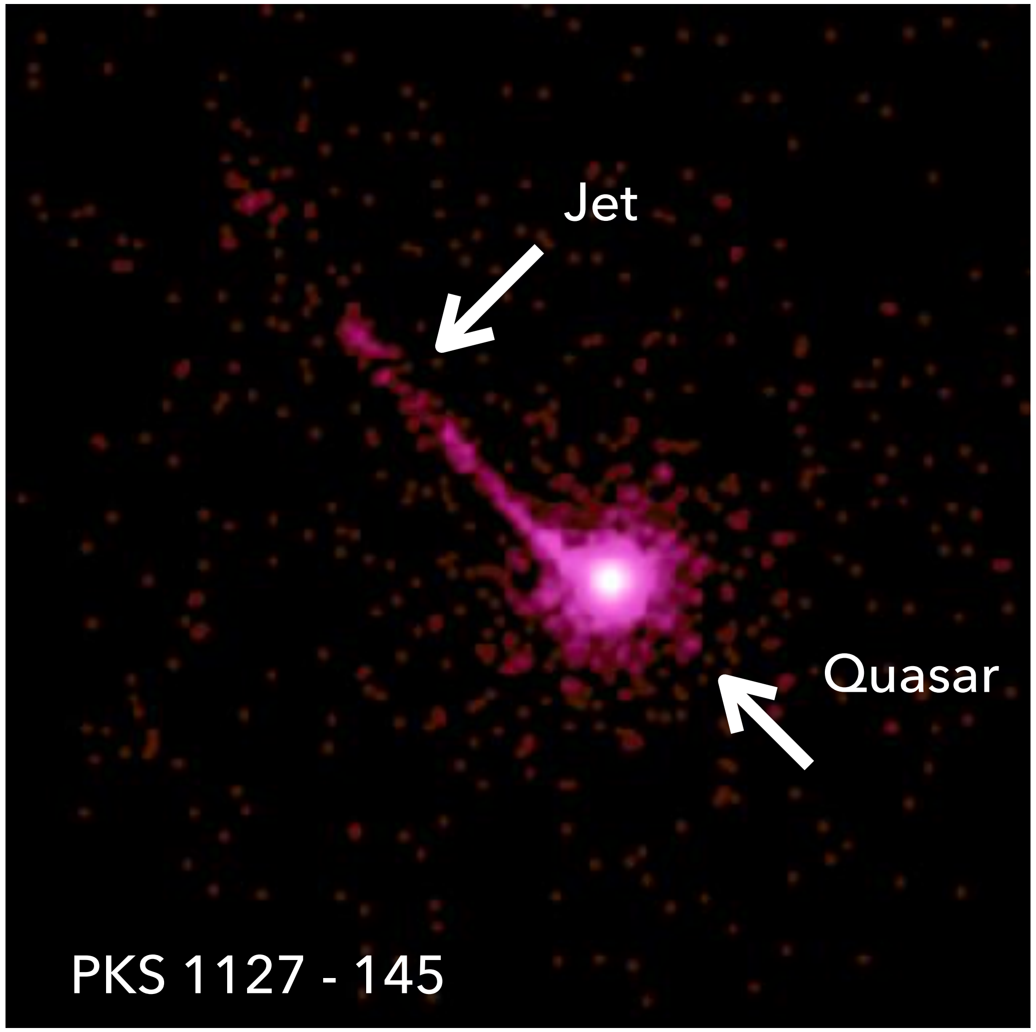
KATY MCKEOUGH & SHIHAO YANG

DEFINING REGIONS THAT CONTAIN X-RAY JETS IN HIGH REDSHIFT QUASARS

SCIENTIFIC MOTIVATION

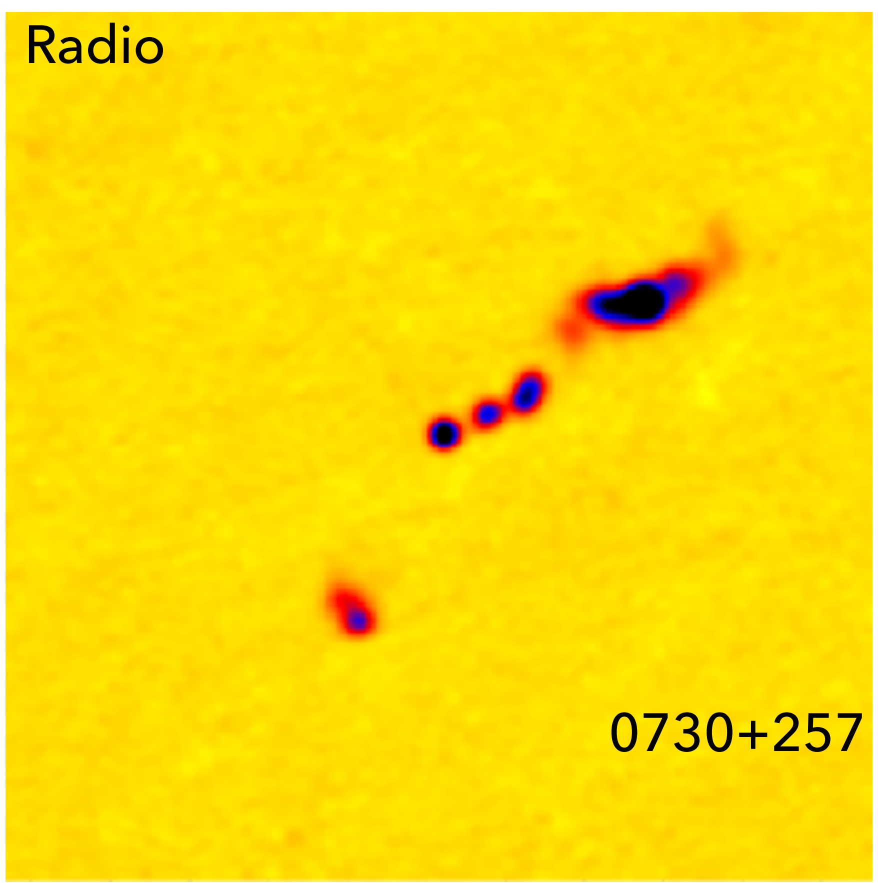
- We are interested in defining an outline around extragalactic jets coming from quasars at high redshift ($z>2.1$) in X-ray images
- Defining this boundary is important for accurate luminosity and flux calculations.
- Detecting jets is difficult because they are diffuse sources (no edges).
- Images of high redshift jets are of low resolution and few X-ray photons

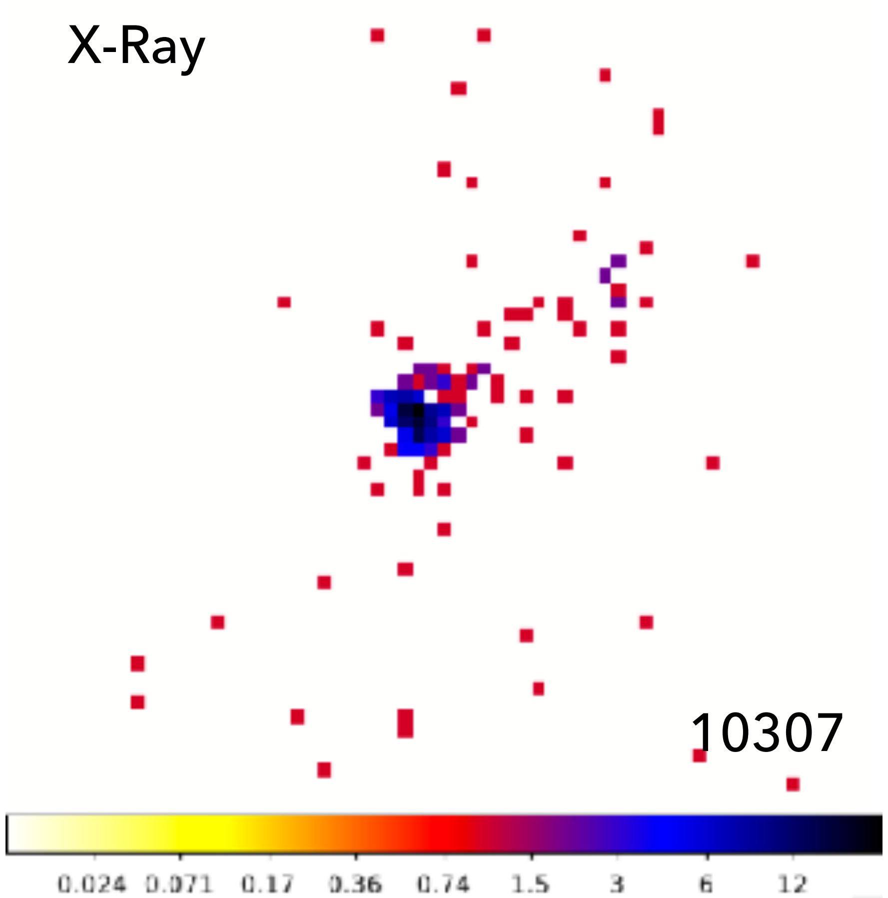


NASA/CXC/A.Siemiginowska(CfA)/J.Bechtold(U.Arizona)

OBSERVATIONAL DATA

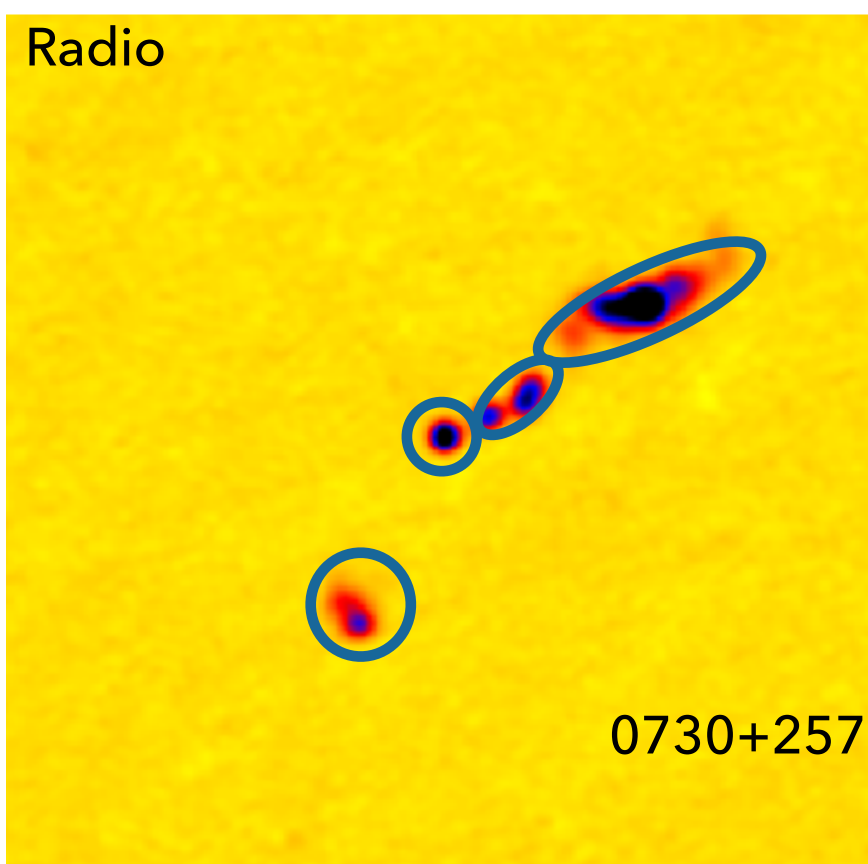
- Chandra X-ray Observatory - ACIS
- 64 x 64 or 128 x 128 pixel image centered on quasar
- High and intermediate redshift (2.10 < z< 4.72)

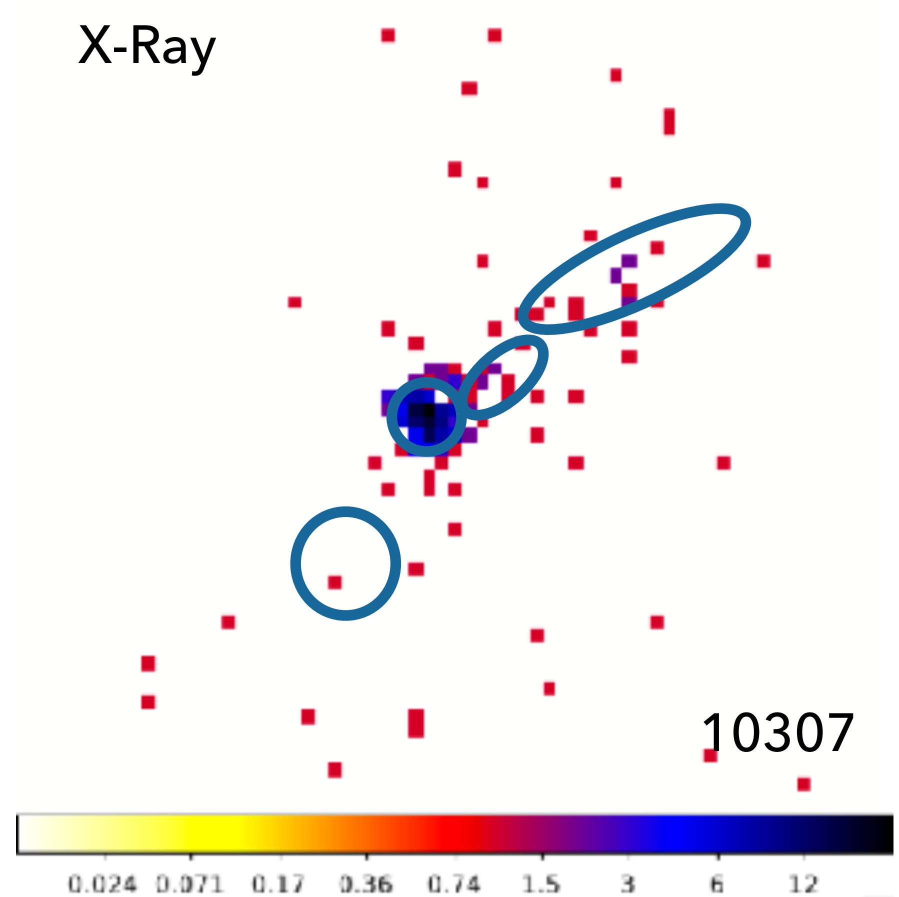




REGION OF INTEREST

- **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)





REGION OF INTEREST

- 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

GOAL

Using only the X-ray observation of a quasar and a jet, we are interested in defining an ROI around the jet.

ROADMAP

- Pre-processing using LIRA
- Establish model for pixel assignments
- Model compatibility
- Draw assignments via Gibbs Sampler
- Results
- Future directions

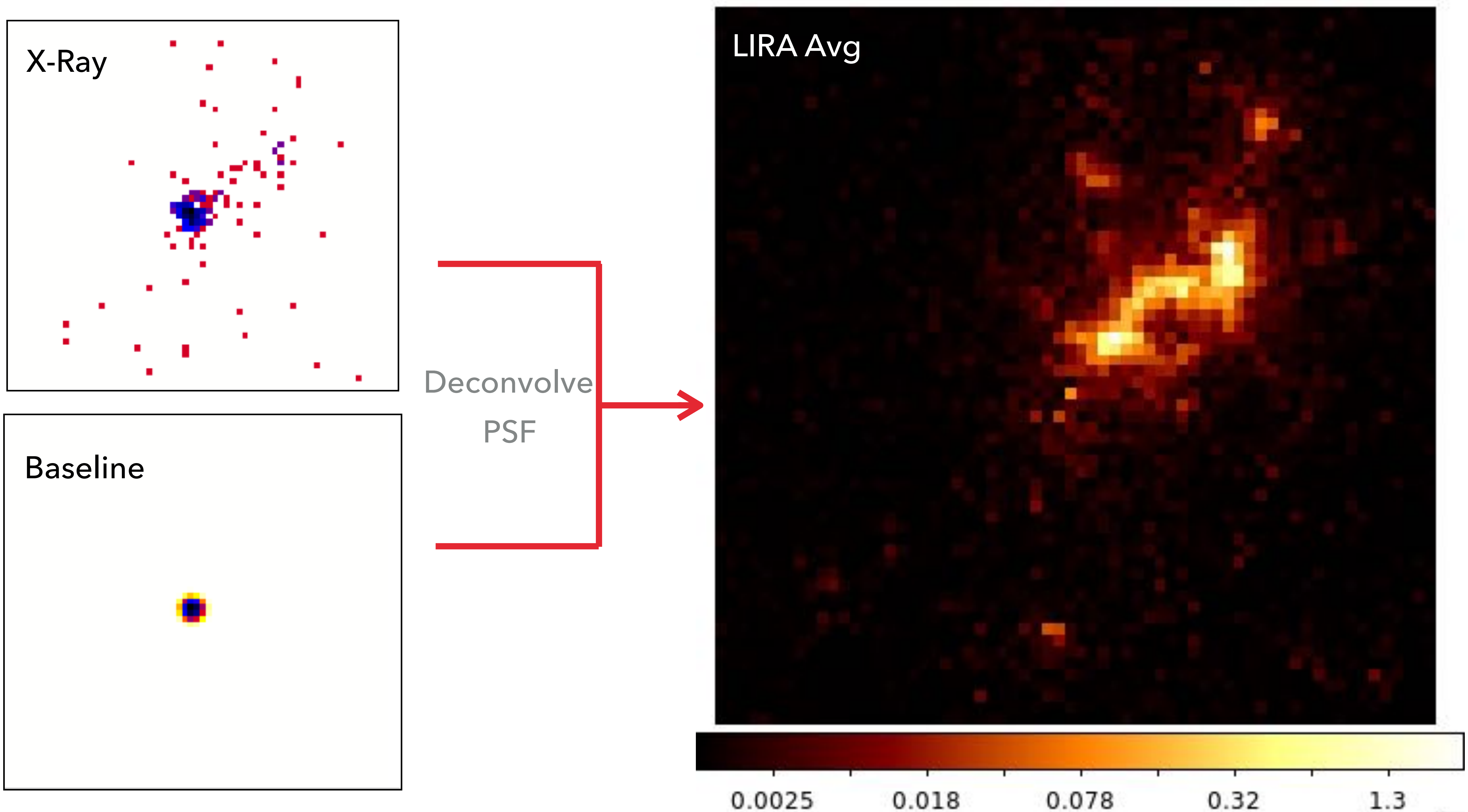
ROADMAP

- **Pre-processing using LIRA**
- Establish model for pixel assignments
- Model compatibility
- Draw assignments via Gibbs Sampler
- Results
- Future directions

LOW COUNT IMAGE RECONSTRUCTION AND ANALYSIS (LIRA)

- Esch et al (2004), Connors & van Dyk (2007)
- Multi-scale Bayesian method
    - Intensity in "splits" of the image rather than individual pixels
- Removes quasar & deconvolve Point Spread Function (PSF)
- Creates posterior 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-processing using LIRA
- **Establish model for pixel assignment**
- Model compatibility
- Draw assignments via Gibbs Sampler
- Results
- Future work

LIKELIHOOD

$$\tilde{\lambda}_{ij}|z, \tau_0, \tau_1, \sigma^2 \sim \text{Log-Normal}(\tau_0, \sigma^2)\mathbb{I}_{z_{ij}=-1} + \text{Log-Normal}(\tau_1, \sigma^2)\mathbb{I}_{z_{ij}=+1}$$

- We are given observation Y from which we draw the LIRA output: $\tilde{\lambda}|Y$
- We want to assign each pixel to either the background (-1) or the ROI (1): $z = \{-1, +1\}$
- Each pixel assignment will have its own average intensity: τ_-, τ_+
- For now, the assignees have the same variances: σ^2

2D ISING PRIOR

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

- Inverse temperature: β
  - Higher β 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.

POSTERIOR

$$p(z|\tilde{\lambda}, \tau_0, \tau_1, \sigma^2, \beta) \propto f(\tilde{\lambda}|z, \sigma^2, \tau_0, \tau_1)p(z|\beta)\pi(\sigma^2, \tau_0, \tau_1, \beta)$$

ROADMAP

- Pre-processing using LIRA
- Establish model for pixel assignments
- **Model compatibility**
- Draw assignments via Gibbs Sampler
- Results
- Future directions

HYPOTHETICAL IDEAL LIRA…

- Curent LIRA output:

$$P(\lambda_{ij}, 1 \leqslant i, j \leqslant 64 | Y_{ij}, 1 \leqslant i, j \leqslant 64)$$

- The missing piece of LIRA is the pixel membership indicator:

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

- An ideal join model (denote using subscript $\mathcal{J}$) would infer λ_{ij} and z_{ij} simultaneously

$$P_{\mathcal{J}}(\lambda, z | Y) \propto f(Y | \lambda, z) \pi_{\mathcal{J}}(\lambda, z)$$

OUR APPROACH IS

- Using LIRA "as is", treating it as pre-processing, trying to infer pixel membership indicator z_{ij} from LIRA output posterior draws of λ_{ij}.
- Essentially a two-step approach:
  - LIRA (model S_1)

$$P_{\mathcal{S}_1}(\lambda|Y) \propto f(Y|\lambda)\pi_{\mathcal{S}_1}(\lambda)$$

  - Ising (model S_2) conditional on ONE draw of from S

$$P_{\mathcal{S}_2}(z|\tilde{\lambda}) \propto P_{\mathcal{S}_2}(\tilde{\lambda}|z)\pi_{\mathcal{S}_2}(z)$$

- So what is this model $S=S_1+S_2$?

$$P_{\mathcal{S}}(\tilde{\lambda}, z|Y) = P_{\mathcal{S}_1}(\tilde{\lambda}|Y)P_{\mathcal{S}_2}(z|\tilde{\lambda})$$

$$P_{\mathcal{S}}(\lambda, z|Y) \propto f(Y|\lambda)\pi_{\mathcal{S}_1}(\lambda)\frac{P_{\mathcal{S}_2}(\lambda|z)\pi_{\mathcal{S}_2}(z)}{P_{\mathcal{S}_2}(\lambda)}$$

THE DIFFERENCE (COMPATIBILITY I)

$$P_{\mathcal{J}}(\lambda, z|Y) \propto f(Y|\lambda, z)\pi_{\mathcal{J}}(\lambda, z)$$

$$P_{\mathcal{S}}(\lambda, z|Y) \propto f(Y|\lambda)\pi_{\mathcal{S}_1}(\lambda)\frac{P_{\mathcal{S}_2}(\lambda|z)\pi_{\mathcal{S}_2}(z)}{P_{\mathcal{S}_2}(\lambda)}$$

- *S*=*J*? Sufficient condition:
  - $f(Y|\lambda) = f(Y|\lambda, z)$
  - $P_{\mathcal{S}_2}(\lambda|z)\pi_{\mathcal{S}_2}(z) = \pi_{\mathcal{J}}(\lambda, z)$, which implies $\pi_{\mathcal{S}_1}(\lambda) = \int \pi_{\mathcal{J}}(\lambda, z)dz = \int P_{\mathcal{S}_2}(\lambda|z)\pi_{\mathcal{S}_2}(z)dz$
- How far is two-step approach from the ideal output?
  - Note that $z \perp\!\!\!\perp Y|\lambda$ for both *S* and $\mathcal{J}$
  - For λ, inference is equivalent

$$P_{\mathcal{J}}(\lambda|Y) = \int P_{\mathcal{J}}(\lambda, z|Y)dz \propto f(Y|\lambda)\int \pi_{\mathcal{J}}(\lambda, z)dz$$

$$P_{\mathcal{S}}(\lambda|Y) = \int P_{\mathcal{S}}(\lambda, z|Y)dz \propto f(Y|\lambda)\pi_{\mathcal{S}_1}(\lambda)$$

THE DIFFERENCE (COMPATIBILITY II)

- For *z*, calculate K-L divergence:

$$
\begin{aligned}
& D_{KL}(P_{\mathcal{J}}(\lambda, z|Y) \| P_{\mathcal{S}}(\lambda, z|Y)) \\
&= \int\int P_{\mathcal{J}}(\lambda, z|Y) \left(\log P_{\mathcal{J}}(\lambda, z|Y) - \log P_{\mathcal{S}}(\lambda, z|Y)\right) d\lambda dz \\
&= \int\int P_{\mathcal{J}}(\lambda|Y) P_{\mathcal{J}}(z|\lambda, Y) \left(\log P_{\mathcal{J}}(z|\lambda, Y) - \log P_{\mathcal{S}}(z|\lambda, Y)\right) d\lambda dz \\
&= \int P_{\mathcal{J}}(\lambda|Y) D_{KL}(P_{\mathcal{J}}(z|\lambda) \| P_{\mathcal{S}}(z|\lambda)) d\lambda
\end{aligned}
$$

- so posterior divergence is bounded by prior divergence.

ROADMAP

- Pre-processing using LIRA
- Establish model for pixel assignments
- Model compatibility
- **Draw assignments via Gibbs Sampler**
- Results
- Future directions

STEP 1 – LIKELIHOOD PARAMETERS

- Drawn through seeded Gibbs sampler via JAGS
- Priors:

$$\tau_-, \tau_+ \sim N(\mu_0, \sigma^2)$$
$$\sigma^2 \sim \mathrm{Inv} - \chi^2(\nu_0, \omega_0^2)$$

STEP 2 – TEMPERATURE PARAMETER

- Drawn through Metropolis Hastings
- Prior:

$$\beta \sim \mathrm{Gamma}(a_\beta, b_\beta)$$

STEP 3- ASSIGNMENTS

- 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|β 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:

$$p(z, d|\tilde{\lambda}, \tau_0, \tau_1, \sigma^2, \beta) \propto \prod_{m=1}^{M} g_m(z_m, d_m|\beta) \prod_{ij} f(\tilde{\lambda}_{ij}|z, \tau_0, \tau_1, \sigma^2)$$

  - Marginal distribution of *z* is equal to our posterior.

    $\sum_d p(z, d|\tilde{\lambda}, \tau_0, \tau_1, \sigma^2, \beta) = p(z|\tilde{\lambda}, \tau_0, \tau_1, \sigma^2, \beta)$
  - Conditional distributions are easy to sample from.

    $p(z|d, \beta, -) \qquad p(d|z, \beta, -)$

COUPLING SPINS TO BONDS

- Bonds can be disconnected (0) or connected (1).

$$d = \{0, 1\}$$

COUPLING SPINS TO BONDS

- Bonds can be disconnected (0) or connected (1).

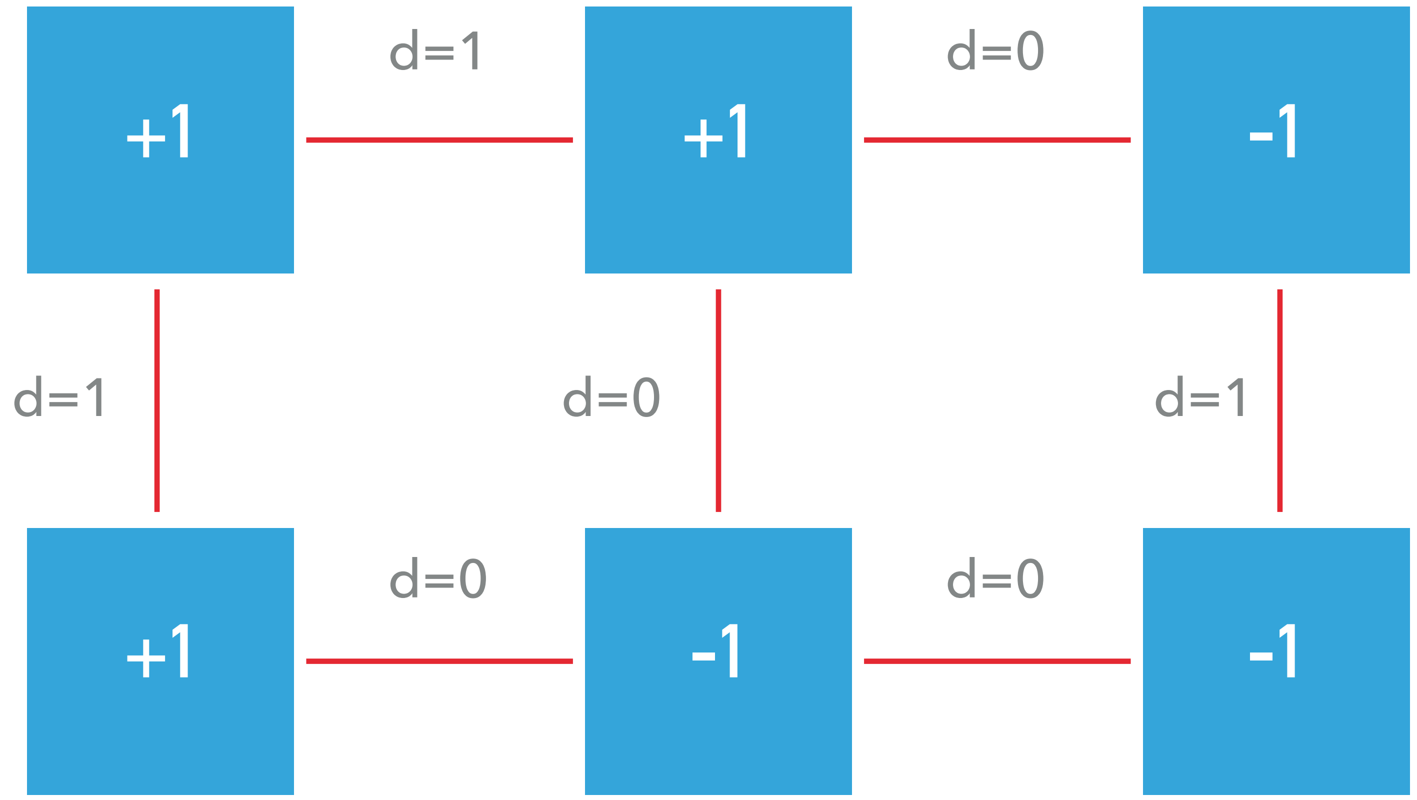
$$d = \{0, 1\}$$

+1	+1	-1
+1	-1	-1

COUPLING SPINS TO BONDS

- Bonds can be disconnected (0) or connected (1).

$$d = \{0, 1\}$$



COUPLING SPINS TO BONDS

- Factor coupling bonds and spins is:

$$g_m(z_m, d_m) = \left\{ \begin{array}{c|cc|cc} & \multicolumn{2}{c}{d_m = 0} & \multicolumn{2}{c}{d_m = 1} \\ & z_{i'j'} = 0 & z_{i'j'} = 1 & z_{i'j'} = 0 & z_{i'j'} = 1 \\ z_{ij} = 0 & e^{-\beta} & e^{-\beta} & e^{\beta} - e^{-\beta} & 0 \\ z_{ij} = 1 & e^{-\beta} & e^{-\beta} & 0 & e^{\beta} - e^{-\beta} \end{array} \right.$$

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

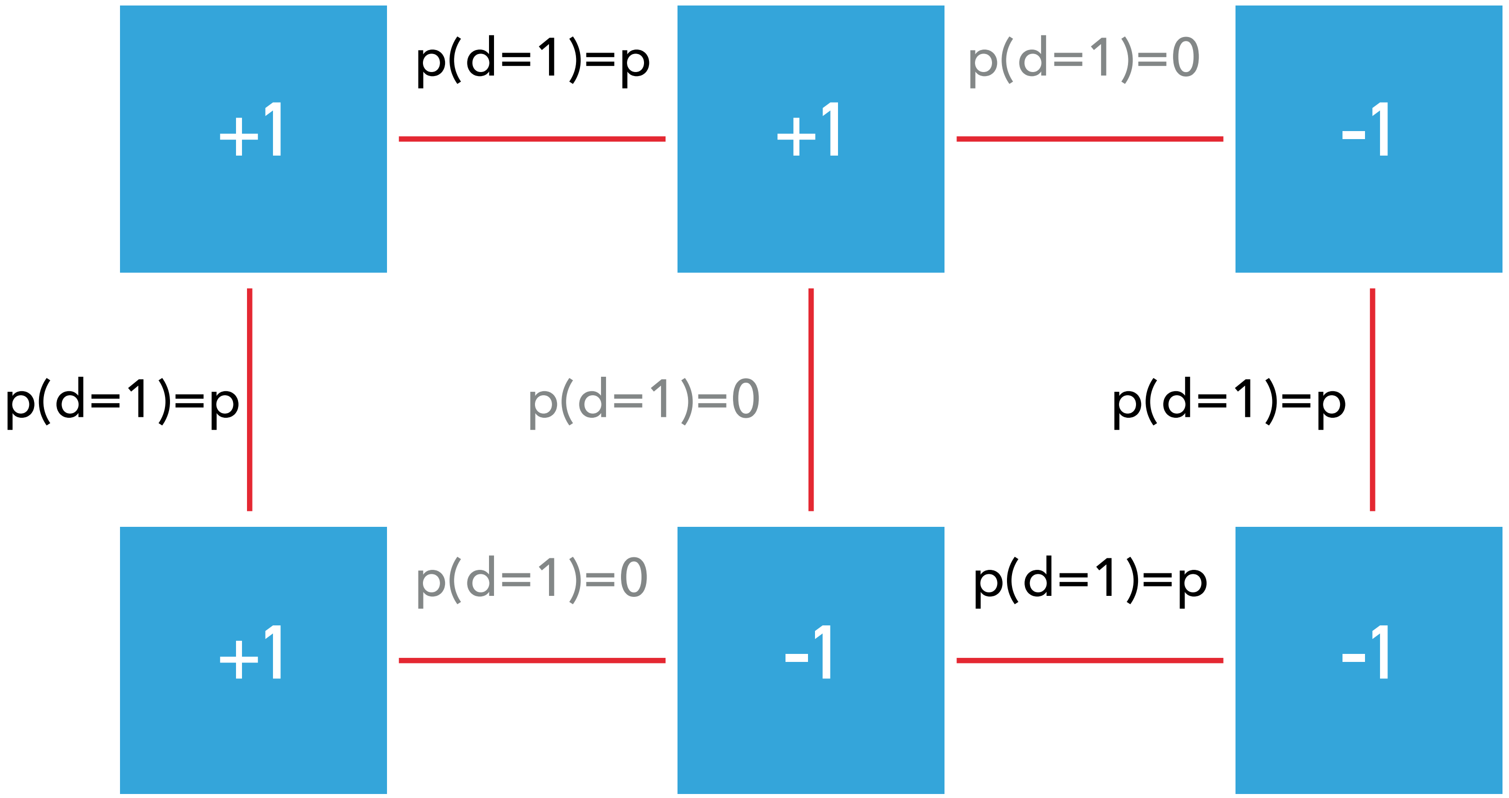
$$\tilde{g}_m(z_m, d_m) = \left\{ \begin{array}{c|cc|cc} & \multicolumn{2}{c}{d_m = 0} & \multicolumn{2}{c}{d_m = 1} \\ & z_{i'j'} = 0 & z_{i'j'} = 1 & z_{i'j'} = 0 & z_{i'j'} = 1 \\ z_{ij} = 0 & 1-p & 1-p & p & 0 \\ z_{ij} = 1 & 1-p & 1-p & 0 & p \end{array} \right.$$

- Therefore:

$$p(z, d|\beta) \propto \prod_m g_m(z_m, d_m|\beta) \prod_{ij} f(\tilde{\lambda}_{ij}|z_{ij}, b, \tau, \sigma^2) \propto \prod_m \tilde{g}_m(z_m, d_m|\beta) \prod_{ij} f(\tilde{\lambda}_{ij}|z_{ij}, b, \tau, \sigma^2)$$

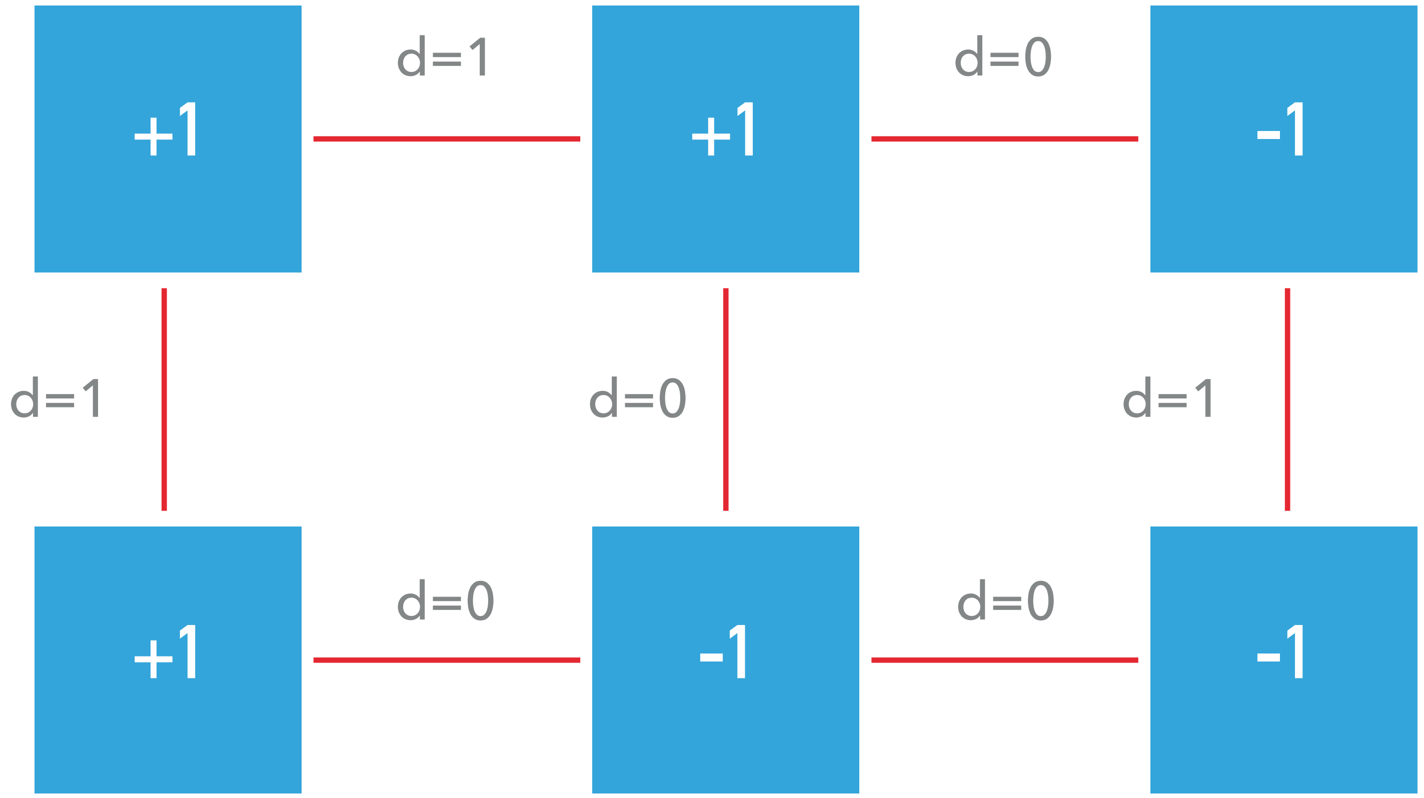
SAMPLING

- 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(-2β)* , and 0 otherwise.



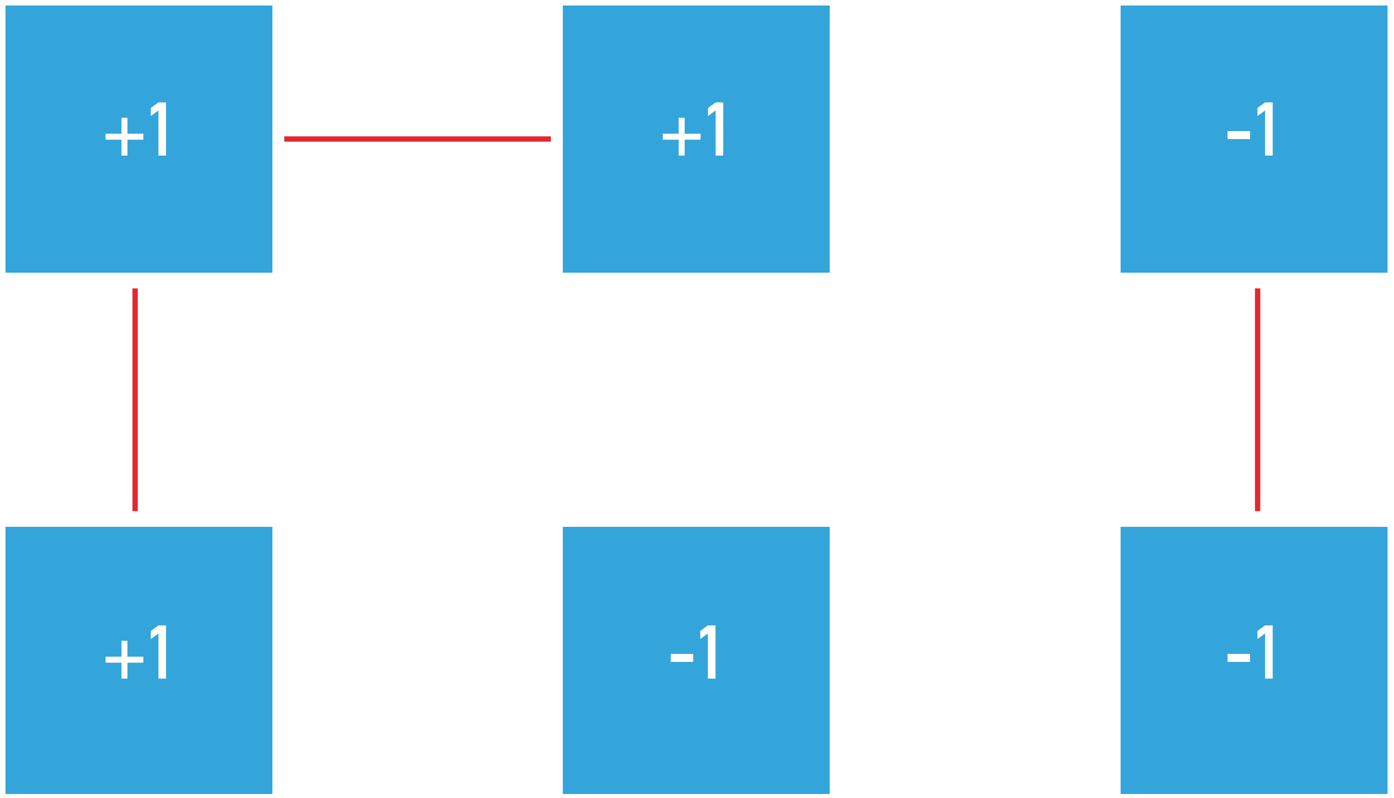
SAMPLING

- 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(-2β)* , and 0 otherwise.



SAMPLING

- 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(-2β)* , and 0 otherwise.



SAMPLING

- 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_+

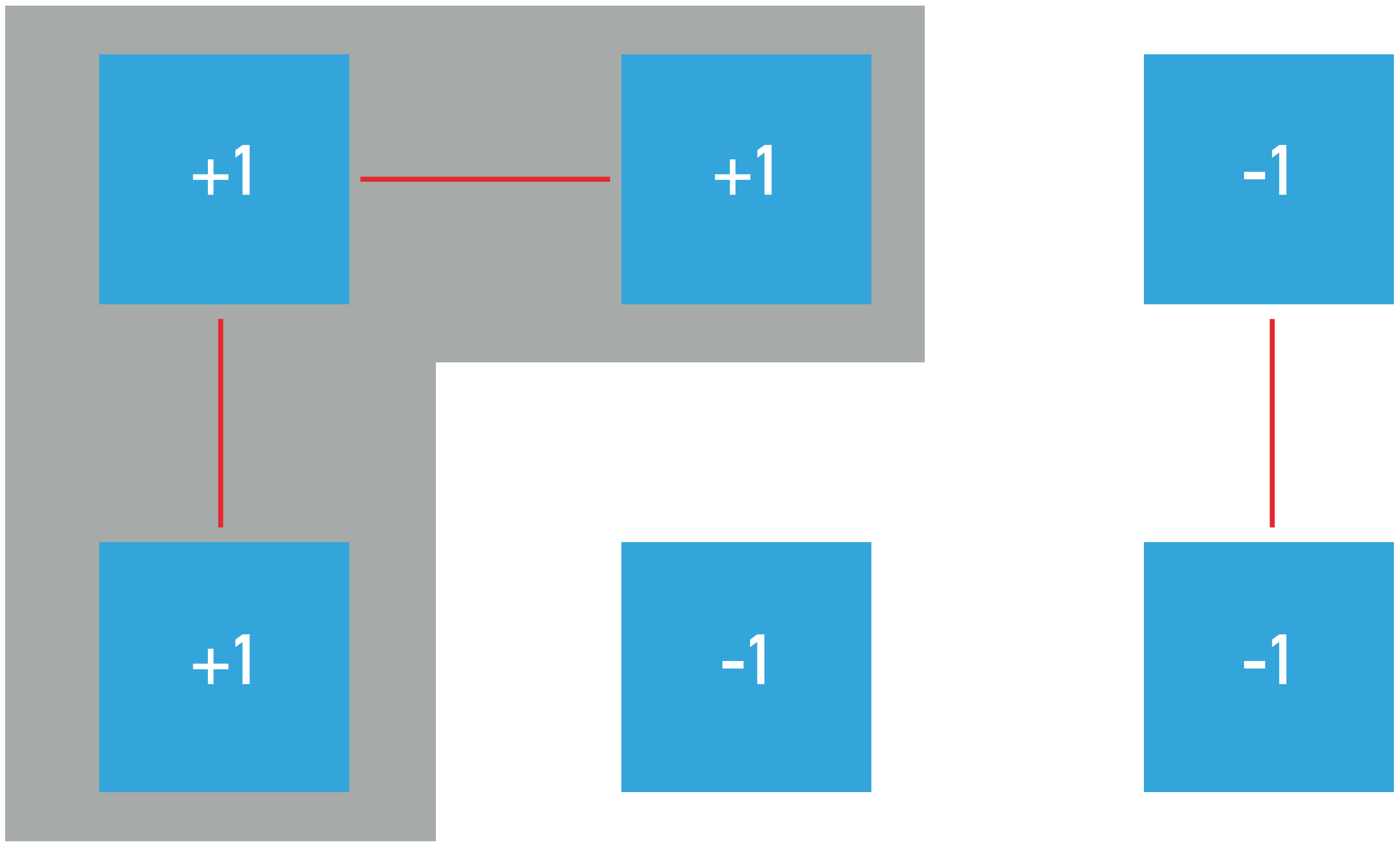
    -1 with probability $p_-=1-p_+$

$$\frac{p_+}{p_-} = \frac{\prod_{ij\in C} f(\tilde{\lambda}_{ij}|, z_{ij} = +1, \tau_1, \tau_2, \sigma^2)}{\prod_{ij\in C} f(\tilde{\lambda}_{ij}|z_{ij} = -1, \tau_1, \tau_2, \sigma^2)}$$

SAMPLING

- Sample from $p(z|d, \beta)$

p(z=+1)=p_+

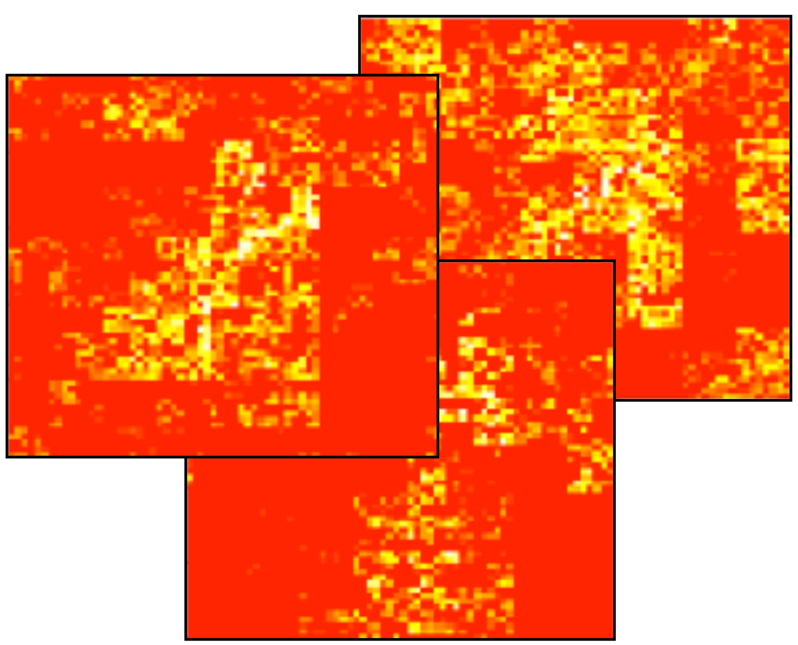


ROADMAP

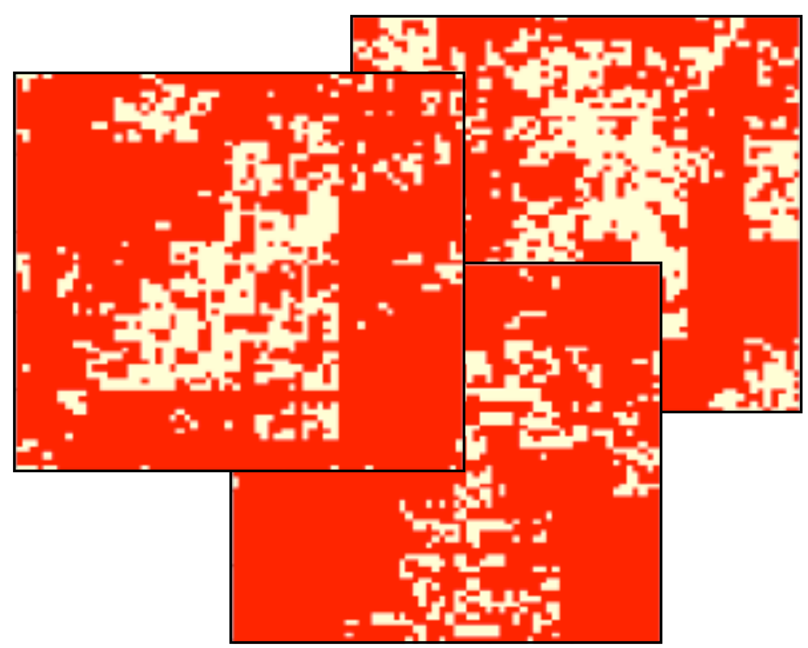
- Pre-processing using LIRA
- Establish model for pixel assignments
- Model compatibility
- Draw assignments via Gibbs Sampler
- **Results**
- Future directions

ISING-LIRA ITERATIONS

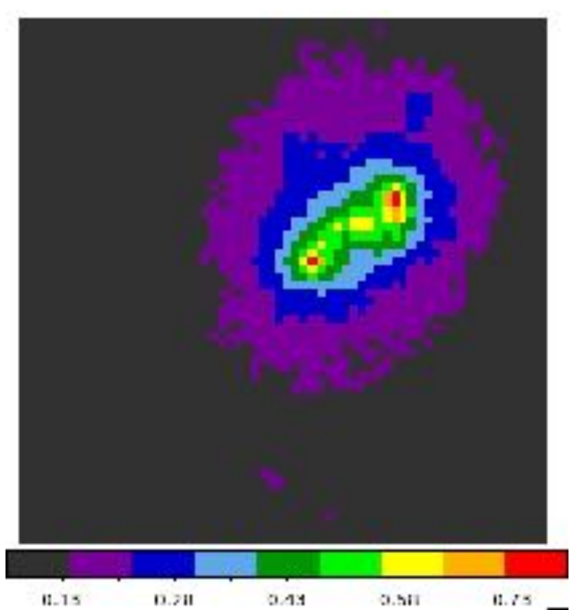
1. Get many posterior draws from LIRA
2. Apply new method to each LIRA draw
3. Average across LIRA-Ising iterations to get probability map.



1



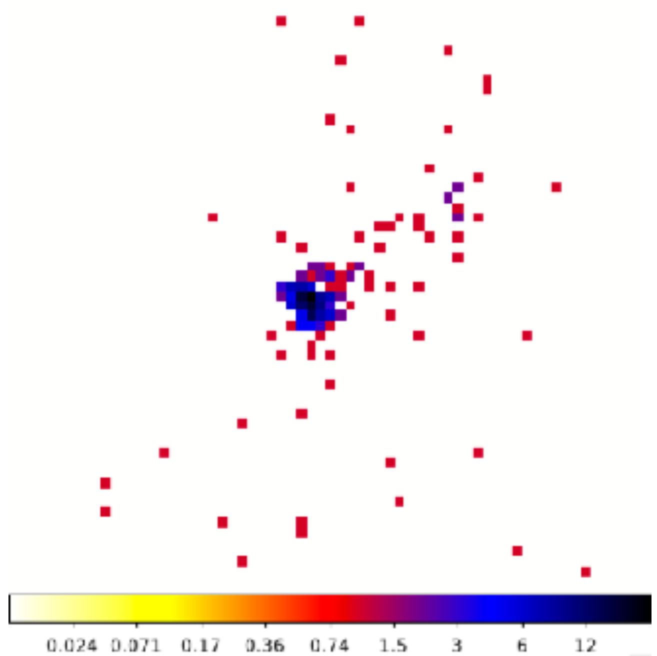
2

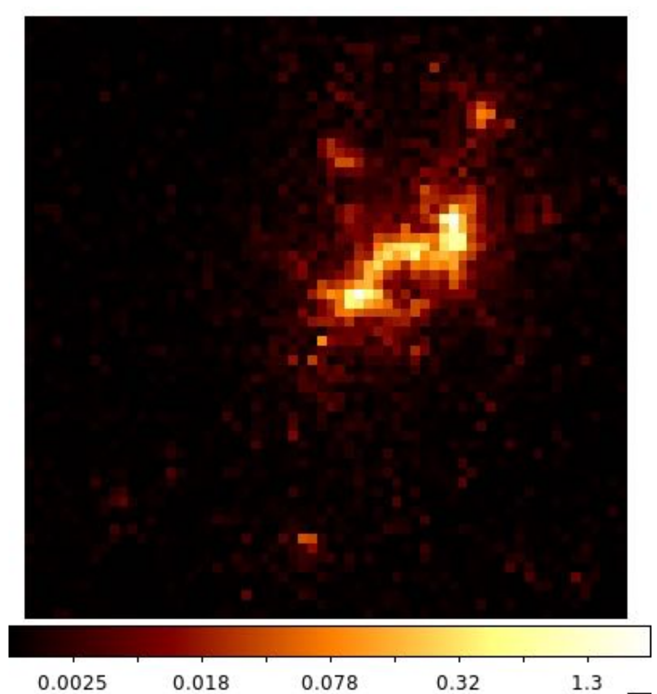


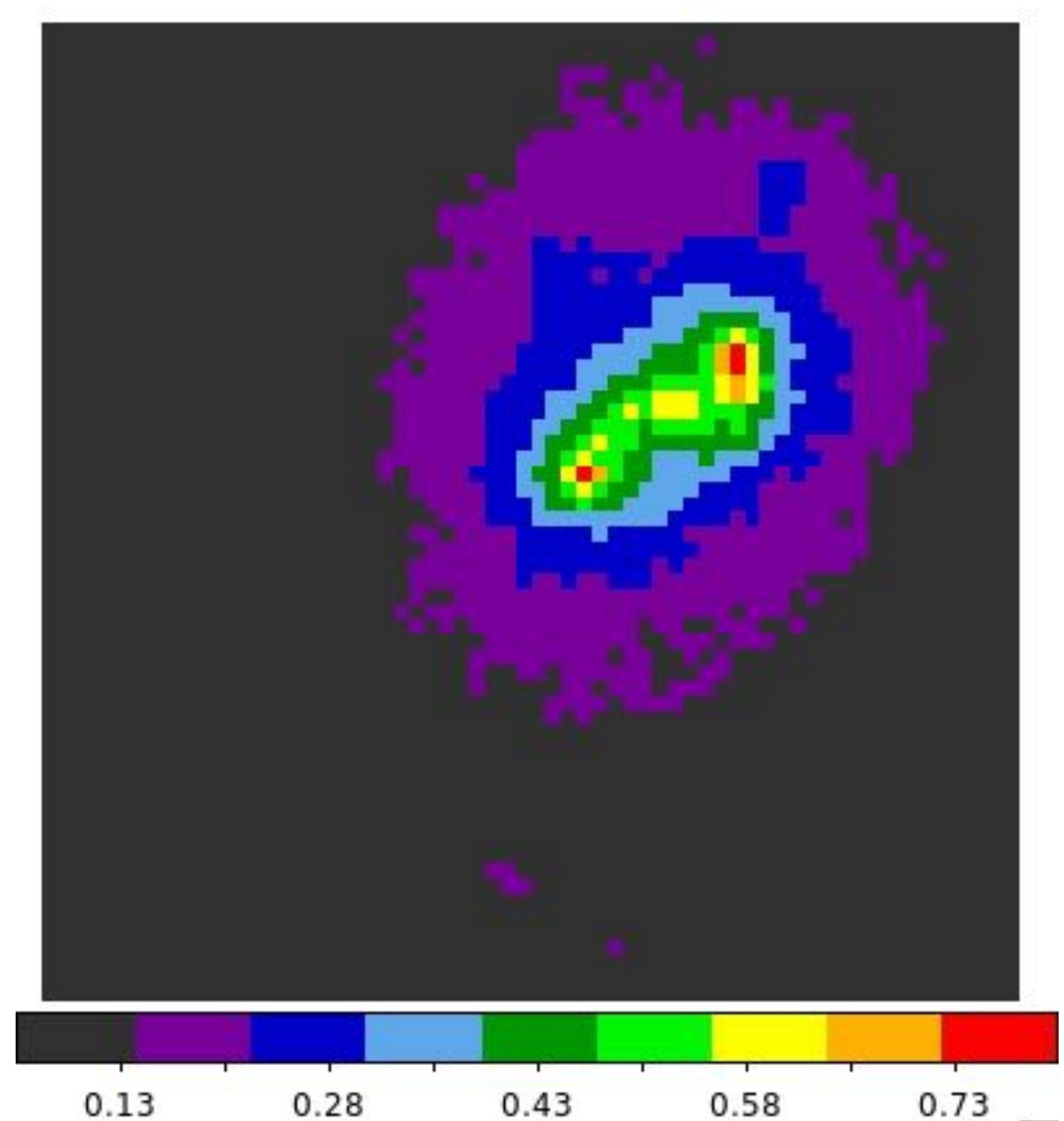
3

PROBABILITY MAP

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





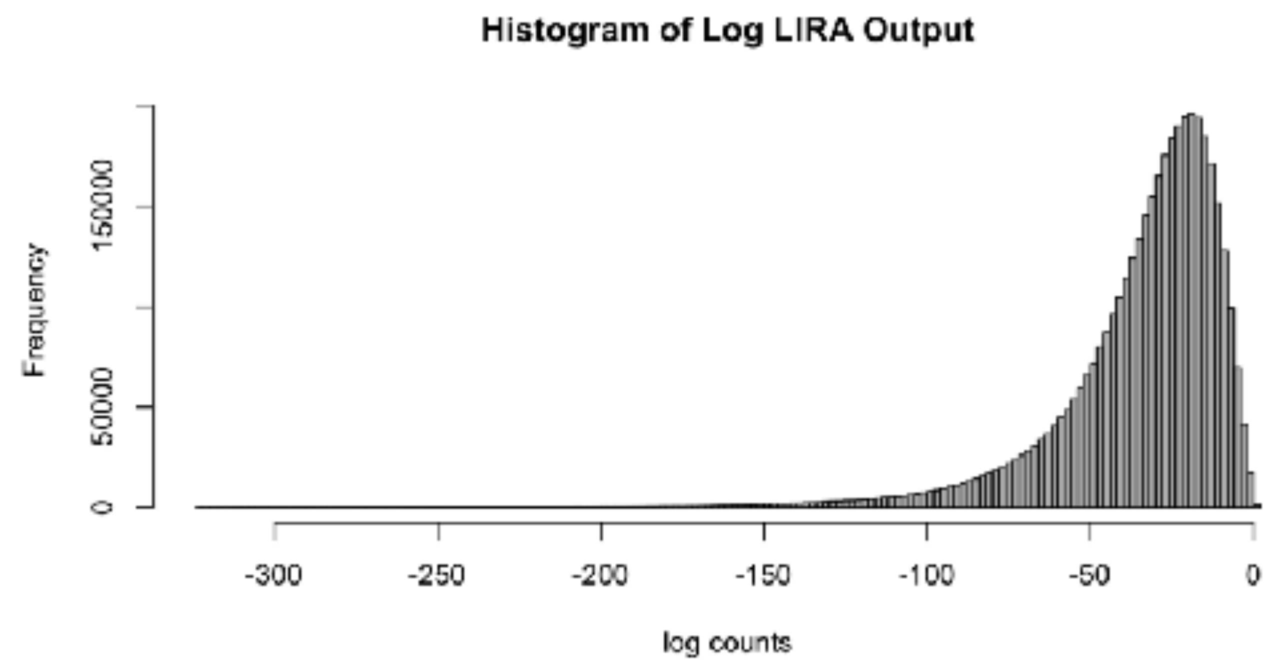


ROADMAP

- Pre-processing using LIRA
- Establish model for pixel assignments
- Model compatibility
- Draw assignments via Gibbs Sampler
- Results
- **Future work**

"UNDERFLOW" ISSUE

- Majority of LIRA output are very small counts ranging as low as $<10^{-300}$.
- This doesn't make sense physically.
- Our model is sensitive to lower ranges induced by multi-scale structure of LIRA algorithm, which overwhelms signal from jet.
- One solution: Machine limitation is 10^{-16} and anything smaller is induced by underflow error.
- After looking at distribution of concatenated pixels from all 1000 iterations, it is clear this is not an underflow issue.
- Prior for LIRA intensity is Gamma weighted for very, very small values.



DIFFERENT VARIANCES

- Pixels in the ROI should have a higher variance than the background since we expect all background pixels to be close to zero.

$$\tilde{\lambda}_{ij} | z, \tau_0, \tau_1, \sigma_0^2, \sigma_1^2 \sim \text{Log-Normal}(\tau_0, \sigma_0^2)\mathbb{I}_{z_{ij}=-1} + \text{Log-Normal}(\tau_1, \sigma_1^2)\mathbb{I}_{z_{ij}=+1}$$

HURDLE MODEL

- Computational limits only produce reliable estimates of the LIRA posterior on the order of 10^-16
- About roughly 70% of the data lies below this
- Hurdle model accounts for this truncation:

$$f(\lambda_{ij}|z_{ij}=+1,\tau_0,\tau_1,\sigma_0^2,\sigma_1^2)=\begin{cases}\Phi(\log x;\tau_1,\sigma_1^2) & ;\ \lambda_{ij}=x\\ g(\lambda_{ij}|\tau_1,\sigma_1^2) & ;\ \lambda_{ij}>x\end{cases}$$

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:

$$z_{ij} = \{0, 1, 2, 3, \dots\}$$

REFERENCES

- McKeough et al., *Detecting Relativistic X-ray Jets in High-Redshift Quasars,* The Astrophysical Journal (2016)
- Stein et al., *Detecting Unspecified Structure in Low-Count Images*, The Astrophysical Journal (2017)
- Connors & van Dyk, *How To Win With Non-Gaussian Data: Poisson Goodness-of-Fit*, SCMA IV (2007)
- Esch et al. , *An Image Restoration Technique with Error Estimates*, The Astrophysical Journal (2004)
- Beale, *Exact Distribution of Energy in the Two-Dimensional Ising Model,* Physical Review Letters (1996)
- Swendsen & Wang, *Nonuniversal Critical Dynamics in Monte Carlo Simulations,* Physical Review Letters (1987)

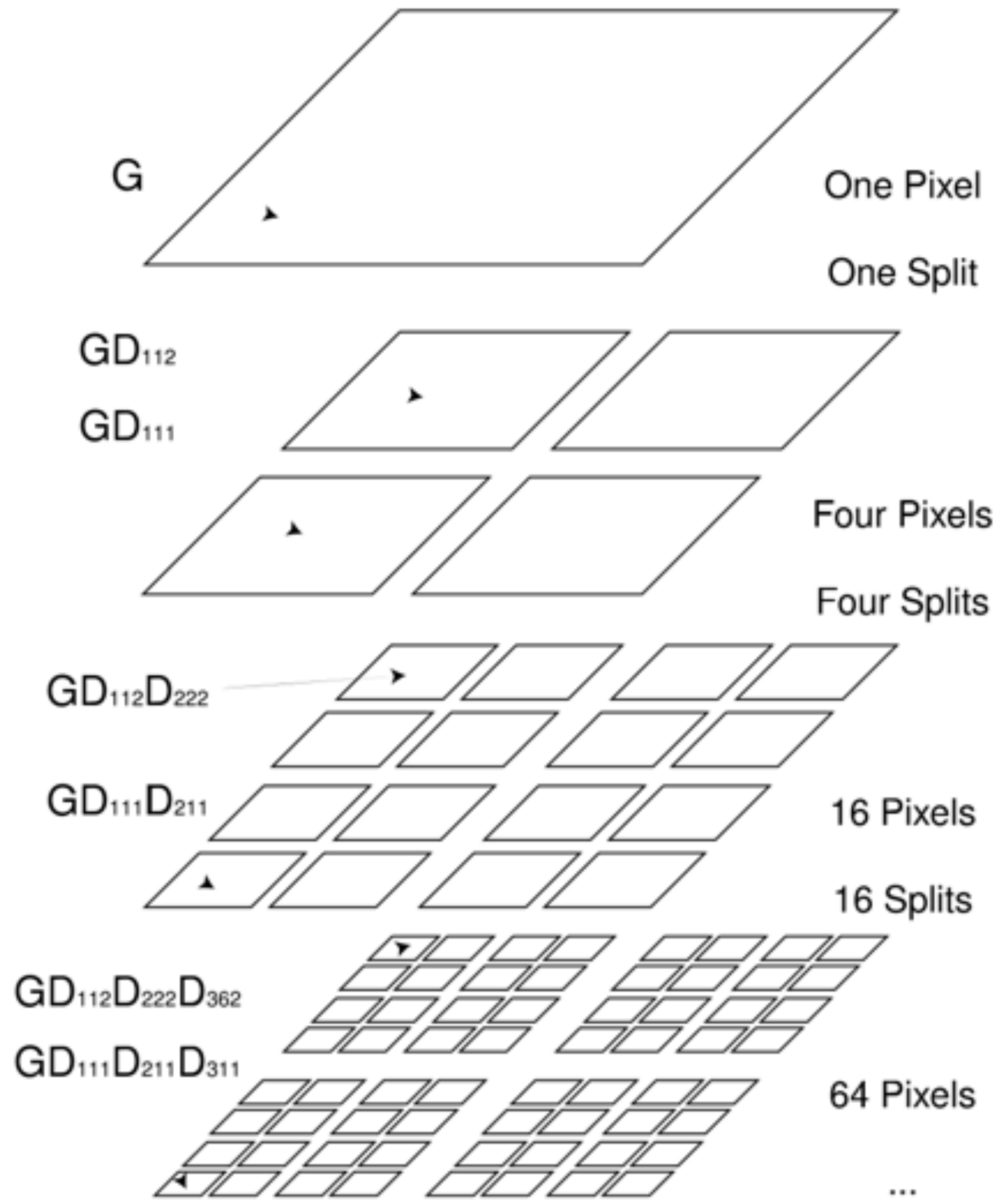
LIRA

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



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*:

$$Y_j|\Lambda, \Lambda^B \overset{d}{\sim} \text{Poisson}\left[\left(\sum_{i \in \mathcal{I}} P_{ij}\Lambda_i\right) + \Lambda_j^B\right]$$

- Suppress background to obtain likelihood:

$$L(\Lambda,\ \Lambda^B|\boldsymbol{Y}) \equiv L(\Lambda|\boldsymbol{Y}) \propto \prod_{j \in \mathcal{I}} p(Y_j|\Lambda).$$

PRIOR

- Prior on total intensity:

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

- Prior on splits:

$$\boldsymbol{D}_{kl} \equiv \{D_{klm},\ m = 1, \ldots, 4\} \overset{\text{d}}{\sim} \text{Dirichlet}(\alpha_k,\ \alpha_k,\ \alpha_k,\ \alpha_k)$$
$$k = 1, \ldots, K, \quad l = 1, \ldots, 4^{k-1}$$

- 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