Solving inverse problems with diffusion models

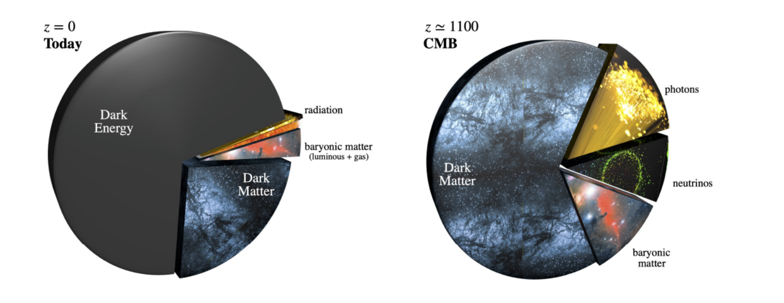
Alexandre Adam

πAIe @ Pisa
Sept. 25th 2024

Université de Montréal et du monde.
Ciela Institute
Mila

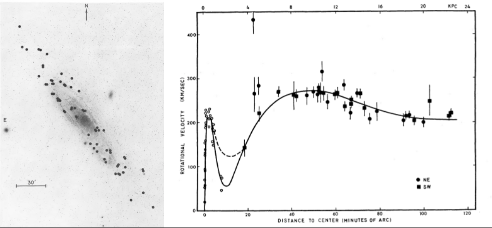
Dark Matter

ΛCDM

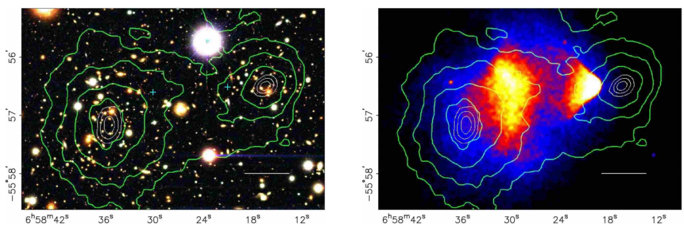


Credit: Cirelli et al. (2024), arxiv:2406.01705

Dark Matter

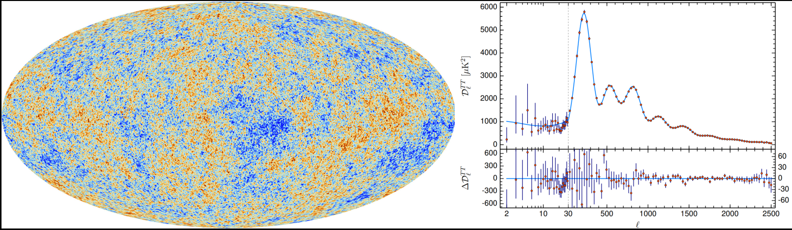


Credit: Rubin et al. (1970), ApJ, vol. 159, p. 279

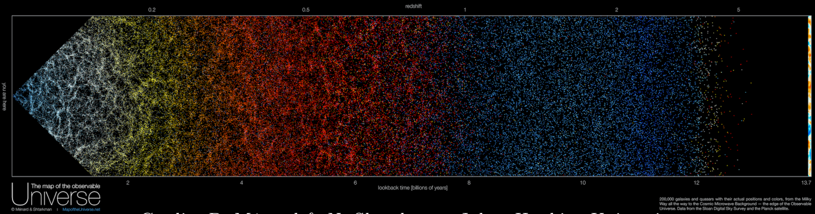


Credit: Clowe et al. (2006), ApJ, vol. 648, p. L109

Dark Matter



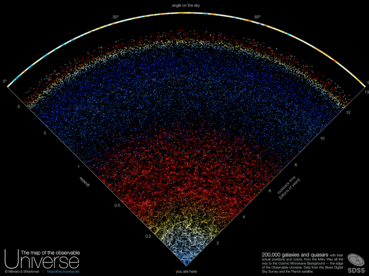
Credit: Planck Collaboration (2018)



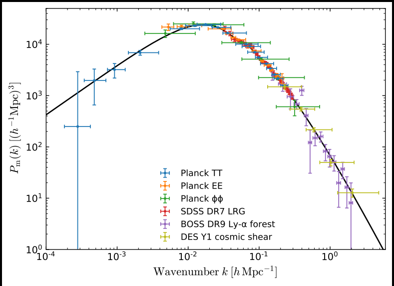
Credit: B. Ménard & N. Shtarkman, Johns Hopkins University

Dark Matter

Large scale observations



Credit: B. Ménard & N. Shtarkman
Johns Hopkins University

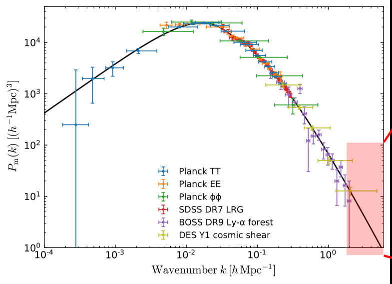


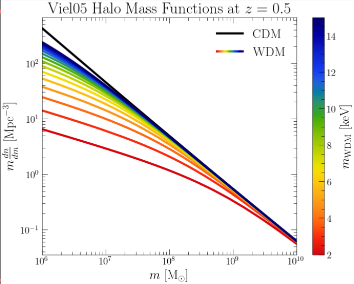
Credit: ESA and the Planck Collaboration

Dark Matter

Small scales?

$P_m(k) \propto k$ $\qquad$ $P_m(k) \propto k^{-3}$



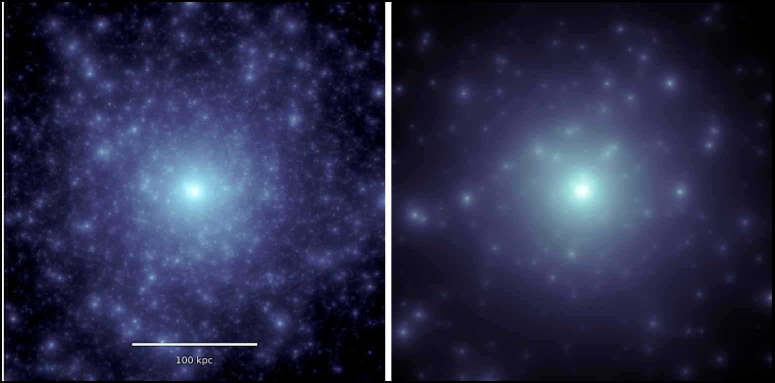


Credit: ESA and the Planck Collaboration

Dark Matter

Substructure

Cold Dark Matter Warm Dark Matter

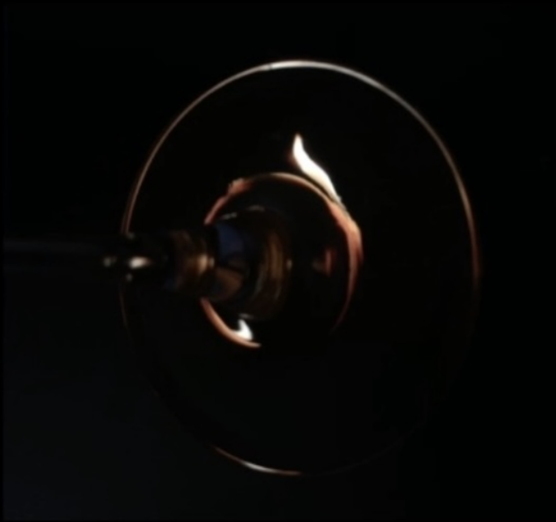


Credit: D. Gilman (2020)

Gravitational Lensing

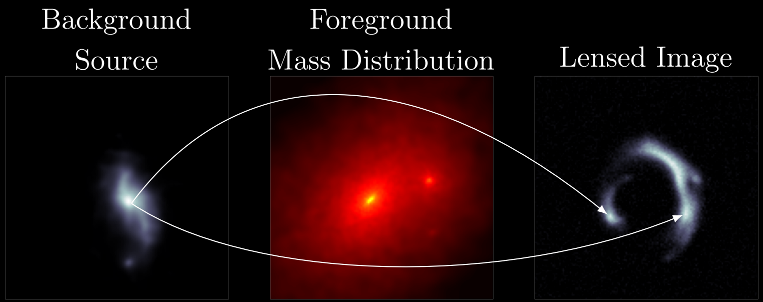
Gravitational Lensing

Gravitational Lensing

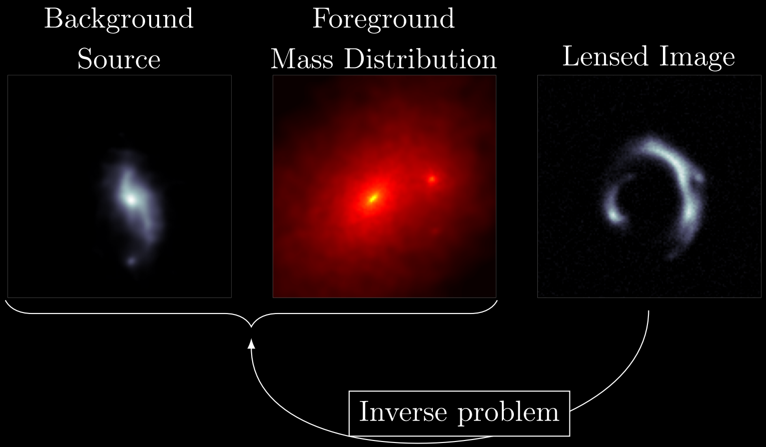


Gravitational Lensing

Gravitational Lensing

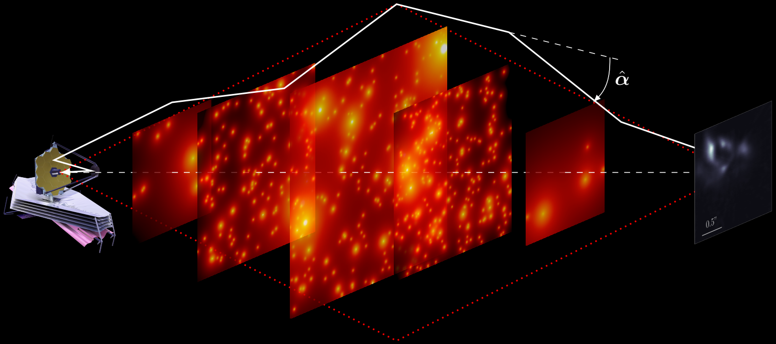


Gravitational Lensing

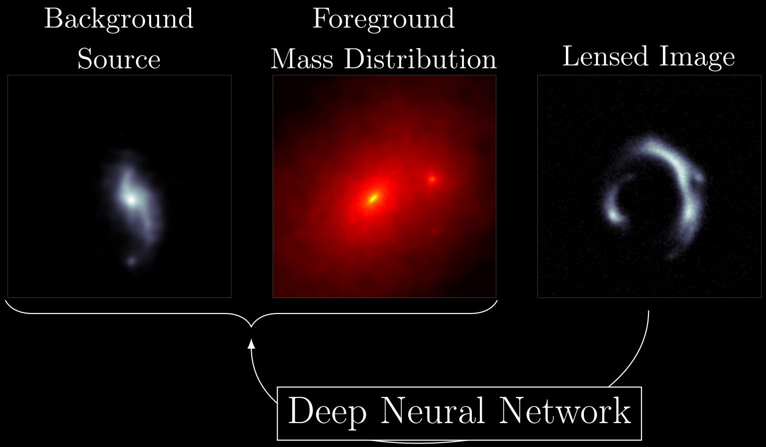


Gravitational Lensing

Dark matter in the line of sight



Can we use machine learning?



Can we use machine learning?


THE ASTROPHYSICAL JOURNAL

OPEN ACCESS

Pixelated Reconstruction of Foreground Density and Background Surface Brightness in Gravitational Lensing Systems Using Recurrent Inference Machines

Alexandre Adam[1,2,3], Laurence Perreault-Levasseur[1,2,3,4], Yashar Hezaveh[1,2,3,4], and Max Welling[5]

Published 2023 June 27 •

The Astrophysical Journal, Volume 951, Number 1

Citation Alexandre Adam et al 2023 ApJ 951 6

357/accf84

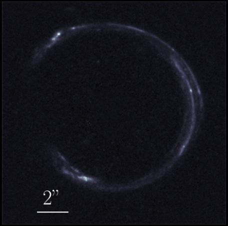

Laurence Perreault-Levasseur


Yashar Hezaveh


Max Welling

Lensed Image

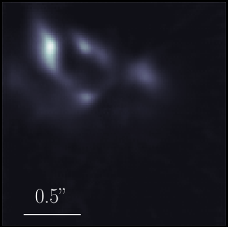
Observation

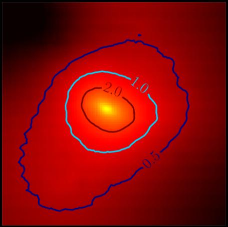


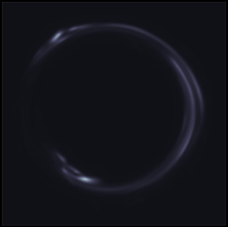
Background

Foreground

RIM Prediction

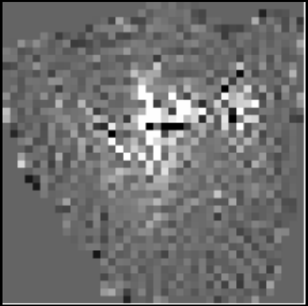






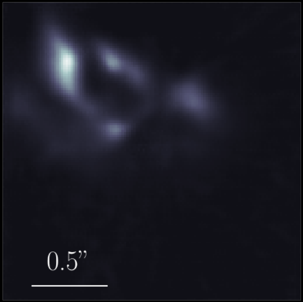
Can we use machine learning?

Traditional method



Schuldt et al. (2019)

ML method



Adam et al. (2023)

Can we use machine learning?

Yes and no

- Not precise enough to detect small substructures
- No notion of uncertainty

Can we use machine learning?

Yes and no

- Not precise enough to detect small substructures
- No notion of uncertainty

Thomas Bayes (1763)

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Score-Based Models

$$\nabla \log p(\mathbf{x})$$

Score-Based Models

The score is local

$$\boldsymbol{\nabla} \log p(\mathbf{x})$$

Gibbs Measure $\qquad p(\mathbf{x}) = \frac{1}{Z} e^{-\beta E(\mathbf{x})}$

Score-Based Models

The score is local

$$\boldsymbol{\nabla} \log p(\mathbf{x})$$

Gibbs Measure $\qquad p(\mathbf{x}) = \frac{1}{\mathcal{Z}} e^{-\beta E(\mathbf{x})}$

$$\log p(\mathbf{x}) = -\beta E(\mathbf{x}) - \log \mathcal{Z}$$

Score-Based Models

The score is local

$$\nabla \log p(\mathbf{x})$$

Gibbs Measure $\quad p(\mathbf{x}) = \frac{1}{\mathcal{Z}} e^{-\beta E(\mathbf{x})}$

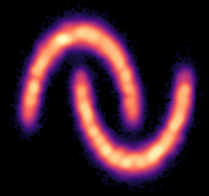
$$\log p(\mathbf{x}) = -\beta E(\mathbf{x}) - \log \mathcal{Z}$$

$$\boxed{\nabla \log p(\mathbf{x}) = -\beta \nabla E(\mathbf{x})}$$

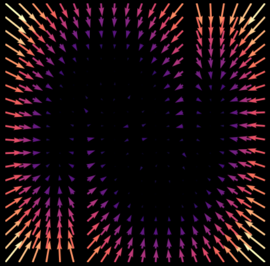
Local and easier to learn

Score-Based Models

$p(\mathbf{x})$

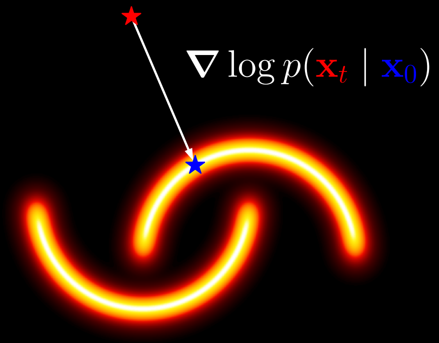


$\nabla \log p(\mathbf{x})$



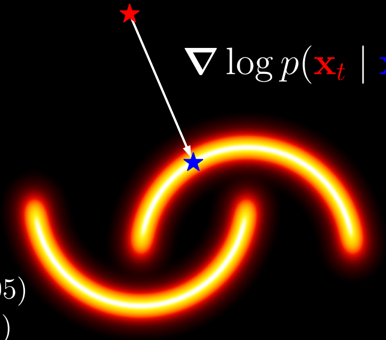
Score-Based Models

Score Matching



Score-Based Models

Score Matching



$$\nabla \log p(\mathbf{x}_t \mid \mathbf{x}_0)$$

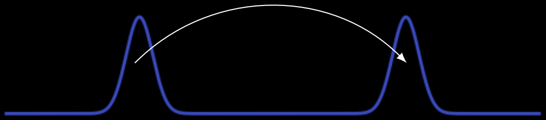
Hyvarinen ('05)
Vincent ('11)
Song+ ('19,'20,'21)

$$\mathcal{L}_\theta \sim \int dt\, \mathbb{E} \|\mathbf{s}_\theta(t, \mathbf{x}_t) - \nabla \log p(\mathbf{x}_t \mid \mathbf{x}_0)\|^2$$

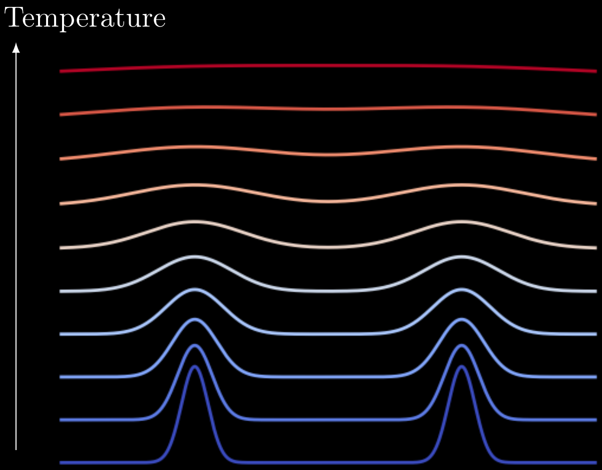
Diffusion Models

Diffusion Models

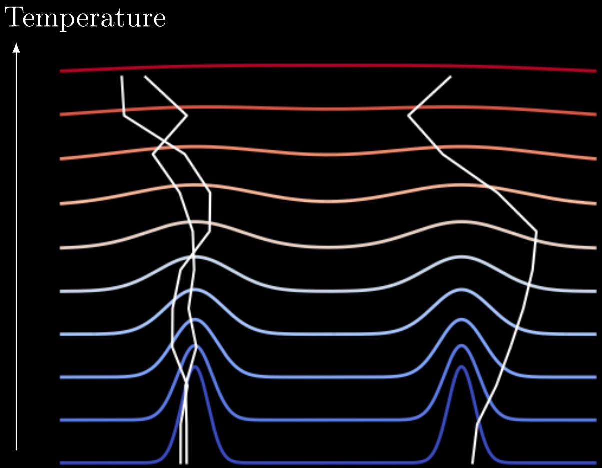
MCMC proprosals often
struggle to jump between modes



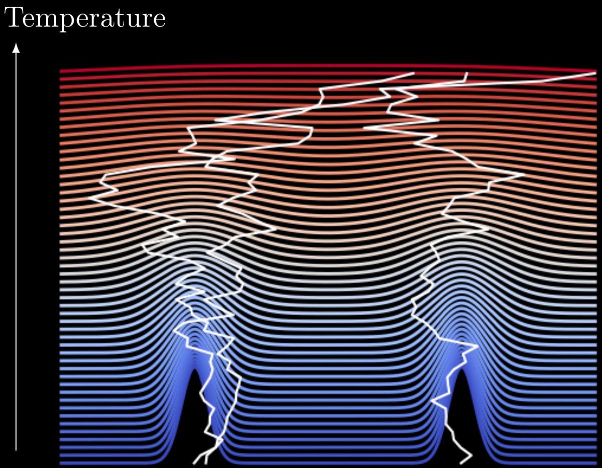
Diffusion Models



Diffusion Models

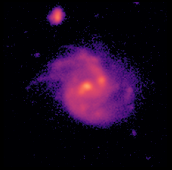


Diffusion Models

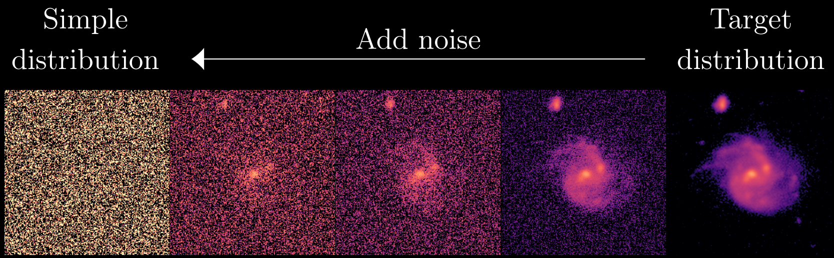


Diffusion Models

Target
distribution



Diffusion Models

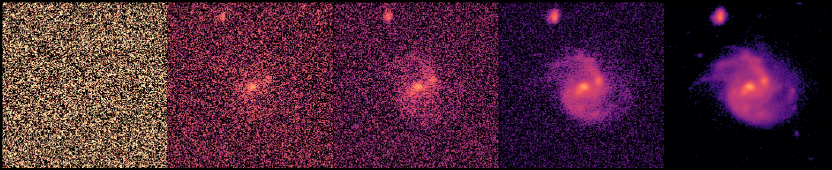


Diffusion Models

Simple distribution

Learn $\boldsymbol{\nabla}_{\mathbf{x}} \log p_t(\mathbf{x})$

Target distribution

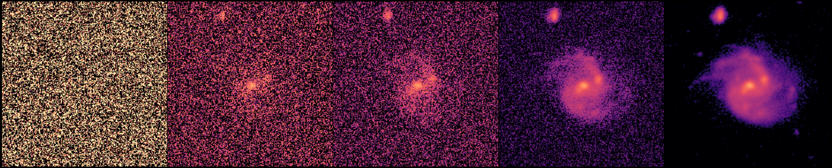


Diffusion Models

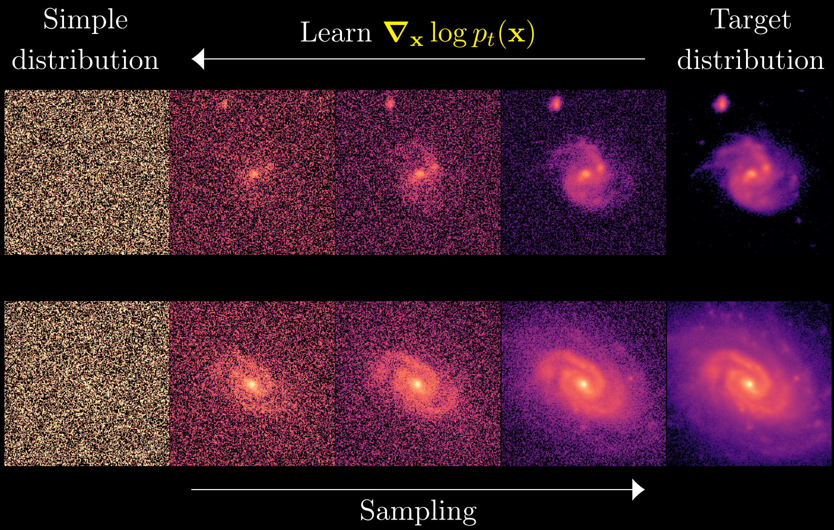
Simple distribution

Learn $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$

Target distribution

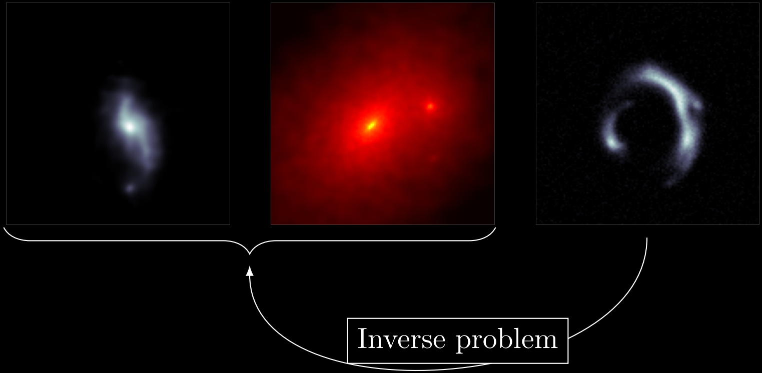


Diffusion Models



Solving inverse problems

Background Source | Foreground Mass Distribution | Lensed Image

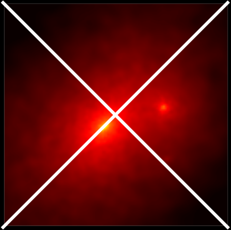


$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

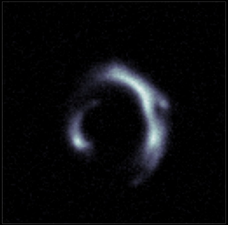
Solving inverse problems

Background Source

Foreground Mass Distribution



Lensed Image



$\boxed{\text{Inverse problem}}$

Posterior $\propto$ Likelihood $\times$ Prior

Solving inverse problems

$$\overbrace{p(\mathbf{x} \mid \mathbf{y})}^{\text{posterior}} = \frac{\overbrace{p(\mathbf{y} \mid \mathbf{x})}^{\text{likelihood}} \overbrace{p(\mathbf{x})}^{\text{prior}}}{p(\mathbf{y})}$$

Solving inverse problems

$$\overbrace{p(\mathbf{x} \mid \mathbf{y})}^{\text{posterior}} = \frac{\overbrace{p(\mathbf{y} \mid \mathbf{x})}^{\text{likelihood}} \overbrace{p(\mathbf{x})}^{\text{prior}}}{p(\mathbf{y})}$$

$$\underbrace{\log p(\mathbf{x} \mid \mathbf{y})}_{\text{posterior}} = \underbrace{\log p(\mathbf{y} \mid \mathbf{x})}_{\text{likelihood}} + \underbrace{\log p(\mathbf{x})}_{\text{prior}} - \log p(\mathbf{y})$$

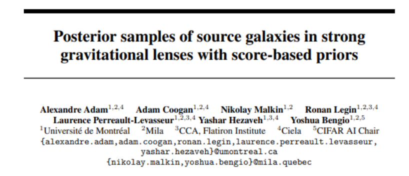
Solving inverse problems

$$\overbrace{p(\mathbf{x} \mid \mathbf{y})}^{\text{posterior}} = \frac{\overbrace{p(\mathbf{y} \mid \mathbf{x})}^{\text{likelihood}}\,\overbrace{p(\mathbf{x})}^{\text{prior}}}{p(\mathbf{y})}$$

$$\underbrace{\log p(\mathbf{x} \mid \mathbf{y})}_{\text{posterior}} = \underbrace{\log p(\mathbf{y} \mid \mathbf{x})}_{\text{likelihood}} + \underbrace{\log p(\mathbf{x})}_{\text{prior}} - \log p(\mathbf{y})$$

$$\underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{x} \mid \mathbf{y})}_{\text{posterior}} = \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{y} \mid \mathbf{x})}_{\text{likelihood}} + \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{x})}_{\text{prior}} - \cancel{\nabla_{\mathbf{x}} \log p(\mathbf{y})}$$

Solving inverse problems



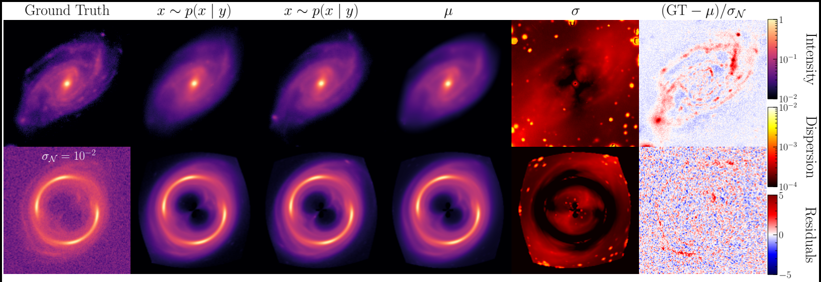
Posterior samples of source galaxies in strong gravitational lenses with score-based priors

Alexandre Adam[1,2,4] **Adam Coogan**[1,2,4] **Nikolay Malkin**[1,2] **Ronan Legin**[1,2,3,4]
Laurence Perreault-Levasseur[1,2,3,4] **Yashar Hezaveh**[1,3,4] **Yoshua Bengio**[1,2,5]
[1]Université de Montréal [2]Mila [3]CCA, Flatiron Institute [4]Ciela [5]CIFAR AI Chair
{alexandre.adam,adam.coogan,ronan.legin,laurence.perreault.levasseur,
yashar.hezaveh}@umontreal.ca
{nikolay.malkin,yoshua.bengio}@mila.quebec

Adam Coogan

Kolya Malkin

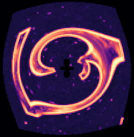
Yoshua Bengio



Solving inverse problems

Misspecified prior

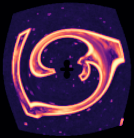
Ground
Truth

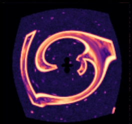


Solving inverse problems

Misspecified prior

Ground
Truth

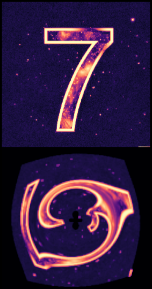


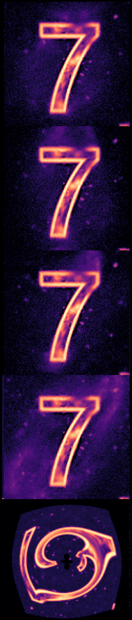


Solving inverse problems

Misspecified prior

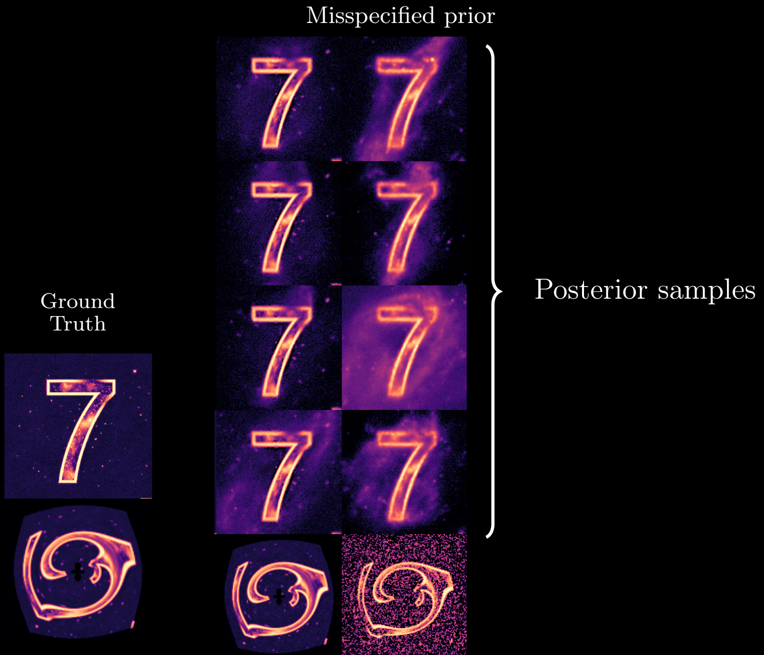
Ground Truth



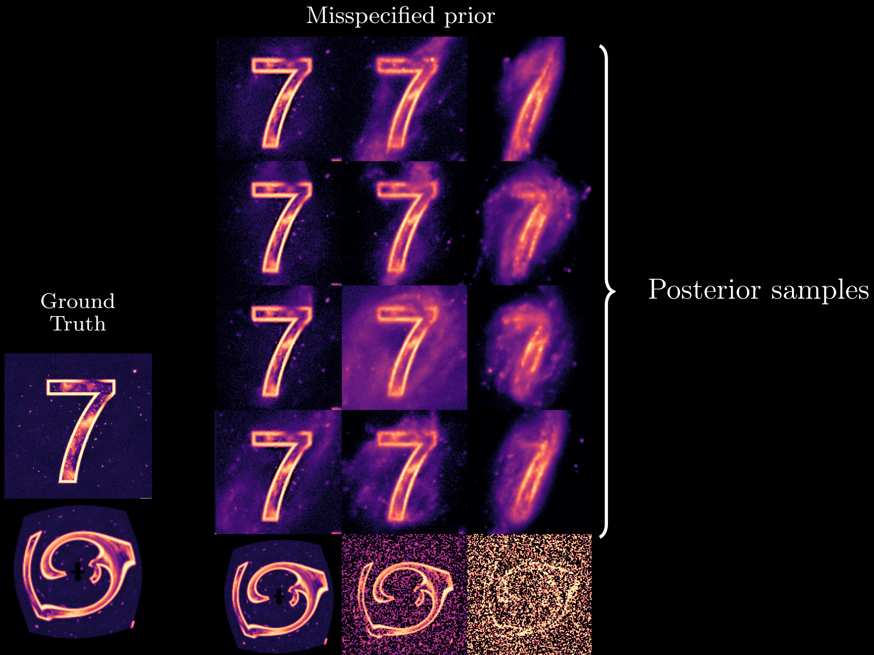


Posterior samples

Solving inverse problems

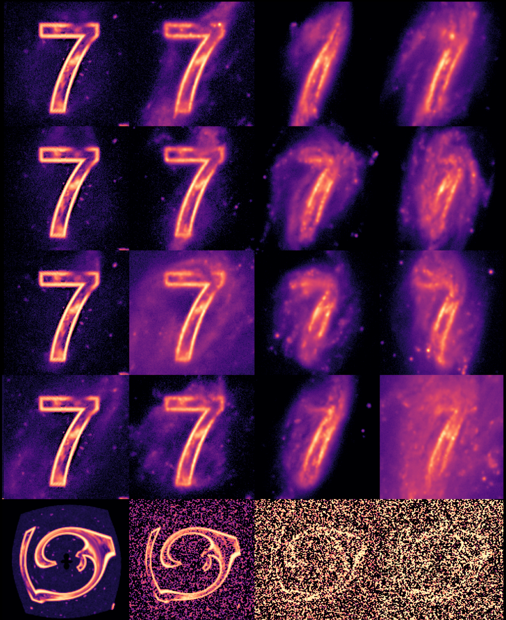


Solving inverse problems

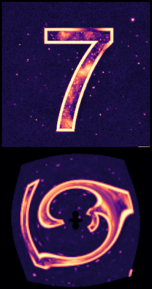


Solving inverse problems

Misspecified prior

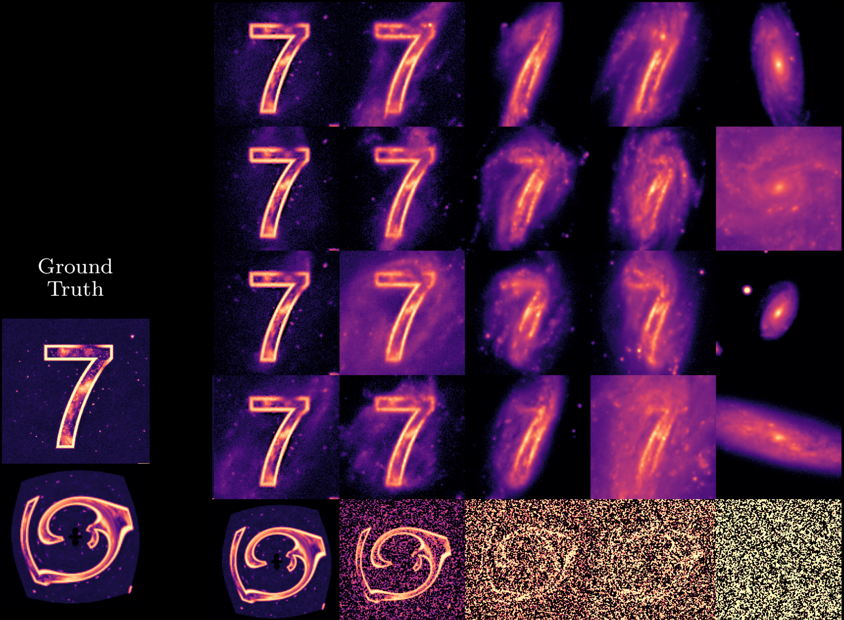


Ground Truth



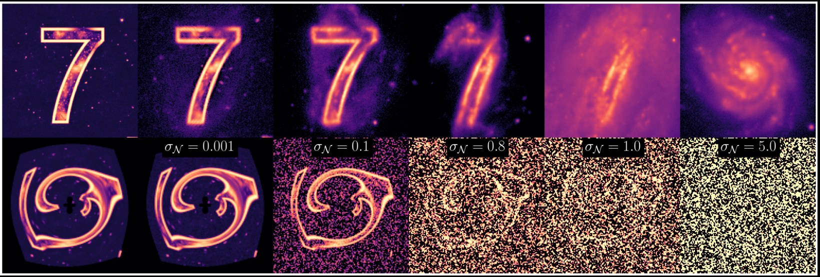
Solving inverse problems

Misspecified prior



Mispecified prior

Ground truth



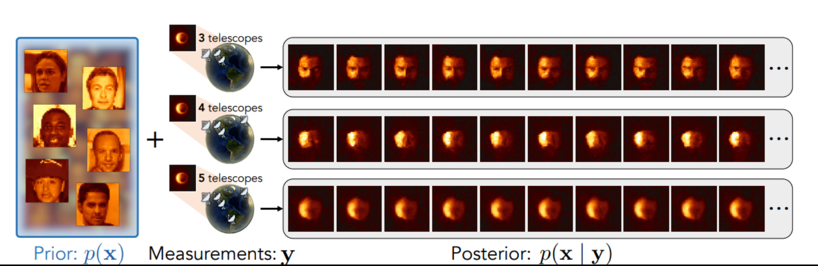
Likelihood less informative

Mispecified prior

Interferometric imaging

Score-Based Diffusion Models as Principled Priors for Inverse Imaging

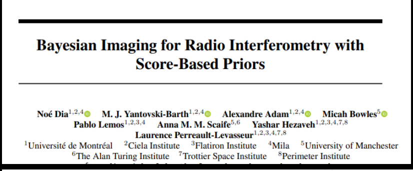
Berthy T. Feng[1*] Jamie Smith[2] Michael Rubinstein[2] Huiwen Chang[2]
Katherine L. Bouman[1] William T. Freeman[2]
[1]California Institute of Technology [2]Google Research



Feng et al., ICCV 2023, arxiv:2304.11751

Mispecified prior

Interferometric imaging



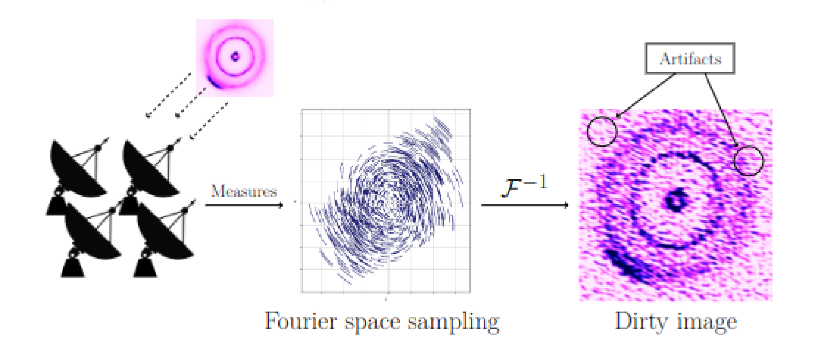
Bayesian Imaging for Radio Interferometry with Score-Based Priors

Noé Dia[1,2,4] M. J. Yantovski-Barth[1,2,4] Alexandre Adam[1,2,4] Micah Bowles[5]
Pablo Lemos[1,2,3,4] Anna M. M. Scaife[5,6] Yashar Hezaveh[1,2,3,4,7,8]
Laurence Perreault-Levasseur[1,2,3,4,7,8]
[1]Université de Montréal [2]Ciela Institute [3]Flatiron Institute [4]Mila [5]University of Manchester
[6]The Alan Turing Institute [7]Trottier Space Institute [8]Perimeter Institute

Noé Dia

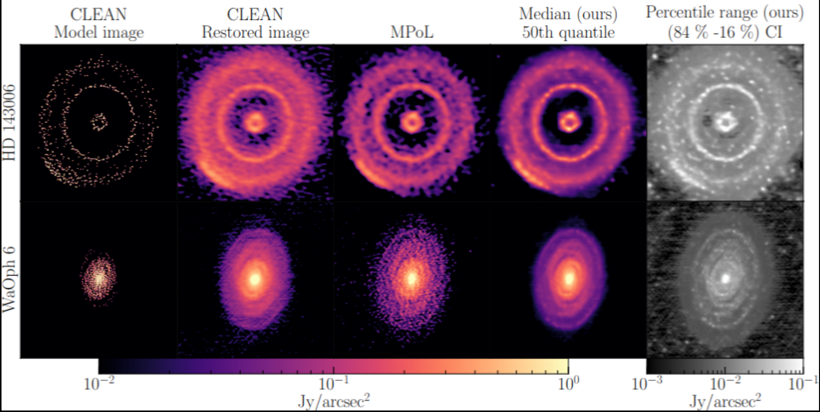
Michael J. Barth

Micah Bowles



Mispecified prior

Interferometric imaging



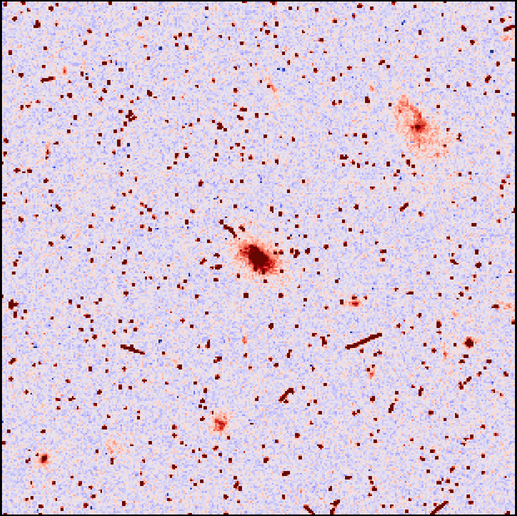
Noé Dia

Bayesian Inference

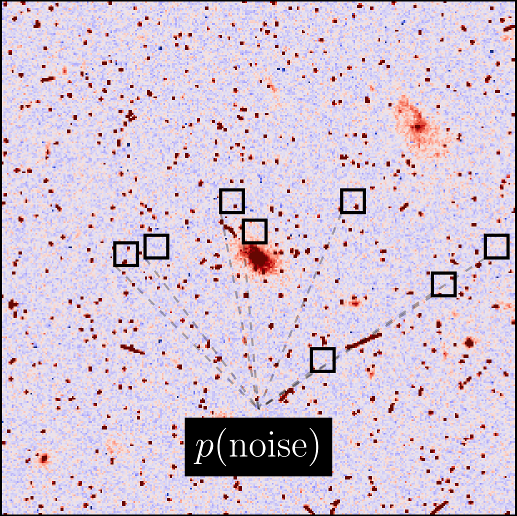
Thomas Bayes (1763)

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

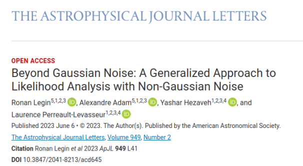
Learning the likelihood



Learning the likelihood



Learning the likelihood


THE ASTROPHYSICAL JOURNAL LETTERS

OPEN ACCESS

Beyond Gaussian Noise: A Generalized Approach to Likelihood Analysis with Non-Gaussian Noise

Ronan Legin[5,1,2,3], Alexandre Adam[5,1,2,3], Yashar Hezaveh[1,2,3,4], and Laurence Perreault-Levasseur[1,2,3,4]

Published 2023 June 6 • © 2023. The Author(s). Published by the American Astronomical Society.

The Astrophysical Journal Letters, Volume 949, Number 2

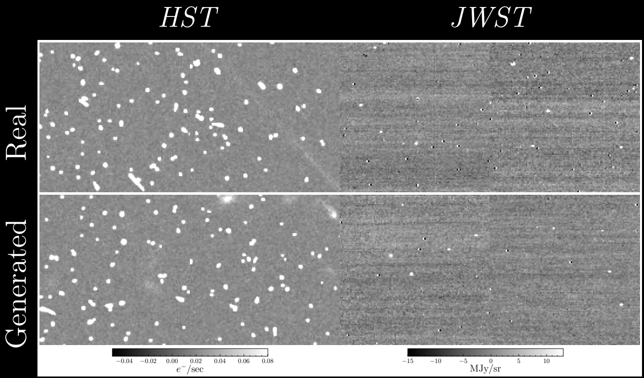
Citation Ronan Legin *et al* 2023 *ApJL* **949** L41

DOI 10.3847/2041-8213/acd645

Ronan Legin

Laurence Perreault-Levasseur

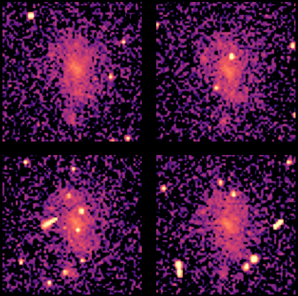
Yashar Hezaveh

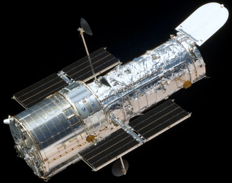


Sampling from the posterior

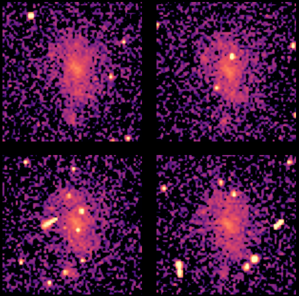
$$\underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{x} \mid \mathbf{y})}_{\text{posterior}} = \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{y} \mid \mathbf{x})}_{\text{likelihood}} + \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{x})}_{\text{prior}}$$

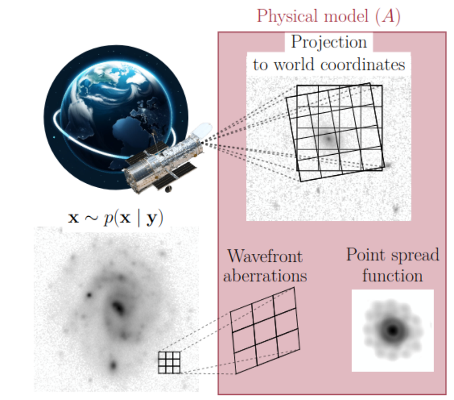
"Raw" HST images

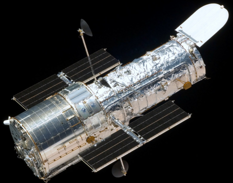




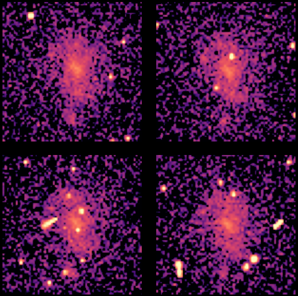
"Raw" HST images



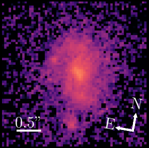


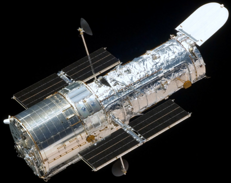


"Raw" HST images

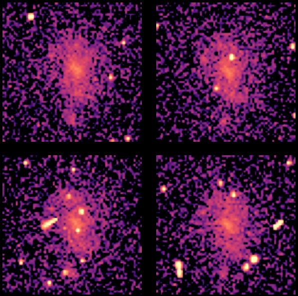


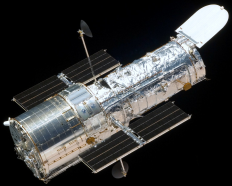
Traditional method



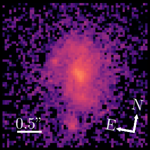


"Raw" HST images

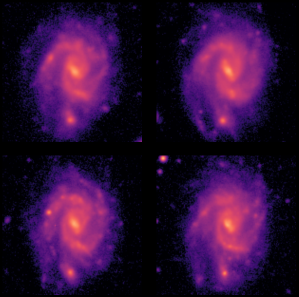


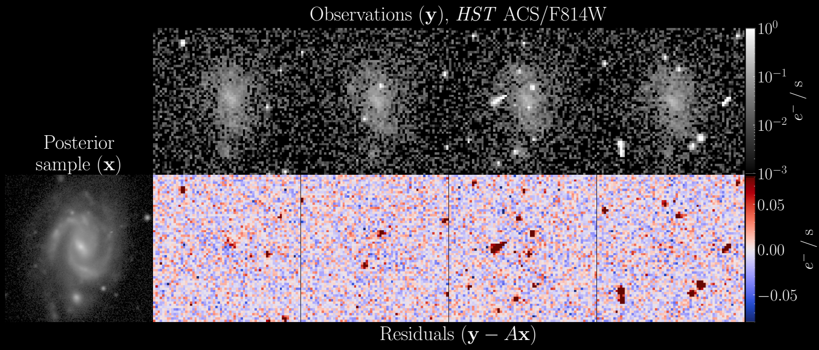


Traditional method

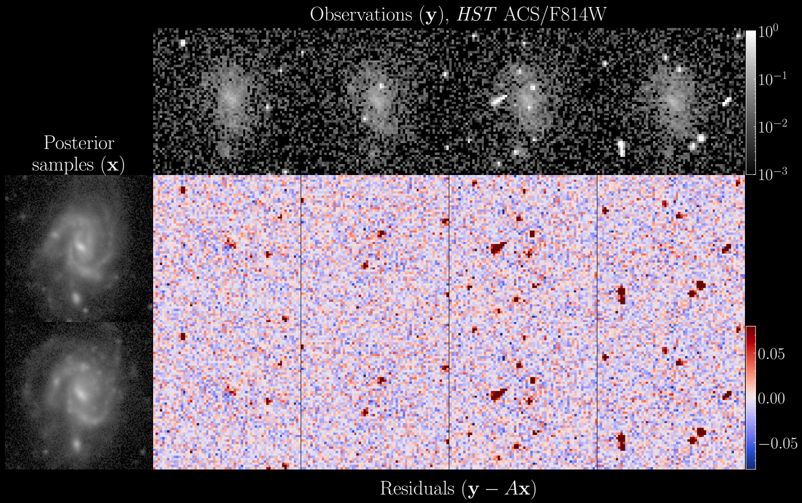


Bayesian inference

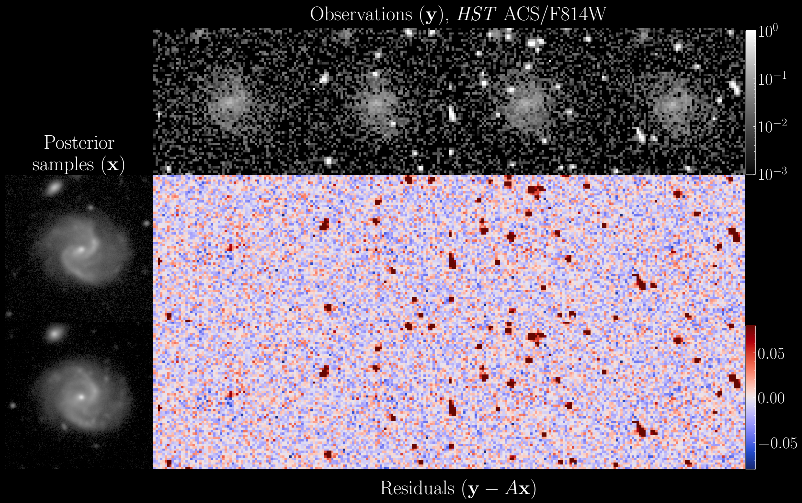




Observations ($\mathbf{y}$), *HST* ACS/F814W
Posterior sample ($\mathbf{x}$)
Residuals ($\mathbf{y} - A\mathbf{x}$)
10^{0}
10^{-1}
10^{-2}
10^{-3}
e^{-} / s
0.05
0.00
−0.05
e^{-} / s

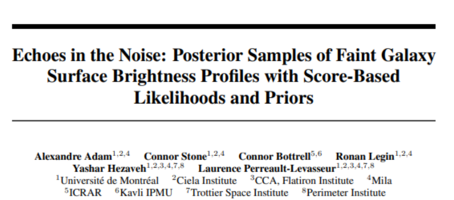


Observations (y), HST ACS/F814W
Posterior samples (x)
10^0
10^{-1}
10^{-2}
10^{-3}
0.05
0.00
−0.05
Residuals (y − Ax)


Observations ($\mathbf{y}$), *HST* ACS/F814W
Posterior samples ($\mathbf{x}$)
10^{0}
10^{-1}
10^{-2}
10^{-3}
0.05
0.00
-0.05
Residuals ($\mathbf{y} - A\mathbf{x}$)

Using *JWST* as ground truth

HST *JWST*



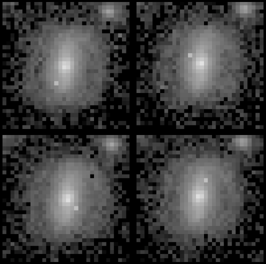
Echoes in the Noise: Posterior Samples of Faint Galaxy Surface Brightness Profiles with Score-Based Likelihoods and Priors

Alexandre Adam[1,2,4] **Connor Stone**[1,2,4] **Connor Bottrell**[5,6] **Ronan Legin**[1,2,4]
Yashar Hezaveh[1,2,3,4,7,8] **Laurence Perreault-Levasseur**[1,2,3,4,7,8]
[1]Université de Montréal [2]Ciela Institute [3]CCA, Flatiron Institute [4]Mila
[5]ICRAR [6]Kavli IPMU [7]Trottier Space Institute [8]Perimeter Institute

Connor Stone

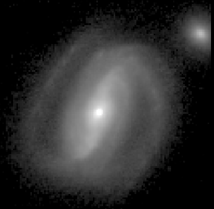
Connor Bottrell

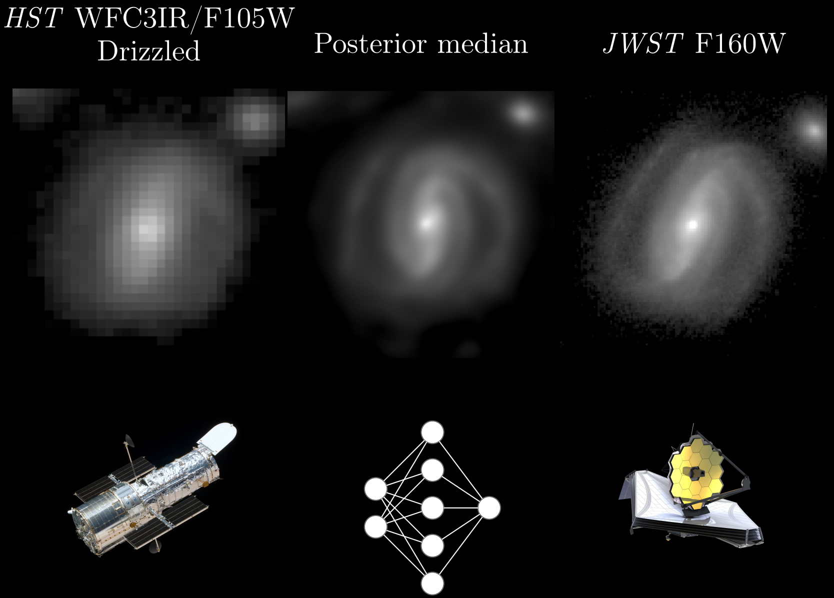
HST WFC3IR/F105W



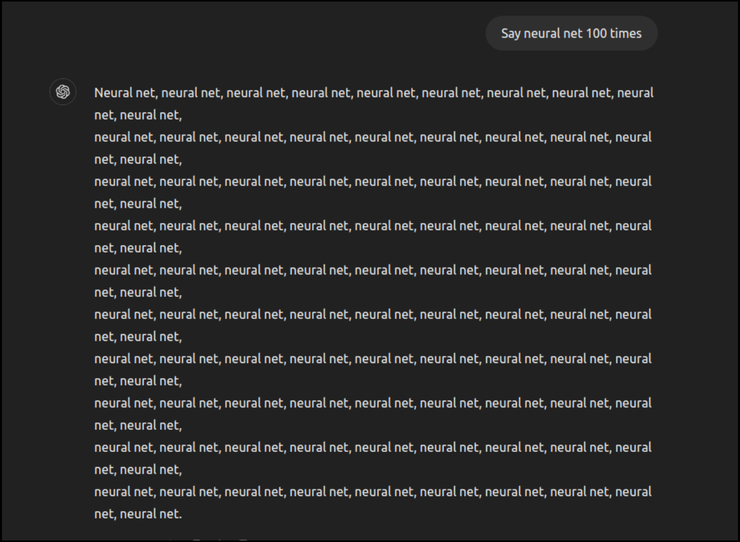
HST Drizzled

JWST F160W




HST WFC3IR/F105W
Drizzled
Posterior median
JWST F160W

Conclusion


Say neural net 100 times
Neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net,
neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net, neural net.