# Detection: Overlapping Sources

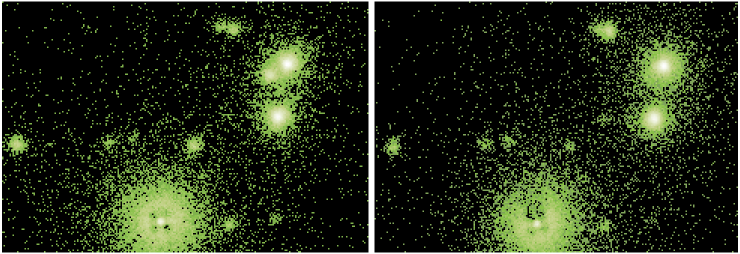
David Jones

Harvard University Statistics Department

November 12, 2013

# Introduction

- X-ray data: coordinates of photon detections, photon energy
- 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 of number of sources, (ii) use spectral information



# Basic Model and Notation

$y_{ij}$ = spatial coordinates of photon $j$ from source $i$
$k$ = # sources (components)
$\boldsymbol{\mu}_i$ = centre of source $i$
$n_i$ = # photons detected from source $i$

$$\begin{aligned}
y_{ij}|\boldsymbol{\mu}_i, n_i, k &\sim \text{PSF centred at } \boldsymbol{\mu}_i \quad j = 1, \ldots, n_i, i = 0, \ldots, k \\
(n_0, n_1, \ldots, n_k)|w, k &\sim \text{Mult}(n; (w_0, w_1, \ldots, w_k)) \\
(w_0, w_1, \ldots, w_k)|k &\sim \text{Dirichlet}(\lambda, \lambda, \ldots, \lambda) \\
\boldsymbol{\mu}_i|k &\sim \text{Uniform over the image} \quad i = 1, 2, \ldots, k \\
k &\sim \text{Pois}(\theta)
\end{aligned}$$

- Component with label 0 is background and its "PSF" is uniform over the image (so its "centre" is irrelevant)
- Reasonably insensitive to $\theta$, 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:

$$
\begin{aligned}
e_{ij}|\alpha_i, \beta_i &\sim \text{Gamma}(\alpha_i, \beta_i) \;\text{ for } i = 1, \ldots, k \text{ and } j = 1, \ldots, n_i \\
e_{0j} &\sim \text{Uniform to some maximum for } j = 1, \ldots, n_0 \\
\alpha_i &\sim \text{Gamma}(a_\alpha, b_\alpha) \\
\beta_i &\sim \text{Gamma}(a_\beta, b_\beta)
\end{aligned}
$$

Using a (correctly) "informative" prior on $\alpha_i$ and $\beta_i$ versus a diffuse prior made very little difference to results.

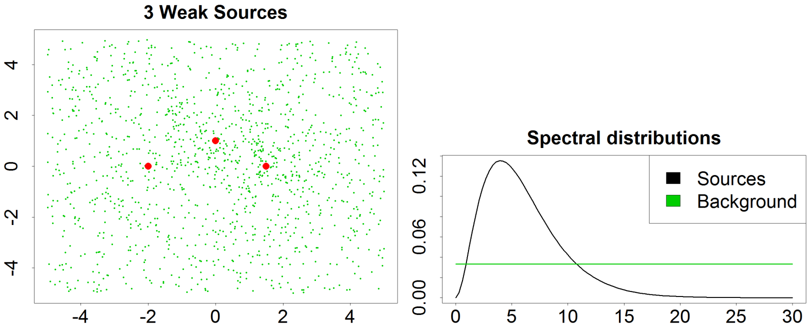
# Computation: RJMCMC

- Similar to Richardson & Green 1997
- Knowledge of the PSF makes things easier
- Insensitive to the prior on $k$ e.g. posterior when $k = 10$ and $\theta = 3$:

| | Posterior of number of sources (k) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| Mean | 0.029 | 0.058 | 0.141 | 0.222 | 0.220 | 0.157 | 0.082 |
| SD | 0.018 | 0.019 | 0.022 | 0.029 | 0.027 | 0.021 | 0.014 |

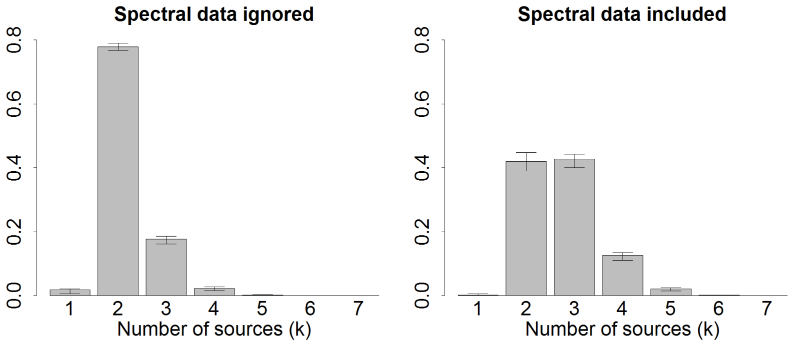
Used posterior probabilities given by 10 chains

# Example




- Region occupied by the three sources (2 SD) is about 28% of the area and contains about 41% of the observations
- Within this sources region around 48% is background
- Positions $(-2, 0), (0, 1), (1.5, 0)$ with intensities $50, 100, 150$ respectively

# Posterior of $k$




- Mean over 10 chains of the posterior probabilities (range indicated)
- When the spectral data is ignored we do not find the faintest source

# Parameter Inference

| | $\mu_{11}$ | $\mu_{12}$ | $\mu_{21}$ | $\mu_{22}$ | $\mu_{31}$ | $\mu_{32}$ | $w_1$ | $w_2$ | $w_3$ | $w_b$ | $\alpha$ | $\beta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Truth** | -2 | 0 | 0 | 1 | 1.5 | 0 | 0.038 | 0.077 | 0.115 | 0.769 | 3 | 0.5 |
| **Spectral data ignored** | | | | | | | | | | | | |
| Mean | -1.266 | 0.839 | 0.401 | 0.549 | 1.798 | -0.054 | 0.049 | 0.067 | 0.086 | 0.798 | NA | NA |
| SD | 0.069 | 0.125 | 0.067 | 0.068 | 0.030 | 0.046 | 0.002 | 0.002 | 0.003 | 0.001 | NA | NA |
| MSE | **0.543** | **0.718** | 0.165 | 0.207 | 0.090 | 0.005 | | | | | NA | NA |
| SD/Mean | | | | | | | 0.050 | 0.027 | 0.032 | 0.001 | NA | NA |
| **Spectral data included** | | | | | | | | | | | | |
| Mean | -1.790 | -0.101 | -0.234 | 1.042 | 1.584 | -0.044 | 0.040 | 0.077 | 0.115 | 0.768 | 2.827 | 0.459 |
| SD | 0.037 | 0.064 | 0.033 | 0.026 | 0.019 | 0.022 | 0.001 | 0.001 | 0.002 | 0.000 | 0.013 | 0.003 |
| MSE | **0.045** | **0.014** | 0.056 | 0.002 | 0.007 | 0.002 | | | | | 0.030 | 0.002 |
| SD/Mean | | | | | | | 0.036 | 0.018 | 0.014 | 0.000 | 0.004 | 0.006 |

- The effects are less pronounced when the sources are more easily distinguished from the background

# Allocation of Photons

Table: Allocation breakdown: (a) ignoring spectral data

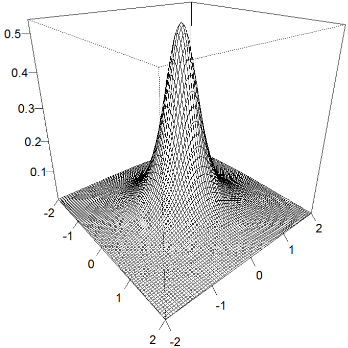
| **Source (intensity)** | **No. Photons** | **Allocation Breakdown** | | | |
|---|---|---|---|---|---|
| | | **Background** | **Left** | **Middle** | **Right** |
| Background (10/sq) | 1015 | 0.876 | 0.035 | 0.040 | 0.049 |
| Left (50) | 38 | 0.798 | 0.121 | 0.067 | 0.014 |
| Middle (100) | 97 | 0.502 | 0.168 | 0.189 | 0.141 |
| Right (150) | 152 | 0.481 | 0.043 | 0.159 | 0.317 |

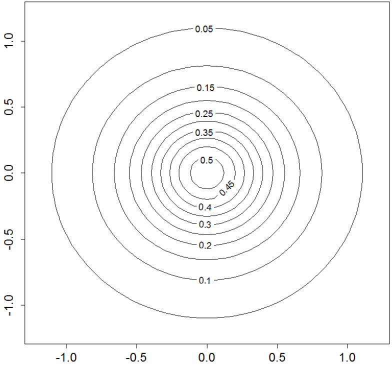
Table: Allocation breakdown: (b) using spectral data

| **Source (intensity)** | **No. Photons** | **Allocation Breakdown** | | | |
|---|---|---|---|---|---|
| | | **Background** | **Left** | **Middle** | **Right** |
| Background (10/sq) | 1015 | 0.894 | 0.024 | 0.038 | 0.045 |
| Left (50) | 38 | 0.531 | 0.278 | 0.165 | 0.026 |
| Middle (100) | 97 | 0.293 | 0.122 | 0.346 | 0.239 |
| Right (150) | 152 | 0.305 | 0.028 | 0.141 | 0.526 |

- Background is more easily distinguished from the sources when we include the spectral data

# 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:

$$n_1 \sim \text{Pois}(1000)$$

- Dim source:

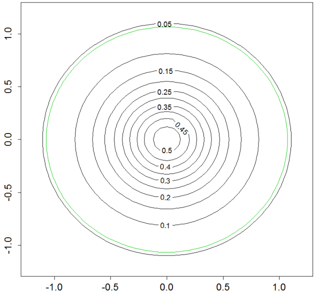
$$n_2 \sim \text{Pois}(1000/r)$$

where $r = 1, 2, 10, 50$ gives the relative intensity

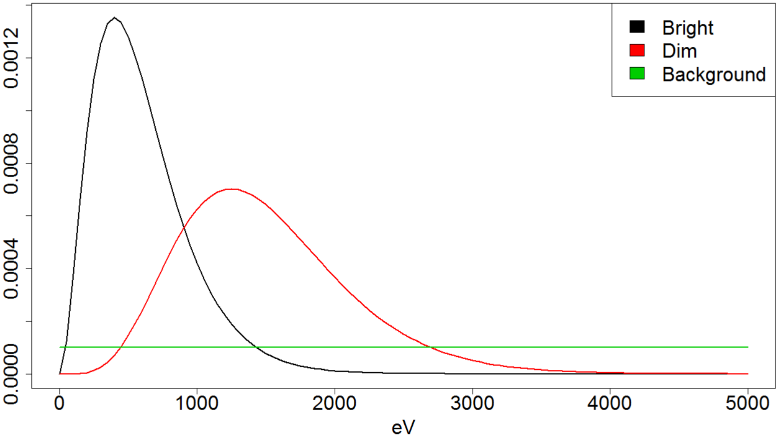
- Background per 'source region':

$$n_0 \sim \text{Pois}(bd1000/r)$$

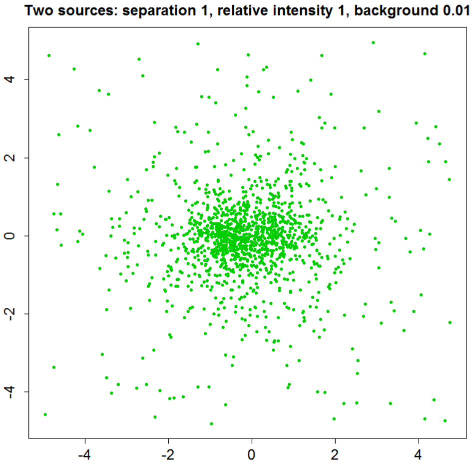
where relative background $b = 0.001, 0.01, 0.1, 1$. Here $d = 0.52$ is the proportion of photons from a source within the region defined by density greater than 10% of the maximum (essentially a circle with radius 1)



# Simulation Study: Data Generation

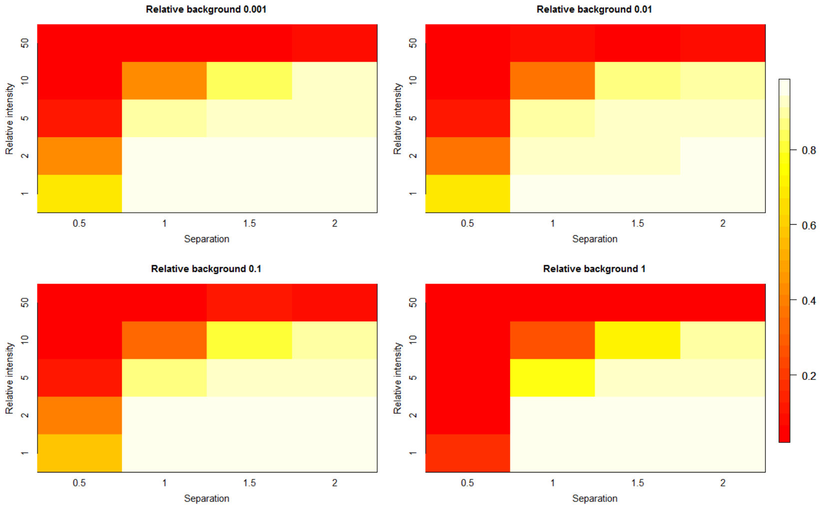


# Simulation Study: Example

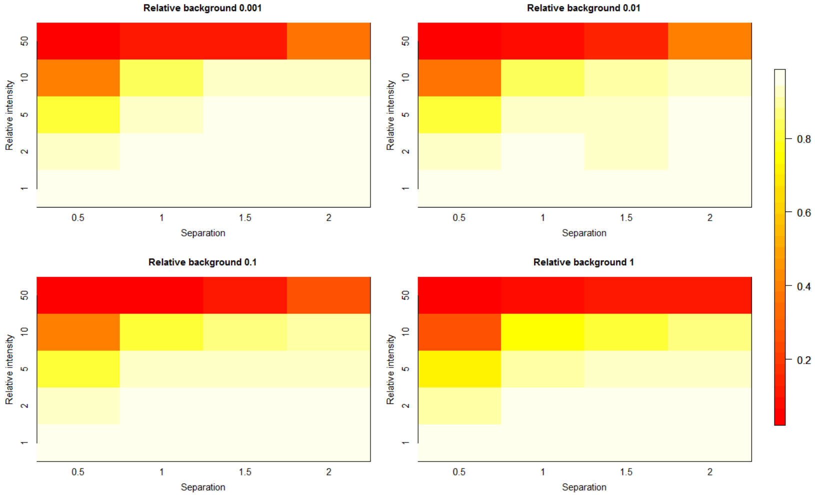



- 50 datasets simulated for each configuration
- Analysis with and without energy data
- Summarize posterior of $k$ by posterior probability of two sources

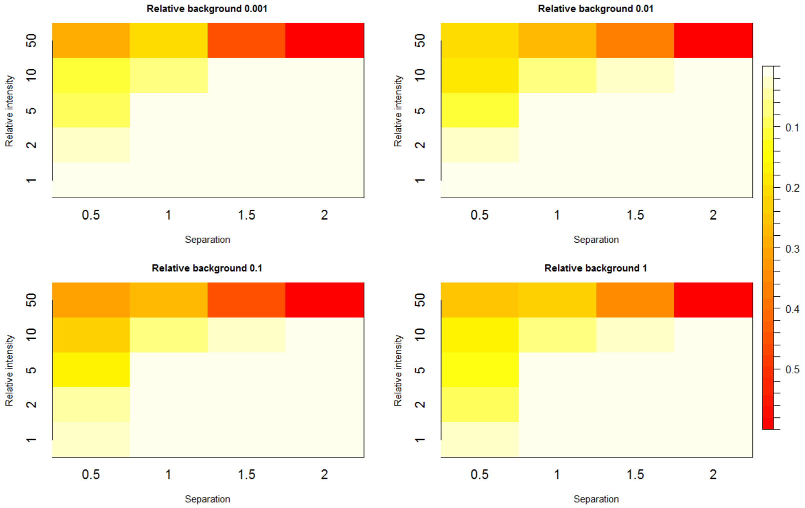
# Posterior Probability at k=2: No Energy



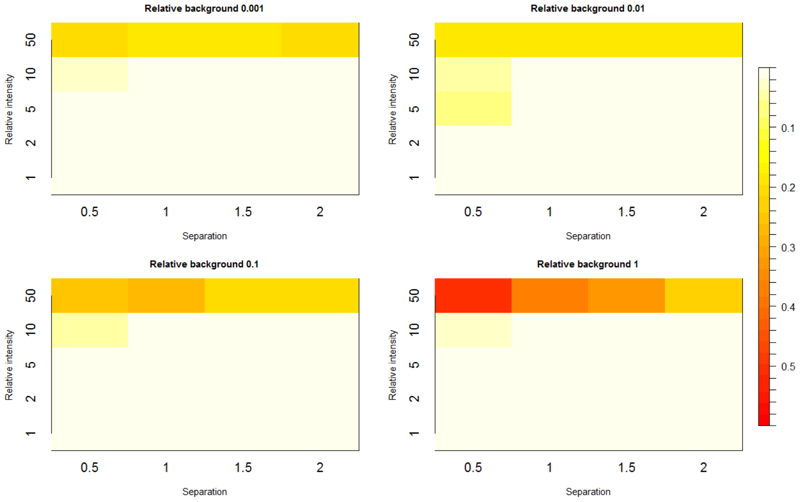
# Posterior Probability at k=2: Energy



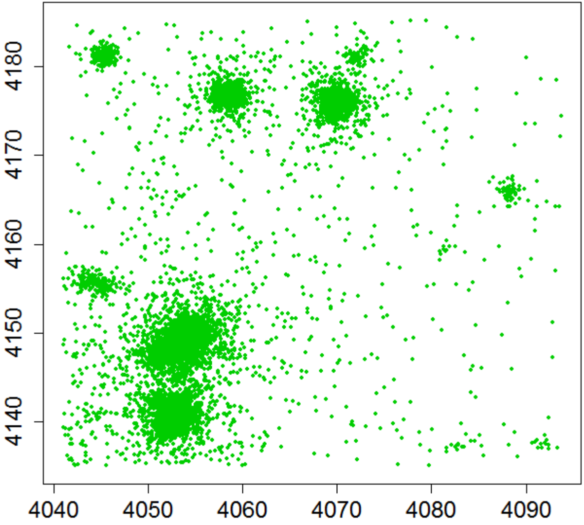
# Average MSE of Positions: No Energy



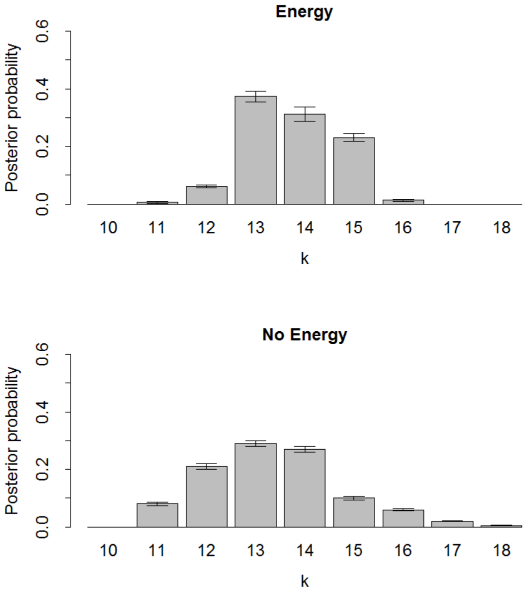
# Average MSE of Positions: Energy



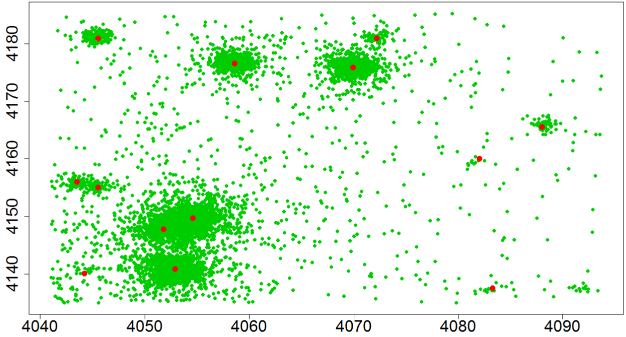
# Chandra Data



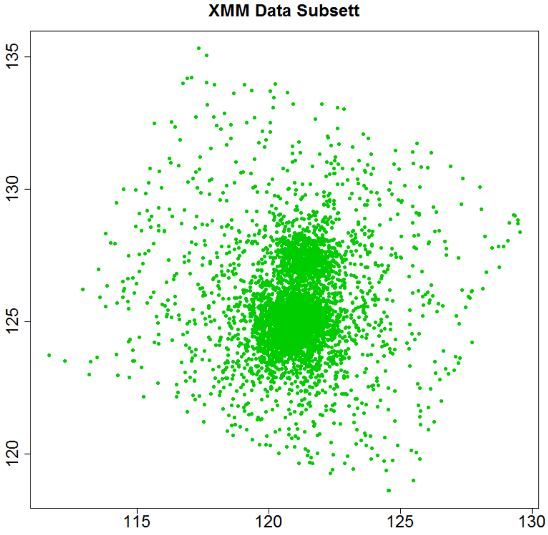
# Chandra $k$ Results



# Locations

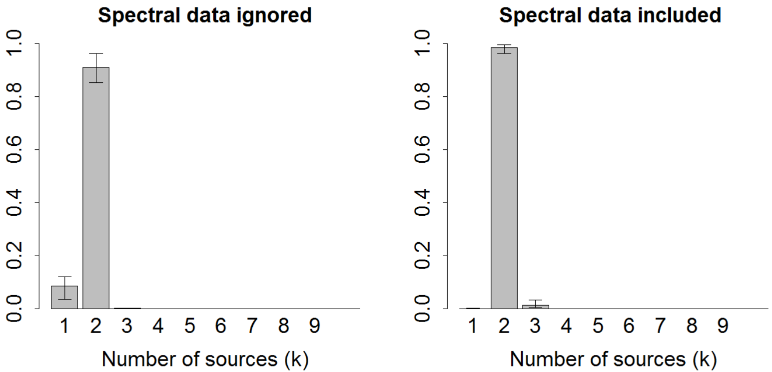


# XMM Data




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

# k posterior



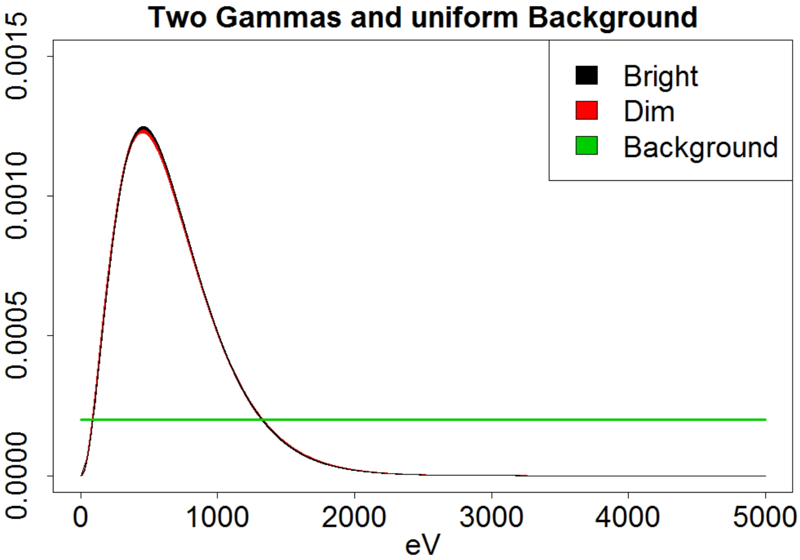

- Mean over 10 chains of the posterior probabilities (range indicated)
- Spectral information focuses posterior on 2 sources

# Parameter Inference

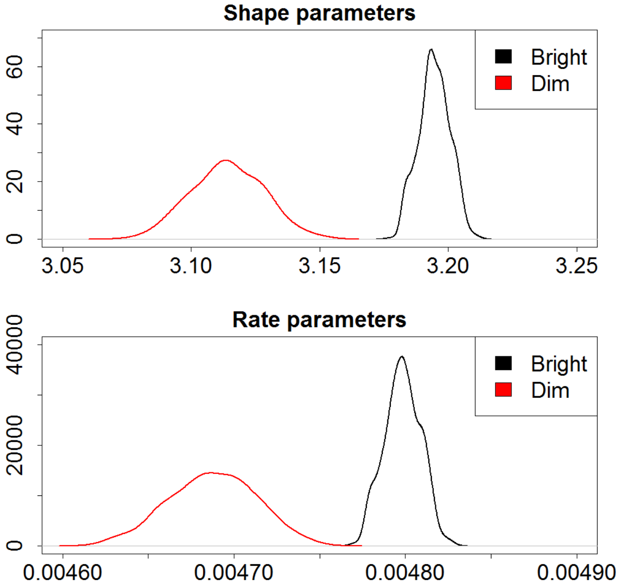
Table: Parameter estimation for FK Aqr and FL Aqr (using spectral data)

| | $\mu_{11}$ | $\mu_{12}$ | $\mu_{21}$ | $\mu_{22}$ | $w_1$ | $w_2$ | $w_b$ | $\alpha$ | $\beta$ |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 120.988 | 124.891 | 121.366 | 127.376 | 0.808 | 0.182 | 0.009 | 3.182 | 0.005 |
| SD | 0.001 | 0.002 | 0.016 | 0.027 | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 |
| SD/Mean | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.005 | 0.011 | 0.000 | 0.000 |

# Componentwise posterior spectral distributions



# Posteriors of source spectral parameters

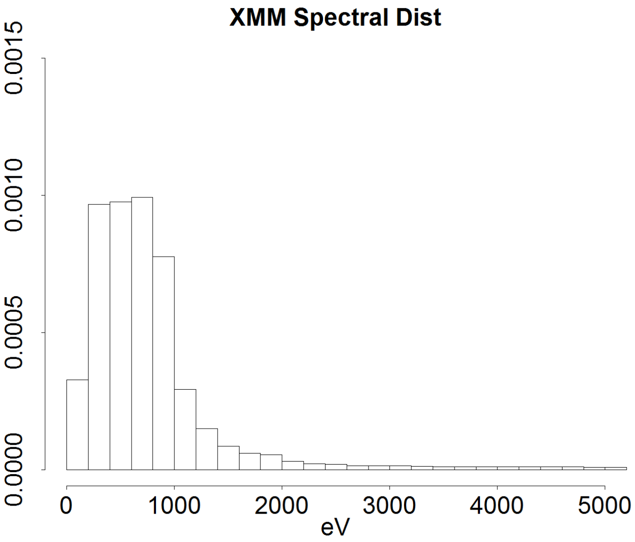


# Summary

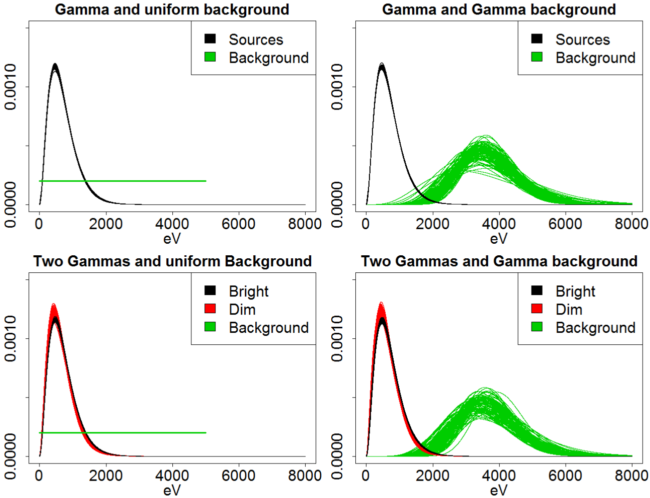
- Coherent method for dealing with overlapping sources that uses spectral as well as spatial information
- Flexibility to include other phenomenon
- How to combine Chandra datsets?
- Other models/computation possible
- Approximation to full method could be desirable

- 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