THE BASICS OF X-RAY TIMING

GianLuca Israel - gianluca@mporzio.astro.it

(INAF – OA Roma)

with thanks to Z. Arzoumanian, C. Markwardt, T. Strohmayer (lectures from previous X-ray school editions)

Why should I be interested?

What are the methods and tools?

What should I do?

Urbino, 31st July 08

PREFACE

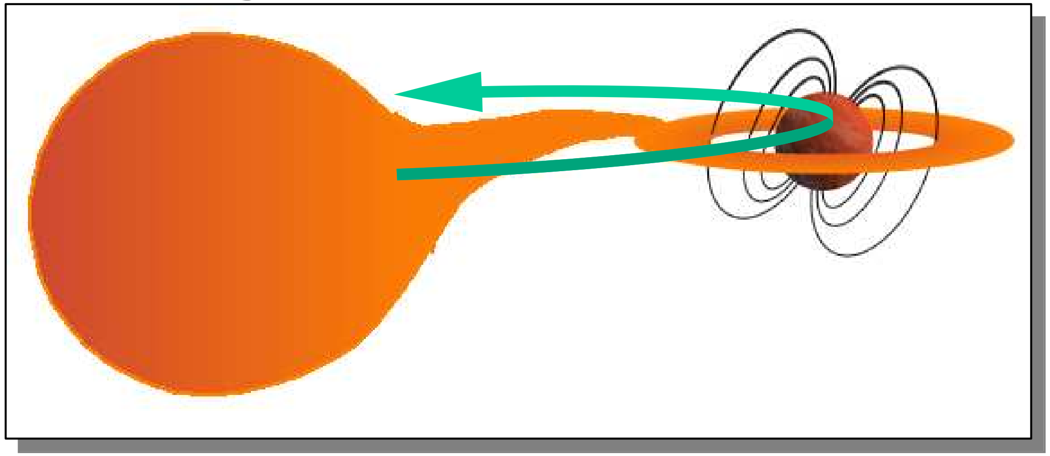
• **Incipit:** time series analysis is a very broad topic, and difficult to cover in one lecture.
• **Goal:** present the most important topics (partially) not discussed in the previous school editions.
• Timing analysis may seem a "magic box", since it can reveal features that are not apparent to the eye in the raw data
• Timing " analysis" is around since a long time: think about day/night, seasons, years, moon phases, etc.

OVERVIEW

• The relevance of timing analysis
• Basic light curve analysis (r.m.s.)
• Fourier power spectral analysis
• Power normalizations and signal searches
• Signal detection, signal UL and Asens
• Search optimization
• A working session example
• Cross-Correlation

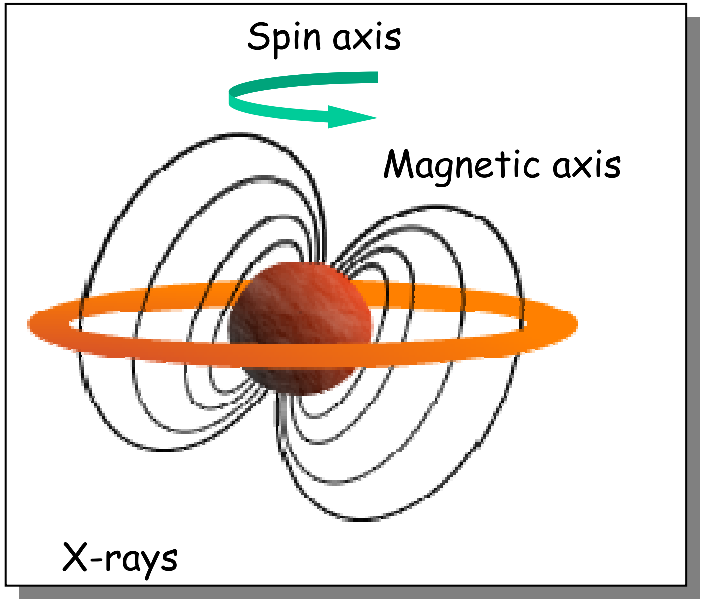
What can timing tell us? (or, why should I be interested?)

Timing => characteristic timescales = PHYSICS
Timing measurements can be extremely precise!!



Binary orbits
- orbital period
- sizes of emission regions and occulting objects
- orbital evolution

Rotation of stellar bodies
- pulsation periods
- stability of rotation
- torques acting on system



Accretion phenomena
- broadband variability
- "quasiperiodic" oscillations (QPOs)
- bursts & "superbursts"
- Energy dependent delays (phase lags)

TYPICAL SOURCES OF X-RAY VARIABILITY

- Isolated pulsars (ms–10 s)
- X-ray binary systems
  - Accreting pulsars (ms–10000 s)
  - Eclipses (10s min–days)
  - Accretion disks (~ms–years)
  - Transients orbital periods (days-months)
- Flaring stars & X-ray bursters
- Cataclysmic Variables (s-days)
- Magnetars (μs-s)
- Pulsating (non-radial) WDs (min-days)

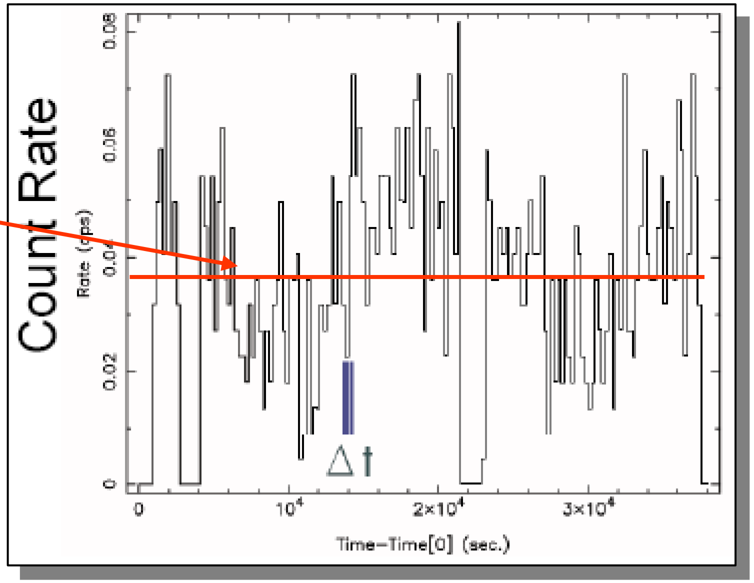
There could be variable serendipitous sources in the field, especially in *Chandra* and *XMM* observations

In short, compact objects (& super-massive black holes?) are, in general, intrinsically variable.

SIMPLEST MEASURE OF VARIABILITY

- The root-mean-square variability (the same as standard deviation):

$$\text{r.m.s.} = \sqrt{\frac{1}{N}\sum_i (\text{RATE}_i - \langle \text{RATE} \rangle)^2}$$



- Also, it is common to quote the fractional r.m.s., r.m.s./<RATE>

Limit: the above def. is bin-size dependent (i.e. You miss any variations smaller than your time bin size)

……. moreover

- We must remember that the light curve has Poisson counting noise (i.e. Some randomness), so we EXPECT some variation even if the source has a constant intrinsic intensity.

CHI-SQUARE TEST

- *Hypothesis: the source is intrinsically constant*
  - Can I reject this hypothesis?
  - Chi-square statistic

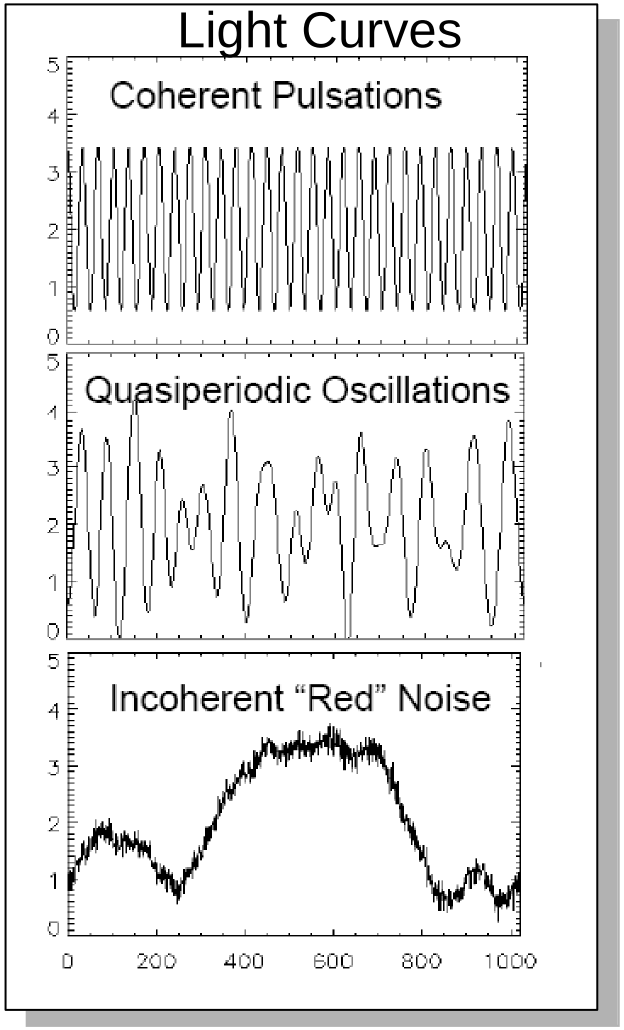
$$\chi^2 = \sum_i \left(\frac{\mathrm{RATE}_i - <\mathrm{RATE}>}{\mathrm{ERROR}_i}\right)^2$$

  - If measurements are gaussian (!), the statistic should have a chi-square distribution with (N-1) degrees of freedom.
  - We can calculate the statistic, compare to tabulated values, and compute confidence in our hypothesis.
- An alternative test for variability is the K-S test

Limits:

- So far, our analysis has focused on the total variability in a light curve.
- This method cannot isolate particular timescales of interest.
- If we are interested in faster time scales (higher frequencies), we must make a light curve with smaller time bins
- The assumption of gaussian statistics eventually fails, when the number of counts per bin is less than ~10, and this method is no longer useful.

SYNTHETIC DATA



Note that all light curves have 50% fractional r.m.s. variability

Implication: TOTAL variability (r.m.s.) does not capture the full information. Its time-scale or (frequency scale) is important as well.

FOURIER ANALYSIS

• This important technique comes from the theorem that any signal can be written as a sum of complex exponentials:

$$f(t_j) = \frac{1}{N} \sum_{k=1}^{N} a_k \exp\left(-2\pi i j k / N\right)$$

• The a_k terms are known as Fourier coefficients (or amplitudes), and can be found by using the Fourier transform (usually a FFT). They are complex-valued, containing an amplitude and phase.

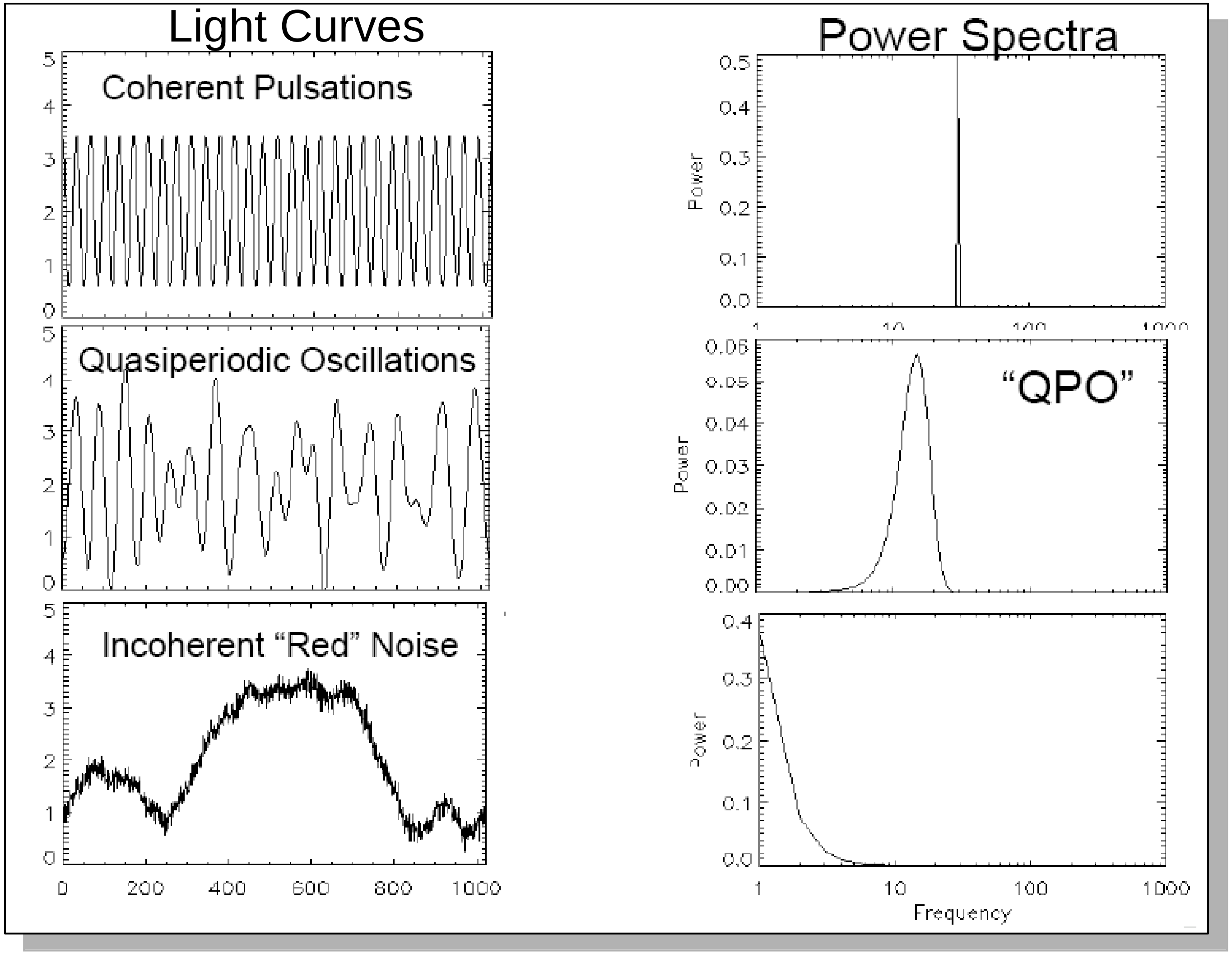
• Once we know the Fourier coefficients, we have divided the time series into its different frequency components, and have entered the frequency "domain."

• Parseval proved that: *Var[f_j]* = $\sum_{j}^{N-1} f_j^2 = \frac{1}{N} \sum_{k}^{N-1} |a_k|^2$

the left hand side is the total (r.m.s.)^2 variance, summed in *time*; the right hand side is the same total variance, summed over *frequencies*.

The values are known as Fourier powers, and the set of all Fourier powers is a POWER SPECTRUM (PSD).

FOURIER ANALYSIS-2



Instrumental noise not included ! When dealing with noise one also need a statistical tool to handle it.

COMPUTING USEFUL POWER SPECTRA

• Power spectra are commonly normalized in two different ways.

• The "Leahy" normalization is useful for computing significances (DETECTION). In the following we will refer to it as the default

• The "density" normalization is useful for computing fractional r.m.s. Variabilities (PHYSICS)

LEAHY NORMALIZATION

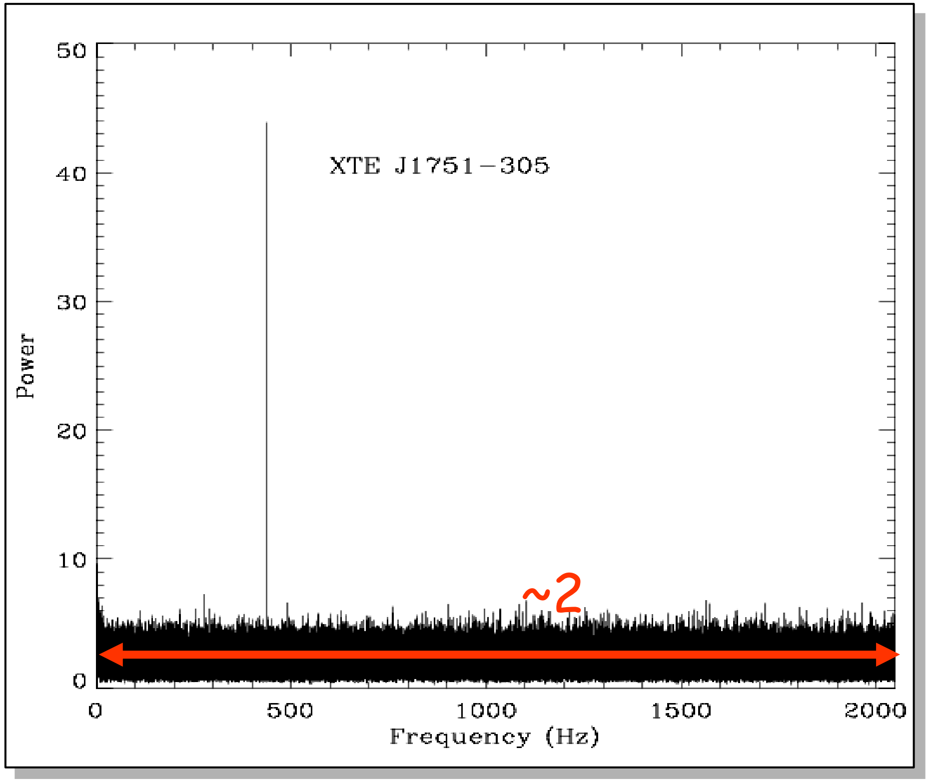
$$P_k = \frac{2}{N_{\mathrm{ph}}} |a_k|^2$$

Expected Poisson Variance

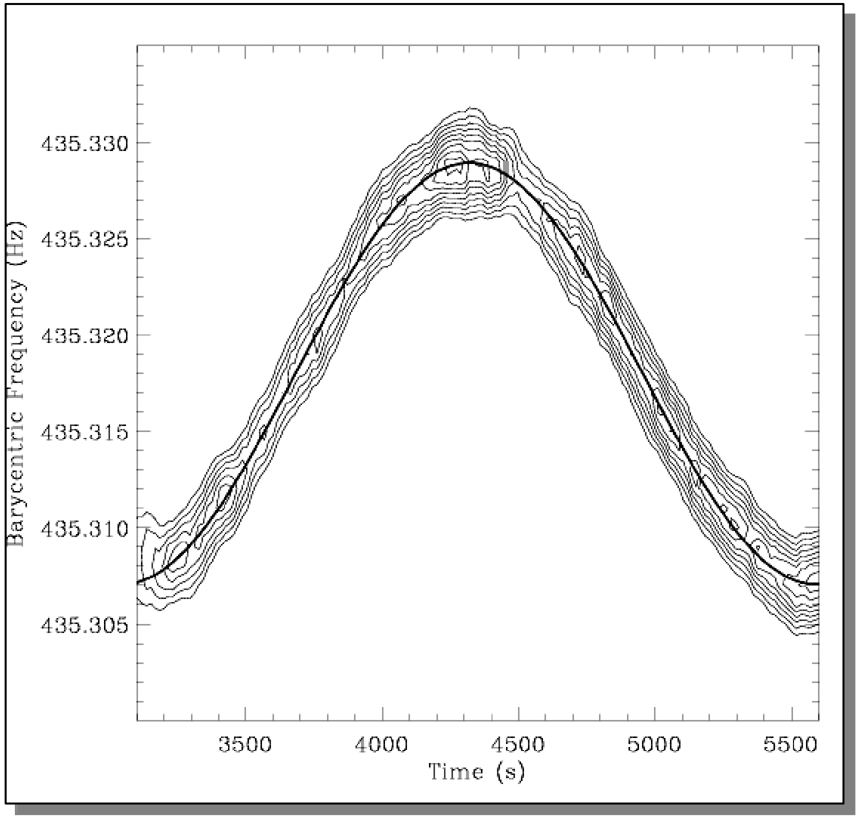
• N_{ph} is the total number of photons
• With this normalization, the Poisson noise level is distributed like a χ^2 with ν =2N_{PSD} degrees of freedom (in units of counts; N_{PSD} is the number of averaged PSD)
- $E[\chi^2|\nu] = \nu$ → 2 for N_{PSD}=1
- $\sigma[\chi^2|\nu]$ = sqrt(2ν) → 2 for N_{PSD}=1 → noisy

EXAMPLE: ACCRETING PULSAR, ORBIT, TIME DELAYS

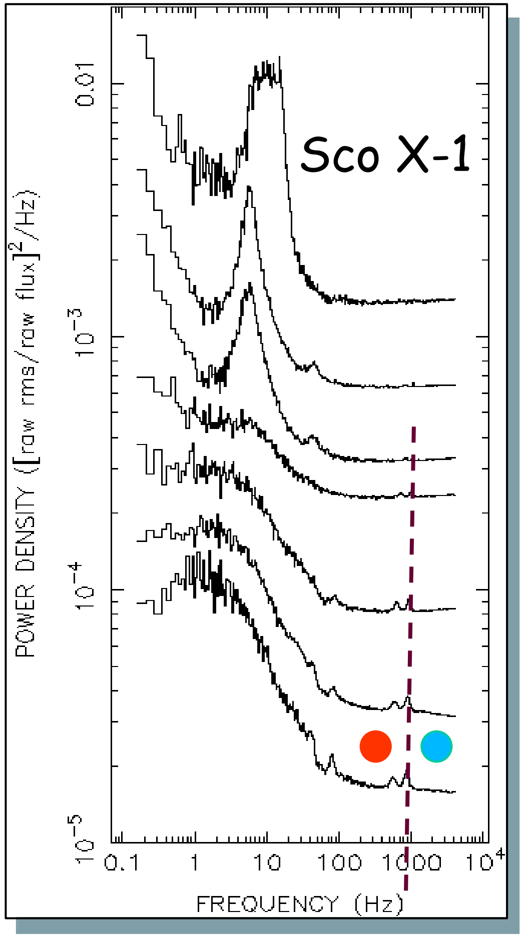
XTE J1751-305: accreting ms pulsar.

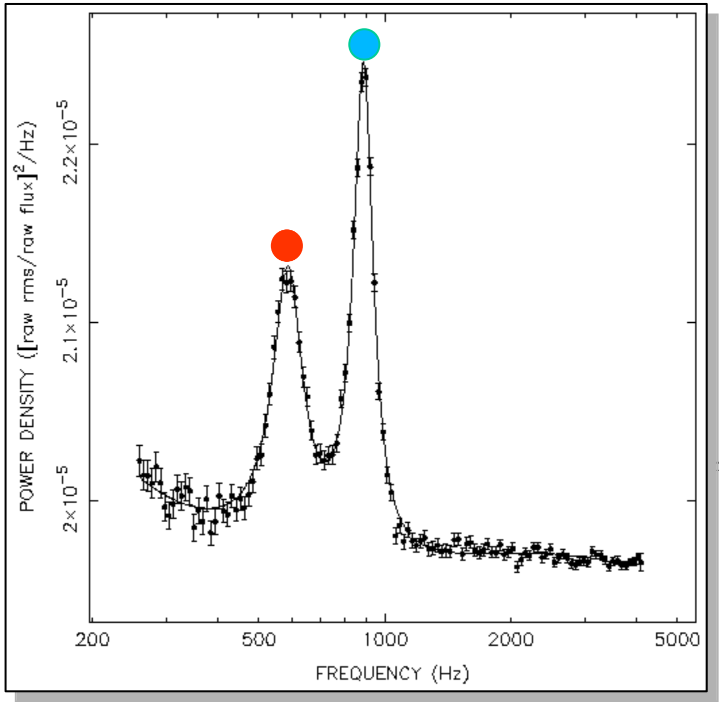


Ex: Period drift testifies of an orbit with period of ~40min

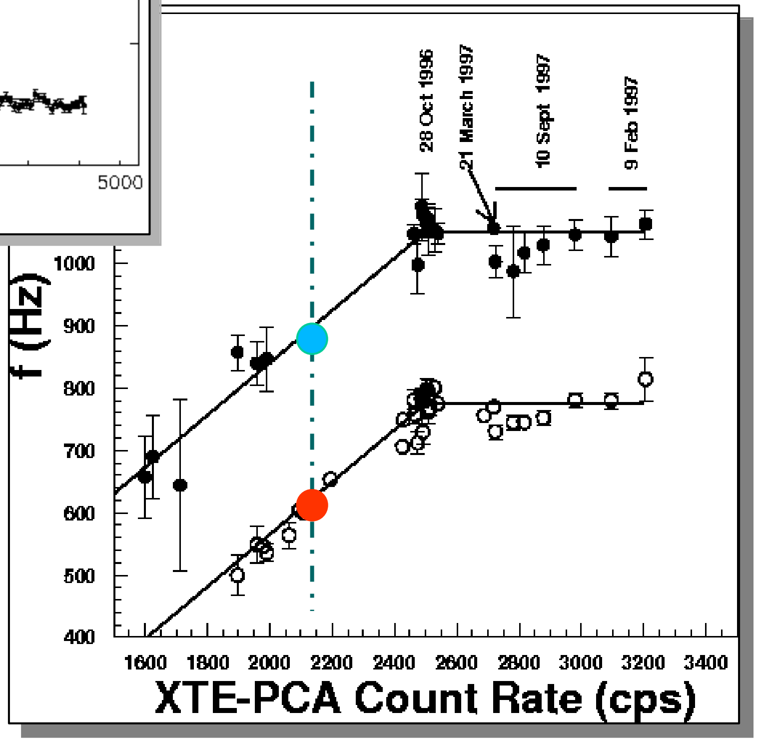


Example: QPOs from NS binaries



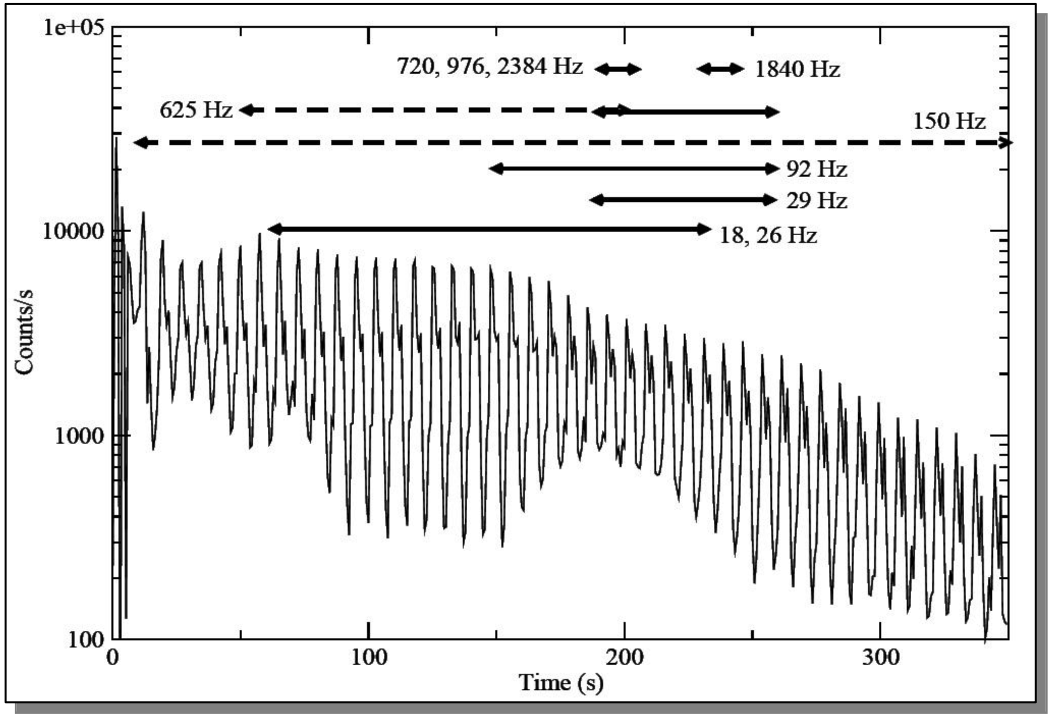


Sub-ms oscillations seen from > 20 NS binaries

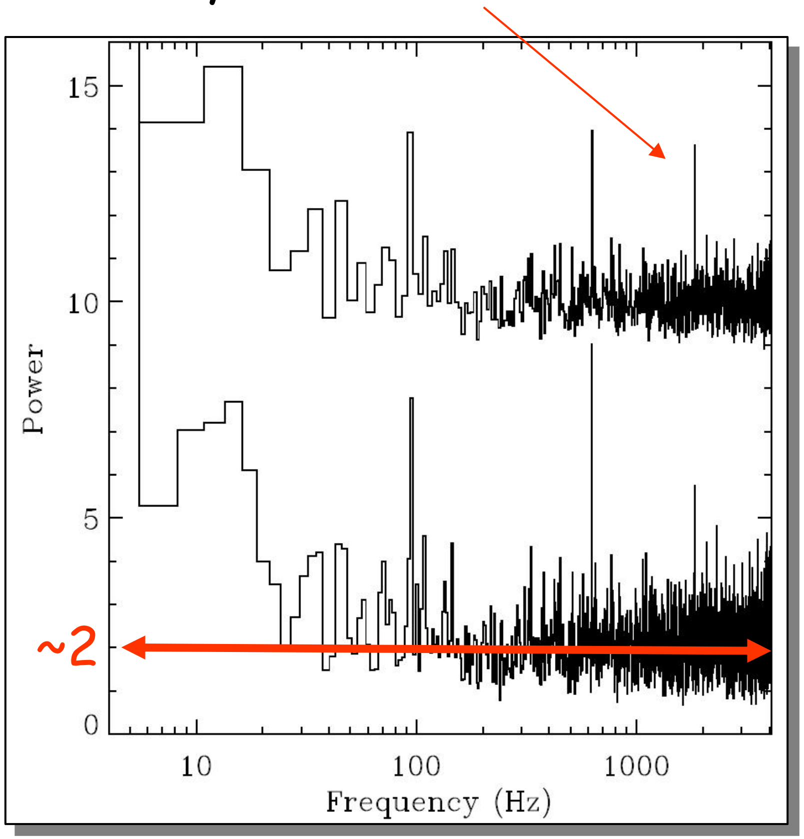


Frequencies saturate to a maximum. This is likely the signature of the ``innermost stable circular orbit'' around a neutron star, a radius predicted by general relativity inside of which matter must inexorably spiral down to smaller radii

Example: Magnetar sub-ms QPOs



The shortest period in (X-ray) astronomy



SGR1806-20 Giant Flare (Dic 2004)

A sequence of QPO frequencies was detected: 18, 29, 92 and 150, 625, and 1840 Hz!

Amplitudes in the 5 – 11% range.

Likely interpretation: A sequence of toroidal modes

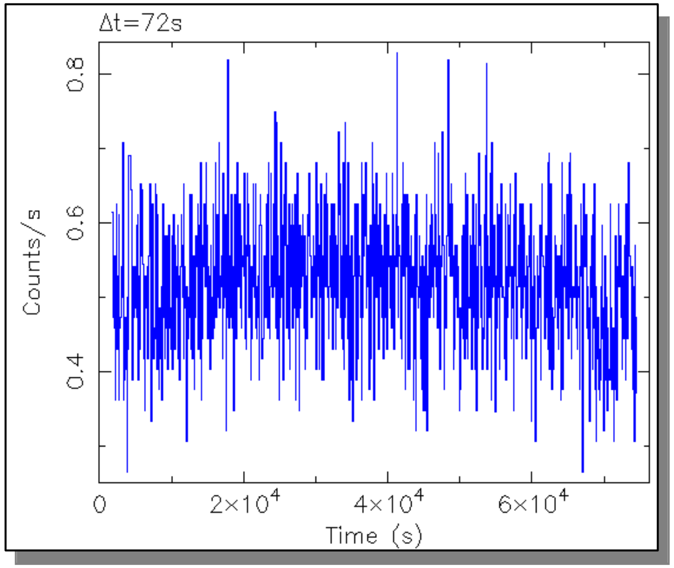
POWER SPECTRUM MAIN PARAMETERS

If your light curve has N bins, with bin size Δt and total duration T, (NOT effective exposure time) then:

• The smallest frequency you can sample is $\nu_{min} = 1/T$: this is also the frequency separation between powers or frequency resolution)

• The largest frequency you can sample is $\nu_{max} = 1/(2\Delta t)$ (this is the "Nyquist" limiting frequency)

• ν_{min} and ν_{max} can be changed arbitrarly in order to study the continuum and narrow (QPOs/coherent signal) components of the PSD



Example:

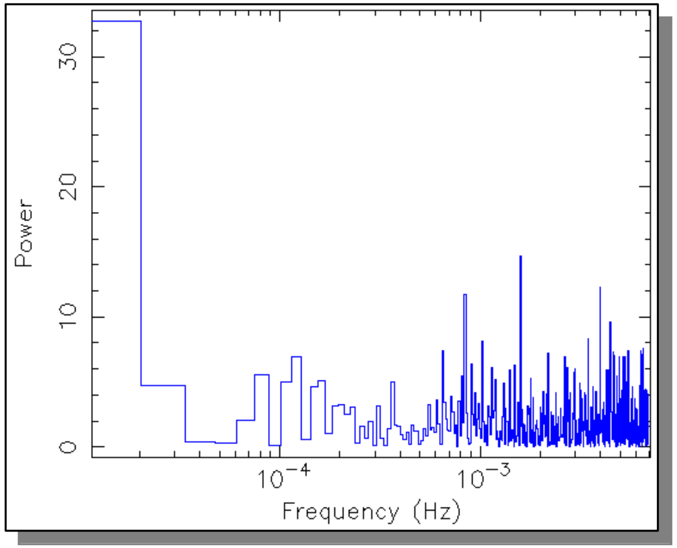
$\Delta t = 72s \rightarrow$

$\nu_{max} = 1/(2\Delta t) \sim 7e\text{-}3$ Hz

$T = 74000s \rightarrow$

$\nu_{min} = 1/T \sim 1.4e\text{-}5$ Hz

$T/\Delta t \sim 1024$ bins



The detection process in a PSD

The process of detecting something in a power spectrum against the background of noise has several steps:

- knowledge of the probability distribution of the noise powers
- The detection level: Number of trials (frequencies and/or sample)
- knowledge of the interaction between the noise and the signal powers (determination of the signal upper limit)
- Specific issues related to the intrinsic source variability (non Poissonian noise)
- Specific issues related to a given instrument/satellite (spurious signals – spacecraft orbit, wobble motion, large data gaps, etc.)

NOISE PROBABILITY DISTRIBUTION

For a wide range of type of noise (including that of counting photon detectors used in X-ray astronomy), the noise powers $P_{j,noise}$ follow a χ^2 distribution with $\nu=2N_{PSD}$ degrees of freedom.

$$Q(\chi^2 \mid \nu) \equiv \left[2^{\nu/2} \Gamma\left(\frac{\nu}{2}\right) \right]^{-1} \int_{P}^{+\infty} P^{\frac{\nu}{2}-1} e^{\frac{P}{2}} dP$$

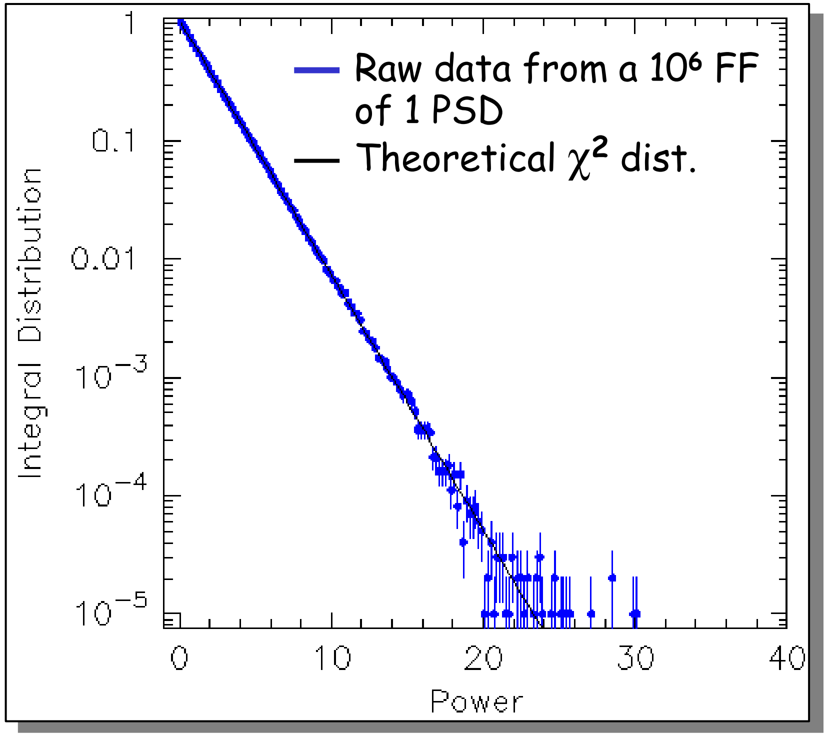
Ex: a power of 44 (in a white noise PSD) has a probability of $e^{-44/2}=3\mathrm{x}10^{-10}$ of being noise.

However, for $\nu=2$ it reduces to

$$Q(\chi^2 \mid 2) = e^{-\frac{P}{2}}$$

Correspondingly, the signal detection process results in defining a P_{thre}, such that the probability of having $P_{j,noise}$ > P_{thres} is small enough (according to the χ^2 probability distribution)

$$\mathrm{Prob}(P_{j,noise} > P_{thres}) = e^{-\frac{P_{thres}}{2}}$$



The search threshold and N_{trials}

- We define *a priori* a confidence level ($1-\varepsilon$) of the search (typically 3.5σ), corresponding to a power $P=P_{thres}$ which has a small probability ε to be exceeded by a noise power

- A crucial consideration, occasionally overlooked, is the number of trial powers N_{trial} over which the search has been carried out
  - o N_{trial} = to the powers in the PSD if all the Fourier frequency are considered;
  - o N_{trial} < than the powers in the PSD if a smaller range of frequencies has been considered;
  - o N_{trial} moltiplied by the number of PSD considered in the project

$$\frac{\varepsilon}{N_{trial} N_{PSD}} = Q(P_{thres} \mid 2) = e^{-\frac{P_{thres}}{2}}$$

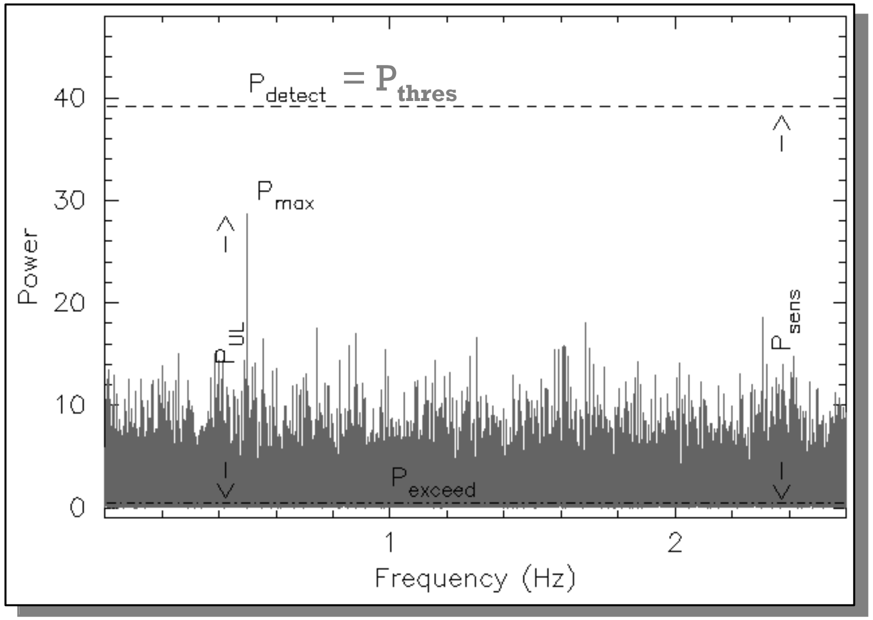
Ex: the previous probability of 3×10^{-10} has to be multiplied by 1.048.000 trial frequencies and 1 PSD
$Prob*N_{trial} = 3\times10^{-10} * 1.048.000 = 3\times10^{-4}$
Still significant!!

ULs AND THE SENSITIVITY TO THE SIGNAL

If no $P_j > P_{thres}$, it is useful to determine an upper limit to any signal power based on the **OBSERVED** properties. This is given by:

$P_{UL} = P_{max} - P_{exceed}$, where P_{max} is the largest actually observed power in the PSD and P_{exceed} is a power level which has a large probability to be exceeded by any $P_{j,noise}$.

It is sometimes useful to predict the capabilities of a planned experiment in terms of sensitivity to signal power. This is calculated based on the **EXPECTED** probability distribution of the noise.



$P_{sens} = P_{thres} - P_{exceed}$

Note that P_{sens} is in a sense the upper limit to P_{UL}.

Consideration: P_{sens} has to be used reported in proposal. P_{UL} is used when reporting a non detection in raw data.

ESTIMATING A_{sens} FOR PROPOSALS

You need the Intensity (cts/s) of the target and the T (s) of obs. → corresponding to net counts N_{ph}. Then, a confidence level has to be set (nσ) → defines P_{thres}

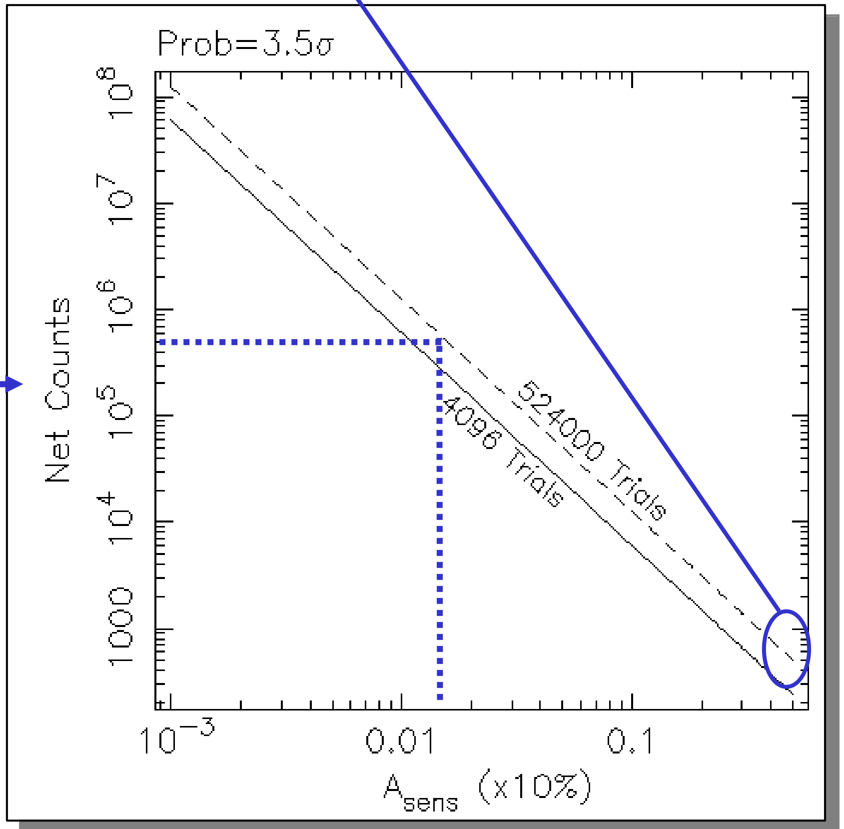
Based on know PSD properties one has :

$$A_{sens} < \left\{ 2.6 \frac{(\pi j / N)}{N_{ph} \sin^2(\pi j / N)} (P_{thres} - P_{exceed}) \right\}^{1/2}$$

$$\approx \left(\frac{2.6}{0.773 N_{ph}} P_{thres} \right)^{1/2}$$

relationship between A_{sens} and the N_{ph} →

Implication: signal detection is not possible for less than ~200 photons !

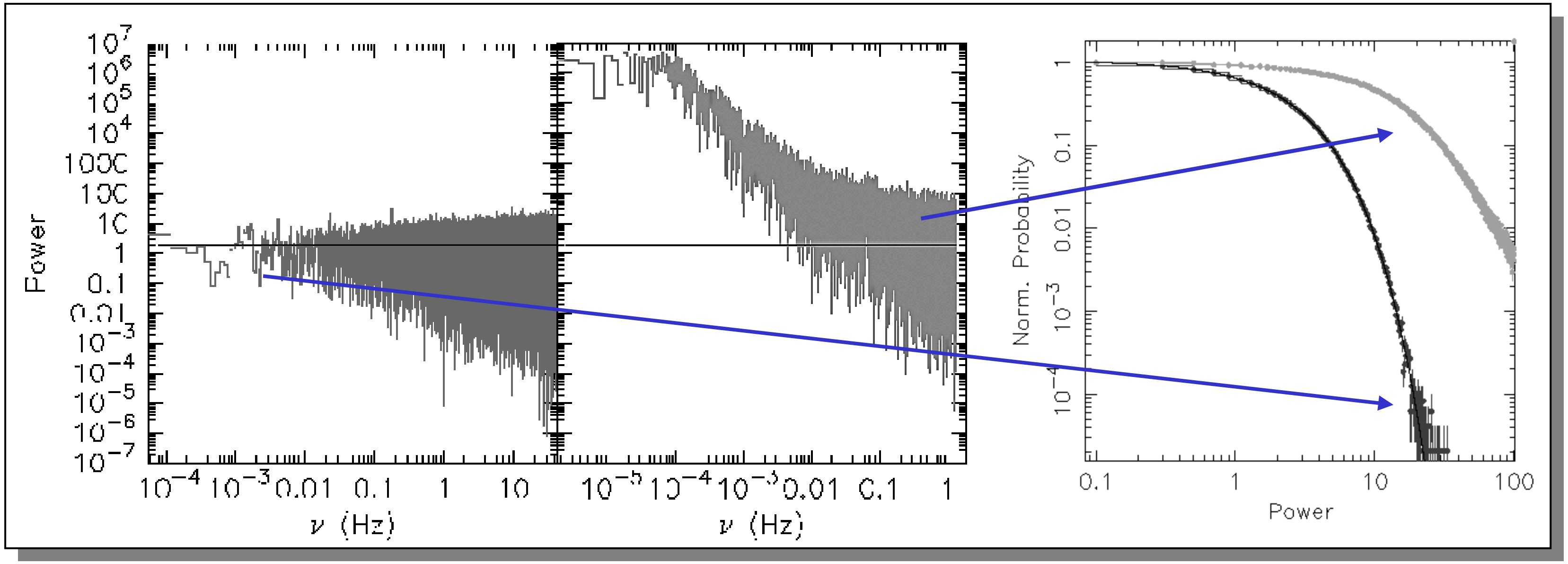


Ex: for a source of Intensity of 5ct/s, an exposure of T=100ks → N_{ph} =5e+5cts and P_{thres}~40 for 3.5σ c.l. (256000 N_{trial})

A_{sens}= [2.6*40/(0.773*5e+5)]^0.5=1.6%

INTRINSIC NON-POISSONIAN NOISE

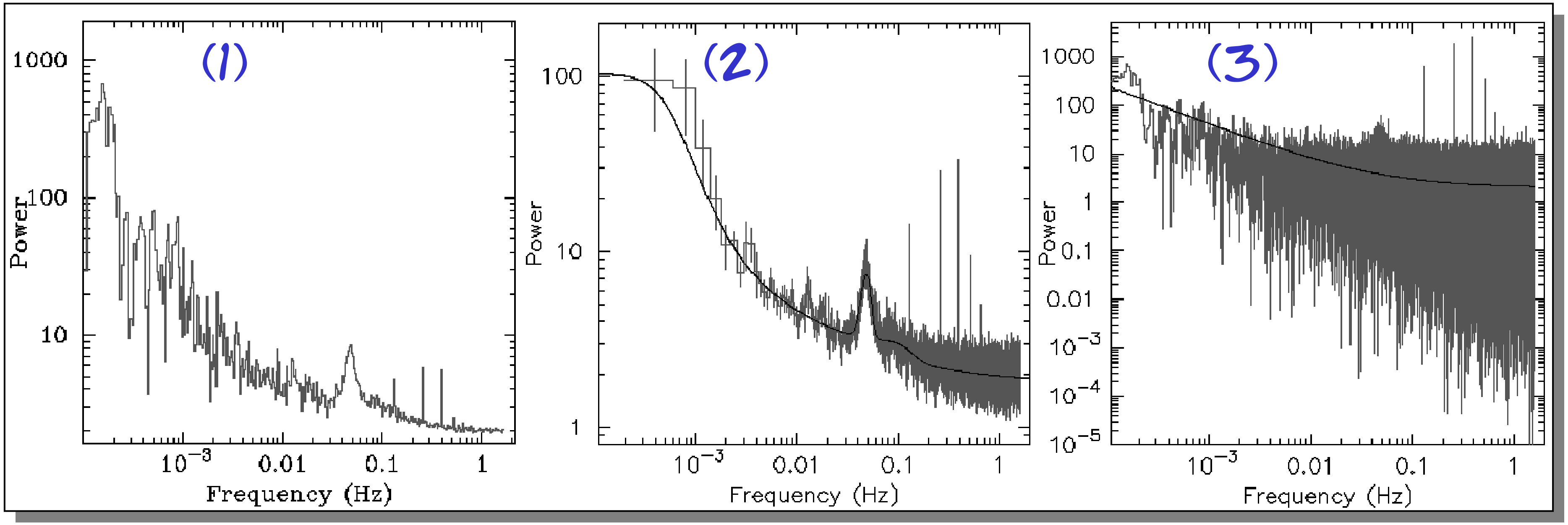
Many different classes of X—ray sources show aperiodic variability which translates into non-Poissonian noises (red-noise, blue-noise, low frequency noise, shot noise, etc.).



Implication: powers are not distributed anymore like a χ^2 with n d.o.f. → **no statistical tools to assess the significance of power peaks.**

INTRINSIC NON-POISSONIAN NOISE-2

Three different but similar approaches: (1) Rebin of the original PSD, (2) Average of more PSD by dividing the light curve into intervals, (3) Evaluation of the PSD continuum through smoothing. The common idea is to use the information of a sufficiently high number of powers such that it is possible to rely upon a known distribution of the new powers and/or continuum level (χ^2 or Gaussian or combination).

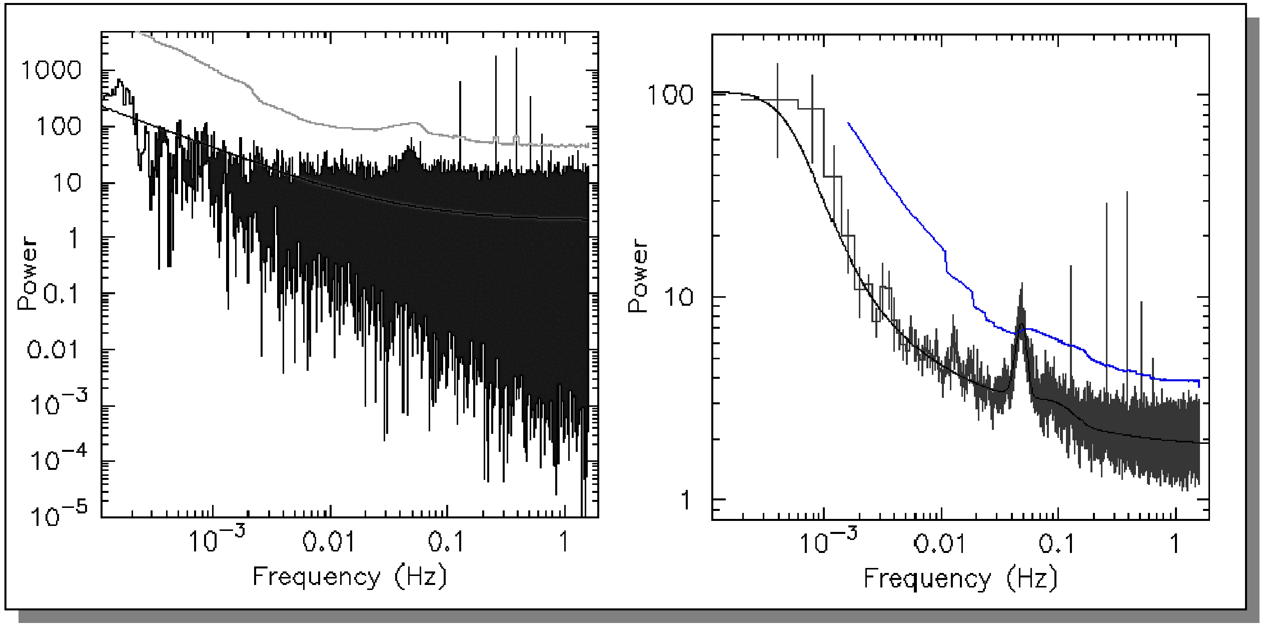


Note that the the processes above modify the PSD Fourier resolution (1/T), but leave unchanged the maximum sampled ν ($1/2\Delta t$)

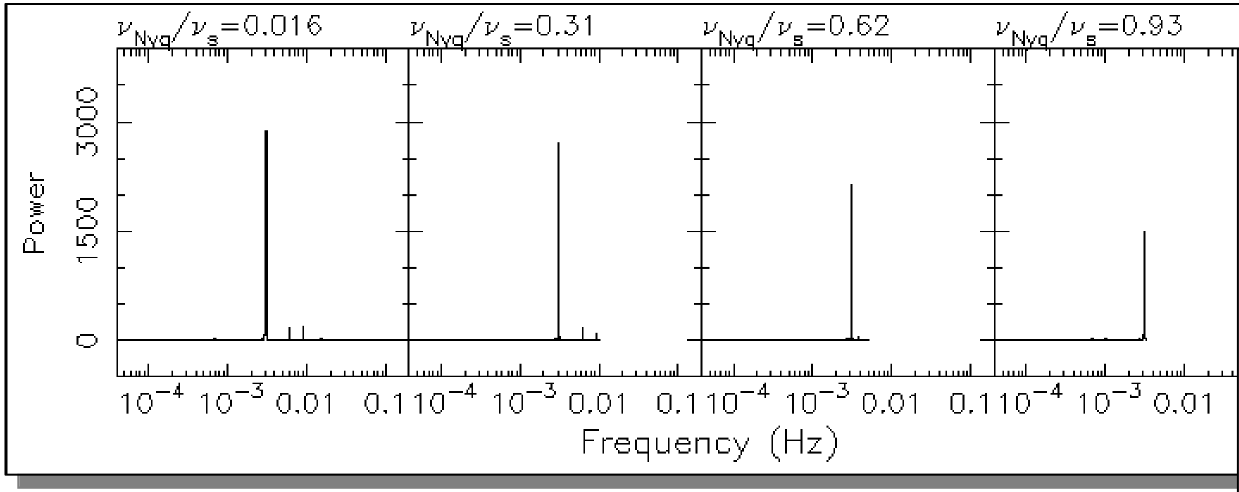
INTRINSIC NON-POISSONIAN NOISE-3

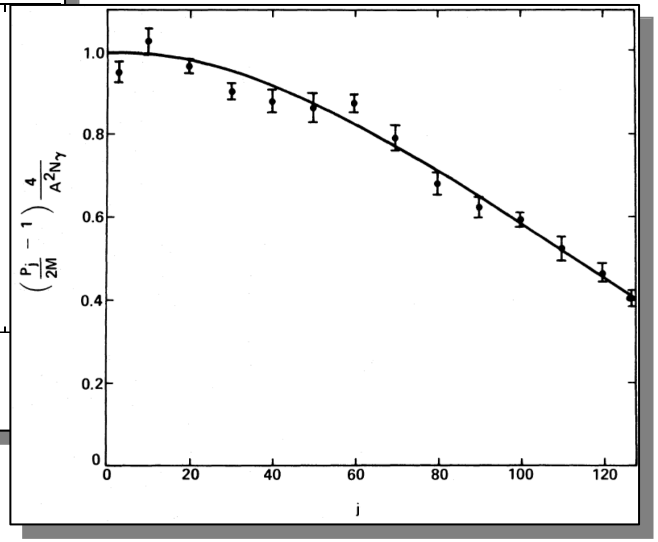
If M spectra are considered and/or W contiguous frequencies are averaged, the new variable (in cases 1 and 2) will be distributed like a rescaled χ^2/MW with 2MW d.o.f. In practice, everything is rescaled in order to have $E[\chi^2|2MW] = 2MW/MW = 2$.
Therefore $\sigma[\chi^2|2MW] = \mathrm{sqrt}(2MW)/MW$ → less noisy !!
Note that for MW>30÷40 the χ^2 → Gaussian

Implications: the noise scatter is largely reduced and faint and "extended" signals may be now detected.



Signal Detection Optimization





$$A = \left\{ \left[\frac{< P_j >}{2MW} - 1 \right] \frac{4}{0.773 N_{ph}} \frac{(\pi j / N)^2}{\sin^2(\pi j / N)} \right\}^{1/2}$$

The presence of the $x^2/\sin^2 x$ term in the amplitude relationship implies a strong correlation between signal power and its location (in terms of Fourier ν_j) with respect to ν_{Nyq}. The power-signal response function Decreases of 60% (from 1 to 0.405) from the 1st and last freq.

Implications: When searching for coherent o quasi-coherent signals It is important to use the original (if binned time series) or minimum (if arrival time series) time resolution → ν_{Nyq} = const.

SIGNAL DETECTION OPTIMIZATION-2

$$A = \left\{ \left[\frac{< P_j >}{2MW} - 1 \right] \frac{4}{0.773 N_{ph}} \frac{(\pi j / N)^2}{\sin^2(\pi j / N)} \right\}^{1/2}$$

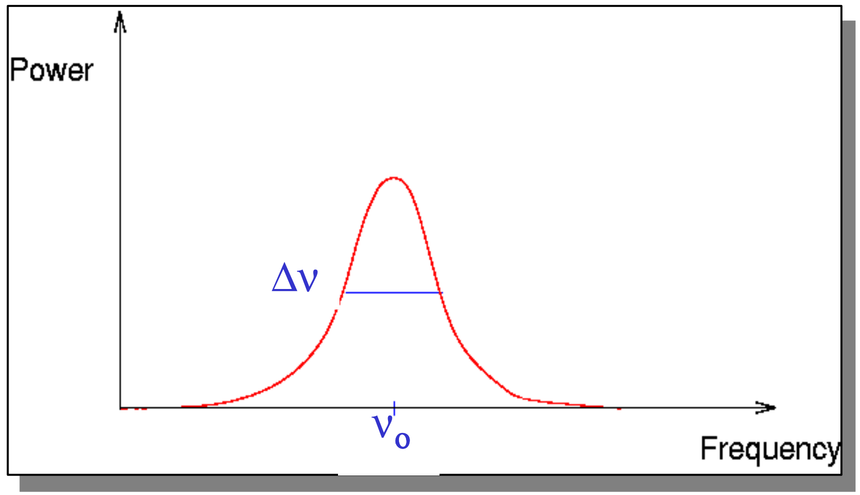
In the greatest part of the cases the signal freq. ν_{sig} is not equal to the Fourier freq. ν_j. The signal power response as a function of the difference between ν_{sign} and the closest ν_j, is again a $x^2/\sin^2 x$ term which varies between 1 and 0.5: for a coherent periodicity 1 means that all the signal power is recovered by the PSD, 0.5 means that the signal power is equally distributed between two adjacent Fourier frequencies ν_j.

Implications: When searching for strictly coherent signals it is important to rely upon the original/maximum Fourier resolution (1/T) → do not divide the observation in time sub-intervals.

Optimizing for the signal shape

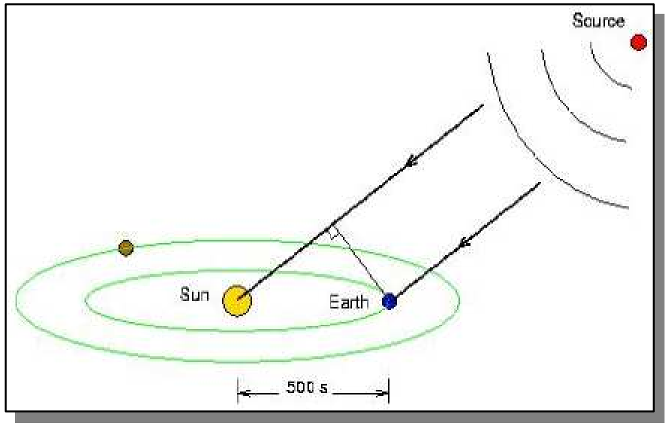
Similar reasoning shows that the signal power for a feature with finite width $\Delta\nu$ drops proportionally to *1/MW* when degrading the Fourier resolution. However, as long as feature width exceeds the frequency resolution, $\Delta\nu > MW/T$, the signal power in each Fourier frequency within the feature remains approx. constant.
When $\Delta\nu < MW/T$ the signal power begins to drop.

Implications: The search for QPOs is a three step interactive process.
Firstly, estimate (roughly) the feature width.
Secondly, run again a PSD by setting the optimal value of MW equal to $\sim T\,\Delta\nu$. Two or three iterations are likely needed.
Finally, use χ^2 hypothesis testing to derive significance of the feature, its centroid and r.m.s.



WHAT TO DO

Step 1. *Barycenter* the data: corrects to arrival times at solar system's center of mass (tools: fxbary/axbary depending on the given mission). Correct for binary orbital param. (if any)

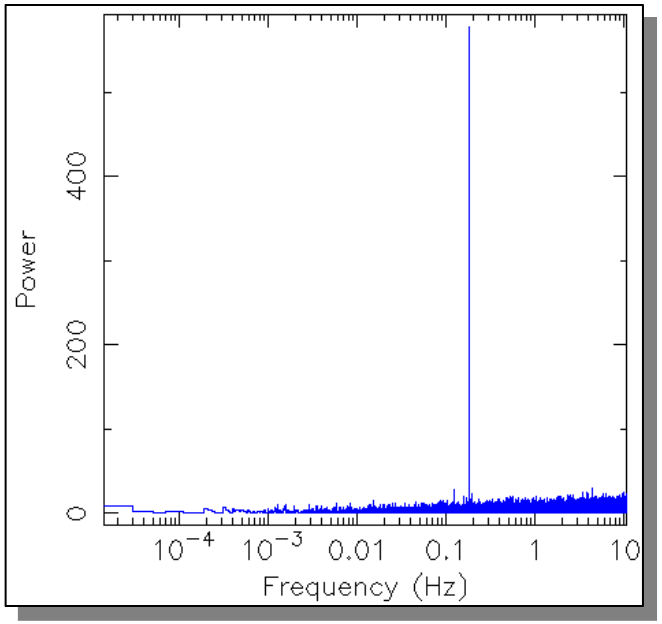


Step 2. Create light curves with lcurve for eachsource in your field of view inspect for features,e.g., eclipses, dips, flares, large long-period modulations.
lcstats give statistical info on the light curve properties (including r.m.s)

Step 3. Power spectrum. Run powspec or equivalent and search for peaks.
If no signal → calculate A_{UL} (or A_{sens})
If a peak is detected → infer ν_{sign}

One peak → likely sinusoidal pulse profile
More peaks → complicated profile

Example: ν_{sign} = 0.18 Hz → P_{sign} = 5.54s
T~48ks

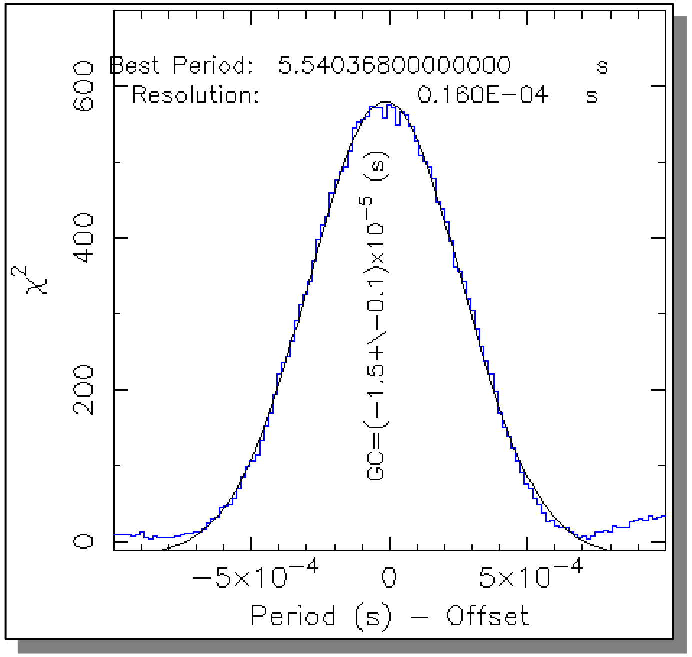


WHAT TO DO-2

Step 4. Use efsearch (P vs χ^2) to refine the period. **Step 1** if you already know the period.

Note that efsearch uses the Fourier period resolution (FPR), $P^2/2T$, as input default.
It depends from P !!!

To infer the best period the FPR has to be overestimated by a factor of several (ex. 20). Fit the resulting peak with a Gaussian and save the central value and ite uncertainty.
OK for period, not good for its uncertainty (which is the FPR)



Example: for a signal at 5.54s and T=48ks → FPR=3.2e-4s
FPR input = 3.2e-4/20=1.6e-5s
GC= $(-1.5\pm0.1)\times10^{-5}$s (1σ c.l.) → **P=**5.540368-0.000015 = **5.540353 s**
For the uncertainty is often used the GC error x 20 (the overestimation factor used in input). **ΔP=** 0.1×10^{-5} x 20 s = $\mathbf{2\times10^{-4}}$ **s**
Final Best Period: **5.5404(2) s** (1σ c.l.)

WHAT TO DO-3

Step 5. Use efold to see the modulation. Fit it withone or more sinusoids. Infer the pulsed fraction (several definitions) and/or the r.m.s. Remove the BG (it works like unpulsed flux).

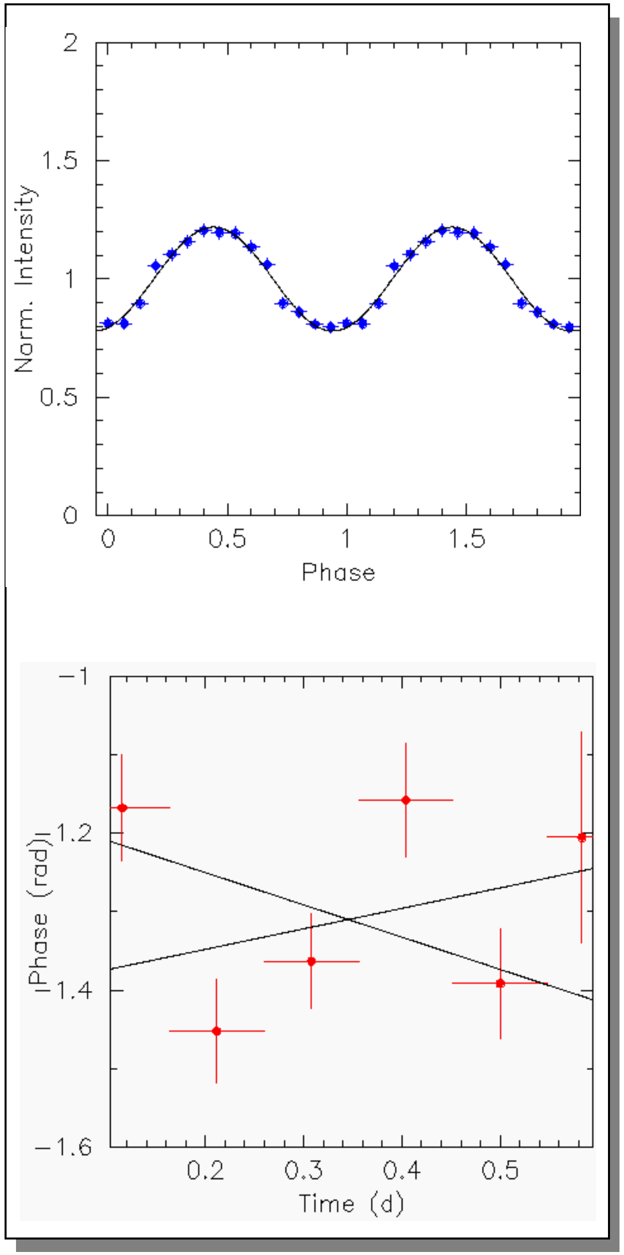
$$PF = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Ex: $PF = \frac{1.22 - 0.78}{1.22 + 0.78} = 0.22$

Step 4b. Apply a phase-fitting technique to your data (if enough photons). Use efold and save the sinusoid phase of pulse profiles obtained in 4 or more time intervals. Plot and fit Time vs Phase with a linear and quadratic component

- If the linear is consistent with 0 the input P is OK
- If a linear component is present the input P is wrong . Correct and apply again the technique.

Example: Best Period: **5.54036(1) s** 1σ c.l.
A factor of ~20 more accurate than efsearch



MORE ON PHASE-FITTING

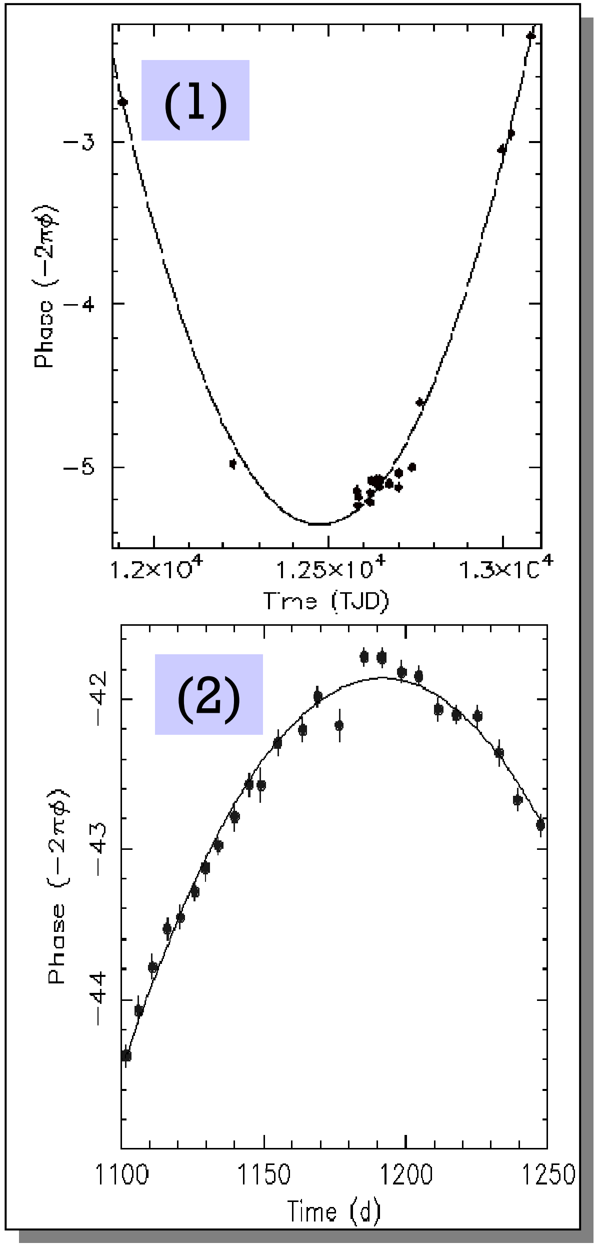
It provides a phase coherent timing solution which can be extended in the future and in the past without loosing the information on the phase, therefore, providing a tool to study small changes of signals on long timescales.

- A negative quadratic term in the phase residuals implies the period is decreasing

- A positive term corresponds to an increasing period

This method is often used in radio pulsar astronomy.

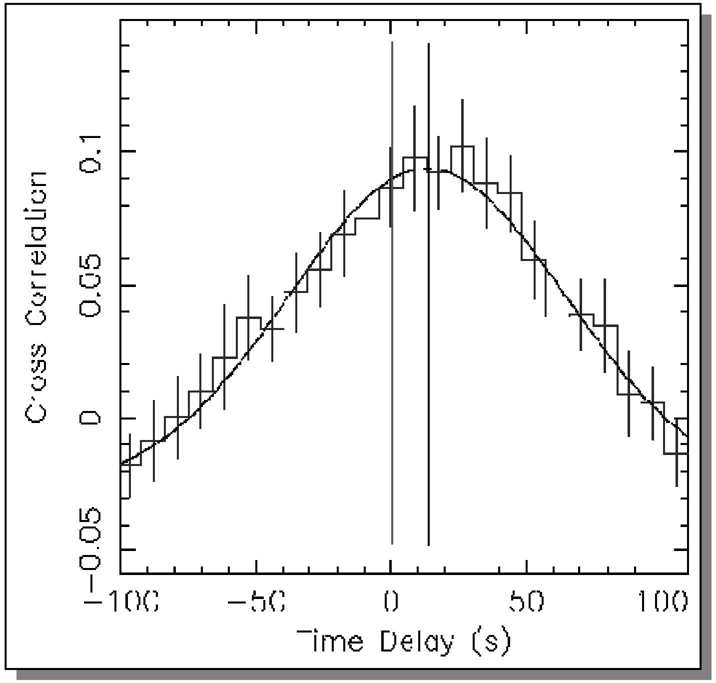
Examples: (1) a shrinking binary – orbital period decreasing at a rate of dP/dt=1ms/yr≈-3x10^{-11}s/s (2) An isolated neutron star spinning down at a rate of dP/dt≈1.4x10^{-11}s/s



CROSS-CORRELATIONS

The cross-correlation measures how closely two different observables are related each other at the same or differing times. It also gives information on possible delays or advances of one variables with respect to the other (in practical cases one deals with times or phases).

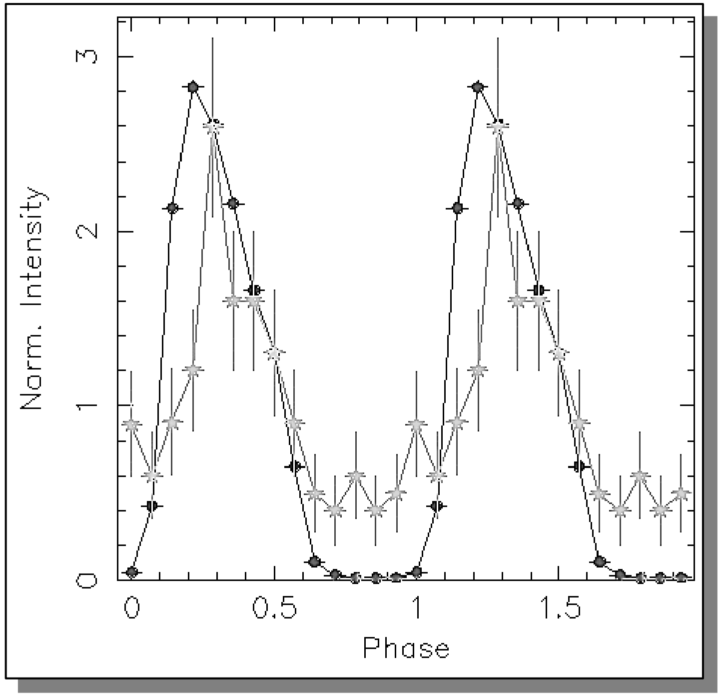
Example: CCF obtained with crosscor. Two simultaneous light curves of a binary system in two different energy intervals (soft and hard). The CCF peaks at positive x and y: the two variables are correlated and the hard variability follows the soft one. $\Delta t = 13 \pm 2$s (1σ c.l.).

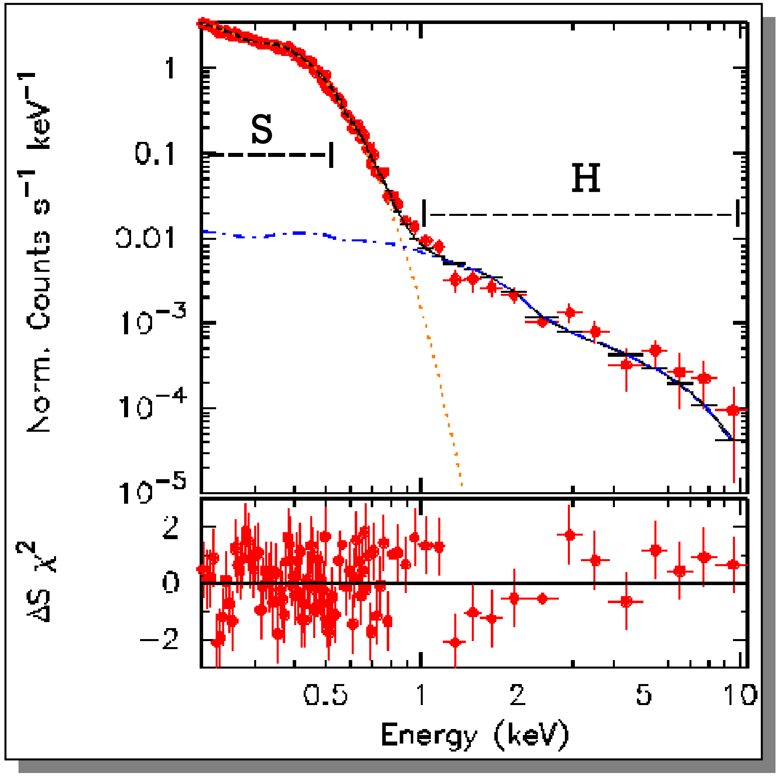


It is often useful to cross check the CCF results with the spectral information or any other useful timing result.

CROSS-CORRELATIONS-2

Example: The folded light curves in the soft (black) and hard (gray) bands confirm the presence of a possible delay



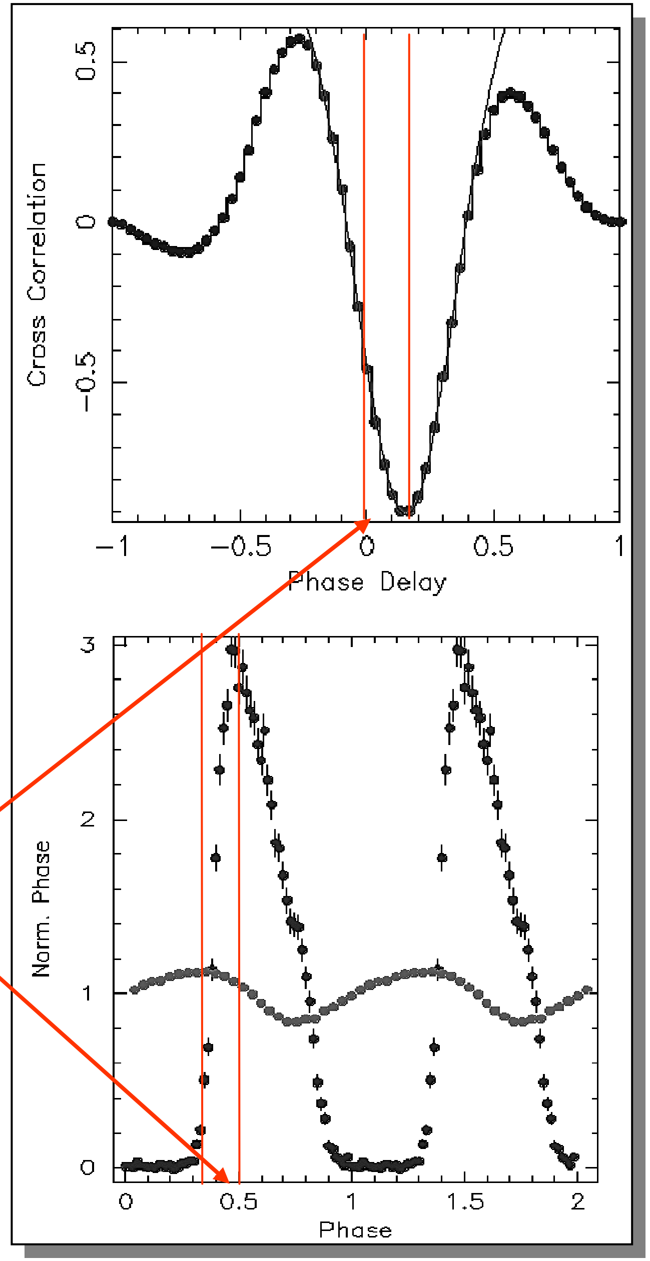


The study of the energy spectrum clearly reveals the presence of two distinct components (BB+PL) in the soft (S) and hard (H) energy bands considered for in the CCF analysis.
The CCF result is reliable/plausible !

CROSS-CORRELATIONS-3

Further considerations: CCF may be also applied to data taken in rather different bands (i.e. optical and X-ray) for a given source.

Example: Same source as before, CCF obtained for the optical and X-ray folded light curves (obtained with efold) over a 4-years baseline. Pseudo-simultaneous data: same phase coherent time solution used.
The CCF peaks at positive x and negative y: the optical and X-ray data are anti-correlated with the optical one proceeding the X—rays by 0.16 in phase.



TIPS

Pulsar (coherent pulsation) searches are most sensitive when *no rebinning* is done (ie., you want the maximum frequency resolution), and when the original sampling time is used (i.e. optimizing the signal power response). Always search in all serendipitous sources (N_{ph}>300)

QPO searches need to be done with *multiple rebinning* scales. In general, you are most sensitive to a signal when your frequency resolution matches (approximately) the frequency width of the signal.

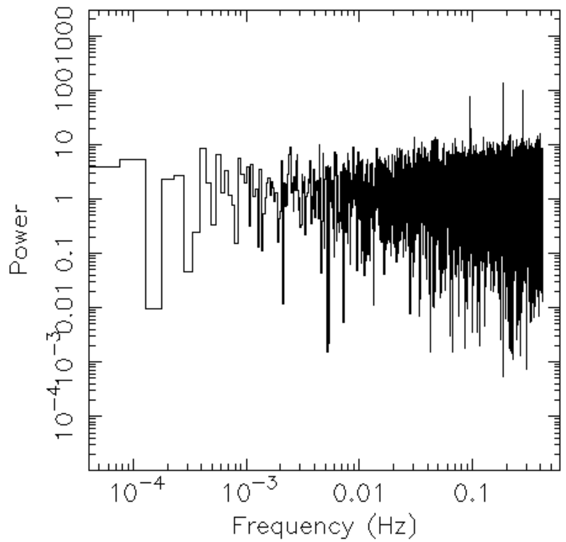
CCF: it is worth using it to study the relation among different energies

Cross-check with spectral information

Beware of signals/effects introduced by

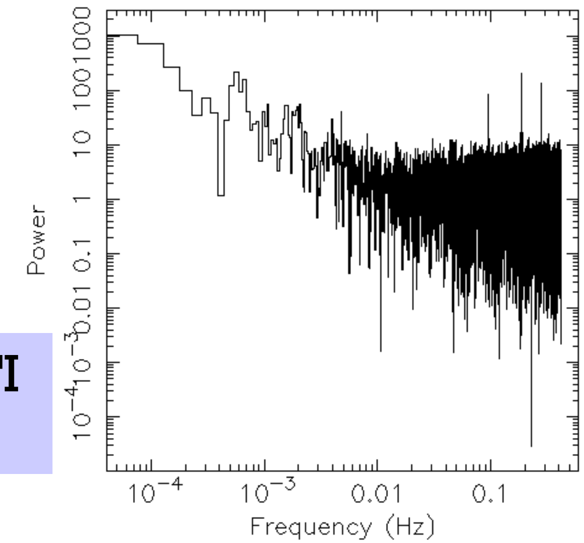
- instrument, e.g., CCD read time (check/add keyword TIMEDEL)
- Dead time
- Orbit of spacecraft
- Telescope motion (wobble,etc.)
- Data gaps
- Pile-up (wash-out the signal)
- Orbital binary motion (")
- The use of uncorrect GTIs (for single and merged simult. light curves)

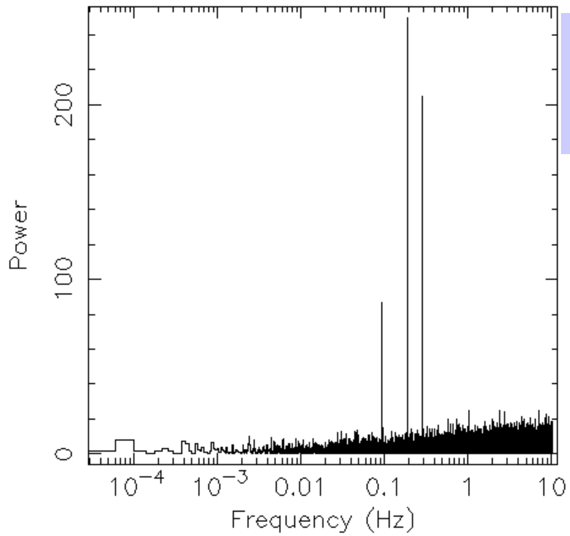
TIPS-2



Right GTI table

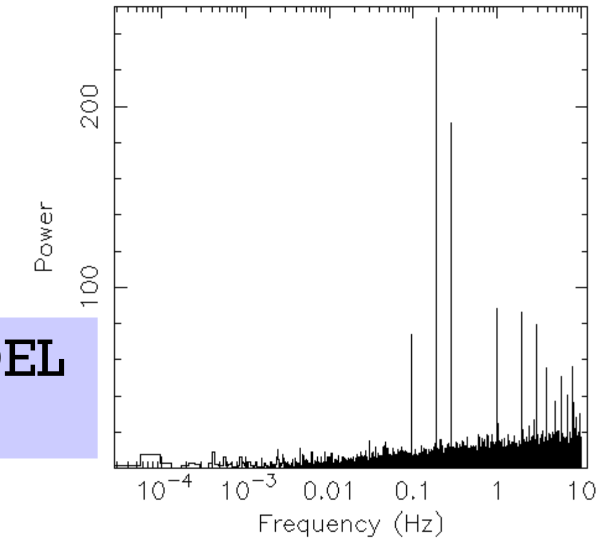
Wrong/no GTI table





Right TIMEDEL keyword

Wrong/no TIMEDEL keyword



Suggested reading

- van der Klis, M. 1989, "Fourier Techinques in X-ray Timing", in *Timing Neutron Stars*, NATO ASI 282, eds. Ögelman & van den Heuvel, Kluwer
  Superb overview of spectral techniques!

- Press et al., "*Numerical Recipes*" - Clear, brief discussions of many numerical topics

- Leahy et al. 1983, ApJ, 266, p. 160 - FFT & PSD Statistics

- Leahy et al. 1983, ApJ, 272, p. 256 - Epoch Folding

- Davies 1990, MNRAS, 244, p. 93 - Epoch Folding Statistics

- Vaughan et al. 1994, ApJ, 435, p. 362 - Noise Statistics

- Israel & Stella 1996, ApJ, 468, 369 – Signal detection in "noisy" PSD

- Nowak et al. 1999, ApJ, 510, 874 - Timing tutorial, coherence techniques

For more questions: gianluca@mporzio.astro.it

Urbino, 31st July 08