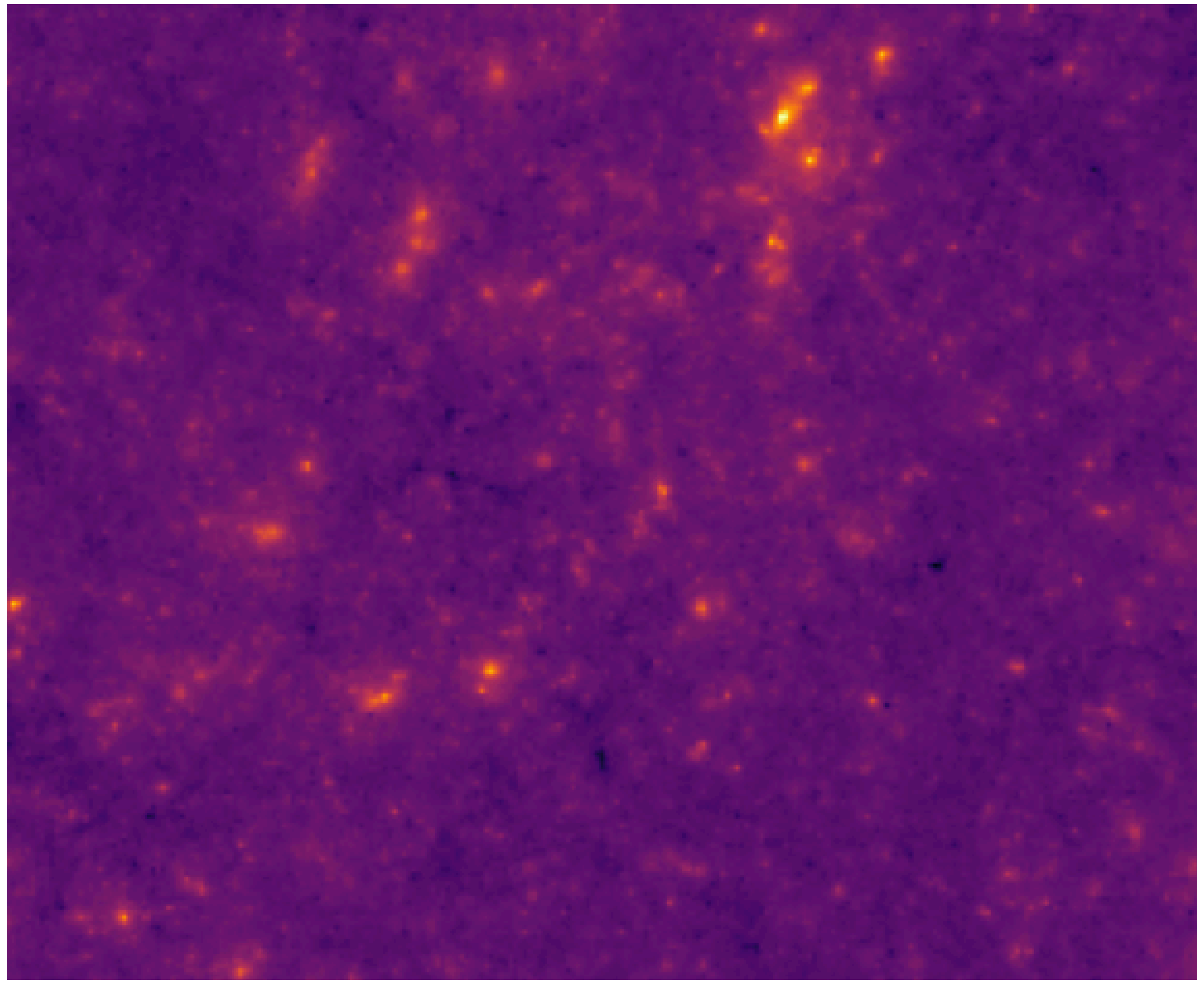


Mass Mapping with Conditional GANs

Jessica Whitney, Tobías Liaudat, Matt Price, Matthijs Mars, Jason D. McEwen

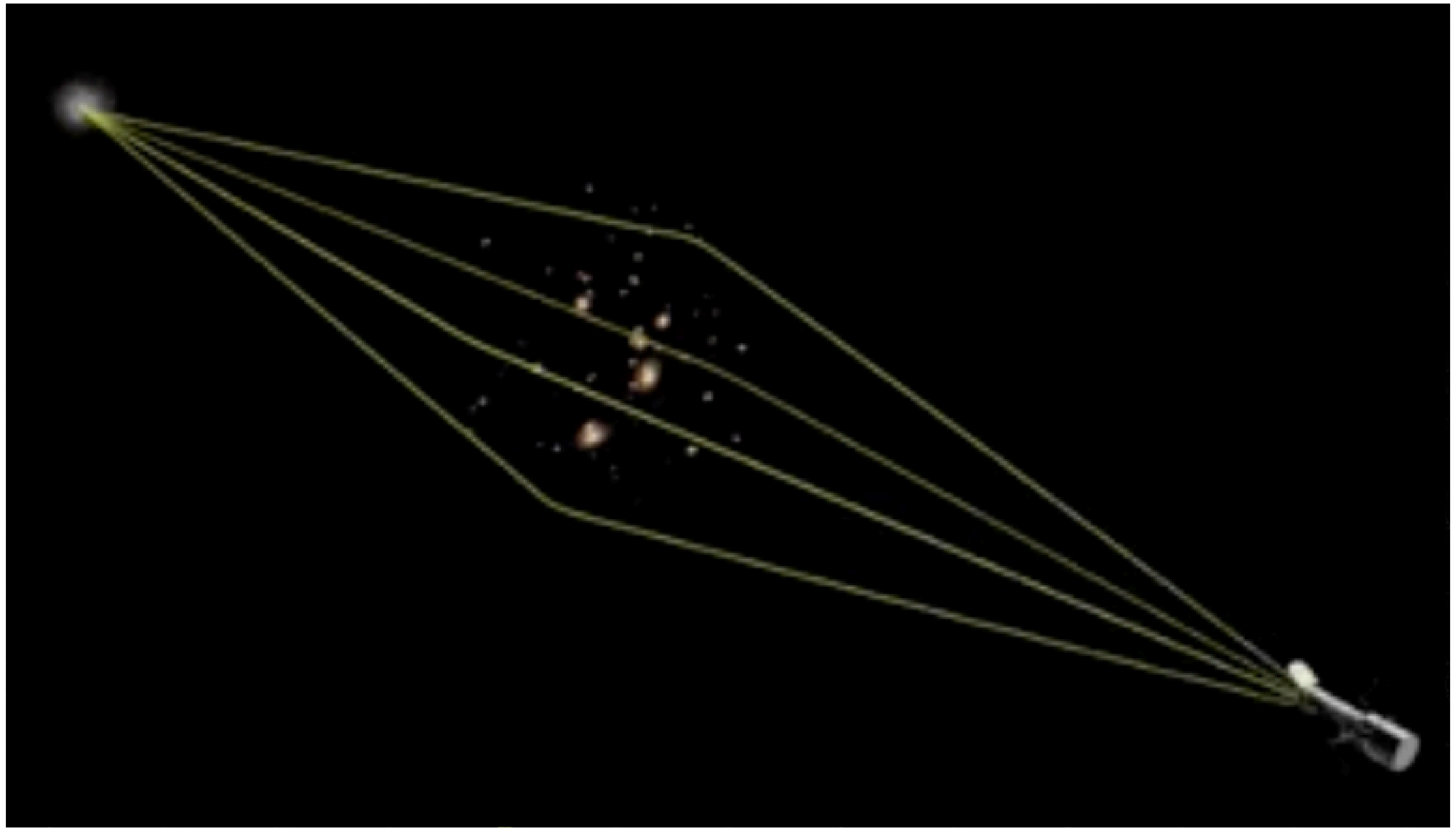
SciAI UCL

Overview

1. What is mass mapping and how do we do it?
2. Our proposed method: using conditional GANs.
3. How cGANs work.
4. Our results (on simulations, and real COSMOS data).

SciAI UCL

Weak Lensing

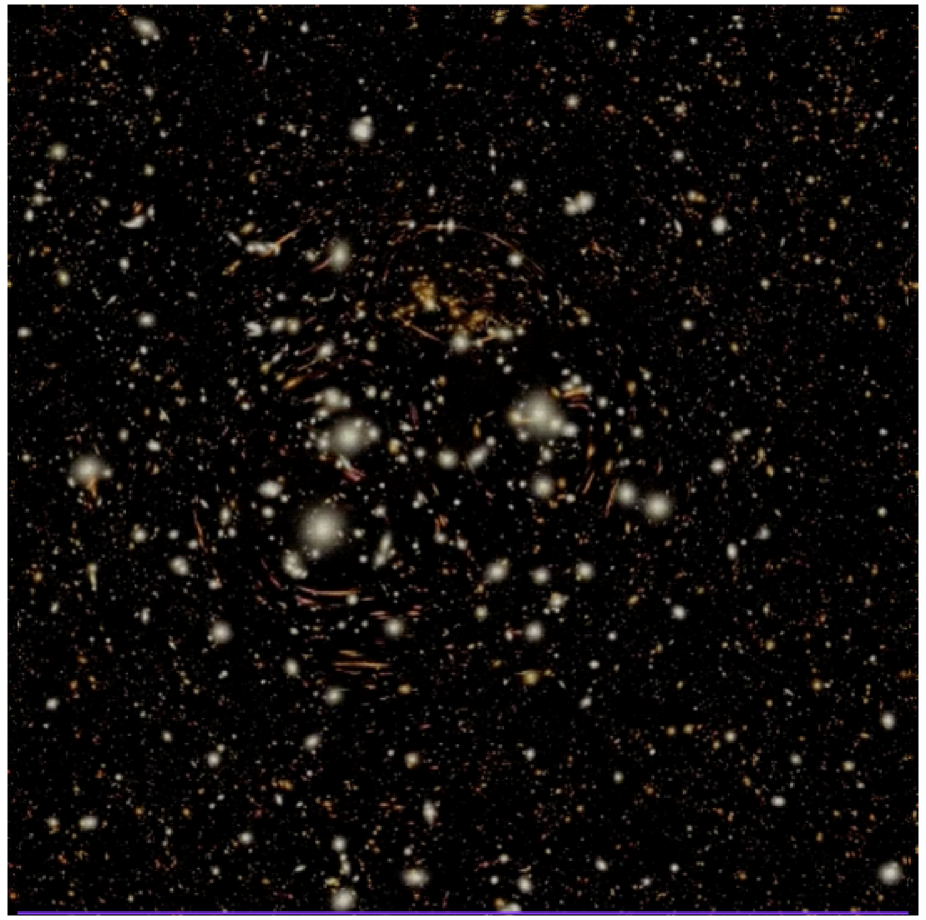


Source: NASA, ESA, and Goddard Space Flight Center/K. Jackson.

Light from distant galaxies is perturbed by the gravitational fields of intervening matter - this causes them to look distorted to the observer.

SciAI UCL

Mass Mapping



Abell 370 galaxy cluster. Source: NASA, ESA, and J. Lotz and HFF Team (STScl).

FORWARD EQUATION

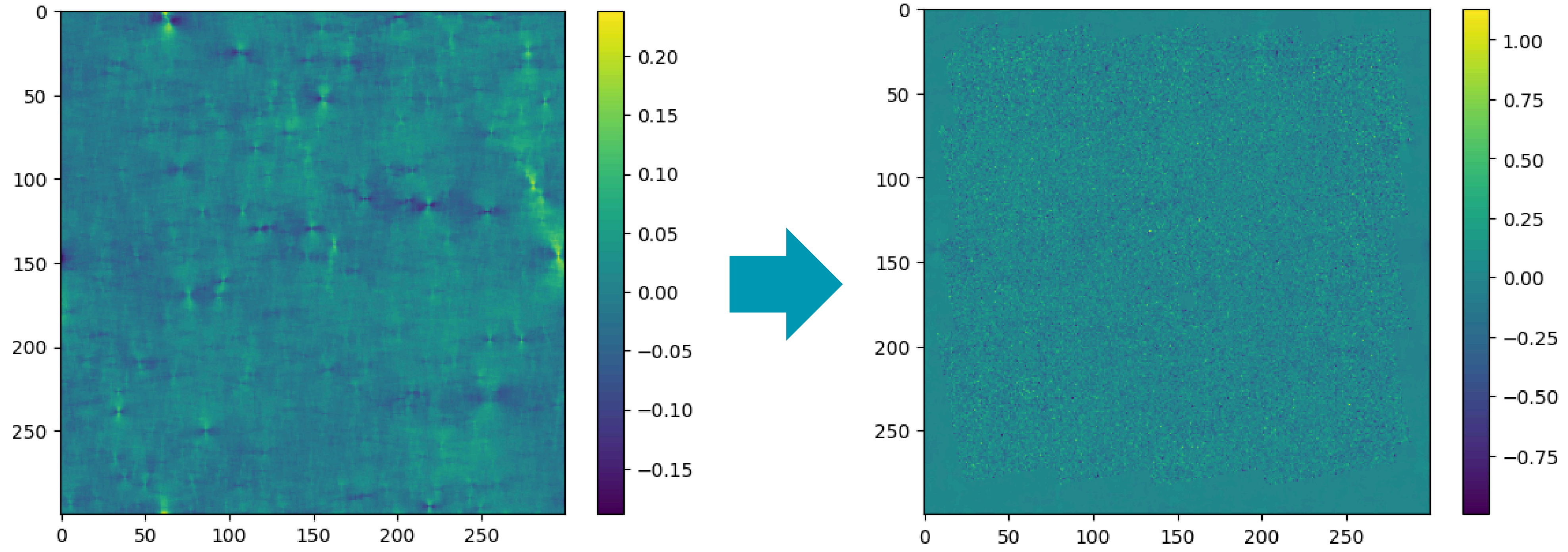
$$\gamma = A\kappa$$

Convergence, κ: apparent magnification

Shear, γ: anisotropic stretching

SciAI UCL

Mass Mapping



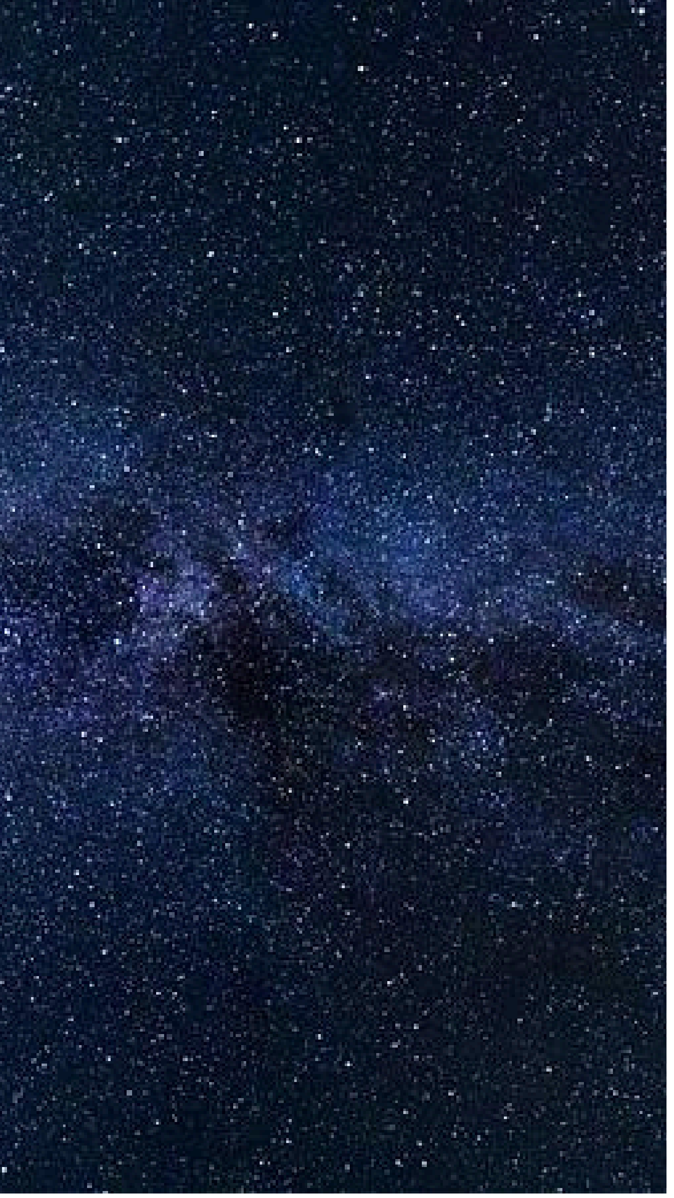
*Real component only of shear shown

FORWARD EQUATION

$$\gamma = A\kappa + n$$

Convergence, κ: apparent magnification

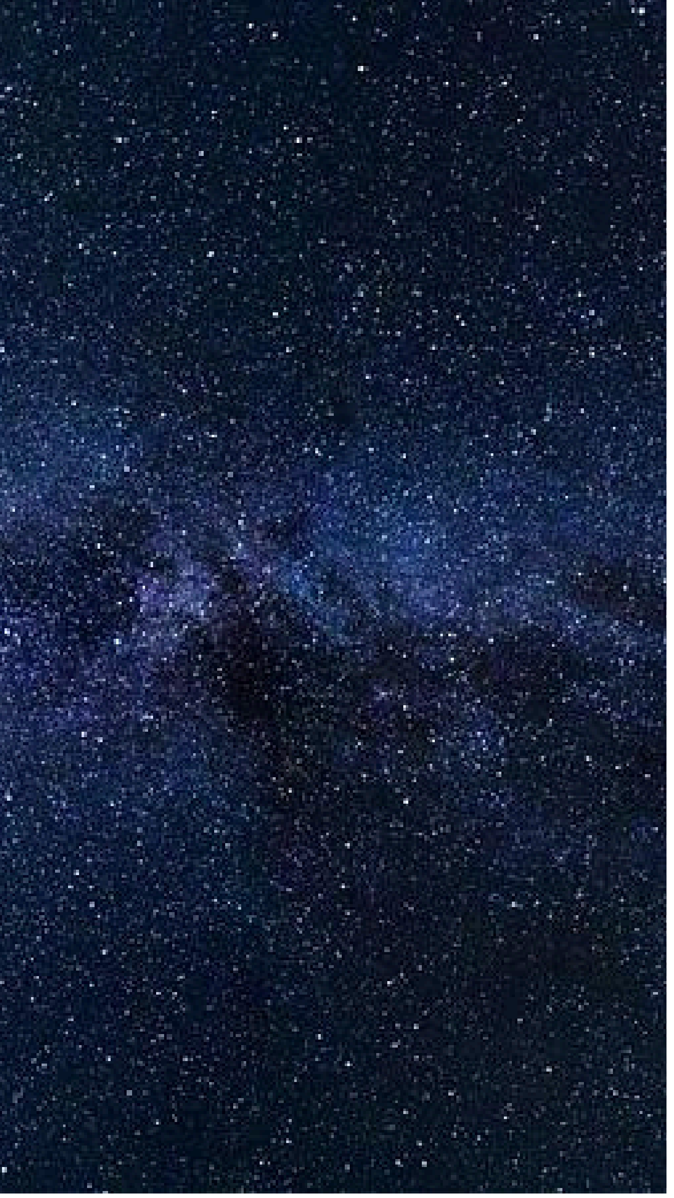
Shear, γ: anisotropic stretching



What do we propose?

Mass mapping is an open problem. Traditional techniques typically require hand-crafted priors, which can limit reconstruction quality.

There's a need for novel techniques that are capable of dealing with noise, that use data-driven priors for better reconstruction quality.



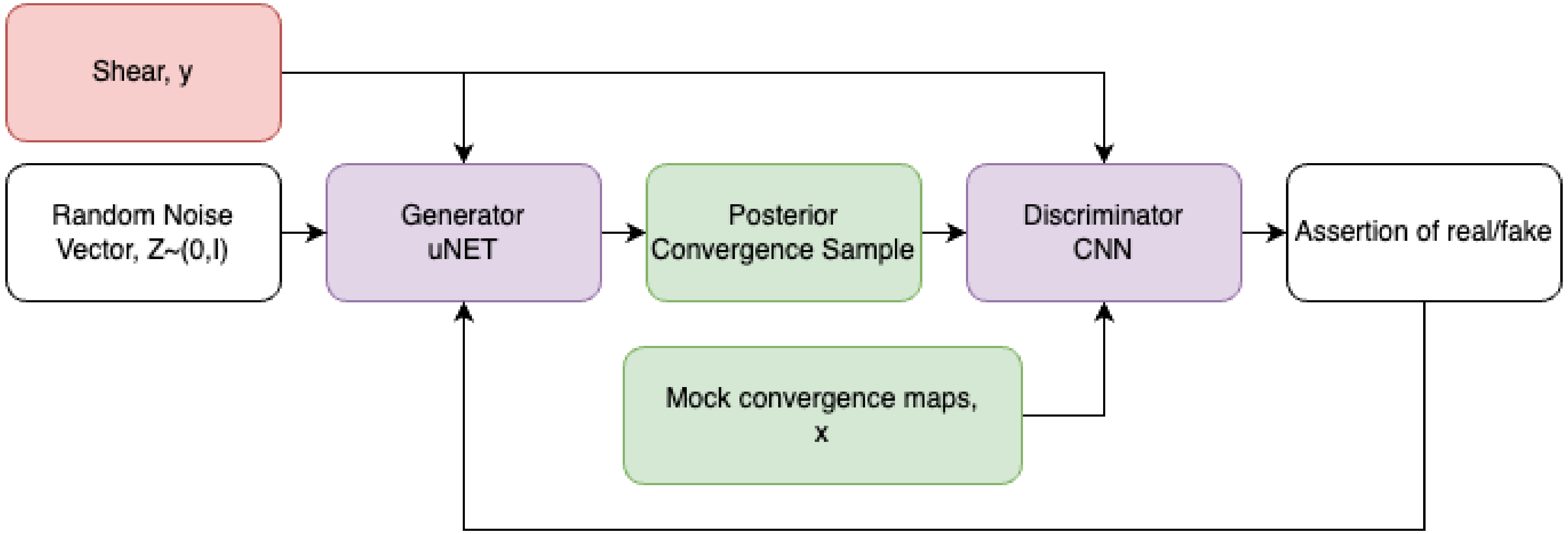
What do we propose?

Mass mapping is an open problem. Traditional techniques typically require hand-crafted priors, which can limit reconstruction quality.
There's a need for novel techniques that are capable of dealing with noise, that use data-driven priors for better reconstruction quality.

Why use GANs*?

Machine learning models have led to methods with more efficient, and higher-quality reconstructions (e.g. DeepMass** and DeepLensingPosterior***).
GANs are **fast**, use **data-driven priors**, and provide **high-fidelity** samples with **uncertainty quantification**.

*GAN = generative adversarial network **Jeffrey et al. (2020) ***Remy et al. (2023)



Conditional GANs

CONDITIONAL GANS

GANs are models containing 2 networks: a generator and a discriminator which train simultaneously.

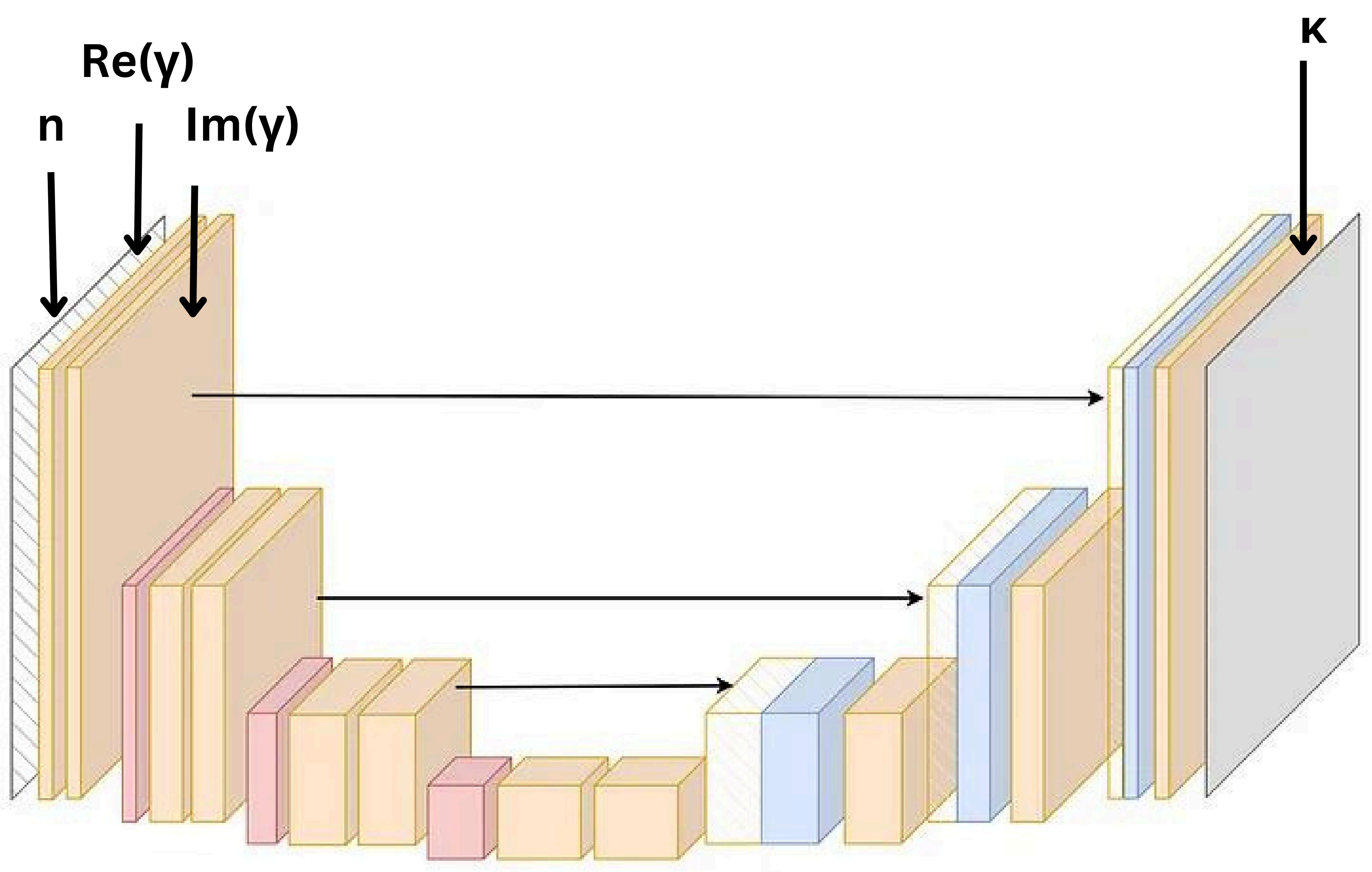
Conditional GANs have an additional ‘conditioning’ label (here the shear)

Challenge:
- Difficult to train
- Mode collapse

Proposed solution:
- Wasserstein GAN* framework
- rcGAN regulariser**

SciAI UCL

*Arjovsky et al. (2017) ** Bendel et al. (2024)



(Source: Conor O'Sullivan, Towards Data Science, March 2023.)

Generator

GOAL

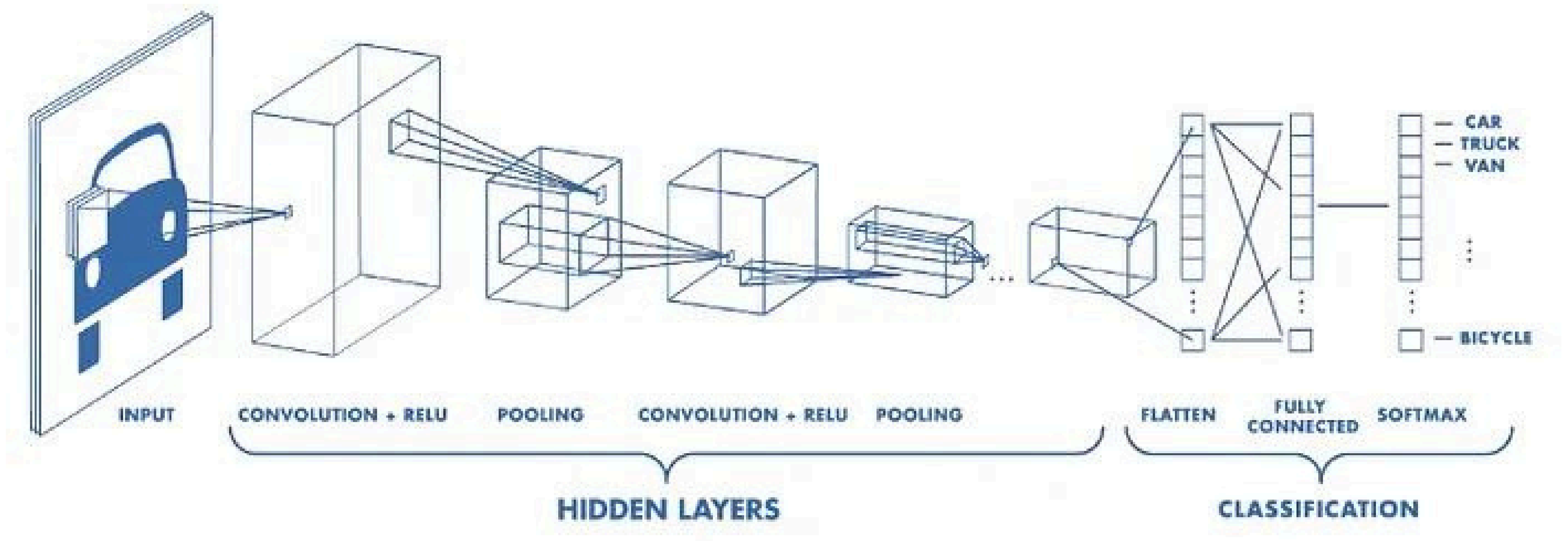
Learn (& match) the posterior distribution of the data as closely as possible.

ARCHITECTURE

UNET: Encoder & Decoder, 4 layers each. 5 residual blocks.

SciAI UCL

Model based on Bendel et al. (2024) architecture.



(Source: UK Mathworks.com)

Discriminator

GOAL

A ‘critic’ to quantify how similar the real and generated data distributions are. This feedback is given to the generator during training.

ARCHITECTURE

Standard CNN.
6 layers + 1 fully connected.
Outputs estimated Wasserstein score.

SciAI UCL

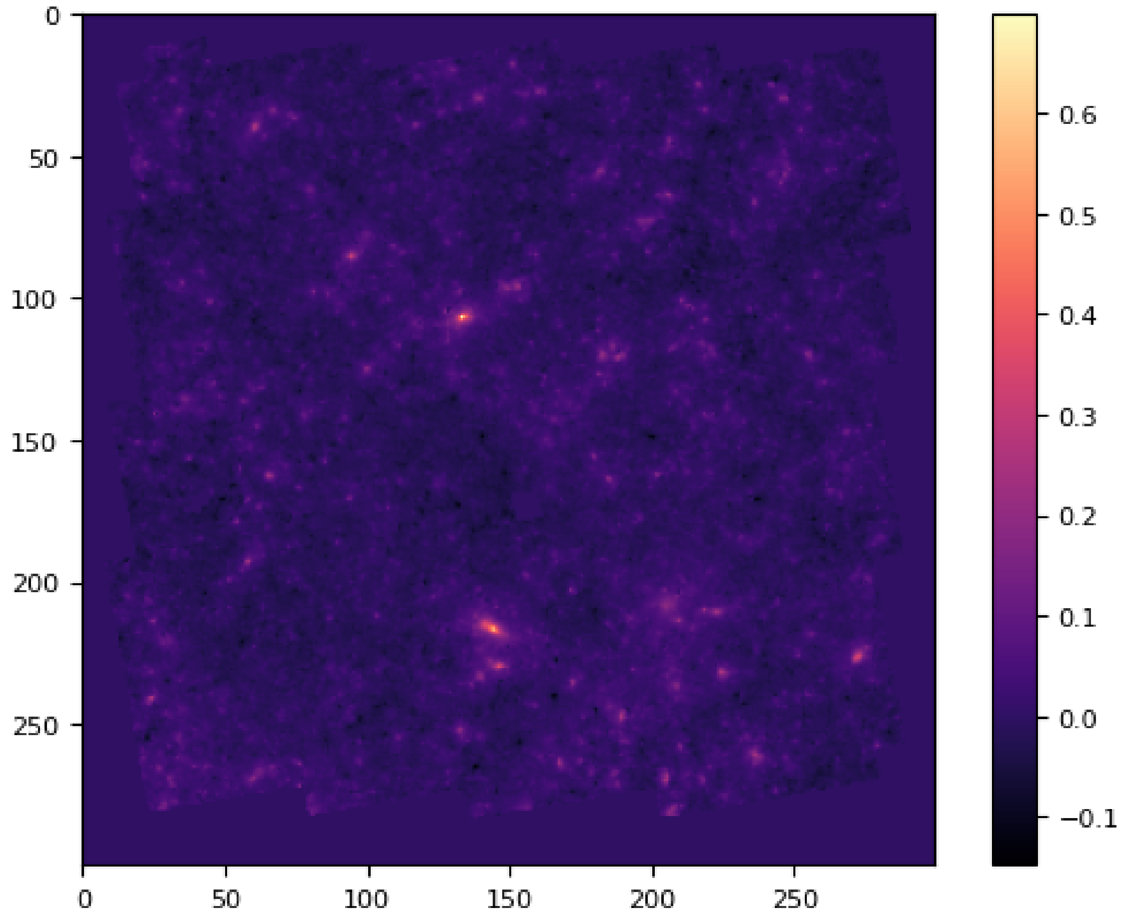
Model based on Bendel et al. (2024) architecture.

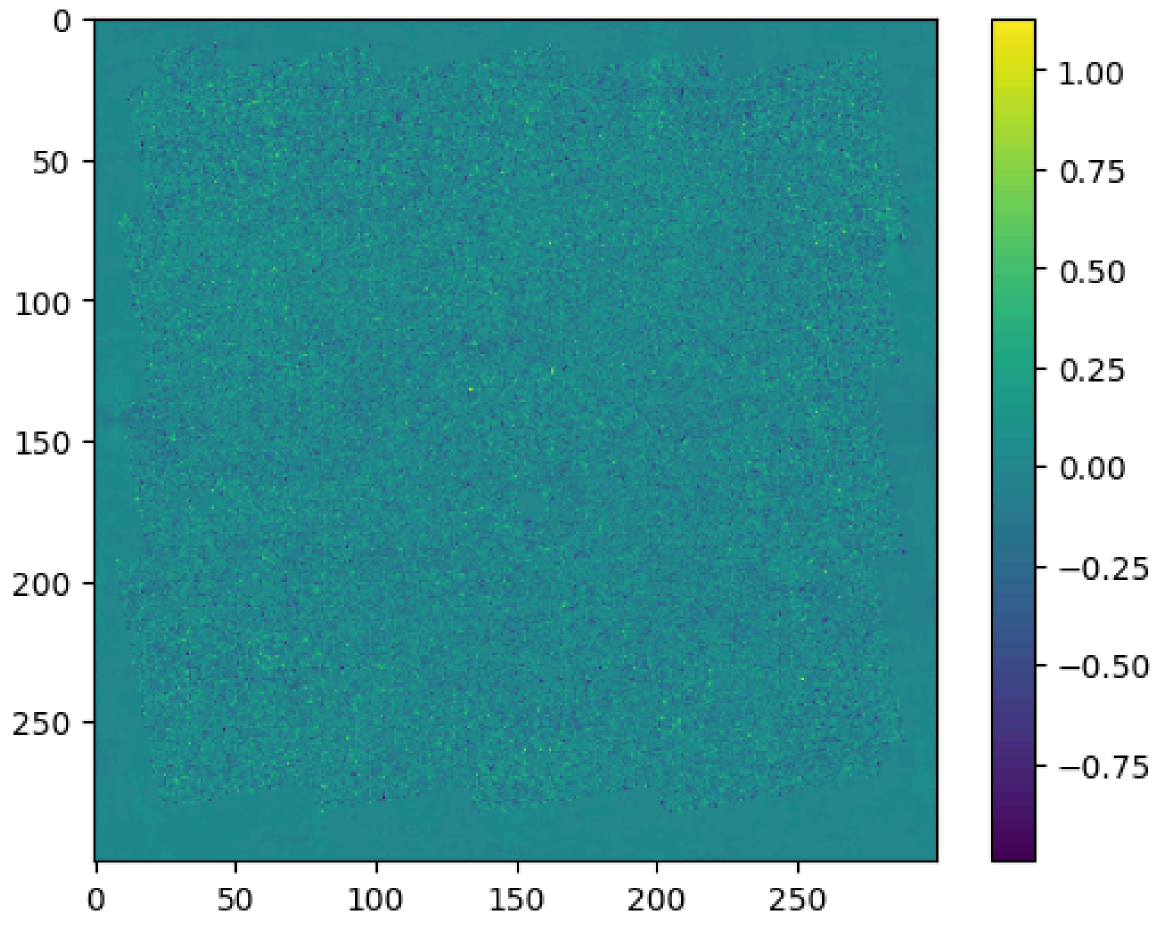
Training Data

κTNG* mock weak lensing suite + COSMOS shape catalog = 10,000 COSMOS-style convergence maps.**

$$\gamma = A\kappa + n$$

- Apply forward model (right) to get mock shear maps

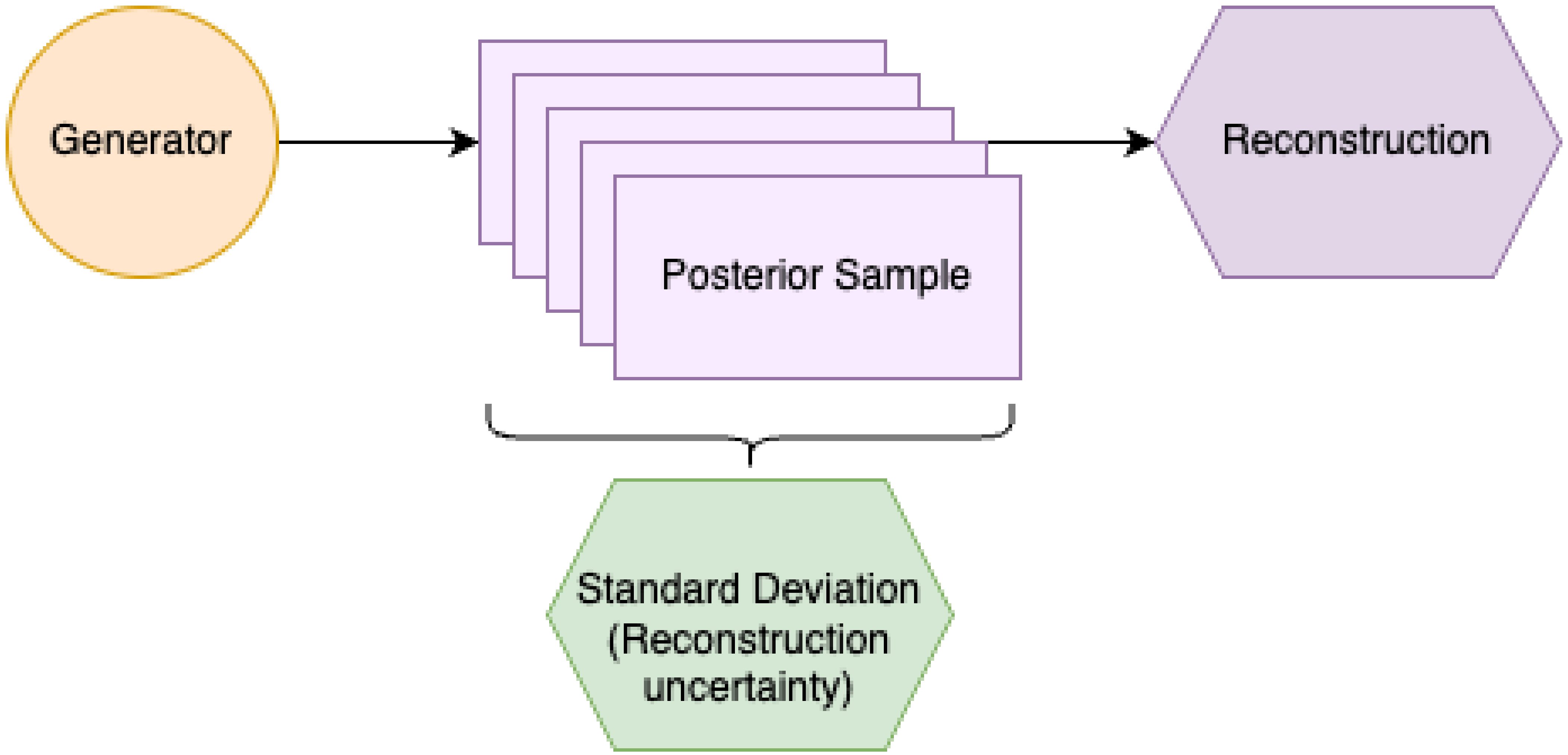




*Osato et al. (2021) http://columbialensing.org/#tng; ** Schrabback et al. (2010)

Reconstructions

- The GAN generates **one** posterior sample every time it is called
- Call the GAN many times to generate many posterior samples
- Average over these posterior samples for a reconstruction
- Standard deviation of these samples is the uncertainty

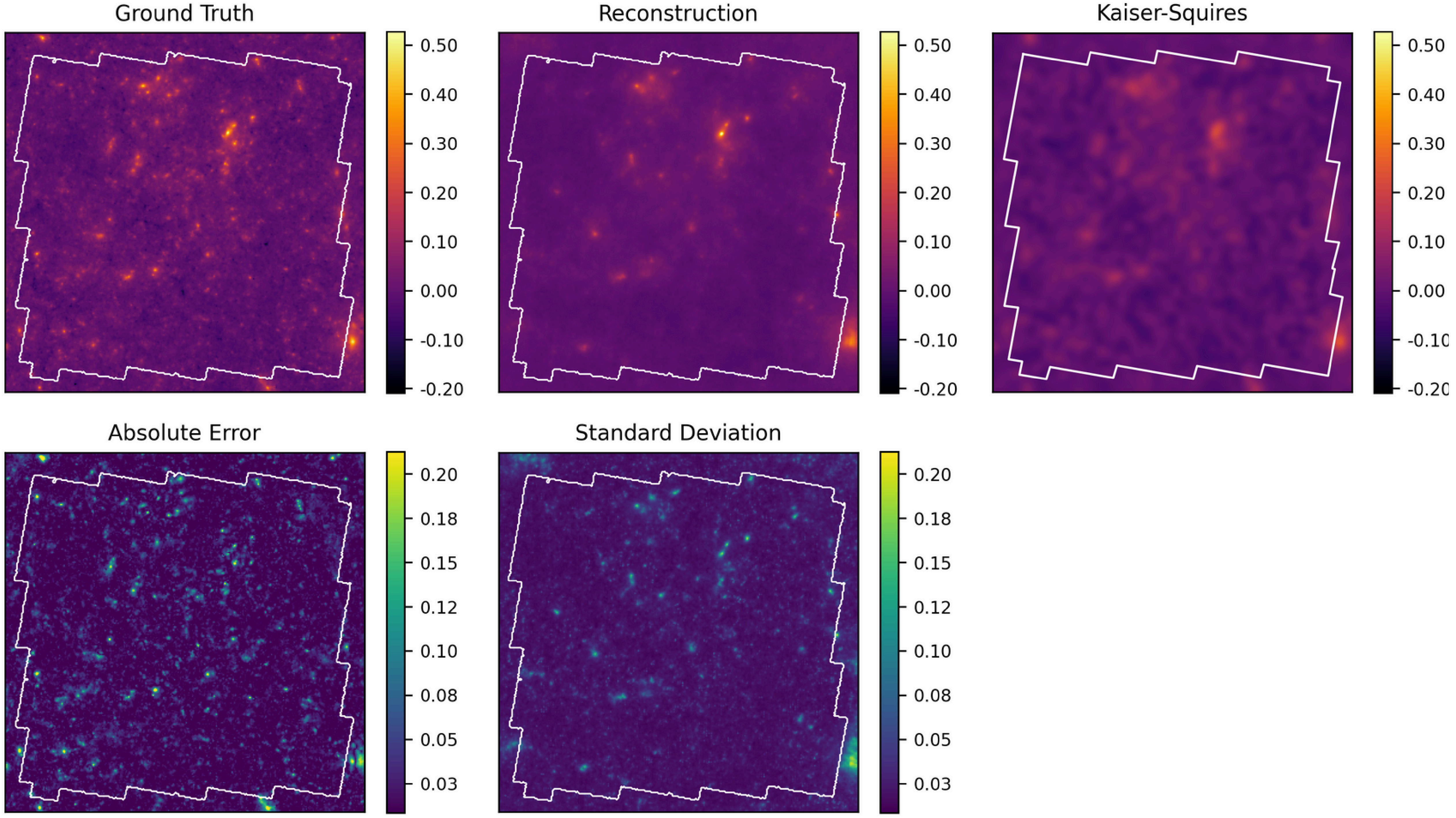


SciAI UCL

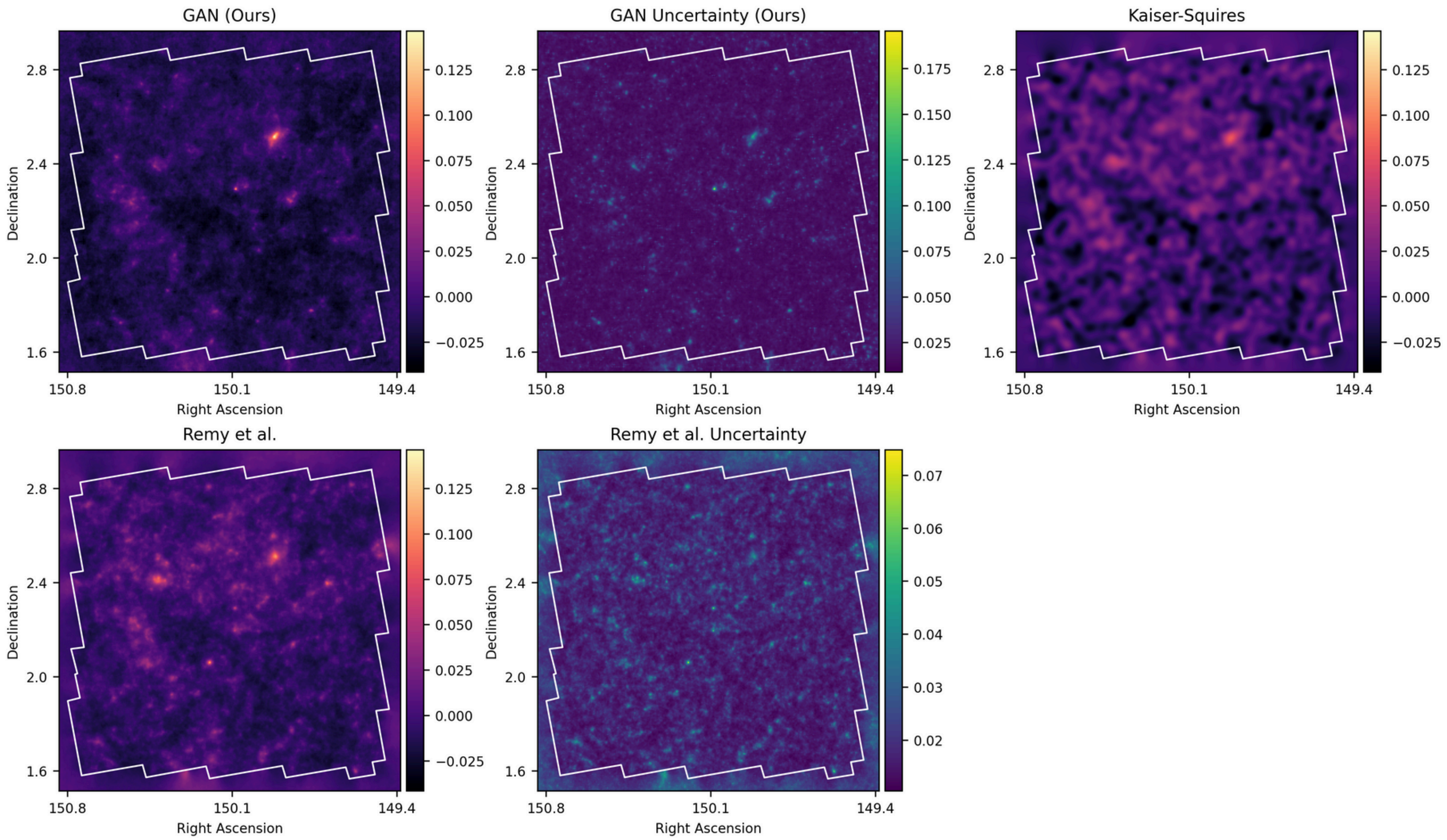
Result - sims

Trained on: 4 A100 GPUs
Time takes: ~ 6 hours

Time to generate sample:
<1 second



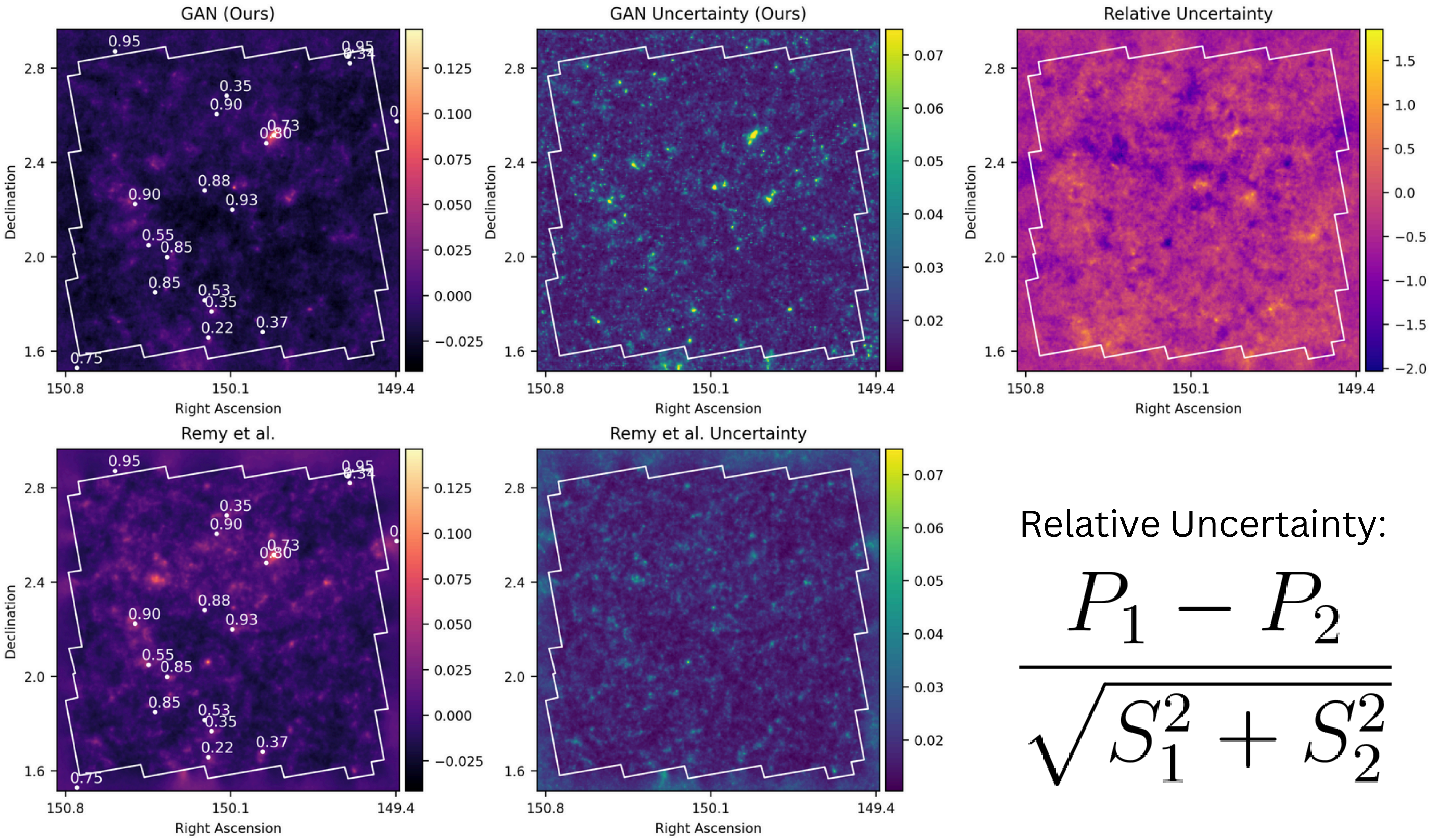
Result - COSMOS



SciAI UCL

DeepPosterior, Remy et al. (2023).

Result - COSMOS



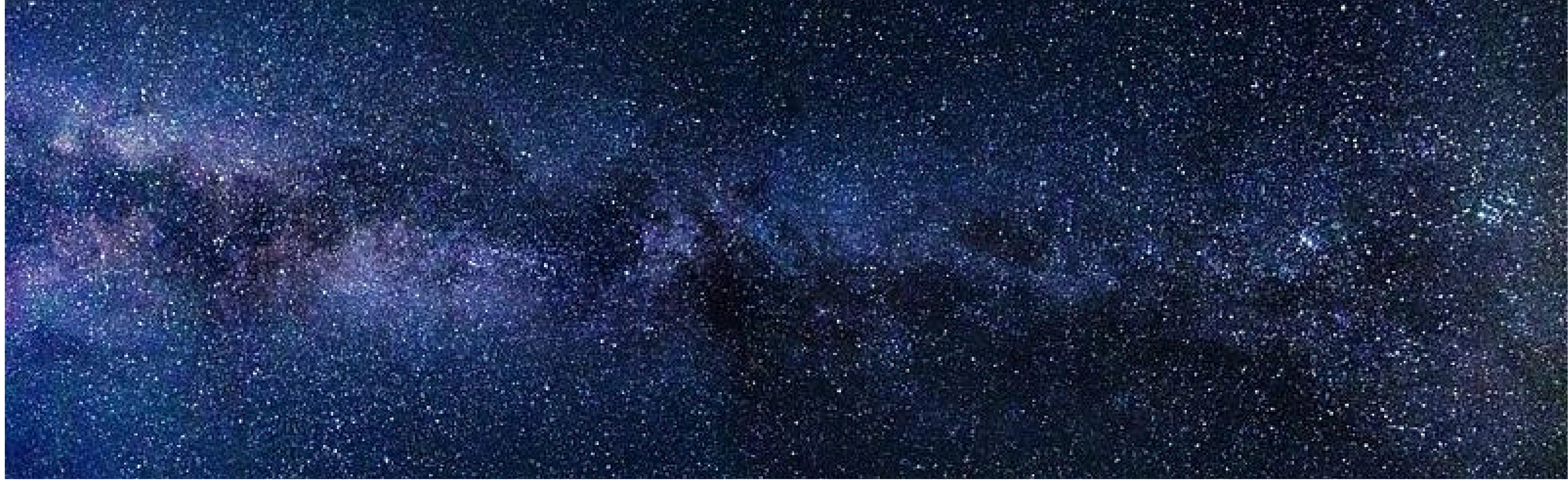
Relative Uncertainty:

$$\frac{P_1 - P_2}{\sqrt{S_1^2 + S_2^2}}$$

X-ray cluster data from Finuguenov et al. (2007) XMM-Newton.

SciAI UCL

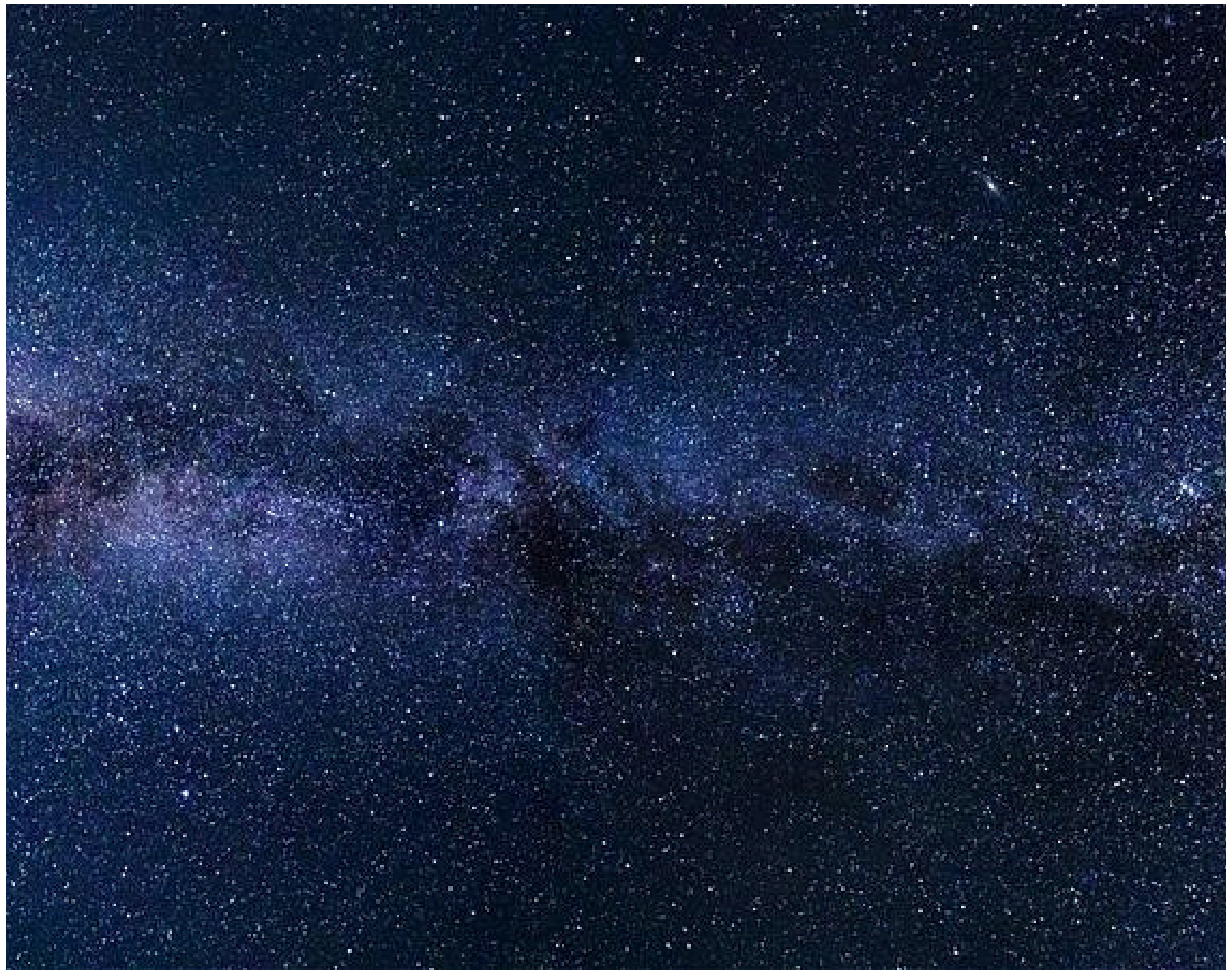
Conclusions



- Developing fast mass mapping techniques with uncertainty quantification is important in preparation for data from Euclid & LSST
- We presented work using conditional generative adversarial networks (cGANs)
- Our method can quickly produce high-quality posterior samples which can be used to make reconstructions and uncertainties
- We validated our results on mock cosmos-style data
- We applied our trained model to the real COSMOS data

SciAI UCL

Thank You For Listening



Presentation by :
JESSICA WHITNEY

SciAI UCL

Regulariser

$$\arg\min_{\theta}\{\beta_{\text{adv}}\mathcal{L}_{\text{adv}}(\theta,\phi) + \mathcal{L}_{1,\text{SD},N_{\text{train}}}(\theta,\beta_{\text{SD}})\}$$

$$\mathcal{L}_{\text{adv}}(\theta,\phi) \triangleq E_{x,z,y}\{D_{\phi}(x,y) - D_{\phi}(G_{\theta}(z,y),y)\},$$

$$\mathcal{L}_{1,\text{SD},N_{\text{train}}}(\theta,\beta_{\text{SD}}) \triangleq \mathcal{L}_{1,N_{\text{train}}}(\theta) - \beta_{\text{SD}}\mathcal{L}_{\text{SD},N_{\text{train}}}(\theta).$$

SciAI UCL

Regulariser

$$\mathcal{L}_{1,\mathrm{SD},N_{\mathrm{train}}}(\theta, \beta_{\mathrm{SD}}) \triangleq \mathcal{L}_{1,N_{\mathrm{train}}}(\theta) - \beta_{\mathrm{SD}} \mathcal{L}_{\mathrm{SD},N_{\mathrm{train}}}(\theta).$$

$$\mathcal{L}_{1,N_{\mathrm{train}}}(\theta) \triangleq E_{x,z_1,\ldots,z_N,y}\{||x - \hat{x}_{(N_{\mathrm{train}})}||_1\}$$

$$\mathcal{L}_{\mathrm{SD},N_{\mathrm{train}}}(\theta) \triangleq \sqrt{\frac{\pi}{2N_{\mathrm{train}}(N_{\mathrm{train}}-1)}} \times \sum_{i=1}^{N_{\mathrm{train}}} E_{z_1,\ldots,z_N,y}\{||\hat{x}_{(i)} - \hat{x}_{(N_{\mathrm{train}})}||_1\}$$

SciAI UCL

Wasserstein GANs

In standard GANs the discriminator usually tries to differentiate real and fake **samples** in a ‘real or fake’ manner.
WGANs try to make this feedback more informative by instead calculating the distance between the generator’s distribution and the true data distribution.

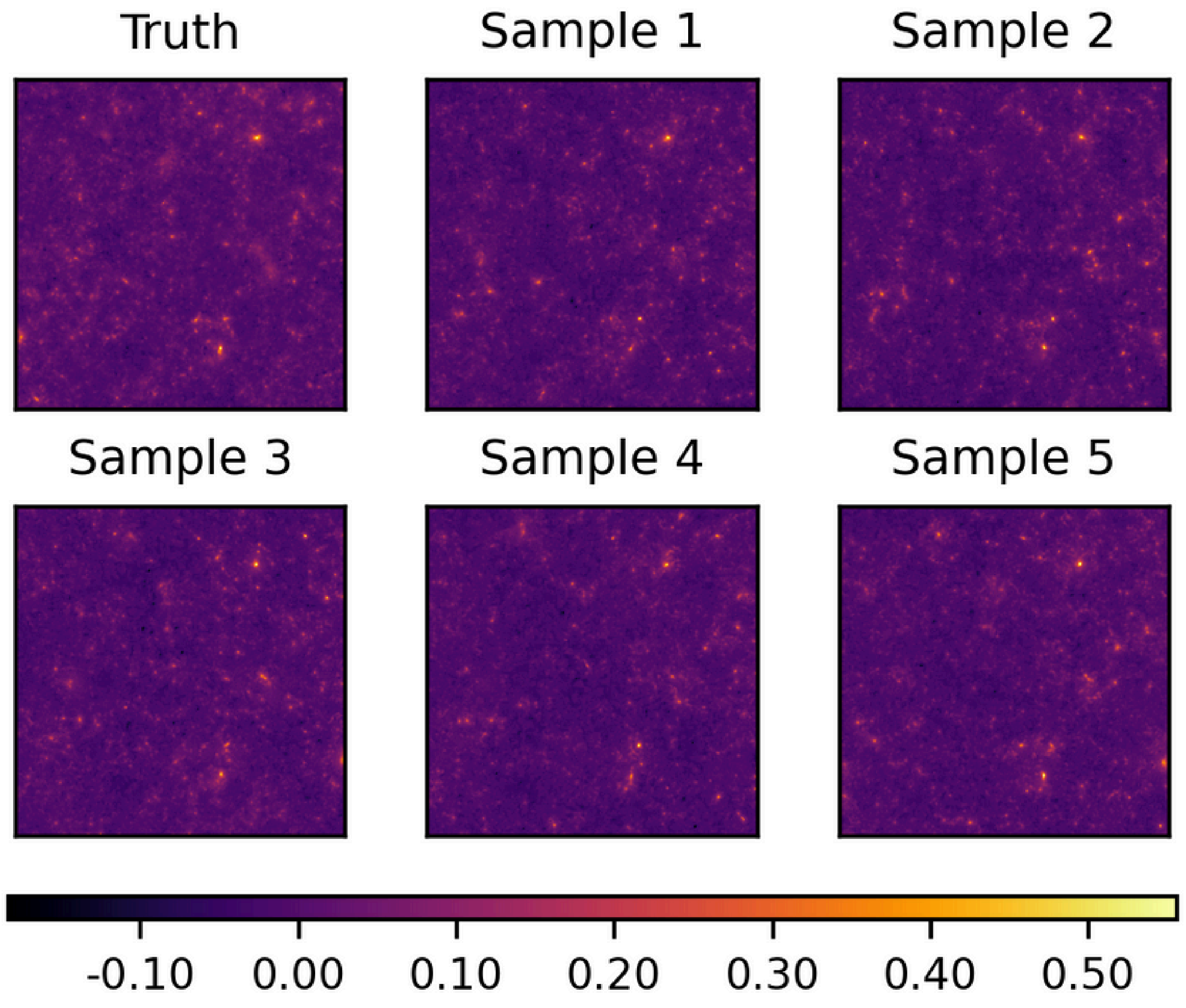
Wasserstein-1 distance (also known as Earth Mover’s distance)

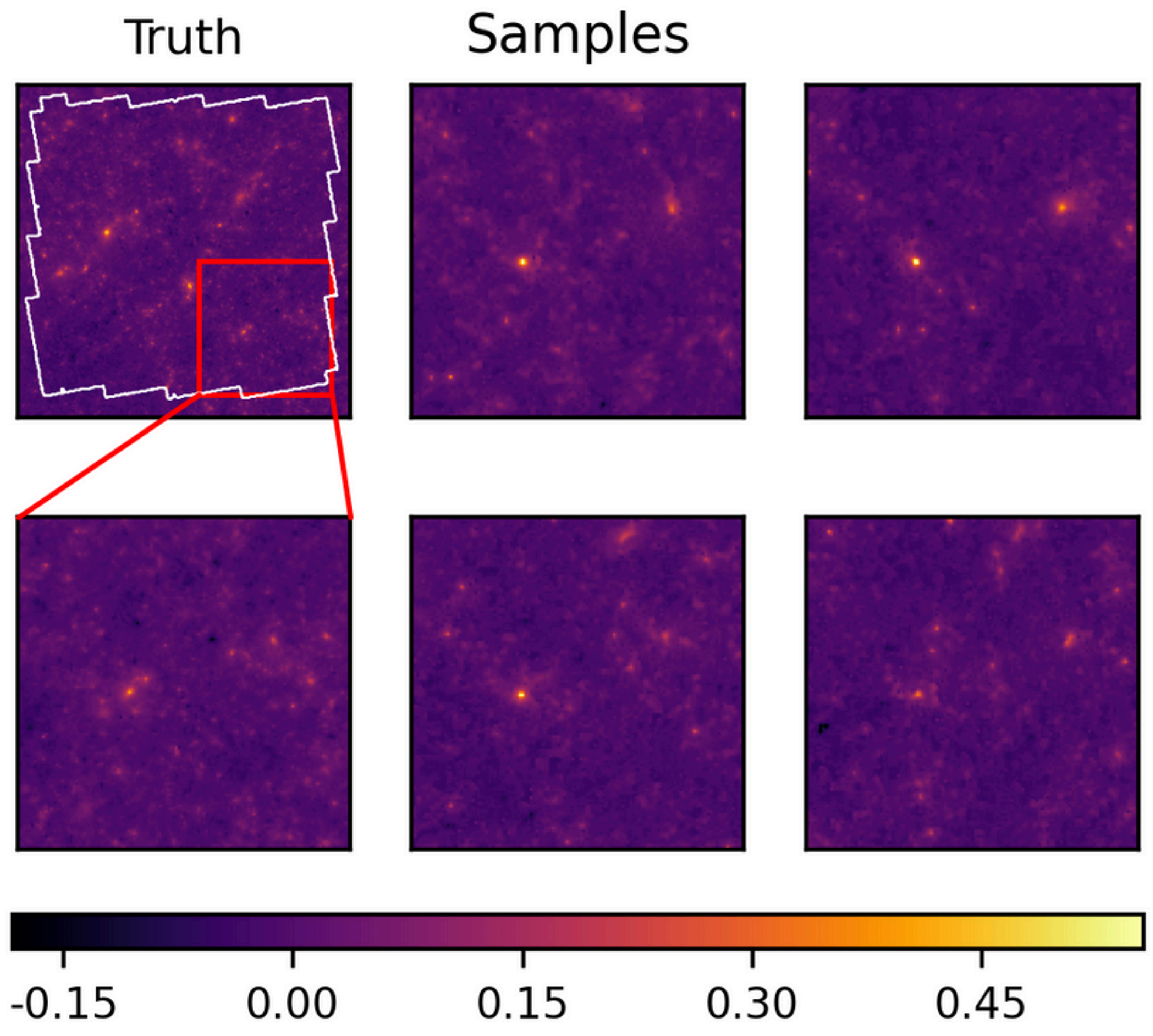
$$W_1(p_{x|y}(\cdot, y), p_{\hat{x}|y}(\cdot, y)) = \sup_{D \in L_1} E_{x|y}\{D(x, y)\} - E_{\hat{x}|y}\{D(\hat{x}, y)\}$$

SciAI UCL

Sims; Variability

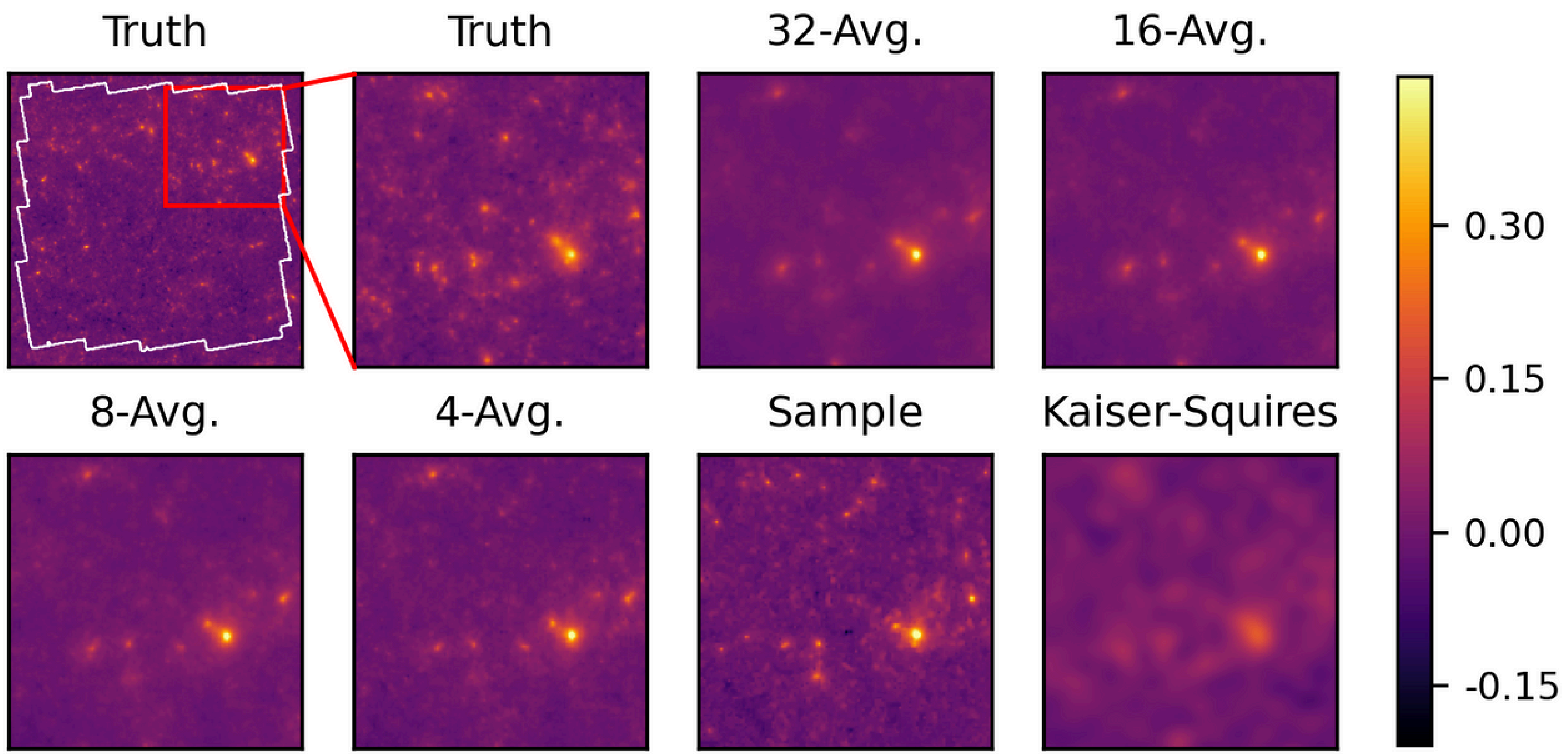
BACKUP SLIDES

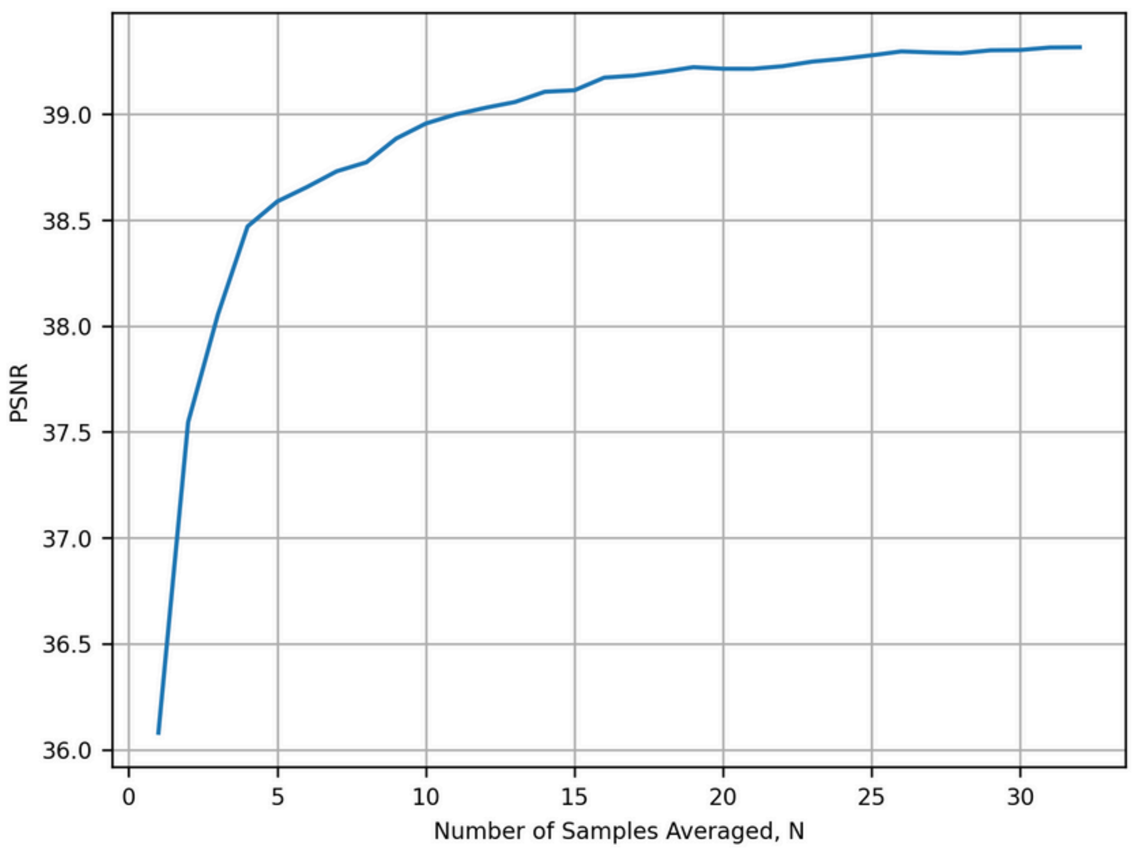




SciAI UCL

Sims; Changing N





SciAI UCL