

BXA

robust parameter estimation & practical model comparison



Johannes Buchner

FONDECYT fellow

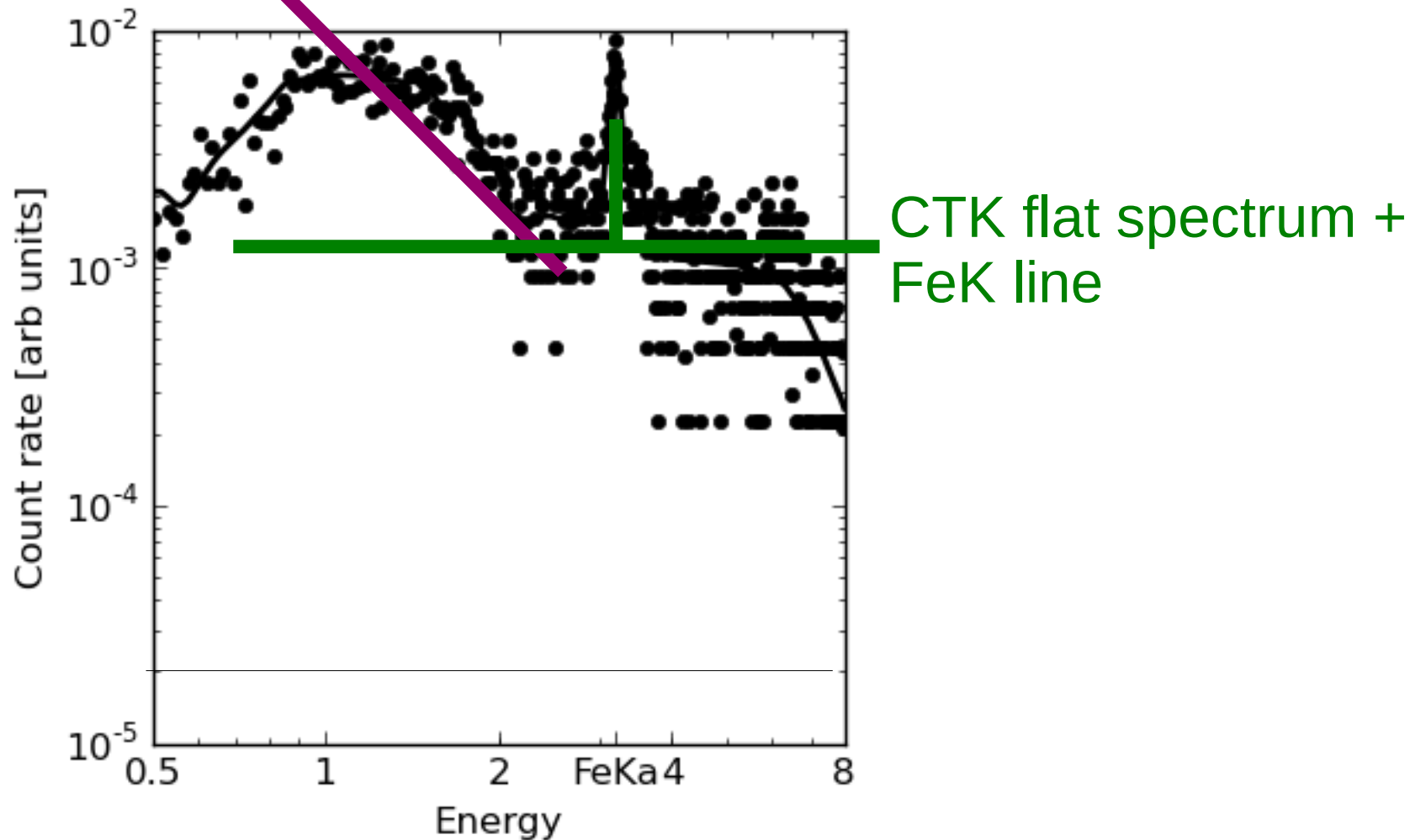
<http://astrost.at/istics/>

BXA – Bayesian X-ray Analysis

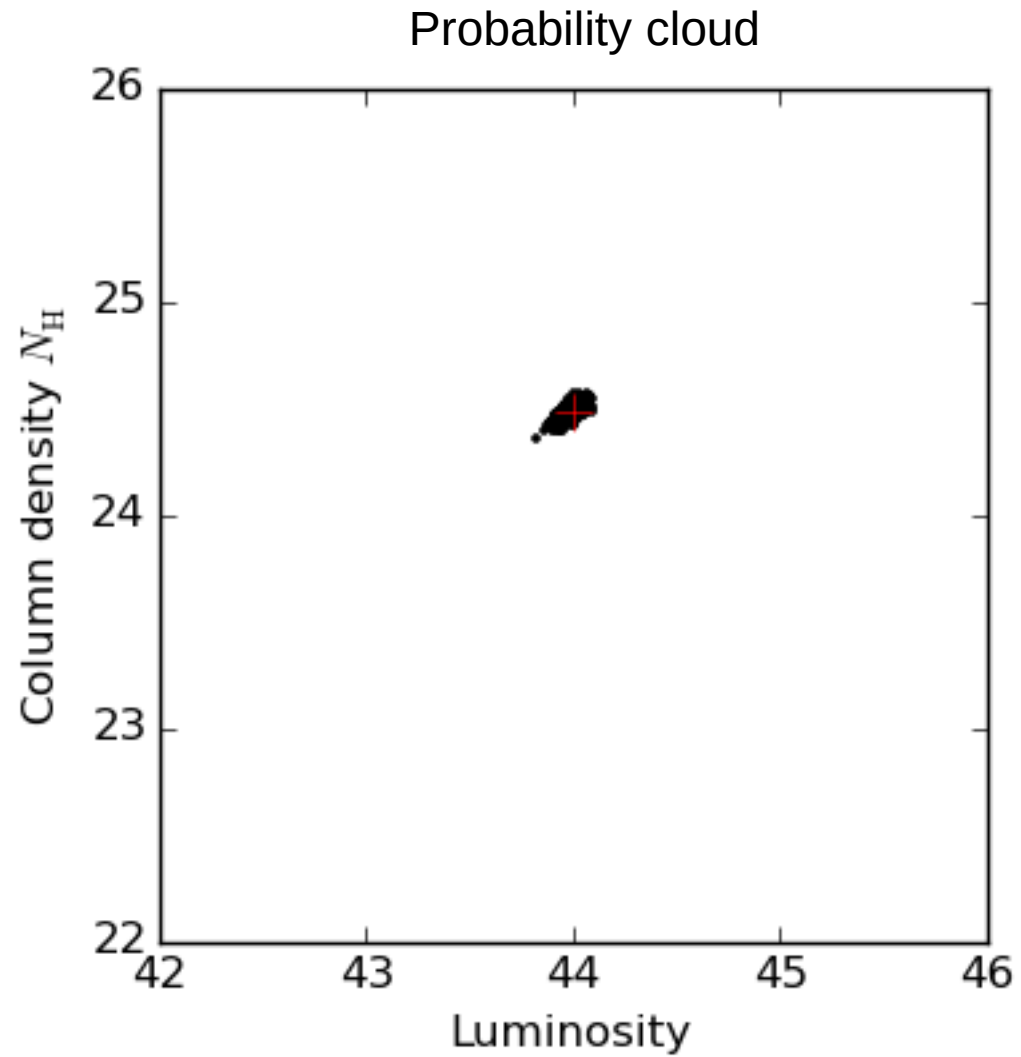
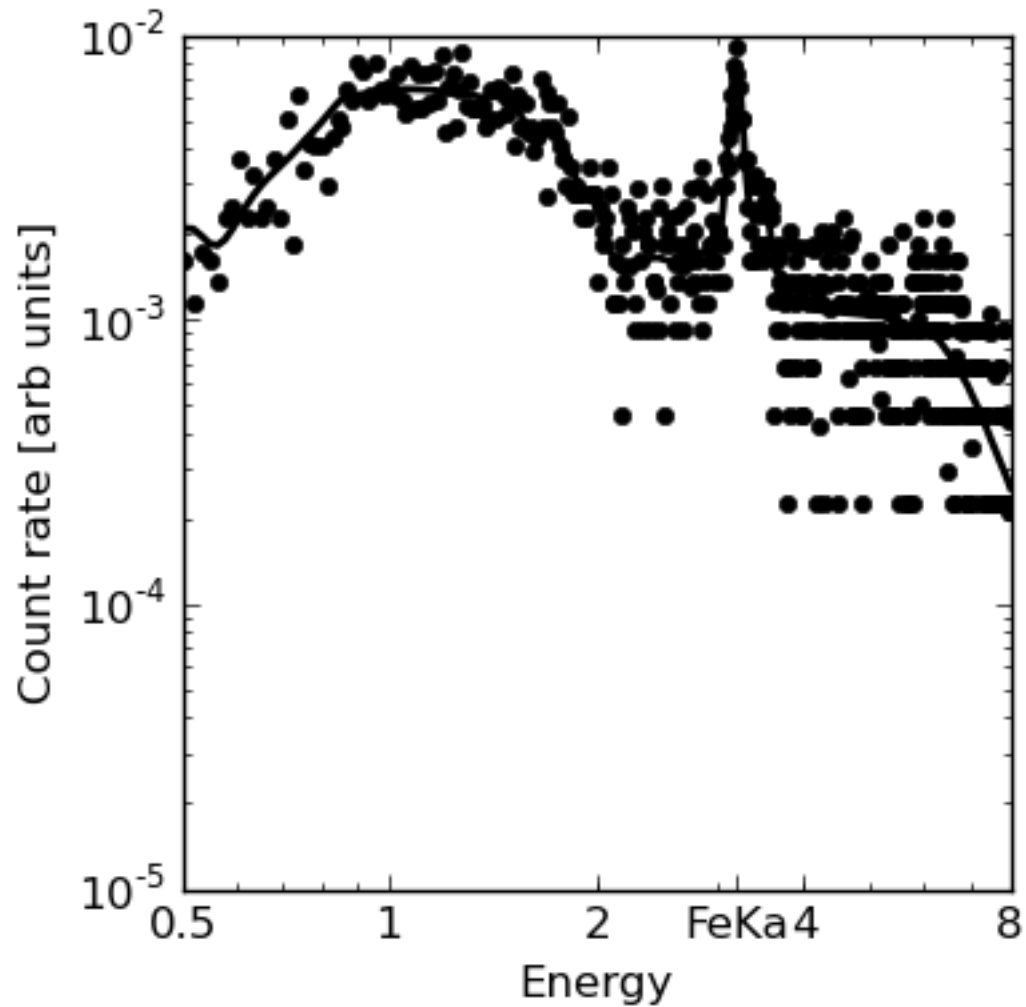
- Plugin connecting xspec/sherpa with MultiNest
<https://github.com/JohannesBuchner/BXA>
- Nested Sampling vs. MCMC for parameter estimation
- Model comparison with evidences/Bayes factors
- Population inference instead of samples statistics

L, N_H from X-ray spectrum

Scattered Powerlaw component

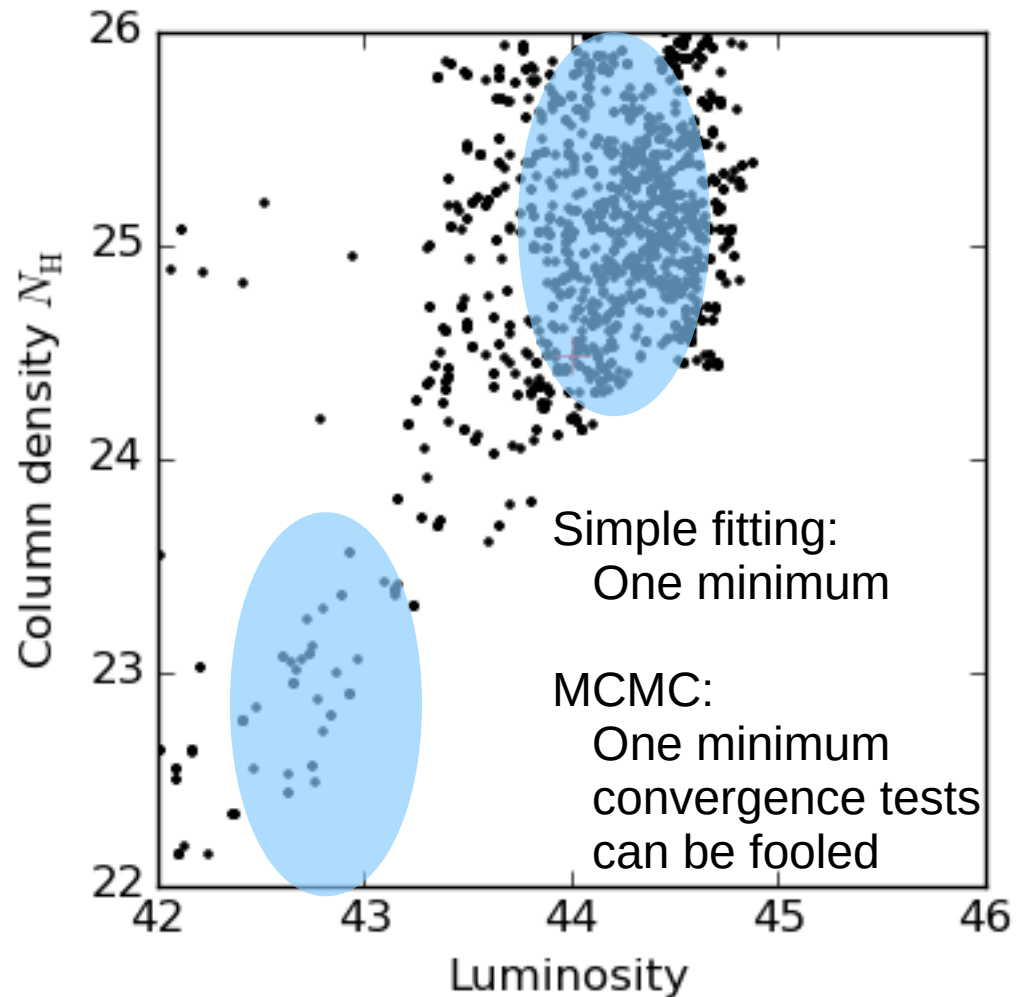


L, N_{H} from X-ray spectrum

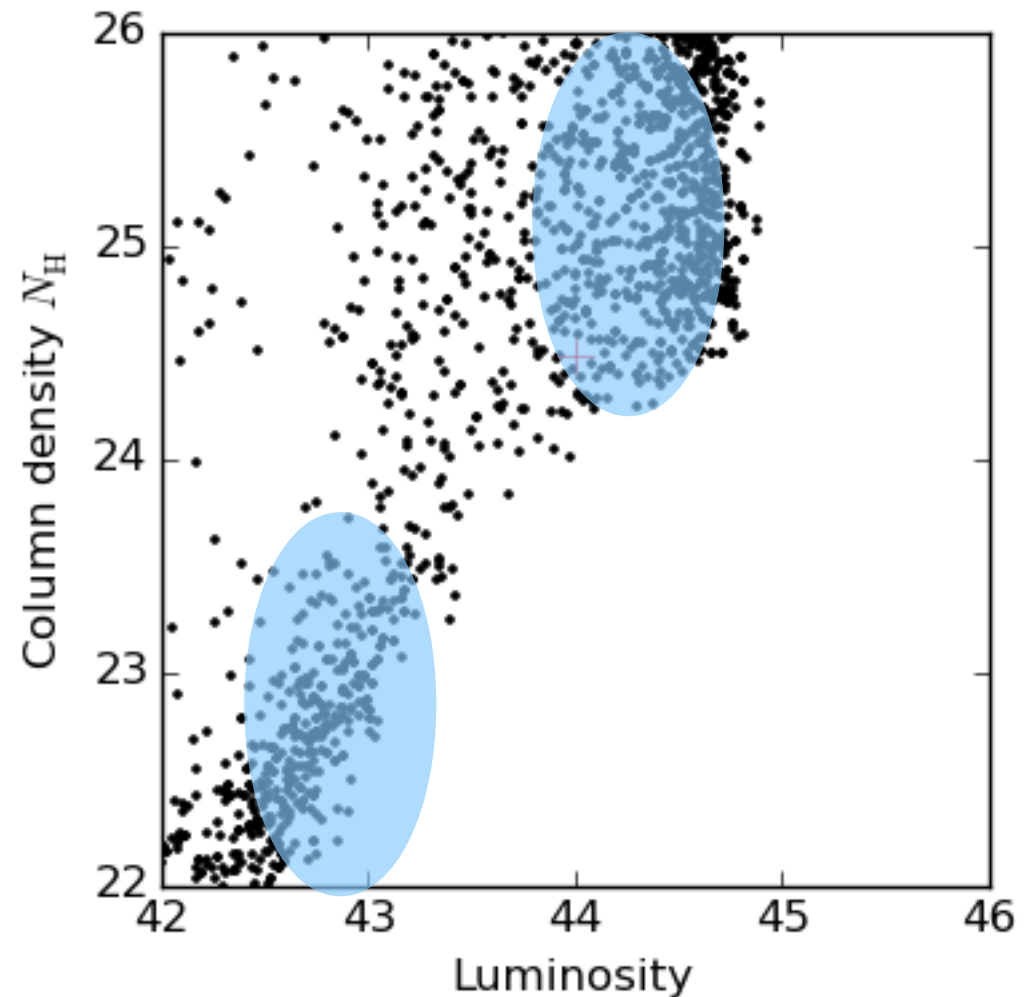


L, N_{H} from X-ray spectrum

MCMC (emcee/GW)

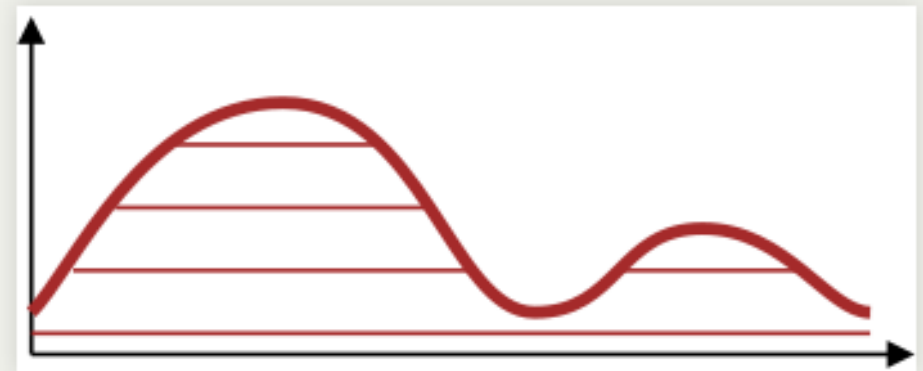
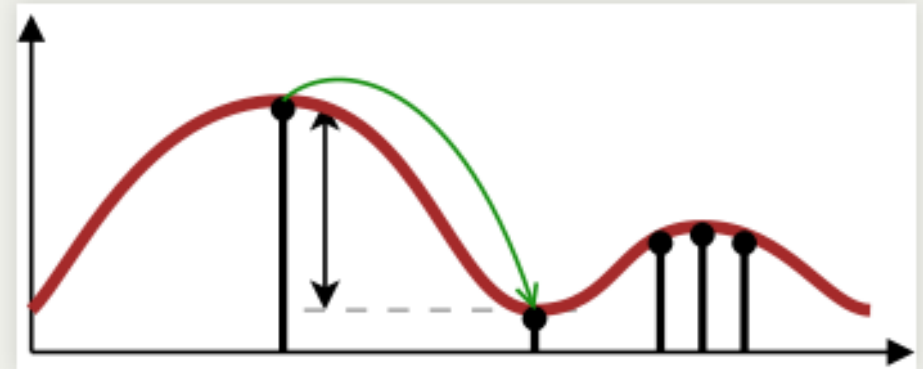


Nested Sampling (BXA/Multinest)



nested sampling idea

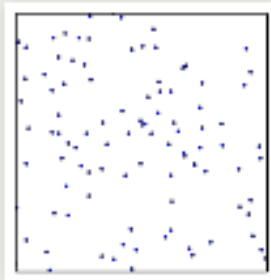
- MCMC: only consider likelihood ratios. Integration by vertical slices
- nested sampling: compute geometric size at various likelihood thresholds
- orthogonal, unique re-ordering of volume by likelihood



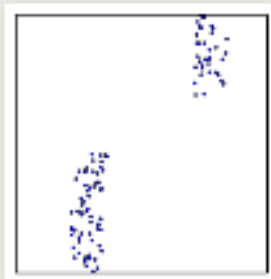
$$\sum \underbrace{\text{Shrinkage} \times \text{Likelihood}}_{\text{Importance of shell}} = Z$$



nested sampling algorithm



- Start with volume 1, draw randomly uniformly 200 points
- remove one, volume shrinks by 1/200.



- draw a new one excluding the removed volume
- Unique ordering of space required: via likelihood

**draw a new uniformly random point,
with higher likelihood**
(the crux of nested sampling)

- Scanning up vertically, done at some point
- converges (flat at highest likelihood)

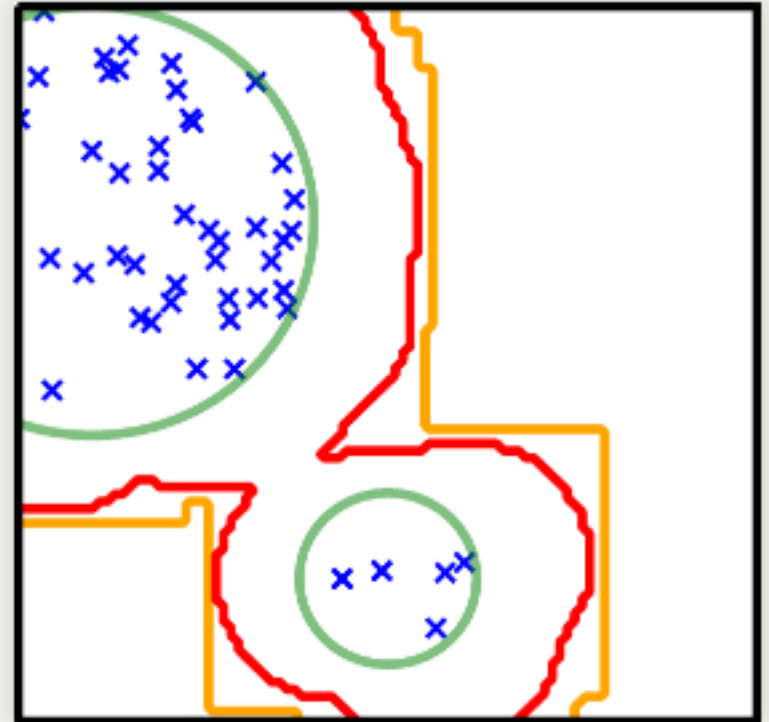


Missing ingredients

- MCMC: Insert tuned transition kernel
- NS: Insert constrained drawing algorithm
 - General solutions: MultiNest, MCMC, HMCMC, Galilean, RadFriends, PolyChord

RadFriends / MultiNest

- Use existing points to guess contour
- Expand contour a little bit
- Draw uniformly from contour
- Reject points below likelihood threshold
- RadFriends: Compute distance at which every point has a neighbor. Bootstrap (Leave out) for safety.
- MultiNest clusters and uses ellipses



Working with BXA

- Do not need to abandon xspec/sherpa
- Specify working parameter space
 - log, linear, etc. ~ priors – or ML MCMC approach
- Run once, be done!
 - No stepping, no convergence worries
 - No wondering about other solutions
- Output: PDF of each parameter

BXA in practice (with pyxspec)

I. Set up spectrum & model

```
s = Spectrum('example-file.fak')  
s.ignore("***"); s.notice("0.2-8.0")  
m = Model("pow")
```

II. Set up parameter space

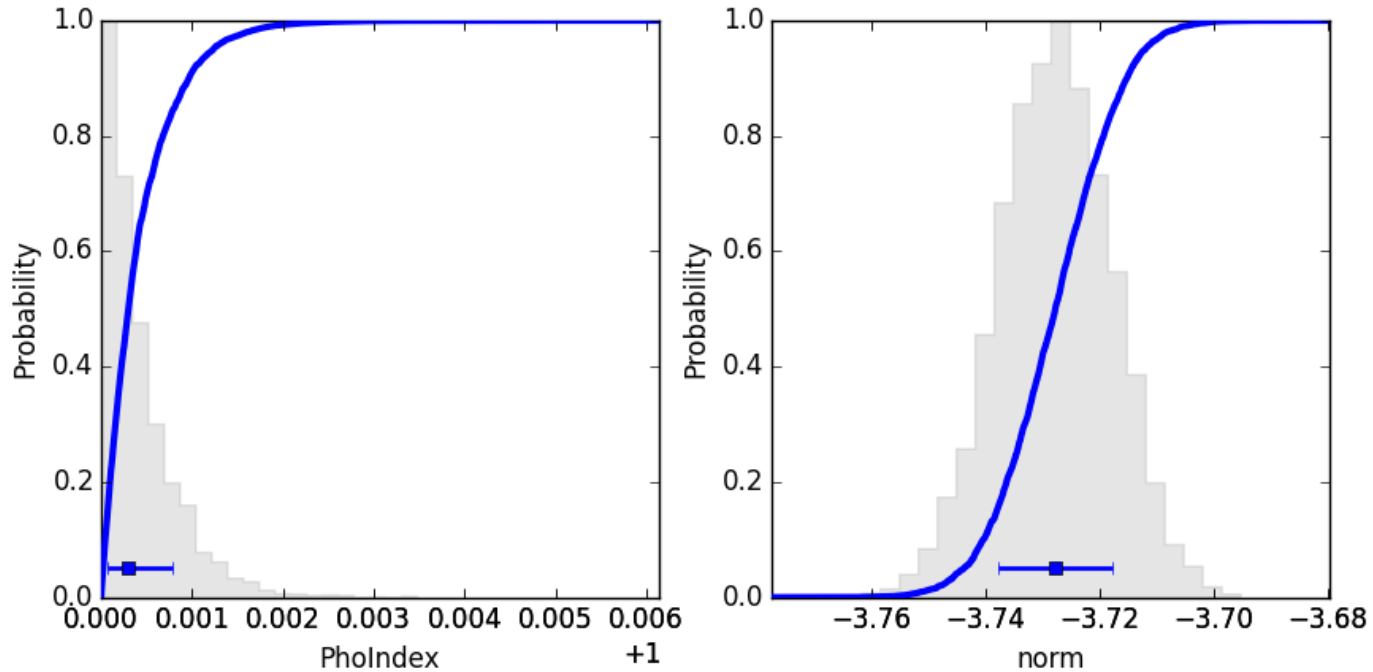
```
m.powerlaw.norm.values = ",,1e-10,1e-10,1e1,1e1" # 10^-10 .. 10  
m.powerlaw.PhoIndex.values = ",,1,1,3,3" # 1 .. 3
```

```
# define prior  
transformations = [  
    bxa.create_uniform_prior_for( m, m.powerlaw.PhoIndex),  
    bxa.create_jeffreys_prior_for(m, m.powerlaw.norm),  
]
```

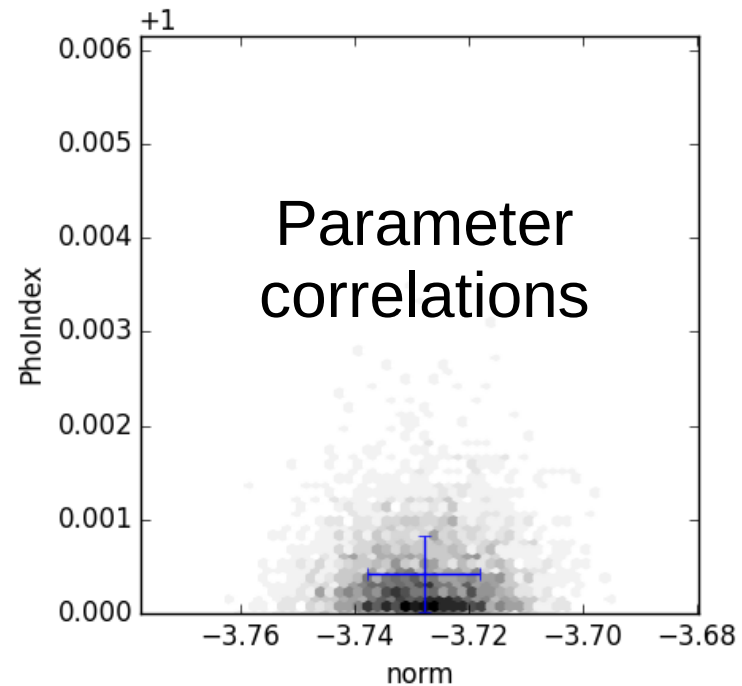
III. Run!

```
bxa.standard_analysis(transformations,  
    outputfiles_basename = 'simple-example')
```

Parameter probability distributions



That was parameter estimation
– now for model comparison!



Model comparison

see Buchner+14

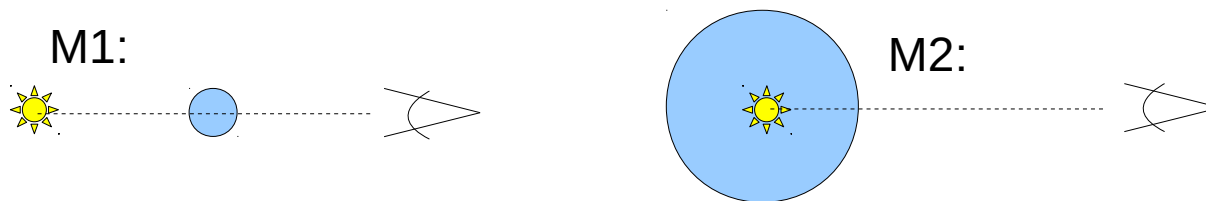
- $Z = p(D|M) =$
 - “marginal likelihood”, “evidence” ...
 - average likelihood over parameter space

- Z_1/Z_2 :

Probability for model M1 compared to M2

- Example:

- M1: wabs*pl – photo-el absorbed powerlaw
- M2: table{sphere}: includes Compton scattering



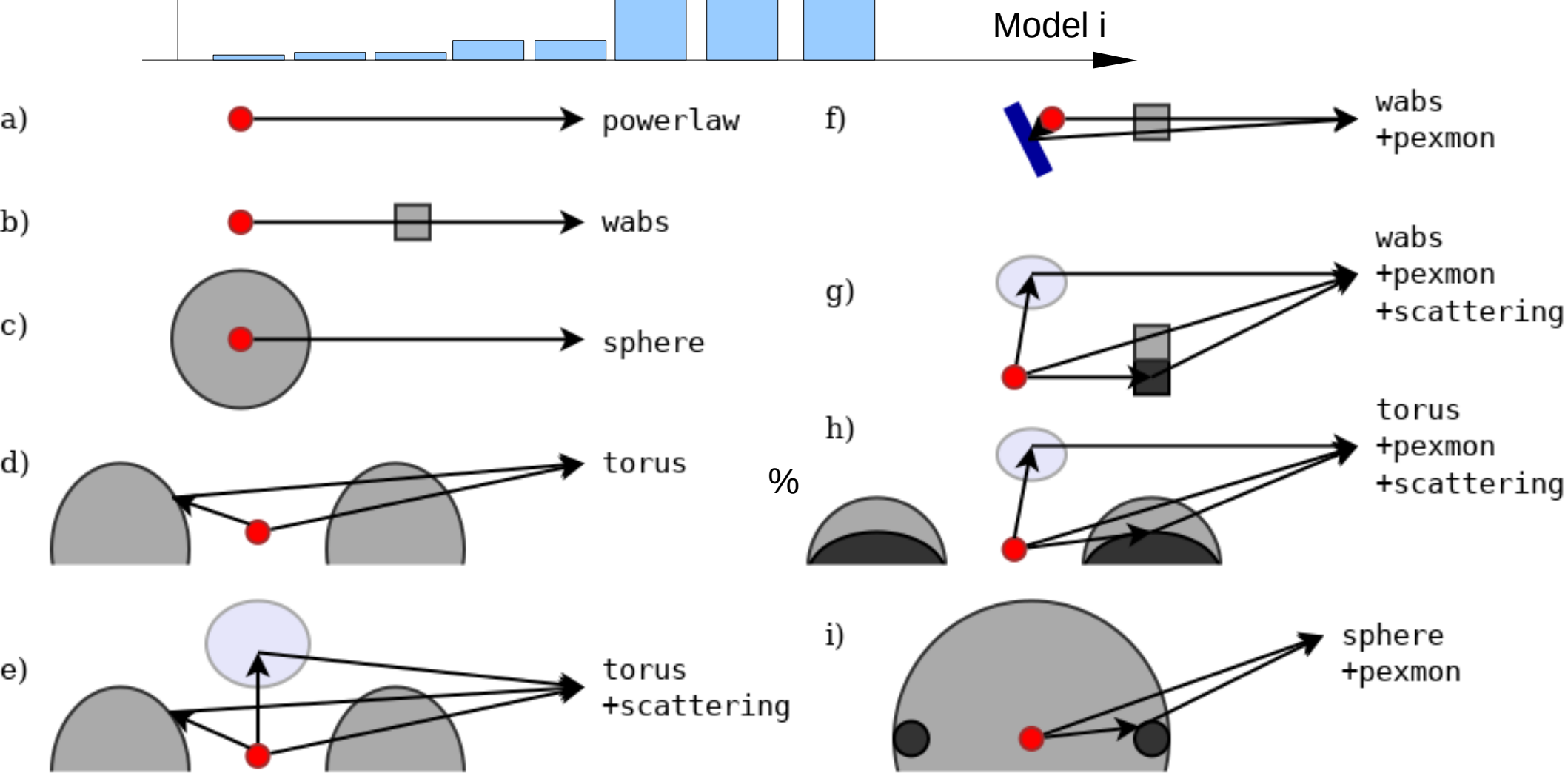
Z_1/Z_2 was 10

CDFS source 179

Comparison of several models

$$Z_i = p(M_i|D)$$

see Buchner+14



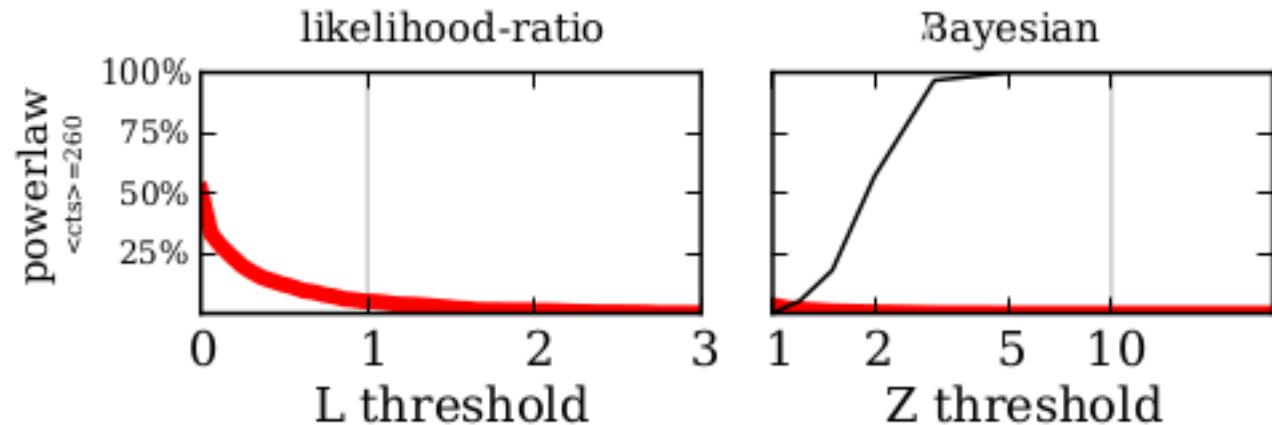
Example applications

- Comparison between obscurer geometries (Buchner+14)
- Identifying relativistic broad lines:
Narrow line model vs.
Relativistic broad line models (Baronchelli+in prep.)
- Threshold for decisions? False selection rates?

False selection rates

- As with any other estimator!
 - Simulate many times without effect
 - measure false decisions
 - Calibrate threshold

False selection rates (powerlaw vs wabs)



- Feasible with BXA

Population inference

- Example: Photon index, N_{H} distribution of population of LGRBs, AGN
- sample statistics
- (selection effects – see Buchner+15, A1)
- Keep uncertainties, Upper limits
- Don't:
 - Plot histogram of means
 - Stack posteriors

Hierarchical Bayesian inference

- Assume flexible population model

Examples: Gaussian, Gaussian mixture, beta distribution

$$M(x|\theta_M) = N(x|\mu, \sigma)$$

- Fold in posterior samples $x_{i,j}$ for object i

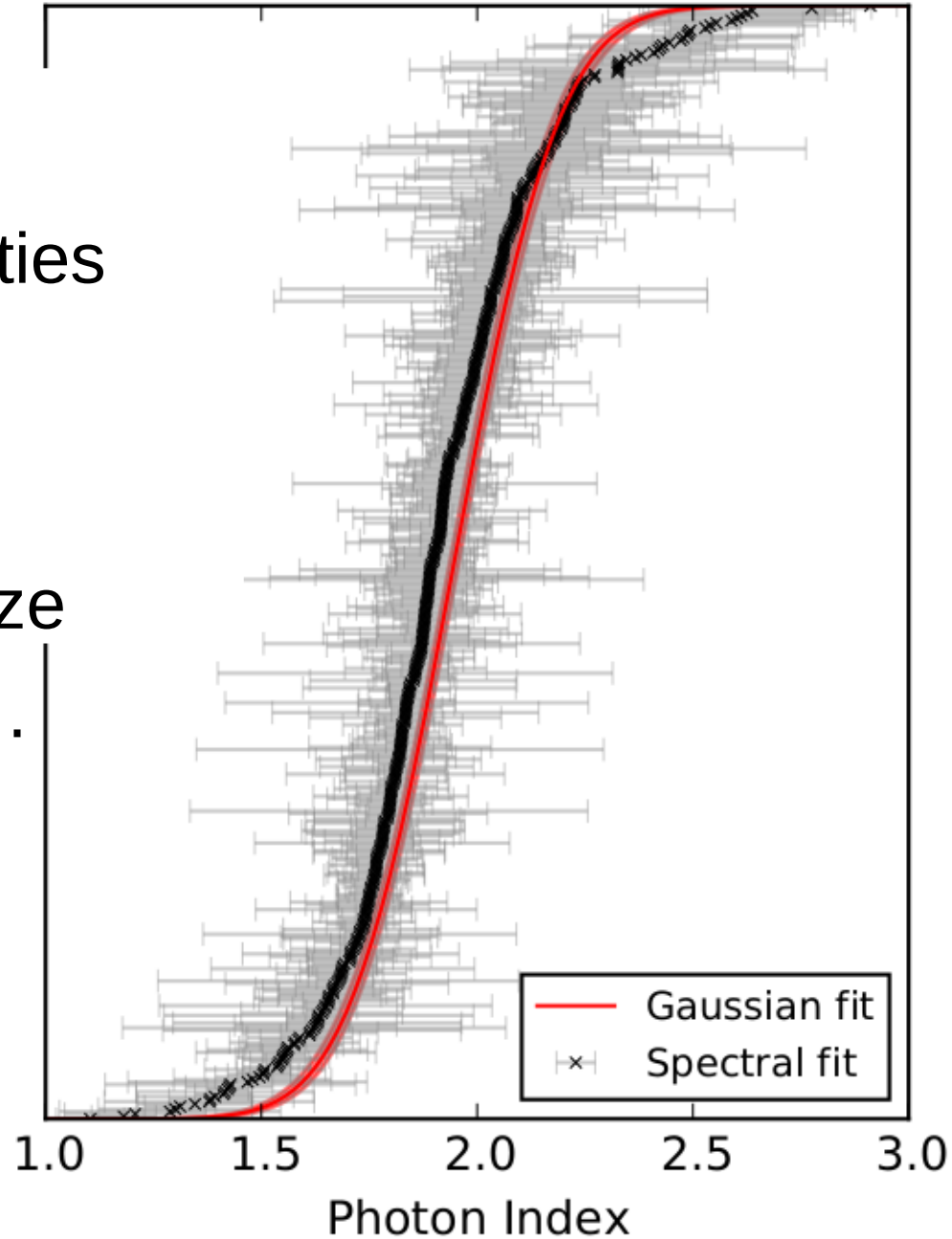
$$L(\theta_M) = \prod_i \sum_{j=1}^N M(x_{i,j}|\theta_M)$$

Product over objects
Mean over samples

- Likelihood for population parameters - MCMC
- Importance sampling from posteriors
 - need flat priors, see Buchner+17a for details

Population inference

- Keeps uncertainties
 - Fold in redshift uncertainties
 - (unlike X-ray stacking!)
- Population uncertainties
 - Include limited sample size
 - Model can be 1d,2d,3d,...
- Special case of luminosity function analysis (Loredo2004)



Summary

- BXA: a MultiNest plugin for xspec/sherpa
- BXA: parameter estimation
 - Robust to multiple solutions, convergence issues, etc.
- BXA: Bayesian model comparison
 - Sensitive; can be interpreted; calibrated to FDR
- Population inference
 - Understand the population from a limited sample
 - Forward-fold all uncertainties

more in: Buchner+14, +15, +17a + their refs