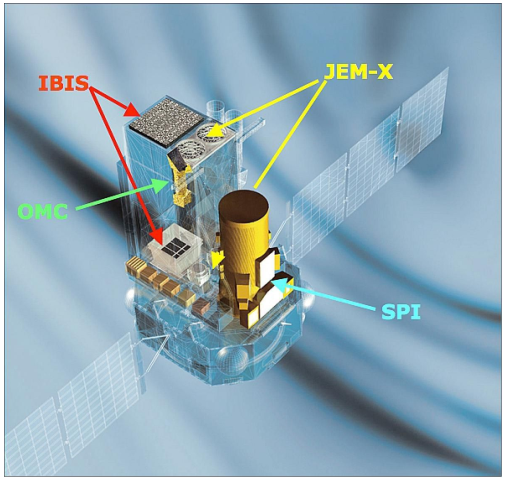
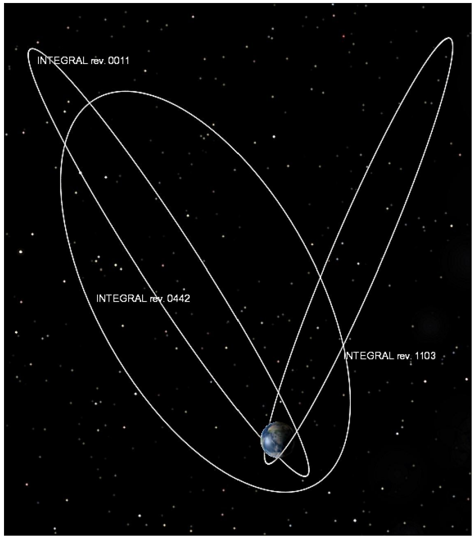
PySPI for Persistent Sources

Julius Möller

About INTEGRAL

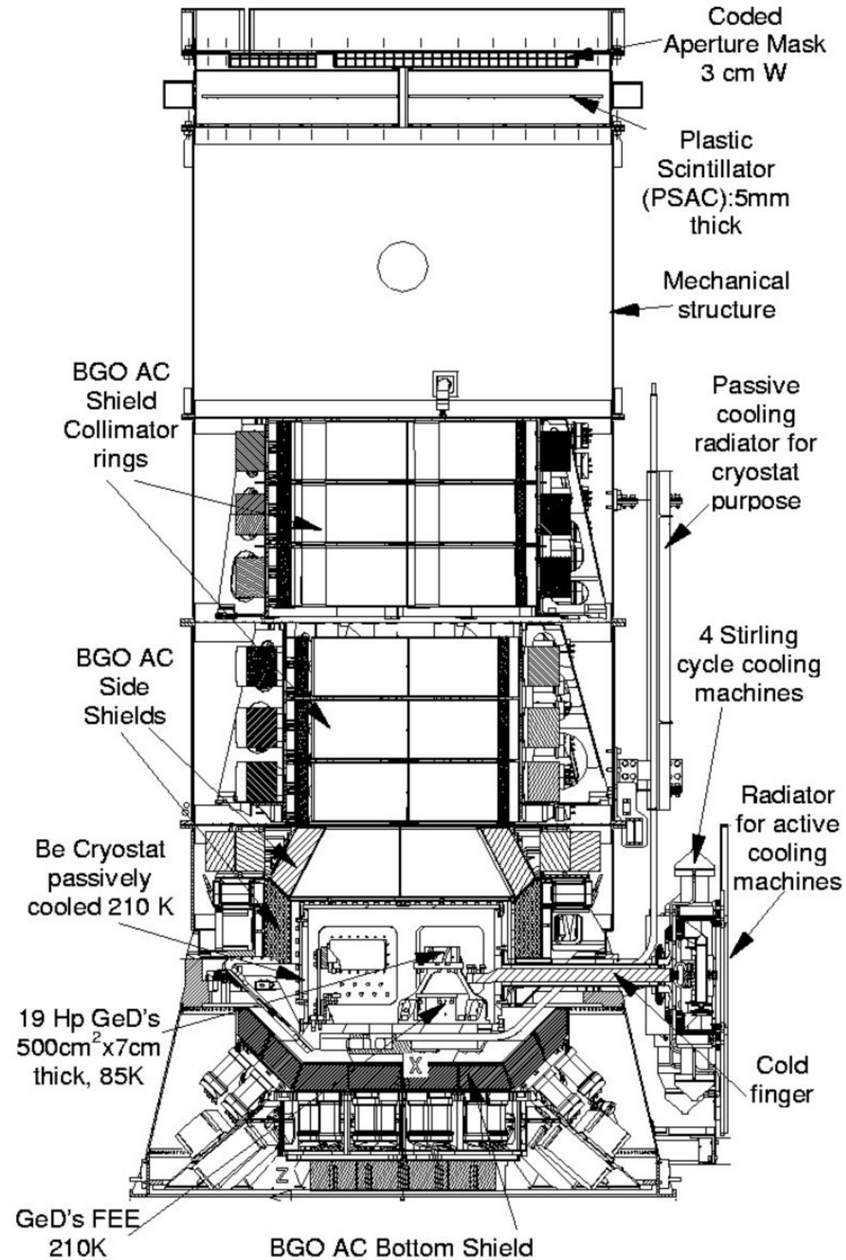
- INTErnational Gamma-Ray Astrophysics Laboratory (INTEGRAL) is an ESA space telescope with contributions from NASA and the RKA
- Launched October 17, 2002
- 3-day highly elliptical orbit
  - 153000km apogee, 9000km perigee
  - Mostly above radiation belts: fewer background count events
- Initial 2+3-year planned lifetime
  - Greatly exceeded due to low fuel consumption
  - Mission was repeatedly extended, currently until end of 2024
  - Reenter Earth's atmosphere in 2029
- INTEGRAL's Instruments:
  - SPectrometer of INTEGRAL (SPI)
  - Imager on-Board the INTEGRAL Satellite (IBIS)
    - 15keV-10MeV, 9°x9° FoV, 12 arcmin angular resolution
  - Joint European X-Ray Monitor (JEM-X)
    - 3-35keV, 4.8° fully illumanted FoV diameter
  - Optical Monitoring Camera (OMC)
    - 500-600nm, 5°x5° FoV





About SPI

- 20keV - 8MeV
- 16° FoV (fully coded, corner to corner)
- Specialized in its high energy resolution: 2.5keV precision at 1.3MeV
- Response Matrices based on extensive ground calibrations and Monte-Carlo simulations
- Notable Systems:
  - Detectors
  - Coded Mask
  - Active Cooling System
  - AC Shield
  - PSD electronics



SPI - Detectors

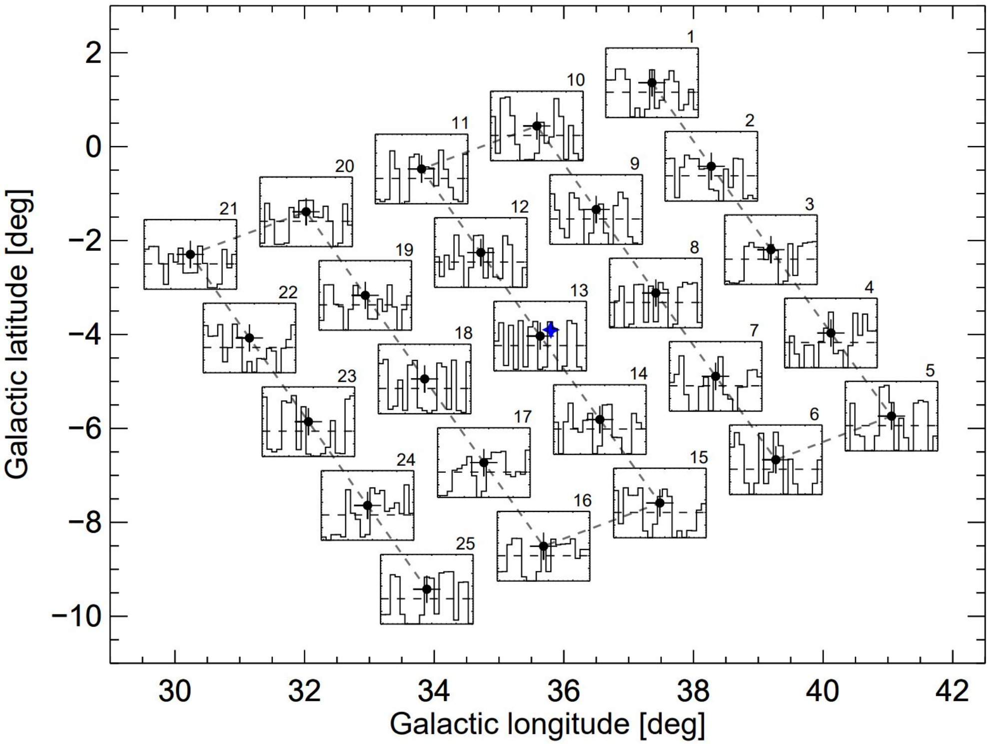
- Hexagonal array of 19 reverse-electrode n-type germanium detectors with 4000V applied to each
  - Ionizing radiation interacts with detector material (Compton scattering, pair creation, photoelectric effect)
  - Electron-hole pair(s) created in semiconductor
  - Swept to p and n electrodes by electric field
  - Measured current fed to preamplifier, allows inference of deposited energy
- Mean weight: 951g, Mean volume: 178cm^3
- Requires below 100K to avoid thermal generation due to low band gap energy of germanium and to reduce radiation damage
- Stirling cycle cryocoolers provide active cooling to detectors
- Radiation damage (hole-trappings) still occurs
- Annealing process heats detectors to 105°C for two days, which undoes radiation damage and restores energy resolution

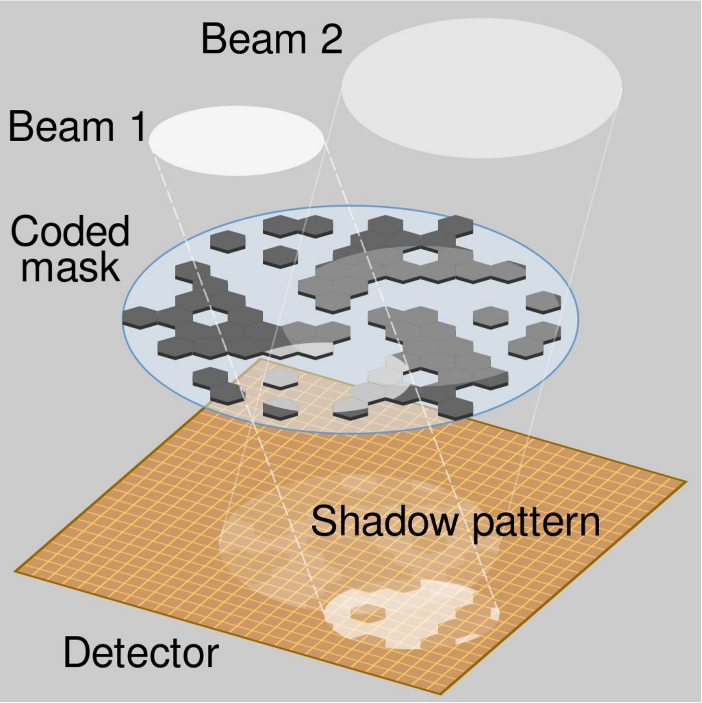
SPI - PSD and ACS

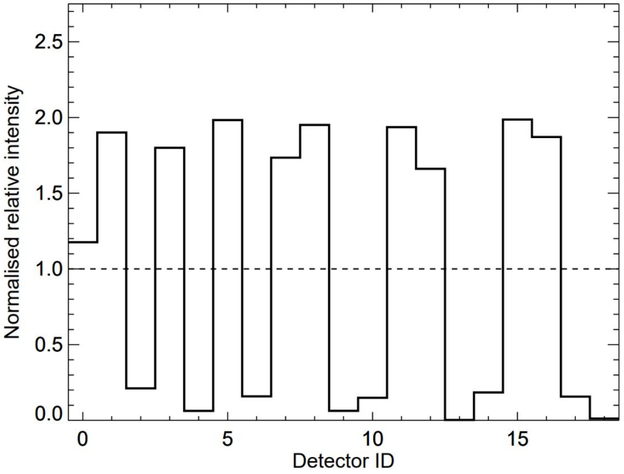
- Each detector is equipped with its own Pulse Shape Discrimination (PSD) electronics
  - Localized Beta Decays in detector material produce single current pulse
  - Regular photon energy deposits involve several compton scatterings resulting in a broader signal
  - “Flag” events depending on type of signal shape to remove activation-decay events
  - Only triggered above variable energy threshold, changed between 400keV and 750keV
- 91 BGO crystals are installed around various SPI components, acting as an Anti-Coincidence Shield (ACS)
  - Whenever an event is simultaneously measured in the ACS and the GeD, it is vetoed

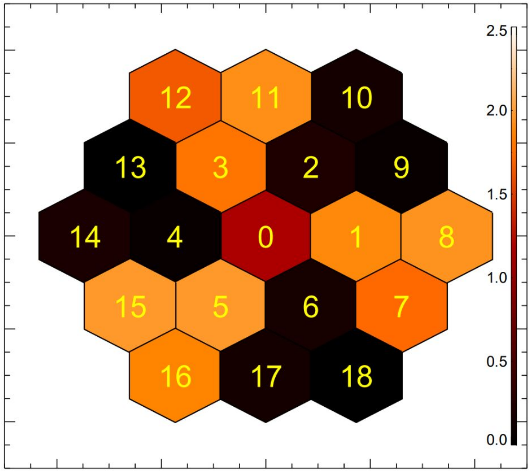
SPI - Coded Mask

- 720mm diameter
- 127 hexagonal tiles (63 opaque)
- 1.71m above detector plane
- 120° rotational symmetry
- 3cm thick tungsten alloy



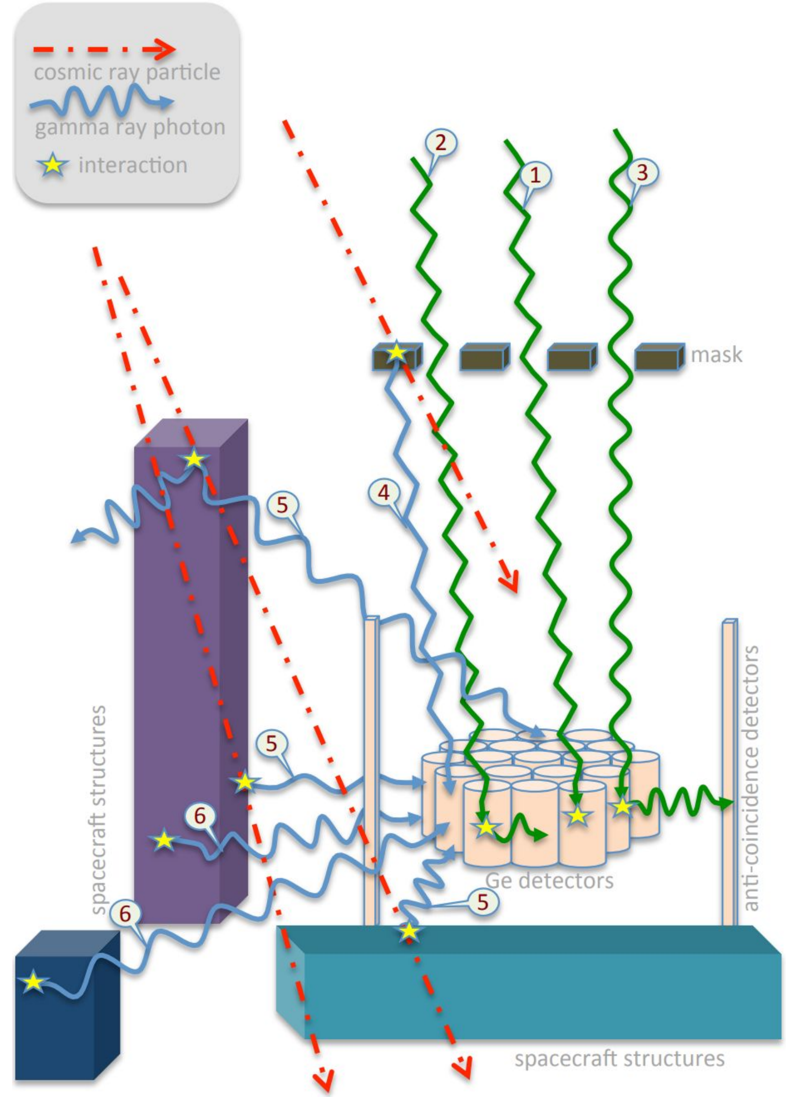






SPI - Events

- Event Types
  - Single Events (SE)
    - Only one GeD measured an energy deposit
  - Multiple Events (ME)
    - Multiple GeDs measured energy deposit simultaneously
  - Pulse Shape Discriminated Events (PE)
    - SE where PSD triggered as good shape
    - PSD can only trigger for SE
    - SE spectrum will have “missing” events that are PE
- Measured photons may or may not be fully absorbed
- ACS may or may not prevent “background” events
- Background counts dominant over all energy ranges for nearly all sources
  - Difficult data analysis
- Spurious Events
  - High-Energy deposits saturate analog electronics, displacing low-energy counts toward high energies
  - Does not occur for PE
  - SE only used for low energies, PE for everything above PSD threshold

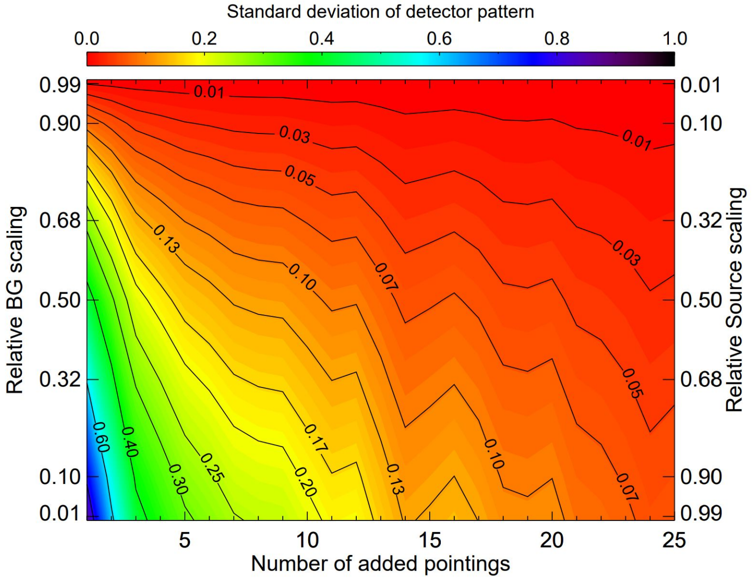


Astrophysics with SPI

- Stellar Evolution and Nucleosynthesis
  - Unstable isotopes created in stellar evolution produce unique energy decay lines
    - Measuring their distribution provides important insight
    - SPIs high energy resolution is excellent for this purpose
- Interstellar Medium
  - Measuring distribution of 511keV line can provides key insights into distribution of interstellar medium, as positrons from the decay of certain isotopes annihilate with electron of the gas
- Compact Objects
  - Crab Nebula (neutron star surrounded by plasma wind)
    - Spectral parameters
    - Temporal flux evolution
  - Measurements may or may not disprove theories on exact processes occurring
- Many more applications

Spimodfit

- SPI source analysis framework developed at MPE
- Uses detailed background and response database to deal with dominant background counts
- Fundamental idea: detector count rate patterns of individual isotopes stay constant on all timescales
- Intensities of different patterns vary, and are fitted to the data in the analysis
- Requires database of different detector patterns
  - Created by combining pointings together
  - Sources washed out, leaving only background patterns
- Used extensively for comparison purposes in this thesis



PySPI and the Goal of this Thesis

- Python analysis framework for GRB data from SPI developed at MPE
- Only transient source analysis is currently implemented
    - Circumenvents problem of dominant background due to time-sensitive nature of sources
- Goal of this thesis:
    - Implement method for fitting persistent sources using PySPI
    - Use profile likelihood to fit source models to data without requiring a background model

Persistent Sources with PySPI

1. Select relevant SCWs based on time and SPI orientation
2. Group nearby SCWs into clusters of predetermined size
   - Assume background count rates per energy bin per detector are constant within clusters
   - Temporal and angular distance of SCWs within Clusters should be minimized
3. Set-up source model to be fitted to the data
4. Set-up likelihood equation
   - Poisson distribution for counts of single bin $P(C|b,s,t) = \frac{(t\,(b+s))^C \exp(-t\,(b+s))}{C!}$
   - For cluster of size 2: $P(C_1, C_2|b, s_1, t_1, s_2, t_2) = P(C_1|b, s_1, t_1) \cdot P(C_2|b, s_2, t_2)$
   - Compute maximum-likelihood background by setting partial derivative to zero:

$$b_M = \frac{1}{2}\left[\frac{C_t}{t_t} - s_t + \sqrt{\left(\frac{C_t}{t_t} - s_t\right)^2 + 4\left(\frac{C_1 s_2 + C_2 s_1}{t_t} - s_1 s_2\right)}\right] \quad C_t = C_1 + C_2,\ s_t = s_1 + s_2,\ \text{and}\ t_t = t_1 + t_2$$

   - Allows likelihood computation depending only on source model:

$$\ln\mathcal{L}(C_1, C_2, b_M, t_1, t_2|s_1, s_2) = \ln P(C_1|b_M, s_1, t_1) + \ln P(C_2|b_M, s_2, t_2)$$

   - Sum for total likelihood: $\ln\mathcal{L}_{\text{total}} = \sum_{\text{Clusters}} \sum_{\text{Detectors}} \sum_{\text{Energy-Bins}} \ln\mathcal{L}(C_1, C_2, b_M, t_1, t_2|s_1, s_2)$

5. Run Bayesian fitting algorithm to obtain source parameter distributions

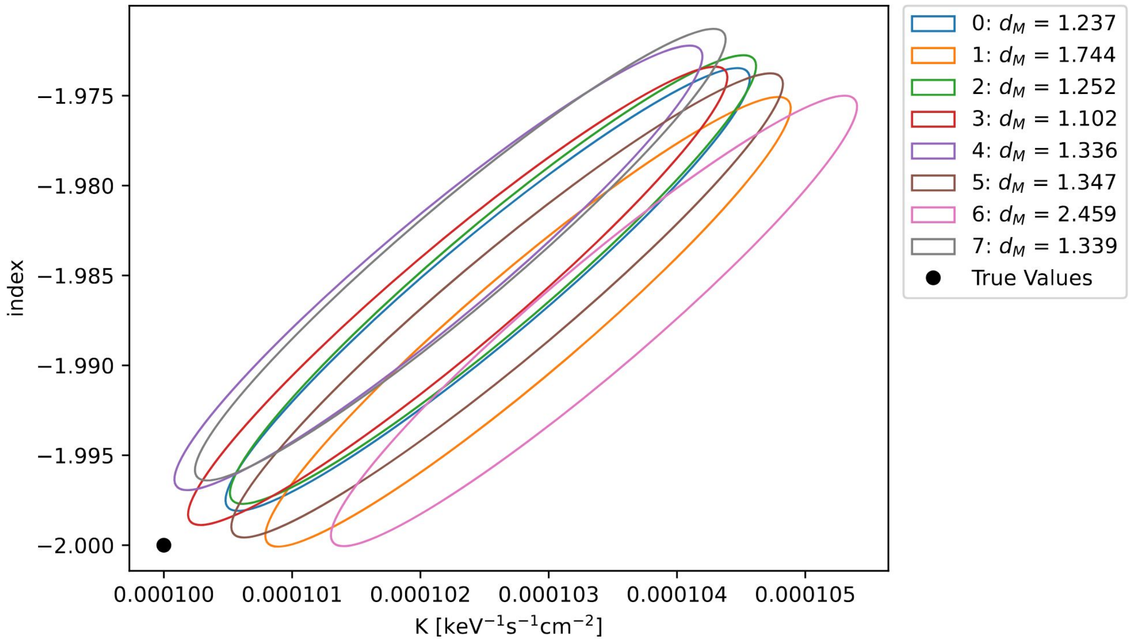
Purely Simulated Source - Consistency

- Test fitting results using purely simulated count data
  - Choose real SPI revolution as baseline
  - Define point source with powerlaw energy spectrum
  - Use IRF to calculate expected source count rates for different SCWs
  - Add (constant) background count rates to expected source count rates
  - Multiply by respective lifetime of SCW and take Poisson sample to obtain count data

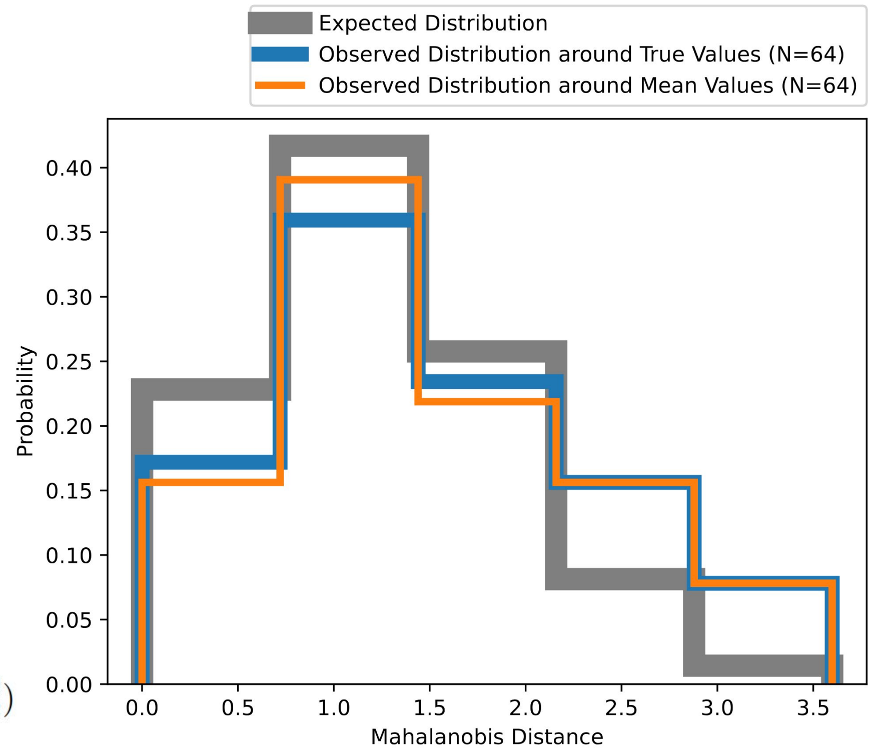
$$f(x) = K \frac{x}{piv}^{index}$$

- Mahalanobis Distance shows how many standard deviations the fit values are from the true values $d_M = \sqrt{(\vec{V}_F - \vec{V}_T)^T S^{-1} (\vec{V}_F - \vec{V}_T)}$

- Fit results shown for exactly identical datasets



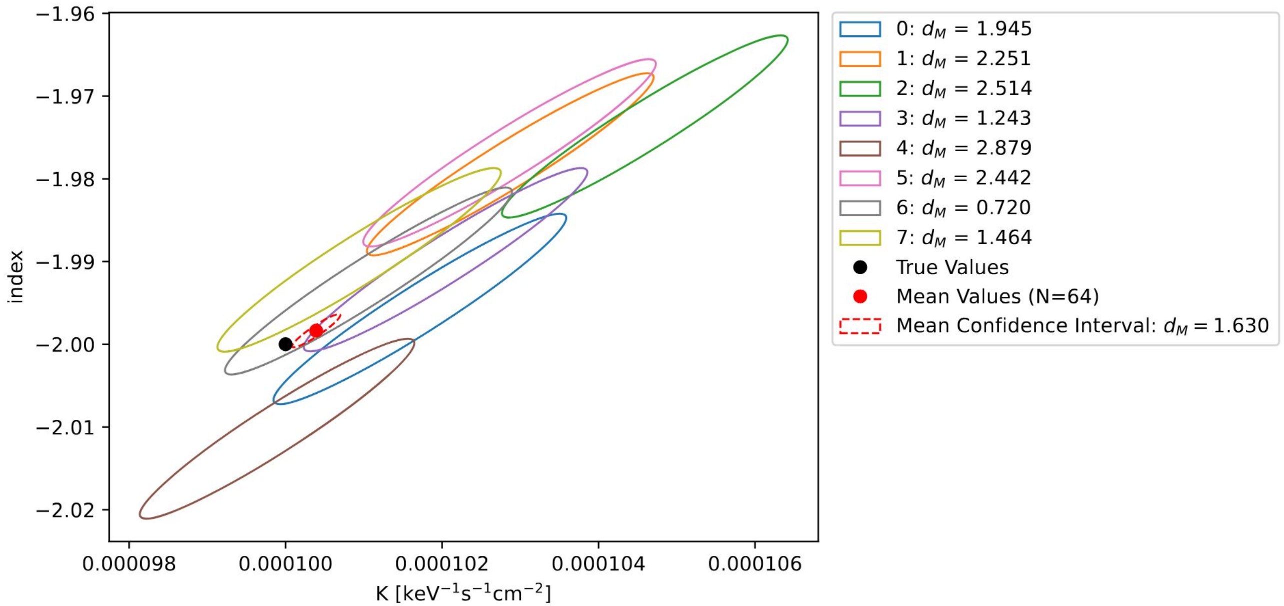
Purely Simulated Source - Accuracy



- Test accuracy by redrawing poisson samples used in generating data spectrum for consecutive fits
- Expected distribution of Mahalanobis Distance for two dimensional Gaussian:

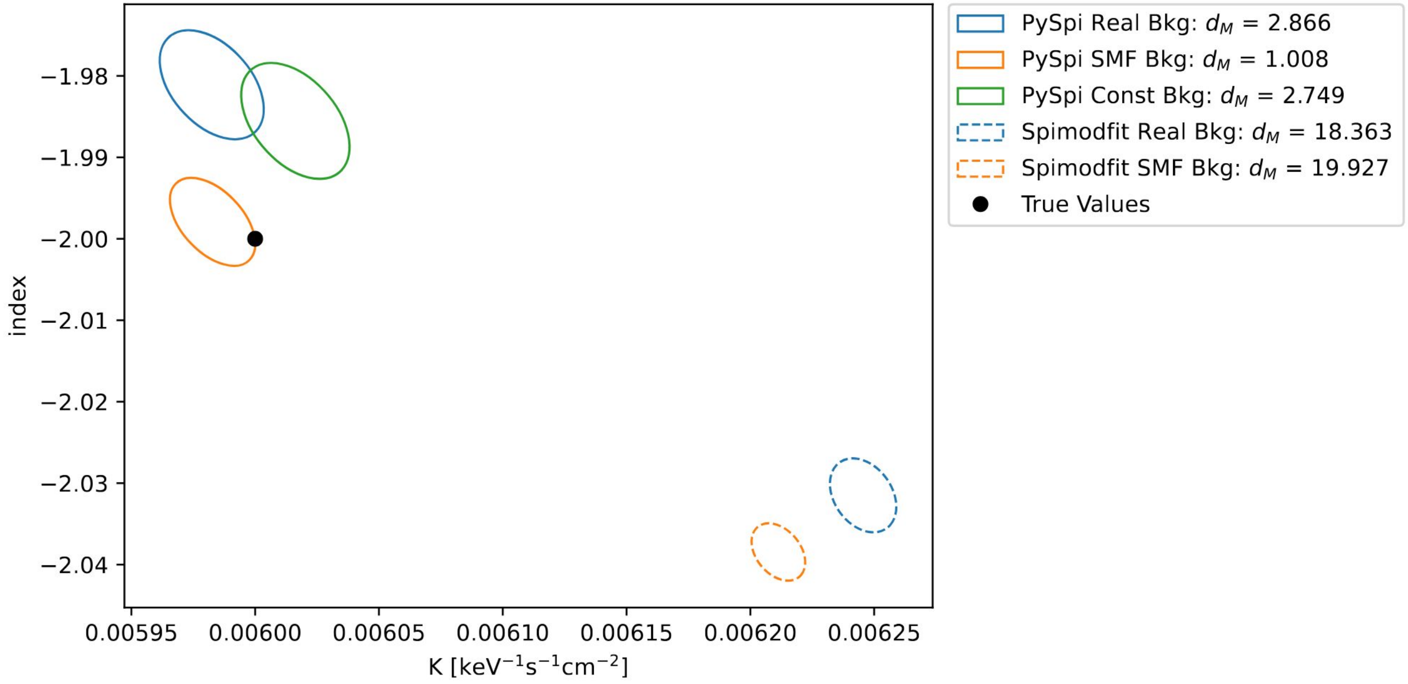
$$P(x_1 < d_M < x_2) = \frac{\int_{x_1}^{x_2} x\exp(-\frac{1}{2}x^2)\mathrm{d}x}{\int_0^{\inf} x\exp(-\frac{1}{2}x^2)\mathrm{d}x} = \exp(-\frac{1}{2}x_1^2) - \exp(-\frac{1}{2}x_2^2)$$

- No statistically significant systematic error
- Small random error



Mixed Simulation - Spimodfit Comparison

- Analogous to pure simulation, except non-constant background counts of real revolution are used alongside simulated source counts
- Important to choose revolution with no bright sources in SPI FoV
- PySPI consistently outperforms Spimodfit



Posterior Predictive Check (PPC) Analysis

- Check if fitted source model actually matches observed data
- Use to remove erroneous SCWs from fit
- Procedure
    a. Fit data to source model
    b. Randomly sample posterior
    c. Calculate expected count rates using IRF
    d. Using respective lifetimes, draw Poisson samples to obtain counts for each bin
    e. Repeat from step b. until posterior predictive distribution is sufficiently detailed
- Problem: background independent from posterior
    - Can only sample source count rates in step c.
- Solution: use maximum likelihood background from real counts
    - Problem: statistical correlation between maximum likelihood background and measured counts
    - Solution: approximate variance around expected values as normal, use numerical sampling to determine variance

PPC - PySPI Application

- Goal: compute distributions:
  - *b* is unknown

$$B_{1m} \sim \mathrm{Pois}(bt_1)$$
$$B_{2m} \sim \mathrm{Pois}(bt_2)$$
$$S_{1m} \sim \mathrm{Pois}(s_1t_1)$$
$$S_{2m} \sim \mathrm{Pois}(s_2t_2)$$

$$C_{1m} = B_{1m} + S_{1m}$$
$$C_{2m} = B_{2m} + S_{2m}$$

- Define quantities that measure difference between desired distributions and known expected values:

$$B_{d1} = B_{1m} - b_M t_1$$
$$B_{d2} = B_{2m} - b_M t_2$$
$$S_{d1} = S_{1m} - s_1 t_1$$
$$S_{d2} = S_{2m} - s_2 t_2.$$

- Desired distributions can be approximated as normal:

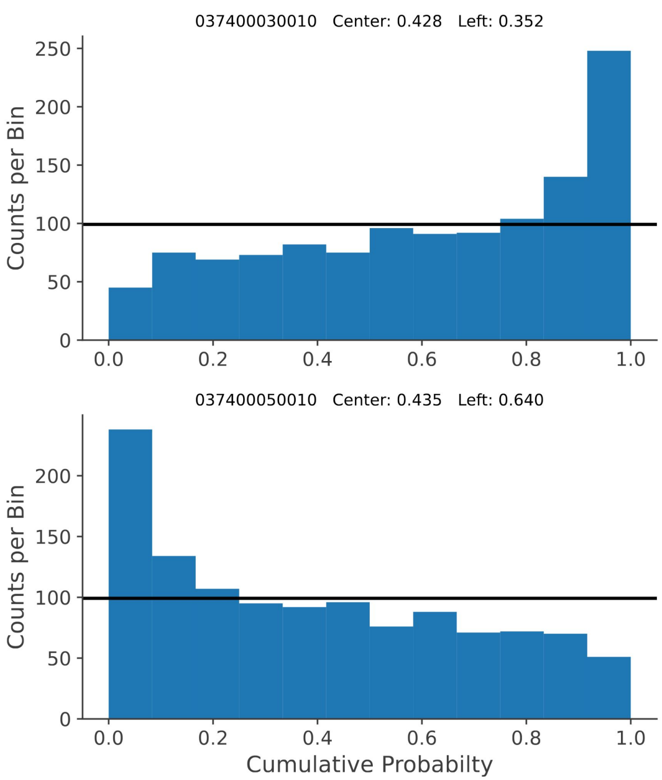
$$\begin{pmatrix} B_{1m} \\ S_{1m} \end{pmatrix} \sim \mathcal{N}(\mu_1, cov_1) \qquad \mu_1 = \begin{pmatrix} b_M t_1 \\ s_1 t_1 \end{pmatrix} \qquad cov_1 = \begin{bmatrix} \text{variance}(B_{d1}) & \text{covariance}(B_{d1}, S_{d1}) \\ \text{covariance}(B_{d1}, S_{d1}) & \text{variance}(S_{d1}) \end{bmatrix}$$

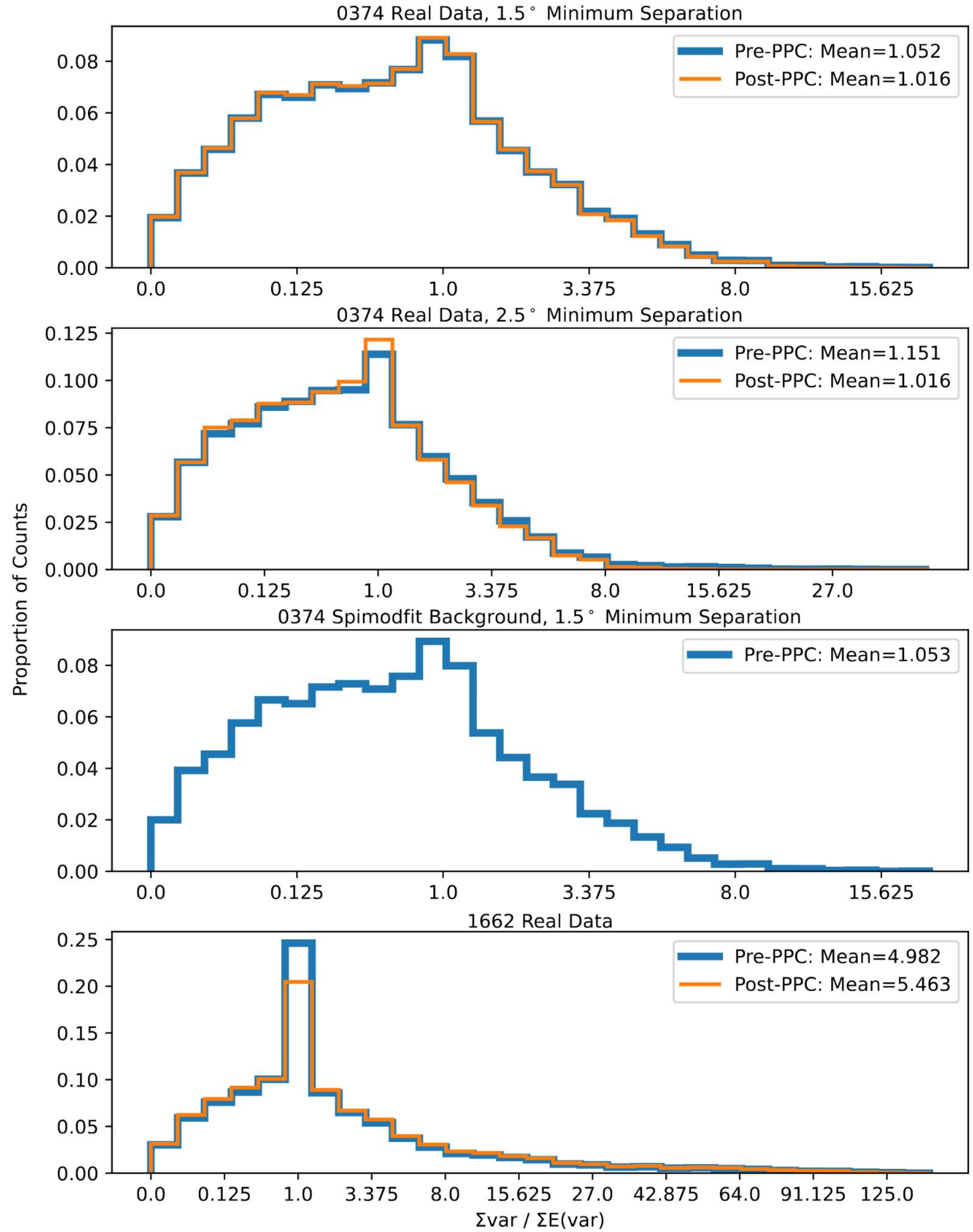
$$\begin{pmatrix} B_{2m} \\ S_{2m} \end{pmatrix} \sim \mathcal{N}(\mu_2, cov_2) \qquad \mu_2 = \begin{pmatrix} b_M t_2 \\ s_2 t_2 \end{pmatrix} \qquad cov_2 = \begin{bmatrix} \text{variance}(B_{d2}) & \text{covariance}(B_{d2}, S_{d2}) \\ \text{covariance}(B_{d2}, S_{d2}) & \text{variance}(S_{d2}) \end{bmatrix}$$

- Estimate covariances using numerical sampling at select values of ($b\ s_1\ s_2\ t_1\ t_2$)
  - Interpolate in between values that were numerically sampled

PPC - Results

- Works well
- Very effective at identifying clusters with non-constant background rates using CDF histograms
- Keep only “good” clusters and repeat fit





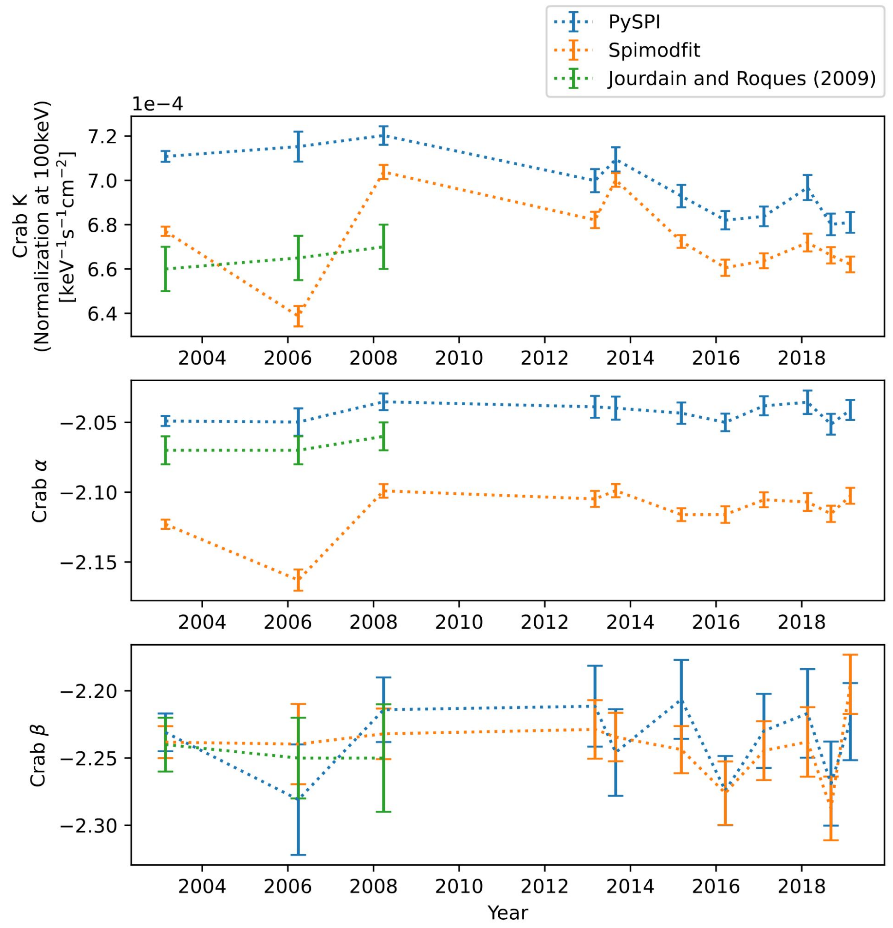
Crab Nebula

- Supernova remnant and pulsar wind nebula in the constellation of taurus
- Well known as one of the few bright sources in the hard x-ray domain with a persistent spectrum
    - Measurements sometimes used for calibration purposes
    - Energy spectrum can be approximated as Powerlaw with *index*=-2
- Be pulsar 1A0535+262 is 4.3° away from the Crab Nebula
    - Usually much dimmer, but can interfere with measurements during flaring periods

Crab Nebula - Comparison to other Analyses

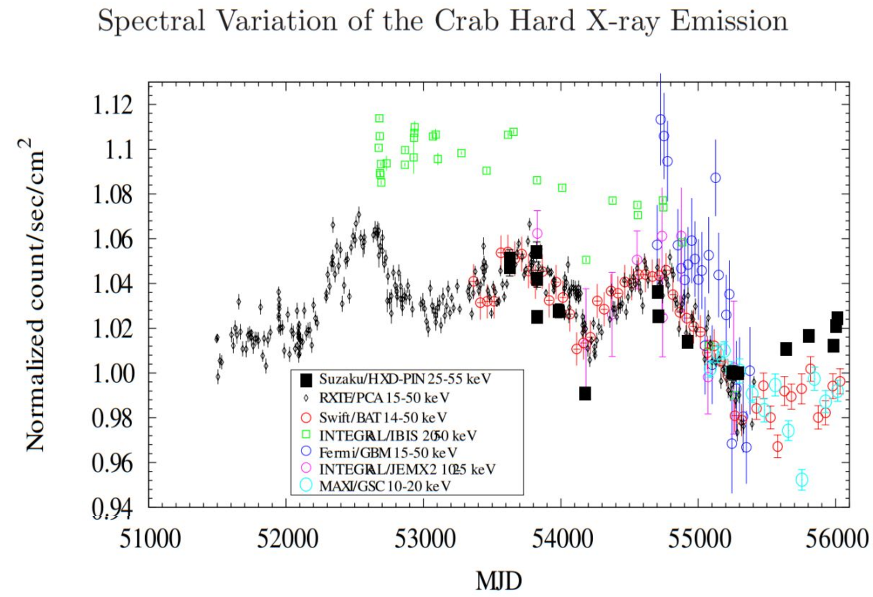
- Comparison of fit withbroken (100keV) Powerlaw as model using SPI data
- Comparison to other Instruments
  - Index not exactly comparable since different energy ranges used in fit
- PySPI measures higher normalization and lower index than most
- Similar temporal trends

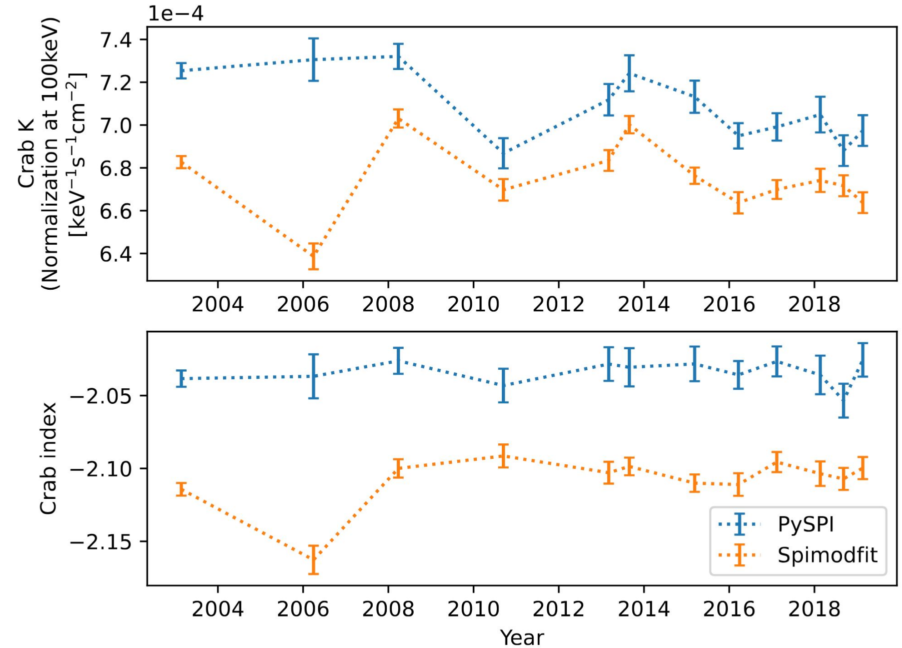
Instrument	Index	50-100keV flux $[s^{-1}cm^{-2}]$
OSO-8	-2.00 ± 0.06	$6.41 \cdot 10^{-2}$
GRIS	-2.15 ± 0.03	$4.52 \cdot 10^{-2}$
CGRO/OSSE	-2.19 ± 0.03	$5.68 \cdot 10^{-2}$
CGRO/BATSE	-2.20 ± 0.01	$6.83 \cdot 10^{-2}$
SAX/PDS	-2.13 ± 0.01	$4.92 \cdot 10^{-2}$
INTEGRAL/SPI - Sizun (2004)	-2.17 ± 0.01	$7.08 \cdot 10^{-2}$
INTEGRAL/SPI - PySPI	-2.08 ± 0.03	$7.26 \cdot 10^{-2}$
INTEGRAL/SPI - Spimodfit	-2.12 ± 0.03	$7.05 \cdot 10^{-2}$



Crab Nebula - Fading

- Small periodic variations in Crab flux are measured by many instruments
- One such “fading” occurred in 2010
- Flux decrease between 5 and 7.5% measured from March 2008 to September 2010, depending on instrument
  - PySPI - 6%
  - Spimodfit - 5%
- Measurements are consistent with previous results





Conclusion

- No statistically significant systematic error
- Slight random error
- Mixed simulations show consistent improvement over Spimodfit
  - Would be very impactful if true, questioning many results published in the last decade
- Slightly different parameters values measured for the Crab Nebula in comparison to other analyses
- Consistent measurement of Crab fading

Outlook

- Fitting only implemented for clusters up to size three
- PPC only implemented for clusters of size two
- Only implemented for SEs, severely limiting available energy range
- Further testing to validate improvement over Spimodfit required
- May be applied to observe many astronomical sources, potentially leading to many scientific discoveries