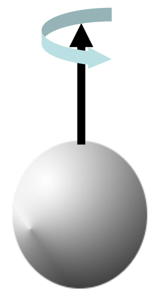
Data Analysis for CW signals

Pia Astone & Cristiano Palomba – INFN Roma

Plan of the lecture



1. **Signal characterization**
2. **Detection theory**
3. **Basic analysis methods**
4. **Targeted search**
5. **Blind search**
6. **Issues when analyzing real data**

Signals characterization - 1

• We classify as “continuous” a GW signal with duration longer than the typical observation time of a detector

• CW signals are expected to be emitted by various sources containing neutron stars: tri-axial, wobbling, accreting, in binary systems,…

• **<u>We KNOW that potential sources of CW exist</u>**: ~2,000 NS are observed in EM (mostly pulsar), 1 billion expected to exist in the Galaxy

• A fraction of these emit in the sensitivity band of ITF

• **<u>We DO NOT KNOW the amplitude of the emitted signals</u>**

Signals characterization - 2

A GW signal can be described through its polarization ellipse.

The polarization ellipse is characterized by the ratio $\eta = \frac{b}{a}$ between its axis (-1≤η≤1) and the polarization angle ψ (direction of the axis a of the ellipse).

Then, the GW is:

$$\mathrm{h}(t) = h_0 \cdot \left(H_+ \cdot \mathbf{e}_\oplus + H_\times \cdot \mathbf{e}_\otimes \right) \cdot e^{\left(j \cdot (\omega_0 t + \gamma) \right)}$$

$$H_+ = \frac{1}{\sqrt{1+\eta^2}} \cdot \left(\cos(2\psi) - j \cdot \eta \cdot \sin(2\psi) \right)$$

$$H_\times = \frac{1}{\sqrt{1+\eta^2}} \cdot \left(\sin(2\psi) + j \cdot \eta \cdot \cos(2\psi) \right)$$

By specifying a particular source model we can express h_0 and η as a function of the source physical parameters.

Signals characterization - 2

For instance, considering an isolated neutron star rotating around a principal axis of inertia we can recover the 'classical' expression of the plus and cross wave components:

$$h_+ = h_0 \cdot \left(\frac{1 + \cos^2 \iota}{2} \right)$$

$$h_\times = h_0 \cdot \cos \iota$$

$$h_0 = \frac{4\pi^2 G}{c^4} \frac{I_{zz} \varepsilon f_{GW}^2}{d}$$

ι : angle between star spin axis and line of sight

$$\varepsilon = \frac{I_{xx} - I_{yy}}{I_{zz}}$$: equatorial ellipticity

$$f_{GW} = 2 f_{rot}$$

Signals characterization - 3

The expected signal maximum frequency is below 2kHz.

The source we are searching for are in the Galaxy, d<O(10kpc), (farther sources are not detectable).

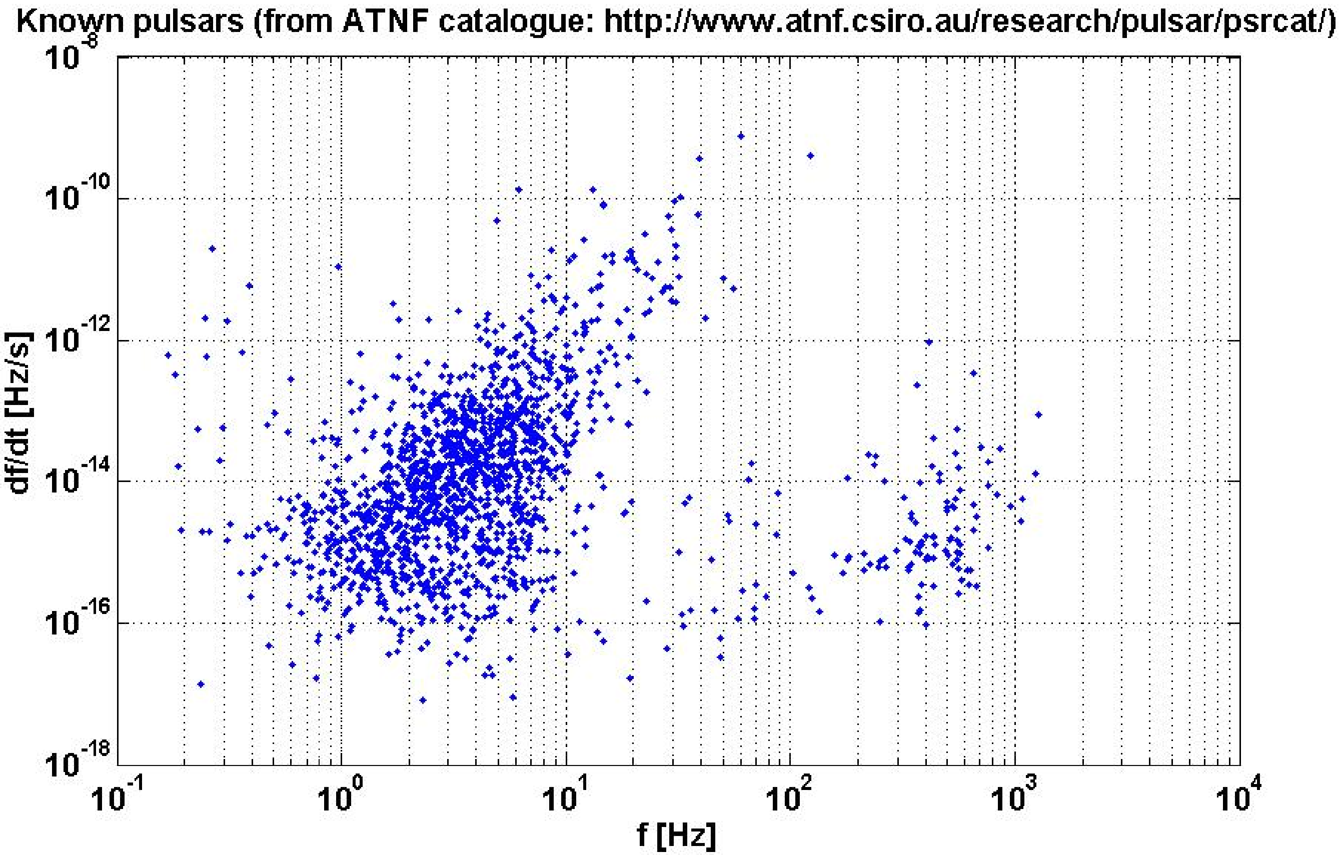
The ellipticity is largely unknown; $\varepsilon_{\max} \approx 10^{-6}$ expected but some exotic EOS foresee $\varepsilon_{\max} \approx 10^{-4}$ or even more.

$$h_0 \cong 10^{-27} \left(\frac{I_{zz}}{10^{38} kg \cdot m^2} \right) \left(\frac{10kpc}{d} \right) \left(\frac{f}{100Hz} \right)^2 \left(\frac{\varepsilon}{10^{-6}} \right)$$

It can reach $\sim 10^{-23}$ for a very near fast spinning NS.

Signals characterization - 4

The rotation frequency, and then the emitted signal frequency, slowly decreases due to the energy loss of the source (EM, GW hopefully…): **spin-down**



Signals characterization - 5

Assuming that the measured spin-down of a source is totally due to the emission of GW, we obtain an upper limit to the signal amplitude:

$$h_{sd} = 8 \cdot 10^{-25} \sqrt{\left(\frac{|\dot{f}|}{10^{-10}\,Hz/s} \right) \left(\frac{f}{100\,Hz} \right)^{-1}} \left(\frac{d}{1kpc} \right)^{-1}$$

spin-down limit

In a search of GW from known NS, we start to enter into a regime of astrophysical interest when the spin-down limit is beaten (for just one object so far: Crab in LIGO S5)

For some pulsars a more stringent limit to the maximum possible amplitude can be computed considering both the EM and GW contribution to the spin-down (Crab, Vela,...).

Signals characterization - 6

The received signal is affected by the **Doppler effect** associated to the relative source-detector motion.

The signal phase evolution respect to the solar system barycenter (SSB), which is approximately an inertial reference frame, is:

$$\phi(T) = \phi_0 + 2\pi \cdot f_0 \left(T - T_0\right)$$

ϕ_0 : signal phase at the time T0

T : SSB time (we neglect source proper motion)

Due to the detector motion and to some relativistic effects, the time at the detector t is not equal to the SSB time T:

Signals characterization - 7

$$T = t + \delta t = t + \frac{\vec{r} \cdot \hat{n}}{c} + \Delta_E + \Delta_S$$

$\vec{r}$: detector position in the SSB

$\hat{n}$: source direction versor in the SSB

Δ_E : Einstein delay (gravitational redshift and time dilation)

Δ_S : Shapiro delay (curvature of space-time in the SSB)

The signal phase at the detector can then be written as

$$\phi(t) = \phi_0(t) + \Delta\phi_{Doppler}(t)$$

$$\Delta\phi_{Doppler} = 2\pi \cdot f_0 \cdot \frac{\vec{r} \cdot \hat{n}}{c} + \text{rel. corr.}$$

Signals characterization - 8

Hence, the Doppler shift formula for the frequency is:

$$f(t) = f_0 \cdot \left(1 + \frac{\vec{V} \cdot \hat{n}}{c}\right) \text{ + relativ. corrections}$$

The detector velocity (in the SSB) is $\vec{V} = \vec{V}_{rot} + \vec{V}_{rev}$

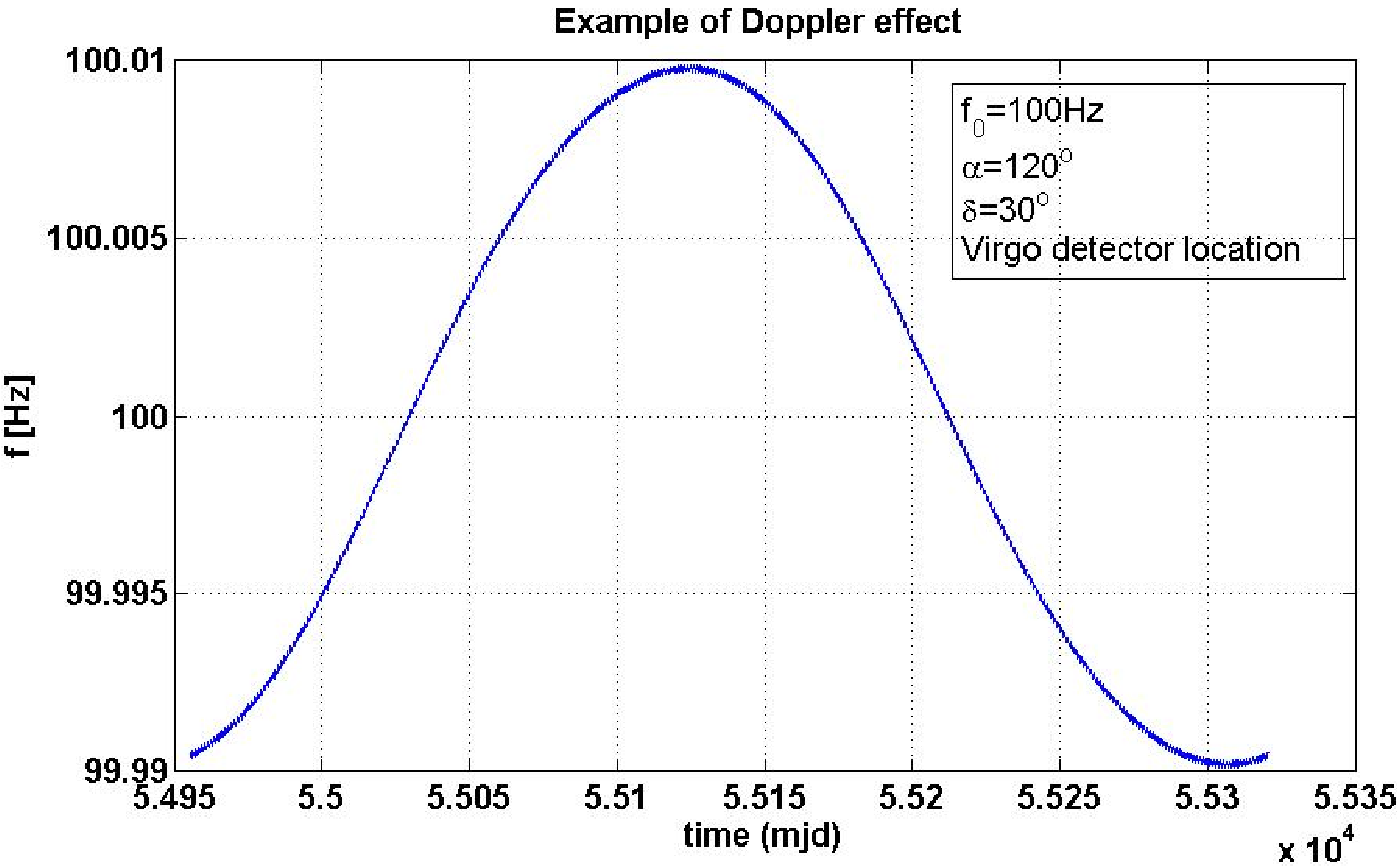
$V_{rev} \approx 30 km/s$ (directed along the ecliptic; P=1 sidereal year)

$V_{rot} \approx 0.32 km/s$ (at 45 deg latitude; tilted by ~23.4 deg respect to the orbital plane; P=1 sidereal day)

The total velocity vector makes small ~0.8 deg oscillations around the ecliptic.

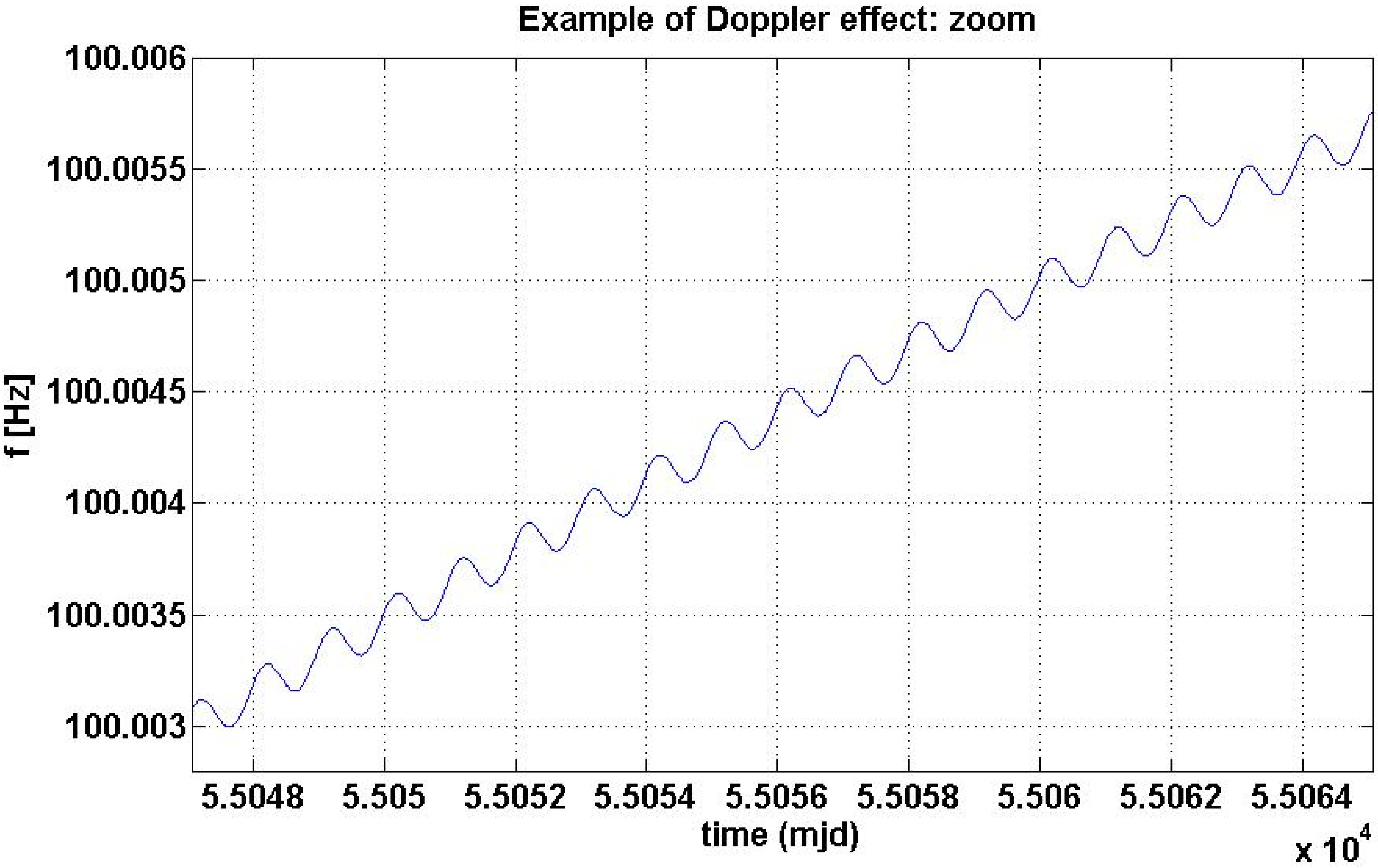
For sources in binary systems there are further terms in the Doppler formula due to the orbital motion.

Signals characterization - 9



The maximum Doppler shift is $\Delta f_{\max} = \frac{|\vec{V} \cdot \hat{n}|_{\max}}{c} \approx 10^{-4} f_0$

Signals characterization - 10



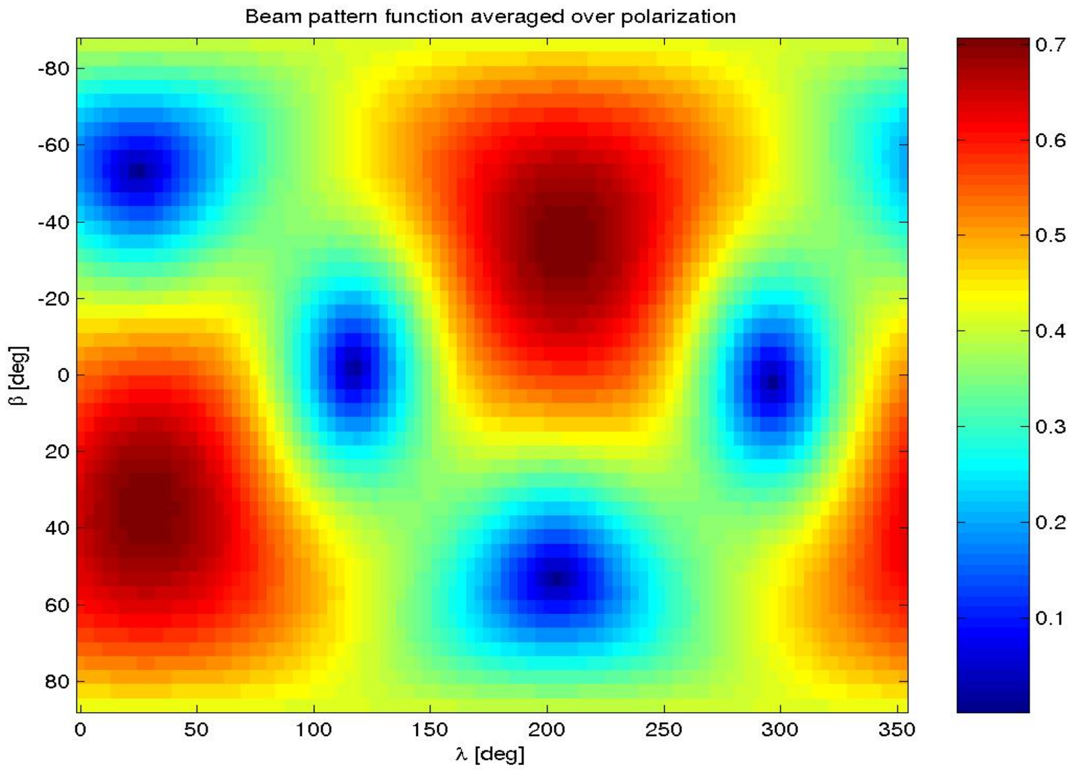
A common mistake: “Over short times the Doppler effect is small”.

Why is this sentence wrong?

Signals characterization - 11

The received signal is **amplitude modulated** due to the non-uniform antenna **sensitivity pattern** of the detector.

Signal at the detector : $h(t) = h_0 \cdot \left(A_+ \cdot H_+ + A_\times \cdot H_\times\right) \cdot \exp\left(j\left(\omega_0 t + \gamma\right)\right)$



Signals characterization - 12

The time-dependent antenna pattern is described by the two functions:

$$A_+ = a_0 + a_{1c}\cdot\cos(\Omega\cdot t) + a_{1s}\cdot\sin(\Omega\cdot t) + a_{2c}\cdot\cos(2\cdot\Omega\cdot t) + a_{2s}\cdot\sin(2\cdot\Omega\cdot t)$$

$$A_\times = b_{1c}\cdot\cos(\Omega\cdot t) + b_{1s}\cdot\sin(\Omega\cdot t) + b_{2c}\cdot\cos(2\cdot\Omega\cdot t) + b_{2s}\cdot\sin(2\cdot\Omega\cdot t)$$

$$a_0 = -\frac{3}{16}(1+\cos 2\delta)(1+\cos 2\lambda)\cos 2a$$

$$a_{1c} = -\frac{1}{4}\sin 2\delta \sin 2\lambda \cos 2a$$

$$a_{1s} = -\frac{1}{2}\sin 2\delta \cos\lambda \sin 2a$$

$$a_{2c} = -\frac{1}{16}(3-\cos 2\delta)(3-\cos 2\lambda)\cos 2a$$

$$a_{2s} = -\frac{1}{4}(3-\cos 2\delta)\sin\lambda \sin 2a$$

$$b_{1c} = -\cos\delta \cos\lambda \sin 2a$$

$$b_{1s} = \frac{1}{2}\cos\delta \sin 2\lambda \cos 2a$$

$$b_{2c} = -\sin\delta \sin\lambda \sin 2a$$

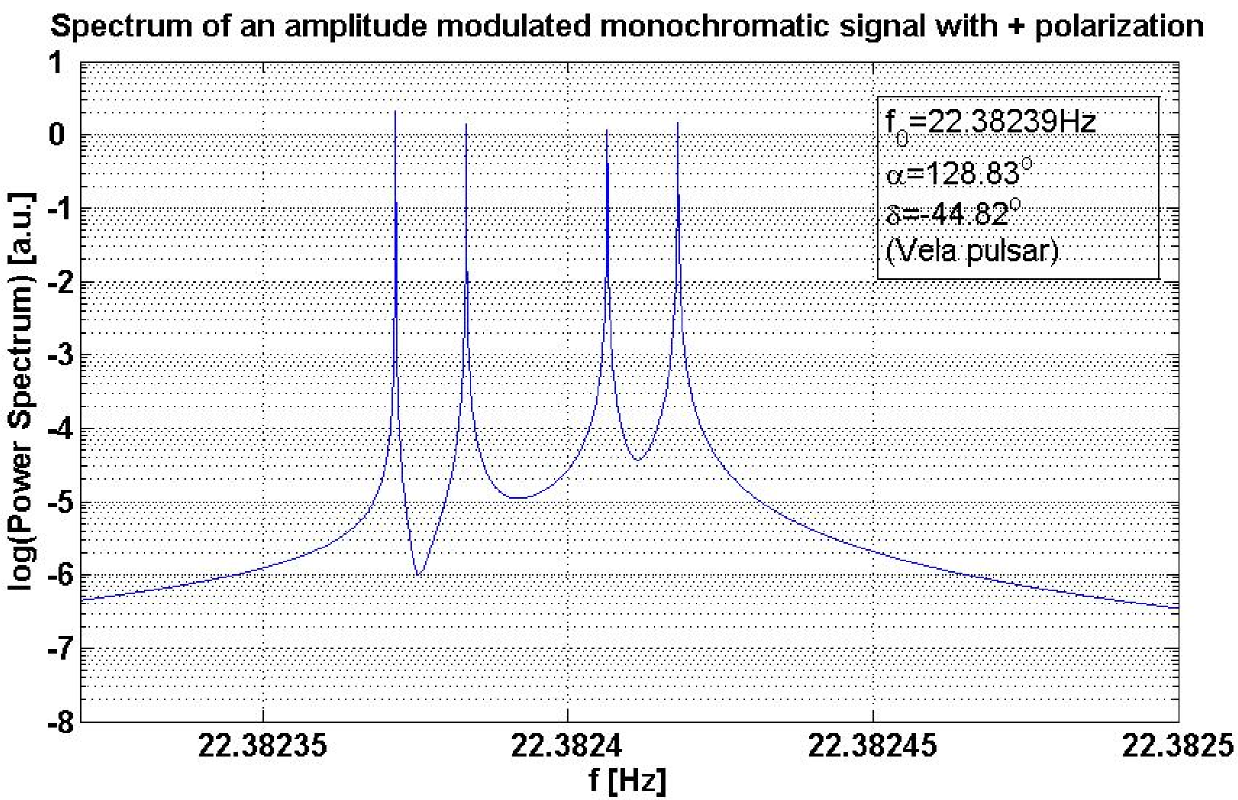
$$b_{2s} = \frac{1}{4}\sin\delta(3-\cos 2\lambda)\cos 2a$$

(α,δ) source coordinates, λ and a the latitude and azimuth of the antenna;
$\Omega\cdot t = \alpha - \Theta$, where Θ is the sidereal time.

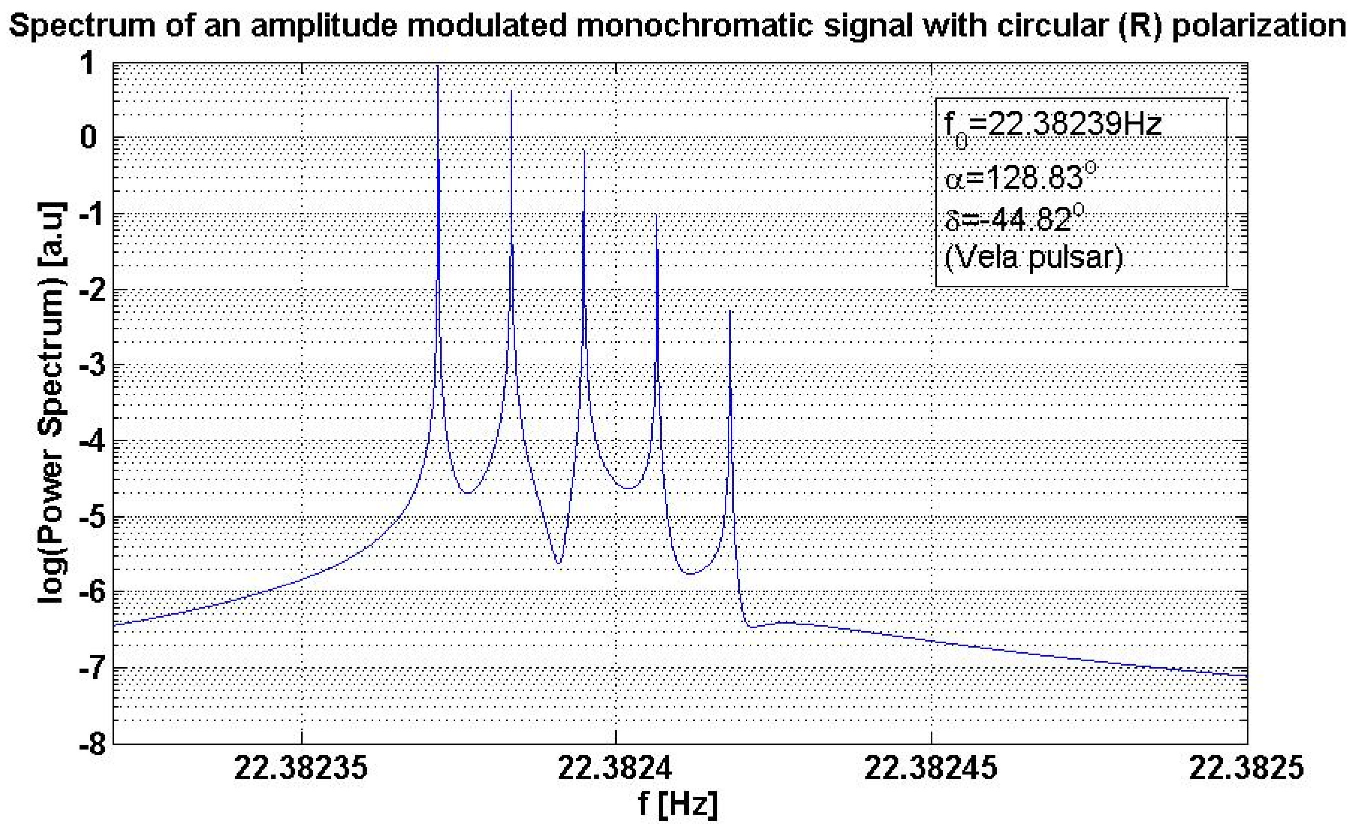
The modulation has period of one sidereal day and does not depends on the signal frequency.

Signals characterization - 13

Looking at the frequency domain, the amplitude modulation produces a spread of the signal power into five bands at frequencies $f_0, f_0 \pm f_{earth,rot}, f_0 \pm 2 f_{earth,rot}$



Signals characterization - 14

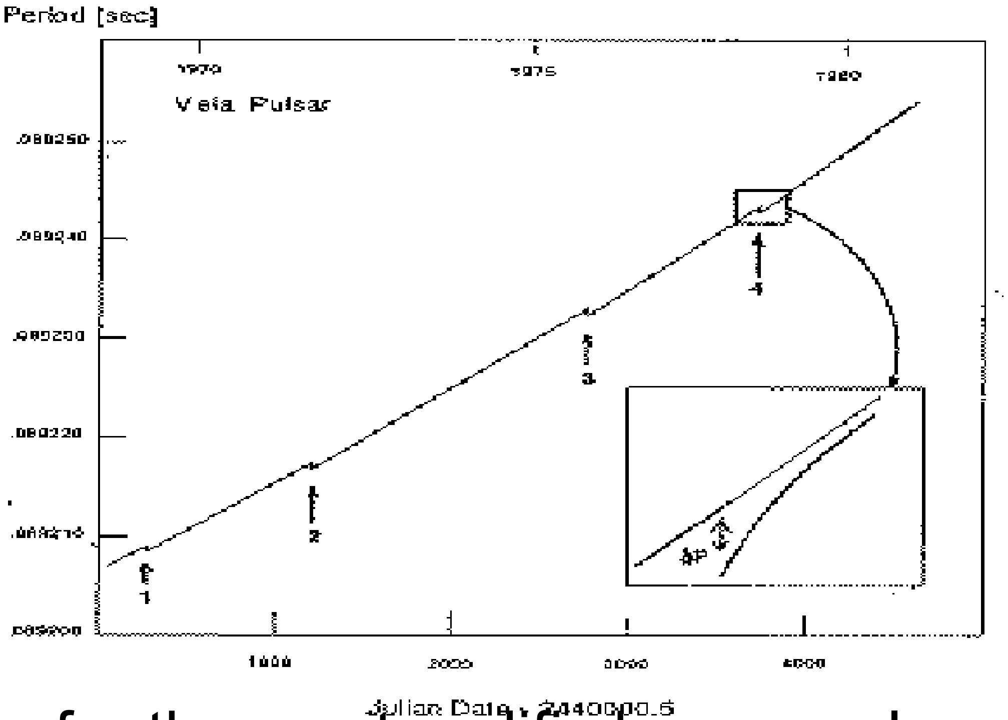


The frequency of the side bands is fixed (respect to the central one). The amount of power in the side bands depends on the source and detector parameters

Signals characterization - 15

There are two further complications that can affect the GW signal from real sources: **glitches** and **timing noise**

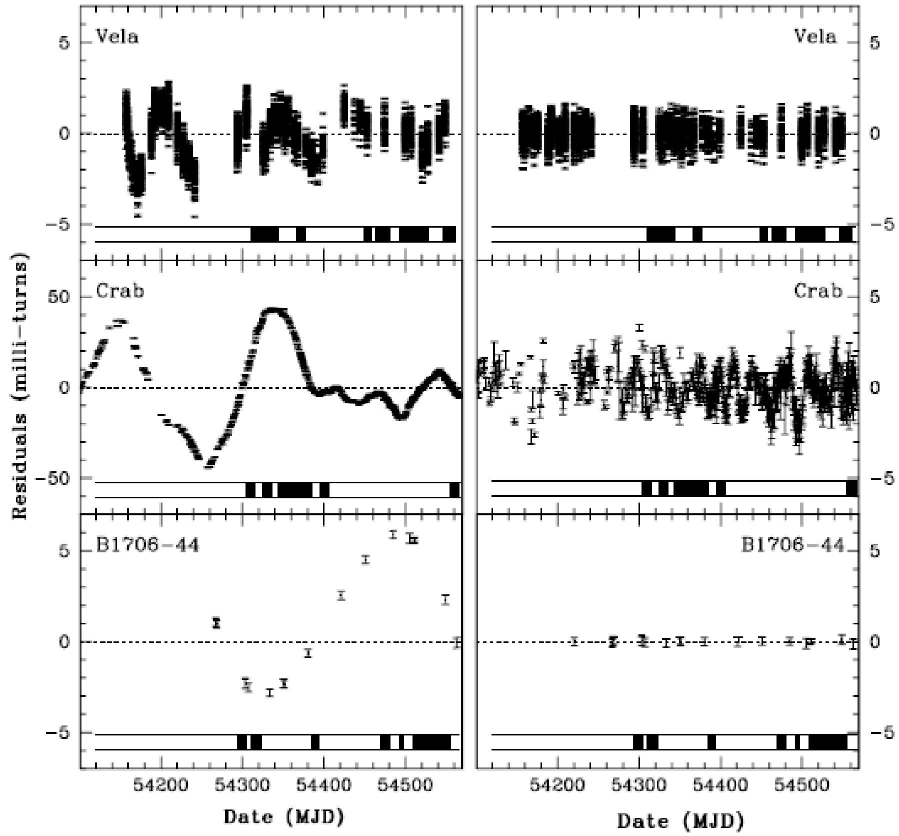
Glitches: sudden abrupt variations of rotational period and period derivative (star-quakes or interaction between crust and superfluid,…)



One every few years for the most prolific known pulsars (Vela, Crab,…)

Signals characterization - 16

Timing noise: random fluctuations of the star rotation frequency or EM signal phase (or both).



Pellizzoni et al, ApJ 691 (2009)

Not clear if it affects also the GW signal

Detection theory - 1

Tipically, **GW signals are buried into the detector noise**. Then, the data at the output of a detector form a stochastic process

$$x(t) = n(t) + h(t)$$

The detection of a signal into the data, and the estimation of its parameters, are a statistical problem.

The key idea behind signal detection is that the presence of a signal changes the probability distribution of the data.

Given our data, we compute a test statistics $\Lambda_{\exp}(x)$, a properly chosen function of the data. This means suitably filtering the data.

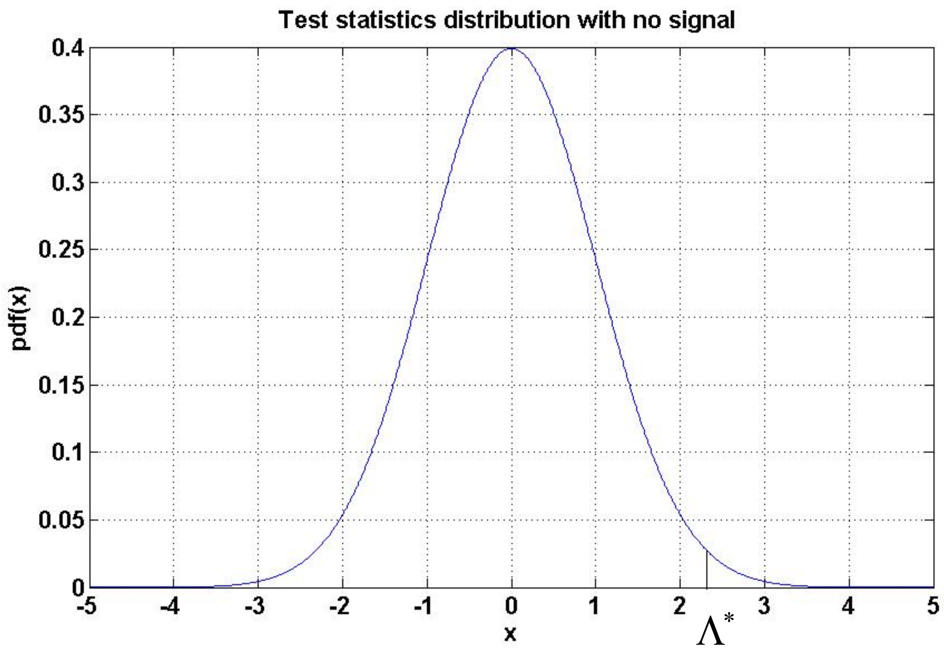
Then, we check if its value is compatible with the pure noise distribution (hypothesis H0) or not (alternative hypothesis H1).

Detection theory - 2

We choose a threshold Λ^* and compare it with $\Lambda_{\exp}(x)$

If $\Lambda_{\exp}(x) < \Lambda^*$ we conclude that no signal is present in the data

If $\Lambda_{\exp}(x) > \Lambda^*$ we reject H0 and conclude that a signal is present



Detection theory - 3

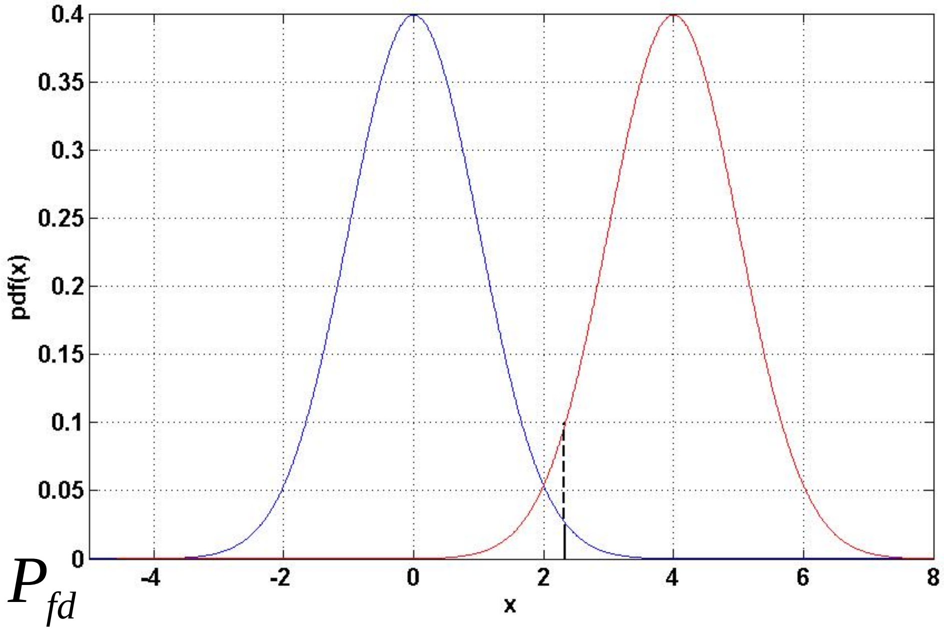
Being the detection, or non detection, claimed on statistical basis, two kinds of error can be done:

Error of type 1: **we claim detection when no signal is present in the data (false alarm)**

Error of type 2: **we claim non detection when a signal is present in the data (false dismissal)**

$$P_{fa} = \int_{\Lambda^*}^{\infty} p(\Lambda(x)\,|\,0)dx$$

$$P_{fd} = \int_{-\infty}^{\Lambda^*} p(\Lambda(x)\,|\,h)dx$$



Detection probability: $P_D = 1 - P_{fd}$

Detection theory - 4

Clearly, by decreasing the f.a.p. the f.d.p. increases (i.e. the detection probability decreases). We need some kind of trade-off.

The Neyman-Pearson lemma says that when we search for a <u>known signal</u> an optimal test exists, such that the detection probability is maximized for a chosen value of the false alarm probability (maximum power test).

This is a likelihood ratio test: $\Lambda(x) = \dfrac{p(x\,|\,h)}{p(x\,|\,0)}$

The threshold Λ^* is computed for a chosen value of P_{fa}

We will see later that this test corresponds to the matched filter, which maximizes the signal-to-noise ratio (SNR) over all linear filters (independently of the probability distribution of the noise).

We will also see that it is not always possible to use the matched filter.

Detection theory - 5

If there is a statistical evidence for the presence of a signal, we can increase its significance by analyzing a longer set of data: if it was really a signal it will be detected with higher and higher SNR (but see the caveat at slides 47-49).

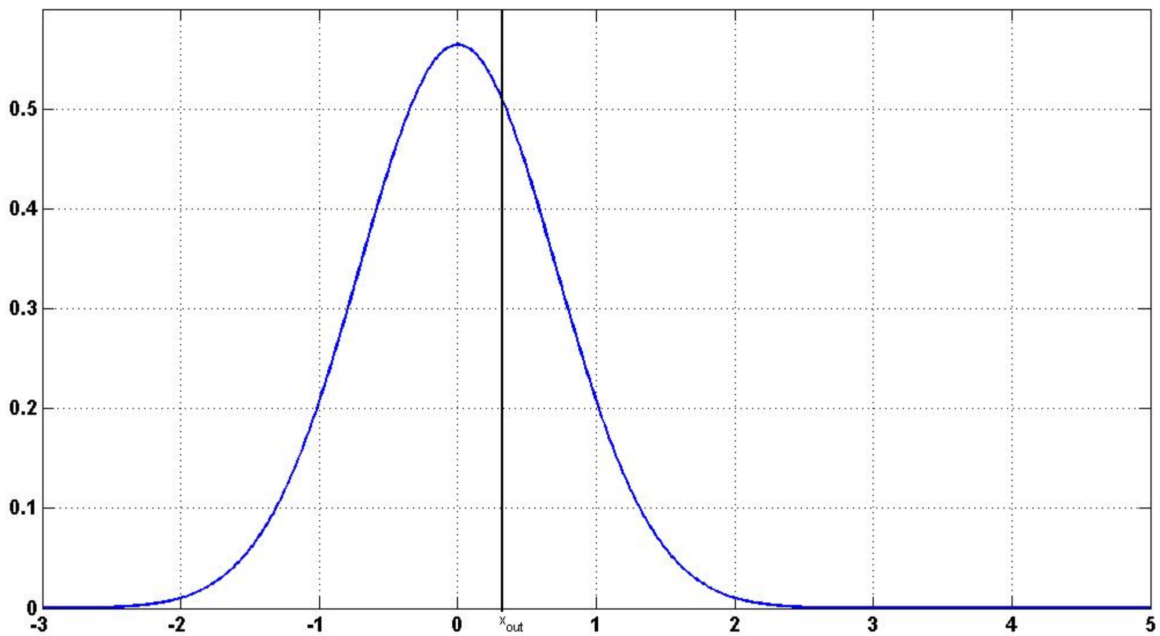
If the output of the analysis is compatible with noise only, we can put an upper limit on the signal strength.

This can be done in different ways. For instance, in a frequentist approach, we can compute the amplitude $h_{0,\min}$ of the signal that, if really present into the data, would produce an analysis output that in say 95% of the cases would be larger than the value actually found.

This requires the knowledge of the filtered data distribution in presence of a signal.

Detection theory - 6

The latter can be computed directly or estimated through the injection in the data of simulated signals.



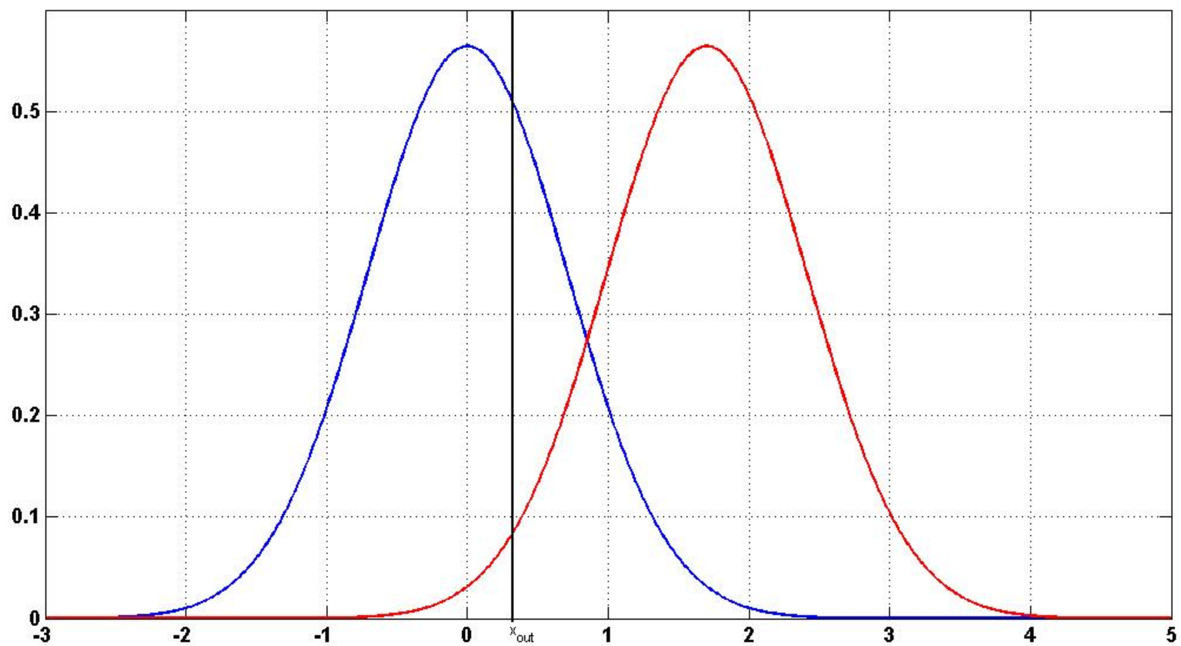
The blue curve is the expected distribution in absence of signals. x_{out} is the value found in the actual analysis.

It corresponds to a FAP of ~40%, then we conclude it is compatible with pure noise.

Detection theory - 7

Then, we search the signal amplitude $h_{0,\min}$ such that if a signal with that amplitude were really present into the data it would produce an output of the analysis larger than x_{out} with a probability of 95%.

In this example $h_{0,\min} \sim 1.8$



Detection theory - 8

Let us assume that the noise is a stationary, gaussian and zero-mean random process.

The logarithm of the likelihood ratio is $\ln(\Lambda) = (x \mid h) - \frac{1}{2}(h \mid h)$

where we have introduced the scalar product

$$(x \mid y) = 4\,\mathrm{Re}\left\{ \int_{-\infty}^{+\infty} \frac{X(f) \cdot Y^*(f)}{S_n(f)}\, df \right\}$$

X, Y : FT of x,y

$S_n(f)$: bi-lateral power spectrum of the noise

For a nearly monochromatic signal, like those in which we are interested, the scalar product can be approximated as

$$(x \mid y) \approx \frac{1}{S_n(f_0)} \int_0^{t_{obs}} x(t) \cdot h(t)\, dt$$

Detection theory - 9

I.e. the likelihood ratio test consist in computing the cross-correlation between the data and the signal and comparing it to a threshold: this is the matched filter (see slides 33-34).

If there are unknown parameters $\vec{\theta}$, a common approach (called Generalized Likelihood Ratio Test) consists in:

a. estimate them by solving the *maximum likelihood equation*

$$\frac{\partial \ln \Lambda(x;\vec{\theta})}{\partial \vec{\theta}} = 0 \Rightarrow \hat{\vec{\theta}}$$

b. Compute the likelihood ratio for $\vec{\theta} = \hat{\vec{\theta}}$ and compare the resulting value to a threshold to establish if a signal is present (with parameters equal to the estimated ones).

For continuous signals this brings to the so-called “F-statistics” invented by Jaranowski, Krolak, Schutz.

Detection theory - 10

The GLRT does not satisfy any optimality criterium. I.e. it is possible, at least in principle, to find a more powerful statistics, that is with a higher detection probability for a fixed FAP.

Indeed, later we will discuss at length a different detection statistics → slides 54 and following

Differences respect to other kind of signals

1. **The signal is always present**: we can confirm or reject a candidate looking at longer and longer data sets -> the false alarm probability can reduced to negligible values

2. The expected signal amplitude is very low, but we can observe it for a long time thus increasing the SNR

3. Coincidences, or joint analysis, with other detectors are desirable but not mandatory

4. Source parameter estimation can be done with extremely high precision

Basic tools - 1

Let us consider a continuous signal buried into noise with bi-lateral power spectrum: $S_n(f)$

Let us start from the simplest model for a continuous signal, a sinusoid, and then add the various complications to make it 'compliant' with a realistic continuous GW signal.

Let us make the following simplifying assumptions:

-The signal adds linearly to the noise: $x(t) = n(t) + h(t)$

- The noise is gaussian;

- The noise spectrum is flat (in the interested band): $S_n(f) = const$

- The noise is stationary;

- The data are contiguous (no 'holes')

Basic tools - 2

The analysis methods can be divided among ‘coherent’, i.e. that take into account the signal phase (e.g. matched filter and correlator) and incoherent, where the signal phase is discarded (e.g. the periodogram, Radon transform, Hough transform).

Typically, incoherent methods are more robust and less computationally demanding but also less sensitive.

What method (or combination of methods) to use depends critically on the information we have on the signal we are searching and on the available computing power.

In this lecture we will focus the attention on coherent methods.

Basic tools - 3

Case 1: the signal is a sinusoid

$$h(t) = h_0 \sin(\omega_0 t + \varphi_0)$$

Let us see how the detection can be done under various hypothesis about what we know of the signal we are searching for.

Basic tools - 4

- ❑ frequency and phase known → **matched filter**

Compute

$$y(t_{obs}) = \int_0^{t_{obs}} x(t) \cdot \sin(\omega_0 t + \varphi_0) dt$$

This is equivalent to pass the data through a filter with impulse response $h(-t)$

The signal output is $y_s(t_{obs}) = h_0 \cdot \frac{t_{obs}}{2}$

The variance of the output noise is $\sigma_n^2 = \frac{S_n(f_0)}{2} \cdot t_{obs}$

The output SNR is

$$SNR = \frac{y_s(t_{obs})}{\sigma_n} = h_0 \sqrt{\frac{t_{obs}}{2S_n(f_0)}}$$

Basic tools - 5

which corresponds to a 'nominal sensitivity' (amplitude of the signal detectable with SNR=1) $h_{SNR=1} = \sqrt{\frac{2S_n(f_0)}{t_{obs}}}$

The output noise has still a Gaussian distribution with $\mu = 0$ and standard deviation σ_n (linear combination of gaussians).

We can compare the filter output with the noise only distribution and establish (statistically) if a signal is present in the data.

E.g. the threshold corresponding to a FAP of 1% is $\vartheta = 2.33 \cdot \sigma_n$

If the output is compatible with noise (for the given threshold) we can set an upper limit.

The distribution of the noise+signal at the output is also a gaussian shifted by the signal amplitude, i.e with mean value $\mu = y_s(t_{obs})$ and standard deviation σ_n.

Basic tools - 6

We can also compute the minimum detectable signal with, say, 1% FAP and 5% FDP, by solving the equation (for Gaussian noise)

$$\vartheta - \mu = -1.64 \cdot \sigma_n$$

from which $h_{0,\min} = 3.97 \cdot \sqrt{\dfrac{2S_n(f_0)}{t_{obs}}}$

This corresponds to SNR=3.97.

The matched filter gives the maximum SNR among all linear filters and this why it is called the 'optimum' filter.

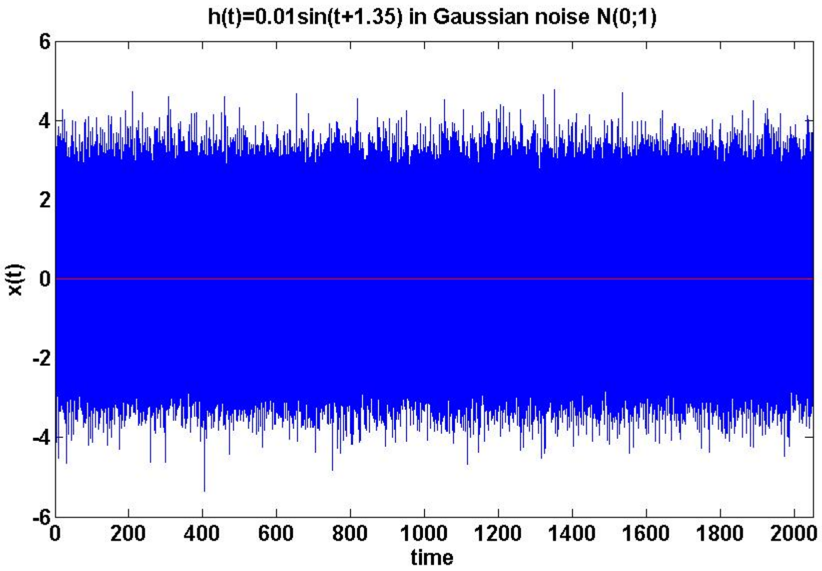
It also maximizes the detection probability for a given value of the FAP (see slide 22)**.**

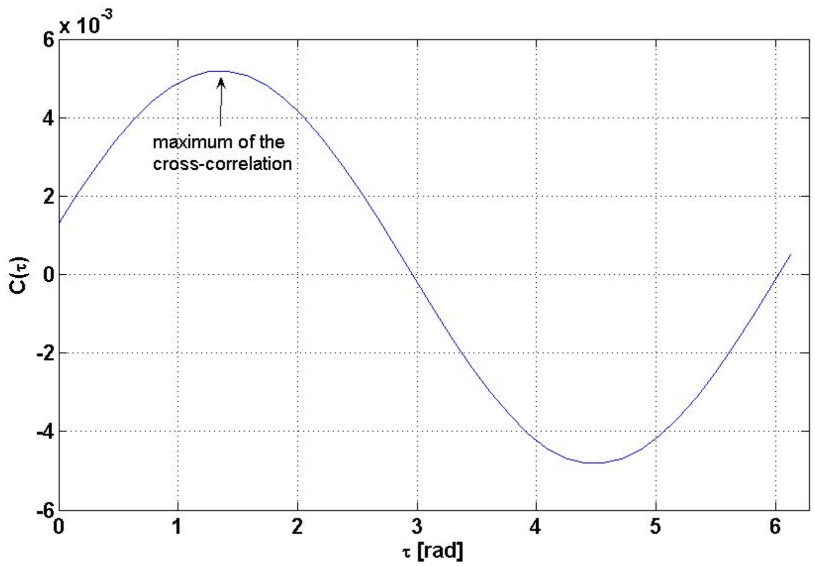
Basic tools - 7

- The frequency is known → **cross-correlation**

Compute
$$C(\tau) = \frac{1}{t_{obs} - \tau} \int_0^{t_{obs}-\tau} x(t) \cdot \sin[\omega_0 (t+\tau)] dt \qquad 0 \le \tau \le \frac{2\pi}{\omega_0}$$

and take the maximum over τ .





Basic tools - 8

In practice, the cross-correlation is computed over a discrete grid of values of τ

The finer is the grid, the better we can estimate the unknown parameter (and, obviously, the higher is the computational time).

If we use a 'mismatched' filter, i.e. use a template slightly different from the signal present into the data, the sensitivity loss can be estimated as

$$L = 1 - \frac{(s \mid s')}{\sqrt{(s \mid s) \cdot (s' \mid s')}} \approx 1 - \cos(\omega_0 \Delta\tau)$$

where $\Delta\tau$ is the phase difference between the signal and the template.

The cross-correlation method can be generalized, in principle, to a signal with a whatever number of unknown parameters but can be computationally very heavy.

Basic tools - 9

- ❑ Unknown frequency → **power spectrum**

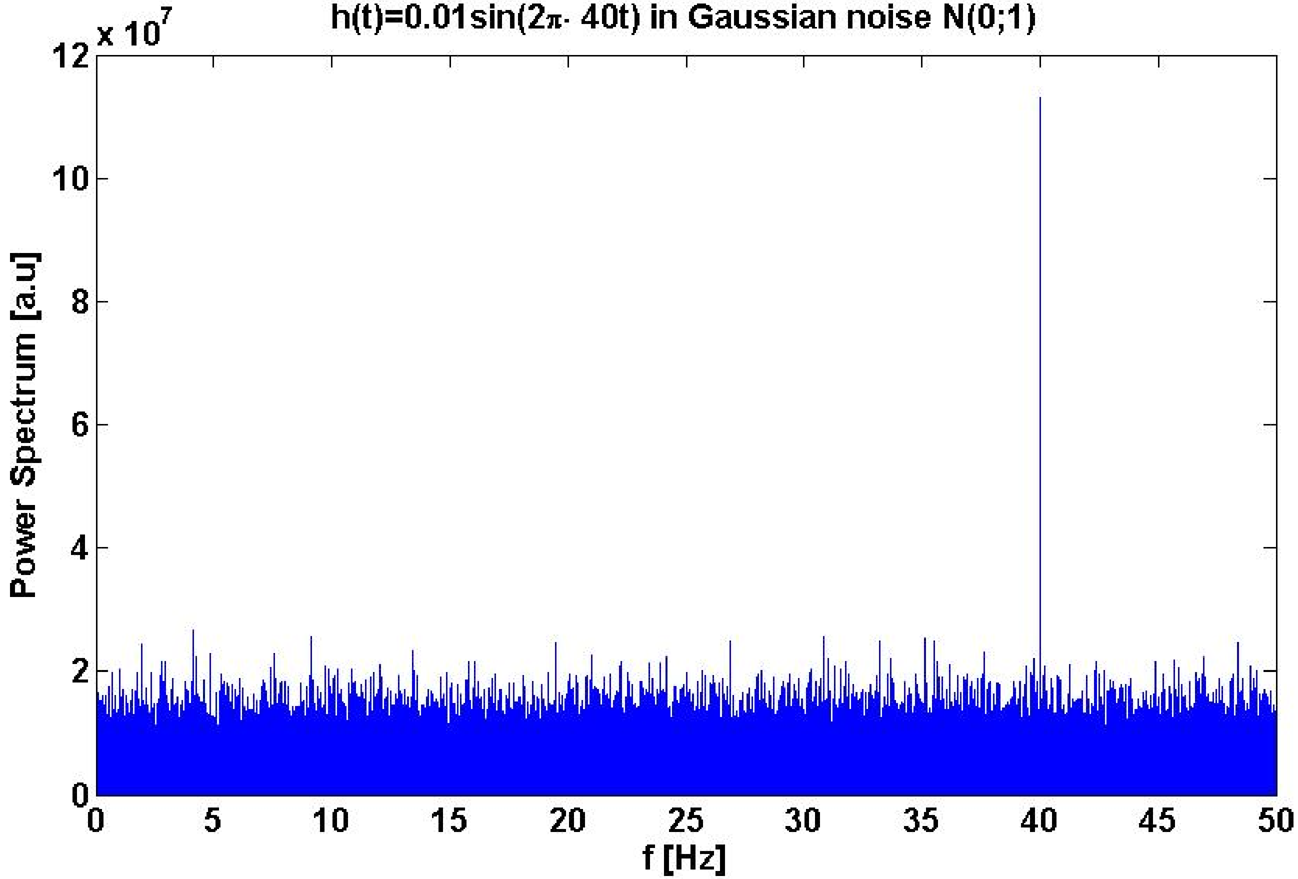
The power spectrum is the Fourier Transform of the autocorrelation of the data.

It can be estimated in several ways. A useful method is the **periodogram**, i.e. the square modulus of the Fourier Transform of the data:

$$X_k = \left| \sum_{i=0}^{N-1} x_i e^{j\frac{k \cdot i}{N}} \right|^2$$

The periodogram can be efficiently computed by the FFT algorithm.

Basic tools - 10



The spectral resolution is $\delta f = 1/t_{obs}$

The power spectrum does not allow to recover the signal phase.

Basic tools - 11

The signal power is $\frac{h_0^2}{2}$

The noise contribution is $2S_n(f_0)\cdot \delta f$

$$\longrightarrow \quad SNR = h_0\sqrt{\frac{t_{obs}}{4S_n(f_0)}}$$

The noise probability distribution for each frequency bin is exponential with variance $2\sigma_n^2$

The probability to have in a given frequency bin a power bigger than a threshold P_{thr} is $P_0 = e^{-P_{thr}/2\sigma_n^2}$

The probability to have a power bigger than P_{thr} in any of N frequency bin is $P_1 = \lambda e^{-\lambda};\ \lambda = N\cdot P_0$ from which the value of the threshold P_{thr} corresponding to, e.g., $P_1 = 0.01$ can be computed.

Basic tools - 12

In presence of a signal of power P_s, the power spectrum distribution is given by $p(P;P_s) = e^{-(P+P_s)} \cdot I_0(2\sqrt{P \cdot P_s})$

Here we assume that the noise average power has been normalized to unity.

$I_0(x)$ is the modified Bessel function of the first kind of order 0.

E.g., assume we have $t_{obs} = 1\,\mathrm{yr}$ and to search for a monochromatic signal over a band of 1000Hz, being the noise spectrum flat ($S_n(f) = const$) for simplicity: $N \approx 3.15 \cdot 10^{10}$

The threshold for a FAP=1% is found to be $P_{thr} = 5.36 \cdot \left(2\sigma_n^2\right)$.

Basic tools - 13

The ‘naked’ periodogram has some problems:

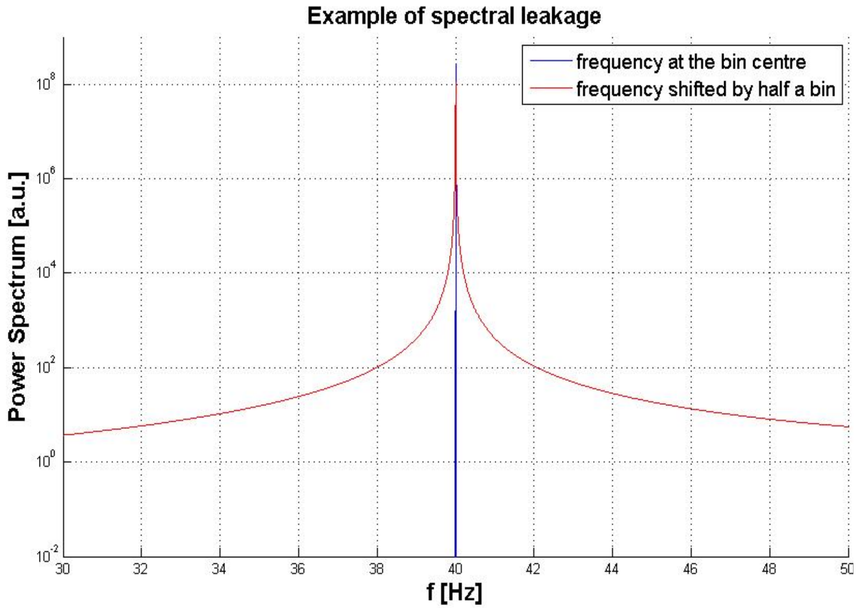
a. If the frequency of a monochromatic signal is not exactly at the centre of a frequency bin, there will be a *leakage* in the adjacent bins

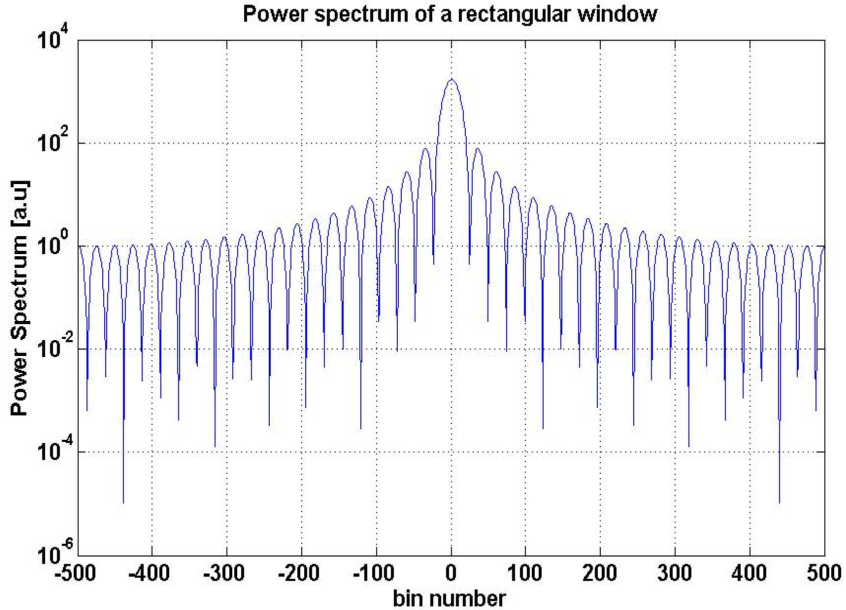
b. Its variance does not decrease with the length of data

c. Often, we cannot do FFT of arbitrarily long pieces of data because of the limited accuracy with which source parameters are known (see later)

Basic tools - 14

An example of a.:





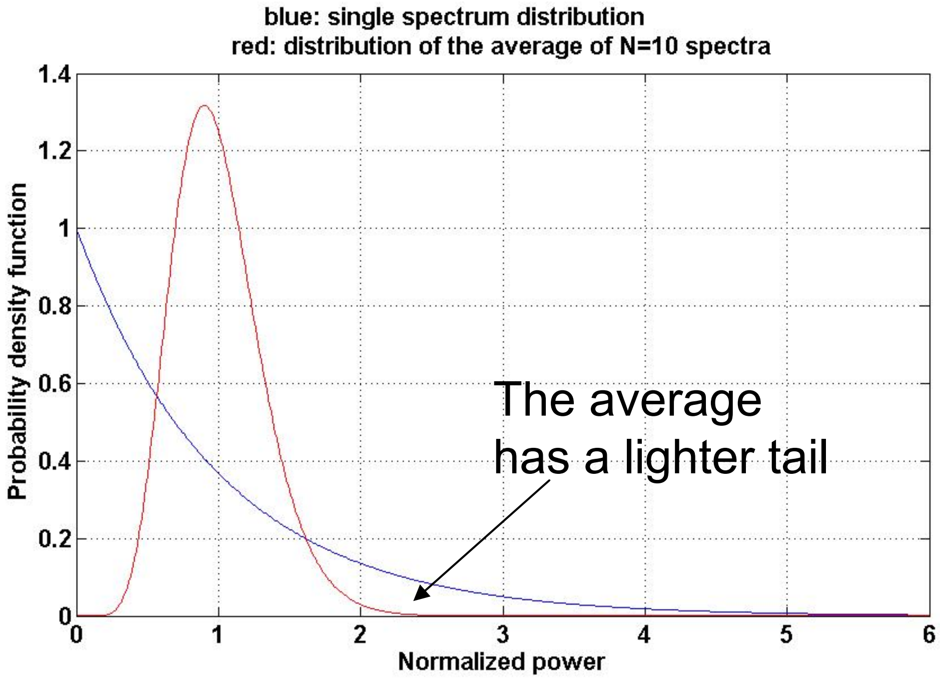
The leakage is an effect of the finite size of the piece of data from which the periodogram is built (it is as the data have been multiplied by a rectangular window).

It can be cured by properly *windowing* the data, i.e. multiplying the data by a smooth function that behaves better in frequency.

Basic tools - 15

Points b. and c. can be cured by dividing the observation time in M intervals, computing the periodogram of each piece and making the average.

The SNR decreases by a factor $\sqrt[4]{M}$ but the distribution becomes a χ^2 with $2M$ degrees of freedom.



Basic tools - 16

Case 2: Doppler effect

Let us consider a sinusoid with a frequency modulated by the Doppler effect.

The Doppler effect can be removed multiplying the data by

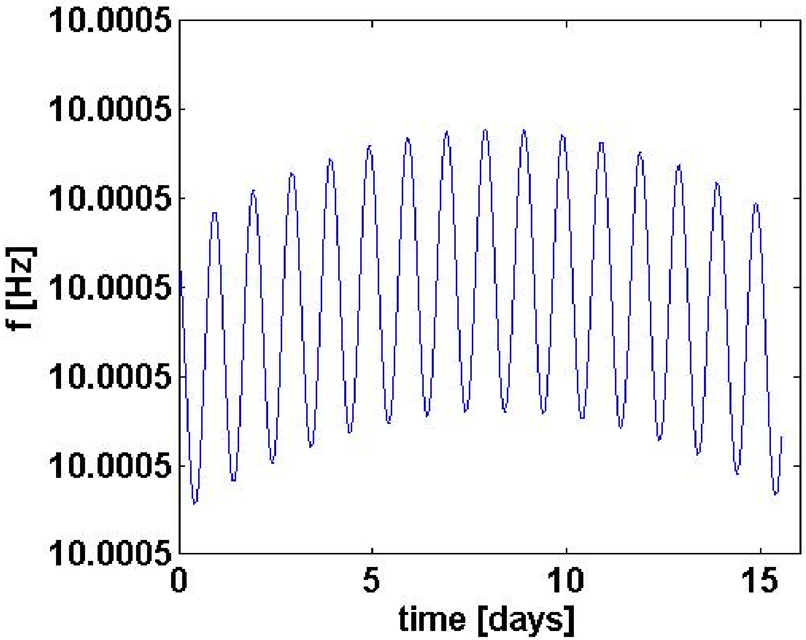
$$e^{-j\Delta\phi_{Doppler}(t)}$$

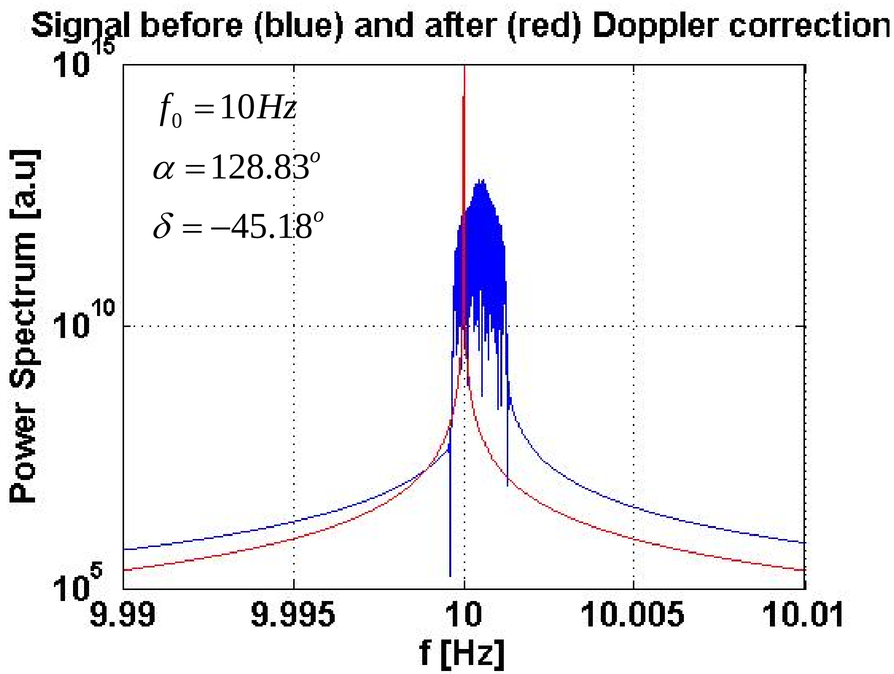
This requires:

1. a very precise knowledge of the detector position in the SSB (the vector $\vec{r}$ in slide 8) as a function of time. This information is produced by various public softwares, like NOVAS.

2. A very precise knowledge of the source position (the vector $\hat{n}$ in slide 9) and frequency: this condition is typically met for pulsars observed in the EM (if T_{obs} is not too large, see slide 71).

Basic tools - 17





Basic tools - 18

Case 3: Spin-down

Let us consider a sinusoidal signal with a steadily decreasing frequency.

By expanding the frequency in a Taylor series around the initial value, the signal phase can be written as

$$\phi(t) = \phi_0 + 2\pi \cdot \sum_n \frac{f_{(n)}}{(n+1)!}(t - t_0)^{n+1}$$

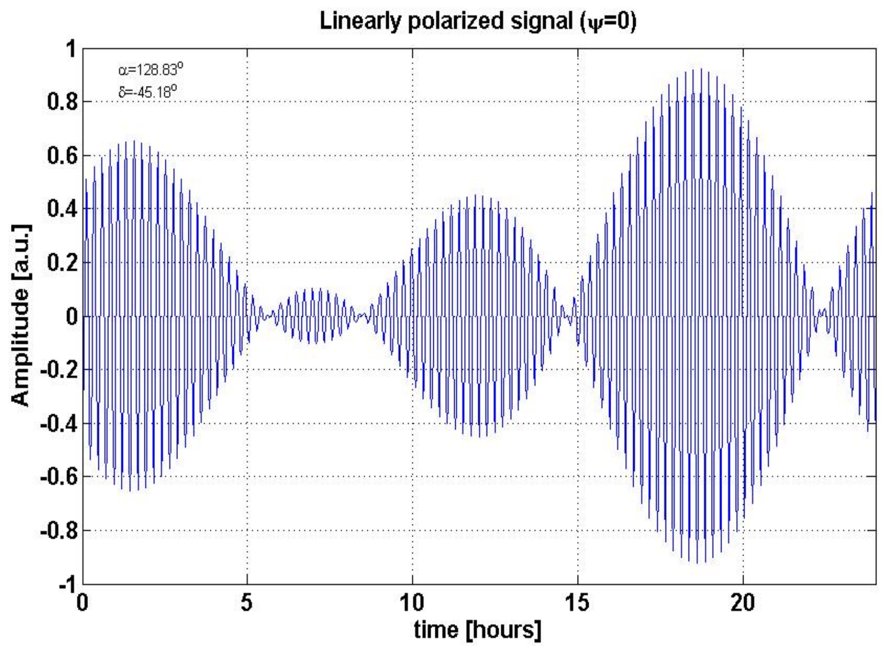
$f_{(n)} = \frac{d^n f_0}{dt^n}$: spin-down parameter of order n

Like in the previous case, if the spin-down parameters of a source are known with high accuracy we can correct by multiplying the data by $e^{-j\Delta\phi_{sd}(t)}$

Basic tools - 19

Case 4: Amplitude modulation

Let us consider a sinusoid modulated in amplitude, as described in slides 13-16: $h(t) = A(t) \cdot e^{j\omega_0 t}$



The signal shape depends on the detector location and source parameters.

Basic tools - 20

We can try to detect this signal using different methods.
E.g. we could apply a bank of templates for the unknown source parameters but this would be extremely heavy (there are 3 parameters which should be taken into account → see slides 36-37).

A better choice would be to use the F-statistics.

A good alternative is to apply the matched filter on the Fourier transform of the signal: in this case each template consists of just five complex numbers (the complex amplitudes of the five spectral lines).

This is the method we will use in the exercise and will be described soon.

Putting all this stuff together…

Let us try to see how all the stuff we have seen up to now can be used in the search of GW signals.

We make a distinction between two cases:

1. The source parameters $(\alpha, \delta, f_0, \dot{f}_0, ...)$ are known. We call this <u>*targeted search.*</u> The nuisance parameters (ψ, ι, ϕ_0) are generally unknown. E.g. the search for GW signals from pulsars observed in the EM belongs to this category.
2. The source parameters are unknown. We call this <u>*blind search.*</u>

Targeted search -1

The received GW signal is not monochromatic and covers a small frequency band around the known emission frequency:

$$\Delta f_{Doppler,\max} \approx 10^{-4} f_0$$

$$\Delta f_{SD} \approx \dot{f}_0 \cdot t_{obs} \qquad \Rightarrow \quad \Delta f_{tot} \quad \text{is a fraction of Hz}$$

$$\Delta f_{AM} \approx 2 \cdot 10^{-5} Hz$$

We can extract the small band of interest and work just on it.

This is the starting point of your exercises.

The calibrated data produced by an ITF are sampled, e.g., at 4096Hz.

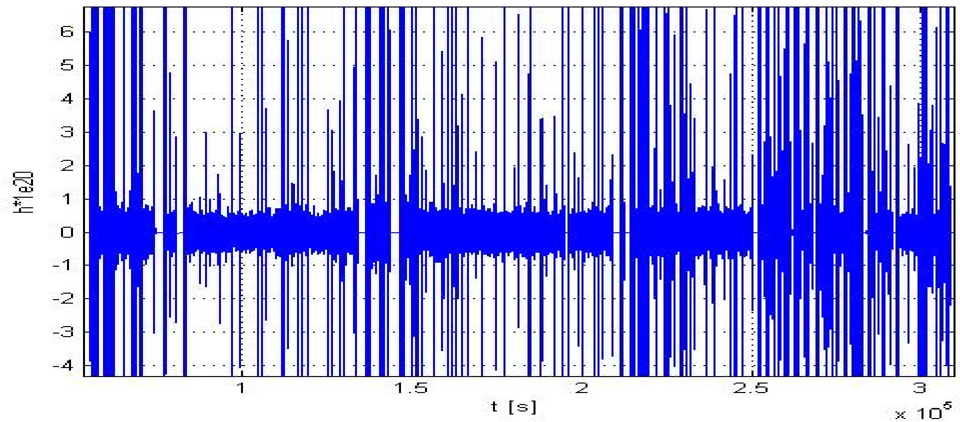
We can simplify our life by down-sampling the data at a much lower rate. This is possible because $\Delta f_{tot} \ll f_{Nyquist} \approx 2\ kHz$

Targeted search -2

Then, **we correct for Doppler and spin-down effects** (as a matter of fact the procedure actually used, based on the construction of the so-called *analytical signal*, allows to down-sample and correct the data at the same time→ see slides 77-81 for more details).

At this point we have just noise plus (if we are lucky) a monochromatic signal with a modulated amplitude .

We remove residual periods which are particularly noisy.



Targeted search -3

We also reduce the effects of noise non-stationarity by weighting the data with the inverse of local variance (Wiener filtering).

Finally, we apply the last filtering stage which takes into account amplitude modulation and estimate the parameters of the GW signal.

The antenna response (slide 13) can be written as

$$h(t) = h_0 \cdot \boldsymbol{A} \cdot \boldsymbol{W} \cdot \exp\left(j \cdot \left(\omega_0 t + \gamma \right) \right)$$

where we have introduced the “generator” 5-vector

$$W_k = e^{jk\Omega t} \quad \text{with} \quad -2 \leq k \leq 2$$

and the signal 5-vector (which completely defines the signal in the antenna)

$$\boldsymbol{A} = H_+ \boldsymbol{A}^+ + H_\times \boldsymbol{A}^\times$$

Targeted search -4

The two polarization signals 5-vectors **A⁺** and **Aˣ** have components:

$$A^{+}_{-2} = \frac{a_{2c}}{2} - j\,\frac{a_{2s}}{2} \qquad A^{\times}_{-2} = \frac{b_{2c}}{2} - j\,\frac{b_{2s}}{2}$$

$$A^{+}_{-1} = \frac{a_{1c}}{2} - j\,\frac{a_{1s}}{2} \qquad A^{\times}_{-1} = \frac{b_{1c}}{2} - j\,\frac{b_{1s}}{2}$$

$$A^{+}_{0} = a_0 \qquad A^{\times}_{0} = 0$$

$$A^{+}_{1} = \frac{a_{1c}}{2} + j\,\frac{a_{1s}}{2} \qquad A^{\times}_{1} = \frac{b_{1c}}{2} + j\,\frac{b_{1s}}{2}$$

$$A^{+}_{2} = \frac{a_{2c}}{2} + j\,\frac{a_{2s}}{2} \qquad A^{\times}_{2} = \frac{b_{2c}}{2} + j\,\frac{b_{2s}}{2}$$

Targeted search - 5

Given the data $x(t)$ the corresponding 5-vector is given by its Fourier components at the five frequencies.

The two signal 5-vectors can, in principle, be computed using eqs. in slide 54. In practice, however, we must build it using the same procedure applied to the data, i.e. by imposing the same cuts, vetoes, Wiener weights.

Once the 5-vectors of the data and of the two signal components have been computed we use them in the matched filtering procedure.

Targeted search - 6

The detection is done by applying a matched filter to each of the two signal components:

$$\hat{h}_+ = \frac{\boldsymbol{X} \cdot \boldsymbol{A}^+}{\left|\boldsymbol{A}^+\right|^2} \qquad \hat{h}_\times = \frac{\boldsymbol{X} \cdot \boldsymbol{A}^\times}{\left|\boldsymbol{A}^\times\right|^2}$$

$\hat{h}_+$ and $\hat{h}_\times$ are, respectively, the estimation of H_+ and $H_\times$ (we use the orthogonality relation $A^+ \cdot A^\times = 0$).

From here we build a detection statistic

$$S = c_+ \cdot \left|\hat{h}_+\right|^2 + c_\times \cdot \left|\hat{h}_\times\right|^2$$

and optimize over the two coefficients.

The "best ROC" statistics is obtained with the choice

$$c_+ = \left|\boldsymbol{A}^+\right|^4 \qquad c_\times = \left|\boldsymbol{A}^\times\right|^4$$

Targeted search - 7

The square modulus of the two basic observables $|\hat{h}_+|^2, \; |\hat{h}_\times|^2$ in case of noise only has an exponential distribution:

$$f(x) = \frac{\left|\boldsymbol{A}^{+/\times}\right|^2}{\sigma_X^2} e^{-\frac{\left|\boldsymbol{A}^{+/\times}\right|^2}{\sigma_X^2}\cdot x}$$

σ_X^2 : variance of the data 5-vector

In case a signal of amplitude h_0 is present in the data, the distribution becomes

$$f(x;h_0) = \frac{k}{2} \cdot e^{-k(x+h_0^2)/2} \cdot I_0\left(\sqrt{k^2 \cdot h_0^2 \cdot x}\right)$$

$$k = 2 \cdot \frac{\left|A^{+/\times}\right|^2}{\sigma_X^2}$$

Modified Bessel function of the 1st kind, order 0

Targeted search - 8

Then, we can derive the distribution functions for our detection statistics (slide 57)

$$S = \left|A^{+}\right|^{4} \cdot \left|\hat{h}_{+}\right|^{2} + \left|A^{\times}\right|^{4} \cdot \left|\hat{h}_{\times}\right|^{2}$$

1. Check if the actual value of S obtained in the analysis is compatible with the noise only distribution;
2. If not claim detection, otherwise set an upper limit using the distribution in presence of a signal (theoretical or built through software injections). → see discussion slides 24-26

Targeted search - 8

If a signal is detected, we can use the two basic observables to estimate signal parameters.

Estimation of the amplitude $\hat{h}_0 = \sqrt{\left|\hat{h}_+\right|^2 + \left|\hat{h}_\times\right|^2}$

Invariants $\hat{h}_+ \cdot \hat{h}_\times' = A + jB$ $\quad \left|\hat{h}_+\right|^2 - \left|\hat{h}_\times\right|^2 = C$

Estimation of η $\eta = \frac{-1+\sqrt{1-4B^2}}{2B}$

Estimation of ψ

$$\cos(4\psi) = \frac{C}{\sqrt{(2A)^2 + B^2}}$$

$$\sin(4\psi) = \frac{2A}{\sqrt{(2A)^2 + B^2}}$$

(these are independent on the absolute phase γ)

Estimation of the absolute phase

$$e^{j\gamma} = \frac{\hat{h}_{+/\times}^{(sperim.)}}{\hat{h}_{+/\times}^{(teor.\ \gamma=0)}}$$

Blind search - 1

The analysis methods we have seen so far are particularly suitable when (most of the) source parameters are known.

Let us see what happens if we want to use them in a blind search.

A plausible minimal range for the source parameters is the following:

$$(\alpha, \delta) \text{ over the whole sky}$$

$$f_0 \in [20, 2000]\, Hz$$

$$\tau = \frac{f_0}{|\dot{f}_0|} > 10{,}000 \text{ yr}$$

While we do not expect CW signals with $f_0 > 2kHz$, we would like to search for signals with τ as small as possible:

Blind search - 2

Smaller τ means higher $\dot{f}_0$. On its turn, this possibly means a younger source, then likely more deformed (higher ellipticity, see slide 4) → higher GW emission.

In order to apply whatever filtering method we need to build a grid in the parameter space.
The dimensionality of the parameter space is 3+s where s is the number of spin-down parameters (we are neglecting the nuisance parameters, which number may depend on the particular analysis method used).

Blind search - 3

Let us indicate with Δt the sampling time.

We can write the following relations:

$$N_f = \frac{t_{obs}}{2 \cdot \Delta t}$$

number of frequency bins

$$N_{DB}(f) = 10^{-4} \cdot f \cdot t_{obs}$$

number of frequency bins in the Doppler band of the frequency f

$$N_{sky}(f) = 4\pi \cdot N_{DB}^2$$

number of points in the sky for the frequency f

$$N_{sky,tot} = 4\pi \cdot 10^{-8} \sum_{i=1}^{N_f} i^2 \approx \frac{4\pi}{3} 10^{-8} \cdot N_f^3$$

total number of points in the sky (all frequencies)

$$N_{SD}^{(j)} = 2N_f \left(\frac{t_{obs}}{\tau_{\min}} \right)^j$$

number of values of spin-down parameter of order j

$$N_{tot} = N_{sky,tot} \cdot \prod_{j:N_{SD}^{(j)} \geq 1} N_{SD}^{(j)}$$

total number of points in the parameter space

Blind search - 4

With observation times O(months) we need to take into account also the second order spin-down parameter and we can write:

$$N_{tot} \approx \frac{10^{-8}\pi}{6} \frac{t_{obs}^8}{\Delta t^5 \cdot \tau_{\min}^3} = 2.28 \cdot 10^{31} \left(\frac{t_{obs}}{4\,\text{months}} \right)^8 \left(\frac{\Delta t}{2.5 \cdot 10^{-4}\, s} \right)^{-5} \left(\frac{\tau_{\min}}{10^4\ \text{years}} \right)^{-3}$$

Let us assume we want to make the analysis by searching for significant peaks in the periodogram. This implies to compute an FFT of length $N_{FFT} \approx 2 \cdot 10^{10}$ (for 4 months of data) for each source position and spin-down value.

The computation of an FFT requires $\approx 5 \cdot N_{FFT} \cdot \log_2(N_{FFT})$ FLOPS.

Then, to make the analysis we would require a computing power

$$CP \approx 3.7 \cdot 10^{21}\ TFLOPS$$

A different approach is necessary!!

Blind search - 5

We would like that an alternative approach satisfies two requirements:

- drastically reduce the computing power needed;

- not loose too much in sensitivity

Both can be satisfied in the so-called **hierarchical search.**

The key idea is that of alternating incoherent and coherent steps.

In the incoherent step a rough exploration of the parameter space is done and some candidates are selected.

In the coherent step each candidate is followed with a more refined search.

Hierarchical searches of different ‘flavours’ exist but we will not discuss them here.

Practical issues with real data - 1

Non-linearity

A. Non linear interaction between signal ans noise
   This is not a problem due to the smallness of expected signal: the linear approximation will always work well

B. Non-linear dynamics of the noise
   The optimum filtering theorem says that the optimum filter is linear in the case of additive gaussian noise. Then, an enhancement by using a non-linear filter is expected only if the noise is not gaussian

Practical issues with real data - 2

Non-gaussianity

A. Statical non-linearities, like saturations, distortions introduced by amplifiers,...

B. Non-linear dymanic of the noise

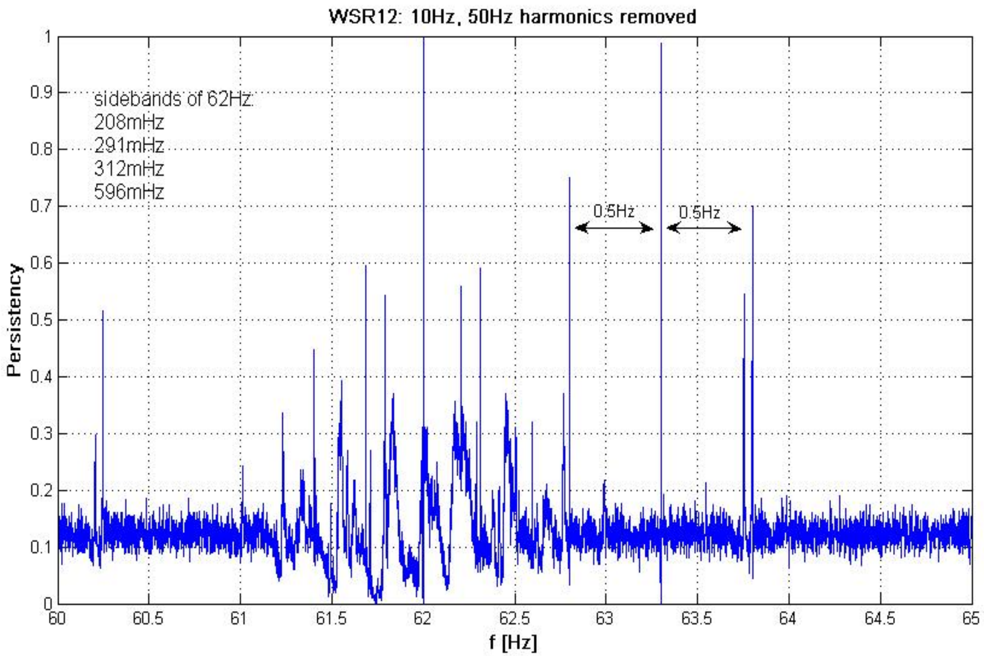
This should be corrected at the level of h-reconstruction. However, a not heavy non-gaussianity is not a big problem for the analysis: according to the central limit theorem the ouput of a linear filter is more gaussian than the input

Practical issues with real data - 3

Non-flatness of the noise spectrum

This can be due to the presence of spectral peaks produced by the instrument.

They must be identified (we must be sure they are not due to a GW!) and removed from the data

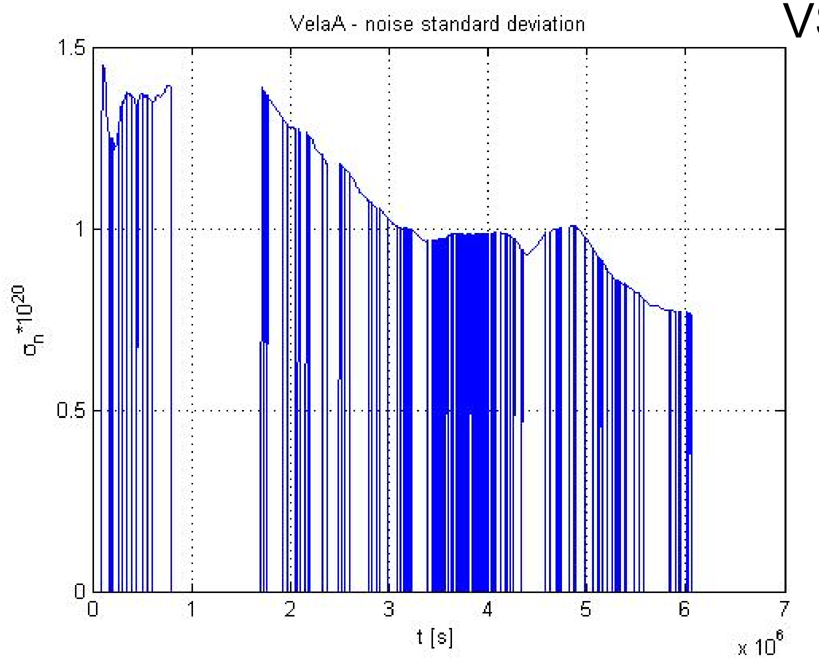


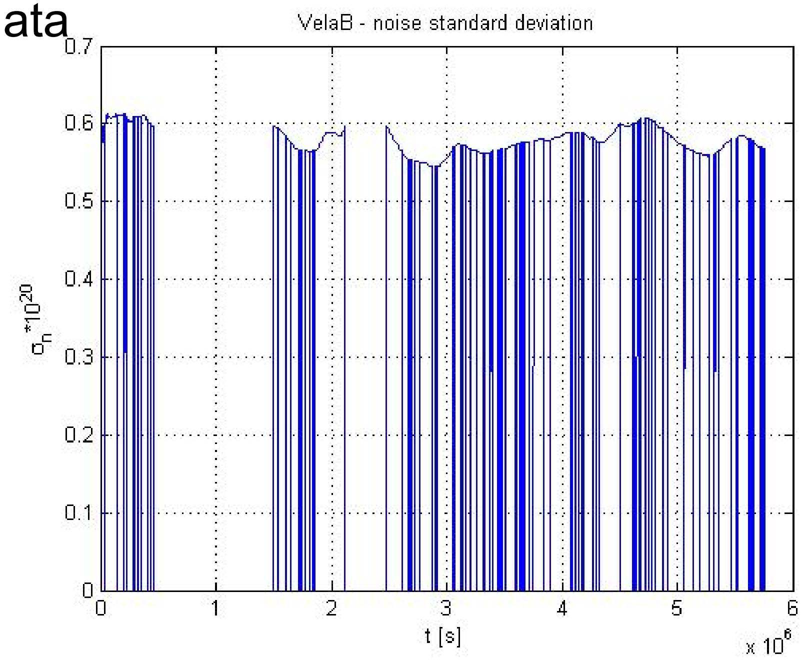
Practical issues with real data - 4

Non-stationarity

A. Slow variation of the noise statistics
This means that the detector sensitivity changes with time and we should use methods to take this into account

VSR1 data





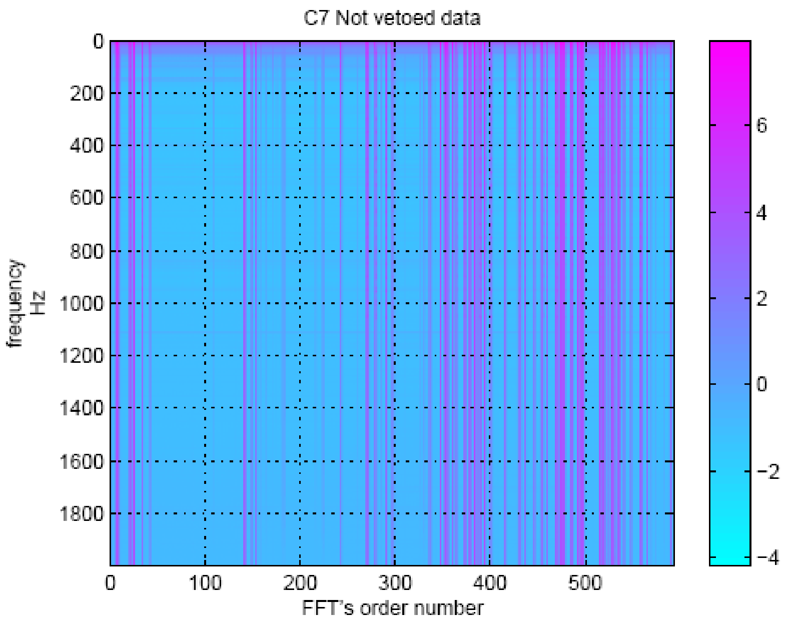
E.g. we can apply a filter to weight less more disturbed periods

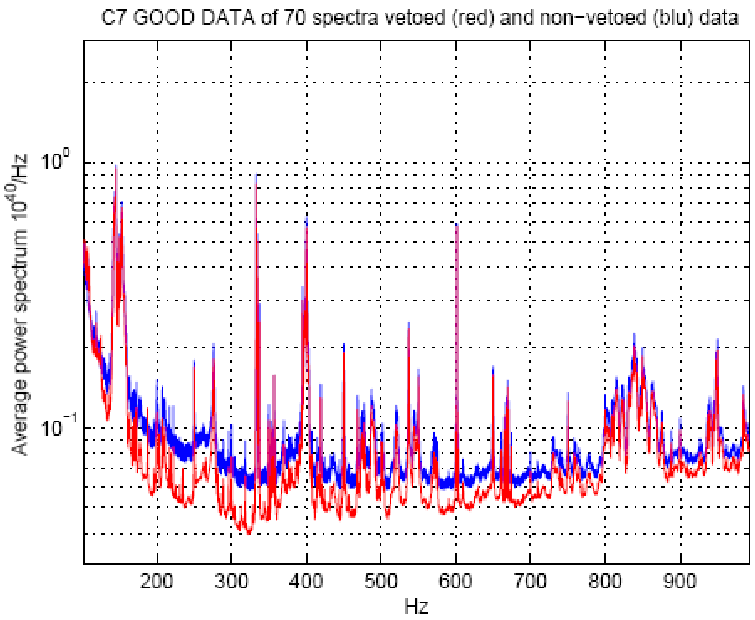
Practical issues with real data - 5

B. Burst disturbances

These are typically short duration spikes that can affect the noise level.

They must be identified and removed.





Practical issues with real data - 6

Holes in the data

Holes are produced, e.g. when the detector is out of lock. The presence of holes has two negative effects.

A. It reduces the energy of the signal that goes into the detector: in the formulae for the SNR we must use the 'effective' observation times.

B. Power leakage: each contiguous set of data is multiplied by a rectangular window.

Practical issues with real data - 7

Targeted searches rely on an accurate knowledge of the source parameters: position, frequency, spin-down.

Even for sources observed in the EM, however, these parameters are known with finite accuracy and this can lead to a loss of sensitivity, especially for very long observation times.

An uncertainty on the frequency and spin-down, $\Delta f_0, \Delta\dot{f}_0$ produce a phase error

$$\Delta\psi = \Delta f_0 \cdot t_{obs} + \frac{1}{2}\Delta\dot{f}_0 \cdot t_{obs}^2 + \ldots$$

Other possible causes of errors are the uncertainty on the source position and the source intrinsic velocity.

Practical issues with real data - 8

Also the presence of NS glitches and timing noise (slides 13-14) can affect the analysis and must be properly taken into account.

A glitch will likely produce a discontinuity in the GW signal phase, other than a jump in frequency and spin-down values.
Astronomical observation can provide information on the occurrence of glitches in known pulsars.

Timing noise, if it affects also the GW signal, will produce a gradual shift of the signal phase respect to the model.
It can be taken into account by building the correction factor $e^{-j\phi(t)}$ using frequently updated ephemeris, which can be provided by astronomical observations.

Interaction with astronomers is very important!

Virgo groups working on DA for CW

Nikhef: Henk Jan Bulten henkjan@nikhef.nl

Pisa: Stefano Braccini stefano.braccini@pi.infn.it

POLGRAW: Andrzej Krolak krolak@impan.pl

Roma: Sergio Frasca sergio.frasca@roma1.infn.it

Bibliography

A. Papoulis, “Probability, random variables and stochastic processes”, McGraw-Hill

M. Maggiore, “Gravitational Waves”, vol. 1, Oxford

R. N. McDonough & A. D. Whalen, “Detection of signals in noise”, 2nd edition, Academic Press

PSS User Guide http:grwavsf.roma1.infn.it/pss/docs

P. Jaranowski, A. Krolak, Living Review of Relativity 8, (2005)

Van Trees (Wiley) “Detection, Estimation and Modulation Theory”, vol 1

E. Groth, Astrophysical Journal Supplement Series, 286, 29:285-302

S. Frasca “Analisi dei segnali” (in italian): http:grwavsf.roma1.infn.it/sp/sp.pdf

Exercises for the hands-on session

• See file vesf_da_exercises_CW.pdf under /data/vesf_school/users/CW

• Two sets of exercises: 1. and 2. allow you to run analysis programs on C7 and simulated data. Exercises 3.-6. are problems to be solved on paper.

• The data files you need for Exercises 1. and 2. are under /data/vesf_school/users/CW

• Each folder contains 4 files:

- sim6_p1s.sbl: contains the data in a .25Hz band around the frequency of interest;
- s0_p1s.sbl: the signal + component (same band)
- s45_p1s.sbl: the signal x component (same band)
- pulsar_1s.m: Matlab file with the pulsar parameters

- Folders p1s-p5s refers to C7 data (each with a pulsar injected)
- Folder p6s has been produced from simulated data (with a pulsar injected)
- Folder psr_j1 is from C7 but with no pulsar injected: the data have been prepared to allow the search GW signals from a real pulsar (Exercise 2). The program filter_dist_mc (under) can be used to setup an UL.

Backup slides

The procedure to extract sub-bands-I

• Take an FFT, from the SFDB. Let $\Delta\nu$ be the frequency resolution, **2N** the number of samples and **B** the bandwidth of the detector;
• take the data from $\mathbf{n_0}$ bins in the frequency band of the actual search; $\mathbf{n_0 = N(\Delta\nu/B)}$;
• build a complex vector that has the following structure:

✓ the first datum equal to zero
✓ the next $\mathbf{n_0}$ data equal to those from the selected bins of the FFT
✓ zeroes from bins $n_0 + 1$ up to the nearest subsequent bin numbered with a power of two. Let us say that this way we have n bins.
✓ zeroes in the next n bins

So, we end up with a vector that is 2n long.

The procedure to extract sub-bands -II

• Take the inverse FFT of the vector.
This is a complex time series that is the “analytical signal” representation of the signal in the band .
It is shifted towards lower frequencies and it is sampled at a sampling rate lower by a factor N/n compared to the original time data.
• Repeat the steps outlined above for all the M FFTs. If they all come from contiguous time stretches simply append one after the other in chronological order. If they are not all contiguous set to zero those stretches where data are missing.

This is a the standard procedure of low-pass filtering for a process.
The analytic signal is zero on the left frequency plane, avoiding aliasing effects in the low-pass operation. The time of the first sample is exactly the same as the first datum used for the data base and the total duration is also that of the original time stretch. There are fewer data because the sampling time is longer.

The procedure to extract sub-bands-III

- The previous are the general features of the procedure
- There are important practical issues of the algorithm implementation, which take into account various aspects, such as the computing time.
- The choice of 1 Hz for the total selected band, having chosen a smaller band for the signal to look for, is reasonable and fast enough for our purposes.

Doppler removal -I

To correct for the Doppler effect a from a given source*, we* multiply each sample of the time sequence by $\exp(-j\,\varphi(t_i))$

(t_i) are the times of the samples

$$\varphi(t_i) = \int \Delta\omega(t)\,dt$$

$\Delta\omega(t)$ is the Doppler correction, in angular frequency, at the time of the i-th sample;
$\Delta\omega(t) = \omega(t) - \omega_0$
is the difference of the observed frequency and the source frequency.

The frequency correction is performed on the sub-sampled data set, thus it is not computing relevant

Doppler removal -II

•Another method is based on a resampling technique
•During its orbit the Earth approaches to or recedes from the source.
•When the distance of the Earth to the source changes by $\pm C \cdot$(1 sample) we add or remove 1 sample to the time sequence

Both the methods can be generalized to include also the spin-down and the Einstein effect

Dither effect -1

Whatever GW signal is present at the input of the interferometer, it is converted to a discrete signal at the output.

Typical ADC have 16 bits and a dynamic of 20V.

The quantization step size, $\frac{20}{2^{nbit-1}}$, corresponds to $h \sim 10^{-21}$

Then, the expected amplitude of typical CW signals is 100 times or more less than the minimum amplitude the ADC can detect.

This means that if we had only signal, the output of the itf would be zero

How can we hope to detect CW signals?

Thanks to the dither effect due to the instrument noise!

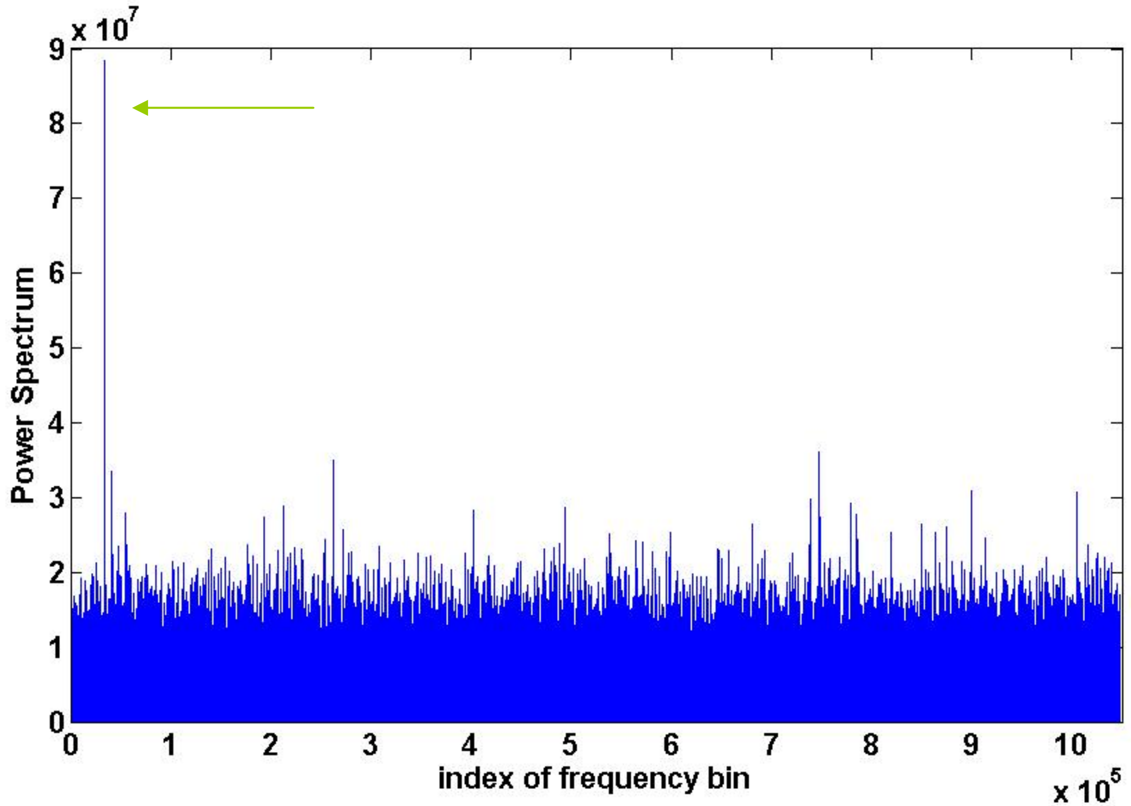
Dither effect -2

Try the following Matlab routine:

```
>> N=2^21;
>> x=(1:N)*0.1;
>> y=0.01*sin(x);             %create a sinusoid with amplitude 0.01
>> n=randn(1,N);              %create gaussian noise
>> yy=round(y+n);             %discretize the sum assuming quantum step =1
>> sp=abs(ffy(yy)).^2;        %compute the power spectrum
>> figure;plot(sp(1:N/2));
```

If n=0 (no noise): yy=0!

Dither effect -3

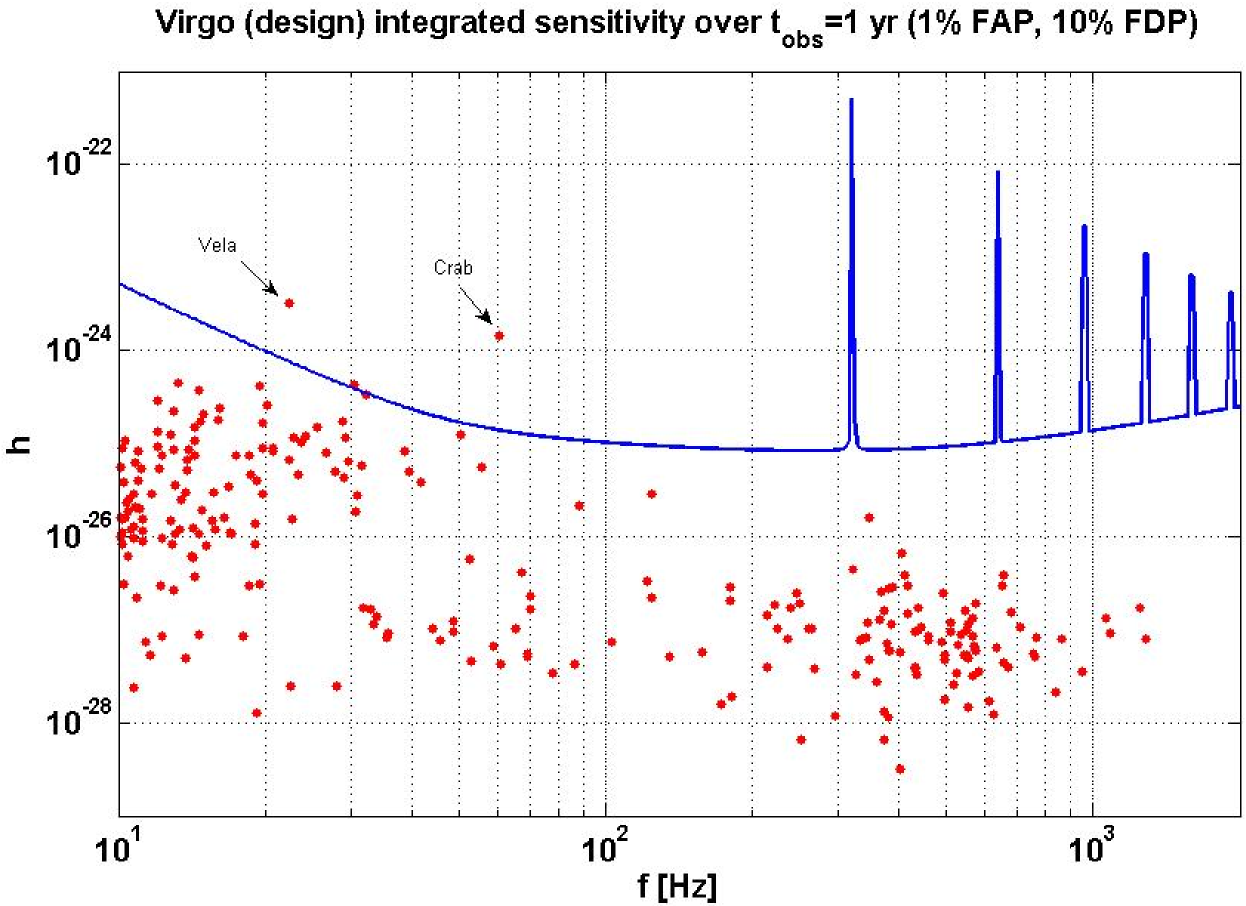


In some cases noise is our friend!

The problems arise because the noise is too much!

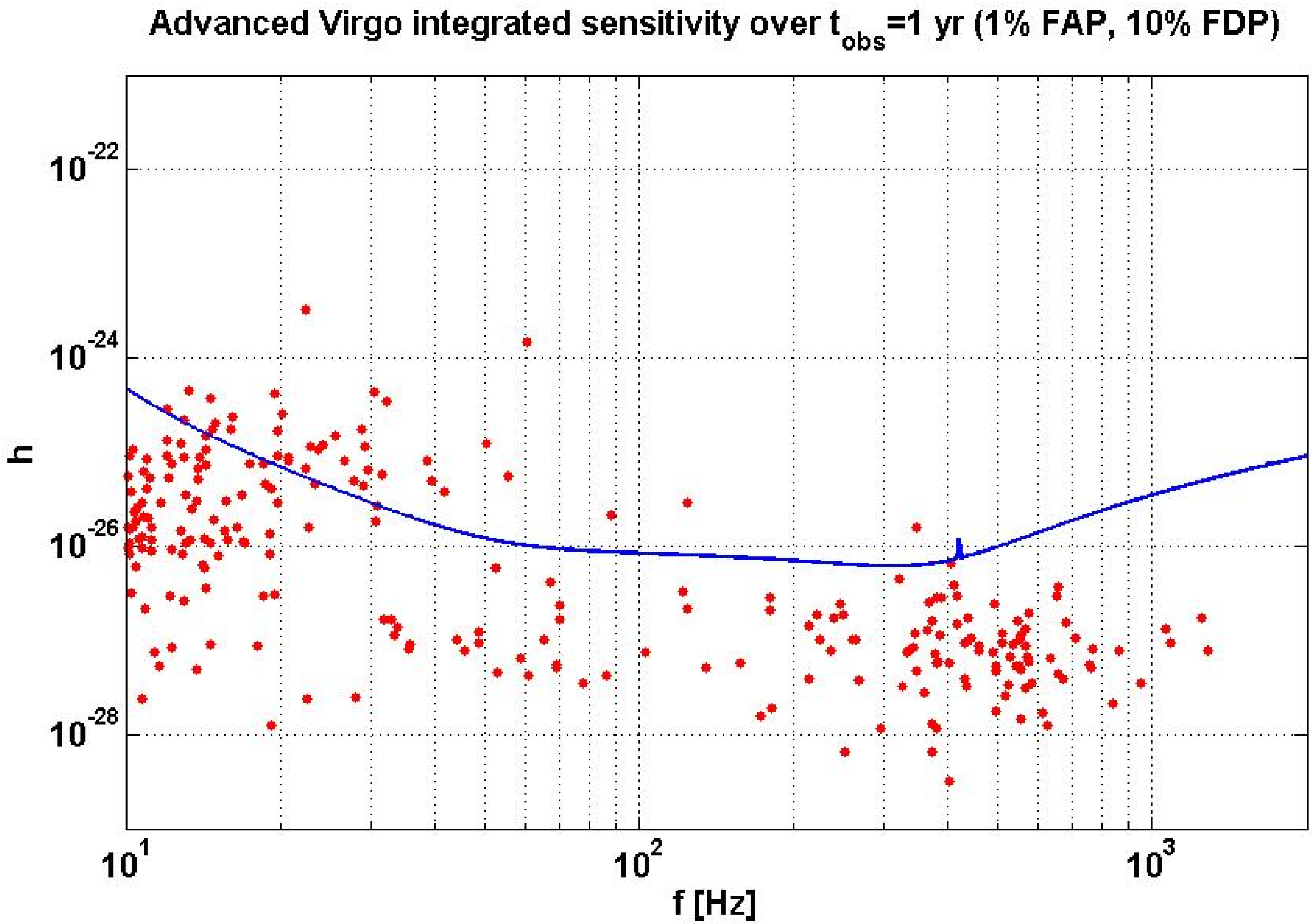
Astrophysics - 1

Let us see what we can do with current and next generation ITF.



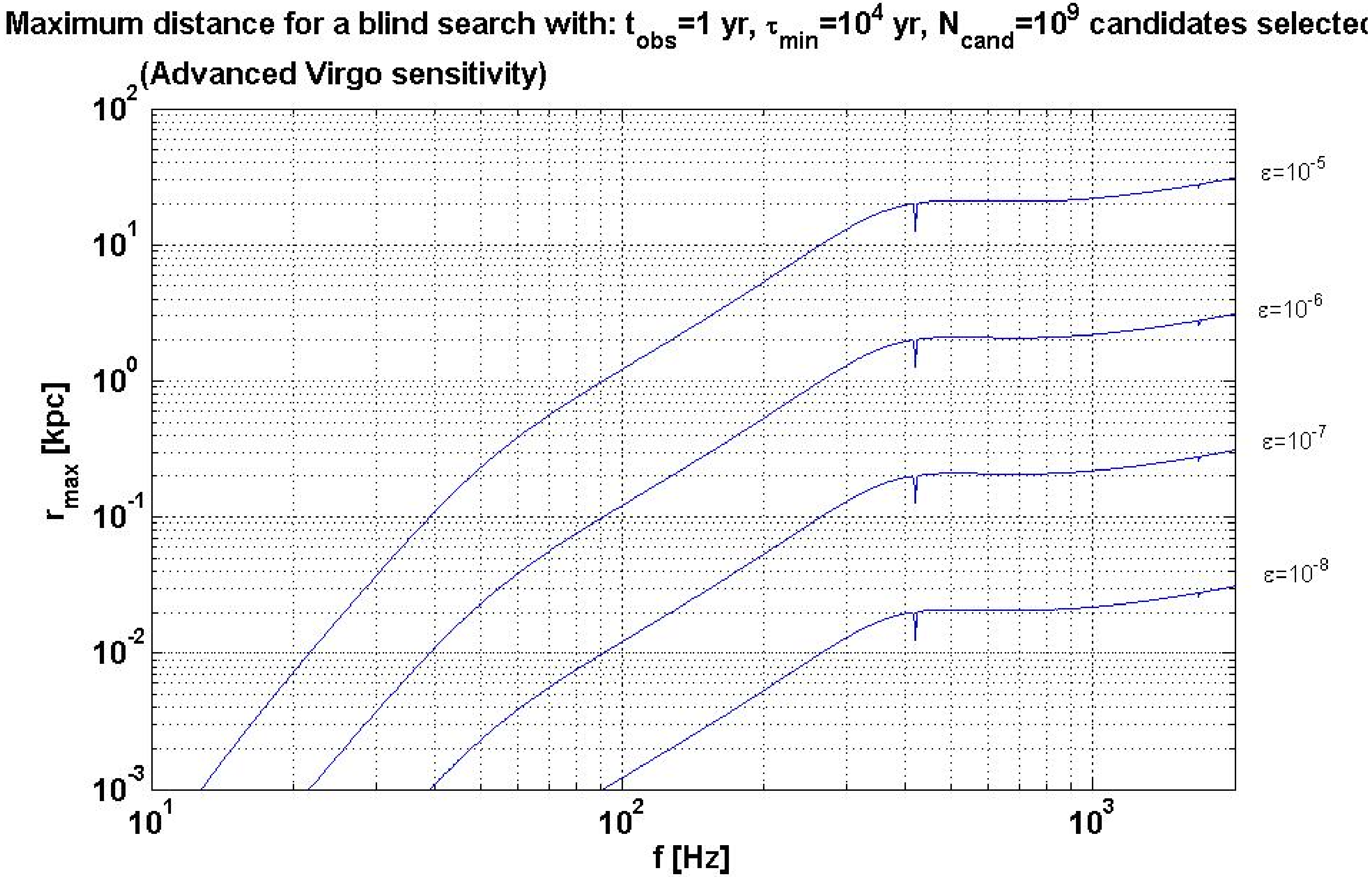
The spin-down limit is assumed for the sources.

Astrophysics - 2



The spin-down limit is beatable for many known pulsars.

Astrophysics - 3



With advanced ITF we can reach the galactic centre at high frequency.