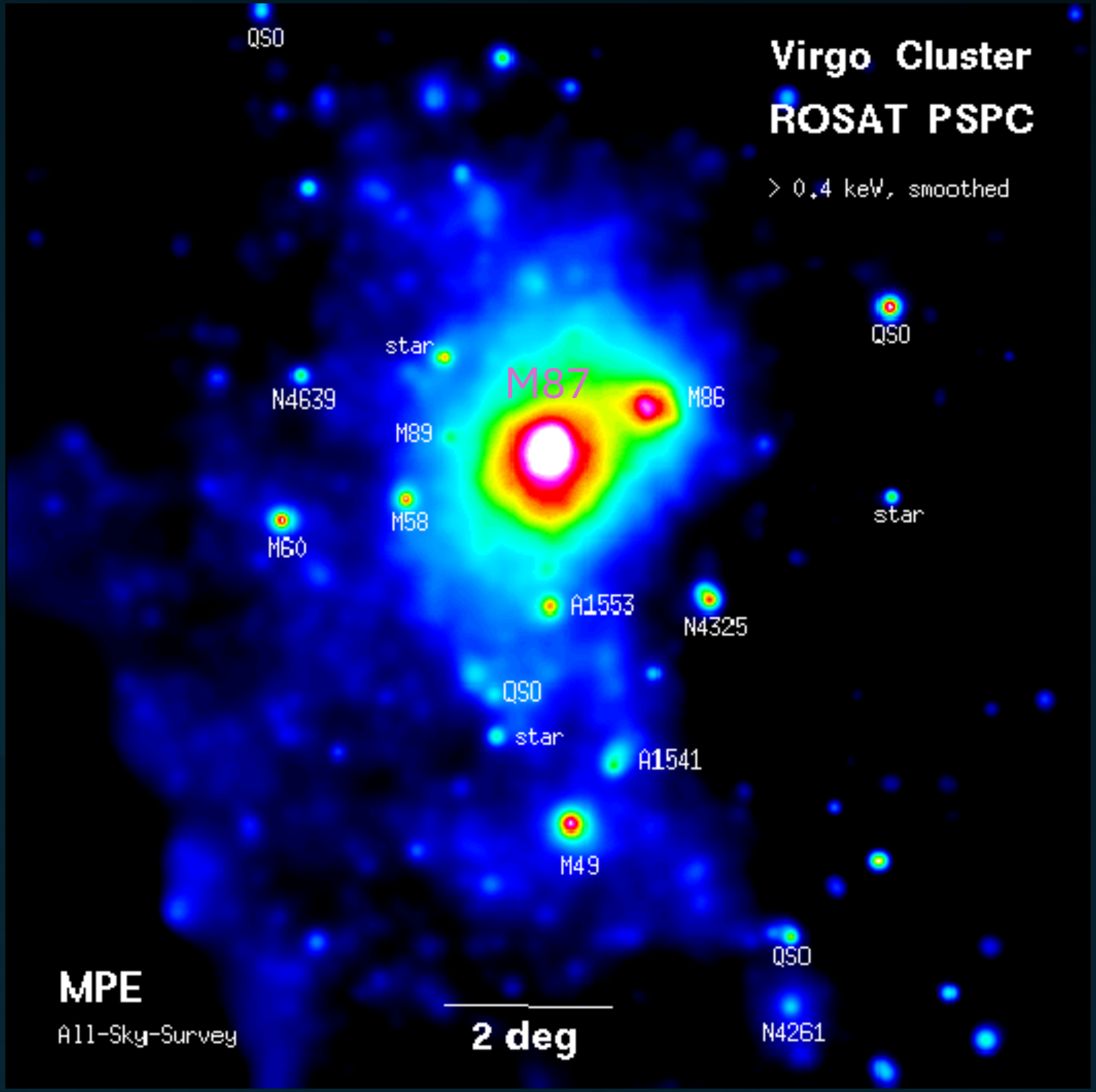
Inferring the past merging history of Galaxy Clusters with Machine Learning

Shera Jafaritabar
Master’s Thesis presentation
phase 1

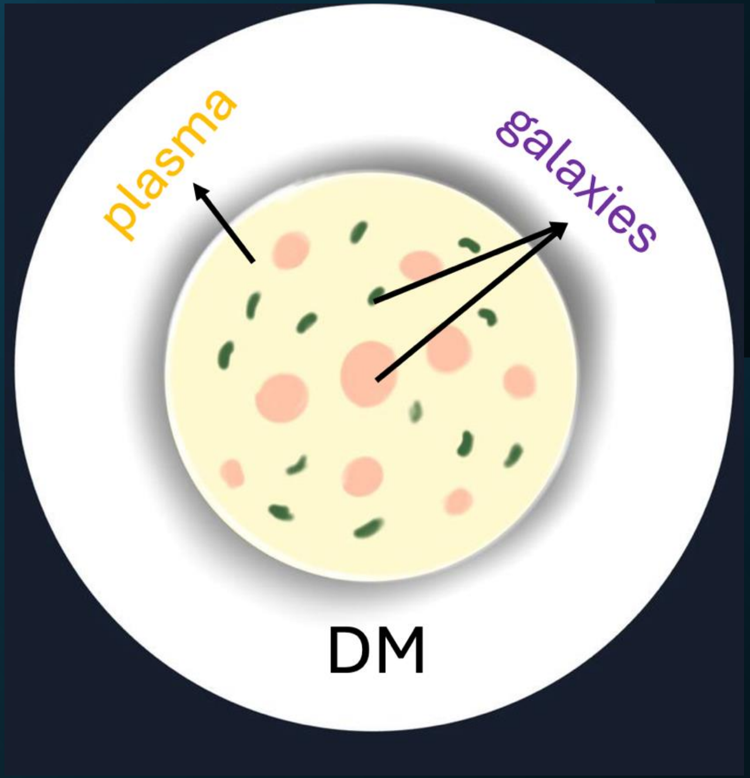
Galaxy Clusters:

- Consists hundreds to thousands of galaxies
- Mass: $M_{200c} \geq 10^{14} M_{\odot}$
- Radius: : $2 \leq R_{200c} \leq 5\ Mpc$
- Higher densities than groups and contains mostly E's and S0s
- They are the largest gravitationally bound
- They may not be complelely virialized



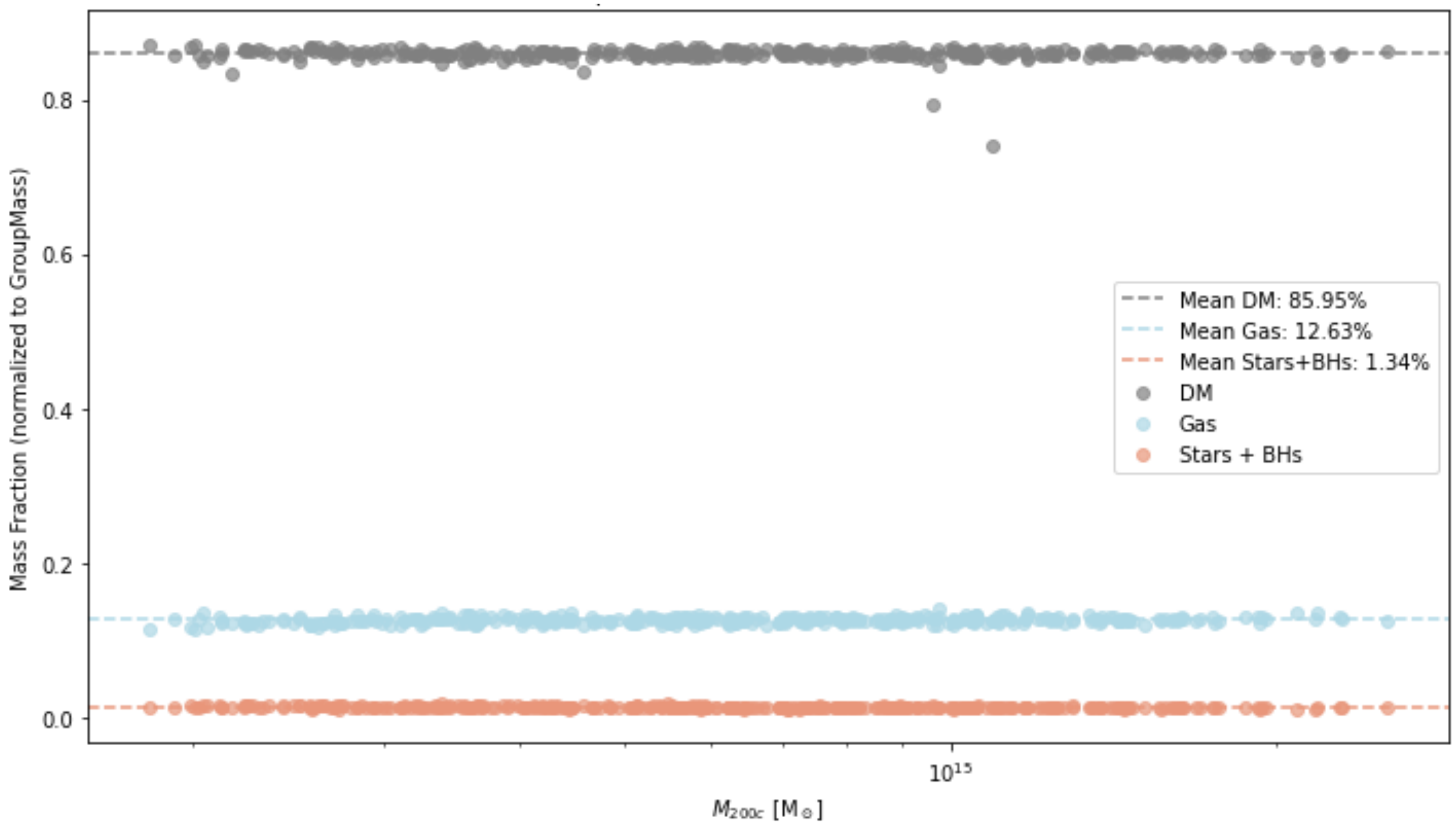
Galaxy Clusters Components:

- DM ~85%
- ICM ~13%
- Galaxies ~2%



How are they Measured in Galaxy clusters

- Gravitational Lensing
- X-ray and Radio
- Optical, IR, and radio



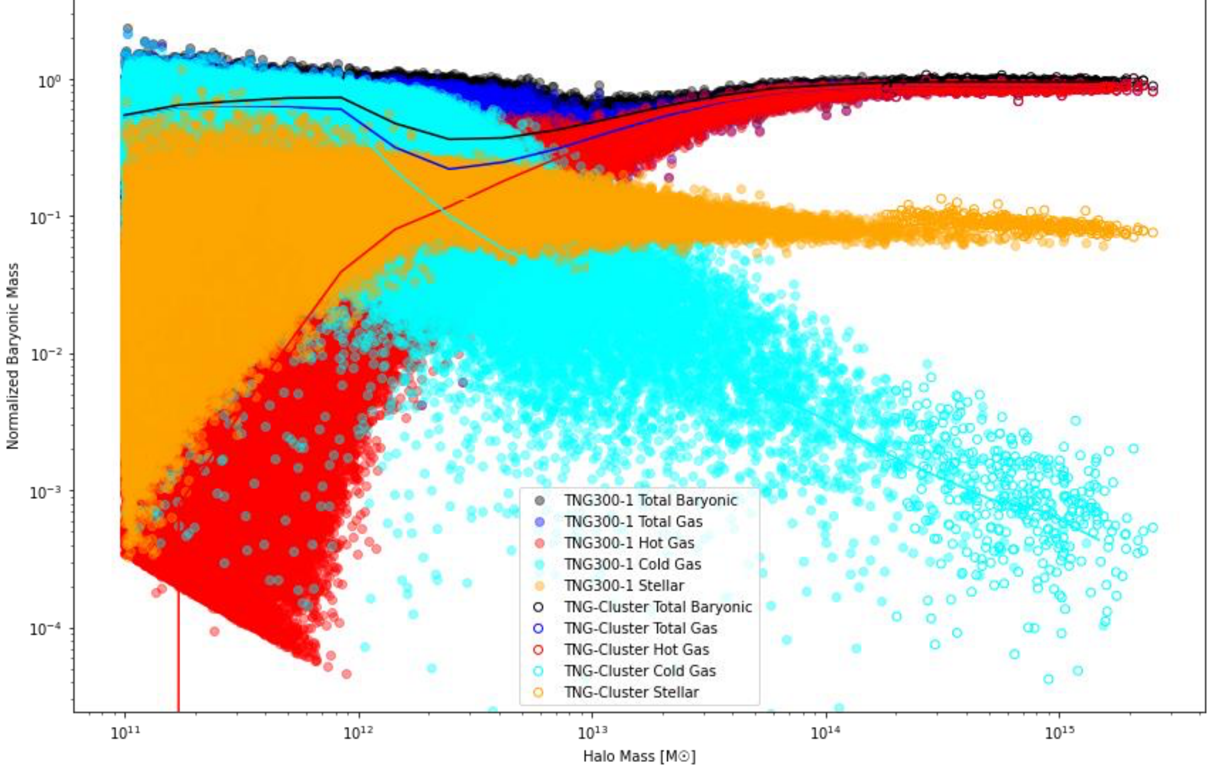
ICM

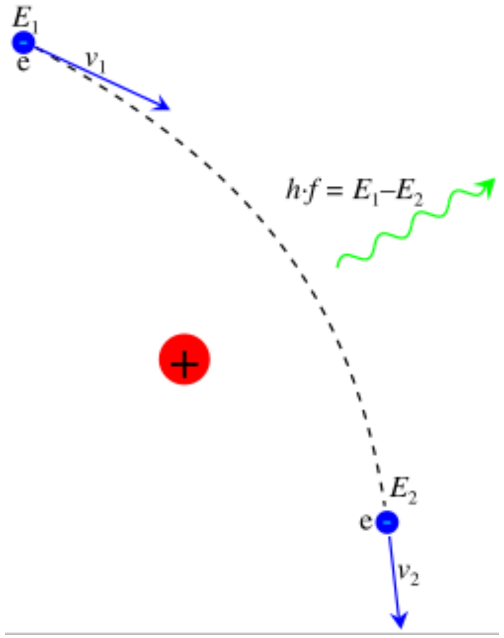
most of the mass of the baryons are
in very hot gas $T > 10^6$
Fully ionized hydrogen and helium

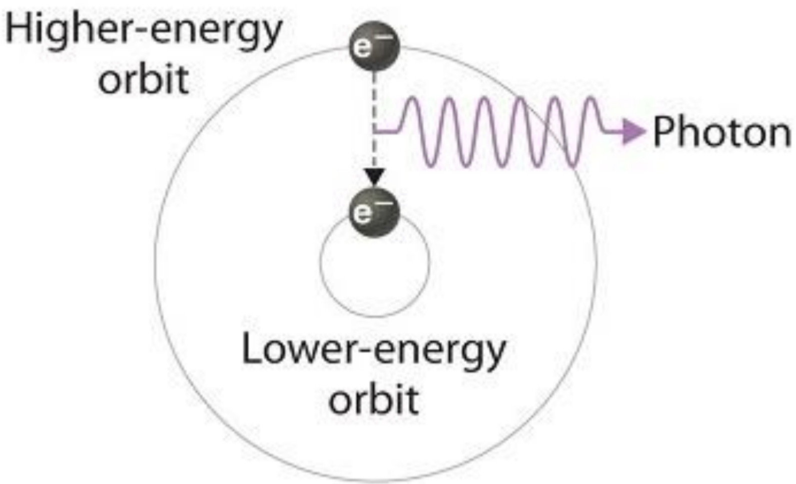
X-ray

1. Bremsstrahlung (Thermal electrons)
2. Metal line emission (such as Fe, Si, O, …_)

Surveys: Chandra, eROSITA, XRISM, …







How Galaxy Clusters are formed?

Dark Energy
Deriving the accelerated expansion of the universe

ΛCDM

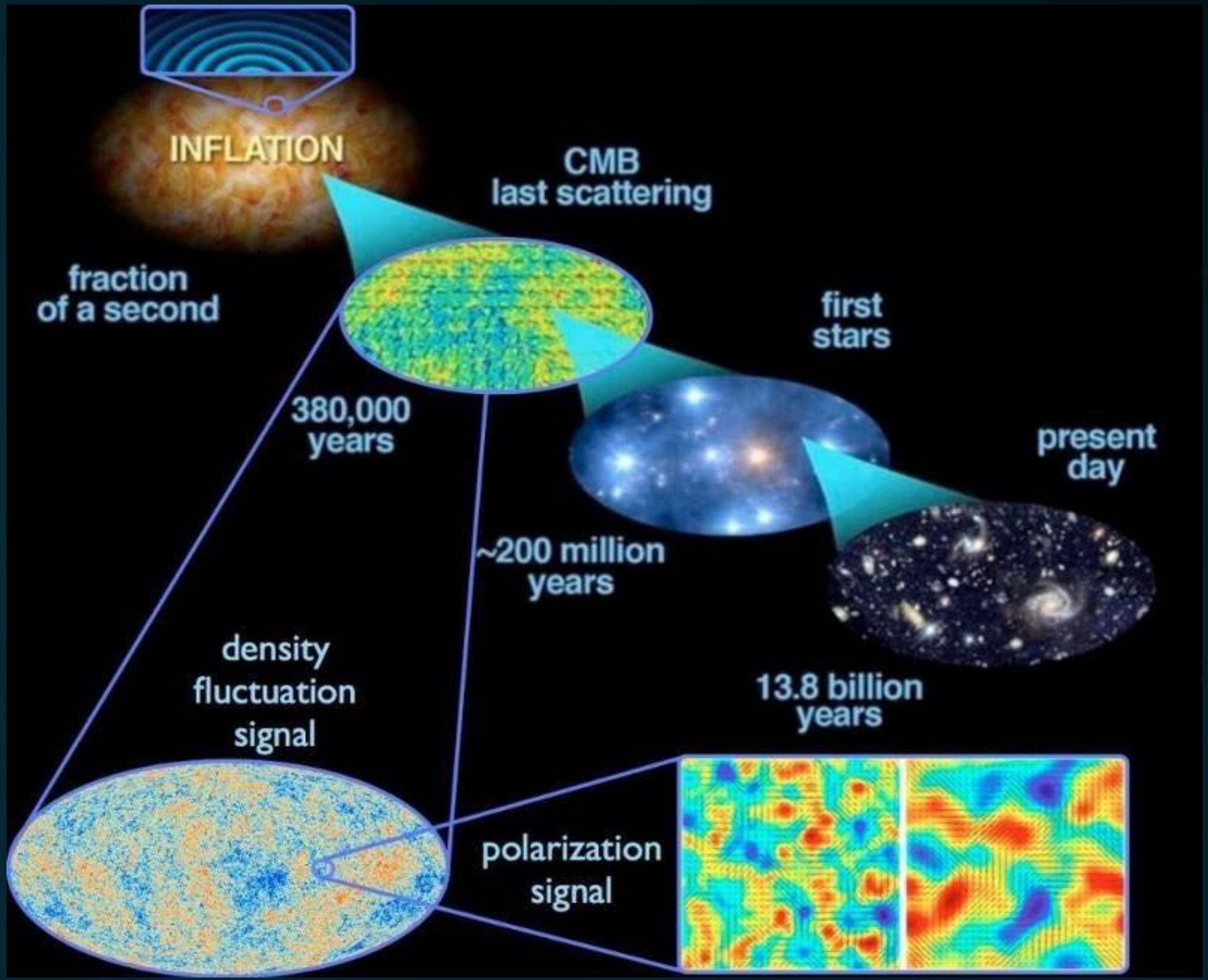
Cold Dark Matter:
slow-moving allowing formation of small structures

1. Initial Density Fluctuations
Tiny quantum fluctuations in the early Universe (seen in the CMB) which grow under gravity

2. Growth of Dark Matter Halos
DM collapses first, forming gravitational wells → Gas (baryons) falls into these wells → forms galaxies

3. Hierarchical Structure Formation
ΛCDM predicts a "bottom-up" scenario
small halos merge hierarchically to form larger halos
→ groups → galaxy clusters

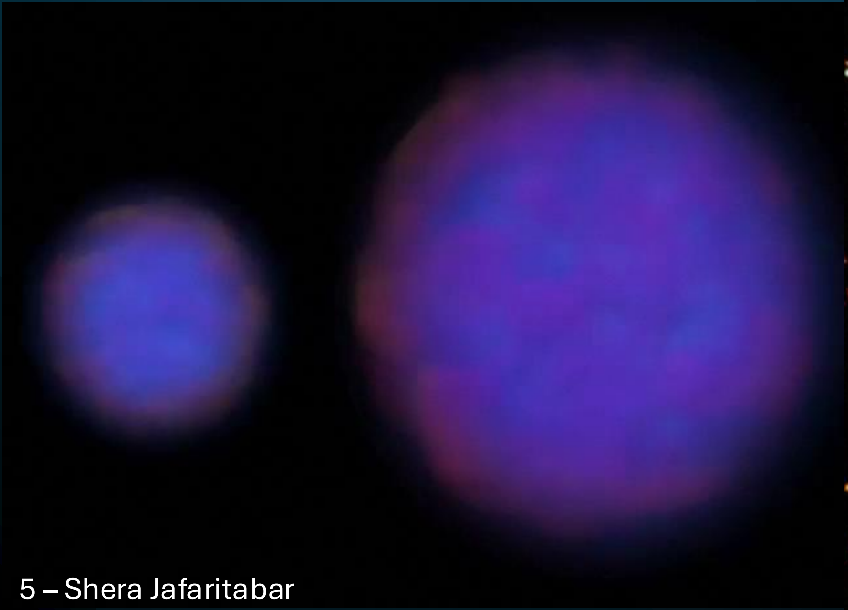
4. Role of Dark Energy (Λ)
•At late times ($z < 1$), dark energy slows down structure growth

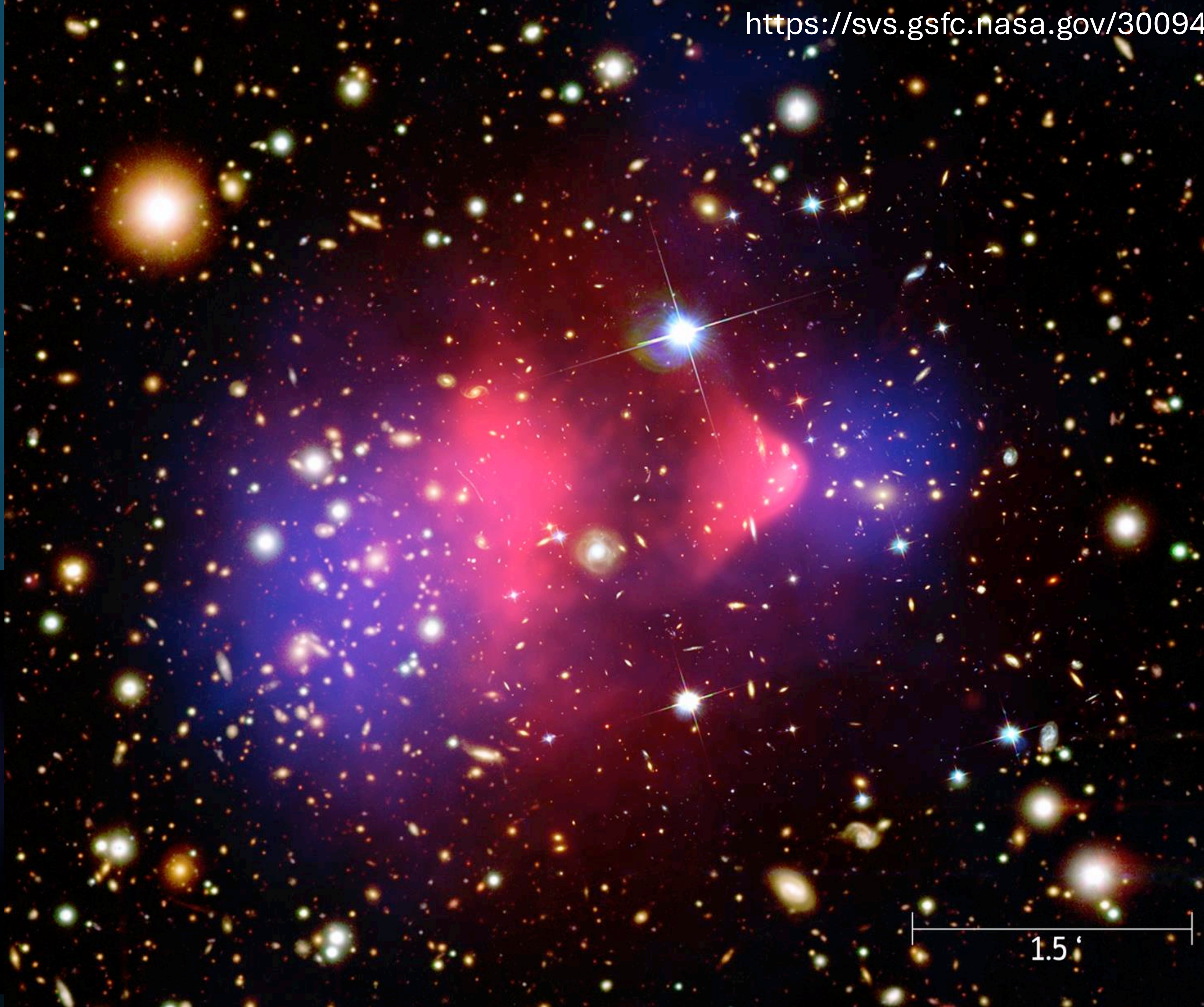


Bullet Cluster

X-ray: hot gas, Chandra
Optical: Magellan and HST
DM: Blue, through weak lensing

During the collision, the hot gas is slowed and distorted by a drag force and resistance, while DM separates from the normal matter.



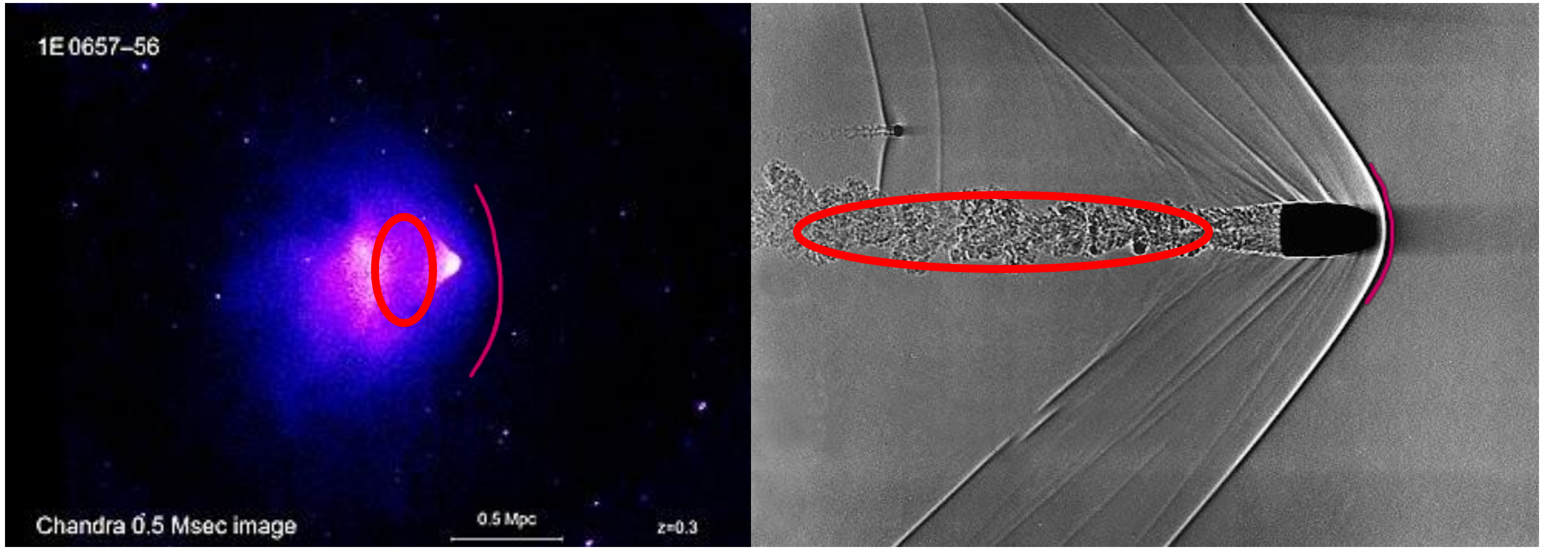


Relativistic electron

Thermal e are accelerated:

1. Merger driven Turbulence

2. Shock acceleration

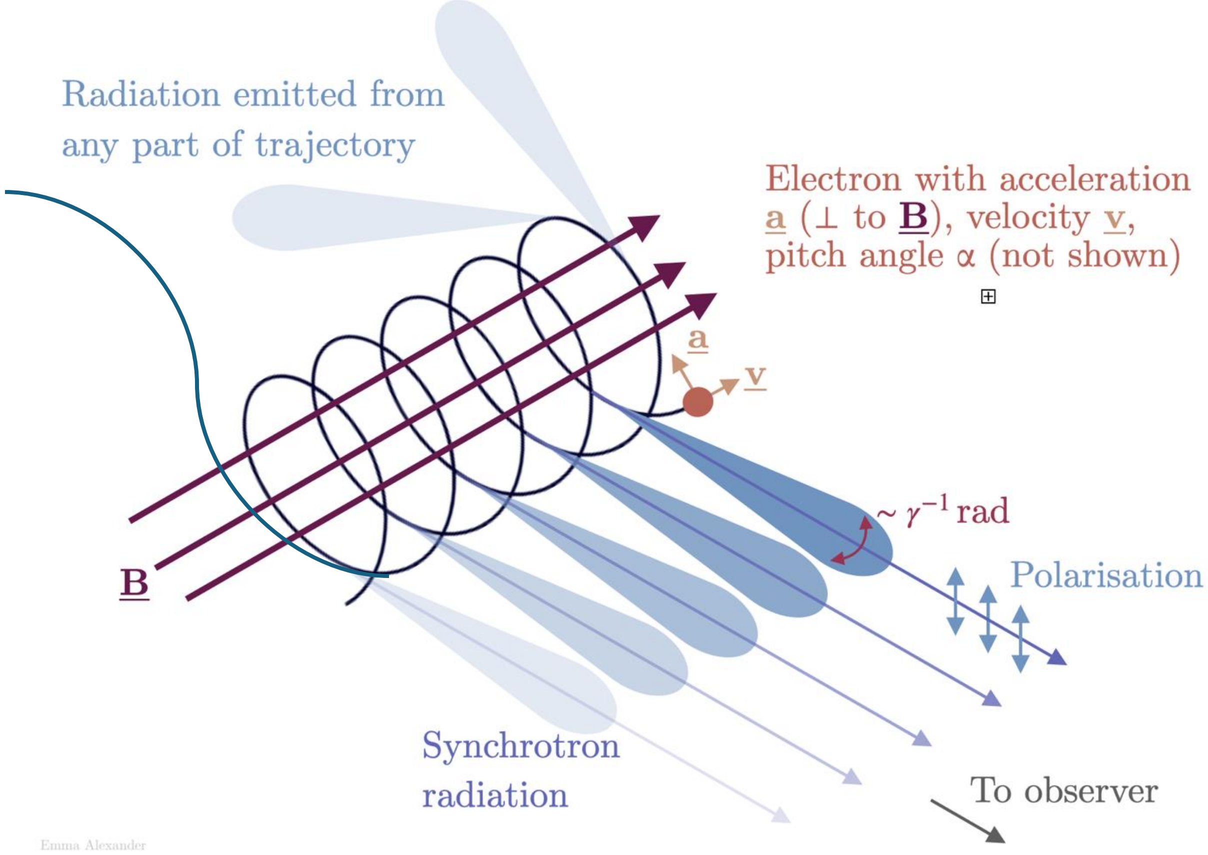


Radio Emission (non-Thermal)

ICM is a plasma filled with MF

Primarily Synchrotron:

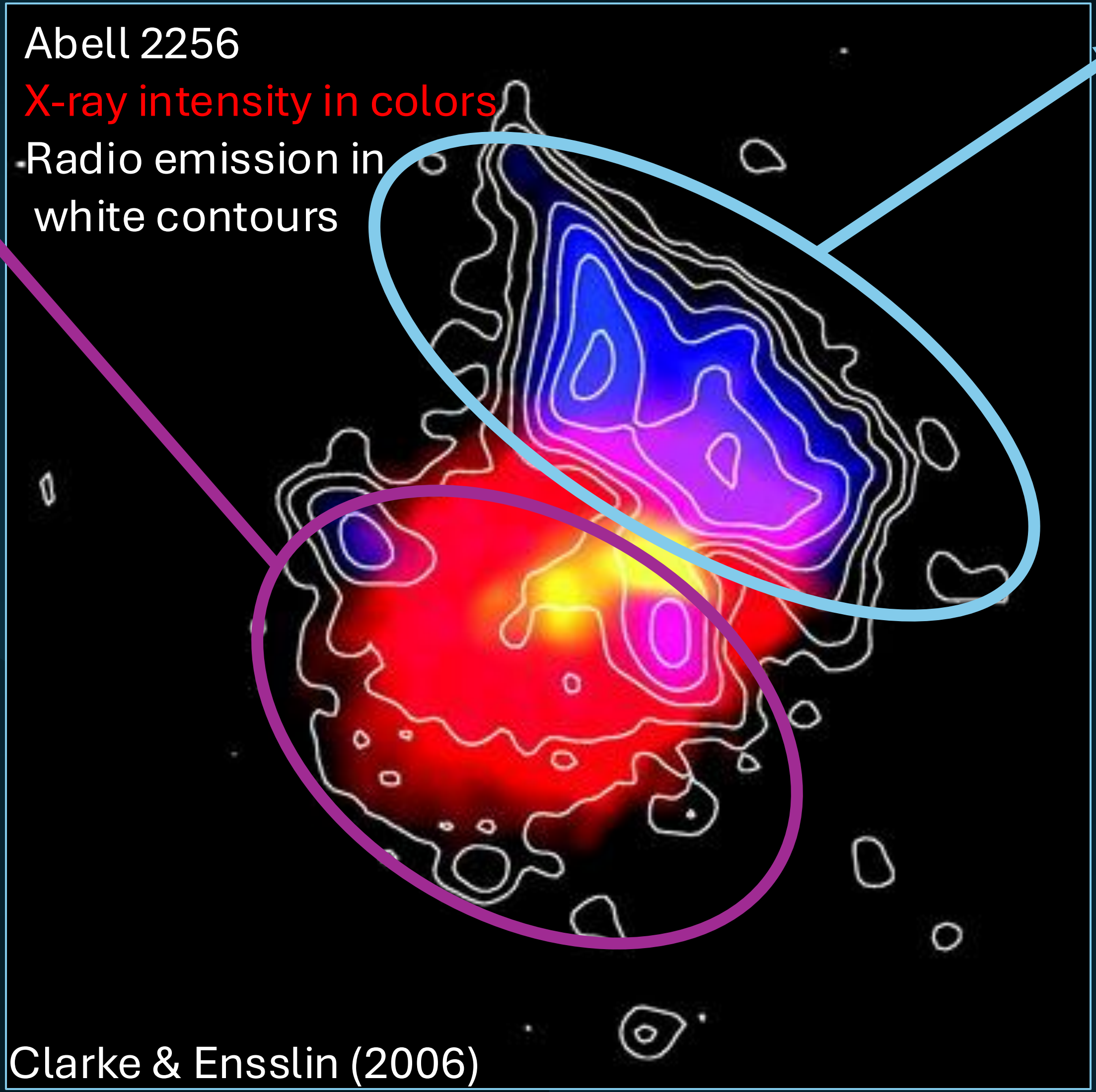
EM radiation emitted by relativistic electron in a magnetic field



Radio Emissions in Clusters

Radio Halo

- Diffuse and located at the cluster centers, unpolarized
- Follows the ICM X-ray distribution
- Formed via turbulent reacceleration of the ICM electrons
- Detected mostly in massive merging clusters



Radio Relics

- Diffuse radio synchrotron emission
- Mpc sized, extended
- In cluster outskirts
- Strongly polarized
- Formed via cluster shocks
- Shocks align MF → strong polarized arc-like relics

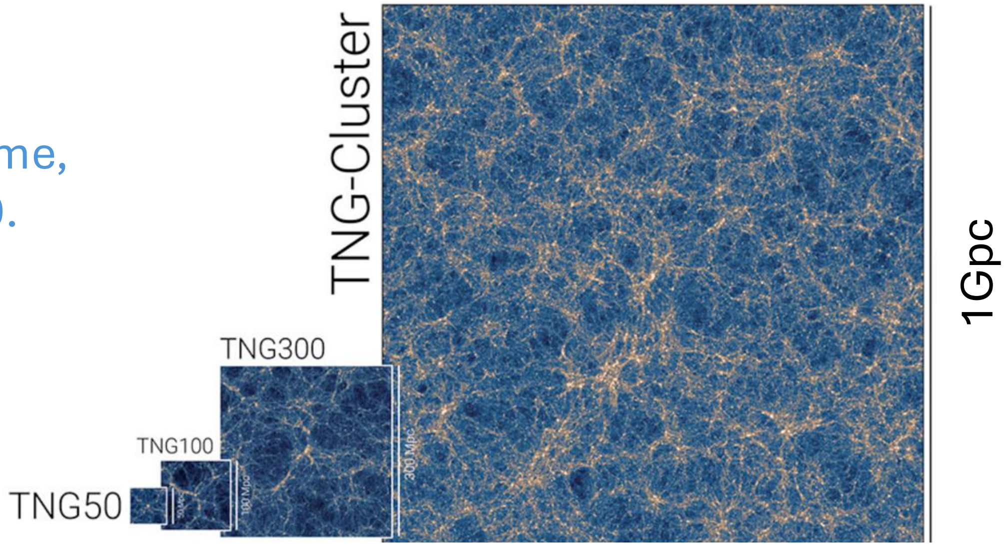
TNG-Cluster
observe clusters at a single point in time
+
Information lost in observations

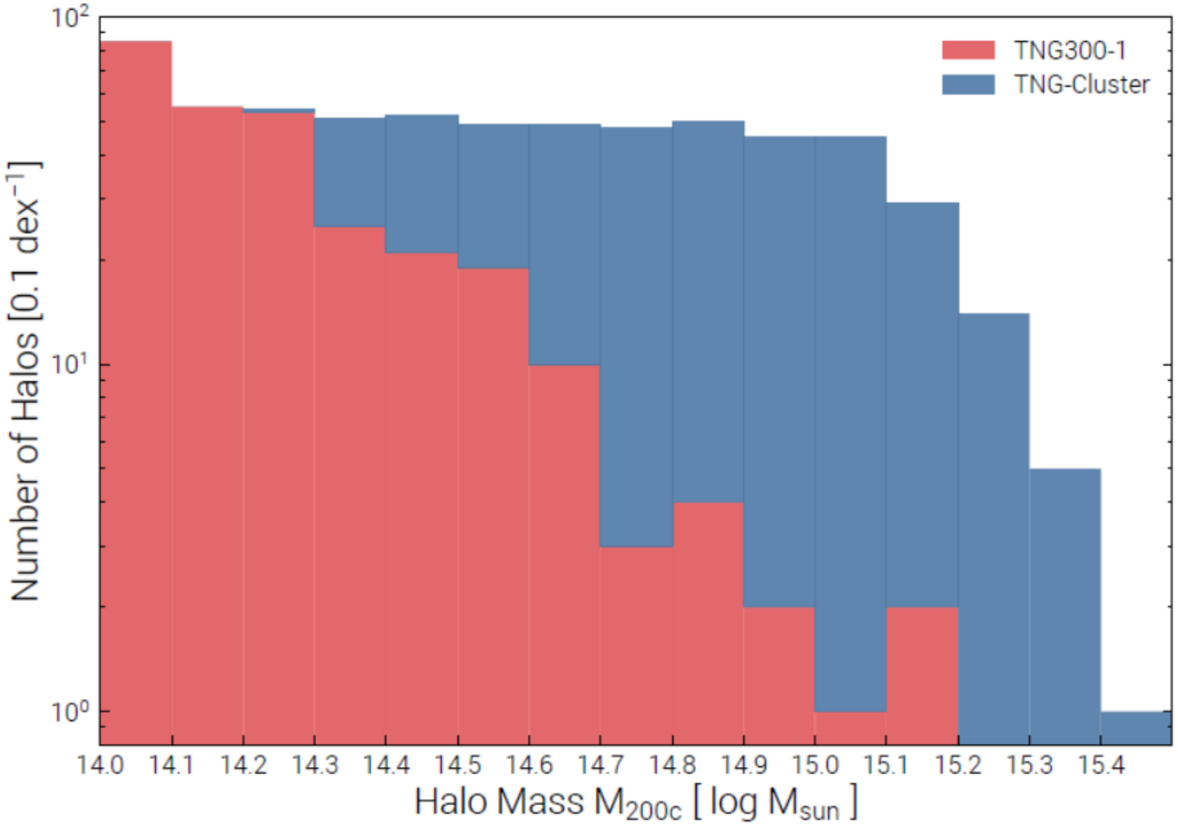
zoom in simulation:

1. selected 352 halo in DM only at z=0
2. trace all DM particles to z=137
3. adaptive Oct-tree around each cluster
4. zoomed-in high resolution region are up to 8192^3 particle ~ TNG300-1
5. Reconstruct a cosmological volume by stitching and shifting the zoom box

Result: 352 high-res cluster regions at z=0 with $\log M_{200c} : 14.3 - 15.4 M_\odot$

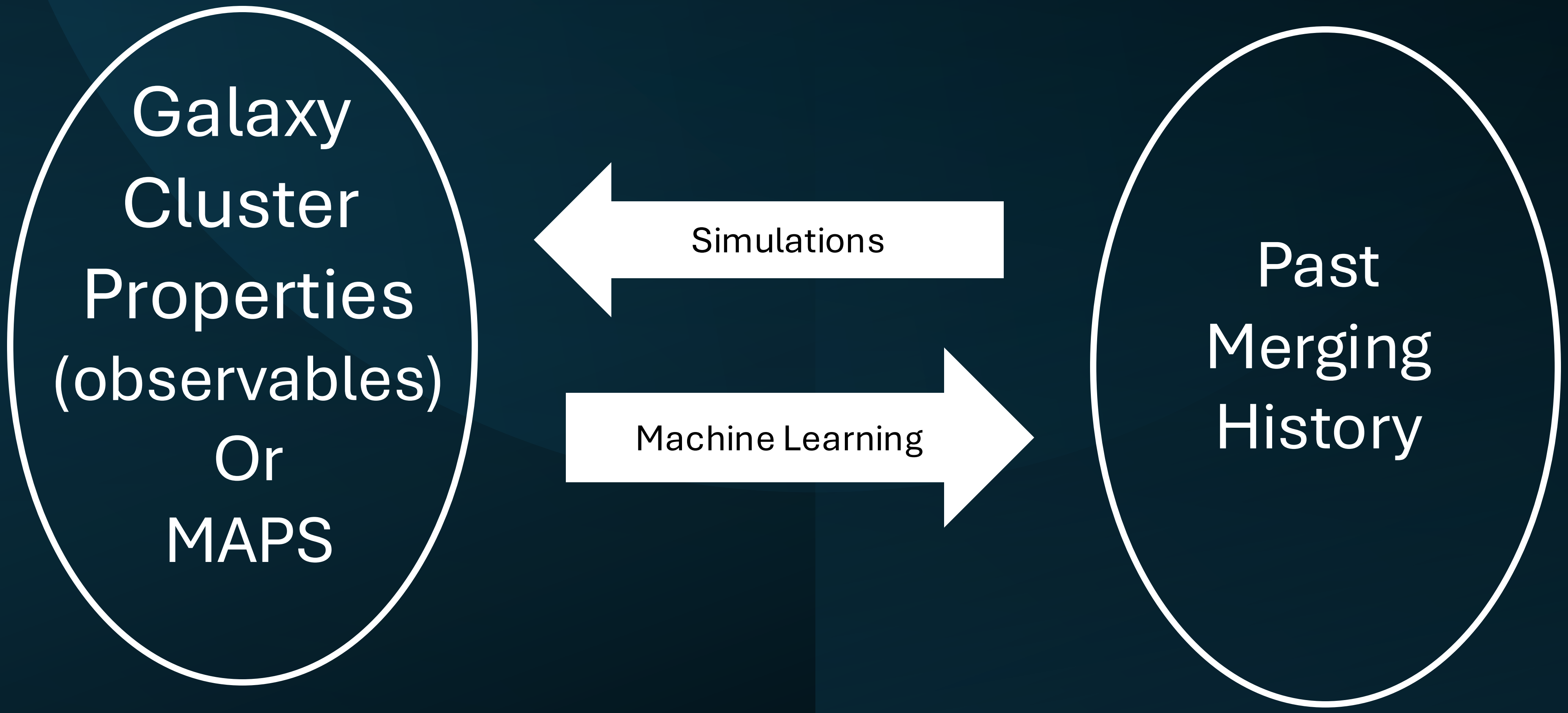
1 Gpc^3 volume, x36 TNG300.





Nelson et al (2023)

Inferring the Past Merging History of Galaxy Cluster



Inferring the Past Merging History of Galaxy Cluster

Galaxy Cluster Properties (observables) Or MAPS

- Scalar Observable Properties
- X-ray maps
- Radio maps

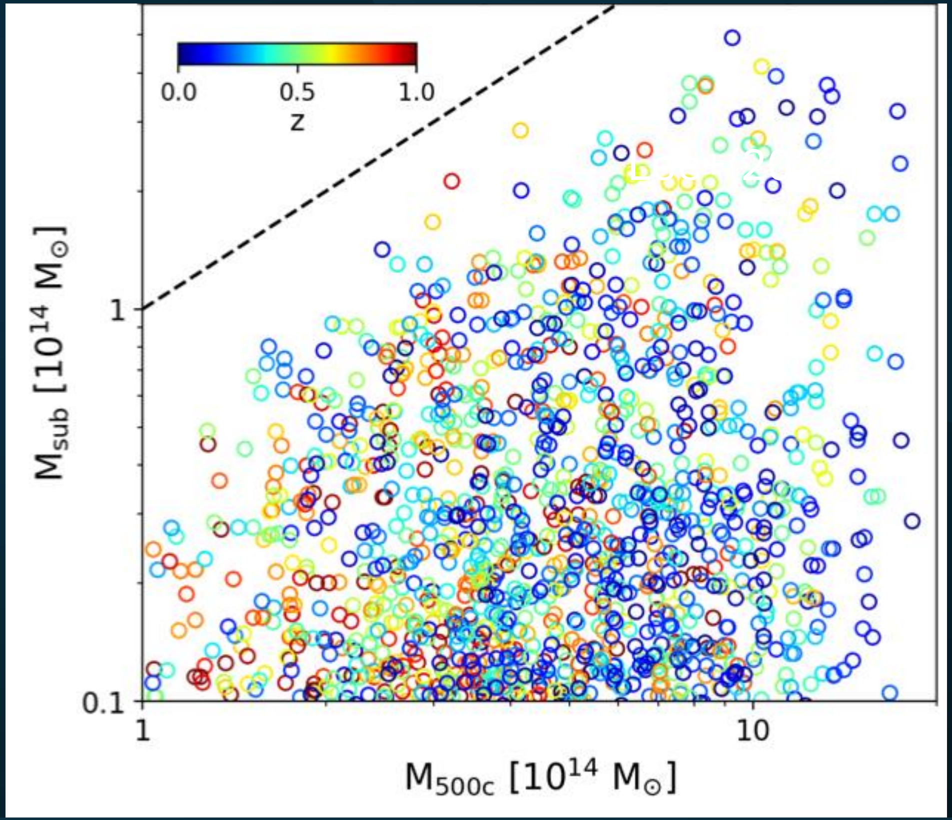
Lee+(2024)

Last Merger History

352 main zoom-in targets

Subcluster mass: $M_{sub} > \ 10^{13} M_{\odot}$

~2000 collisions with 0<z<1



Time of Collision

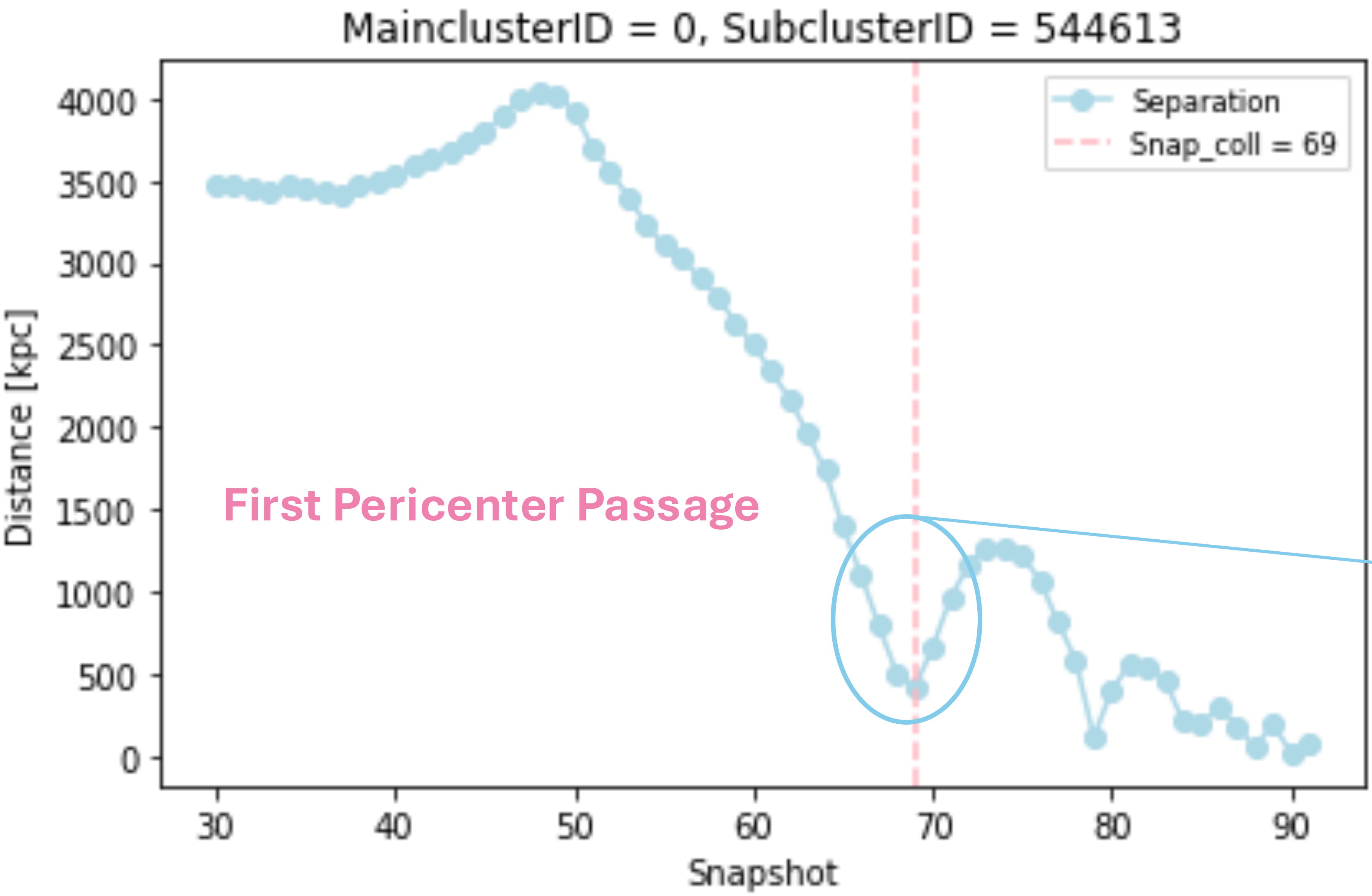
Velocity of collision

Peri-center distance

Main Halo's M_{500c}

Mass Ratio

Sub-cluster mass

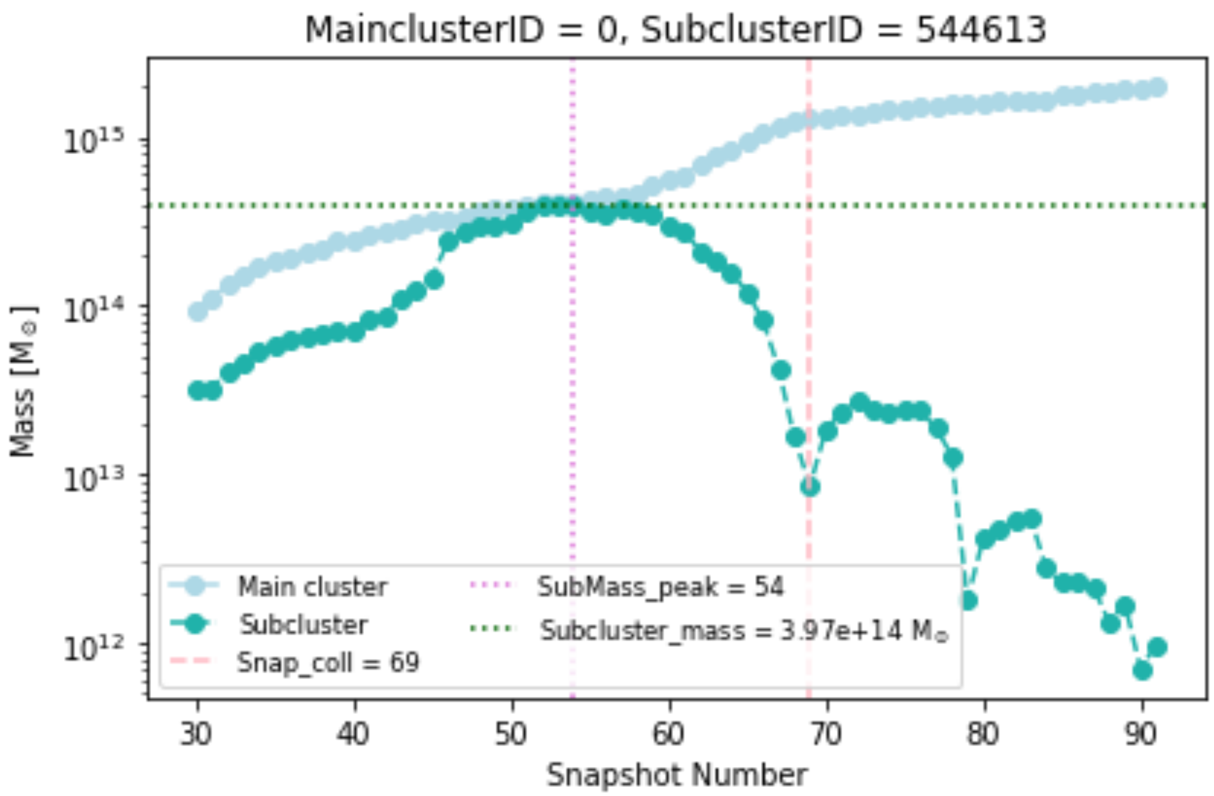


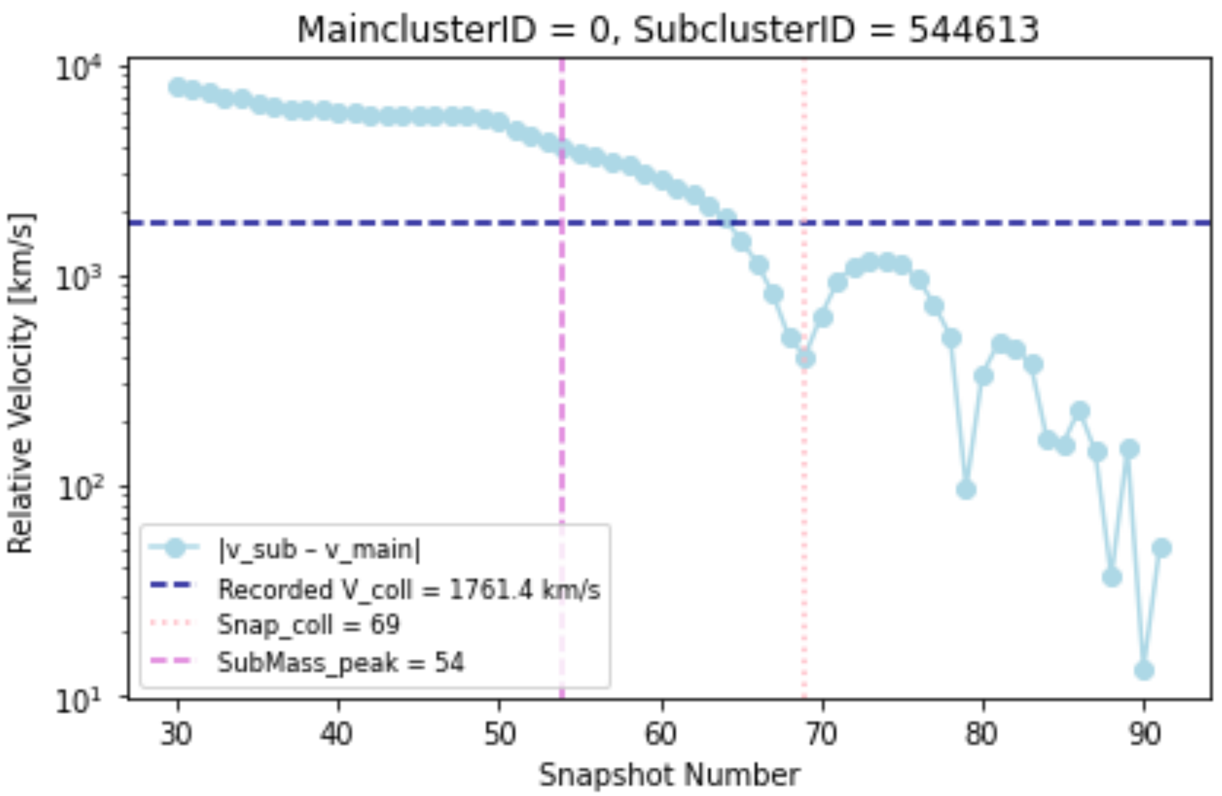
Properties measured:

Time of Collision
Pericenter Distance
M_{500c} of the main cluster

Fit a quadratic function for exact:

1. Collision Time
2. Pericenter Distance





Mass ratio
Subcluster Mass

Where subcluster mass peaks

Collision velocity: d(separation)/dt at?

Centers: SubhaloPos

But how?

Conditional Invertible Neural Networks

$$p(x|c) = \frac{p(x,c)}{p(c)} = \frac{p(c|x)p(x)}{p(c)}$$

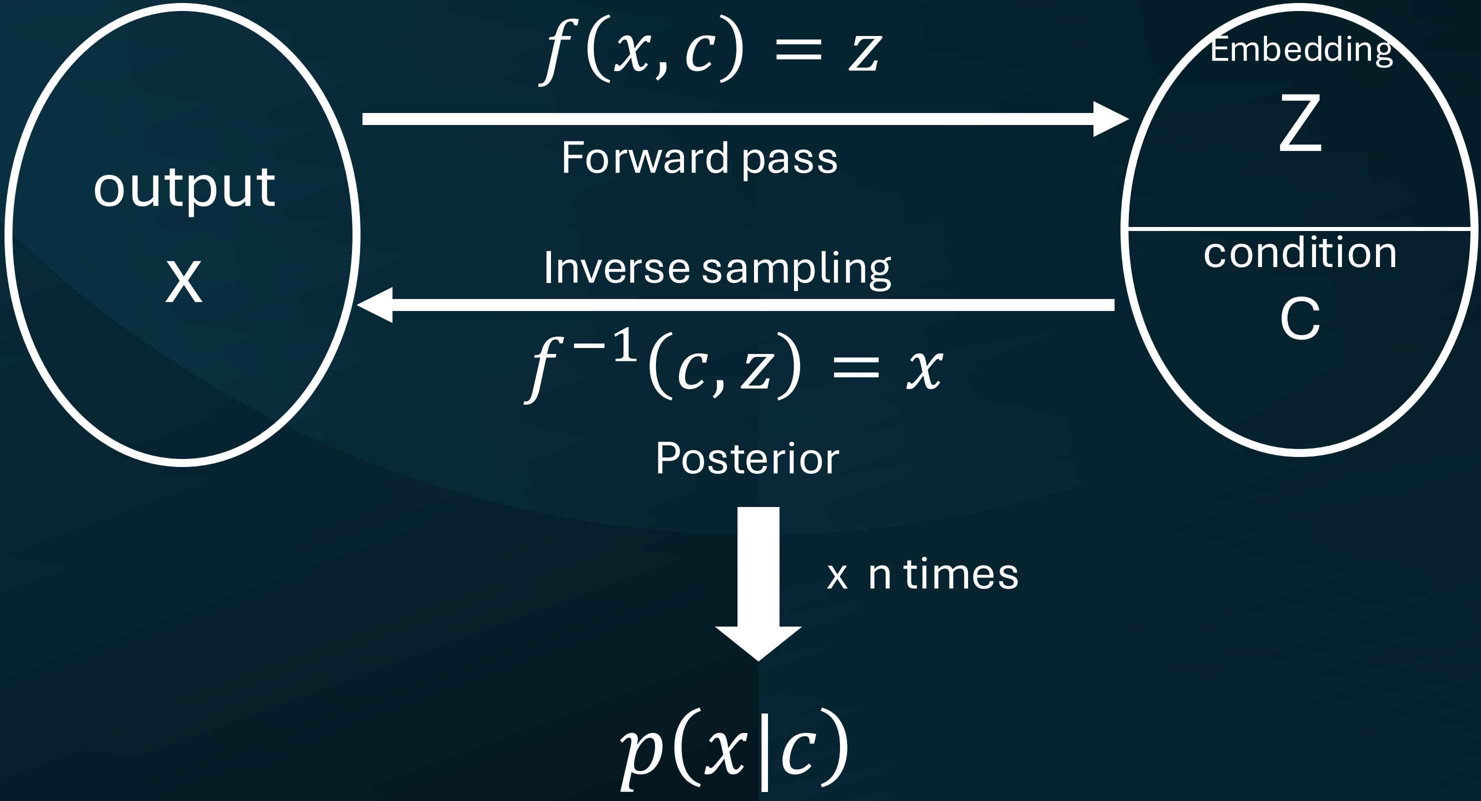
X : last merger history properties
C : condition (inputs)

P(x): prior
P(c): evidence (marginal distribution)

p(c|x): likelihood
p(x|c): conditional probability distribution
= posterior

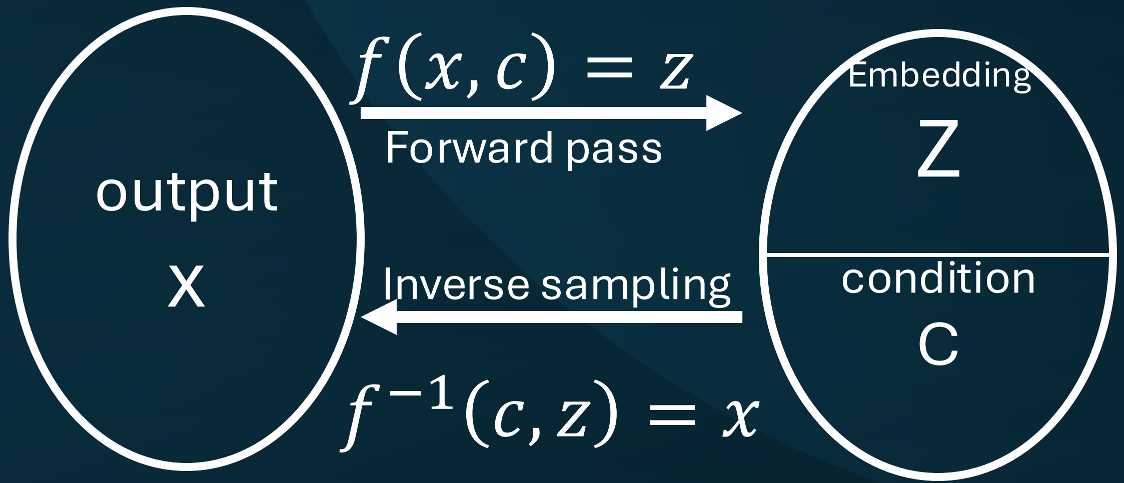
But how?

Conditional Invertible Neural Networks



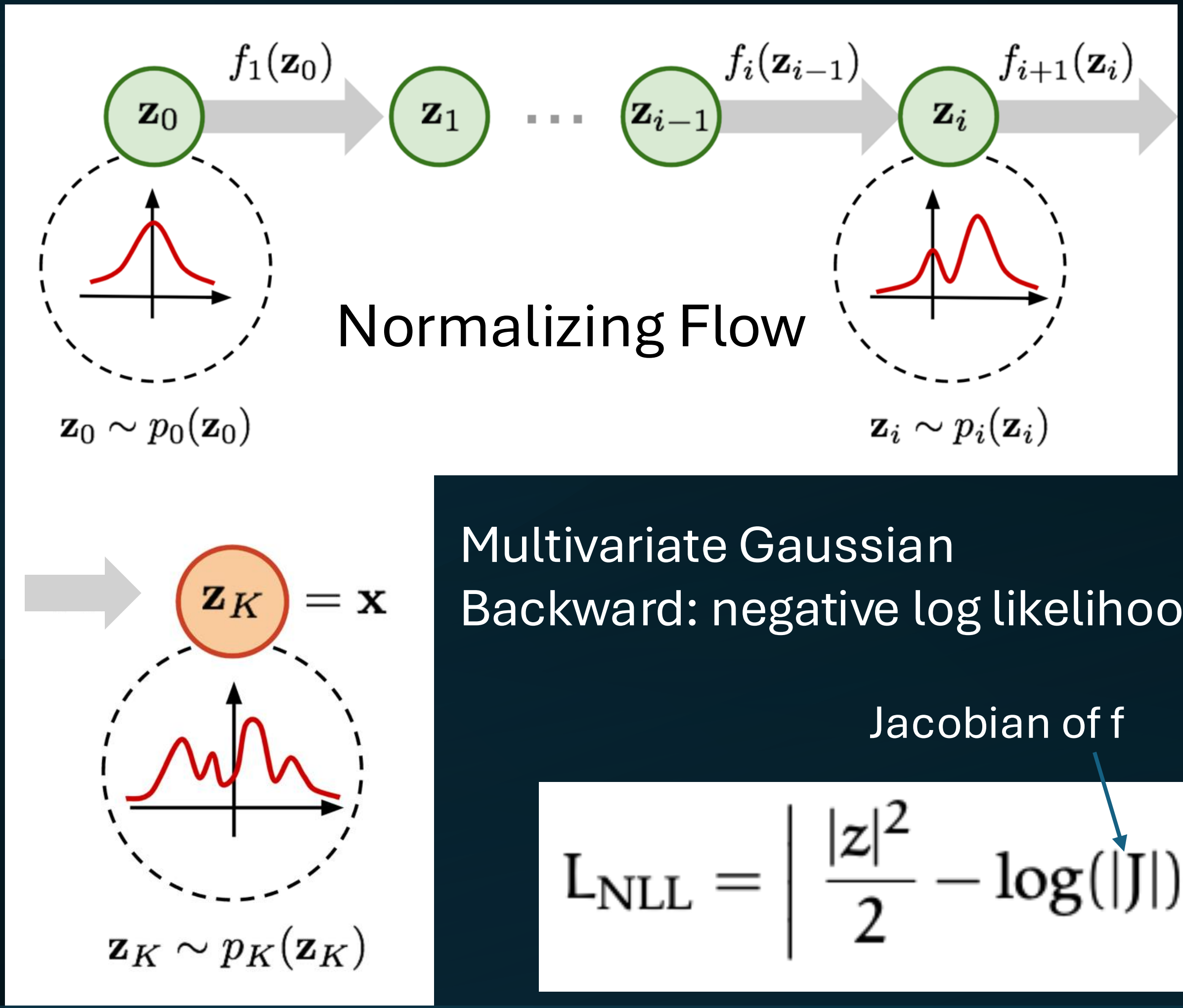
But how?

Conditional Invertible Neural Networks



Posterior

$p(x|c)$



Multivariate Gaussian
Backward: negative log likelihood

Jacobian of f

$$L_{NLL} = \left| \frac{|z|^2}{2} - \log(|J|) \right|$$

Galaxy Cluster Properties observables Or maps

- Scalar Observable Properties
- X-ray maps
- Radio maps

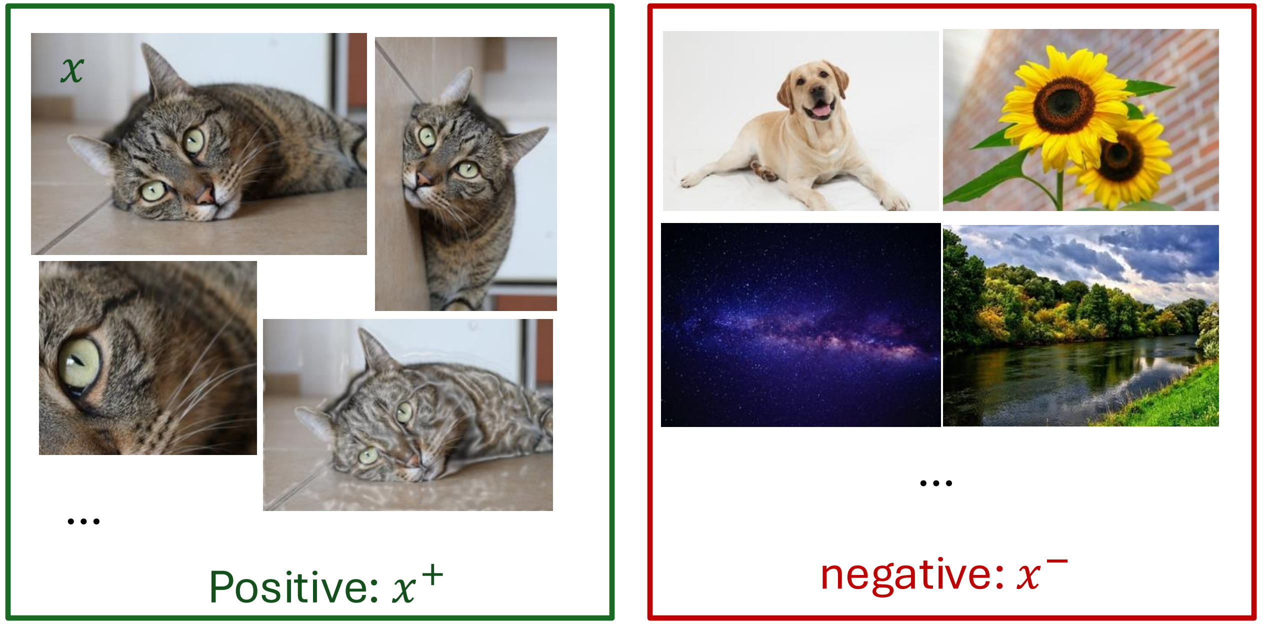
Options for X-ray or Radio maps:

- Putting the map directly into the CINN (D: n x n)
- Reduce its dimension by learning a **representation space (D: $m < n^2$)**

How? ↓

Self Supervised Learning Methods e.g. contrastive learning

Contrastive Learning



$$\text{score}(f(x), f(x^+)) >> \text{score}(f(x), f(x^-))$$

InfoNCE loss

$$L = -\mathbb{E}_X \left[\log \frac{\exp(s(f(x), f(x^+)))}{\exp(s(f(x), f(x^+))) + \sum_{j=1}^{N-1} \exp(s(f(x), f(x_j^-)))} \right]$$

Score of the positive pair

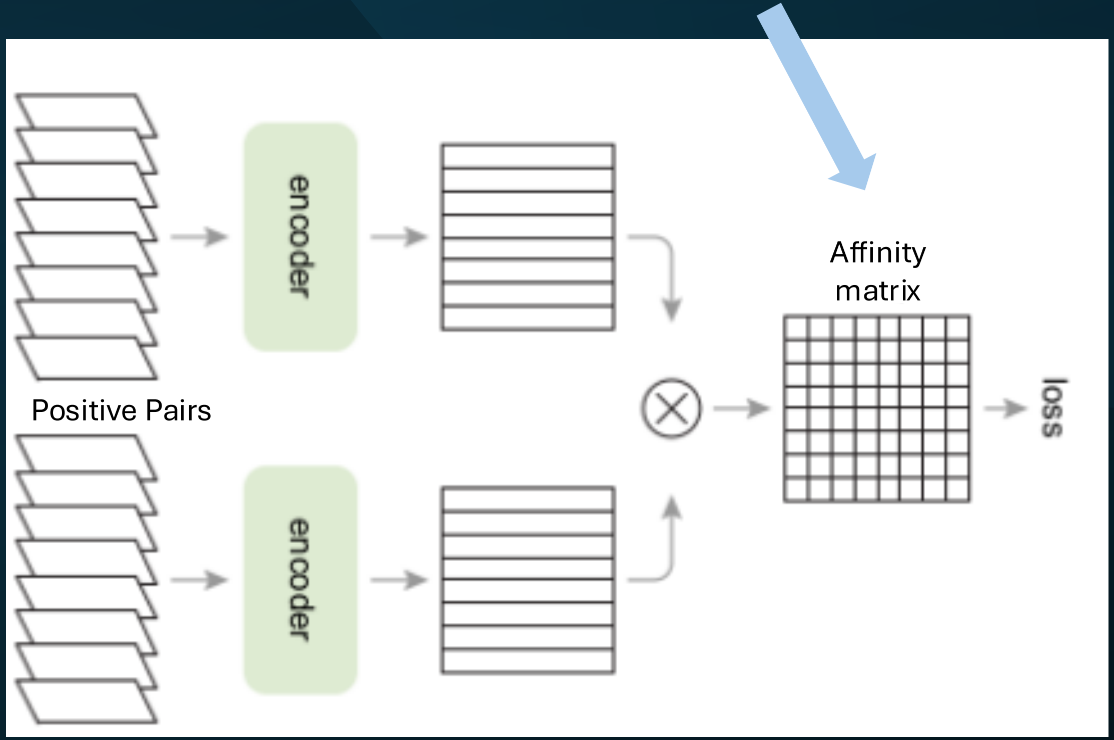
Score of the N-1 Negative pairs

SimCLR:

A Simple Framework for Contrastive Learning

Score function: Cosine similarity

$$s(u, v) = \frac{u^T v}{||u||||v||}$$



SimCLR:

A Simple Framework for Contrastive Learning

Score function: Cosine similarity

$$s(u, v) = \frac{u^T v}{||u|| ||v||}$$

In the end:

You will have a representation space with similar pictures close to each other in m dimensions (256, 512, ...)

In case of having access to labels, representation space can be tested directly via nearest neighboring, UMAPS, or hexbin color-coded charts by the quantitive labels

Road Map

Goal: Inferring the merger history of Galaxy Clusters in TNG-Cluster

Use CINN to obtain full posteriors on merger parameters → Input: Learned Representation Space → X-Ray maps / Radio maps

Questions

What yields the most accurate inference of merger properties ?

Radio maps or X-ray maps alone, Radio + X-ray?

Summary

1. Galaxy clusters are the ultimate result of hierarchical structures in the ΛCDM cosmological frames.
2. Most of the mass of the ICM is hot gas which results in Bremstrehlung process producing X-ray emission.
3. Cluster mergers rank among the most energetic events since the Big Bang . This process accelerates the thermal electrons in the MF of the ICM, resulting in Synchrotron emission:

   mergers { shocks → Polarised Radio Relics; Turbulence → Diffuse Radio Halos }

4. By applying machine learning on the simulation data, we can learn complex mappings between the observable galaxy clusters properties or maps, to their underlying merger histories.
5. X-ray and radio maps are first passed through a contrastive-learning encoder, building a low dimensional representation space, distilling high-resolution maps into feature vectors that capture the relevant structure.
6. Using either the scalar observables or the learned image embeddings as inputs, CINNs perform fast, exact Bayesian inference: returning full posterior distributions over the unobservable merger parameters.