

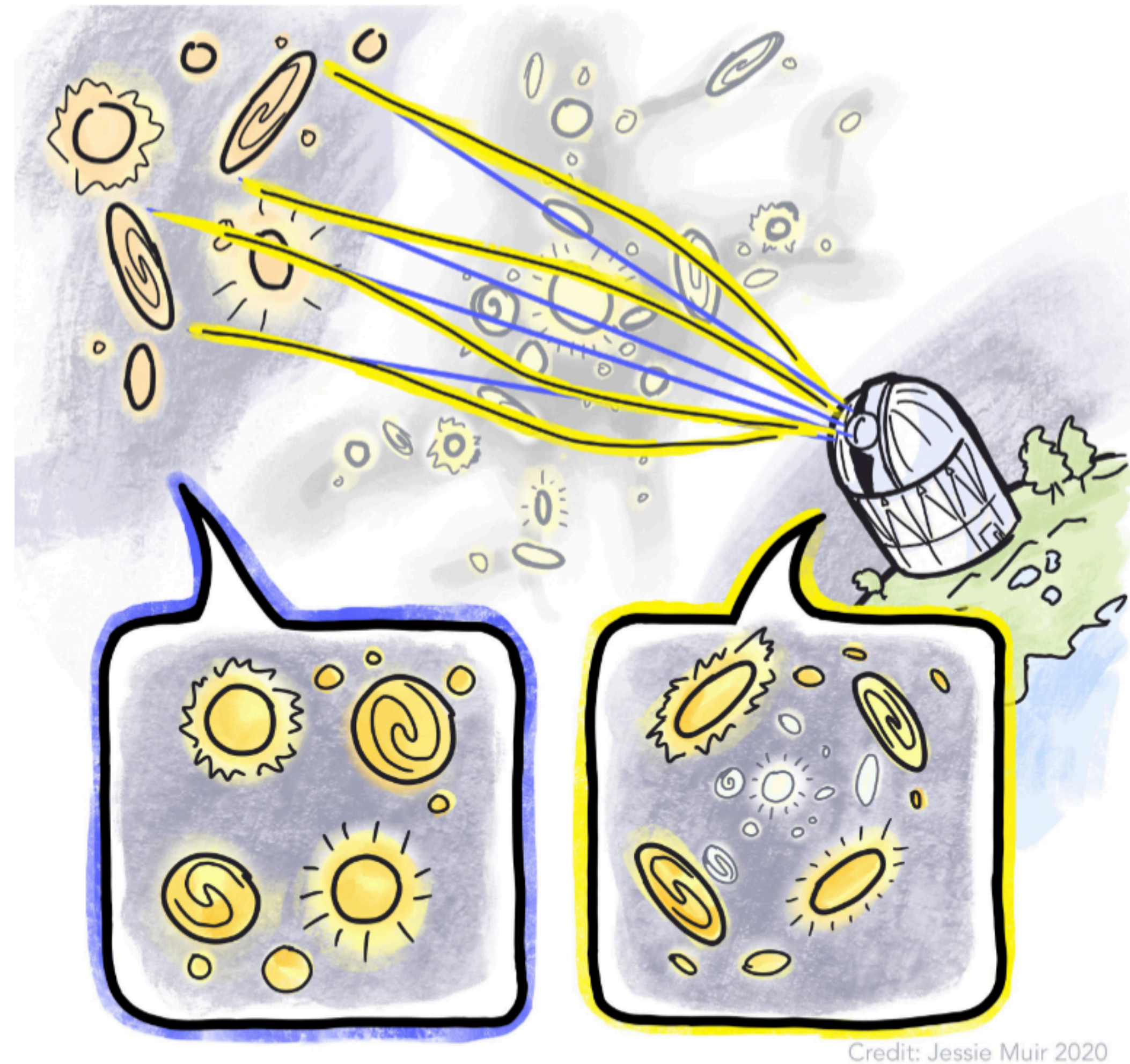
Simulation-Based Inference for Probabilistic Galaxy Detection and Deblending

Ismael Mendoza

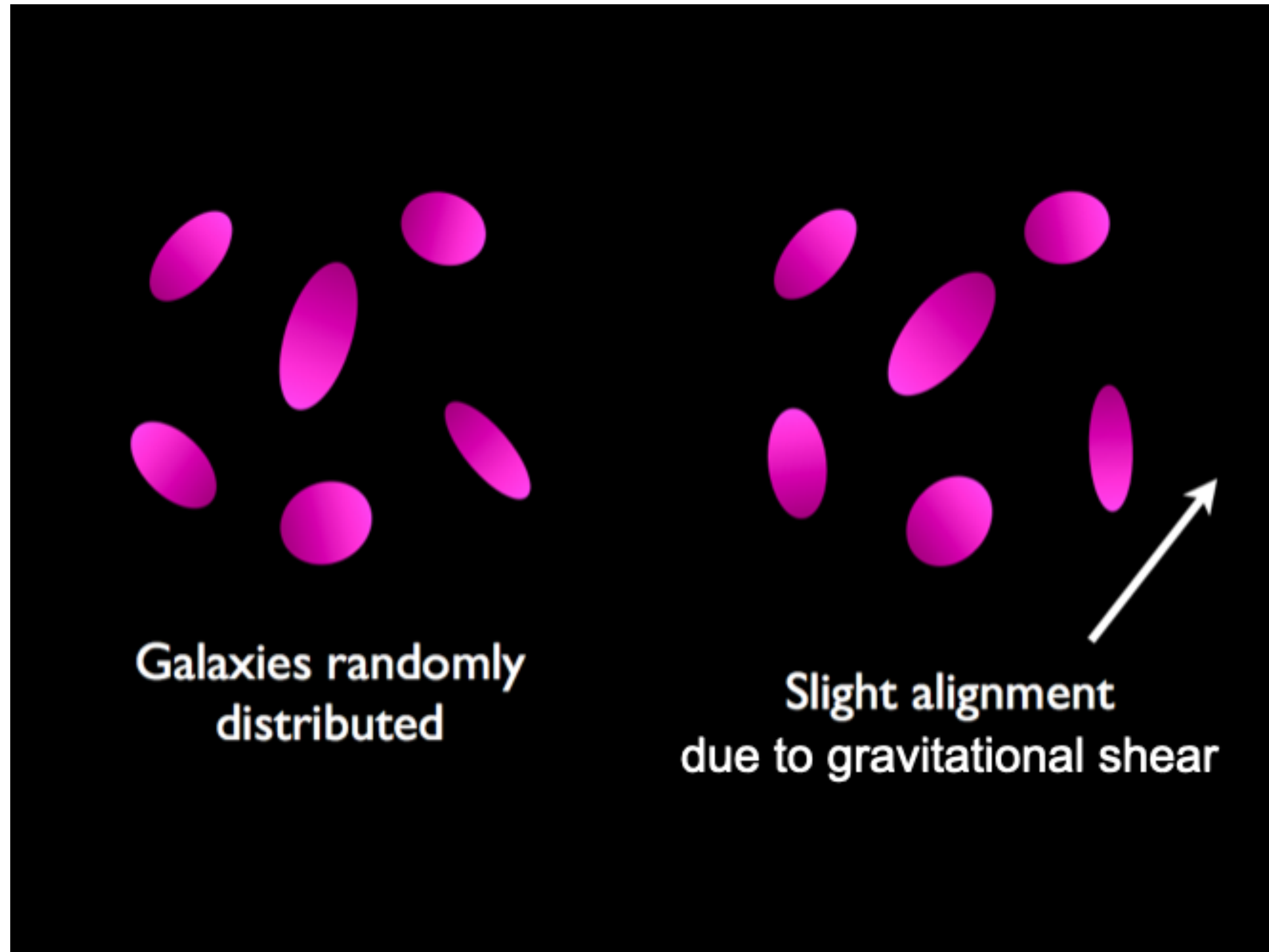
Postdoc @ University of Maryland

Context

Weak Lensing: Cosmic structures distort shapes of distant light sources

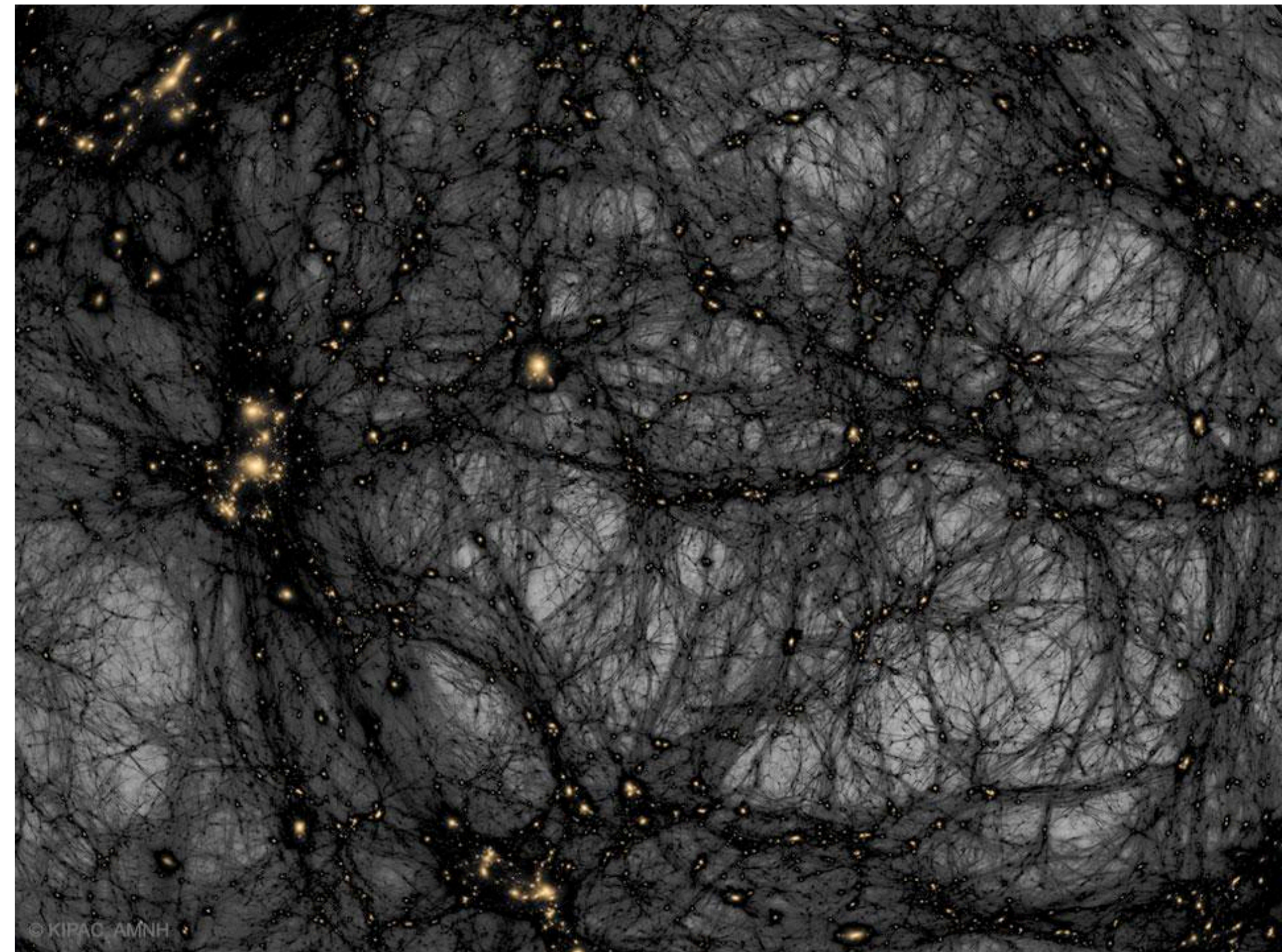


Cosmic Shear: Spatial correlation of galaxy shapes



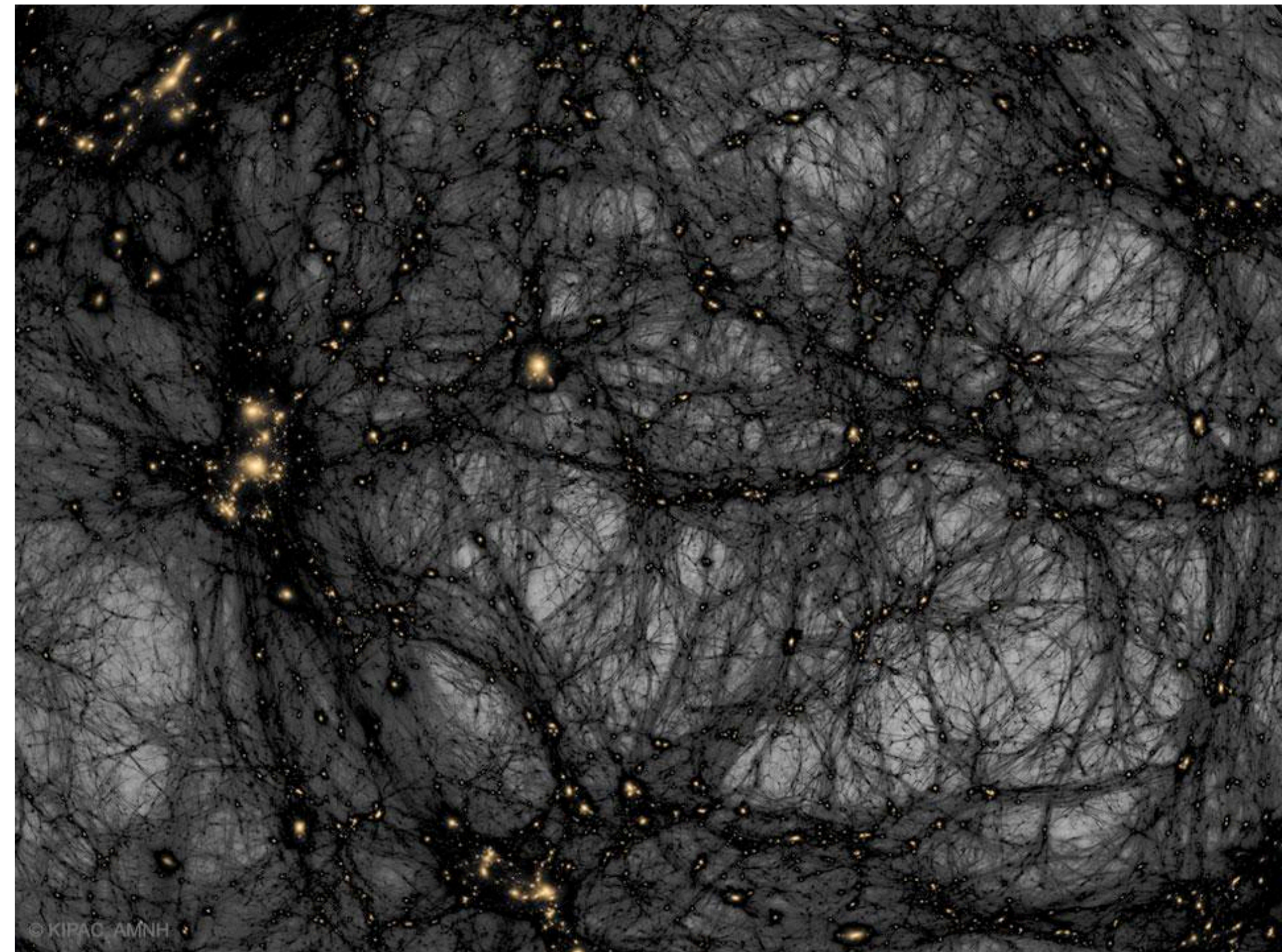
Cosmic shear traces the structure in our Universe

- Shear directly depends on the amount of mass between us and a distant light source.



Cosmic shear traces the structure in our Universe

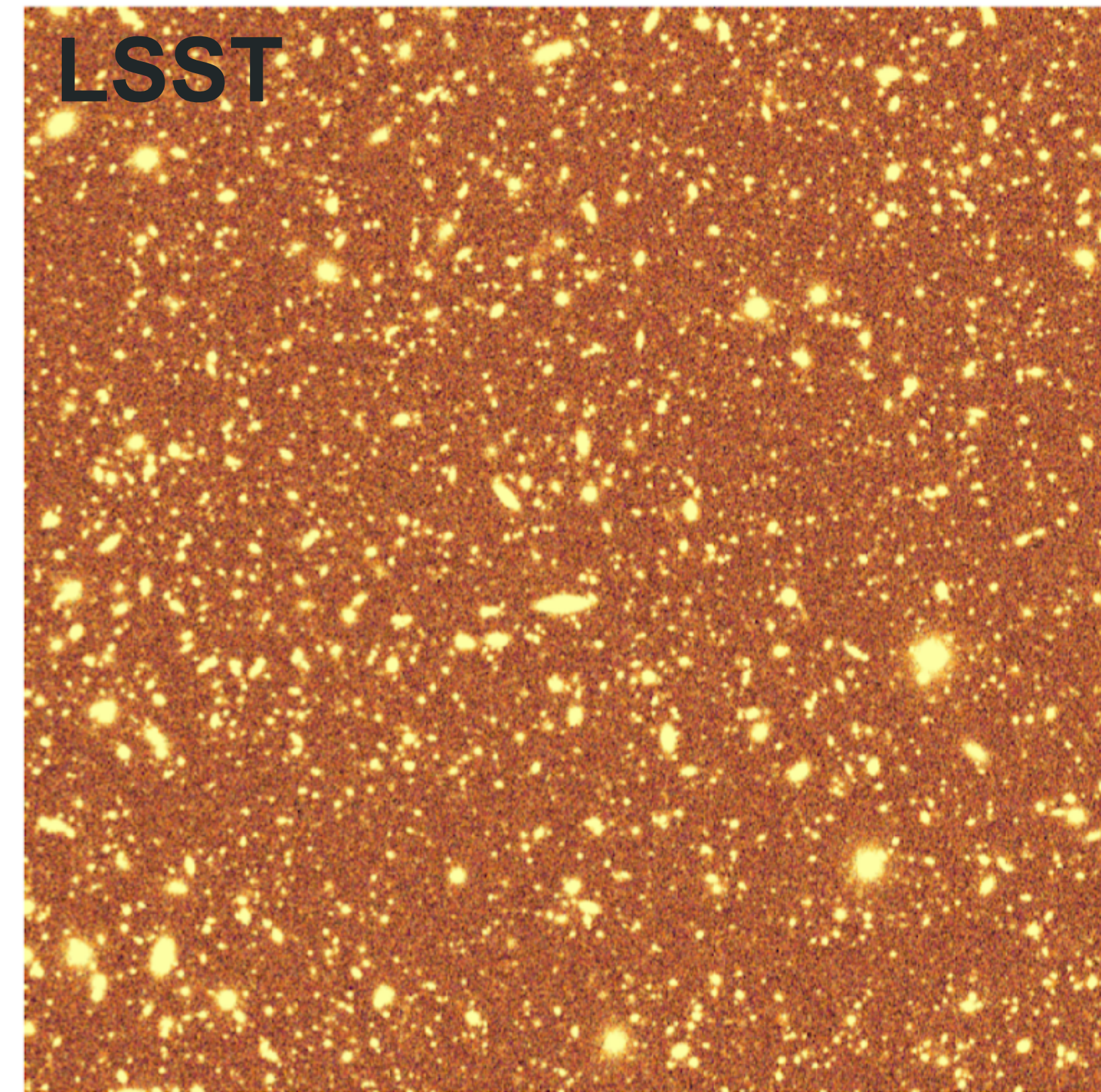
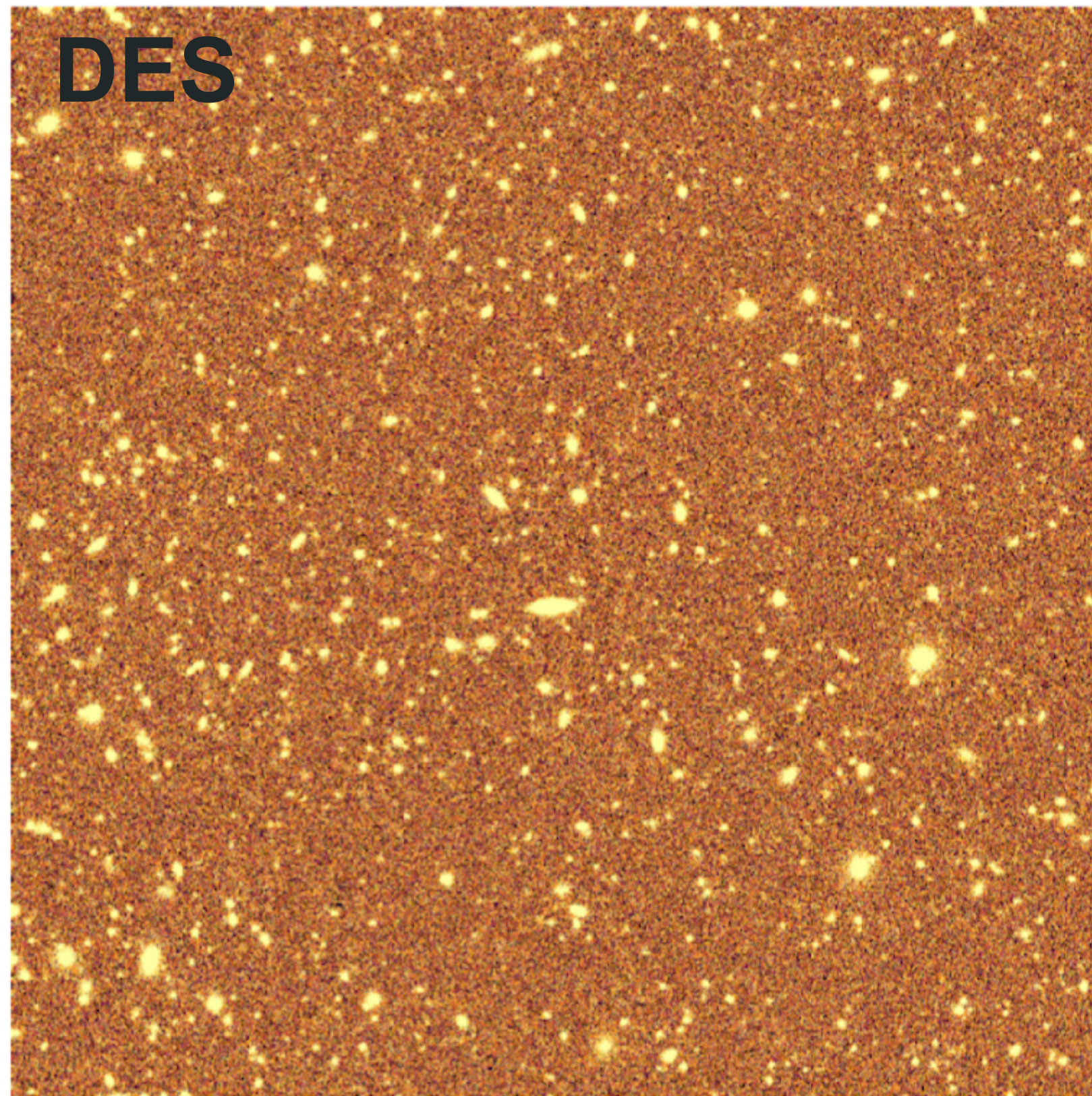
- Shear directly depends on the amount of mass between us and a distant light source.



Accurately measuring galaxy shapes is essential for cosmology!

Blending: the visual overlap of light sources

- Approximately 60% of objects used for lensing in **LSST** will be blended.
(Sanchez, [Mendoza](#), et al. 2021)

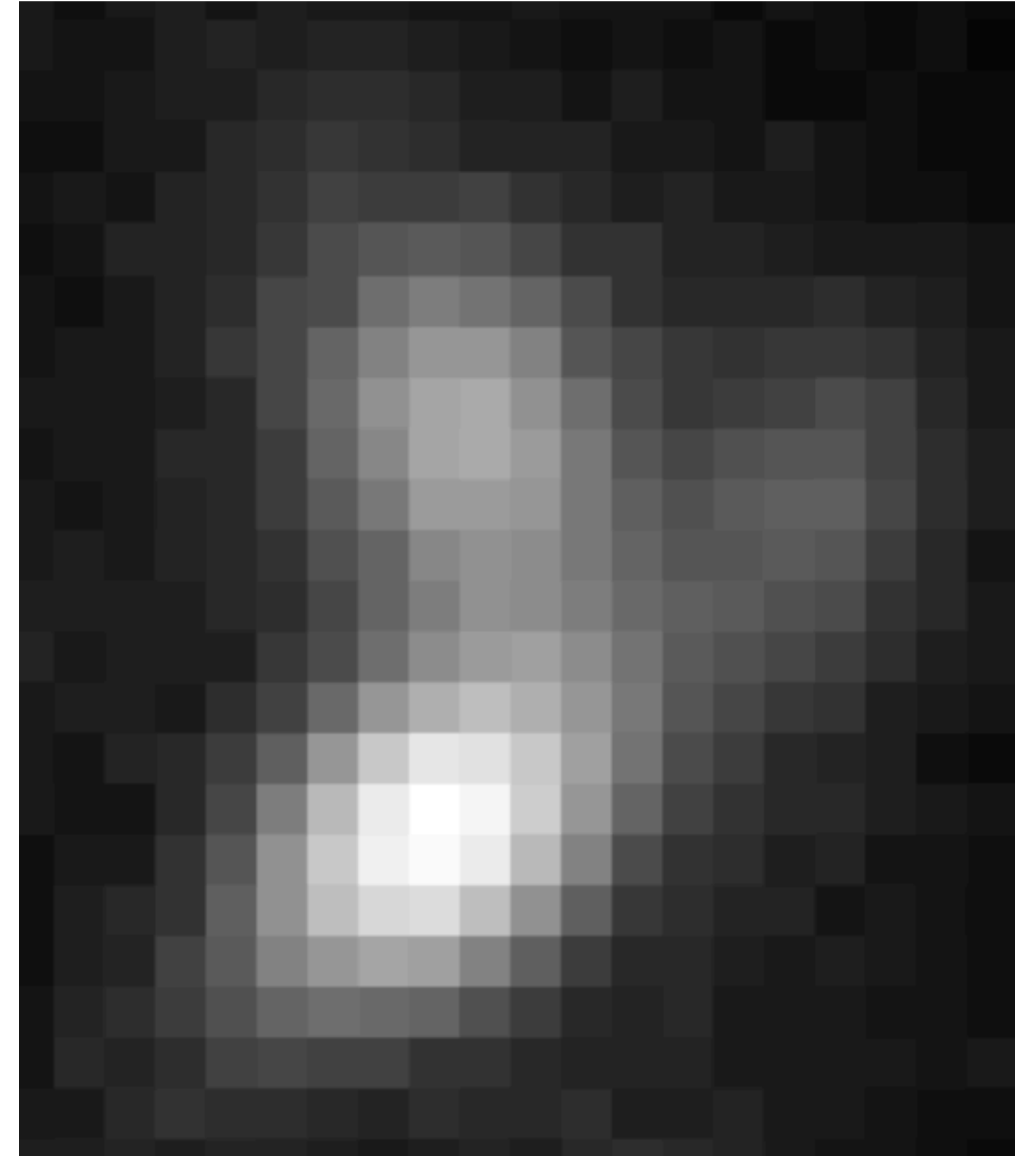


Blending is one of the leading systematics in modern surveys

Bayesian Light Source Separator (BLISS)

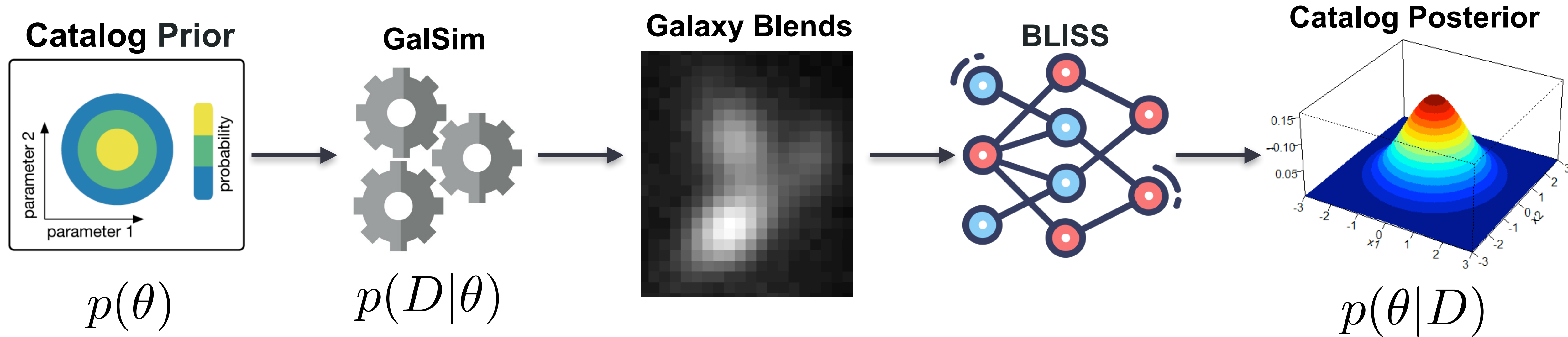
Motivation: Propagating uncertainty in measurements of blended galaxies

- Blending naturally introduces **significant uncertainty** in the galaxy measurement process:
 - How many sources are present?
 - Where are the sources located within the blend?
 - How should we assign the flux to each source?
- How do we characterize this uncertainty?
- How can we leverage this additional information?



Bayesian Light Source Separator (BLISS)

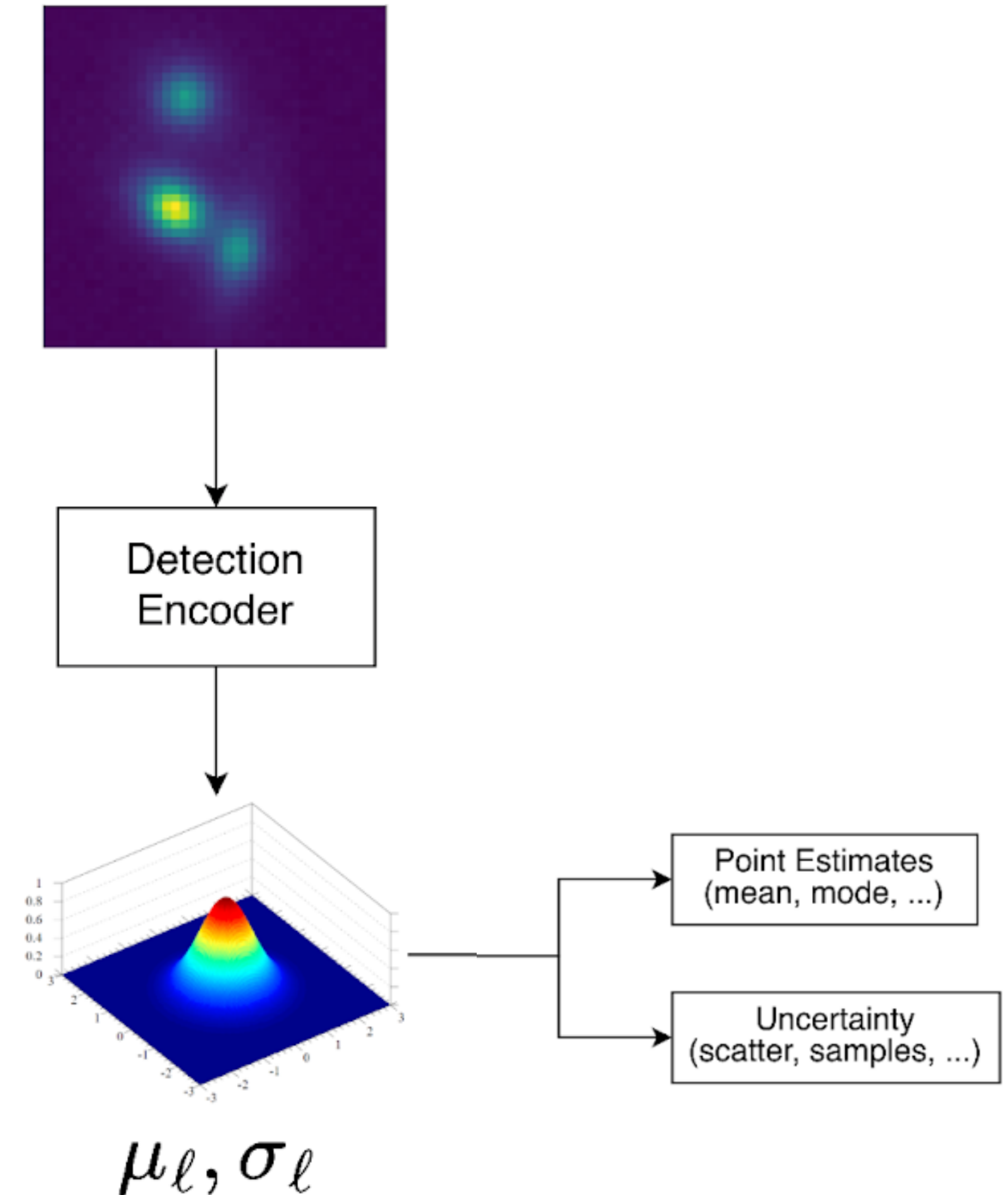
- Train a supervised ML model to develop a **probabilistic association** between data and underlying parameters (**SBI**).



- BLISS uses SBI to quantify the uncertainty in galaxy properties due to blending.

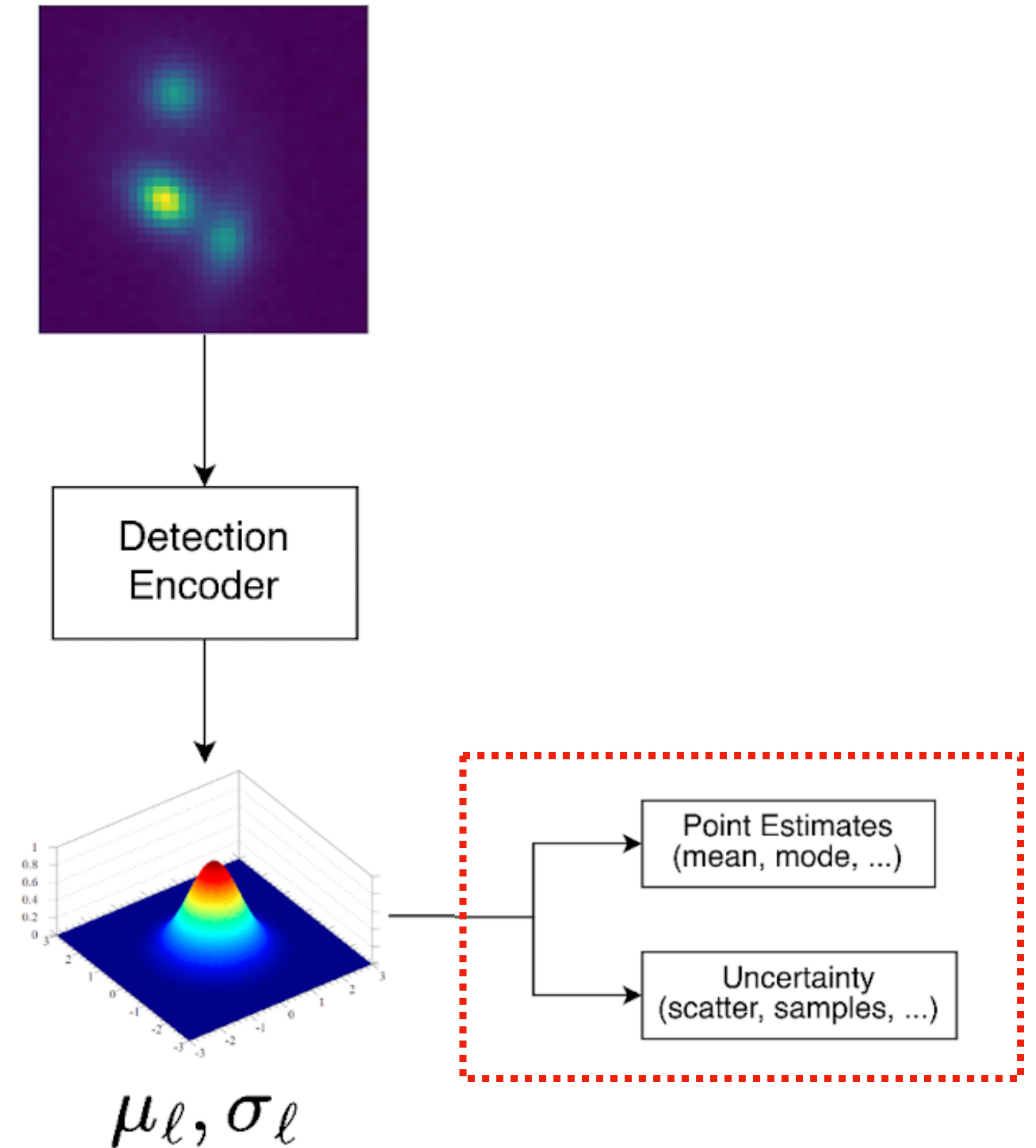
Networks parametrize distributions on galaxy properties

- Counts + Locations / Centroids
- Classification (Galaxy vs. Star)
- Galaxy morphology
 - Noiseless reconstructions of individual galaxies



Networks parametrize distributions on galaxy properties

- Counts + Locations / Centroids
- Classification (Galaxy vs. Star)
- Galaxy morphology
 - Noiseless reconstructions of individual galaxies

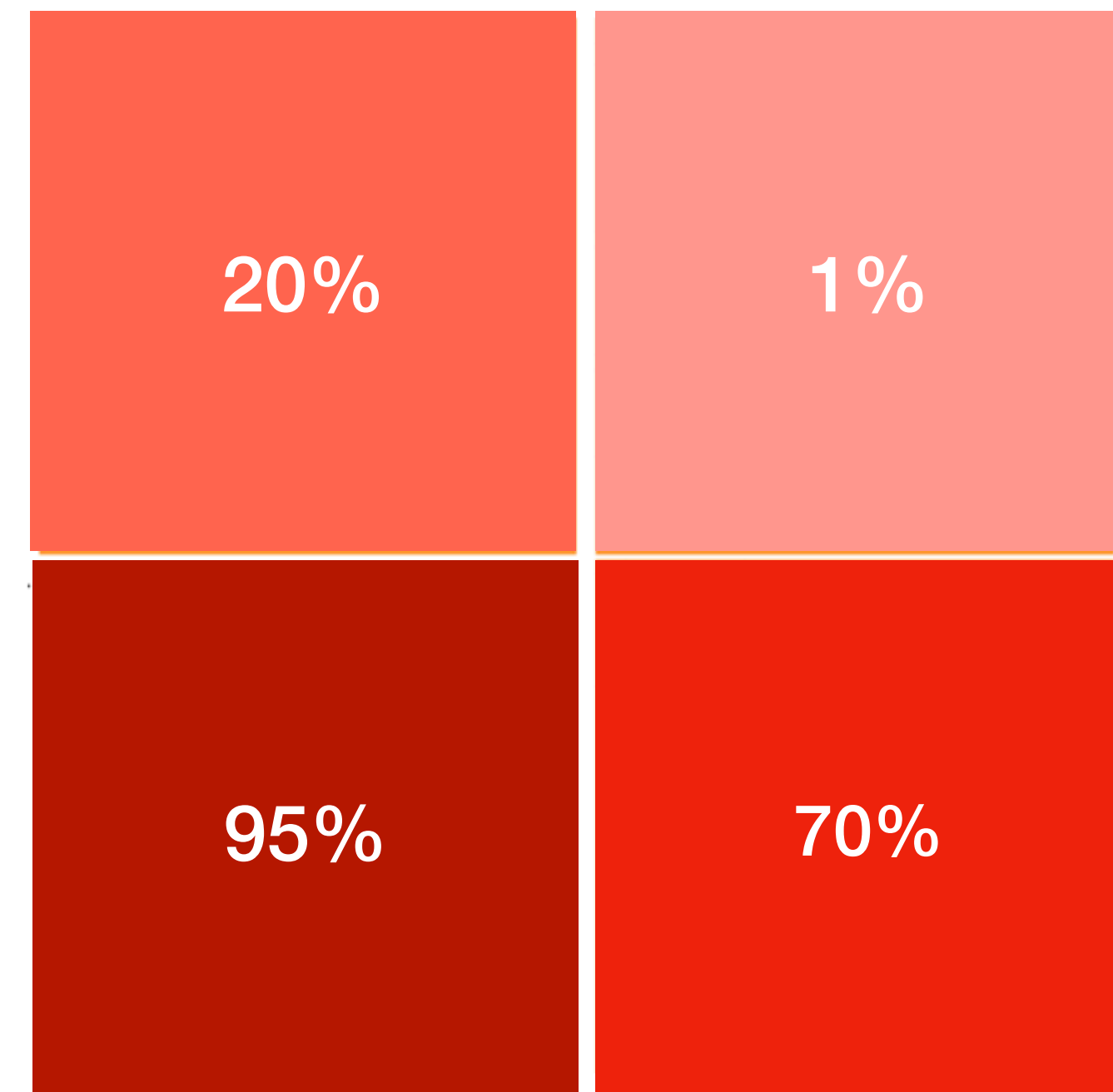
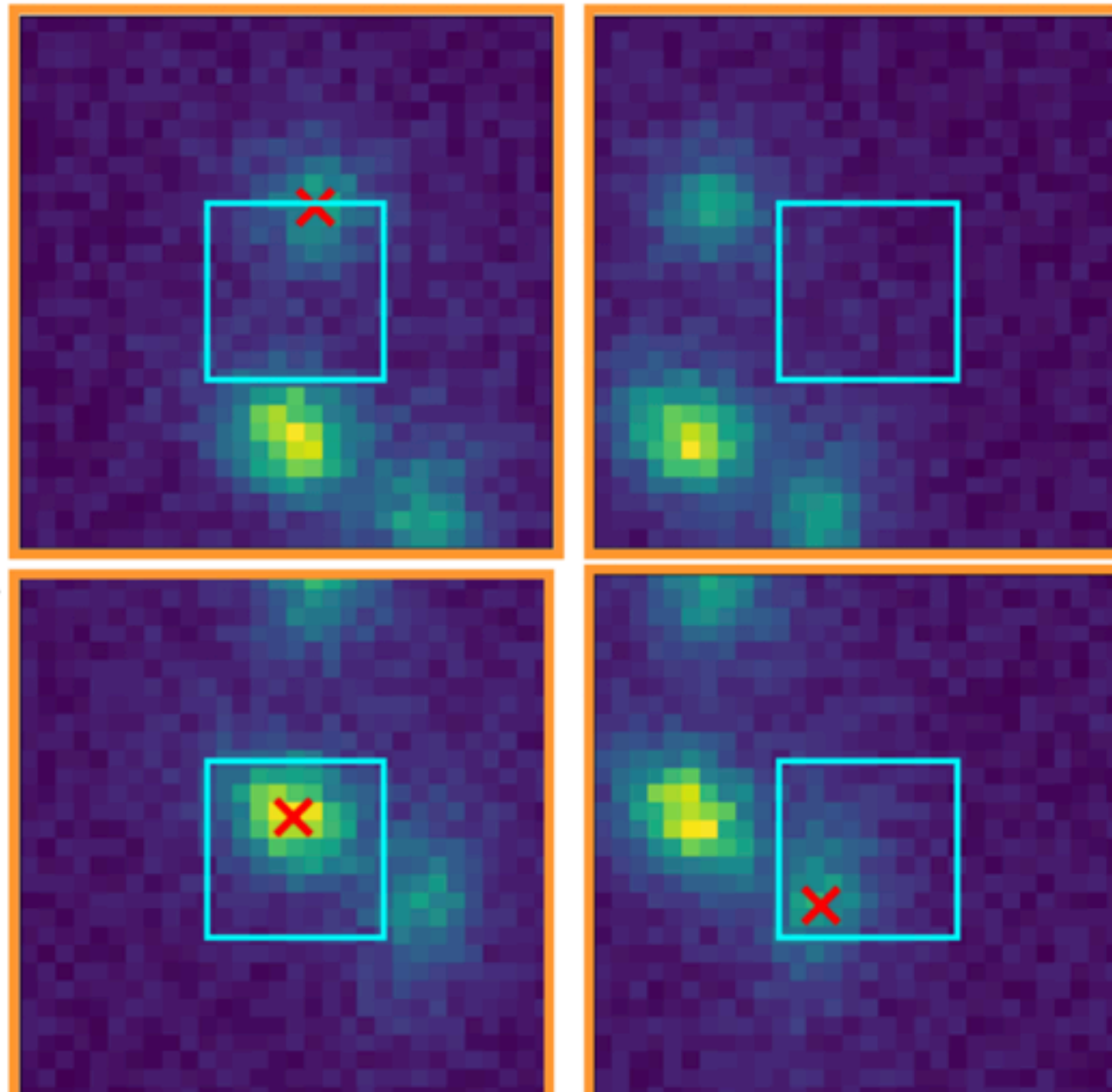


BLISS assigns a detection probability per tile

X True Centroids

□ Tiles

□ Padded Tiles

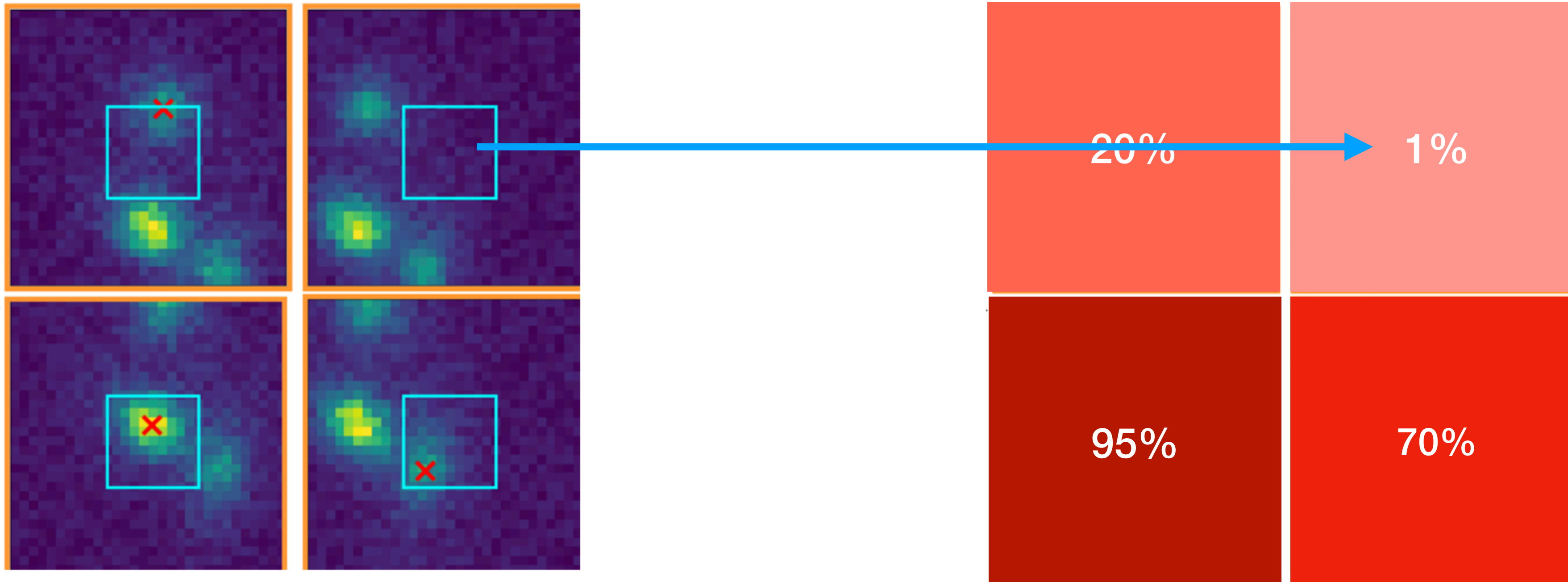


BLISS assigns a detection probability per tile

X True Centroids

□ Tiles

□ Padded Tiles

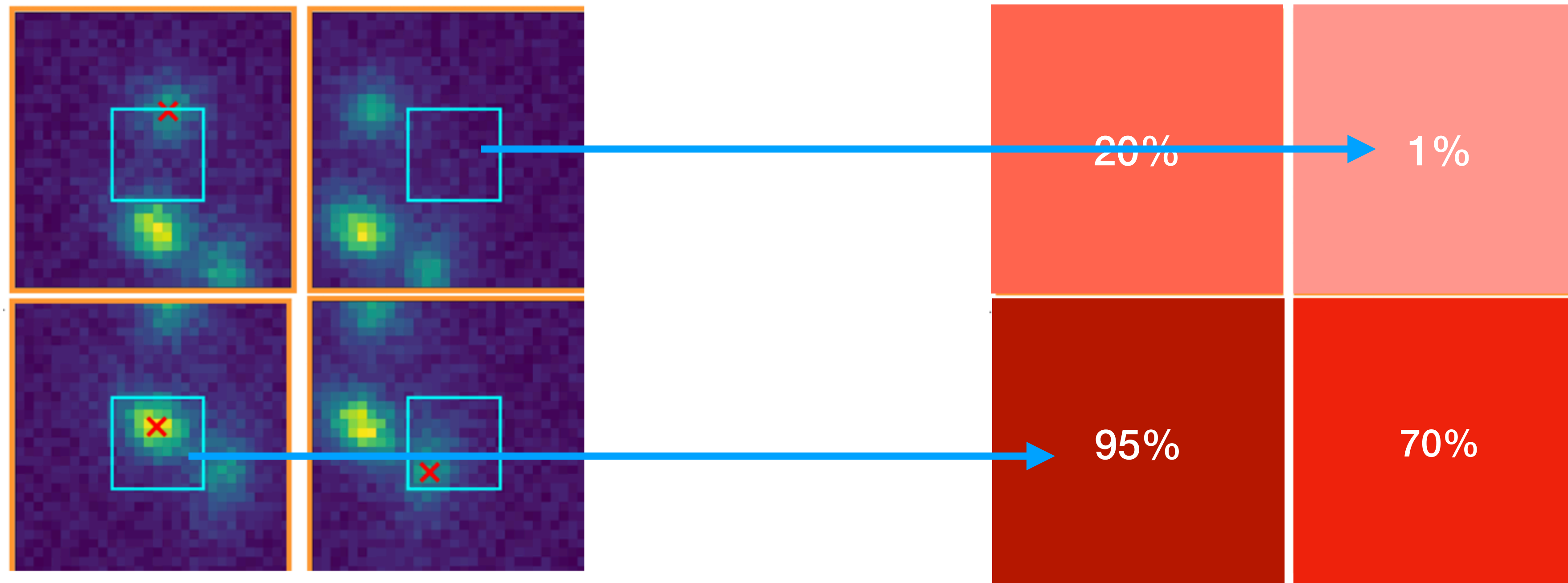


BLISS assigns a detection probability per tile

X True Centroids

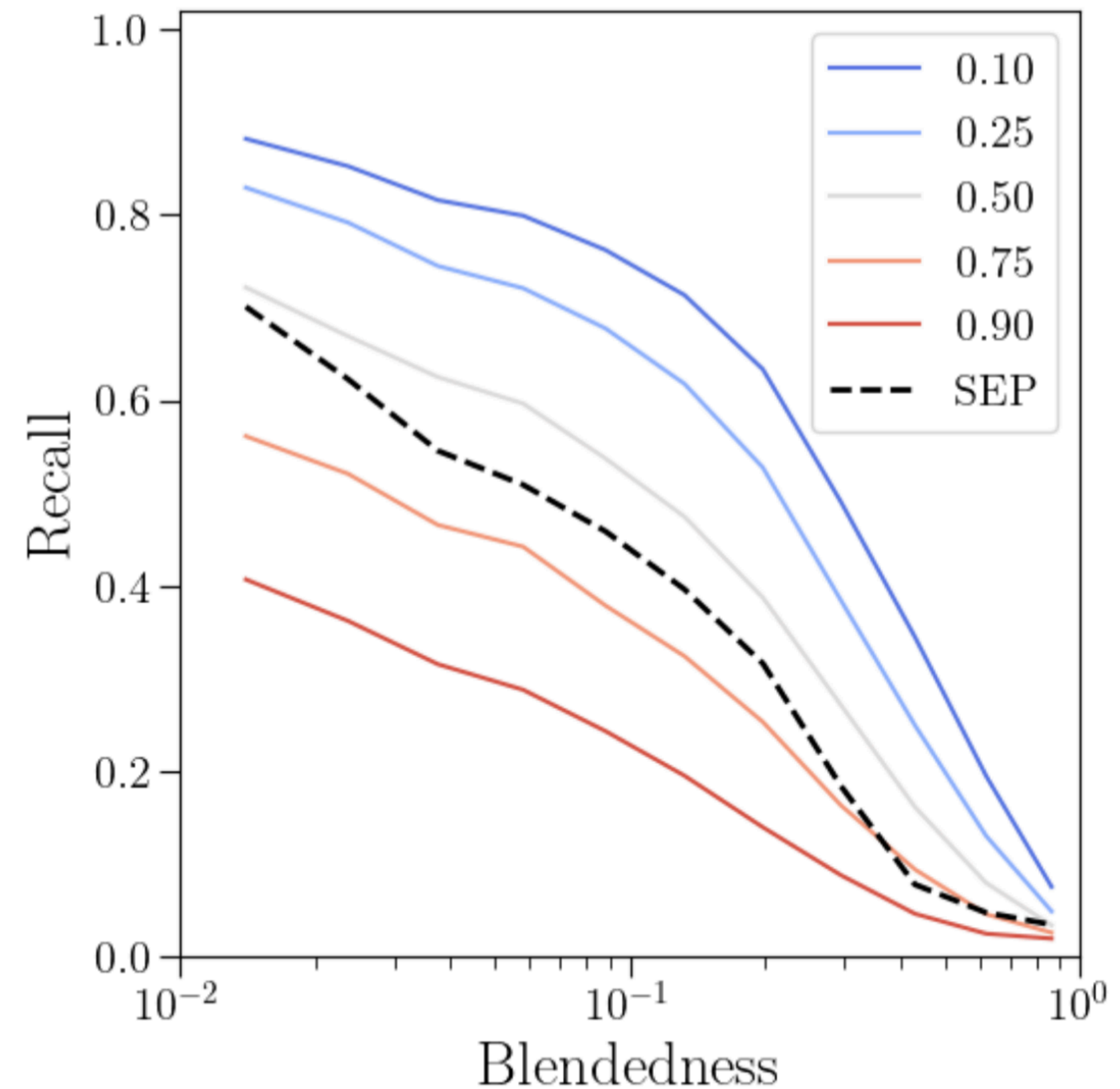
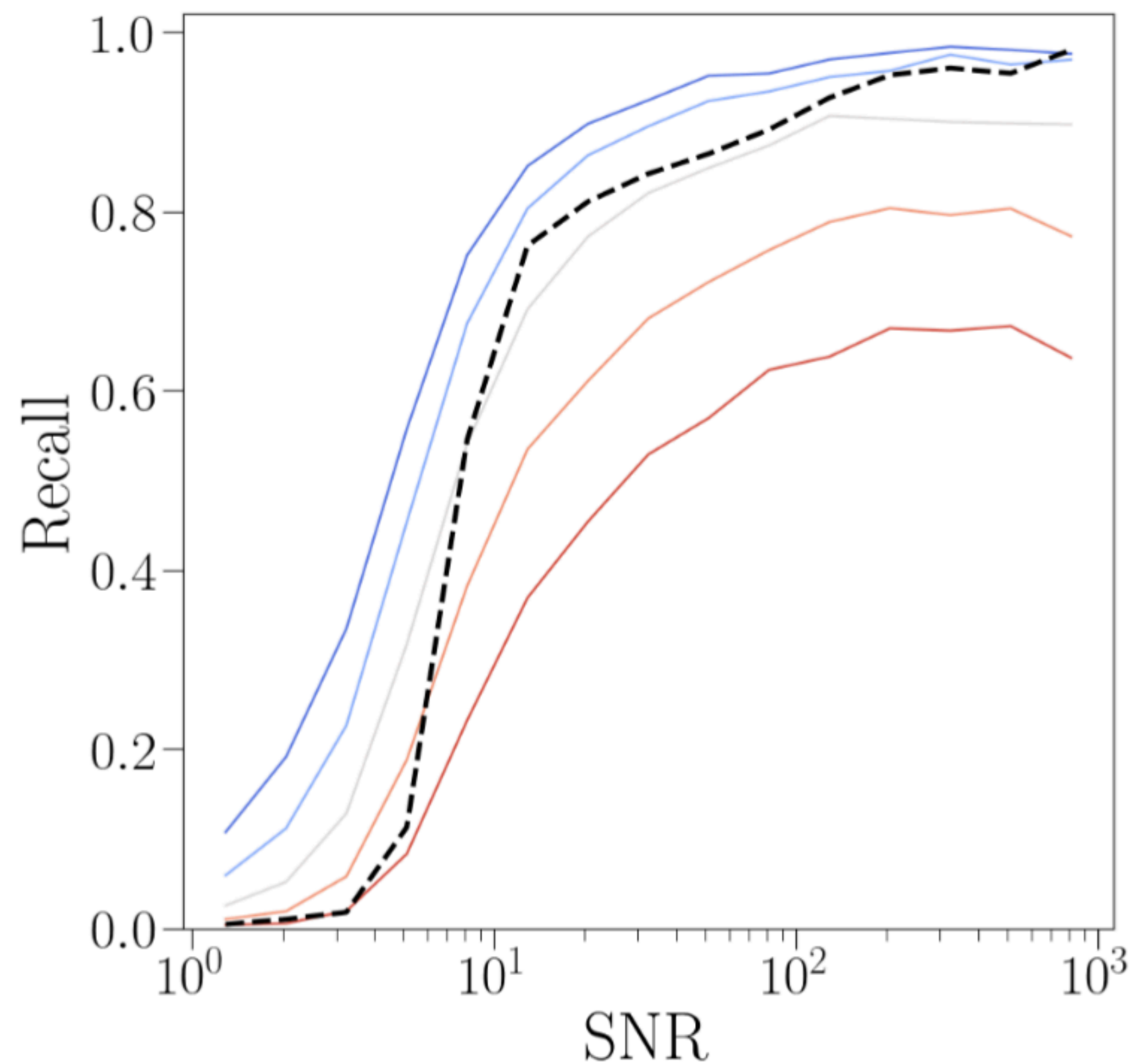
□ Tiles

□ Padded Tiles



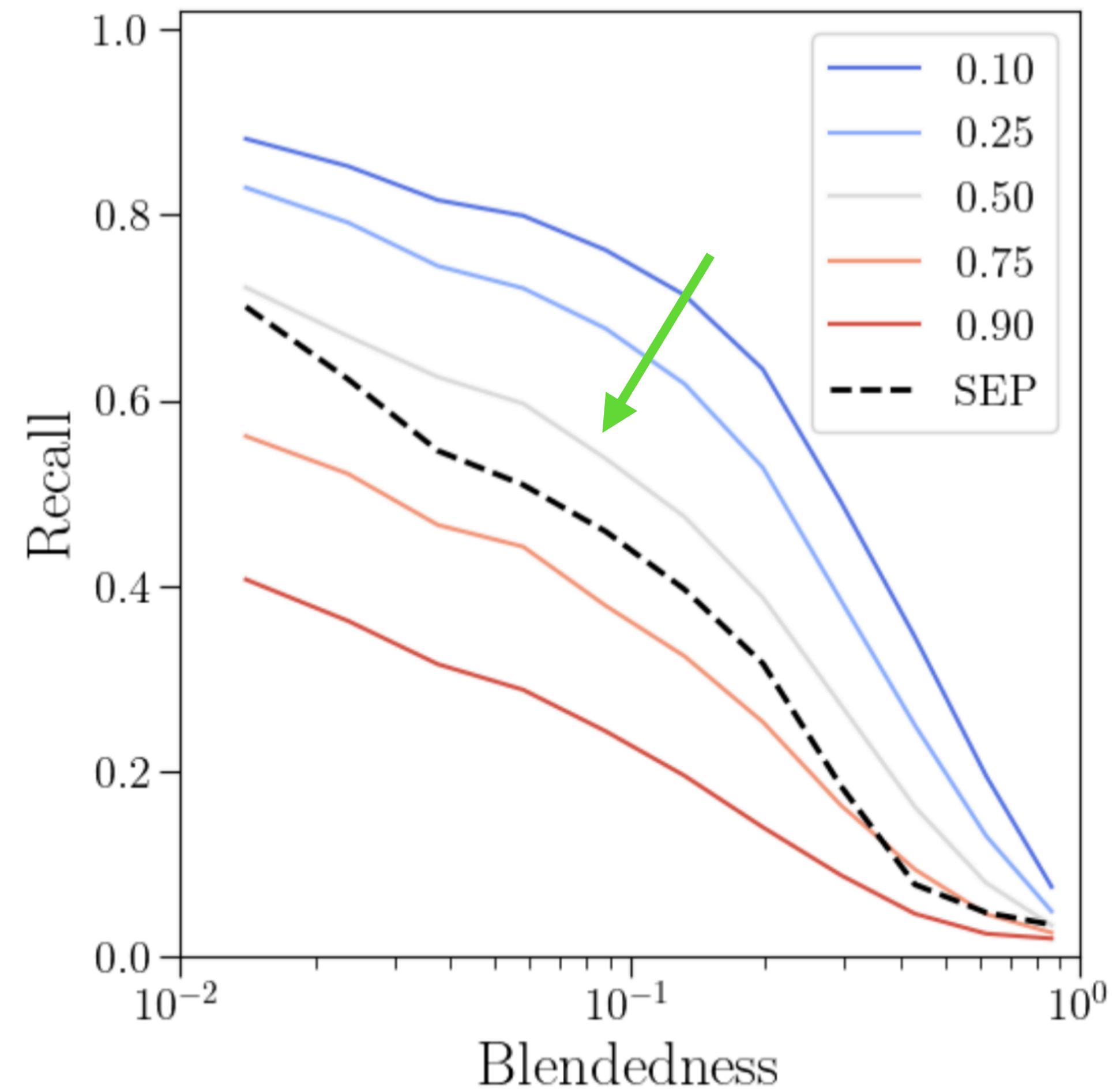
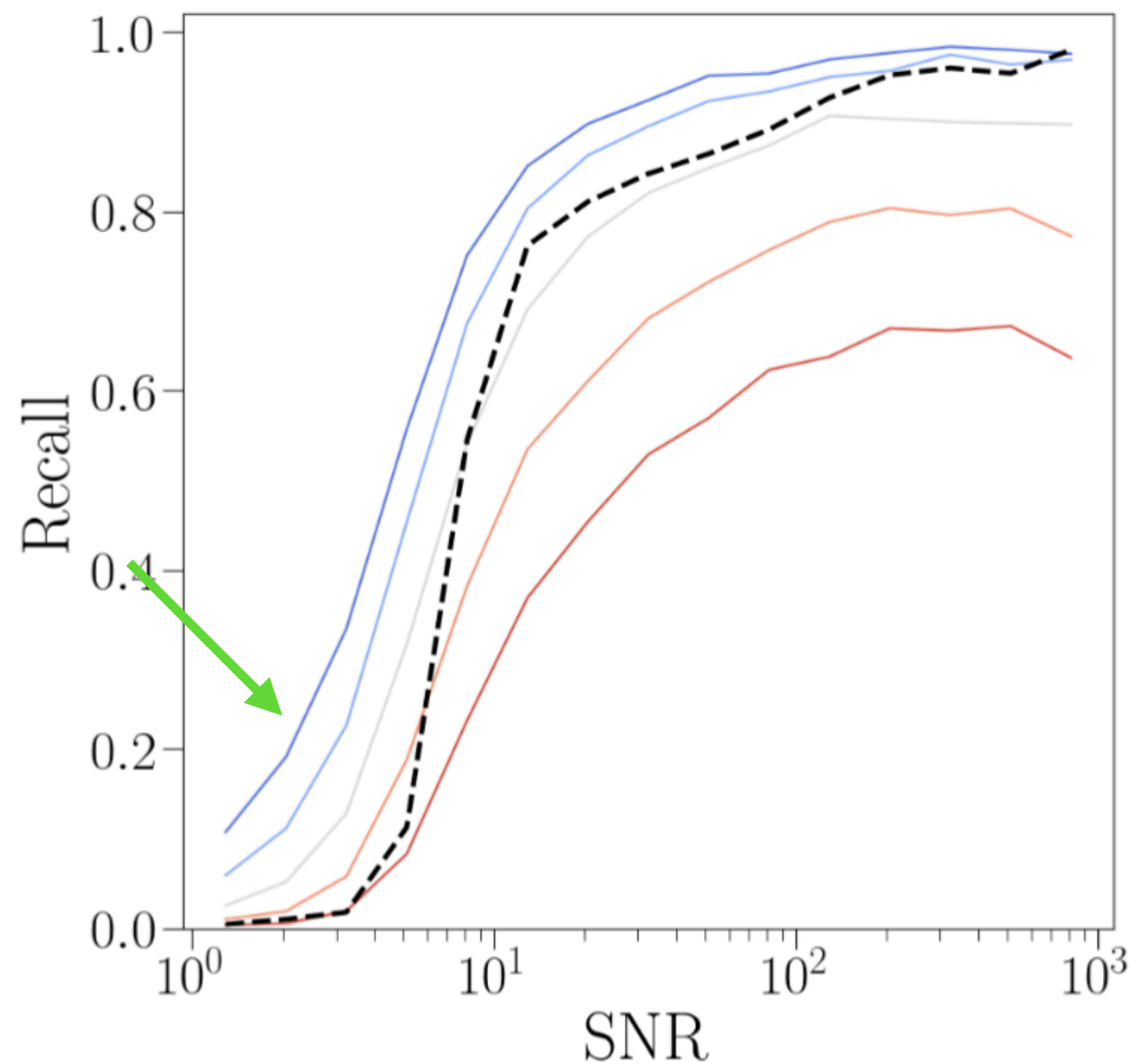
Detection “completeness” (recall)

- **Dataset:** LSST-like image simulations using parametric galaxies.
- BLISS detects more dim and blended sources than SourceExtractor (with chosen settings).



Detection “completeness” (recall)

- **Dataset:** LSST-like image simulations using parametric galaxies.
- BLISS detects more dim and blended sources than SourceExtractor (with chosen settings).

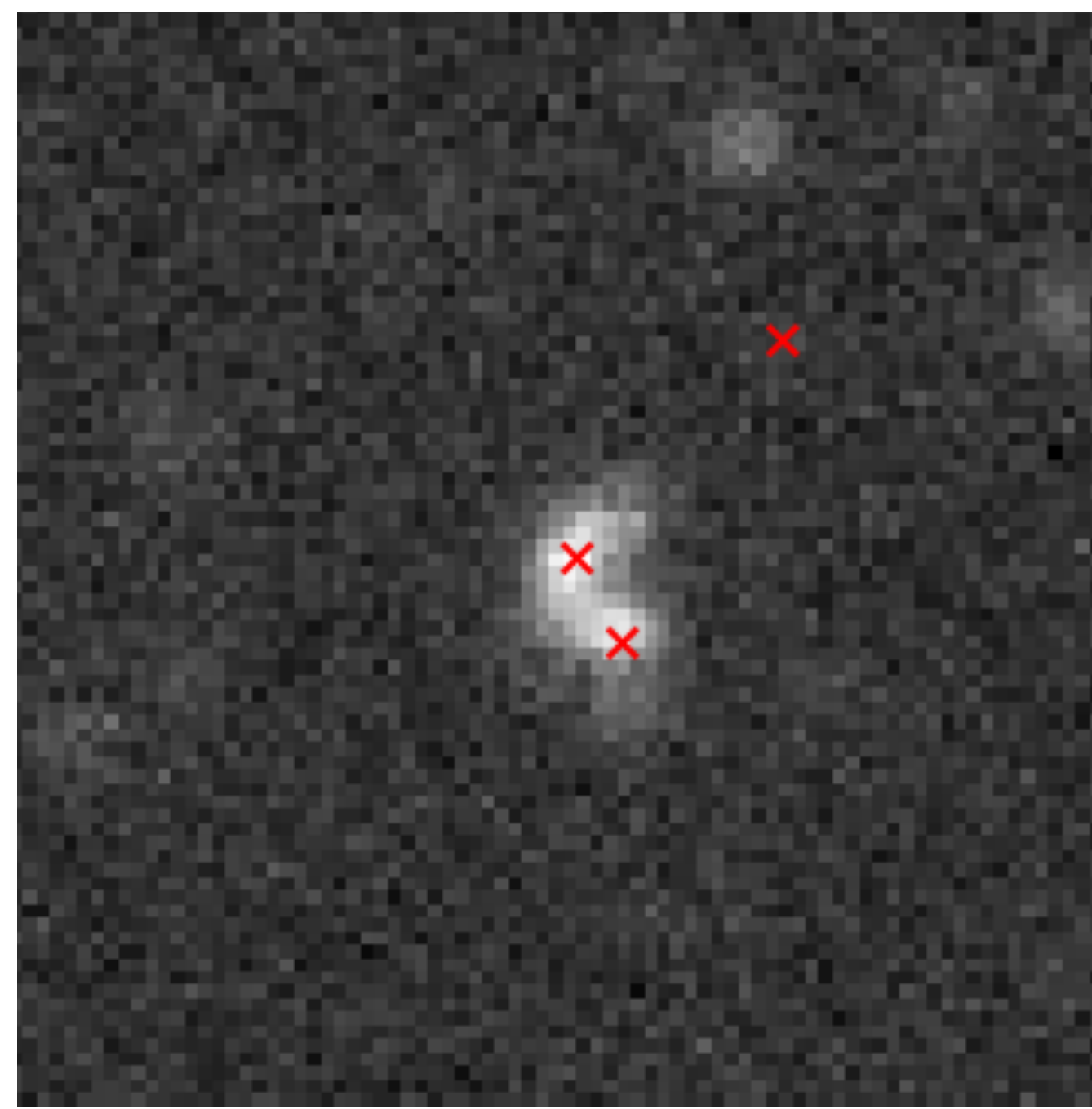


Deblending Procedure for measuring flux

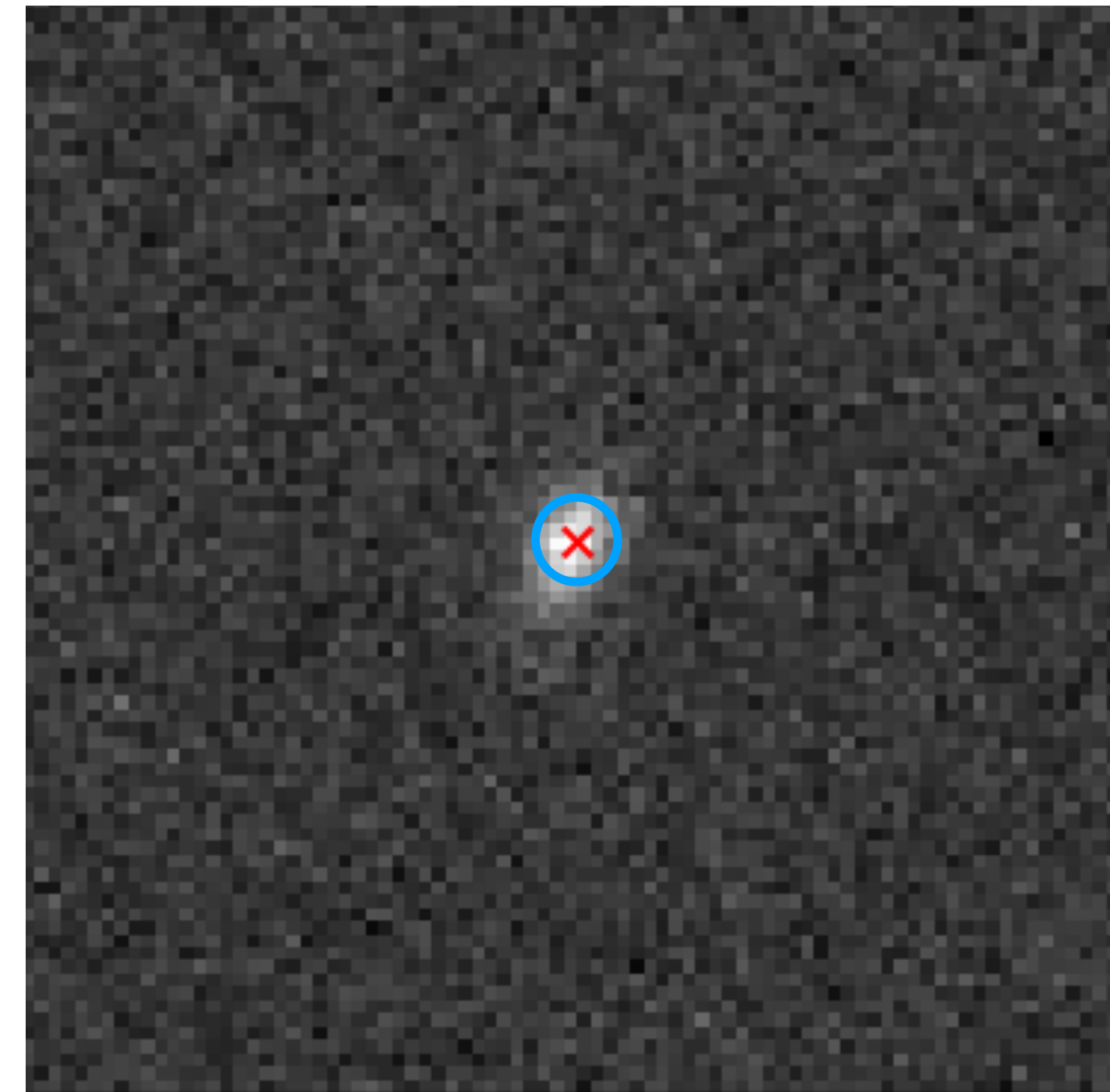
- Deblending for flux measurement involves reconstructing all detected galaxies and subtracting all except one from the original image.



Original blend image

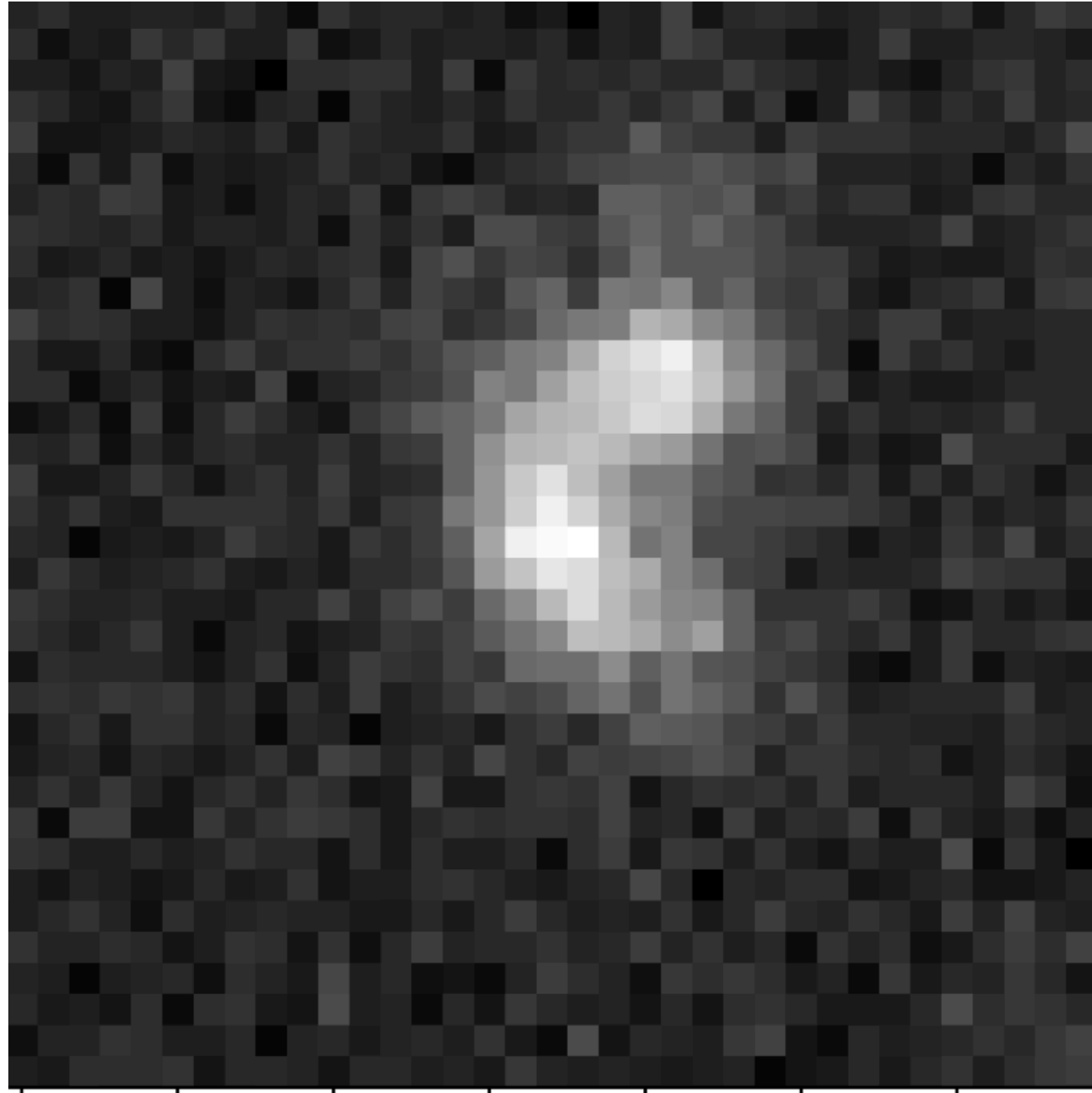


Run detection
Run deblender with
detections as input

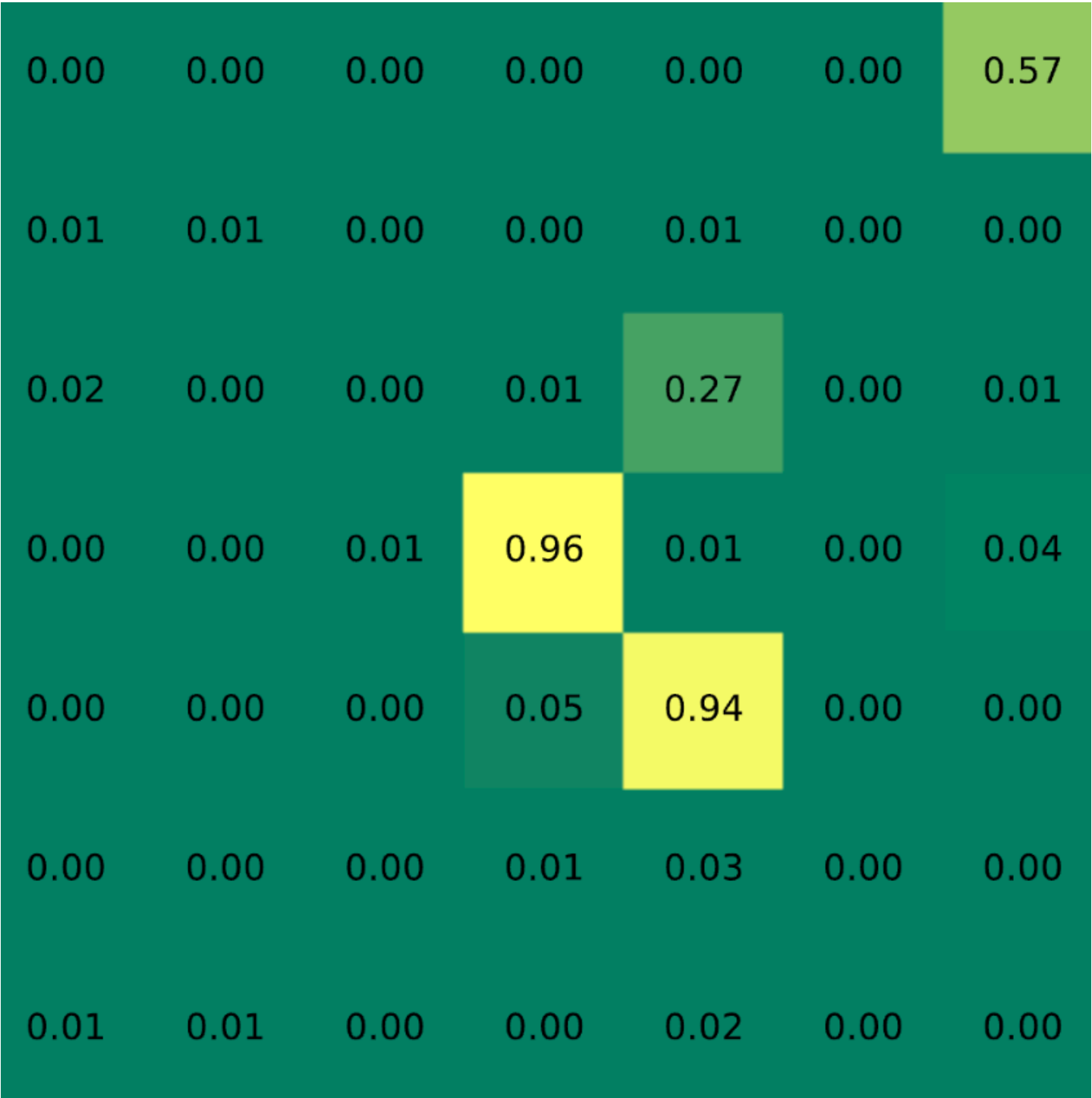
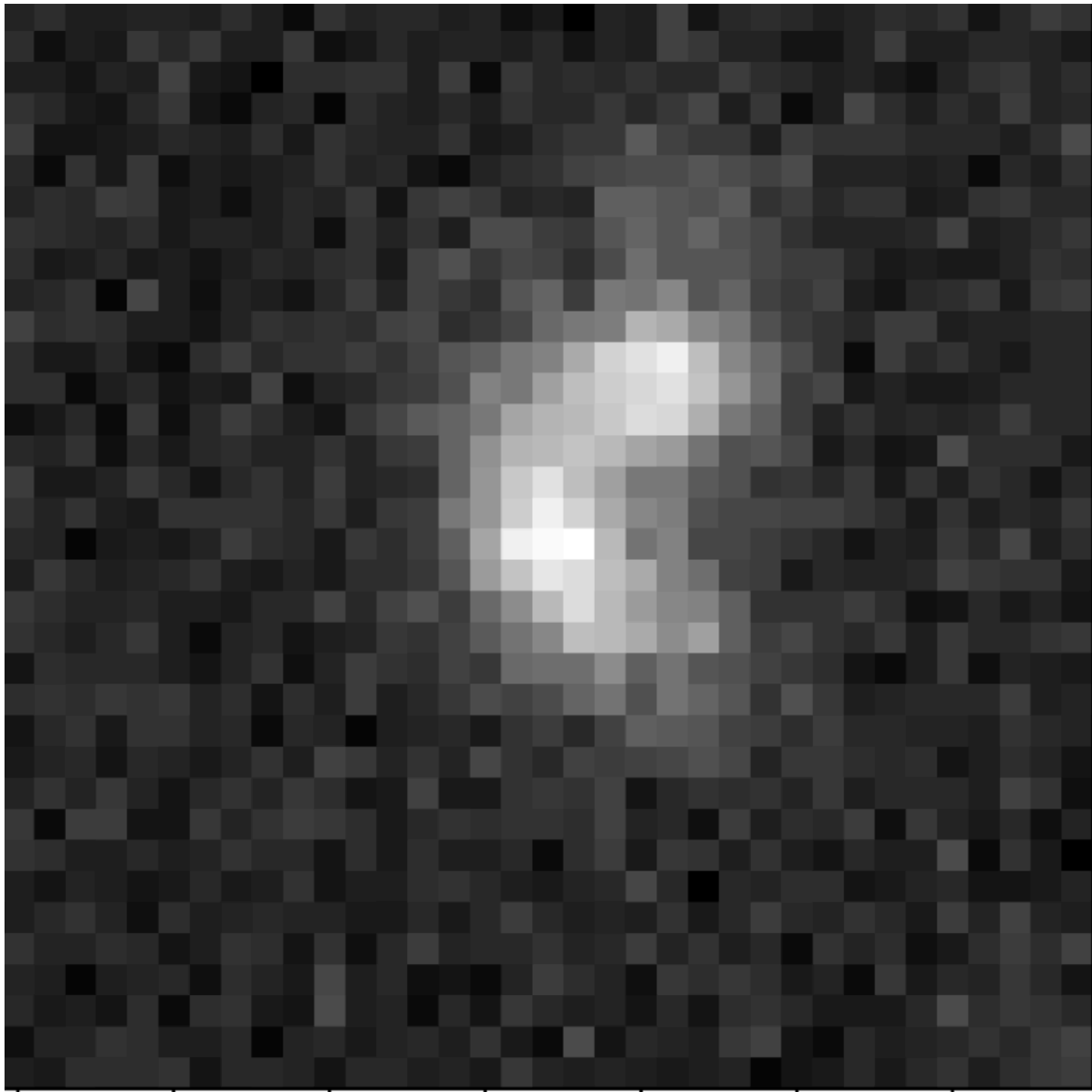


Remove other sources
using deblender models
Measure aperture flux on
“subtracted image”

Example of probabilistic detection output



Example of probabilistic detection output

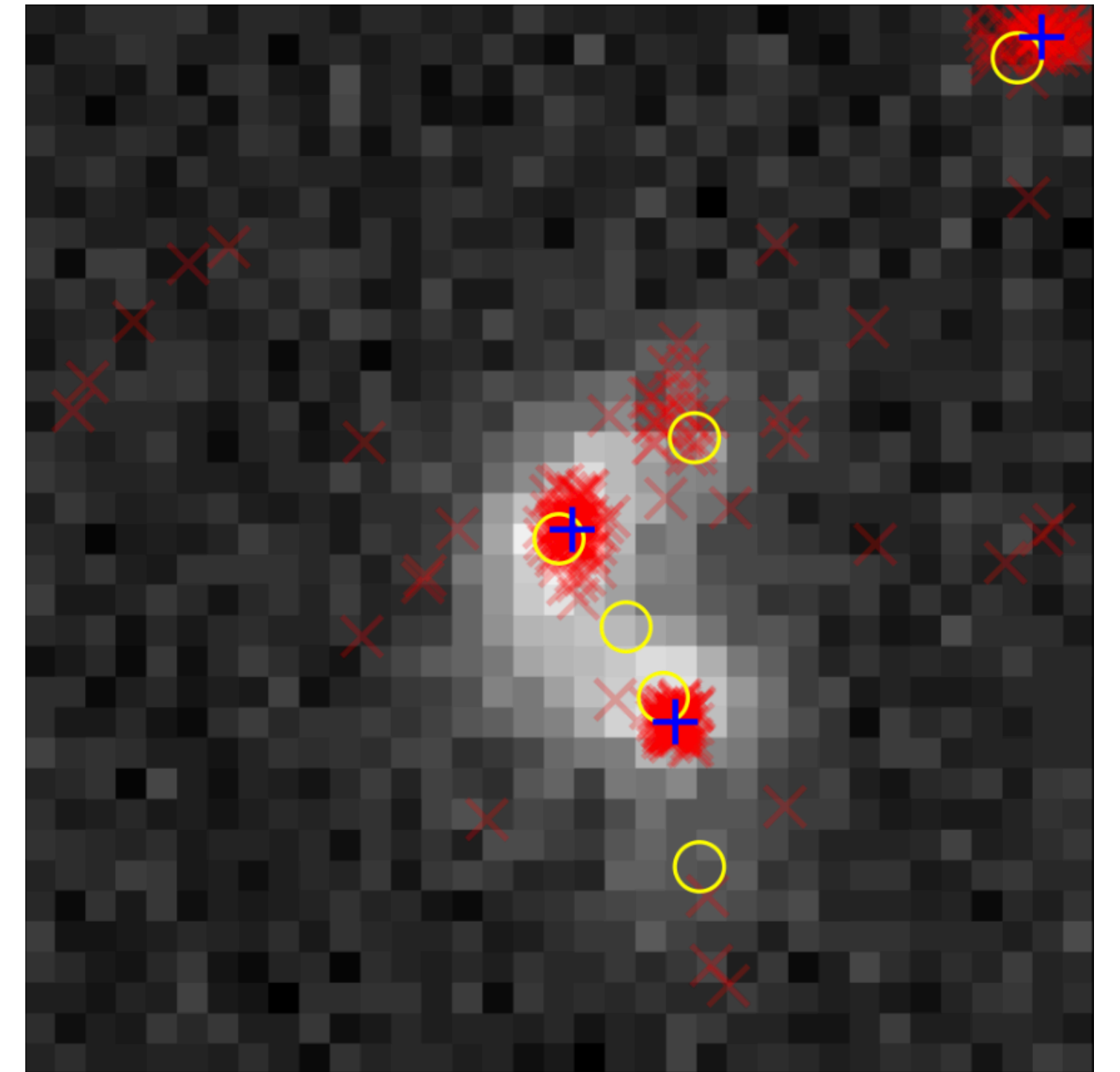
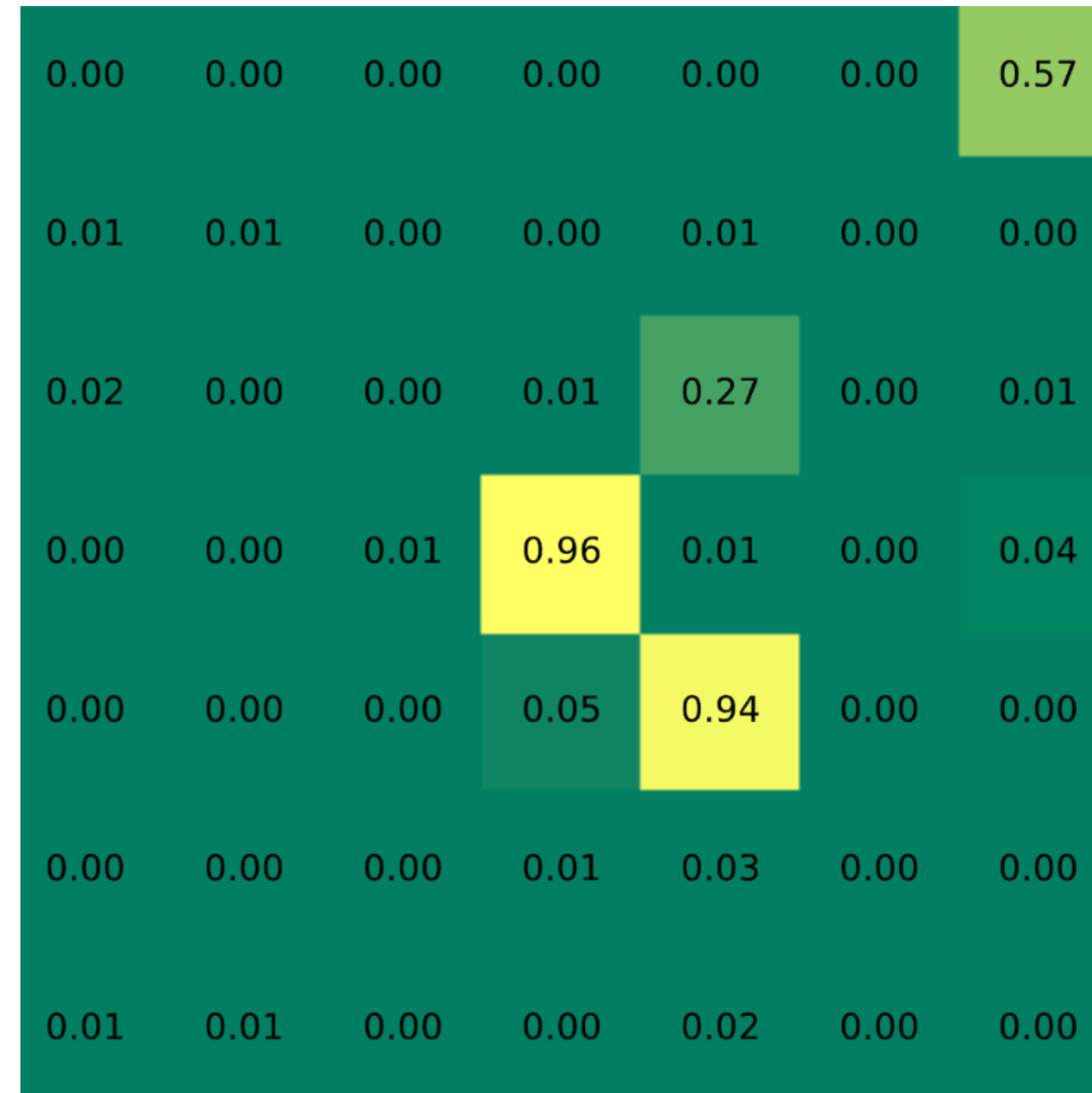
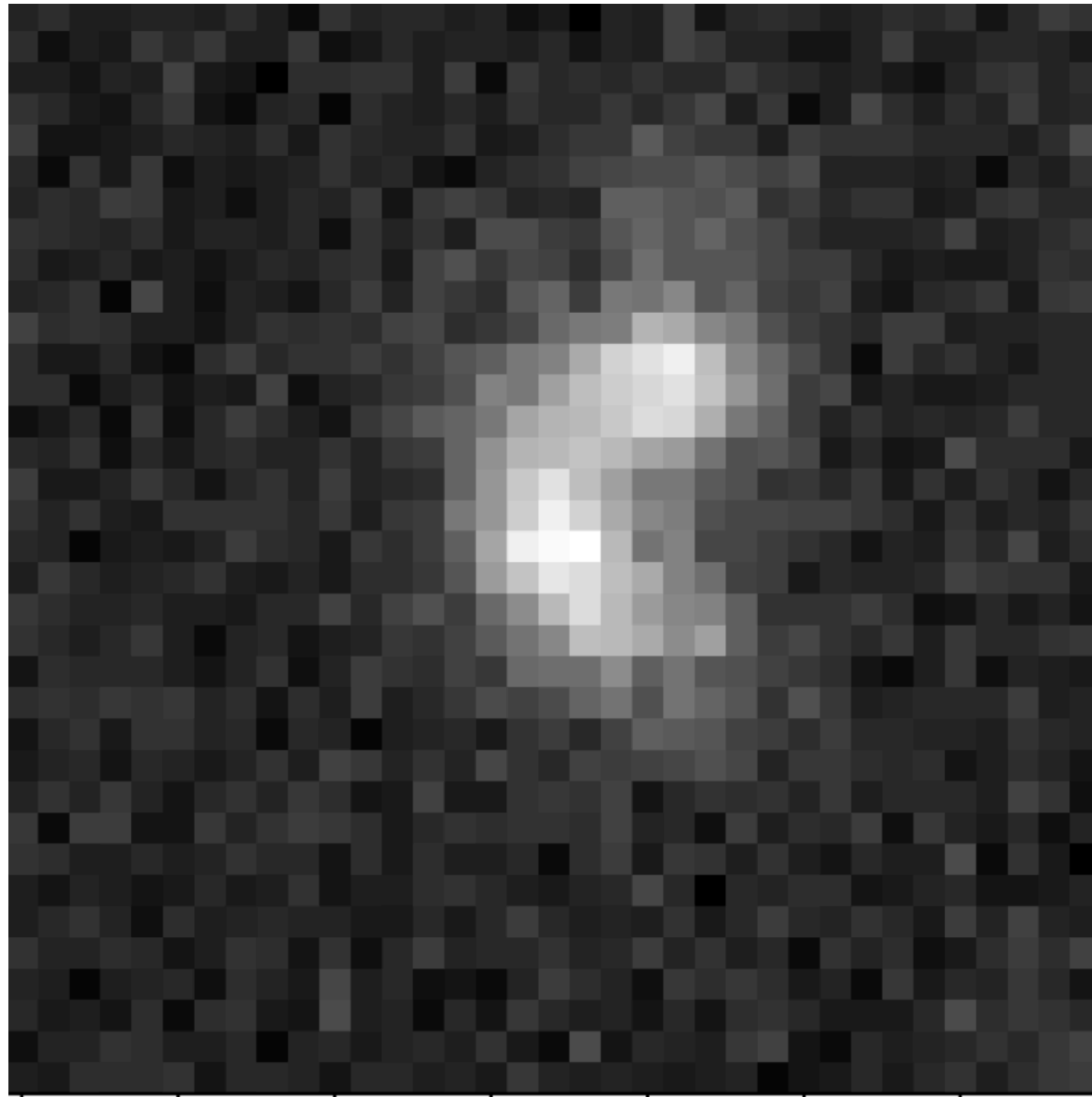


Example of probabilistic detection output

+ Threshold Detections $> 50\%$ (“MAP”)

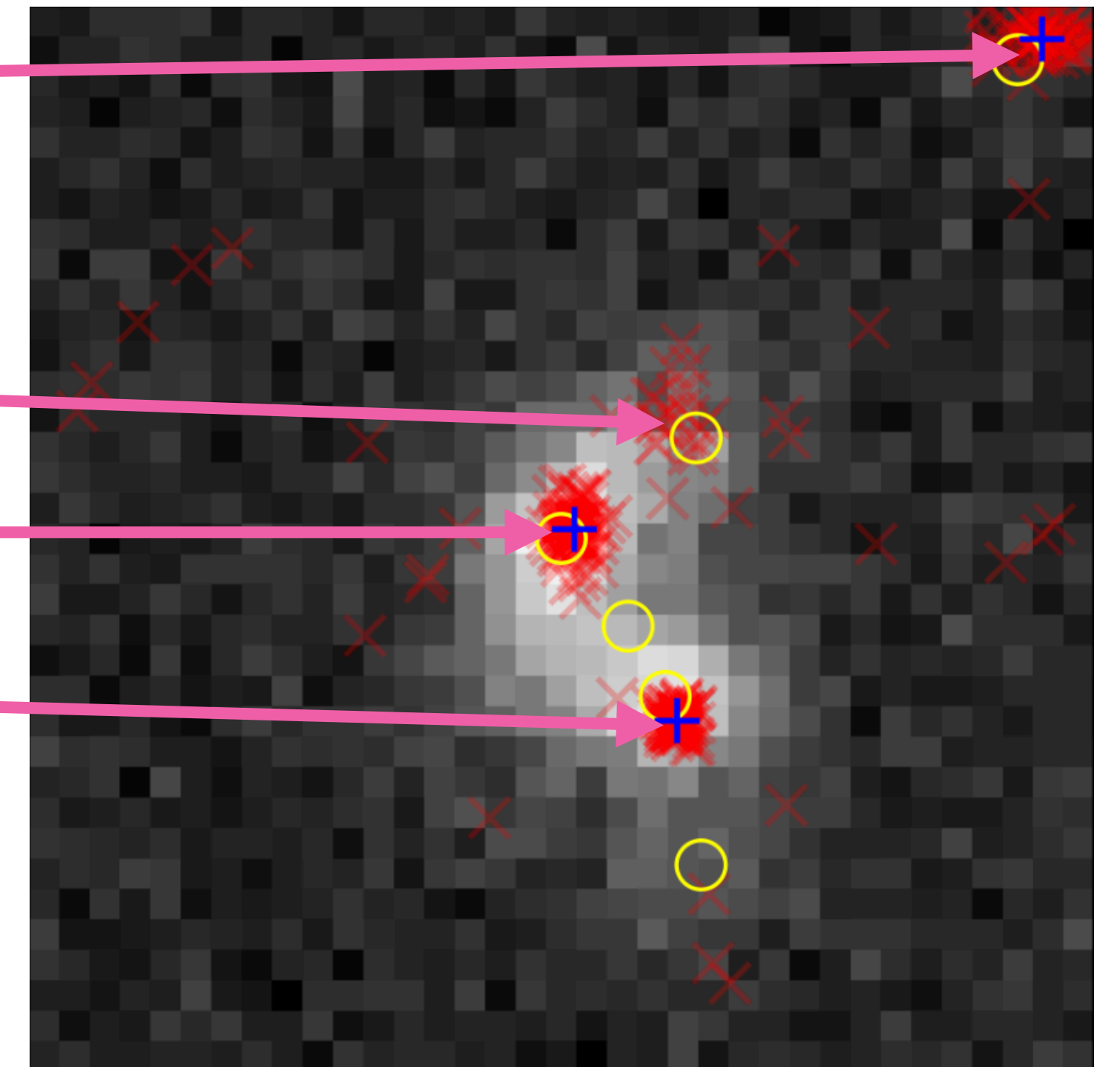
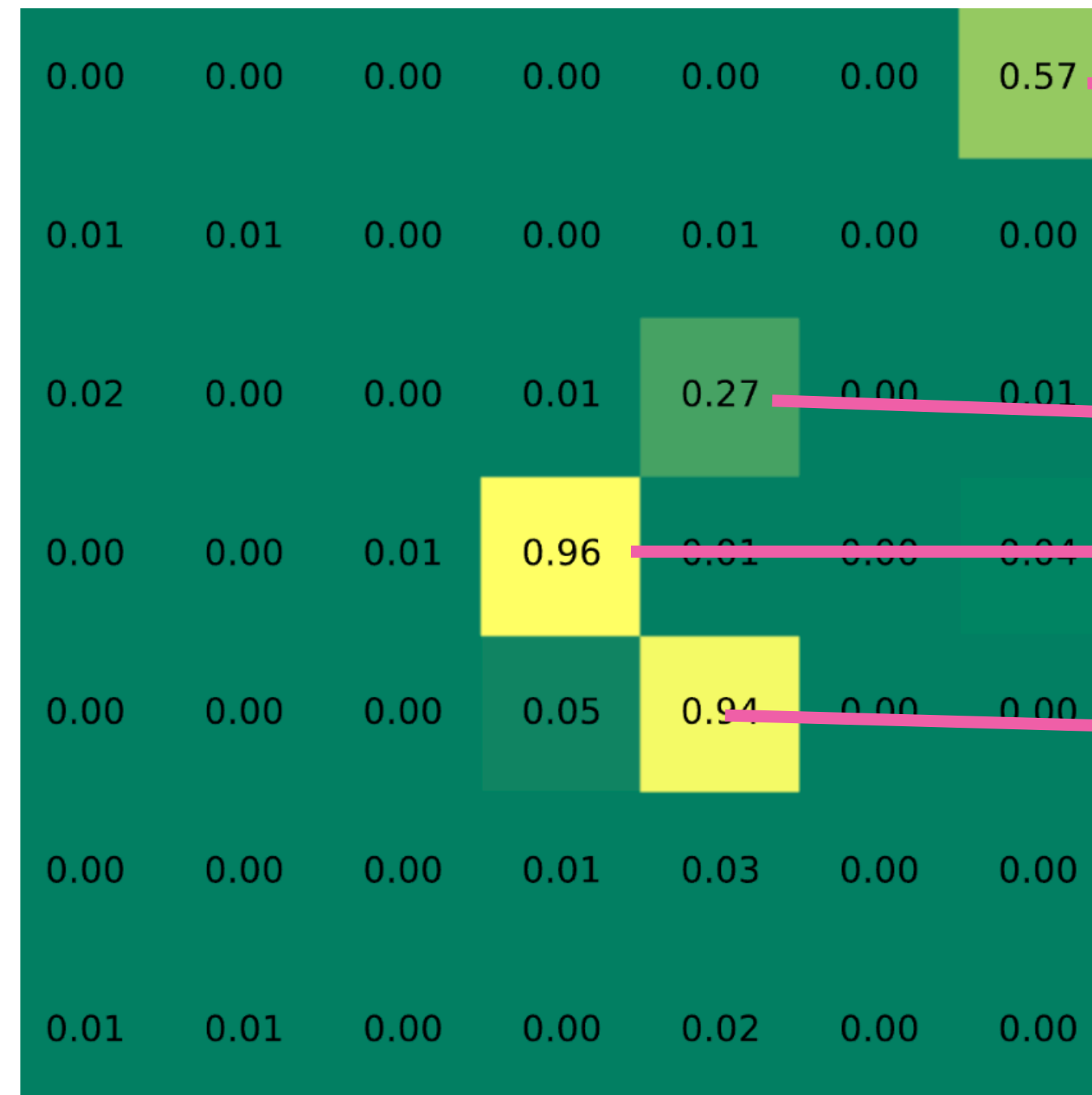
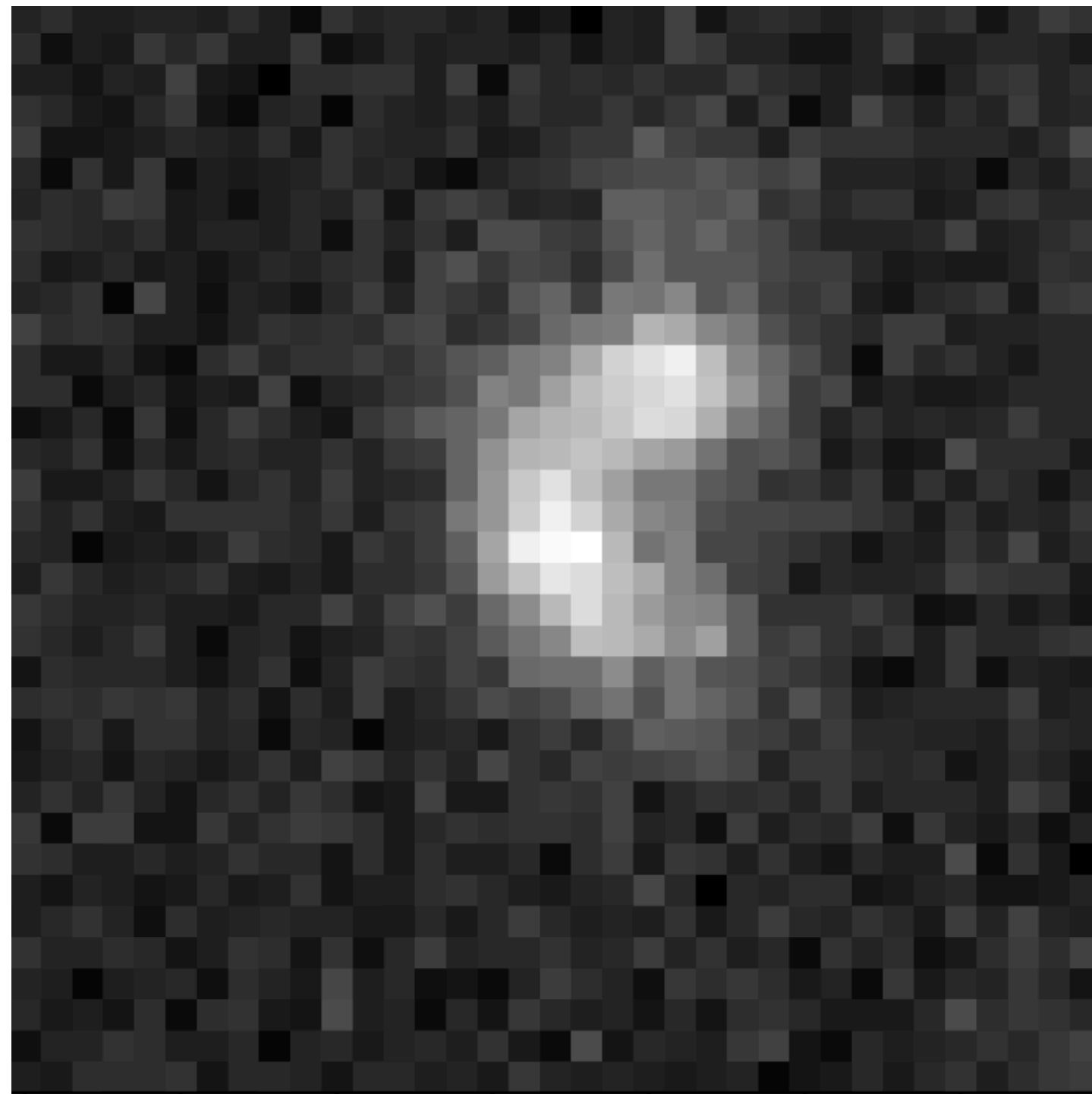
X Detection samples

O Truth



Example of probabilistic detection output

- + Threshold Detections $> 50\%$ (“MAP”)
- X Detection samples
- O Truth



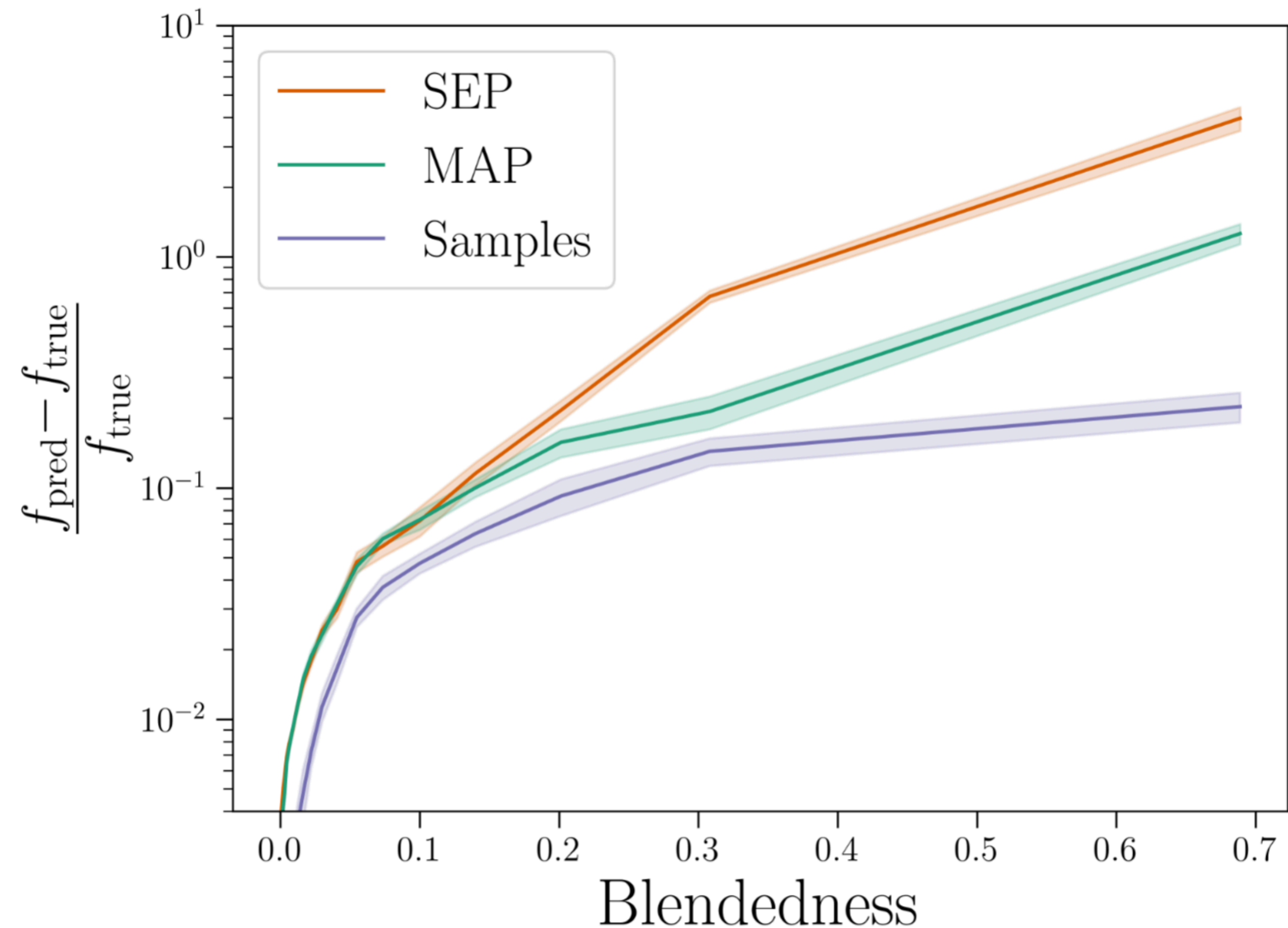
Probabilistic vs Deterministic aperture flux measurement

- Using samples for flux measurement reduces average residual by ~100%

SEP: SourceExtractor detections

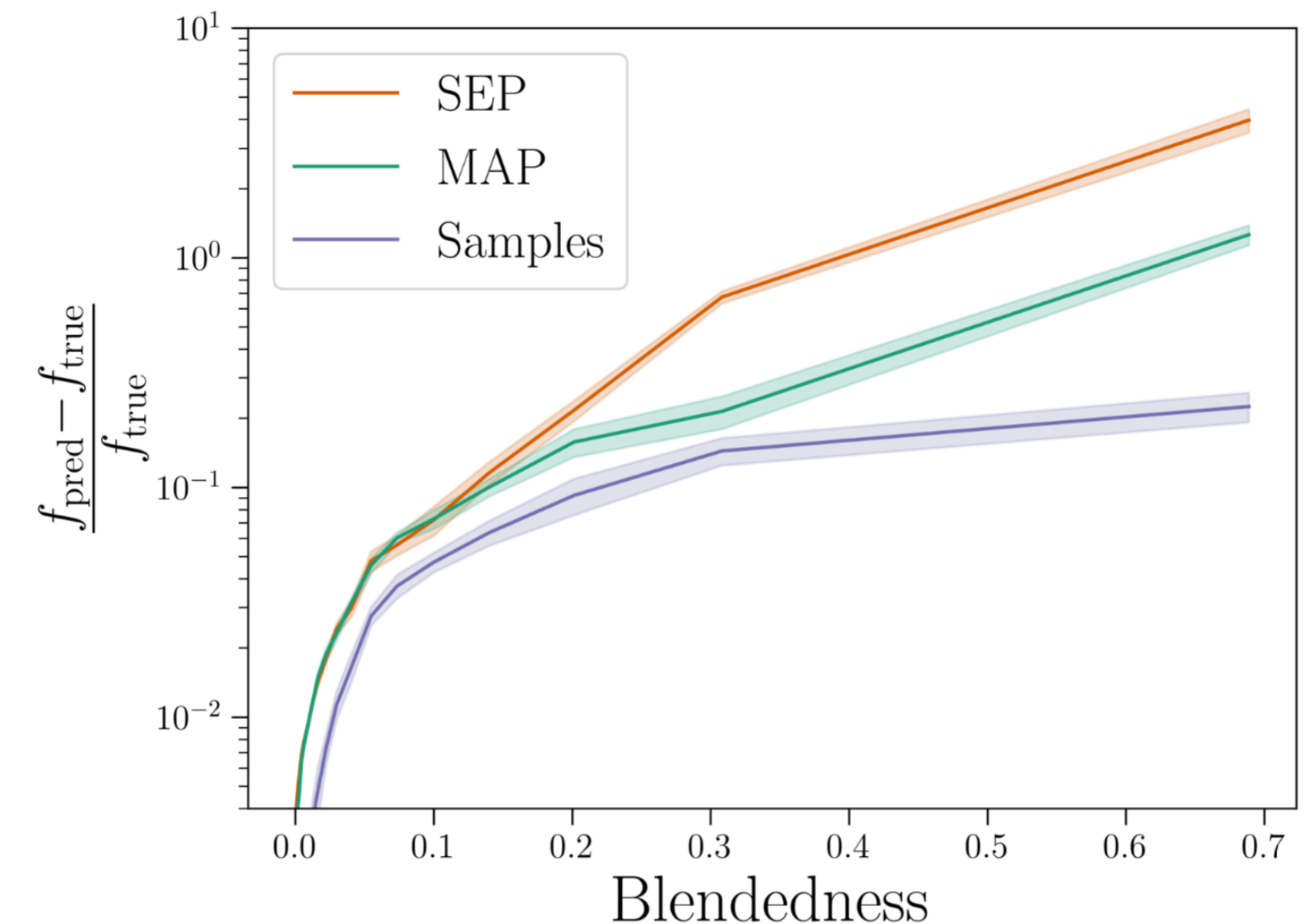
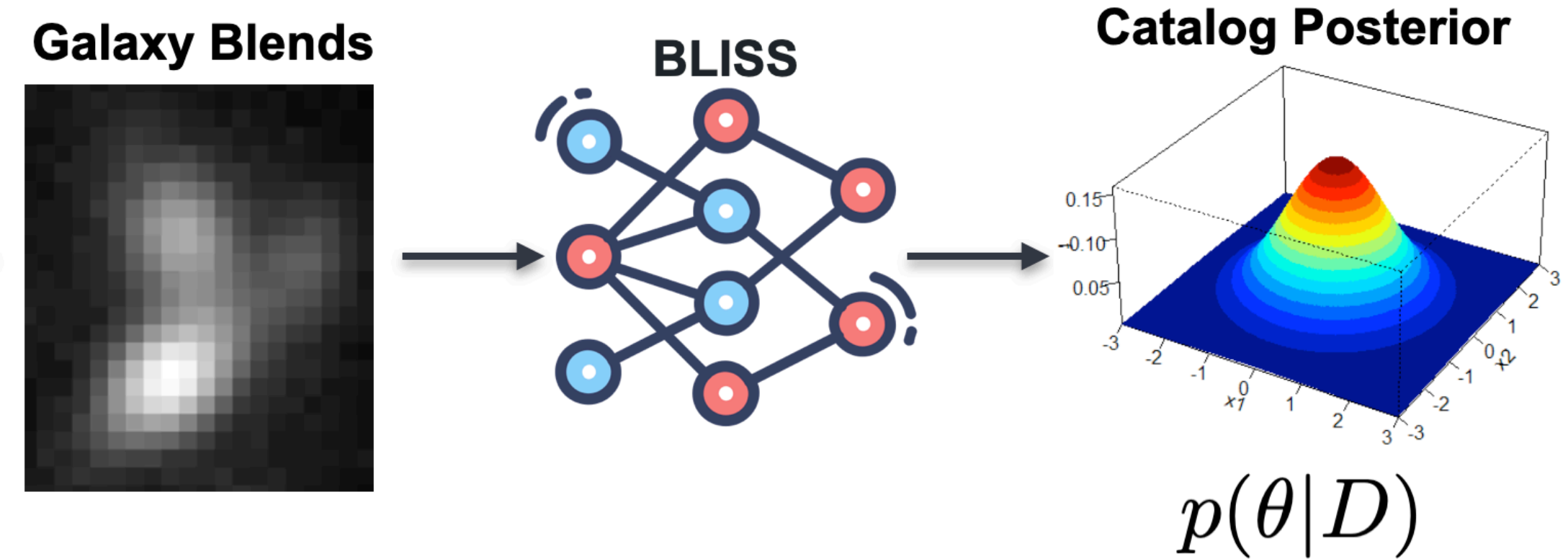
MAP: Most probable BLISS detections

Samples: Aggregate flux measurement across detection samples from encoder



Summary and outlook

- BLISS can **detect dimmer and more blended galaxies** than classical detection algorithms such as SourceExtractor.
- Probabilistic detection + deblending leads to **large reductions in measured flux bias**.
- **Future work:**
 - How robust is BLISS to OOD data?
 - Extend BLISS to multi-band images and evaluate improvement on photometric redshift estimation.



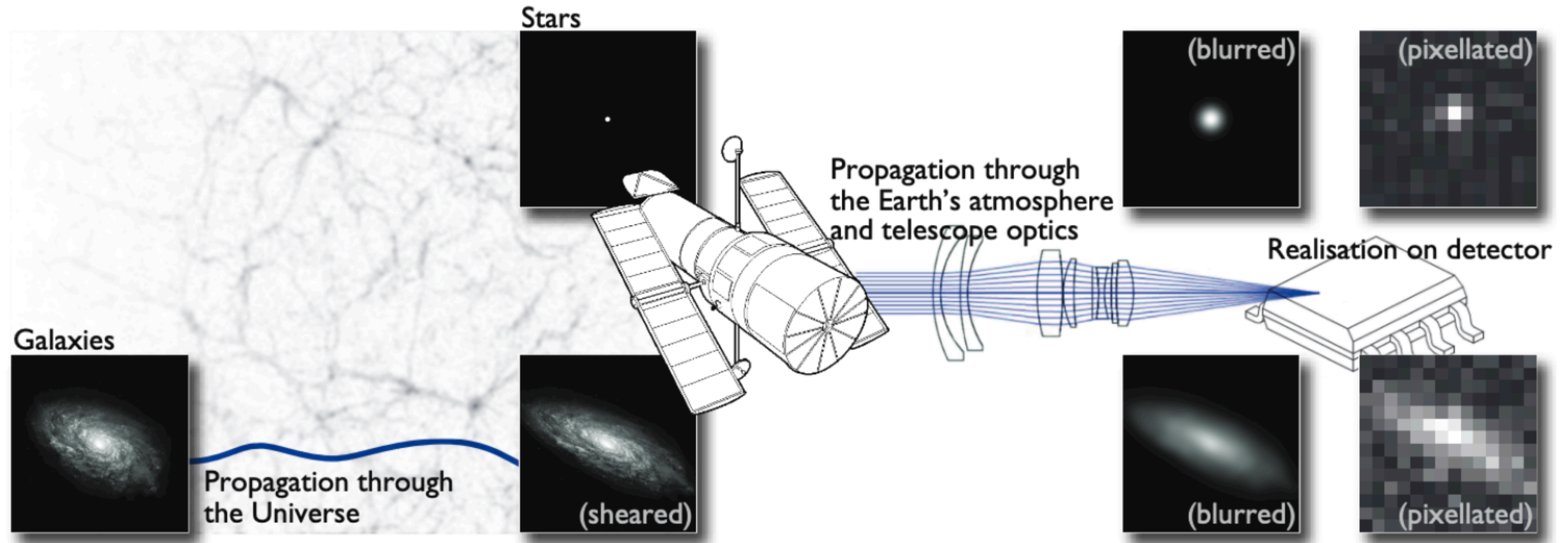
Extra Slides

How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!

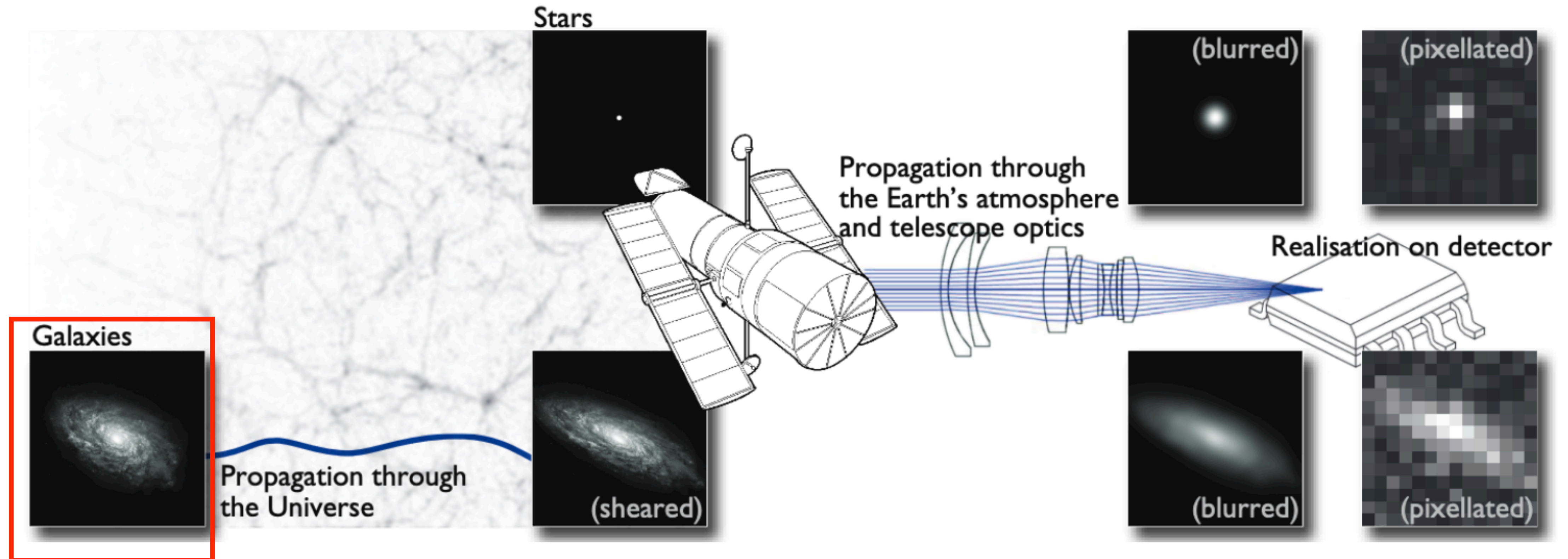
How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!



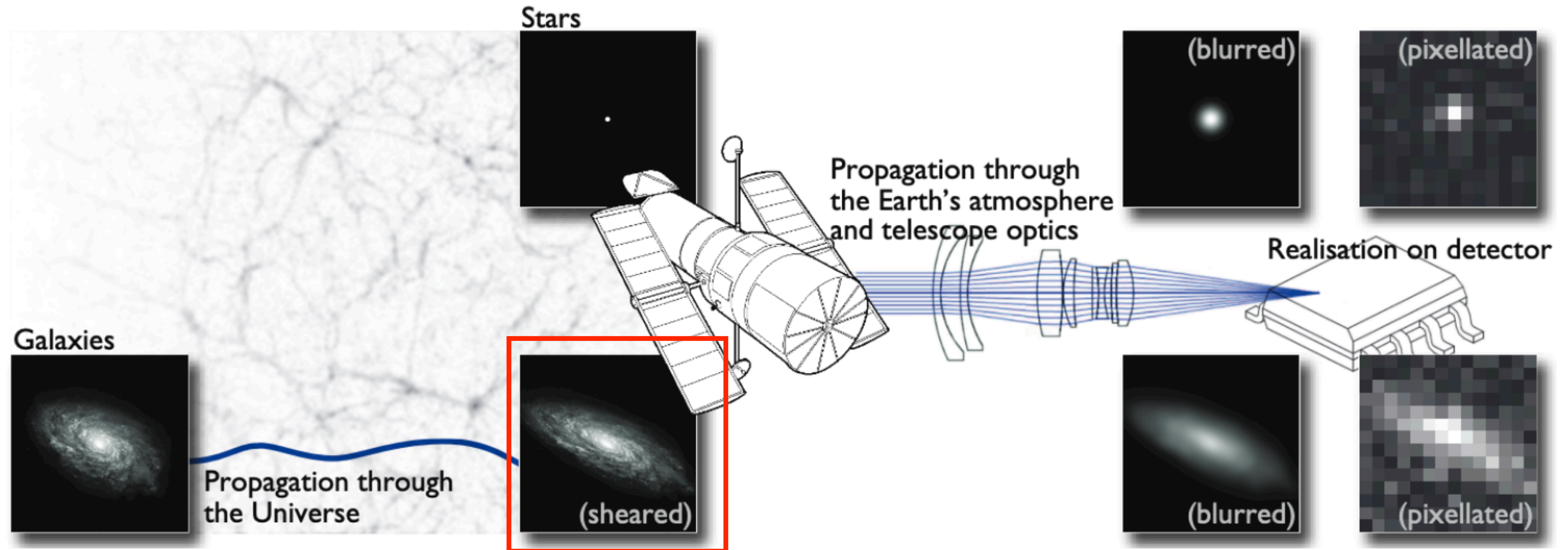
How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!



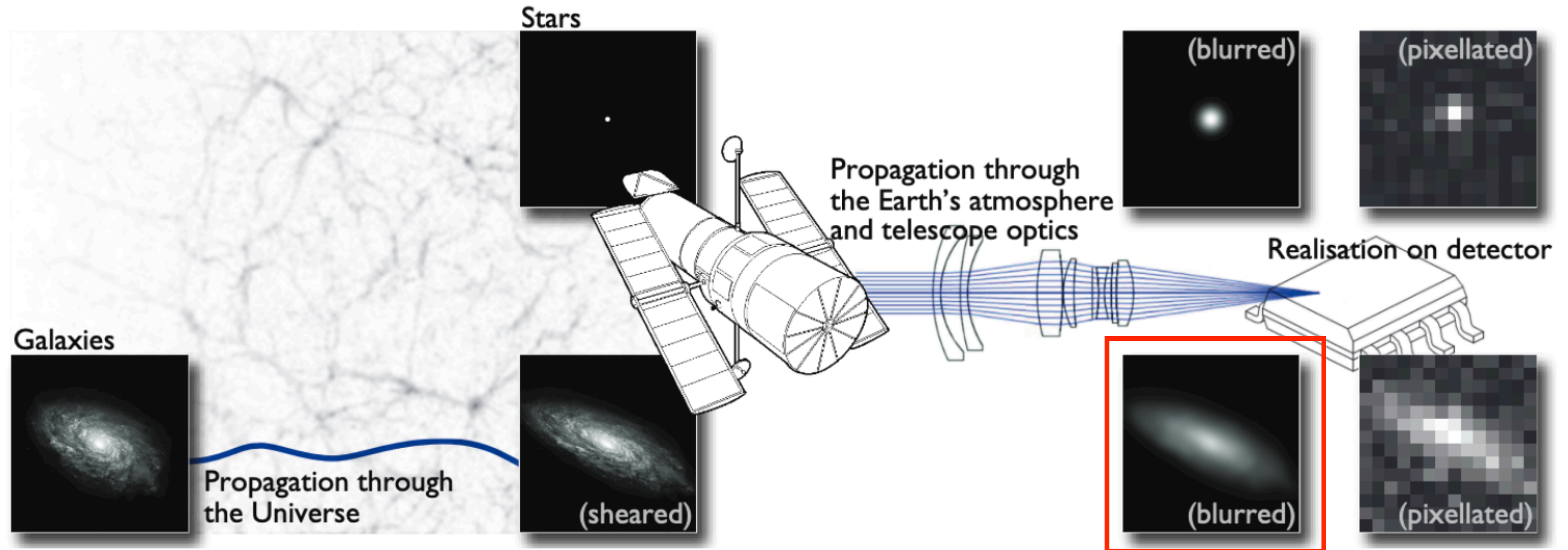
How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!



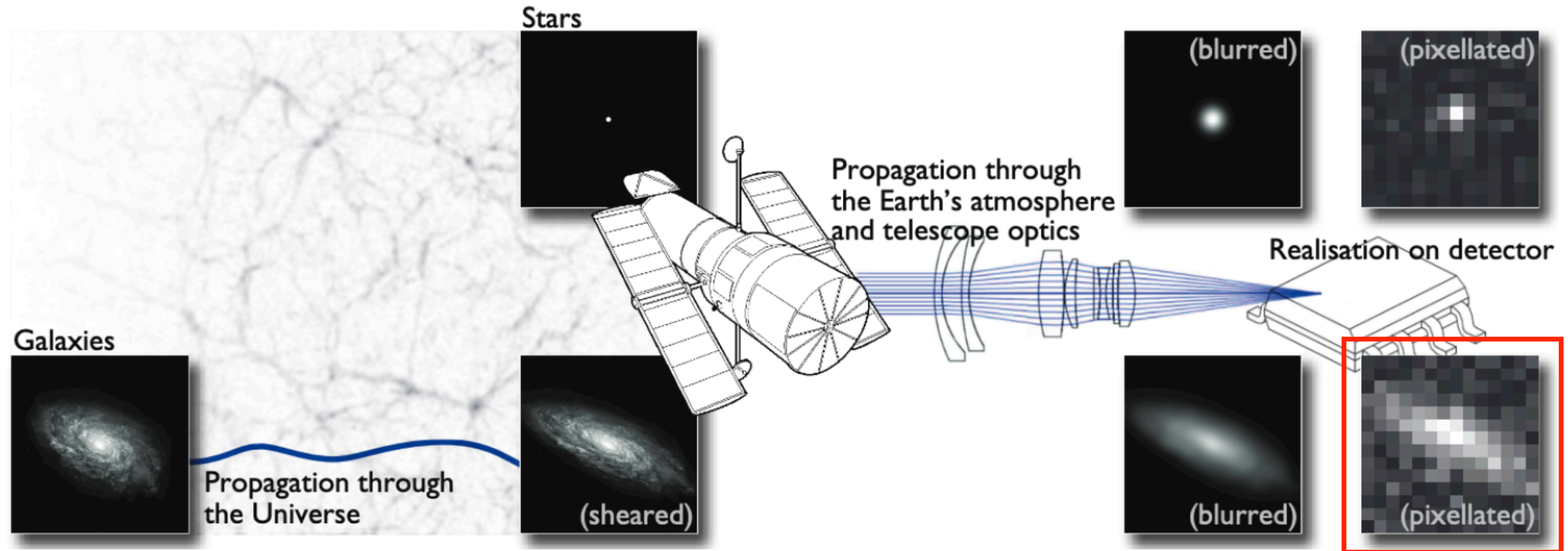
How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!



How difficult is it to measure cosmic shear?

- Shear is typically a **2% effect** change in the galaxy's axis ratio.
- LSST requirement is to measure this effect with errors **below 1 part in 1000**.
=> systematics need to be very well controlled!

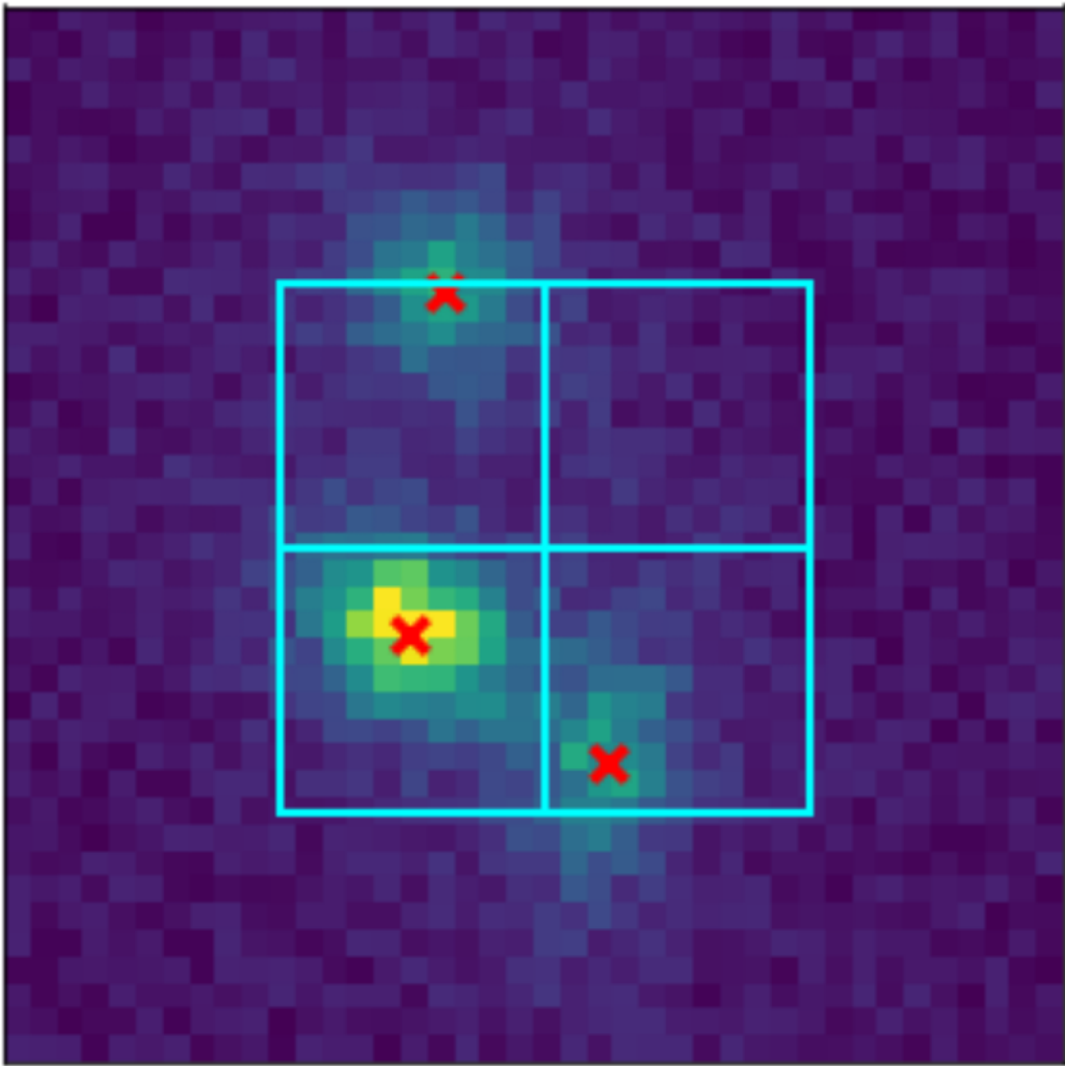


BLISS makes predictions per tile

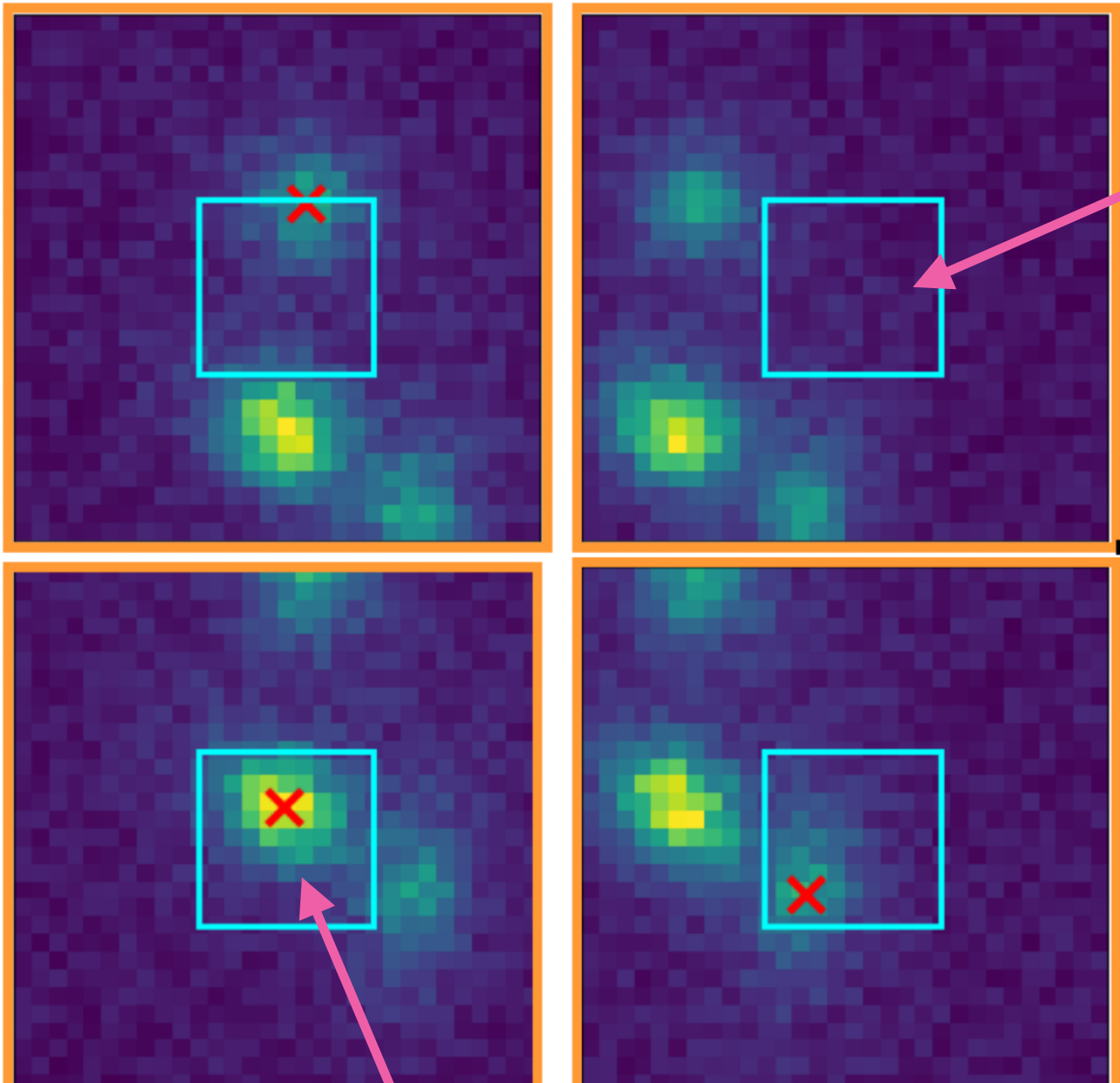
X True Centroids

□ Tiles

□ Padded Tiles

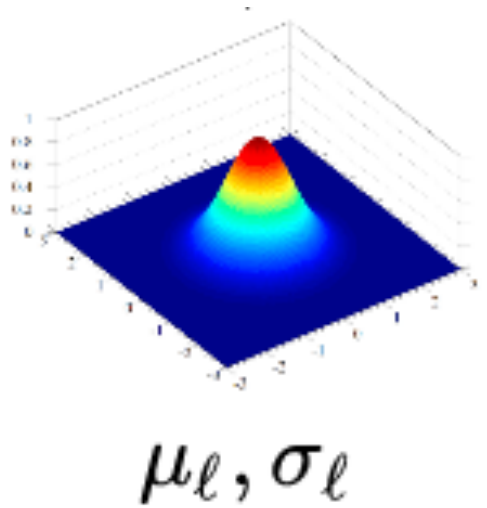


Split into
tiles +
border padding



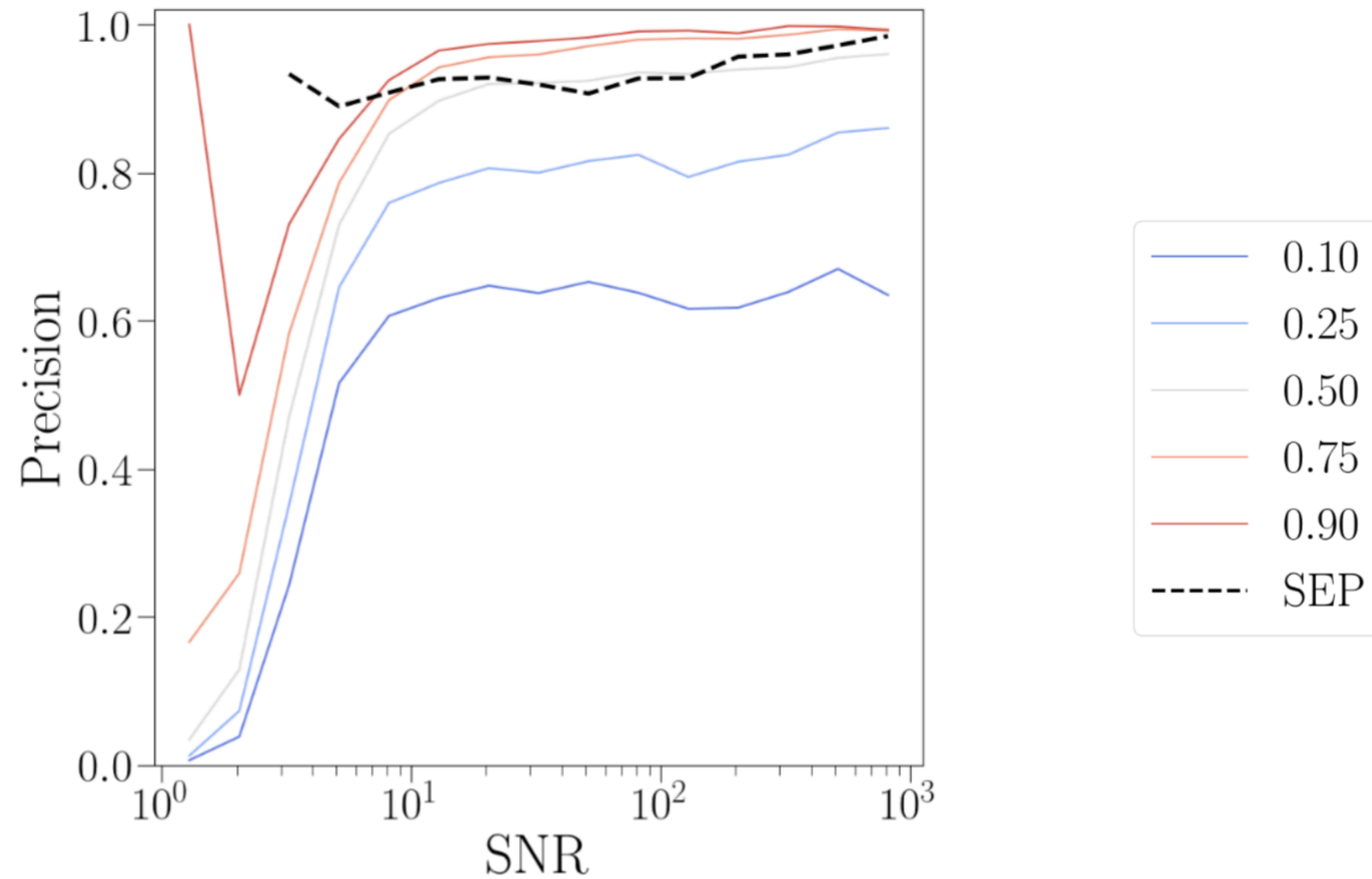
0 objects detected

1 object detected
=> output posterior
location/classification



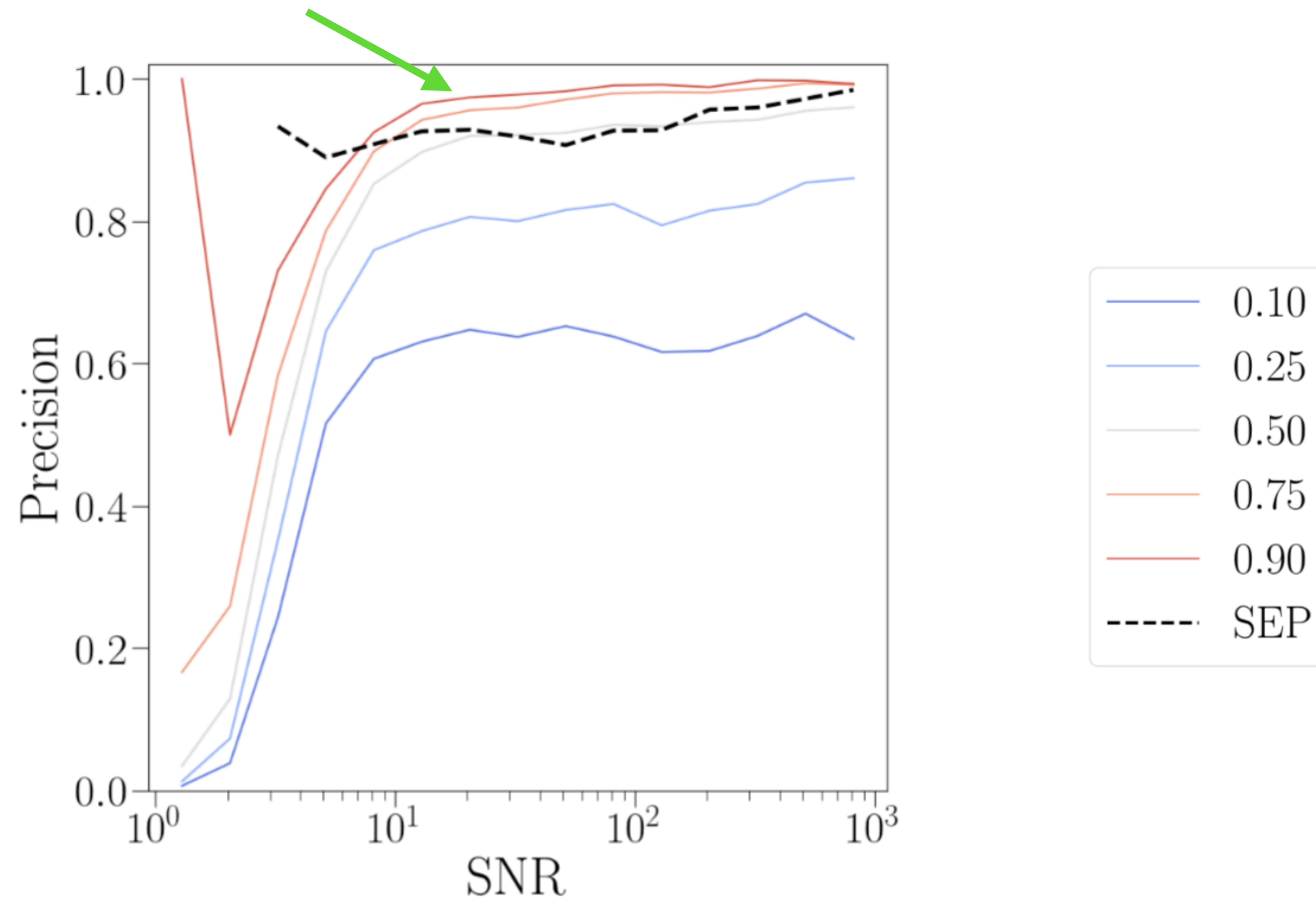
Detection “purity” (precision)

- BLISS can be customized to match the precision in SourceExtractor.



Detection “purity” (precision)

- BLISS can be customized to match the precision in SourceExtractor.



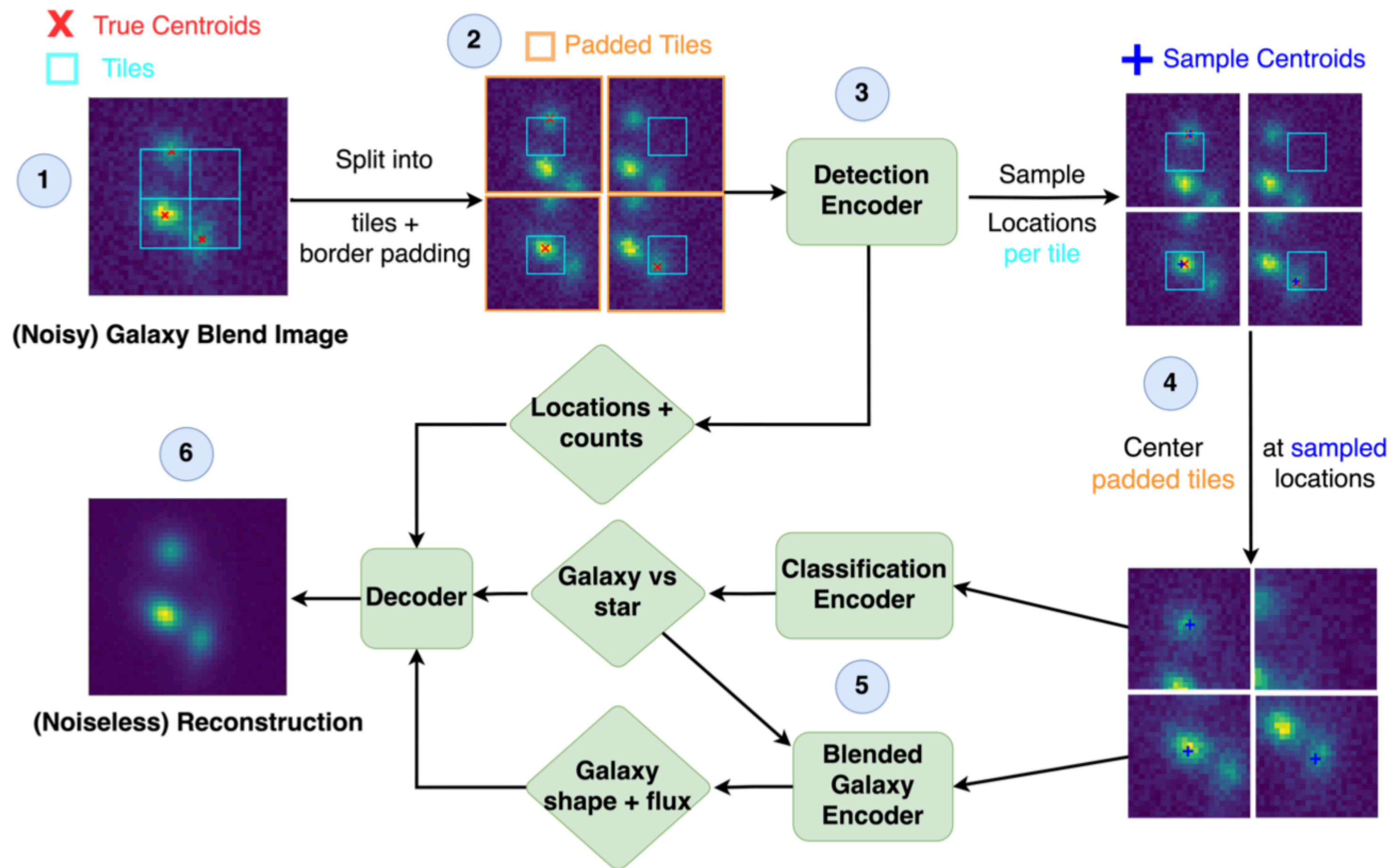


Figure 3.1: **BLISS inference pipeline outline**. Schematic demonstrating the procedure for detecting, deblending, and reconstructing galaxies with BLISS. The procedure starts with an astronomical image of a galaxy blend, which is split into padded tiles (orange border) that are passed through the detection encoder. The detection encoder outputs centroid locations for each detected galaxy that can be used to center each padded tile on its corresponding source. The center padded tiles are then passed to the classification and deblender encoder. Finally, the outputs of each encoder can be used to construct a full posterior catalog of the galaxy blend, and a corresponding noiseless reconstruction of the original image. For more details on each of the steps (blue circles) in this outline, see Section 3.4.5.

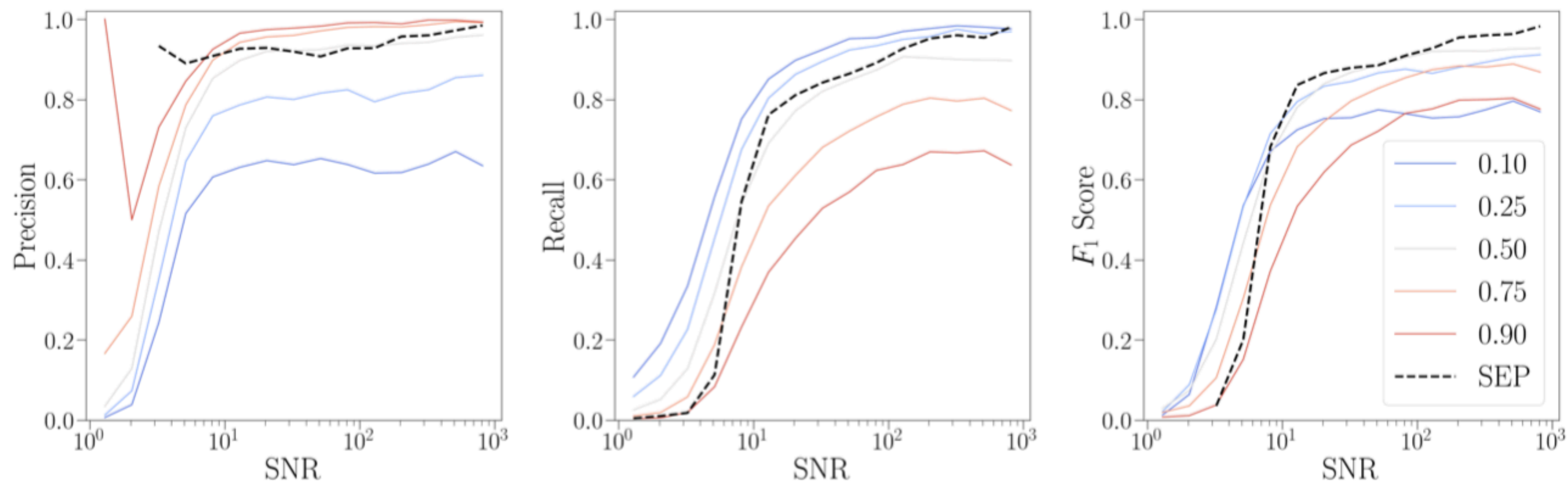


Figure 3.2: **Detection performance as a function of SNR.** In this figure we present the performance of the trained BLISS detection encoder using different probability thresholds (color gradient, blue to red) and SEP (dashed, black) on our testing set of galaxy blends. We show the precision, recall, and F_1 score in the same equally-log-spaced SNR bins. For more details on this figure see Section 3.5.2.

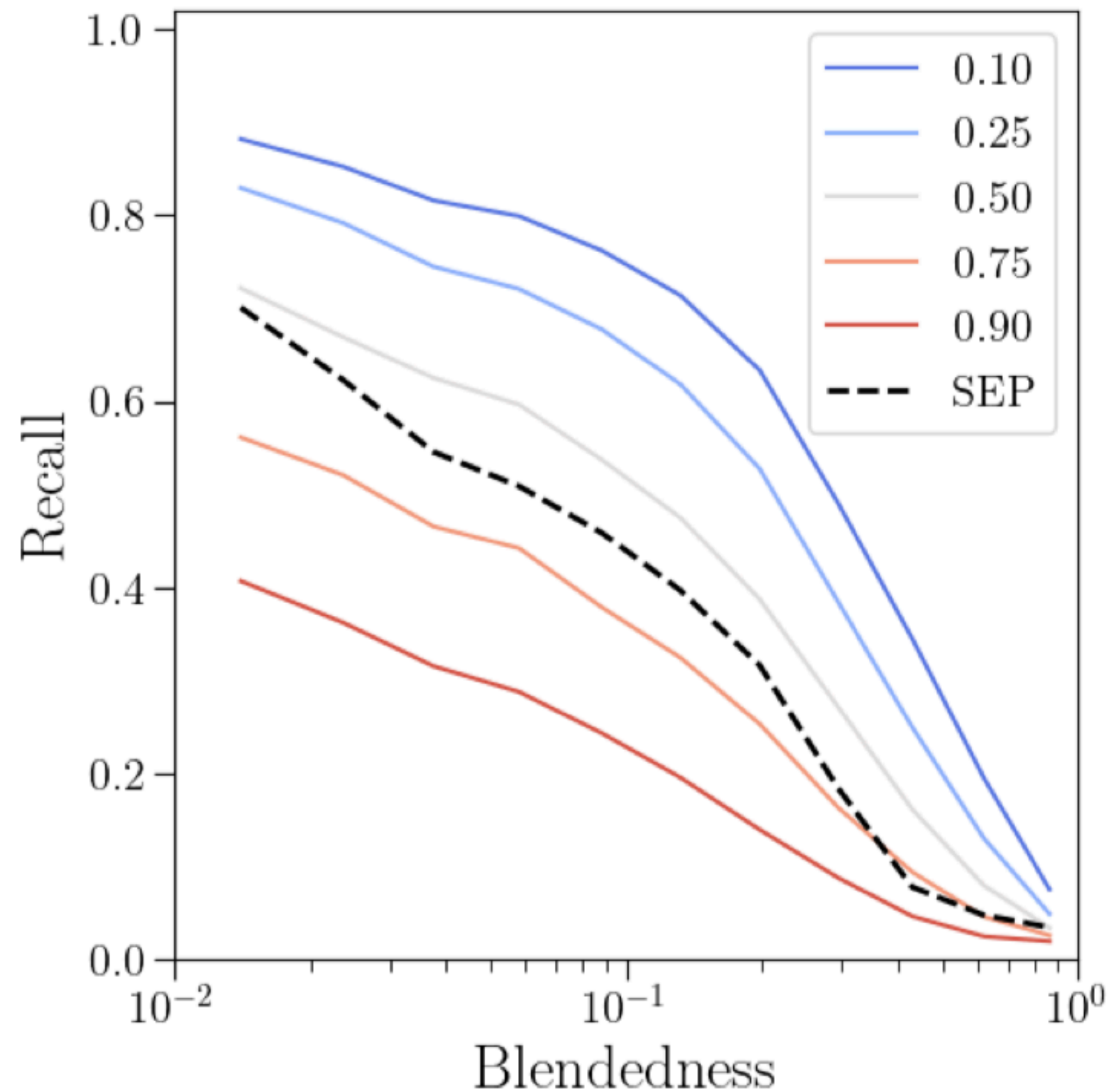


Figure 3.3: **Completeness as a function of blendedness.** This figure shows the fraction of detected true objects (i.e., recall) as a function of their blendedness for SEP and the BLISS detection encoder. The BLISS detection encoder’s recall for different probability thresholds is shown in the blue to red gradient curves, while SEP’s is shown with a black dashed curve. See Section 3.5.2 for details.

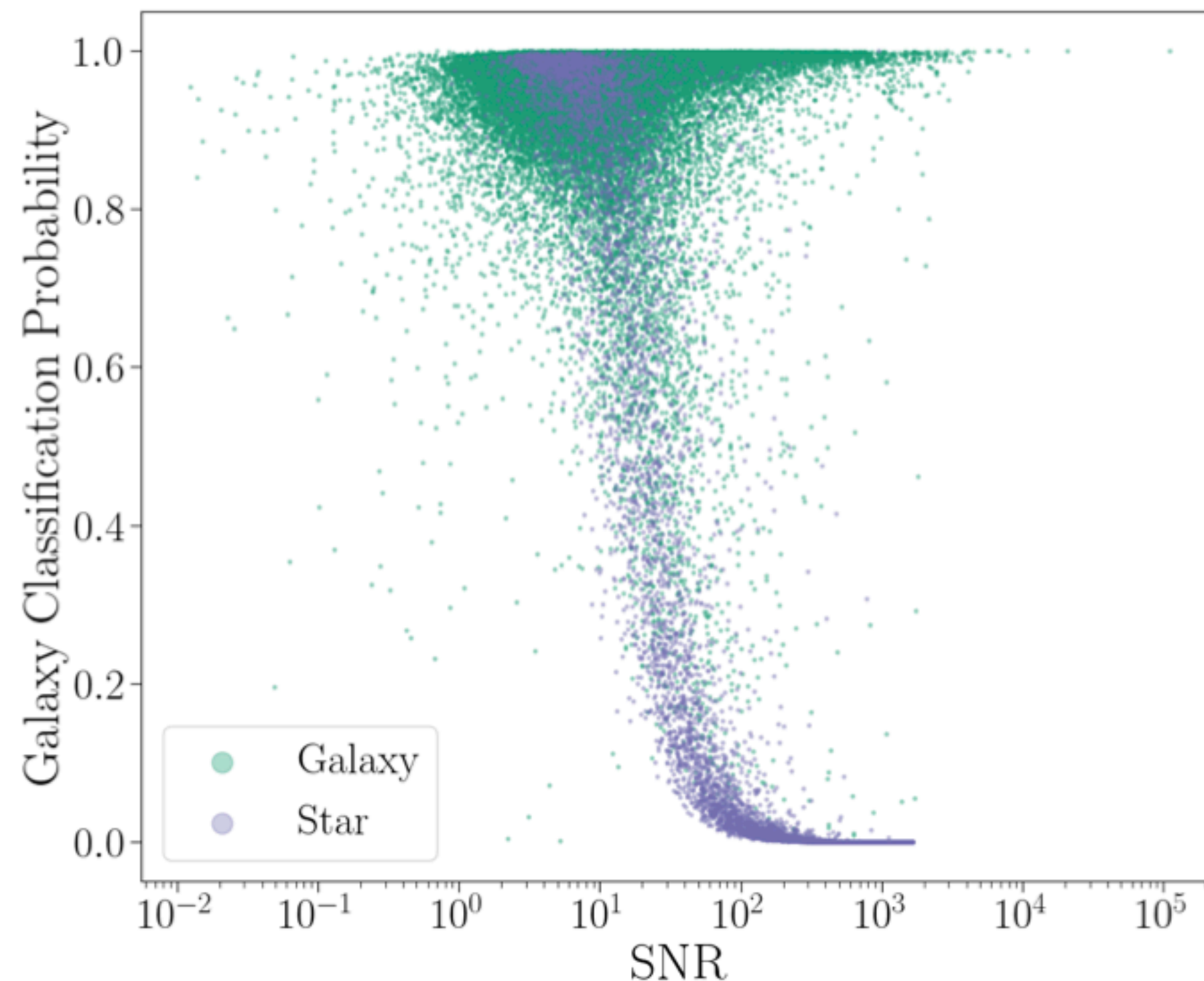


Figure 3.4: **Classification probability as a function of SNR.** In this figure we show the classification probability of a given source being a galaxy as a function of true SNR, for every source in the testing `dataset-blend`. This probability is the output from the binary encoder conditioned on true counts and centroids of every source, as in Figure 3.5. Each dot is color coded with its true galaxy (green) or star (purple) classification. For more details on this figure, see Section 3.5.3.

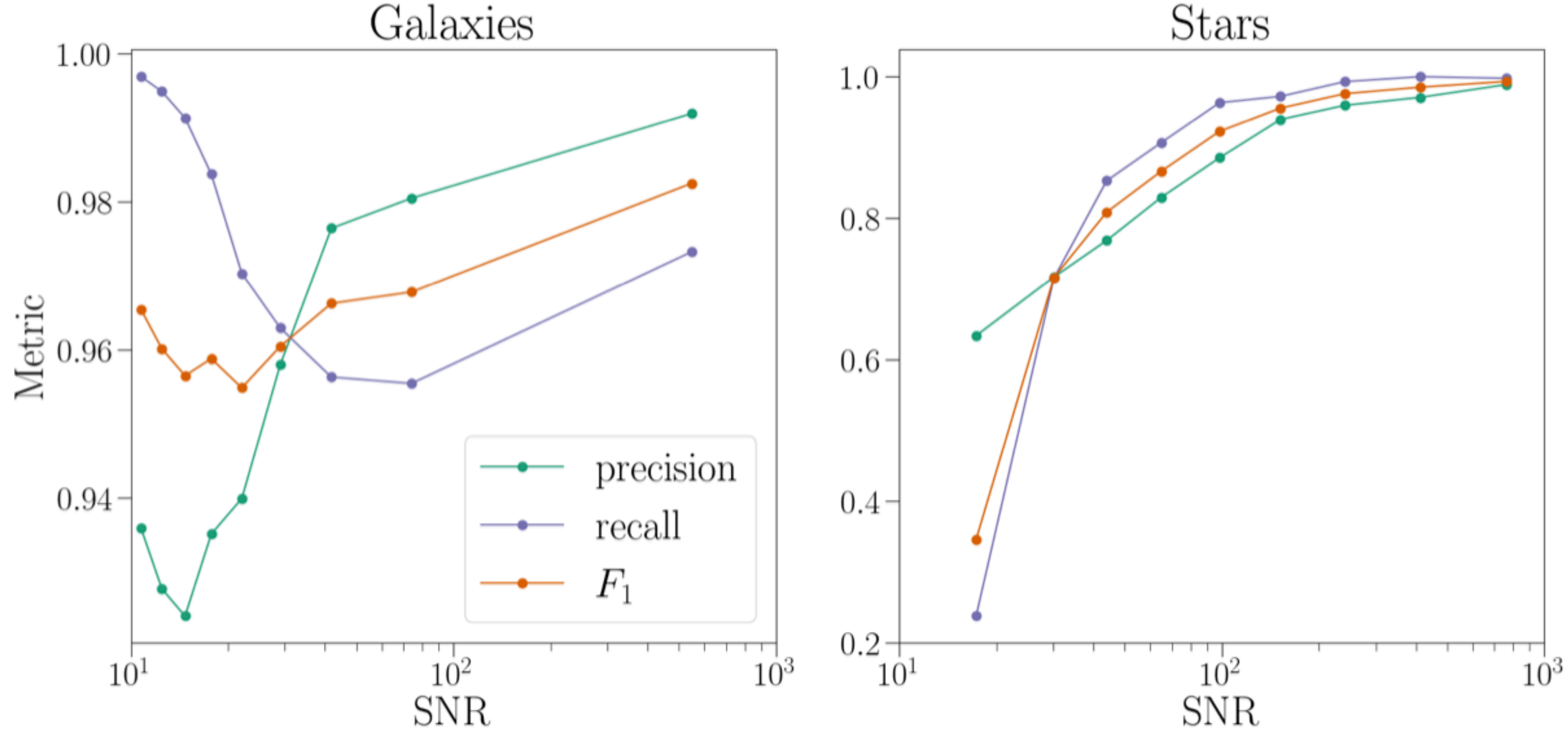


Figure 3.5: **Classification performance as a function of SNR for galaxies and stars.** This figure shows the precision (green), recall (purple), and F_1 score (orange) as a function of true SNR obtained by applying the binary encoder to our testing dataset-blend of galaxies and star blends. The SNR bins in each case are chosen so that there are an equal number of light sources on each bin. We condition the binary encoder on the true counts and centroids of every source in this dataset. For more details on this figure, see Section 3.5.3.

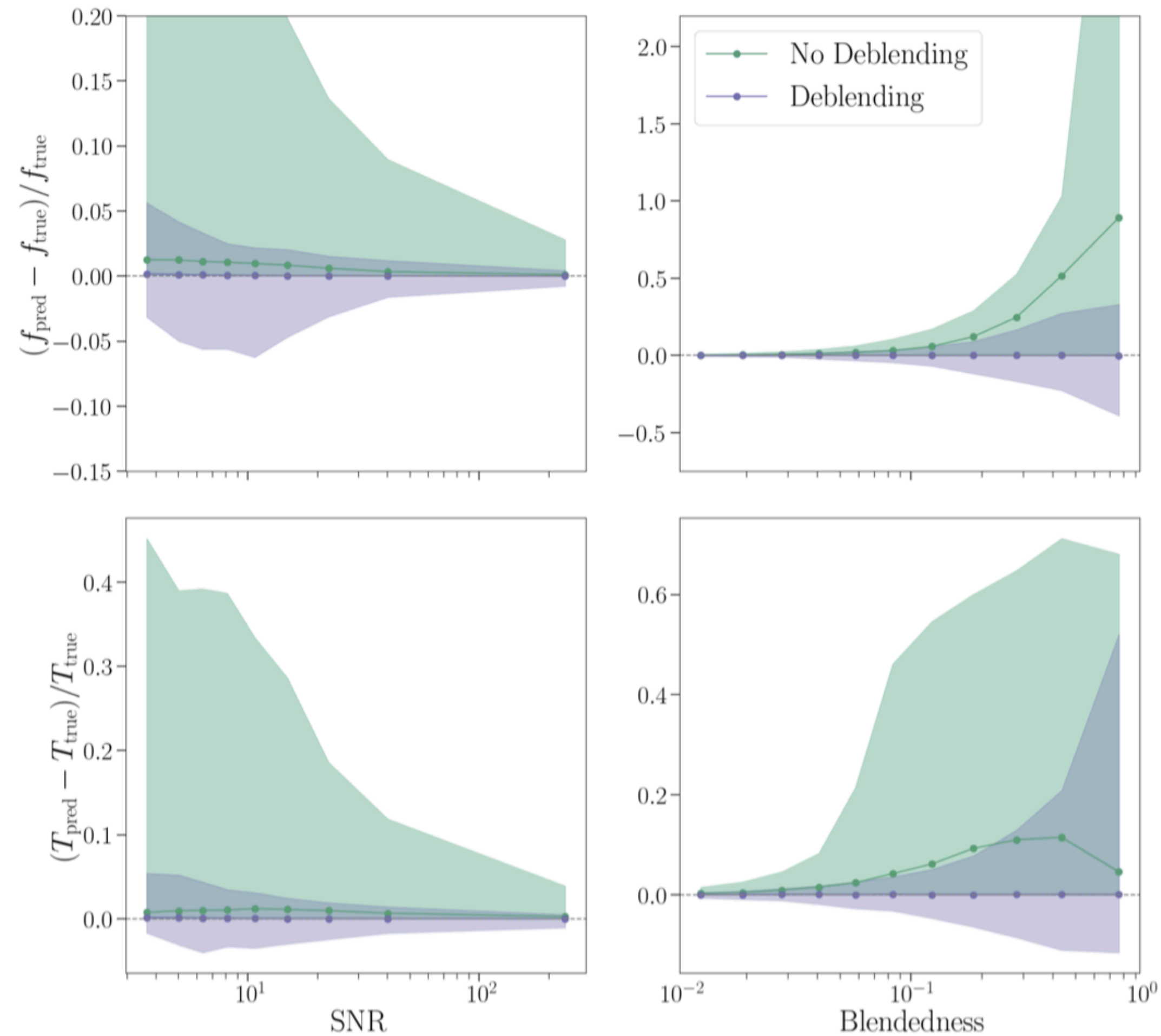


Figure 3.6: **Residual flux and size measurements as a function of SNR and blendedness.** In this plot we show the median residual flux and size measurements on galaxy blended images using true galaxy centroids. The flux (f) measurement is performed via aperture photometry with an aperture of size 5 pixels. The size ($T = |\det \mathbf{M}|^{1/4}$) The green color curves corresponds to the measurements with no deblending performed, but using the true centroid of each galaxy. The purple color curves use the reconstructed models from the deblender encoder to remove every other galaxy in the image before performing the measurement. The shaded regions correspond to quantiles of the 1σ deviation from a Gaussian distribution (i.e., 0.159 and 0.841 respectively). For more details on this figure, see Section 3.5.4.

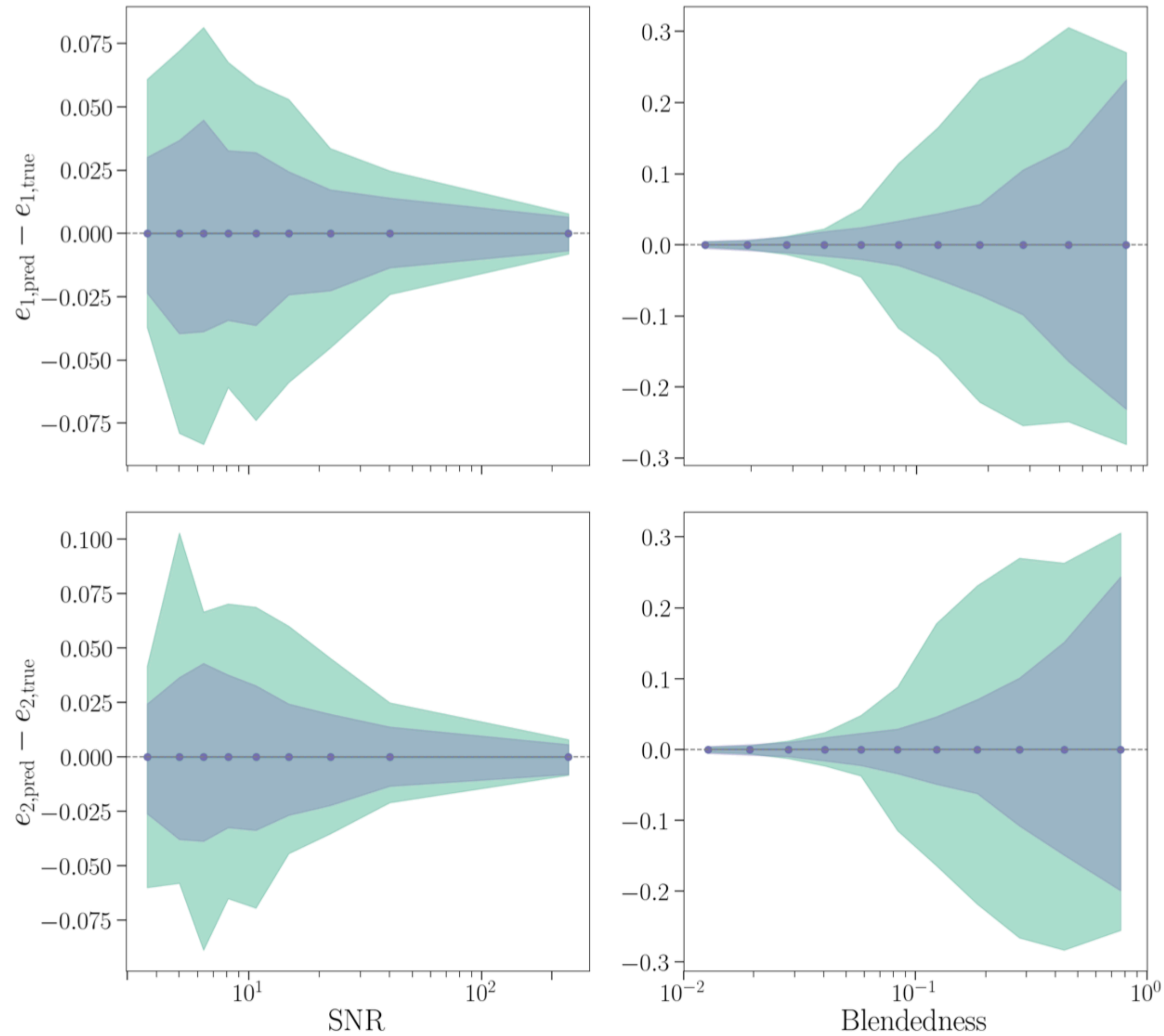


Figure 3.7: **Residual ellipticity measurements as a function of SNR and blendedness.** This figure is the same as Figure 3.6 except that the residual ellipticities are reported. The ellipticities (e_1 , e_2) are measured using the GALSIM adaptive moments routine. For more details on this figure, see Section 3.5.4.

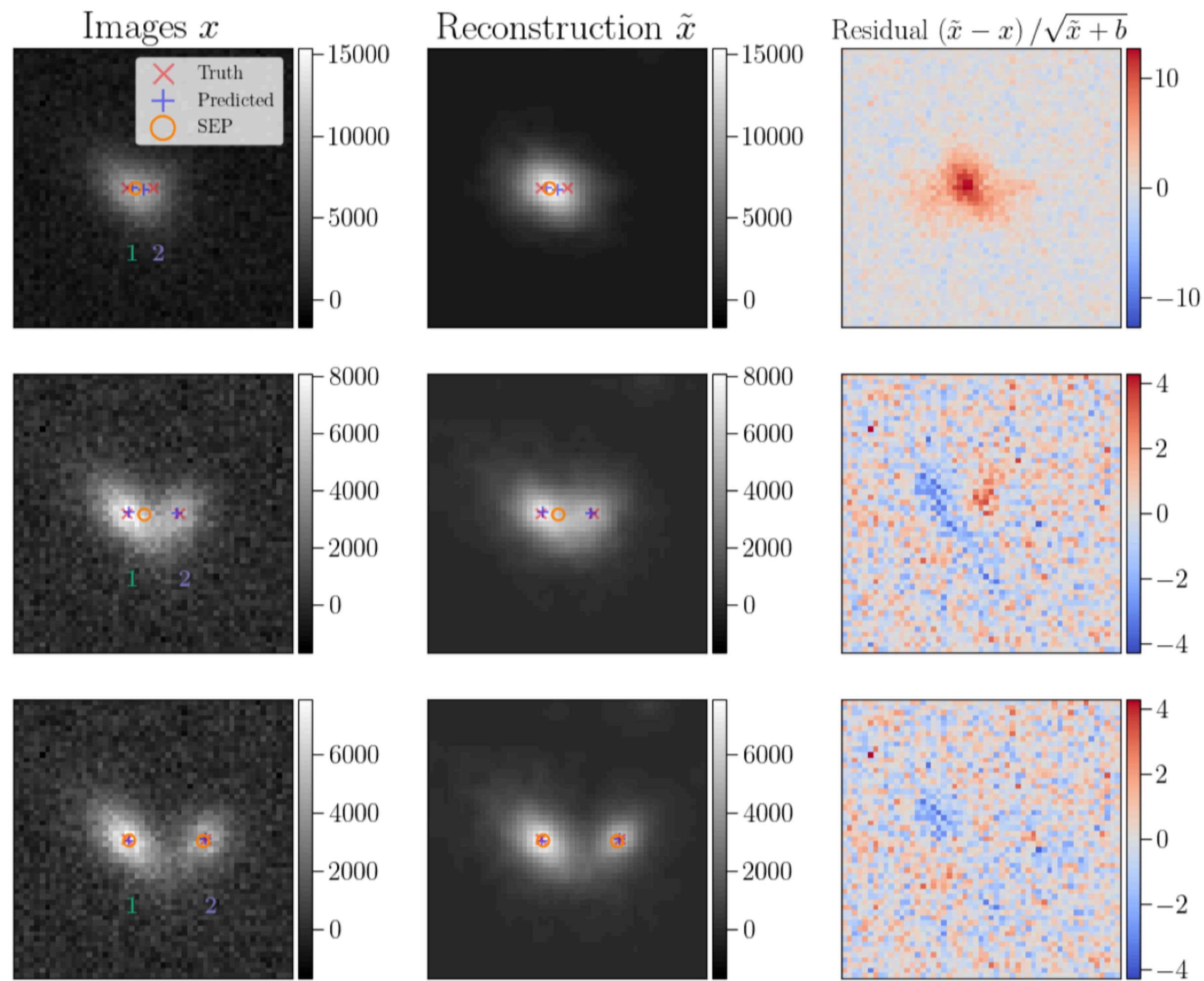


Figure 3.8: **Image residuals and detections on galaxy pairs at three separations.**

In this plot we show the pair of example galaxies used throughout Section 3.5.5 at three different separations: 5, 10, and 15 pixels. The galaxy on the left (labeled with “1”, in green) is fixed at the center of the image. The galaxy on the right (labeled with “2”, in purple) is shifted horizontally at different separations from Galaxy 1. In the first column, we see the true noisy image at each separation with the true centroid of each galaxy in red and the predicted centroid from BLISS in blue. The SEP prediction for centroids is also shown with an orange circle as a reference. The second column shows the noiseless reconstruction from BLISS with the same centroids plotted as in the first column. The last column shows the residual image.

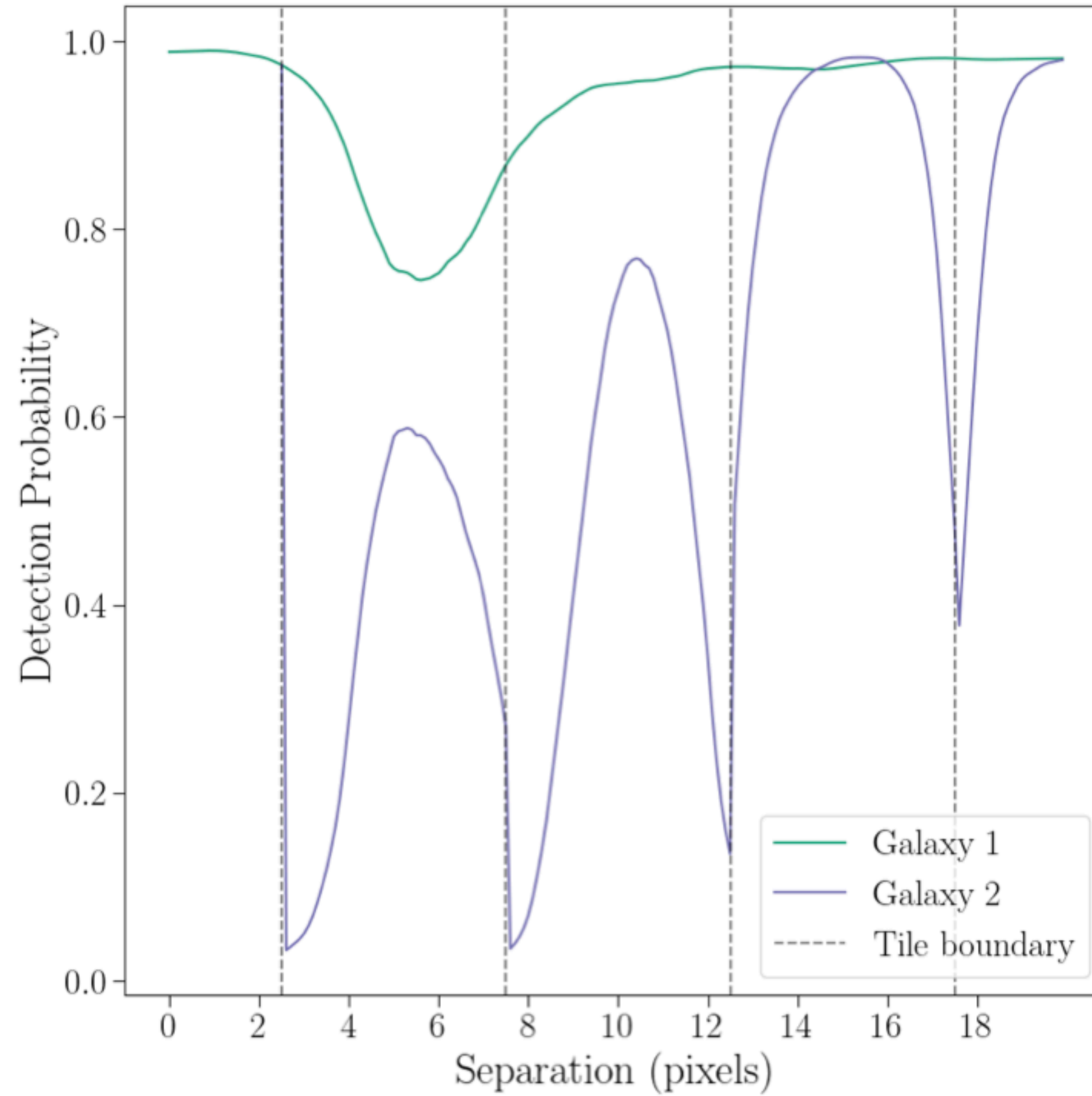


Figure 3.9: **Detection probability of each galaxy as a function of separation.** This plot shows the detection probability that BLISS assigns to the tile containing Galaxy 1 (green) and 2 (purple) as a function of the separation between them (as described in Figure 3.8). The dashed vertical lines correspond to the separations at which Galaxy 2’s centroid lands in a tile boundary. See Section 3.5.5 for more details on this figure.

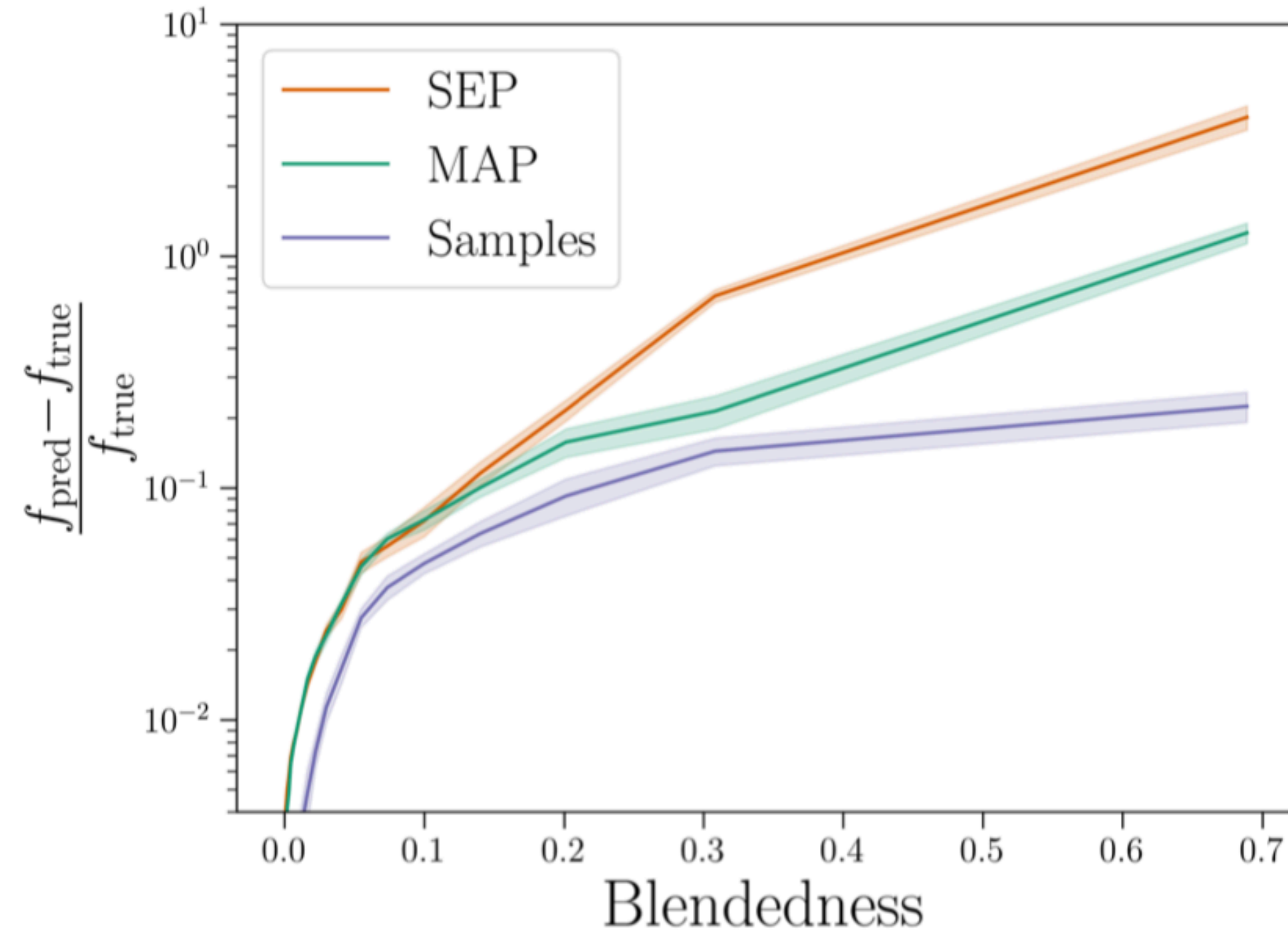


Figure 3.10: **Flux measurement in BLISS using probabilistic detections as a function of blendedness.** In this figure, we compare the (fractional) median flux residual for three different approaches to flux measurement using BLISS in the dataset `dataset-central`. In all three cases, the flux is measured on the central galaxy of each image using aperture photometry after deblending is performed using the same BLISS deblender encoder, but different detections are used. Detections are matched within 2 pixels of the true galaxy centroid on the center of the image. If no matches with the central galaxy, no flux measurement is performed for those set of detections. In the first case (MAP), the most likely detections (above 50% probability) from the BLISS detection encoder are used for deblending and aperture photometry. In the second case (SEP), detections from SEP are used instead. Finally, in ‘Samples’, we draw 100 detection samples from the BLISS detection encoder and used these as input to the deblender encoder. We take the mean over these measured fluxes as the flux estimator for each image. See Section 3.5.6 for more discussion on this figure.

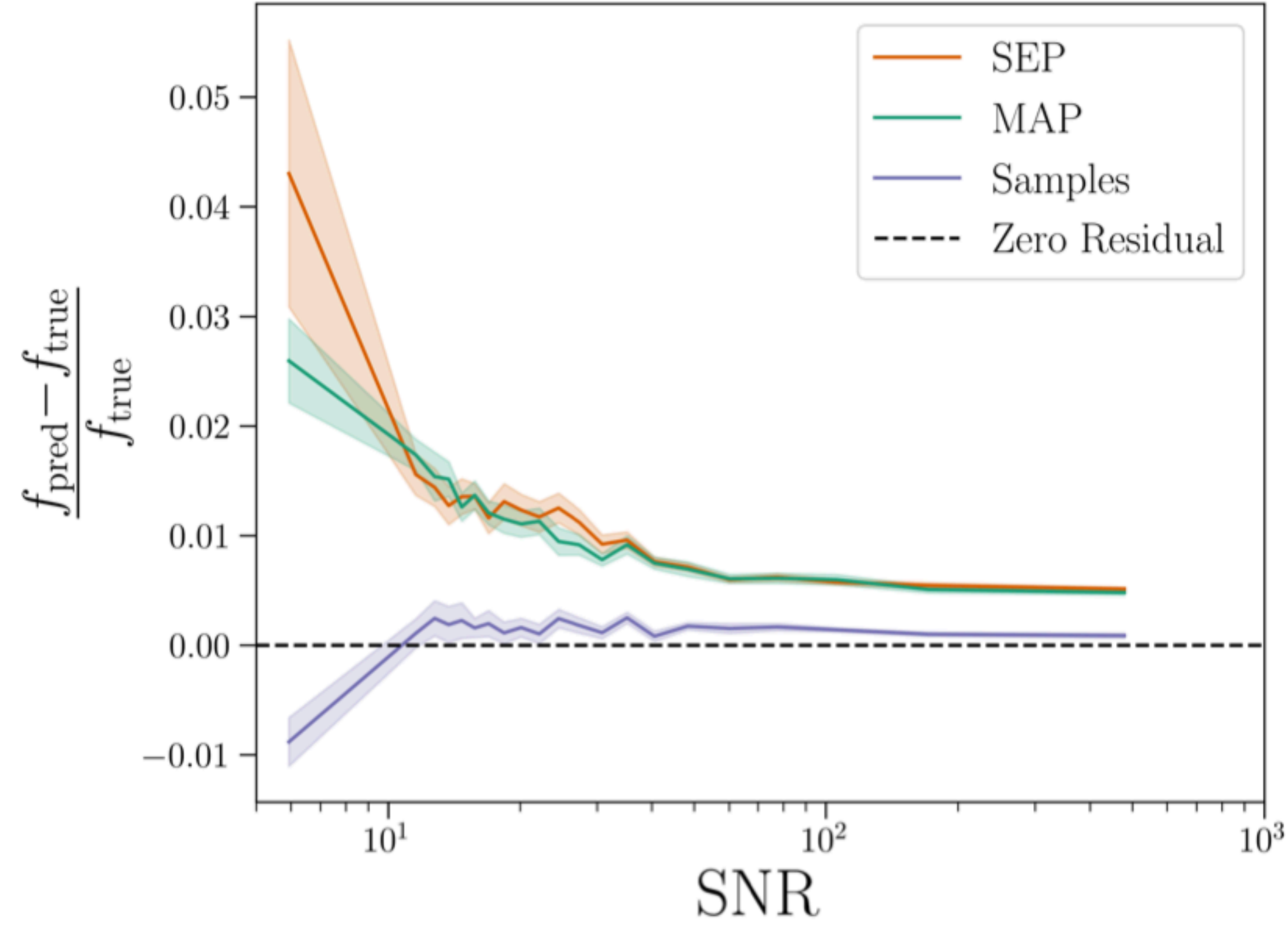


Figure 3.11: **Flux measurement in BLISS using probabilistic detections as a function of SNR.** This figure presents the same results as Figure 3.3, but residuals are split into SNR bins based on the true SNR of the central source of each image. The bins are chosen so that there are an equal number of images in each. See Section 3.5.6 for additional discussion.

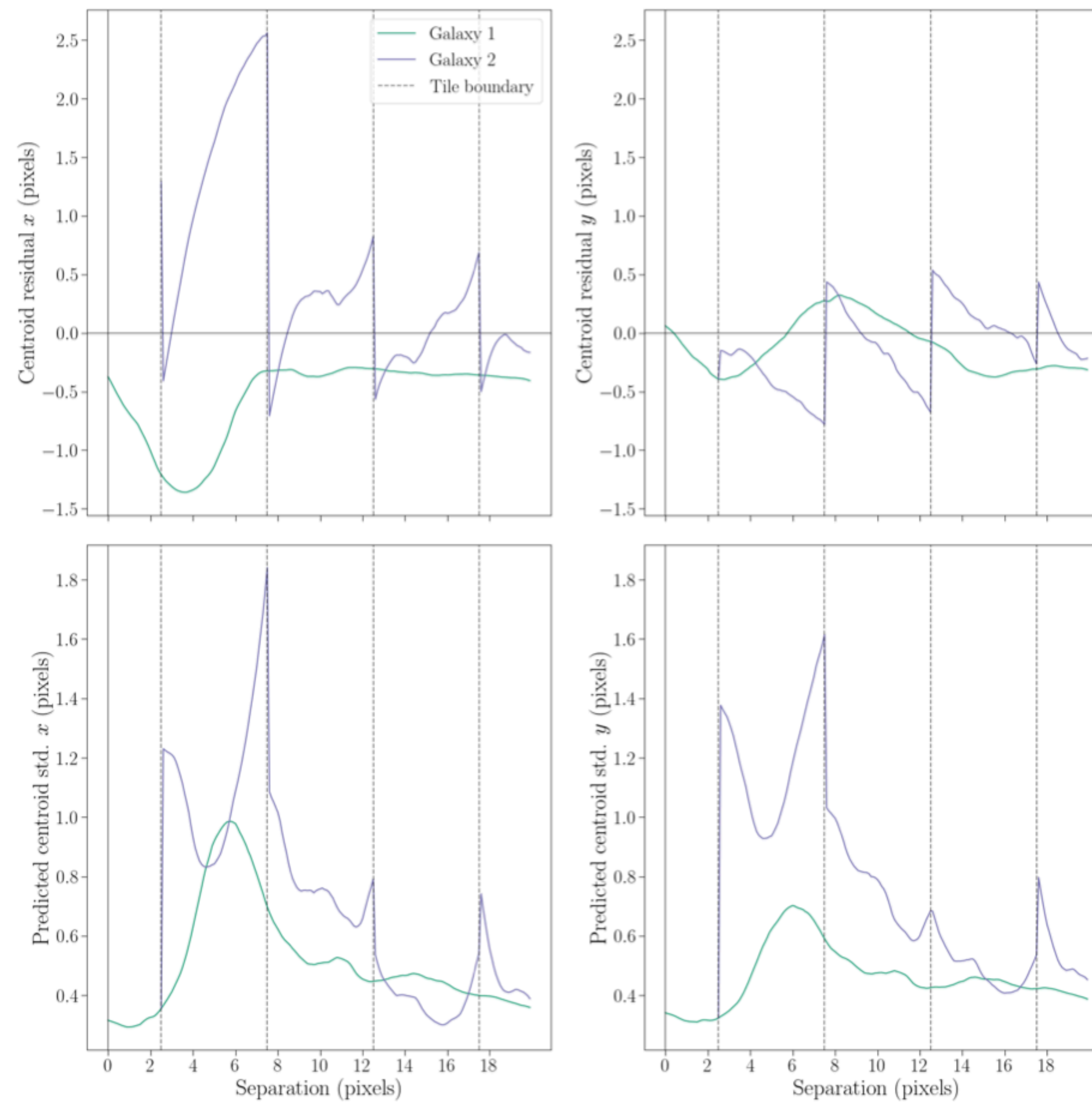


Figure 3.12: **Detection residuals and errors as a function of separation between galaxies.** In this plot we show the centroid location residual and BLISS predicted uncertainties in the x and y directions for the pair of galaxies in Section 3.5.5. Specifically, the top row shows the difference between the mean centroid predicted by the detection encoder for the tile containing each galaxy and the true centroid as a function of separation. The bottom row shows the predicted uncertainty on the centroid as a function of separation.

Signal to noise ratio (SNR)

We use the aperture photometry module in SEP ([K. Barbary, 2016](#)) to measure the signal-to-noise ratio (SNR) defined as:

$$\sqrt{\sum_{p \in \text{aperture}} \frac{I_p^2}{\sigma_{n,p}^2}}, \quad (3.12)$$

where I_p is the intensity of the image in pixel p , and $\sigma_{n,p}^2$ is the sky noise variance. We assign an SNR to individual galaxies. When galaxies are blended, we draw each member individually and use this equation to compute their SNR. We use an aperture radius of 5 pixels throughout our experiments.

Blendedness

Quantity that represents the fraction of the flux of an individual object that overlaps with other objects in a blend. Takes into account the distance between objects and their relative fluxes

We define blendedness using the definition in [P. Melchior et al. \(2018\)](#). For a galaxy with intensity $\mathbf{I}$, *blendedness* is

$$B = 1 - \frac{\langle \mathbf{I}, \mathbf{I} \rangle}{\langle \mathbf{I}, \mathbf{I}_{\text{total}} \rangle}, \quad (3.13)$$

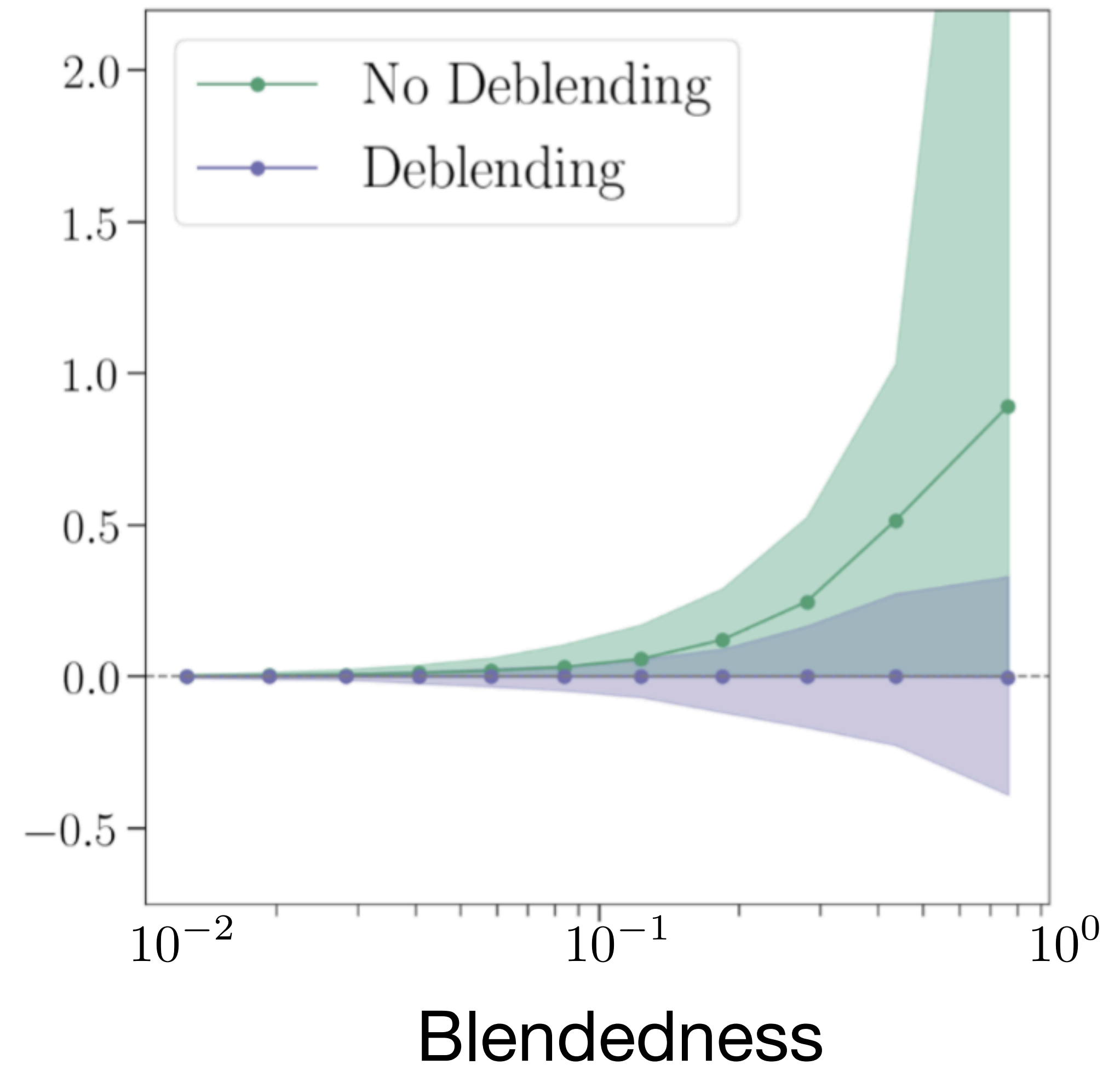
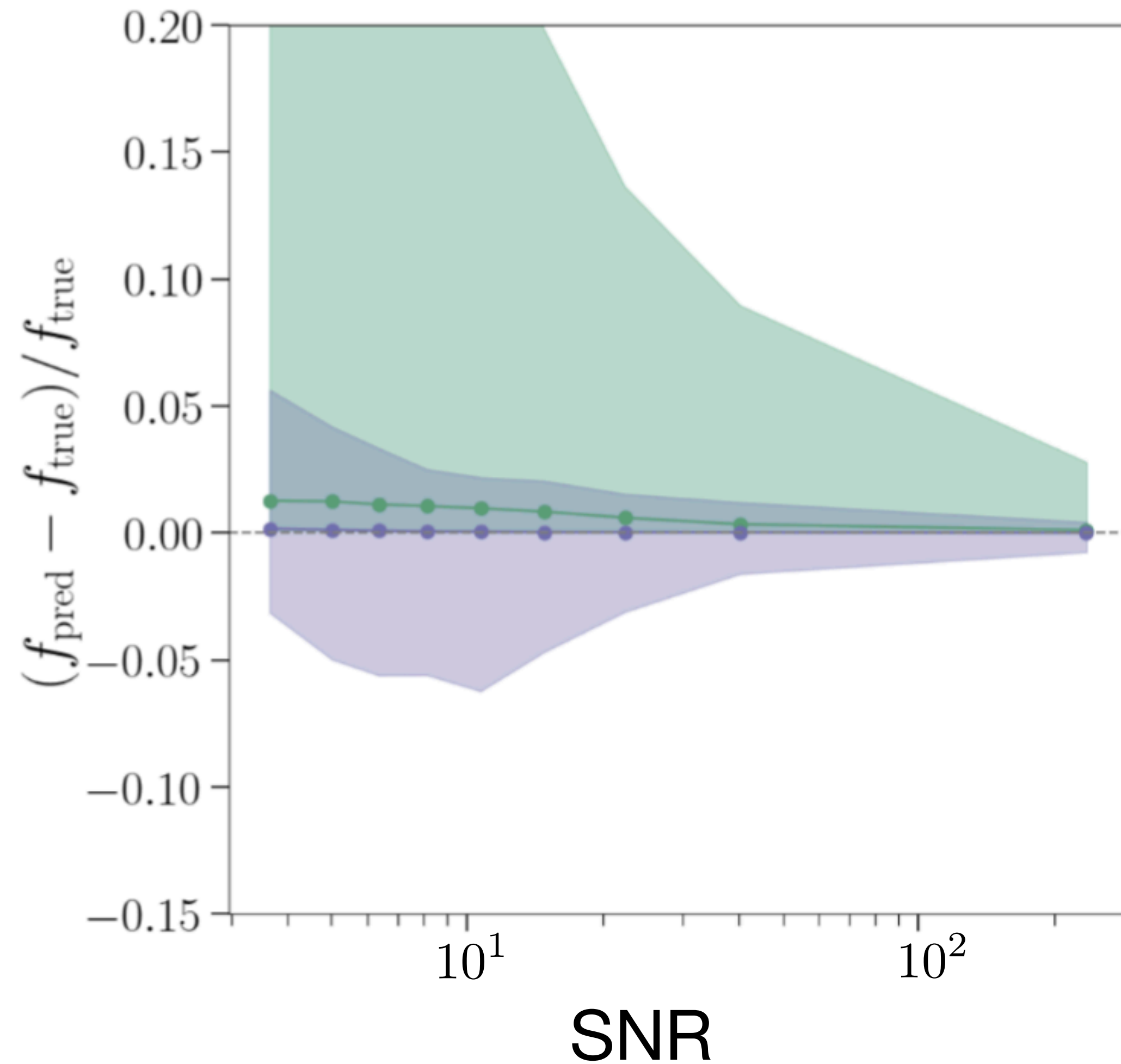
where $\mathbf{I}_{\text{total}}$ is the total intensity from every source in the blend and, for intensities $\mathbf{I}_1, \mathbf{I}_2$,

$$\langle \mathbf{I}_1, \mathbf{I}_2 \rangle = \sum_{p \in \text{pixels}} I_{1,p} I_{2,p}. \quad (3.14)$$

By construction, $B \in [0, 1]$, with $B = 0$ corresponding to an isolated galaxy and $B = 1$ to a completely overlapping galaxy.

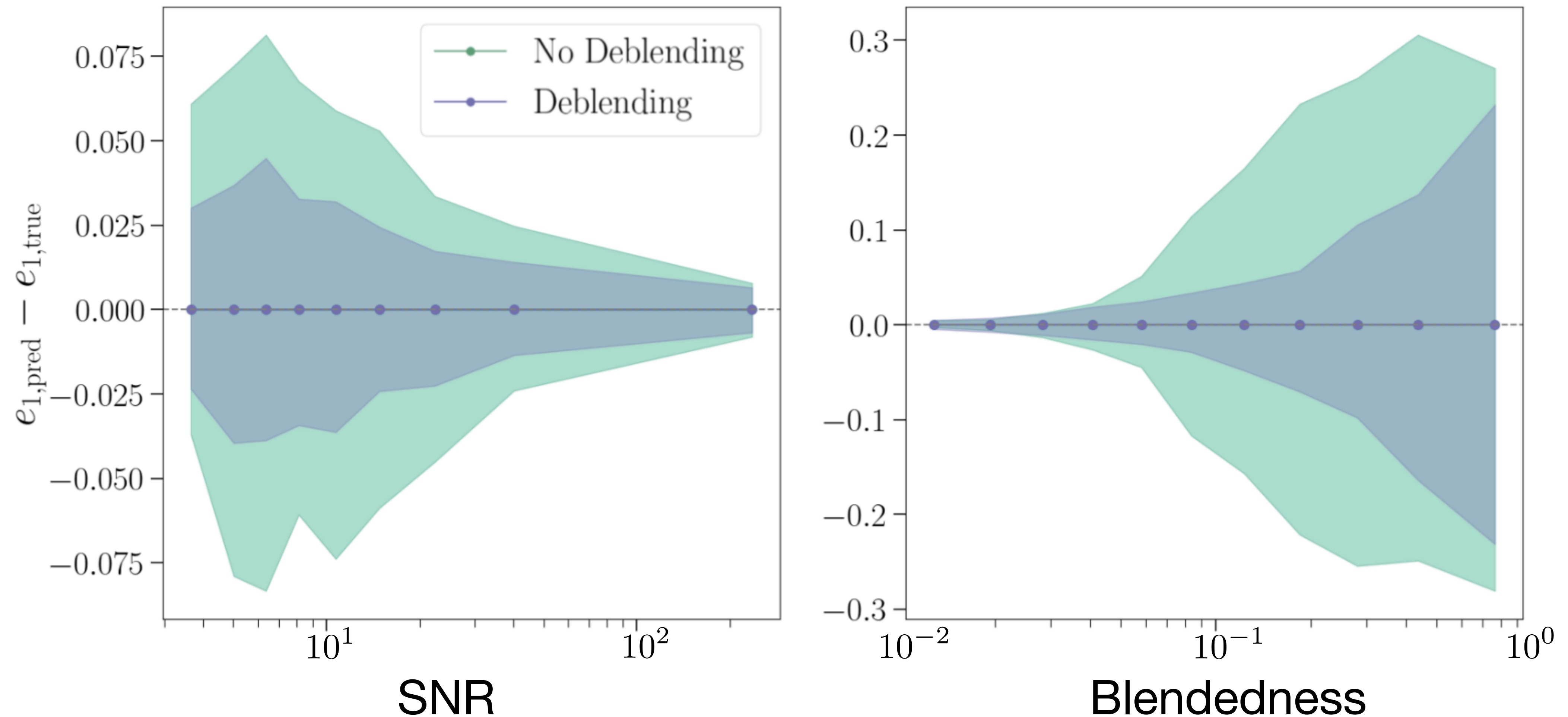
Deblending with true centroids

- Deblending significantly reduces the bias and scatter in **flux** estimates.



Deblending with true centroids

- Deblending significantly reduces the scatter in **ellipticity** estimates.

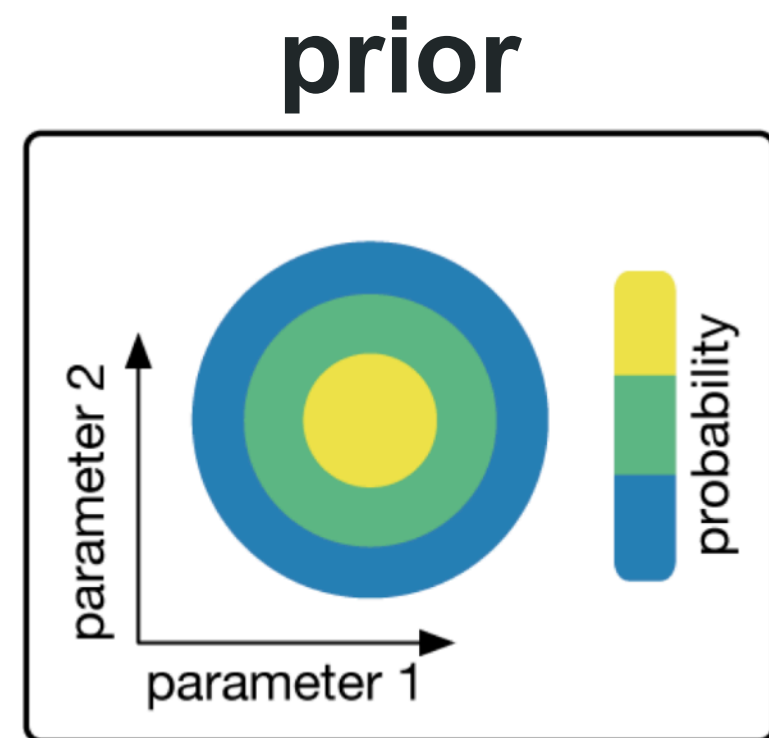


Simulation-Based inference (SBI)

- Train a ML model to develop a **probabilistic association** between data and underlying parameters.
- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Simulation-Based inference (SBI)

- Train a ML model to develop a **probabilistic association** between data and underlying parameters.

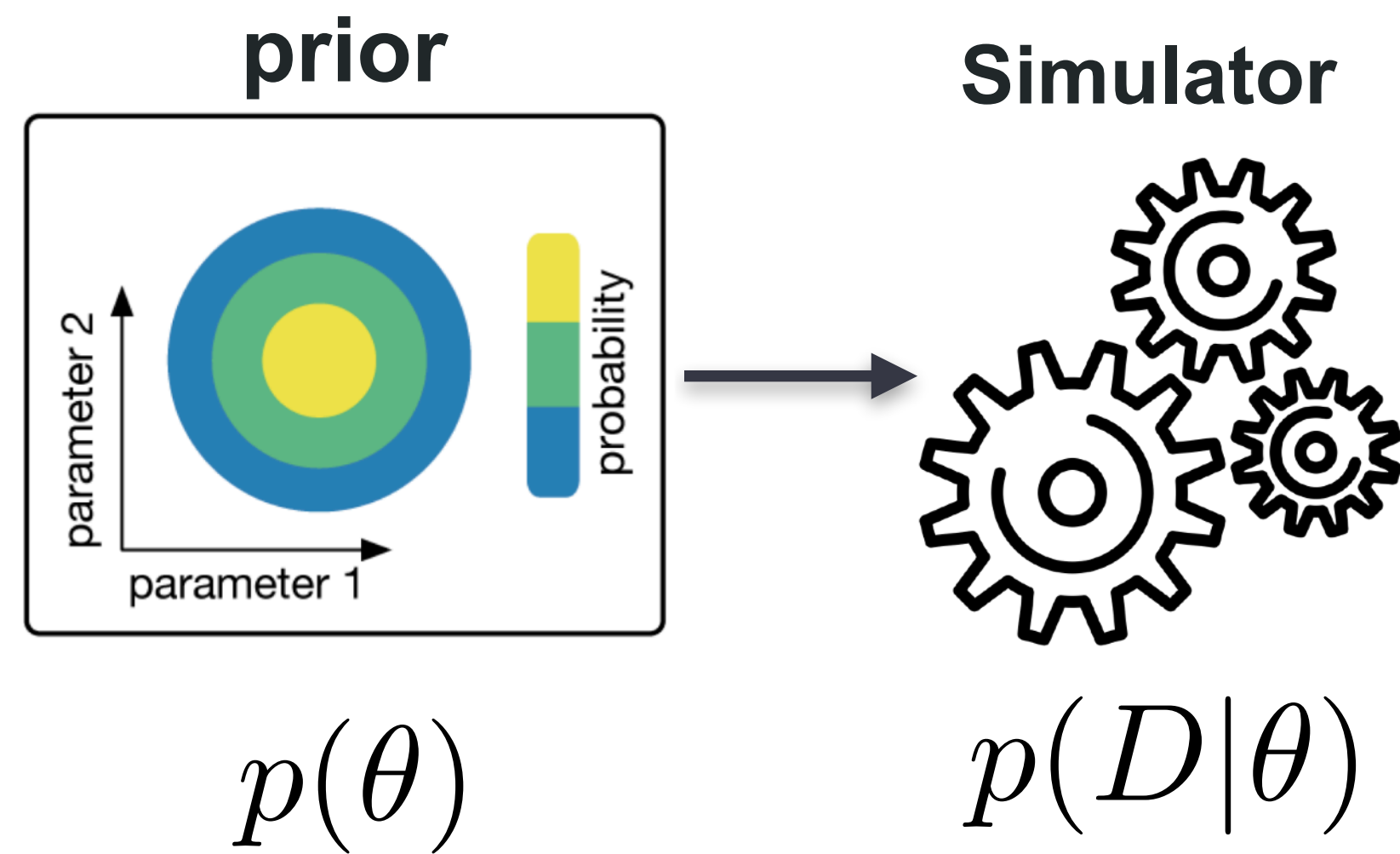


$$p(\theta)$$

- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Simulation-Based inference (SBI)

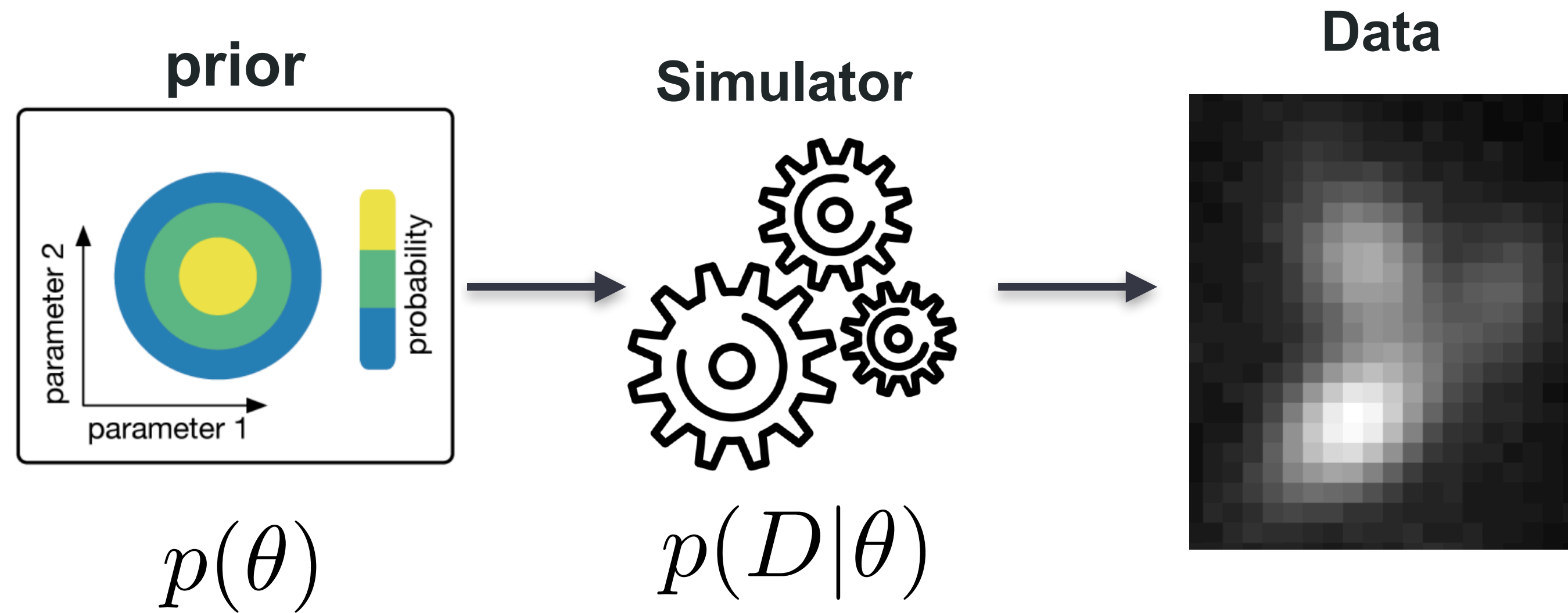
- Train a ML model to develop a **probabilistic association** between data and underlying parameters.



- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Simulation-Based inference (SBI)

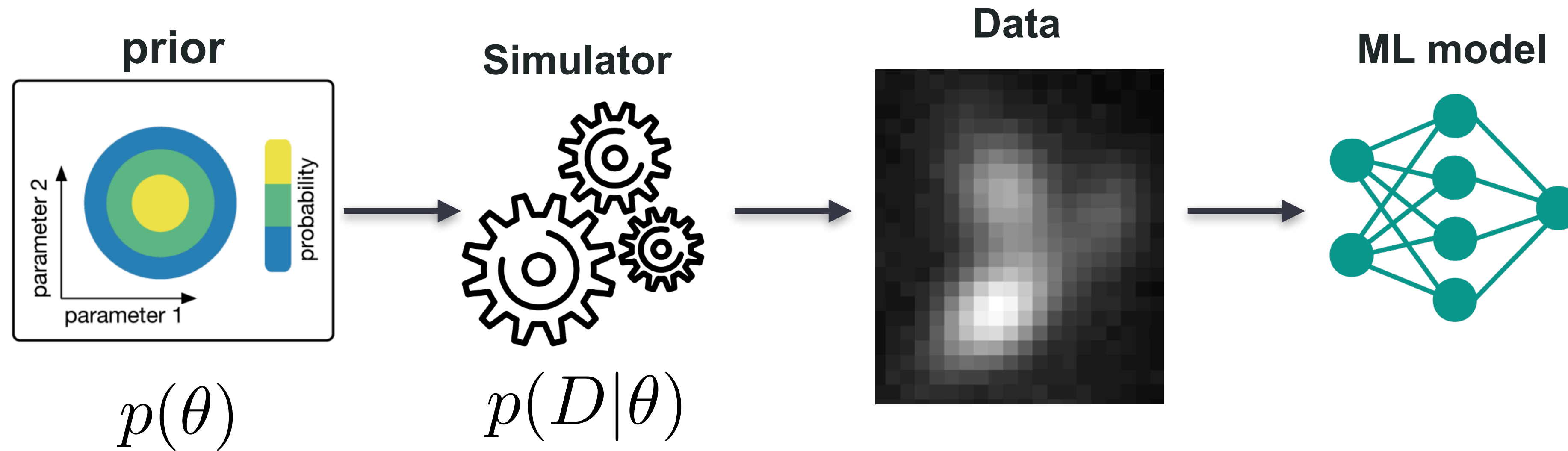
- Train a ML model to develop a **probabilistic association** between data and underlying parameters.



- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Simulation-Based inference (SBI)

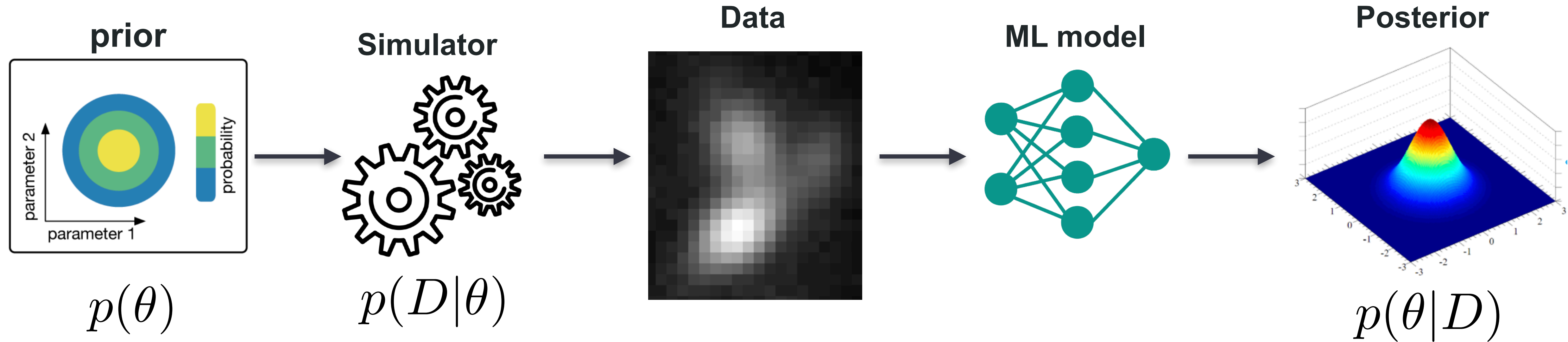
- Train a ML model to develop a **probabilistic association** between data and underlying parameters.



- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Simulation-Based inference (SBI)

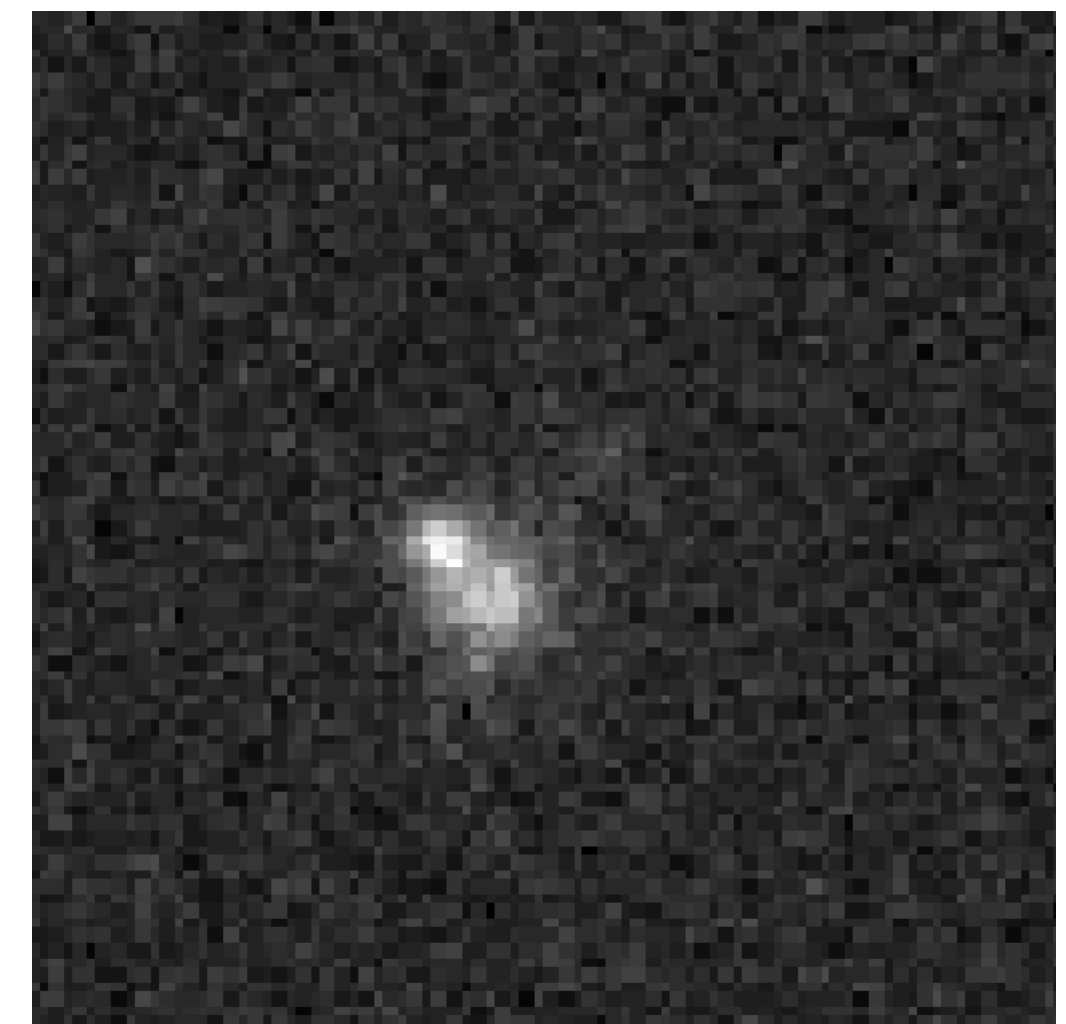
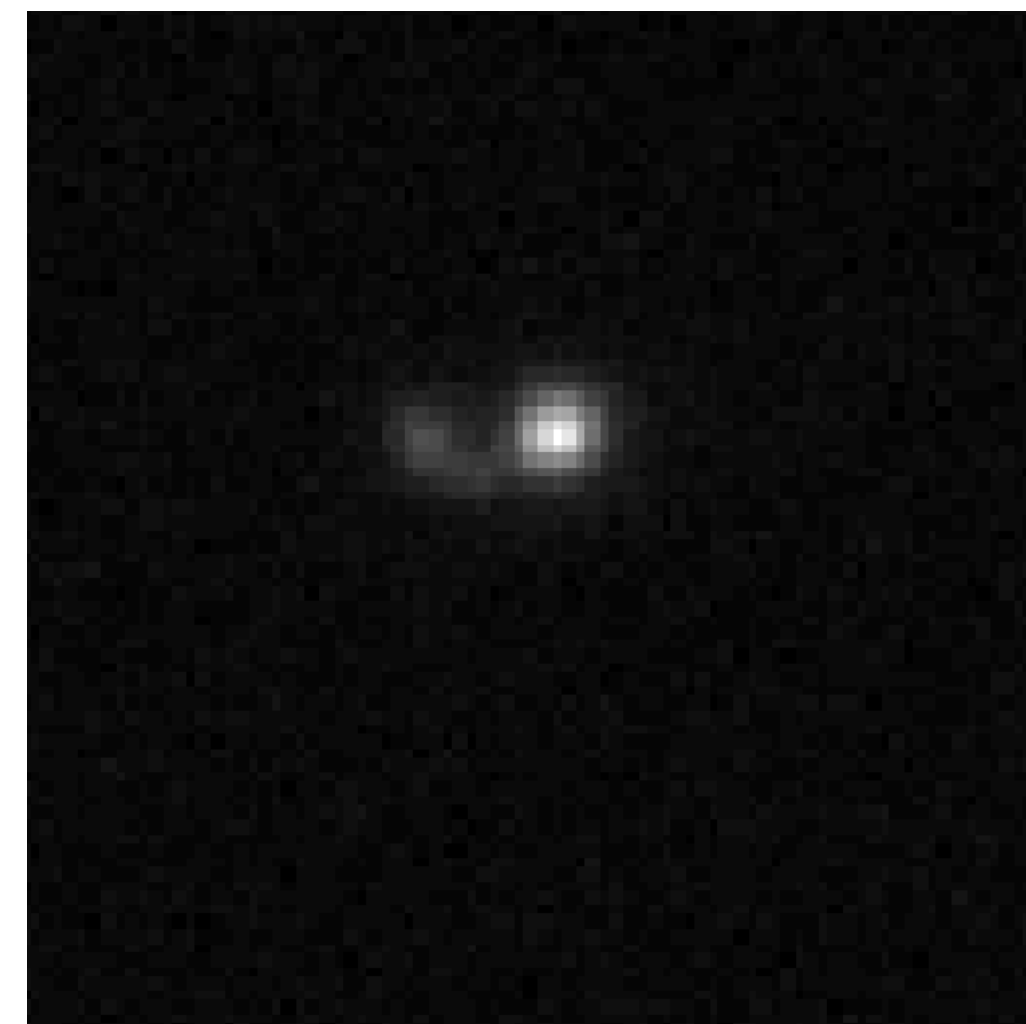
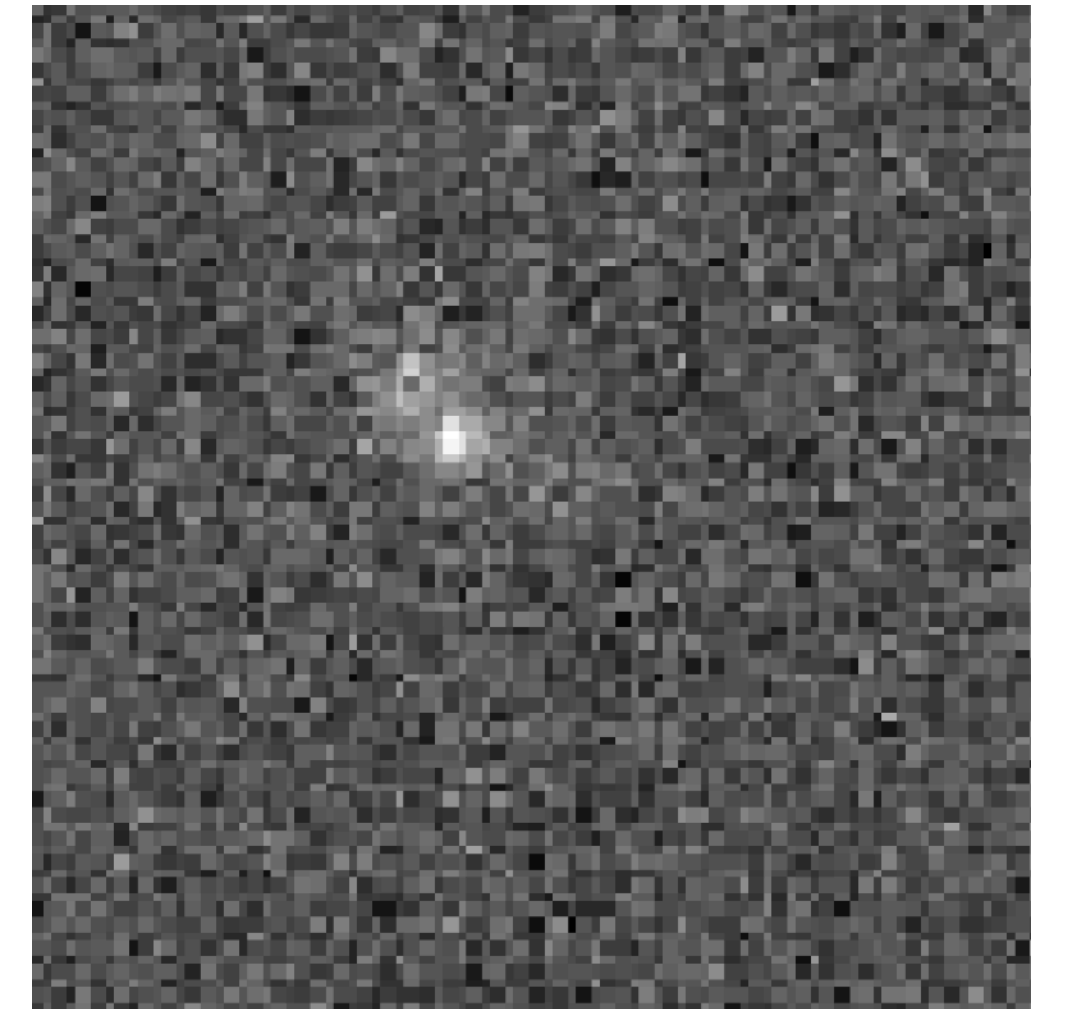
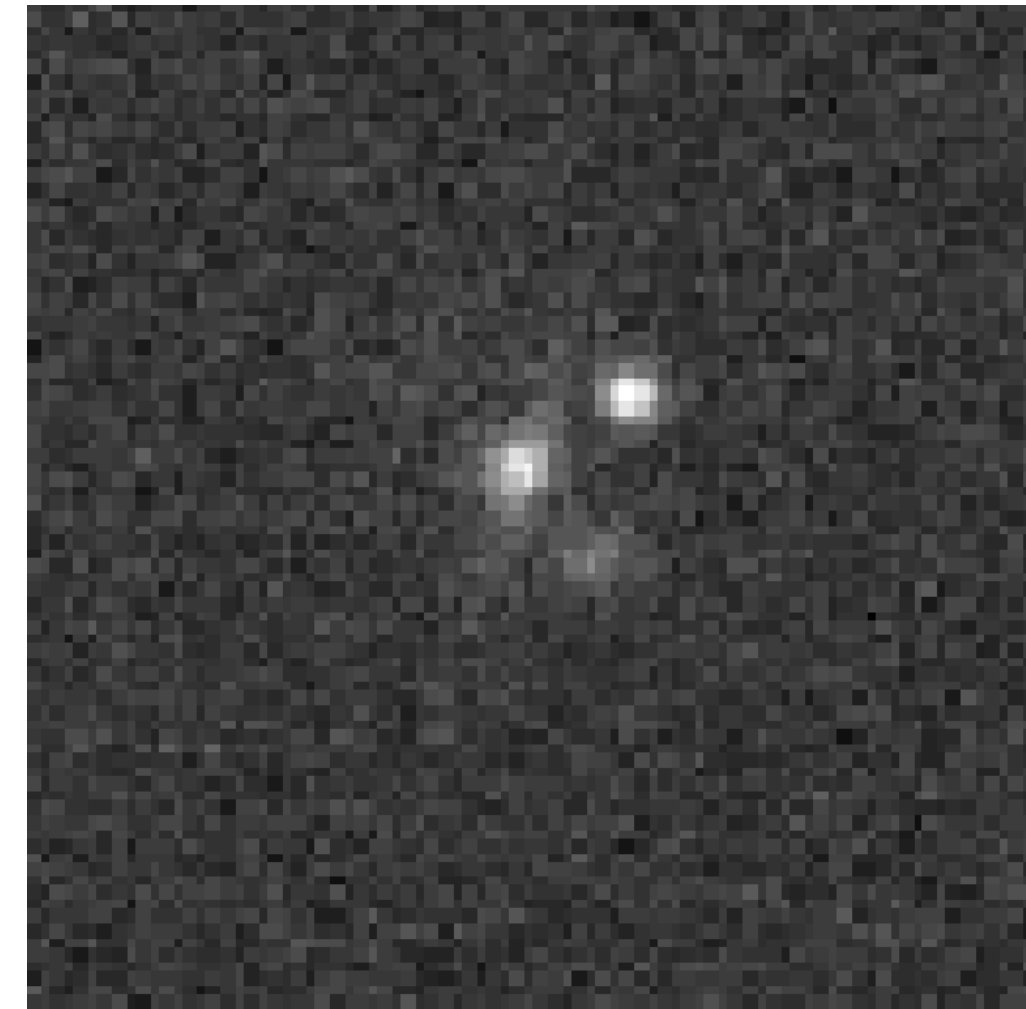
- Train a ML model to develop a **probabilistic association** between data and underlying parameters.



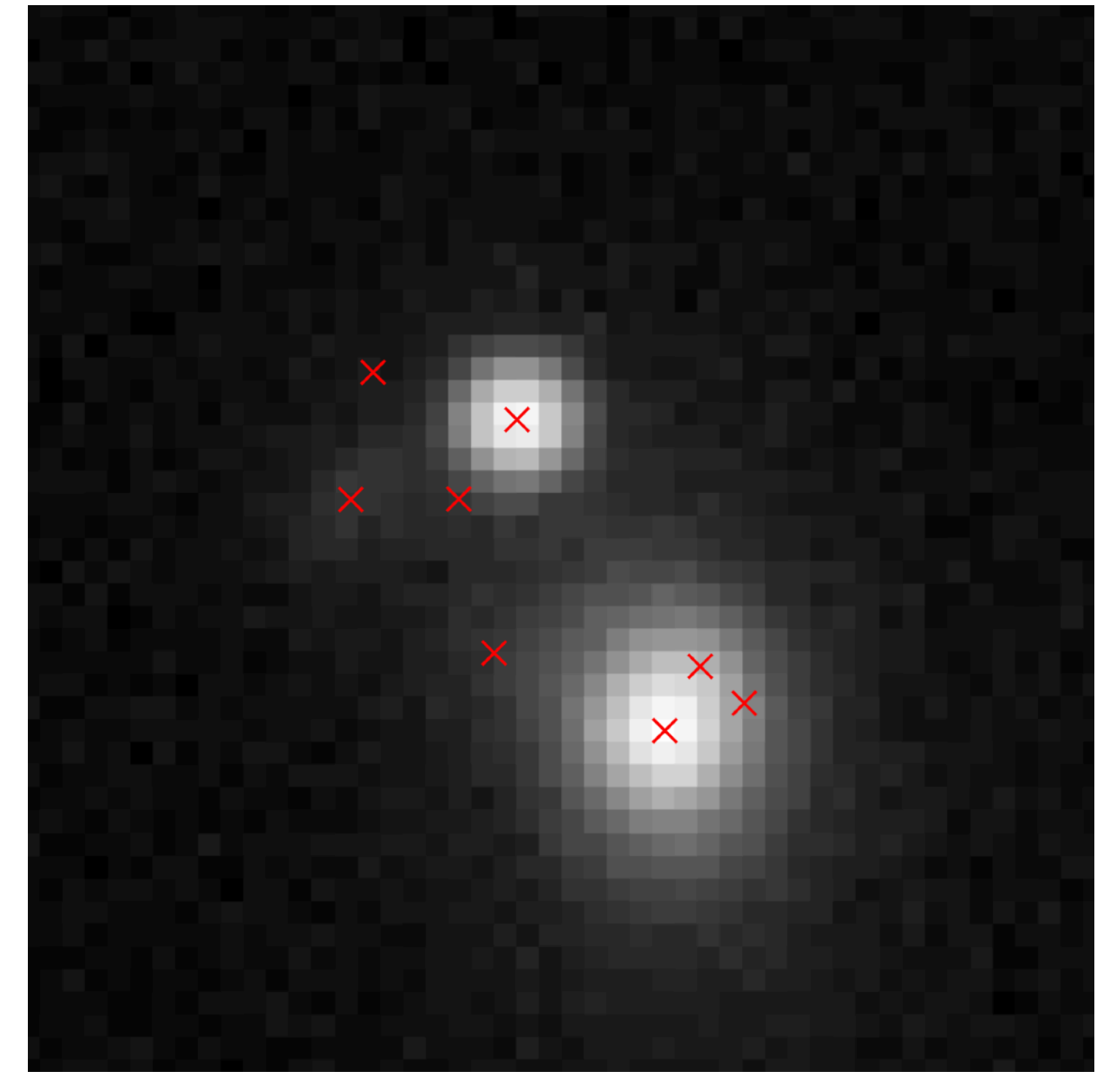
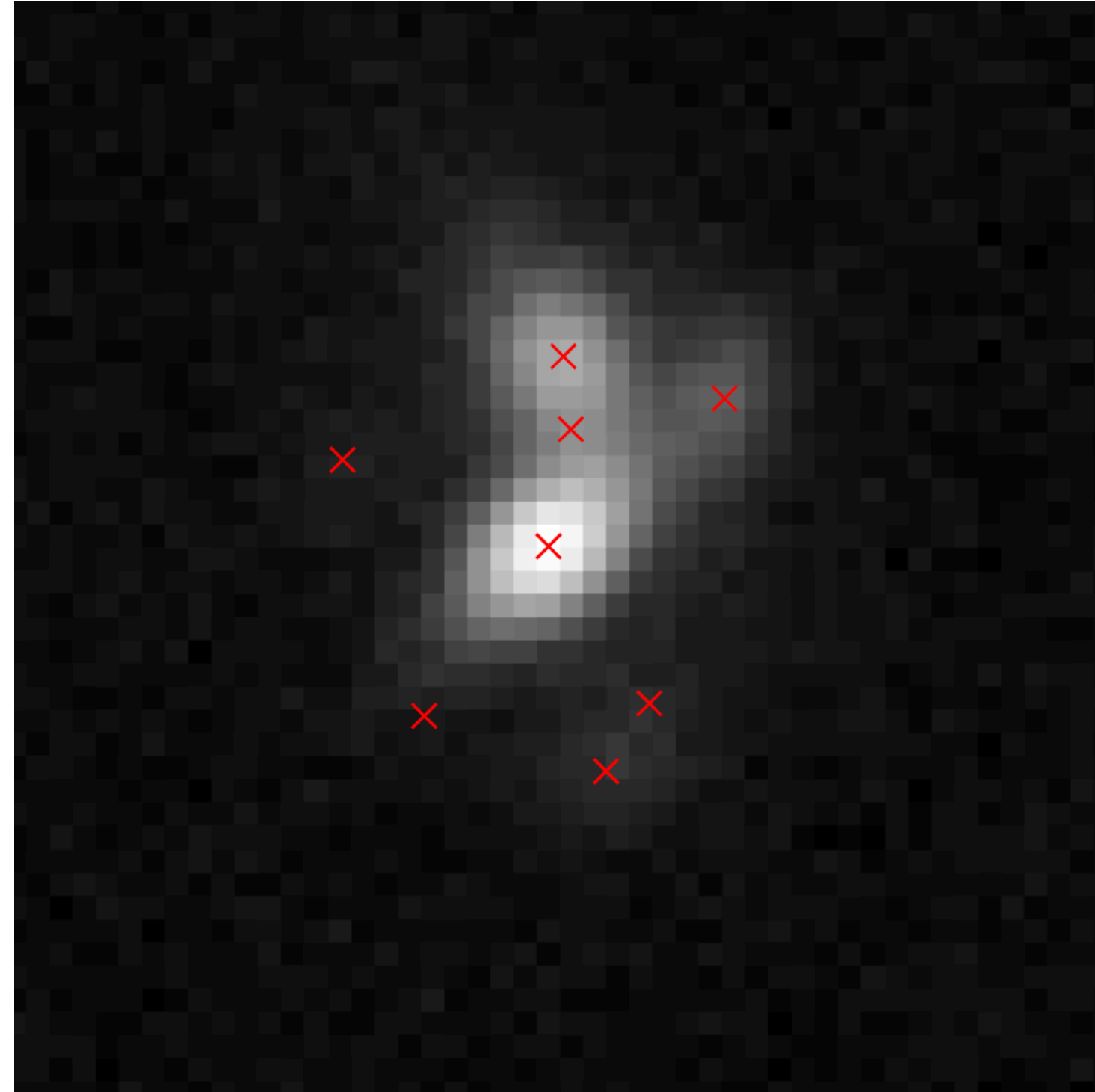
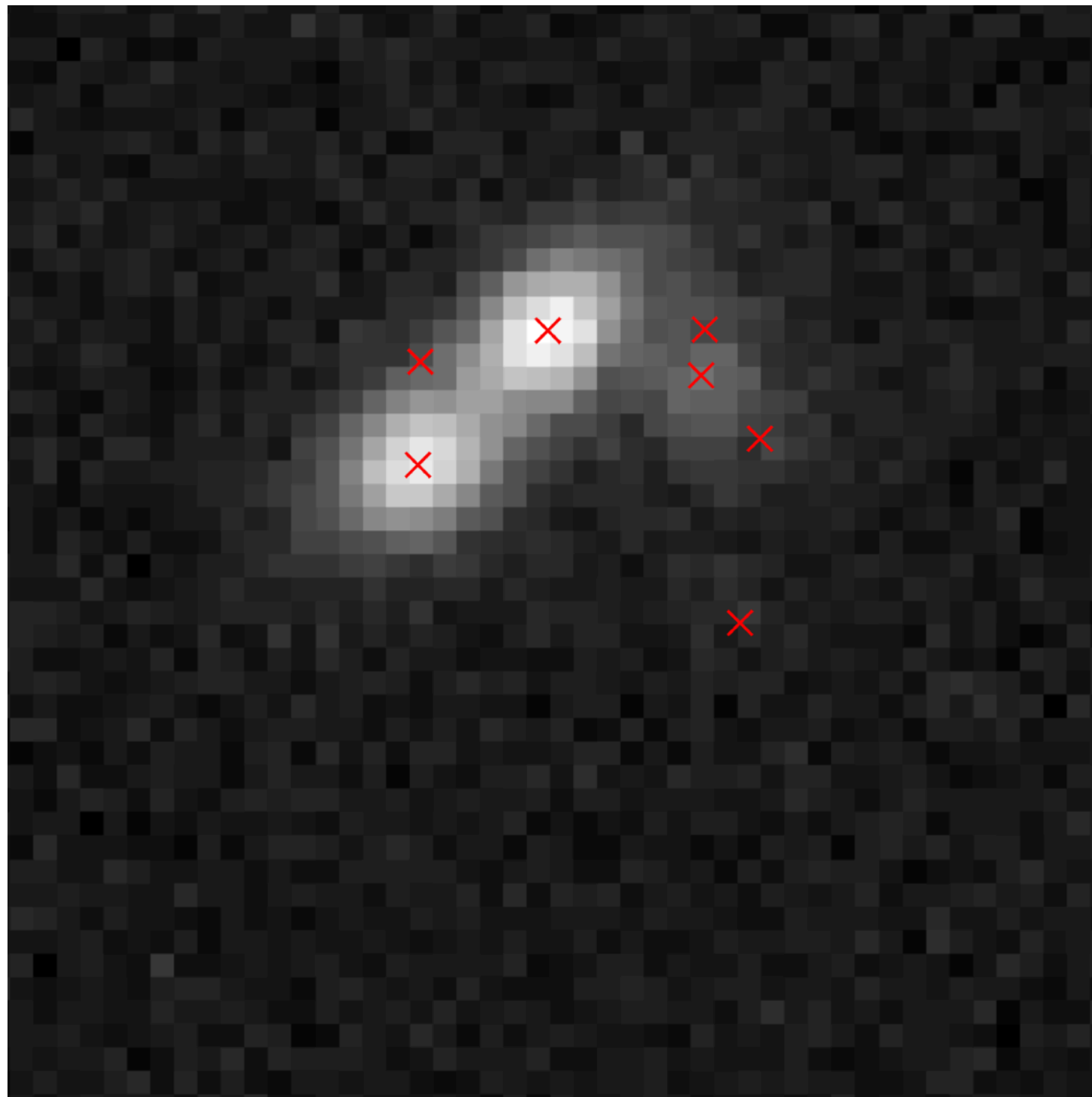
- If simulations are realistic + ML model is flexible $\Rightarrow$ characterize measurement uncertainty

Dataset

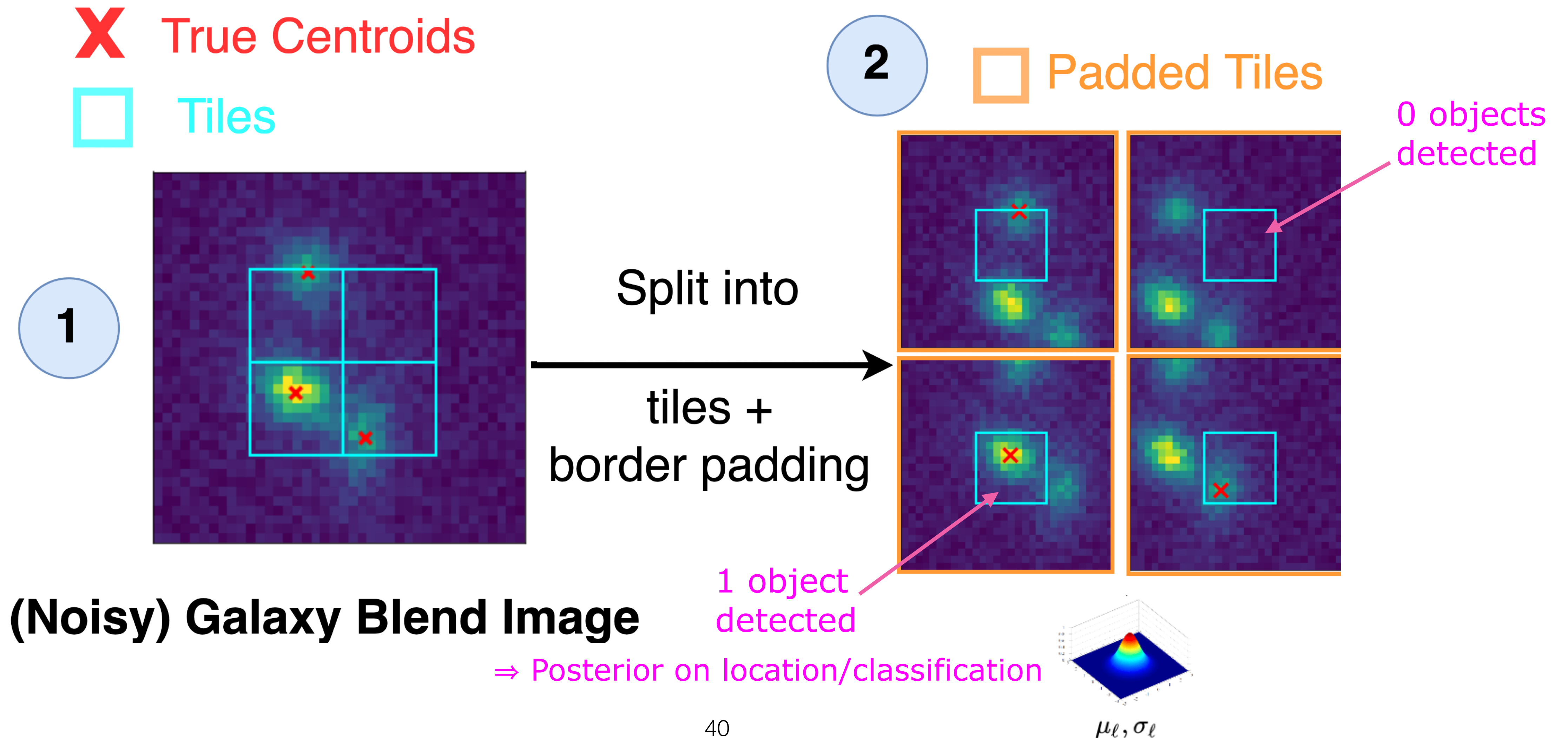
- LSST-like image simulations of galaxy blends
- Parametric Bulge+Disk galaxies
- Realistic source density and noise levels
- Catalog of galaxy properties based on N-body simulation and semi-analytic baryon model (CATSIM).



Galaxy blends example

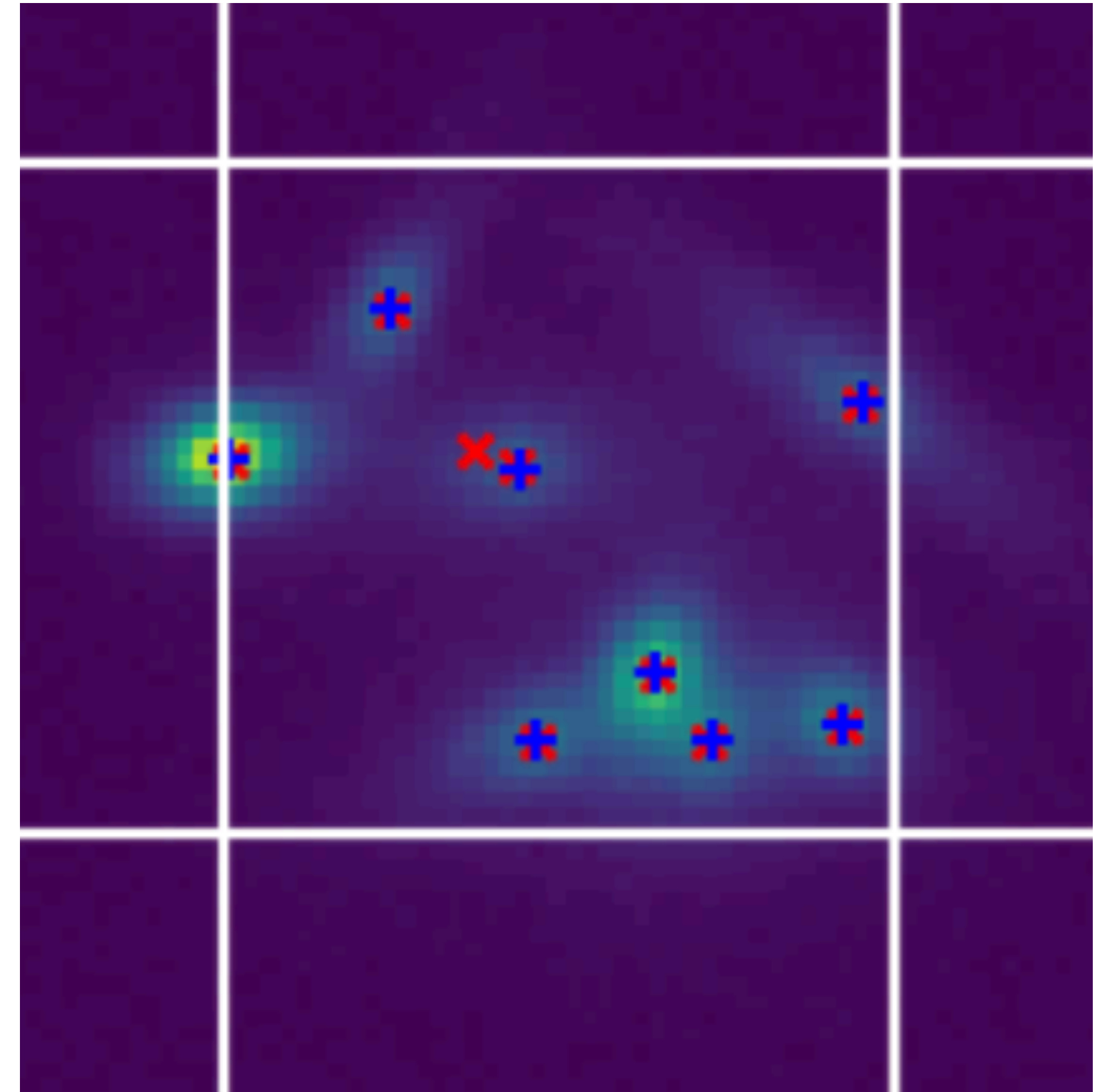
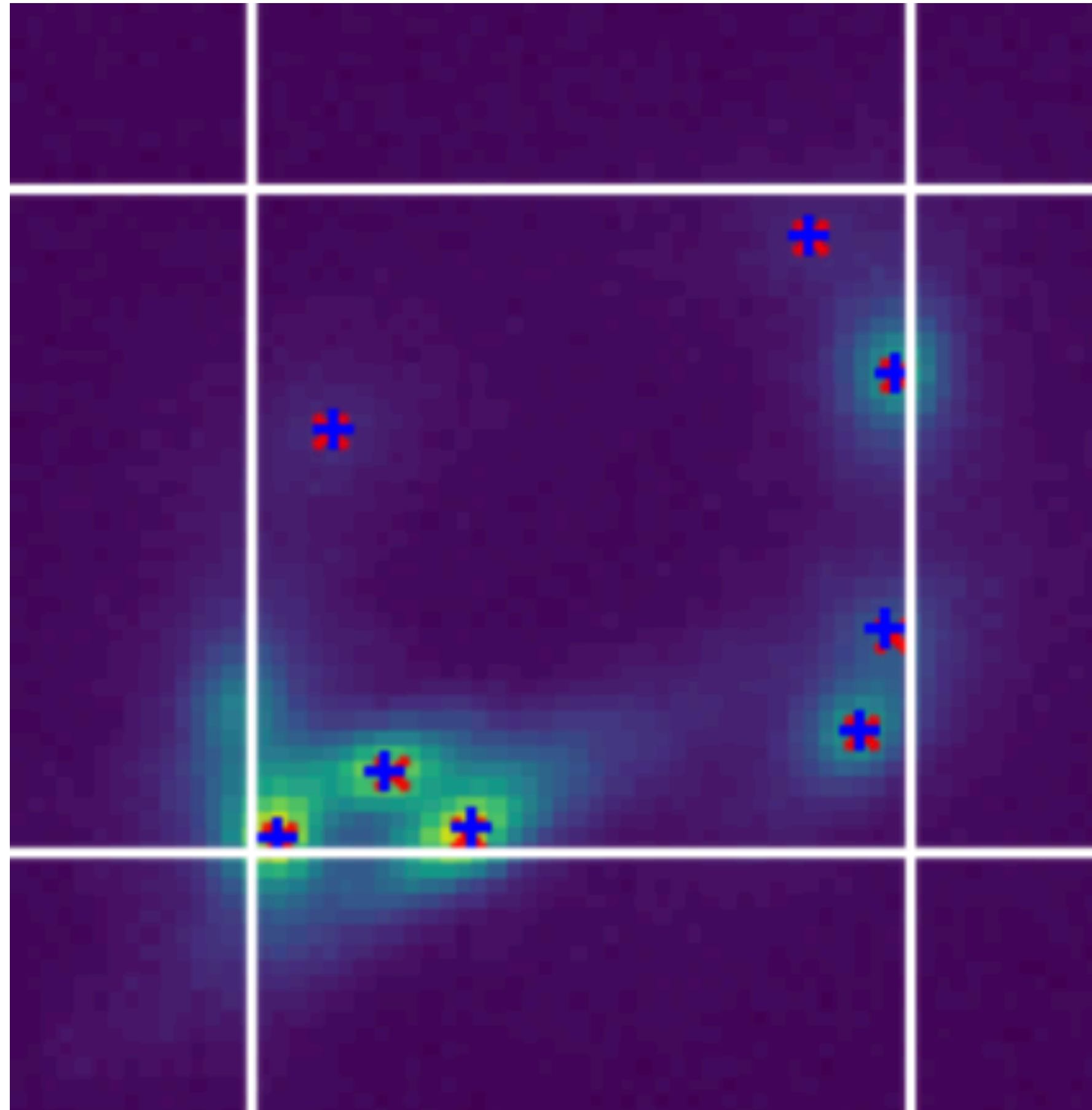


BLISS predictions are done “per tile”



Detection results on GalSim simulated blends

✕ Truth + BLISS



Example of detection output

+ Threshold Detection $> 50\%$ (“Point predictions”)

X Centroid samples

O Truth

