Seeing the Low-Count Sky

EGRET Examples with an Eye to GLAST

A. Connors, D. van Dyk, J. Chiang,
for:
CHASC

SAMSI06 Working Group: Source and Feature Detection

Bottom line: "I think we've cracked it."
(i.e. no more χ^2 for low-count Poisson; plus non-parametric comparisons of observations; plus....)

OUTLINE:

1/ Who: Cumulative work of many many people

* *Your help (priorities, acronym, ... ?)*

2/ What/Why:

* *Example: CGRO/EGRET All-sky Diffuse*
* *(3C279 VP 0030, 0110)*

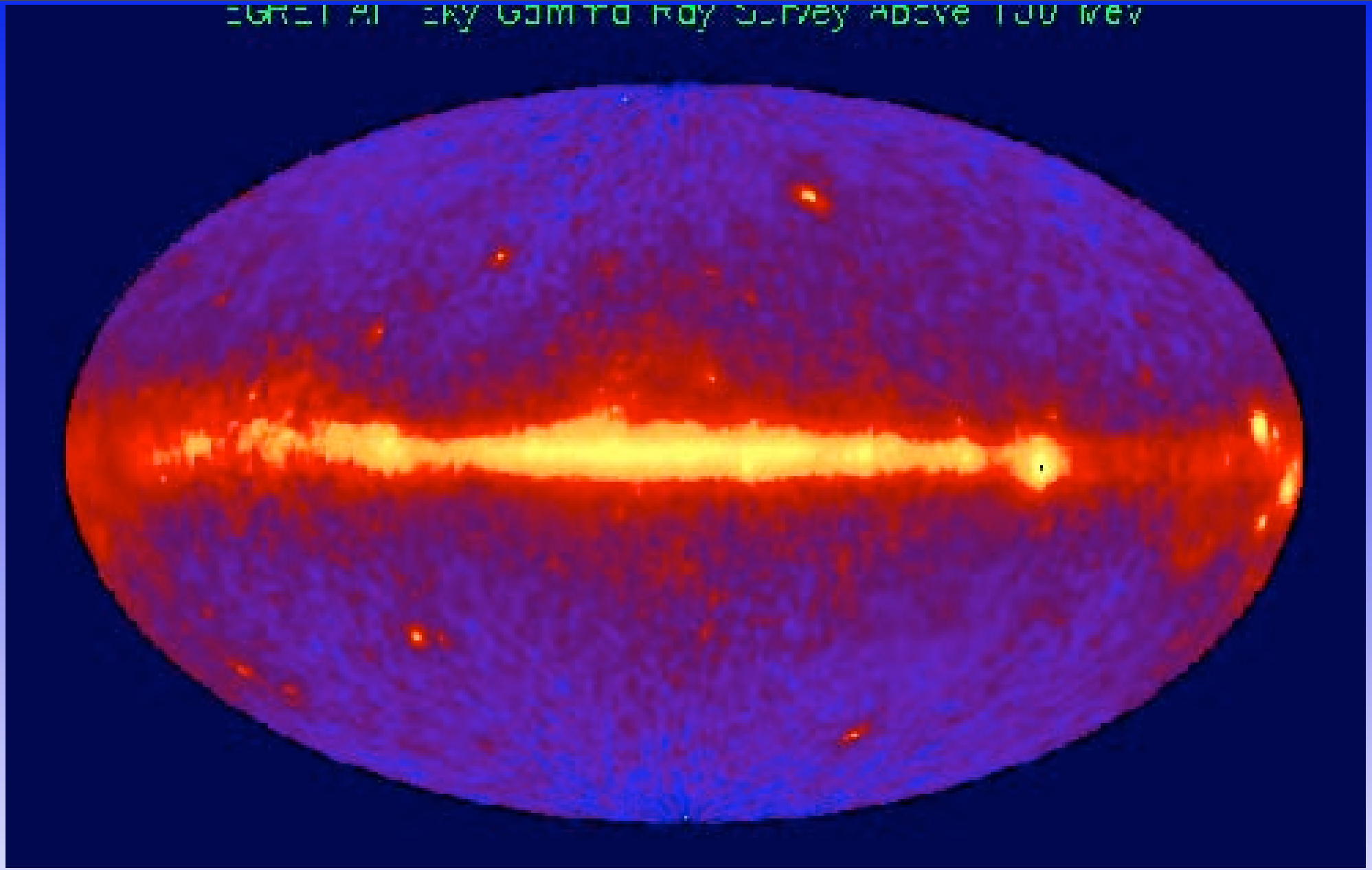
3/ How (You almost know already)

* *The Rules, and Breaking Them*
* *Replace χ^2 With "Capture A Multi-Scale Difference"*
* *Use Full Likelihood Analysis, + Simulate H0*

4/ Examples?

* *CGRO/EGRET Diffuse (simulations) Models vs Data*
* *Examples: CGRO/EGRET Diffuse Models vs Data*

5/ It Works ---> Future?

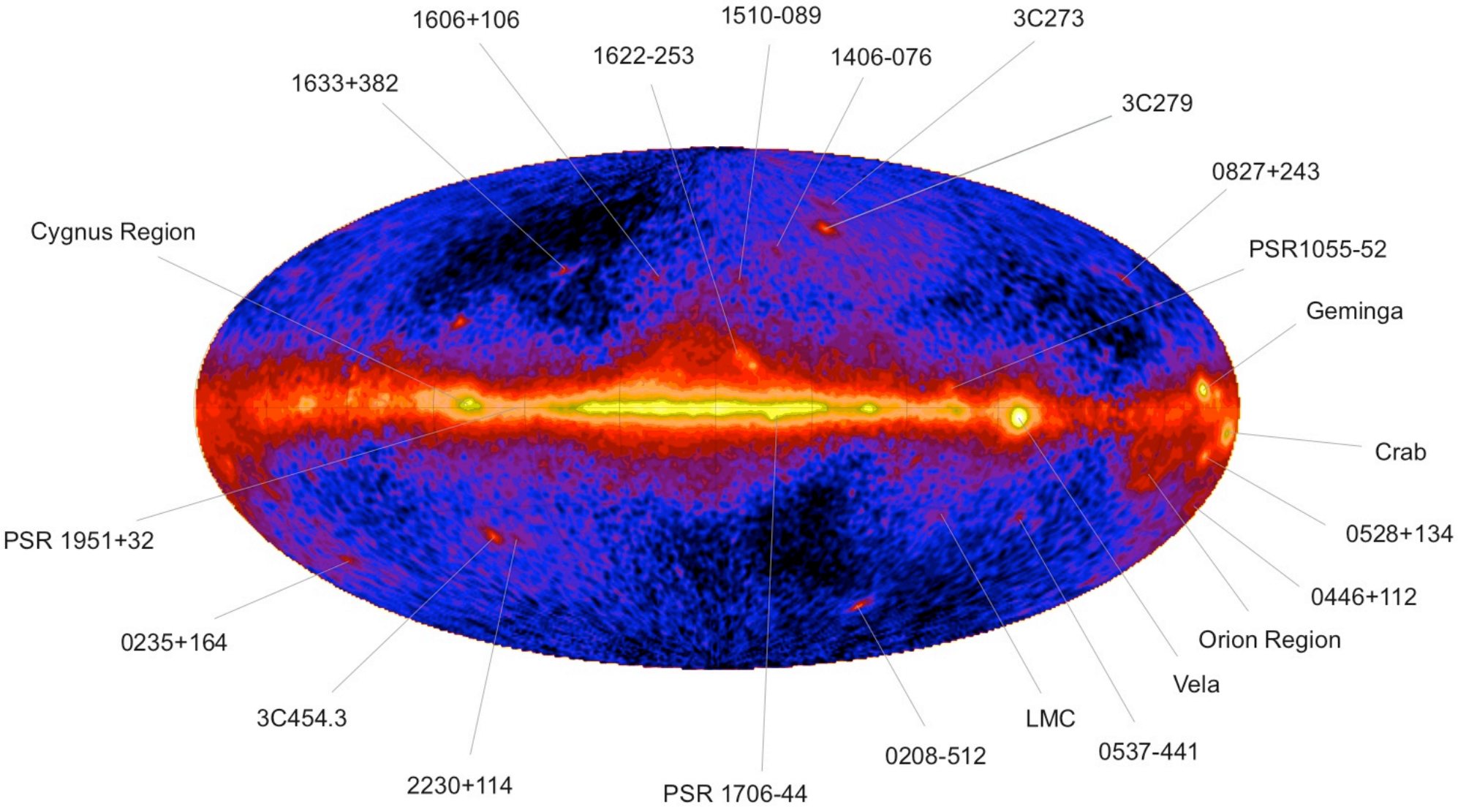


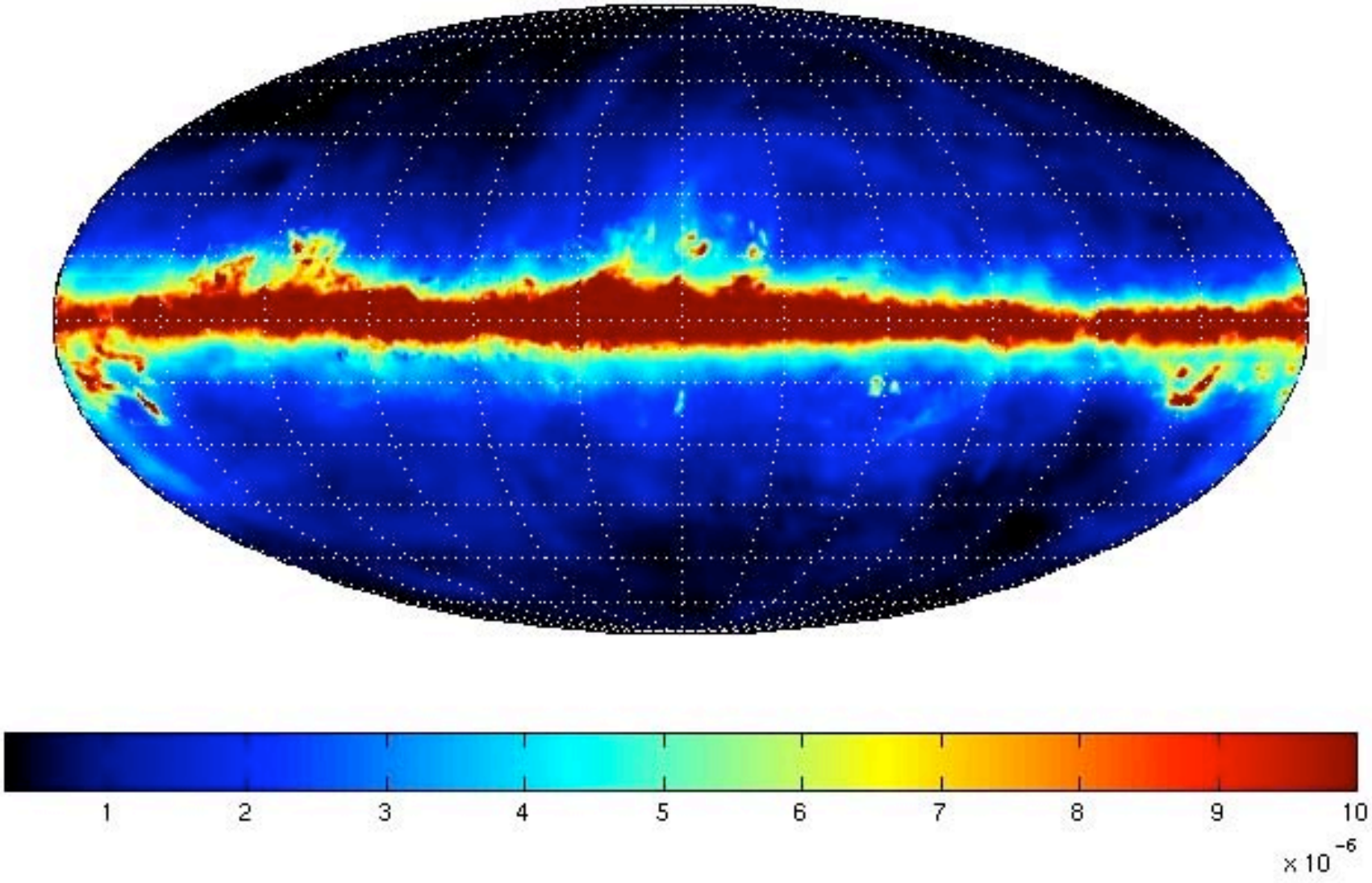
All-Sky Map > 0.1 GeV Photons
9 years of CGRO/EGRET

EGRET

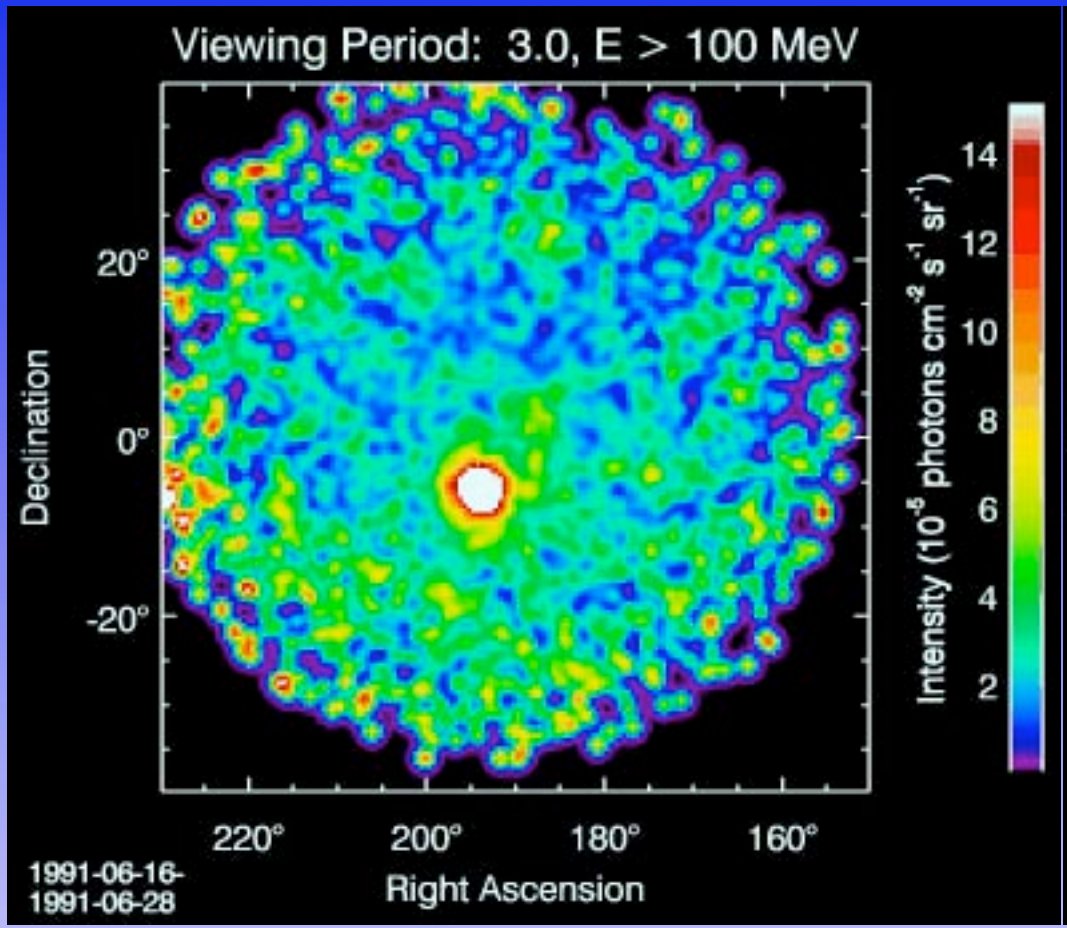
Skymap E > 100 MeV

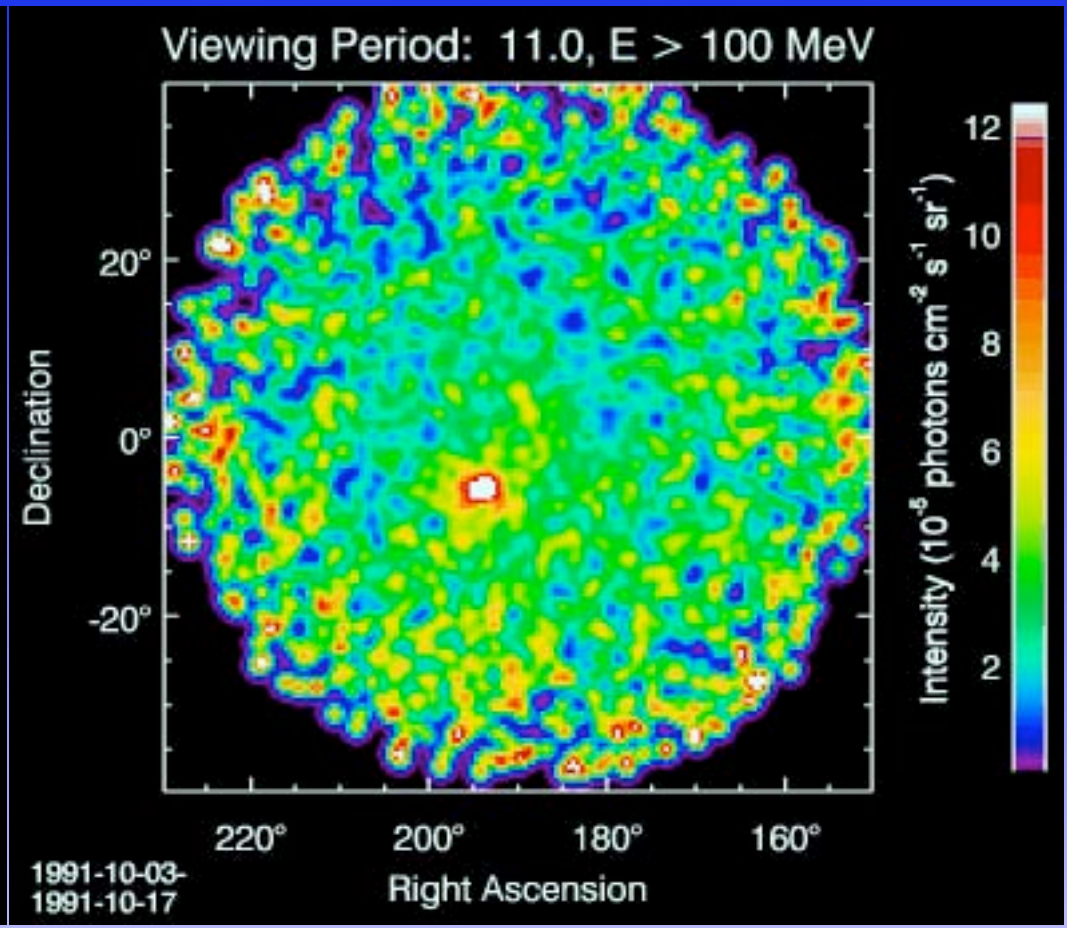
Phase 1 - 4



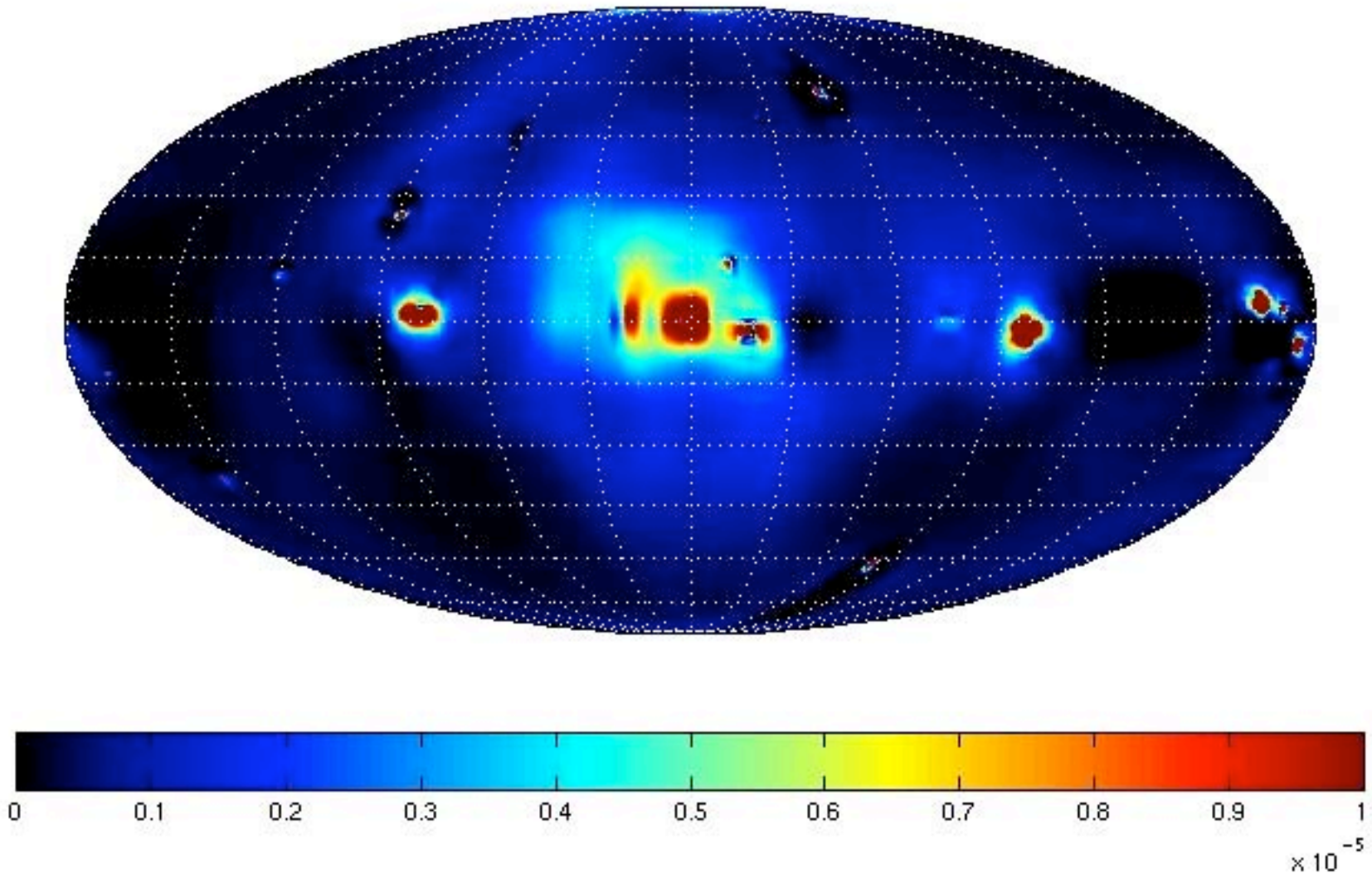


Model of Diffuse Emission along Galactic Plane:
Strong, Moskalenko, Reimer (GALPROP)





Can we do "Model-free" (actually, non-parametric) comparison of, say, CGRO VP 0030 and 0110 ?



What is that excess glow around the Milky Way?
(Dixon, Hartman, Kolaczyk, et al:
Poisson-tailored Haar Wavelet Thresholding)

How We Began (SAMSI06 SaFDe):

Dixon, Hartman, Kolaczyk, et al 1998:
New Astronomy 3 (1998) 539.

`The immediate question arises as to the statistical significance of this feature. Though we are able to make rigorous statements about the coefficient-wise and level-wise FDR, similar quantification of object-wise significance (e.g., "this blob is significant at the n sigma level") are difficult.'

NOW (POST SCMA IV):

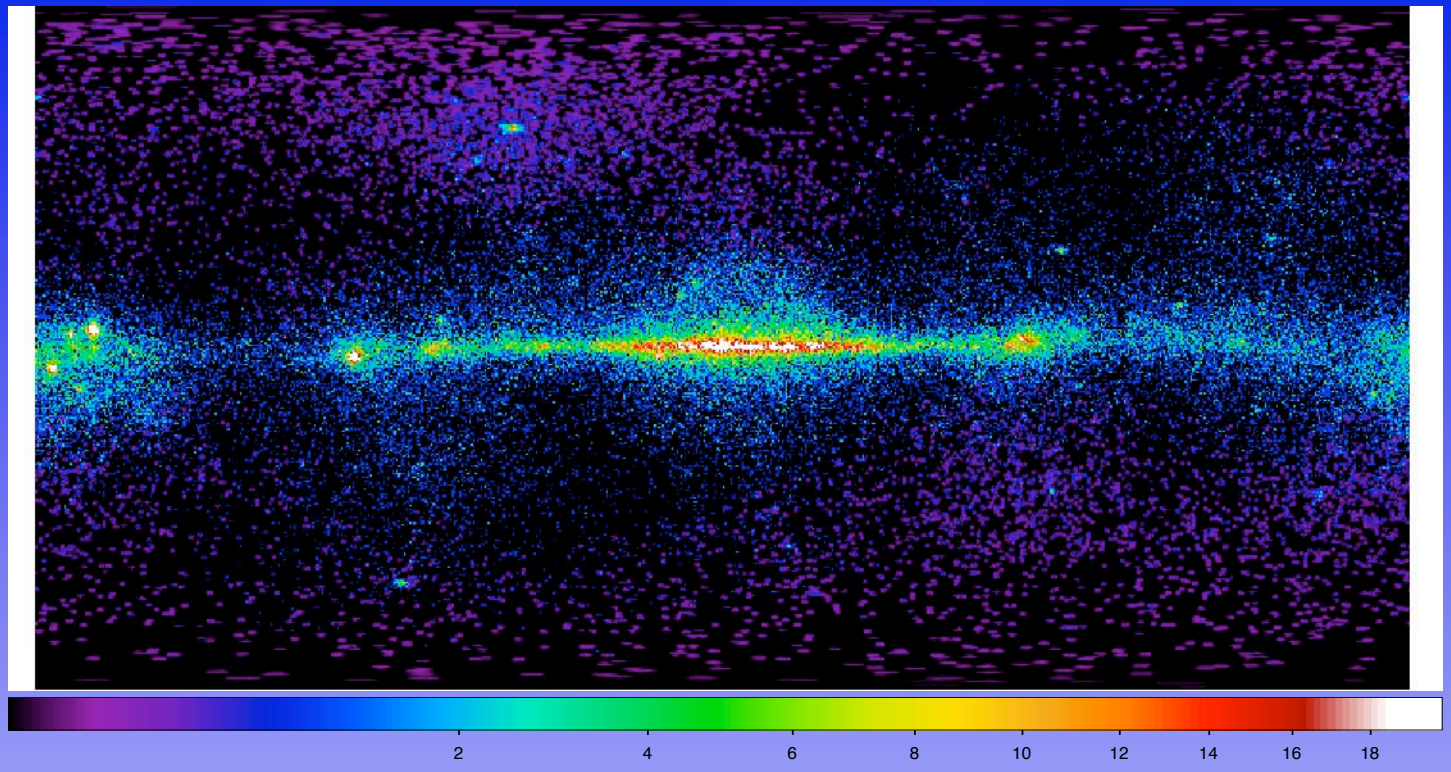
TRICK 1: Multi-scale Explicitly for Poisson "MMI" (Kolaczyk, Nowak; see Willett also).

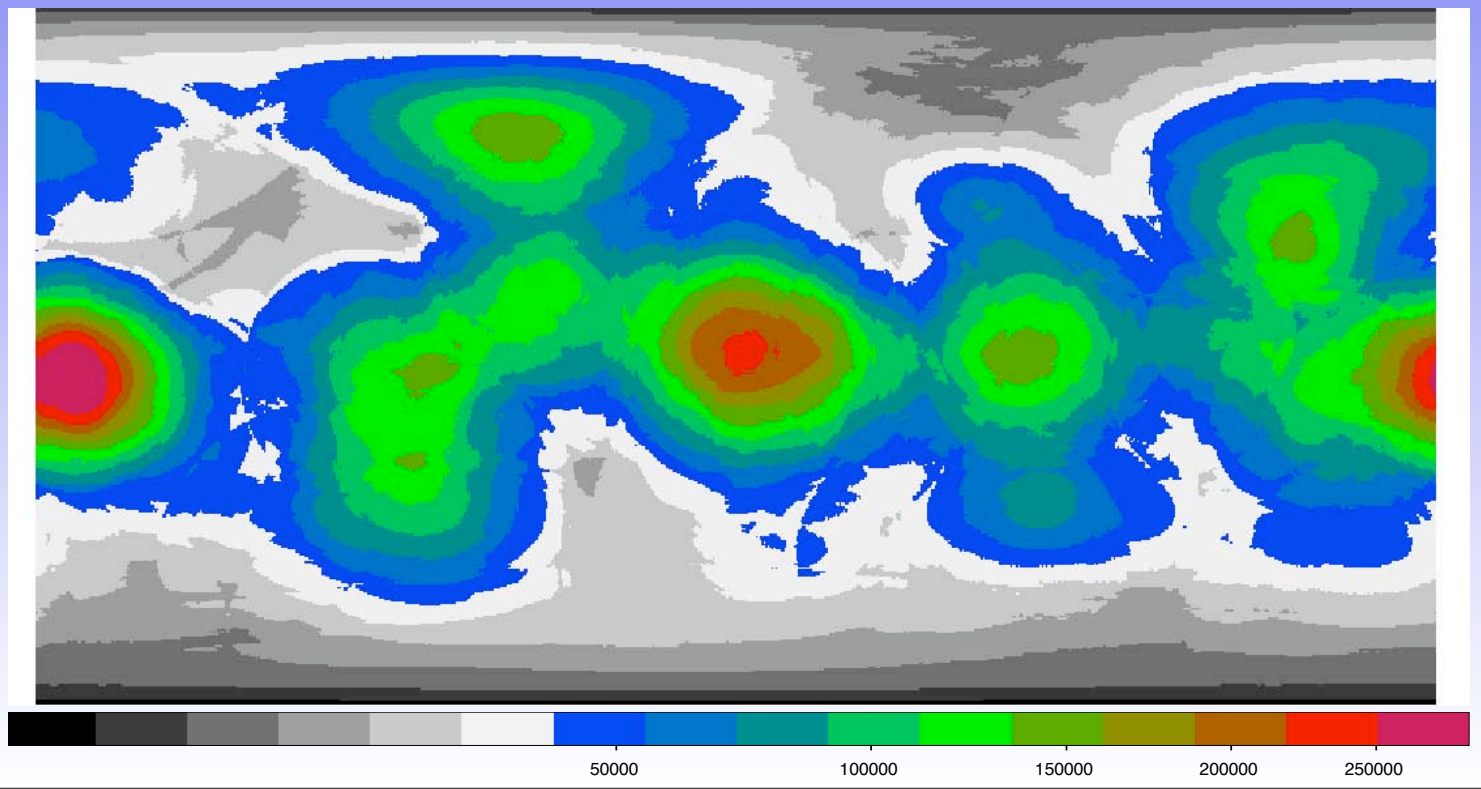
TRICK 2: Embedded in Full Poisson Likelihood - needs MCMC to "fit", as in EMC2:
Esch et al, van Dyk et al.,

TRICK 3: Compare Results on Data vs NULL

Q: What do we GET by being so careful about the statistics? What does mathematical elegance GET us?

A: Known Convergence; Error-handling; and *Speed*.


2
4
6
8
10
12
14
16
18


50000
100000
150000
200000
250000

METHOD:

I. MODEL the "mis-match" with flexible,complete, basis such as Multi-scale, MRF, etc:
"Capture A Multi-scale Residual"

CAMR

II. EMBED in Full Likelihood Analysis (Esch et al., van Dyk et al., MCMC)

FLA

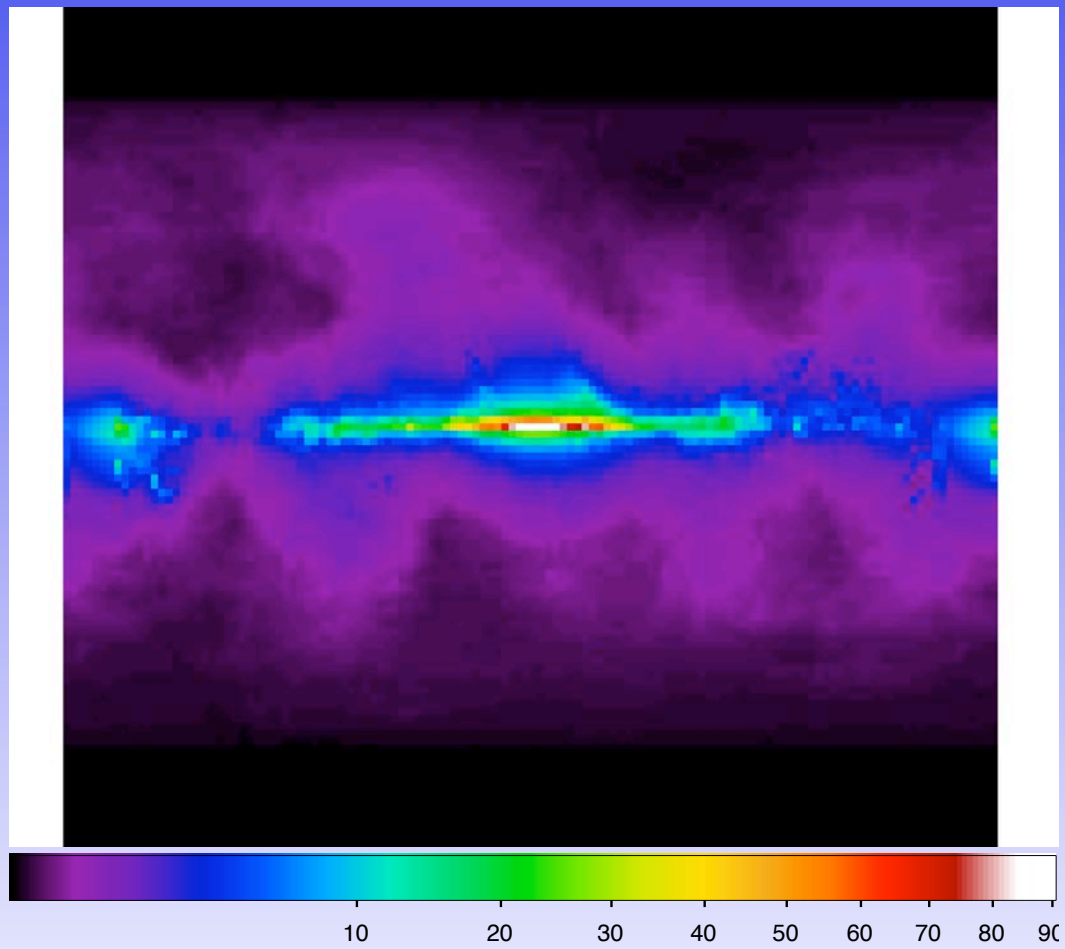
+

III. COMPARE results on data to results on simulations of null hypothesis

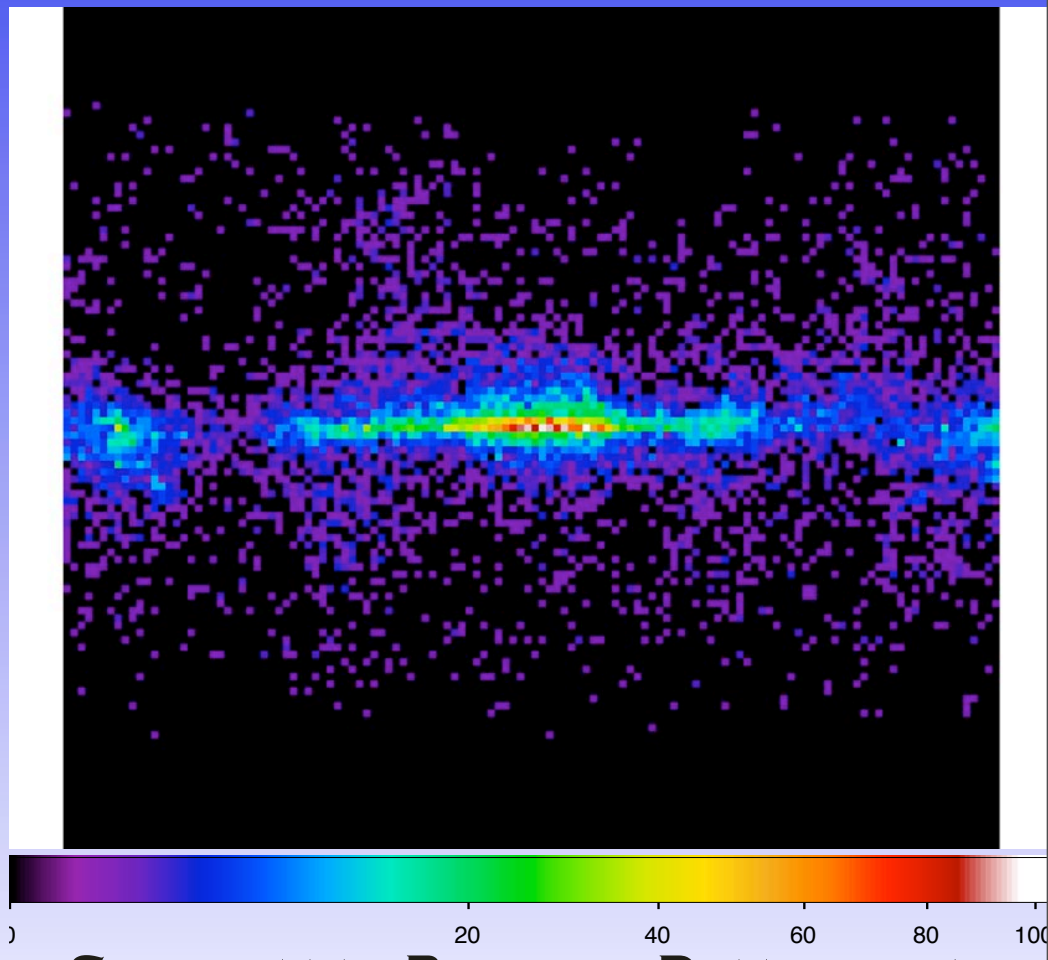
SH_0

IV. Toss $\chi 2$ .....

What does "nothing" (Data=null) look like, with our method?

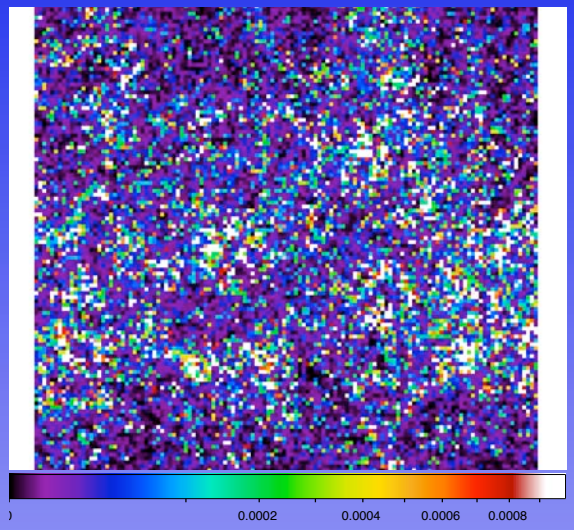


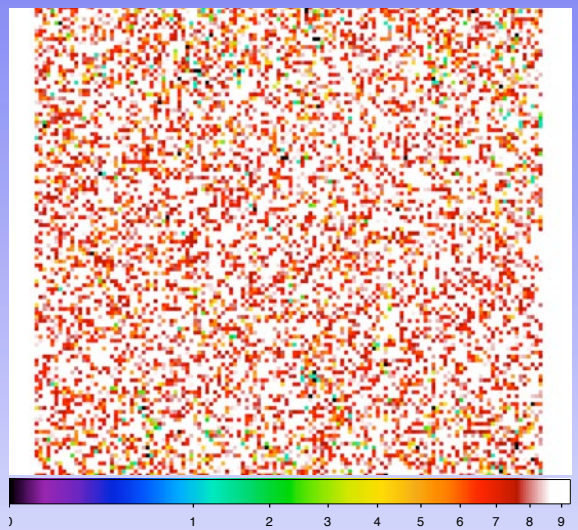
Background (i.e. physics) model --- GALPROP IC, Bremss, Pion and CGRO/EGRET >1GeV Exposure



Simulated Poisson Data based on GALPROP+CGRO/EGRET model I.E. **NO** Model/Data Mis-Match

What does "nothing" (Data=Null)look like?

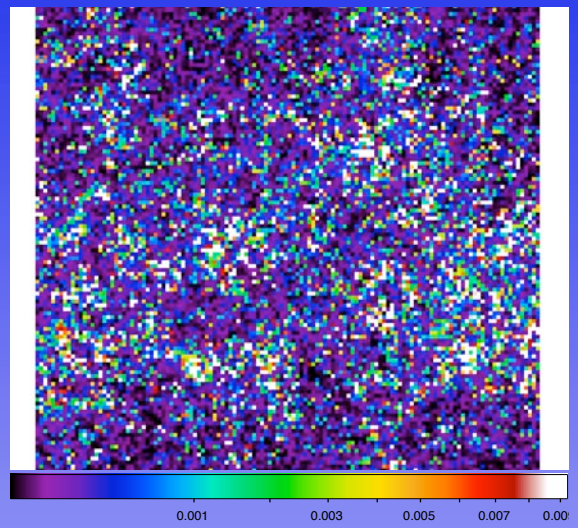


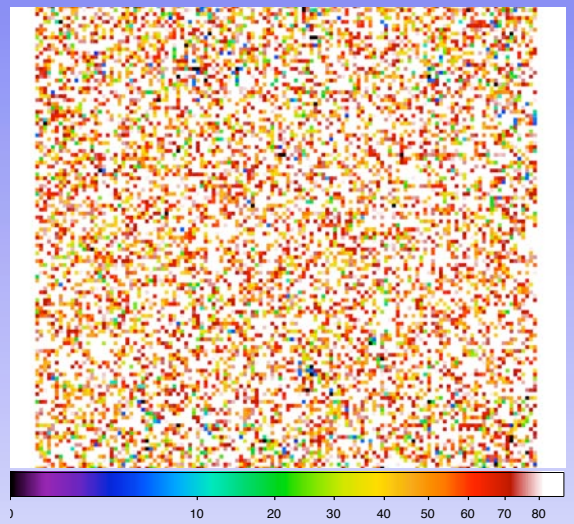


Mean, SKEW

(of MCMC fit to EMC2)

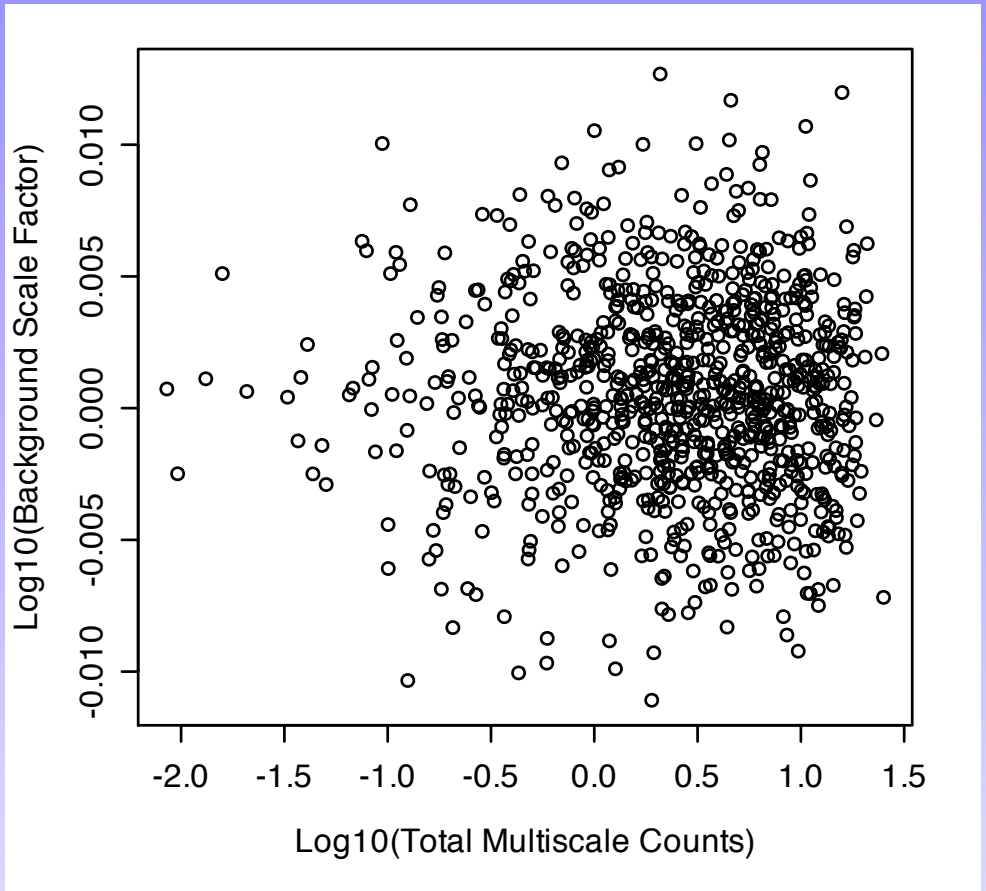
After ~10^3 draws

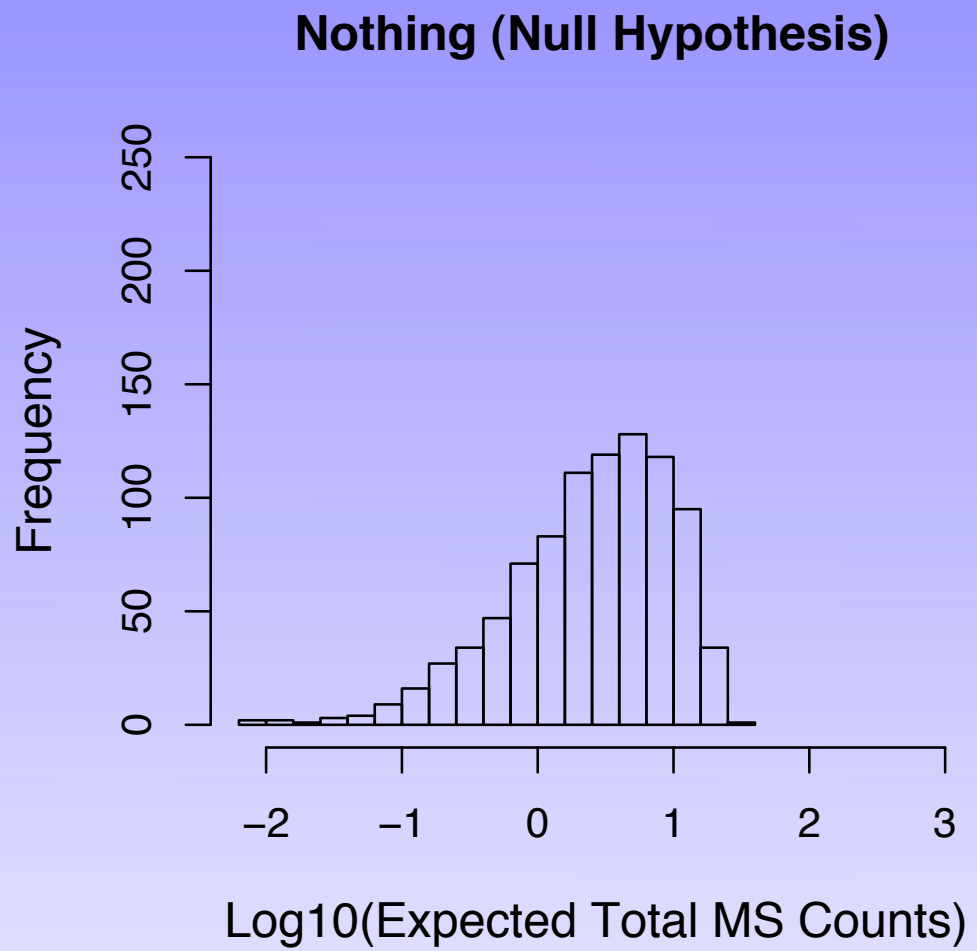




Sigma, Kurtosis

What does "nothing" (Data=Null) look like?

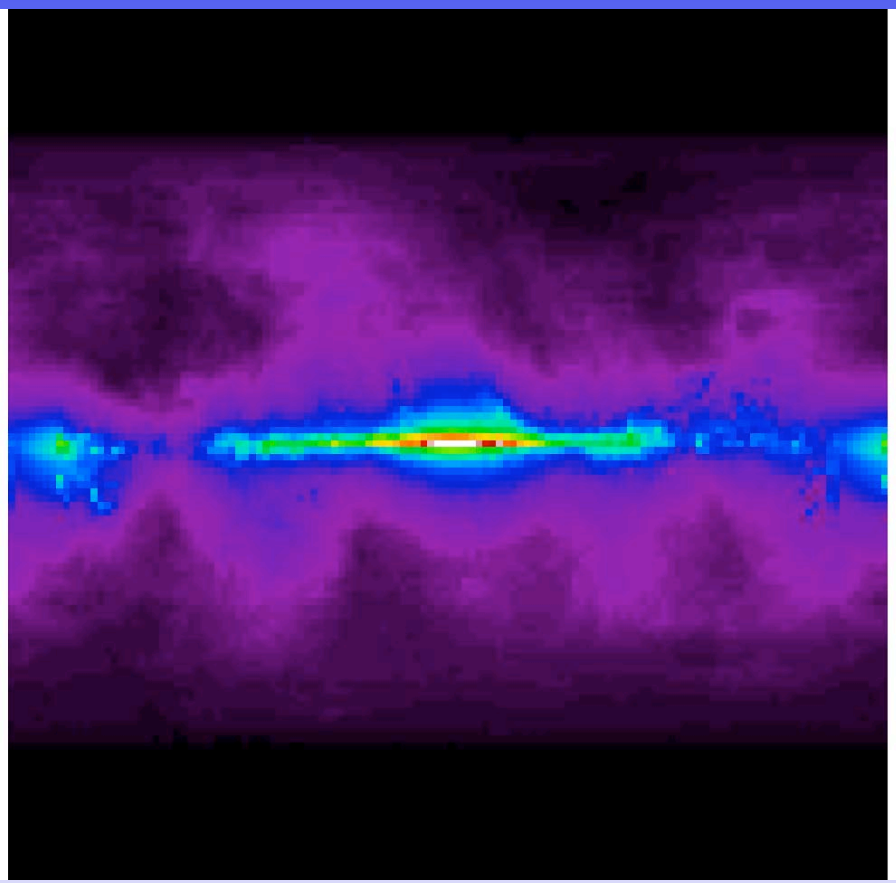


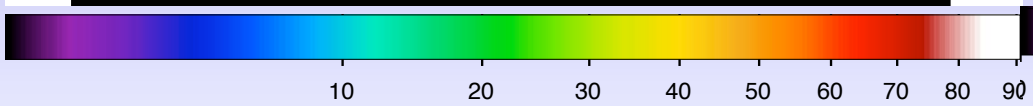


Scatter Plot

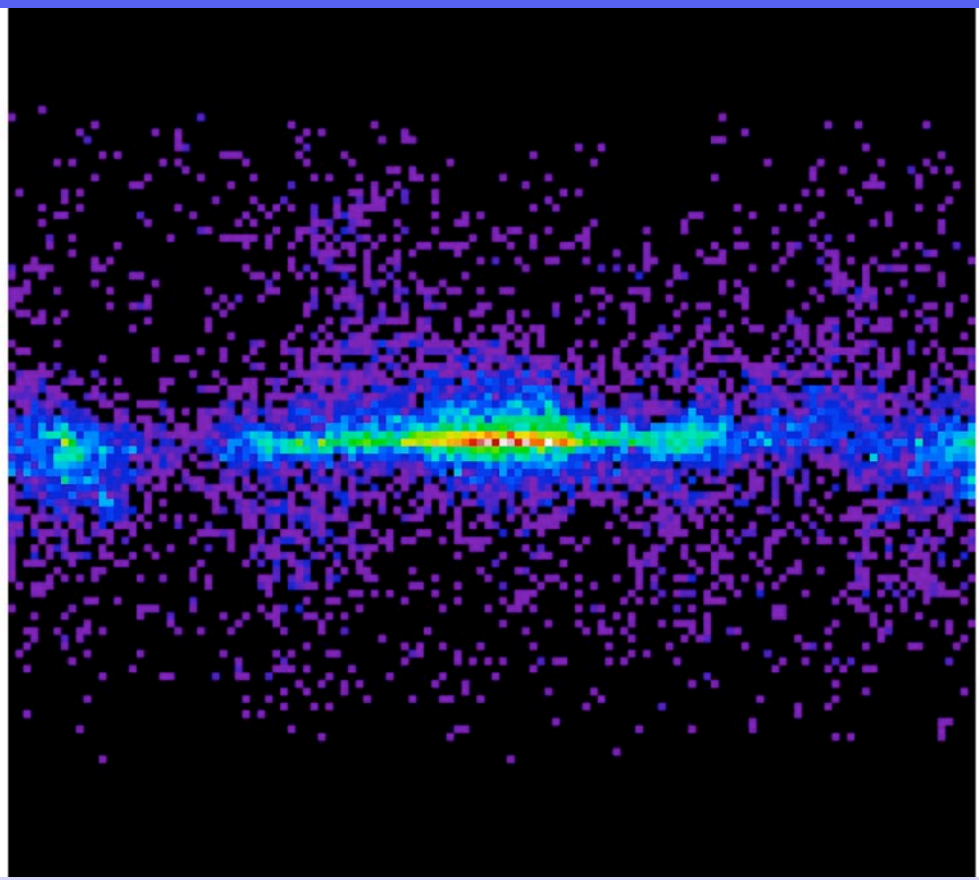
This is what we use to estimate significance.

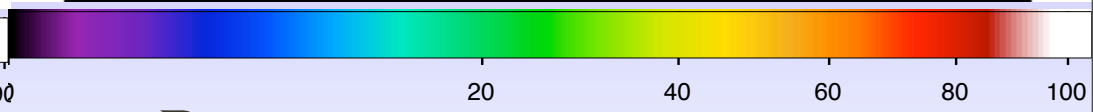
What does Small Model Shape Mistake look like, with our method?





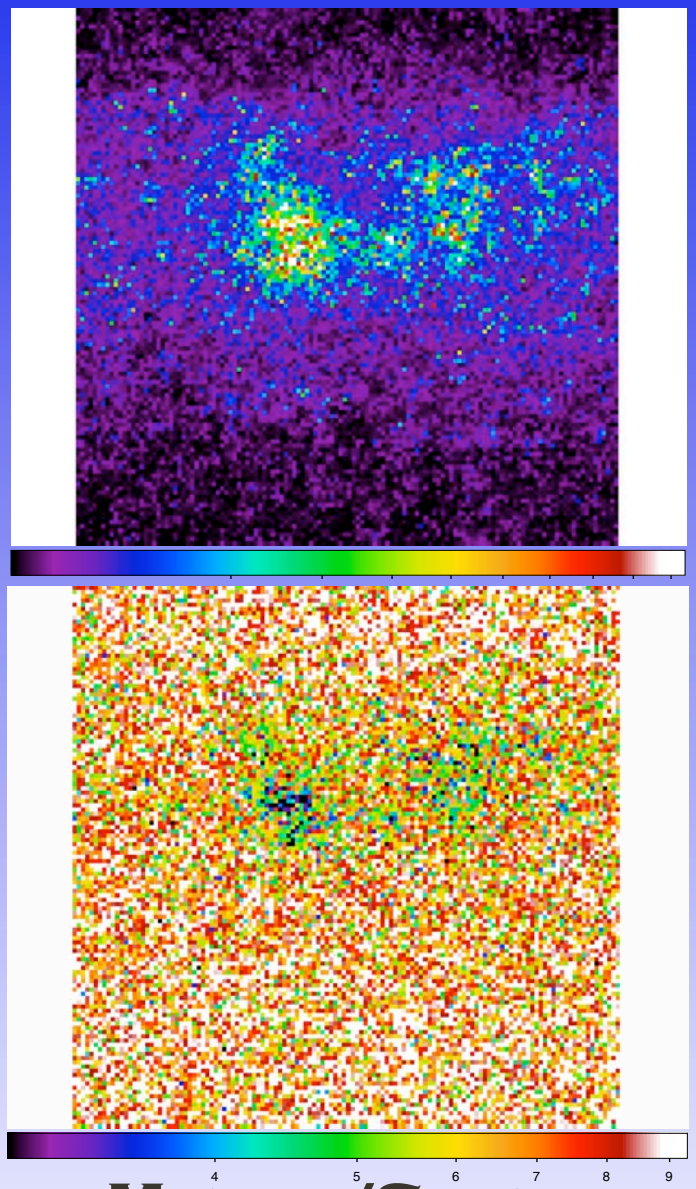
Fit Model (i.e. background) now has Inverse Compton suppressed above plane



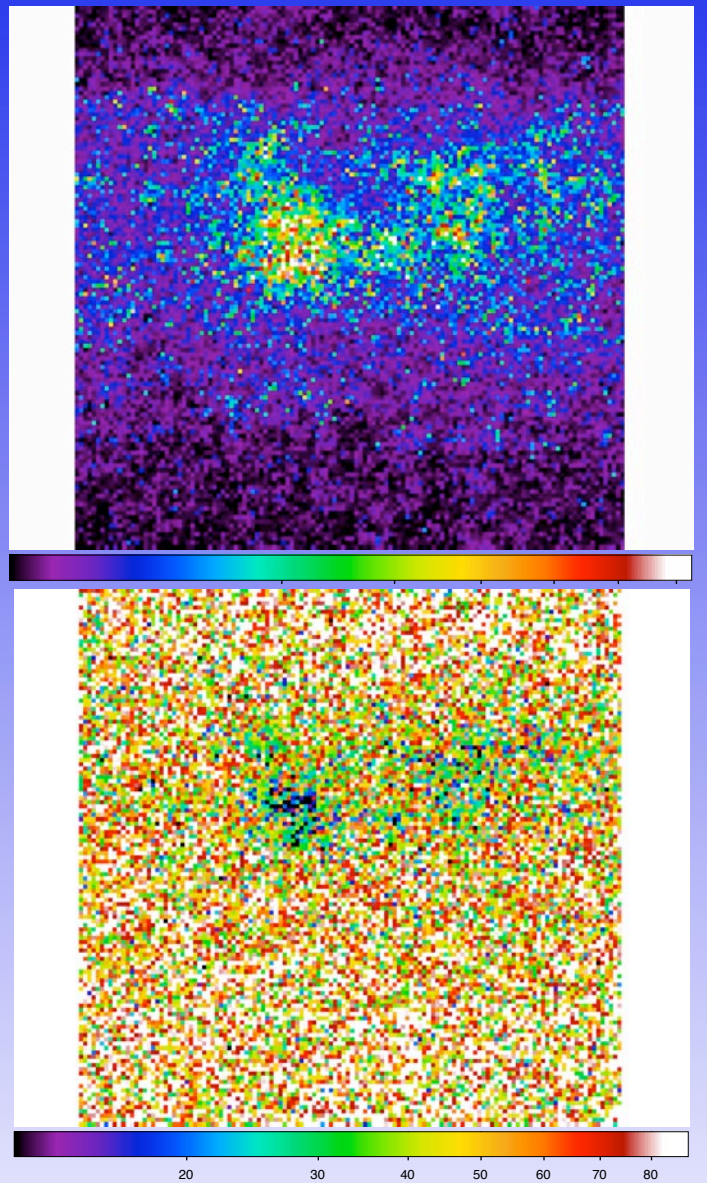


Data are the same as in correct null model (from 1st example)

What does small model shape mistake look like?



Mean /Skew

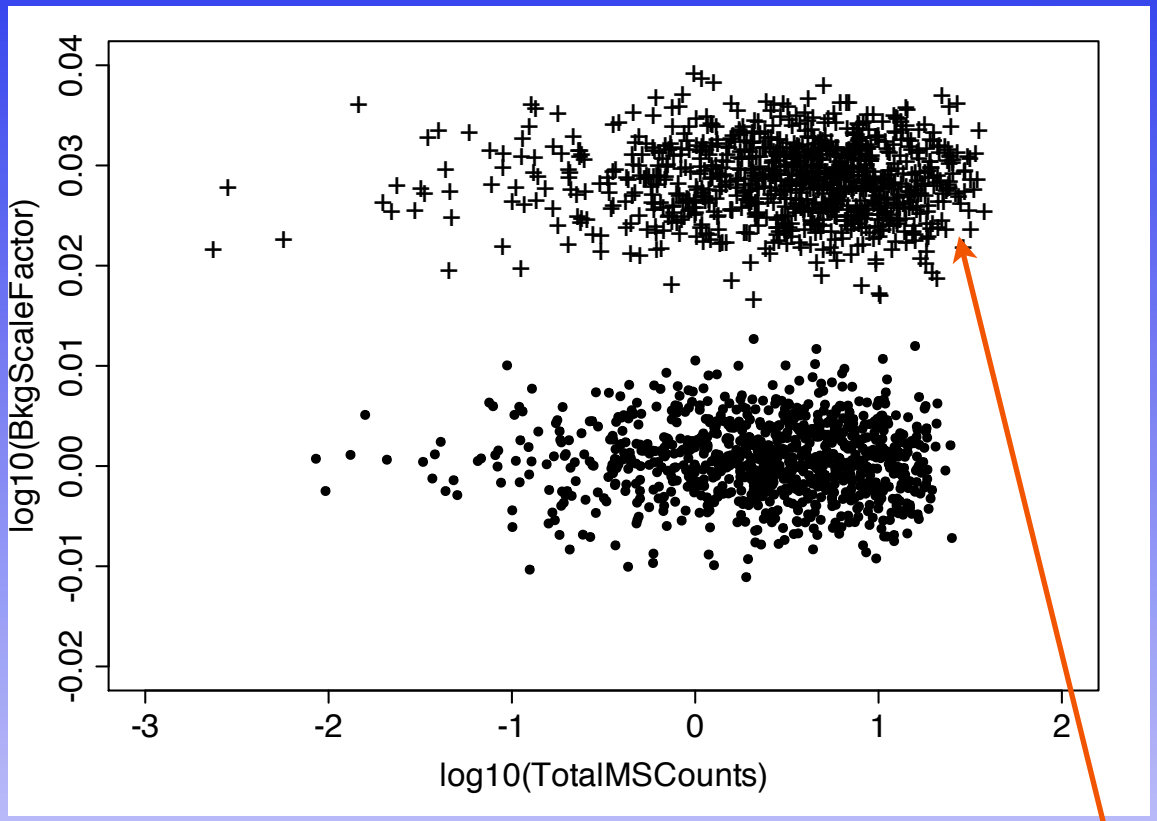


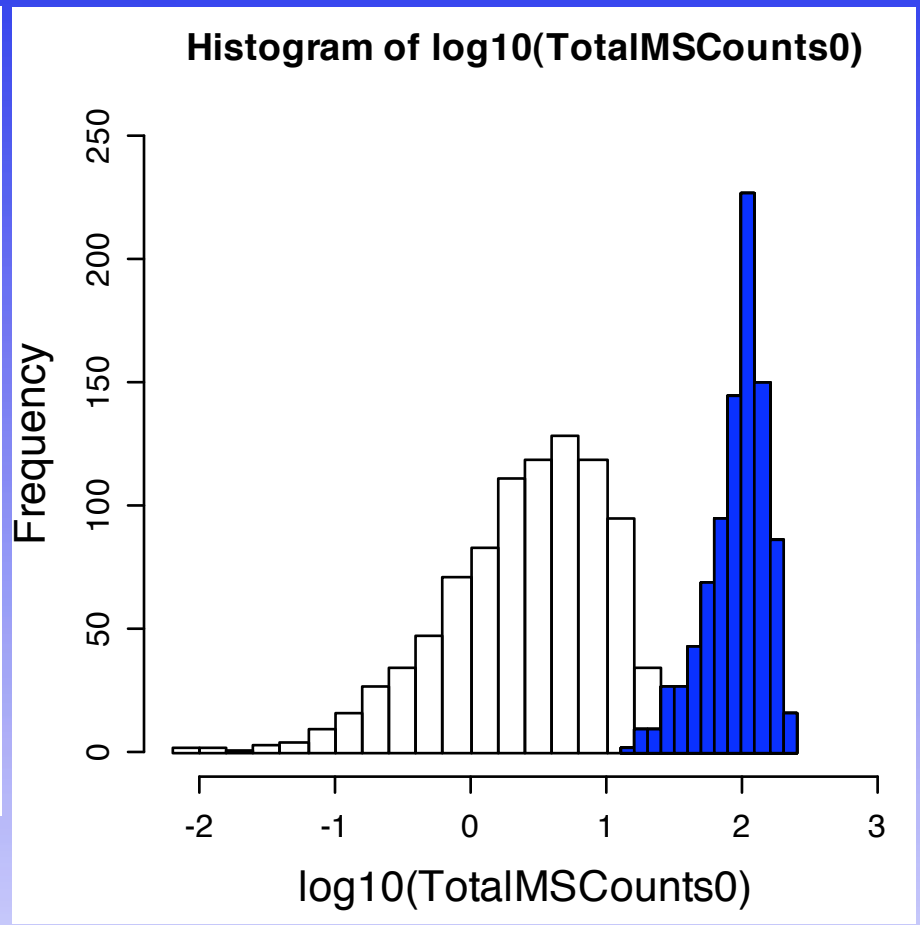
Sigma, Kurtosis

(i.e. Sqrt(Variance))

(Recall these give shape, not significance)

What does small model shape mistake look like?





Combined Scatter Plot

background scale factor **FREE**

Note: excess is put into larger scale factor, rather than Multi-Scale component

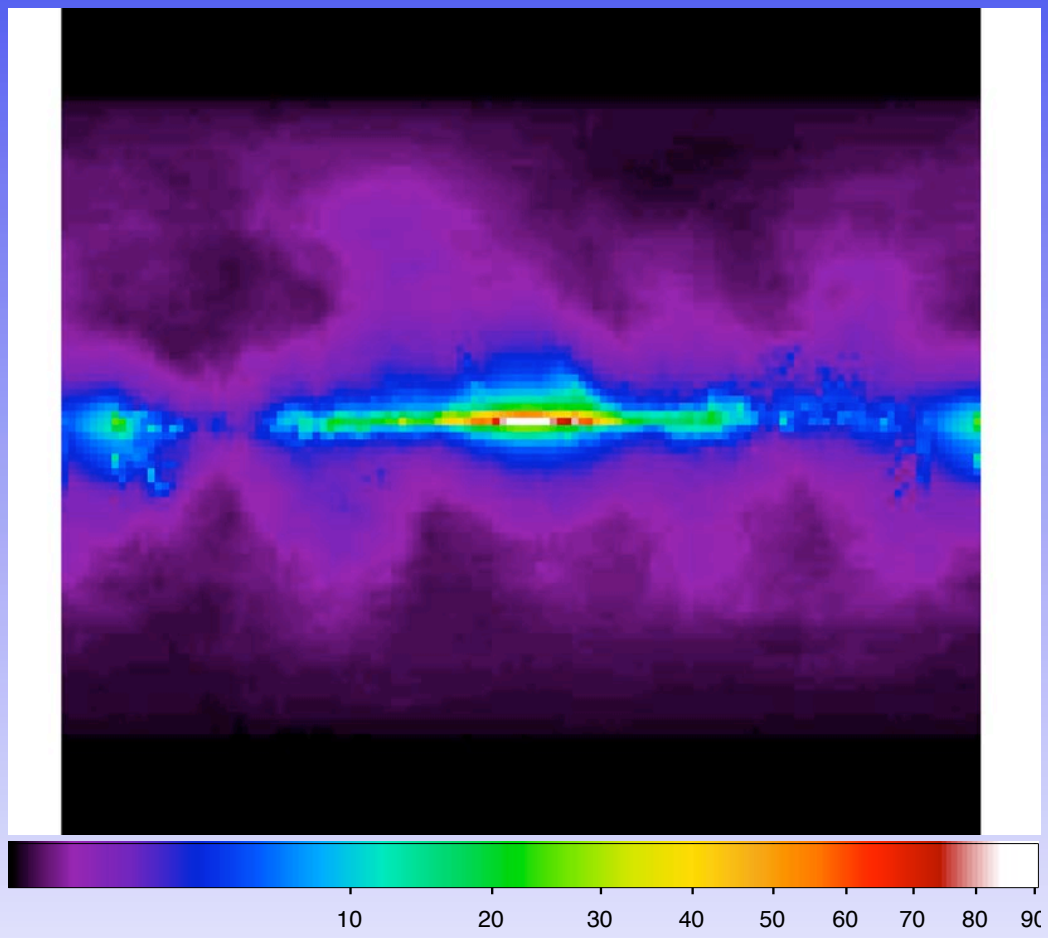
.=null, +=data

Combined Histogram

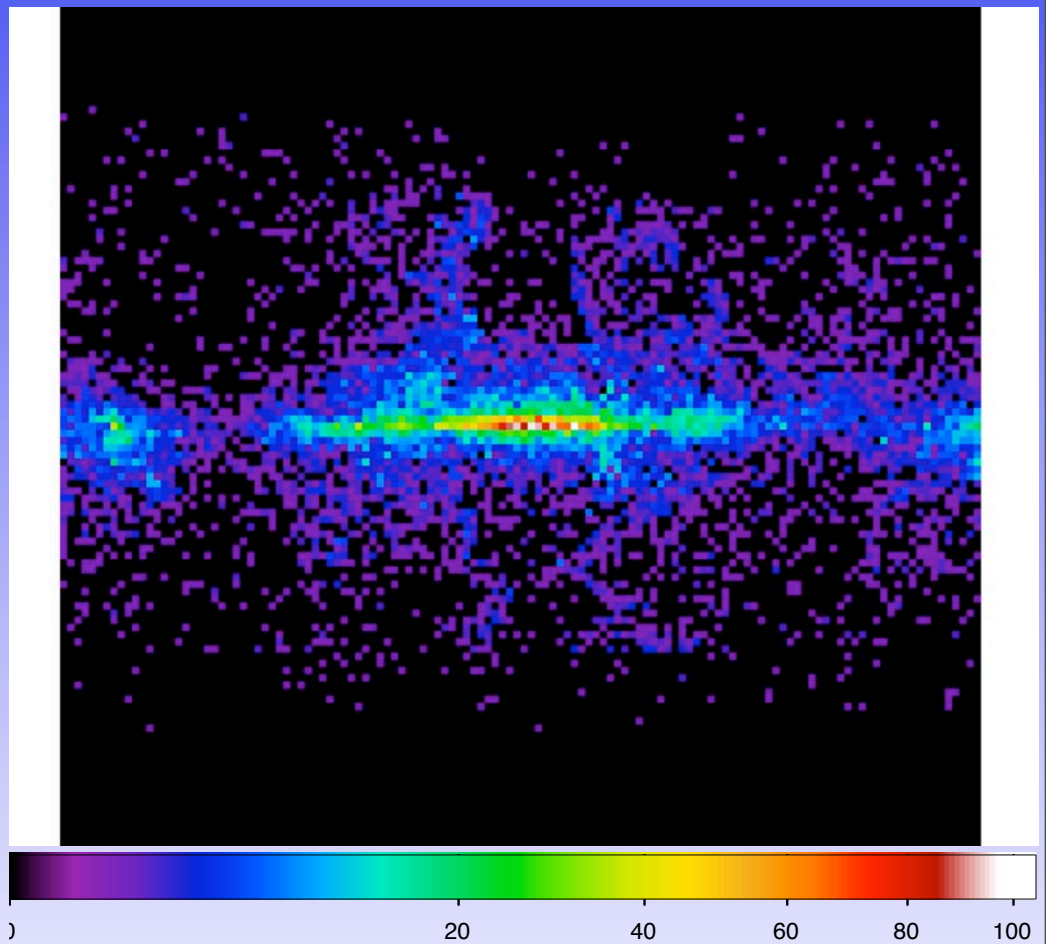
for background scale factor **FIXED**

white=null, blue=data

What does Bright discontinuous source look like, with our method?

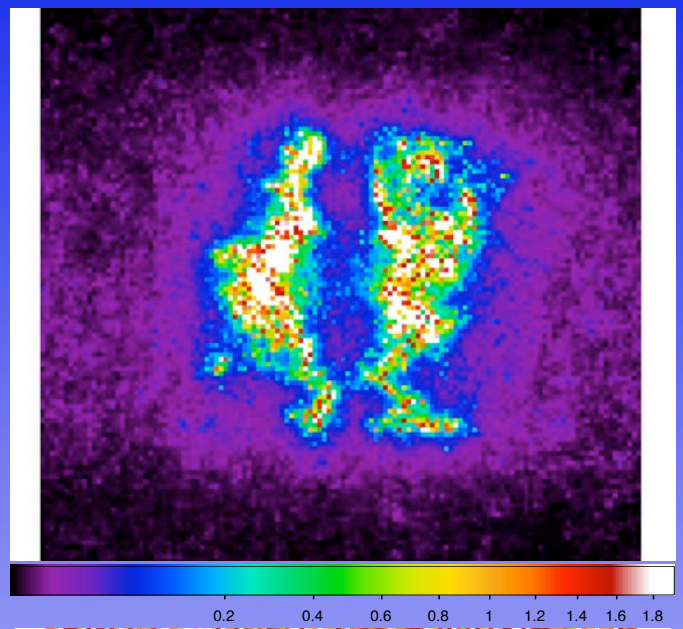


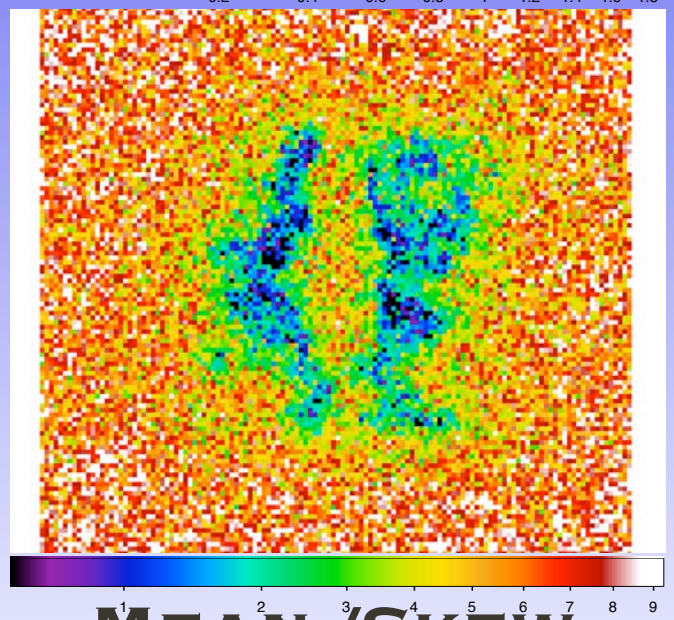
Null Model (i.e. background)



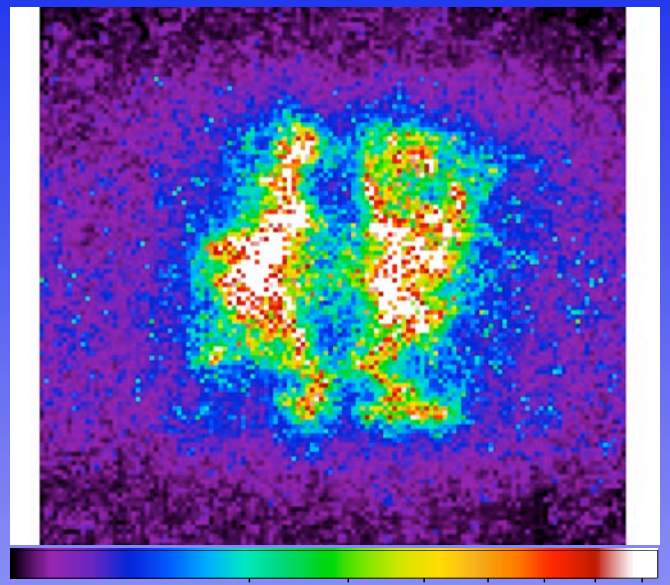
Simulated Data
(extra fingers of gas)

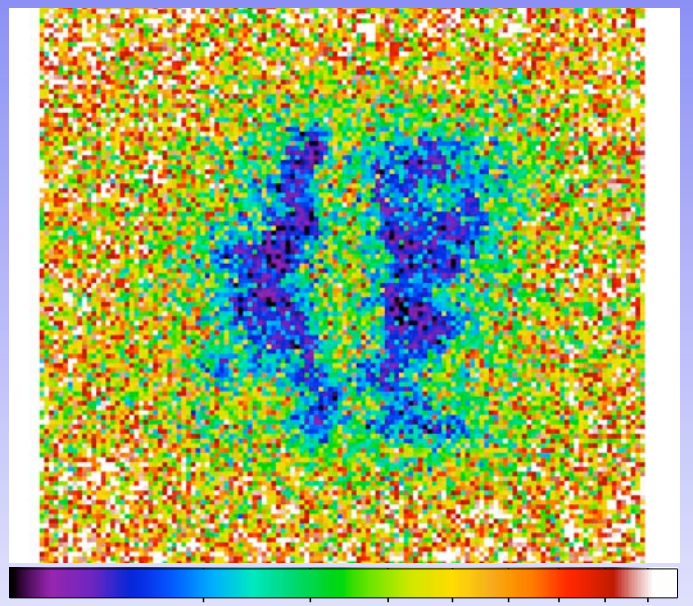
What does bright discontinuous source look like?





Mean /Skew



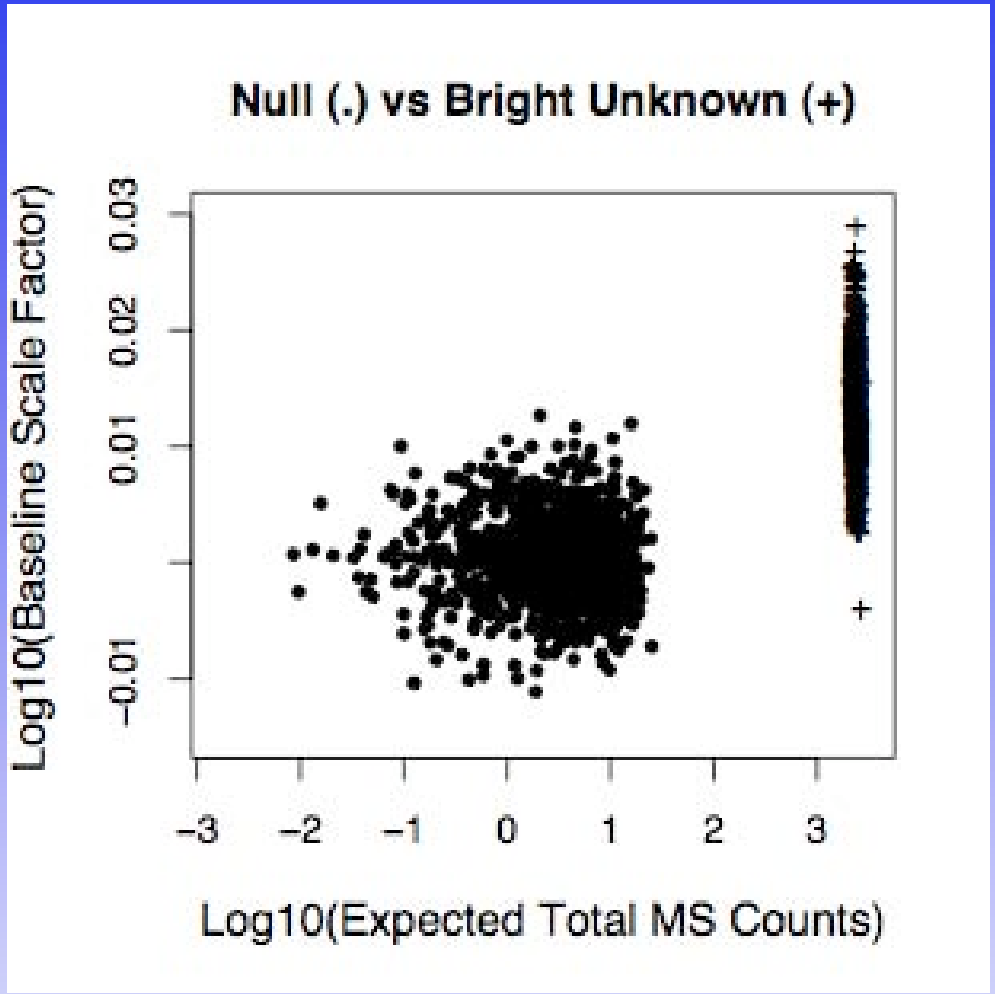


Sigma, Kurtosis

(i.e. Sqrt(Variance))

(Recall these give shape, not significance)

What does bright discontinuous unknown look like?



Combined Scatter Plot

background scale factor **FREE**

Note: some of the excess is put into larger scale factor, rather than Multi-Scale component

So far, looks good!

Help us set priorities!

The authors and CHASC acknowledge support from NSF, Chandra X-Ray Center, AISRP ``Python Tools'', & numpy/pyfits efforts. Special thanks to the organizers of the SAMSI06 Special Program in Astrostatistics, and SCMA IV.

CGRO/EGRET Data courtesy of HEASARC
cossc.gsfc.nasa.gov/docs/cgro/egret/
GALPROP courtesy A. Strong:
www.gamma.mpe.mpg.de/\~~aws/aws.html

AC thanks T. Loredo, M. Karovska, R. Willett, C.A. Young, E. Kolaczyk for wide-ranging discussions of imaging methods in astronomy and statistics.

FOUR RULES:

RESPECT THE DATA

No "binning up" - you lose information and *you don't need to* (see Scargle etc).
No (or minimal) filtering or pre-processing
No subtracting ("model out")
Cut the 'cuts'
Data exploration/visualization is DIFFERENT than inference
(And both are useful

RESPECT THE UNDERLYING DISTRIBUTION (LIKELIHOOD-BASED)

(COS B, SAS 2, CGRO, ...)
IF you want to know uncertainties BUT beware model incompleteness

RESPECT YOUR KNOWNS

IE Be aware of, and use, assumptions in your model;
Put in the actual knowledge you have

RESPECT YOUR UNKNOWNS

i.e. Respect your uncertain background and calibration 'constants' (EffArea, etc.)
They have a distribution, too.