X-ray Timing Analysis

*Michael Nowak - Chandra X-ray Science Center/MIT

The Questions That We’d Like to Answer:

Does My Source Vary?

On What Time Scales Does it Vary?

Are the Variations Periodic or Aperiodic?

How Do Different Energy Bands Relate to One Another?

* (With Some Judicious Stealing of Slides from Z. Arzoumanian’s 2003 X-ray Astronomy School Talk)
Characteristic Time Scales:

\[ \tau \geq \frac{R}{V}, \quad V \leq c, \quad R \geq 2 \frac{GM}{c^2} \]

- \( \tau \geq 1000 \text{ sec} \) \( 10^8 \, M_\odot \) (AGN)
- \( \tau \geq 100 \, \mu\text{sec} \) \( 10 \, M_\odot \) (BHC)
- \( \tau \geq 75 \, \mu\text{sec} \) \( 1.4 \, M_\odot \) (NS)

These are the Fastest Achievable Time Scales. In Reality, There Can be Variability on a Range of Time Scales.
X 1820-303 (11 minute orbit)

180 Day Superorbital Period

msec QPO
Rotational Periods:
  msec - sec for NS/WD
  hr - days for Stars

Accretion Time Scales:
  Dynamical, Thermal, Viscous Time Scales
  msec - days for NS/BHC
  minutes - years for AGN

Orbital Time Scales:
  minutes to days for NS/BHC
  Suber-orbital periods:
  weeks to months
What are the Tools of the Trade?

- Spectra: XSPEC; Sherpa, ISIS - A Few Hardy Souls Run Their Own
- Timing: Xronos - Which Some People Use
- Most People “Roll Their Own”
  - Custom Fortran/C Code
  - IDL or MATLAB
  - Me: Converting over to S-lang run Under ISIS/Sherpa
  
  (http://space.mit.edu/CXC/analysis/SITAR
  - contributions welcome!)
Timing Starts with a Lightcurve

- Different Spacecraft can have different tools for creating Lightcurves
- ftools, dmtools, xselect
- Always choose integer multiple of “natural” time unit for binning
- Don’t bin any more than you have to - save it for subsequent analysis

Example: 
```
dmextract infile="4u2129_chandra.fits [EVENTS] [sky=region(source.reg)] [bin time::1.14104]"
outfile=4u2129_ps.fits opt=ltc1
```
Length & Binning Determine Limits

Lowest Frequency: \( f_{\text{long}} = \frac{1}{T} \)

Highest Frequency: Nyquist Frequency, \( f_{\text{Nyq}} = \frac{1}{2 \triangle t} \)

Basic Question, is the Variance:
\[
\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2
\]
Greater than Expected from Poisson Noise?

\( \sigma = \text{Root Mean Square Variability} \)
Variability Test I: Excess Variance

Binned Lightcurve with Values: \( X_i \pm \sigma_i \) and mean: \( \mu \)

\[
\sigma_{\text{rms}}^2 = \frac{1}{N \mu^2} \sum_{i=1}^{N} \left[ (X_i - \mu)^2 - \sigma_i^2 \right]
\]

\[
\Delta \sigma_{\text{rms}}^2 = s_D / (\mu^2 \sqrt{N})
\]

\[
s_D^2 = \frac{1}{N - 1} \sum_{i}^{N} \left( \left[ (X_i - \mu)^2 - \sigma_i^2 \right] - \sigma_{\text{rms}}^2 \mu^2 \right)^2
\]

Test II: Kolmogorov-Smirnov

- Technique for determining whether two cumulative distributions are the same.

- Example: Is cumulative arrival time consistent with constant rate?

- Could have instead done distribution of times inbetween events.

- See Press et al., “Numerical Recipes”, plus lots of other better statistics books.

- Only answers whether there is variability - doesn’t characterize it.

- Significance = $8 \times 10^{-5}$

- $D$ = Maximum Deviation

- Observed Arrival Time Distribution

- Uniform Rate Distribution

- Cumulative Probability

- Time-Time[0] (sec.)
Bayesian Methods Don’t Require Binning
(Case below: event times only!)


Bayesian Blocks (J. Scargle, in prep.) - Determines Optimal Non-uniform Binning. (S-lang Version on SITAR page)

Drawbacks: No ‘Frequentist’ Significance Levels. Only ‘Odds Ratios’ or ‘Penalty Factors’.

t = fits_read_col("4u2129_chandra.fits","time");
cell = sitar_make_data_cells(t,2,0.7,1.14104,min(t),max(t));
ans = sitar_global_optimum(cell,3.5,2);
Fourier Transform Methods

The Workhorse of the Timing World

How is Variability Power Distributed as a Function of Frequency?
Fast Fourier Transform (FFT)

\[ X_j \equiv \sum_{k=0}^{N-1} x_k \exp\left(\frac{2\pi i j k}{N}\right), \quad j = [-N/2, \ldots, 0, \ldots, N/2] \]

\[ P_j = 2|X_j|^2 / (\text{Rate}^2 \times T_{\text{total}}) \]  
\( \text{"One Sided" RMS Normalization} \)

\[ P_j = 2|X_j|^2 / (\text{Rate} \times T_{\text{total}}) \]  
\( \text{"One Sided" Leahy Normalization} \)

Lightcurve with: N bins, Comprised of Counts, \( x_i \), becomes Power Spectrum, with \( N/2+1 \) independent Amplitudes, and \( N/2-1 \) independent Complex Phases (for Real Inputs)

Good FFTs Usually Optimized for \( N = \text{Power of 2} \) (RXTE Clock Runs in Powers of 2!)

Know Your Normalization!!! Various FFT Routines Have Different Ones!

Power Spectrum is the Squared Fourier Amplitude, Properly Normalized

Power Spectrum is Throwing Out Information! Not Unique!
**Ninja Topic: Convolution/Cross Correlation Theorem**

\[ h \star g (t) \equiv \int_{-\infty}^{\infty} h(\tau)g(t + \tau) \, d\tau \]

\[ G(f) \equiv \mathcal{F}[g(t)] \equiv \int_{-\infty}^{\infty} g(t) e^{-2\pi if t} \, dt \]

\[ \mathcal{F}[h \star g] = F^*(f)G(f) \]

---

- **Power Spectral Density (PSD, or Power Density Spectrum/PDS)** is just the Fourier Transform of the Auto-correlation Function (i.e., \( h(t) = g(t) \)).

- **Cross Power Spectral Density (CPD)** is just the Fourier Transform of the Cross-correlation Function.
- Leahy: Poisson Noise Level = 2, Intrinsic Power Scales as Rate
- RMS: Intrinsic Power Independent of Rate, Noise Level = 2/Rate
- Integral of PSD is Measure of Root Mean Square Variability

\[ A = \int P_{\text{rms}} \, df = \sum_j P_{\text{rms}}^j \Delta f \quad , \quad \Delta f = 1/T \]

\[ \sqrt{A} = \text{rms/mean} = \left( \frac{\langle x^2 \rangle - \langle x \rangle^2}{\langle x \rangle^2} \right)^{1/2} \]

Pulsed Fraction (Coherent Oscillation): \[ f_P = \sqrt{\frac{2(P_{\text{Leahy}} - 2)}{\text{Rate}}} \]

PSD Normalizations are Often Plotted as \((\text{RMS})^2/\text{Hz}\)
PSD Statistics

Leahy Noise Level is 2 +/- 2 (Distributed as $\chi^2$ with 2 DoF)

Increasing Lightcurve Length Doesn’t Help - Distributes Noise Among More Frequency Bins!

“Statistically Stationary Processes” Have Power = $P_i +/- P_i$

Reduce Noise by Averaging PSD from Individual Lightcurve Segments, as Well as Over (Usually Logarithmically Spaced) Adjacent Frequency Bins

Errors Reduced by Factor of: $\sqrt{N_{avg}}$
Example: Cyg X-1 (RXTE)

"Band Limited Noise"

White Noise

Red Noise

Effective Noise Level \( \propto 1/\sqrt{f} \)

With: \[ P'_j = (P_j - P_{\text{noise}}) \pm P_j / \sqrt{N_{\text{avg}}} \]

You Can Fit Models

Note: Total RMS = Incoherent Sum of Components, i.e., \( \left( \sum_i \text{RMS}_i^2 \right)^{1/2} \)

Advice: Fit Models that Average over Frequency Bin Widths
Quasi-Periodic Oscillations (QPO)

Q-value (Coherence) = \( \frac{f}{\Delta f_{\text{FWHM}}} \)

Width Can Come From: Finite Length of Data Segment, Finite Duration of Signal, Random Walk in Phase (e.g., Damped, Driven Oscillators), Random Walk in Frequency, ...
Ninja Topic: Deadtime

Detector ‘Deadtime’ = When a Photon Event Prevents Subsequent Events from Being Detected (‘Paralyzable’ / ‘Non-paralyzable’ is When an Event During the Deadtime Does/Does Not Increase the Length of the Deadtime), or ...

When the Detector Does Not Take Data, e.g., during Readout (e.g., Chandra), or ...

Deadtime Modifies the Power Spectrum of Poisson Noise from the Expected $P_{\text{Leahy}} = 2$ (Usually to Something $< 2$)

Proposal Estimates

Detecting Broad Band Noise at the $n_\sigma$ confidence level:

$$\text{RMS}_{\text{limit}}^2 \approx 2n_\sigma \sqrt{\Delta f} / \sqrt{\text{Rate}^2 \times T_{\text{total}}}$$

Detecting Coherent Pulsations:

$$f_p^{\text{limit}} = 4n_\sigma / (\text{Rate} \times \text{Time})$$

For Broad Band Timing, You Win More with Rate than Time

Searches for Coherent Pulsations (e.g., Pulsars) are Best Done Unbinned
Coherent Pulsations:

- Barycentering the data (fxbary/axbary) important.
- Short Data Segments, to Search for (Binary) Orbital Variation
Ninja Topic: Aliasing!

Signal Can Appear at Sum and Difference Frequencies of Primary Signals
This is True Whether the Signal is “Real” or “Fake” (e.g., Sampling Periods)
Beware Characteristic Times! Spacecraft orbits, dither time scale, 1 year, ...
Example: RXTE-All Sky Monitor - Many sources show periods at 24 hours +/- a small bit. This is the beating of a large power 1/Many Year Secular Change with a 24 hour sample Period (e.g., from AGN monitoring).
Ninja Topic: Phase Info

There are Statistics That Also Deal with Fourier Phase - Cross Correlations!
Gives the Frequency Dependent Time-lag between Hard and Soft Components
Ninja Topic: Cross Correlations

Complex Phase is Called the “Phase Lag”, Divided by $2\pi f$ is Called the “Time Lag”

Keep track of signs! Depending upon Algorithm, and Whether You Use Forward or Backward Transform, that Can Alter the Sign. (See Nowak et al. for Associating this with Lag/Lead.)

$\Upsilon^2(f)$ is the “Coherence Function” (Distinct from Coherence, Q!). Measures Degree of Linear Correlation.

\[
\langle CPD \rangle = \frac{\langle H^* G \rangle}{\left(\langle |H|^2 \rangle \langle |G|^2 \rangle \right)^{1/2}} , \quad \Upsilon^2(f) \equiv \frac{|\langle H^* G \rangle|^2}{\langle |H|^2 \rangle \langle |G|^2 \rangle}
\]
Epoch Folding & Period Searches

Good for Non-sinusoidal Variations
Good for When there are Data Gaps or Complicated Window Functions
Not Good for Aperiodic Variability

```
event = sitar_readasm("xa_x1820-303_d1",,1.2);
fld = sitar_epfold_rate(event.time,event.rate,10,500,20,2000);
xlabel("Trial Period"); ylabel("L Statistic");
plot(fld.prd,fld.lstat);
```

Xronos has epoch folding, various IDL routines can be found on the web.

Read the literature on significance levels!
Reiterating Words of Advice:

- Bin the Lightcurve on Integer Multiples of “Natural” Time Scales
- Do FFTs with Evenly Spaced Bins (Lomb-Scargle for Unevenly Spaced Bins), and Avoid Data Gaps (see literature if dealing with Gaps)
- Beware of Signals that Appear on Characteristic Time Scales (of Spacecraft, Earth, etc.)
- Large Literature with Many Techniques for Those with Strong Kung Fu

Press et al., “Numerical Recipes” (Discussions Only! Better Code Exists on the Web!)


