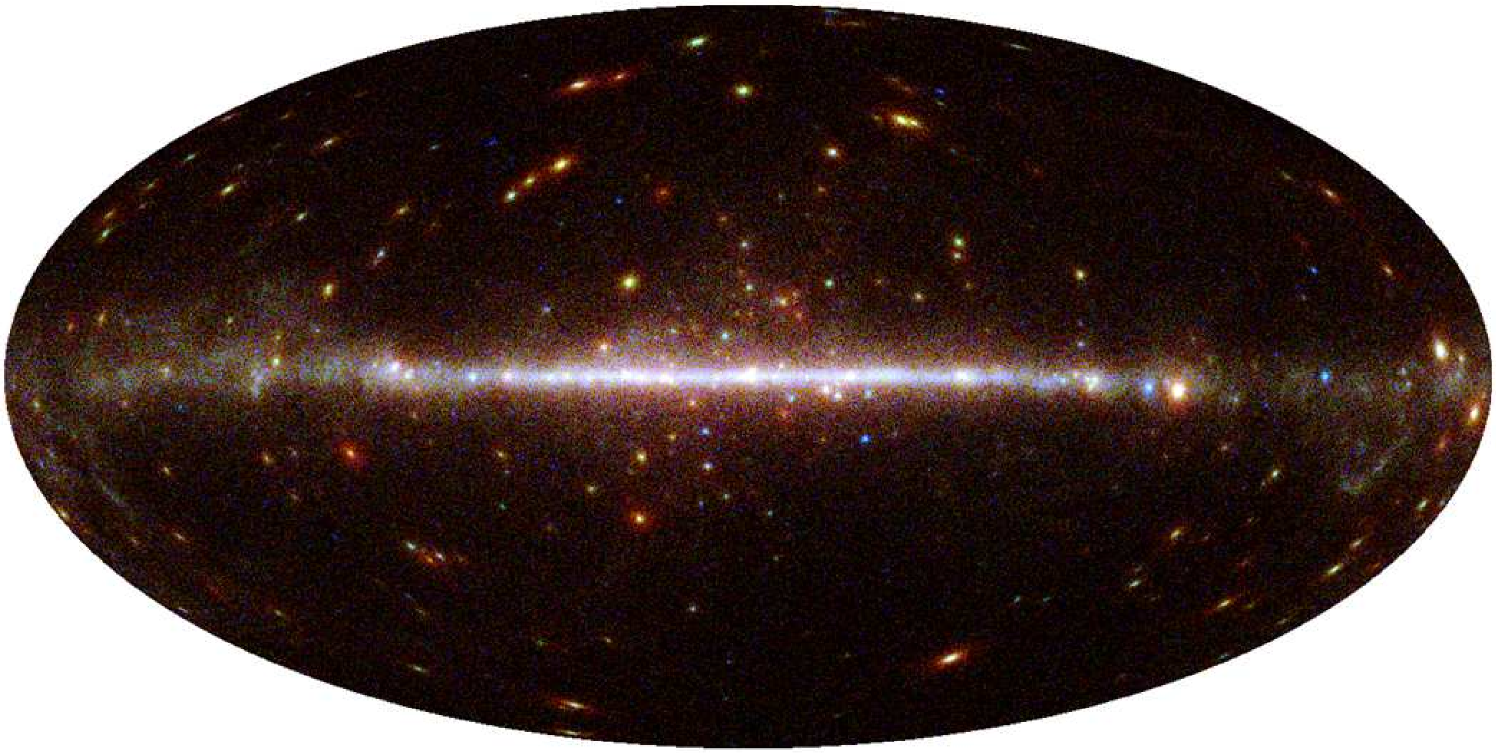
Image Reconstruction for GLAST/LAT Data Using EMC2†

Jim Chiang (SLAC)
Current Collaborators: David van Dyk, Adam Roy (UC Irvine), Alanna Connors (Eureka)
Past Participants: D. Esch (Harvard), M. Korovska (CfA)

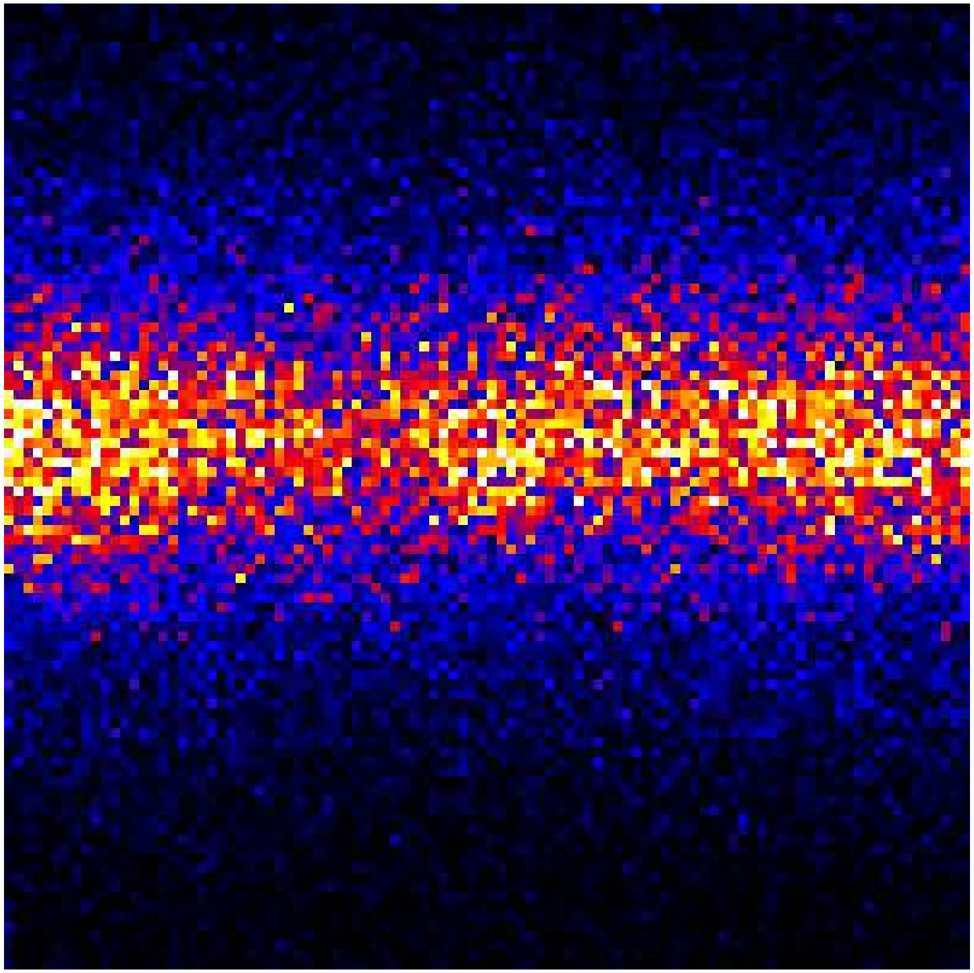
†Expectation through Markov Chain Monte Carlo

The Gamma-Ray Sky as (may be) seen by GLAST



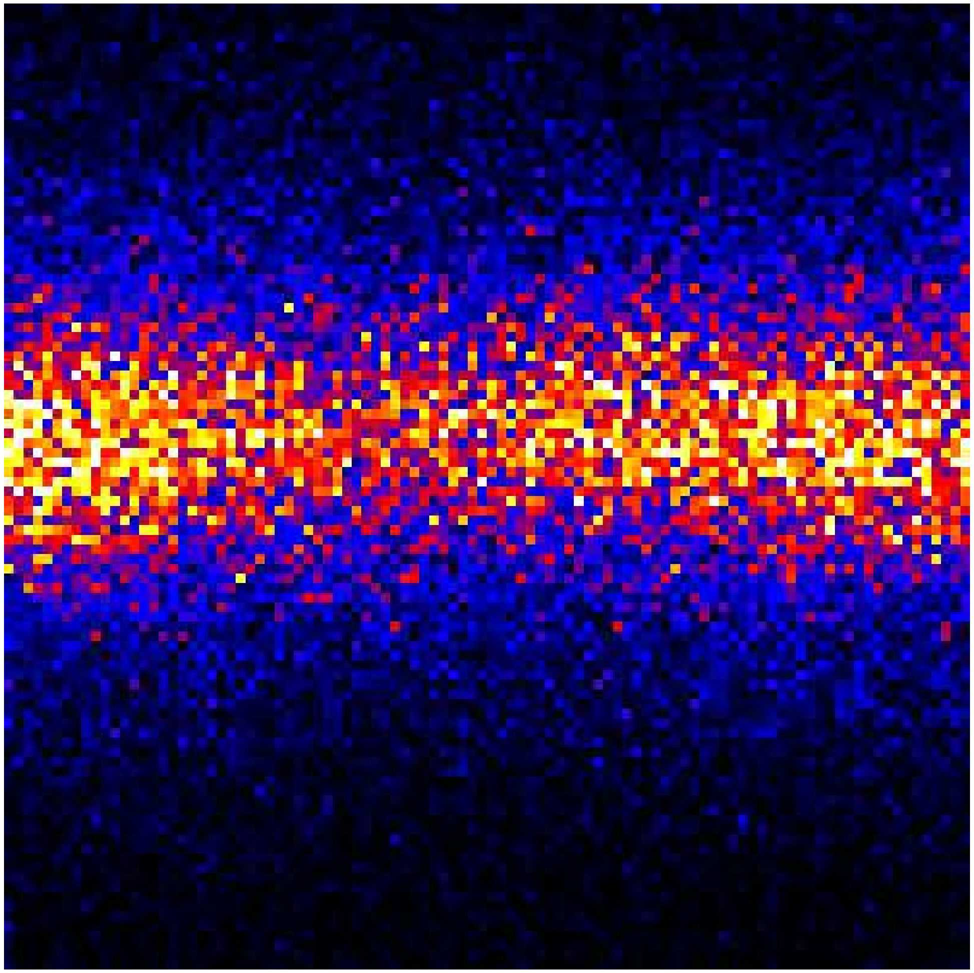
55 day simulation from GLAST Data Challenge II

Find the SNR in These Data!



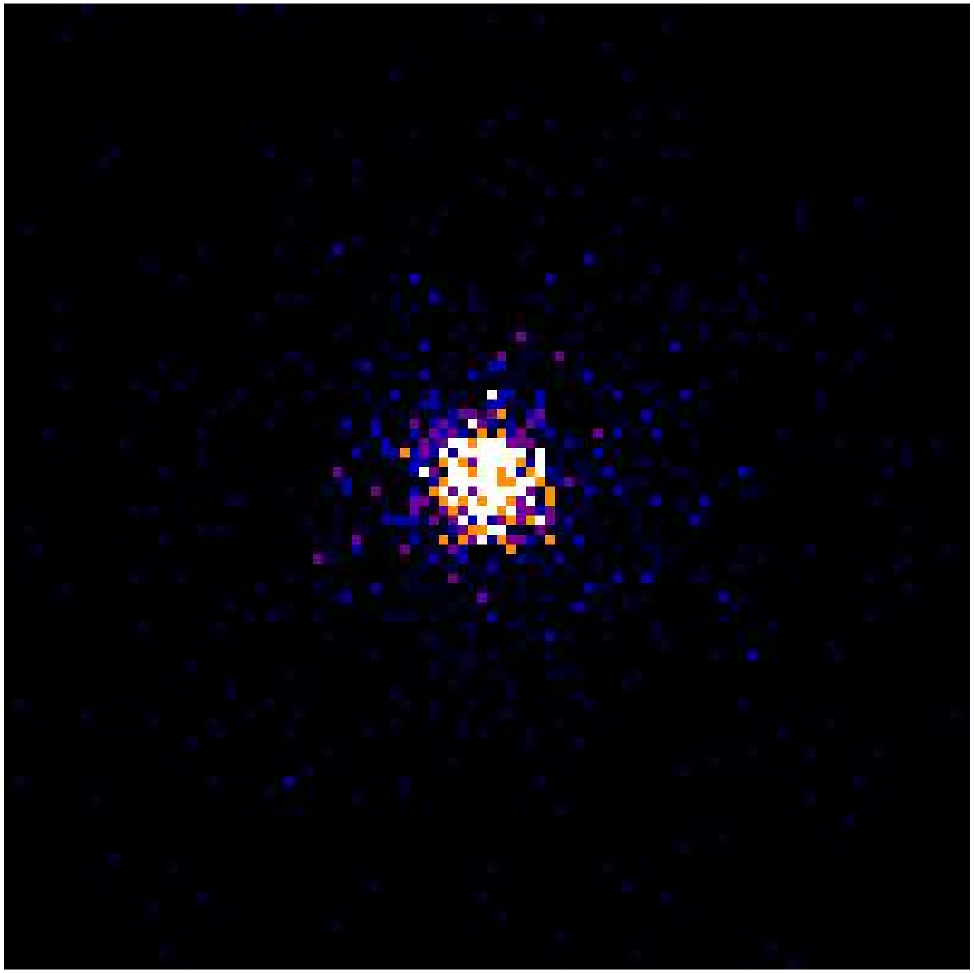
GALPROP diffuse + RX J1713.7−3946, 1 year LAT sim, $10^\circ \times 10^\circ$ field

Find the SNR in These Data!



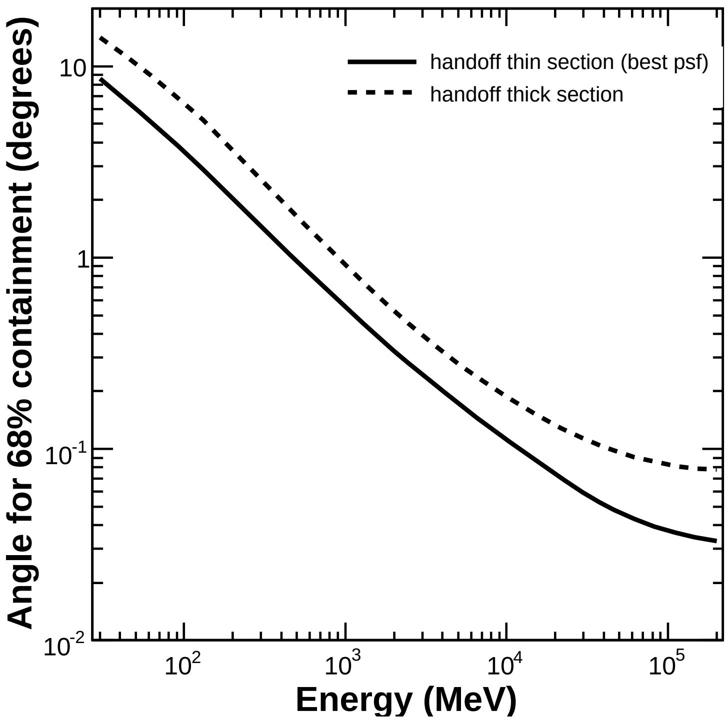
GALPROP diffuse, 1 year LAT sim, $10^\circ \times 10^\circ$ field

Find the SNR in These Data!

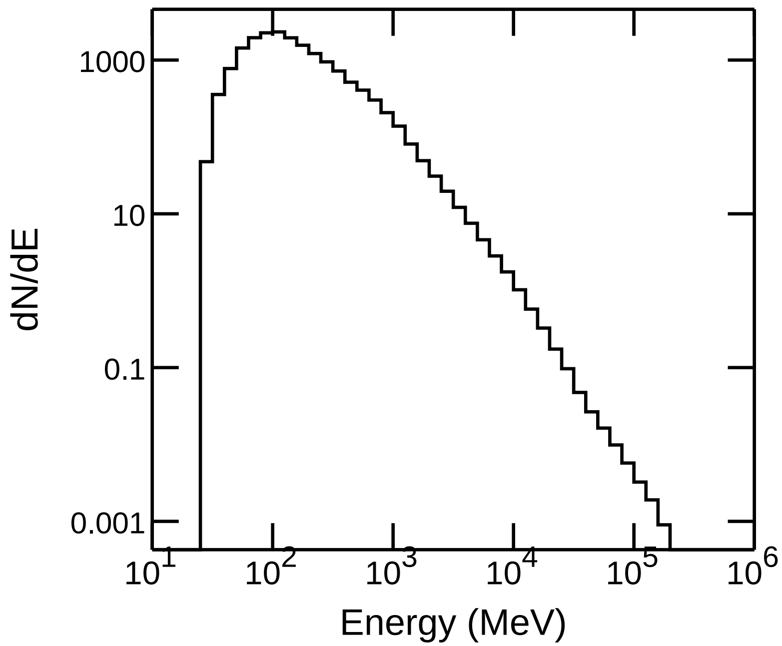


RX J1713.7−3946, 1 year LAT sim, $10^\circ \times 10^\circ$ field

LAT PSF and Counts Spectra



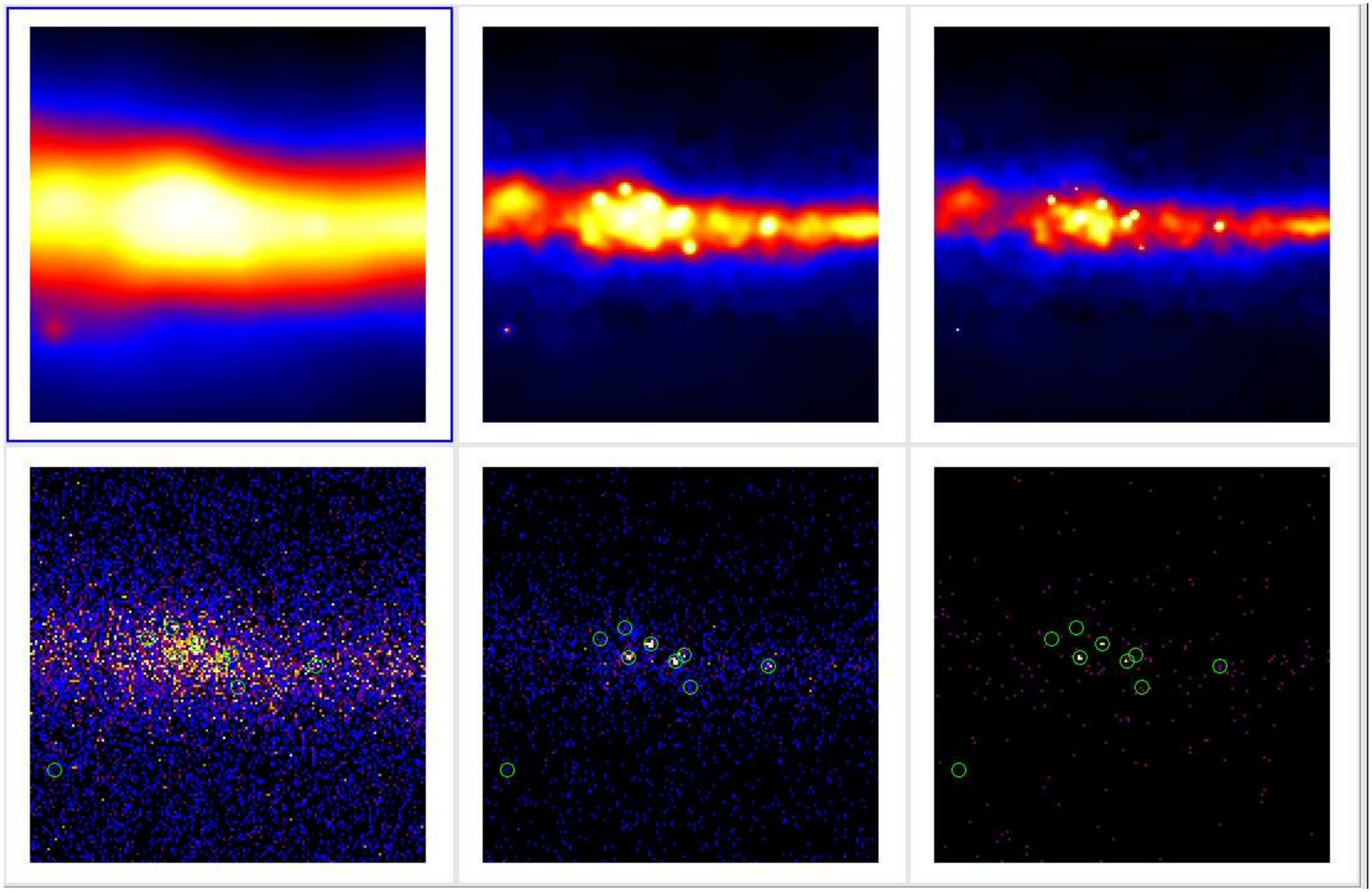
$\theta_{68} \sim E^{-0.8}$



$dN/dE \sim E^{-2}$ for $E > 100$ MeV

Energy-dependent Blurring in Poisson Limit

GALPROP + 3EG sources in Cygnus Region



100 MeV 1 GeV 10 GeV

Richardson-Lucy: Expectation-Maximization for Image Reconstruction

Consider a model image $\Lambda = \{\lambda_i\}$, where λ_i is the model count in pixel i. For image reconstruction, we formulate the problem in terms of "missing data",

$$x_i \overset{d}{\sim} \text{Poisson}(\lambda_i) = e^{-\lambda_i} \lambda_i^x / x! \tag{1}$$

Here $X = \{x_i\}$ are the source counts detected in each pixel for an ideal instrument, i.e., without effects from PSF, pile-up, etc..

In the case of PSF blurring, $A = A_{ij}$, the *observed* data, $Y = \{y_j\}$, are given by

$$y_j \overset{d}{\sim} \text{Poisson}(\sum_i A_{ij} \lambda_i) \tag{2}$$

We could write the log-likelihood in terms of Y,

$$\log \mathcal{L}(\Lambda|Y) = \sum_j \left[-\sum_i A_{ij}\lambda_i + y_j \log \left(\sum_i A_{ij}\lambda_i \right) \right] + \text{terms independent of } \Lambda, \tag{3}$$

and optimize wrt $\{\lambda_i\}$, but it is easier to write the likelihood in terms of x_i and apply EM:

$$\log \mathcal{L}(\Lambda|X) = \sum_i (-\lambda_i + x_i \log \lambda_i) \tag{4}$$

EM Implementation

E-Step: Given a $\Lambda = \Lambda^{(t)}$, at iteration t, replace X in the log-likelihood by its **conditional expectation**:

$$x_i \rightarrow E(x_i|Y, \Lambda^{(t)}) \tag{5}$$

where $E(\cdots)$ is the average of x_i weighted by the conditional probability $P(x_i|Y, \Lambda^{(t)})$. This probability can also include priors on Λ.

M-Step: Given X, update Λ by maximizing the resulting log-likelihood:

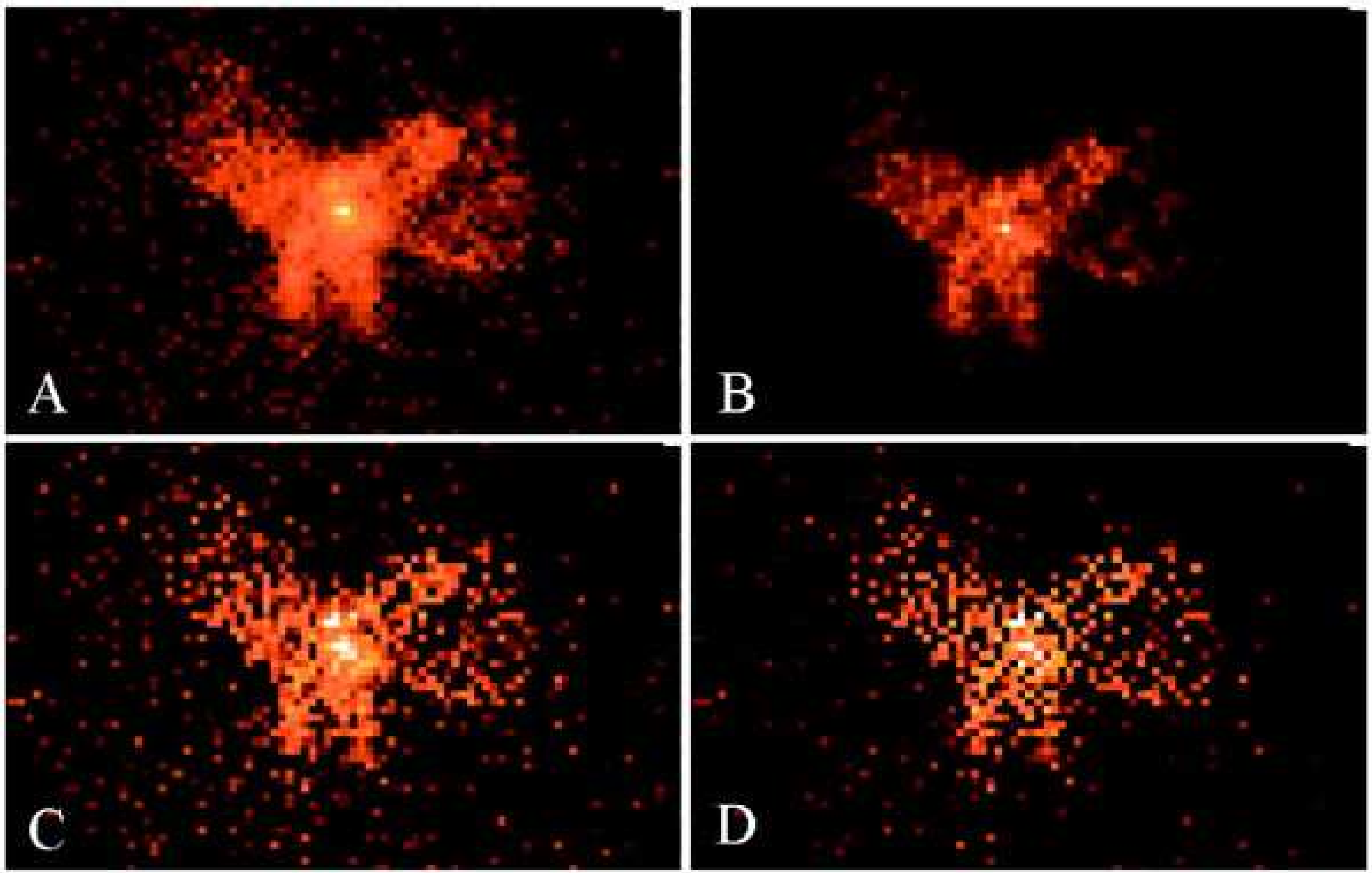
$$\frac{\partial \mathcal{L}(\Lambda|X)}{\partial \lambda_i} = -1 + E(x_i|Y, \Lambda^{(t)})/\lambda_i = 0 \tag{6}$$

$$\Rightarrow \lambda_i^{(t+1)} = E(x_i|Y, \Lambda^{(t)}) \tag{7}$$

Richardson (1972) and Lucy (1974) use Bayes's Theorem to compute $E(x_i|Y, \Lambda^{(t)})$:

$$\lambda_i^{(t+1)} = \lambda_i^{(t)} \sum_k A_{ik} \frac{y_k}{\sum_j A_{jk} \lambda_j^{(t)}} \tag{8}$$

RL Applied to Chandra Data of NGC 6240



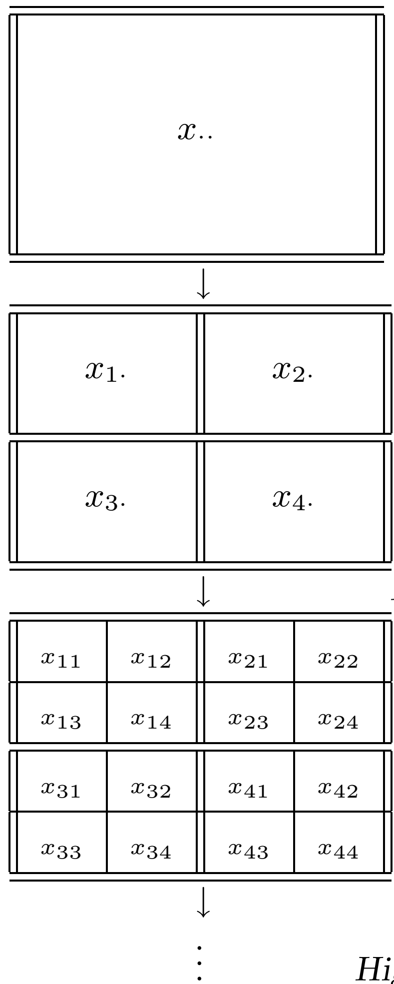
A) Chandra data, B) EMC2 reconstruction, C) RL at 20 iterations, D) RL at 100 iterations.

Motivations for a Multiscale Poisson Framework

- RL tends to amplify Poisson fluctuations, especially for low count data. This can be ameliorated by including regularization or smoothing in the reconstruction. Often this takes the form of a prior distribution.
- Most image data have spatial structure on a variety of different scales, with different amounts of intrinsic smoothness depending on the underlying physics. This information can also be folded in to a prior distribution.
- Wavelet-based regularization methods capture the multiscale aspects of the data, but are better-suited to Gaussian problems.
- Nowak & Kolaczyk (NK, 2000) have developed a multiscale framework that is inspired by wavelet methods, but is suited to Poisson statistics. Furthermore, it is amenable to an EM algorithm wherein the maximization step can be expressed as a closed form solution and so is computationally efficient.

Multiscale Representation and Prior Distributions

Low Resolution



$$x_{\cdot\cdot} \overset{d}{\sim} \mathrm{Poisson}(\lambda)$$
$$\lambda \overset{d}{\sim} \mathrm{Gamma}\{(\alpha_0, \beta)\}$$

$$x_{,\cdot}|x_{\cdot\cdot} \overset{d}{\sim} \mathrm{Multinomial}(p_1)$$
$$\rho_1 \overset{d}{\sim} \mathrm{Dirichlet}\{(\alpha_1, \alpha_1, \alpha_1, \alpha_1)\}$$

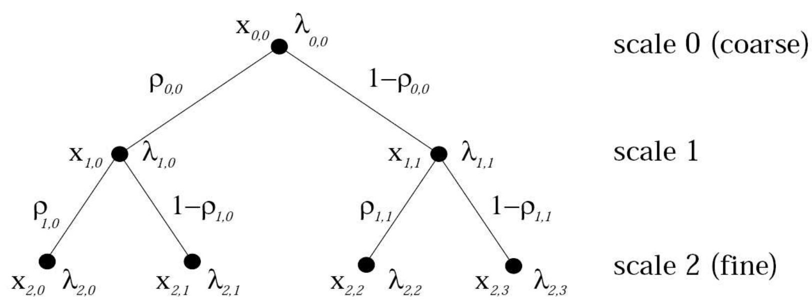
Prior shrinks toward Smooth images

$$x_{i,}|x_{i\cdot} \overset{d}{\sim} \mathrm{Multinomial}(p_{2i})$$
$$\rho_{2i} \overset{d}{\sim} \mathrm{Dirichlet}\{(\alpha_2, \alpha_2, \alpha_2, \alpha_2)\}$$

High Resolution

Splitting Factors (in 1D)

Splitting factors are introduced via a prior distribution that controls the amount of smoothing for each level of resolution. These factors determine the assignment of model counts in going from a coarser level to the next level of refinement. In 1D, the effect of these factors can be depicted as



The ρ_{ji}s are drawn from a Beta distribution,

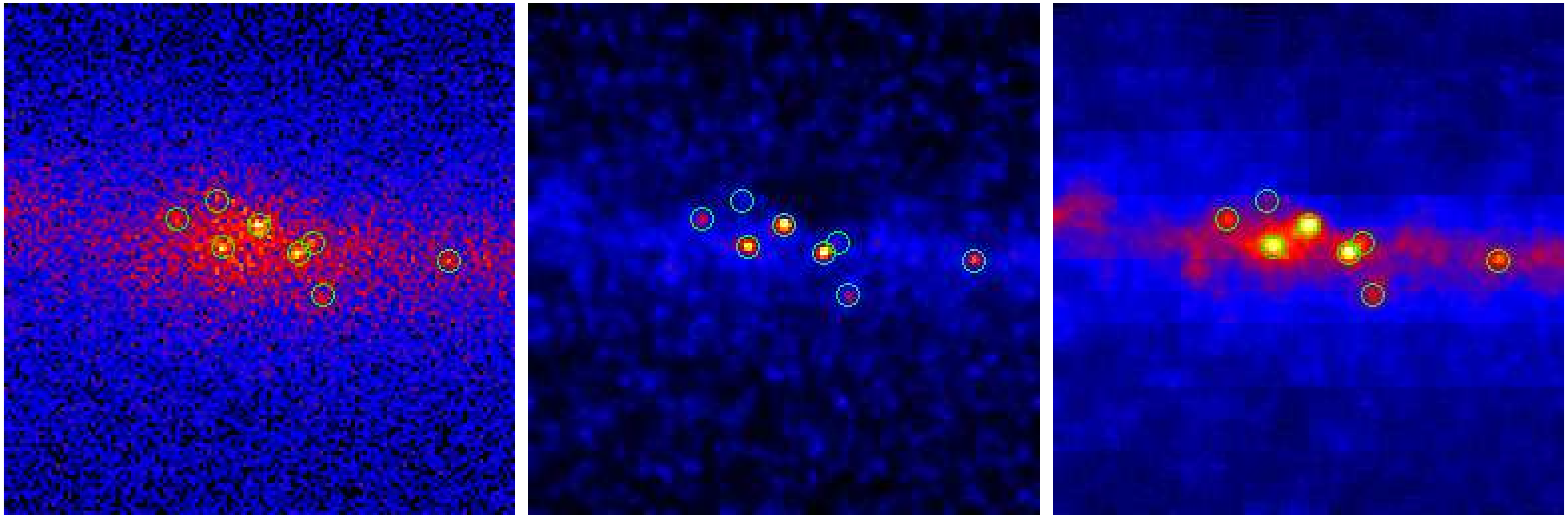
$$\rho_{j,m} \overset{d}{\sim} \rho^{\alpha_j - 1}(1-\rho)^{\beta_j - 1}/B(\alpha_j, \beta_j) \tag{9}$$

where j indicates the resolution level, m the pixel index at that level, and $B(\alpha, \beta)$ is the beta function. If we take $\alpha_j = \beta_j$, then the maximum *a posteriori* estimate (M-step) of the splits are

$$\hat{\rho}_{j,m} = \frac{x_{j+1,2m} + \alpha_j - 1}{x_{j,m} + 2(\alpha_j - 1)} \tag{10}$$

The α_j determine the amount of smoothing: $\hat{\rho}_{j,m} \to 1/2$ for $\alpha_j \to \infty$; and $\hat{\rho}_{j,m} = \frac{x_{j+1,2m}}{x_{j,m}}$ for $\alpha_j = 1$.

RL and NK Applied to LAT Simulations



Simulated counts map — Richardson-Lucy — Multiscale Poisson (Nowak & Kolaczyk 2000)

"Cycle spinning" can be applied to remove the blocky structures from the NK reconstruction.

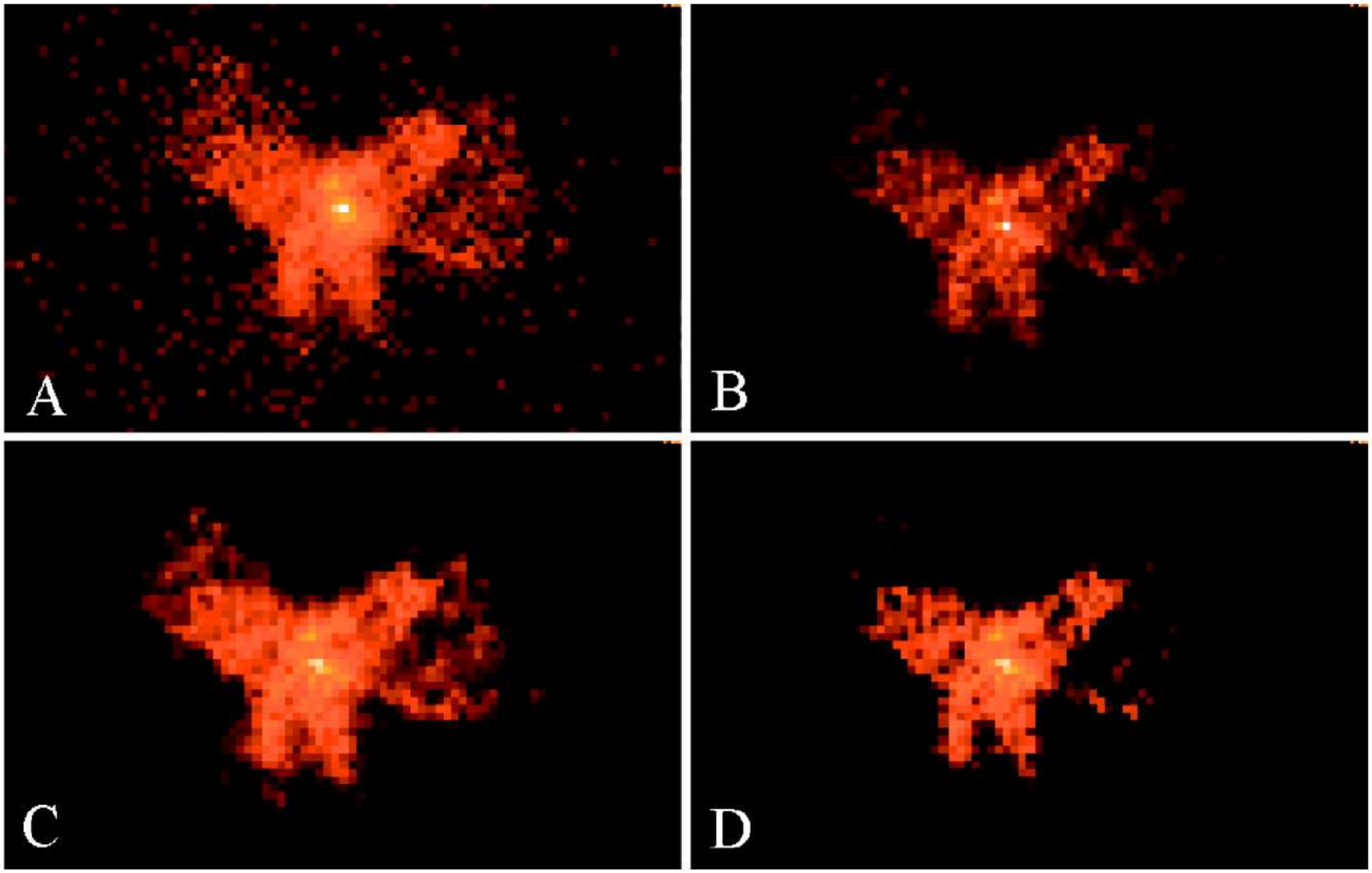
EMC2 (Expectation through Markov Chain Monte Carlo)

Esch et al. 2004, ApJ, 610, 1213; Connors & van Dyk 2006, SCMA, 2006

- Combines EM algorithm with a Markov Chain to provide information on the posterior distribution of model pixel intensities $\Rightarrow$ feature significance and uncertainties.
- "Hyper-prior" distribution for α_js
- Alternatively, one can use multiwavelength data and/or physical models to set α_js to appropriate values.
- Allows for a background or null model to be specified.

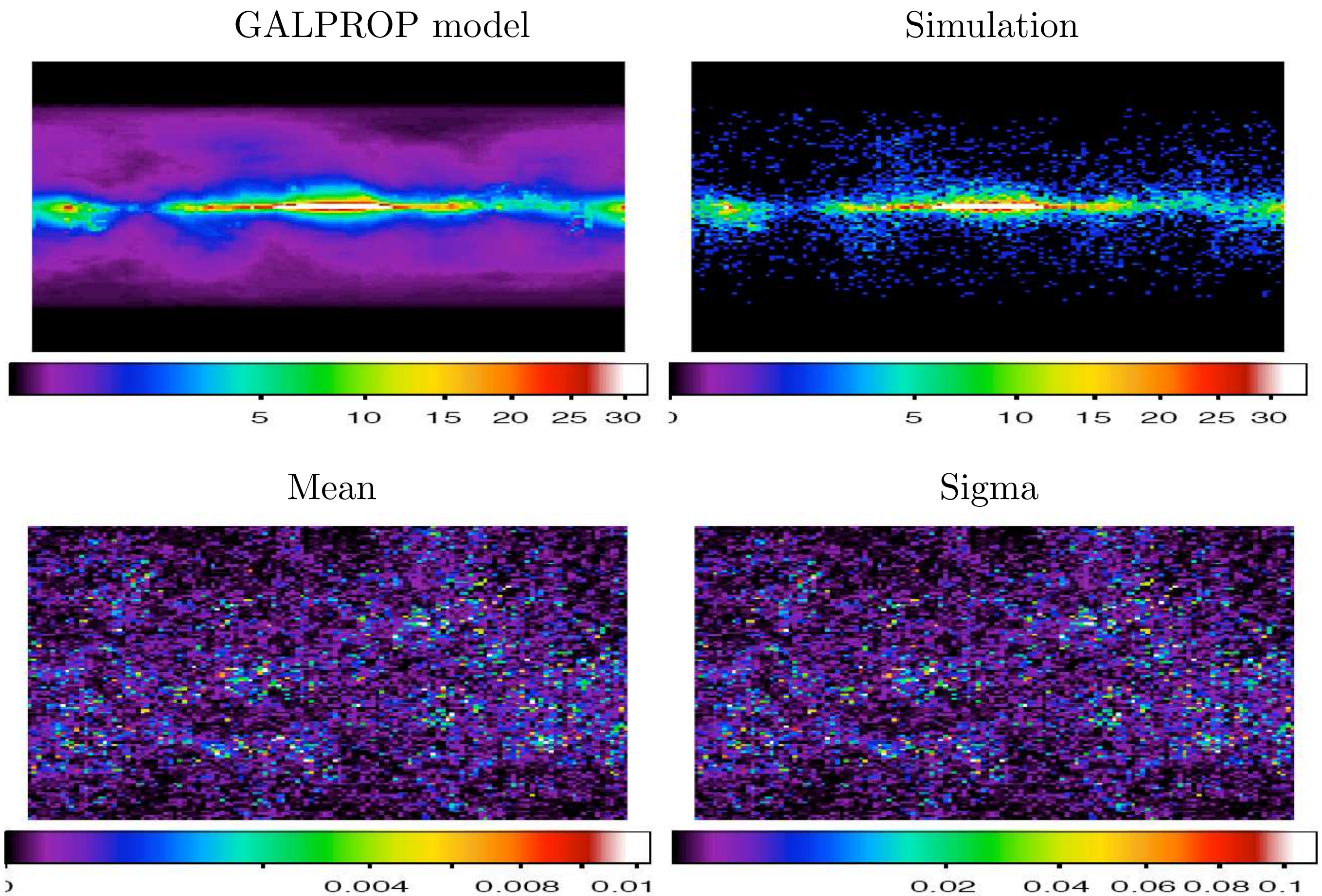
– Extra computational cost because of MCMC steps.

EMC2 Applied to Chandra Data

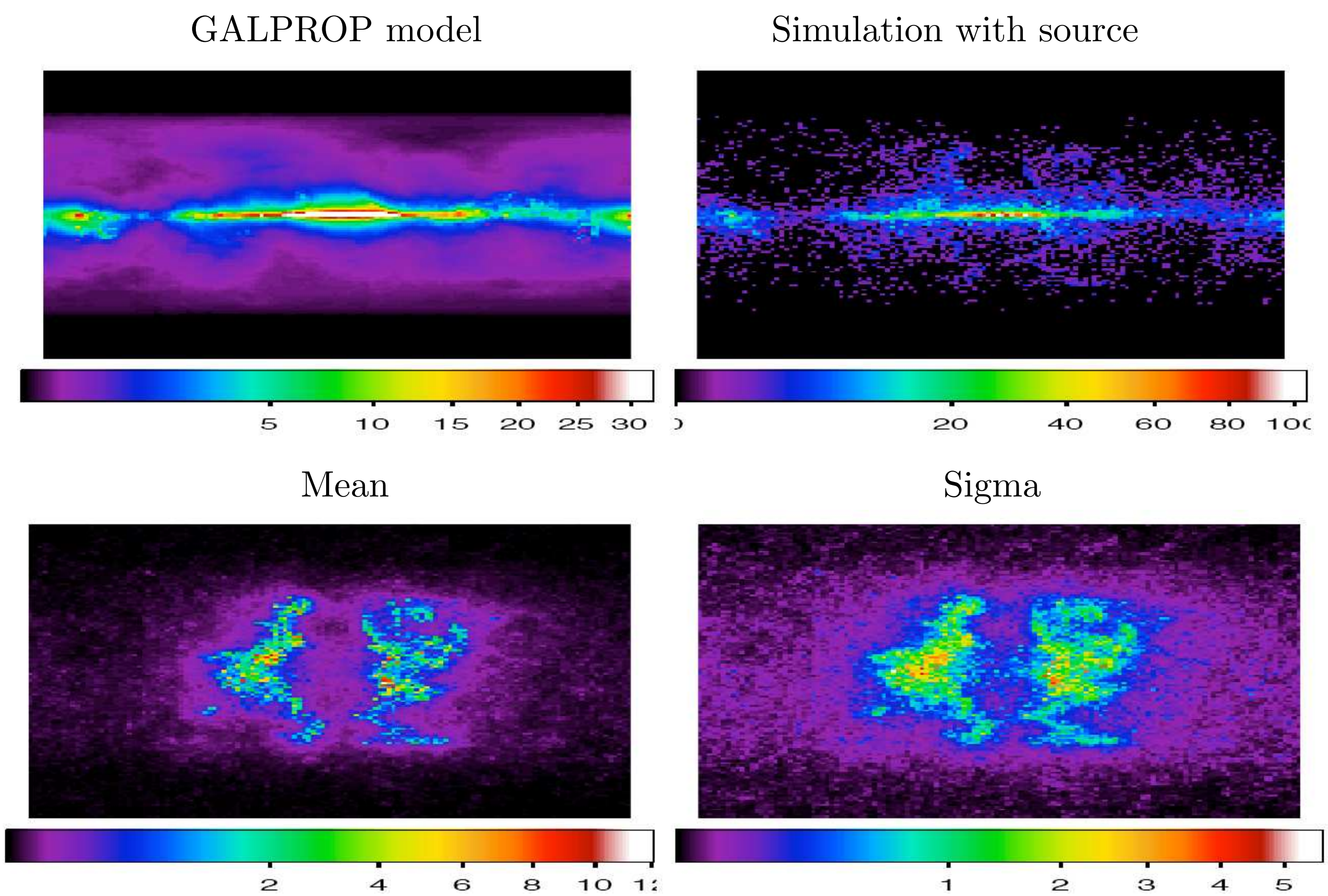


A) Chandra data, B) EMC2 reconstruction, C) 1-sigma significance map, D) 3-sigma map.

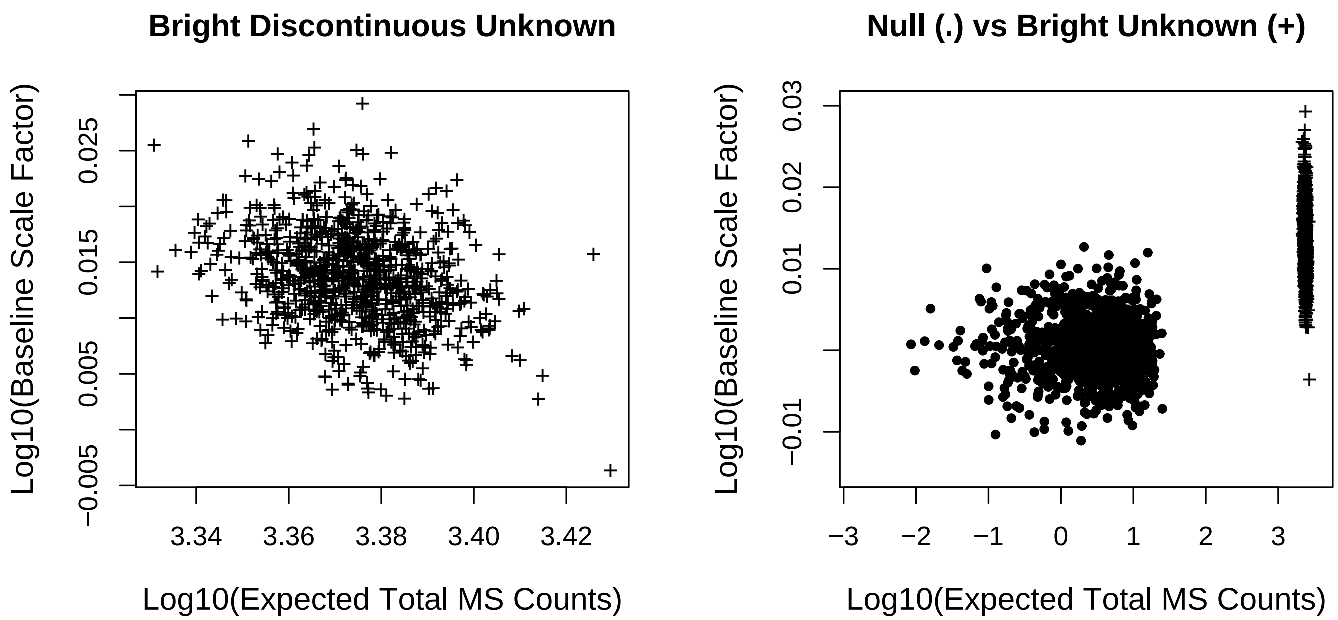
Including a Null Model: EGRET Analysis



Reconstruction with a Null Model



Assessing the Significance of an Extra Component



Extensions for GLAST/LAT Data

- Use radio survey data of gas distributions and GALPROP predictions to inform choice of smoothing parameters α_js.
- Use energy-dependent PSF information through joint prior distributions inferred from reconstructions in different energy bands.
- HEALPix representation for all-sky reconstructions: