Deep Learning Models to Infer Mass Maps from SZ, X-ray and Galaxy Members Observations in Galaxy Clusters.

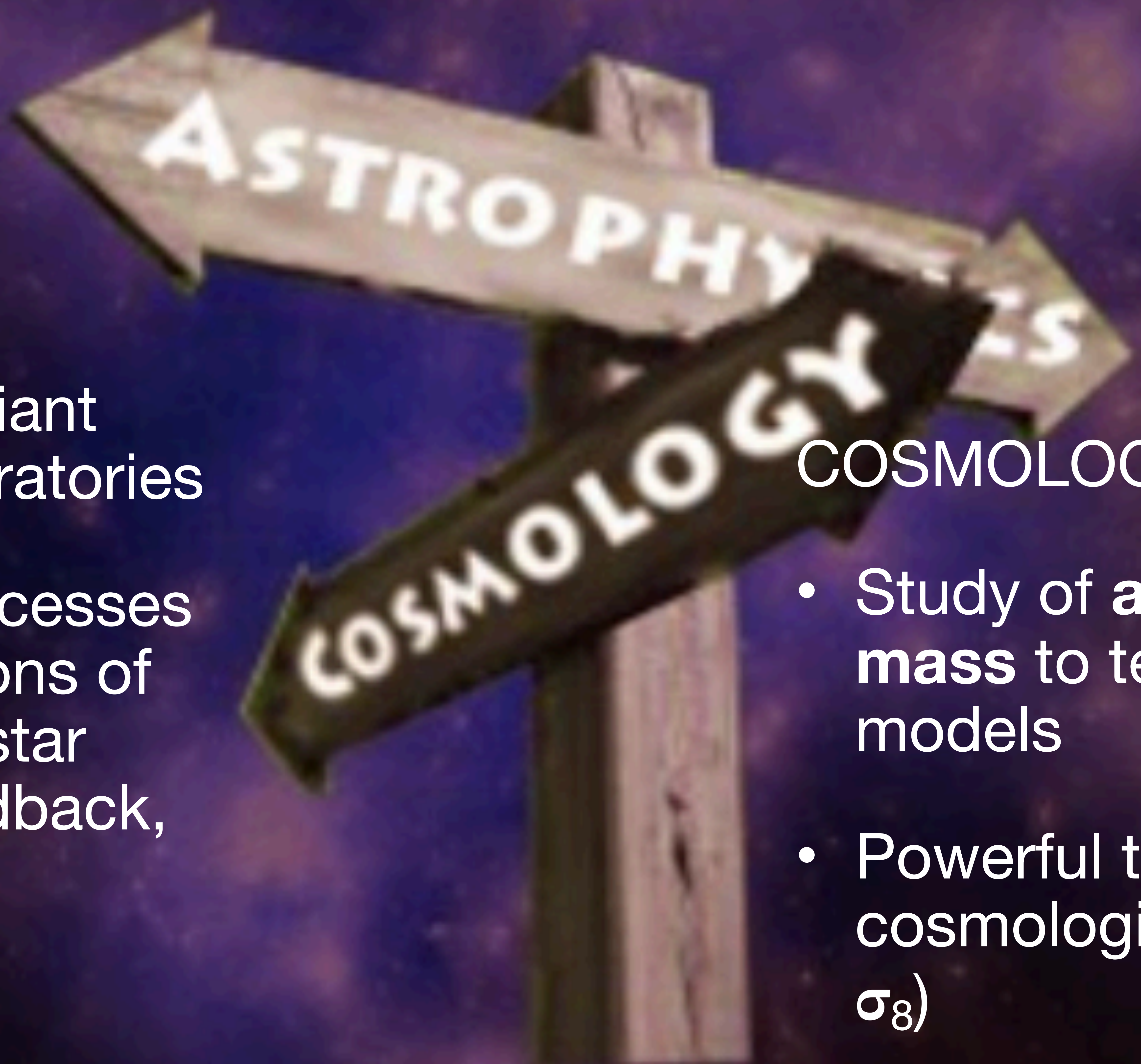
Daniel de Andrés daniel.deandres@uam.es

Gustavo Yepes, Weiguang Cui, Marco de Petris, Antonio Ferragamo, Gianmarco Aversano.

Universidad Autónoma de Madrid

ML-IAP/CCA-2023

Galaxy Clusters: the crossroad



ASTROPHYSICS:

- Isolated system: giant astrophysical laboratories
- Many physical processes involving the baryons of the ICM: cooling, star formation, SN feedback, AGN feedback etc

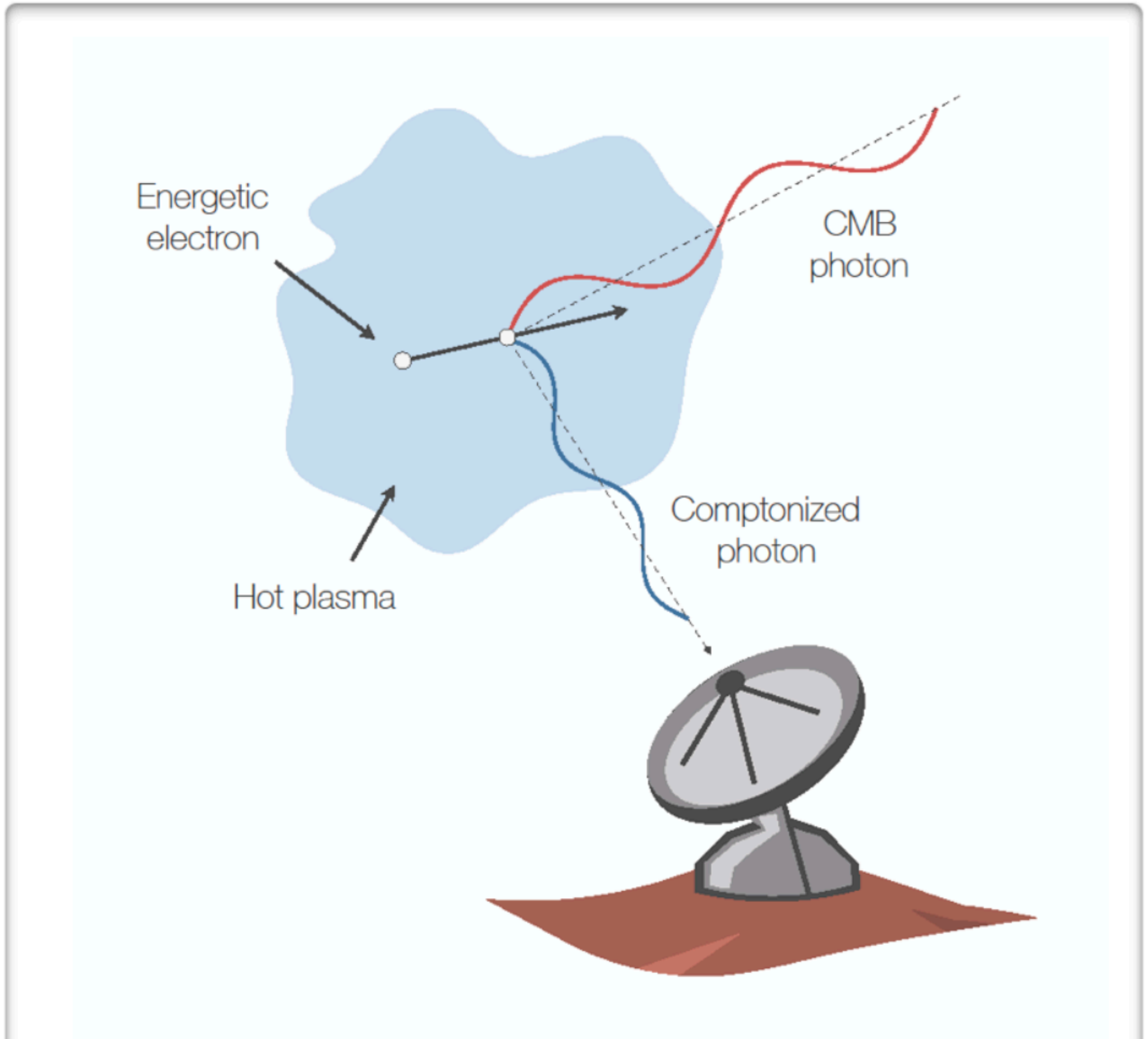
COSMOLOGY:

- Study of **abundance and mass** to test cosmological models
- Powerful tool to estimate cosmological parameters (Ω_m , σ_8)

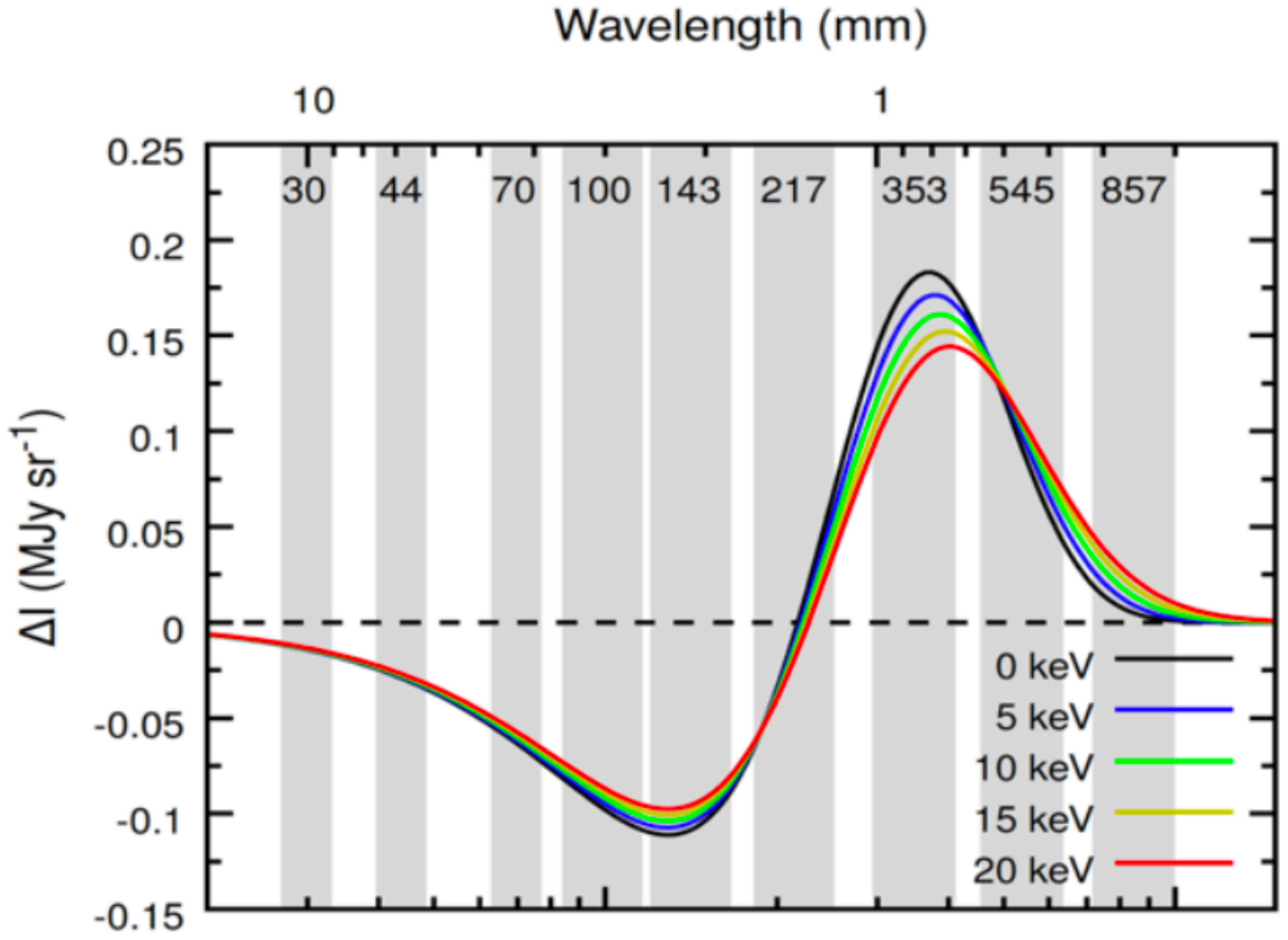
Motivation

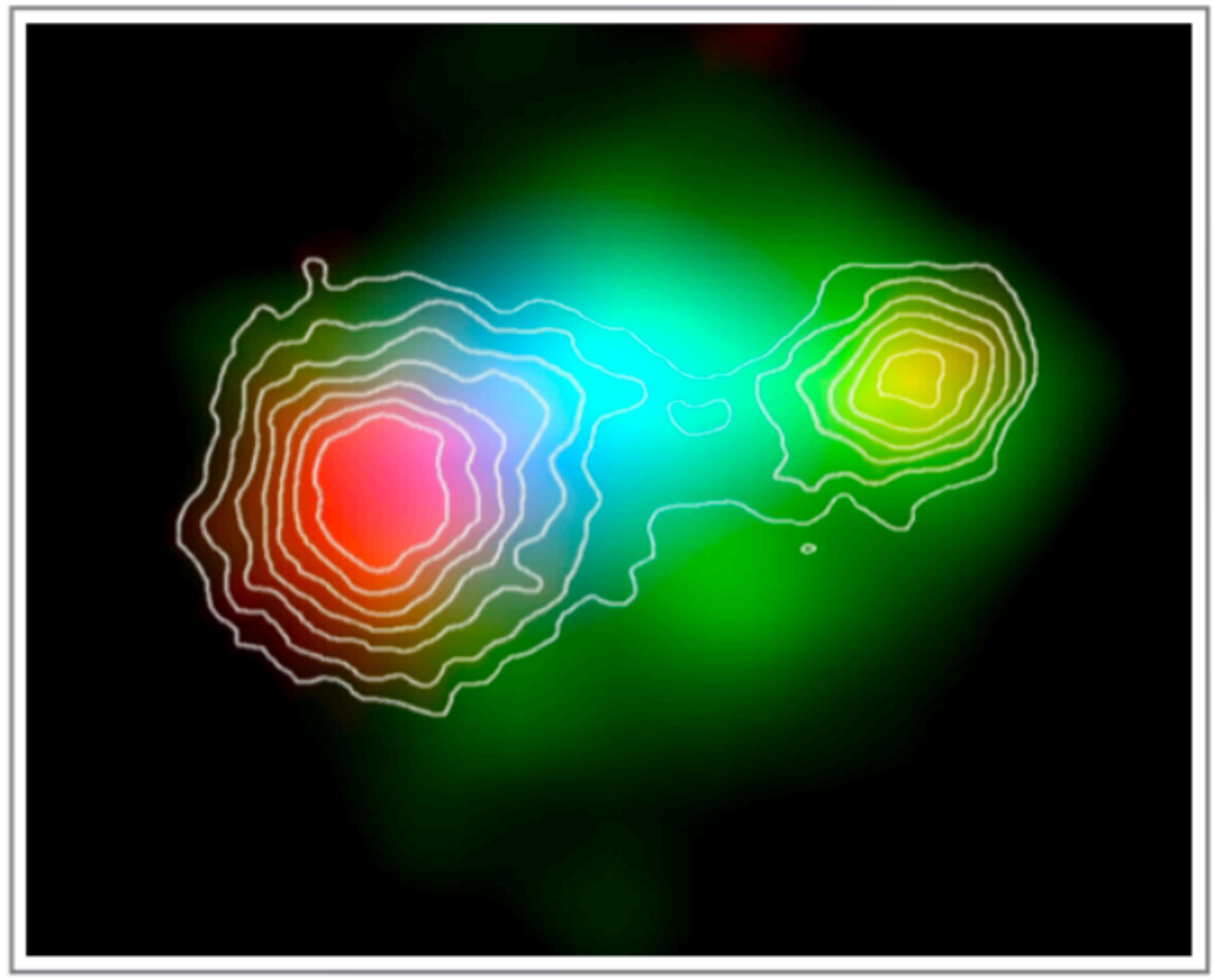
- In de Andres et al 2022, we have predicted the mass of galaxy clusters from SZ observations using deep learning, particularly for the Planck PSZ2 catalog.
- We want to generalise this approach to infer projected mass density maps from an observation, e.g. from tSZ we aim at inferring the mass map in 2D.
- Weak lensing traces projected mass density, but WL surveys are scarce (tens of clusters) compared to Compton-y and X-ray whose surveys observe hundreds or thousands of galaxy clusters.

Gas: SZ effect



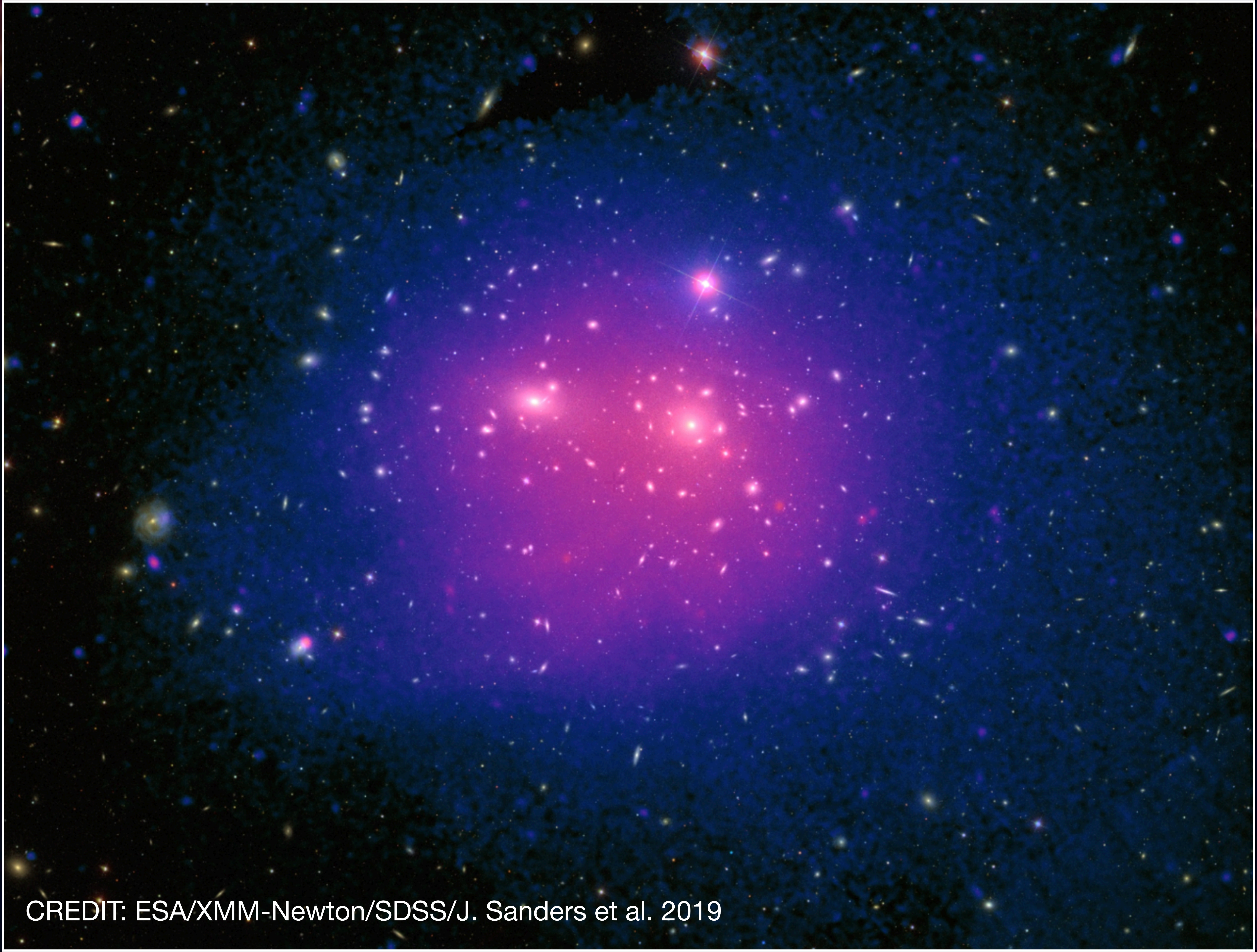
Credit: Kaustuv Basu et al





https://astro.uni-bonn.de/en/research/mm-submm-astronomy/projects-1/sz-effect-and-cosmology

Gas and galaxies: X-ray and optical



Final goal: Simulation based inference

De Andres et al 2022

nature > nature astronomy > articles > article

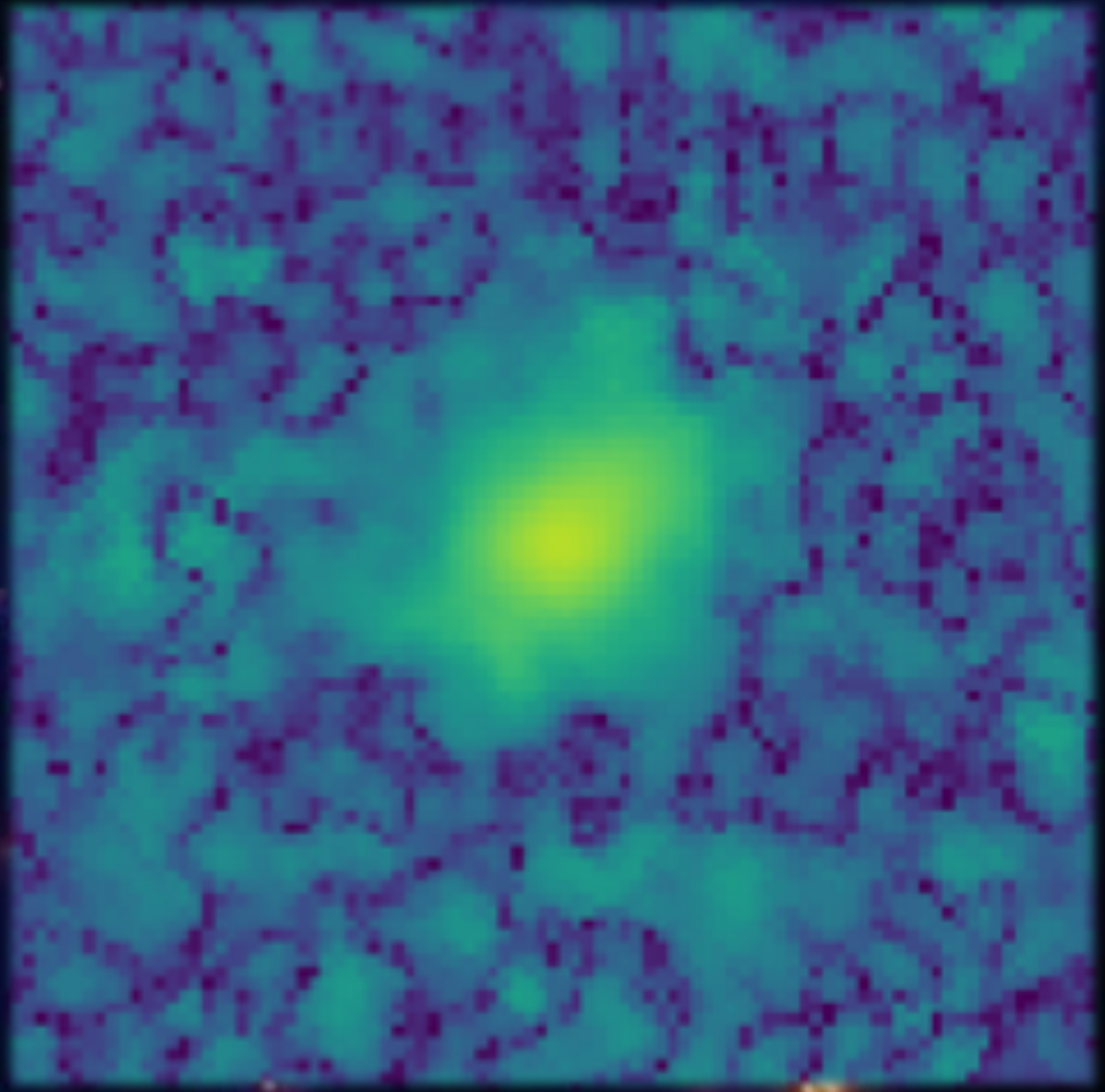
Article | Published: 17 October 2022

A deep learning approach to infer galaxy cluster masses from Planck Compton-*y* parameter maps

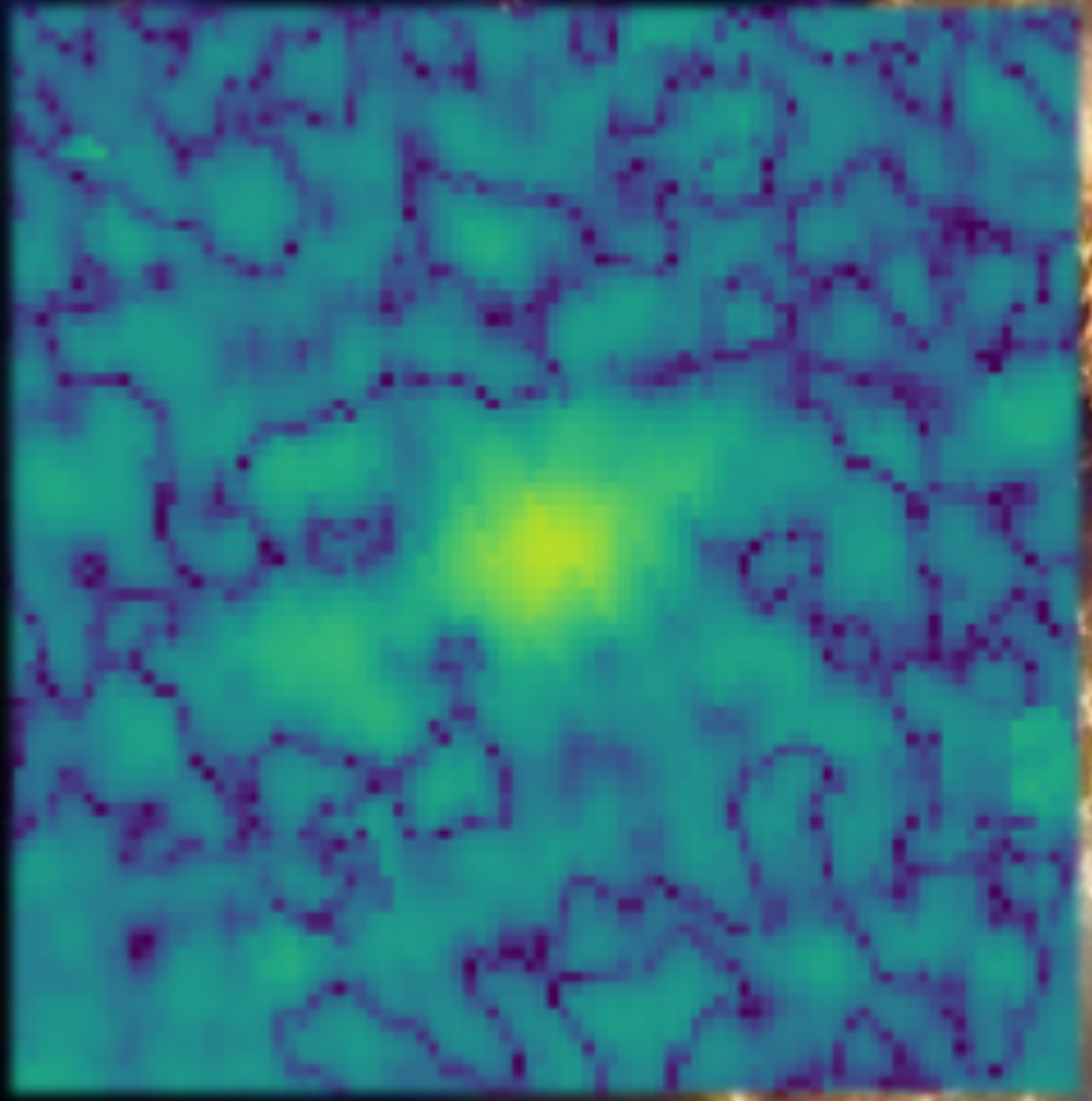
Daniel de Andres, Weiguang Cui, Florian Ruppin, Marco De Petris, Gustavo Yepes, Giulia Gianfagna, Ichraf Lahouli, Gianmarco Aversano, Romain Dupuis, Mahmoud Jarraya & Jesús Veg
Ferrero

Nature Astronomy **6**, 1325–1331 (2022) | Cite this article

Simulation of a Compton-y observation



Real Compton-y Planck observation



Training on simulated data from cosmological simulations to predict properties of real surveys

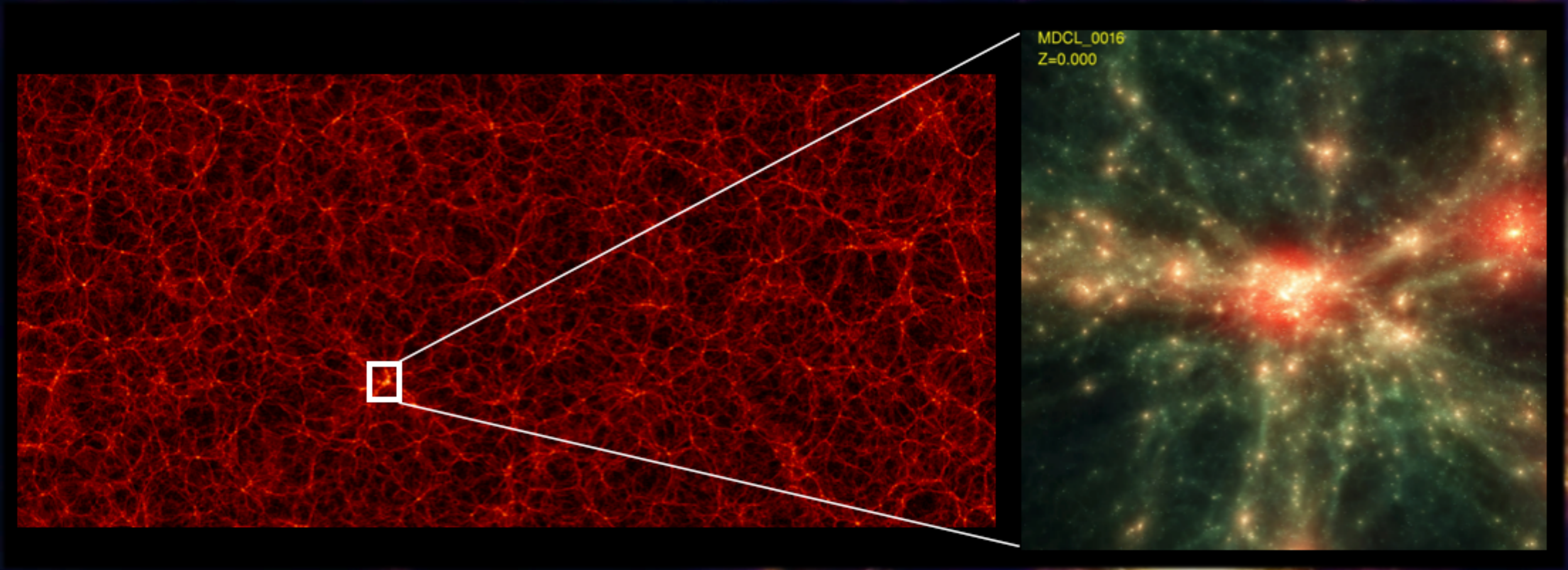
Outline

- Cosmological simulations: The Three Hundred Project.
- Mock dataset of images: SZ, Xray and stars.
- Deep learning model and results on the inference of mass maps.

- **Machine Learning group:** Daniel de Andrés, Weiguang Cui, Gustavo Yepes, Marco De Petris, Florian Ruppin, Federico De Luca, Giulia Gianfagna, Jesús Vega Ferrero, Alejandro Jiménez (+EURANOVA people: Gianmarco Aversano).

The Three Hundred Project

A set of **Cosmological hydrodynamical simulations**: Zoom-in simulations of $15/h$ Mpc radius around the 324 most massive clusters of the full $1/h$ Gpc MultiDark N-Body simulation.



The Three Hundred Project

- A set of Cosmological hydrodynamical simulations: Zoom-in simulations of $15/h$ Mpc radius around the 324 most massive clusters of the full $1/h$ Gpc MultiDark N-Body simulation. **Galaxy clusters of $M \sim 10^{15} h^{-1} M_{\odot}$, with particle resolution of $\sim 10^{9} h^{-1} M_{\odot}$.**

- DATA SAMPLE: **3 different versions** of the 324 simulations with different physics: GADGET-MUSIC (SN feedback, stellar winds), GADGET-X (+AGN Feedback), GIZMO-SIBMA (+stronger AGN Dave's model).

- Mock observations: X-ray (XMM, Athena), t-SZ, CCD (SDSS bands), lensing maps. Participate in Check-Mate and NIKA2 LPSZ as simulation providers.

Deep Learning Models to Infer Mass Maps

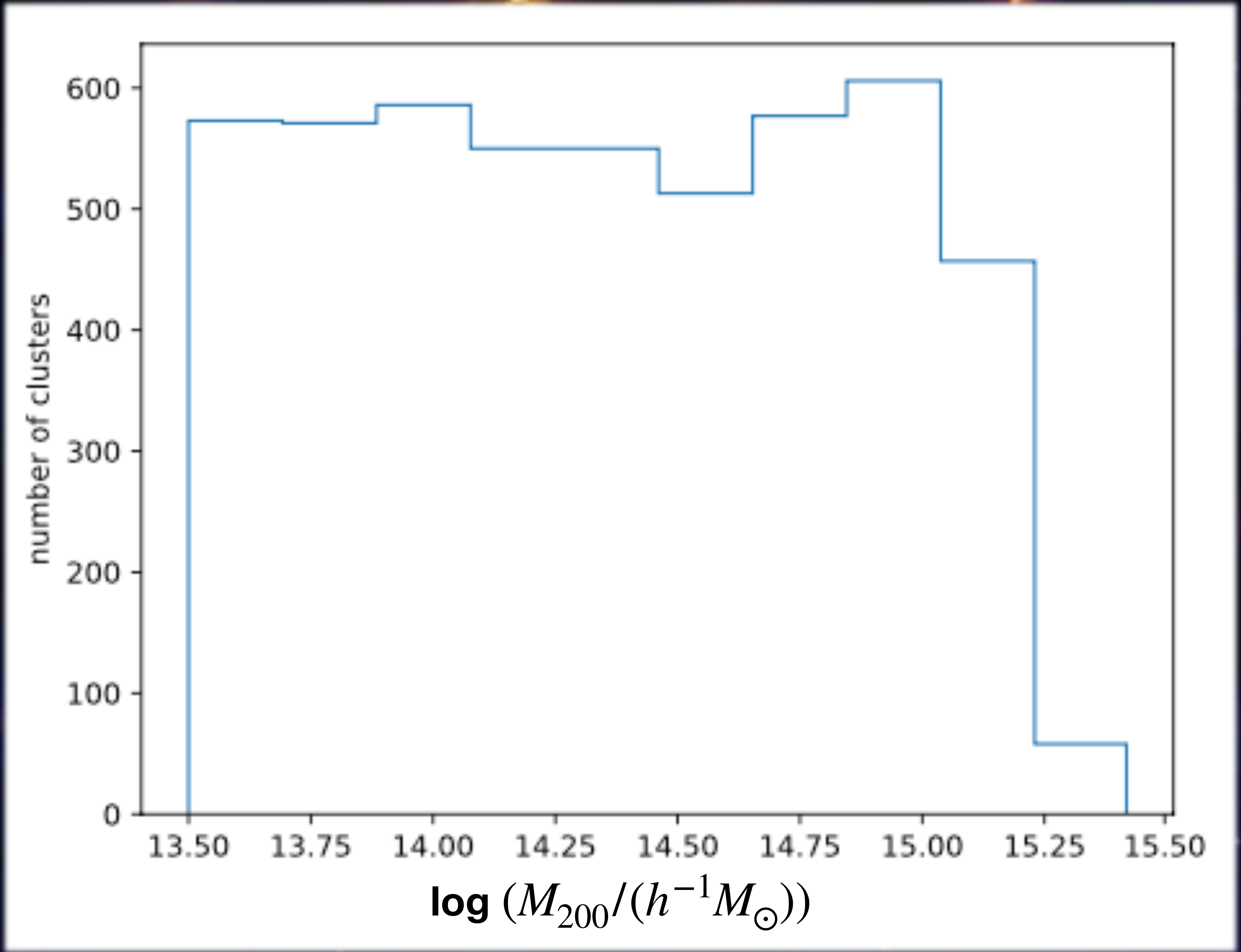
Mock data images

Mock data images

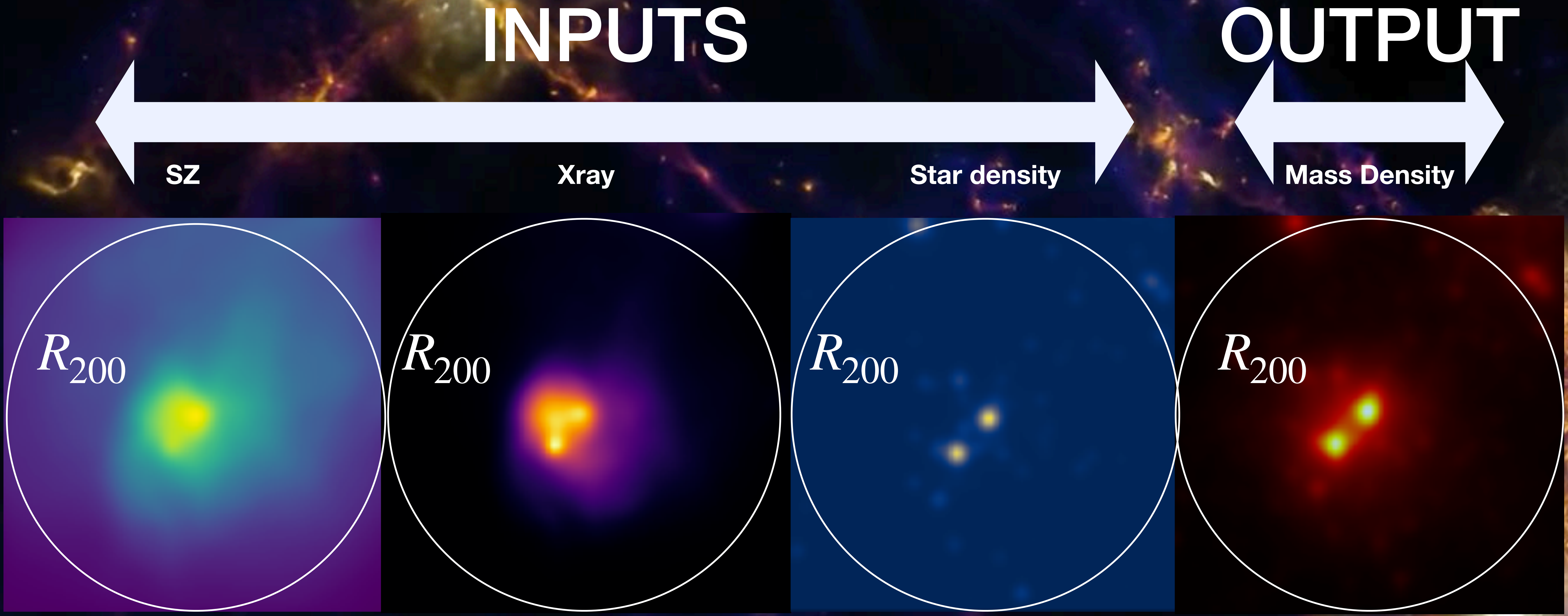
"Theoretical dataset" free from contamination/noise and no telescope's impact to test Deep learning models:

- **Compton-y parameter maps** PYMSZ, https://github.com/weiguangcui/pymsz
- **Bolometric X-ray surface brightness** estimated by emulating the X-ray energies by thermal bremsstrahlung in the hot intra-cluster medium using a wrapper of AtomDB https://atomdb.readthedocs.io/en/master/ , https://github.com/rennehan/xraylum
- **star density maps** are generated by projecting the sum of the masses of the star particles in the observer's line of sight. That value is divided by the surface area of a pixel.
- **mass density maps** are generated by projecting the sum of the masses of all the particles, i.e., gas, star, dark matter and black holes particles in the observer's line of sight

DATASET



- Only halos with $M_{200} > 10^{13.5}h^{-1}M_{\odot}$ are considered following a flat distribution in mass at redshift z~0.
- 29 l.o.s. projections and 5040 different halos and thus, ~**146 000 mock** images to train deep learning models.
- Density maps are sized such that the number of pixels is $2R_{200} = N_{pix}$= 80 .
- Maps are Gaussian-smoothed with a beam FWHM of ~$0.01R_{200}$.



The statistical inference problem is stated as follows: How much information is available on the input maps to reconstruct the output mass density map?

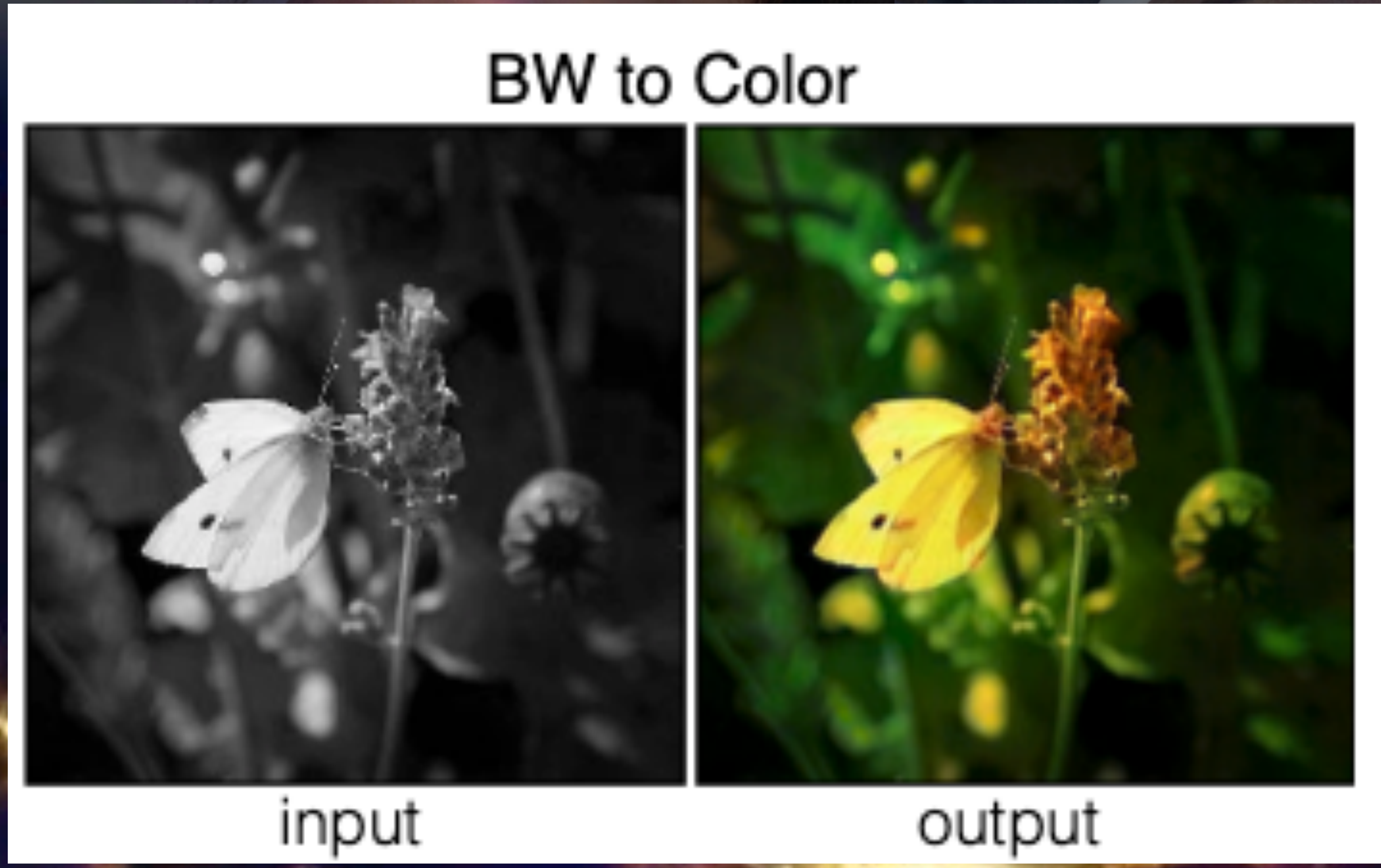
Deep learning models and results

Deep Learning is performing very well in other fields….

Image translation

Pix2Pix: Phillip Isola et al. 2018

These people do not exist



Our model

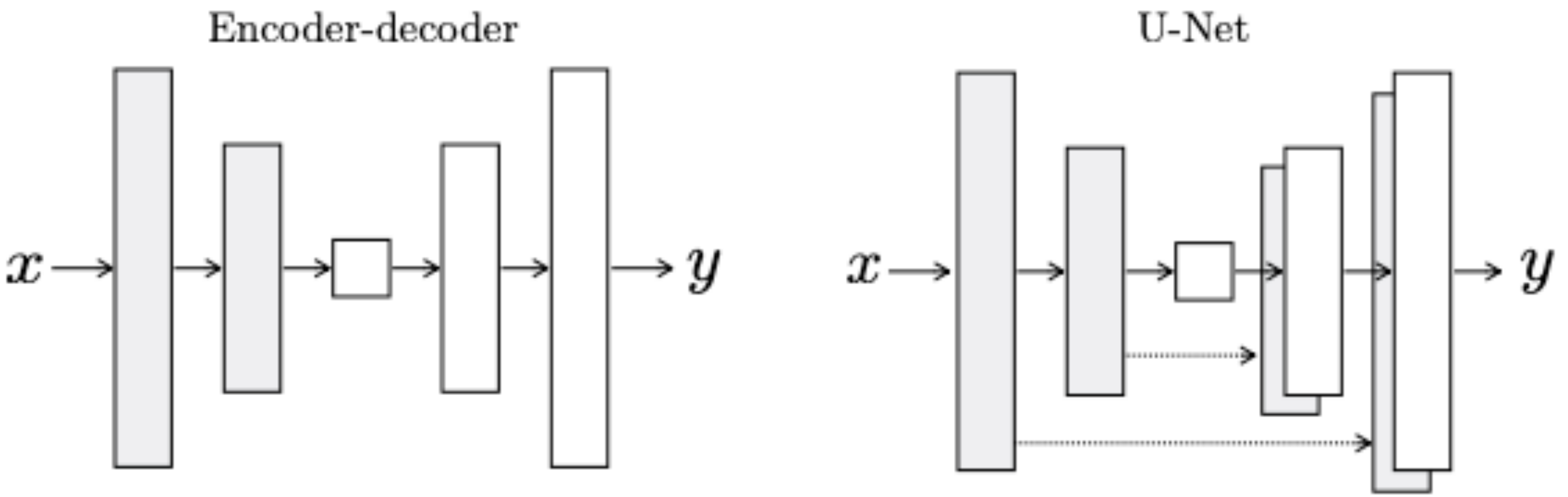


Figure 3: Two choices for the architecture of the generator. The "U-Net" [50] is an encoder-decoder with skip connections between mirrored layers in the encoder and decoder stacks.

Phillip Isola et al. 2018

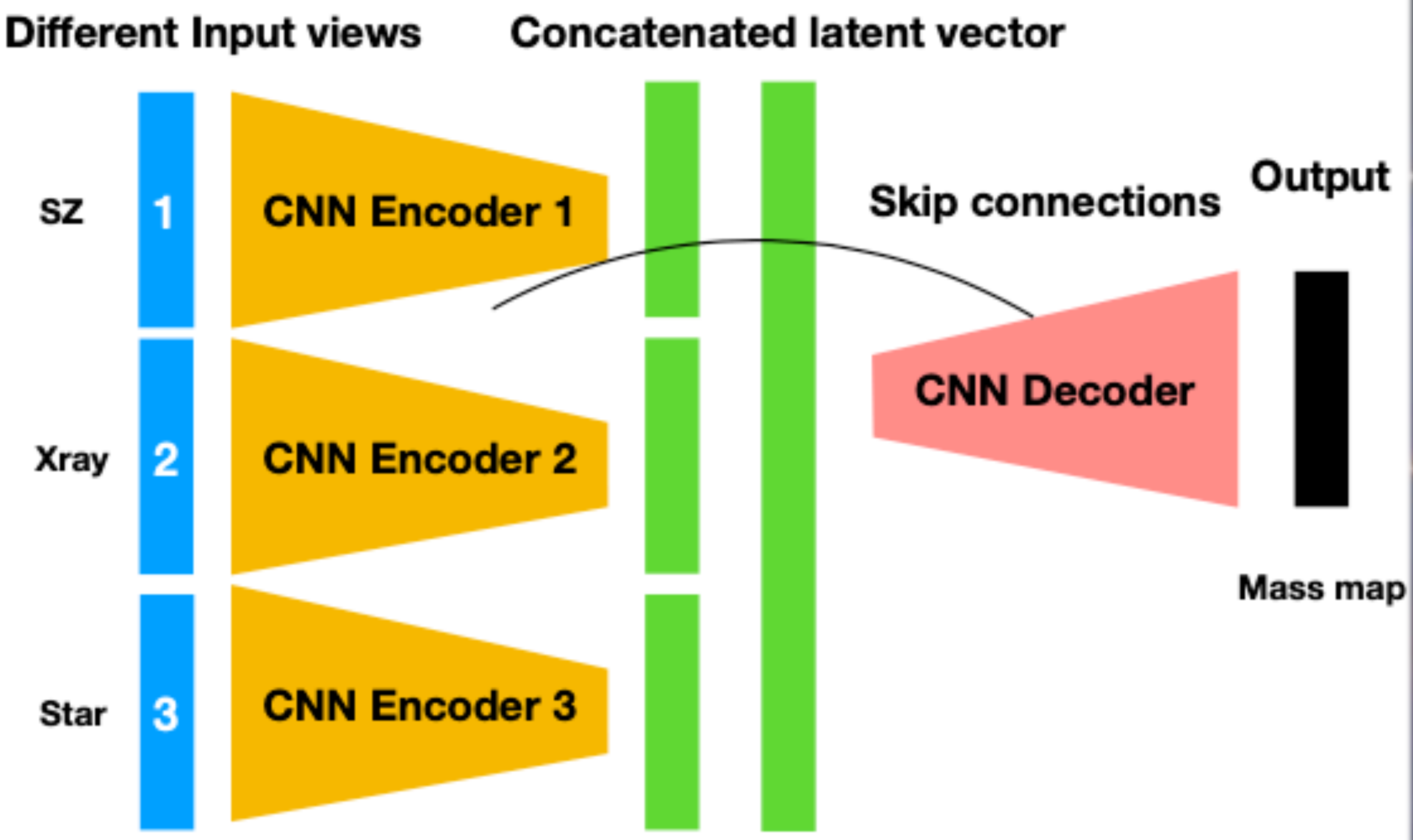
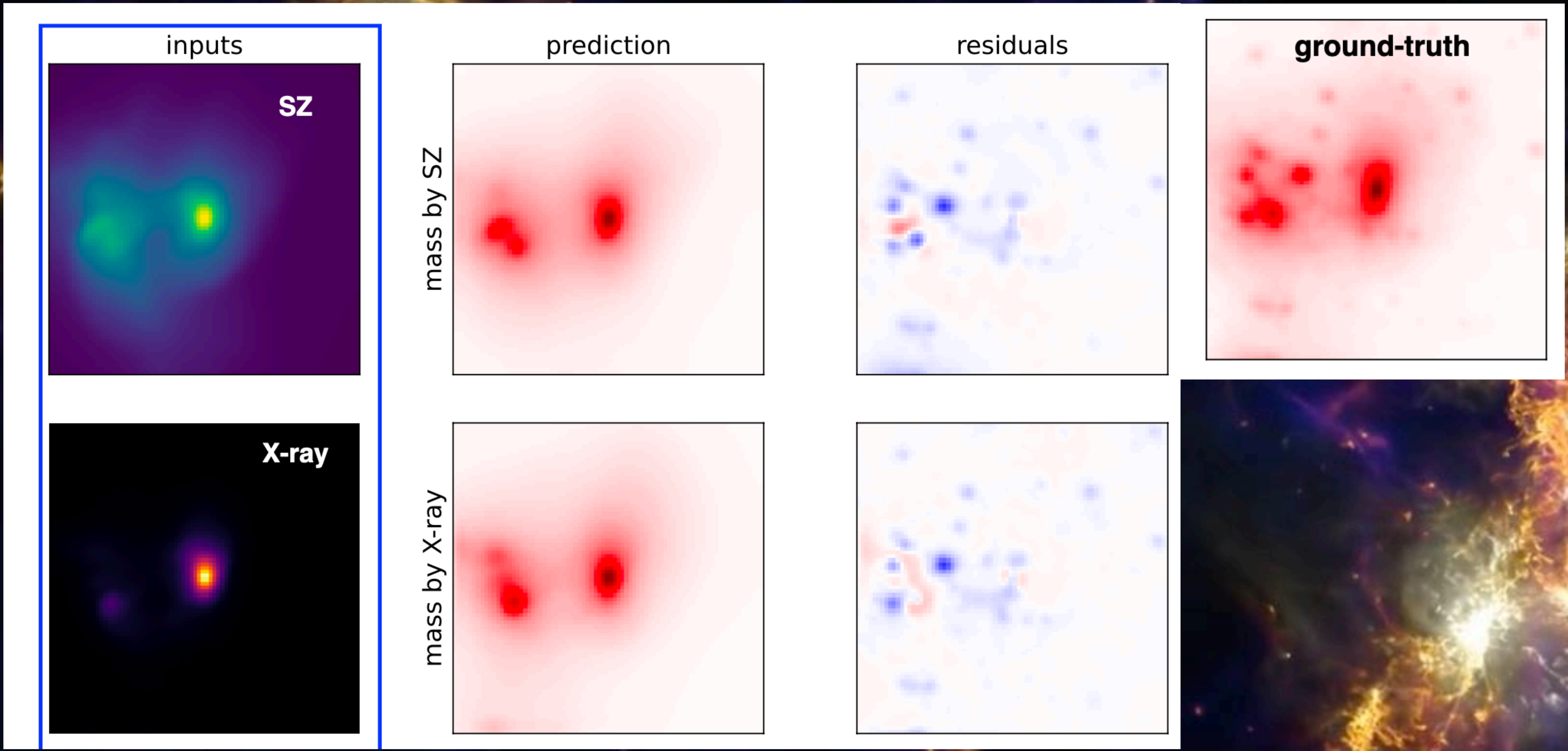


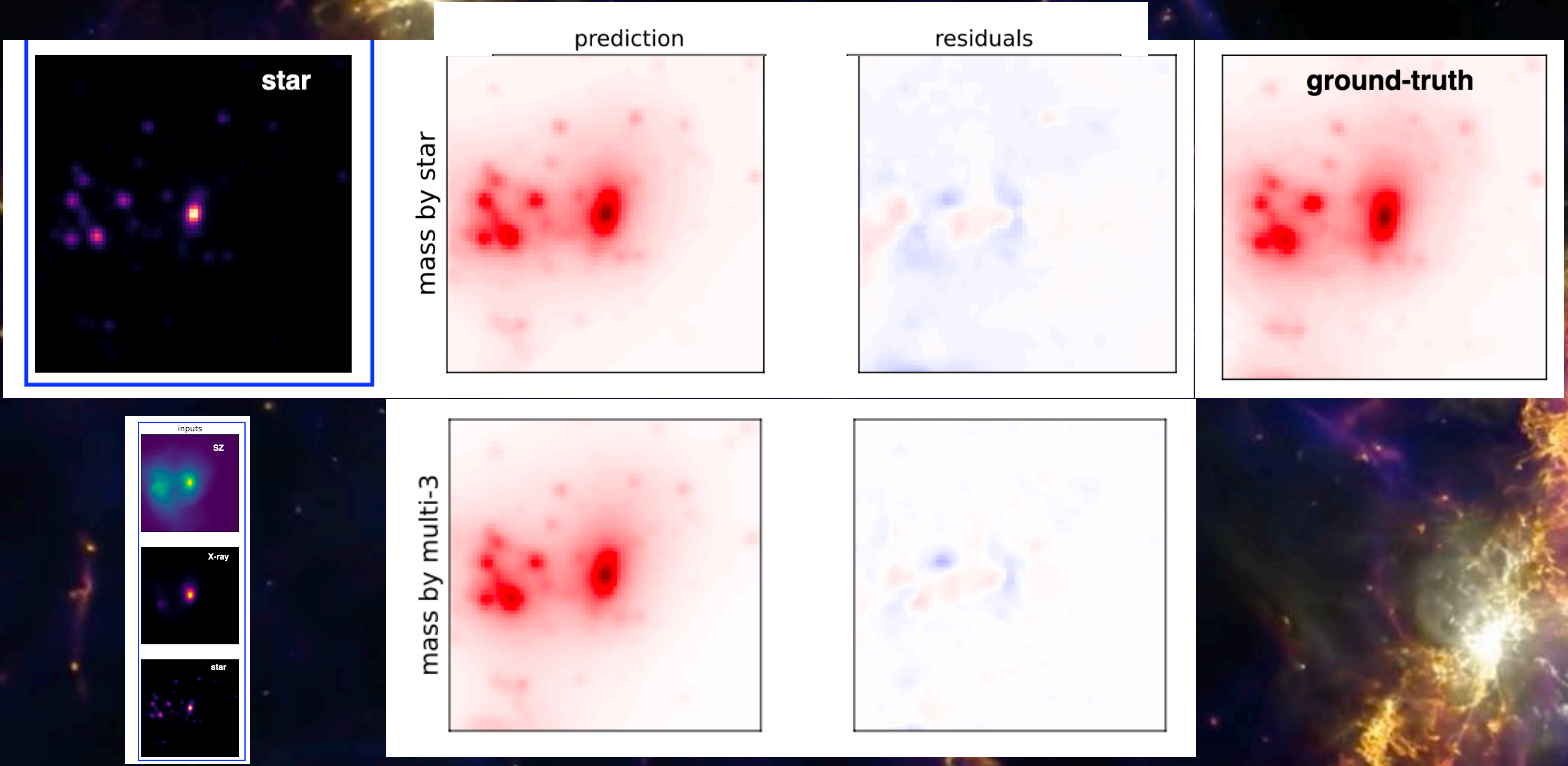
Figure 4. Multi-view approach

Daniel de Andres et al 2023 submitted

- One of the main advantages of this model is that it can make use of several input views to reconstruct the output density mass map.
- We train separately **4 UNETS varying the input view: SZ, Xray, star and multi view.**
- Models are trained with the L1 loss function (also tried a conditional GAN)
- Total 10 layers, with ~7M parameters.

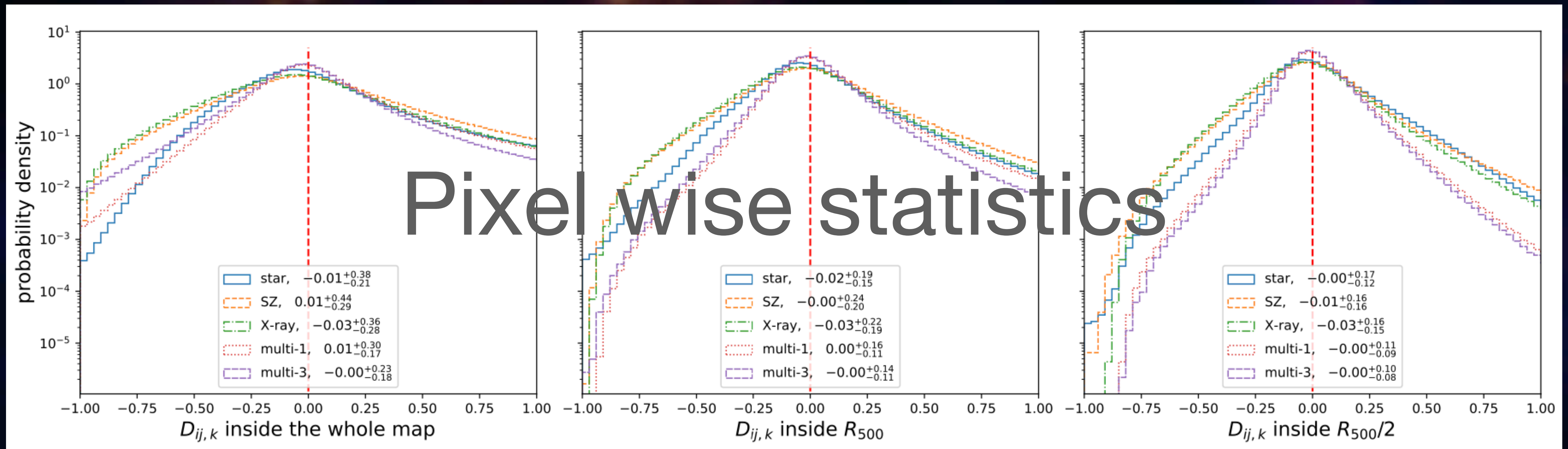


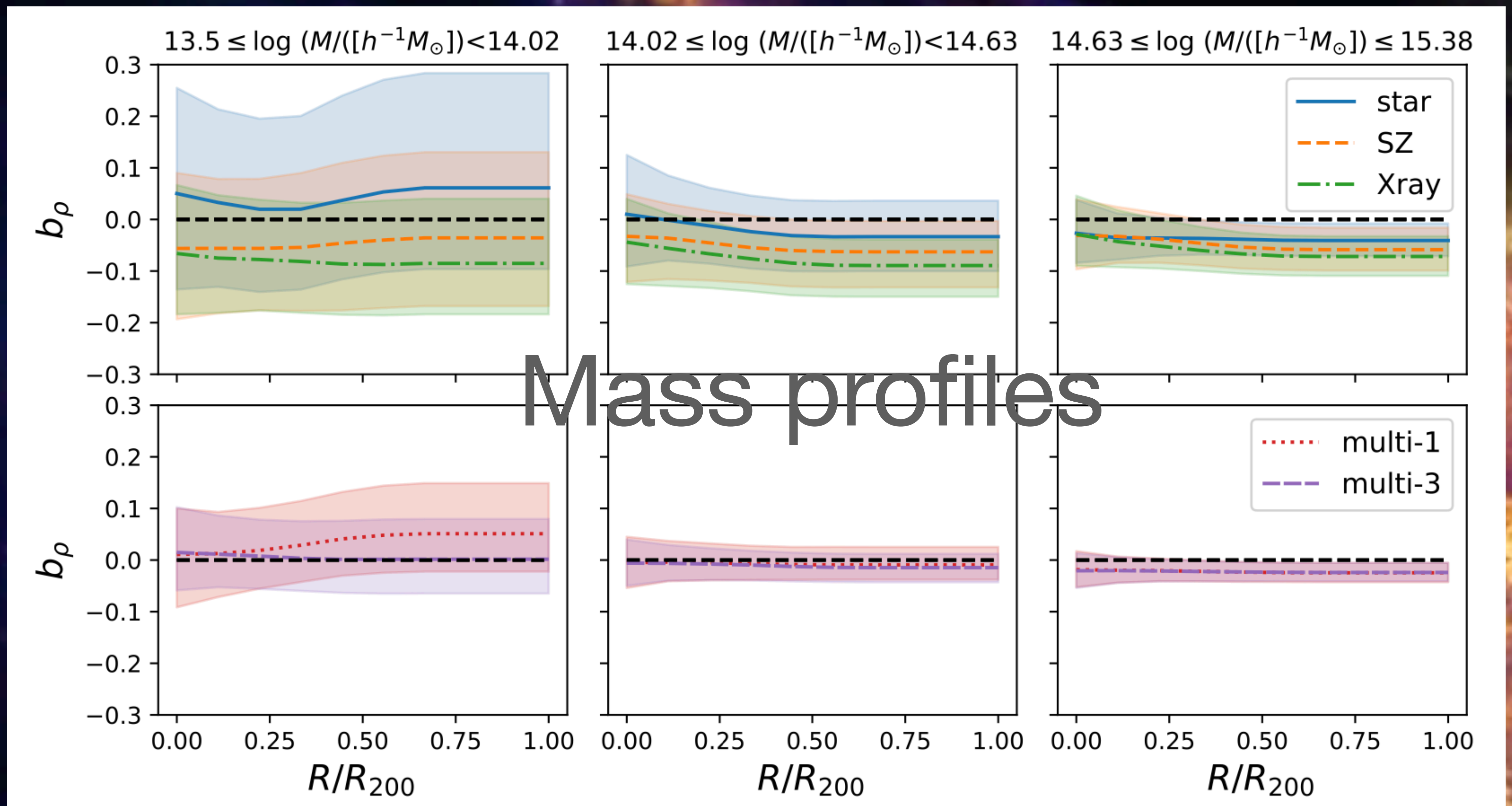
From gas tracers (SZ and X-ray) the residuals are the missing high spatial frequencies. But overall the mass distribution is well predicted.

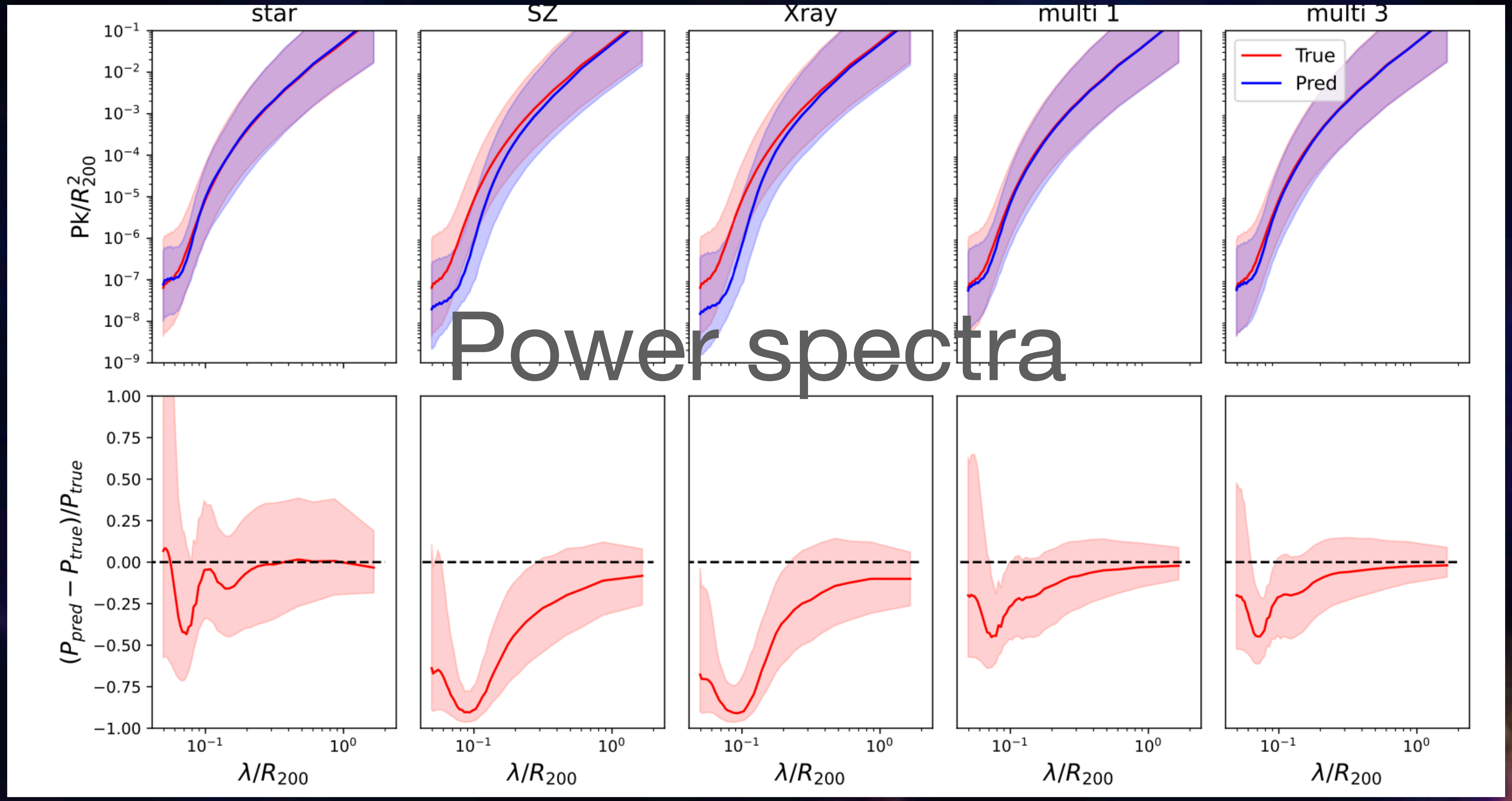


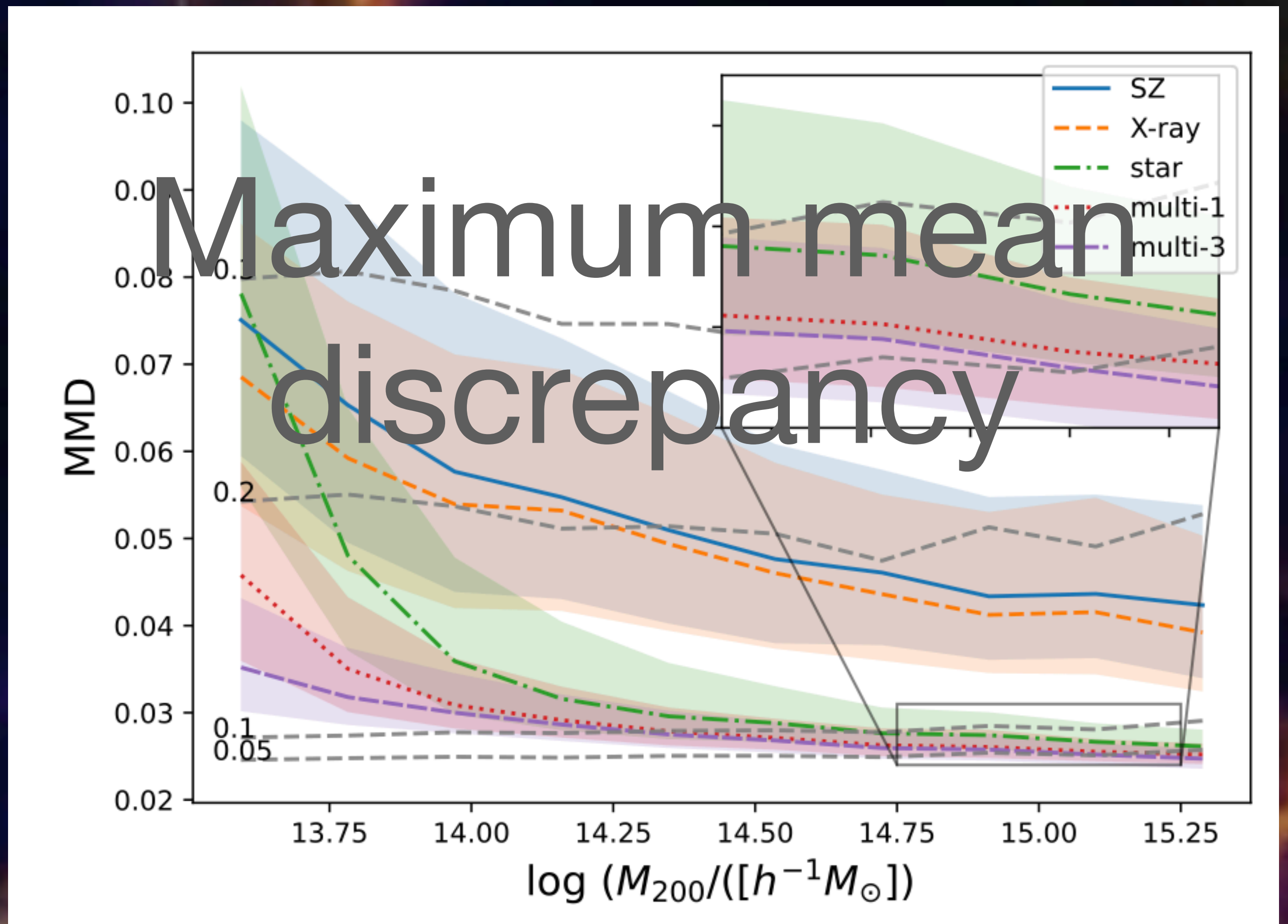
From star and multi-input the residuals are the missing small spatial frequencies. But overall the mass distribution is well predicted. By eye the ground-truth is the same as the generated maps.

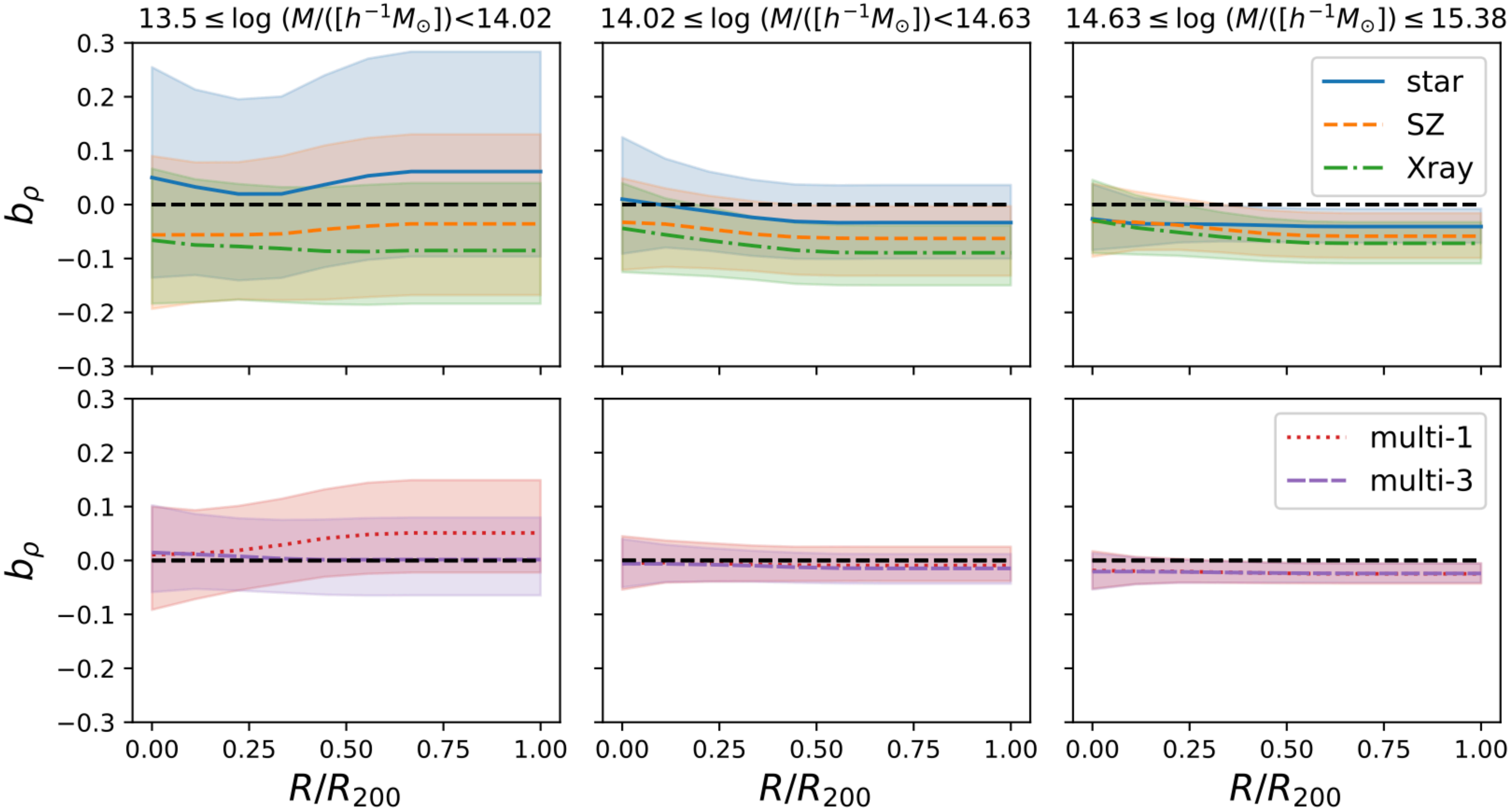
The main problem of deep learning is to what extent the generated mass maps are precise











$$b_\rho = \frac{\hat{M}_\rho - M_\rho}{M_\rho} \,.$$

Figure 7. Radial profiles of the mass bias, see Eq.(9). From left to right, we show the bias corresponding to different mass intervals as indicated in Figure 1: Interval 1 corresponds to $13.5 \leq \log M/\,h^{-1}\mathrm{M}_\odot < 14.02$, the interval 2 range is $14.02 \leq \log M/\,h^{-1}\mathrm{M}_\odot < 14.63$ and the interval 3 is $14.63 \leq \log M/\,h^{-1}\mathrm{M}_\odot \leq 15.38$. The top panels show the single-input view models (star, SZ, X-ray) and the bottom panels the multiview-1 and -3 models. The lines represent the median values per bin and the shaded regions cover the 16^{th} and 84^{th} percentiles. Furthermore, the particular enclosure mass bias at R_{200} is presented in Table 4.

Take home message

- The Three Hundred simulations provides a very good dataset (with hundreds of massive **galaxy clusters**) to train ML models. **Cosmological simulations are invaluable** tools to train ML models

- Deep learning models can be used for generating the 2D mass density distribution from observational SZ, X-ray, star data.

- This method has been tested with 'theoretical' simulated mock data.

- The objective is performing simulation based inference -> we are testing our models to inferred 2D mass density maps of NIKA2, SPT, eRosita, SDSS.

- The model architecture is flexible so that different luminosity bands can be combined. This application can be used for different photometric surveys.

Just-in-case slides

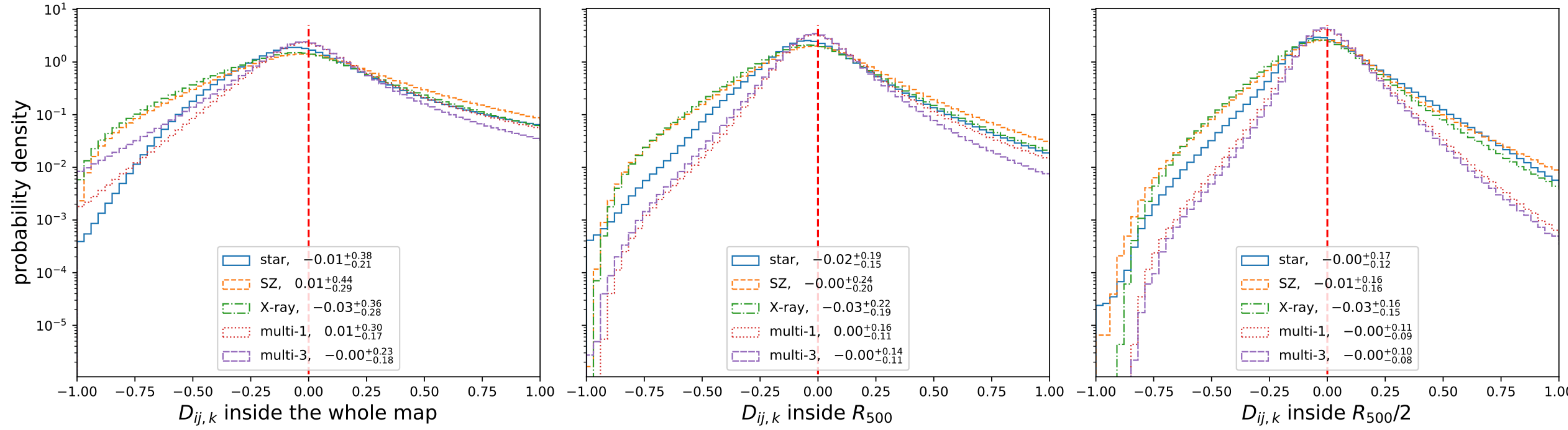


Figure 6. Pixel-wise relative difference $D_{ij,k}$ (see Eq.(6)) between the predicted mass density maps $\hat{I}_{ij,k}$ and the ground-truth mass density maps $I_{ij,k}$. Different lines represent the use of different input views to predict the mass density map: star, SZ, X-ray, multi-1 and multi-3. The legend also shows the median value of the distributions with the 16^{th} and 84^{th} percentiles as $\text{median}^{+|84^{\text{th}}-\text{median}|}_{-|\text{median}-16^{\text{th}}|}$. From left to right, $D_{ij,k}$ is computed for pixels inside various circular apertures: the whole map, inside R_{500} and inside $R_{500}/2$.

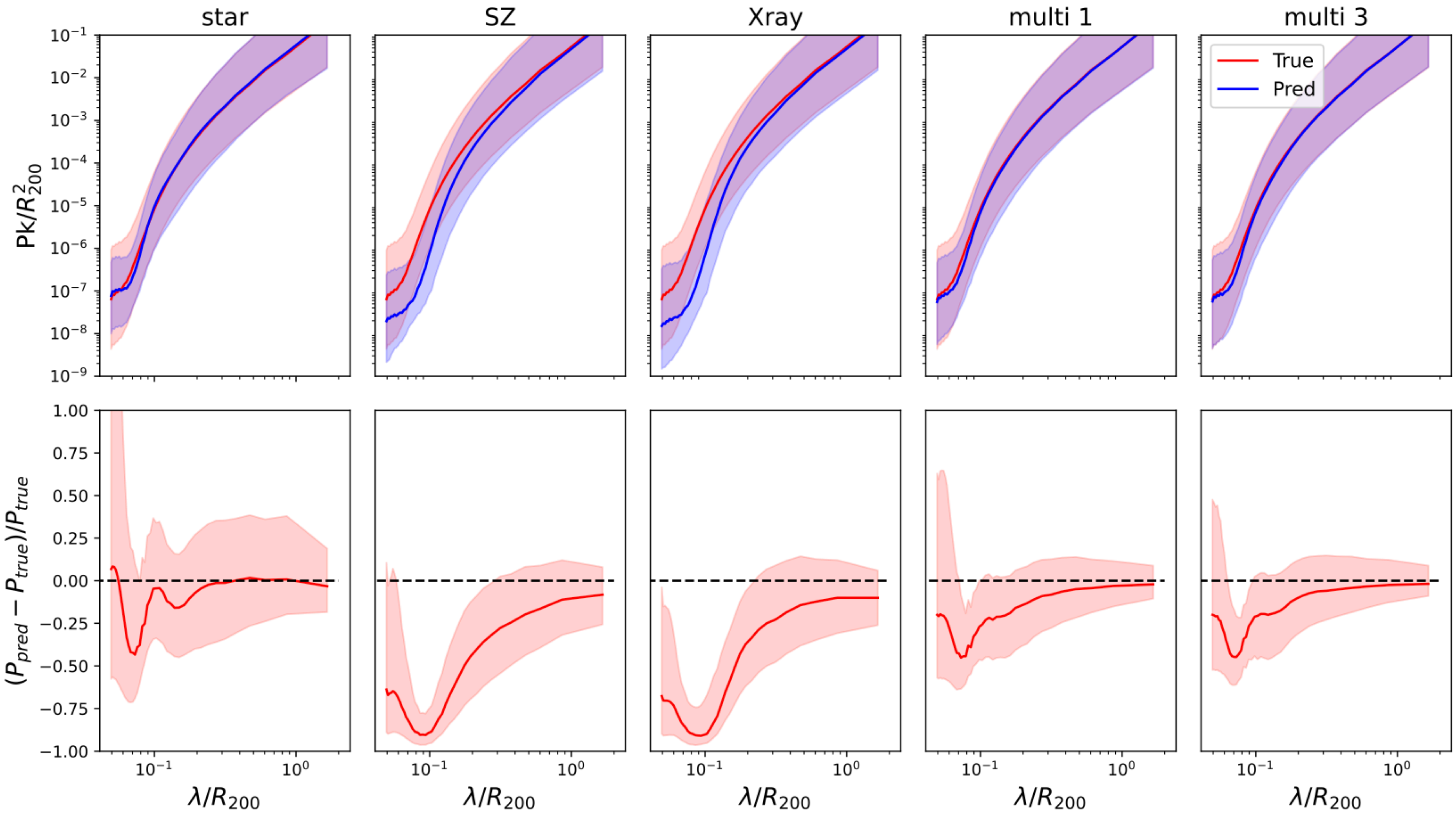


Figure 8. Top: Power spectrum corresponding to our ground-truth (red) and the predicted (blue) mass density maps as a function of the spatial length $L = 2\pi/k$ for our different inputs: star, SZ, X-ray, multi-1 and multi-3, are displayed in different columns. **Bottom:** We show the relative difference $(P_{\rm pred} - P_{\rm true})/P_{\rm true}$ of the predicted power spectrum $P_{\rm pred}$ and the ground-truth power spectrum $P_{\rm true}$ of our mass density maps. The dashed black line depicts the perfect prediction where the difference is zero. The solid lines correspond to the median values while the shaded regions represent the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles.

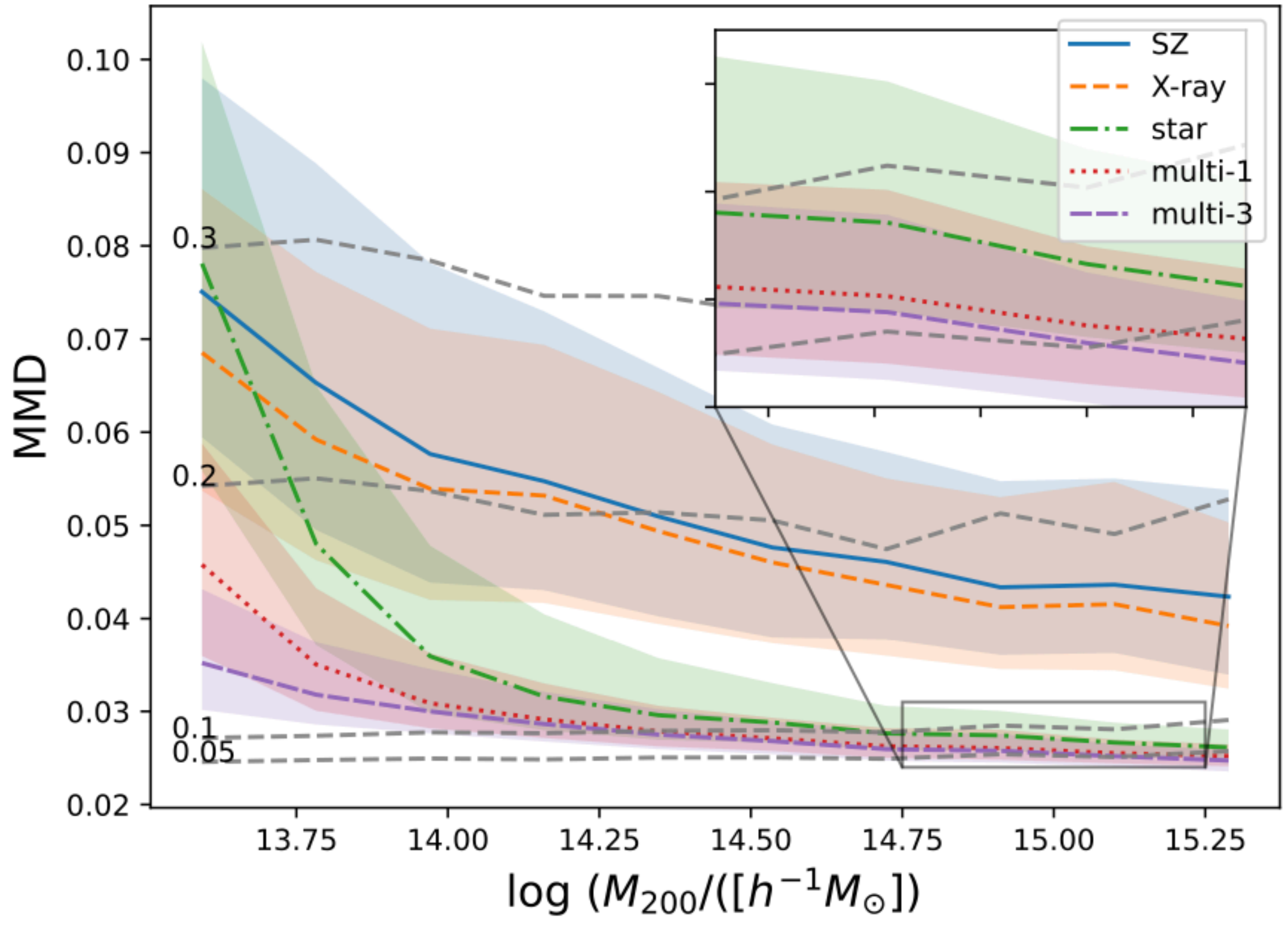


Figure 9. Maximum Mean Discrepancy, see Eq.(10), between predicted mass density maps and the ground-truth mass density maps as a function of the cluster mass M_{200}. Different colours represent the predictions when the model is trained with only SZ data (blue), X-ray data (orange), stellar data (green), multi-1 (red) and multi-2 (purple). Horizontal grey dashed lines represent the calibration of the MMD values and numbers written in black colour on the left of these lines correspond to the noise intensity σ used for the calibration of the MMD values. The inset highlights the results at higher halo mass where the multiview results are compatible with $\sigma \sim 0.05$.

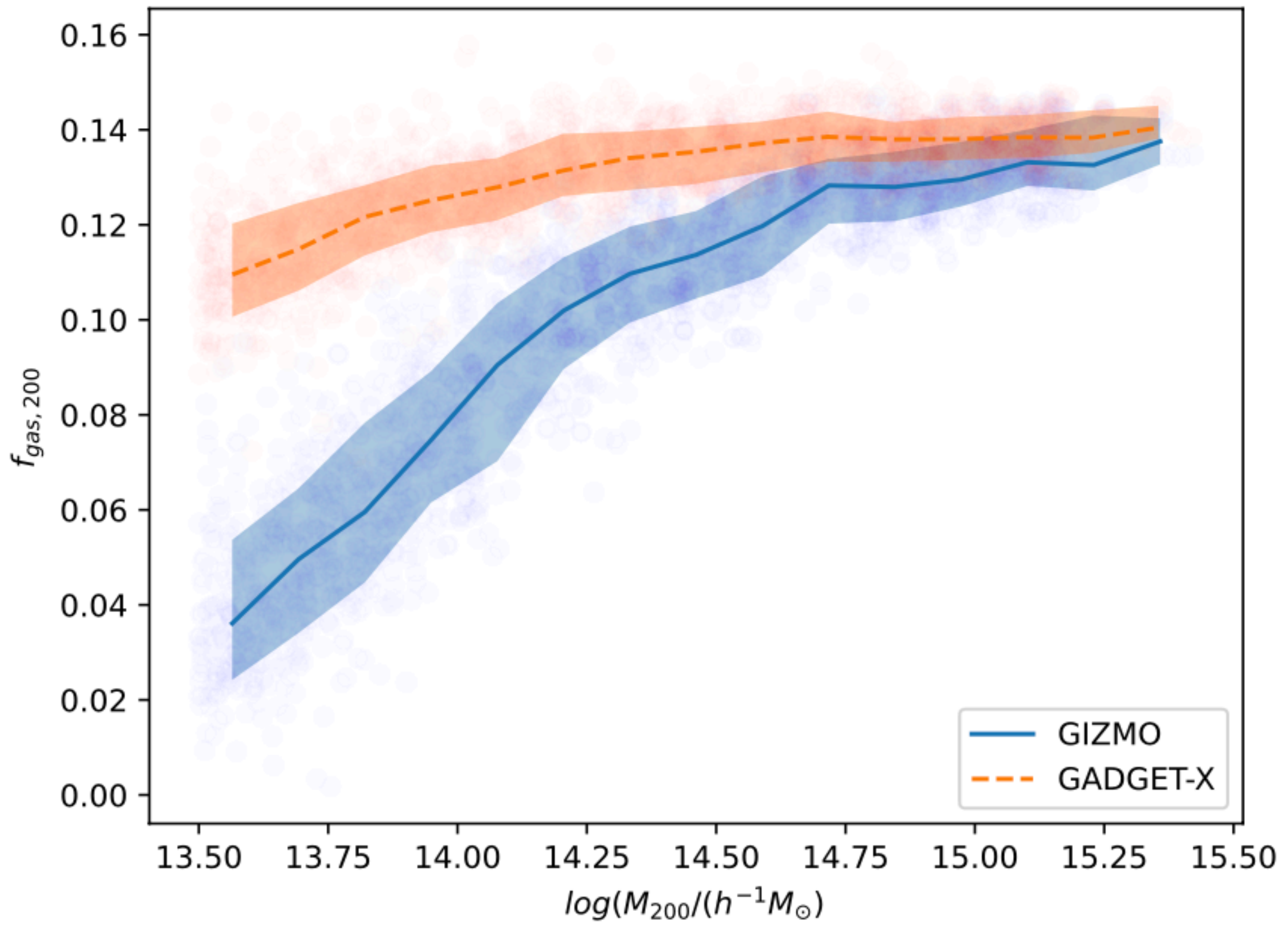


Figure A1. Median gas fraction as a function of the mass for the GIZMO and GADGET-X simulations. The shaded regions correspond to the 16th and 84th percentiles.

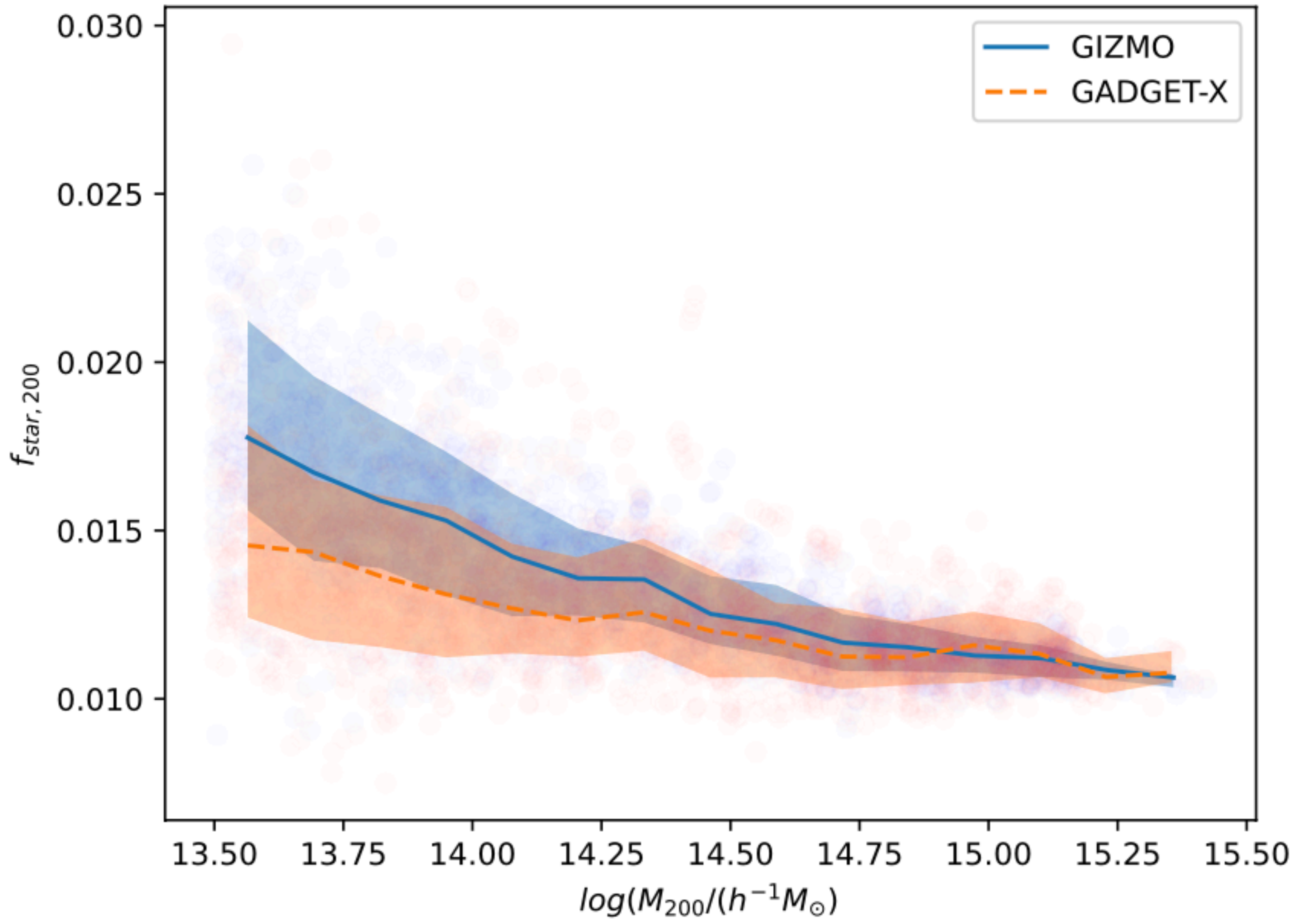


Figure A2. Median star fraction as a function of the mass for the GIZMO and GADGET-X simulations. The shaded regions correspond to the 16th and 84th percentiles.

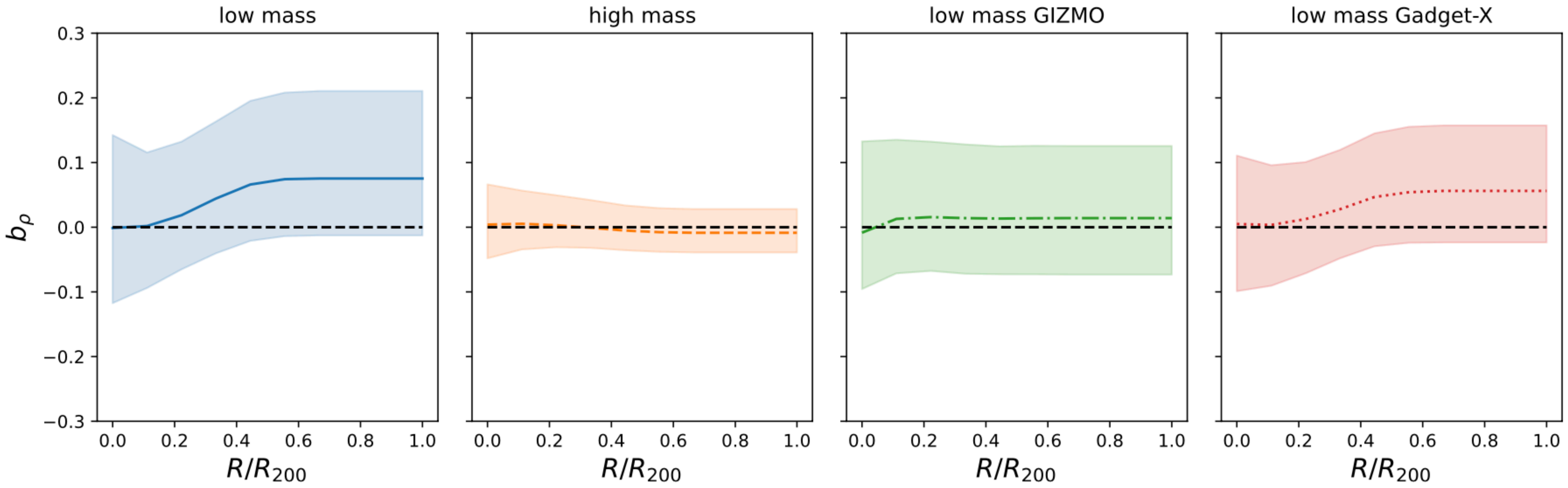


Figure A3. Mass bias see Eq.(9) for the mass profiles as a function of the radius R/R_{200} for the experiments previously mentioned in this appendix. From left to right, we show the bias corresponding to the different experiments: Train with low-mass objects with data from both simulations, train only with high-mass objects with data from both simulations, train low-mass objects with data only from GIZMO and train with low-mass objects with data only from GADGET-X. The lines represent the medium values per bin and the shaded regions cover the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentiles.