Learning from galactic rotation curves: a neural network approach

based on B.Dave & G. Goswami, arXiv:2412.03547

Bihag Dave
Ahmedabad University

AI/ML Applications in Astronomy & Astrophysics

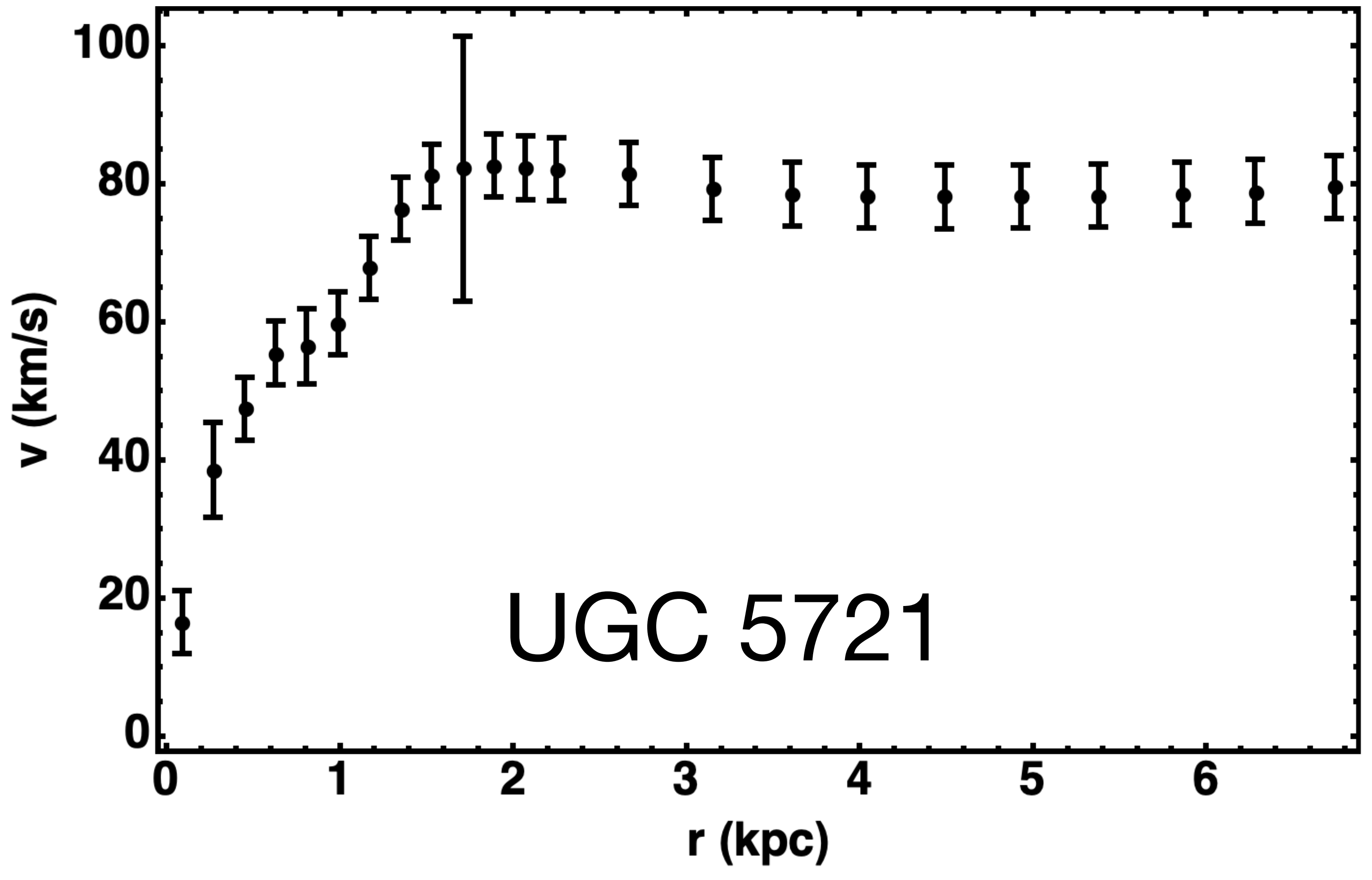
Ahmedabad University

10th January 2025

IUCAA

Learning from rotation curves

Rotation curves are an important set of observations at galactic scales:



SPARC, Lelli et al. (2016)

Orbital velocity of gas or stars around the centre of a galaxy as a function of radius.

Understand and constrain the galactic scale effects of theoretical models

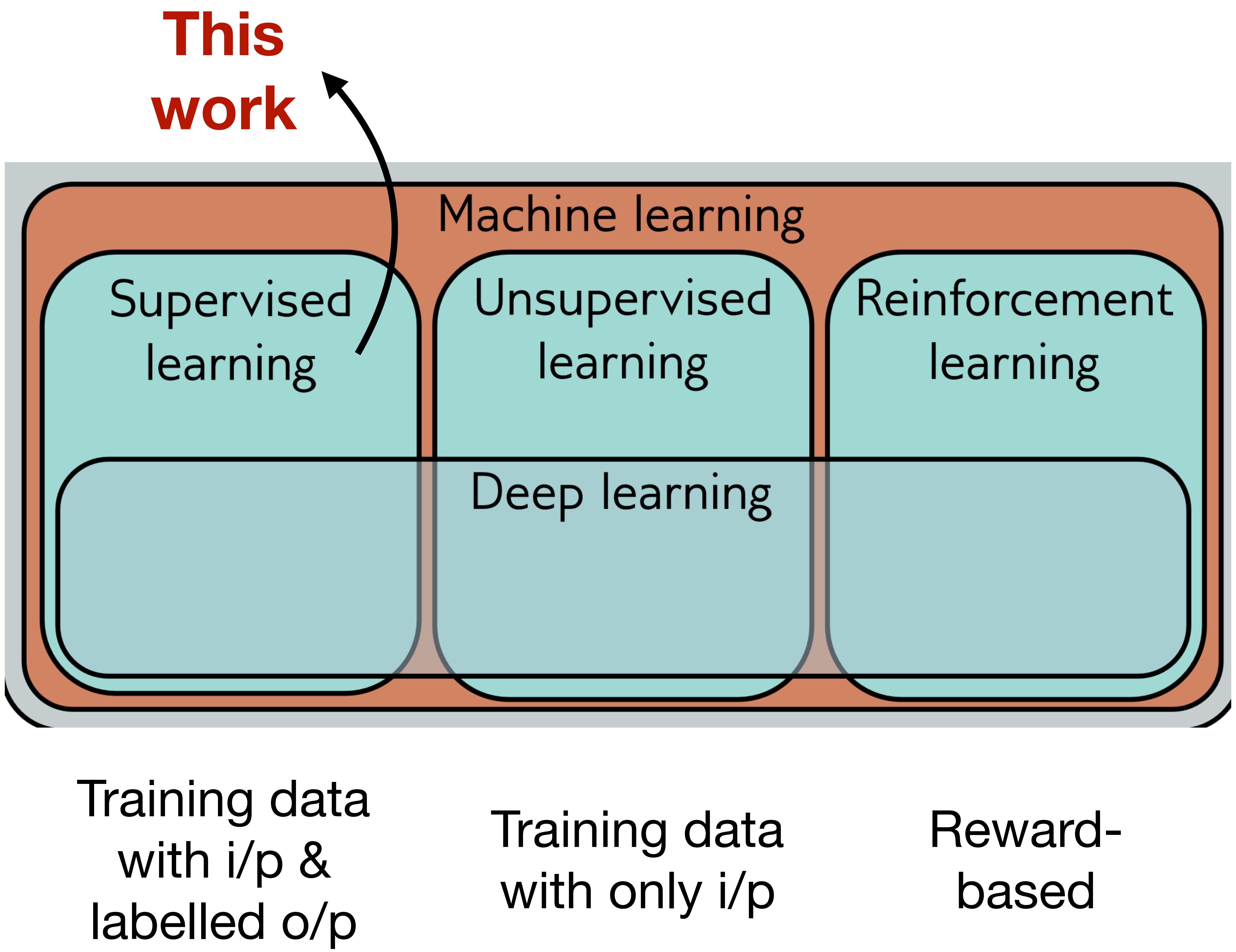
e.g. various models of dark matter and modified gravity

Usually done using Bayesian inference

i.e. using Bayes rule to obtain a posterior distribution of parameters given a likelihood and priors.

Can we use machine learning techniques like neural networks to learn parameters from the observations?

Machine learning



Supervised learning - Neural networks, decision trees, random forests, etc.

Fitting mathematical models to observed data, such that the model generalises to unseen data

Neural networks

Suppose there exists some (potentially complex) relationship between the input $\mathbf{x}$ and output $\mathbf{y}$

Given some observed samples: $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{I}$

Input

Weights (and biases)

$$\mathbf{f}_i = \mathbf{f}(\mathbf{x_i}; \mathbf{\Omega})$$

Find optimum values of $\mathbf{\Omega}$ such that $\mathbf{f}_i$ is 'close' to $\mathbf{y}_i$, i.e. $\mathscr{L}(\mathbf{\Omega})$ is minimised

Loss/cost function

Usual applications in Cosmology and Astrophysics

- Speeding up Bayesian inference - Emulate instead of solving Boltzmann equations to obtain CMB spectra during likelihood computation

A. Spurio-Mancini et al. (2022)

- Non-parametric reconstruction - Rotation curves, H(z), etc.

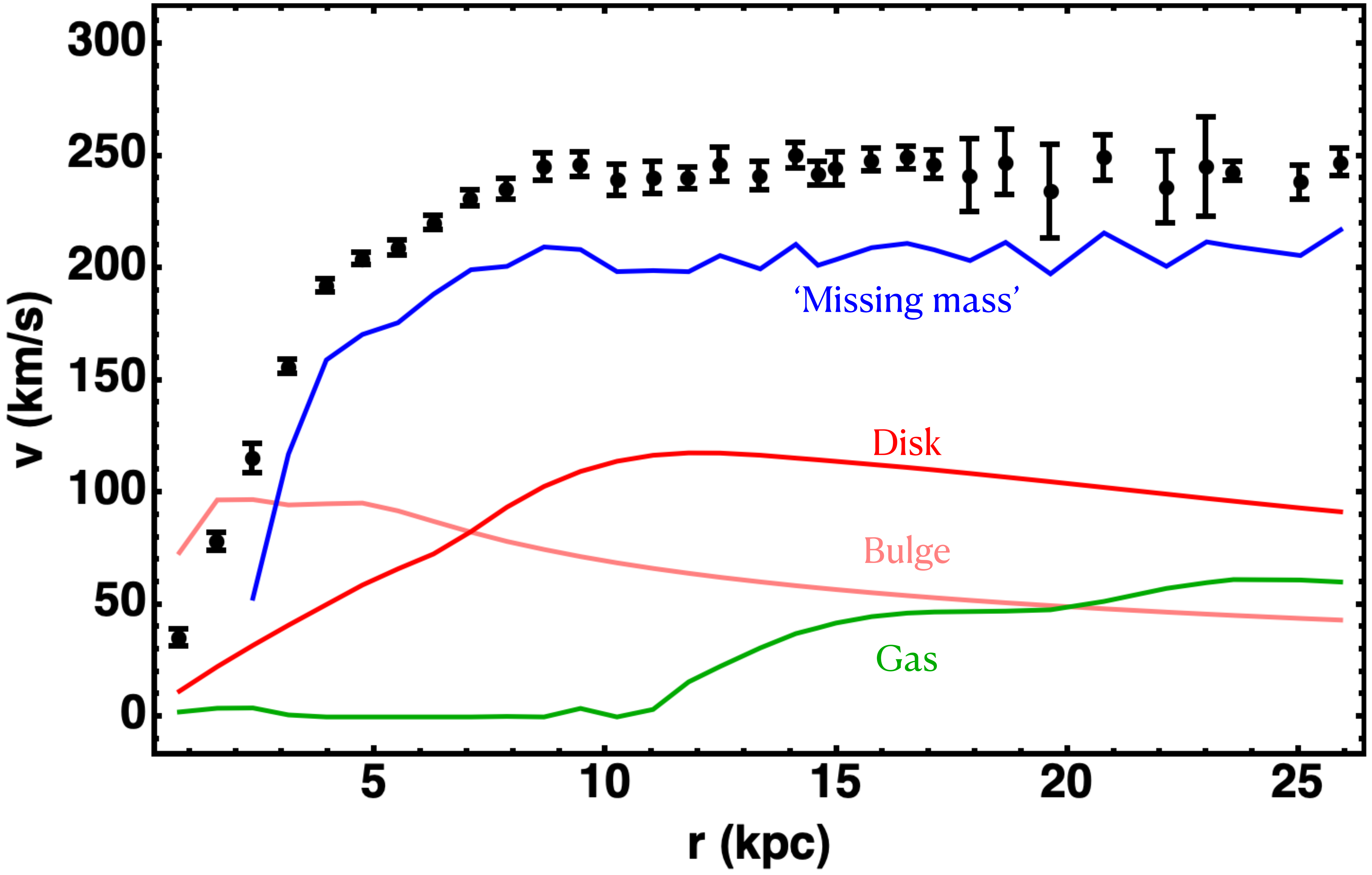
G. Garcia-Arroyo et al. (2024), R. Shah et al. (2024), G. J. Wang et al. (2020)

- Perform direct parameter estimation - Cosmological parameters from H(z) data, Axion string network parameters from CMB birefringence maps, etc.

S. Pal and R. Saha (2023), G. J. Wang et al. (2020), Ray Hagimoto et al. - arXiv: 2411.05002

Galactic rotation curves

Luminous mass in a galaxy is insufficient to describe observed rotation curves



IC4202 from SPARC

Lelli et al. (2016)

$$V_{baryons}^2 = \Upsilon_d V_{disk}^2 + \Upsilon_b V_{bulge}^2 + V_{gas}|V_{gas}|$$

$$V_{baryons} < V_{obs}$$

Remaining contribution from new physics, e.g. DM s.t.

$$V_{DM}^2 + V_{baryons}^2 = V_{obs}^2$$

Nature of DM

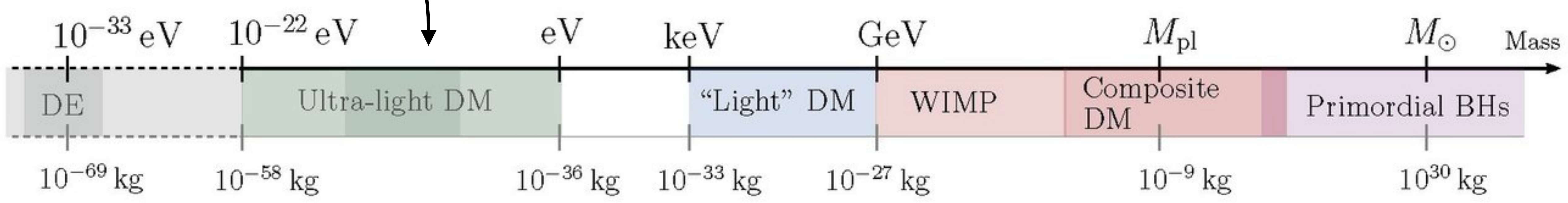
What do we know?

- Non-relativistic $P \ll \rho$
- $\Omega_{DM} \sim 0.2$
- Small or negligible couplings with Standard Model (except gravitationally)

What do we not know?

This talk

- Mass of DM particle? Huge range:



- Spin (Boson or Fermion)?
- Are self-interactions important? ($\lambda \lesssim 1\mathrm{cm}^2/g$)

Observations

$m n_{DM} \sim 0.3\ \mathrm{GeV/cm^3}$

$\mathrm{v} \sim 10^{-3}$

<u>Scalar Field Dark Matter</u>

Spin-zero light boson

$m \ll \mathcal{O}(10)$ eV → $\lambda_{dB} \gg d$ → Describe using a **classical scalar field!**

Scalar field dark matter

$$S_\varphi = \frac{1}{16\pi G}\int d^4xR\sqrt{-g} + \int d^4x\sqrt{-g}\left(-\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - U(\varphi)\right)$$

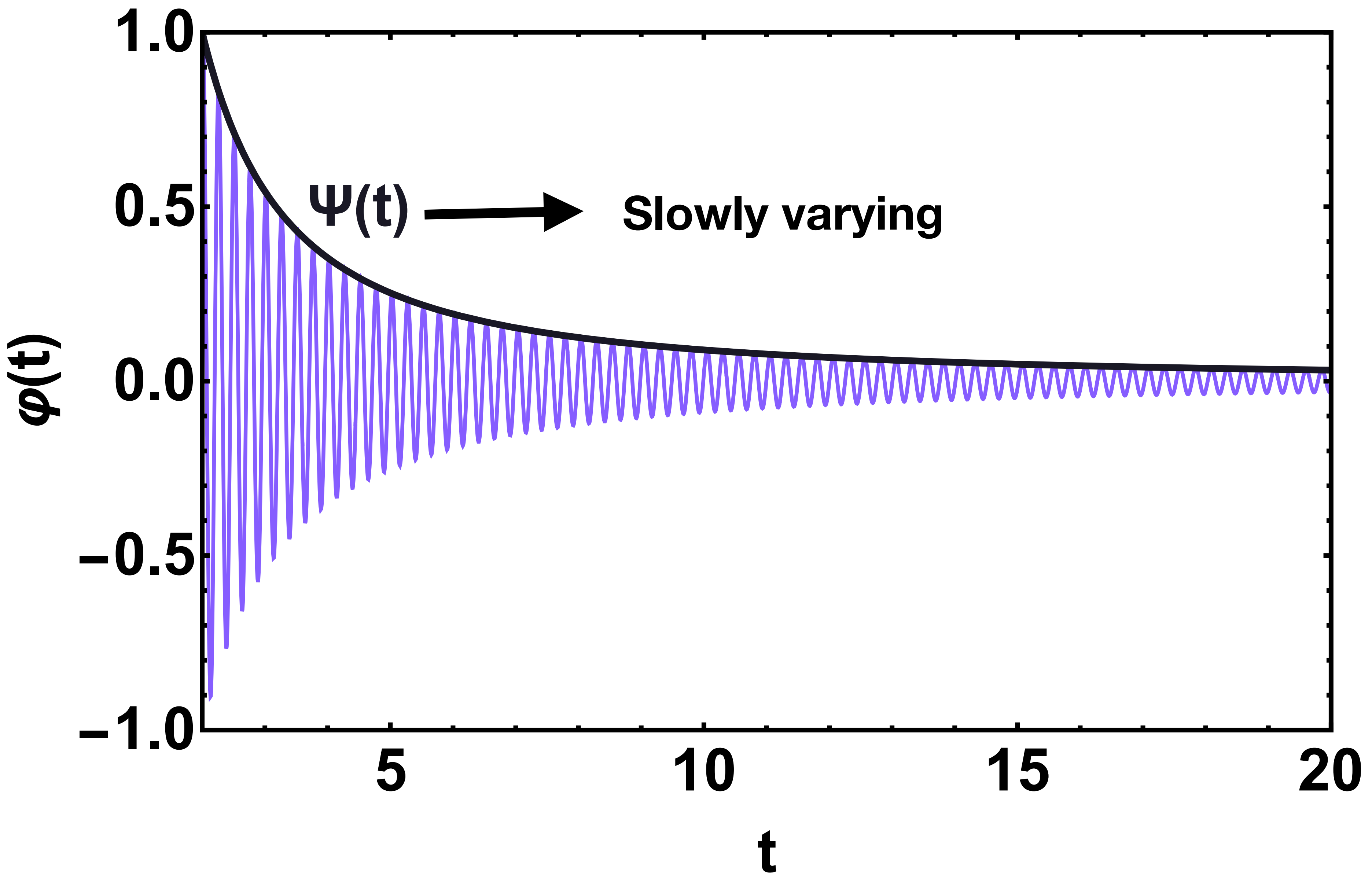
$$U(\varphi) = \frac{1}{2}m^2\varphi^2 + \cdots$$

Can have higher order terms as well

Field redefinition:

$$\varphi \equiv \frac{1}{\sqrt{2m}}\left[e^{-imt}\,\Psi + e^{imt}\,\Psi^*\right]$$

Non-relativistic limit: $\Psi \gg \dot{\Psi}/m \gg \ddot{\Psi}/m^2$



Shrödinger-Poisson Equations

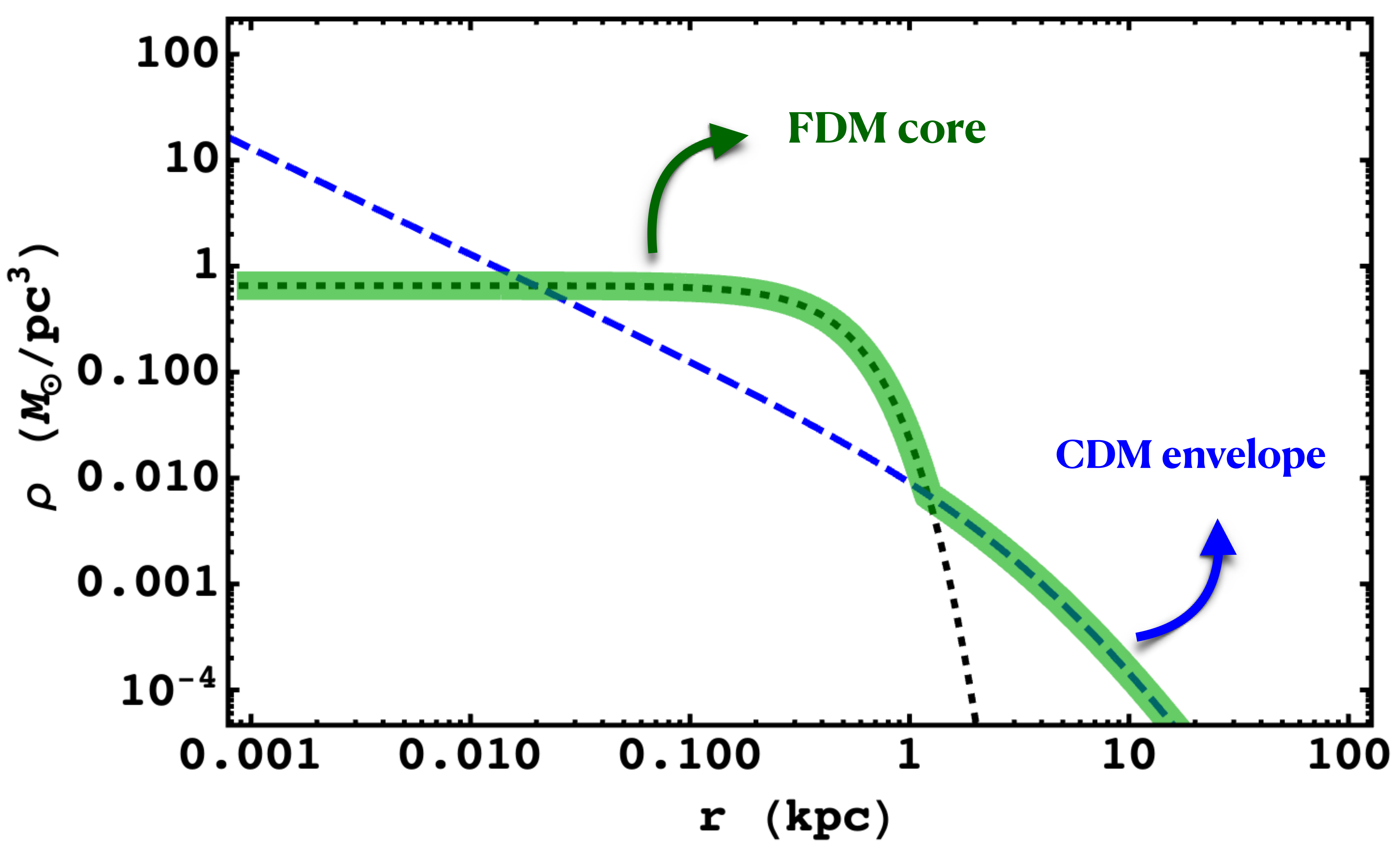
$$i\,\dot{\Psi} = -\frac{1}{2m}\nabla^2\Psi + m\Phi\Psi$$

$$\nabla^2\Phi = 4\pi G|\Psi|^2$$

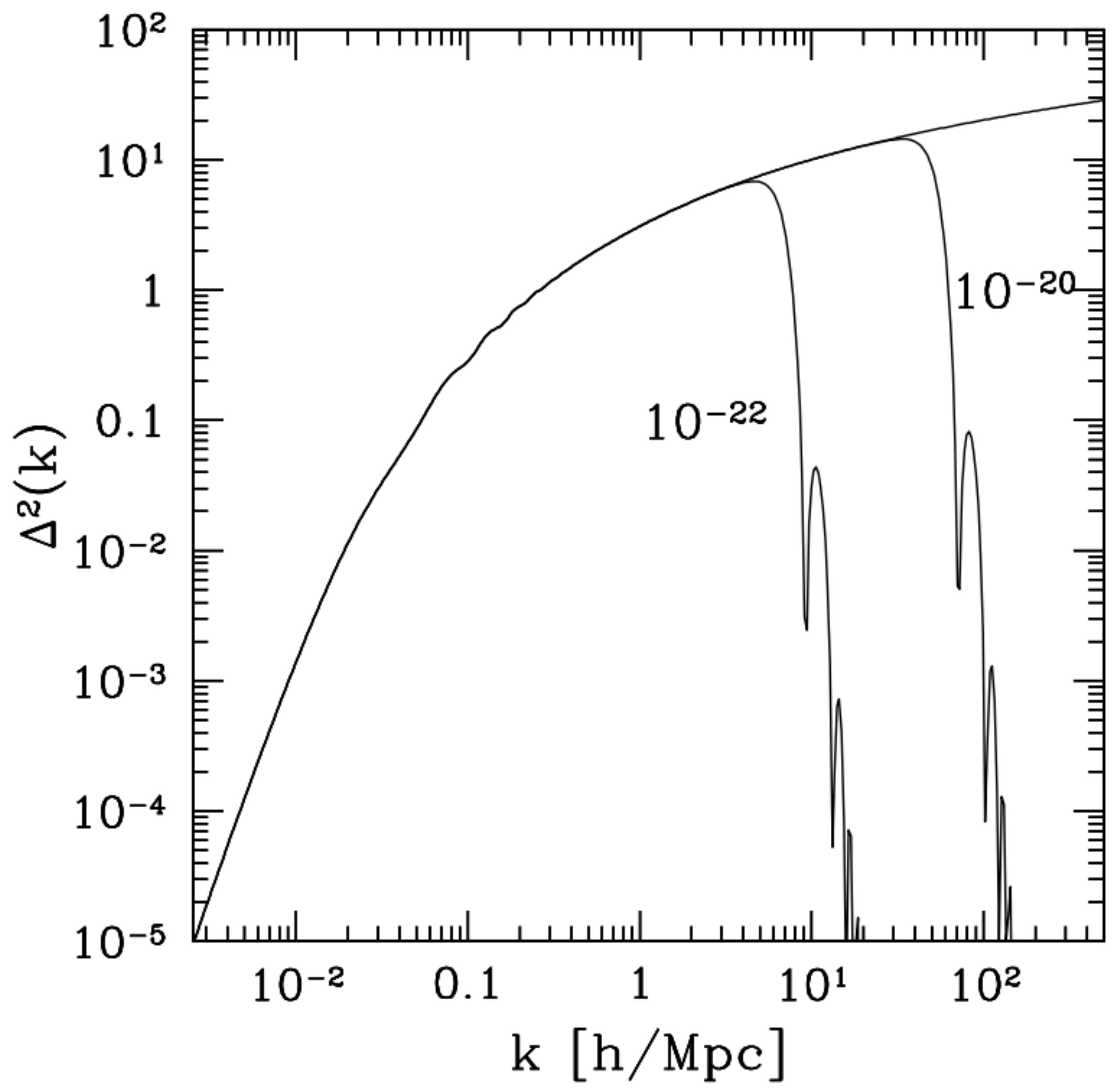
Mass density

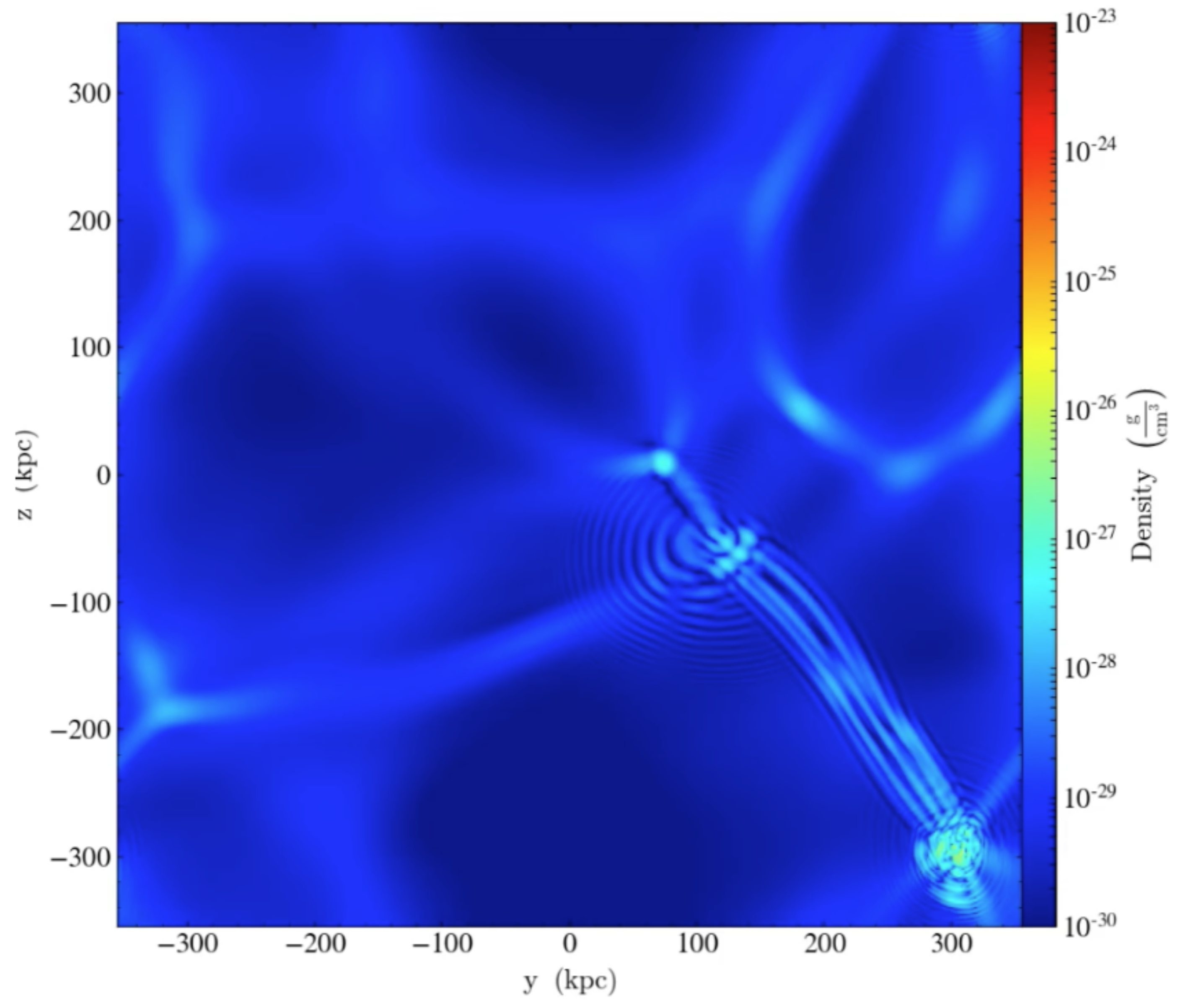
Fuzzy Dark Matter (FDM)

- $m \sim 10^{-22}$ eV $\implies \lambda_{dB} \sim \mathcal{O}(\text{kpc})$
- Interesting effects at galactic scales ⟶ **Flat-density cores, interference, power suppression…**



Core-halo structure expected from simulations by *Schive at al. 2014*

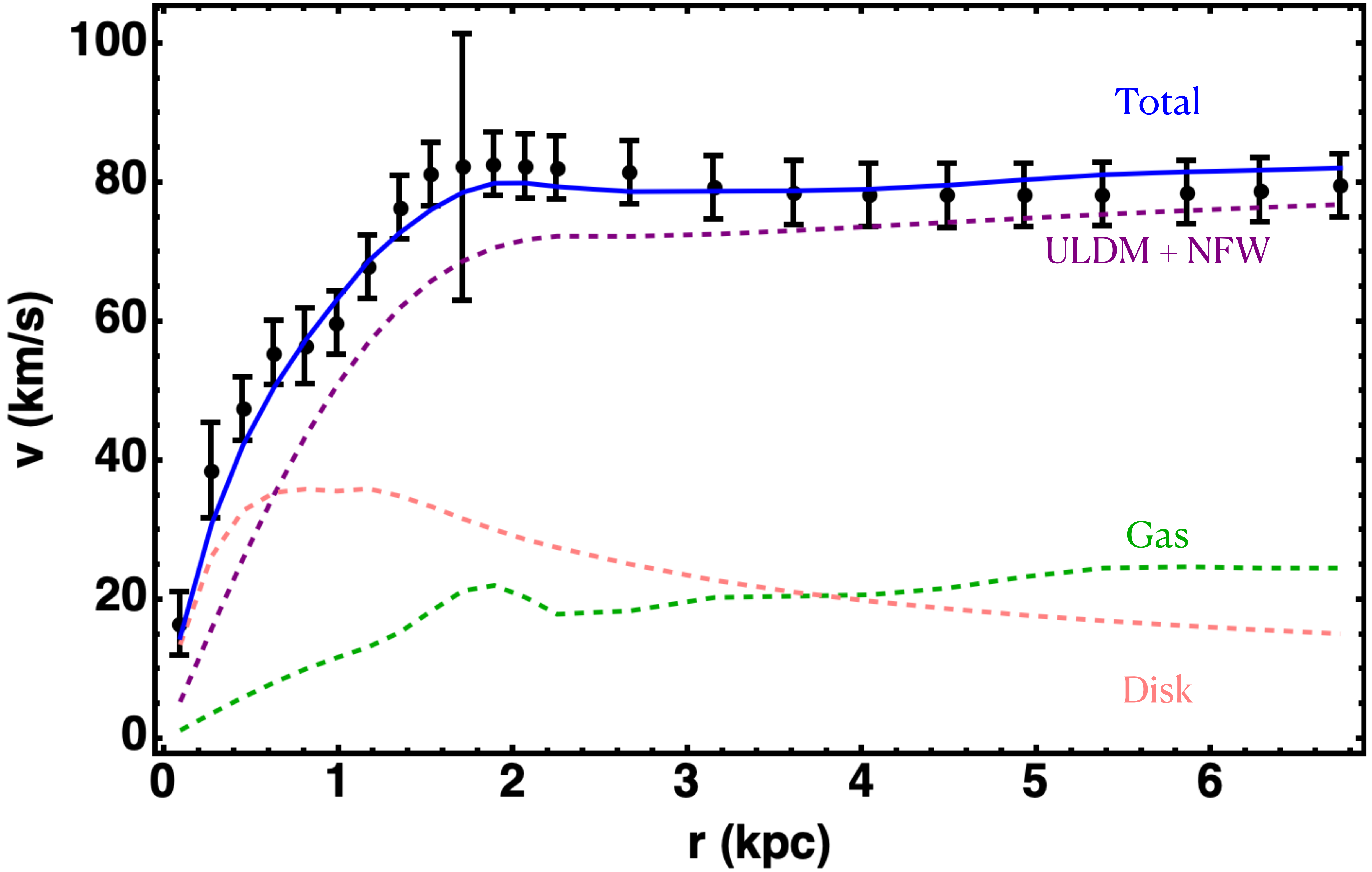




Hui, Wave Dark Matter (2021)

Parameters of Interest

UGC 5721 from SPARC



$$V_{tot}^2 = V_{DM}^2 + \Upsilon_d V_{disk}^2 + |V_g| V_g$$

Dark Matter — Disk — Gas

Baryonic contribution can be tuned using

$$\Upsilon_* \equiv \Upsilon_d$$

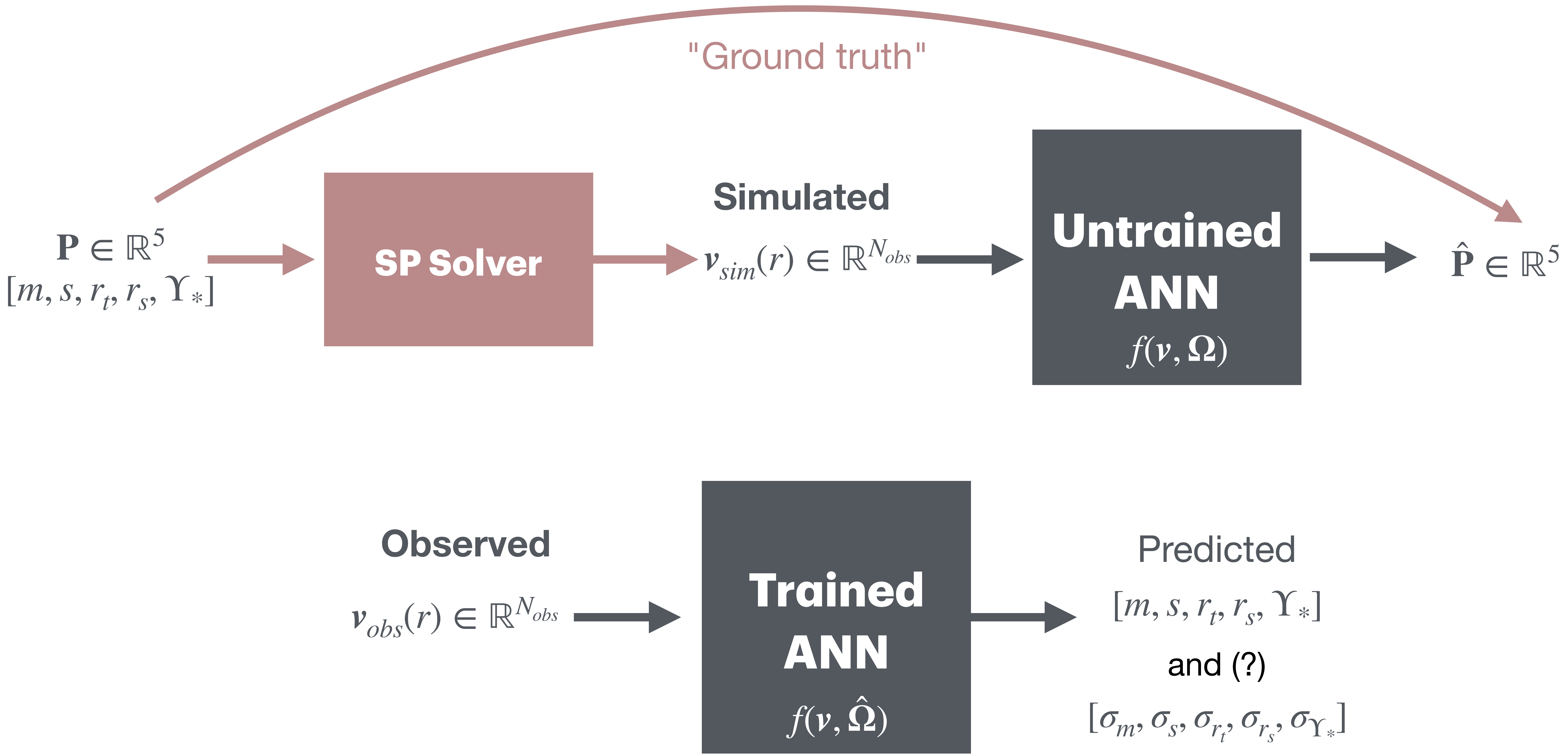
DM contribution can be tuned using

$$\rho_{DM} = \Theta(r_t - r)\rho_{ULDM}(r) + \Theta(r - r_t)\rho_{NFW}(r)$$

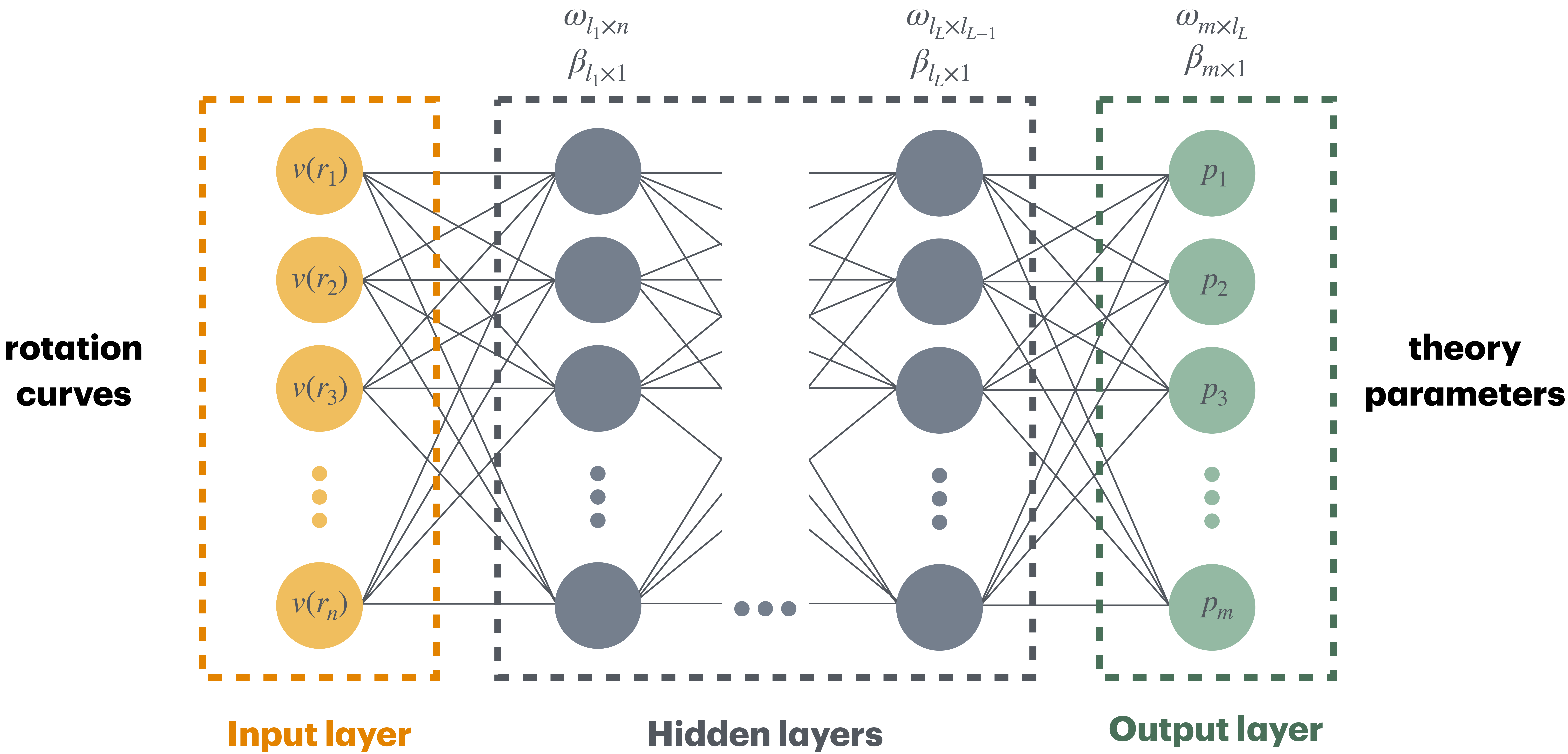
$$\{m, s, r_t, r_s\}$$

Goal: ***Obtain values of parameters given an observed rotation curve.***

How can neural networks help?



Neural network



Generating simulated rotation curves

An appropriate uniform distribution of parameters → For every choice of parameter vector $\{m, s, r_t, r_s, \Upsilon_*\}$

FDM Schive et al. (2014)

$$\rho_{FDM}(r) \approx \frac{0.019\left(\frac{m}{10^{-22}\ eV}\right)^{-2}\left(\frac{r_c}{kpc}\right)^{-4}}{1+\left(0.091\left(\frac{r}{r_c}\right)^2\right)^8}$$

$$r_c = 0.8242\left(\frac{s}{10^4}\right)\left(\frac{m}{10^{-22}\ \text{eV}}\right)^{-1}\ \text{kpc}$$

NFW

$$\rho_{NFW}(r) = \frac{\rho_s}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}$$

$$\rho_{FDM}(r_t) = \rho_{NFW}(r_t)$$

Baryons

Gas V_g and disk V_d contributions modelled by SPARC

↓

$$\rho_{DM} = \Theta(r_t - r)\rho_{FDM}(r) + \Theta(r - r_t)\rho_{NFW}(r)$$

$$V_{DM}^2(r) = \frac{GM(r)}{r} = \frac{4\pi G\int_0^r \rho_{DM}(r')r'^2 dr'}{r}$$

→

$$V_{sim}(r) = \sqrt{V_{DM}^2(m, s, r_t, r_s) + \Upsilon_* |V_d| V_d + |V_g| V_g}$$

Explored parameter space

Galaxy	Parameter ranges				
	m (10^{-23} eV)	s (10^3)	r_t (kpc)	r_s (kpc)	$\Upsilon_*(M_\odot/L_\odot)$
DDO 154	$[1, 10]$	$[3, 9]$	$[1, 5.99]$	$[1, 15]$	$[0.3, 0.8]$
ESO444-G084	$[1, 10]$	$[2, 9]$	$[1, 4.44]$	$[1, 15]$	$[0.3, 0.8]$
UGC 5721	$[1, 10]$	$[1.5, 5]$	$[1, 6.74]$	$[1, 15]$	$[0.3, 0.8]$
UGC 5764	$[1, 10]$	$[2, 9]$	$[1, 3.62]$	$[1, 15]$	$[0.3, 0.8]$
UGC 7524	$[1, 10]$	$[1, 9]$	$[1, 10.69]$	$[1, 15]$	$[0.3, 0.8]$
UGC 7603	$[1, 10]$	$[2, 7]$	$[1, 4.11]$	$[1, 15]$	$[0.3, 0.8]$
UGC A444	$[1, 10]$	$[2, 9]$	$[1, 2.55]$	$[1, 15]$	$[0.3, 0.8]$

Hyperparameters

Hyperparameters	Value
Number of hidden layers	2
Number of neurons per layer	200
Epochs	250
Batch size	32
Dropout probability	0.2
Optimizer	ADAM
Learning rate	0.0001

Grid search for 4 galaxies (number of hidden layers, number of neurons per layer)

Trial and error (epochs, batch size, dropout probability, optimizer, learning rate)

80:10:10 train-test-validation split

Loss function: Mean-squared error

$$\mathscr{L} = \frac{1}{np} \sum_{i=1}^{n} \left(\sum_{j=1}^{p} |y_{ij} - f_{ij}|^2 \right)$$

n is the number of samples

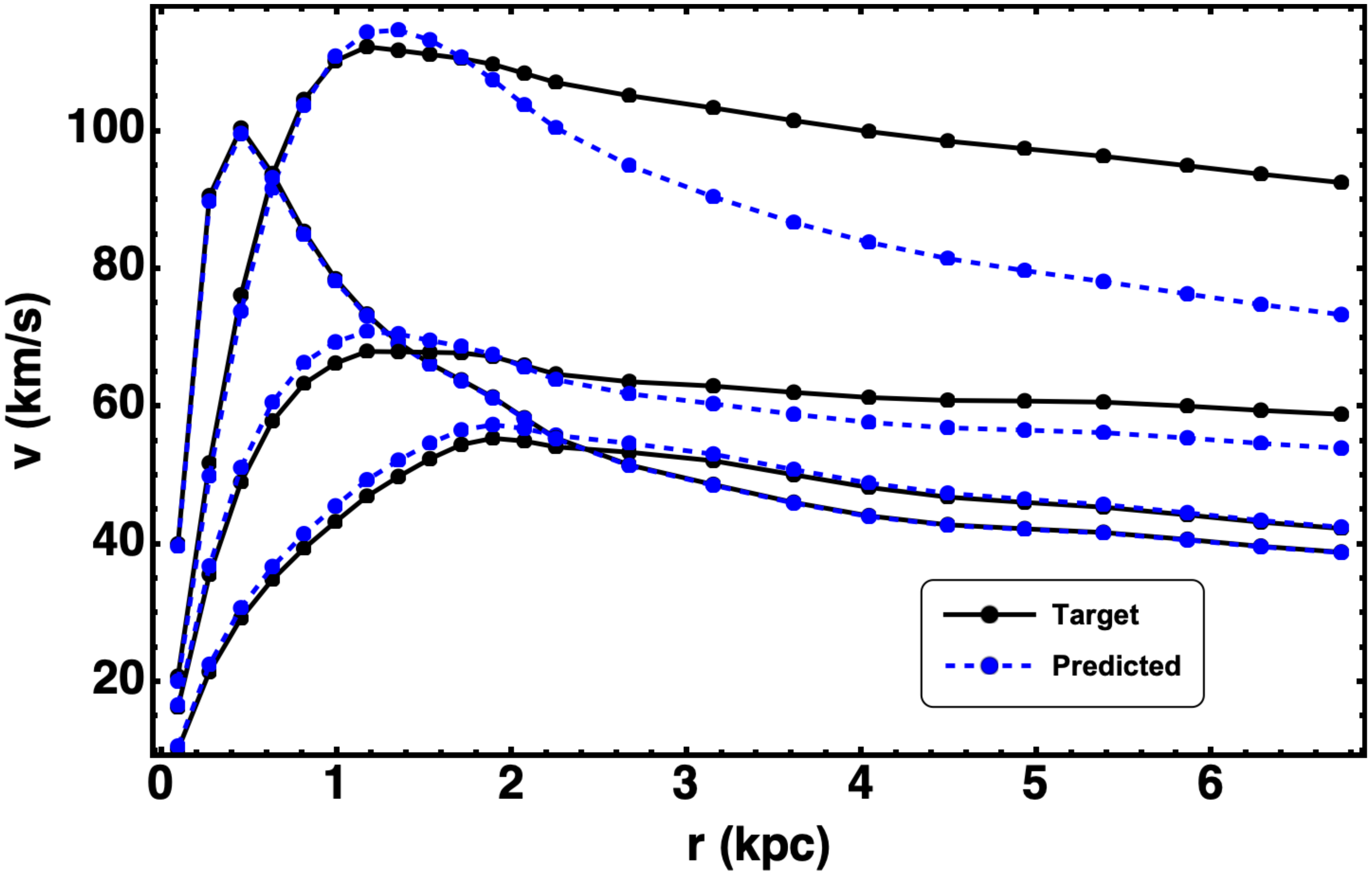
p is the size of the output vector

$y_{ij} \rightarrow$ ground truth

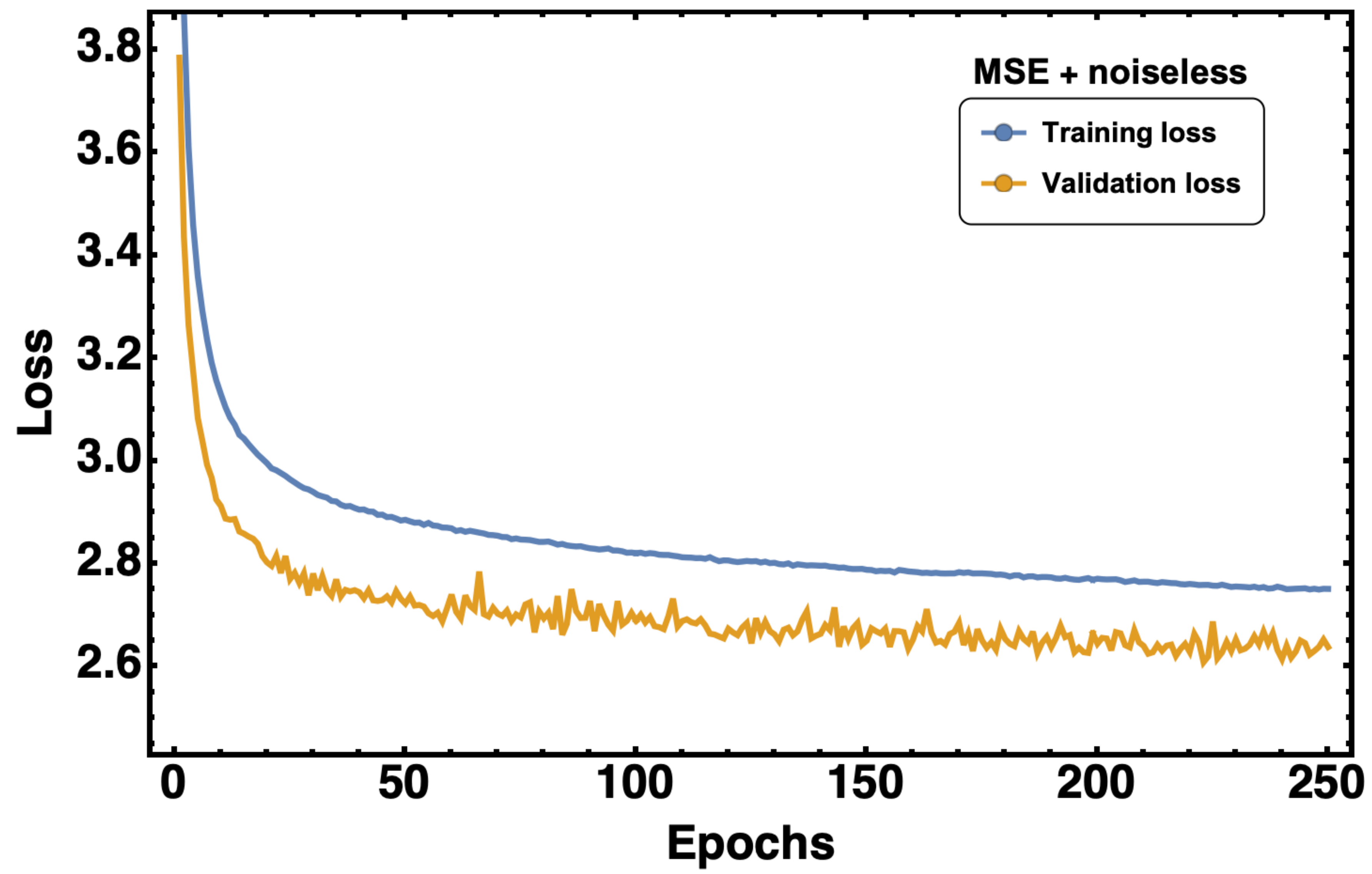
$f_{ij} \rightarrow$ neural network o/p

Testing

4 randomly chosen test samples

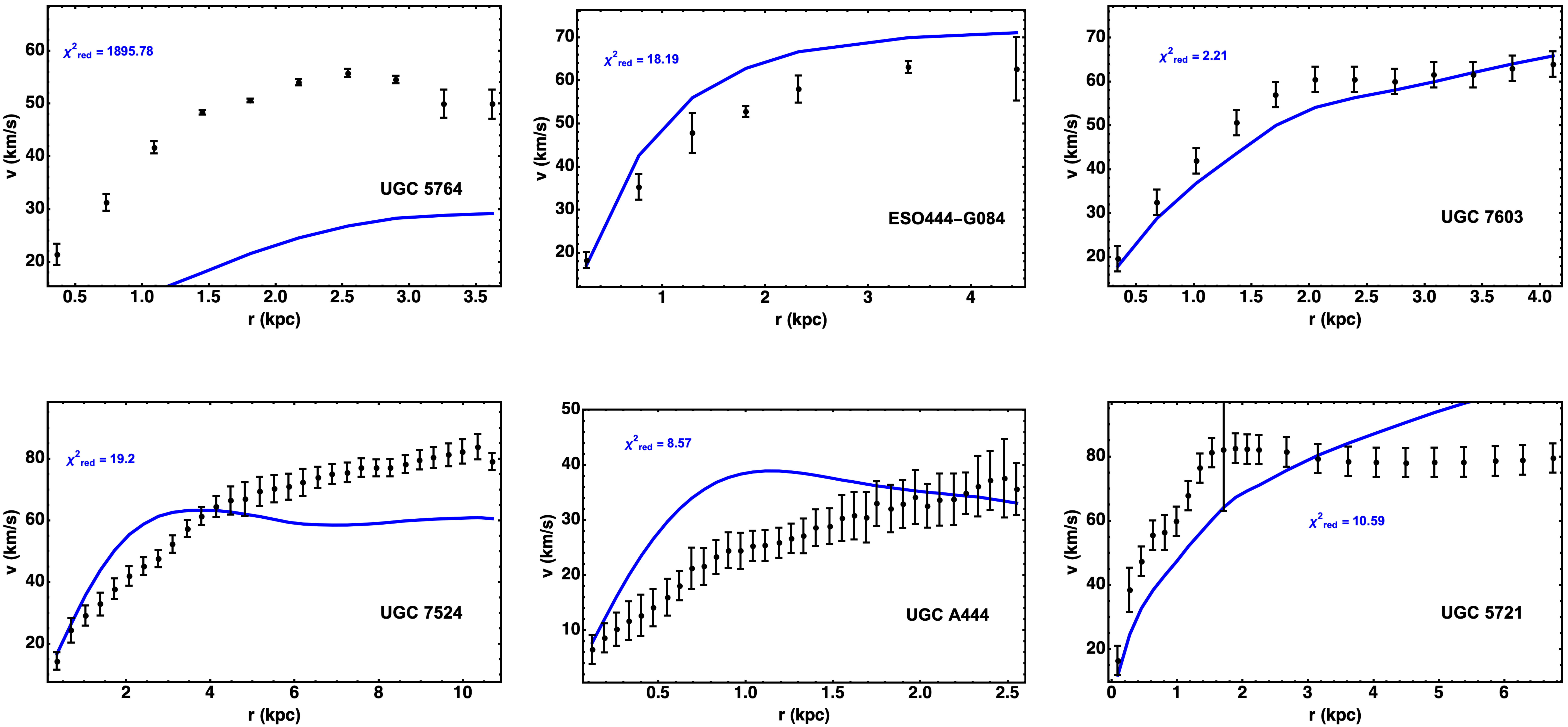


For most rotation curves in the test dataset, the ANN does reasonably well.

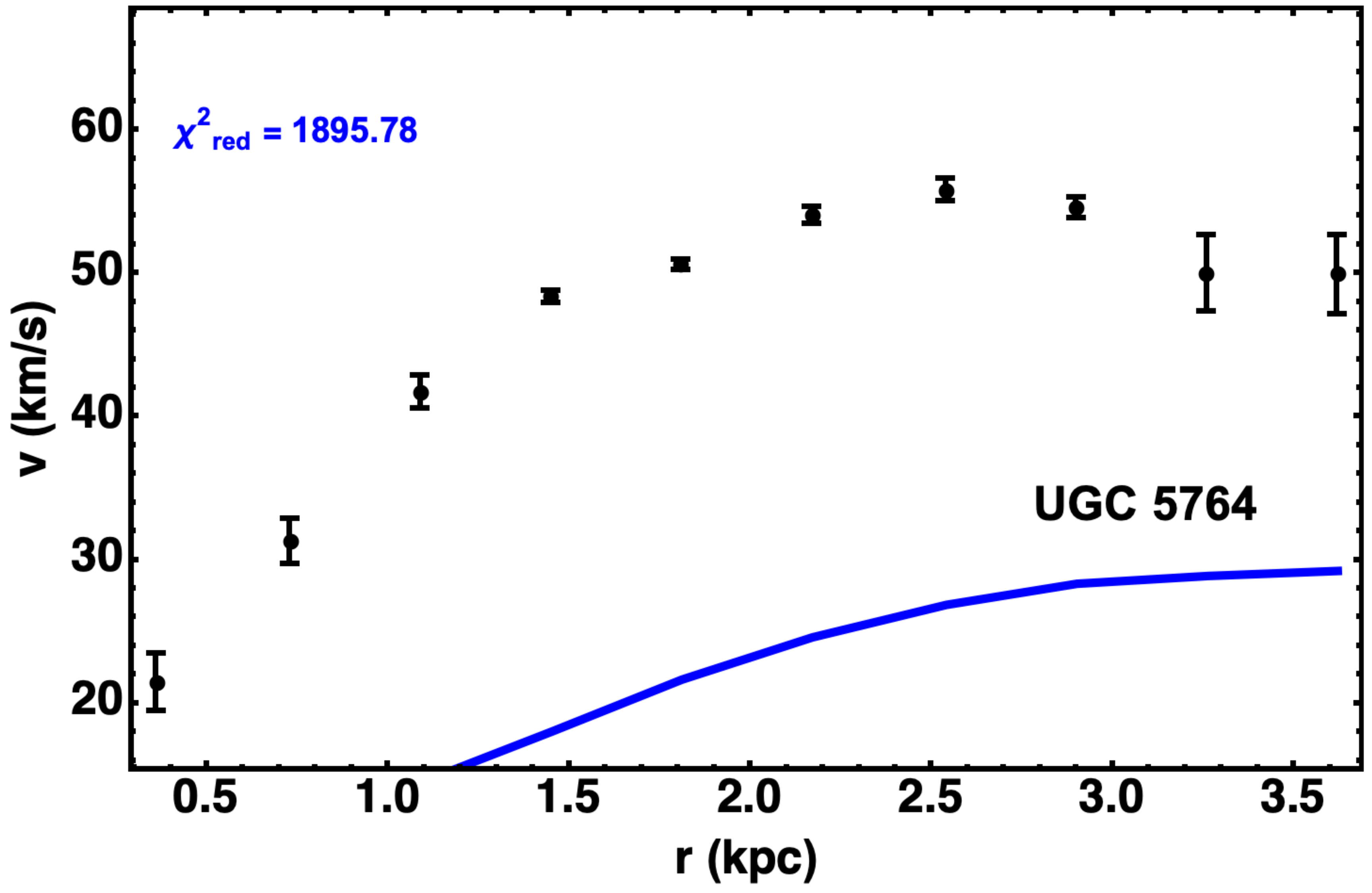


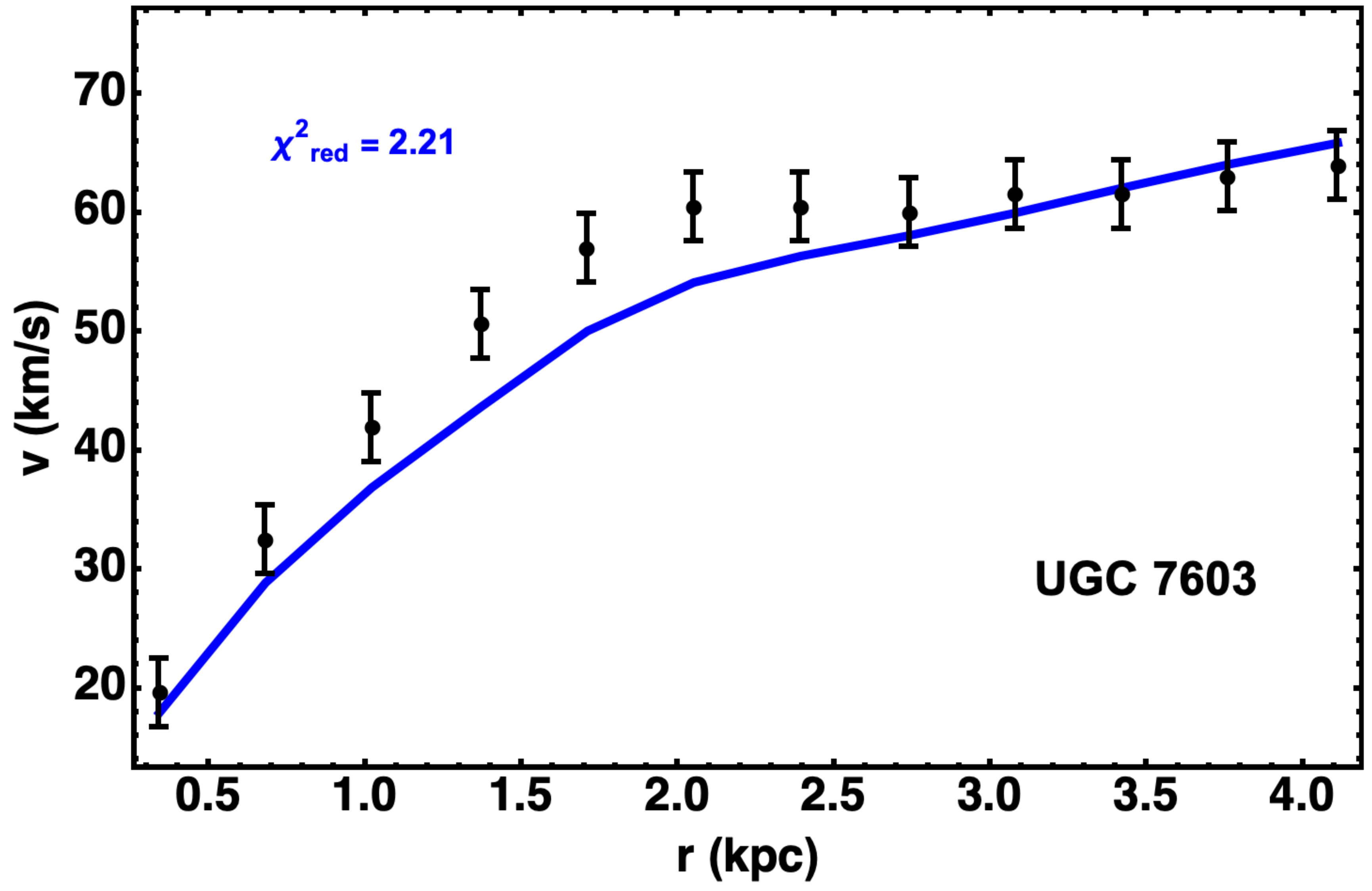
Loss also converges by 250 epochs

But, when confronted by observations

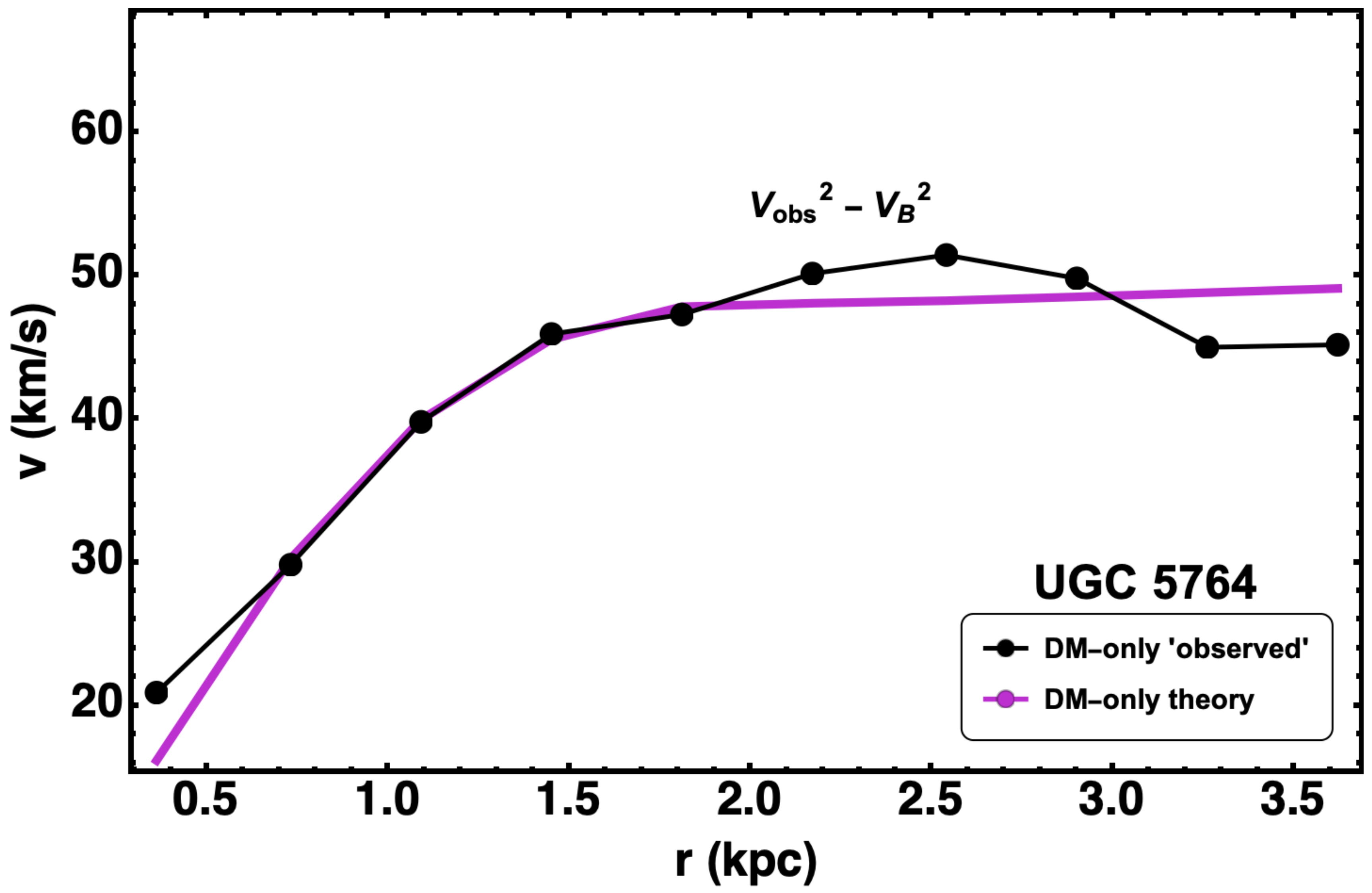


Explanation

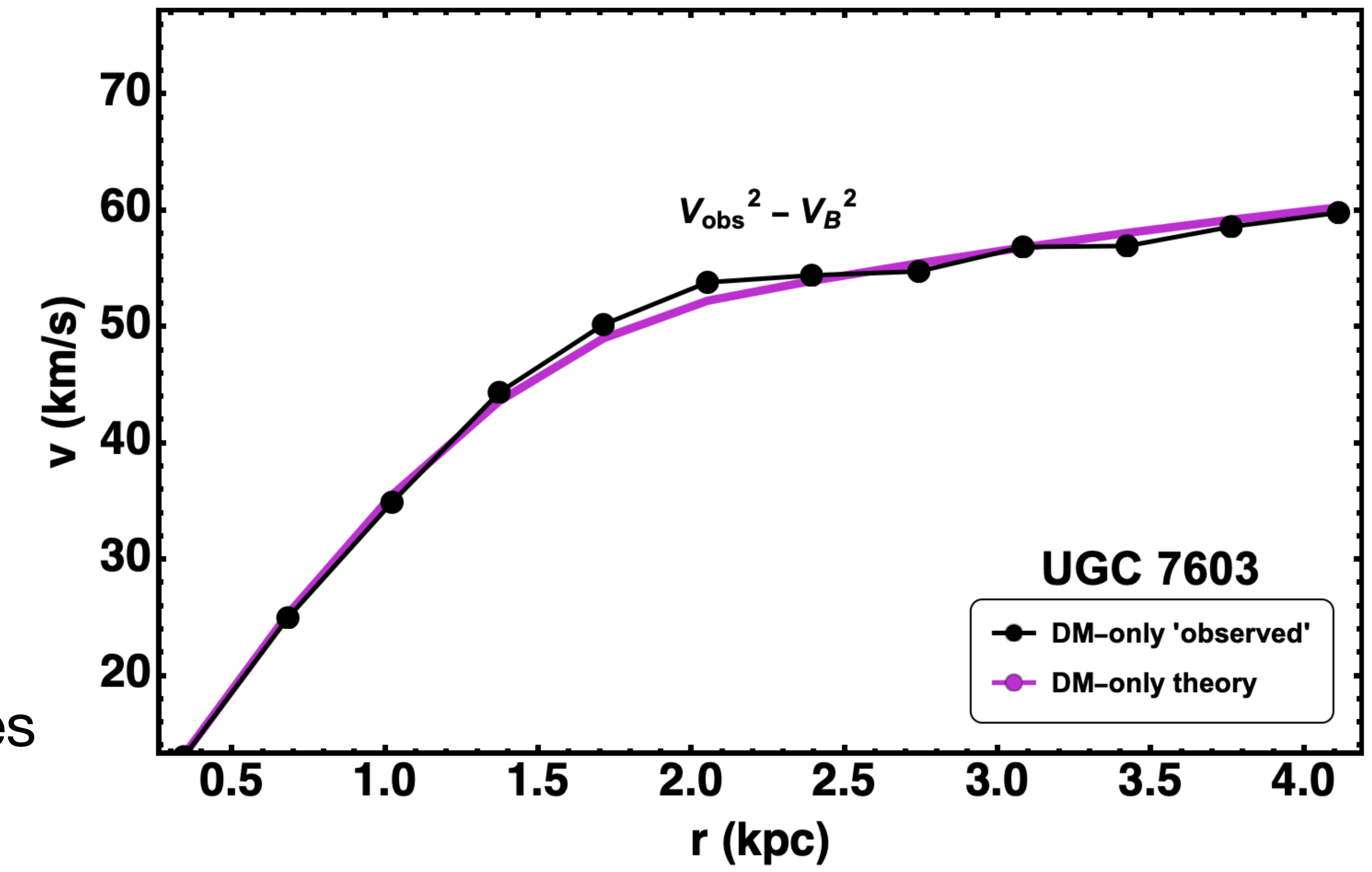




"observed" DM contribution is not smooth, and not seen by the network before...



parameters predicted for smoother curves do fine...

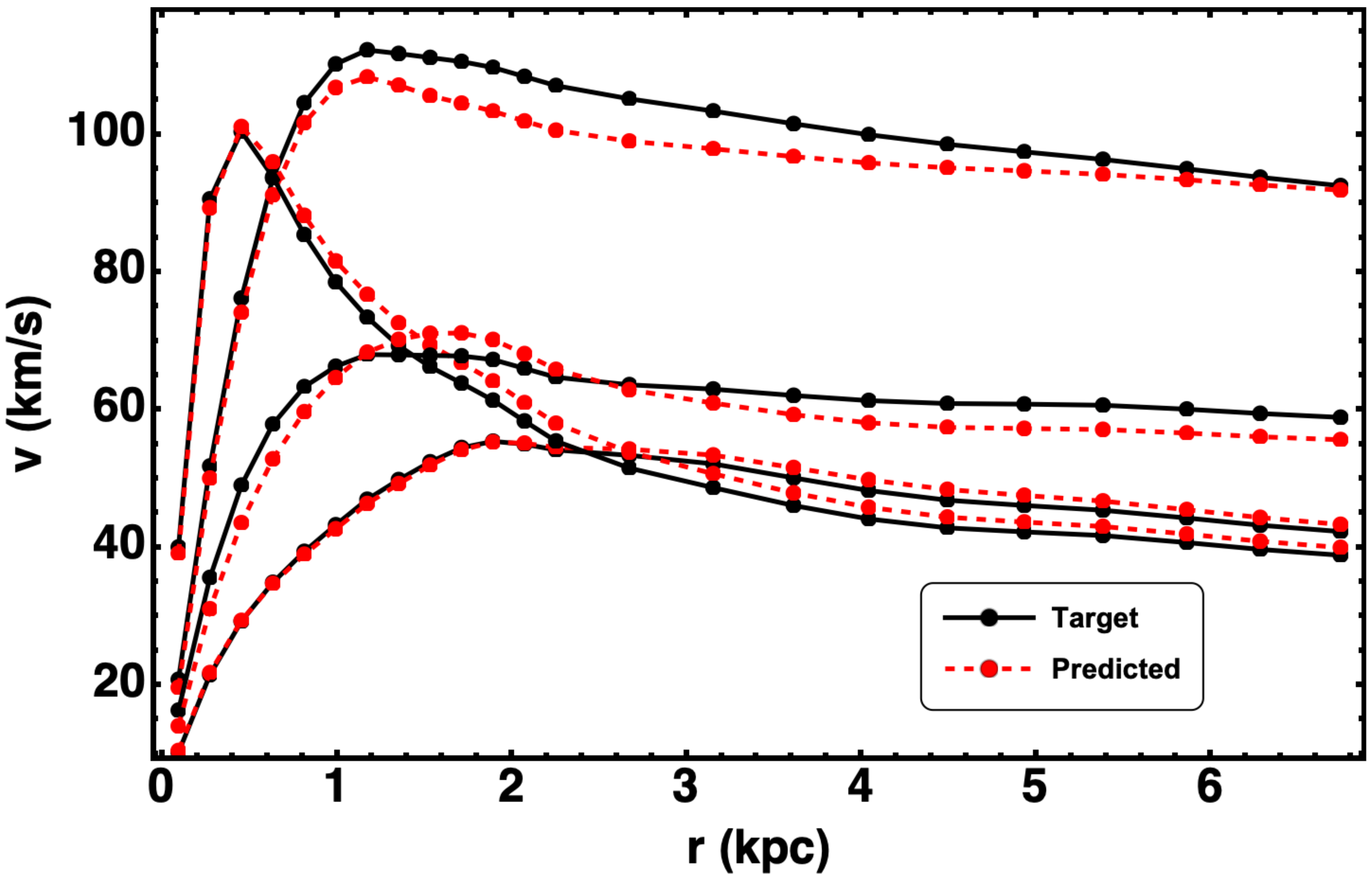


Adding noise

Since observed rotation curves have errors associated with them, i.e. $v_{obs}(r_k) \pm \sigma_{obs}(r_k)$

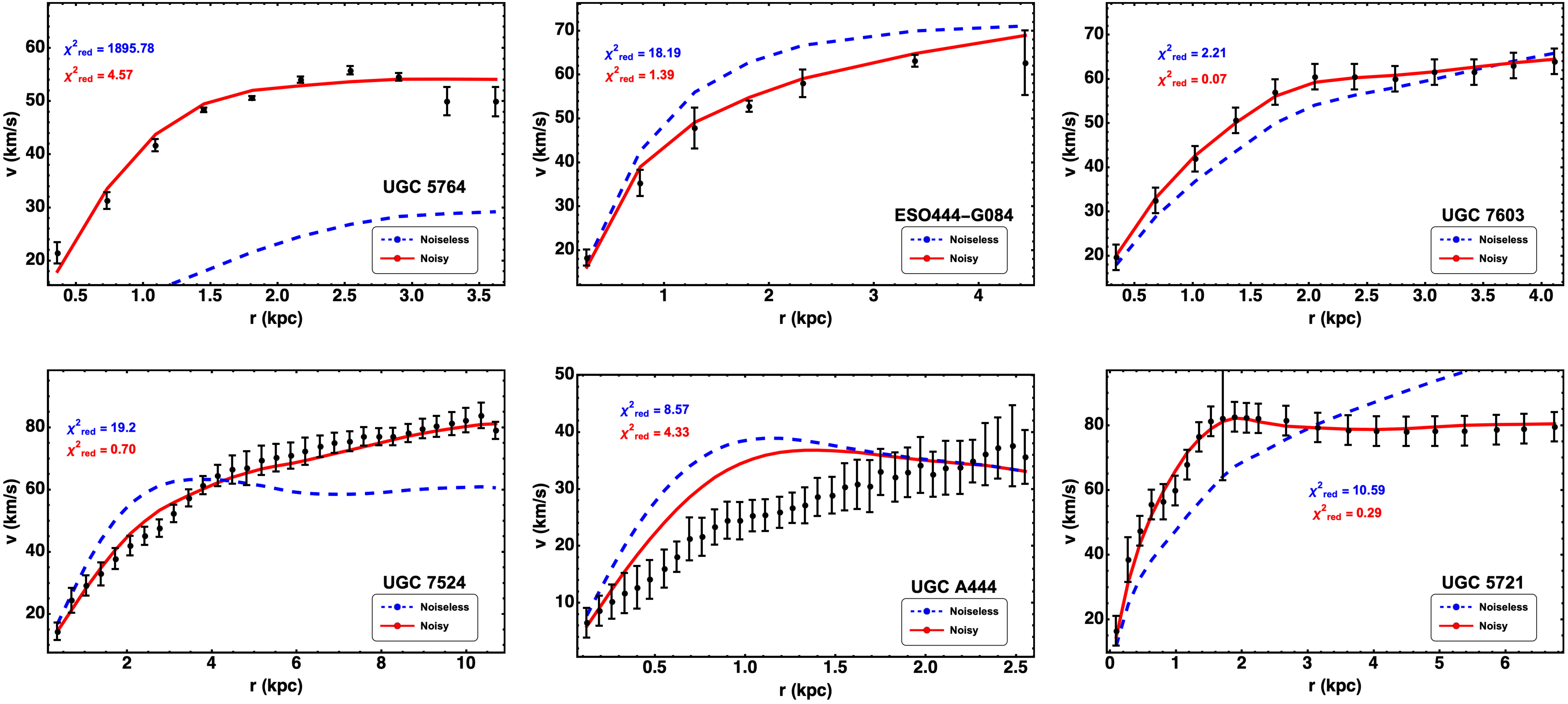
$$v_{sim}(r_k) \to v_{sim}(r_k) + \text{Err.}$$

Err. is a random draw from $\mathcal{N}(0,\ \sigma^2_{obs}(r_k))$



Performs just as well on noisy test data

With noisy inputs



B. Dave *and G. Goswami, arXiv:2412.03547 [astro-ph.CO]*

Multiple realizations of observed rotation curves

Observed rotation velocity for a galaxy at radius r_k:

$$v_{obs}(r_k) \pm \sigma_{obs}(r_k)$$

To simulate multiple observations, we draw from

$$\mathcal{N}(v_{obs}(r_k), \sigma^2_{obs}(r_k))$$

1000 draws → 1000 observations

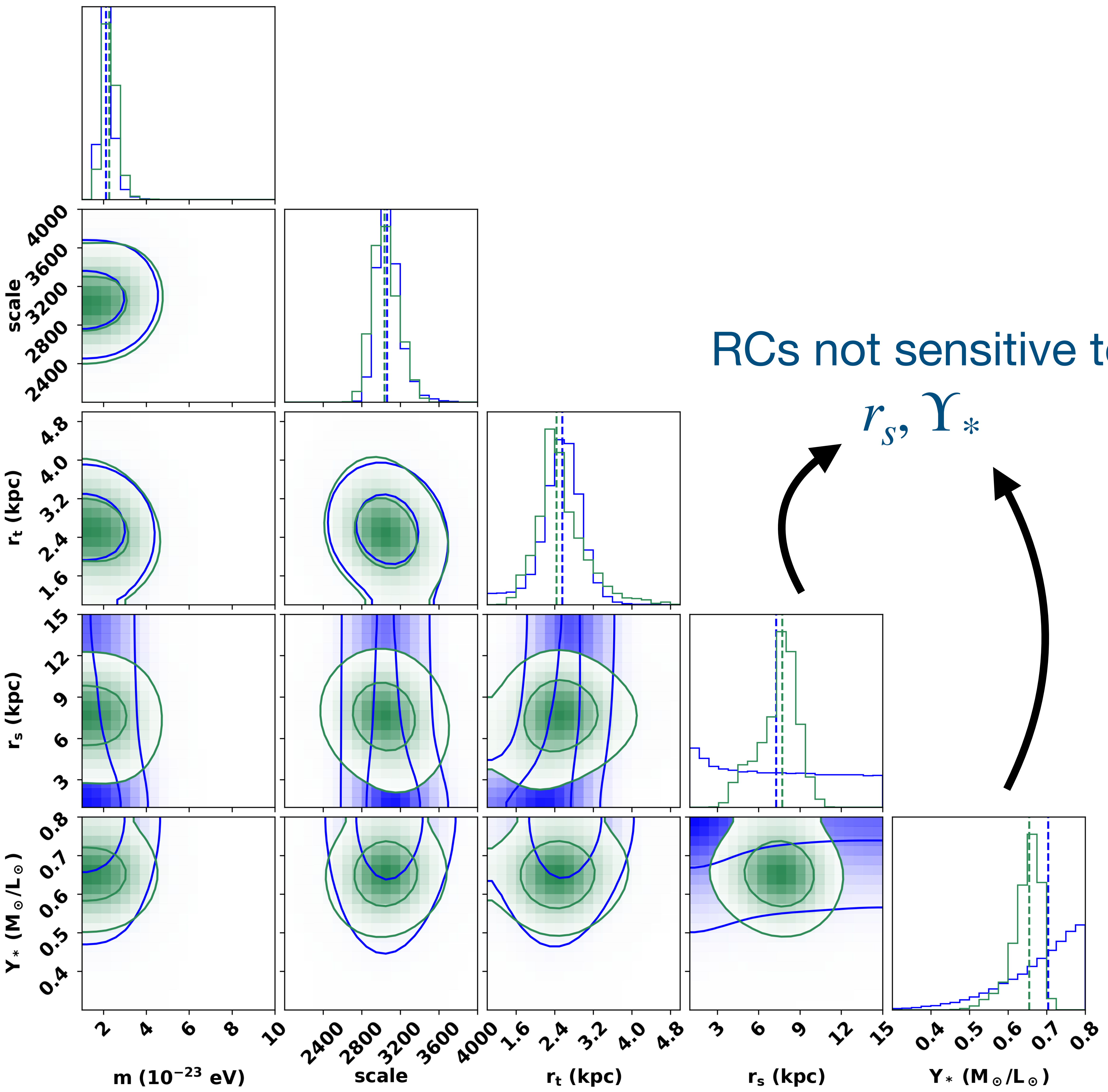
↓

Make predictions for each observation
(trained model with MSE loss)

One can think of this as sampling curves from the vicinity of the observed rotation curve while being consistent with observations.

Similar approach (for cosmological parameters) in G. J. Wang et al. (2020)

Comparison with MCMC



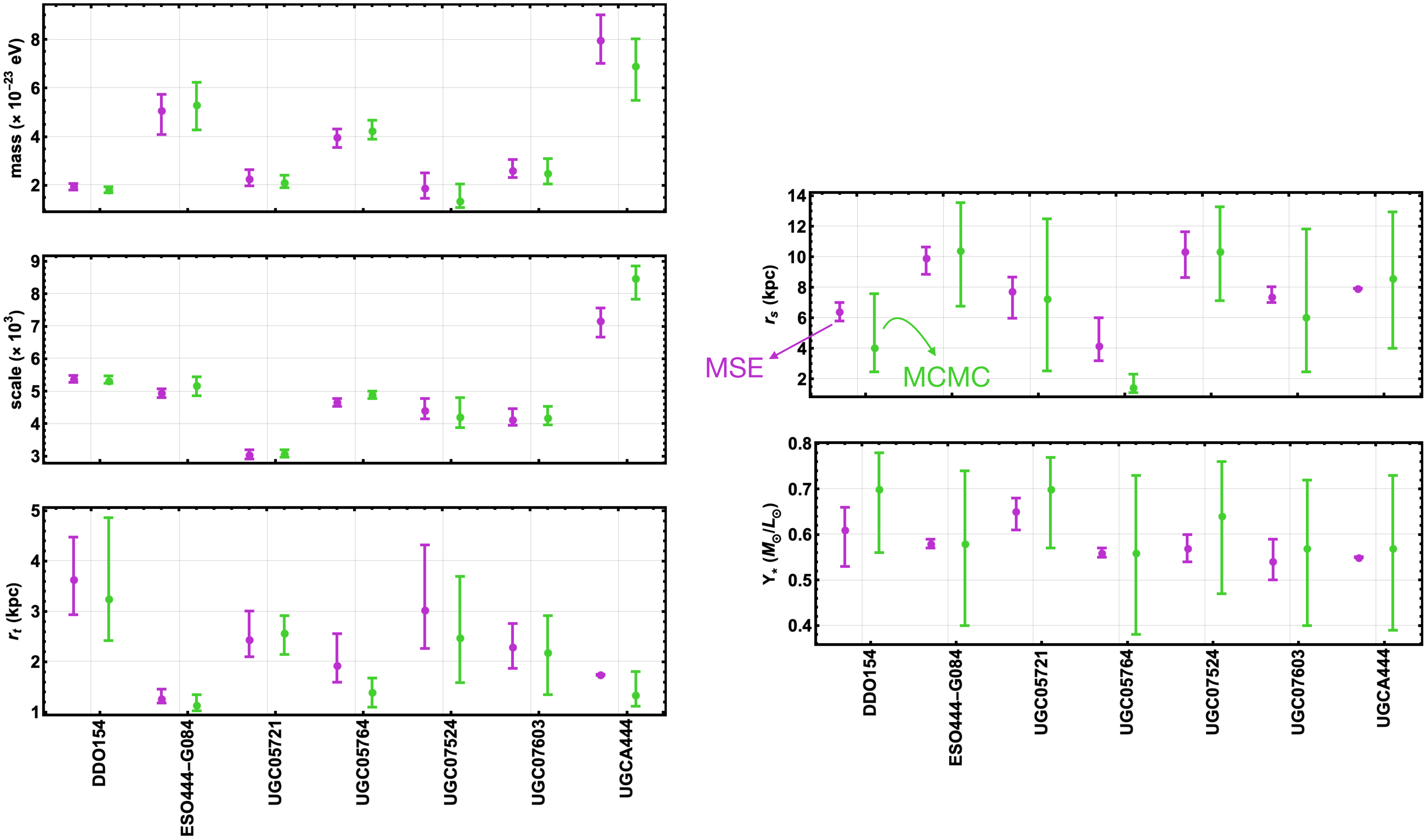
RCs not sensitive to r_s, Υ_*

- The contours obtained using MSE+MR captures the appropriate posteriors for m, s, r_t
- MSE+MR fails for 'hard to learn' parameters r_s, Υ_*
- Predicted uncertainties too small compared to an MCMC approach.

Need a different approach...

B. Dave and G. Goswami, arXiv:2412.03547 [astro-ph.CO]

MSE vs. MCMC



B. Dave *and G. Goswami, arXiv:2412.03547 [astro-ph.CO]*

'Learning' uncertainties

Simulated training data does not have errors in the outputs (i.e. parameters)

Heteroscedastic loss function allows the model to learn uncertainty implicitly!

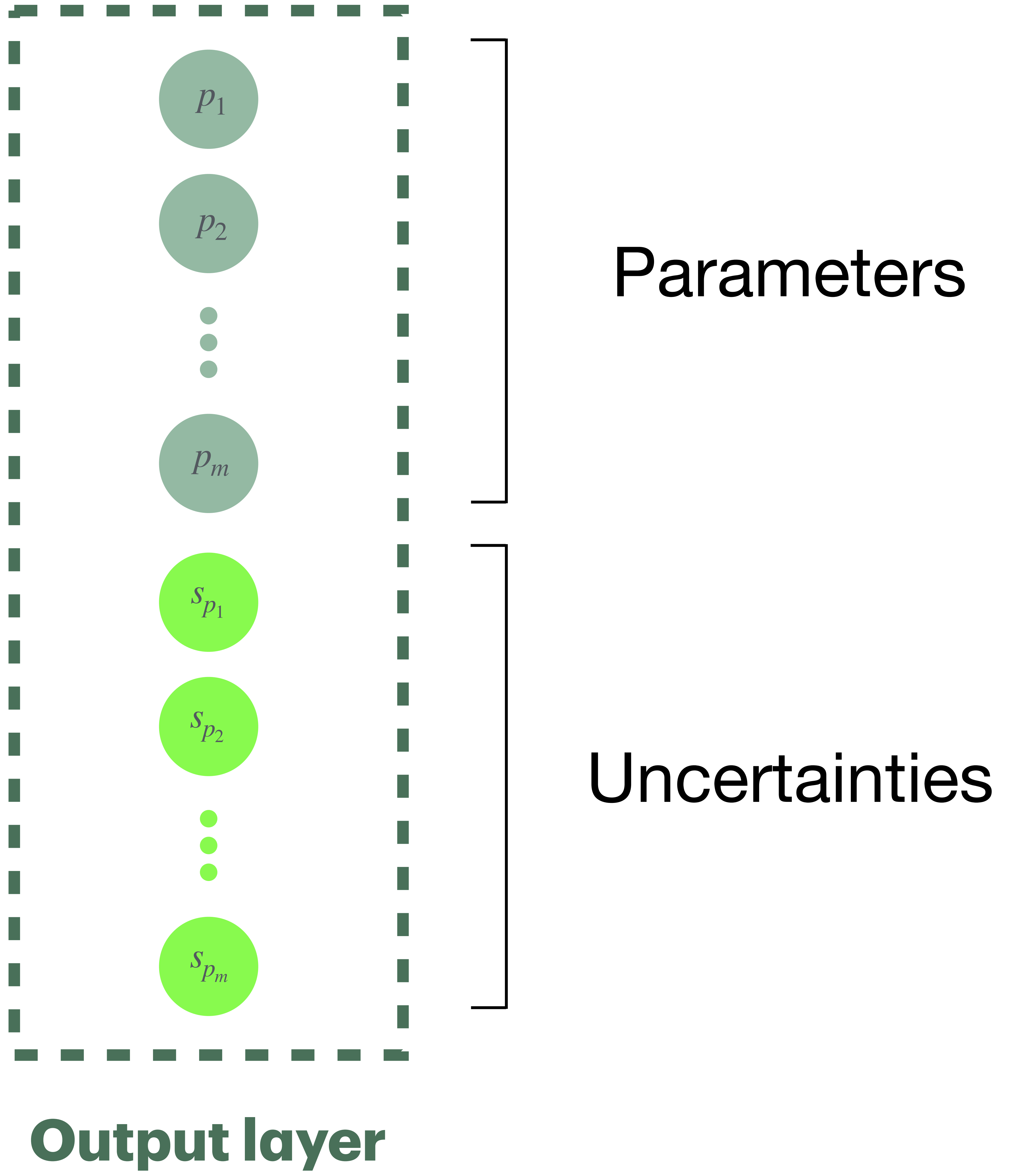
Ground truth ANN o/p

$$\mathscr{L}_{HS} = \frac{1}{2np}\sum_{i=1}^{n}\left(\sum_{j=1}^{p}\left[\frac{(y_{ij}-f_{ij})^2}{\sigma_{ij}^2} + \log\left(\frac{1}{\sqrt{2\pi\sigma_{ij}^2}}\right)\right]\right)$$

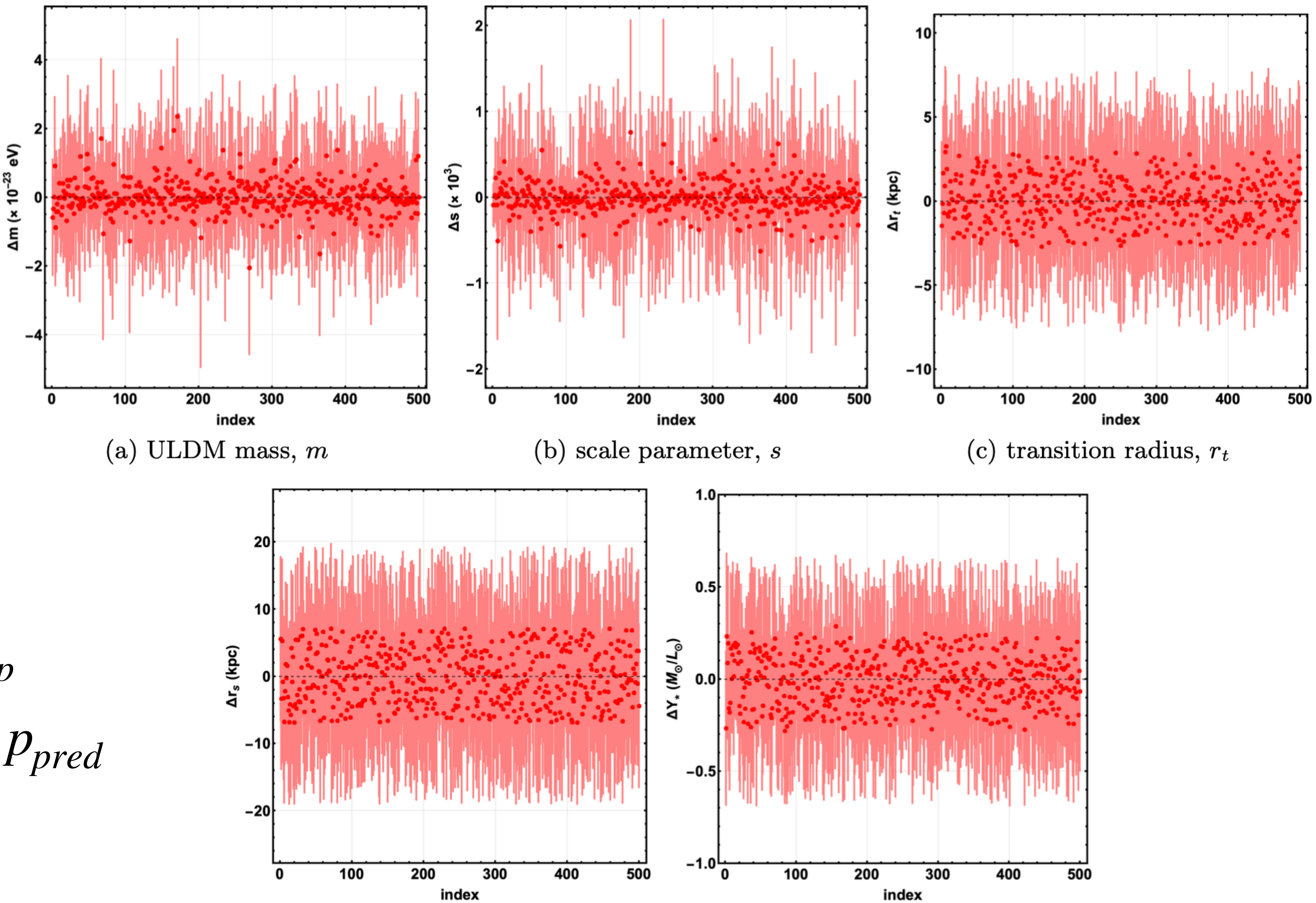
$$s_{ij} = \log \sigma_{ij}^2$$

$$\mathscr{L}_{HS} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2p}\sum_{j=1}^{p}\left(e^{-s_{ij}}(y_{ij}-f_{ij})^2 + s_{ij}\right)\right]$$

A. Kendall and Y. Gal *NIPS'17* (2017), 5580-5590



'Learning' uncertainties



(a) ULDM mass, m

(b) scale parameter, s

(c) transition radius, r_t

(d) scale radius, r_s

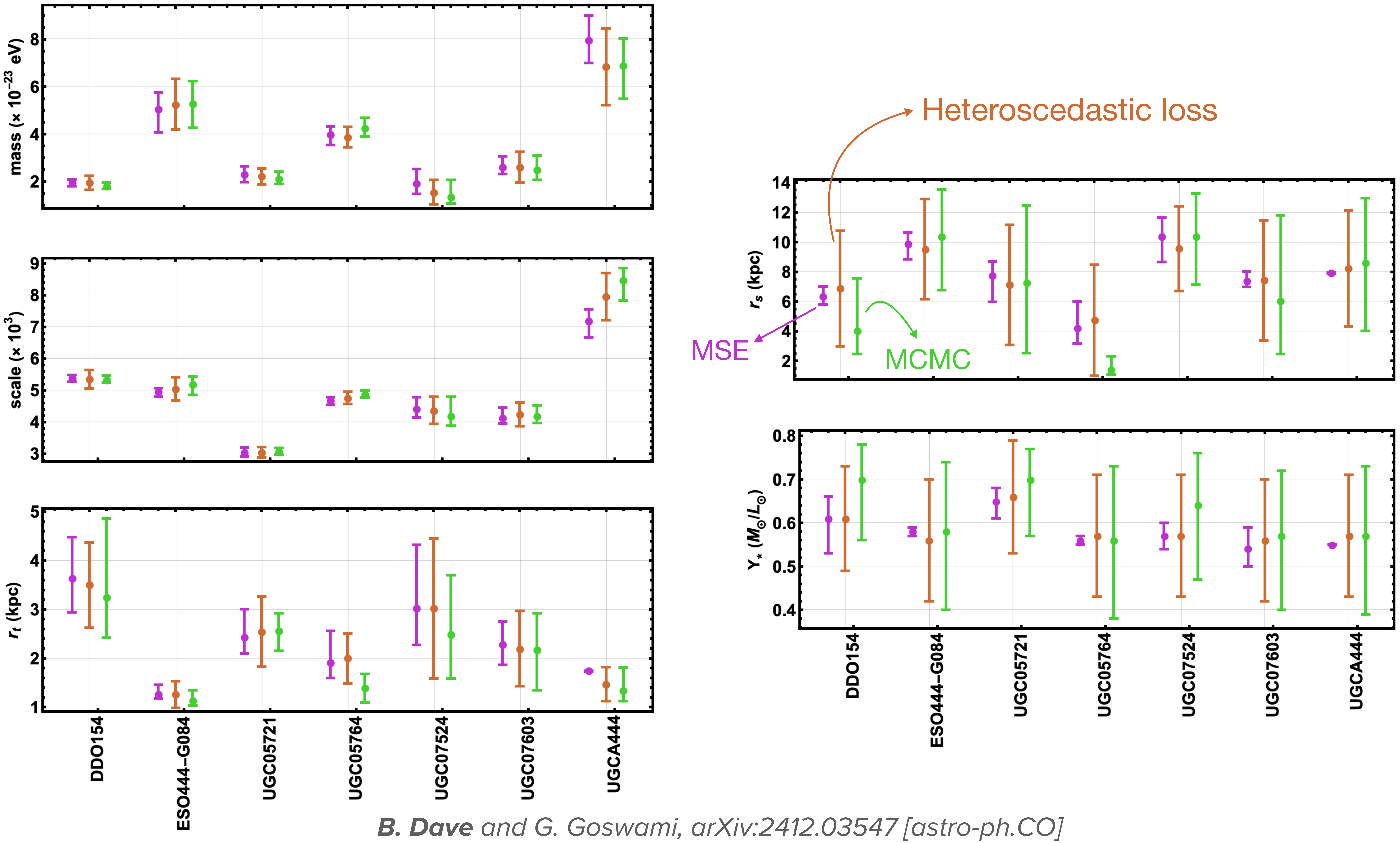
(e) stellar mass-to-light ratio, Υ_*

$$\Delta p \pm 3\sigma_p$$

$$\Delta p = p_{true} - p_{pred}$$

When given observations

MSE vs. Heteroscedastic vs. MCMC



***B. Dave** and G. Goswami, arXiv:2412.03547 [astro-ph.CO]*

Constraining other models

Recall

$$V_{obs}^2 > V_{baryons}^2$$

It is possible that[1, 2]

Modified gravity

$$V_{tot} = V_{baryons}^2 + V_{DM}^2 + V_{mg}^2$$

Neural networks can be used to learn parameters from rotation curves for such a model as well!

Velocity curve for Renormalisation Group correction to General Relativity (RGGR)[2] model (w/o DM)

$$V_{RGGR}^2 = V_{baryons}^2 \left(1 - \frac{c^2 \bar{\nu}}{\Phi_{baryons}}\right) \stackrel{?}{=} V_{obs}^2$$

RGGR contribution can be tuned using $\bar{\nu}$

$$\Phi_{baryons} = \Upsilon_d \Phi_{disk} + \Upsilon_b \Phi_{bulge} + \Phi_{gas}$$

$$V_{baryons}^2 = \Upsilon_d V_{disk}^2 + \Upsilon_b V_{bulge}^2 + |V_{gas}| V_{gas}$$

Parameters: $\{\Upsilon_d, \Upsilon_b, \bar{\nu}\}$

Work ongoing!

1. A. Almeida et al. JCAP 08 (2018) 012
2. E. Bhatia et al. arXiv:2403.00531 [gr-qc]

Thank you