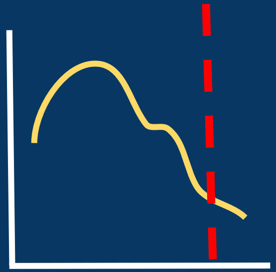
Learning Cosmology with Graph Neural Networks

Natalí S. M. de Santi

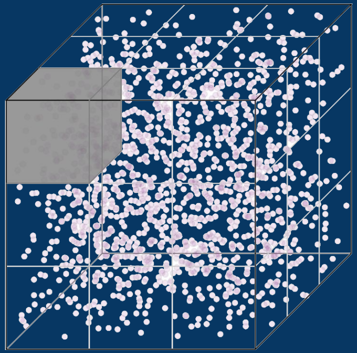
Pos-doc @BCCP

October 7th, 2025

Is there an optimal way to infer the cosmological parameters?

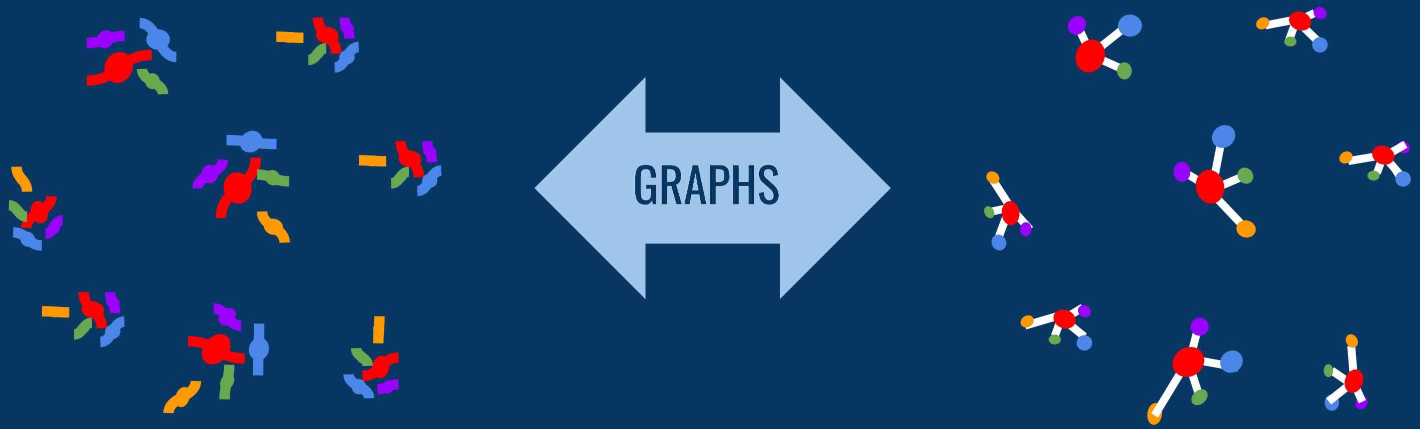


Traditional Methods

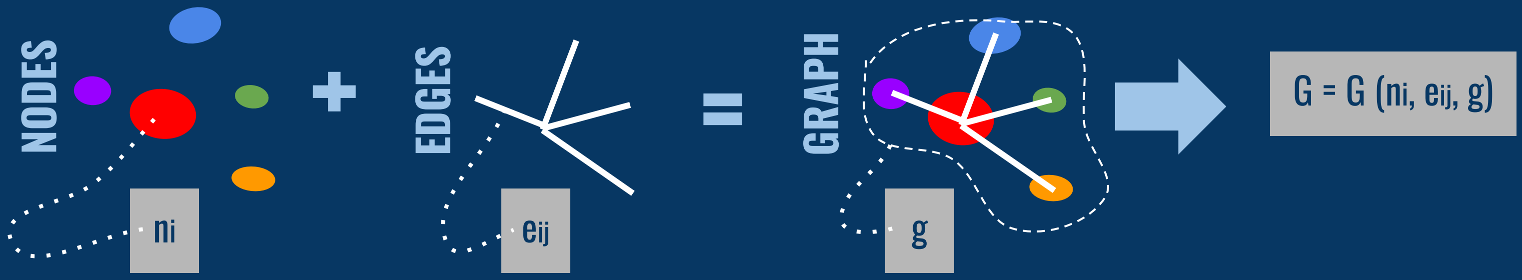


SBI with CNNs

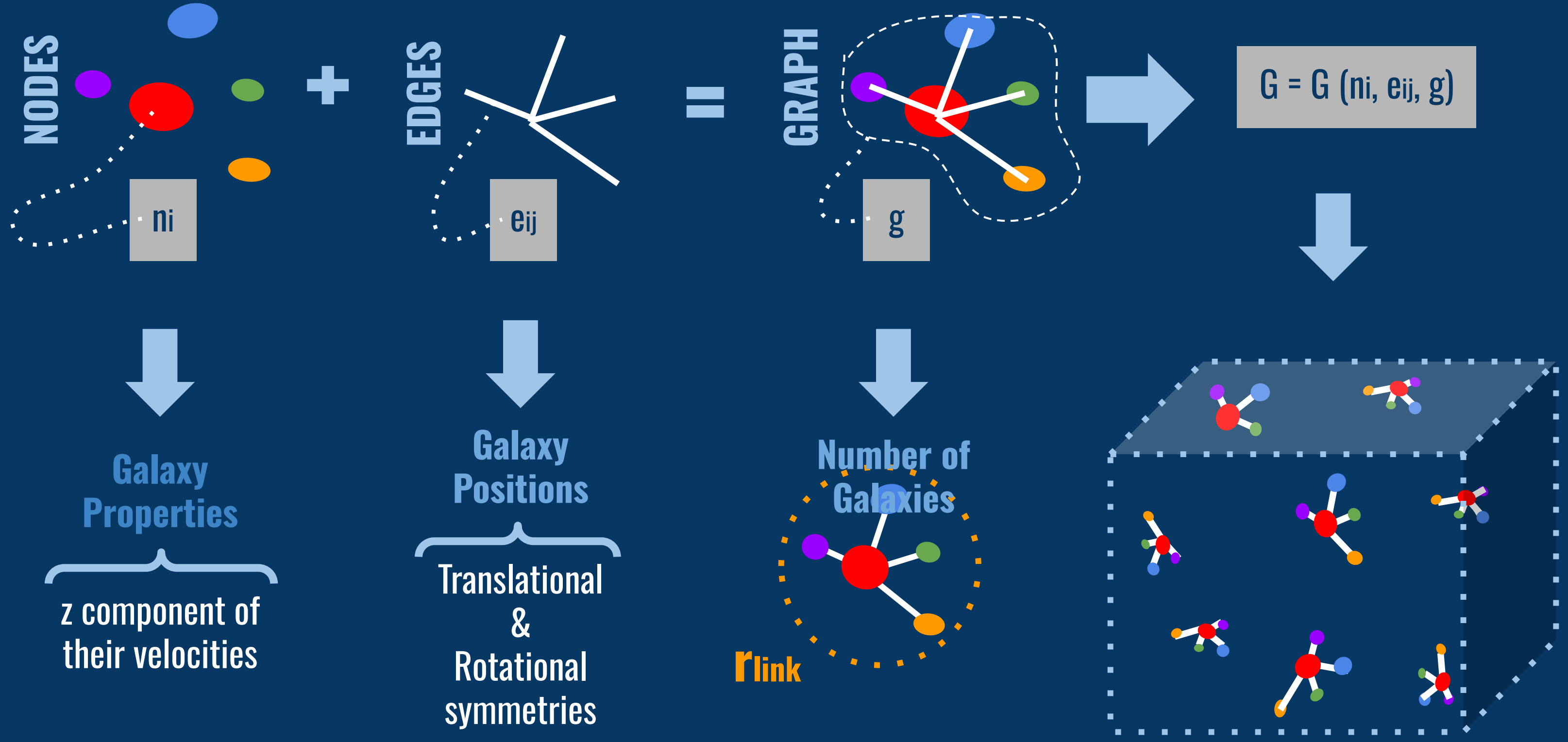
Is there a mathematical structure that naturally represents the cosmic-web?



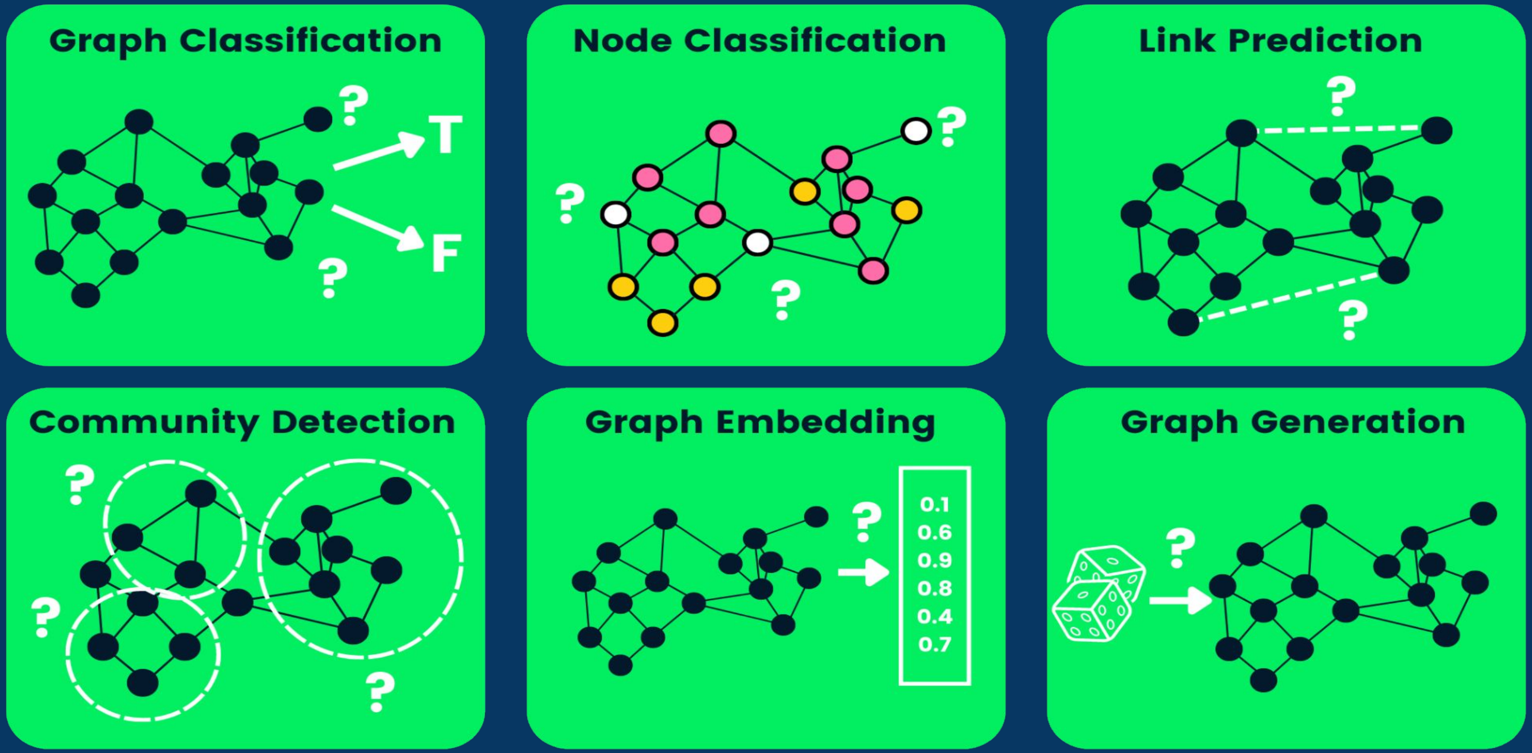
What are graphs?



How can galaxy catalogs be translated into graphs?

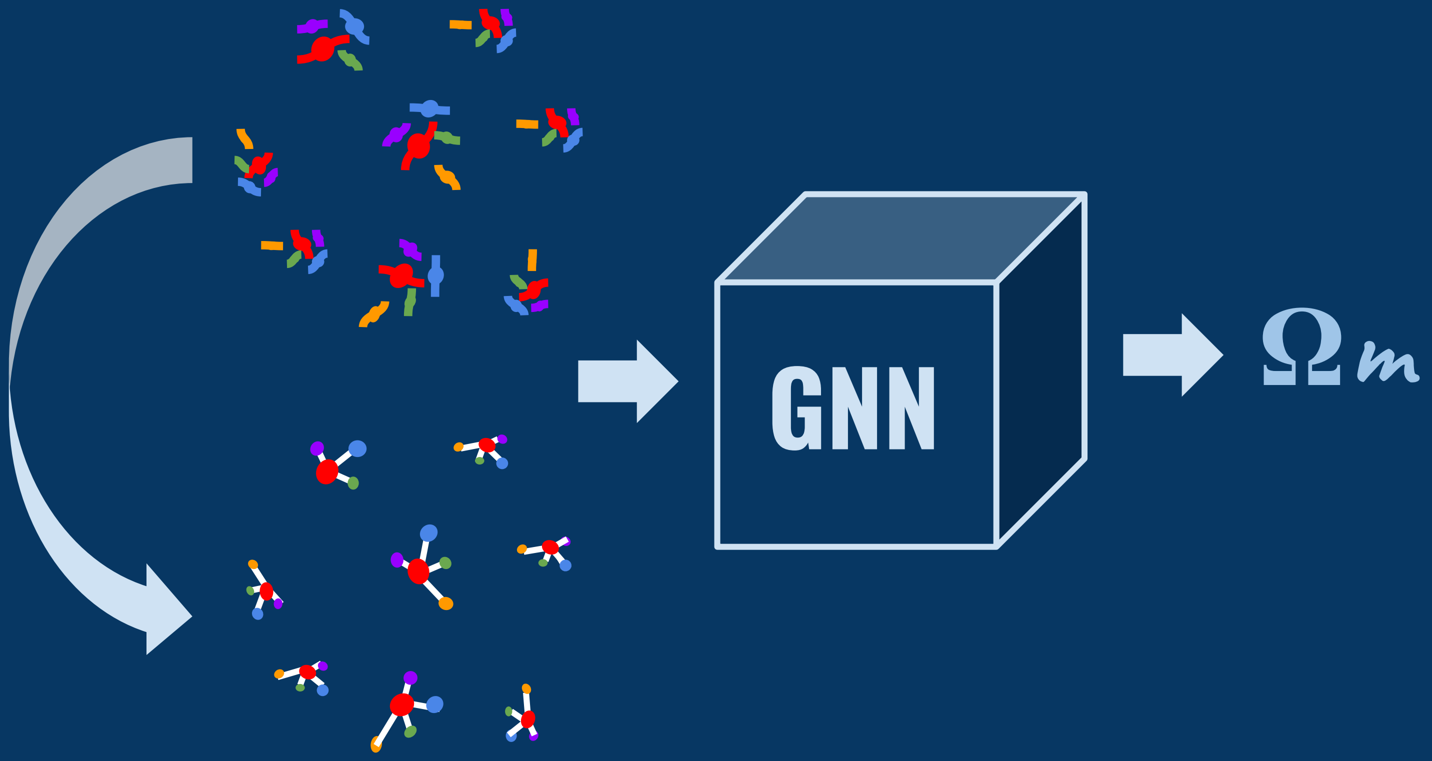


What can we use graphs for?

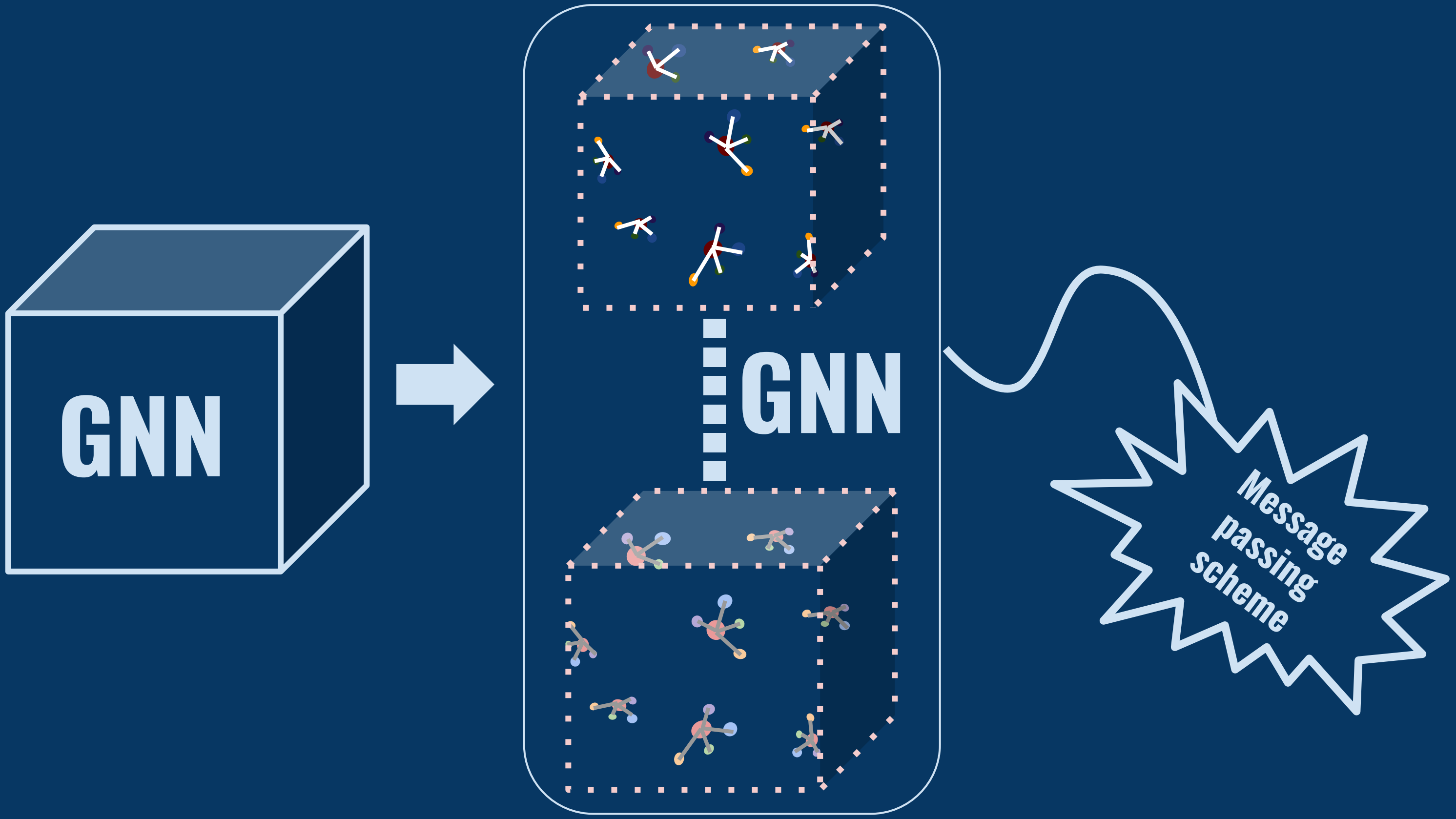


https://images.datacamp.com/image/upload/v1658404112/Types_of_Graph_Neural_Networks_fd300394e8.png

How can we learn cosmology with graph neural networks?



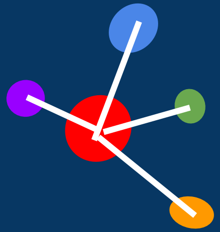
What is the mechanism behind the scenes?



Different Graph Neural Network Architectures

PyTorch Geometric

ni



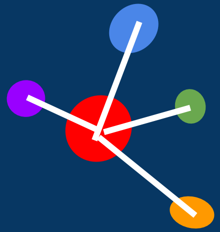
GRAPH CONVOLUTIONAL NETWORKS

GRAPH SAGE NETWORKS

ni

eij

g



GRAPH NEURAL NETWORKS FOR COSMOLOGY

Different Graph Neural Network Architectures

Graph Convolutional Networks

Designed to perform a convolution operation (matrix multiplication) on graphs using node attributes

Graph Convolutional Network Layer

$$\mathbf{n}_i^{\ell+1} = \sum_{j \in \mathfrak{N}_{(i)} \cup i} \frac{1}{\sqrt{\deg(i)} \cdot \sqrt{\deg(j)}} \cdot \underbrace{\left[\mathbf{W}^T \cdot \mathbf{n}_j^{(\ell)}\right]}_{\text{Learnable weight matrix}} + \underbrace{\mathbf{b}}_{\text{Additive bias}}$$

Edge Convolutional Layer

$$\mathbf{n}_i^{(\ell+1)} = \underbrace{\max}_{\text{Aggregation Function}}{}_{j \in \mathfrak{N}_j} \underbrace{\mathcal{H}^{(\ell+1)}}_{\text{MLP}} \left(\underbrace{\left[\mathbf{n}_i^{\ell}, \mathbf{n}_j^{\ell} - \mathbf{n}_i^{\ell}\right]}_{\text{"Edge" information}} \right)$$

Different Graph Neural Network Architectures

Graph Sage networks

Designed to update node information on graphs based on the neighborhood of each node

SAGE Convolutional Layer

$$\mathbf{n}_i^{\ell+1} = \mathbf{W}_1 \mathbf{n}_i^{\ell} + \mathbf{W}_2 \cdot \mathbf{mean}_{j \in \mathfrak{N}_i} \mathbf{n}_j^{\ell}$$

Learnable
weight matrices

Different Graph Neural Network Architectures

Graph Neural Networks for Cosmology

Meta Layer

Edge model

$$\mathbf{e}_{ij}^{(\ell+1)} = \underbrace{\mathcal{E}^{(\ell+1)}}_{\text{MLP}}\left(\left[\mathbf{n}_i^{(\ell)}, \mathbf{n}_j^{(\ell)}, \mathbf{e}_{ij}^{(\ell)}\right]\right)$$

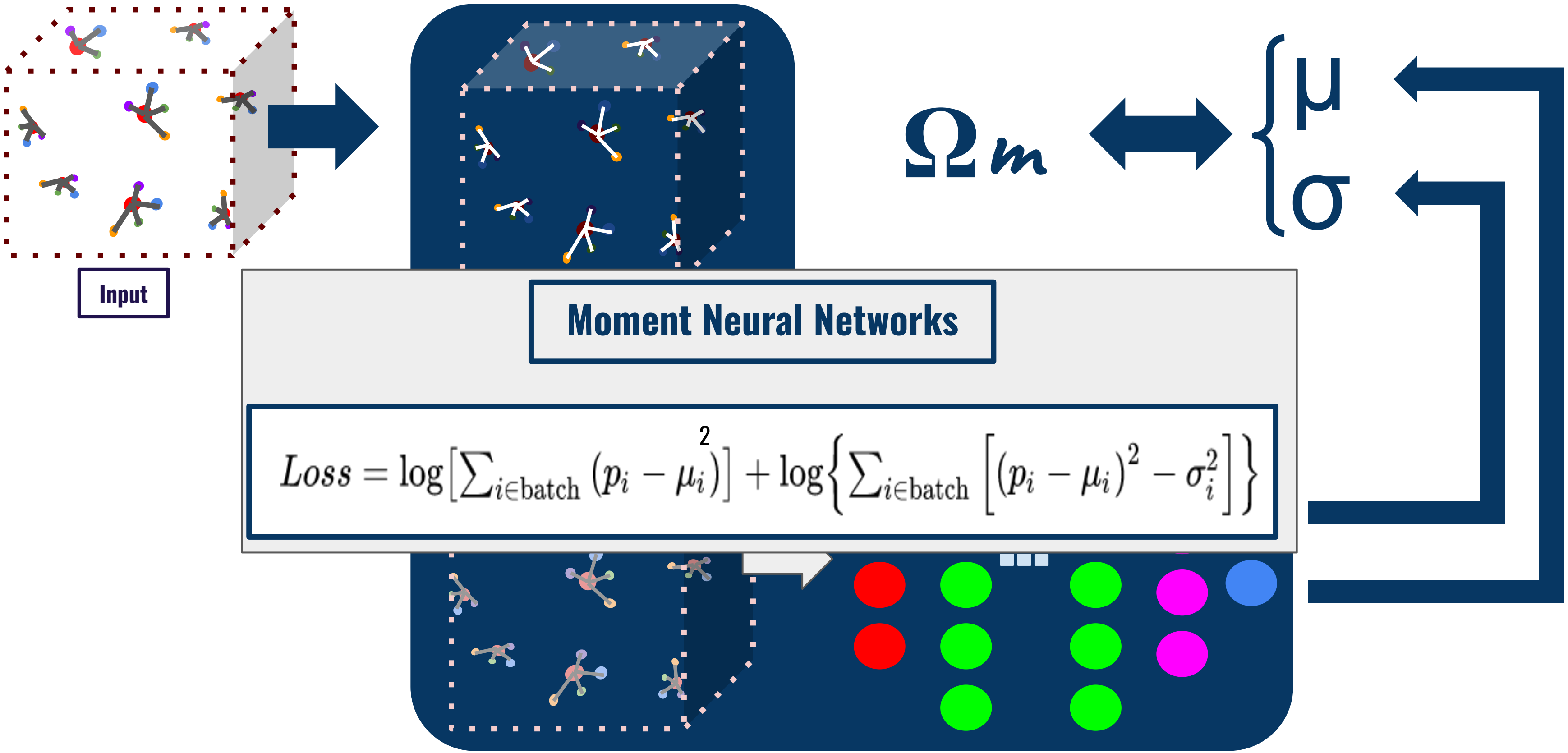
Node model

$$\mathbf{n}_i^{(\ell+1)} = \underbrace{\mathcal{N}^{(\ell+1)}}_{\text{MLP}}\left(\left[\mathbf{n}_i^{(\ell)}, \bigoplus_{j\in\mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)}, \mathbf{g}\right]\right)$$

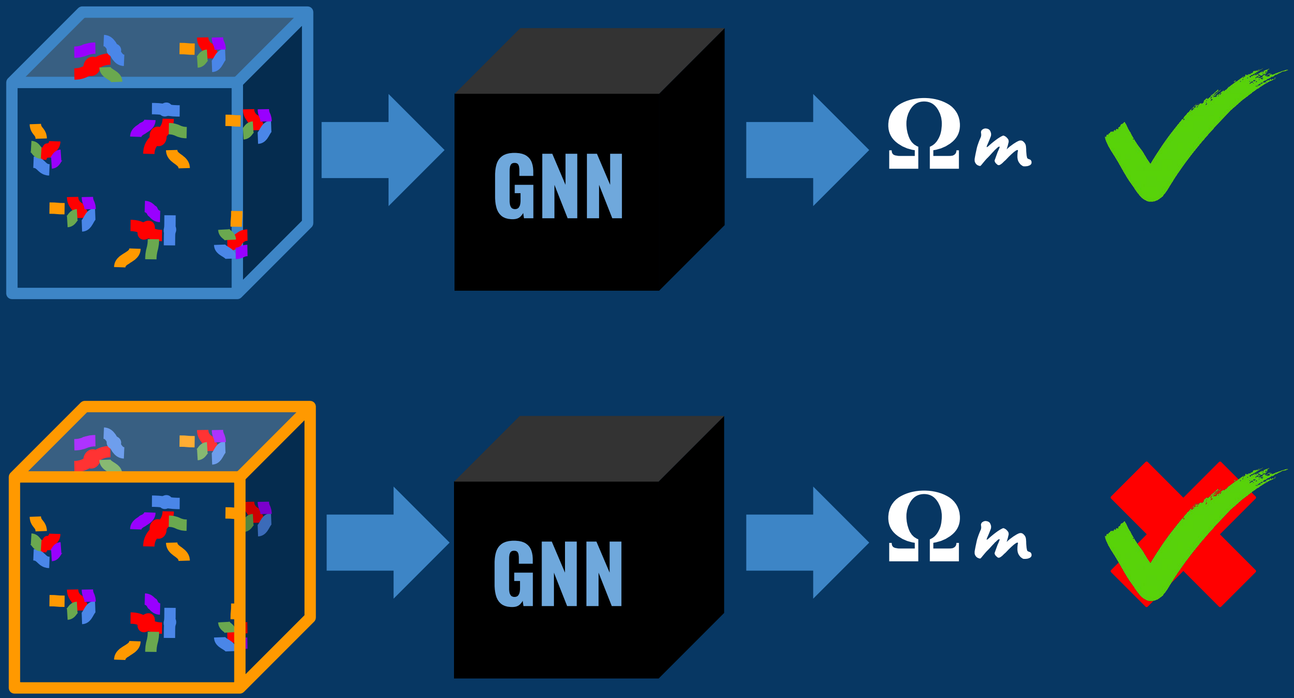
Multi Pooling Operation

$$\bigoplus_{j\in\mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)} = \left[\max_{j\in\mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)}, \sum_{j\in\mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)}, \frac{\sum_{j\in\mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)}}{\sum_{j\in\mathfrak{N}_i}}\right]$$

Graph Neural Networks for Cosmology

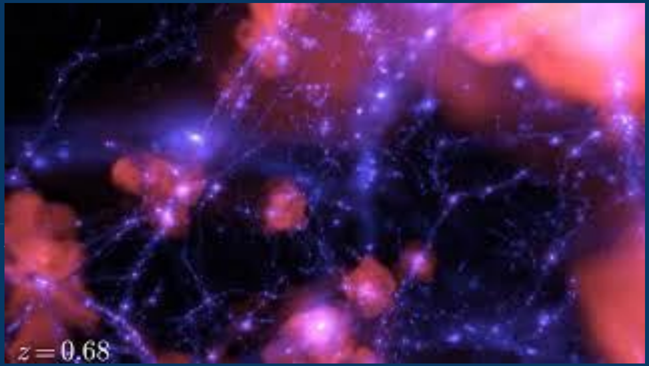


Can we build a robust model?



Why do we need a robust model?

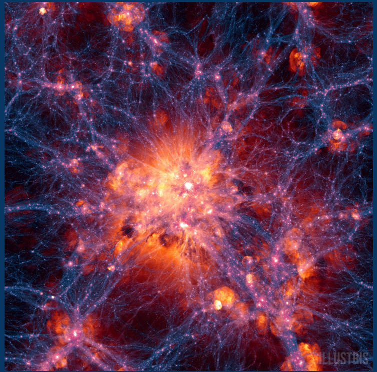
[Webb's First Deep Field]



[IllustrisTNG evolution]

Ultimate goal!

First step: Hydrodynamical Simulations

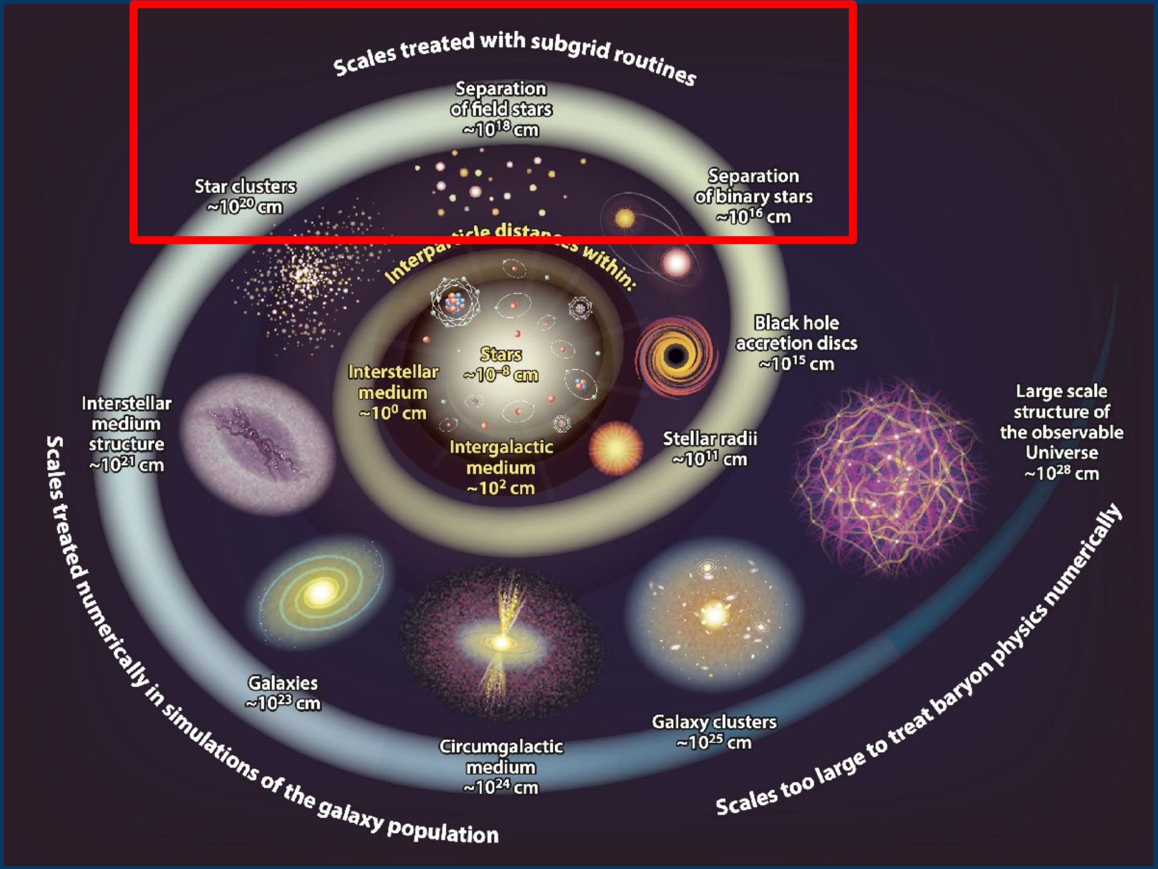


Illustris: Dark Matter + gas ⇋ halos and galaxies

Lagrangian formulation

$$\frac{D\rho}{Dt} = -\rho\nabla\cdot\mathbf{v},$$
$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho}\nabla P,$$
$$\frac{De}{Dt} = \frac{1}{\rho}\nabla\cdot p\mathbf{v},$$

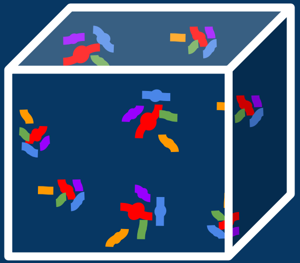
Scheme of the range of scales into galaxy formation and evolution: arXiv: 2309.17075

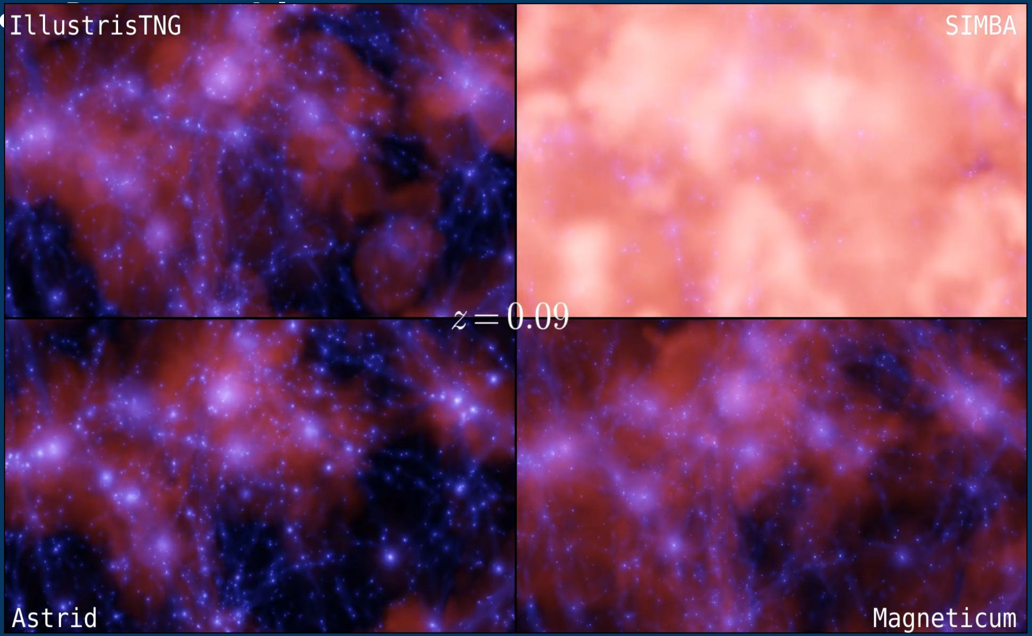


Subgrid physical processes

First step: Hydrodynamical Simulations

CAMELS





- **Parameters:**
  - **Cosmological:** Ω_m, σ_8
  - **Astrophysical:** A_{SN1}, A_{SN2}, A_{AGN1}, A_{AGN2}

- **Sets:**

LH
Vary all parameters (6)
Different random seeds

→ Train & Test

SIMBA

SB28

SWIFT-EAGLE

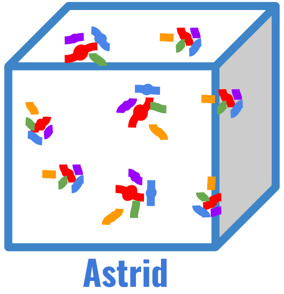
CAMELS visualizations

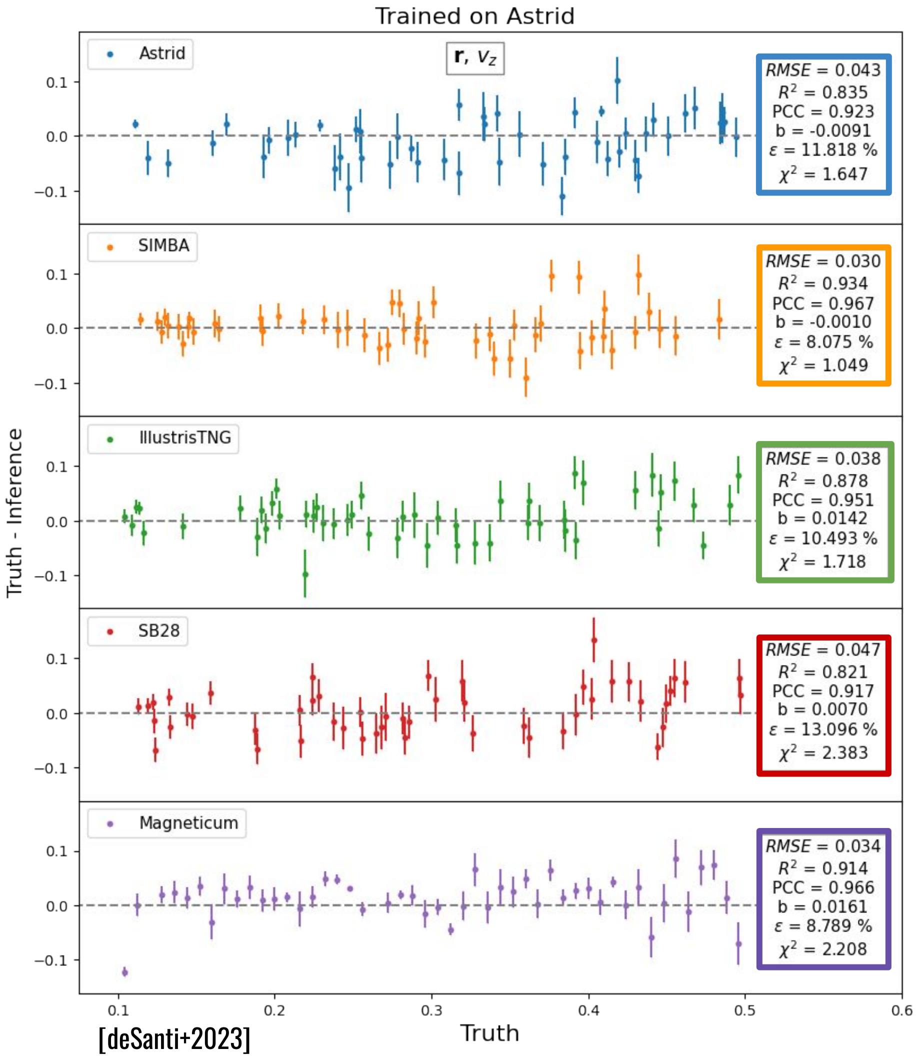
Vary 28 parameters
Different random seeds

Uses VELOCIraptor

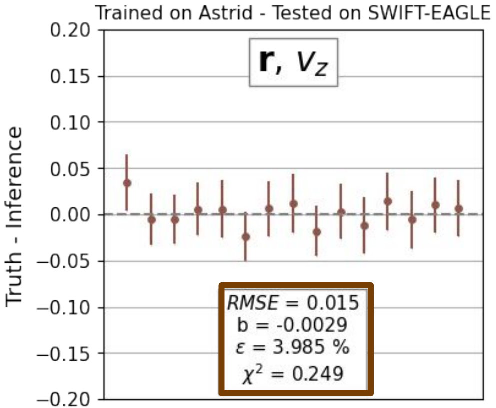
Galaxy Information

3D Positions
z component of the velocity

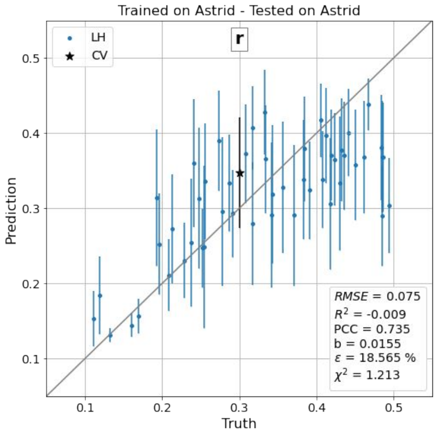


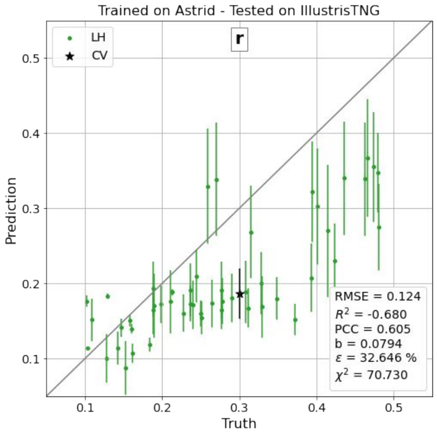


[deSanti+2023]



Where does the information come from?





Positions

Velocities

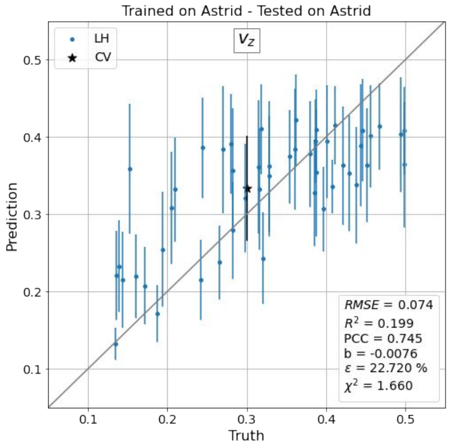
GNN without node information

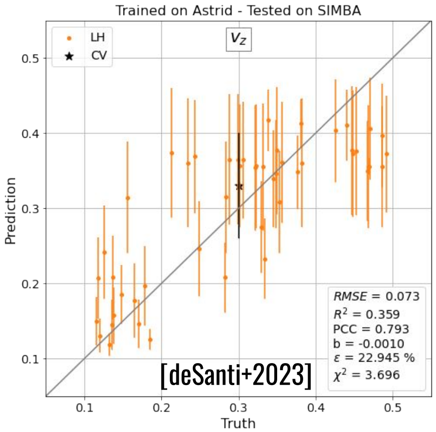
We need both!

Deep Set

Sanity check — Error bars increased

Robustness — Breaks down

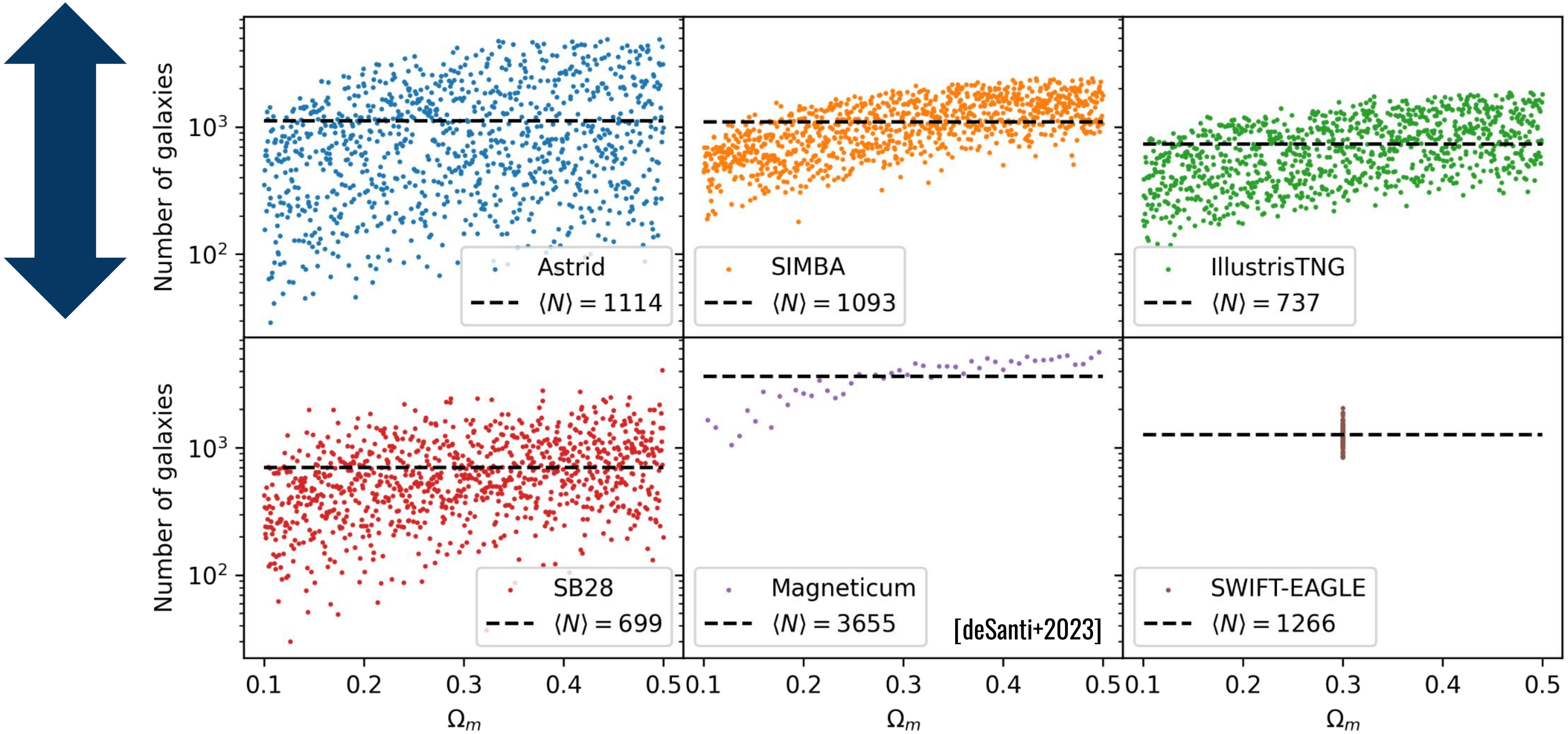




[deSanti+2023]

Where does the information come from?

It is important to train the ML algorithm on a dataset which contain broader variations in the galaxy properties: **NUMBER OF GALAXIES**

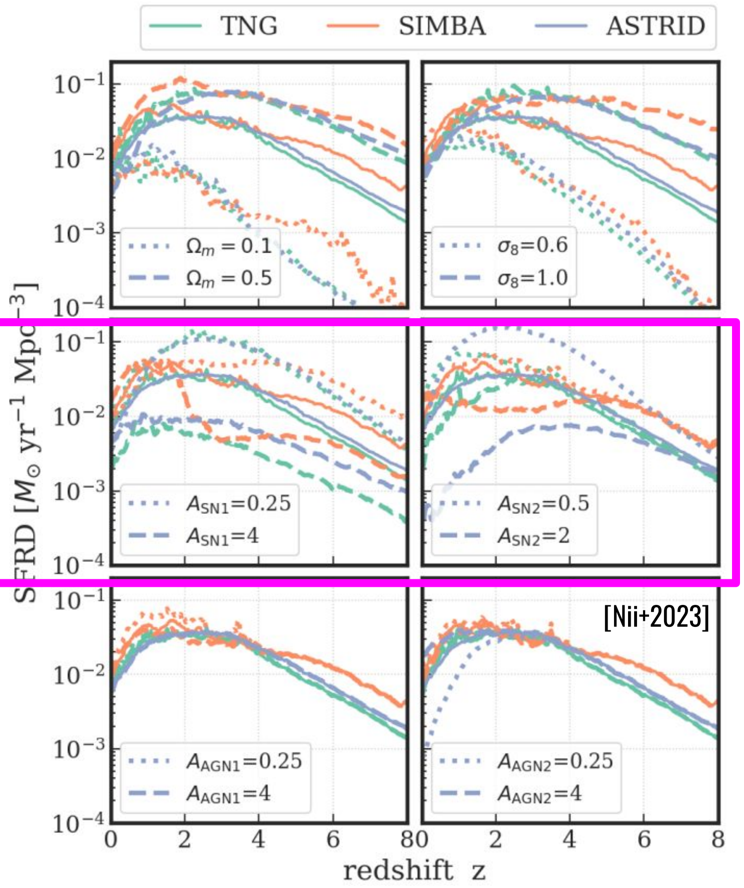


[deSanti+2023]

Where does the information come from?

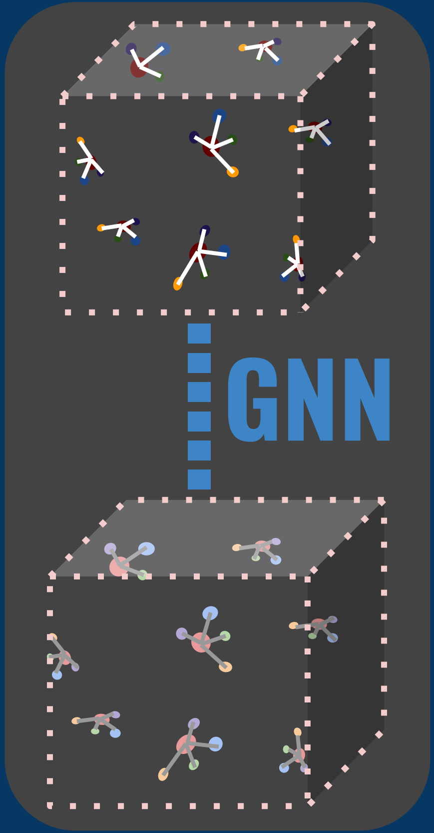
- Broader variation in the **galaxy population** in the ASTRID;
- A_{SN} parameters drive larger variations in star formation;
- A_{SN2} modulates the speed of galactic winds;
- **Star formation** in the ASTRID model turns out to be more sensitive to the SN wind speed than in TNG and SIMBA

Broader variation
SN wind speed



[Nii+2023]

Second step: Is there an universal equation to translate the GNNs?



Edge Model

$$\mathbf{e}_{ij}^{(\ell+1)} = \mathcal{E}^{(\ell+1)}\left(\left[\mathbf{n}_i^{(\ell)}, \mathbf{n}_j^{(\ell)}, \mathbf{e}_{ij}^{(\ell)}\right]\right)$$

→ Edge Equation

Node Model

$$\mathbf{n}_i^{(\ell+1)} = \mathcal{N}^{(\ell+1)}\left(\left[\mathbf{n}_i^{(\ell)}, \bigoplus_{j \in \mathfrak{N}_i} \mathbf{e}_{ij}^{(\ell+1)}, \mathbf{g}\right]\right)$$

→ Node Equation

Final MLP

$$\mathbf{y} = \mathcal{F}\left(\left[\bigoplus_{i \in \mathfrak{F}} \mathbf{n}_i^{N}, \mathbf{g}\right]\right)$$

→ Final MLP Equation

Training on halos

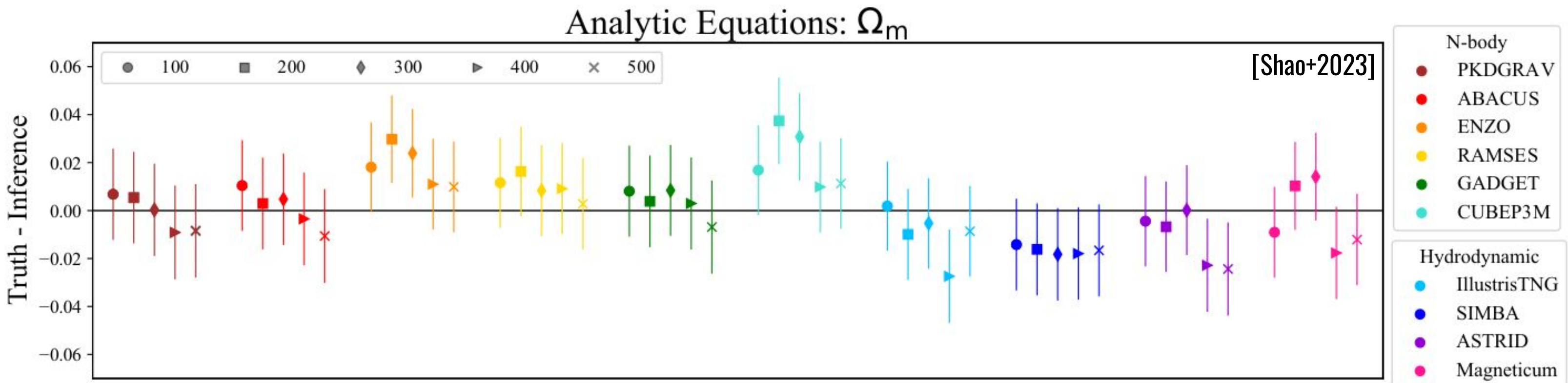
A Universal Equation to predict Ω_m from Halos & Galaxies

GNN Component	Formula	RMSE
Edge Model: $e_1^{(1)}$	$1.32\lvert v_i - v_j + 0.21\rvert + 0.12(v_i - v_j) - 0.12(\gamma_{ij} + \beta_{ij} - 1.73)$	0.03
Edge Model: $e_2^{(1)}$	$\lvert 1.62(v_i - v_j) + 0.45\rvert + 1.98(v_i - v_j) + 0.55$	0.04
Node Model: $v_1^{(1)}$	$1.21^{v_i}(0.77^{3.29\sum_{j\in\mathcal{N}_j} e_1^{(1)} + \sum_{j\in\mathcal{N}_j} e_2^{(1)}}) + 0.12$	0.02
Node Model: $v_1^{(1)} + v_2^{(1)}$	$0.78 - \sqrt{\log(0.16^{\sum_{j\in\mathcal{N}_j} e_2 + \sum_{j\in\mathcal{N}_j} e_1 - 0.41 v_i - 1.05})} + 1.45$	0.03
Final MLP: μ_{Ω_m}	$4\times 10^{-4}\cdot(-5.5\sum_{i\in\mathcal{G}} v_2^{(1)} + 2.21\sum_{i\in\mathcal{G}} v_1^{(1)} + \lvert 0.96\sum_{i\in\mathcal{G}} v_2^{(1)} + 0.82\sum_{i\in\mathcal{G}} v_1^{(1)}\rvert) - 0.103$	0.03

$$\gamma_{ij} = \underbrace{\frac{|\boldsymbol{d}_{ij}|}{r_{\mathrm{link}}}}_{\text{Density of populations}} \qquad \beta_{ij} = \underbrace{\frac{\boldsymbol{r}_i - \boldsymbol{c}}{|\boldsymbol{r}_i - \boldsymbol{c}|}}_{\text{Shape/filamentary distribution}} \cdot \frac{\boldsymbol{d}_{ij}}{|\boldsymbol{d}_{ij}|}$$

A Universal Equation to predict Ω_m from Halos & Galaxies

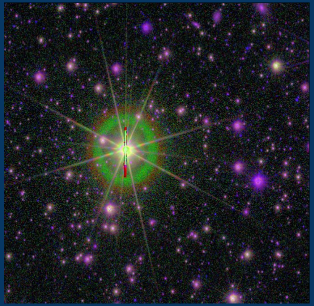
GNN Component	Formula	RMSE
Edge Model: $e_1^{(1)}$	$1.32\lvert v_i - v_j + 0.21\rvert + 0.12(v_i - v_j) - 0.12(\gamma_{ij} + \beta_{ij} - 1.73)$	0.03
Edge Model: $e_2^{(1)}$	$\lvert 1.62(v_i - v_j) + 0.45\rvert + 1.98(v_i - v_j) + 0.55$	0.04
Node Model: $v_1^{(1)}$	$1.21^{v_i}(0.77^{3.29\sum_{j\in\mathcal{N}_j} e_1^{(1)} + \sum_{j\in\mathcal{N}_j} e_2^{(1)}}) + 0.12$	0.02
Node Model: $v_1^{(1)} + v_2^{(1)}$	$0.78 - \sqrt{\log(0.16^{\sum_{j\in\mathcal{N}_j} e_2 + \sum_{j\in\mathcal{N}_j} e_1 - 0.41 v_i - 1.05})} + 1.45$	0.03
Final MLP: μ_{Ω_m}	$4\times 10^{-4}\cdot(-5.5\sum_{i\in\mathcal{G}} v_2^{(1)} + 2.21\sum_{i\in\mathcal{G}} v_1^{(1)} + \lvert 0.96\sum_{i\in\mathcal{G}} v_2^{(1)} + 0.82\sum_{i\in\mathcal{G}} v_1^{(1)}\rvert) - 0.103$	0.03



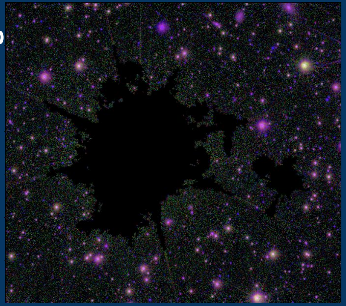
Third step: Can we take into account some systematics?

Mask effects

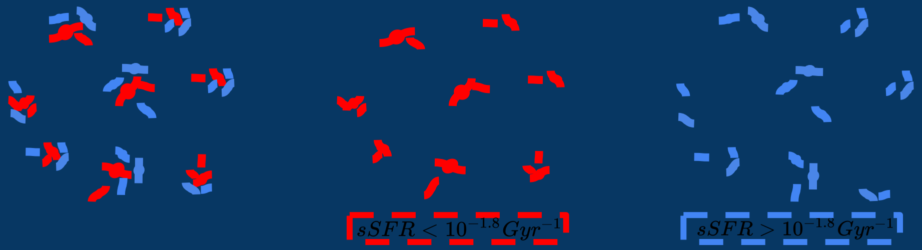
S-PLUS original image



S-PLUS masked image



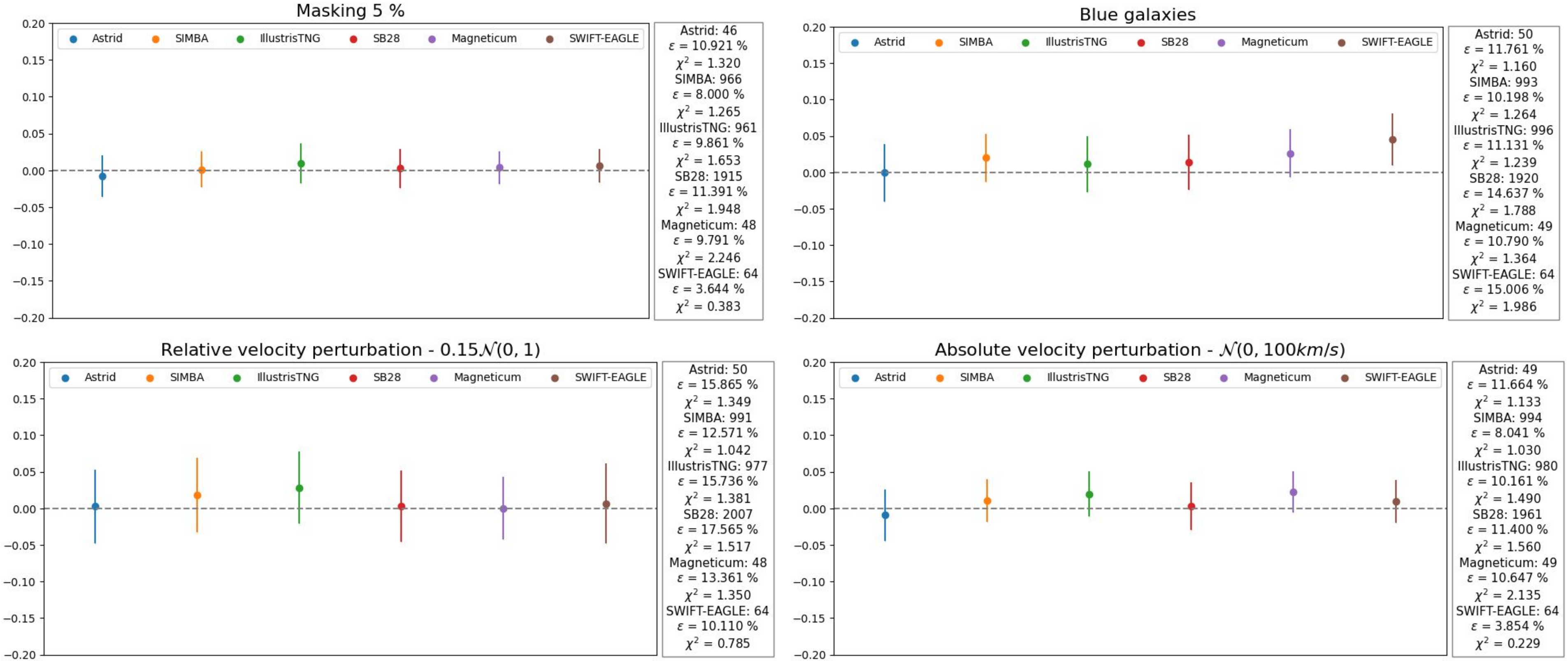
Color Selection



Velocity errors

Relative error	$v_z \Rightarrow v_z[1 + p\mathcal{N}(0,1)]$	Uncertainties in redshift
Absolute error	$v_z \Rightarrow v_z + \mathcal{N}(\mu, \sigma)$	Velocity dispersion

The Impact of Systematic Effects



[deSanti+2025]

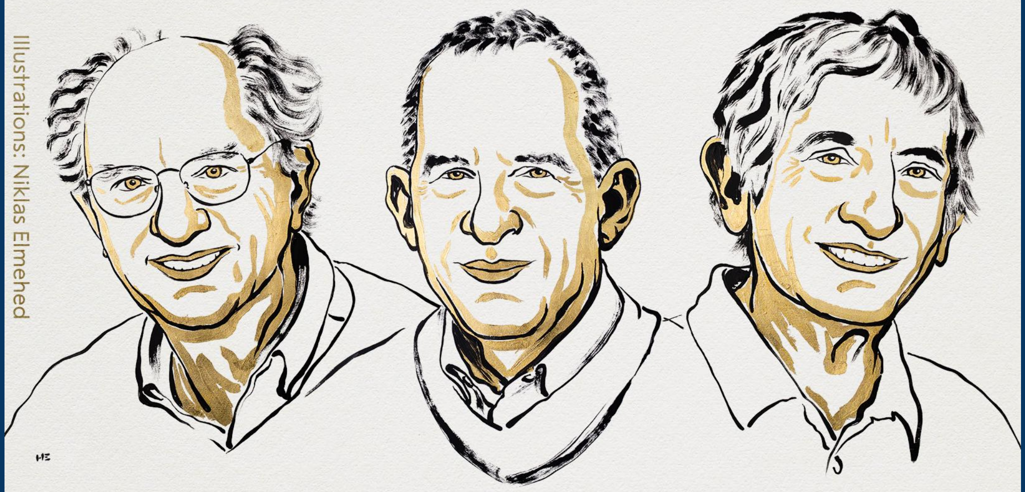
Take Away Messages

- We only need galaxy: positions and velocities;
- We got the first robust model across:
  - 5 different hydrodynamical simulations
  - Different halo/subhalo finders
  - Different variations in the cosmological/astrophysical parameters

- Equations show that the GNN make use of the phase-space information
- The model can deal with:
  - Masking effects
  - Uncertainties in the peculiar velocities and radial distances
  - Different galaxy selections

- These are the first steps before applying these techniques to real data!

THE NOBEL PRIZE IN PHYSICS 2025



John Clarke | **Michel H. Devoret** | **John M. Martinis**

"for the discovery of macroscopic quantum mechanical tunnelling and energy quantisation in an electric circuit"

THE ROYAL SWEDISH ACADEMY OF SCIENCES