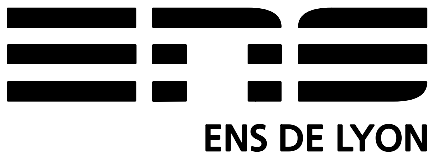
Infering cosmological velocity field with graph neural network

Albert Bonnefous-Abiven

Internship with Hideki Tanimura



Presentation

- Studying at the ENS de Lyon
- 2nd year of master in Physics
- Actually : Intern at IPMU

Main interest in Physics : Theory, Cosmology and Condensed matter
Other interest : Doing art, Learning japanese, Music

Introduction

- ΛCDM model tells us that the 84% of matter is Dark Matter, not observable
- Galaxies are nested in dark matter halo, forming a cosmic web

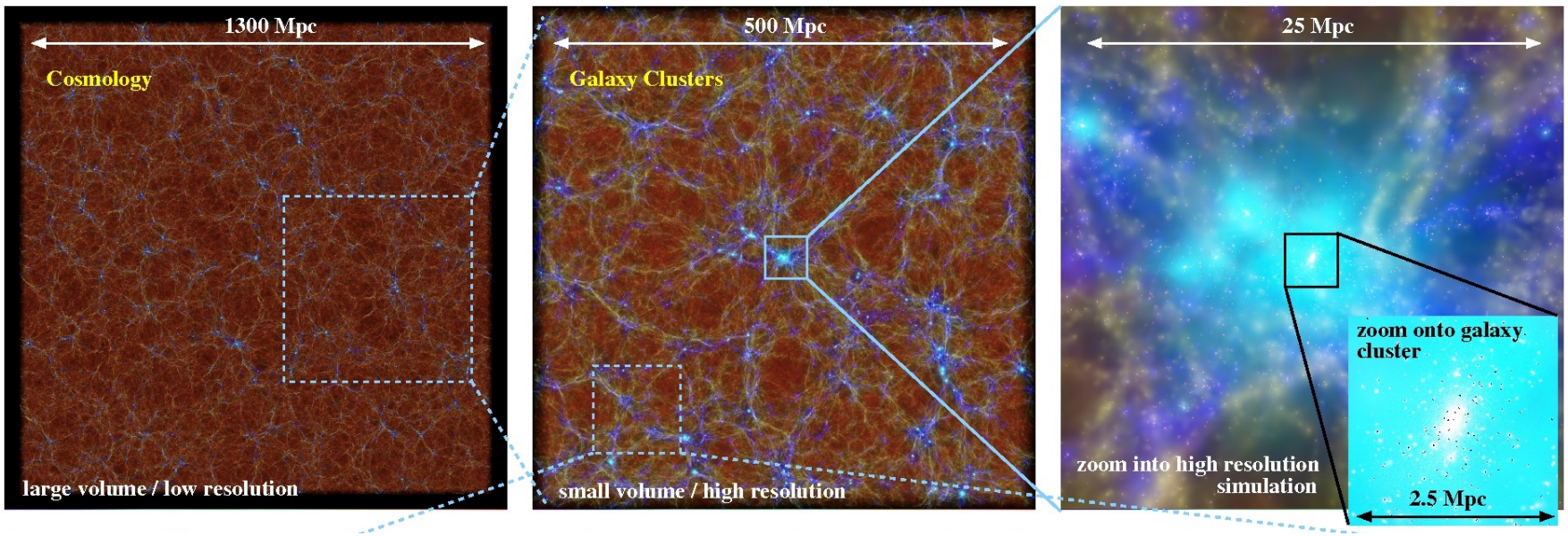


Figure: The cosmic web, in the Magneticum simulation

Introduction

- Goal of this internship : finding the velocity of these halo

Many applications, for example : correcting the Sunyaev-Zel'dovich effect on CMB observation

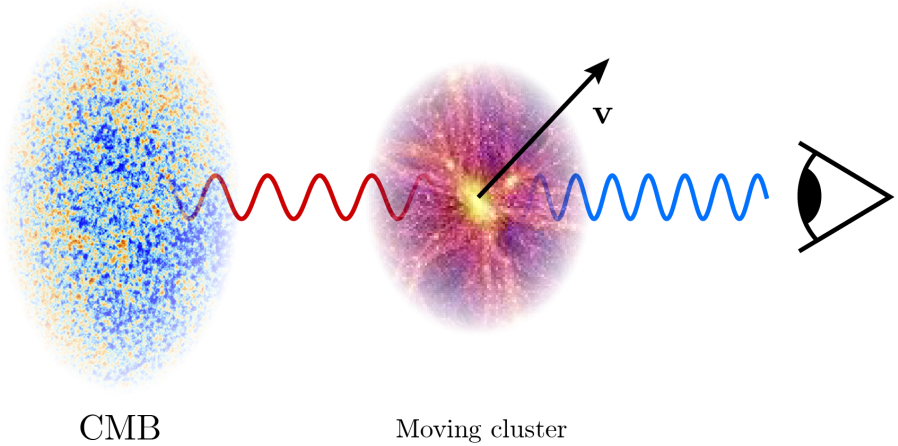


Figure: The Sunyaev-Zel'Dovich effect

But, problem :

- We cannot observe these halos
- Measuring peculiar velocity is difficult

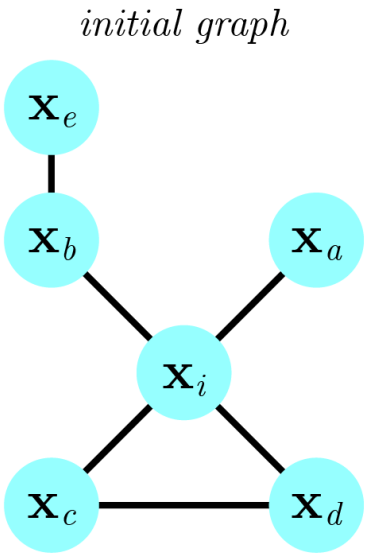
To remedy to that : let's try to use GNN

What is Graph Neural Network ?

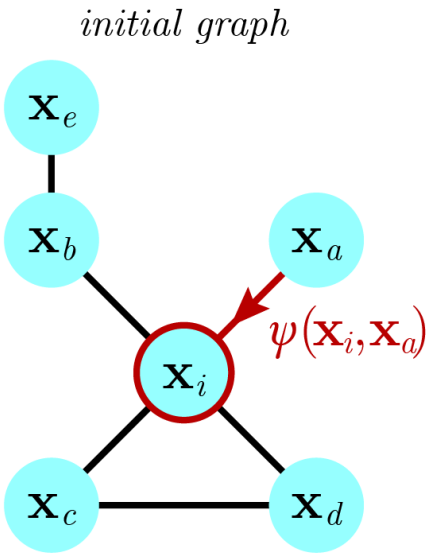
- Graph : a set of nodes linked by edges, how we want to model the cosmic web
- Neural Network : a function $\mathbb{R}^n \to \mathbb{R}^m$ with a set of parameters that can be optimised

Idea : training a GNN model on simulation (Magneticum) to find that velocity, in order to apply it to real data

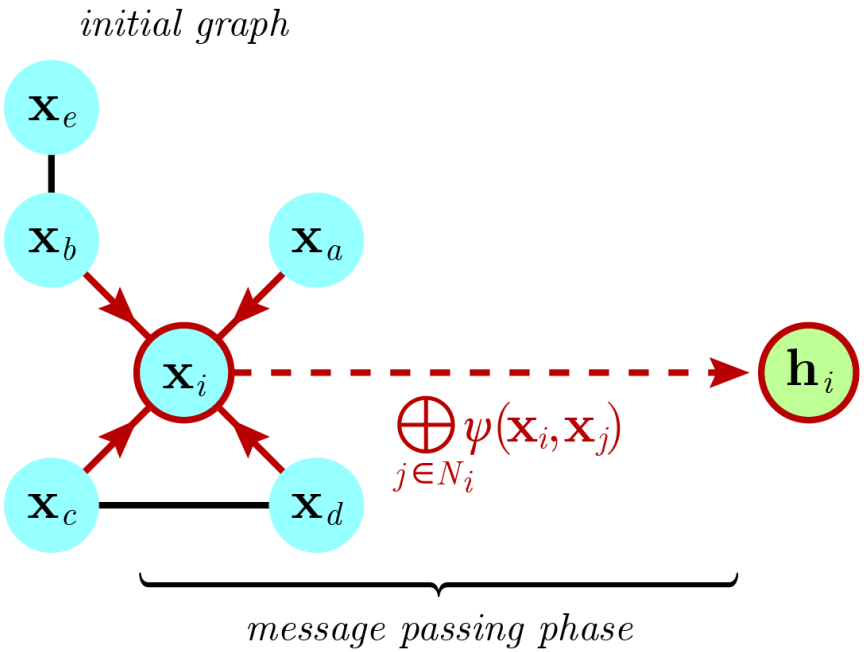
How does a Graph Neural Network works



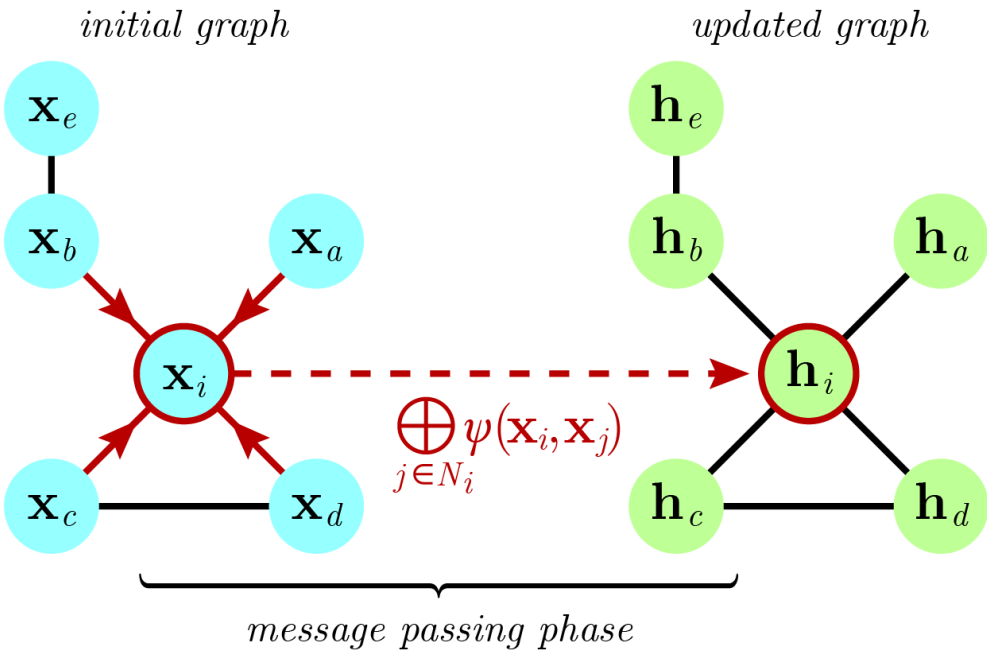
How does a Graph Neural Network works



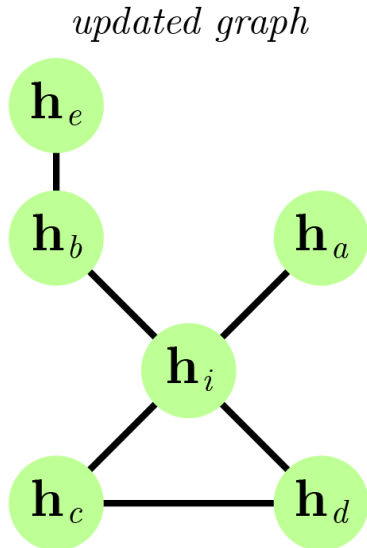
How does a Graph Neural Network works



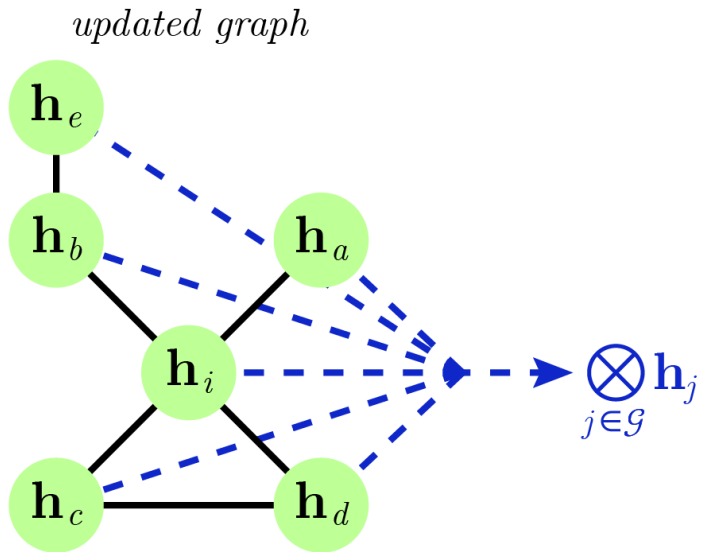
How does a Graph Neural Network works



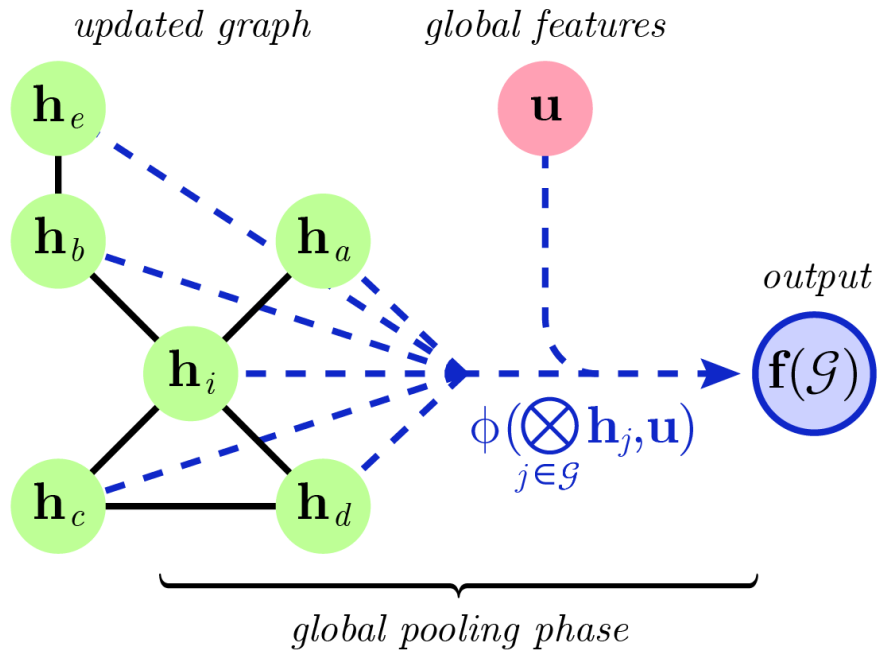
How does a Graph Neural Network works



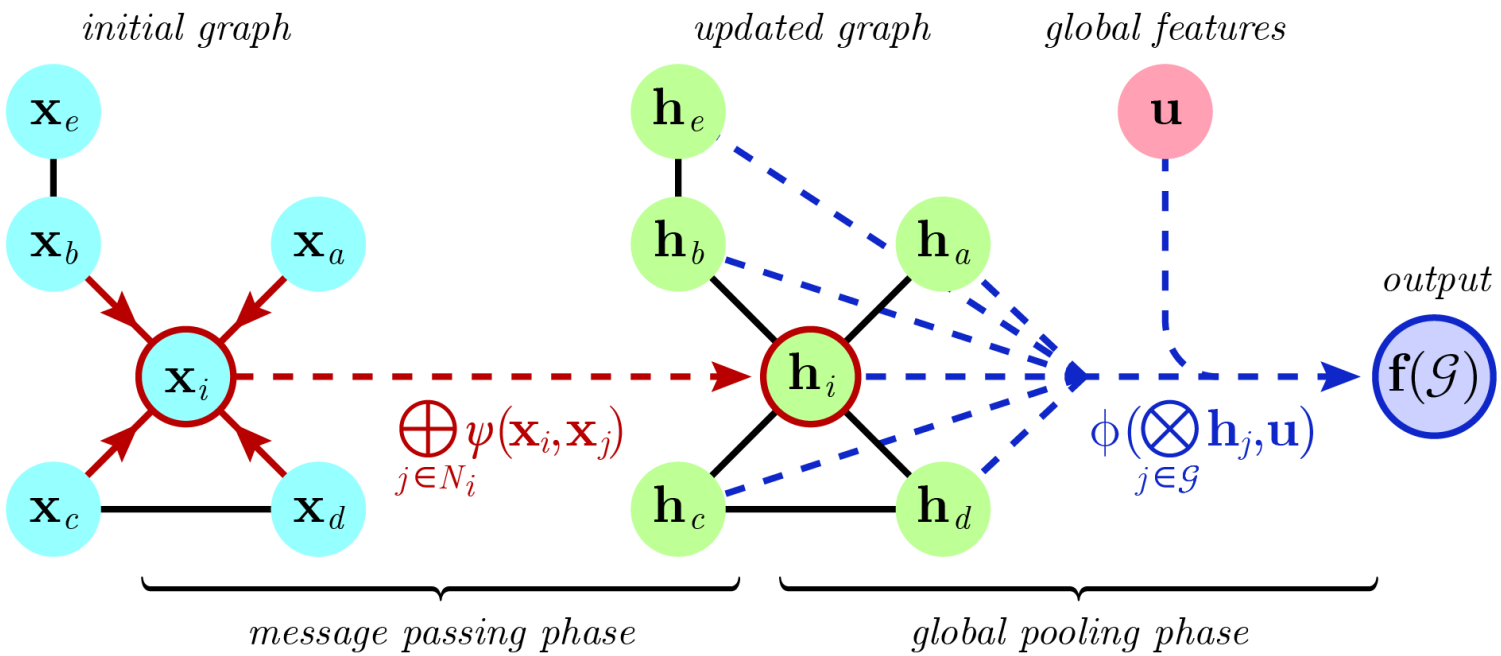
How does a Graph Neural Network works



How does a Graph Neural Network works



How does a Graph Neural Network works



Applying GNN to clusters of galaxies

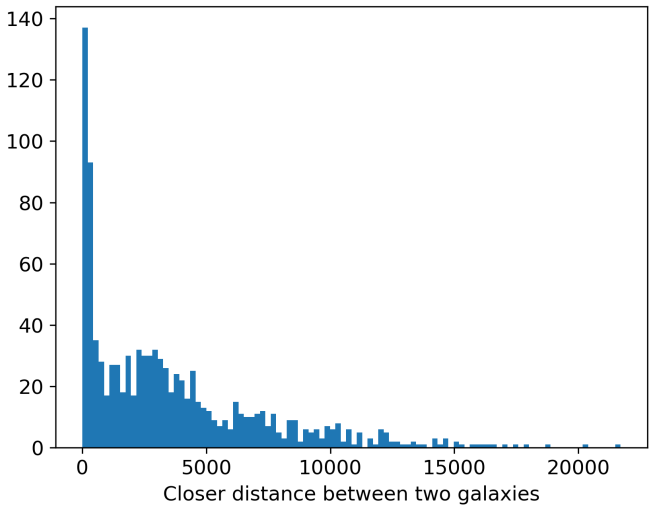
- Nodes : all galaxies within R_{max} of center halo
- Node features : position, and mass
- Global features : total number of galaxies + total mass
- Edges : two galaxies are linked if they are close enough $d < d_n$

Problem of GNN $\rightarrow$ MANY hyperparameters

- learning rate, weight decay, n_{layers}, $n_{neurons}$, n_{mp}, size of the hidden vectors
- plus choice of aggregation functions $\bigoplus$ and $\bigotimes$
- plus d_n, R_{max}

Applying GNN to clusters of galaxies

How to choose d_n :



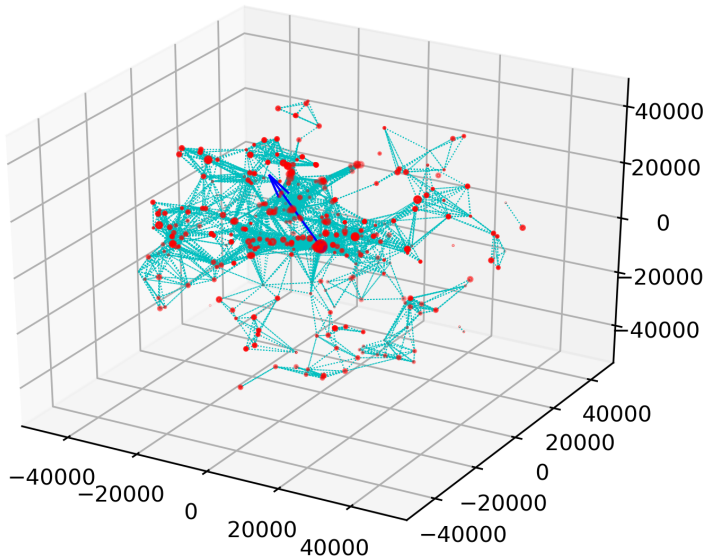


Figure: Plot of the distribution of the closest distance between two galaxies, and the graph obtained with $d_n = 15$ Mpc

Results

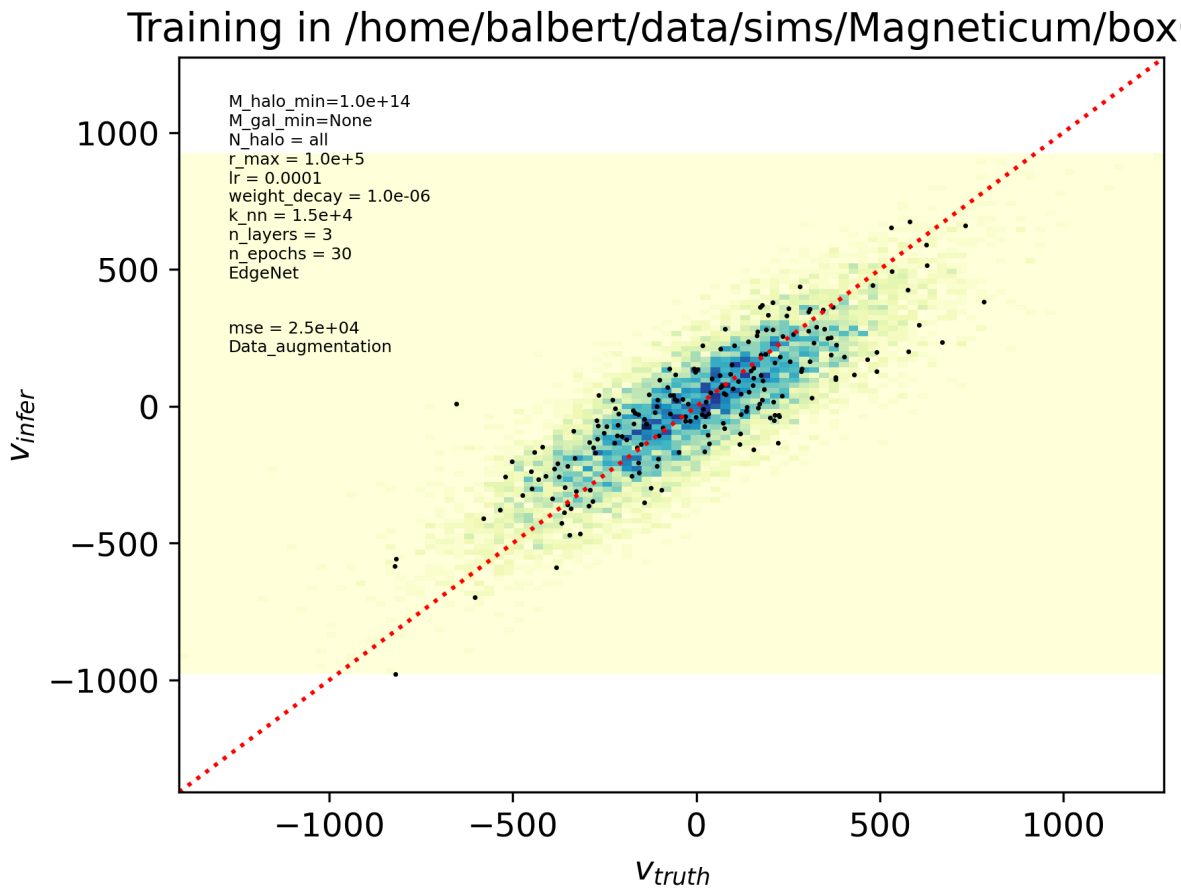


Figure: Accuracy of the tuned GNN

And now ?

- Trying to compare the accuracy with other features
- Trying to see if there is a way to interpret the GNN