Overlapping Astronomical Sources: Utilizing Spectral Information

David Jones

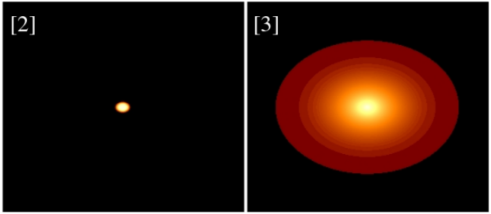
Advisor: Xiao-Li Meng

Collaborators: Vinay Kashyap (CfA) and David van Dyk (Imperial College)
CHASC Astrostatistics Group

April 15, 2014

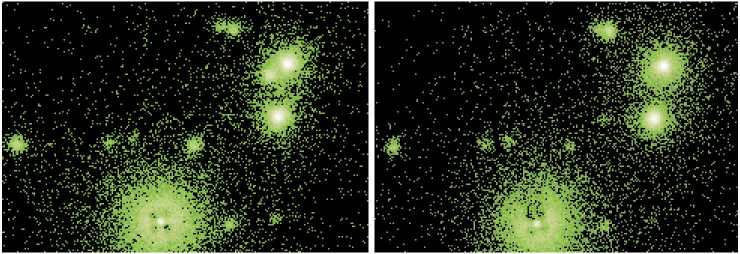
Introduction

- X-ray telescope data:
  - spatial coordinates of photon detections
  - photon energy
- Instrument error: diffraction in the telescope means recorded photon positions are spread out according to the point spread function (PSF)



Introduction

- PSFs overlap for sources near each other
- Aim: inference for number of sources and their intensities, positions and spectral distributions
- Key points: (i) obtain posterior distribution of number of sources, (ii) use spectral information



Basic Model and Notation

(x_i, y_i) = spatial coordinates of photon i
k = # sources
$\boldsymbol{\mu}_j$ = centre of source j
s_i = latent variable indicating which source photon i is from
$n_j = \sum_{i=1}^{n} 1_{\{s_i=j\}} =$ # photons detected from component $j \in \{0, 1, \ldots, k\}$

$$\begin{aligned}
(x_i, y_i)|s_i = j, \boldsymbol{\mu}_j, k &\sim \text{PSF}_j \text{ centred at } \boldsymbol{\mu}_j \text{ for } i = 1, \ldots, n \\
(n_0, n_1, \ldots, n_k)|w, k &\sim \text{Mult}(n; (w_0, w_1, \ldots, w_k)) \\
(w_0, w_1, \ldots, w_k)|k &\sim \text{Dirichlet}(\lambda, \lambda, \ldots, \lambda) \\
\boldsymbol{\mu}_j|k &\sim \text{Uniform over the image } \ j = 1, 2, \ldots, k \\
k &\sim \text{Pois}(\theta)
\end{aligned}$$

- Component with label 0 is background and its "PSF" is uniform over the image (so its "centre" is irrelevant)
- Reasonably insensitive to θ, the prior mean number of sources

3rd Dimension: Spectral Data

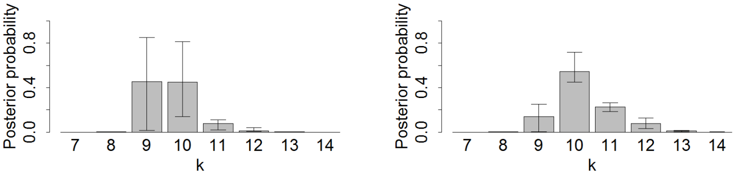
We can distinguish the background from the sources better if we jointly model spatial and spectral information:

$$\begin{aligned} e_i | s_i = j, \alpha_j, \beta_j &\sim \text{Gamma}(\alpha_j, \beta_j) \text{ for } j \in \{1, \ldots, k\} \\ e_i | s_i = 0 &\sim \text{Uniform to some maximum} \\ \alpha_j &\sim \text{Gamma}(a_\alpha, b_\alpha) \\ \beta_j &\sim \text{Gamma}(a_\beta, b_\beta) \end{aligned}$$

Using a (correctly) "informative" prior on α_{s_i} and β_{s_i} versus a diffuse prior made very little difference to results.

Computation: RJMCMC

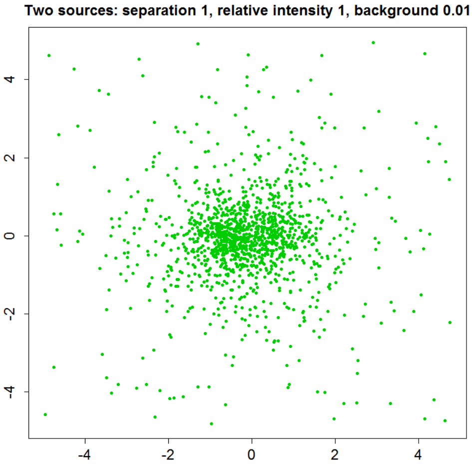
- Similar to Richardson & Green 1997
- Knowledge of the PSF makes things much easier
- Insensitive to the prior $k \sim \text{Pois}(\theta)$ e.g. posterior when $k = 10$:



(a) $\theta = 1$ (b) $\theta = 10$

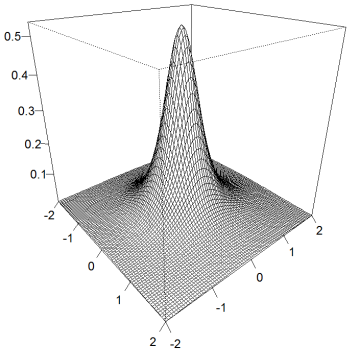
Figure: Average posterior probabilities of each value of k across ten datasets

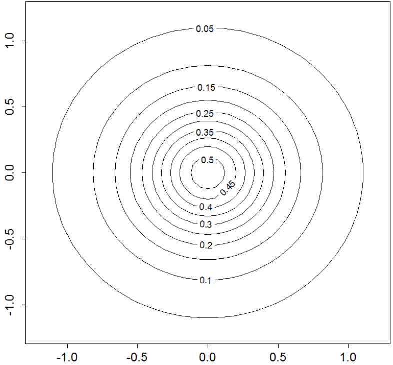
Simulation Study: Example



- 100 datasets simulated for each configuration
- Analysis with and without energy data
- Summarize posterior of k by posterior probability of two sources

Simulation Study: PSF (King 1962)



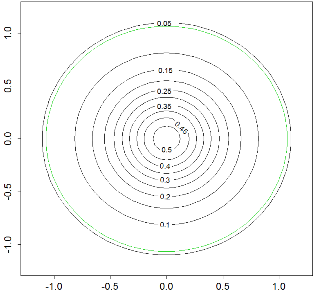


- King density has Cauchy tails
- Gaussian PSF leads to over-fitting in real data

Simulation Study: Data Generation

- Bright source:
$$n_1 \sim \text{Pois}(m_1 = 1000)$$
- Dim source:
$$n_2 \sim \text{Pois}(m_2 = 1000/r)$$
where $r = 1, 2, 10, 50$ gives the relative intensity
- 'Source region': the region defined by PSF density greater than 10% of the maximum (essentially a circle with radius 1)

$d =$ the probability a photon from a source falls within its own region



Simulation Study: Data Generation

- Background per 'source region':

$$\text{Pois}(bdm_2)$$

where relative background $b = 0.001, 0.01, 0.1, 1$

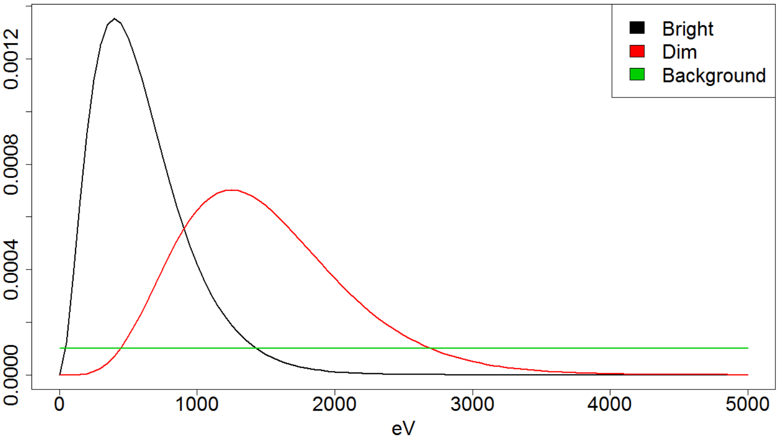
Overall background

$$n_0 \sim \text{Pois}\left(\frac{\text{image area}}{\text{source region area}} bdm_2\right)$$

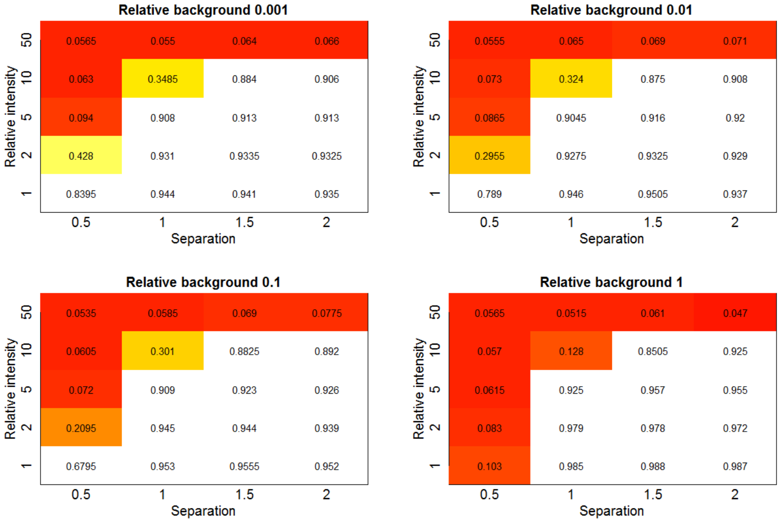
- Separation: the distance between the sources. Values: 0.5, 1, 1.5, 2

Simulation Study: Data Generation

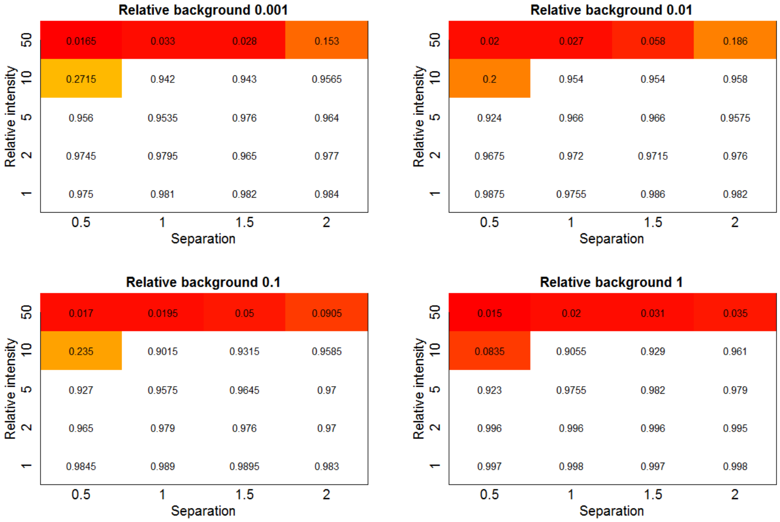
Note: units should be pulse invariant (PI) channel



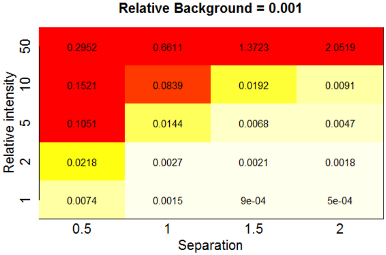
Median Posterior Probability at k=2: No Energy

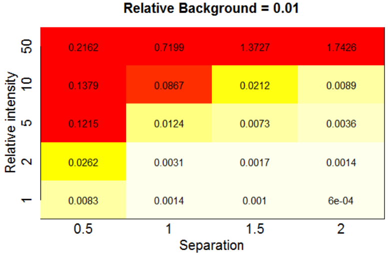


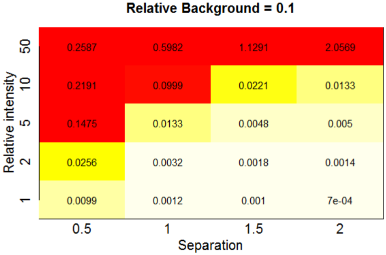
Median Posterior Probability at k=2: Energy

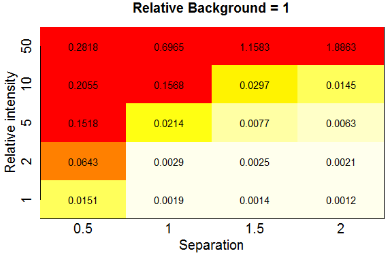


Median SE of Dim Source Posterior Mean Position: No Energy

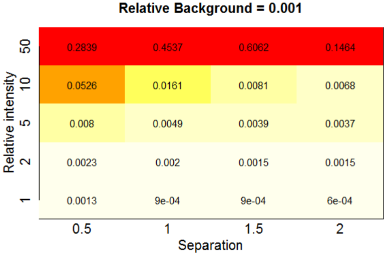


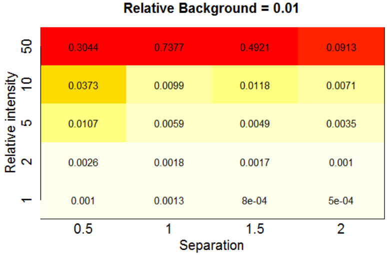


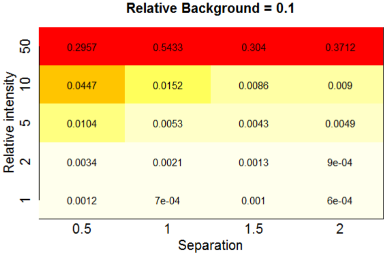


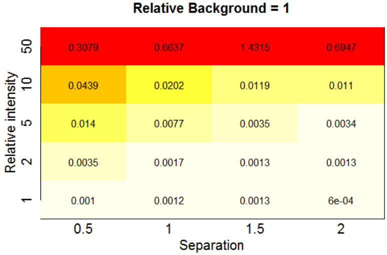


Median SE of Dim Source Posterior Mean Position: Energy

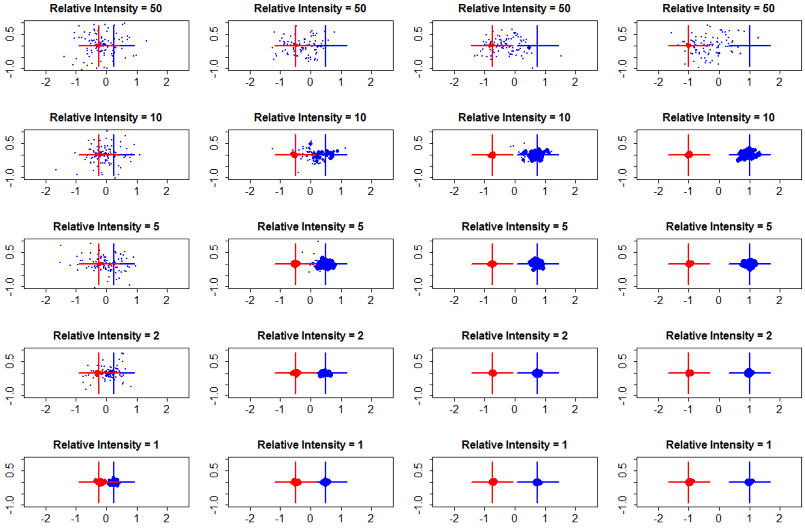




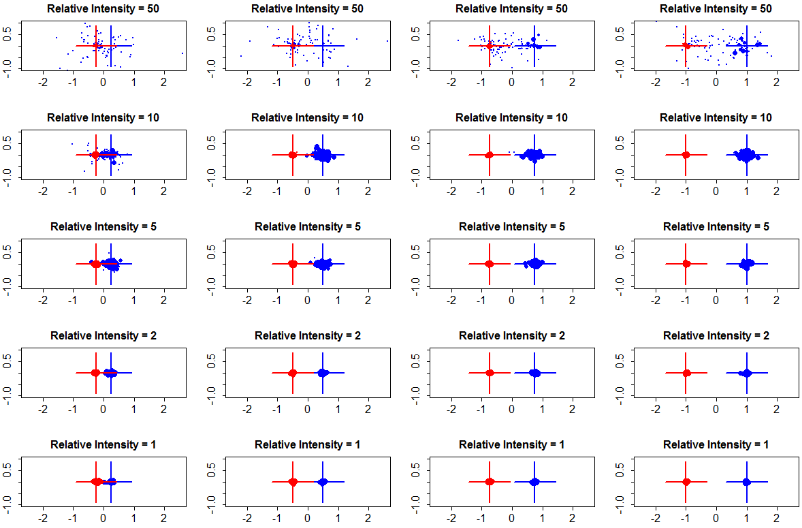




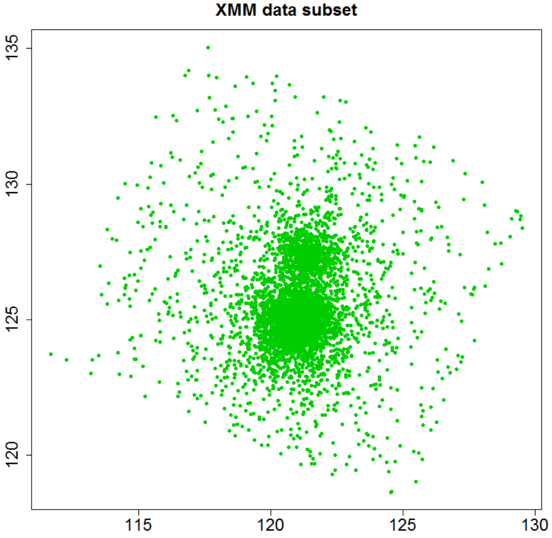
Strong Background Posterior Mean Positions: No Energy



Strong Background Posterior Mean Positions: Energy



XMM Data



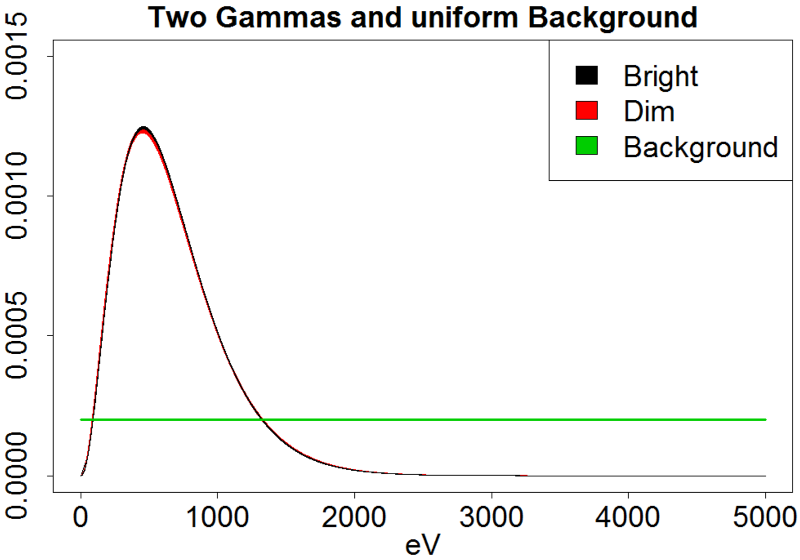
- Additional question: how do the spectral distributions of the sources compare?

Parameter Inference

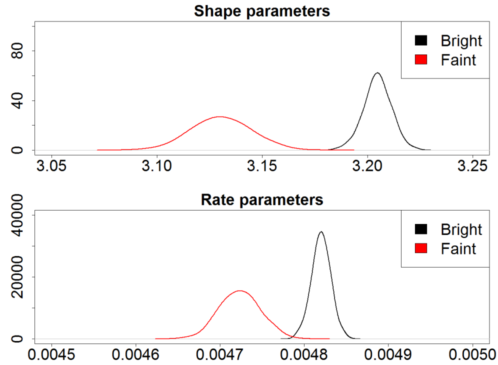
Table: Posterior parameter estimation for FK Aqr and FL Aqr (using spectral data)

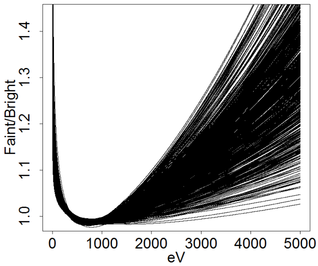
	μ_{11}	μ_{12}	μ_{21}	μ_{22}	w_1	w_2	w_b	α_1	α_2	β_1	β_2
Mean	120.973	124.873	121.396	127.326	0.732	0.189	0.079	3.195	3.121	0.005	0.005
SD	0.001	0.001	0.002	0.002	0.001	0.001	0.000	0.008	0.014	0.000	0.000
SD/Mean	0.000	0.000	0.000	0.000	0.001	0.003	0.004	0.002	0.004	0.002	0.005

Componentwise posterior spectral distributions

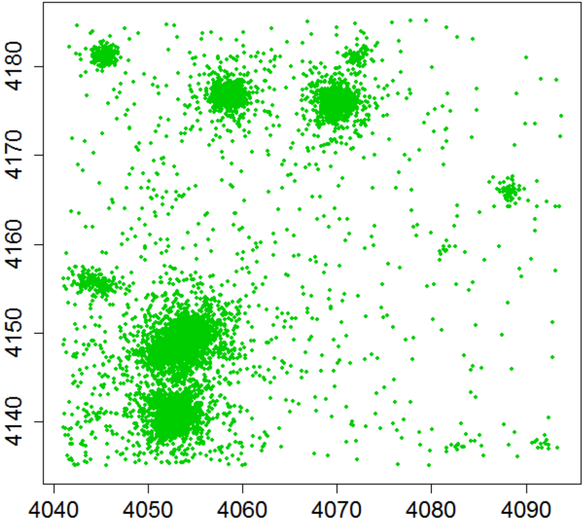


Posteriors of source spectral parameters

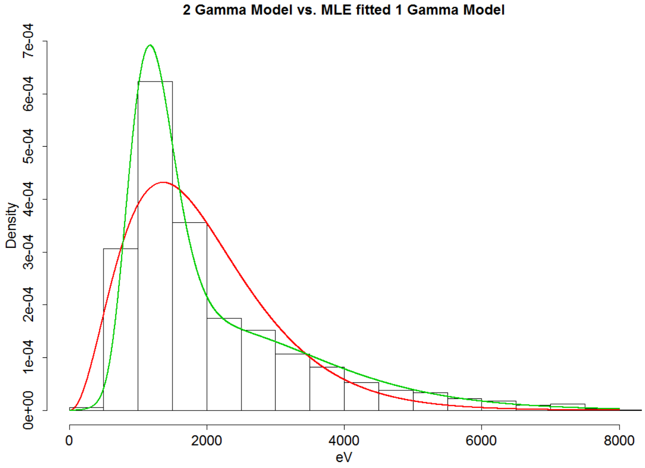




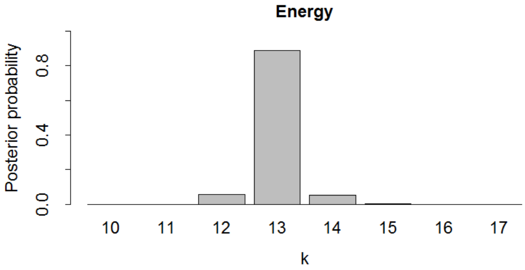
Chandra Data

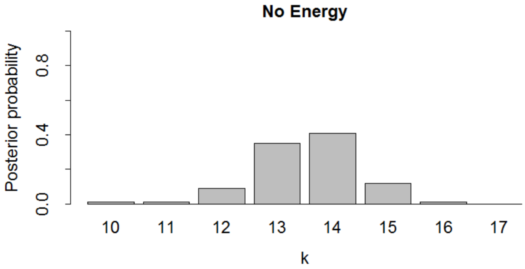


Gamma Mixture Spectral Model

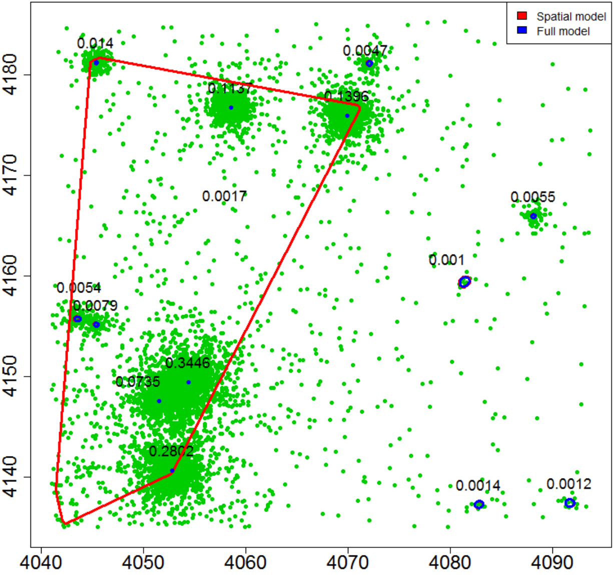


Chandra *k* Results





Locations: 90% credible regions



Chandra data spectral model contribution

1. Spectral model gives some constraints on the spectral distributions helping us to infer source properties more precisely
   - Posterior standard deviations are smaller with the spectral model
   - Without it some fainter sources are occasionally not found

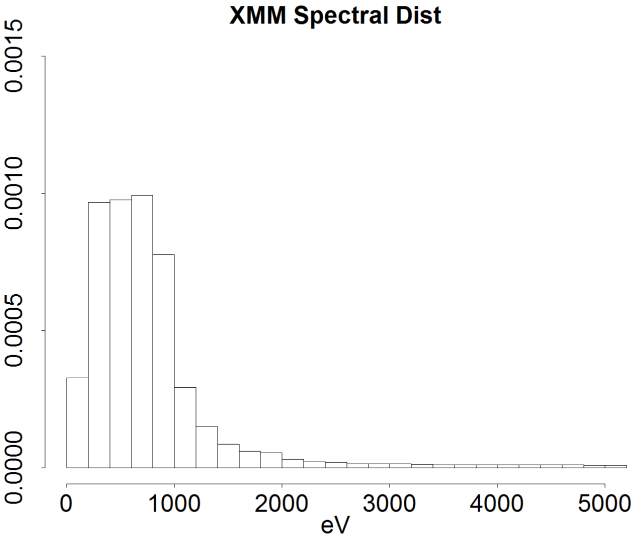
2. It also offers some robustness to chance or systematic variations in the PSF and background
   - Background is more easily distinguished from sources
   - Spectral model plays a large role in the likelihood so a strange shaped source will be unlikely to be split unless the spectral data also supports two sources

Summary and extensions

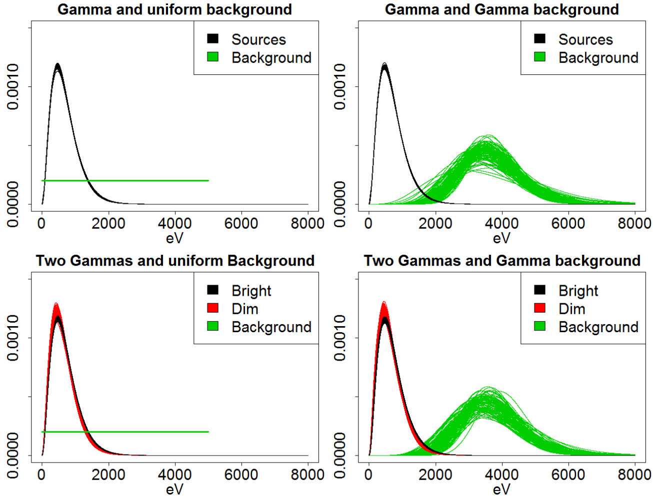
- Coherent method for dealing with overlapping sources that uses spectral as well as spatial information
- Flexibility to include other phenomenon
- Temporal model? Flares and other activity change the intensity and spectral distribution of sources over time
- Approximation to full method could be desirable in some cases

- S. Richardson, P. J. Green *On Bayesian analysis of mixtures with an unknown number of components* (with discussion), J. R. Statist. Soc. B, 59, 731792, 1997; corrigendum, 60 (1998), 661.
- I. King, *The structure of star clusters. I. An empirical density law*, The Astronomical Journal, 67 (1962), 471.
- C. M. Bishop, N. M. Nasrabadi, *Pattern recognition and machine learning*, Vol. 1. New York: springer, 2006.
- A. P. Dempster, N. M. Laird, D. B. Rubin. *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society, Series B (Methodological) (1977): 1-38.
- S. P. Brooks, A. Gelman, *General Methods for Monitoring Convergence of Iterative Simulations*, Journal of Computational and Graphical Statistics, Vol. 7, No. 4. (Dec., 1998), pp. 434-455.

XMM data spectral distribution



Four models


Gamma and uniform background
Sources
Background
0.0000
0.0010
0
2000
4000
6000
8000
eV
Gamma and Gamma background
Sources
Background
0.0000
0.0010
0
2000
4000
6000
8000
eV
Two Gammas and uniform Background
Bright
Dim
Background
0.0000
0.0010
0
2000
4000
6000
8000
eV
Two Gammas and Gamma background
Bright
Dim
Background
0.0000
0.0010
0
2000
4000
6000
8000
eV