Disambiguating Sources II
Valentine’s Day Edition

Luis Campos (with David Jones)

February 14, 2017
Illustrative Example: HBC515
Illustrative Example: HBC515

- Discovered in 2012
- Part of a system containing multiple young stellar objects (YSOs)
- Difficult to Study: One study published in 2017 Principe, et.al. ‘The Multiple Young Stellar Objects of HBC 515: An X-ray and Millimeter-wave Imaging Study in (Pre-main Sequence) Diversity’
Illustrative Example: HBC515

- Discovered in 2012
- Part of a system containing multiple young stellar objects (YSOs)
- Difficult to Study: One study published in 2017 Principe, et.al. ‘The Multiple Young Stellar Objects of HBC 515: An X-ray and Millimeter-wave Imaging Study in (Pre-main Sequence) Diversity’

Methods:

- “The point spread functions of the two binary components overlap significantly, however, complicating their photometric and spectral decomposition.”
- “Hence, we used two spectral extraction regions for each of the component sources”
Illustrative Example: HBC515
Problem Setup

Given:

- \((x_i, y_i)\): photon-level location information
- \(E_i\): photon-level Energy information
- \(t_i\): photon-level time-arrival info.
- \(S\): number of sources (assume known for now)

Can we
- allocate photon \(i\) to one of the sources?
- approximate the source centers? \((\mu_x(s), \mu_y(s))\)

From these, we can then go on and calculate more complex things,
- source intensities
- distance between sources
- better models for time arrival (O-U process, flares, etc)...
Problem Setup

Given:

- \((x_i, y_i)\): photon-level location information
- \(E_i\): photon-level Energy information
- \(t_i\): photon-level time-arrival info.
- \(S\): number of sources (assume known for now)

Can we

- allocate photon \(i\) to one of the sources? \(z_i = s\)
- approximate the source centers? \((\mu_x(s), \mu_y(s))\)
Problem Setup

Given:

- $\left( x_i, y_i \right)$: photon-level location information
- $E_i$: photon-level Energy information
- $t_i$: photon-level time-arrival info.
- $S$: number of sources (assume known for now)

Can we

- allocate photon $i$ to one of the sources? $z_i = s$
- approximate the source centers? $(\mu_x(s), \mu_y(s))$

From these, we can then go on and calculate more complex things,

- source intensities
- distance between sources
- better models for time arrival (O-U process, flares, etc)...
Problem Setup

Given:

- \((x_i, y_i)\): photon-level location information
- \(t_i\): photon-level time-arrival information
- \(E_i\): photon-level Energy info. (/ht/ D. Jones #BASCS)
- \(S\): number of sources (assume known for now)

Can we

- allocate photon \(i\) to one of the sources? \(z_i = s\)
- approximate the source centers? \((\mu_x(s), \mu_y(s))\)

From these, we can then go on and calculate more complex things,

- source intensities
- distance between sources
- better models for time arrival (O-U process, flares, etc)…
Illustrative Example: HBC515
Illustrative Example: HBC515
Illustrative Example: HBC515
Using time-arrival?

Distribution of Time Arrivals

Distribution of Time Arrivals
How do we model this?

We’re going to model the location and time arrival conditionally independent given $z_i$, i.e. $(x_i, y_i) \perp \perp t_i | z_i$
How do we model this?

We’re going to model the location and time arrival conditionally independent given $z_i$, i.e. $(x_i, y_i) \perp \perp t_i|z_i$

$$p(z_i = s|x_i, y_i, t_i = t) \propto p(x_i, y_i, t_i = t|z_i = s)p(z_i = s)$$

$$= p(x_i, y_i|z_i = s)p(t_i = t|z_i = s)p(z_i = s)$$
How do we model this?

We’re going to model the location and time arrival conditionally independent given $z_i$, i.e. $(x_i, y_i) \perp \perp t_i | z_i$

$$p(z_i = s | x_i, y_i, t_i = t) \propto p(x_i, y_i, t_i = t | z_i = s)p(z_i = s)$$

$$= p(x_i, y_i | z_i = s)p(t_i = t | z_i = s)p(z_i = s)$$

- $p(x_i, y_i | z_i = s)$: Can use the King Profile (a 2-d Cauchy)
- $p(z_i = s)$: Can use a Dirichlet distribution.
- $p(z_i = s | t_i = t)$?
A Simple Time Model

Model the time-arrival as piece-wise constant. For each source \( (s) \):

- \( n_s \): number of photons from source \( s \)
- Break time into \( K \) bins (fixed)
- ... at fixed locations \( (b_0, b_1, ..., b_K) \)
- \( n_{s1}, n_{s2}, ..., n_{sK} \): photon counts in each bin.
- For now, \( (b_0, b_1, ..., b_K) \) are shared across sources.
A Simple Time Model

Model the time-arrival as piece-wise constant. For each source $(s)$:

- $n_s$: number of photons from source $s$
- Break time into $K$ bins (fixed)
- ... at fixed locations $(b_0, b_1, ..., b_K)$
- $n_{s1}, n_{s2}, ..., n_{sK}$: photon counts in each bin.
- For now, $(b_0, b_1, ..., b_K)$ are shared across sources.

$$(\alpha_{s1}, ..., \alpha_{sK})|n_s \sim \text{Dir}(n_{s1} + \tilde{\alpha}_s, n_{s2} + \tilde{\alpha}_s, ..., n_{sK} + \tilde{\alpha}_s)$$

- Time arrival parameters: $(\alpha_{s1}, ..., \alpha_{sK})$, for $s = 1, .., S$
- Prior on arrival dist: $\alpha_s$, for $s = 1, .., S$
A Simple Time Model

- \( p(z_i = s | t_i = t) \)
A Simple Time Model

- \( p(z_i = s | t_i = t) \)?

We’ll assign the probabilities as follows, for each photon:
- Find \( k \) such that \( b_{k-1} < t_i < b_k \)
- Assign the probabilities

\[
\begin{align*}
  p(z_i = 1 | t_i = t) &= \frac{\alpha_{1k}}{\sum_{s=1}^{S} \alpha_{sk}} \\
  p(z_i = 1 | t_i = t) &= \frac{\alpha_{2k}}{\sum_{s=1}^{S} \alpha_{sk}} \\
  &\vdots \\
  p(z_i = S | t_i = t) &= \frac{\alpha_{Sk}}{\sum_{s=1}^{S} \alpha_{sk}}
\end{align*}
\]
Simulated Data

We want to first study a few simple things:

1. If we incorporate time, will we do better than not incorporating?
2. If (1), at what point do our gains fall apart?
3. A simple example for (2), how does the distance between the sources affect our ability to distinguish them?

Simulation:

1. Two sources with background
2. Source separation (0.5, 1, 1.5, 2)
3. Ideal time arrival distributions (for now)
Simulated Data (separation = 2)

Two Overlapping Sources

Bright Source

Dim Source

Background
Simulated Data (separation = 0.5)
What we have to work with (separation = 2)
What we have to work with (separation = 0.5)
Source Location (separation = 2)

(a) Location and Time Model

(b) Location Only Model
**Source Location (separation = 0.5)**

(c) Location and Time Model

(d) Location Only Model
Source Separation (with replicates)

Posterior CI: Separation

Location and Time

Location Only

Posterior: Separation

Separation

Separation
Source Intensity (separation = 2)

(e) Location and Time Model

(f) Location Only Model
Source Intensity (separation = 0.5)

(g) Location and Time Model
(h) Location Only Model
Source Intensity: Bright Source (with replicates)
Source Intensity: Dim Source (with replicates)
Source Intensity: Background (with replicates)
Average Correct Source Allocation (with replicates)
Average Correct Source Allocation by Source (with replicates)

% Correct Allocation

**Bright Source**

- Location and Time
- Location Only

**Dim Source**

- Location and Time
- Location Only

**Background**

- Location and Time
- Location Only
Where do we go from here?

What we’ve done:

▶ We’ve shown that using time can help disambiguate sources.
▶ Even simple models (constant functions) can prove useful.

Future directions:

▶ Real light curve shapes (mine were too simple)
▶ More complex time models?
  ▶ vary cut-offs per source
  ▶ more complex models per source
Light Curves with Spikes (Separation = 1)

(i) Source Centers

(j) Light Curve Posterior (bright)
Light Curves with Spikes (Separation = 1)

Posterior: Light Curves for Dim Source

Posterior: Light Curves for Background

(k) Light Curve Posterior (dim)  (l) Light Curve Posterior (bg)
Light Curves with Spikes (Separation = 0.5)

(m) Source Centers

(n) Light Curve Posterior