TIME-DOMAIN ASTRONOMY

Lectures 4: Lomb-Scargle Periodogram

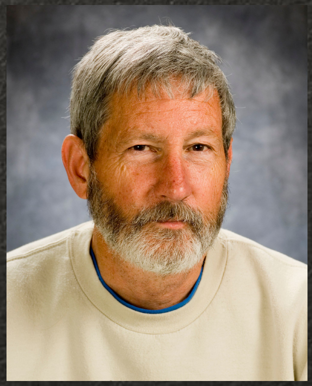
Stefano Covino
INAF / Brera Astronomical Observatory

Unevenly Sampled Data

- Harmonic analysis of unevenly spaced data is problematic due to the loss of information and increase in aliasing.
- A solution to this problem was proposed in astronomy and it is now widely used and known as “Lomb-Scargle periodogram”.
- We will see that in case of regular sampling it is equivalent to the Fourier periodogram.

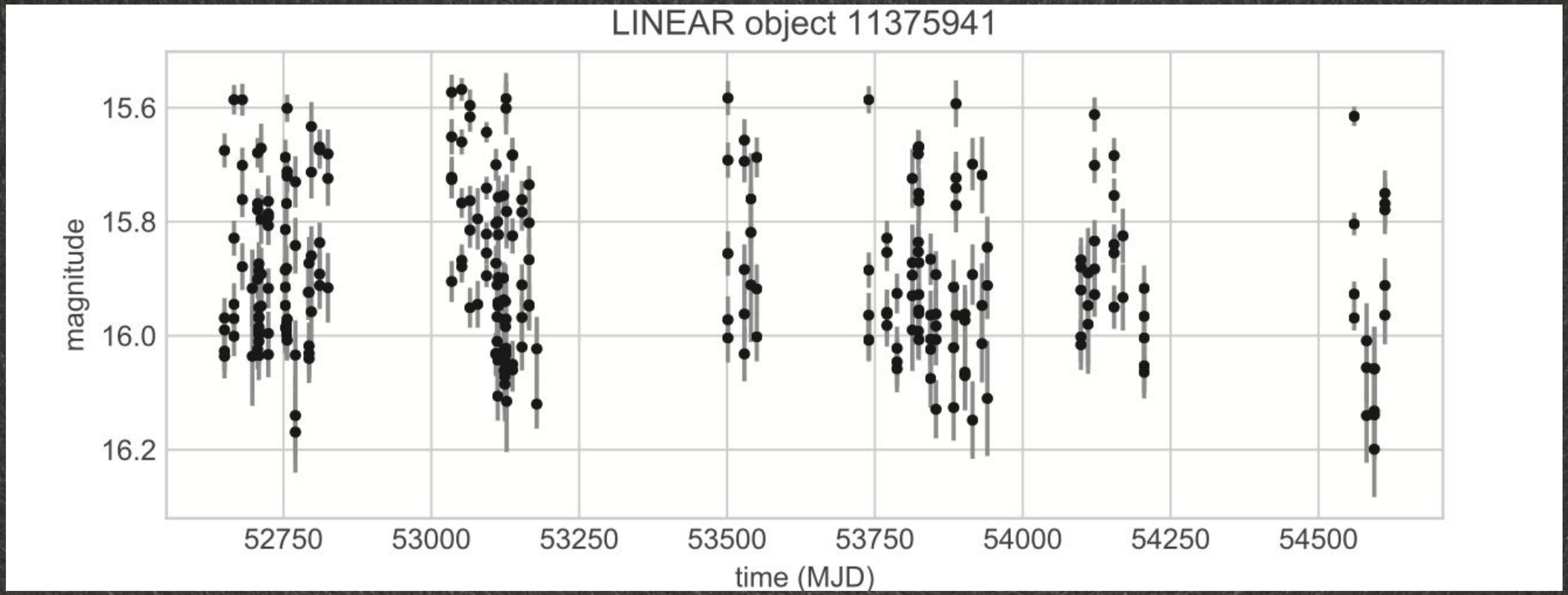
Nick Lomb

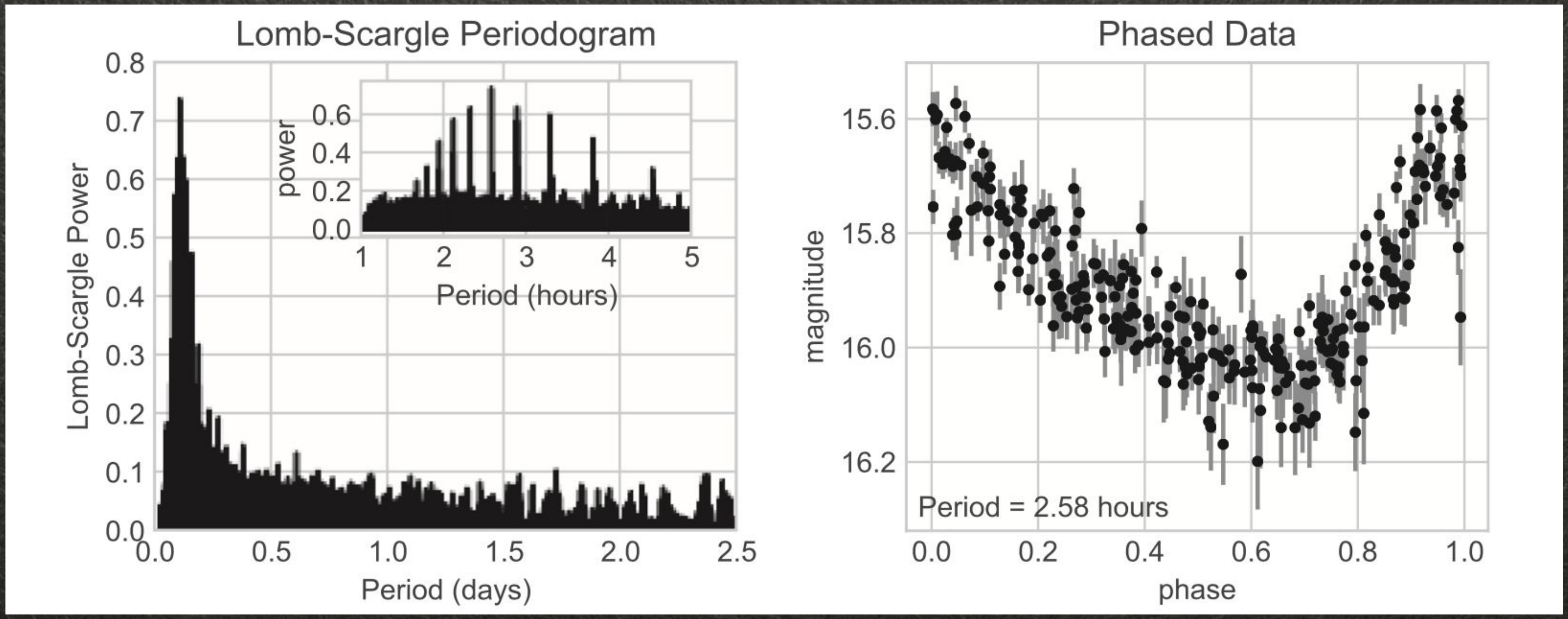
Jeffrey Scargle



Example

- Optical observation of the LINEAR object 11375941





Many possible techniques

- Fourier methods:
  - are based on the Fourier transform, power spectra and correlation functions (Schuster periodogram, Lomb-Scargle periodogram, correlation-based methods, wavelet transform, etc.)
- Phase-folding methods:
  - depend on folding observations as a function of phase (String Length, Analysis of Variance, Phase Dispersion Minimization, Gregory & Loredo method, conditional entropy method, correntropy methods, etc.)
- Least-squares methods:
  - involve fitting a model to the data at each candidate frequency (Lomb-Scargle periodogram, Supersmoother approach, orthogonal polynomial fits, etc.)
- Bayesian approaches:
  - apply Bayesian probability theory to the problem (Lomb-Scargle generalisation, Gregory & Loredo, Gaussian process models, stochastic process models, etc.)

A reminder of CFT

- given a continuous signal g(t) the Fourier transform and its inverse are:

$$\hat{g}(f) \equiv \int_{-\infty}^{\infty} g(t) e^{-2\pi i f t} dt.$$

$$g(t) \equiv \int_{-\infty}^{\infty} \hat{g}(f) e^{+2\pi i f t} df.$$

- We can define the Fourier transform operator:

$$\mathcal{F}\{g\} = \hat{g},$$
$$\mathcal{F}^{-1}\{\hat{g}\} = g.$$

- g and ĝ are know as a Fourier pair: $g \Longleftrightarrow \hat{g}.$

FT properties

- FT is a linear operator:

$$\mathcal{F}\{f(t) + g(t)\} = \mathcal{F}\{f(t)\} + \mathcal{F}\{g(t)\}$$
$$\mathcal{F}\{Af(t)\} = A\mathcal{F}\{f(t)\}.$$

- The FT of a sinusoid with frequency f_0 is a sum of delta functions at $\pm f_0$ $\left(\delta(f) \equiv \int_{-\infty}^{+\infty} e^{-2\pi i t f}\, dt\right)$:

$$\mathcal{F}\{e^{2\pi f_0 t}\} = \delta(f - f_0).$$

$$\mathcal{F}\{\cos(2\pi f_0 t)\} = \frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)]$$
$$\mathcal{F}\{\sin(2\pi f_0 t)\} = \frac{1}{2i}[\delta(f - f_0) - \delta(f + f_0)].$$

From Euler's formula: $e^{\mathrm{ix}} = \cos(x) + i\sin(x)$

FT properties

- A time shift imparts a phase in the FT:

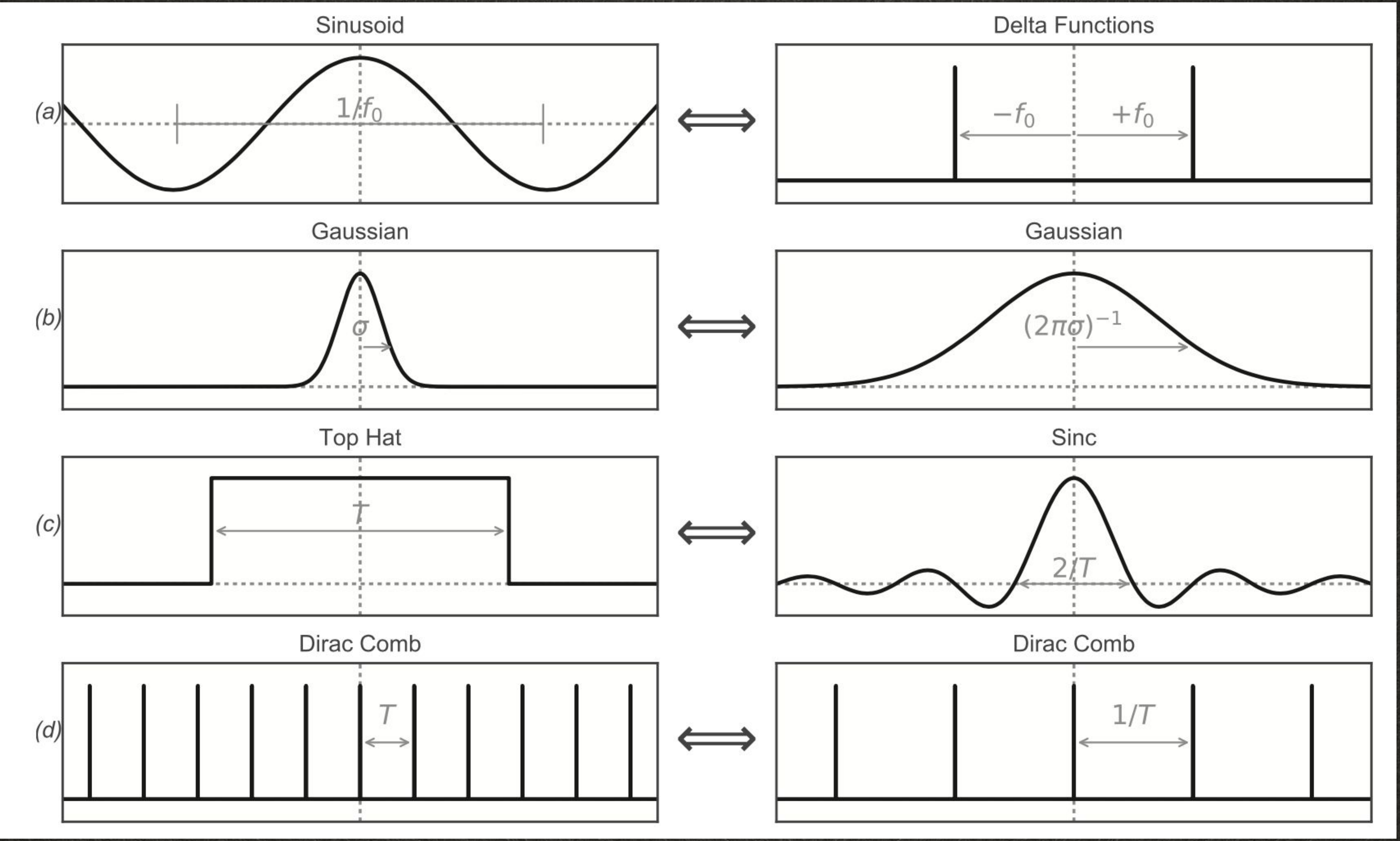
$$\mathcal{F}\{g(t-t_0)\} = \mathcal{F}\{g(t)\}e^{-2\pi i f t_0}.$$

- As we know, the squared amplitude of the FT of a continuous signal is known as the power spectral density (PSD):

$$\mathcal{P}_g \equiv |\mathcal{F}\{g\}|^2.$$

- if g is real-valued $P_g(f) = P_g(-f)$

Useful FT pairs



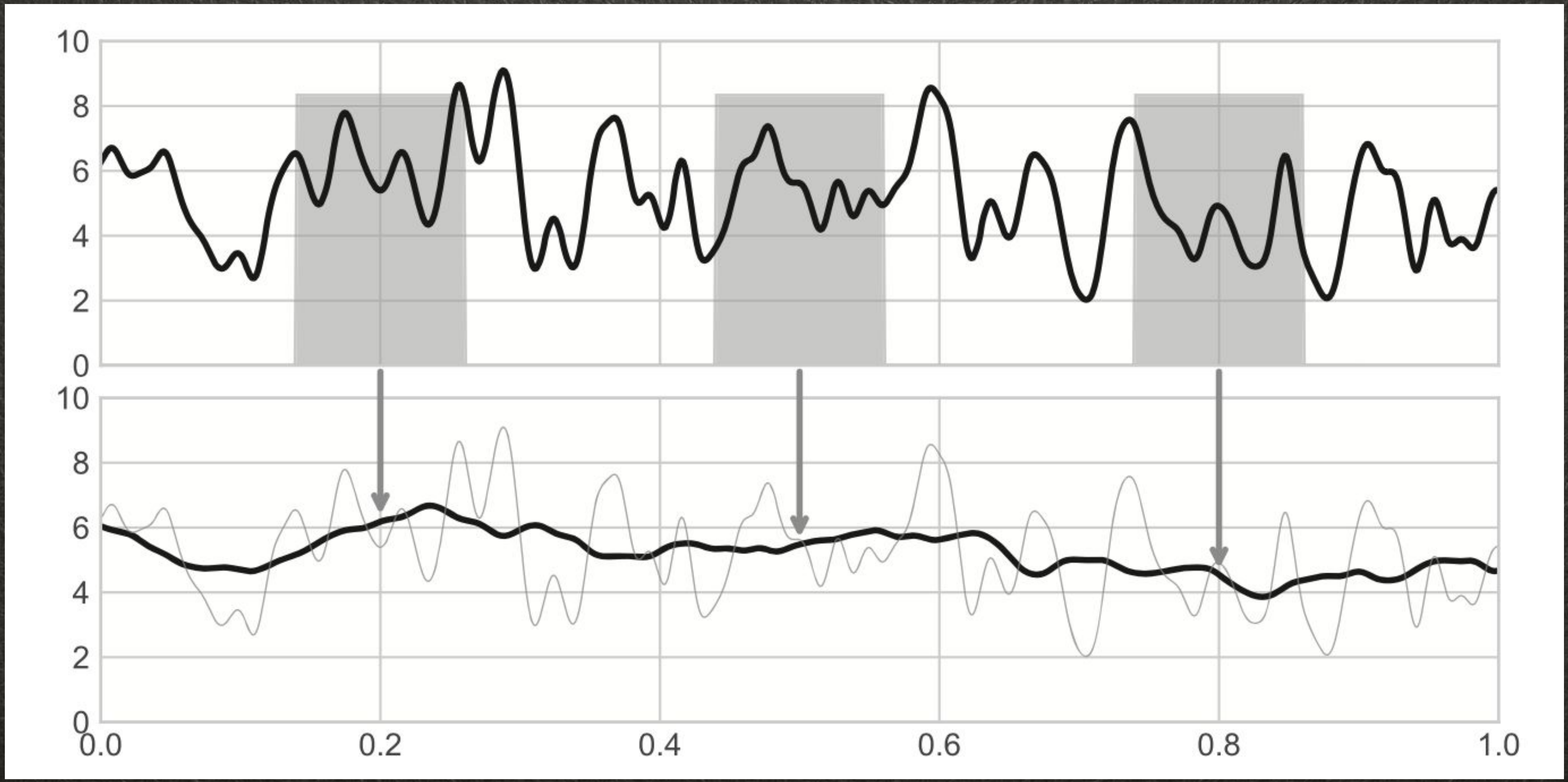
Convolution

- We know already these properties:

$$[f * g](t) \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau.$$

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}.$$

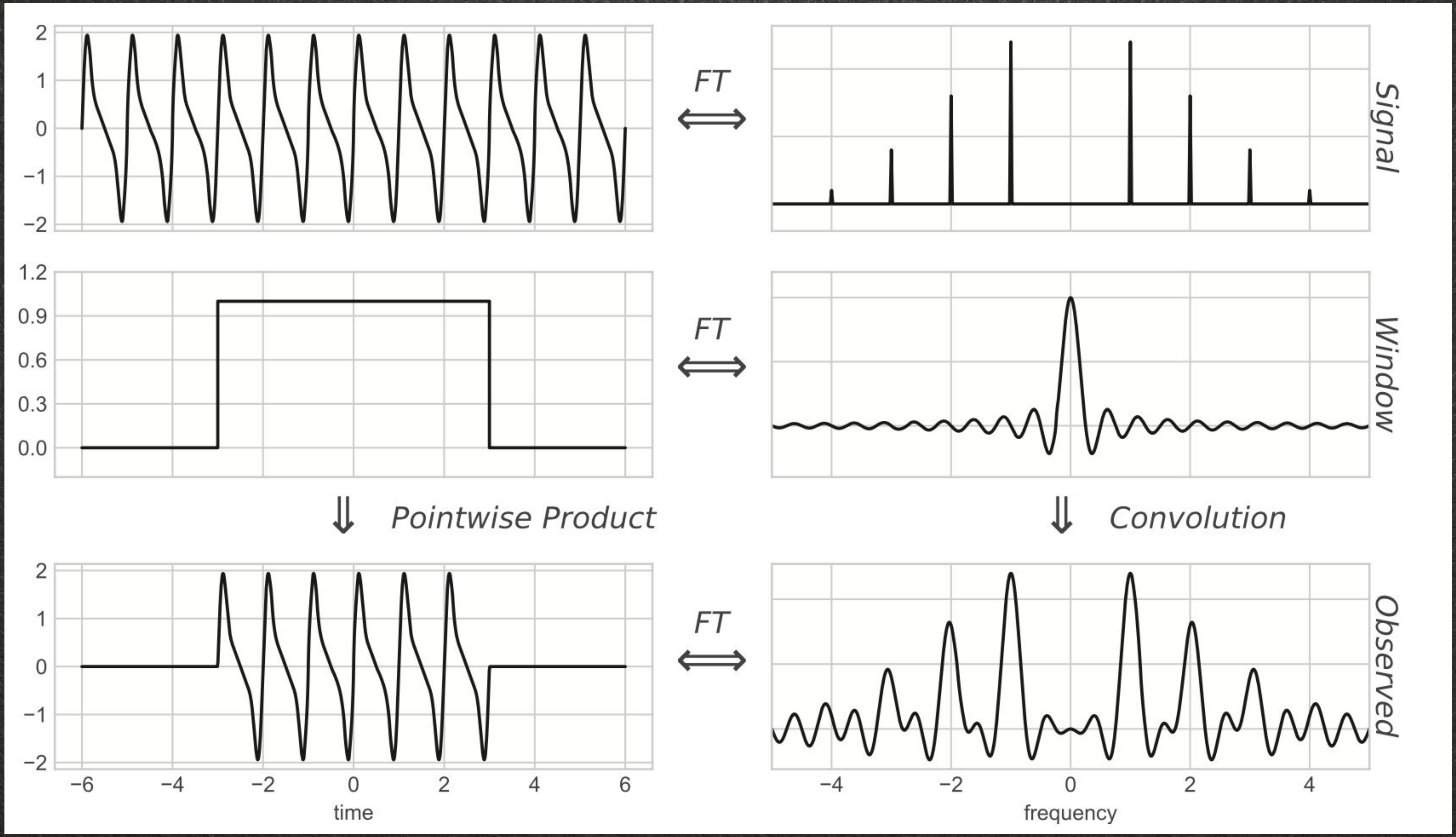
$$\mathcal{F}\{f \cdot g\} = \mathcal{F}\{f\} * \mathcal{F}\{g\}.$$



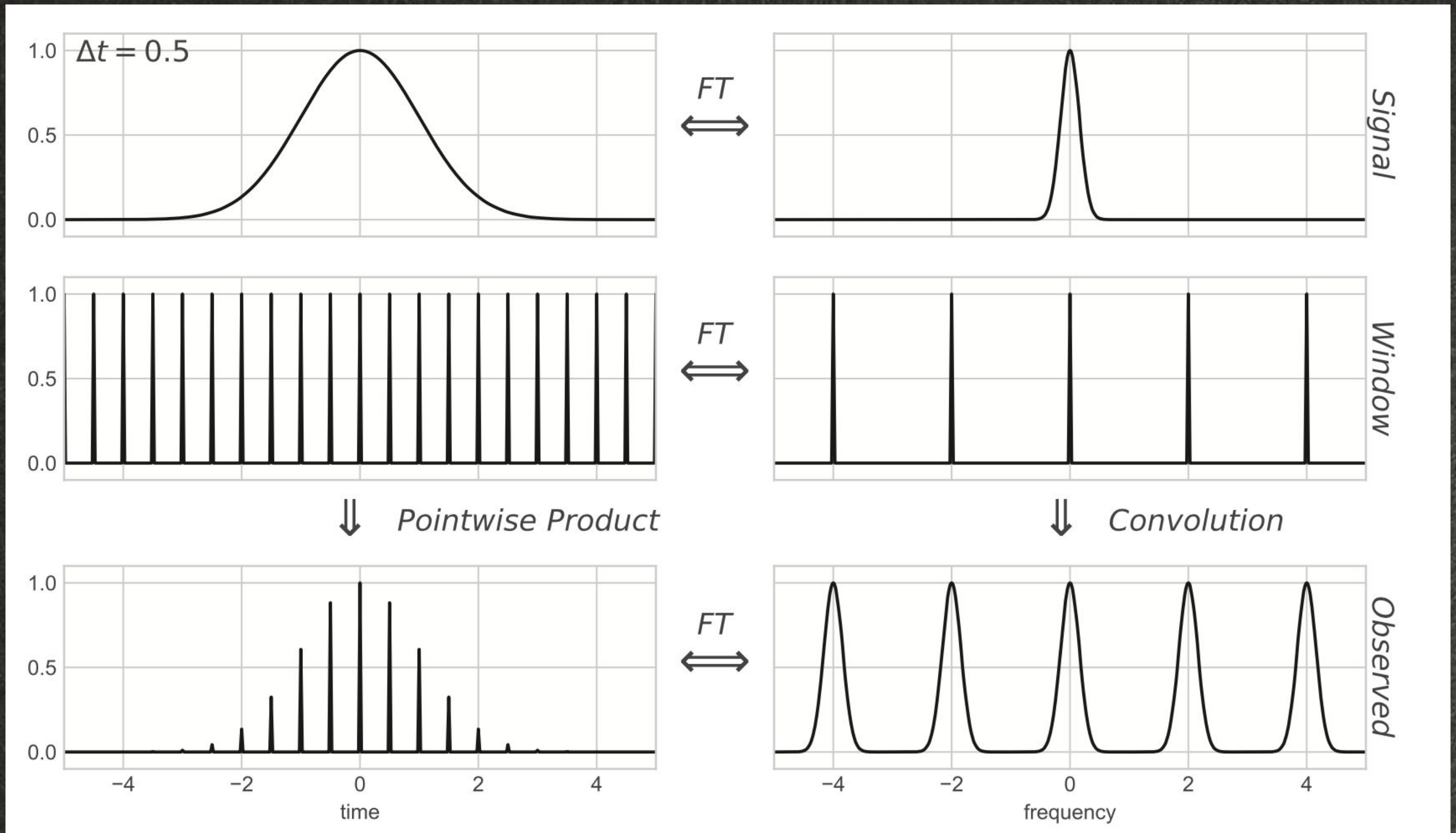
Observing Windows

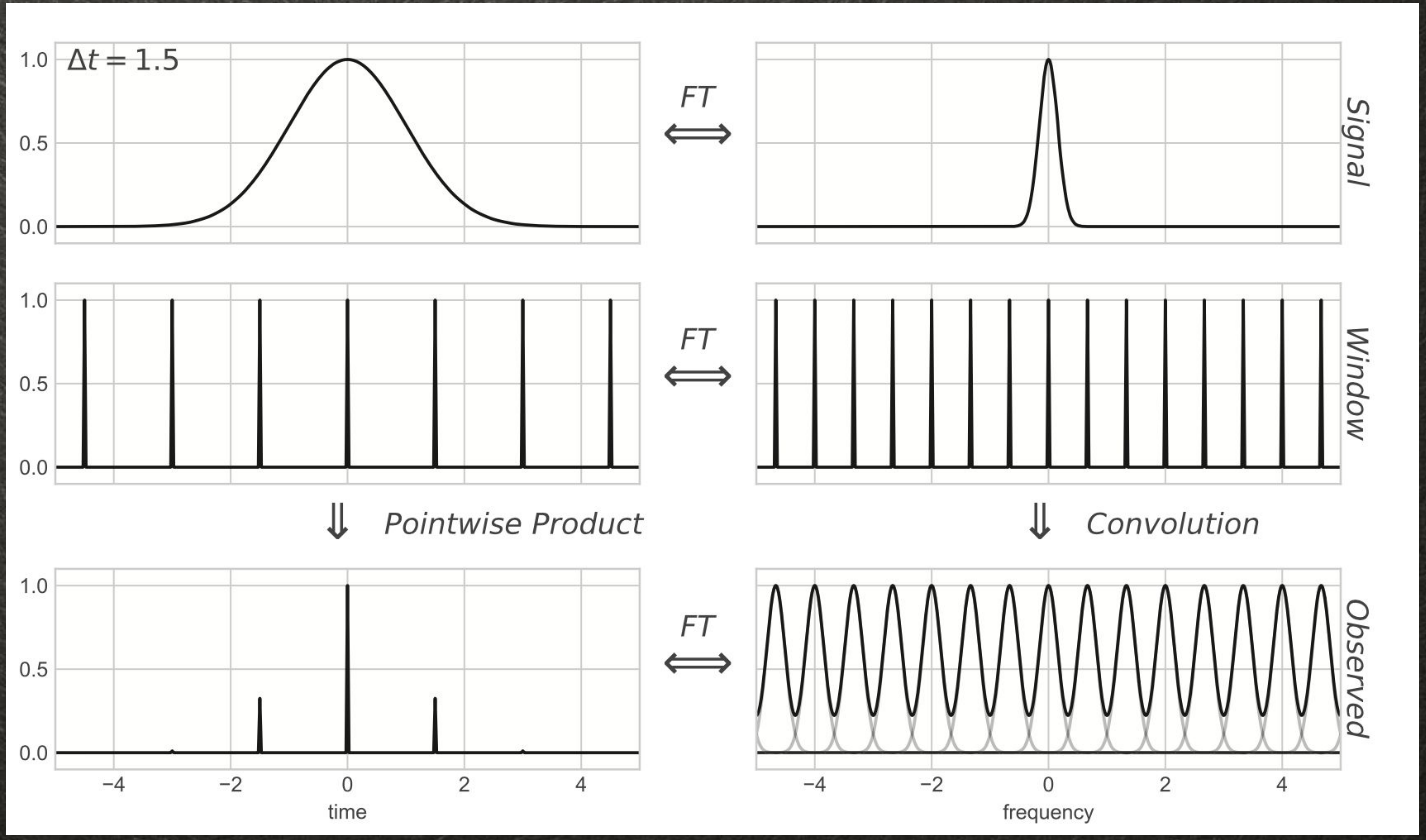
$$g_{\rm obs}(t) = g(t)W(t),$$

$$\mathcal{F}\{g_{\rm obs}\} = \mathcal{F}\{g\} * \mathcal{F}\{W\}.$$



Sampling Rates





The DFT

- Given and infinitely long and continuous signal g(t) observed on a regular grid with spacing Δt we have:

$$g_{\mathrm{obs}} = g(t)\mathrm{III}_{\Delta t}(t)$$

$$\hat{g}_{\mathrm{obs}}(f) = \sum_{n=-\infty}^{\infty} g(n\Delta t)e^{-2\pi i f n \Delta t},$$

- In the real world, we never have an infinite number of points... and defining $g_n = g(n\Delta t)$ we have:

$$\hat{g}_{\mathrm{obs}}(f) = \sum_{n=0}^{N} g_n e^{-2\pi i f n \Delta t}.$$

- or, with $\Delta f = 1/(N\Delta t)$ and $\hat{g}_k \equiv \hat{g}_{\mathrm{obs}}(k\Delta f)$:

$$\hat{g}_k = \sum_{n=0}^{N} g_n e^{-2\pi i k n/N},$$

The Classical Periodogram

- Applying the definition of Fourier spectrum $P_g \equiv |F\{g\}|^2$ we then have:

$$P_S(f) = \frac{1}{N} \left| \sum_{n=1}^{N} g_n e^{-2\pi i f t_n} \right|^2 .$$

- This is the periodogram, which is an "estimator" of the true power spectrum of the underlying continuum function g(t).
- As we know, it is not a consistent estimator since it suffers from intrinsic variance even in the limit of an infinite number of observations.

Nonuniform sampling

- As in the uniform sampling case we can define an observing window as:

$$W_{\{t_n\}}(t) = \sum_{n=1}^{N} \delta(t - t_n).$$

- that applied to a continuous signal g(t) gives:

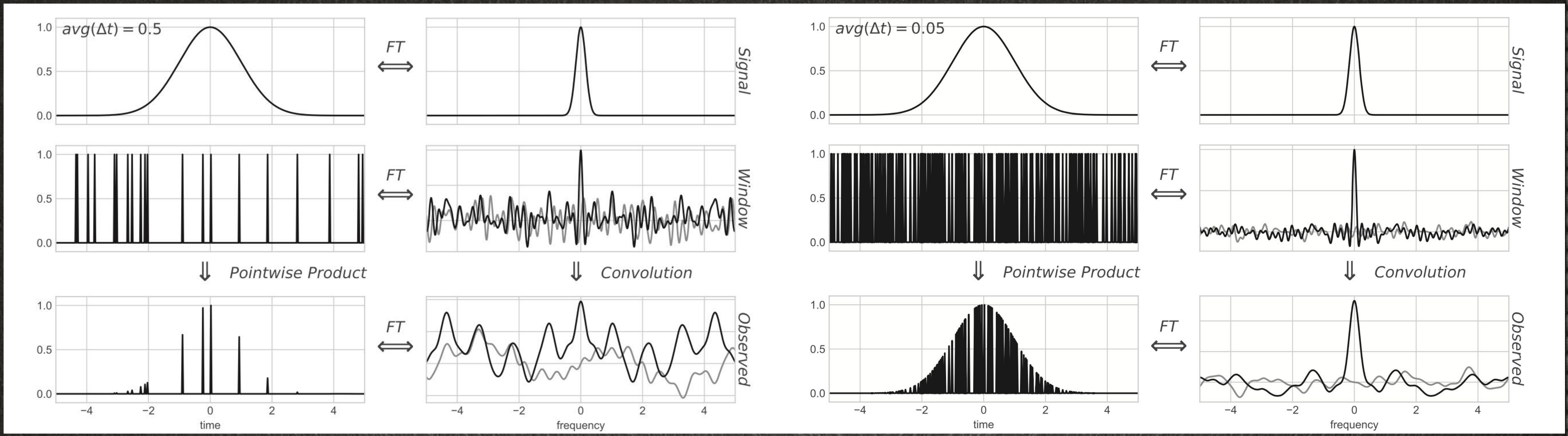
$$\begin{aligned} g_{\text{obs}}(t) &= g(t) W_{\{t_n\}}(t) \\ &= \sum_{n=1}^{N} g(t_n) \delta(t - t_n). \end{aligned}$$

- with FT:

$$\mathcal{F}\{g_{\text{obs}}\} = \mathcal{F}\{g\} * \mathcal{F}\{W_{\{t_n\}}\}.$$

- However, now the window transform will not be in general a simple symmetric function.

Nonuniform sampling



- The FT of a non-uniformly spaced delta functions looks like random noise, and partly it is.
- It reflects the "random" distribution of the sampling time.
- A denser sampling helps, but to some extent "noise" is unavoidable.

No Nyquist limit!

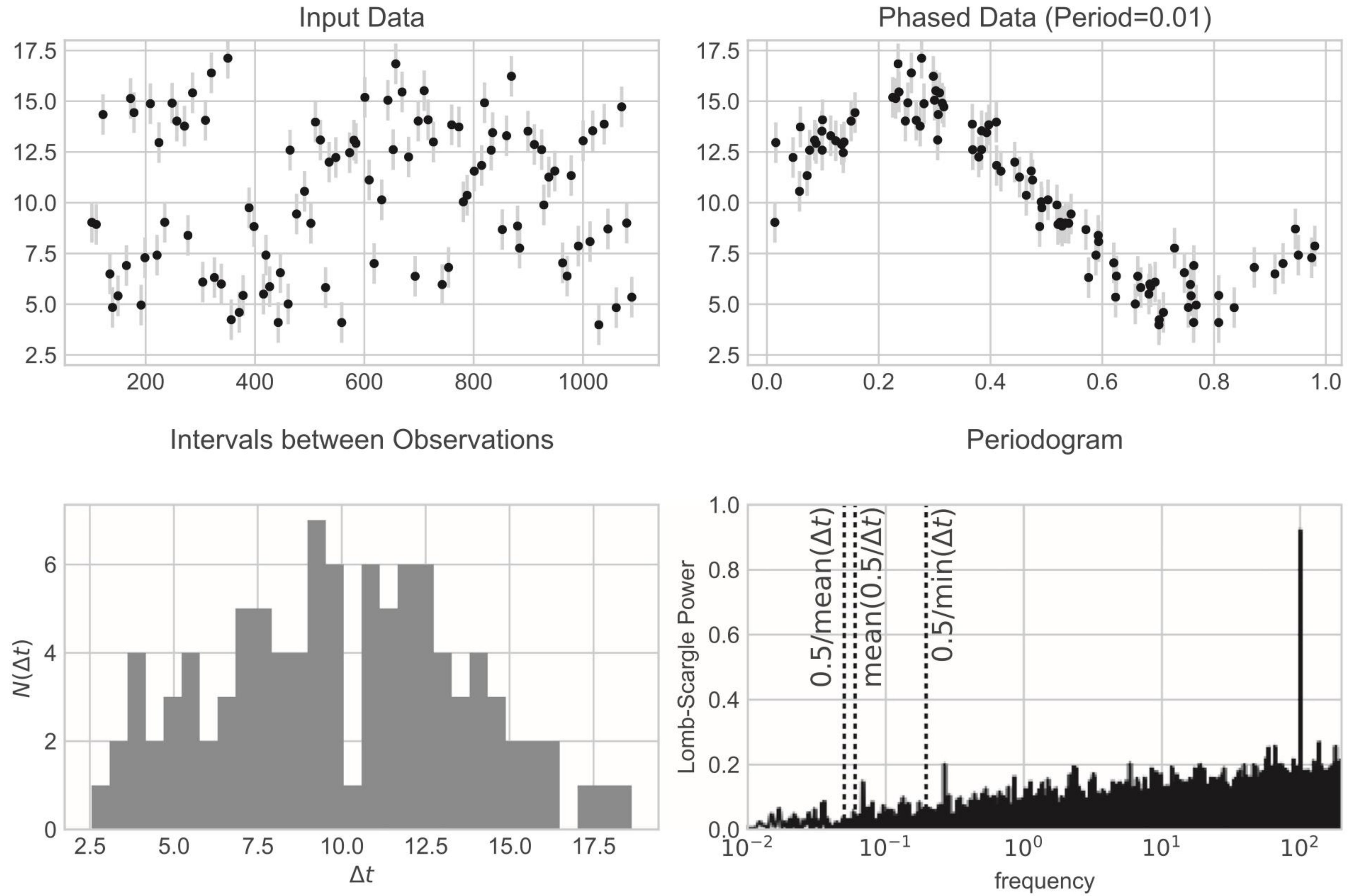


Figure 11. Example of data for which the various poorly motivated "pseudo-Nyquist" approaches outlined in Section 4.1 fail spectacularly. The top panels show the data, a noisy sinusoid with a frequency of 100 (i.e., a period of 0.01). The bottom left panel shows a histogram of spacings between observations: the minimum spacing is 2.55, meaning that the signal has *over 250 full cycles* between the closest pair of observations. Nevertheless, the periodogram (bottom right) clearly identifies the correct period, though it is orders of magnitude larger than pseudo-Nyquist estimates based on an average or minimum sampling rate.

The Lomb-Scargle Periodogram

- The classical periodogram can be rewritten as:

$$P(f) = \frac{1}{N} \left| \sum_{n=1}^{N} g_n e^{-2\pi i f t_n} \right|^2$$
$$= \frac{1}{N} \left[\left(\sum_n g_n \cos(2\pi f t_n) \right)^2 + \left(\sum_n g_n \sin(2\pi f t_n) \right)^2 \right].$$

- In principle this formula could be used for non-uniform sampling too, yet the obtained periodogram does not offer, in general, most of the useful statistical properties of the even sampling case.

The Lomb-Scargle Periodogram

- The problem was addressed by proposing a generalized expression:

$$P(f) = \frac{1}{N}\left|\sum_{n=1}^{N} g_n e^{-2\pi i f t_n}\right|^2 = \frac{1}{N}\left[\left(\sum_n g_n \cos(2\pi f t_n)\right)^2 + \left(\sum_n g_n \sin(2\pi f t_n)\right)^2\right].$$

$$P(f) = \frac{A^2}{2}\left(\sum_n g_n \cos(2\pi f\,[t_n - \tau])\right)^2 + \frac{B^2}{2}\left(\sum_n g_n \sin(2\pi f\,[t_n - \tau])\right)^2,$$

- where A, B and τ are arbitrary functions of frequencies and observing times.

- A, B and τ are found so that:
  - the periodogram reduces to the classical form in the case of equally spaced observations;
  - the periodogram distribution is analytically computable;
  - the periodogram is insensitive to global time shifts in the data.

The Lomb-Scargle Periodogram

- A, B and τ satisfying the previously mentioned requirements are:

$$P_{\mathrm{LS}}(f) = \frac{1}{2}\left\{\left(\sum_n g_n \cos(2\pi f\,[t_n - \tau])\right)^2 \Big/ \sum_n \cos^2(2\pi f\,[t_n - \tau]) + \left(\sum_n g_n \sin(2\pi f\,[t_n - \tau])\right)^2 \Big/ \sum_n \sin^2(2\pi f\,[t_n - \tau])\right\},$$

$$\tau = \frac{1}{4\pi f}\tan^{-1}\left(\frac{\sum_n \sin(4\pi f t_n)}{\sum_n \cos(4\pi f t_n)}\right).$$

The Lomb-Scargle Periodogram

- The LS and Schuster periodogram are different to the extent that the denominators differ from N/2, the expected value in case of even sampling.

$$P_{\rm LS}(f) = \frac{1}{2}\left\{\left(\sum_n g_n \cos(2\pi f\,[t_n - \tau])\right)^2 \Big/ \sum_n \cos^2(2\pi f\,[t_n - \tau]) + \left(\sum_n g_n \sin(2\pi f\,[t_n - \tau])\right)^2 \Big/ \sum_n \sin^2(2\pi f\,[t_n - \tau])\right\},$$

$$\tau = \frac{1}{4\pi f}\tan^{-1}\left(\frac{\sum_n \sin(4\pi f t_n)}{\sum_n \cos(4\pi f t_n)}\right).$$

$$P(f) = \frac{1}{N}\left|\sum_{n=1}^{N} g_n e^{-2\pi i f t_n}\right|^2 = \frac{1}{N}\left[\left(\sum_n g_n \cos(2\pi f t_n)\right)^2 + \left(\sum_n g_n \sin(2\pi f t_n)\right)^2\right].$$

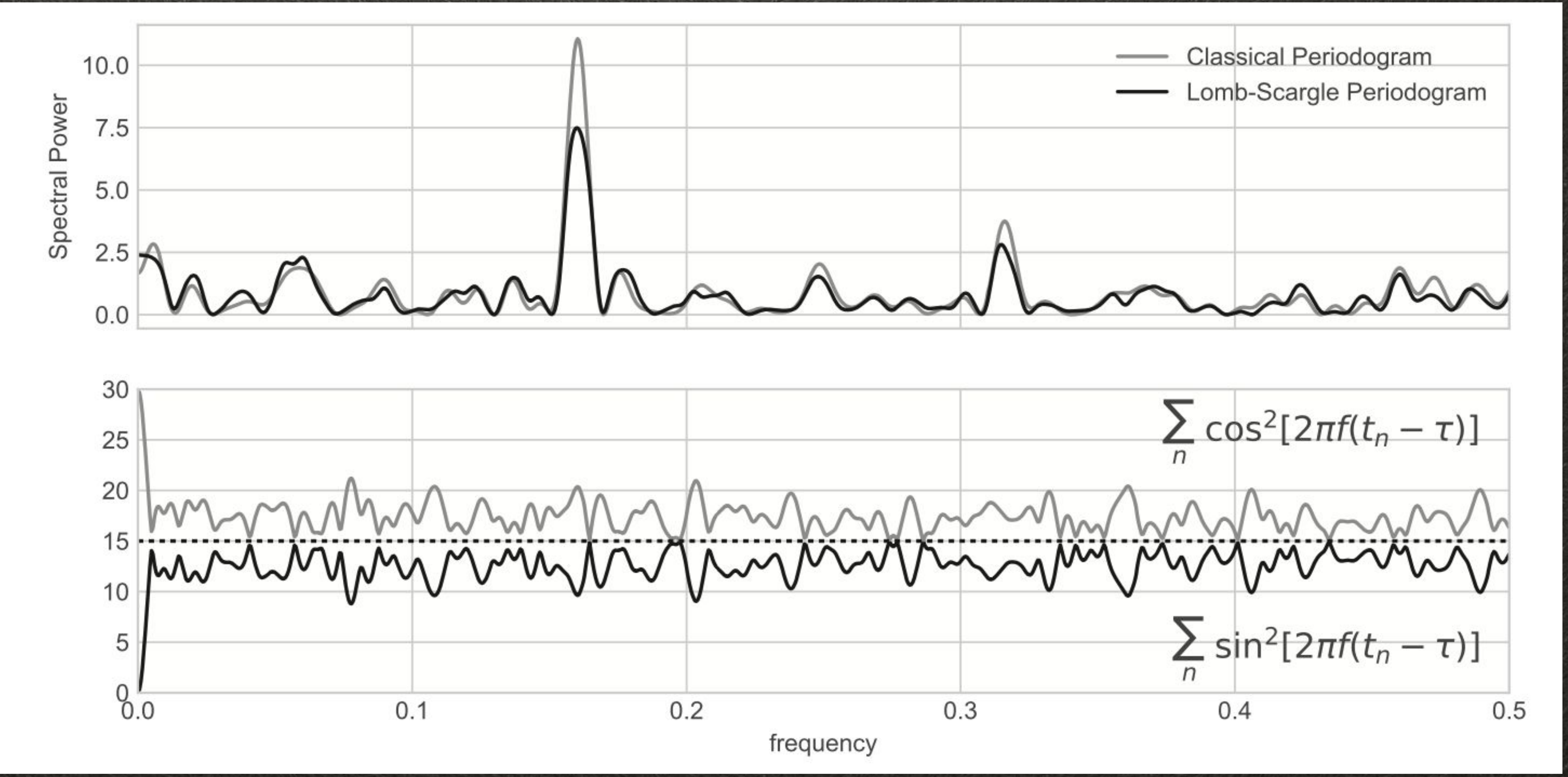
- An important result is that the LS periodogram could be obtained fitting a simple sinusoidal model to the data at each given frequency, and deriving the periodogram from the goodness of fit (χ^2).

- In any case several of the various considerations derived for the Fourier periodograms in general holds for the LS too.

Exercises

- Useful notebooks:

1. LombScargleVsClassical



LS Periodogram Extensions

- Considering the LS formula as the result of a regular fitting procedure allows one for interesting generalizations.

$$y(t; f) = A_f \sin(2\pi f (t - \phi_f)),$$

- where A_f and φ_f can vary as a function of frequency.

- The model is fit to the data minimizing the χ^2 wrt A_f and φ_f:

$$\chi^2(f) \equiv \sum_n (y_n - y(t_n; f))^2.$$

- and it can be also be shown that the LS periodogram is equivalent to:

$$P(f) = \frac{1}{2}[\hat{\chi}_0^2 - \hat{\chi}^2(f)],$$

LS Periodogram Extensions

- Basing on this view a trivial yet fundamental extension is to include in the fitting procedure the (Gaussian) errors on the data:

$$\chi^2(f) \equiv \sum_n \left(\frac{y_n - y_{\mathrm{model}}(t_n; f)}{\sigma_n} \right)^2.$$

- and it is also possible to consider more complex noise structure (e.g. correlation) in the data. If Σ is the noise covariance matrix:

$$\boldsymbol{y} = [y_1, y_2, \cdots y_n]^T$$
$$\boldsymbol{y}_{\mathrm{model}} = [y_{\mathrm{model}}(t_1), y_{\mathrm{model}}(t_1), \cdots y_{\mathrm{model}}(t_n)]^T,$$

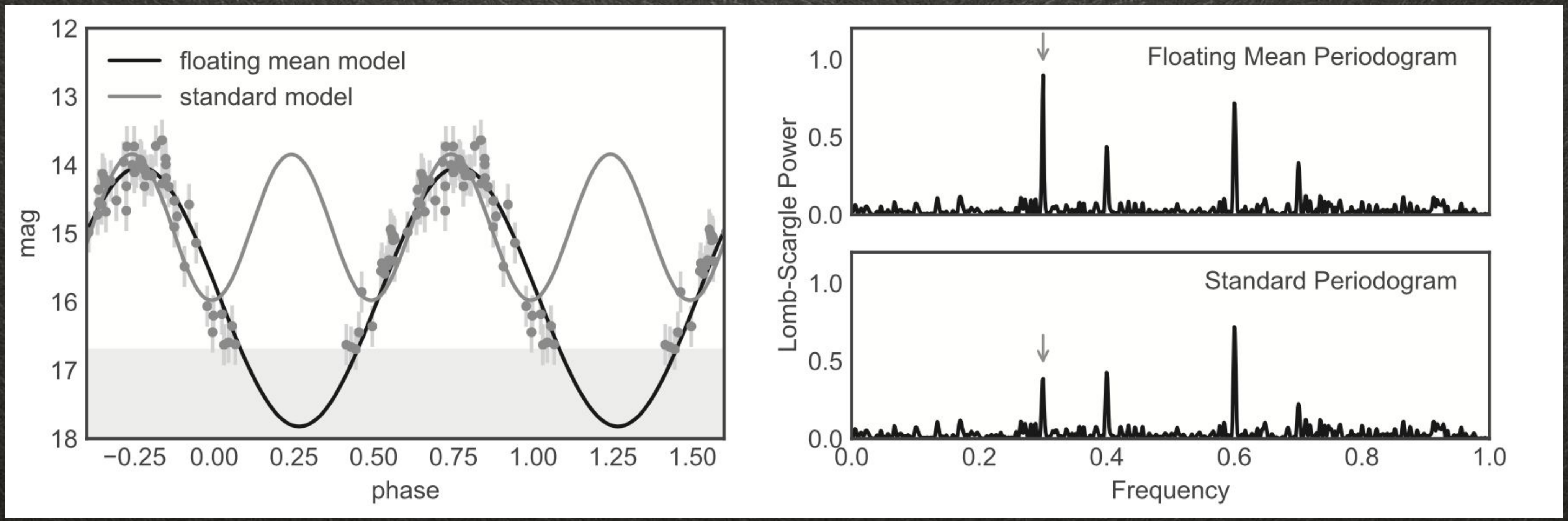
$$\chi^2(f) = (\boldsymbol{y} - \boldsymbol{y}_{\mathrm{model}})^T \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{y}_{\mathrm{model}}),$$

LS Periodogram Extensions

- Another important extension involves of an offset term to the sinusoidal model at each frequency:

$$y_{\text{model}}(t; f) = y_0(f) + A_f \sin(2\pi f (t - \phi_f)).$$

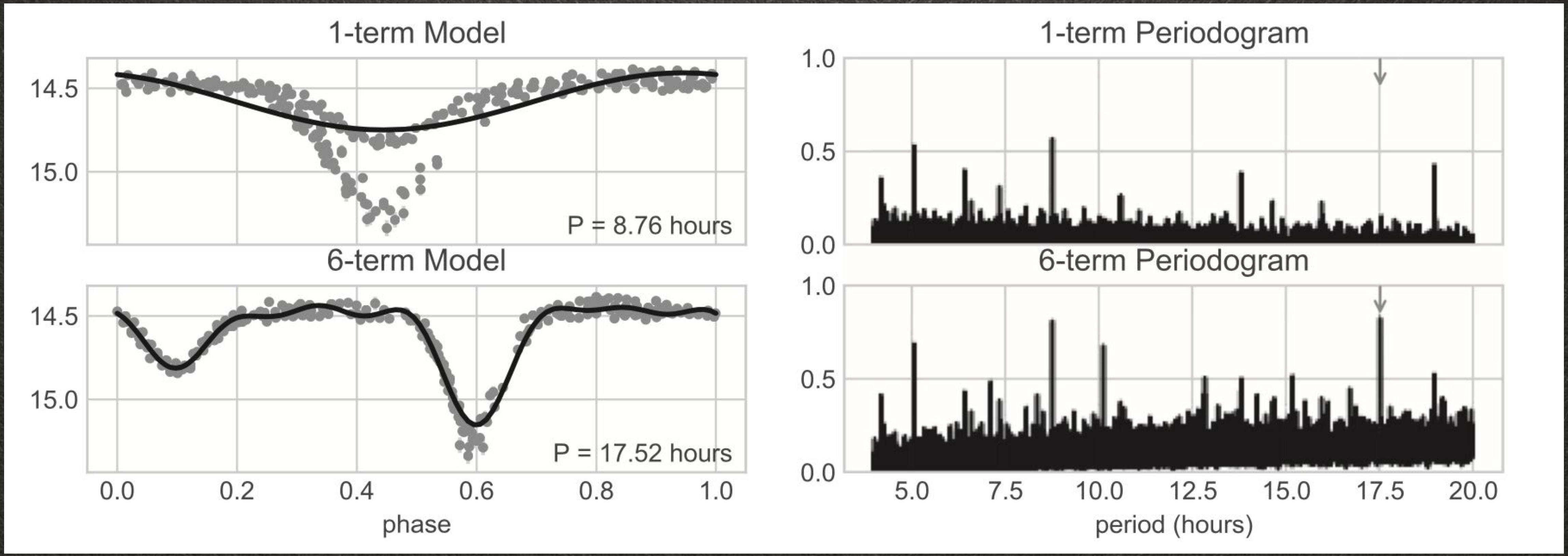
- This avoids the need to "zero-center" the light-curve, particularly important for incomplete (e.g. flux-limited) list curve.



LS Periodogram Extensions

- The interpretation of a periodogram as a regular fit procedure opens the way to even more extensions.

- It is possible, for instance, to model the curve with multiple Fourier components, allowing a greater flexibility.



LS Periodogram Extensions

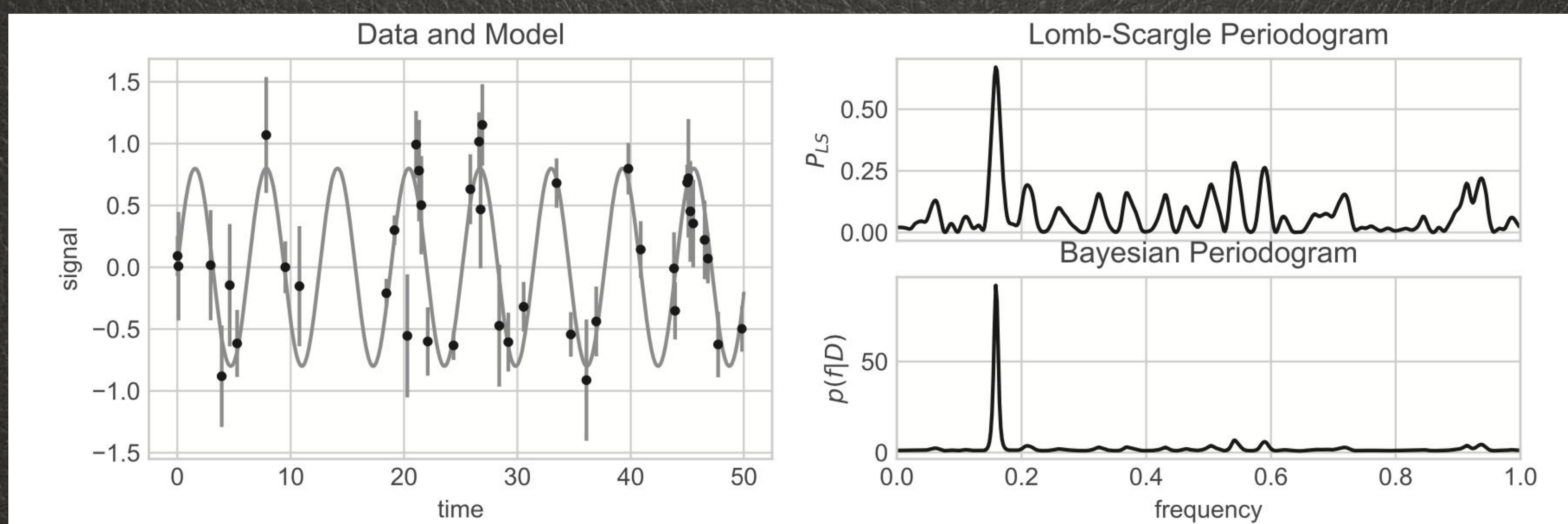
- In principle it is not even necessary to think to sinusoidal terms only. If a more complex signal shape is physically motivated, there are no intrinsic limitations.
- Model comparison is better carried out in a full Bayesian framework, with properly formalized prior knowledge, etc.
- Quite interestingly, it can be shown that the LS periodogram is in fact the optimal statistics for detecting a stationary sinusoidal signal in the presence of Gaussian noise.

Bayesian LS Periodogram

- For the standard, simple-sinusoidal model, the Bayesian periodogram is given by the posterior probability of frequency f given the data D and sinusoidal model M:

$$p(f|D, M) \propto e^{P_{\mathrm{LS}}(f)},$$

  - where P_{LS}(f) is the Lomb–Scargle power.

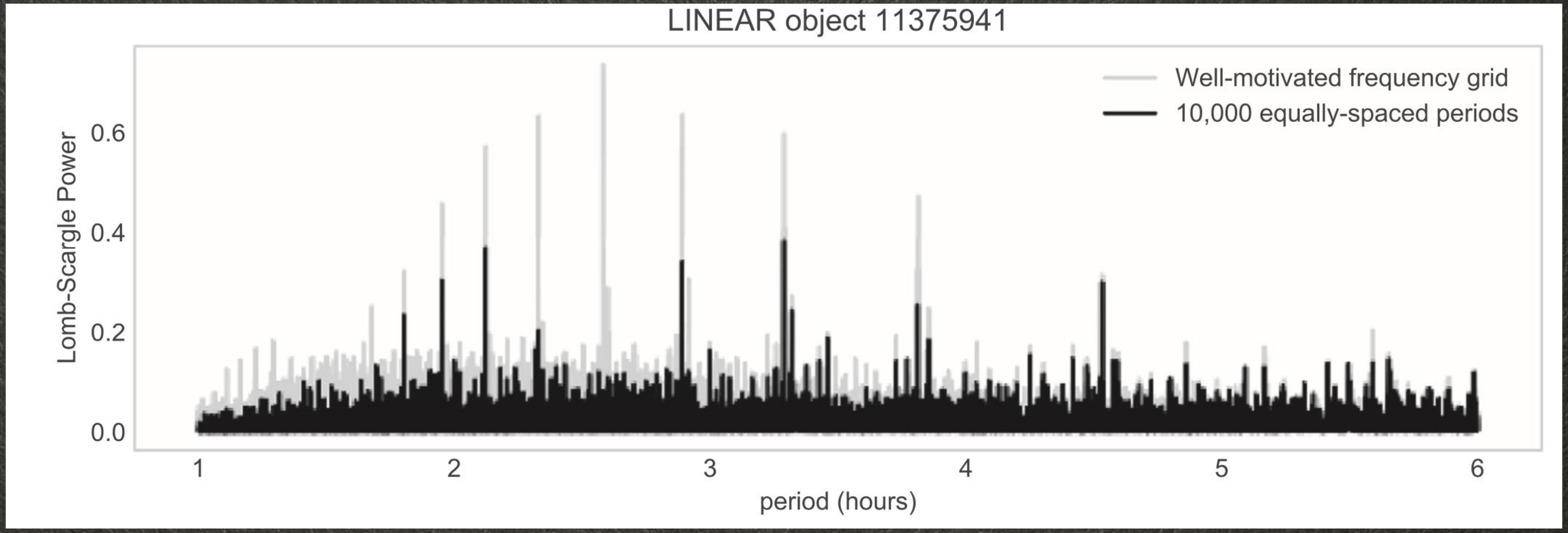


- Be aware that p(f | D,M) is the probability data are drawn from a sinusoidal model. Not the probability that data are periodic in general.

LS frequency grid

- At variance with the DFT, the frequency grid for LS analysis is not determined by the sampling.

- The most important warning is to pay attention not to choose a too sparse grid, with the risk to miss important features.



- In general, for an observation length T, we have sinc-shaped peaks of width ~1/T. A rule of the thumb advise is typically to oversample the grid by a factor ~5.

The Window Function

- It is clear that a careful evaluation of the window function is crucial in interpreting the results of a LS analysis.

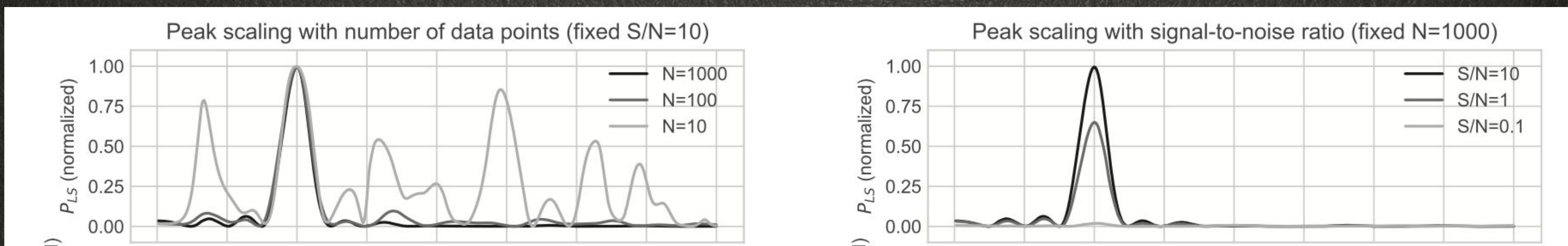
- The LS periodogram can offer a sufficiently reliable way to compute it:

$$\mathcal{P}_W(f;\ \{t_n\}) = \left| \sum_{n=1}^{N} e^{-2\pi i f t_n} \right|^2 .$$

- It is, in its essence, the periodogram for data g_n =1 at all times t_n. No floating-mean model should be used in this case.

LS Peak Uncertainty

- Be also careful in quoting period uncertainties basing on the peak width:



- In the Bayesian interpretation one can derive an approximate expression relating the uncertainty to the number of samples, N, and the average S/N ratio, Σ:

$$\sigma_f \approx f_{1/2}\sqrt{\frac{2}{N\Sigma^2}}.$$

LS Peak Significance

- Even for the LS periodogram, and pure Gaussian noise, it can be proved that the values of the unnormalized periodogram follow a distribution with two degrees of freedom.

- If $Z=P(f_0)$ is the periodogram value then the cumulative probability to observe a value less then Z is:

$$P_{\text{single}}(Z) = 1 - \exp(-Z)$$

- As for the DFT, we are generally not interested in the distribution of one particular randomly chosen frequency, but rather the distribution of the highest peak of the periodogram.

LS Peak Significance

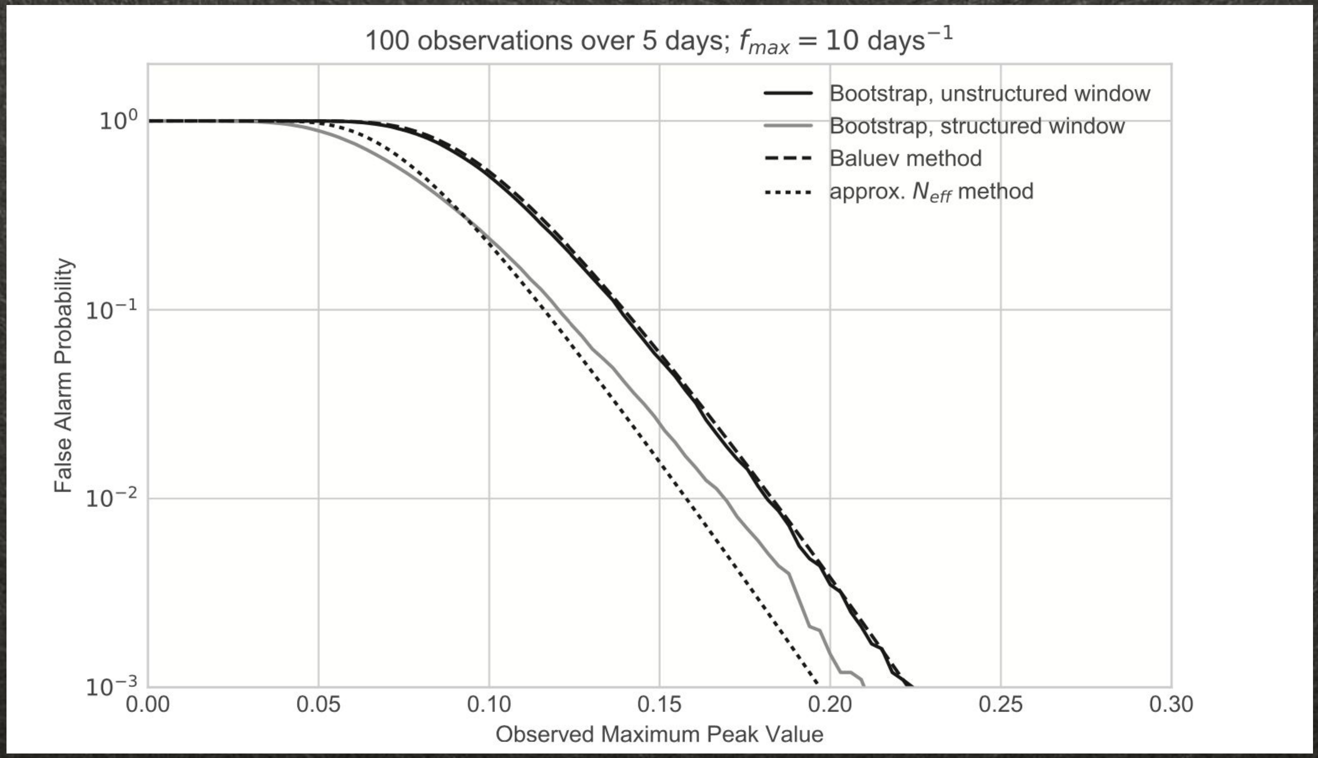
- The problem is not simple to solve here since the value at one frequency is correlated with the value at other frequencies in a way that is quite difficult to analytically express. These correlations come from the convolution with the survey window.
- One common approach is to assume that it can be modeled on some “effective number” of independent frequencies N_{eff}, so that the False Alarm Probability can be estimated as:

$$FAP(z) \approx 1 - [P_{\rm single}(z)]^{N_{\rm eff}}.$$

Independent frequencies

- A very simple estimate for N_{eff} can based on the arguments about the expected peak width, $\delta f = 1/T$. In this approximation, the number of independent peaks in a range $0 \le f \le f_{max}$ is assumed to be $N_{eff} = f_{max} T$.

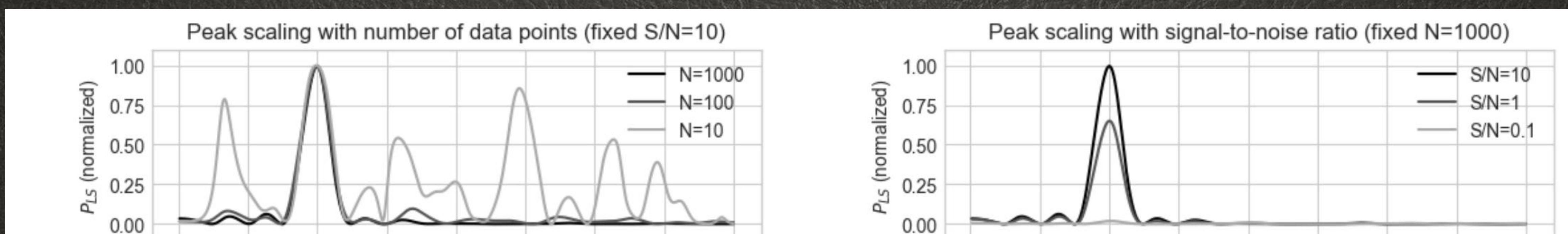
- Finding a reliable way to compute N_{eff} analytically is still and open problem. More often the problem is solved computationally by, e.g., a bootstrap procedure.

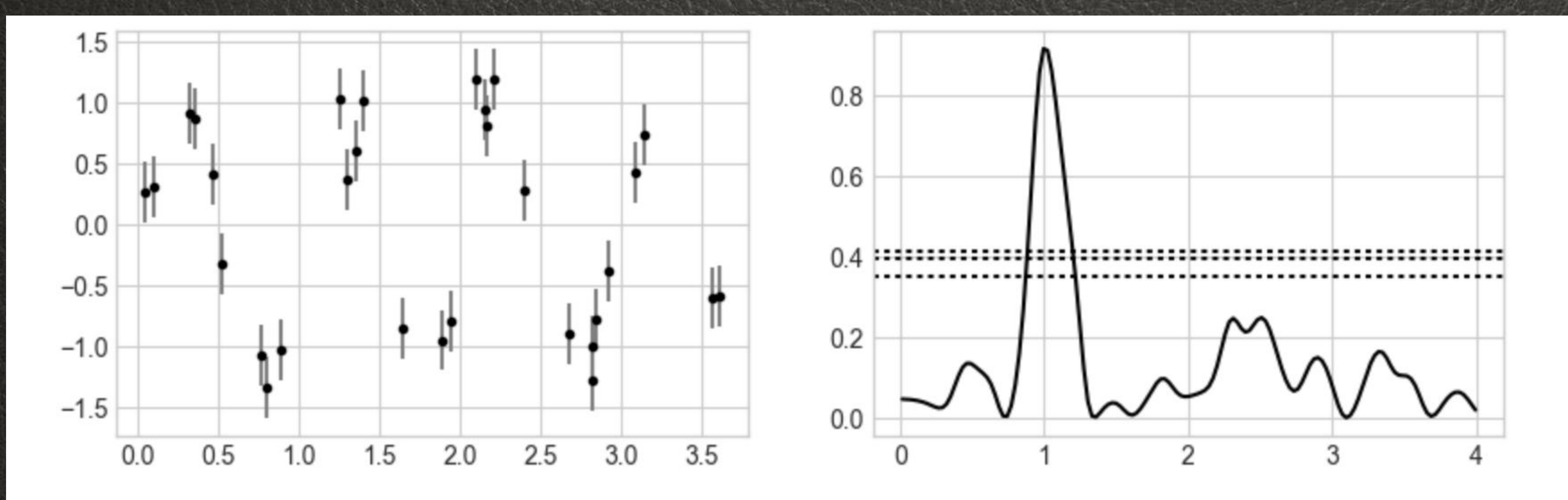


Exercises

- Useful notebooks:

1. Uncertainties





We'll see the role of data number and uncertainties, a simple application of a bootstrap algorithm.

PSD Normalization

- When considering the periodogram from the Fourier perspective, it is useful to normalize the periodogram such that in the special case of equally spaced data it recovers the standard Fourier power spectrum.

- This is the so-called 'psd' normalization, and the equivalent least-squares expression is:

$$P(f) = \frac{1}{2}[\hat{\chi}_0^2 - \hat{\chi}^2(f)].$$

- For equally spaced data this becomes:

$$P(f) = \frac{1}{N}\,|\mathrm{FFT}(y_n)|^2;$$

PSD Normalization

- With the 'psd' normalization periodogram units are "unit(y)^2".
- And can be interpreted as squared amplitudes of the Fourier component at each frequency.
- However, if uncertainties are included in the analysis, the periodogram becomes unitless: it is essentially a measure of periodic content in signal-to-noise ratio rather than in signal itself.

PSD Normalization

- In the least-squares view of the periodogram, the periodogram is interpreted as an inverse measure of the goodness of fit for a model.

$$P(f) = \frac{1}{2}[\hat{\chi}_0^2 - \hat{\chi}^2(f)],$$

- If the sinusoidal model *perfectly* fits the data at some frequency f_0, then $\hat{\chi}^2(f_0) = 0$, and the periodogram is maximized at a value $\hat{\chi}_0^2/2$.

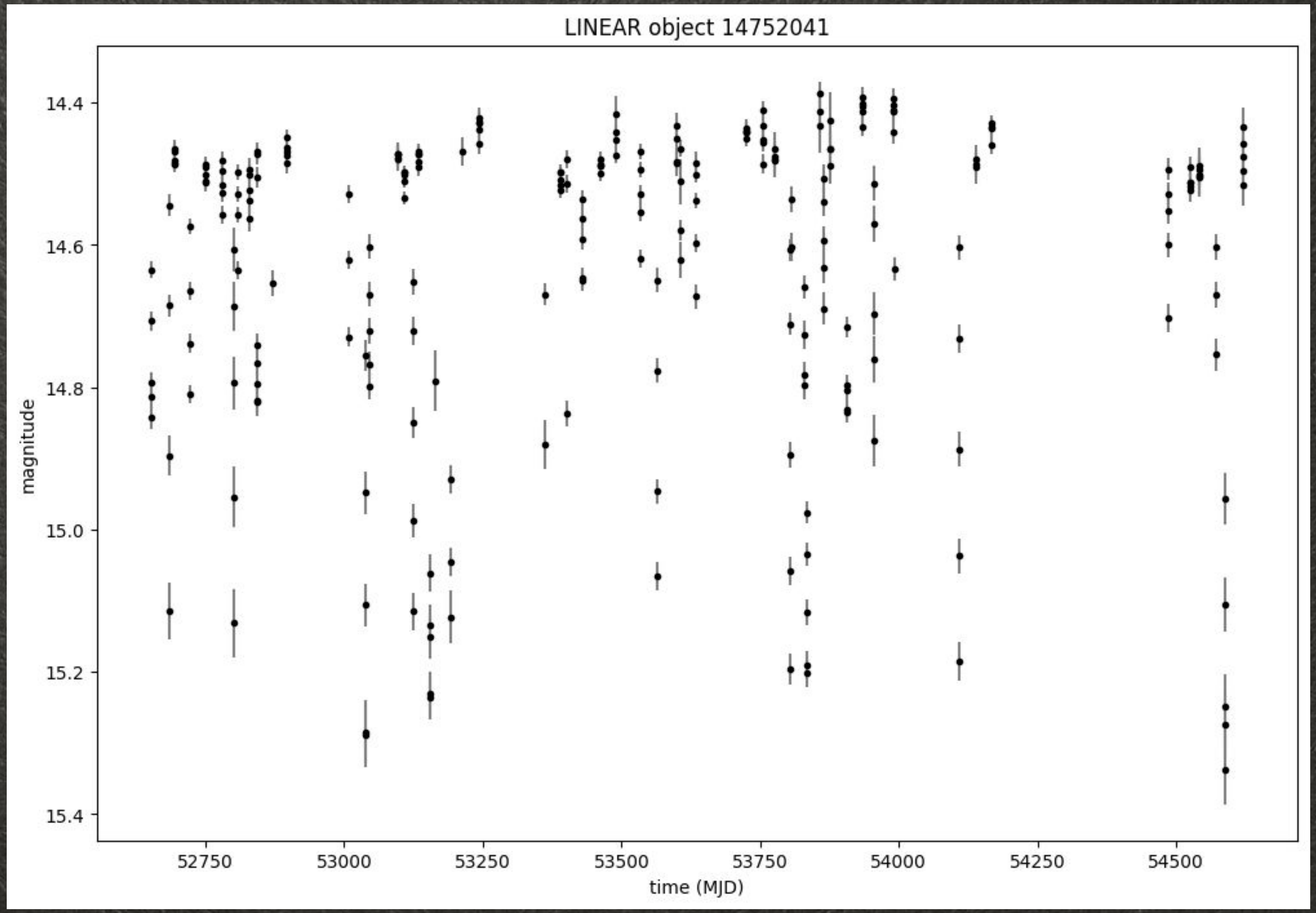
- The minimum value of the periodogram can only be 0, and therefore a possible normalization that keep the (unitless) values between 0 and 1 is:

$$P_{\mathrm{norm}}(f) = 1 - \frac{\hat{\chi}^2(f)}{\hat{\chi}_0^2}.$$

Exercises

- Useful notebooks:

1. LINEAR_object_14752041



REFERENCES AND DEEPENING

Understanding the Lomb–Scargle Periodogram

Jacob T. VanderPlas
University of Washington, eScience Institute, 3910 15th Ave NE, Seattle WA 98195

Received 2017 August 26; revised 2017 December 6; accepted 2017 December 7; published 2018 May 11

Abstract

The Lomb–Scargle periodogram is a well-known algorithm for detecting and characterizing periodic signals in unevenly sampled data. This paper presents a conceptual introduction to the Lomb–Scargle periodogram and important practical considerations for its use. Rather than a rigorous mathematical treatment, the goal of this paper is to build intuition about what assumptions are implicit in the use of the Lomb–Scargle periodogram and related estimators of periodicity, so as to motivate important practical considerations required in its proper application and interpretation.

Key words: methods: data analysis – methods: statistical

Jacob VanderPlas