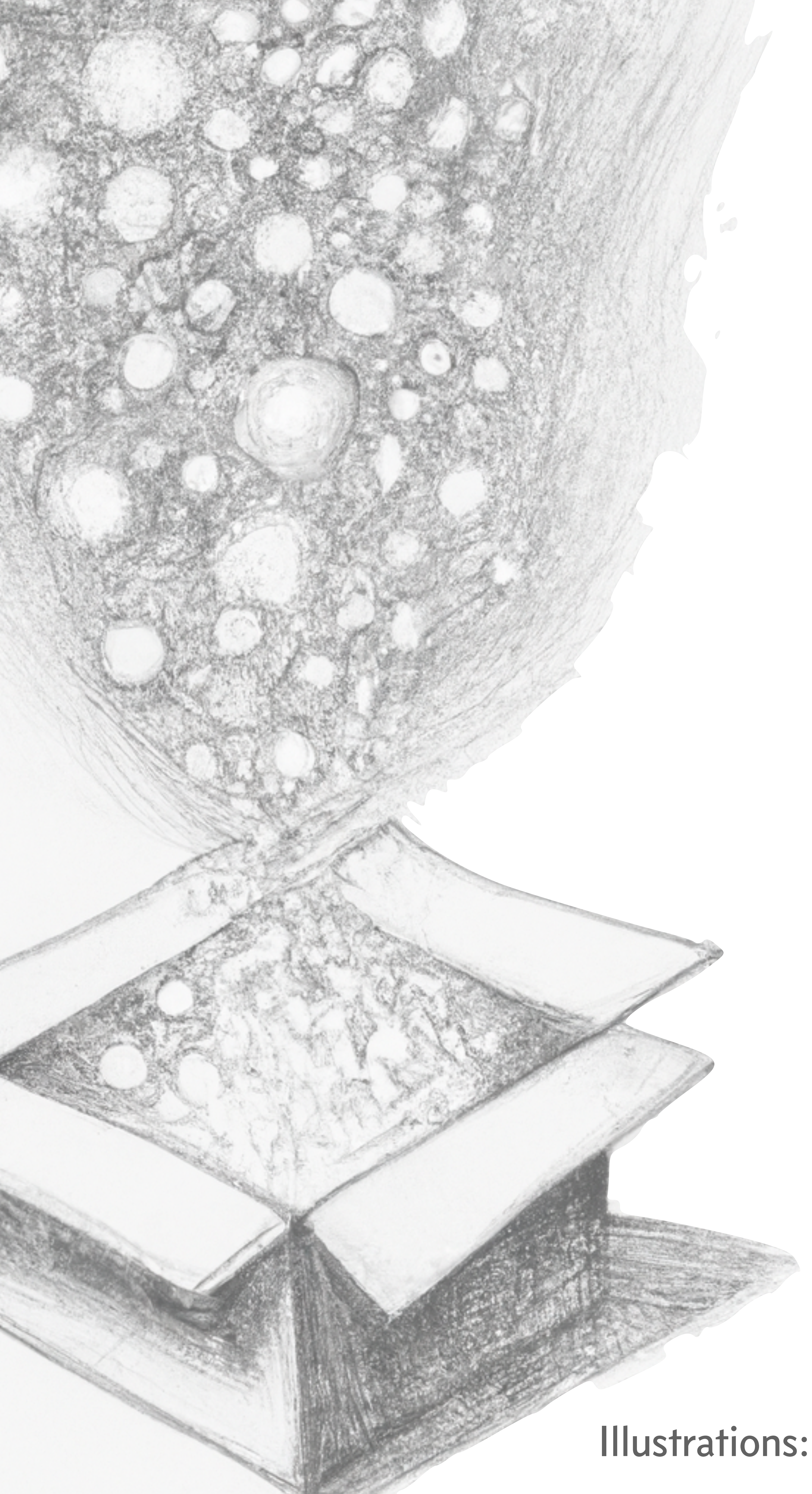


Astrophysical searches for dark matter with neural simulation-based inference

Based on 2208.12825; Nguyen, SM, Williams, Necib
arXiv:2110.06931; SM, Cranmer

Siddharth Mishra-Sharma

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

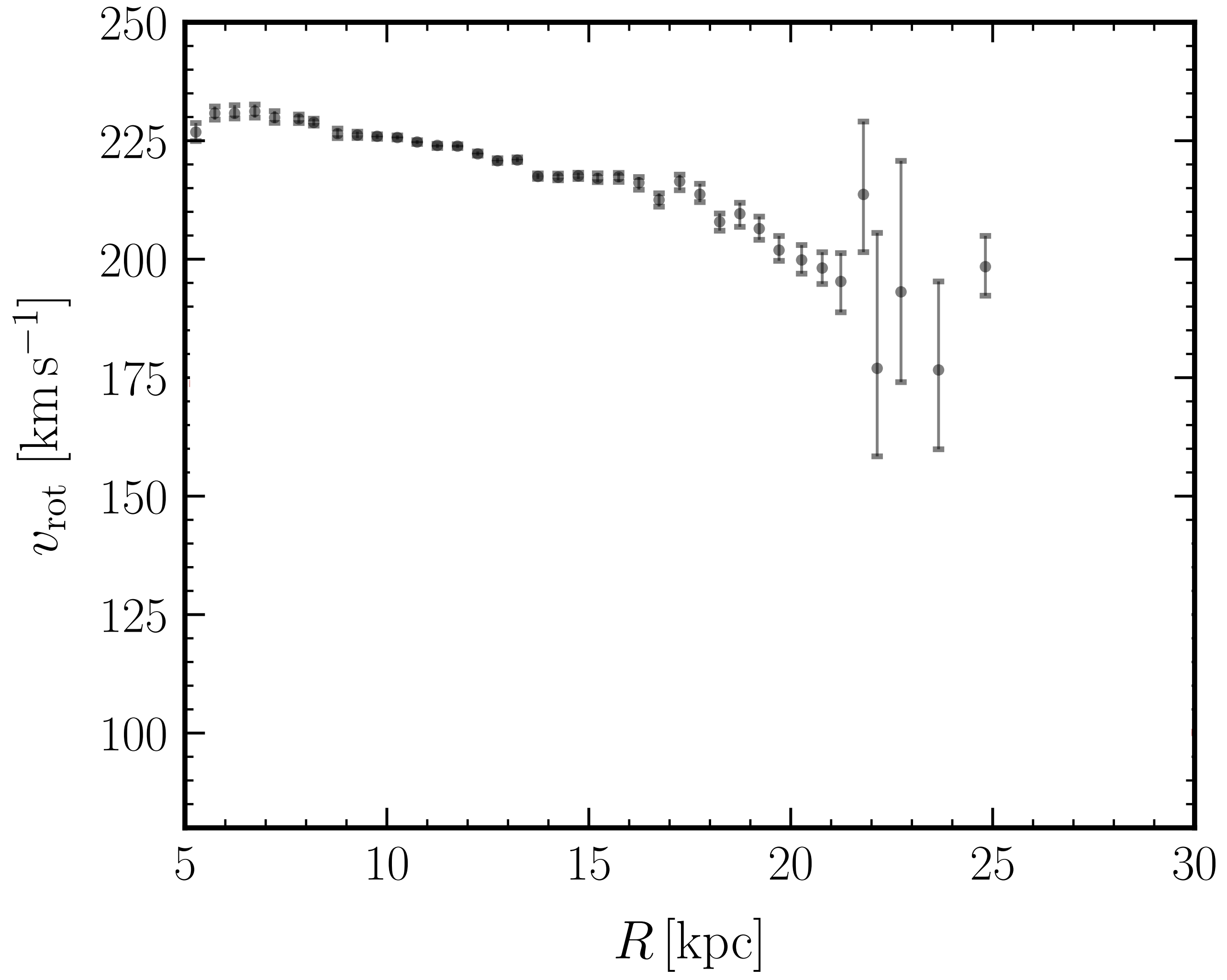
Aspen Winter 2023
March 29, 2023

Illustrations: *Stable Diffusion*

Inference in astrophysics

Parameters of interests, θ

Data, x
Observations



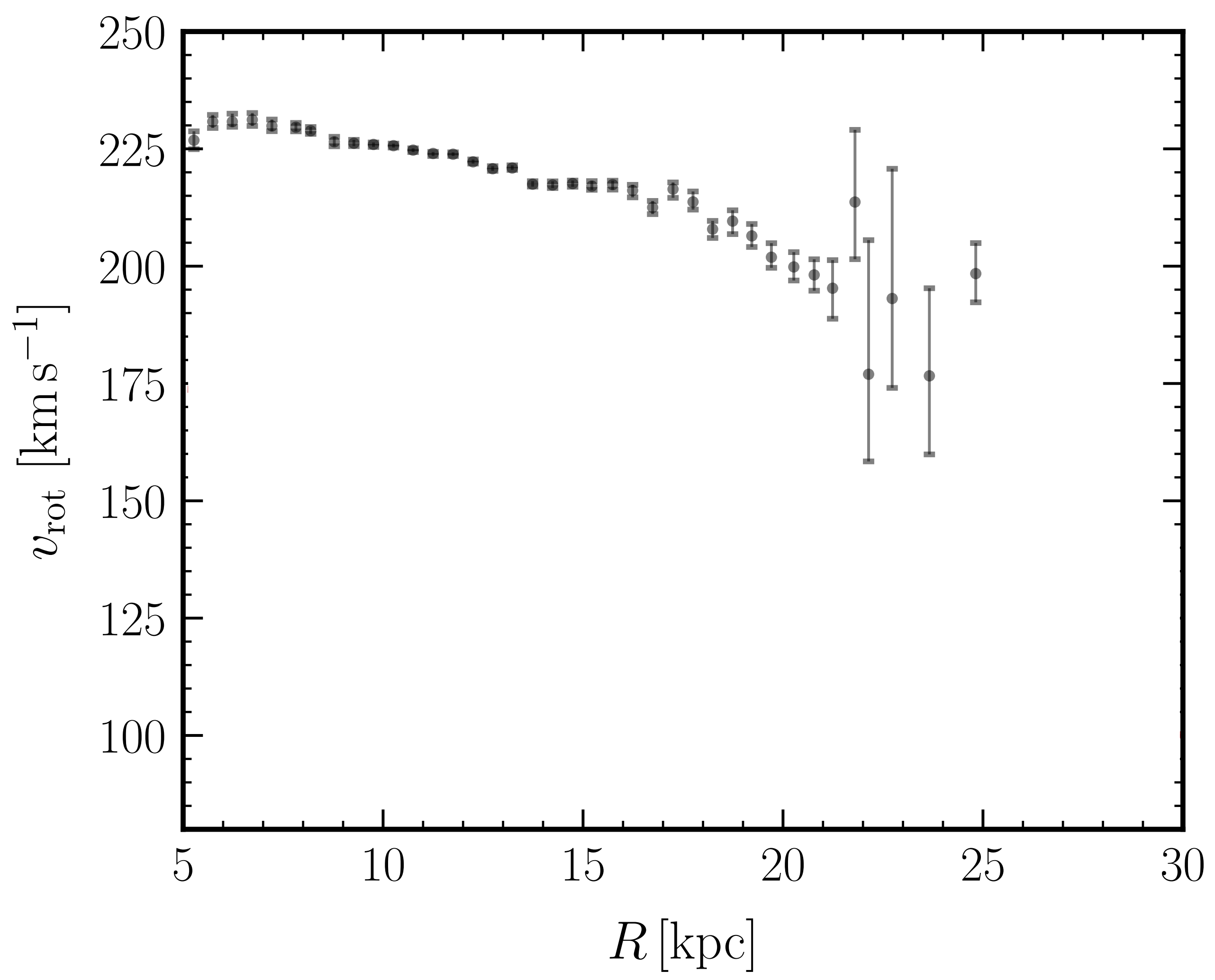
Inference in astrophysics

Parameters of interests, θ

Data, x
Observations

$p(x \mid \theta)$

Likelihood function



Inference in astrophysics

Parameters of interests, θ

Latent variables, z

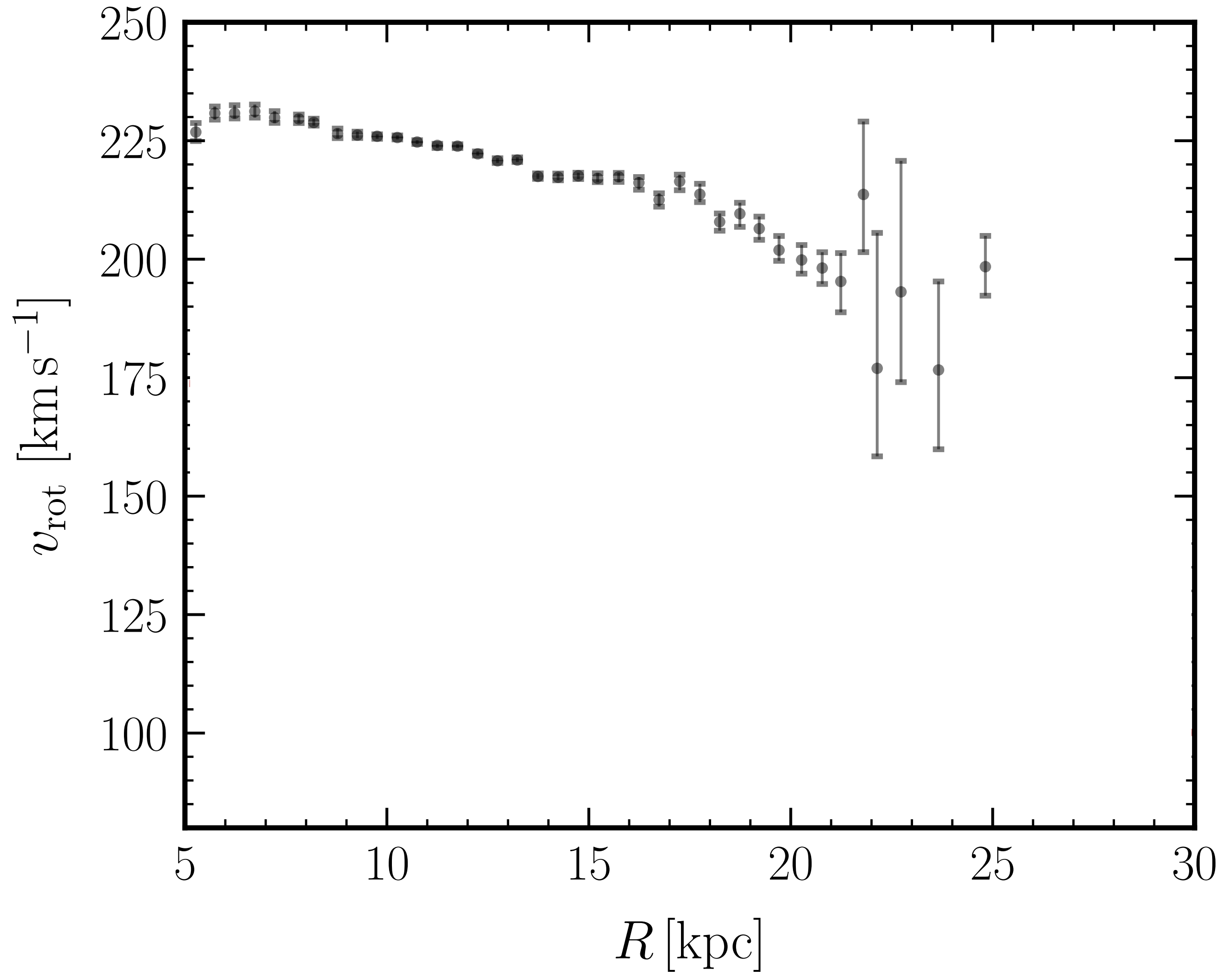
(Modeled) Parameters other than θ which participate in the data-generation process

Data, x

Observations

$$p(x \mid \theta) = \int \mathrm{d}z \, p(x, z \mid \theta)$$

Likelihood function



Inference in astrophysics

Parameters of interests, θ

Latent variables, z

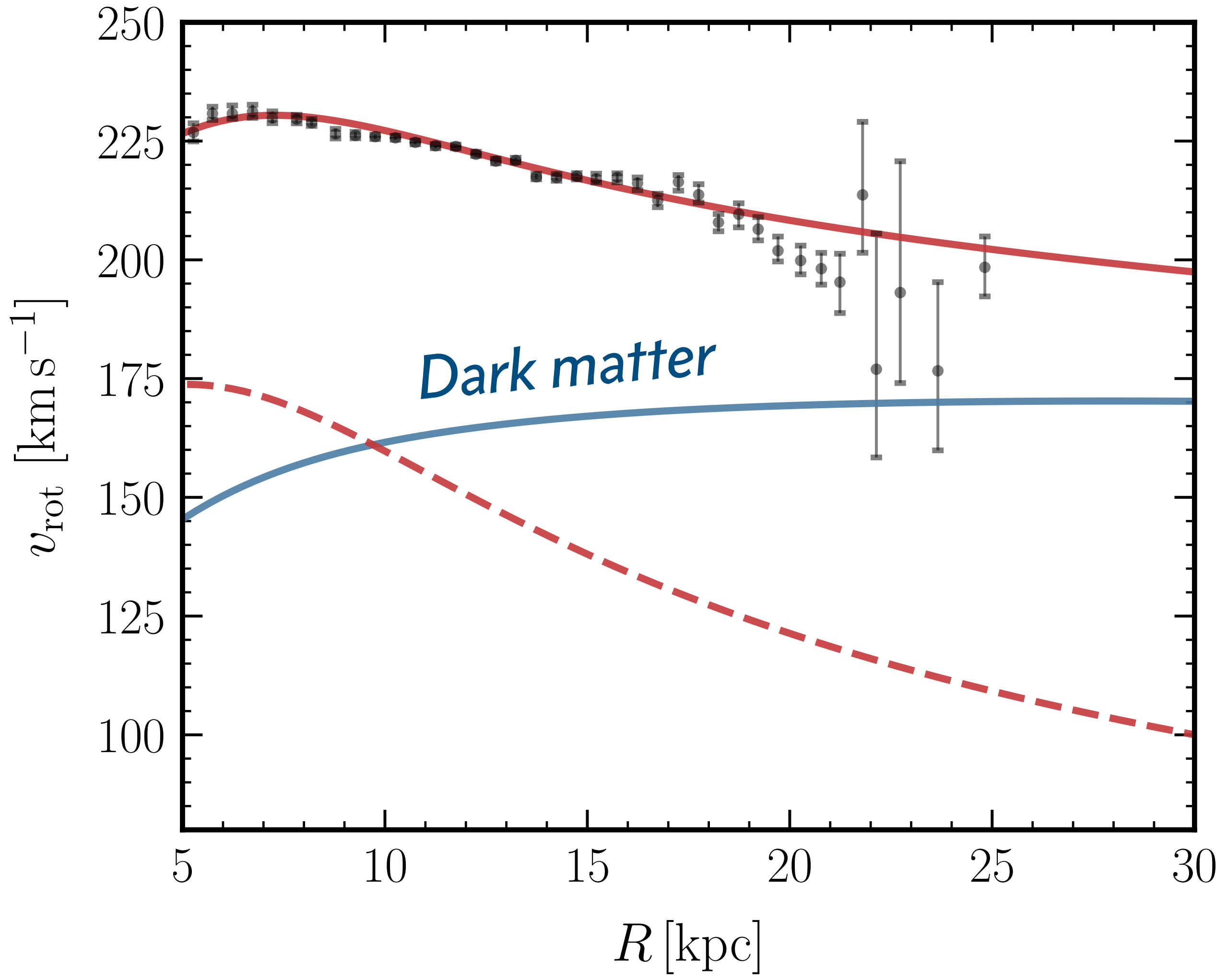
(Modeled) Parameters other than θ which participate in the data-generation process

Data, x

Observations

$$p(x \mid \theta) = \int \mathrm{d}z \, p(x, z \mid \theta)$$

Likelihood function



$$p(x \mid \theta) \propto e^{\left(-\frac{v_{\rm obs} - v_{\rm model}(\theta)}{\sigma_{v_{\rm obs}}}\right)^2}$$

Inference in astrophysics

Hypothesis testing

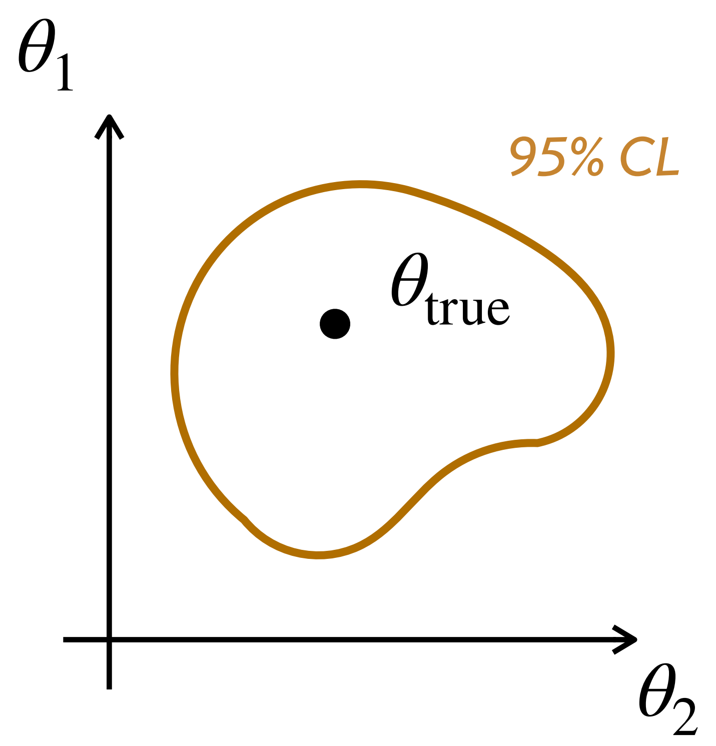
Parameters of interests, θ

Latent variables, z

(Modeled) Parameters other than θ which participate in the data-generation process

Data, x

Observations

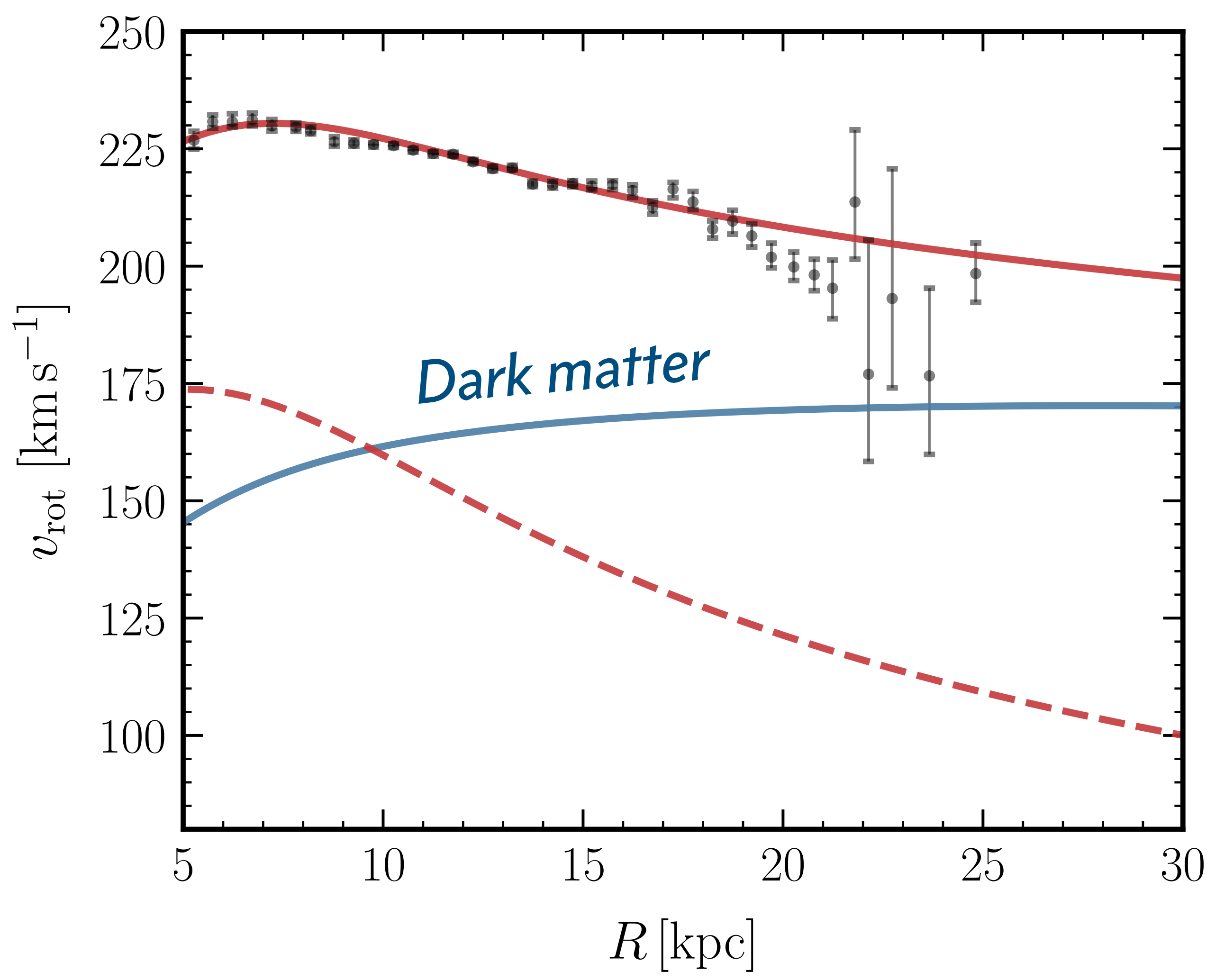


Confidence sets

$\int \mathrm{d}\theta\, p(\theta|x) = 0.95$

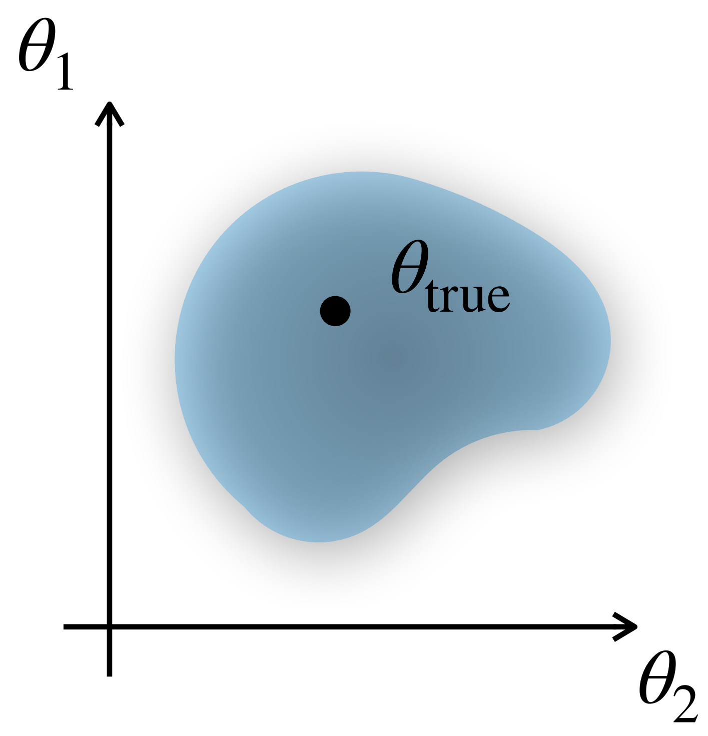
$$p(x\,|\,\theta) = \int \mathrm{d}z\, p(x, z\,|\,\theta)$$

Likelihood function



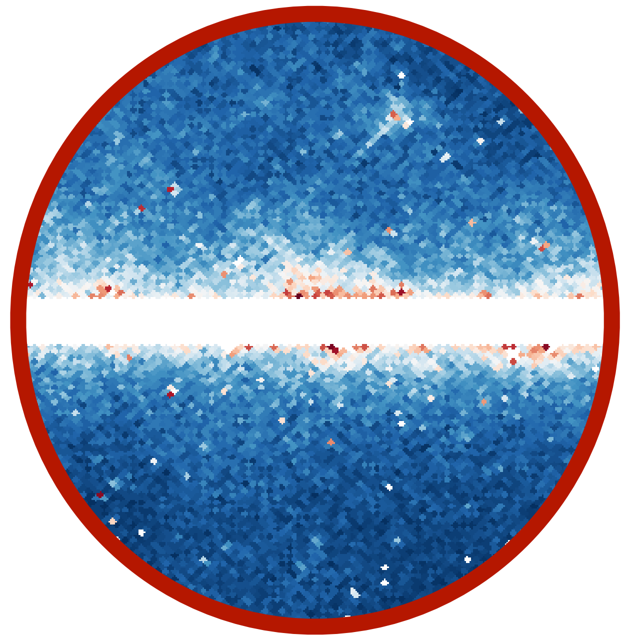
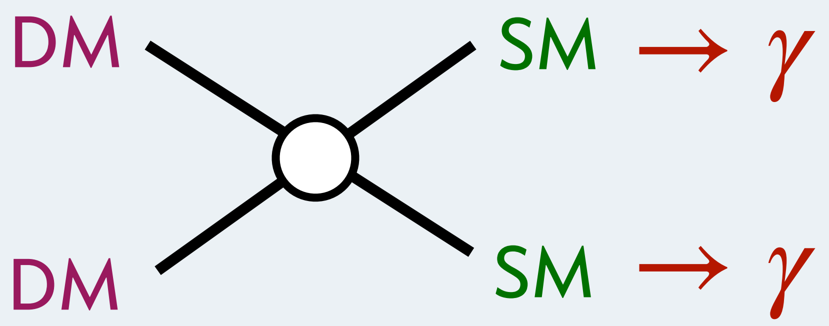
Posterior distributions

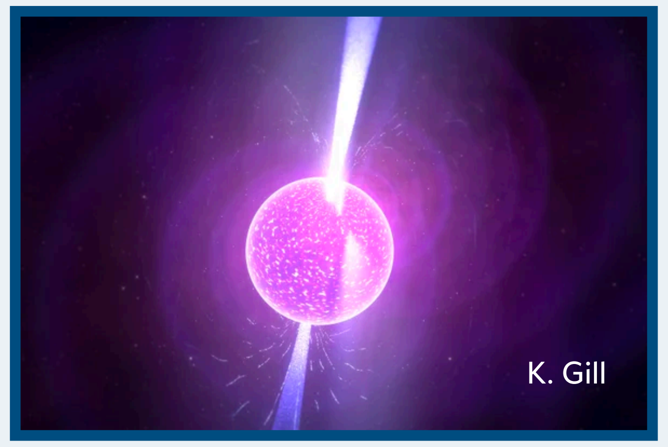
$p(\theta|x) \sim p(x\,|\,\theta) \cdot p(\theta)$



$$p(x\,|\,\theta) \propto e^{\left(-\frac{v_{\text{obs}} - v_{\text{model}}(\theta)}{\sigma_{v_{\text{obs}}}}\right)^2}$$

Likelihood is often intractable…

$p($  $|$ 



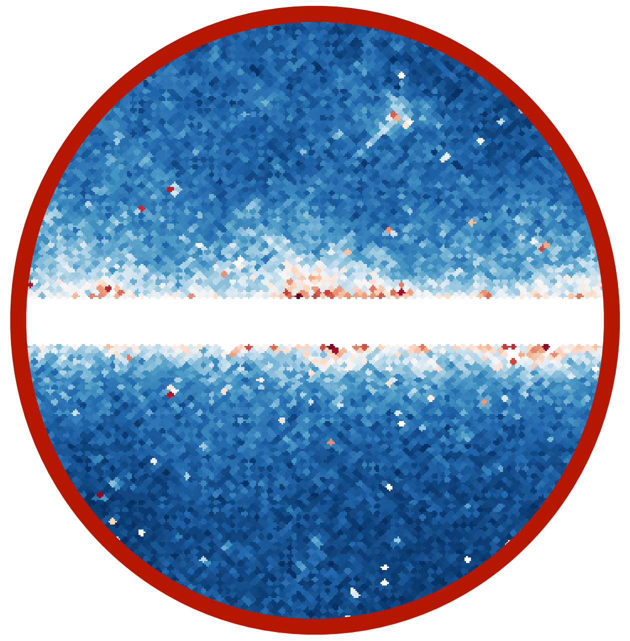
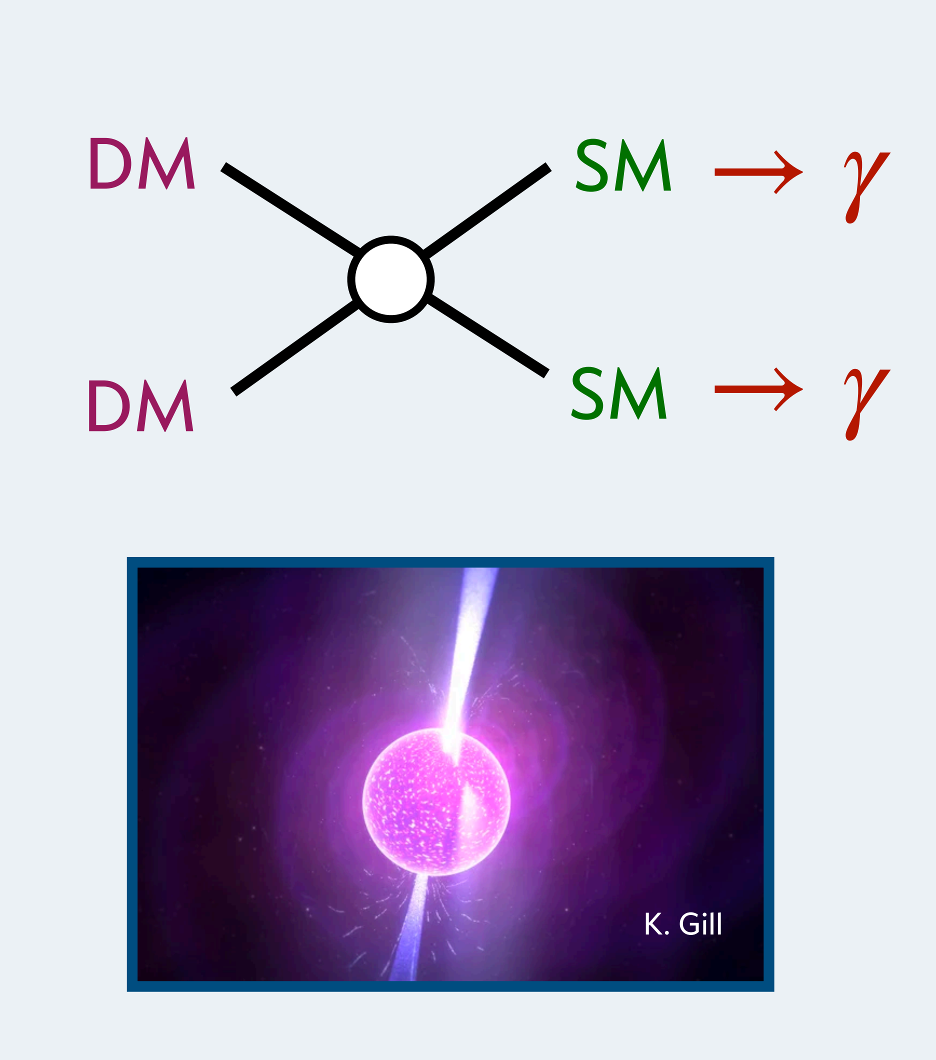
$)$

γ-ray data

Pulsars or DM

Data analysis typically requires simplifying assumptions

Likelihood is often intractable…

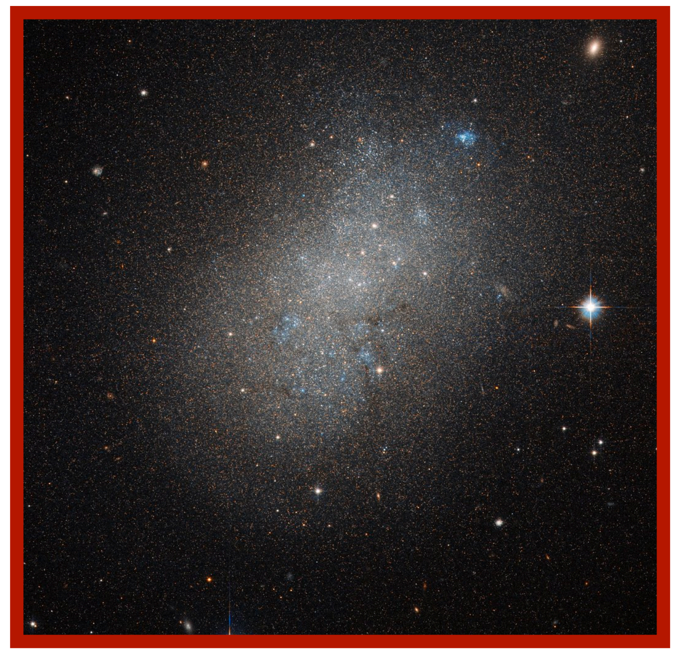
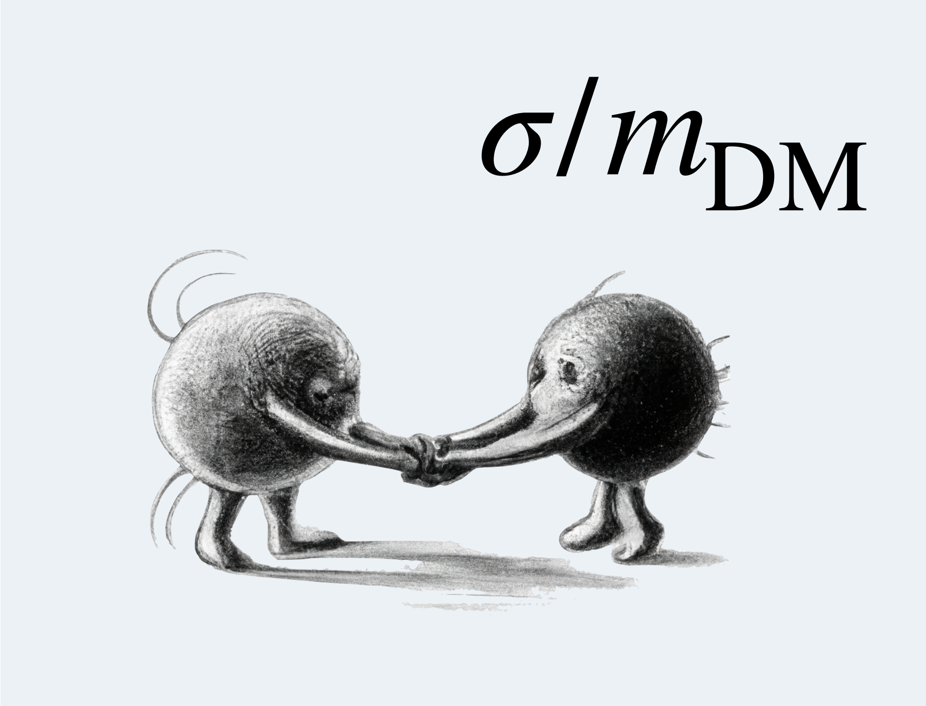
$p($  $|$ 

$)$

γ-ray data

Pulsars or DM

Data analysis typically requires simplifying assumptions

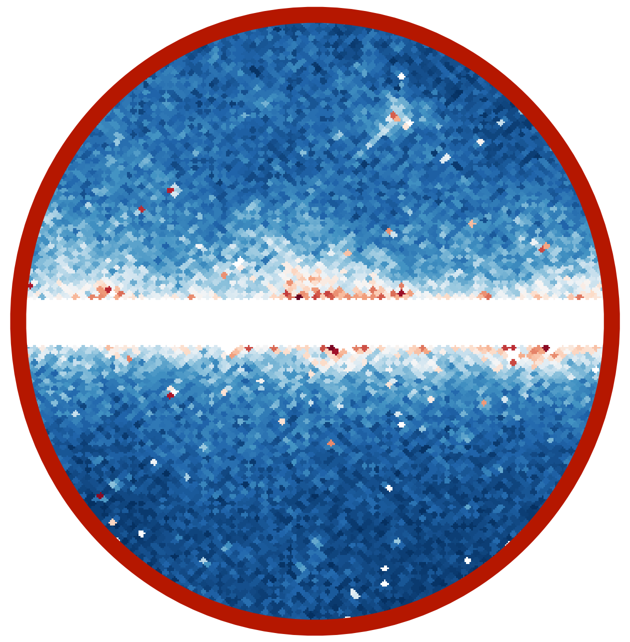
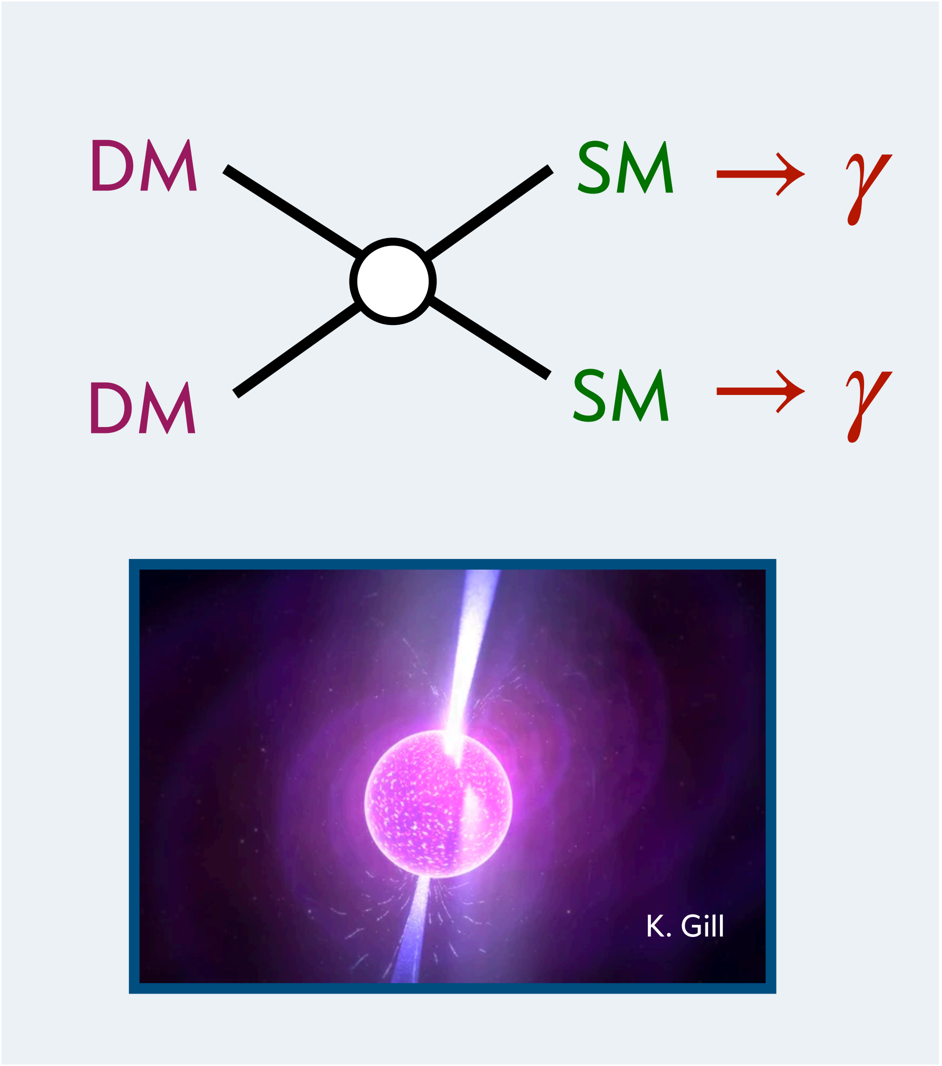
$p($  $|$ 

$)$

Dwarf galaxies

Self-interactions

Likelihood is often intractable…

$p\Big($  $\Big|$ 

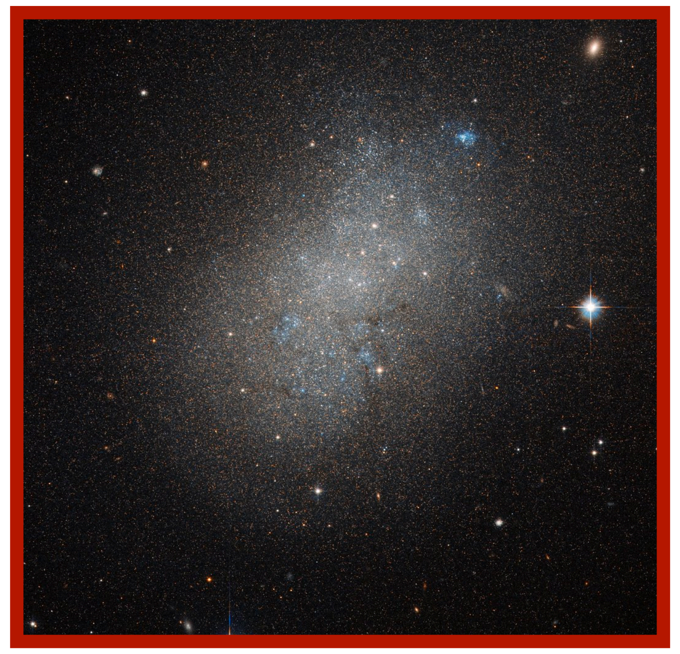
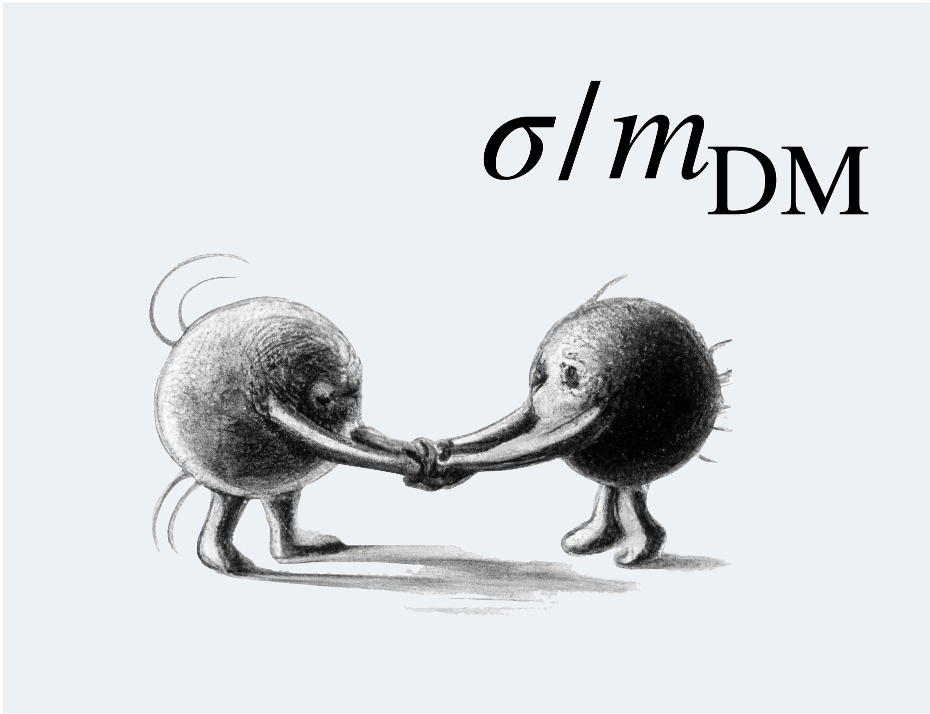
$\Big)$

γ-ray data

Pulsars or DM

Data analysis typically requires simplifying assumptions

How can we do inference without compromise?

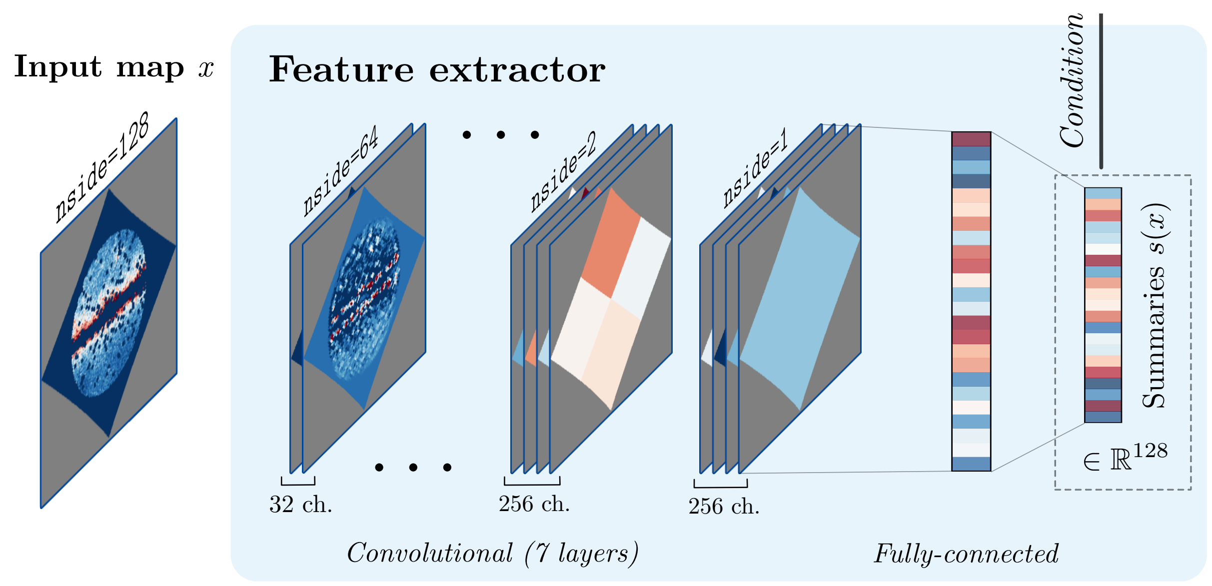
$p\Big($  $\Big|$ 

$\Big)$

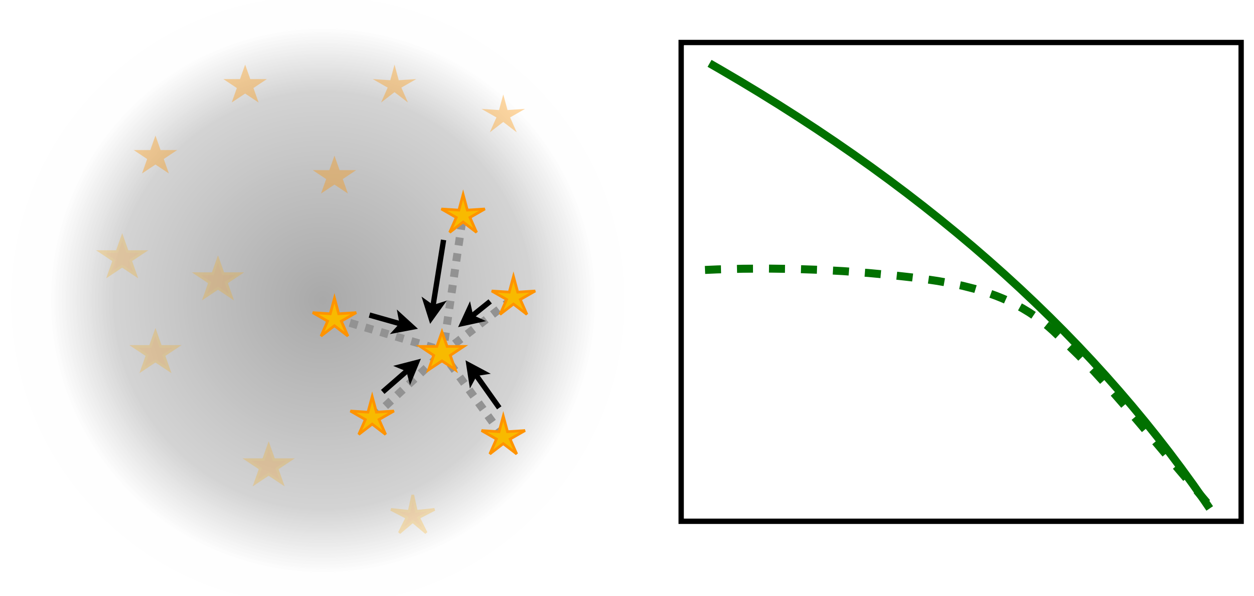
Dwarf galaxies

Self-interactions

Outline

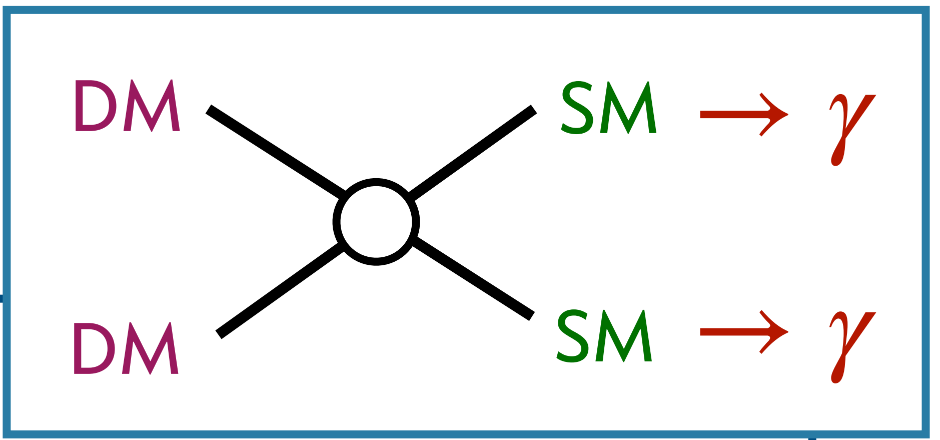


Characterizing the Galactic Center Excess



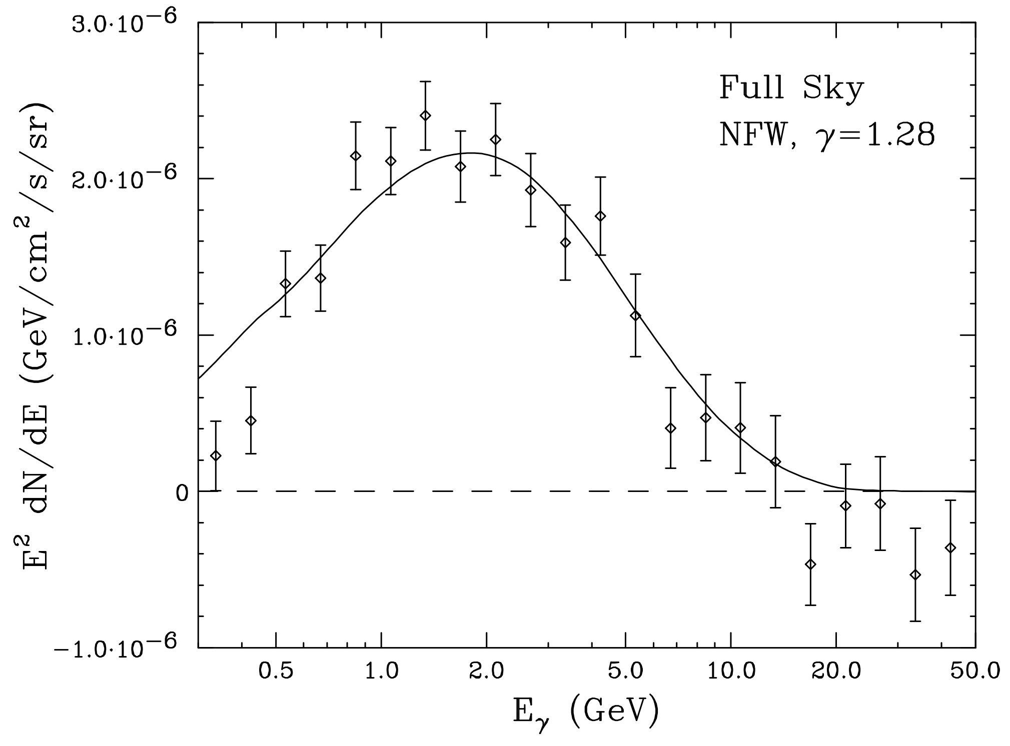
Inferring dark matter halo shapes in dwarf galaxies

Possible explanations

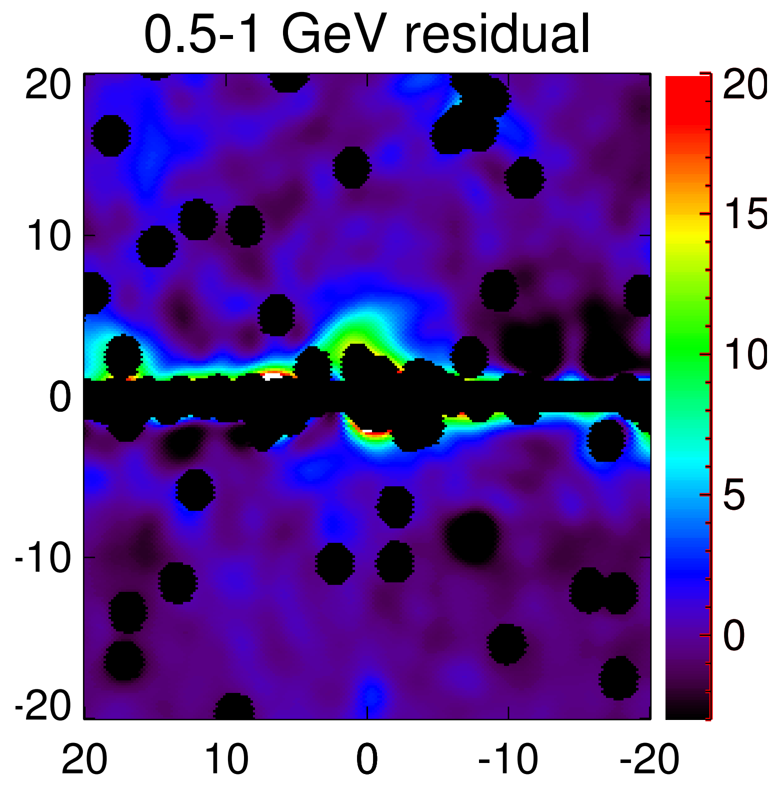


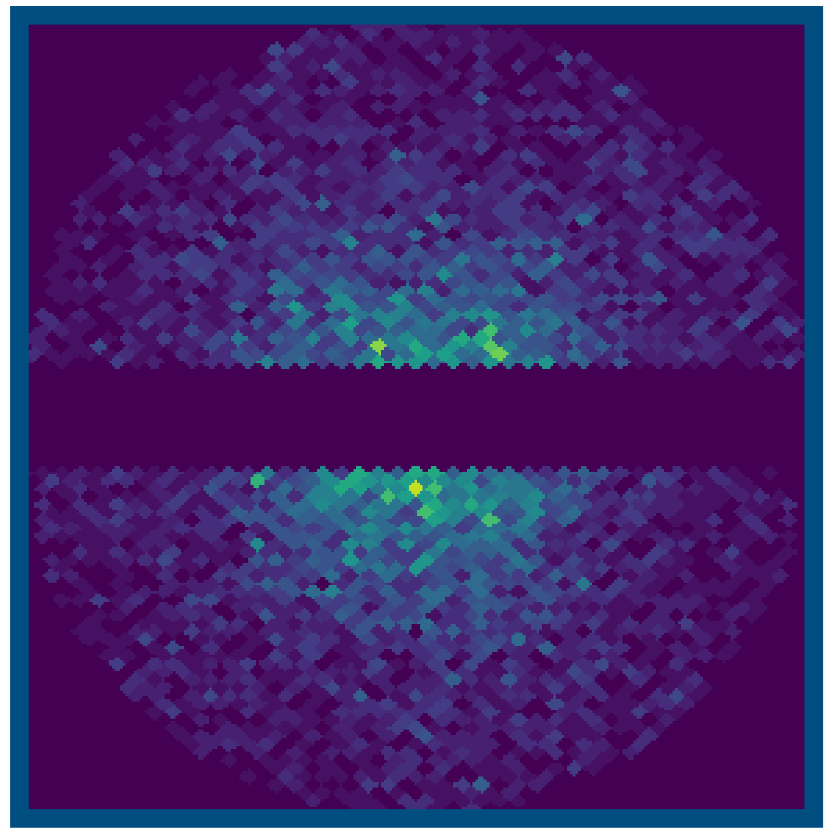
Dark Matter

Spectrum and morphology consistent with DM expectation



Daylan et al [PDU 2016]

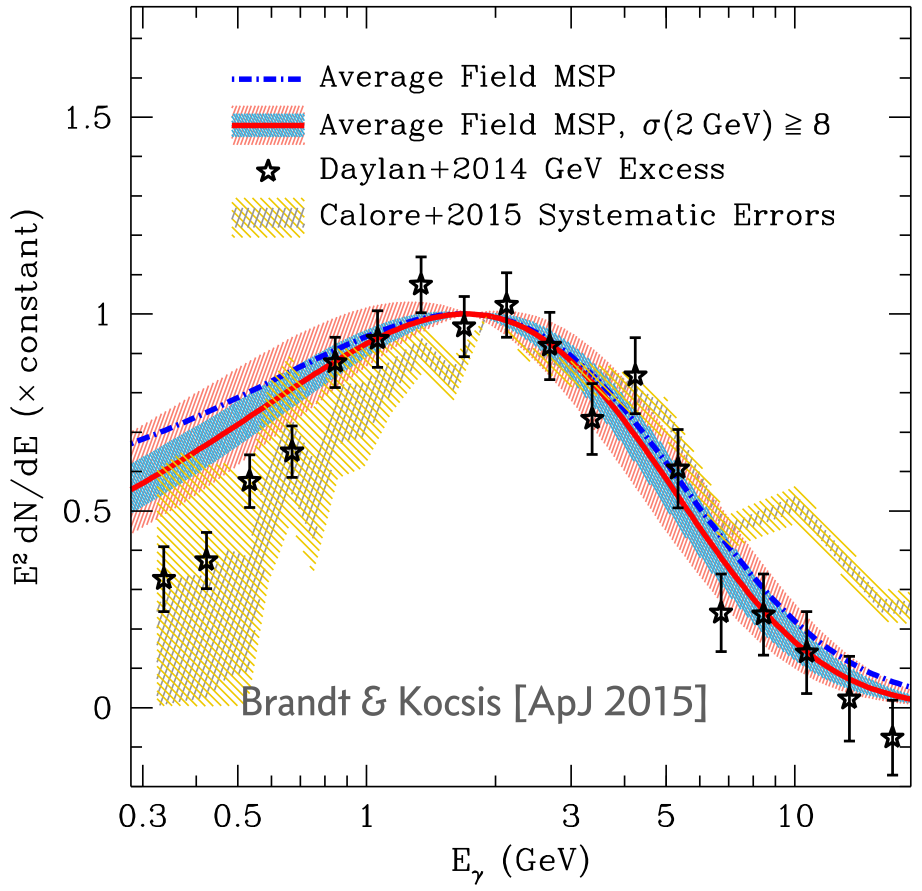


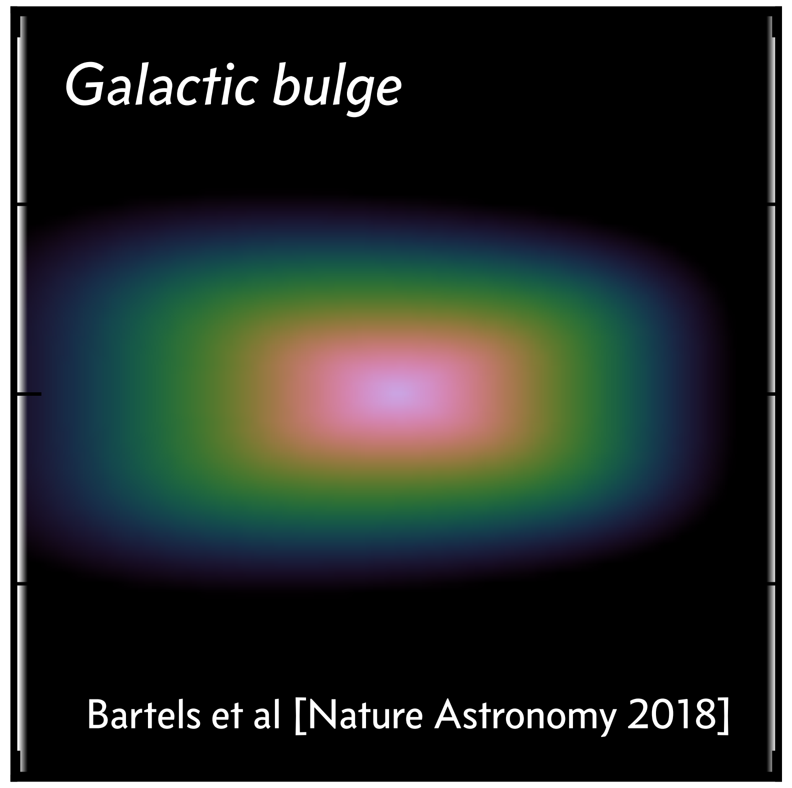


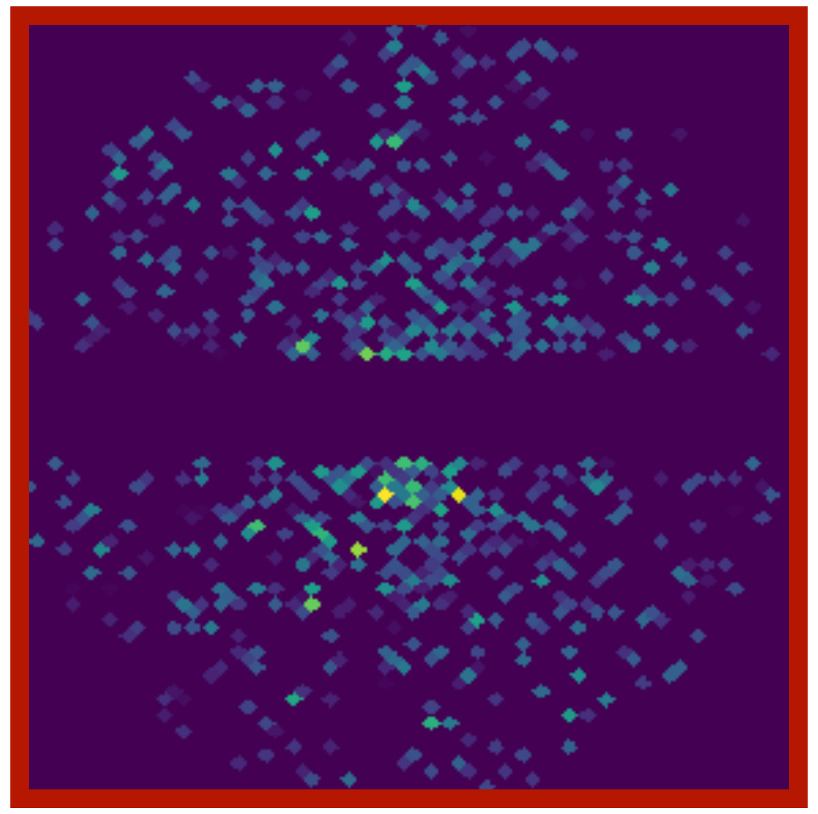
"Smoother" signal

Astrophysics

Spectrum and morphology consistent with millisecond pulsar expectation







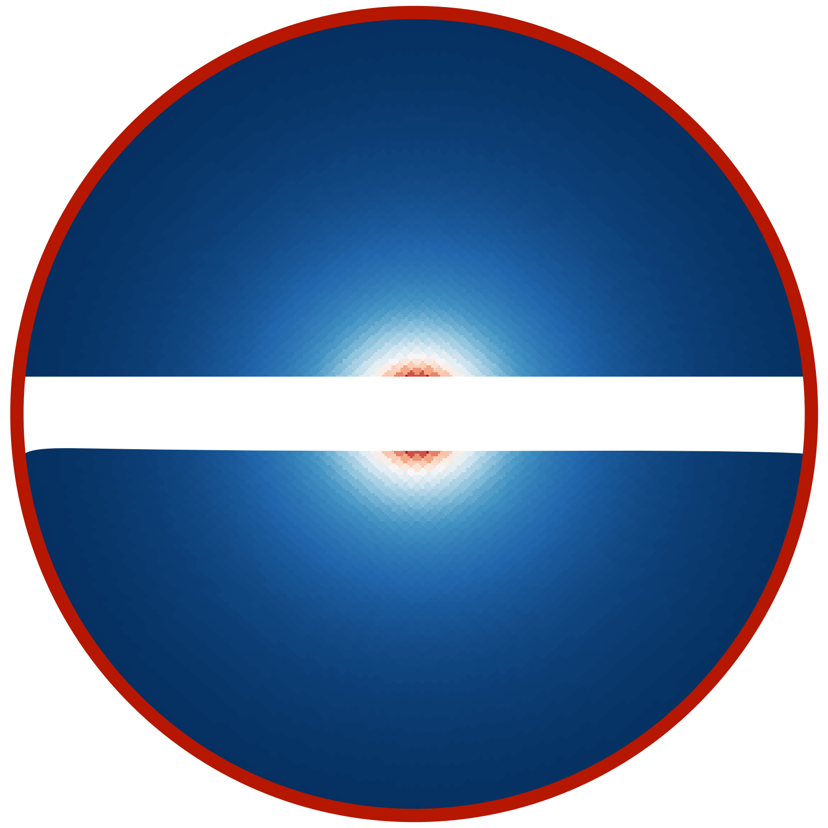
"Clumpier" signal

Modeling PS populations: ingredients

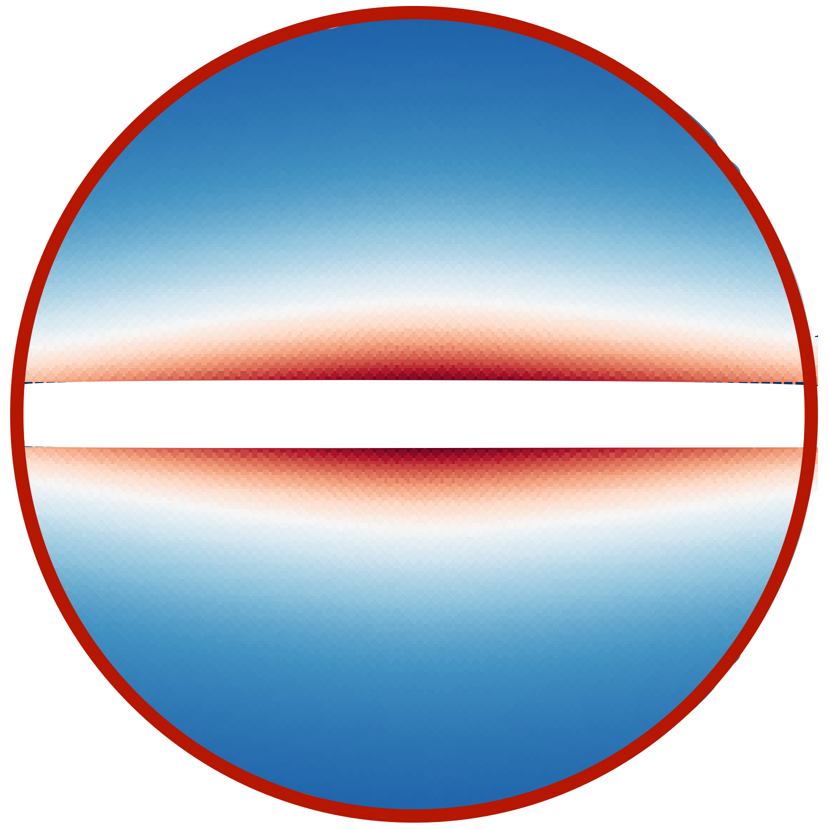
Spatial distribution of PSs

Specified by template

Galactic Center Excess



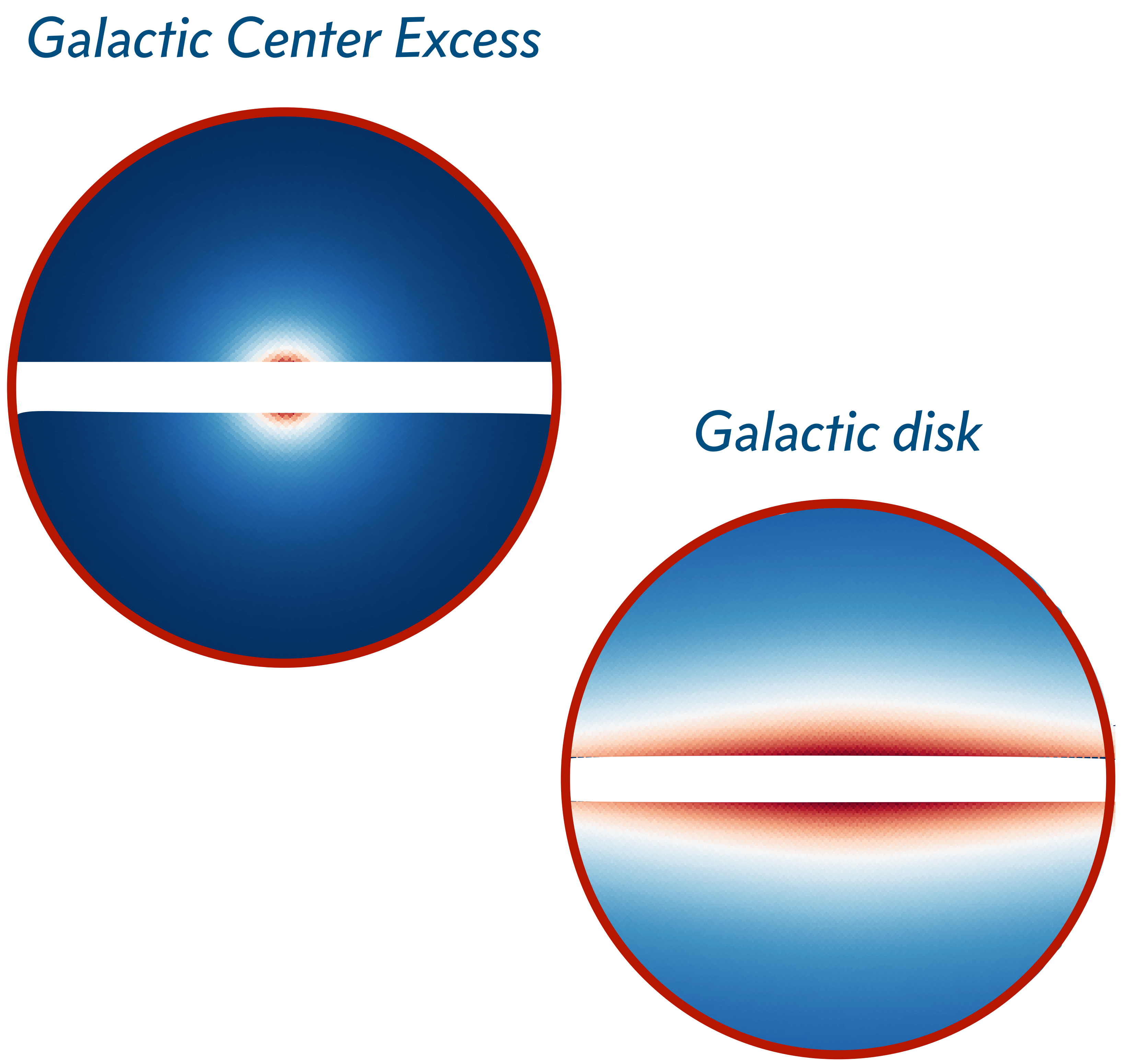
Galactic disk



Modeling PS populations: ingredients

Spatial distribution of PSs

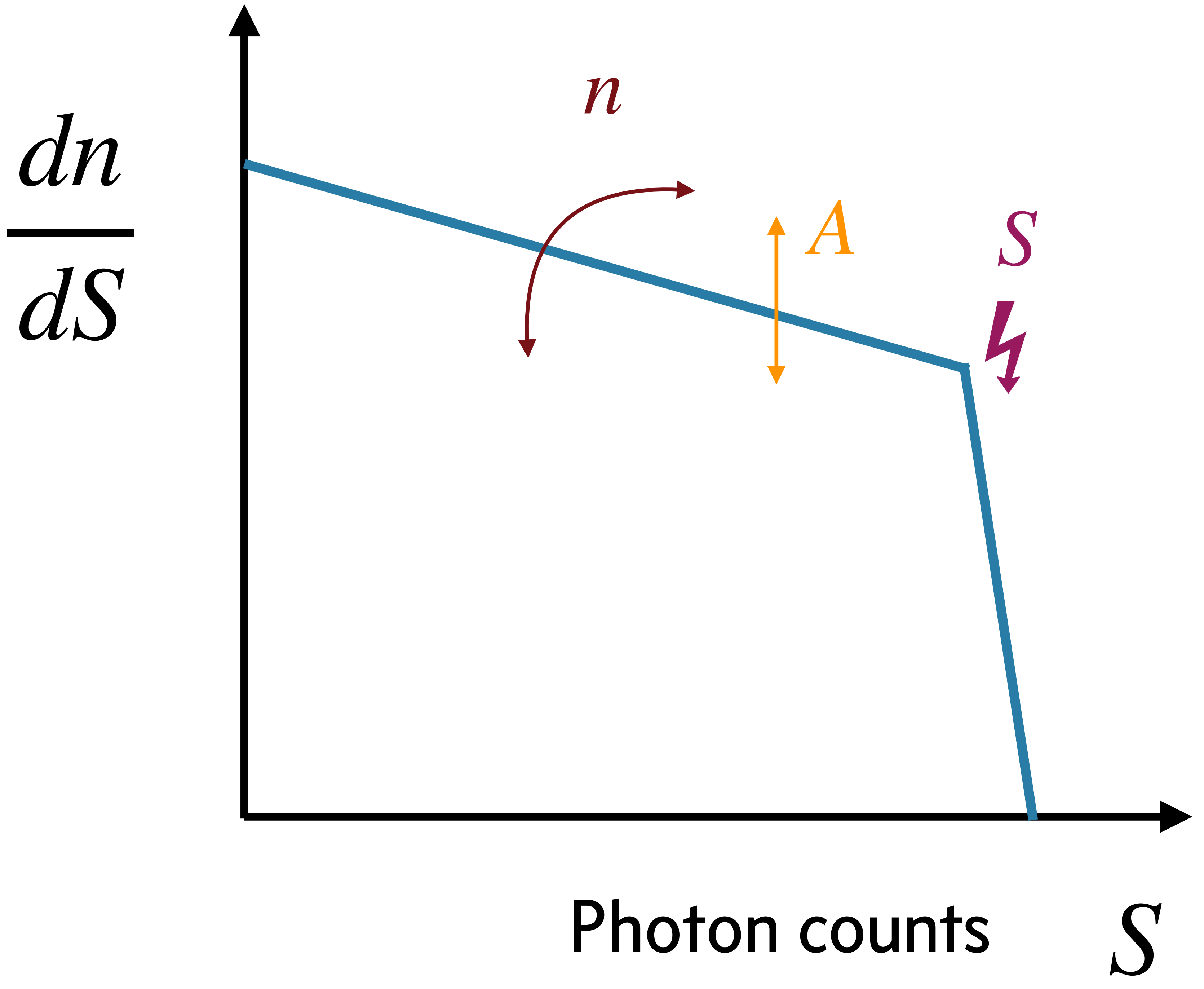
Specified by template



Counts distribution from PSs

Parameterize as a power law with breaks

$$\theta_{\mathrm{PS}} = \{A_{\mathrm{PS}}, n_1, n_2, n_3, S_1, S_2\}$$

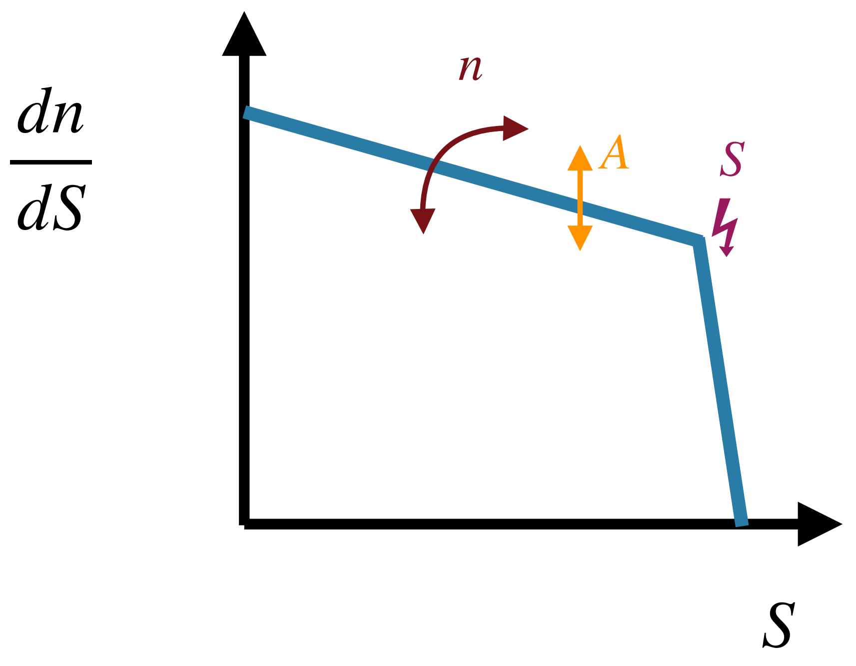


The likelihood function

Parameters of interest

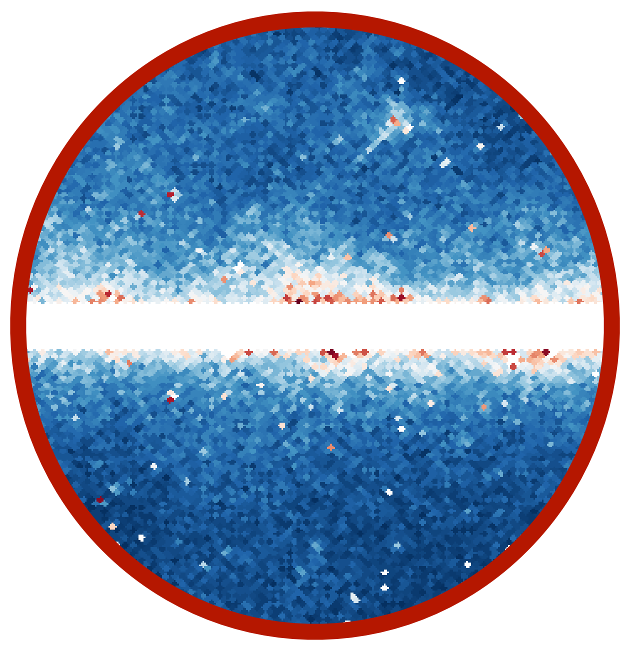
PS population parameters

$$\theta = \{A, n, S\}$$



Observables

γ-ray map x

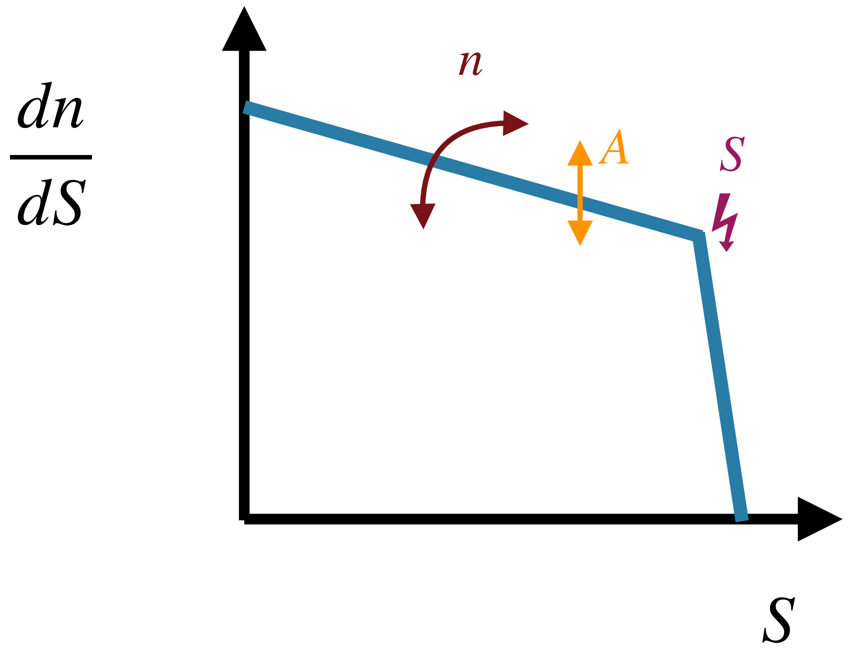


The likelihood function

Parameters of interest

PS population parameters

$$\theta = \{A, n, S\}$$



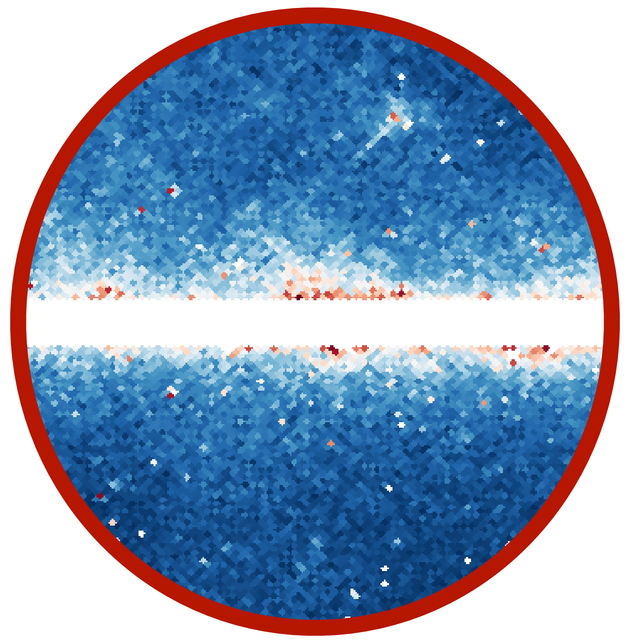
Latent variables

Individual PS properties

$$n_{\mathrm{PS}}, \{z_{\mathrm{PS},\,i}\}$$

Observables

γ-ray map x



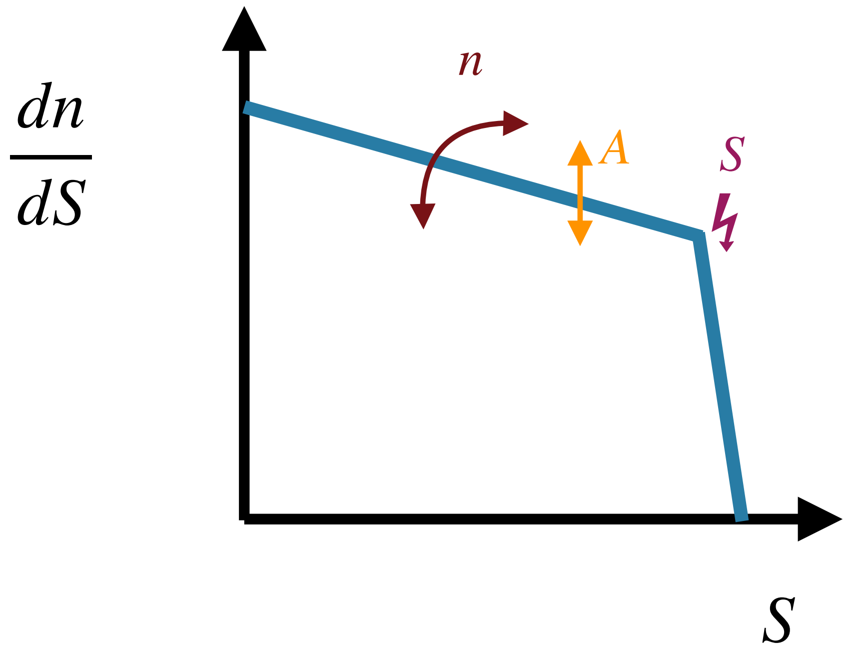
$$p\left(n_{\mathrm{PS}} \mid \theta_{\mathrm{PS}}\right) \prod_{i}^{n_{\mathrm{PS}}} p\left(z_{\mathrm{PS},\,i} \mid \theta_{\mathrm{PS}}; T_{\mathrm{PS}}\right) \quad \times \quad p\left(x \mid \theta_{\mathrm{smooth}}, \left\{z_{\mathrm{PS},\,i}\right\}\right)$$

The likelihood function

Parameters of interest

PS population parameters

$$\theta = \{A, n, S\}$$



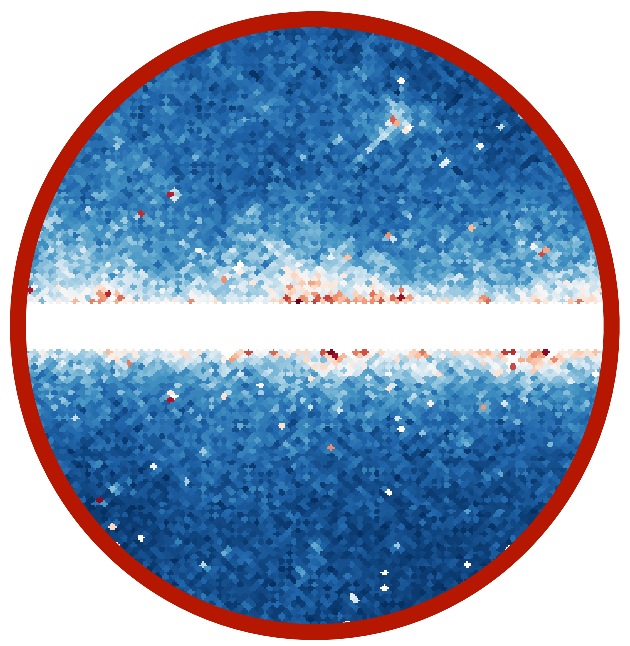
Latent variables

Individual PS properties

$$n_{\mathrm{PS}}, \{z_{\mathrm{PS},\,i}\}$$

Observables

γ-ray map x



We can easily write a simulator to sample from

$$p(x, z \mid \theta) = p\left(n_{\mathrm{PS}} \mid \theta_{\mathrm{PS}}\right) \prod_{i}^{n_{\mathrm{PS}}} p\left(z_{\mathrm{PS},\,i} \mid \theta_{\mathrm{PS}}; T_{\mathrm{PS}}\right) \times p\left(x \mid \theta_{\mathrm{smooth}}, \left\{z_{\mathrm{PS},\,i}\right\}\right)$$

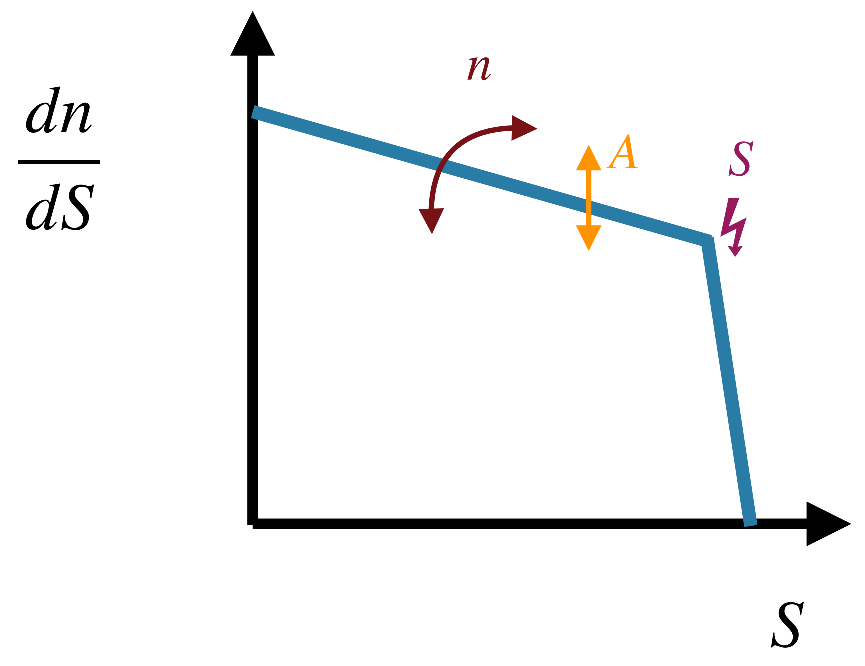
Prediction (Simulation)

The likelihood function

Parameters of interest

PS population parameters

$\theta = \{A, n, S\}$



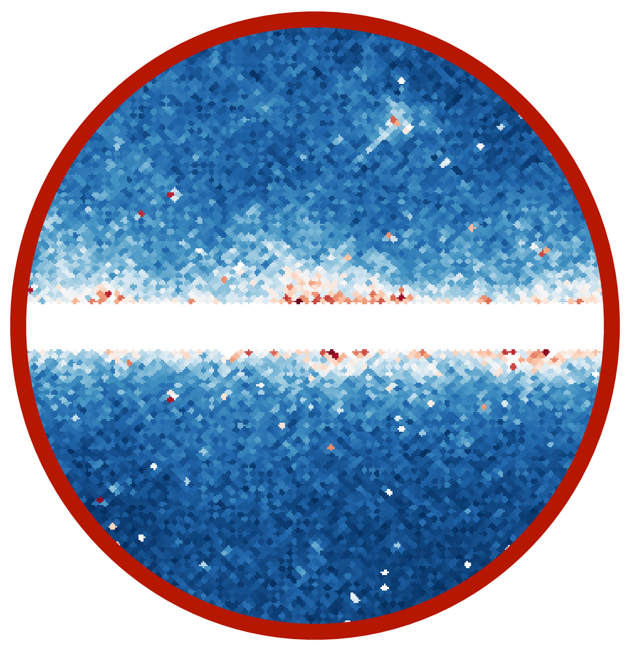
Latent variables

Individual PS properties

$n_{\mathrm{PS}}, \{z_{\mathrm{PS},i}\}$

Observables

γ-ray map x



$$p(x \mid \theta) = \sum_{n_{\mathrm{PS}}} \int \mathrm{d}^{n_{\mathrm{PS}}} z_{\mathrm{sub}} \quad p\left(n_{\mathrm{PS}} \mid \theta_{\mathrm{PS}}\right) \prod_{i}^{n_{\mathrm{PS}}} p\left(z_{\mathrm{PS},i} \mid \theta_{\mathrm{PS}}; T_{\mathrm{PS}}\right) \quad \times \quad p\left(x \mid \theta_{\mathrm{smooth}}, \left\{z_{\mathrm{PS},i}\right\}\right)$$

The key quantity for inference is the marginal likelihood

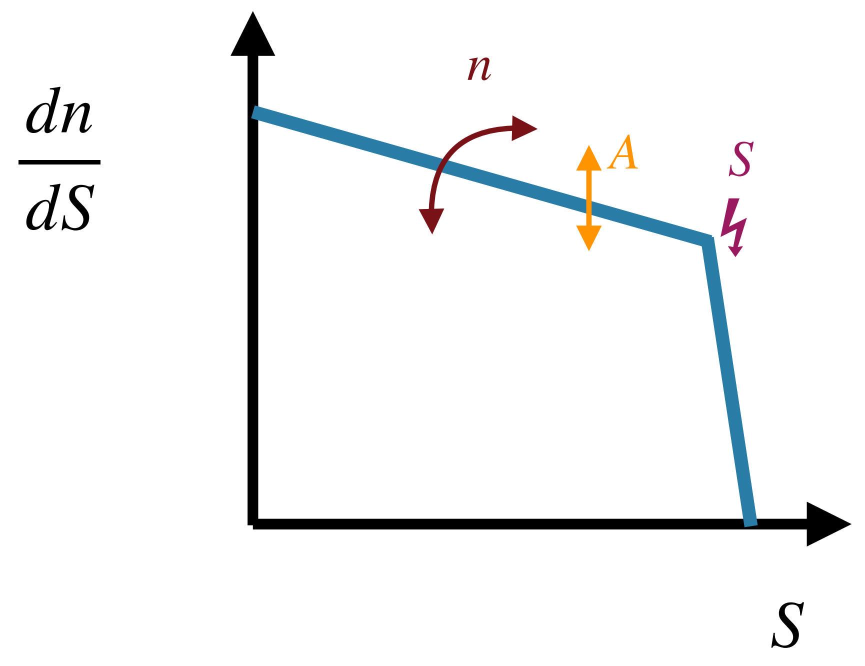
Inference

The likelihood function

Parameters of interest

PS population parameters

$$\theta = \{A, n, S\}$$



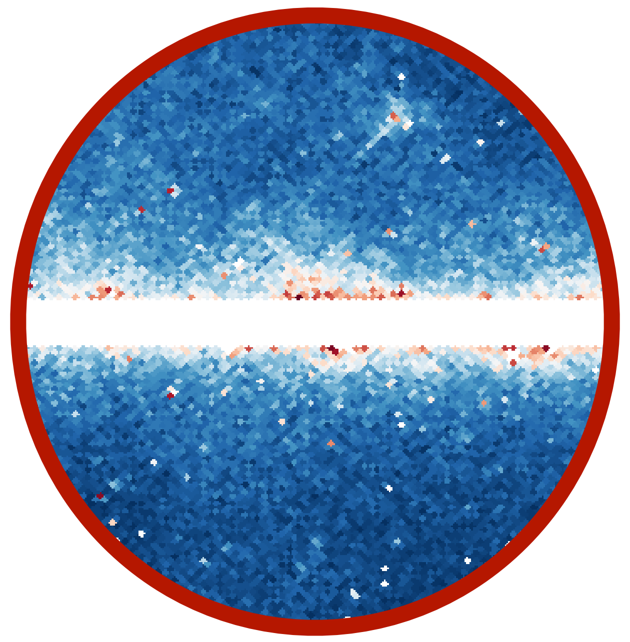
Latent variables

Individual PS properties

$$n_{\mathrm{PS}}, \{z_{\mathrm{PS},\,i}\}$$

Observables

γ-ray map x



$$\boxed{p(x \mid \theta) = \sum_{n_{\mathrm{PS}}} \int \mathrm{d}^{n_{\mathrm{PS}}} z_{\mathrm{sub}}} \; p\left(n_{\mathrm{PS}} \mid \theta_{\mathrm{PS}}\right) \prod_{i}^{n_{\mathrm{PS}}} p\left(z_{\mathrm{PS},\,i} \mid \theta_{\mathrm{PS}}; T_{\mathrm{PS}}\right) \times p\left(x \mid \theta_{\mathrm{smooth}}, \left\{z_{\mathrm{PS},\,i}\right\}\right)$$

The key quantity for inference is the marginal likelihood

Inference

Simplifying the problem: pixel-wise conditional independence

Assume pixel-wise conditional independence $\Longrightarrow$ model photon counts PDF as a doubly-stochastic Poisson process

NPTF = *Non-Poissonian Template Fitting*

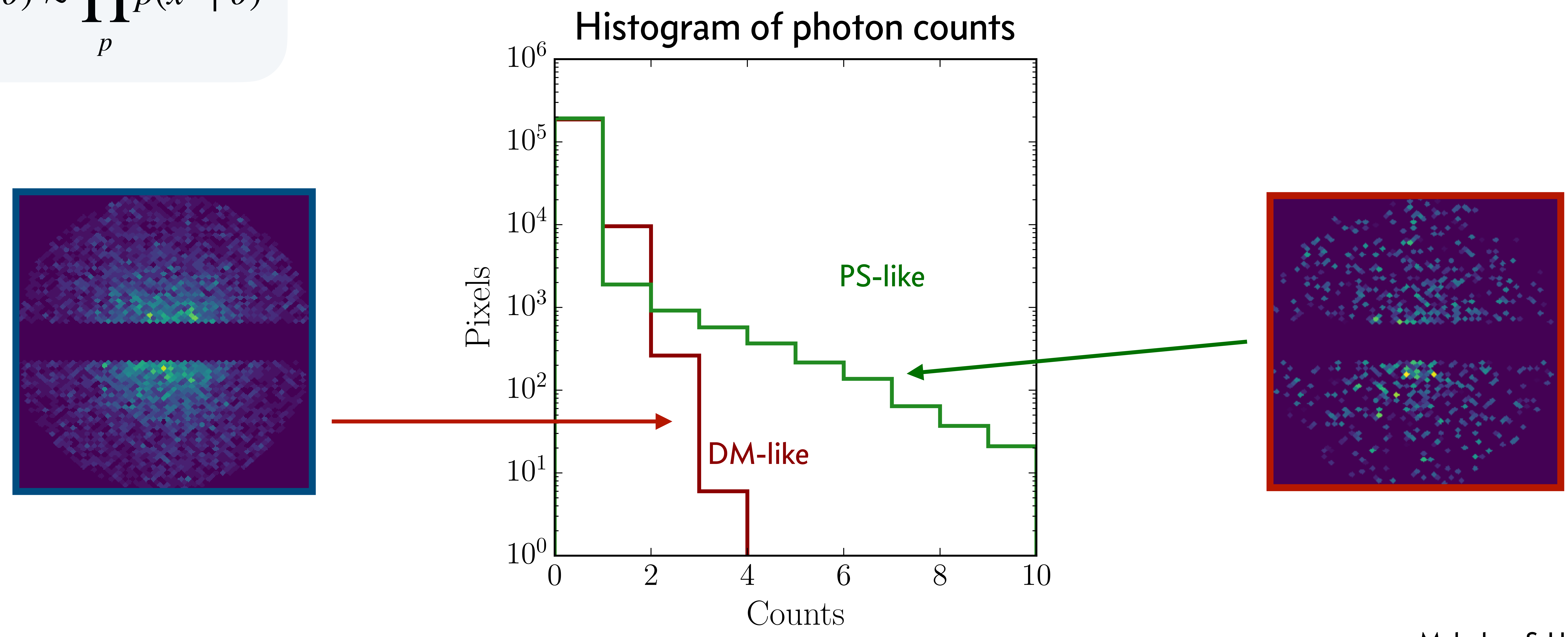
$$p(x \mid \theta) \approx \prod_p p(x^p \mid \theta)$$

Simplifying the problem: pixel-wise conditional independence

Assume pixel-wise conditional independence $\Longrightarrow$ model photon counts PDF as a doubly-stochastic Poisson process

NPTF = *Non-Poissonian Template Fitting*

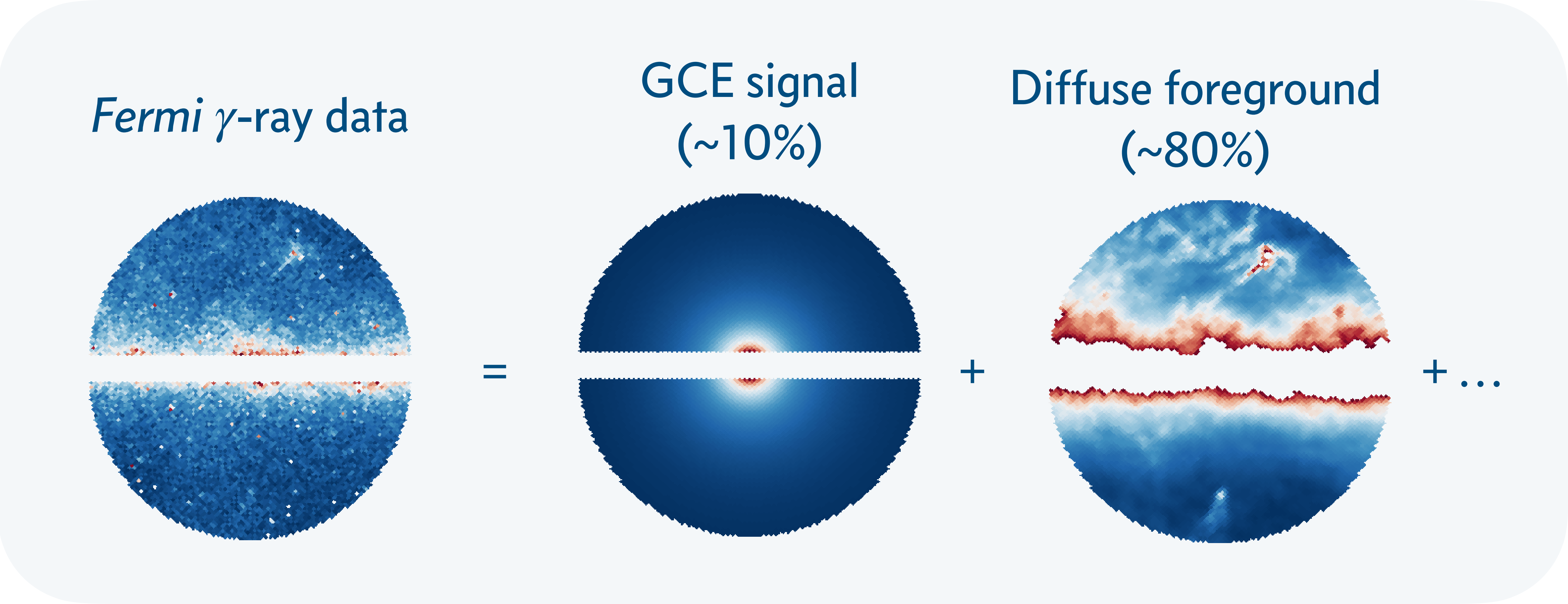
$$p(x \mid \theta) \approx \prod_p p(x^p \mid \theta)$$



Malyshev & Hogg [ApJ 2011]
Lee et al [JCAP 2014]

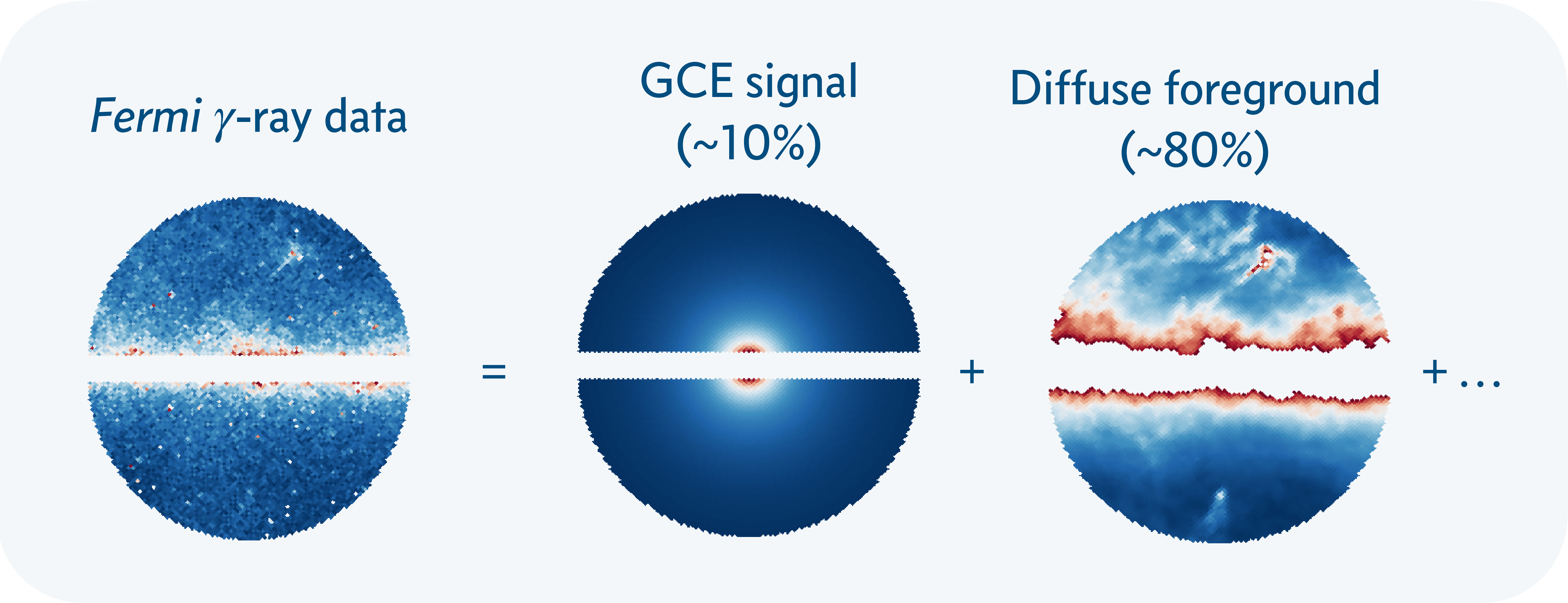
What could go wrong?

Leane & Slatyer [PRL 2019]
+ Leane & Slatyer [PRL 2020, PRD 2020]

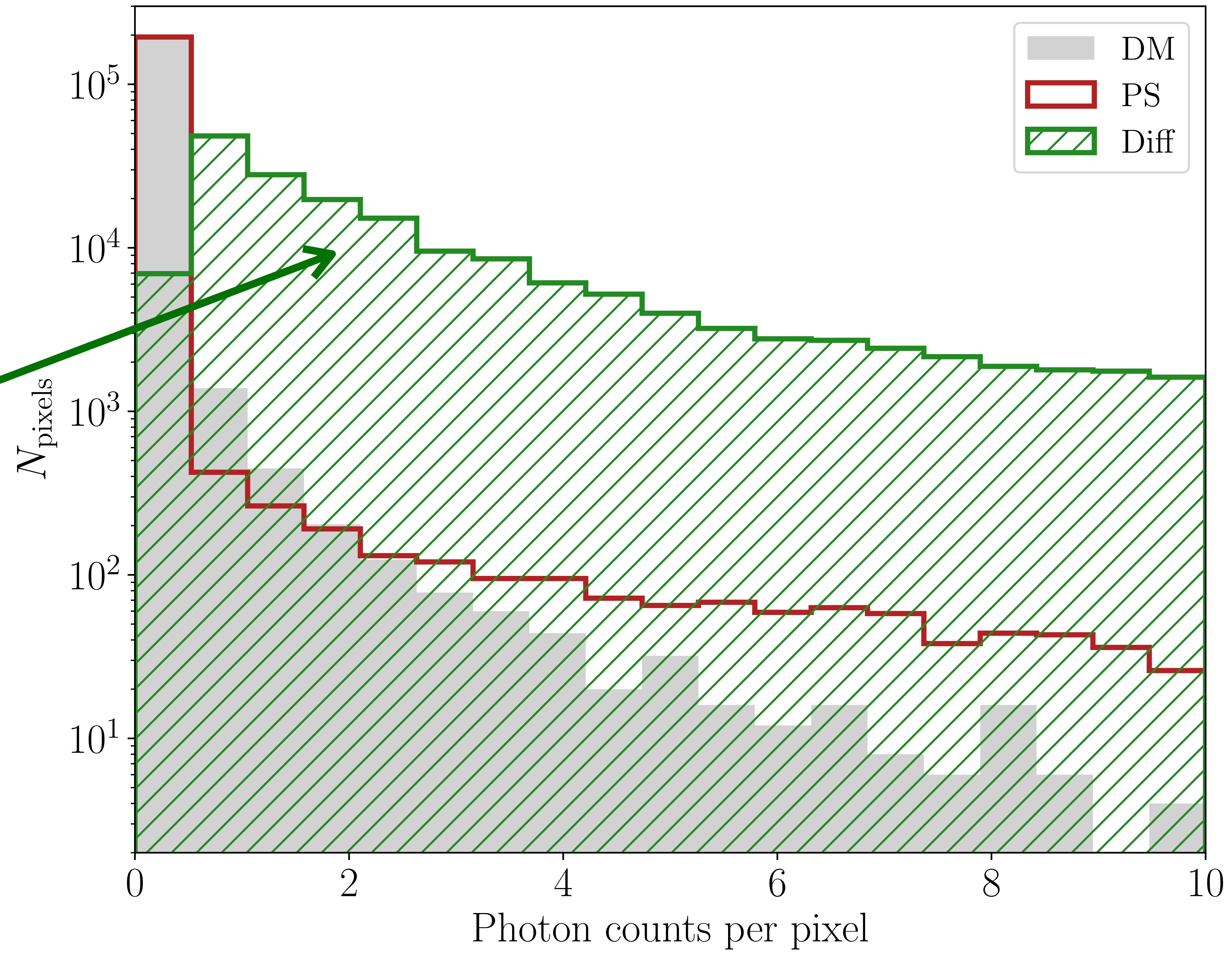


What could go wrong?

Leane & Slatyer [PRL 2019]
+ Leane & Slatyer [PRL 2020, PRD 2020]

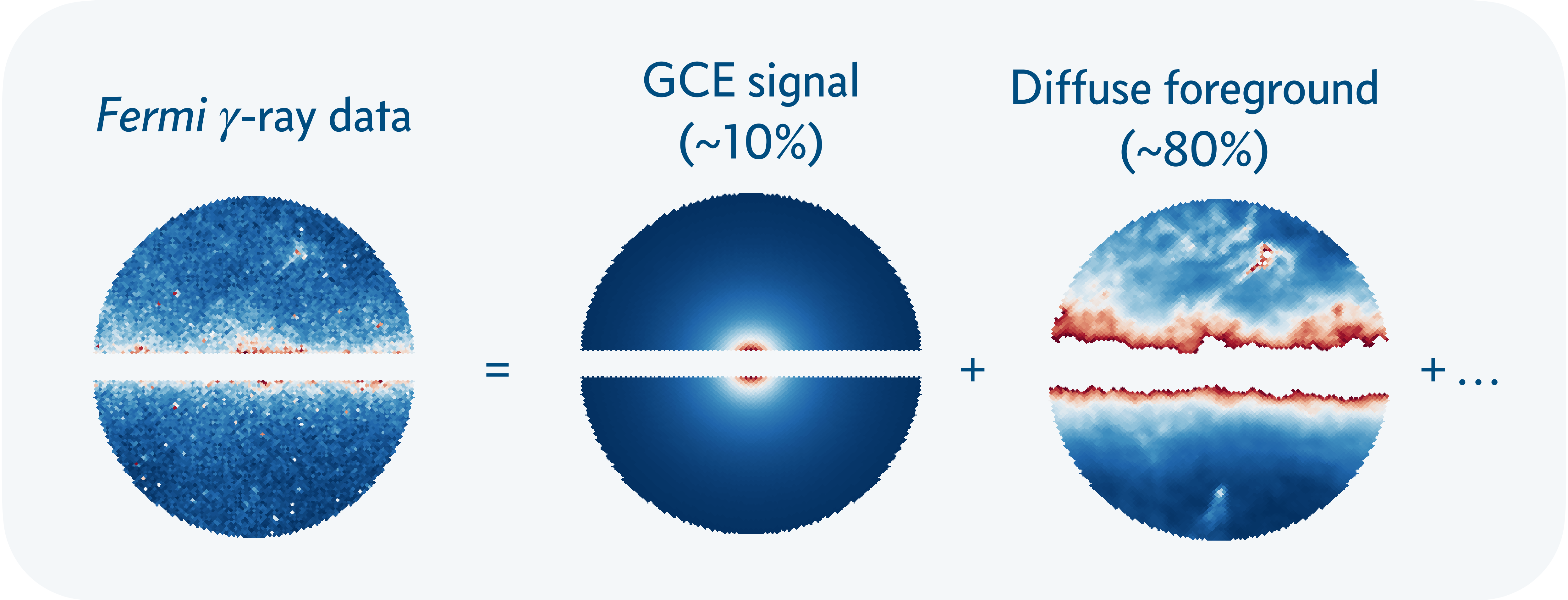


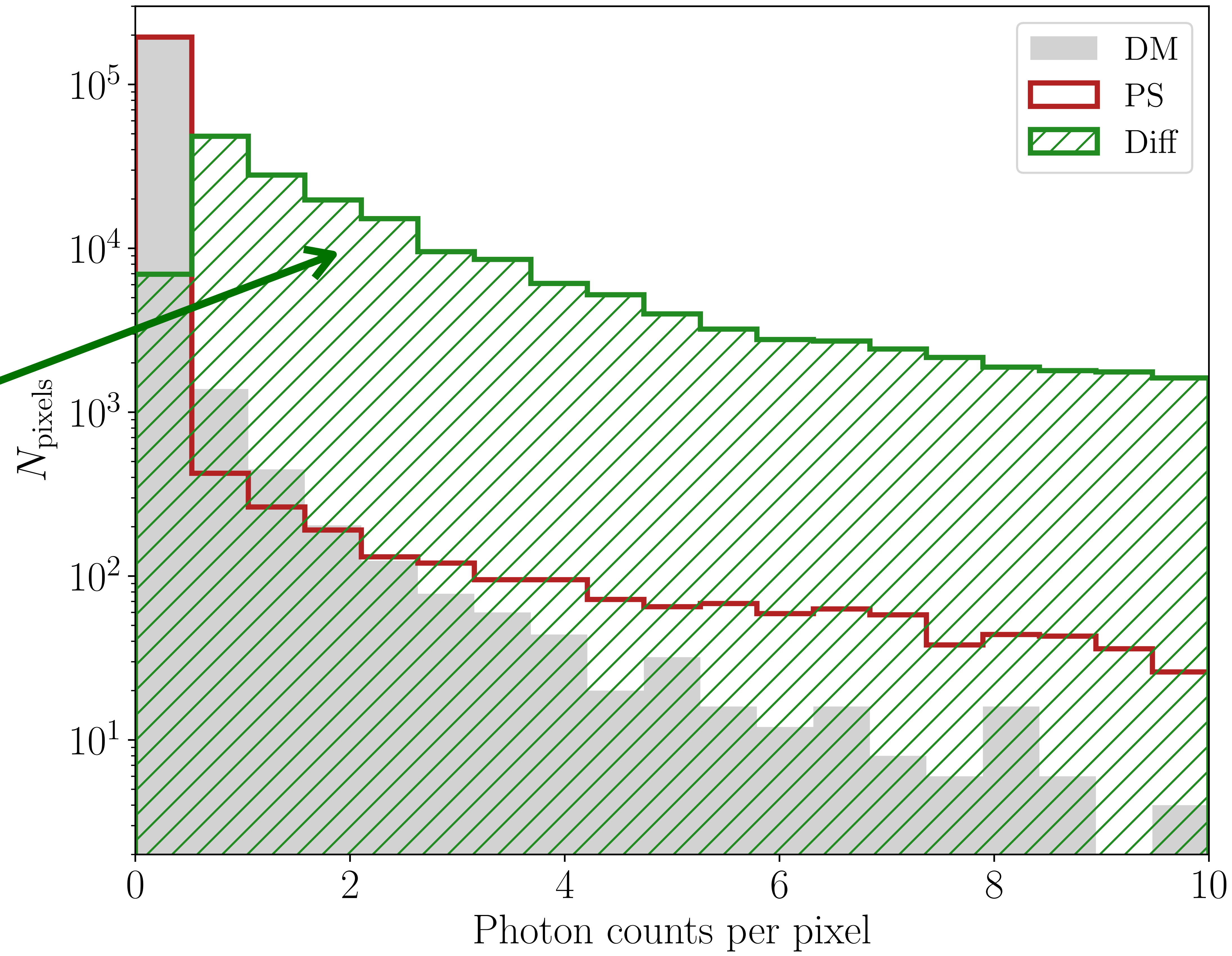
Diffuse foregrounds make up most of the observed emission in the Galactic Center



What could go wrong?

Leane & Slatyer [PRL 2019]
+ Leane & Slatyer [PRL 2020, PRD 2020]





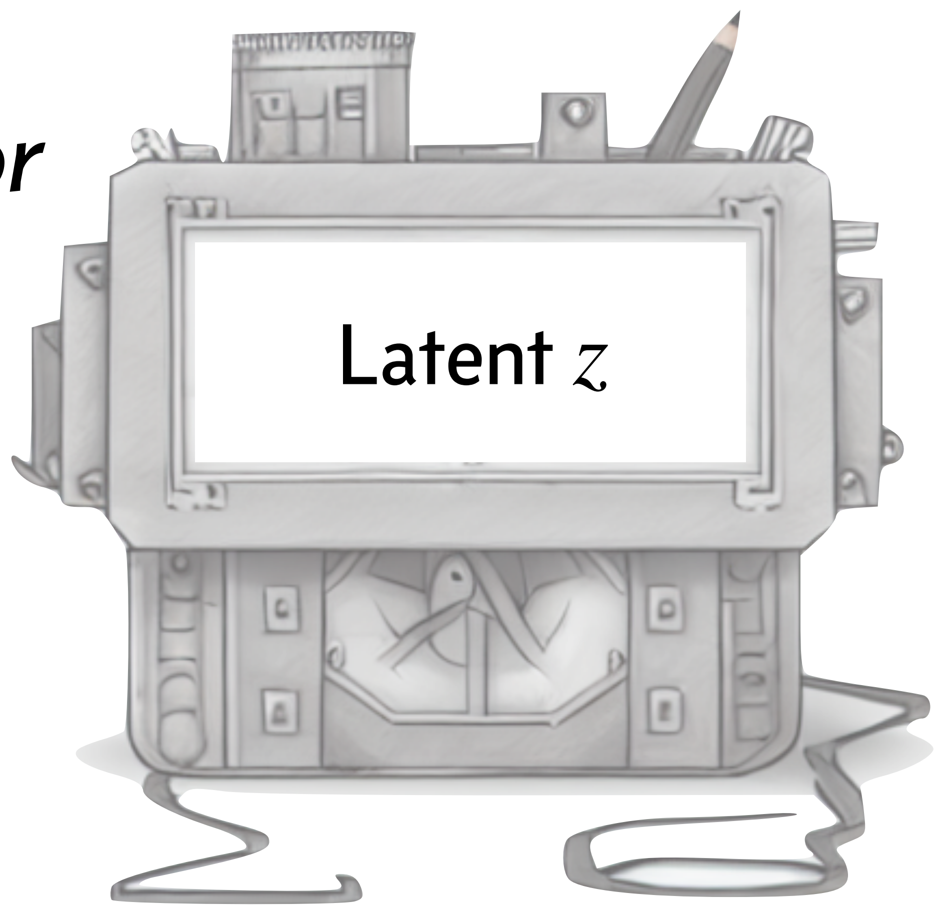
Diffuse foregrounds make up most of the observed emission in the Galactic Center

Promising direction: *build/apply better diffuse models*

Buschmann et al [PRD 2020]
Pohl et al [ApJ 2022], Macias et al [JCAP 2019]

Simulation-based inference (SBI)

Simulator

Parameters θ →  → Observations x

Prediction:

- Well-motivated mechanistic, causal model
- 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
- *Inference is challenging*

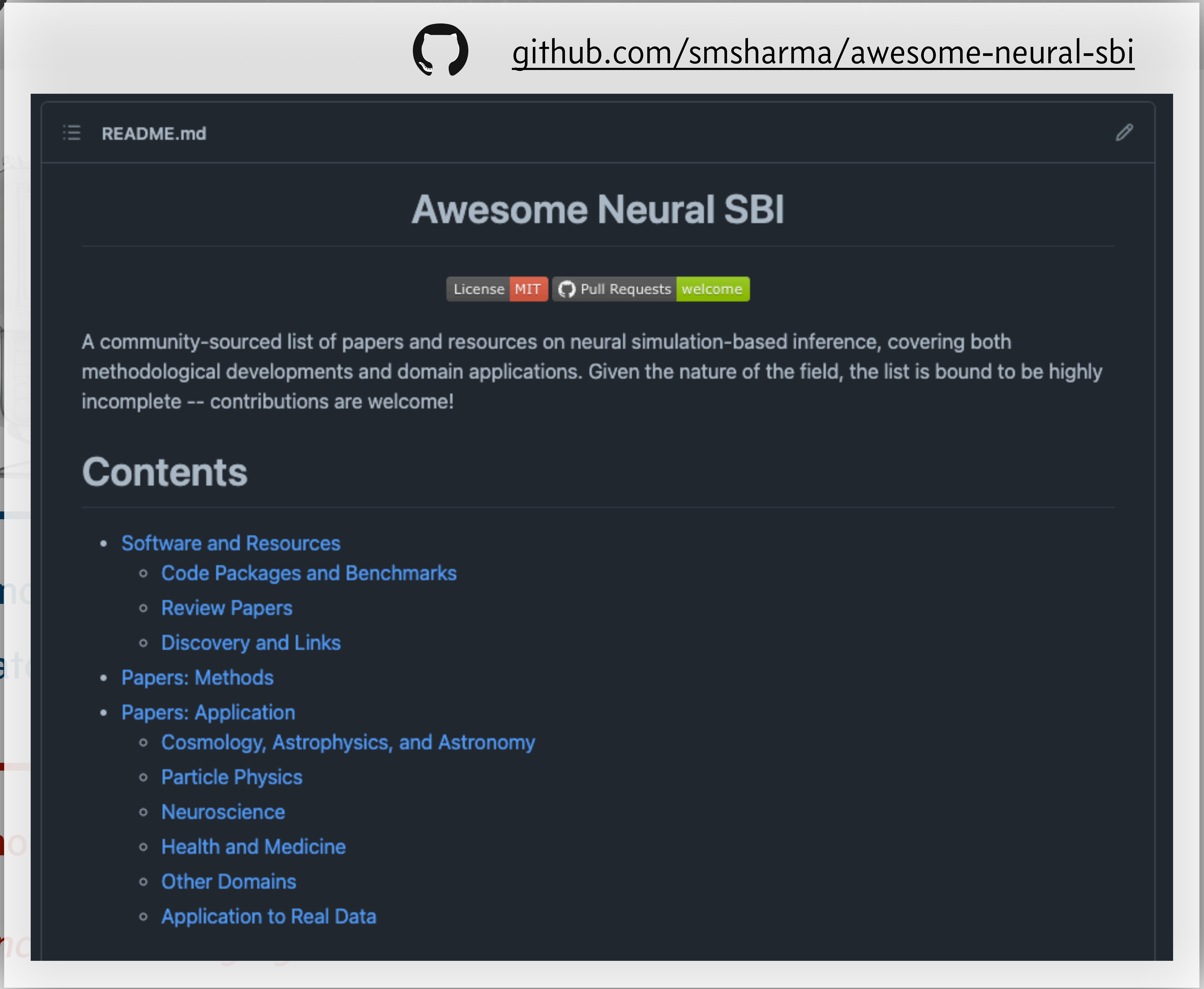
Simulation-based inference (SBI)

Simulator

Parameters
θ

Prediction:
- Well-m
- Simula

Inference:
- Likelih
- *Inferen*


github.com/smsharma/awesome-neural-sbi

README.md

Awesome Neural SBI

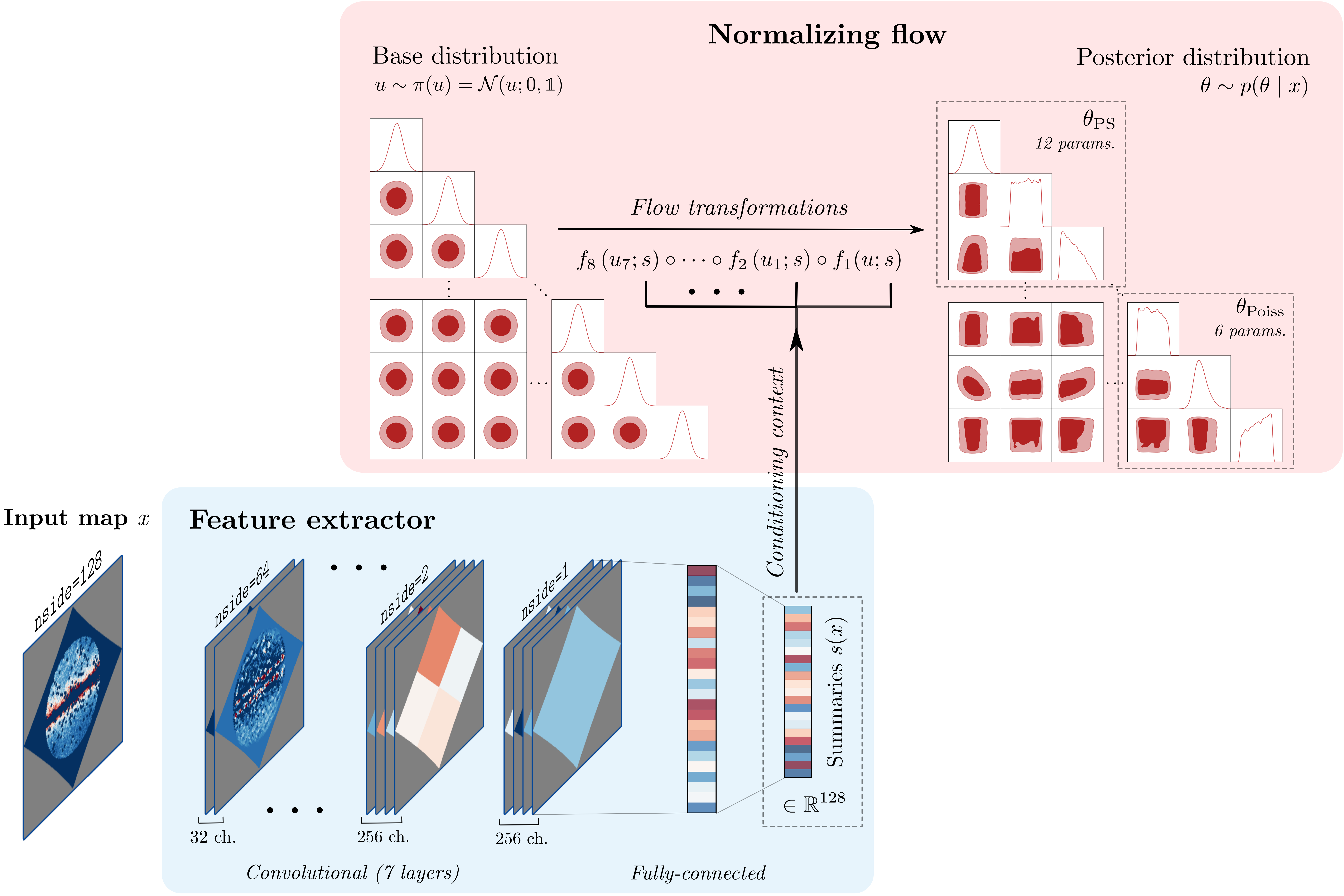
License MIT | Pull Requests welcome

A community-sourced list of papers and resources on neural simulation-based inference, covering both methodological developments and domain applications. Given the nature of the field, the list is bound to be highly incomplete -- contributions are welcome!

Contents

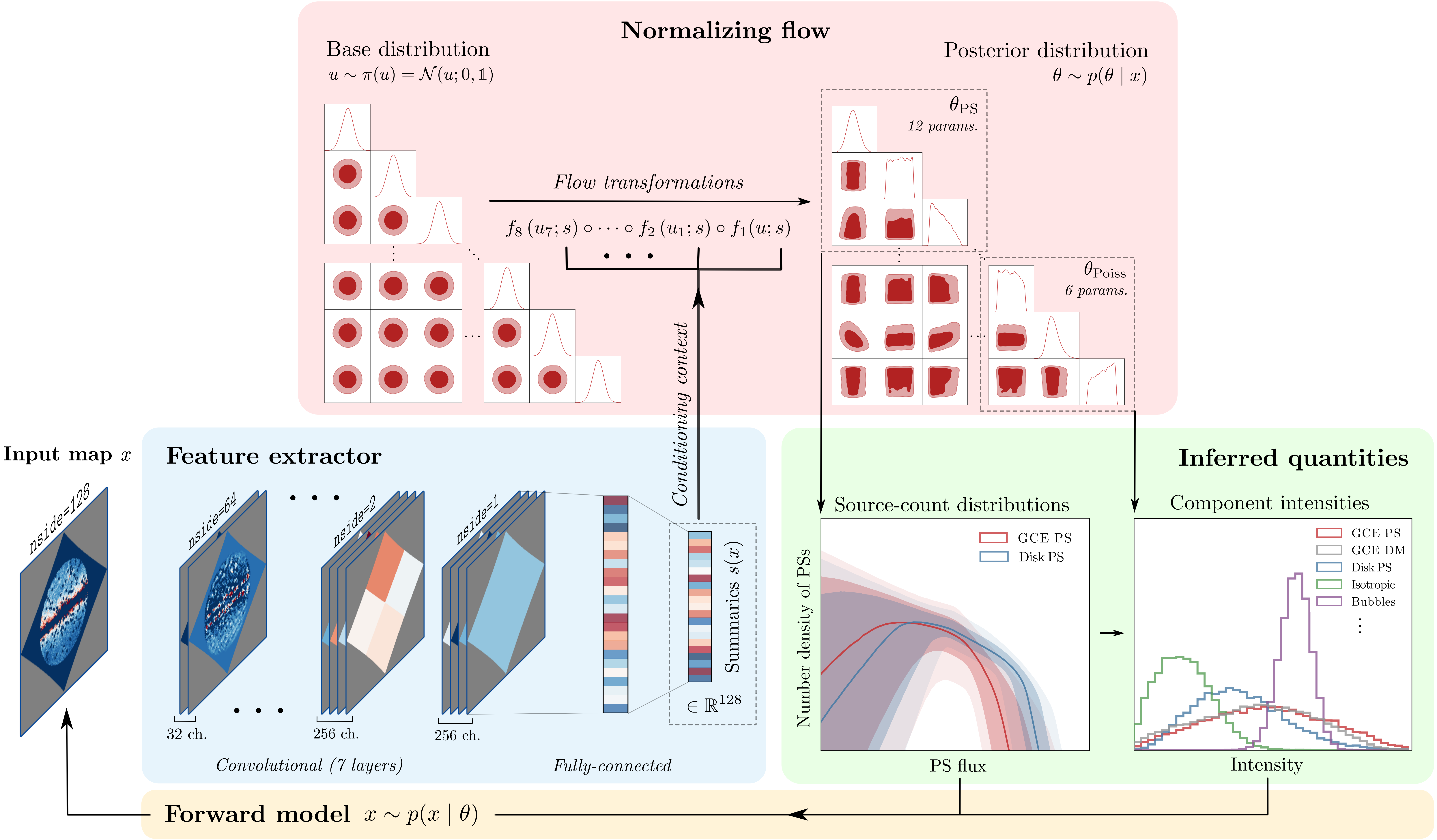
Going beyond the counts PDF: *neural posterior estimation*

SM, Cranmer [PRD 2022]



Going beyond the counts PDF: *neural posterior estimation*

SM, Cranmer [PRD 2022]



Going beyond the counts PDF: *neural posterior estimation*

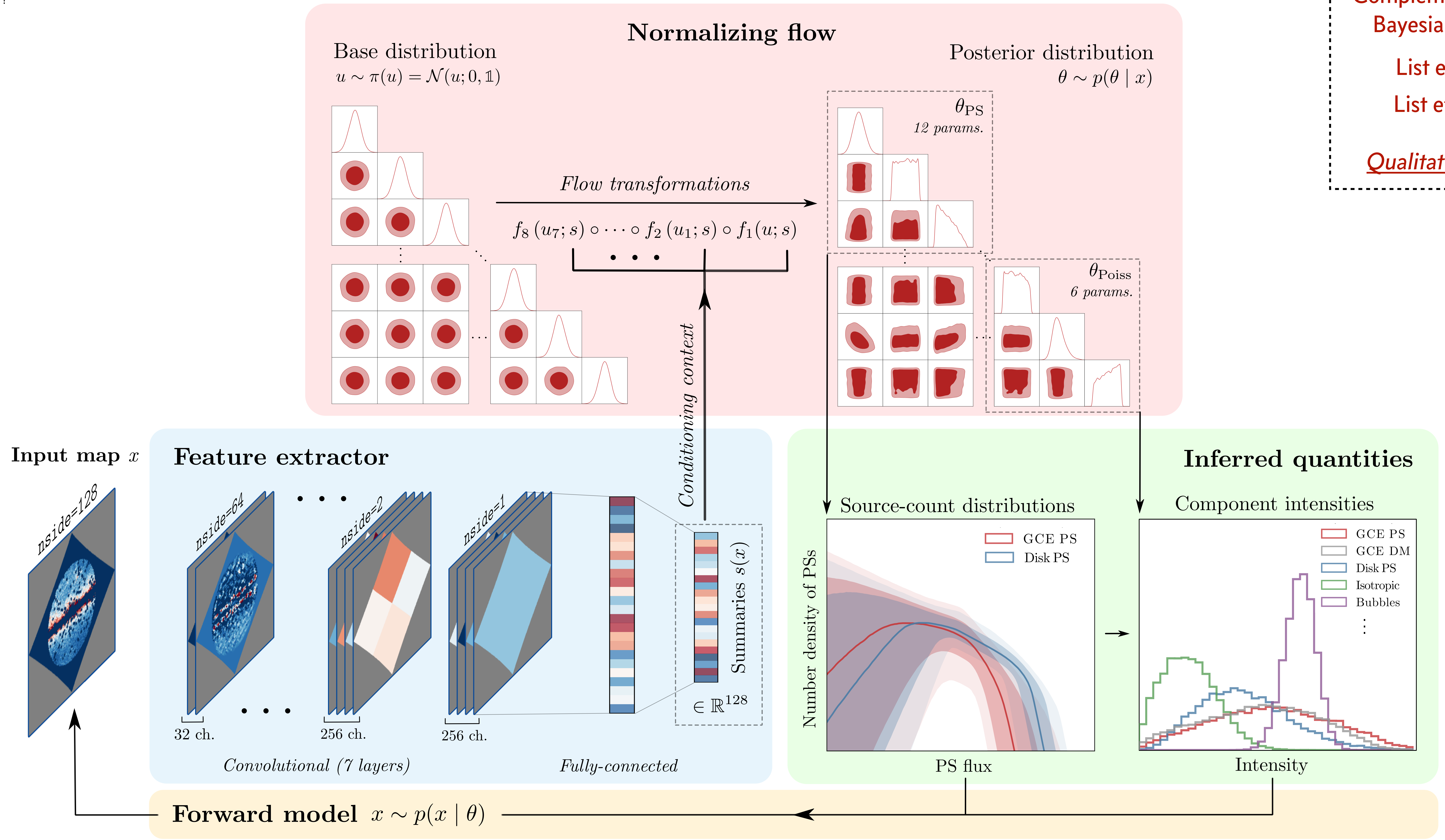
SM, Cranmer [PRD 2022]

Complementary method using Bayesian neural networks:

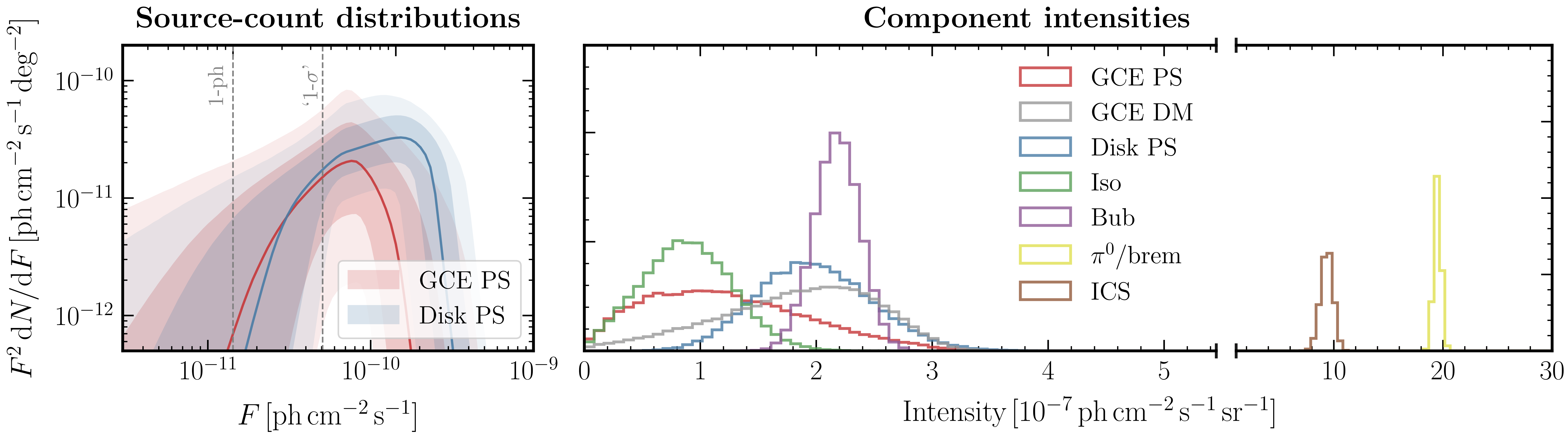
List et al [PRL 2021]

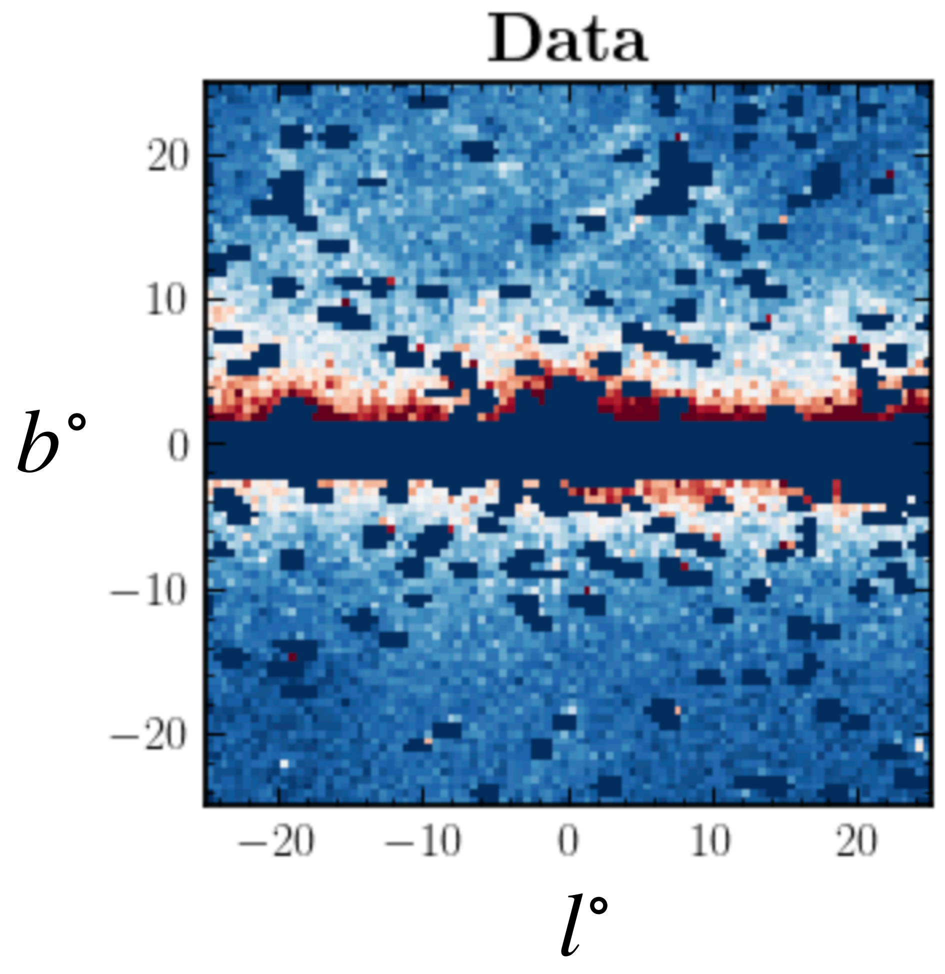
List et al [PRD 2021]

Qualitatively similar results

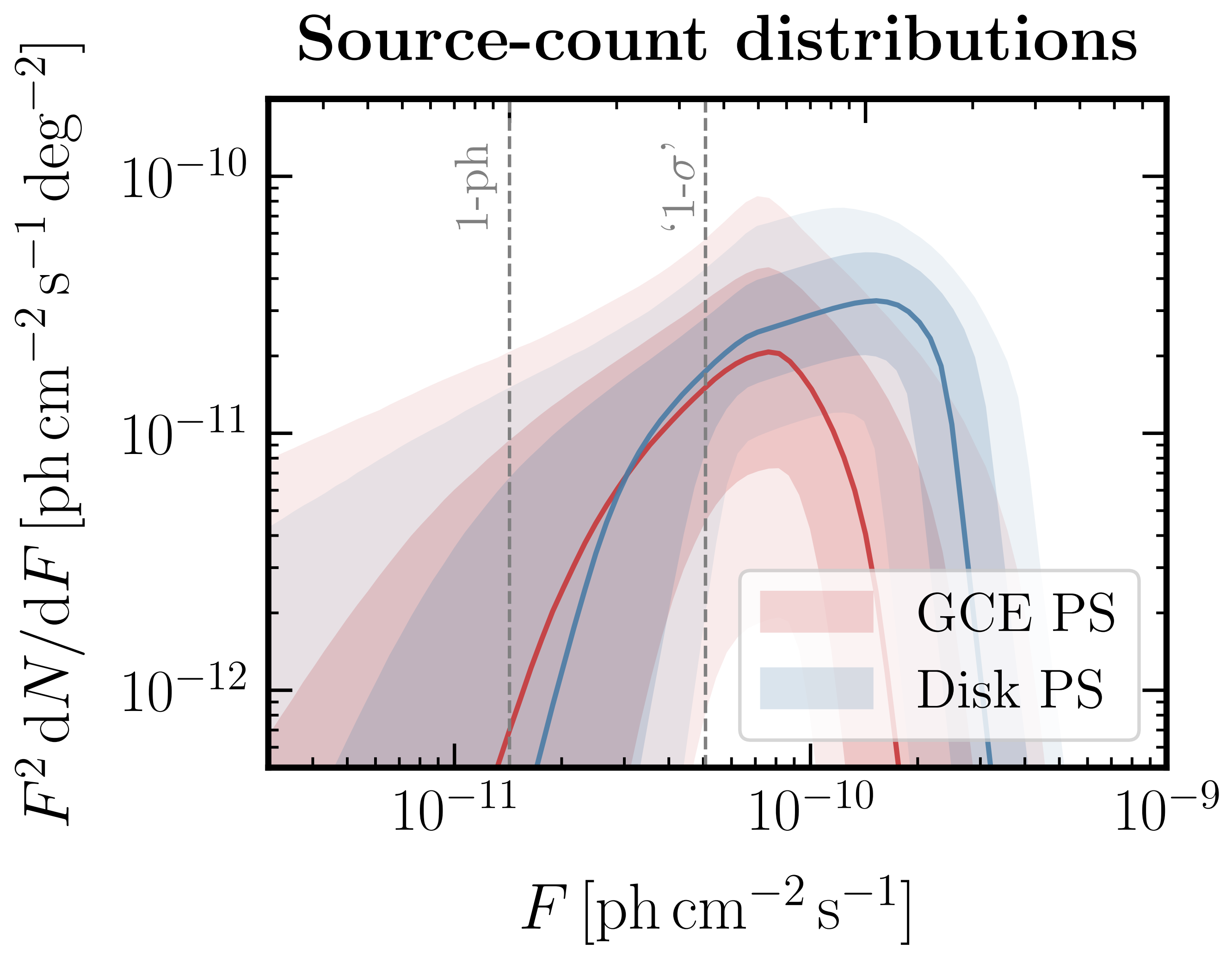


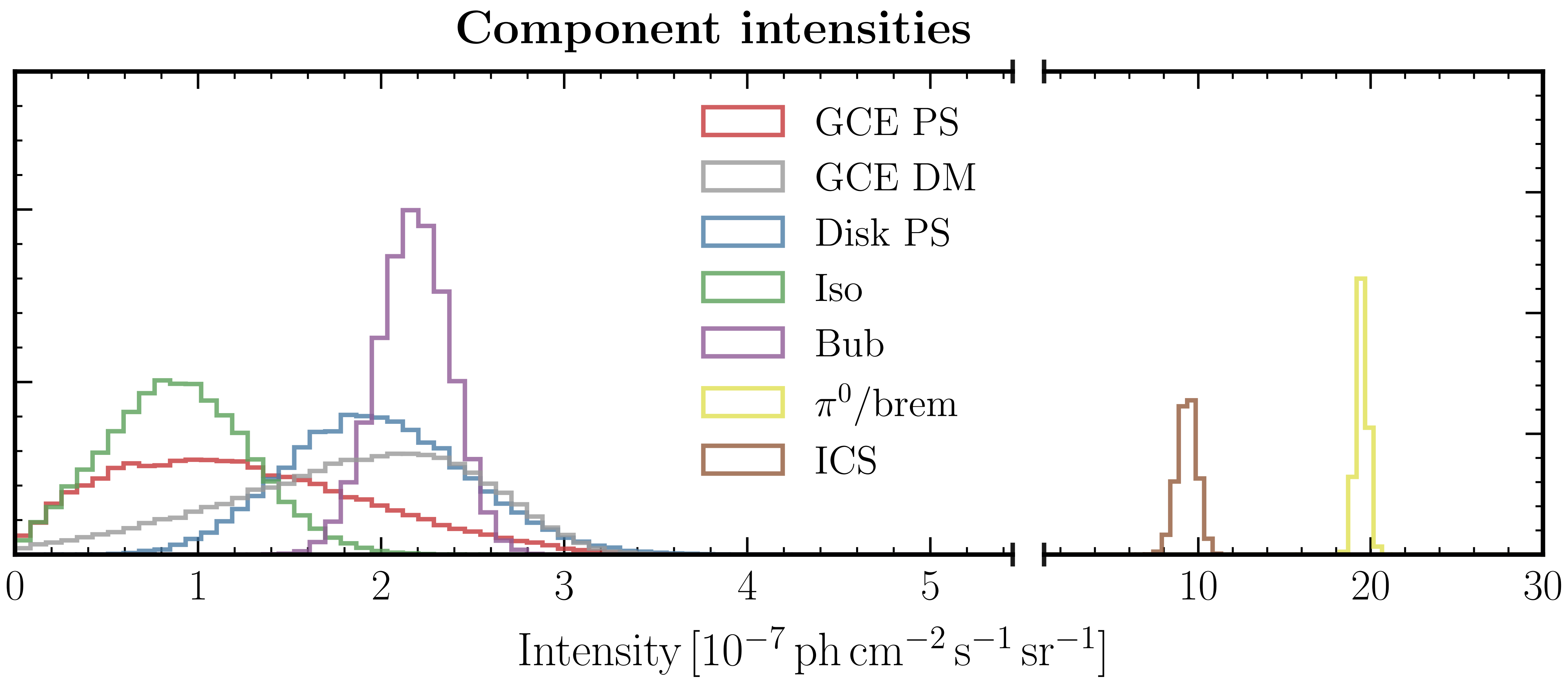
Application to *Fermi* γ-ray data

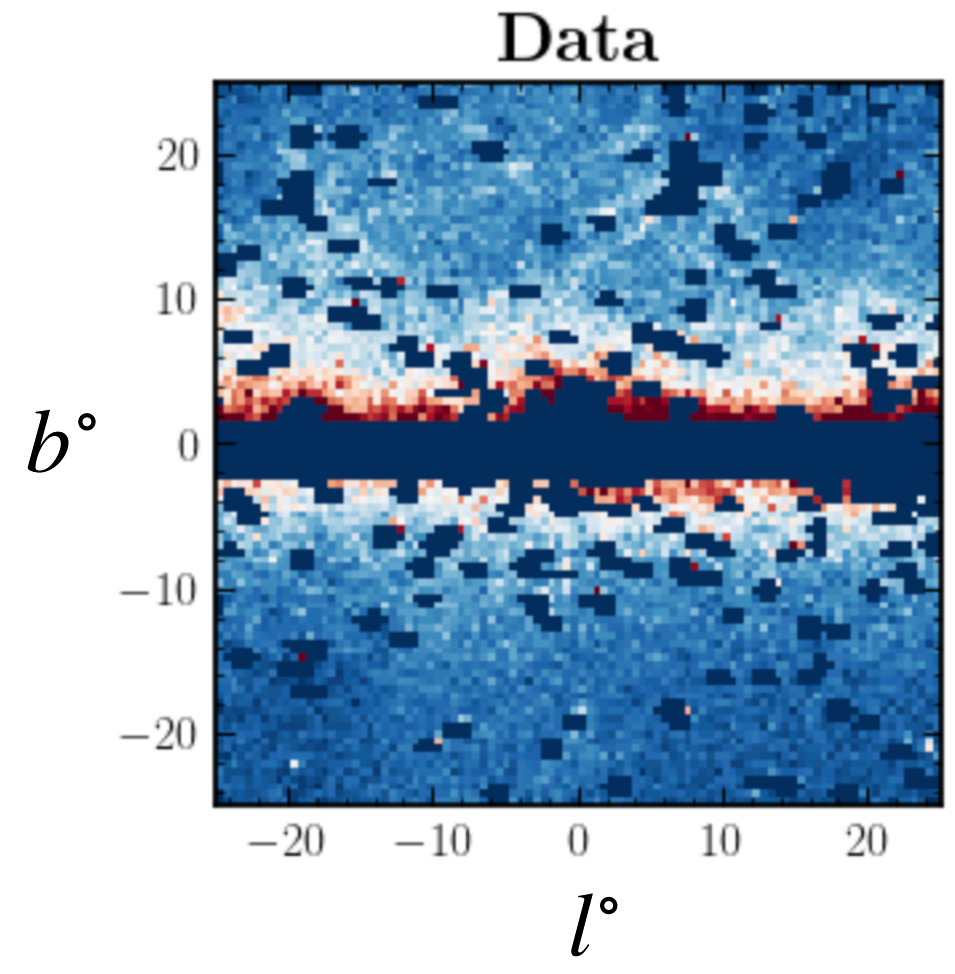


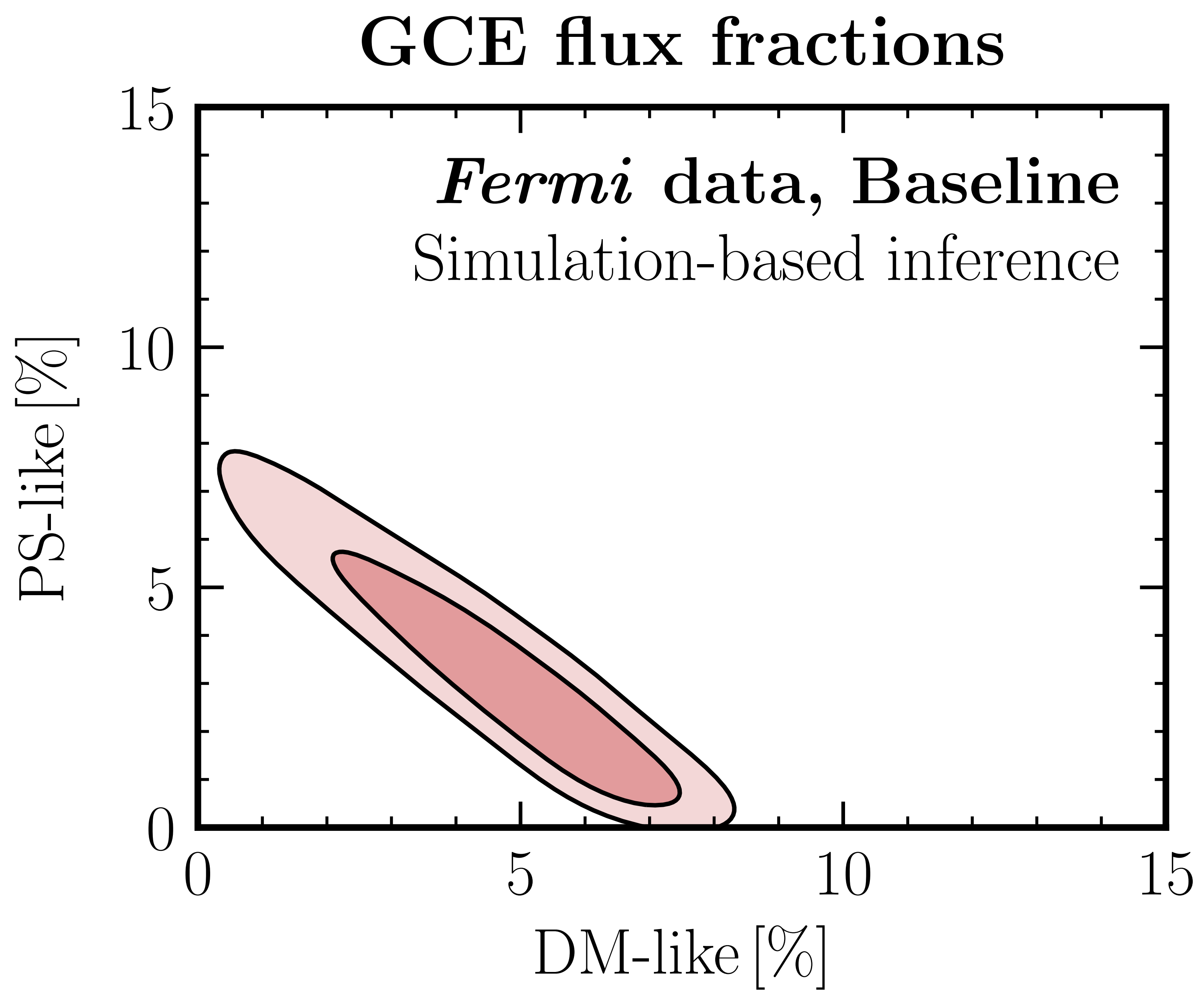


Application to *Fermi* γ-ray data





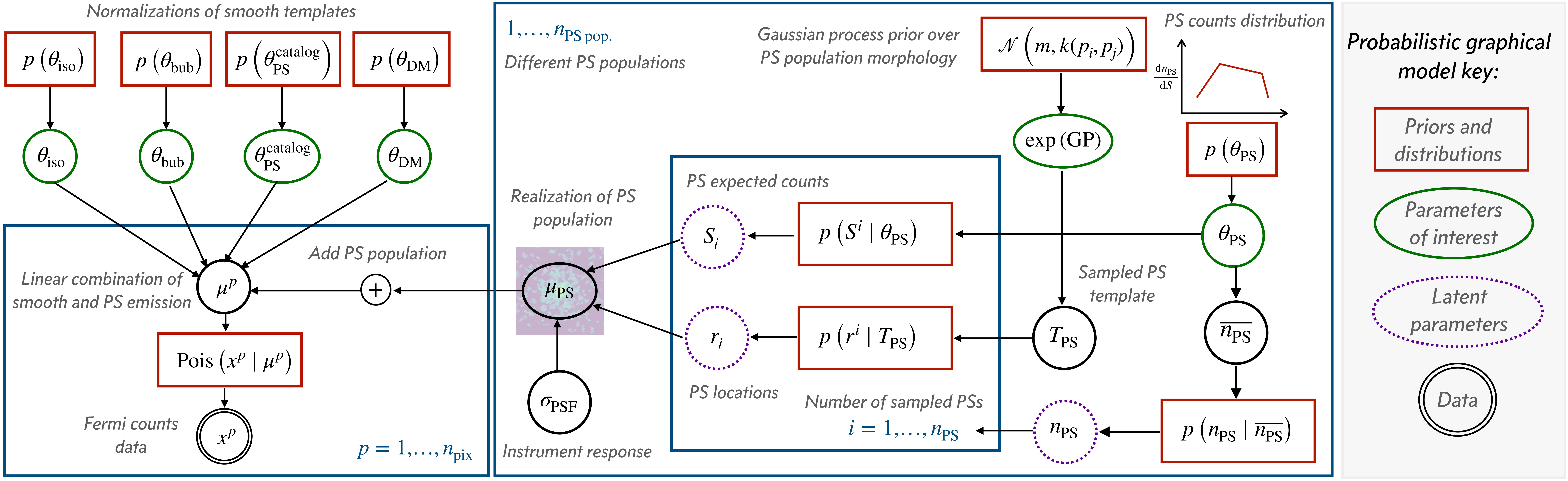




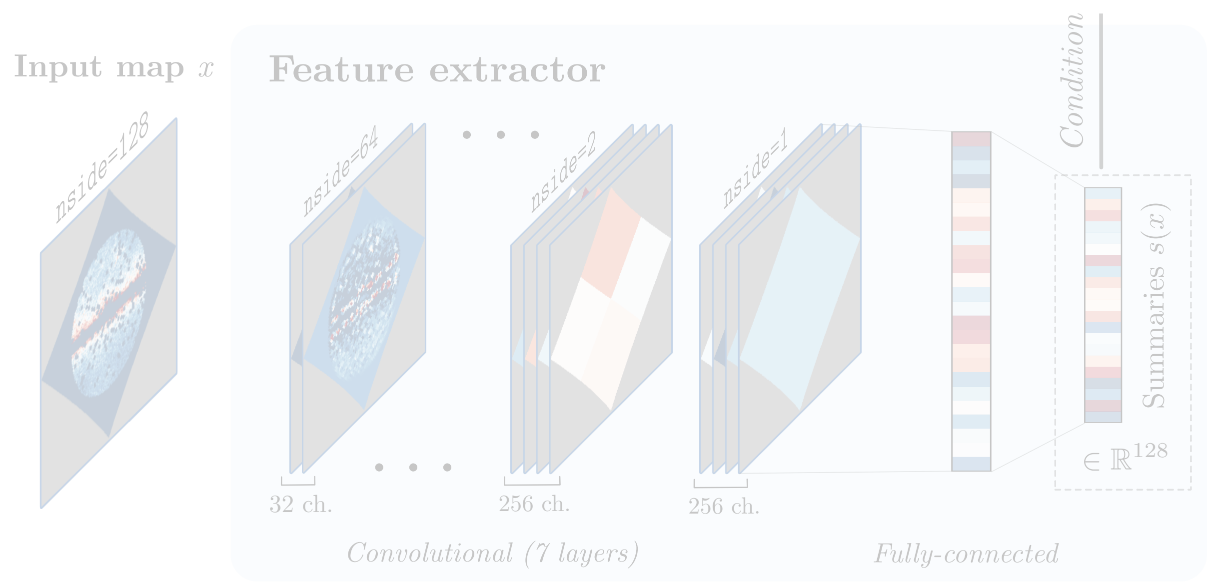
Exploiting more information in the γ-ray maps results in smaller PS component compared to NPTF

Differentiable probabilistic programming: *the future?*

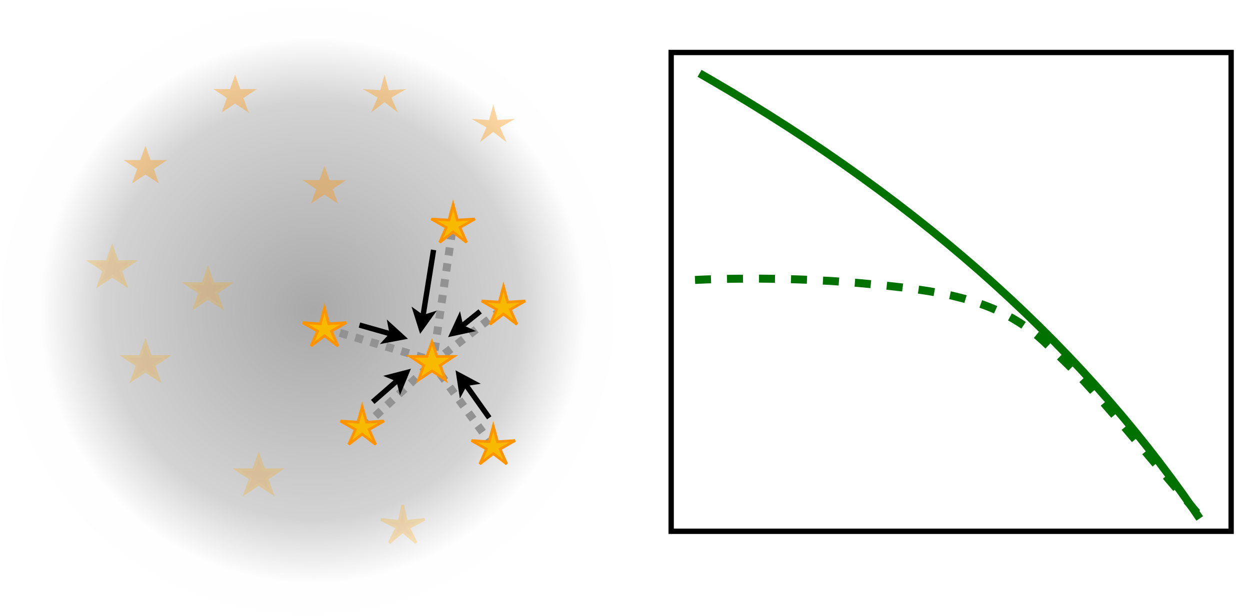
Differentiable probabilistic programming: *the future?*



Outline



Characterizing the Galactic Center Excess

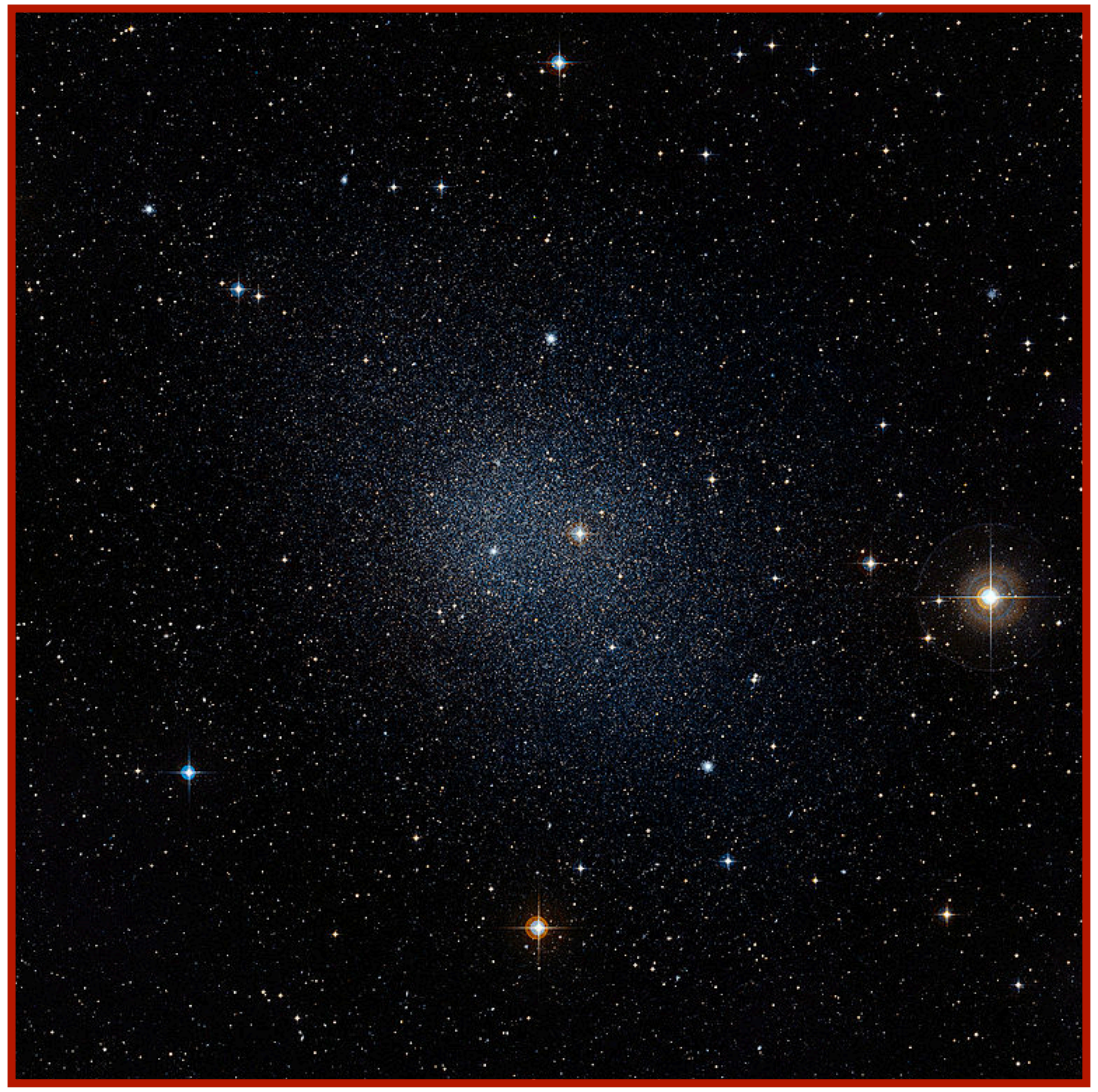


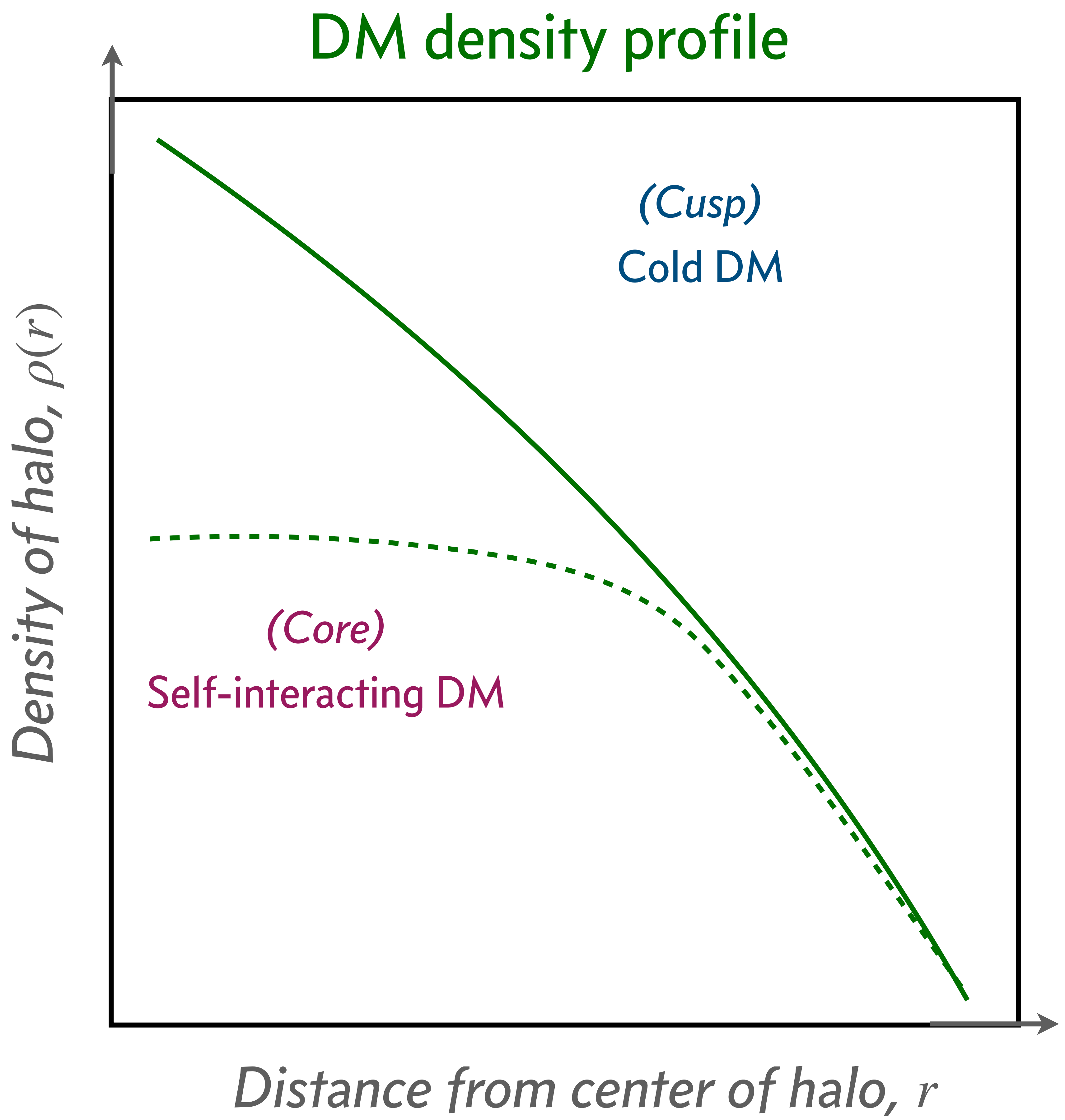
Inferring dark matter halo shapes in dwarf galaxies

Dwarf galaxies and halo shapes

Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

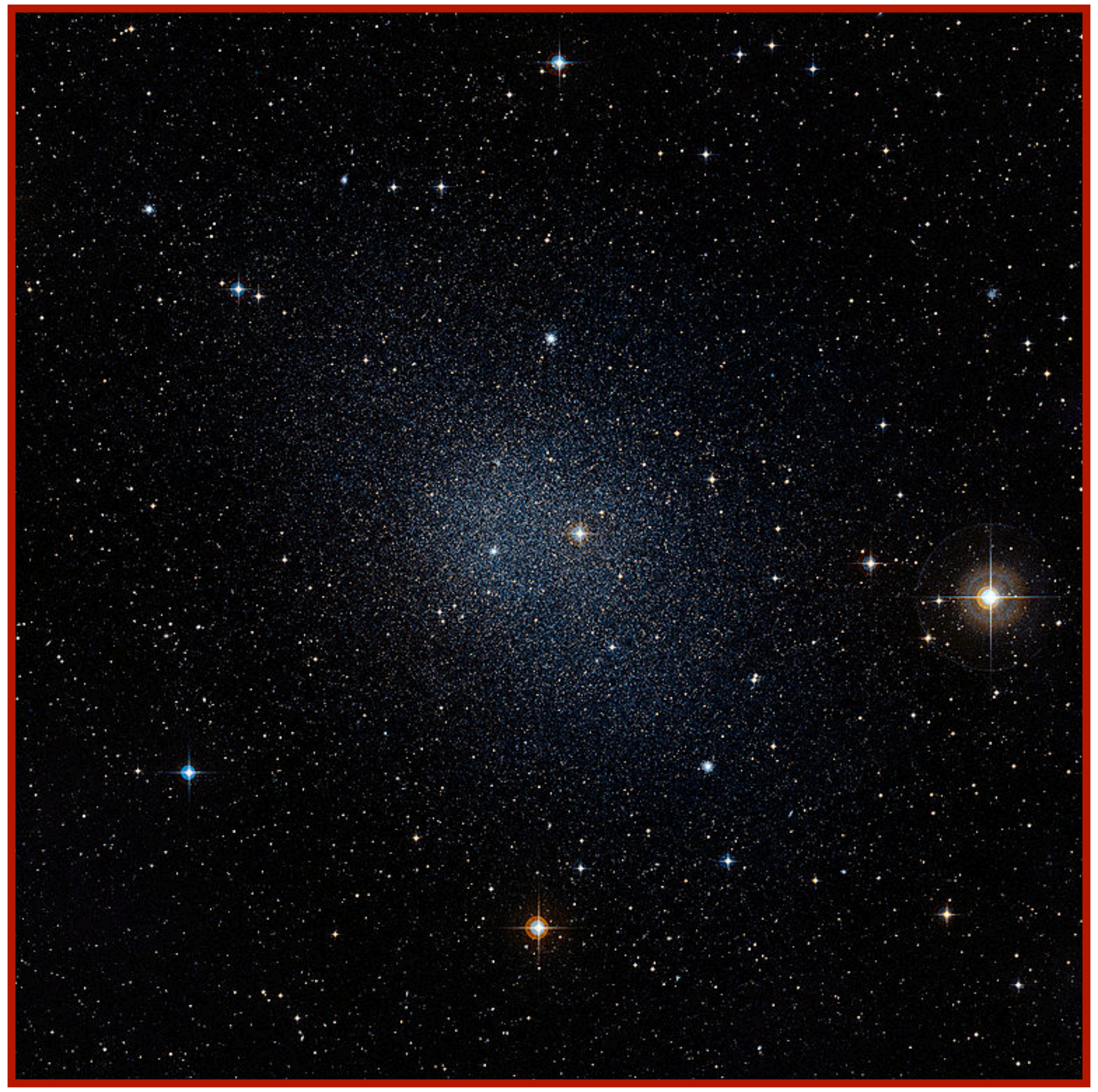


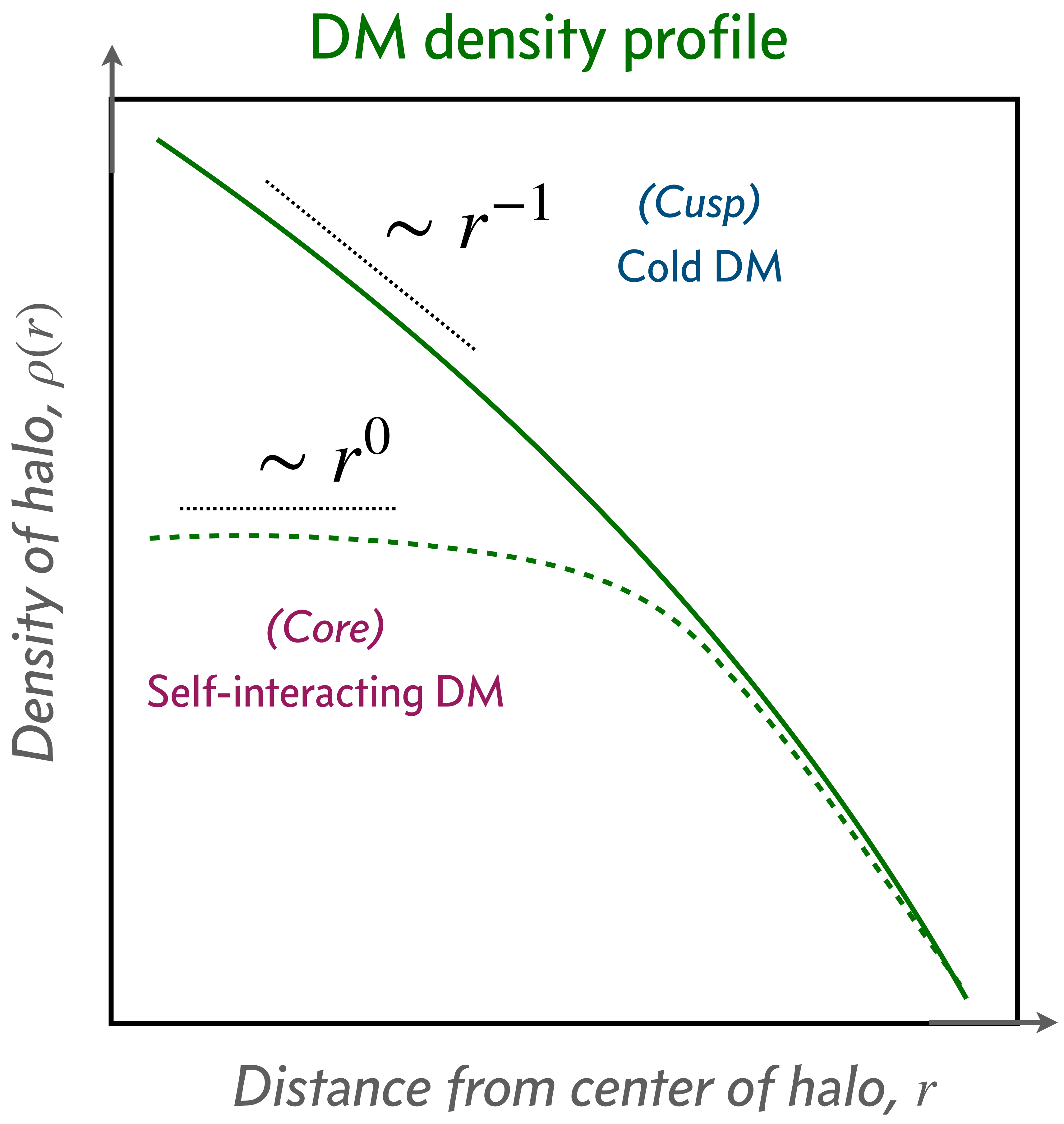


Dwarf galaxies and halo shapes

Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

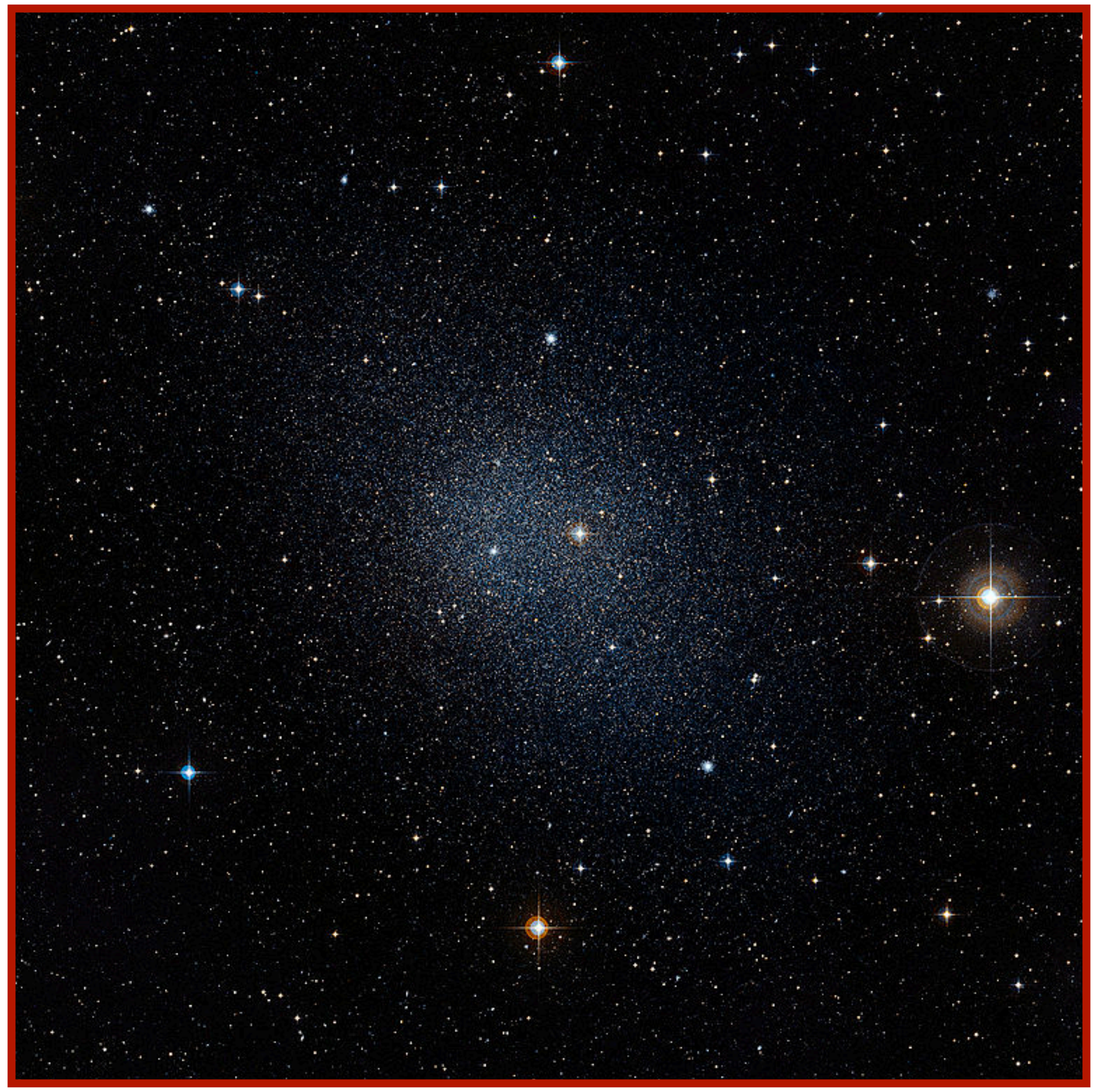


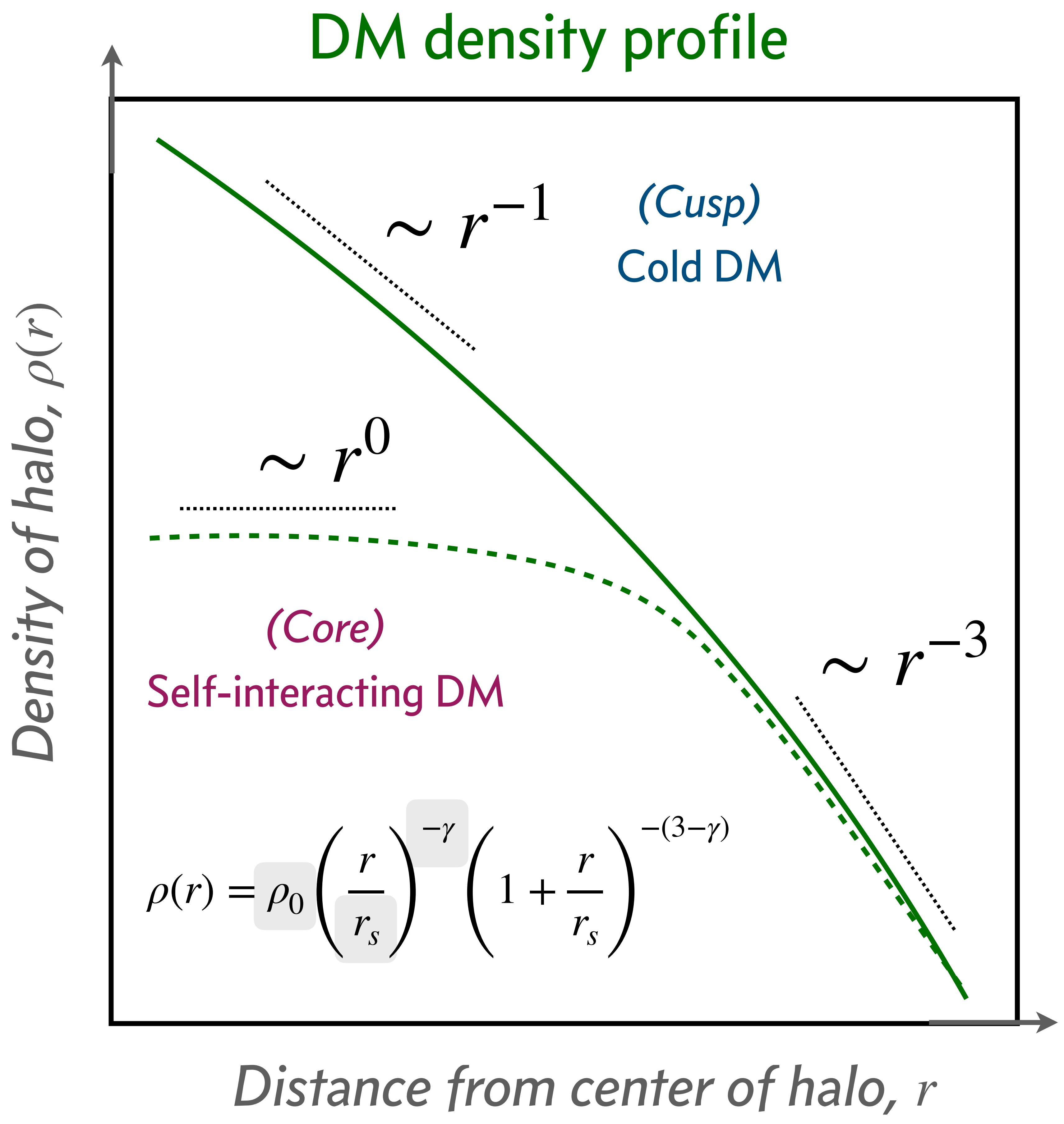


Dwarf galaxies and halo shapes

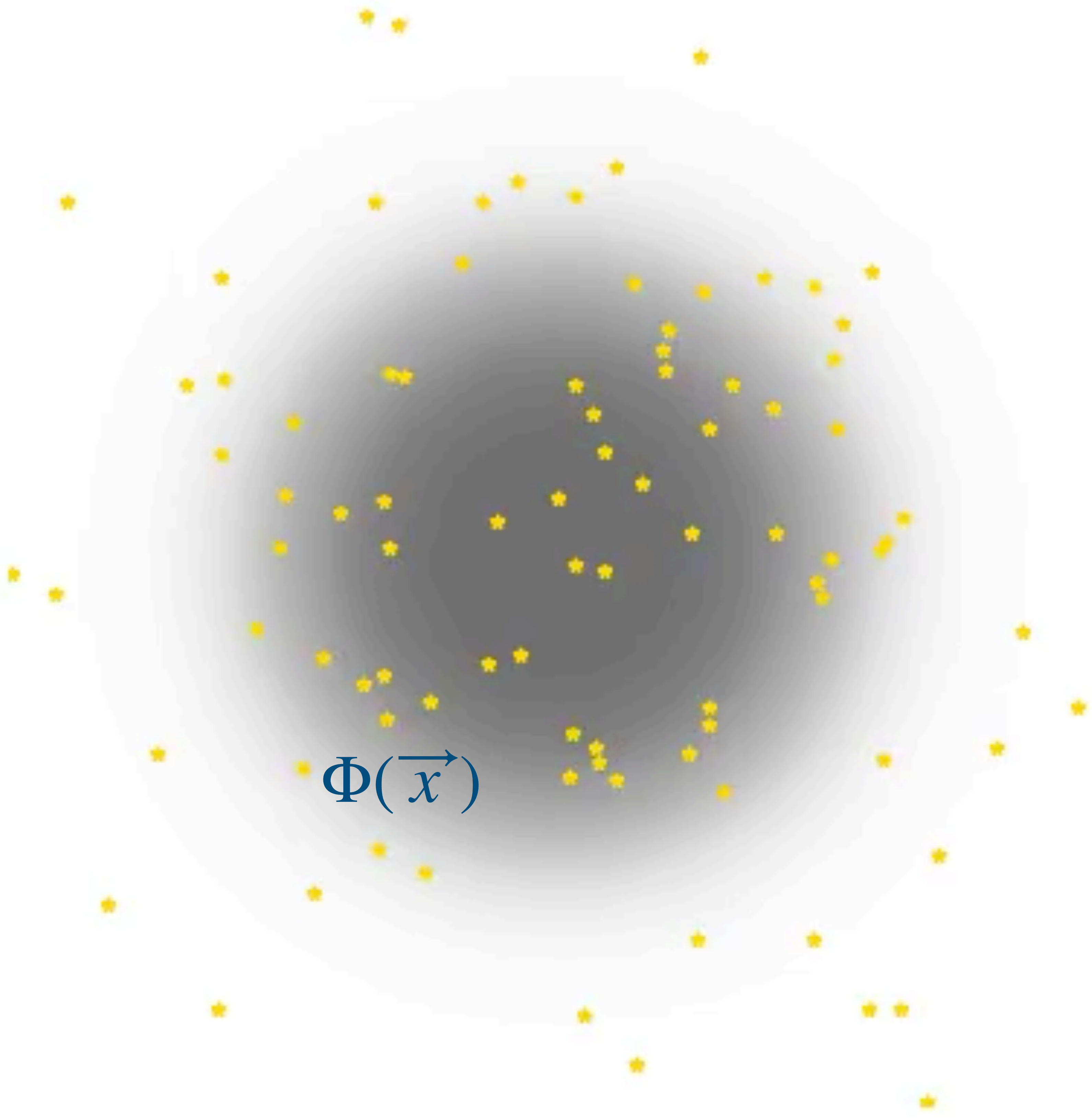
Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

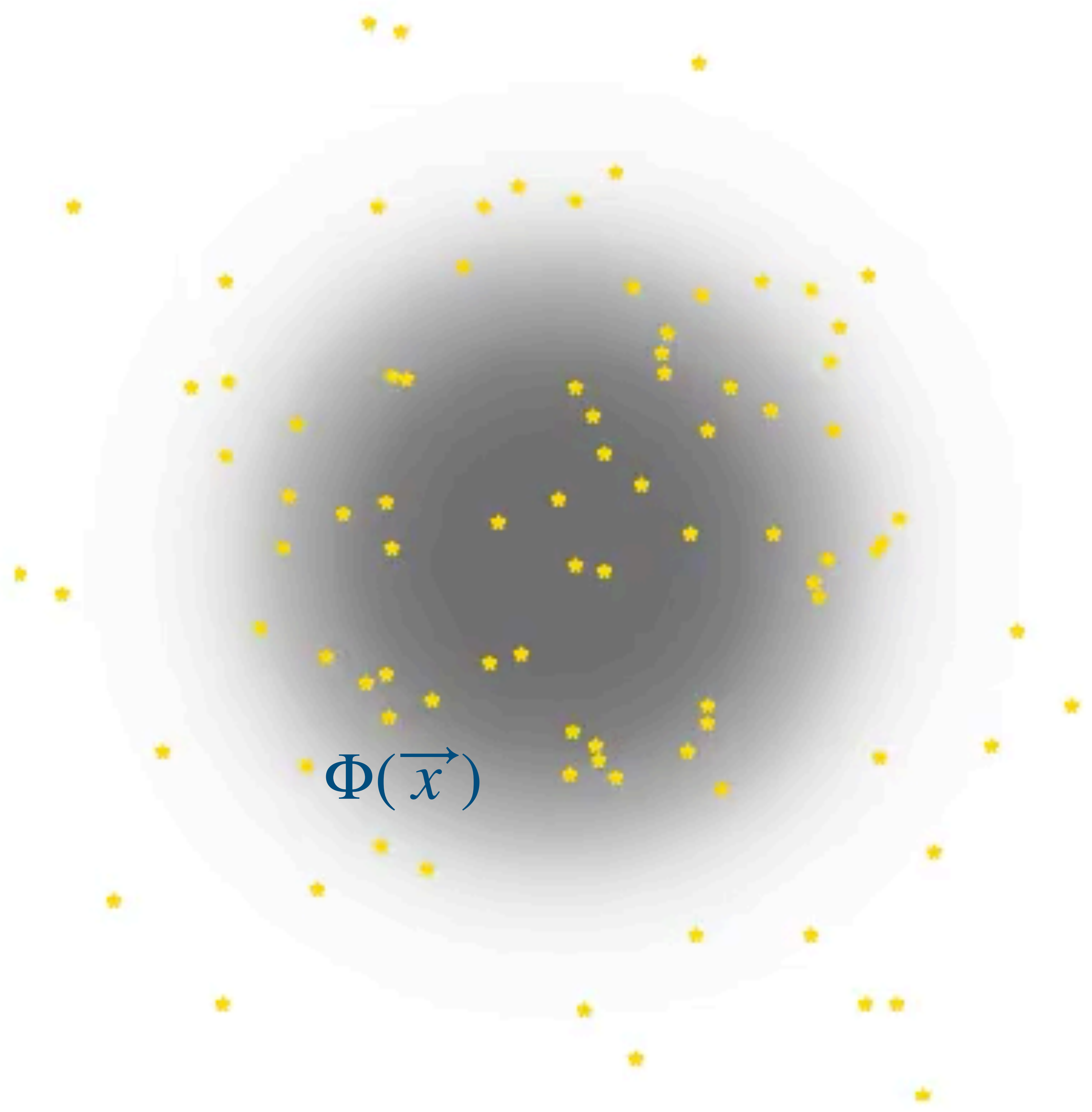




From stellar kinematics to halo shapes: *Jeans modeling*



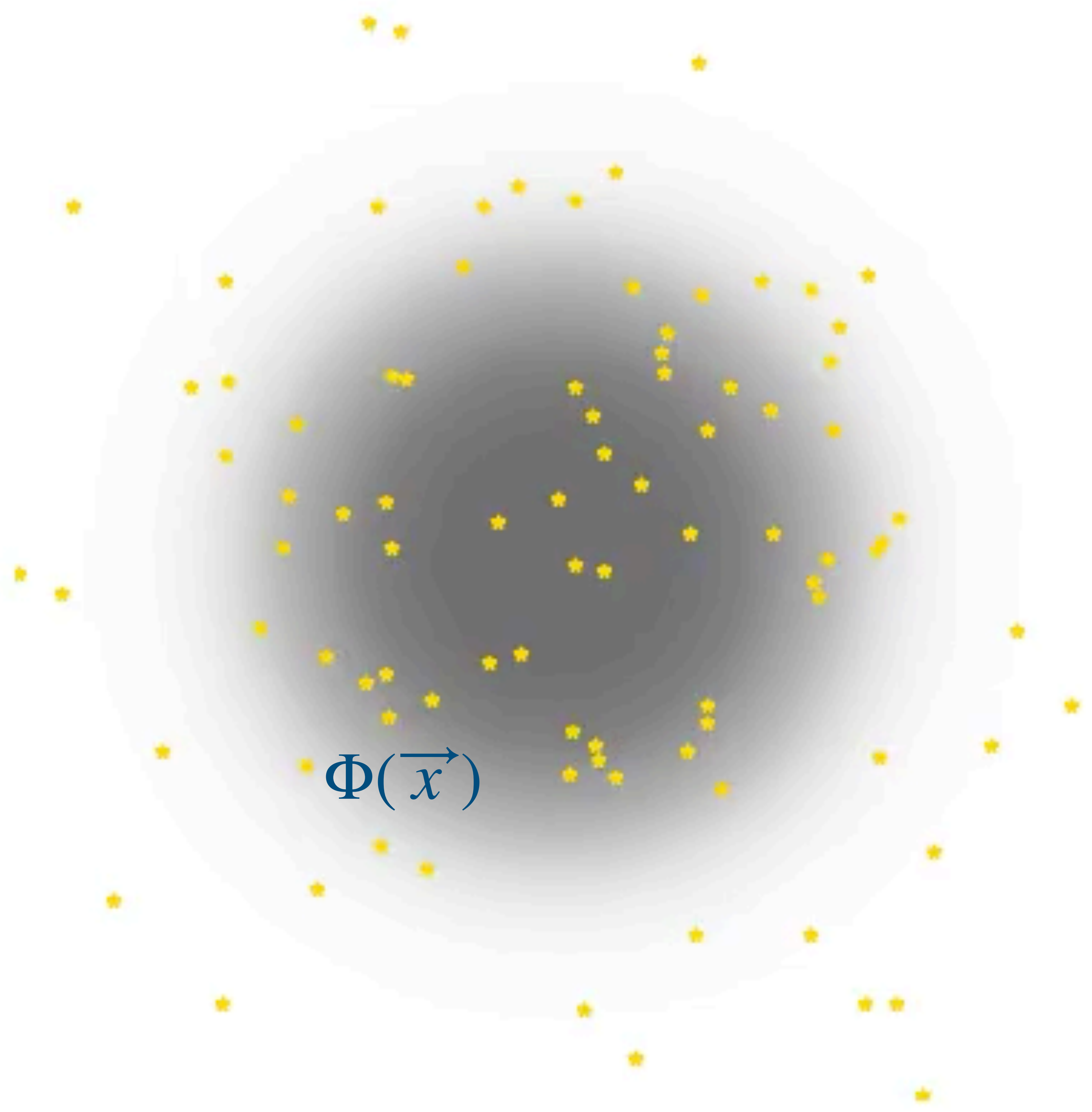
From stellar kinematics to halo shapes: *Jeans modeling*



From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

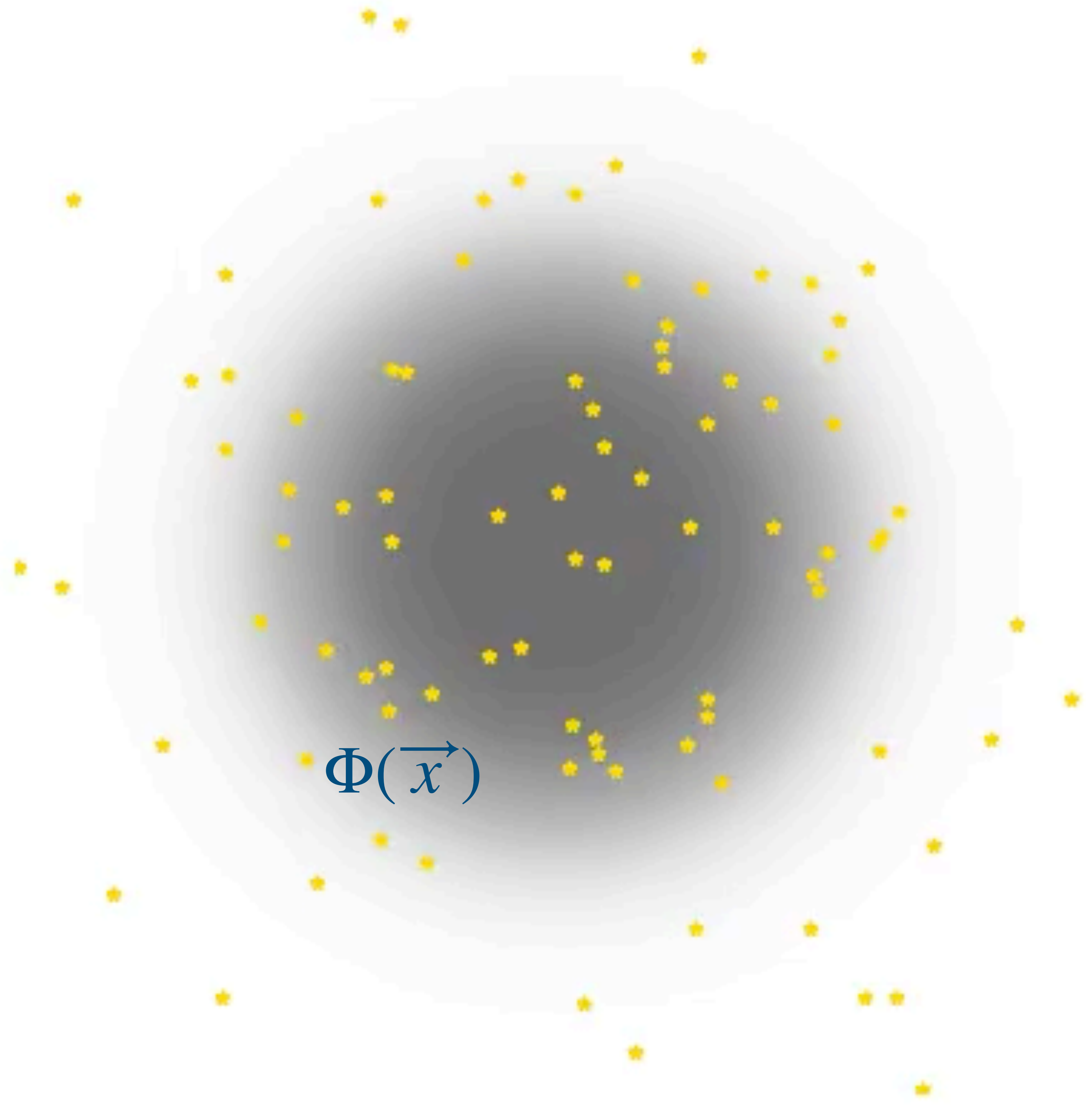
$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



Phase space density and its *moments*

$$n(\vec{x}) = \int \mathrm{d}^3v\, f(\vec{x}, \vec{v})$$

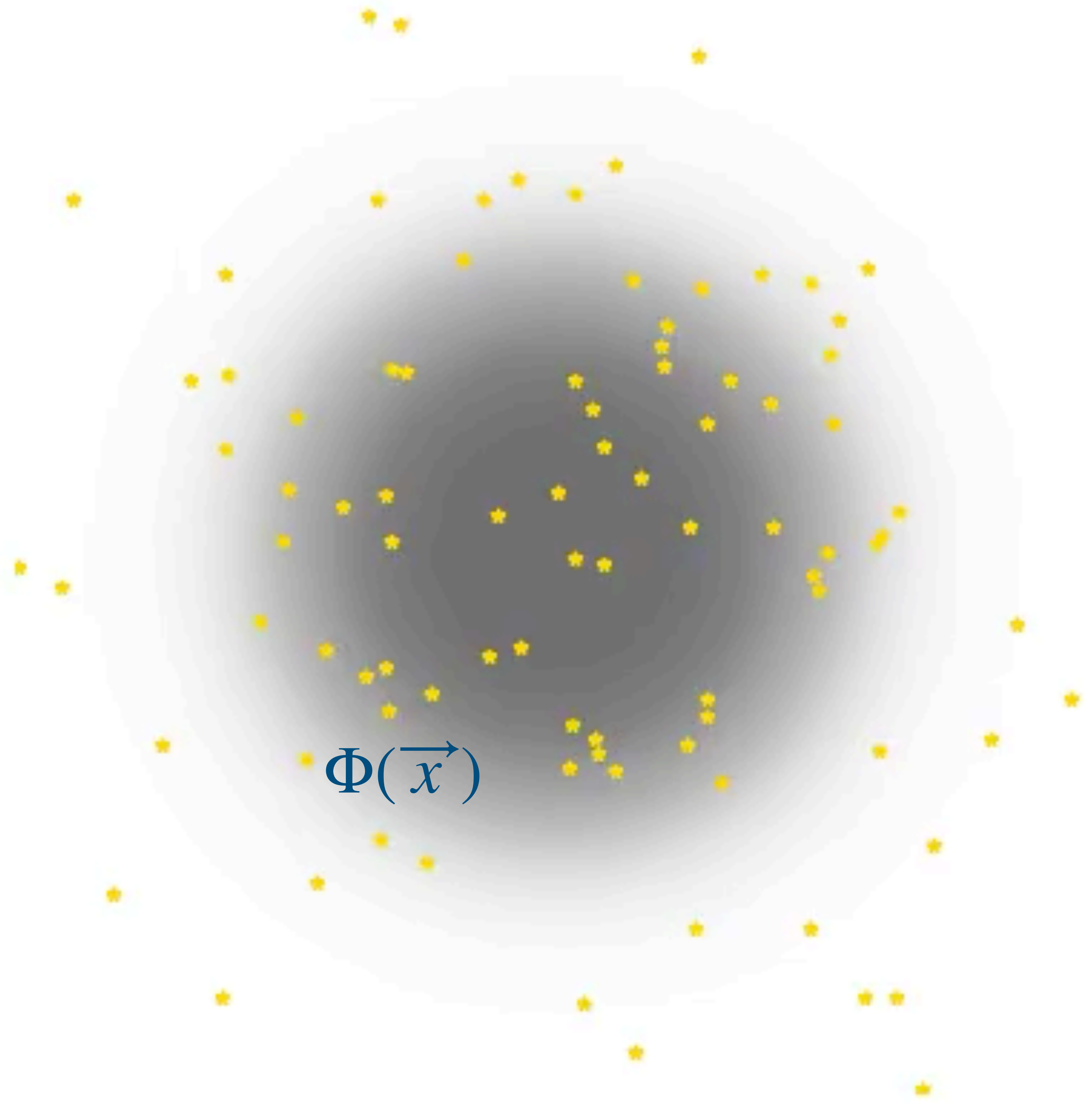
$$\langle v_i(\vec{x}) \rangle = \int \mathrm{d}^3v\, v_i\, f(\vec{x}, \vec{v})$$

$$\sigma_{ij}(\vec{x}) = \int \mathrm{d}^3v\, (v_i - \bar{v}_i)(v_j - \bar{v}_j)\, f(\vec{x}, \vec{v})$$

From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



Phase space density and its *moments*

$$n(\vec{x}) = \int \mathrm{d}^3v\, f(\vec{x}, \vec{v})$$

$$\langle v_i(\vec{x}) \rangle = \int \mathrm{d}^3v\, v_i\, f(\vec{x}, \vec{v})$$

$$\sigma_{ij}(\vec{x}) = \int \mathrm{d}^3v\, (v_i - \bar{v}_i)(v_j - \bar{v}_j) f(\vec{x}, \vec{v})$$

Jeans equations connect moments of $f(\vec{x}, \vec{v})$ to $\Phi(\vec{x})$

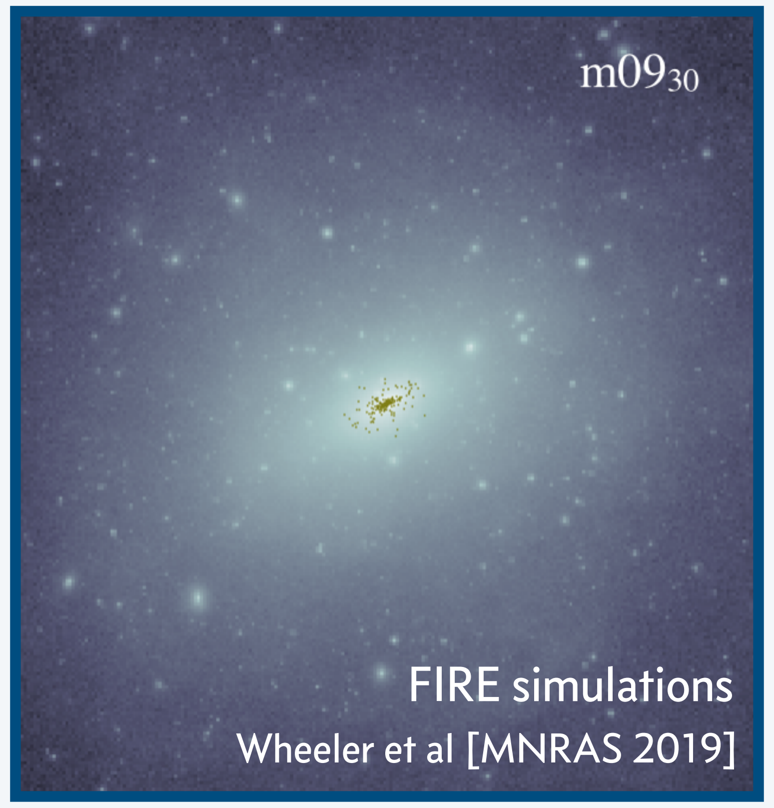
$$n\langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i} + n \frac{\partial \Phi}{\partial x_j} + \frac{\partial \left[n\sigma_{ij}^2 \right]}{\partial x_i} = 0$$

Limitations of Jeans modeling

Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential



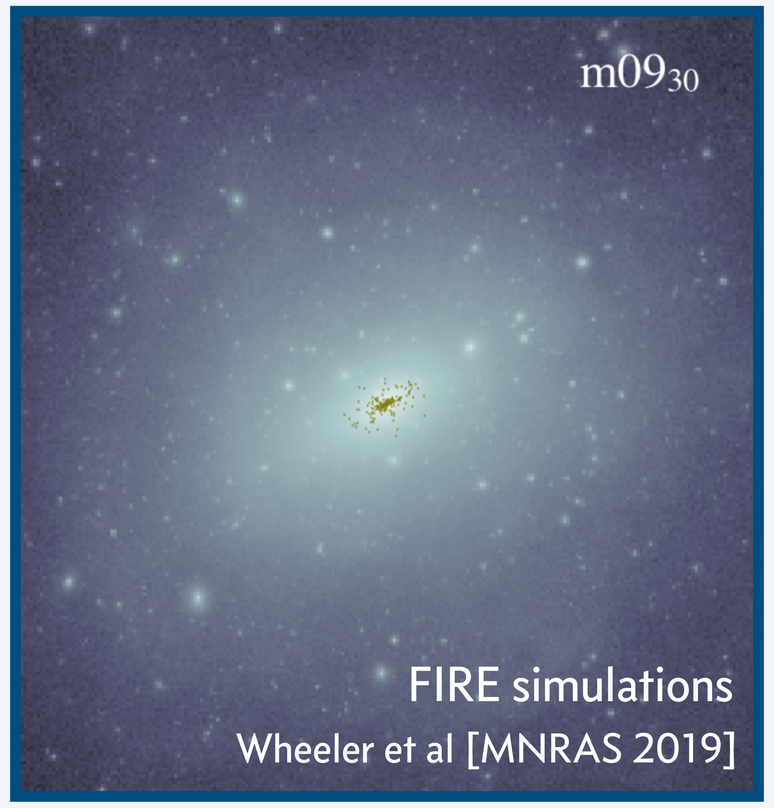
FIRE simulations
Wheeler et al [MNRAS 2019]

Limitations of Jeans modeling

Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential



FIRE simulations
Wheeler et al [MNRAS 2019]

Reliance on moments of $f(\vec{x}, \vec{v})$

$$n\langle v_i\rangle \frac{\partial \langle v_j\rangle}{\partial x_i} + n\frac{\partial \Phi}{\partial x_j} + \frac{\partial \left[n\sigma_{ij}^2\right]}{\partial x_i} = 0$$

- Simplified description of the data = *loss of information*
- Typically only *3* phase-space coordinates available:

$$\{\vec{r}, \vec{v}\} \longrightarrow \{\vec{r}_\perp, \vec{v}_{\rm los}\}$$

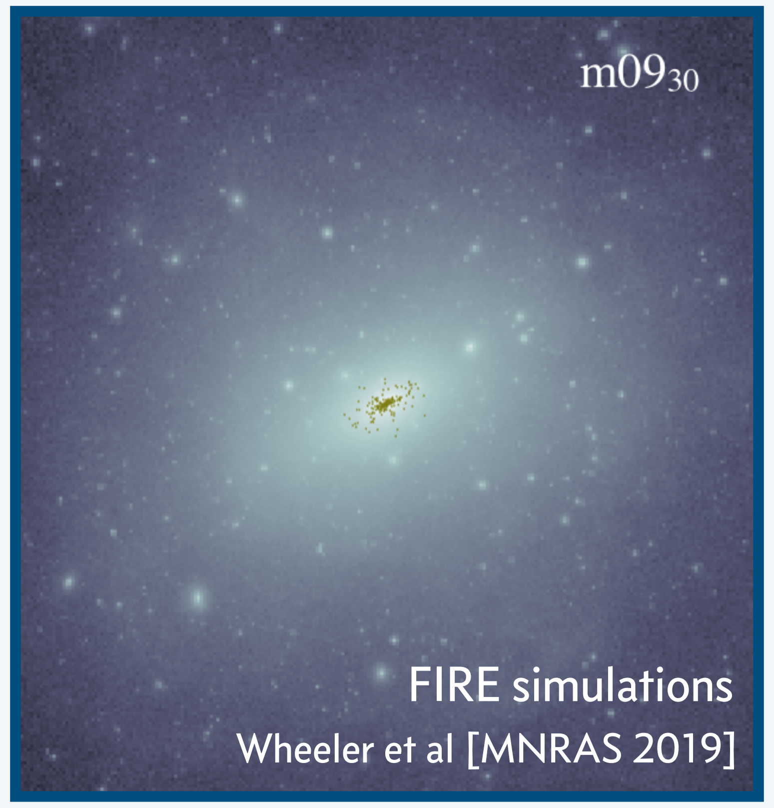
- Noisy estimates of $\sigma_r^2(r)$, $n(r)$ and derivatives

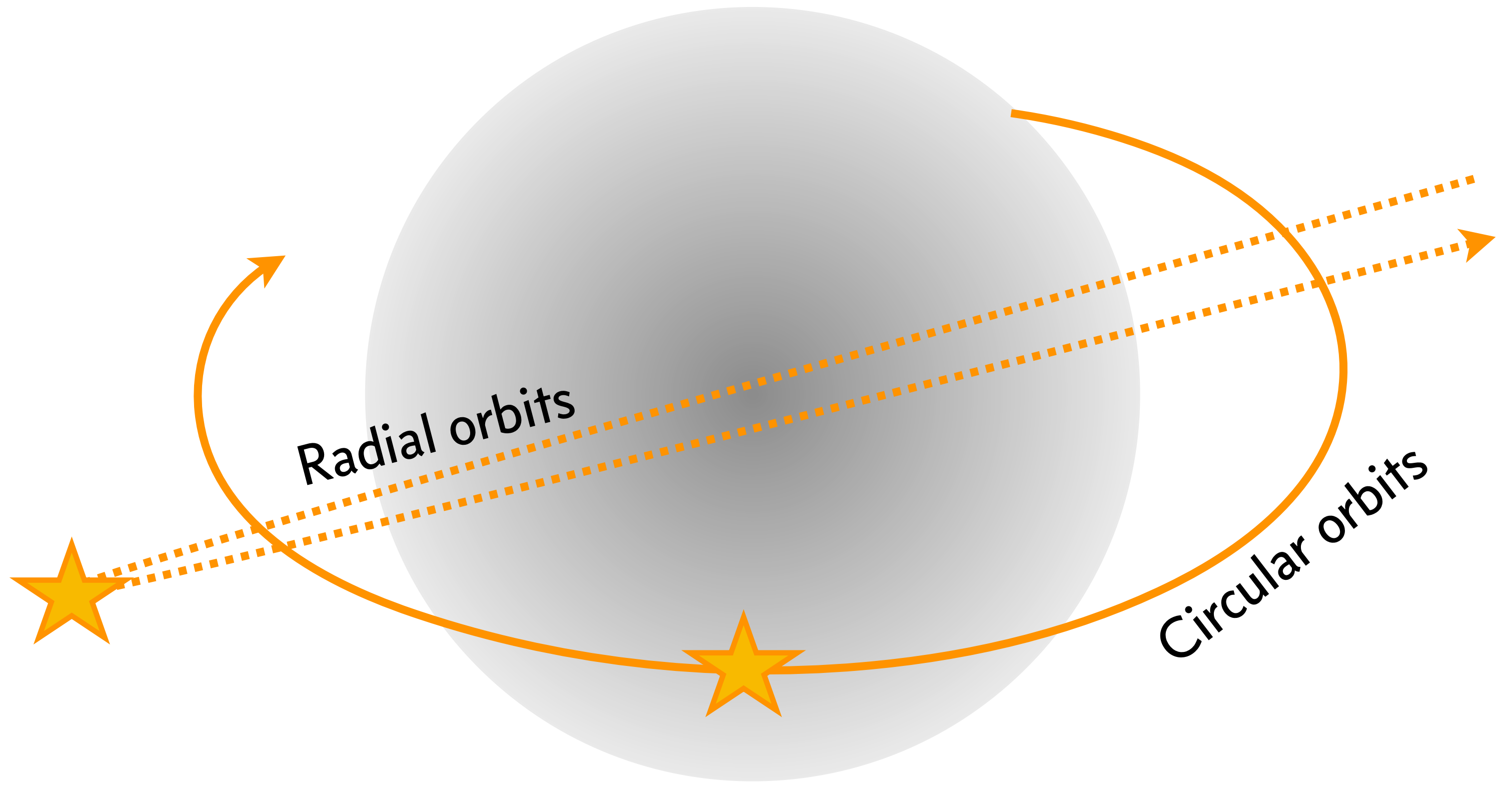
Limitations of Jeans modeling

Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential





Reliance on moments of $f(\vec{x}, \vec{v})$

$$n\langle v_i\rangle\frac{\partial\langle v_j\rangle}{\partial x_i}+n\frac{\partial\Phi}{\partial x_j}+\frac{\partial\left[n\sigma_{ij}^2\right]}{\partial x_i}=0$$

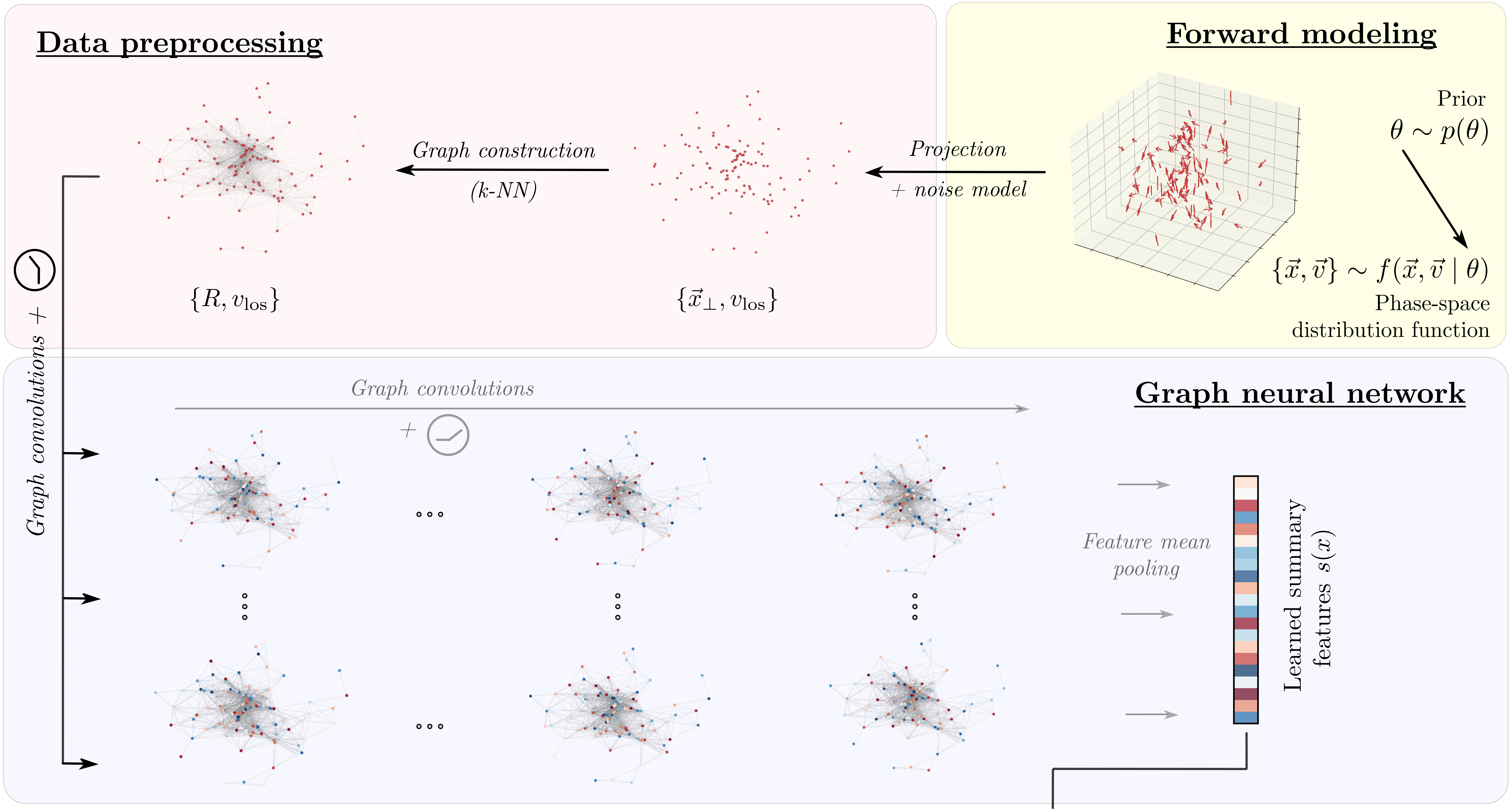
- Simplified description of the data = *loss of information*
- Typically only *3* phase-space coordinates available:

$$\{\vec{r},\vec{v}\}\longrightarrow\{\vec{r}_\perp,\vec{v}_{\rm los}\}$$

- Noisy estimates of $\sigma_r^2(r)$, $n(r)$ and derivatives

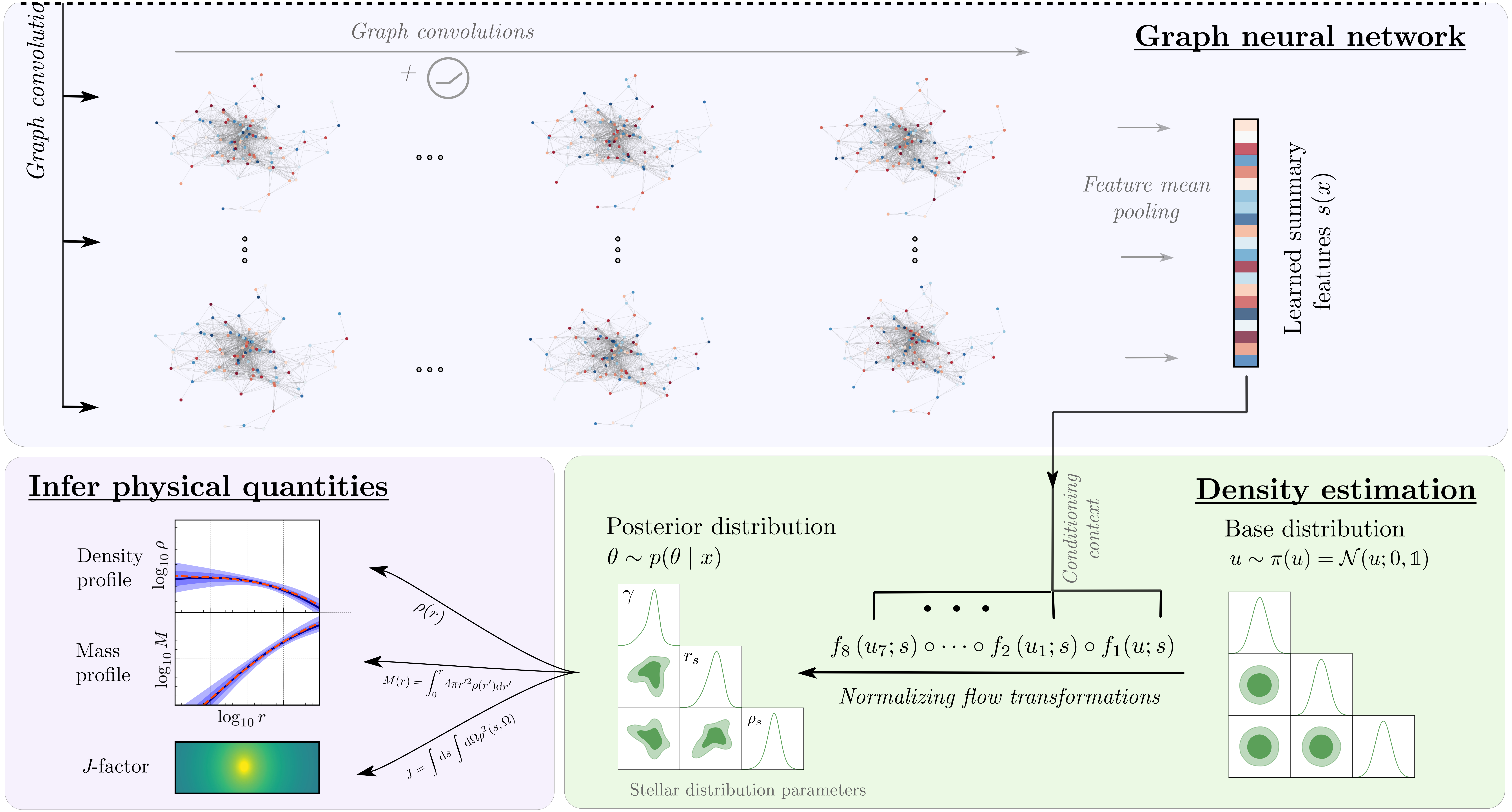
Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]



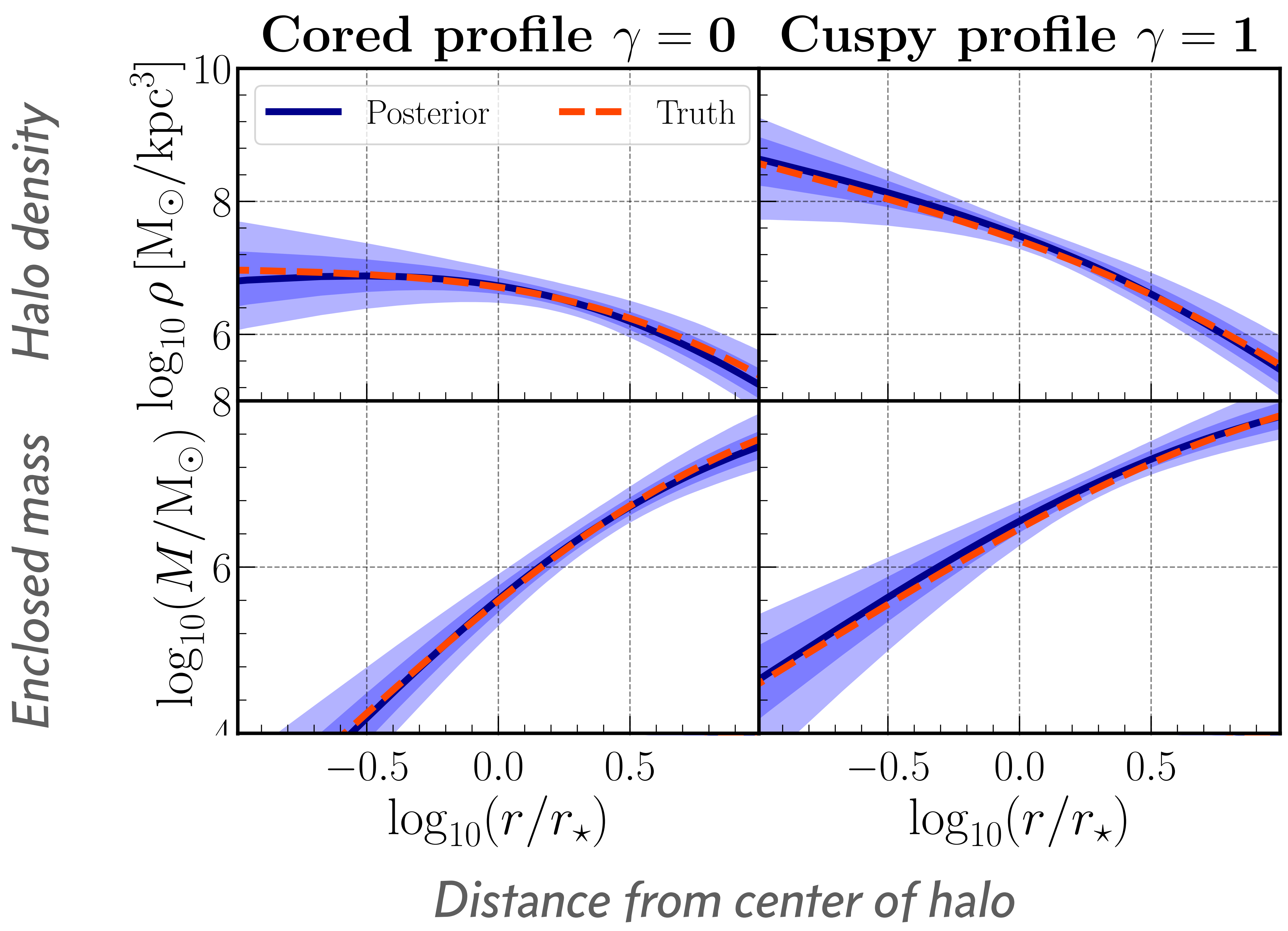
Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]



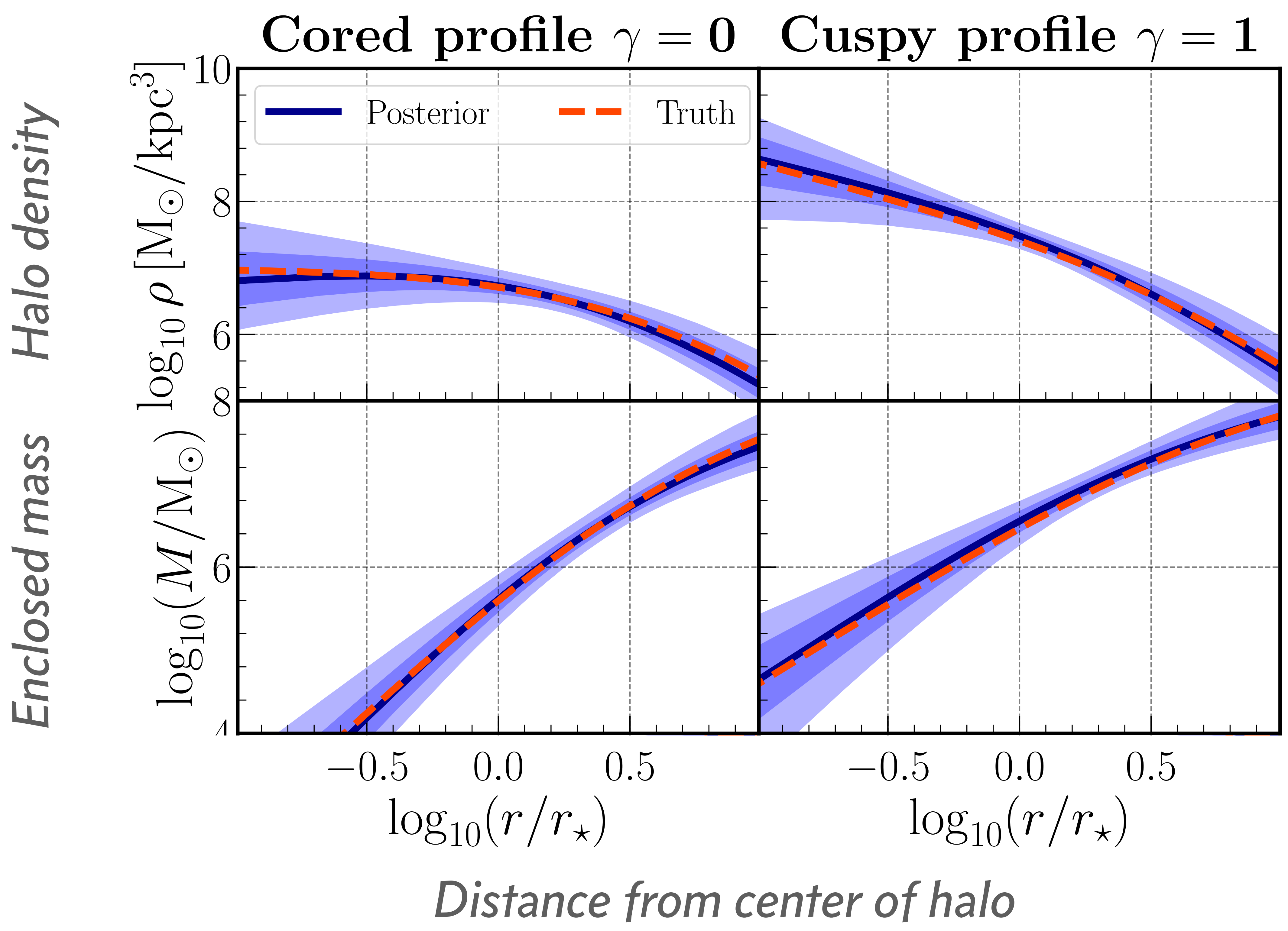
Inferring the dark matter profile

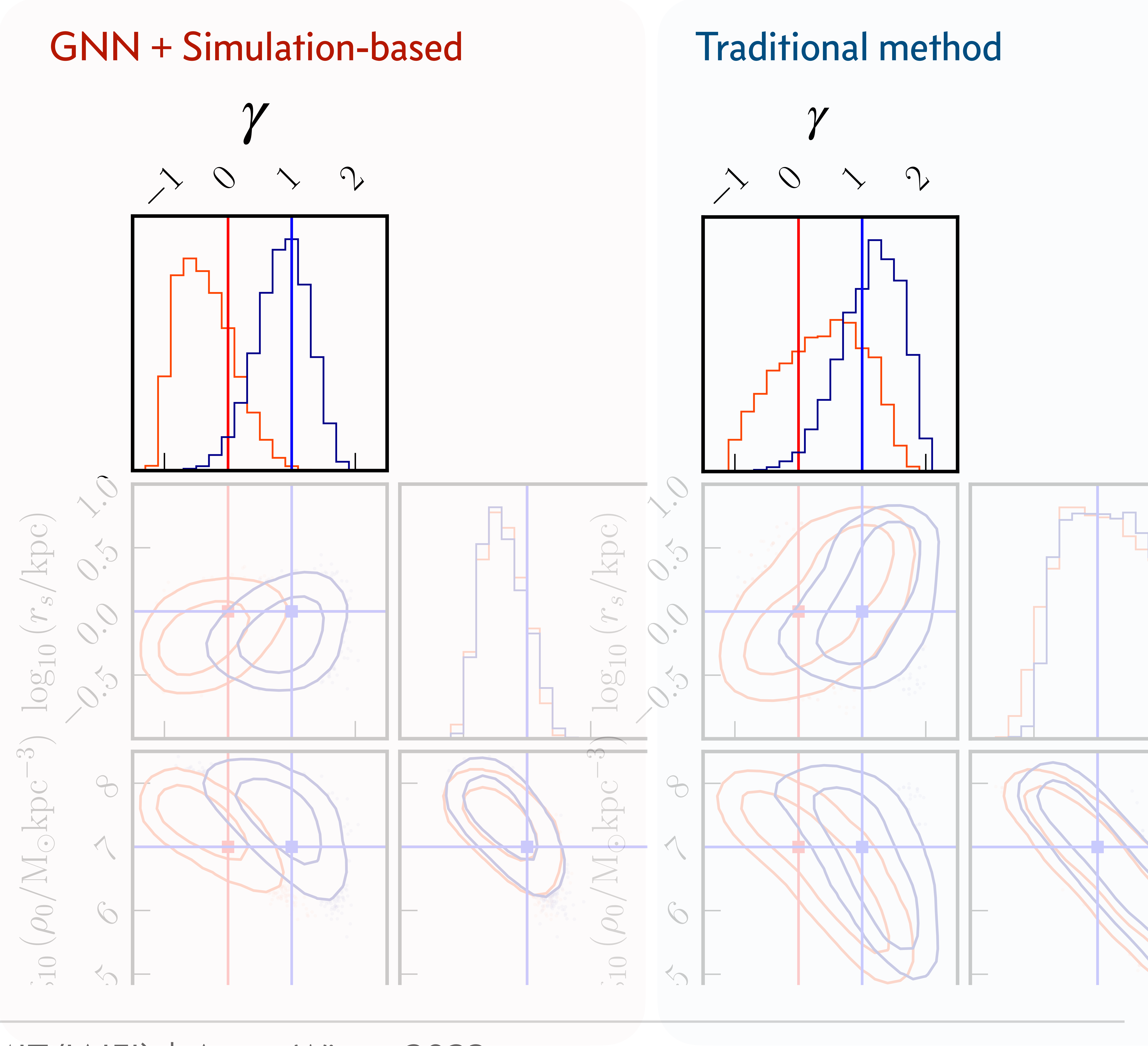
Nguyen, SM et al [PRD 2023]



Inferring the dark matter profile

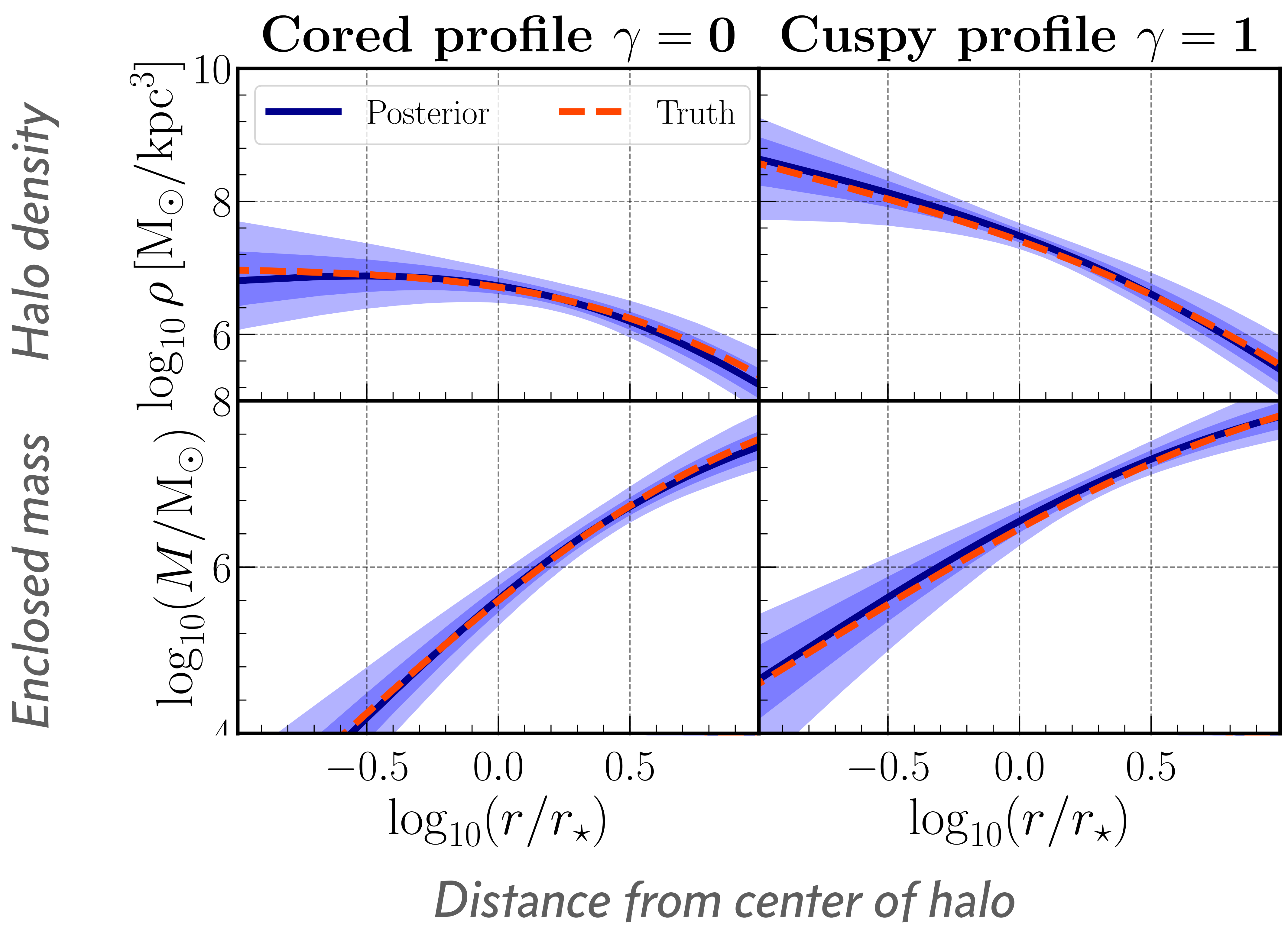
Nguyen, SM et al [PRD 2023]

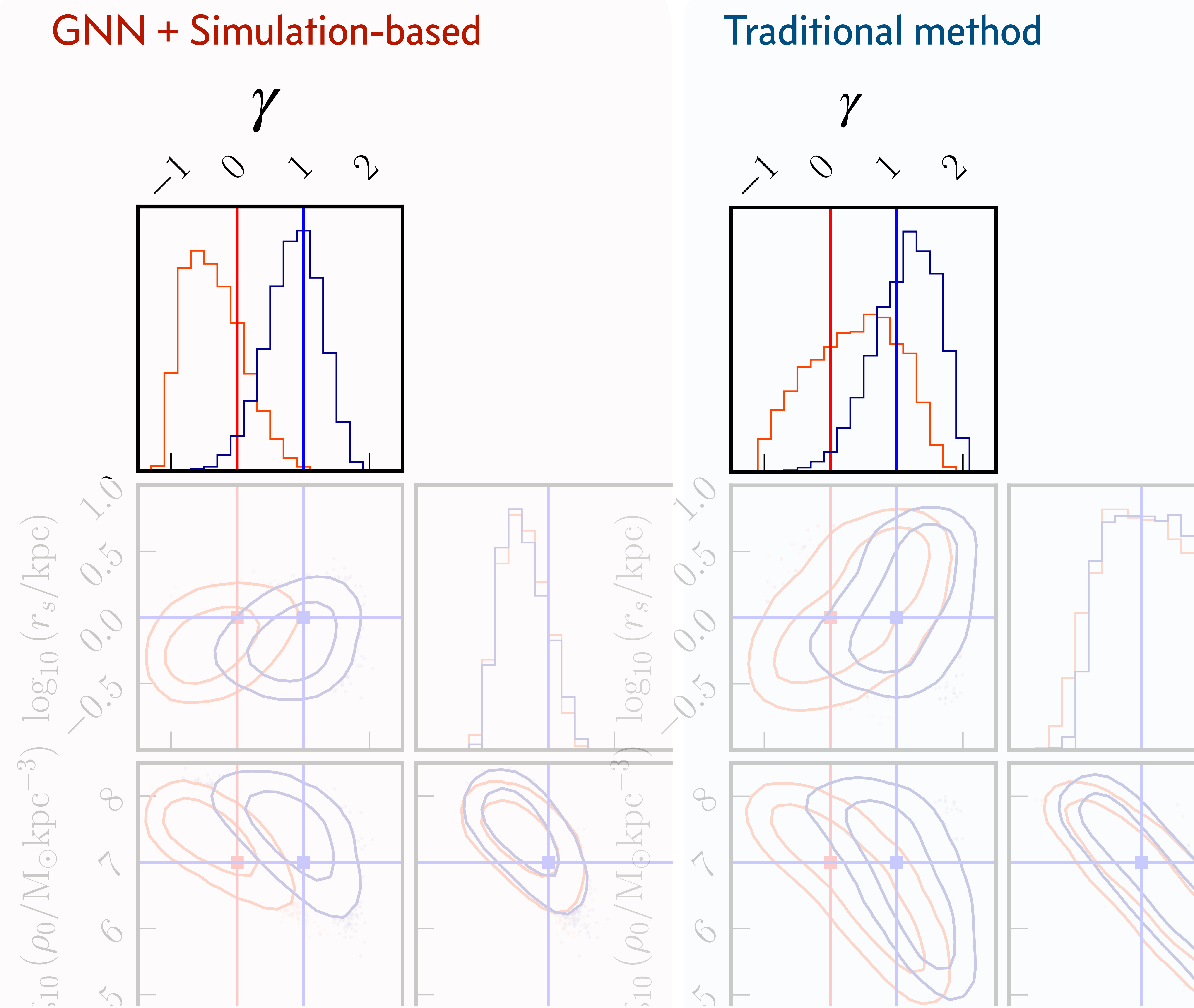




Inferring the dark matter profile

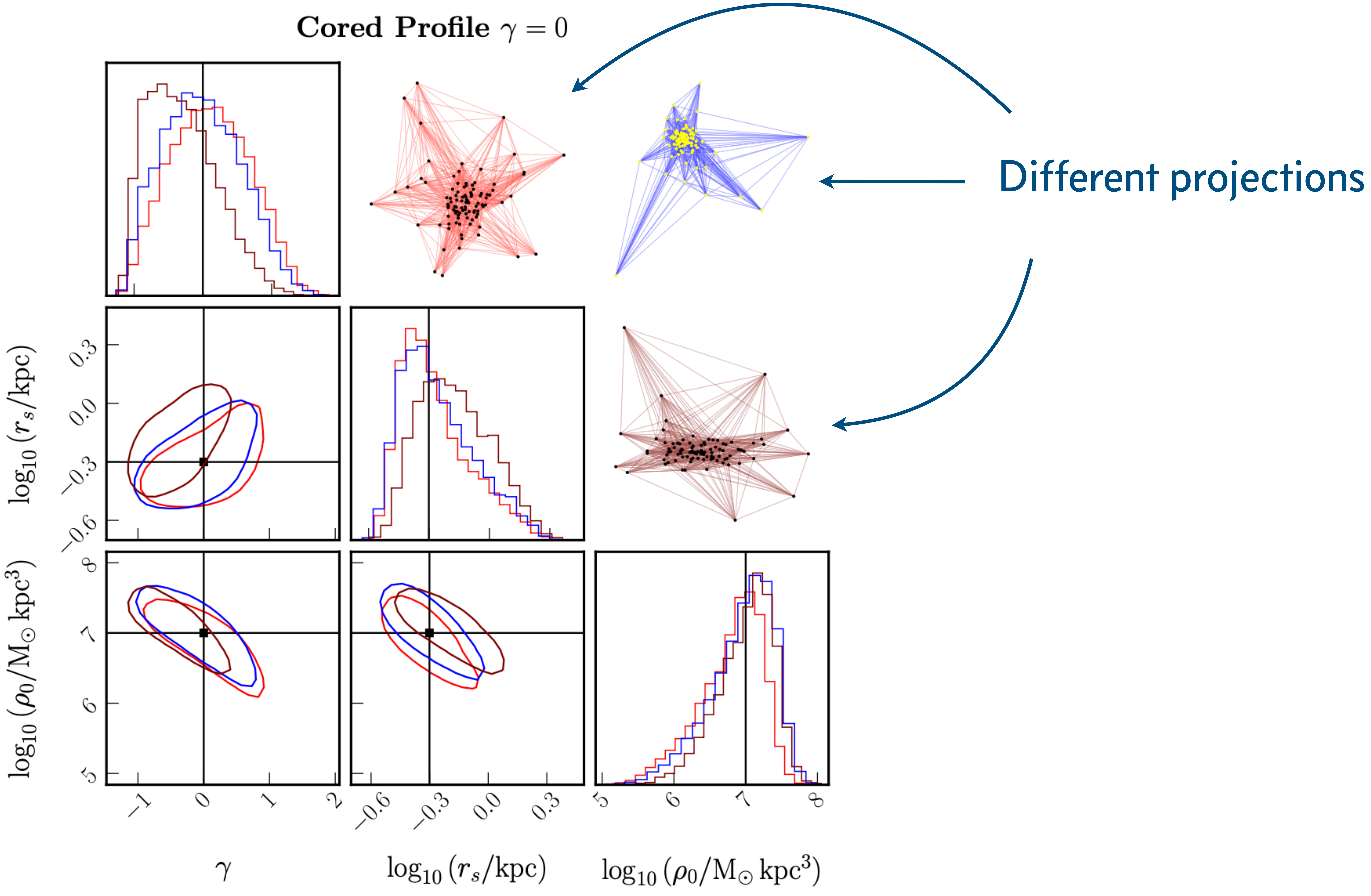
Nguyen, SM et al [PRD 2023]





- *Leverage more information → greater sensitivity*
- *Fewer assumptions → more flexible*
- *Significantly faster analysis*

Sensitivity to projection



Applications to hydrodynamic simulations

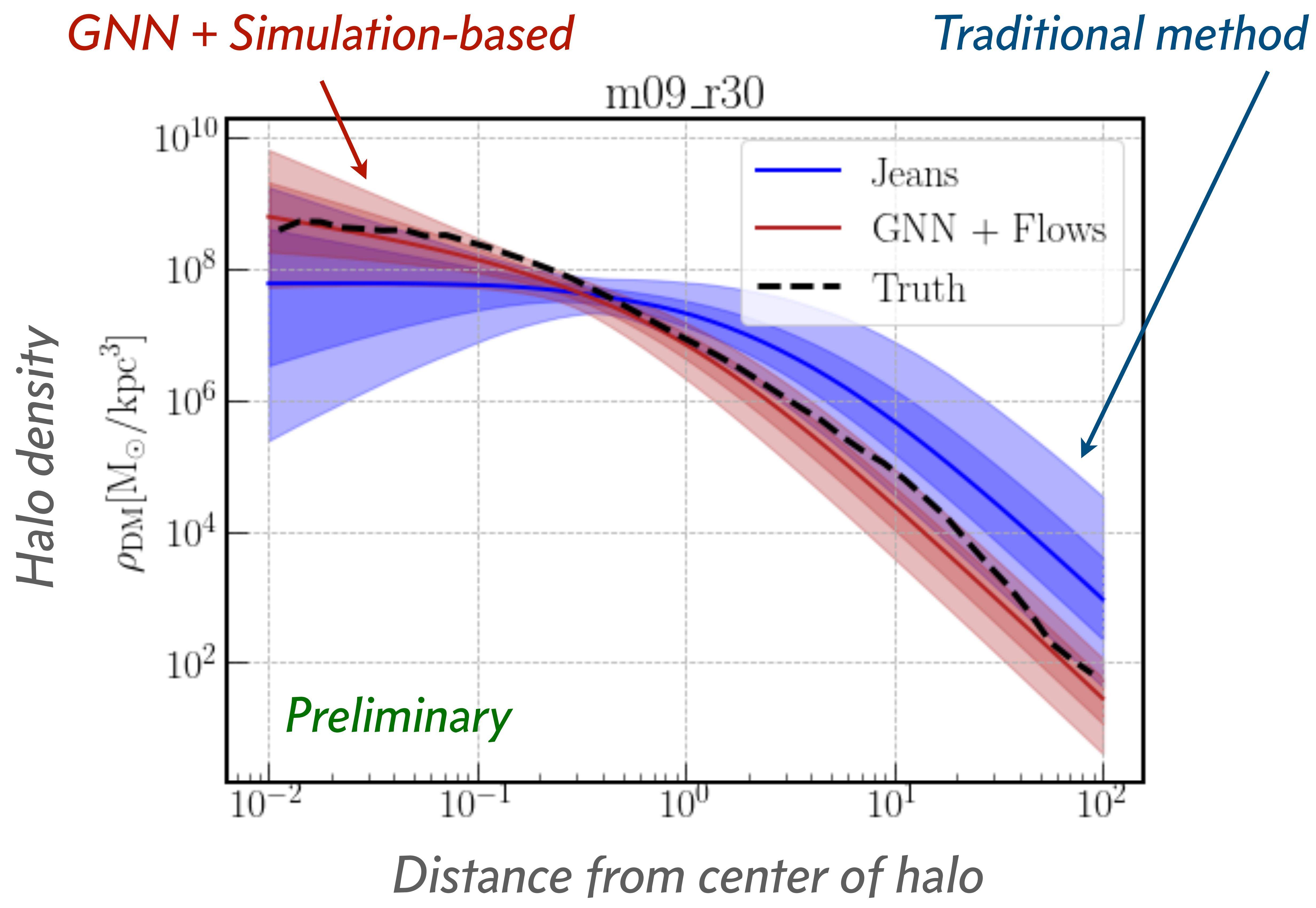
Nguyen, SM et al [In prep]

Wheeler et al [MNRAS 2019]

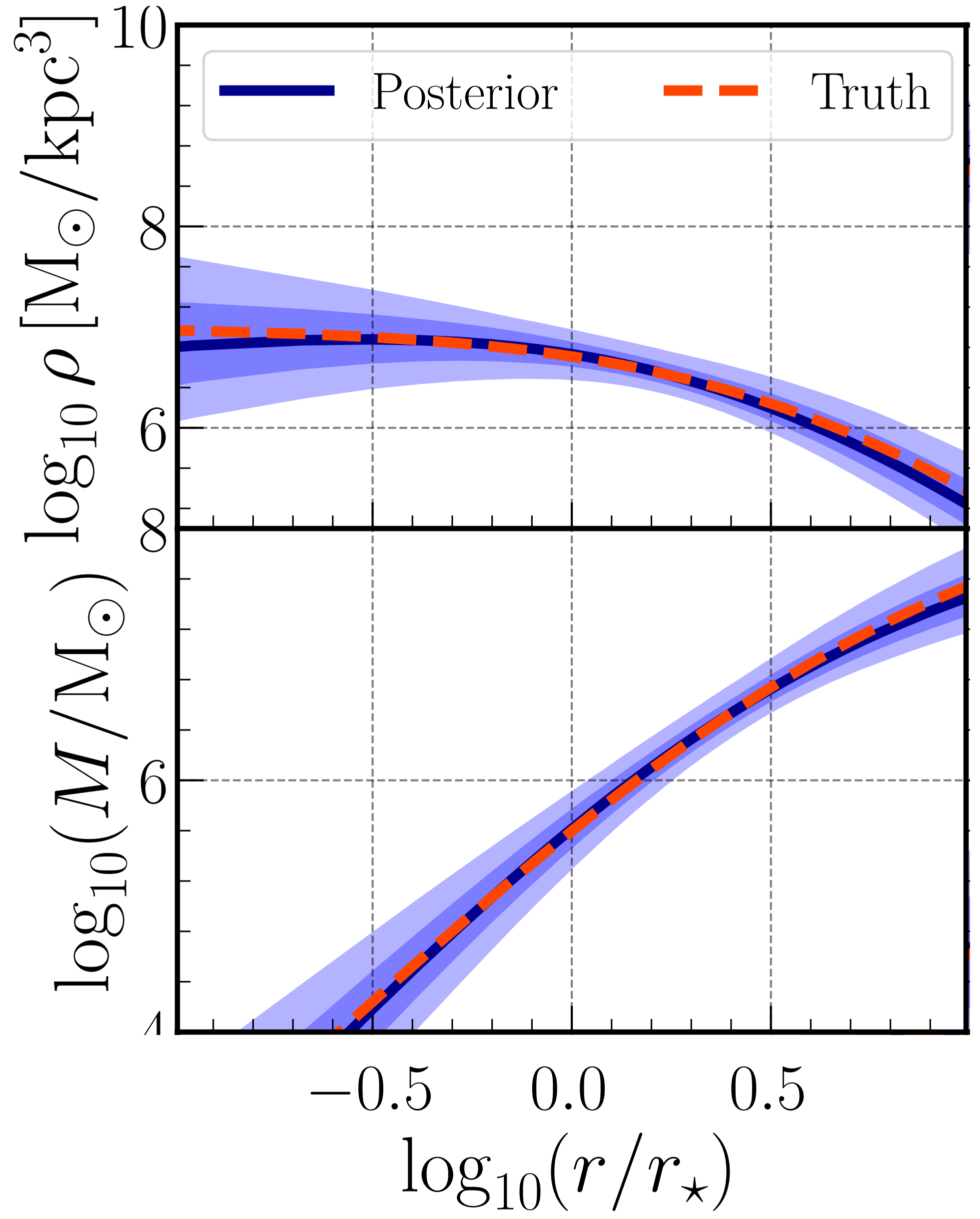
Applications to hydrodynamic simulations

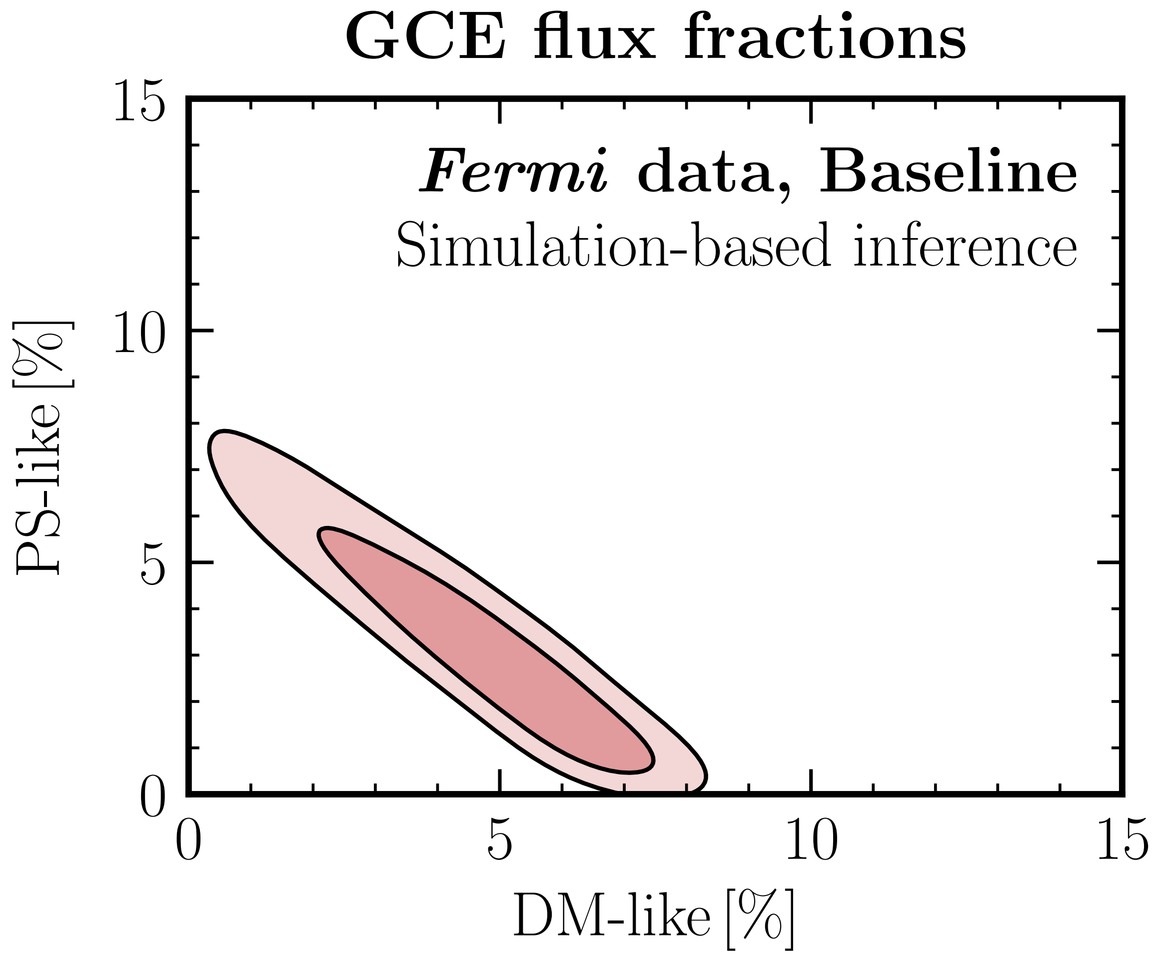
Nguyen, SM et al [In prep]

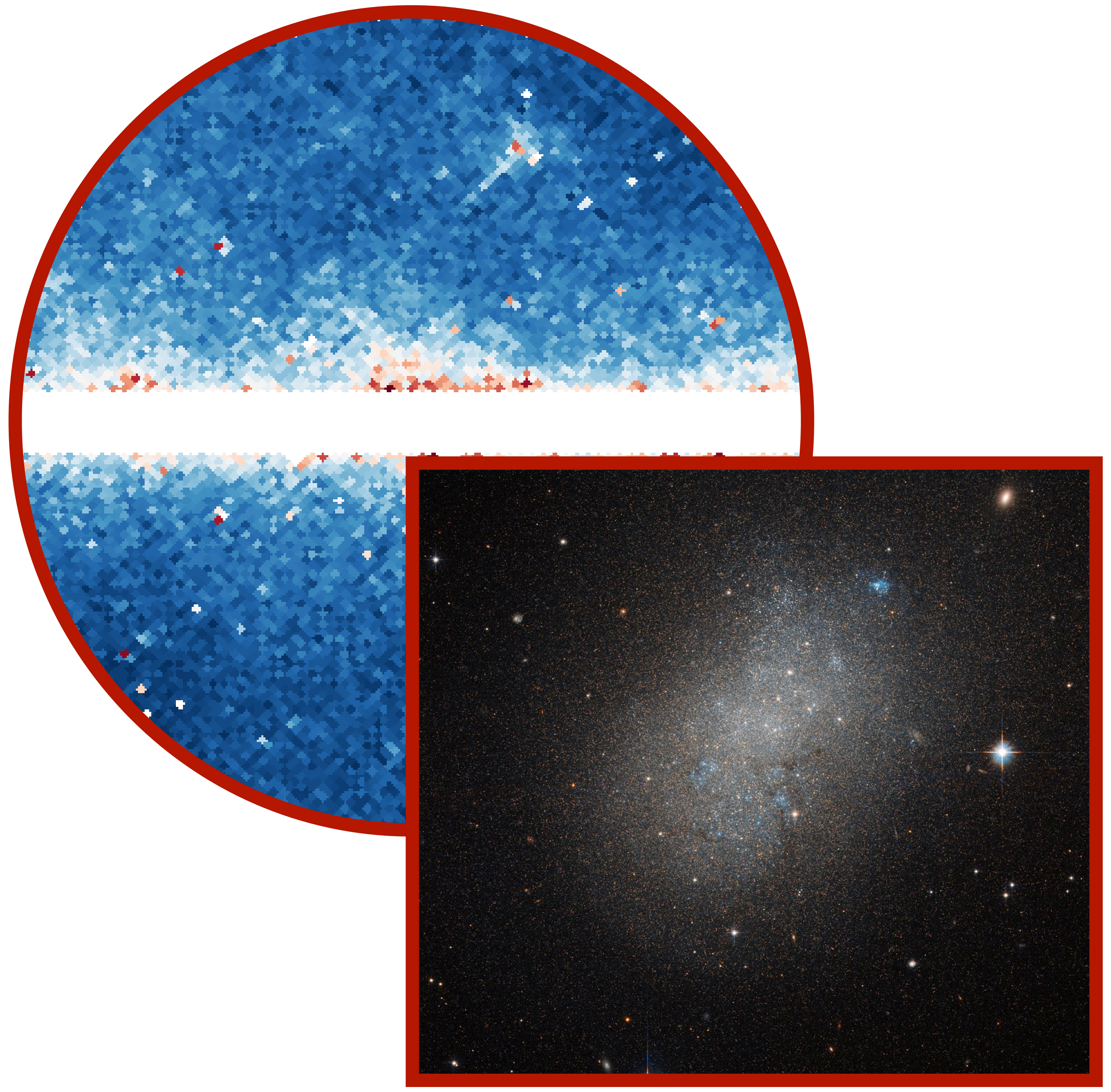
Wheeler et al [MNRAS 2019]



Conclusions



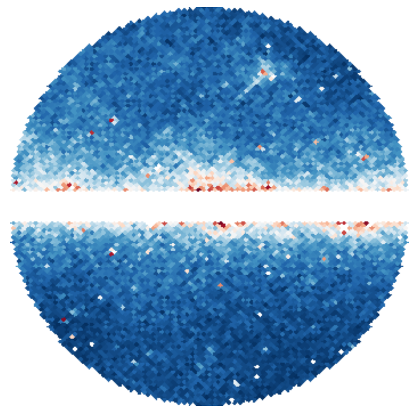


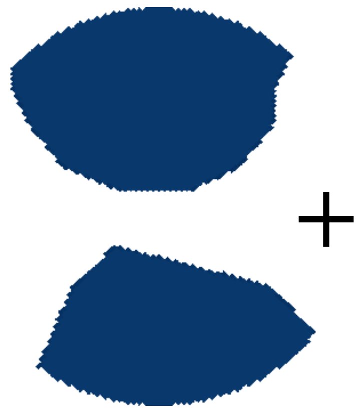
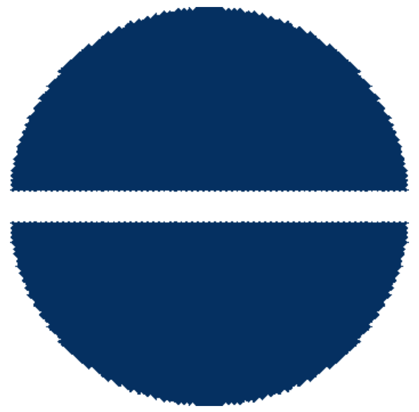
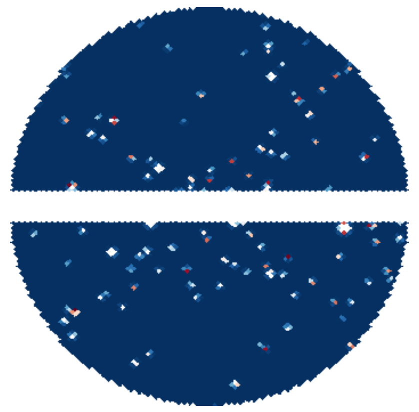
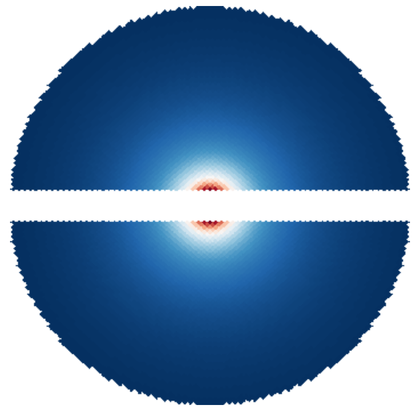
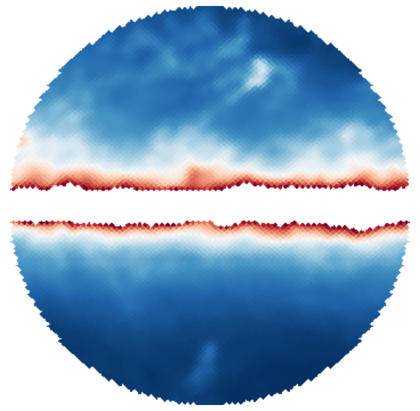


Additional slides

Modeled components

Data modeled as a Poisson realization of a linear combination of spatial templates

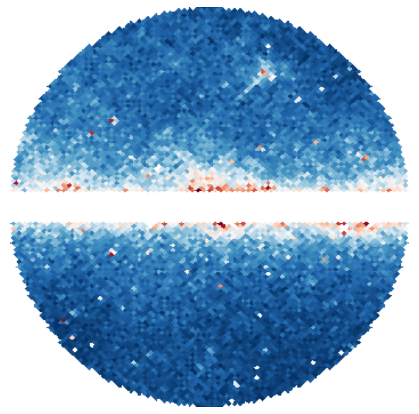


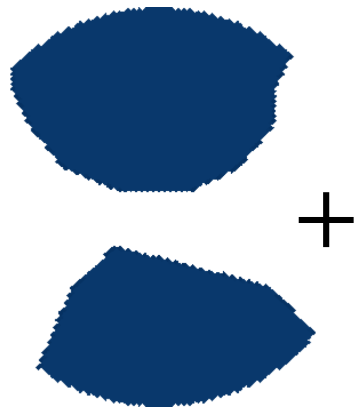
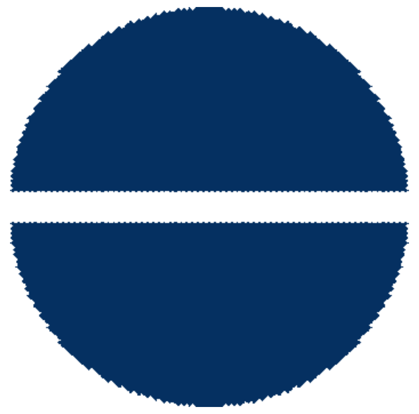
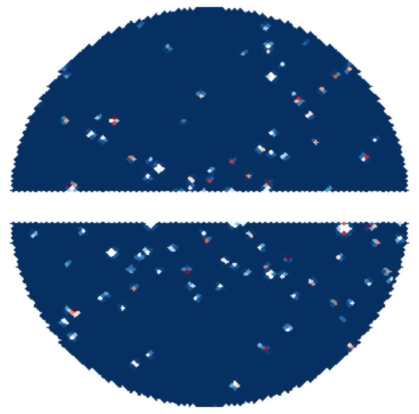
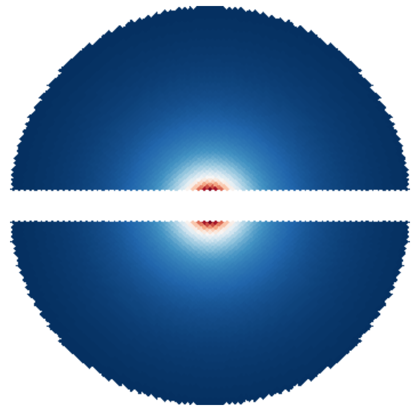
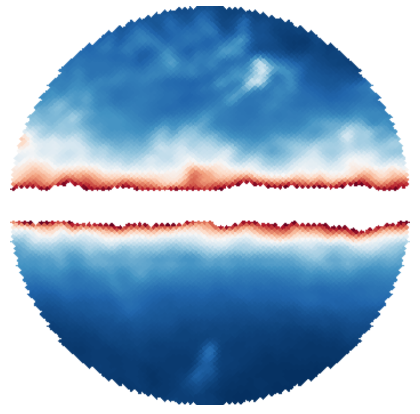
$= \mathrm{Pois}\Big(\theta_{\mathrm{bub}} \times$  $+\ \theta_{\mathrm{iso}} \times$  $+\ \theta_{\mathrm{PS}} \times$  $+\ \theta_{\mathrm{DM}} \times$  $+\ \theta_{\mathrm{dif}} \times$  $\Big)$

~10% ~80%

Modeled components

Data modeled as a Poisson realization of a linear combination of spatial templates

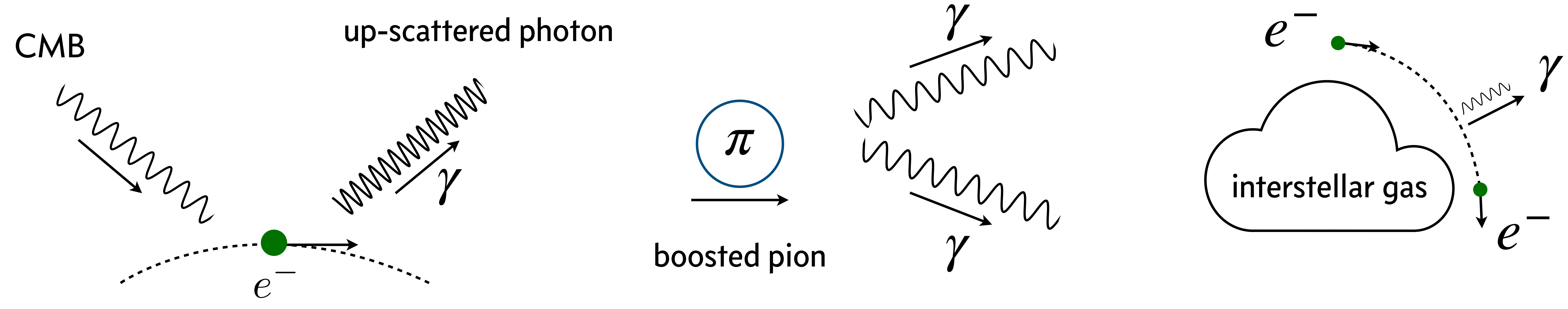


$= \mathrm{Pois}\Big(\theta_{\mathrm{bub}} \times$  $+\ \theta_{\mathrm{iso}} \times$  $+\ \theta_{\mathrm{PS}} \times$  $+\ \theta_{\mathrm{DM}} \times$  $+\ \theta_{\mathrm{dif}} \times$  $\Big)$

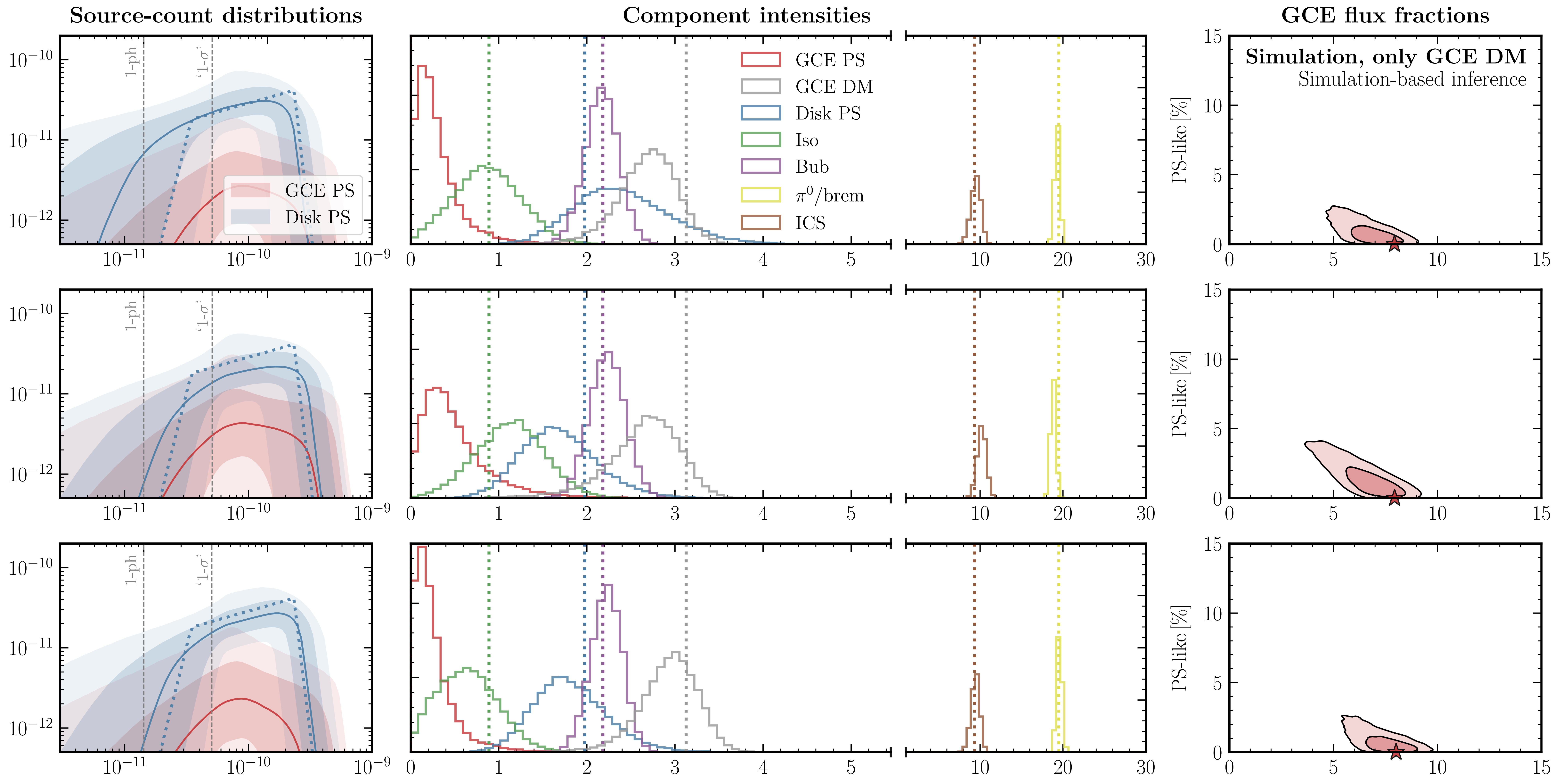
~10% (DM)

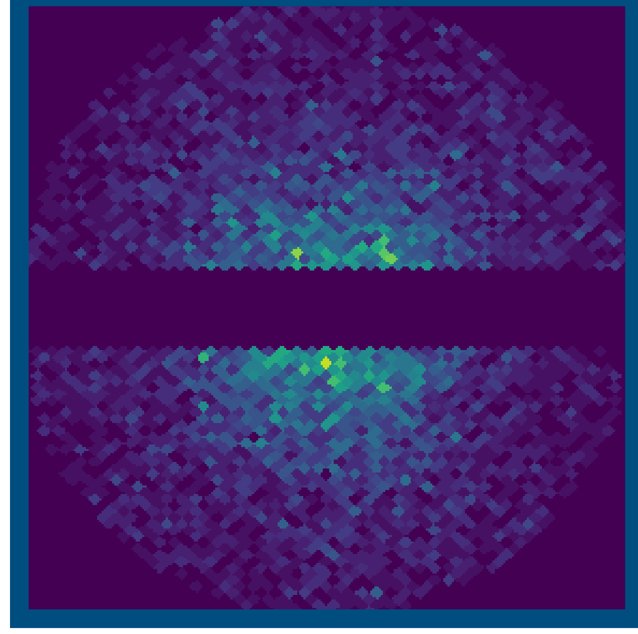
~80% (dif)

Modeling the diffuse Galactic background

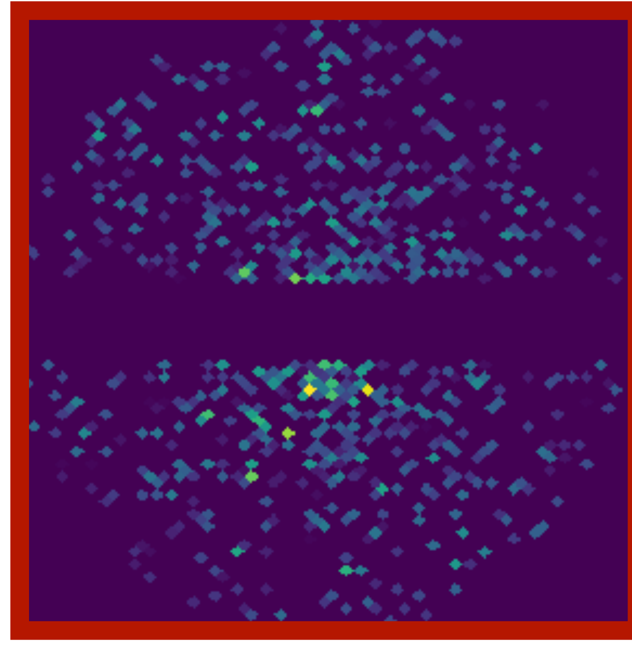


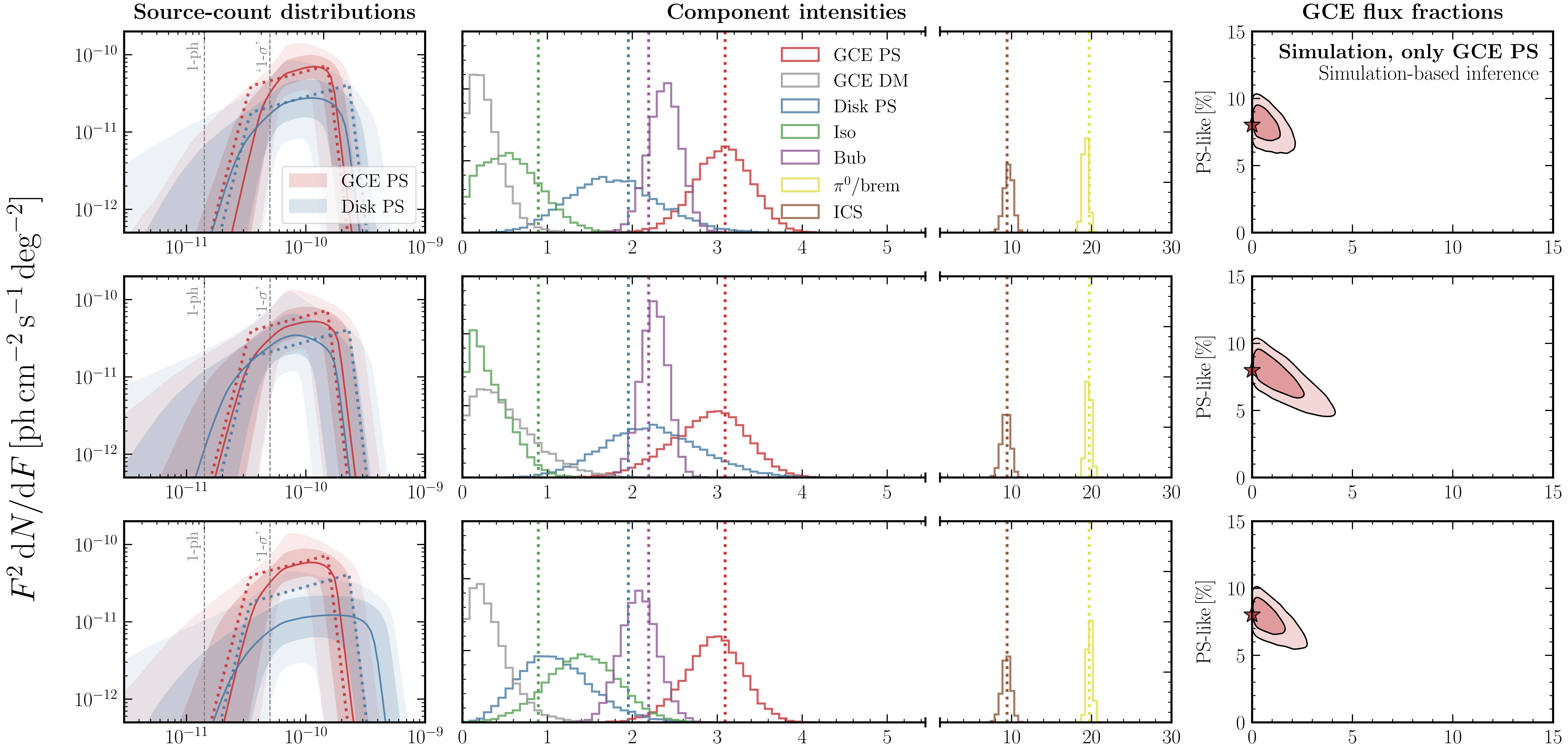
Tests on simulations: dark matter signal





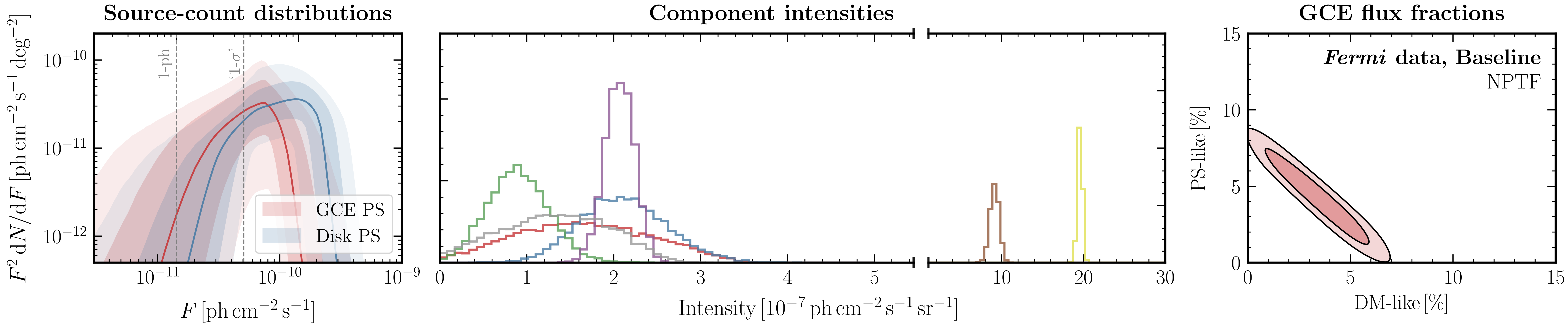
Tests on simulations: point source signal





NPTF vs

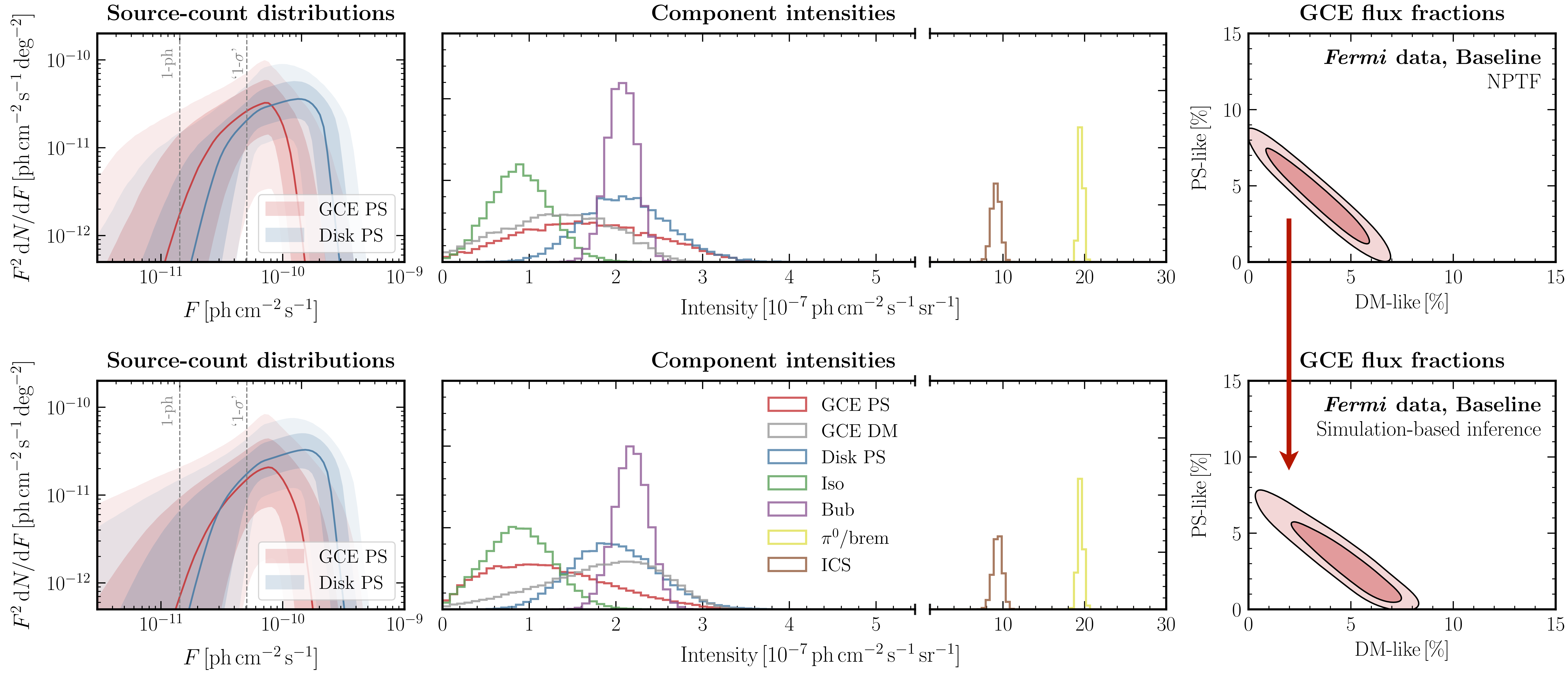
Counts PDF



NPTF vs

Counts PDF

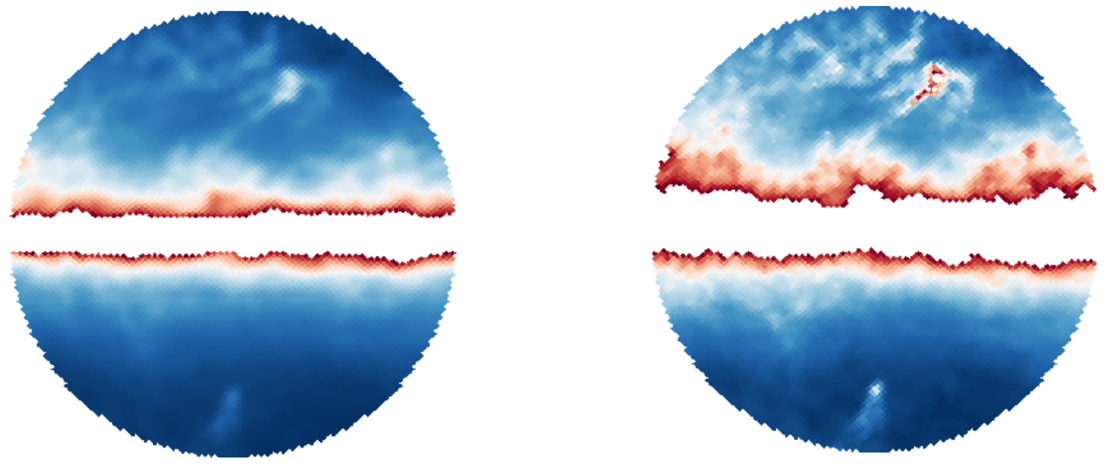
SBI

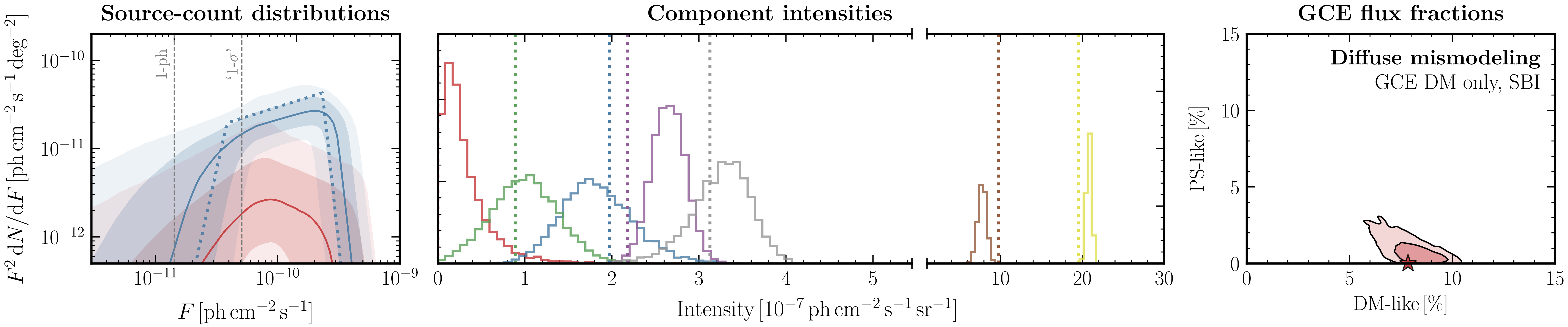


Exploiting more information in the γ-ray maps results in smaller, but still significant PS-like component

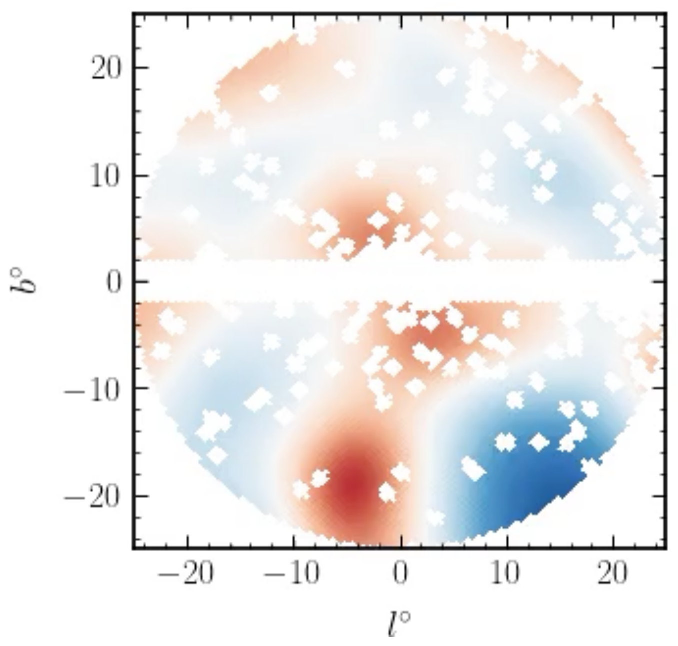
Robustness tests

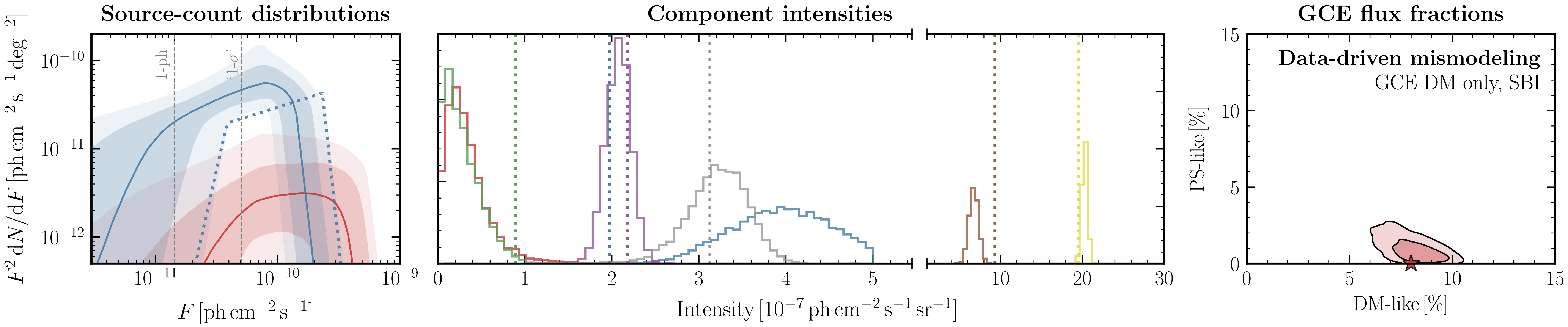
Train with one diffuse model / test with another





Test on GP-modulated background

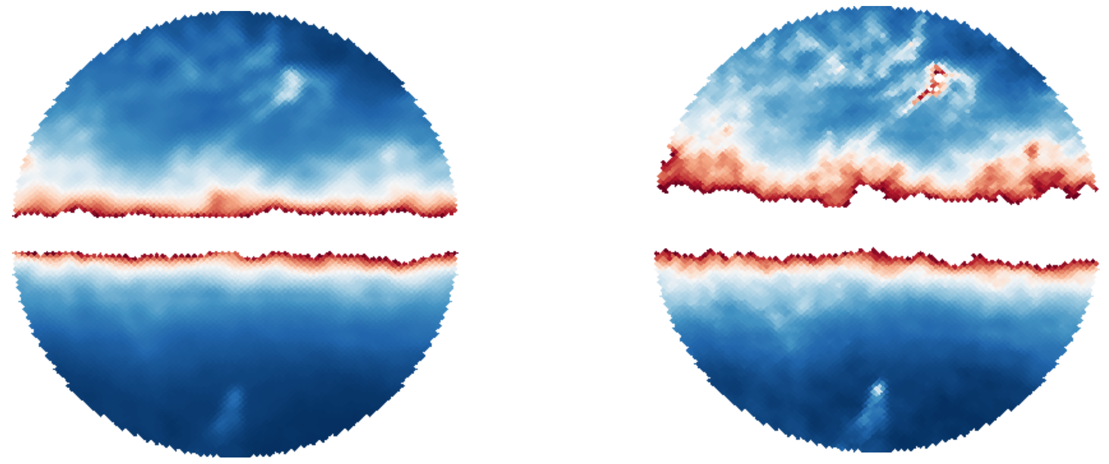


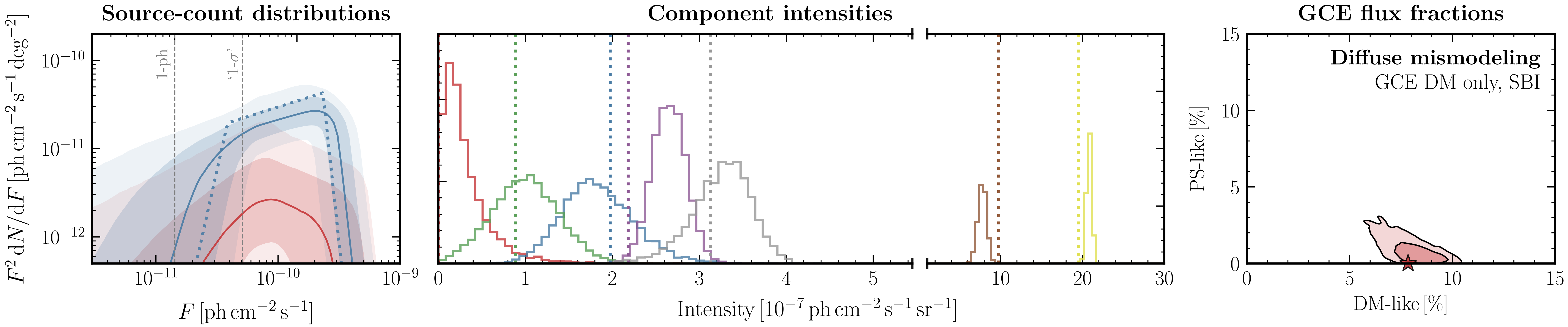


Generally well-behaved under known forms of systematic mismodeling

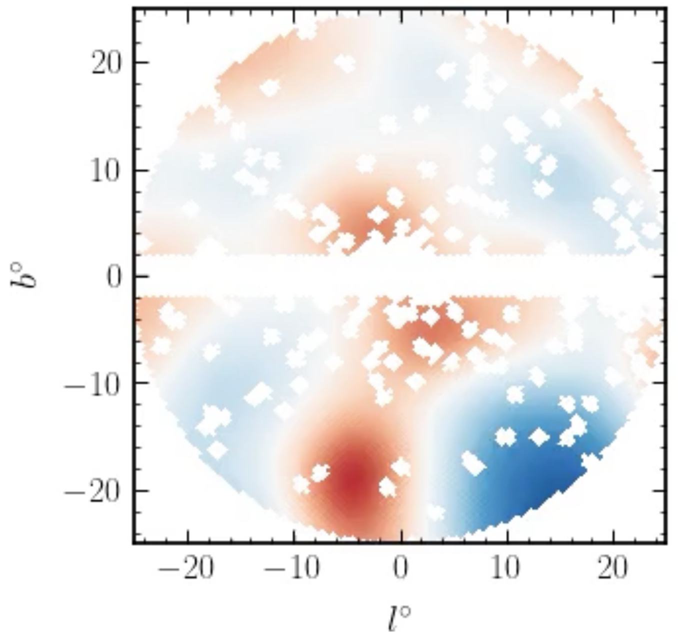
Robustness tests

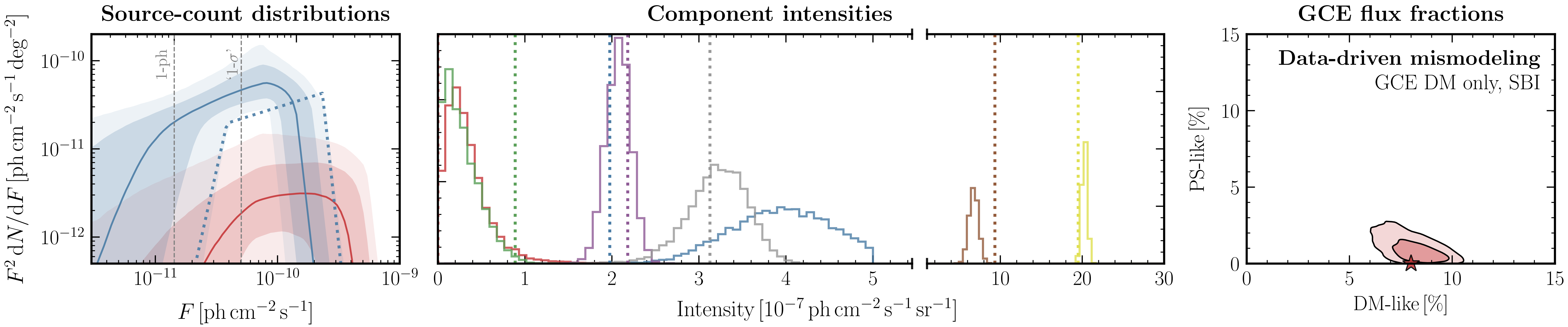
Train with one diffuse model / test with another





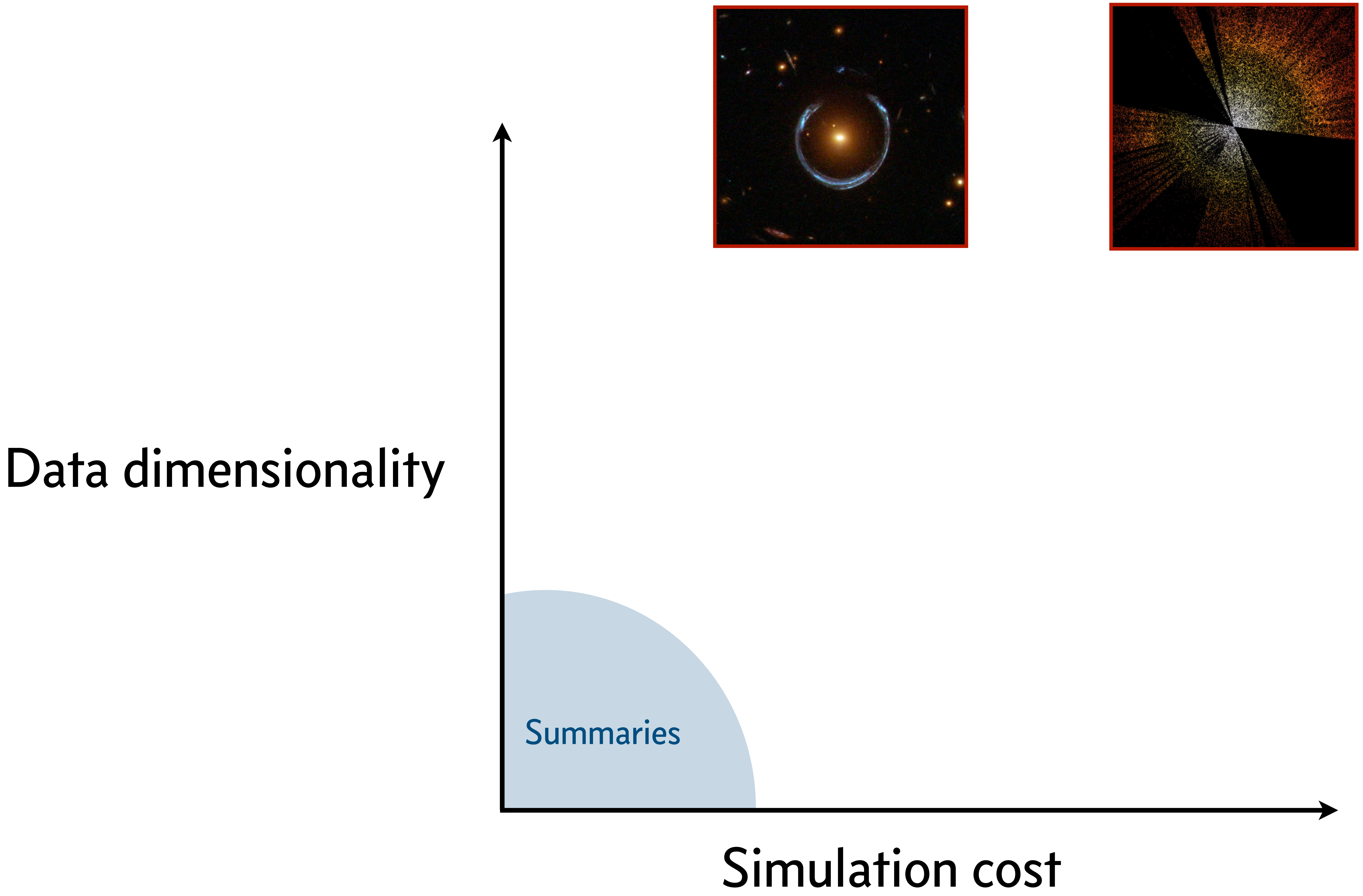
Test on GP-modulated background



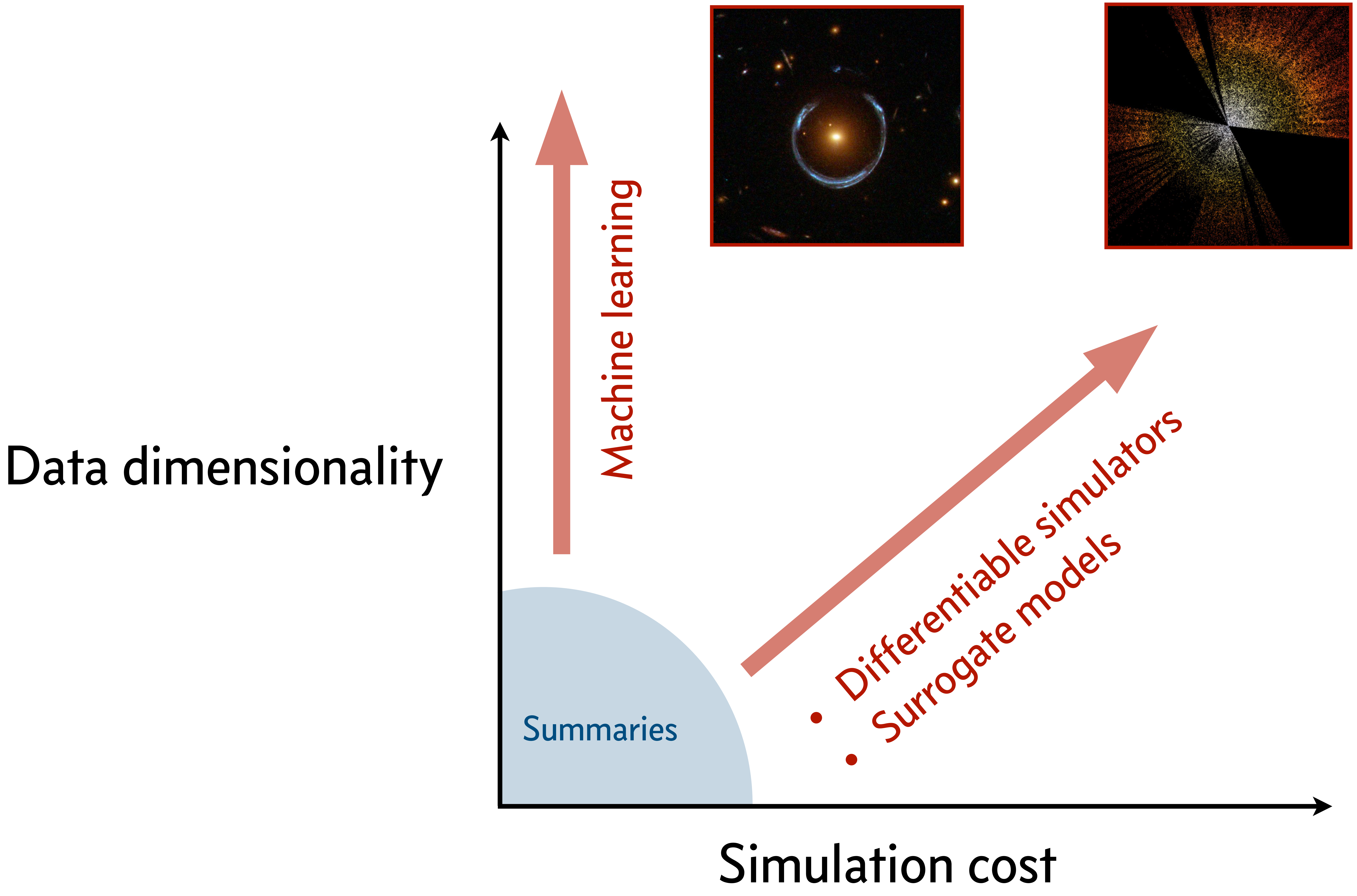


Generally well-behaved under known forms of systematic mismodeling

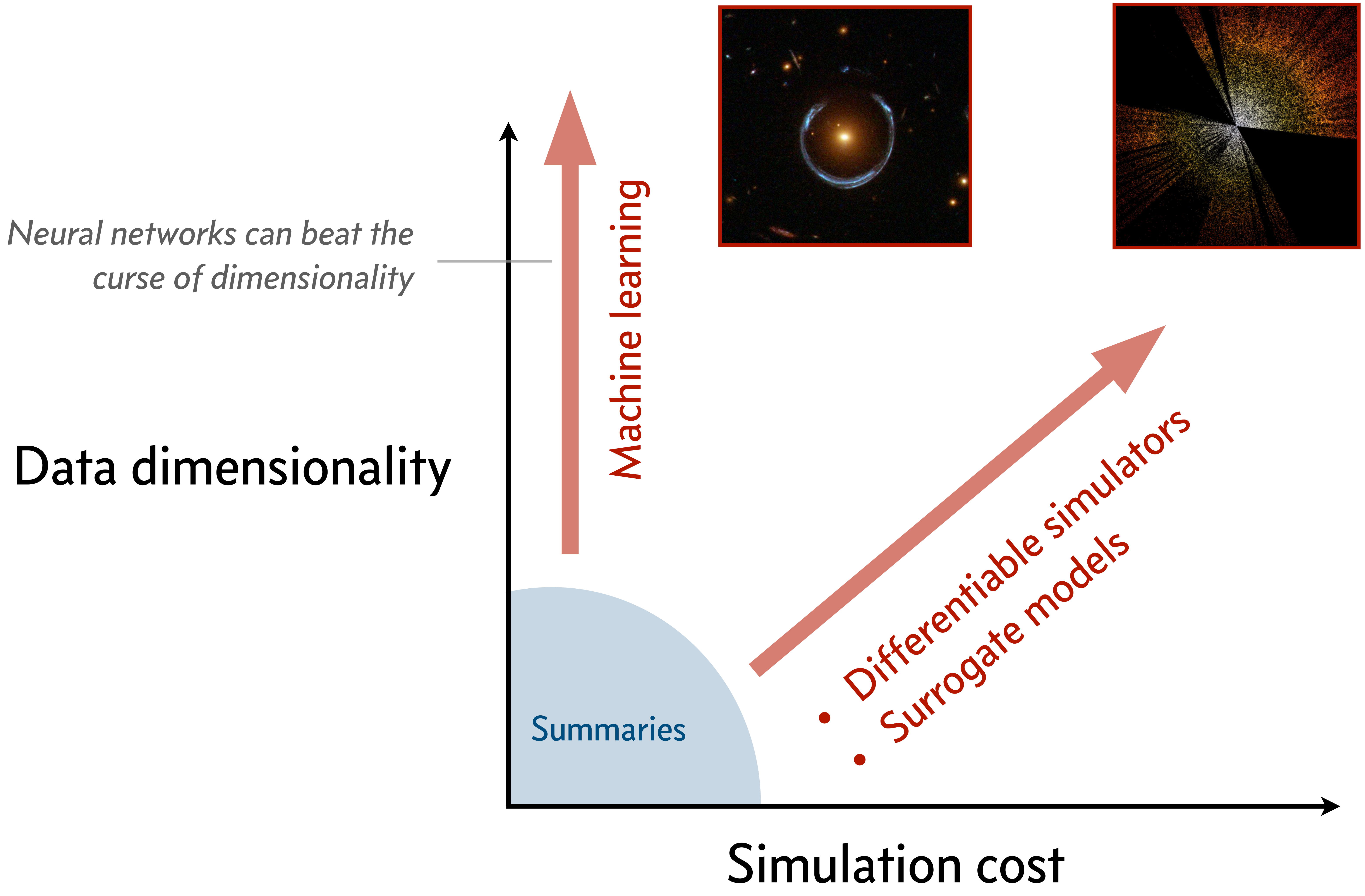
Neural simulation-based inference



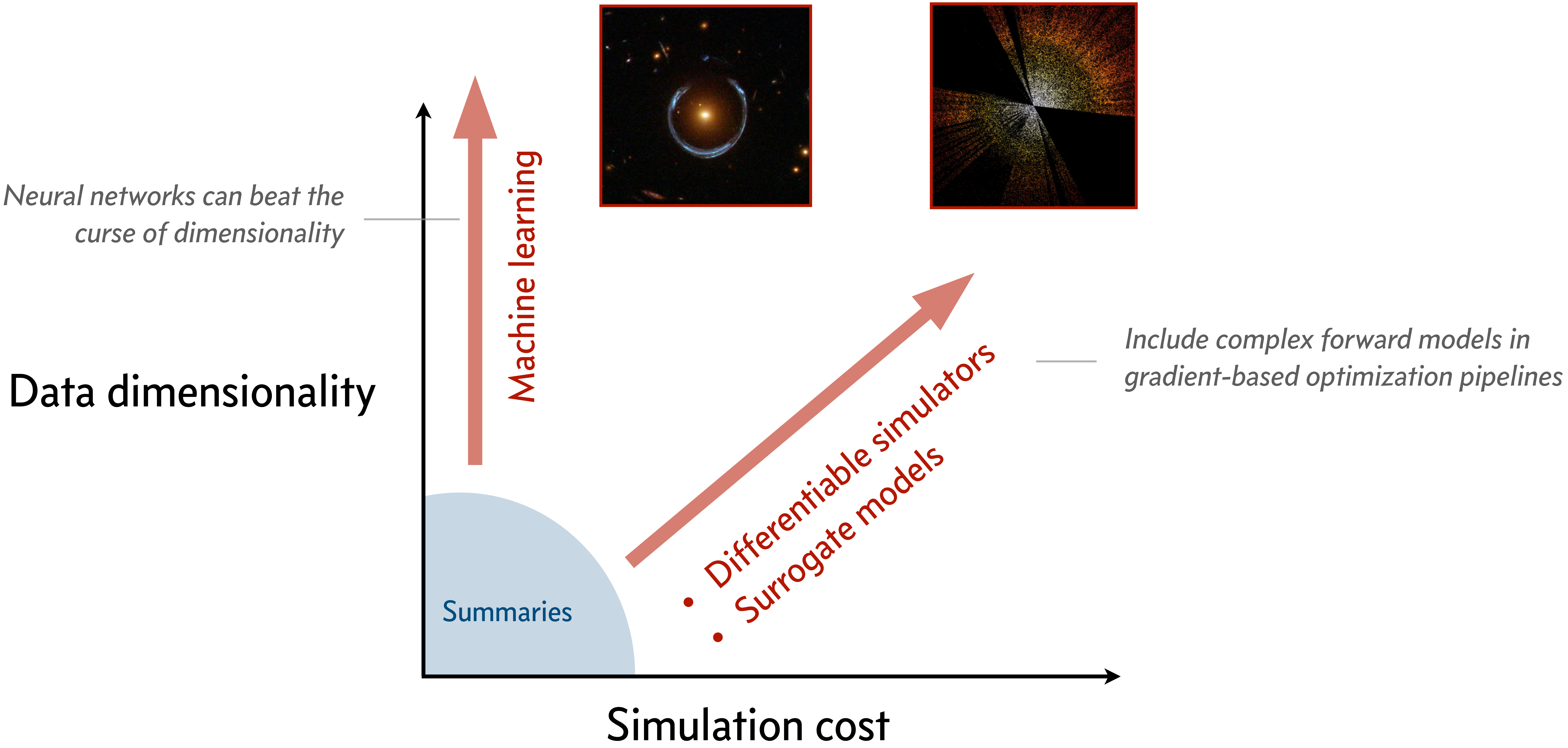
Neural simulation-based inference



Neural simulation-based inference



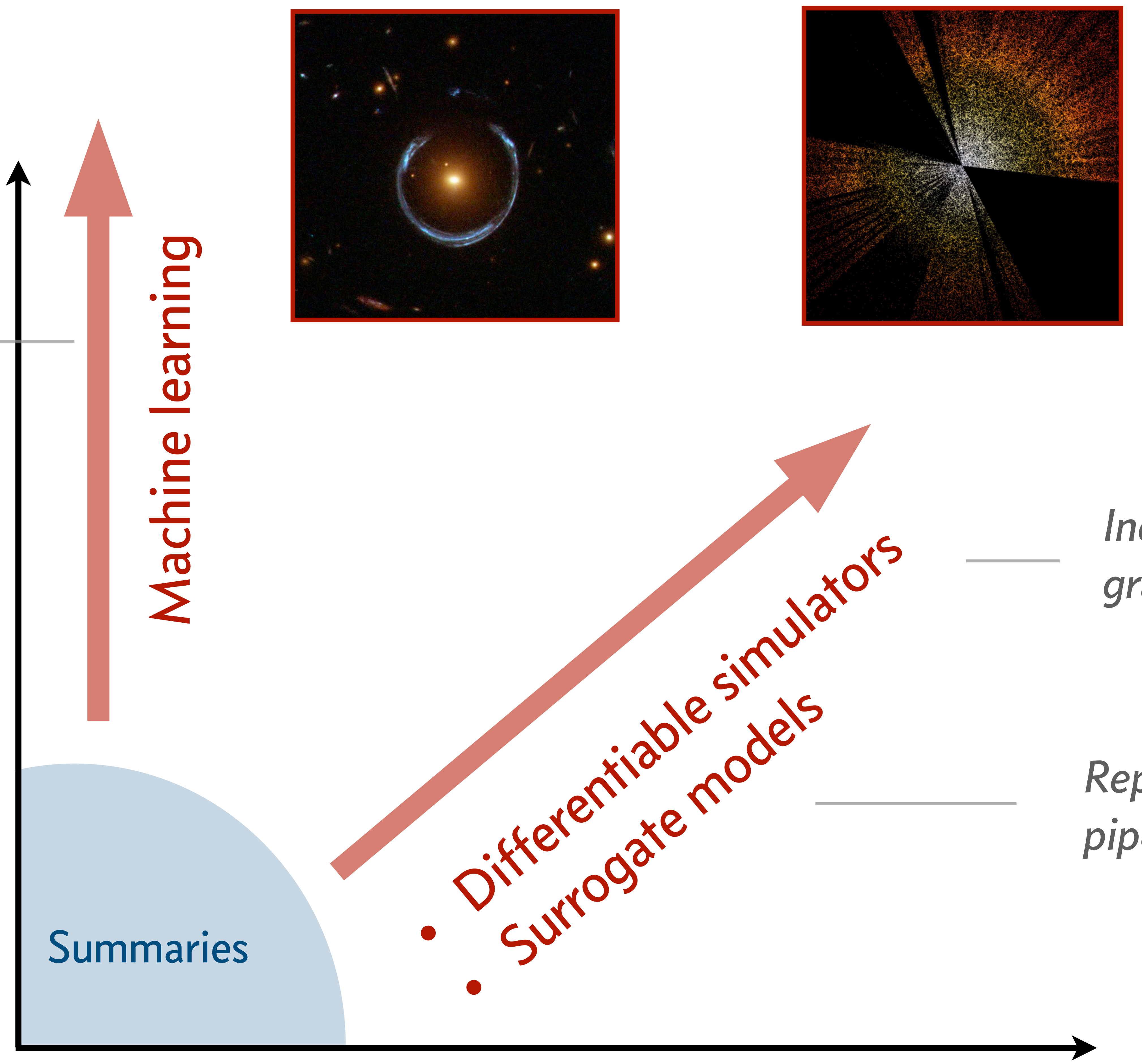
Neural simulation-based inference



Neural simulation-based inference

Neural networks can beat the curse of dimensionality

Data dimensionality



Simulation cost

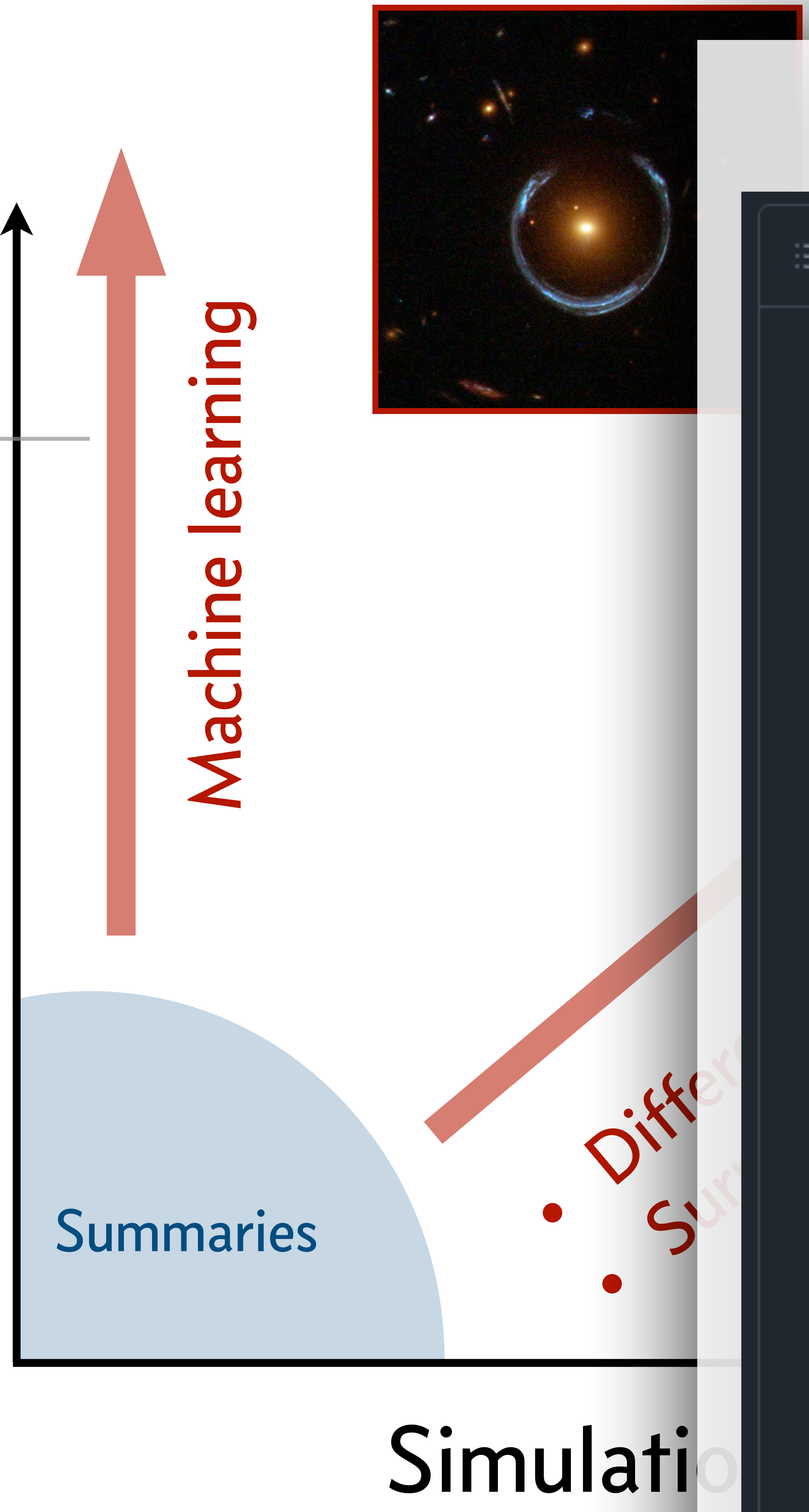
Include complex forward models in gradient-based optimization pipelines

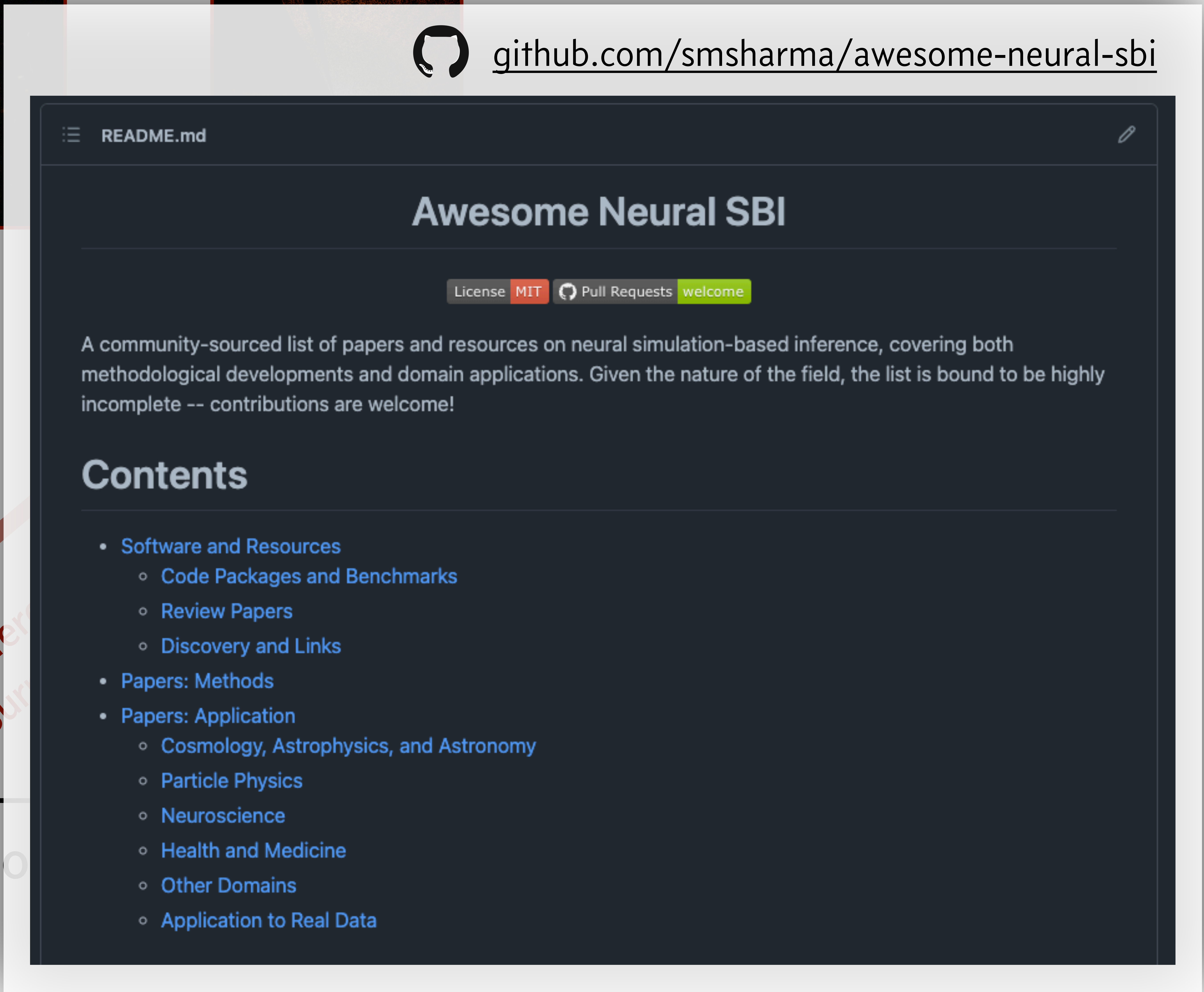
Replace costly parts of simulation pipeline with proxies/emulators

Neural simulation-based inference

Neural networks can beat the curse of dimensionality

Data dimensionality





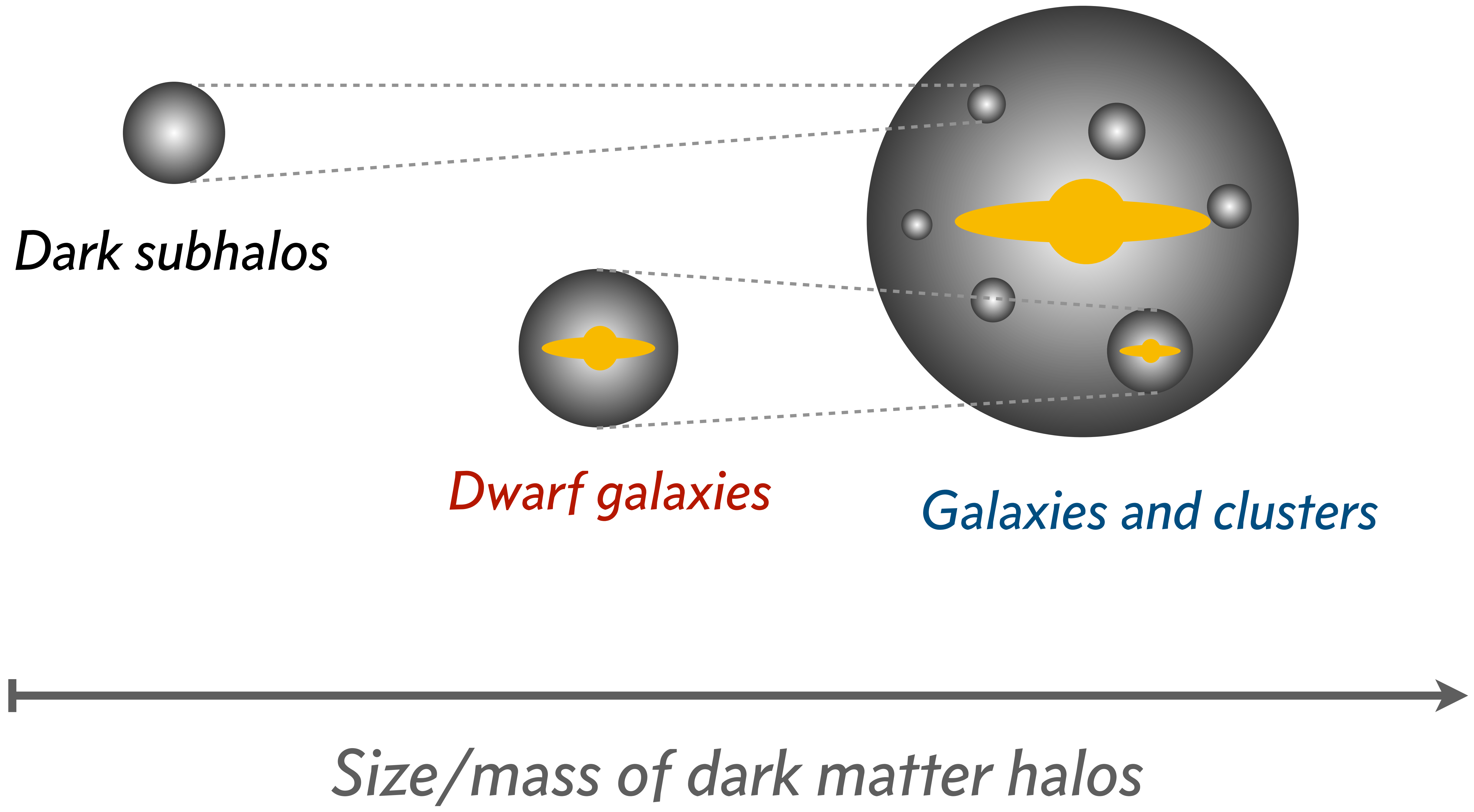
Awesome Neural SBI

License MIT | Pull Requests welcome

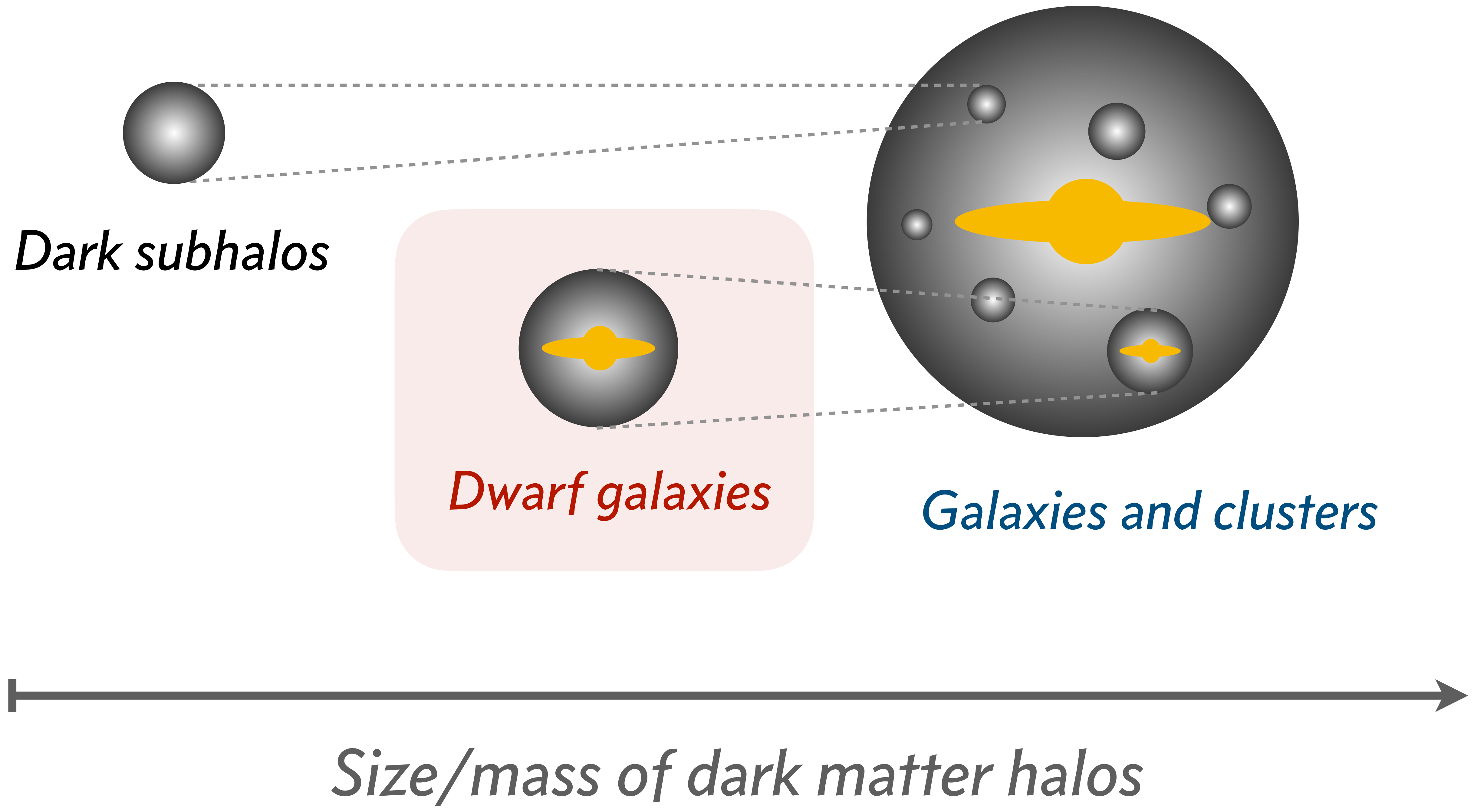
A community-sourced list of papers and resources on neural simulation-based inference, covering both methodological developments and domain applications. Given the nature of the field, the list is bound to be highly incomplete -- contributions are welcome!

Contents

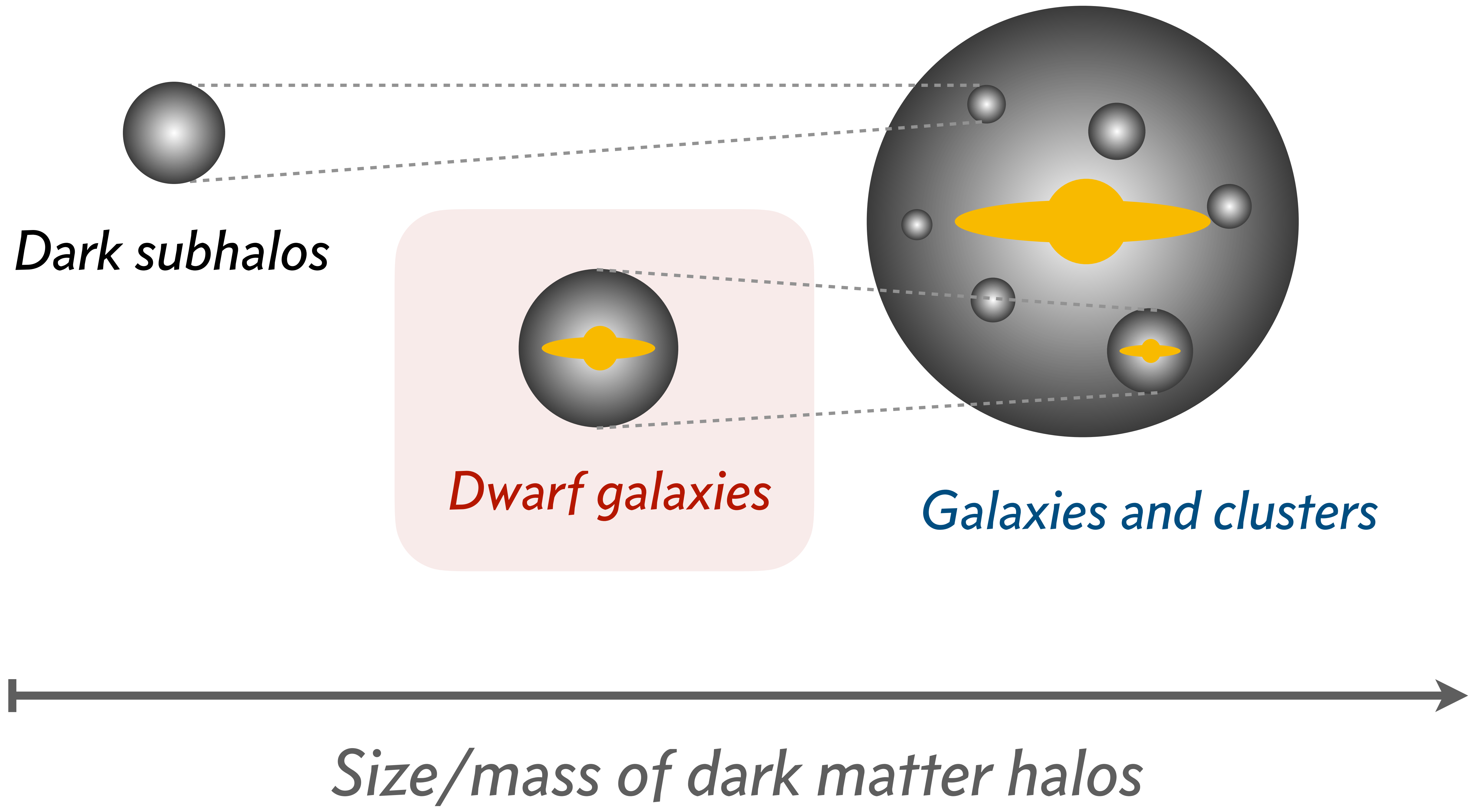
Dwarf spheroidal galaxies



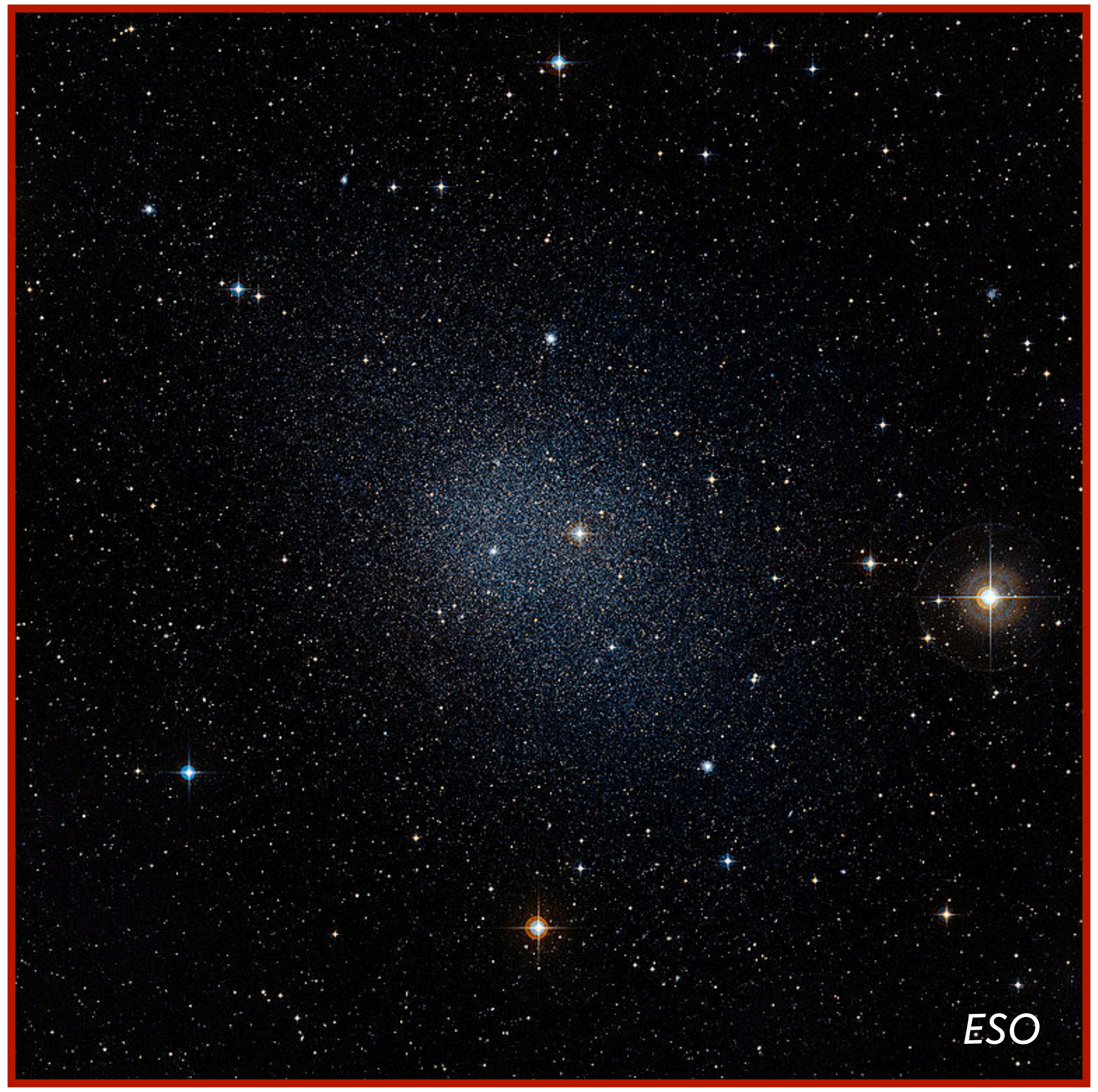
Dwarf spheroidal galaxies



Dwarf spheroidal galaxies



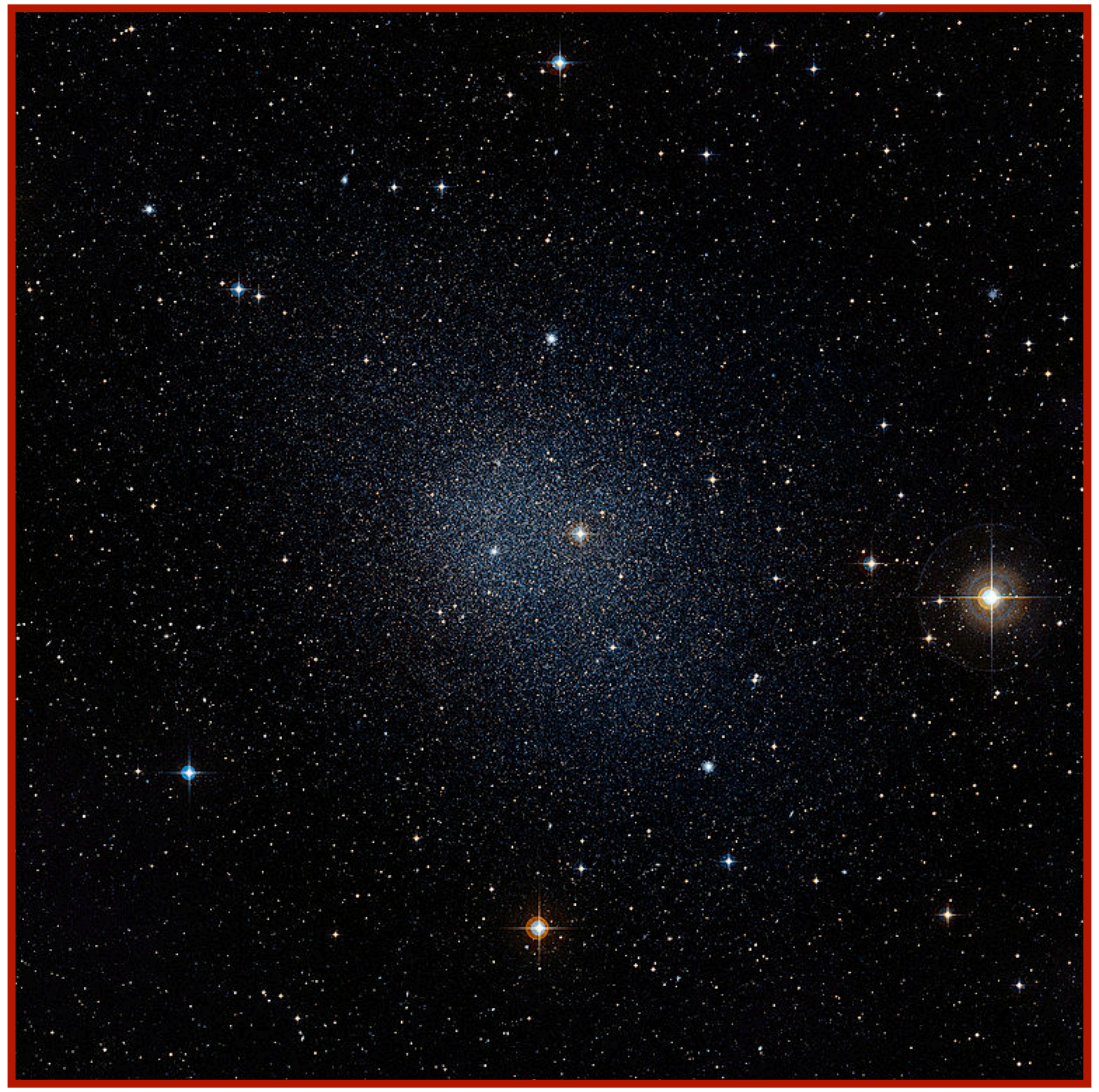
Fornax dwarf galaxy

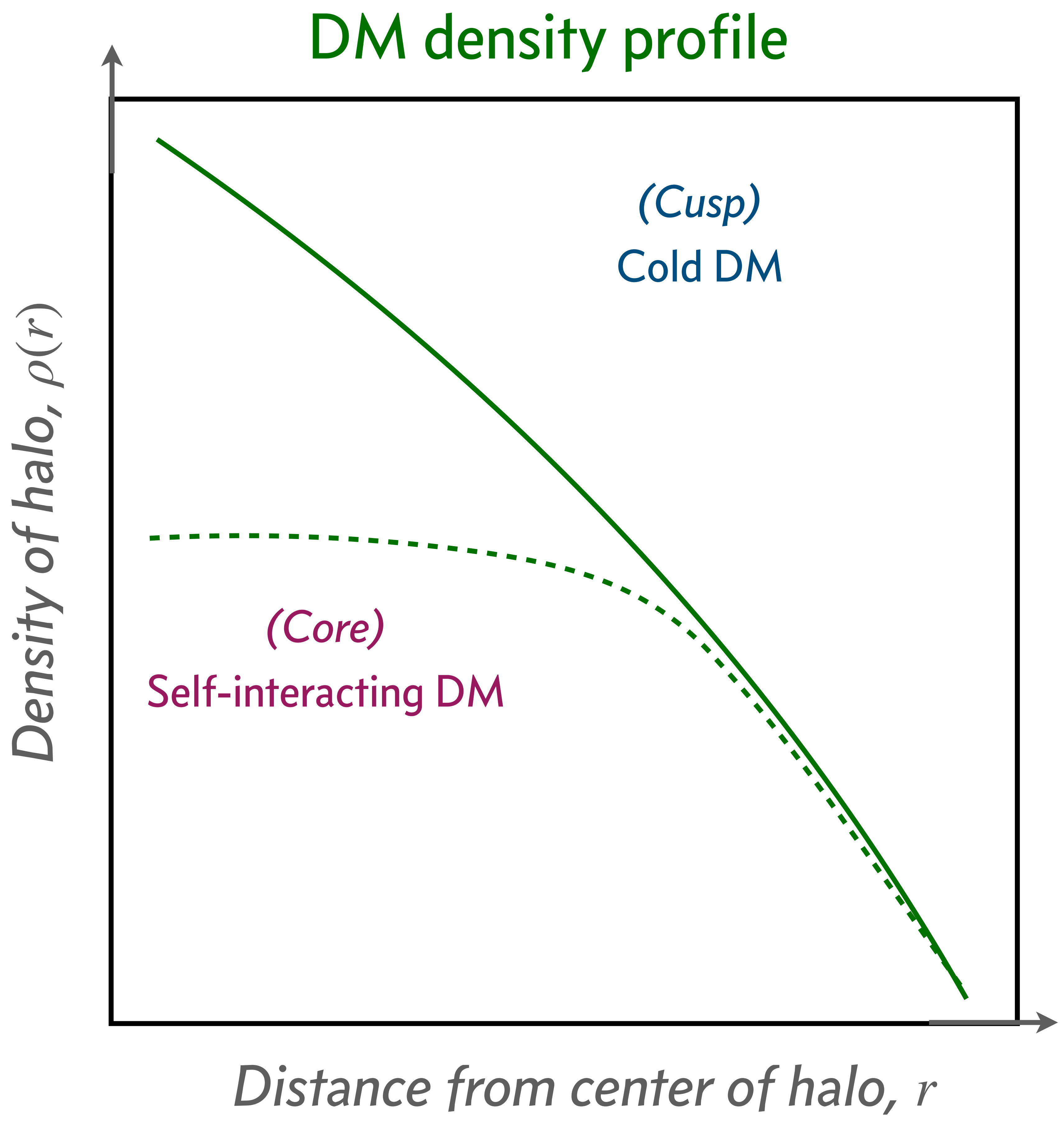


Dwarf galaxies and halo shapes

Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

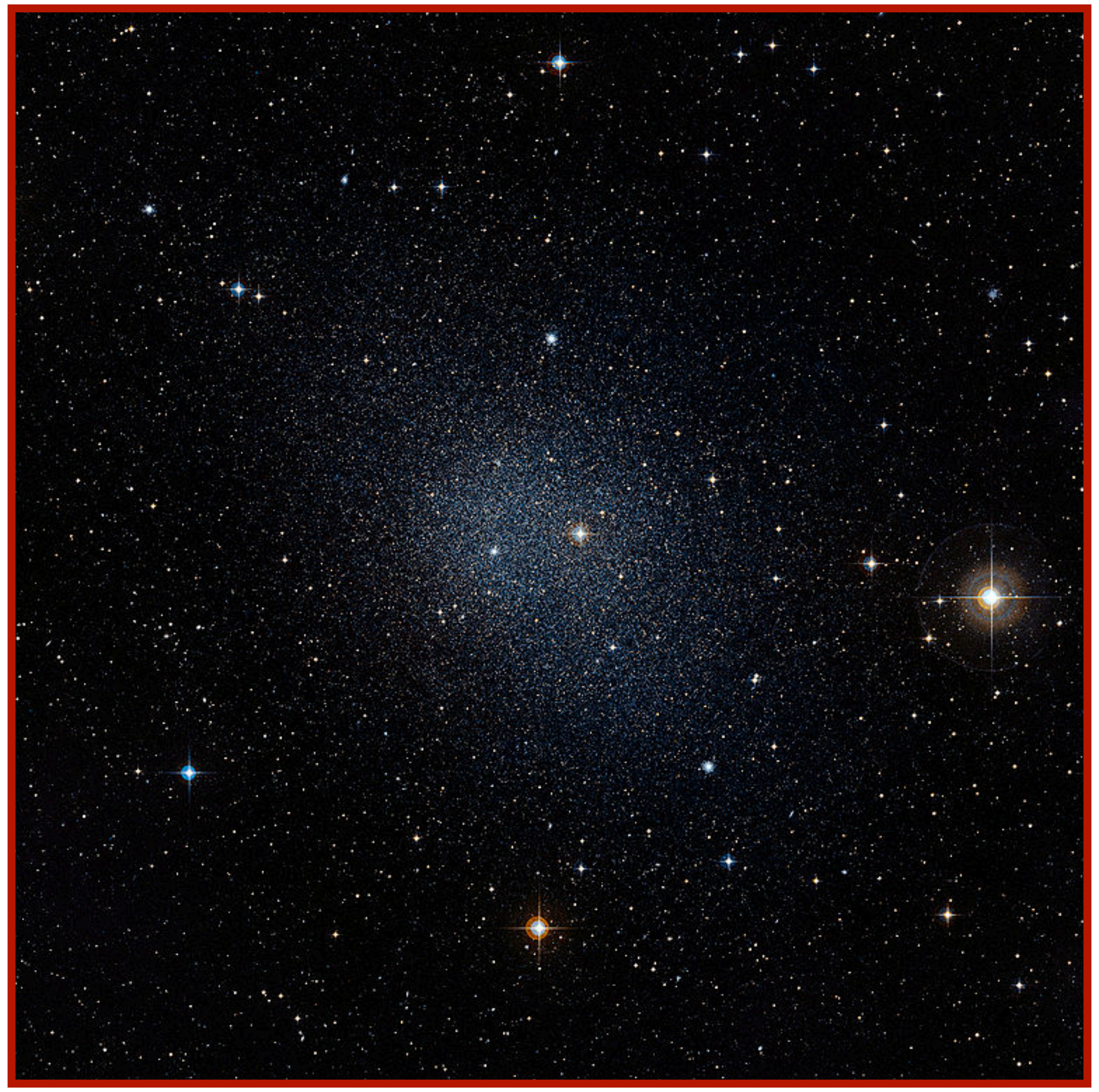


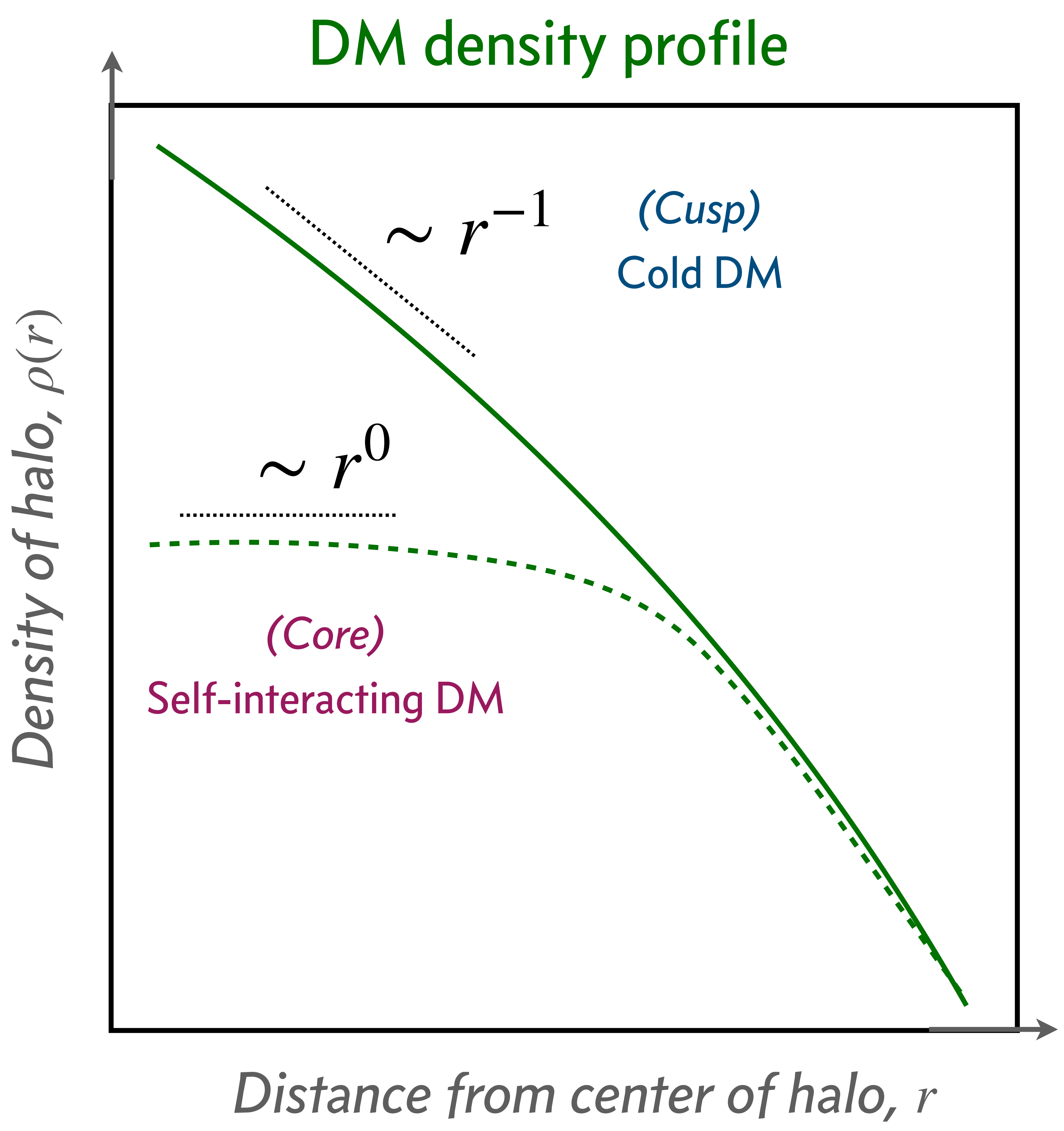


Dwarf galaxies and halo shapes

Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

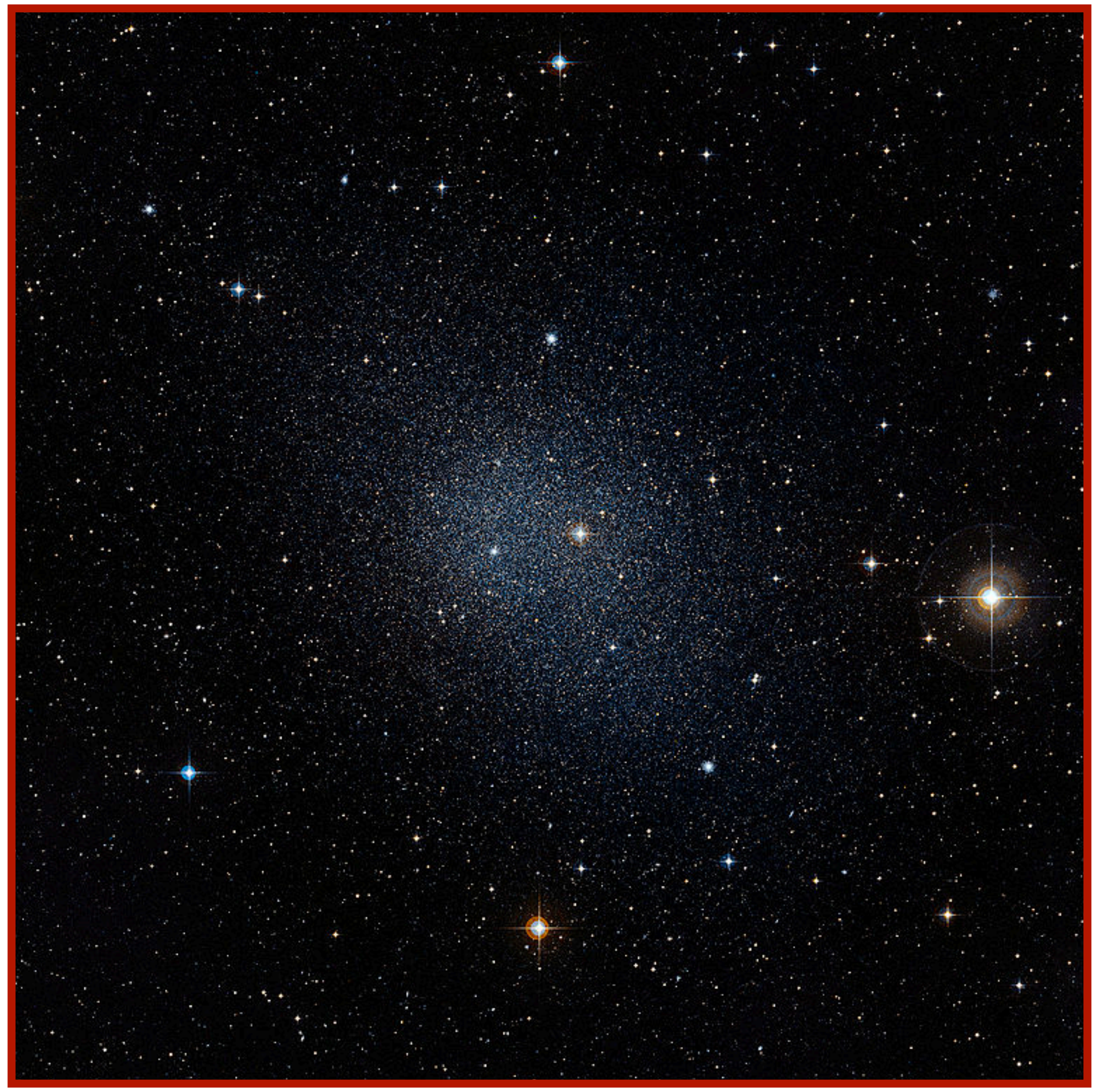


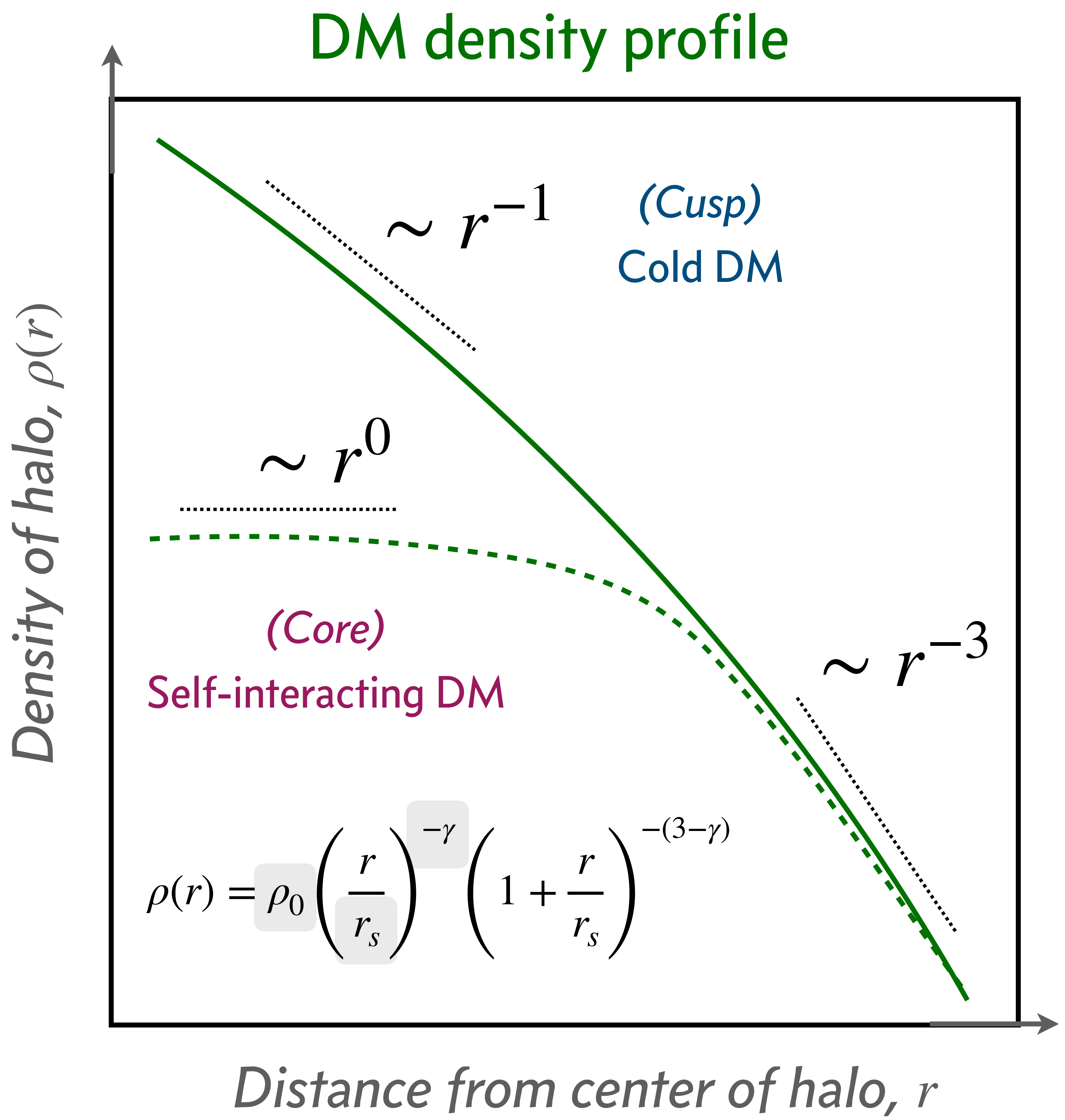


Dwarf galaxies and halo shapes

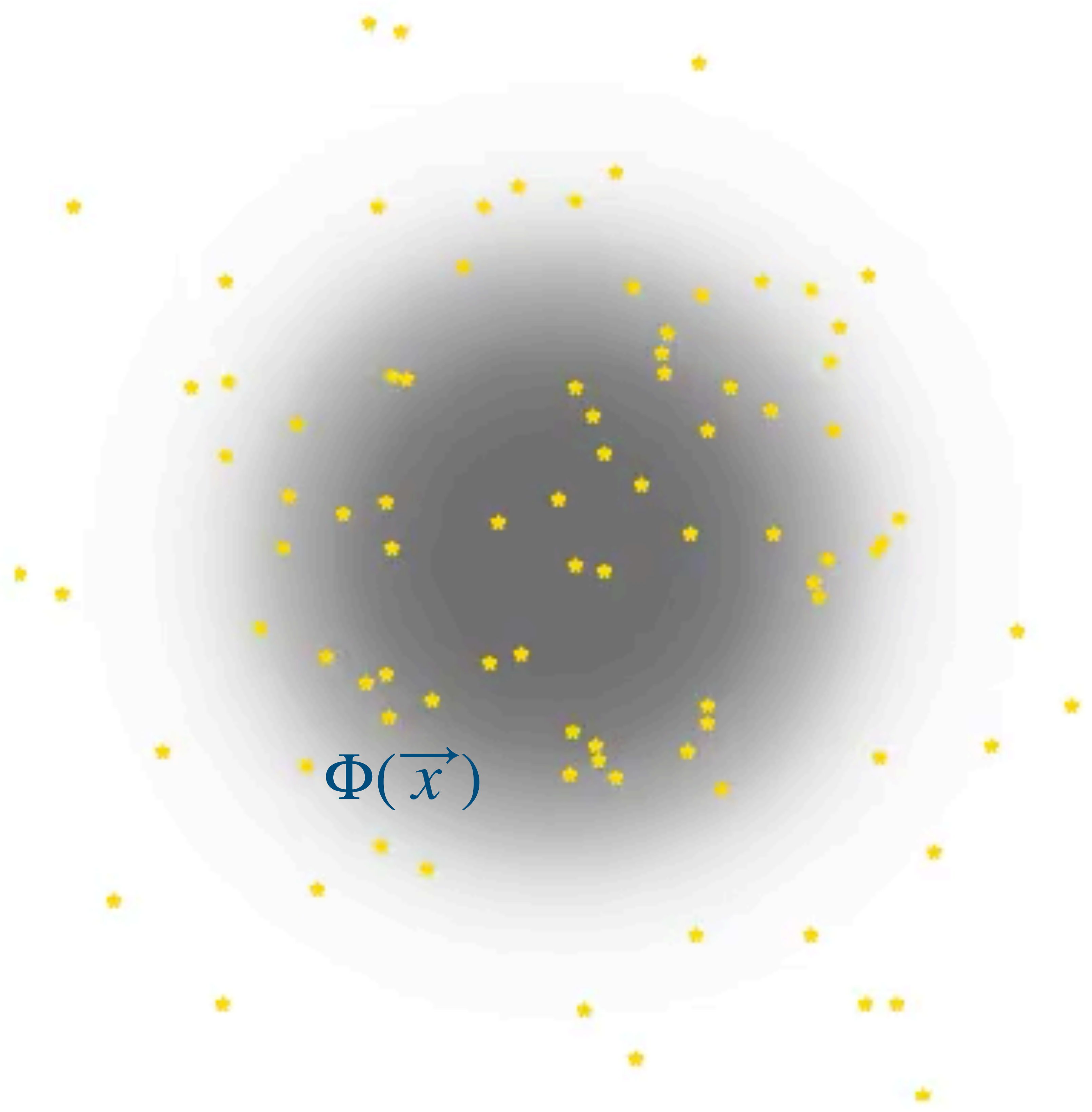
Dwarf galaxies are ideal targets for probing the shapes of DM halos

Fornax dwarf galaxy

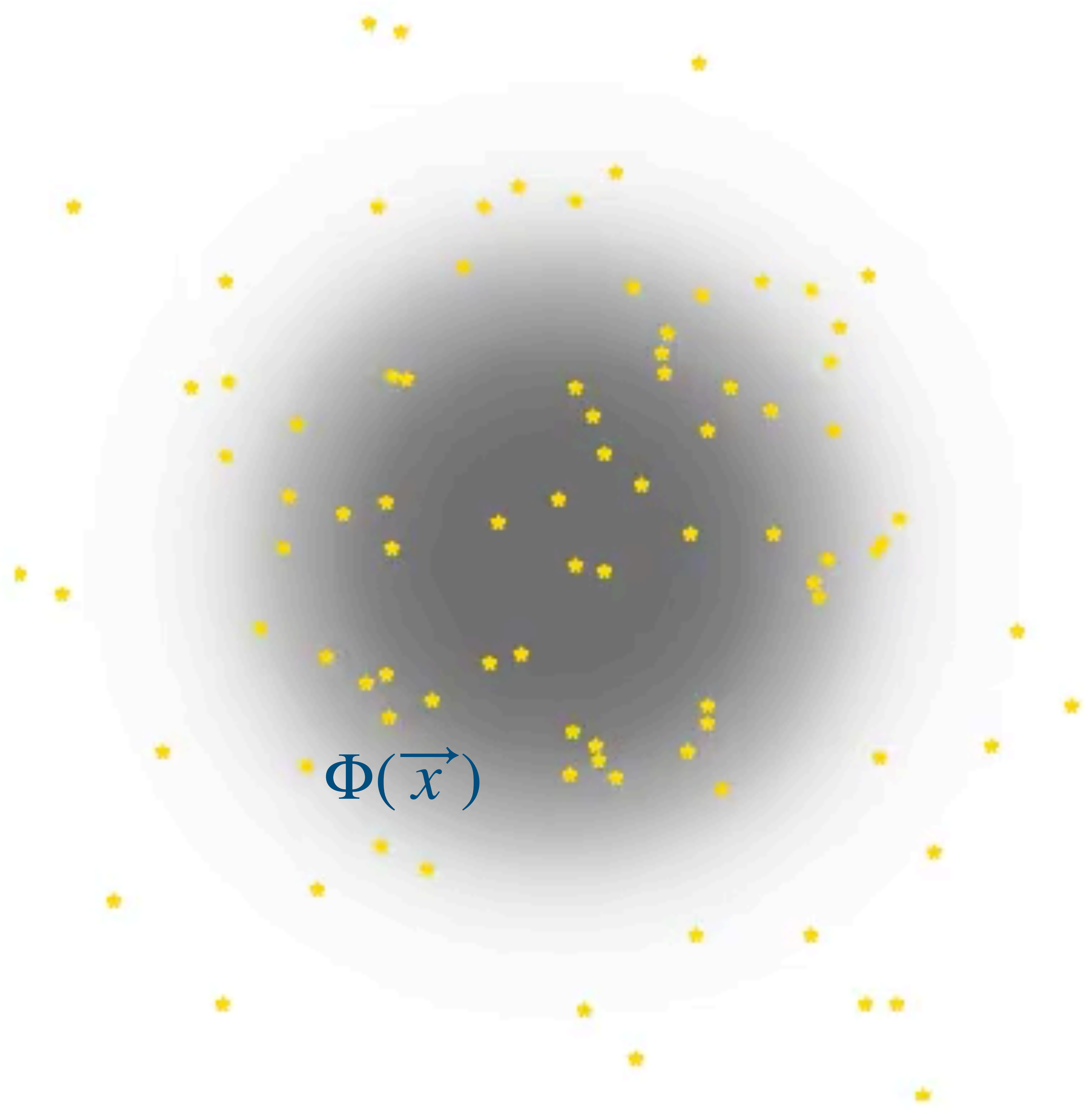




From stellar kinematics to halo shapes: *Jeans modeling*



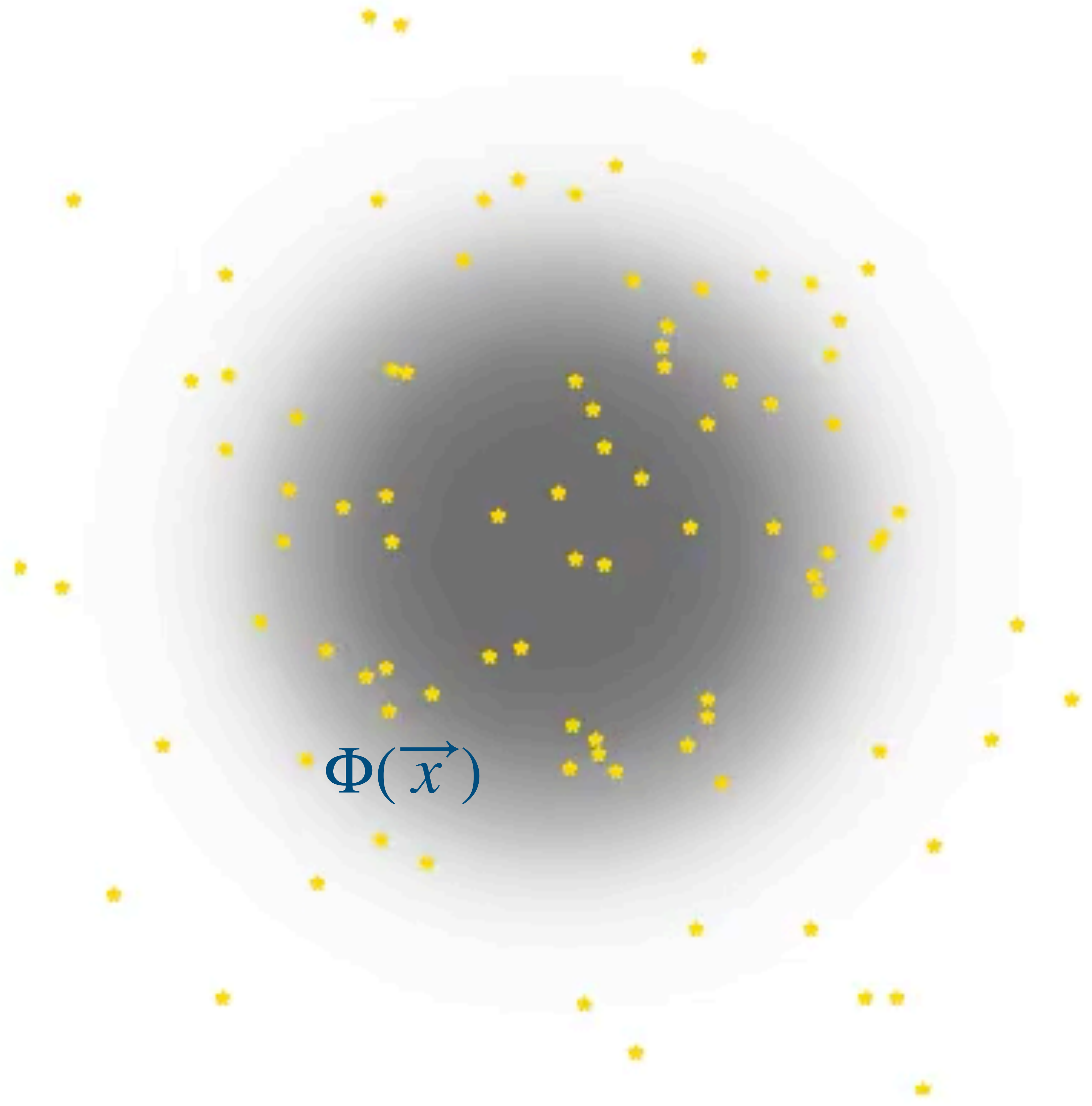
From stellar kinematics to halo shapes: *Jeans modeling*



From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

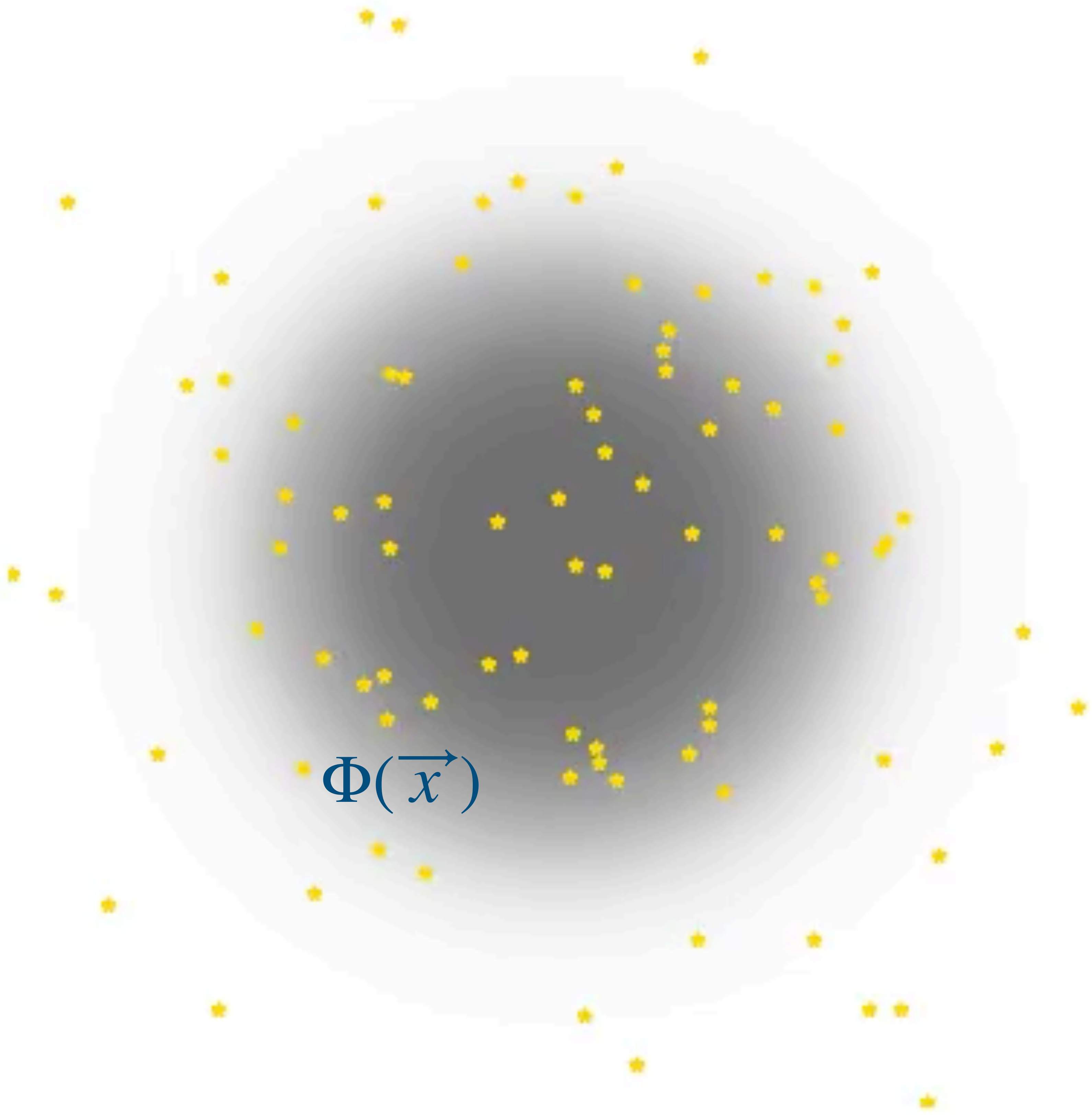
$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



Phase space density and its *moments*

$$n(\vec{x}) = \int \mathrm{d}^3v\, f(\vec{x}, \vec{v})$$

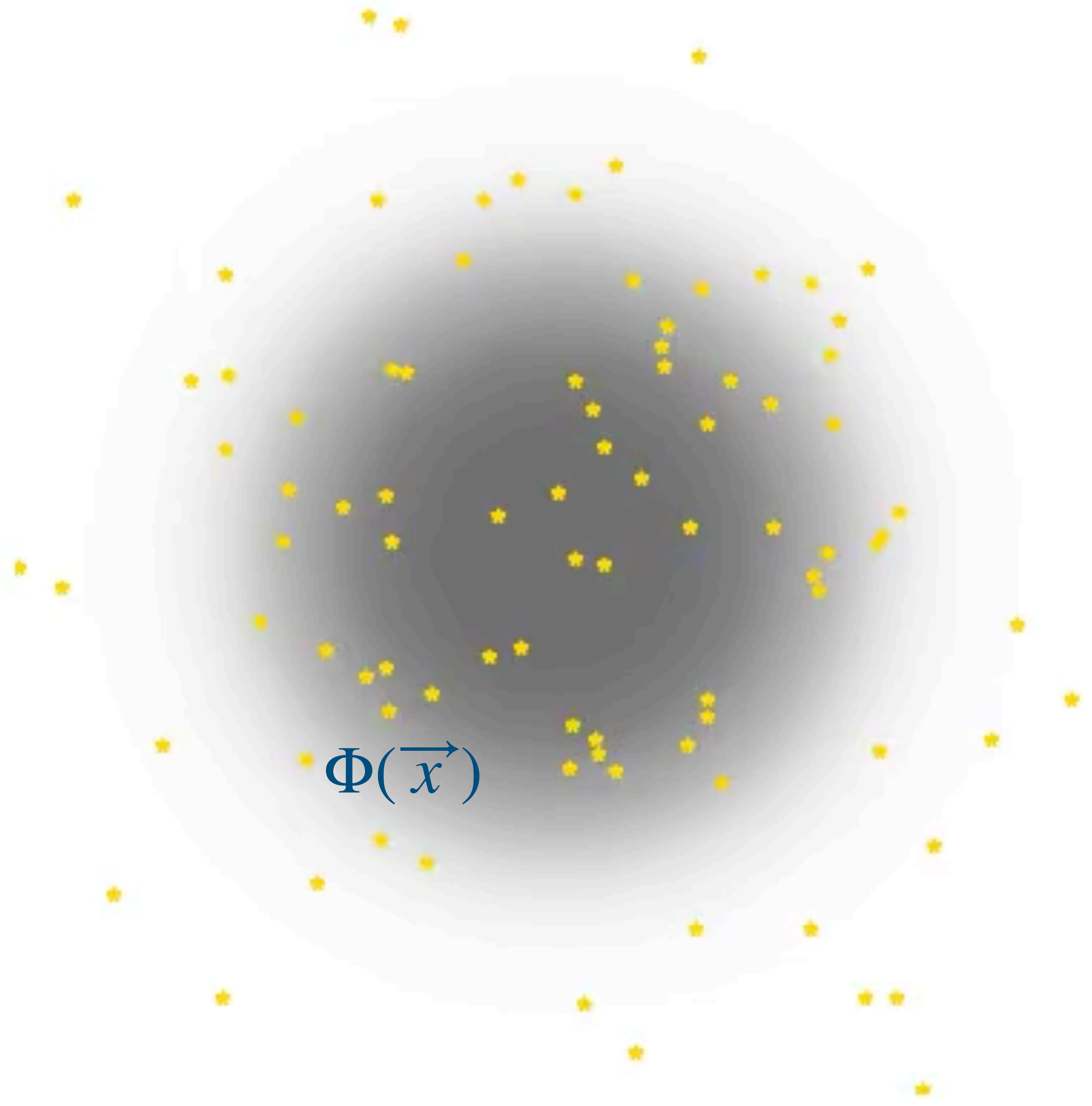
$$\langle v_i(\vec{x}) \rangle = \int \mathrm{d}^3v\, v_i\, f(\vec{x}, \vec{v})$$

$$\sigma_{ij}(\vec{x}) = \int \mathrm{d}^3v\, (v_i - \bar{v}_i)(v_j - \bar{v}_j) f(\vec{x}, \vec{v})$$

From stellar kinematics to halo shapes: *Jeans modeling*

Phase-space density

$$\mathrm{d}n(\vec{x}, \vec{v}) \propto f(\vec{x}, \vec{v})\, \mathrm{d}^3x\, \mathrm{d}^3v$$



Phase space density and its *moments*

$$n(\vec{x}) = \int \mathrm{d}^3v\, f(\vec{x}, \vec{v})$$

$$\langle v_i(\vec{x}) \rangle = \int \mathrm{d}^3v\, v_i\, f(\vec{x}, \vec{v})$$

$$\sigma_{ij}(\vec{x}) = \int \mathrm{d}^3v\, (v_i - \bar{v}_i)(v_j - \bar{v}_j) f(\vec{x}, \vec{v})$$

Jeans equations connect moments of $f(\vec{x}, \vec{v})$ to $\Phi(\vec{x})$

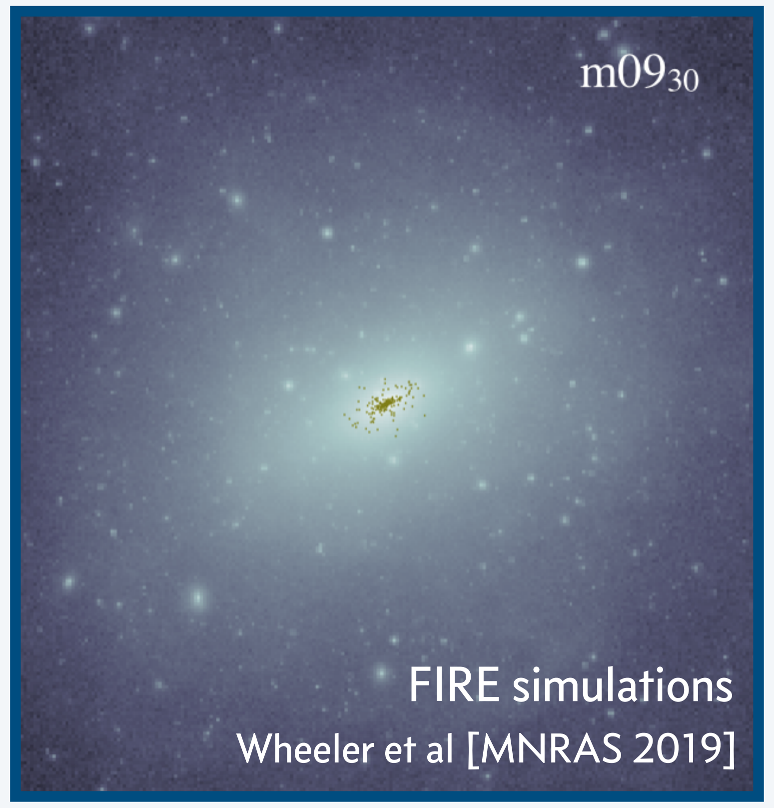
$$n\langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i} + n \frac{\partial \Phi}{\partial x_j} + \frac{\partial \left[n\sigma_{ij}^2 \right]}{\partial x_i} = 0$$

Limitations of Jeans modeling

Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential



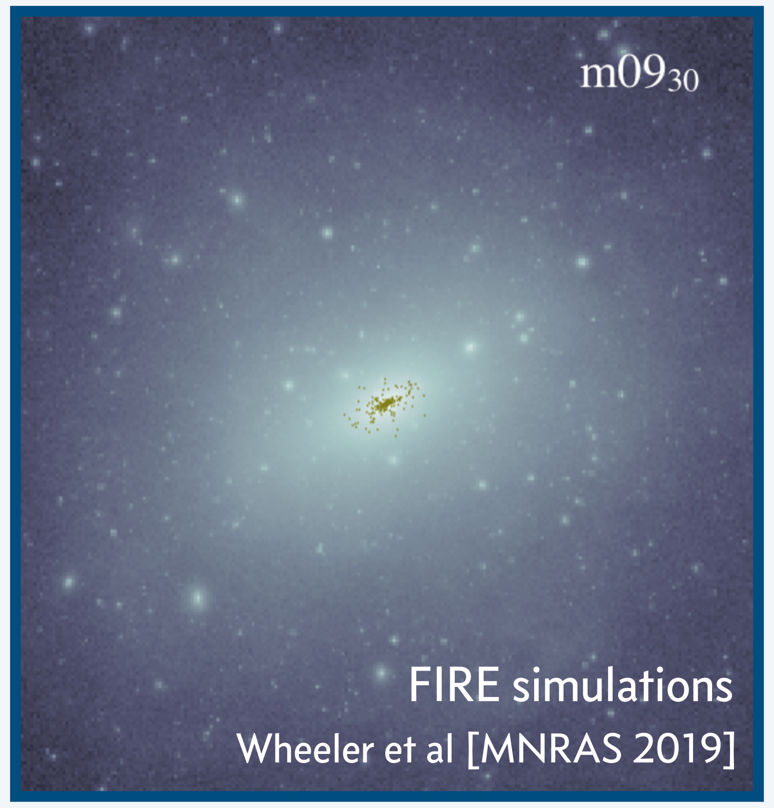
FIRE simulations
Wheeler et al [MNRAS 2019]

Limitations of Jeans modeling

Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential



FIRE simulations
Wheeler et al [MNRAS 2019]

Reliance on moments of $f(\vec{x}, \vec{v})$

$$n\langle v_i\rangle \frac{\partial \langle v_j\rangle}{\partial x_i} + n\frac{\partial \Phi}{\partial x_j} + \frac{\partial \left[n\sigma_{ij}^2\right]}{\partial x_i} = 0$$

- Simplified description of the data = *loss of information*
- Typically only *3* phase-space coordinates available:

$$\{\vec{r}, \vec{v}\} \longrightarrow \{\vec{r}_\perp, \vec{v}_{\rm los}\}$$

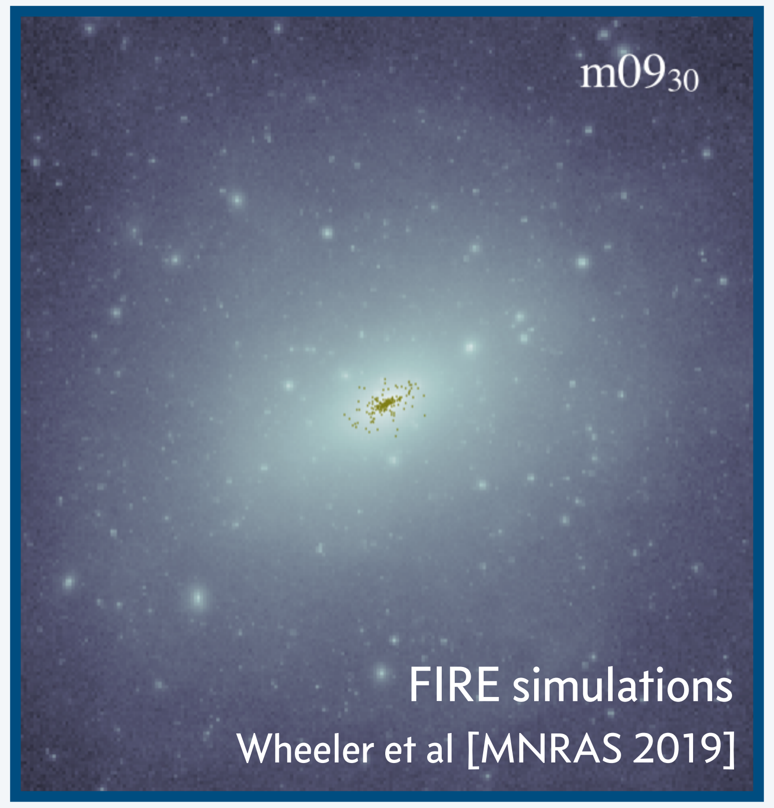
- Degeneracy between DM density profile and anisotropy configuration of stellar orbits
- Noisy estimates of $\sigma_r^2(r)$, $n(r)$ and derivatives

Limitations of Jeans modeling

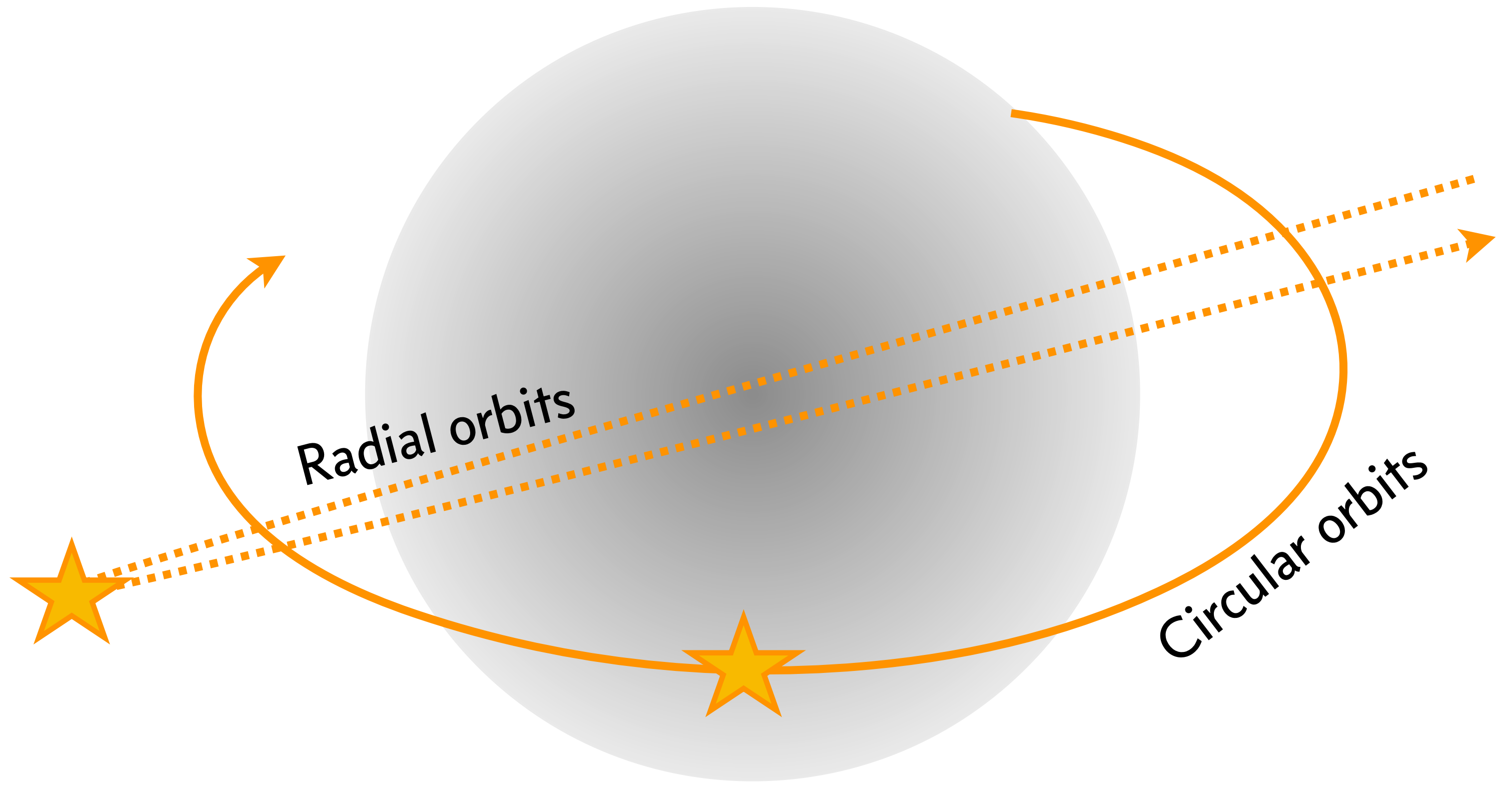
Assumptions about the data-generating process

Challenging to include:

- Non-equilibrium effects
- Asphericity
- Baryonic feedback
- Host potential



FIRE simulations
Wheeler et al [MNRAS 2019]



Reliance on moments of $f(\vec{x}, \vec{v})$

$$n\langle v_i\rangle \frac{\partial \langle v_j\rangle}{\partial x_i} + n\frac{\partial \Phi}{\partial x_j} + \frac{\partial \left[n\sigma_{ij}^2\right]}{\partial x_i} = 0$$

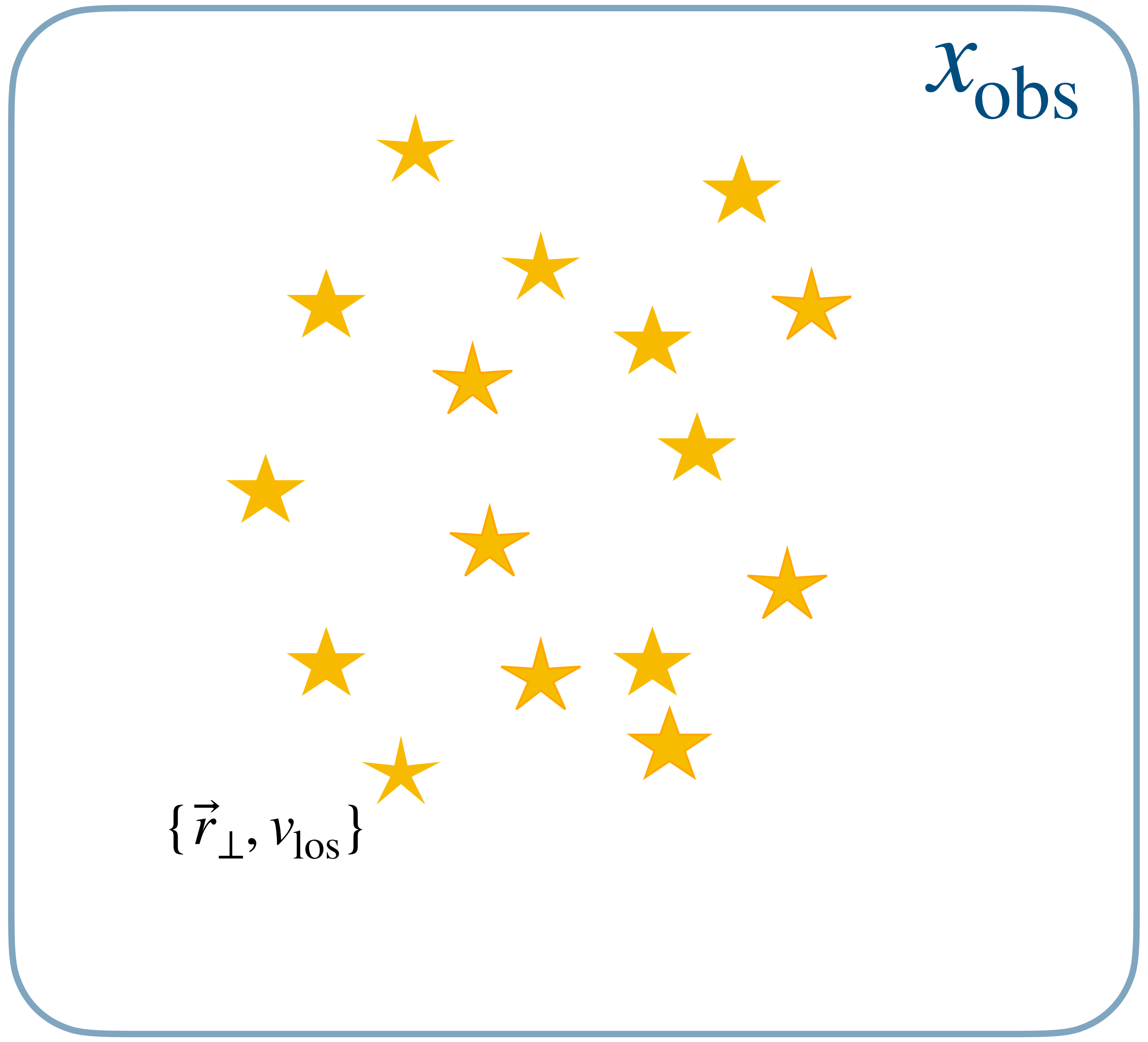
- Simplified description of the data = *loss of information*
- Typically only *3* phase-space coordinates available:

$$\{\vec{r}, \vec{v}\} \longrightarrow \{\vec{r}_\perp, \vec{v}_{\rm los}\}$$

- Degeneracy between DM density profile and anisotropy configuration of stellar orbits
- Noisy estimates of $\sigma_r^2(r)$, $n(r)$ and derivatives

Simulation-based inference for dwarf galaxies

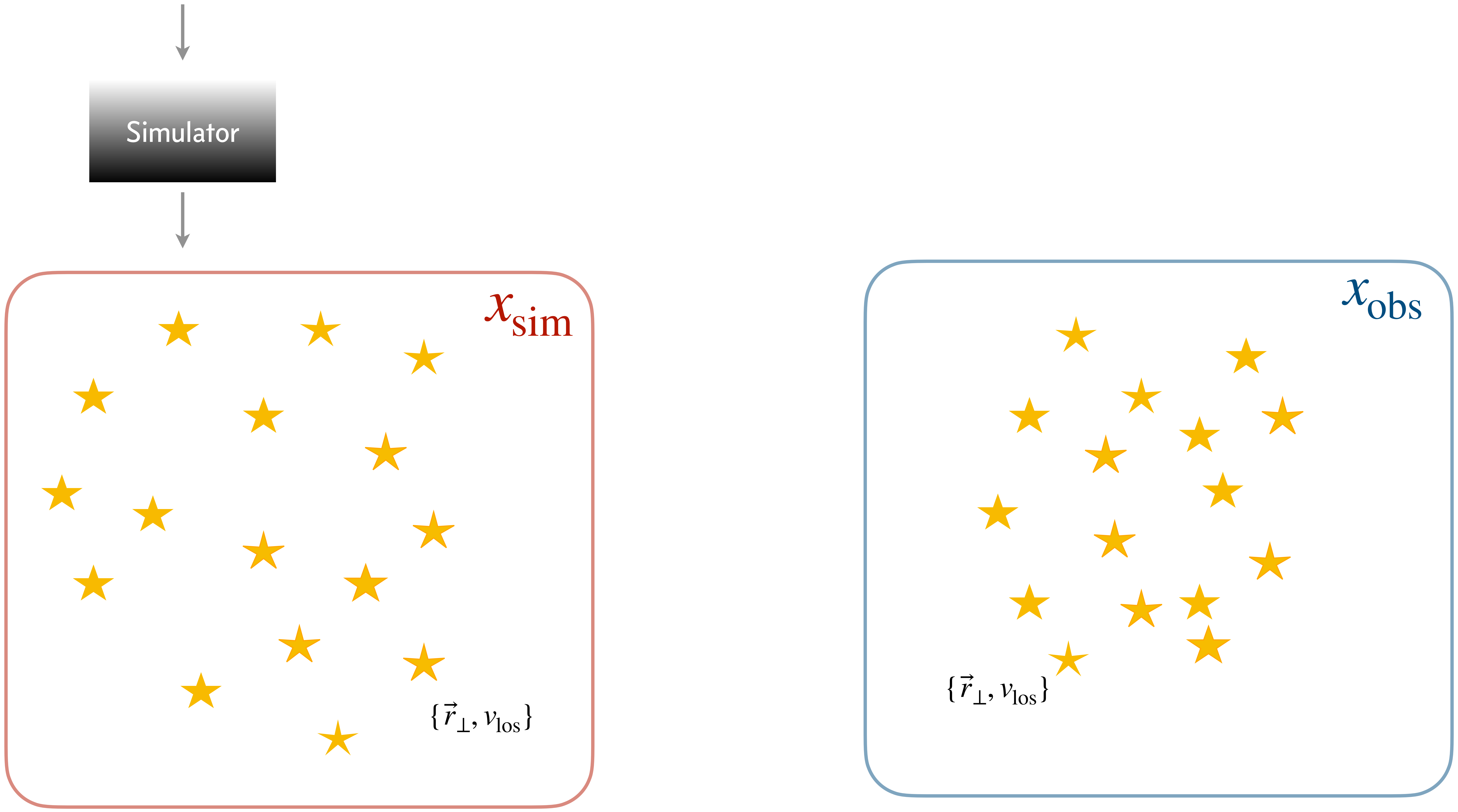
Nguyen, SM et al [PRD 2023]



Simulation-based inference for dwarf galaxies

Nguyen, SM et al [PRD 2023]

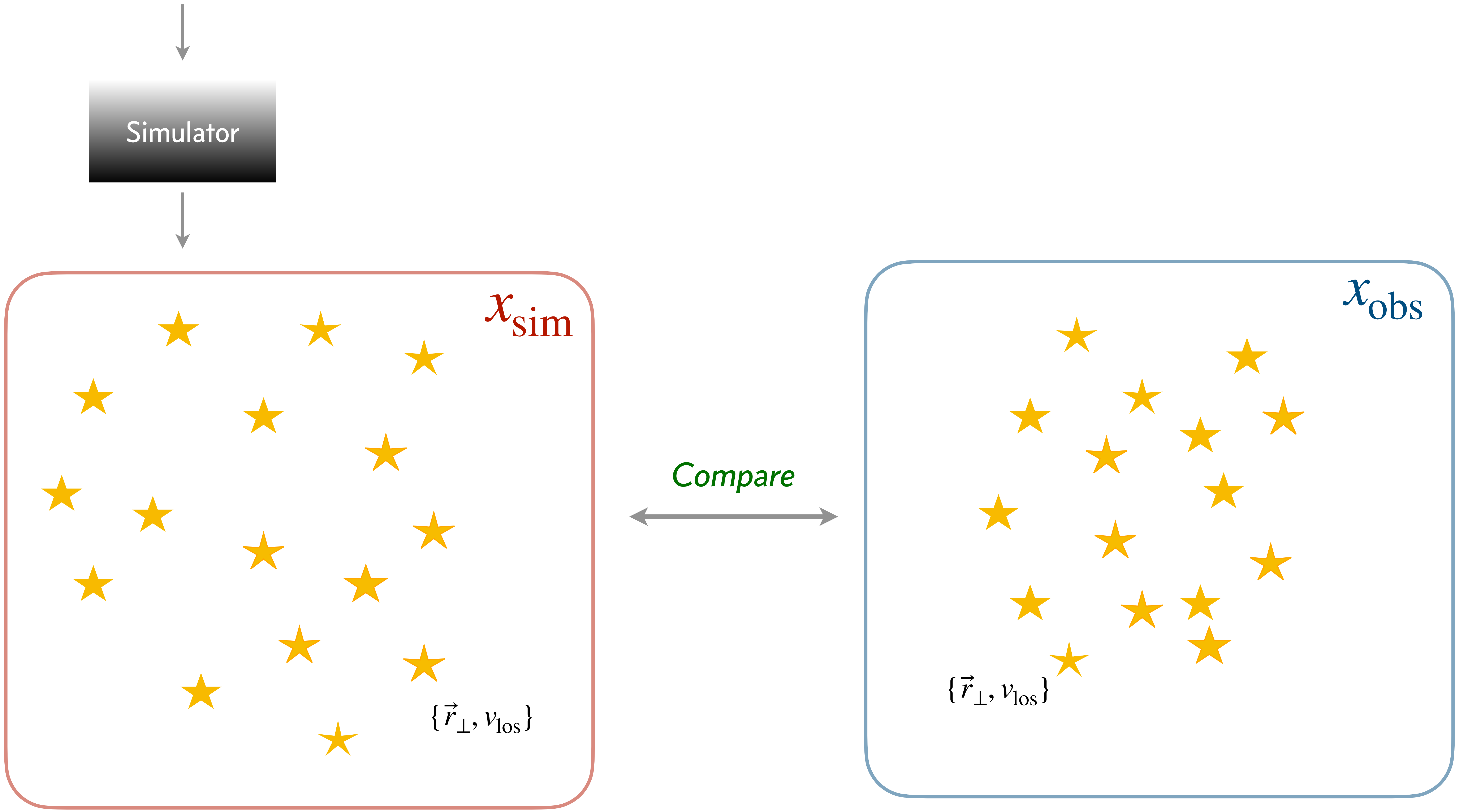
DM + stellar parameters θ



Simulation-based inference for dwarf galaxies

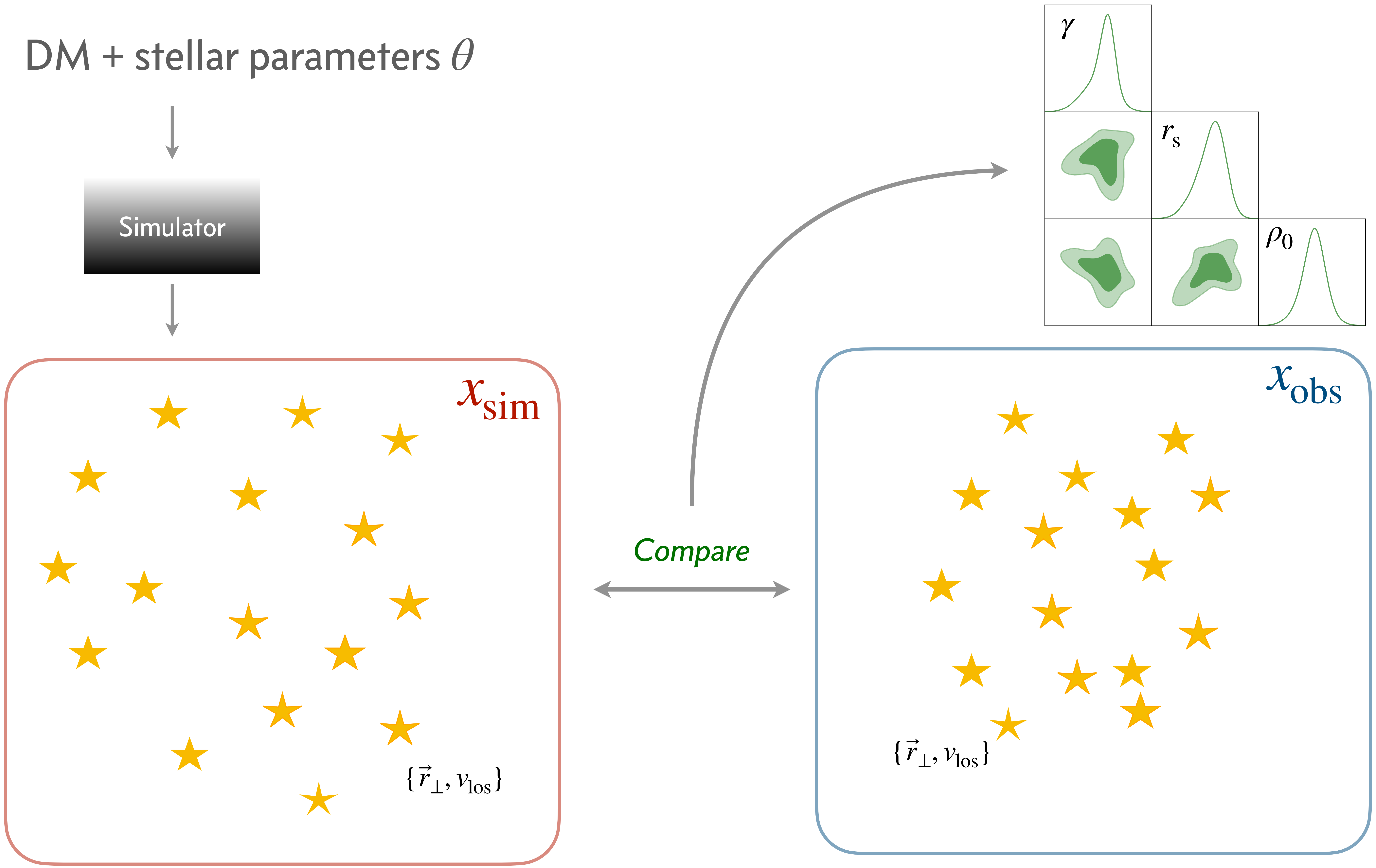
Nguyen, SM et al [PRD 2023]

DM + stellar parameters θ



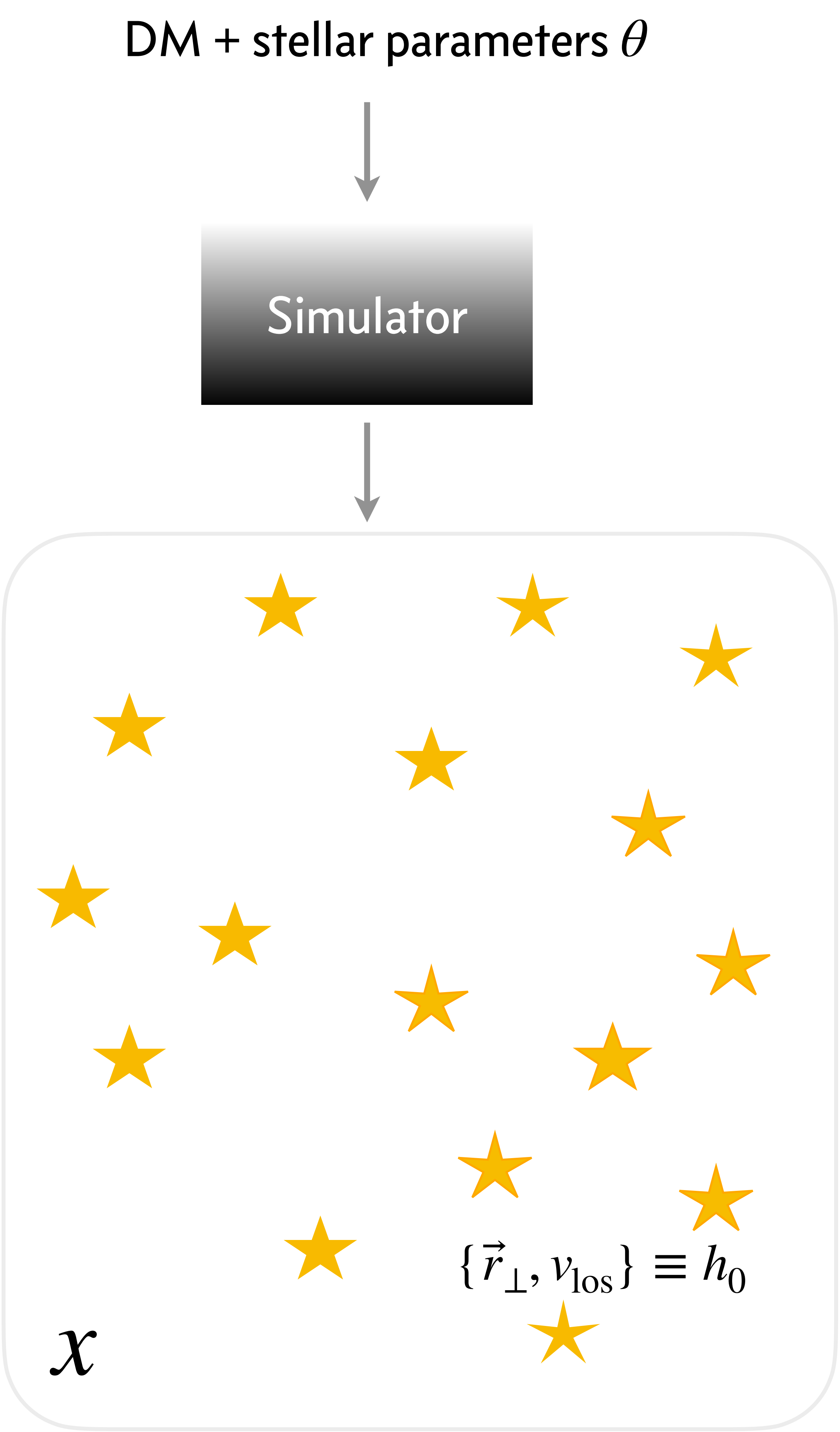
Simulation-based inference for dwarf galaxies

Nguyen, SM et al [PRD 2023]



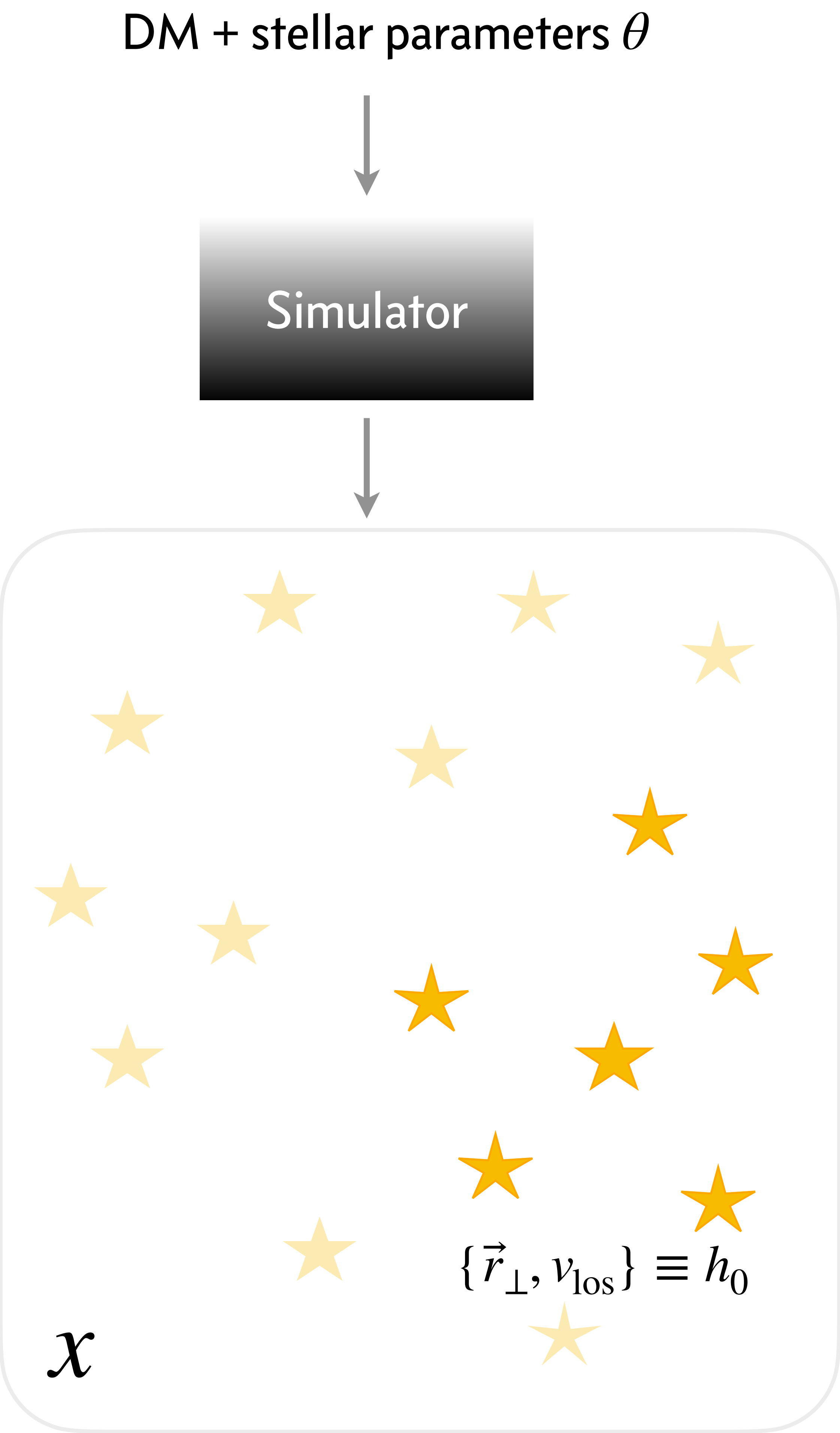
Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]



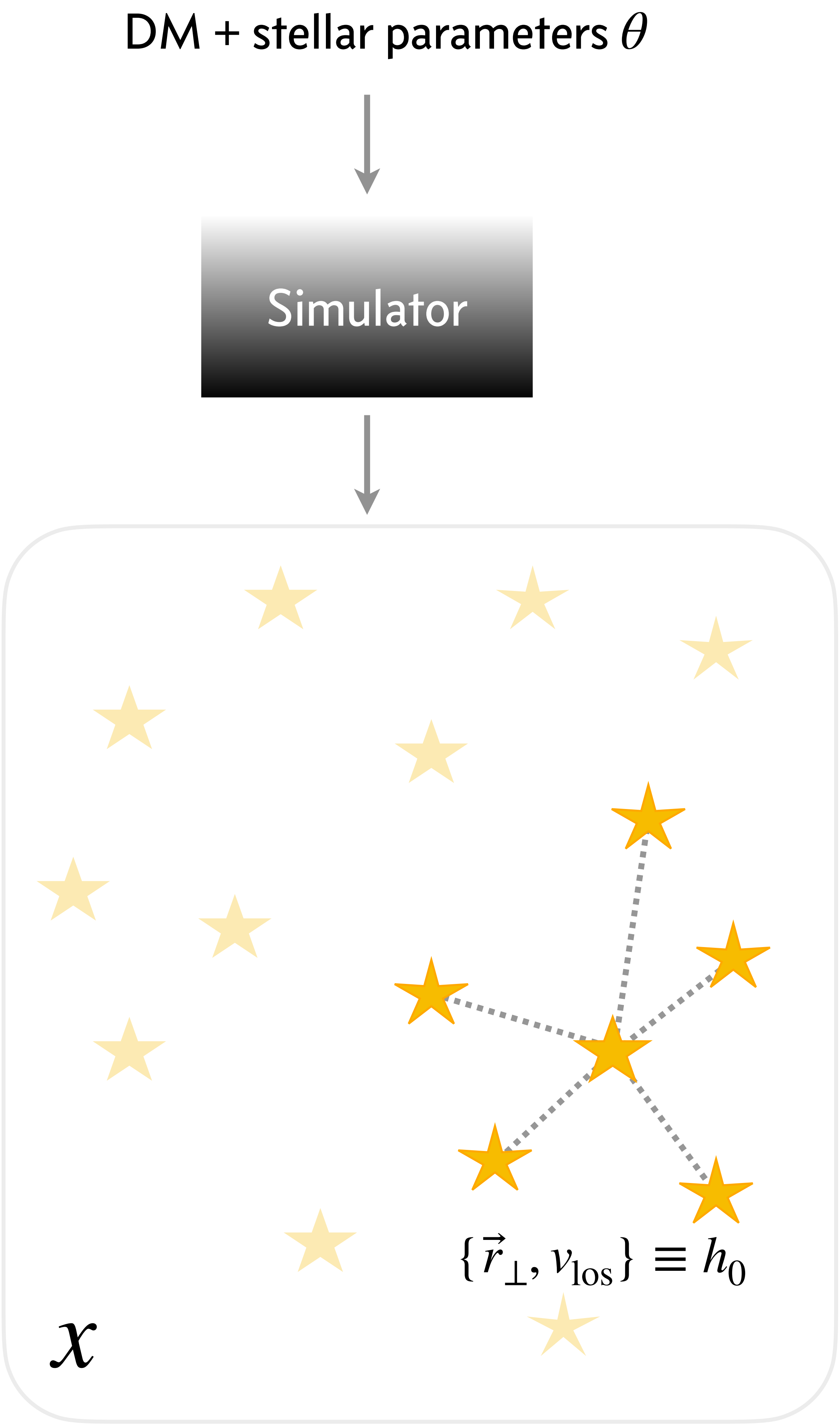
Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]



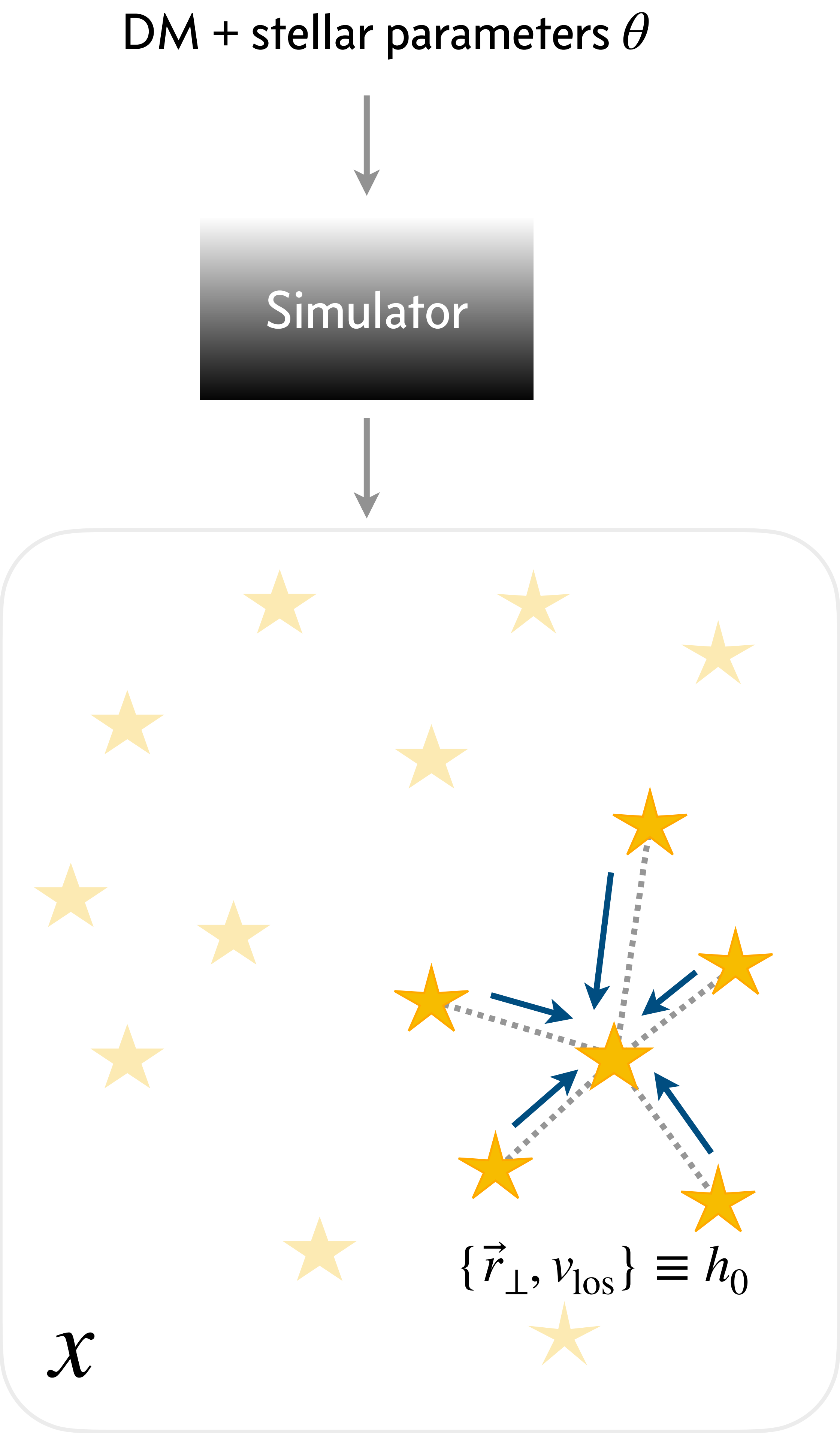
Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]



Graph neural networks for stellar kinematics

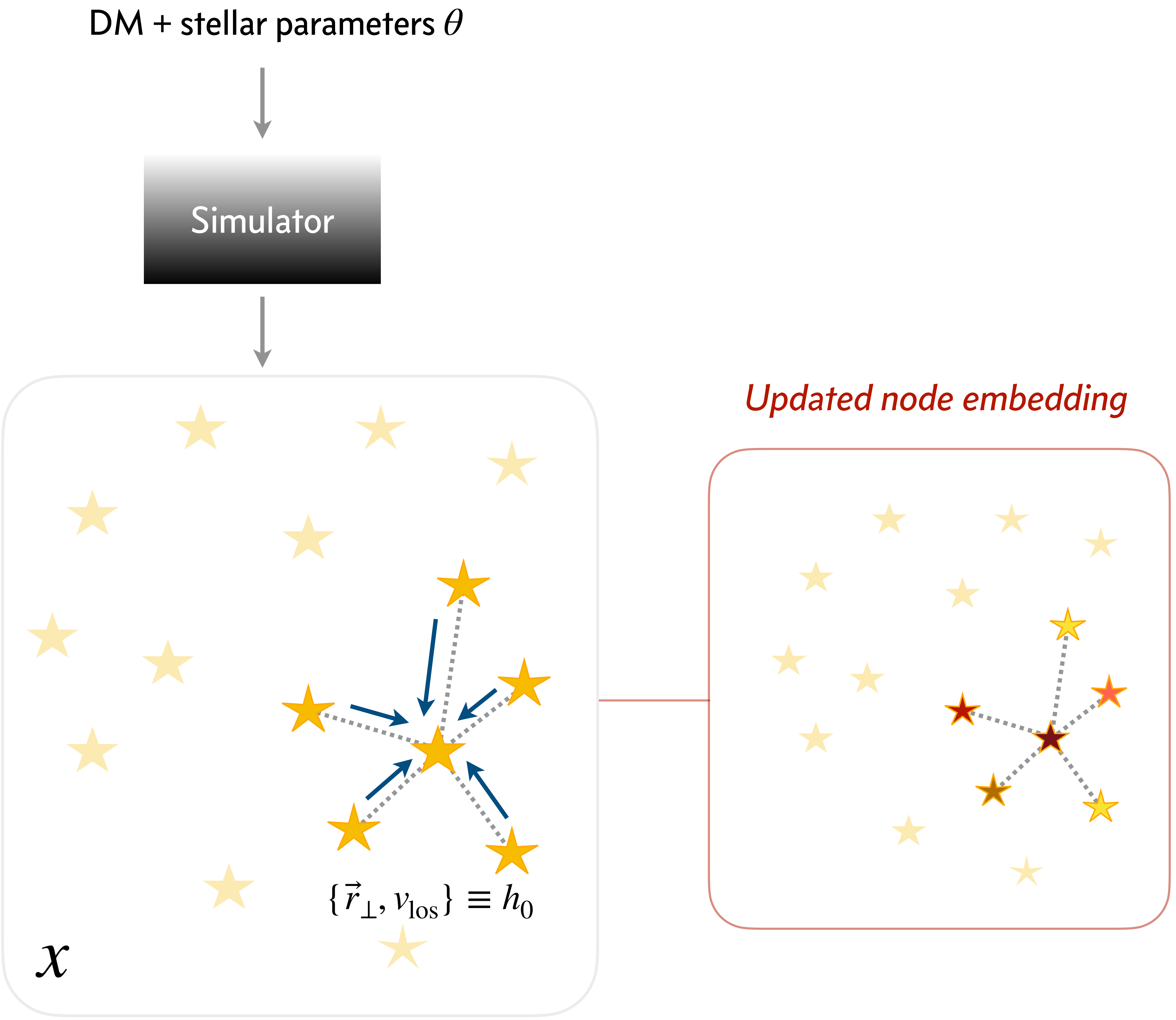
Nguyen, SM et al [PRD 2023]



Node update rule $h_{i+1} = \sigma\left(W^T \sum_{j \in \mathcal{N}(i)} h_j\right)$

Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]

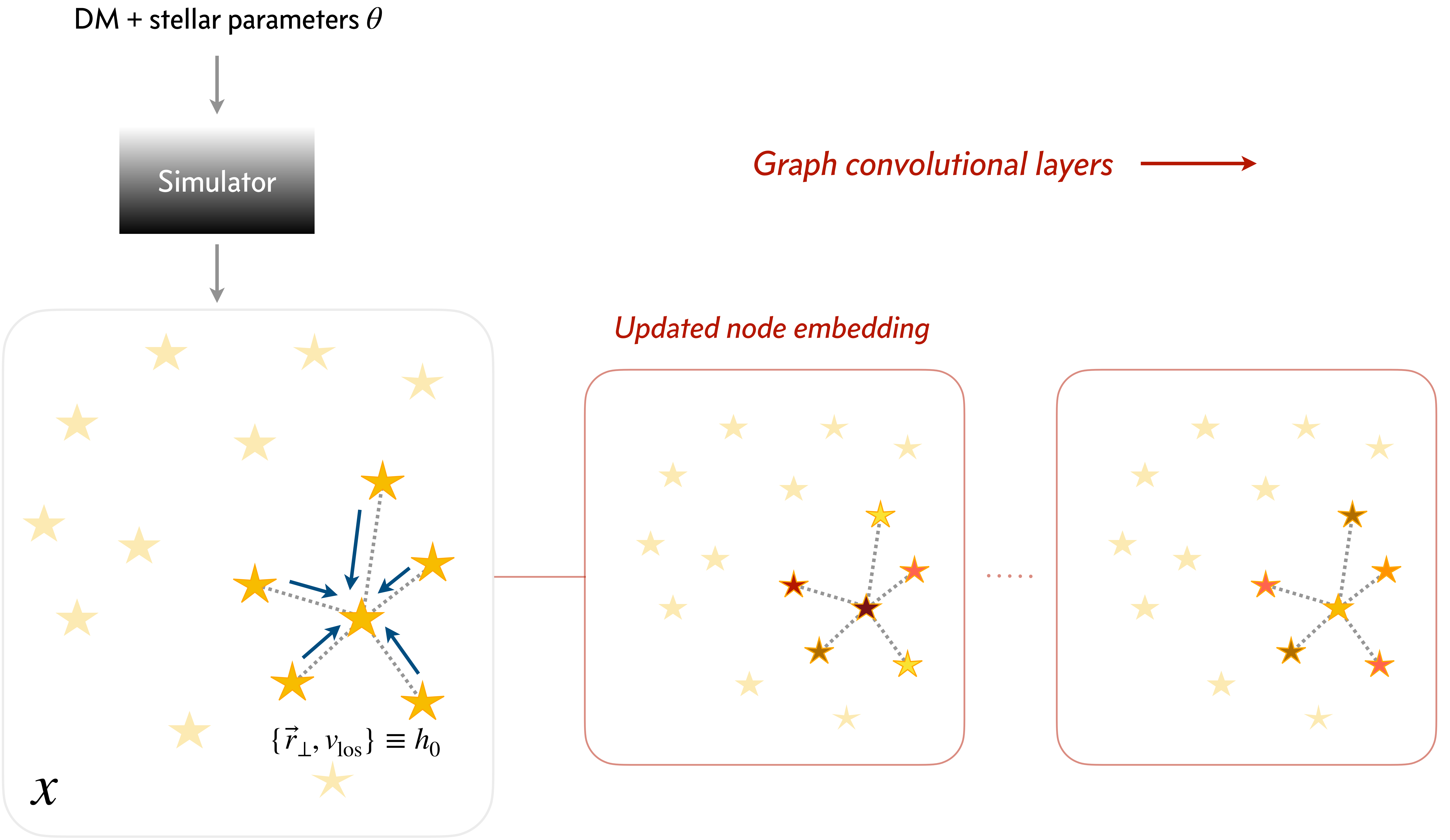


Node update rule

$$h_{i+1} = \sigma\left(W^T \sum_{j \in \mathcal{N}(i)} h_j\right)$$

Graph neural networks for stellar kinematics

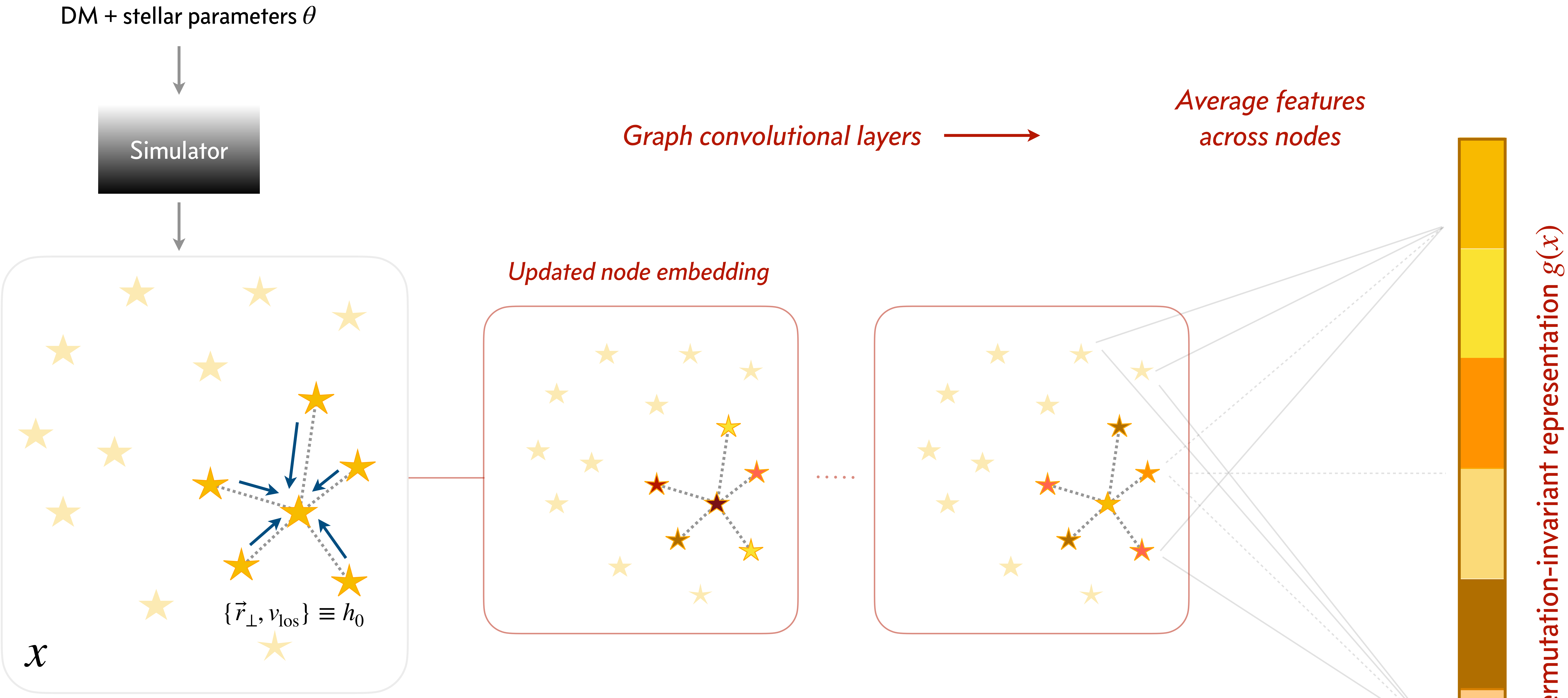
Nguyen, SM et al [PRD 2023]



Node update rule $h_{i+1} = \sigma\left(W^T \sum_{j \in \mathcal{N}(i)} h_j\right)$

Graph neural networks for stellar kinematics

Nguyen, SM et al [PRD 2023]

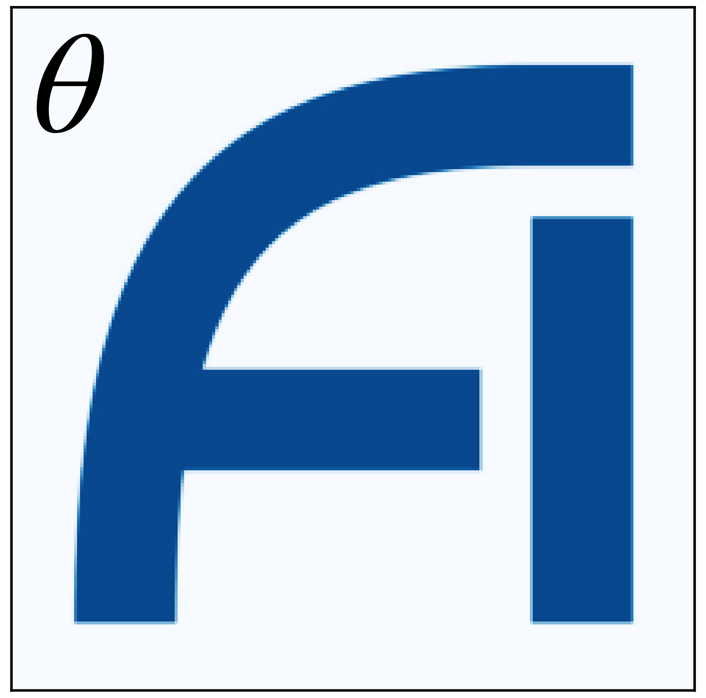


Node update rule $h_{i+1} = \sigma\left(W^T \sum_{j \in \mathcal{N}(i)} h_j\right)$

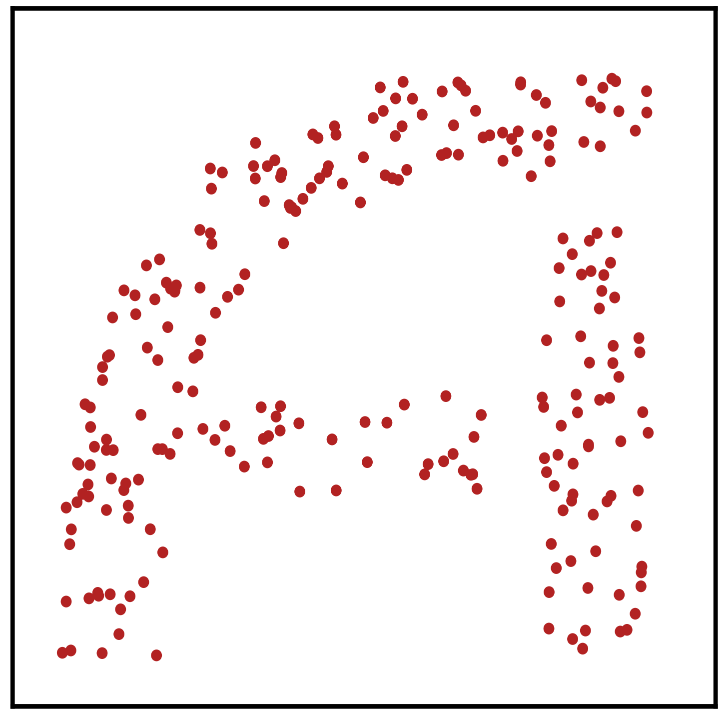
Normalizing flows for density estimation and sampling

Goal: *model* $p(\theta)$

Target density



Samples

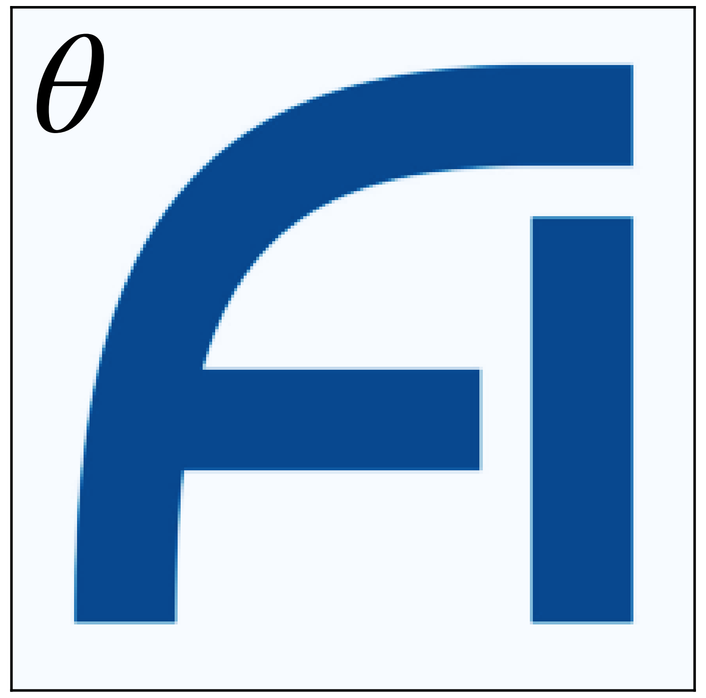


Rezende & Mohamed [ICML 2015]

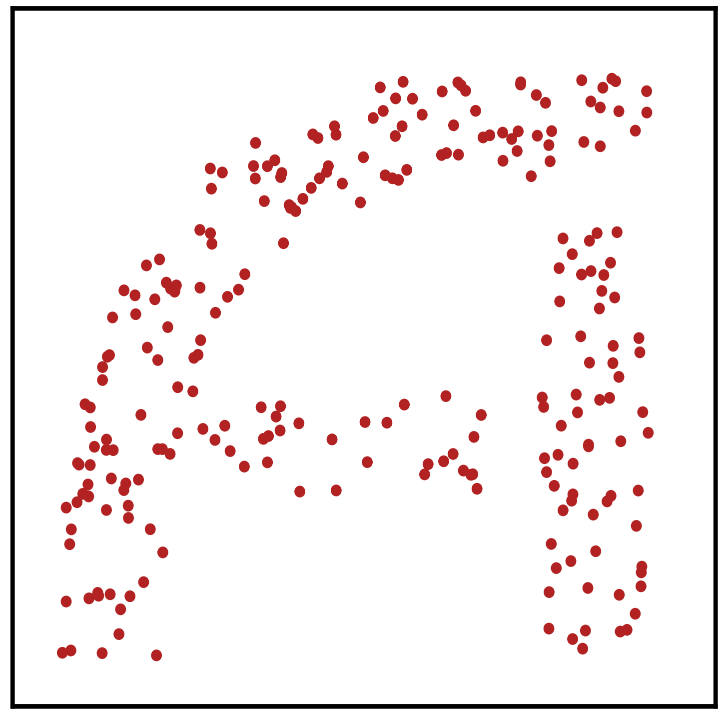
Normalizing flows for density estimation and sampling

Goal: *model* $p(\theta)$

Target density



Samples

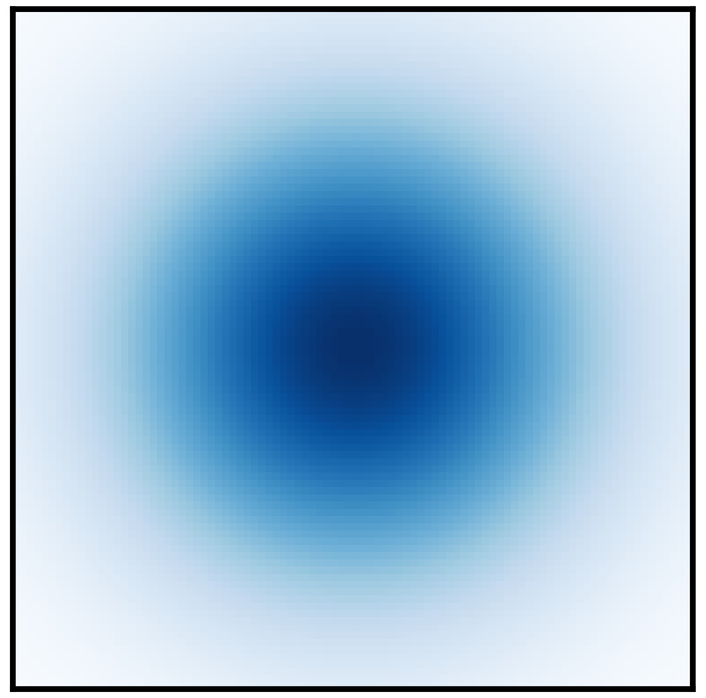


Normalizing flows for density estimation and sampling

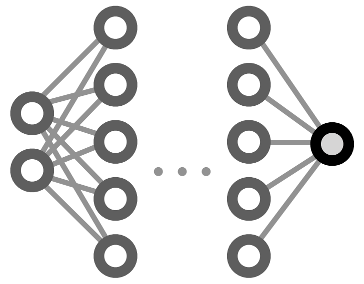
Rezende & Mohamed [ICML 2015]

Goal: ***model*** $p(\theta)$

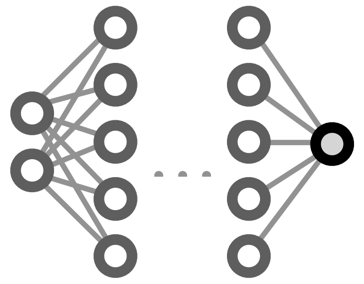
Base density



$\mathcal{N}(u)$



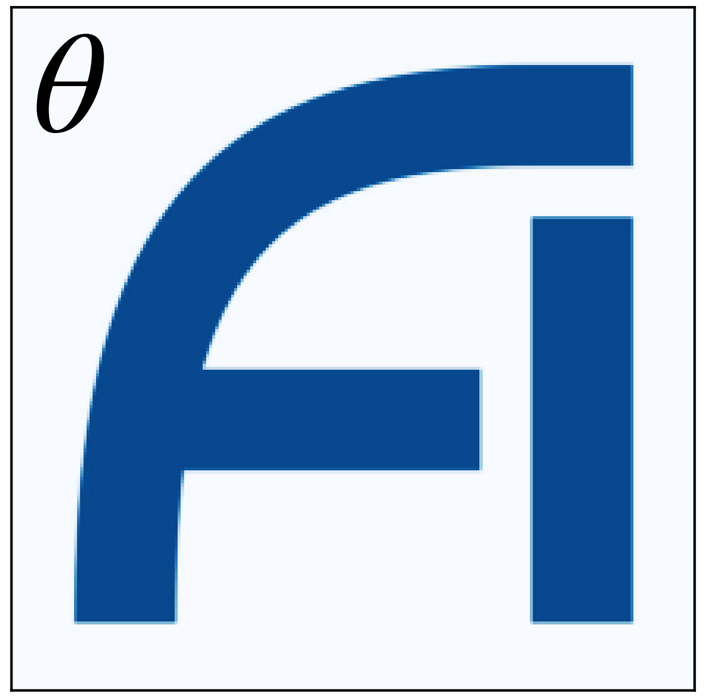
$\theta = f(u)$



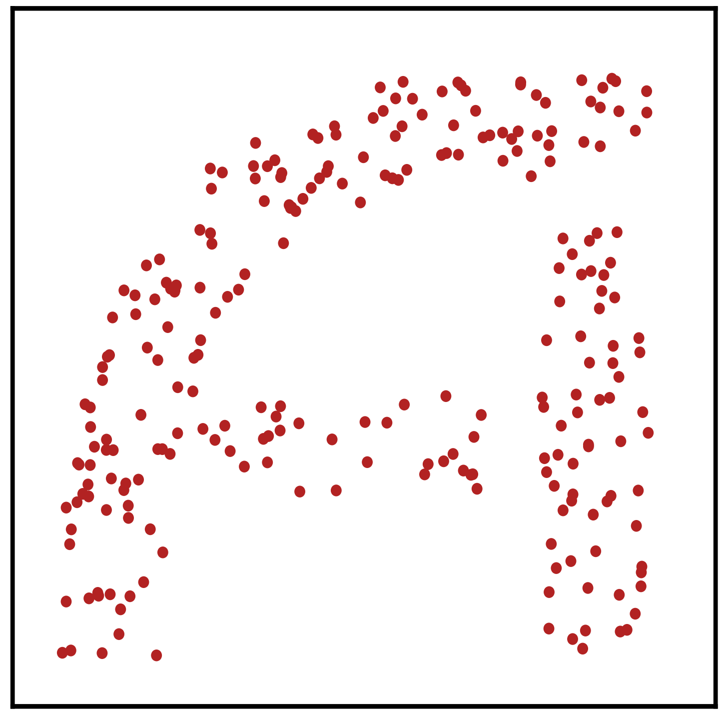
One-to-one transformation

Tractable f^{-1} *and* $\det \nabla f$

Target density



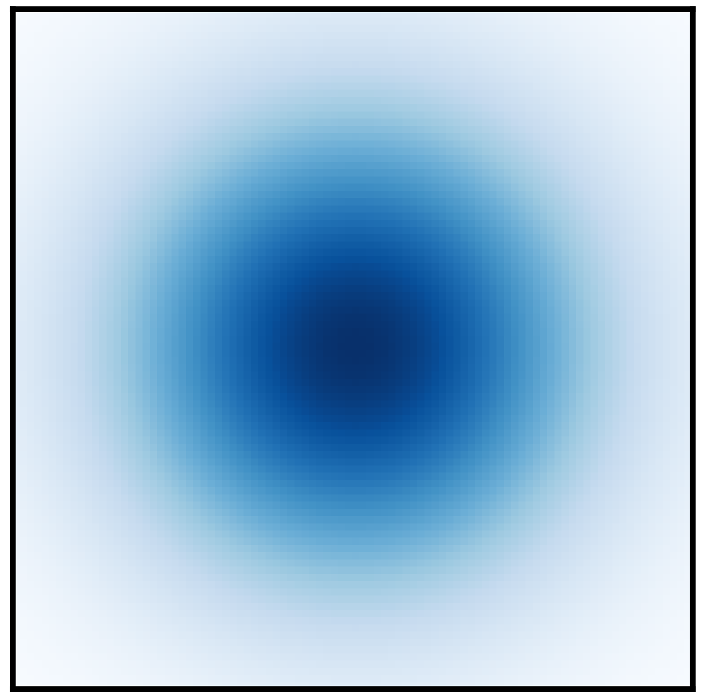
Samples



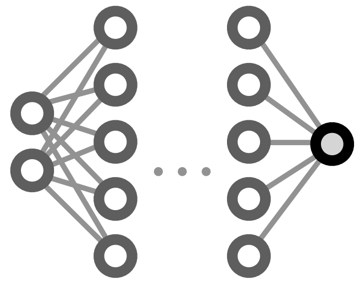
Normalizing flows for density estimation and sampling

Goal: ***model*** $p(\theta)$

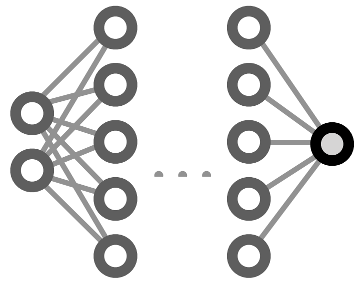
Base density



$\mathcal{N}(u)$



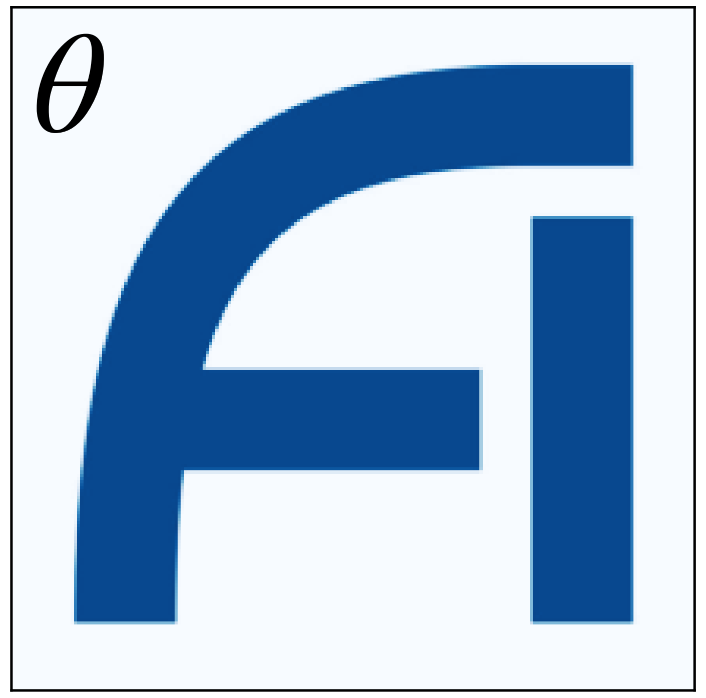
$\theta = f(u)$



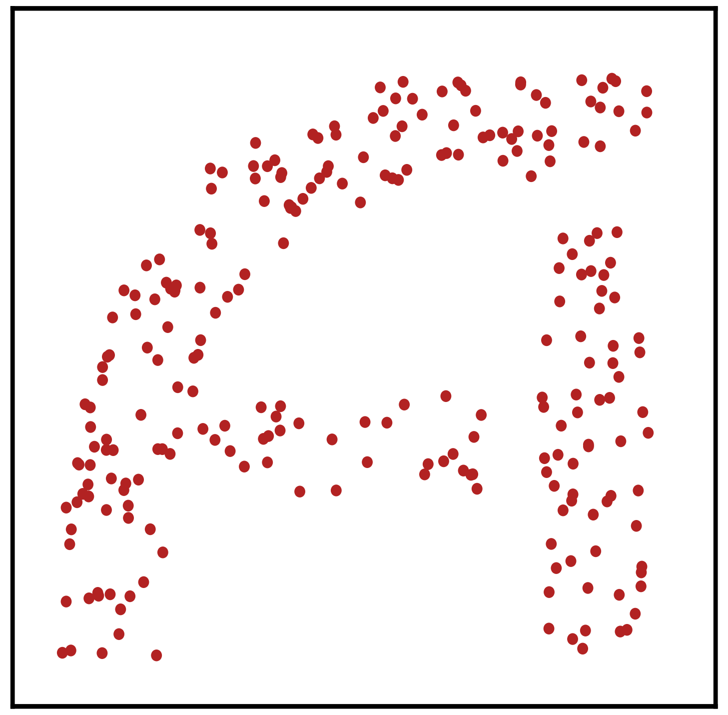
One-to-one transformation

Tractable f^{-1} *and* $\det \nabla f$

Target density



Samples



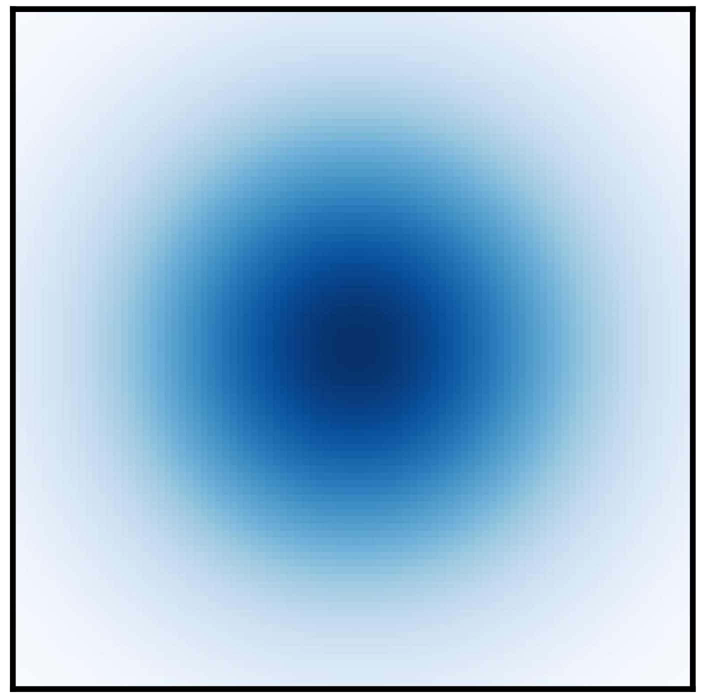
$$p(\theta) = \mathcal{N}\left(f^{-1}(\theta)\right) |\det \nabla f|^{-1}$$

Normalizing flows for density estimation and sampling

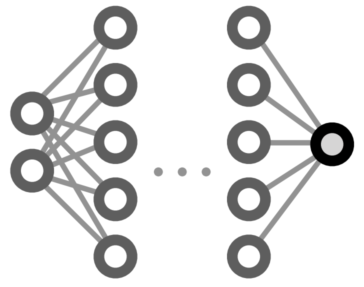
Rezende & Mohamed [ICML 2015]

Goal: *model* $p(\theta)$

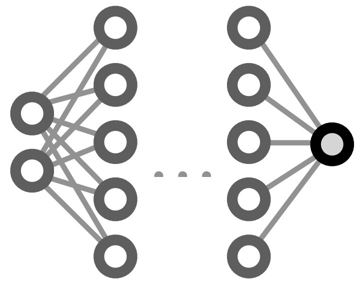
Base density



$\mathcal{N}(u)$



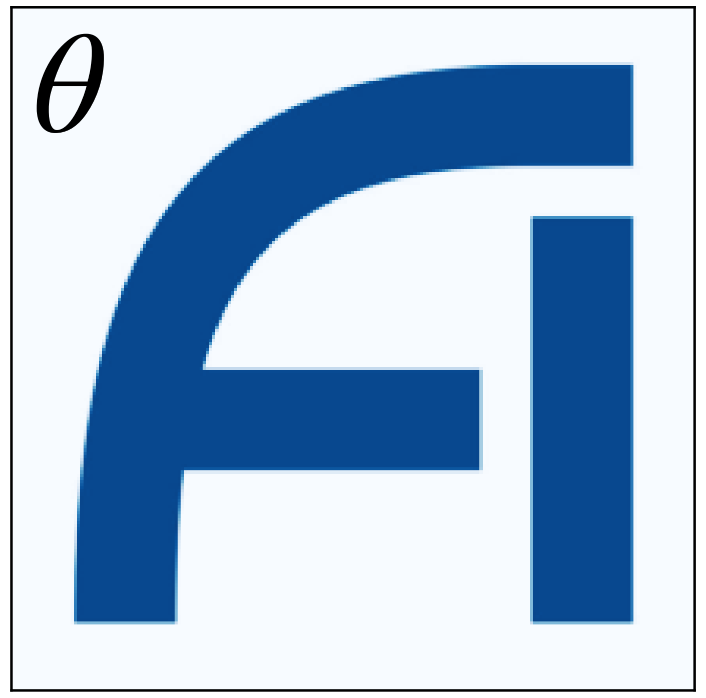
$\theta = f(u)$



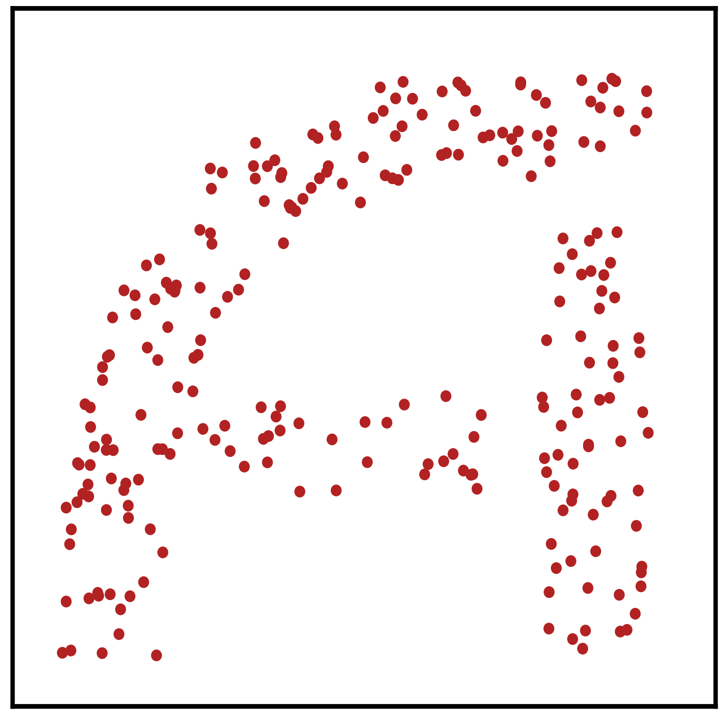
One-to-one transformation

Tractable f^{-1} *and* $\det \nabla f$

Target density

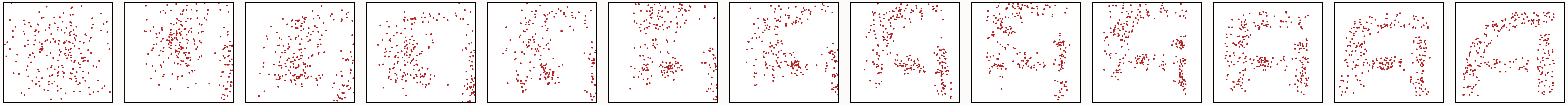


Samples



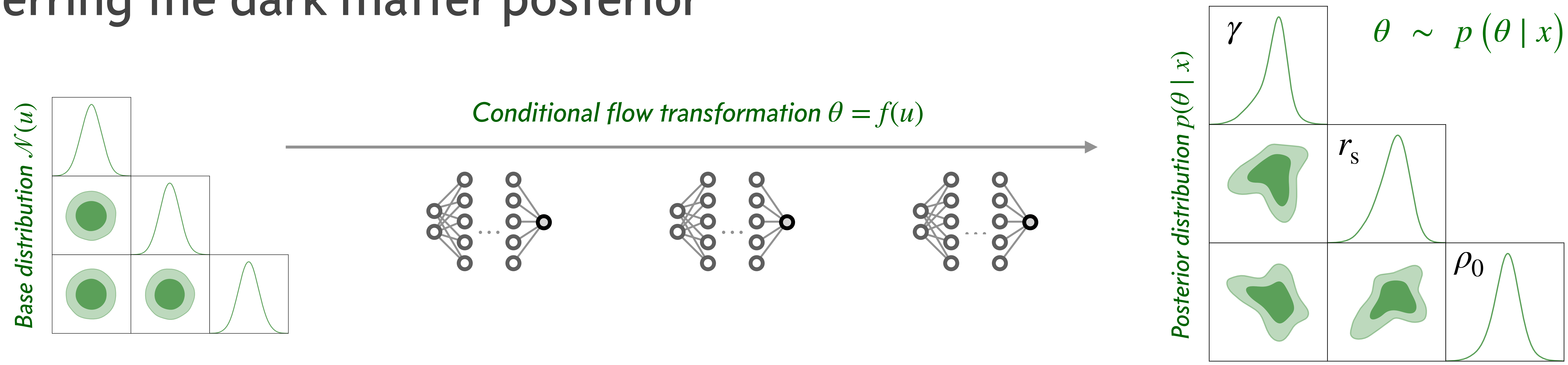
$$p(\theta) = \mathcal{N}\left(f^{-1}(\theta)\right) |\det \nabla f|^{-1}$$

Efficient sampling and density estimation



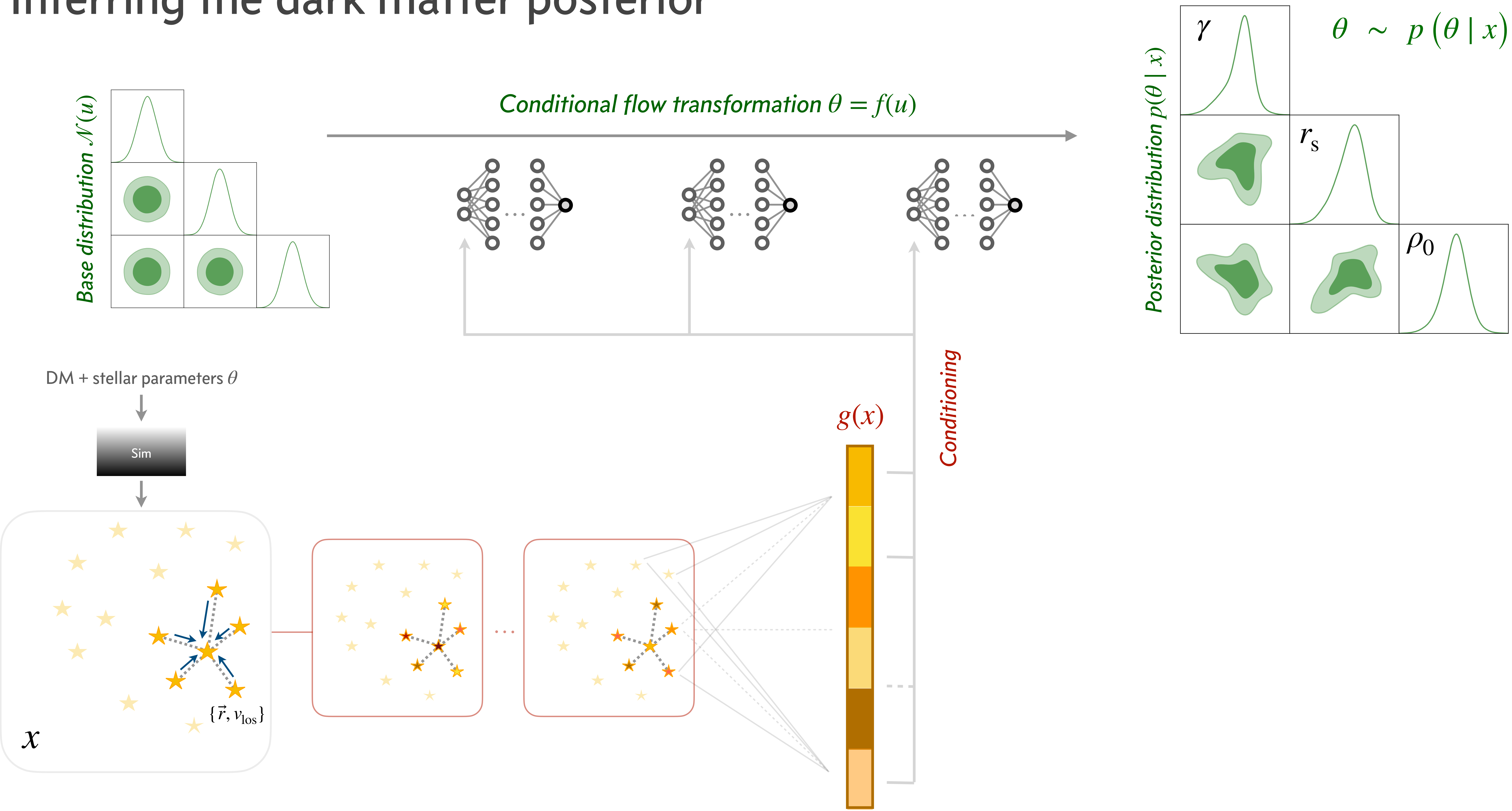
Inferring the dark matter posterior

Nguyen, SM et al [PRD 2023]



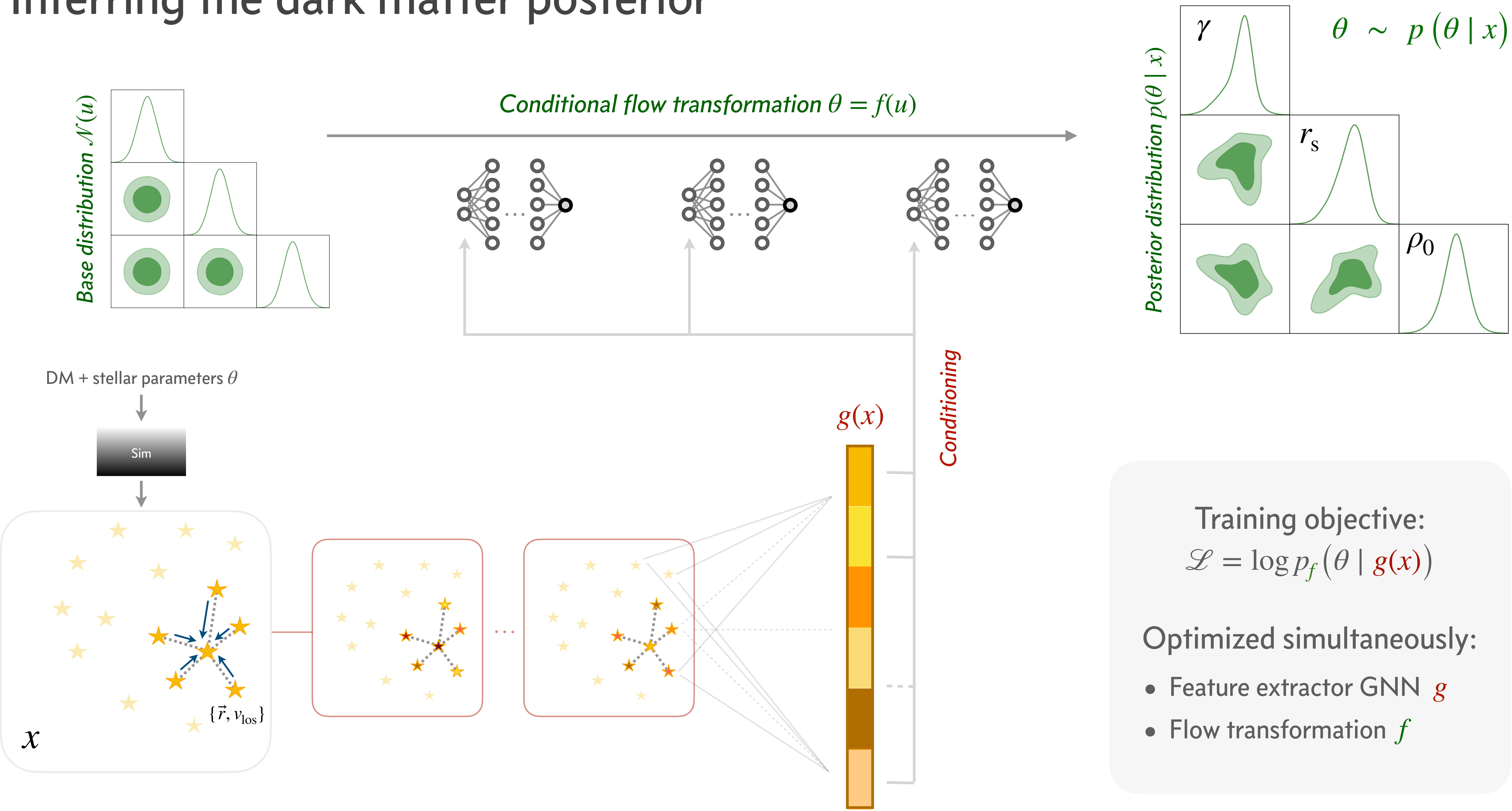
Inferring the dark matter posterior

Nguyen, SM et al [PRD 2023]



Inferring the dark matter posterior

Nguyen, SM et al [PRD 2023]



Training objective:

$\mathscr{L} = \log p_f\left(\theta \mid g(x)\right)$

Optimized simultaneously:

- Feature extractor GNN g
- Flow transformation f

Applications to hydrodynamic simulations

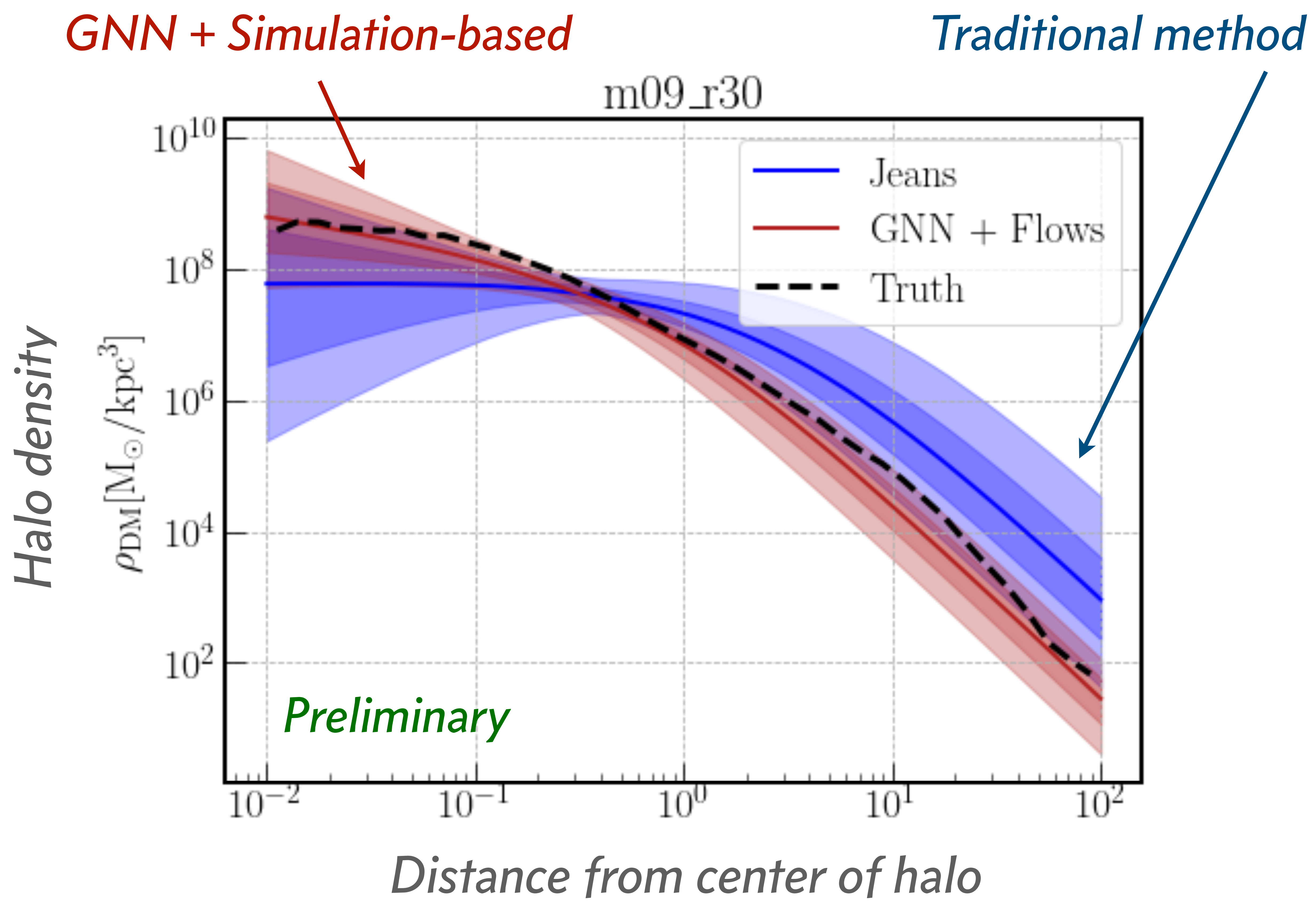
Nguyen, SM et al [In prep]

Wheeler et al [MNRAS 2019]

Applications to hydrodynamic simulations

Nguyen, SM et al [In prep]

Wheeler et al [MNRAS 2019]



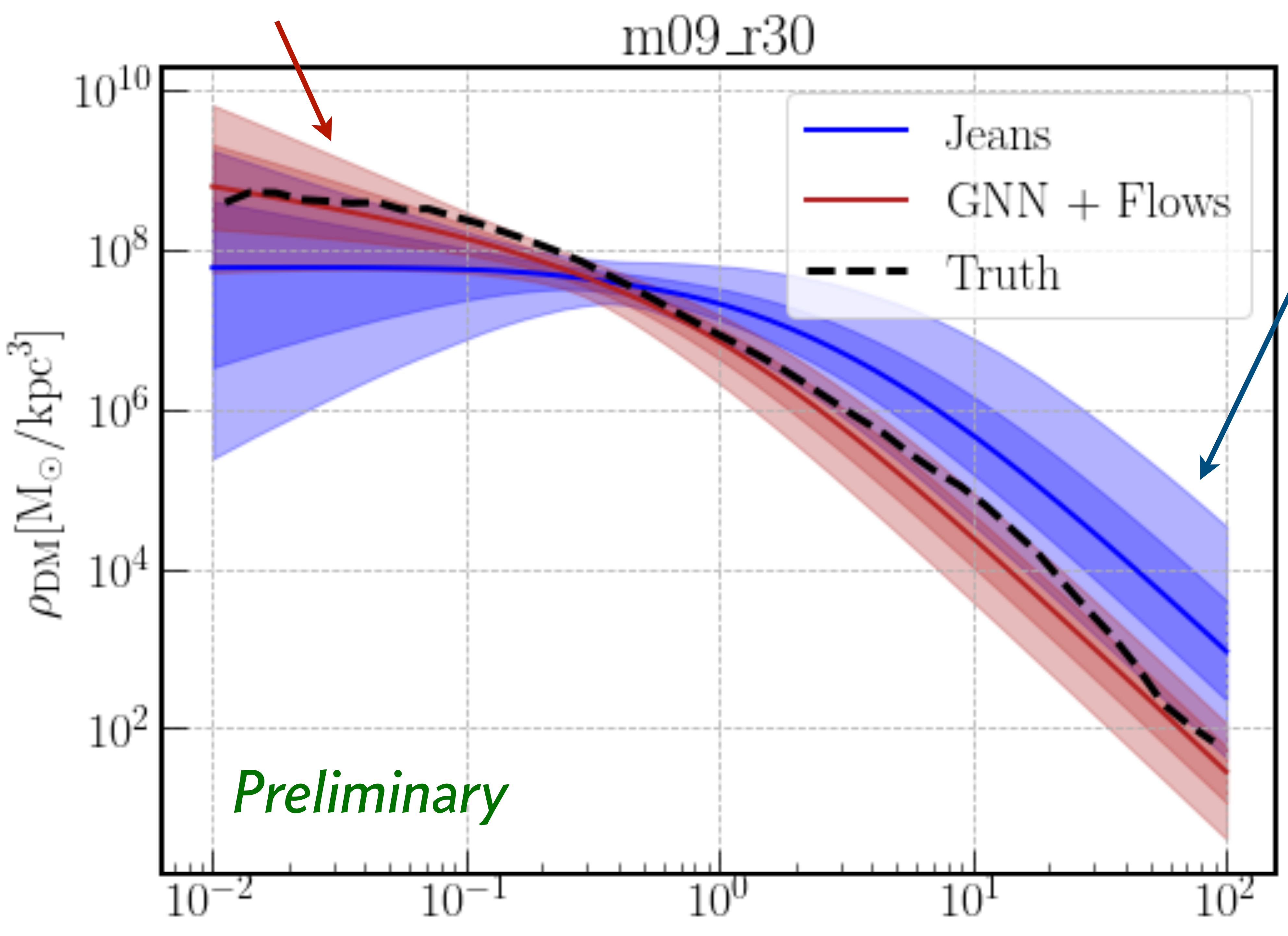
Applications to hydrodynamic simulations

Nguyen, SM et al [In prep]

Wheeler et al [MNRAS 2019]

GNN + Simulation-based

Traditional method

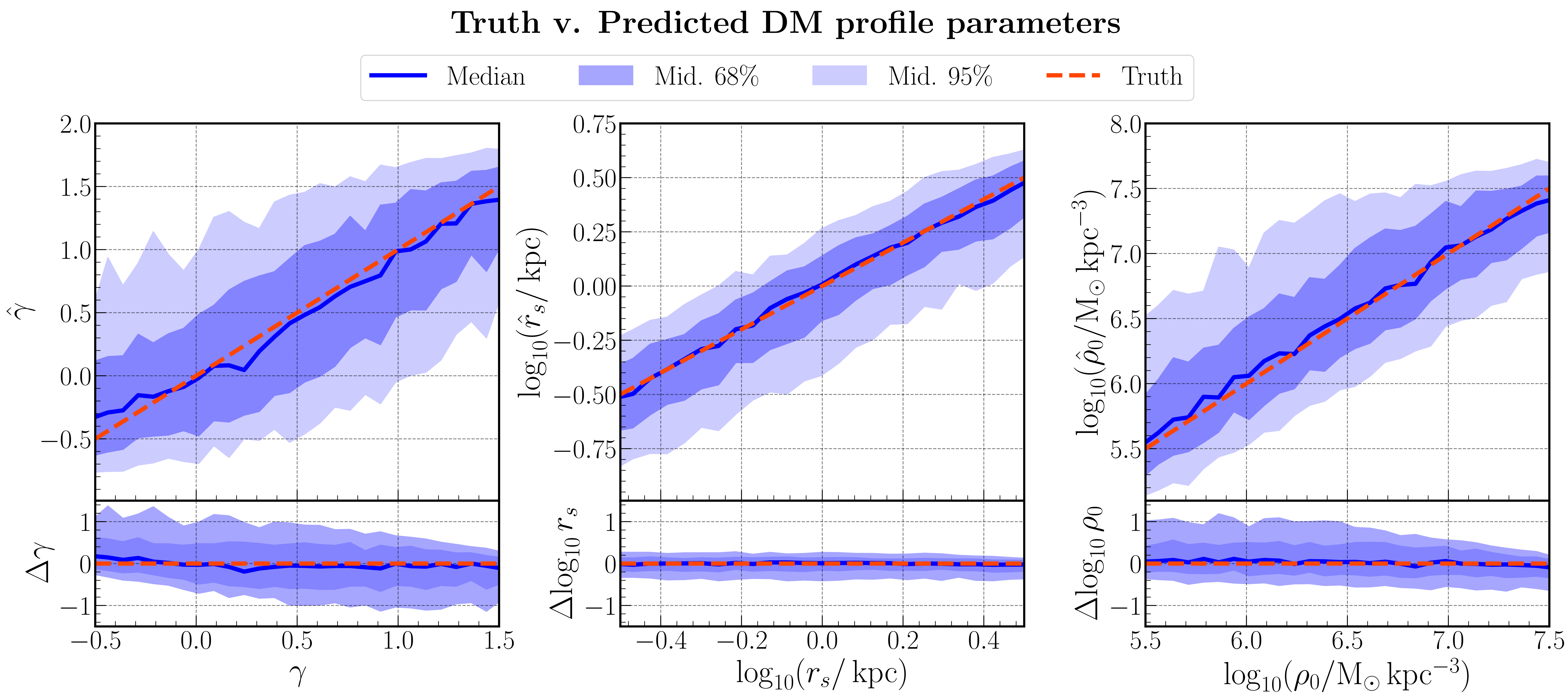


Halo density

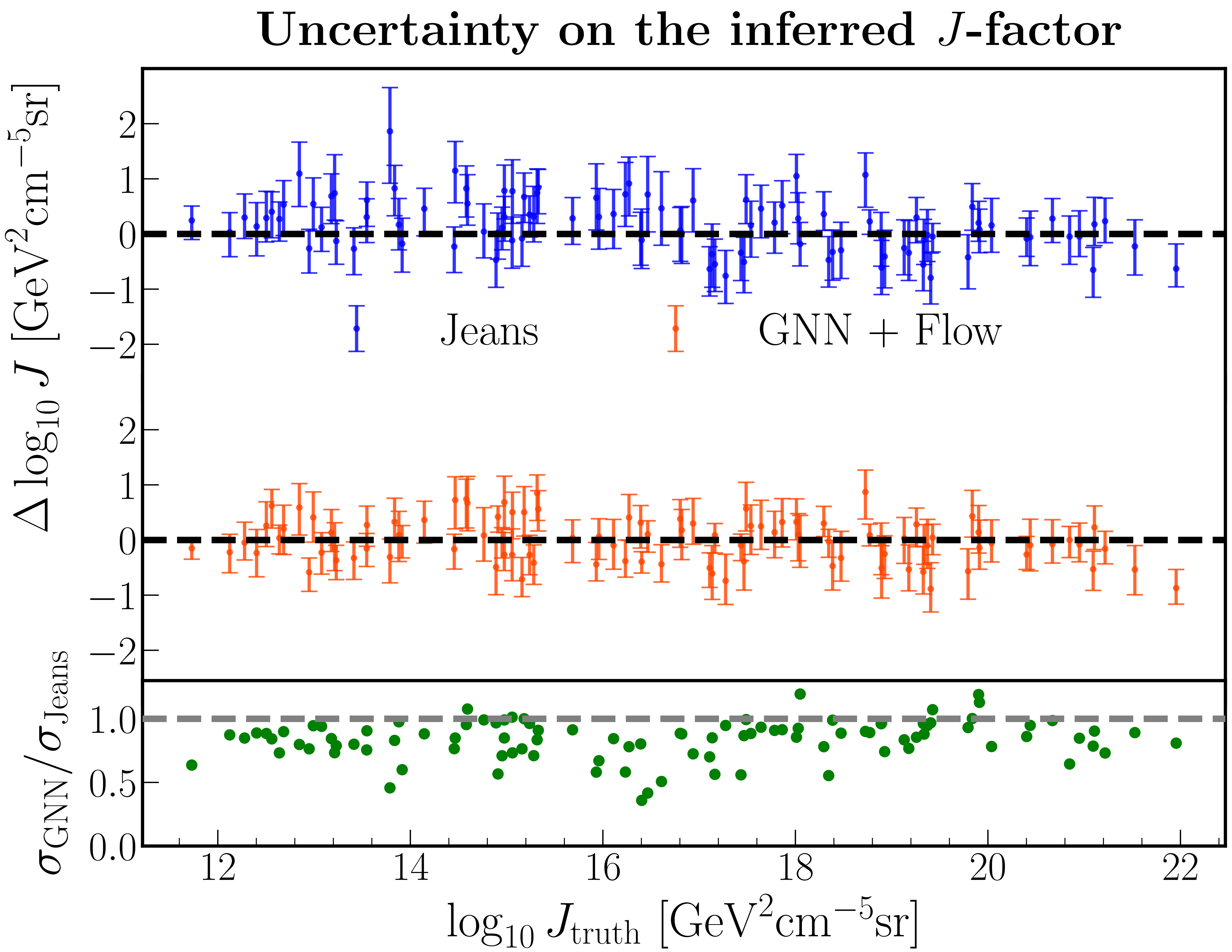
Distance from center of halo

Significantly better performance!

True vs predicted DM parameters

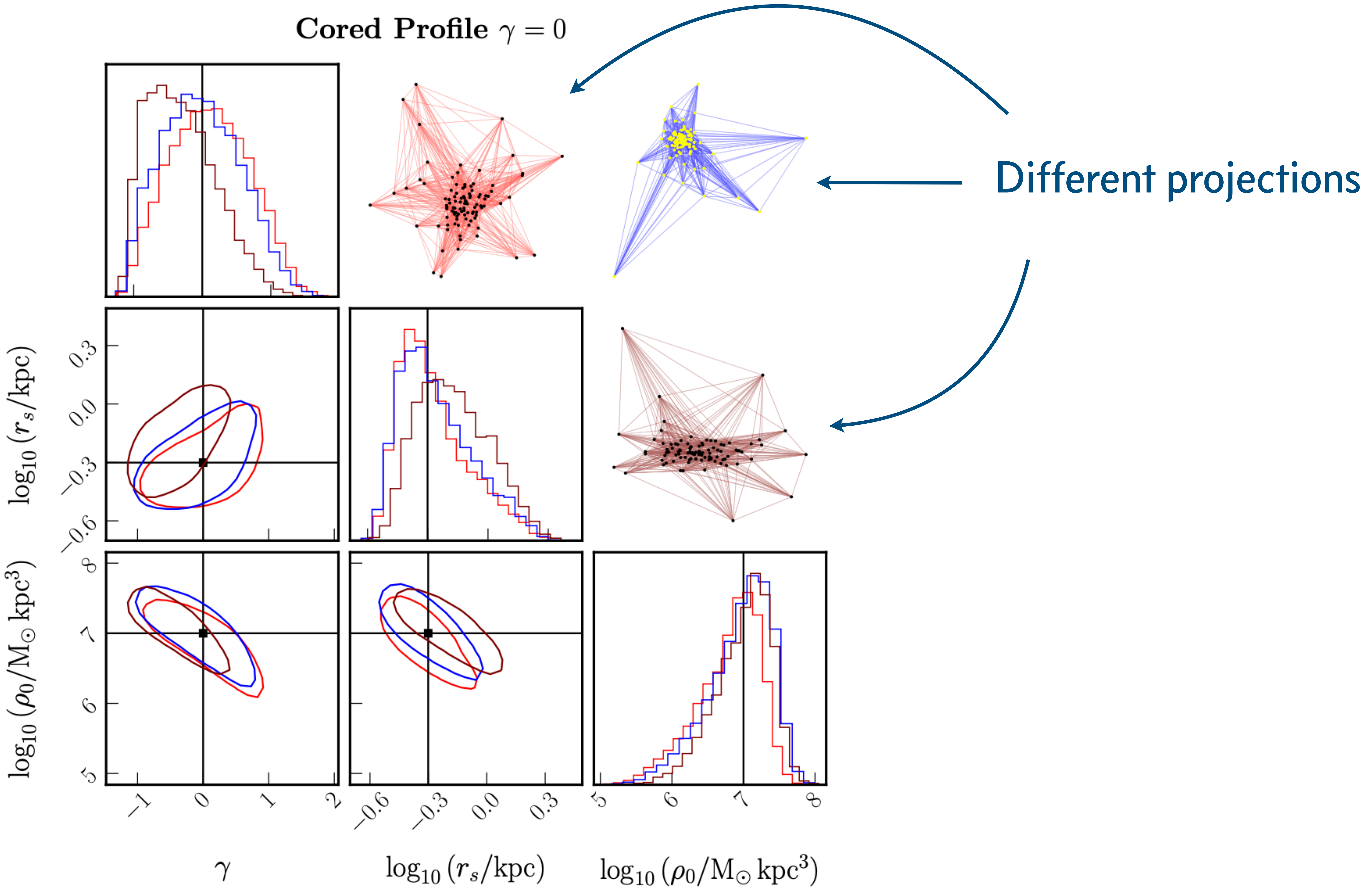


J-factors

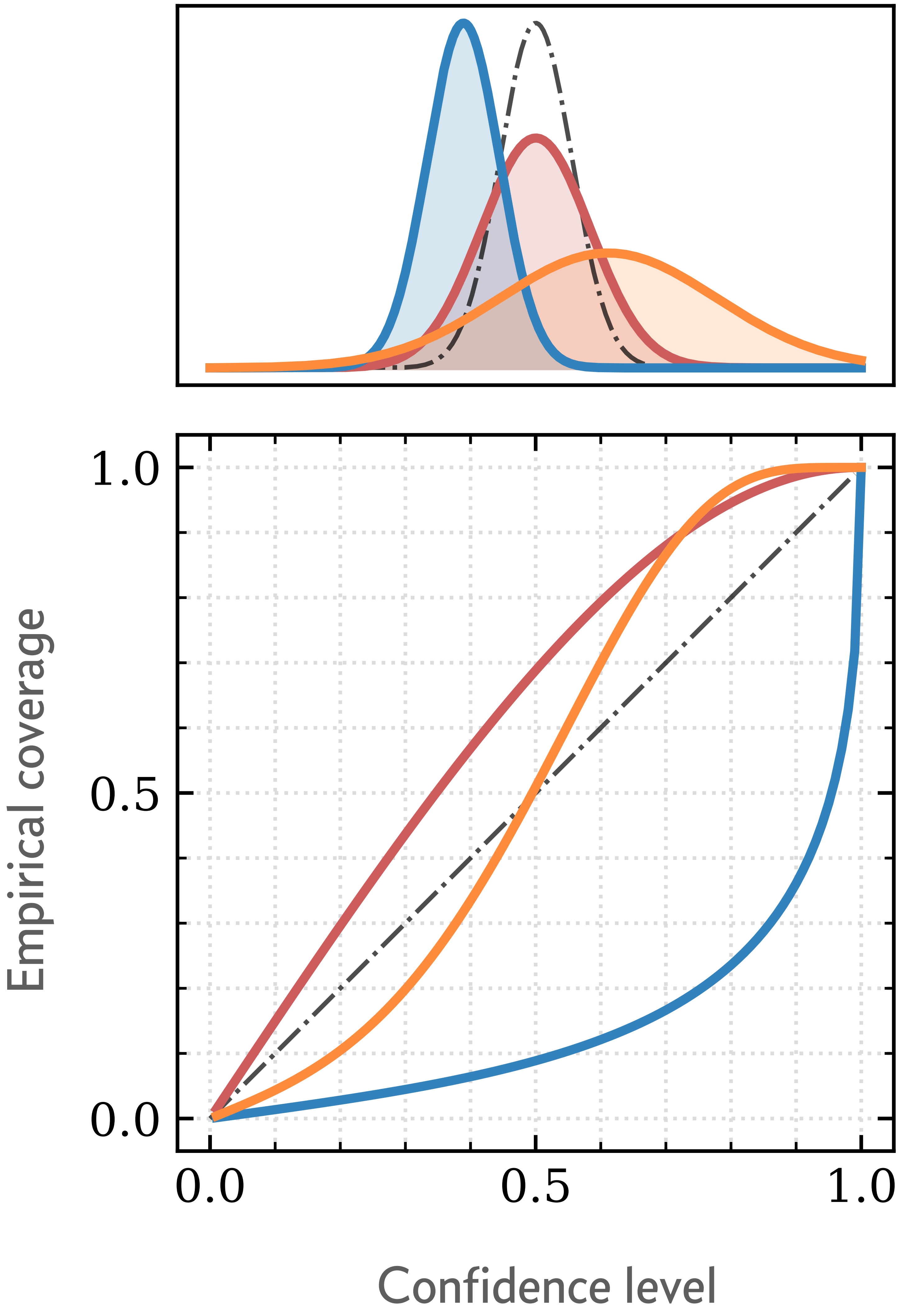


Uncertainty on the inferred J-factor
Jeans
GNN + Flow
$\Delta \log_{10} J$ [GeV^{2}cm^{-5}sr]
$\sigma_{\mathrm{GNN}}/\sigma_{\mathrm{Jeans}}$
$\log_{10} J_{\mathrm{truth}}$ [GeV^{2}cm^{-5}sr]
2
1
0
−1
−2
1.0
0.5
0.0
12
14
16
18
20
22

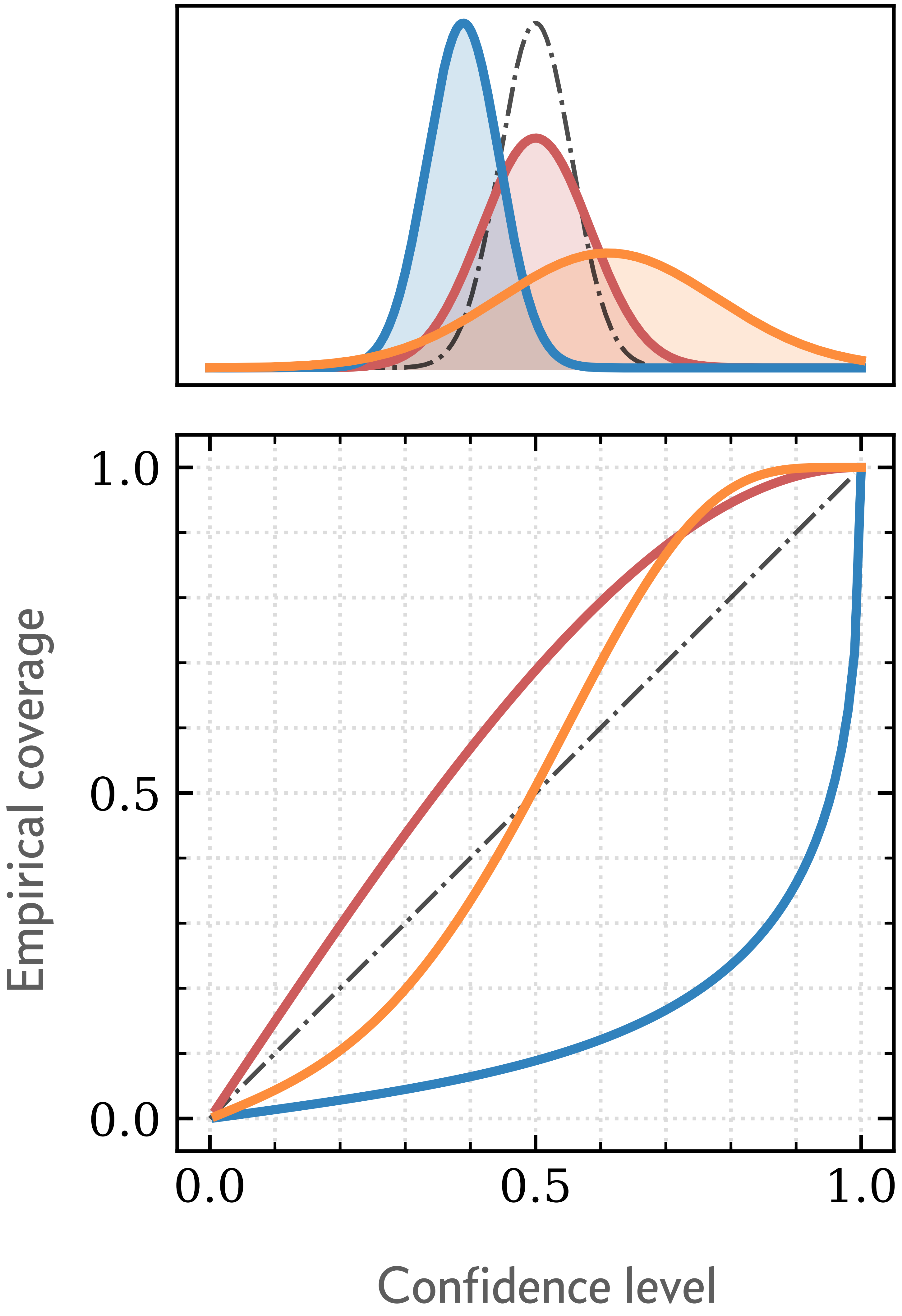
Sensitivity to projection

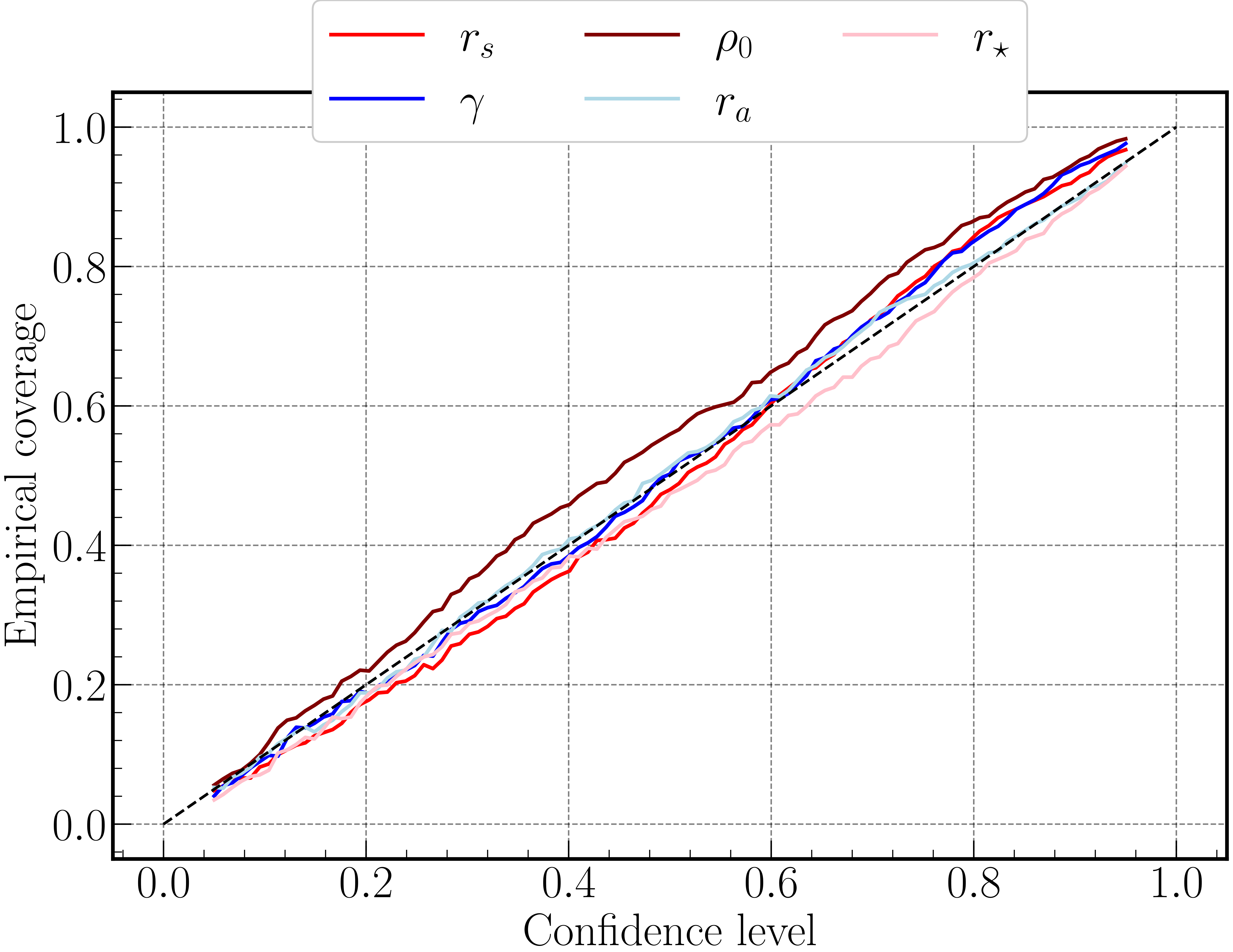


Statistical coverage

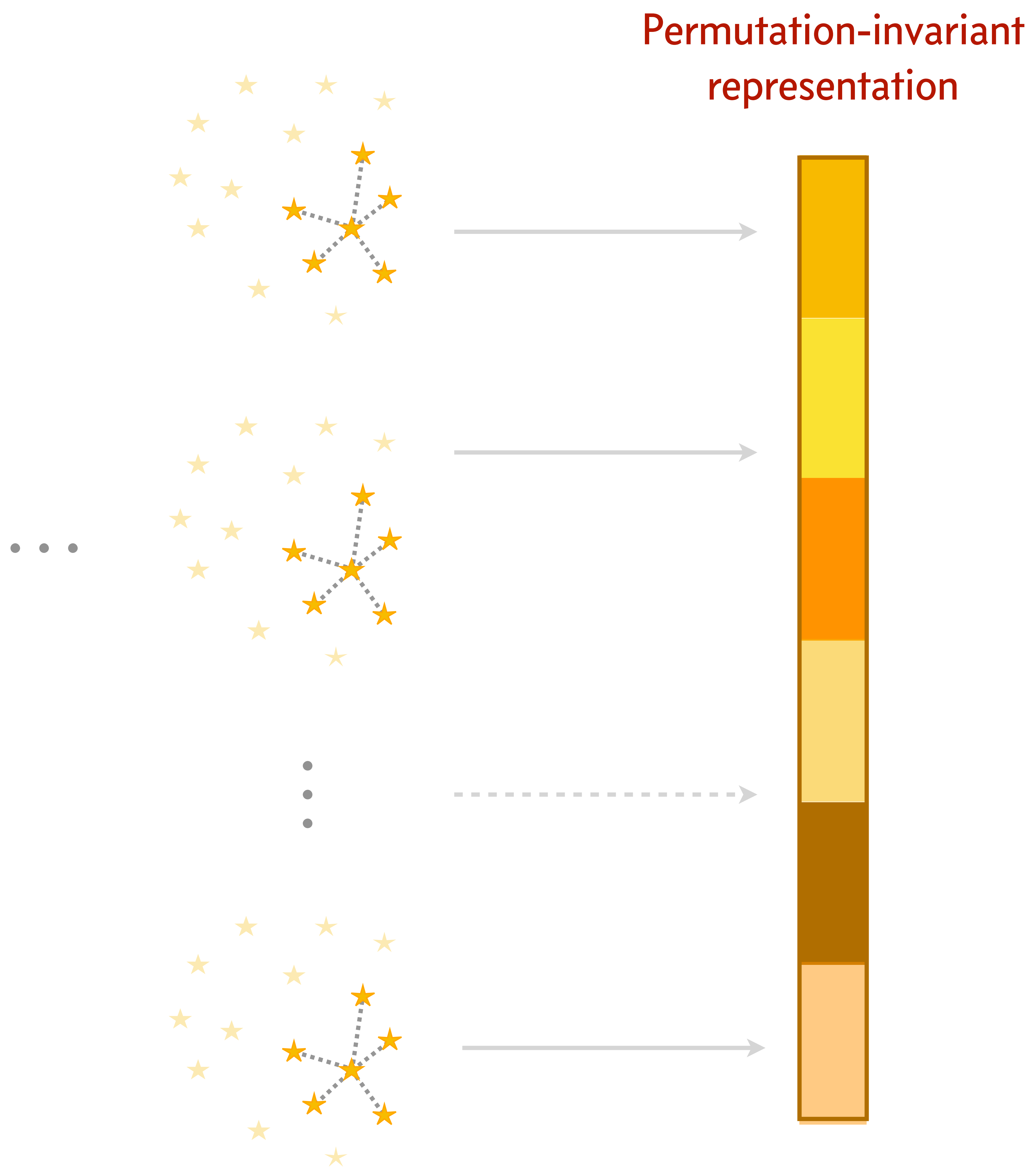


Statistical coverage

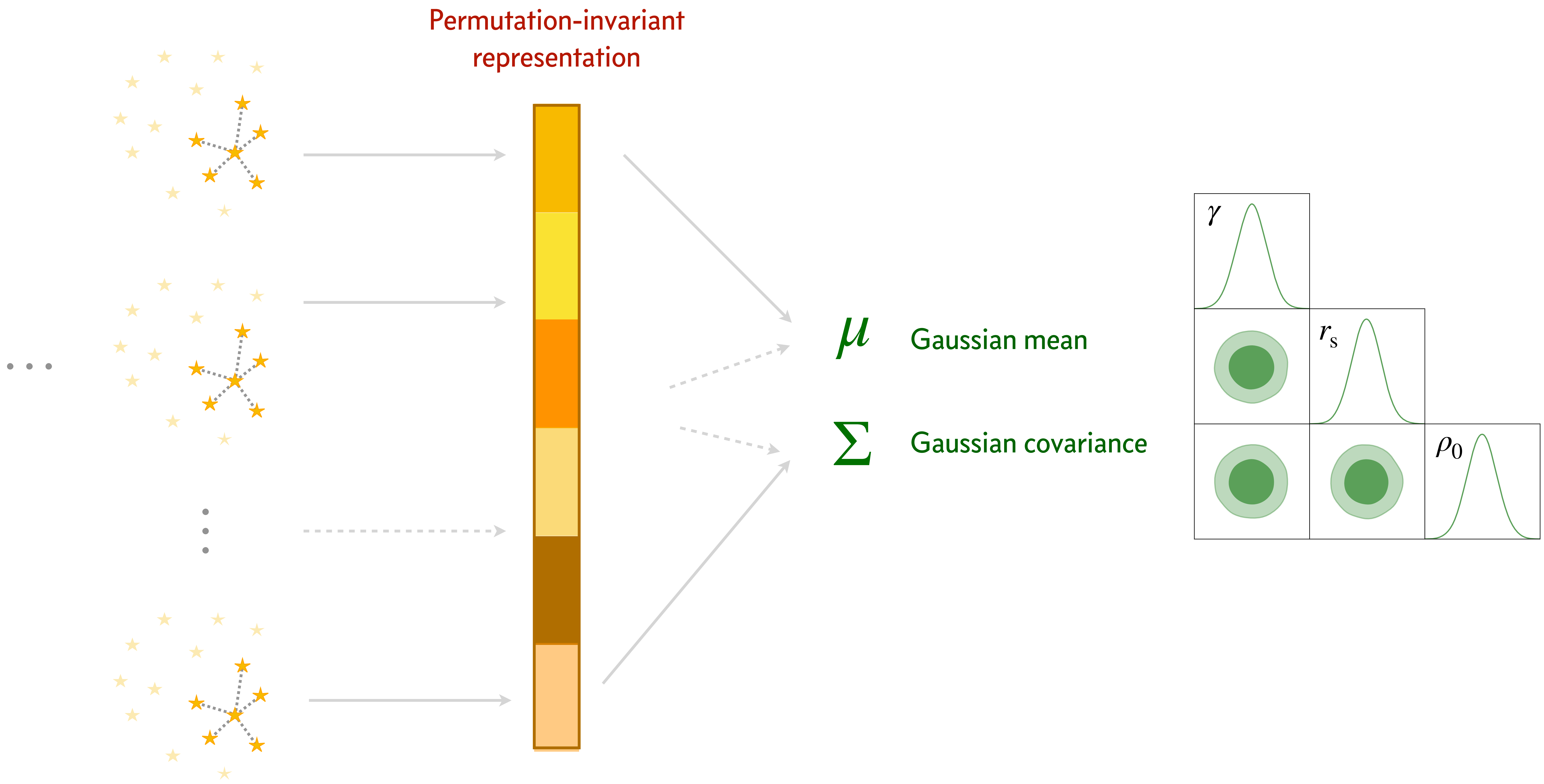




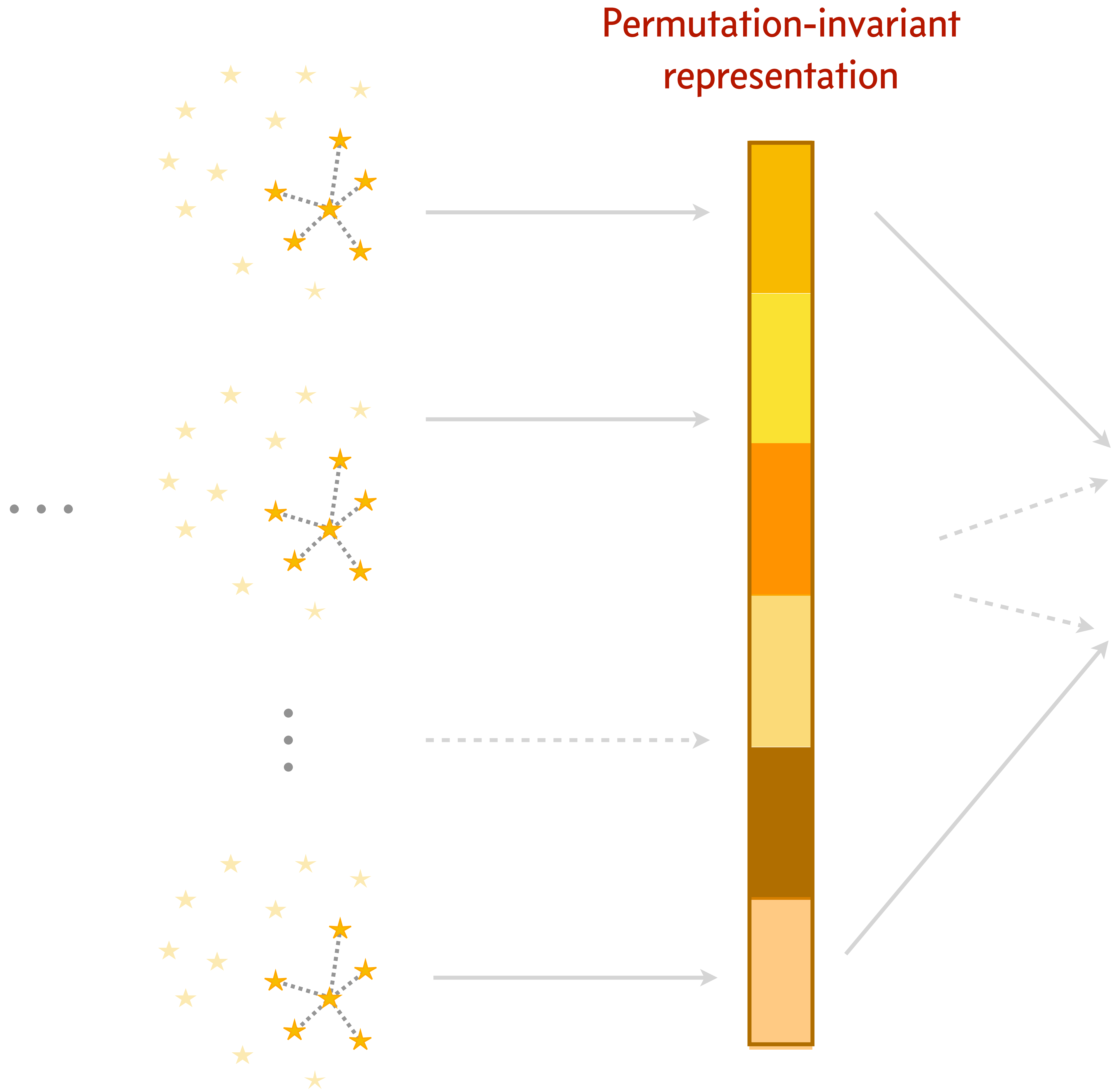
Posterior density estimation



Posterior density estimation

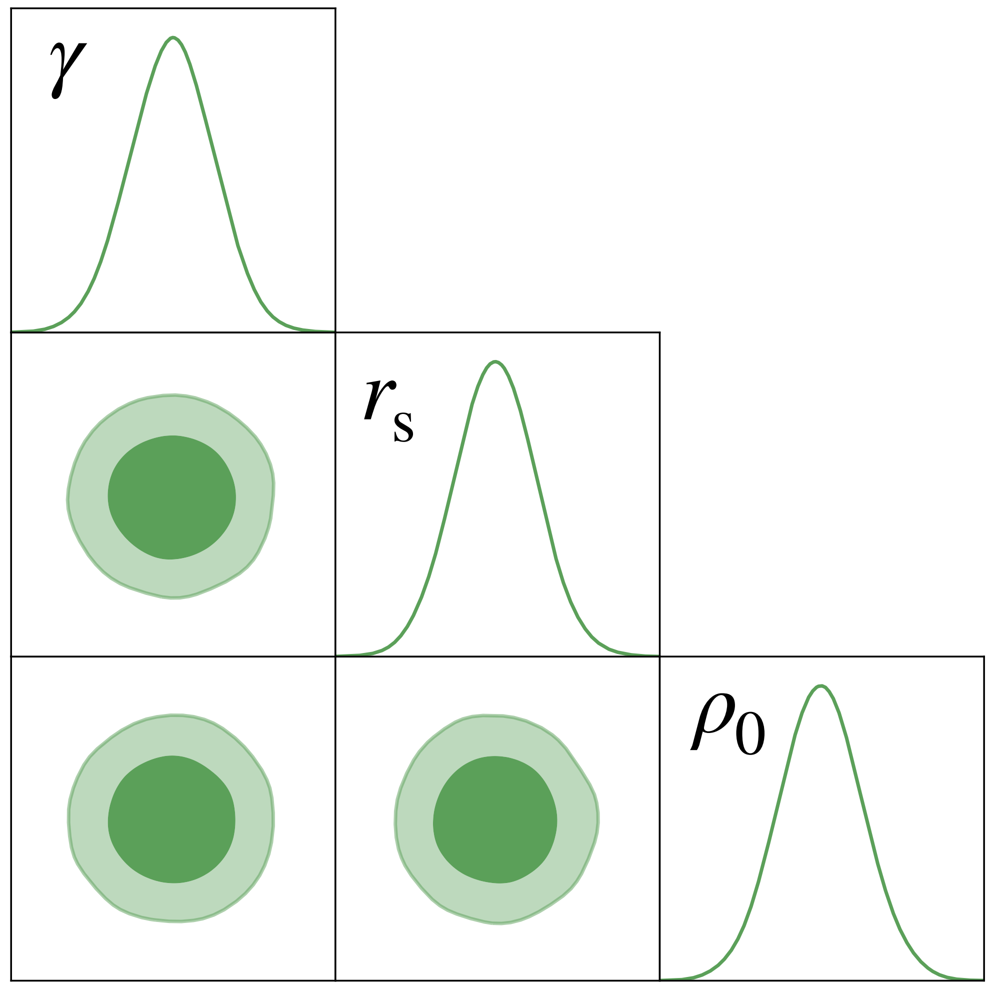


Posterior density estimation



μ Gaussian mean

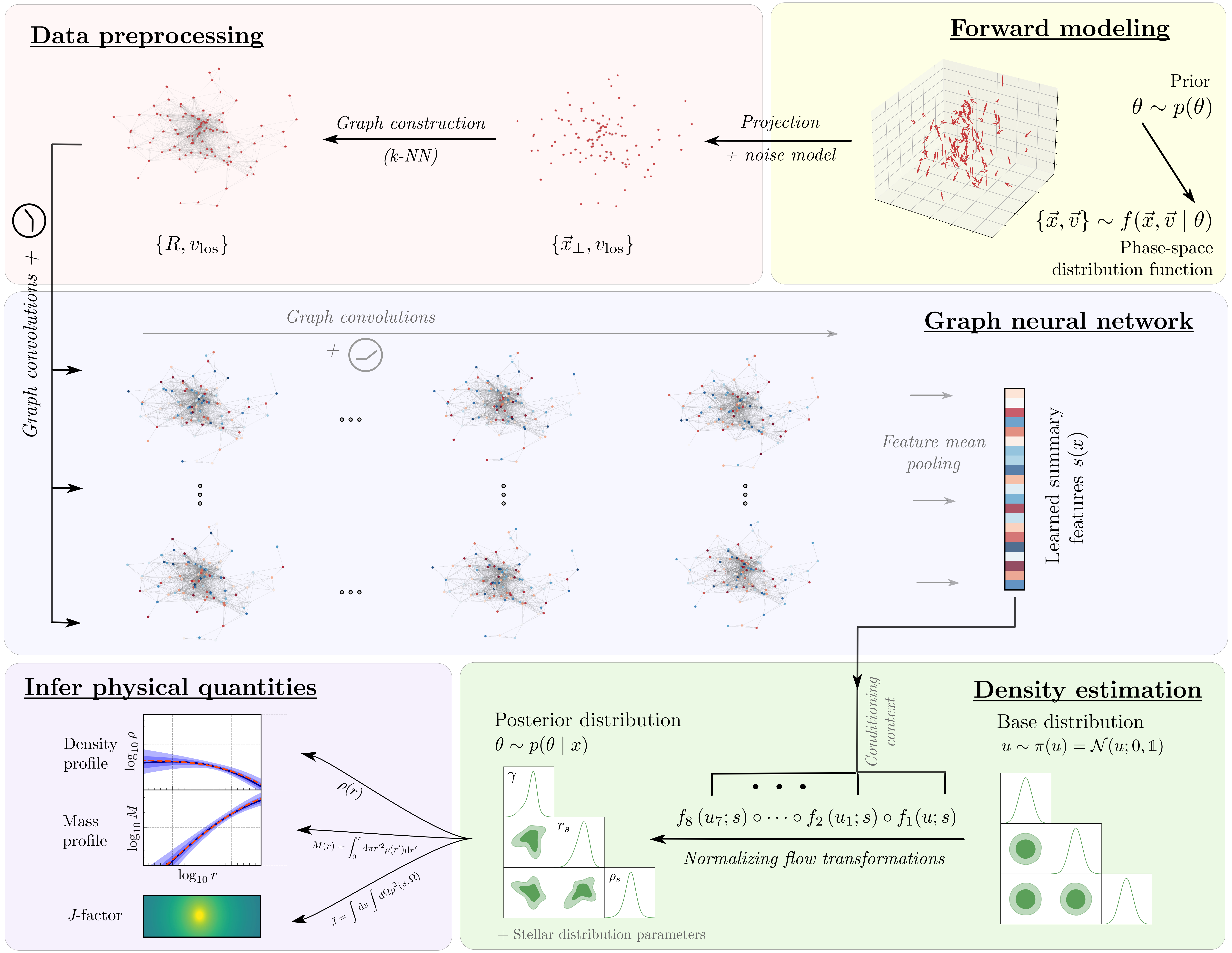
Σ Gaussian covariance



$$\mathscr{L} = \log \mathcal{N}(\theta \mid \mu, \Sigma)$$

Training objective:
Gaussian log-density

Pipeline

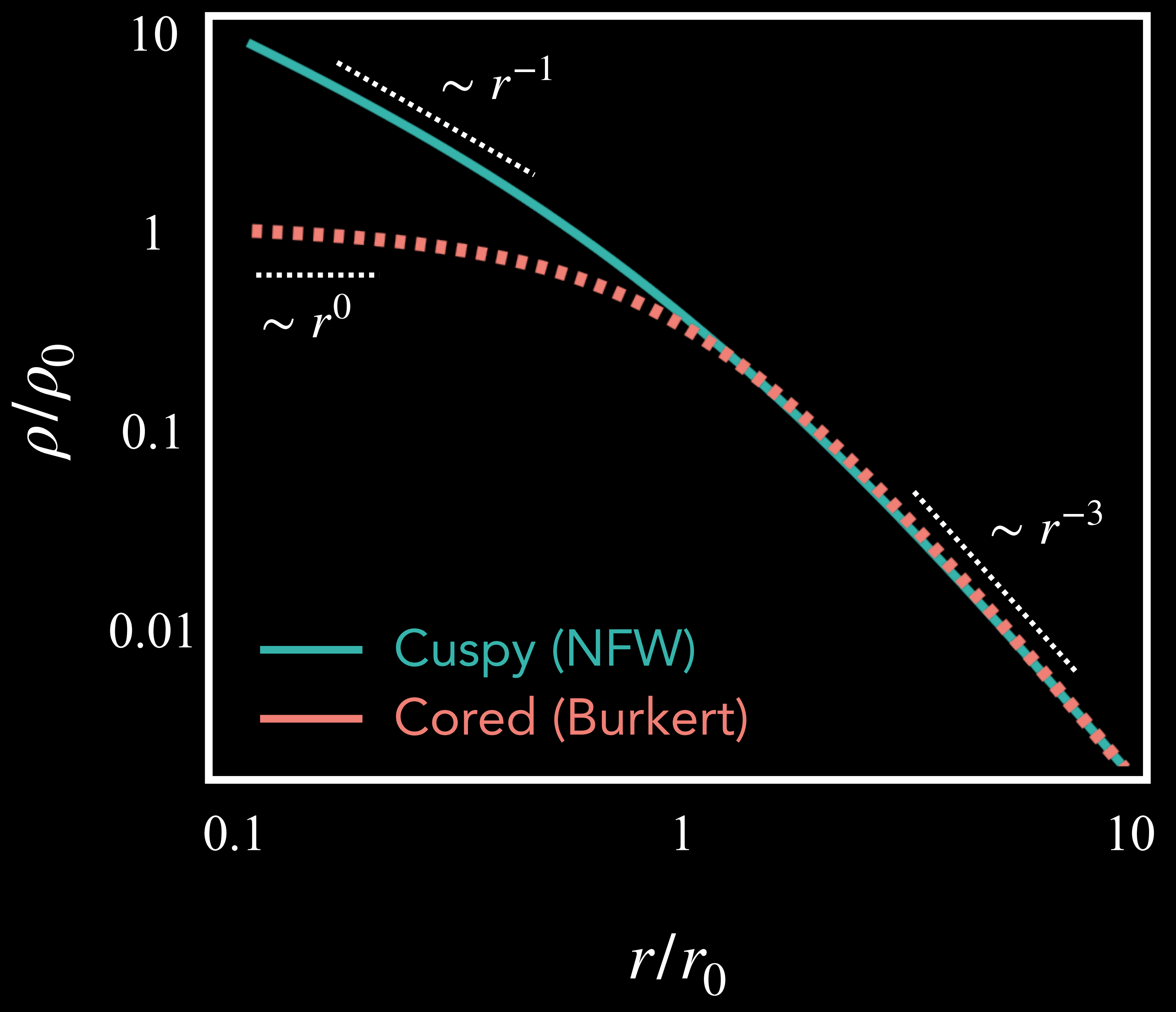


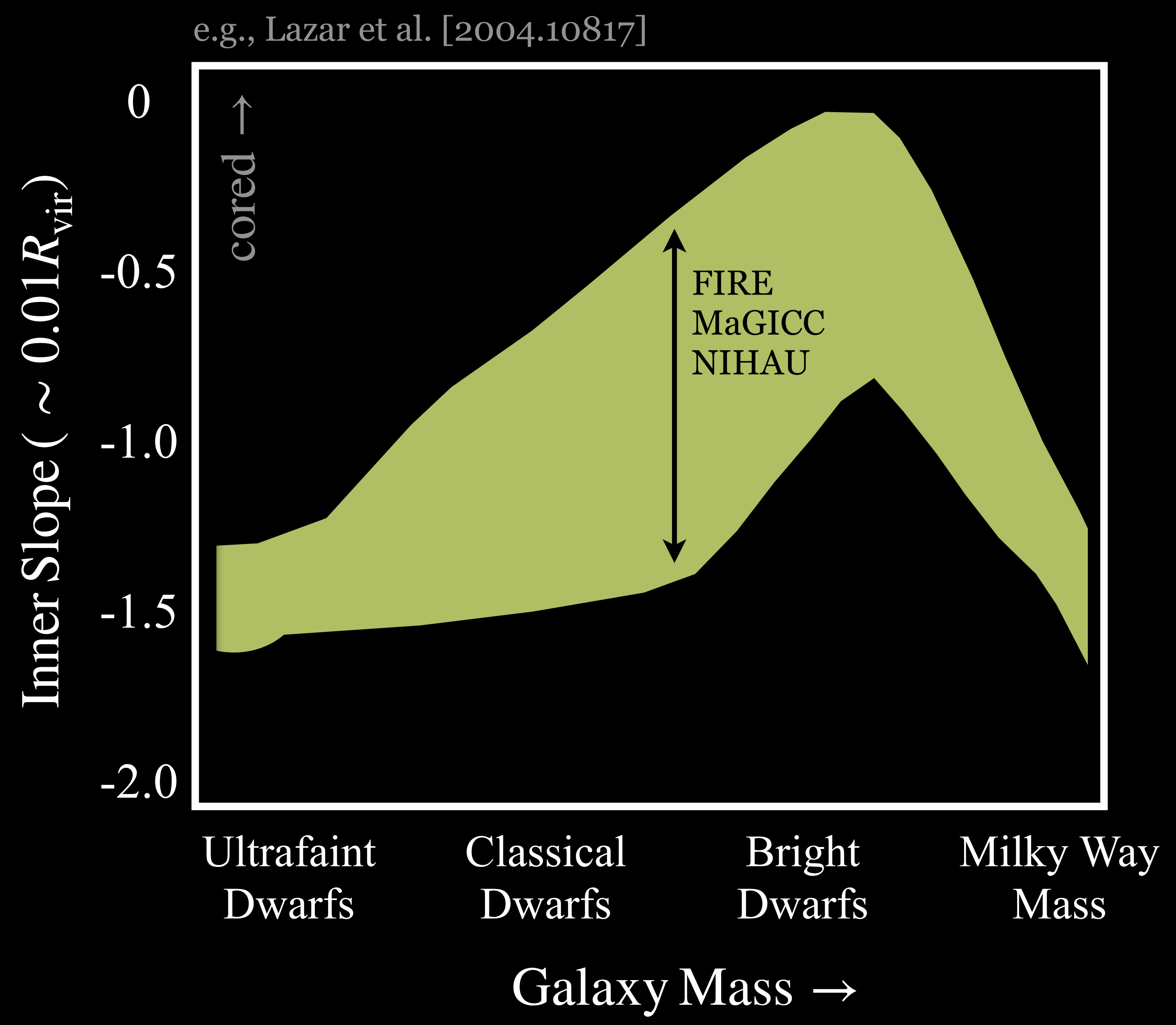
Data preprocessing
Graph construction
(k-NN)
Projection
+ noise model
$\{R, v_{\rm los}\}$
$\{\vec{x}_\perp, v_{\rm los}\}$
Forward modeling
Prior
$\theta \sim p(\theta)$
$\{\vec{x}, \vec{v}\} \sim f(\vec{x}, \vec{v} \mid \theta)$
Phase-space
distribution function
Graph convolutions +
Graph convolutions
Graph neural network
Feature mean
pooling
Learned summary
features $s(x)$
Conditioning
context
Infer physical quantities
Density estimation
Posterior distribution
$\theta \sim p(\theta \mid x)$
Base distribution
$u \sim \pi(u) = \mathcal{N}(u; 0, \mathbb{1})$
$f_8(u_7; s) \circ \cdots \circ f_2(u_1; s) \circ f_1(u; s)$
Normalizing flow transformations
γ
r_s
ρ_s
+ Stellar distribution parameters
Density
profile
$\log_{10} \rho$
Mass
profile
$\log_{10} M$
$\log_{10} r$
J-factor
$\rho(r)$
$M(r) = \int_0^r 4\pi r'^2 \rho(r') dr'$
$J = \int ds \int d\Omega \rho^2(s, \Omega)$

Baryonic Physics

CDM halos of all masses typically have ``cuspy" density profiles

However, baryonic feedback can ``core" the inner region of a CDM halo





Internal Halo Properties

Dark matter self interactions can transfer heat throughout halo, redistributing matter distribution