Tiana_Athriel

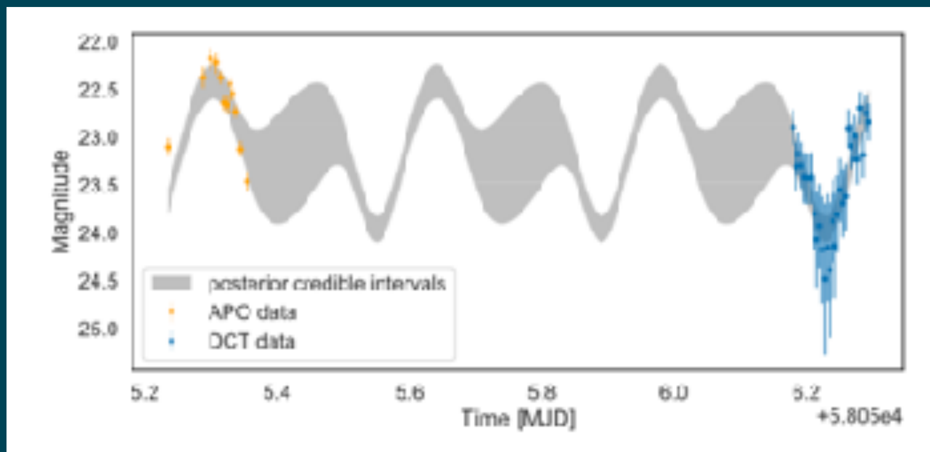
dhuppenkothen

Fun Statistics with Fourier Spectra

Daniela Huppenkothen

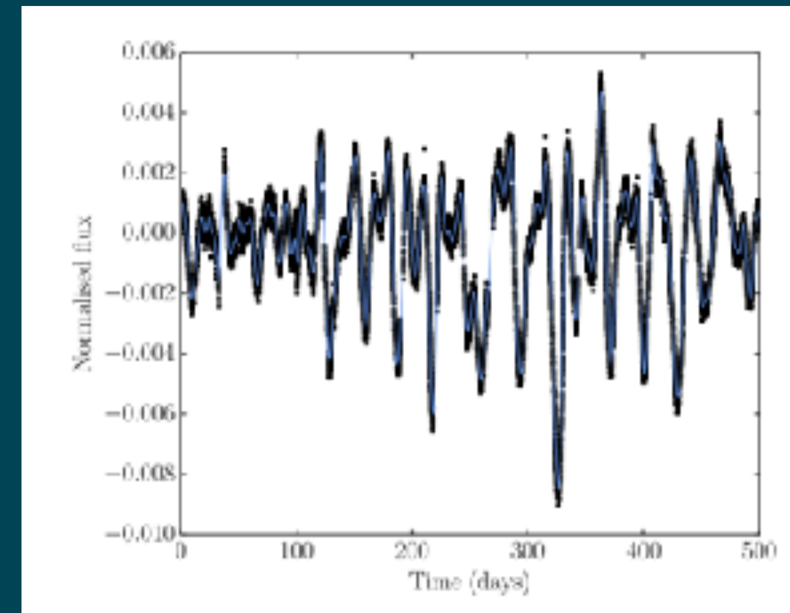
DIRAC Institute, UW Seattle

asteroids



Bolin et al (2017)

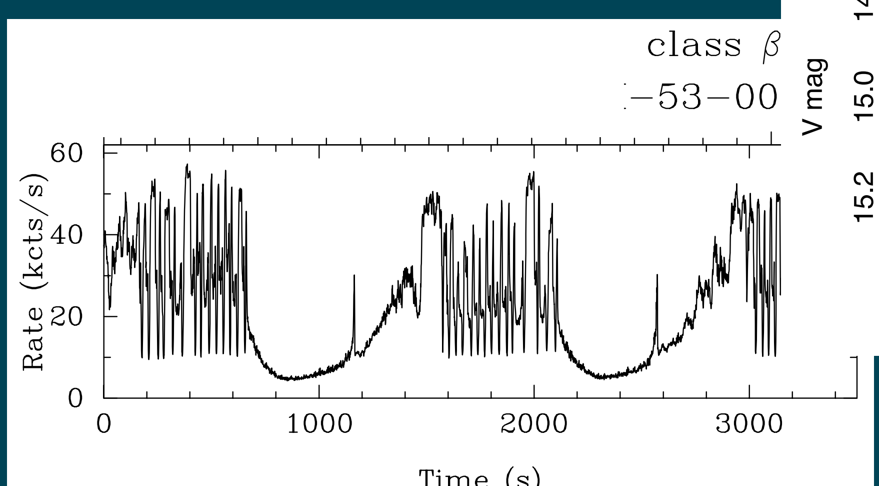
stars



Angus et al (2018)

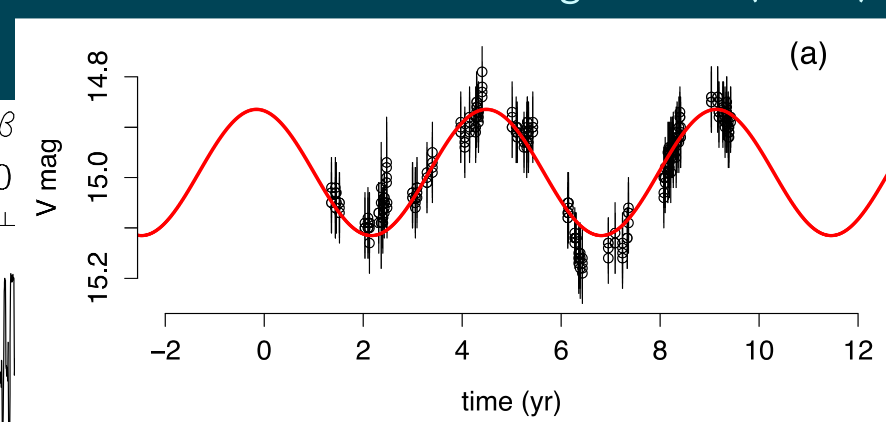
Almost all things in the universe are variable

black holes



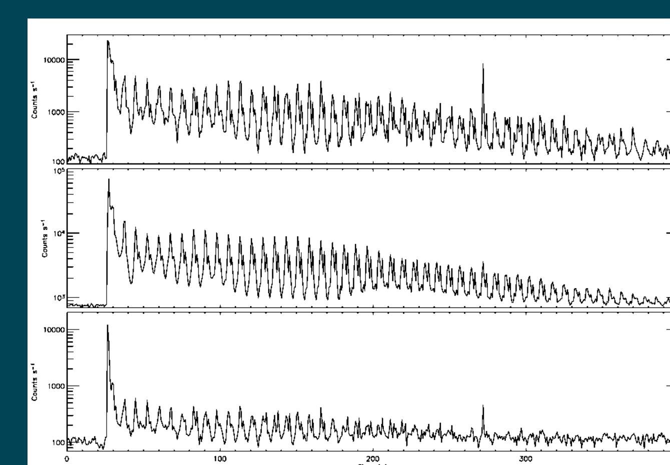
Belloni et al (2000)

Vaughan et al (2016)



AGN

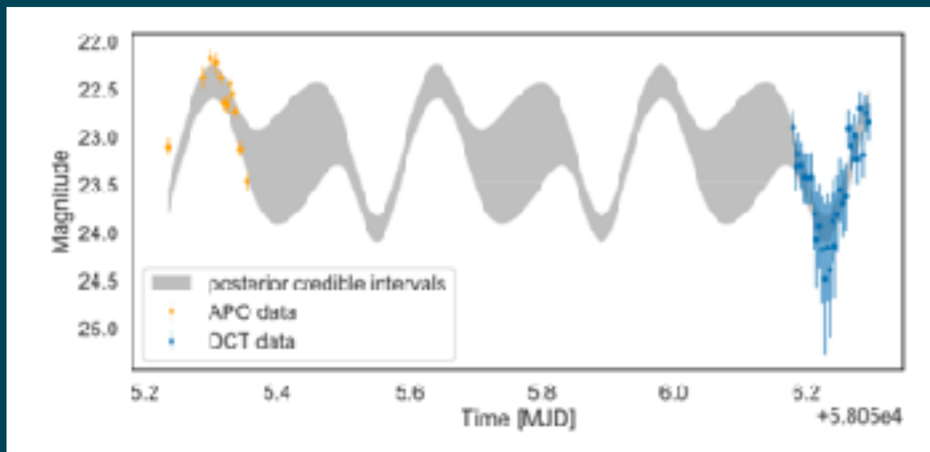
Strohmayer & Watts (2005)



neutron stars

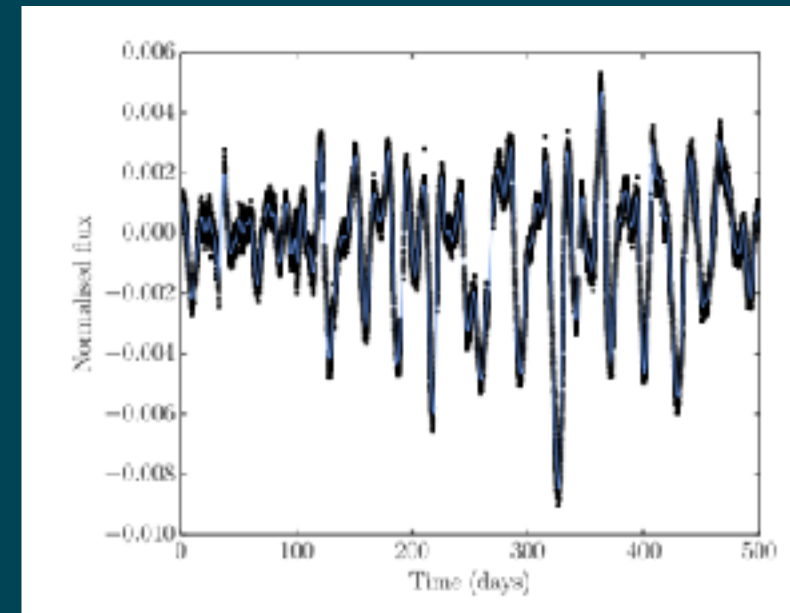
Almost all things in the universe are variable

asteroids



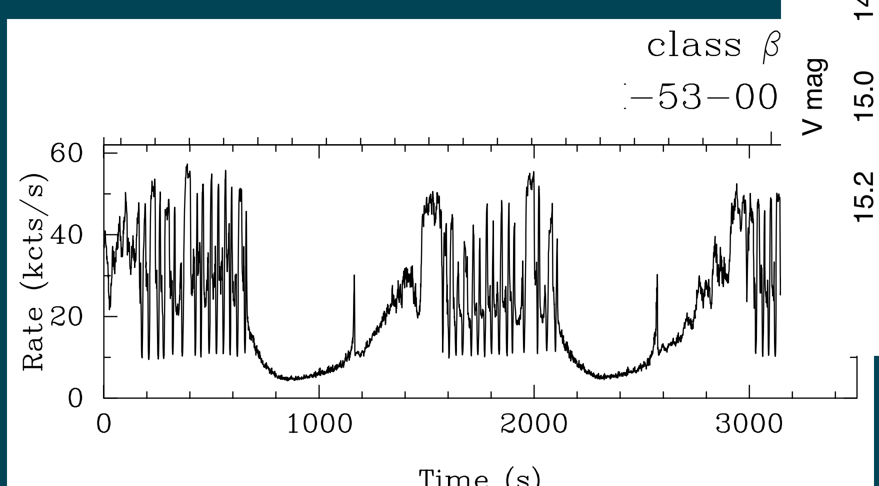
Bolin et al (2017)

stars



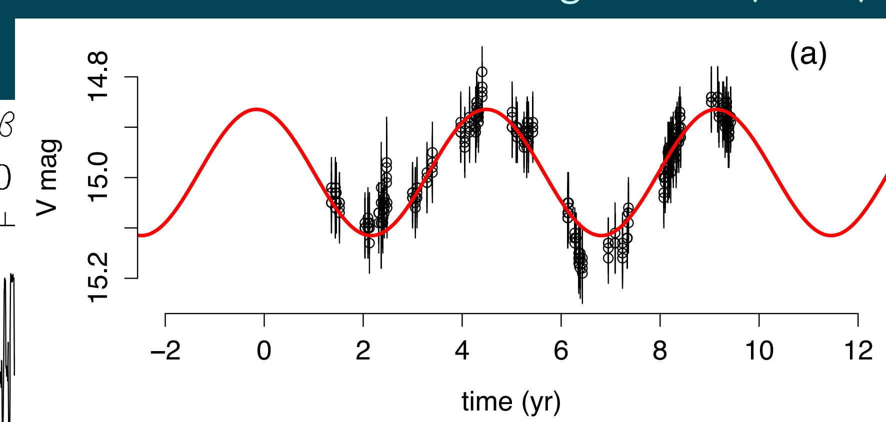
Angus et al (2018)

black holes



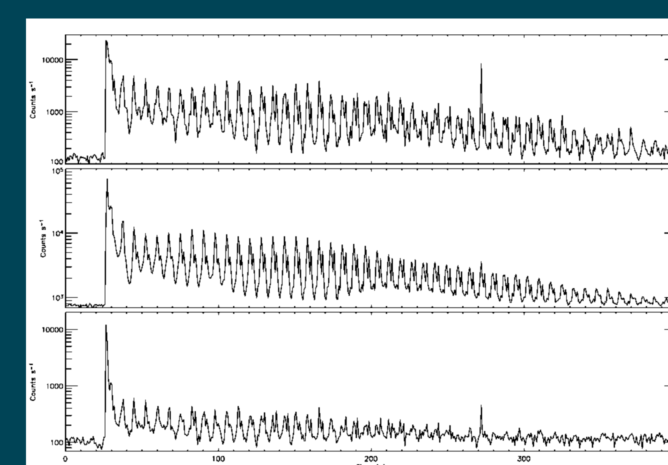
Belloni et al (2000)

Vaughan et al (2016)

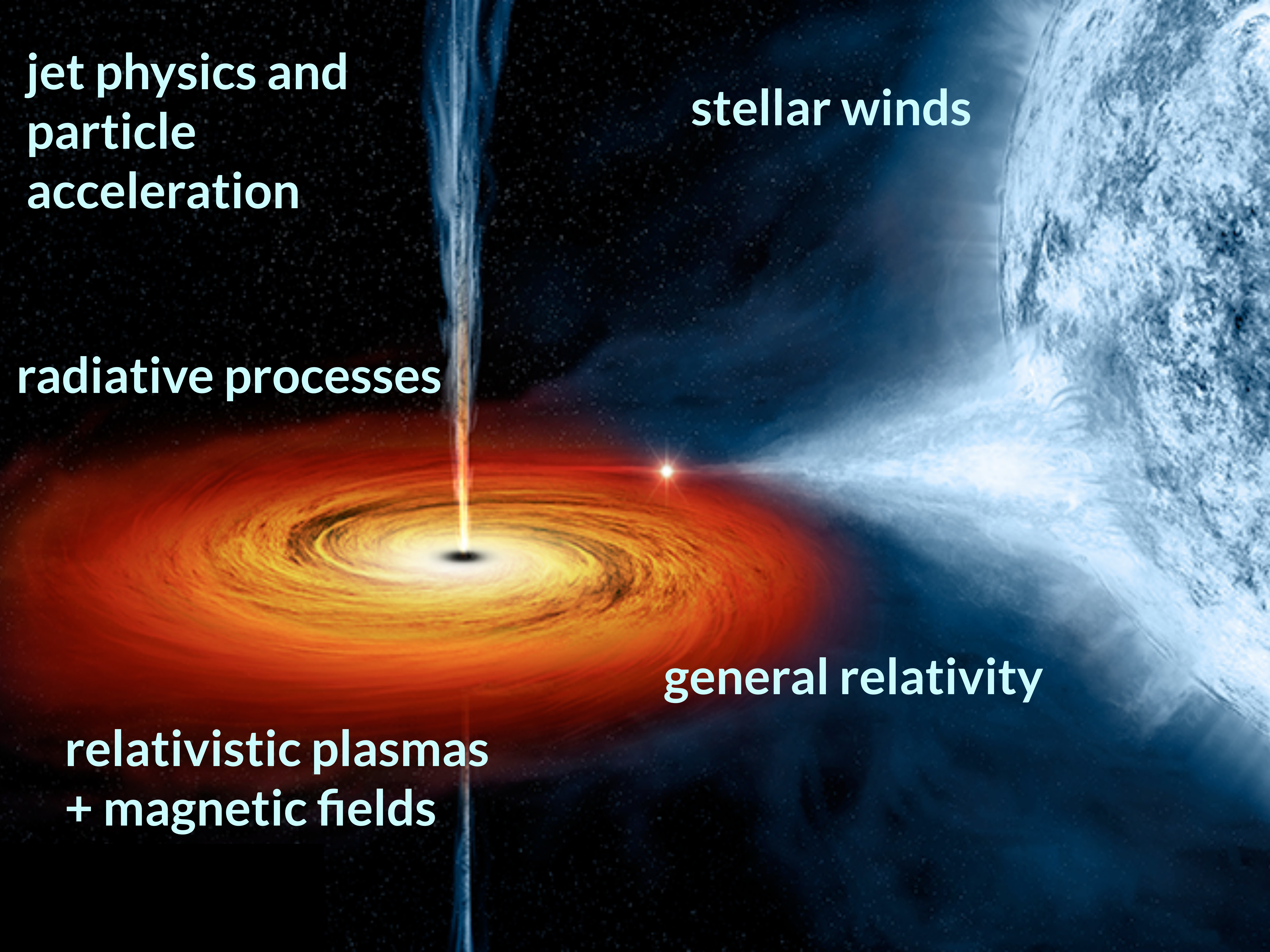


AGN

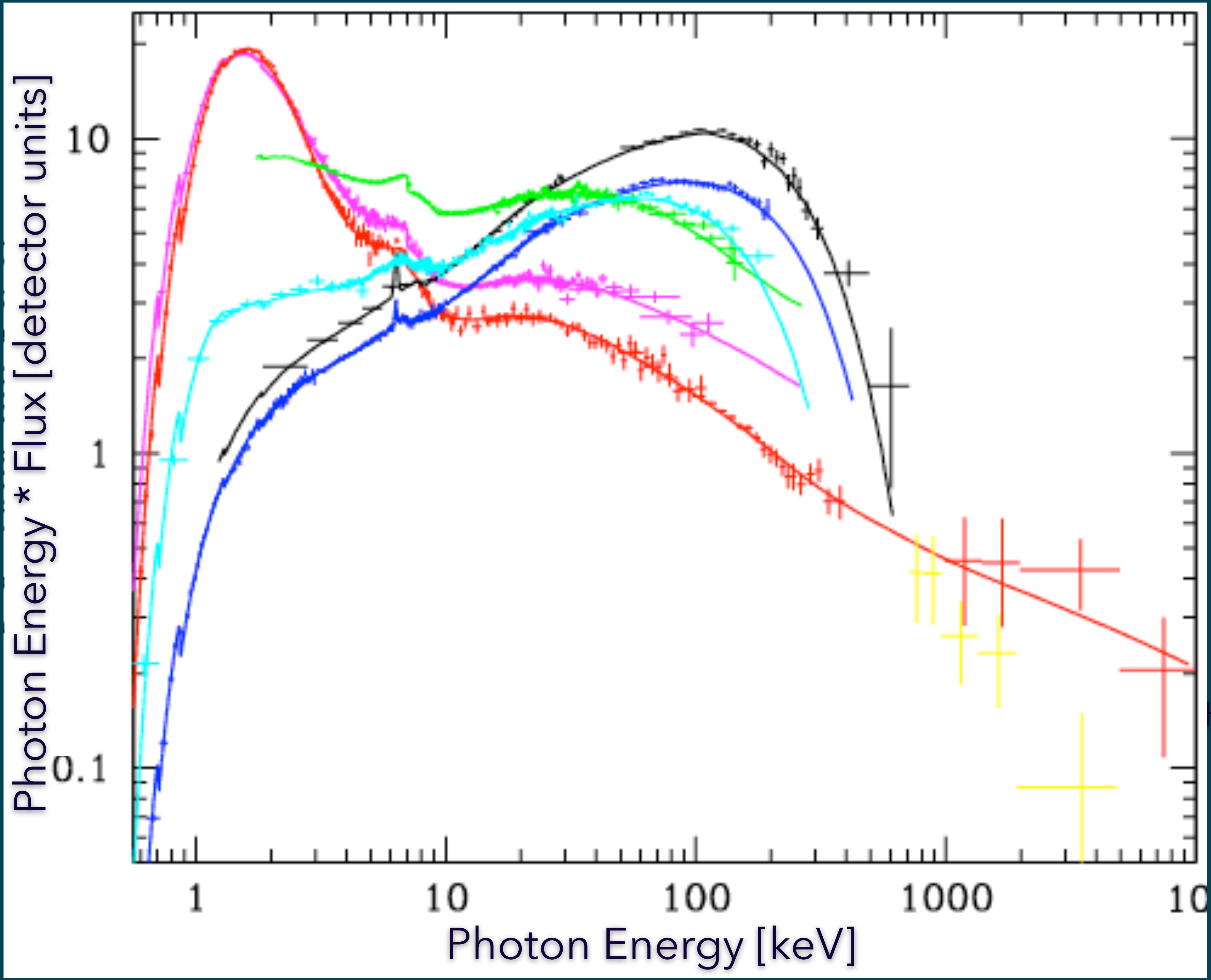
Strohmayer & Watts (2005)



neutron stars

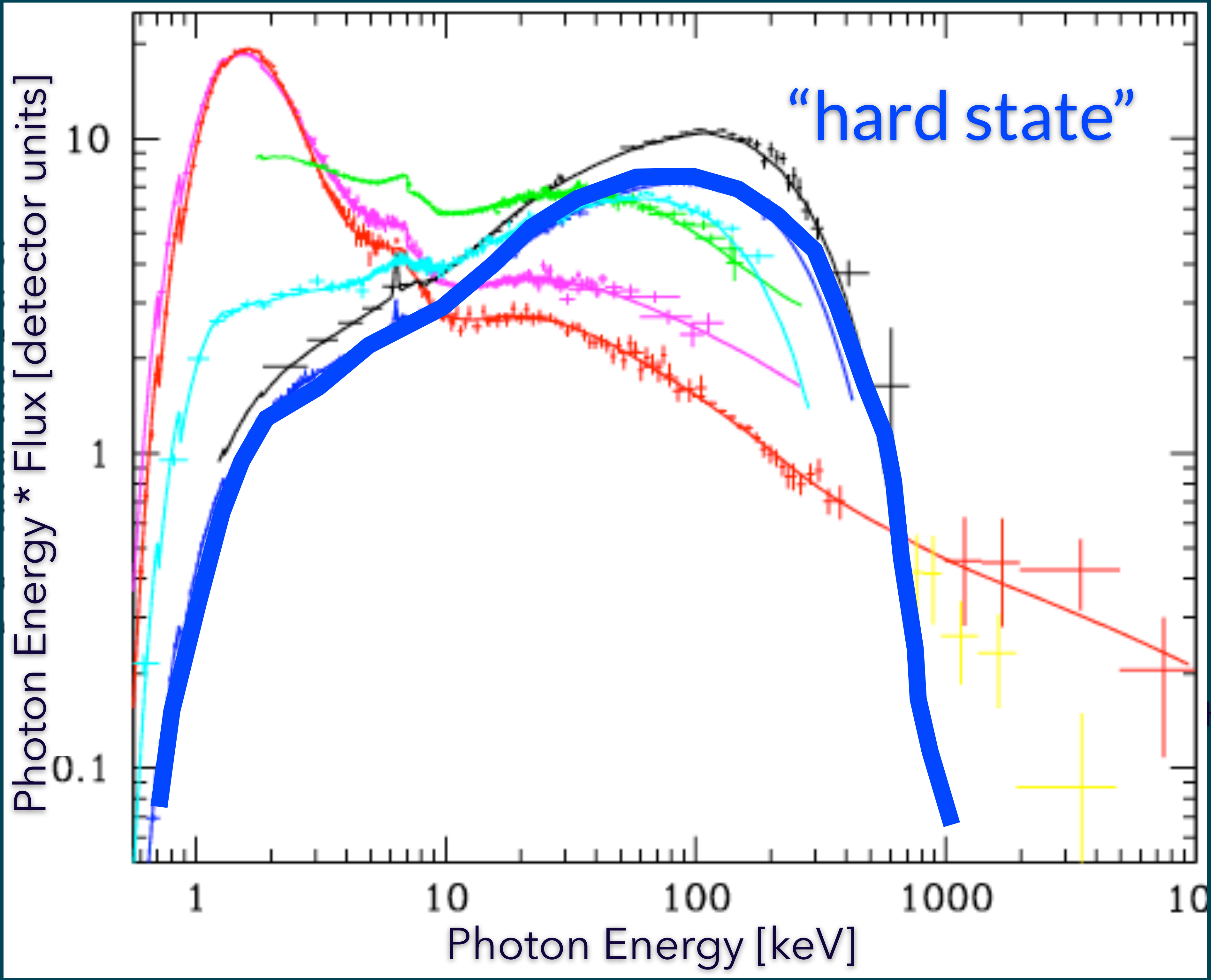

jet physics and
particle
acceleration
stellar winds
radiative processes
general relativity
relativistic plasmas
+ magnetic fields

X-ray Spectra Vary with Time



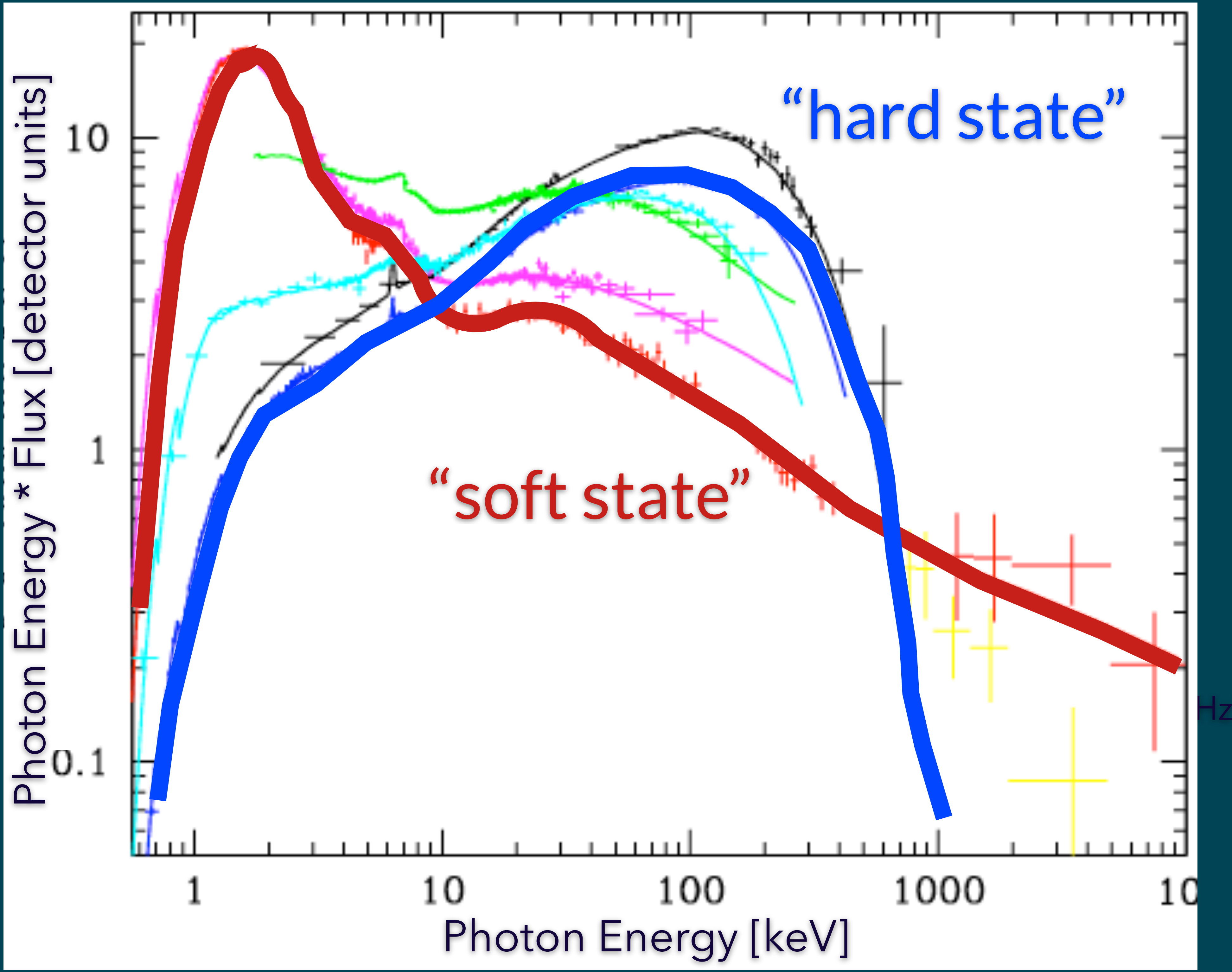
Malzac (2008)

X-ray Spectra Vary with Time



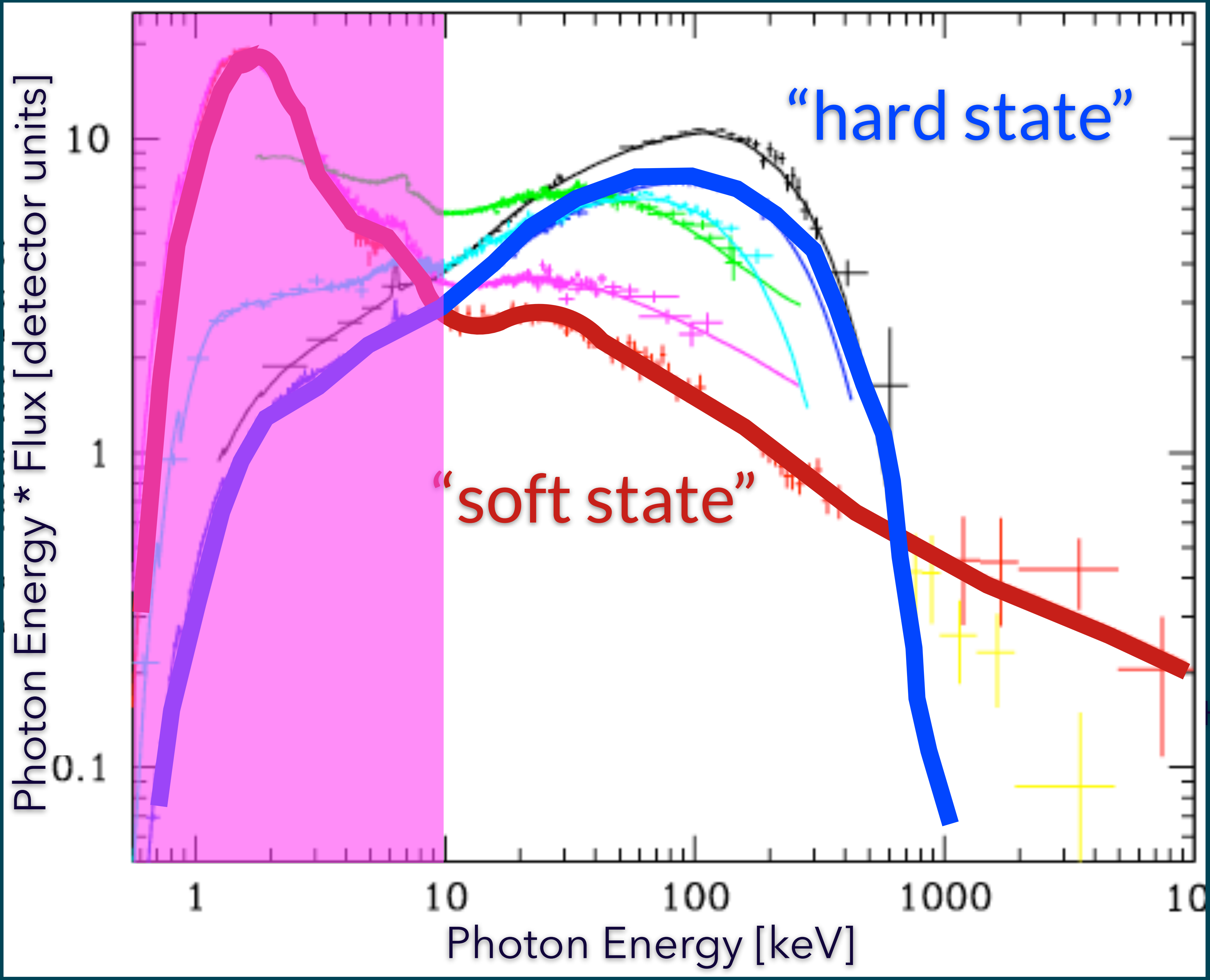
Malzac (2008)

X-ray Spectra Vary with Time



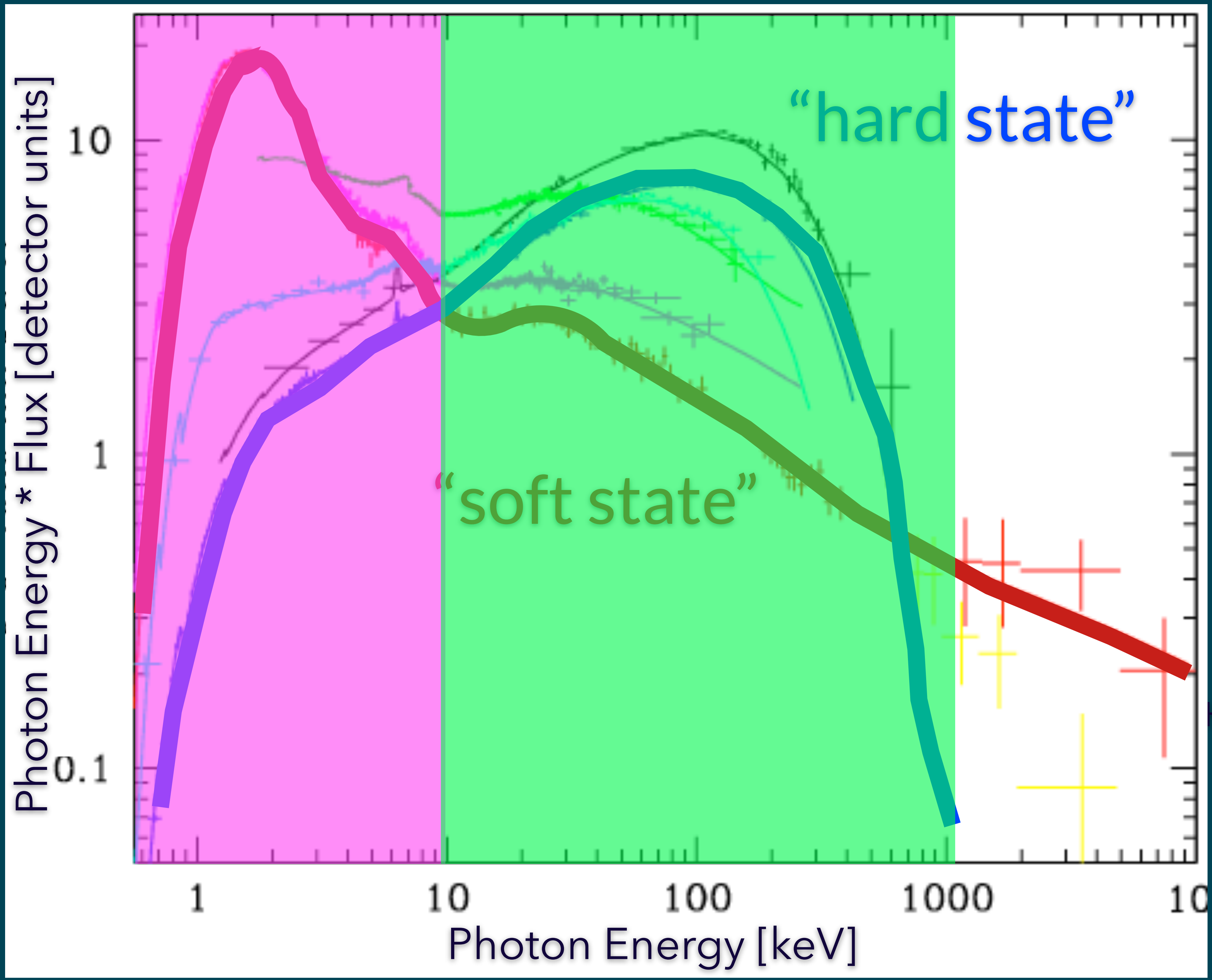
Malzac (2008)

X-ray Spectra Vary with Time



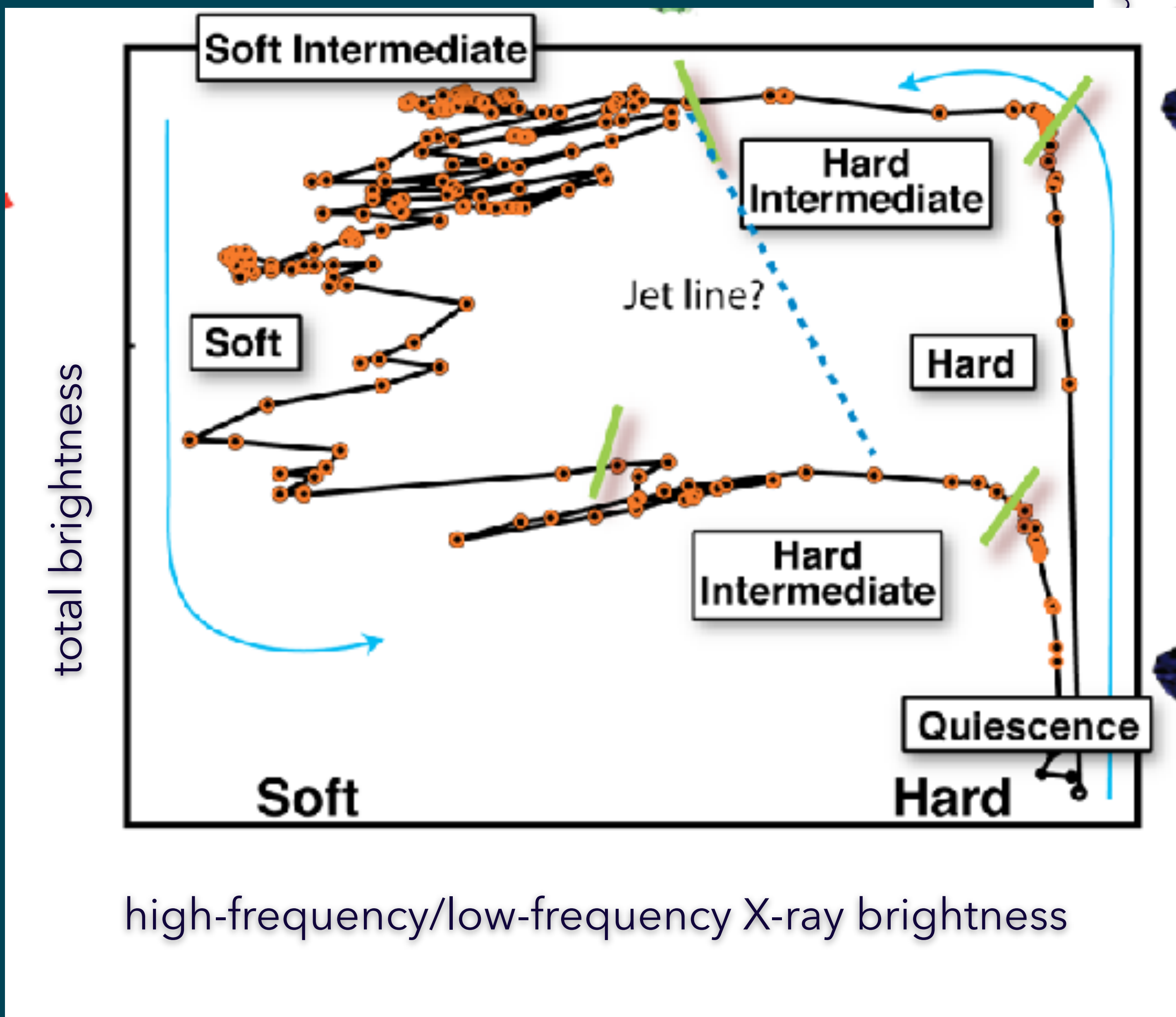
Malzac (2008)

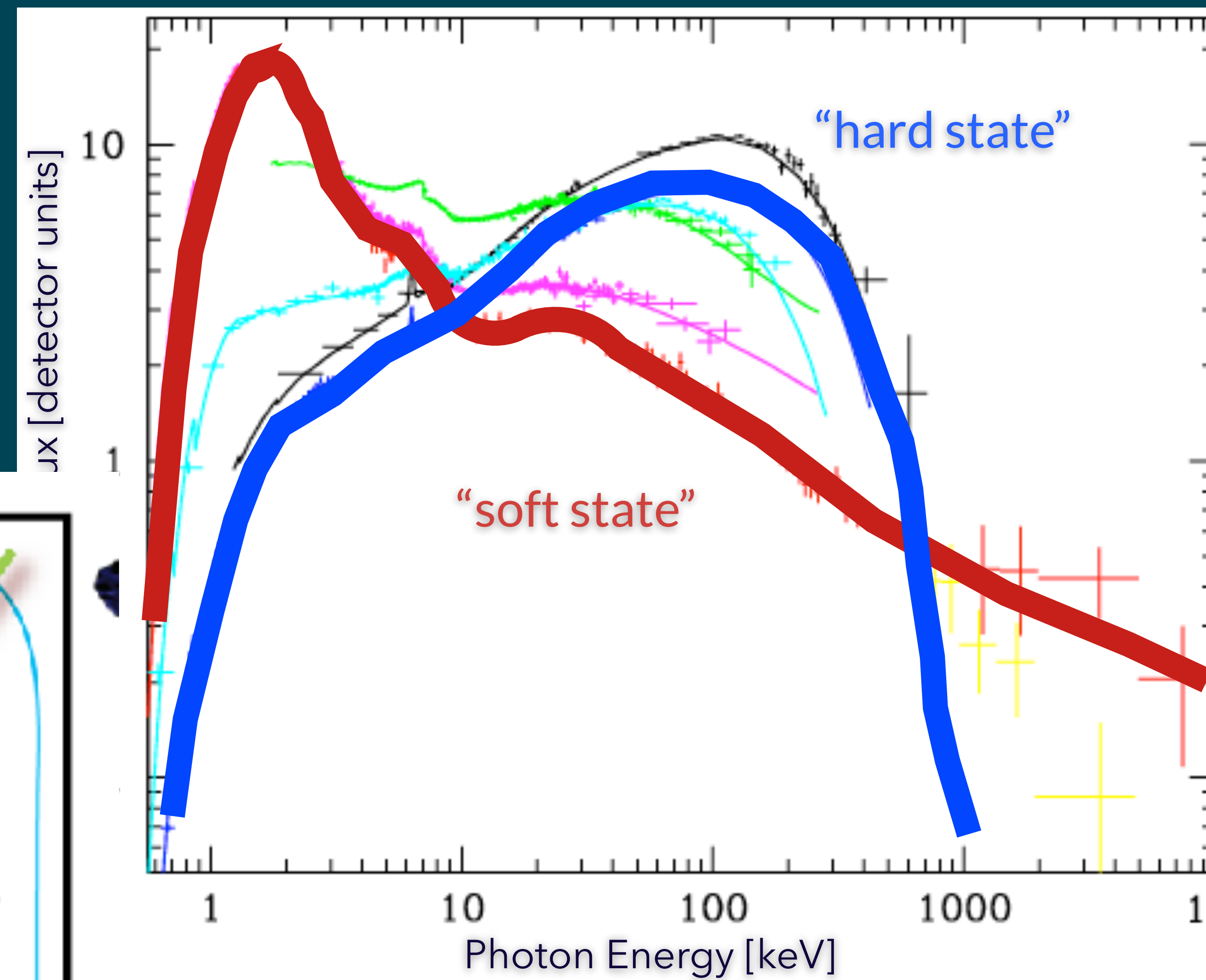
X-ray Spectra Vary with Time



Malzac (2008)

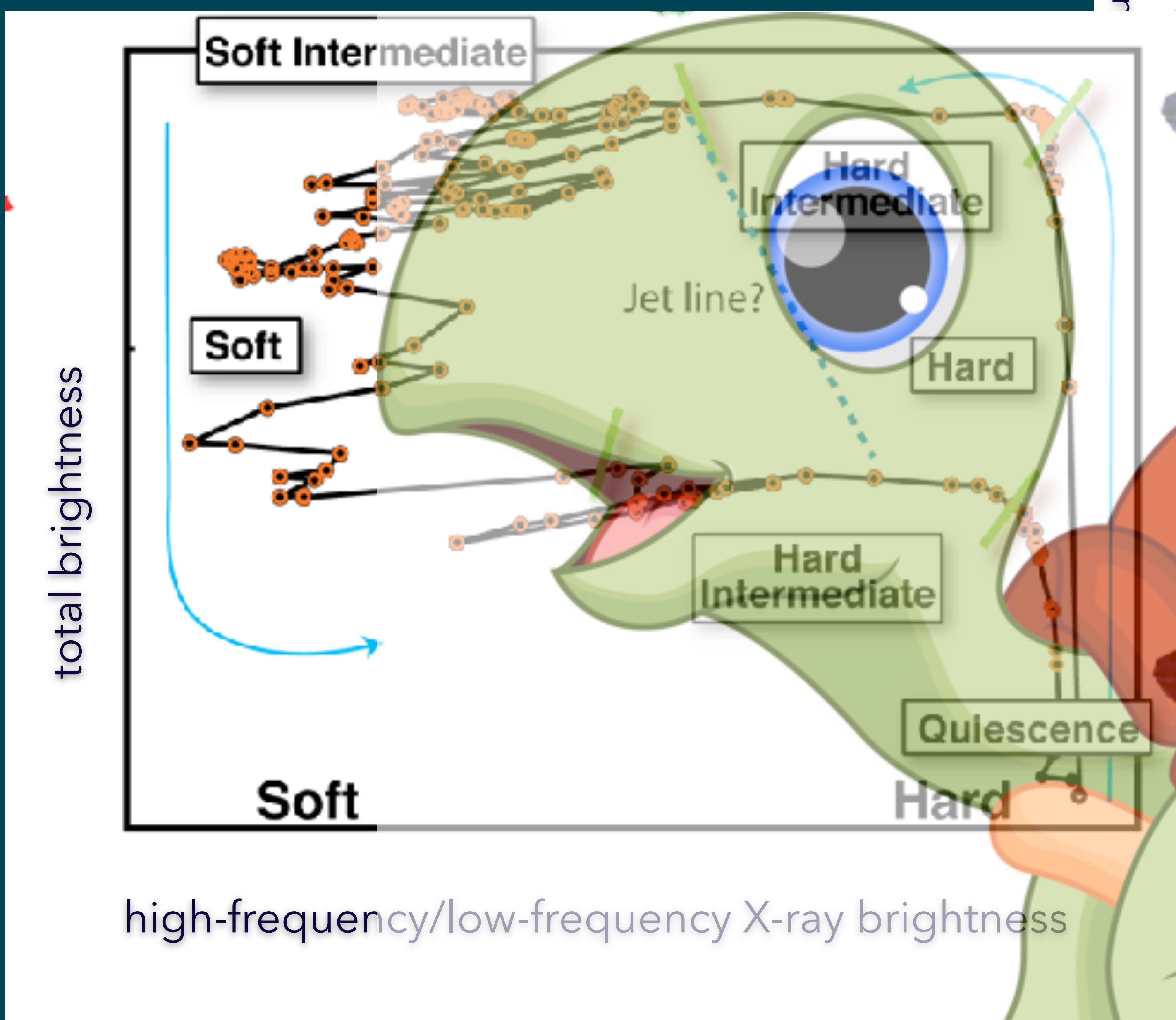
X-ray Spectra Vary with Time

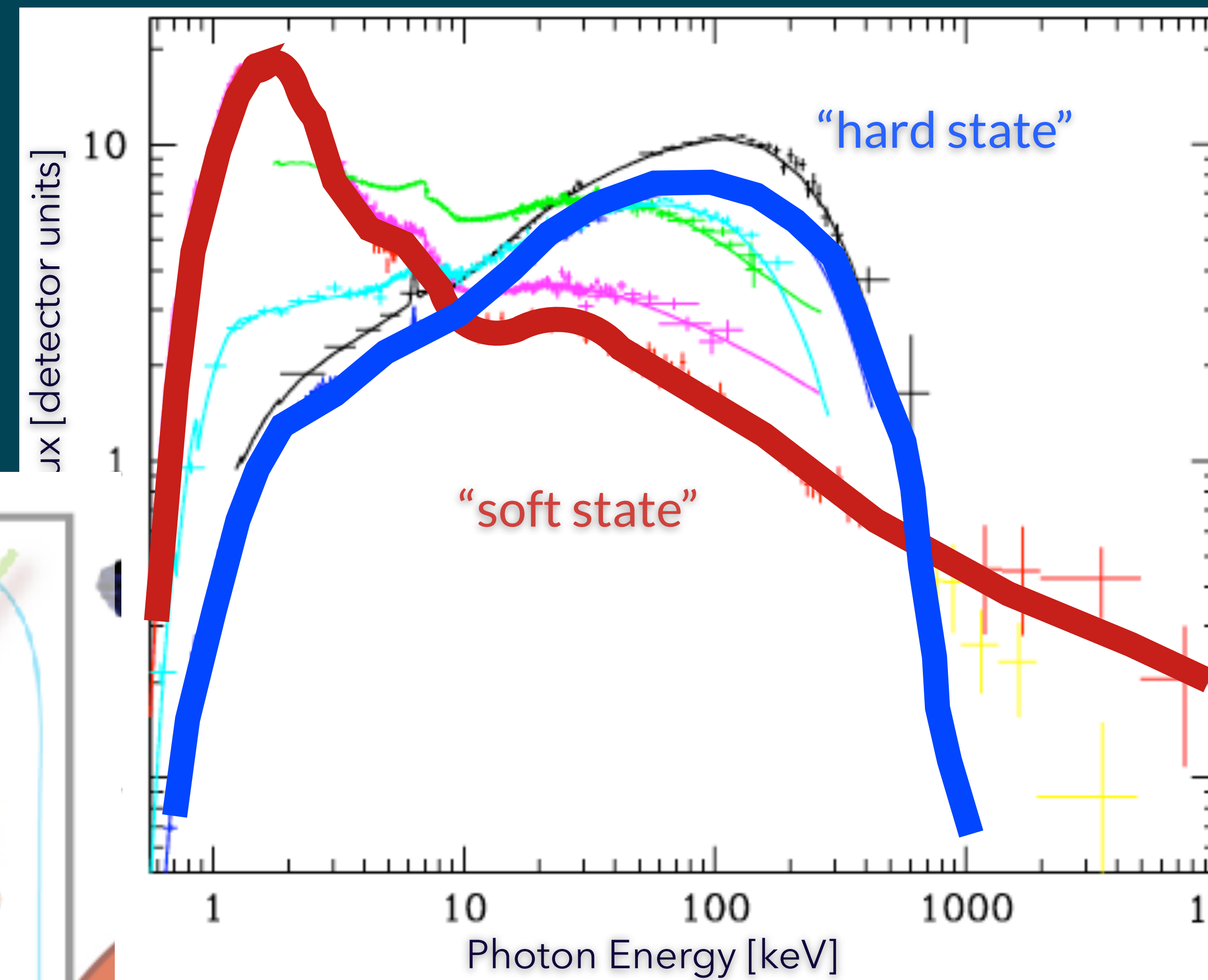




Malzac (2008)

X-ray Spectra Vary with Time





Malzac (2008)

Fourier Analysis tells us about variability!*

Fourier Analysis tells us about variability!*

*for evenly sampled time series

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i j k/N}$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t) .$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i jk/N}$$

$$A_j = \sum_k x_k \cos\omega_j t_k$$

$$B_j = \sum_k x_k \sin\omega_j t_k ,$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi ijk/N}$$

$$A_j = \sum_k x_k \cos\omega_j t_k$$

$$B_j = \sum_k x_k \sin\omega_j t_k,$$

$$a = \sqrt{A^2 + B^2}$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i jk/N}$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N} \sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N} \sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t) \,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i j k/N}$$

periodogram:

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i jk/N}$$

periodogram:

$$P_j \equiv \frac{2}{N_{ph}}|a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N} \sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N} \sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t) .$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i j k/N}$$

useful normalization

periodogram:

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N} \sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N} \sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t) \,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i j k/N}$$

periodogram:

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

$$a_j = \sum_{k=0}^{N-1} x_k e^{2\pi i j k/N}$$

periodogram:

$$P_j \equiv \frac{2}{N_{ph}}|a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

statistical distribution?

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t)\,.$$

$$P_j \equiv \frac{2}{N_{ph}}|a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N} \sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N} \sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t) \, .$$

assume many data points

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

assume many data points

~ Gaussian

$$P_j \equiv \frac{2}{N_{ph}}|a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N} \sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N} \sum_j (A_j \cos \omega_j t + B_j \sin \omega_j t)\,.$$

assume many data points

~ Gaussian

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

$$|a_j|^2 = A^2 + B^2$$

Fourier Analysis tells us about variability!*

*for evenly sampled time series

$$x(t) = \frac{1}{N}\sum_j a_j \cos(\omega_j t - \phi_j) = \frac{1}{N}\sum_j (A_j \cos\omega_j t + B_j \sin\omega_j t)\,.$$

assume many data points

~ Gaussian

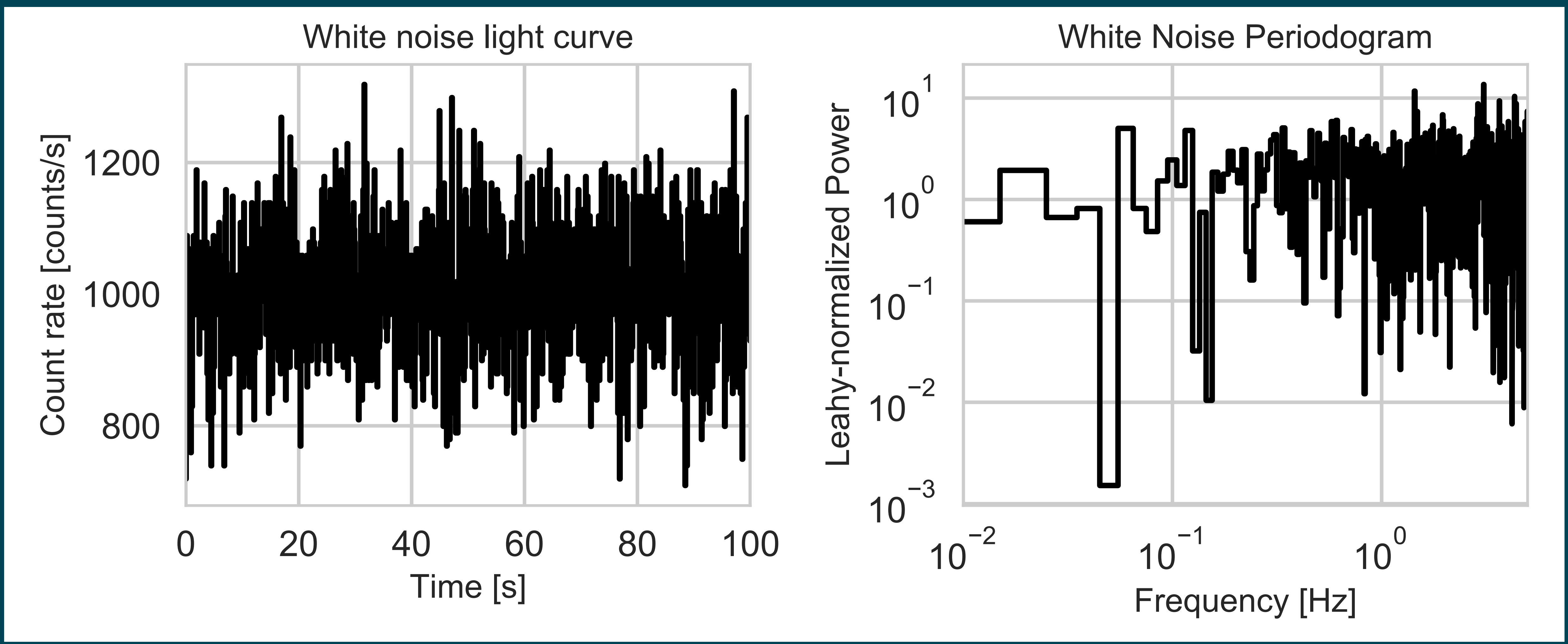
$$P_j \equiv \frac{2}{N_{ph}}|a_j|^2 \qquad j = 0, \ldots, \frac{N}{2},$$

$$|a_j|^2 = A^2 + B^2$$

χ^2 with 2 degrees of freedom

Fourier Analysis tells us about variability!*

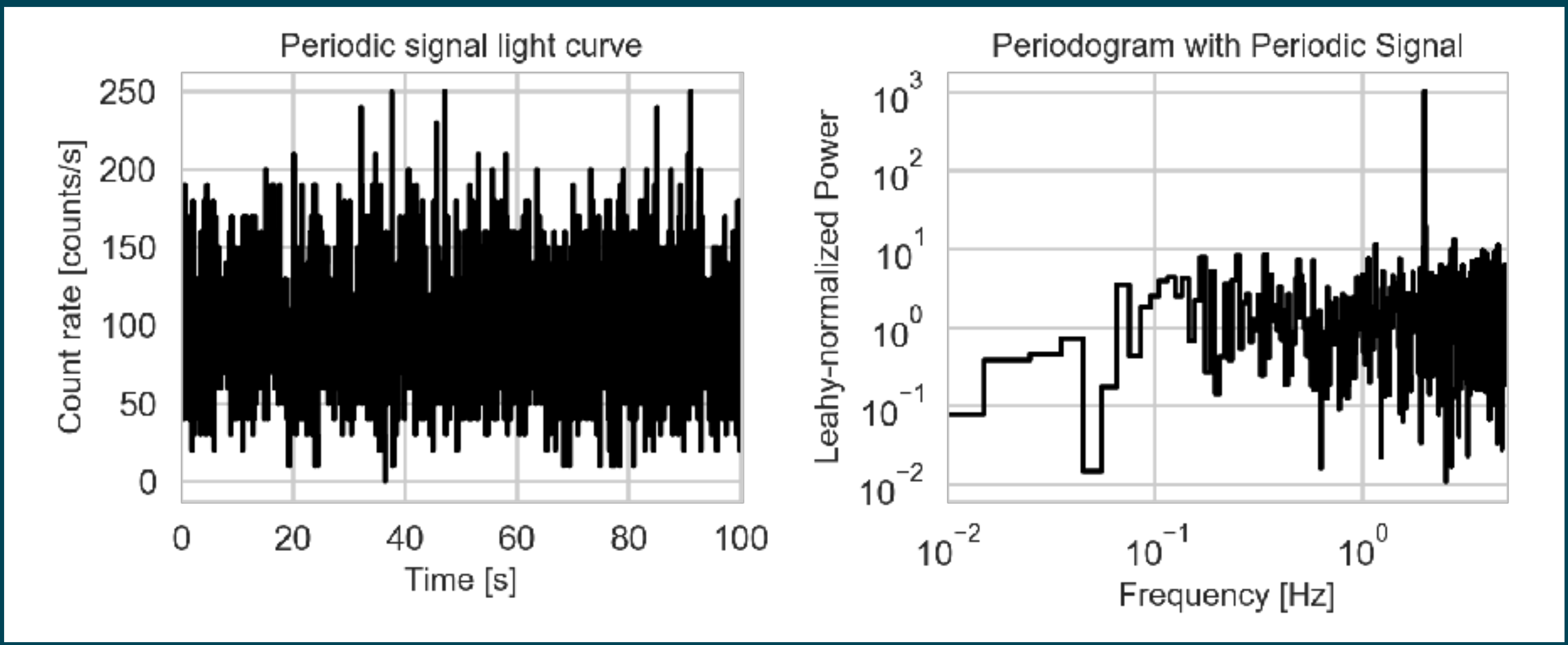
*for evenly sampled time series



white noise

Fourier Analysis tells us about variability!*

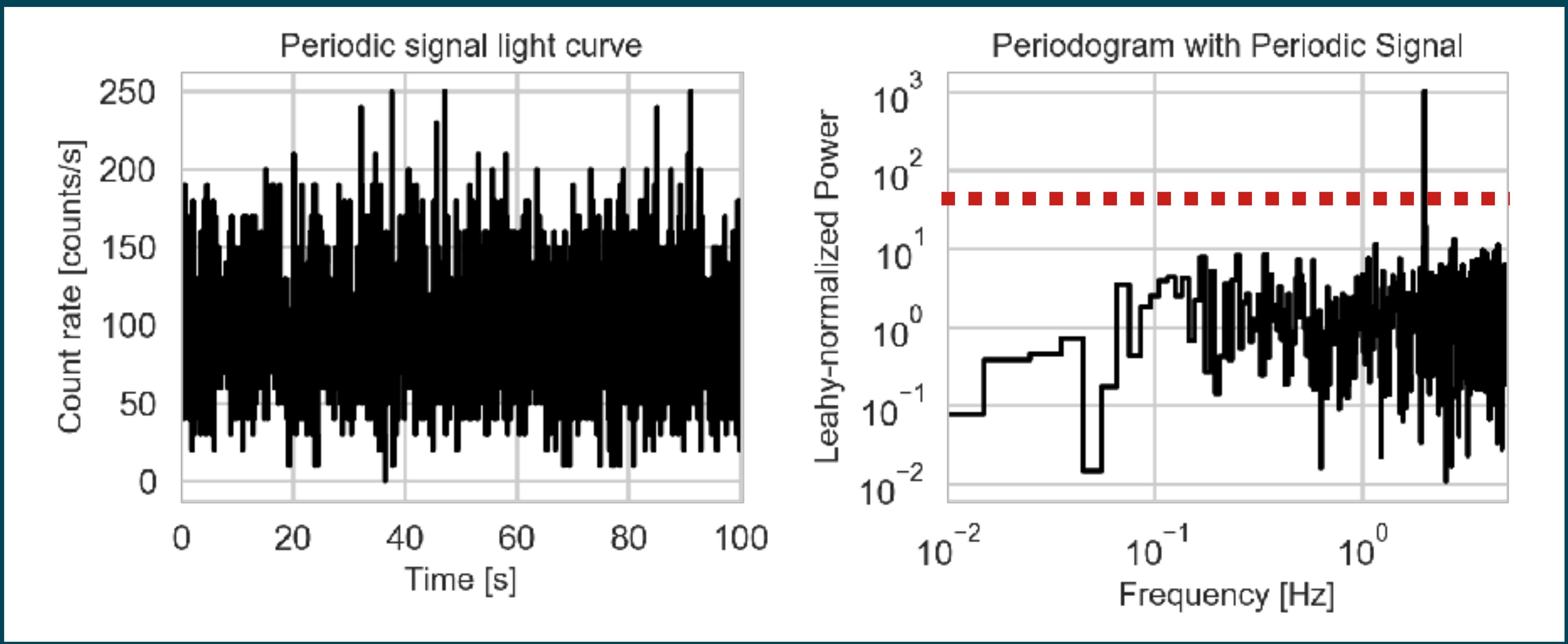
*for evenly sampled time series



periodic signal

Fourier Analysis tells us about variability!*

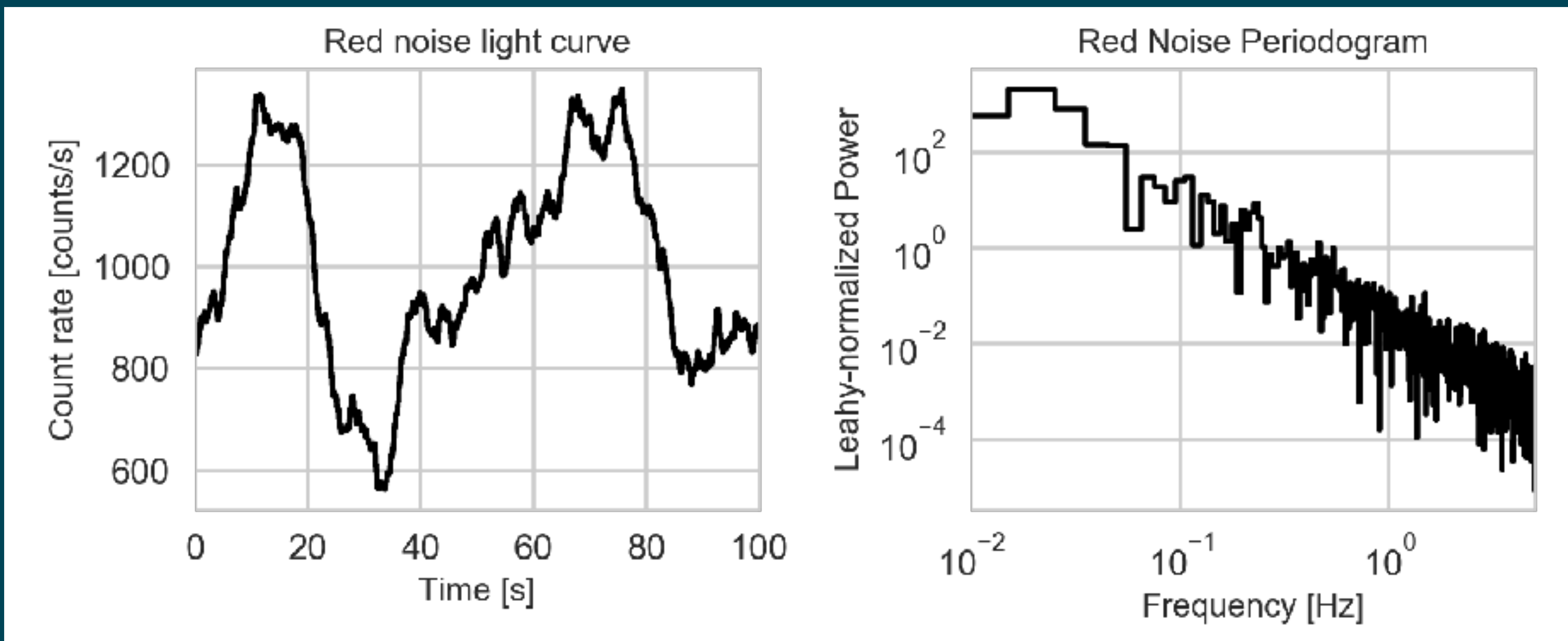
*for evenly sampled time series



periodic signal

Fourier Analysis tells us about variability!*

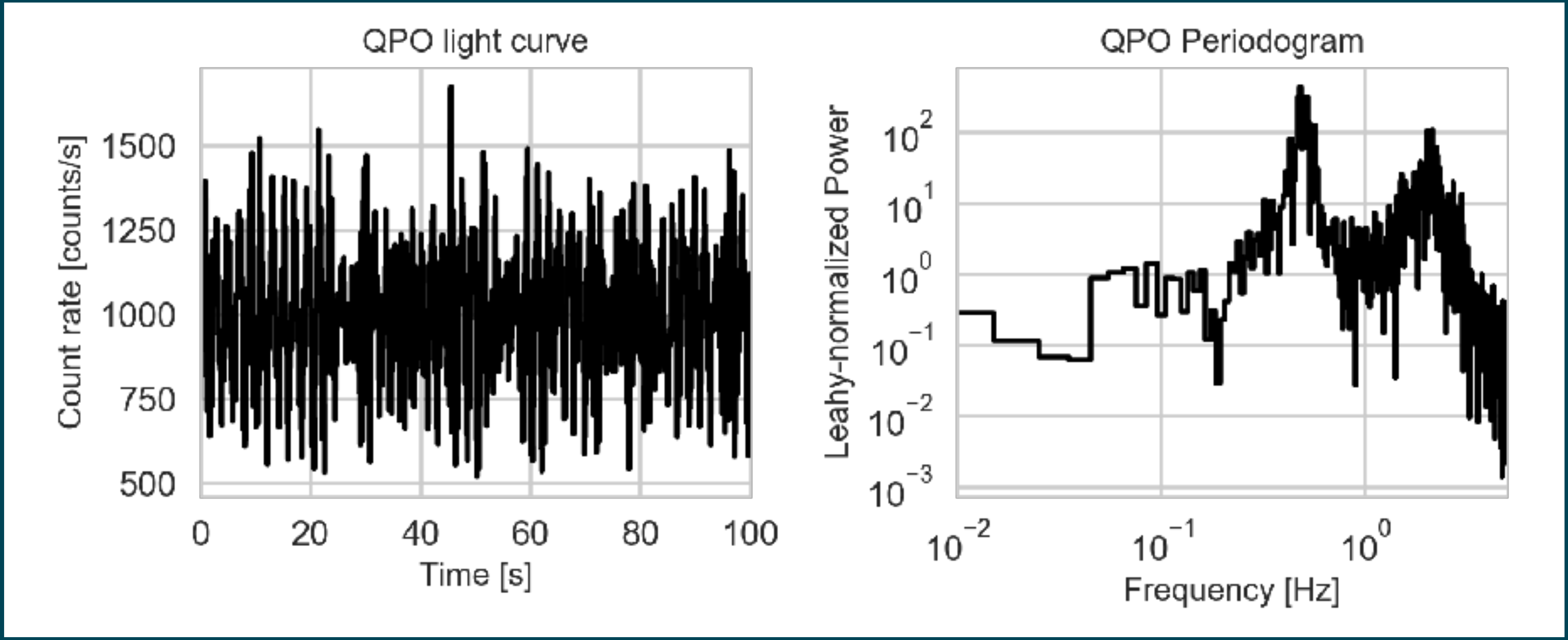
*for evenly sampled time series



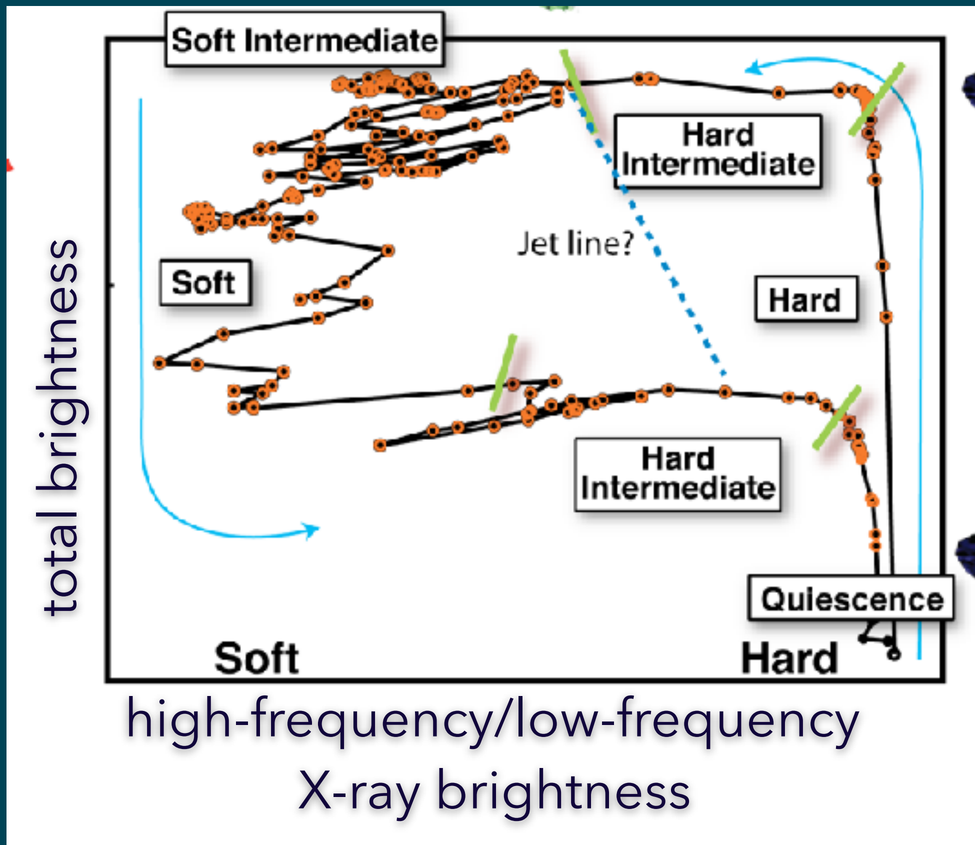
red noise

Fourier Analysis tells us about variability!*

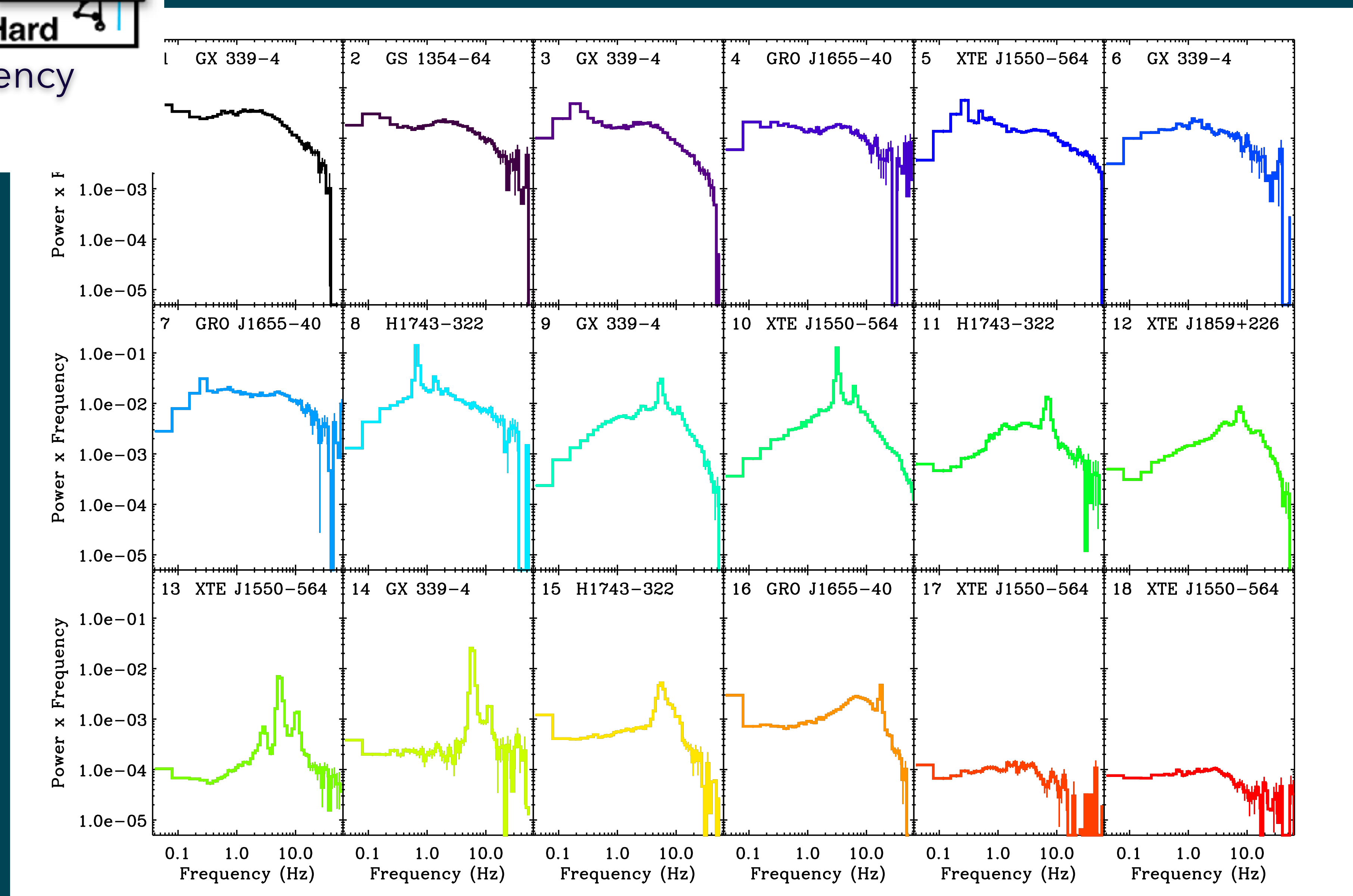
*for evenly sampled time series



quasi-periodic oscillations



Credit: Sera Markoff

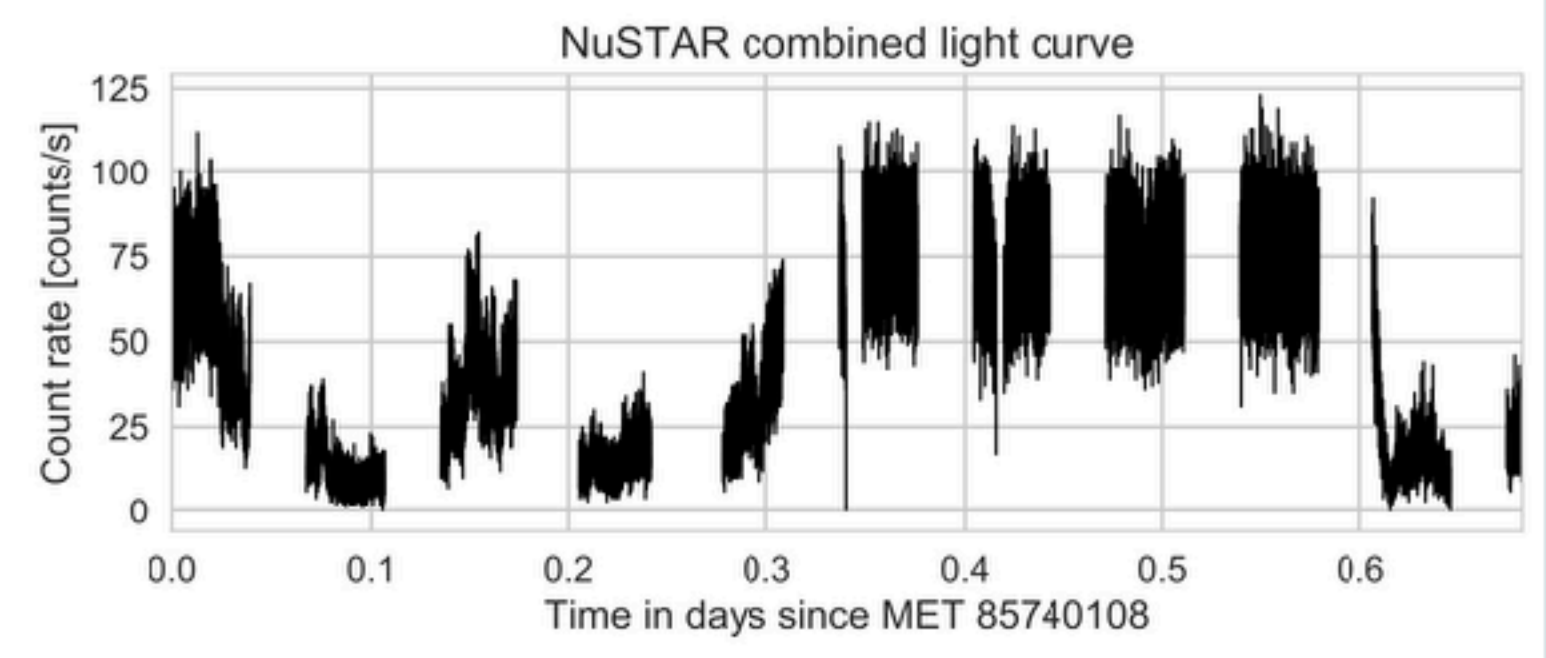


Heil et al 2014

So, we’re done, right?

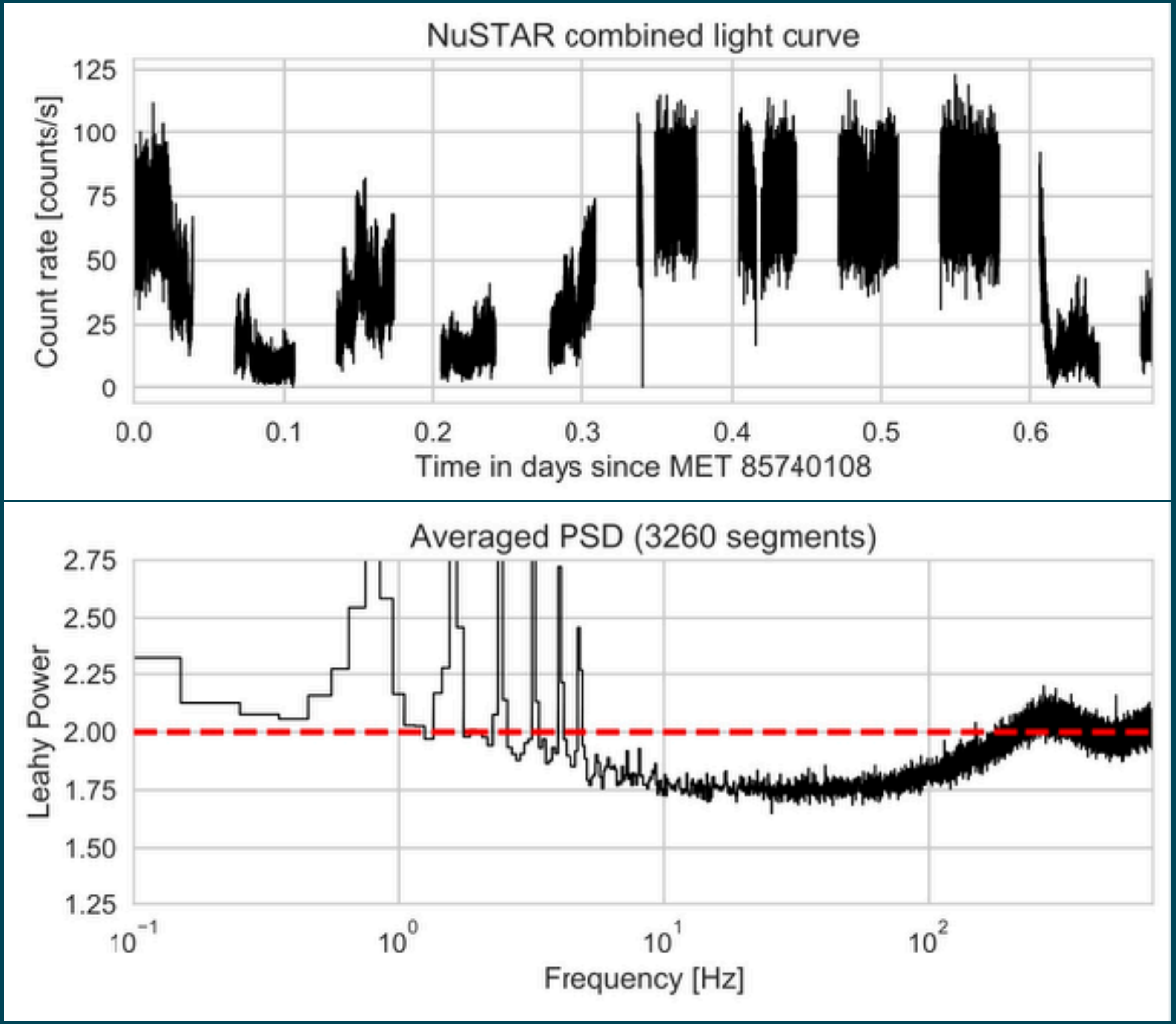
Hercules X-1

Huppenkothen & Bachetti, in press



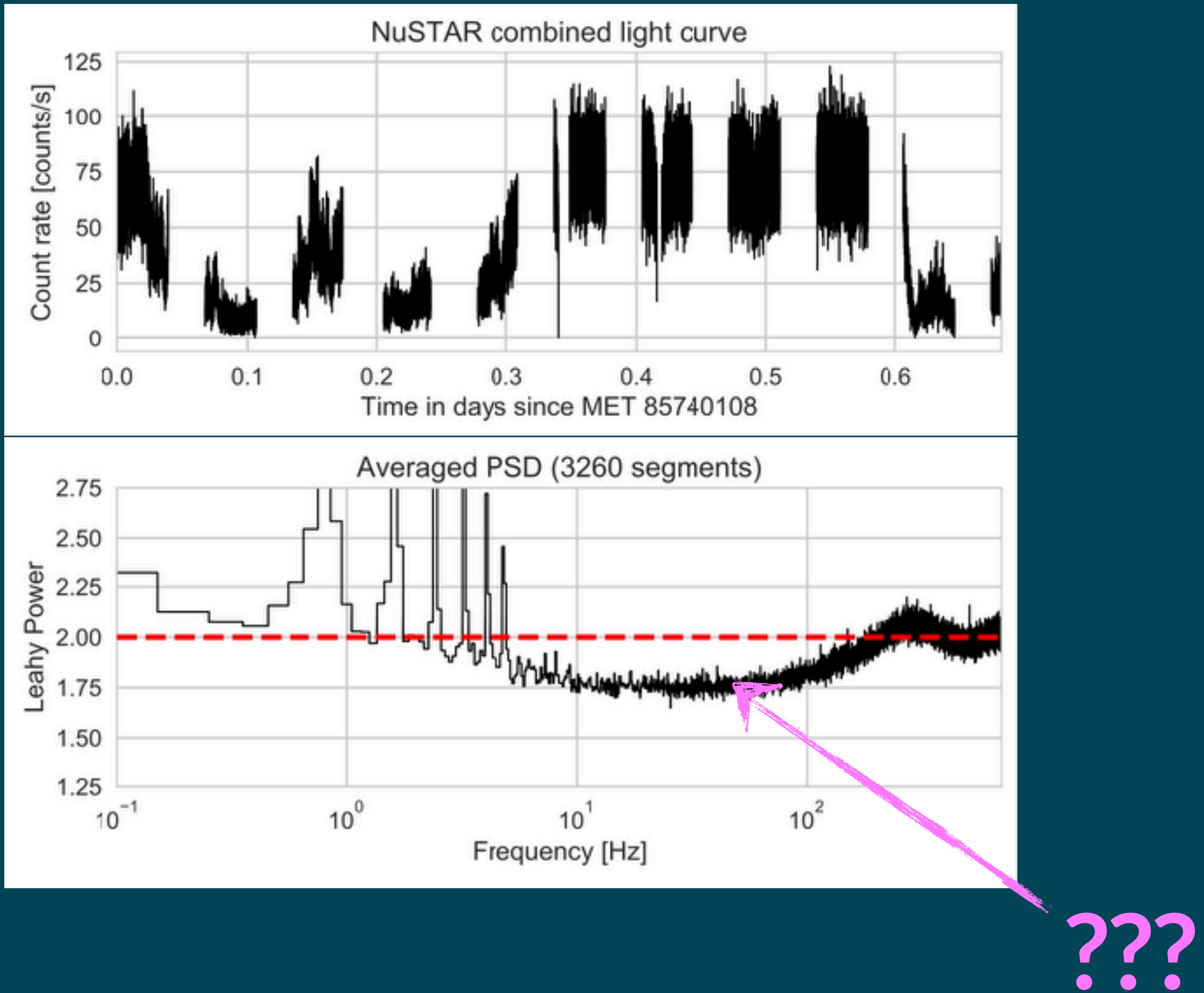
Hercules X-1

Huppenkothen & Bachetti, in press

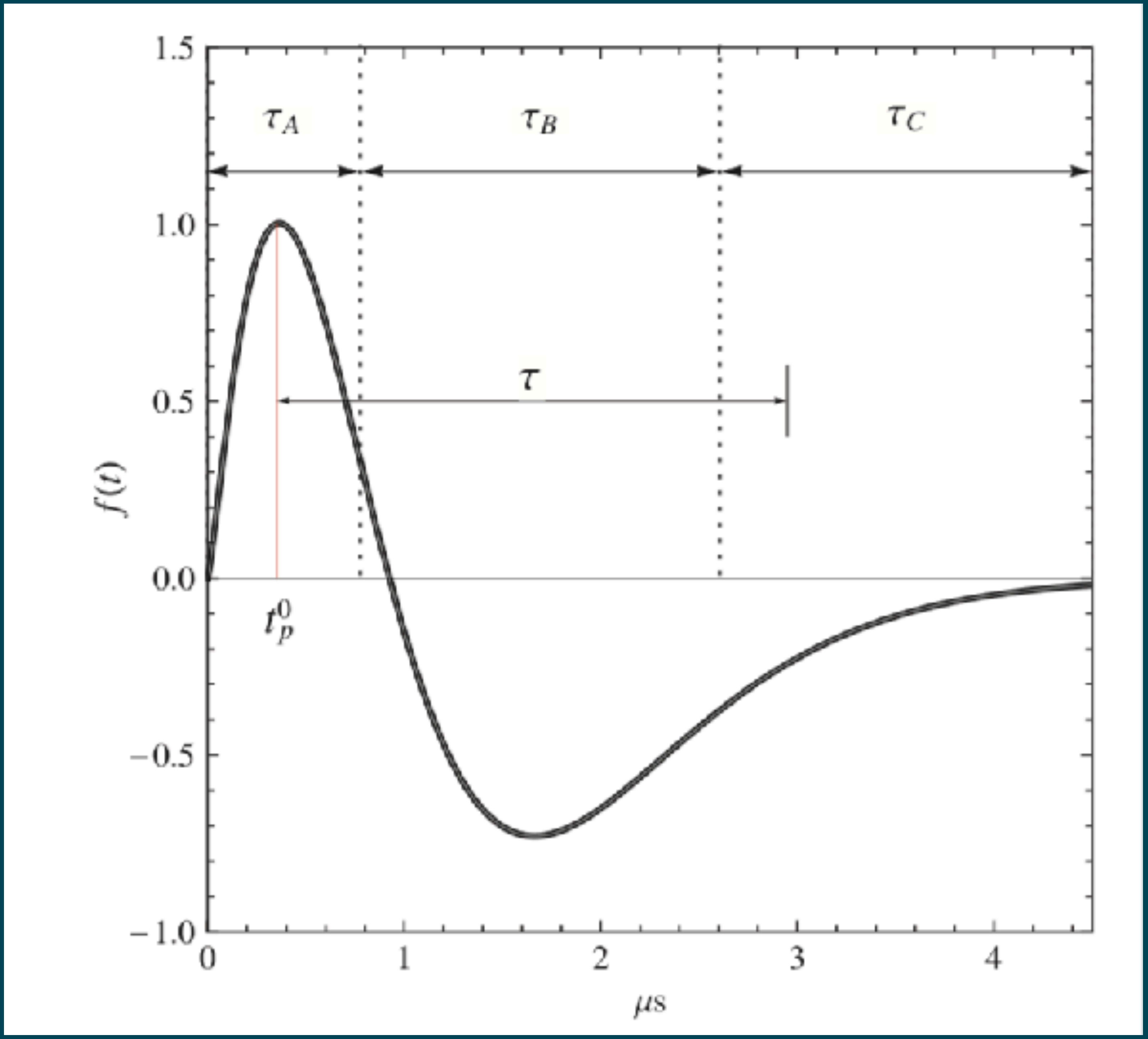


Hercules X-1

Huppenkothen & Bachetti, in press

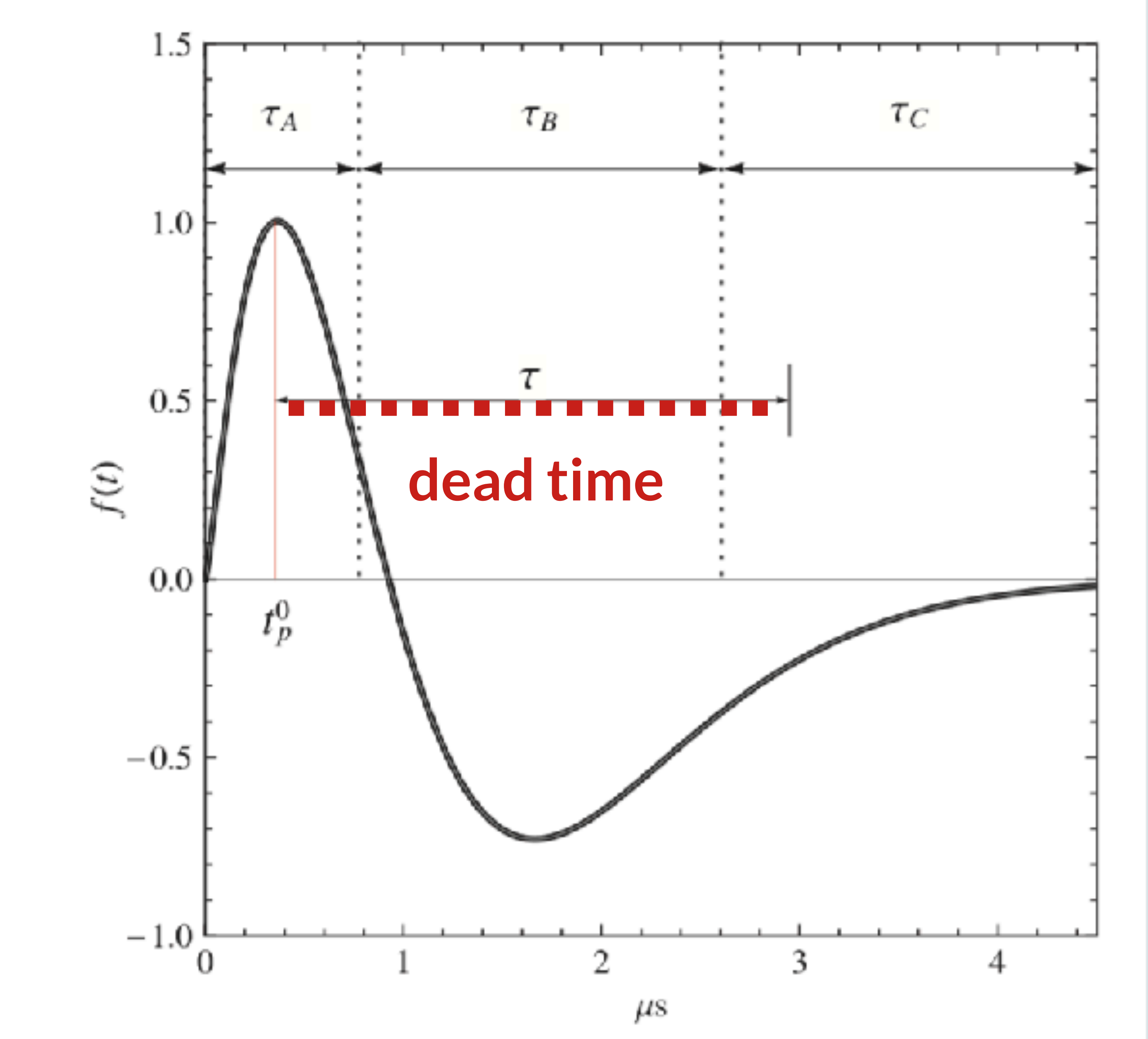


After detection of a photon, the detector is "dead" for a short interval



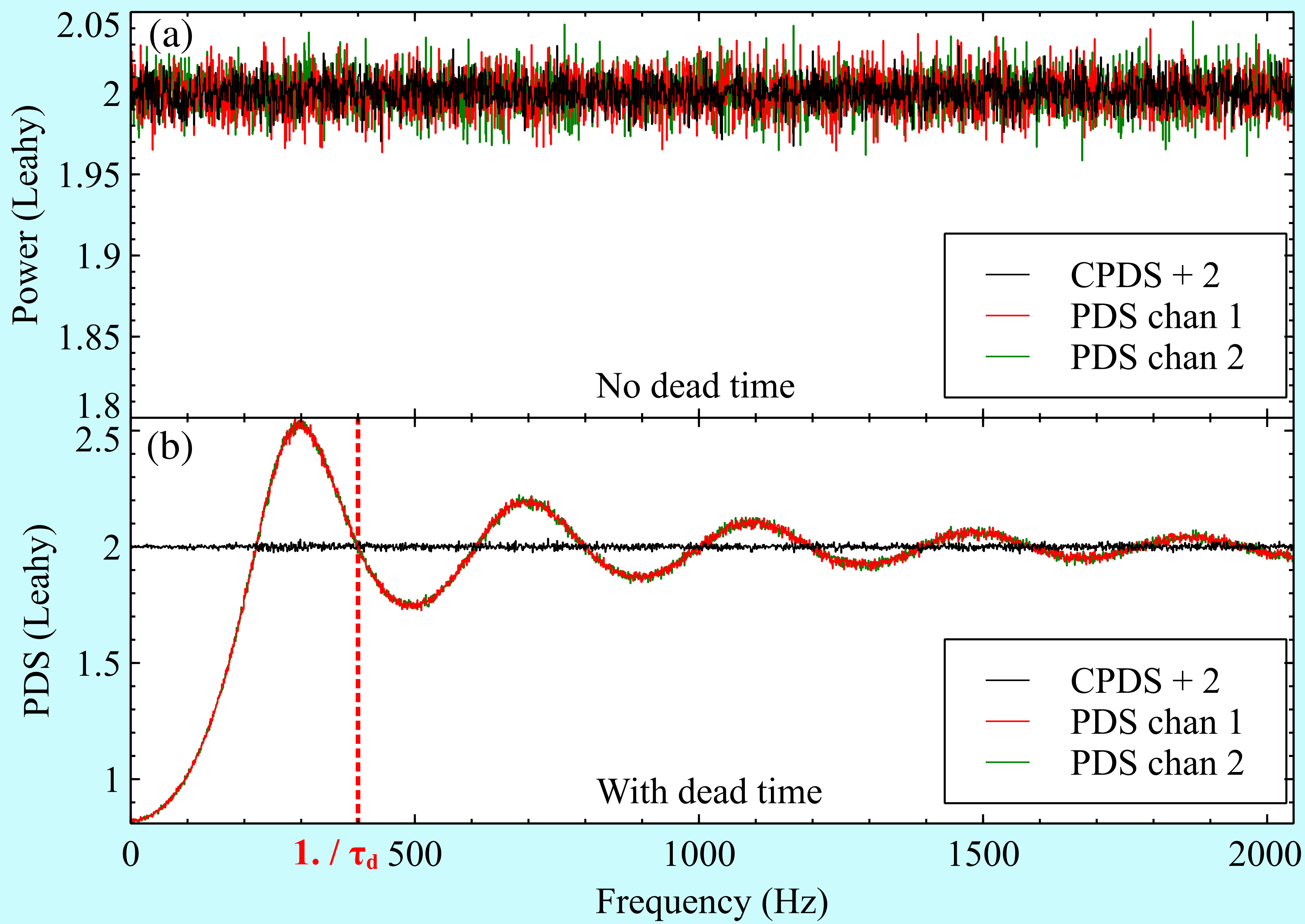
Chaplin et al (2012)

After detection of a photon, the detector is "dead" for a short interval



Chaplin et al (2012)

Dead time affects mean and variance of the periodogram

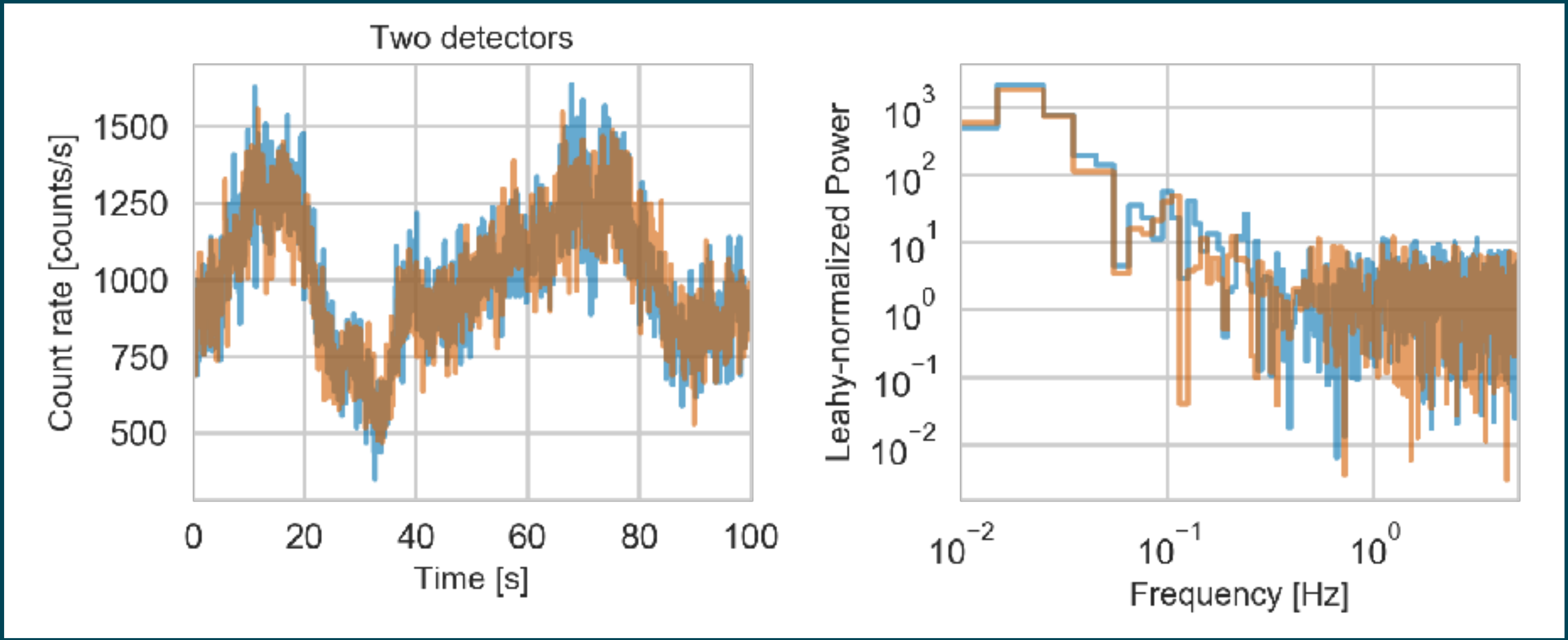


Bachetti et al (2015)

Can we fix this?

Yes (ish)


Having Two Detectors Helps!



Signal is the same, but the measurement noise is different!

Compute cross spectrum instead of periodogram

$$\mathcal{F}_x(j)\mathcal{F}_y^*(j) = \frac{1}{2}\{(A_{xj}A_{yj} + B_{xj}B_{yj}) + i(A_{xj}B_{yj} - A_{yj}B_{xj})\}$$

Compute cross spectrum instead of periodogram

$$\mathcal{F}_x(j)\mathcal{F}_y^*(j) = \frac{1}{2}\left\{(A_{xj}A_{yj} + B_{xj}B_{yj}) + i(A_{xj}B_{yj} - A_{yj}B_{xj})\right\}$$

phase/time lag

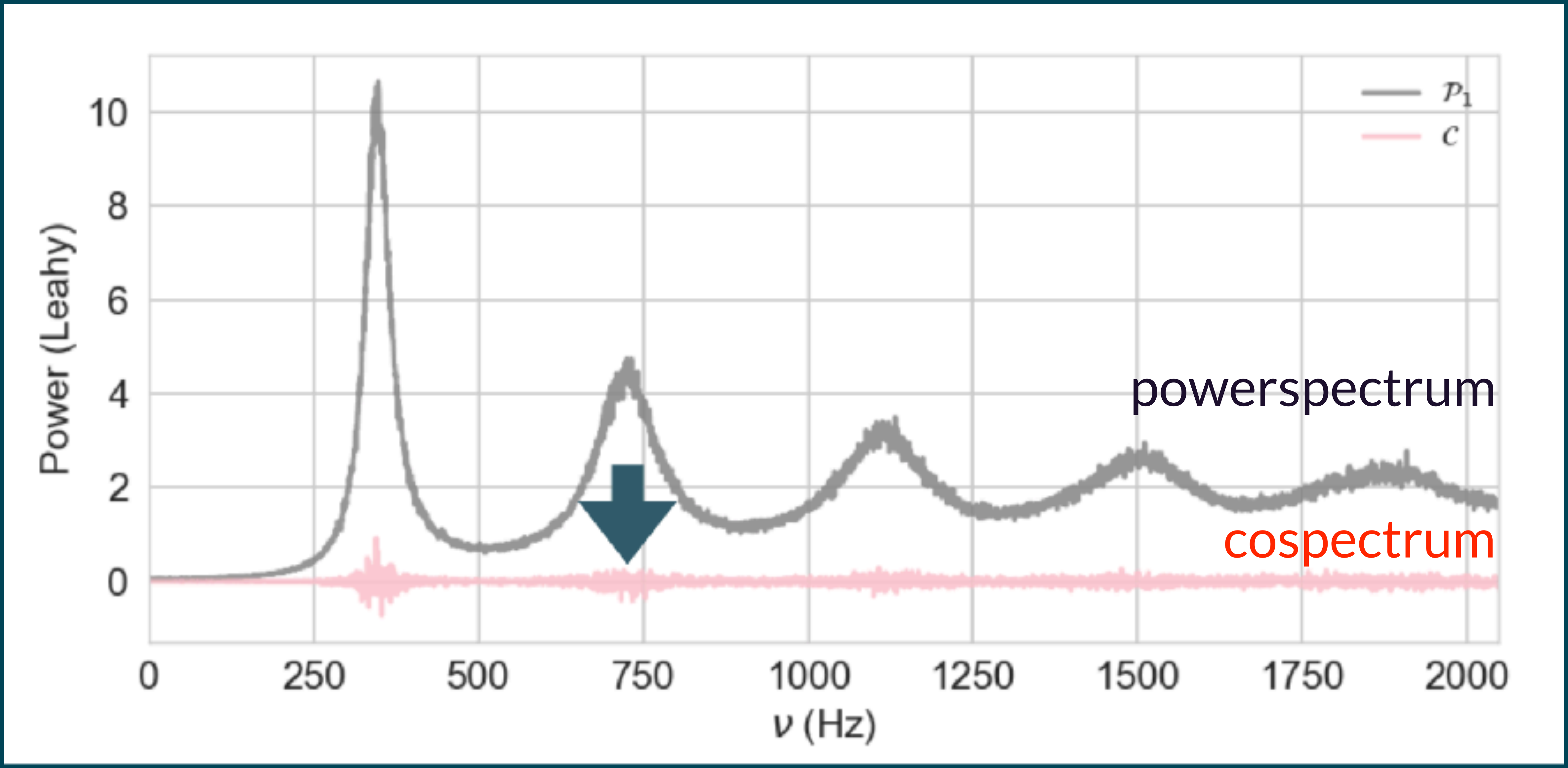
Compute cross spectrum instead of periodogram

$$\mathcal{F}_x(j)\mathcal{F}_y^*(j) = \frac{1}{2}\{(A_{xj}A_{yj} + B_{xj}B_{yj}) + i(A_{xj}B_{yj} - A_{yj}B_{xj})\}$$

phase/time lag

$$C_j = \frac{1}{2}(A_{xj}A_{yj} + B_{xj}B_{yj}) \, .$$

cospectrum



Power (Leahy)
10
8
6
4
2
0
0
250
500
750
1000
1250
1500
1750
2000
ν (Hz)
$\mathcal{P}_1$
$\mathcal{C}$
powerspectrum
cospectrum

What is the statistical distribution of the cospectrum?

periodogram

periodogram

$$|a_j|^2 = A^2 + B^2$$

periodogram

$$|a_j|^2 = A^2 + B^2$$

← χ^2 distributed

periodogram

$$|a_j|^2 = A^2 + B^2$$

← χ^2 distributed

cospectrum

$$C_j = \frac{1}{2}(A_{xj}A_{yj} + B_{xj}B_{yj})\,.$$

periodogram

$$|a_j|^2 = A^2 + B^2$$

← χ^2 distributed

cospectrum

$$C_j = \frac{1}{2}(A_{xj}A_{yj} + B_{xj}B_{yj}) \,.$$

$$A_{xj}^2 \neq A_{xj}A_{yj}$$

periodogram

$$|a_j|^2 = A^2 + B^2$$

← χ^2 distributed

cospectrum

$$C_j = \frac{1}{2}(A_{xj}A_{yj} + B_{xj}B_{yj})\,.$$

$$A_{xj}^2 \neq A_{xj}A_{yj}$$

← not χ^2 distributed

we know

$$A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$$

we know

$$A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$$

set

$$Z = A_{xj}A_{yj} \ , \ Q = B_{xj}B_{yj}$$

we know

$$A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$$

set

$$Z = A_{xj}A_{yj} \ , \ Q = B_{xj}B_{yj}$$

then the PDF is:

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x \sigma_y}\right)}{\pi \sigma_x \sigma_y},$$

Watson (1922); Wishart & Bartlett (1932)

we know

$$A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$$

set

$$Z = A_{xj}A_{yj} \;,\; Q = B_{xj}B_{yj}$$

Bessel function

then the PDF is:

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y},$$

Watson (1922); Wishart & Bartlett (1932) we know $A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$

set $Z = A_{xj}A_{yj}\ ,\ Q = B_{xj}B_{yj}$

Bessel function

then the PDF is: $$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y},$$

Watson (1922); Wishart & Bartlett (1932)

the distribution for $Z + Q$

we know $A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$

set $Z = A_{xj}A_{yj}\ ,\ Q = B_{xj}B_{yj}$

Bessel function

then the PDF is:

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x \sigma_y}\right)}{\pi \sigma_x \sigma_y},$$

Watson (1922); Wishart & Bartlett (1932)

the distribution for $Z + Q$

is the convolution of PDFs: $p_{Z+Q}(z) = p_Z * p_Q(z)$

combine

$$p_{Z+Q}(z) = p_Z * p_Q(z)$$

combine

$$p_{Z+Q}(z) = p_Z * p_Q(z)$$

and

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x \sigma_y}\right)}{\pi \sigma_x \sigma_y},$$

combine

$$p_{Z+Q}(z) = p_Z * p_Q(z)$$

and

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y},$$

to get

$$p_{Z+Q}(z) = \int_{-\infty}^{+\infty} \frac{K_0\left(\frac{|t|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y} \frac{K_0\left(\frac{|z-t|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y} dt$$

combine

$$p_{Z+Q}(z) = p_Z * p_Q(z)$$

and

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y},$$

to get

$$p_{Z+Q}(z) = \int_{-\infty}^{+\infty} \frac{K_0\left(\frac{|t|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y} \frac{K_0\left(\frac{|z-t|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y} dt$$

“The sum of two random variables is equivalent to the multiplication of its moment-generating functions.

— no astronomer, ever

moment-generating function: $M_Z(t) := \mathbb{E}[e^{tZ}]$

moment-generating function: $M_Z(t) := \mathbb{E}[e^{tZ}]$

multiply MGFs of two variables: $M_S(t) = M_Z(t)M_Q(t)\,.$

moment-generating function: $M_Z(t) := \mathbb{E}[e^{tZ}]$

multiply MGFs of two variables: $M_S(t) = M_Z(t) M_Q(t)\,.$

for the cospectrum: $M_C(t) = M_Z(t) M_Q(t) = \frac{1}{1 - t^2 \sigma_x^2 \sigma_y^2}\,.$

Seijas-Macías & Oliveira (2012)

moment-generating function:

$$M_Z(t) := \mathbb{E}[e^{tZ}]$$

multiply MGFs of two variables:

$$M_S(t) = M_Z(t) M_Q(t) \, .$$

for the cospectrum:

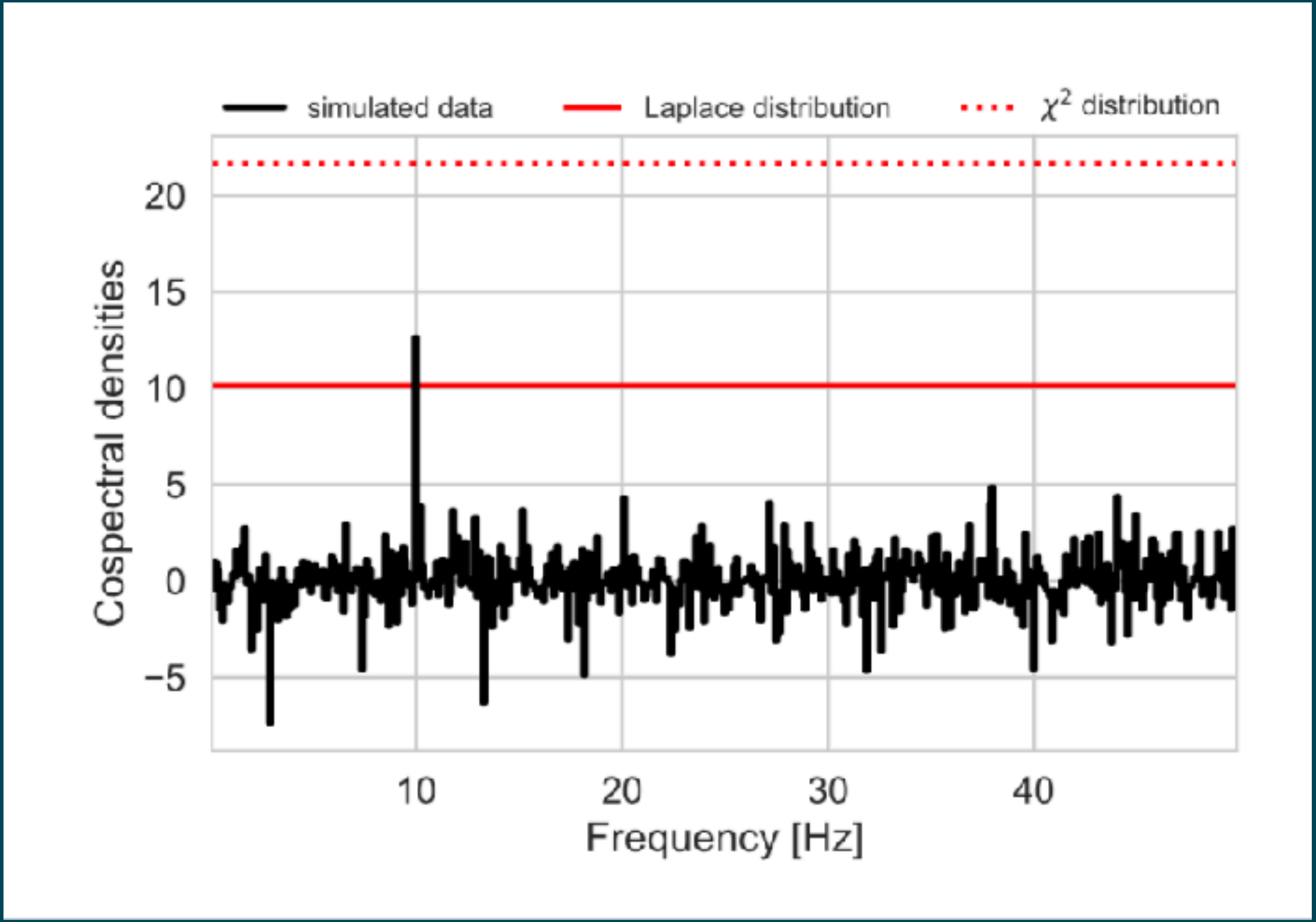
$$M_C(t) = M_Z(t) M_Q(t) = \frac{1}{1 - t^2 \sigma_x^2 \sigma_y^2} \, .$$

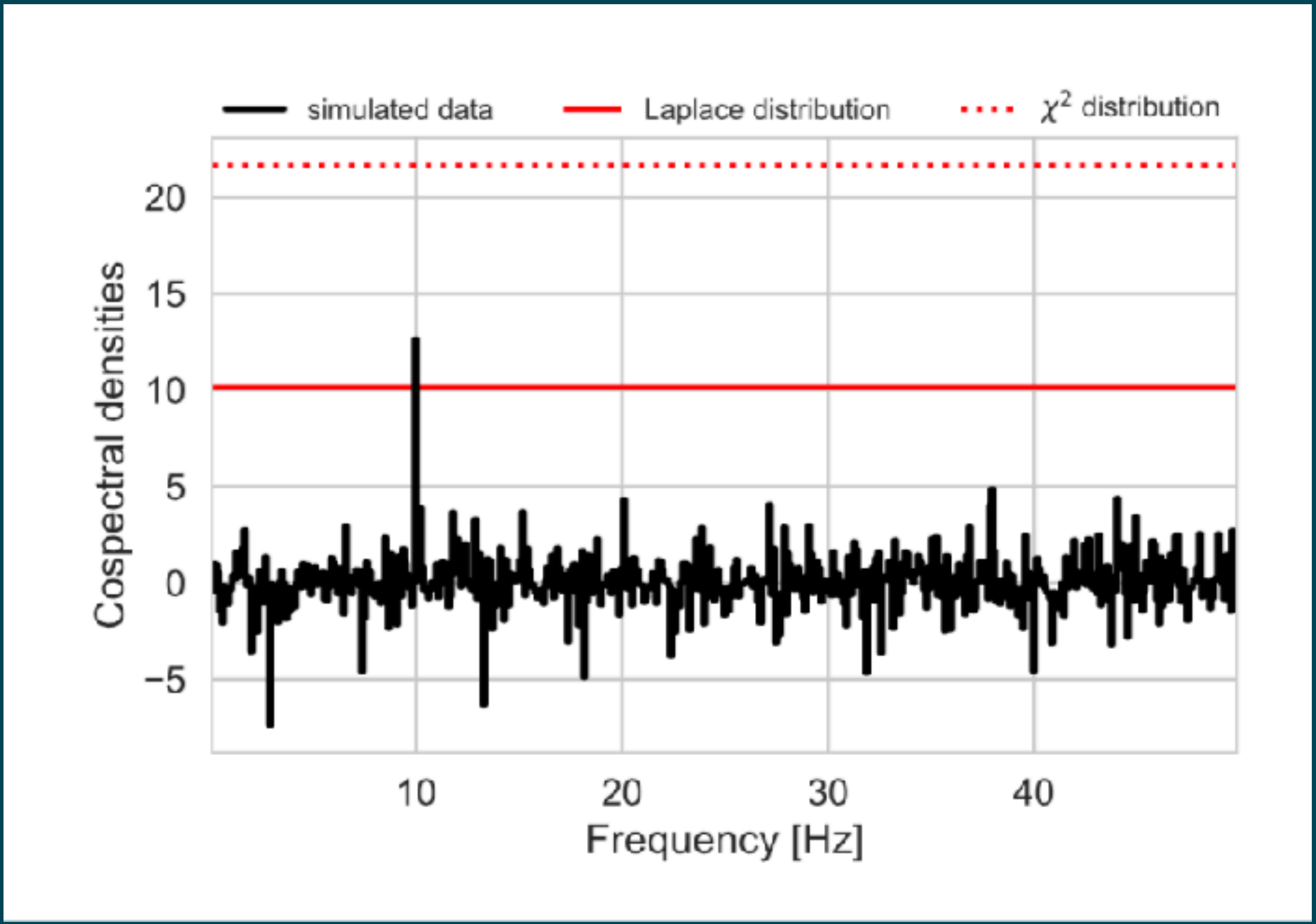
Seijas-Macías & Oliveira (2012)

This is a Laplace distribution!

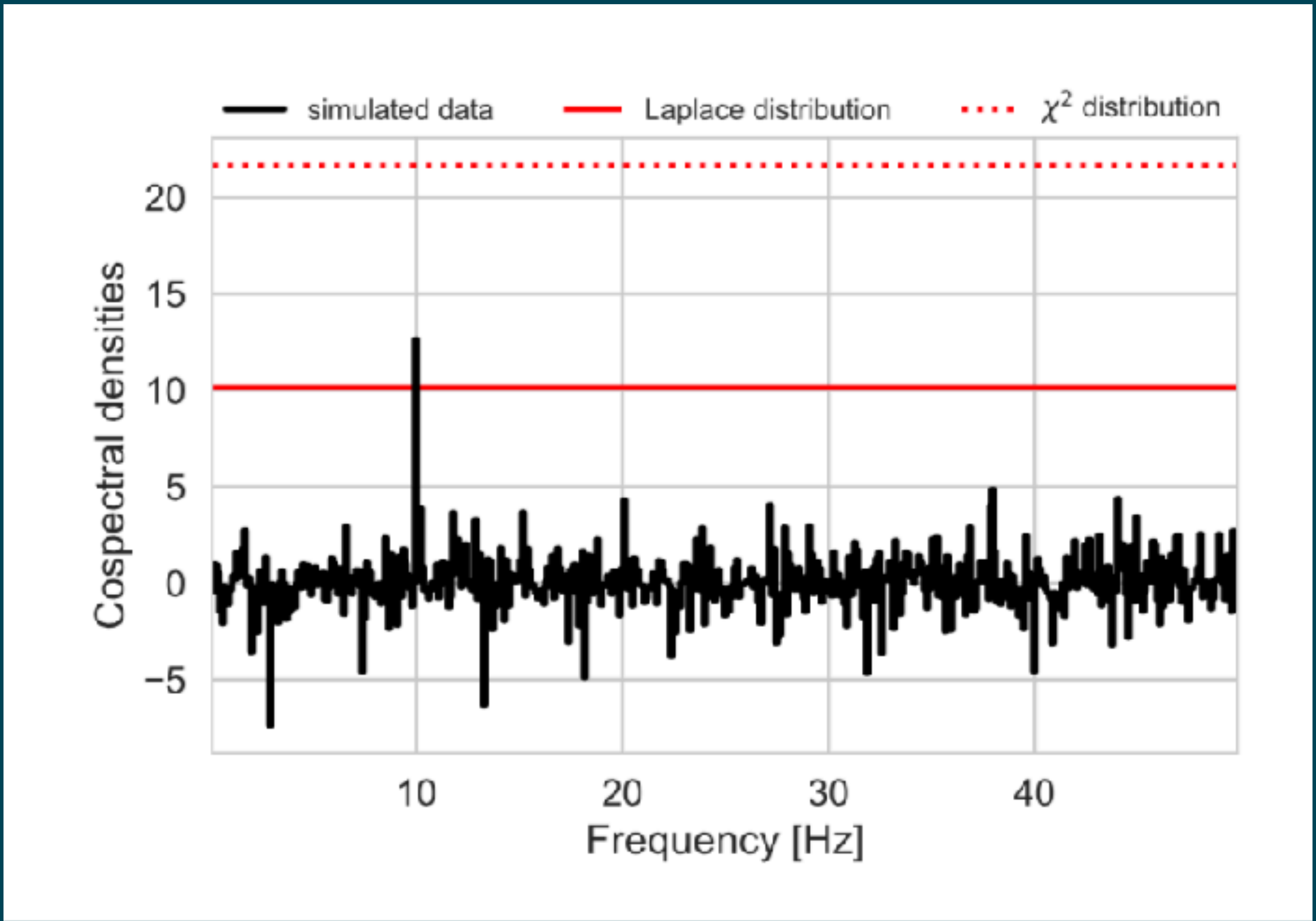
$$p(C_j | 0, \sigma_x \sigma_y) = \frac{1}{\sigma_x \sigma_y} \exp\left(- \frac{|C_j|}{\sigma_x \sigma_y} \right)$$

… why are we doing this again?


simulated data
Laplace distribution
χ^2 distribution
Cospectral densities
20
15
10
5
0
−5
10
20
30
40
Frequency [Hz]

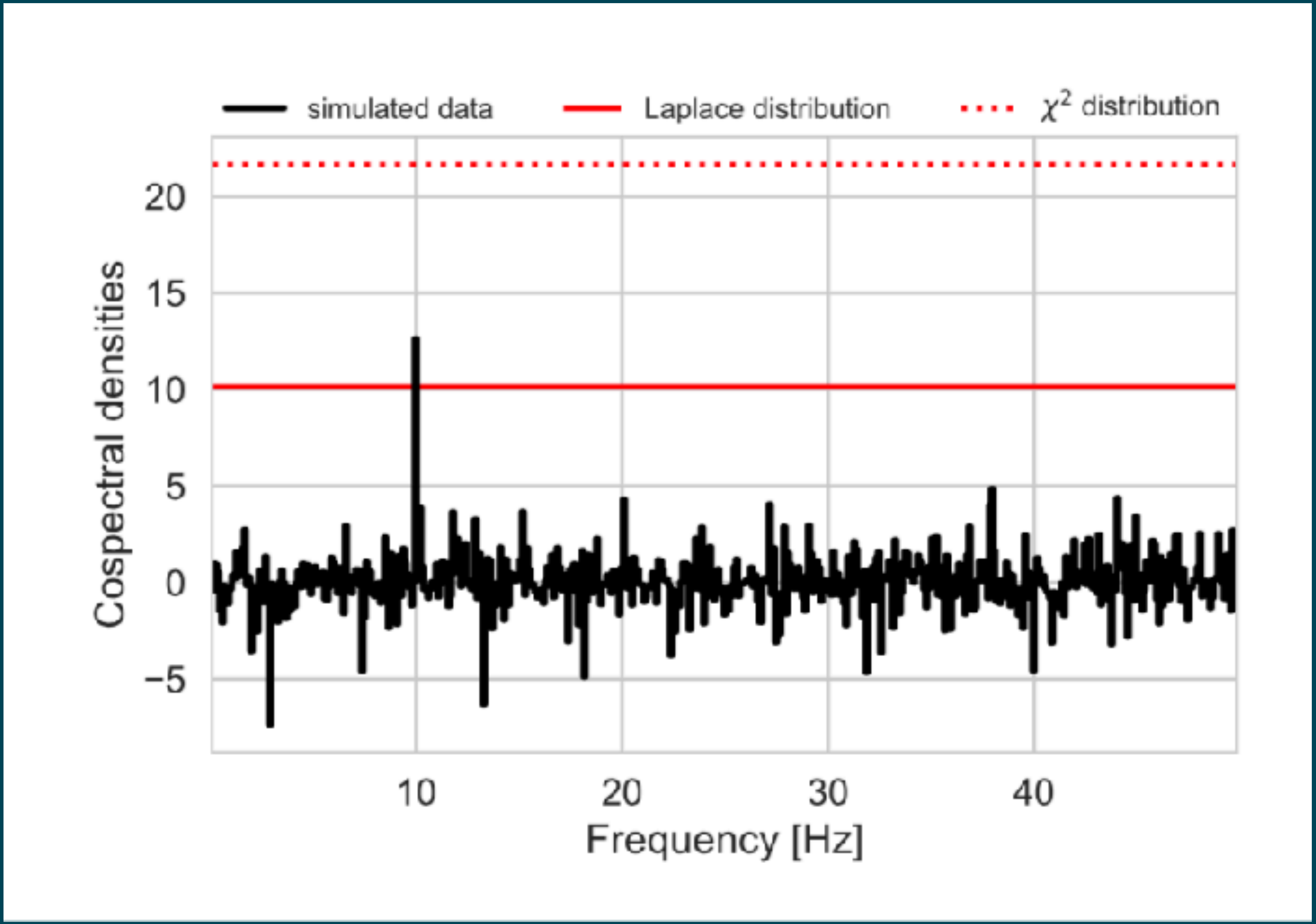


χ^2 distribution



χ^2 distribution

Laplace distribution



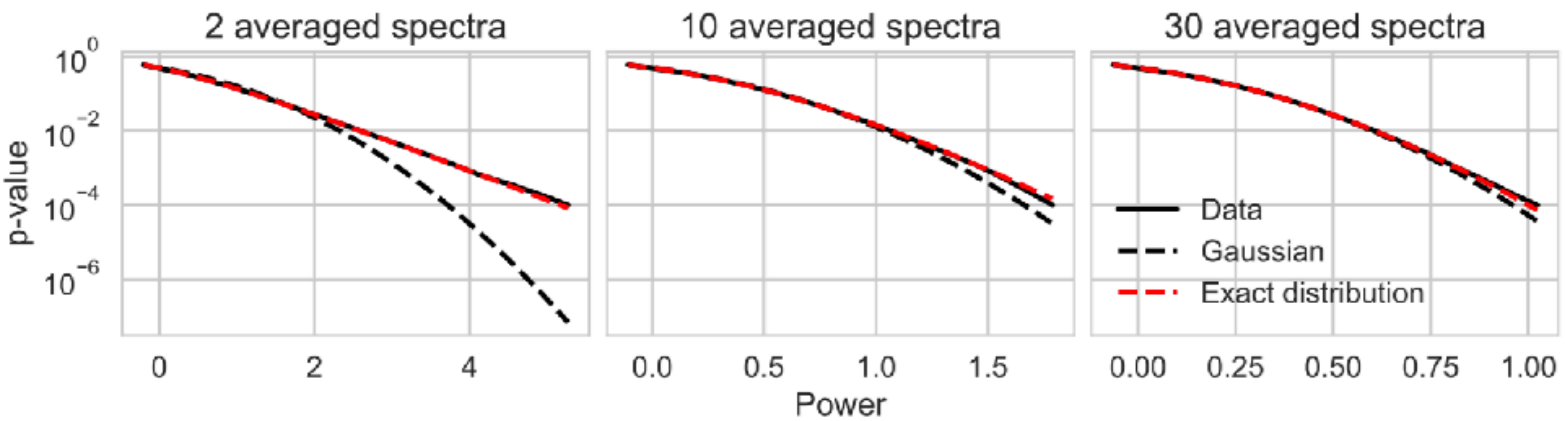
χ^2 distribution

Laplace distribution

significance threshold matters!

what about averaged cospectra?

$$f_{\overline{X}_n}(x) = \frac{ne^{-|nx|}}{(n-1)!2^n} \sum_{j=0}^{n-1} \frac{(n-1+j)!}{(n-1-j)!j!} \frac{|nx|^{n-1-j}}{2^j} \quad , x \in R \, .$$



Gaussian approximation works for large N

so now we're done, right?

so now we're done, right?

... not quite!

1) Equations so far only work for white noise

2) The cospectrum only fixes the mean in the dead time case, not the variance!

What about stochastic variability?

What about stochastic variability?

$$M_{C_j}(t) = \frac{1}{[1 - (1 + r)t][1 + (1 - r)t]}\,.$$

What about stochastic variability?

$$M_{C_j}(t) = \frac{1}{[1-(1+r)t][1+(1-r)t]} \, .$$

correlation coefficient
between Fourier amplitudes

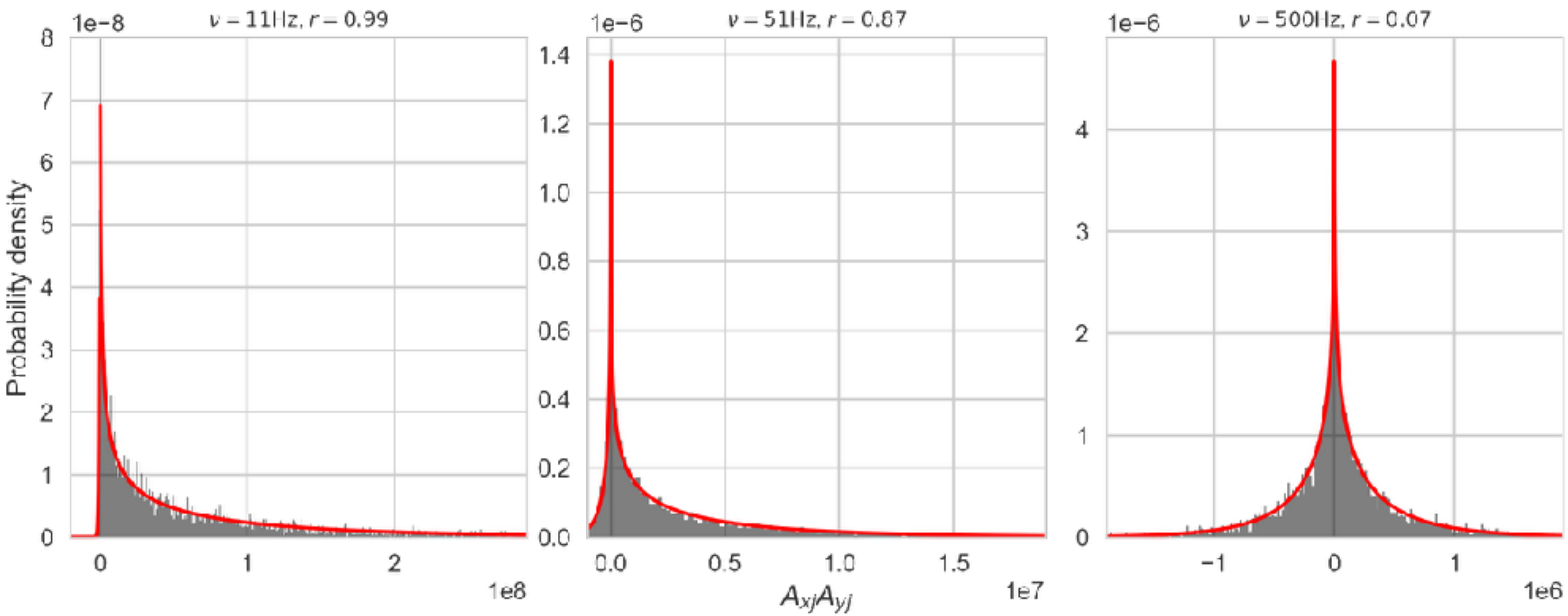
What about stochastic variability?

$$M_{C_j}(t) = \frac{1}{[1-(1+r)t][1+(1-r)t]} \, .$$

↑

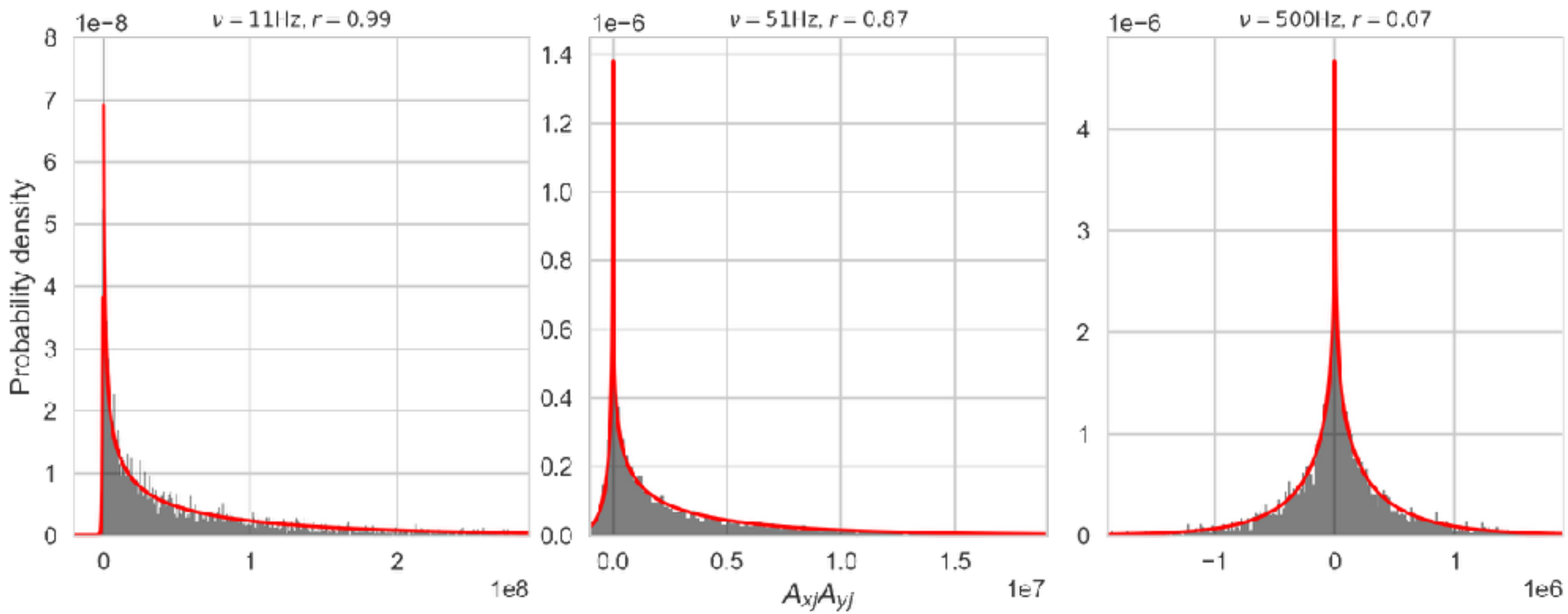
correlation coefficient
between Fourier amplitudes

r = 0 for white noise
r = 1 for power spectra

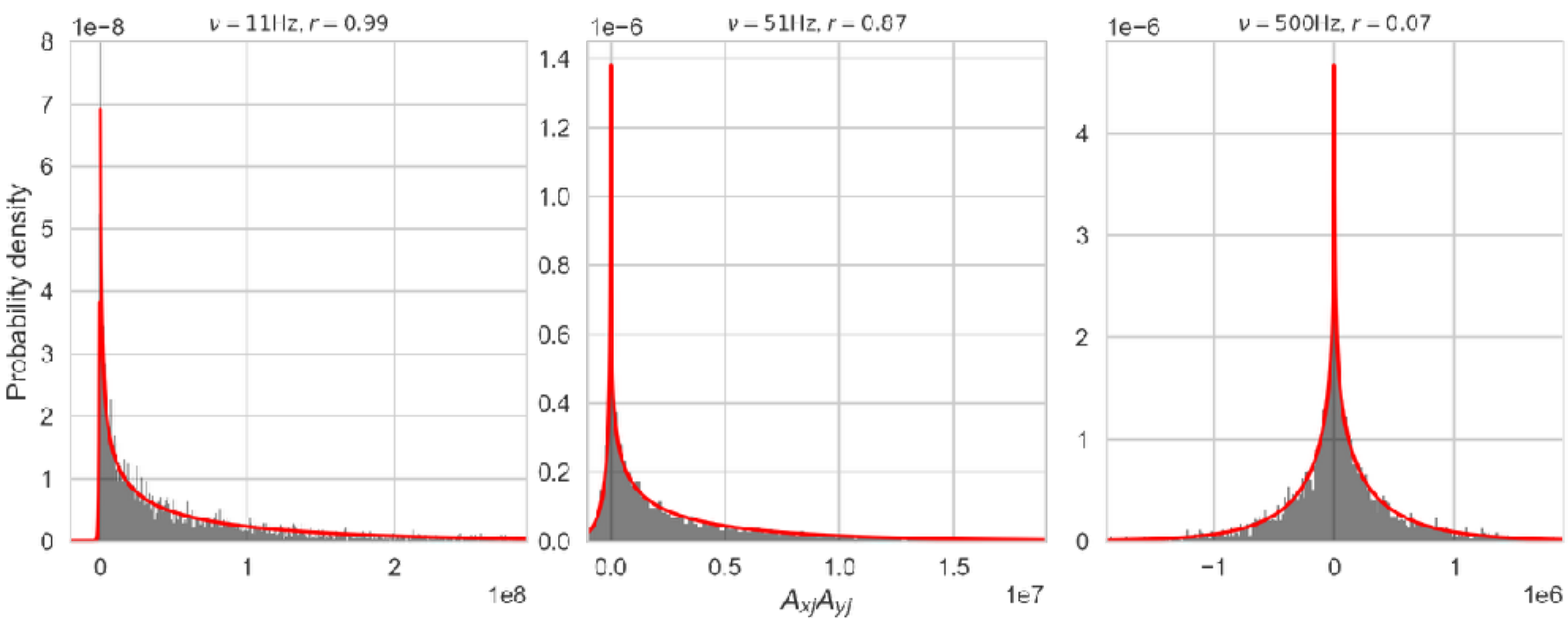


1e−8
ν = 11Hz, r = 0.99
Probability density
1e8
1e−6
ν = 51Hz, r = 0.87
$A_{xj}A_{yj}$
1e7
1e−6
ν = 500Hz, r = 0.07
1e6

statistical distribution depends on r, which depends on power spectral shape!

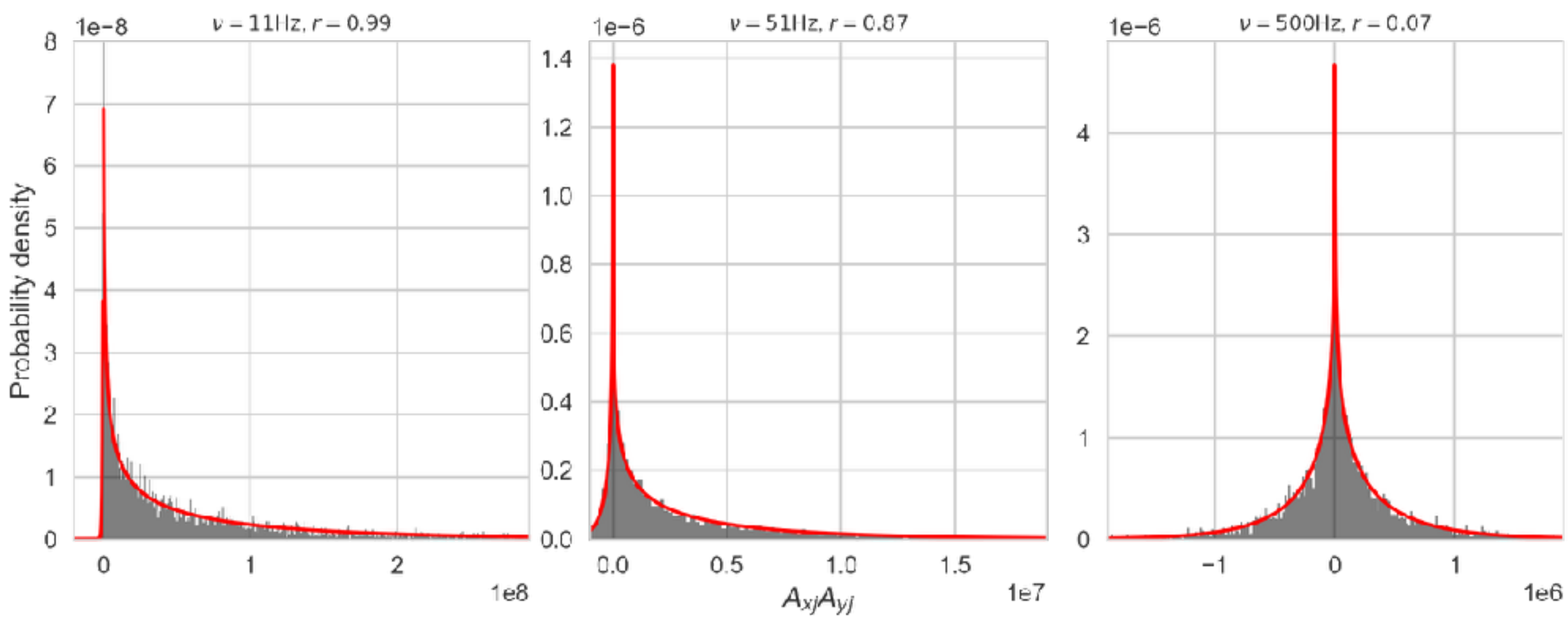


statistical distribution depends on r,
which depends on power spectral shape!



also, there is no existing closed-form
solution for the PDF

statistical distribution depends on r, which depends on power spectral shape!



also, there is no existing closed-form solution for the PDF

Can we model the Fourier amplitudes directly?

Can we model the Fourier amplitudes directly?

$$p(\{A_{xj}, A_{yj}, B_{xj}, B_{yj}\}_{j=0}^{N/2} | \theta, \lambda_{\mathrm{phot}}) = \sum_{j=0}^{N/2} \left[-\log(2\pi|C|) - ([A_{xj}, A_{yj}]^T C^{-1} [A_{xj}, A_{yj}]) - ([B_{xj}, B_{yj}]^T C^{-1} [B_{xj}, B_{yj}]) \right]$$

$$C = \begin{bmatrix} \sigma_{sj} + \sigma n & \sigma_{sj} \\ \sigma_{sj} & \sigma_{sj} + \sigma n \end{bmatrix}$$

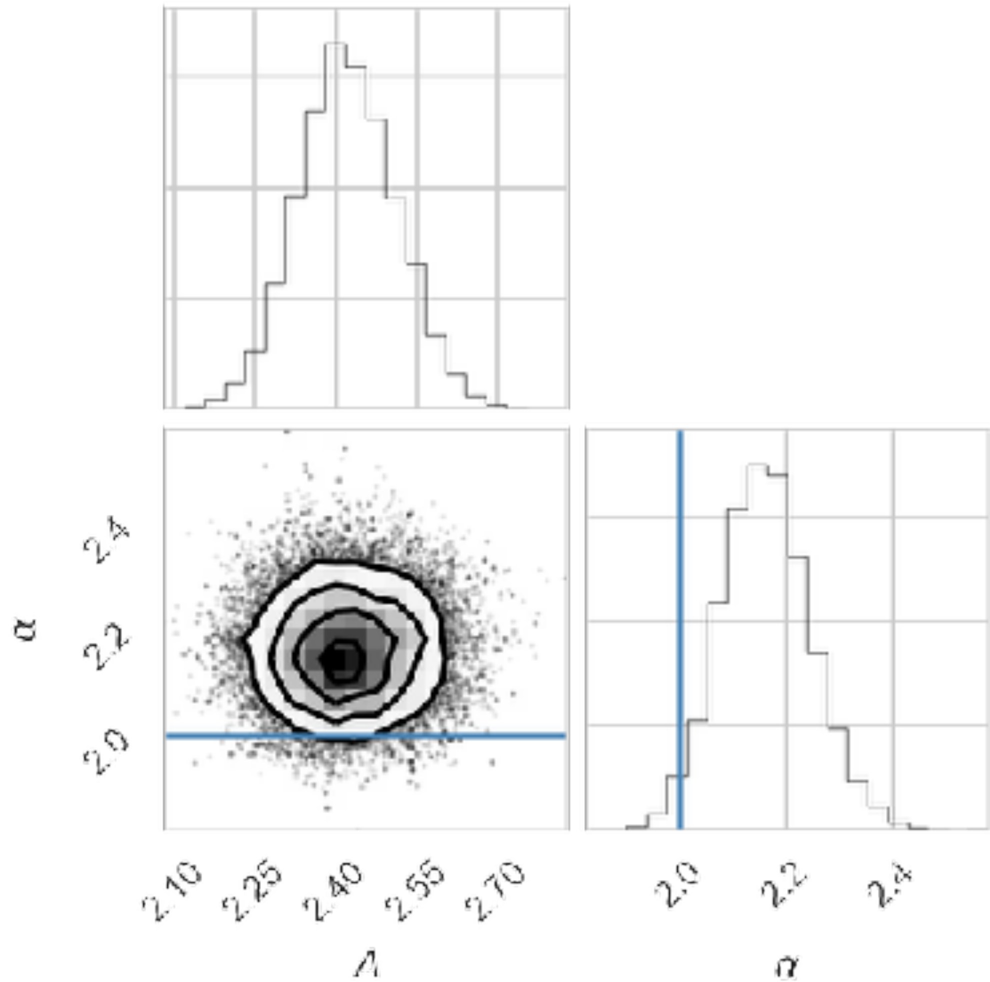
Can we model the Fourier amplitudes directly?

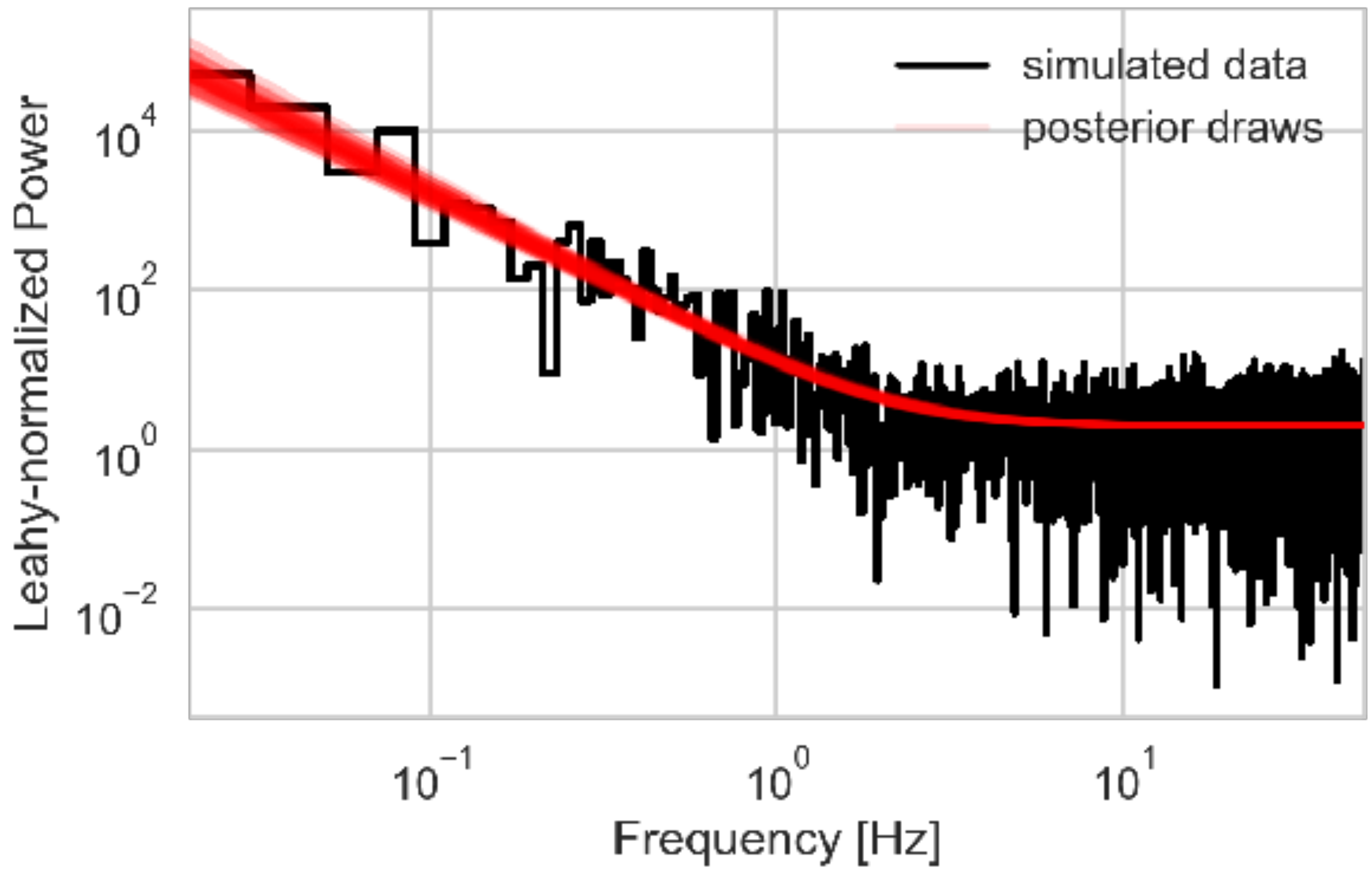
$$p(\{A_{xj}, A_{yj}, B_{xj}, B_{yj}\}_{j=0}^{N/2} | \theta, \lambda_{\mathrm{phot}}) = \sum_{j=0}^{N/2} \left[-\log(2\pi|C|) - ([A_{xj}, A_{yj}]^T C^{-1} [A_{xj}, A_{yj}]) - ([B_{xj}, B_{yj}]^T C^{-1} [B_{xj}, B_{yj}]) \right]$$

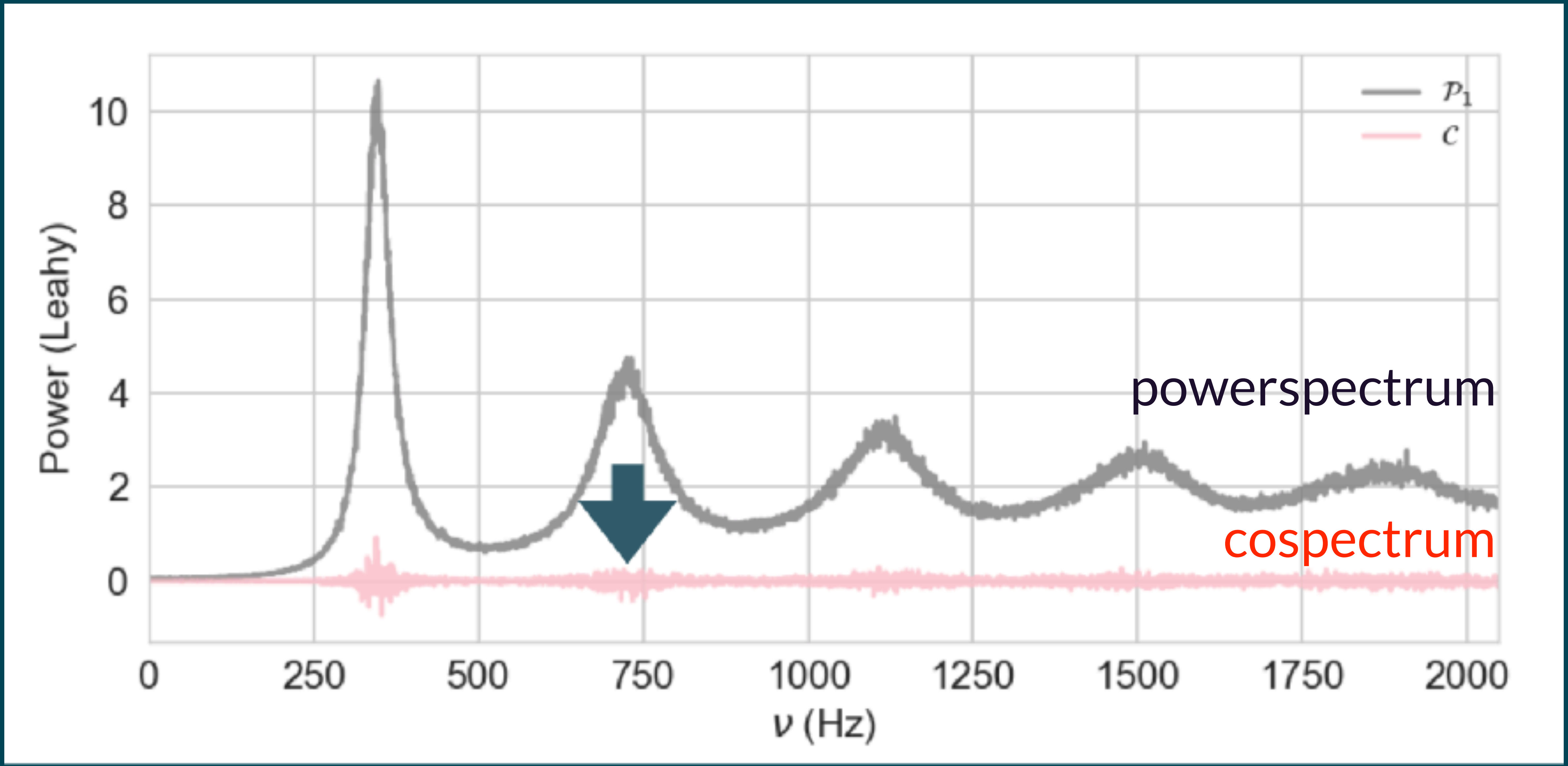
$$C = \begin{bmatrix} \sigma_{sj} + \sigma n & \sigma_{sj} \\ \sigma_{sj} & \sigma_{sj} + \sigma n \end{bmatrix}$$

depends on P(ν)

[work in progress!]





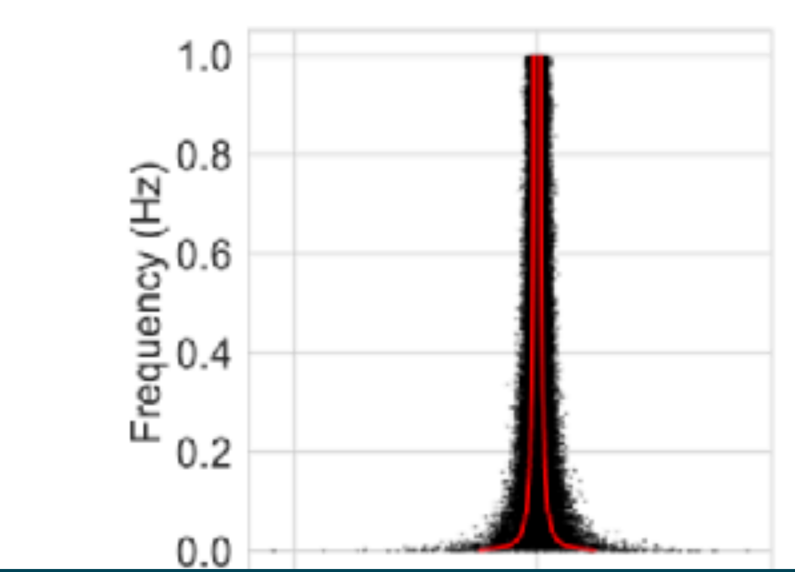


Power (Leahy)
10
8
6
4
2
0
0
250
500
750
1000
1250
1500
1750
2000
ν (Hz)
$\mathcal{P}_1$
$\mathcal{C}$
powerspectrum
cospectrum

But: we can **correct** the periodogram (and the cospectrum) in some cases!

Fourier Amplitude Difference Correction

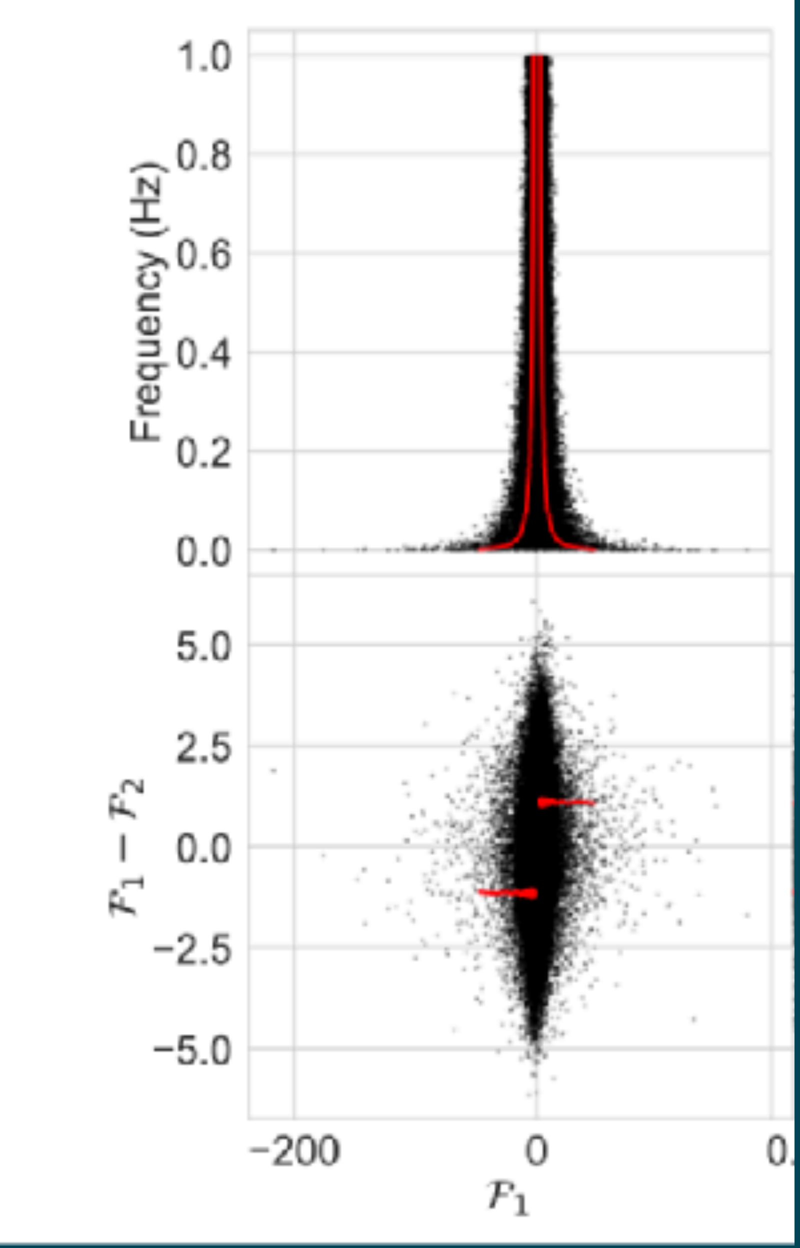
red noise, no dead time



Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

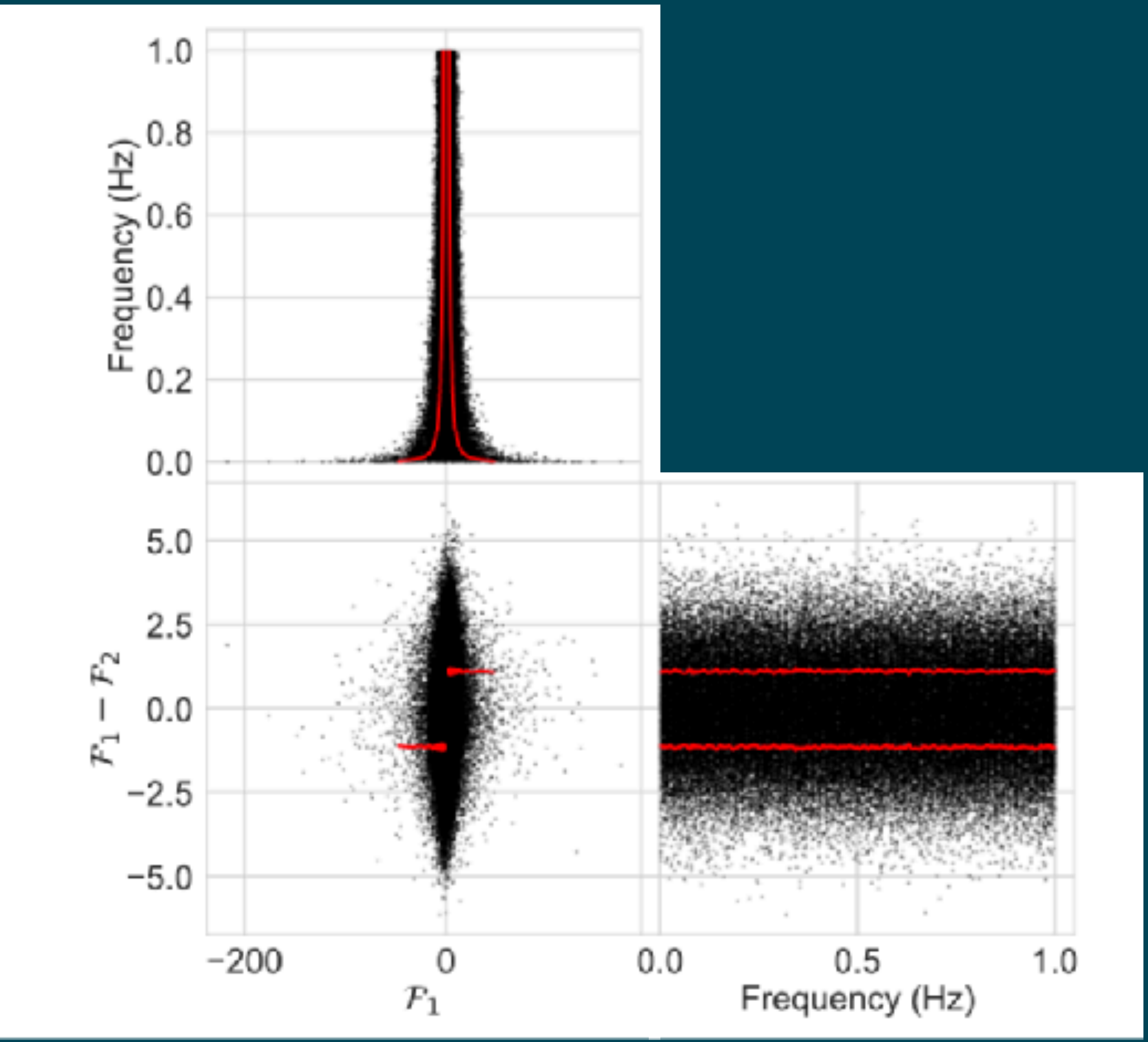
red noise, no dead time



Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

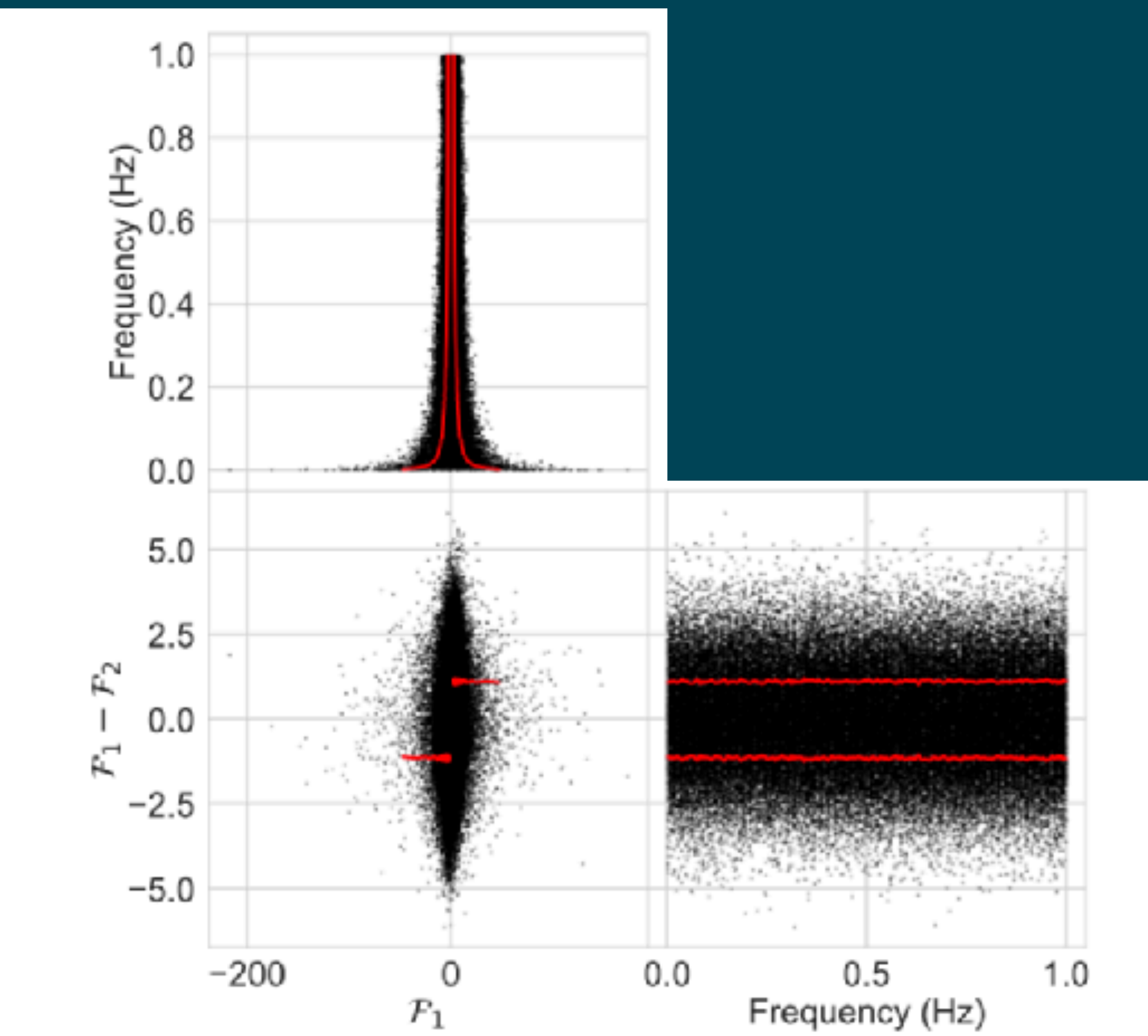
red noise, no dead time



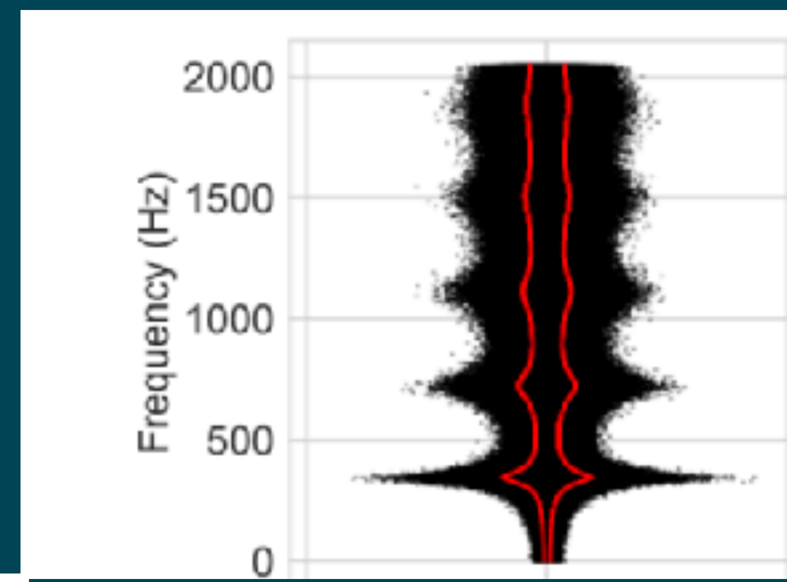
Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

red noise, no dead time



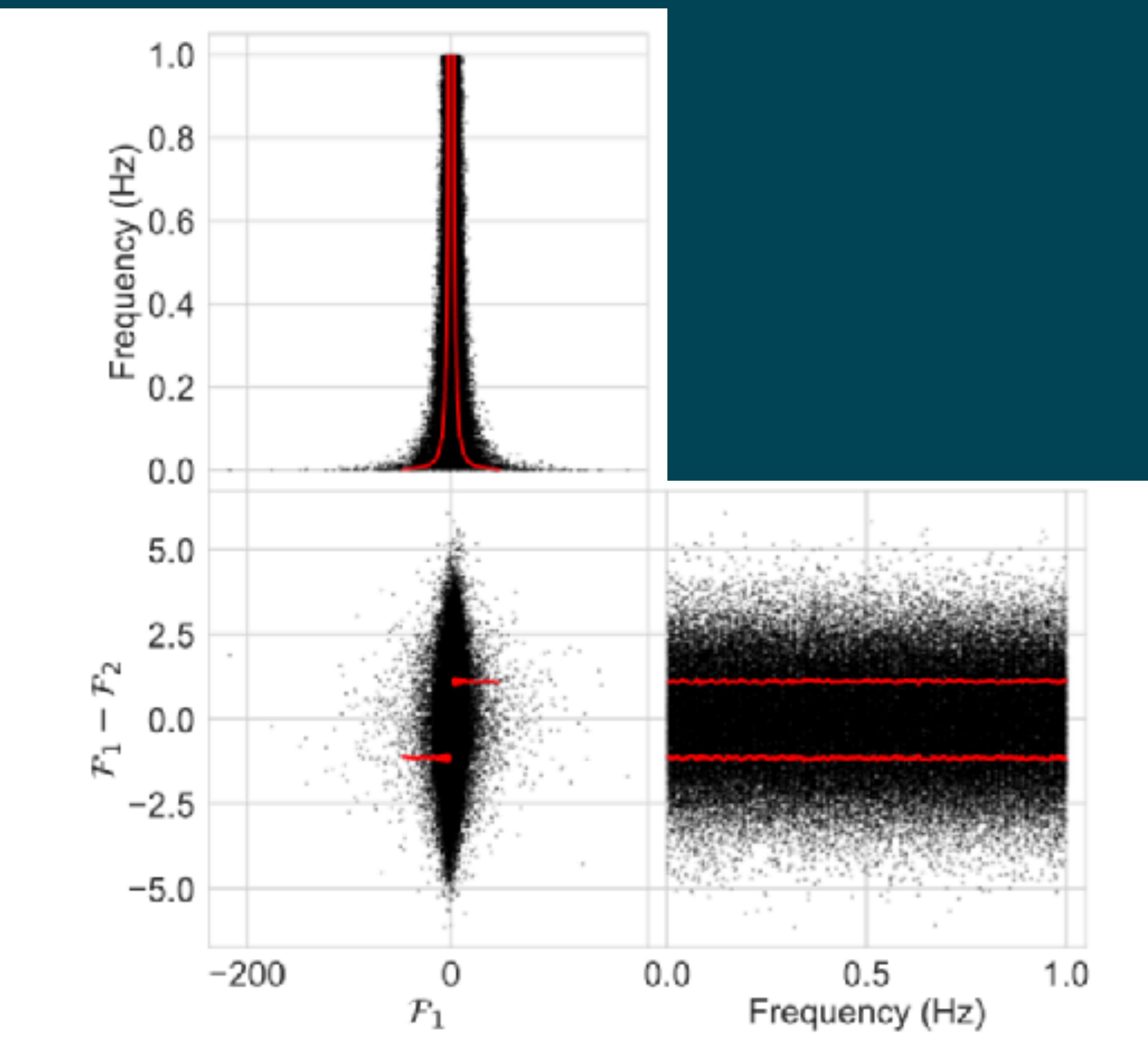
white noise, dead time



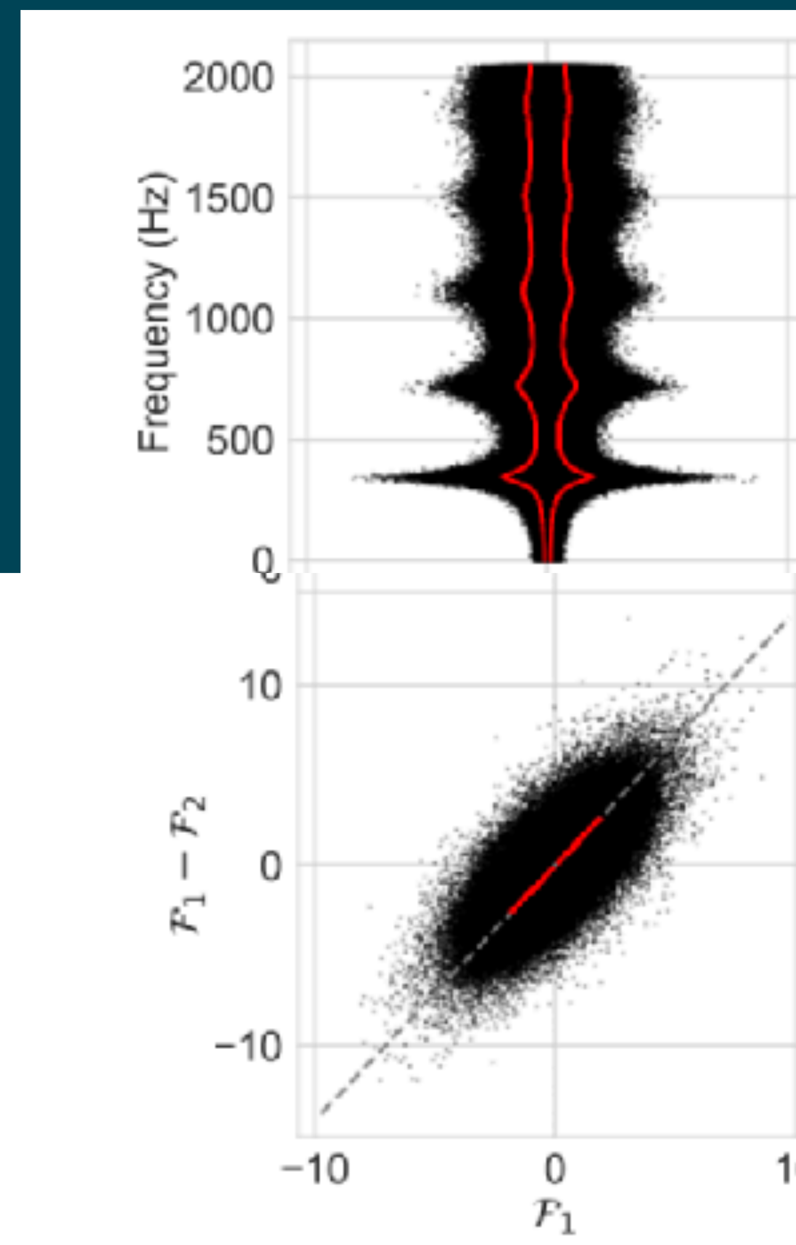
Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

red noise, no dead time



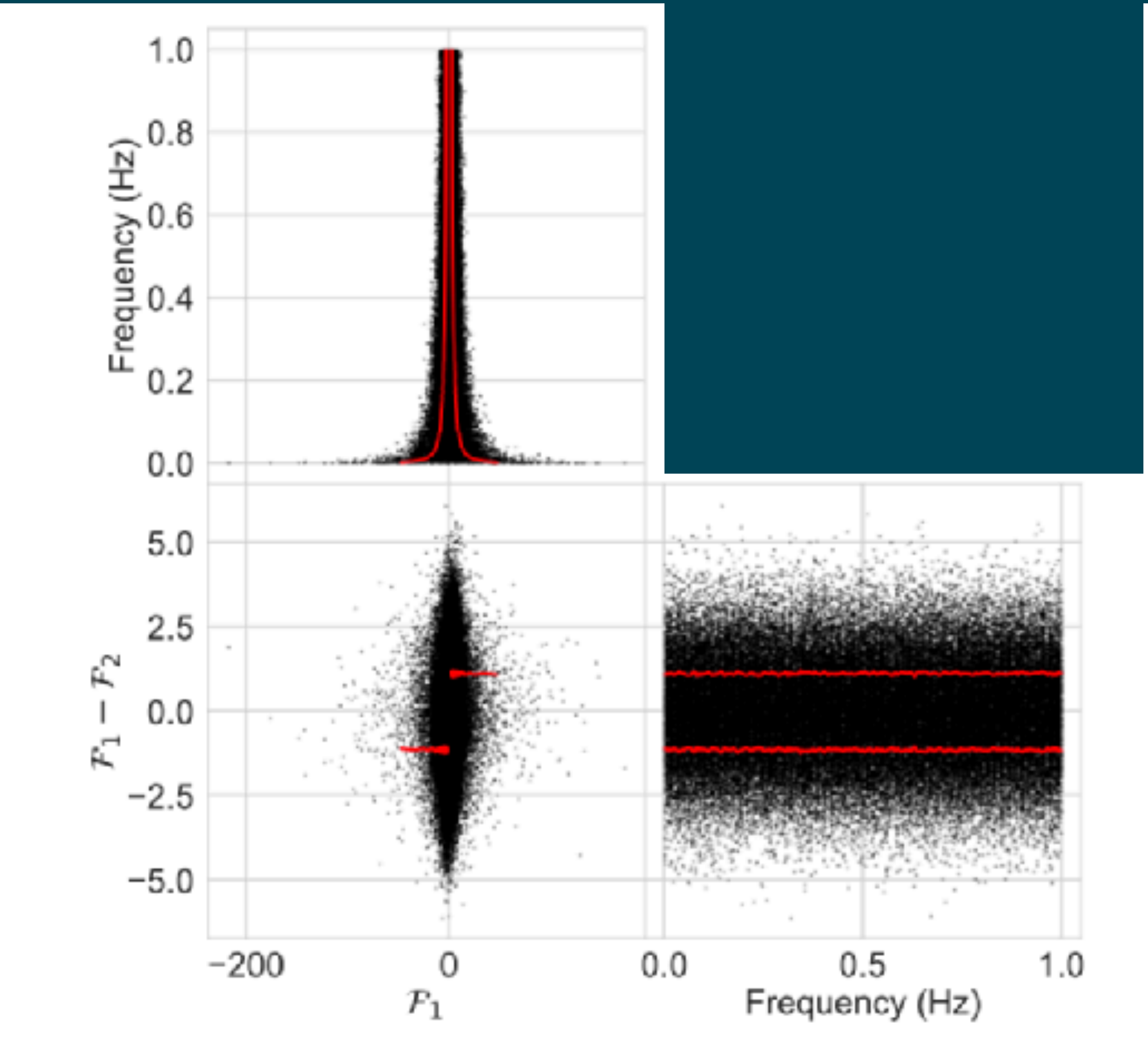
white noise, dead time



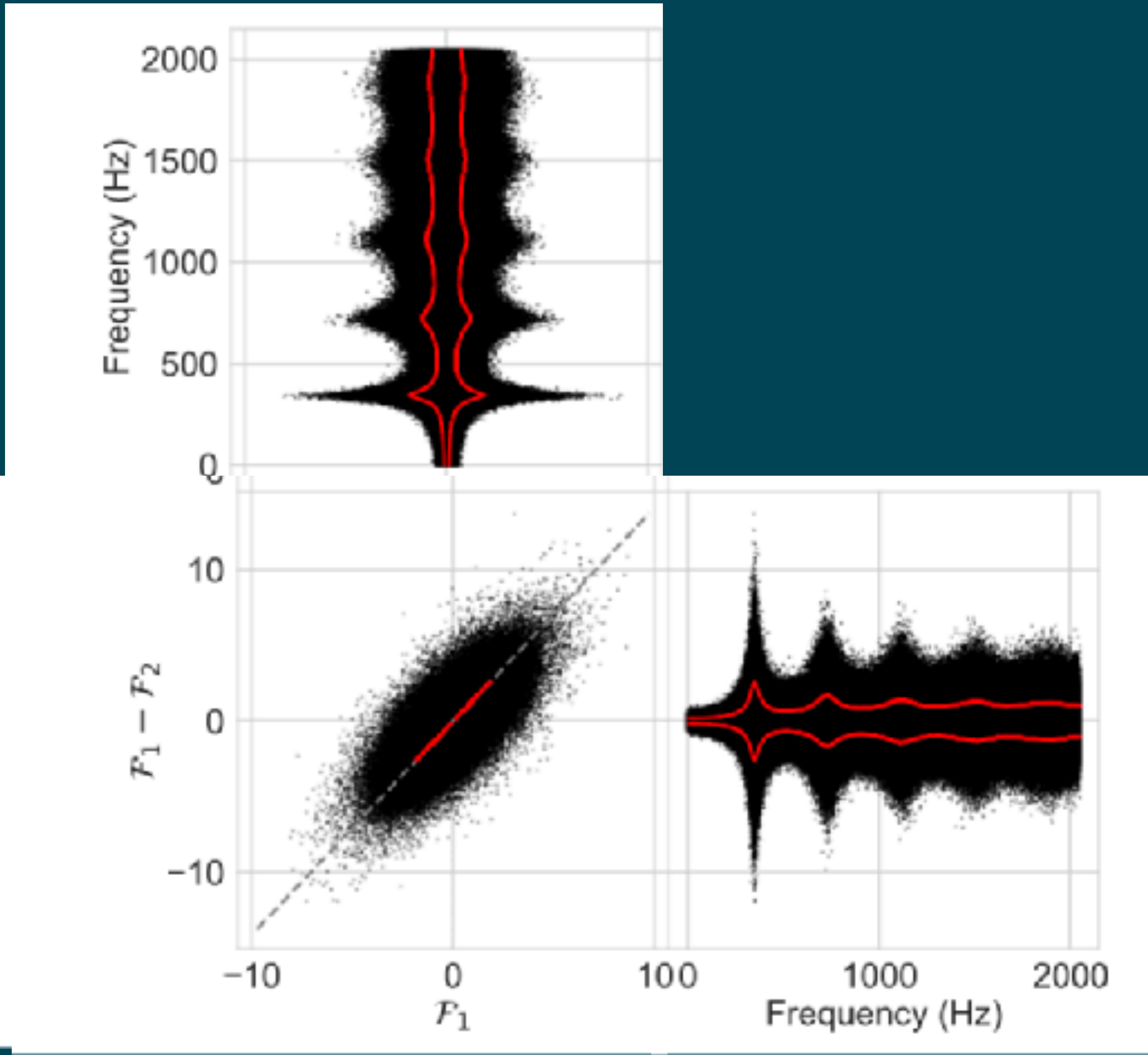
Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

red noise, no dead time



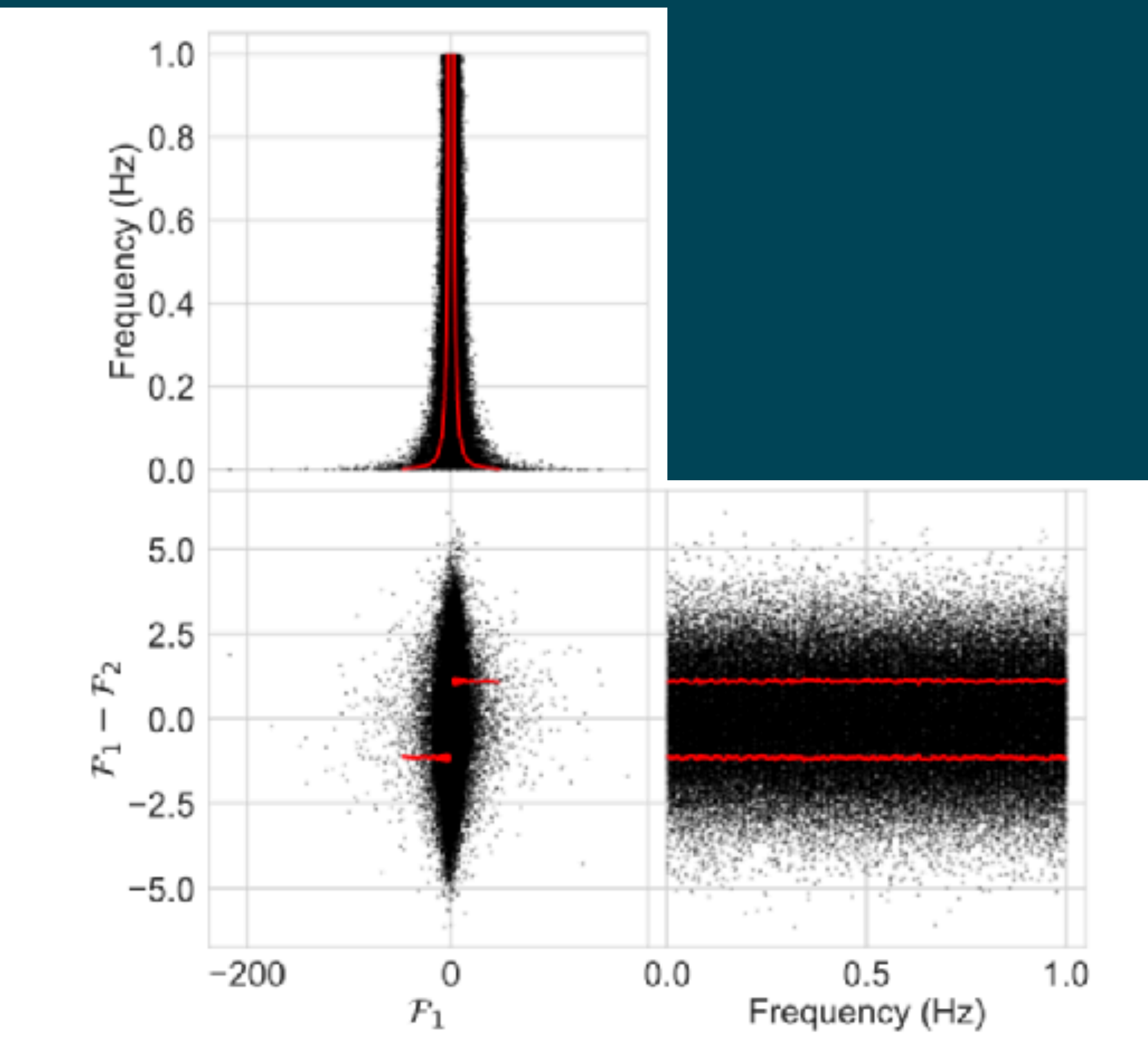
white noise, dead time



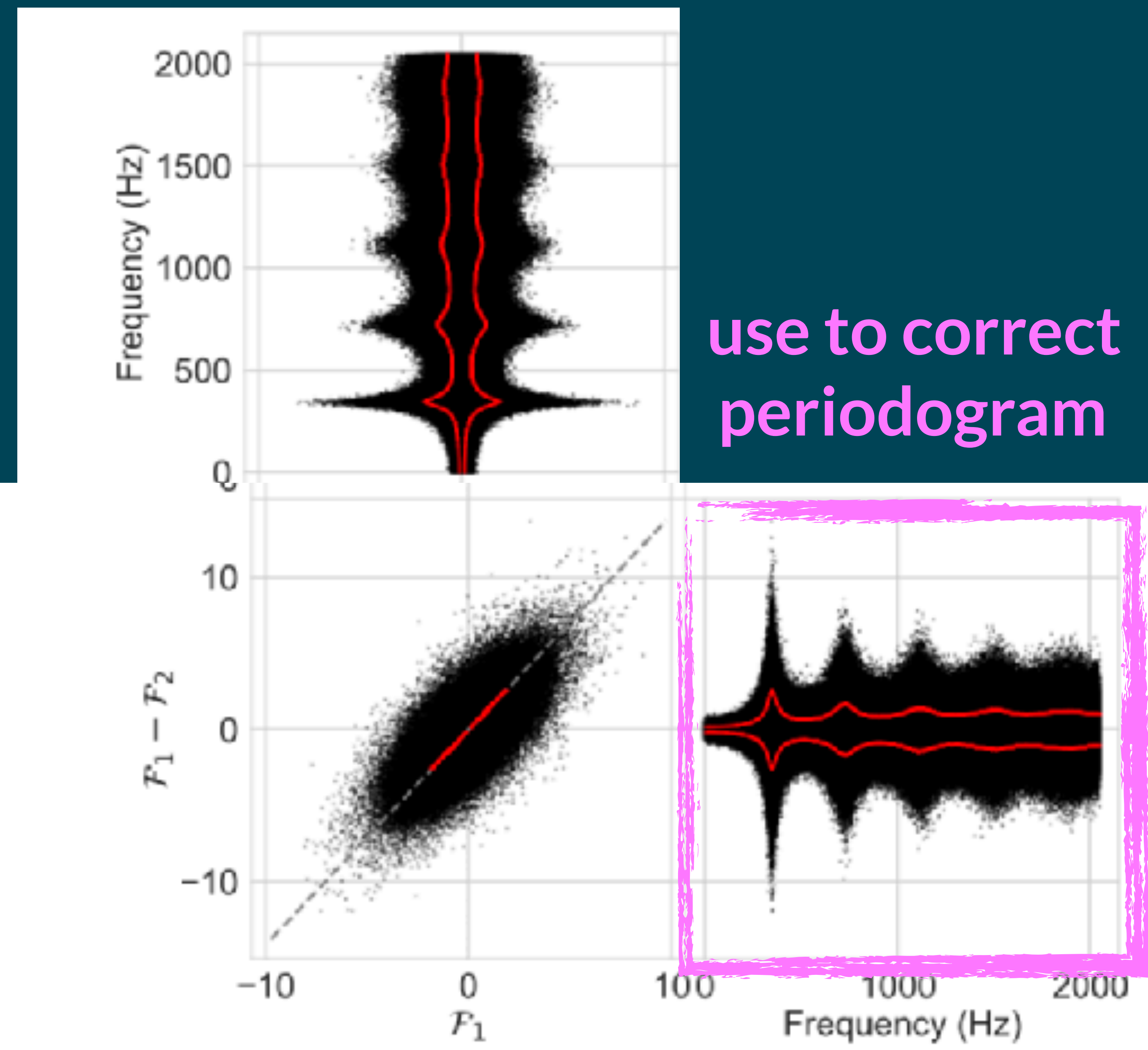
Bachetti & Huppenkothen (2018)

Fourier **A**mplitude **D**ifference Correction

red noise, no dead time

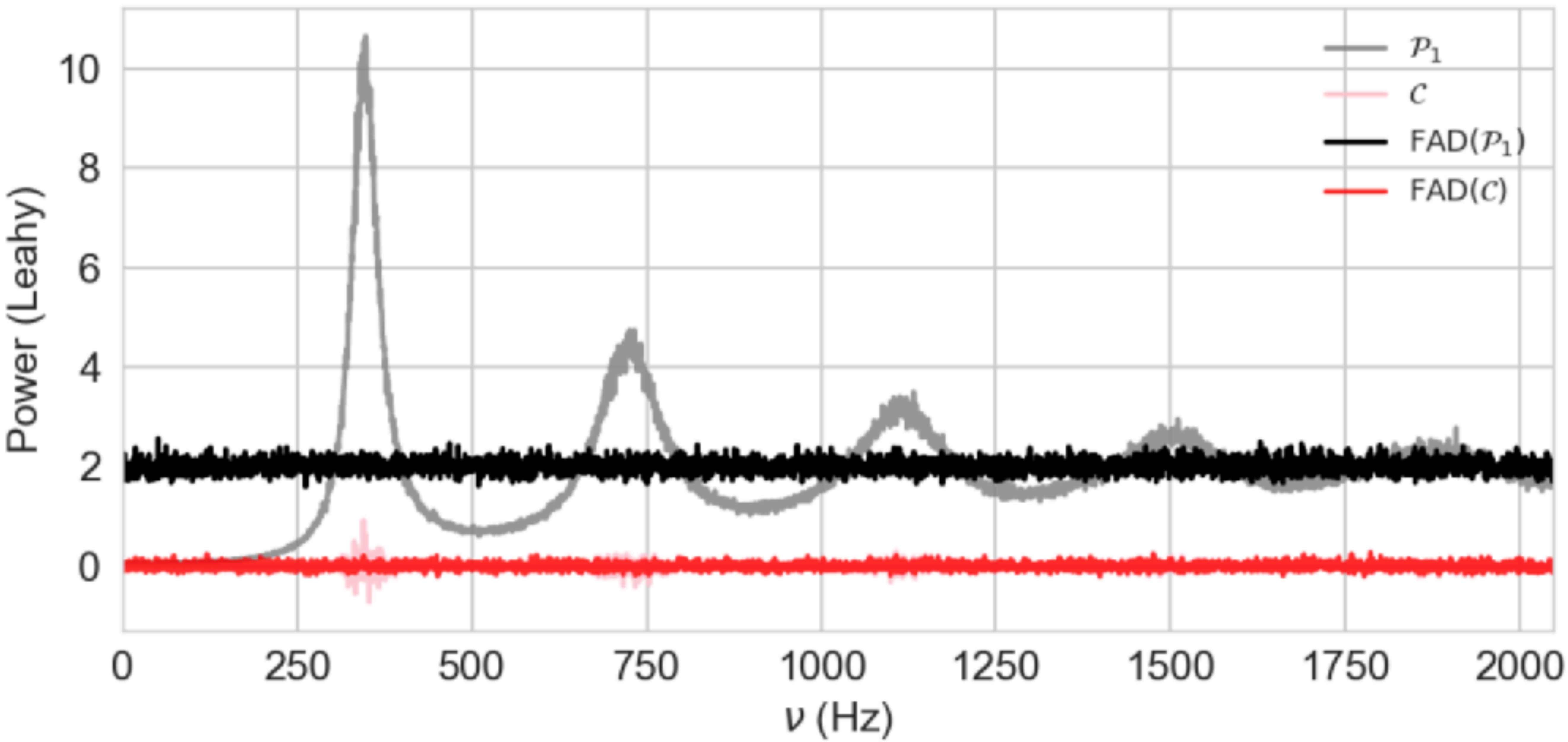


white noise, dead time



Bachetti & Huppenkothen (2018)

Fourier Amplitude Difference Correction

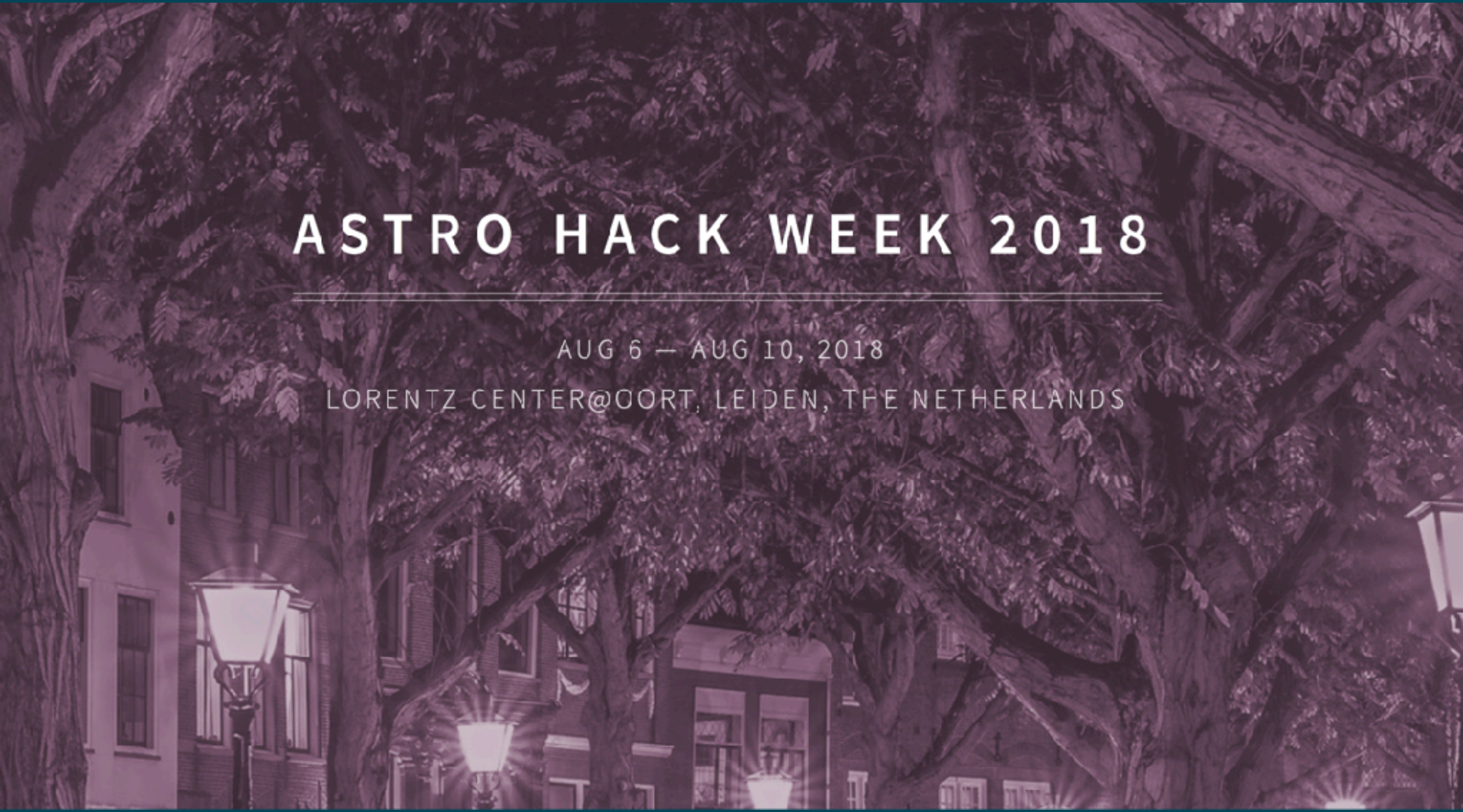


Bachetti & Huppenkothen (2018)

Caveat: this **overestimates** the rms amplitude when both flux and rms are very large

Conclusions

- statistics with Fourier spectra is fun!
- use the cospectrum to do timing of bright sources in the presence of dead time when more than one detector is available (Bachetti+, 2015)
- the (averaged) cospectrum requires different statistical distributions for significance testing (Huppenkothen+Bachetti, arXiv:1709.09666)
- there is currently no closed-form solution for red noise cospectra (future work)
- but red noise periodograms can be corrected using the FAD technique (Bachetti+Huppenkothen, arXiv:1709.09700)



www.astrohackweek.org

Google Summer of Code

Python API for XSPEC

Develop a modular python API to use XSPEC (a popular X-ray astronomy tool) in python workflows

dhuppenkothen
abigailStev

GSOC

timelab

Optimize Stingray for Large Datasets

Optimize tools in the Stingray library for use on large datasets from new X-ray space missions

matteobachetti
pbalm abigailStev

GSOC

timelab

Phase-resolved oscillations

Implement method to calculate the phase of oscillatory phenomena with non-constant frequency, and calculate phase-resolved spectra

abigailStev
matteobachetti

GSOC

timelab

http://openastronomy.org/gsoc/gsoc2018/