Learning the energy dependence of a possible point source explanation of the Galactic Center Excess

| Eve Schoen | August 26th, 2024 | TeVPA

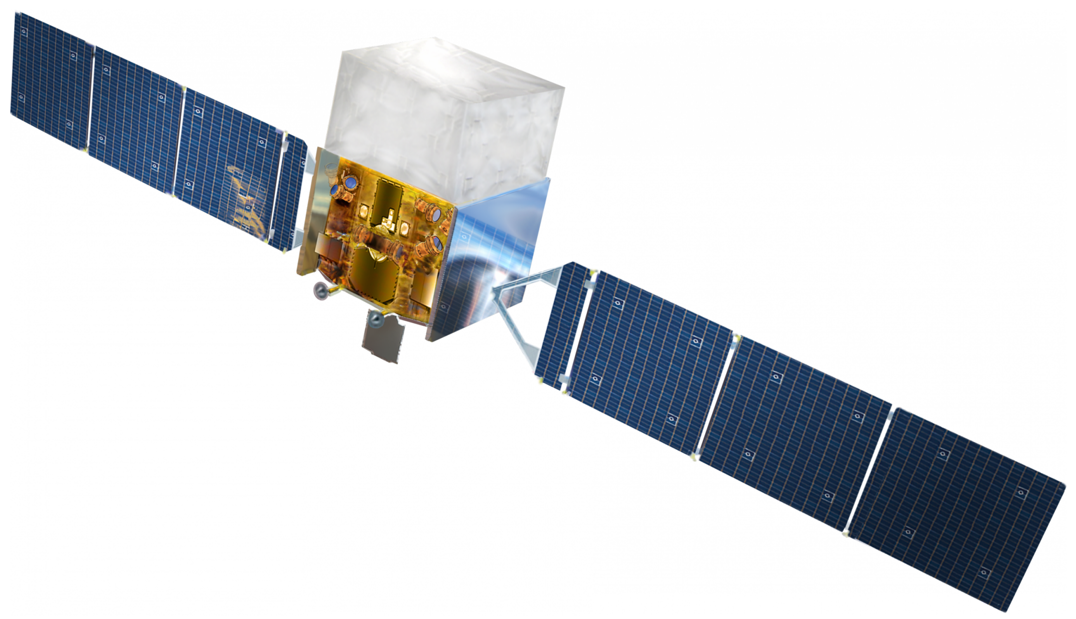
Collaborators: Florian List, Yujin Park, Nick Rodd, Florian Wolf

Talk Outline

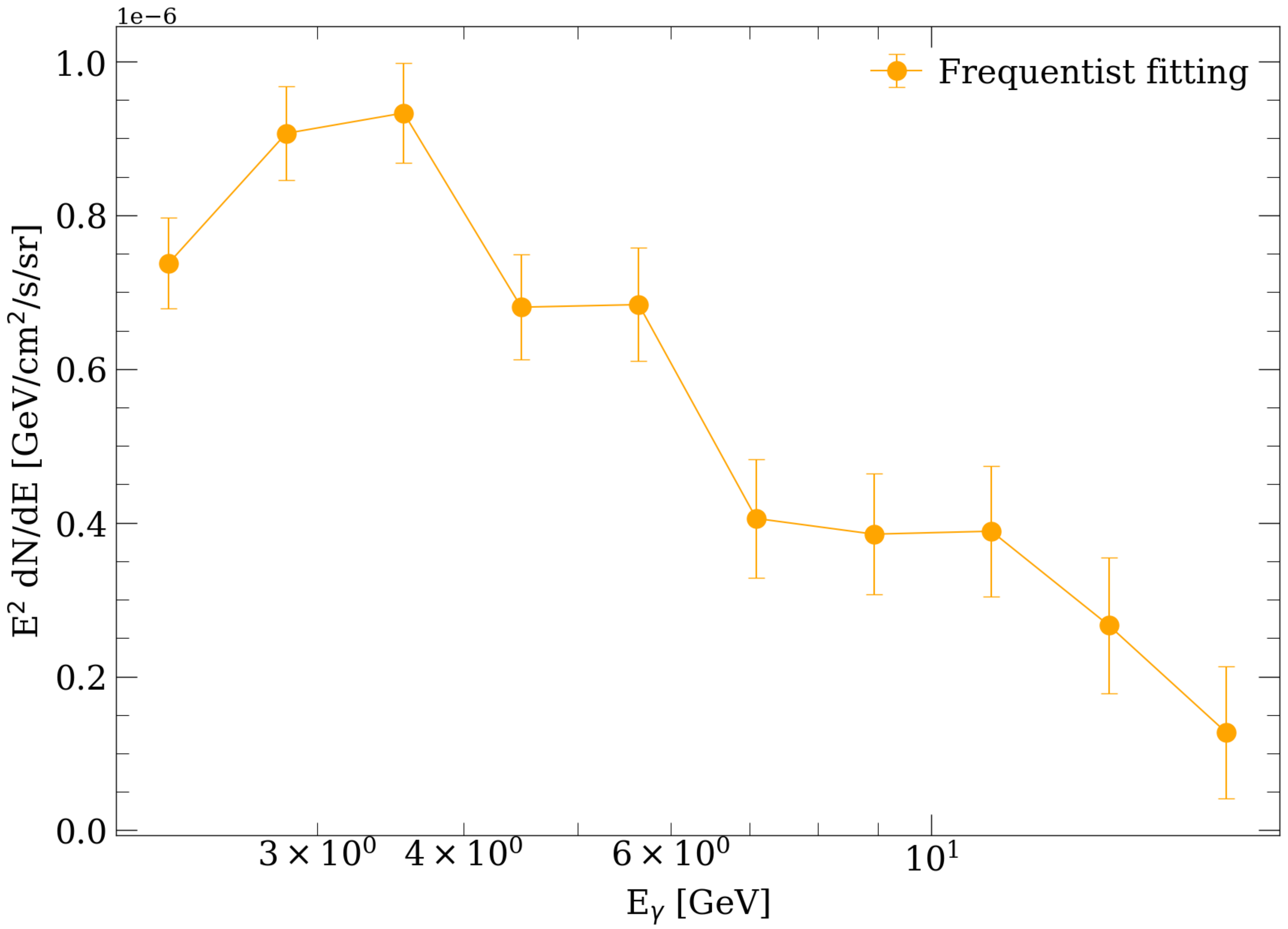
- What is the Galactic center excess (GCE)?
- How might we tell if it is dark matter or millisecond pulsars?
- Why would a neural network (NN) be good for this job?
- How is our NN currently doing on test data?
- **Present the results of NN on the GCE, accounting for the energy spectra information for the first time.** And comparison to previous works.
- How does the NN do on injected signals and with mismodeling?

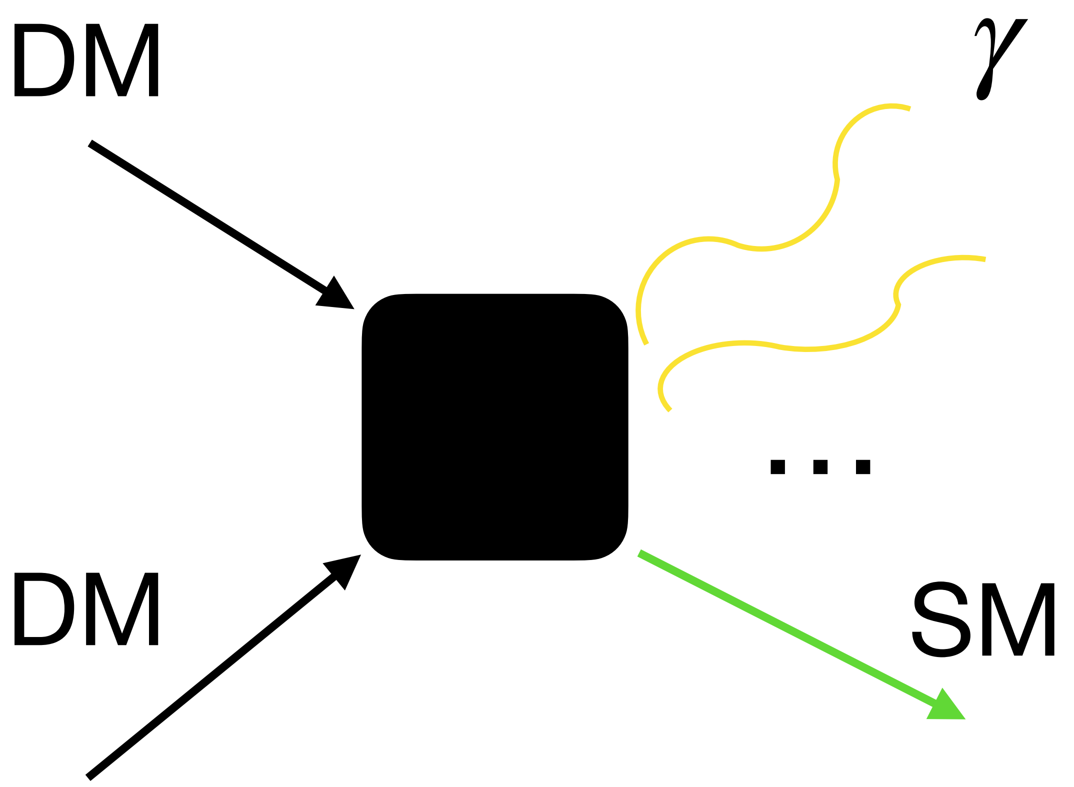
The Galactic Center Excess:
A potential signal of dark matter

Fermi LAT

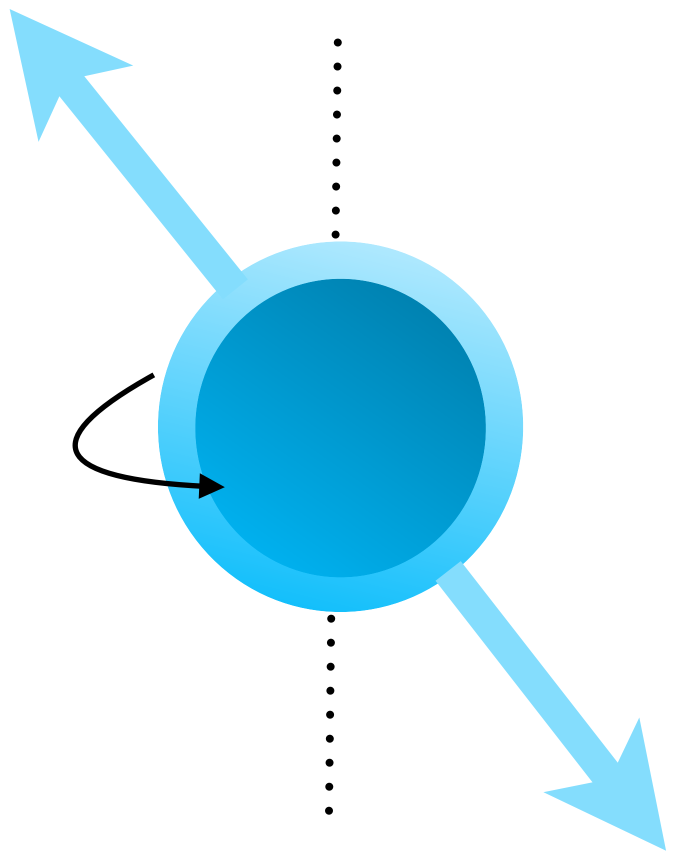


NASA



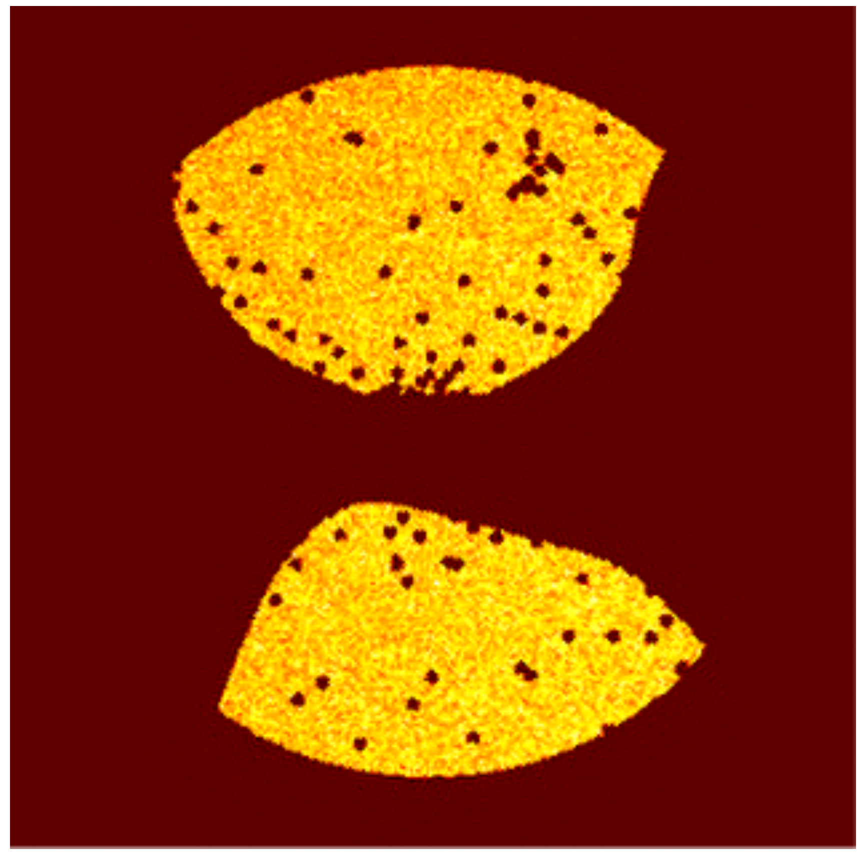


MS Pulsars

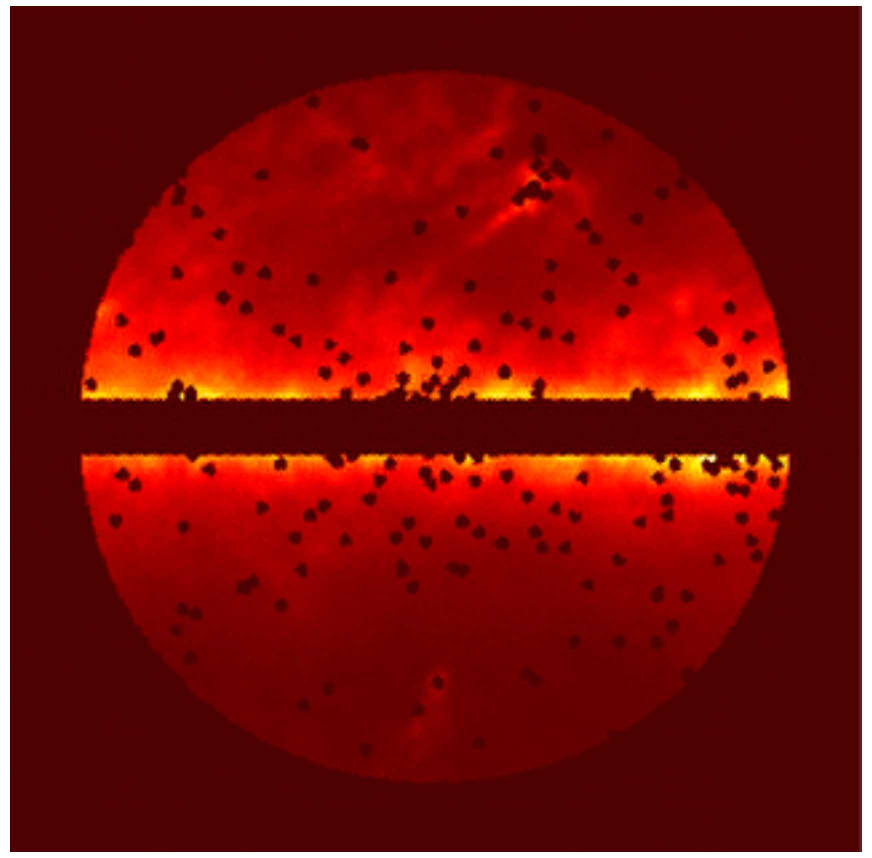


(1) What is the neural network learning exactly? → Energy Dependent Template Fitting

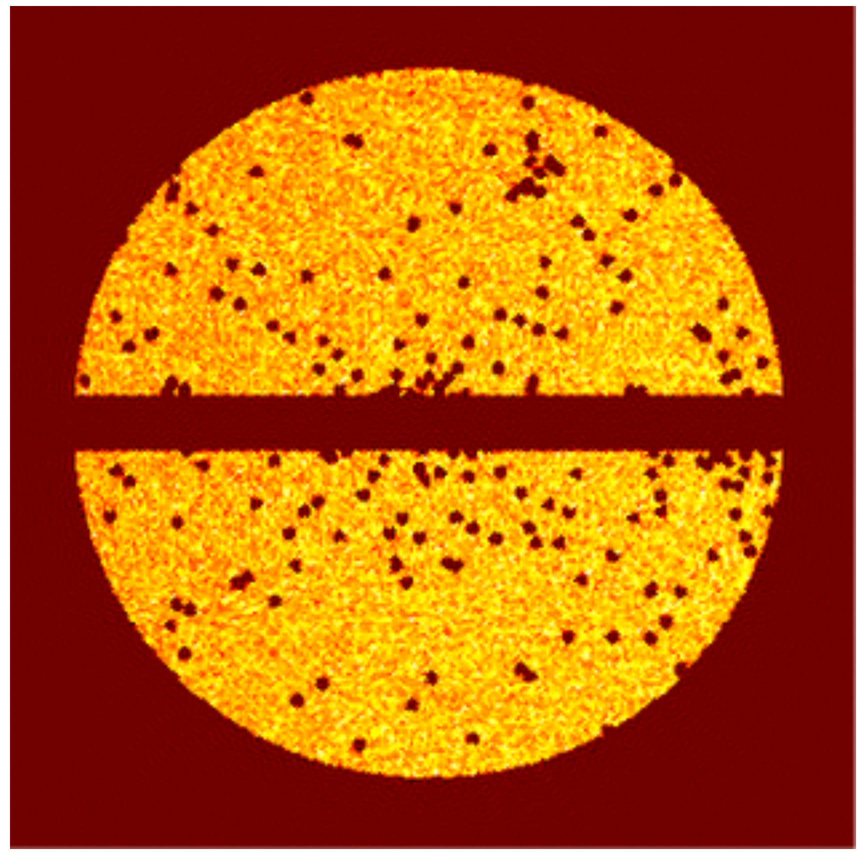
Fermi bubbles



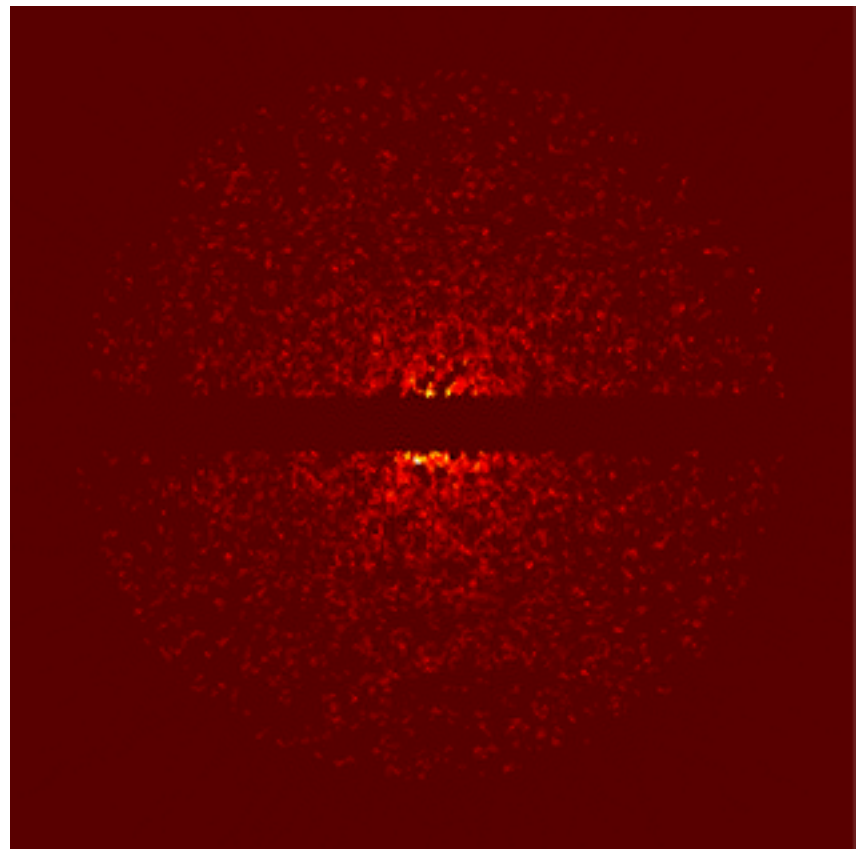
Diffuse



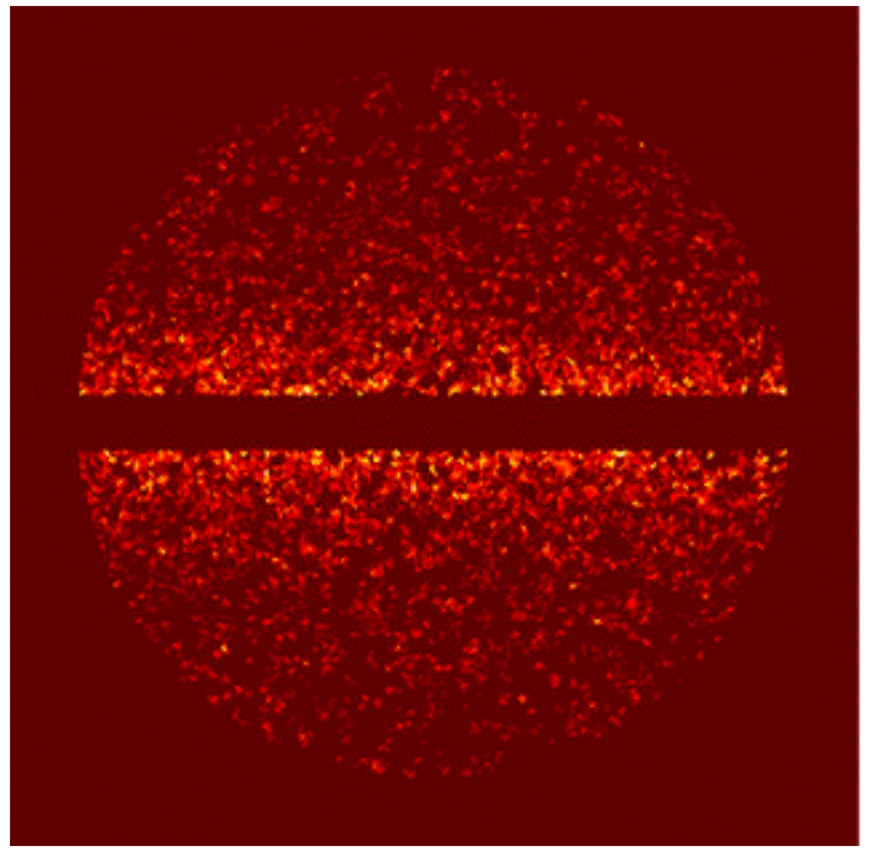
Isotropic



GCE

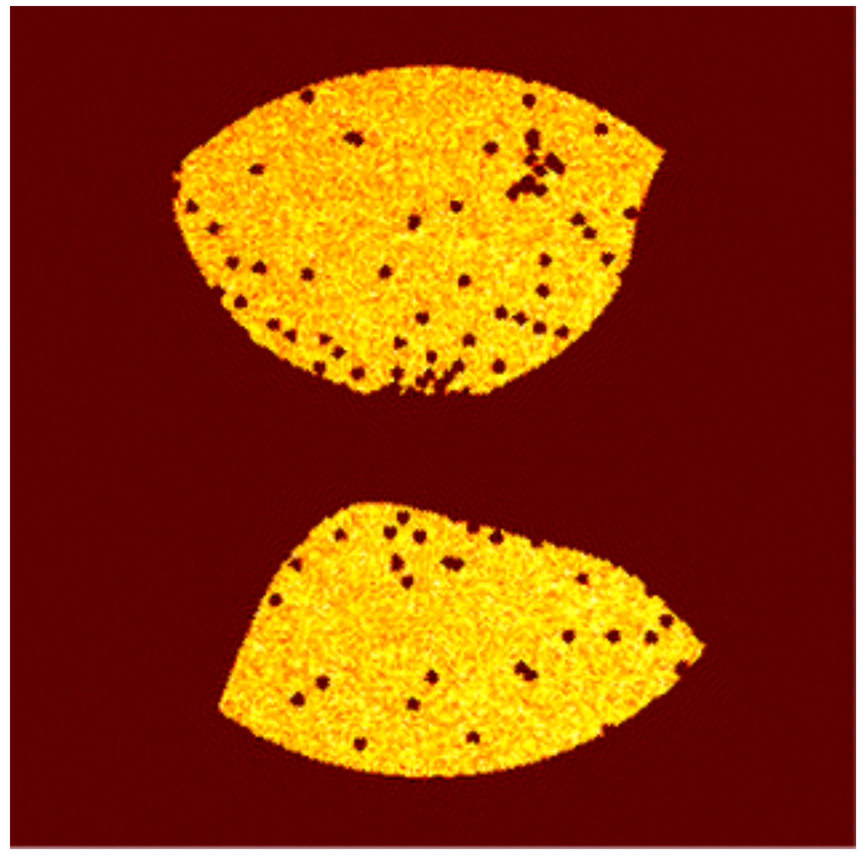


Disk

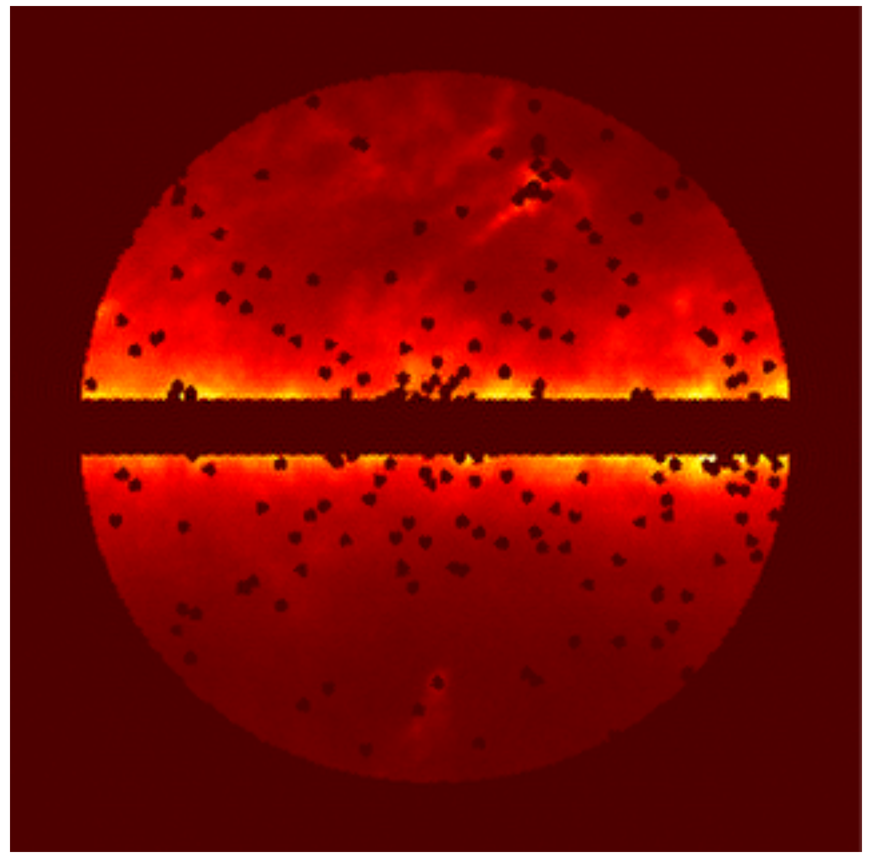


(1) What is the neural network learning exactly? → Energy Dependent Template Fitting

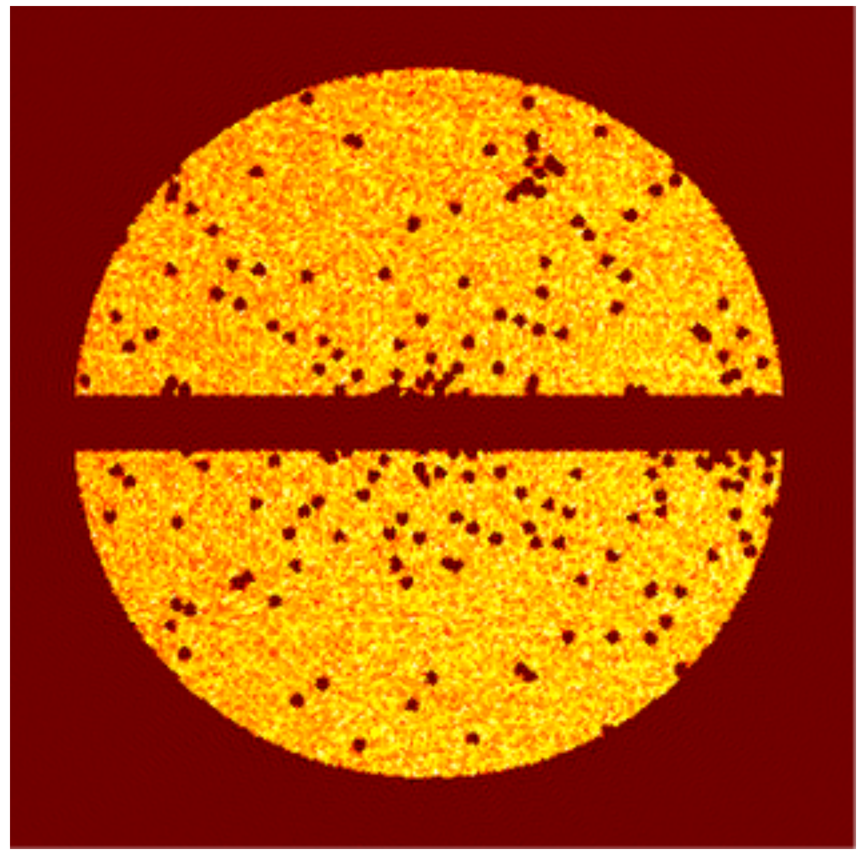
Fermi bubbles



Diffuse

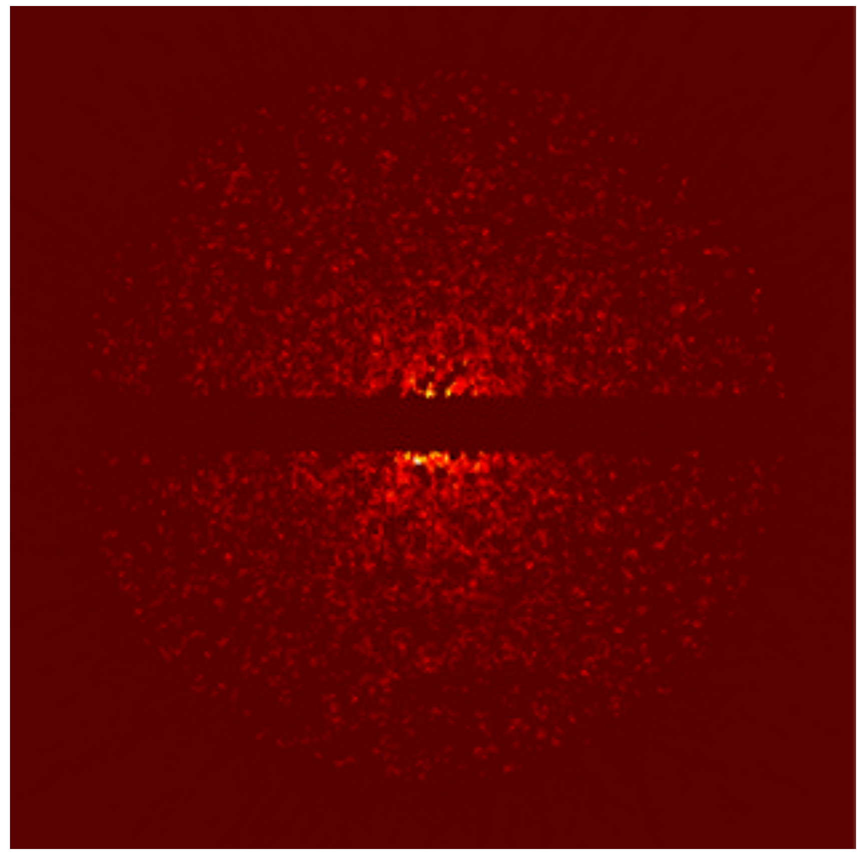


Isotropic

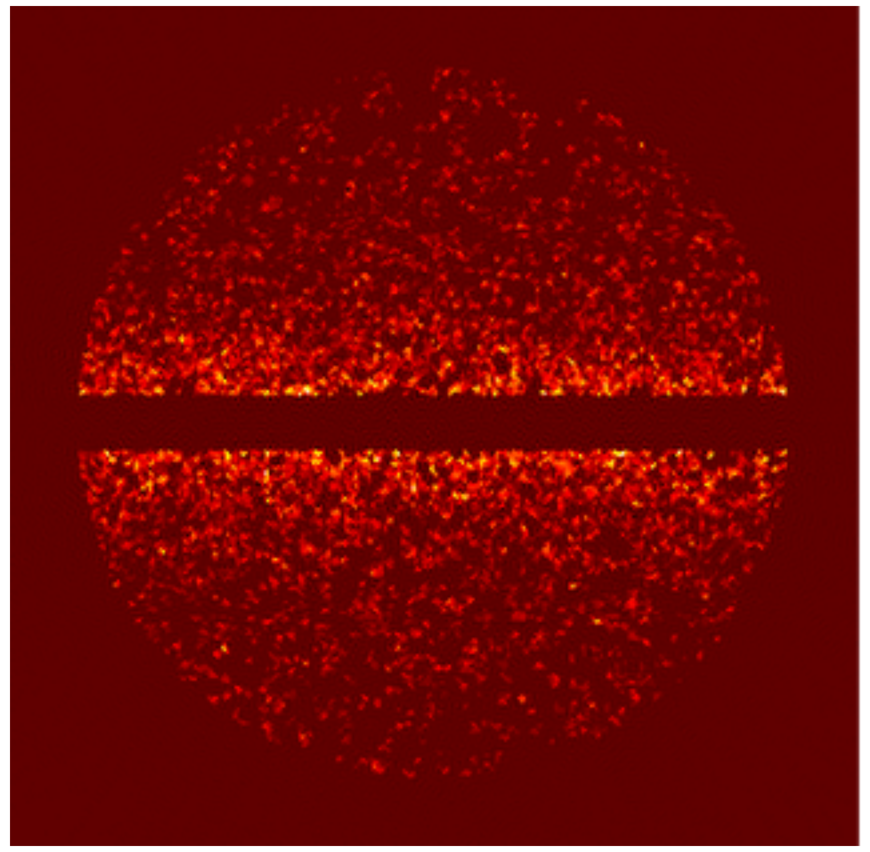


x 10

GCE

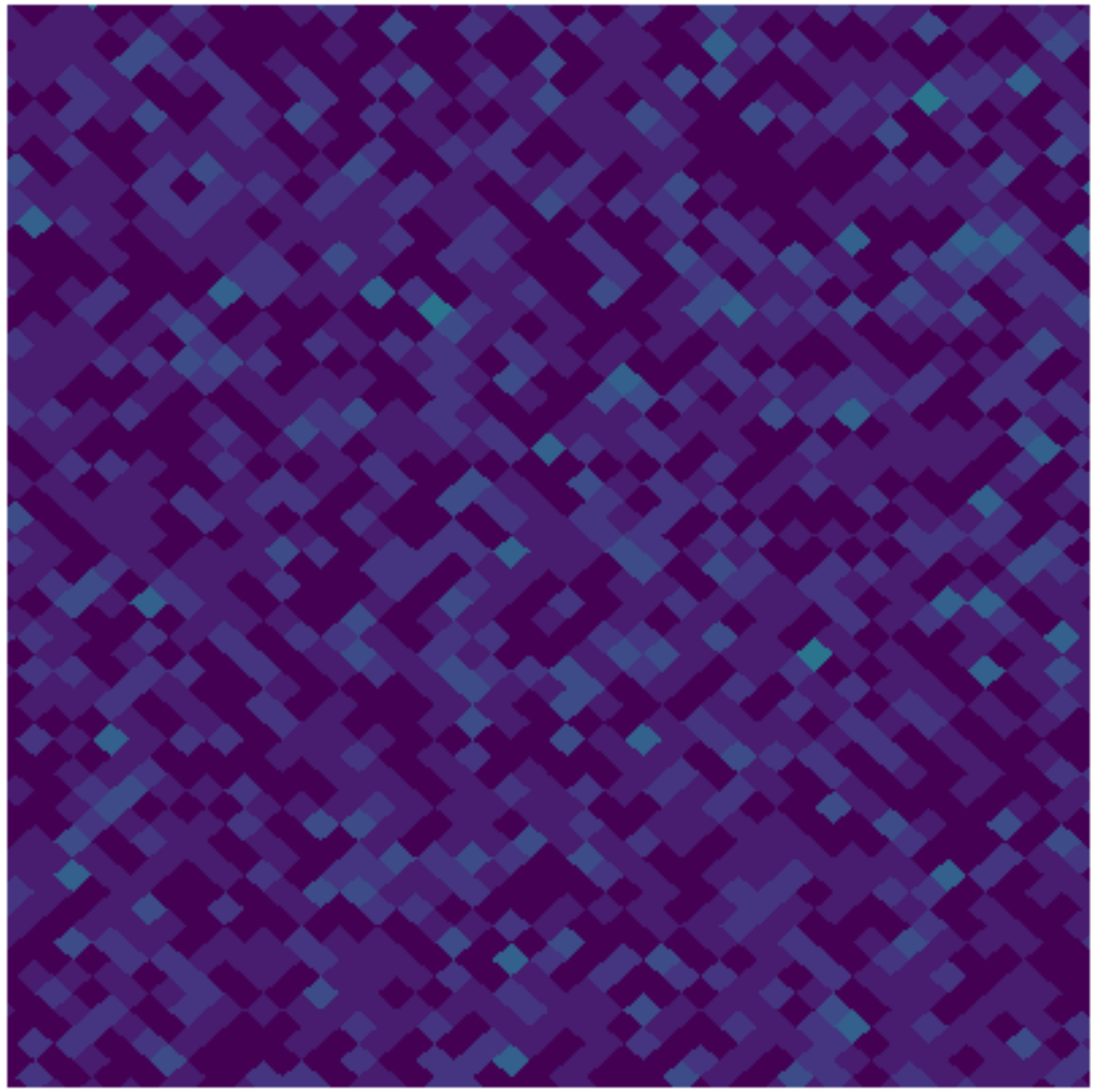


Disk



x 10

(2) What is the neural network learning exactly? → Point source statistics



N = avg number of sources

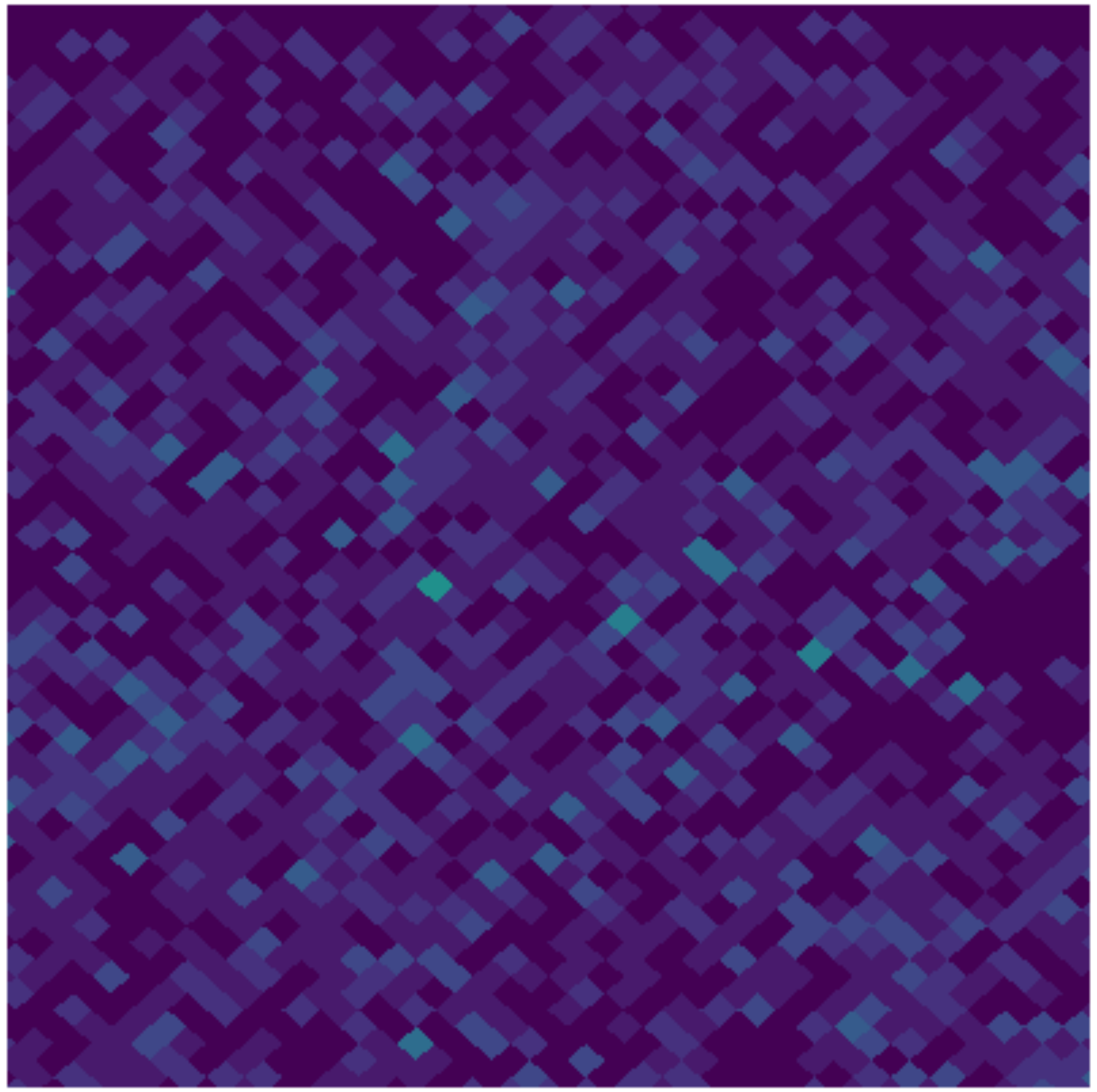
k = avg number of photons per source

$$\sigma^2_{\text{PS}} = Nk(1 + k)$$

(Limit $k \to 0$)

$$\sigma^2_{\text{Poisson}} = \mu = Nk$$

(2) What is the neural network learning exactly? → Point source statistics



N = avg number of sources

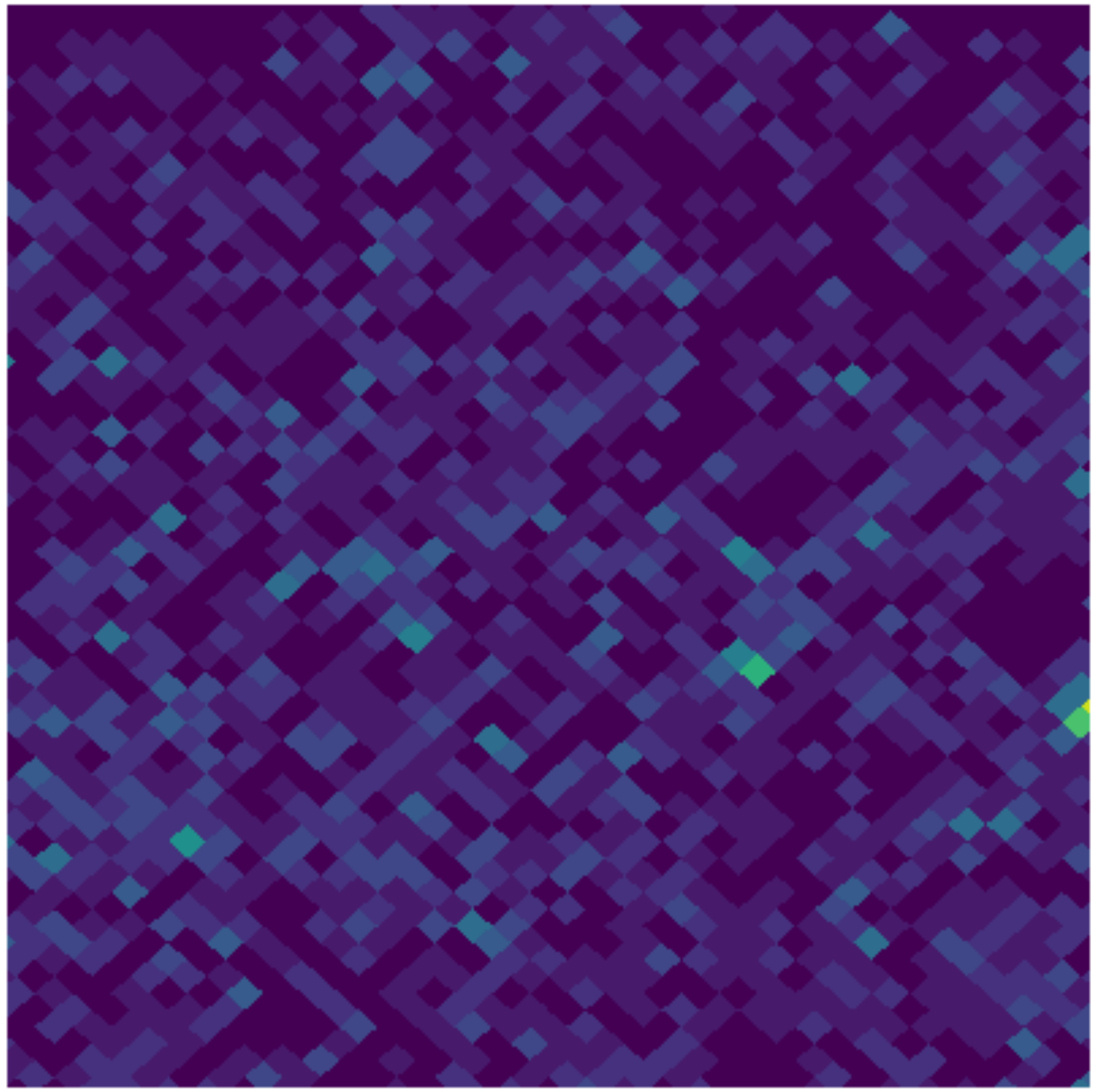
k = avg number of photons per source

$$\sigma^2_{\text{PS}} = Nk(1+k)$$

(Limit $k \to 0$)

$$\sigma^2_{\text{Poisson}} = \mu = Nk$$

(2) What is the neural network learning exactly? → Point source statistics



N = avg number of sources

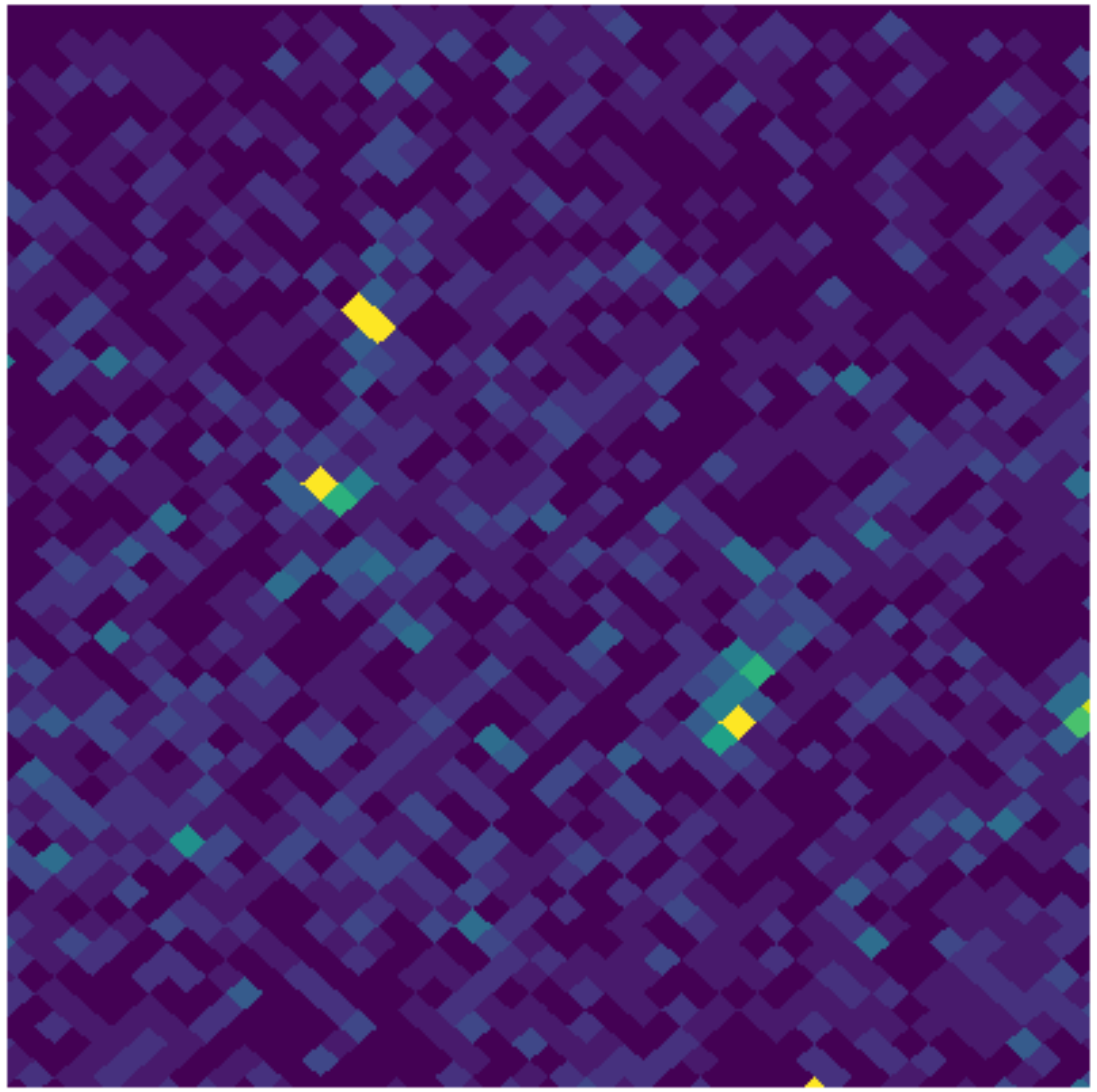
k = avg number of photons per source

$$\sigma^2_{\mathrm{PS}} = Nk(1+k)$$

(Limit $k \to 0$)

$$\sigma^2_{\mathrm{Poisson}} = \mu = Nk$$

(2) What is the neural network learning exactly? → Point source statistics



N = avg number of sources

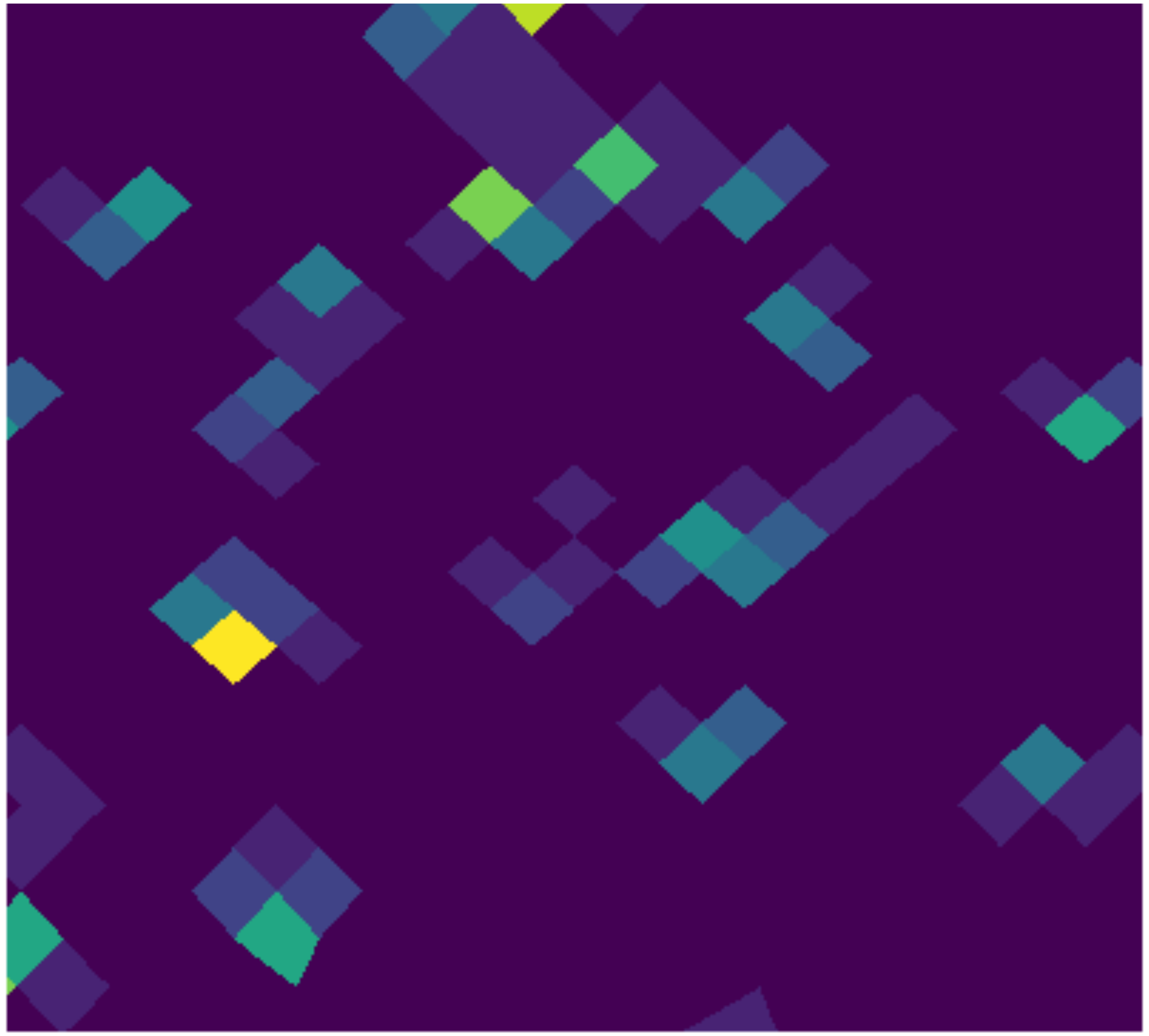
k = avg number of photons per source

$$\sigma^2_{\text{PS}} = Nk(1 + k)$$

$$\Big\downarrow \text{(Limit } k \to 0\text{)}$$

$$\sigma^2_{\text{Poisson}} = \mu = Nk$$

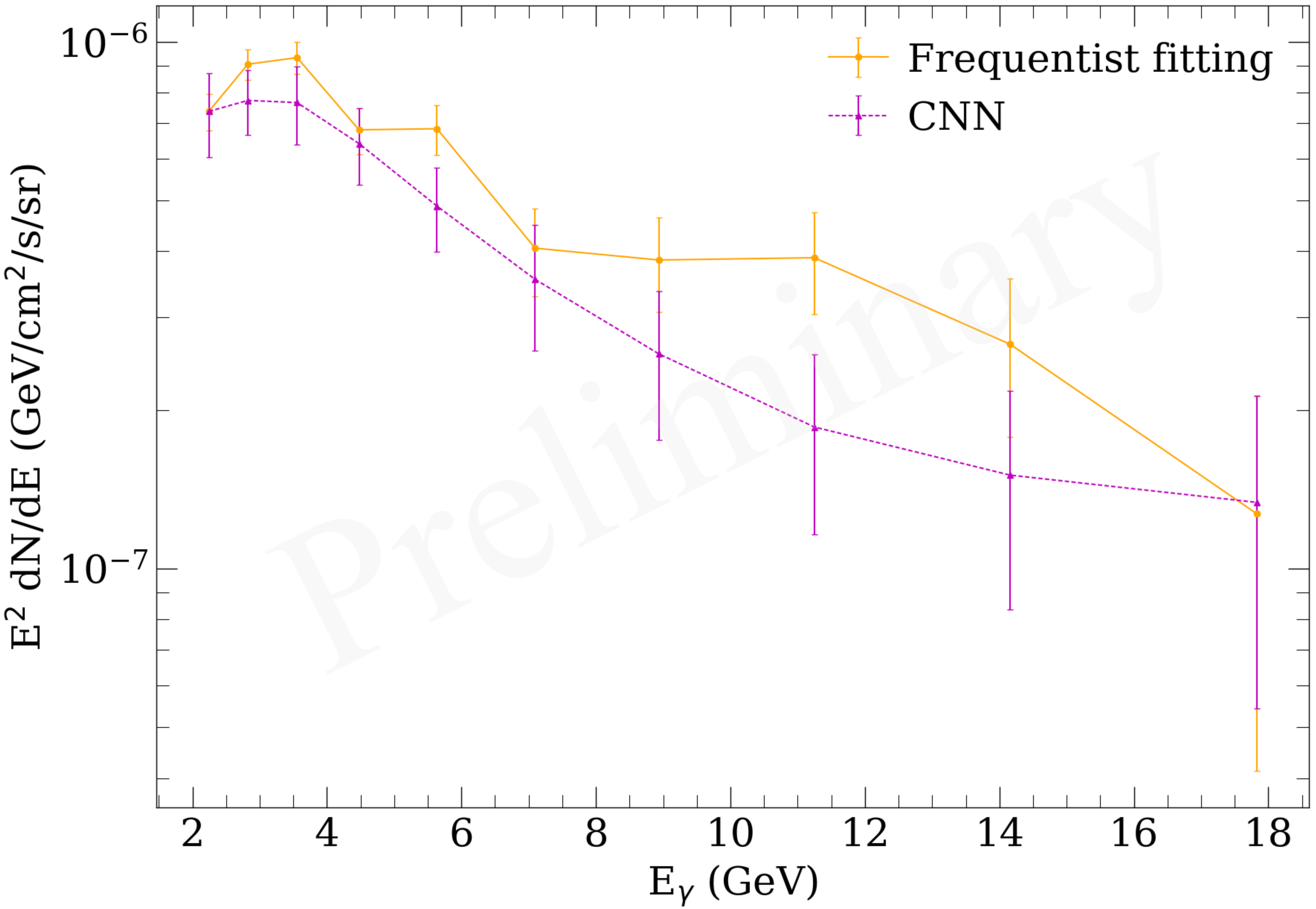
(2) What is the neural network learning exactly? → Point source statistics



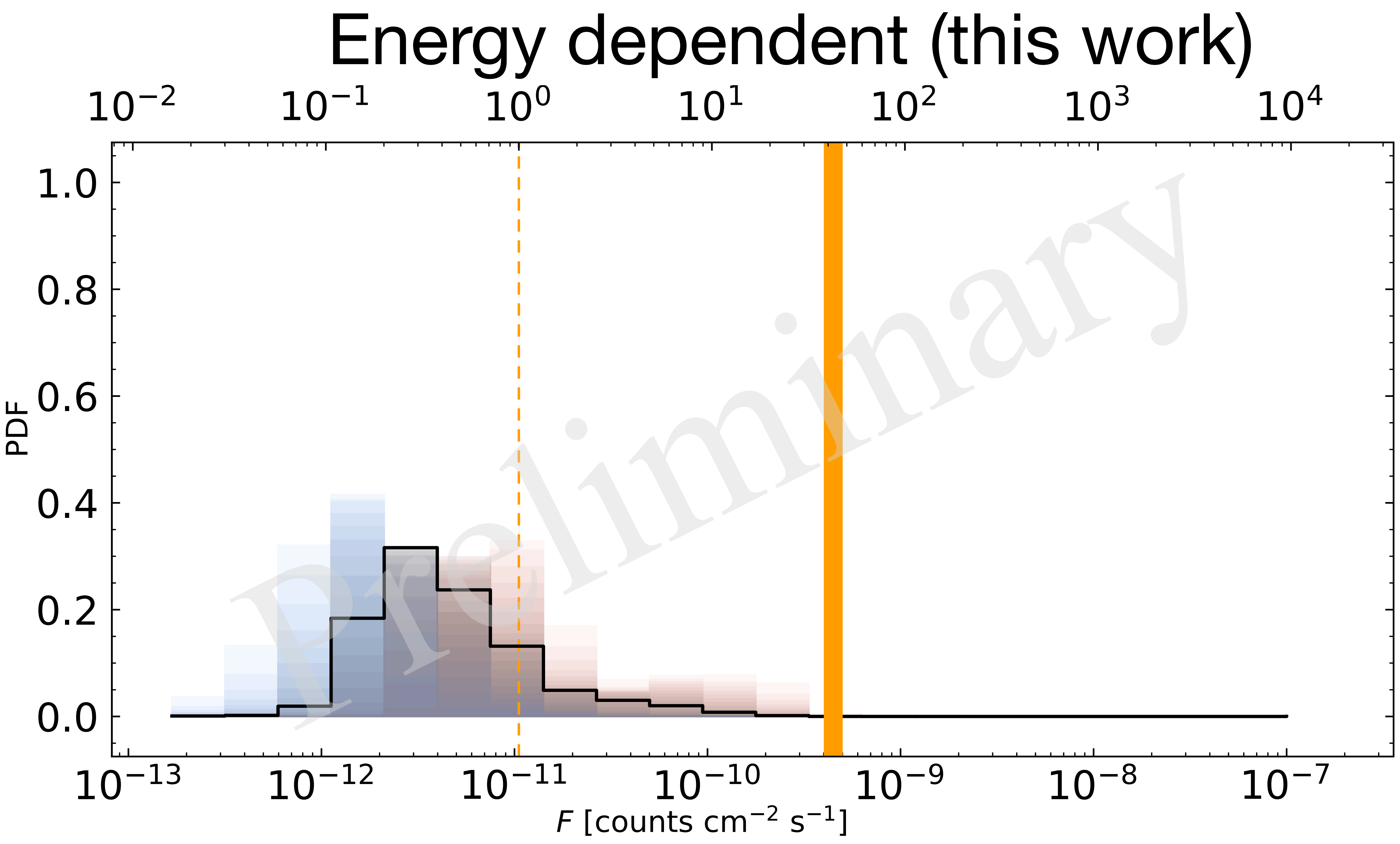
Computation complicated by the PSF, the next talk will probably talk about this more!

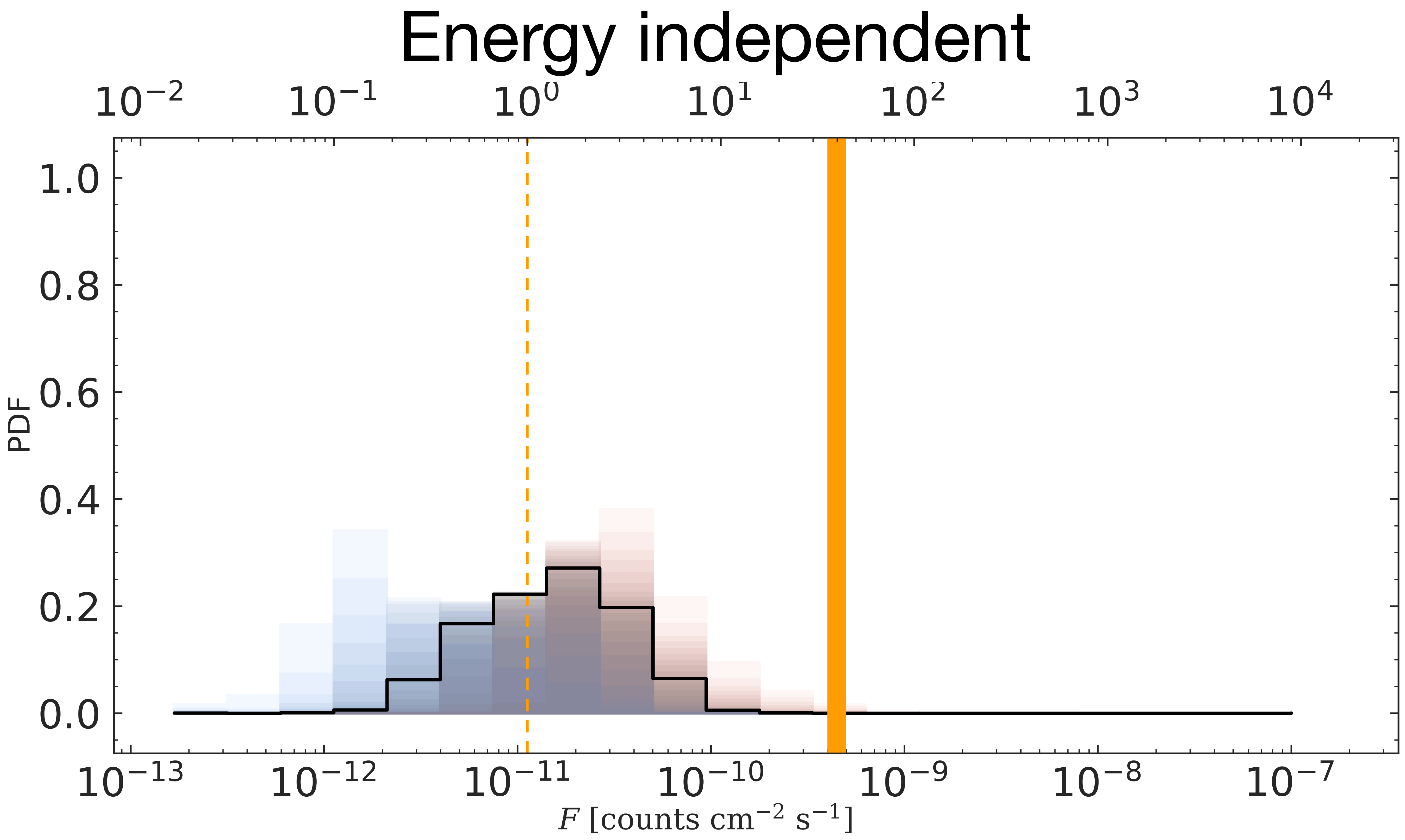
Current performance of the NN on the GCE of Test Maps

The GCE energy spectra

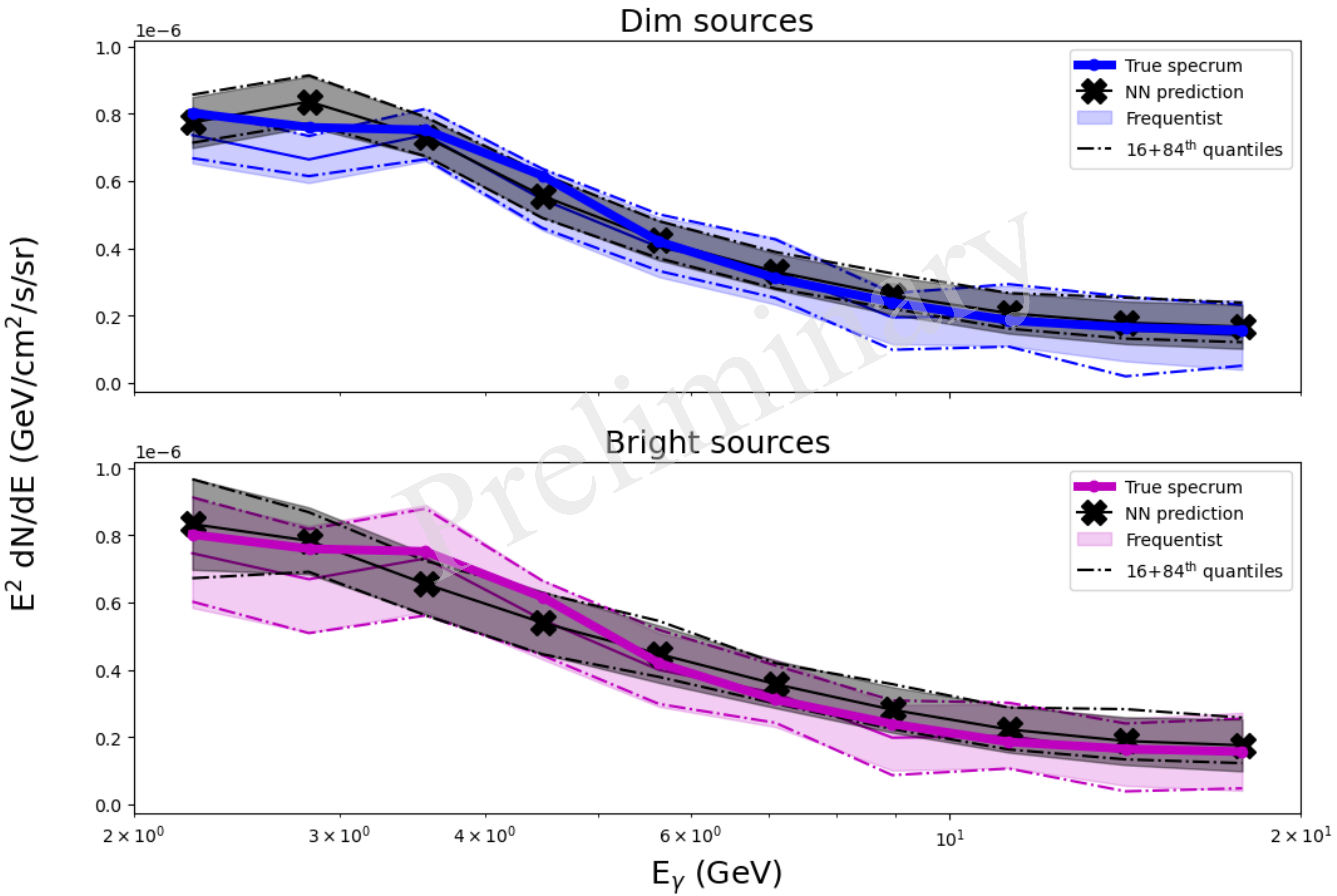


Source count distribution of the GCE predicted



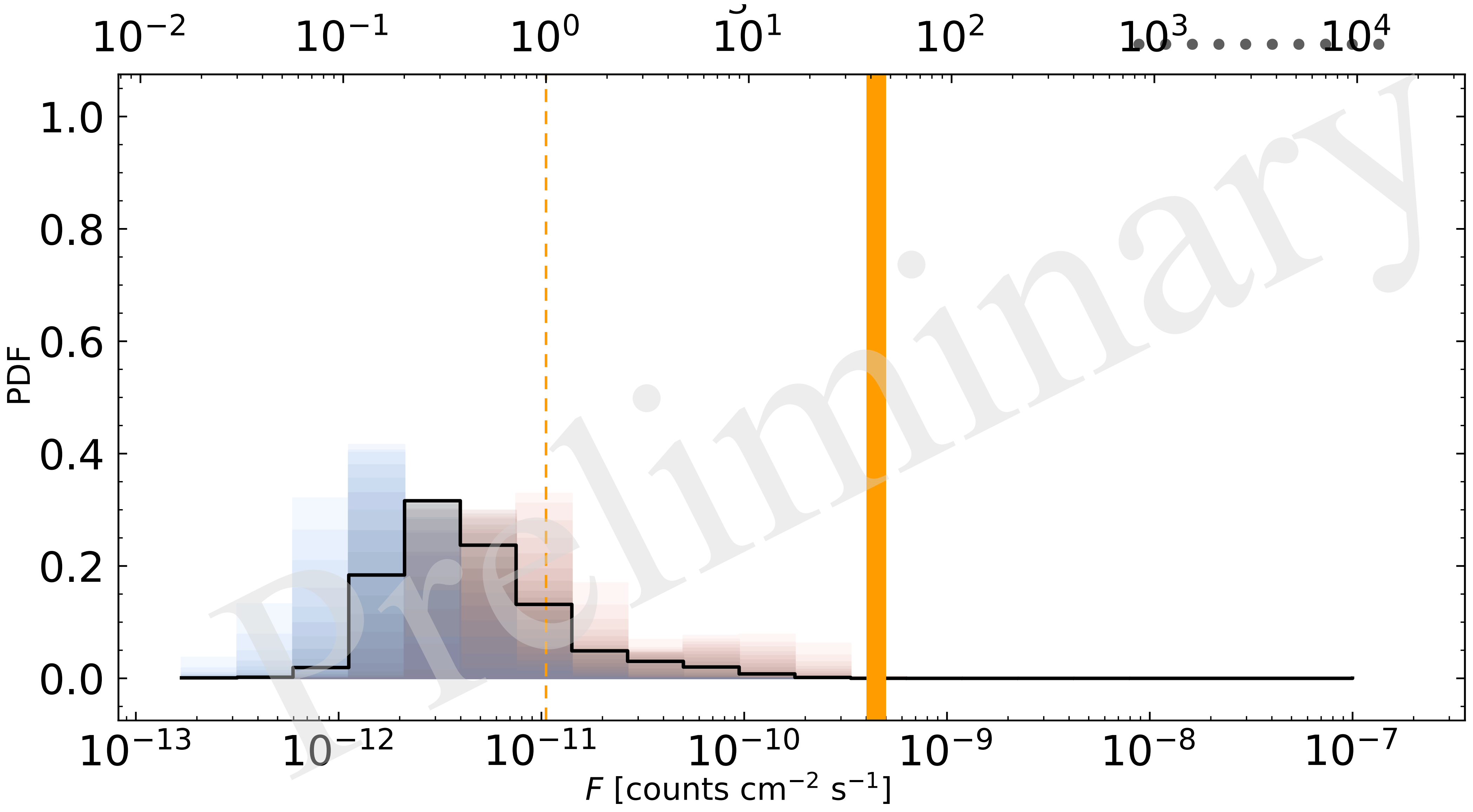


Testing NN and frequentist method on injected data sets

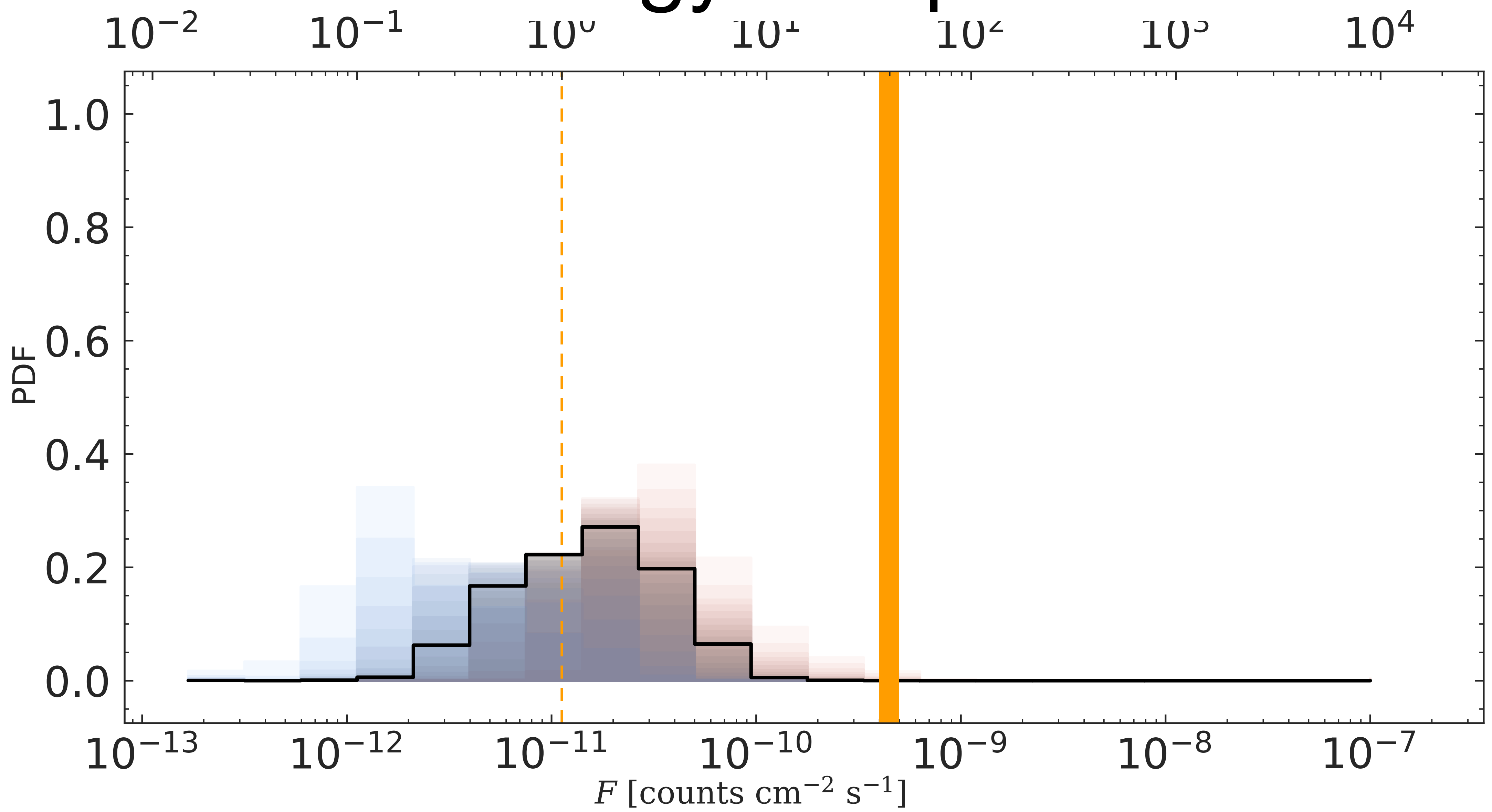


What is driving the new result? Likely not energy dependence of PSF

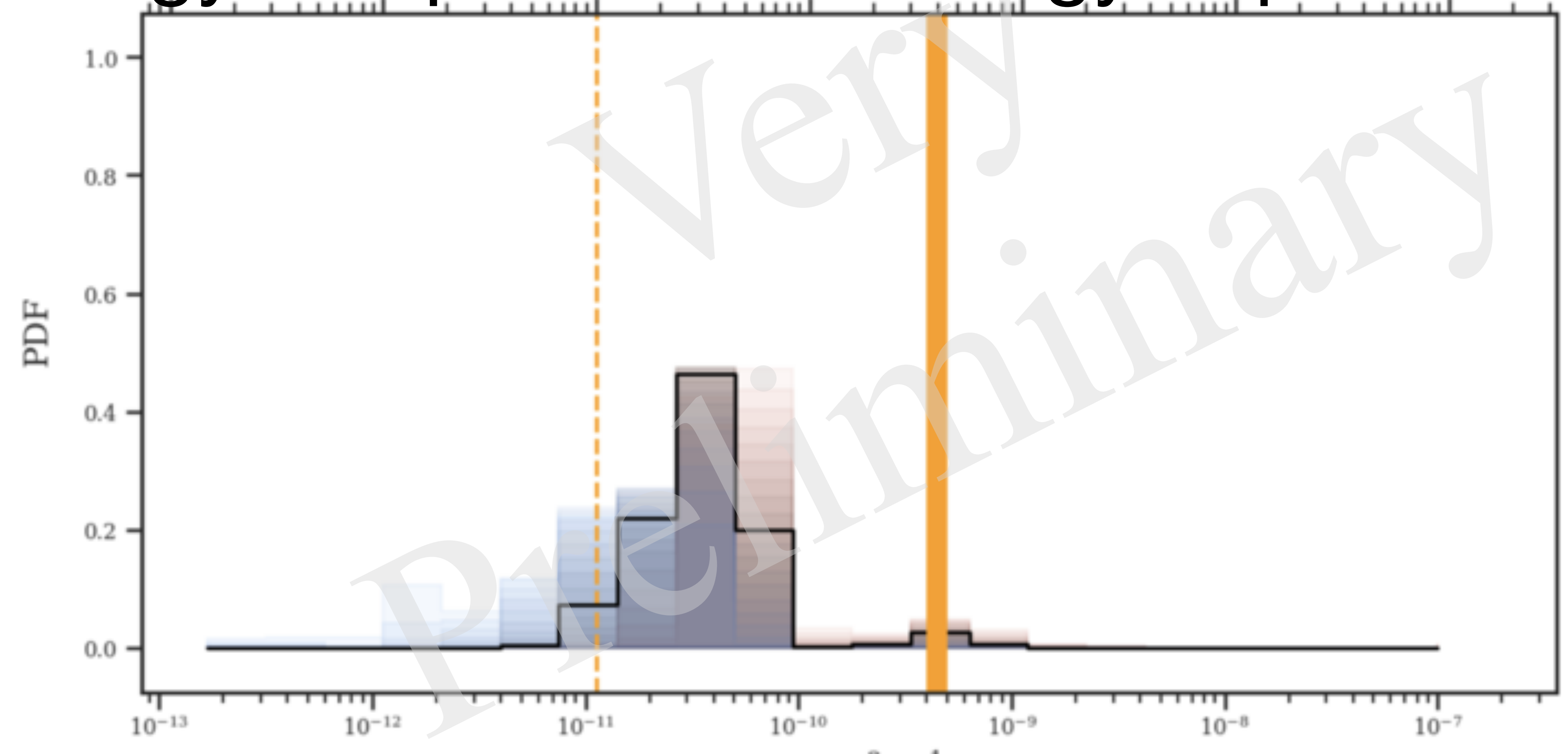
Energy dependent (this work)



Energy independent

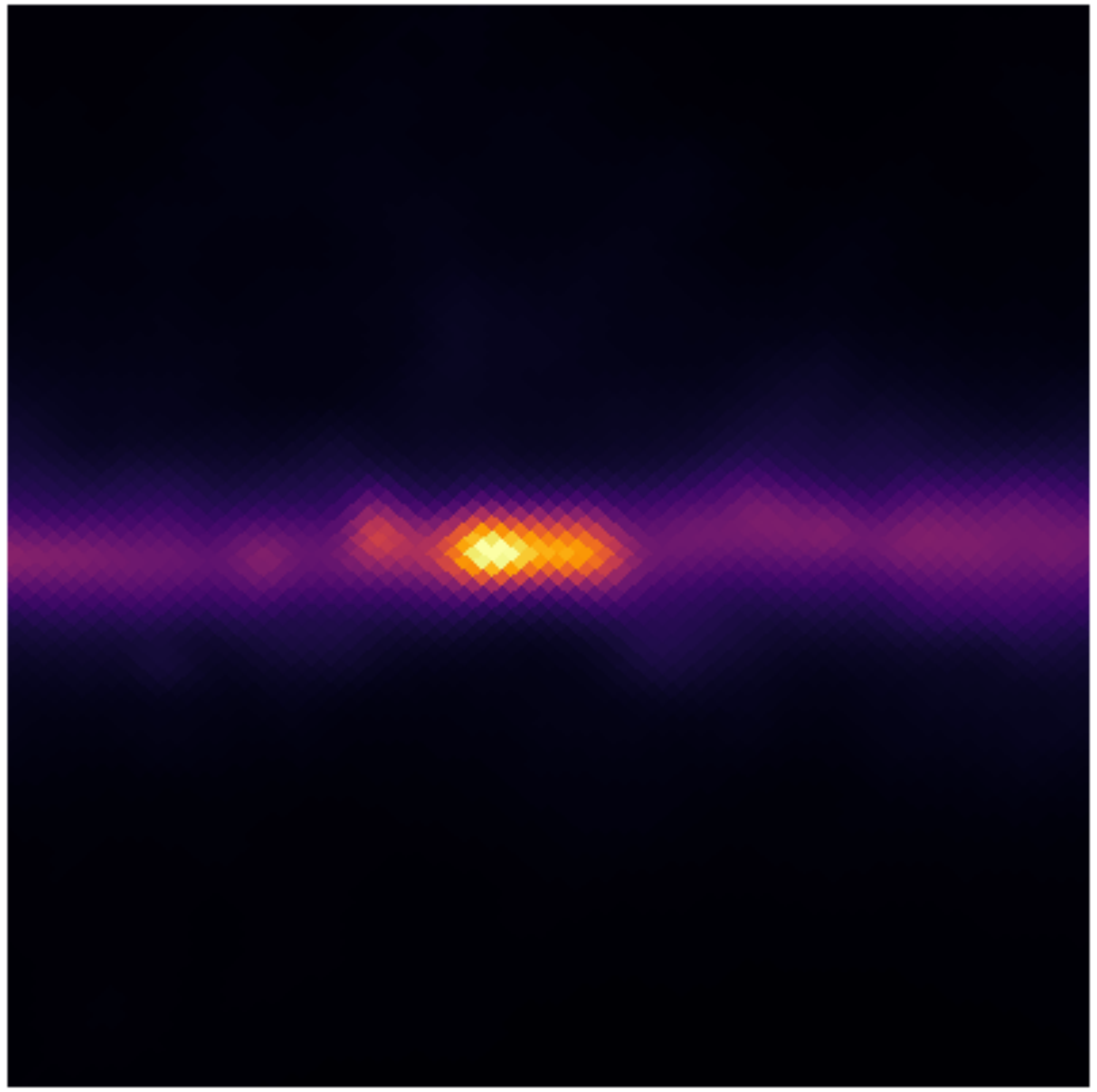


Energy independent on energy dependent data

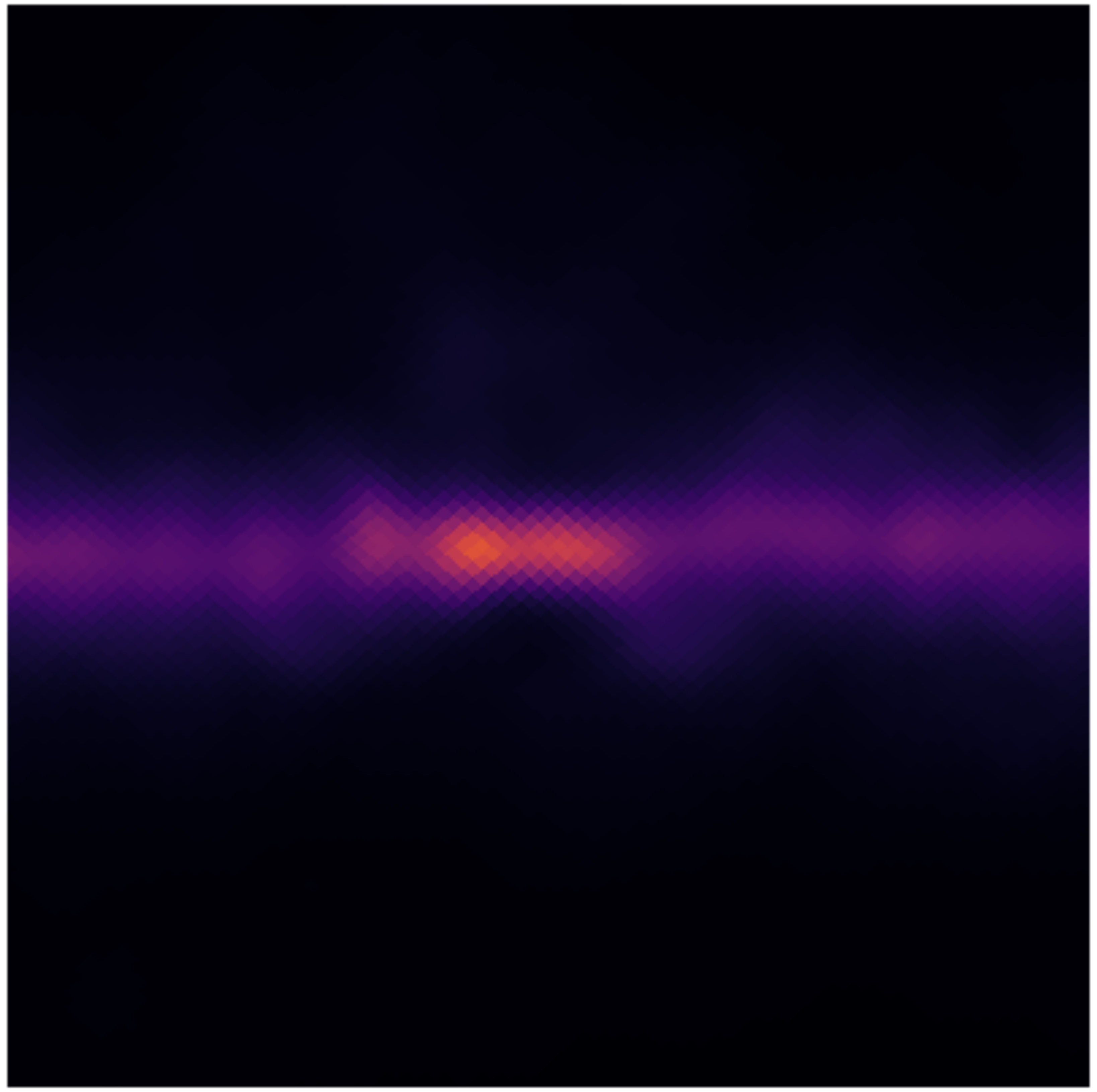


Ongoing testing of mismodeling of diffuse emission

Model O



Model F



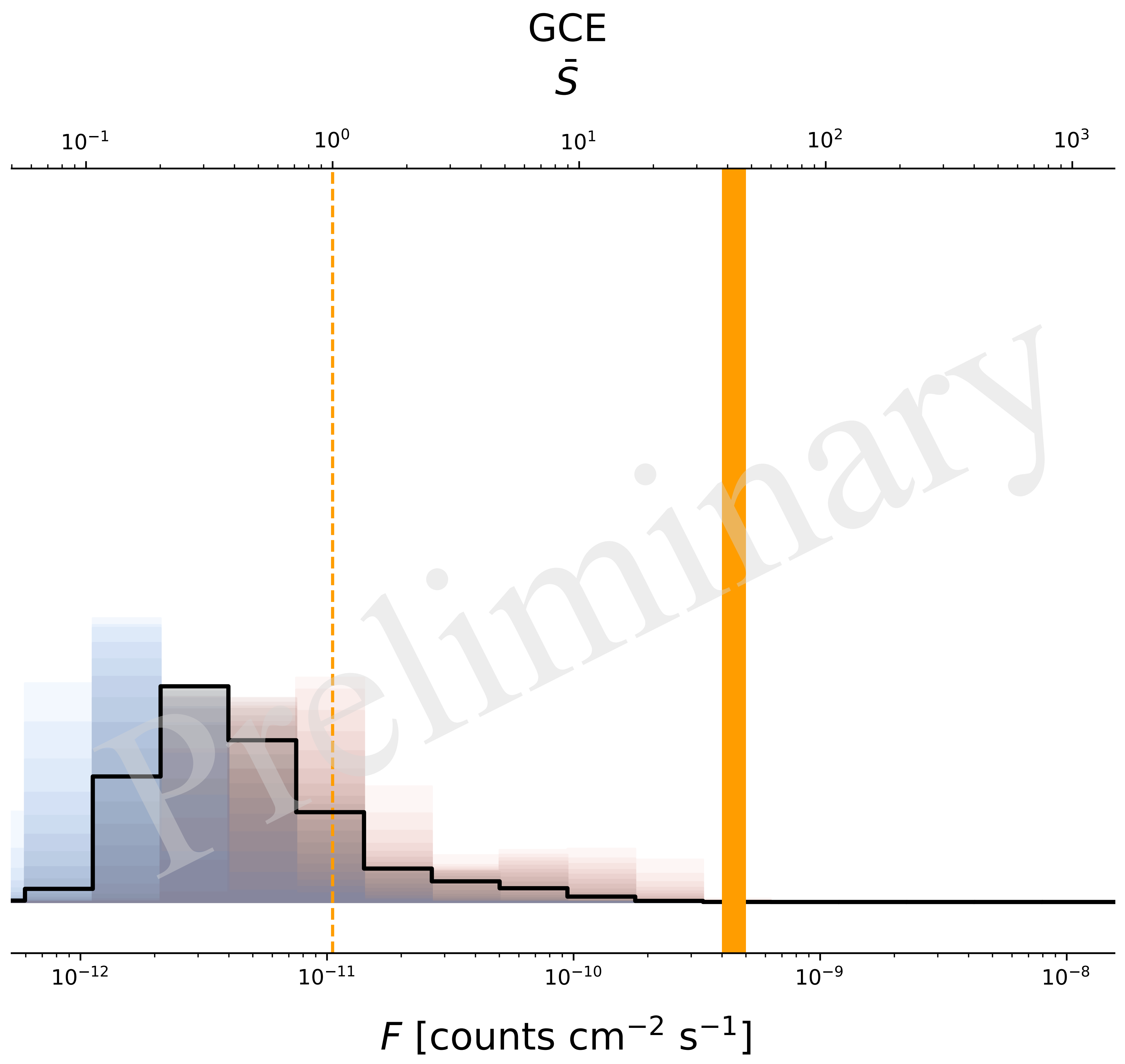
Next Steps

1. Further systematic checks
2. Is the CNN the best NN for the task?
    - Consider a transformer neural network
3. Other ways to validate the NN method
    - Test in the high latitude sky

Key Take Aways

- Adding energy dependent information to the NN gives a similar energy spectrum as energy independent NN but…

- the predicted average brightness of sources is dimmer than in the independent case.



Extra Slides

The full loss functions used to train the NN

Spectra:

$$\mathscr{L}(\tilde{\mathbf{y}}_\theta(\mathbf{x}), \mathbf{y}) = \sum_{t=1}^{T} \sum_{e=1}^{E} \left[\frac{1}{2(\tilde{\sigma}_\theta^{t,e}(\mathbf{x}))^2} (\tilde{y}_\theta^{t,e}(\mathbf{x}) - y^{t,e})^2 + \frac{\ln[(\tilde{\sigma}_\theta^{t,e}(\mathbf{x}))^2]}{2} \right]$$

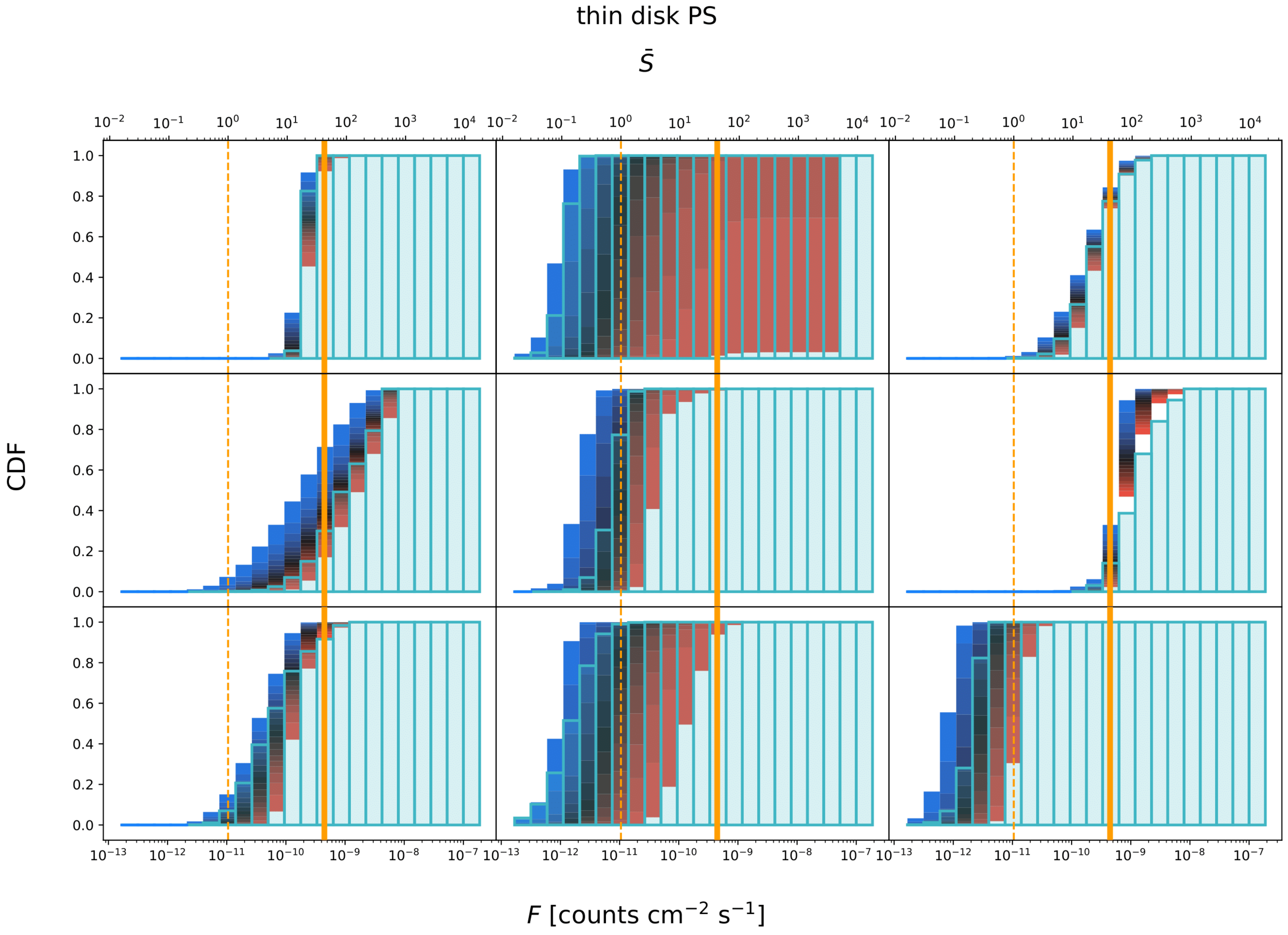
Histograms:

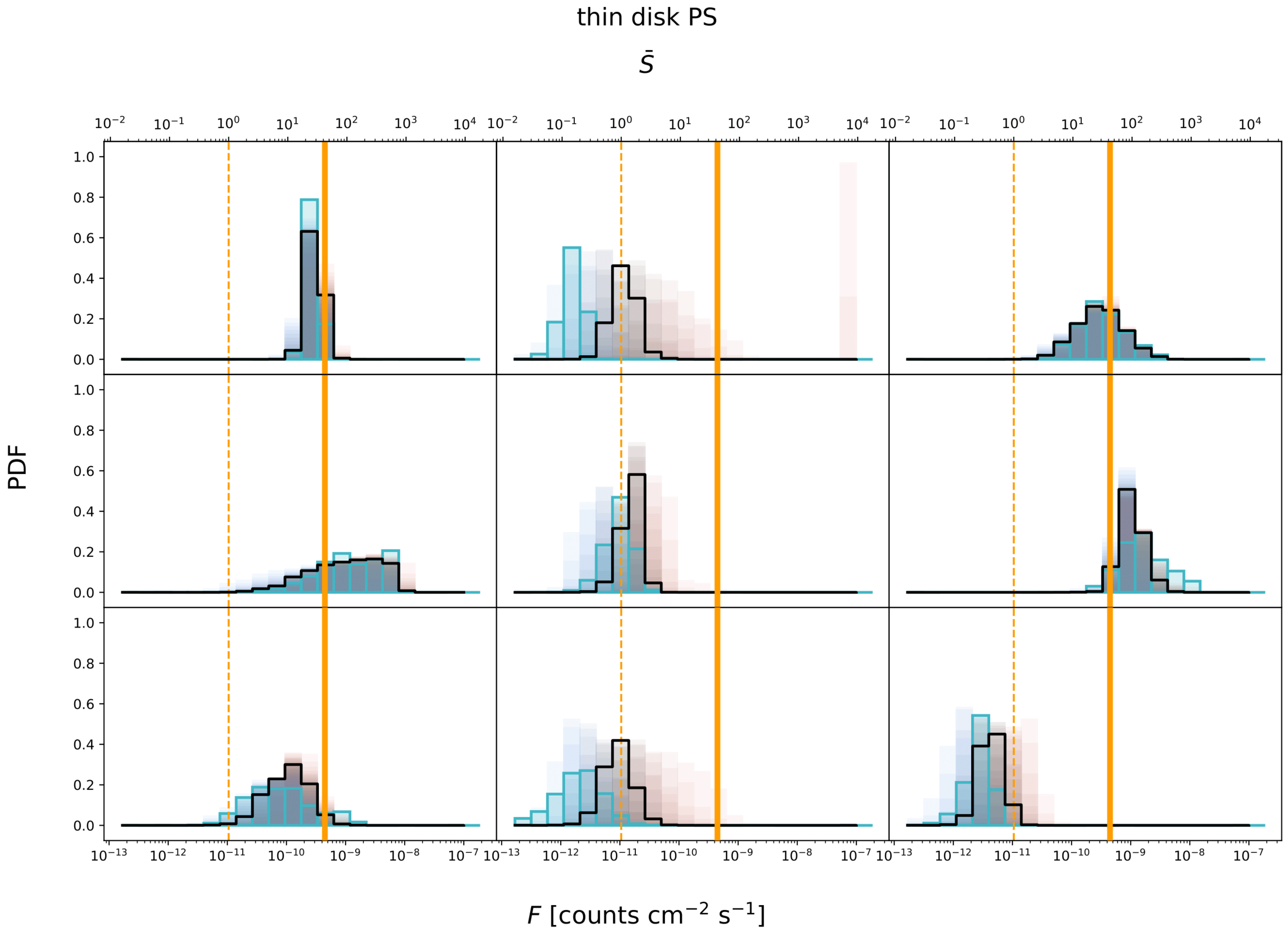
$$\mathscr{L}^\tau(\tilde{\mathbf{u}}_\theta(\mathbf{x}, \tau), \mathbf{u}) = \frac{1}{N} \sum_{n=1}^{N} [(\tilde{U}_\theta^n(\mathbf{x}, \tau) - U^n)(\tau - I[\tilde{U}_\theta^n(\mathbf{x}, \tau) < U^n])]$$

e=energy bin t=template $\tilde{\mathbf{y}}_\theta(\mathbf{x})$ = estimated SCD of all models $\mathbf{y}$= true SCD of all models	$\tilde{\sigma}_\theta^{t,e}(\mathbf{x})$= statistical uncertainty $\tilde{y}_\theta^{t,e}$= estimated counts per model, per bin $y^{t,e}$= true counts per model per bin	n= the probability bin τ= histogram quantile level $\tilde{\mathbf{u}}_\theta^n$= estimated SCD histogram for all models $\mathbf{u}$= true SCD histogram for all models	$\tilde{U}_\theta^n(\mathbf{x}, \tau)$= Estimated cumulative histogram U^n= True cumulative histogram $I[\,..\,]$= the indicator function

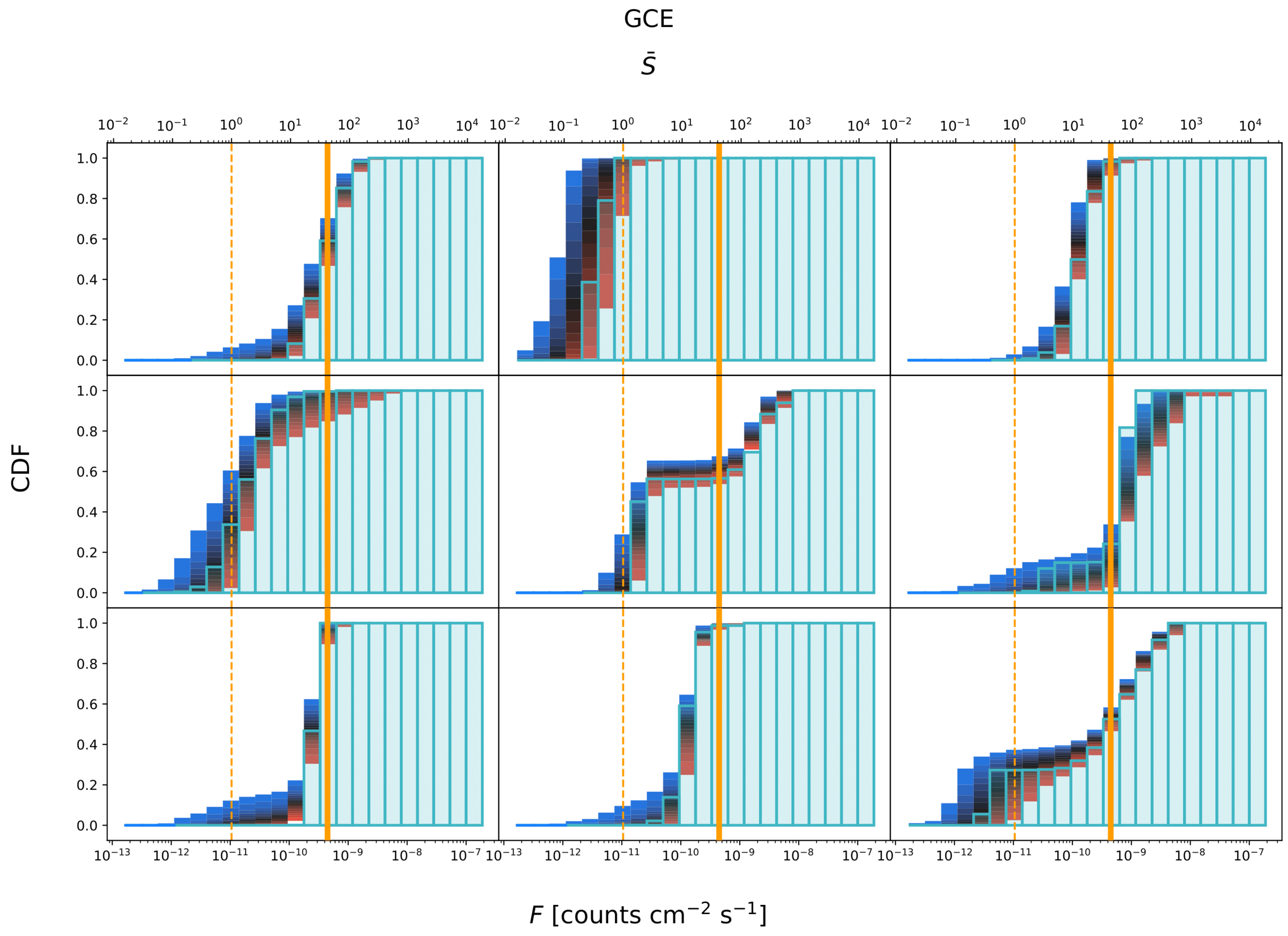
Florian List. *The earth mover's pinball loss: Quantiles for histogram-valued regression*, 2021.

The NN fits for point sources in the disk, in addition to in the GCE.





NN actually learns in CDF space, current results from the NN



Related Works

- Florian List. The earth mover's pinball loss: Quantiles for histogram-valued regression, 2021.
- Florian List, Nicholas L. Rodd, and Geraint F. Lewis. Extracting the galactic center excess'
  source-count distribution with neural nets. Physical Review D, 104(12), December 2021.
- Florian Wolf, Florian List, Nicholas L. Rodd, and Oliver Hahn. A deep learning framework for
  jointly extracting spectra and source-count distributions in astronomy, 2024.