Astroparticle Physics and AI

A (Biased) Summary

Siddharth Mishra-Sharma

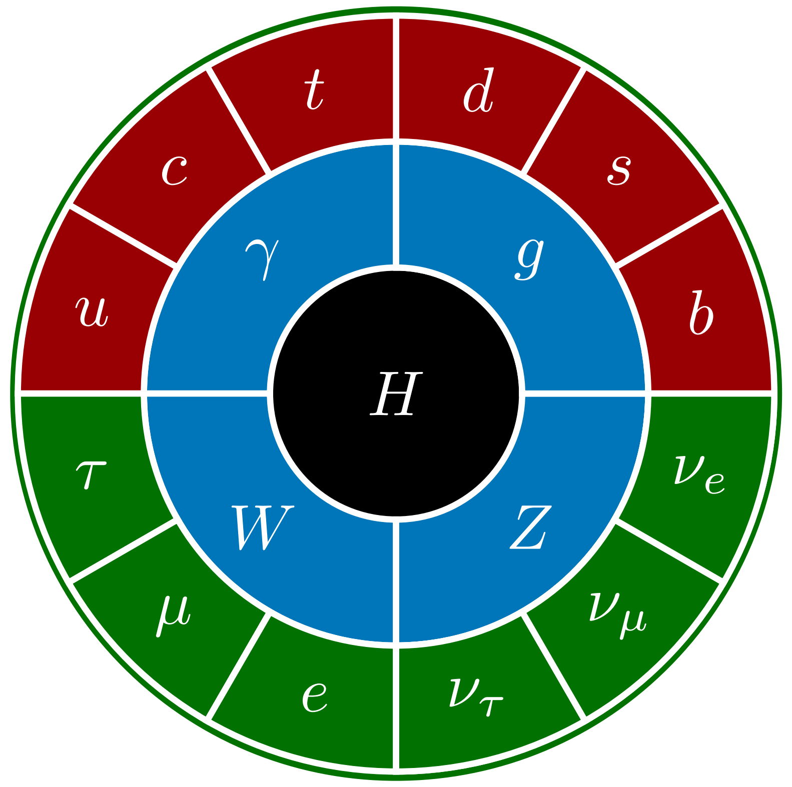
NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI)

European AI for Fundamental Physics Conference (EuCAIFCon)

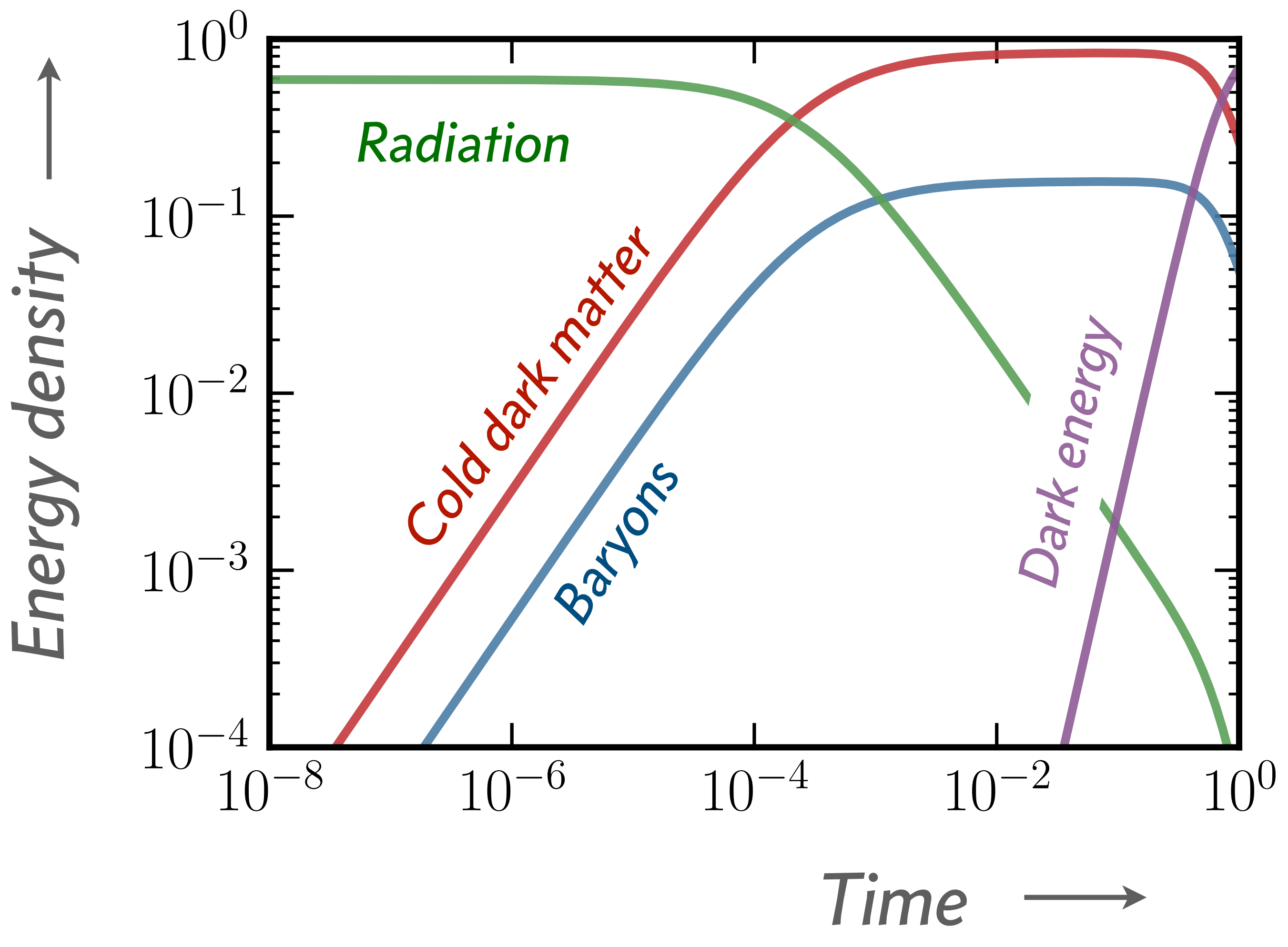
May 3, 2024

Models and Observations

Standard Model

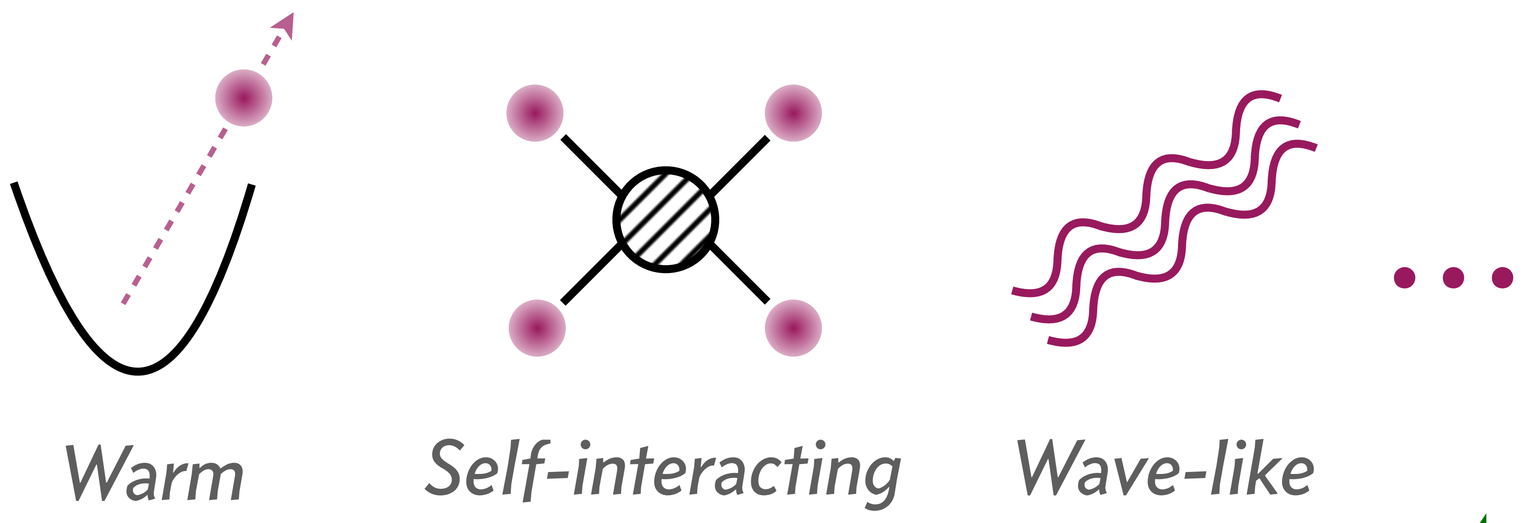


Standard Cosmology

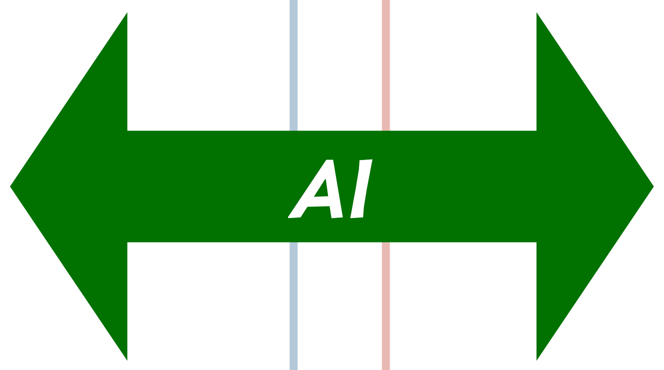


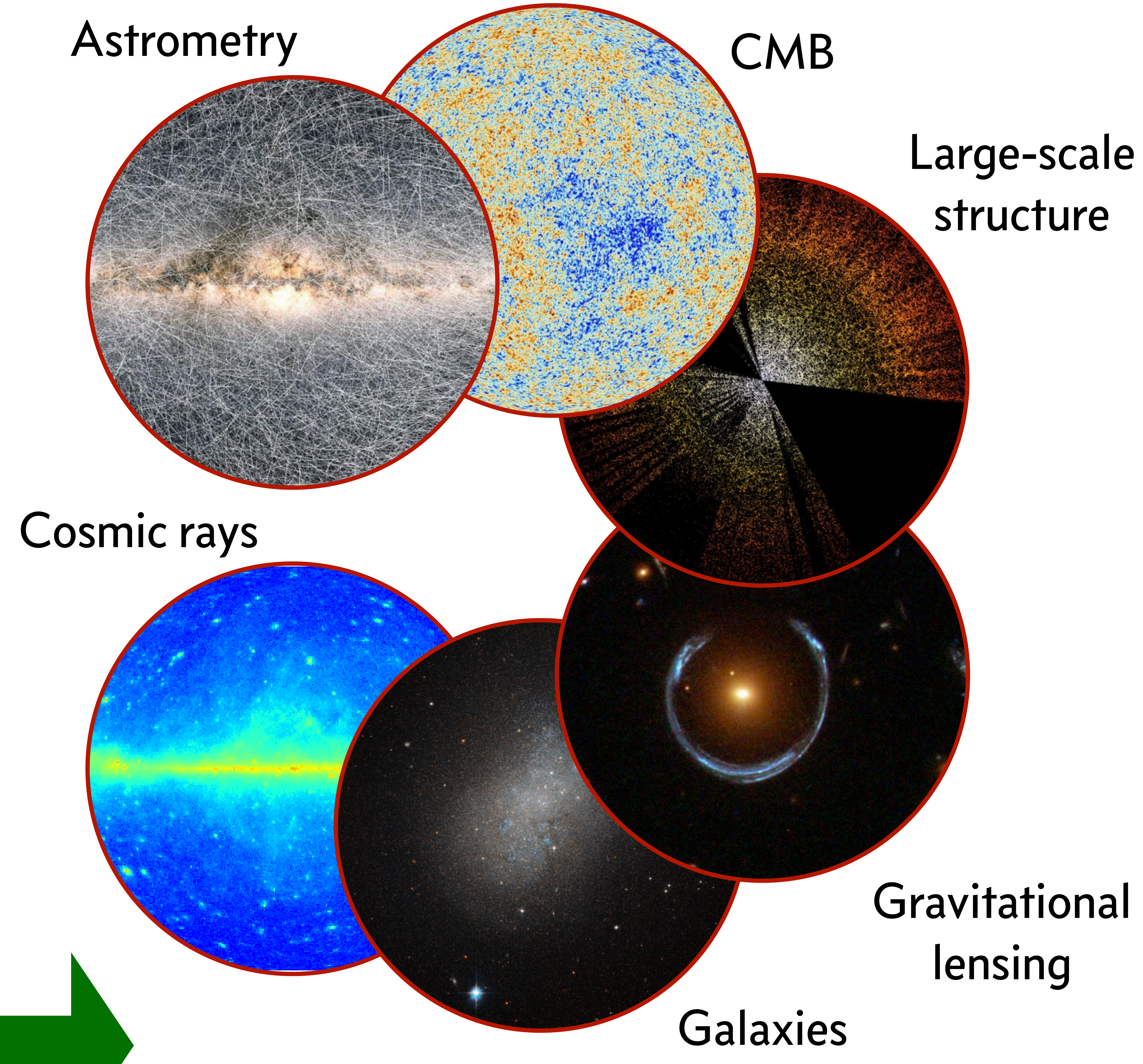
New Physics

(e.g. particle dark matter)



Models





Observations

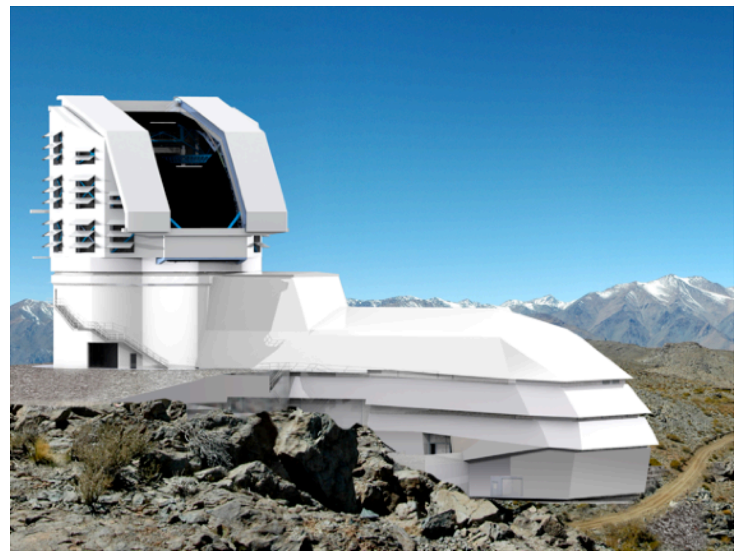
e.g.

Lots of data is on the way…

DESI

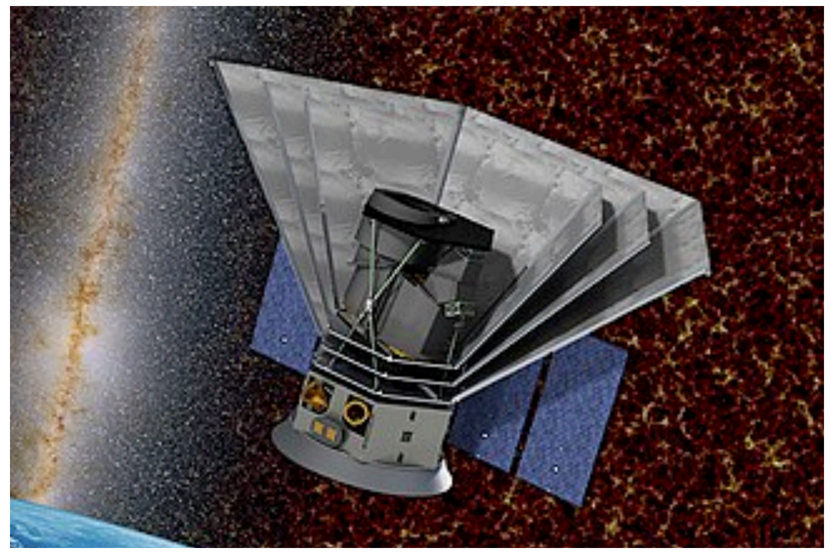
Spectroscopic survey

Vera Rubin Observatory



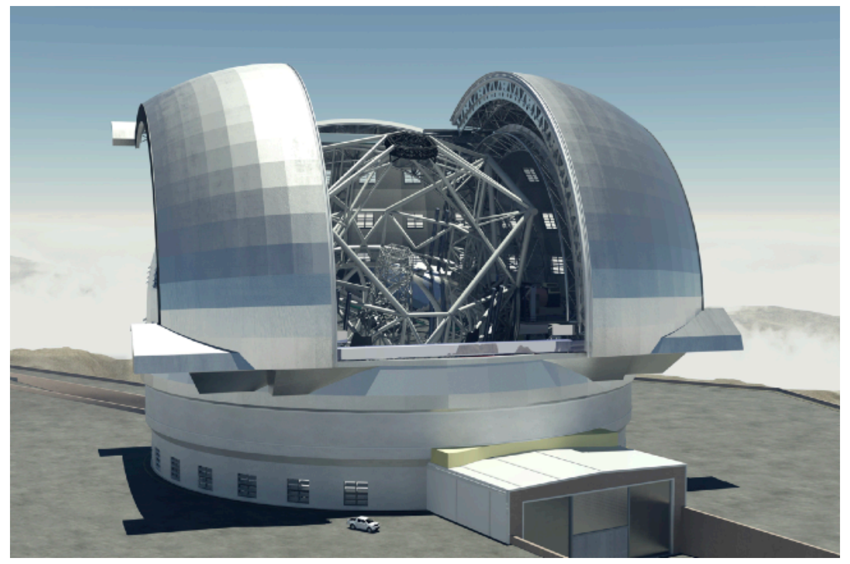
Deep optical survey

SPHEREx



Near-IR survey

Extremely Large Telescope



Optical/Near-IR

2022

2024

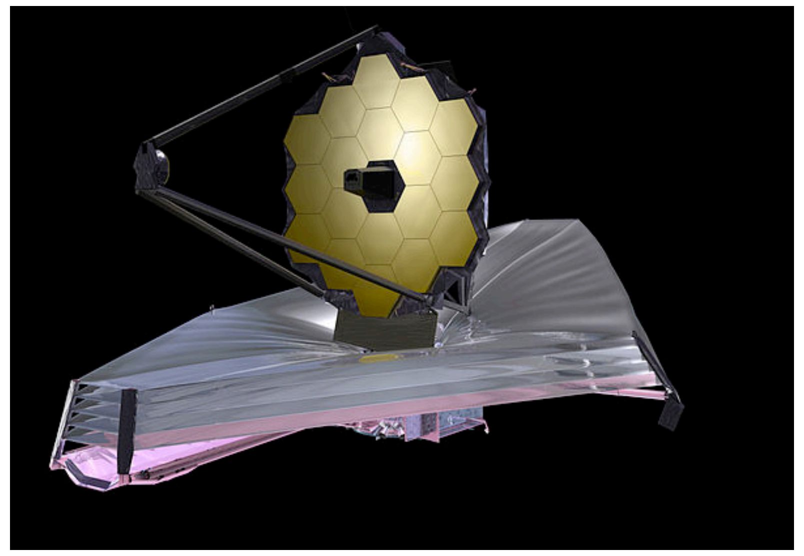
2026

2023

2025

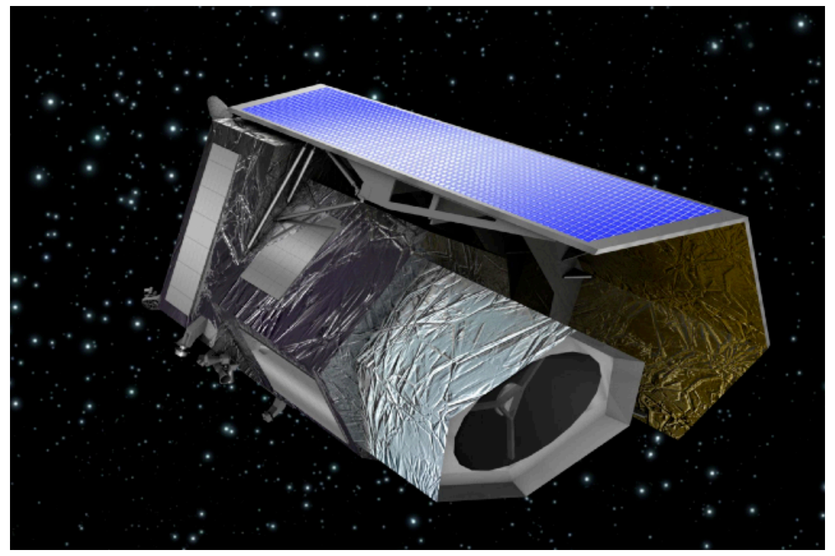
2027

JWST



Optical/Near-IR survey

Euclid

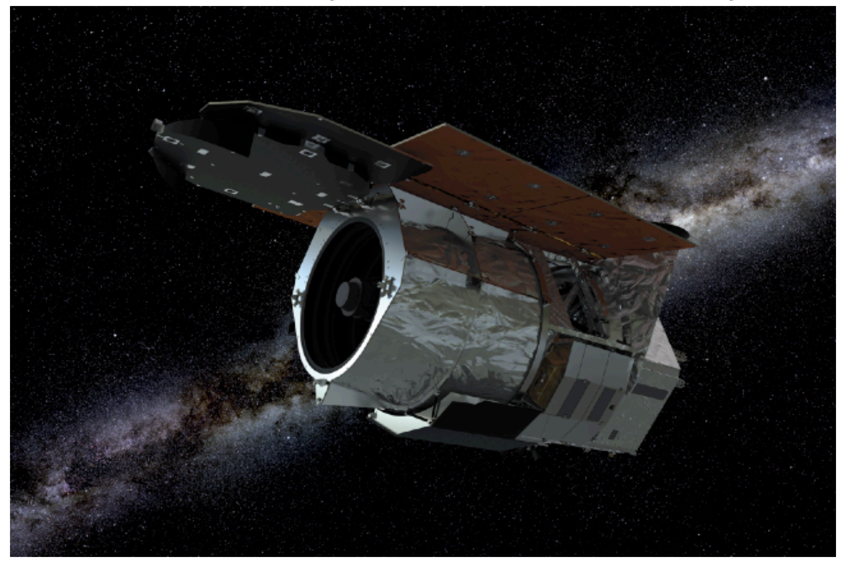


Deep optical survey

Simons Observatory

Next-generation CMB observatory

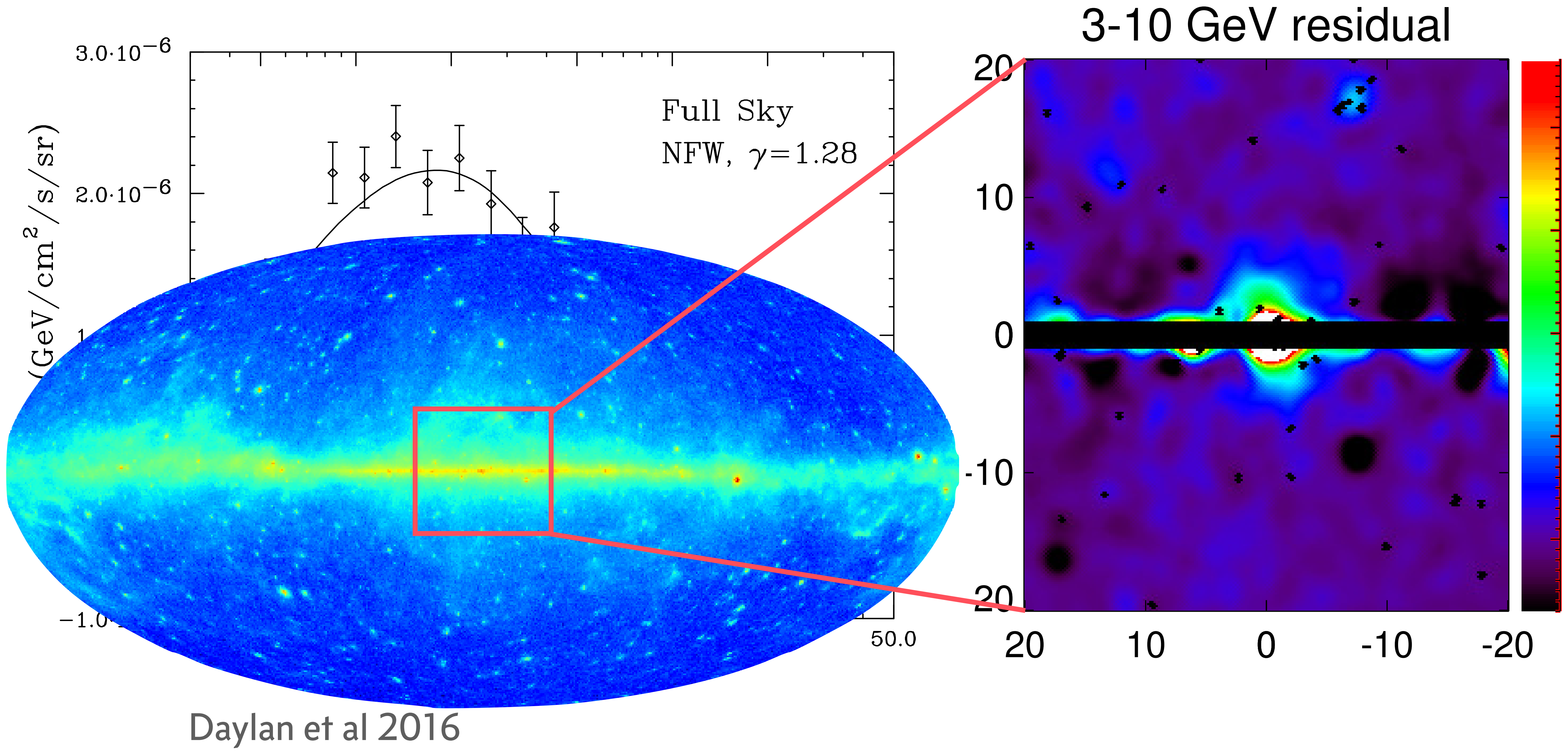
Roman Space Telescope



Wide optical survey

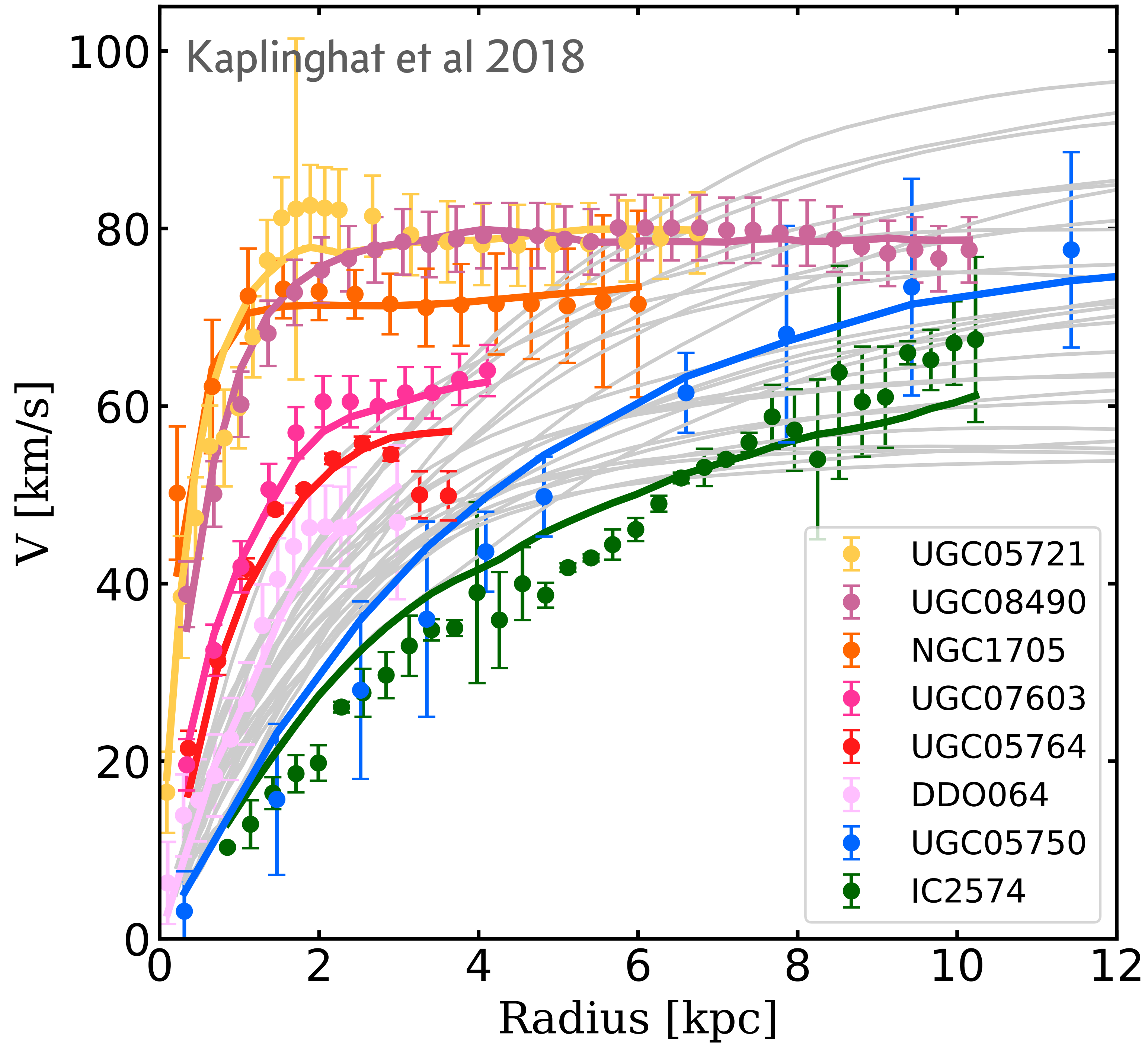
Signals and impasses

Fermi Galactic Center Excess



Annihilating dark matter? Millisecond pulsars?

Diversity of dark matter halo shapes

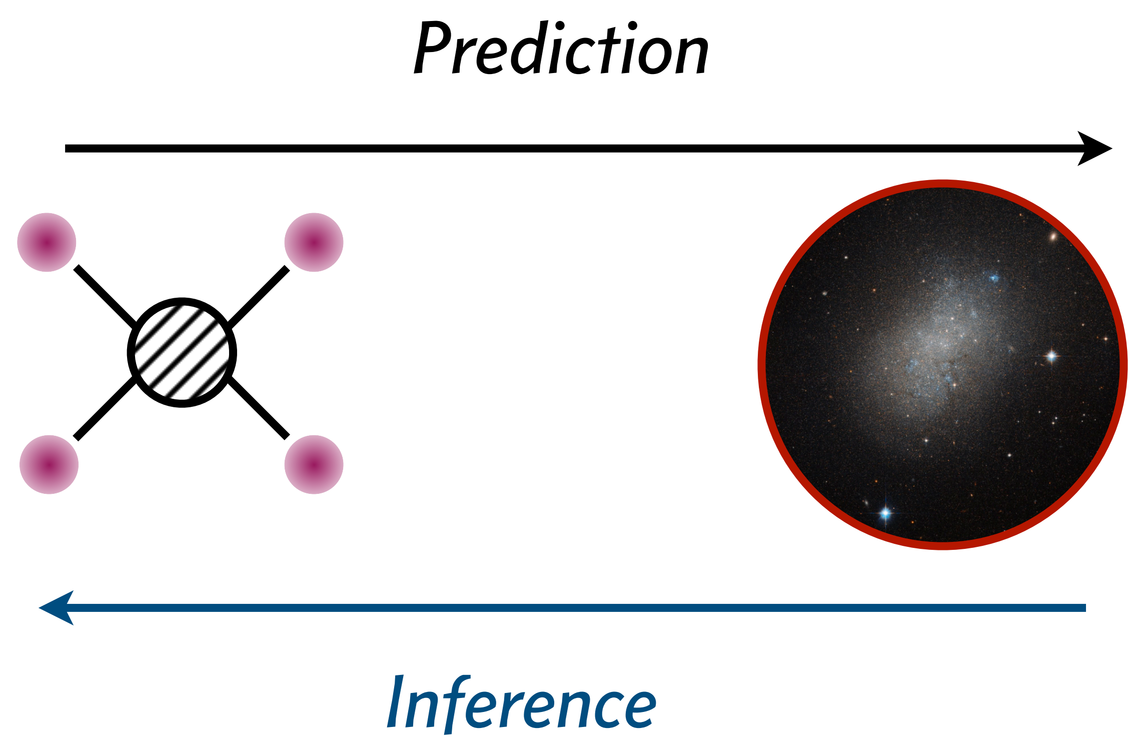


New fundamental physics? Baryonic effects?

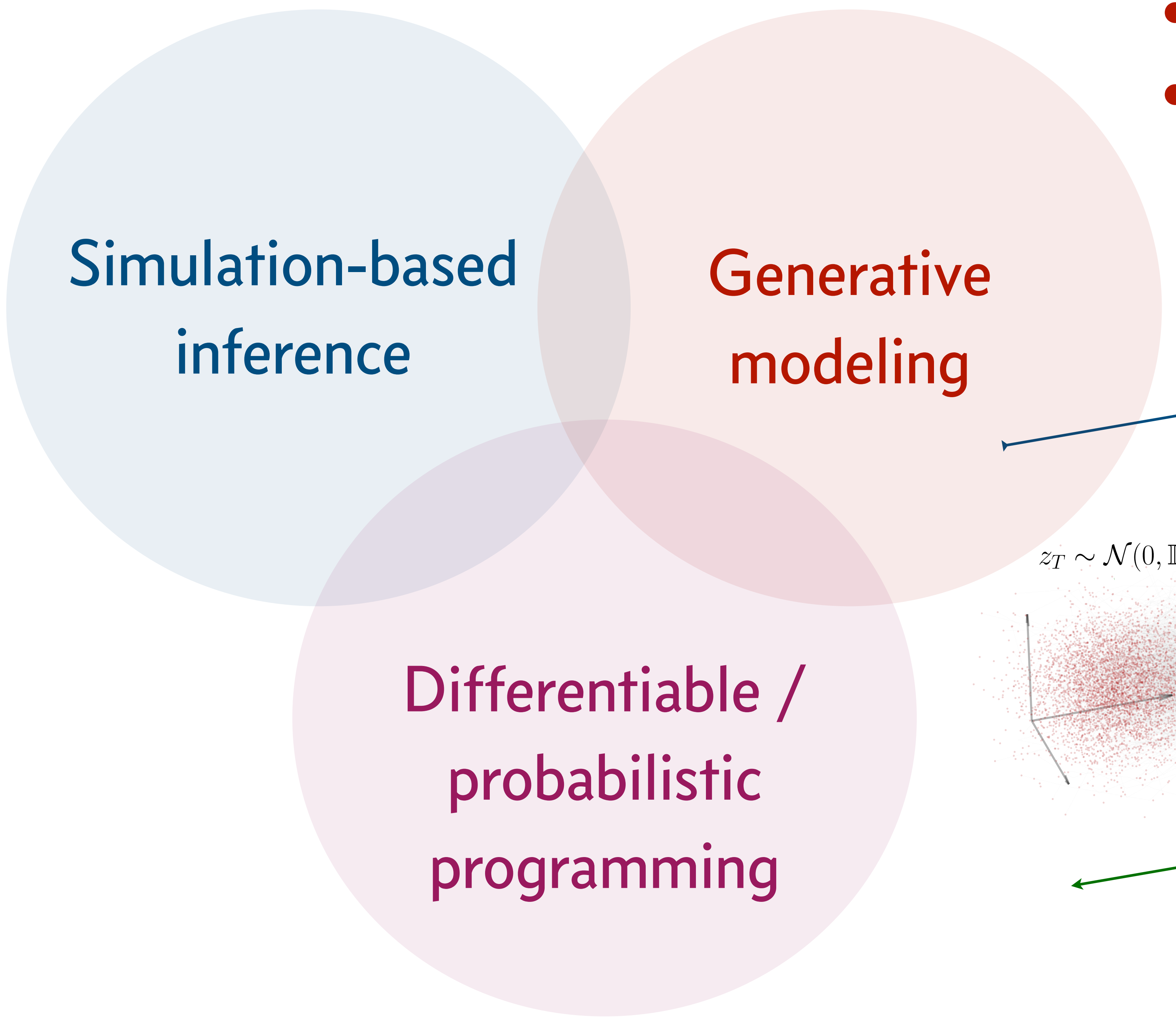
Ability to make robust conclusions is often limited by by challenges in connecting theory to data

Broad methodological directions*

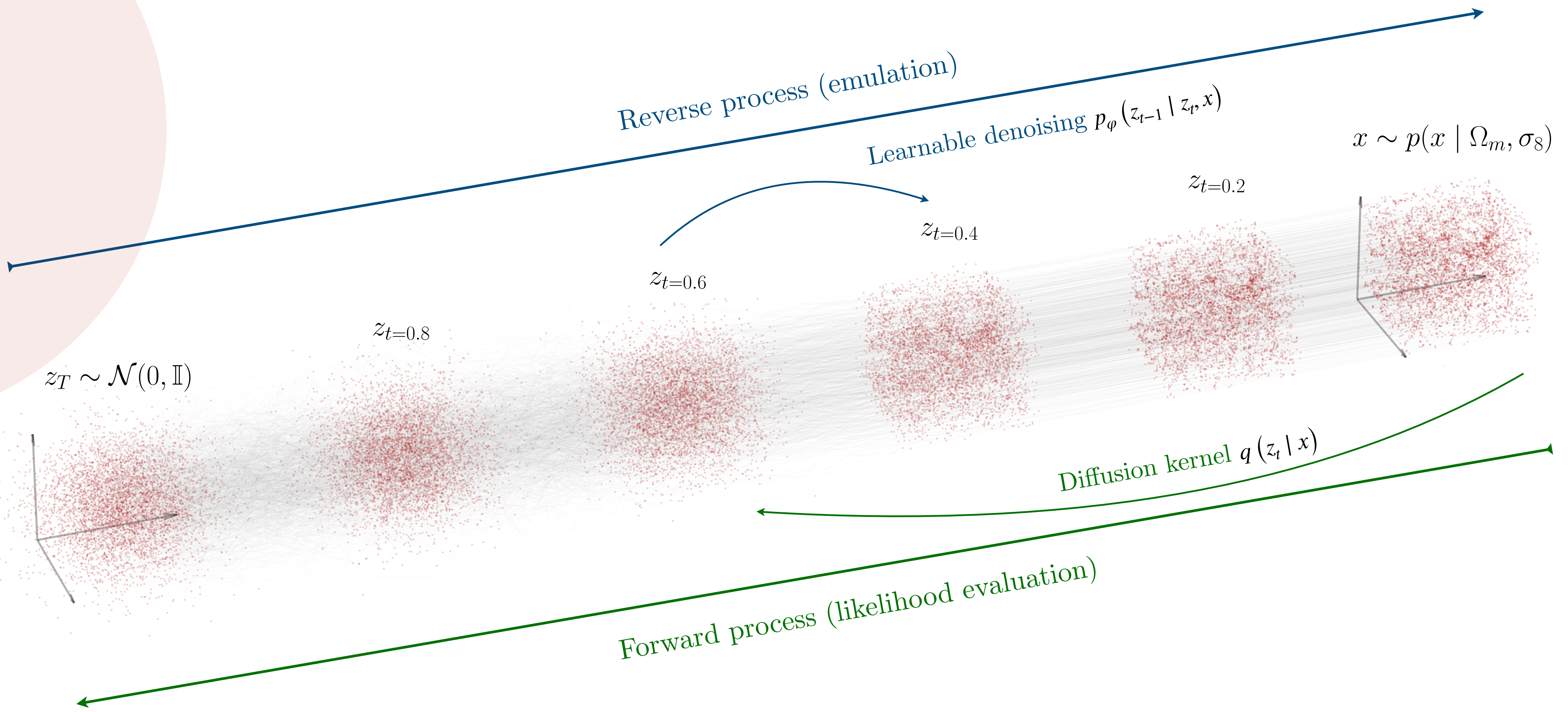
*Not exclusive or exhaustive!

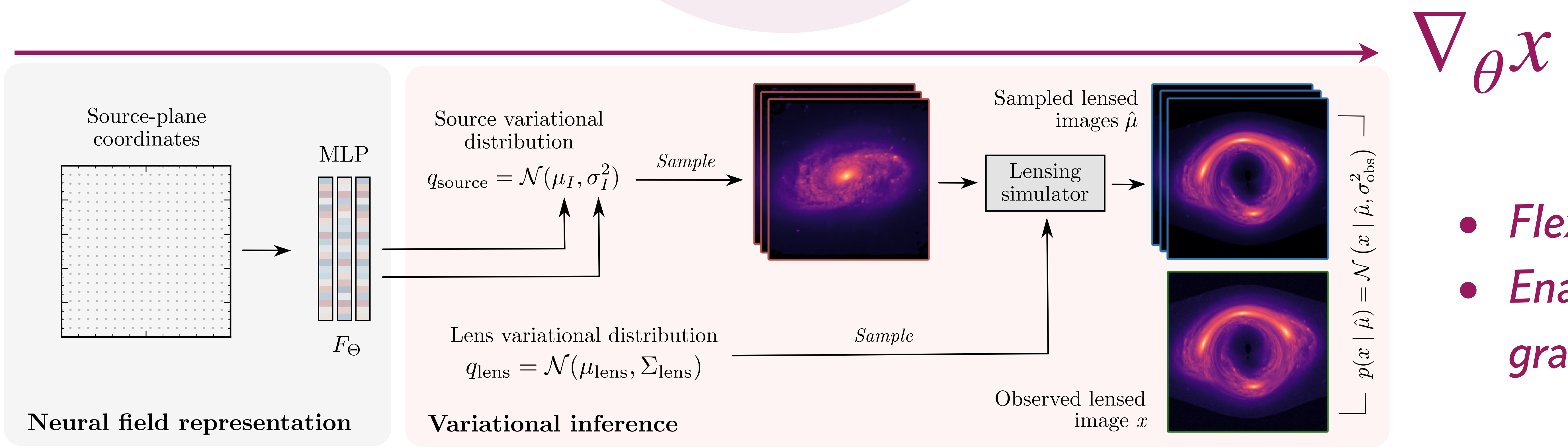


- *Invert complex physical simulators*
- *Directly work with high-dim data*



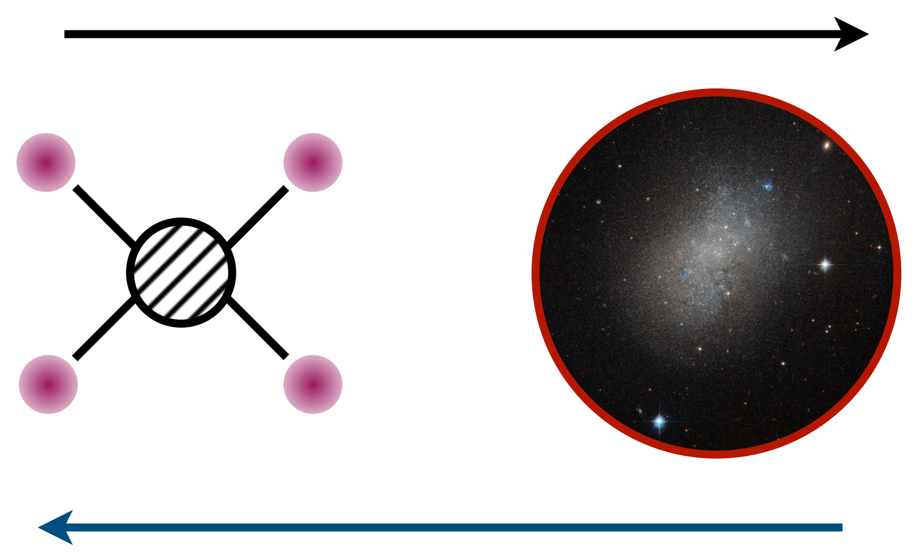
- *Encode complex physical distribution*
- *Use end-to-end or as physical prior*





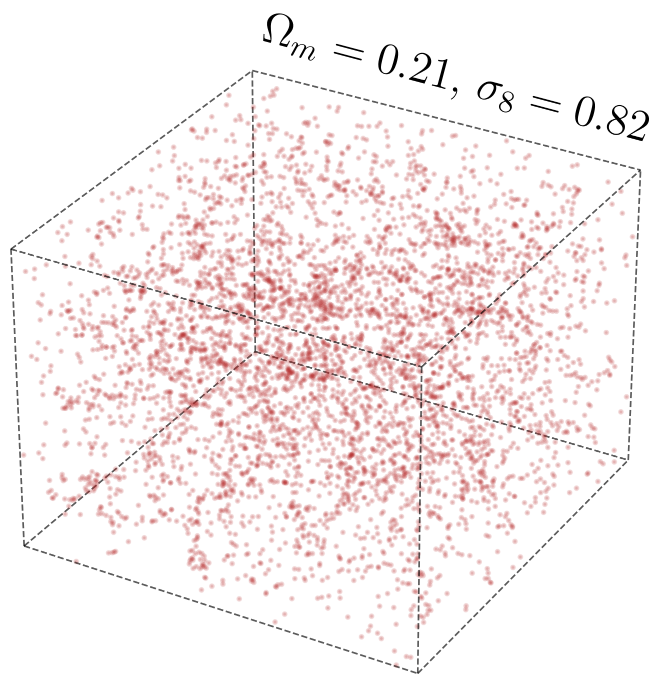
- *Flexible specification of model components*
- *Enable high-dimensional optimization using gradient-based inference techniques*

Outline



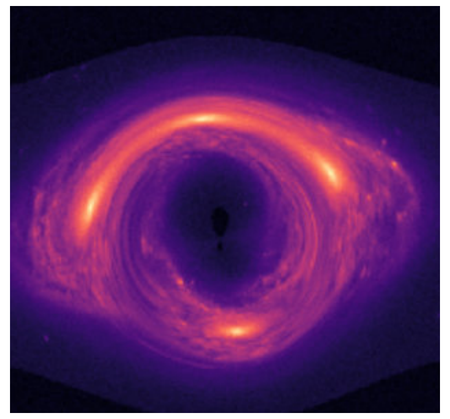
Simulation-based inference

Inverting complex physical simulators



Generative modeling

Capturing the distribution of complex data for emulation and inference

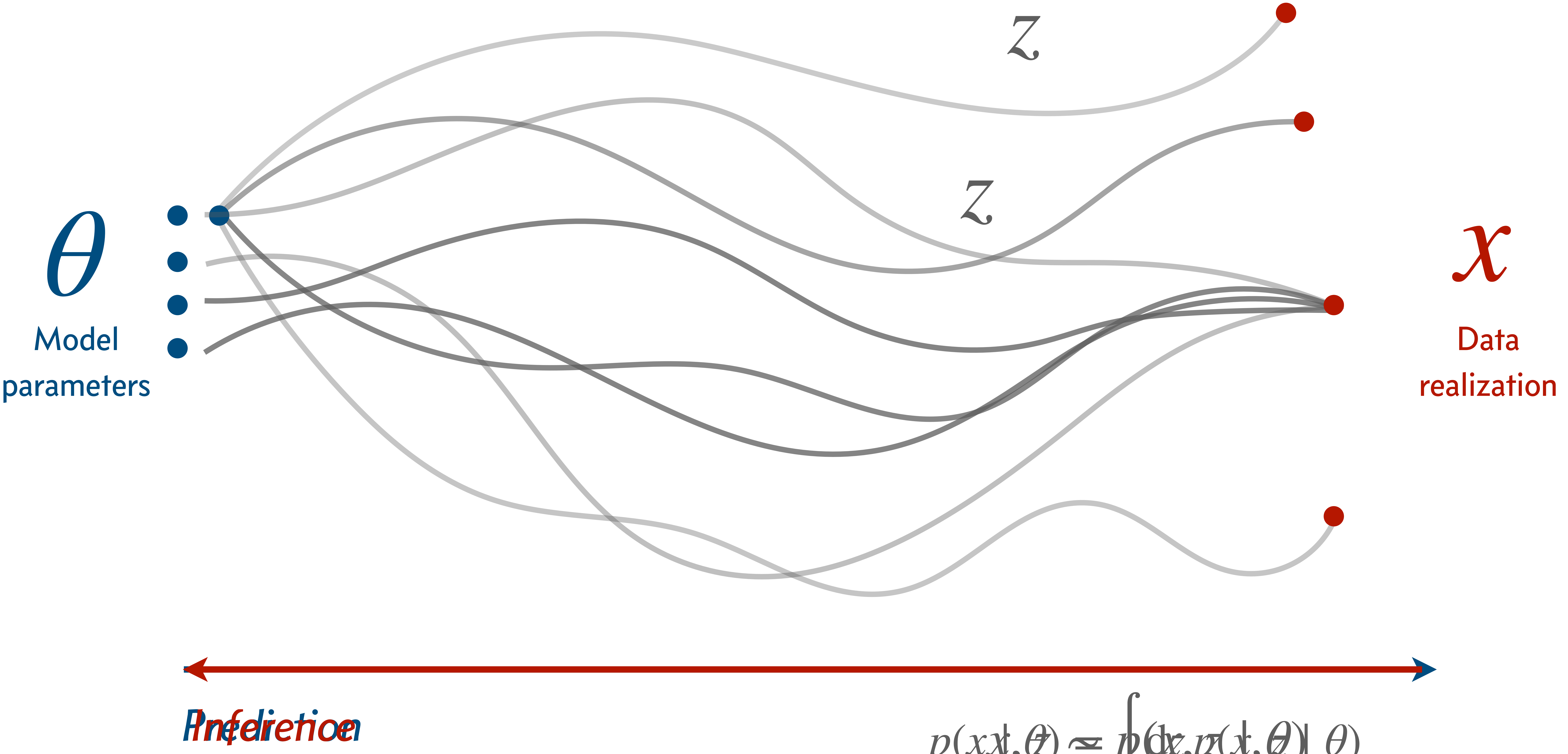
∇_θ 

Differentiable and probabilistic programming

Specifying models with autodiff capabilities and enabling flexible inference

Simulators are great for prediction tasks

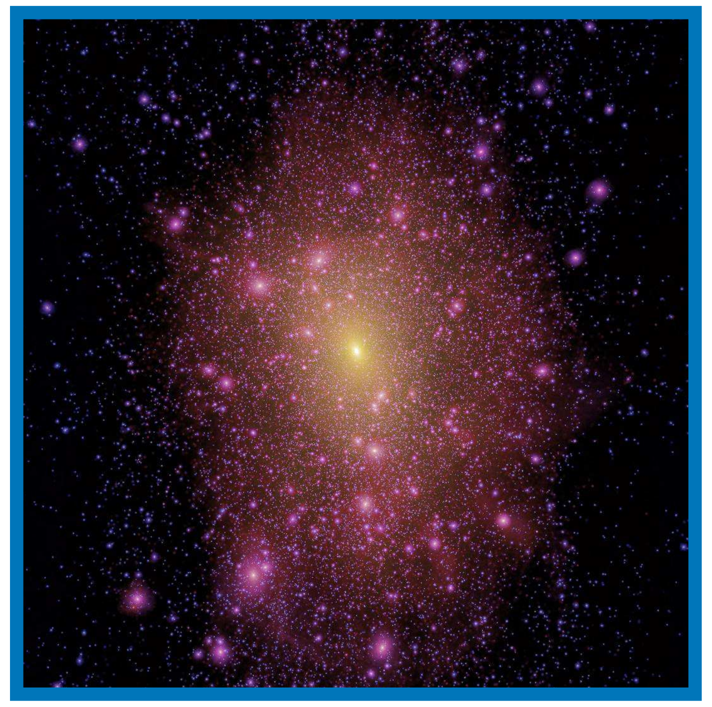
But poorly suited for inference...



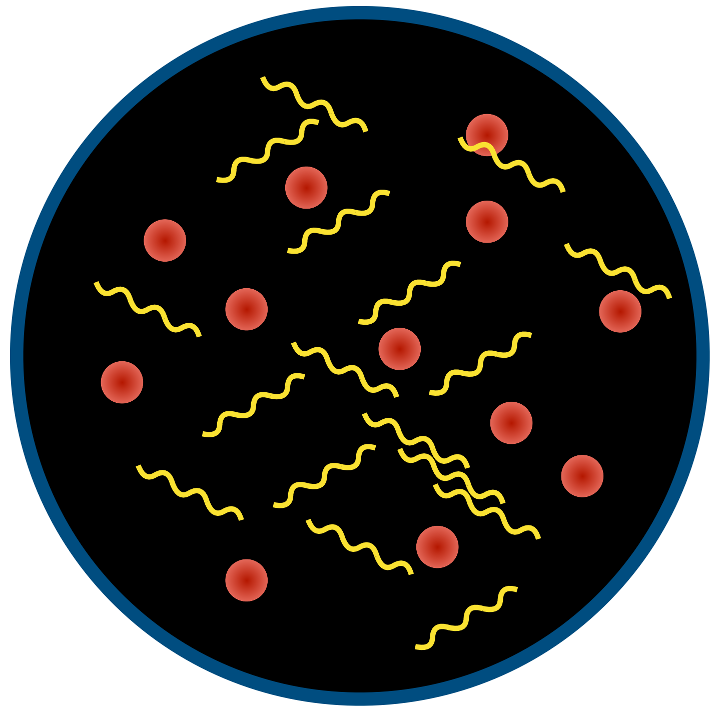
$$p(x \mid \theta) = \int p(x, z \mid \theta)\, dz$$

Inference with summary statistics

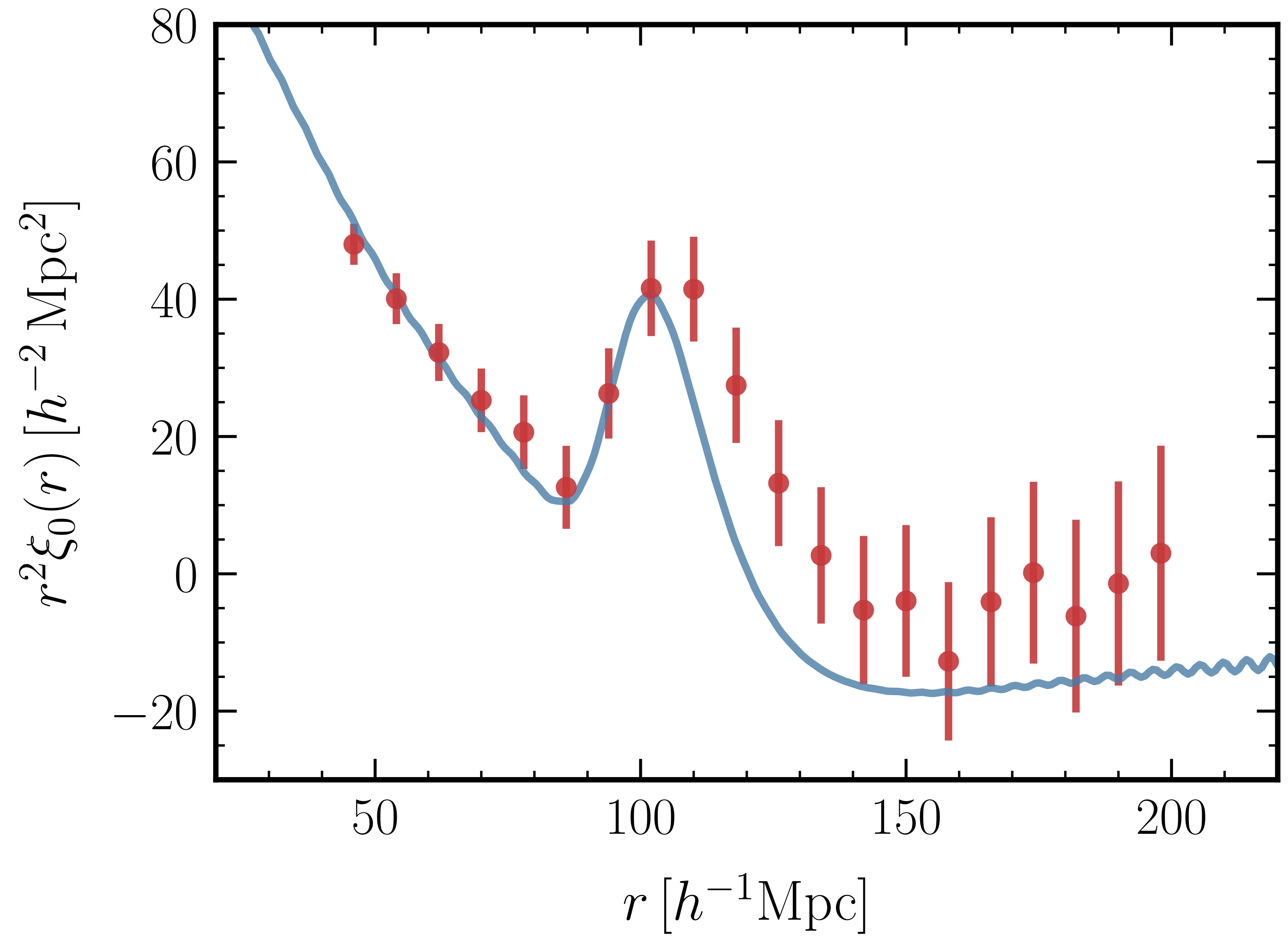
Simulations



(Semi)-analytic models

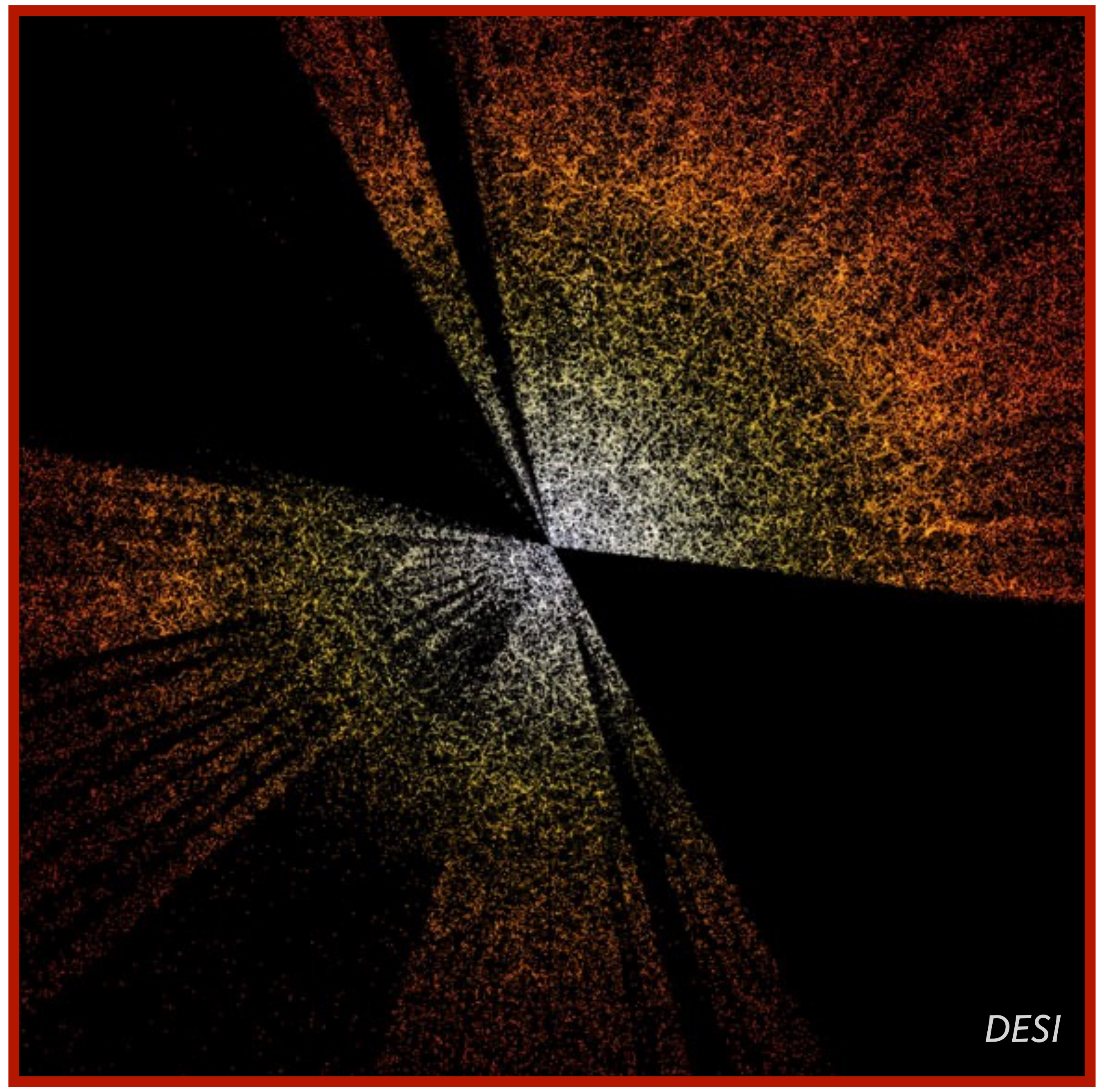


Summary



We typically rely on *simplified summaries* like correlation functions

Observations



Data is complex and high-dimension

We'd like to use observations and models to their full complexity

SBI in astro(particle) physics

~ 70% of applications in astro/cosmology!

SBI is well-suited to many problems in astro/cosmology

Simulation-based inference

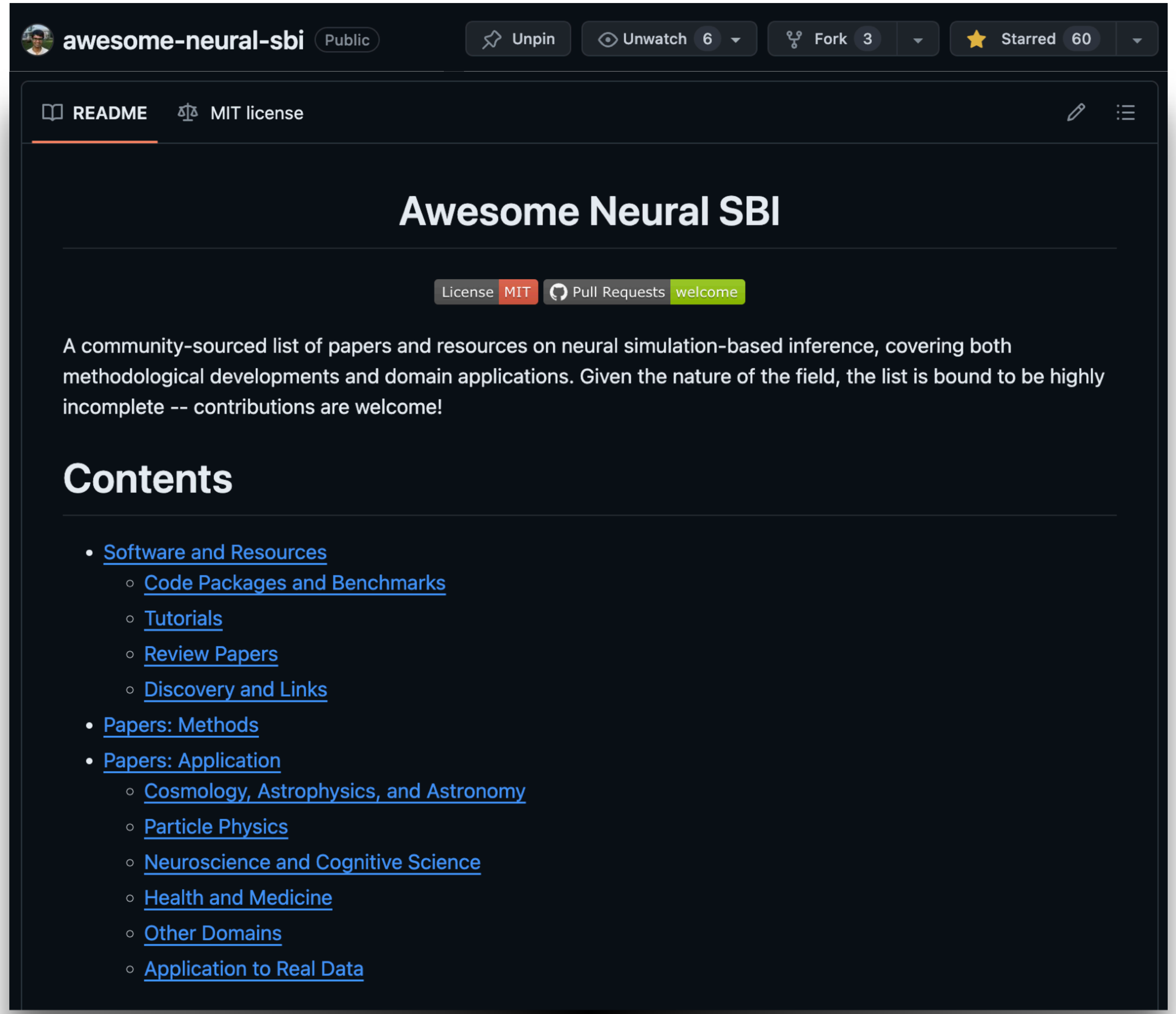
Astrophysics/ cosmology

Problems in astro/cosmology are an opportunity to drive methodological developments in SBI

Proliferation also driven by investment in tools:

`swyft`, `sbi`

https://github.com/smsharma/awesome-neural-sbi



See also https://simulation-based-inference.org/

Astrophysical dark matter searches: microphysics from macrophysics

Signs of new physics can show up in the *macroscopic distribution* of matter

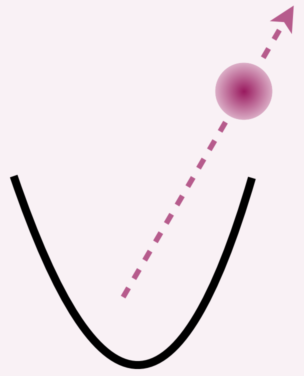
Distribution of dark matter

Cold dark matter (m_{DM} ~ GeV)

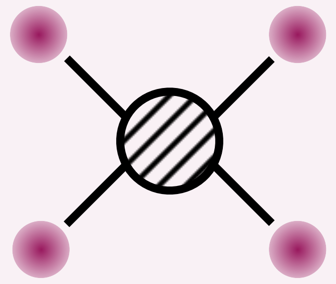
Warm dark matter (m_{DM} ~ keV)

Microphysical models

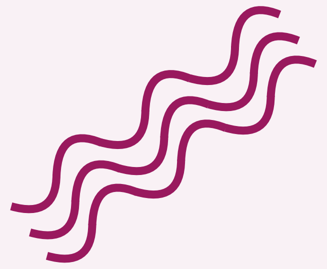
Warm dark matter



Self-interacting dark matter

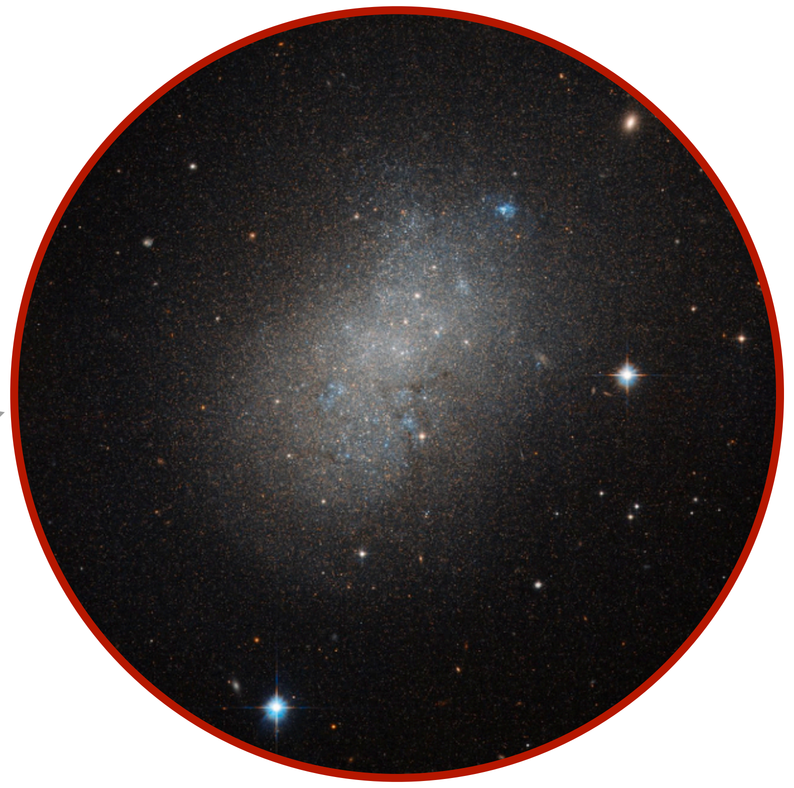


Fuzzy (wave-like) dark matter



From matter distribution to observations

Dark matter distribution

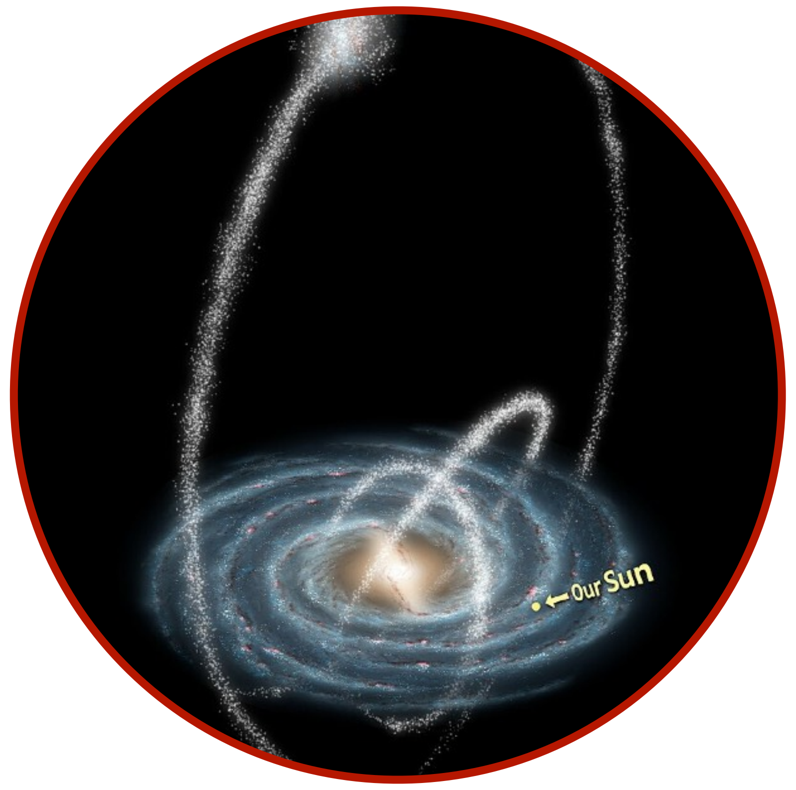


Motions of gravitationally bound stars

Inferring shapes of DM clumps

Gravitational lensing of background galaxies

Inferring (mass) distribution of DM clumps

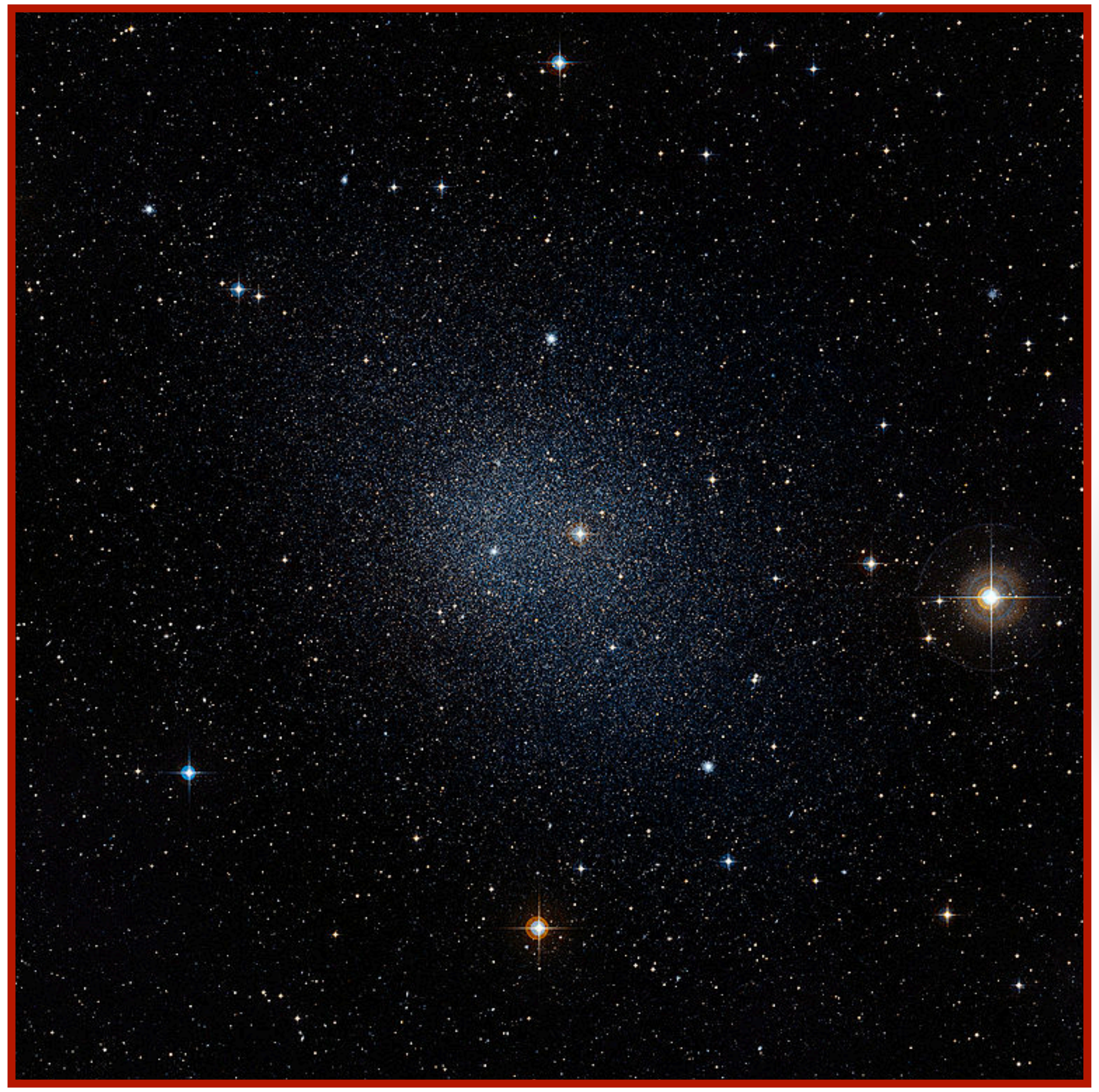


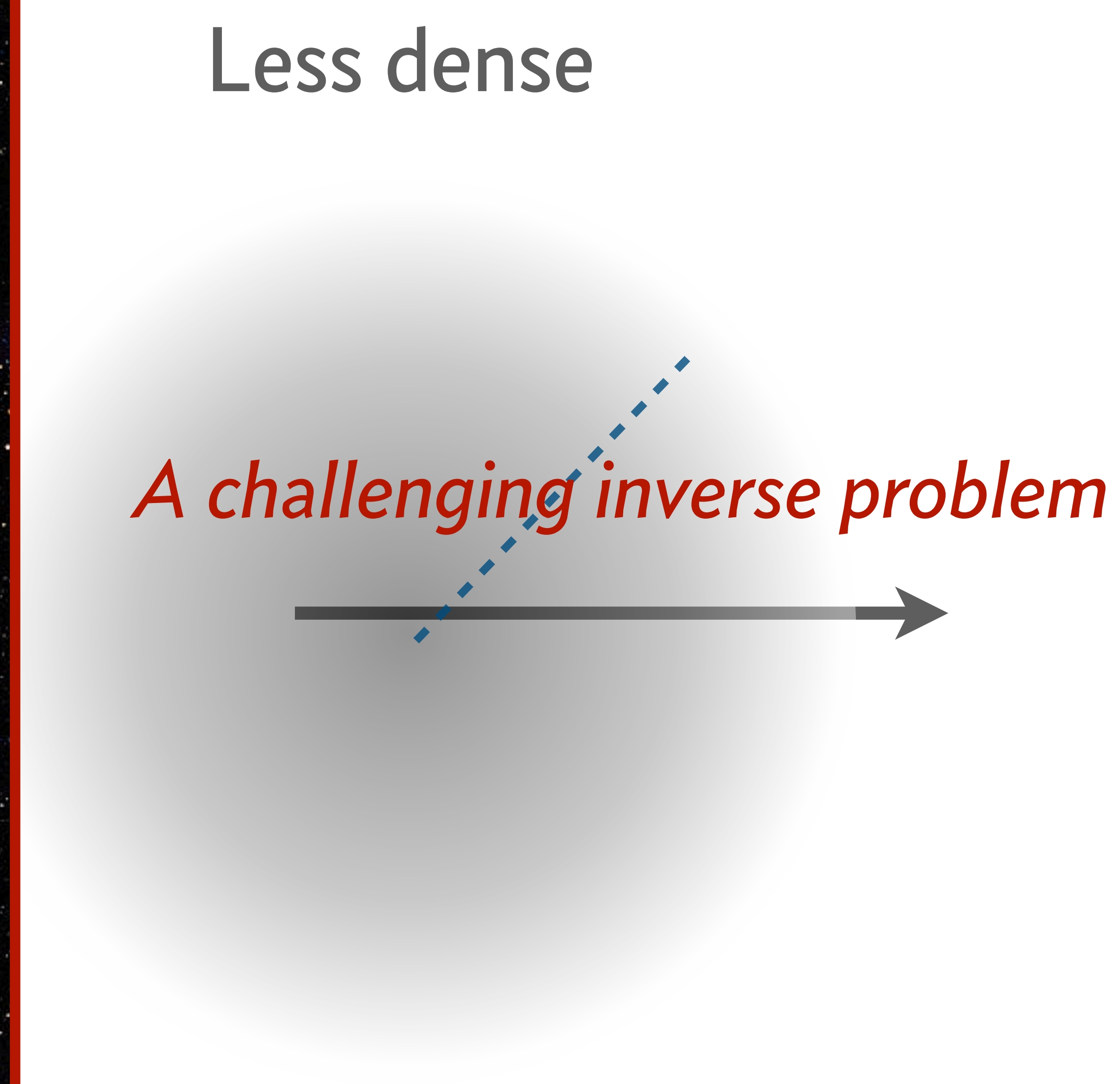
Perturbations of stellar streams

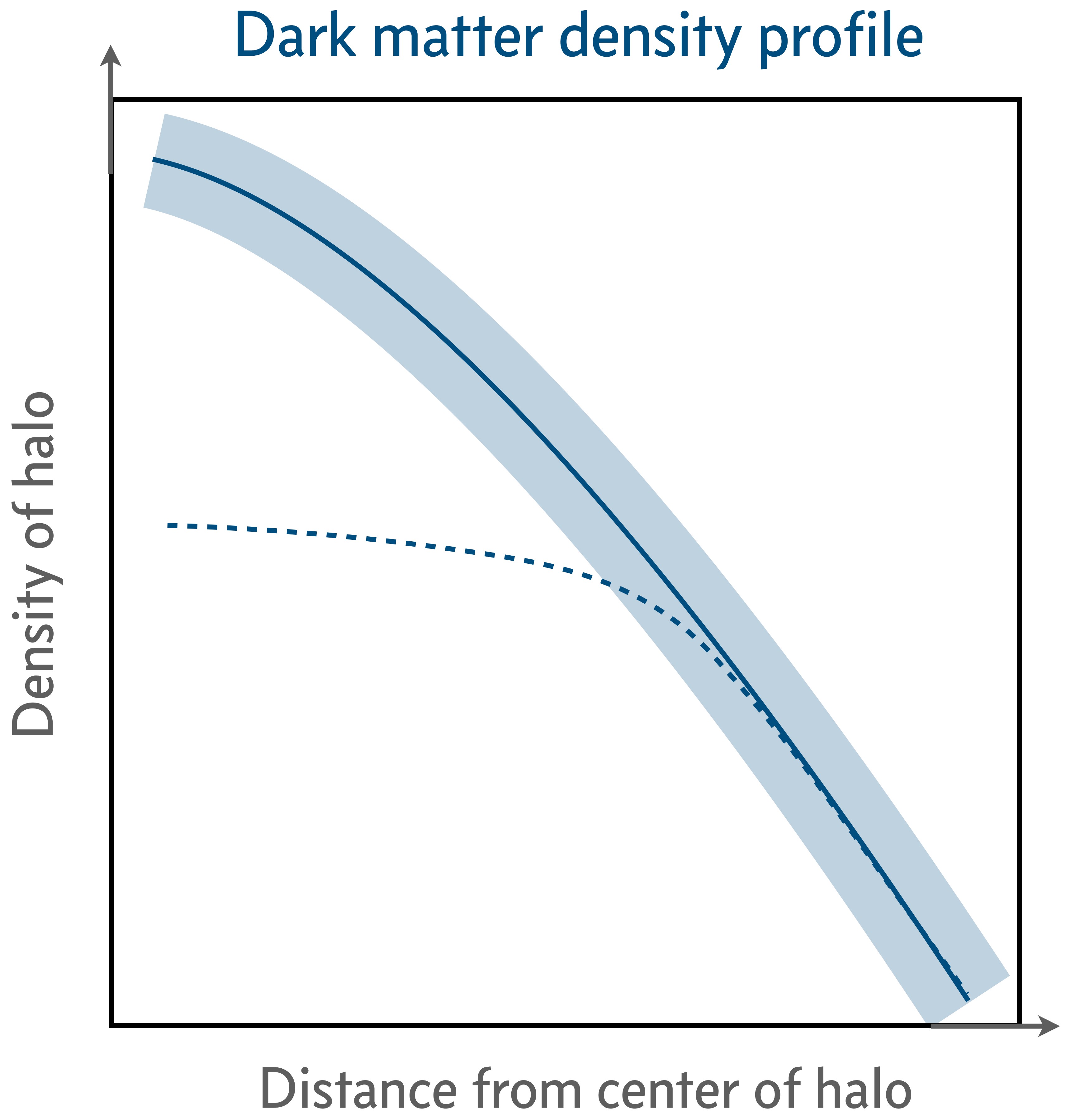
Alvey, Gerdes, Weniger [MNRAS 2023]
Hermans et al [MNRAS 2021]

Example: Learning the shape of the dark matter halos

Fornax dwarf galaxy





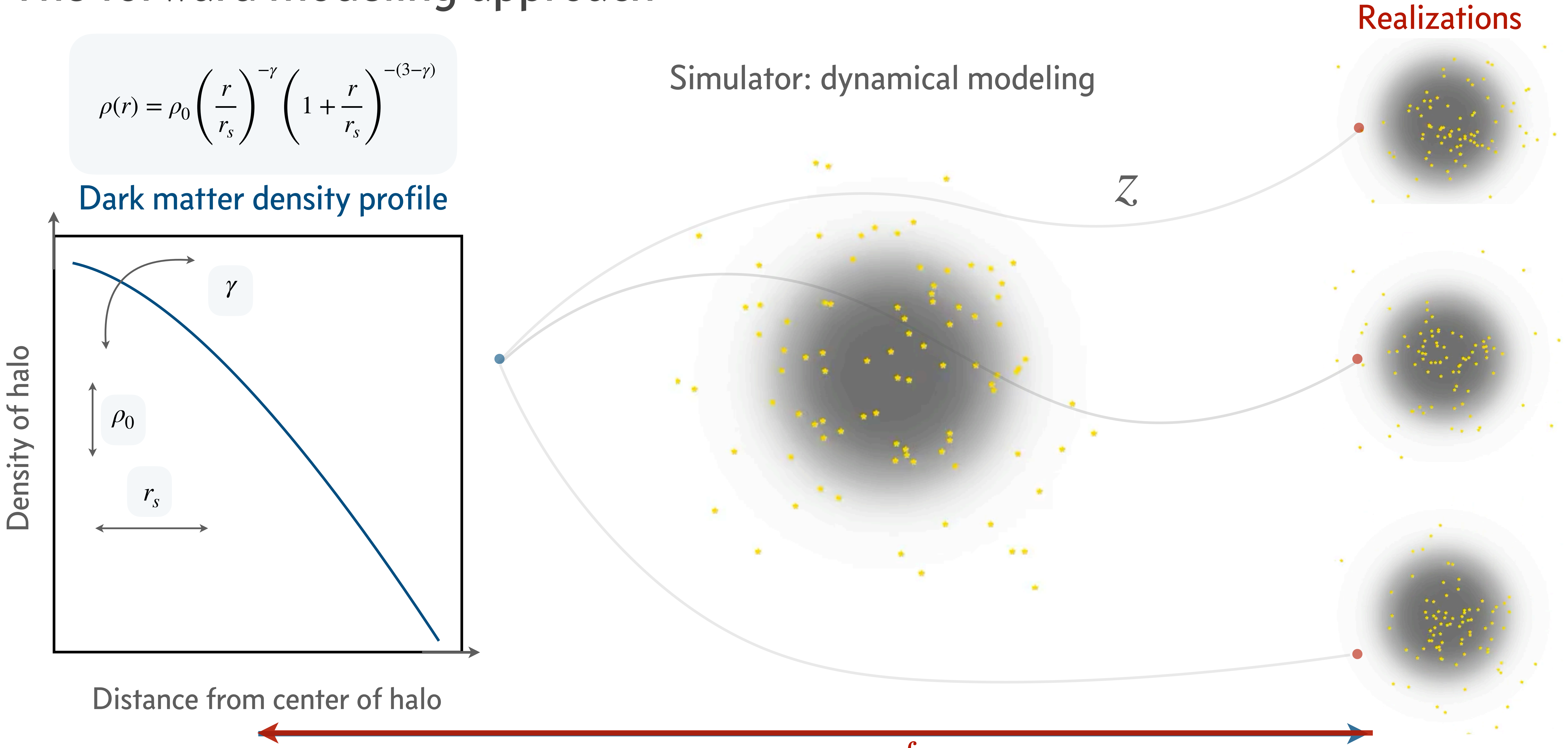


Traditional method: equilibrium dynamical modeling with low-order velocity moments (~summaries)

The forward modeling approach

Nguyen, SM et al [PRD 2023]

$$\rho(r) = \rho_0 \left(\frac{r}{r_s}\right)^{-\gamma} \left(1 + \frac{r}{r_s}\right)^{-(3-\gamma)}$$



Prediction: Simulator can generate samples $x \sim p(x \mid \theta)$

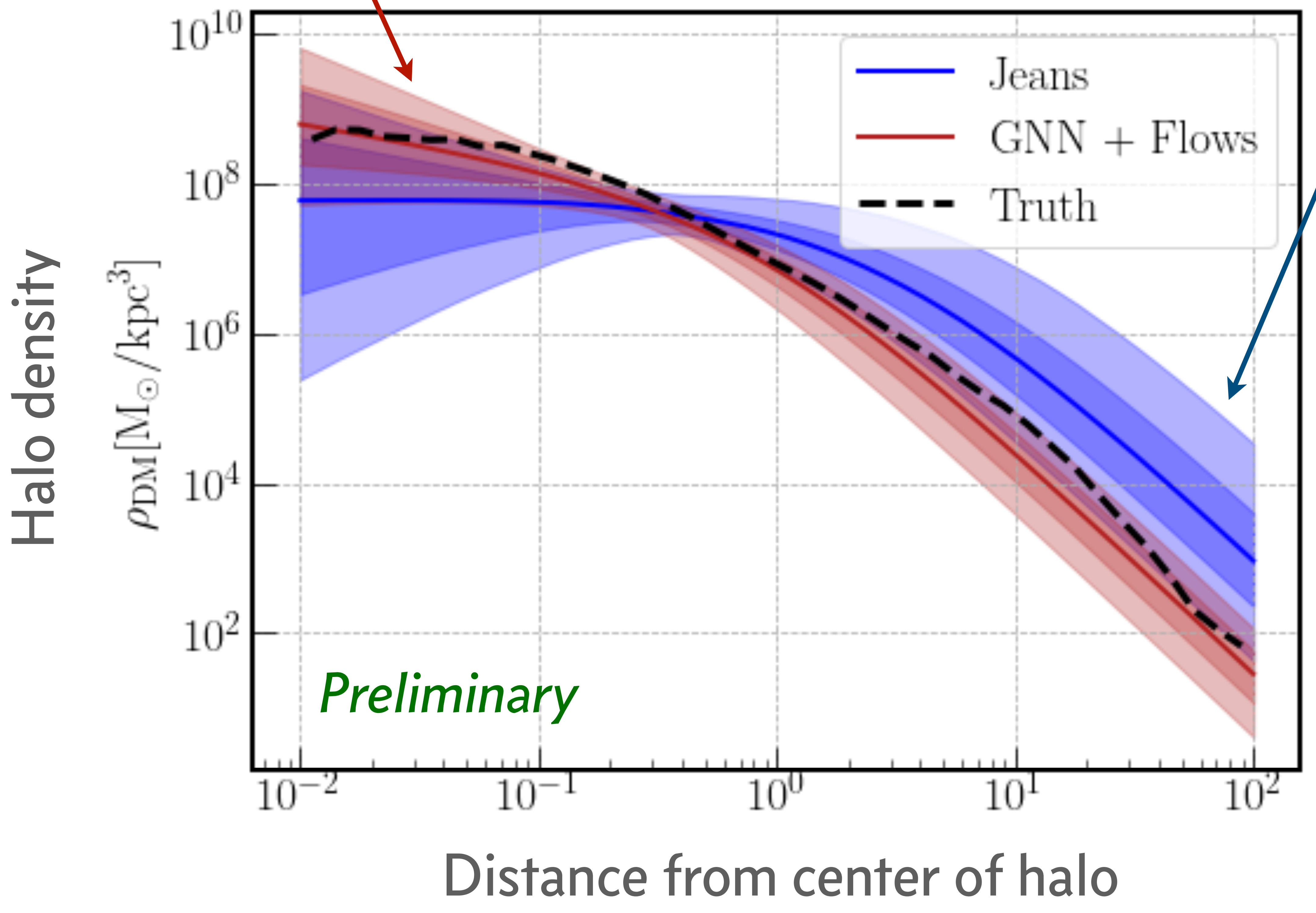
Inference: Likelihood $p(x \mid \theta) = \int \mathrm{d}z\, p(x, z \mid \theta)$ is intractable

Applications to realistic simulations

Wheeler et al [MNRAS 2019]

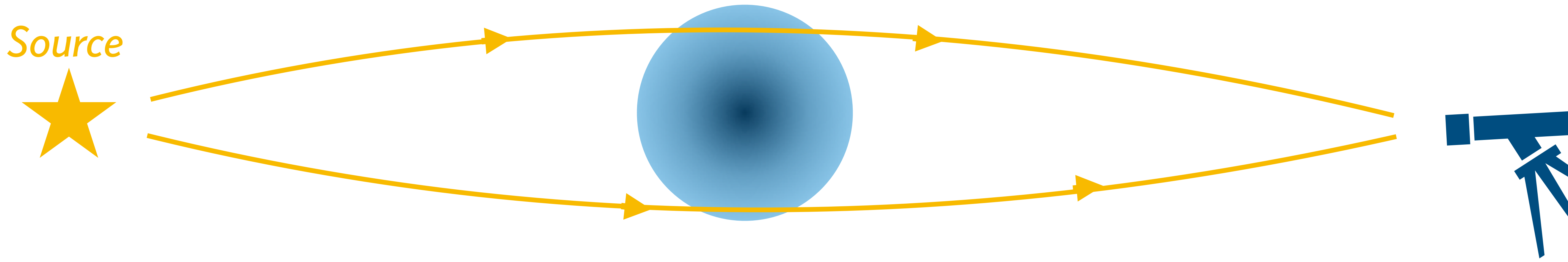
GNN + Simulation-based

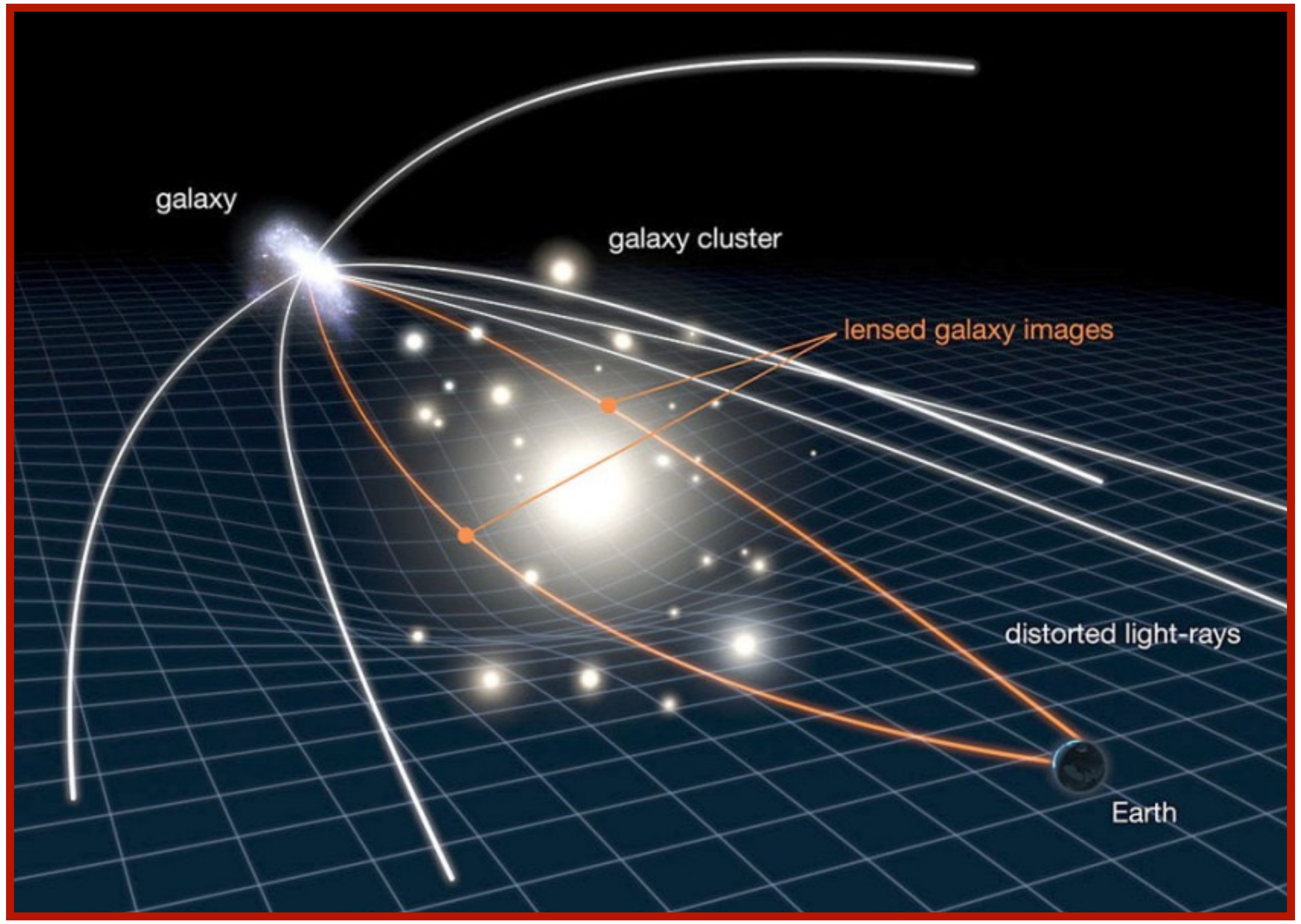
Traditional method
(Velocity moments)



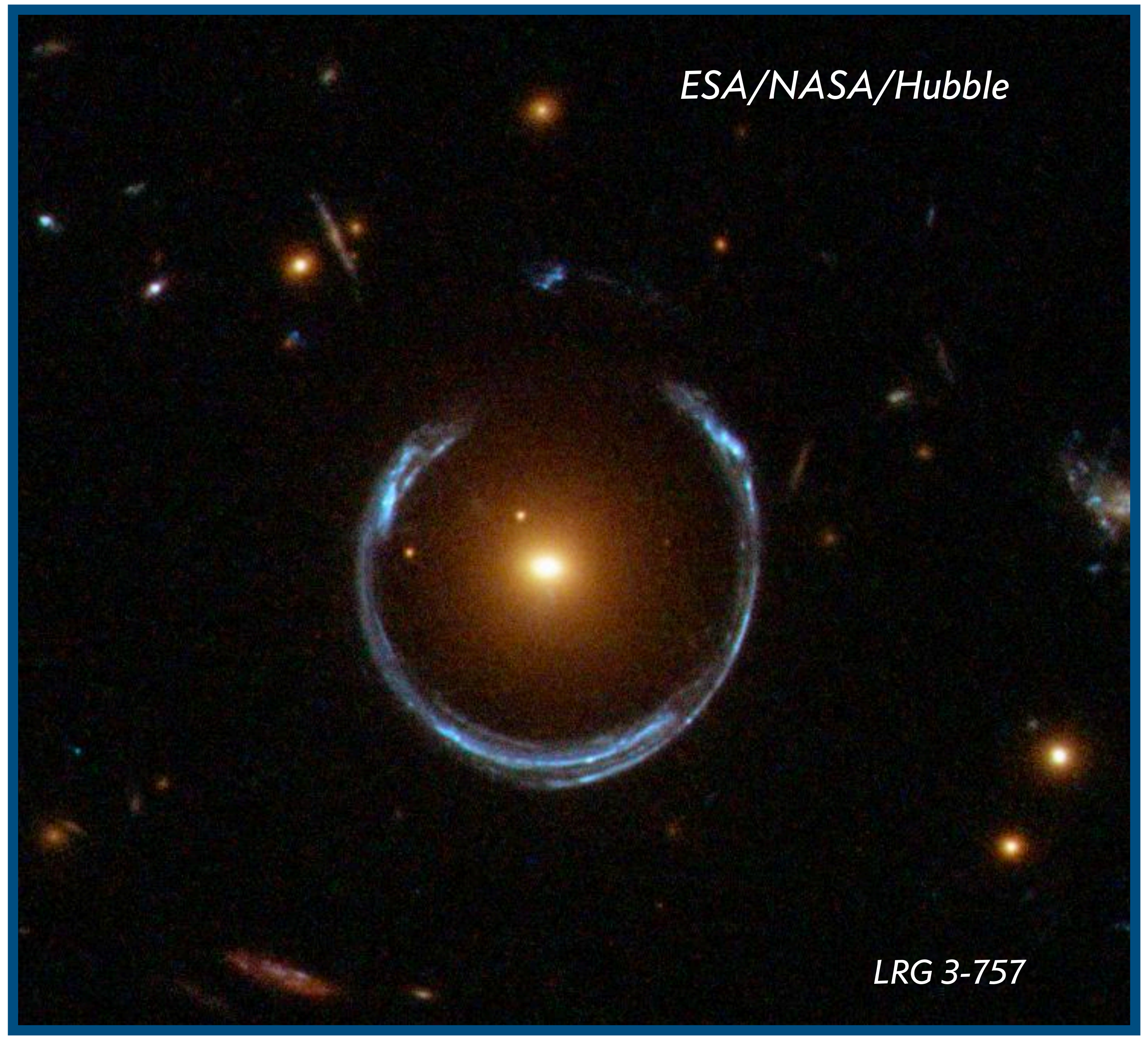
More robust, with fewer assumptions

Example: Gravitational lensing

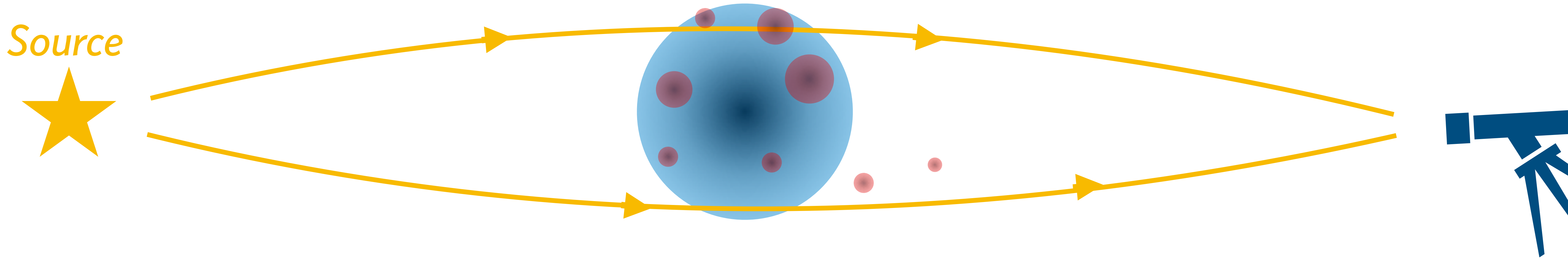


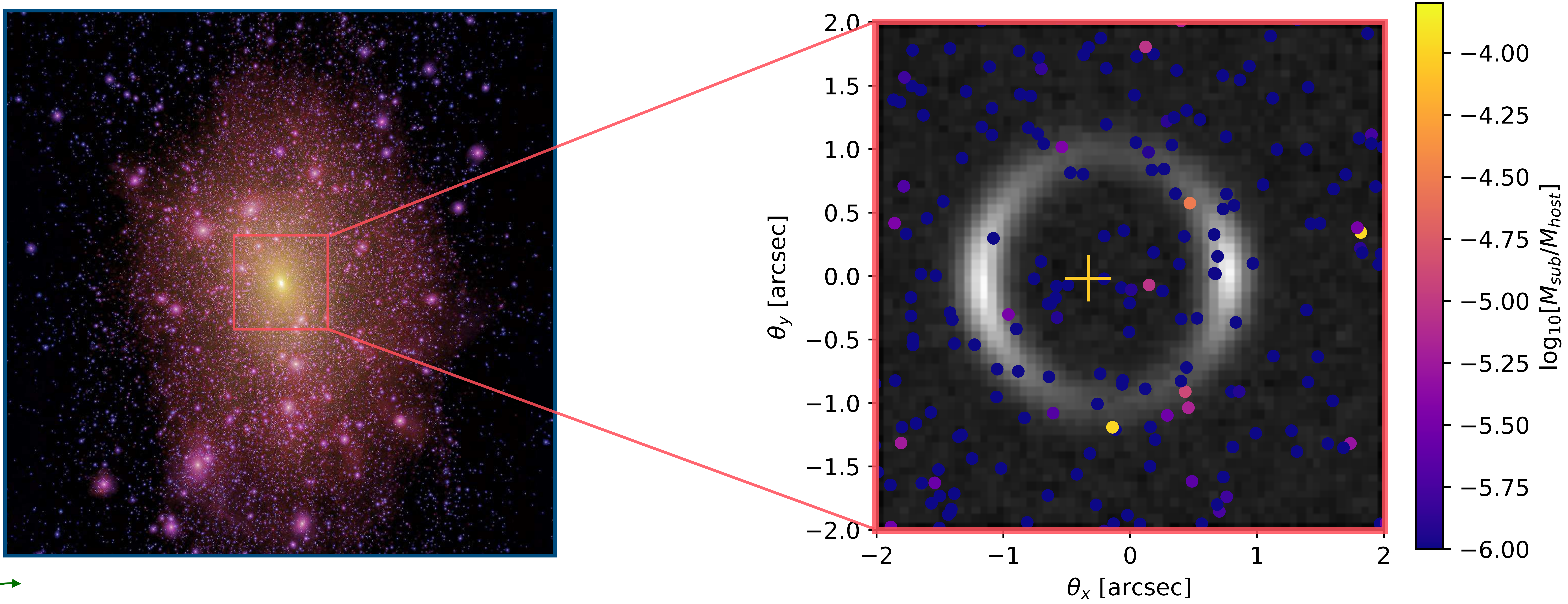


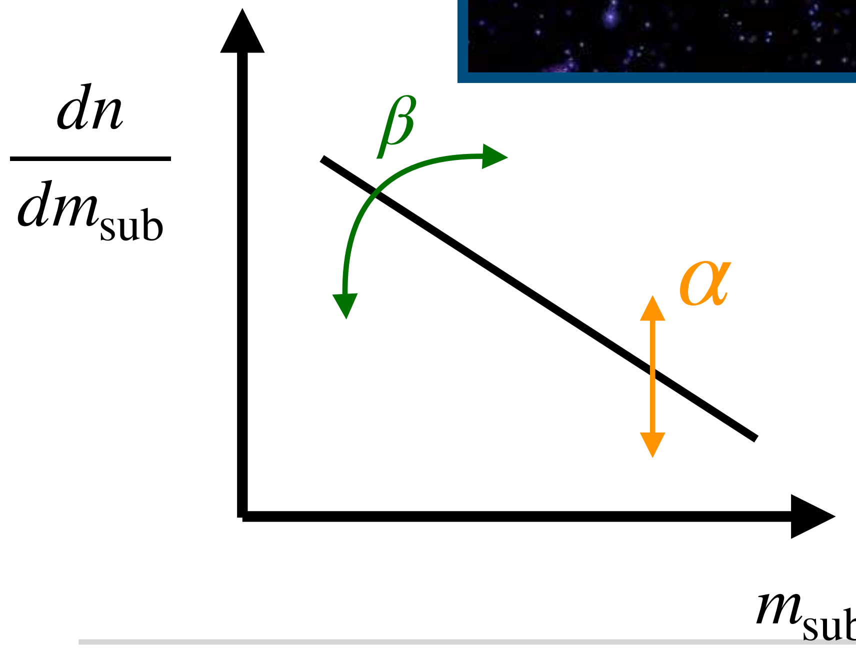
Intervening mass causes a deflection in light from a background source



Strong lensing: effect of subhalos

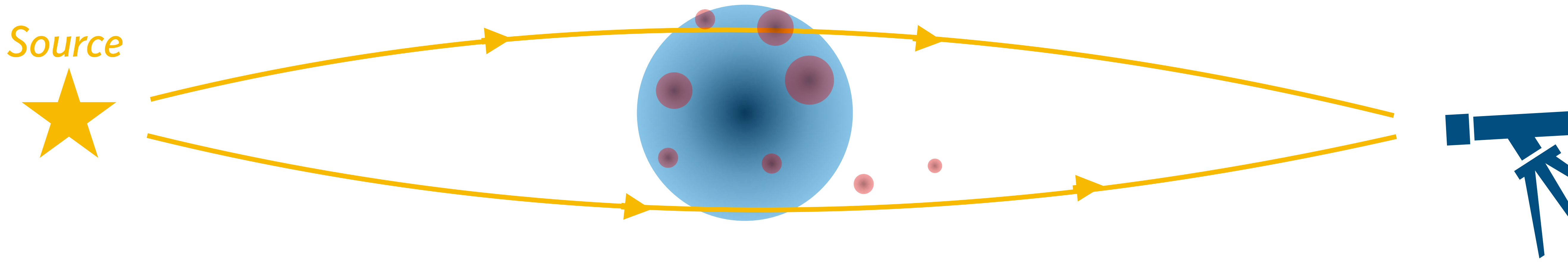




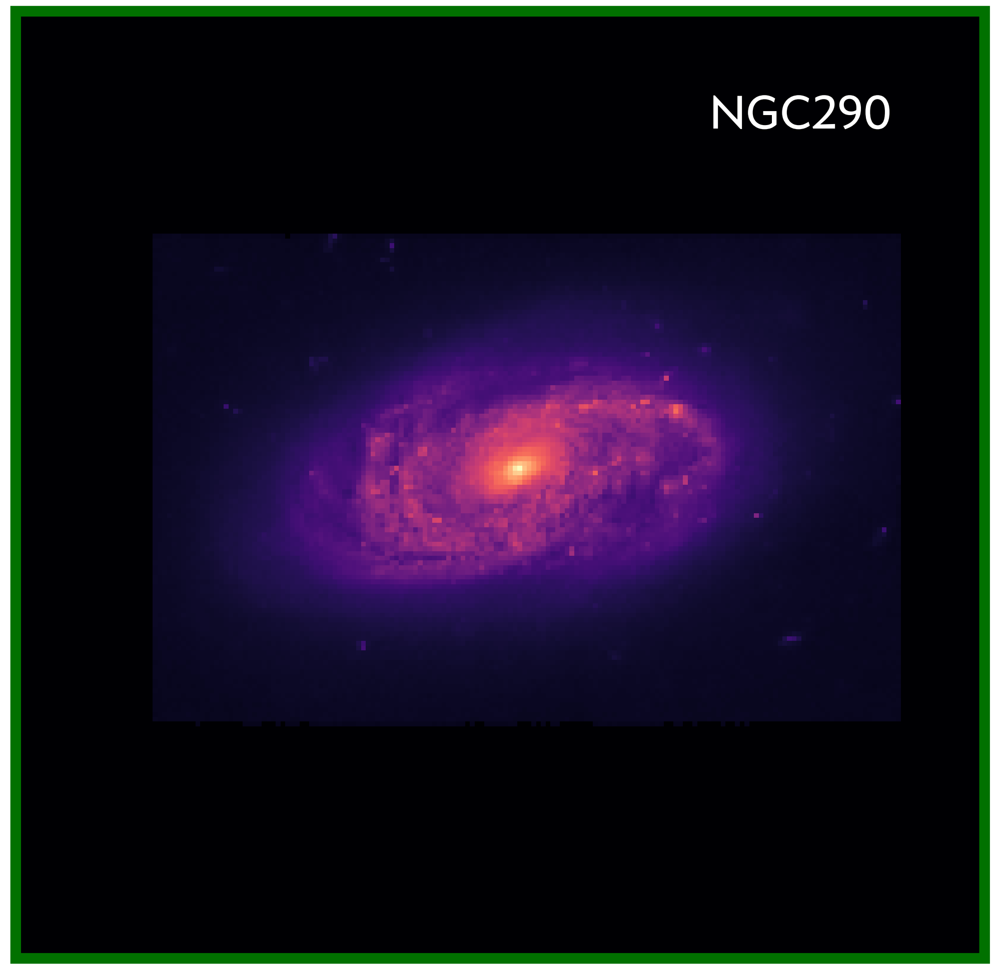


Subhalos causes *percent-level shifts* in strongly lensed images

A challenging inference problem



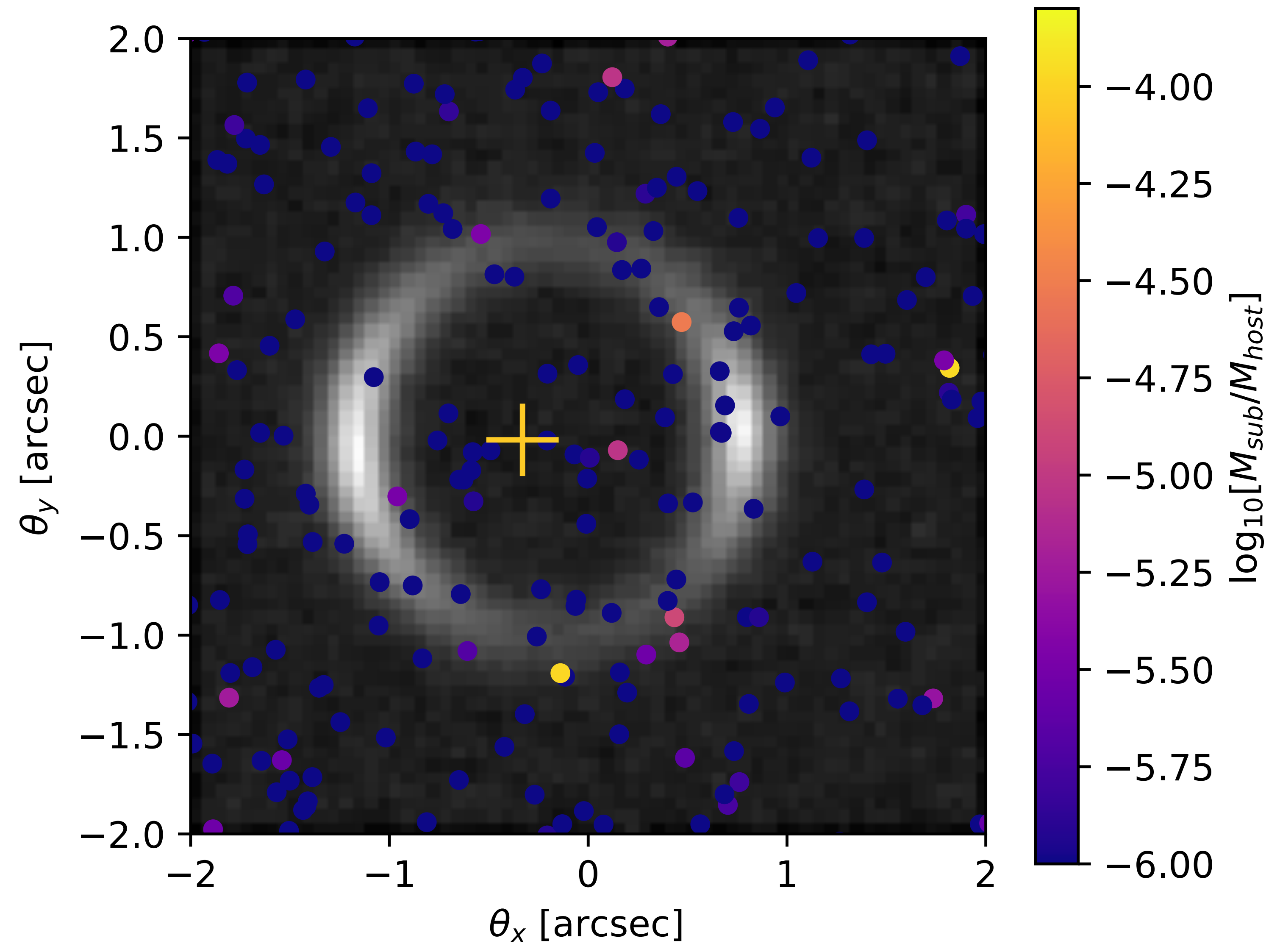
Complex background model



+ host lens

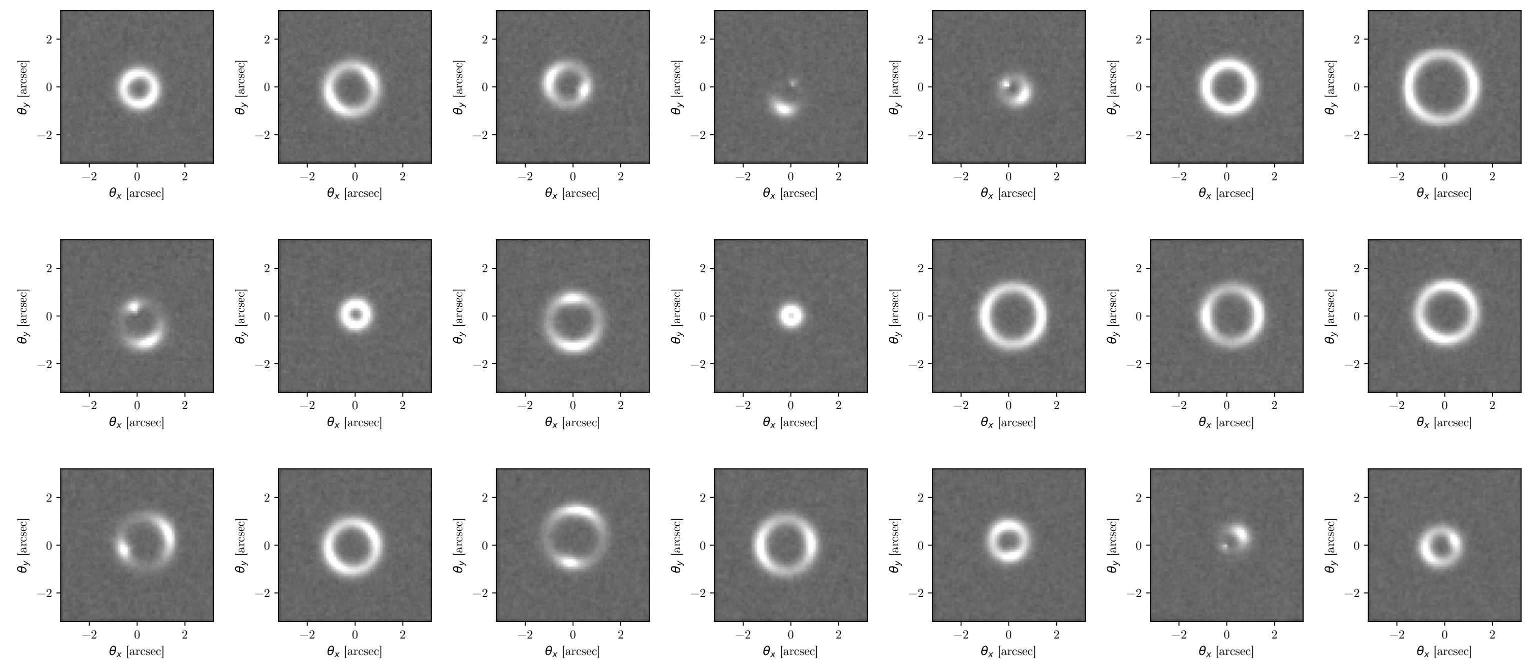
$$\{z_{\rm src}, z_{\rm lens}\}$$

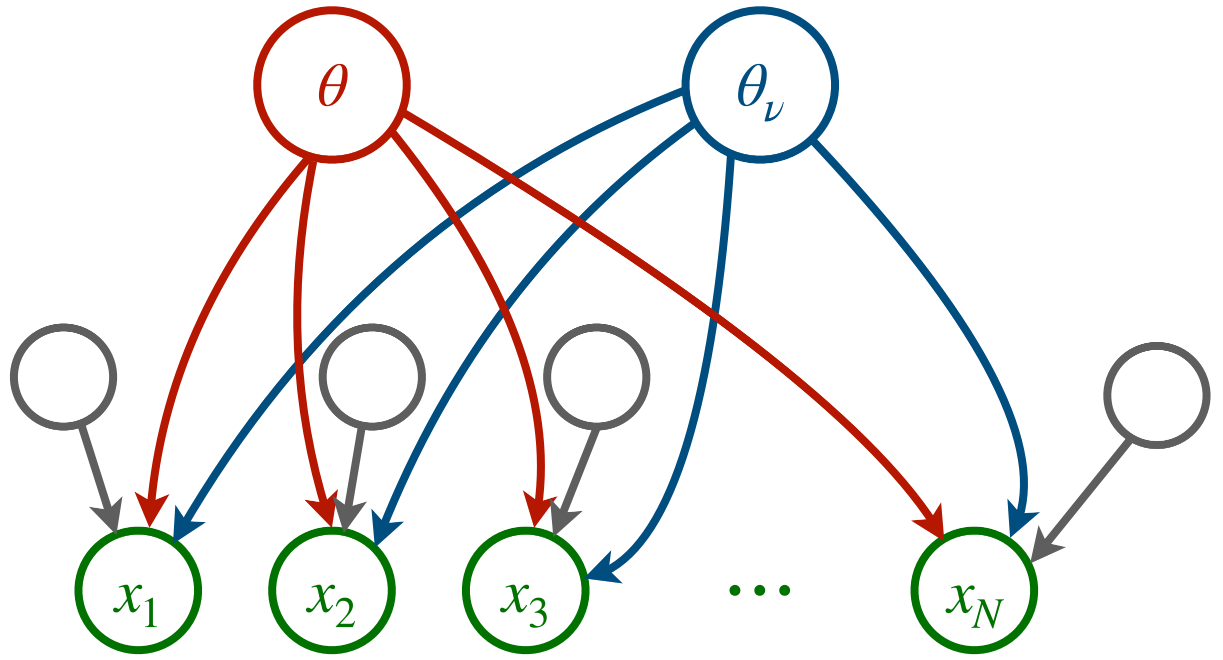
High-dim signal latents



$$\{\vec{r}_{{\rm sub},i}\,, M_{{\rm sub},i}\}_{i=1}^{N_{\rm sub}}$$

Hierarchical structure

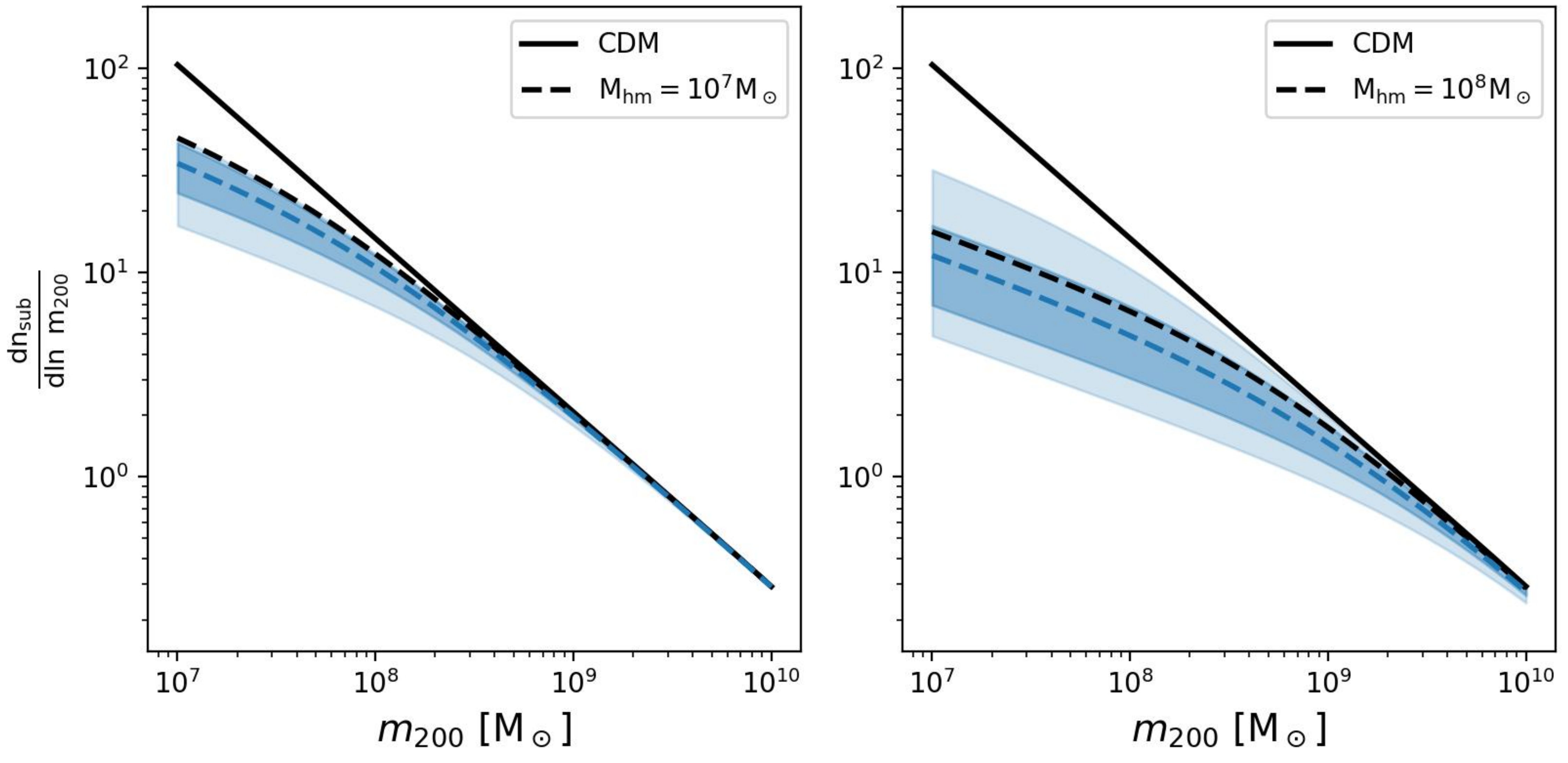




Analyzing an ensemble of gravitational lenses

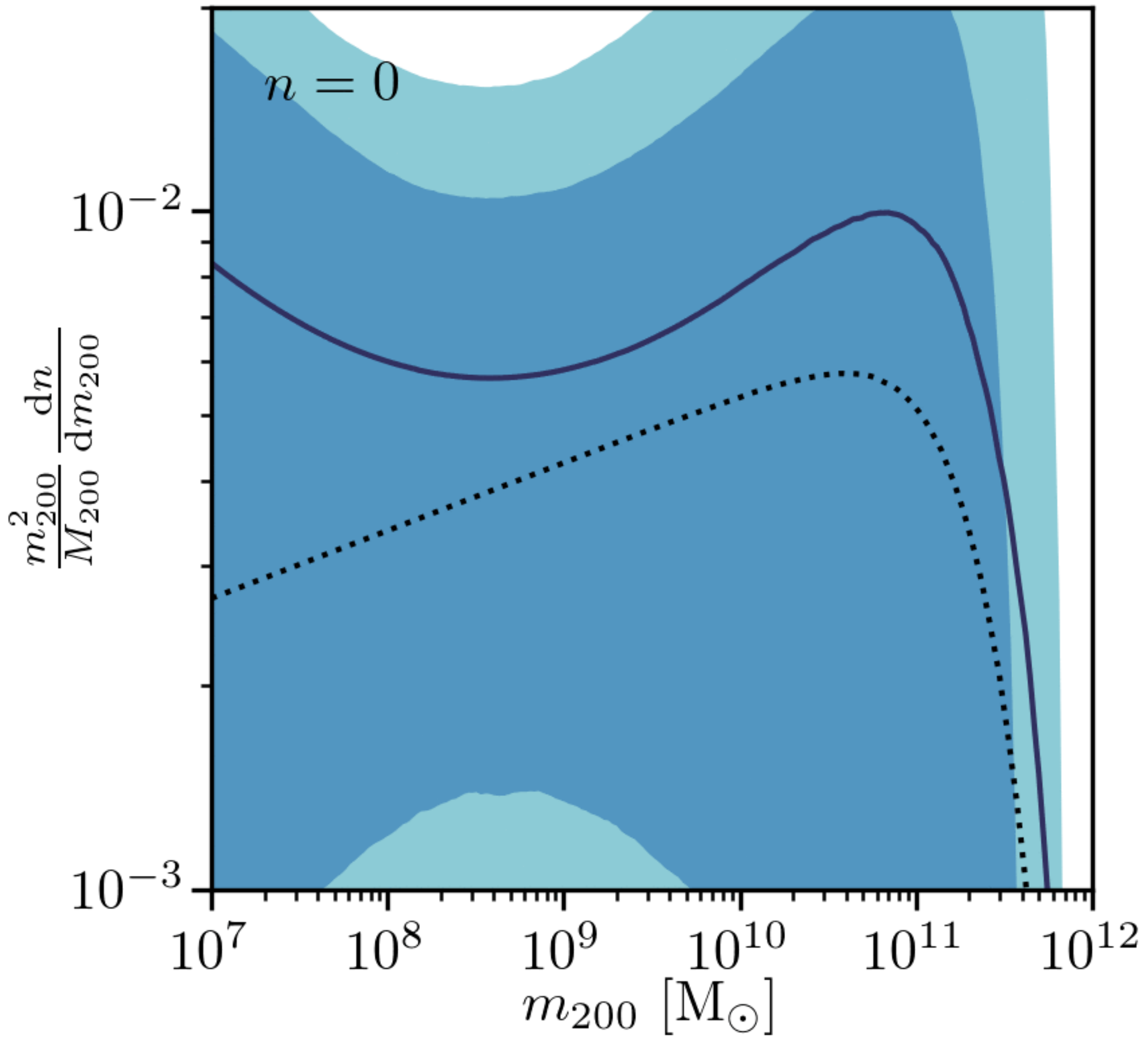
SM*, Brehmer*, et al [ApJ 2019]

Estimating warm dark matter mass



Lens sample

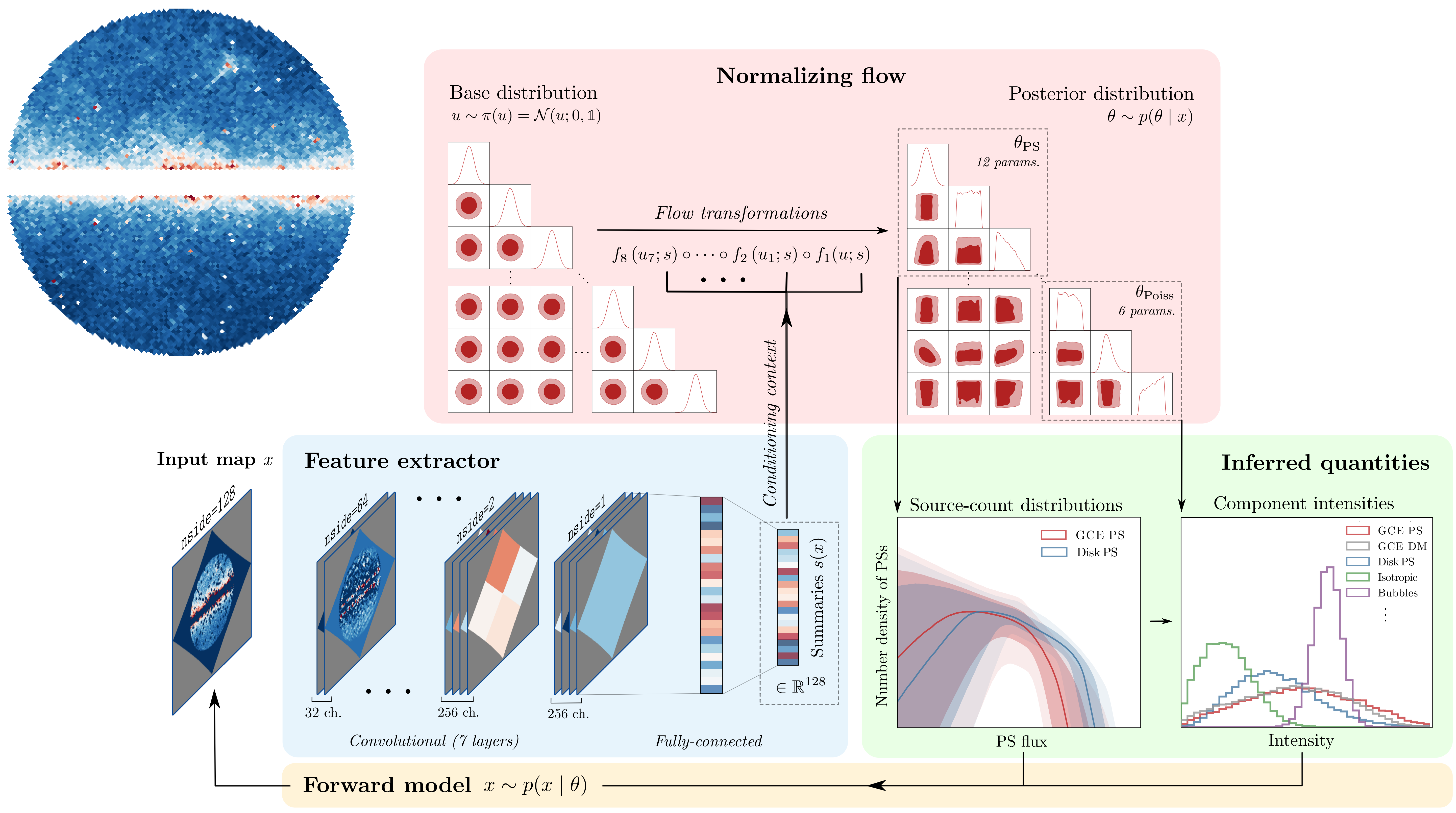
Mass function posterior

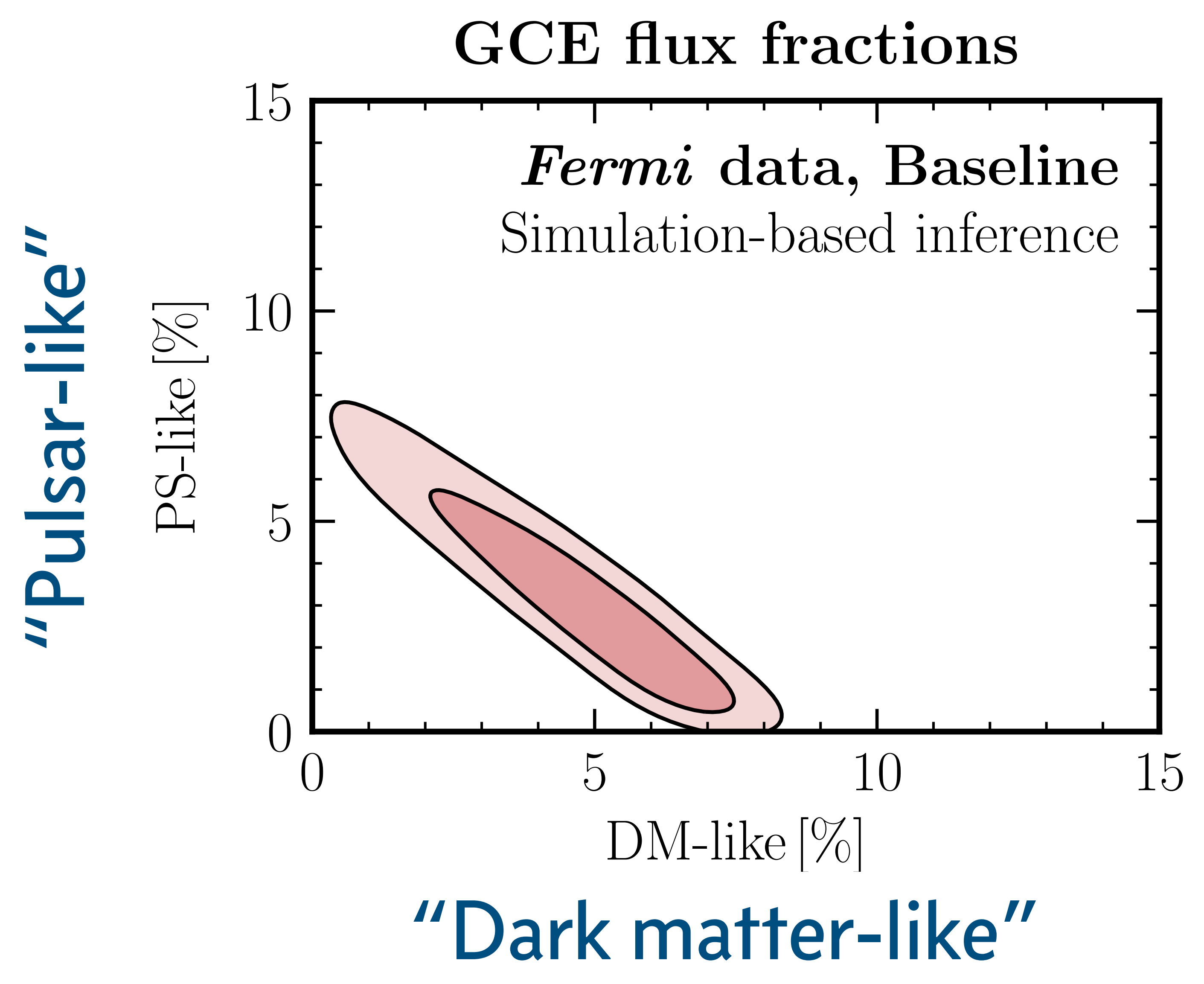


Anau Montel et al [MNRAS 2022]: Warm DM mass inference
Coogan et al [NeurIPS ML4PS 2020]: Targeted inference
Wagner-Carena et al [ApJ 2023]: Inference using realistic background galaxies
Wagner-Carena et al [2024]: Targeted population-level inference
Coogan et al [2022]: Effect of perturber populations

Example: Weighing in on the Galactic Center Excess

SBI pipeline for characterizing Excess signal including pulsar contribution





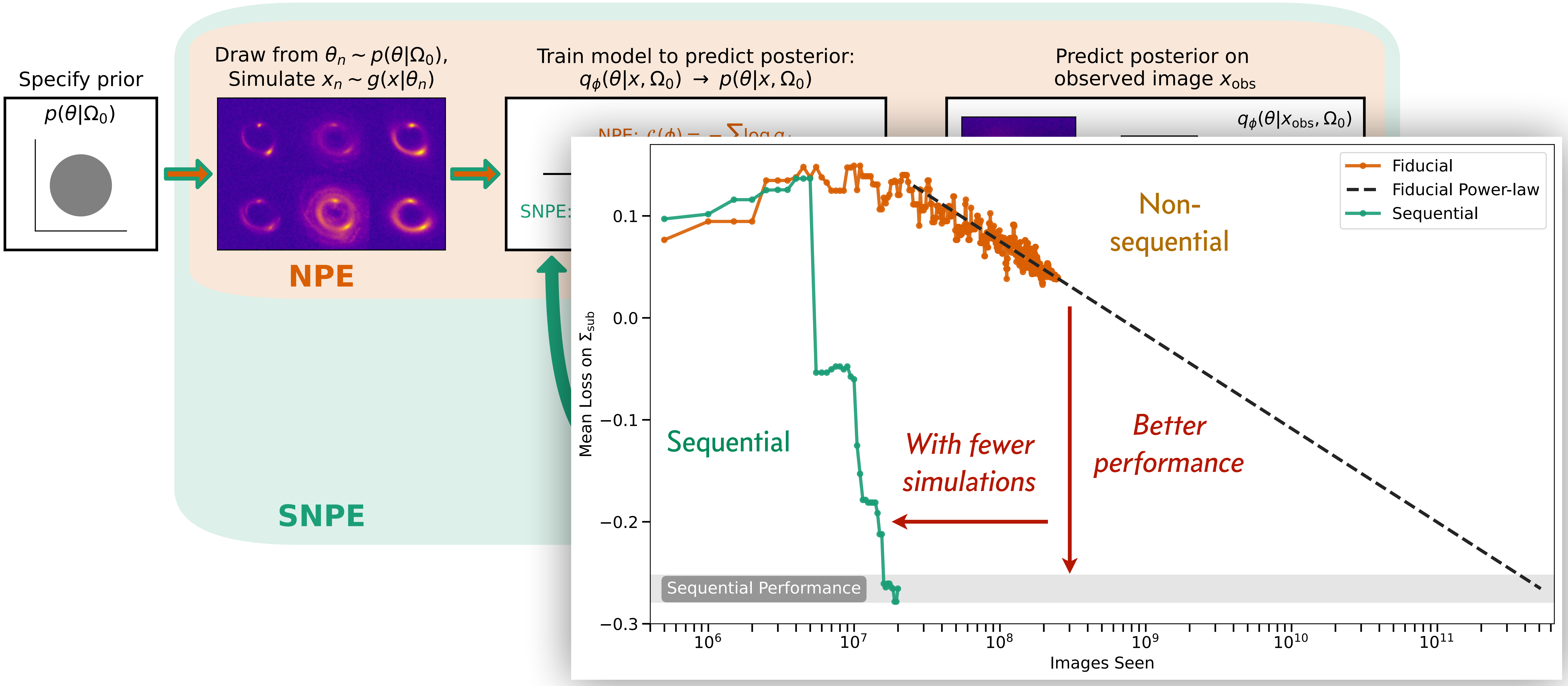
More robust to background mis-modeling than traditional (counts-PDF) approach —
But more work to be done, e.g.

- Data-driven background modeling
- Exploiting spectral information
- …

Recent trends and challenges: Sequential methods

Wagner-Carena et al [2024]

Specialize to particular data at the cost of amortization

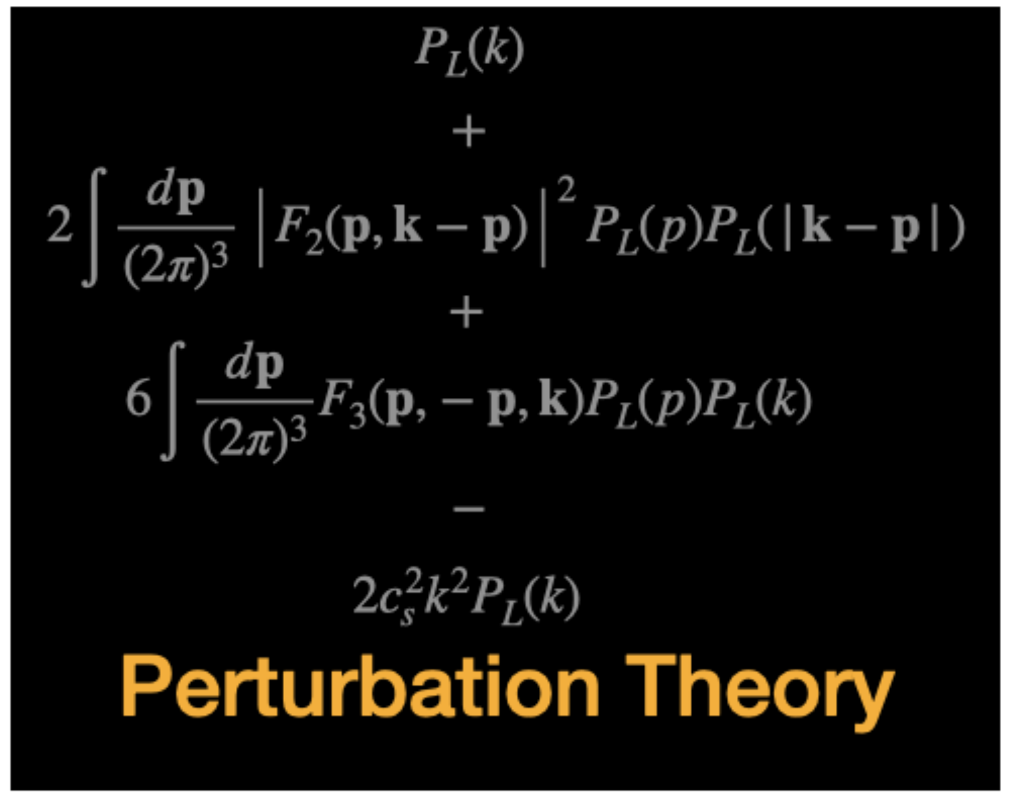


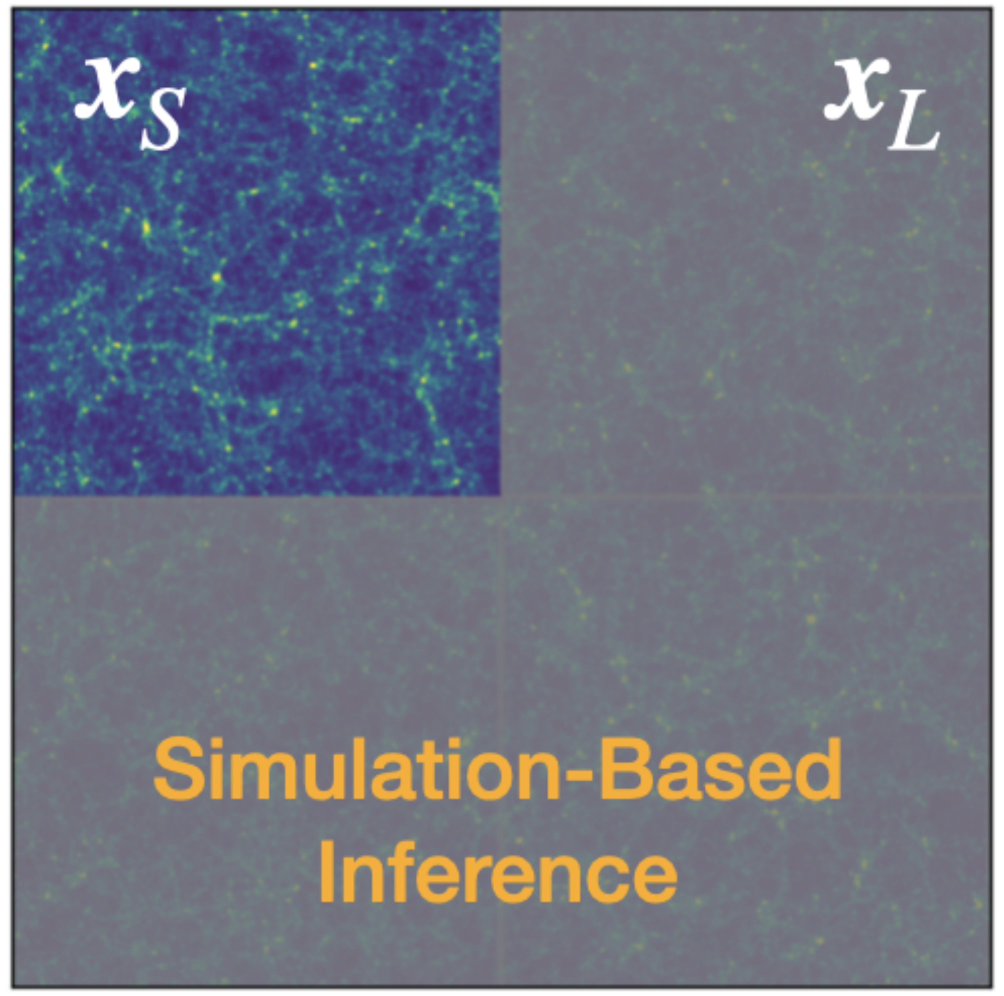
Recent trends and challenges: Hybrid methods

Modi & Philcox [2024]

Inject domain knowledge where possible for better robustness and simulation-efficiency

Model large / mildly non-linear scales with perturbation theory



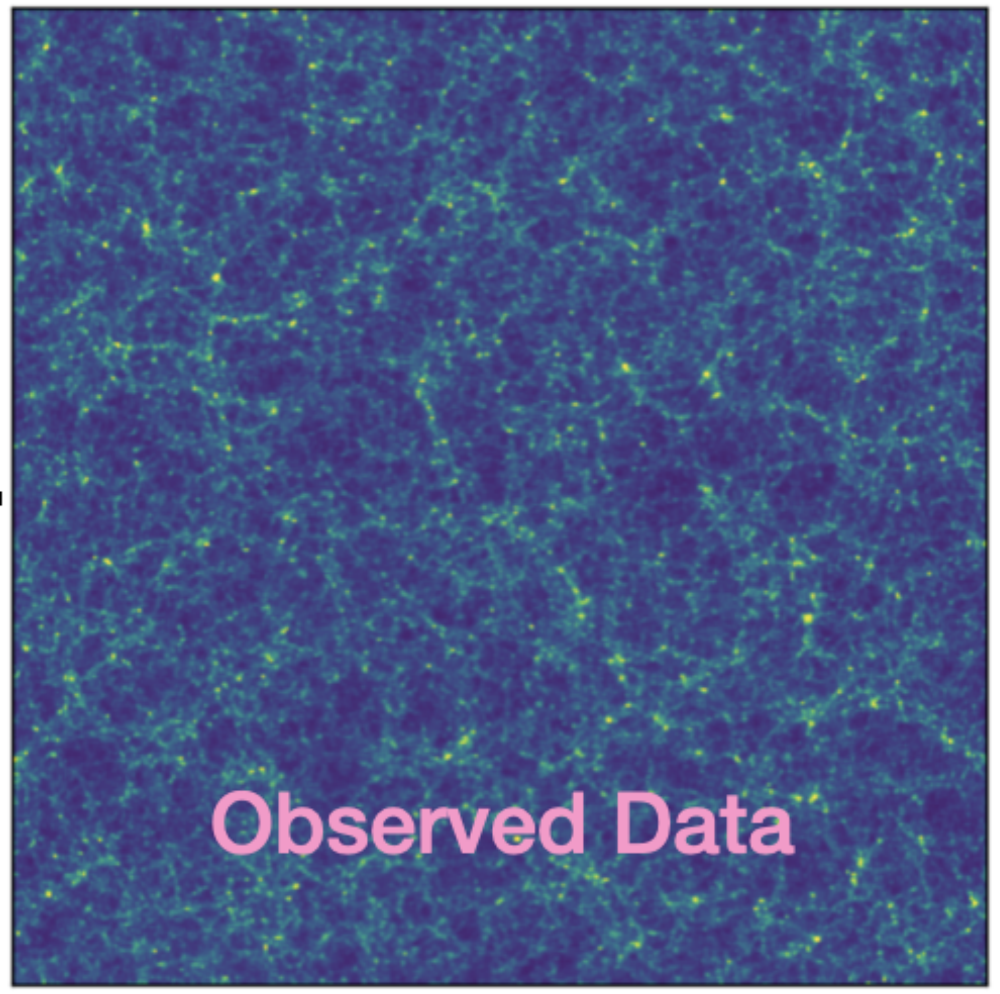


Large scales are modeled analytically

$$x_L : P(k), B(\{k_i\}), \ldots \rightarrow P(x_L \mid \theta)$$

$$P(x_L, x_S \mid \theta)$$

Joint likelihood



$$x_S : P(k), \mathrm{WST}, \ldots \rightarrow P(x_S \mid x_L, \theta)$$

Small scales are modeled numerically

Model small (non-linear) scales with SBI

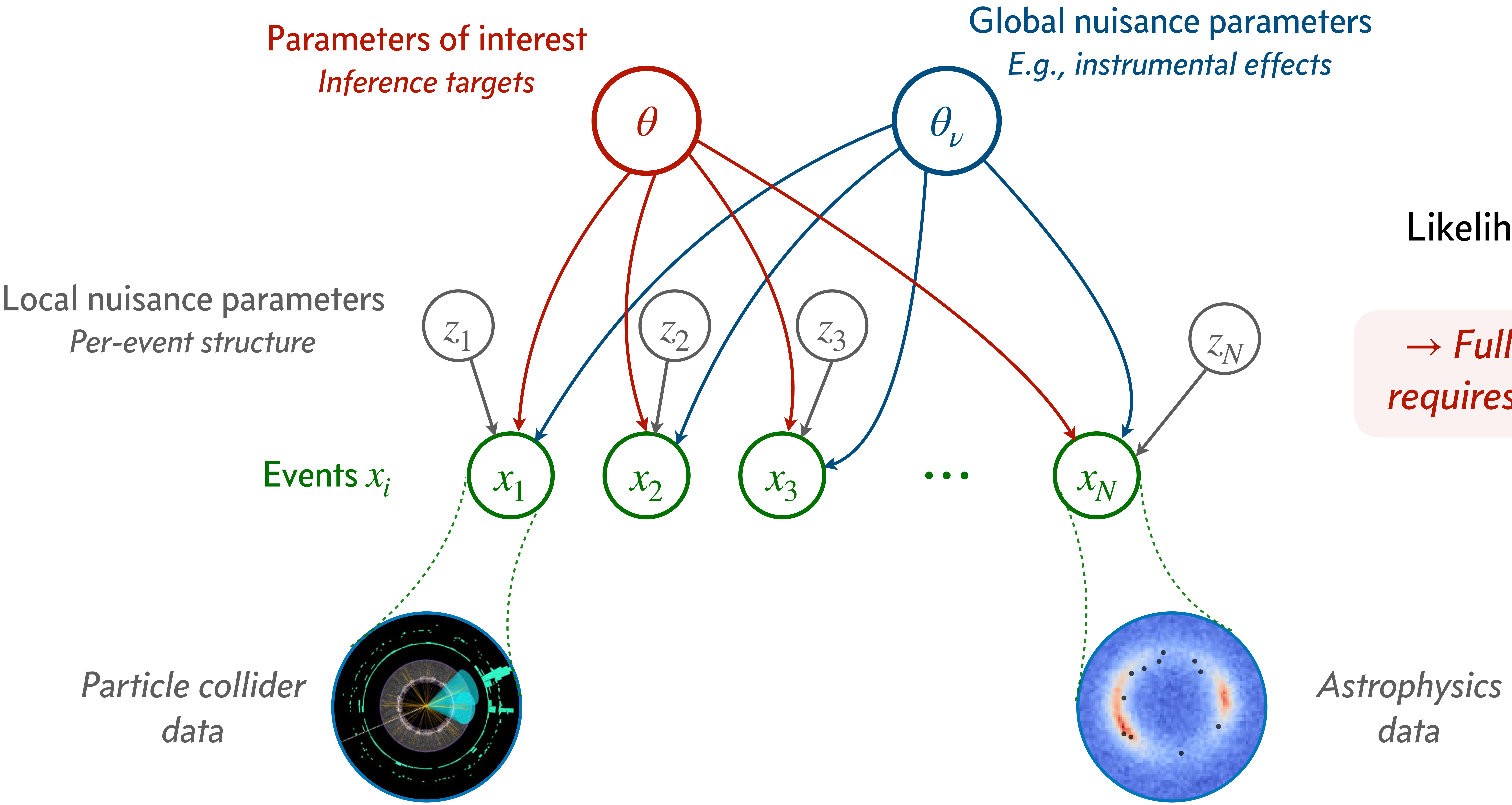
Also: Ivanov et al + SM [2024]

Slide: Philipp Windischhofer

Heinrich*, SM* et al [TMLR 2024]

Recent trends and challenges: Hierarchical models

Many problems in astroparticle physics have hierarchical structure

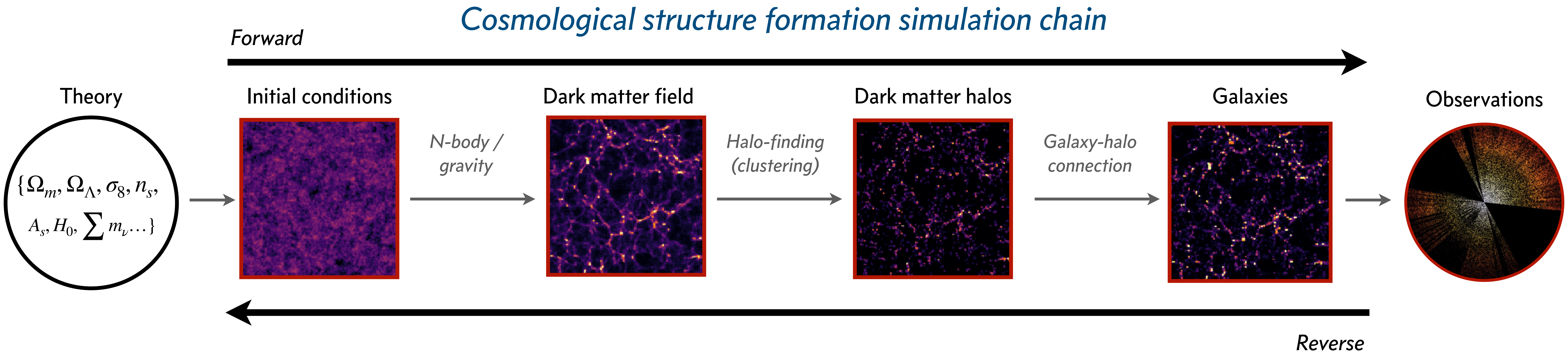


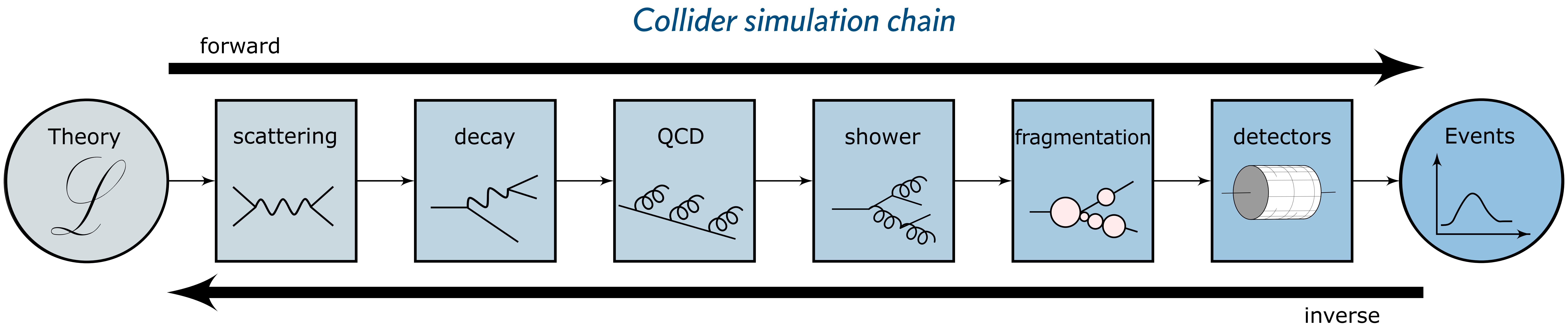
Likelihood doesn't factorize over events

→ Fully capitalizing on data requires hierarchical approach

Collider and astro/cosmo — many commonalities!

AI bringing communities together! Common goals + transferable methods



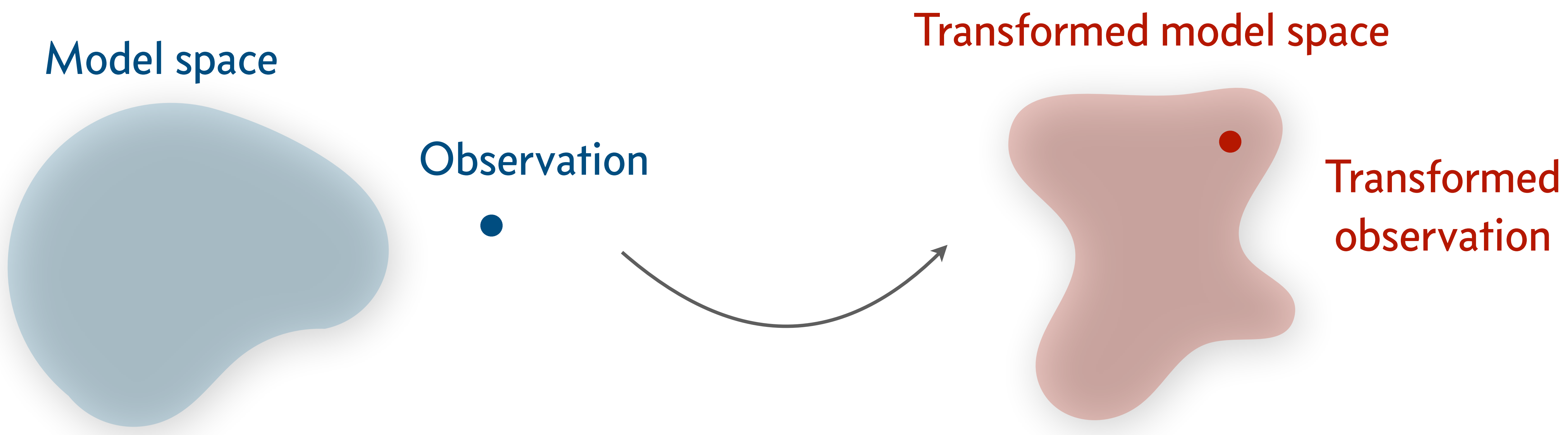


Butter, Plehn et al [SciPost2023]

Recent trends and challenges: Model misspecification

Using simulations to leverage more information can be a double-edged swords: Methods can be sensitive to aspects of the simulation that are mis-specified, which would otherwise be "washed over" when using summaries

- Methods to detect the {*degree/source*} of model mis-specification
- Methods to *correct* for model mis-specification

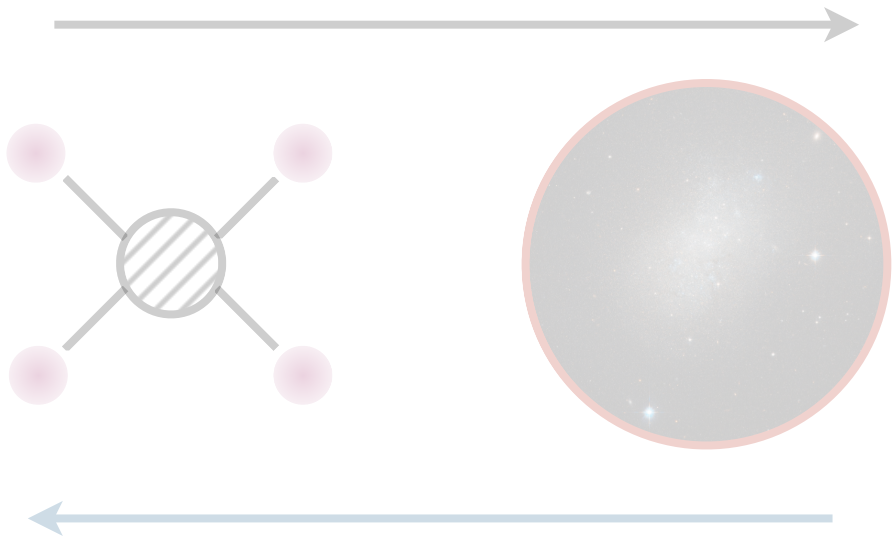


e.g., Huang et al [NeurIPS 2023], Gao et al [NeurIPS 2023]

+ methods towards better calibration, simulation efficiency, high-dim posteriors…

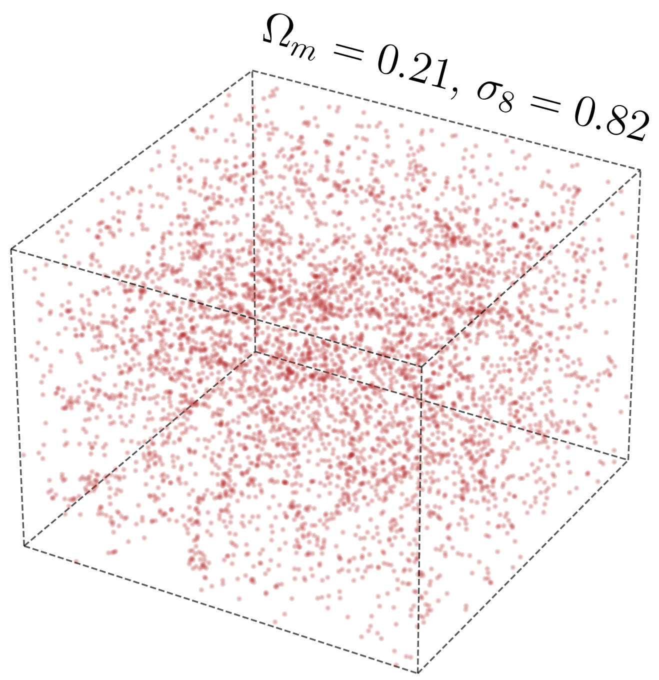
See https://github.com/smsharma/awesome-neural-sbi!

Outline



Simulation-based inference

Inverting complex physical simulators



Generative modeling

Capturing the distribution of complex data for emulation and inference

∇_θ

Differentiable and probabilistic programming

Specifying models with autodiff capabilities and enabling flexible inference

Generative modeling

Generative models are simulators of the data

Goal: learn a probability distribution $p_\vartheta(x)$ that is as close as possible to the true underlying data distribution $p(x)$

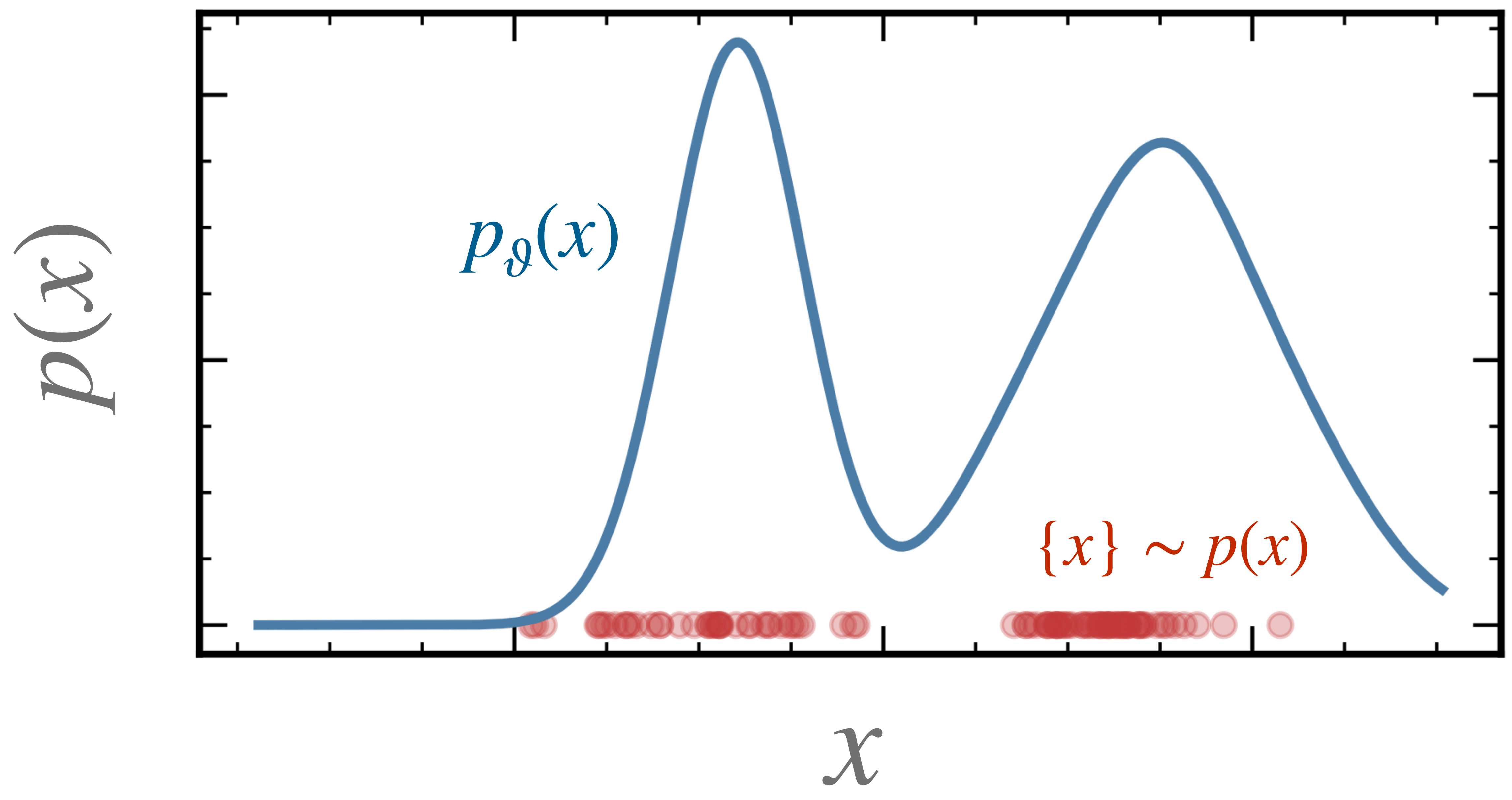
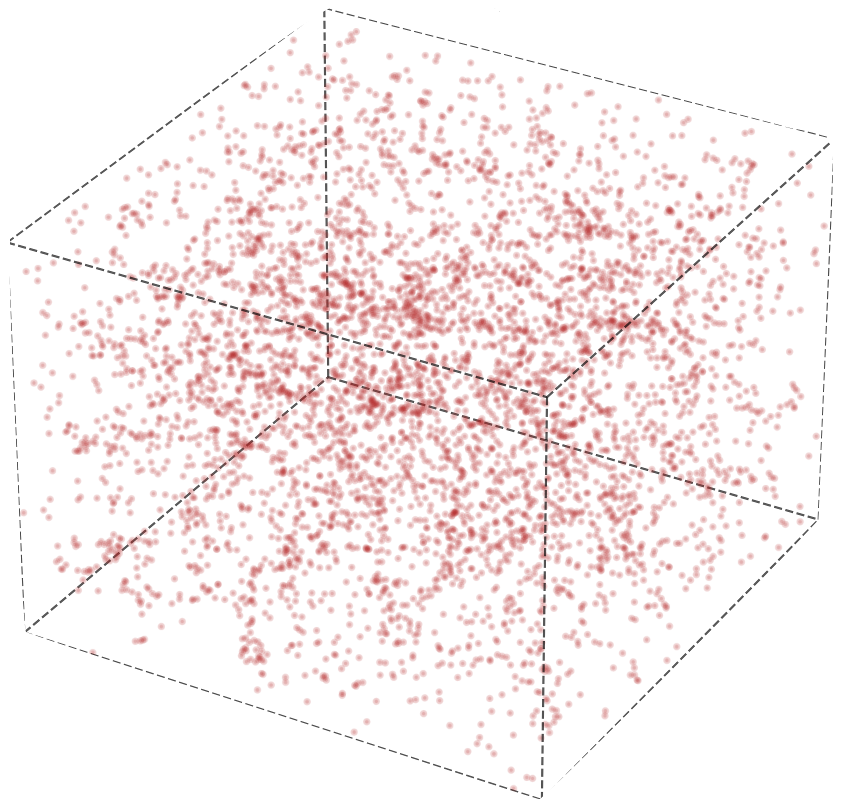
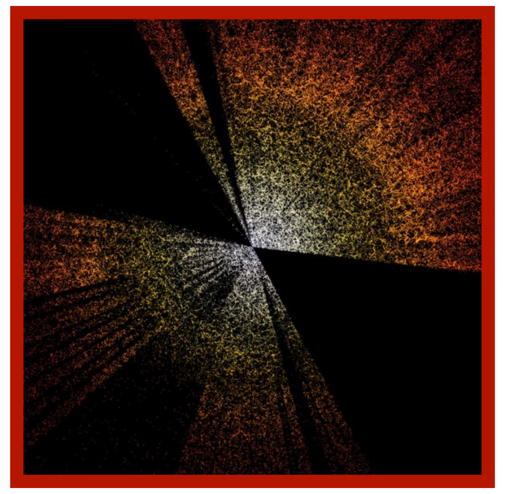


Image generation
(Midjourney v5)

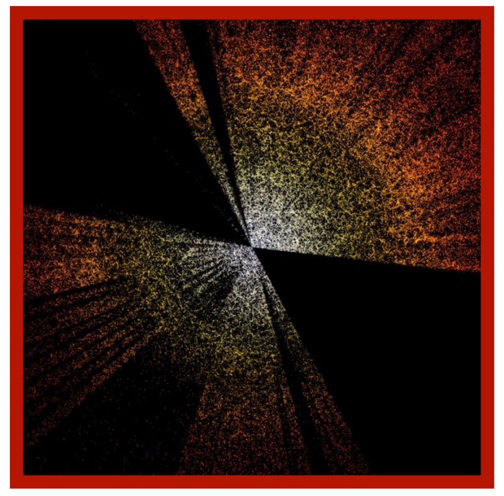
Generative modeling — capabilities

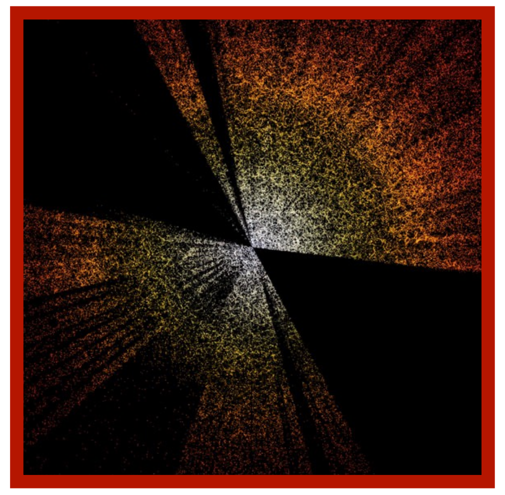
Emulation

 $\sim p($  $|$ Cosmology $)$

Model evaluation

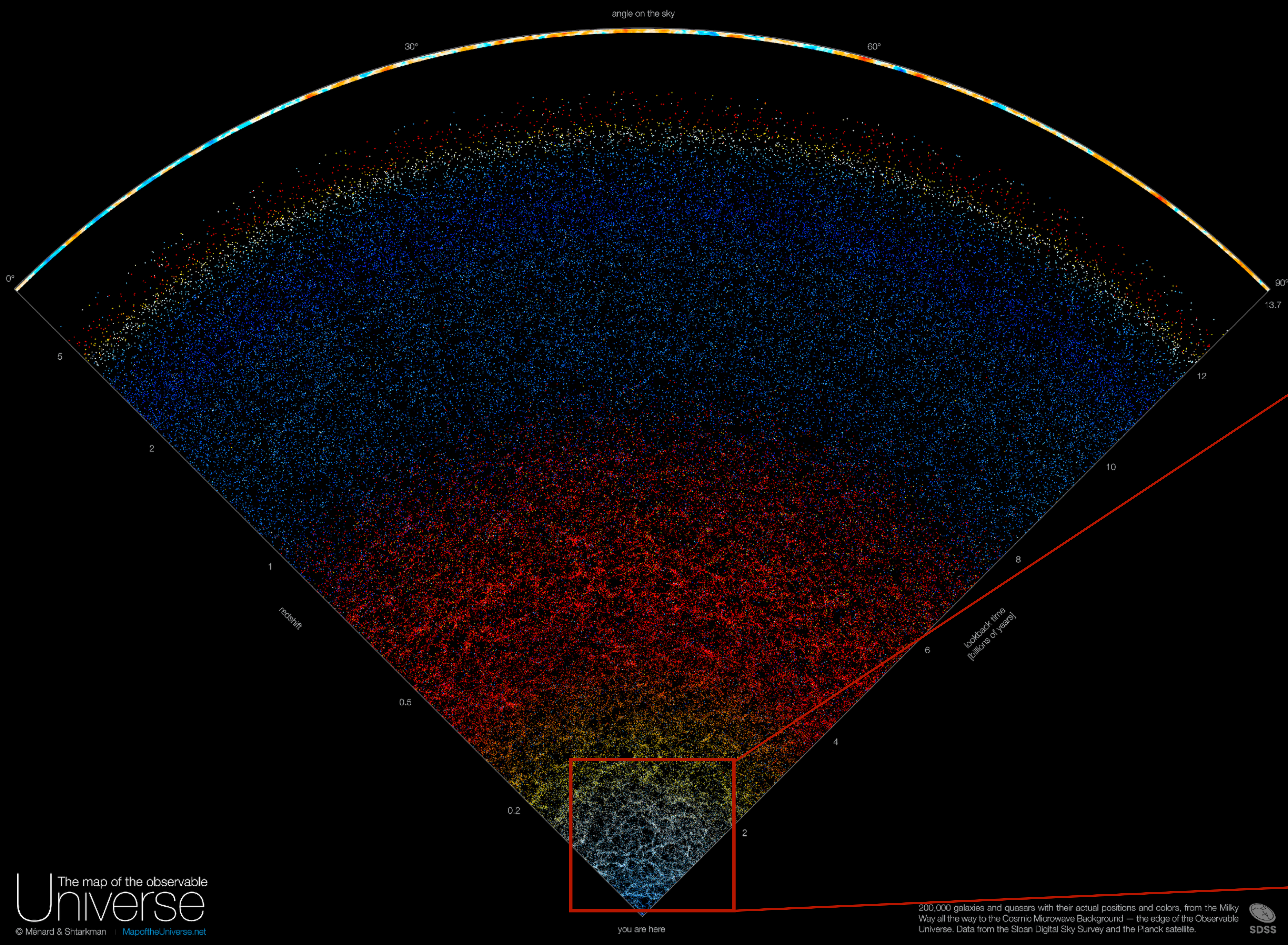
Parameter inference

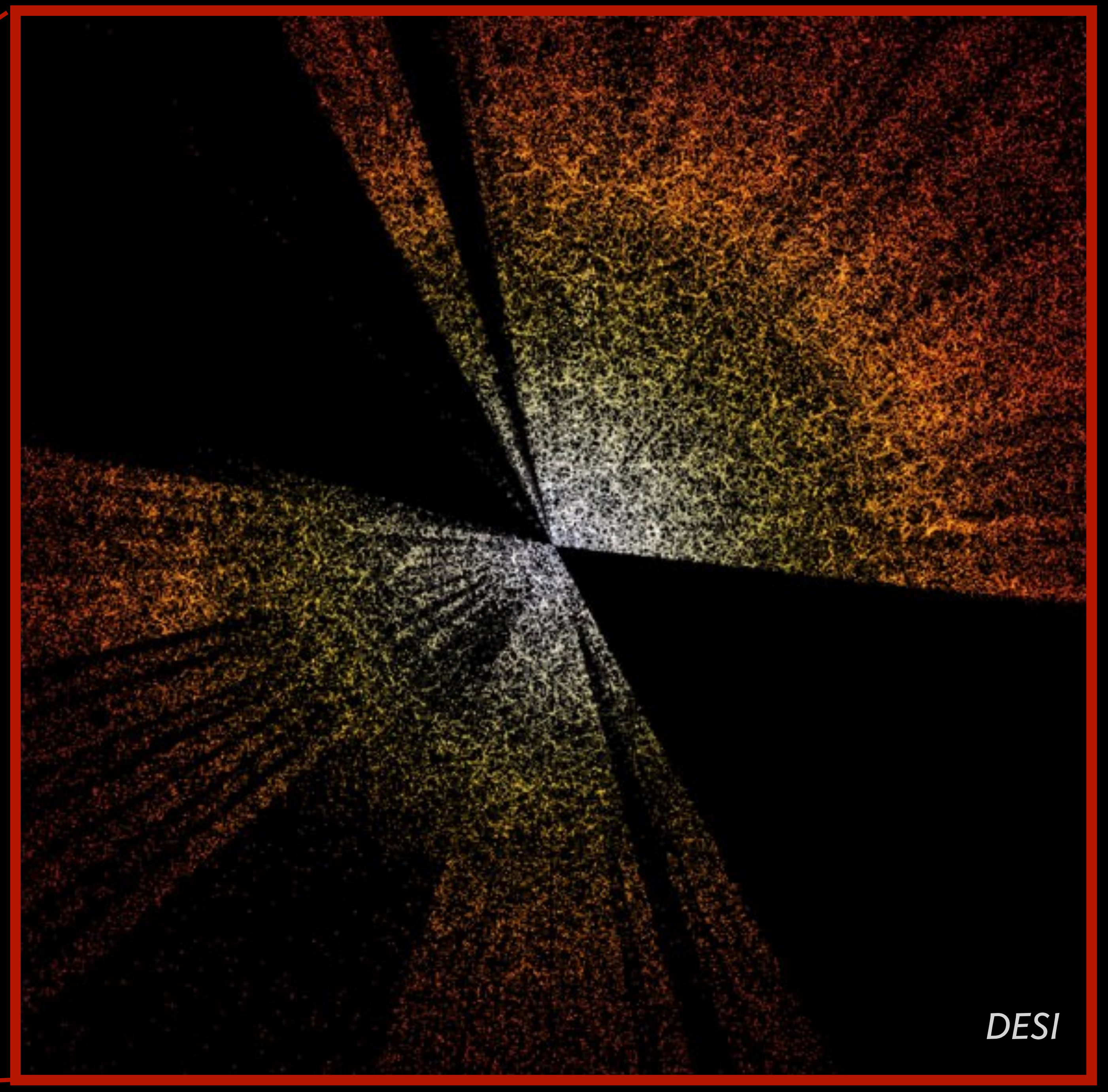
$p($ Cosmology $|$  $)$

$p($  $)$

Example: galaxy clustering

Galaxy clustering: the statistical distribution of galaxies in the Universe





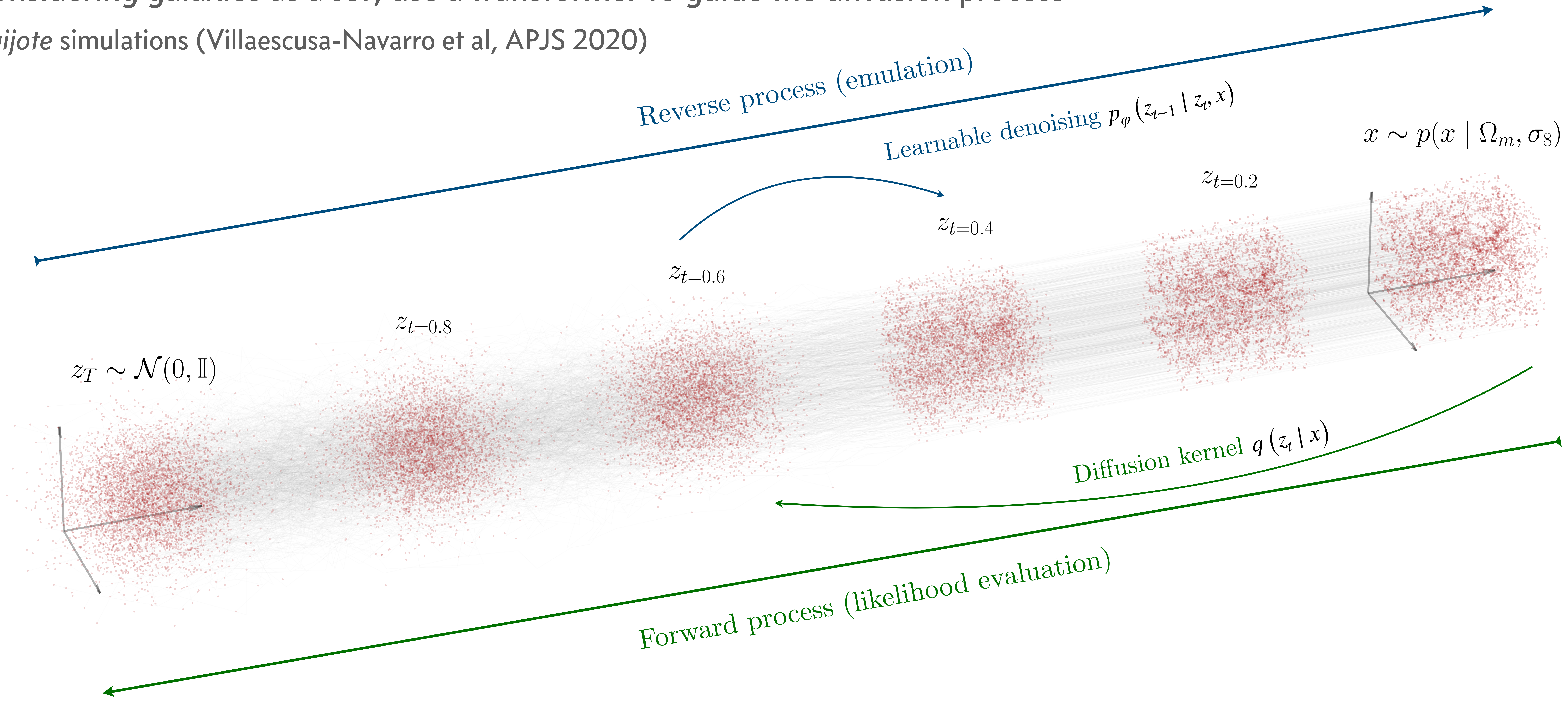
Transformer-guided diffusion on galaxies

SM*, Cuesta-Lazaro*

[ICML ML4Astro 2023 Spotlight]

Considering galaxies as a *set*, use a transformer to guide the diffusion process

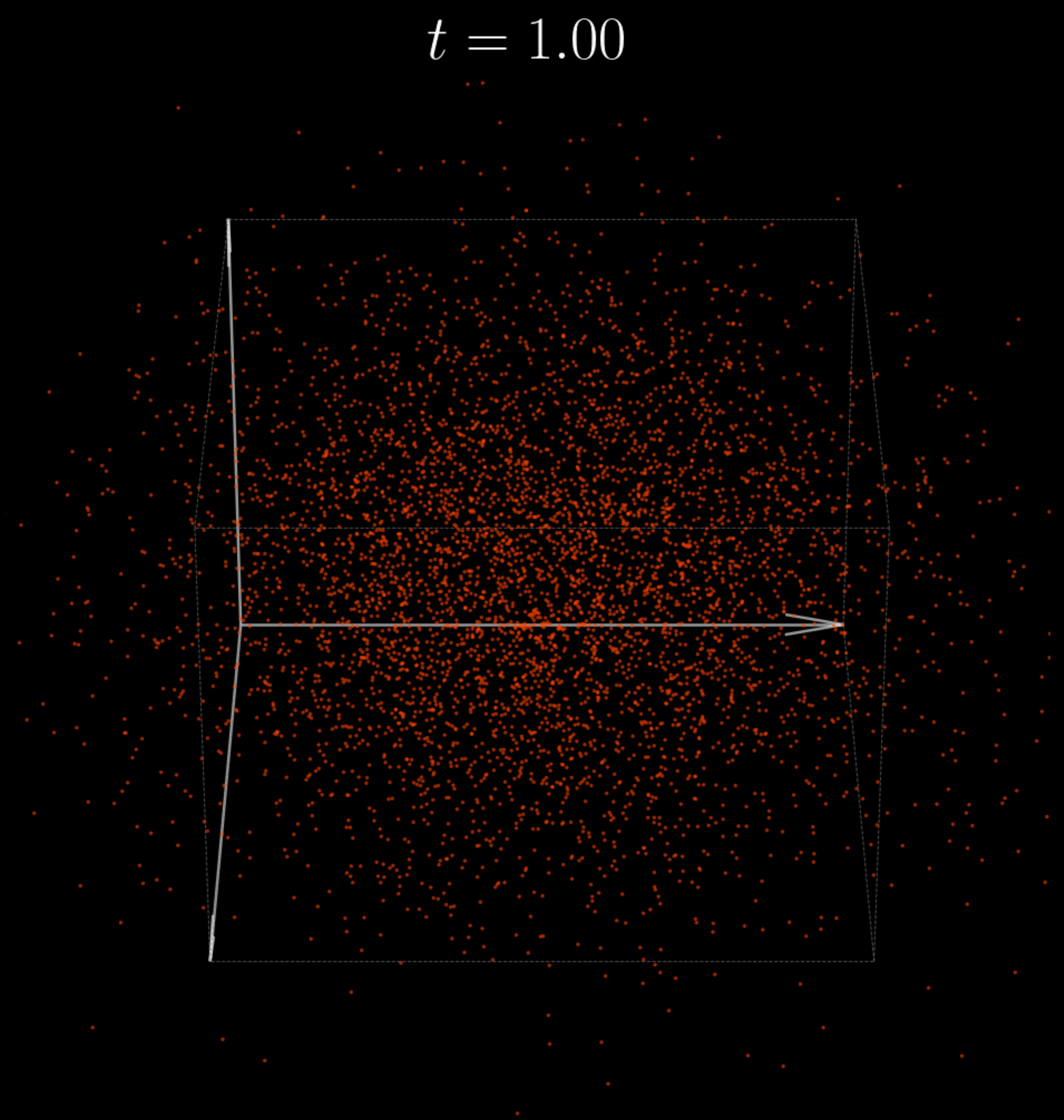
Quijote simulations (Villaescusa-Navarro et al, APJS 2020)



Diffusion on galaxies

SM*, Cuesta-Lazaro*

[ICML ML4Astro 2023 Spotlight]

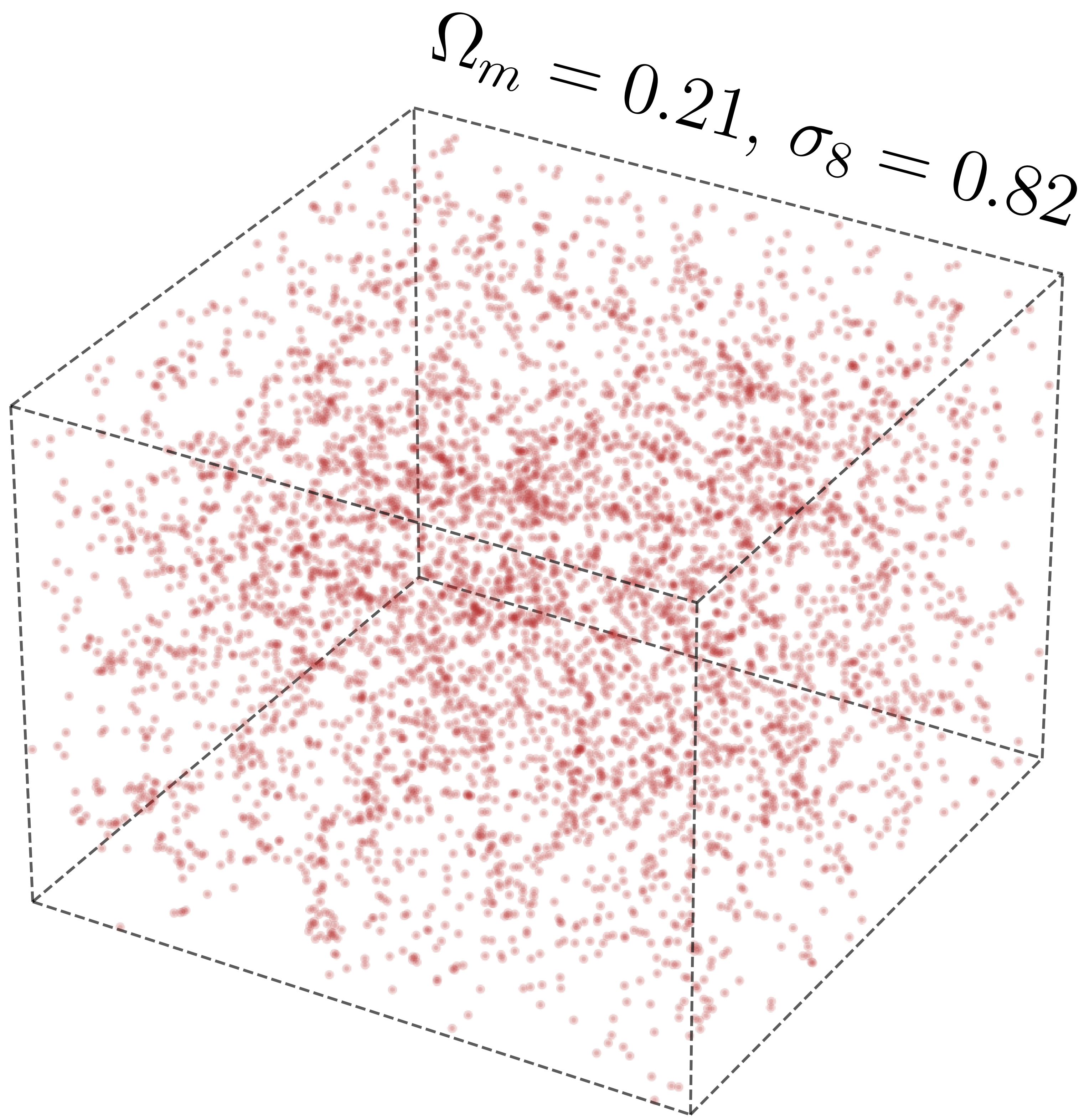


- Fast (~seconds) generation of galaxy fields
- Accurately captures cosmological dependence of generated field

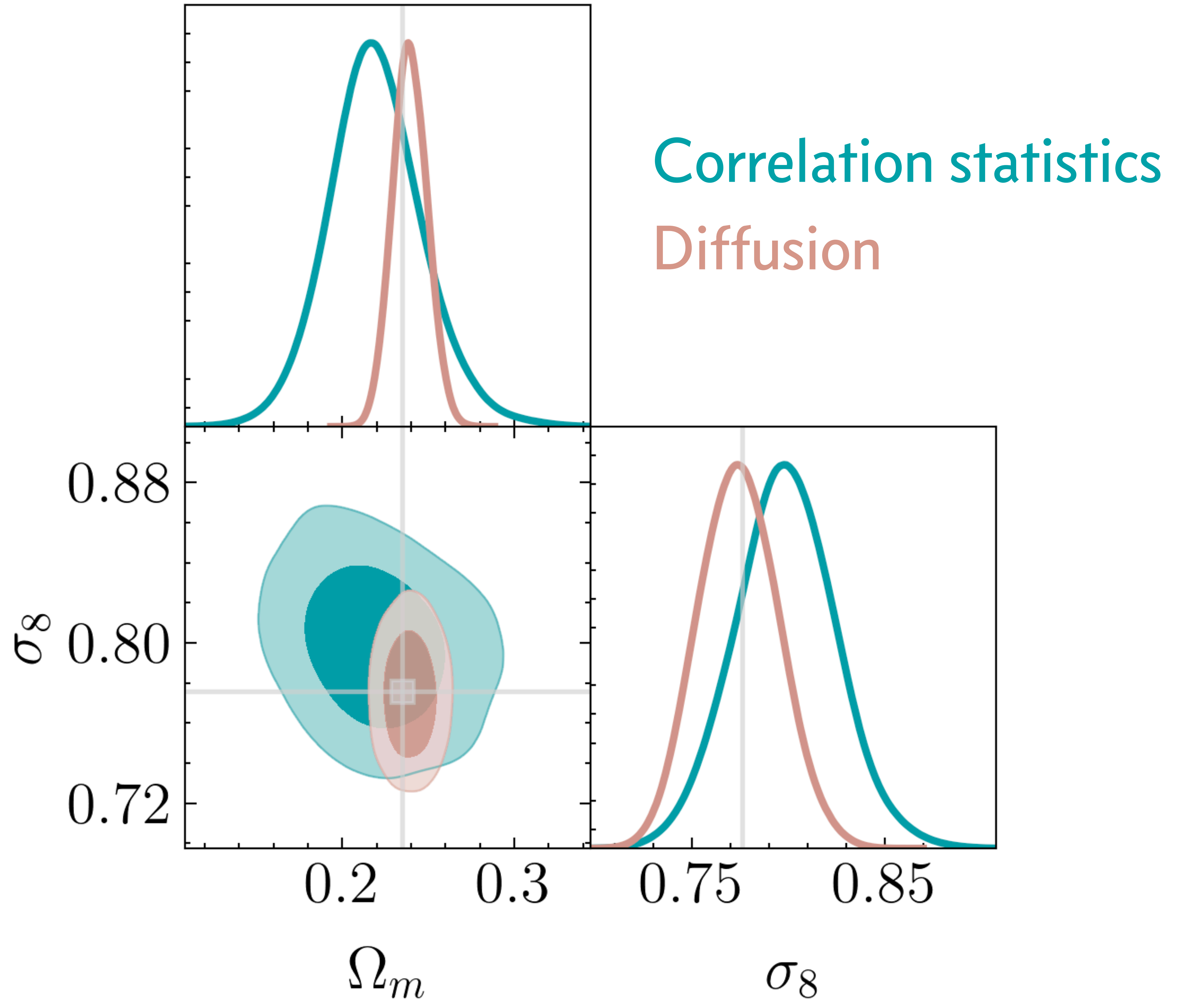
SM*, Cuesta-Lazaro*
[ICML ML4Astro 2023 Spotlight]

Parameter estimation

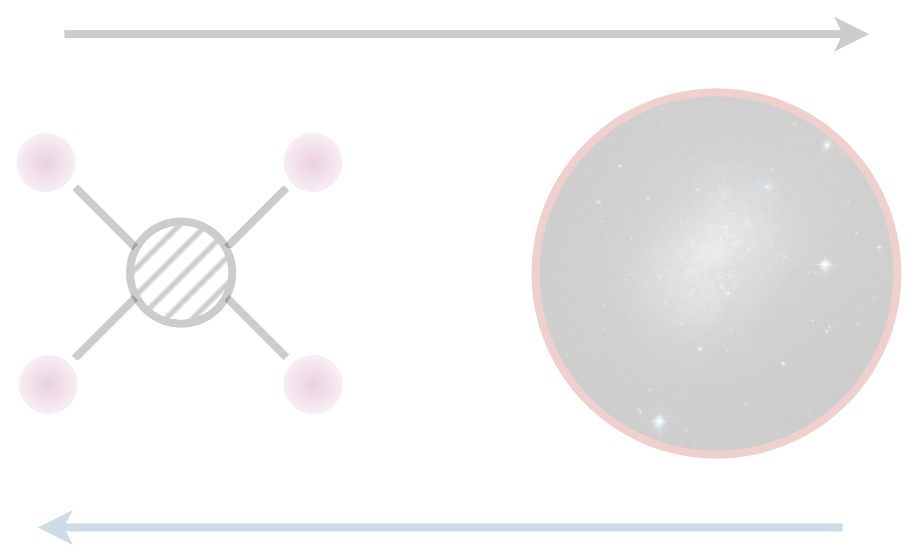
Construct differentiable likelihood $p(x \mid \theta)$ for posterior parameter inference



Inference →



Outline

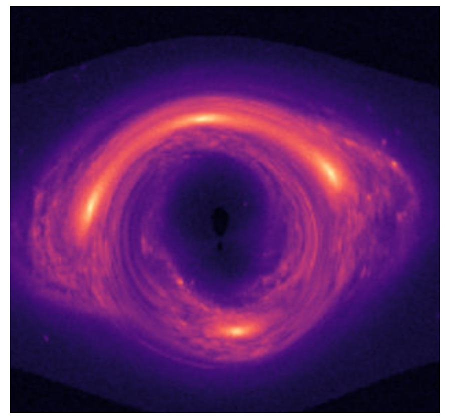


Simulation-based inference

Inverting complex physical simulators

Generative modeling

Capturing the distribution of complex datasets for emulation and inference

∇_θ 

Differentiable and probabilistic programming

Specifying models with autodiff capabilities and enabling flexible inference

Forward models / simulations

We have a access to a bunch of simulations

We can run new simulations

The simulations are fast ⚡

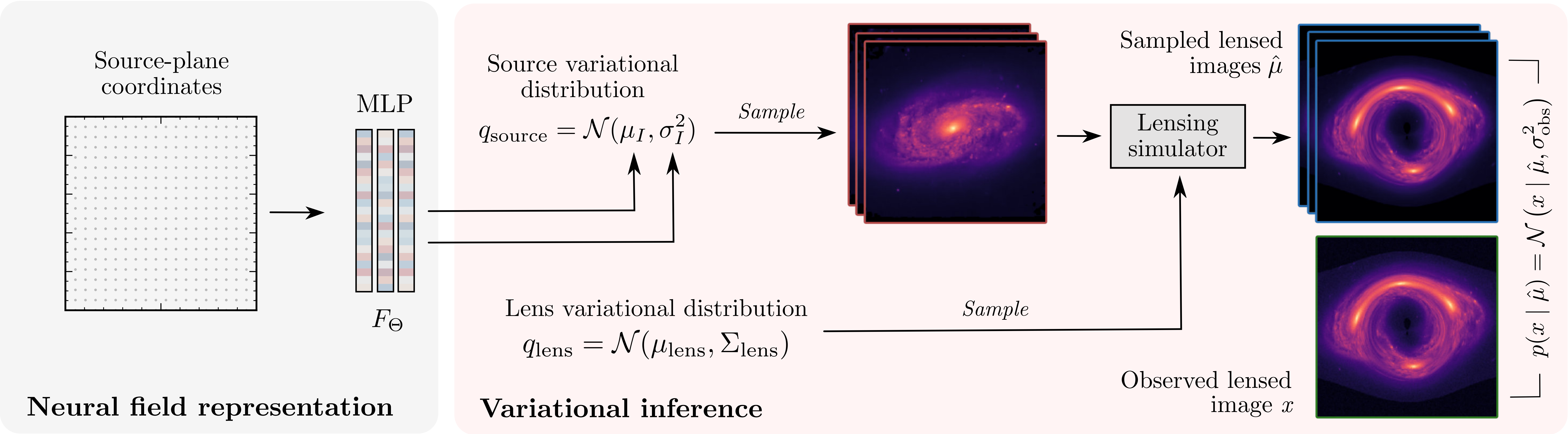
We can easily add our favorite new physics scenario

The simulations are differentiable

- *Efficient gradient-based optimization*
- *Inclusion of flexible models (sparse GPs, NNs, ...) as part of pipeline*

Example: Flexible inference via differentiable lensing

End-to-end gradient-based optimization using a differentiable, GPU-accelerated lensing simulator

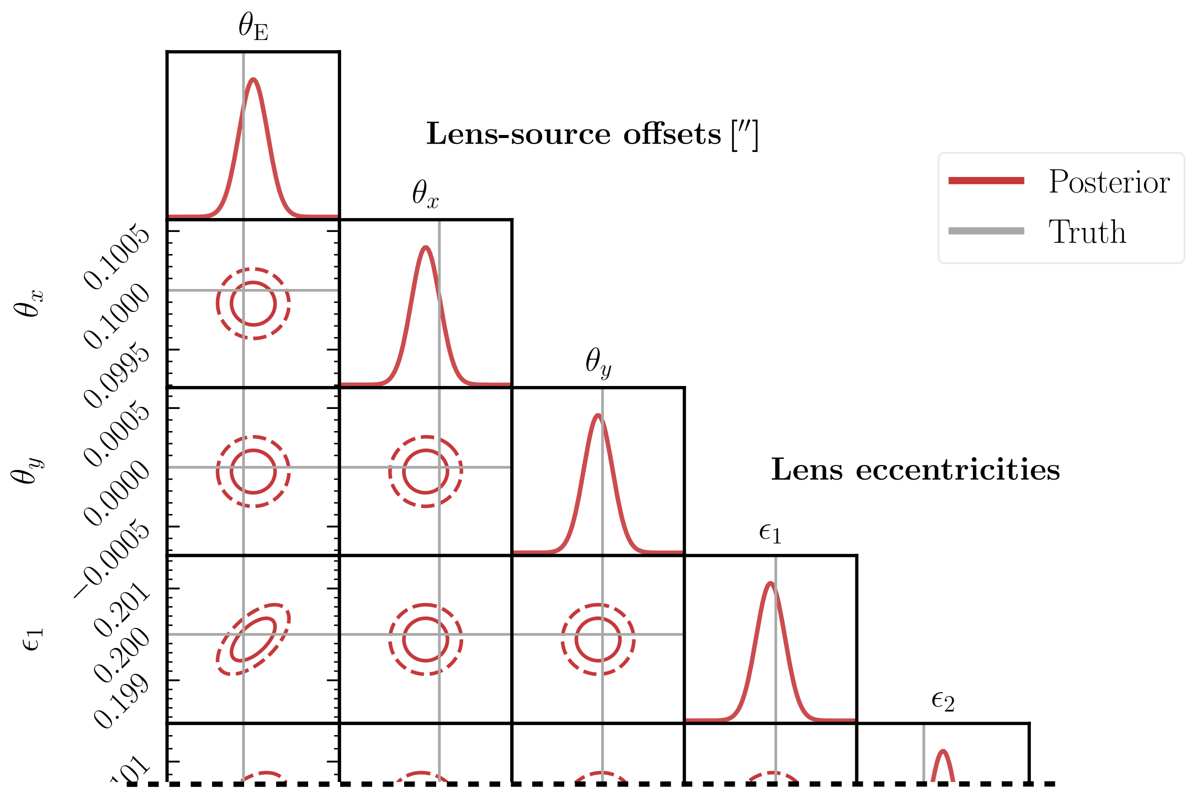


Likelihood-based analysis

Model complex part of the problem with a machine learning model

Chianese et al [MNRAS 2020]: Variational autoencoders
Karchev et al [MNRAS 2023]: Variational Gaussian Processes
Karchev et al [NeurIPS ML4PS 2022]: Diffusion models
Legin et al [ApJ 2023]: Diffusion models

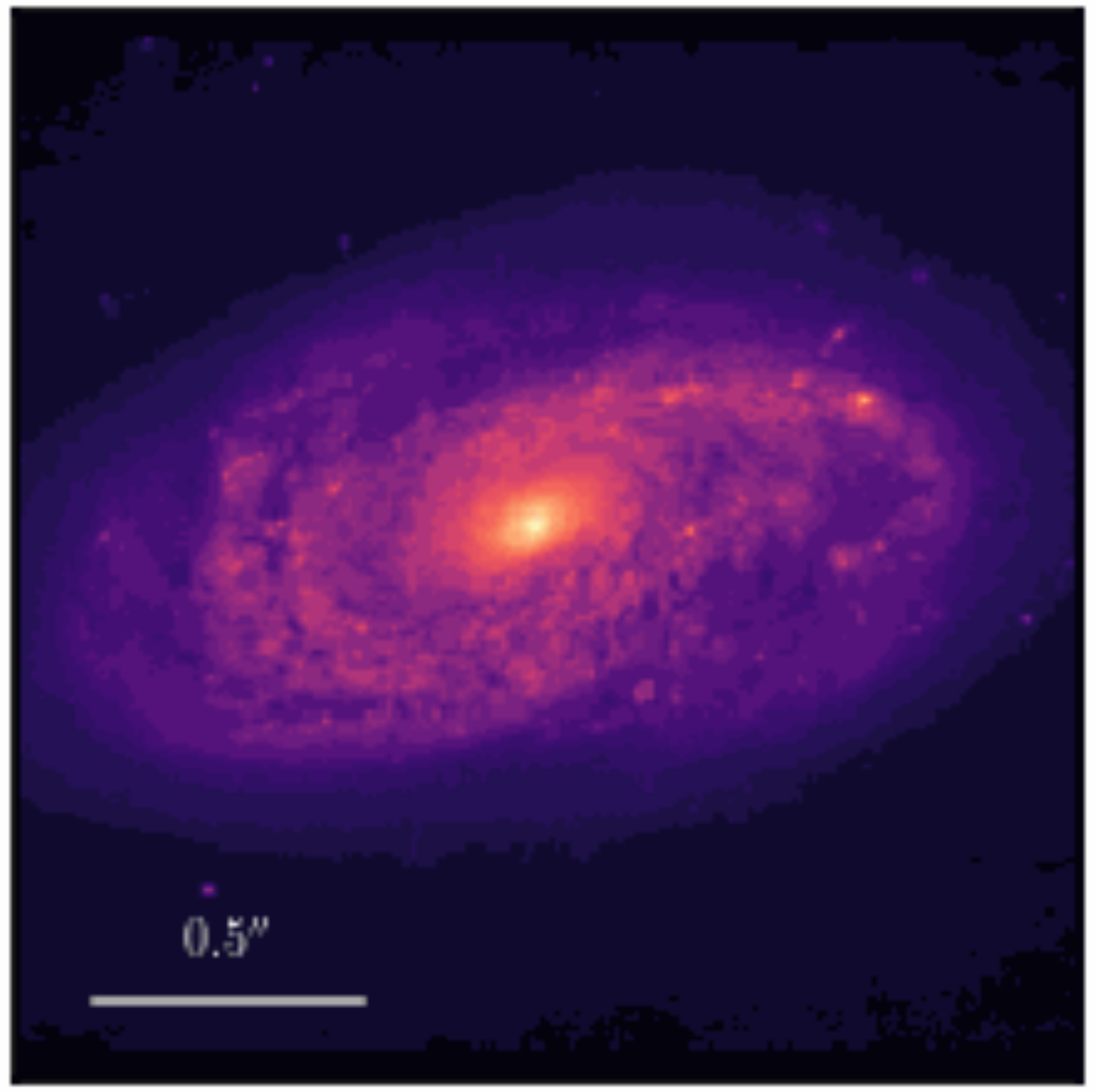
+ Differentiable lensing



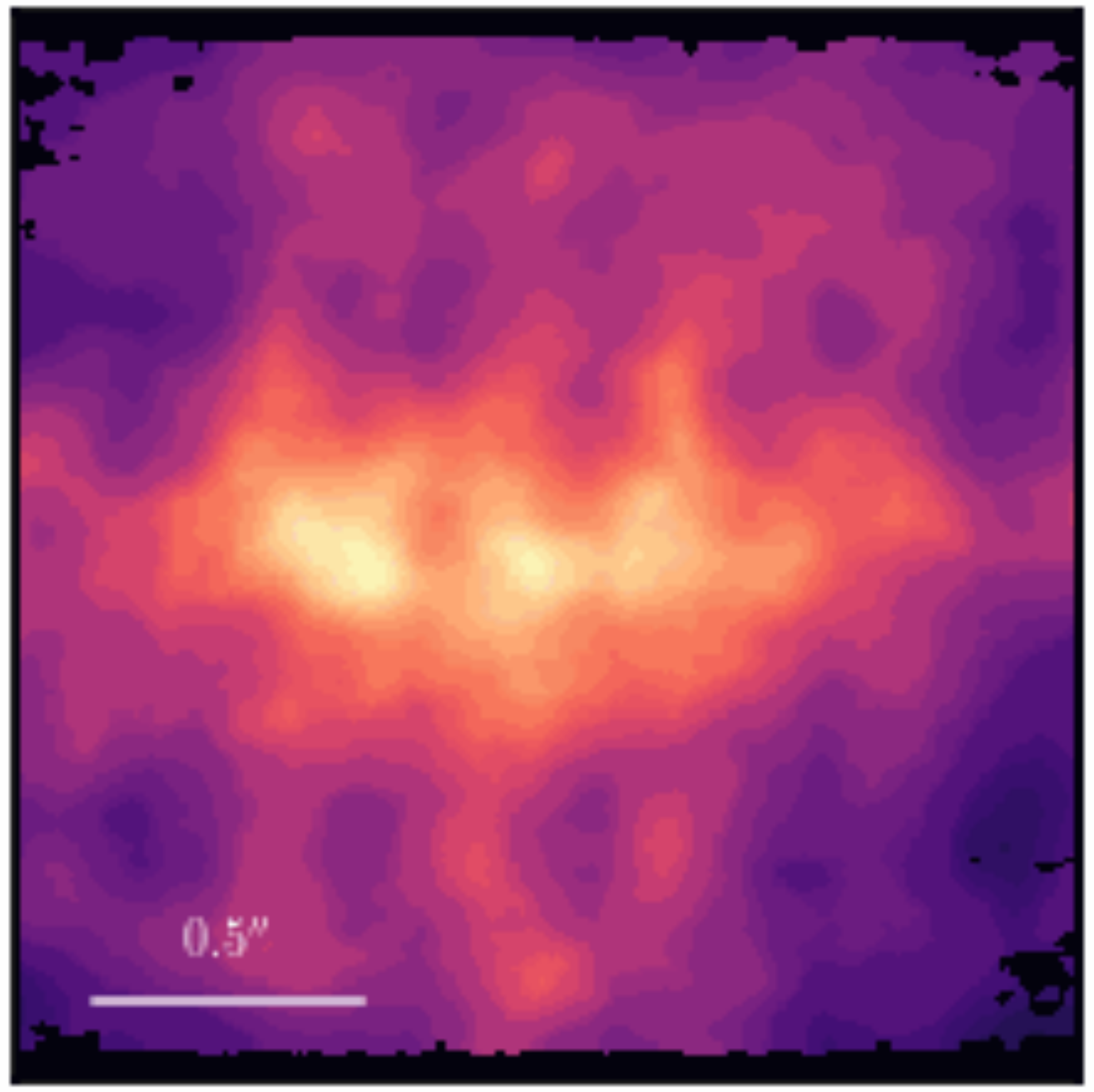
Simultaneously reconstruct posterior on lens model

Complex source reconstruction

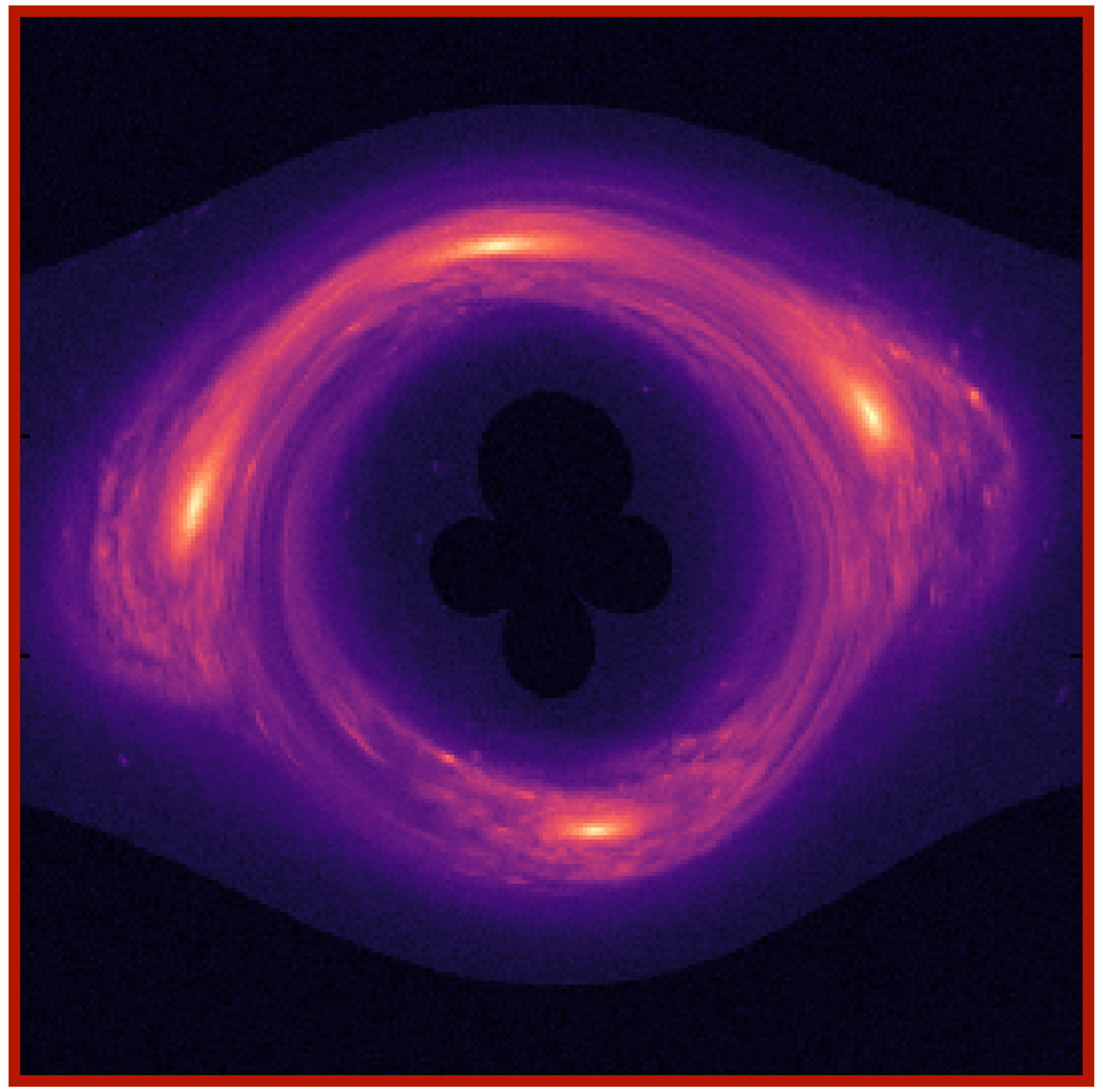
Ground truth

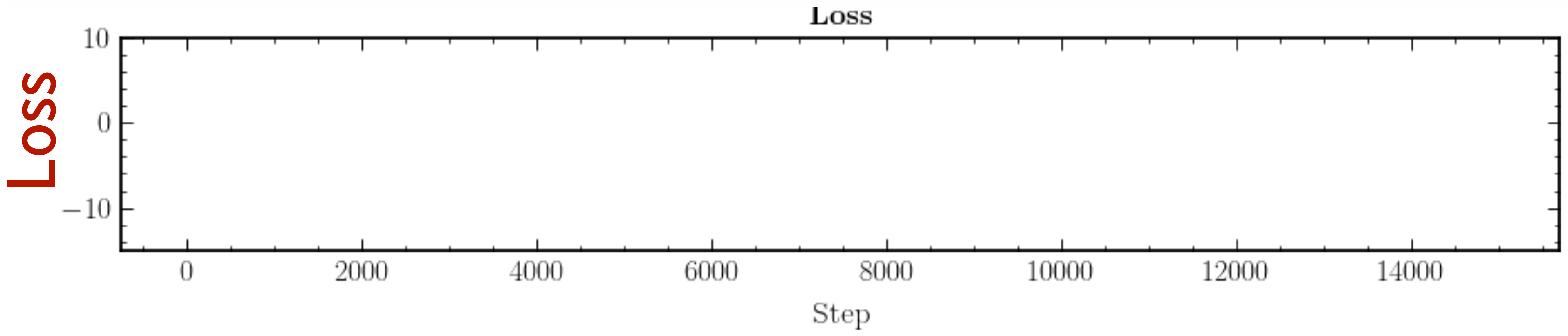


Reconstruction



Lensed image



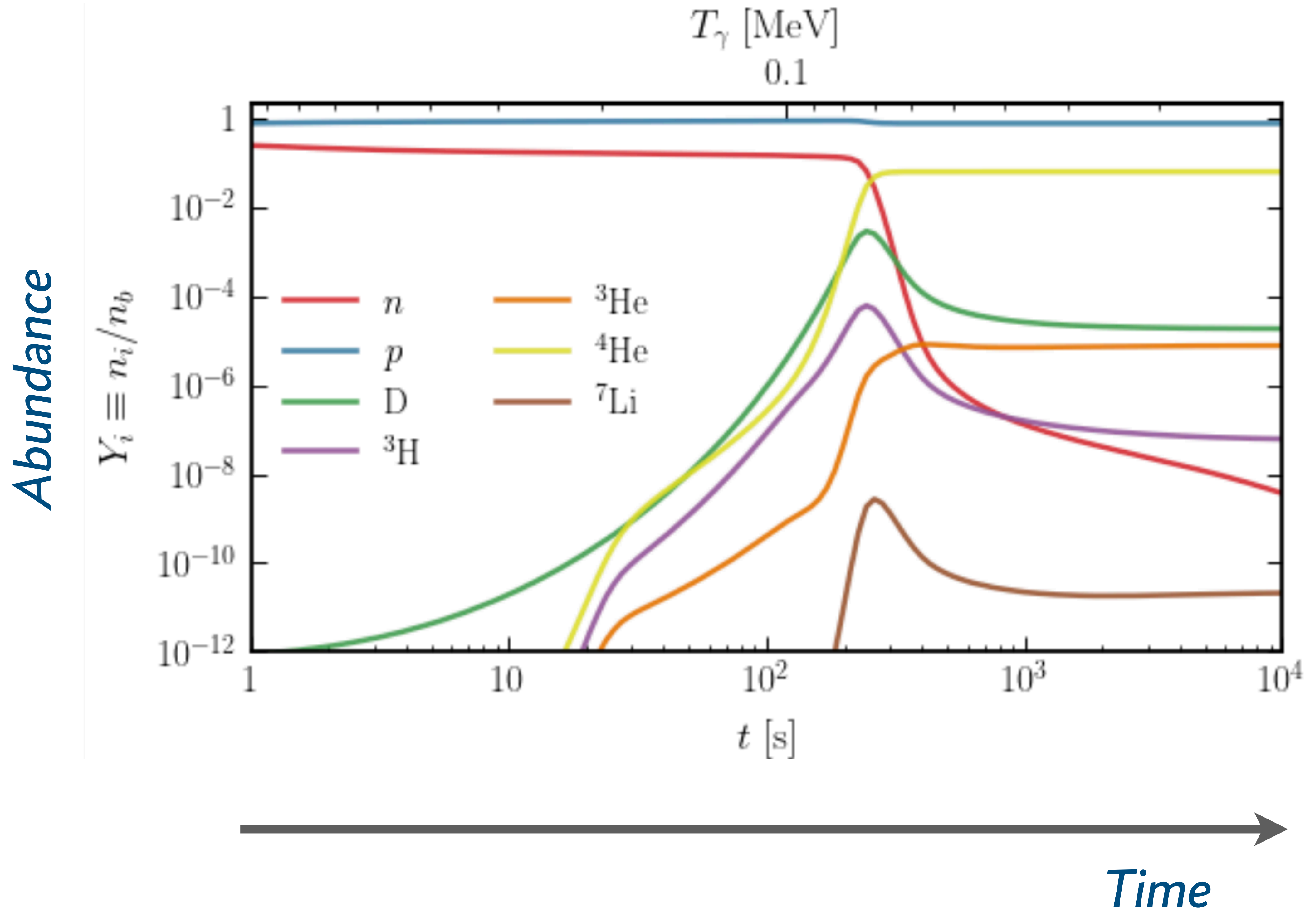


Probabilistic reconstruction of high-resolution source galaxy + dark matter lens!

Example: Differentiable big-bang nucleosynthesis

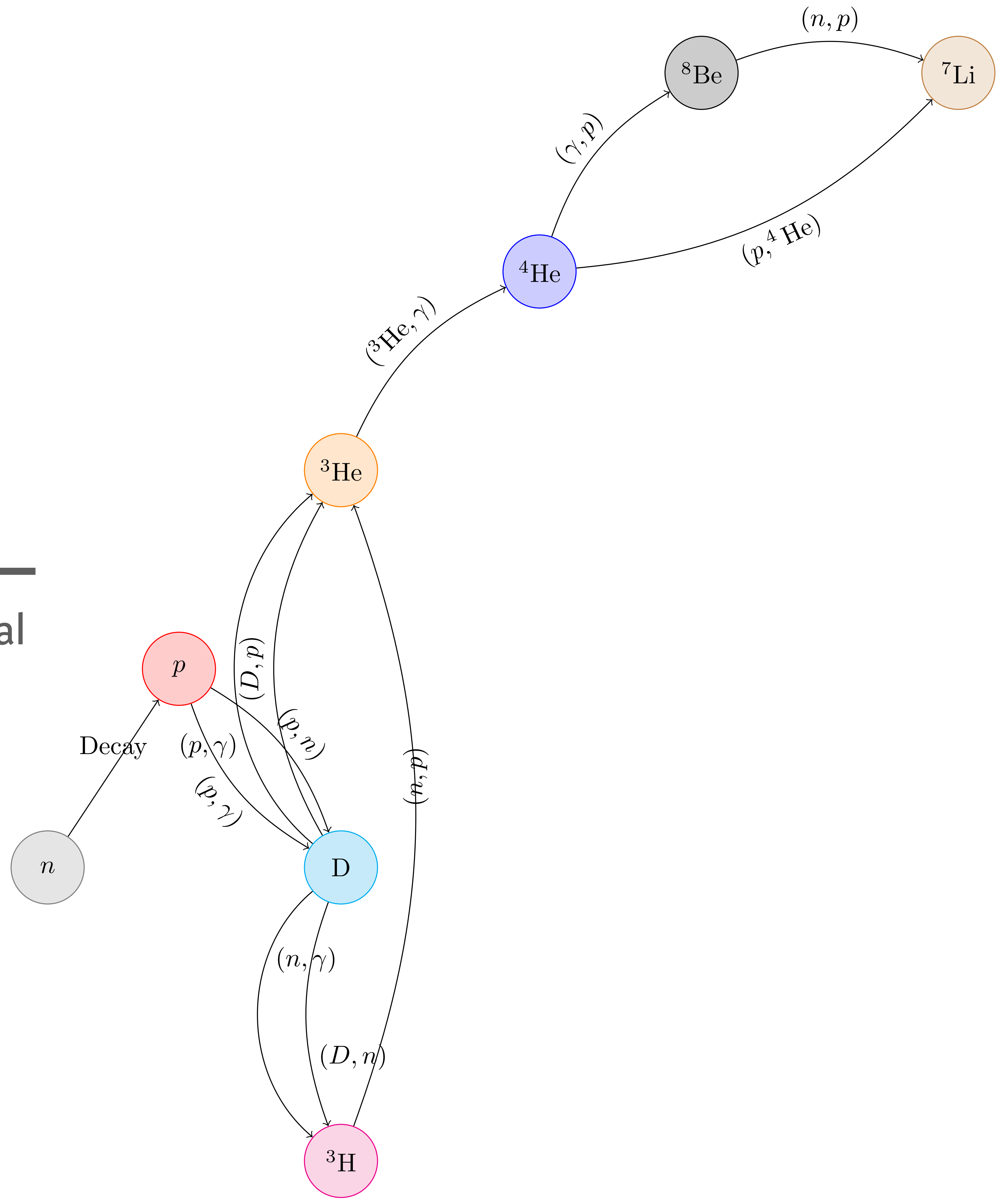
Giovanetti, Lisanti, Liu, SM, Ruderman [In prep]

Nuclear reactions active in the early hot Universe gradually shut off as temperature cools, freezing out the abundance of light elements



Nuclear reaction rates network

(Coupled differential equations)



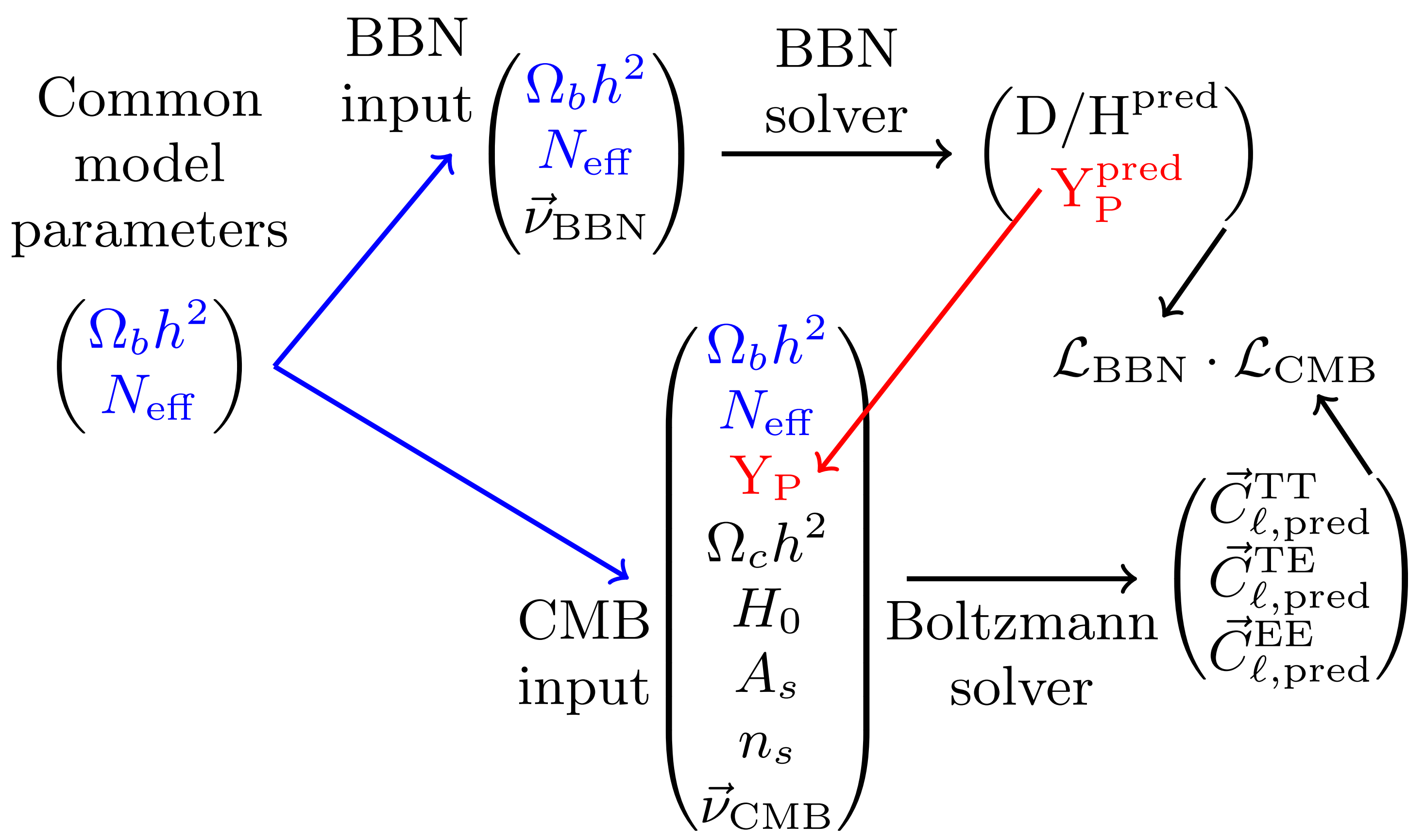
Measure this abundance in pristine gas clouds in early Universe, and compare to predictions — sensitive to MeV scale physics

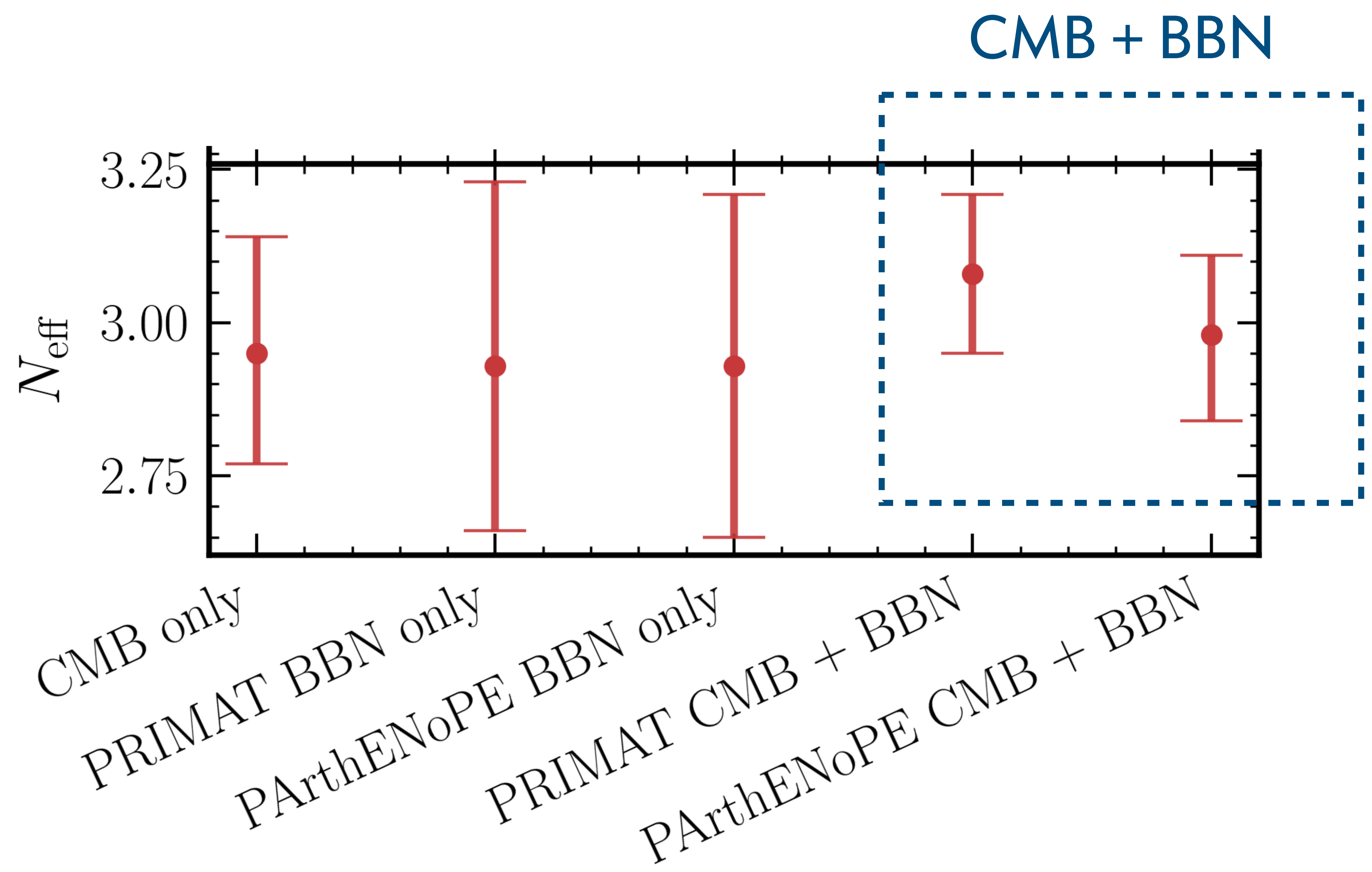
Example: Differentiable big-bang nucleosynthesis

Giovanetti, Lisanti, Liu, SM, Ruderman [In prep]

LINX: Light Isotope Nucleosynthesis with JAX

- Predict BBN observables in $\mathcal{O}(0.1)$ s without compromises
- Fully differentiable → amenable to variational inference and gradient-based sampling
- Easily extensible for new physics scenarios
- Puts BBN on the same footing as CMB, and allows for principled combinations





Using differentiable CMB emulator: `cosmopower`

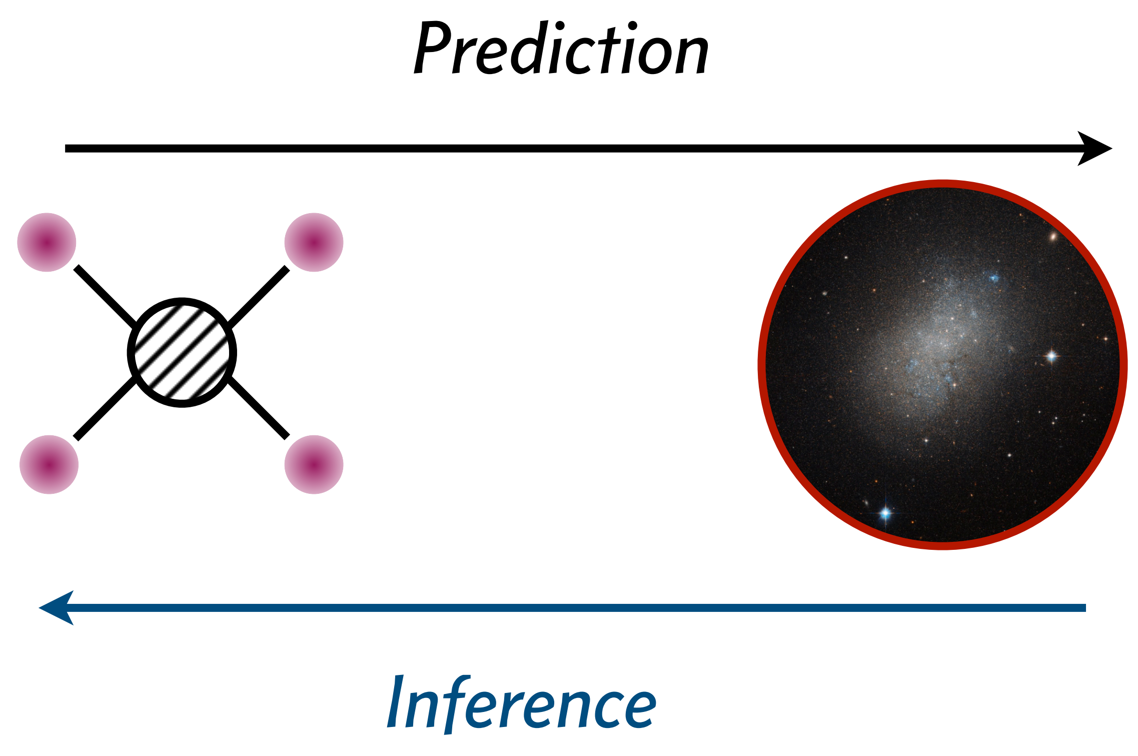
Spurio Mancini et al [MNRAS 2022]

See DISCO-DJ: DIfferentiable Simulations for COsmology — Done with JAX

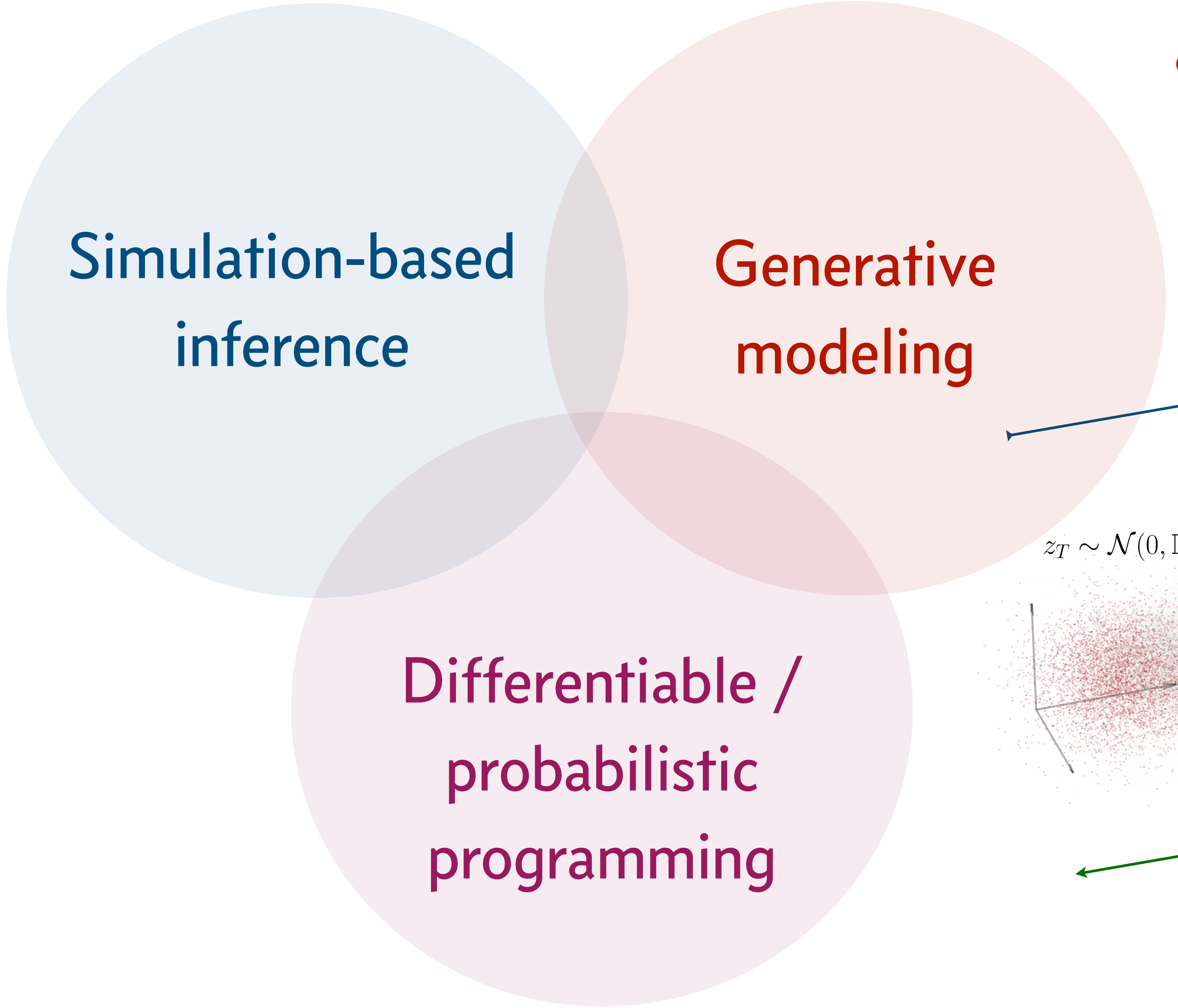
Hahn, List, Porqueres [2023]

The future is differentiable!

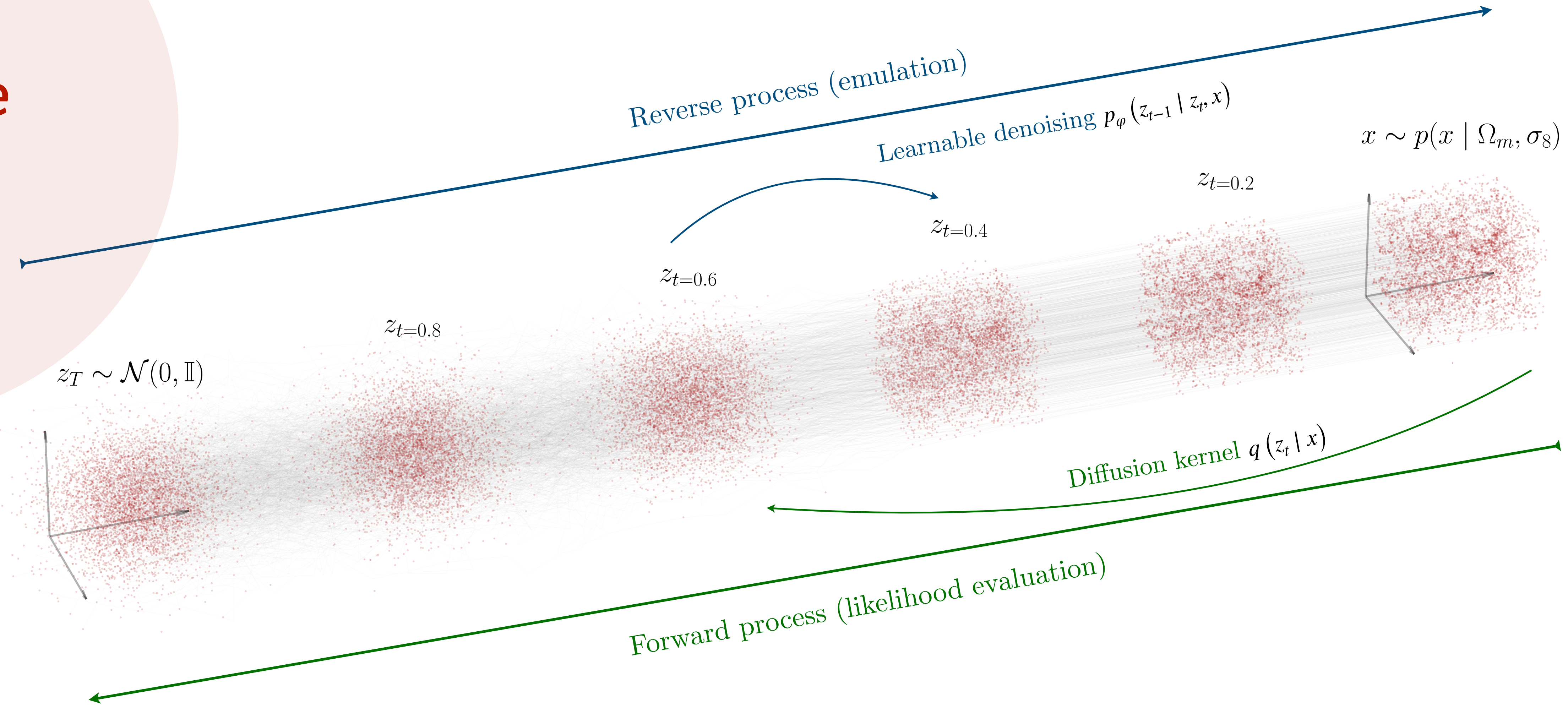
Conclusions



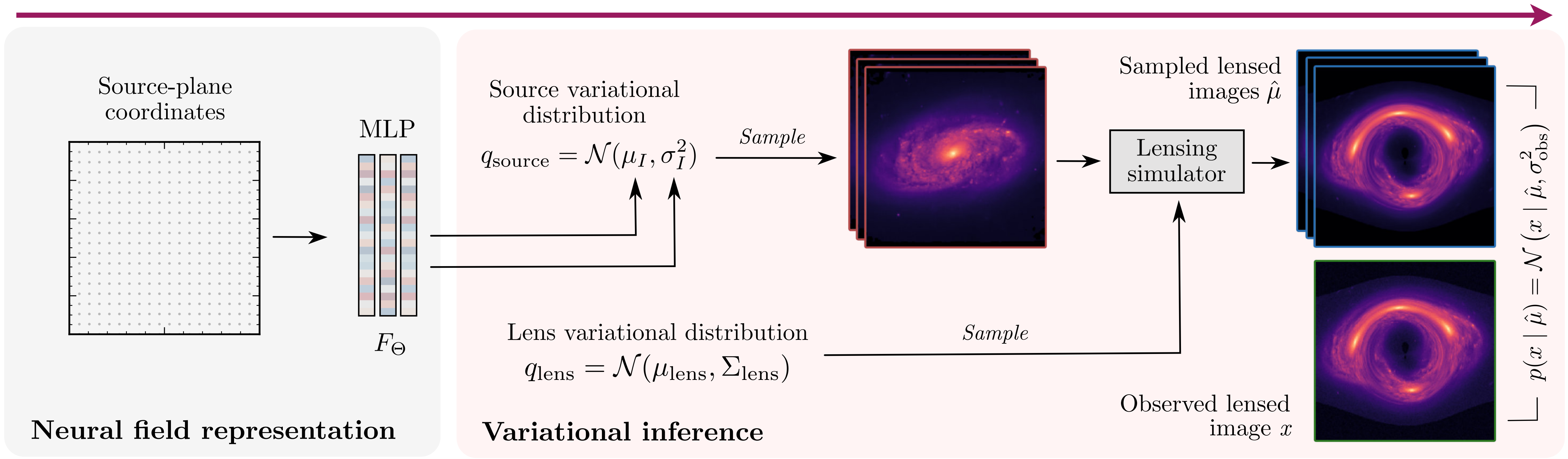
- *Invert complex physical simulators*
- *Directly work with high-dim data*



- *Encode complex physical distribution*
- *Uses end-to-end or as physical priors*
- *Compute data-sim compatibility*



$\nabla_\theta x$



- *Flexible specification of model components*
- *Enable high-dimensional optimization using gradient-based inference techniques*