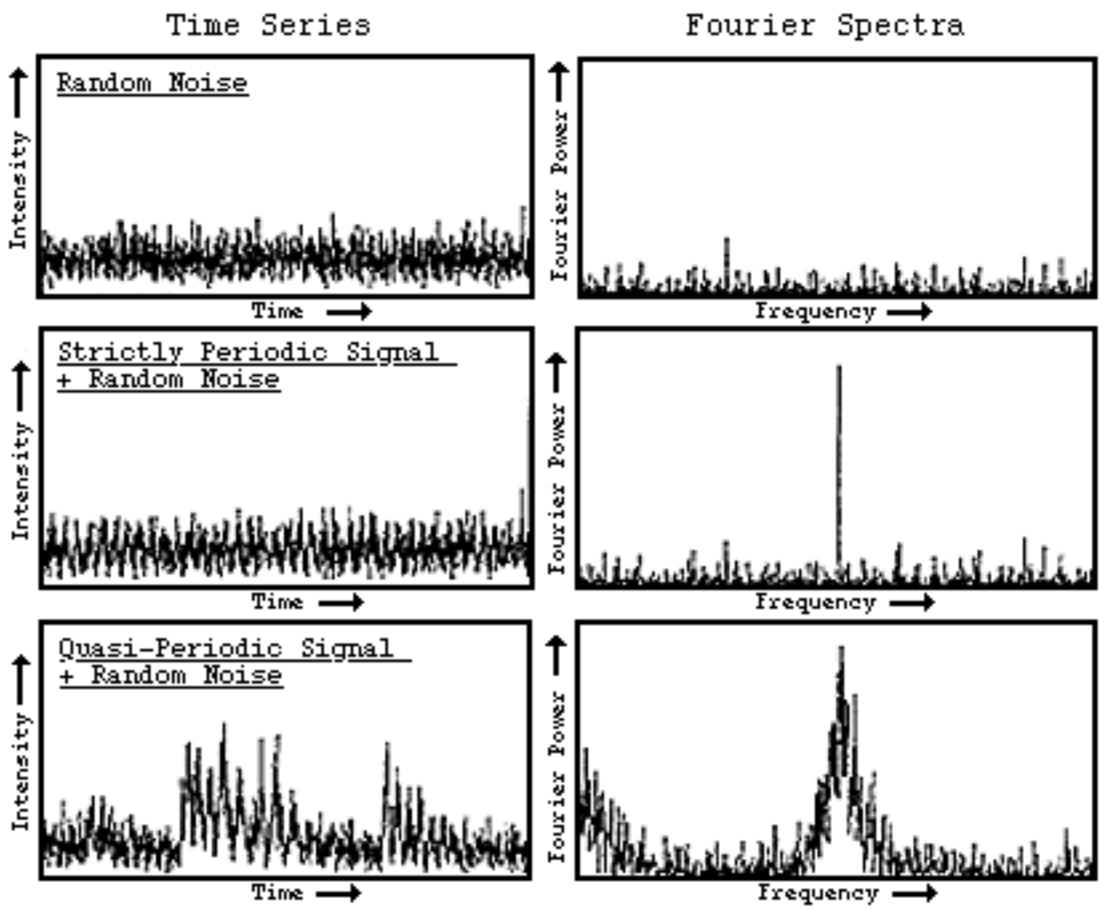
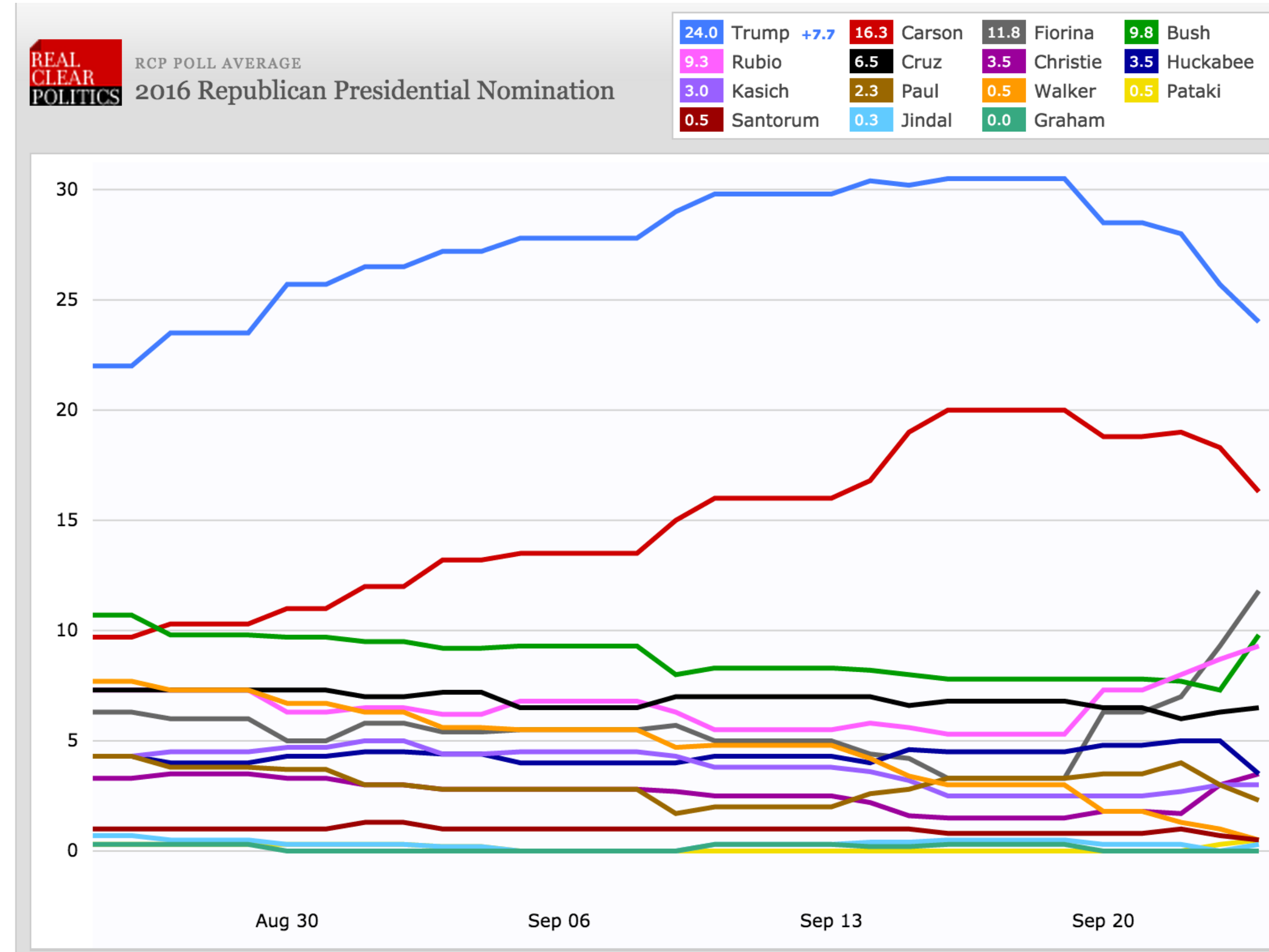




Astrophysics with Novel (Statistical) Observables

Nachiketa Chakraborty
MPIK, Heidelberg

Alexander von Humboldt Stiftung/Foundation

TeVPA, 8th August, 2017
Columbus, Ohio, USA

Motivation - Individual Source Physics

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)

- Standard SED modeling, morphology, "eyeballing" lightcurves insufficient - extract more from MWL LCs ?

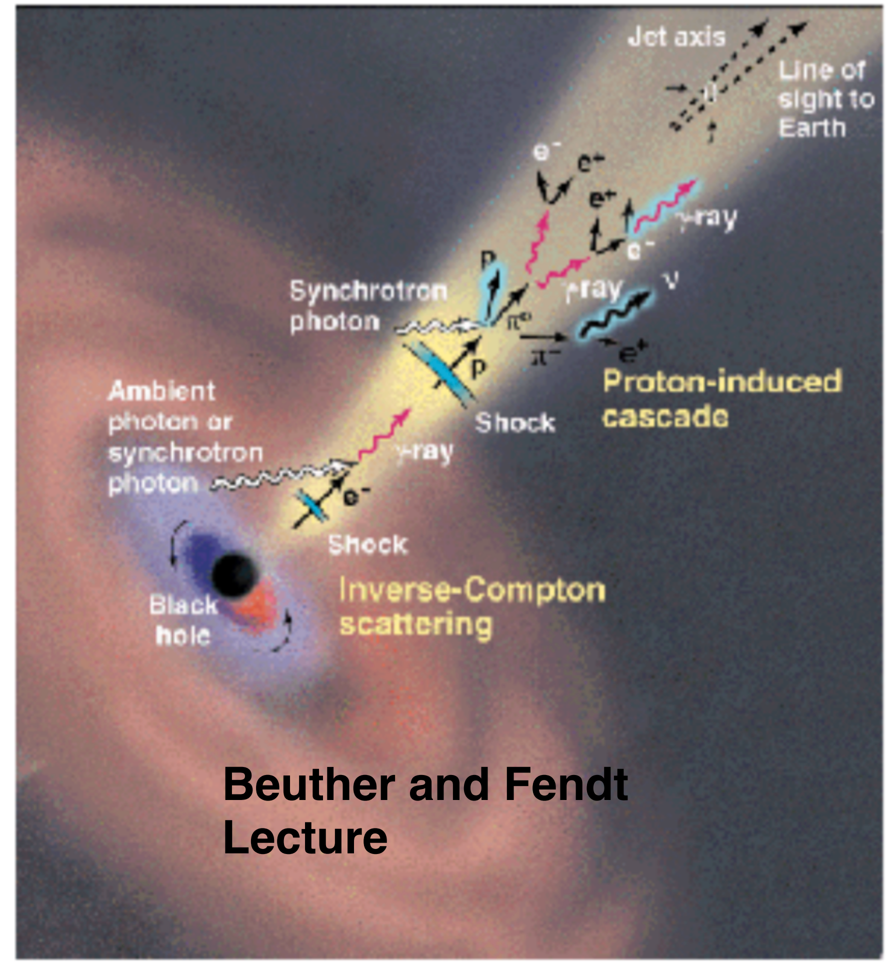
- Need **newer and novel "observables"** for sharper understanding -> PSD, PDF, Polarisation

- Large datasets <=> Statistical Methods (both individual and population) e.g. time series methods

- Better statistics per obs, more sources

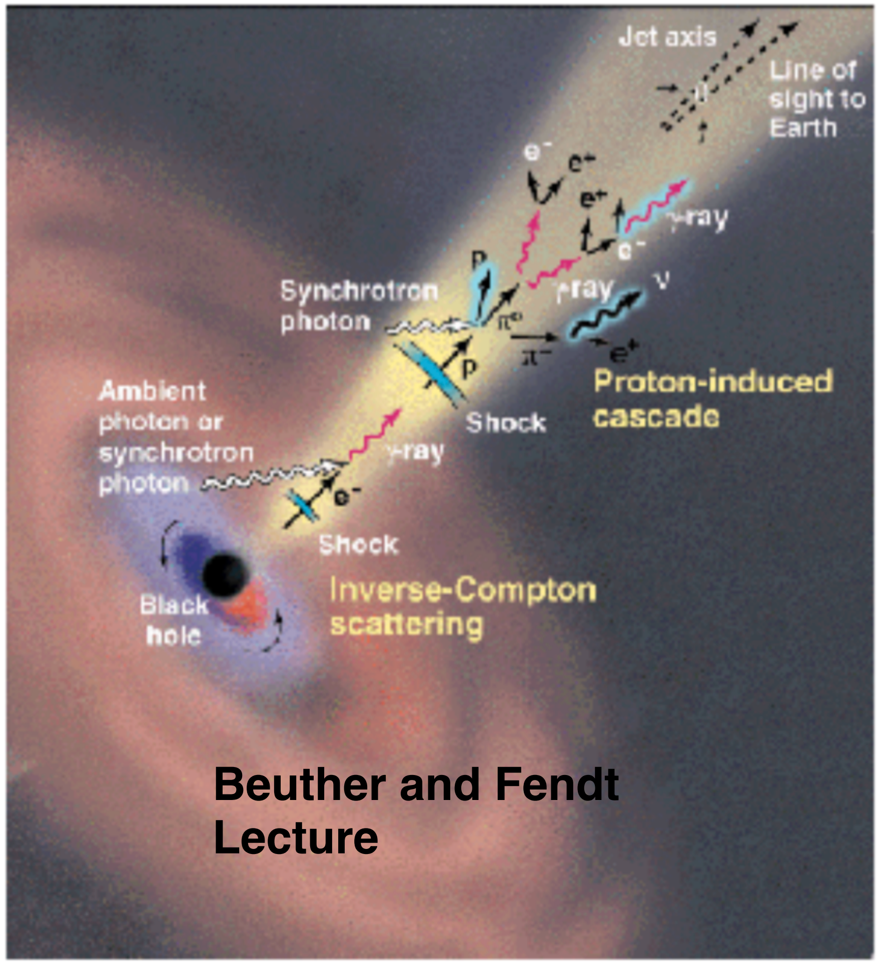
Motivation - Individual Source Physics

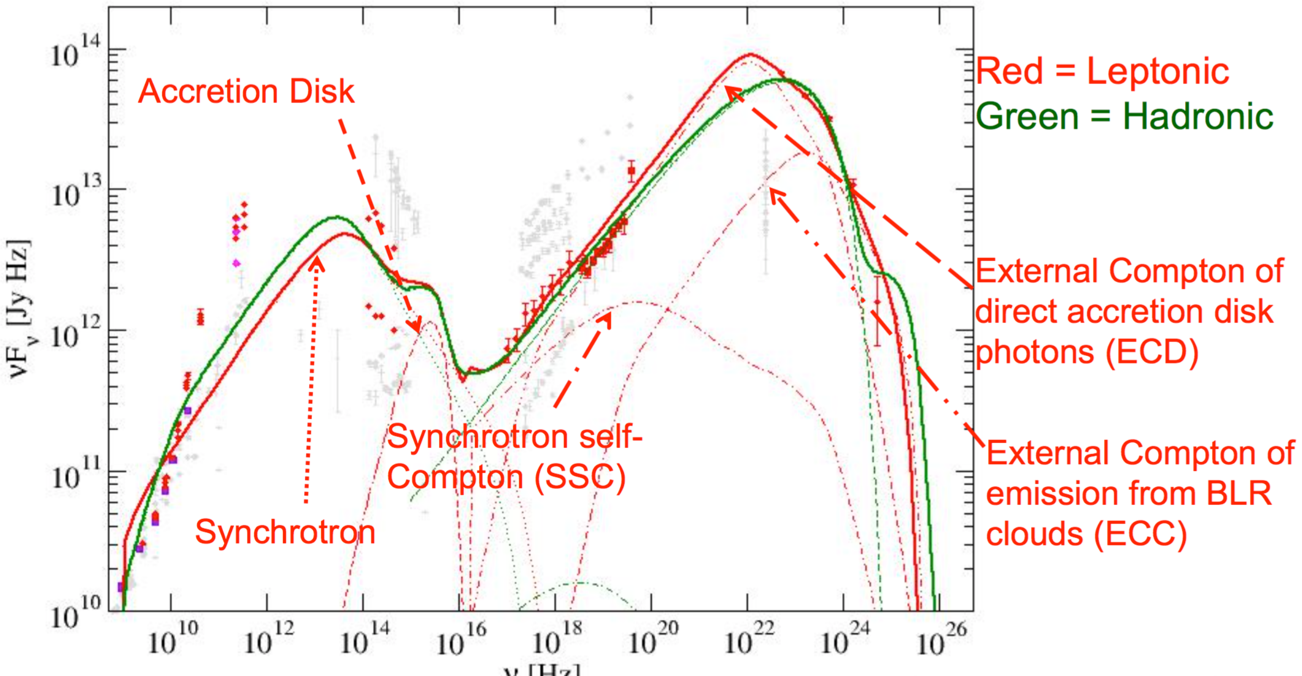
- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)
- Standard SED modeling, morphology, “eyeballing” lightcurves insufficient - extract more from MWL LCs ?
- Need **newer and novel “observables”** for sharper understanding -> PSD, PDF, Polarisation
- Large datasets <=> Statistical Methods (both individual and population) e.g. time series methods
- Better statistics per obs, more sources



Motivation - Individual Source Physics

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)
- Standard SED modeling, morphology, "eyeballing" lightcurves insufficient - extract more from MWL LCs ?
- Need **newer and novel "observables"** for sharper understanding -> PSD, PDF, Polarisation
- Large datasets <=> Statistical Methods (both individual and population) e.g. time series methods
- Better statistics per obs, more sources

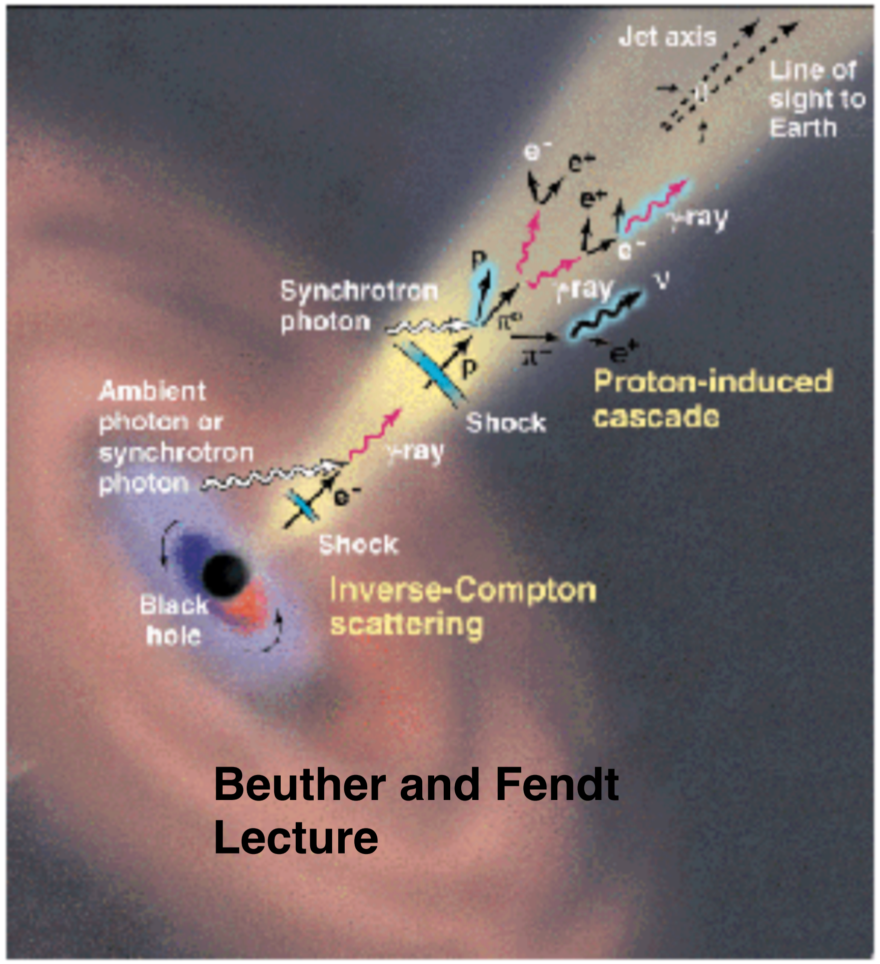


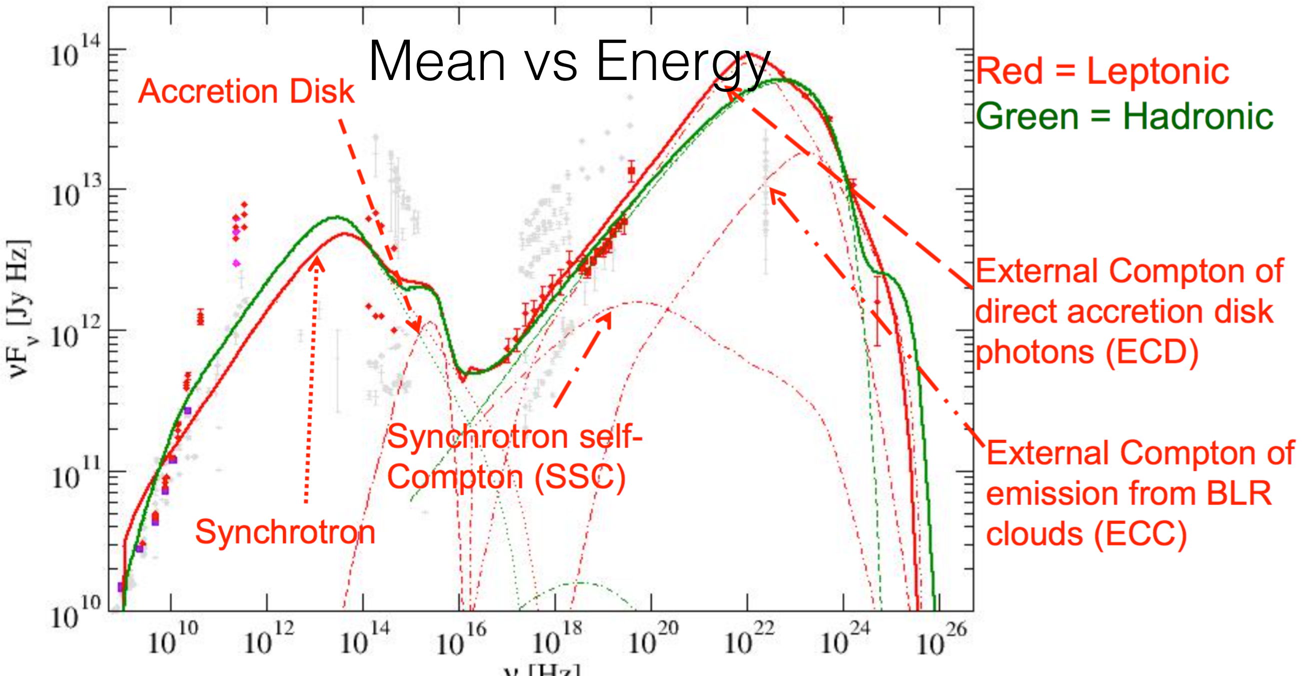


(Böttcher et al. 2013)

Motivation - Individual Source Physics

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)
- Standard SED modeling, morphology, “eyeballing” lightcurves insufficient - extract more from MWL LCs ?
- Need **newer and novel “observables”** for sharper understanding -> PSD, PDF, Polarisation
- Large datasets <=> Statistical Methods (both individual and population) e.g. time series methods
- Better statistics per obs, more sources

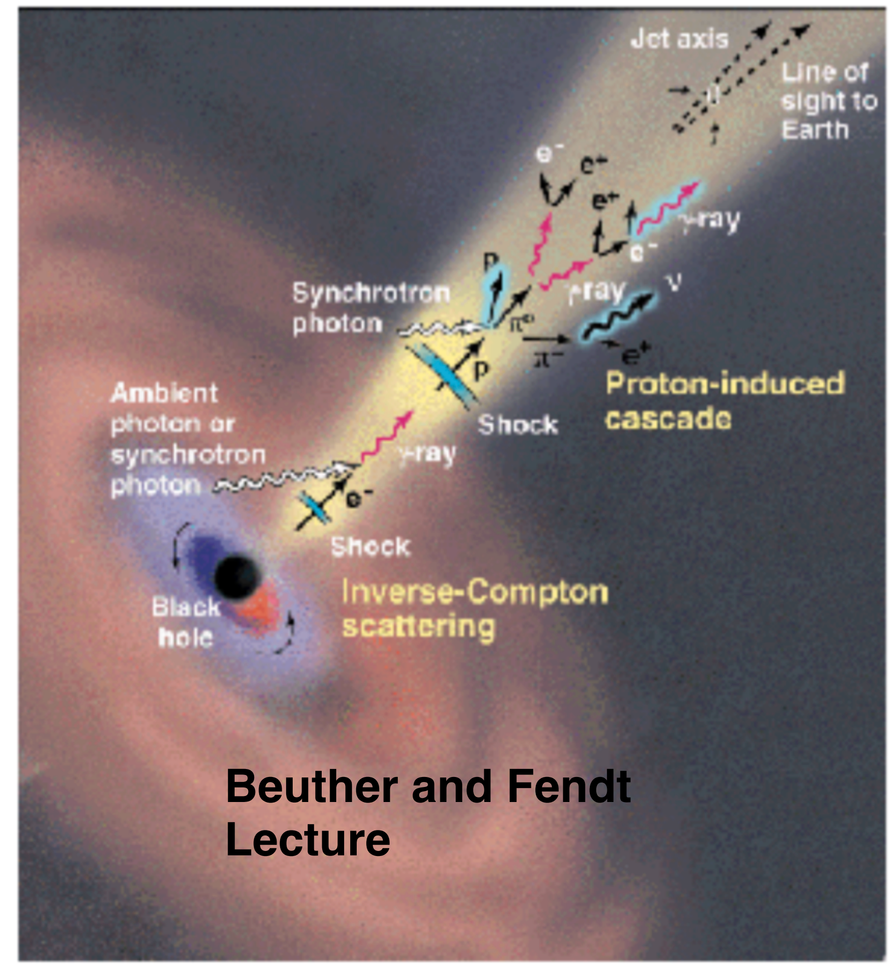


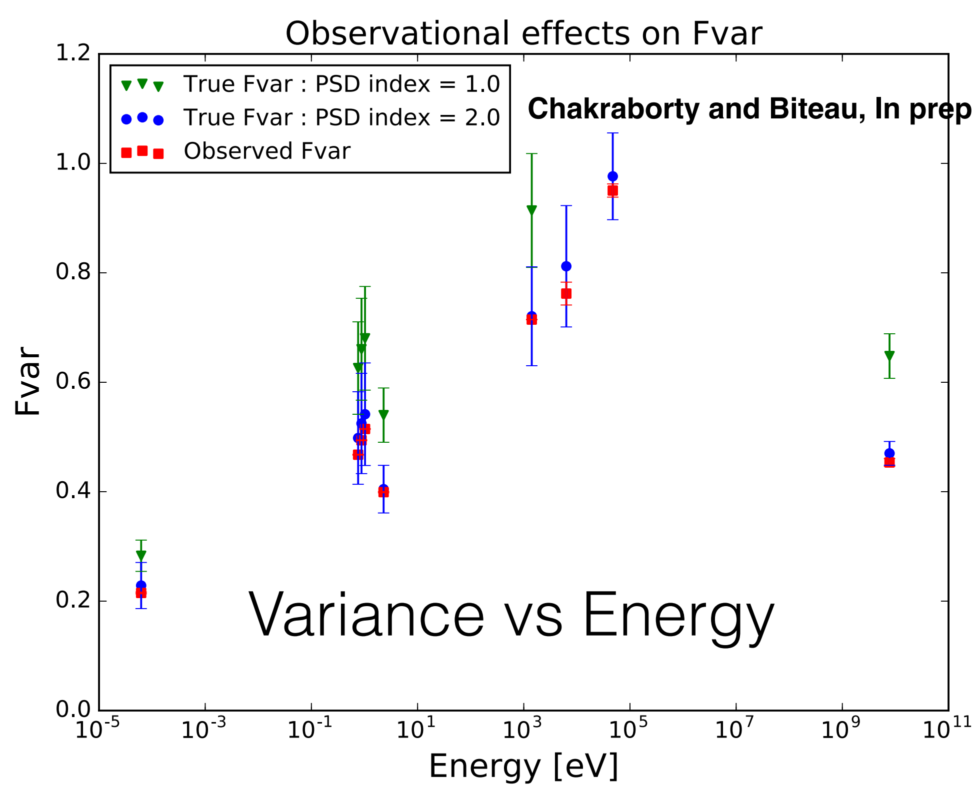


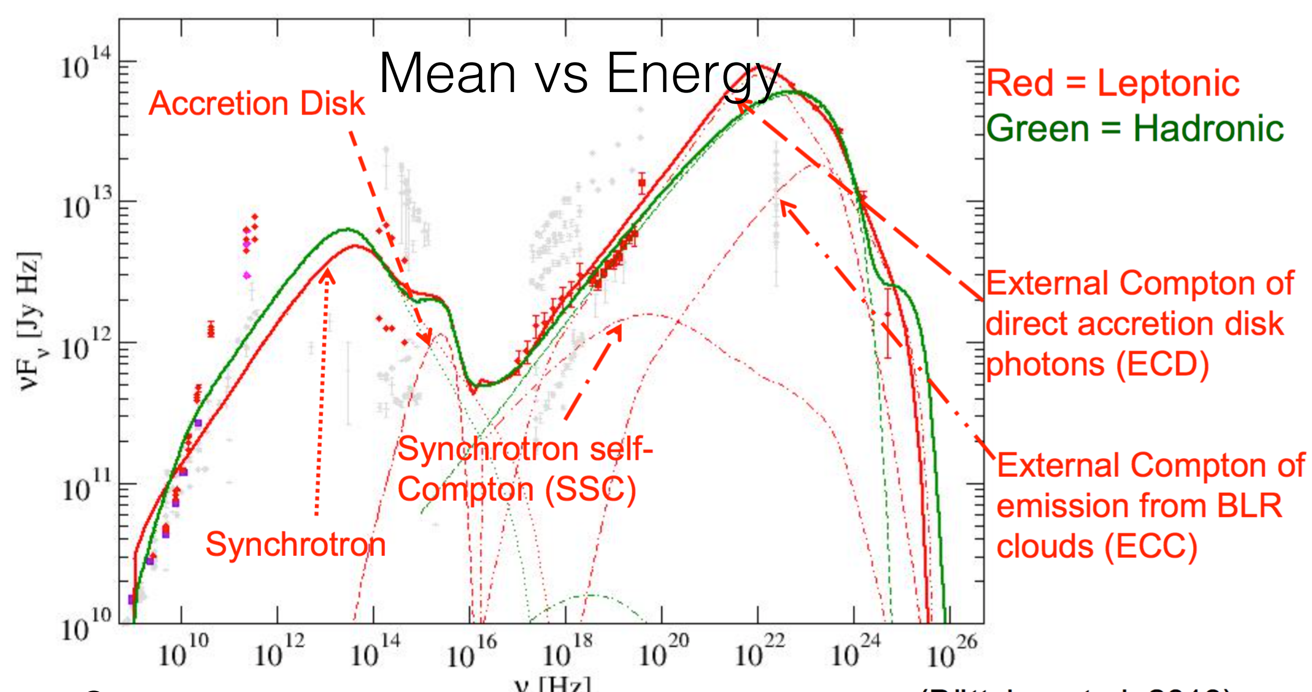
(Böttcher et al. 2013)

Motivation - Individual Source Physics

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)
- Standard SED modeling, morphology, “eyeballing” lightcurves insufficient - extract more from MWL LCs ?
- Need **newer and novel “observables”** for sharper understanding -> PSD, PDF,



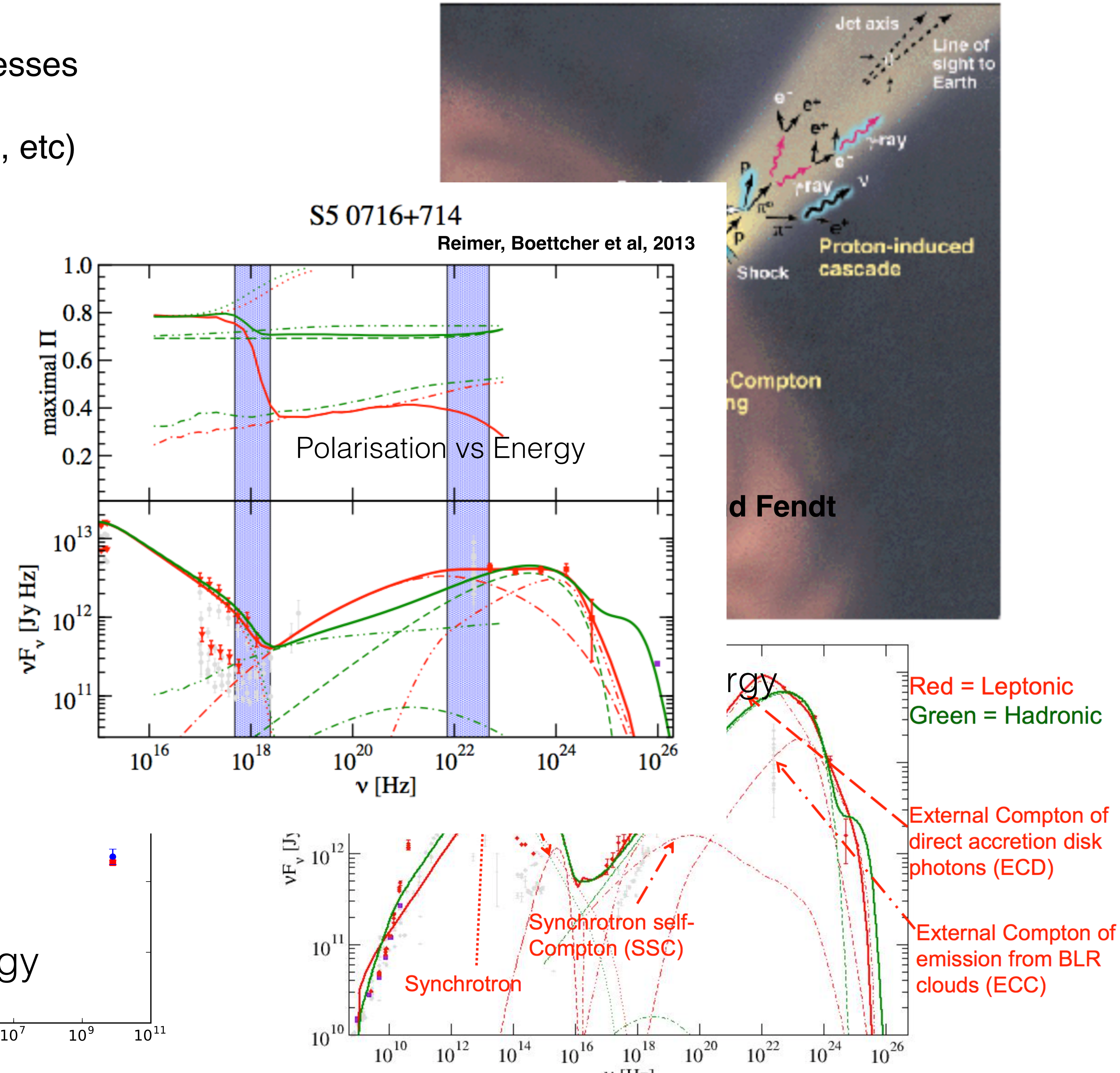


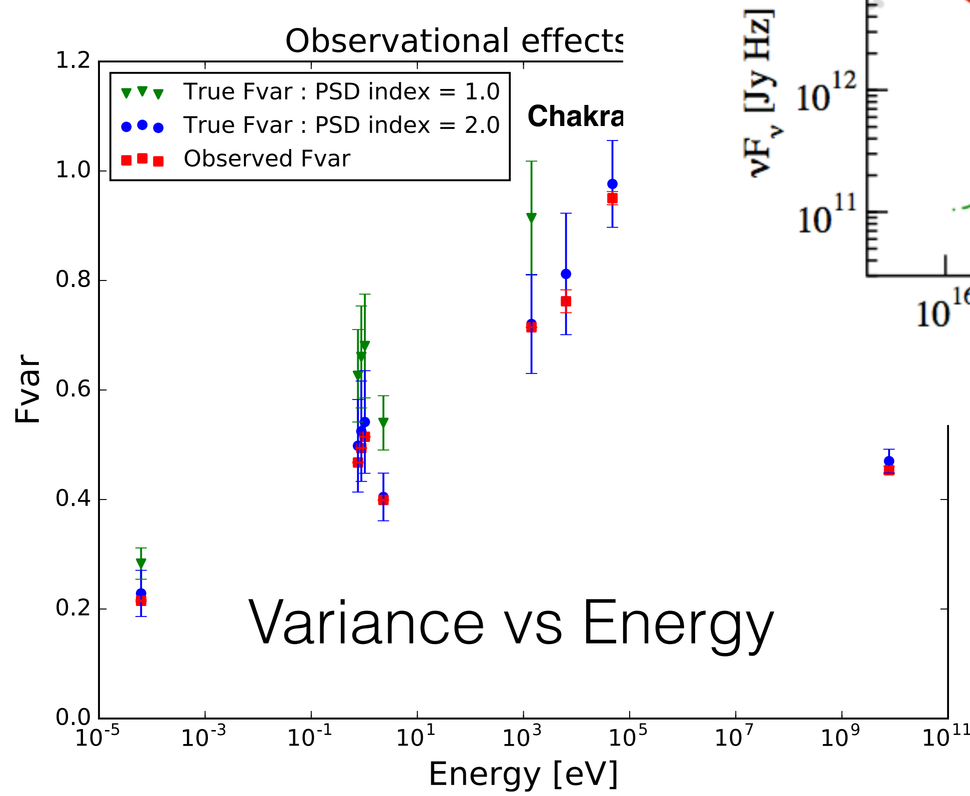


(Böttcher et al. 2013)

Motivation - Individual Source Physics

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)
- Standard SED modeling morphology, “eyeballing insufficient - extract more LCs ?
- Need **newer and novel “observables”** for sharp understanding -> PSD, F





Motivation - Population / Diffuse Background

- Complexity of physical processes and environment lead to degeneracies (AGNs, GRBs, etc)

- Standard SED modeling, morphology, "eyeballing" lightcurves insufficient - extract more from MWL LCs ?

- Need **newer and novel "observables"** for sharper understanding -> PSD, PDF, Polarisation

- Large datasets <=> Statistical Methods (both individual and population) e.g. time series methods

- Better statistics per obs, more sources

Motivation - Population / Diffuse Background

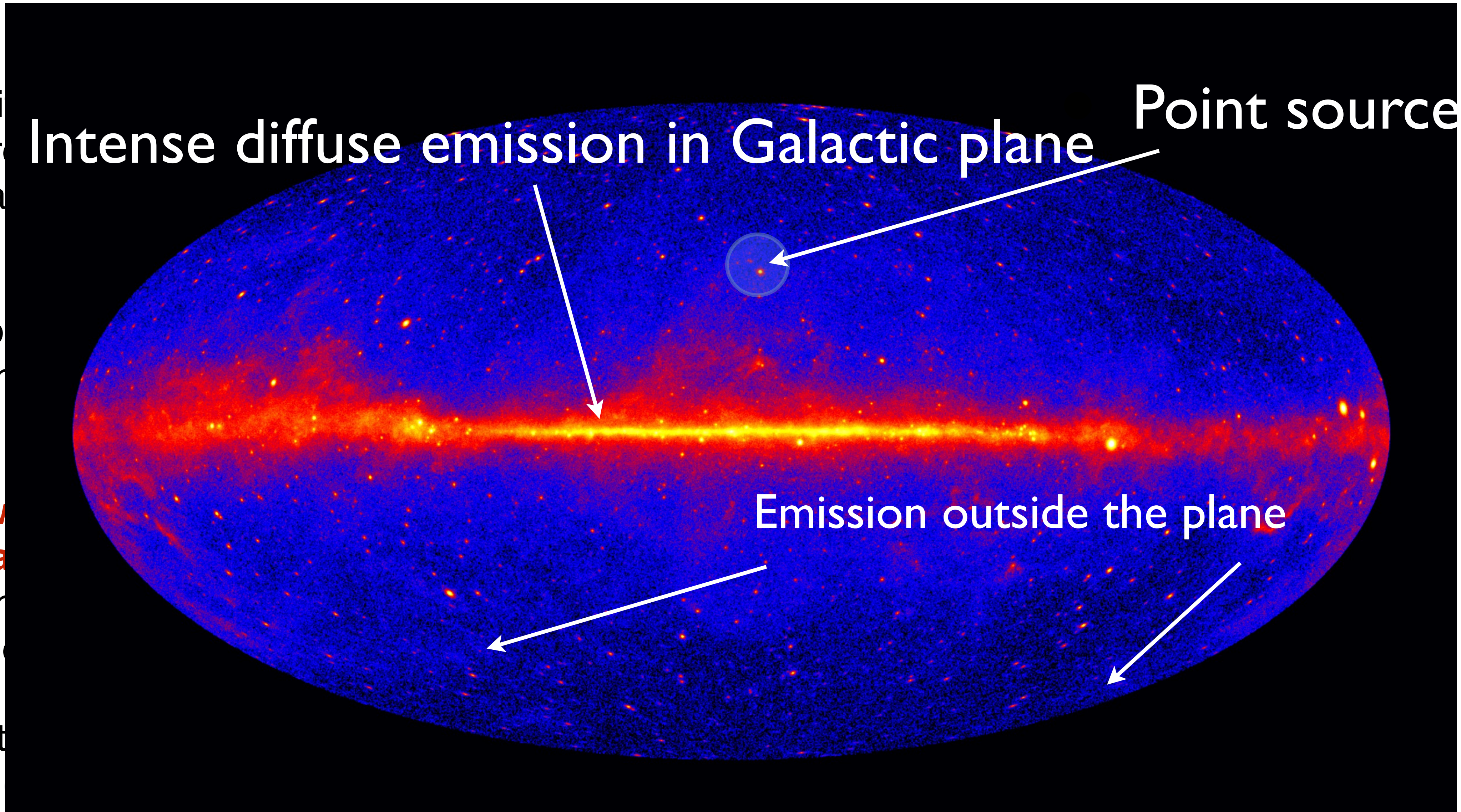
- Complexi
  and envir
  degenera

- Standard
  morpholo
  insufficien
  LCs ?

- Need **nev**
  “observa
  understar
  Polarisati

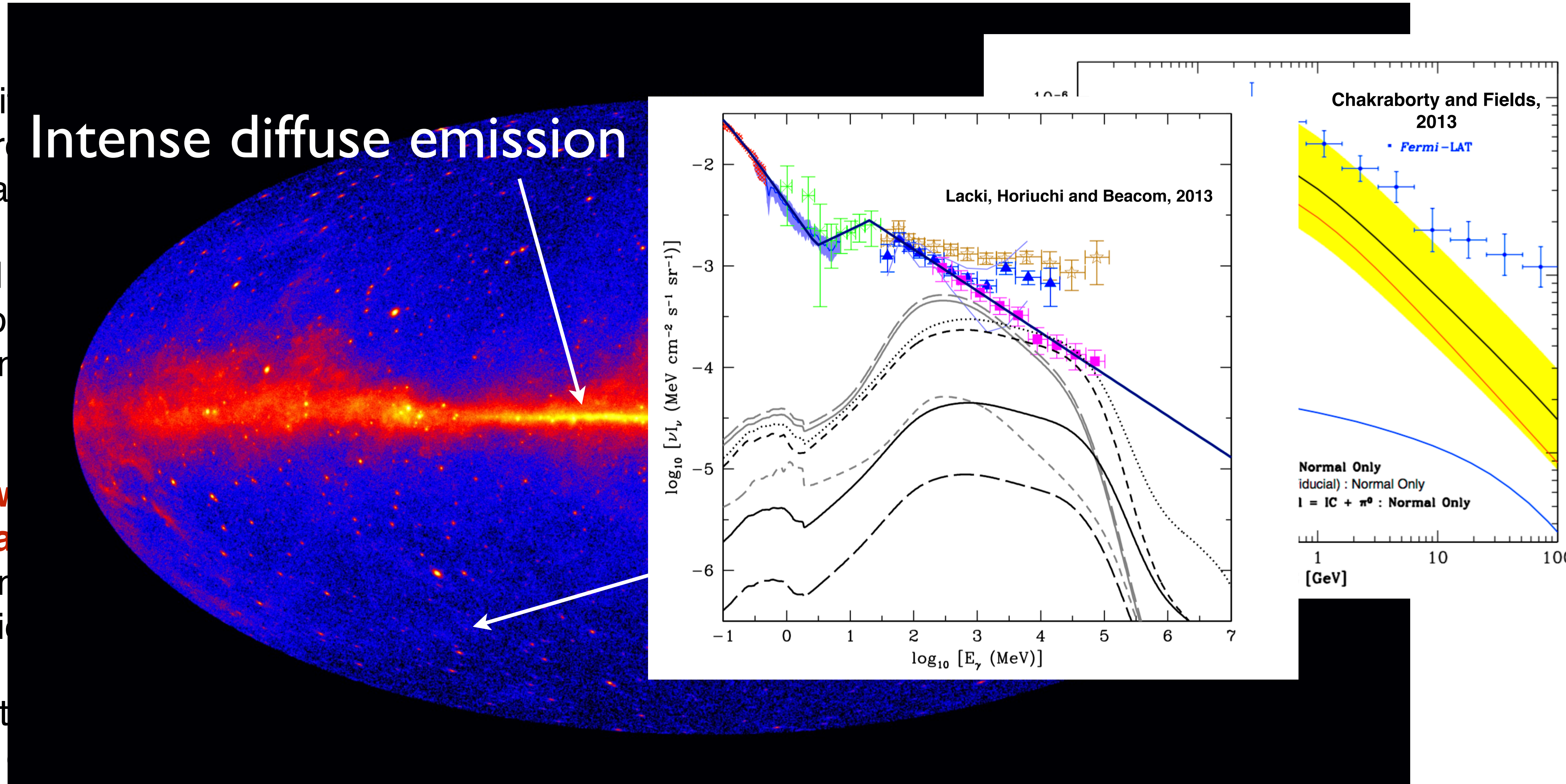
- Large dat
  Methods
  population) e.g. time series methods

- Better statistics per obs, more sources



Motivation - Population / Diffuse Background

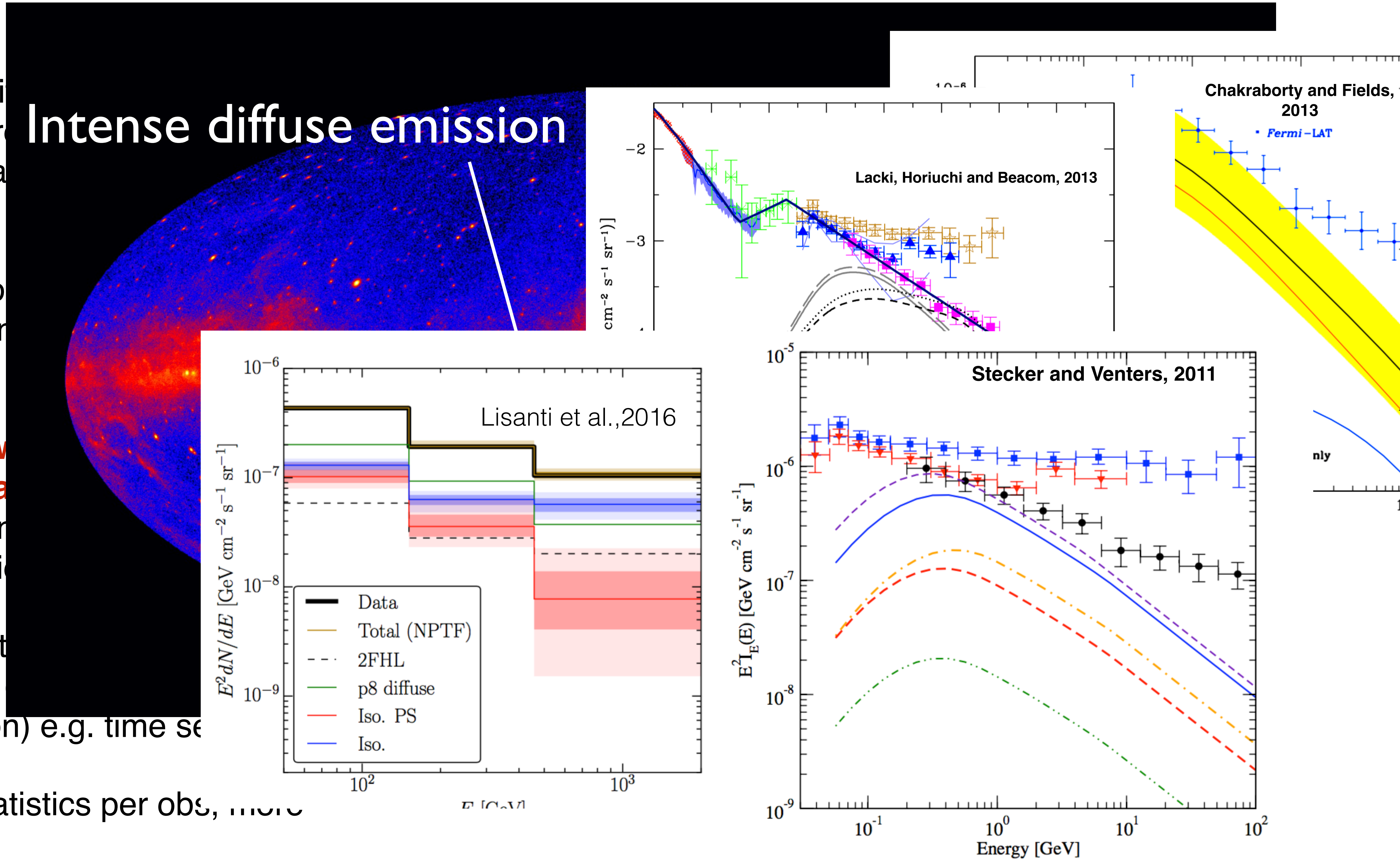
- Complexi
  and envir
  degenera
- Standard
  morpholo
  insufficier
  LCs ?
- Need **new**
  “observa
  understar
  Polarisati



- Large dat
  Methods
  population) e.g. time series methods
- Better statistics per obs, more sources

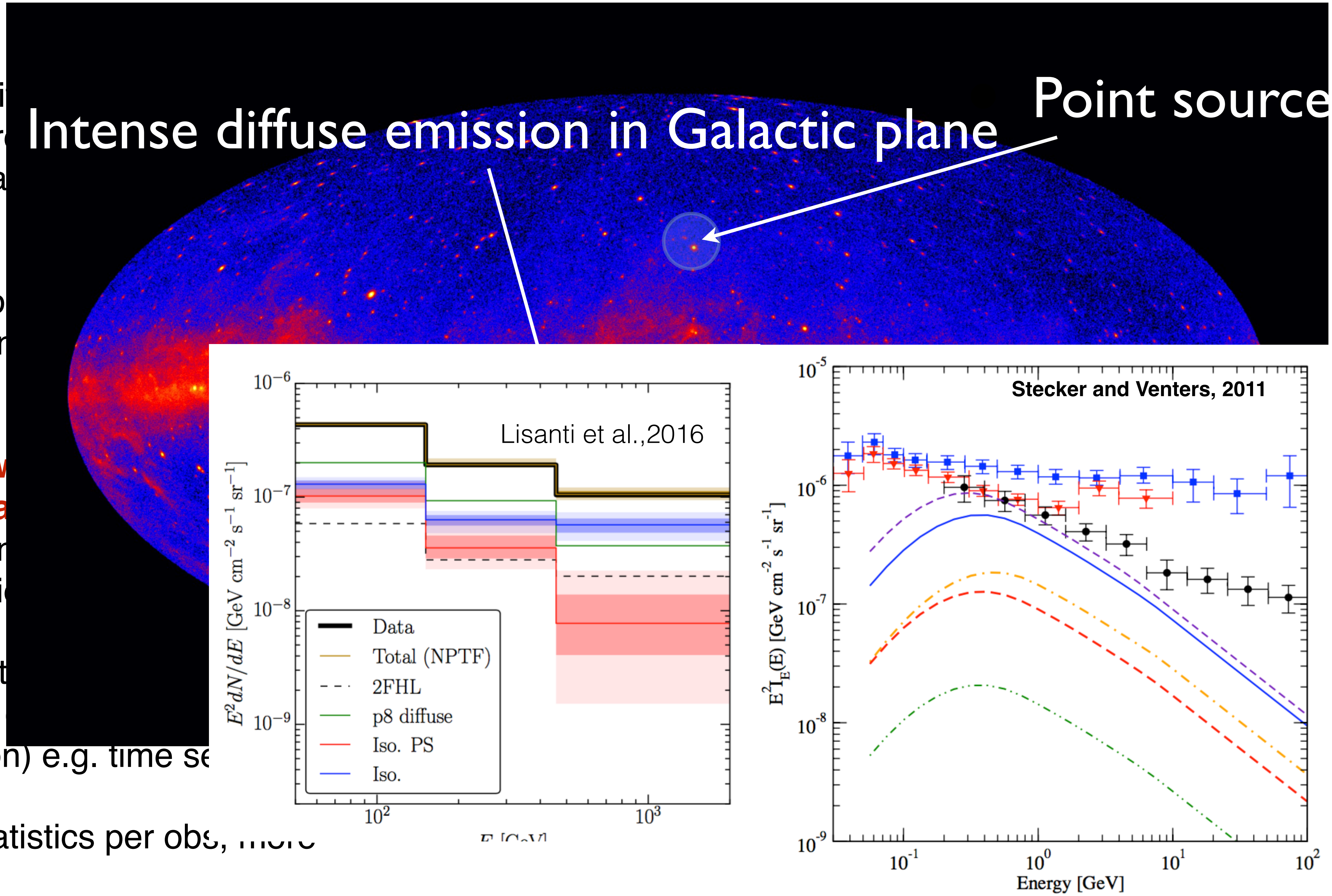
Motivation - Population / Diffuse Background

- Complexi
  and envir
  degenera
- Standard
  morpholo
  insufficien
  LCs ?
- Need **new**
  “observa
  understan
  Polarisati
- Large dat
  Methods
  population) e.g. time se
- Better statistics per obs,
  sources



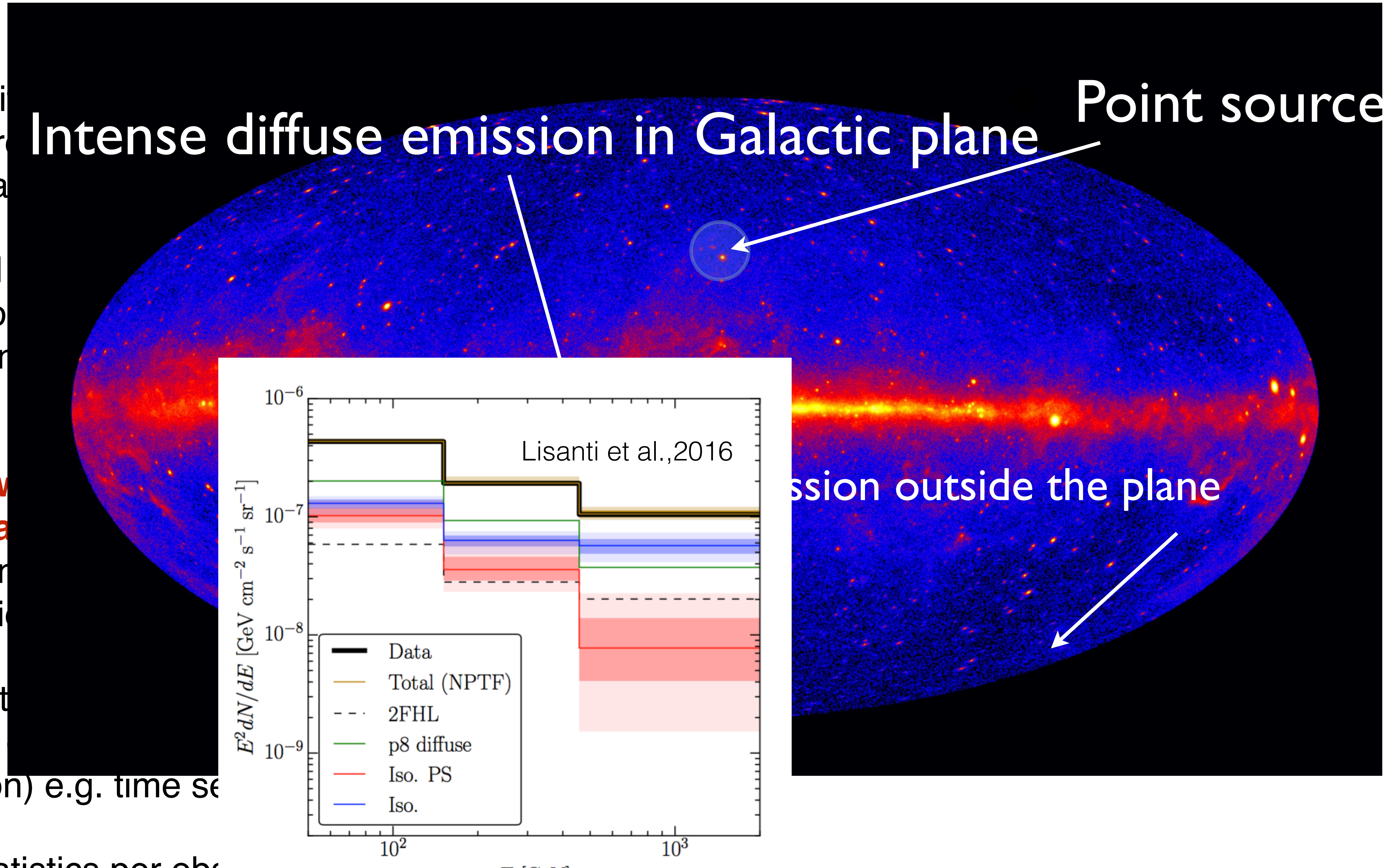
Motivation - Population / Diffuse Background

- Complexi
  and envir
  degenera
- Standard
  morpholo
  insufficien
  LCs ?
- Need **new**
  “observa
  understar
  Polarisati
- Large dat
  Methods
  population) e.g. time se
- Better statistics per obs,
  sources



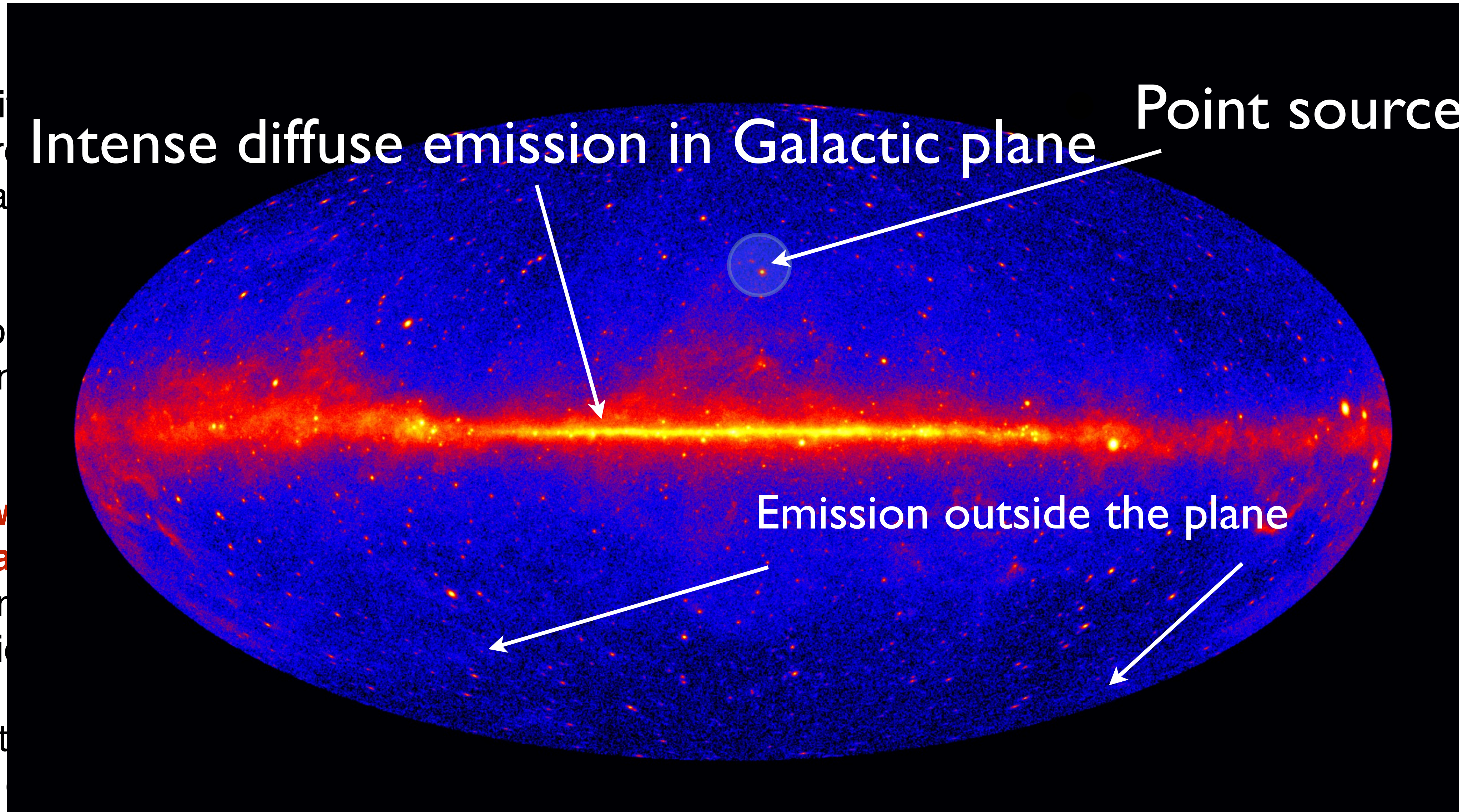
Motivation - Population / Diffuse Background

- Complexi
  and envir
  degenera
- Standard
  morpholo
  insufficier
  LCs ?
- Need **new**
  “observa
  understar
  Polarisati
- Large dat
  Methods
  population) e.g. time se
- Better statistics per obs, more
  sources



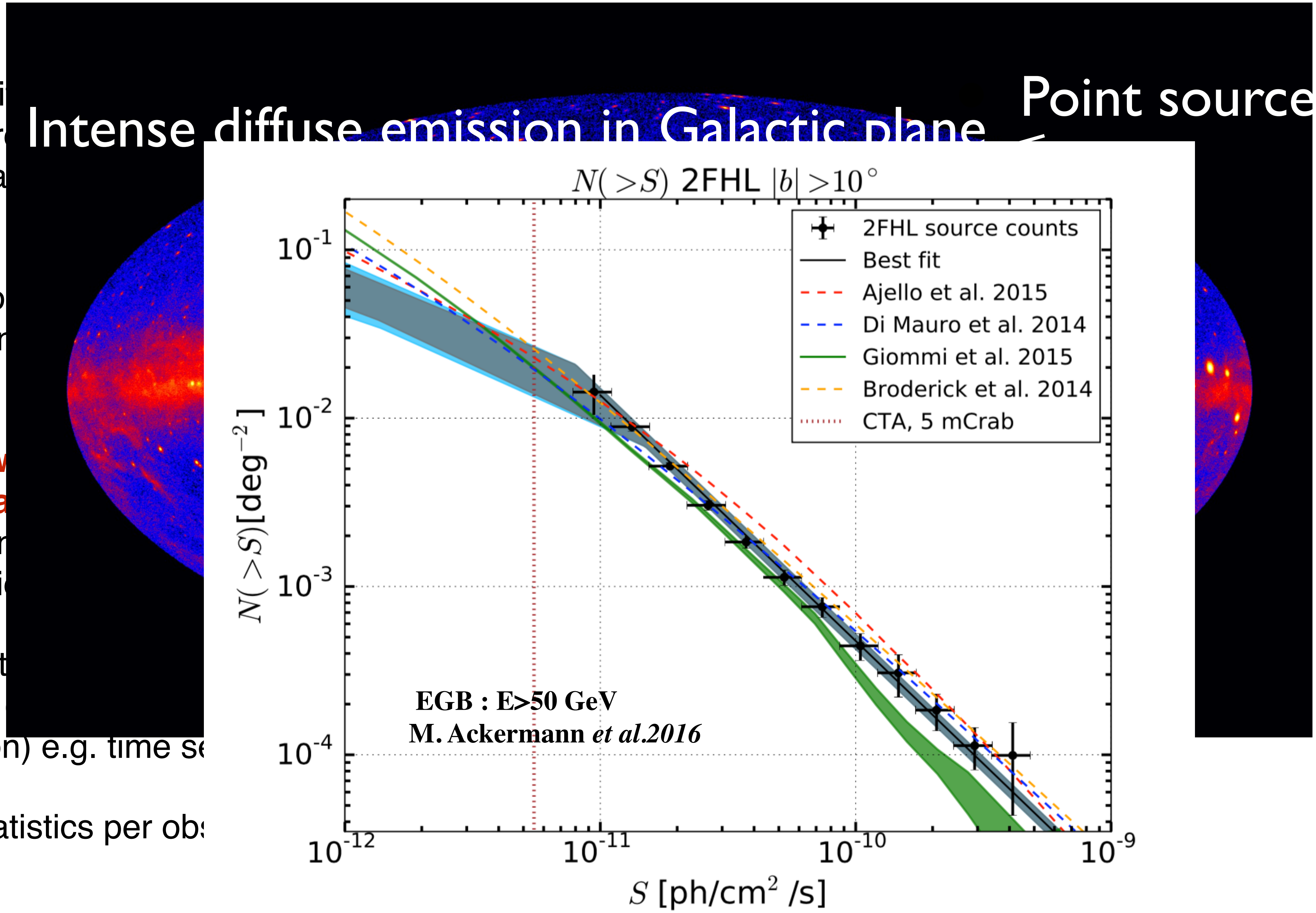
Motivation - Population / Diffuse Background

- Complexi
  and envir
  degenera
- Standard
  morpholo
  insufficier
  LCs ?
- Need **new**
  “observa
  understar
  Polarisati
- Large dat
  Methods
  population) e.g. time series methods
- Better statistics per obs, more sources

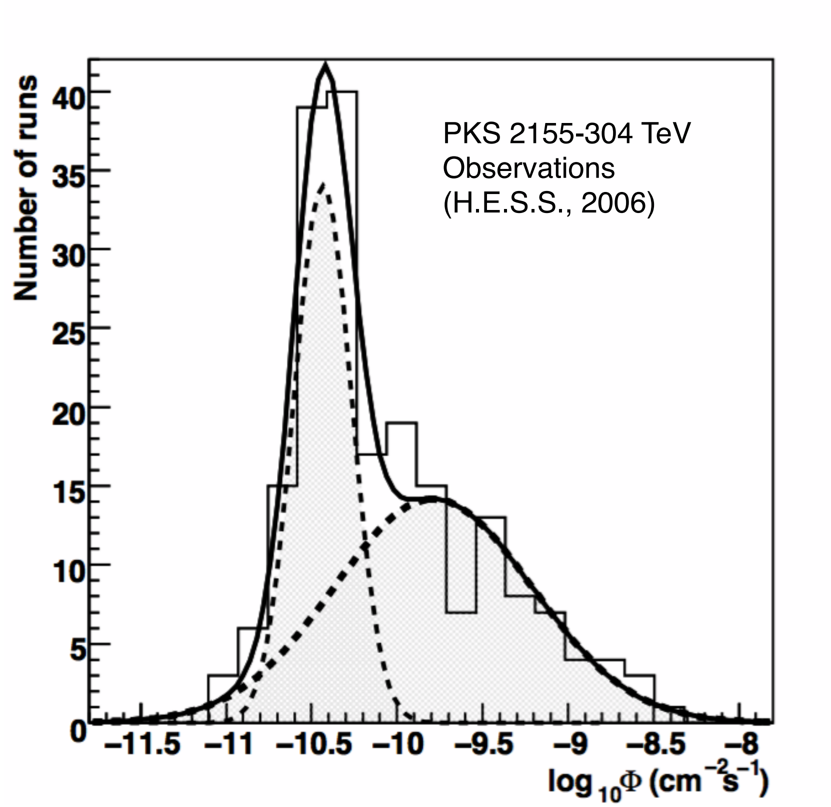


Motivation - Population / Diffuse Background

- Complexi and envir degenera
- Standard morpholo insufficien LCs ?
- Need **new** **“observa** understar Polarisati
- Large dat Methods population) e.g. time se
- Better statistics per obs sources



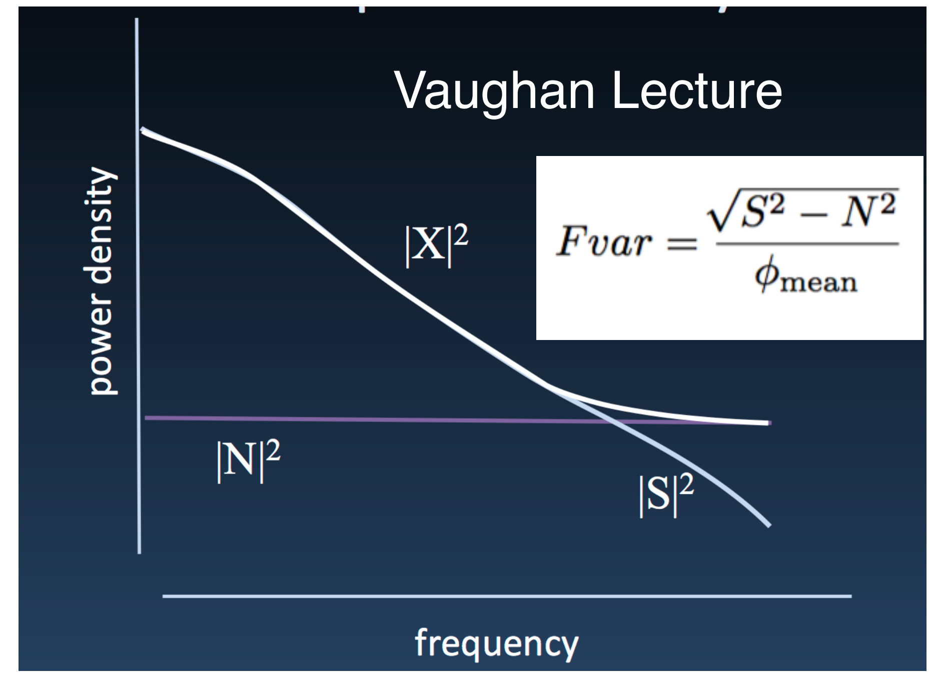
Additional (statistical) Observables : PSD and PDF



- **Distribution of fluxes (or PDF) probes the fundamental form of the physical processes**

- Default assumption is Gaussian ; evidence for lognormality => Multiplicative **(Lyubarskii 97, Uttley et a., 2005)** or Cascade like processes (exception see Biteau and Giebels, 2012)

- Contains the skewness and kurtosis of the underlying data



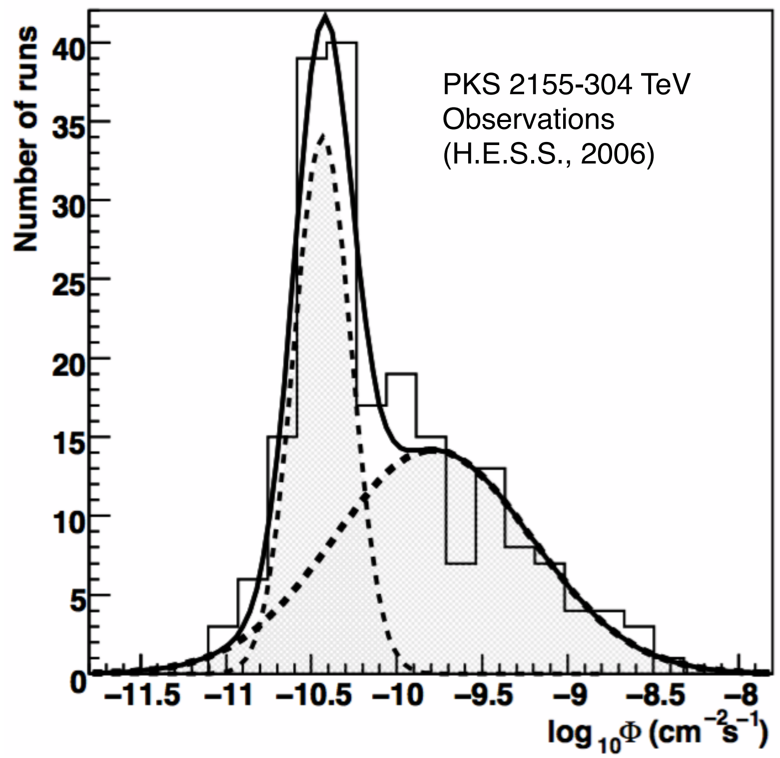
- **Distribution of timescales" (or PSD) encodes temporal structure**

- Time : x = s + n (Vaughan Lecture)
Fourier : X = S + N
$|X|^2 = |S|^2 + |N|^2 +$ Cross
PSD(f) $= \langle |S|^2 \rangle = \langle |X|^2 \rangle - \langle |N|^2 \rangle$

- Formally (for AGNs and others)
Time : Lightcurve(t) = Dynamical(t) x Acceleration(t) x Radiation(t) x Observation(t) [**Product**]

- First 2 moments - mean and variance

Additional (statistical) Observables : PSD and PDF



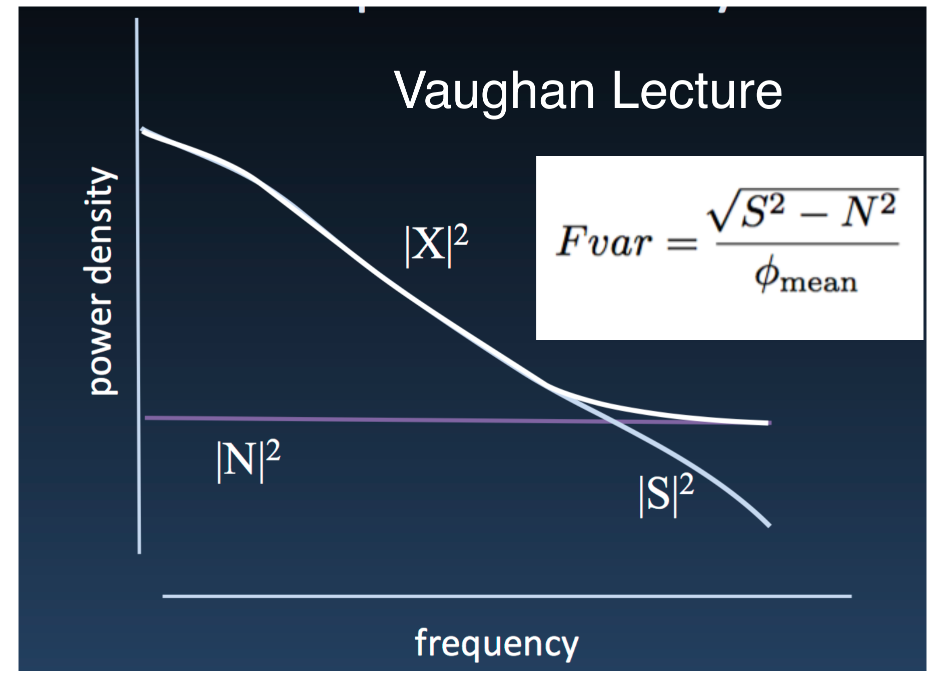
- **Distribution of fluxes (or PDF) probes the fundamental form of the physical processes**

$$\sum f_i(t) = f_1(t) + f_2(t) + ... \xrightarrow{\text{Central Limit Theorem}} \frac{e^{-\left(\frac{f-\mu_f}{\sigma_f}\right)^2}}{2\,\pi\,\sigma_f^2}$$

or Cascade like processes (exception

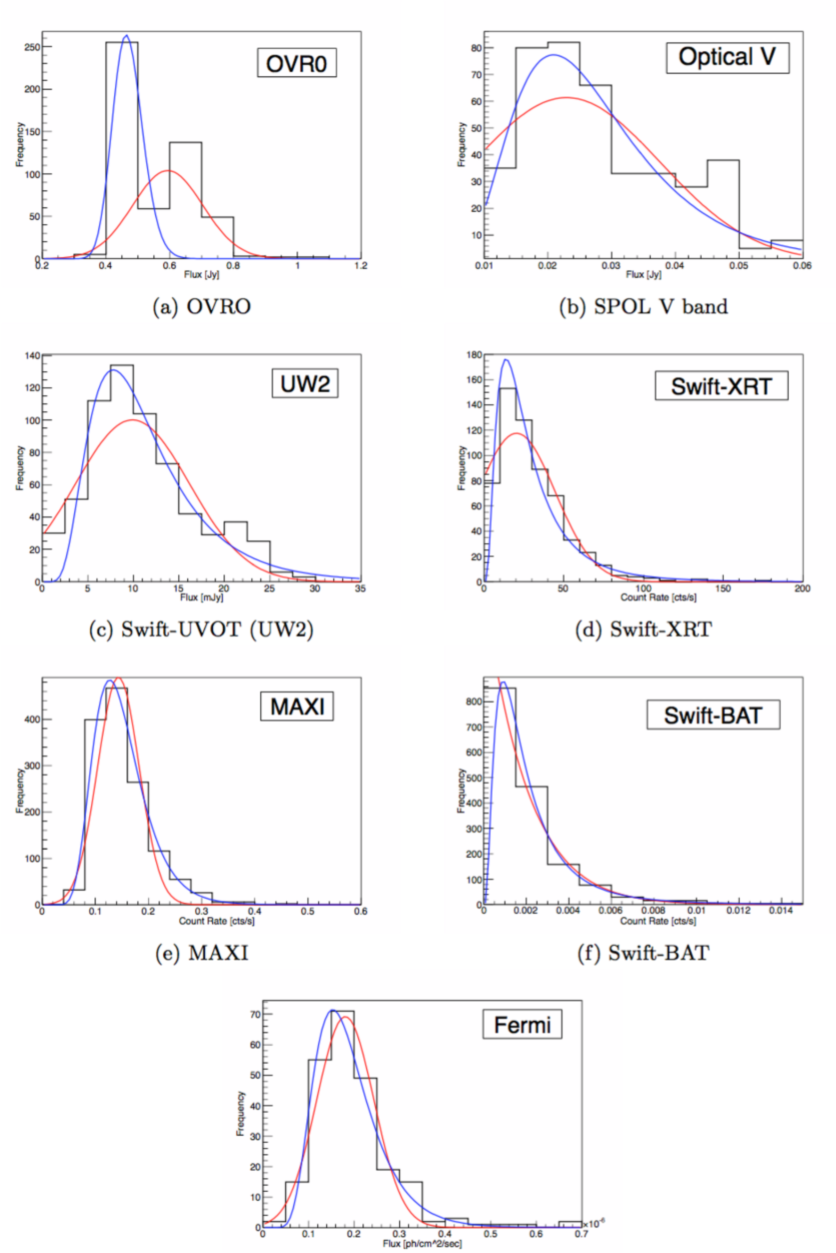
$$\prod f_i(t) \xrightarrow{\log} \log\left[f_1(t)\right] + \log\left[f_2(t)\right] + ... \xrightarrow{\text{Central Limit Theorem}} \frac{e^{-\left(\frac{\log f-\mu_{lf}}{\sigma_{lf}}\right)^2}}{2\,\pi\,\sigma_{lf}^2}$$

- Contains the skewness and kurtosis of the underlying data

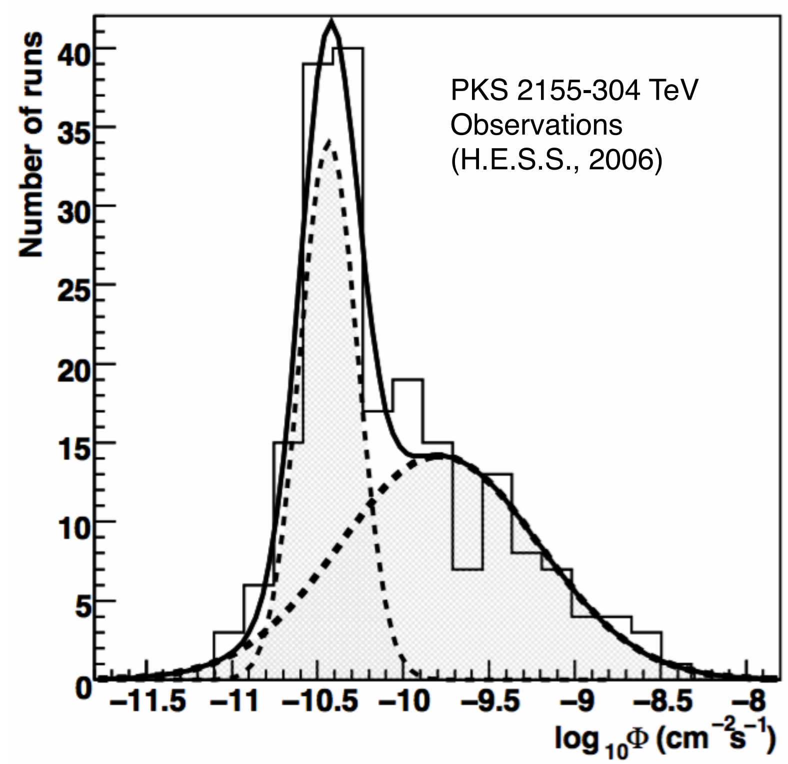


- **Distribution of timescales” (or PSD) encodes temporal structure**
- Time : x = s + n (Vaughan Lecture)
  Fourier : X = S + N
  $|X|^2 = |S|^2 + |N|^2$ + Cross
  PSD(f) $= \langle |S|^2 \rangle = \langle |X|^2 \rangle - \langle |N|^2 \rangle$
- Formally (for AGNs and others)
  Time : Lightcurve(t) = Dynamical(t) x Acceleration(t) x Radiation(t) x Observation(t) [**Product**]
- First 2 moments - mean and variance

PDF : Observational precedence



Mrk 421 : Sinha et al., 2016



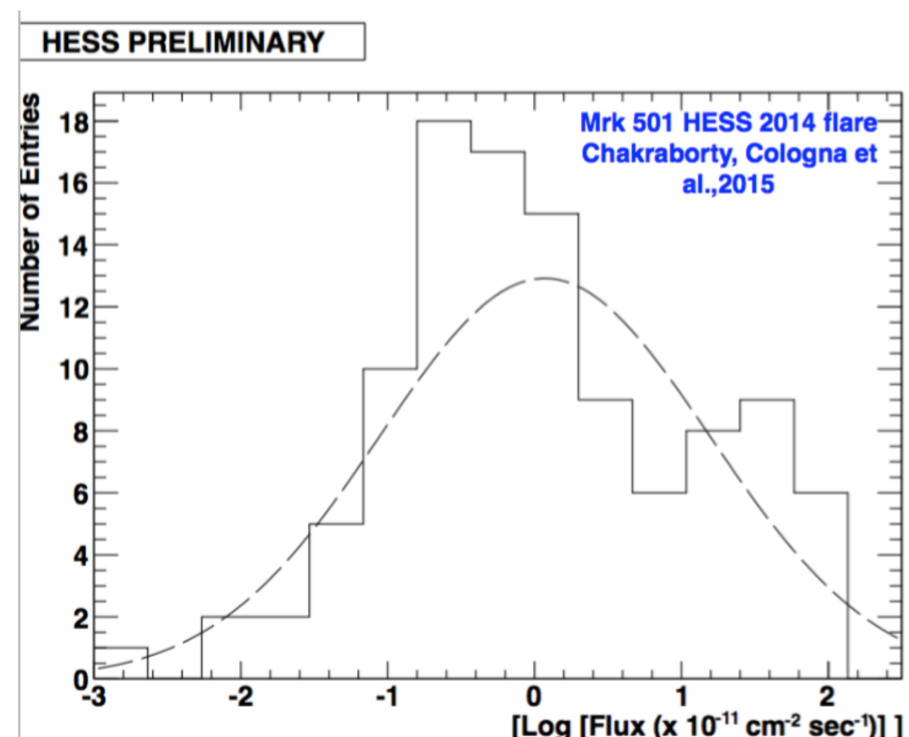
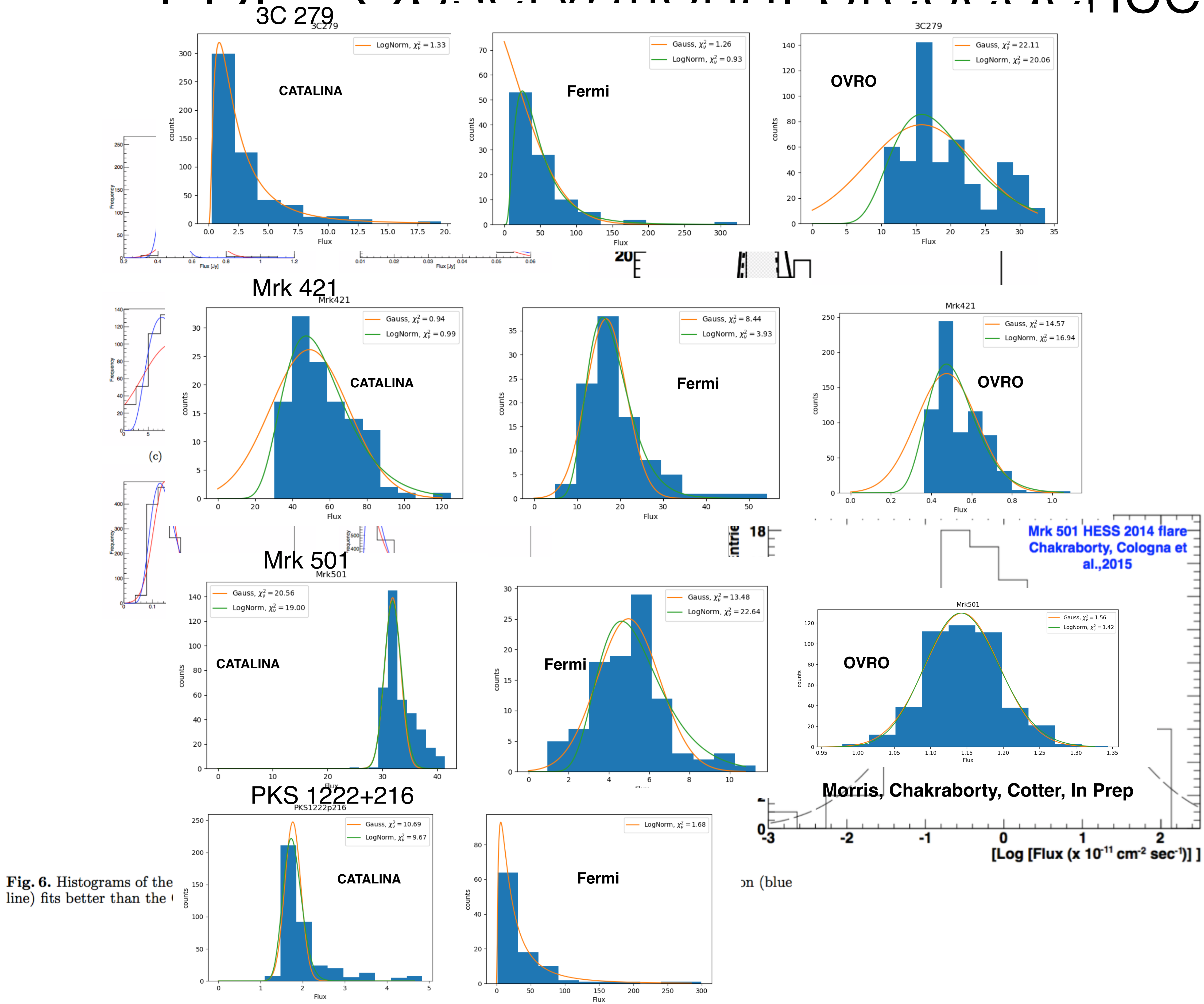


Fig. 6. Histograms of the fluxes (shown in black) at different wavebands. In all the cases, a lognormal distribution (blue line) fits better than the Gaussian distribution (red line). The reduced chi-squares are given in Table 3.

PDF : Observational precedence

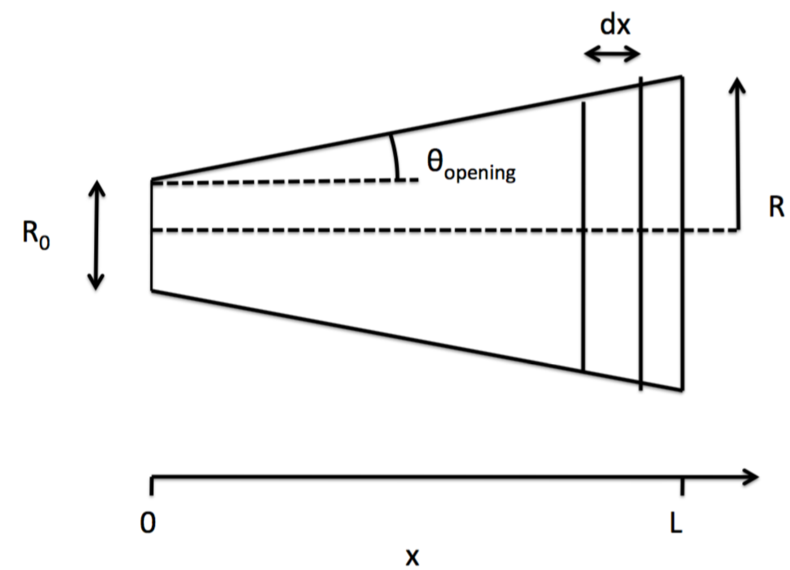


Origin of Lognormality ?

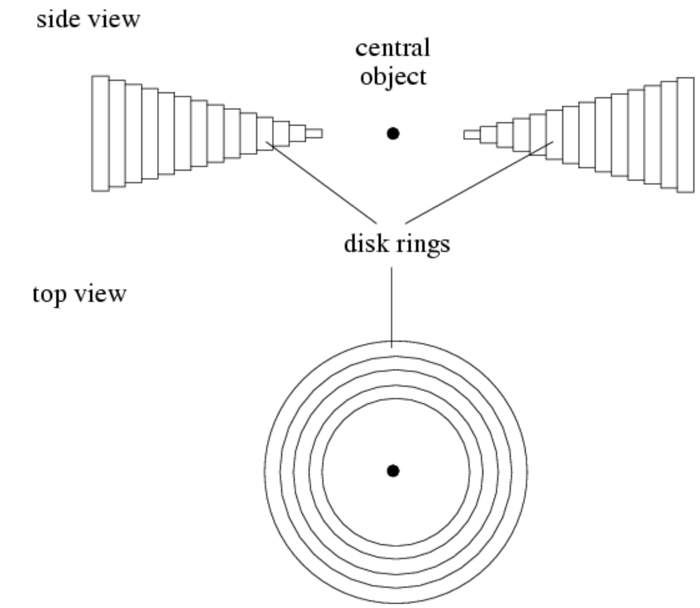
- Multiplicative process <=> Lognormality

- Lyubarskii's accretion disk => Fluctuations propagating from outer to inner rings => Multiplicative

- Analogous picture for jets



Potter & Cotter.



Nagel et al.,2004

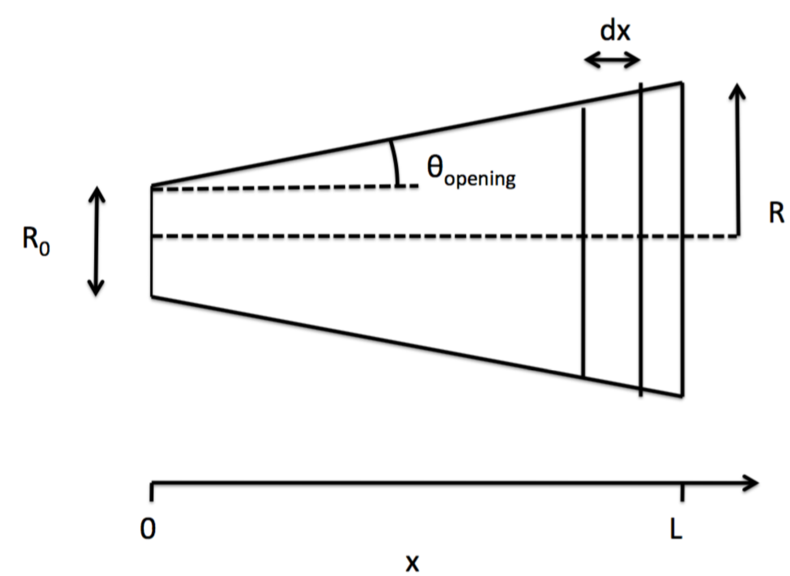
Lyubarskii, 1997 (flicker noise in accretion)

$$\dot{M} = \dot{M}_0[1 + \dot{m}(r, t)],$$

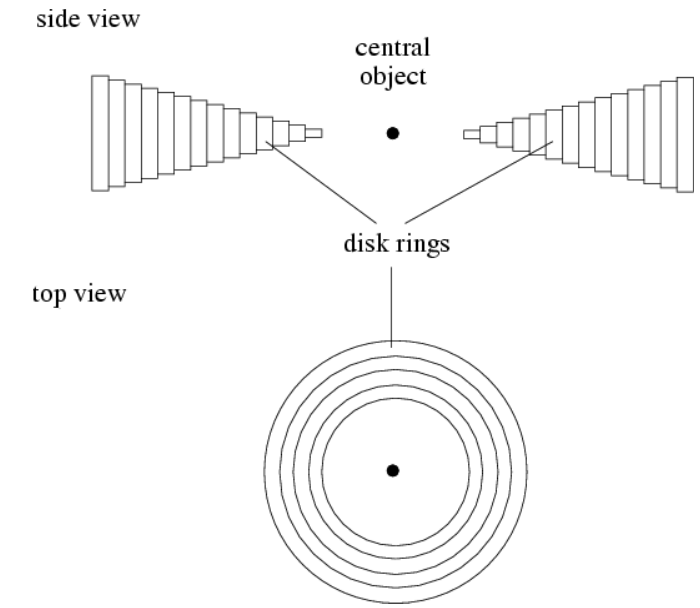
Chakraborty, Morris, Cotter, In Prep

Origin of Lognormality ?

- Multiplicative process <=> Lognormality
- Lyubarskii's accretion disk => Fluctuations propagating from outer to inner rings => Multiplicative
- Analogous picture for jets



Potter & Cotter.



Nagel et al.,2004

Lyubarskii, 1997 (flicker noise in accretion)

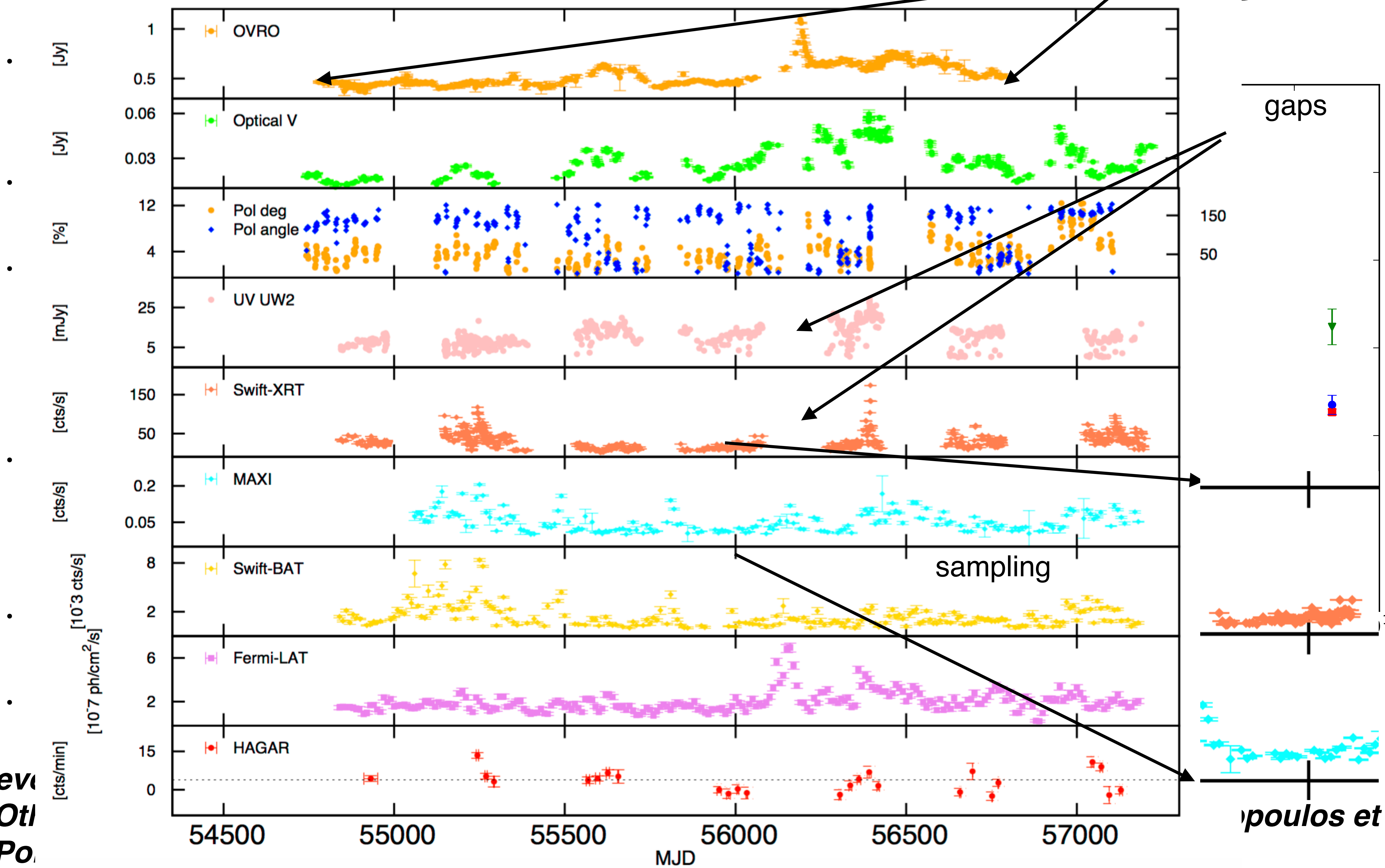
$$\dot{M} = \dot{M}_0[1 + \dot{m}(r, t)],$$

$$\sum f_i(t) = f_1(t) + f_2(t) + ... \xrightarrow{\text{Central Limit Theorem}} \frac{e^{-\left(\frac{f-\mu_f}{\sigma_f}\right)^2}}{2\ \pi\ \sigma_f^2}$$

$$\prod f_i(t) \xrightarrow{\log} \log\left[f_1(t)\right] + \log\left[f_2(t)\right] + ... \xrightarrow{\text{Central Limit Theorem}} \frac{e^{-\left(\frac{\log f - \mu_{lf}}{\sigma_{lf}}\right)^2}}{2\ \pi\ \sigma_{lf}^2}$$

Chakraborty, Morris, Cotter, In Prep

Variability Energy Distribution



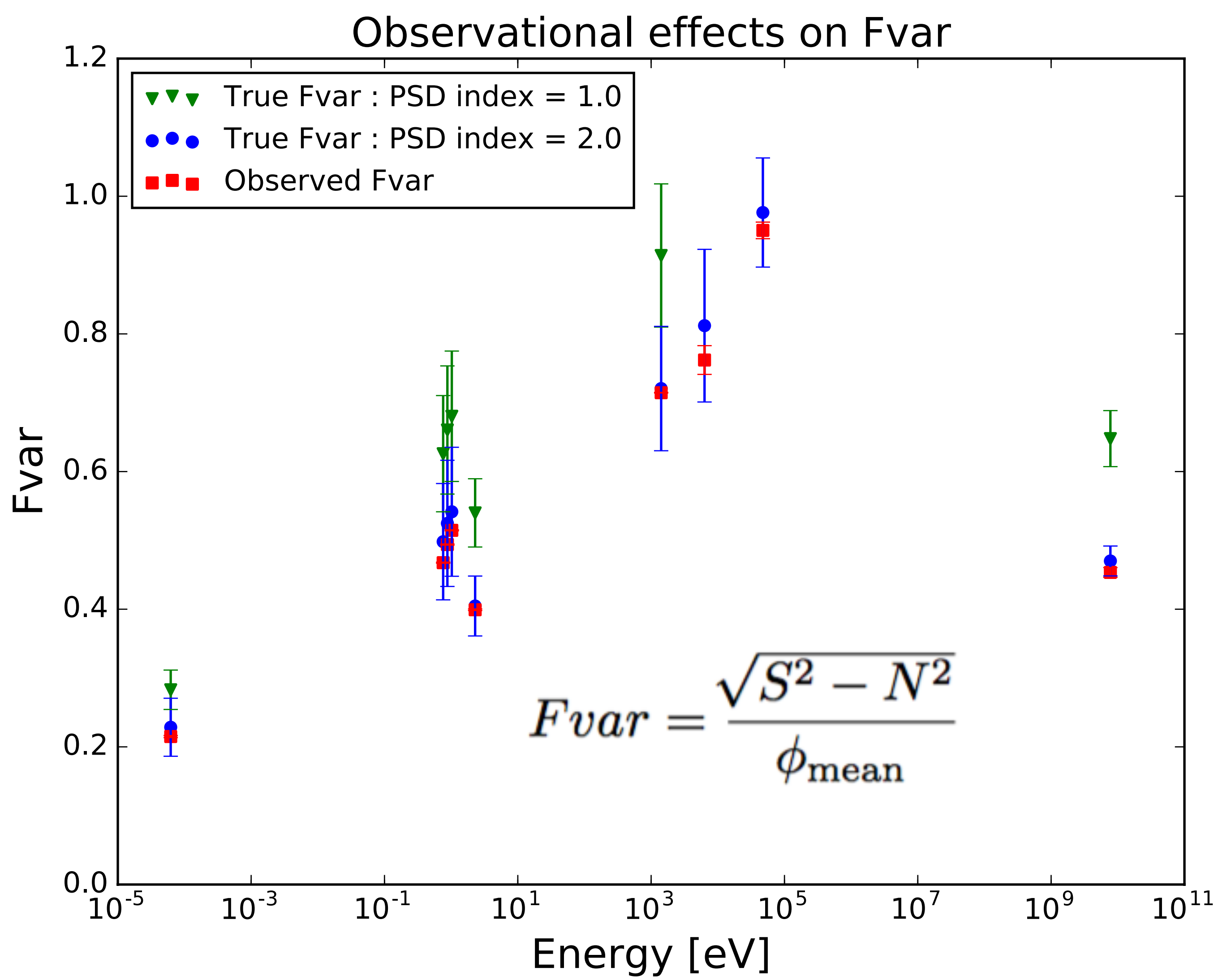
- *Seve*
- *Otl*
- *Po*
- *Estimation of flaring in AGNs*

etc…..

poulos et 13)

Variability **E**nergy **D**istribution

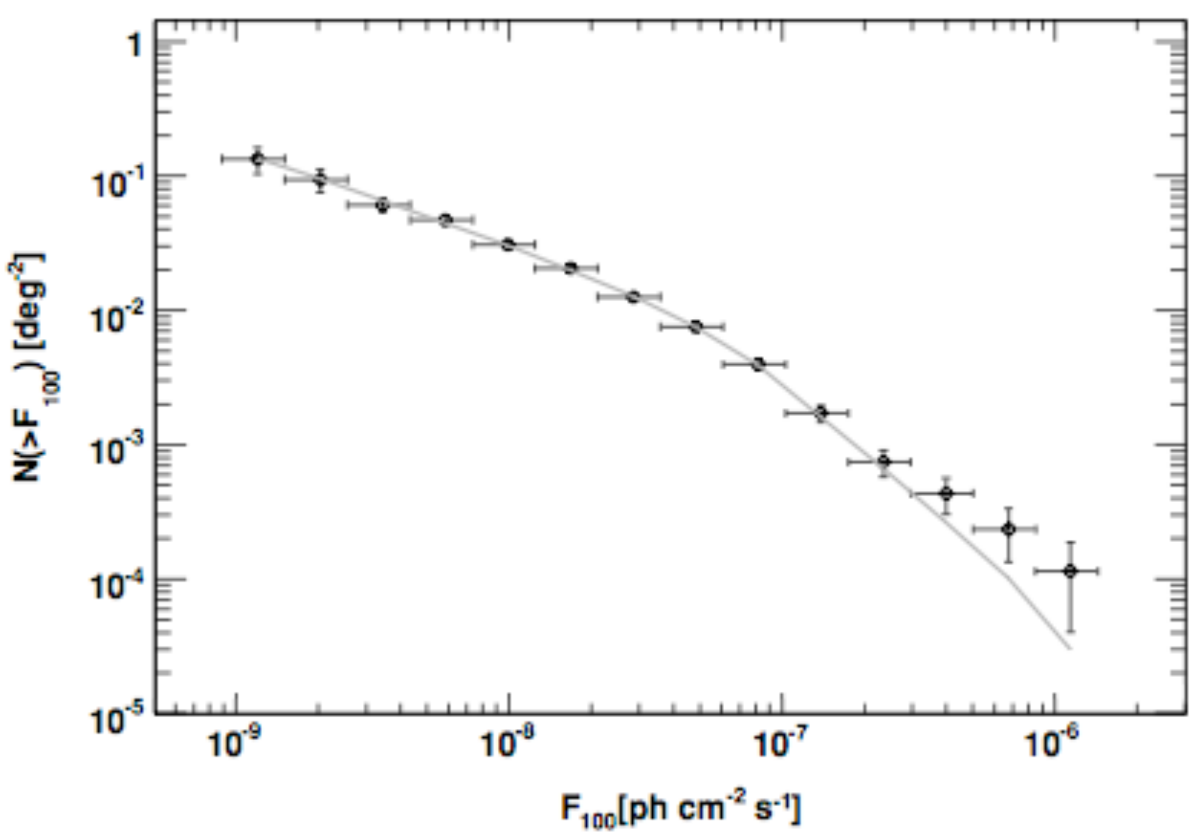
- Both the **uncertainty and bias due to observational effects are non-trivial**
- Simple yet **not unbiased estimator**
- Correct estimate of variability necessitates incorporating observational cadence **(uneven and sparse sampling)** errors and biases
- **Errors on flux bins less important** than errors due to non-accounting of gaps and sampling limits
- Account with simulations as shown
- Then model VED along with SED



- ***Several applications with simulated LCs***
  - Other estimators like CCF, doubling times, etc. (previous talks, Vaughan, Emmanuoulopoulos et 13)
  - Polarisation variability (Blinov - RoboPol first season results)
  - Estimation of flaring in AGNs
  etc…..

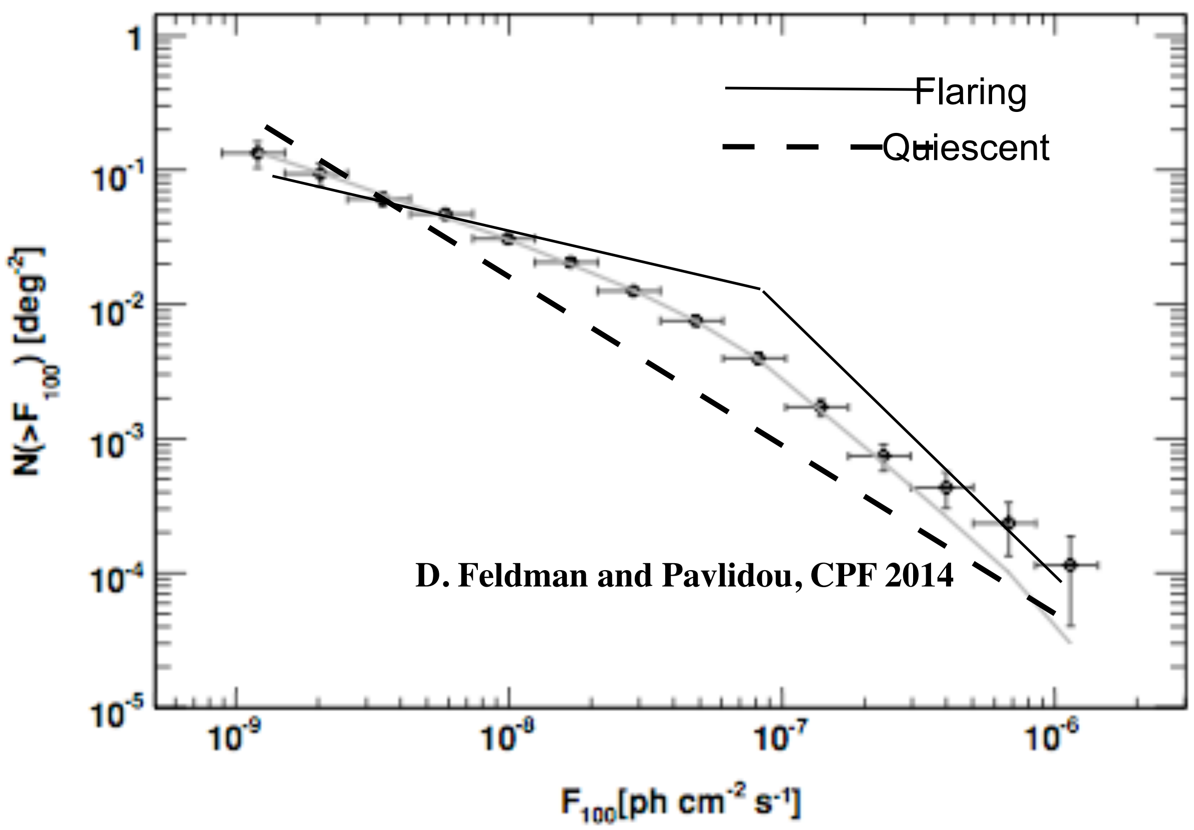
Variable Population : Source Flux / Counts Distribution

Fermi -LAT Collaboration



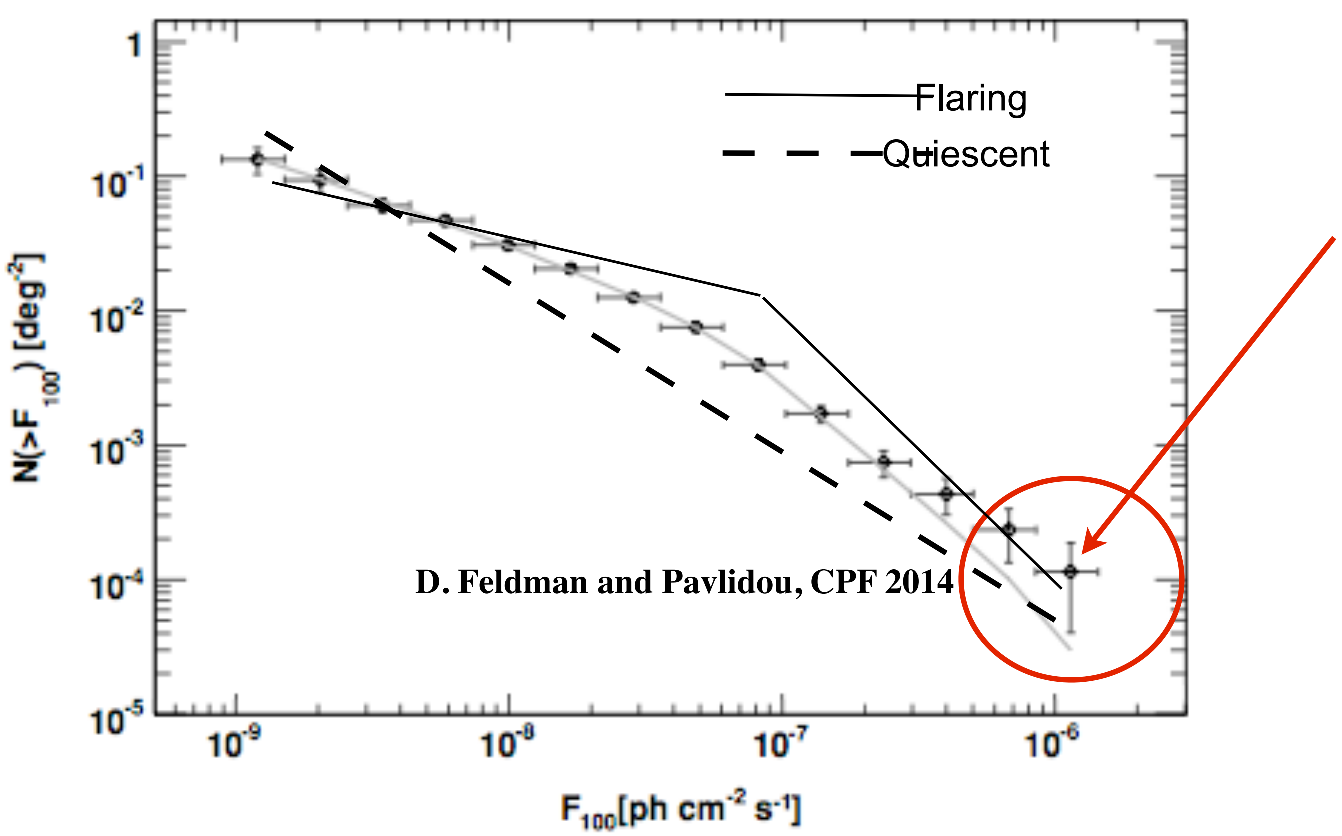
Variable Population : Source Flux / Counts Distribution

Fermi -LAT Collaboration



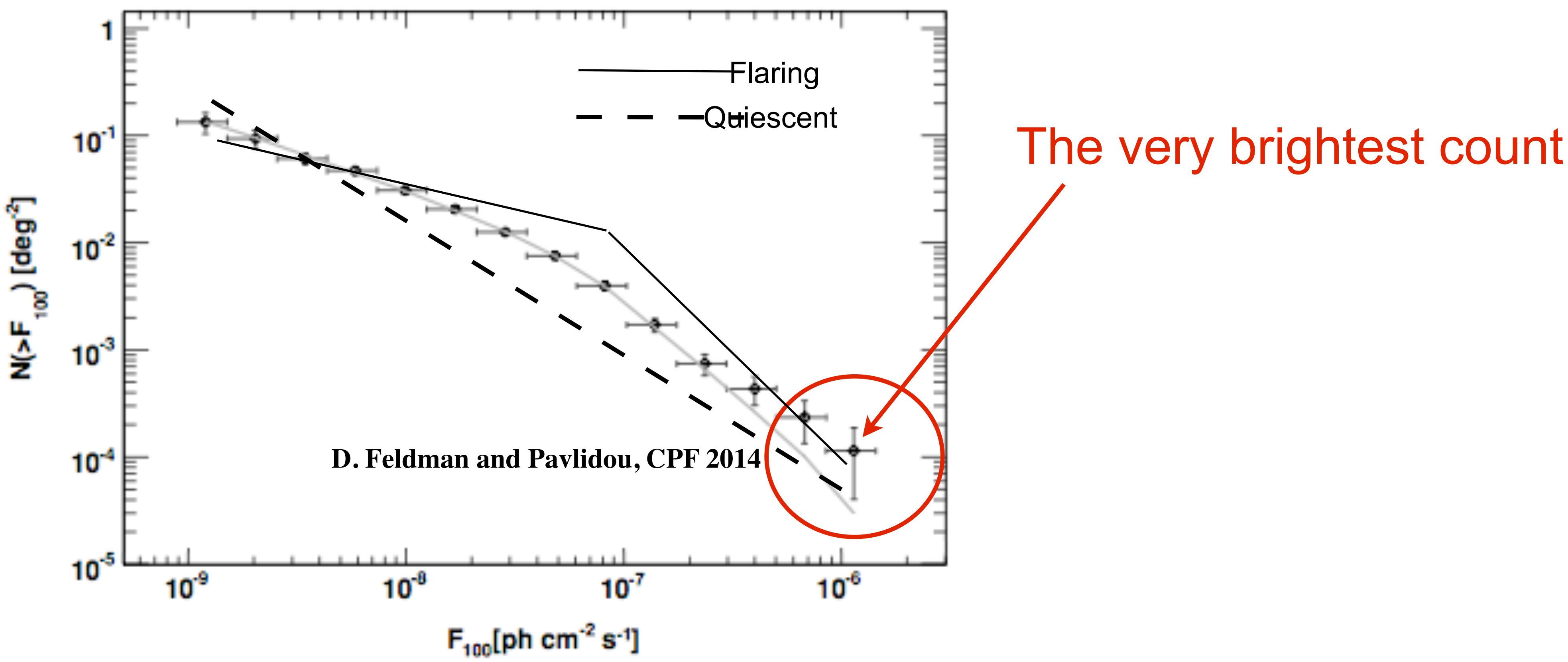
Variable Population : Source Flux / Counts Distribution

Fermi -LAT Collaboration



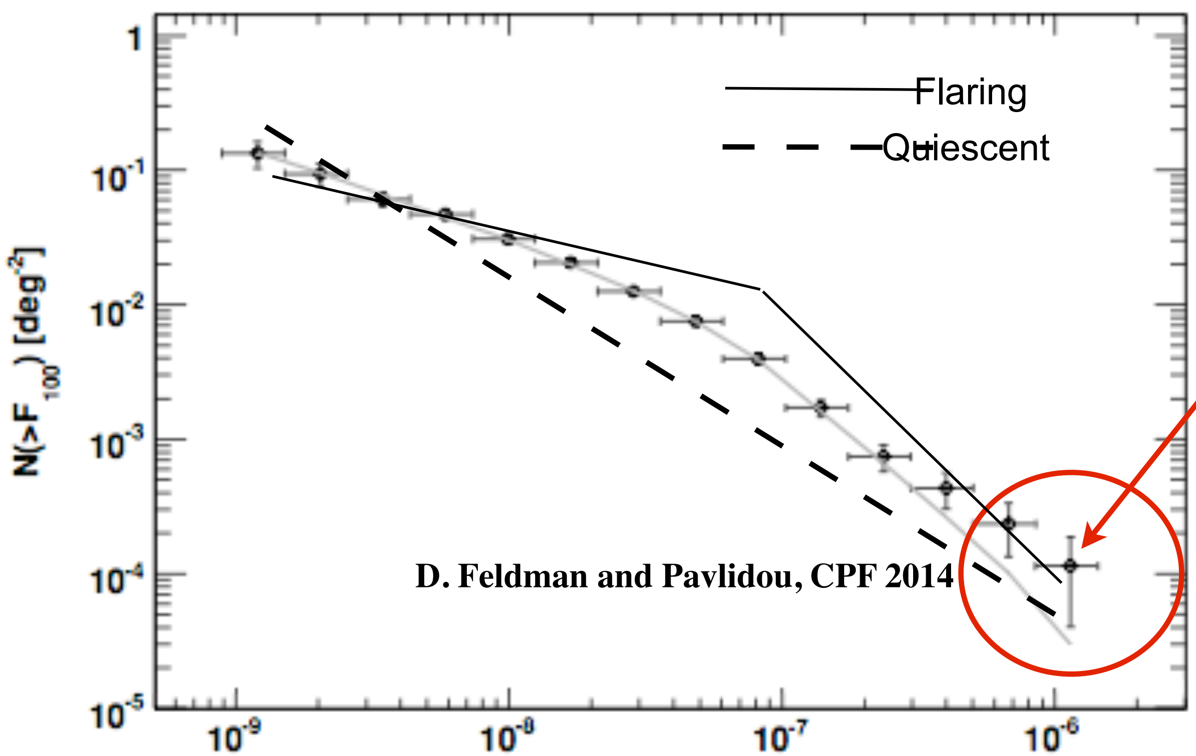
Variable Population : Source Flux / Counts Distribution

Fermi -LAT Collaboration



Variable Population : Source Flux / Counts Distribution

Fermi -LAT Collaboration



The very brightest count

For flaring timescales of a week
=> Handful to atleast 2 x Handful

Chakraborty, Pavlidou and Fields, 2014

Conclusions

- Complexity of processes and environment of individual sources like AGNs necessitates “novel observables”

- Also relevant for population studies and diffuse backgrounds

- Model degeneracies can be lifted in both cases

- Better instruments => Better data / statistics => Statistical observables (PSD, PDF)

- Improve theoretical understanding of observables in terms of physical processes

Thank you !!!

Acknowledgments
Jonathan Biteau (IPNO)
Paul Morris (Oxford)
Garret Cotter (Oxford)
Frank Rieger (MPIK)
HESS Collaboration

Supplementary

PSD : Simulations
General Approach

- Observed Emission
  - Function of **time (lightcurve)**, space (morphology), energy (energy spectrum) **How tells us why**
  - Individual sources : physical mechanisms at emission sites
  - Population : general trends

- **Timing analysis** : Observed light curve is 1 sample or realisation -> we need to "repeat" to get significant results
  (Timmer and Koenig, 1995, Emmanoulopoulos, McHardy and Papadakis, 2013)

- Signal coupled with noise
  - Either disentangle deterministic signal from random fluctuations (for eg. detecting periodic/QPOs)
  - Or the interesting signals are random fluctuations themselves (for eg. flaring vs quiescence)

- Observational Irregularities : Allocation, satellite cycles, visibility, competing targets, etc
  - gaps
  - coarse or uneven sampling
  - length of observation limited

Emmanoulopoulos et al., 2013, Allevato et al., 2013, Chakraborty & Biteau (In prep)

PSD : Simulations
General Approach

- Observed Emission
  - Function of **time (lightcurve)**, space (morphology), energy (energy spectrum) **How tells us why**
  - Individual sources : physical mechanisms at emission sites
  - Population : general trends

- **Timing analysis** : Observed light curve is 1 sample or realisation -> we need to "repeat" to get significant results
  (Timmer and Koenig, 1995, Emmanoulopoulos, McHardy and Papadakis, 2013)

- Signal coupled with noise
  - Either disentangle deterministic signal from random fluctuations (for eg. detecting periodic/QPOs)
  - Or the interesting signals are random fluctuations themselves (for eg. flaring vs quiescence)

- Observational Irregularities : Allocation, satellite cycles, visibility, competing targets, etc
  - gaps
  - coarse or uneven sampling
  - length of observation limited

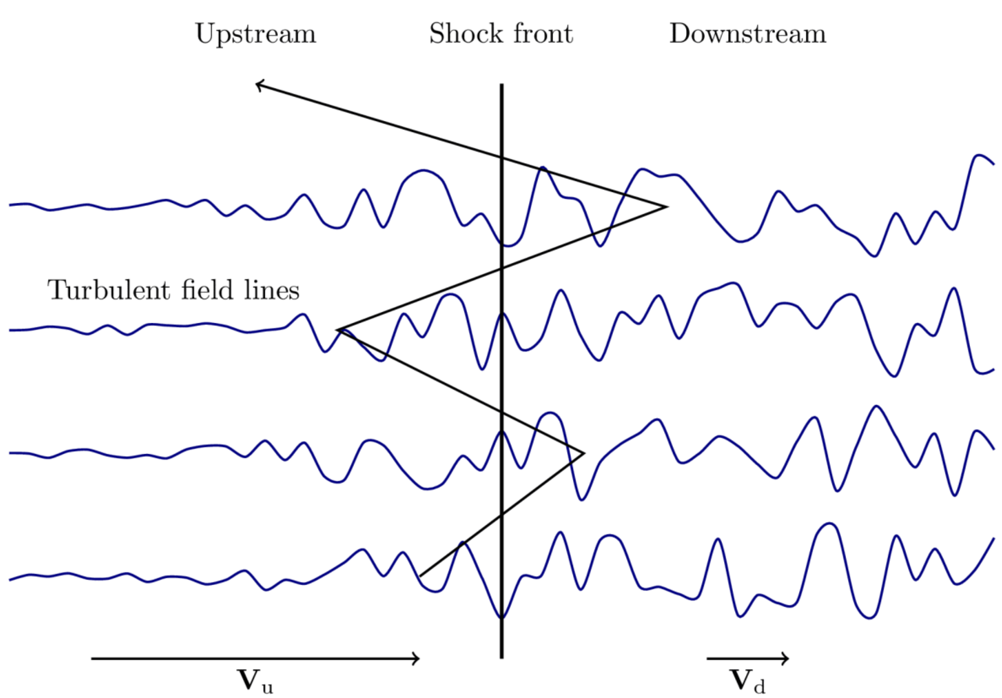
Emmanoulopoulos et al., 2013, Allevato et al., 2013, Chakraborty & Biteau (In prep)

Start with "fundamental observables" : PSD and PDF ⇒ Generate realisations / surrogate data (for e.g. lightcurves TK95,DE13, etc) ⇒ Impose observational conditions (cadence, cuts, uncertainties, etc)

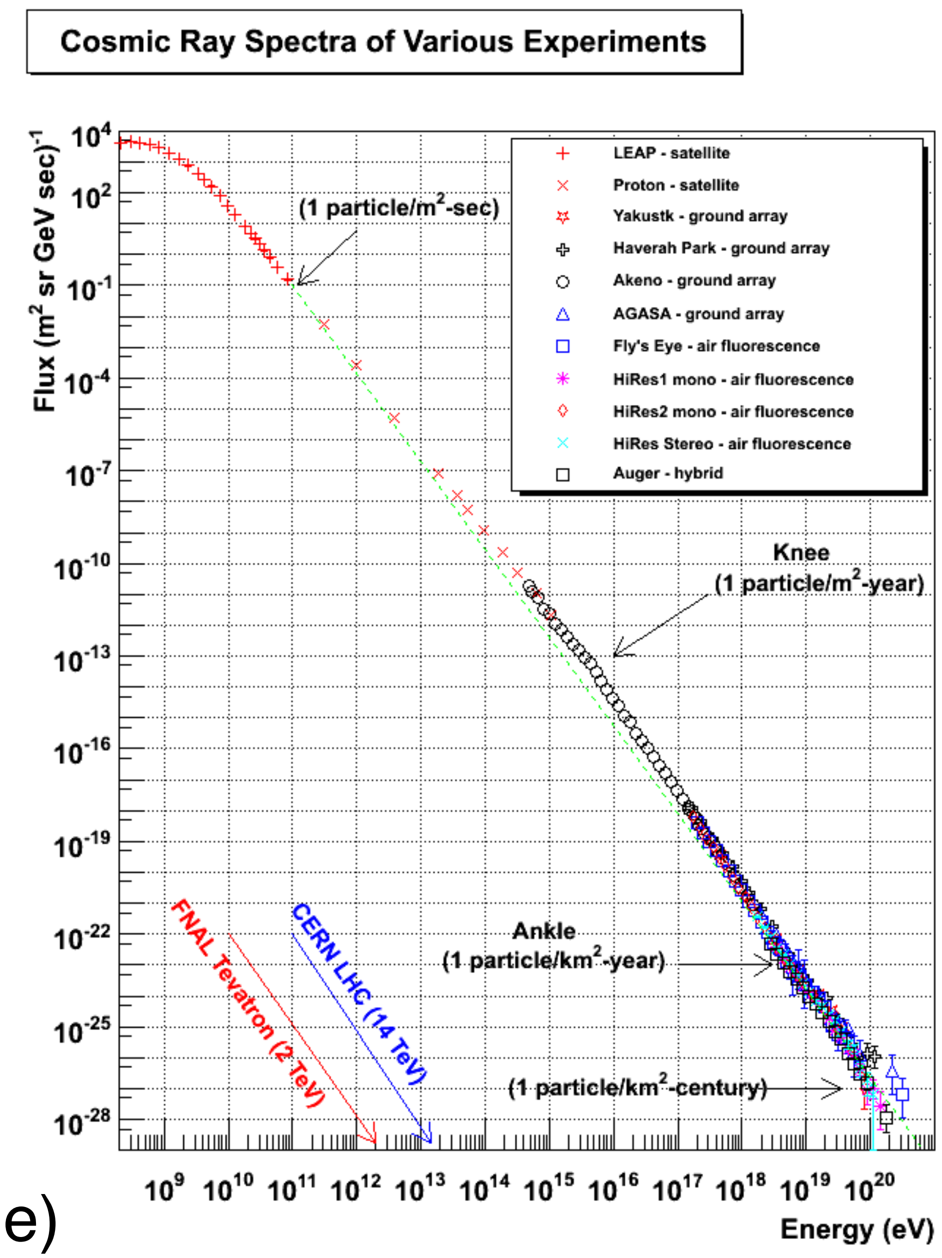
⇒ Evaluate estimators (Fvar, CCF, Spectral shape, etc.) for each realisation ⇒ Likelihood Analysis as function of the "fundamental parameters"

Particle Acceleration -> Power-Laws

- Lightcurve(f) = Dynamical(f) * **Acceleration**(f) * Radiation(f) * Observation(f)



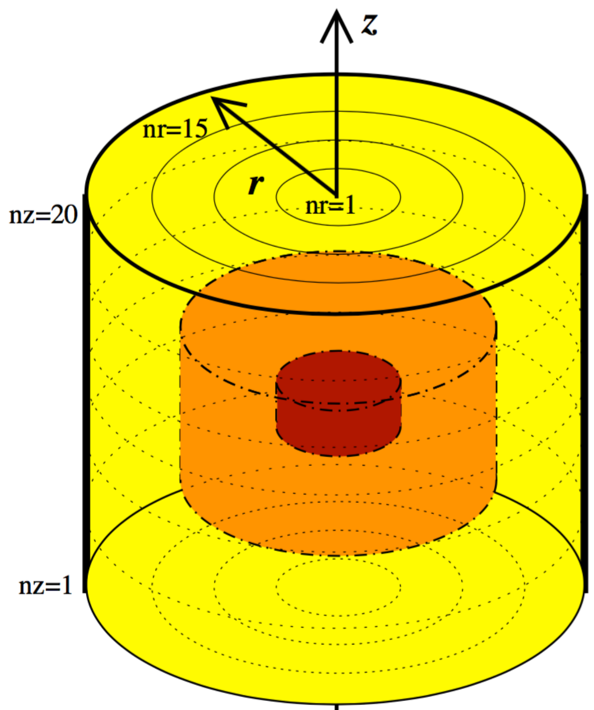
=>



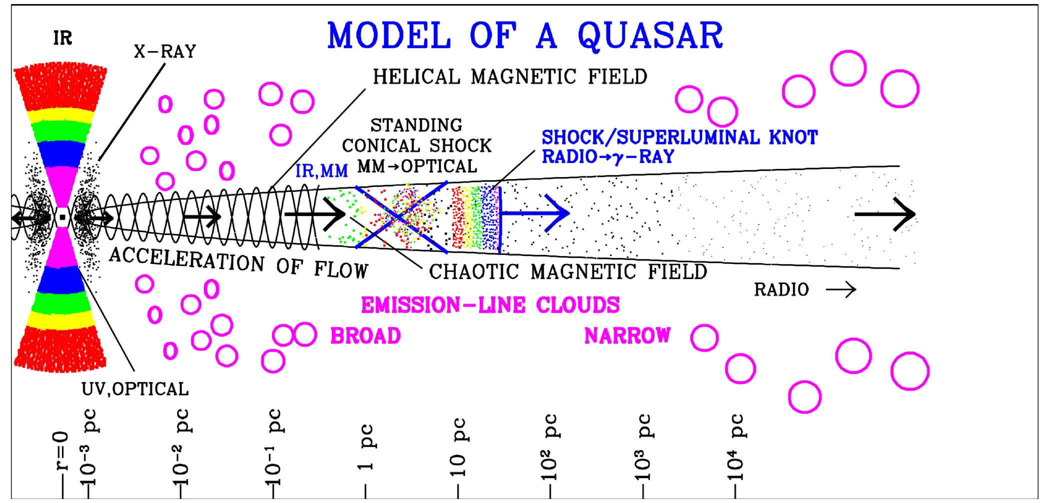
Could they have same origin ?

(Analytical / simulations with Simone Giacche)

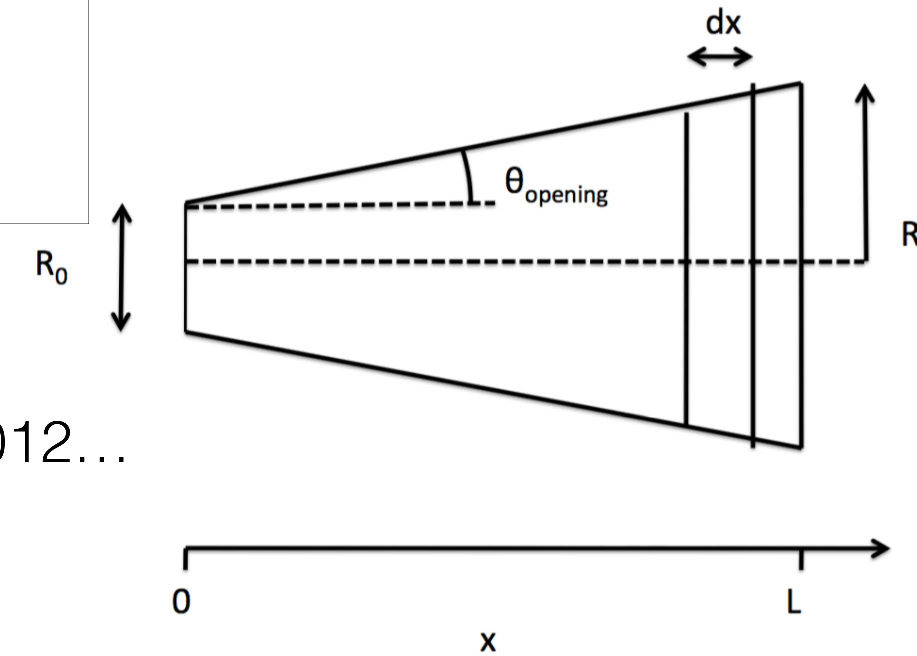
Physics of individual sources - AGN jets



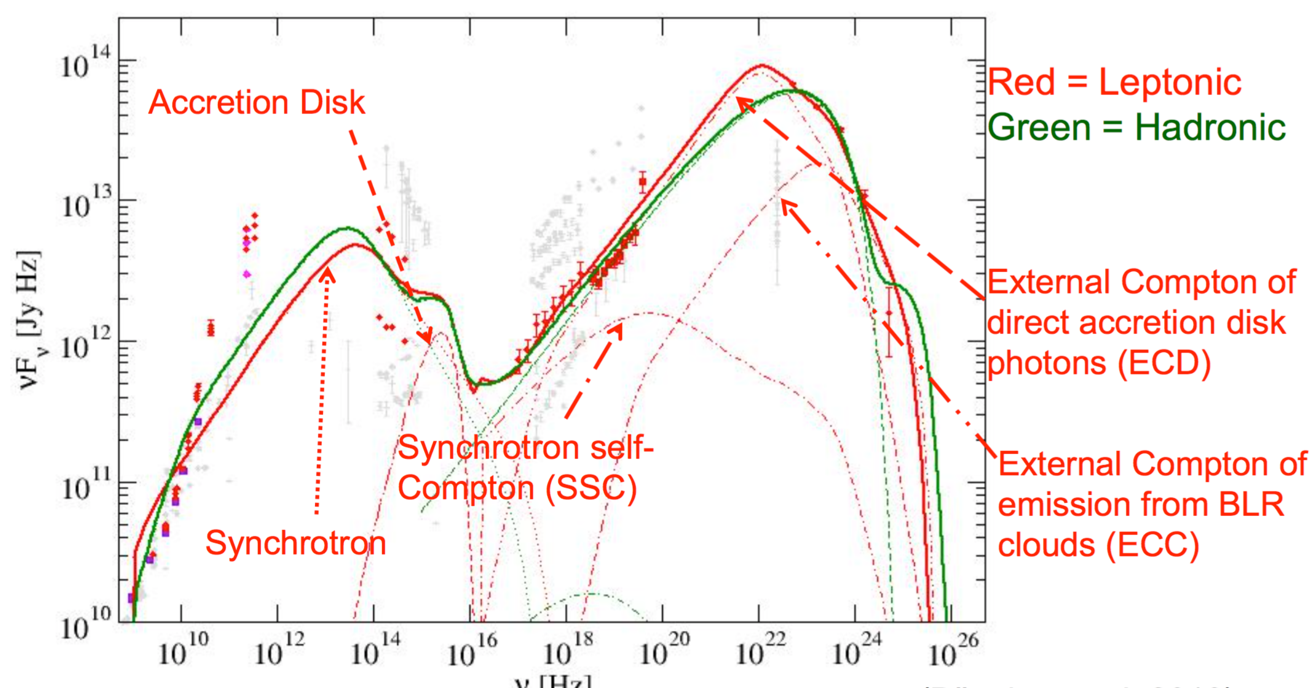
Chen, Botcher, Pohl



Marscher



Potter & Cotter, 2012…



(Böttcher et al. 2013)

Start with “fundamental observables” : PSD (**PL**) and PDF (**Lognormal**)

→ Generate realisations / surrogate data (for e.g. lightcurves **TK95**,DE13, etc)

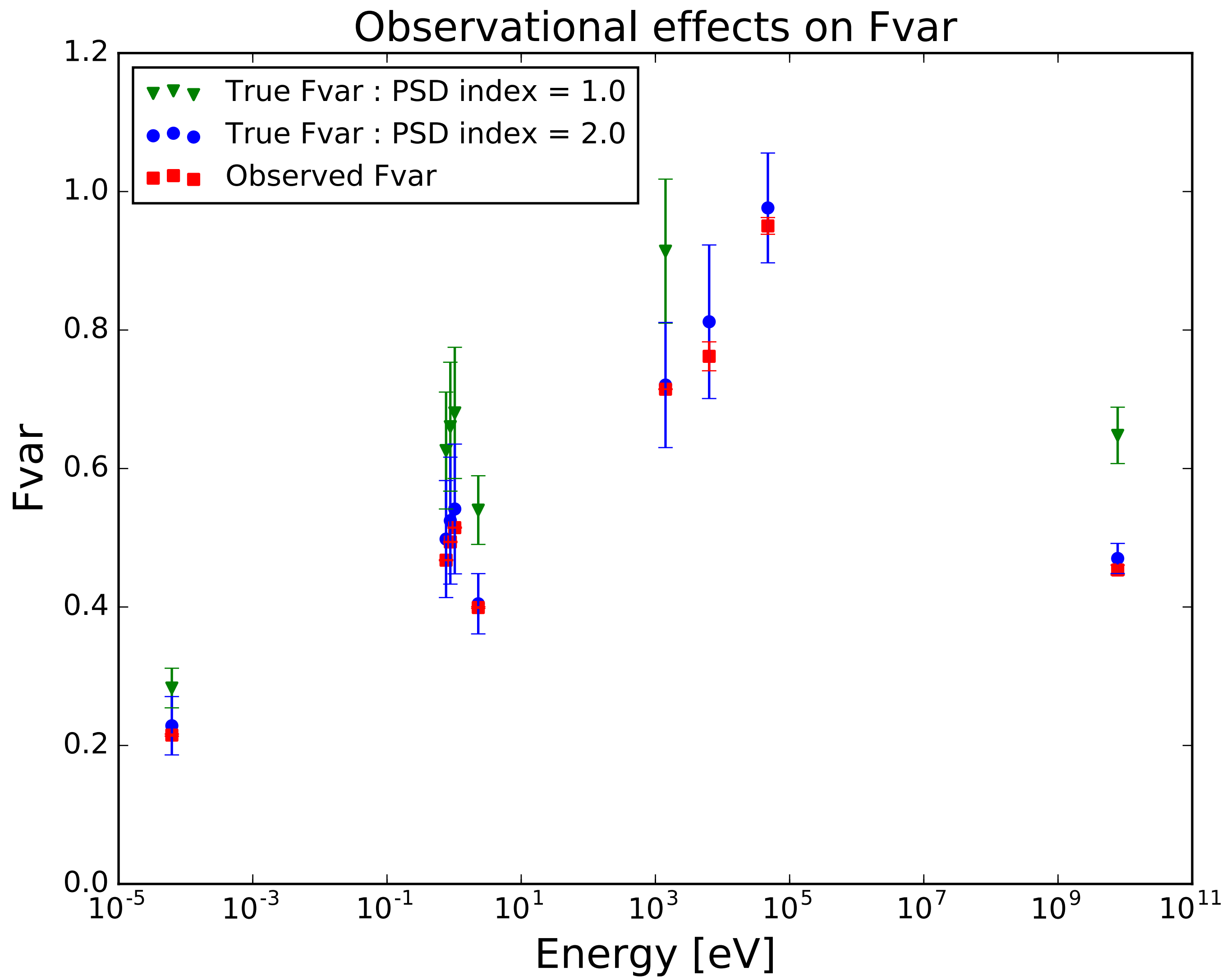
→ Impose observational conditions (cadence, cuts, uncertainties, etc)

→ Evaluate estimators (**Fvar**, CCF, Spectral shape, etc.) for each realisation

→ Likelihood Analysis as function of the “fundamental parameters”

Variability **E**nergy **D**istribution

- Even with red noise, both the **uncertainty and bias due to observational effects are non-trivial**
- Correct estimate of variability necessitates incorporating these systematic uncertainties
- Crucial to have coordinated observational cadence across wavelengths
- Further work - **non-Gaussian PDFs**, tests for stationarity



- ***Several applications with simulated LCs***
  - Other estimators like CCF, doubling times, etc. (previous talks, Vaughan, Emmanuoulopoulos et 13)
  - Polarisation variability (Blinov - RoboPol first season results)
  - Estimation of flaring in AGNs
  etc…..

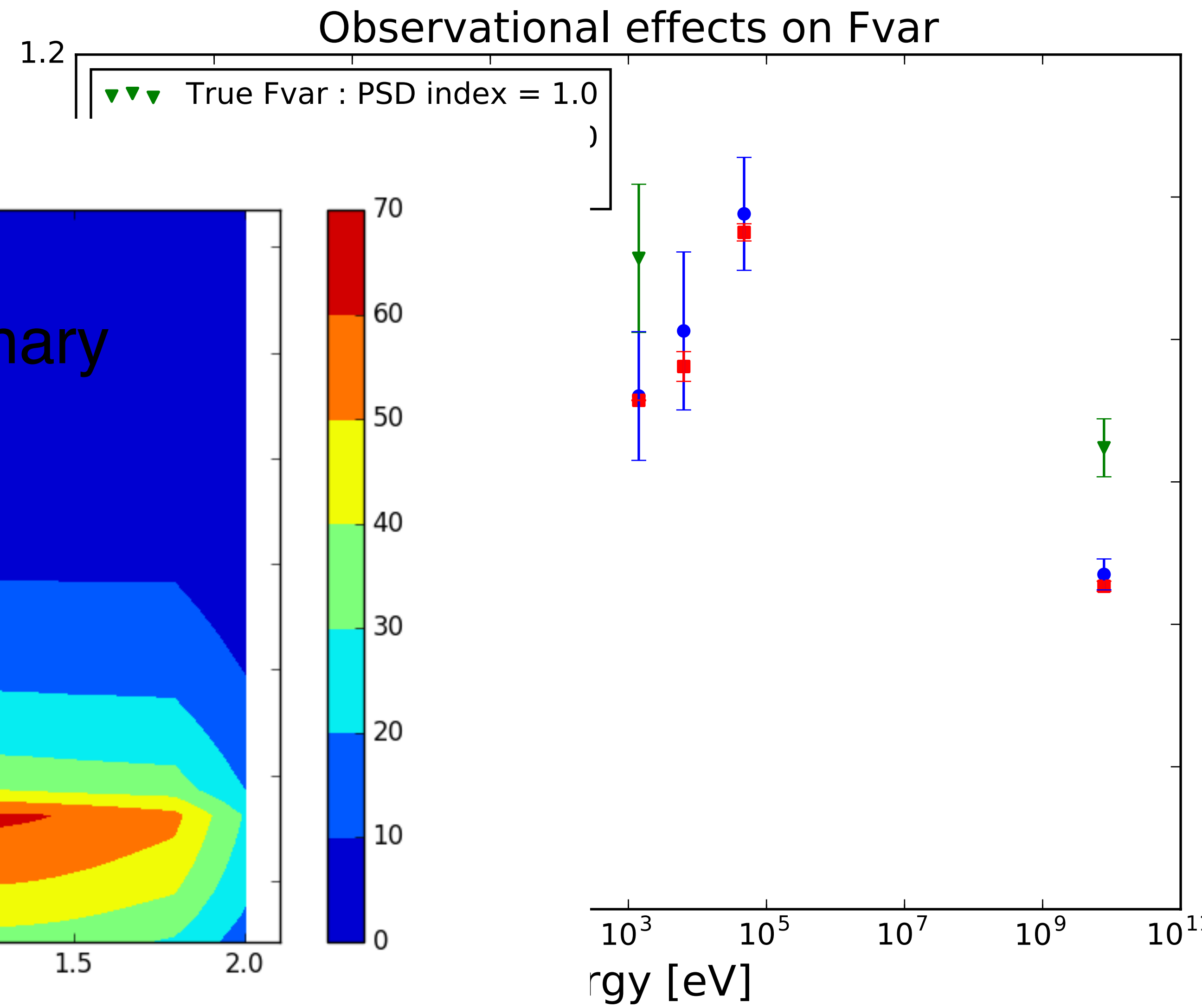
Variability **E**nergy **D**istribution

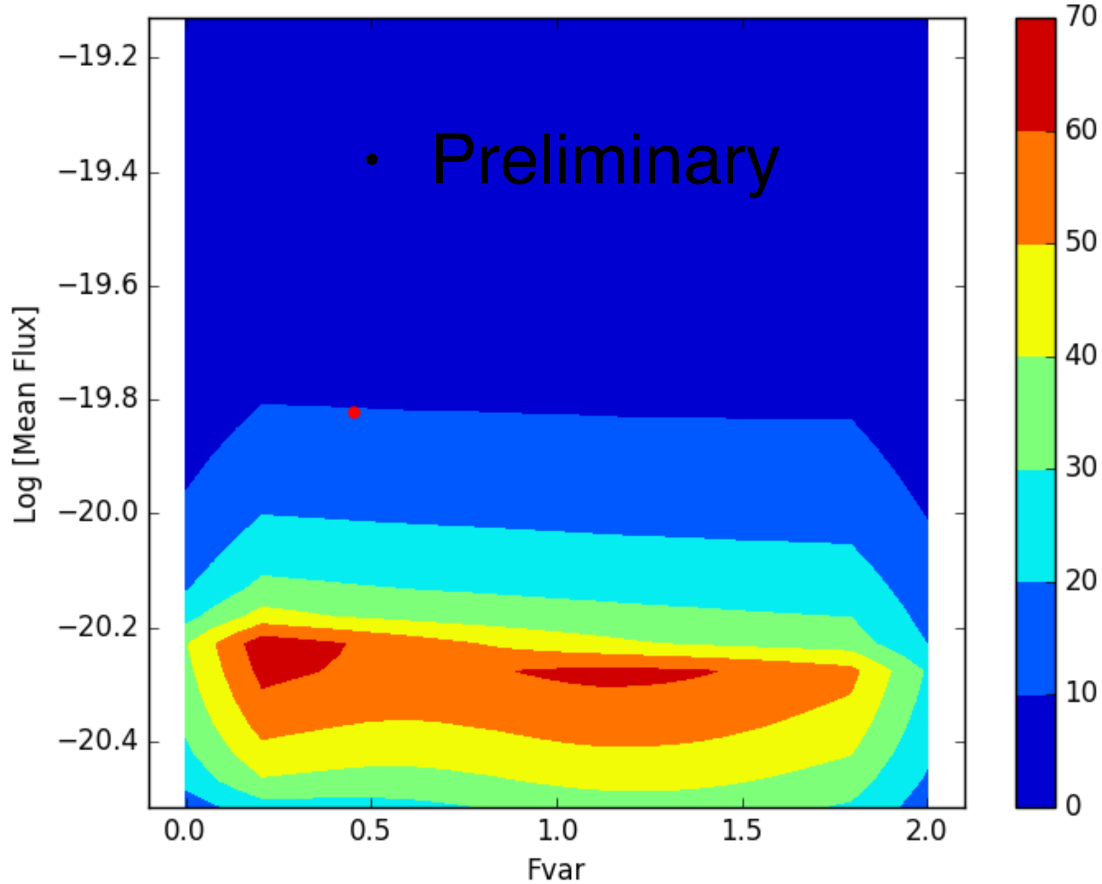
- Even with red noise, both the **uncertainty and bias due to observationa trivial**

- Correct estim necessitates i systematic un

- Crucial to hav observational wavelengths

- Further work - **PDFs**, tests for stationarity





- ***Several applications with simulated LCs***
  - Other estimators like CCF, doubling times, etc. (previous talks, Vaughan, Emmanuoulopoulos et 13)
  - Polarisation variability (Blinov - RoboPol first season results)
  - Estimation of flaring in AGNs
  etc…..

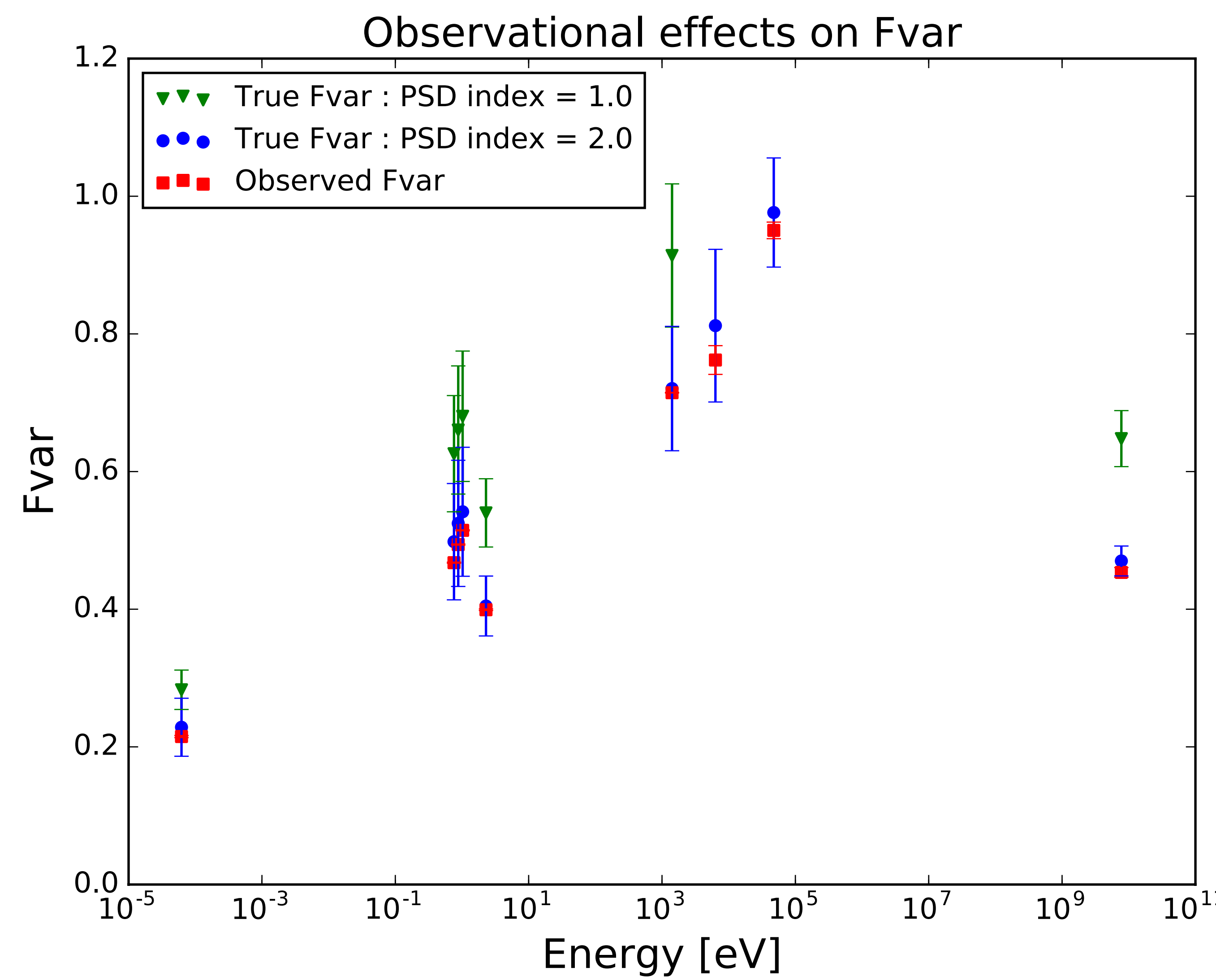
Variability **E**nergy **D**istribution

- Even with red noise, both the **uncertainty and bias due to observational effects are non-trivial**

- Correct estimate of variability necessitates incorporating these systematic uncertainties

- Crucial to have coordinated observational cadence across wavelengths

- Further work - **non-Gaussian PDFs**, tests for stationarity

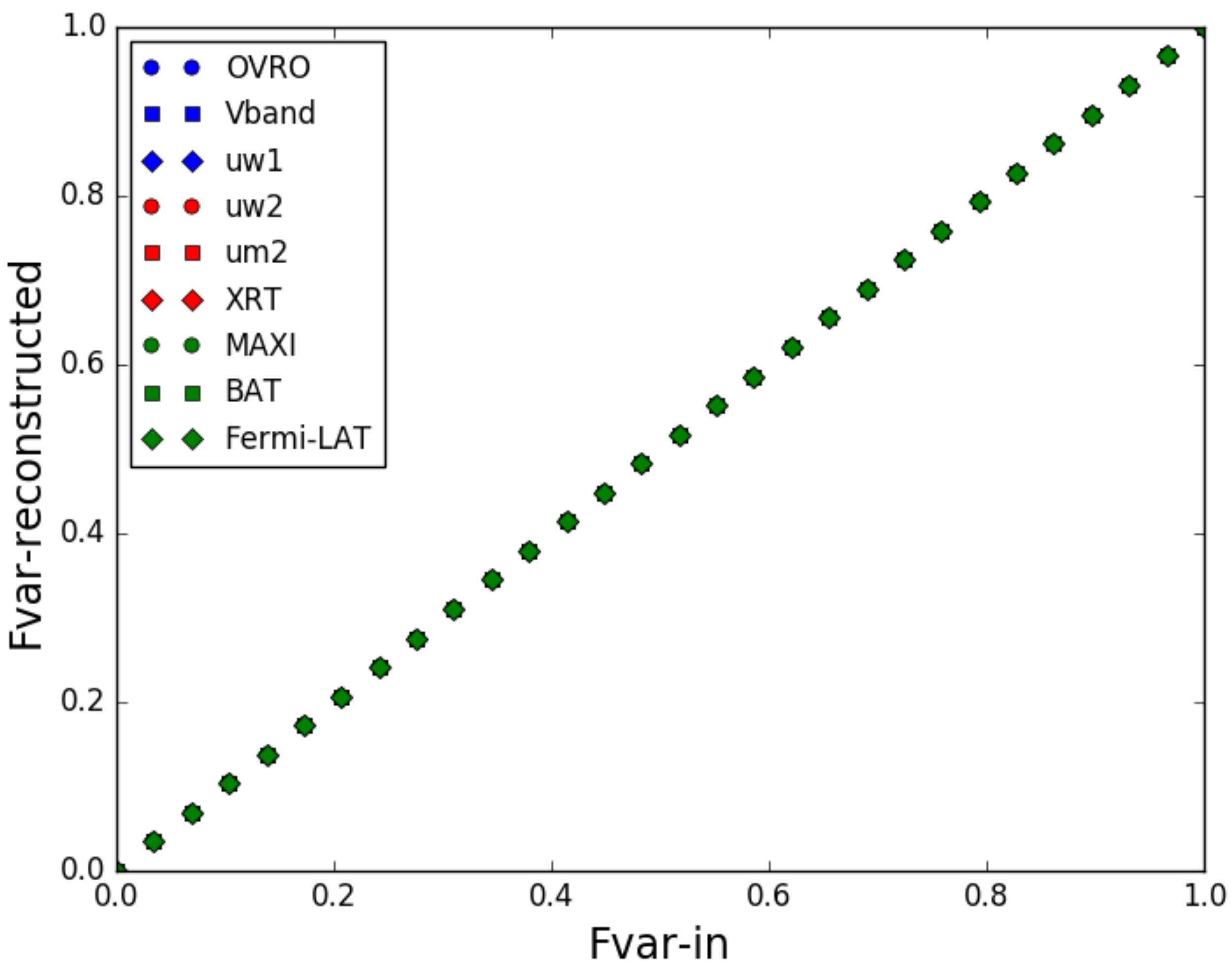


- ***Several applications with simulated LCs***
  - Other estimators like CCF, doubling times, etc. (previous talks, Vaughan, Emmanuoulopoulos et 13)
  - Polarisation variability (Blinov - RoboPol first season results)
  - Estimation of flaring in AGNs
  etc…..

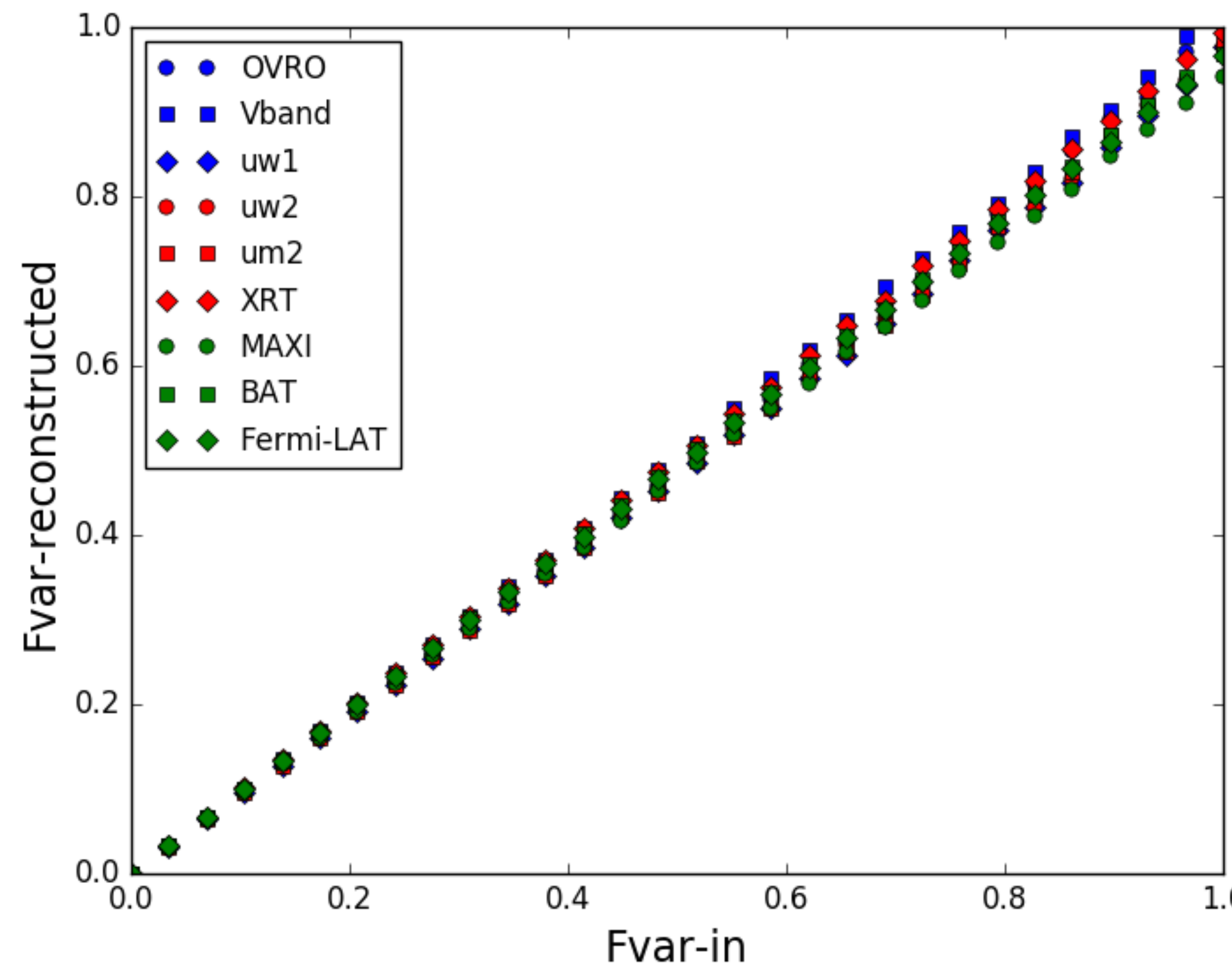
Fvar reconstruction

Red Noise : Index = 2.0

"Ideal" Cadence



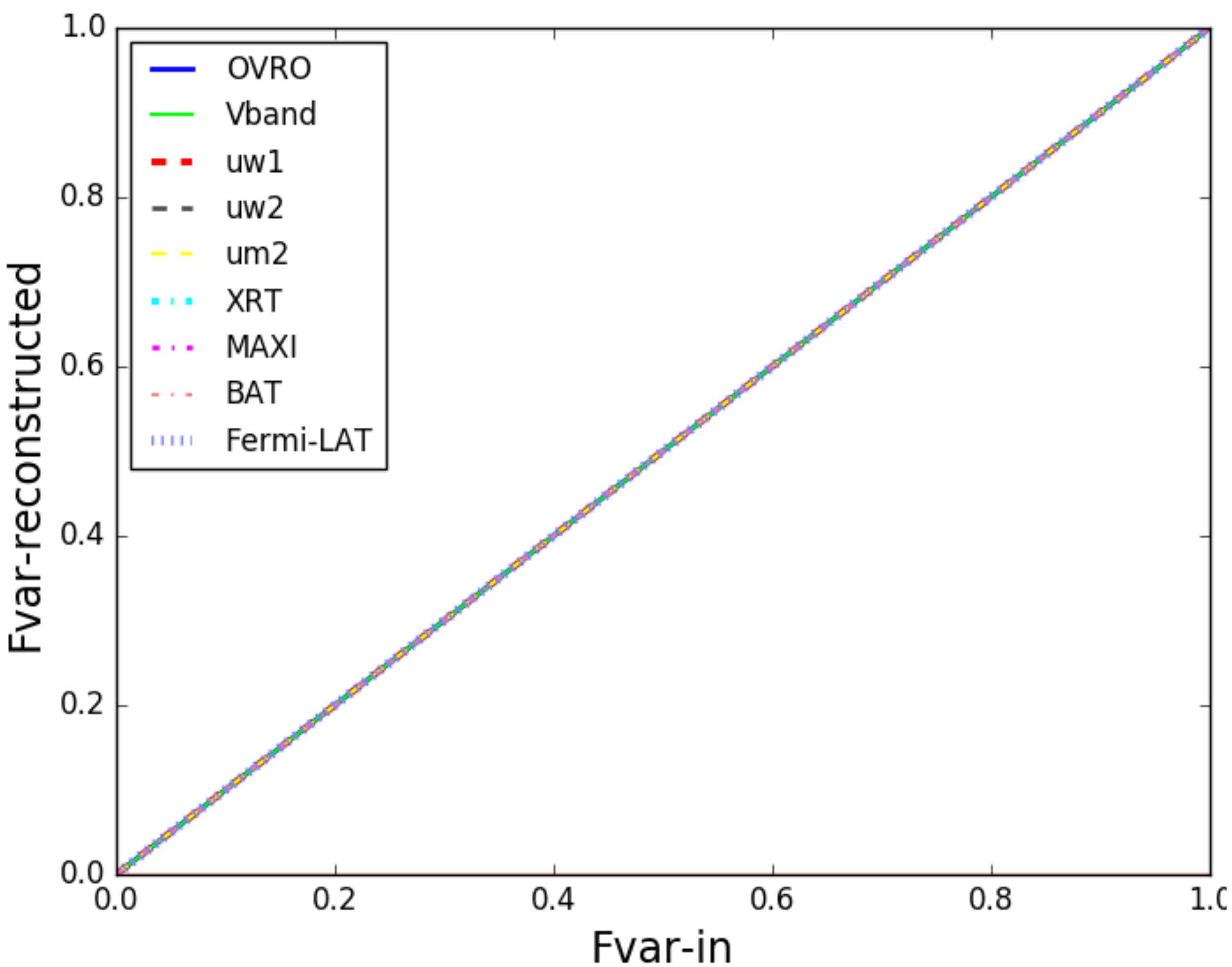
Observed Cadence



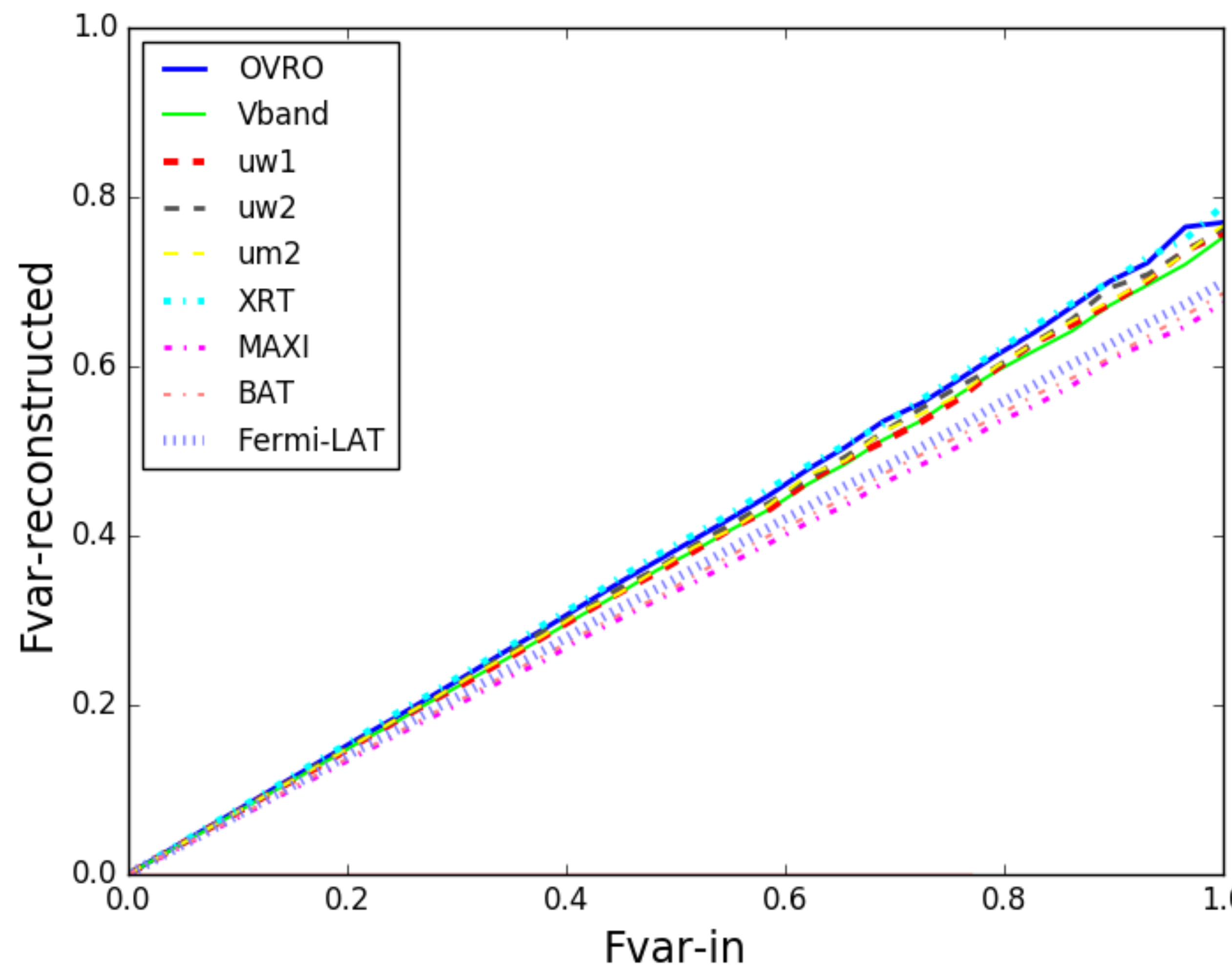
Fvar reconstruction

Pink Noise : Index = 1.0

“Ideal” Cadence



Observed Cadence

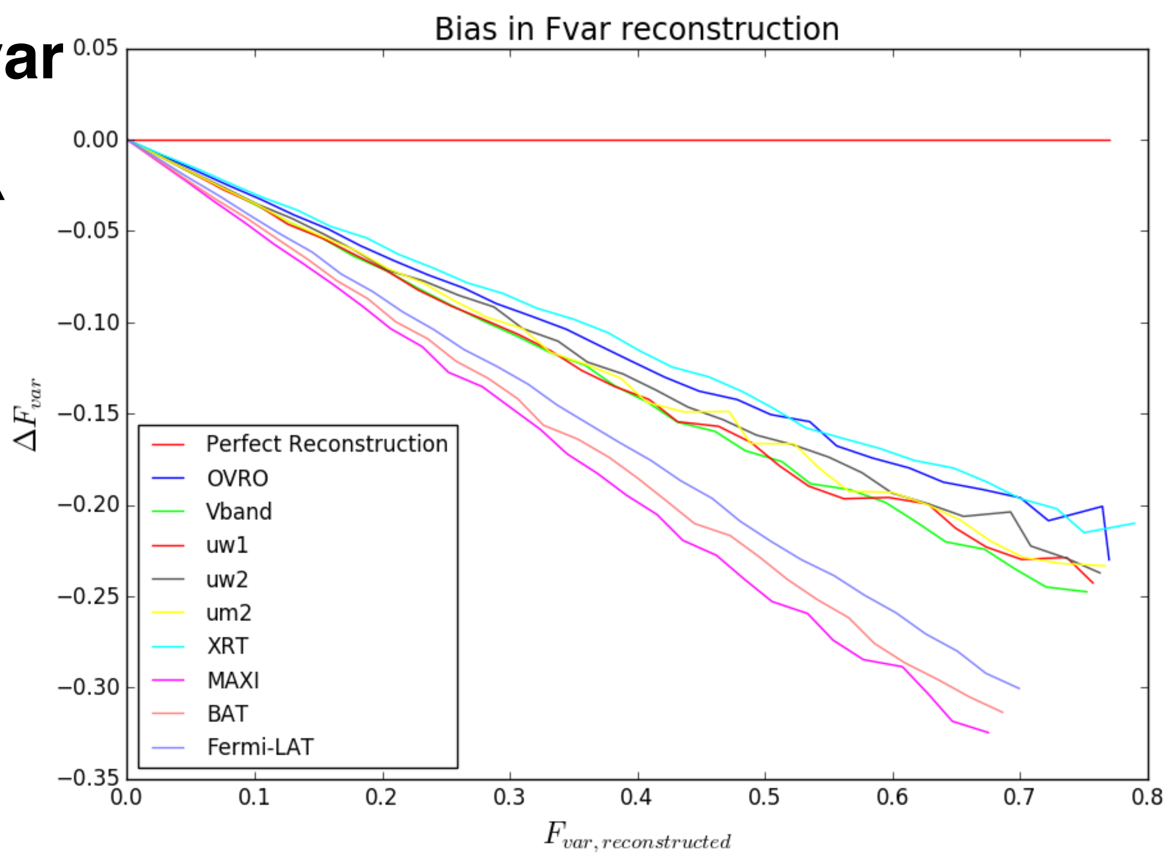


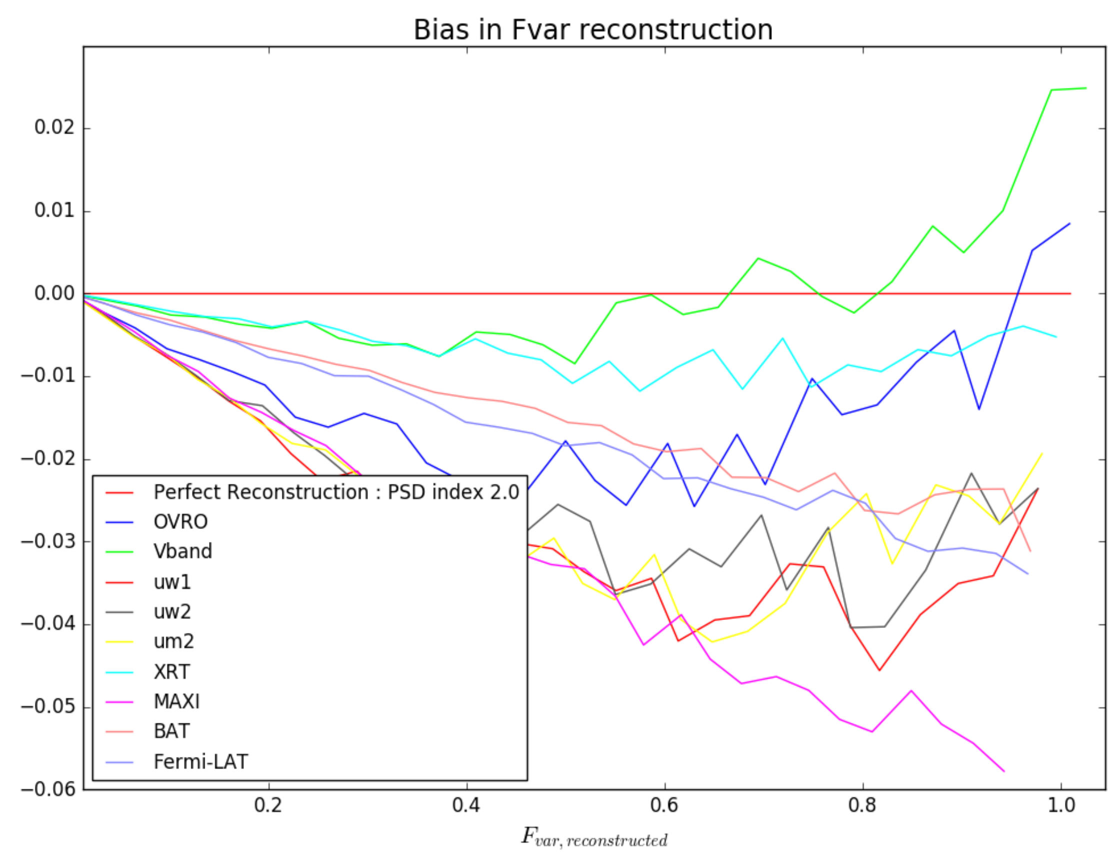
Fvar reconstruction

simulation PSD index = 1.0

simulation PSD index = 2.0

Δ Fvar

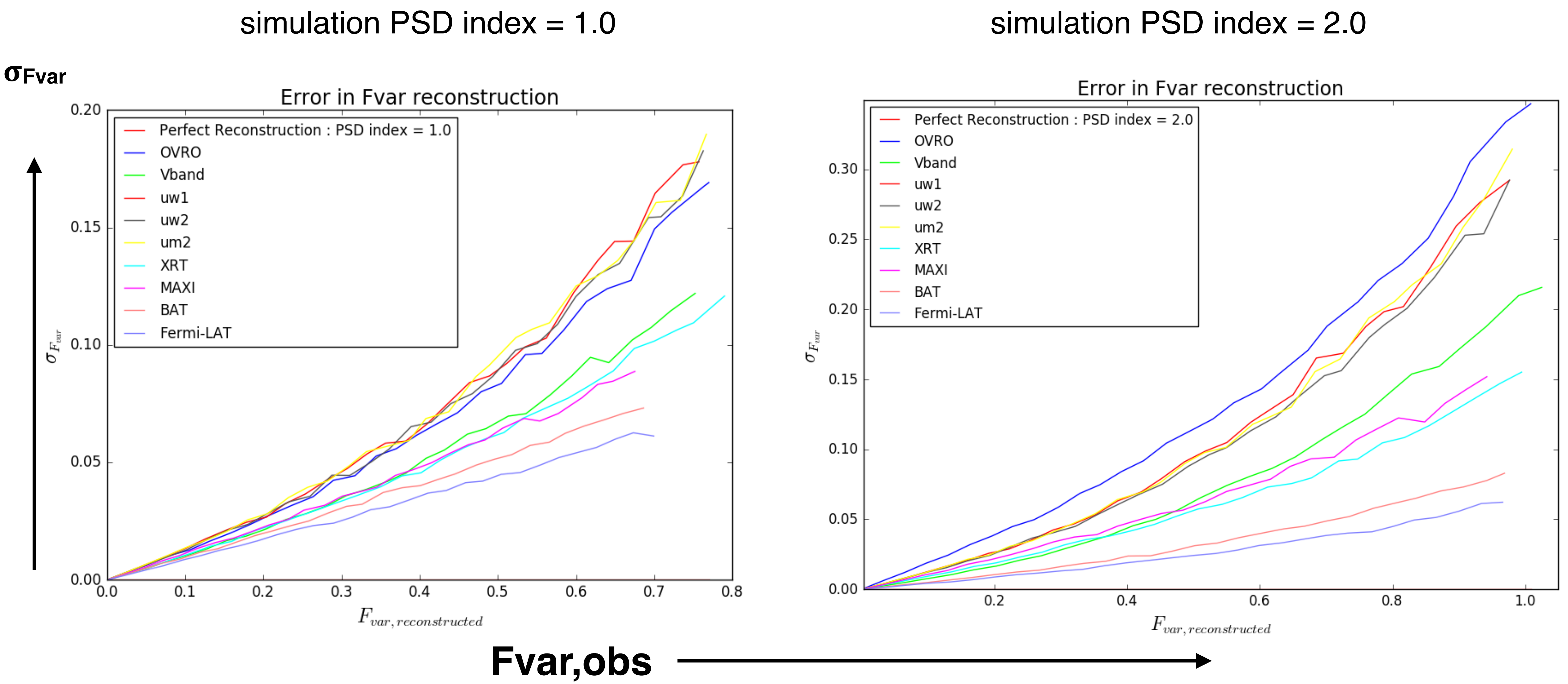




Fvar,obs

- Bias due to observational effects - larger for harder PSD (brown vs red) -> sampling effects ?
- Relatively, best reconstructions for finely sampled, least “gapped” (OVRO, Vband)
- Length of observational window less important for long enough durations and slow variations

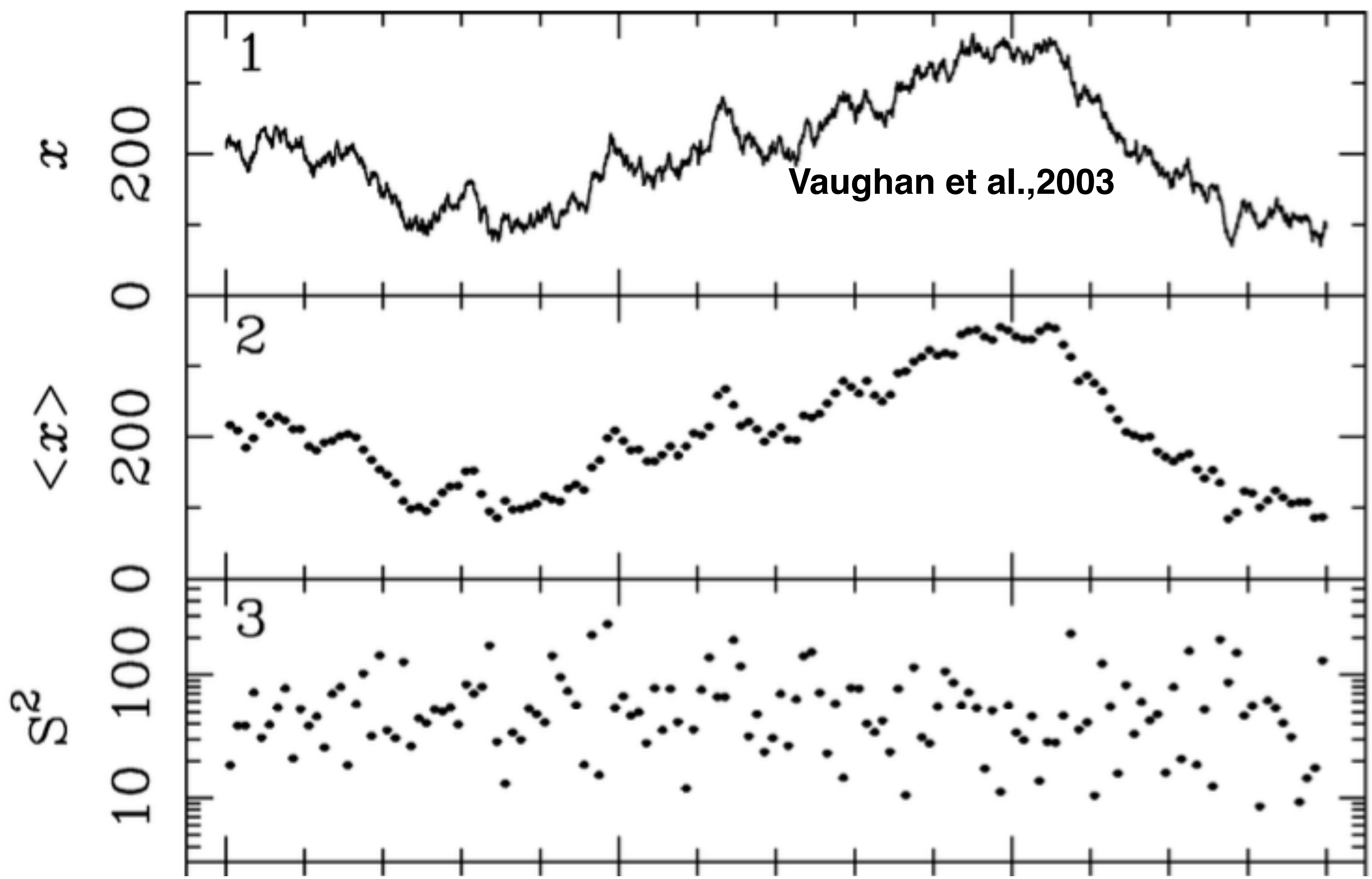
Fvar reconstruction….



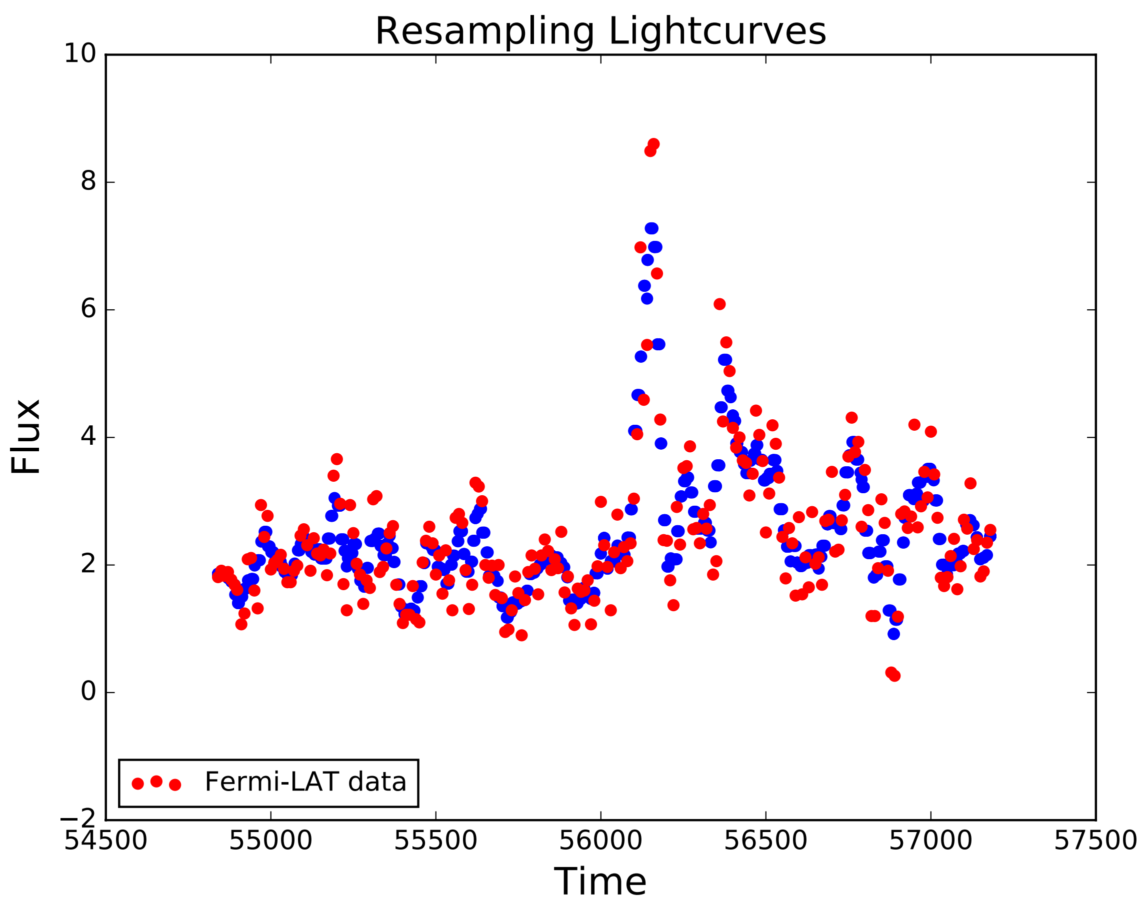
- Uncertainty in Fvar comparable for brown vs red noise

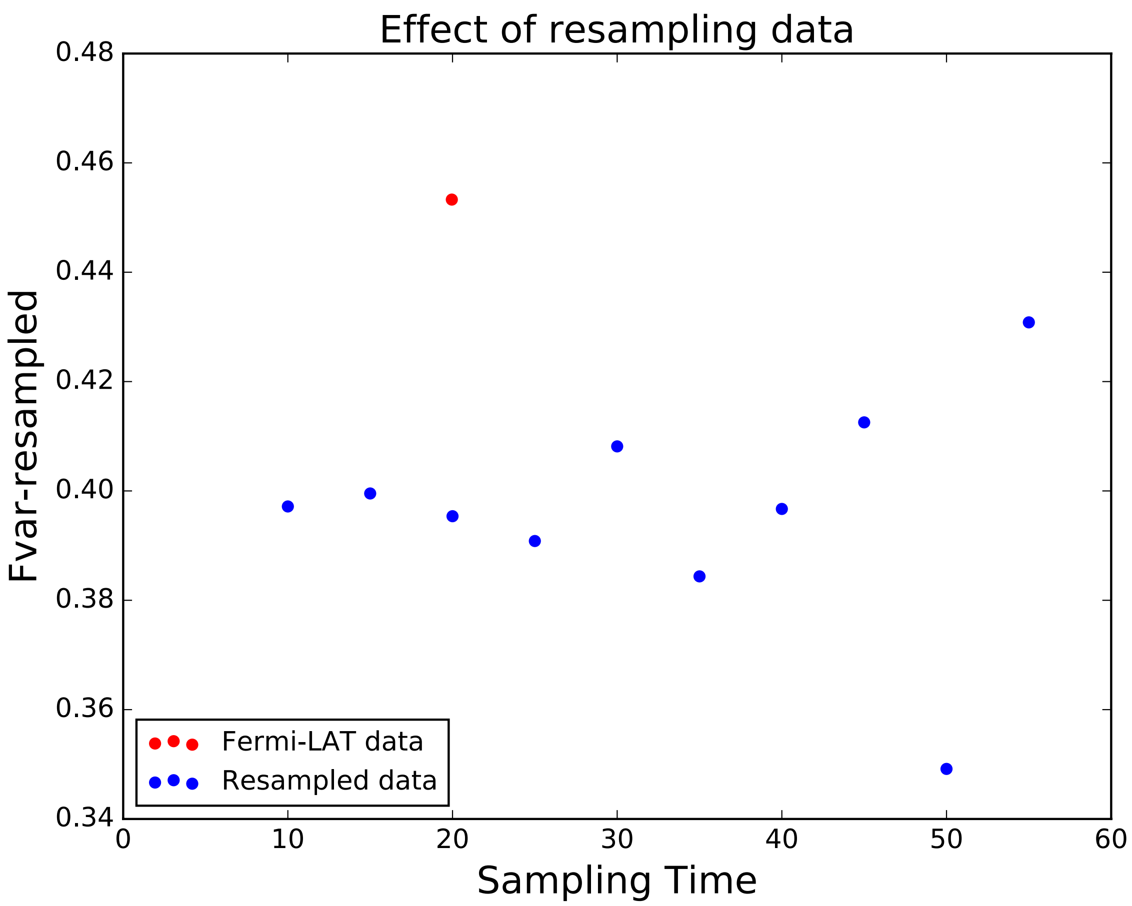
- Relative uncertainty larger for longer wavelengths - larger dispersion (**σ** does not include flux errors)

Backup

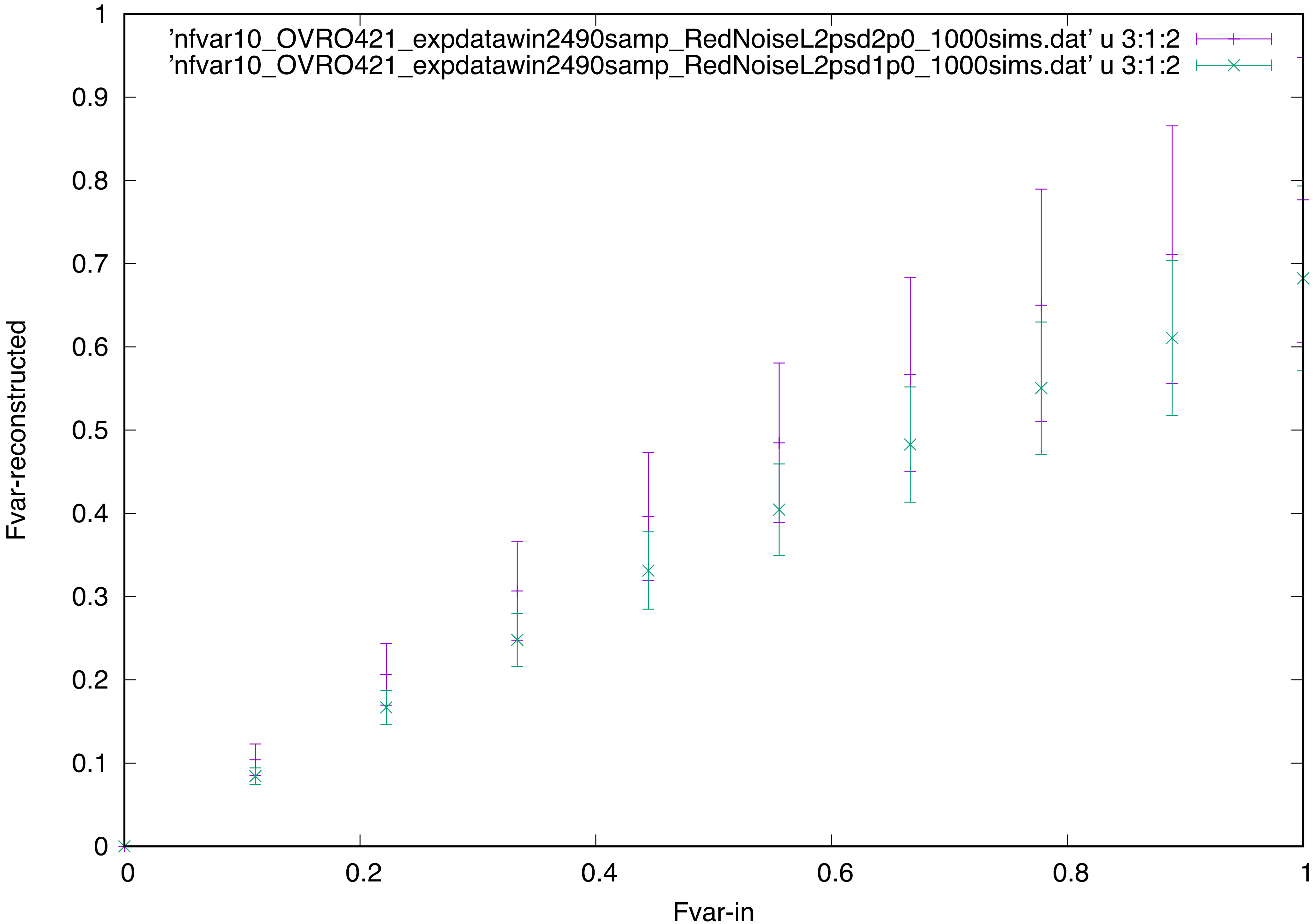


Simple Resampling Effects



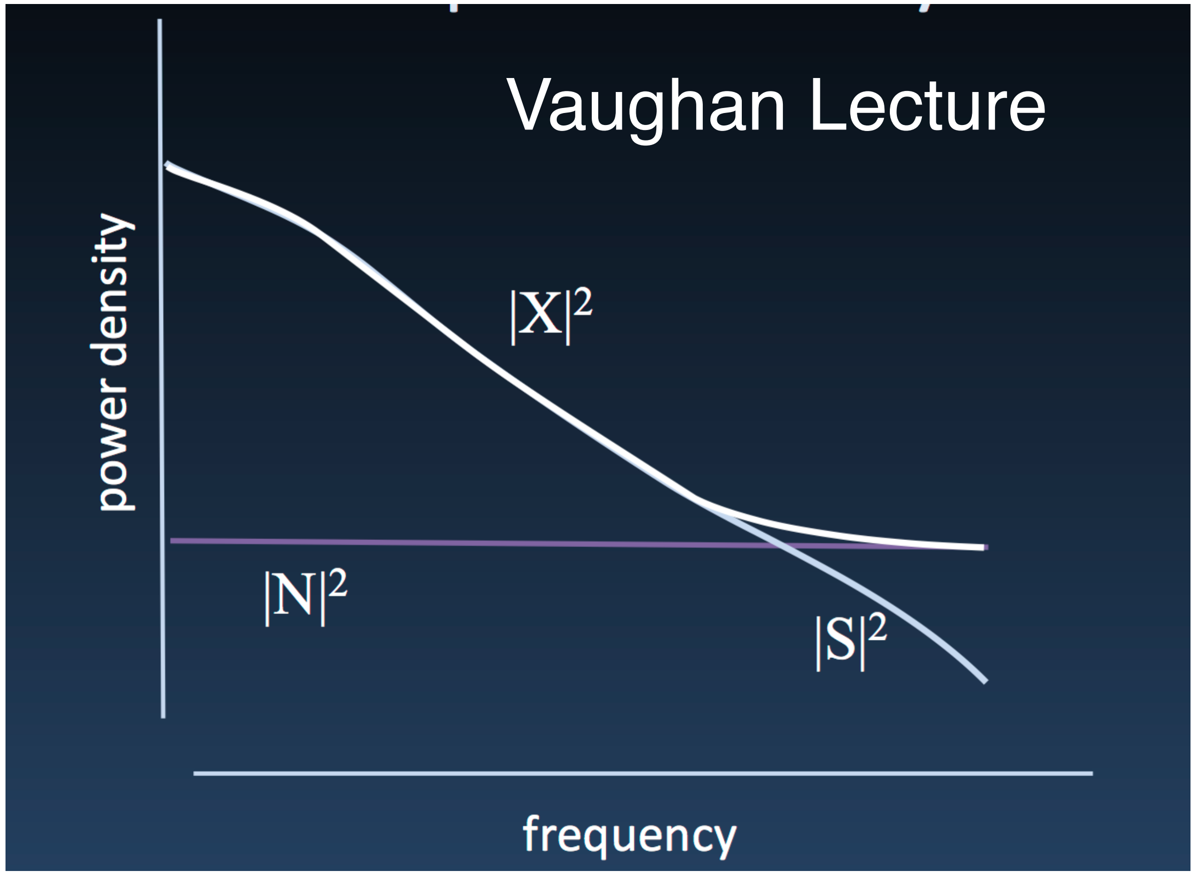


Non-gaussian PDF



Power Spectral Density

- Power spectral density or PSD is the "distribution of timescales"
- Frequency <-> timescales
- Time : $x = s + n$
  Fourier : $X = S + N$
  $|X|^2 = |S|^2 + |N|^2 + \text{Cross}$
  $PSD(f) = \langle |S|^2 \rangle = \langle |X|^2 \rangle - \langle |N|^2 \rangle$
  <=> Related to the variance
- Formally (for AGNs and others)
  Time : Lightcurve(t) = Dynamical(t) x Acceleration(t) x Radiation(t) x Observation(t) [**Product**]
- Fourier : Lightcurve(f) = Dynamical(f) * Acceleration(f) * Radiation(f) * Observation(f) [**Convolution**]

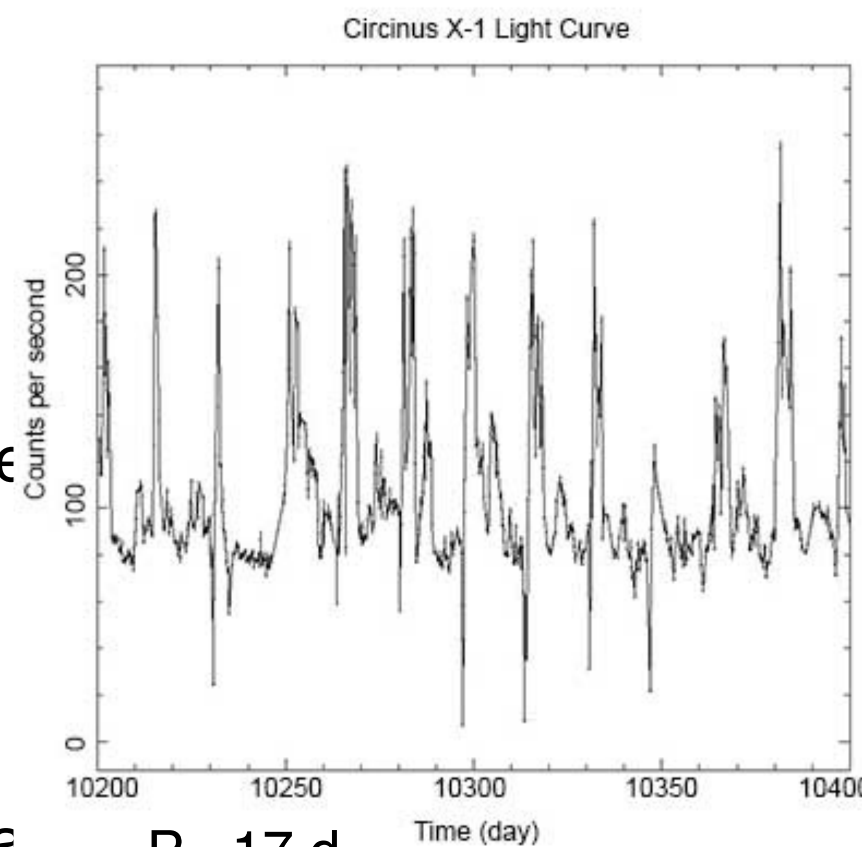


- Dynamical -> Periodic, slow variations
  Acceleration -> Stochastic / Shocks
  (Sironi et al., 2015, Giacche and Chakraborty, in progress)
  Radiation-> (LC simulations <-> "Observables")
  Observation-> Potential (CTCs, others)

Types of lightcurves

- **Periodic** - differentiate deterministic from noisy background
- **Transient** - differentiate deterministic from noisy background
- **Stochastic** - noisy signal (from noisy background ?!)

Types of lightcurves



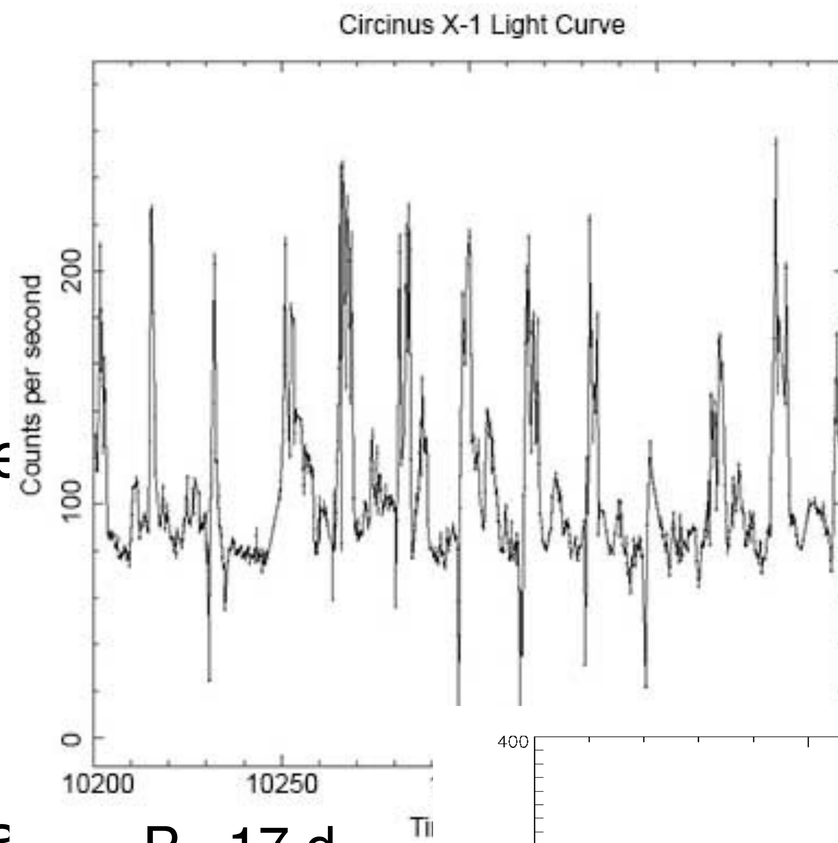
P ~17 d

http://imagine.gsfc.nasa.gov/science/toolbox/timing2.html

- **Periodic** - differentiate deterministic from noisy background
- **Transient** - differentia deterministic from noisy background
- **Stochastic** - noisy signal (from noisy background ?!)

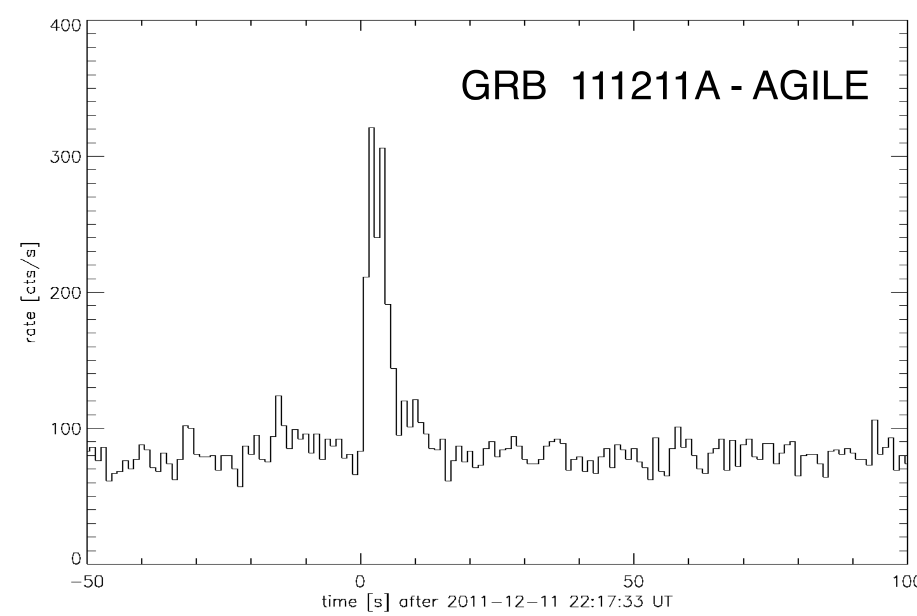
Types of lightcurves

- **Periodic** - differentiate deterministic from noisy background
- **Transient** - differentia deterministic from noisy background
- **Stochastic** - noisy signal (from noisy background ?!)



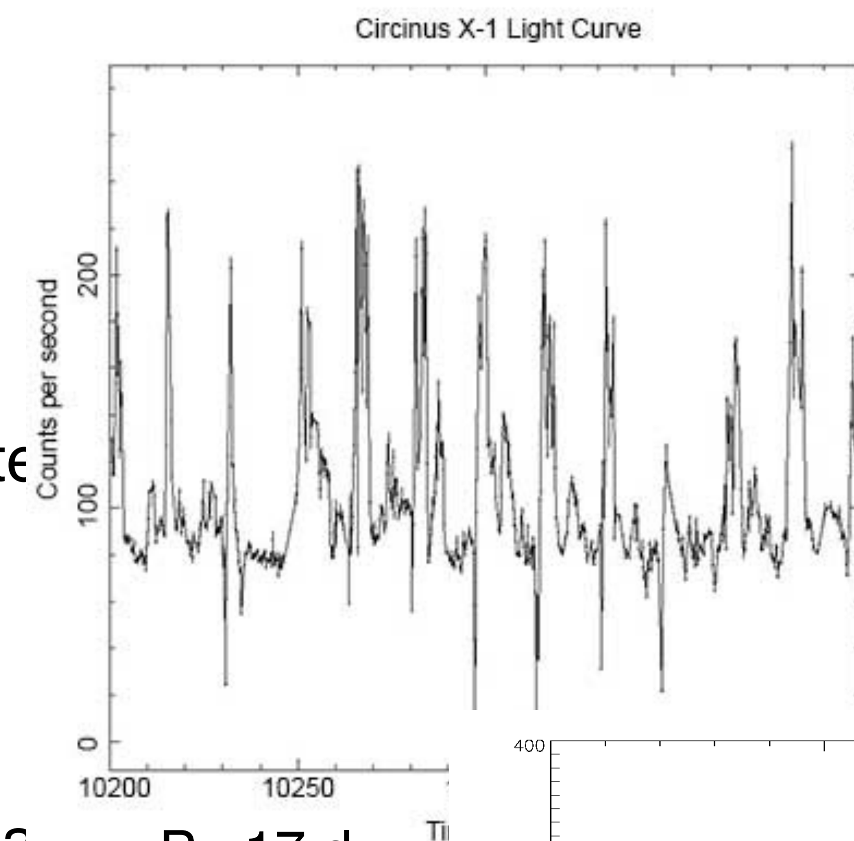
P ~17 d

http://imagine.gsfc.nasa.



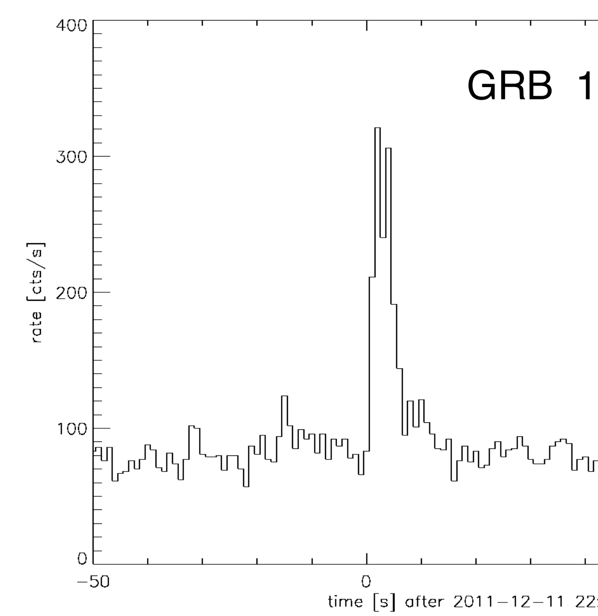
Types of lightcurves

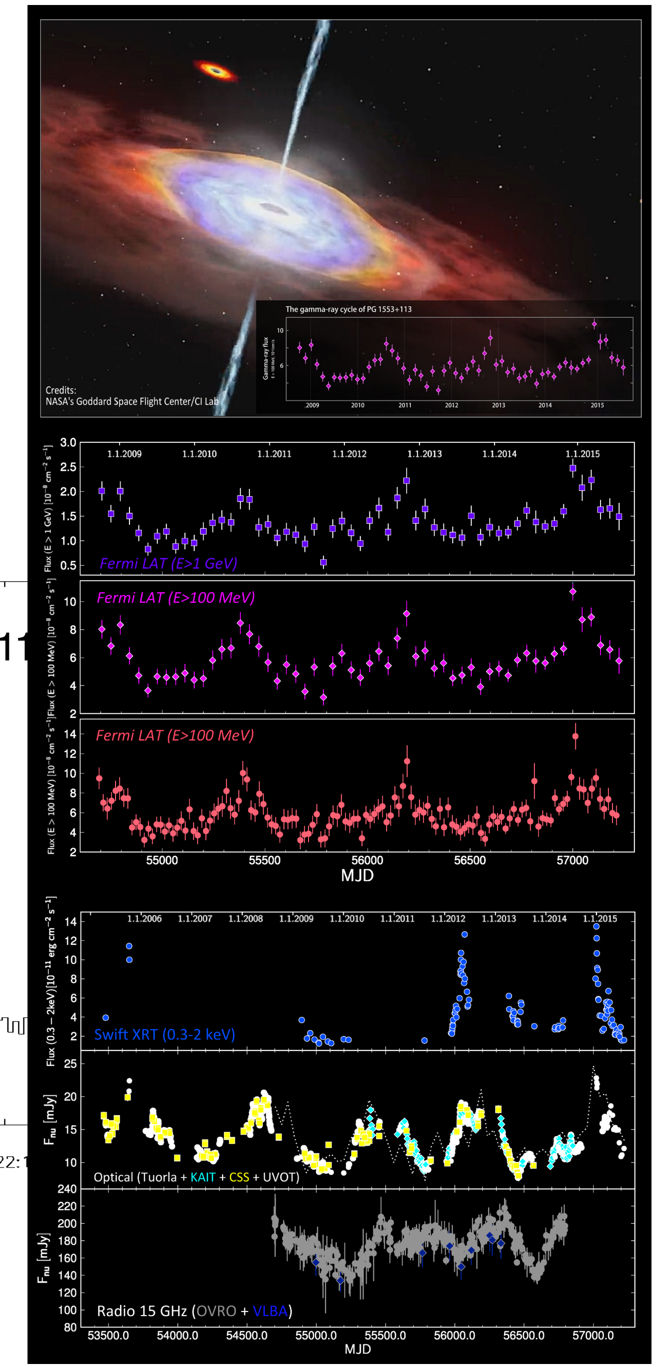
- **Periodic** - differentiate deterministic from noisy background
- **Transient** - differentia... deterministic from noisy background
- **Stochastic** - noisy signal (from noisy background ?!)



P ~17 d

http://imagine.gsfc.nasa.

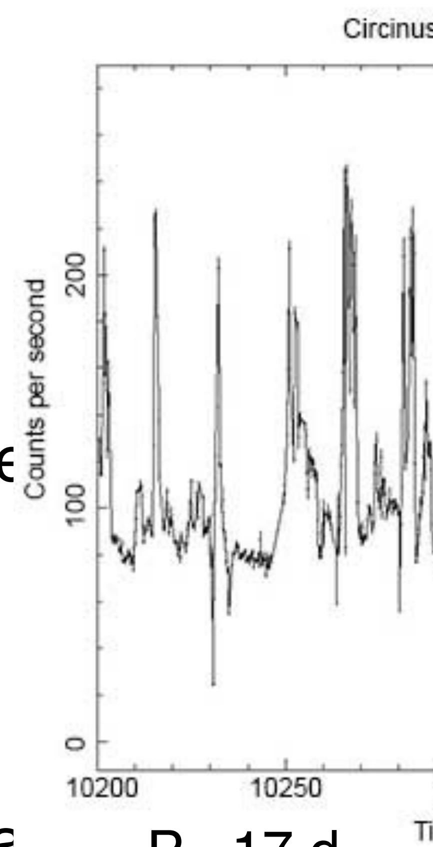




PG 1553+113

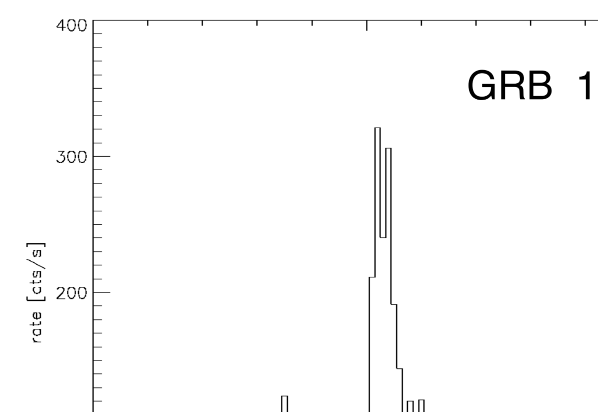
Types of lightcurves

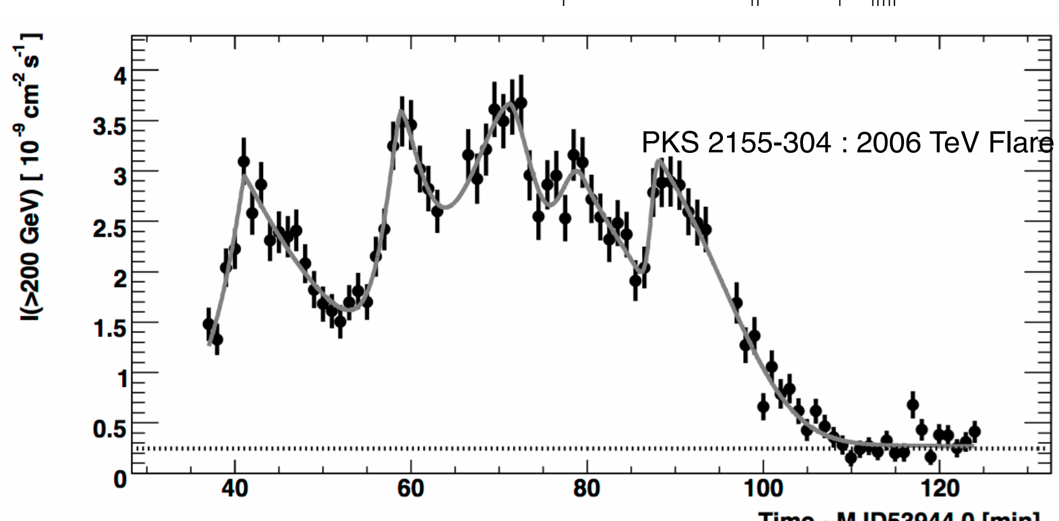
- **Periodic** - differentiate deterministic from noisy background
- **Transient** - differentiate deterministic from noisy background
- **Stochastic** - noisy (from noisy background ?

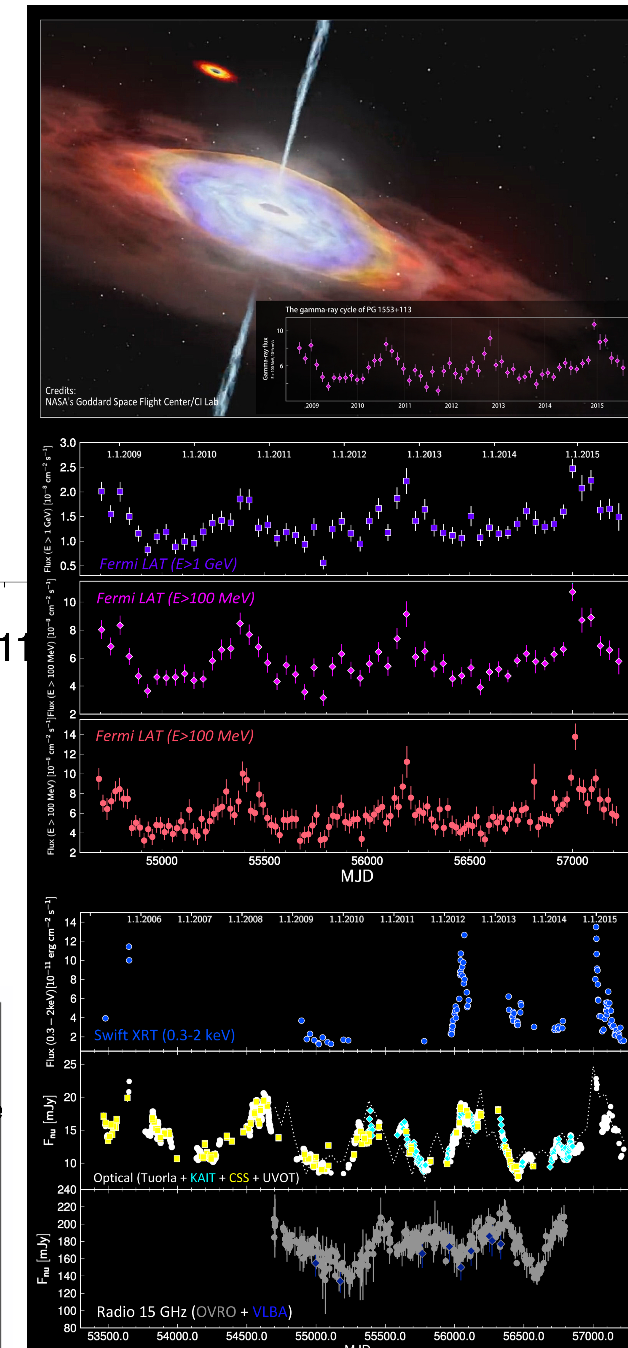


P ~17 d

http://imagine.gsfc.nasa.







PG 1553+113

Types of data analyses

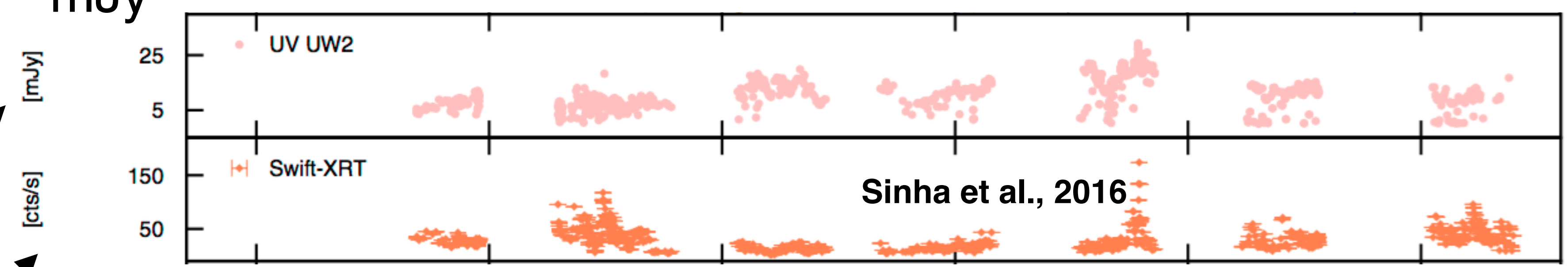
- Depending on the wavelength
  - (quasi-)continuous, binned signals or fluxes
  - discrete : time tagged events
  - discrete : counts per time bins

- Naturally analyses methods are also different(ly used)

- Temporal vs Fourier Analyses ; mixed

Types of data analyses

mJy



cts/sec

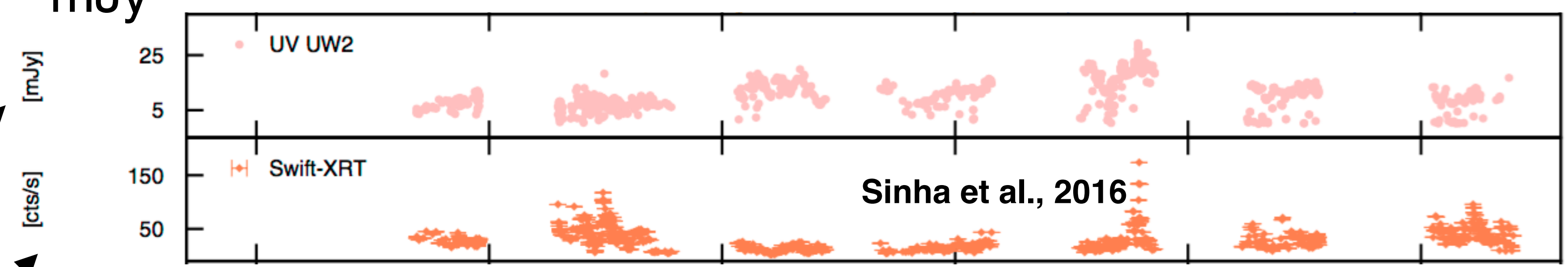
- Depending on the wavelength
  - (quasi-)continuous, binned signals or fluxes
  - discrete : time tagged events
  - discrete : counts per time bins

- Naturally analyses methods are also different(ly used)

- Temporal vs Fourier Analyses ; mixed

Types of data analyses

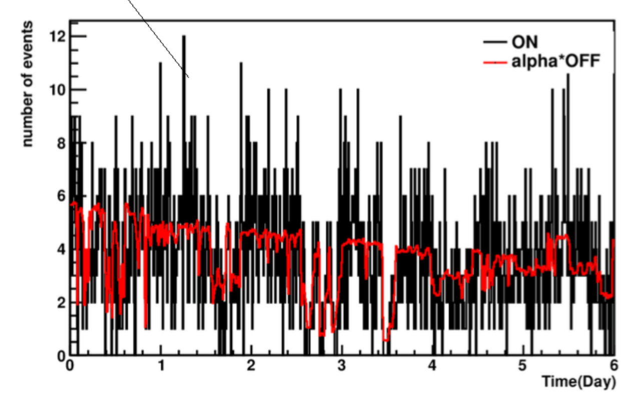
mJy

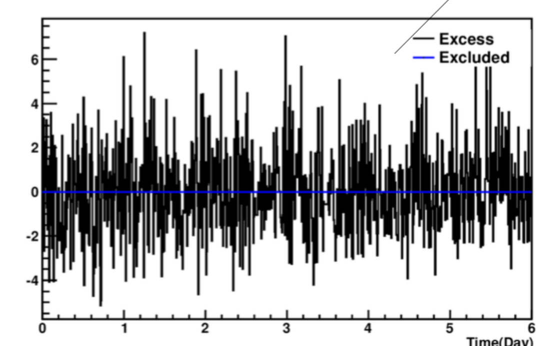


cts/sec

ON/OFF events

Excess





Helen Poon : GC Analysis

- Depending on the wavelength
  - (quasi-)continuous, binned signals or fluxes
  - discrete : time tagged events
  - discrete : counts per time bins

- Naturally analyses methods are also different(ly used)

- Temporal vs Fourier Analyses ; mixed

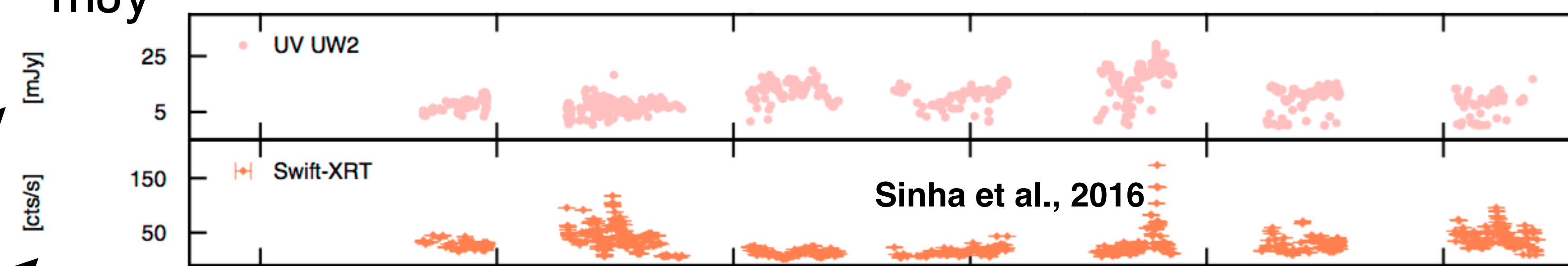
Types of data analyses

- Depending on the wavelength
  - (quasi-)continuous, binned signals or fluxes
  - discrete : time tagged events
  - discrete : counts per time bins

- Naturally analyses methods are also different(ly used)

- Temporal vs Fourier Analyses ; mixed

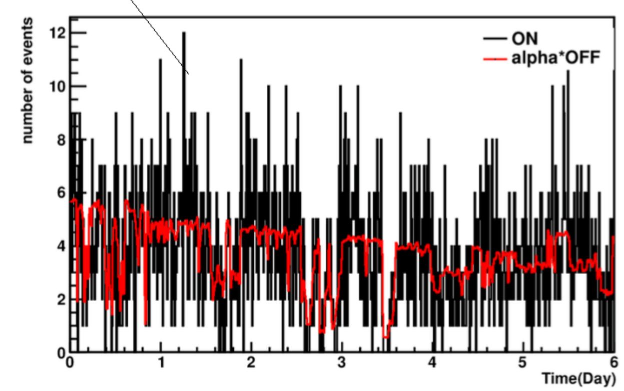
mJy

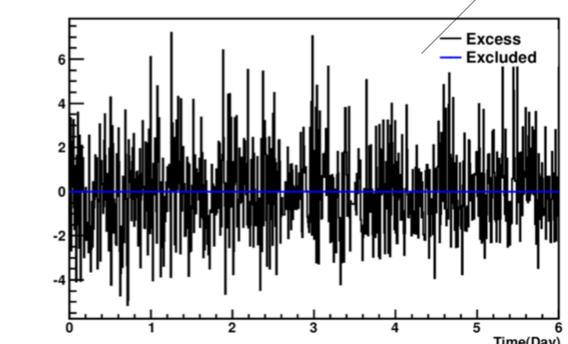


cts/sec

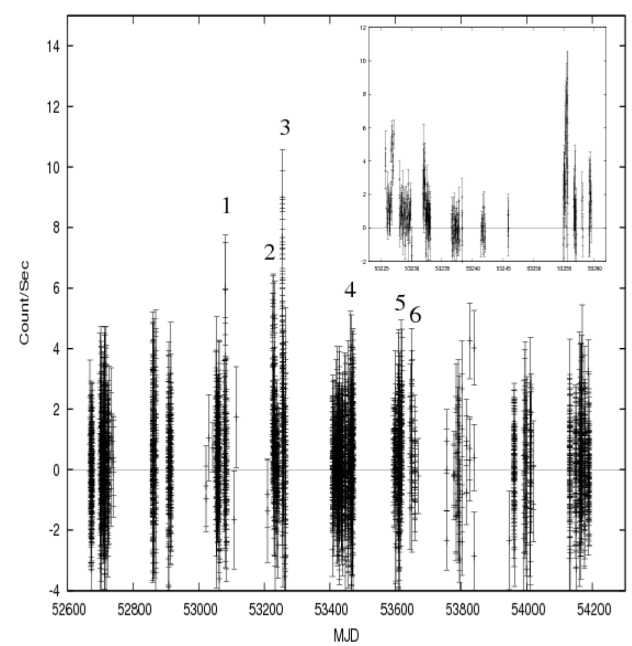
ON/OFF events

Excess





Helen Poon : GC Analysis



IBIS/ISGRI long term LC (18-60\,keV) of IGR J16465--4507

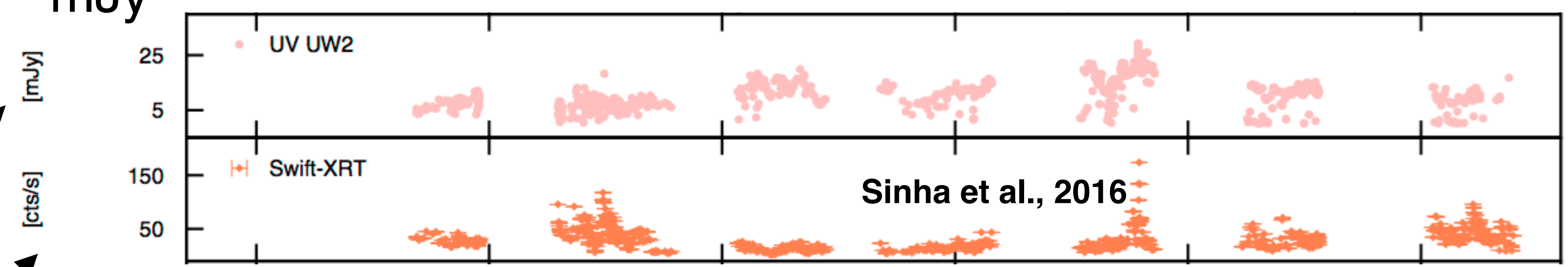
Types of data analyses

- Depending on the wavelength
  - (quasi-)continuous, binned signals or fluxes
  - discrete : time tagged events
  - discrete : counts per time bins

- Naturally analyses methods are also different(ly used)

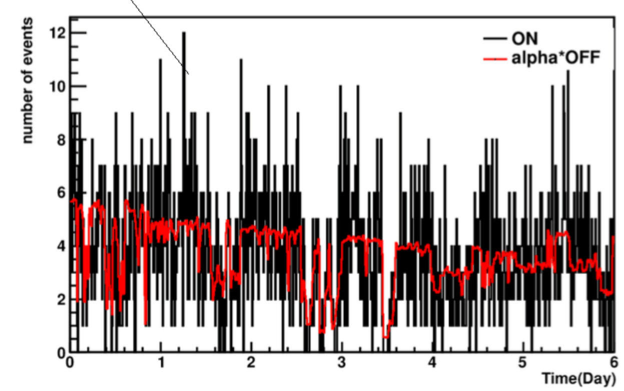
- Temporal vs Fourier Analyses ; mixed

mJy

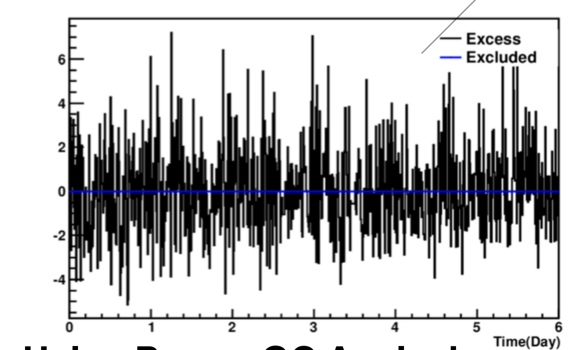


cts/sec

ON/OFF events

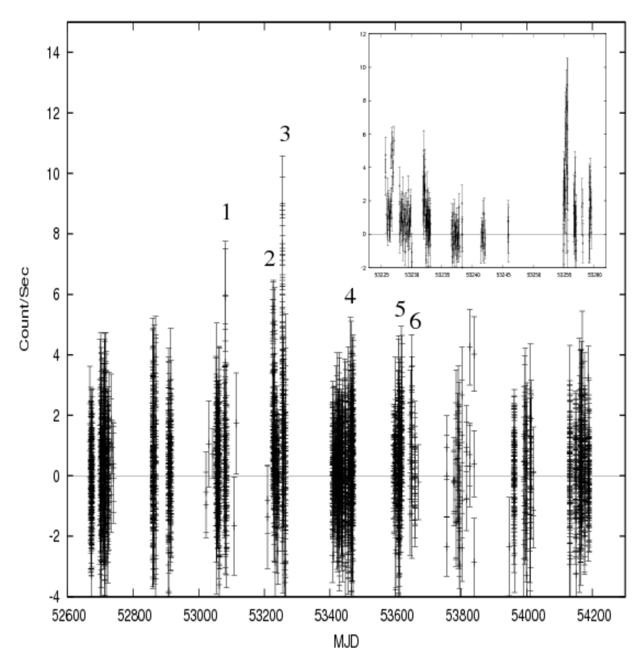


Excess

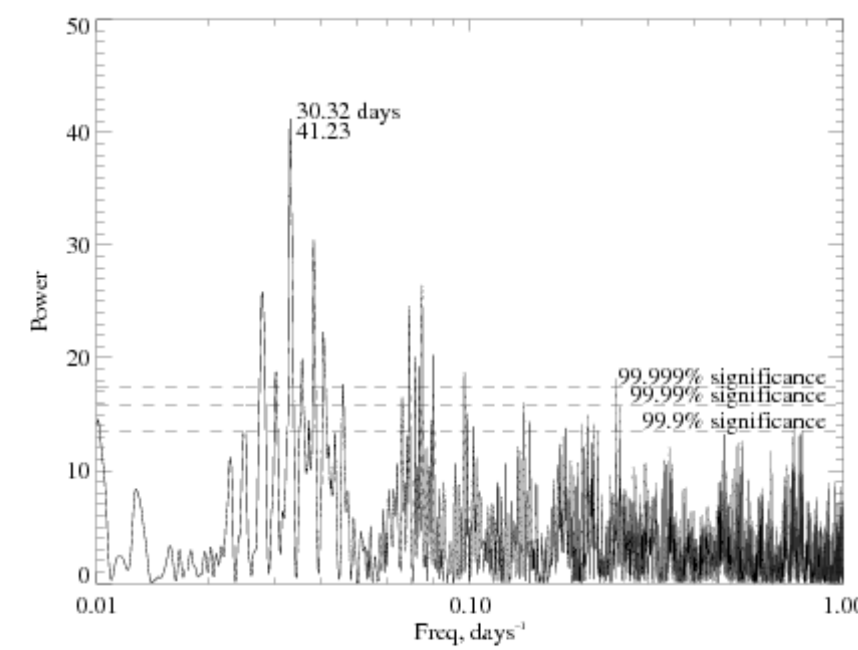


Helen Poon : GC Analysis

Time vs Frequency domain



<=>



IBIS/ISGRI long term LC (18-60\,keV) of IGR J16465--4507

CPF2014

$$N_{\rm f,det} = \int_{\rm F_{sens}/\eta_{flare,eff}}^{\infty} dF \frac{dN}{dF} \propto \left(\frac{F_{\rm sens}/\eta_{\rm flare,eff}}{F_{\rm b}} \right)^{1-\beta_1} . \quad (15)$$