Effects of red noise on gravitational lensing experiments

(work still in progress!)

Achille A. Nucita[1]

D. Vetrugno[1], S. Carpano[2], F. De Paolis[1], G. Ingrosso[1]

1) **Department of Physics, University of Salento, and *INFN*, via per Arnesano, 73100 Lecce, Italy**
2) **European Space Agency (ESA), RSSD, Nordwijck, The Netherlands**

White Noise vs Red Noise (or coloured)

Implications in astrophysical light curves

Implications in gravitational lensing experiments

Some experiments for a full characterization

Definitions

When we want to maximize our chance to detect a weak signal, the choice of a test statistic is very important. It follows that the characterization of the noise is crucial in evaluating the signal's associated chance probability

By "chance probability" we mean the probability that the apparent signal is due to pure background fluctuation (the underlying noise process). Assigning a probability to a given value can only be done in reference to the distribution of the variable under consideration.

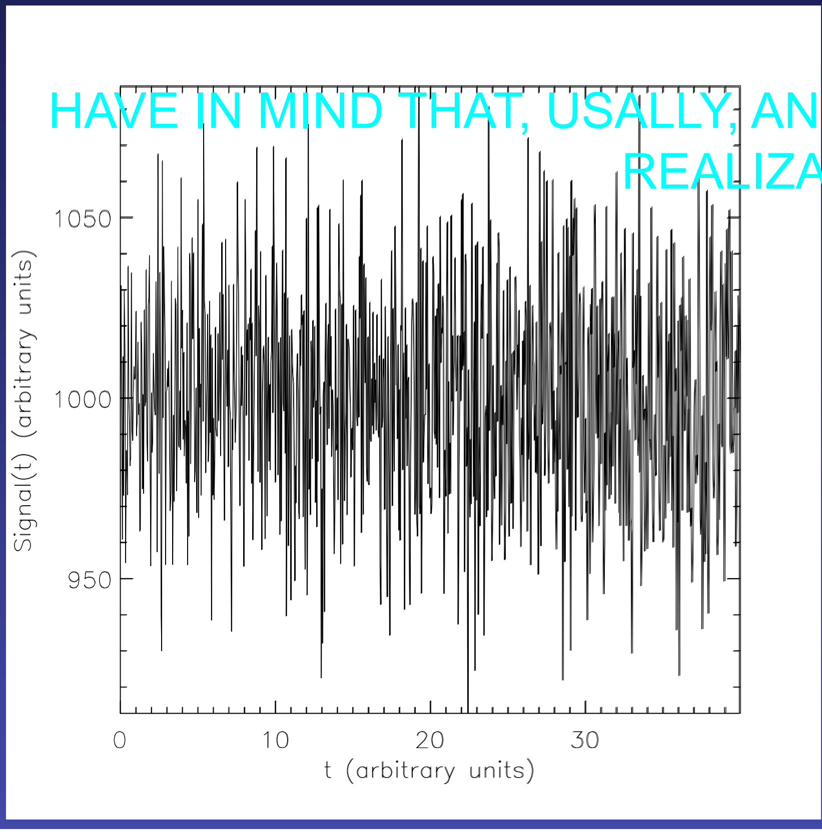
More in details, the chance probability of obtaining a value x, is the ratio of the area under the curve of the distribution from x to ∞, to the total area under the curve.

Normalizing the area of the probability distribution to 1 yields what is referred to as the **Probability Distribution Function** (PDF). An example: the probability of getting a 3σ result is 1.35×10^{-3}, when the random variable measured is distributed according to the Normal or Gaussian distribution.

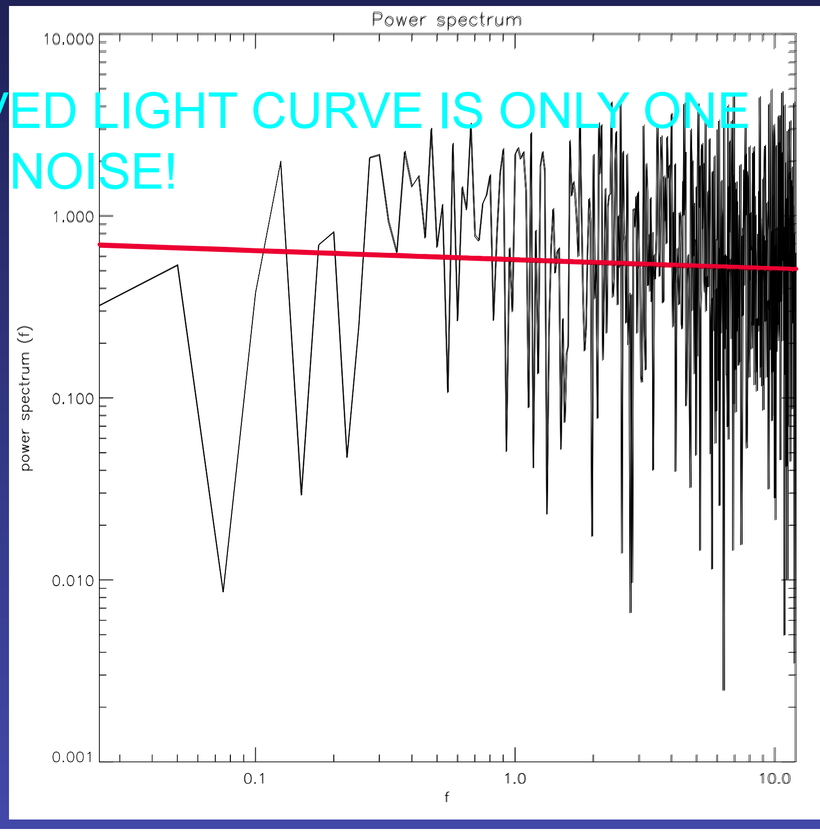
A test statistic is also a random variable with a specific PDF.

However, the shape of the PDF depends on the type of data to which we apply the test. For example, in Fourier analysis, a noise process that follows Poisson statistics will lead to a power spectrum with equal power at all frequencies, naturally referred to as *white noise.*

Given the well studied and predictable properties of such a noise process, it is in general possible to derive the PDF of a test statistic under the assumption that the background is indeed white noise.

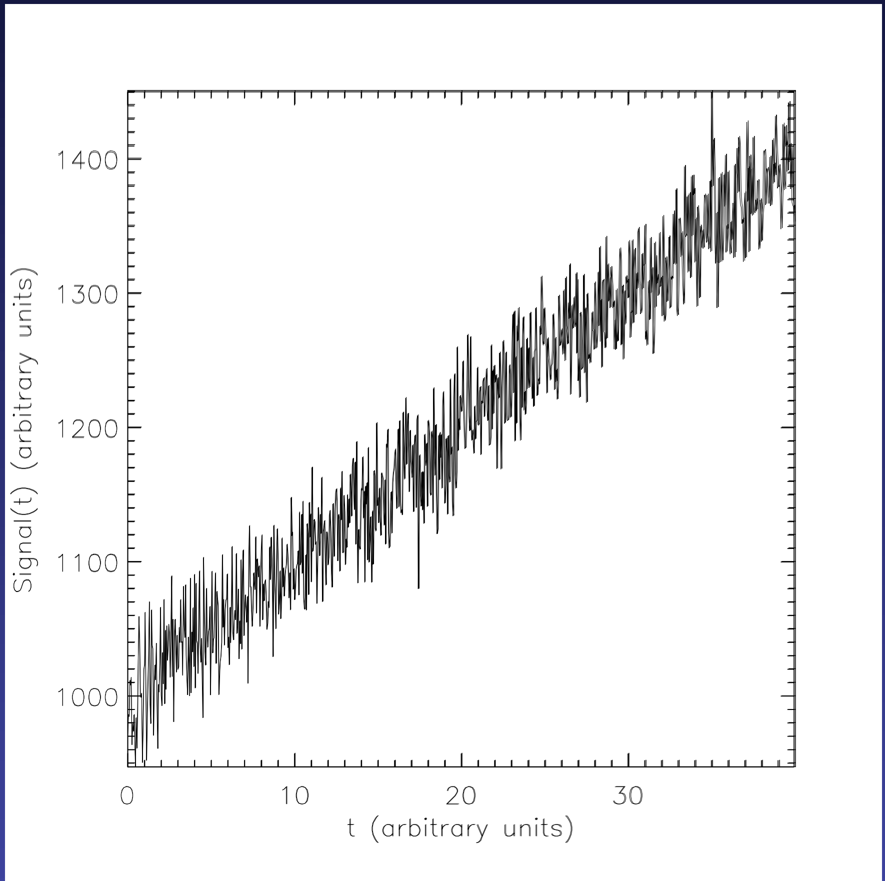


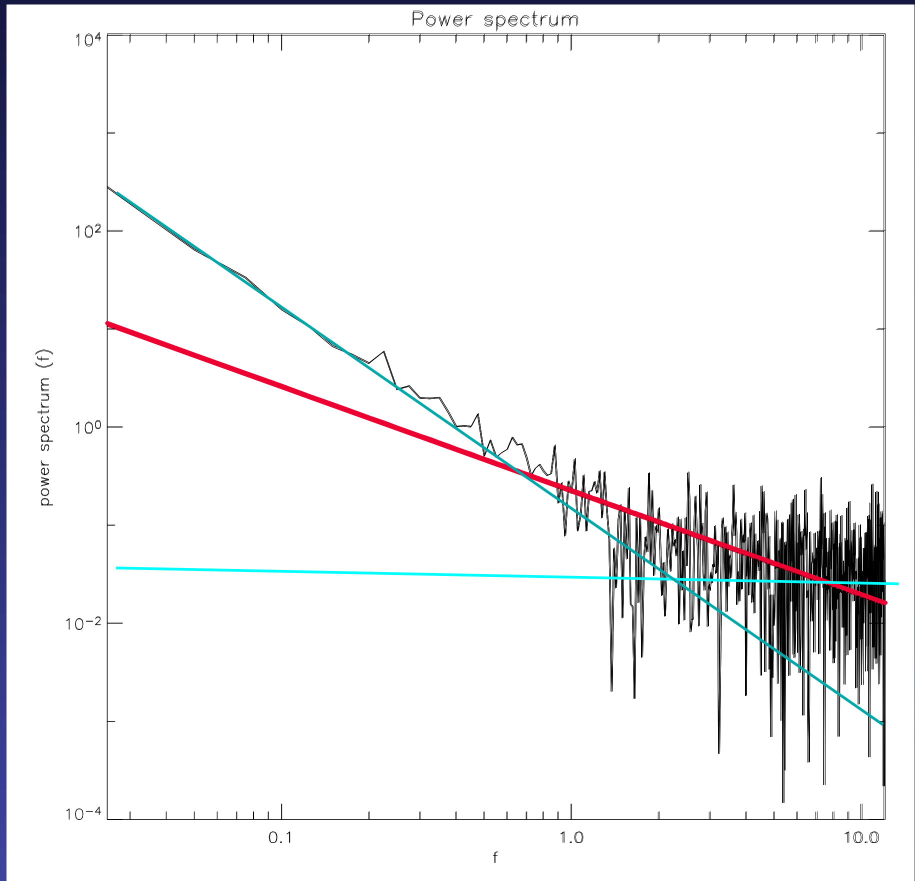
HAVE IN MIND THAT, USALLY, AN OBSERVED LIGHT CURVE IS ONLY ONE REALIZATION OF NOISE!



If the random noise process exhibits even a small amount of correlation in time, its properties deviate from Poissonian. Such random correlations will commonly be present over a wide range of frequencies, and will translate into a spectrum where the power is greater as lower frequencies.

In this case, the test statistic will not be distributed as it would be in the presence of white noise, and its PDF becomes in general not predictable.





It is the case of a gradient in the data, i.e. correlated noise

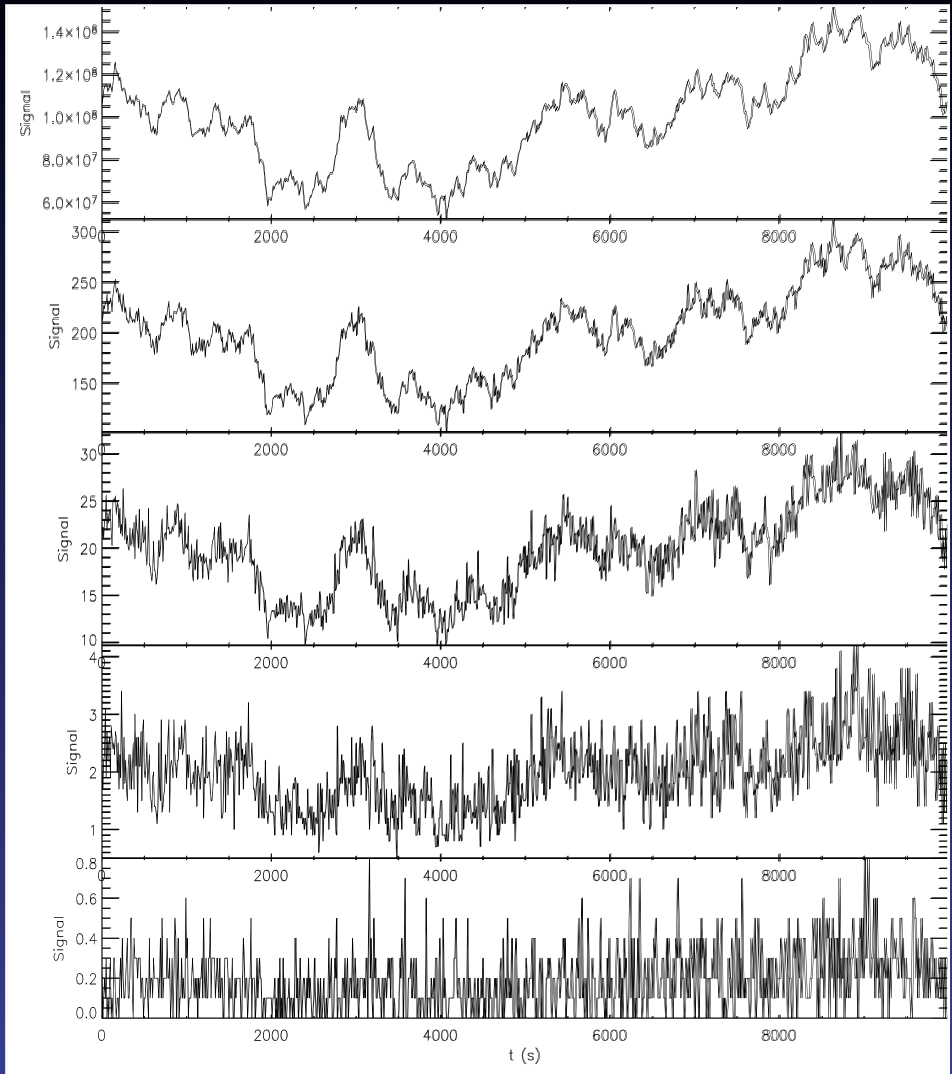
When the range of time-correlations in the noise is such that the power spectrum follows a power law with a negative spectral index, the noise process is referred to as *red noise* ($1/f^{\alpha}$)

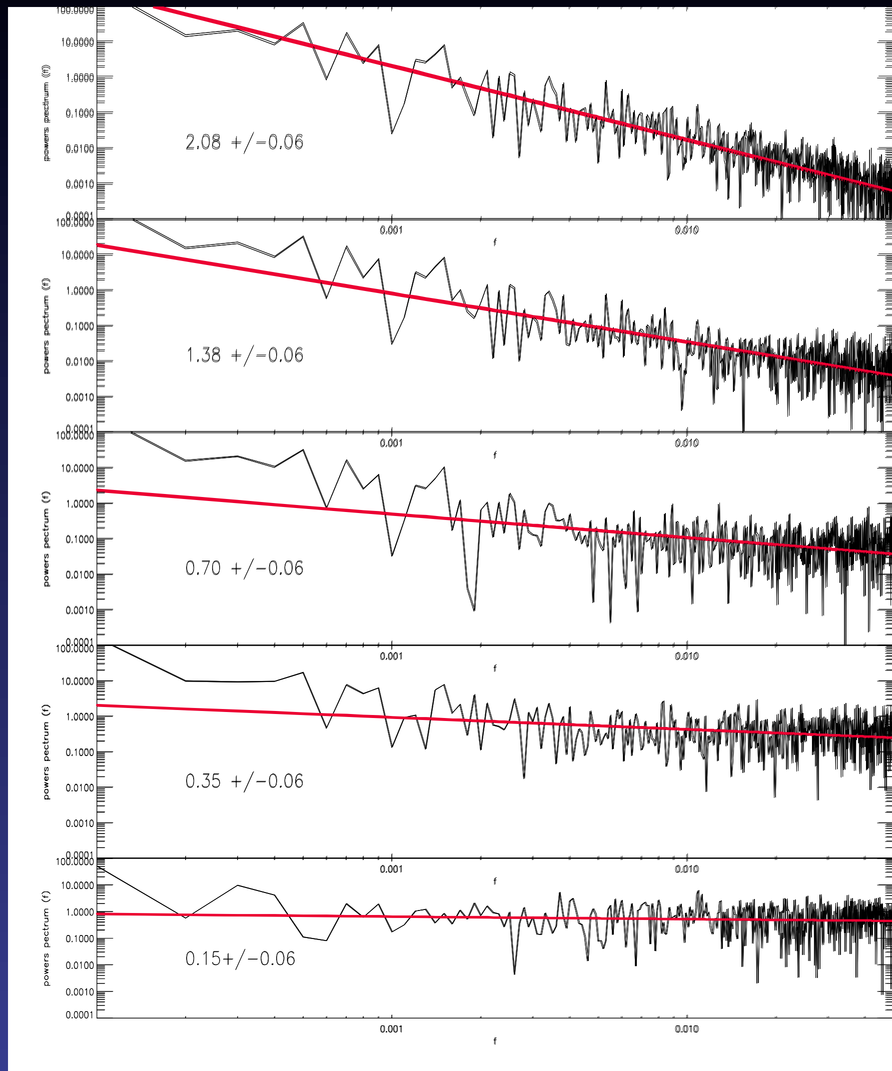
The general properties of a signal can be determined by looking at the overall shape of the distribution of powers as a function of frequency, i.e. we can always construct the power spectral density of a signal (or part of it) via FFT

Note also that a PDF from the FFT (Leahy normalized) of a white noise signal follows a χ^2. This is not the case for red noise!

The amount of red noise detected depends on several properties. The signal that you detect is what emitted from the source (possibly the same signal is correlated). Part of it travels for kpc in your direction and is colleced by your instrument (sensitivity). It is binned (integration time which naturally works against correlation). In case of a ground experiment, trends and systematics are always present (source correlation, change in air mass, telescope tracking, flat field errors..). Gaps in the data window (variability depends on the time slice → Excess Variance).

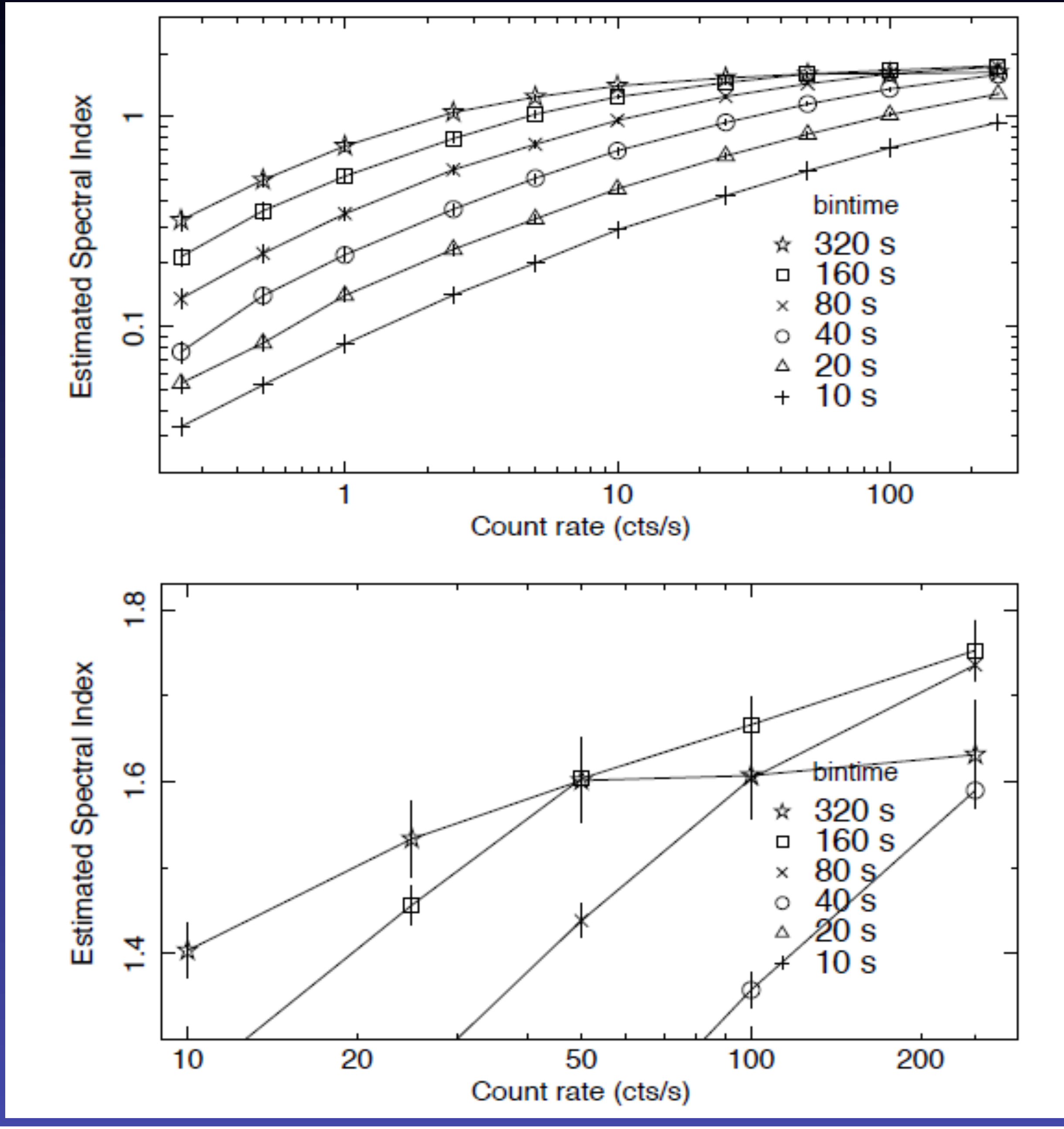
Red noise depends on the source flux (se e.g. Belanger et al., 2009)



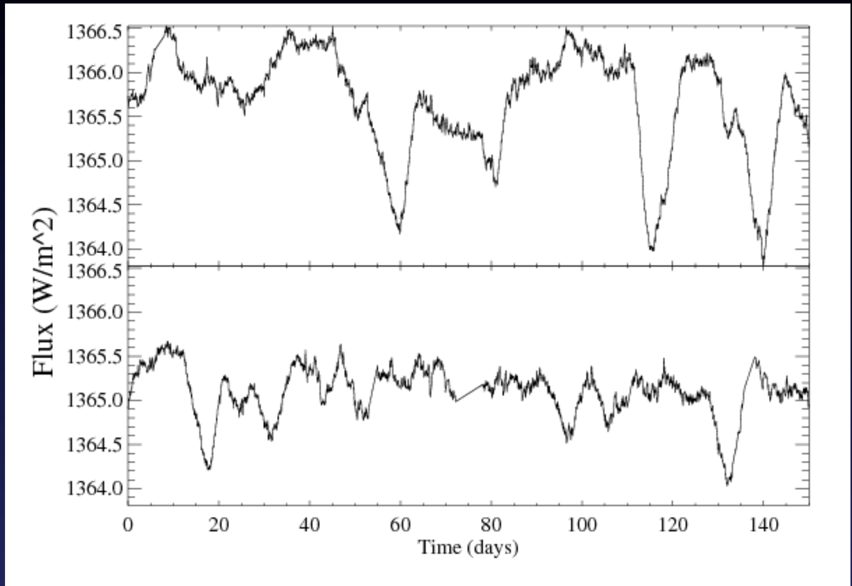


A time domain signal is degraded by the process of detecting a given number of events (flux) from the source. Each observed data point at time t is a realization of a signal (at the same time) extracted from a Poissonian distribution centered on a given mean. This degradation is reflected in our ability to estimate the spectral index.

Red noise depends on the binsize (integration time). Of course, if the used binsize is very large compared to the observation duration then any information is lost.



High timing resolution space based observations.



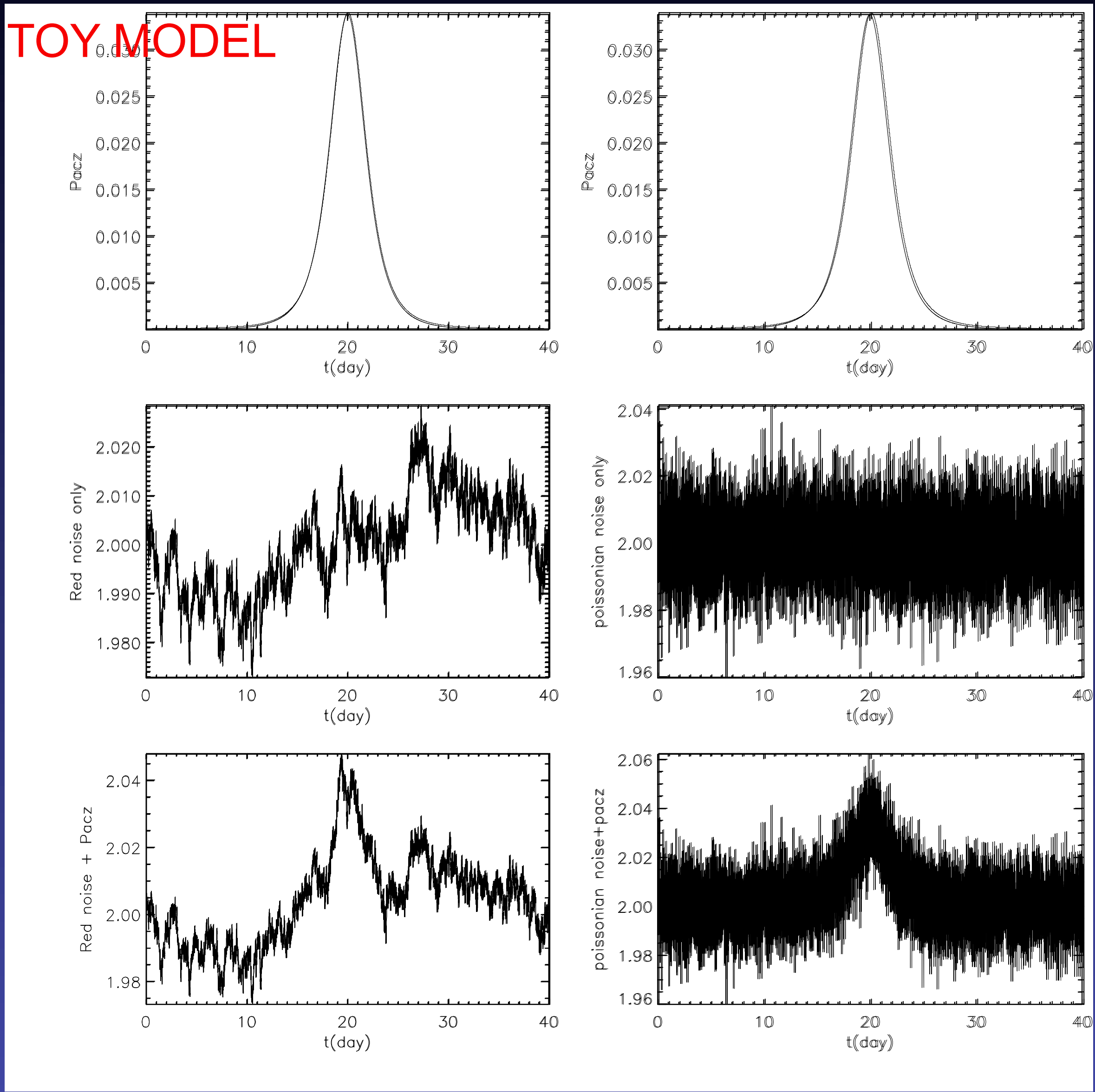
Virgo/PM06 data from the SoHO satellite from years 2000 (top) and 2006 (bottom)

Already considered in space based experiments searching for planetary transits (Smith et al., 2006 for a description of the SuperWasp collaboration and the implications of red noise for finding extra-solar planets, but see also Pont et al. MNRAS, 2006, 373, 231). WITHENING and DE-TRENDING introduce spurious effects, particular importan when one wants to estimate the transit parameters.

Corot, Kepler and the next generation satellite Plato.

We think that red noise is also important for lensing searches, particularly when we are dealing with low SN ratios, i.e. low amplifications. A careful analysis should be performed in order to disentangle between real events and noise artifacts in this regime.

As pointed out by Udalsky (this conference) low signal-to-noise microlensing (planetary) events could be missed from observations. This is particularly true for high sampled noise-coulored light curves (as observed from space)



A dedicated Monte Carlo study can help us in understanding this effect. The main questions are:

a) How the estimated event fit parameters are affected by red-noise?

b) Which is the minimum amplification that can be detected in case of red noise regime?

CoRoT satellite: 7689 monochromatic light curves of the LRc01. Binsize = 512 sec. Hot pixels and trend removed.

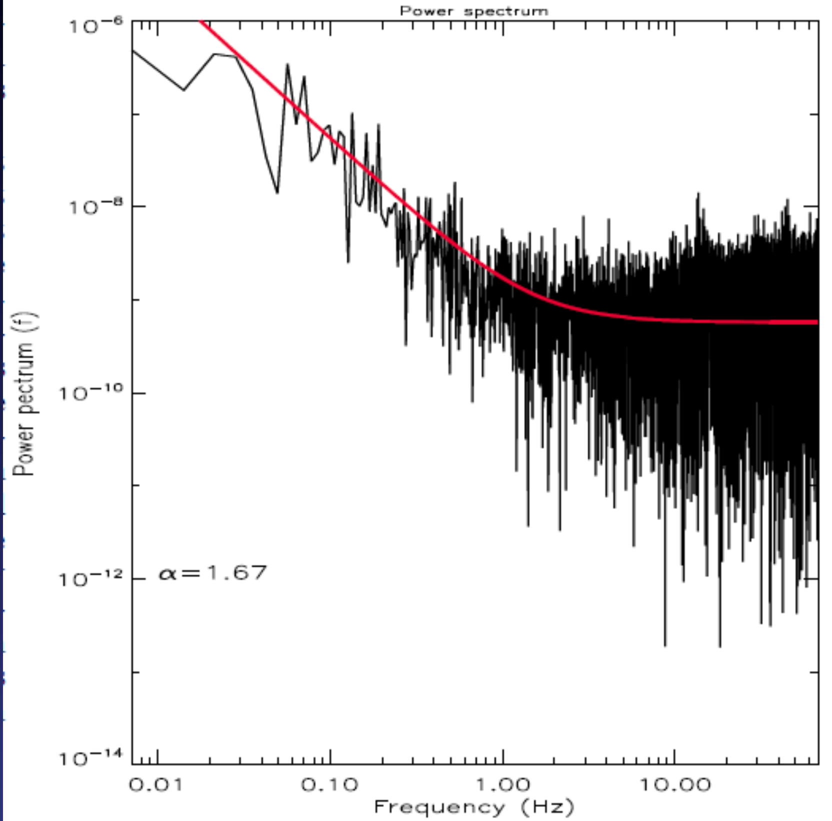


Fig. 1. Example of power spectrum of a randomly selected CoRoT light curve. The red line represents the best fit with the model described in the text.

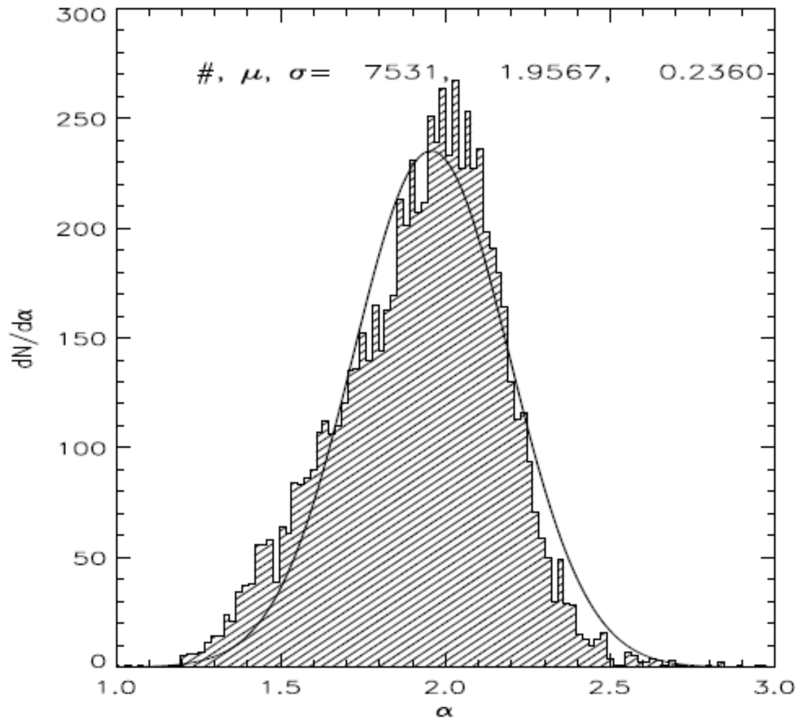
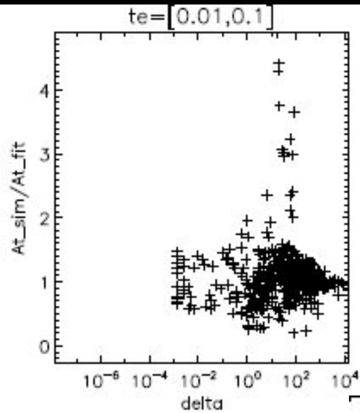
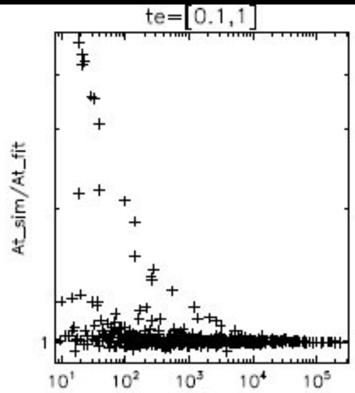


Fig. 2. Distribution of the spectral index α for the 7531 CoRoT light curves detected in the first Long Run towards the Galactic center (LRc01). The continuous line is the best fit curve obtained with a Gaussian function.

$$\mathcal{F}[x(t)](f) = \beta\ f^{-\alpha} + \gamma$$

1. $t_{\mathrm{E,sim}} = [0.01, 0.1]$
2. $t_{\mathrm{E,sim}} = [0.1, 1]$
3. $t_{\mathrm{E,sim}} = [1, 10]$
4. $t_{\mathrm{E,sim}} = [10, 100]$





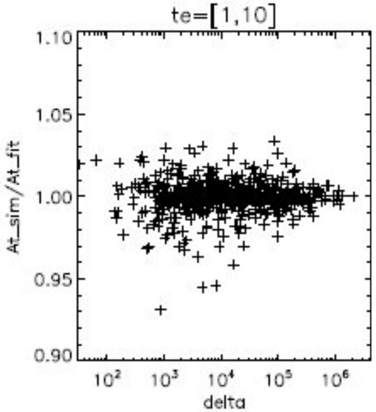
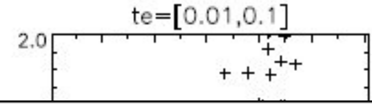
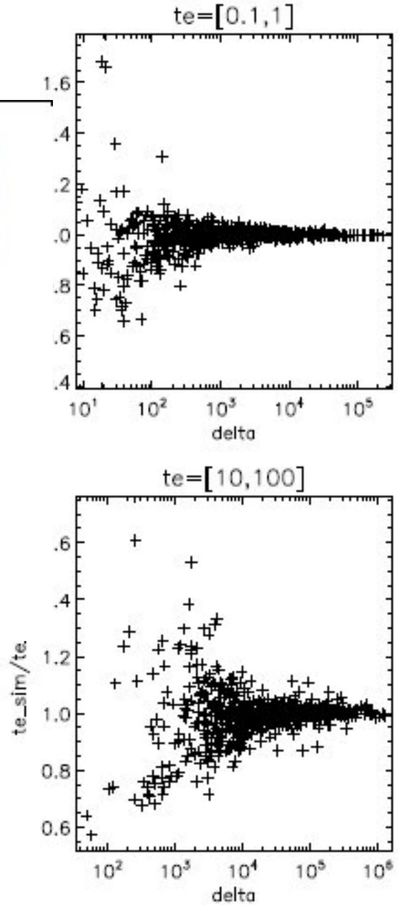


Fig. 5. From left-upper pane... for the values of the simula... fication versus Δ have been reported for simulated parameters ranges, respectively, in $\{t_E = [0.01, 0.1], u_0 = [0.01, 1]\}$, $\{t_E = [0.1, 1], u_0 = [0.01, 1]\}$, $\{t_E = [1, 10], u_0 = [0.01, 1]\}$, $\{t_E = [10, 100], u_0 = [0.01, 1]\}$.

Table 1. The obtained standard deviations σ for the distribution of $t_{\mathrm{E,sim}}/t_{\mathrm{E,fit}}$ and $A_{\mathrm{sim}}/A_{\mathrm{fit}}$ are reported. Each column shows the value of standard deviation obtained for $t_{\mathrm{E,sim}}$ in the given range (in days). In all simulation $u_{0,\mathrm{sim}}$ was randomly extracted in the range $[0.01, 1]$.

$t_{\mathrm{E,sim}}$	$[0.01, 0.1]$	$[0.1, 1]$	$[1, 10]$	$[10, 100]$
$\sigma(t_{\mathrm{E,sim}}/t_{\mathrm{E,fit}})$	0.1190	0.0124	0.0108	0.0191
$\sigma(A_{\mathrm{sim}}/A_{\mathrm{fit}})$	0.2030	0.0035	0.0041	0.0123





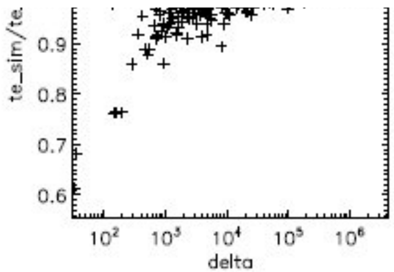


Fig. 6. From left-upper panel, moving clockwise, scatter plots for the values of the simulated over fitted t_E versus Δ have been reported for simulated parameters ranges, respectively, in $\{t_E = [0.01, 0.1], u_0 = [0.01, 1]\}$, $\{t_E = [0.1, 1], u_0 = [0.01, 1]\}$, $\{t_E = [1, 10], u_0 = [0.01, 1]\}$, $\{t_E = [10, 100], u_0 = [0.01, 1]\}$.

$$\Delta = \frac{\chi^2_{\mathrm{cst}} - \chi^2_{\mathrm{ml}}}{\chi^2_{\mathrm{ml}}/N_{\mathrm{dof}}} \frac{1}{\sqrt{2N_{\mathrm{dof}}}}$$

Correlated time series exist, and in their presence a careful analysis has to be performed in order to disentangle real features (planetary transits, periodicities, micro-lensing planetary signatures) from noise artefacts.

The magic receipt could be…MC simulations, e.g.

a) Find the typical noise power index of the time series
b) Simulate correlated time series (no signal in it) as much you can (use, e.g. Vaughan's prescription)
c) Use the same algorithm used in extracting the feature in which you are interesting on each of the simulated lcs.
d) Calculate the number of false-positive of getting a given feature and associate it a probability.