Disentangling Overlapping Astronomical Sources using Spatial, Spectral, and Temporal Information

David Jones

SAMSI and Duke University

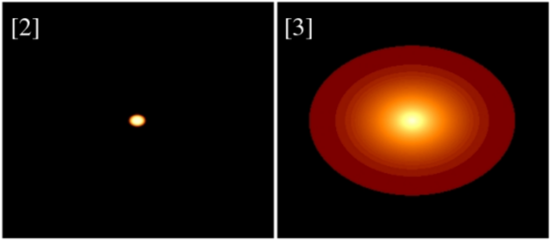
Collaborators: Vinay Kashyap (CfA), David van Dyk (Imperial College Statistics), Luis Campos (Harvard)

International CHASC Astrostatistics Center

SAMSI ASTRO Workshop

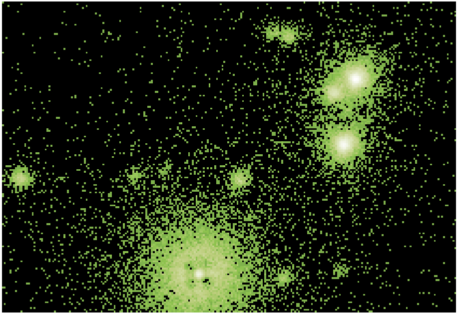
August 21, 2016

- Chandra and XMM-Newton X-ray telescope data:
  - spatial coordinates of photon detections
  - photon energy (PI channel)
- Telescope response: recorded photon positions are spread out according to the point spread function (PSF)

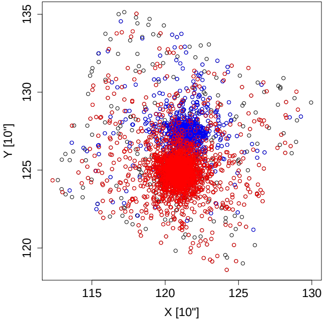


Source: XMM-Newton release notes

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



Chandra observation of the Orion Nebula Cluster



Example photon assignment for XMM observation of FK and FL Aqr

Spatial Data: Bayesian Data Generating Model

Spatial Data: Bayesian Data Generating Model

Observed quantities:

N = total # photons (fix)

(x_i, y_i) = spatial coordinates of photon i

Spatial Data: Bayesian Data Generating Model

Observed quantities:

N = total # photons (fix)

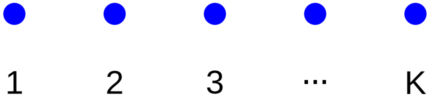
(x_i, y_i) = spatial coordinates of photon i

Model:

Spatial Data: Bayesian Data Generating Model

Observed quantities:
N = total # photons (fix)
(x_i, y_i) = spatial coordinates of photon i

Model:
 # sources: $K \sim \text{Pois}(\kappa)$



Spatial Data: Bayesian Data Generating Model

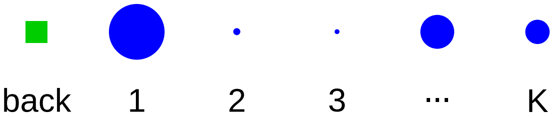
Observed quantities:

N = total # photons (fix)

(x_i, y_i) = spatial coordinates of photon i

Model:

True relative brightness: $(w_0, \ldots, w_K)|K \sim$ Dirichlet$(1, \ldots, 1)$



Spatial Data: Bayesian Data Generating Model

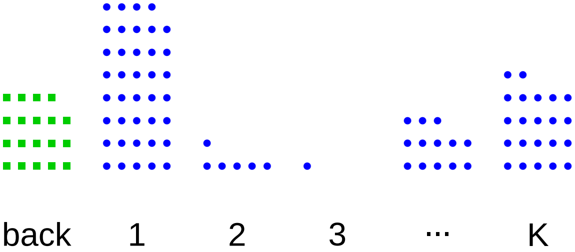
Observed quantities:

N = total # photons (fix)

(x_i, y_i) = spatial coordinates of photon i

Model:

Counts: $(n_0, \ldots, n_K) | w, N \sim$ Multinomial$(N; (w_0, \ldots, w_K))$



Spatial Data: Bayesian Data Generating Model

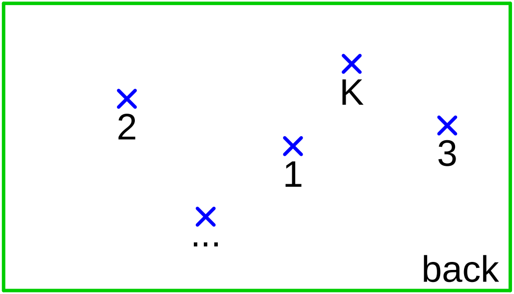
Observed quantities:

N = total # photons (fix)

(x_i, y_i) = spatial coordinates of photon i

Model:

Positions: $\boldsymbol{\mu}_j | K \sim$ Uniform over image, for $j = 1, 2, \ldots, K$



Spatial Data: Bayesian Data Generating Model

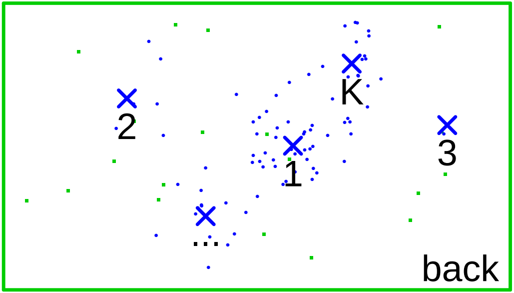
Observed quantities:

N = total # photons (fix)

(x_i, y_i) = spatial coordinates of photon i

Model:

Photon coords:	(x_i, y_i)|source $j, \boldsymbol{\mu}_j$	$\sim$	PSF$_j$ centred at $\boldsymbol{\mu}_j$
Background:	(x_i, y_i)|background	$\sim$	Uniform over the image



Spectral Data Model

Additionally observed:

e_i = PI channel of photon i

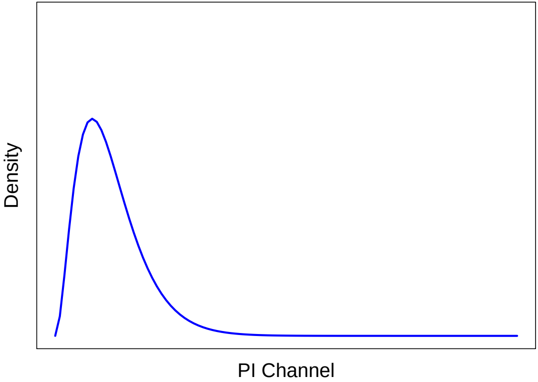
Spectral Data Model

Additionally observed:

e_i = PI channel of photon i

Model:

Photon energy: $e_i |$source $j, \alpha_j, \beta_j \sim$ Gamma(α_j, β_j)



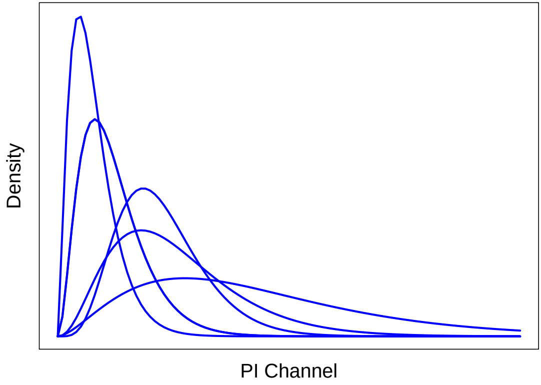
Spectral Data Model

Additionally observed:

e_i = PI channel of photon i

Model:

Photon energy: $e_i | \text{source } j, \alpha_j, \beta_j \sim \text{Gamma}(\alpha_j, \beta_j) \quad \text{for } j = 1, 2, \ldots, K$



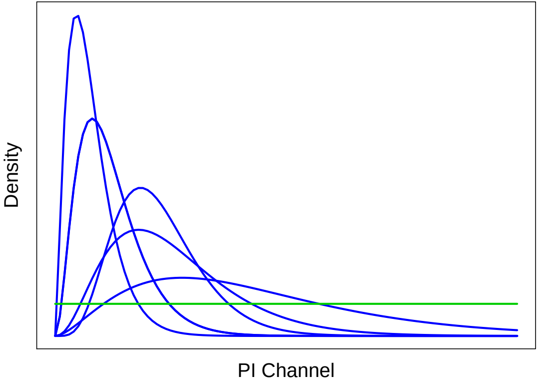
Additionally observed:

e_i = PI channel of photon i

Model:

Photon energy: $e_i|\text{source } j, \alpha_j, \beta_j \sim \text{Gamma}(\alpha_j, \beta_j)$ for $j = 1, 2, \ldots, K$

Background: $e_i|\text{background} \sim \text{Uniform}(0, E_{\max})$

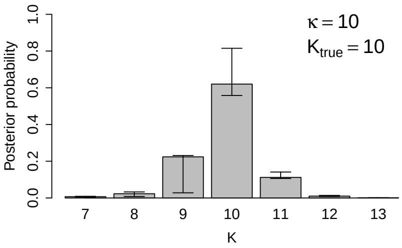


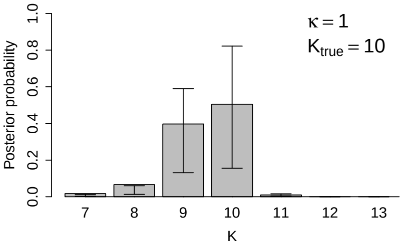
Computation: RJMCMC

- **Final output:** joint posterior distribution of all the parameters
- **How we get there:** RJMCMC that combines algorithms of Richardson & Green 1997 and Wiper et al. 2001

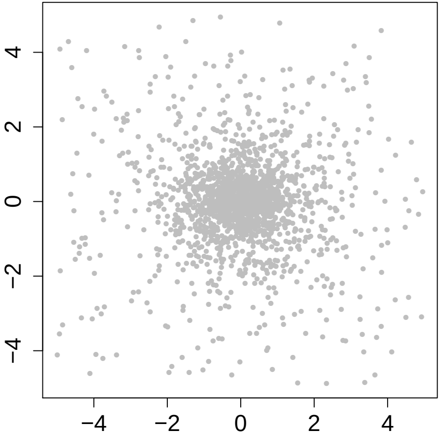
Is this tractable?

- Knowledge of the PSF makes things much easier
- Inference is insensitive to the prior $K \sim \text{Pois}(\kappa)$ e.g. for 20 simulated datasets each with 10 sources we have the following mean posterior distributions . . .



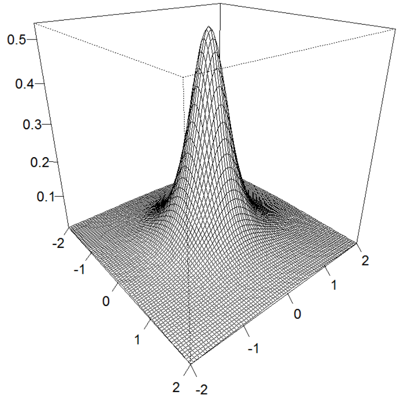


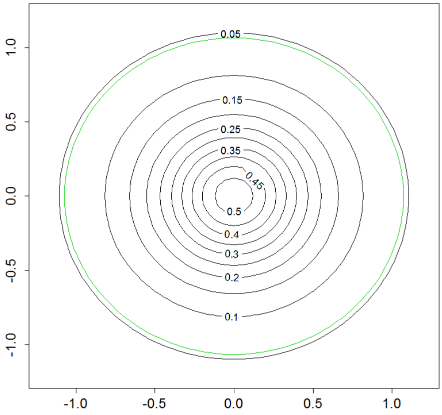
Simulation Study: Example



- 100 datasets simulated for each configuration
- Analysis with spatial-only model and full model

Simulation Study: PSF (King 1962)

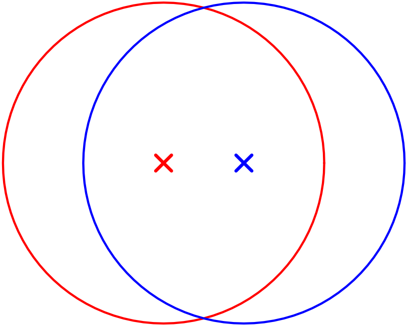




- King density has Cauchy tails
- Gaussian PSF leads to over-fitting in real data
- 'Source region': the region defined by PSF density greater than 10% of the maximum (essentially a circle with radius 1)

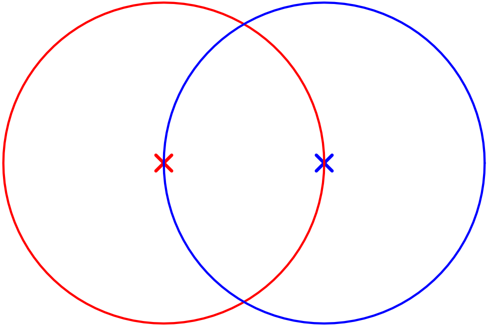
Simulation Study: Spatial Data

Source separation: $d = 0.5, 1, 1.5, 2$



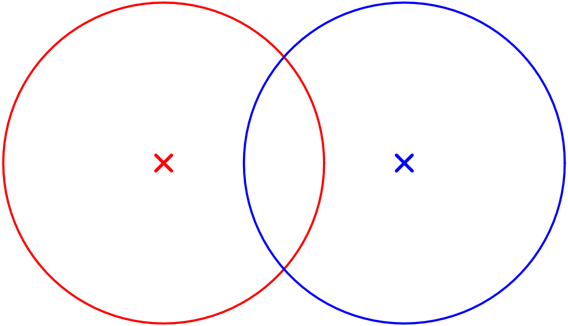
Simulation Study: Spatial Data

Source separation: $d = 0.5, 1, 1.5, 2$

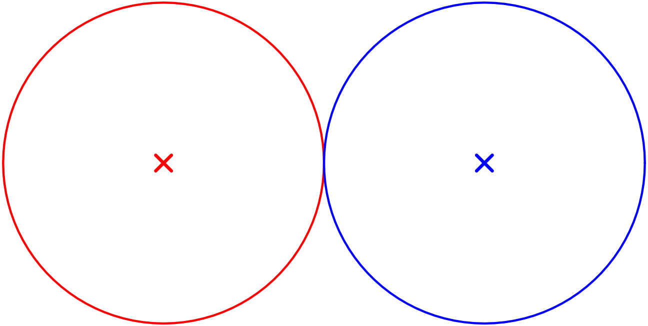


Simulation Study: Spatial Data

Source separation: $d = 0.5$, 1, 1.5, 2



Source separation: $d = 0.5$, 1, 1.5, 2



Simulation Study: Spatial Data

Relative intensity:

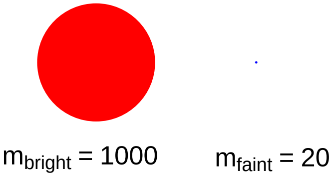
- Bright source:

$$n_1 \sim \text{Pois}(m_{\text{bright}} = 1000)$$

- Faint source:

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

where $r = 50, 10, 5, 2, 1$



Relative intensity:

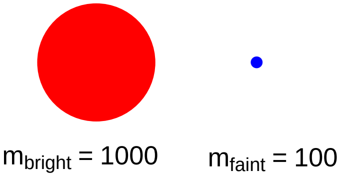
- Bright source:

$$n_1 \sim \text{Pois}(m_{\text{bright}} = 1000)$$

- Faint source:

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

where $r = 50$, 10, 5, 2, 1



Simulation Study: Spatial Data

Relative intensity:

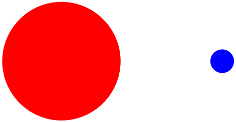
- Bright source:

$$n_1 \sim \text{Pois}(m_{\text{bright}} = 1000)$$

- Faint source:

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

where $r = 50, 10, 5, 2, 1$



$m_{\text{bright}} = 1000$ $m_{\text{faint}} = 200$

Simulation Study: Spatial Data

Relative intensity:

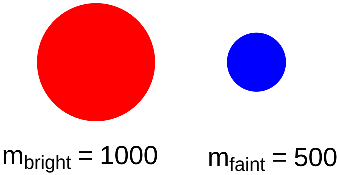
- Bright source:

$$n_1 \sim \text{Pois}(m_{\text{bright}} = 1000)$$

- Faint source:

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

where $r = 50, 10, 5, 2, 1$



Relative intensity:

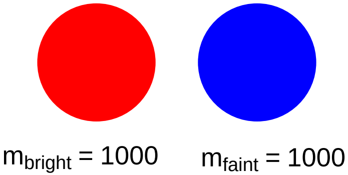
- Bright source:

$$n_1 \sim \text{Pois}(m_{\text{bright}} = 1000)$$

- Faint source:

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

where $r = 50, 10, 5, 2, 1$

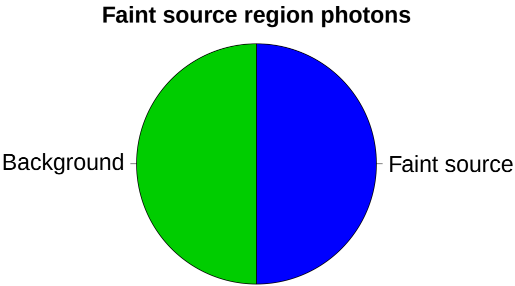


Simulation Study: Spatial Data

Relative background:

$$n_0 \sim \text{Pois}\left(b \times \text{avg \# faint source photons in faint source region} \times \frac{\text{image area}}{\text{source region area}} \right)$$

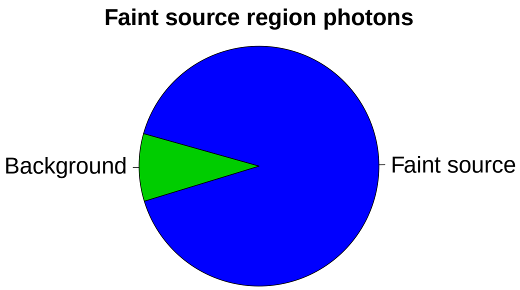
$b = 1, 0.1, 0.01, 0.001$



Relative background:

$$n_0 \sim \text{Pois}\left(b \times \text{avg \# faint source photons in faint source region} \times \frac{\text{image area}}{\text{source region area}}\right)$$

$b = 1, 0.1, 0.01, 0.001$

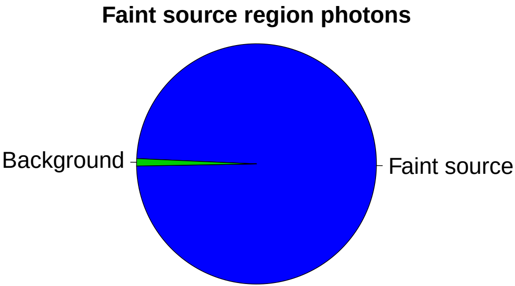


Simulation Study: Spatial Data

Relative background:

$$n_0 \sim \text{Pois}\left(b \times \text{avg \# faint source photons in faint source region} \times \frac{\text{image area}}{\text{source region area}}\right)$$

$b = 1, 0.1, 0.01, 0.001$

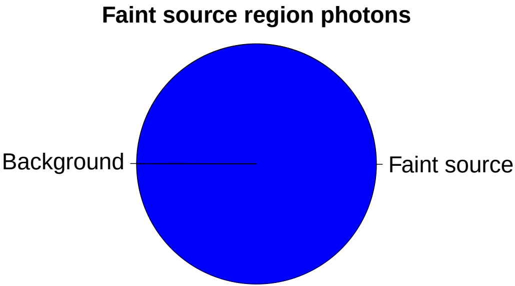


Simulation Study: Spatial Data

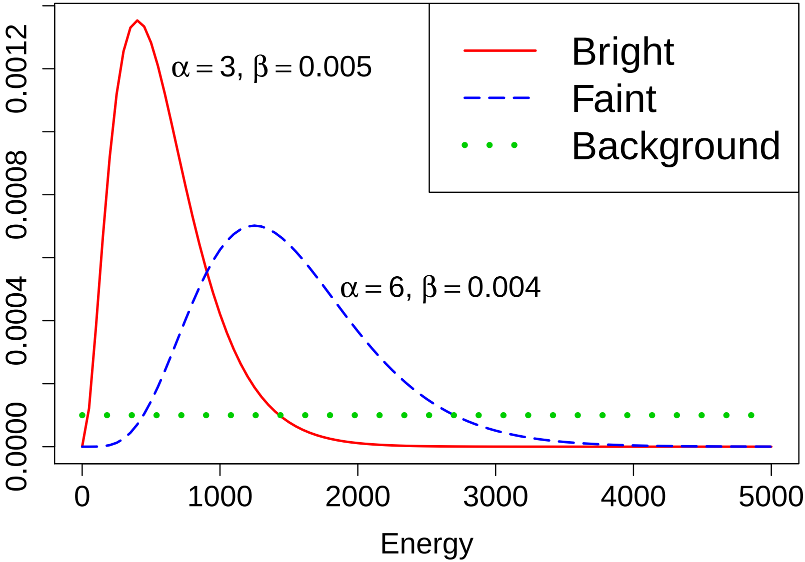
Relative background:

$$n_0 \sim \text{Pois}\left(b \times \text{avg \# faint source photons in faint source region} \times \frac{\text{image area}}{\text{source region area}} \right)$$

$b = 1, 0.1, 0.01, 0.001$

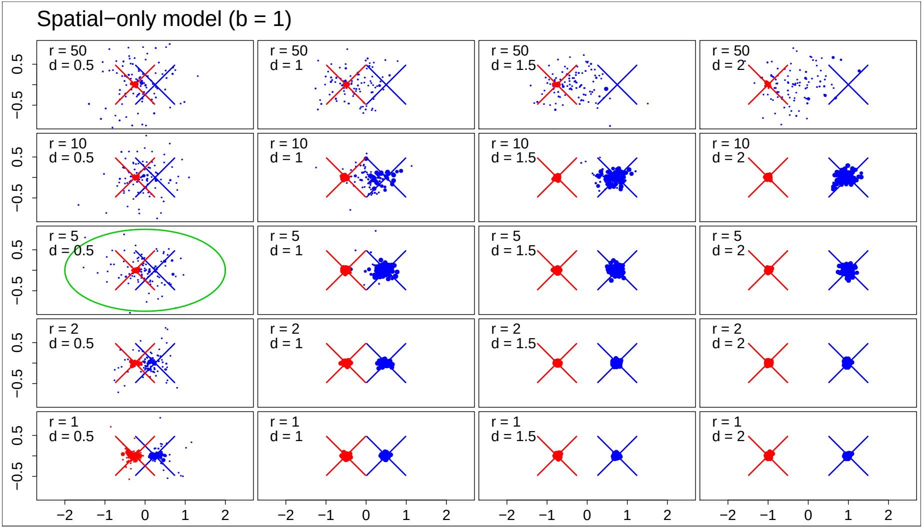


Simulation Study: Spectral Data



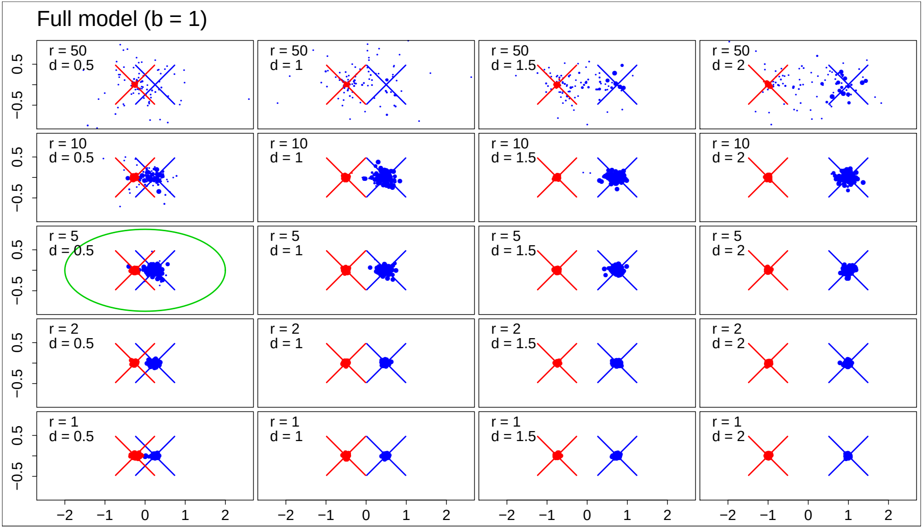
Mean Posterior Positions (Strong Background, b=1)

- Red = bright sources, blue = faint source
- d = separation, r = relative intensity
- Size of dots $\propto$ posterior probability of two sources



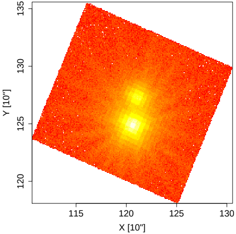
Mean Posterior Positions (Strong Background, b=1)

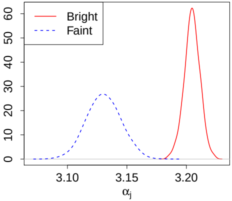
- Red = bright sources, blue = faint source
- d = separation, r = relative intensity
- Size of dots $\propto$ posterior probability of two sources

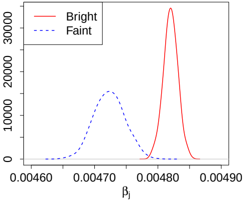


Paper gives two data analyses:

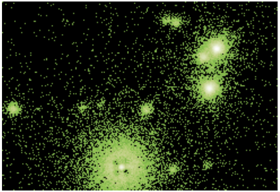
- **Briefly:** XMM data – binary source, FK and FL Aqr





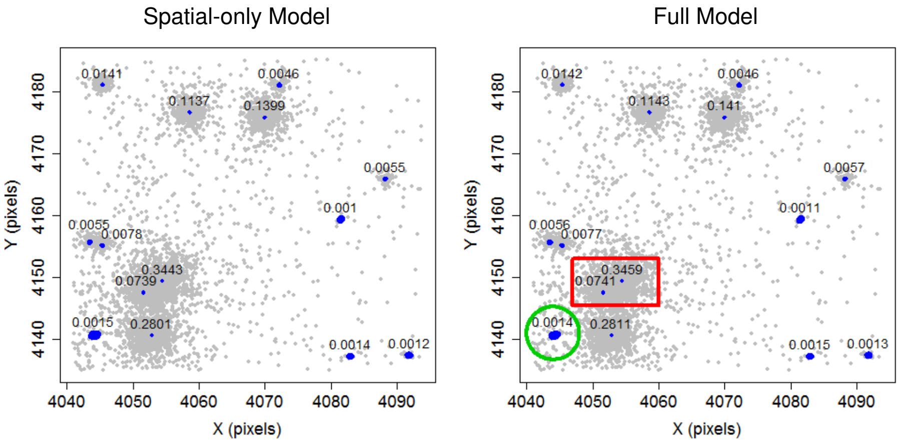


- **Focus:** Chandra image

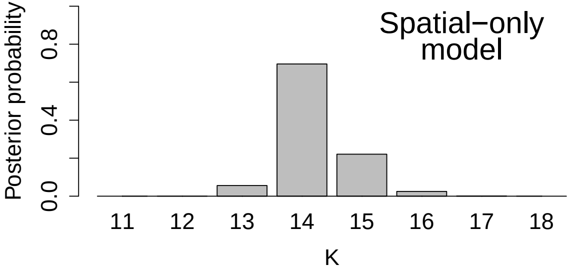


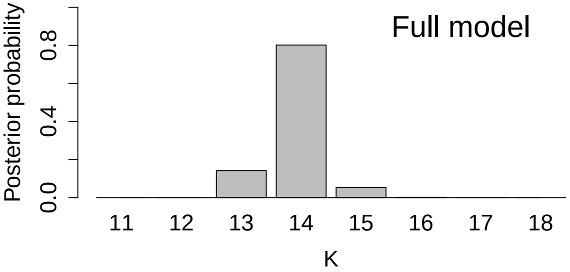
Chandra observation of the Orion Nebula Cluster

Chandra data



- Part of Chandra observation of the Orion Nebula Cluster (distorted source cut out)
- Approximately 25” $\times$ 25” in size


Spatial−only
model
Posterior probability
0.0
0.4
0.8
11
12
13
14
15
16
17
18
K

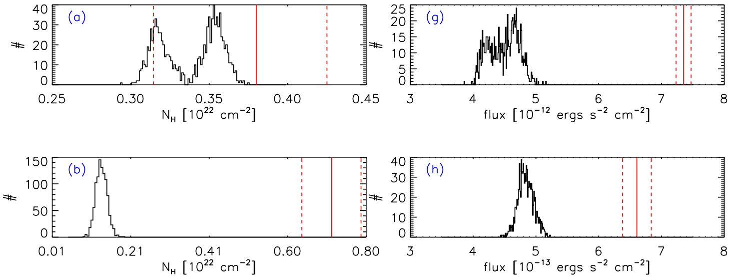

Full model
Posterior probability
0.0
0.4
0.8
11
12
13
14
15
16
17
18
K

Follow-up Spectral Analysis using CIAO/Sherpa v4.6

- Each iteration of our algorithm probabilistically assigns every photon to a source or the background
- Our assignments can be used to repeatedly perform more detailed spectral analysis
- The resulting histogram of spectral parameter fits enable us to quantify uncertainty

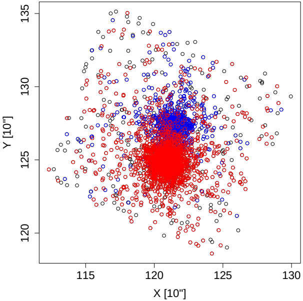
Bright source:

Faint source:

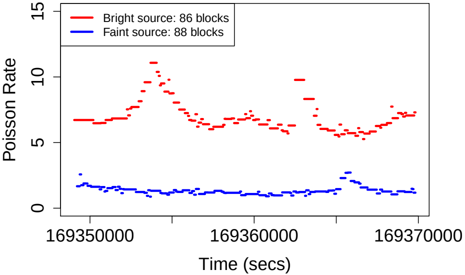


Temporal Extension

- **Concept:** for variable sources, modeling temporal data should further help separation
- Need a simple but flexible lightcurve model
- One idea: Poisson process with piecewise constant rate . . . as in Bayesian Blocks (Scargle 1998, Scargle et al. 2013):



Example photon assignment for XMM observation of FK and FL Aqr

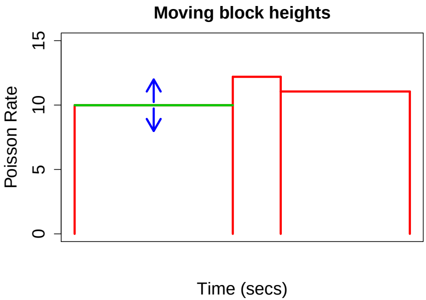


Bayesian blocks fits of the corresponding lightcurves

How to implement the MCMC?

To allocate photons, we need to take account of our **uncertainties** about the underlying lightcurves. Computationally:

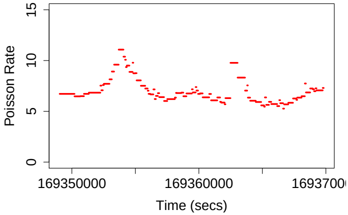
- MCMC iterations must update lightcurve models
- How to propose "nearby" models? Starting approach:
  1. Run Bayesian blocks on all data and then fix the breakpoints
  2. Set priors on the block heights and then update the heights in each MCMC iteration

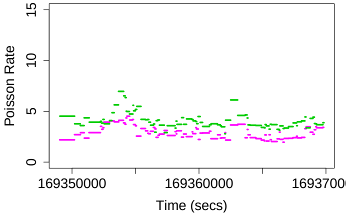


- More in the spirit of Bayesian blocks, we could also allow the breakpoints to move left or right

How to implement the RJMCMC?

- RJMCMC will add further challenges
- When we “split” an existing source we will need to split its lightcurve model into two (stochastically)





Extensions

Scalability:

- Divide up image into sub-images e.g. Safarzadeh et al. (2014)
- Sample sub-images multiple times and combine posterior estimates . . . or the posterior distributions themselves e.g. Minsker et al. (2014)

Additional improvements / directions:

- Instrument effects e.g. varying PSF
- Separation of extended sources and point sources
- Binning and LSST data

Thanks!