DEFINING REGIONS THAT CONTAIN COMPLEX ASTRONOMICAL STRUCTURES
We are interested in defining an outline around extragalactic jets coming from quasars at high redshift ($z > 2.1$) in X-ray images. Defining this boundary is important for accurate luminosity and flux calculations. Detecting jets is difficult because they are diffuse sources (no edges, or center) and dim compared to the quasar. Images of high redshift jets are of low resolution and few X-ray photons.
OBSERVATIONAL DATA

- Chandra X-ray Observatory - ACIS
- 64 x 64 or 128 x 128 pixel image centered on quasar
- High to intermediate redshift (2.10 < z< 4.72)
**BACKGROUND**

**REGION OF INTEREST**

- **Region of Interest (ROI)** - region containing the jet or a partition of the jet (e.g. node or lobe)

- Previous work tests whether or not a jet exists in a predefined ROI (McKeough et al. 2016, Stein et al. 2015)
Ability to detect jet is sensitive to fit of ROI

Issues with previous methods:

- Region is defined using radio imaging
  - Not always available
  - Not always aligned with X-ray imaging
- Region definition relies on human interaction
  - Inefficient and source of potential error
Define a boundary around the ROI of an irregularly shaped, diffuse source.

Give a measure of uncertainty.
ROADMAP

- Pre-Process Image
- Pixel Assignments
- Boundary of ROI
LOW COUNT IMAGE RECONSTRUCTION AND ANALYSIS (LIRA)

- Multi-scale Bayesian method
  - Intensity in “splits” of the image rather than individual pixels
- Removes quasar & deconvolve Point Spread Function (PSF)
- Creates posterior draws for residual pixels as a series of images that capture the emission that is present in excess of the quasar (i.e. the jet)
LOW COUNT IMAGE RECONSTRUCTION AND ANALYSIS (LIRA)
ROADMAP

Pre-Process Image

Pixel Assignments

Boundary of ROI
We are given observation $Y$ from which we draw the LIRA output:

$\tilde{\lambda}_{ij} | Z, \tau_\pm, \sigma^2_\pm \sim \text{Normal}(\tau_-, \sigma^2_-) \mathbb{I}_{z_{ij} = -1} + \text{Normal}(\tau_+, \sigma^2_+) \mathbb{I}_{z_{ij} = +1}$

We want to assign each pixel to either the background (-1) or the ROI (+1):

$z_{ij} = \{-1, +1\}$

Each pixel assignment will have its own average intensity:

$\tau_-, \tau_+$

We suspect the variance of the source will be greater than the background:

$\sigma^2_-, \sigma^2_+$
 ESTABLISH MODEL

2D ISING PRIOR

\[
p(z|\beta) = \frac{\exp(\beta \sum_{i,j,i',j' \in |i-j, j'|=1} z_{ij} z_{i'j'})}{\tilde{Z}(\beta)}
\]

- Inverse temperature: \(\beta\)
  - Higher \(\beta\) induces more correlation between pixels

- Partition function: \(\tilde{Z}(\beta)\)
  - Estimated via Beale (1996) assuming periodic structure

- Commonly used in modeling ferromagnetism.

- Induces spatial correlation; adjacent pixels will tend to have the same assignment.
REMINDER: MODEL SETUP

- **Likelihood:**

\[
\sqrt{\lambda_{ij}} | Z, \tau_\pm, \sigma_\pm^2 \sim \text{Normal}(\tau_- , \sigma_-^2) \mathbb{I}_{z_{ij} = -1} + \text{Normal}(\tau_+ , \sigma_+^2) \mathbb{I}_{z_{ij} = 1}
\]

- **Prior:**

\[
p(z|\beta) = \frac{\exp(\beta \sum_{ij, i'j' \in|ij-i'j'|=1} z_{ij} z_{i'j'})}{\tilde{Z}(\beta)}
\]
STEP 1 – LIKELIHOOD PARAMETERS

- Draw from posterior directly:

- Priors:

  \[ \tau_{\pm} \sim \text{Normal}(\mu_0, \sigma^2_{\pm}) \]

  \[ \sigma^2_{\pm} \sim \text{Inv-\chi}^2(\nu_0, \omega^2_0) \]

STEP 2 – TEMPERATURE PARAMETER

- Drawn through Metropolis Hastings

- Prior:

  \[ \beta \sim \text{Gamma}(a_\beta, b_\beta) \]
A well established way to draw the spin state given a specific temperature is Swendsen & Wang (1987).

The S-W method takes a spin system \( z | \beta \) and induces a bigger system that contains the original \( N \) spin variables and \( M \) additional bond variables, denoted by \( d \).

Define joint distribution that couples spins to bonds:

\[
p(z, d | \tilde{\lambda}, \tau_\pm, \sigma^2_\pm, \beta) \propto \prod_{m=1}^{M} g_m(z_m, d_m | \beta) \prod_{ij} f(\tilde{\lambda}_{ij} | z, \tau_\pm, \sigma^2_\pm)
\]

Marginal distribution of \( z \) is equal to our posterior.

\[
\sum_d p(z, d | \tilde{\lambda}, \tau_\pm, \sigma^2_\pm, \beta) = p(z | \tilde{\lambda}, \tau_\pm, \sigma^2_\pm, \beta)
\]

Conditional distributions are easy to sample from.

\[
p(z|d, \beta-) \quad p(d|z, \beta-)
\]
Bonds can be disconnected (0) or connected (1).

\[ d = \{0, 1\} \]
Sample from $p(d|z, \beta)$

If two spins connected to bond are equal, set the bond $d_m$ equal to 1 with probability $p=1-\exp(-2\beta)$, and 0 otherwise.
Sample from $p(d|z, \beta)$

If two spins connected to bond are equal, set the bond $d_m$ equal to 1 with probability $p=1-\exp(-2\beta)$, and 0 otherwise.
Sample from $p(z|d, \beta)$

- Bonds connect spins into $C$ cluster.
- **Cluster** - all pixels that are connected by a bond $d_m=1$
- Each cluster will take spin $+1$ with probability $p_+$
- $-1$ with probability $p_- = 1 - p_+$

$$p_\pm \propto \prod_{ij \in C} f(\tilde{\lambda}_{ij} | z_{ij} = \pm 1, \tau_\pm, \sigma^2_\pm)$$
Sample from $p(z|d, \beta)$

$p(z=+1) = p_+$
1. Get many posterior draws from LIRA
2. Apply Ising step to each LIRA draw
3. Average across LIRA-Ising iterations to get probability map.
PROBABILITY MAP & BOUNDARY ESTIMATE

PROBABILITY MAP

- Probability each pixel is a member of the ROI:
ROADMAP

Pre-Process Image

Pixel Assignments

Boundary of ROI
Maximize posterior predictive:

\[ P(Z|Y) = \int P(Z, \theta, \lambda|Y)d\theta d\lambda \]
Maximize posterior predictive:

\[ P(Z|Y) = \int P(Z, \theta, \lambda|Y) d\theta d\lambda \]

Ideally we could approximate this as:

\[ \hat{P}(Z|Y) = \frac{1}{N} \sum_{k=1}^{N} P(Z|\theta^{(k)}, \lambda^{(k)}) \]
Maximize posterior predictive:

\[ P(Z|Y) = \int P(Z, \theta, \lambda|Y) d\theta d\lambda \]

Ideally we could approximate this as:

\[ \hat{P}(Z|Y) = \frac{1}{N} \sum_{k=1}^{N} P(Z|\theta^{(k)}, \lambda^{(k)}) \]

... but this is very difficult.
MAXIMIZE POSTERIOR RATIO

- Compare two different Z states:

\[
\frac{\hat{P}(Z_1|Y)}{\hat{P}(Z_2|Y)} = \frac{\sum_{k=1}^{N} \exp(\log P_k(Z_1))}{\sum_{k=1}^{N} \exp(\log P_k(Z_2))}
= \sum_{k=1}^{N} w_k \exp(\log \frac{P_k(Z_1)}{P_k(Z_2)})
\]
Neighborhood statistic:

\[
N_{ij} = \frac{\sum_{i' j' \in |i j - i' j'| = 1} z_{ij} z_{i' j'}}{\sum_{i' j' \in |i j - i' j'| = 1} |i j - i' j'|}
\]

- \(N_{ij} = 0.75\)
- \(N_{ij} = 0\)
- \(N_{ij} = 1\)
**BOUNDARY OF ROI**

**OPTIMIZATION SPACE**

- Average $N_{ij}$ across all posterior draws
- Rank $N_{ij}$ from highest to lowest
- Build space to optimize over:
  - For the zip with the highest corresponding $N_{ij}$, set to 1 and the remainder to -1
  - Repeat including the next highest $N_{ij}$ until all pixels are 1
Maximize across all $Z$ created using the neighborhood statistic and all $Z$ drawn from the posterior

We will always compare the new $Z$ with the $Z$ at the current maximum

To build a confidence interval take more posterior iterations and repeat the process (TBD)
BOUNDARY OF ROI
RESULTS

- **Original**
  - Pixel Assignments

- **Pre-Process Image**

- **Pixel Assignments**
  - Prob

- **Boundary of ROI**
  - sqrt(Count)

- **Boundary of ROI**
  - sqrtCounts
Future Work
ADJACENT PIXEL DEFINITION

- Could be modified to the 8 nearest pixels instead of 4.
- Modified to include pixels beyond just the adjacent pixels
- Correlation as a function of distance
**Potts Model**

- Want to identify multiple partitions of the jet (e.g. nodes)
- Potts is a more generalized version of the Ising model allows for more than two spin assignments:

\[ z_{ij} = \{0, 1, 2, 3, \ldots \} \]

**Different Likelihoods**

- **Hurdle model** - Account for many of the background pixels in the LIRA output being zero.
LIRA has been successful in analyzing low count images and extracting noisy structure.

- No way to define a ROI
- No correlation structure between pixels
- Utilized an Ising distribution and corresponding techniques to create a probabilistic ROI.
Model Compatibility
“IDEAL” MODIFICATION TO LIRA

- Current LIRA output:
  \[ P(\tilde{\lambda}|Y) \]

- The missing piece of LIRA is the pixel membership indicator:
  \[ z_{ij} = \{-1, +1\} \]

- An ideal joint model (denote using subscript \( \mathcal{J} \)) would infer \( \lambda_{ij} \) and \( z_{ij} \) simultaneously
  \[ P_{\mathcal{J}}(\tilde{\lambda}, z|Y) \propto f(Y|\tilde{\lambda}, z)\pi_{\mathcal{J}}(\tilde{\lambda}, z) \]
MODEL COMPATIBILITY

OUR APPROACH

- Two-step approach:
  - LIRA “as is” (model $S_1$)
    \[ P_{S_1}(\tilde{\lambda}|Y) \propto f(Y|\tilde{\lambda})\pi_{S_1}(\tilde{\lambda}) \]
  - Ising (model $S_2$) conditional on ONE draw of from $S_1$
    \[ P_{S_2}(z|\tilde{\lambda}) \propto P_{S_2}(\tilde{\lambda}|z)\pi_{S_2}(z) \]
  - Combine to get desired model:
    \[
    P_S(\tilde{\lambda}, z|Y) = P_{S_1}(\tilde{\lambda}|Y)P_{S_2}(z|\tilde{\lambda})
        \propto f(Y|\tilde{\lambda})\pi_{S_1}(\tilde{\lambda}) \frac{P_{S_2}(\tilde{\lambda}|z)\pi_{S_2}(z)}{P_{S_2}(\tilde{\lambda})}
    \]
SUFFICIENT CONDITIONS

\[ P_J(\tilde{\lambda}, z|Y) \propto f(Y|\tilde{\lambda}, z) \pi_J(\tilde{\lambda}, z) \iff P_S(\tilde{\lambda}, z|Y) = P_{S_1}(\tilde{\lambda}|Y) P_{S_2}(z|\tilde{\lambda}) \]
\[ \propto f(Y|\tilde{\lambda}) \pi_{S_1}(\tilde{\lambda}) \frac{P_{S_2}(\tilde{\lambda}|z) \pi_{S_2}(z)}{P_{S_2}(\tilde{\lambda})} \]

- Assignment information does not effect distribution of photon counts:

\[ f(Y|\tilde{\lambda}) = f(Y|\tilde{\lambda}, z) \]
MODEL COMPATIBILITY

**SUFFICIENT CONDITIONS**

\[ P_S(\lambda, z|Y) = P_{S_1}(\lambda|Y)P_{S_2}(z|\lambda) \]

\[ \propto f(Y|\tilde{\lambda}) \pi_{S_1}(\tilde{\lambda}) \]

- Assignment information does not effect distribution of photon counts:

\[ f(Y|\tilde{\lambda}) = f(Y|\tilde{\lambda}, z) \]

- LIRA prior on photon counts is compatible with Ising model prior on assignments:

\[ \pi_{S_1}(\tilde{\lambda}) = \int \pi_{\mathcal{J}}(\tilde{\lambda}, z) dz = \int P_{S_2}(\tilde{\lambda}|z)\pi_{S_2}(z) dz \]
HOW FAR OFF ARE WE?

\[ P_J(\tilde{\lambda}, z|Y) \propto f(Y|\tilde{\lambda}, z)\pi_J(\tilde{\lambda}, z) \quad P_S(\tilde{\lambda}, z|Y) = P_{S_1}(\tilde{\lambda}|Y)P_{S_2}(z|\tilde{\lambda}) \]
\[ \propto f(Y|\tilde{\lambda})\pi_{S_1}(\tilde{\lambda})\frac{P_{S_2}(\tilde{\lambda}|z)\pi_{S_2}(z)}{P_{S_2}(\tilde{\lambda})} \]

- Inference for \( \lambda \) is equivalent:

\[ P_J(\lambda|Y) \propto f(Y|\lambda) \int \pi_J(\lambda, z)dz = f(Y|\lambda)\pi_{S_1}(\lambda) \propto P_S(\lambda|Y)dz \]
**HOW FAR OFF ARE WE?**

Inference for $\lambda$ is equivalent:

\[
P_J(\tilde{\lambda}, z | Y) \propto f(Y | \tilde{\lambda}, z) \pi_J(\tilde{\lambda}, z)
\]

\[
P_S(\tilde{\lambda}, z | Y) = P_{S_1}(\tilde{\lambda} | Y) P_{S_2}(z | \tilde{\lambda})
\]

\[
\propto f(Y | \tilde{\lambda}) \pi_{S_1}(\tilde{\lambda}) \frac{P_{S_2}(\tilde{\lambda} | z) \pi_{S_2}(z)}{P_{S_2}(\tilde{\lambda})}
\]

- Inference for $\lambda$ is equivalent:

\[
P_J(\lambda | Y) \propto f(Y | \lambda) \int \pi_J(\lambda, z) dz = f(Y | \lambda) \pi_{S_1}(\lambda) \propto P_S(\lambda | Y) dz
\]

- Posterior inference is bounded by the prior divergence (which can be calculated)

\[
D_{KL}(P_J(\lambda, z | Y), P_S(\lambda, z | Y)) = \int P_J(\lambda | Y) D_{KL}(P_J(z | \lambda), P_S(z | \lambda)) d\lambda
\]
REFERENCES

MULTI-SCALE IMAGE REPRESENTATION

- Stores total intensities and series of four way split proportions such that the product recovers original pixel intensities
- Pixel Intensity
  \[ \Lambda = \{ \Lambda_i, I = 1 \ldots N \} \]
- Splits
  \[ D_{k, l_k(i), m_k(i)} \]
  - Split proportion at scale \( k \) corresponding to group \( i \)

\[
\Lambda_i = G \prod_{k=1}^{K} D_{k, l_k(i), m_k(i)}
\]
MULTI-SCALE IMAGE REPRESENTATION

G

One Pixel

One Split

GD_{112}

Four Pixels

Four Splits

GD_{111}D_{222}

16 Pixels

16 Splits

GD_{112}D_{222}D_{362}

64 Pixels

GD_{111}D_{211}D_{311}

...
LIRA CONT.

LIKELIHOOD

- Probability photon originating in pixel $i$, is observed in pixel $j$ (PSF):

  $P_i = \{P_{ij}, j = 1, \ldots N\}$

- Observed pixel counts:

  $Y = \{Y_i, i = 1, \ldots N\}$

- Distribution of $Y$:

  $Y_j | \Lambda, \Lambda^B \sim \text{Poisson} \left( \left( \sum_{i \in I} P_{ij} \Lambda_i \right) + \Lambda^B_j \right)$

- Suppress background to obtain likelihood:

  $L(\Lambda, \Lambda^B | Y) \equiv L(\Lambda | Y) \propto \prod_{j \in I} p(Y_j | \Lambda)$
Prior on total intensity:

\[ G \sim \text{Gamma}(\gamma_0, \gamma_1) \]

Prior on splits:

\[ D_{k1} \equiv \{D_{k1m}, m = 1, \ldots, 4\} \overset{d}{\sim} \text{Dirichlet}(\alpha_k, \alpha_k, \alpha_k, \alpha_k) \]

\[ k = 1, \ldots, K, \quad l = 1, \ldots, 4^{k-1} \]

Hyperprior favors smoother image:

\[ p(\alpha_k) \propto \exp(-\delta \alpha^3 / 3) \]
CYCLE SPINNING

- Multiscale format produces checkerboard-like patterns

- Solution:
  - Shift center of image randomly before making splits
  - Splits wrap around edges of image to induce translation invariance
SWENDSEN-WANG
COUPLING SPINS TO BONDS

- Factor coupling bonds and spins is:

\[
g_m(z_m, d_m) = \begin{cases} 
  d_m = 0 \\
  z_{i'j'} = -1 \\
  e^{-\beta} \\
  d_m = 1 \\
  z_{i'j'} = +1 \\
  e^{-\beta} \\
  z_{ij} = -1 \\
  e^{-\beta} \\
  z_{ij} = +1 \\
  e^{-\beta}
\end{cases}
\]

- Rescale by constant factor: \( p = 1 - e^{-2\beta} \)

\[
\tilde{g}_m(z_m, d_m) = \begin{cases} 
  d_m = 0 \\
  z_{i'j'} = -1 \\
  1 - p \\
  d_m = 1 \\
  z_{i'j'} = +1 \\
  1 - p \\
  z_{ij} = -1 \\
  1 - p \\
  z_{ij} = +1 \\
  1 - p
\end{cases}
\]