Astro-Statistics:
What is it good for?

Aneta Siemiginowska and Vinay Kashyap
Harvard-Smithsonian Center for Astrophysics
ICHASC

Astro-Statistics

DvD: “Statistics applied to Astronomy”

Astro perspective: develop algorithms to infer astronomical truth
Keep astronomers from fooling themselves

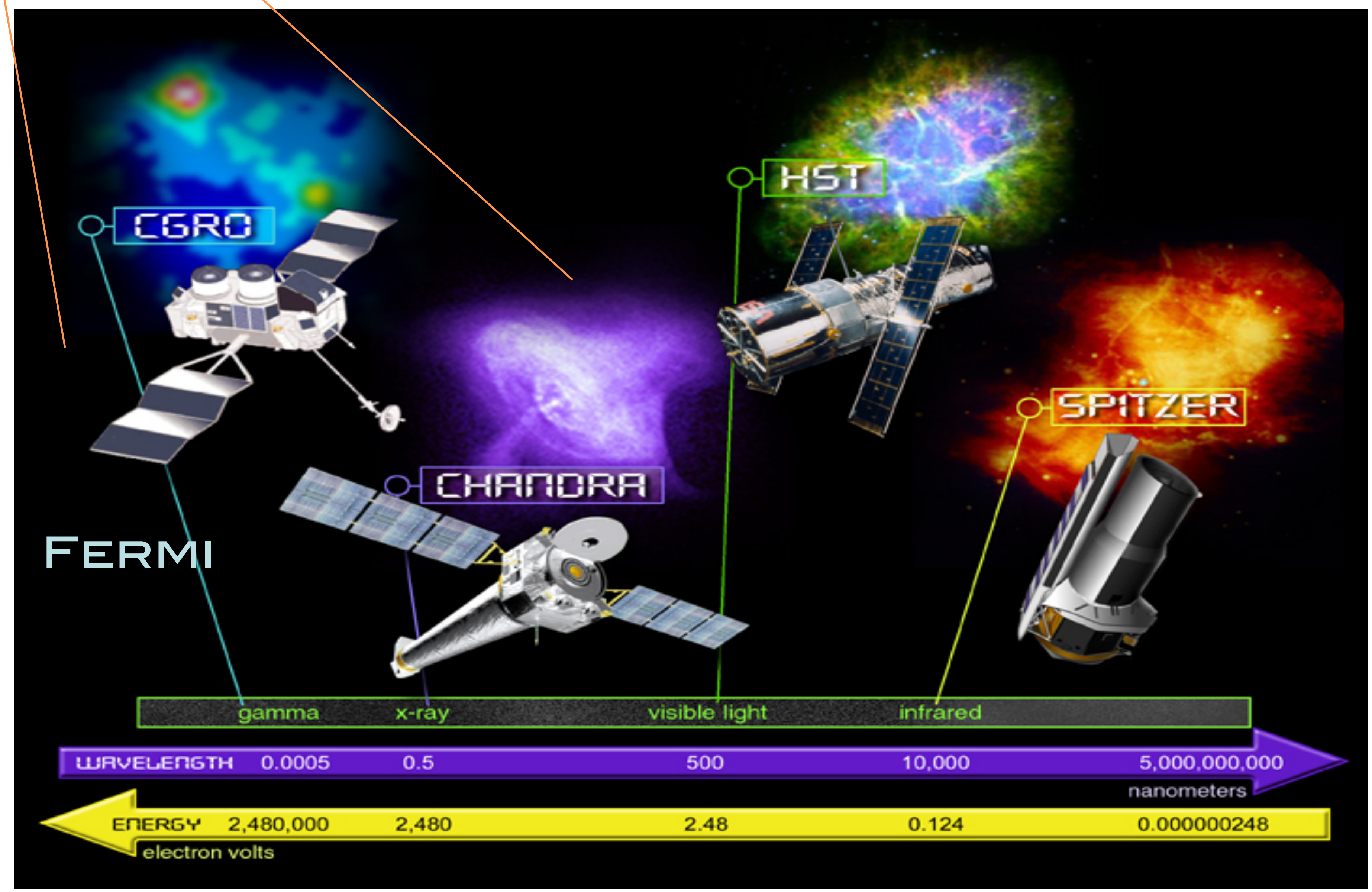
Astronomical data are generally cleaner – there is less uncorrectable bias
Loads of BIG data
Some unique circumstances like well-defined calibration
One-shot experiments that require Bayesian analyses
High-energy datasets doubly simpler, being recordings of a Poisson point process

Google

Astronomy

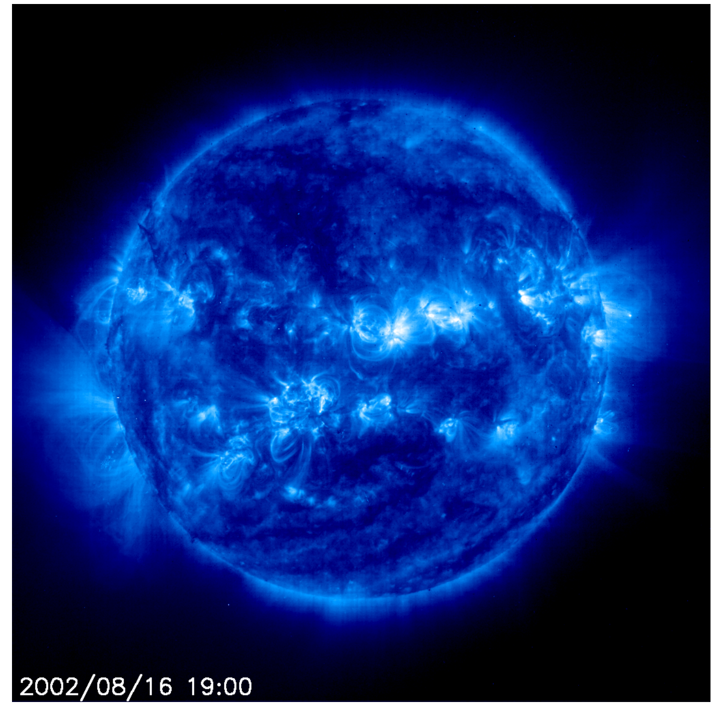
High-Energy Astrophysics

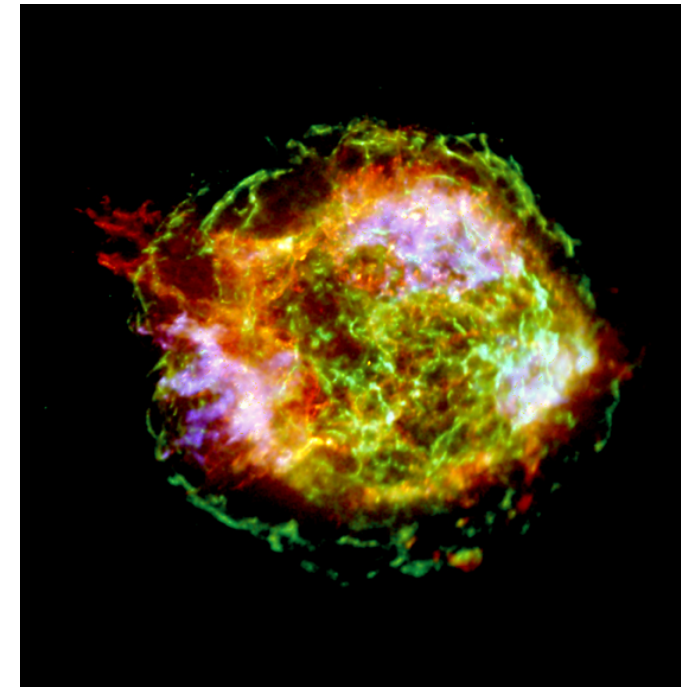
X-rays and Gamma-rays < 10^{-6} cm or > 2×10^{16} Hz
Not visible from the ground - Space-based observations

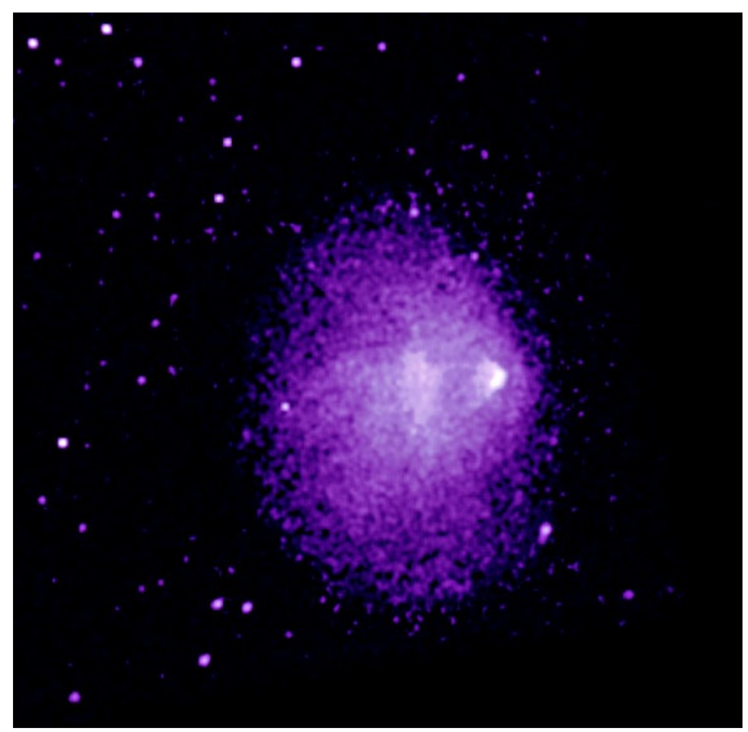


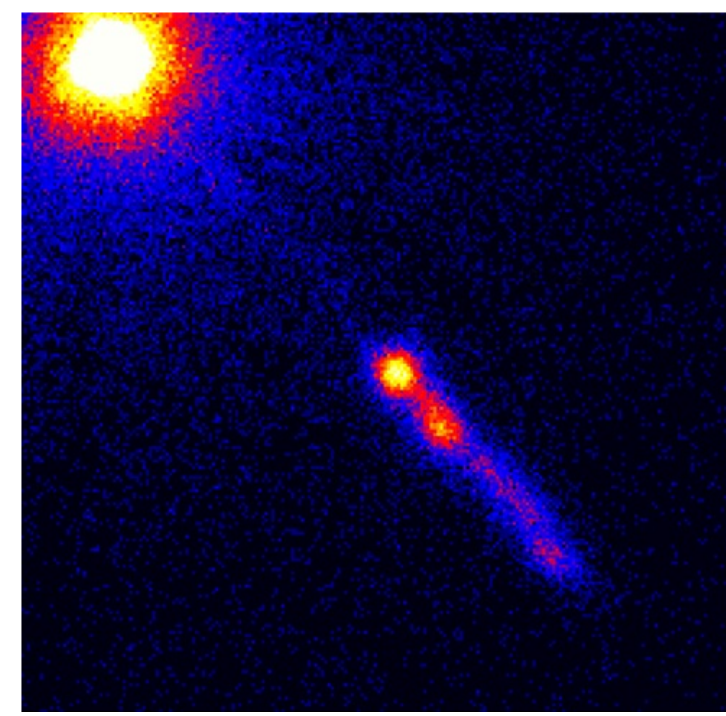
Sources of High-Energy Radiation

- Stellar Coronae
- Supernova remnants
- Galactic outflows
- Clusters of galaxies
- Compact objects:
  neutron stars,
  accreting black holes,
  supermassive black holes
- Relativistic jets
- GRBs
- etc…



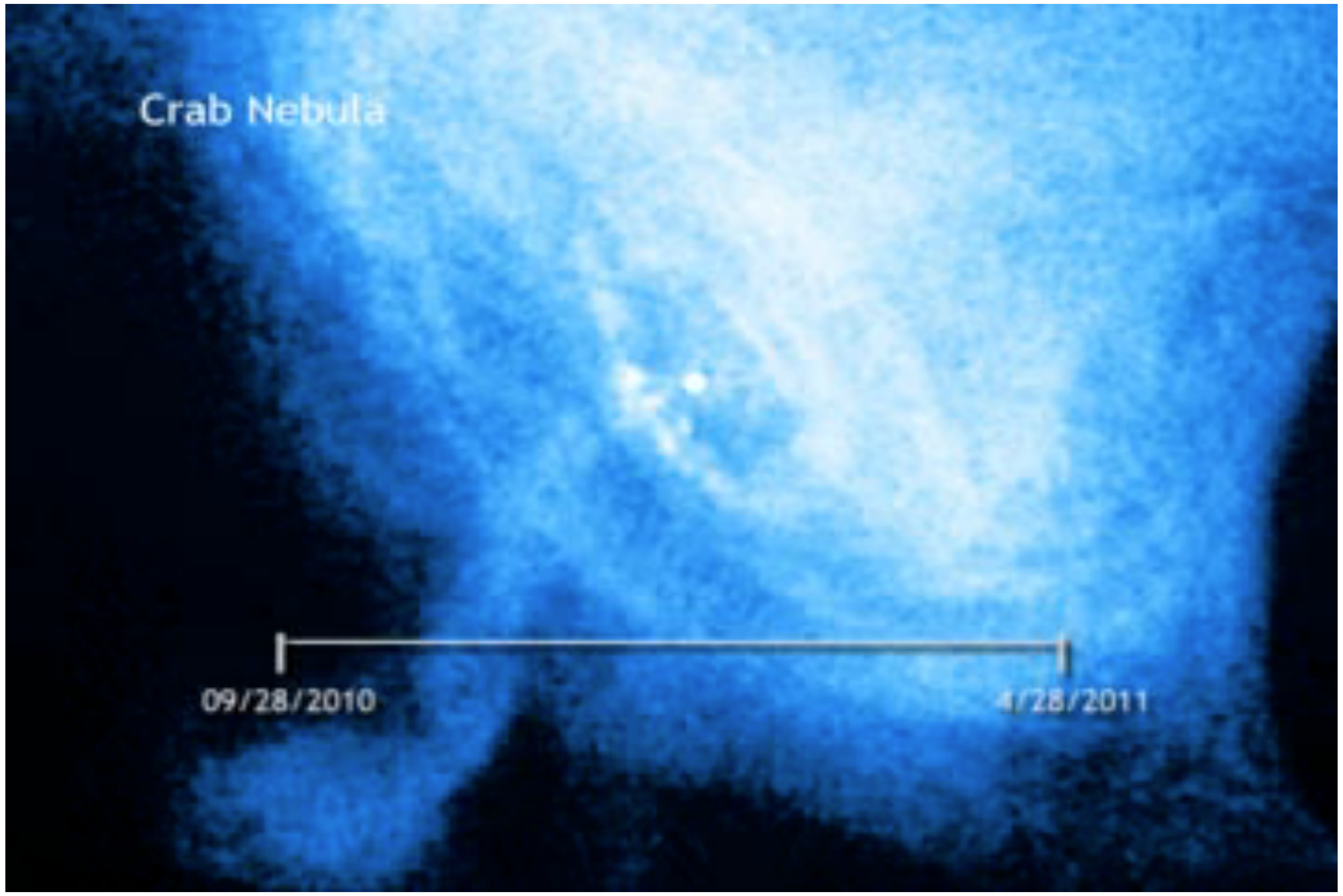






Data in High-Energy Astrophysics

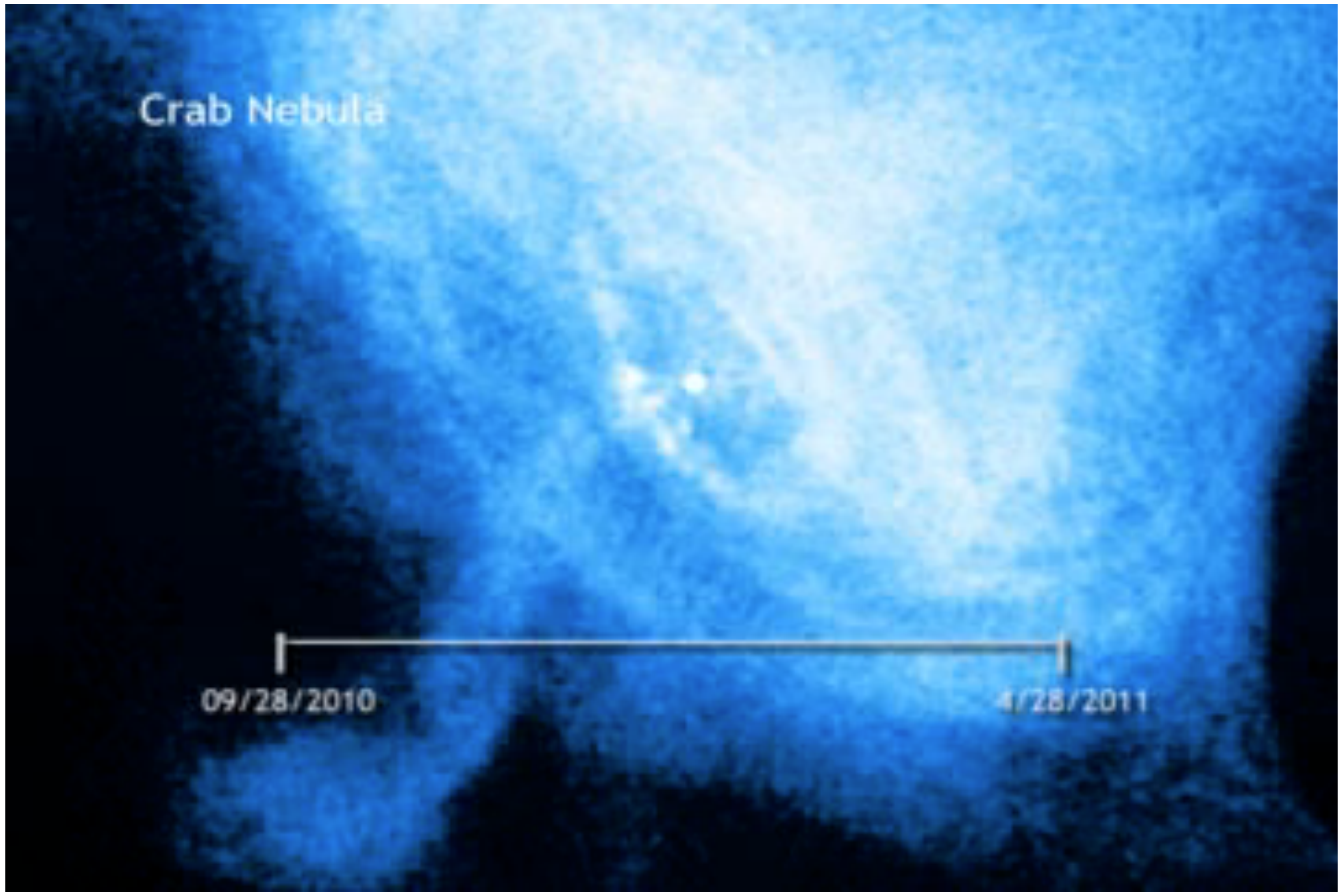
- X-ray and γ-ray data count photons => Poisson in nature
- Complex physics and data collection
- Data may exhibit Spectral, Temporal and Spatial variations



Crab Nebula - variations during 6 month of snap-shot observations with Chandra X-ray Observatory

Data in High-Energy Astrophysics

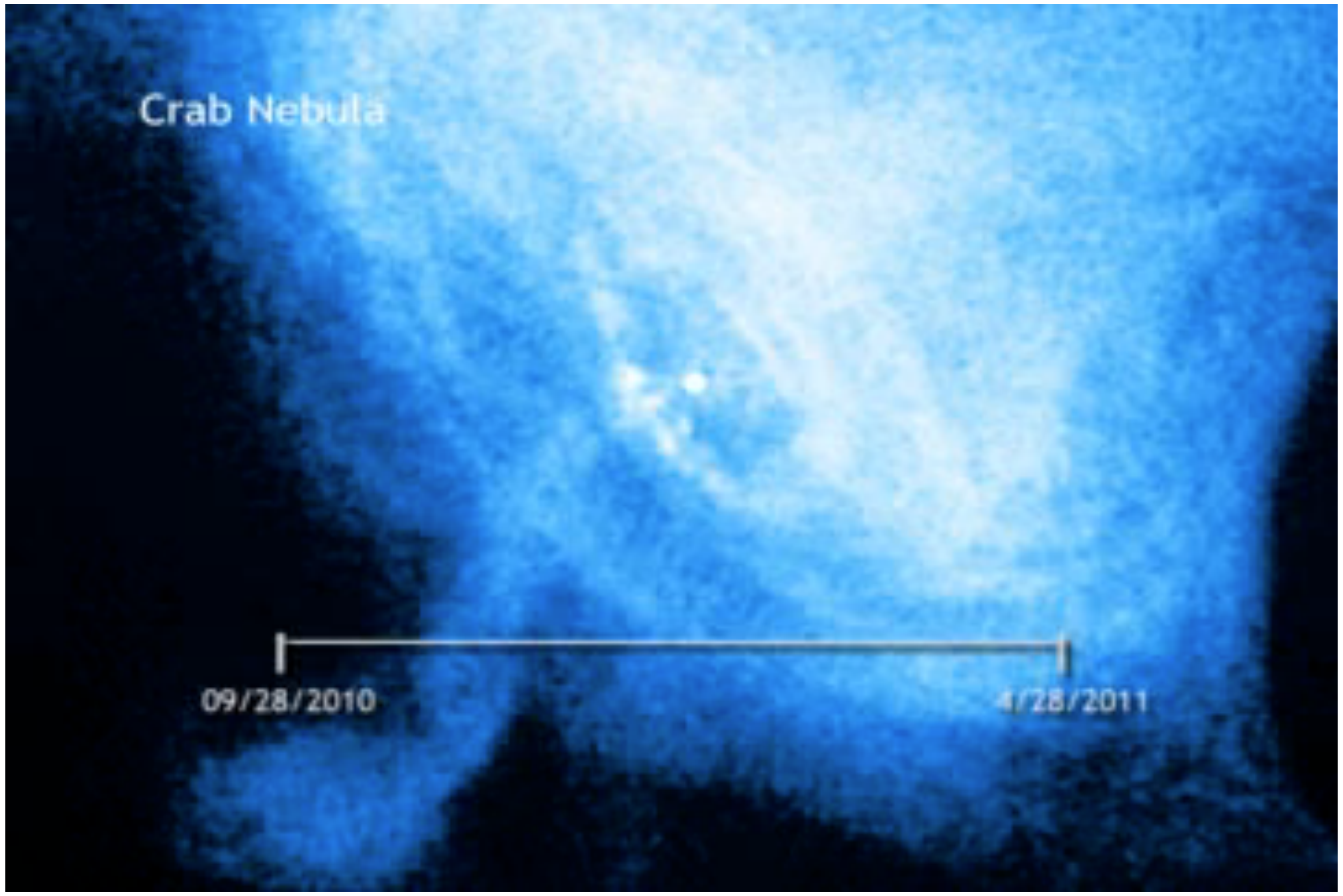
- X-ray and γ-ray data count photons => Poisson in nature
- Complex physics and data collection
- Data may exhibit Spectral, Temporal and Spatial variations



Crab Nebula - variations during 6 month of snap-shot observations with Chandra X-ray Observatory

Data in High-Energy Astrophysics

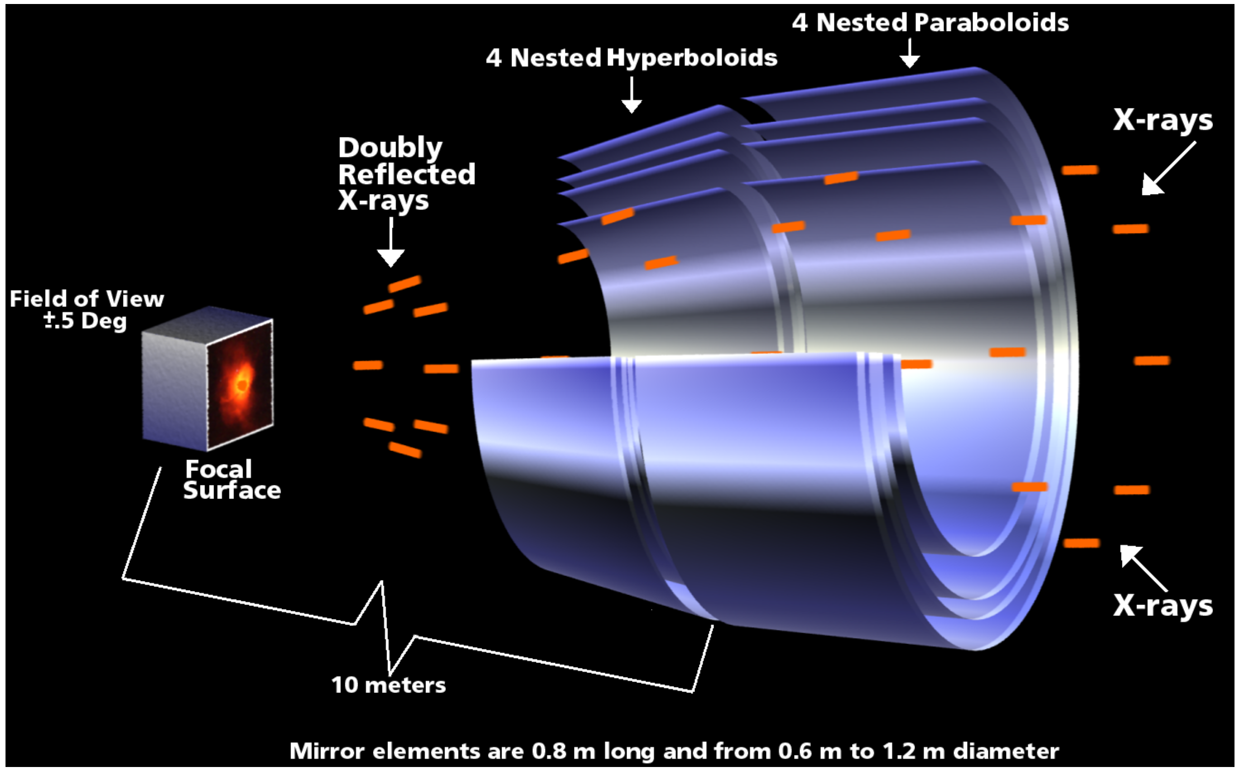
- X-ray and γ-ray data count photons => Poisson in nature
- Complex physics and data collection
- Data may exhibit Spectral, Temporal and Spatial variations

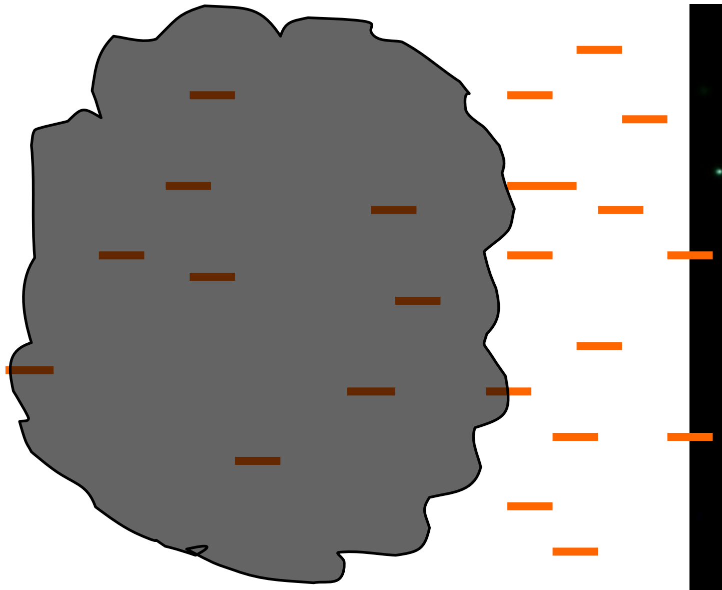


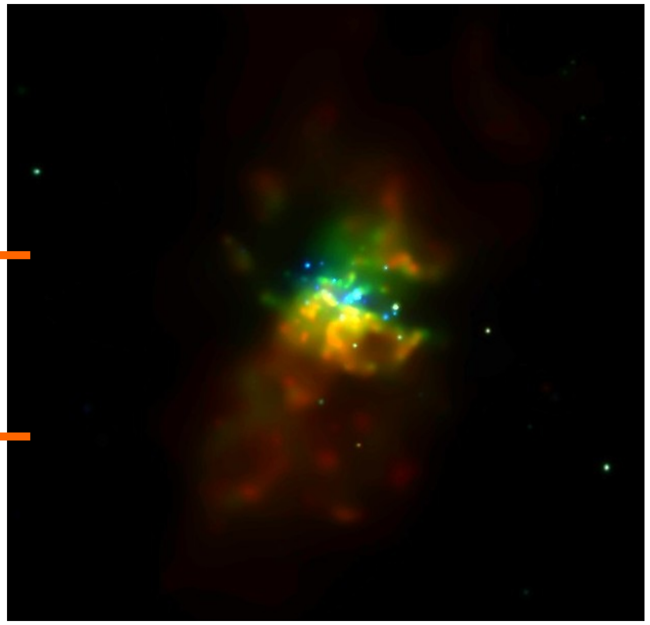
Crab Nebula - variations during 6 month of snap-shot observations with Chandra X-ray Observatory

Data Collection in Space

Chandra X-ray Observatory







Telescope + Detectors

Interstellar Medium

Astronomical Object

Measurement Process
Inefficient data collection Process
Instrument characteristics
Instrument Calibration

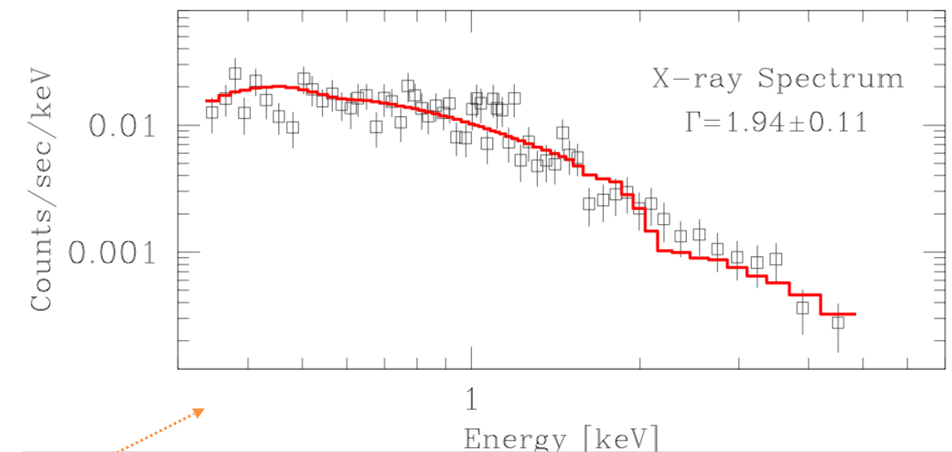
Loss of signal
but also imprints
information

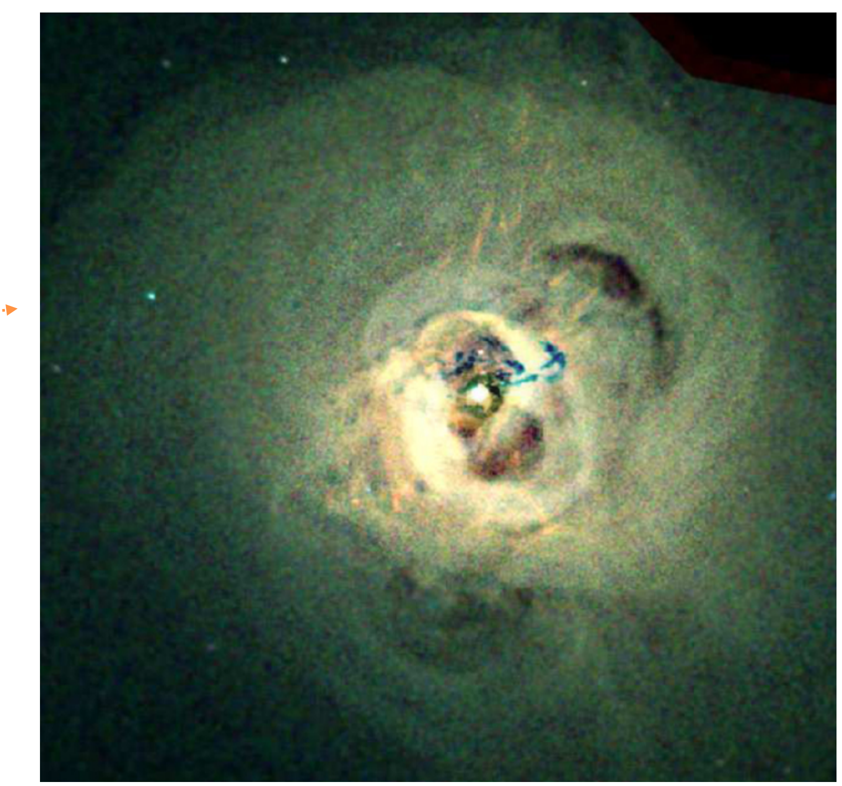
Physics

Data Collection

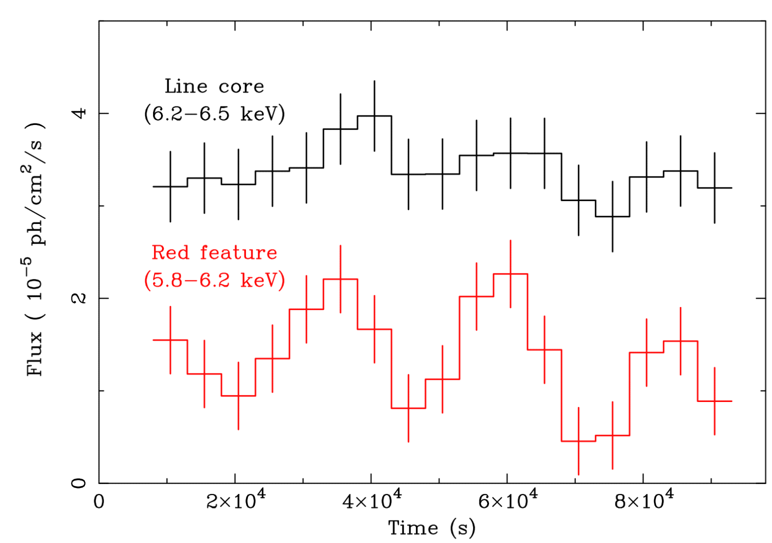
- Data are recorded for each arriving photon:
  - the (2-dimensional) location - sky coordinates
  - the photon energy
  - the arrival time
- All variables are discrete
  - High resolution -> finer discretization,
  - e.g., 4096 x 4096 spatial or up to 16384 spectral bins
- Table with photon counts for:
  - Spectral analysis - 1D
  - Spatial analysis - 2D
  - Timing analysis - 1D

Energy Spectra 1D





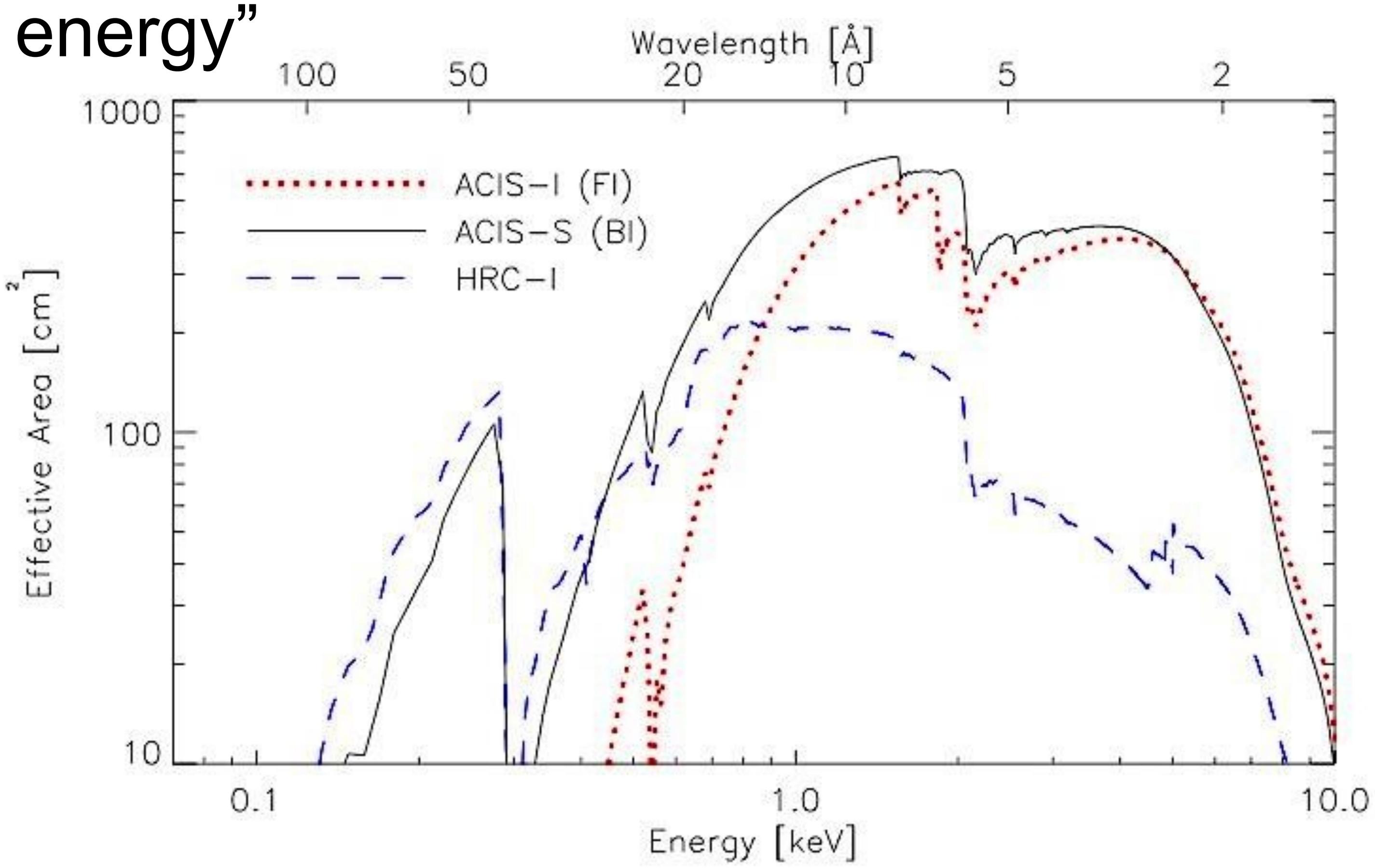
Chandra X-ray Image



Instrumental Effects: Recording inefficiency

- Image:
  - exposure map
  "sensitivity to photons per area"

- Spectrum:
  - effective area (ARF)
  "sensitivity to photons per energy"

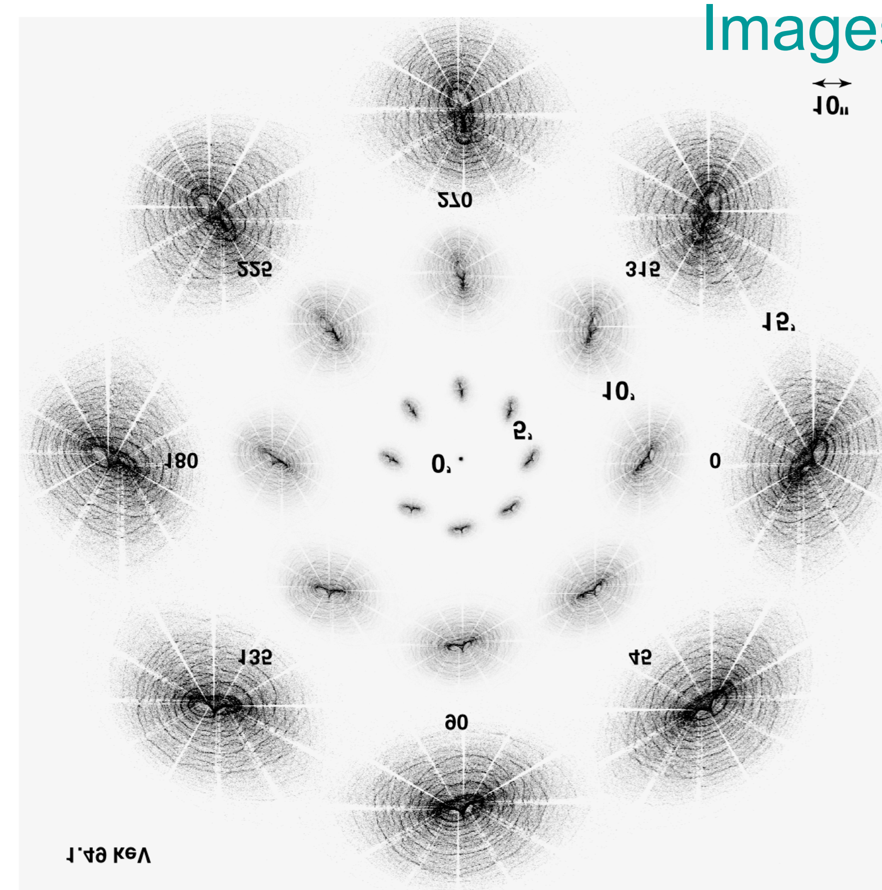


Instrumental Effects: Blurring

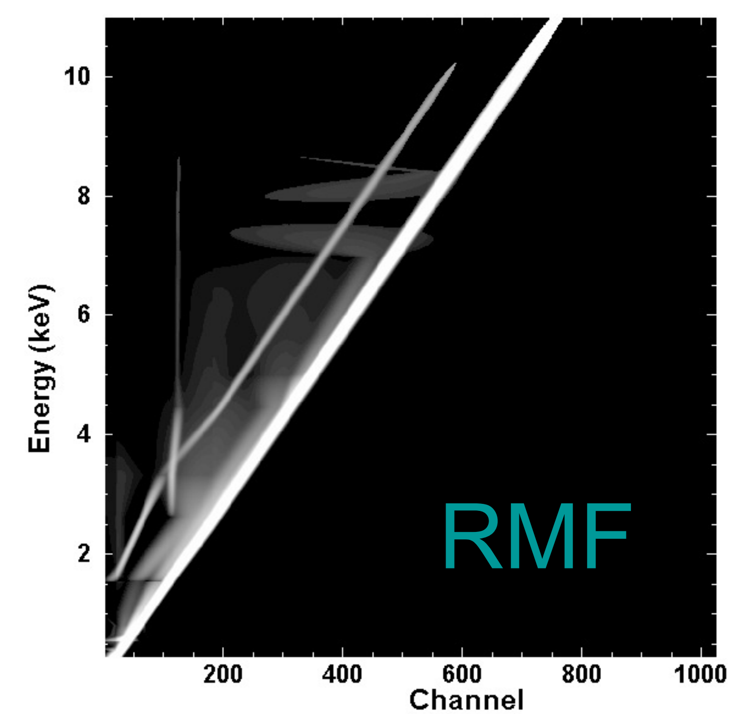
- Image
  - point source observed size depends on the source location on the detector
  - "blurring" is described by a point spread function (PSF)

- Spectrum
  - photon energy is "blurred"
  - probability of detecting photon at given energy in given detector channel is described by a redistribution matrix (RMF)

PSF Simulated Images



Energy

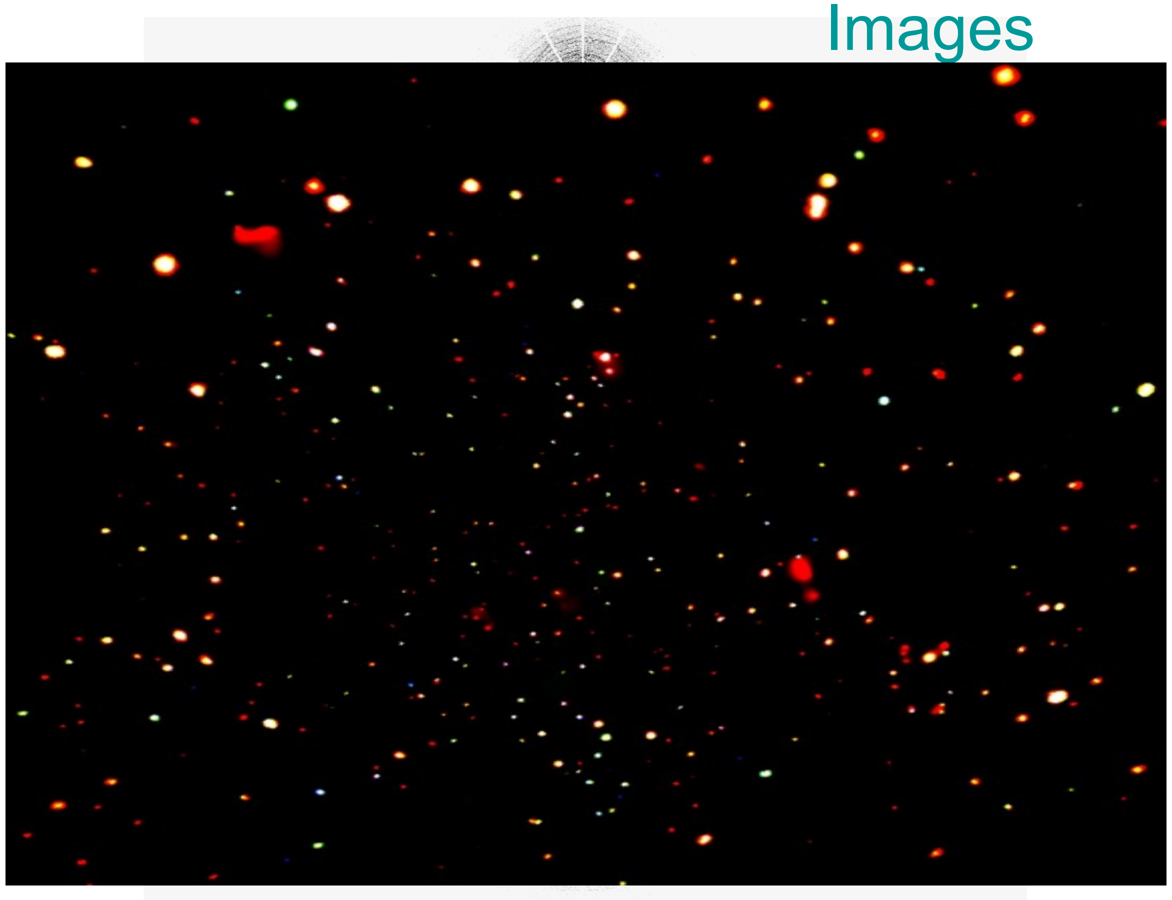


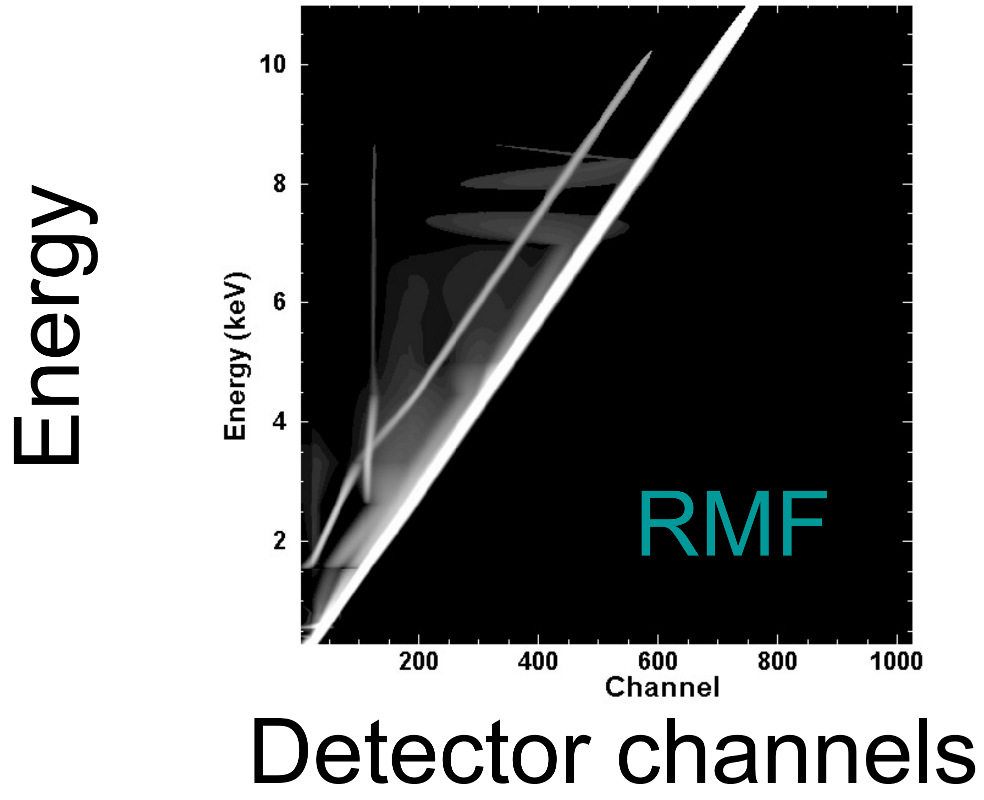
Detector channels

Instrumental Effects: Blurring

- Image
  - point source observed size depends on the source location on the detector
  - “blurring” is described by a point spread function (PSF)
- Spectrum
  - photon energy is “blurred”
  - probability of detecting photon at given energy in given detector channel is described by a redistribution matrix (RMF)

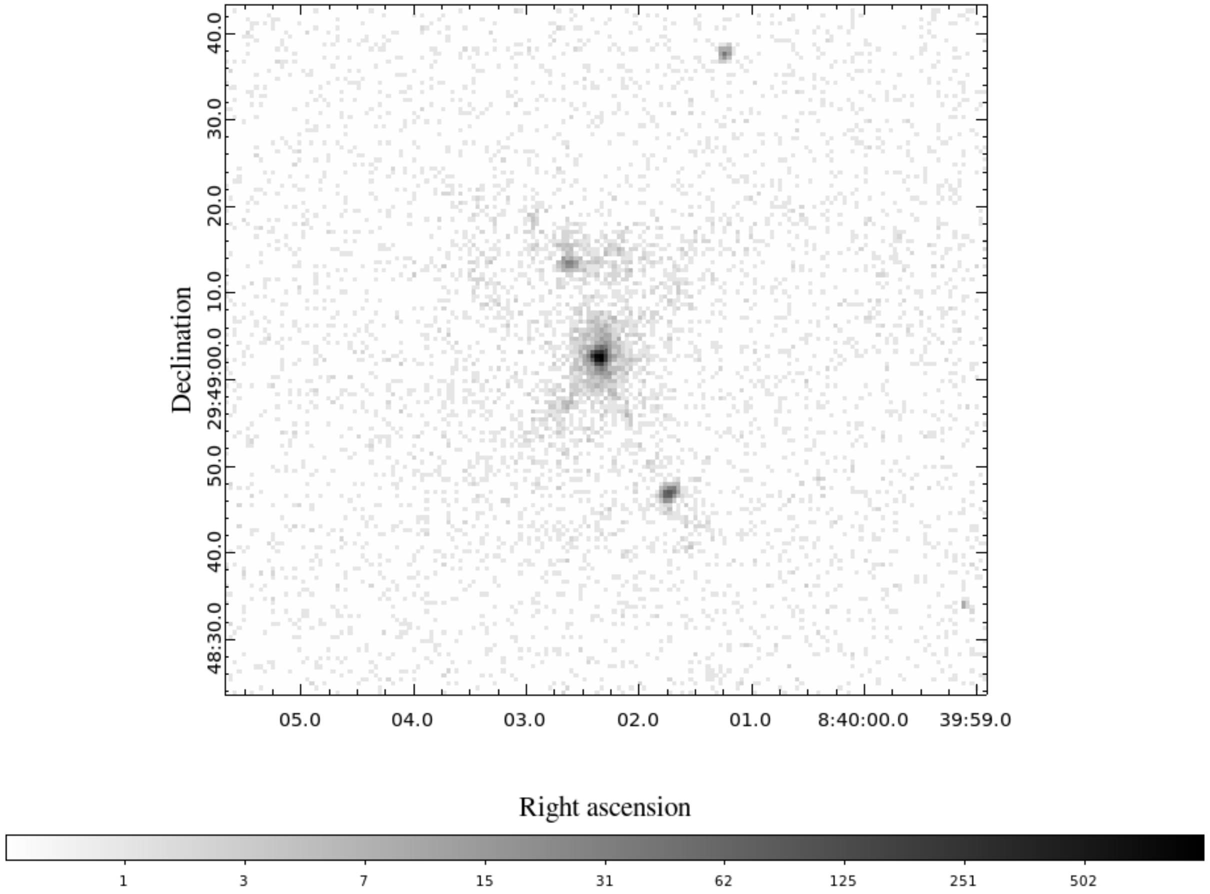
PSF Simulated Images





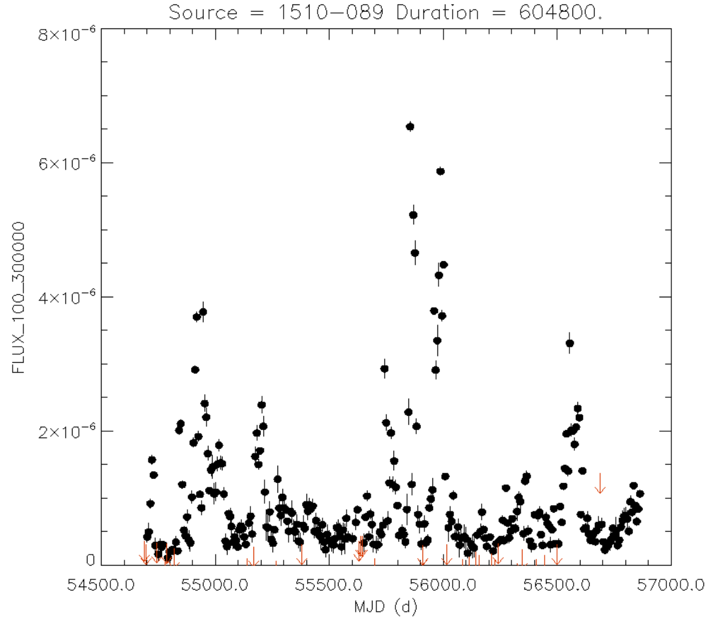
Challenges

- Sparse, locally saturating, Poisson data
- Instrumental effects
- Source Detection in Deep Images
- Irregular extended structures
- Source boundaries
- Complex physical models
- Non-periodic, stochastic variability



Chandra X-ray Observatory

Fermi LAT



International CHASC Astro-Statistics Collaboration

This page lists resources of specific interest to astronomers. For detailed descriptions and reports of C-BAS/ICHASC activities, see
www2.imperial.ac.uk/~dvandyk/astrostat.php

Software | Activities | Bibliography | Astro jargon | Stat jargon | People | Mailing-List | Internal

ostat-announce GoogleGroup | GoogleCalendar | AstroStat Slog Archive

Search site

Faculty/Researchers

Statisticians
- David van Dyk, Imperial College London
- Nathan Stein, University of Pennsylvania
- Paul Baines, University of California, Davis
- Thomas Lee, University of California, Davis
- Xiao-Li Meng, Harvard
- Yaming Yu, University of California, Irvine

Astronomers
- Andreas Zezas, Crete
- Aneta Siemiginowska, Harvard-Smithsonian Center for As
- Kaisey Mandel, Harvard-Smithsonian Center for Astrophy
- Vinay Kashyap, Harvard-Smithsonian Center for Astrophy

Associates

Alex Young, NASA-GSFC
Pavlos Protopapas, Harvard
Peter Freeman, Carnegie Mellon
Taeyoung Park, Yonsei

PhD Students

Dan Cervone (Harvard)
David Jones (Harvard)
David Stenning (UC Irvine)
Hyungsuk Tak (Harvard)
Irina Udaltsova (UC Davis)
Lazhi Wang (Harvard)
Minjie Fan (UC Davis)
Qi Gao (UC Davis)
Shijing Si (Imperial)
Vasileios Stampoulis (Imperial)
Xiyun Jiao (Imperial)

CHASC was founded in 1997

Former Students

CJ Zijin Shen (Harvard 2000), Head of Options Trading, Jump Trading, LLC
Chris Hans (Harvard AB 2001), Associate Professor, Ohio State University
Rostislav S. Protassov (Harvard 2003) Director, Citigroup
David Esch (Harvard 2003), Director of Research, New Frontier Advisors
Hosung Kang (Harvard 2005), Quantitative Analyst, Graham Capital Management
Yaming Yu (Harvard 2005), Associate Professor, Univ of California, Irvine
Taeyoung Park (Harvard 2006), Associate Professor, Yonsei Univ., Korea
Alan Burton Lenarcic (Harvard 2009), Manhattan Securities and Exchange Commis
Paul David Baines (Harvard 2010), Assistant Professor, Univ of California, Davis
Xianchao Xie (Harvard 2011), Two Sigma Investments
Li Zhu (Harvard 2012), Getco LLC
Jingchen Liu (Harvard 2008), Assistant Professor, Columbia
Victoria Liublinska (Harvard 2013), College Fellow in Statistics, Harvard
Nathan Stein (Harvard 2013), Visiting Assistant Professor, University of Pennsylva
Alex Blocker (Harvard 2013), Google
Jin Xu (UC Irvine 2014), Adobe
Shandong Zhao* (UC Irvine 2014), Apple
Raymond Wong* (UC Davis 2014), Assistant Professor, Iowa State University

Former Faculty / Researchers / Associates

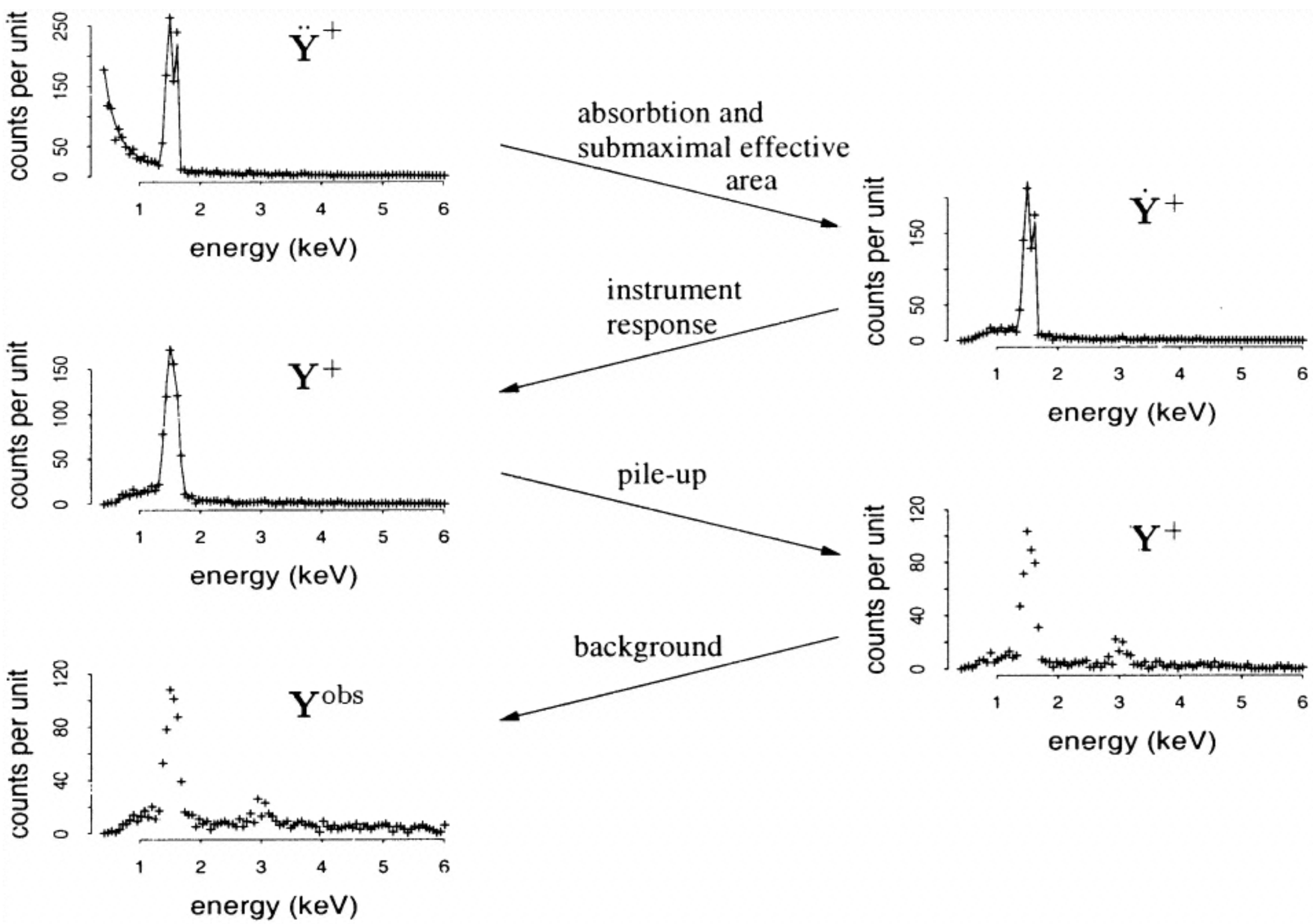
Alanna Connors
Eric Kolatcyk, Boston University
James Chiang, Stanford
Rima Izem

Former Post Docs

Brandon Kelly (Hubble Fellow), UCSB
Hyunsook Lee, Korea Institute of S&T Evaluation and Planning

BLoCXS

CJ Shen / Chris Hans / Rostislav Protassov / Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

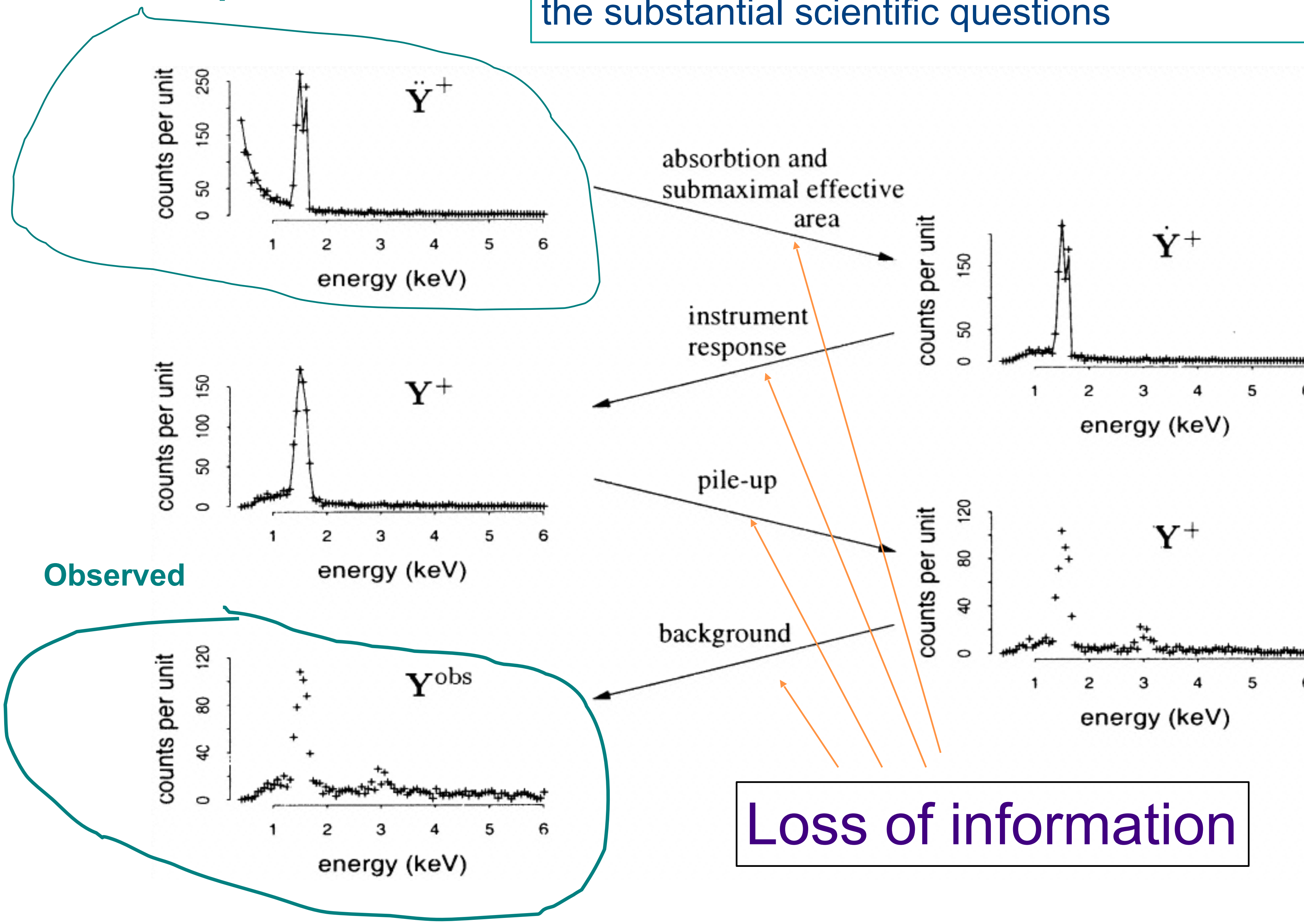


van Dyk, D. A., Connors, A., Kashyap, V. L., Siemiginowska, A. (2001)
Analysis of Energy Spectra with Low Photon Counts via Bayesian Posterior Simulation.
The Astrophysical Journal , 548, 224-243.

Highly Structured Statistical Models

Model directly the source and data collection, and include statistical procedure to fit the resulting highly structured models and address the substantial scientific questions

Emitted Spectrum



van Dyk et al. 2001

Bayesian Inference

- Complex data collection needs to be included in the statistical model:

Physical Source Model

$$\lambda = I\{As\} + \beta$$

Expected photon count

Instrumental "blurring"
* from calibration

Non-homogenous stochastic censoring
* from calibration

Background Contamination
* from separate observation

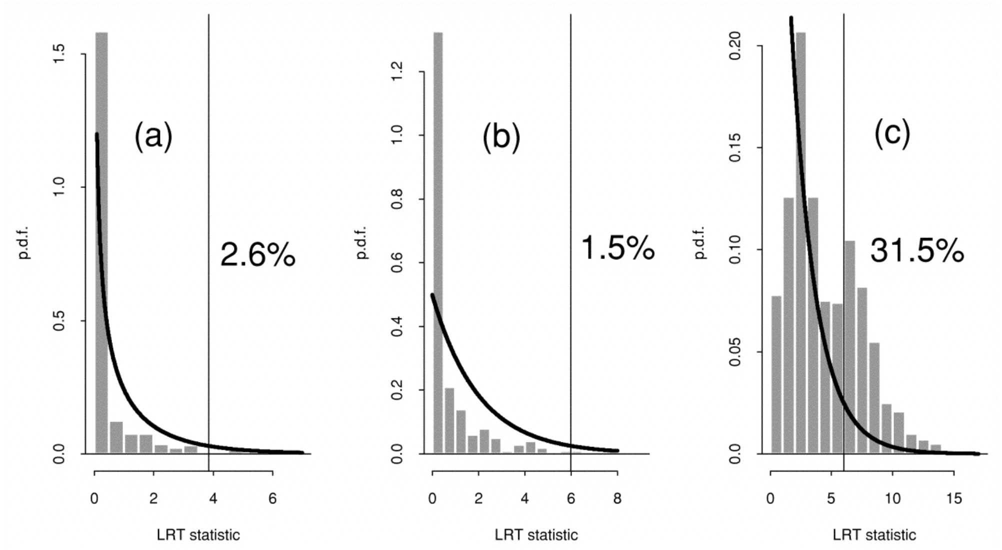
Observed counts are modeled as independent Poisson variables with λ mean

BLoCXS / ppp

Rostislav Protassov / Yaming Yu / Taeyoung Park

Protassov LRT

- plot of LRT distributions
line detection



F-test was being commonly misused in astro analyses
because of a lack of appreciation
of the asymptotic conditions under which it was valid.

posterior predictive p-values for LRTs

Protassov+ 2002, became our most famous paper
has been cited 301 times

Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L. and Siemiginowska, A. (2002). *Statistics: Handle with Care, Detecting Multiple Model Components with the Likelihood Ratio Test.* ApJ, 571, 545-559.

Park, T. van Dyk, Siemiginowska, A. (2008) -*Searching for Narrow Emission Lines in X-ray Spectra: Computation and Methods,* ApJ. 688, 807

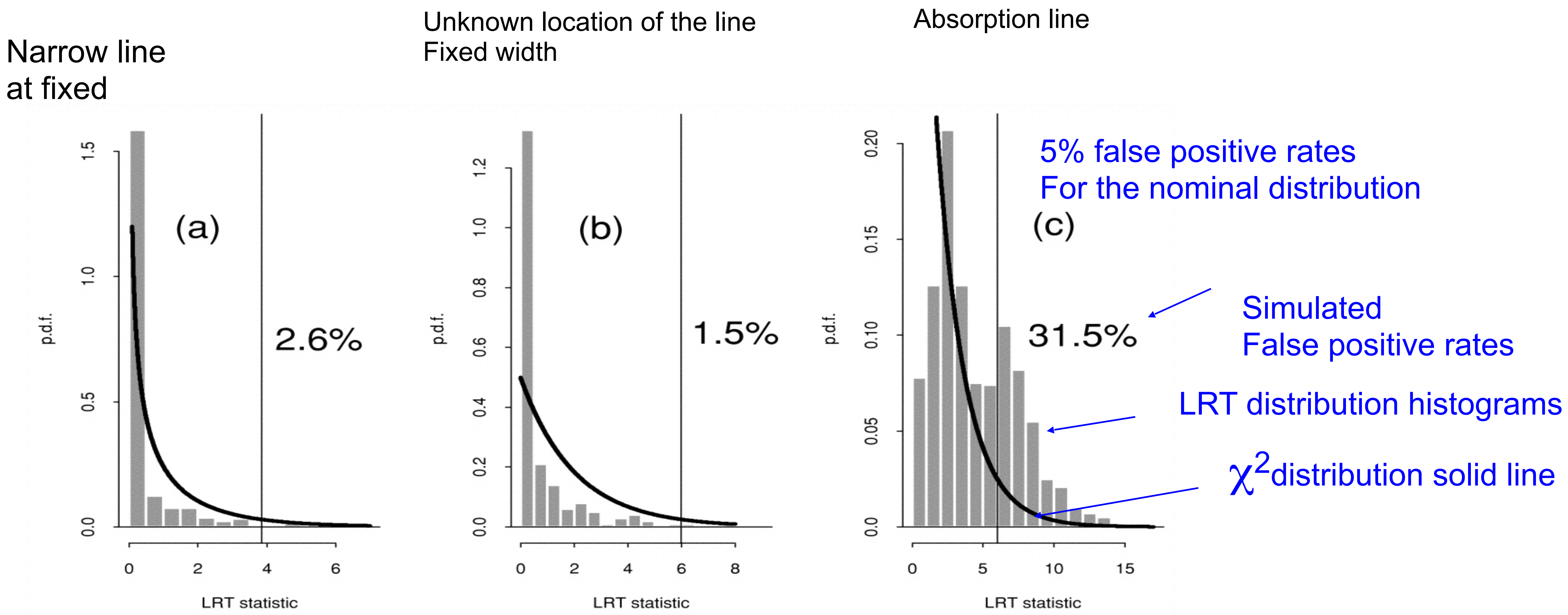
LRT

» Assumptions of the Likelihood Ratio Test statistics:
  - The null hypothesis must be a special case of the alternative
  - The parameter space of the null must be interior of the alternative parameter space.

» The second assumption fails when testing for a spectral emission line:
  - When there is no line, the line intensity is zero, it may not be negative.
  - The line locations and width of the line do not exist when there is no line. They have no values.

LRT

IMPORTANT! We do not know the true distribution of the test statistics.



» Results of three tests compared to the nominal χ^2 distribution

Protassov et al. (2002)

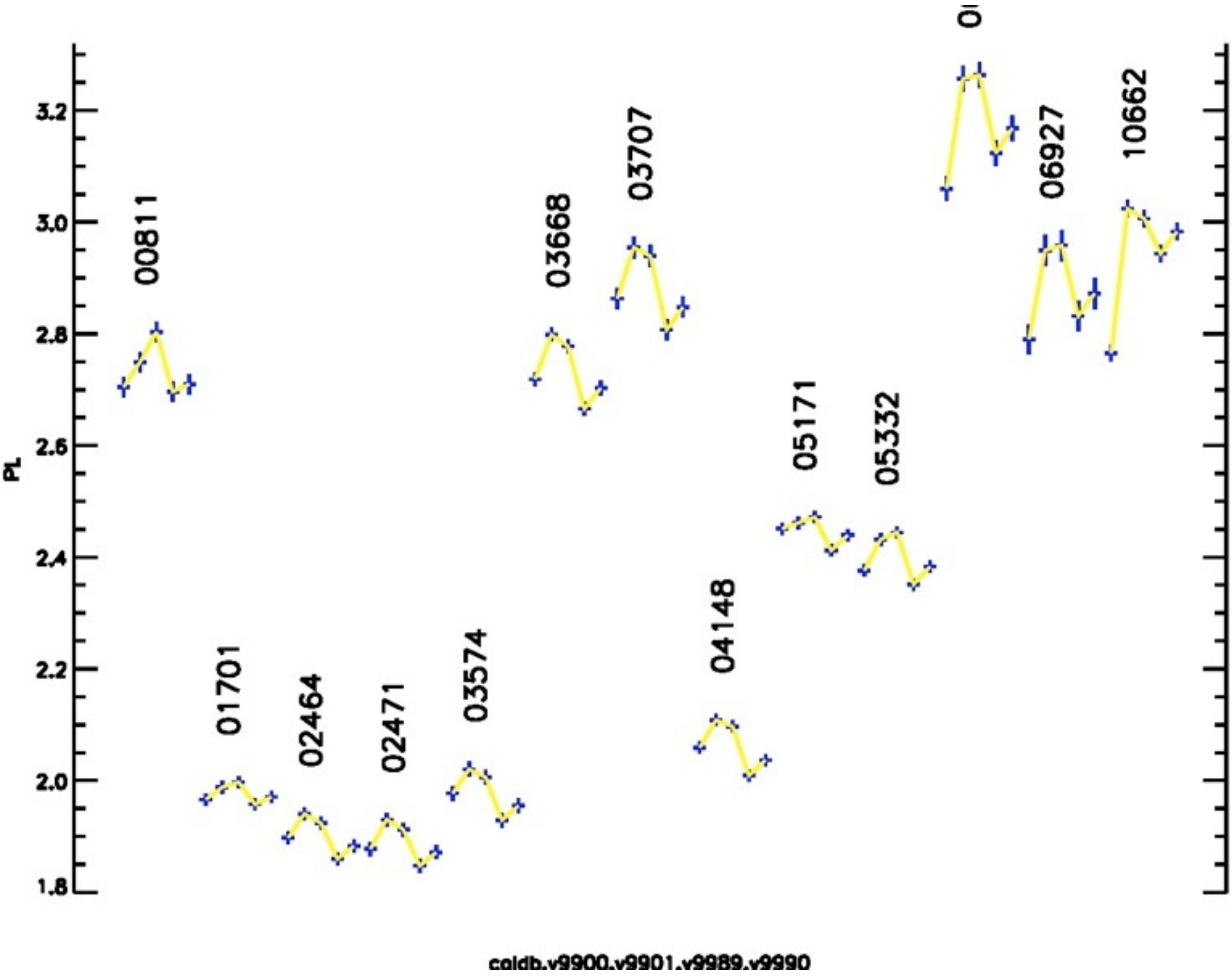
pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu

Foundations of Astronomical inference:
Measurement
Significance
Calibration

Calibration is not perfect, it has known statistical and systematic errors, and unknown errors that are only guessed at.

Drake, J.J., et al. 2006, *"Monte Carlo processes for including Chandra instrument response uncertainties in parameter estimation studies"*, SPIE Proc. 6270, 49
Kashyap, V.L., et al. 2008, *"How to handle calibration uncertainties in high-energy astrophysics"*, SPIE Proc. 7016, 21
Lee, H., et al. 2011, *"Accounting for Calibration Uncertainties in X-ray Analysis: Effective Areas in Spectral Fitting"*, ApJ, 731, 126
Xu, J., et al. 2014, *"A Fully Bayesian Method for Jointly Fitting Instrumental Calibration and X-ray Spectral Models"*, ApJ, in press

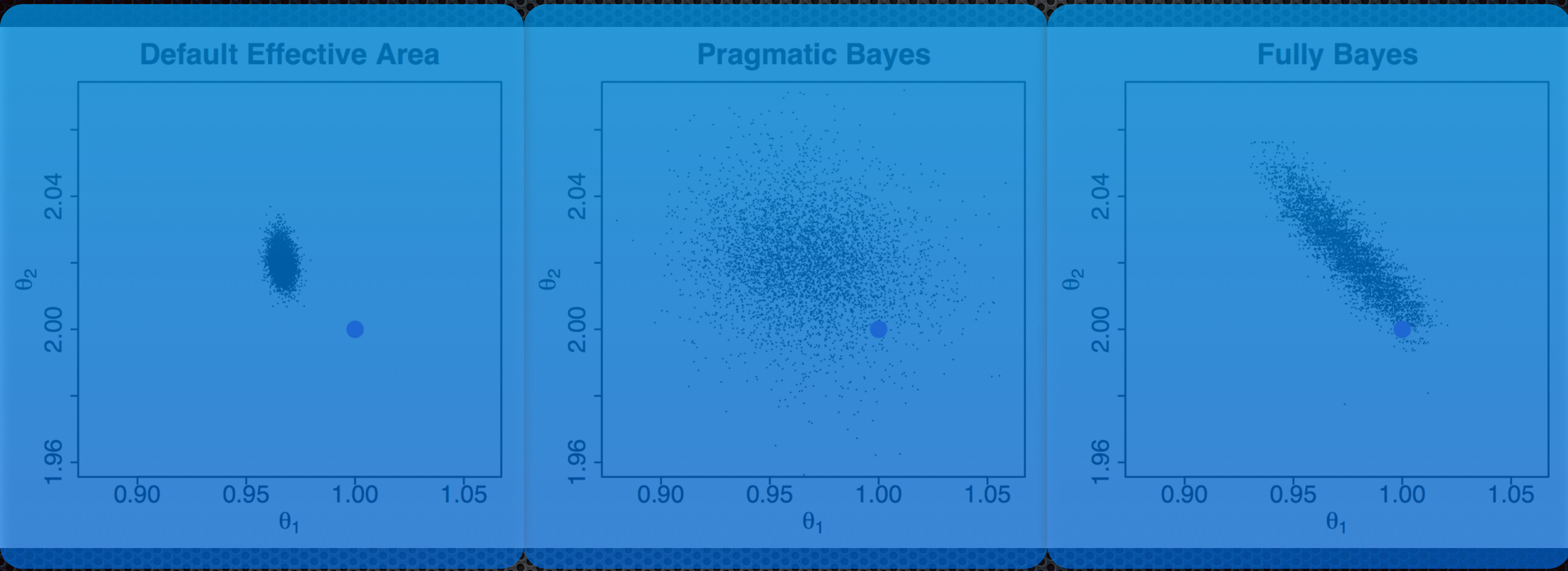

3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
PL
00811
01701
02464
02471
03574
03668
03707
04148
05171
05332
06927
10662
coldb.v9900.v9901.v9989.v9990

pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$

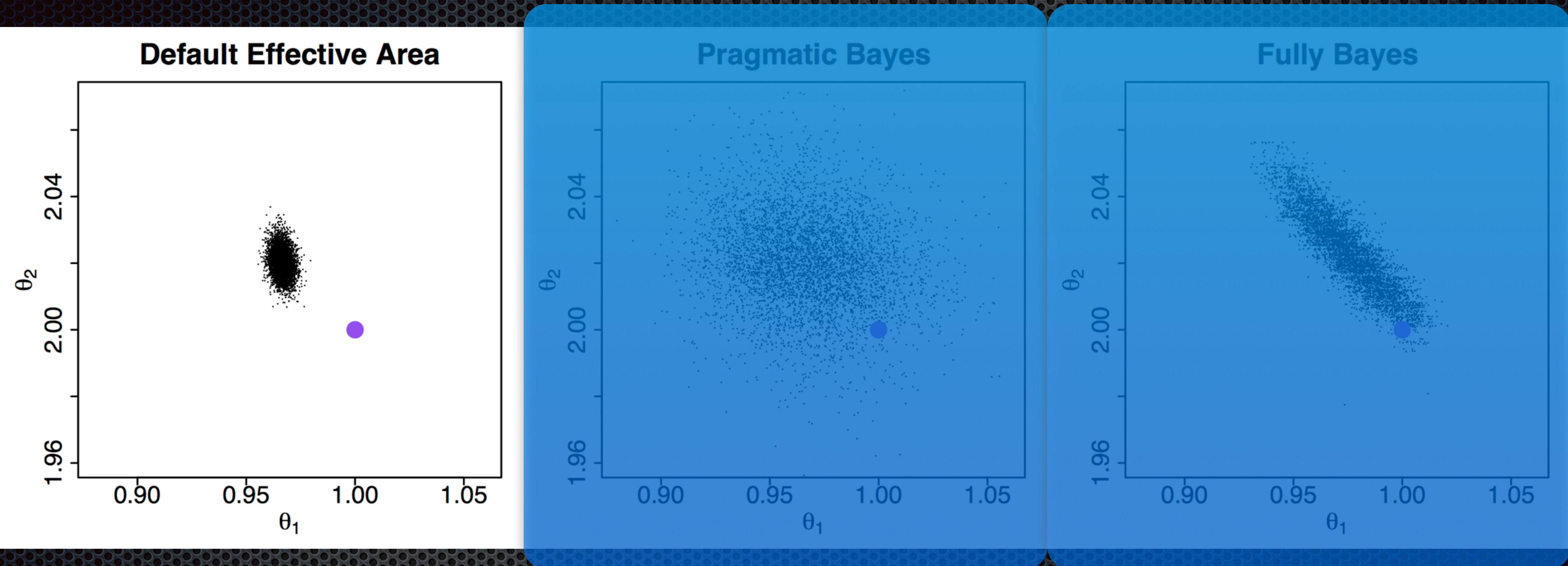


pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$



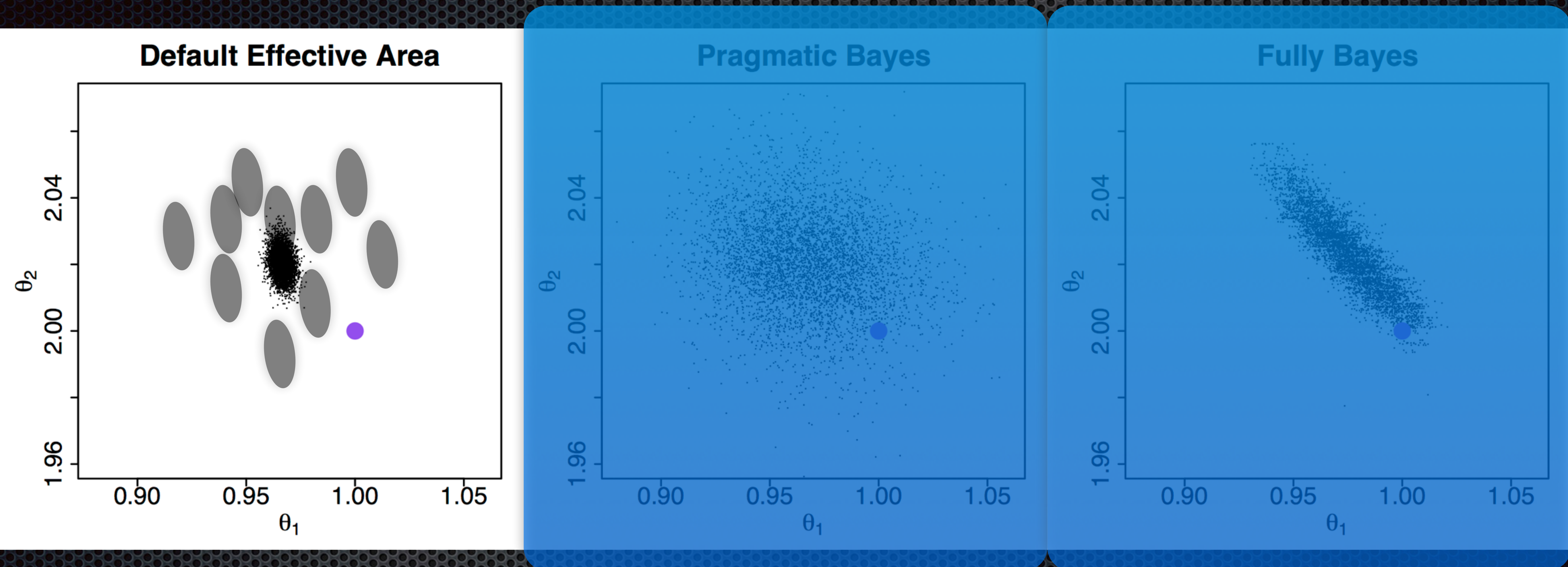
$p(\theta|D,A_0)$

pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$



$p(\theta|D,A_0)$

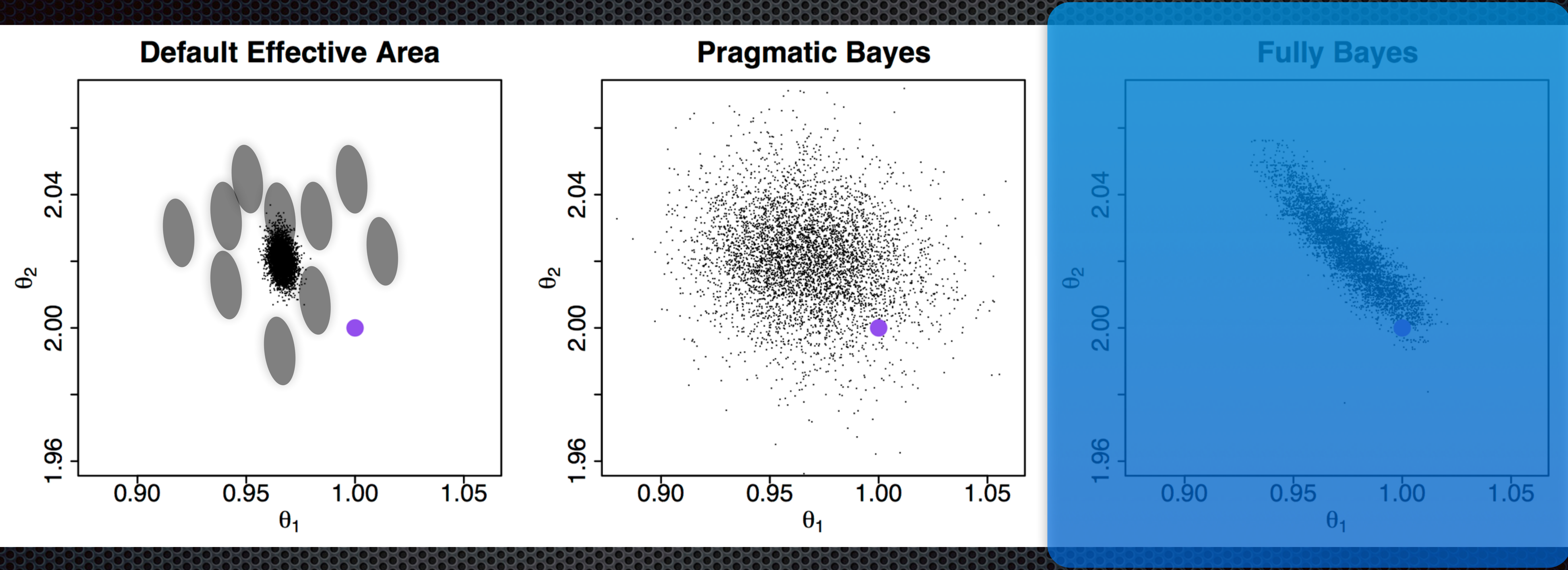
$p(\theta|D,A_i)$

pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$



$p(\theta|D,A_0)$

$p(A)\ p(\theta|D,A)$

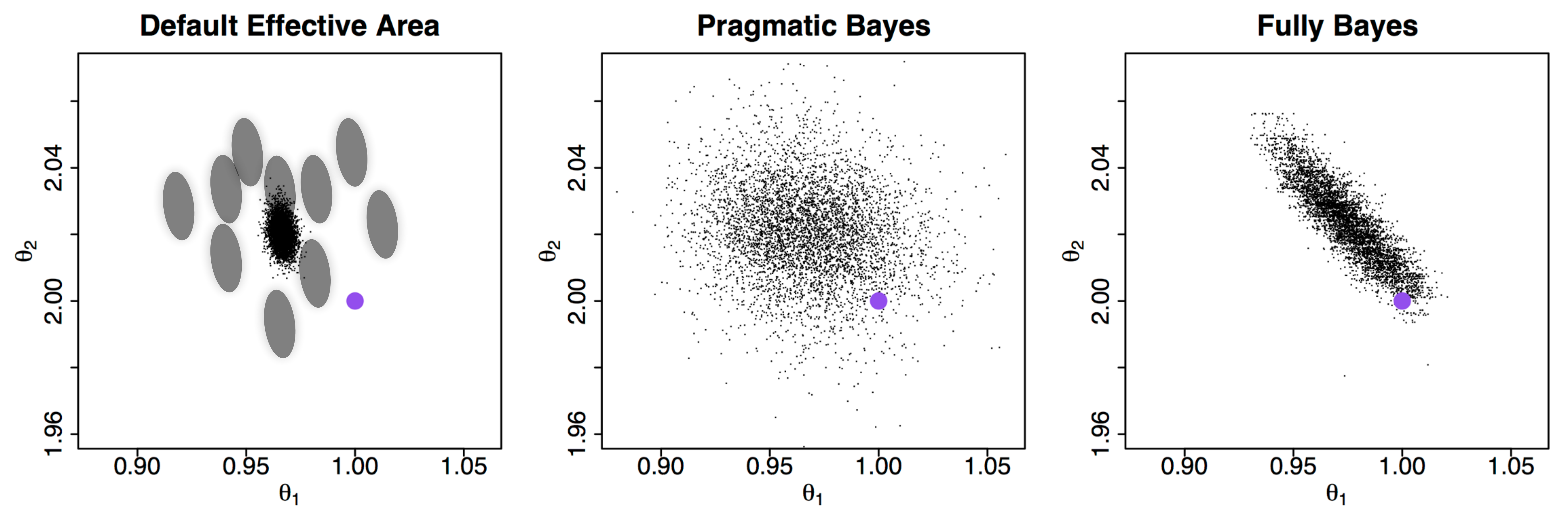
$p(\theta|D,A_i)$

pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$



$p(\theta|D,A_0)$ $p(A)\ p(\theta|D,A)$ $p(A,\theta|D)$

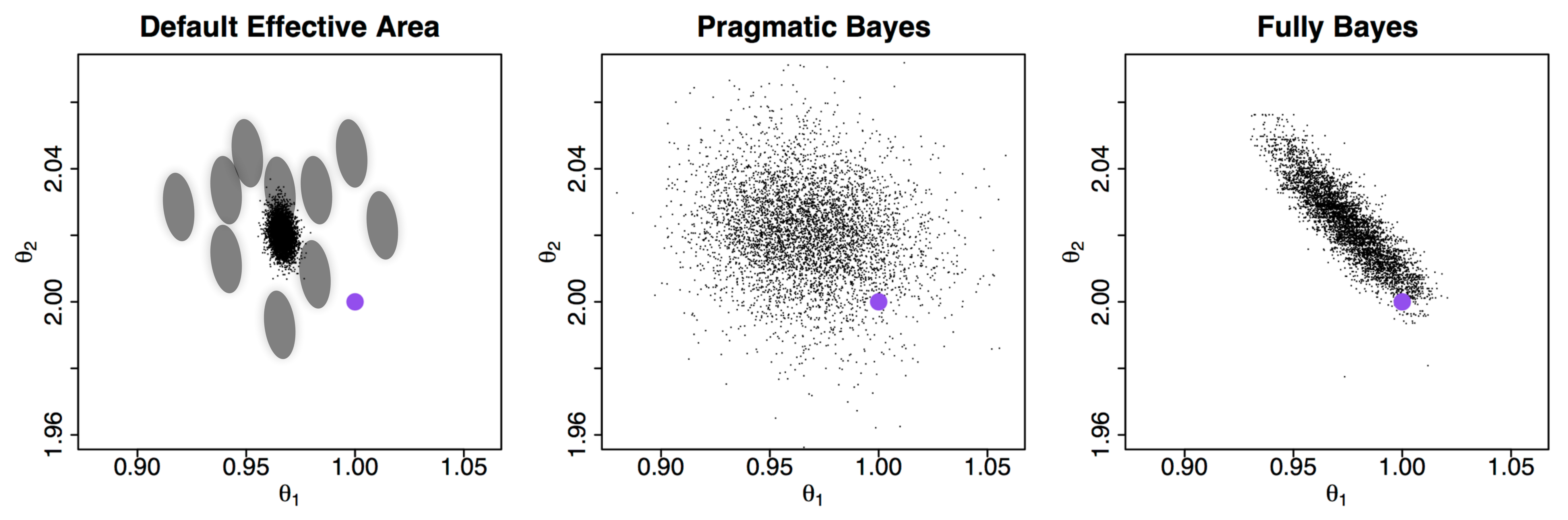
$p(\theta|D,A_i)$

pyBLoCXS / Calibration

Yaming Yu / Taeyoung Park / Hyunsook Lee / Jin Xu / Shandong Min

fitting to simulated data

$$f(\varepsilon;\theta) = \theta_3\, \varepsilon^{-\theta_1}\, e^{-\theta_2\, \sigma(\varepsilon)}$$



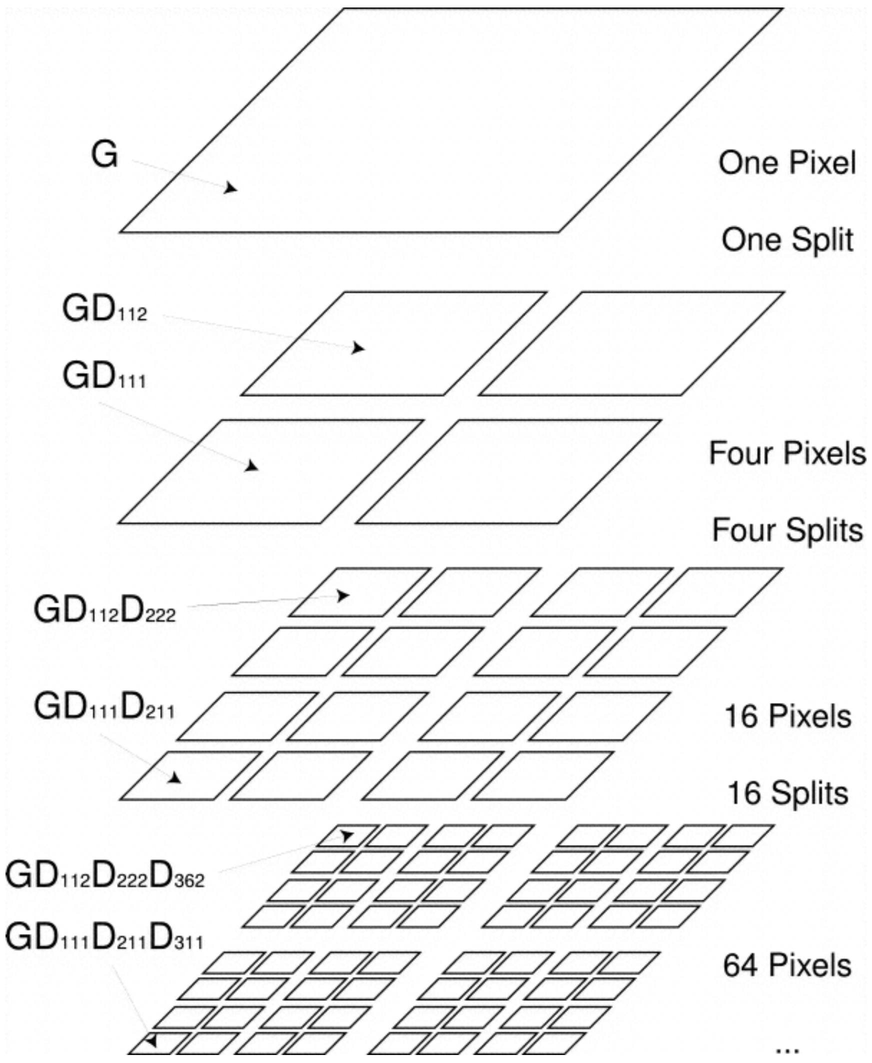
$p(\theta|D,A_0)$ $\quad$ $p(A)\ p(\theta|D,A)$ $\quad$ $p(A,\theta|D)$

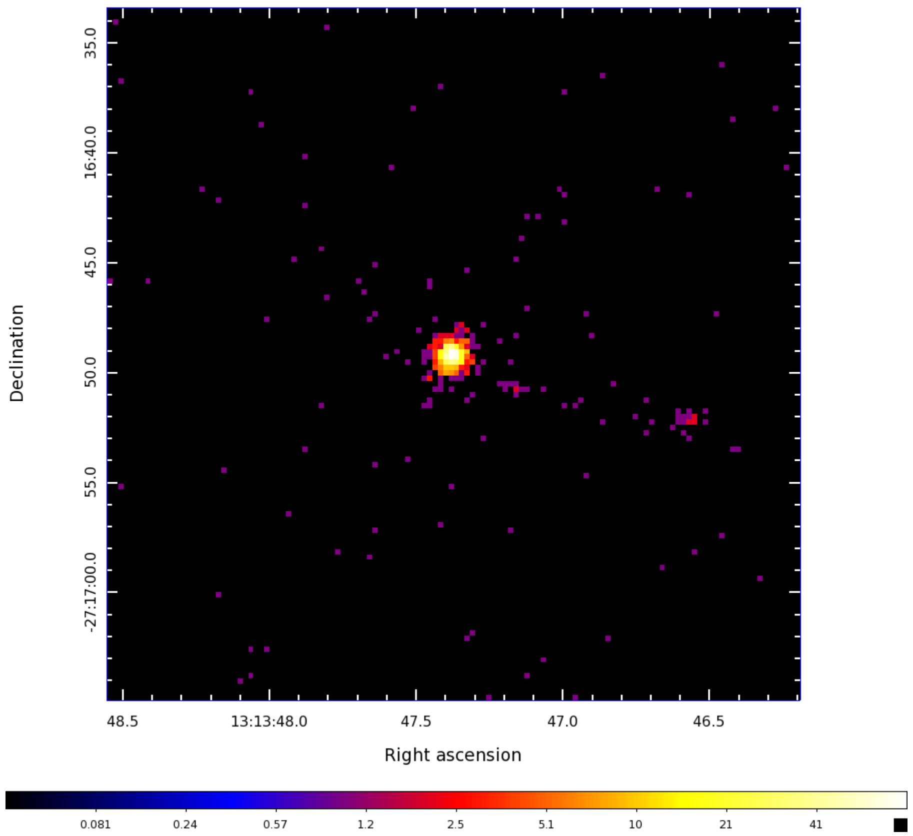
$p(\theta|D,A_i)$ $\quad$ $p(A(\theta'),\theta|D)$

EMC2

David Esch

Bayesian Multi-scale reconstruction of low-counts images

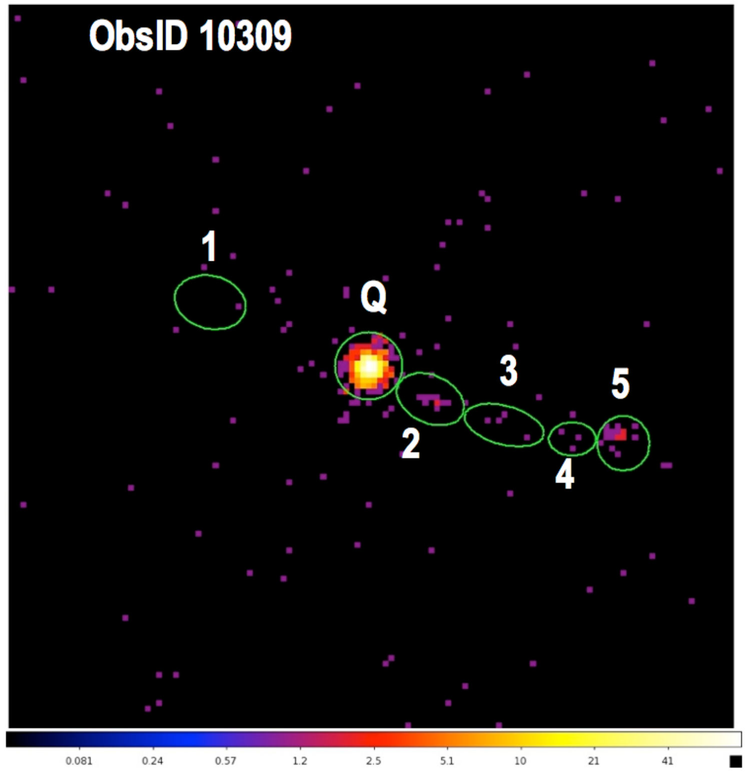


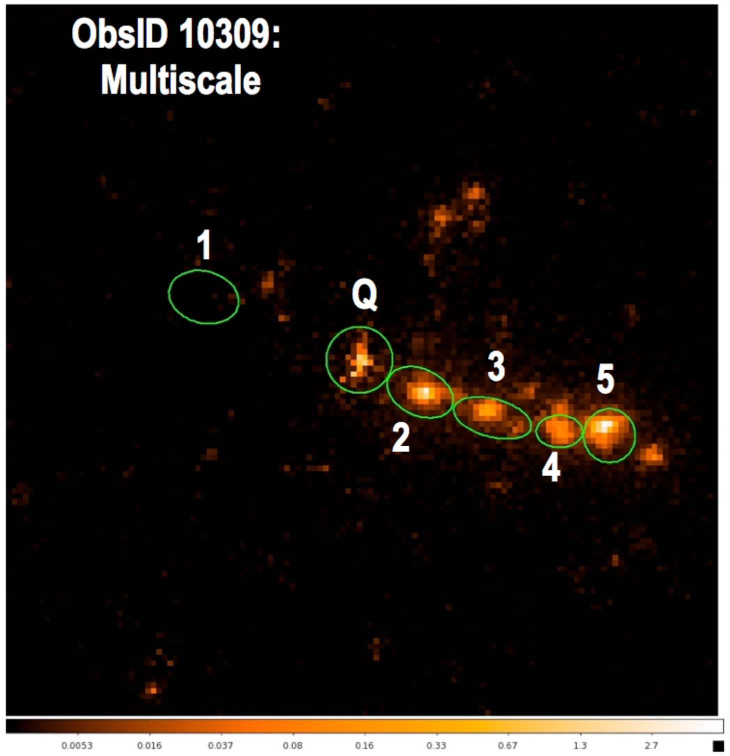


Esch, D.N., Connors, A., Karovska, M., & van Dyk, D.A., 2004,
An Image Restoration Technique with Error Estimates ApJ, 610

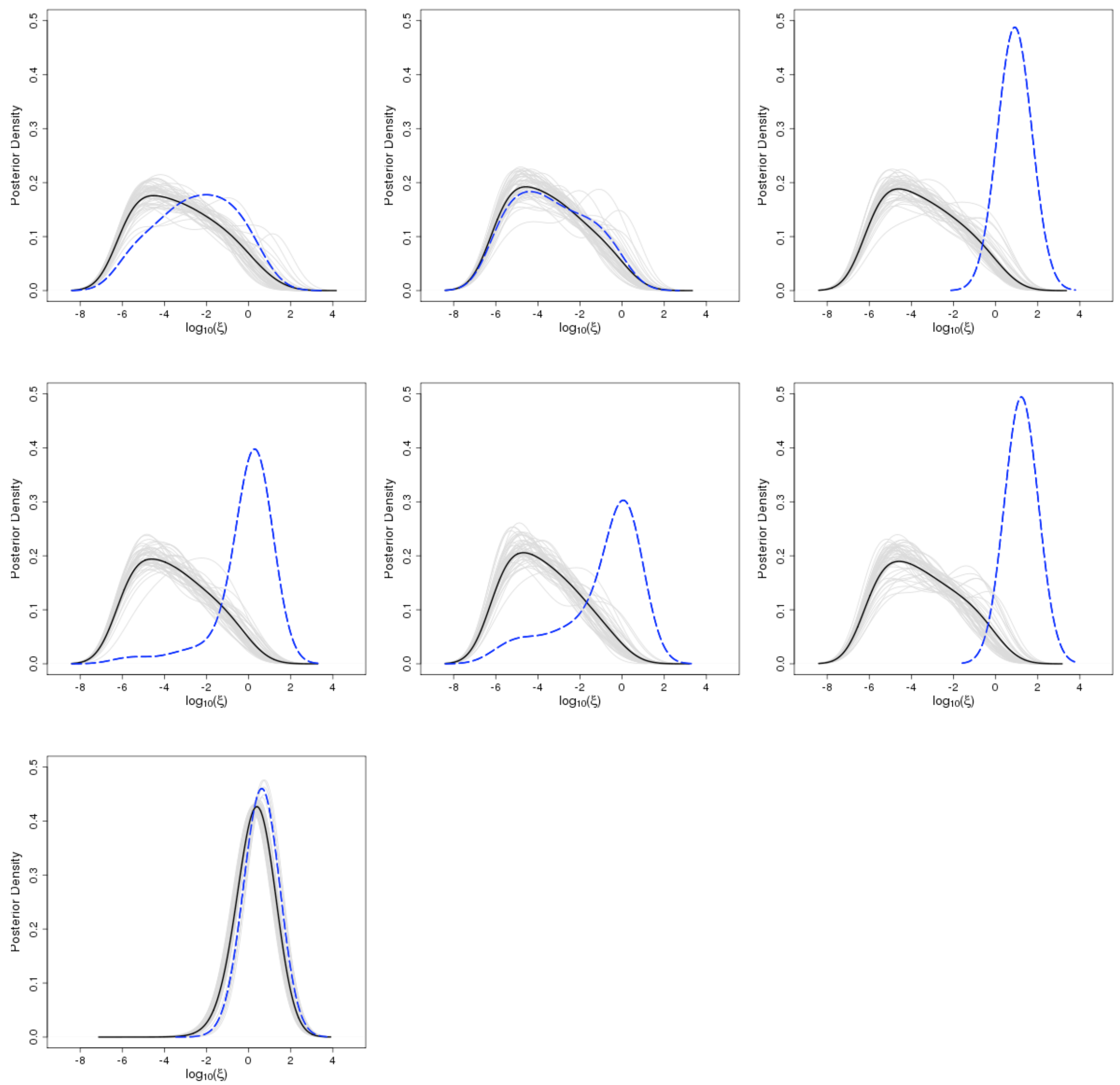
LIRA

Nathan Stein / Katy McKeough





Significance of irregular structure



McKeough +2014, Stein+ 2014

Hardness Ratios

Chris Hans / Yue Wu / Taeyoung Park

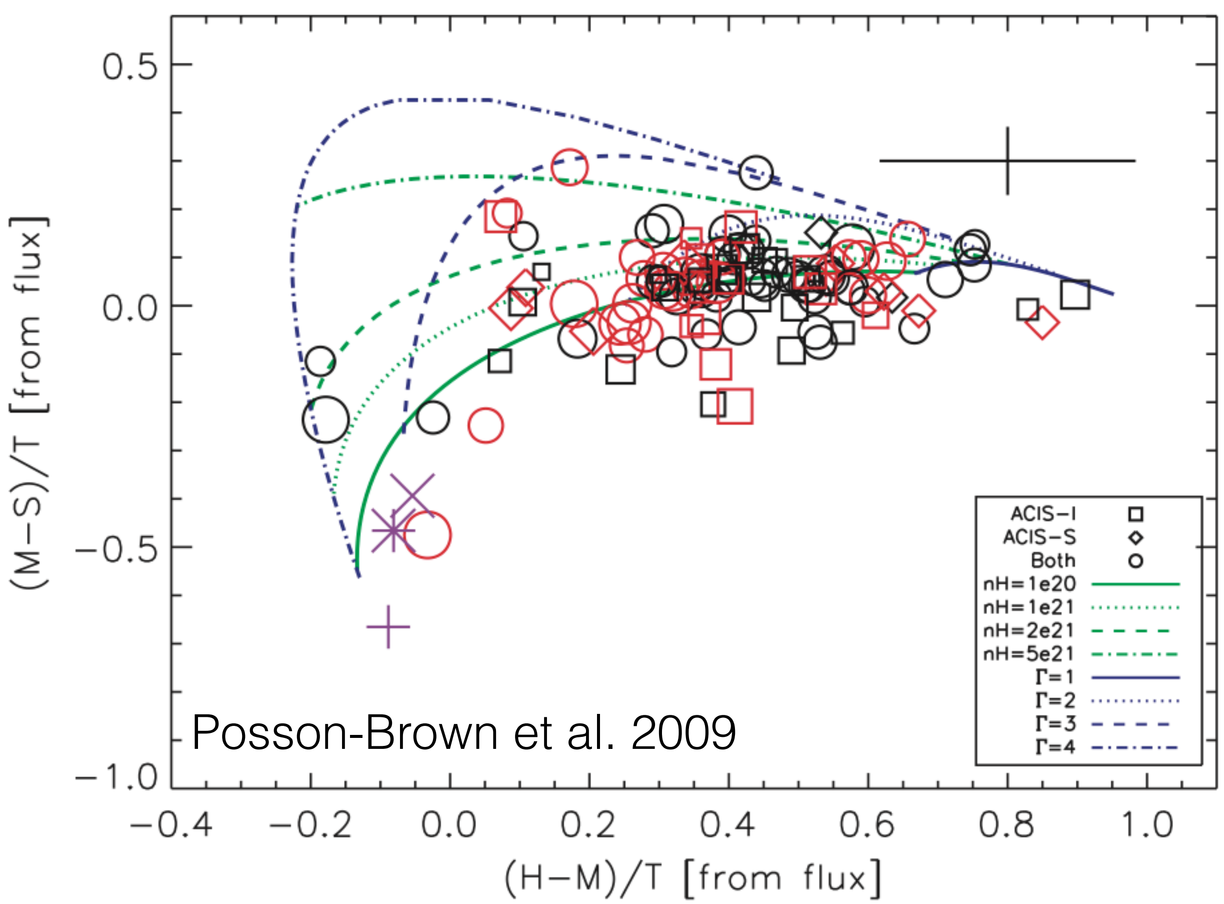
Simplest measure of shape of a spectrum.

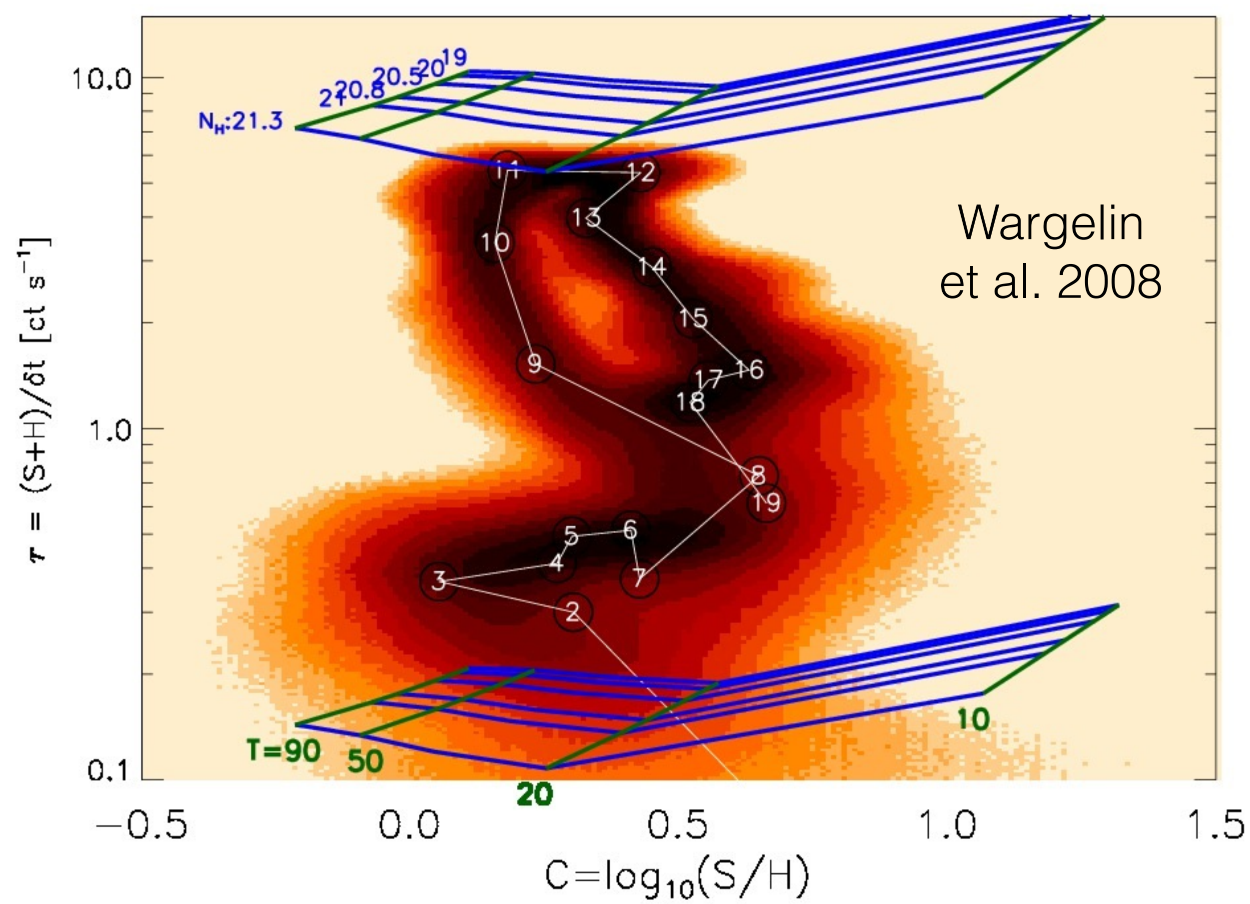
Use counts in passbands, $C_i \sim Pois(a_i\lambda_i + b_i)$, to compute

$$p(R=\lambda_1/\lambda_2 \mid C_i)$$

$$p(HR=(\lambda_1-\lambda_2)/(\lambda_1+\lambda_2) \mid C_i)$$

$$p(C=\log(R) \mid C_i)$$



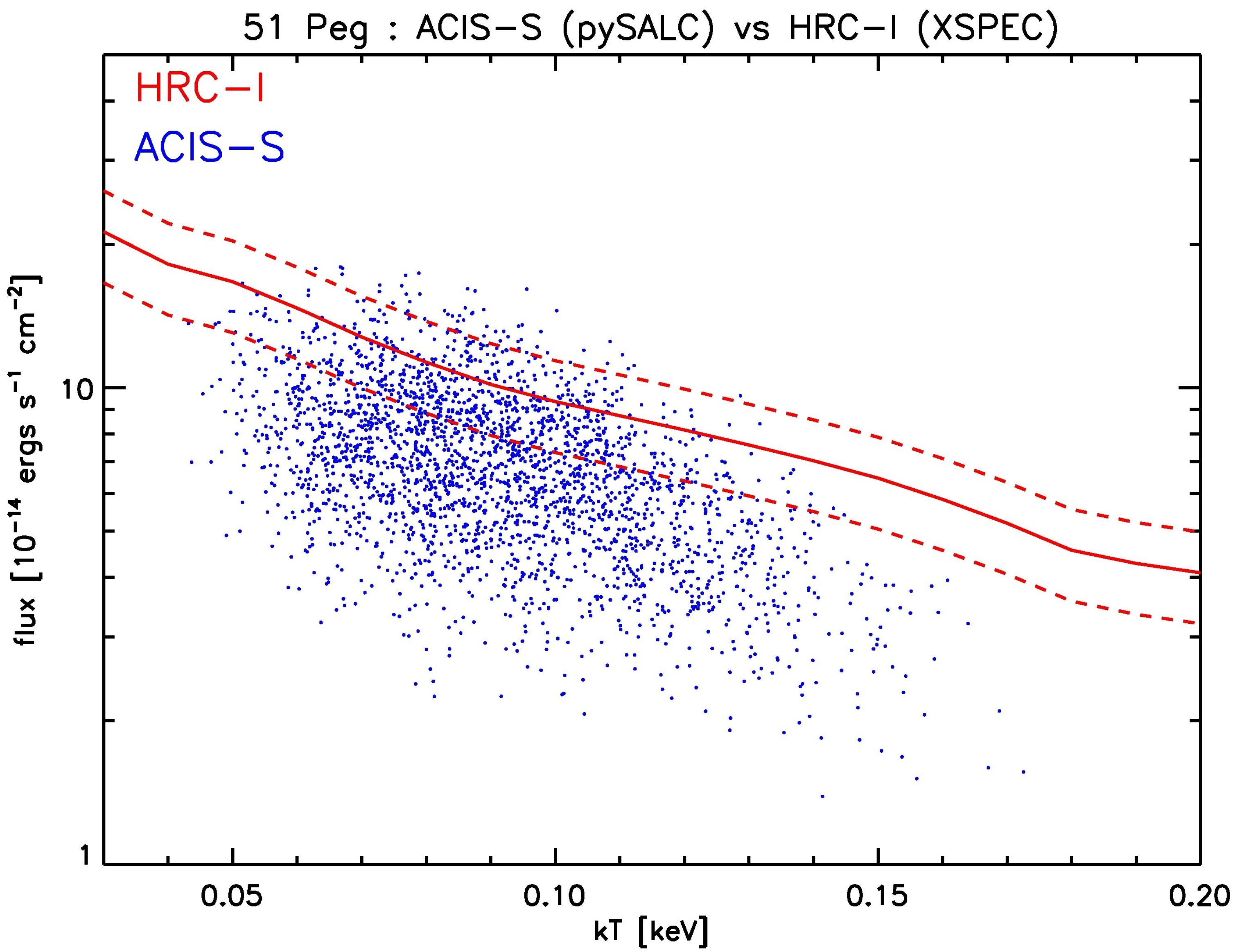


Park, T., et al. 2006, *"Bayesian Estimation of Hardness Ratios: Modeling and Computations"*, ApJ, 652, 610

Hardness Ratios

Taeyoung Park

- ✦ BEHR is used in the *Chandra* Source Catalog
- ✦ Next step: pySALC – infer spectral model parameters directly

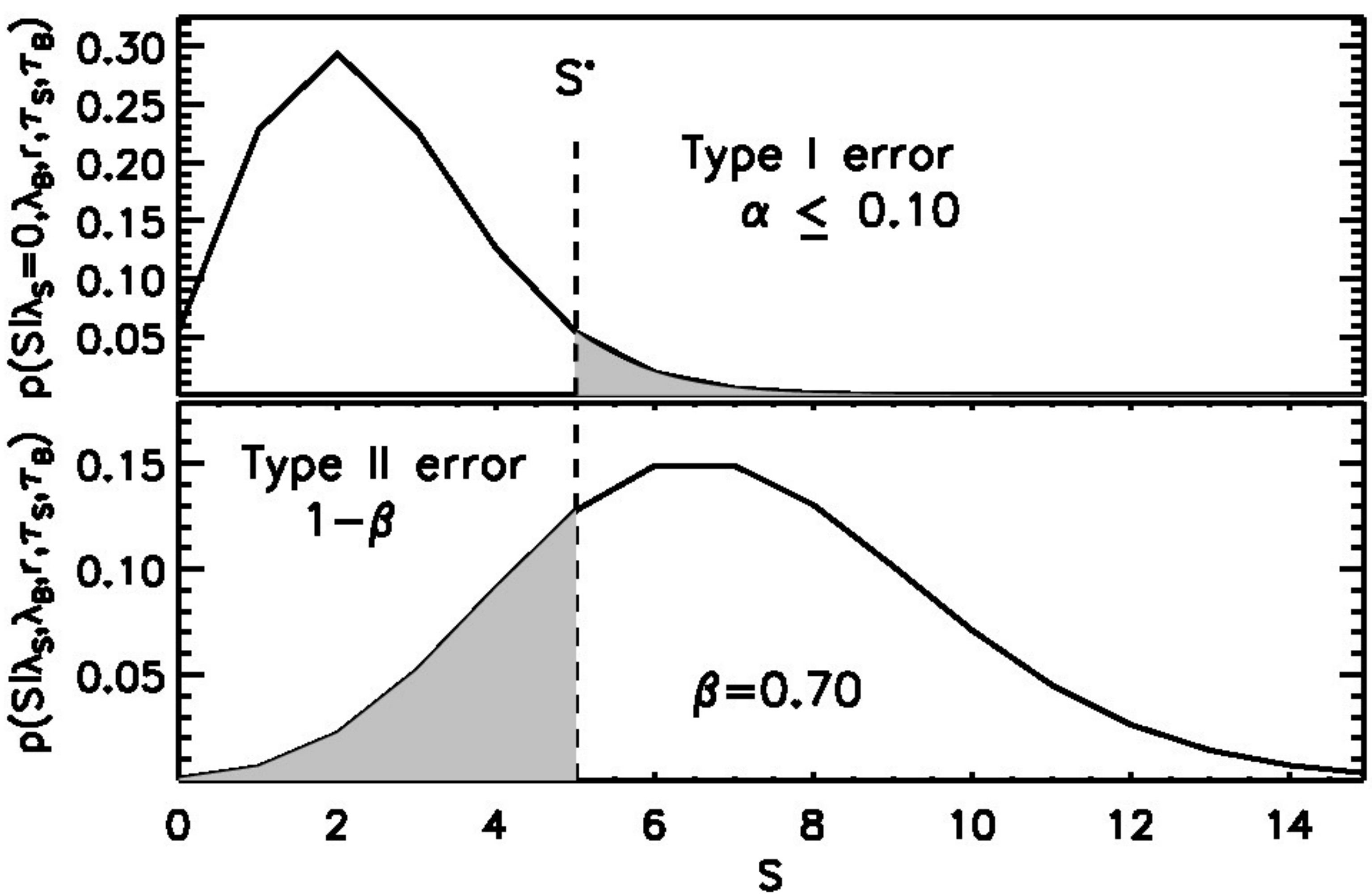


Upper Limits

Jin Xu

Bounds or Limits?

- Bounds: Confidence/Credible range estimates on parameters
- Limits: You make an observation, find nothing, and ask at what brightness would the source have been detected, and conclude it must be dimmer than that.



Kashyap, V.L., et al. 2010, *"On Computing Upper Limits to Source Intensities"*, ApJ, 719, 900

Differential Emission Measure

Hosung Kang / Viktoria Liublinska / Nathan Stein

$$f_{\mathrm{ul};\lambda} = \int d\log T \; G_{\mathrm{ul};\lambda}(T, n_e) \; A_Z \quad n_e^2 \; dV/d\log T$$

Kashyap, V. & Drake, J.J., 1998, *"Markov-Chain Monte Carlo Reconstruction of Emission Measure Distributions: Application to Solar Extreme-Ultraviolet Spectra"*, ApJ, 503, 450
Kang, H., et al., 2003, *"A Response Matrix Approach to the Reconstruction of Differential Emission Measure"*, AAS/SPD 34, 02.01, BAAS 35, p807
Kang, H., et al. 2004, *"Reconstructing Stellar DEMs from X-ray Spectra"*, AAS/HEAD 8.0501
Kang, H, et al. 2005, *"Incorporating Atomic Data Errors in Stellar DEM Reconstruction"*, in X-Ray Diagnostics of Astrophysical PlasmasL Theory, Experiment, and Observation, AIP Conf. Proc., v774, p373

Differential Emission Measure

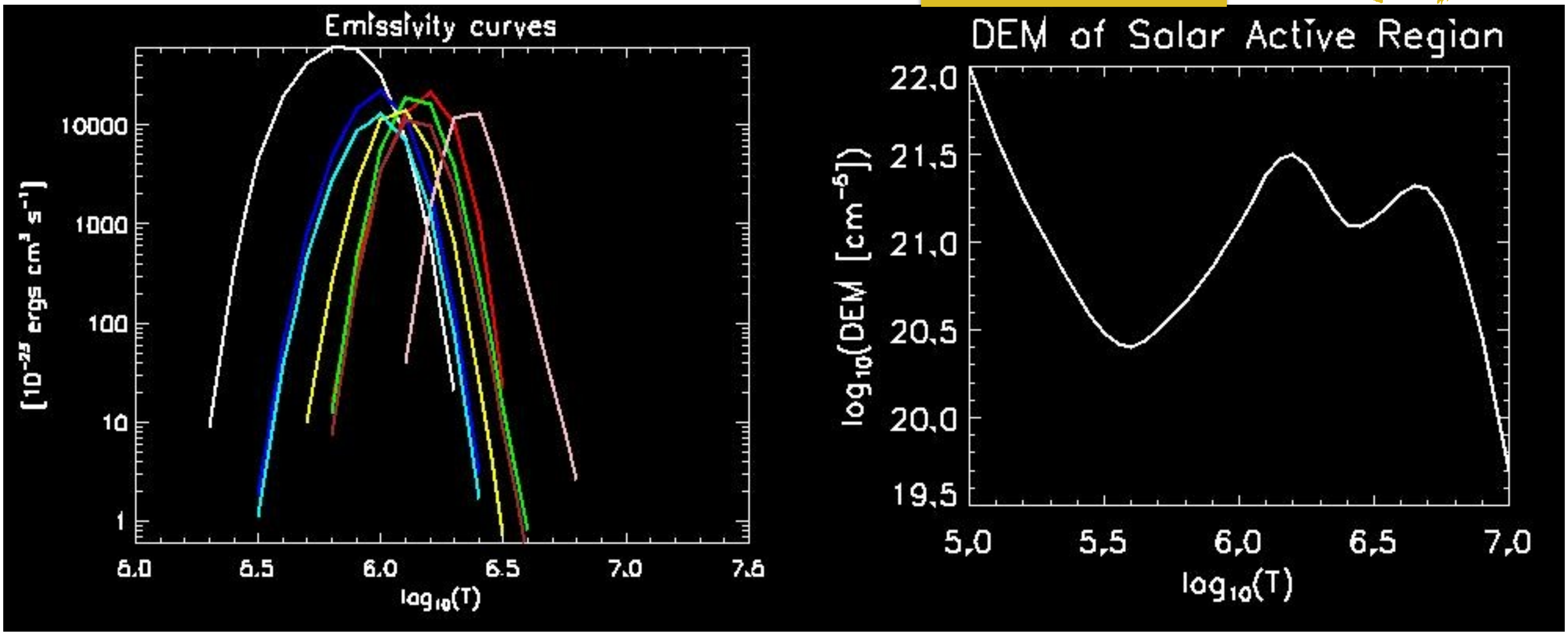
Hosung Kang / Viktoria Liublinska / Nathan Stein

$$f_{\mathrm{ul};\lambda} = \int d\log T\ G_{\mathrm{ul};\lambda}(T,n_e)\ A_Z\ \boxed{n_e^2\ dV/d\log T}\ \text{DEM(T)}$$

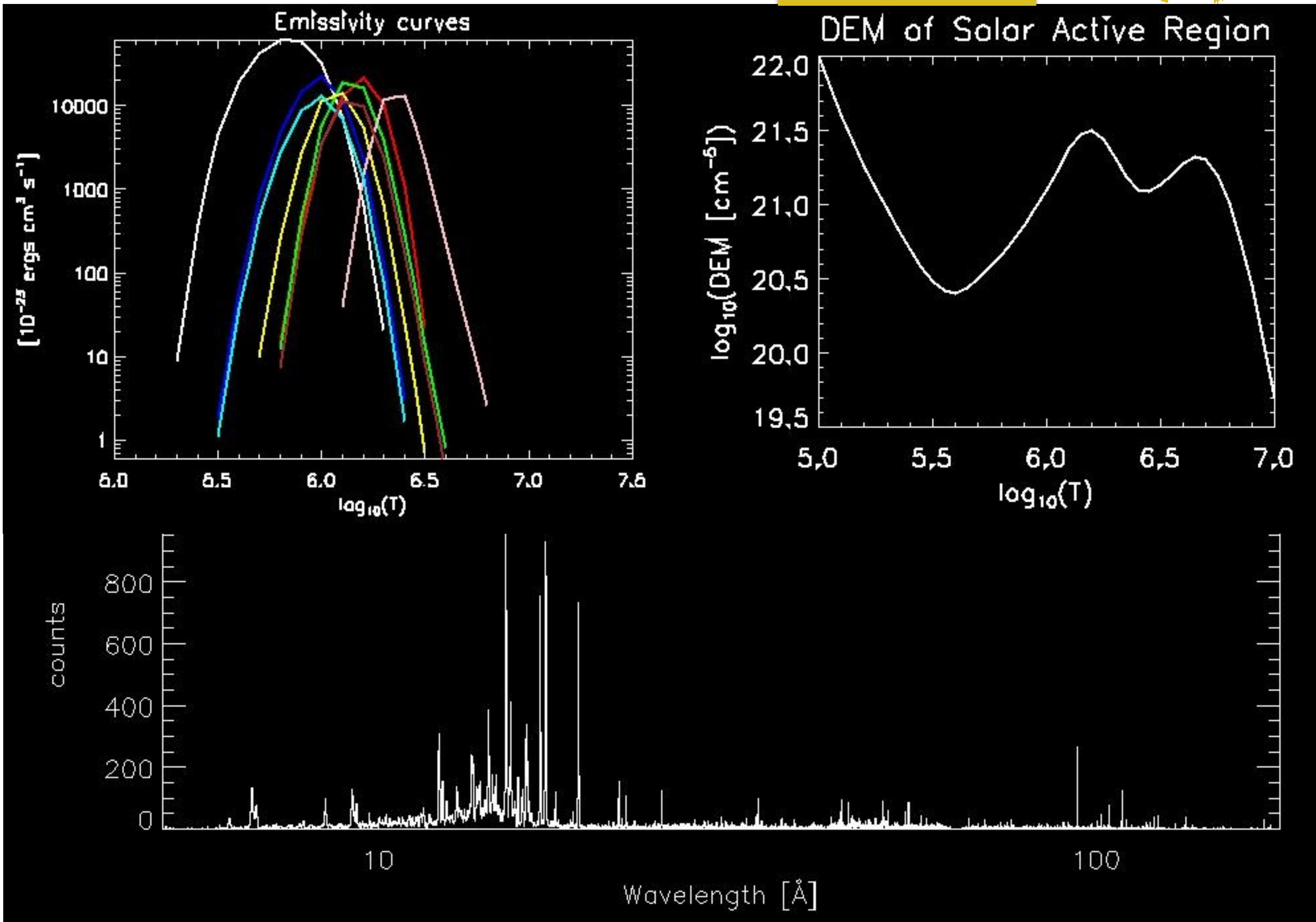
Kashyap, V. & Drake, J.J., 1998, *"Markov-Chain Monte Carlo Reconstruction of Emission Measure Distributions: Application to Solar Extreme-Ultraviolet Spectra"*, ApJ, 503, 450
Kang, H., et al., 2003, *"A Response Matrix Approach to the Reconstruction of Differential Emission Measure"*, AAS/SPD 34, 02.01, BAAS 35, p807
Kang, H., et al. 2004, *"Reconstructing Stellar DEMs from X-ray Spectra"*, AAS/HEAD 8.0501
Kang, H, et al. 2005, *"Incorporating Atomic Data Errors in Stellar DEM Reconstruction"*, in X-Ray Diagnostics of Astrophysical PlasmasL Theory, Experiment, and Observation, AIP Conf. Proc., v774, p373

$$f_{\mathrm{ul};\lambda} = \int d\log T\; G_{\mathrm{ul};\lambda}(T,n_e)\; A_Z\; \boxed{n_e^2\; dV/d\log T}$$ DEM(T)

$$f_{ul;\lambda} = \int d\log T \, G_{ul;\lambda}(T,n_e)\, A_Z \, \boxed{n_e^2 \, dV/d\log T}$$ DEM(T)



$$f_{ul;\lambda} = \int d\log T \, G_{ul;\lambda}(T,n_e)\, A_Z \, \boxed{n_e^2 \, dV/d\log T}$$ DEM(T)



Solar DEMs

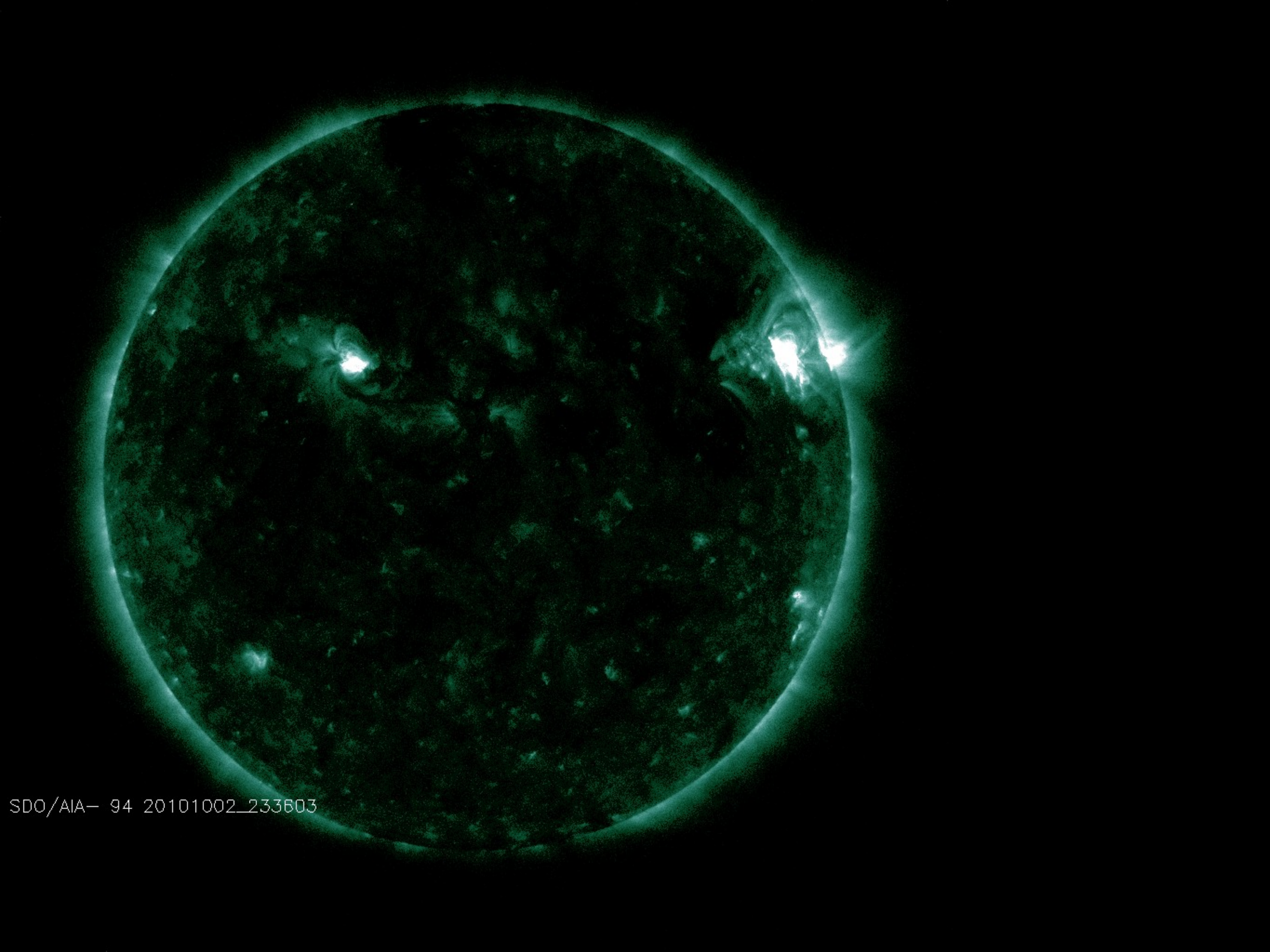
Nathan Stein

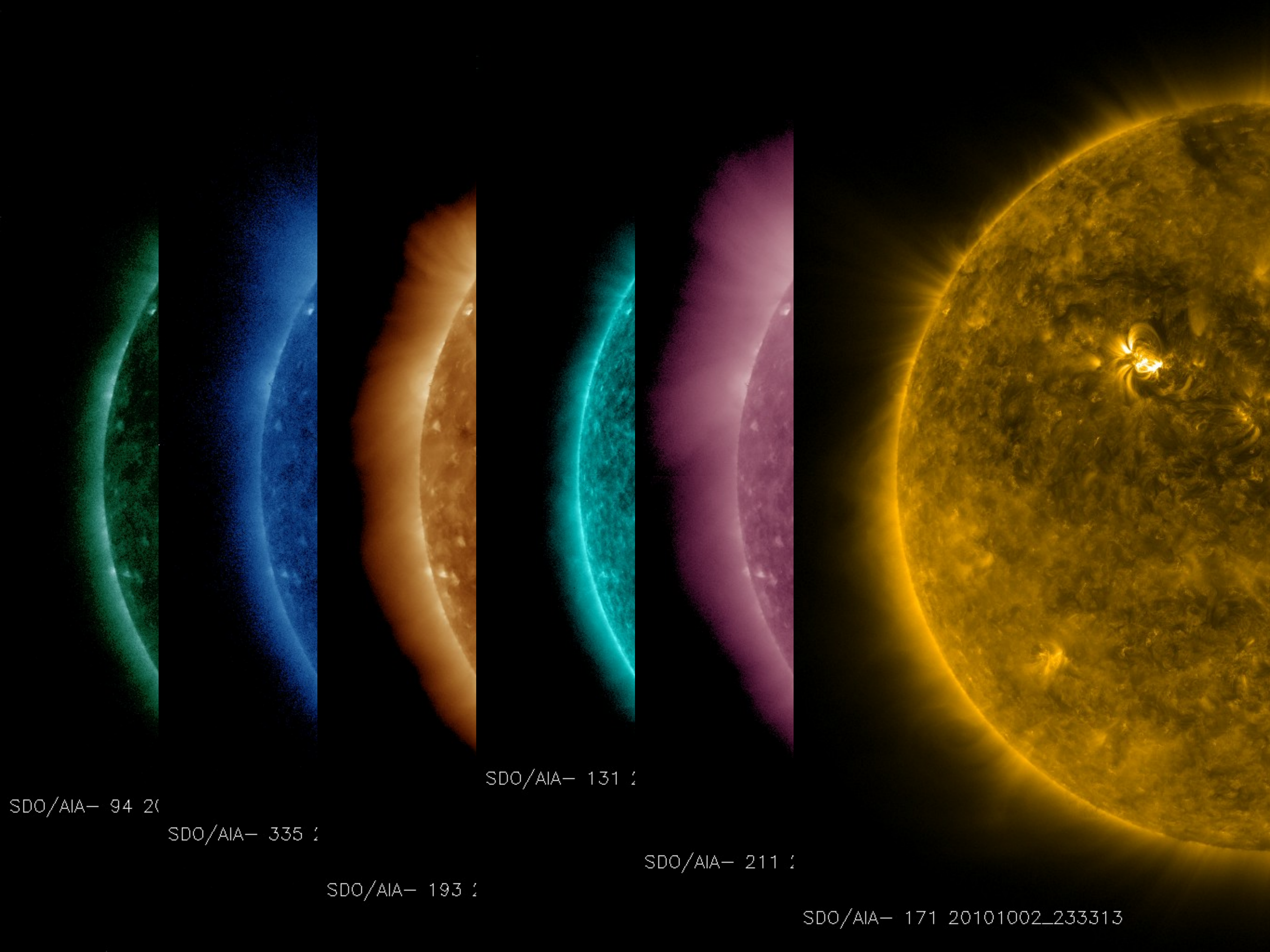
Unlike stellar gratings data, Solar data have high spatial resolution, low spectral and temporal resolutions. Also, high data rate. Large images in multiple filters at ≈12 sec cadence.

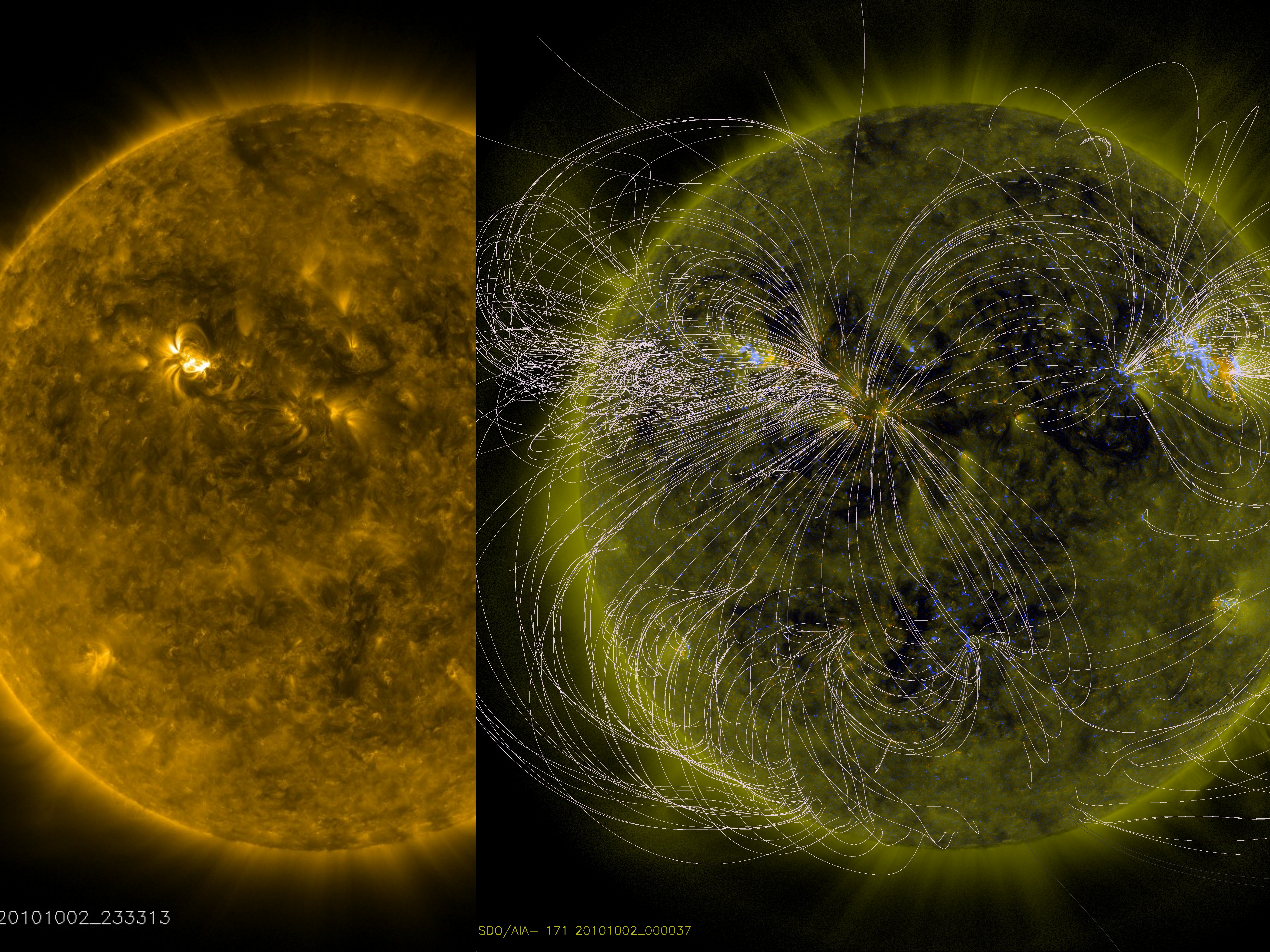
Bypass DEM generation and compute thermal segmentation directly from the data.

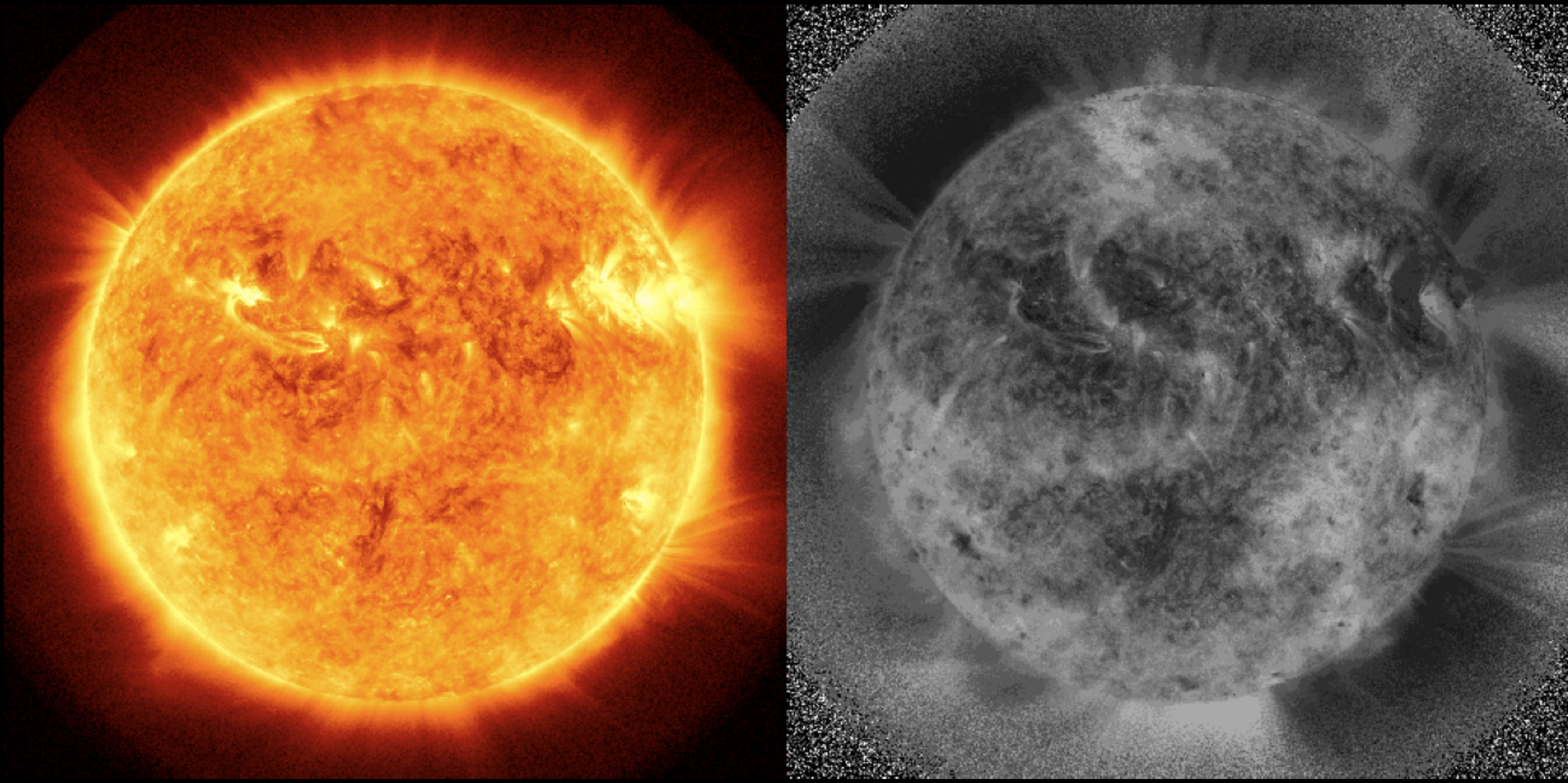
Stein, N.M., et al. 2012, *"H-means Image Segmentation to Identify Solar Thermal Features"*, In IEEE International Conference on Image Processing (ICIP). (student paper award finalist)

Stein, N.M., 2014, *"Detecting Thermal Features in Massive Streams of Solar Images"*, in Big Data in Astro Statistics, Section on Statistical Learning and Data Mining, JSM

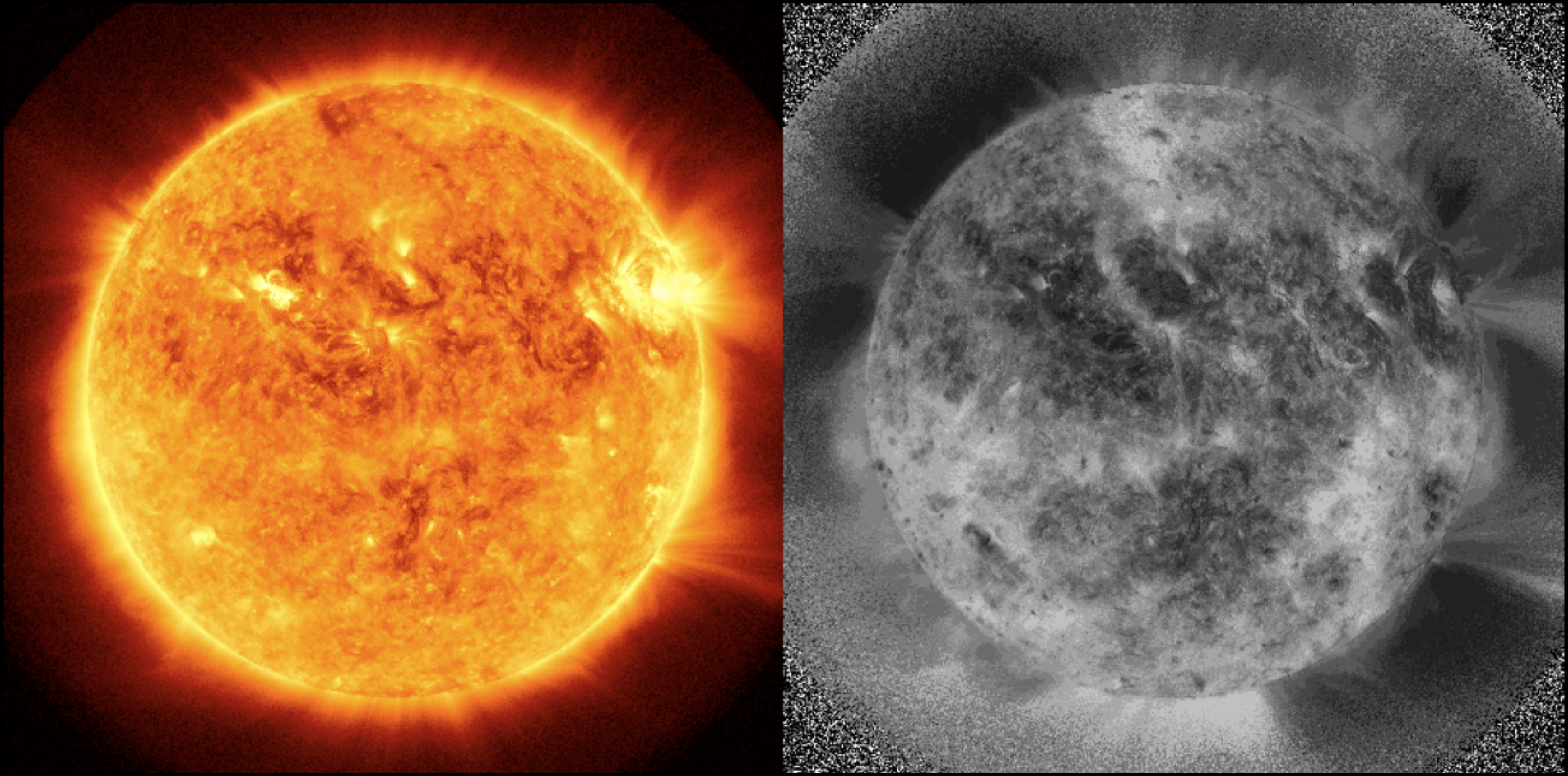

SDO/AIA- 94 20101002_233603


SDO/AIA– 94 2(
SDO/AIA– 335 2
SDO/AIA– 193 2
SDO/AIA– 131 2
SDO/AIA– 211 2
SDO/AIA– 171 20101002_233313


20101002_233313
SDO/AIA– 171 20101002_000037



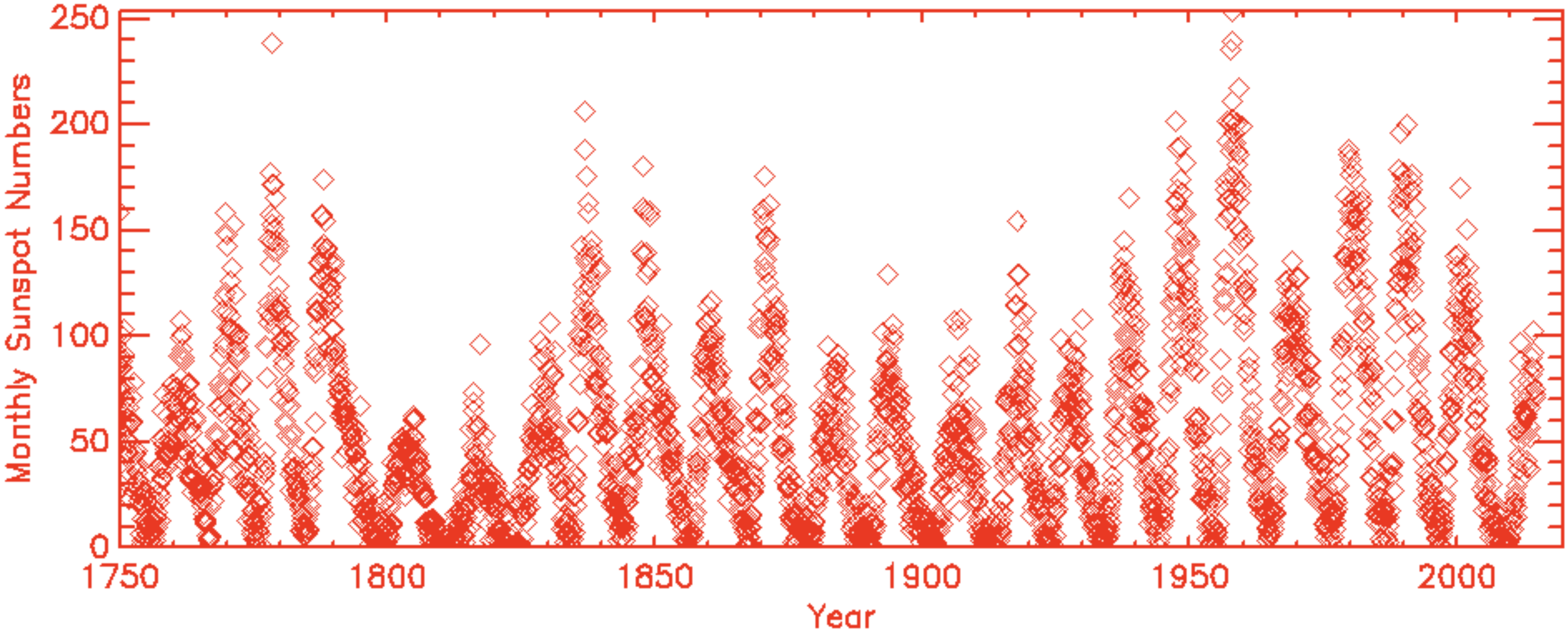
SDO/AIA 2010 Oct 2 05:57



SDO/AIA 2010 Oct 2 18:43

Sunspots: Cycle

Yaming Yu / David Stenning



$$\text{amplitude}_{(\text{next})} \sim 4.1 + 3.9\ (\text{amplitude} / \text{fall time})$$

$$\text{time to maximum}_{(\text{next})} \sim 8.5 - 0.43\ \text{amplitude}_{(\text{next})}$$

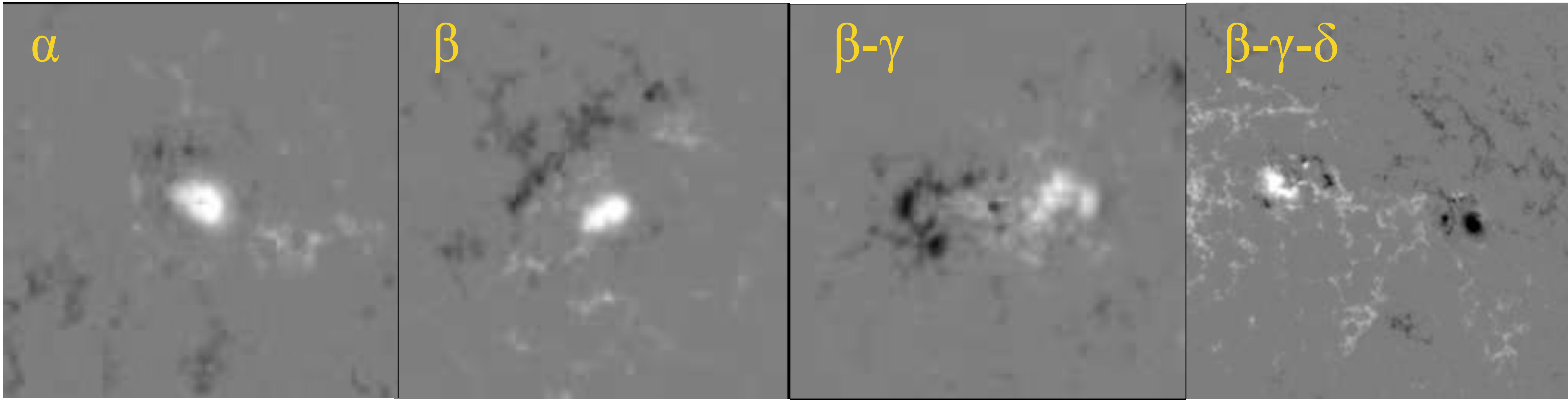
$$\text{fall time}_{(\text{next})} \sim 4.3 + 0.43\ \text{amplitude}_{(\text{next})}$$

Yu, Y., et al. 2012, *"A Bayesian Analysis of the Correlations Among Sunspot Cycles"*, Solar Physics, 281, 847

Stenning, D., et al. 2014, *"A Bayesian Analysis of the Solar Cycle Using Multiple Proxy Variables"*, Current Trends in Bayesian Methodology with Applications, Editors: S. Upadhyay, D.K. Dey, U. Singh and A. Loganathan, Chapman & Hall/CRC Press, in press

Sunspots: Classification

David Stenning

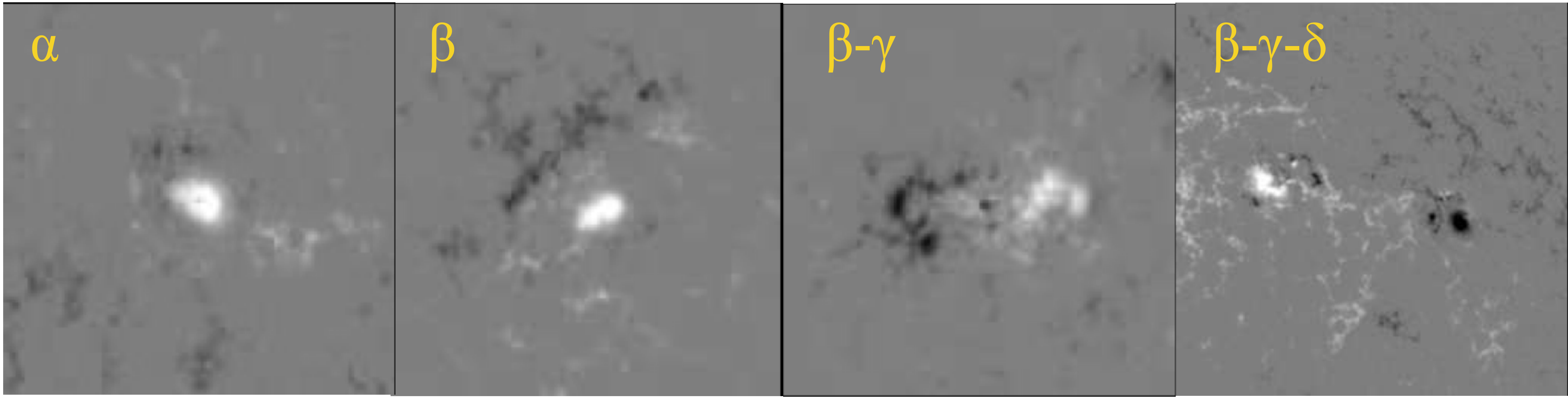


Stenning, D., et al. 2012, *"Morphological Image Analysis and its Application to Sunspot Classification"*, Statistical Challenges in Modern Astronomy V (Editors: G.J. Babu and E.D. Feigelson), Springer Verlag, New York, 2012

Stenning, D., et al., 2013, *"Morphological feature extraction for statistical learning with applications to solar image data"*, in Statistical Analysis and Data Mining, DOI: 10.1002/sam.11200

Sunspots: Classification

David Stenning

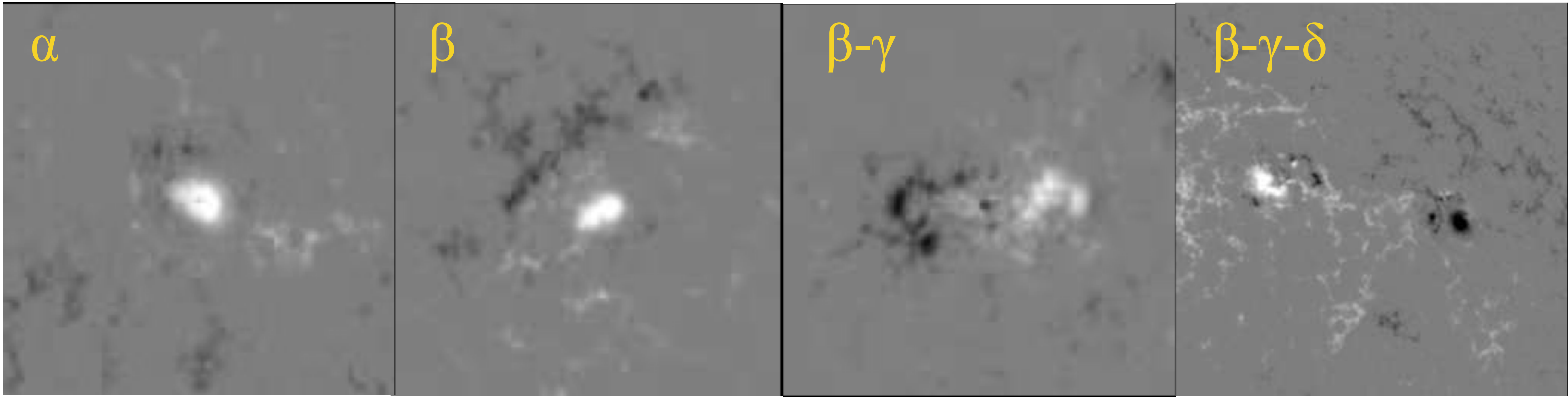


Extract scatter for both polarities, relative strength, curvature of separator, and penumbral polarity overlaps

Stenning, D., et al. 2012, *"Morphological Image Analysis and its Application to Sunspot Classification"*, Statistical Challenges in Modern Astronomy V (Editors: G.J. Babu and E.D. Feigelson), Springer Verlag, New York, 2012
Stenning, D., et al., 2013, *"Morphological feature extraction for statistical learning with applications to solar image data"*, in Statistical Analysis and Data Mining, DOI: 10.1002/sam.11200

Sunspots: Classification

David Stenning



Extract scatter for both polarities, relative strength, curvature of separator, and penumbral polarity overlaps

Table 1. Confusion matrix of the random forest predictions on out-of-bag data.

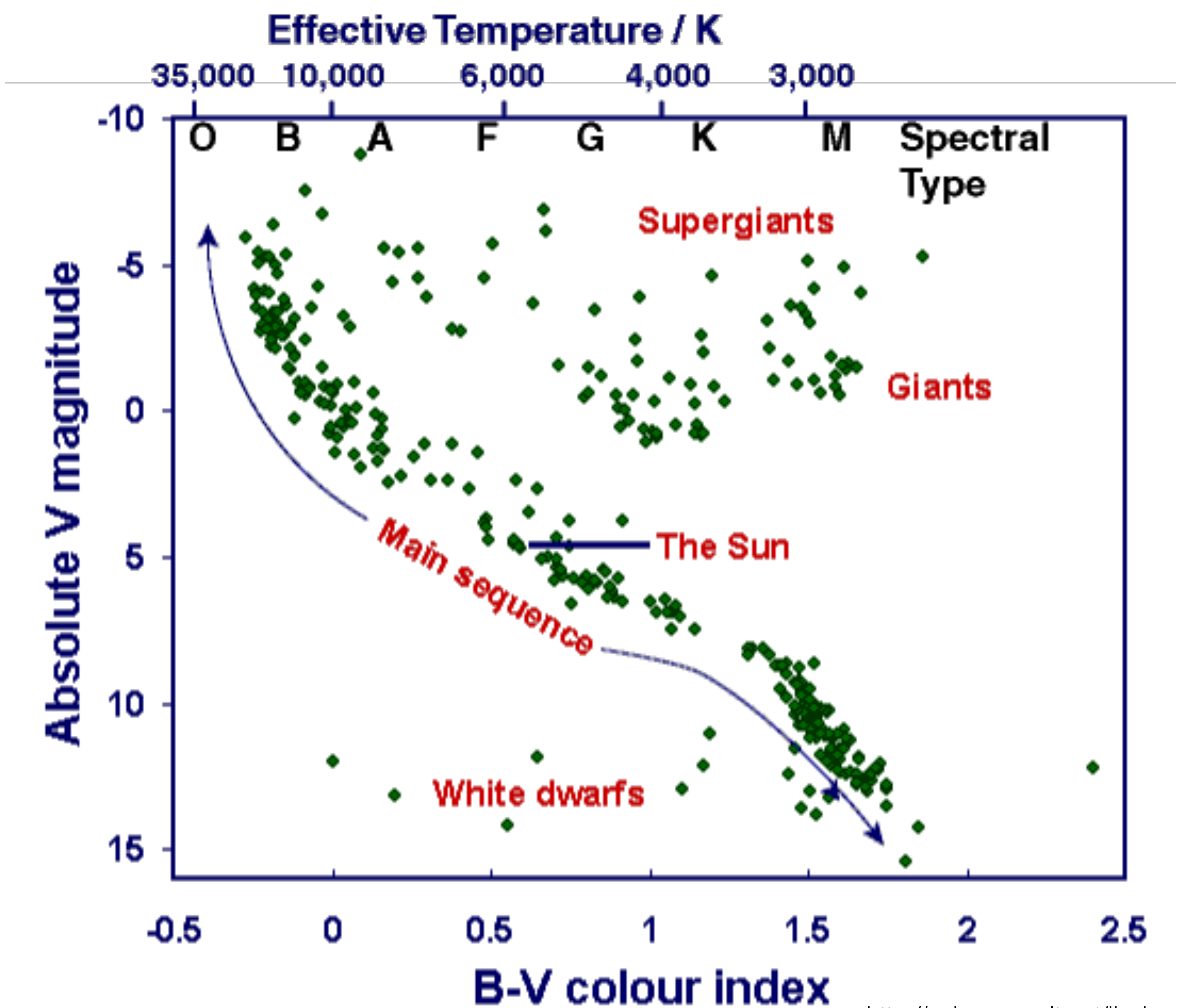
		Manual classification			
		α	β	$\beta\gamma$	$\beta\gamma\delta$
Automatic classification	α	25	1	0	0
	β	2	63	5	0
	$\beta\gamma$	0	1	11	1
	$\beta\gamma\delta$	0	0	2	8

Stenning, D., et al. 2012, *"Morphological Image Analysis and its Application to Sunspot Classification"*, Statistical Challenges in Modern Astronomy V (Editors: G.J. Babu and E.D. Feigelson), Springer Verlag, New York, 2012
Stenning, D., et al., 2013, *"Morphological feature extraction for statistical learning with applications to solar image data"*, in Statistical Analysis and Data Mining, DOI: 10.1002/sam.11200

Color-Magnitude Diagrams

Paul Baines / Nathan Stein

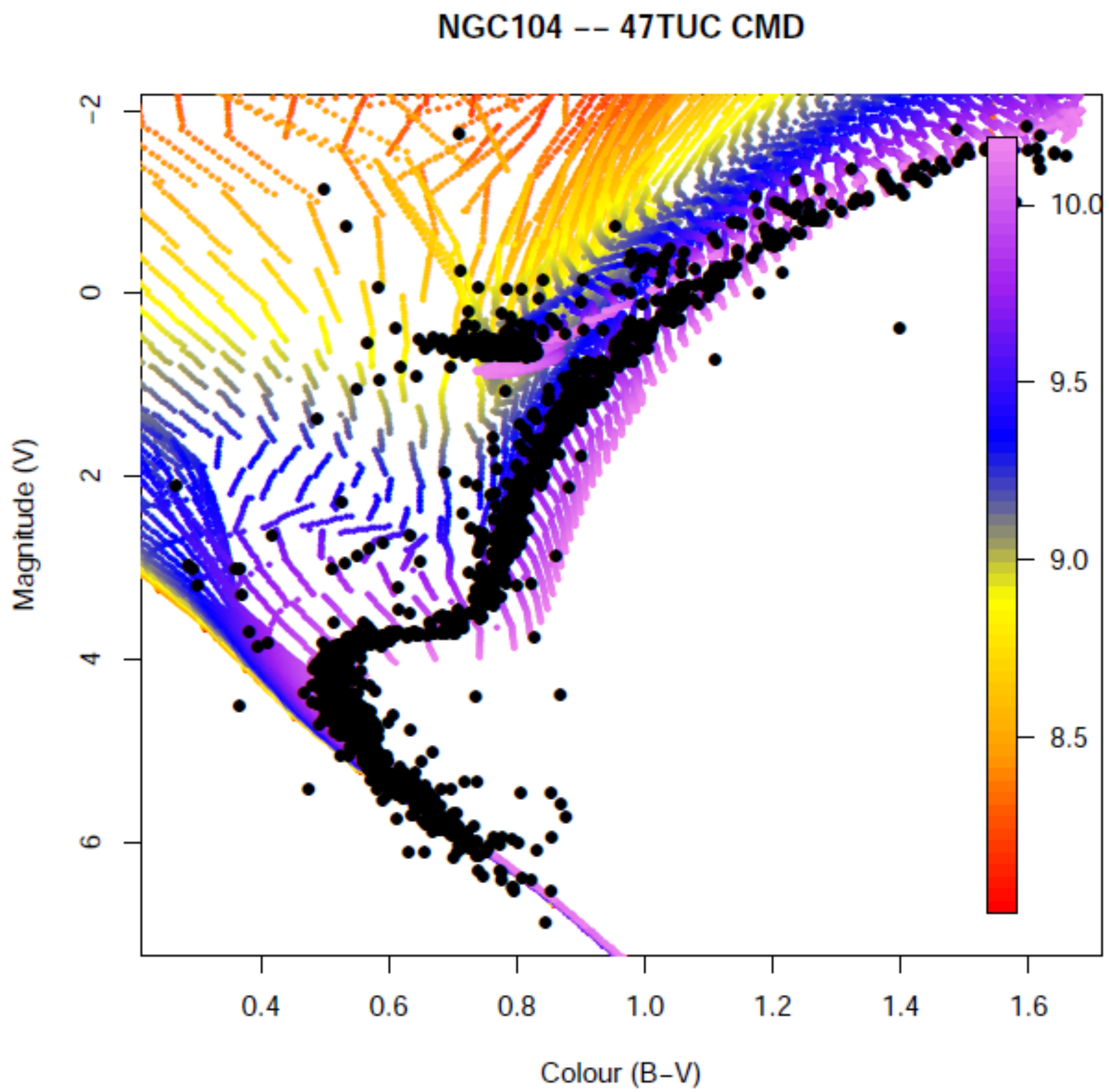
Magnitude = –*log*(brightness) (relative to a standard)
Color = difference in Magnitudes at different wavelengths

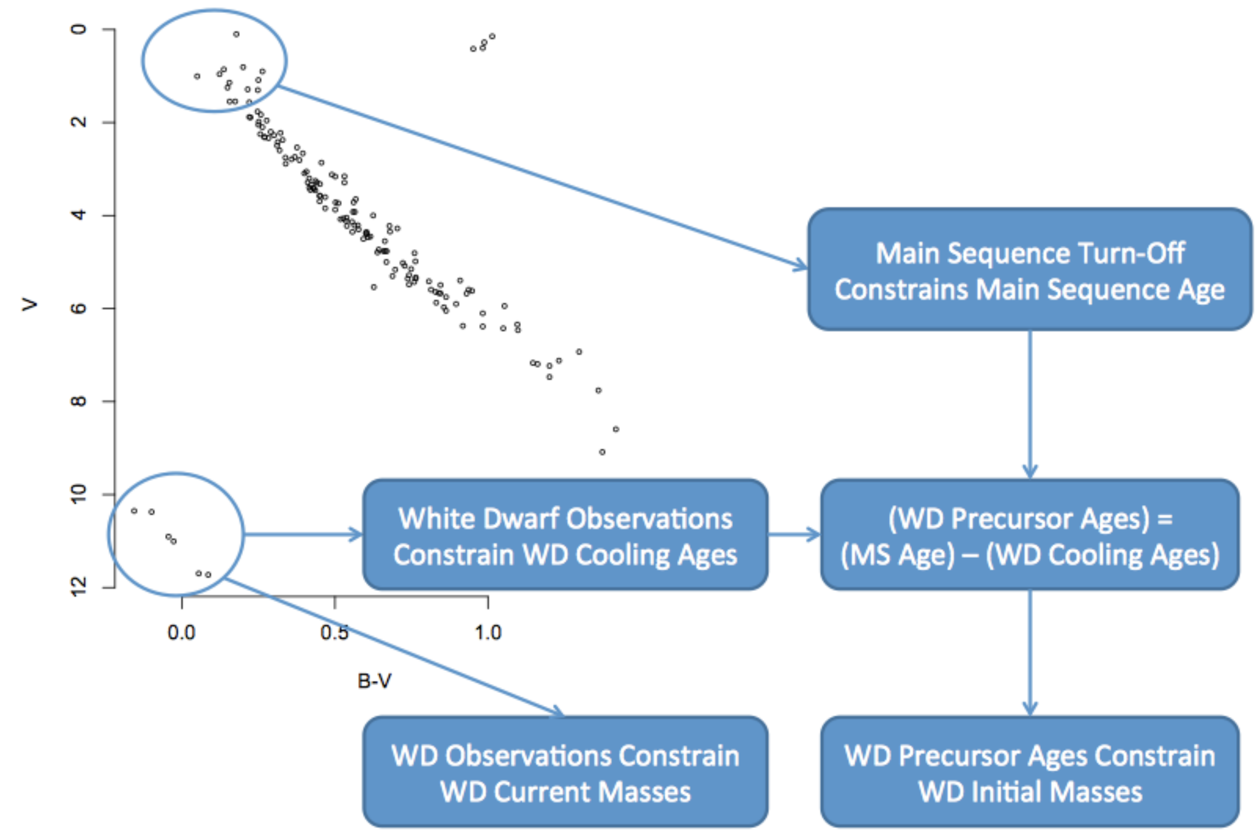


http://sciencevault.net/ibphysics/astrophysics/pics/hrdiagram1.gif

Color-Magnitude Diagrams

Paul Baines / Nathan Stein

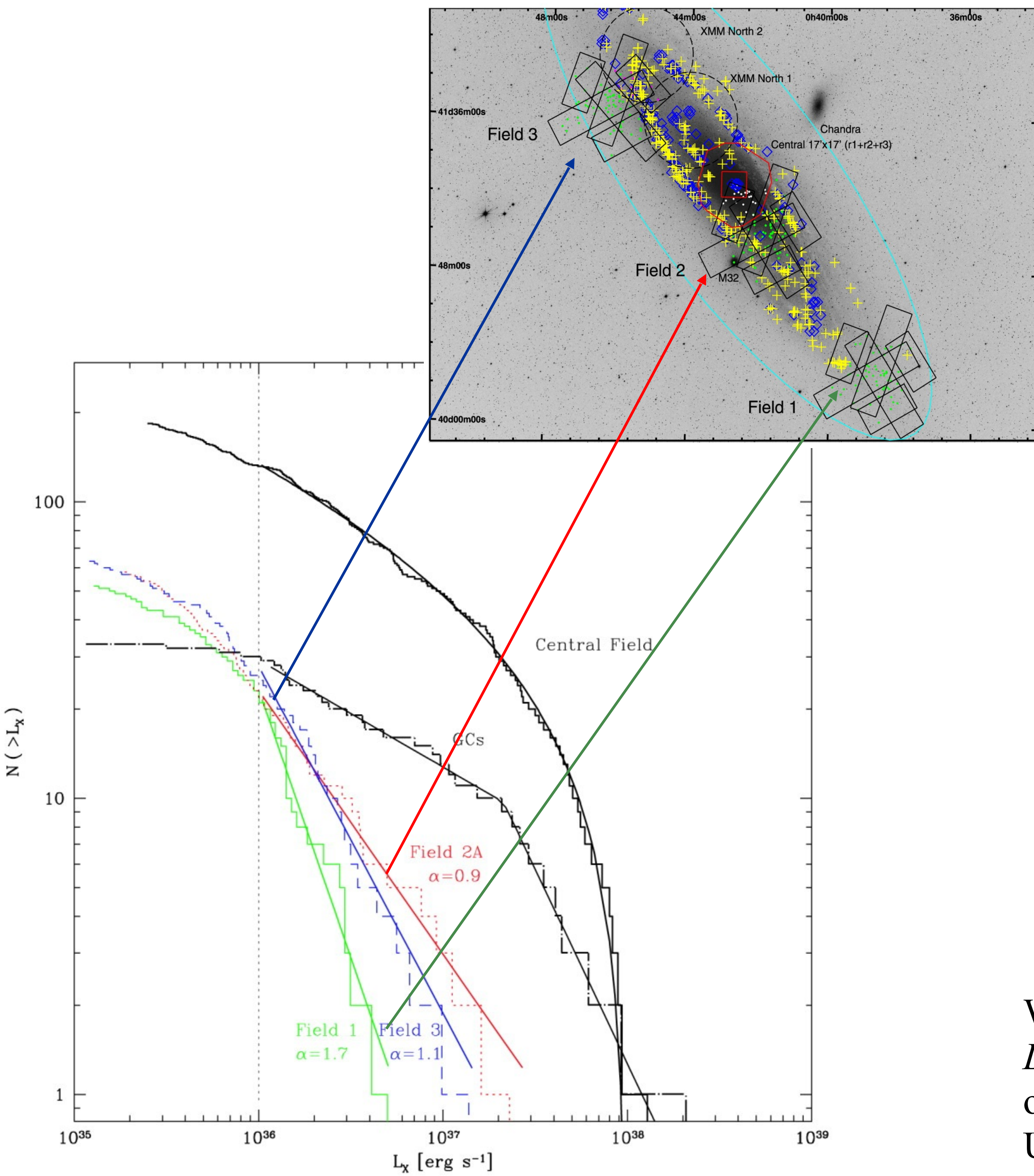




DeGennaro, S., et al. 2008, *"Inverting Color-Magnitude Diagrams to Access Precise Star Cluster Parameters: A New White Dwarf Age for the Hyades"*, ApJ, 696, 12
van Dyk, D.A., et al. 2009, *"Statistical Analysis of Stellar Evolution"*, Annals of Applied Statistics, 3, 117
Jeffery, E.J., et al., 2011, *"The White Dwarf Age of NGC 2477"*, ApJ, 730, 35
Stein, N.M., et al. 2013, *"Combining Computer Models to Account for Mass Loss in Stellar Evolution"*, Statistical Analysis and Data Mining, 6, 34

log*N*-log*S*

Nondas Sourlas / Paul Baines / Irina Udaltsova / Raymond Wong



Kong et al. 2003, via A. Zezas

$$\frac{dN}{dS} = KS^{-\alpha-1}$$

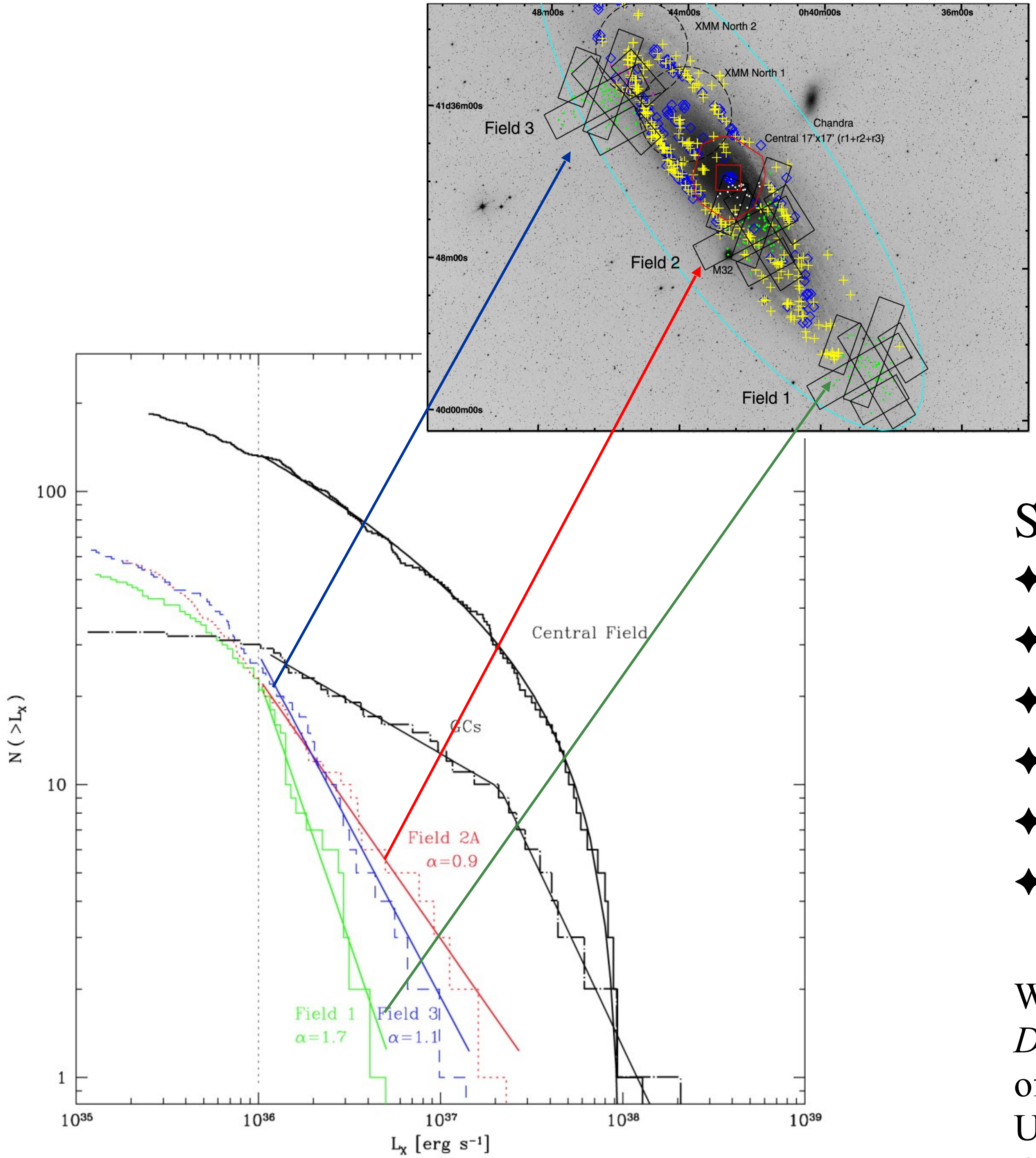
$$N(>S) = KS^{-\alpha}$$

$$N(>S) = \begin{cases} K_1 S^{-\alpha_1} & S \geq S_b \\ K_2 S^{-\alpha_2} & S_0 < S < S_b \end{cases}$$

Wong, R., et al. 2014, *"Automatic Estimation of Flux Distributions of Astrophysical Source Populations"*, Annals of Applied Statistics, in press
Udaltsova, I., et al., 2011, *"log(N)-log(S): A Measuring Stick for the Universe"*, SCMA V
Udaltsova, I., et al., 2014a,b, in preparation

log$N$-log$S$

Nondas Sourlas / Paul Baines / Irina Udaltsova / Raymond Wong



Kong et al. 2003, via A. Zezas

$$\frac{dN}{dS} = KS^{-\alpha-1}$$

$$N(>S) = KS^{-\alpha}$$

$$N(>S) = \begin{cases} K_1 S^{-\alpha_1} & S \geq S_b \\ K_2 S^{-\alpha_2} & S_0 < S < S_b \end{cases}$$

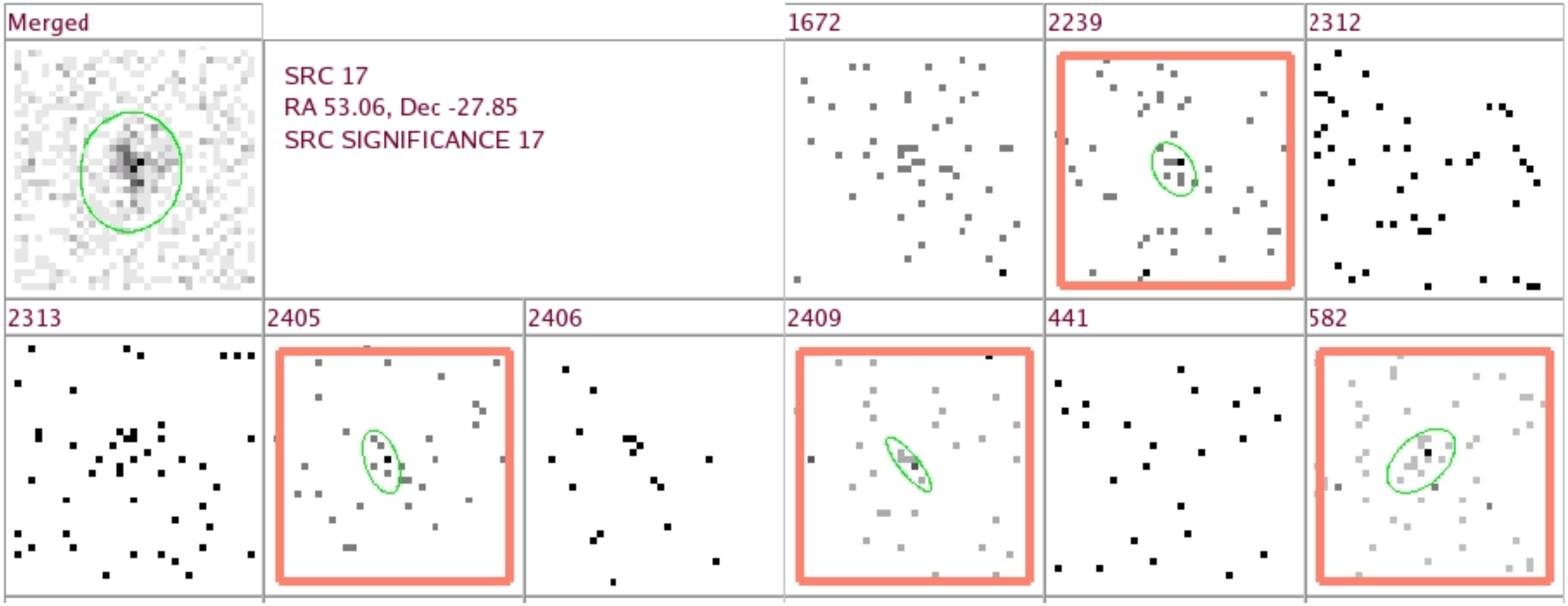
Some recent capabilities:

✦ ability to determine number of segments
✦ hierarchical Bayesian modeling
✦ allow for detection efficiency
✦ multiple-Pareto models
✦ posterior predictive p-value checks
✦ base sensitivity limit

Wong, R., et al. 2014, *"Automatic Estimation of Flux Distributions of Astrophysical Source Populations"*, Annals of Applied Statistics, in press
Udaltsova, I., et al., 2011, *"log(N)-log(S): A Measuring Stick for the Universe"*, SCMA V
Udaltsova, I., et al., 2014a,b, in preparation

Luminosity Functions

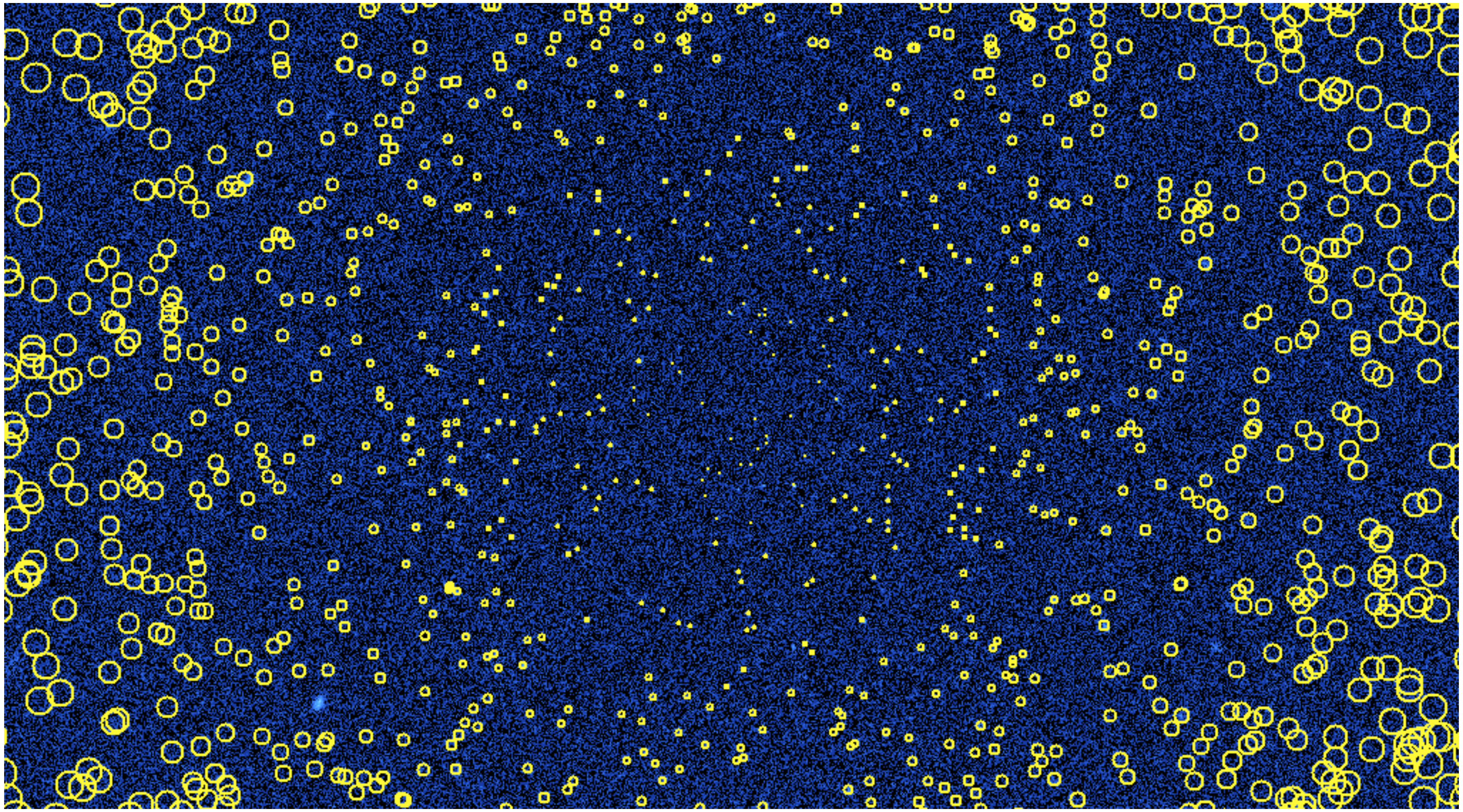
Alex Blocker



Blocker et al. 2009*, X-ray Stacking for the Analysis of Faint Sources: A Bayesian Alternative,*
Proceedings of the conference held 22-25 September, 2009 in Boston, Chandra

Luminosity Functions

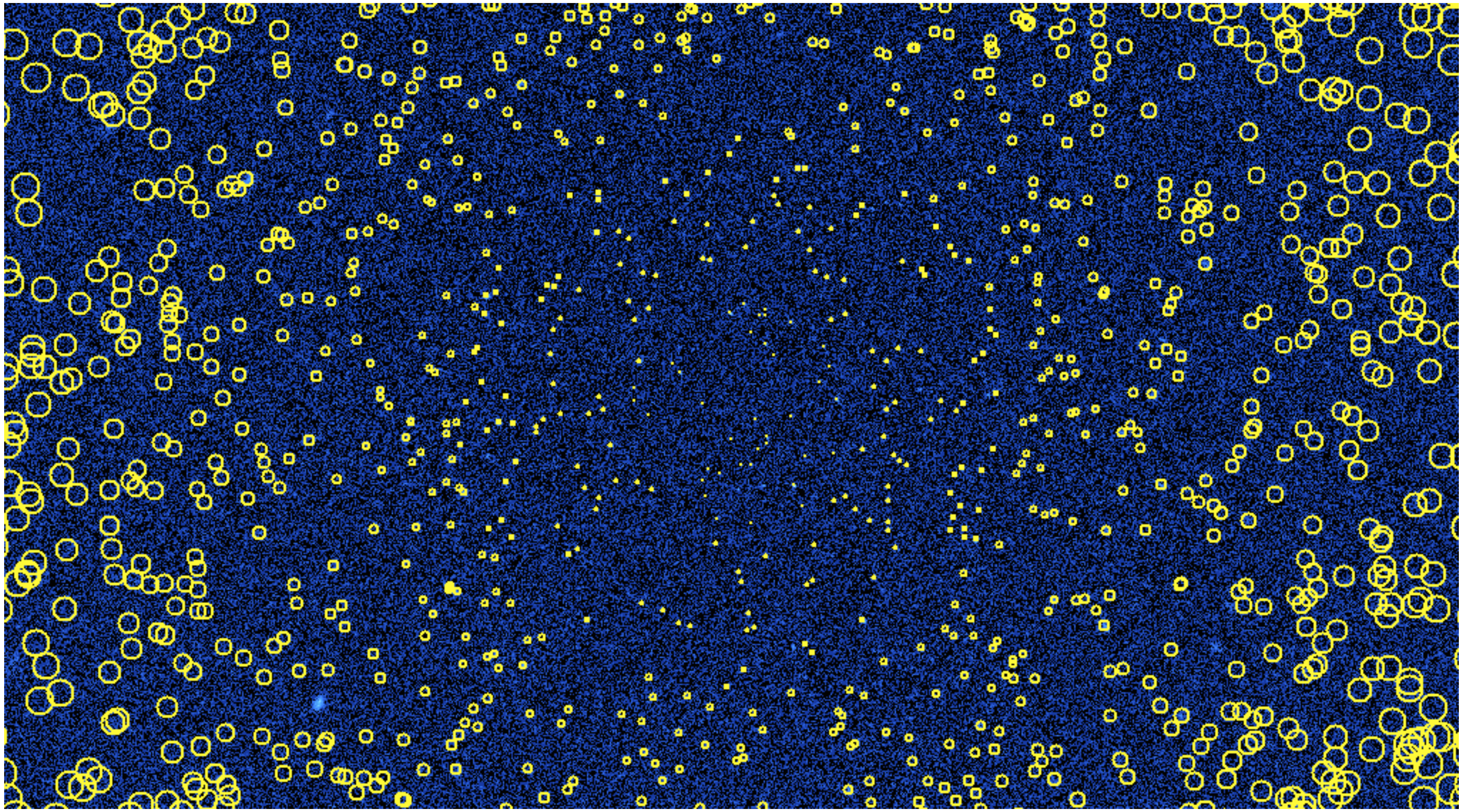
Lazhi Wang



Wang, L., et al. 2014, *"Catalog-based X-ray Population Modeling"*, in preparation

Luminosity Functions

Lazhi Wang



∃ independently derived catalogs of sources (from optical/IR/radio observations)

Question: Can we bypass X-ray detection process to infer properties of the underlying luminosity function? Are detectable X-ray sources representative of the whole sample?

Wang, L., et al. 2014, *"Catalog-based X-ray Population Modeling"*, in preparation

Luminosity Functions

Lazhi Wang

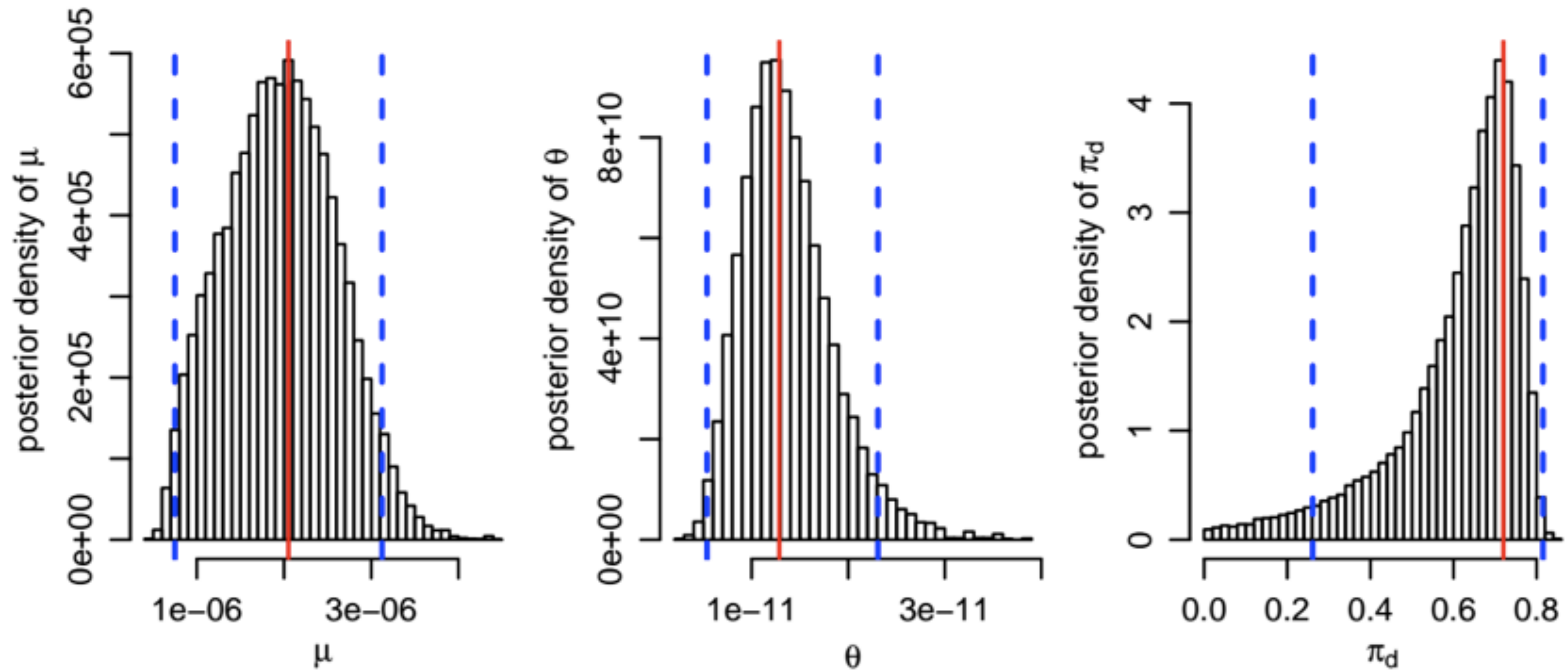
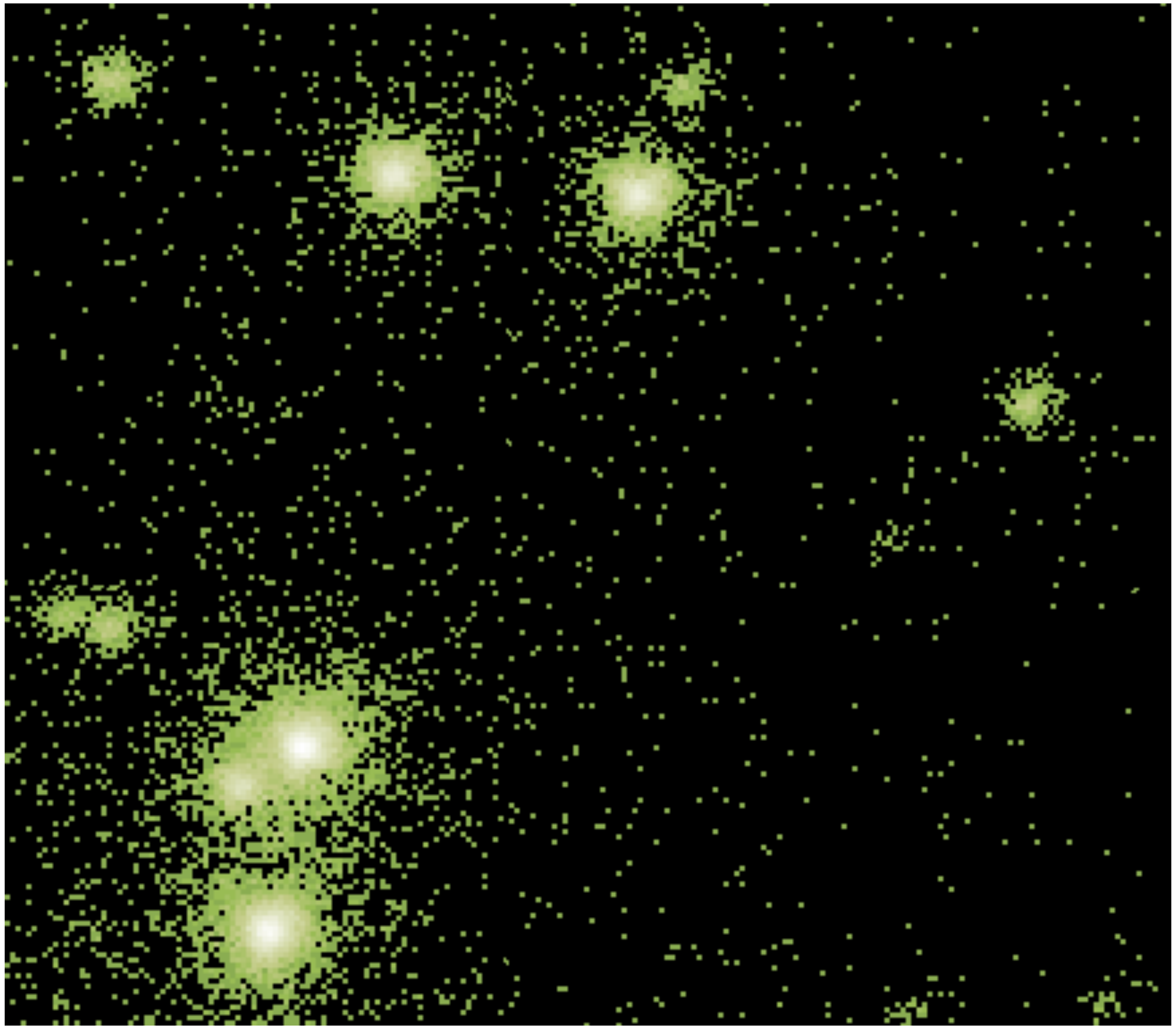


Figure 9: The histograms of the posterior draws of μ (left), θ (middle) and π_d (right). The dataset includes only the 649 sources within 6-arcmin from the center of the field. The red solid lines are the posterior mode estimators and the blue dash lines show the lower and upper bound of the 95% HPDI of the parameters.

Wang, L., et al. 2014, *"Catalog-based X-ray Population Modeling"*, in preparation

Spatio-Spectral Disentangling

David Jones

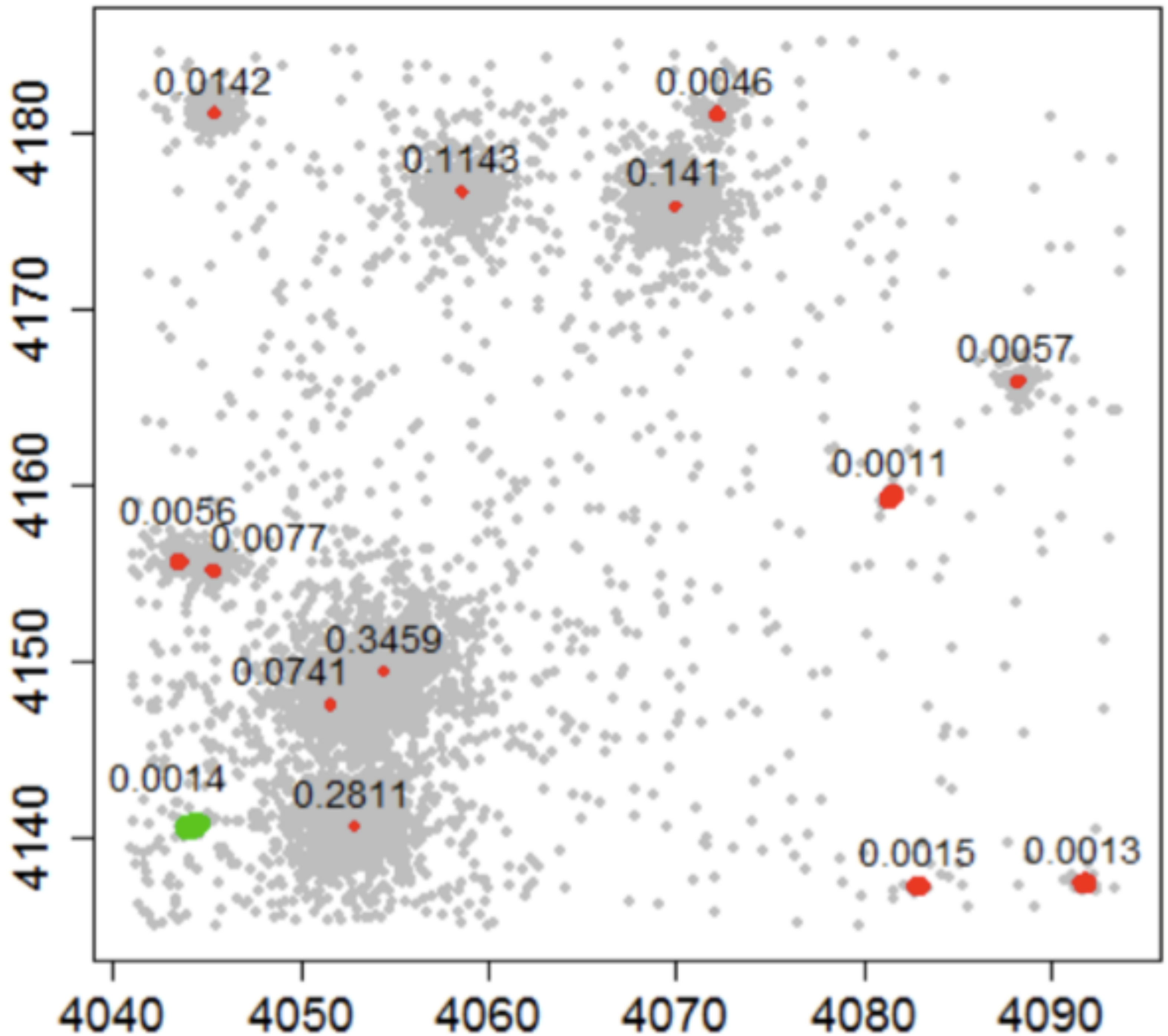


Orion Nebula Cluster
ObsID 1522

Jones, D., et al. 2014, *"Disentangling Overlapping Astronomical Sources Using Spatial and Spectral Information"*, in preparation

Spatio-Spectral Source Disentangling

David Jones



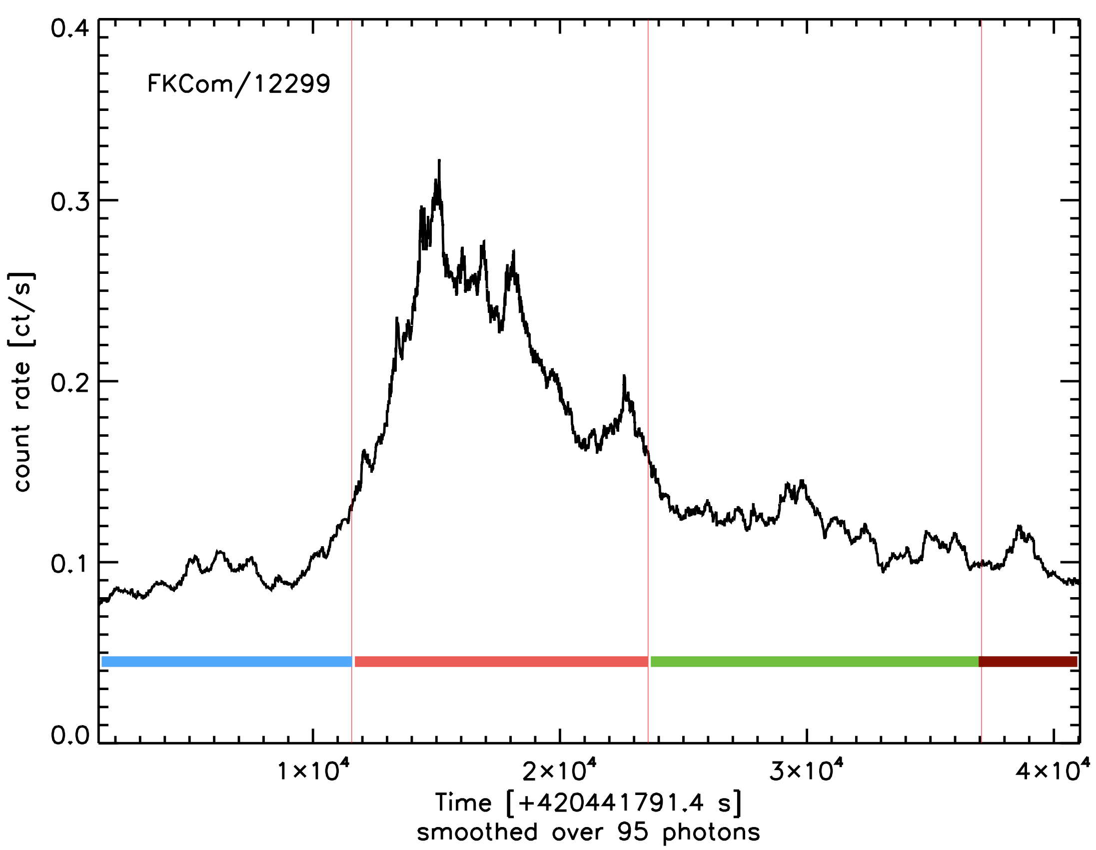
Orion Nebula Cluster
ObsID 1522

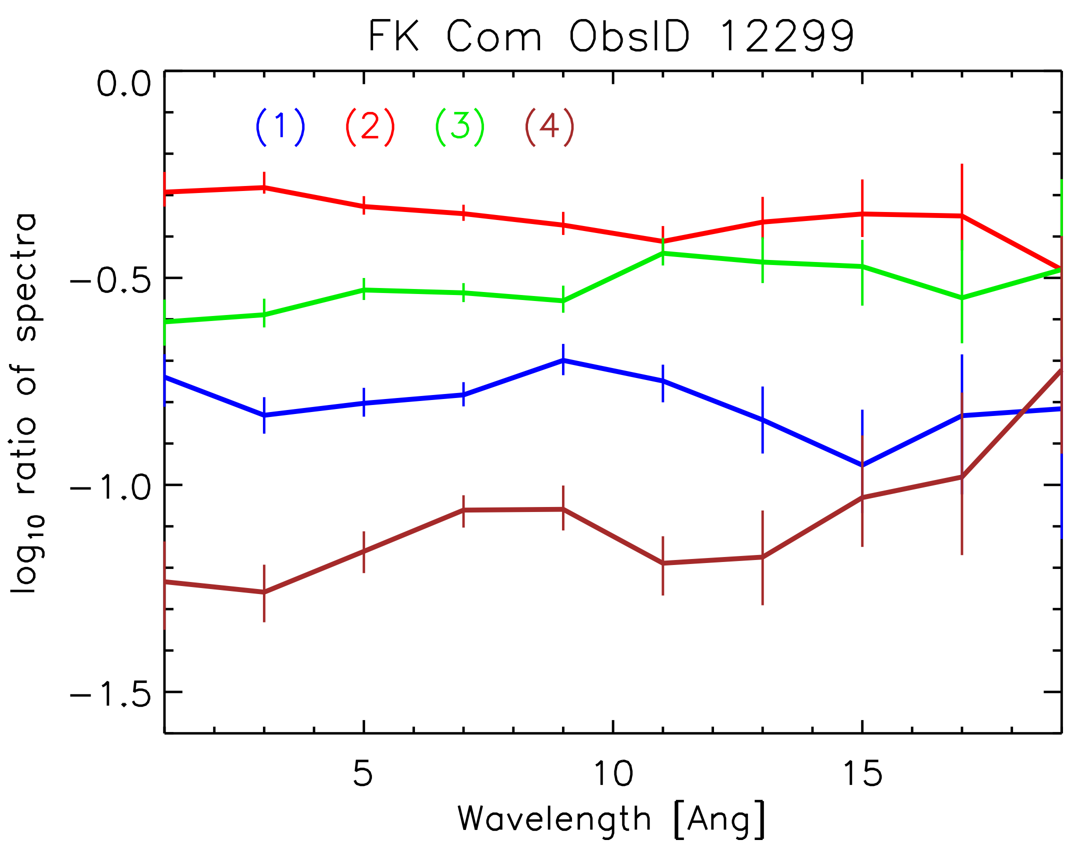
Jones, D., et al. 2014, *"Disentangling Overlapping Astronomical Sources Using Spatial and Spectral Information"*, in preparation

Spectro-temporal Partitioning

Raymond Wong

With high resolution available in both spectral and temporal regimes, compute a 2D Bayesian Blocks like partitioning of the data. Segment data at points where both intensity and spectral shape change significantly. Fit lines+continuum model to spectra in small time bins, compute likelihood that fitted spectra are different, group time bins with similar spectra.

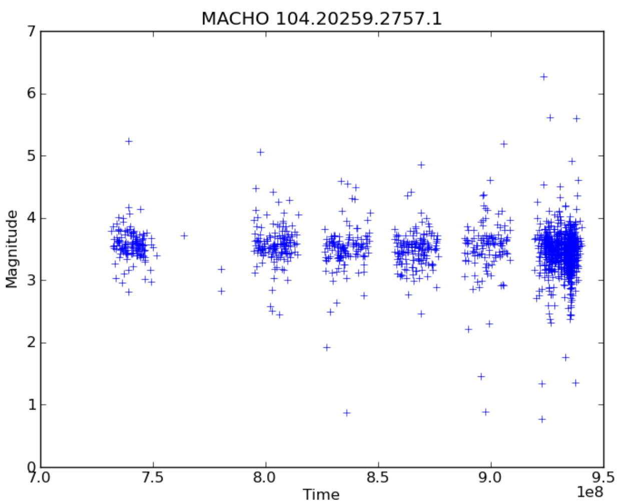


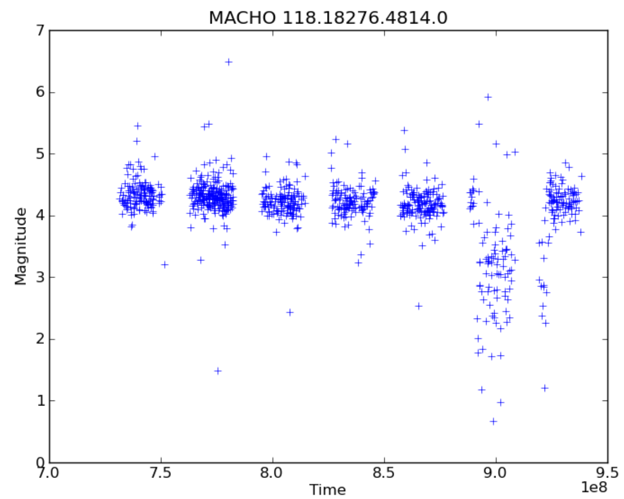


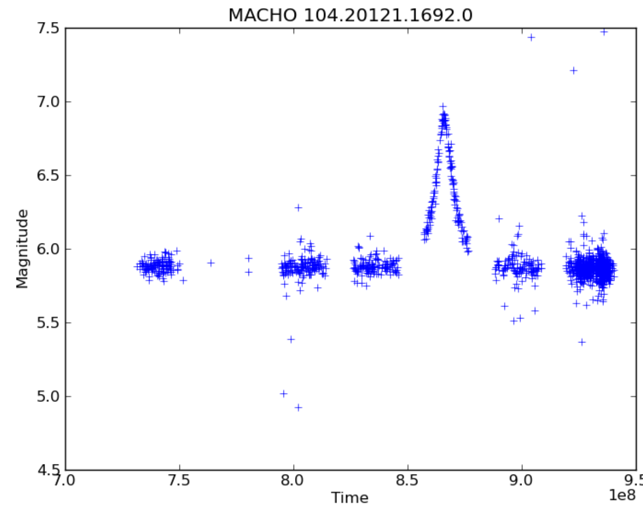
Wong, R., et al. 2014, *"Detecting Abrupt Changes in the Spectra of High-Energy Astrophysical Sources"*, in preparation

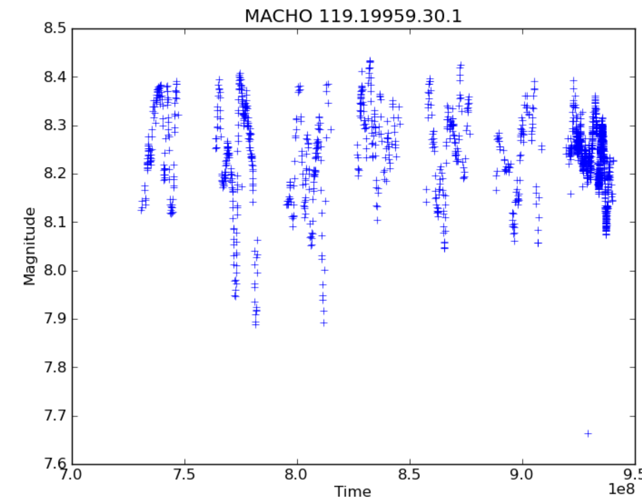
Event Detection in Time Series

Alex Blocker





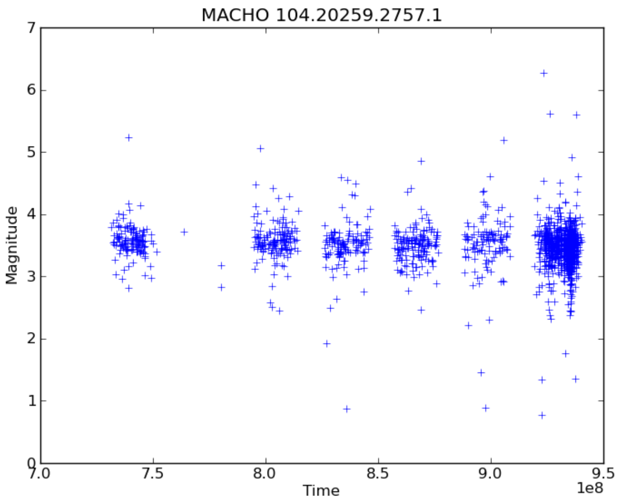


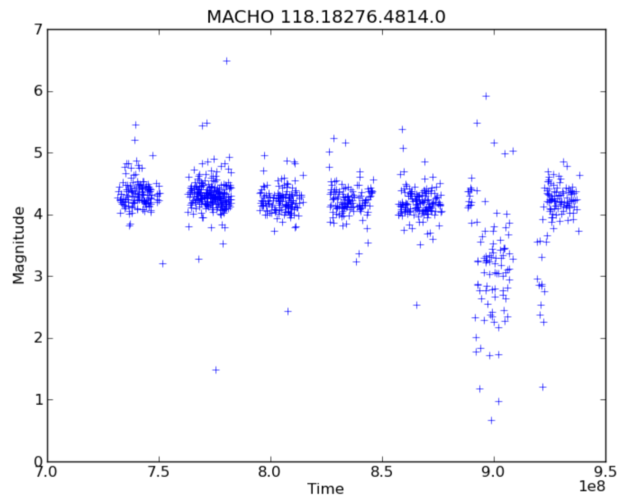


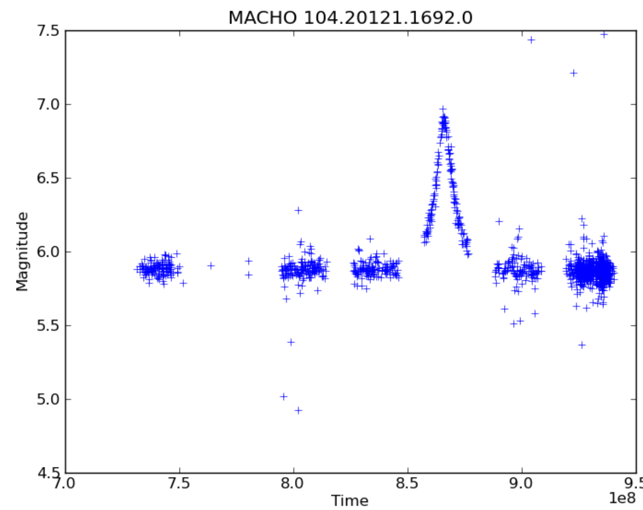
Blocker, A., and Protopapas, P., 2013, *"Semi-parametric Robust Event Detection for Massive Time-Domain Databases"*, SCMA V, Springer-Verlag, p177; http://arxiv.org/abs/1301.3027

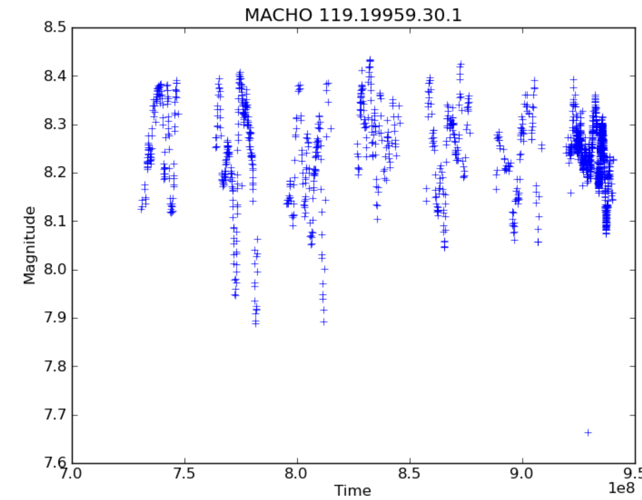
Event Detection in Time Series

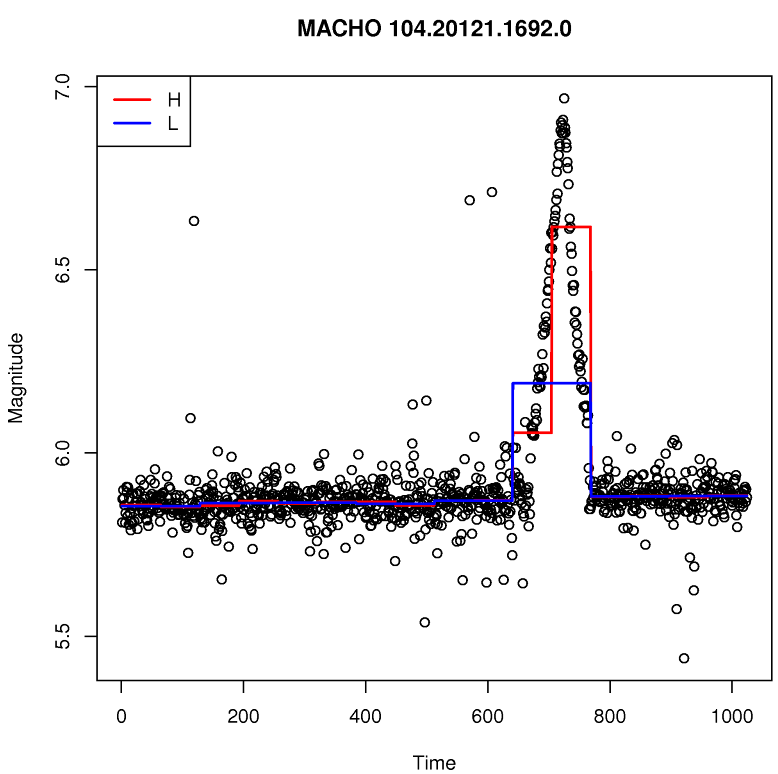
Alex Blocker







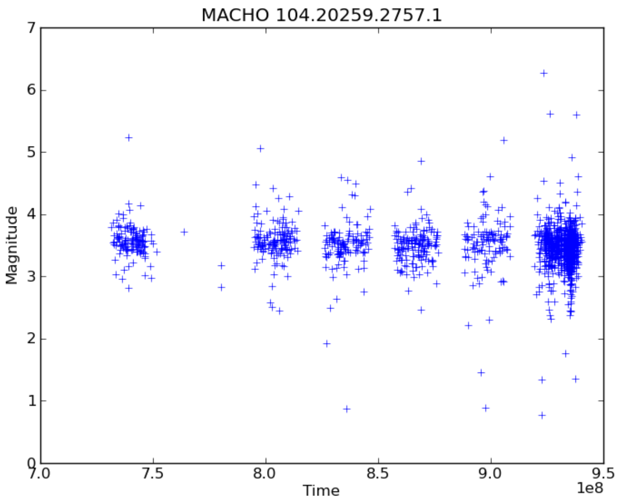


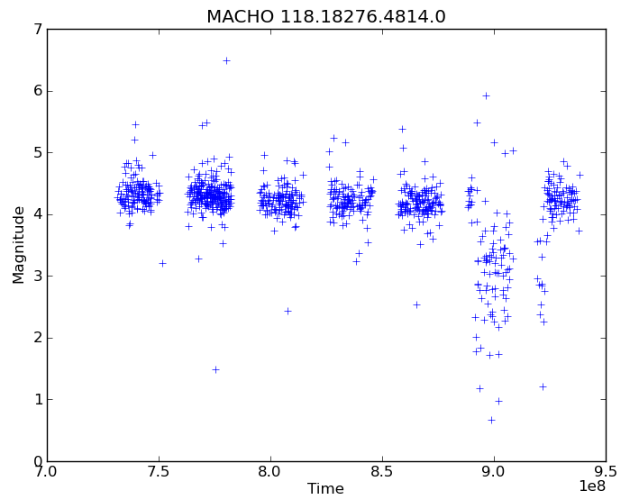


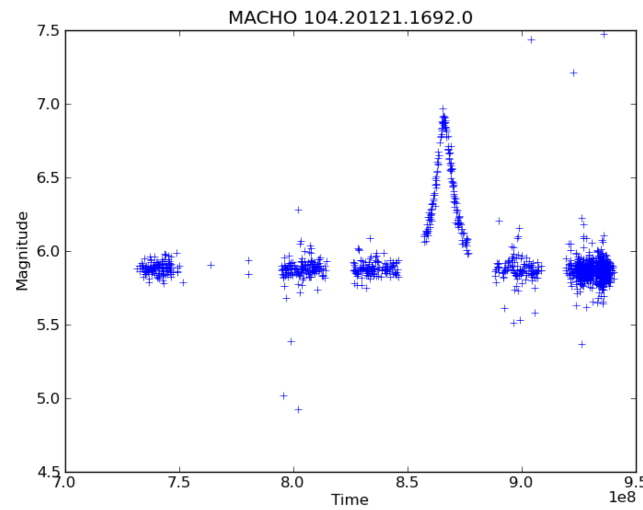
Blocker, A., and Protopapas, P., 2013, *"Semi-parametric Robust Event Detection for Massive Time-Domain Databases"*, SCMA V, Springer-Verlag, p177; http://arxiv.org/abs/1301.3027

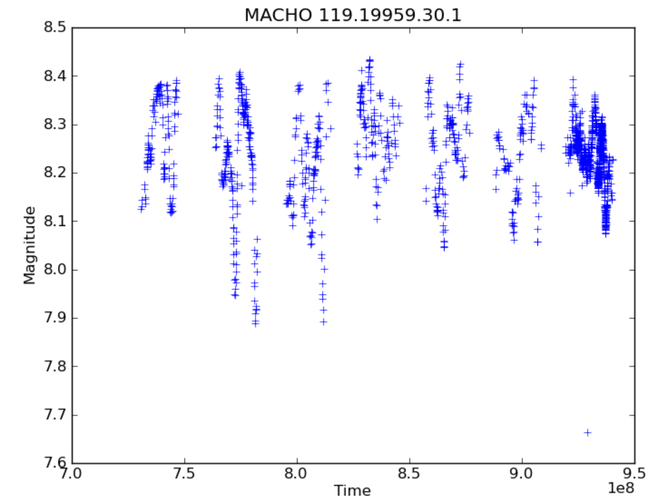
Event Detection in Time Series

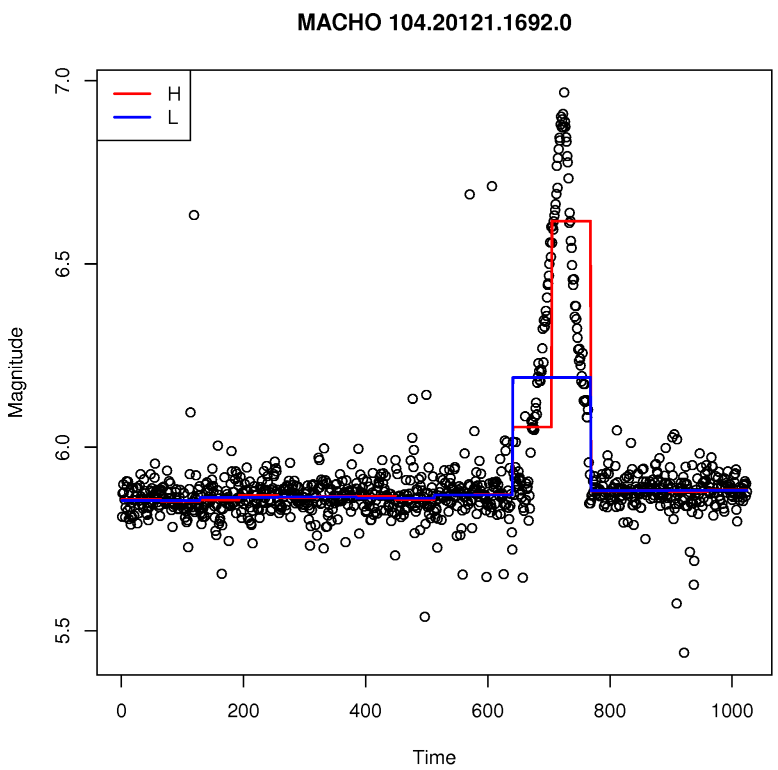
Alex Blocker











Multi-stage analysis, first with simpler tools, then with model checking and uncertainty assessment.
Separate medium-frequency "candidates" from low-frequency trends and high-frequency noise via wavelet-based statistical fitting.

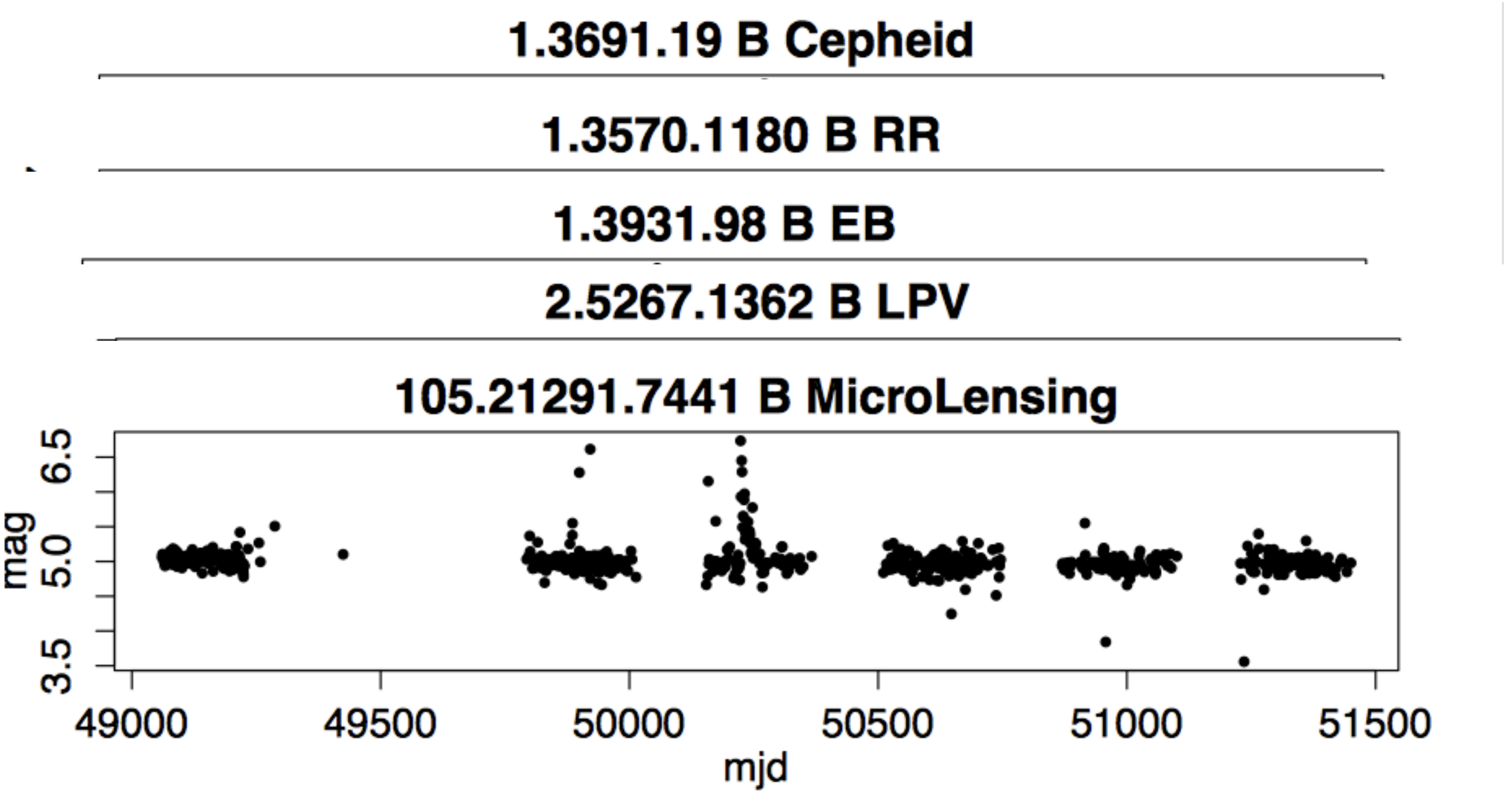
Blocker, A., and Protopapas, P., 2013, *"Semi-parametric Robust Event Detection for Massive Time-Domain Databases"*, SCMA V, Springer-Verlag, p177; http://arxiv.org/abs/1301.3027

Real-time Light Curve Classification

Dan Cervone

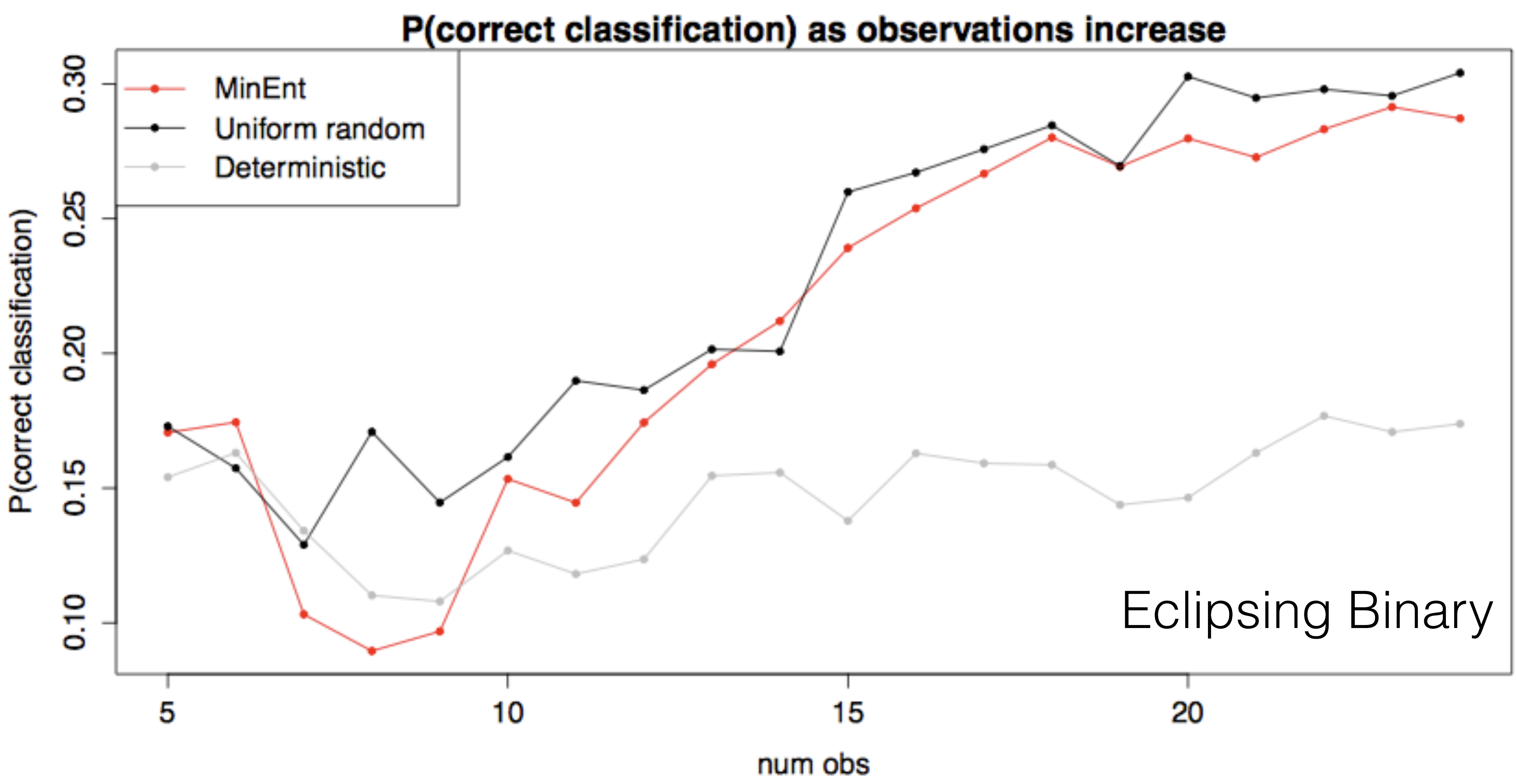
Real-time Light Curve Classification

Dan Cervone



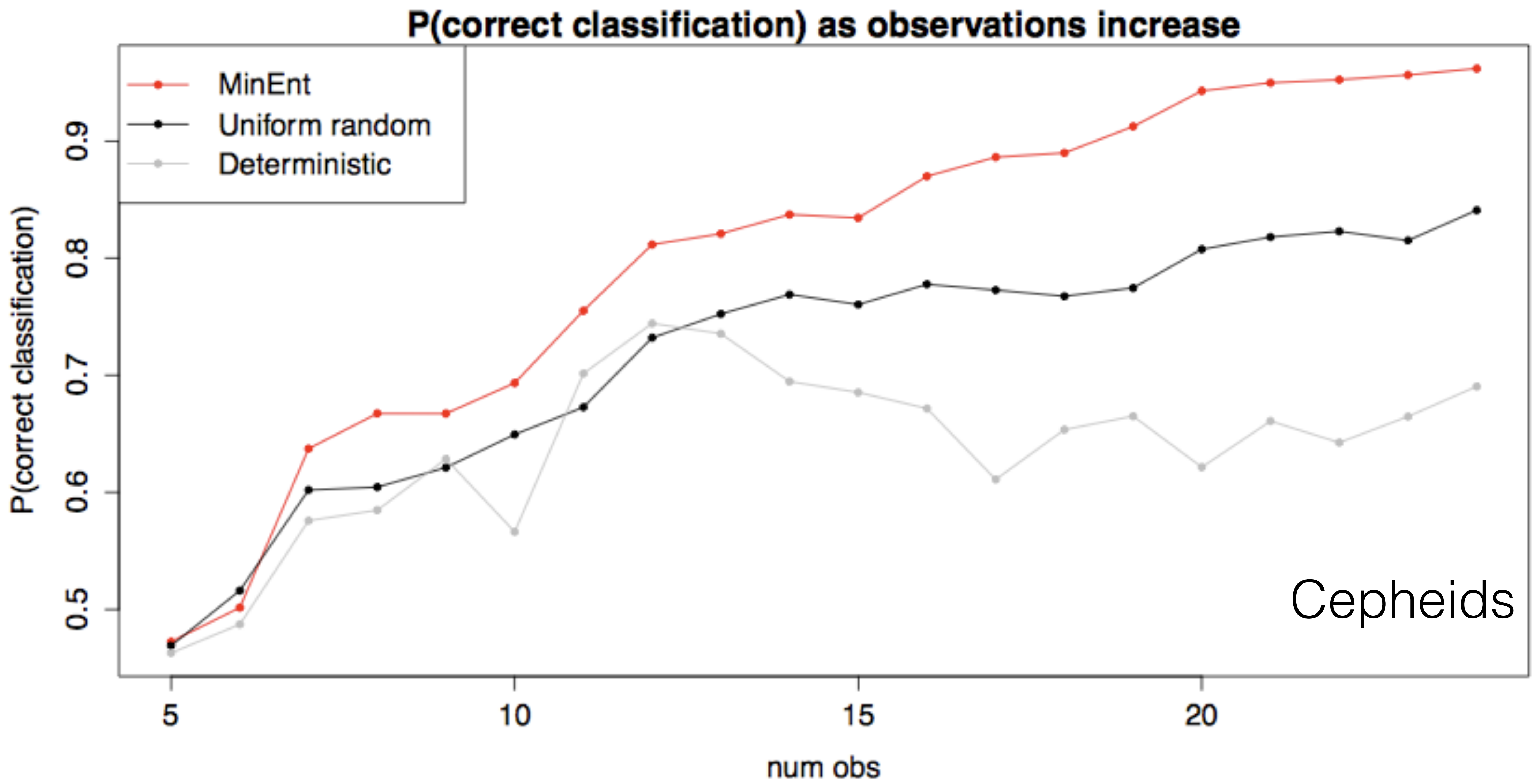
Real-time Light Curve Classification

Dan Cervone



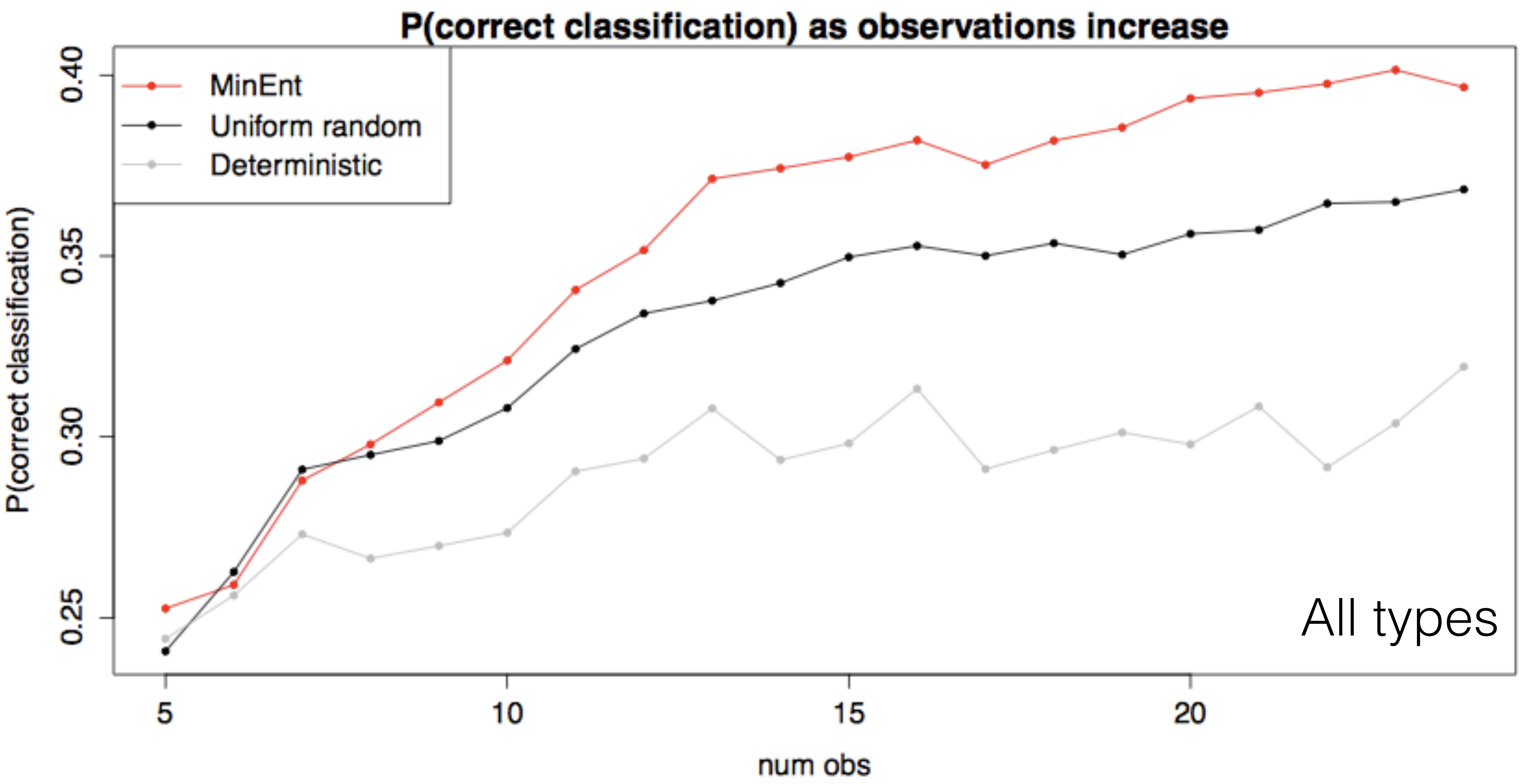
Real-time Light Curve Classification

Dan Cervone



Real-time Light Curve Classification

Dan Cervone



Challenges
Banff Higgs
Paul Edlefsen

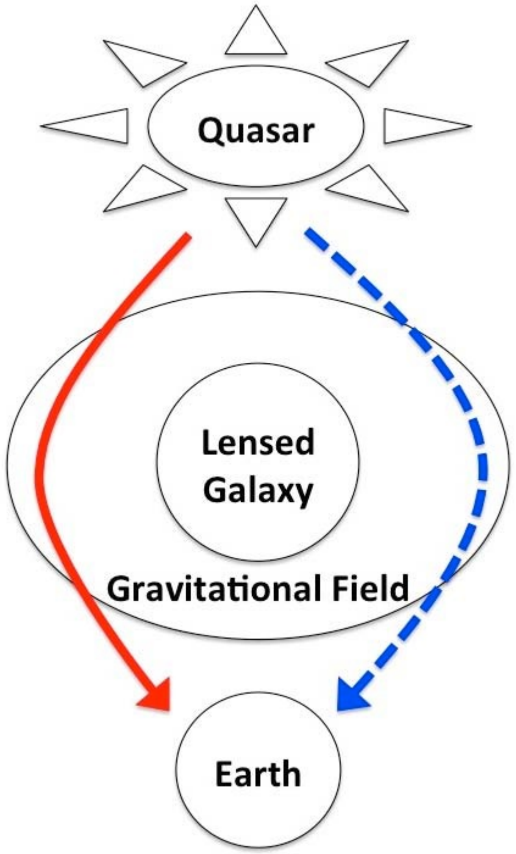
Introduction. This article addresses the problem of estimating **a Poisson rate in the presence of additive and multiplicative noise**, when additional measurements provide data for estimating the nuisance parameters. The problem of estimating rates from noisy or censored Poisson counting data arises in various subdisciplines of physics. In astrophysics, for example, the counts are photons emitted from a distant source. In particle accelerator experiments, the counts are indirect measurements of the number of particles produced by a high-energy collision. In such contexts the observed counts typically include some additive background noise, such as ambient particles. In many cases there is also multiplicative noise, caused, for example, by photon censoring or particle decay, which further complicates the process of estimating the rate of interest.

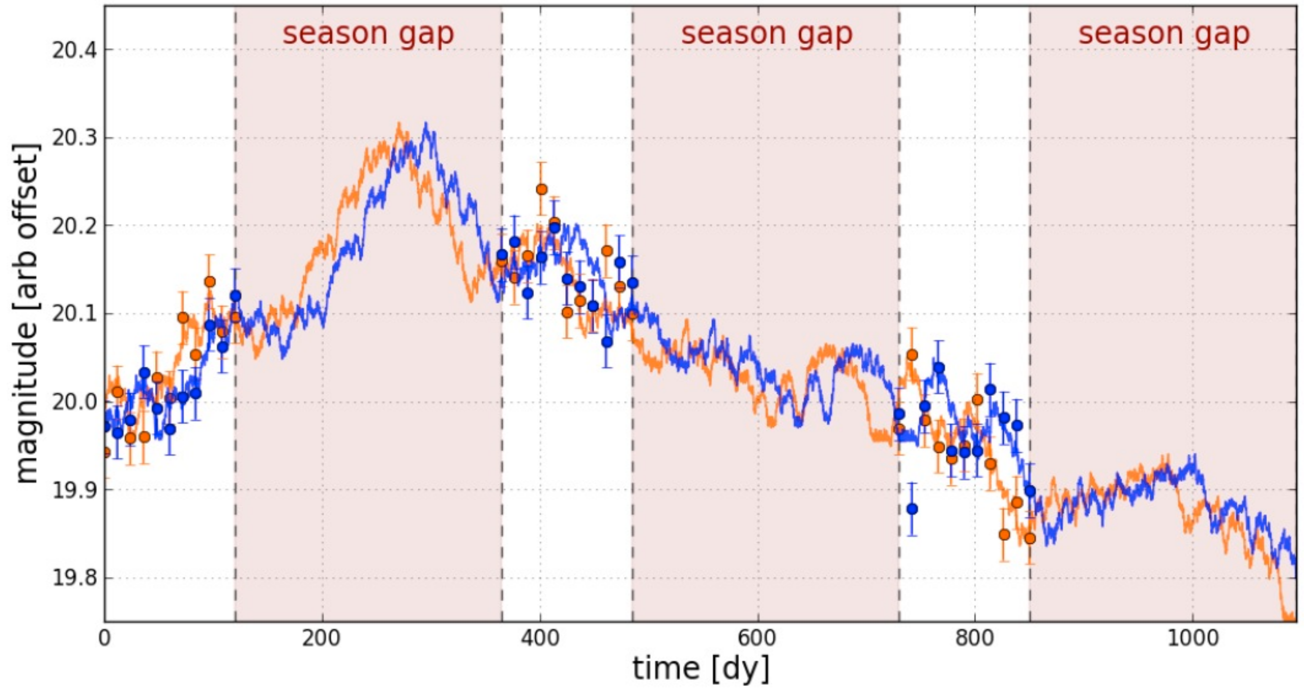
$$n \sim Pois(\varepsilon \cdot s + b)$$
$$y \sim Pois(t \cdot b)$$
$$z \sim Pois(u \cdot \varepsilon)$$

Edlefsen Liu, Dempster, *"Estimating Limits from Poisson counting data using Dempster–Shafer analysis", 2009 The Annals of Applied Statistics, Vol.3, 764, astro-ph 0812.1690*

Challenges
Strong Lens Time Delay Challenge

Hyungsuk Tak





Tak et al 2014 *Bayesian Approach to Time Delay Estimation*