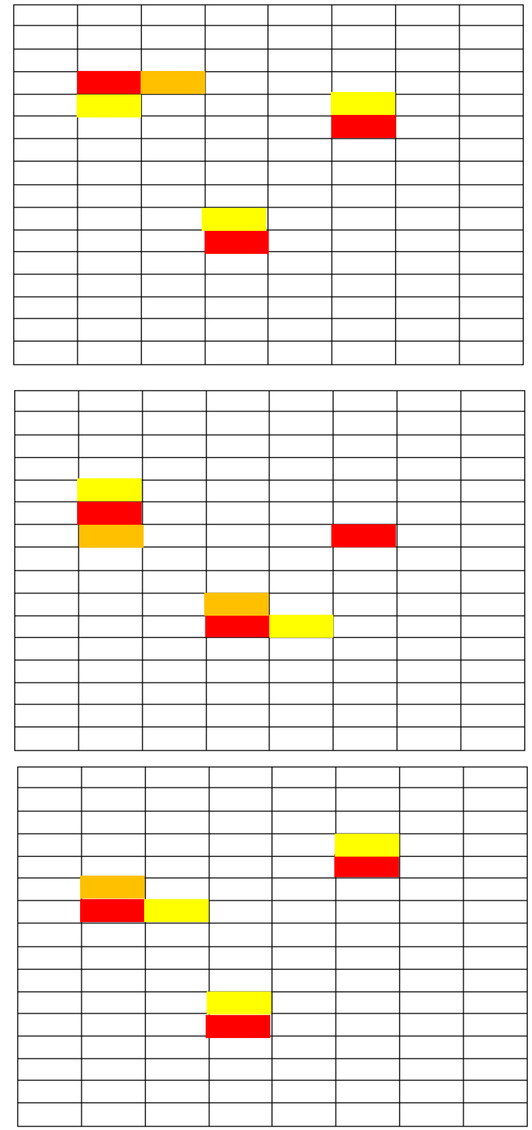
$$E_{core} > E_0 / 2$$

AND

$$\sqrt{E_{core} E_{halo} / 2} > E_0 / 2$$

Clustering

- TF pixels are selected according to coherence between detectors
  - Coherence verify if the energy of the pixels overcome a threshold
- Coincident TF pixels from different detectors are combined to form a cluster
- The cluster identify an event
- Cluster for each TF map are combined to form a supercluster
- Likelihood is calculated on the supercluster



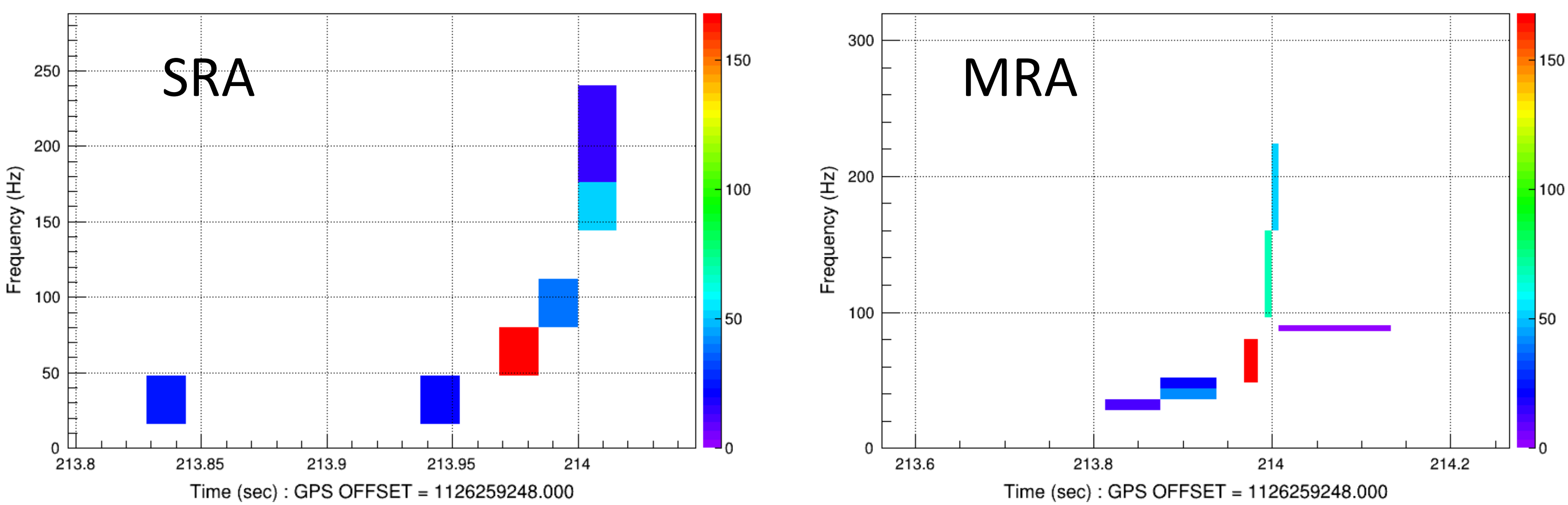
Principal Component Analysis

Using multiple TF trasform allows to find what is the optimal resolution for a given signal (Single-Resolution Analysis).

Signal could show variable behaviour along the TF plane (two compact object coalesce)

Multi-Resolution Analysis: include in the likelihood TF pixels with different resolution

Binary System with two Black Holes of 25 & 60 Solar Masses

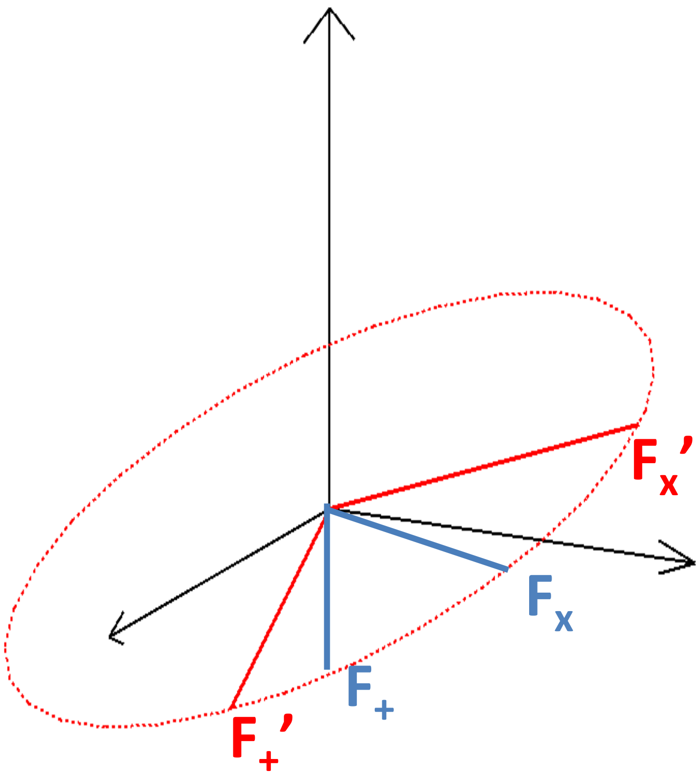


Dominant Polarization Frame

- Defining a multi-dimensional space where axes are defined by detectors
- Likelihood is invariant if we apply any rotation on the Wave Frame coordinate
  - "Rotation" on the antenna pattern

$$F_+^{'} = F_+ \cos(\Psi) - F_\times \sin(\Psi)$$

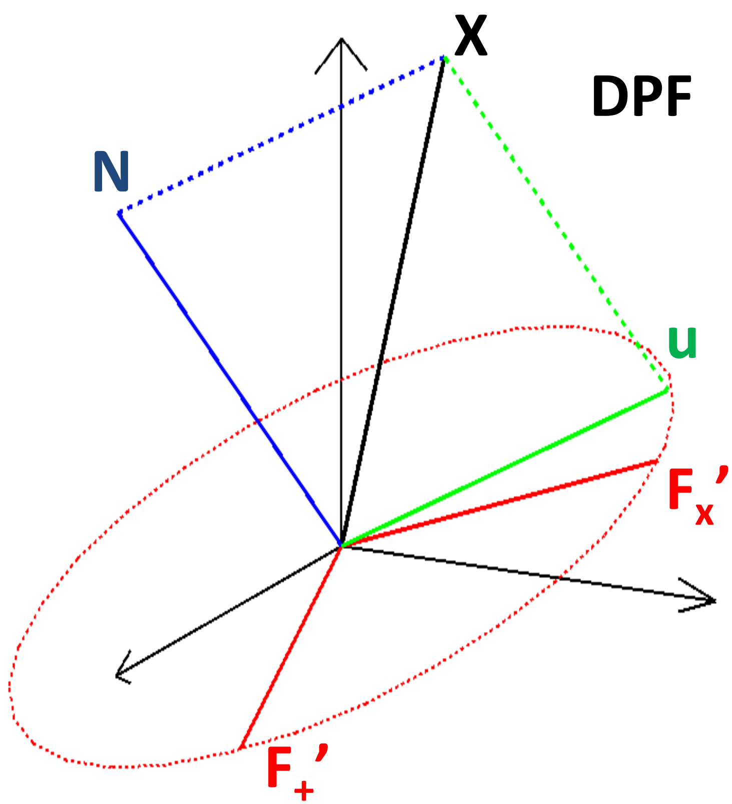
$$F_\times^{'} = F_+ \sin(\Psi) + F_\times \cos(\Psi)$$



- The Dominant Polarization Frame is that angle which gives the following properties to the Antenna patterns
  - Orthogonality $\quad F_+ \cdot F_\times = 0$
  - F_+ is dominant $\quad |F_+| \geq |F_\times|$

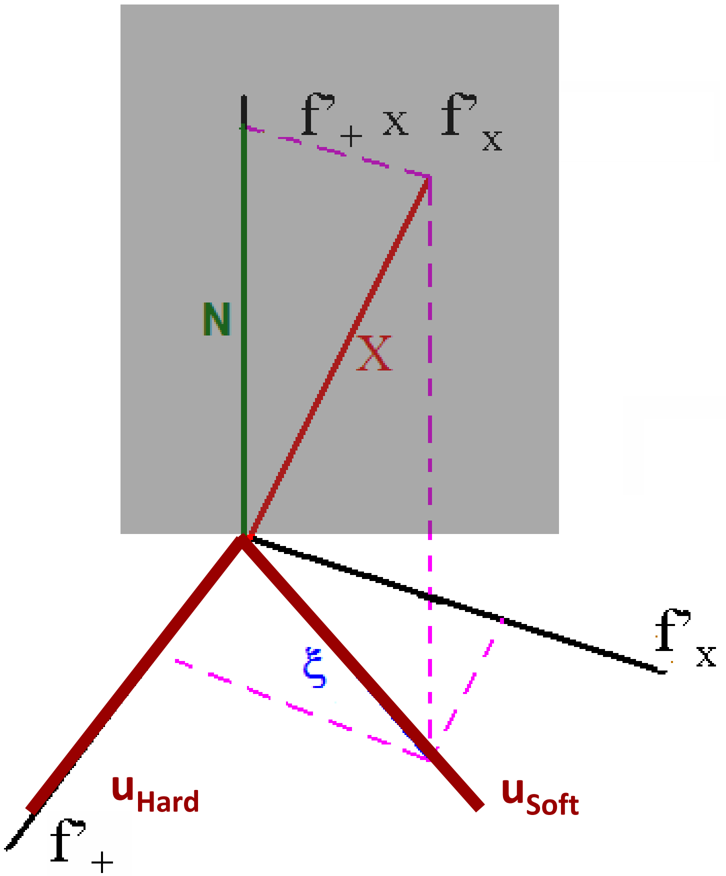
Maximum Likelihood

- In the DPF, detector response corresponding to maximum likelihood is the projection of X vector on f'_+, f'_x plane
- Null stream (N) described the noise after the likelihood subtraction
- Noise could be also in the f'_+, f'_x plane: use of regulator
- The likelihood projection on the axes is the contribution for each detector to the event energy, the rest is the coherent energy



Two detector case

- For two detector, maximum likelihood approach does not distinguish properly the data from the noise
  - Only two dimension: projection remains the original data
- Regulator approach: project the data on the F+ vector in the DPF
  - Hard regulator



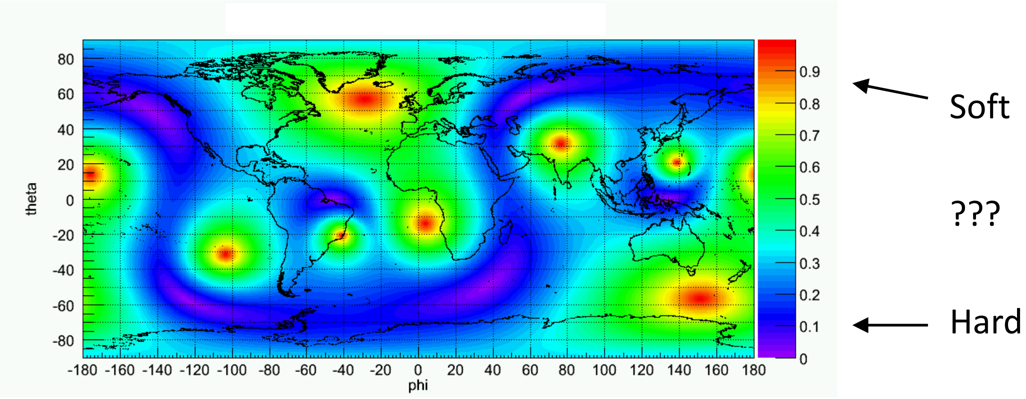
Regulators

- Detectors are not equally sensitive over all the sky
- Where a detector is blind, the network is lacking of detectors
  - Applying regulator

- Define a quantity over the sky which gives the value of involved detectors in the network for each sky position

  - Apply Hard regulator for sky position that have N<=2
  - Apply maximum likelihood (Soft regulator) for the rest

L1H1V1 $|f_x|/|f_+|$ in the DPF

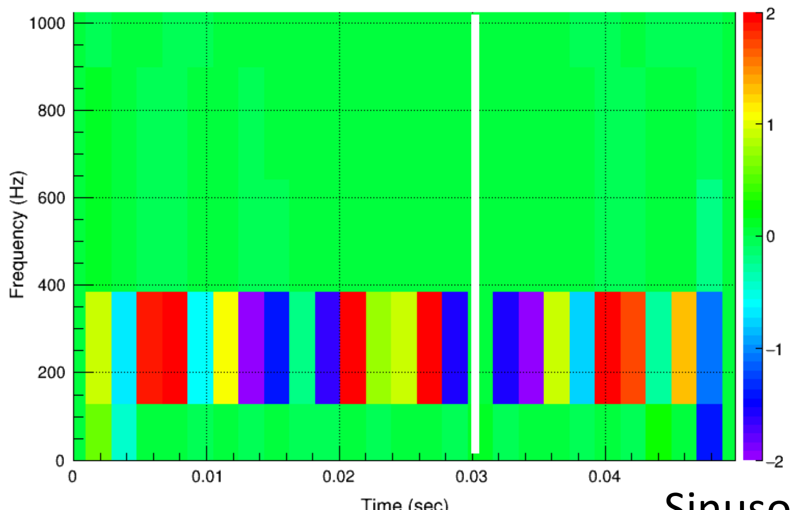


"Best" regulator depend on the sky position

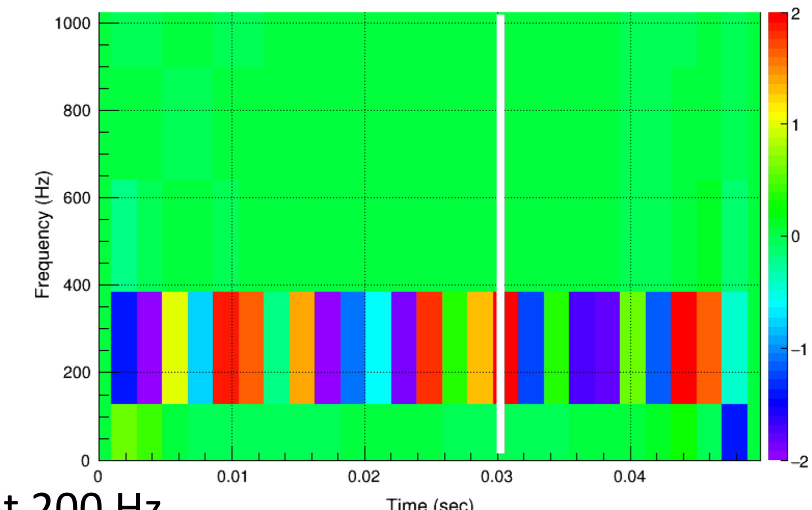
Dual stream analysis

- Apply a phase shift of 90° to data
  - Obtain an indipendent and complete characterization
- Pixel selection: it is not assured that same pixels are above the threshold
  - Considering two data stream: more complete description of the signal
- Combine 0° likelihood and 90° likelihood

0° phase



90° phase



Sinusoide at 200 Hz

Dual Stream Phase Transform

For a polarized wave each time-frequency pixel can be parametrized as :

- $s = F_+ * q + e * F_x * Q$ 0^o phase response
- $S = -F_+ * Q + e * F_x * q$ 90^o phase response

s, S are the DSP components, e is the ellipticity

The ***DSP Transform*** is defined as :

- $w = s * c_{dsp} + S * s_{dsp}$
- $W = S * c_{dsp} - s * s_{dsp}$

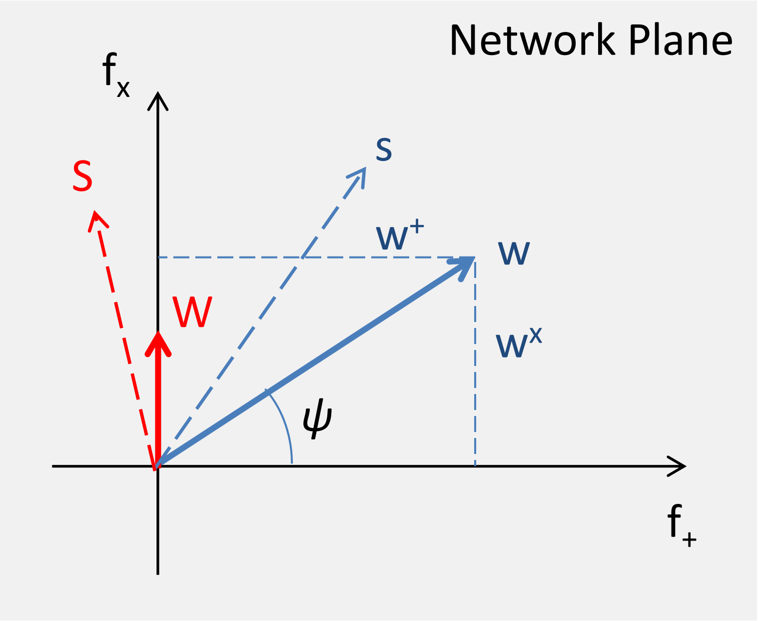
where

- $c_{dsp} = N*(\mathbf{f_+},s)/|f_+|^2$ and $s_{dsp} = N*(\mathbf{f_+},S)/|f_+|^2$ (x,y) is the scalar product and N is the sin,cos normalization factor : $N^2 = 1/(c^2_{dsp} + s^2_{dsp})$

w,W define the DSP pattern and are expressed as:

- $w^+ = \mathbf{f_+} * \{(1+e^2) + (1-e^2) * \cos[2*(d-p)]\} * N/2$
- $w^x = -\mathbf{f_x} * (e^2-1) * \sin[2*(d-p)] * N/2$
- $W = \mathbf{f_x} * e * N$

W is aligned with $\mathbf{f_x}$, $N = (q^2+Q^2)^{1/2}$



The angle ψ is defined by :

- the *DPF angle d*
- the *polarization angle p*
- *the ellipticity e* : |e|<=1

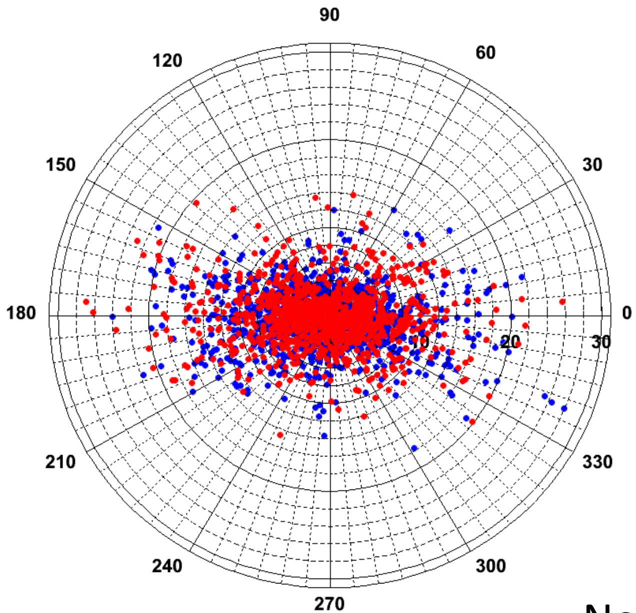
For a polarized wave *e and p* are fixed but *d* can be different for each pixel because PSD change with frequency

- This could produce a dispersion of the ψ angle

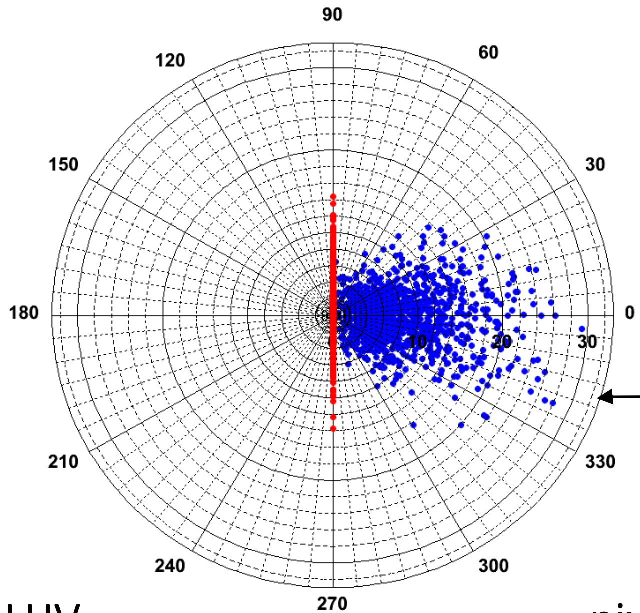
Simple example

Dual Stream Data

phase 0° : phase 90°



Dual Stream Phase (DSP) Transform



Network : ADV-LHV

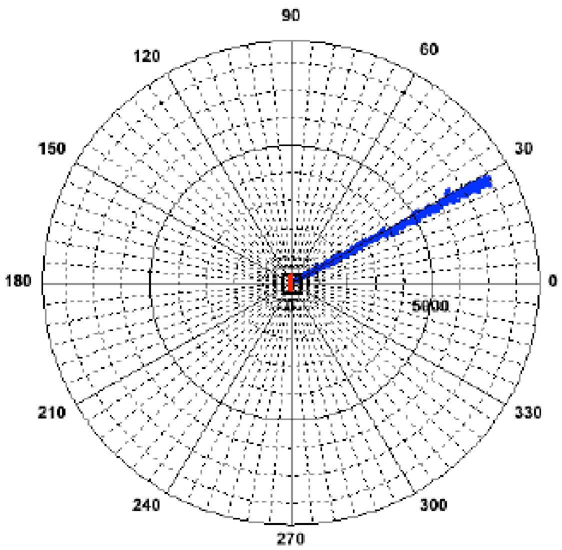
Injection : WNB100_100_0d100 SNR=100

pixels are randomly distributed

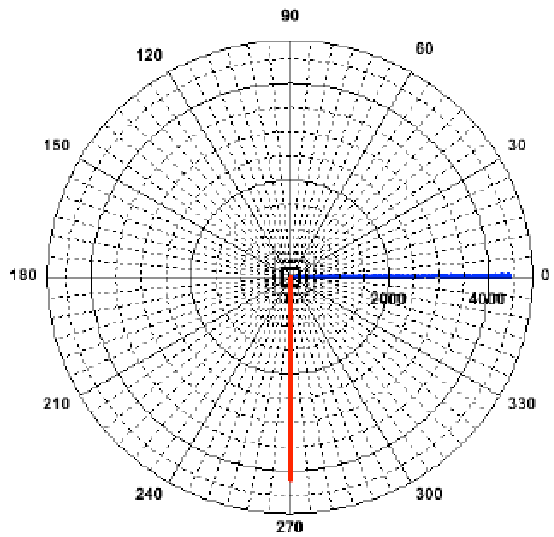
- No particular polarization, ellipticity of the pixels is random and no patterns are produced

Polarization constraint

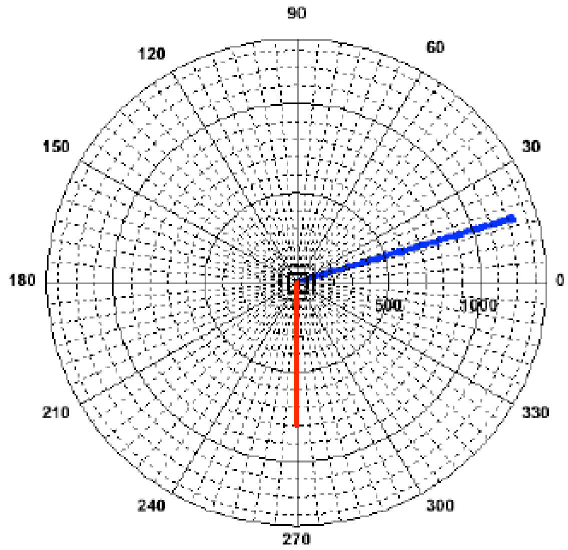
- The use of dual stream allows to define a search focusing on the expected polarization of the signal
- Applying the DSP transform, it is possible to characterize the likelihood of 0° and 90° according to the polarization pattern



Linear



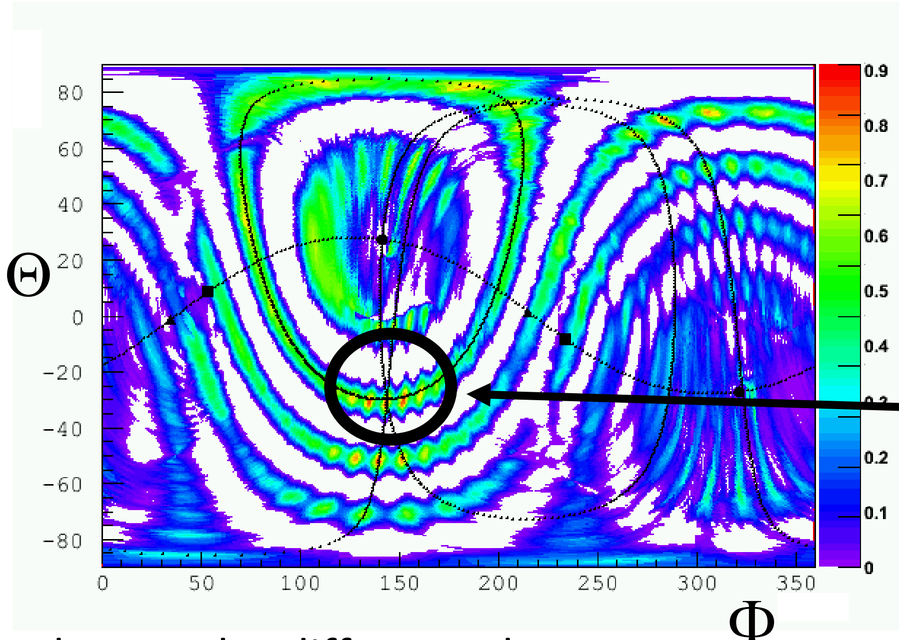
Circular



Elliptical

More you introduce restrictions more the glitches rejection is efficient, but the waveform detection is improved only for restricted signals

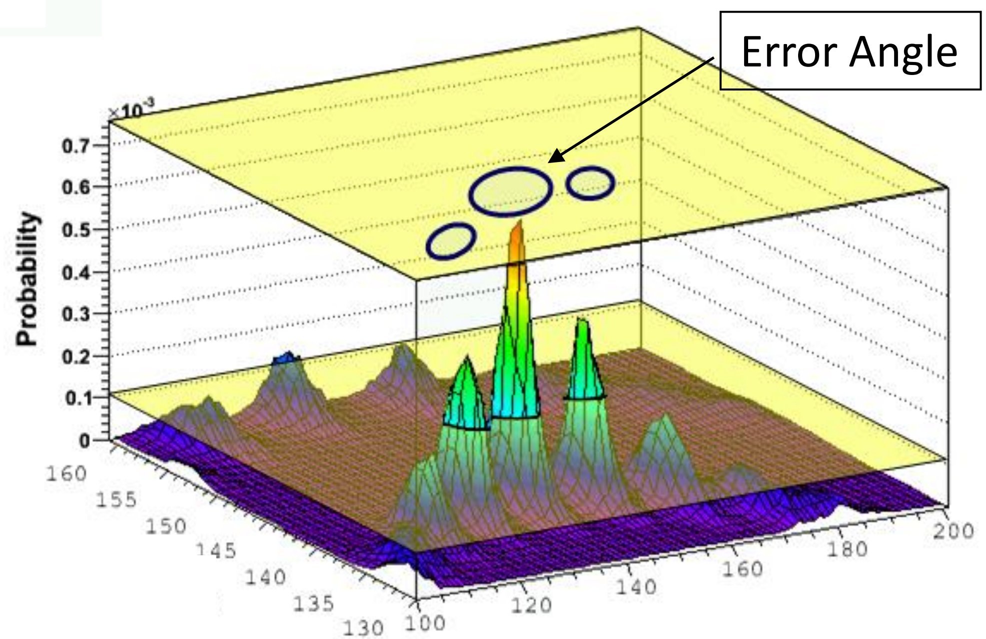
Coordinate reconstruction



Likelihood Sky Map shows how consistent are reconstructed waveforms and time delays as a function of Θ, Φ.

Maximum likelihood point to reconstructed direction

- The angular difference between injected and reconstructed position gives an estimation of the reconstruction error
- Error Angle: sum of sky pixel with likelihood greater than injected position
  - Likelihood is used as a ranking parameter
  - May be composed of disjoint areas in the sky

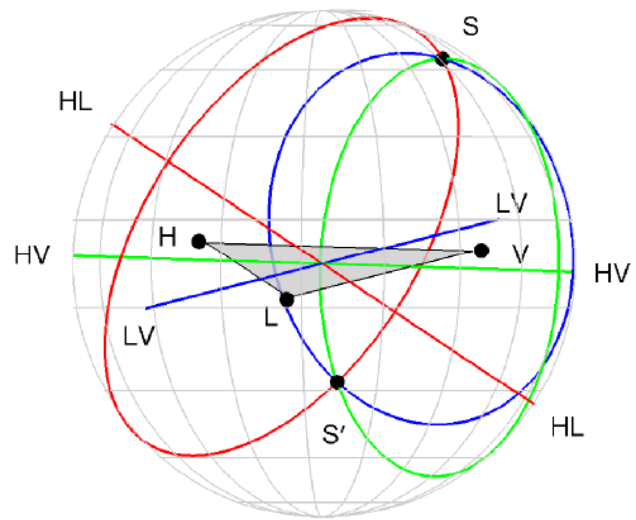


Direction Issues

- Reconstruction of source direction is affected by many problems:
  - Network Geometry
  - Sky segmentation
  - Antenna pattern

Direction Issues

- Reconstruction of source direction is affected by many problems:
  - Network Geometry
  - Sky segmentation
  - Antenna pattern

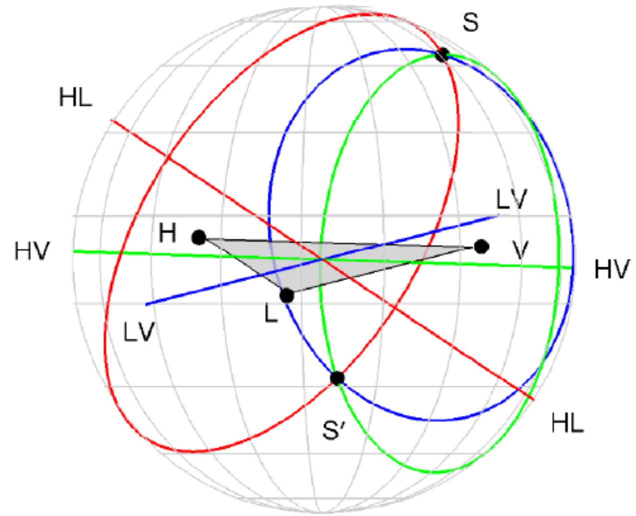


Time delays between three locations give an ambiguity of two possible locations

Can be resolved using antenna pattern informations

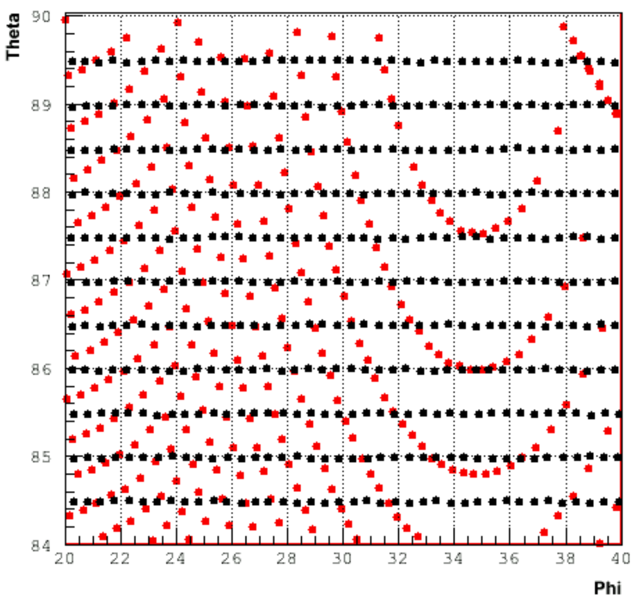
Direction Issues

- Reconstruction of source direction is affected by many problems:
  - Network Geometry
  - Sky segmentation
  - Antenna pattern



Time delays between three locations give an ambiguity of two possible locations

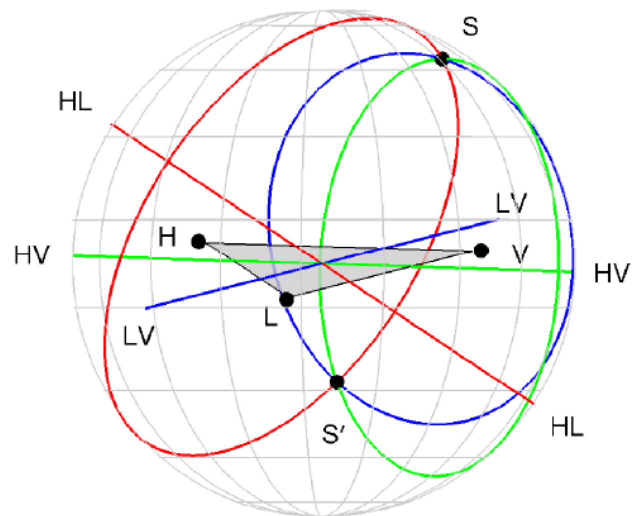
Can be resolved using antenna pattern informations



The discretization is sfavoureable for some zones in the sky

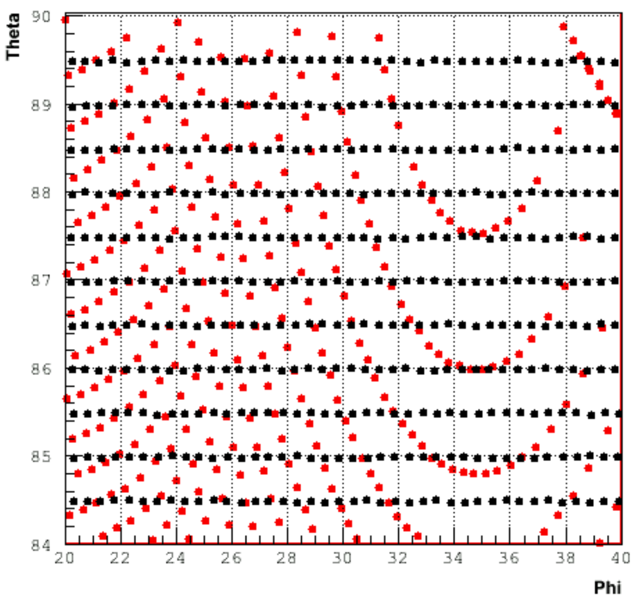
Direction Issues

- Reconstruction of source direction is affected by many problems:
  - Network Geometry
  - Sky segmentation
  - Antenna pattern



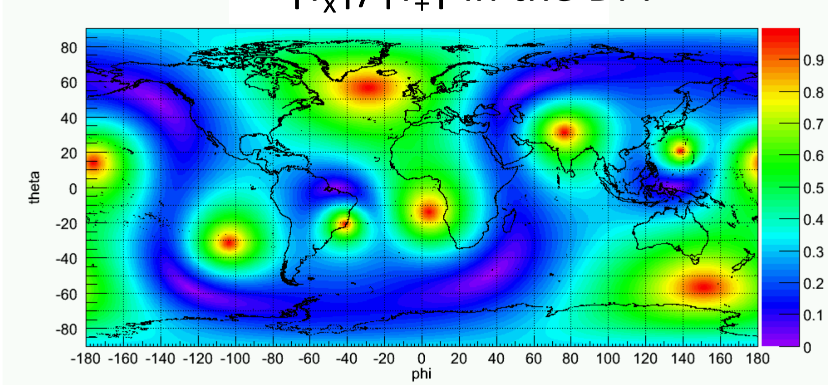
Time delays between three locations give an ambiguity of two possible locations

Can be resolved using antenna pattern informations



The discretization is sfavoureable for some zones in the sky

$|f_x|/|f_+|$ in the DPF

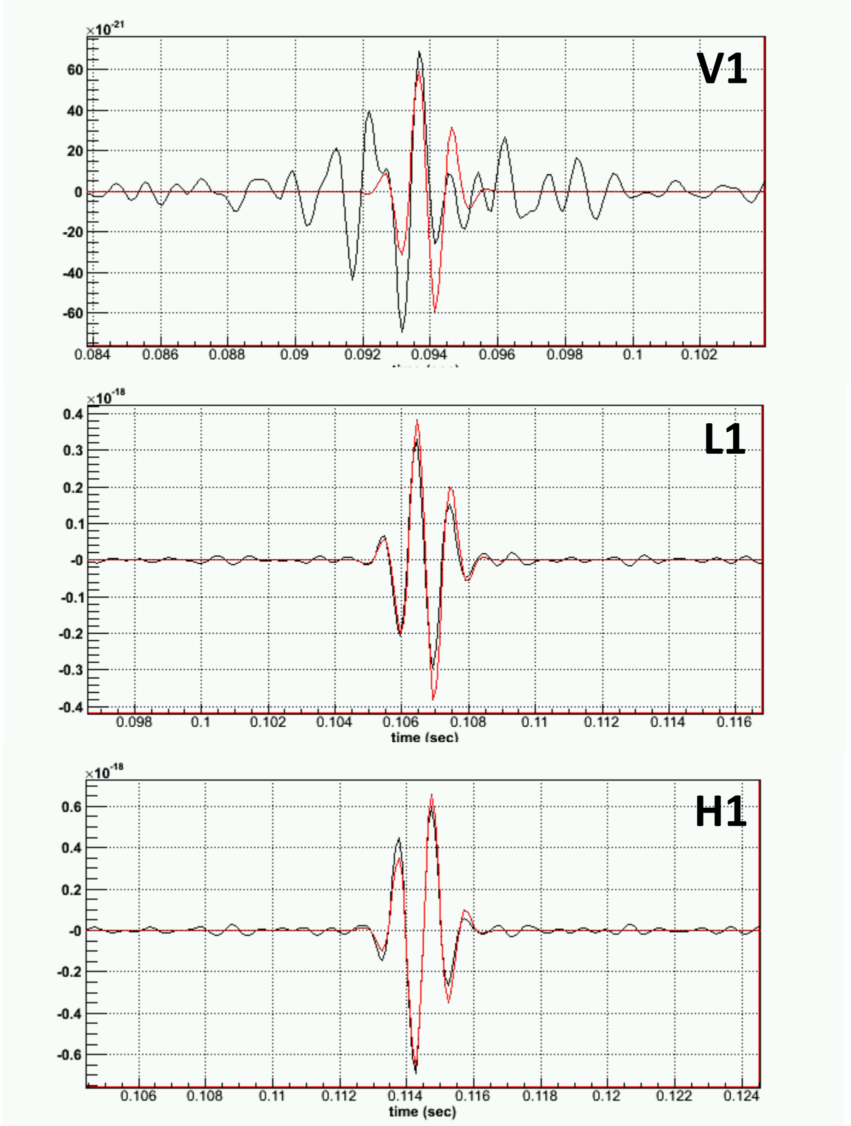


Where detector is blind is not possible to see

Waveform reconstruction

- The detector response vector in the DPF frame gives our solution
- From this solution we can recover the original detector response of each detector
  - We reconstruct the GW signal for each detector
- Detector response can be confronted with source models for extraction of the source parameters

SG2226Q9 injection



Post Production Analysis

- Necessity of a decision rule to identify a reconstructed event as a possible candidate

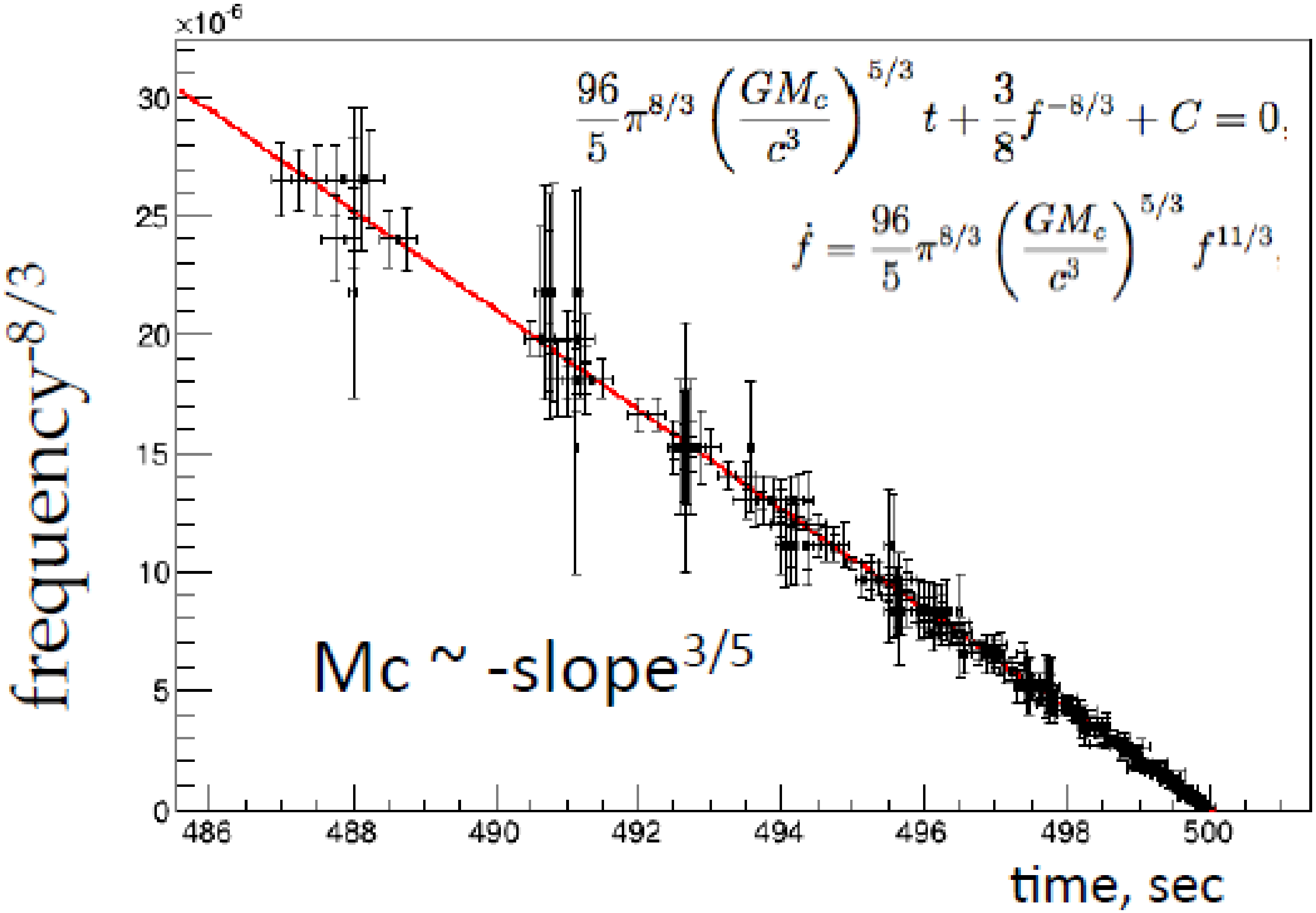
- Event selection
  - Effective correlated SNR: $\rho > \rho_{th}$
    This is the effective correlated SNR of the signal, averaged over all detectors. Used for the False Alarm Rate estimation
  - Network correlation coefficient: $cc > cc_{th}$
    This is a correlation test based on the coherent (Ec) and null (Null) energies.

$$cc = \frac{Ec}{Null + |Ec|}$$

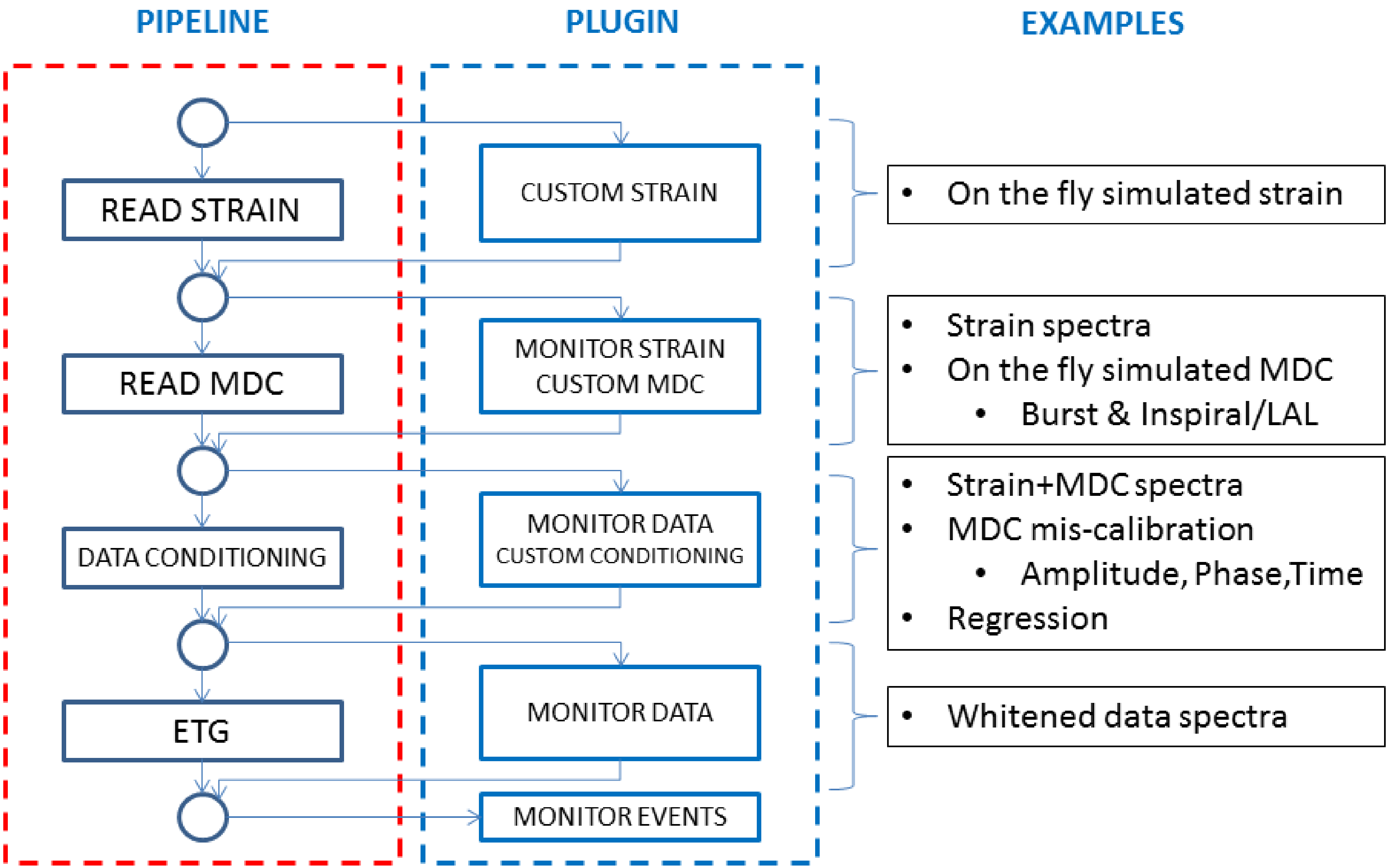
$$\rho \approx \sqrt{\frac{SNR}{2 \cdot N_{\mathrm{det}}}}$$

Signal classification

- Chirp mass can be estimated from the TF data without detail knowdledge of the waveforms

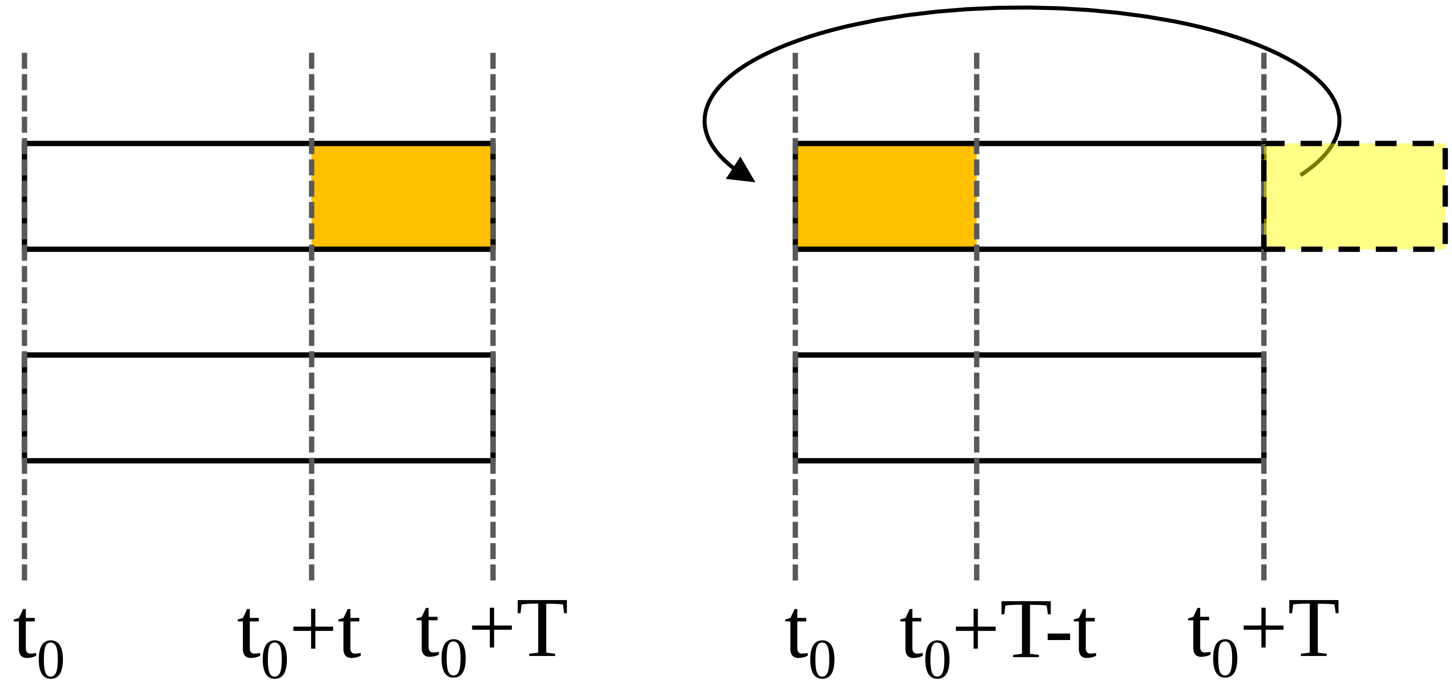


Plugin


PIPELINE
PLUGIN
EXAMPLES
READ STRAIN
CUSTOM STRAIN
On the fly simulated strain
READ MDC
MONITOR STRAIN
CUSTOM MDC
Strain spectra
On the fly simulated MDC
Burst & Inspiral/LAL
DATA CONDITIONING
MONITOR DATA
CUSTOM CONDITIONING
Strain+MDC spectra
MDC mis-calibration
Amplitude, Phase,Time
Regression
ETG
MONITOR DATA
Whitened data spectra
MONITOR EVENTS

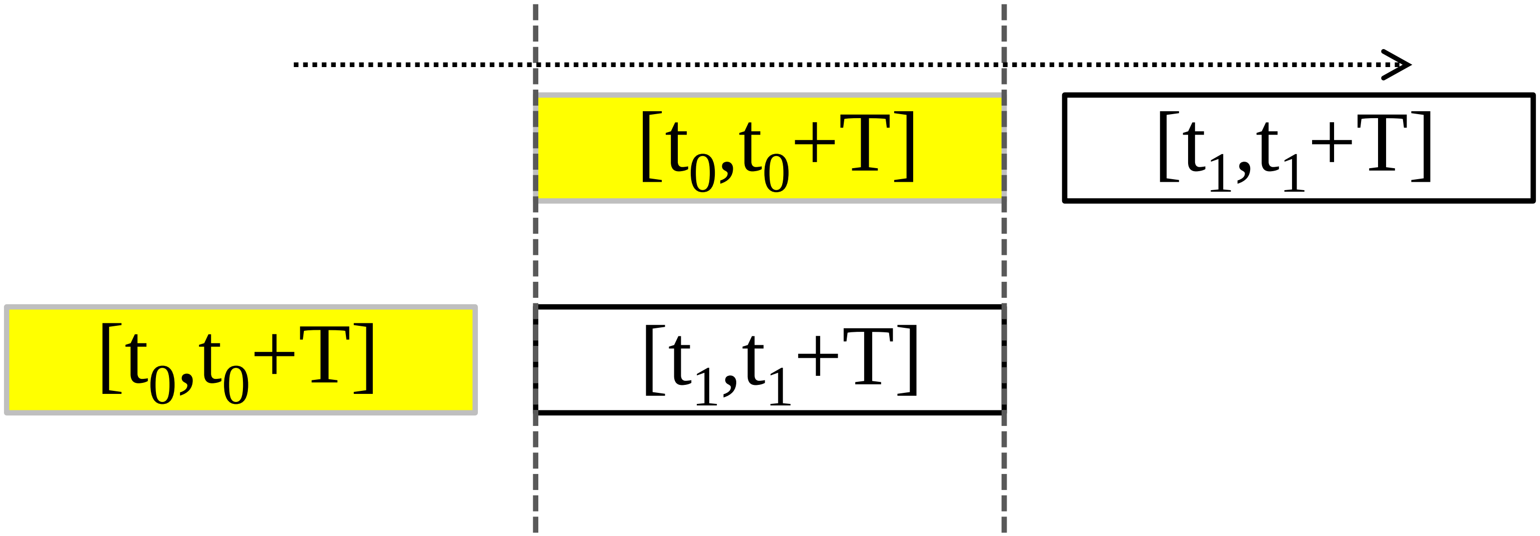
Background

- To characterize noise, we perform time shift between detectors
  - Minimum shift greater than maximum delay between detector
  - Triggers from shifted data cannot be due to GW
- Time shift is applied circularly
  - It can be applied to all detectors
  - Each detector has its independent shift
  - No losing time

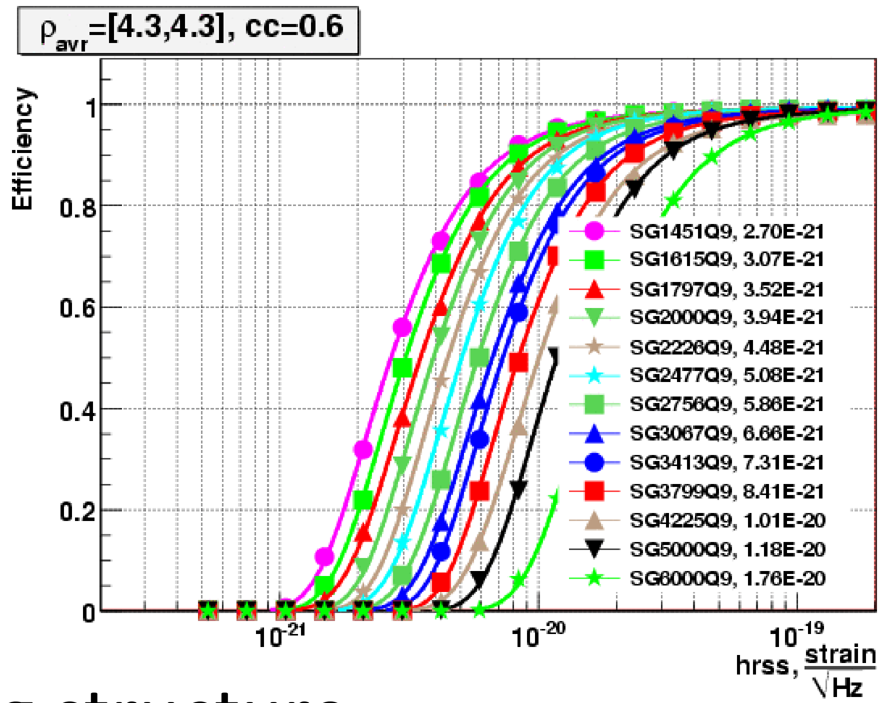


Background (2)

- Maximum shift allowed is equal to the segment length of analysis
- It could be necessary to increase statistic
- Super-lags: consider as “coincident” segment chunks that have a time shift greater than the segment lenght
- In principle any time shift is possible.



Injections engine



- General idea: Calculate search sensitivity with efficiency curve injecting waveforms at different amplitude

- cWB has various injecting structure:
  - Discrete hrss values (SNR vary along the sky, according to antenna pattern)
  - Discrete SNR values (hrss is adjusted along the sky, according to antenna pattern
  - User-defined hrss distribution

Signal classes

- cWB internal engine:
  - SineGaussian (SG)
  - Gaussian (GA)
  - White Noise Bursts (WNB)
  - Ring Down (RD)
  - Eccentric Binary Black Holes

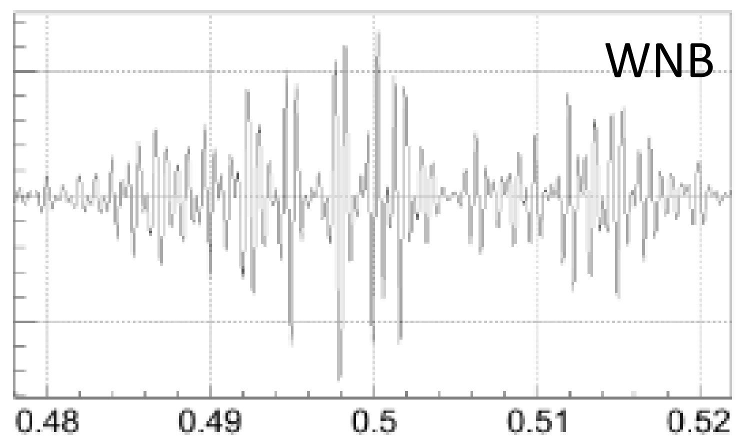
$$h_{SG} = e^{-(t/\tau)^2} \sin(2\pi f t)$$

$$h_{GA} = e^{-(t/\tau)^2}$$

$$h_{RD} = e^{-(t/\tau)} \cos(2\pi f t)$$

- LAL waveforms:
  - Compact Binary Coalescence, with mass from NS-NS to BH-BH

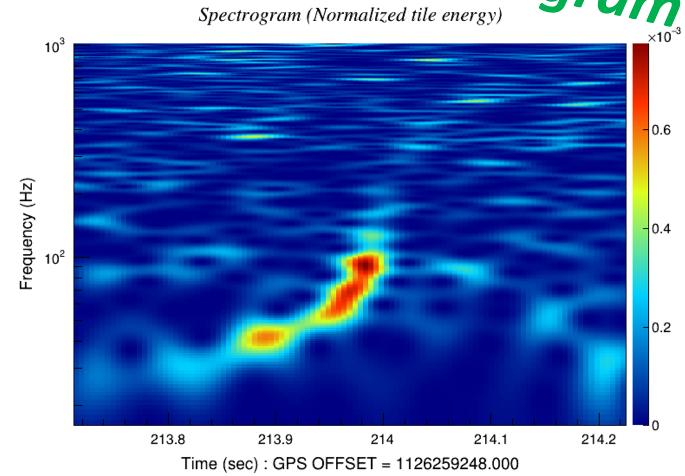
- User defined waveforms
  - Giving h_+ and h_x to the pipeline



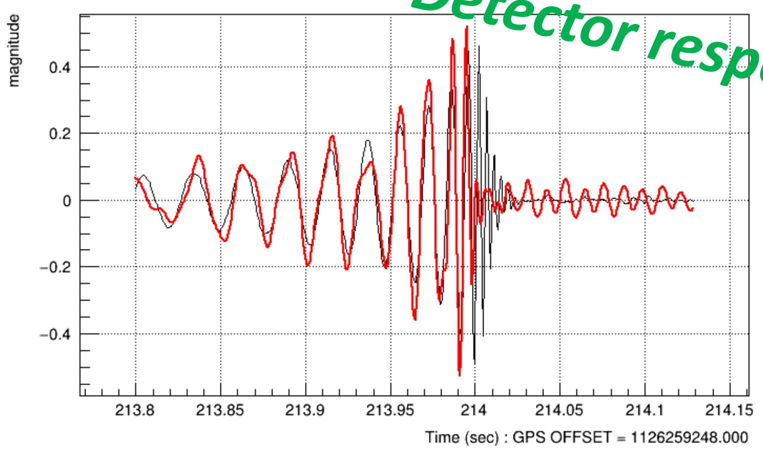
Coherent Event Display

- CED is a detailed study of a particular event reporting more information than usual analysis

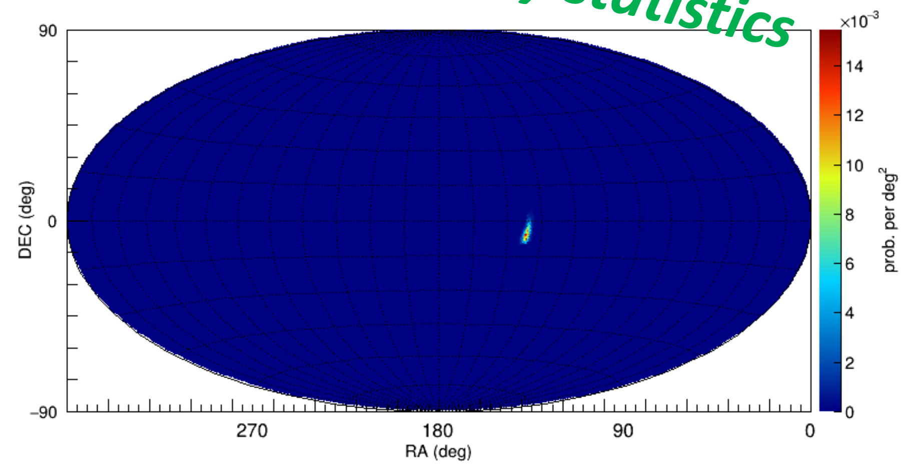
Spectrogram



Detector response



Sky statistics



And more…

Bibliography

- Virtual box
- cWB flowchart
  - Regression: Paper (arXiv)
  - Whitening
  - TF map: Paper
  - Clustering
  - Dual stream and regulators
  - Chirp mass: Paper (arXiv)
- Background (superlags)
- Injection engine
- CED
- Plugins

Exercise

- Consider the example on the manual and repeat step by step, understanding the various stages and looking at the result