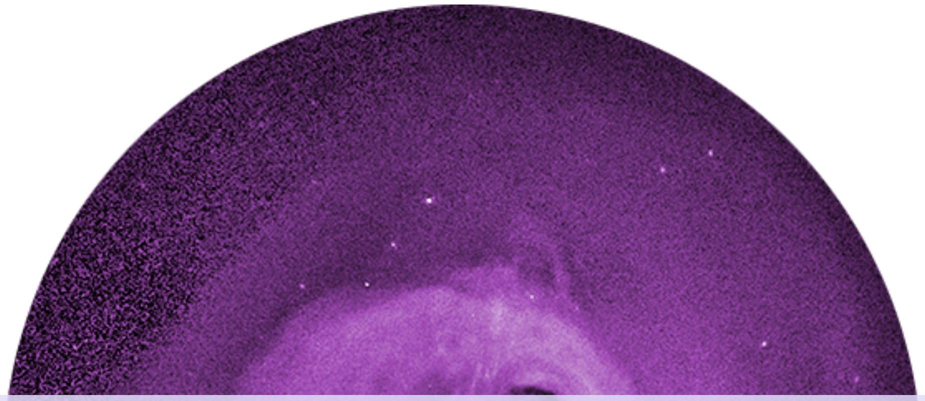


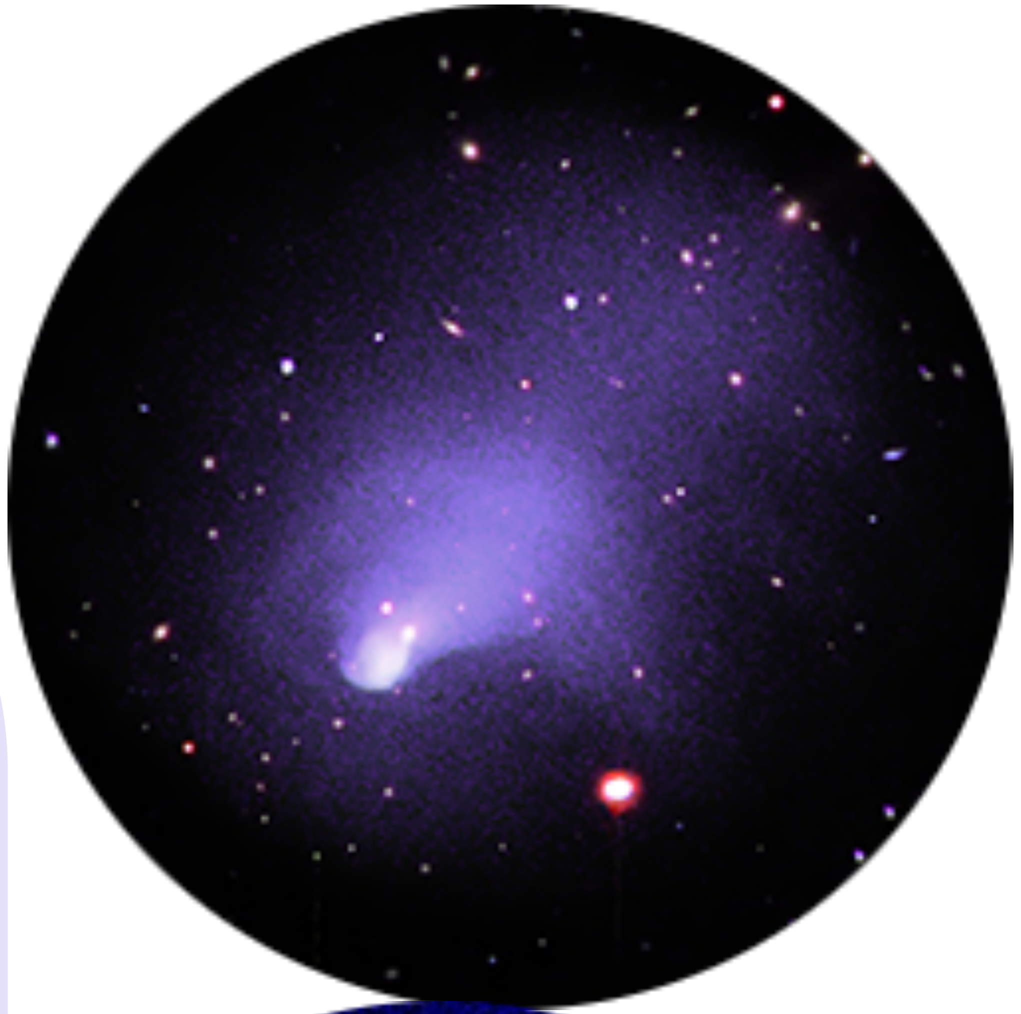


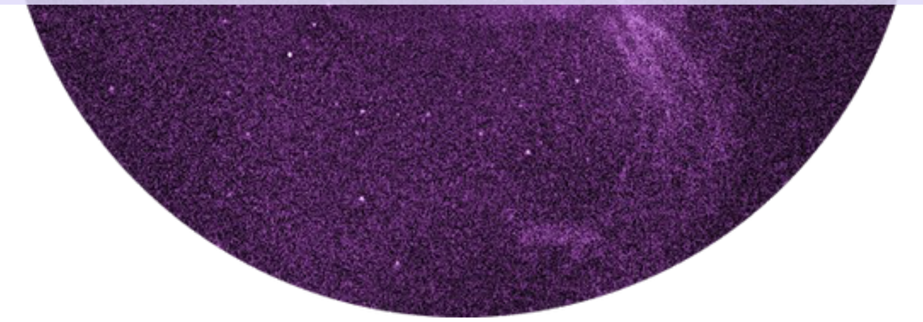
IMAGE-LEVEL INFERENCE OF COSMOLOGY & ASTROPHYSICS WITH GALAXY CLUSTERS

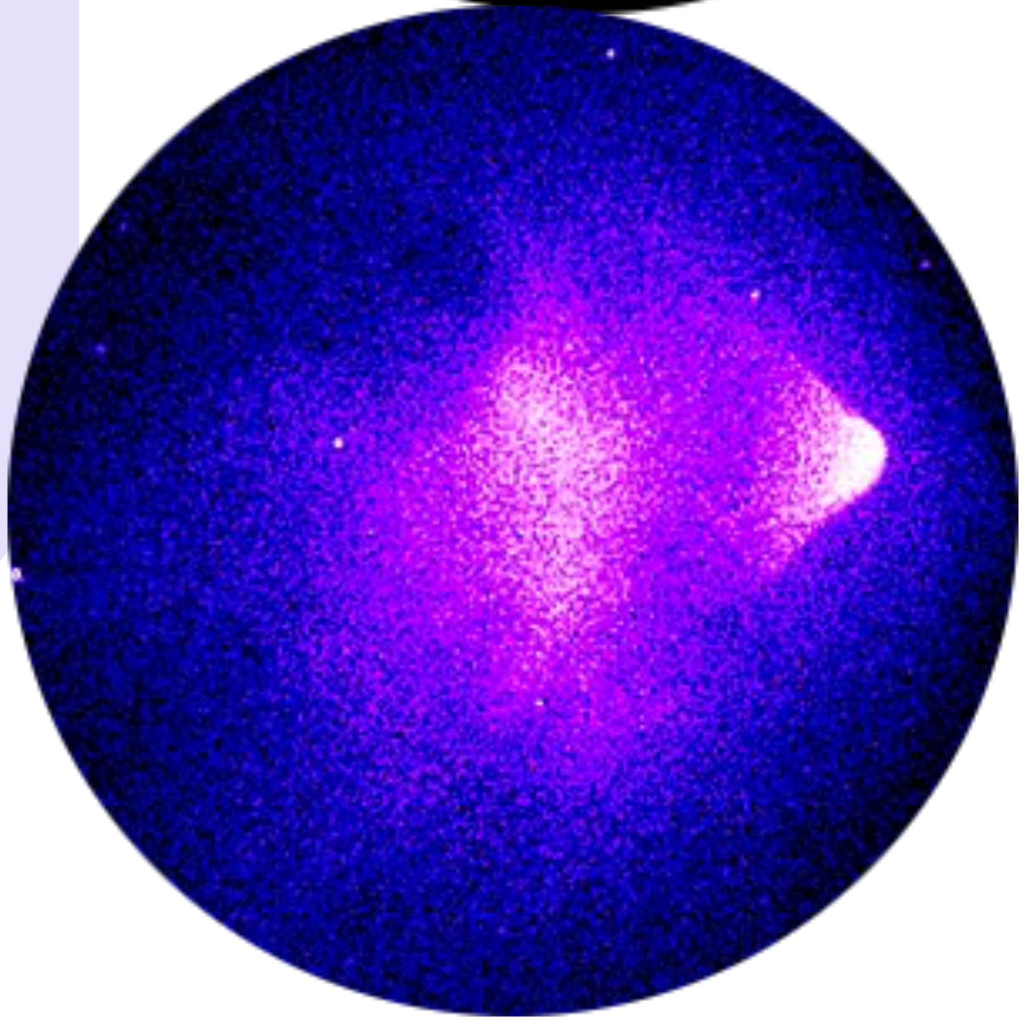
Urmila Chadayammuri

Postdoctoral Fellow | MPIA

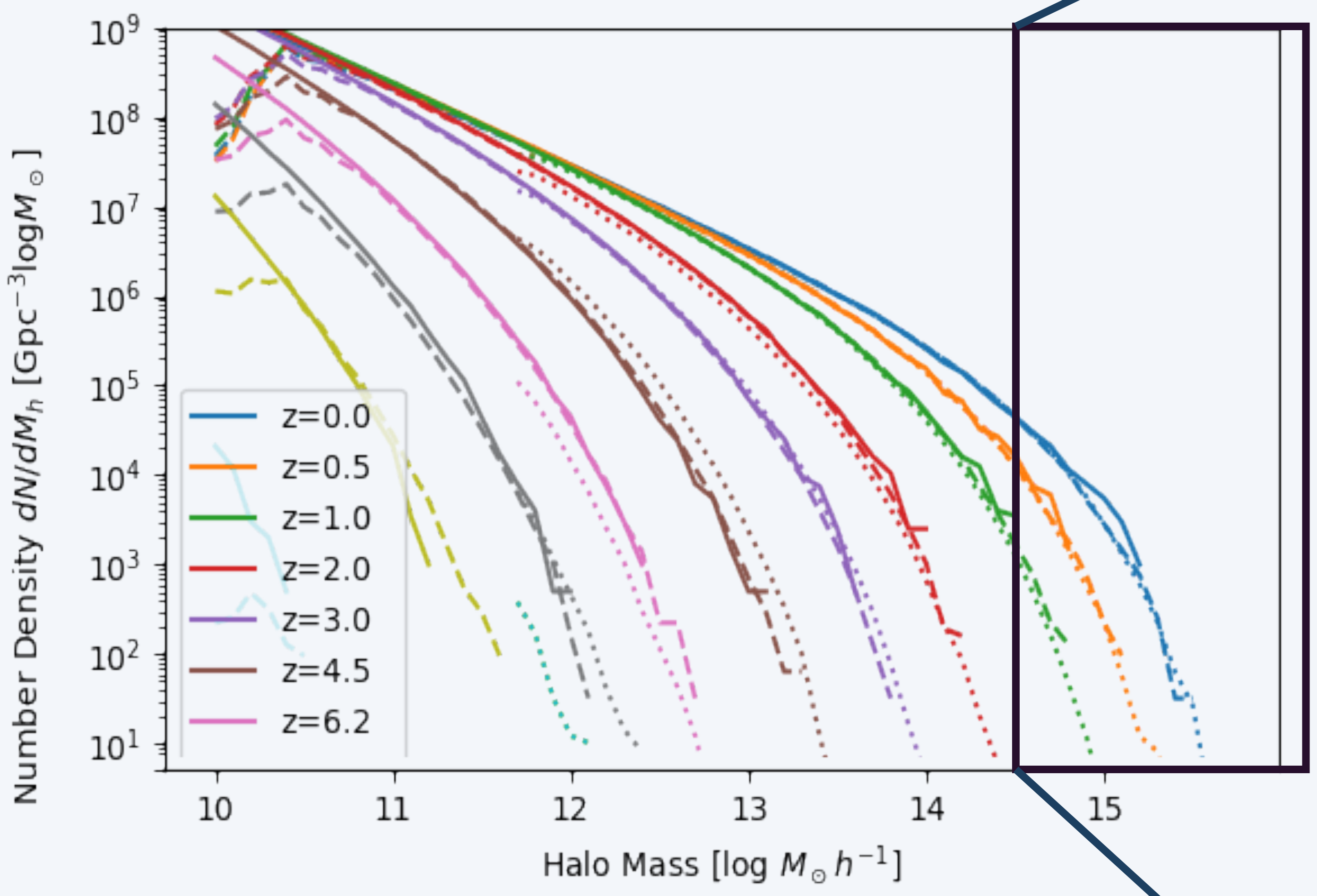
ML4Astro II | Catania, Italy

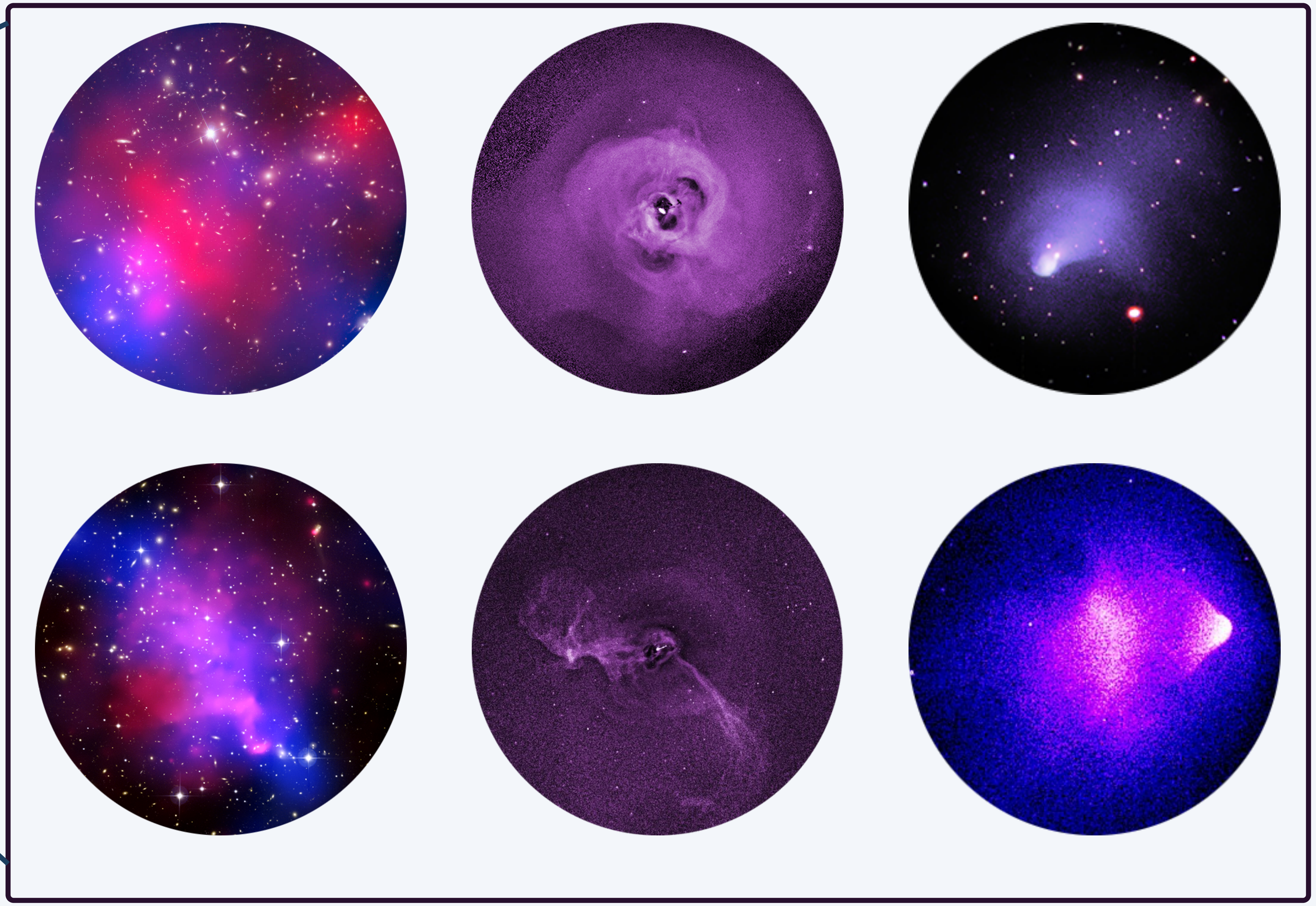
July 10, 2024





Galaxy Clusters are the biggest things at z = 0





Used to measure cosmology

And astrophysics

But the inference pipeline is complex and degenerate

X-ray SB, kT

Y_{SZ}

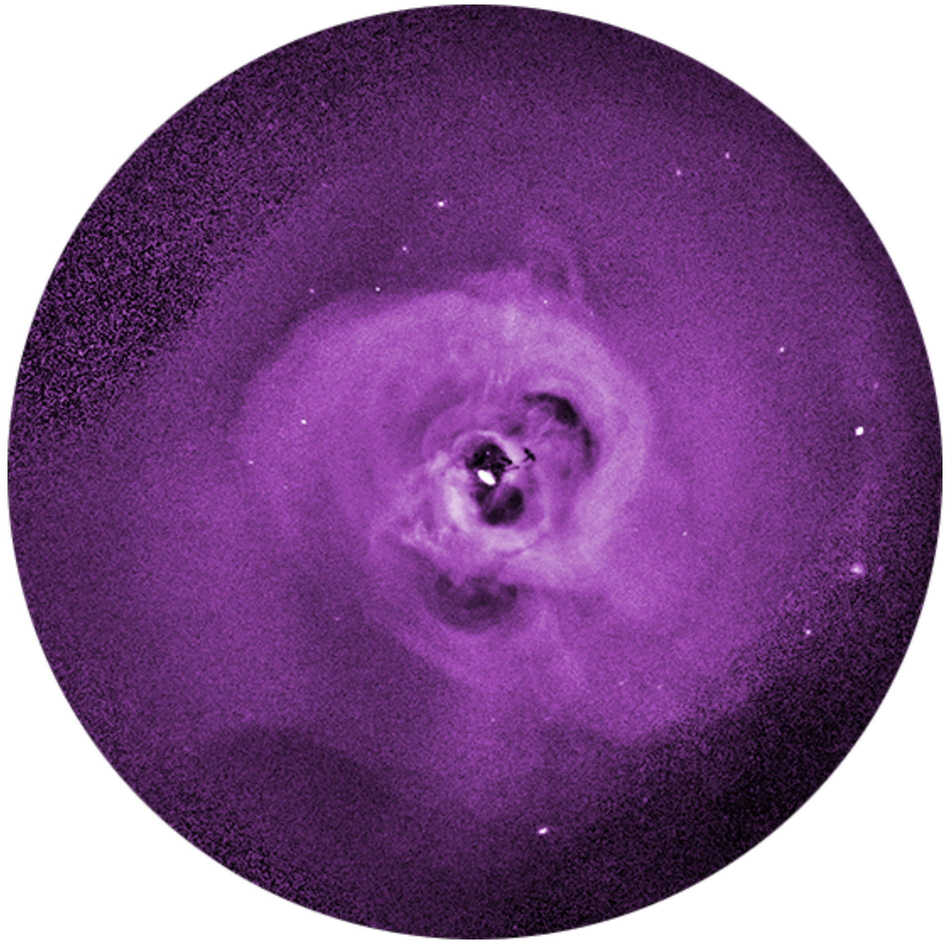
Galaxy spectra/colours

Lensing κ

Analytic arguments (assume HSE) / Parameter searches with idealised simulations

$$\left\{ \begin{matrix} M_{gas} \\ M_{vir} \\ kT \\ Z \\ R, t_{merge} \end{matrix} \right\}$$

But the inference pipeline is complex and degenerate



X-ray SB, kT

Y_{SZ}

Galaxy spectra/colours

Lensing $\varkappa$

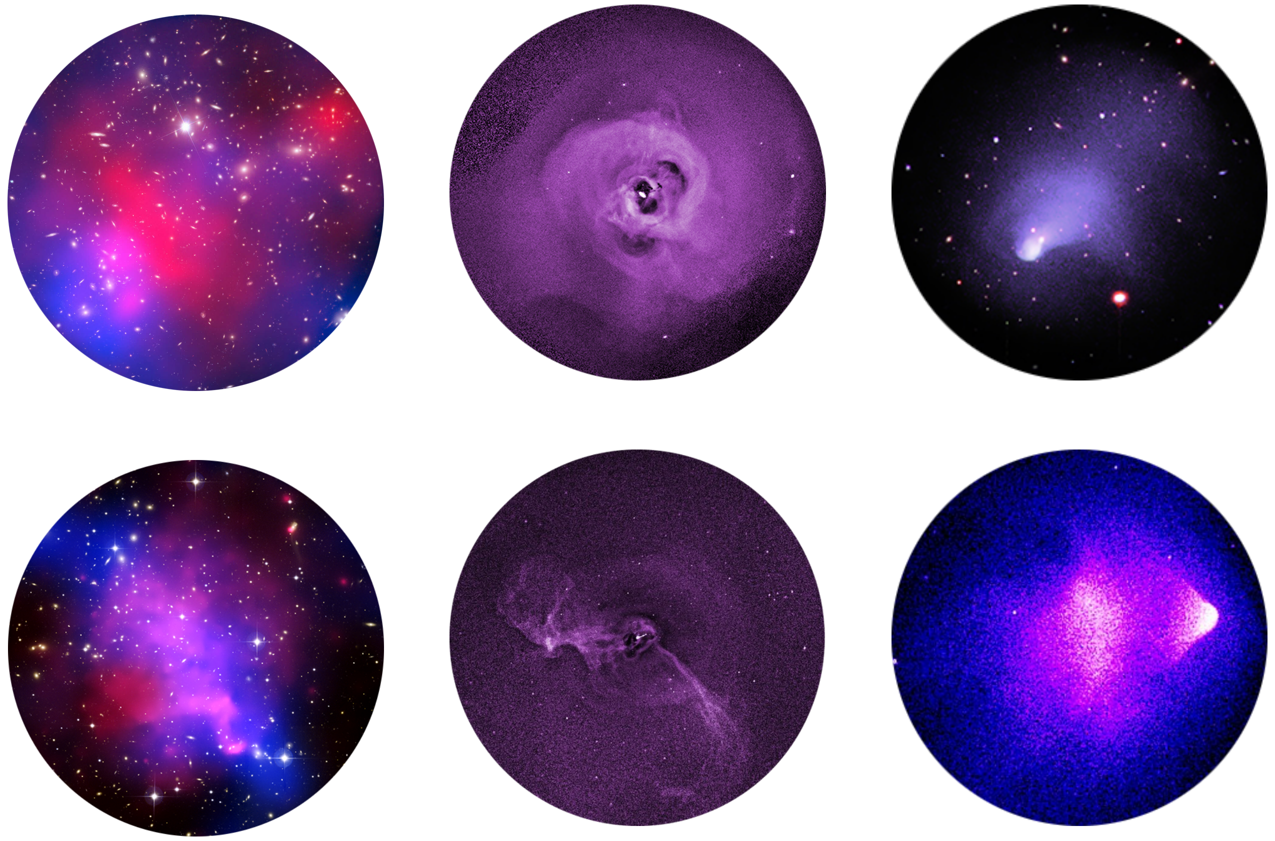
Analytic arguments (assume HSE)

/

Parameter searches with idealised simulations

$$\left\{ \begin{array}{l} M_{gas} \\ M_{vir} \\ kT \\ Z \\ R, t_{merge} \\ t_{AGN}, E_{AGN} \end{array} \right\}$$

But the inference pipeline is complex and degenerate



Galaxy cluster sample

Analytic arguments (assume HSE) / Parameter searches with idealised simulations

$\{M_{vir}\}$

Library of HMFs from N-body sims / analytic models

$\{\Omega_m, \sigma_8, H_0, w_\Lambda \ldots\}$

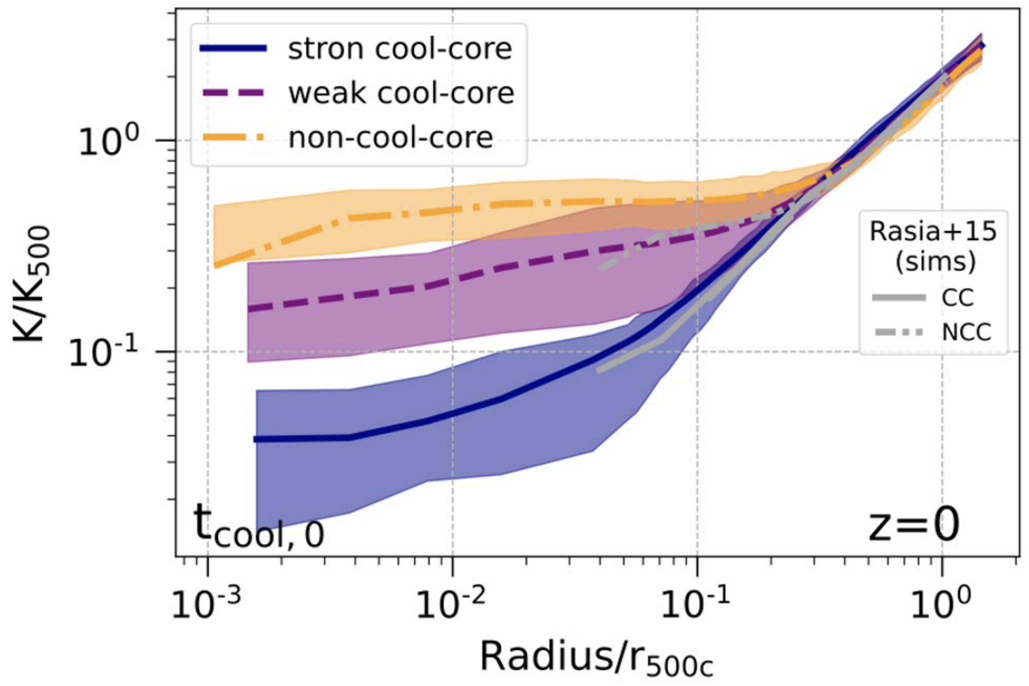
It requires us to reduce complex images to scalars or azimuthally averaged profiles

X-ray SB, kT

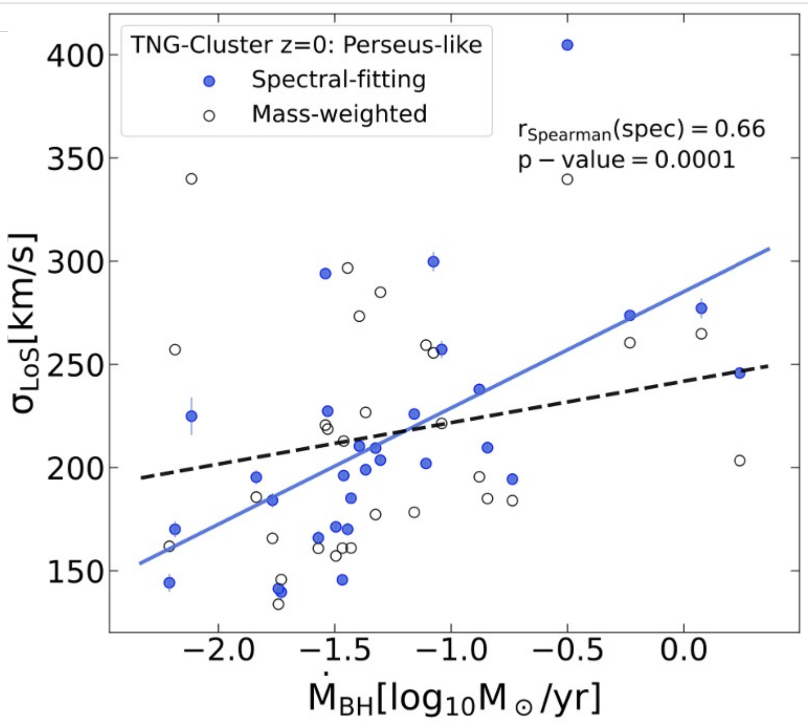
Y_{SZ}

Galaxy spectra/colours

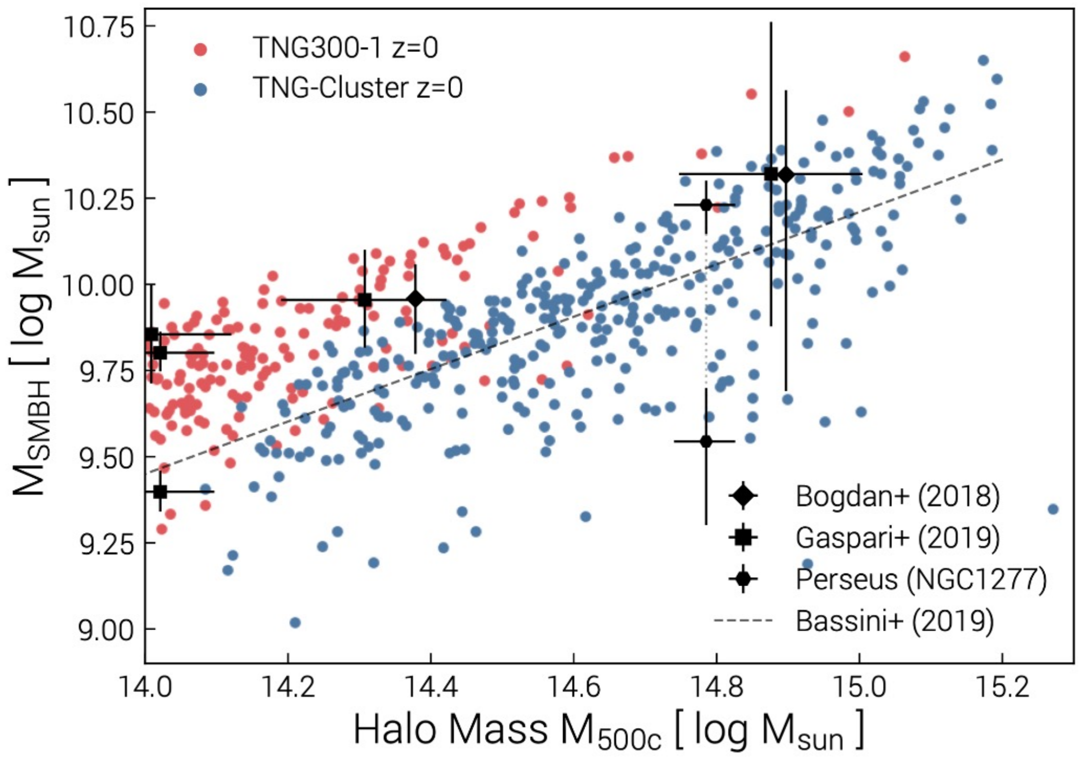
Lensing $\varkappa$



Lehle+ 2024



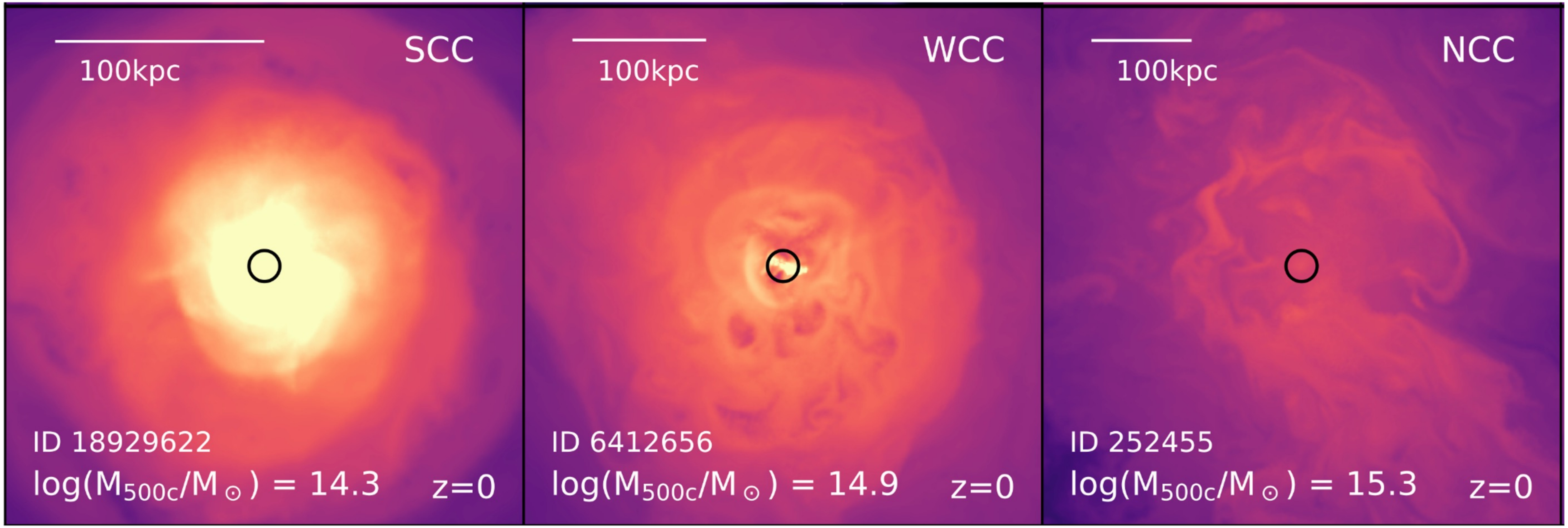
Truong+ 2024



Nelson+ 2024

When really, clusters are diverse

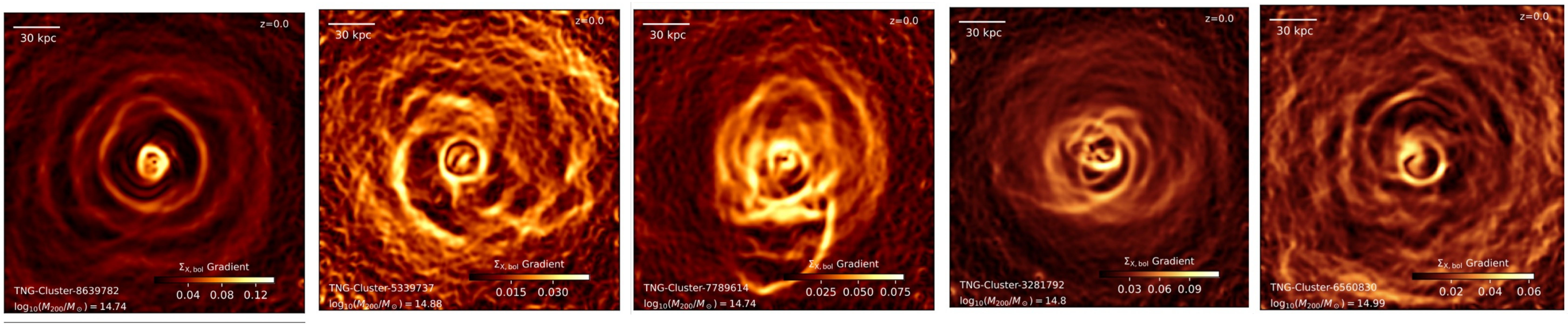
They have a wide variety of core thermodynamic profiles



Lehle+ 2024

When really, clusters are diverse

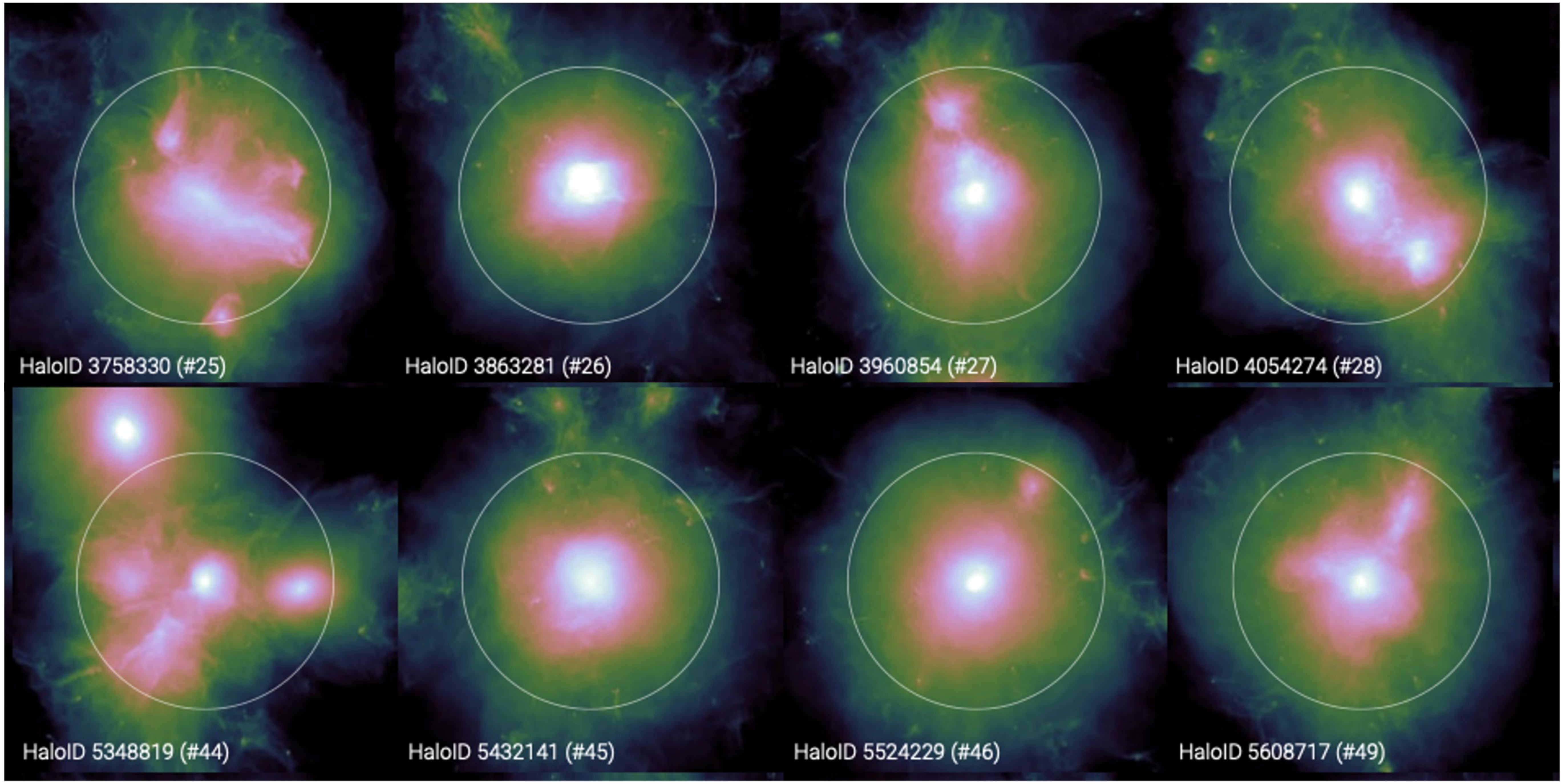
They have a wide variety of histories of AGN activity, in different phases



Truong+ 2024

When really, clusters are diverse

They can be in very different stages of assembly

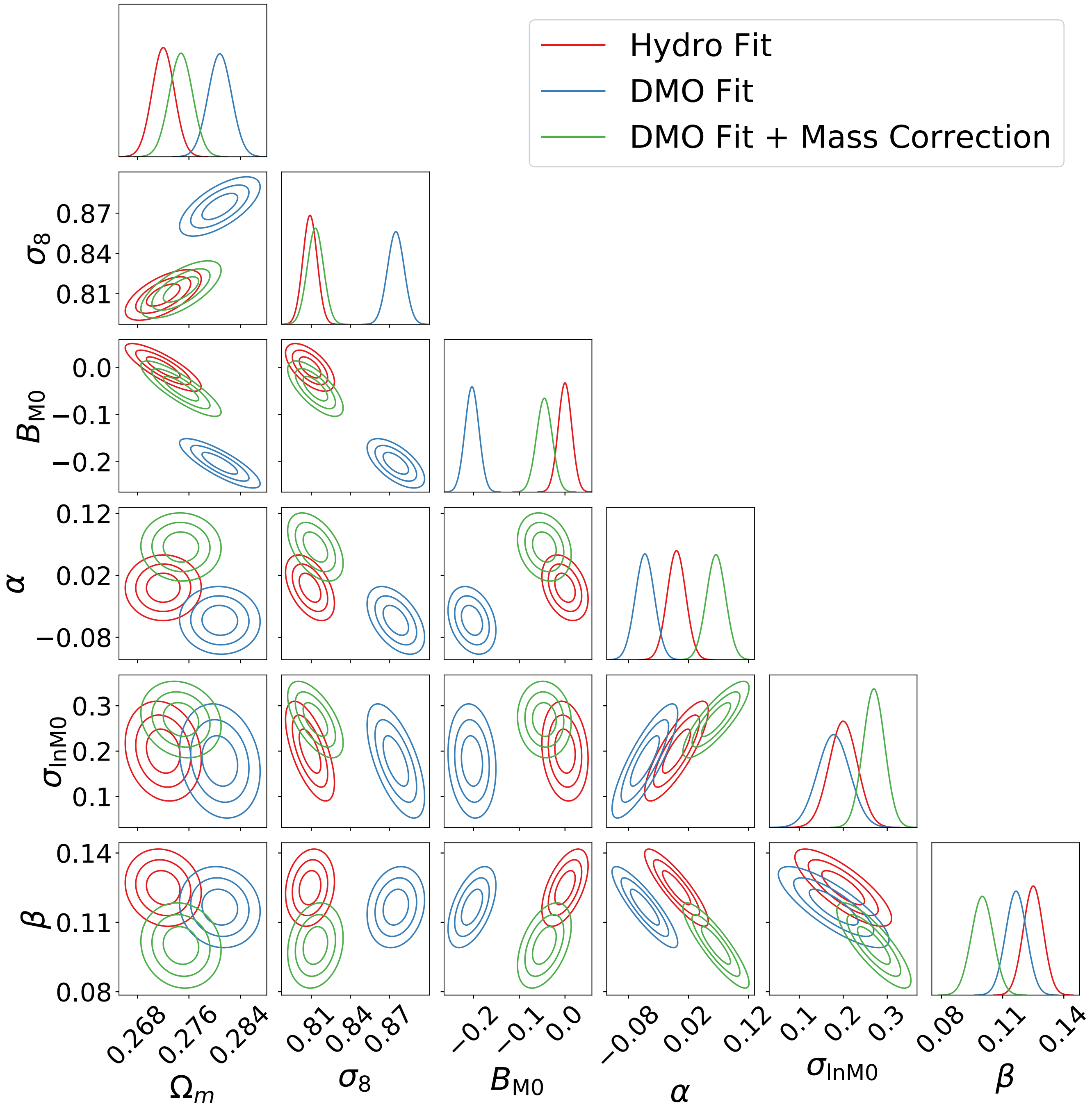


Nelson+ 2024

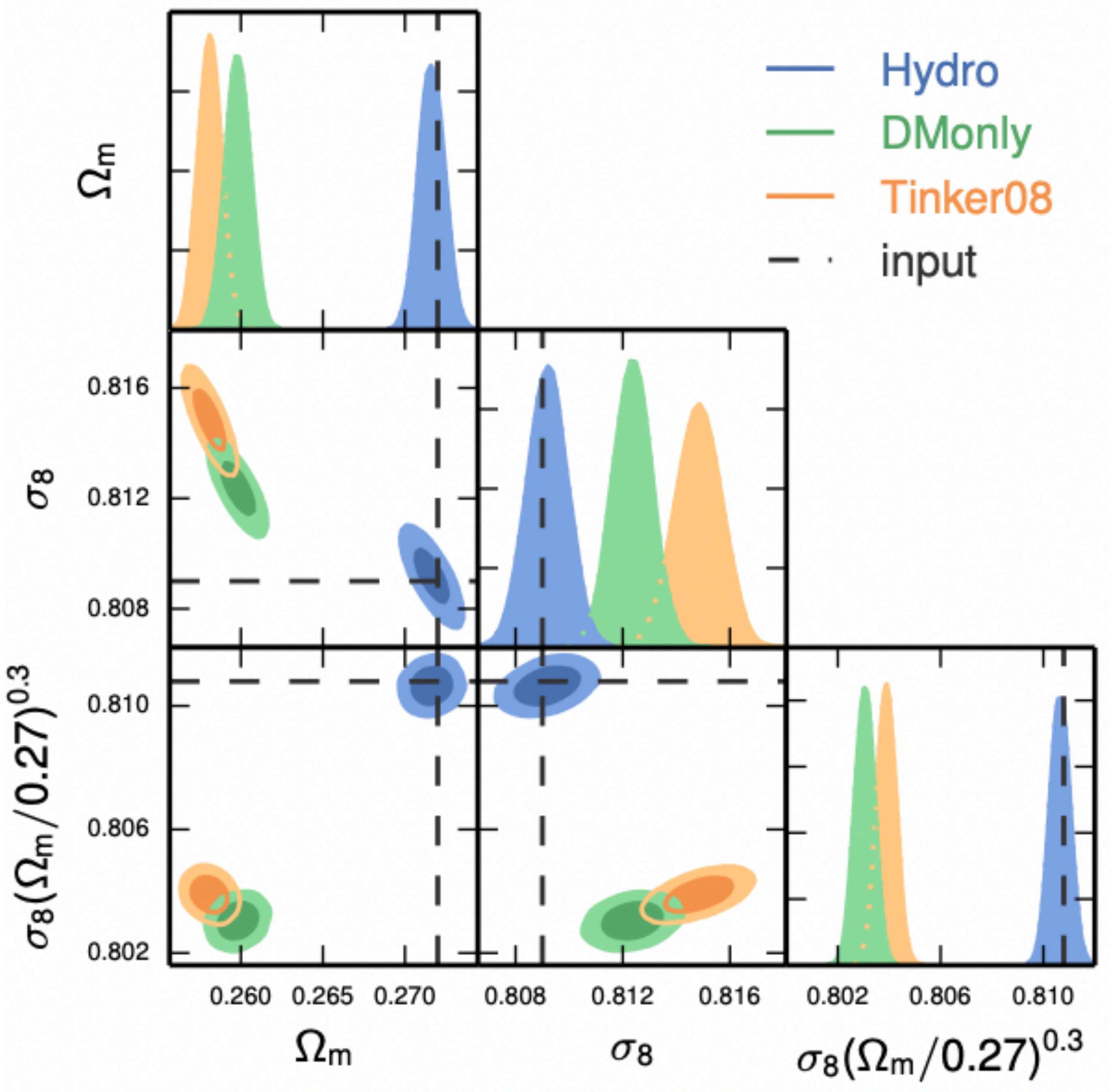
How can we **infer cosmological and astrophysical parameters from images** of the intracluster medium with **minimal information loss?**

1. Paint baryons onto N-body predictions

Because Halo Mass Function not identical with and without baryons



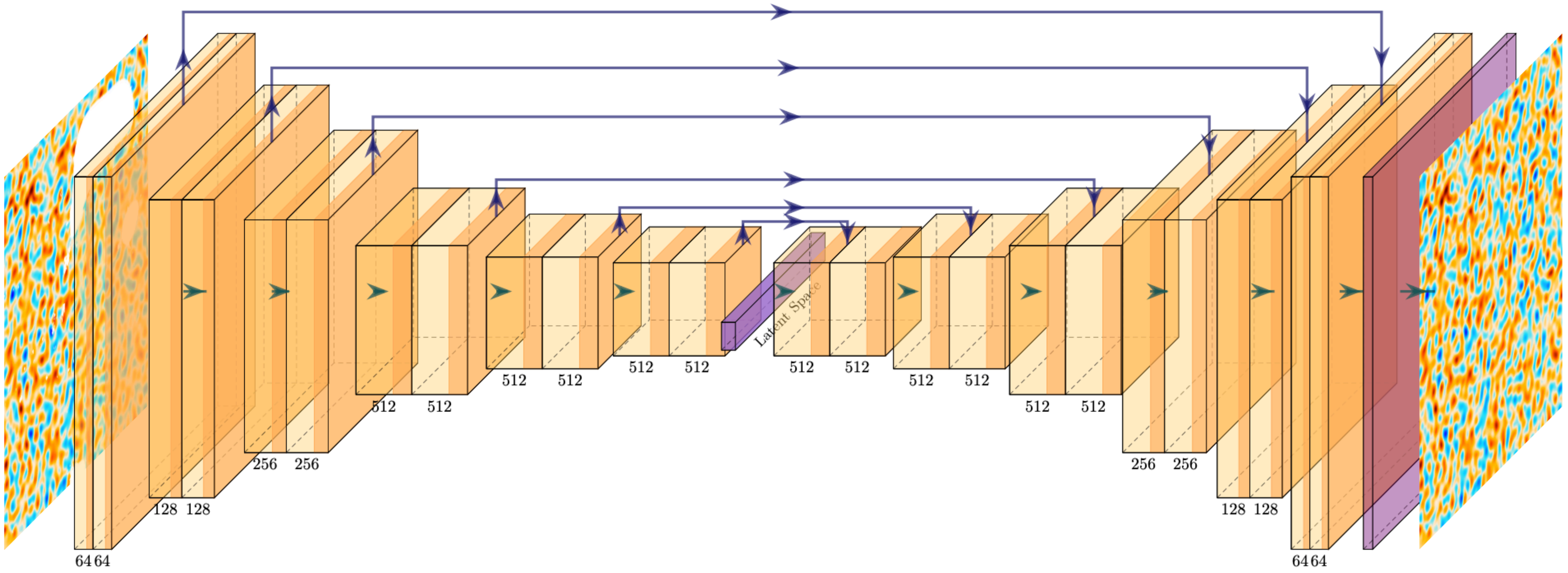
Castro+ 2021 (Euclid mocks)



Bocquet+ 2016 (eROSITA mocks)

Start simple: CNNs

- Traditionally used for classification tasks: input image —> scalar output
- But can just as easily be used to go image —> image! This is in fact the “autoencoder” or “U-Net” architecture used for image colorisation, etc.



Training sample: TNG300

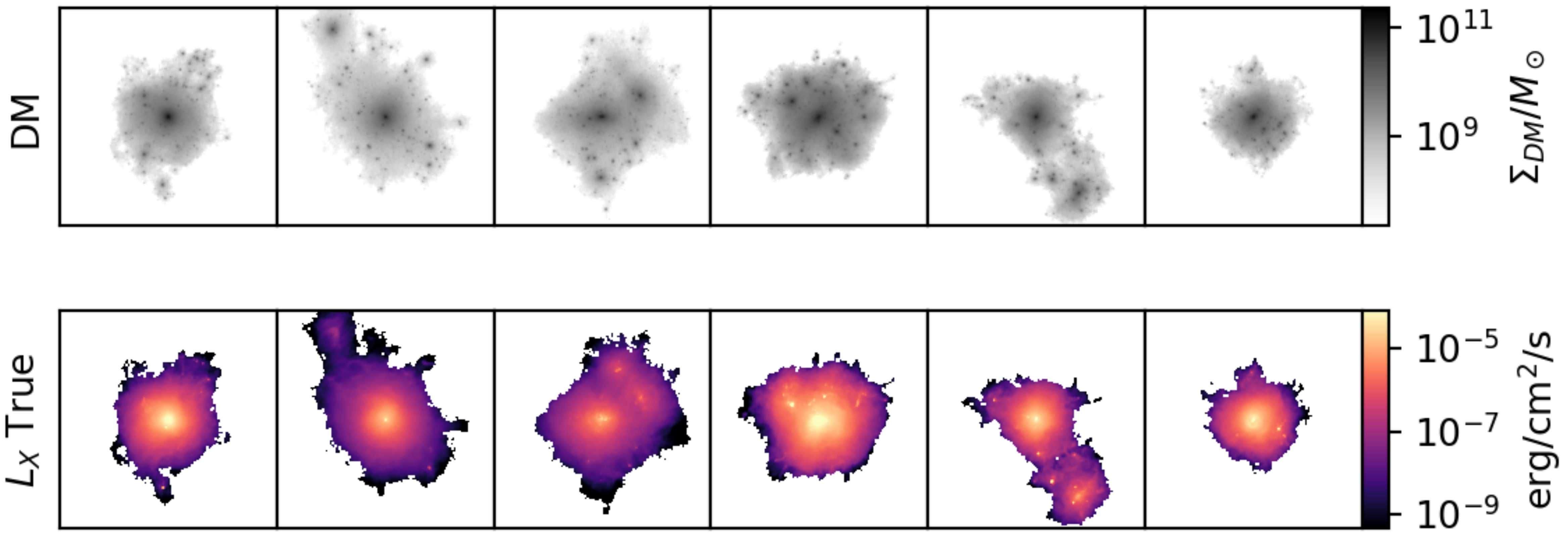
- Everything with $M_{FoF} > 10^{14} M_{\odot}$ in FP
- DM mass → gas column density,

  → projected temperature

  → X-ray surface brightness

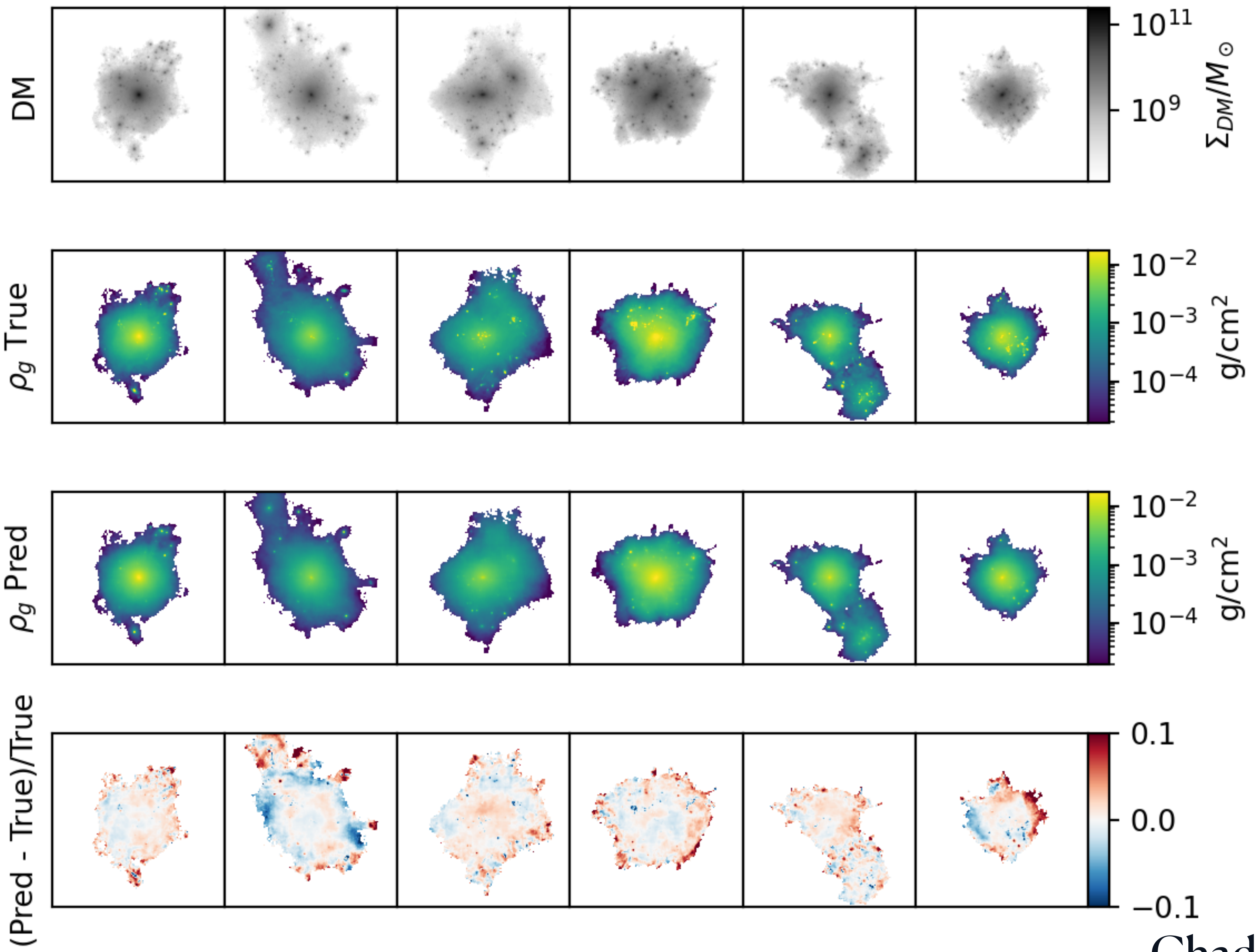
from TNG300-1

(highest resolution, full-physics)



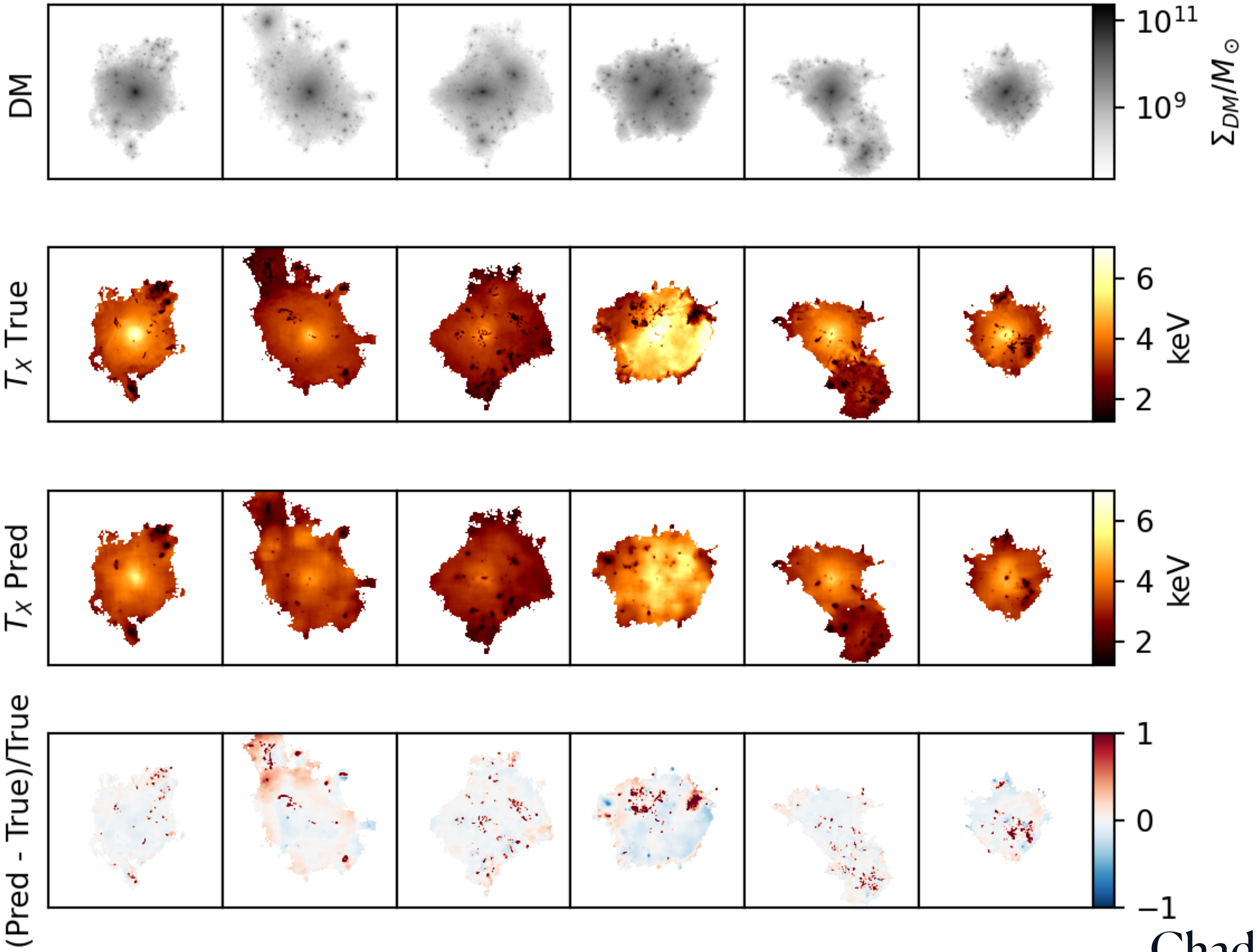
Chadayammuri+ 2023

$$M_{DM} \rightarrow \Sigma_g$$



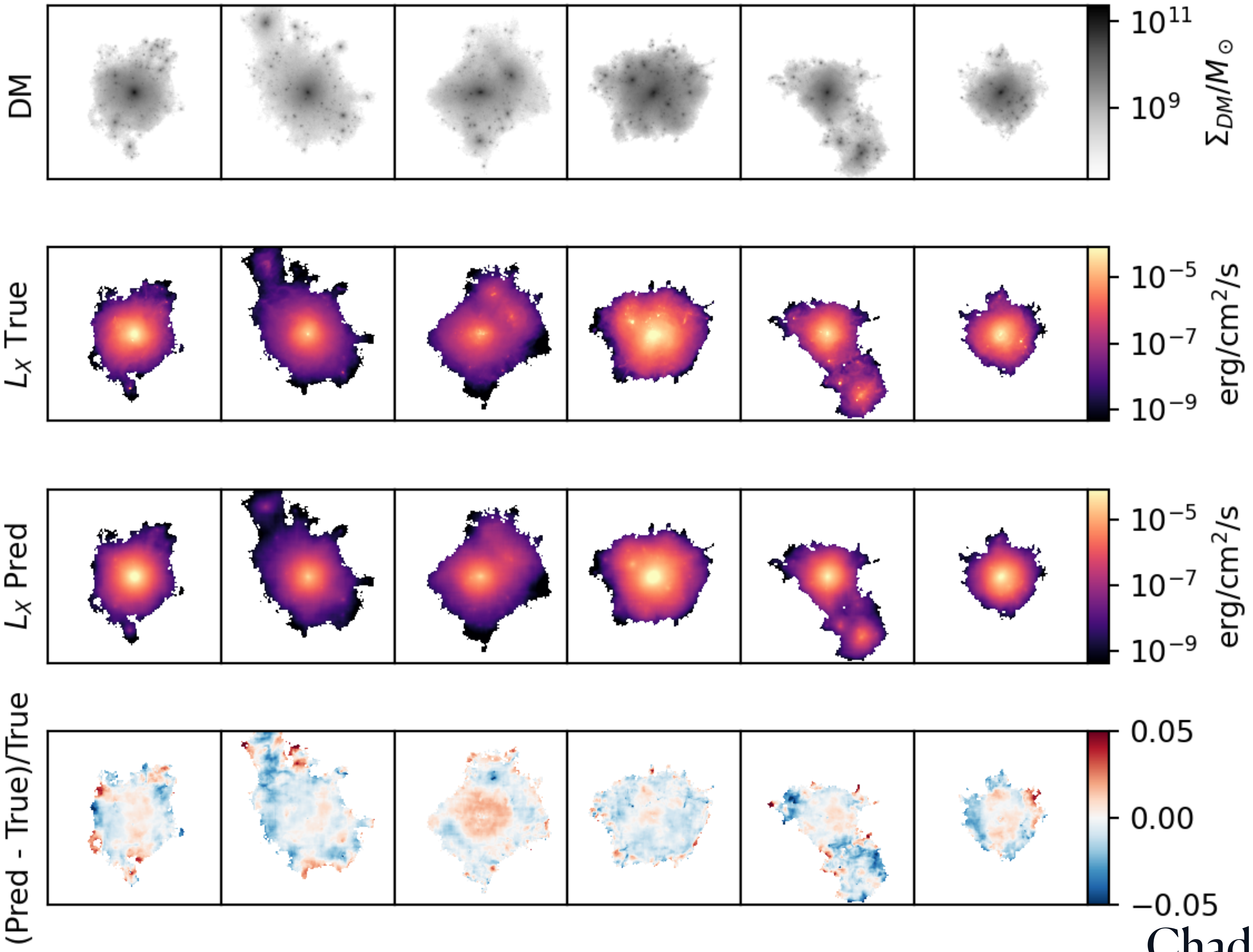
Chadayammuri+ 2023

$$M_{DM} \rightarrow T_X$$



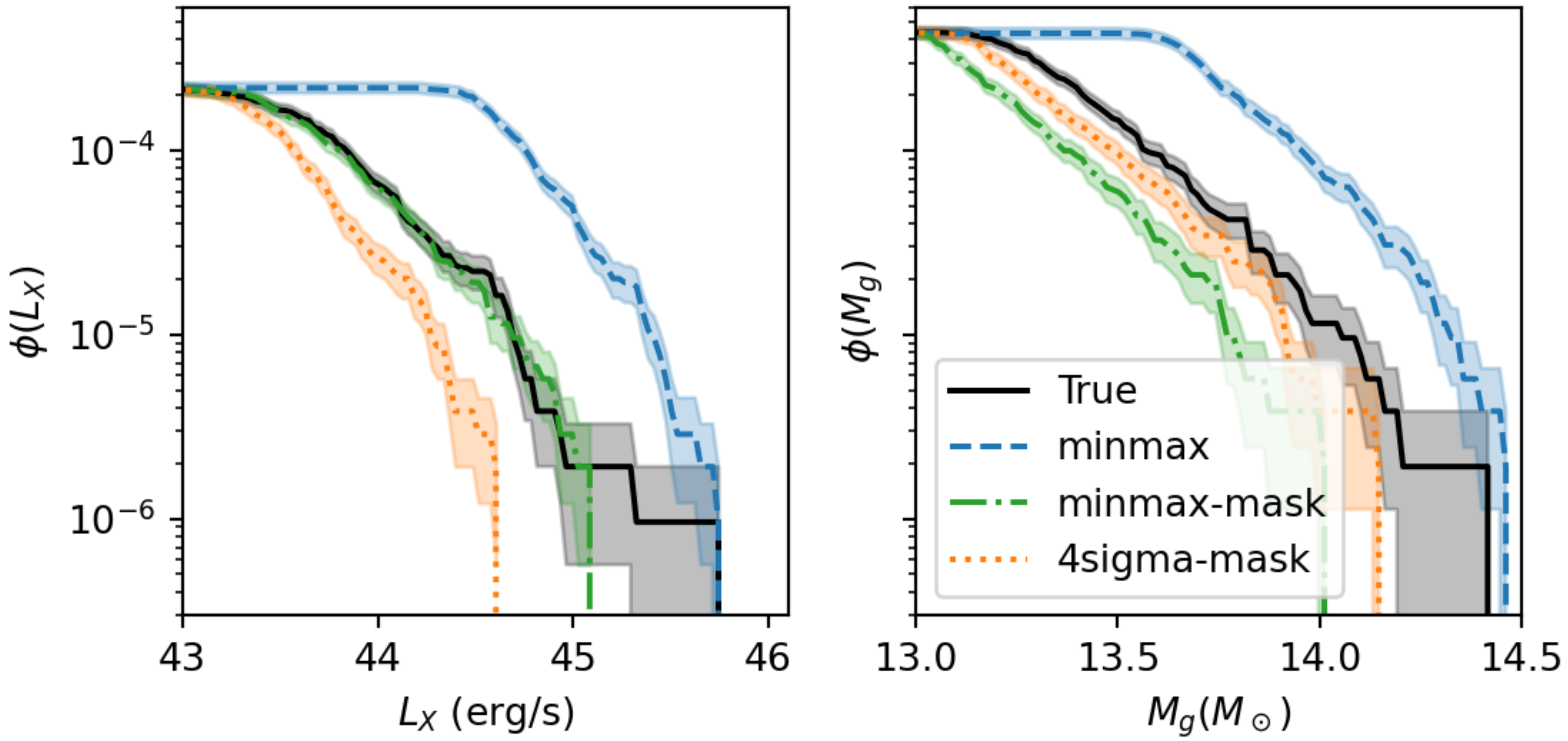
Chadayammuri+ 2023

$M_{DM} \rightarrow L_X$



Chadayammuri+ 2023

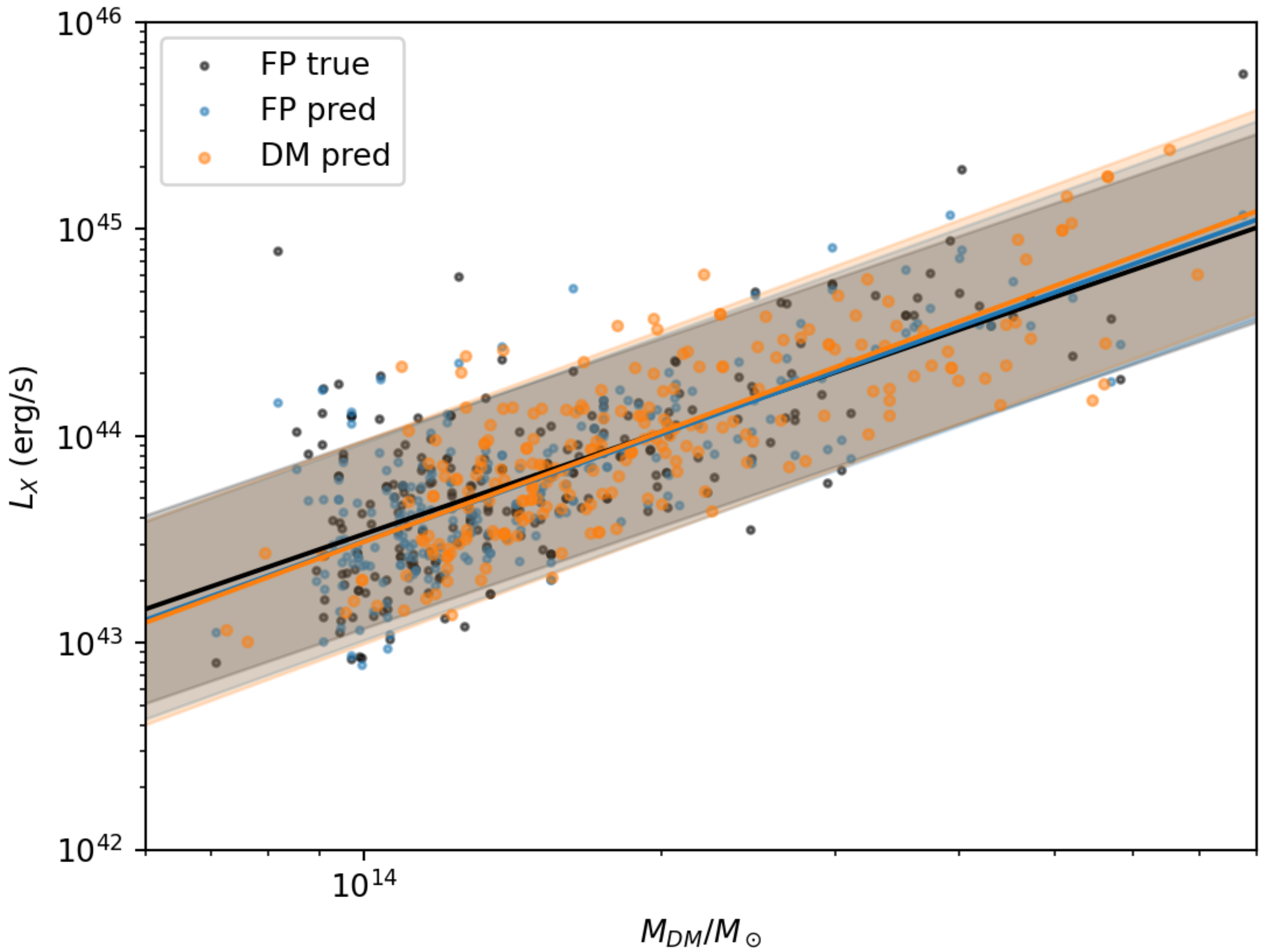
Training on image level, but population statistics recovered



L_x predicted over 3 OOM for 1 OOM in halo mass

Chadayammuri+ 2023

Training on image level, but population statistics recovered



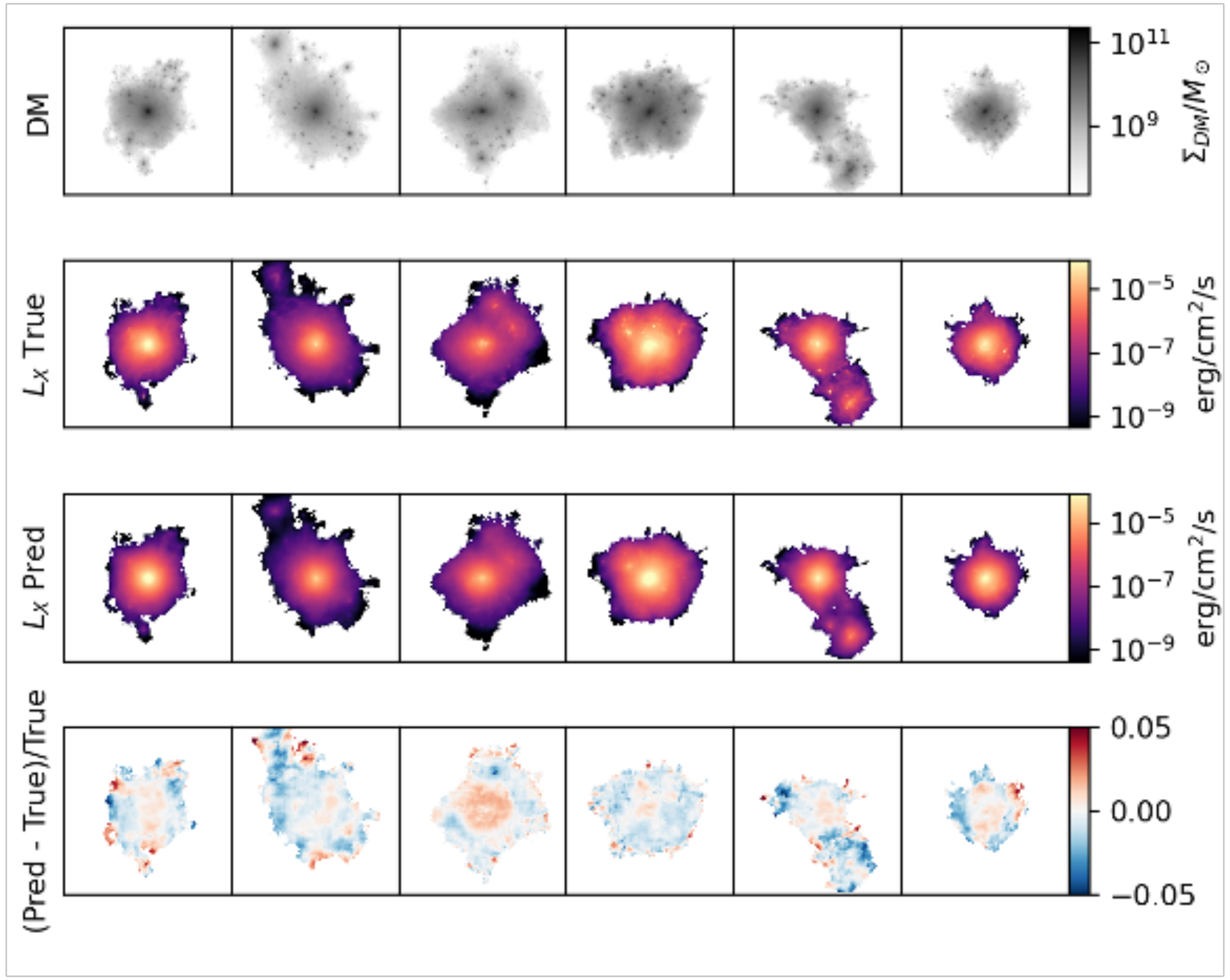
L_X predicted over 3 OOM for 1 OOM in halo mass

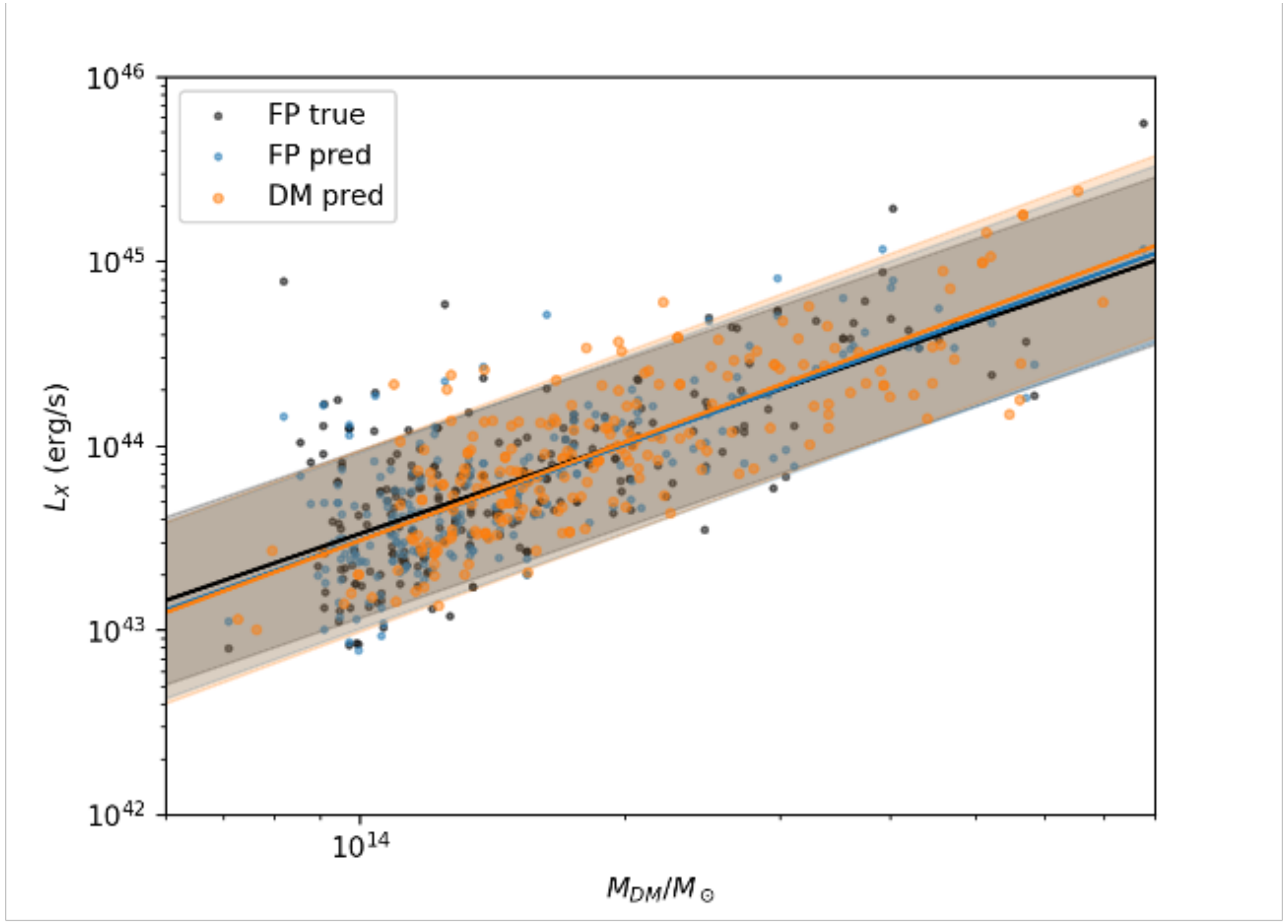
+ we recover the scatter in scaling relations!

Chadayammuri+ 2023

So what can we learn from galaxy cluster images?

1. CNNs allow us to predict ICM observables from DM-only simulations at the % level, and reproduce scatter in scaling relations





Chadayammuri+ 2023

So what can we learn from galaxy cluster images?

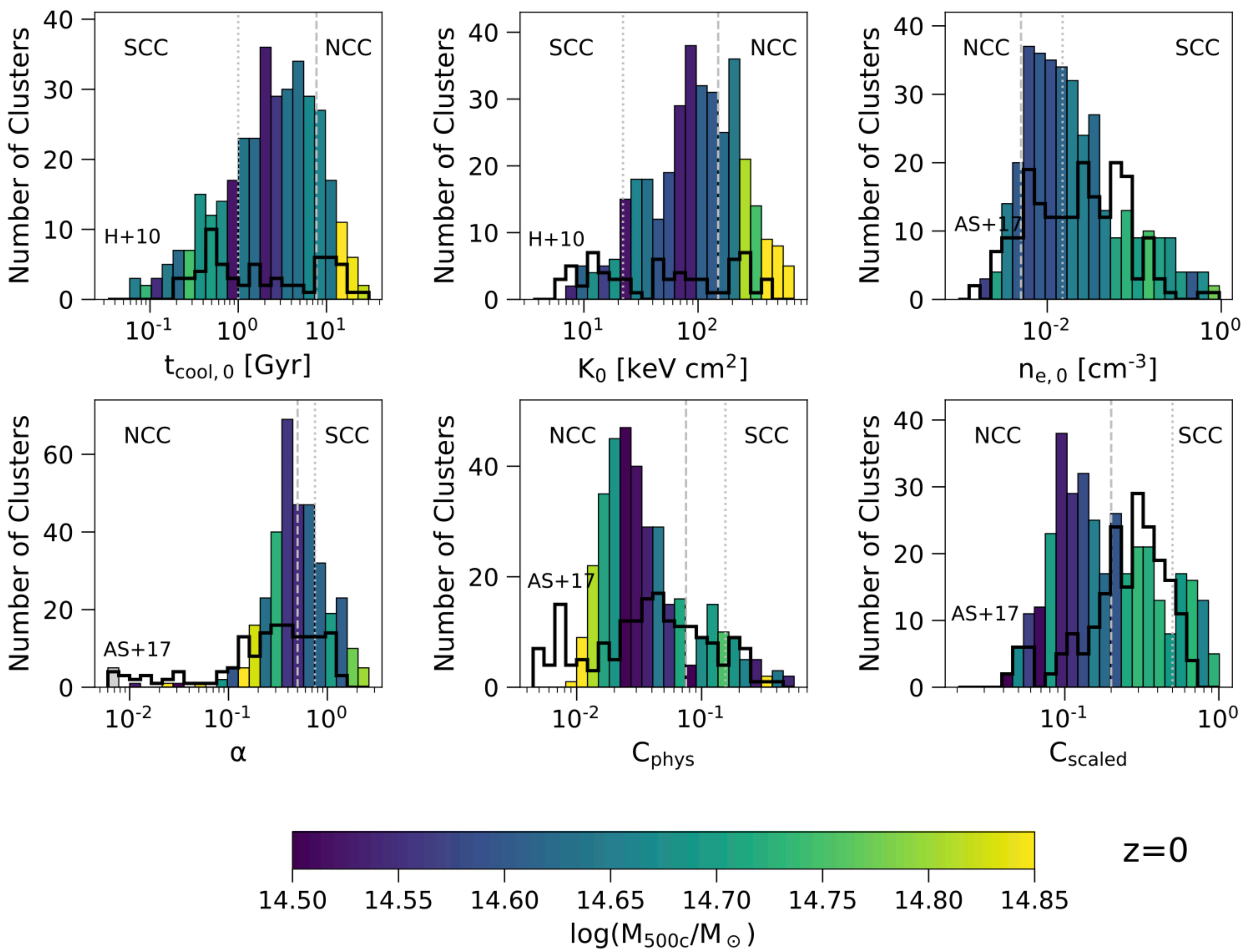
(We can also tackle the inverse problem)

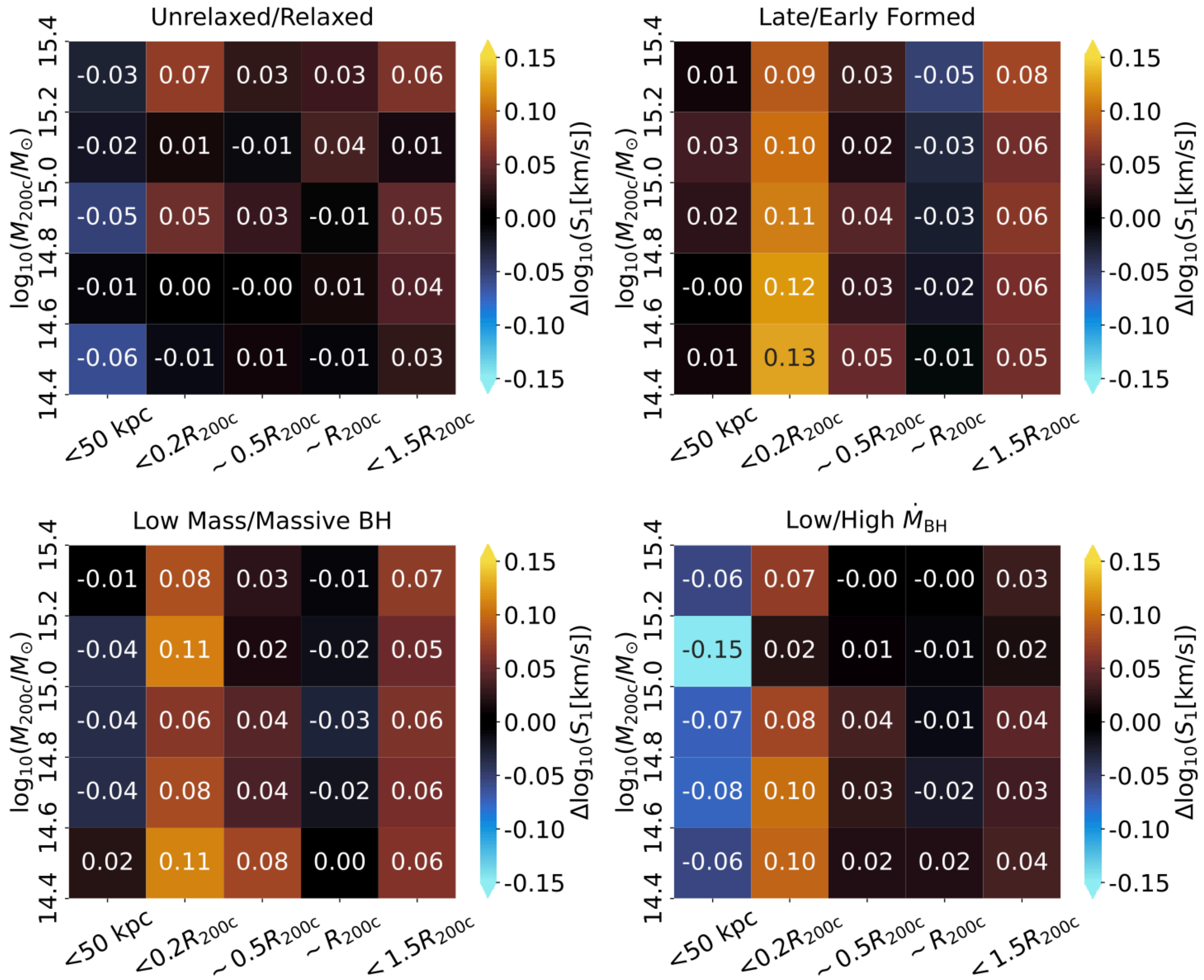
The Three Hundred Project: Mapping The Matter Distribution in Galaxy Clusters Via Deep Learning from Multiview Simulated Observations

Daniel de Andres[1,2★], Weiguang Cui[3,1,2†], Gustavo Yepes[1,2], Marco De Petris[4], Antonio Ferragamo[4,5], Federico De Luca[4], Gianmarco Aversano[6] and Douglas Rennehan[7]

De Andres+ 2024

Next: how do we capture the diversity of galaxy clusters?





Quantifying "cool core-ness"

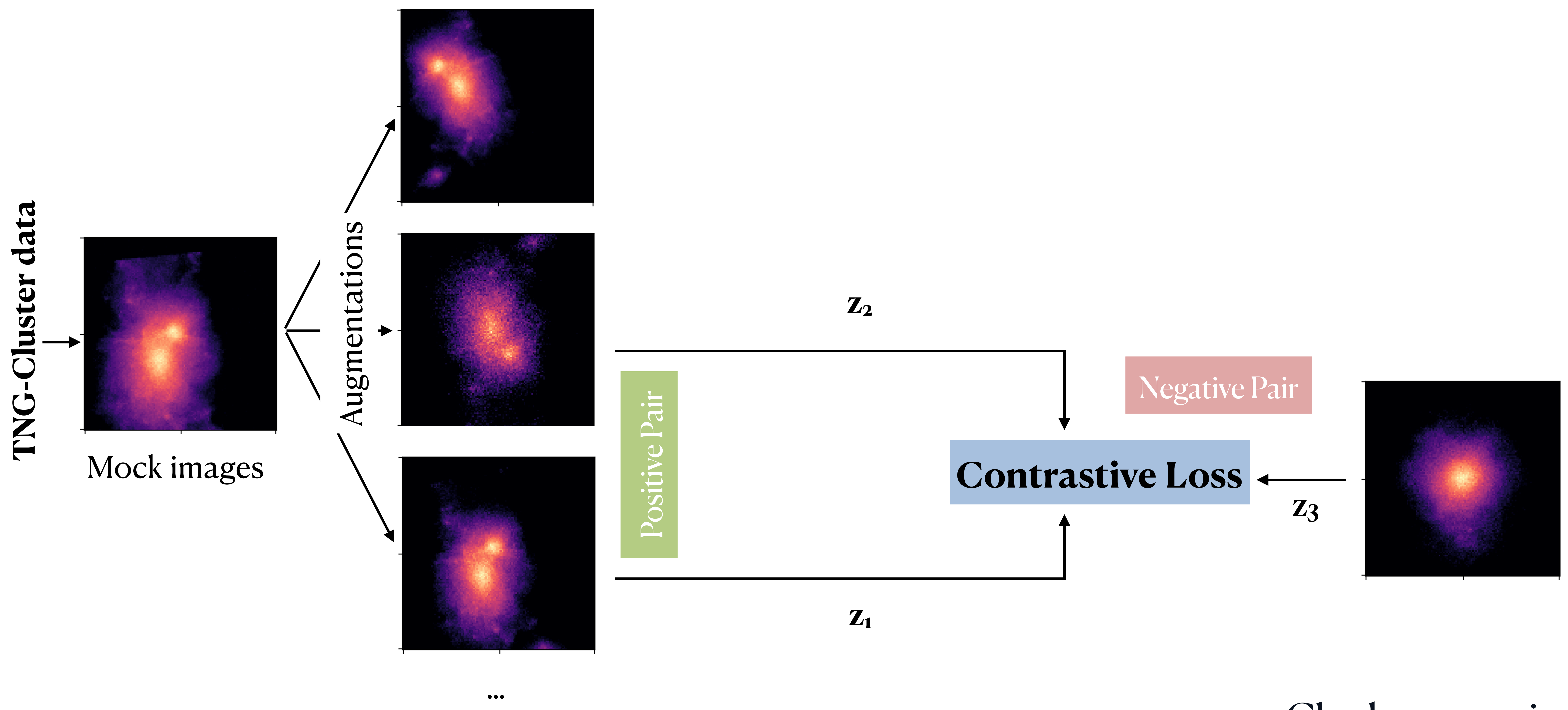
Lehle+ 2024

Quantifying turbulence via VSF

Ayromlou+ 2024

Next: how do we capture the diversity of galaxy clusters?

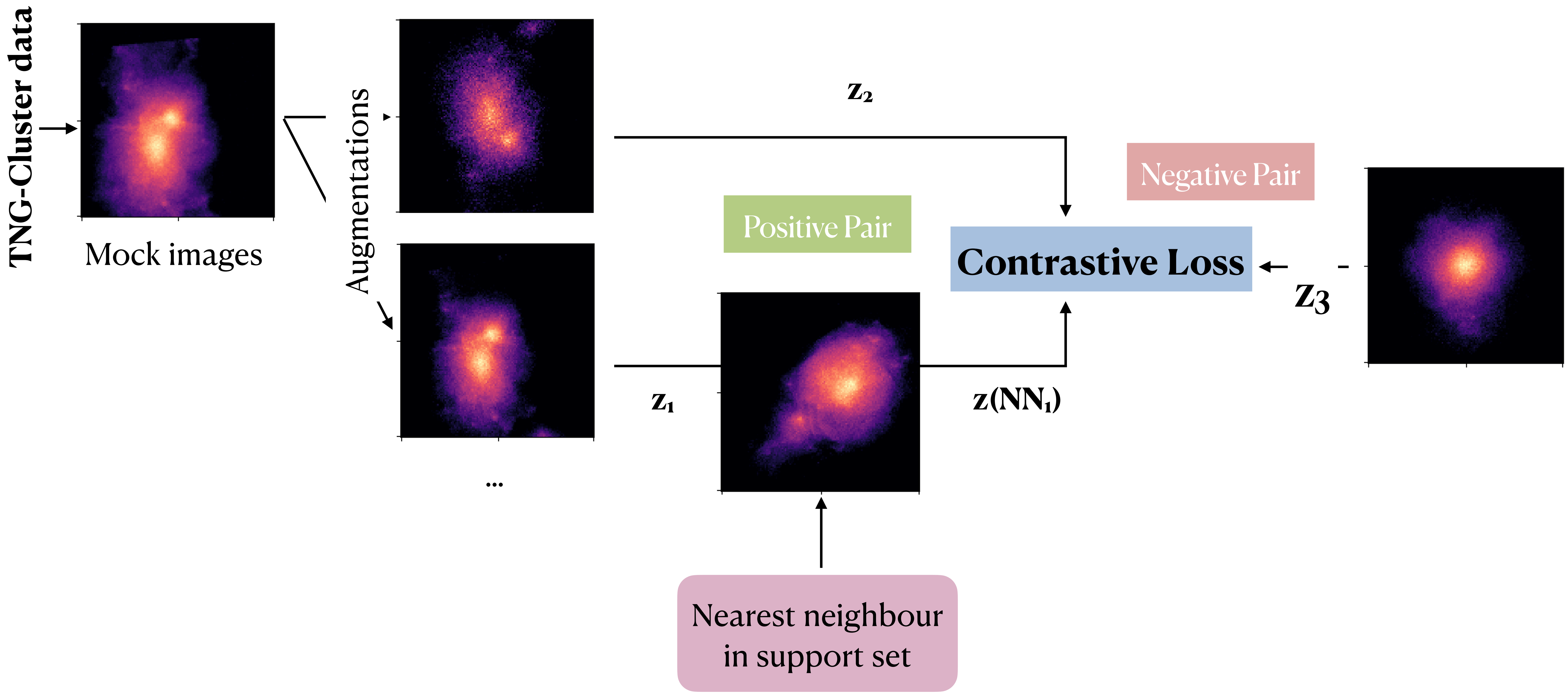
Contrastive learning: self-supervised sorting of images



Chadayammuri+ in prep.

Next: how do we capture the diversity of galaxy clusters?

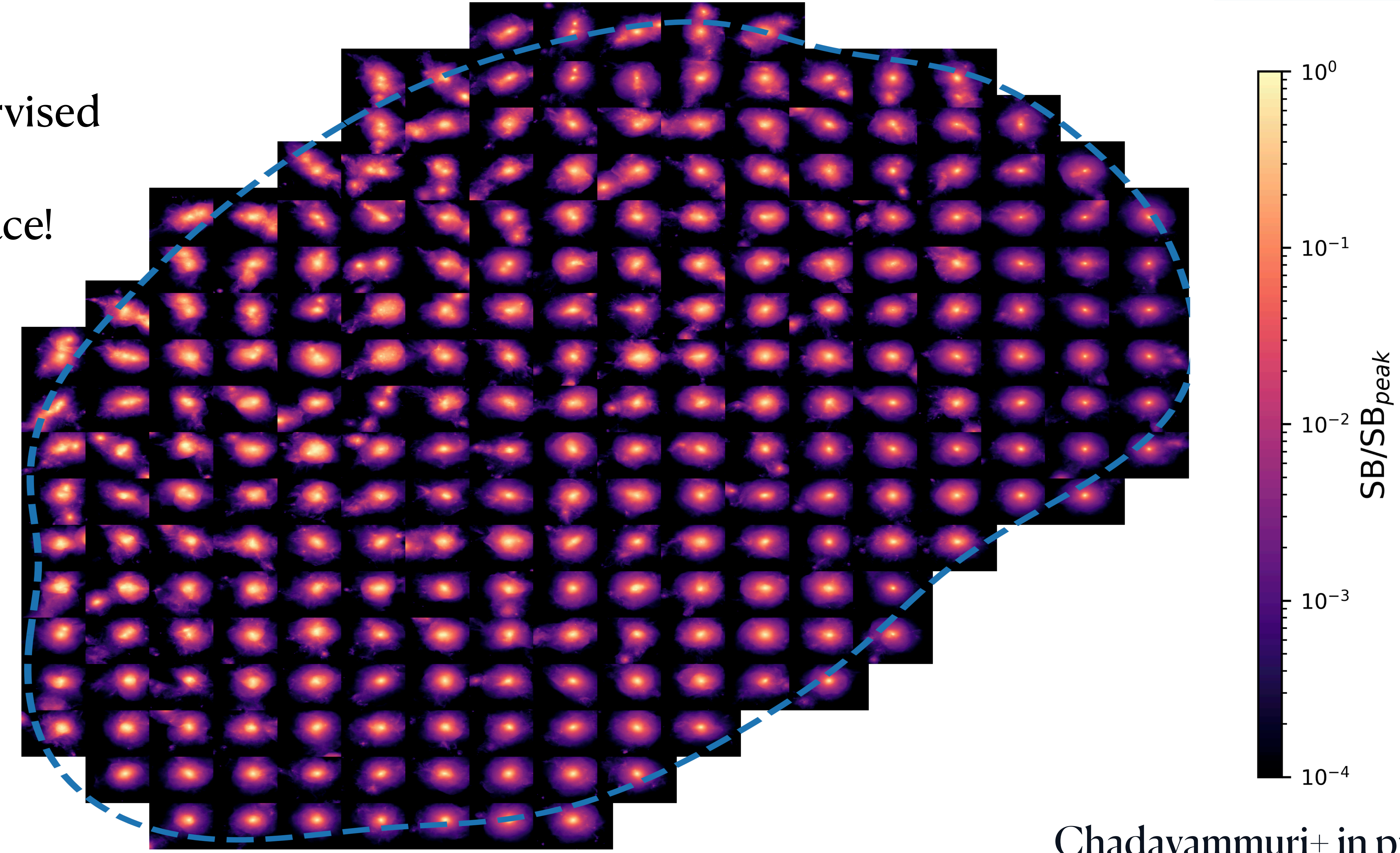
Nearest Neighbour Contrastive Learning (NNCLR, Dwibedi+ ’21)



Chadayammuri+ in prep.

Next: how do we capture the diversity of galaxy clusters?

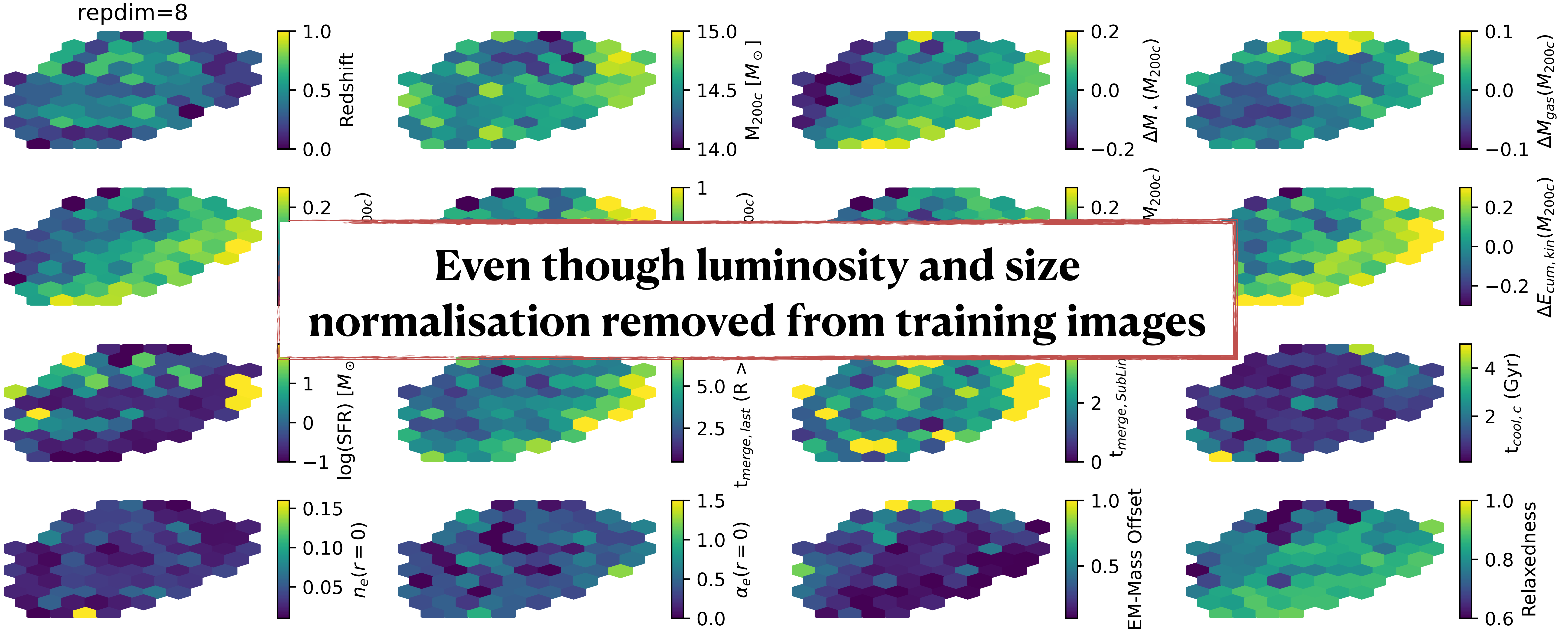
Self-supervised sorting in image space!



Chadayammuri+ in prep.

So what can we learn from galaxy cluster images?

2. Median values of key galaxy cluster properties in bins of (self-supervised!) representation space ***<u>in narrow bins of halo mass</u>***

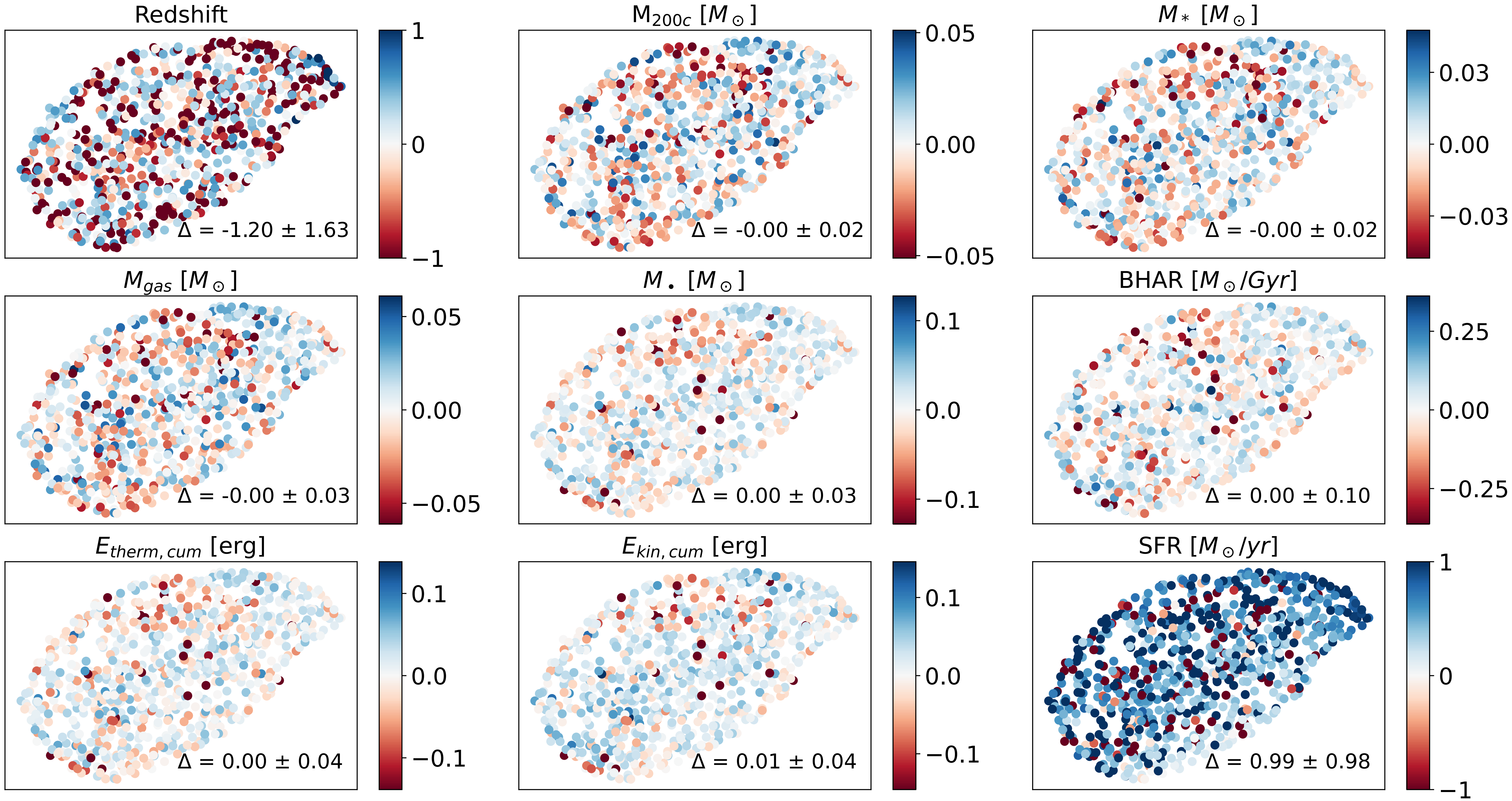


Even though luminosity and size normalisation removed from training images

Chadayammuri+ in prep.

So what can we learn from galaxy cluster images?

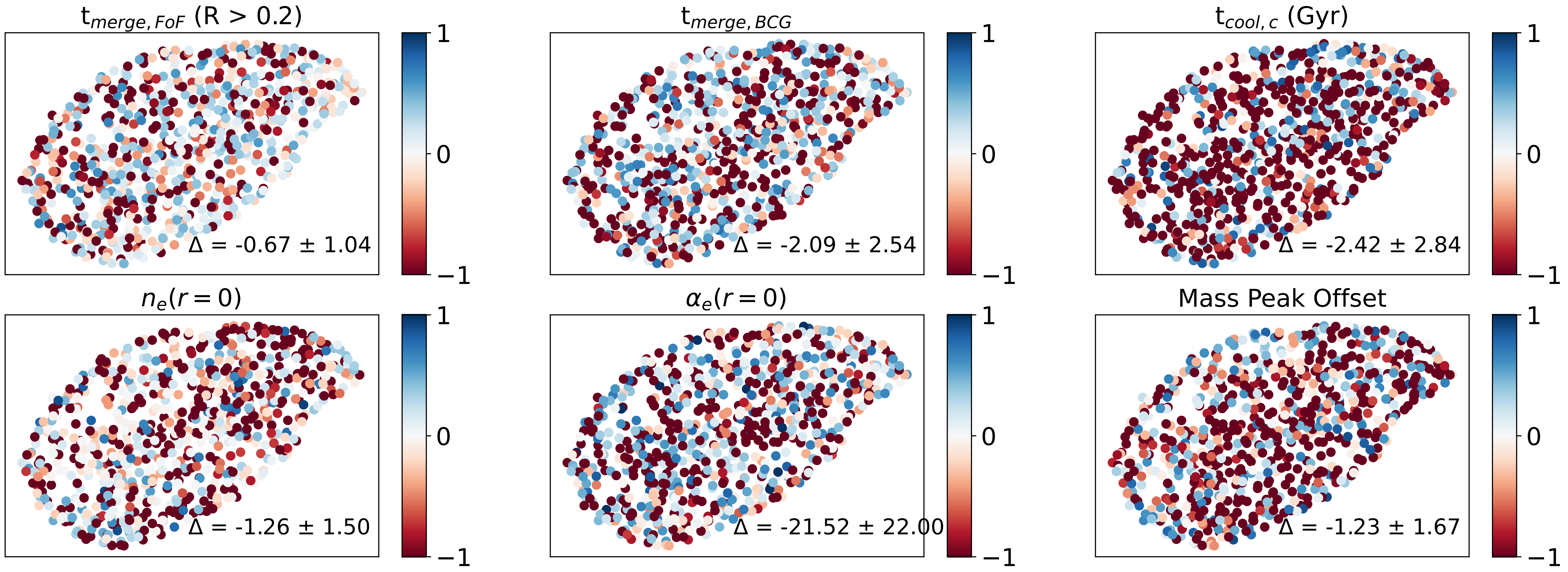
3. We can predict many galaxy cluster properties using just 2D representation of image



Chadayammuri+ in prep.

So what can we learn from galaxy cluster images?

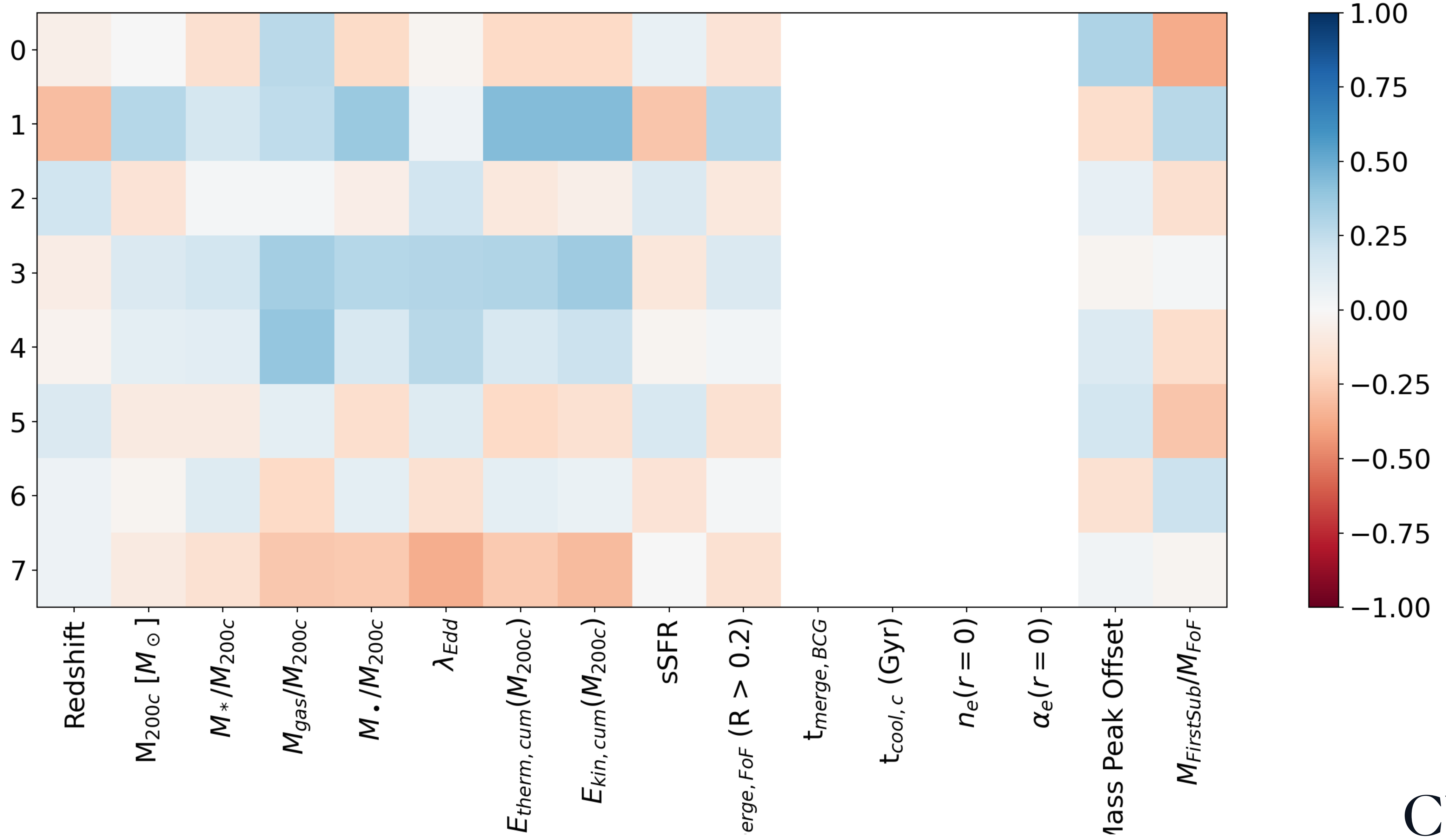
3. We can predict many galaxy cluster properties using just 2D representation of image (but not all)



Chadayammuri+ in prep.

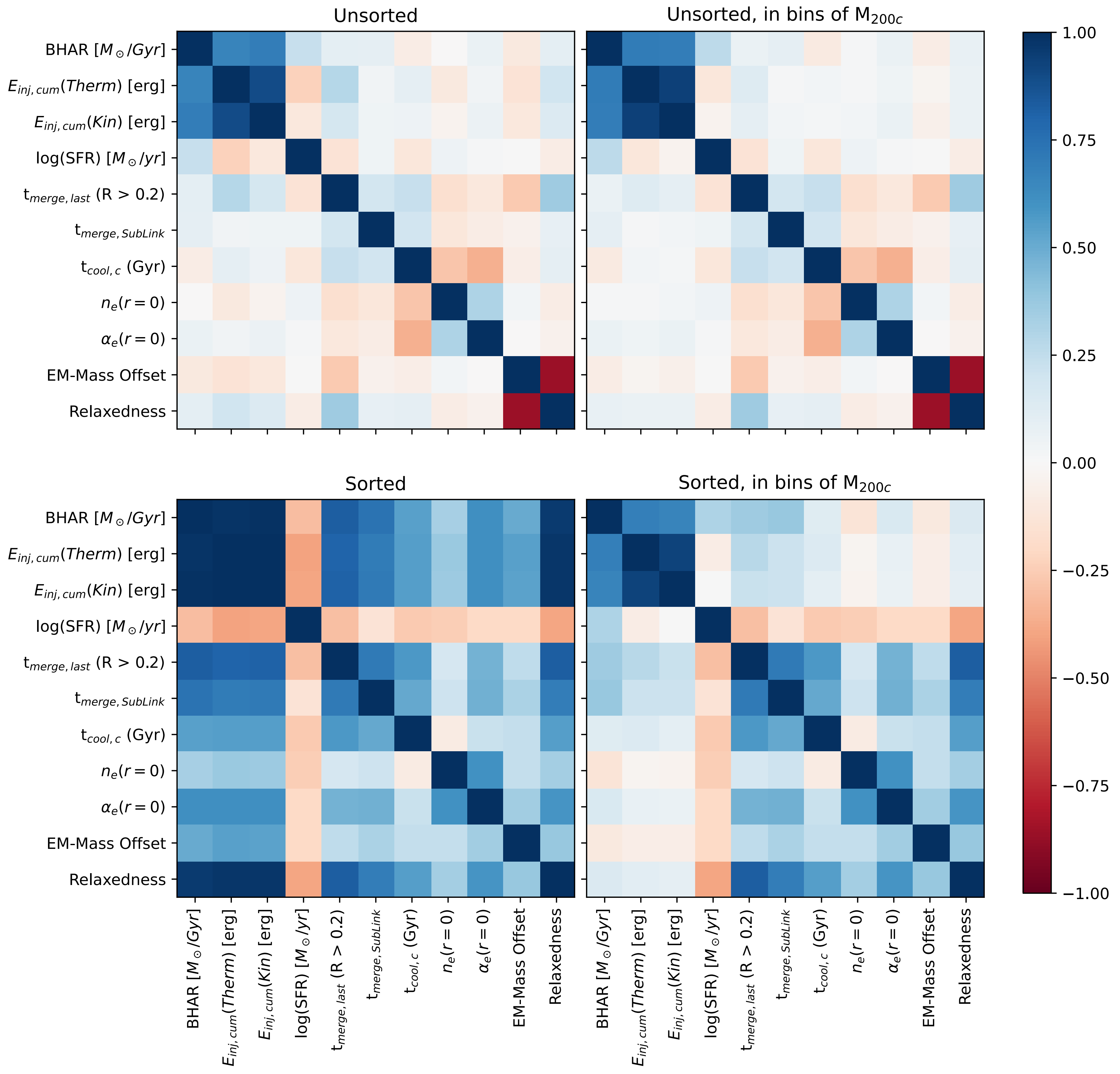
So what can we learn from galaxy cluster images?

In fact, the properties we predict poorly are the ones that show least correlation with any of the representation parameters



Chadayammuri+ in prep.

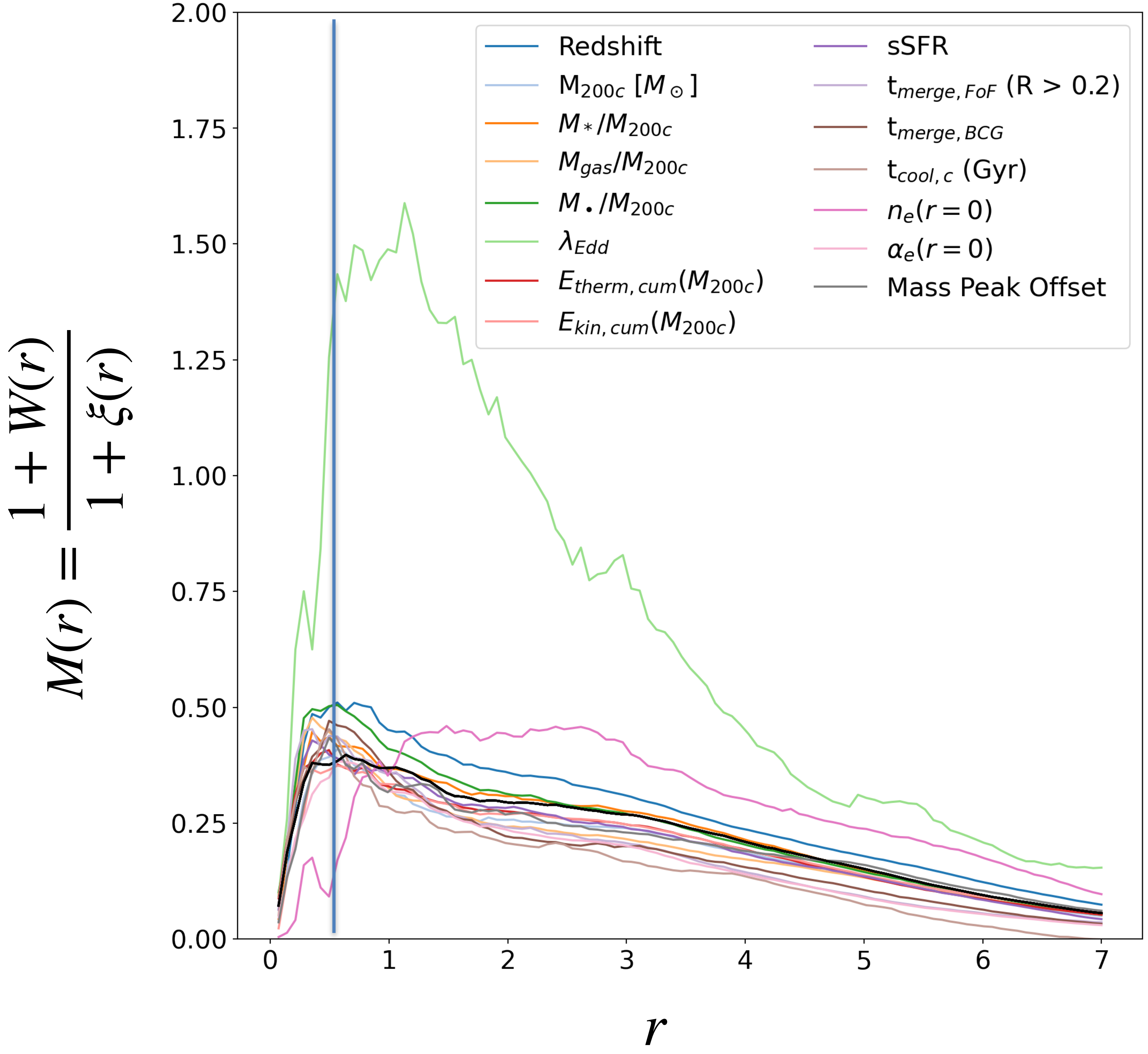
So what can we learn from galaxy cluster images?



4. We can quantify relationships between physical properties for different populations in the image space

Chadayammuri+ in prep.

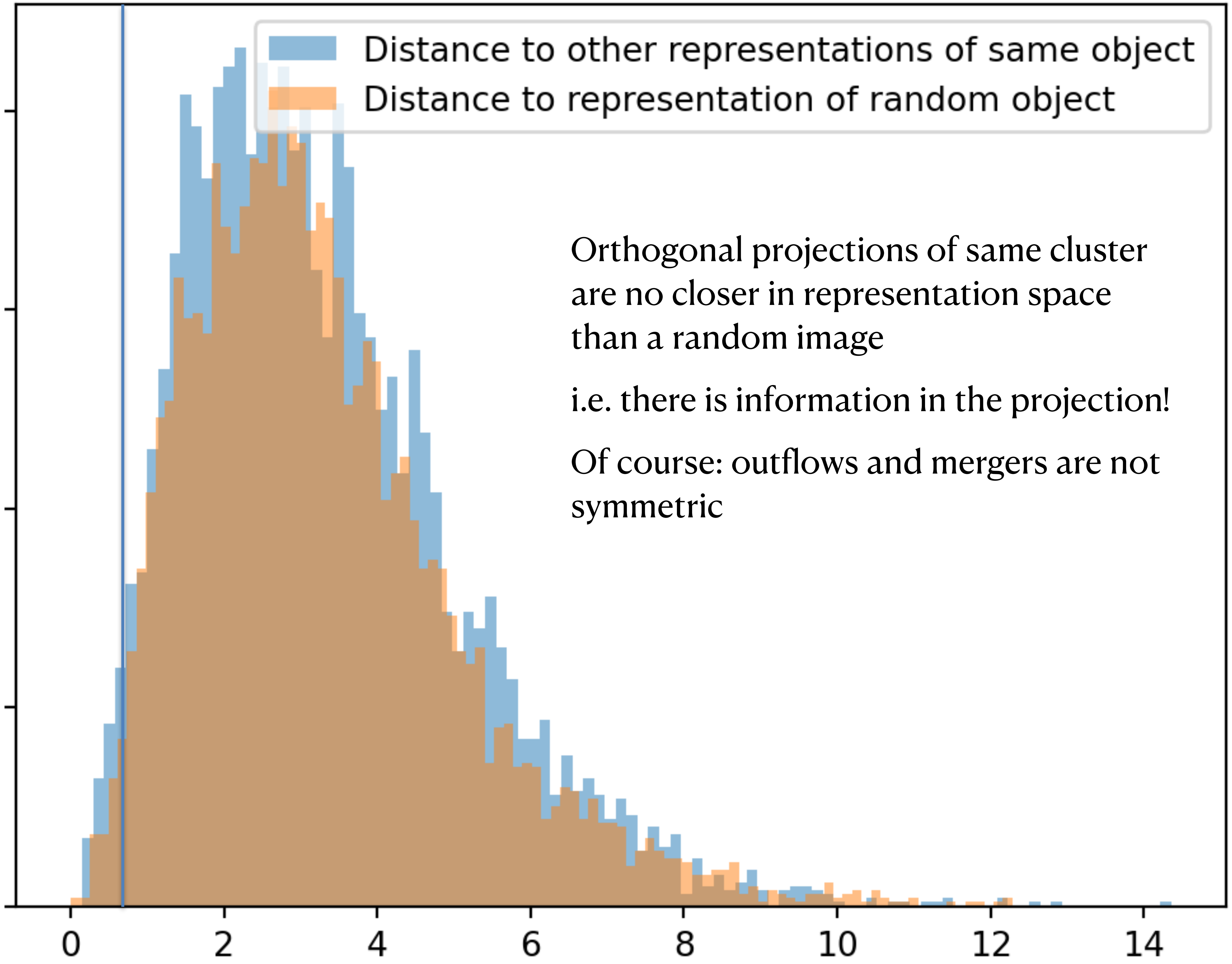
So what can we learn from galaxy cluster images?



5. We can identify a "neighbourhood scale" in the (unphysical) representation space within which physical properties are more correlated to each other than the overall cluster population

Chadayammuri+ in prep.

So what can we learn from galaxy cluster images?



Orthogonal projections of same cluster are no closer in representation space than a random image

i.e. there is information in the projection!

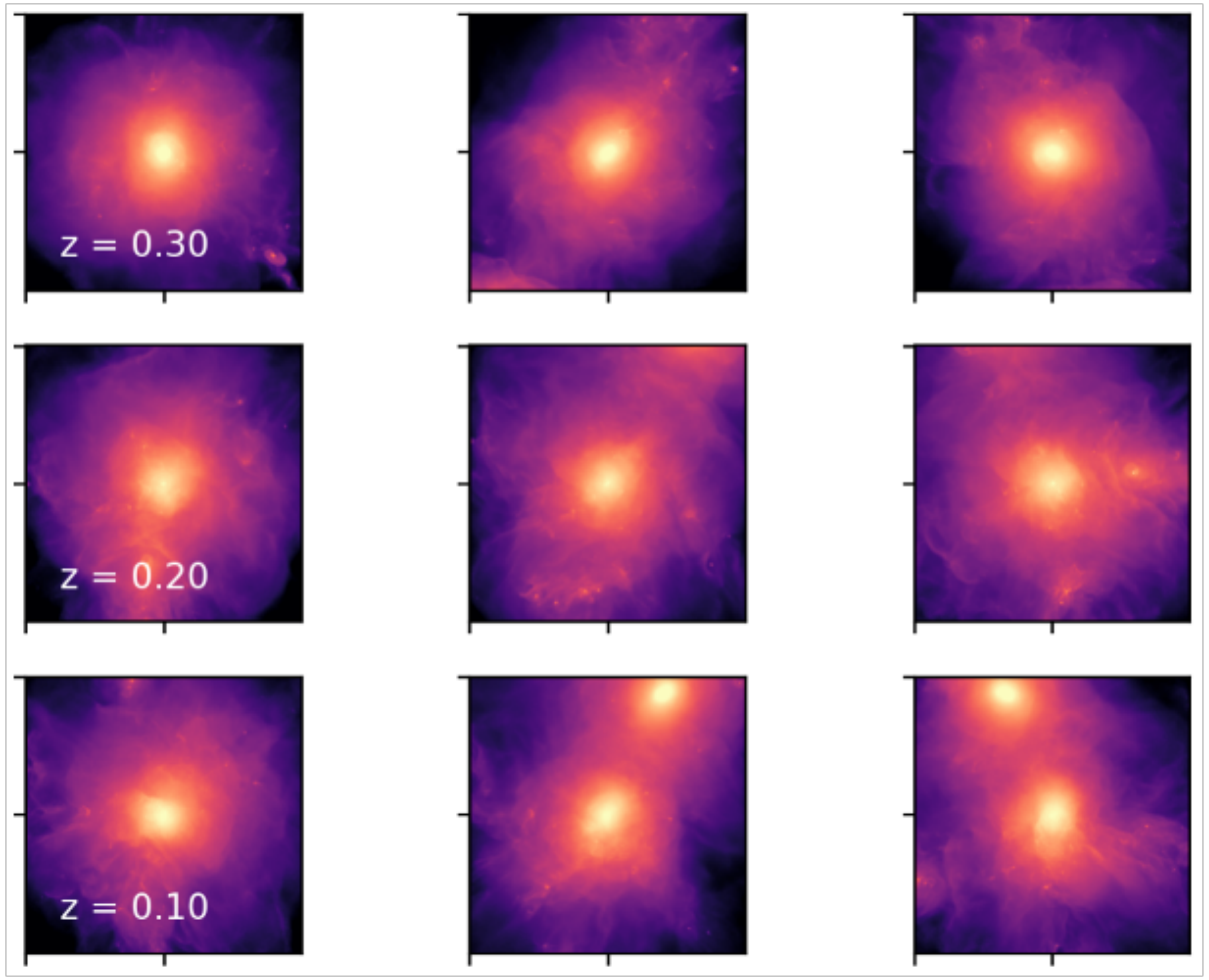
Of course: outflows and mergers are not symmetric

Chadayammuri+ in prep.

Nice side-effect: This is an analogue finder!

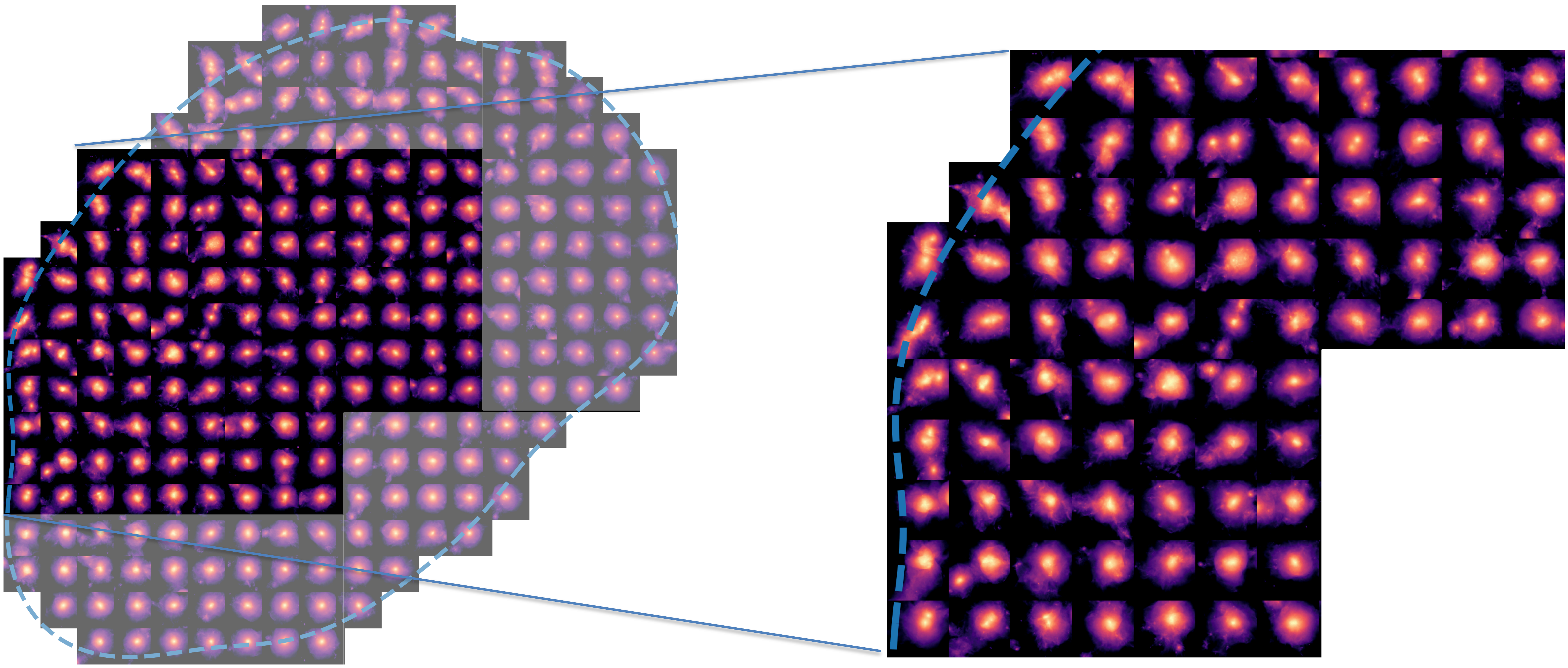
Classic approach: apply cuts to mass ratio, impact parameter in merger tree

But this can easily fail, because assigning particles to separate halos during mergers is very challenging

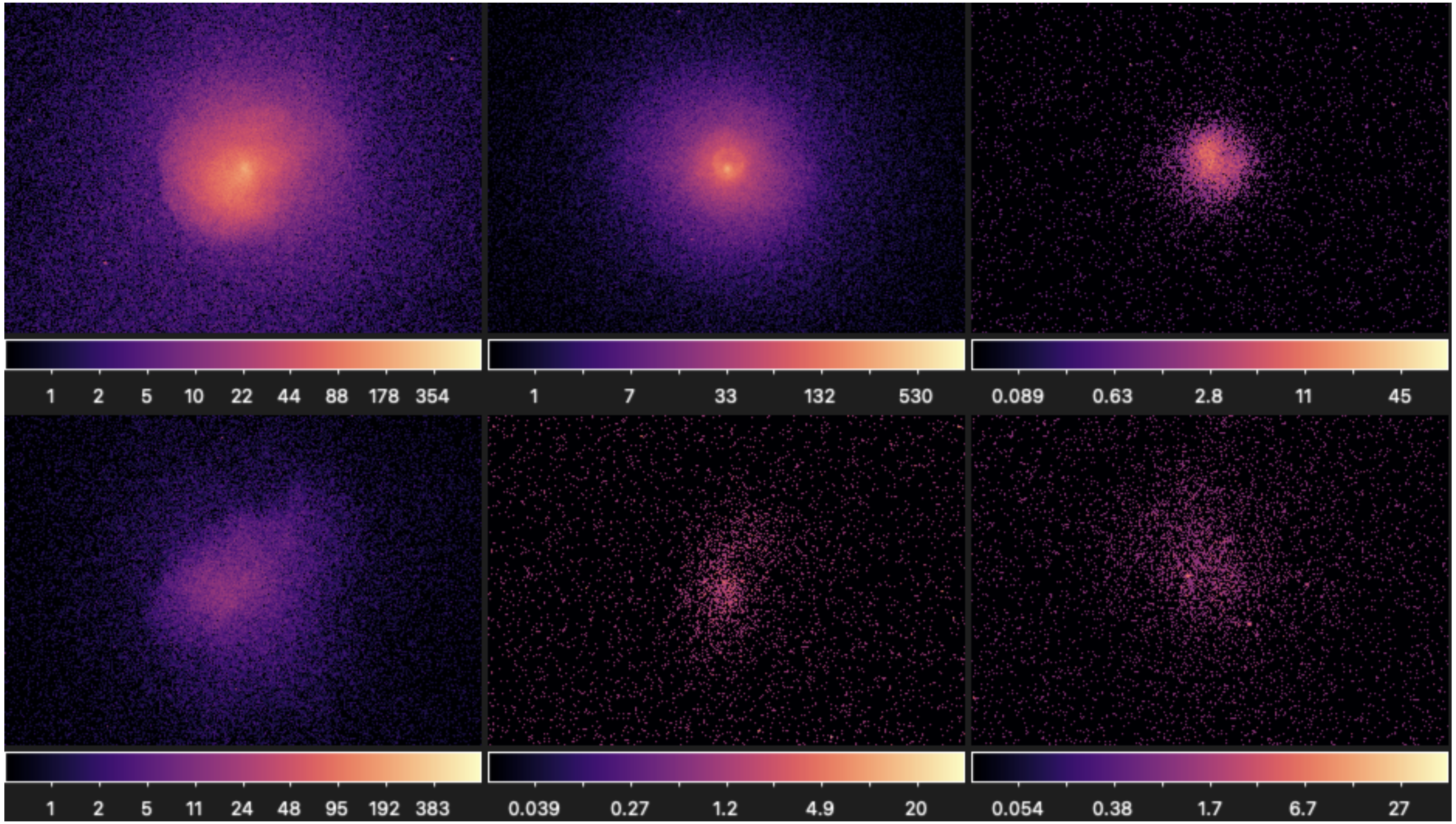


Nice side-effect: This is an analogue finder!

Instead: just find neighbours in representation space



Next step: How well does this work with realistic images?



Prunier, UC+ in prep.

Conclusions

- Galaxy clusters are powerful tools for both cosmology and astrophysics

Requires mapping between baryons and dark matter

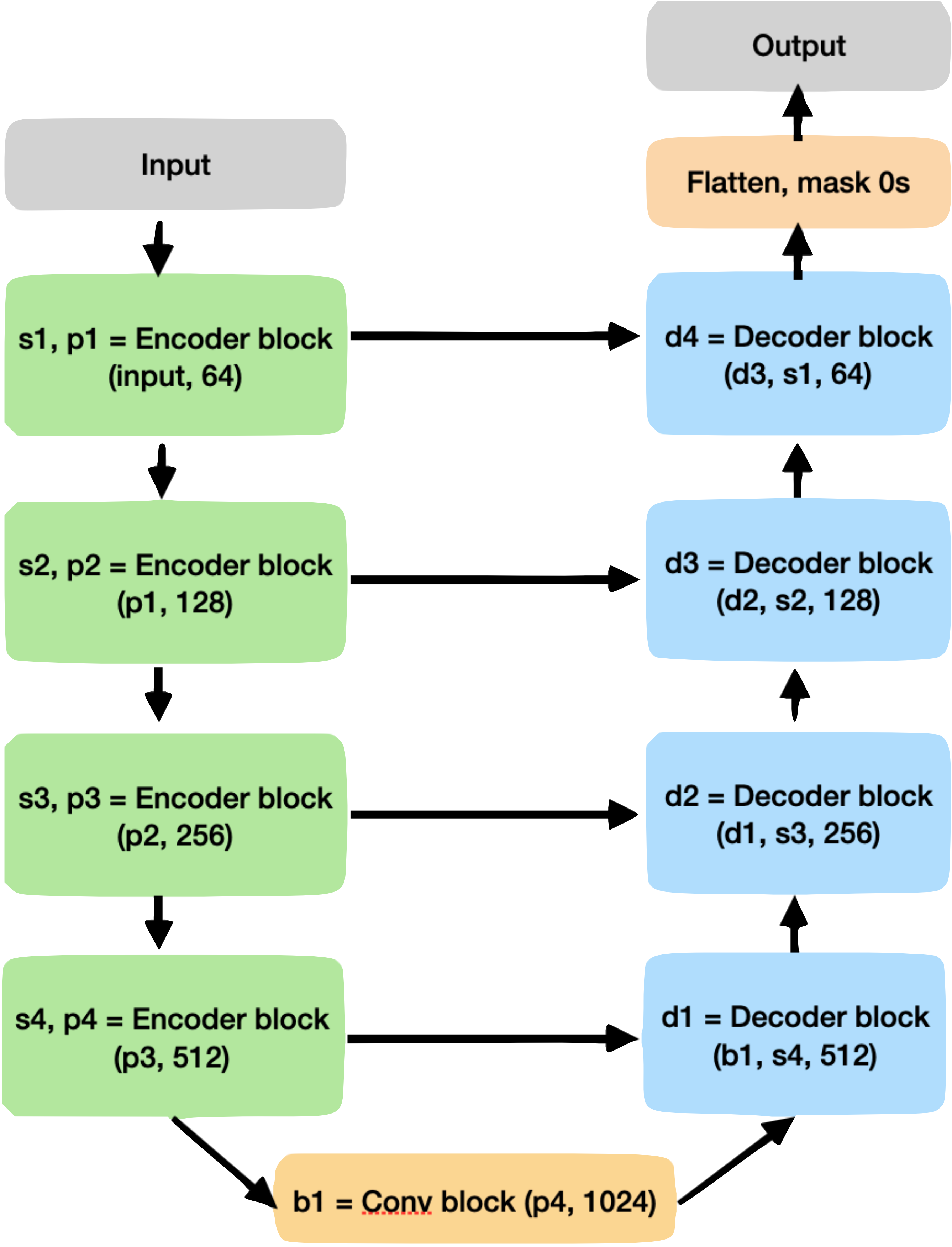
U-Nets can do this at the <1% level

Requires finding analogues matched in assembly history + dynamics

Contrastive learning can reduce images to compact representation spaces where similar objects live nearby

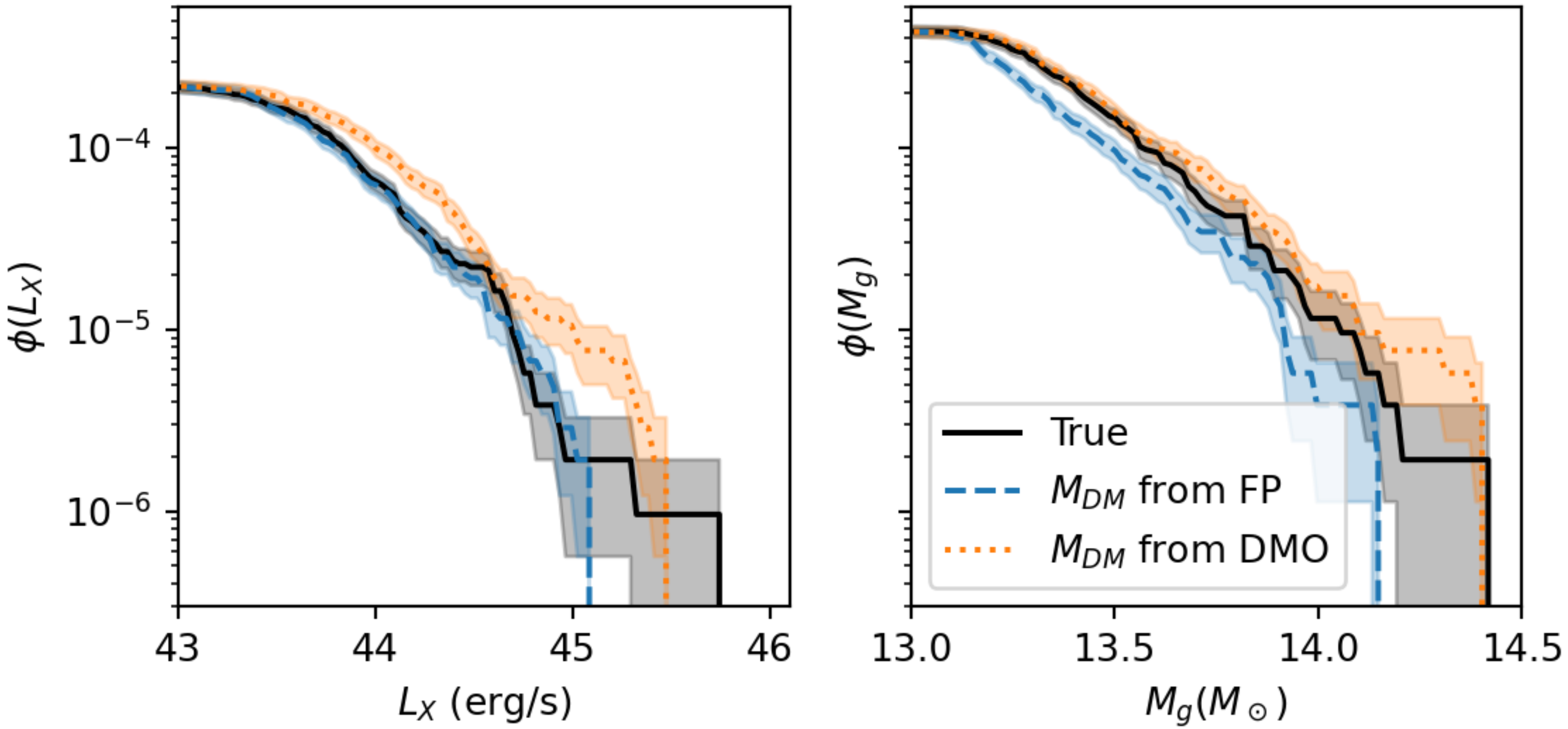
- **Image-based machine learning** trained on **cosmological simulations** allows us to **capture information that is lost in summary statistics** and **increase the robustness of** cosmology + astrophysics **inference** from galaxy clusters

Architecture: U-Net



Chadayammuri+ 2023

Training on FP → apply to DMO

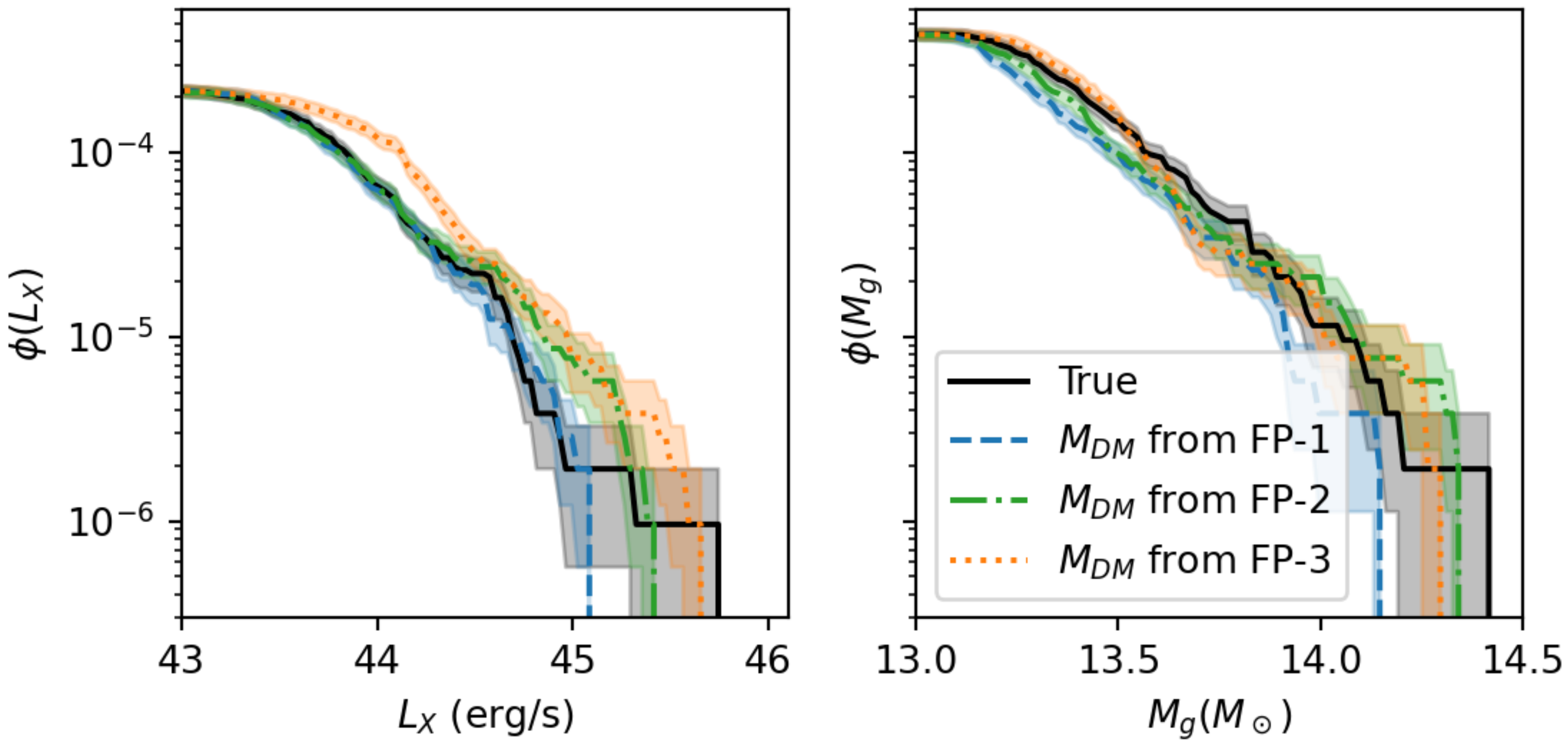


Systematic differences between the DM distribution in FP and DMO simulations → predictions biased high

But can be calibrated

Chadayammuri+ 2023

Training on high res → apply to low res



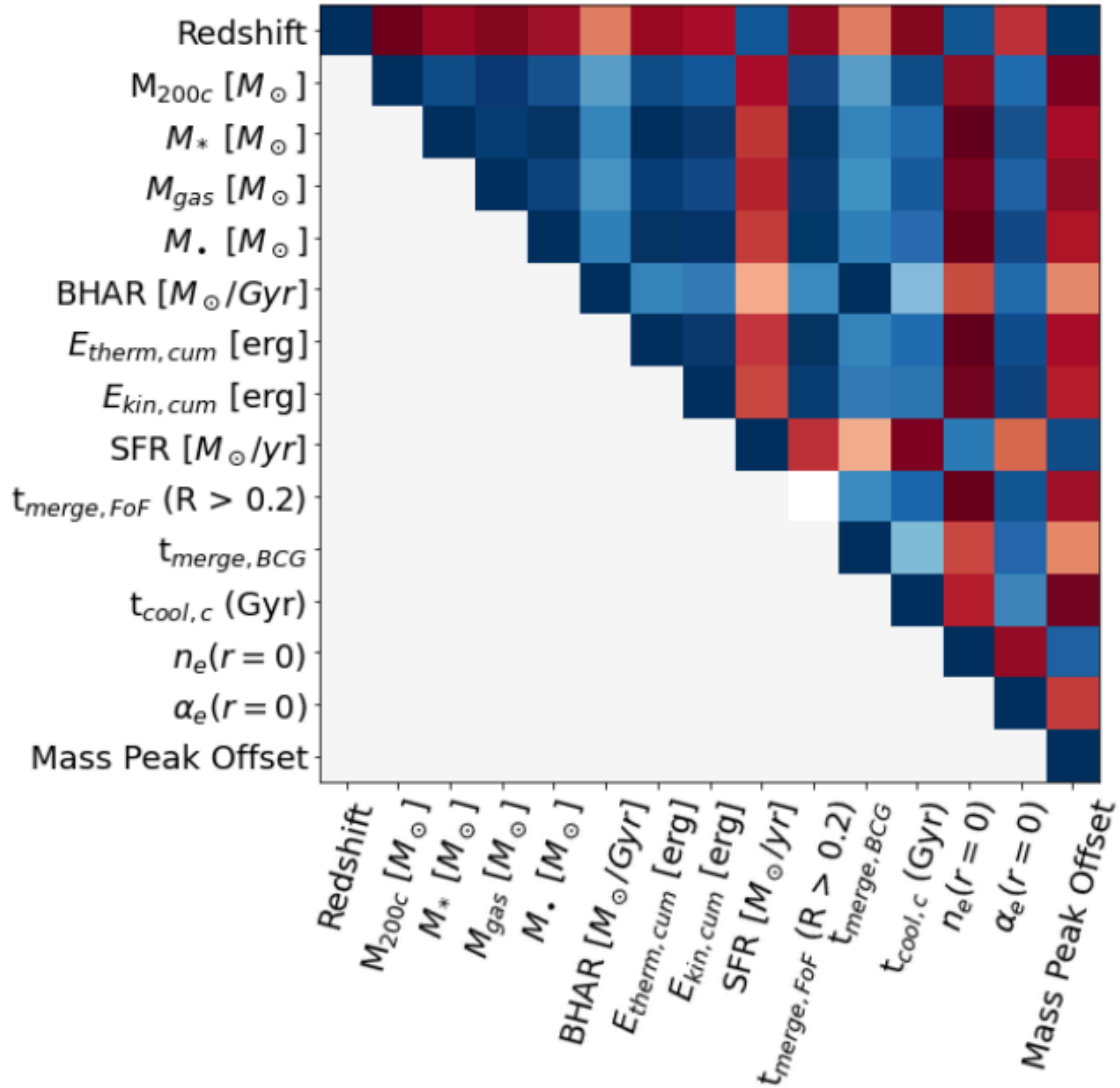
If resolution degraded by more than 8x → predictions biased high

As with DMO, this is because of differences in spatial structure of DM maps

This, too can be calibrated

Chadayammuri+ 2023

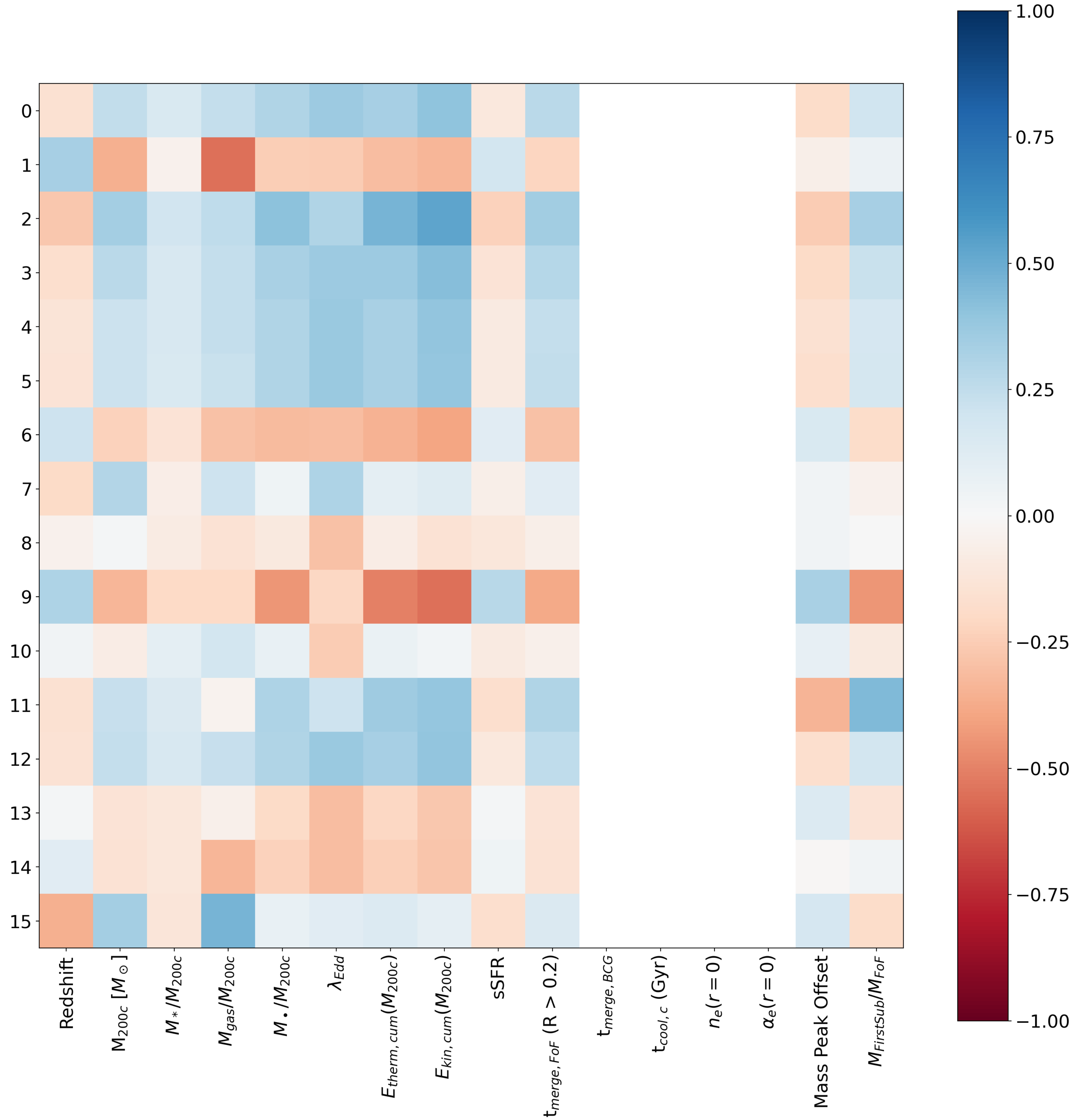
So what can we learn from galaxy cluster images?



Quantifying relationships between physical processes in the representation space

Chadayammuri+ in prep.

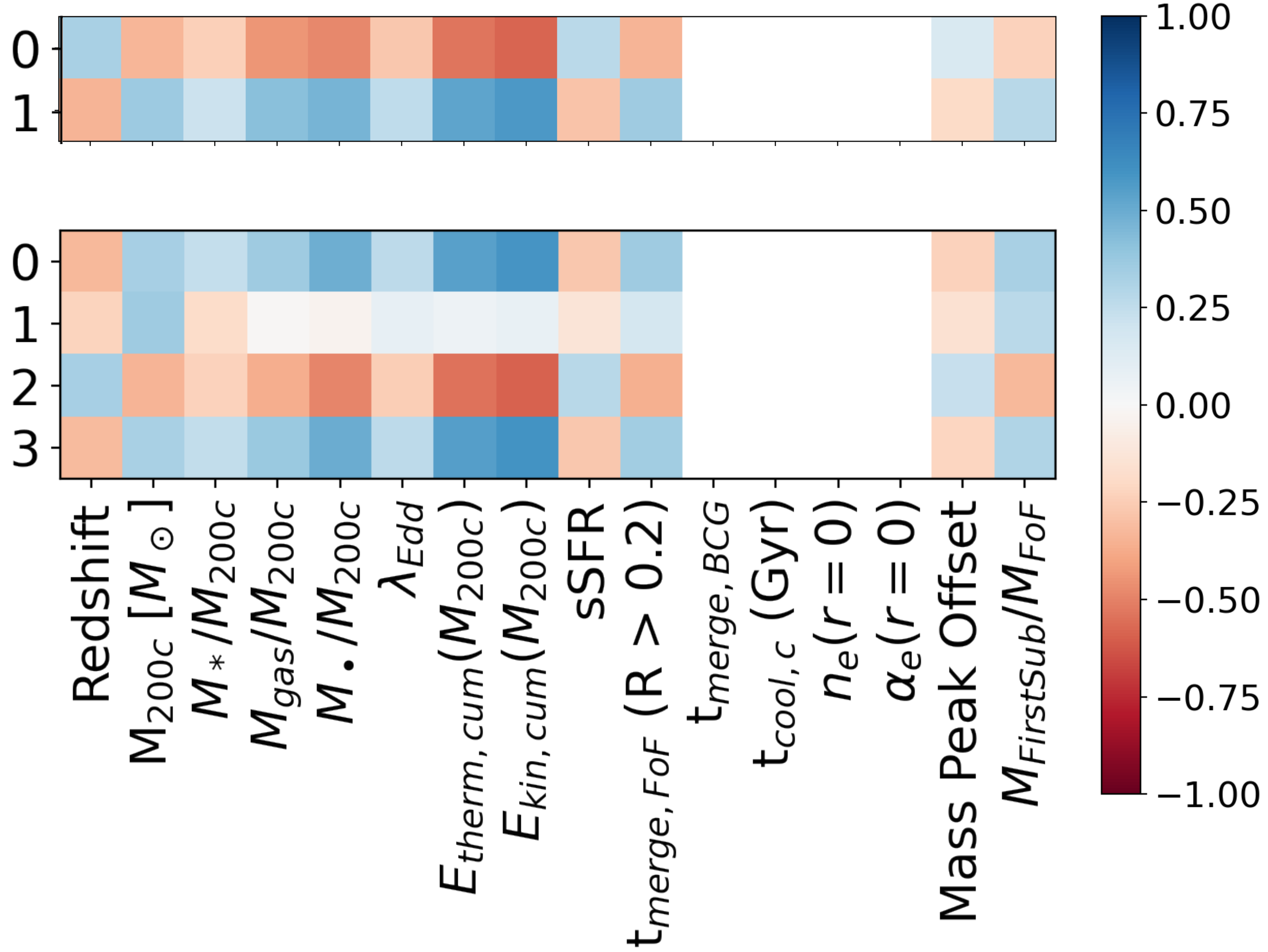
What can we learn from the representation space?



Adding more representation dimensions results in a few being strongly correlated with halo mass, BH properties, and time since merger, and most being unrelated

Chadayammuri+ in prep.

What can we learn from the representation space?



Using < 8 dimensions focuses on BH trends (which in turn correlate with total mass)

Chadayammuri+ in prep.