Time Series Analysis
In the Dynamic Universe

Jeff Scargle
jeffscargle@gmail.com

PHYSTAT-Gamma 2022 September 29, 2022

Time Series Analysis

- TSA in a Zwicky Morphological Box
- Elementary Operations
- Fourier Transform of Unevenly Sampled Data
- Time-Domain Segmentation: Bayesian Blocks
- Time-Domain Models

**** The line is similar to a length of time,***
and as the points are the beginning and end of the line,
so the instants are the endpoints of any given extension of time.

Leonardo da Vinci, Codex Arundel, folio 190v., c. 1500

Definition: A TIME SERIES IS …

- A set of data elements, each of which conveys some information about the underlying signal, or function of time

- Special case: *Sequential Data*, when the elements can be ordered in time

MORPHOLOGICAL BOX (after F. ZWICKY)

	A	**B**	C
V			
W			
X			
Y			
Z			

Local

Time Domain

Regression

Segmentation

Global

Frequency Domain

Scale Domain

Power Spectrum

Scalegram

Local ⊗ Global

Time-Frequency Time-Scale Distributions

Wavelet Scalogram, Spectrogram, Synchro-squeeze

(GLOBAL + LOCAL)

TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

DATA / RESULT	TIME SERIES	AMPLITUDES	SPECTRUM
TIME DOMAIN REPRESENTATION			
TIME-FREQUENCY / TIME-SCALE DISTRIBUTION			
FREQUENCY /SCALE DOMAIN REPRESENTATION			
MODEL			
STATISTIC			

TIME SERIES MORPHOLOGICAL BOX (circa 1968)

DATA / RESULT	TIME SERIES	AMPLITUDES	SPECTRUM
TIME DOMAIN REPRESENTATION			
TIME-FREQUENCY / TIME-SCALE DISTRIBUTION			
FREQUENCY /SCALE DOMAIN REPRESENTATION	POWER SPECTRUM		
MODEL			
STATISTIC			

TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

DATA / RESULT	TIME SERIES	AMPLITUDES	SPECTRUM
TIME DOMAIN REPRESENTATION	REGRESSION (parametric/non) SEGEMENTATION SMOOTHING DENOISING LOCAL MAX or MIN WATERSHEDS SIGNAL RECOVERY		RECONSTRCTION AUTO-CORRELATION
TIME-FREQUENCY / TIME-SCALE DISTRIBUTION	SPECTROGRAM WIGNER-VILLE SCALOGRAM WAVELETS CHIRPLETS OPTIMAL BASIS PURSUIT		
FREQUENCY /SCALE DOMAIN REPRESENTATION	POWER/PHASE SPECTRUM SCALEGRAM Z-TRANSFORM (MULTI)-TAPERED SPECTRA WELCH SPECTRA CEPSTRUM		SMOOTHING AVERAGING FILTERING
MODEL	AUTOREGRESSIVE MOVING AVERAGE ORNSTEIN UHLENBECK MARKOV DAMNED RANDOM WALK	HISTOGRAM BB HISTOGRAM STANDARD FORMULA	PERIODICITY QUASI PERIODICITY $f^{-\alpha}$ NOISE BROKEN POWER LAW CHIRP
STATISTIC	MOMENTS (MEAN, VAR, KURTOSIS) NUMBER OF PEAKS / VALLEYS "RMS-FLUX" RELATION BOOTSTRAP	QUIET BACKGROUND DYNAMIC RANGE BOOTSTRAP	POWER LAW INDEX BREAK FREQUENCY NOISE FLOOR

TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

DATA RESULT	TIME SERIES	AMPLITUDES	SPECTRUM
TIME DOMAIN REPRESENTATION	REGRESSION (parametric/non) SEGEMENTATION SMOOTHING DENOISING LOCAL MAX or MIN WATERSHEDS SIGNAL RECOVERY		RECONSTRCTION AUTO-CORRELATION
TIME-FREQUENCY / TIME-SCALE DISTRIBUTION	SPECTROGRAM WIGNER-VILLE SCALOGRAM WAVELETS CHIRPLETS OPTIMAL BASIS PURSUIT		
FREQUENCY /SCALE DOMAIN REPRESENTATION	POWER/PHASE SPECTRUM SCALEGRAM Z-TRANSFORM (MULTI)-TAPERED SPECTRA WELCH SPECTRA CEPSTRUM		SMOOTHING AVERAGING FILTERING
MODEL	AUTOREGRESSIVE MOVING AVERAGE ORNSTEIN UHLENBECK MARKOV DAMNED RANDOM WALK	HISTOGRAM BB HISTOGRAM STANDARD FORMULA	PERIODICITY QUASI PERIODICITY $f^{-\alpha}$ NOISE BROKEN POWER LAW CHIRP
STATISTIC	MOMENTS (MEAN, VAR, KURTOSIS) NUMBER OF PEAKS / VALLEYS "RMS-FLUX" RELATION	QUIET BACKGROUND DYNAMIC RANGE	POWER LAW INDEX BREAK FREQUENCY NOISE FLOOR

TIME-RESOLVED VERSION OF ANY STATISTIC!

TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

INPUT DOMAIN / OUTPUT DOMAIN	TIME	AMPLITUDES	FREQUENCY
	REGRESSION		RECONSTRUCTION
TIME-FREQUENCY	TIME-FREQ DIST SPECTROGRAM SYNCHROSQUEEZE SCALOGRAM WAVELETS CHIRPLETS OPTIMAL BASES	?	TIME-RESOLVED VERSION OF ANY STATISTIC
FREQUENCY	POWER SPECTRUM PHASE SPECTRUM SCALEGRAM PERIOD FOLDING Z-TRANSFORM (MULTI)-TAPERED SPECTRA WELCH SPECTRA CEPSTRUM	HISTOGRAM BB HISTOGRAM	SMOOTHING AVERAGING (E.G. WELCH)

$\delta(t-t_n)$	$x(t_n)$	$x_n p_t(t-t_n)$	$p_x(x-x_n)$	$\delta(t-t_n)$ $p_x(x-x_n)$ $p_t(x-x_n)$	N_n	ACGT
time-tag	point	distributed in time	distributed in amplitude	distributed In time and amp	counts in bins	category

BATSE trigger 00551

N = 2894 photons

1
2
3
458
740
1084
1628
2050
2446
2770
3058
3388
4078
4284

x_n

t_n

TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

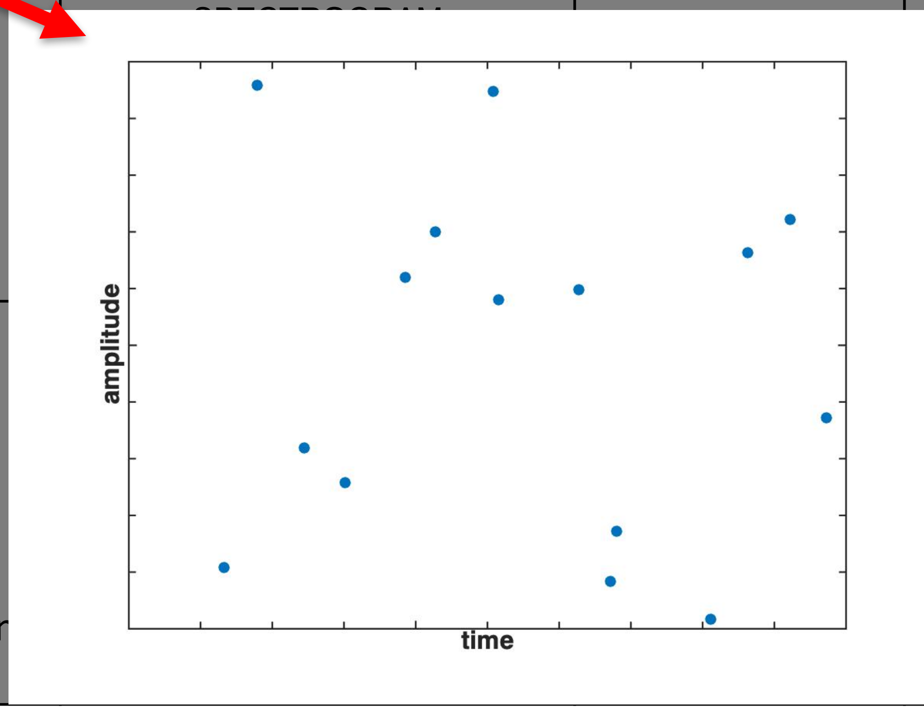
INPUT DOMAIN / OUTPUT DOMAIN	TIME	AMPLITUDES	FREQUENCY
	REGRESSION		RECONSTRUCTION
TIME-FREQUENCY	TIME-FREQ DIST SPECTROGRAM		TIME-RESOLVED VERSION OF ANY STATISTIC
FREQUENCY			SMOOTHING AVERAGING (E.G. WELCH)

$\delta(t-t_n)$ $x(t_n)$ $x_n p_t(t-t_n)$ $p_x(x-x_n)\delta(t-t_n)$ $p_x(x-x_n)$ $p_t(x-x_n)$ N_n ACGT

time-tag | point | distributed in time | distributed in amplitude | distributed In time and amp | counts in bins | category

x_n

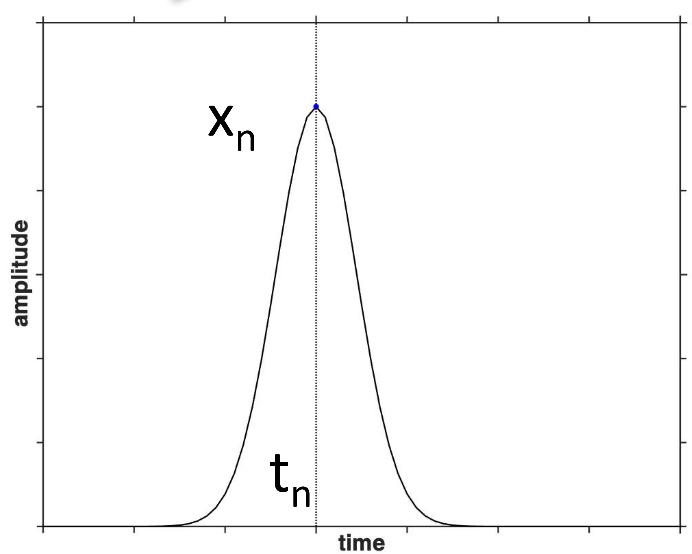
t_n

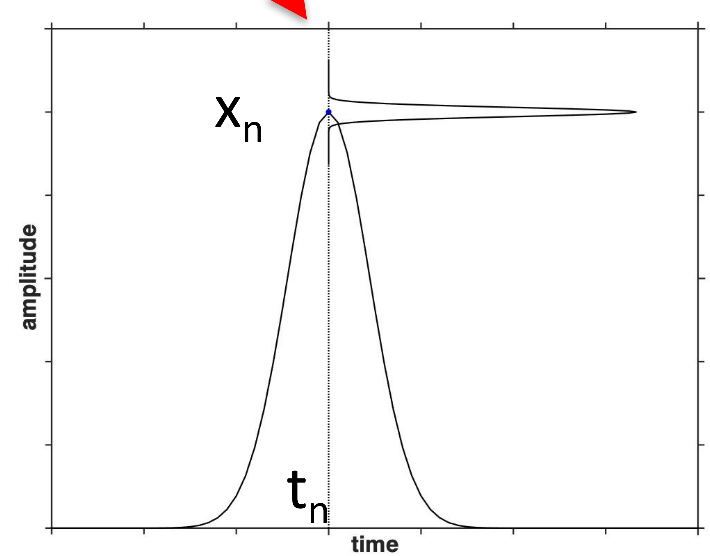


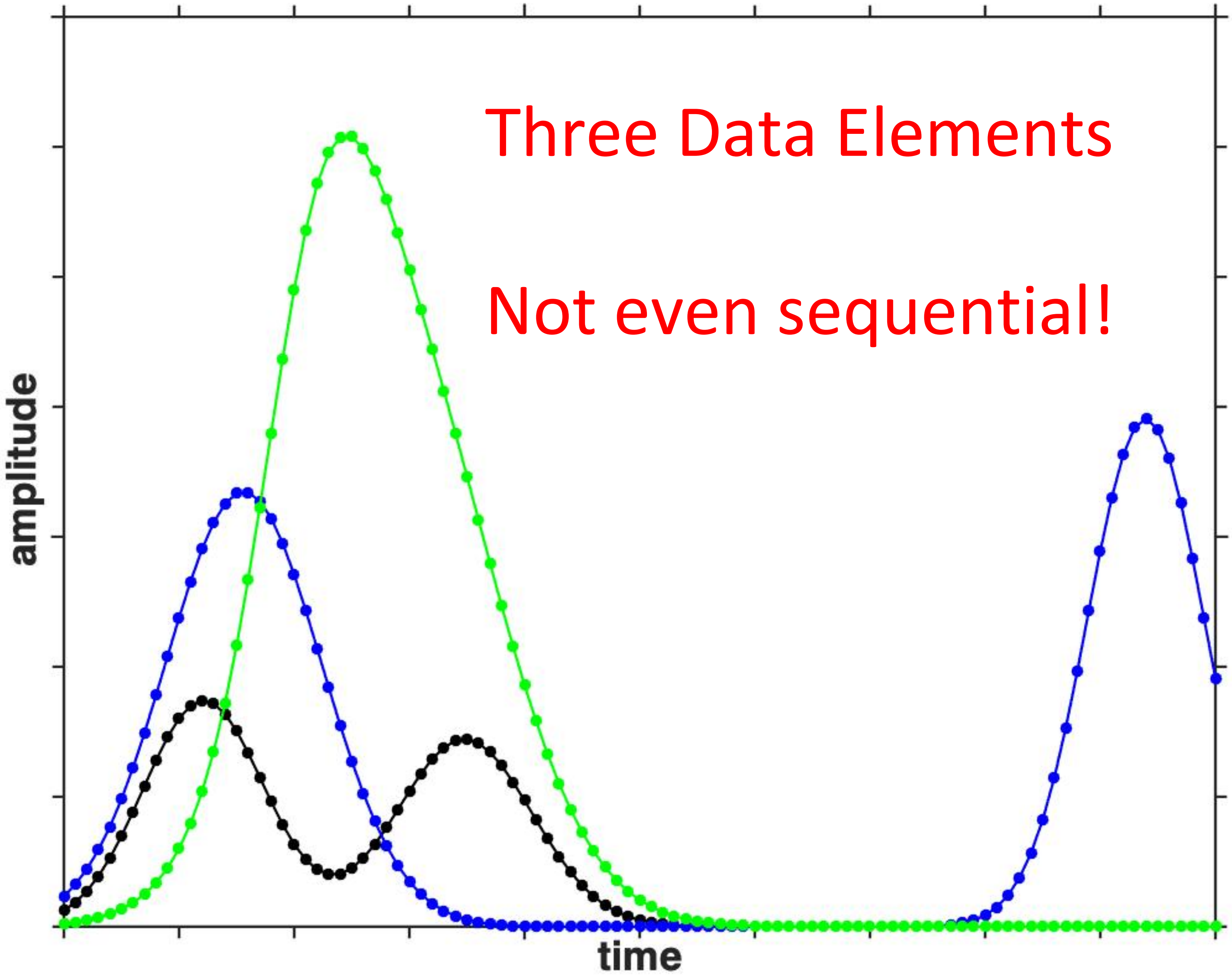
TIME SERIES MORPHOLOGICAL BOX (after F. ZWICKY)

INPUT DOMAIN / OUTPUT DOMAIN	TIME	AMPLITUDES	FREQUENCY
	REGRESSION		RECONSTRUCTION
TIME-FREQUENCY	TIME-FREQ DIST SPECTROGRAM SYNCHROSQUEEZE SCALOGRAM	?	TIME-RESOLVED VERSION OF ANY STATISTIC

$\delta(t - t_n)$	$x(t_n)$	$x_n p_t(t - t_n)$	$p_x(x-x_n)\delta(t - t_n)$	$p_x(x-x_n)\, p_t(x-x_n)$	N_n	ACGT
time-tag	point	distributed in time	distributed in amplitude	distributed In time and amp	counts in bins	category






Three Data Elements
Not even sequential!
amplitude
time

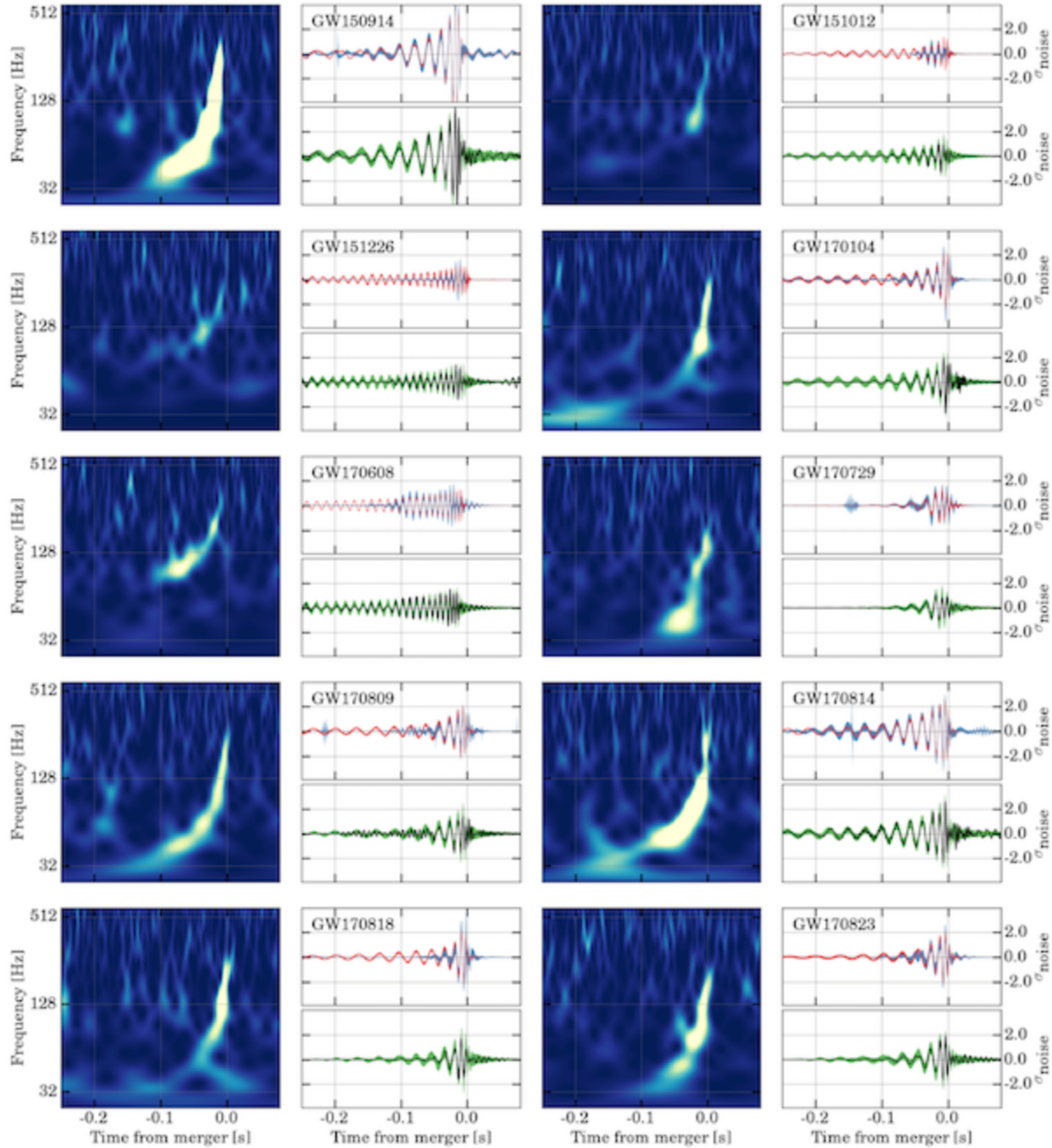


Image adopted from arXiv:1811.12907 by the LIGO Scientific Collaboration and the Virgo Collaboration.

Solar Cycle Variability and Surface Differential Rotation from Ca II K-Line Time Series Data

Jeffrey Scargle, Stephen Keil, Pete Worden

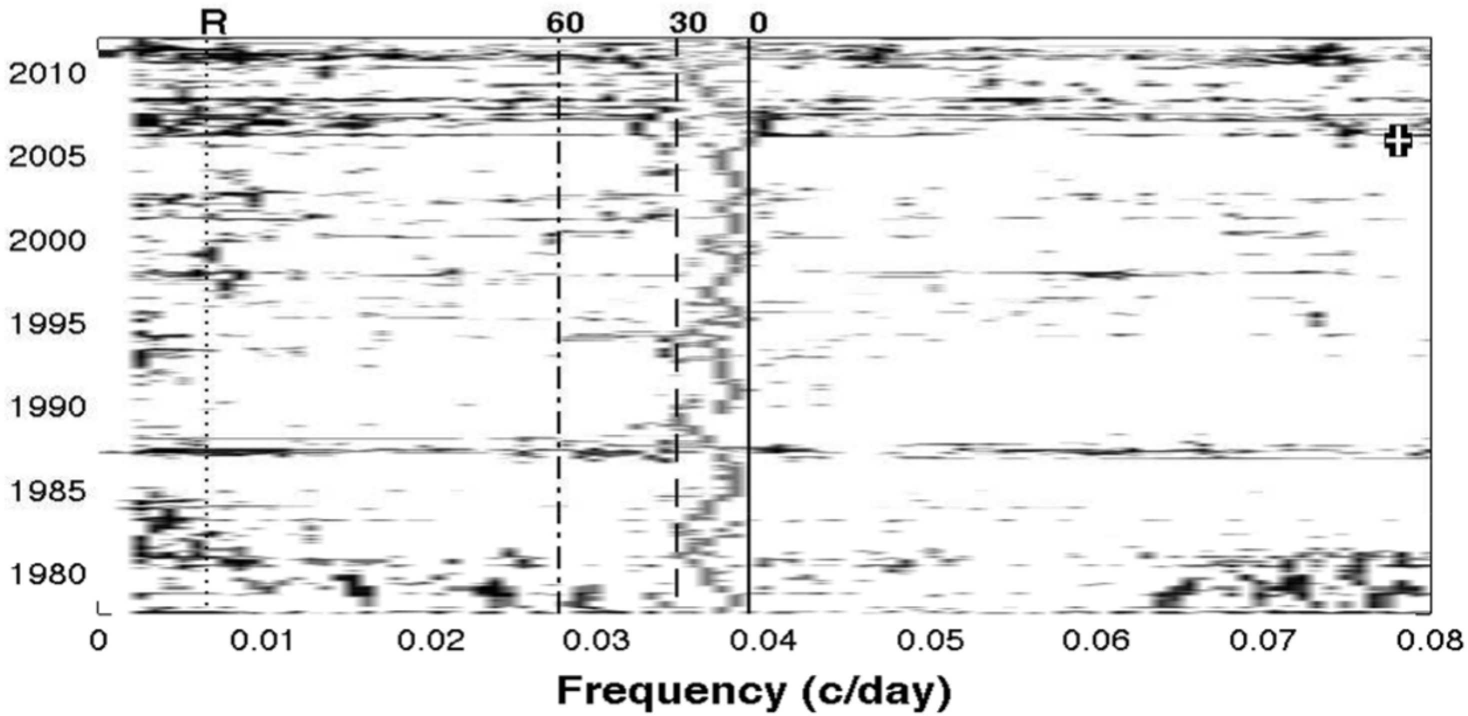
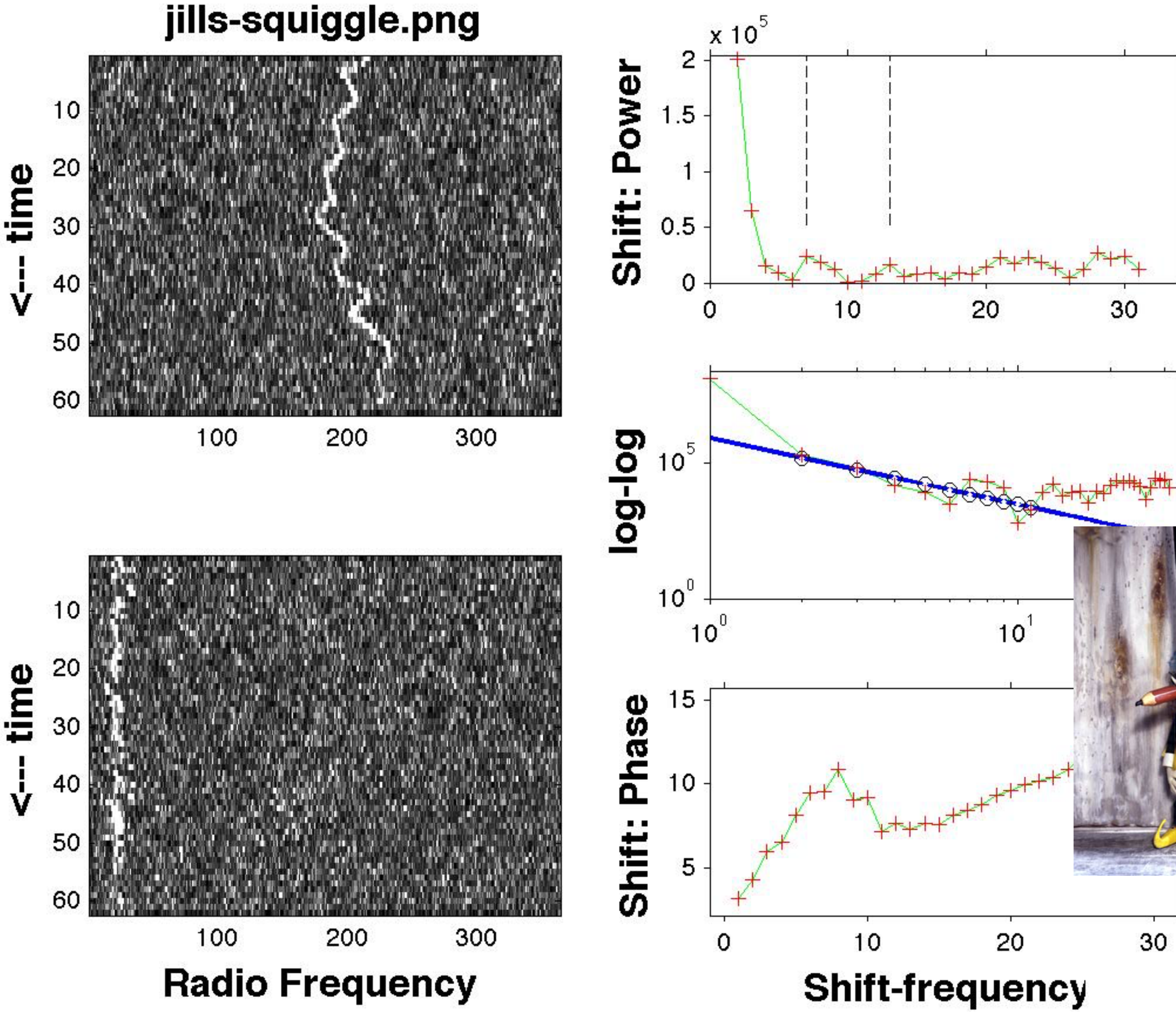
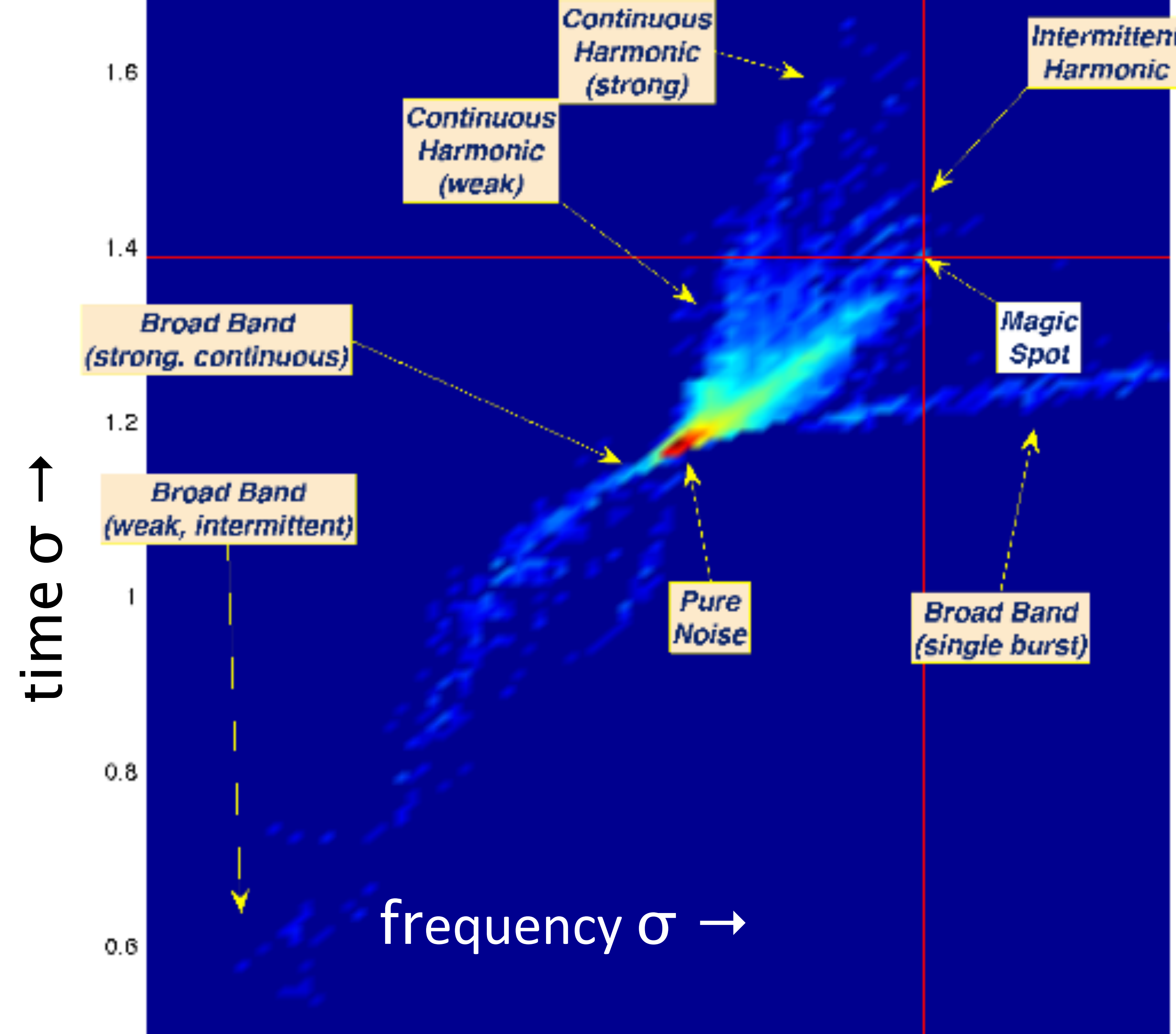
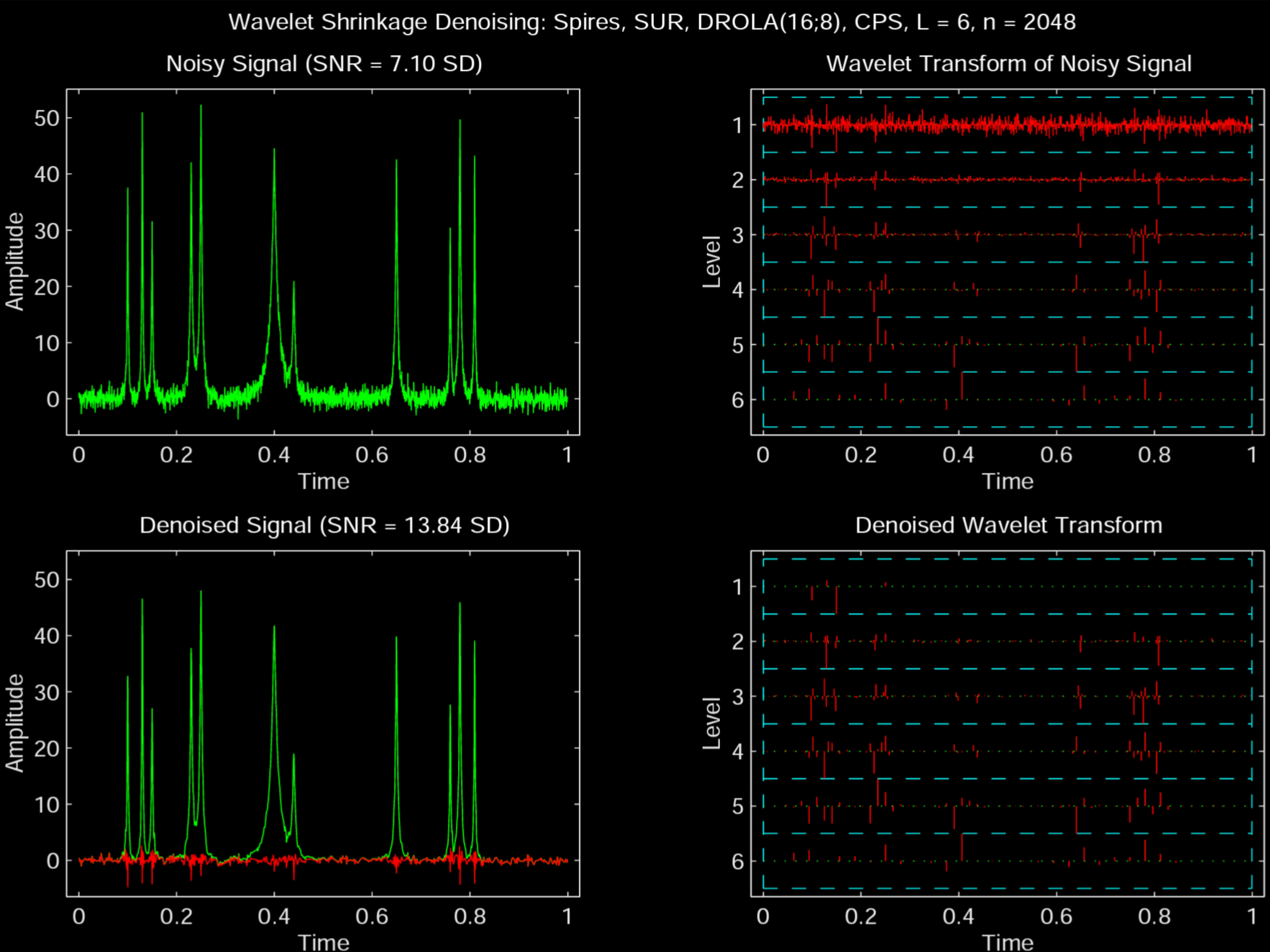


Fig. 10.— Time-Frequency distribution of **EMDX** residuals. Bottom, renormalized to make constant the power between the frequencies corresponding to 0 and 30 degrees, bringing out behavior during solar minimum that is otherwise lost. Vertical lines denote latitudes 0, 30, and 60 deg based on Eq. (1). The dotted line at .0082 c/d roughly corresponds to quasi-periodicities discussed in §4. Power below .00176 c/day is divided by 10 to improve the display.


jills-squiggle.png
<--- time
Radio Frequency
Shift: Power
x 10^5
log-log
Shift: Phase
Shift-frequency


Continuous Harmonic (strong)
Intermittent Harmonic
Continuous Harmonic (weak)
Broad Band (strong. continuous)
Magic Spot
Broad Band (weak, intermittent)
Pure Noise
Broad Band (single burst)
1.6
1.4
1.2
1
0.8
0.6
time σ →
frequency σ →


Wavelet Shrinkage Denoising: Spires, SUR, DROLA(16;8), CPS, L = 6, n = 2048
Noisy Signal (SNR = 7.10 SD)
Amplitude
Time
Wavelet Transform of Noisy Signal
Level
Time
Denoised Signal (SNR = 13.84 SD)
Amplitude
Time
Denoised Wavelet Transform
Level
Time

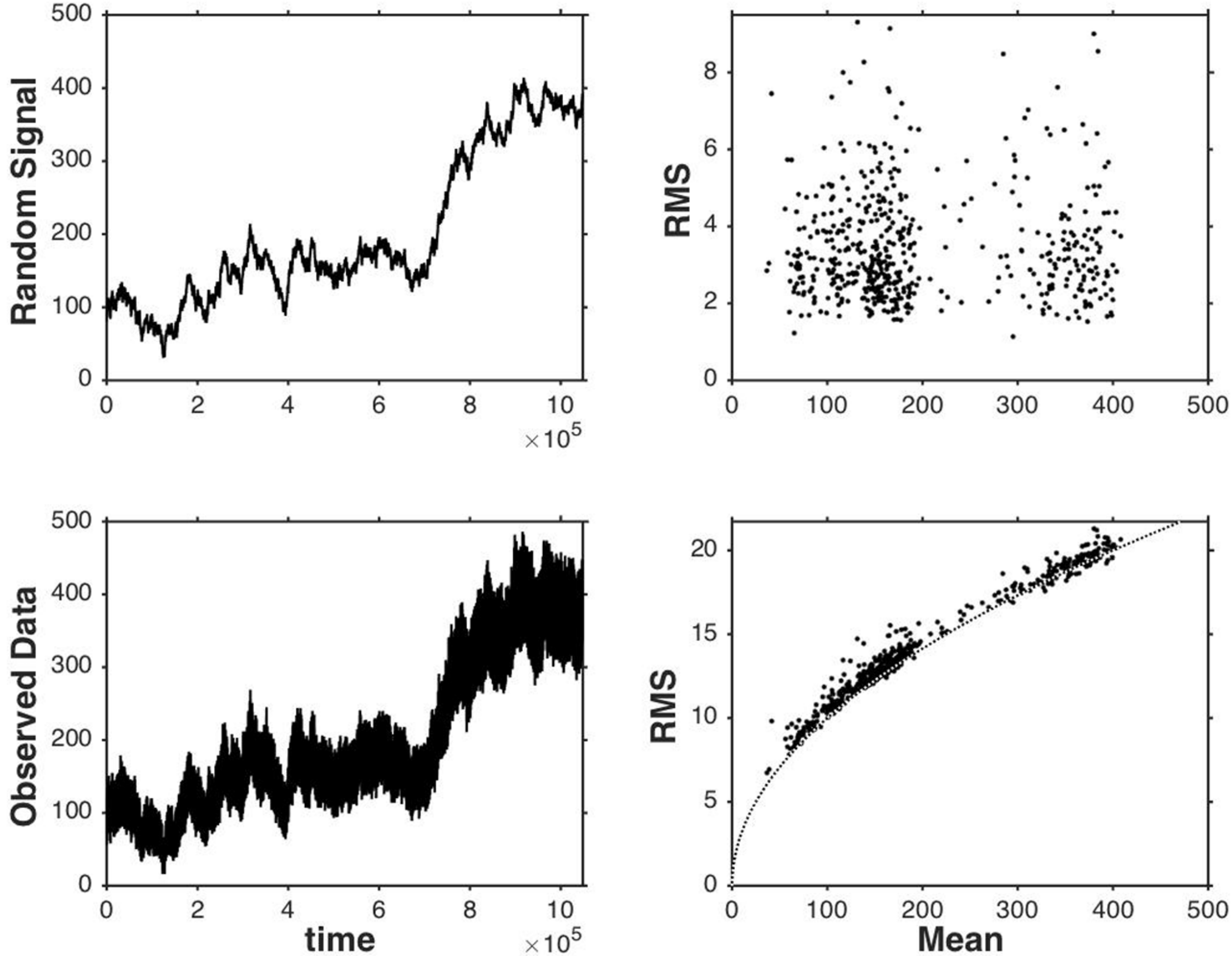


Fig. 2.— A simulated random walk time series. Upper-Left: cumulative sum of slightly biased normally distributed variables, renormalized. Lower-Left: The same data processed by applying the Poisson operator. Right Panels: scatter plots of root-mean-square flux vs. mean flux, derived from the corresponding data on the left-hand side. The solid line in the bottom-right panel is the theoretical square-root relation for the Poisson distribution. This figure demonstrates the effects of photon counting fluctuations, but for the most part these are accounted for in the data processing.

Fourier Transform of Unevenly Spaced Time Series

weights (extensive)

(intensive) amplitudes

time tags

The Measurable Quantities of Physics
Tolman, 1917, Phys. Rev. 9, 237–253

frequencies
(arbitrary)

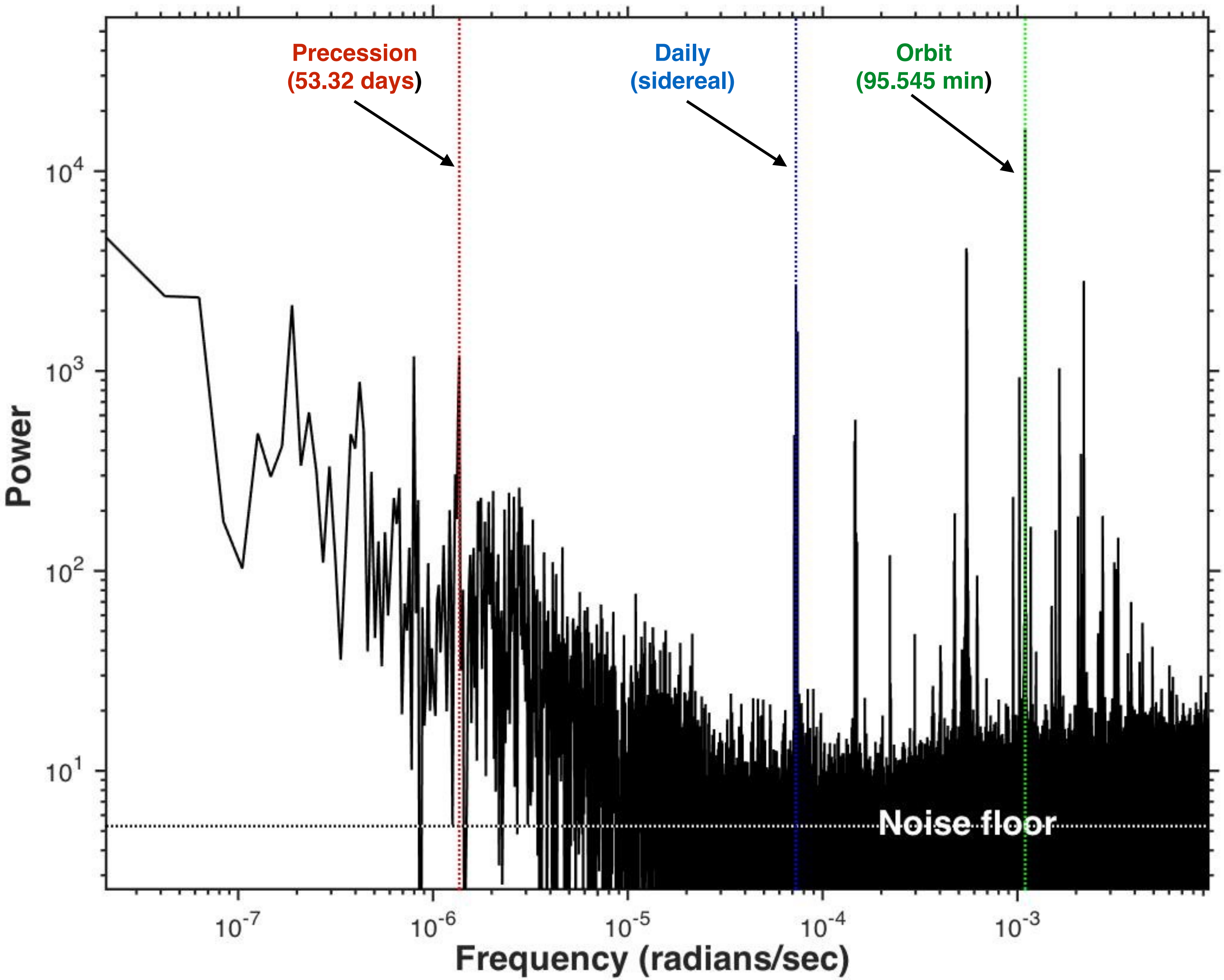
$$FT_X(\omega) = F_0 \left\{ \frac{\sum_n w_n x_n \cos \omega(\, t_n - \tau(\omega)\,)}{[\, \sum_m w_m \cos^2 \omega(\, t_m - \tau(\omega)\,)\,]^{1/2}} + i\, \frac{\sum_n w_n x_n \sin \omega(\, t_n - \tau(\omega)\,)}{[\, \sum_m w_m \sin^2 \omega(\, t_m - \tau(\omega)\,)]^{1/2}} \right\}$$

where

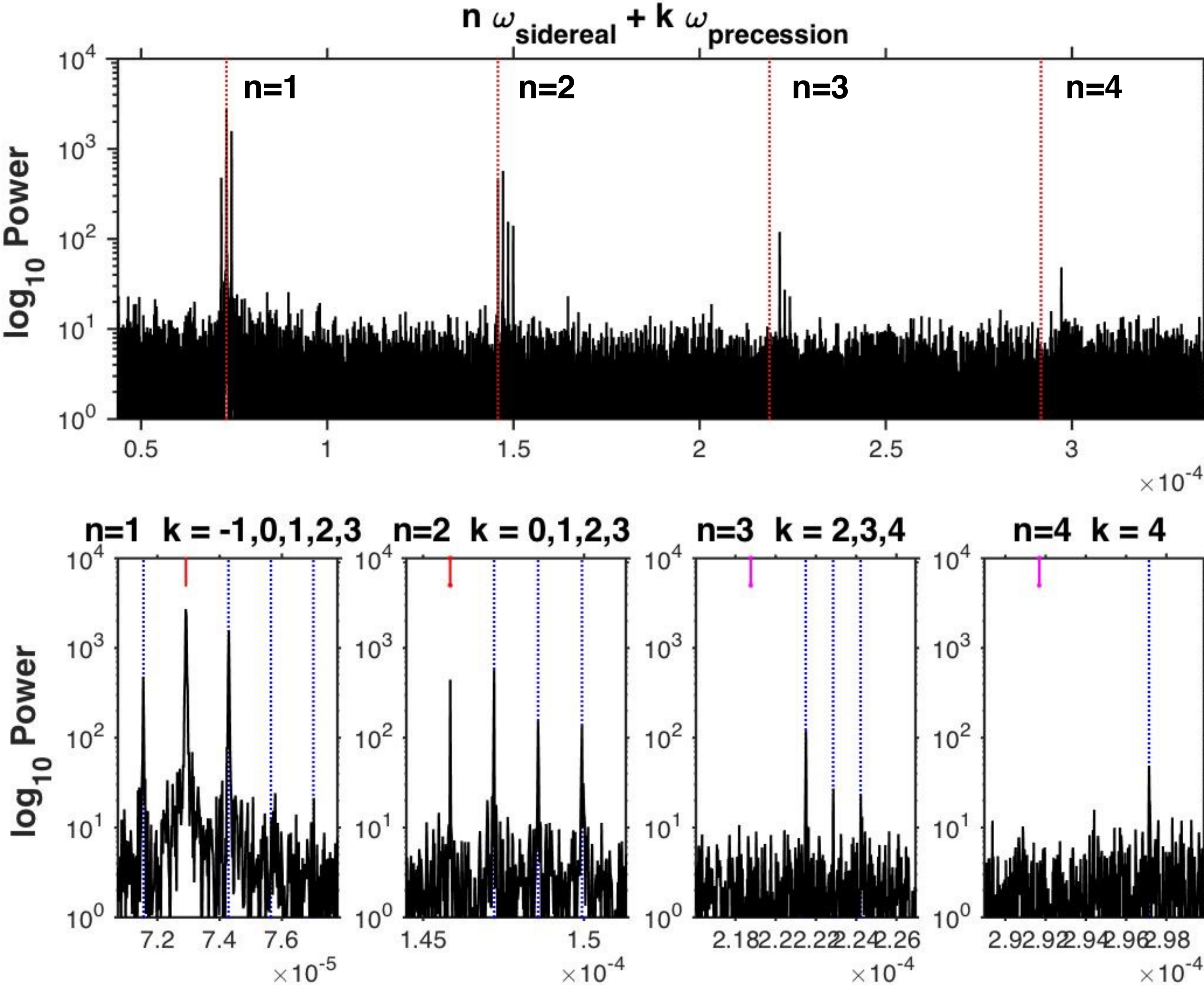
$$\tau(\omega) \;=\; \frac{1}{2\omega} \arctan[\, (\sum w_n \sin 2\omega t_n) \,/\, (\sum w_m \cos 2\omega t_m)\,]$$

$$\boxed{\text{Treat event (photon) data as } \boldsymbol{\delta}(t - t_n) \quad \ldots \text{ that is, } x_n = 1}$$

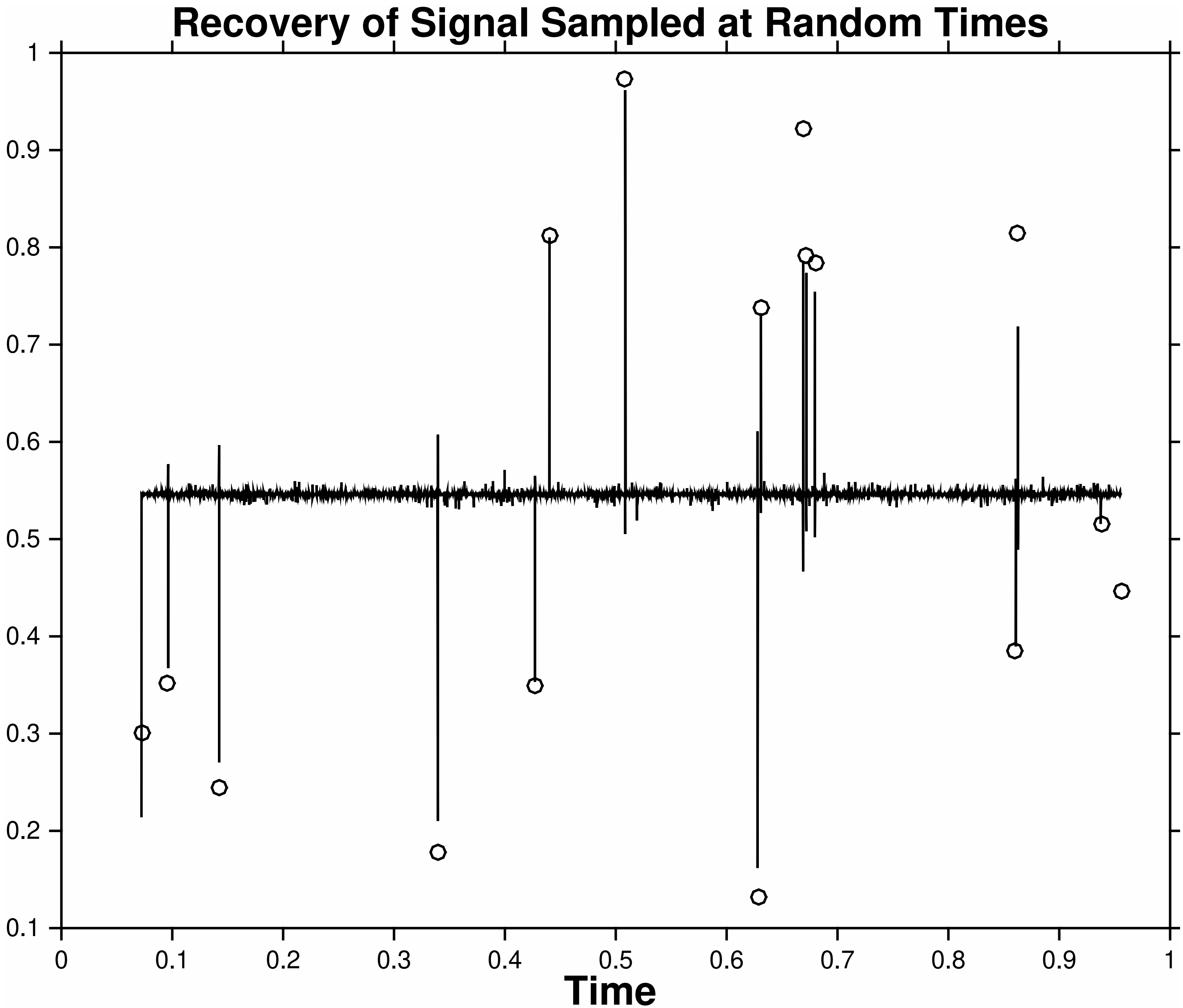
Studies in astronomical time series analysis. III.
Fourier Transforms, Autocorrelation Functions, and Cross-Correlation Functions of Unevenly Spaced Data.
JS, ApJ 343, 874-887



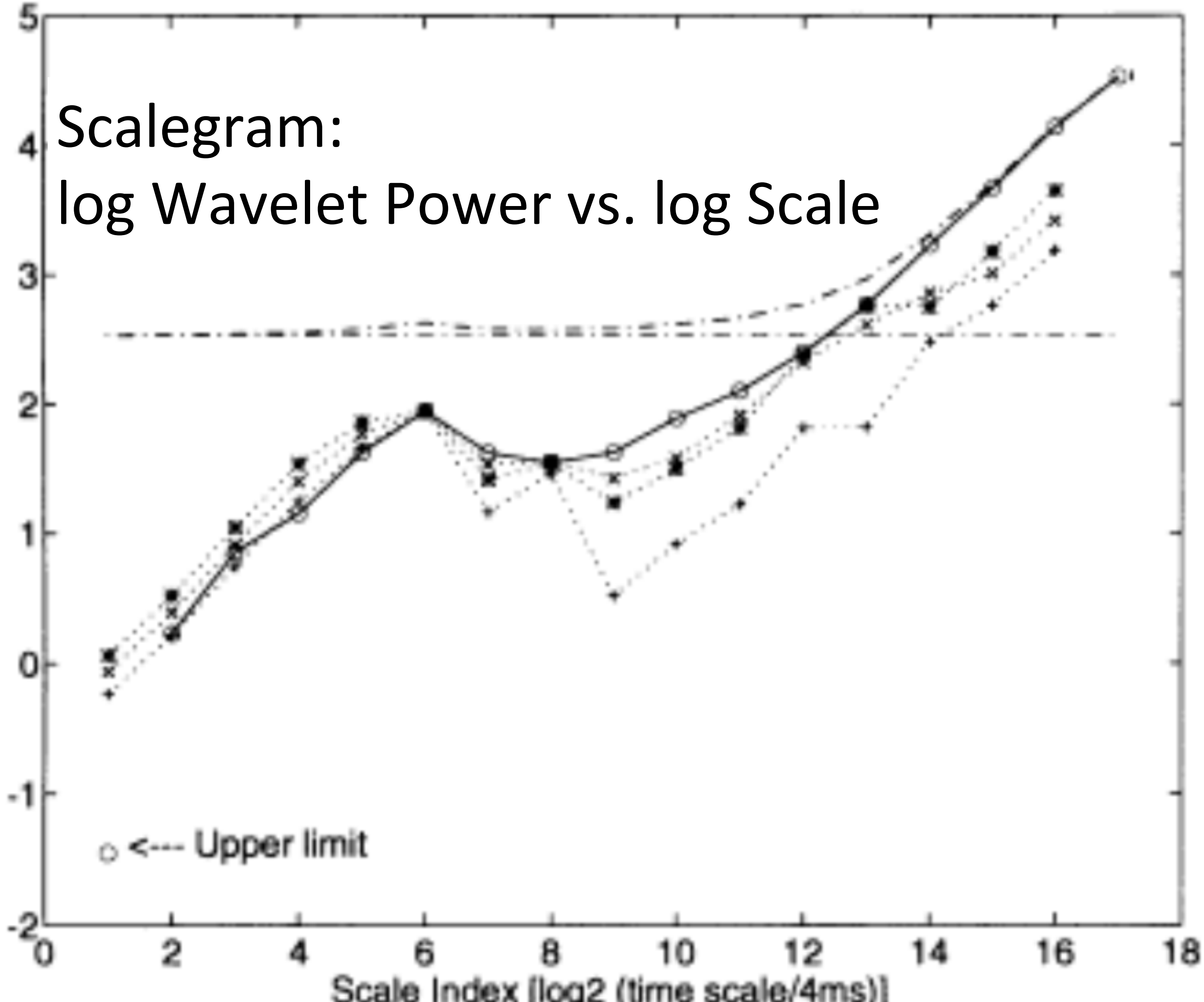
Precession
(53.32 days)
Daily
(sidereal)
Orbit
(95.545 min)
Noise floor
Power
Frequency (radians/sec)



n ω sidereal + k ω precession
n=1
n=2
n=3
n=4
log10 Power
×10-4
n=1 k = -1,0,1,2,3
n=2 k = 0,1,2,3
n=3 k = 2,3,4
n=4 k = 4
×10-5


Recovery of Signal Sampled at Random Times
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time

Scalegram:
log Wavelet Power vs. log Scale



The quasi-periodic oscillations and very low frequency noise
of Scorpius X-1 as transient chaos - A dripping handrail?
JS, Steiman-Cameron, Young, Donoho, Crutchfield, Imamura, ApJ, 411, L91

Moral Imperatives of Data Analysis

("Information Husbandry?")

The Truth: Extract valid information from the data

The Whole Truth: ***avoid the sin of omission by not***

- discarding information (data smoothing, binning, rounding, cutting)
- hiding information (Fourier phase spectra, publication bias)
- resorting to easy, standard, greedy suboptimal methods (unprincipled tapers, multi-point correlation functions, the "data microscope")

Nothing But the Truth: ***avoid the sin of lying by***

- understanding and accounting for random and systematic errors, biases
- avoiding corruption of true information (data smoothing, binning, rounding)
- not introducing false information (interpolation, gap filling)
- cherry picking (trials factor, publication bias, "physical values")

Essential Features of Bayesian Blocks:

Best possible step-function fit to data
(Exact Global Optimum)

The truth, the whole truth, nothing but the truth.

No loss of information or resolution by pre-binning

Single parameter: prior on number of blocks
(penalty for model complexity)
Mediates the bias-variance trade-off

- *Data*

  Points ordered in time

- *Blocks (sets of consecutive data points)*

  Block Shape Model

  Block Fitness (objective function)

- *Segmented Model: partition into blocks*

  Total Fitness = Sum of Block Fitness

  Optimize Total Fitness over all 2^N possible partitions

  (Optimum number of blocks automatically determined!)

- ***One parameter,** a penalty constant,*

  derived from the prior on the number of blocks;

  calibrate to an acceptable false positive rate.

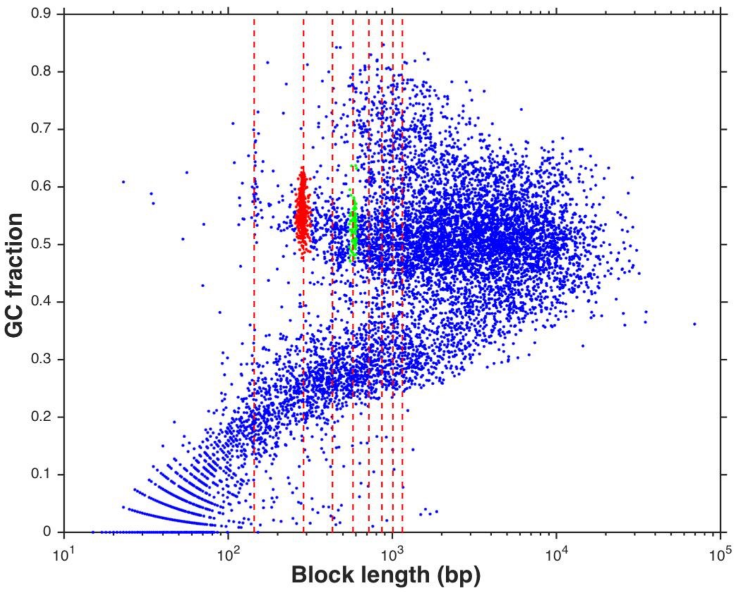
Bayesian Blocks can use ANY Data Mode

- point measurements: $X(t_n)$
- time-tagged events: t_n
- time-to-spill
- categorical
- data with gaps
- circular data
- uneven sampling
- exposure variation
- real-time or retrospective
- arbitrary mixtures of data modes
- multivariate data (with or without the constraining the change-points to be the same)
- higher dimensions
- auxiliary information

Bayesian Blocks can use ANY Data Mode

- *point measurements: $X(t_n)$*
- *time-tagged events: t_n*
- *time-to-spill*
- *categorical*
- *data with gaps*
- *circular data*
- *uneven sampling*
- *exposure variation*
- *real-time or retrospective*
- *arbitrary mixtures of data modes*
- *multivariate data (with or without the constraining the change-points to be the same)*
- *higher dimensions*
- *auxiliary information*

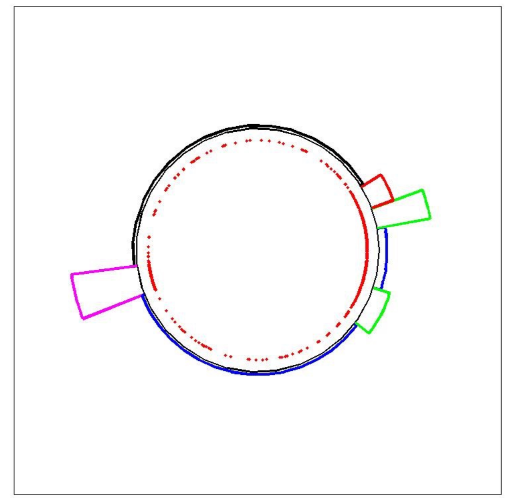
GC Islands in the Human Genome



*Bayesian Blocks can use **ANY** Data Mode*

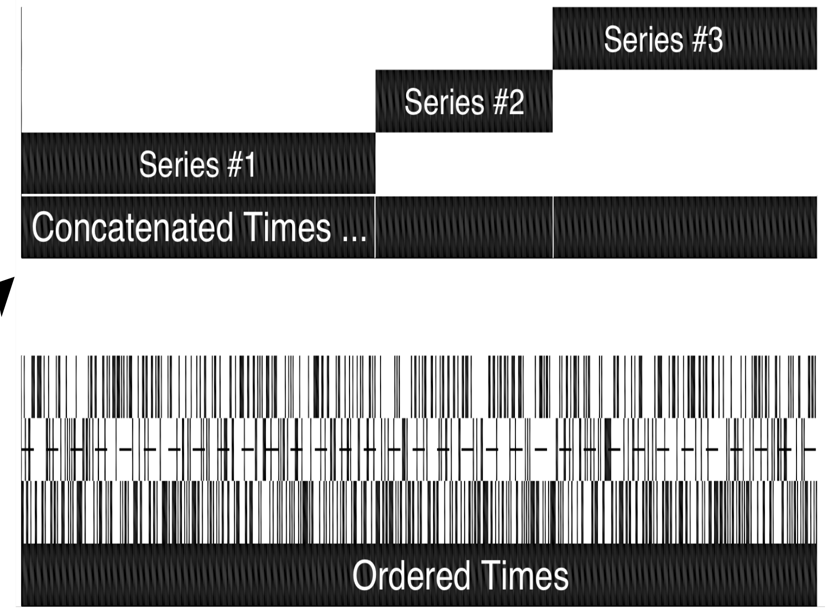
- *point measurements: $X(t_n)$*
- *time-tagged events: t_n*
- *time-to-spill*
- *categorical*
- *data with gaps*
- *circular data*
- *uneven sampling*
- *exposure variation*
- *real-time or retrospective*
- *arbitrary mixtures of data modes*
- *multivariate data (with or without the constraining the change-points to be the same)*
- *higher dimensions*
- *auxiliary information*

Optimal Partitioning
of
Data on the Circle



Bayesian Blocks can use ANY Data Mode

- point measurements: $X(t_n)$
- time-tagged events: t_n
- time-to-spill
- categorical
- data with gaps
- circular data
- uneven sampling
- exposure variation
- real-time or retrospective
- arbitrary mixtures of data modes
- multivariate data (with or without the constraining the change-points to be the same)
- higher dimensions
- auxiliary information



What good are Segmented Time Series Representations?

Detect and characterize statistically significant variability supported by the data.

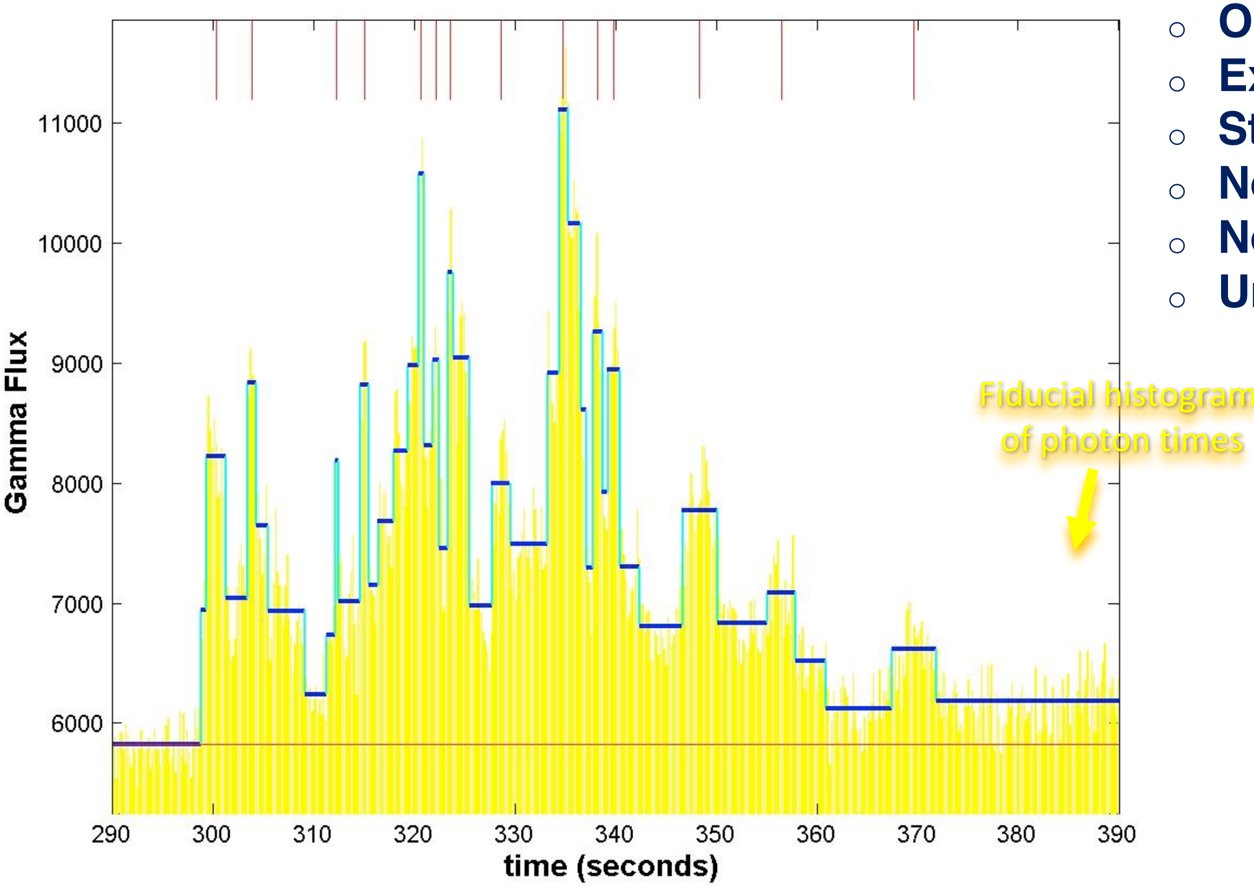
Detect and characterize, without bins or smoothing:

- *Pulses (aka "flares")*
- *Pulse shapes (including the Arrow of Time)*
- *Variability index*
- *Variability time scales (min, max, distribution, …)*
- *Transient event triggers (real-time mode)*

… and implement:

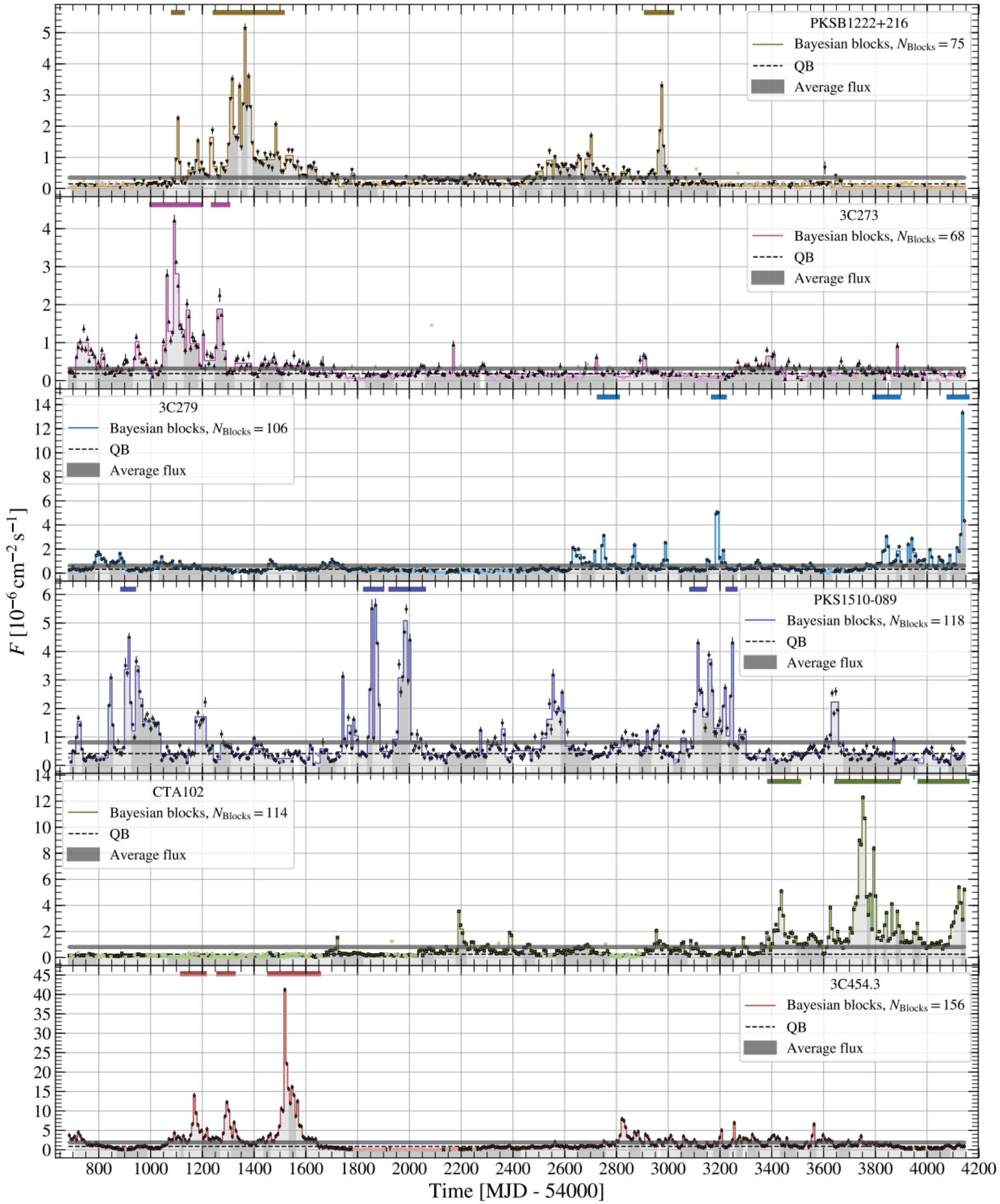
- *Exploratory Data Analysis*
- *Time series classification*
- *Noise suppression*
- *Visual displays*
- *Data compression*
- *Data adaptive histograms*

Studies in Astronomical Time Series Analysis. VI. Bayesian Block Representations
Jeffrey D. Scargle, Jay P. Norris, Brad Jackson, James Chiang arxiv.org/abs/1207.5578

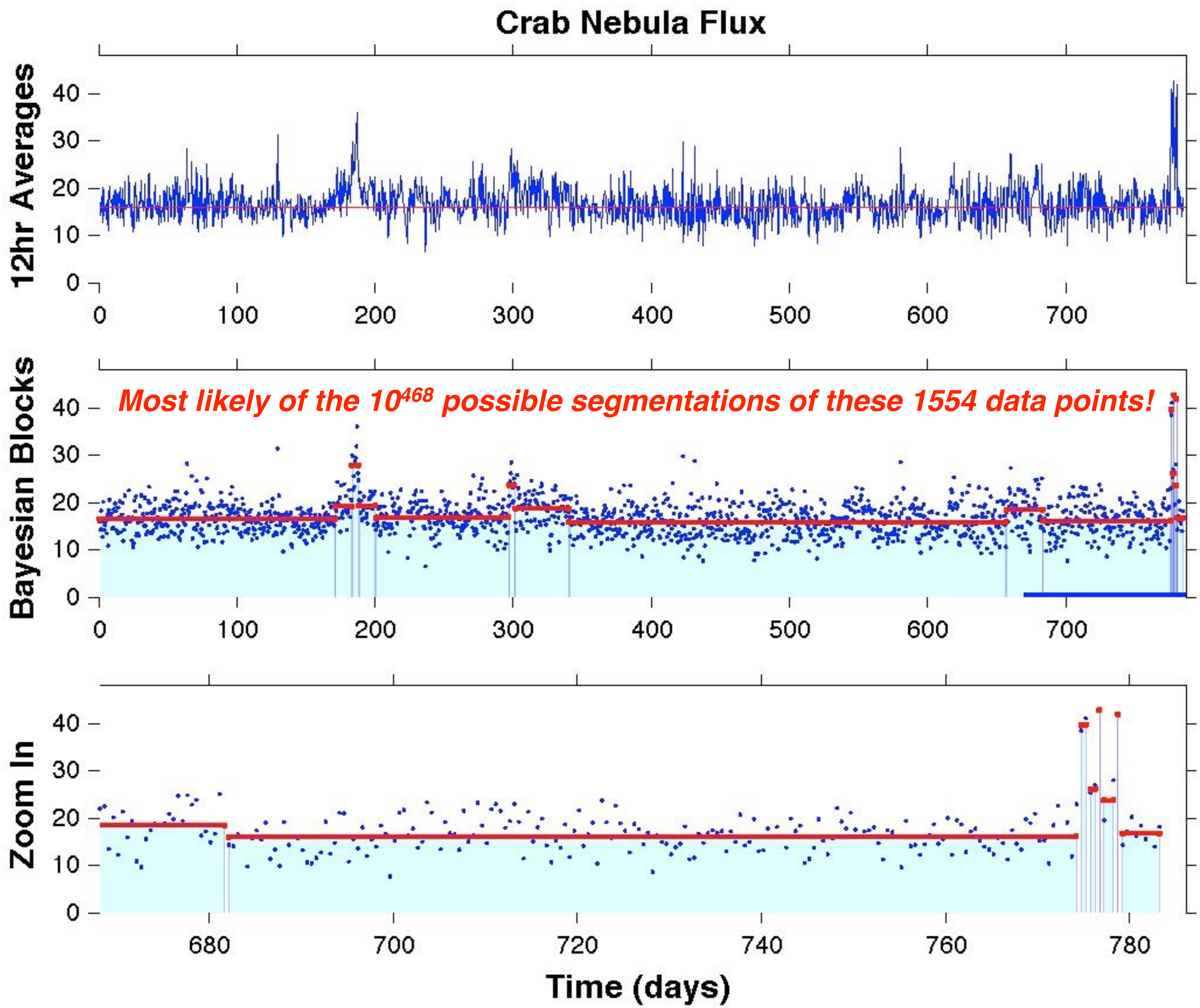


- **Flexible w.r.t. data mode**
- **O(N²) or faster**
- **Exact global optimum**
- **Statistically significant features**
- **Noise suppressed**
- **No smoothing**
- **Unlimited dynamic range**

Bellman, R. 1961, On the approximation of curves by line segments using dynamic programming, Communications of the ACM, 4, 284.


PKSB1222+216
Bayesian blocks, $N_{\rm Blocks} = 75$
QB
Average flux
3C273
Bayesian blocks, $N_{\rm Blocks} = 68$
QB
Average flux
3C279
Bayesian blocks, $N_{\rm Blocks} = 106$
QB
Average flux
PKS1510-089
Bayesian blocks, $N_{\rm Blocks} = 118$
QB
Average flux
CTA102
Bayesian blocks, $N_{\rm Blocks} = 114$
QB
Average flux
3C454.3
Bayesian blocks, $N_{\rm Blocks} = 156$
QB
Average flux
$F\ [10^{-6}\mathrm{cm^{-2}s^{-1}}]$
Time [MJD - 54000]

Crab Nebula rises above the status of a constant calibration source!



Standard Methods to Fix the Number and Width of Bins in Histograms …

	Method formula bins	Method formula width
Square-root	$\sqrt{n}$	$\frac{\max(values) - \min(values)}{\sqrt{n}}$
Sturges 1926	$\mathrm{ceil}(\log_2 n) + 1$	$\frac{\max(values) - \min(values)}{\mathrm{ceil}(\log_2 n) + 1}$
Rice 1944	$2 * \sqrt[3]{n}$	$\frac{\max(values) - \min(values)}{2 * \sqrt[3]{n}}$
Scott 1979	$\frac{\max(values) - \min(values)}{3.5 * \frac{stdev(values)}{\sqrt[3]{n}}}$	$3.5 * \frac{stdev(values)}{\sqrt[3]{n}}$
Freedman-Diaconis 1981	$\frac{\max(values) - \min(values)}{2 * \frac{IQR(values)}{\sqrt[3]{n}}}$	$2 * \frac{IQR(values)}{\sqrt[3]{n}}$

BATSE trigger 00551

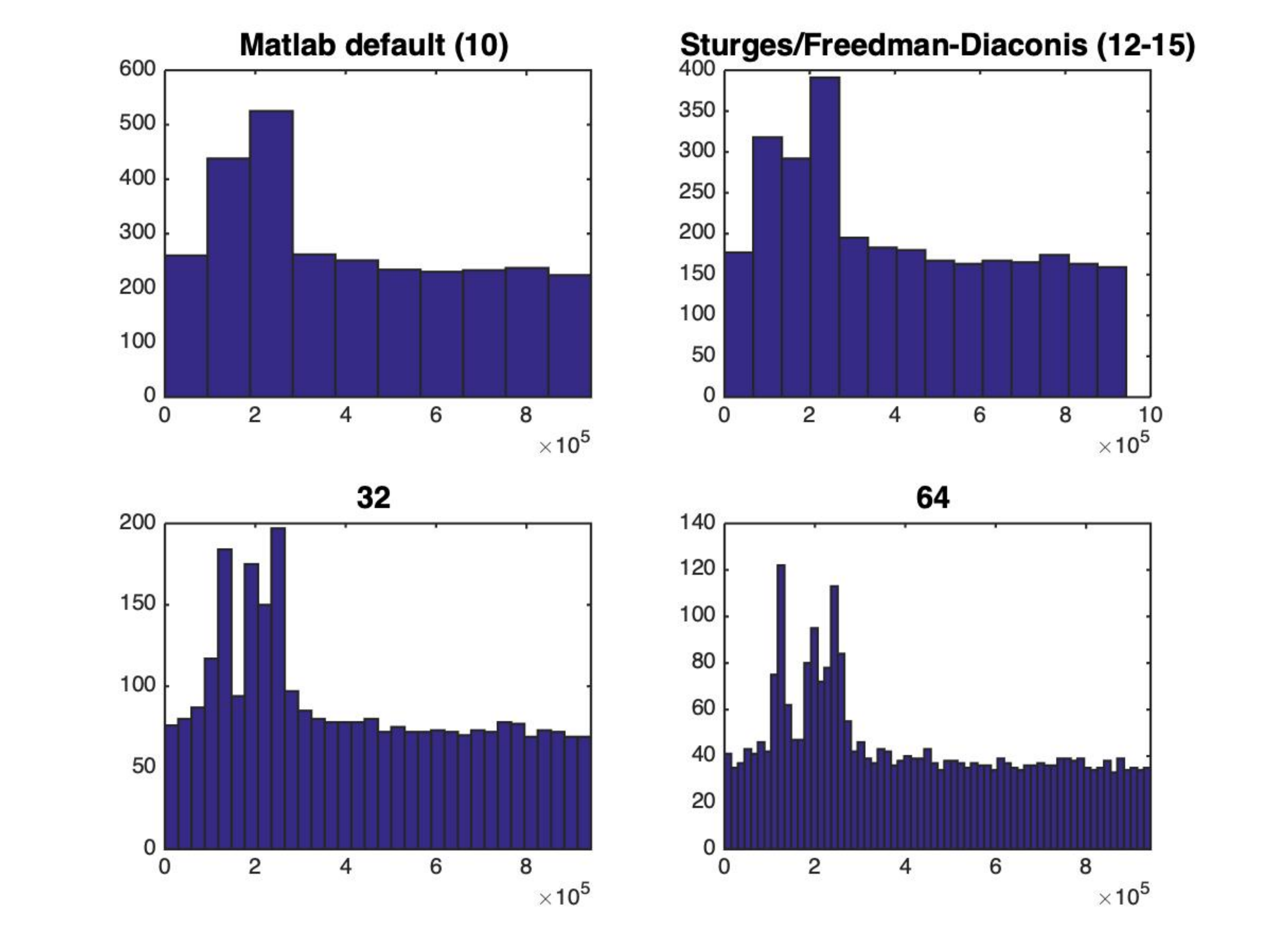
N = 2894 photons

1
2
3
458
740
1084
1628
2050
2446
2770
3058
3388
4078
4284
4660
5028
5336
5648
5860
6230
6538
6728
7198
7542
…
941816
942270
942644

The problem with Sturges' rule for constructing histograms, Rob J Hyndman (1995)

https://robjhyndman.com/publications/sturges/

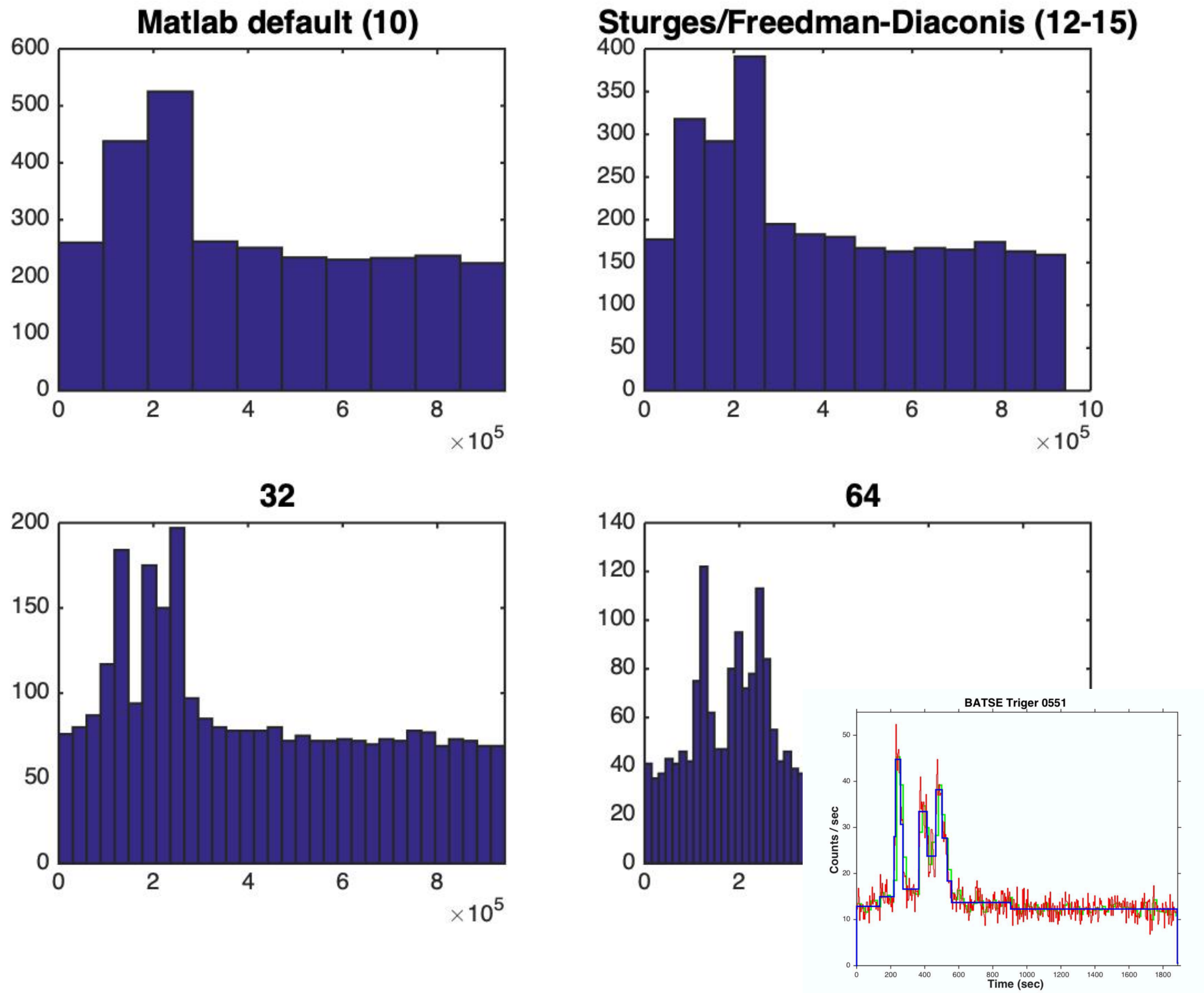
"It is known that Sturges' rule leads to oversmoothed histograms, but Sturges' derivation of his rule has never been questioned. … the argument leading to Sturges' rule is wrong …"



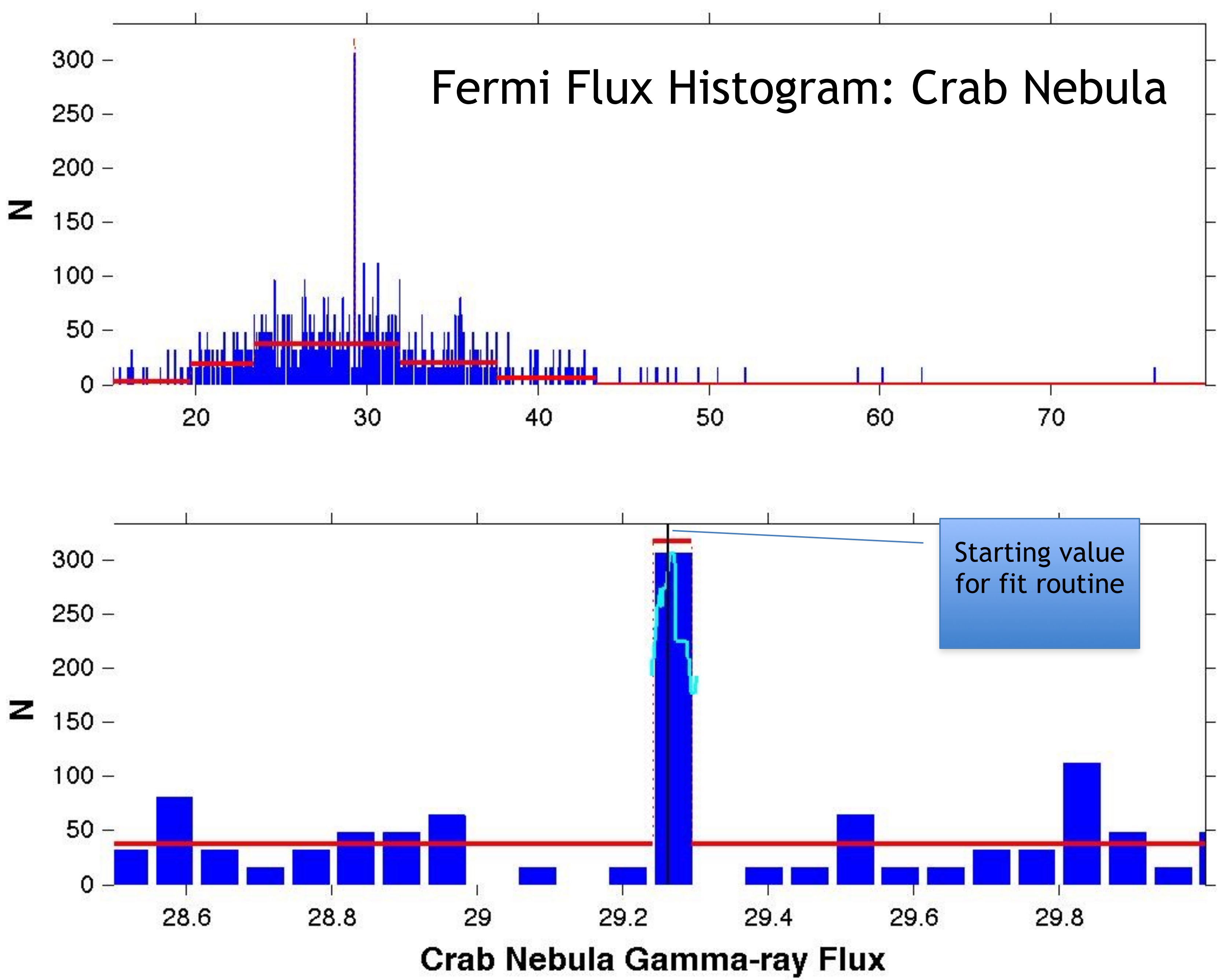
BATSE trigger 00551
N = 2894 photons

So what is the best number of bins?
Scott, K. Knuth, etc. … ?

1
2
3
458
740
1084
1628
2050
2446
2770
3058
3388
4078
4284
4660
5028
5336
5648
5860
6230
6538
6728
7198
7542
…
941816
942270
942644

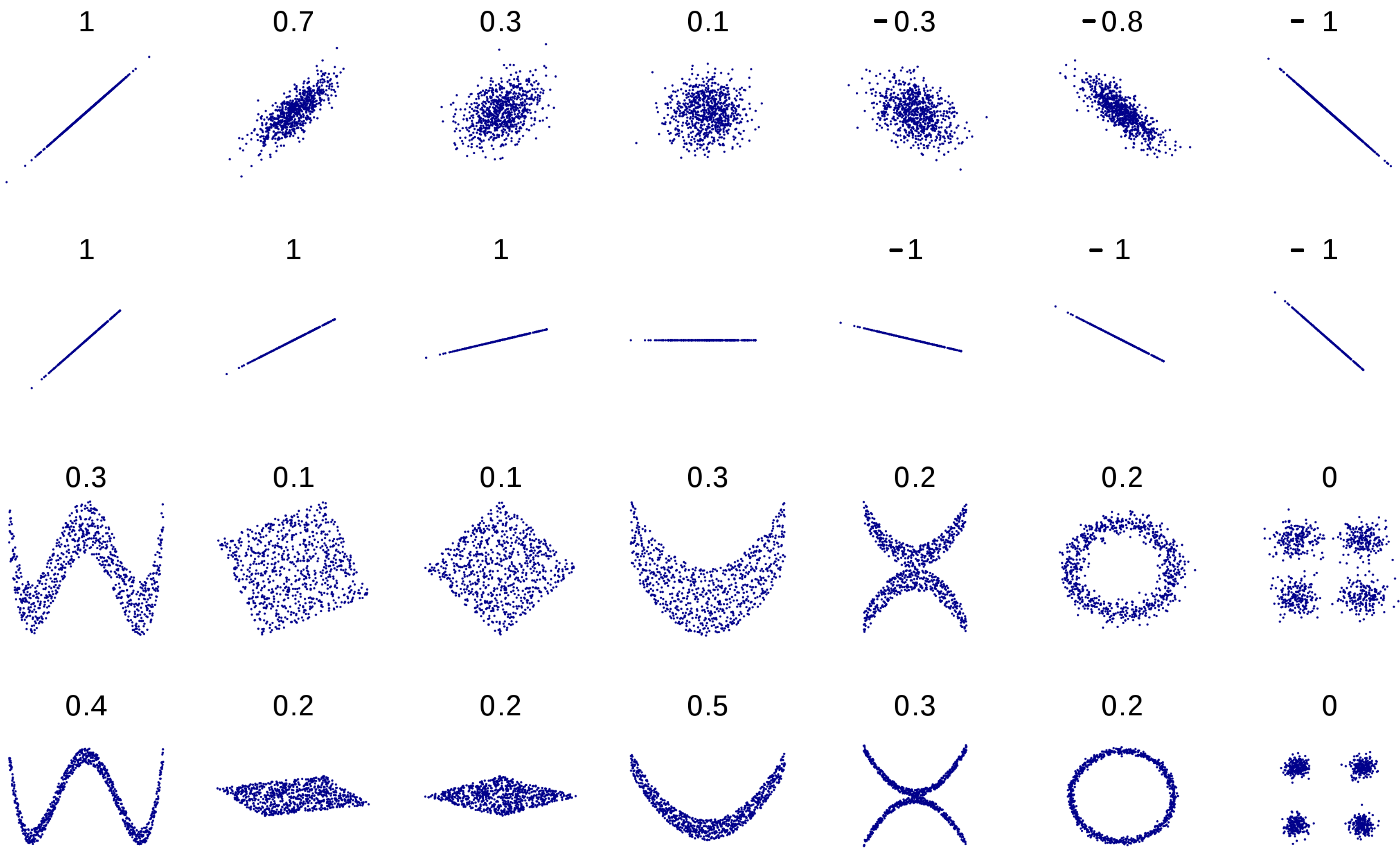


Better yet, discard constraint of equal bins: *Bayesian Blocks*


Fermi Flux Histogram: Crab Nebula
Starting value for fit routine
N
Crab Nebula Gamma-ray Flux

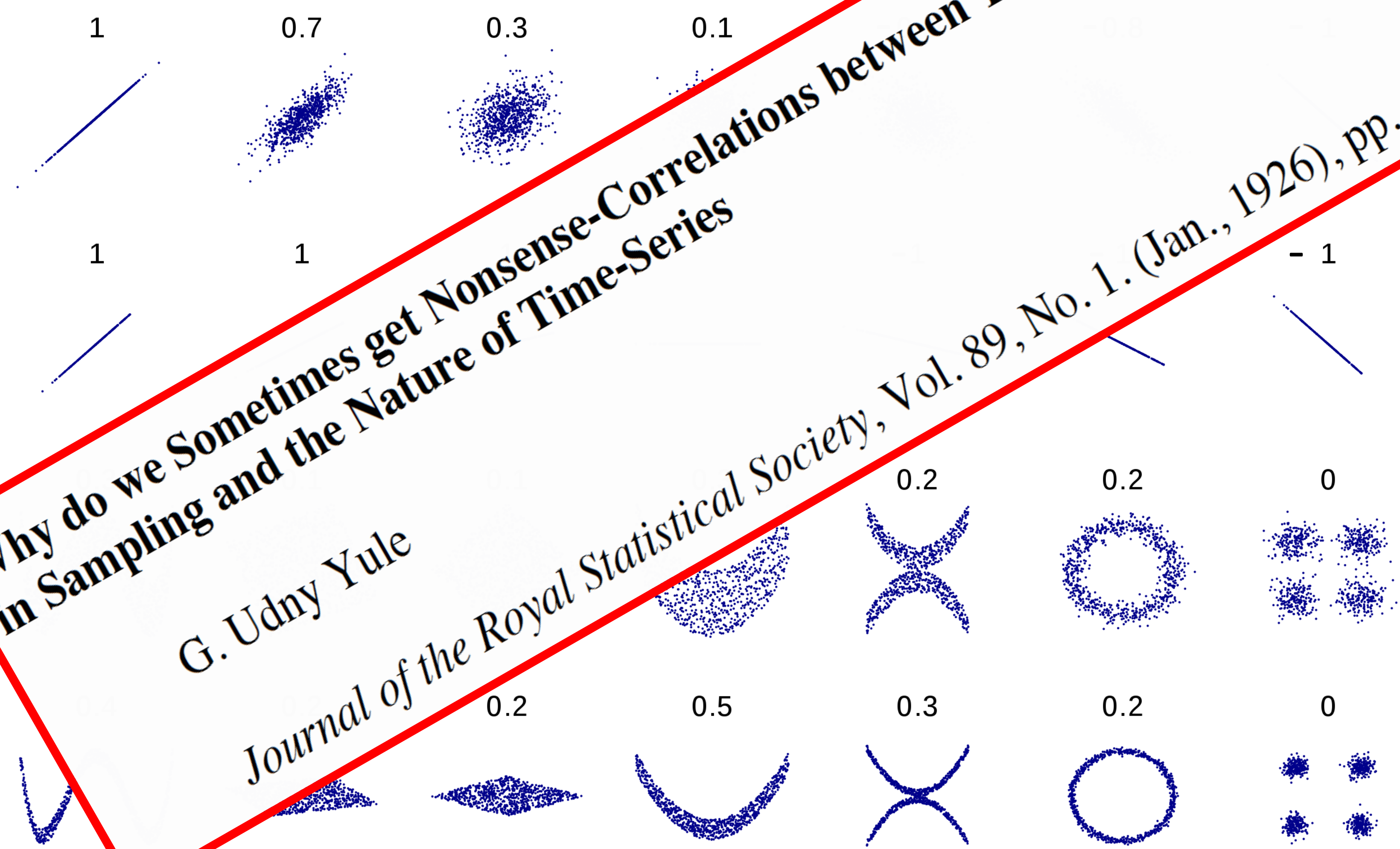
Given two different flux time series: how related are they?

The ubiquitous **correlation coefficient** characterizes linear relations … but is insensitive to non-linear ones.



Given two different flux time series: how related a[illegible]hey?

The ubiquitous **correlation coefficient** ch[illegible]
linear relations … but is insensitive to [illegible]



Why do we Sometimes get Nonsense-Correlations between Time-Series?--A Study in Sampling and the Nature of Time-Series

G. Udny Yule

Journal of the Royal Statistical Society, Vol. 89, No. 1. (Jan., 1926), pp. 1-63

Hierarchy of Degrees of Relationship between two processes X and Y

(Inner Product)

Uncorrelated: $E[\,X(t)\,Y(t)\,] = 0$

(Expectations)

Martingale Property: $E[\,X(t) \mid Y(t)] = E[\,X(t)\,]$

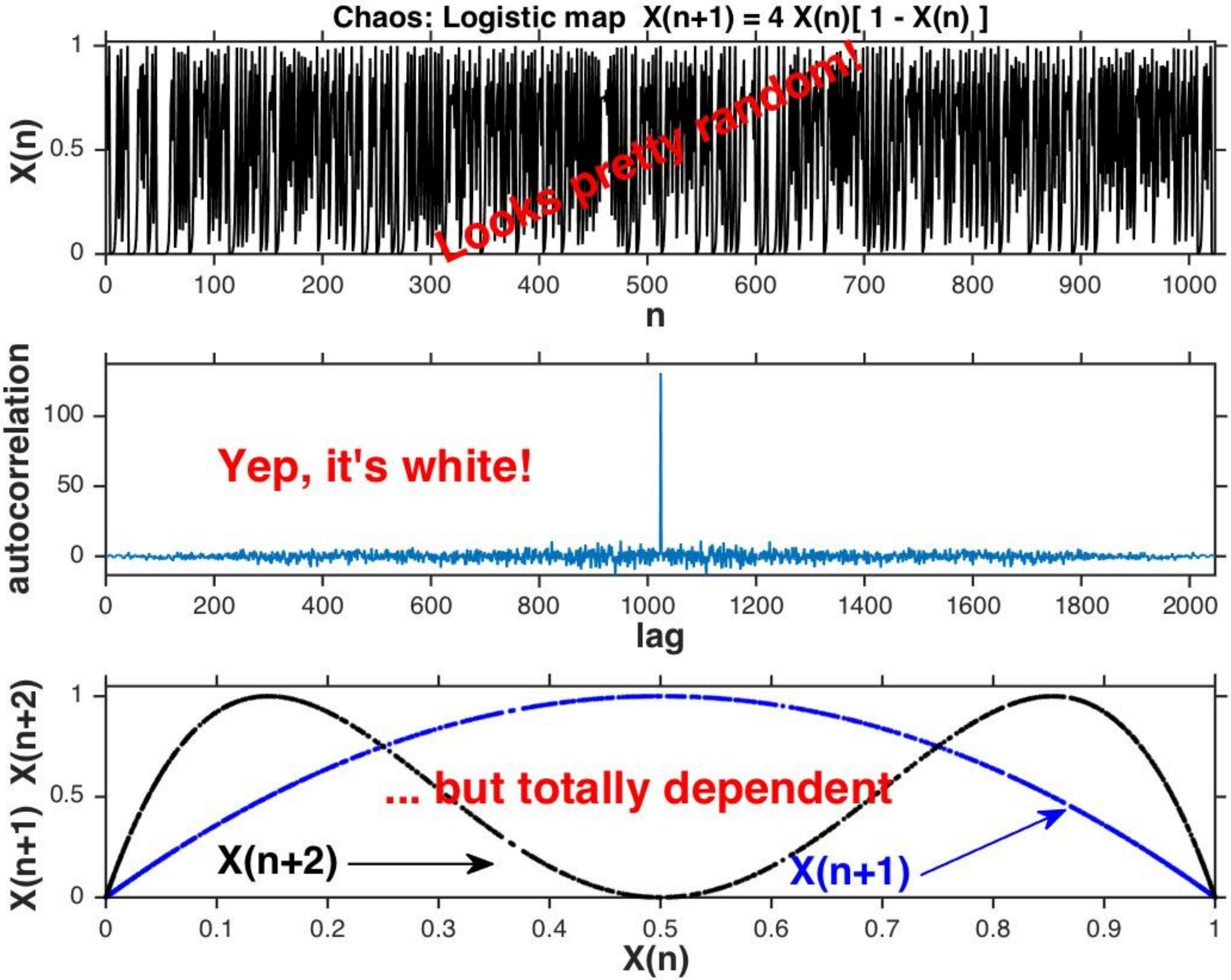
$E[\,Y(t) \mid X(t)] = E[\,Y(t)\,]$

(Distributions)

Independent: $F(X,Y) = F(X)\,F(Y)$

(F: PDF, CDF, characteristic function)

Example: a white noise process for which correlation is useless



Measure (in)dependence between two time series

$$D(X,Y) == < F(X,Y) \, , \, F(X) \, F(Y) >$$

F can be:

- Differential Probability Distribution
- Cumulative Probability Distribution
- Characteristic Function …

$< \; , \; >$ can be:

- Mean Square Difference
- Mutual Information
- D.Wolpert’s Bayesian Histogram Comparator
- Earth-Mover’s Distance
- Copula, etc., etc., etc.

Computing cumulative distributions F(X,Y) , F(X), F(Y)

Given data: $x(t_n)$, $y(t_n)$, $n = 1, 2, \ldots, N$

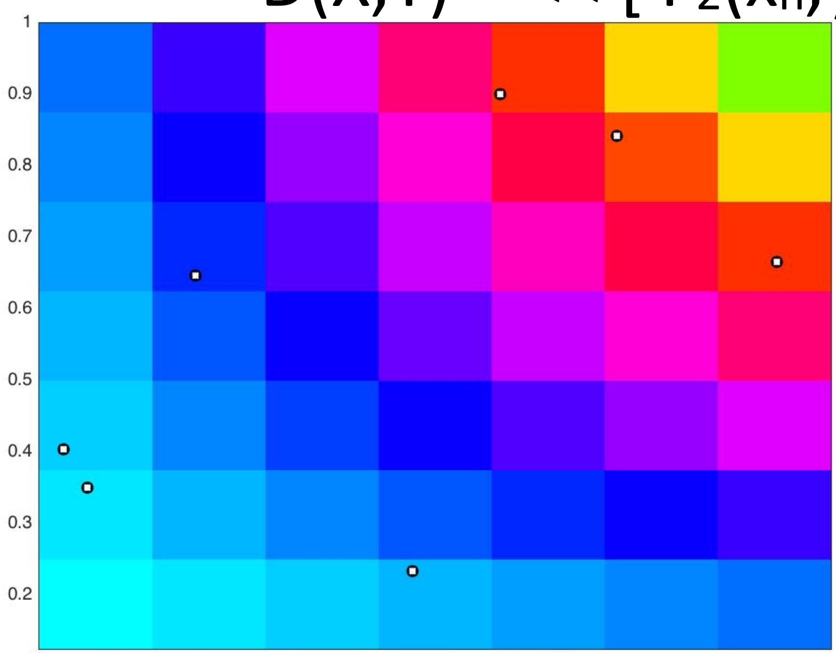
(t_n need not be evenly spaced)

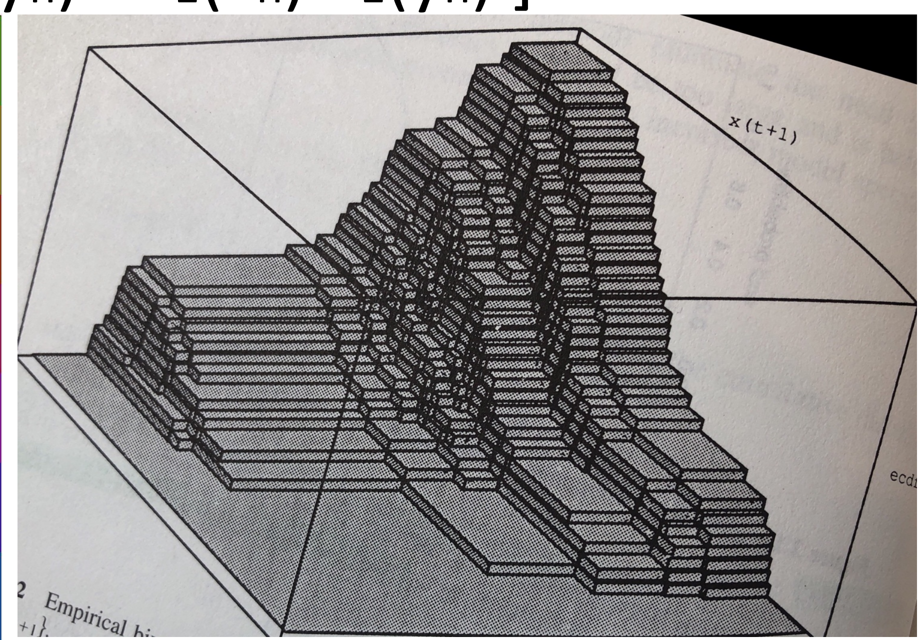
$$F_2(x_n,y_n) = \#(x <= x_n \text{ and } y <= y_n) / N$$

$$F_1(x_n) = \#(x <= x_n) / N$$

$$F_1(y_n) = \#(y <= y_n) / N$$

$$D(X,Y) = << [\, F_2(x_n,y_n) - F_1(x_n)\, F_1(y_n) \,]^2 >>$$





Why are Dependence Measures not Used as much as Correlations?

Correlations (and Correlation Functions)

- Bi-Linear in the Data
- Directly computable from the time series
- Widely used standard tool (the Lemming effect)
- Easily computable

Dependence Measures (and Functions)

- Not Bi-Linear in the Data
- Require estimates of probability distributions

(But bin-free estimates are straightforward)

- Many ways to measure F(X,Y) vs. F(X)F(Y)
- Theoretical analysis difficult

Dependence in the Statistics Literature

Goodman (1997)
Statistical Methods … the Midway View of Nonindependence
The Practice of Data Analysis, Essays in Homor of John W. Tukey, Princeton U. Press

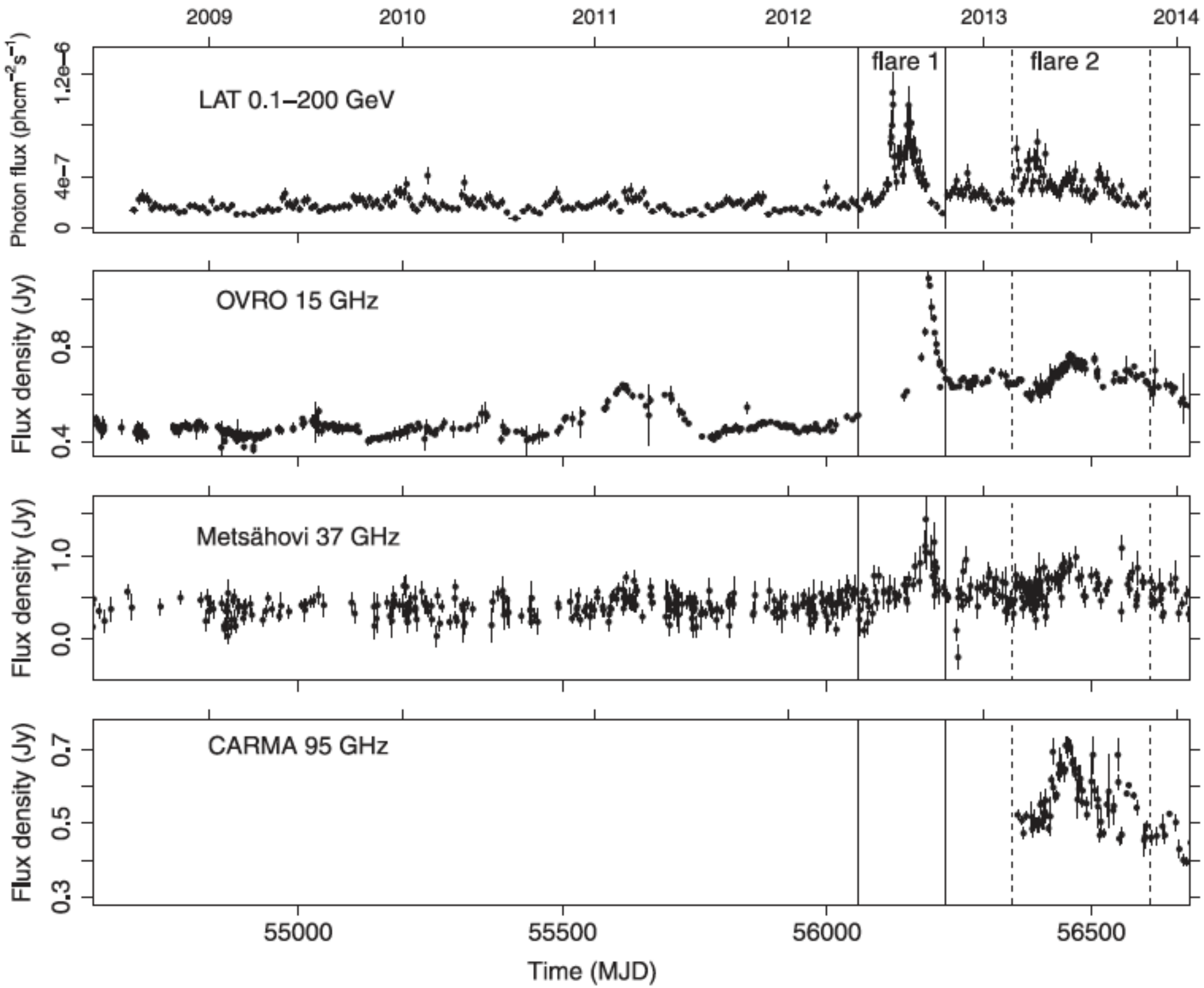
Szekely, Rizzo and Bakirov (2007)
Measuring and Testing Dependence by Correlation of Distances
Annals of Statistics, 35, 2769-2794.

Ding and Li (2015)
Copula Correlation: An Equitable Dependence Measure and Extension of Pearson's Correlation 1312.7214

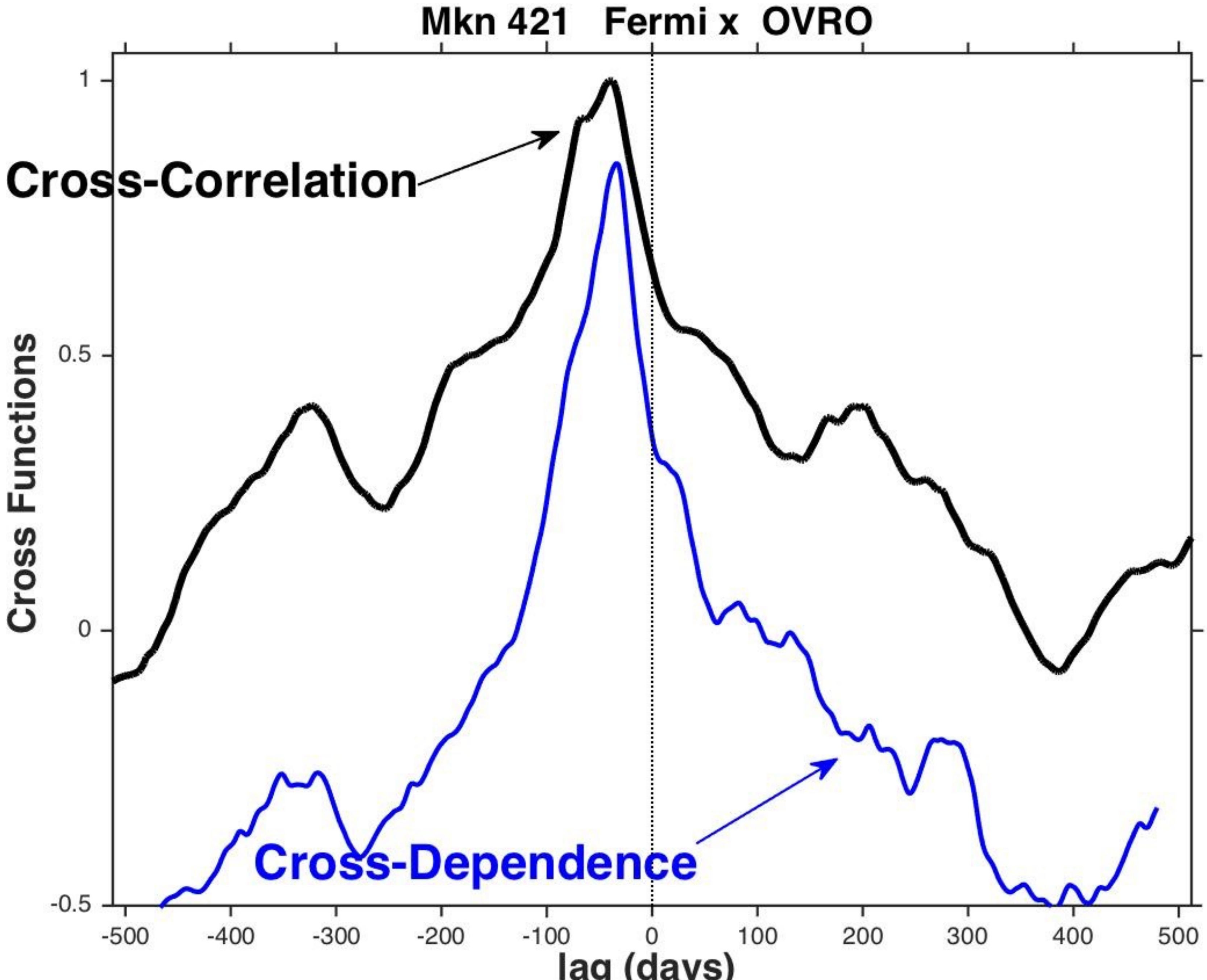
Yakir et al (2016)
Measuring Dependence Powerfully and Equitably 1505.02213

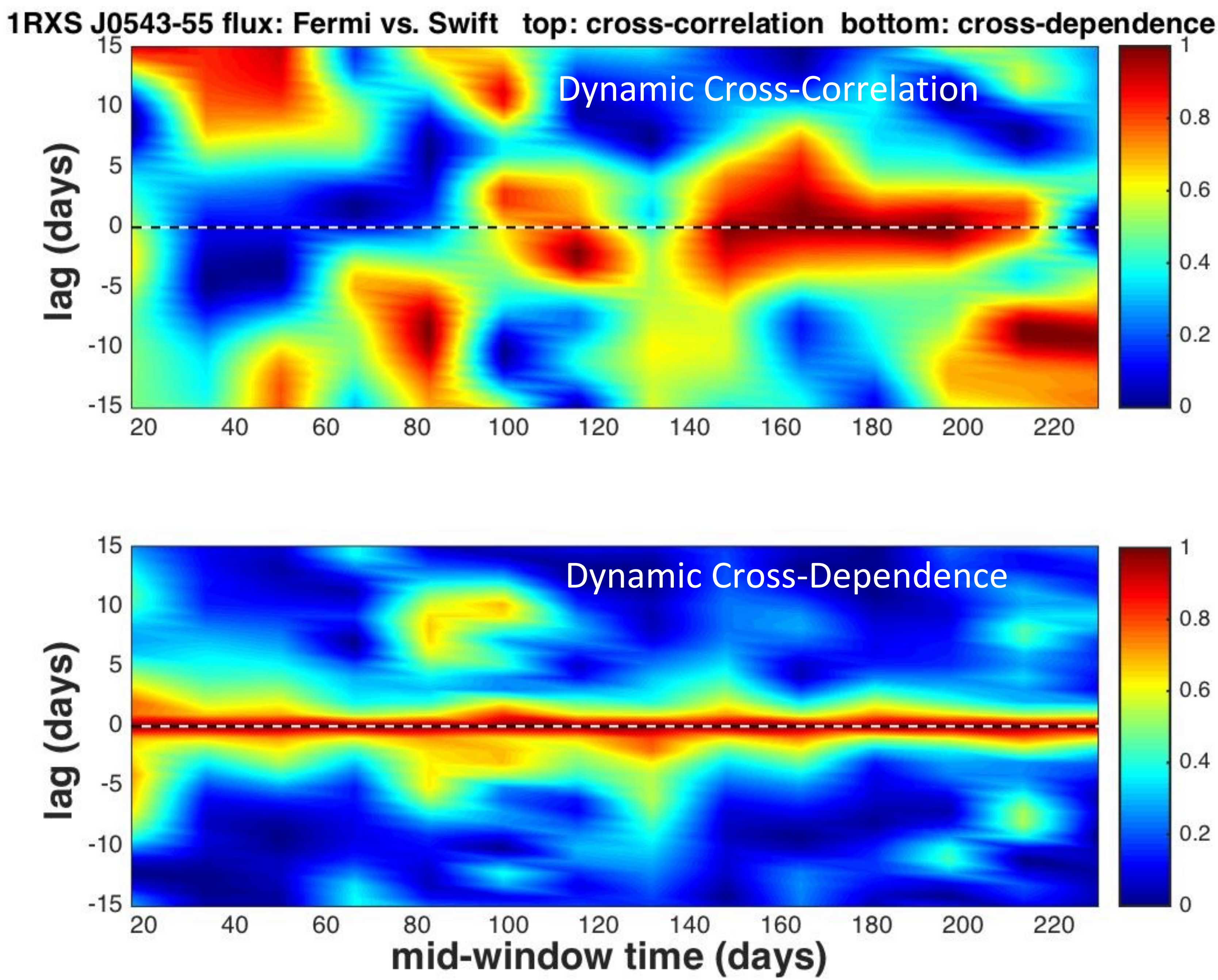
Richards (2017)
Distance Correlation: A New Tool for Detecting Association and Measuring Correlation Between Data Sets 1709.06400

Kagan and Szekely (2019)
Calibrating Dependence between Random Elements 1903.04663



A combined radio and GeV γ -ray view of the 2012 and 2013 flares of Mrk 421
T. Hovatta et al., MNRAS 448, 3121–3131 (2015)


Mkn 421 Fermi x OVRO
Cross-Correlation
Cross-Dependence
Cross Functions
lag (days)
1
0.5
0
-0.5
-500
-400
-300
-200
-100
0
100
200
300
400
500


1RXS J0543-55 flux: Fermi vs. Swift top: cross-correlation bottom: cross-dependence
Dynamic Cross-Correlation
Dynamic Cross-Dependence
lag (days)
mid-window time (days)
15
10
5
0
-5
-10
-15
20
40
60
80
100
120
140
160
180
200
220
1
0.8
0.6
0.4
0.2
0

Bayesian Blocks Bibliogrpahy

Studies in Astronomical Time Series Analysis.
VI. Bayesian Block Representations
Scargle, J., Norris, J., Jackson, B. and Chiang, J.
The Astrophysical Journal, 764, 167 (2013)
Our Blog: http://bayesianblocks.blogspot.com/

Jake Vanderplas' Blog ***Dynamic Programming in Python: Bayesian Blocks***
http://jakevdp.github.com/blog/2012/09/12/dynamic-programming-in-python/

Starship Asterisk* APOD and General Astronomy Discussion Forum
Bayesian Blocks: Detecting local variability in time series
http://asterisk.apod.com/viewtopic.php?f=35&t=29458

An algorithm for optimal partitioning of data on an interval
Jackson, Scargle, Barnes, Arabhi, Gioumousis, Gwin, Sangtrakulcharoen, Tan, Tun Tao Tsai
IEEE Signal Processing Letters, 2005, 12, 105

Studies in Astronomical Time Series Analysis.
V. Bayesian Blocks, a New Method to Analyze Structure in Photon Counting Data
Scargle, 1998, Astrophysical Journal, 504, 405

Obsolete algorithm

The Wold Theorem: Any Stationary Process has an exact Moving Average (and/or Auto-regressive) Representation

$$X = C * R + D$$

(random + deterministic)

Gaussian R —> AR = Gauss-Markov = OU

Auto-regressive (AR): $X(n) = \sum_{k=1}^{\infty} A(k)X(n-k) + R(n)$

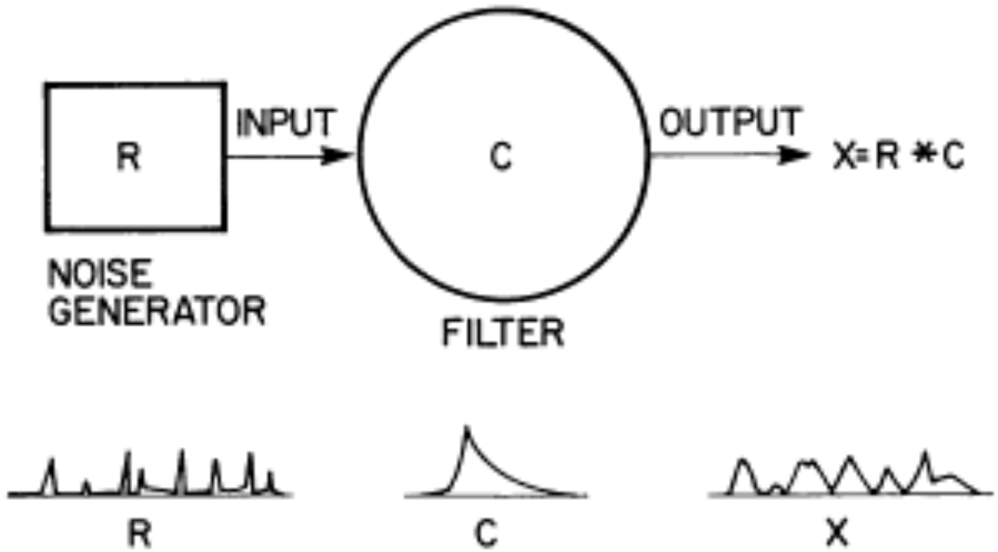
Memory Random Driver

Moving Average (MA): $X(n) = \sum_{k=0}^{\infty} C(k)R(n-k)$

Shot (Filtered) Noise

AR/MA Equivalence: $\sum A(k)X(n-k) \leftrightarrows \sum C(k)R(n-k)$

The Moving Average as a Shot Noise Process



In the Wold Representation the filter C (pulse) is

- minimum delay
- causal
- constant in time

!

The Arrow of Time?

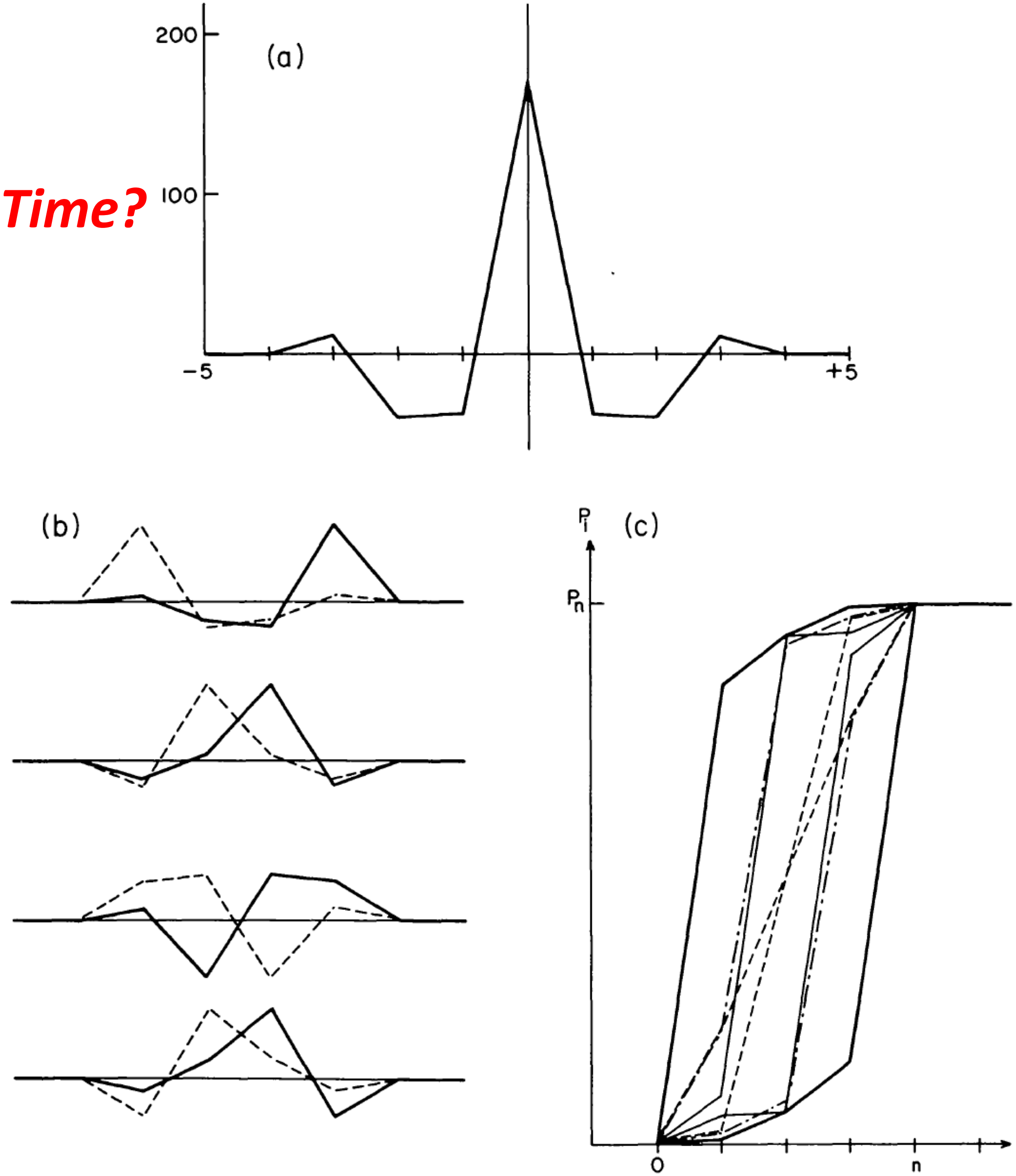


FIG. 14.—The concepts of minimum and maximum delay. (a) A short autocorrelation function. (b) The set of eight pulses which share this autocorrelation. (c) a plot of the eight corresponding partial energy curves: the uppermost curve corresponds to the minimum delay pulse (*dashed line*, topmost part of [b]) and the lowest curve corresponds to the maximum delay pulse (*solid line*, topmost part of [b]).

Convenient algebra of pulse shapes Implemented via the z-transform.

$$B_m = \sum_k C_k D_{m-k}. \tag{45}$$

Thus, the action of two filters in succession (series) can be completely represented by a single filter, called the *convolution* of the two, written as

$$B = C * D. \tag{46}$$

It is readily verified that *the Z-transform of the convolution of two filters is the product of their Z-transforms*:

$$B(z) = C(z)\,D(z). \tag{47}$$

Generalized Wold Theorem: Any Stationary Process has a family of equivalent, exact AR and/or MA Representations

$$\text{Family of MA representations:} \quad X(n) = \sum_{k=-\infty}^{\infty} C^m(k) R^m(n-k)$$

If the Moving Average has M coefficients
there are 2^M C's (with different, but white, R's)
that yield ***exactly equivalent representations of X***.
One minimum delay, one maximum delay; others mixed delay.

Conjecture: ***the unique R of the "correct" MA is IID.***
—> minimum dependence blind deconvolution!

Two Possible Solutions to the Arrow of Time:
(1) two-sided (acausal) models + new fitness measures
(2) "local" models, such as Bayesian Blocks

All this, and more, in JS, Studies in Astronomical Time Series Analysis:
I. Modeling Random Processes in the Time Domain, ApJS. 1981, 45, 1-71