Disentangling Overlapping Point Sources

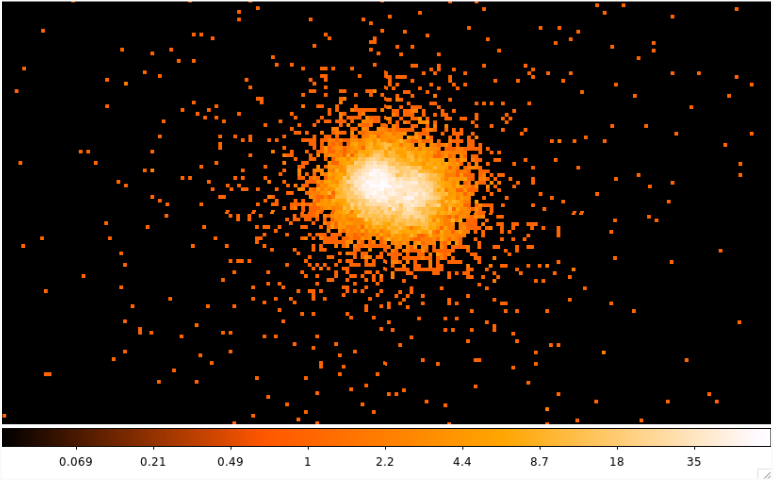
Using Spatial, Spectral, and Temporal Information

Luis F. Campos[1]

with David Jones,[2] Aneta Siemiginowska,[3] Vinay Kashyap,[3]
Xiao-Li Meng,[1] David van Dyk[4]

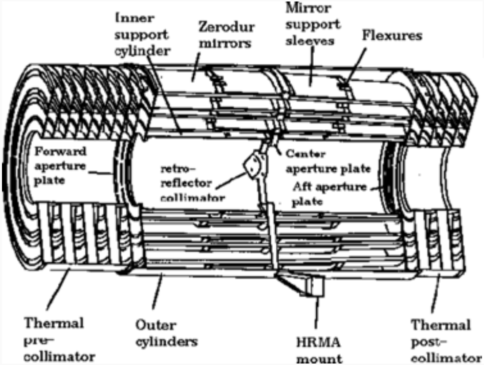
[1] *Harvard University,*
[2] *Texas A&M University,*
[3] *Center for Astrophysics, Harvard and Smithsonian*
[4] *Imperial College London*

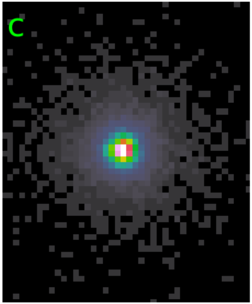
What are overlapping sources?



Why do we see this?

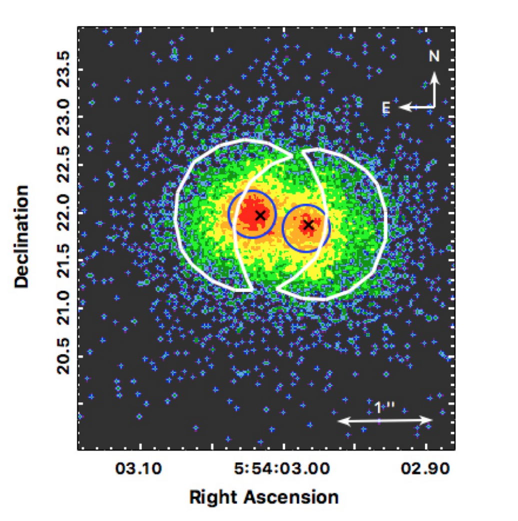
- Close proximity and instrument effects:





Chandra X-ray observatory and point spread function (psf)

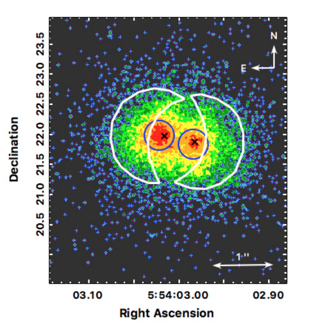
How are overlapping sources typically analyzed?



- cores/wings are defined spatially for each source
- separate events into sources
- continue analysis separately

Principe, et.al. (2017) 'The multiple young stellar objects of HBC 515: An X-ray and millimeter-wave imaging study in (pre-main sequence) diversity'

How are overlapping sources typically analyzed?



This leaves much to be desired

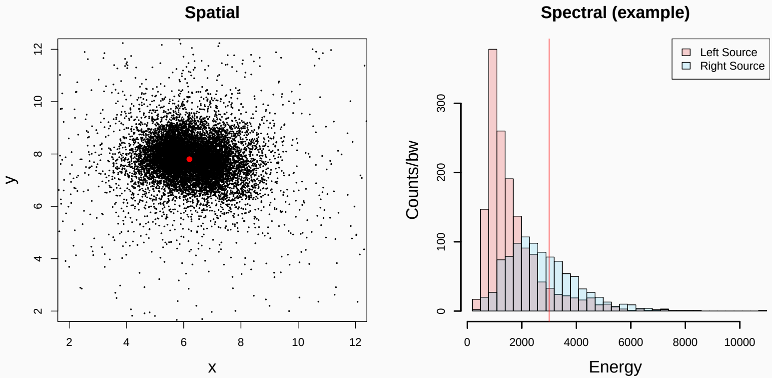
- discards lots of data
- overestimates our certainty

Certainty?

- Sources: 1.2 arcsec apart
- Core regions: 0.5 arcsec radius
- Left source: 1.6x brighter
- About about 13% events in dimmer core will be misclassified

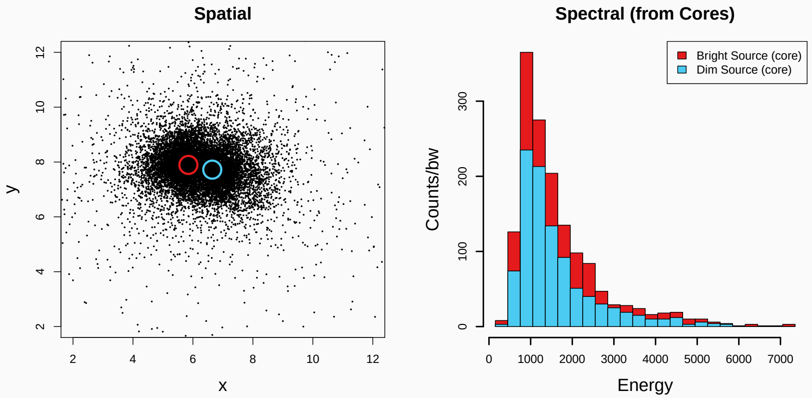
Using models to capture uncertainty

Jones, Kashyap, van Dyk (2015) "Disentangling Overlapping Astronomical Sources using Spatial and Spectral Information." ApJ



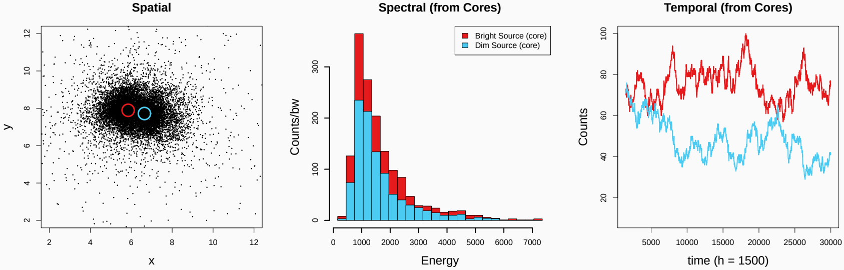
- Define $p(source|location, energy)$
- Based on spatial, equally likely L or R – energy?
- **Key**: you need *differing energy distributions*.

A complication for HBC515A



- HBC515Aab: the energy distributions don't seem to differ.

Extending the model to supplement



- Would it benefit to use temporal information (t_i) to supplement?

Setting A Joint Data Model

Consider the following factorization of likelihood for model parameters $\mathbf{\Theta}$

$$
\begin{aligned}
p(x_i, y_i, E_i, t_i \mid z_i = j, \mathbf{\Theta}) = \; & p(x_i, y_i \mid E_i, t_i, z_i = j, \mathbf{\Theta}) \\
& p(t_i \mid z_i = j, \mathbf{\Theta}) \\
& p(E_i \mid t_i, z_i = j, \mathbf{\Theta})
\end{aligned}
$$

Some modeling assumptions/decisions:

1. Spatial model: $(x_i, y_i | z_i = s) \sim f_{\mu_j}$ point-spread function
2. Time model: Multinomial-Dirichlet model with fixed time-breaks
3. Energy model: Gamma distributions that vary across source **and** time.

Parameters and model fitting

Parameters:

- Source Intensities: $\boldsymbol{\pi} = (\pi_0, \pi_1, ..., \pi_S)$
- Source locations: $(\mu_1, \mu_2, ..., \mu_S)$
- Time-varying Intensities: $(\boldsymbol{\lambda}_1, ..., \boldsymbol{\lambda}_S)$
- Time and source-varying Energy distributions:

$$(\alpha_{jk}, \beta_{jk}), \quad j = 1, ..., J, k = 1, ..., S$$

- Event Allocation: $(z_1, z_2, ..., z_n)$

Fitting Procedures:

- Gibbs sampling and Metropolis-Hastings MCMC

Bayesian Mixture Model

Allocation Output: For each iteration, r, we have

$$\boldsymbol{z}^{(r)} = \left(z_1^{(r)}, z_2^{(r)}, ..., z_n^{(r)}\right)$$

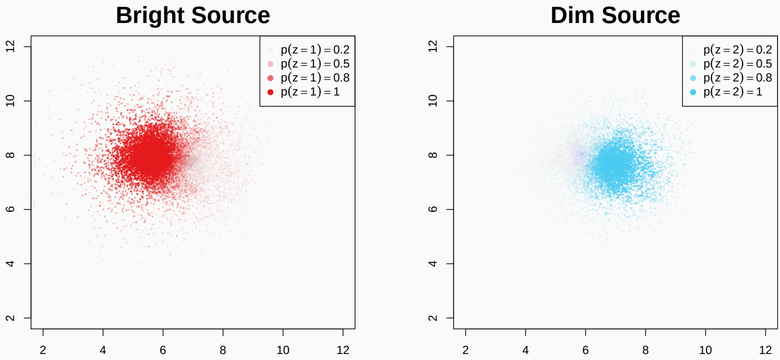
Events list for source k:

$$\text{Subset events such that } z_i^{(r)} = k$$

Allocation Probabilities:

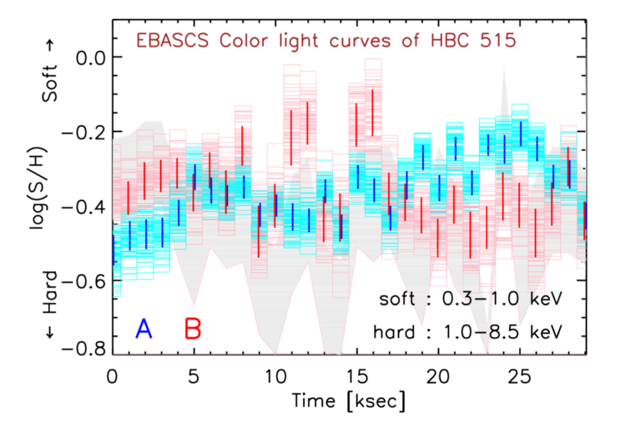
$$\Pr(z_i = k | x_i, y_i, E_i, t_i) \approx \frac{1}{R} \sum_{i=1}^{R} 1\{z_i^{(r)} = k\}$$

Results for HBC515ab



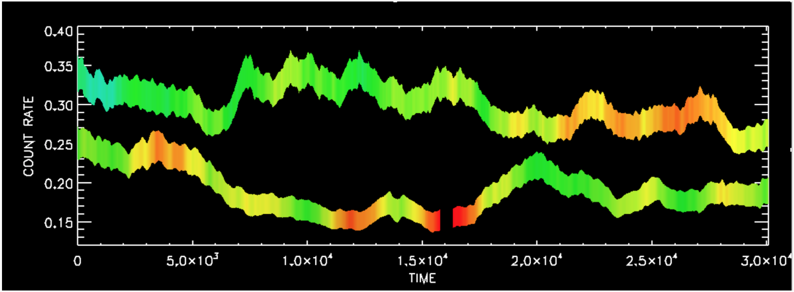
Allocation probabilities for sources as an alternative to core/wing extraction

HBC515A a/b: Hardness ratio $log \frac{S}{H}$ light curves



- Sources can be treated as if *isolated* for each allocation ($z^{(r)}$)
- Spectra vary and differ at times.

HBC515A a/b: Light-Energy Curves (sliding window)



- Red areas indicate source softened, blue = hardened
- Spectra are changing – especially dimmer source

Conclusions

- **Deterministic** allocation rule $\rightarrow$ **probabilistic** allocation rule

$$z_i = 1 | x_i, y_i \rightarrow p(z_i = 1 | x_i, y_i, e_i, t_i)$$

- Quantifying uncertainty like this **utilizes more data** and **more closely reflects reality**
- Enables more **honest down-stream analyses** by reflecting uncertainty