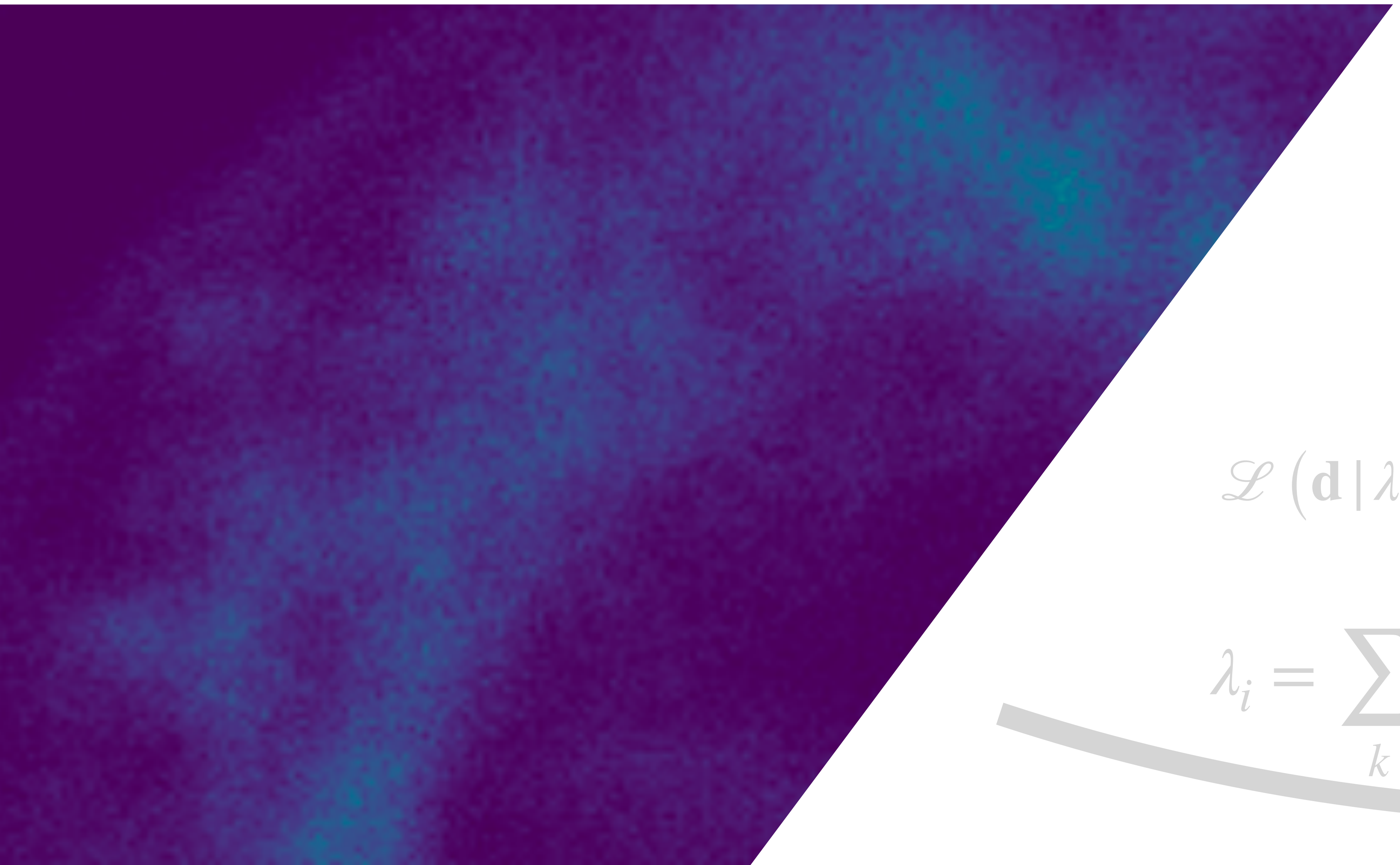
Joint Likelihood Deconvolution of Astronomical Images with Poisson Noise

Axel Donath

CHASC Astrostat Seminar

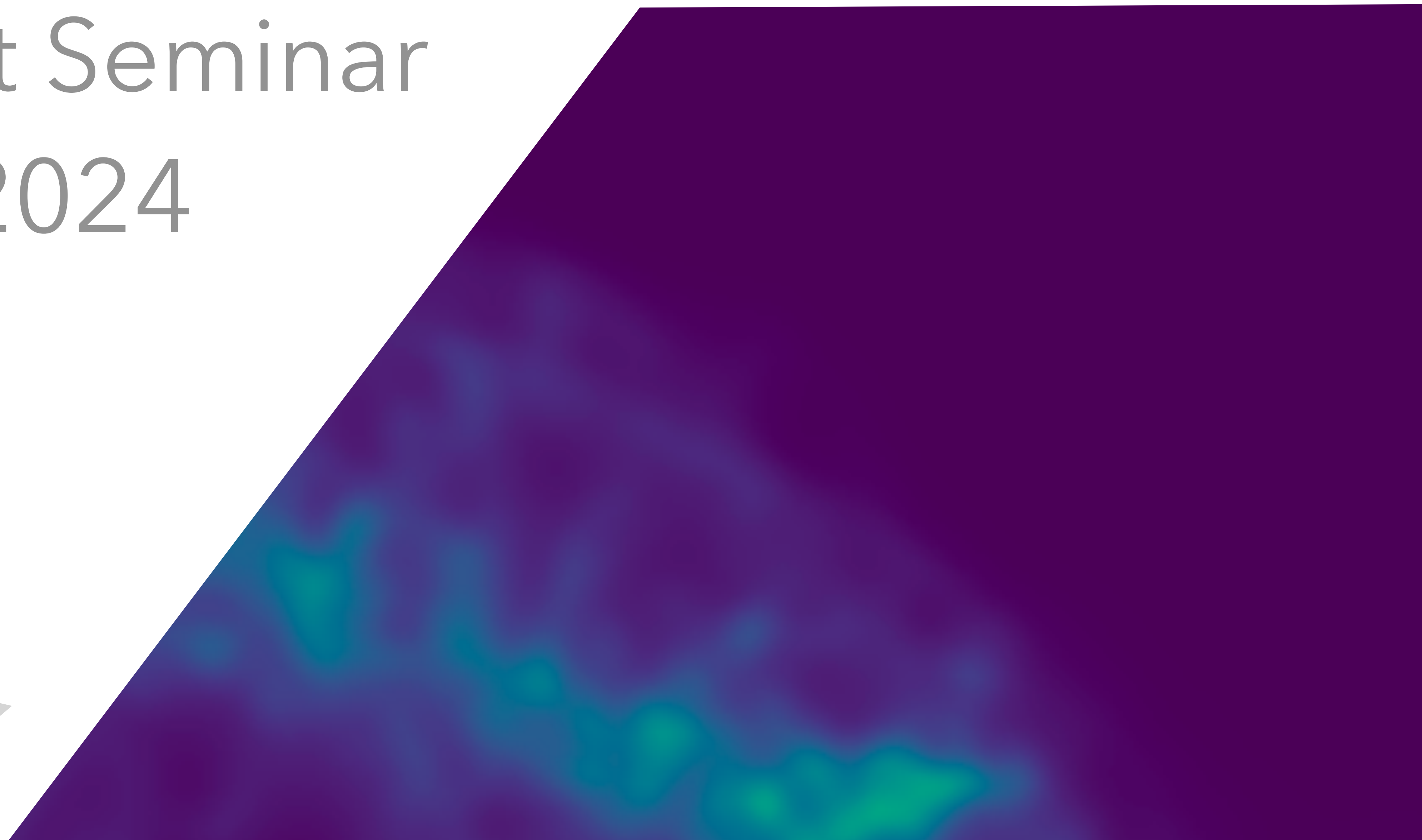
Feb 14th, 2024



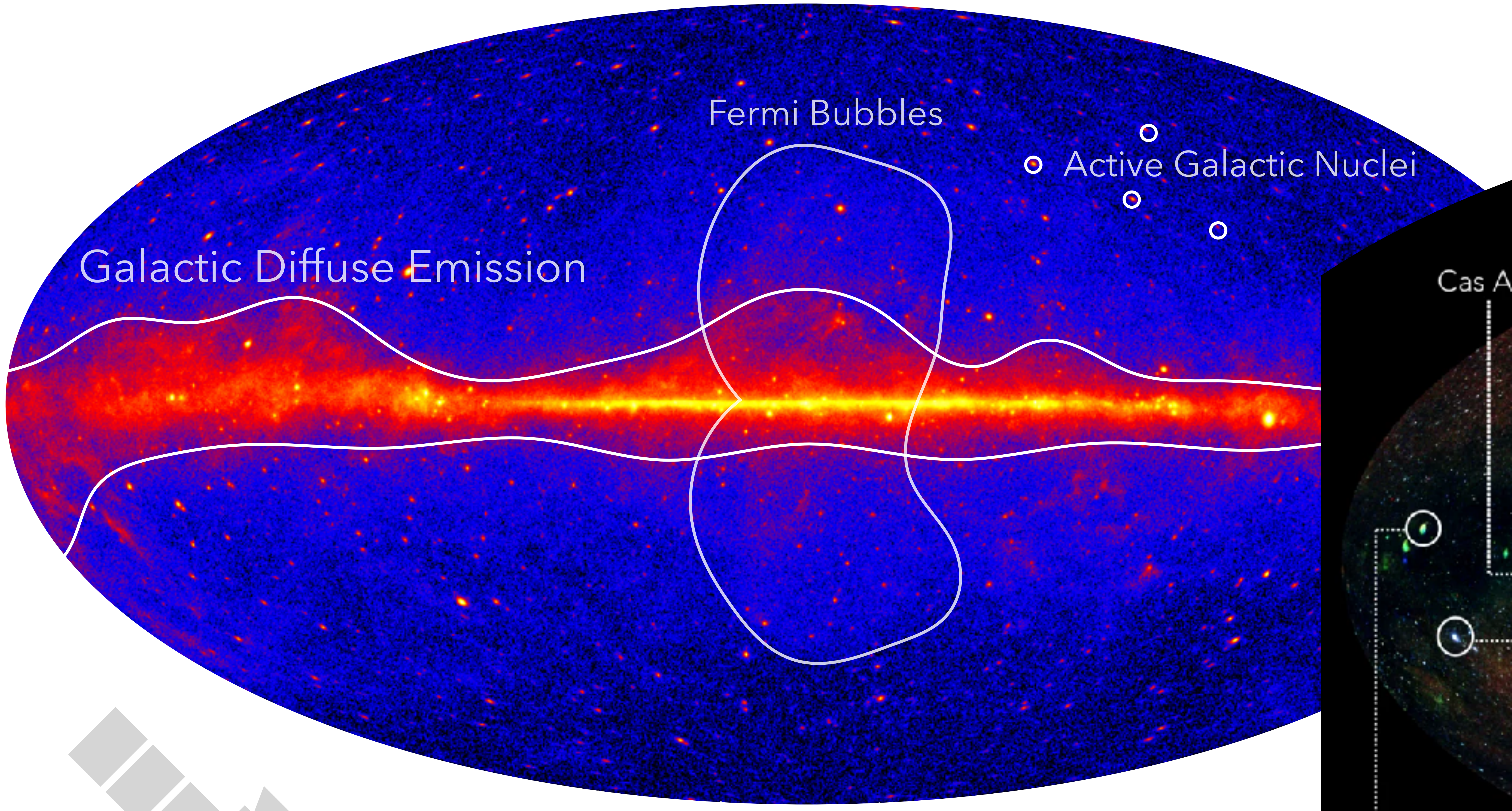
$$\mathscr{L}\left(\mathbf{d} \mid \lambda\right) = \prod_{i}^{N} \frac{\lambda_i^{d_i} \mathrm{e}^{-d_i}}{d_i!}$$

$$\lambda = \mathbf{x} \circledast \mathrm{PSF}$$

$$\lambda_i = \sum_{k} \left(e_i \cdot (x_i + b_i)\right) p_{i-k}$$

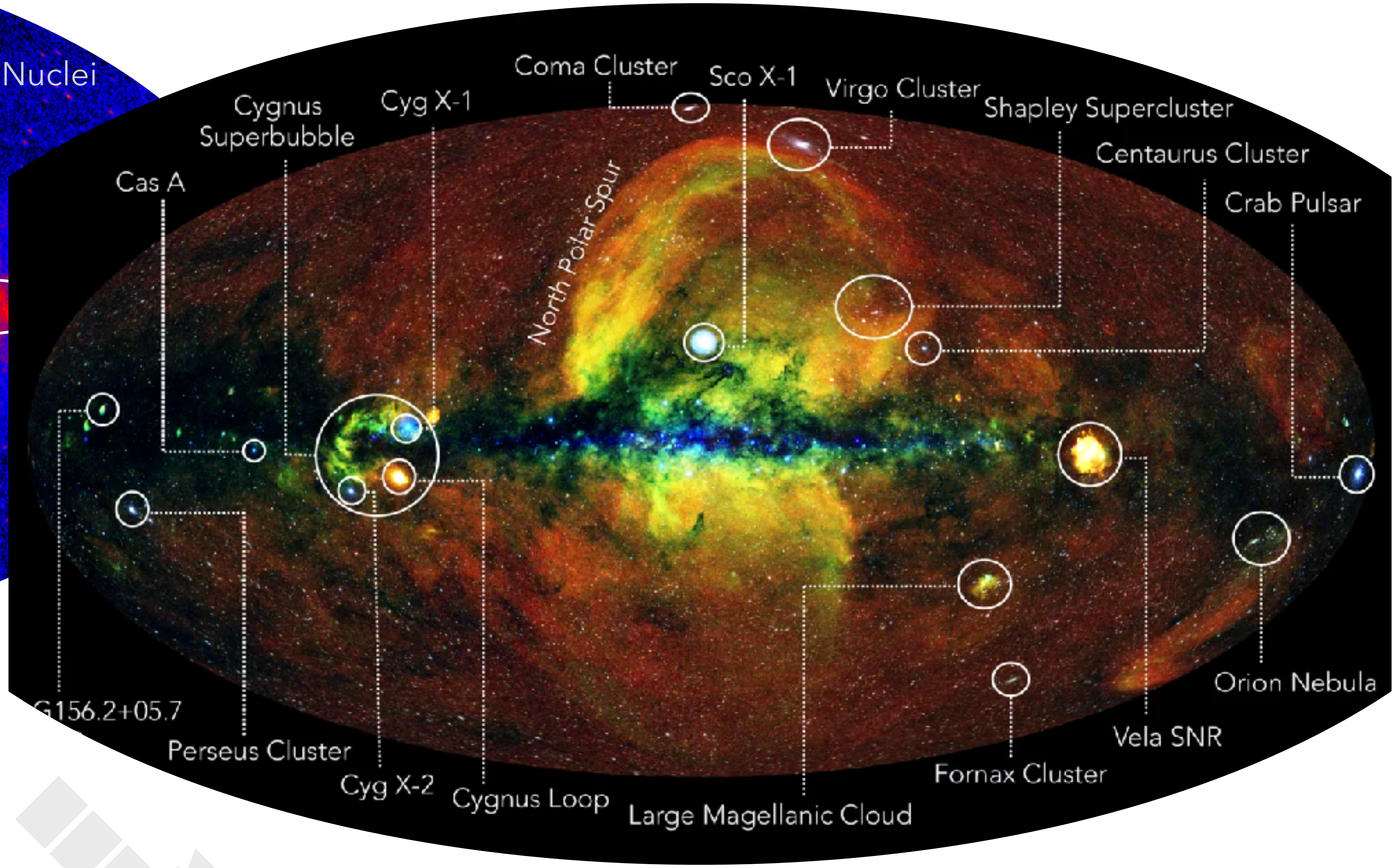


X-Ray / Gamma-Ray Astronomy



Fermi-LAT all-sky counts map > 1 GeV

But also [H.E.S.S.], [HAWC], [CTA]…

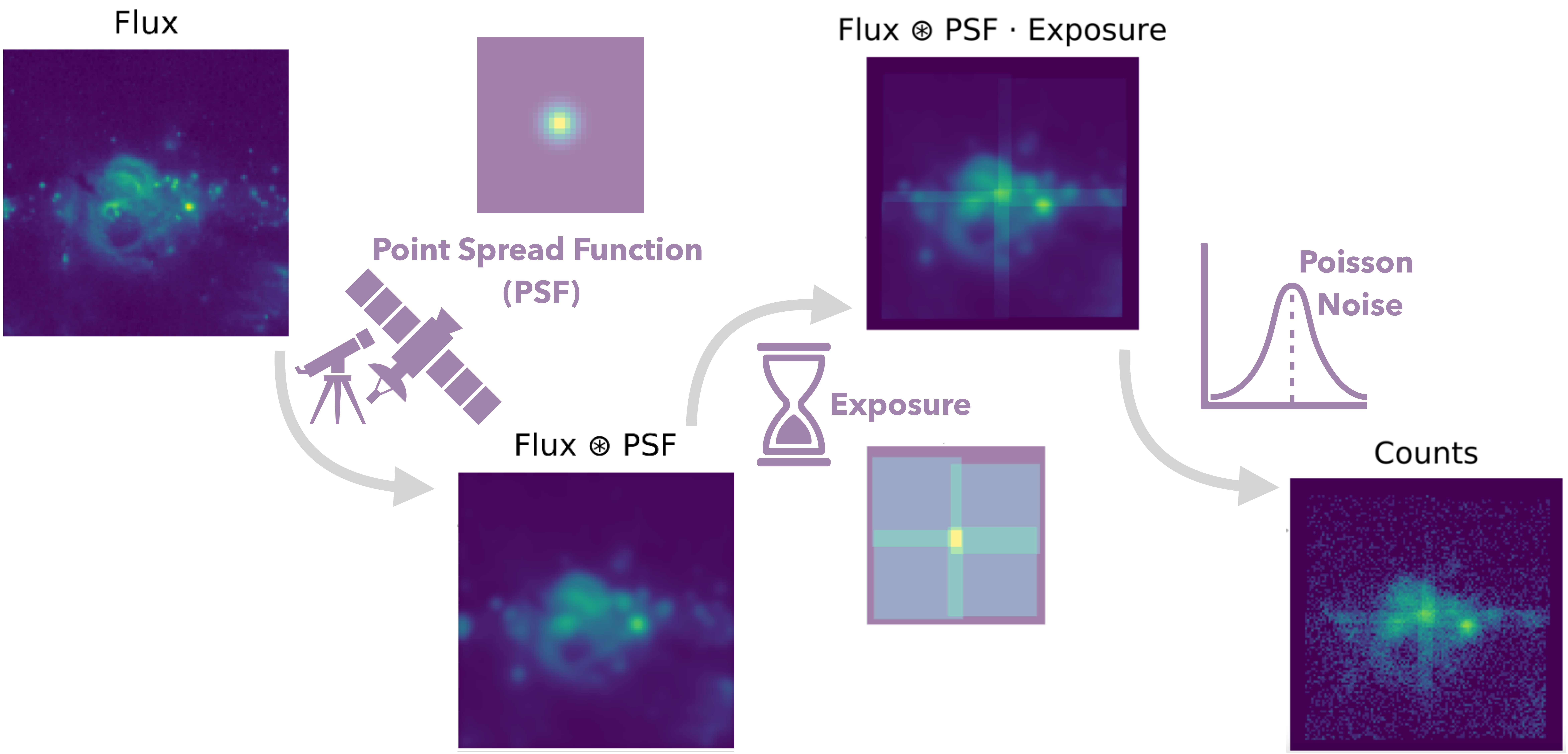


eROSITA all-sky counts map between 0.3 and 2.3 KeV

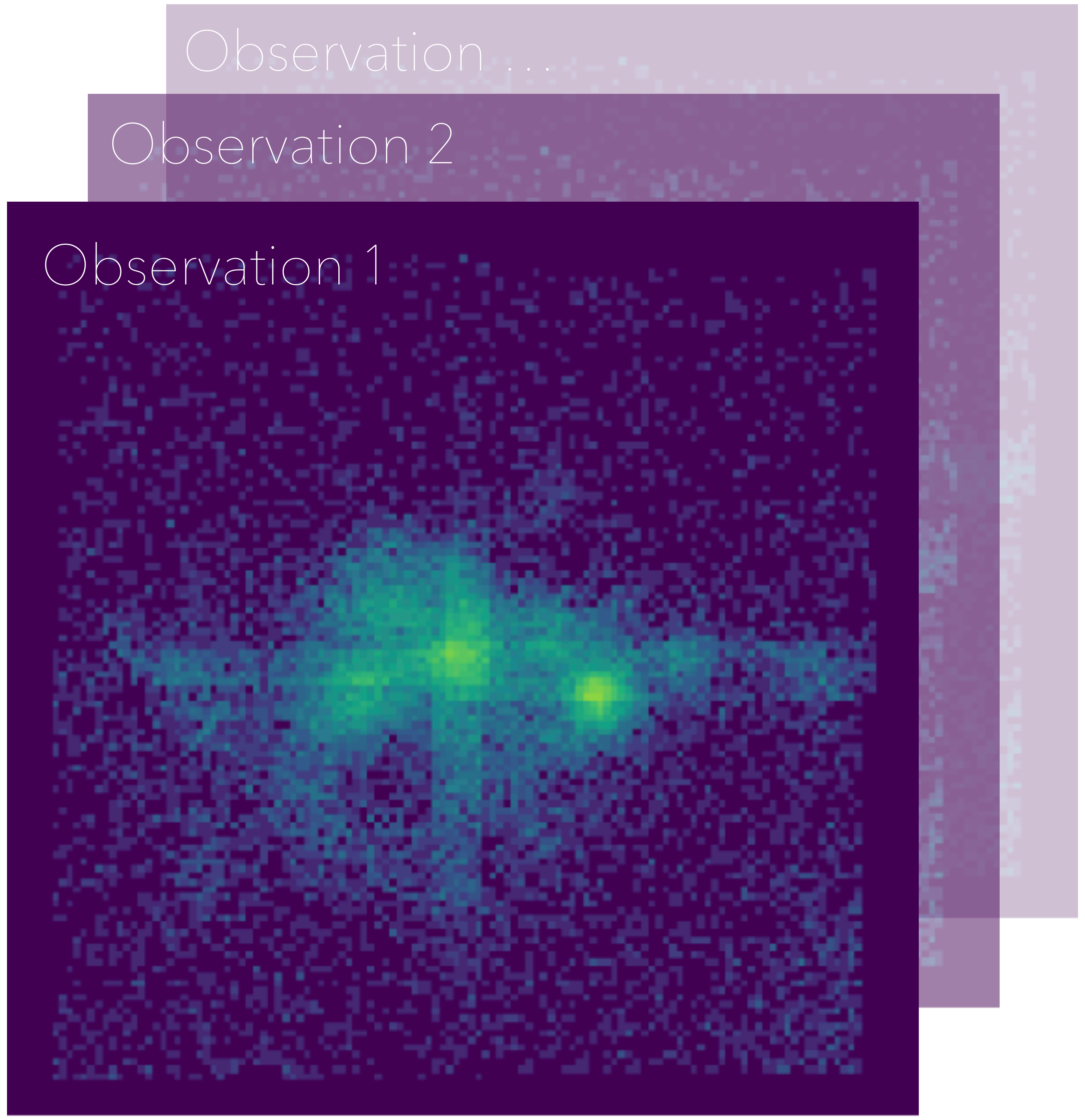
But also [Chandra], [XMM], [Nustar]…

Brightnesses (flux) in the x-ray / gamma-ray range are really low. Observing the universe in this range means counting single photon events and requires long exposure times…

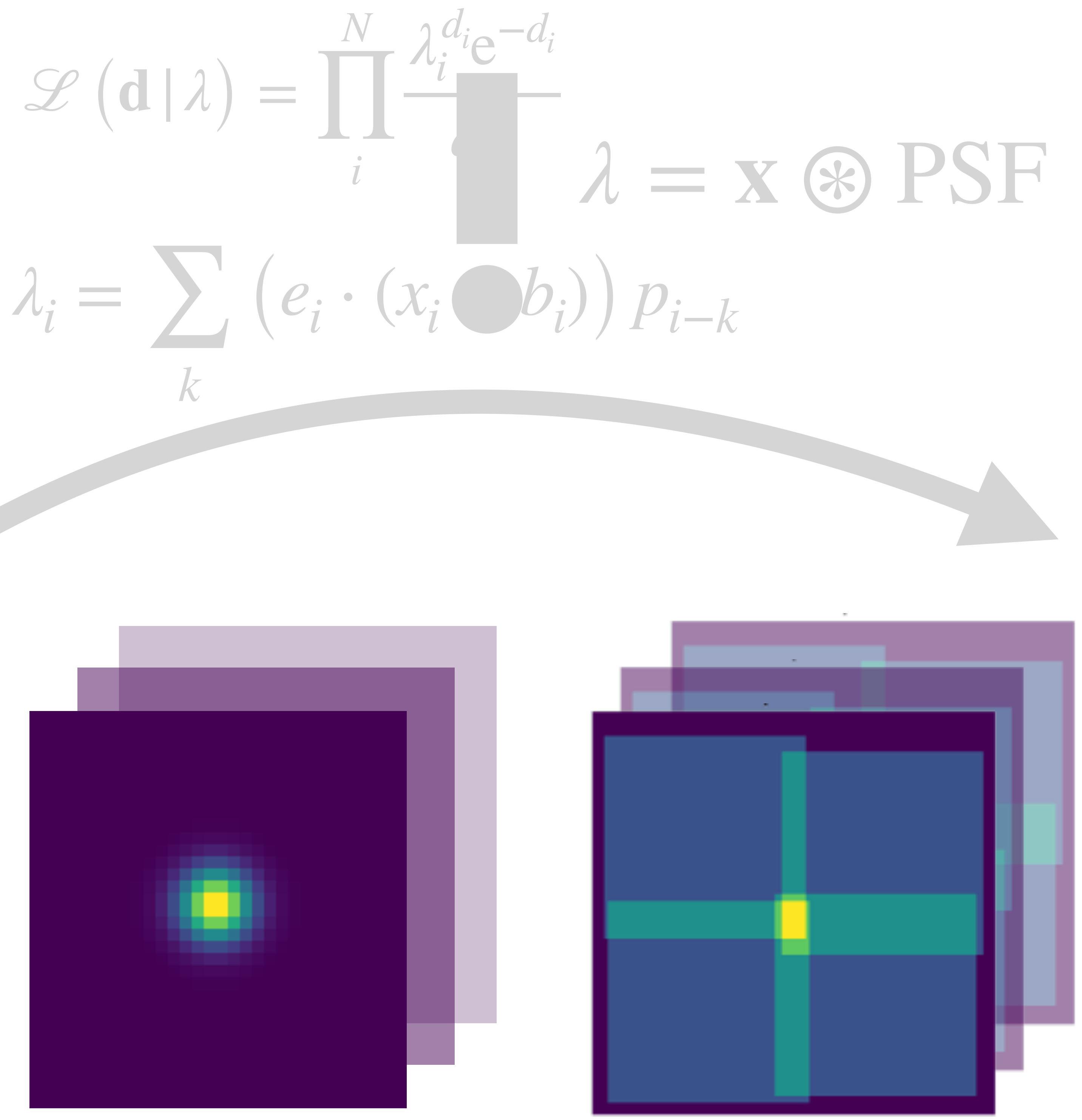
The “low counts” imaging process



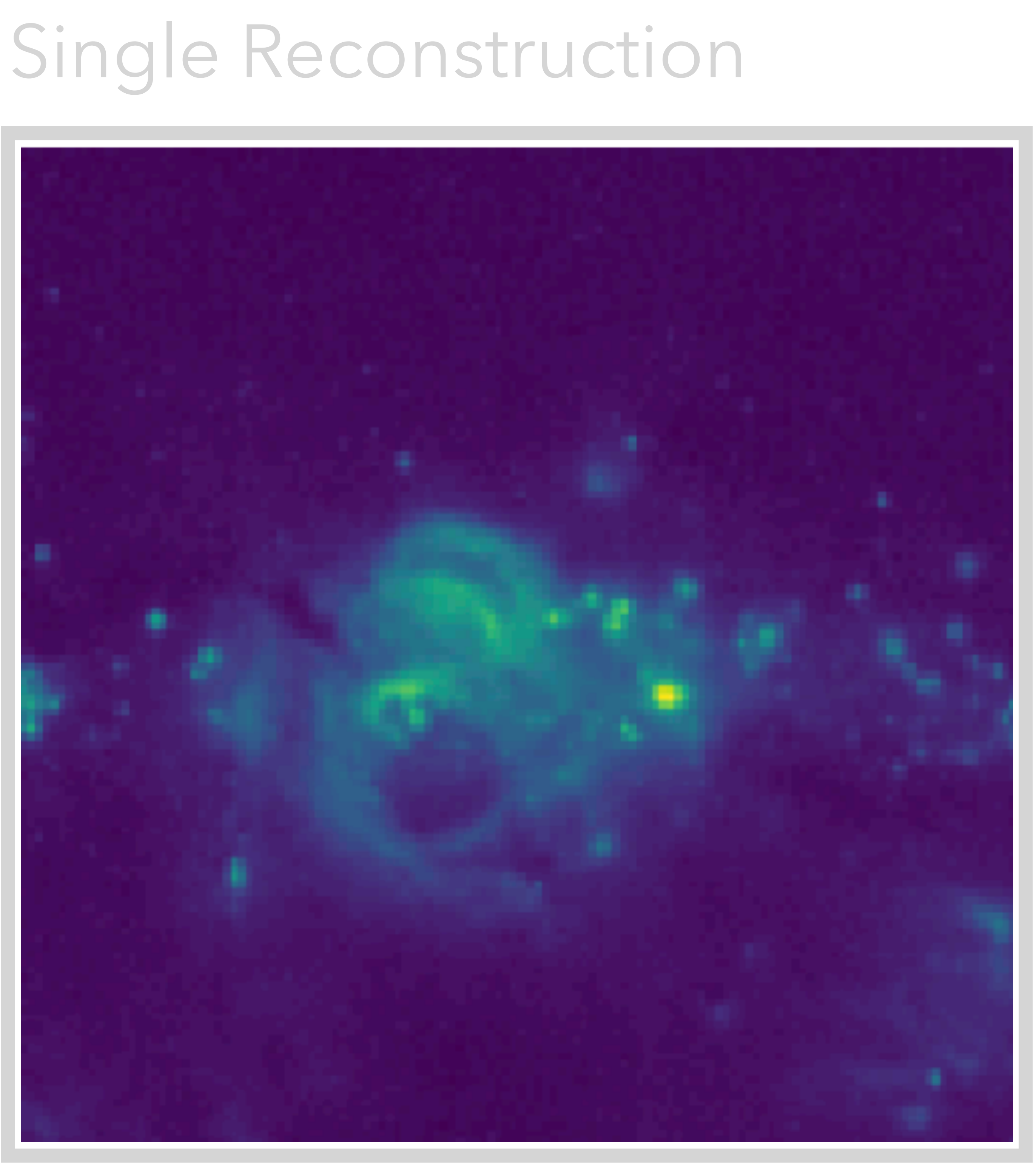
“Unblind” joint deconvolution



Multiple low counts astronomical images from different observations or instruments



“Unblind”: PSF and exposure are known or can be simulated



A “joint” reconstructed flux image using statistical methods

Some math...

For the counts image assume per pixel **Poisson** likelihood:

$$L\left(\mathbf{d} \,|\, \lambda\right) = \prod_{i}^{N} \frac{\lambda_i^{d_i} \mathrm{e}^{-d_i}}{d_i!}$$

Take the **negative log-likelihood**, in Astronomy often call "Cash" statistics

$$\mathscr{C}\left(\mathbf{d} \,|\, \lambda\right) = \sum_{i}^{N} \left(\lambda_i - d_i \log \lambda_i\right)$$

λ are the "model counts"

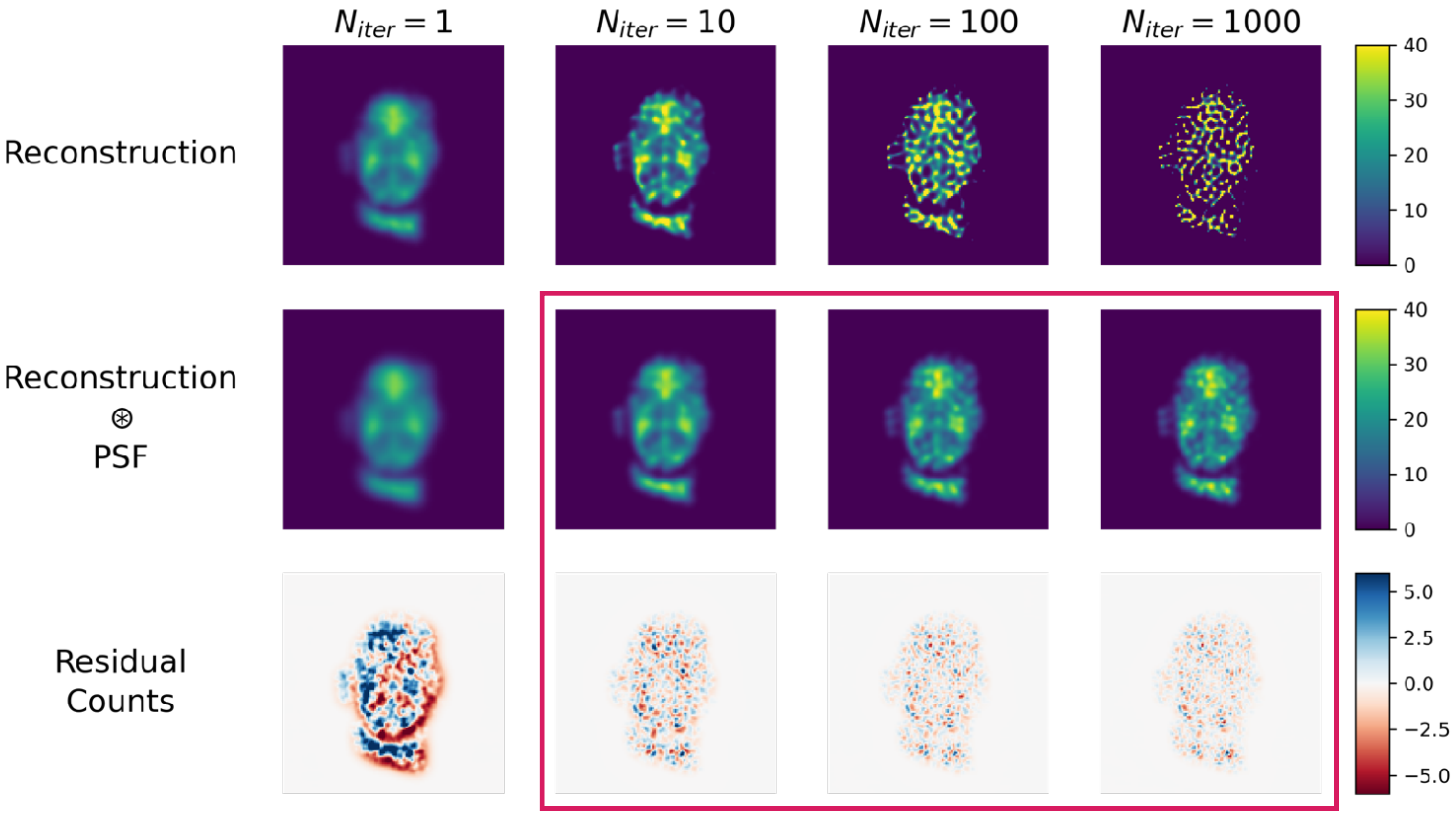
$$\lambda = \mathbf{x} \circledast \mathrm{PSF}$$

x is the reconstructed image we are looking for. Consider each pixel x_i as independent parameter in the model...

Minimize e.g. by **Expectation Maximisation (EM** Proposed by [Richardson 1972] & [Lucy 1974] (RL)

$$\mathbf{x}_{n+1} = \mathbf{x}_n \frac{\mathbf{d}}{\mathbf{x}_n \circledast \mathrm{PSF}} \circledast \mathrm{PSF}^{\mathrm{T}}$$

RL reconstruction quality



All show good residuals and model counts. But reconstructions very different…

Bayes to the "rescue" / priors

Objective function **extended by a log-prior term,** only depending on **x**:

$$\mathscr{L}\left(\mathbf{x} \,|\, \mathbf{D}\right) = \sum_{j=1}^{J} \mathscr{C}\left(\mathbf{D}_j \,|\, \mathbf{x}\right) - \mathscr{P}(\mathbf{x})$$

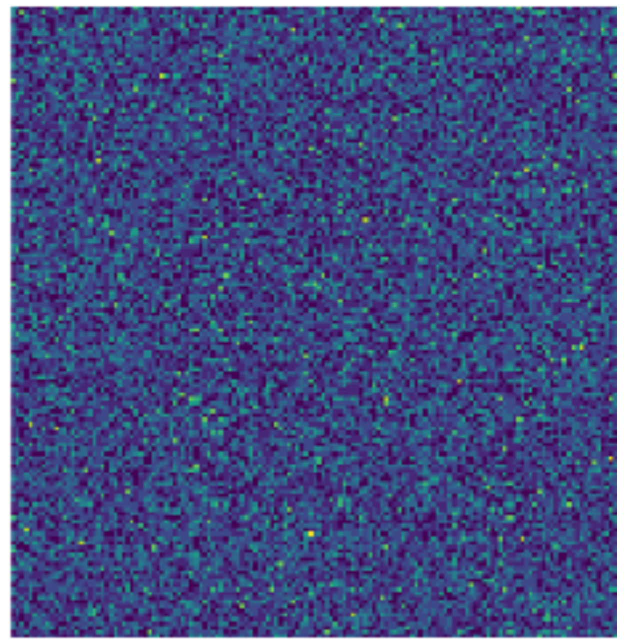
Log-Posterior

Log-Likelihood

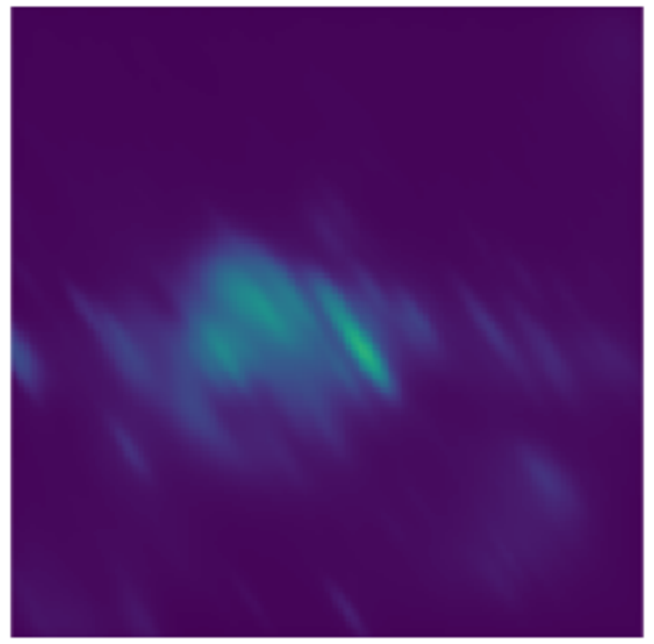
Log-Prior

- Intuitively we **humans have a good understanding of what an actual astronomical image should look like**, because we learned it from seeing many images
- **Learning the full probability distribution of an astronomical image using "deep learning" is hard**, there is not enough training data, we have no ground truth etc. what remains is "transfer learning"…

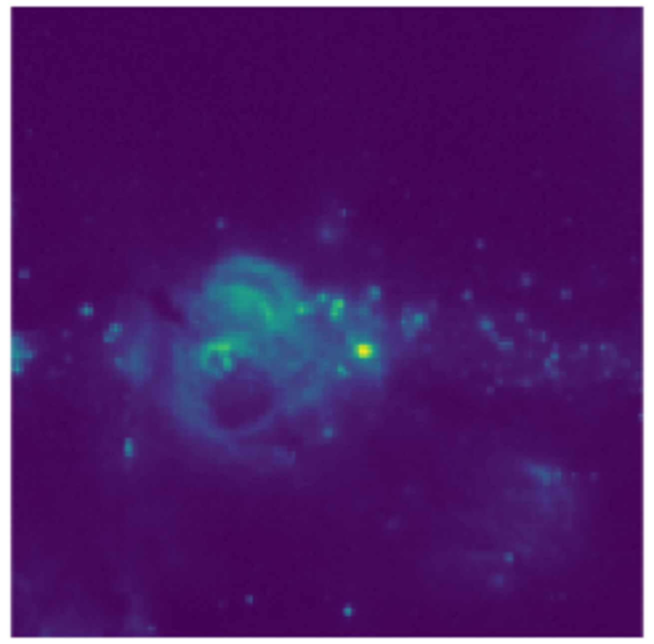
Represents our prior knowledge…



An unlikely image



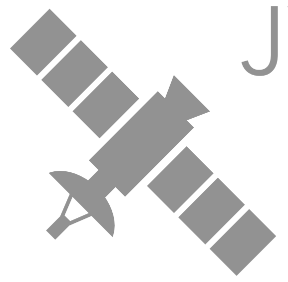
A more likely image



A likely image

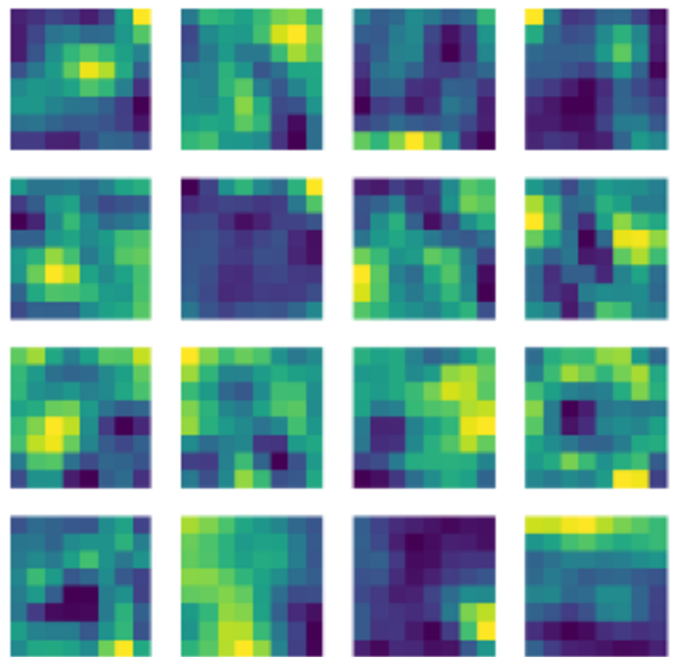
Patch based image prior

$$\mathscr{P}(\mathbf{x}) = \log\left(\sum_{k=1}^{K} \pi_k N(x_i; \mu_k, \sigma_k^2)\right)$$

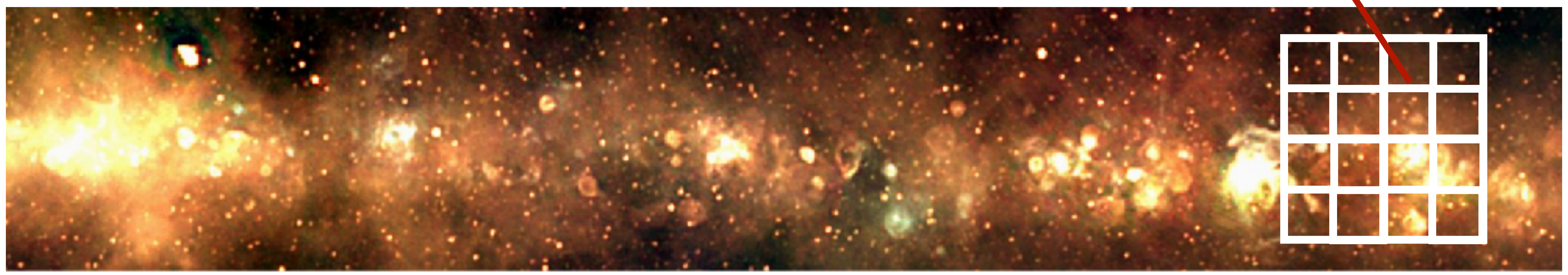


JWST Cas-A

Some example patches from an astronomical image…

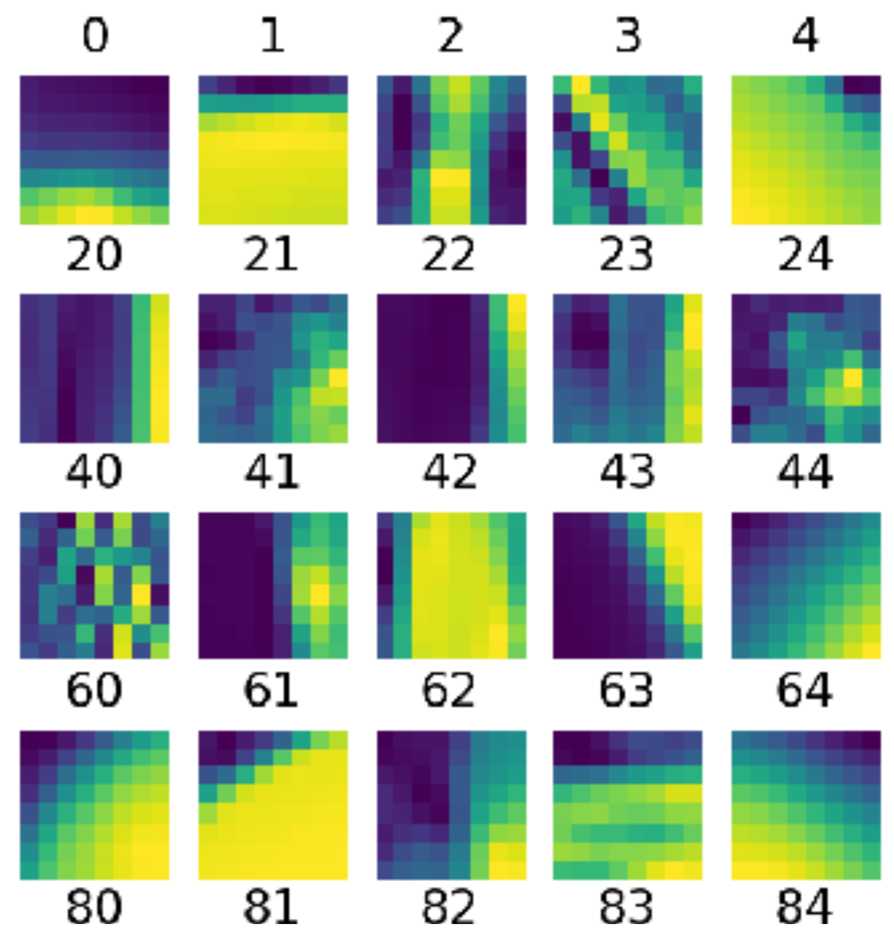


GLEAM Radio Survey



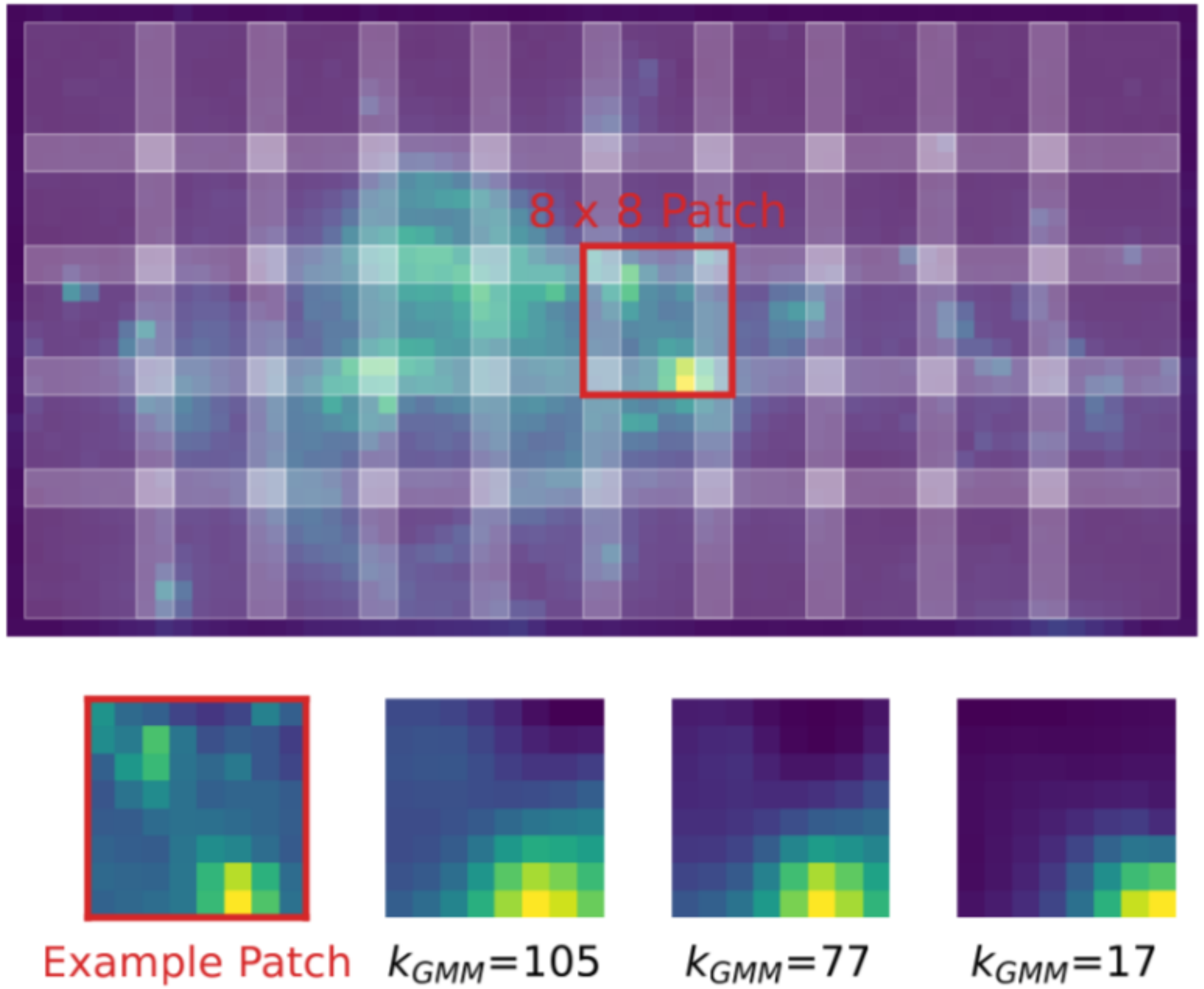
- Split training images(s) into "patches" of a given size, e.g. 8x8 pixels
- Learn a 64 dimensional **Gaussian Mixture Model (GMM) with N components** on the distribution of patches.
- Compute the likelihood for a GMM and train them using EM

Means of the GMM model



- GMM as clustering algorithm: **patches are grouped in different "base" structures such as edges, curves, lines, etc.**
- Initial idea by [Zoran et al. 2011] GMM trained on "every day" images, cars, people landscapes, etc.
- Also used in EHT reconstruction "CHIRP" algorithm [Bouman et al. 2016]

Reconstruction with the patch prior



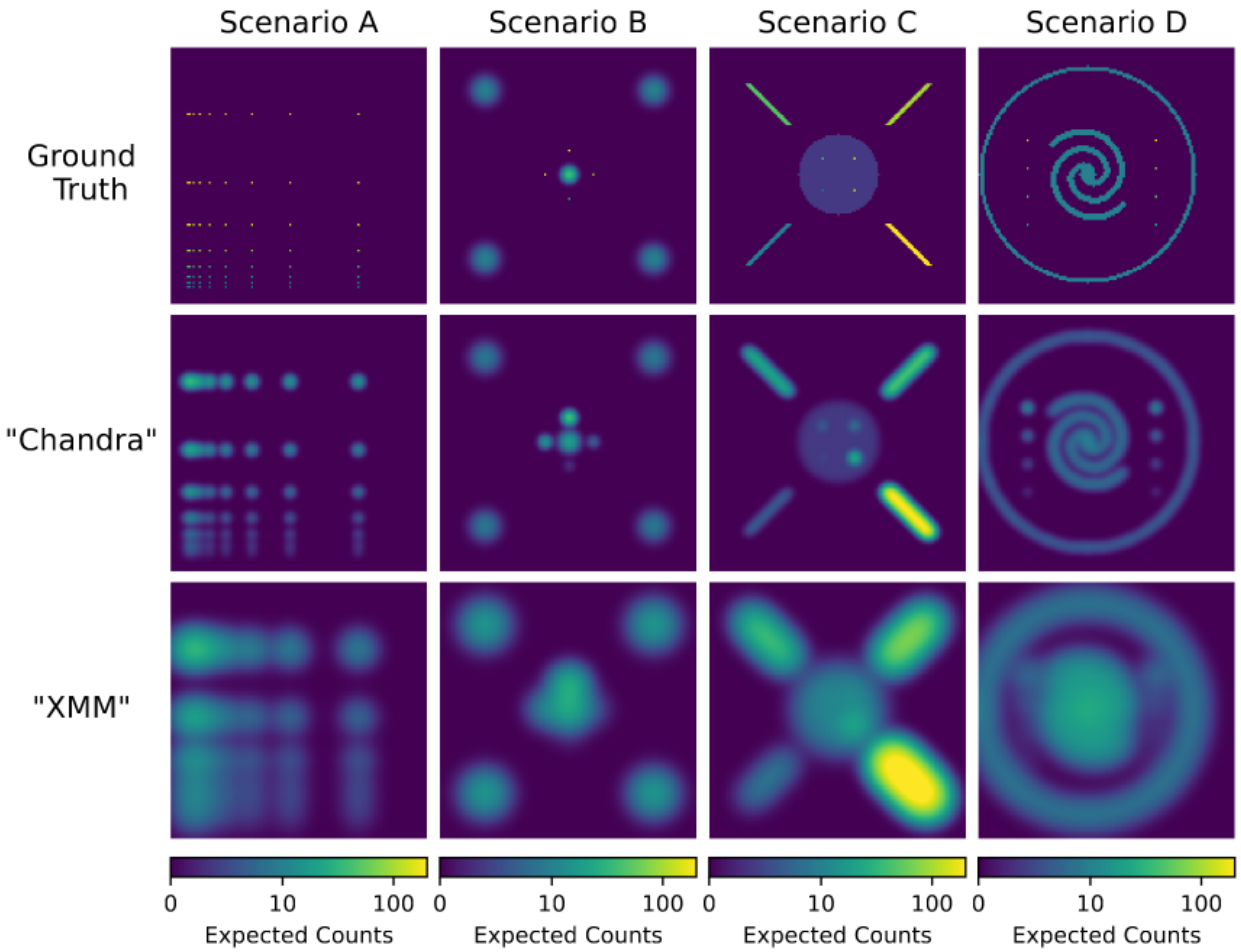
Intuition: the GMM effectively works as a "patch denoiser" and draws the solution towards the most likely structure

- The learned GMM can then be used to build an approximate prior for the image reconstruction
- In each iteration split the current estimation for the reconstructed image into overlapping patches
- For each patch evaluate the GMM and choose the component with the highest log-likelihood
- Sum up these "best" log-likelihood values for all patches in the image to compute the total log-prior value

$$\mathscr{P}(\mathbf{x}) = \sum_i \log p_{\hat{k},GMM}(\mathbf{P}_i\mathbf{x})$$

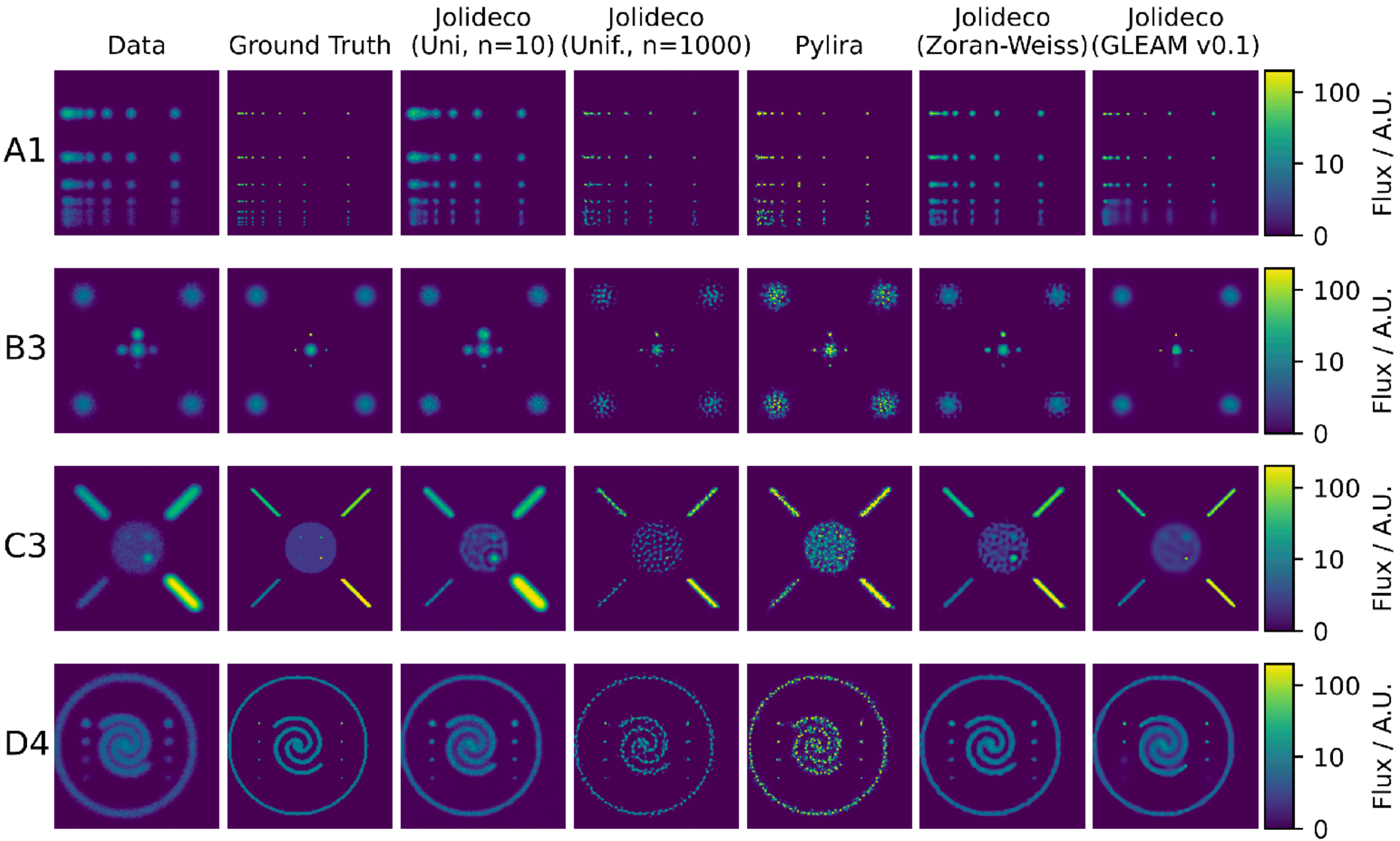
- Optimize the log-posterior to get a **Maximum A Posteriori (MAP) estimate** for the reconstructed image!

Test datasets

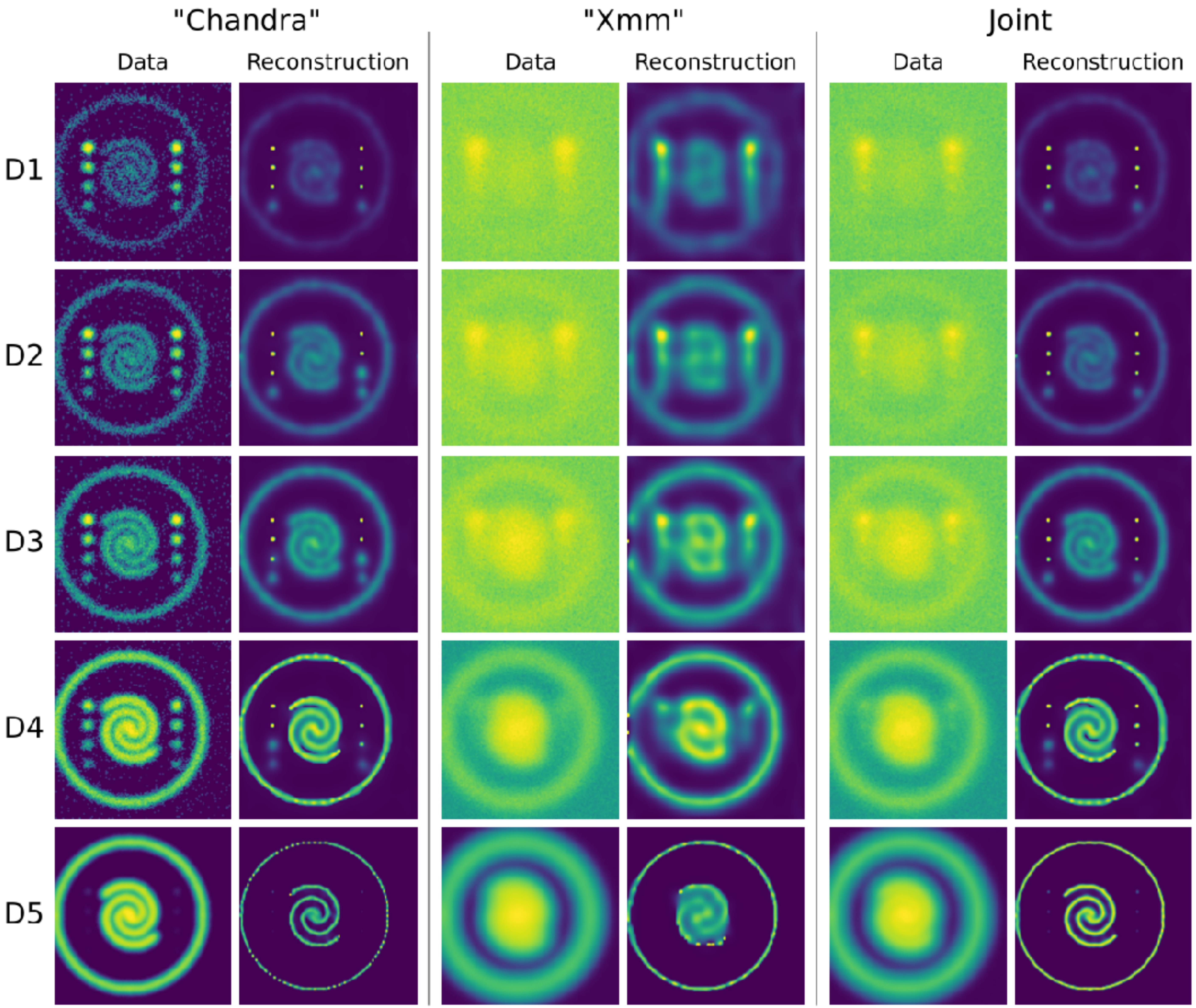


- Simulated dataset, with multiple scenarios A, B, C and D
- Dataset aims to cover a variety of semi-realistic astrophysical emission structures: **point sources, extended "center filled" sources, jet like features, disk shaped and spiral shaped structures.**
- Source are overlapping, and / or confused with nearby sources. **Designed to challenge astronomical deconvolution algorithms!**
- Two instrument scenarios for the joint likelihood case:
    - **"Chandra":** good angular resolution (Gaussian PSF of $\sigma = 2$ pix), smaller effective area (unity)
    - **"XMM":** worse angular resolution (Gaussian PSF of $\sigma = 6$ pix), larger effective area (by a factor of 5 compared to "Chandra")

Results for “Chandra” scenario

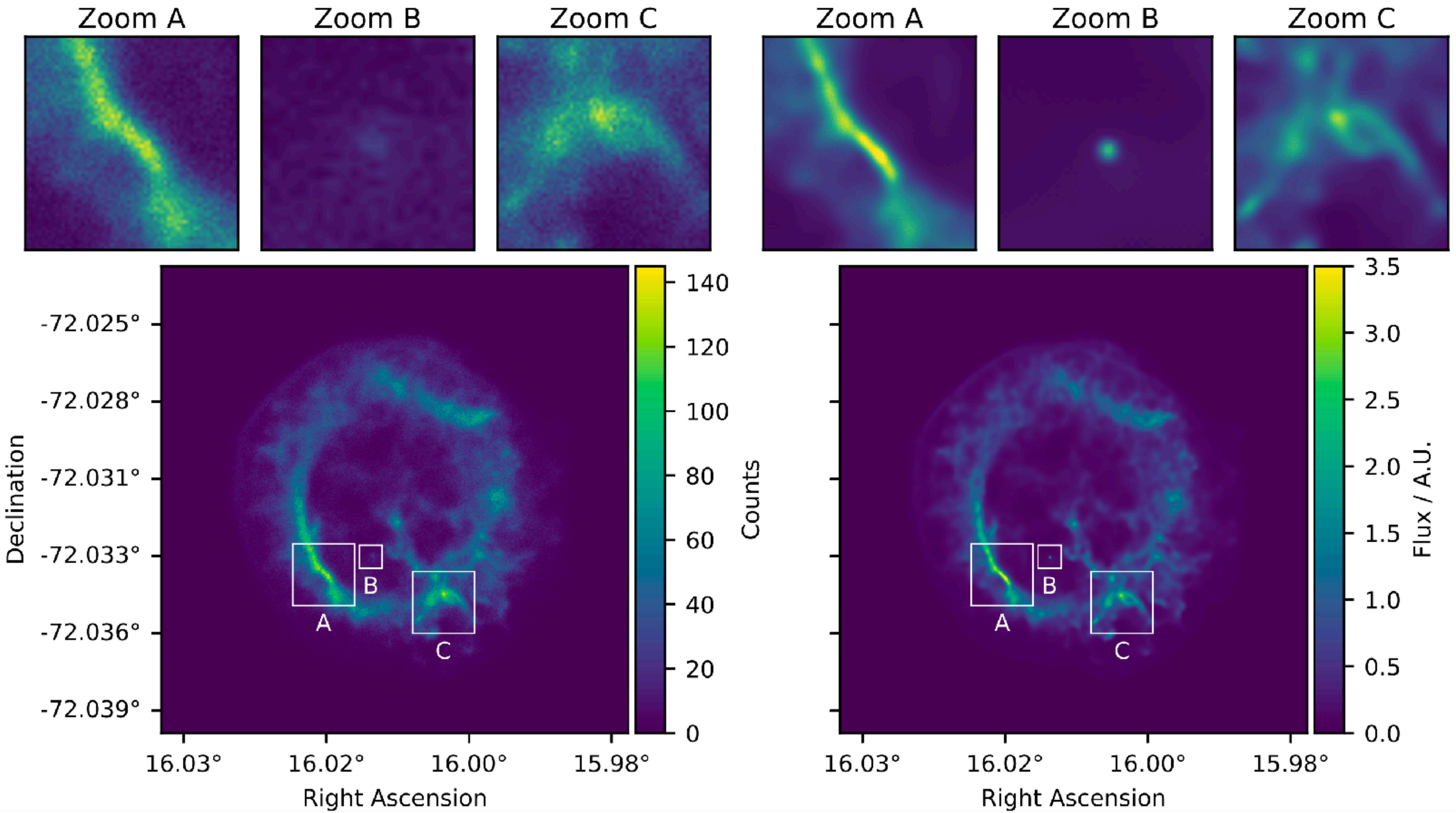


Varying instrument scenarios

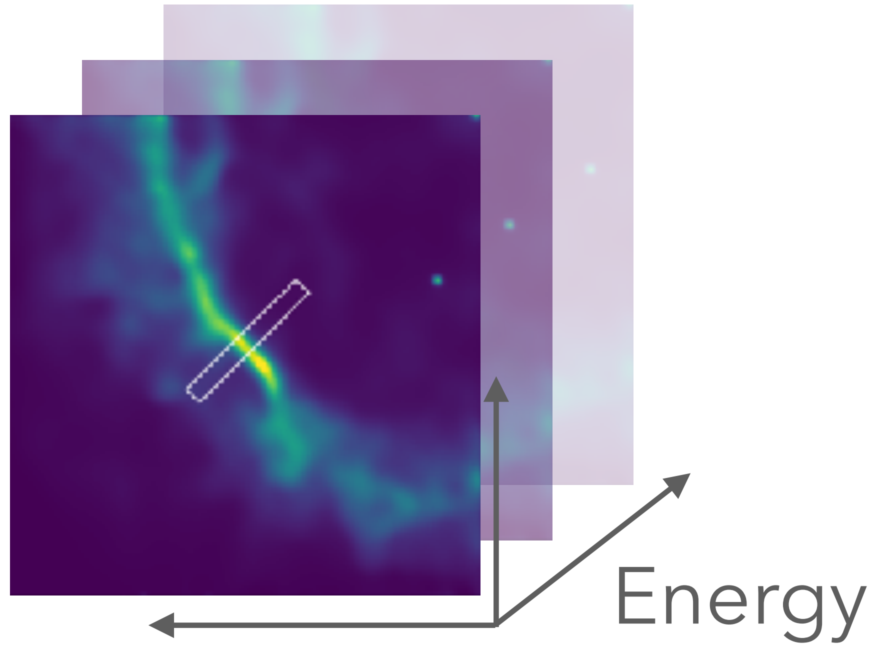


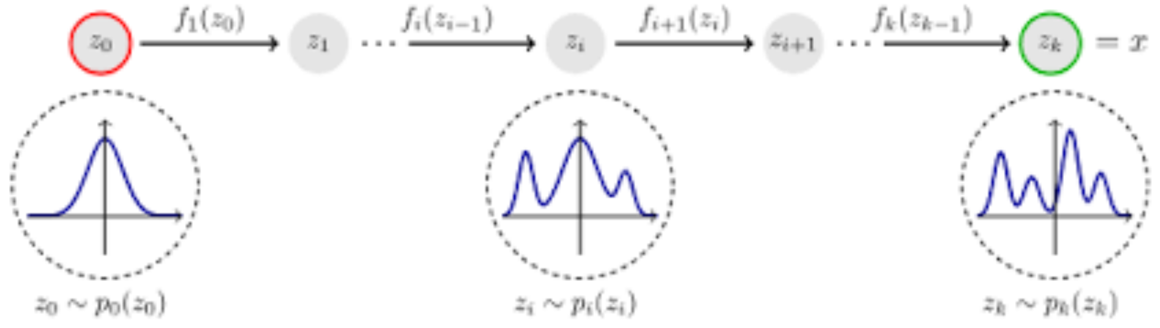
- **Including "lower quality" data always improves the result**. Never makes the result worse!
- The joint result are close to "Chandra", but there is still **improved reconstruction in the highest S/N** regime.
    - The larger "Xmm" PSF maintains the pixel correlations between more distant pixels and improves the "speckle" effect
    - However this might not be true, if one of the datasets is affect by large systematics…albeit Jolideco can correct some of those.

Example: Chandra E0102



Summary & ideas





- Jolideco is a **method for joint likelihood deconvolution** for astronomical images in the presence of Poisson noise
- It uses a **patch based image prior to reconstruct a flux image** from multiple observations of the same region of the sky. Taking individual instrument response such as PSF, exposure and instrumental background into account.
- Both the patch prior as well as accounting for multiple observations **improve the reconstruction quality**!

- Jolideco currently uses a **Maximum A-Posteriori approach** to get a point estimate. Get a **generic error estimate from the likelihood function or change to sampling from the posterior instead.**
- Extend method to **handle the spectral dimension** as well. I.e. provide "deconvolution" as well as spectral "unfolding" at the same time.
- Extend method to **handle multiple flux components** at the same time. E.g. treat point sources as independent model component?
- Try **different approach for patch modeling such as Normalizing Flows, other mixture models…**