Plan de Recuperación, Transformación y Resiliencia

Financiado por la Unión Europea
NextGenerationEU

GOBIERNO DE ESPAÑA
MINISTERIO DE UNIVERSIDADES

SAPIENZA UNIVERSITÀ DI ROMA

Universidad de La Laguna

A. FERRAGAMO

(Post-doc at Instituto de Astrofisica de Canarias
Visiting Researcher at Sapienza Università di Roma)

D. DE ANDRES, A. SBRIGLIO , W. CUI, M. DE PETRIS,
G. YEPES, R. DUPUIS, M. JARRAYA,I. LAHOULI,
F. DE LUCA, G. GIANFAGNA, E. RASIA

Ferragamo et al. (2023) MNRAS, 520, 4000. doi:10.1093/mnras/stad377

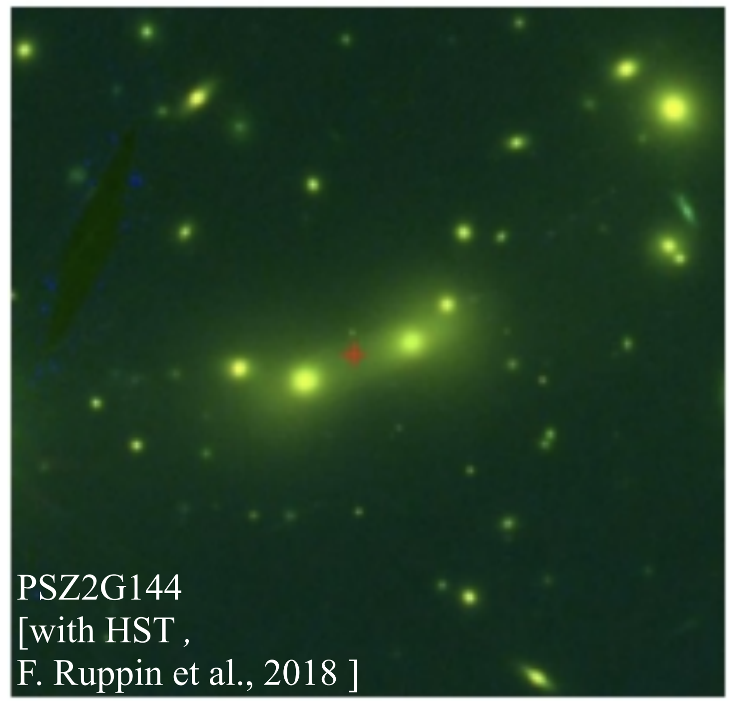
Machine Learning approach to Sunyaev-Zel'dovich cluster radial mass profiles

Antonio Ferragamo — mmUniverse@NIKA2 — Grenoble 26-30 June 2023

Cluster observables and mass estimate

Self similarity (scale invariance) → observable physical quantities are related to the cluster mass

Optical/Infrared:
Stellar light of the galaxies

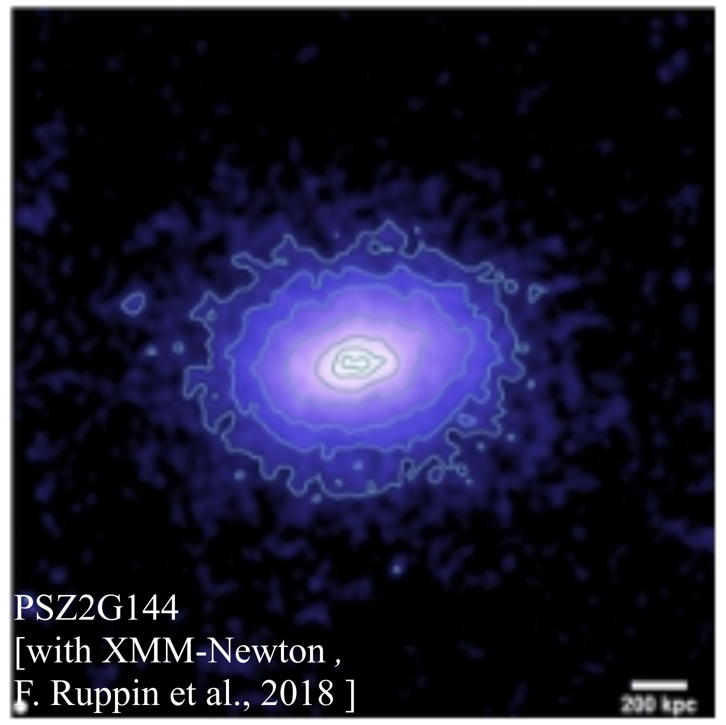


PSZ2G144
[with HST ,
F. Ruppin et al., 2018]

- Lensing of background galaxies
- Galaxy members richness and
- velocity dispersion

→ Different **mass** estimates

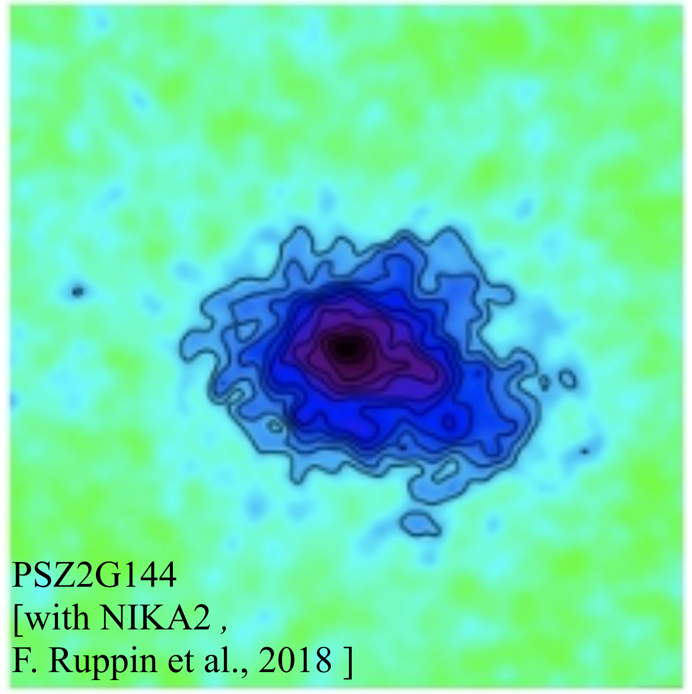
X-ray:
the hot ICM emission in X-ray



PSZ2G144
[with XMM-Newton ,
F. Ruppin et al., 2018]

- Spectral line emission of metals
  → Temperature
- Bremsstrahlung of electrons
  → Electron density

Millimeter wavelengths:
Sunyaev-Zel'dovich (SZ)



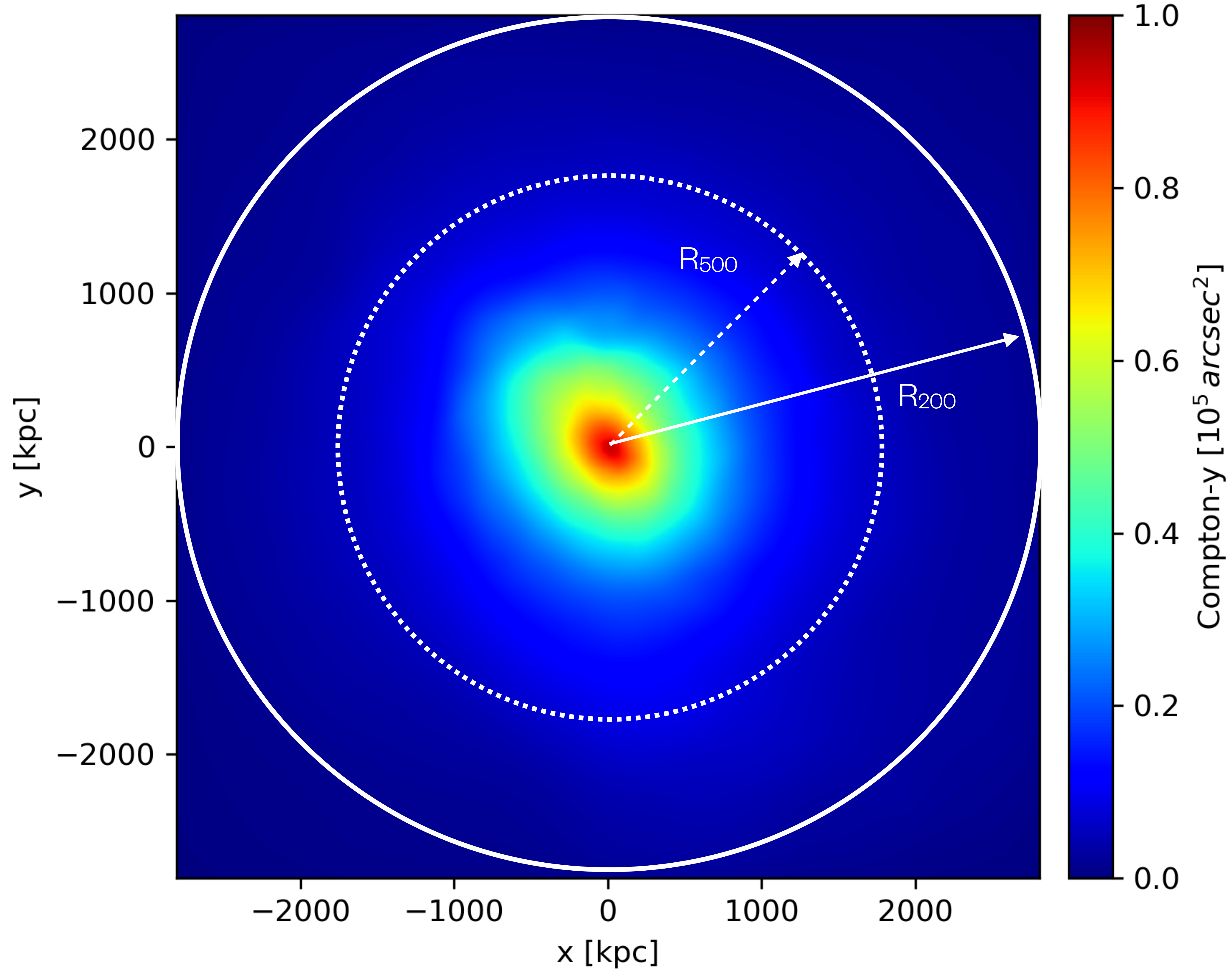
PSZ2G144
[with NIKA2 ,
F. Ruppin et al., 2018]

→ Thermal pressure of the ICM

→ **Hydrostatic mass** estimates from X-ray and SZ

Different mass estimates affected by **different systematic** uncertainties. **Combining observables** may help building a consistent picture of the cluster physics to gain accuracy on the mass estimates

Hydrostatic mass profile



$$y_\Delta = \frac{\sigma_T}{m_e\, c^2\, D_A^2} \int_0^{R_\Delta} P_e(r)\; 4\pi r^2 \mathrm{d}r$$

$$\Delta = \frac{\langle \rho\,(<r)\rangle}{\rho_c(z)}$$

Thermal pressure of the ICM
from SZ effect

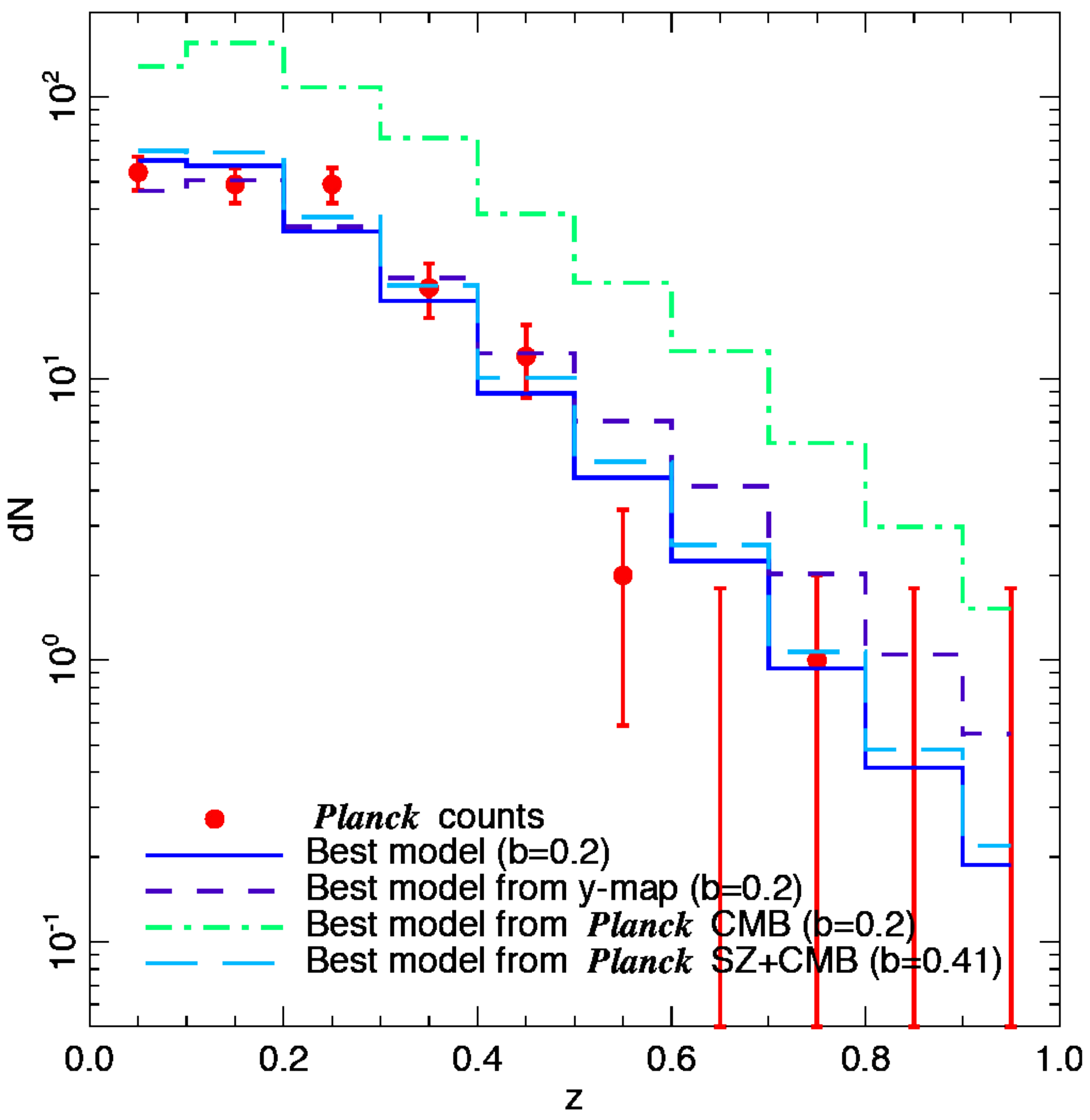
$$\boxed{M_{\mathrm{HE}}\,(<r) = -\frac{r^2}{G\mu m_p n_e(r)} \frac{\mathrm{d}P_e(r)}{\mathrm{d}r}}$$

IT IS WELL KNOWN THAT THE HYDROSTATIC MASS ESTIMATES IS BIASED

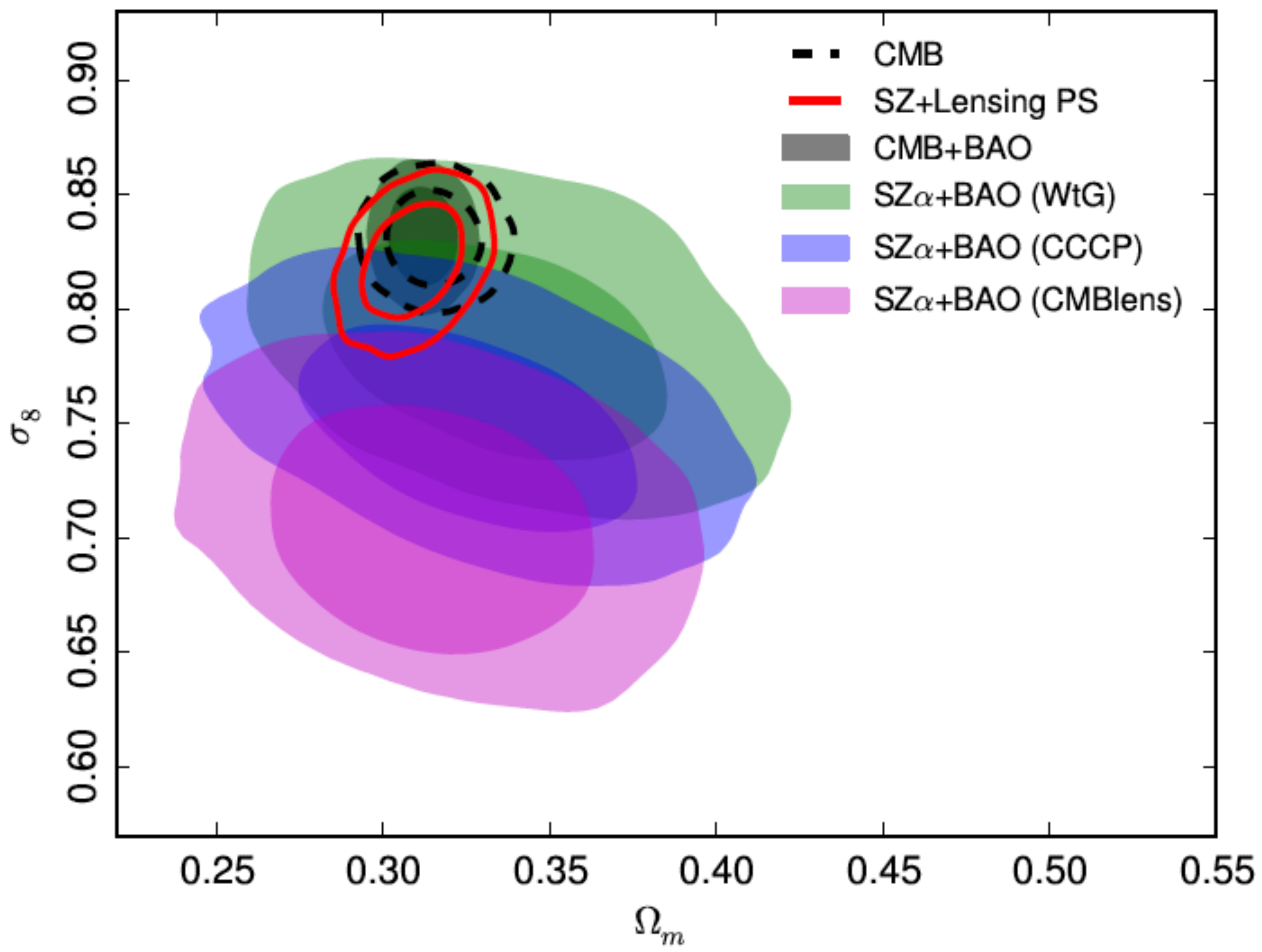
Cosmology with galaxy clusters: number counts

Theory predicts more clusters than observed

Planck Collaboration XX (2013)



Planck Collaboration XXIV



TENSION IN COSMOLOGICAL PARAMETERS

Unbiased cluster mass estimate can alleviate the tension

Machine Learning

Machine Learning algorithms enable data analysis and prediction without assuming any previously known behavior

Pros:

- Uncorrupted by hypotheses/ assumptions
- Programmable and able to faithfully reproduce particular attitudes
- Huge masses of data —> full potential of new impending technologies

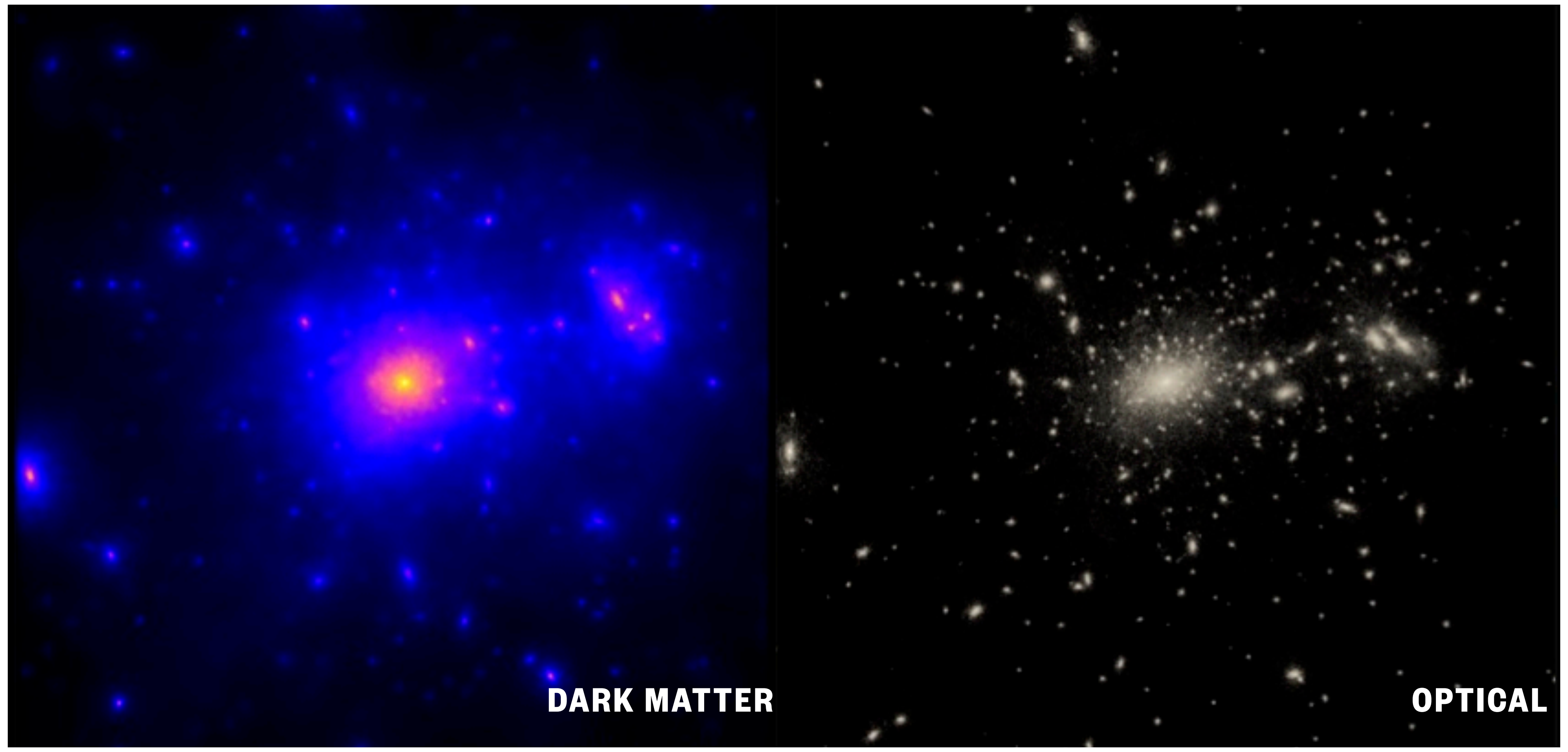
Cons:

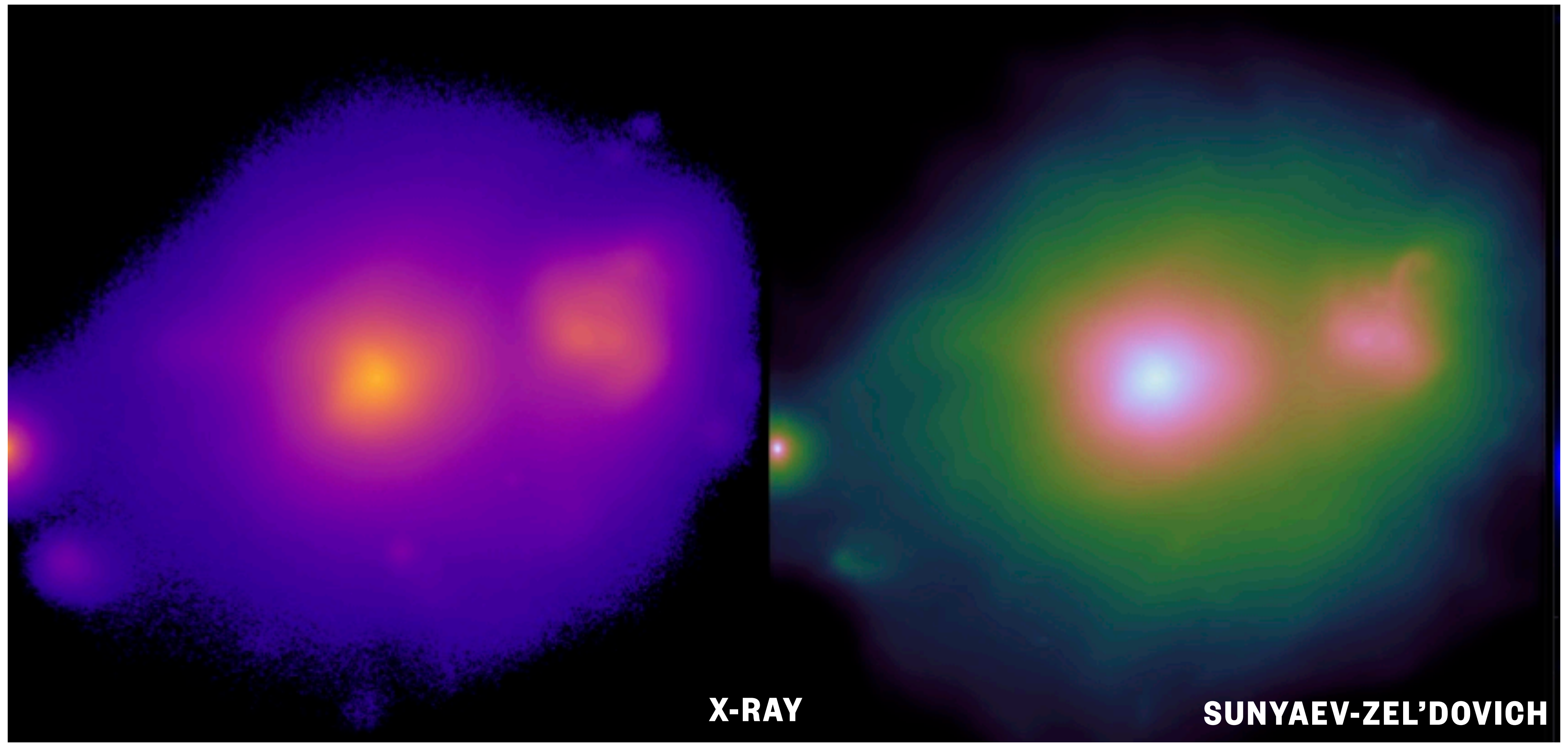
- Internal mechanisms are often black boxes —> complexity in interpretations
- Possibile overfitting

THE THREE HUNDRED Project

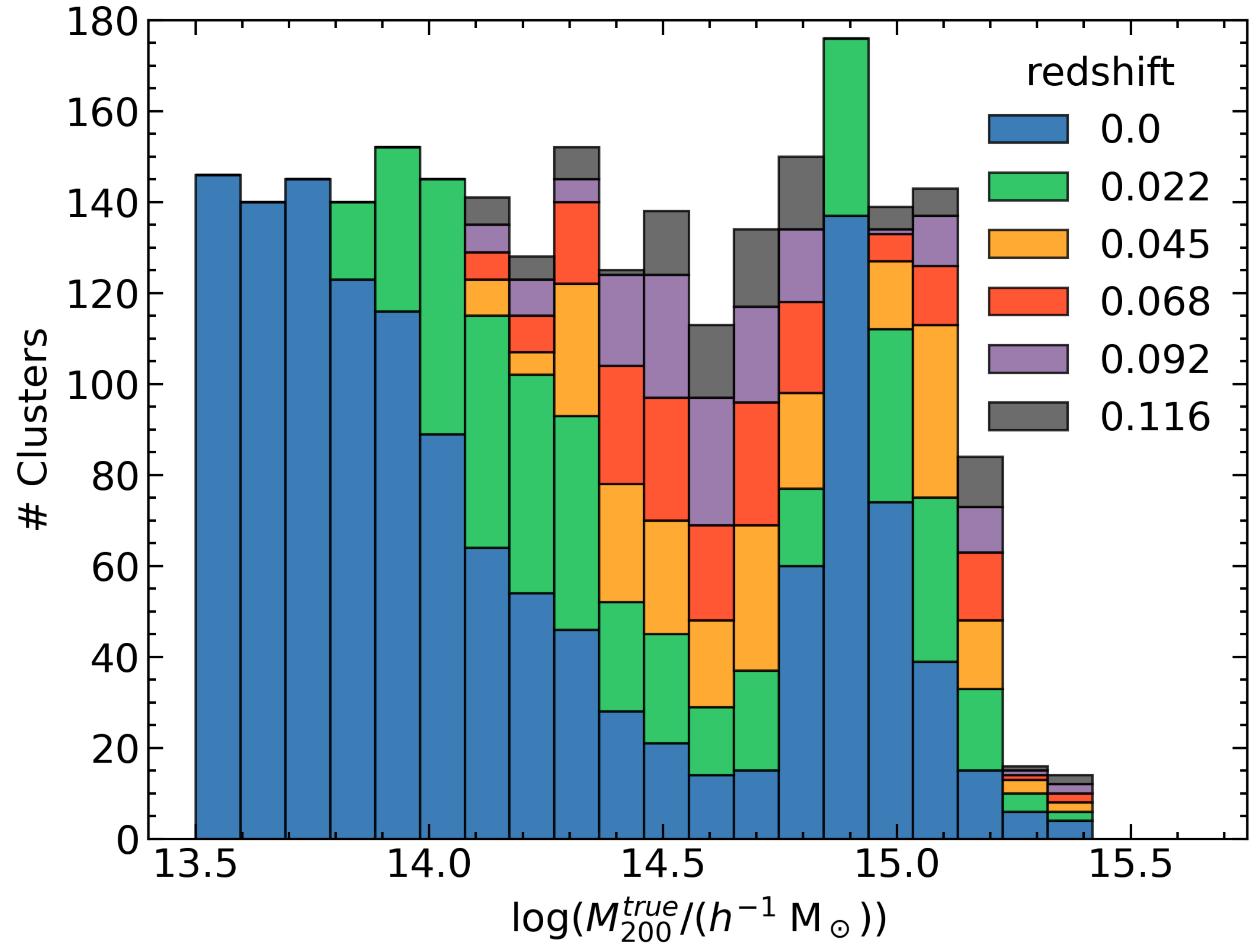
Spherical Zoomed regions of 15 h^{-1} Mpc radius centred around the 324 most massive clusters of the Multidark-Planck simulation formed at z=0 (Mass: $8x10^{14}$ – 3.2x1015 h^{-1} $M_{\odot}$)

- 4 large spherical regions (R>30 h^{-1} Mpc) "void" of clusters (M <10^{13})
- A few clusters will be resimulated at **high resolution**, with up to 7680^3 effective particles ($2x10^8$ $M_{\odot}$ per dark matter particle)
- **DATA SAMPLE:**
  - 3 different hydrodynamical runs of the same objects (324 regions)+ 4 void regions
    - GADGET-MUSIC (standard SPH, SN Feedback, Stellar winds)
    - GADGET-X (modern SPH, AGN feedback, Trieste Model)
    - GIZMO-SIMBA (modern SPH + AGN feedback Dave's Model)
- **SAM results** from the Multidark simulation @3840^3 mass resolution for the same regions:
  - Galacticus, SAG, SAGE
- **Mock observations provided:**
  - X-ray (Chandra, Athena, XMM) , tSZ(y-maps), CCD (SDSS bands), lensing maps.





Data-set



2580 GADGET-X galaxy clusters

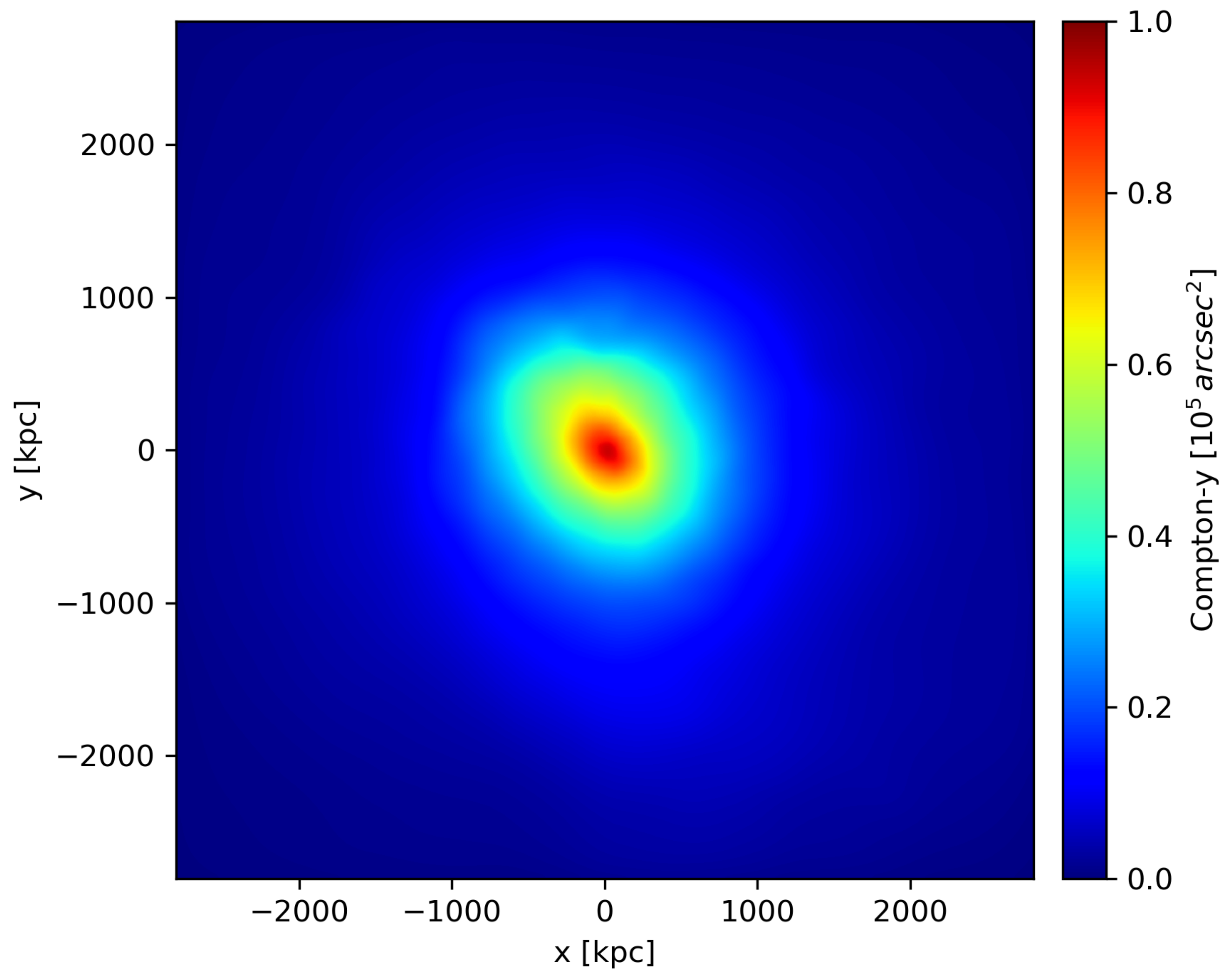
$$3\times 10^{13} \leq M_{200}\ h^{-1}M_\odot \leq 2\times 10^{15}$$

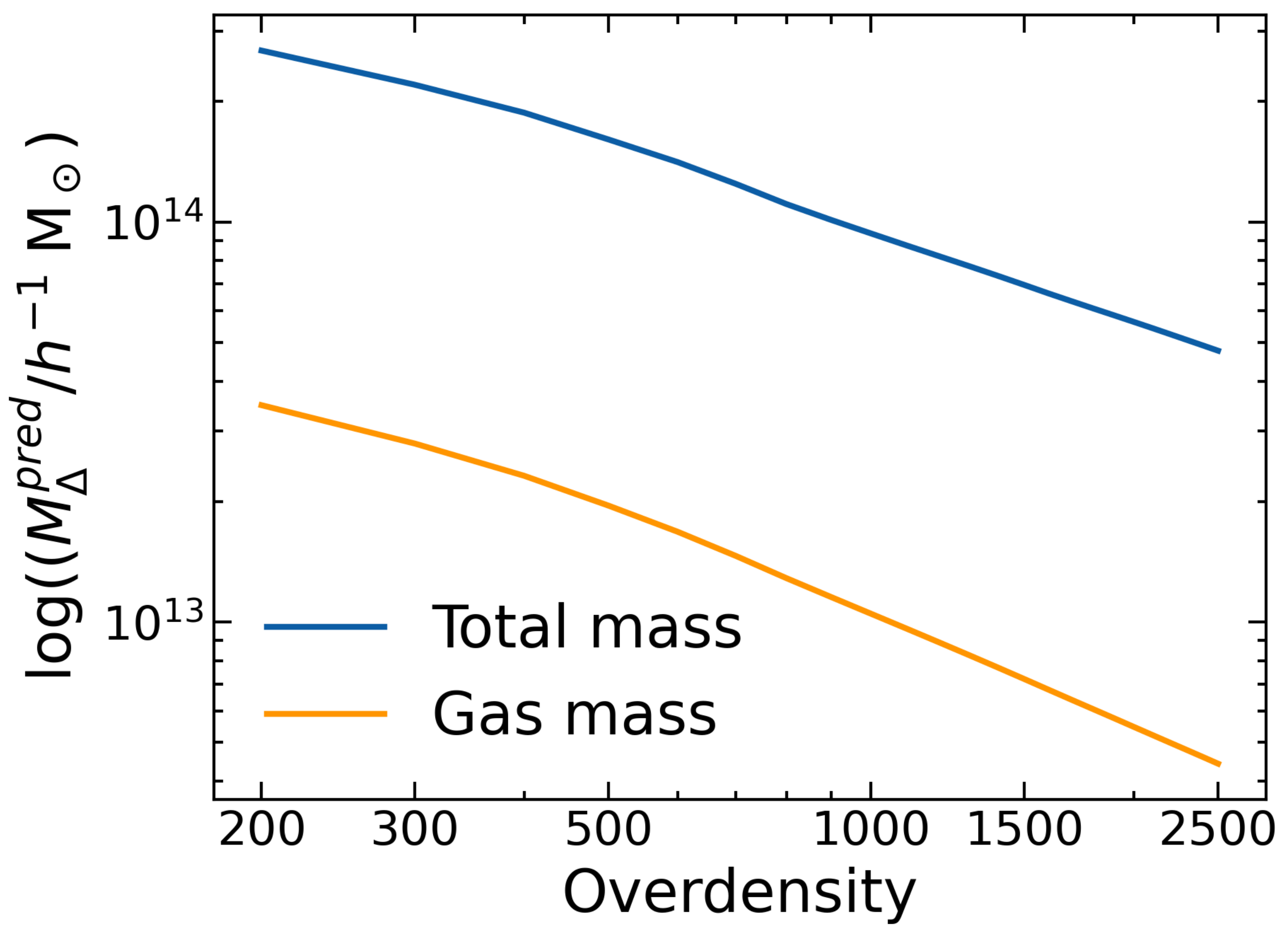
$$0.0 \leq z \leq 0.116$$

SZ maps along 29 l. o. s.

128x128 pixels

$$2R_{200}\times 2R_{200}$$





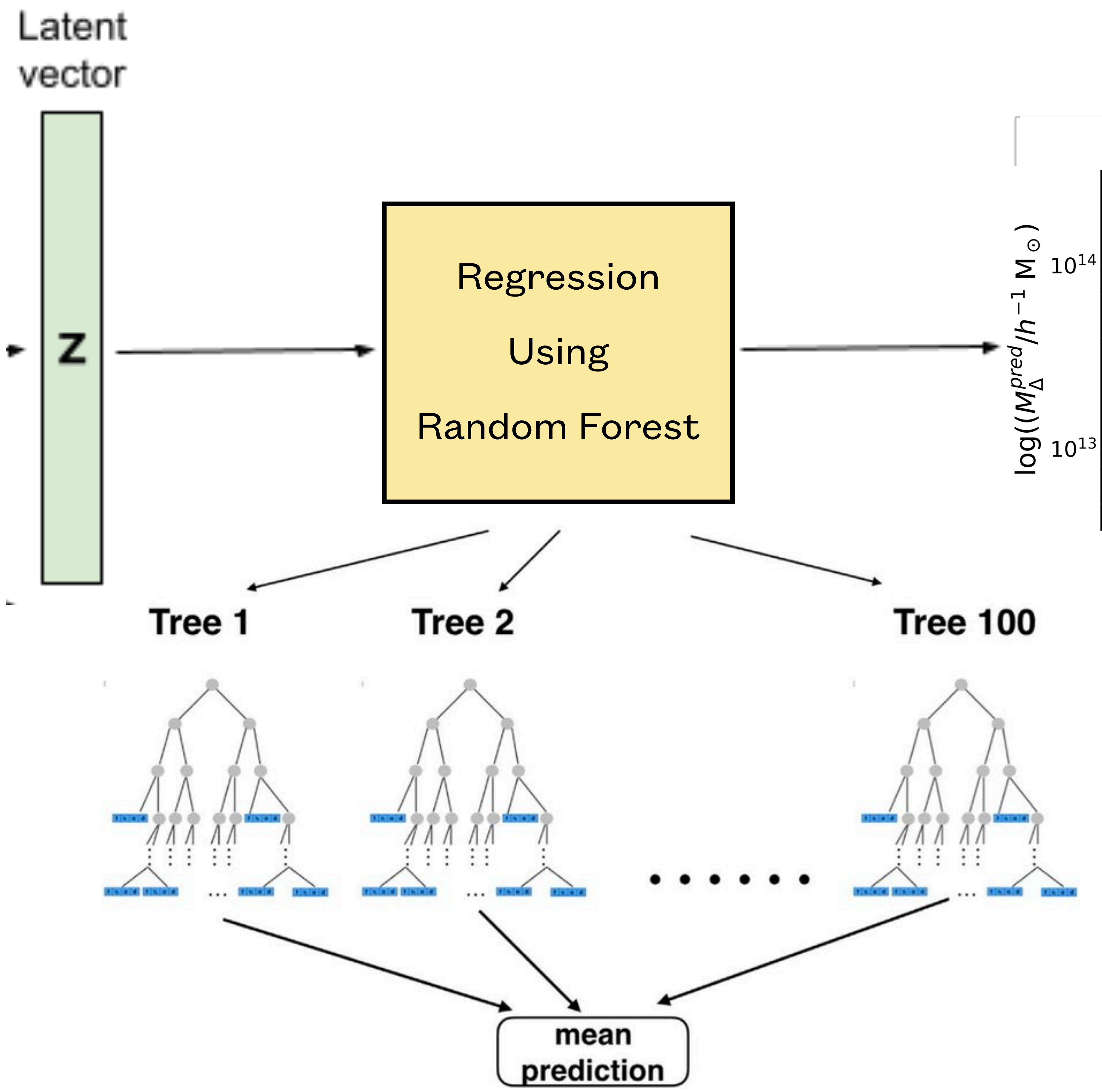
3D radial total and gas mas profiles

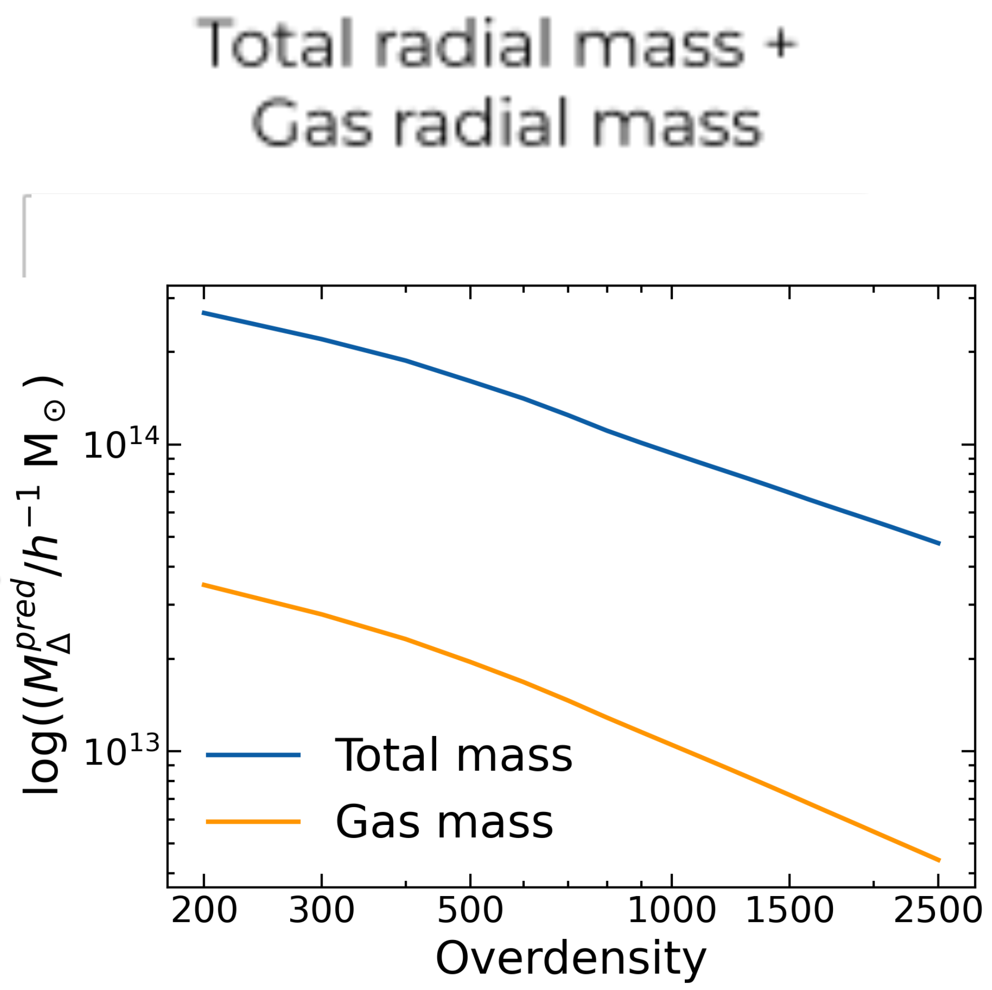
Encoding SZ maps

We used an **Autoencoder** unsupervised learning convolutional neural network

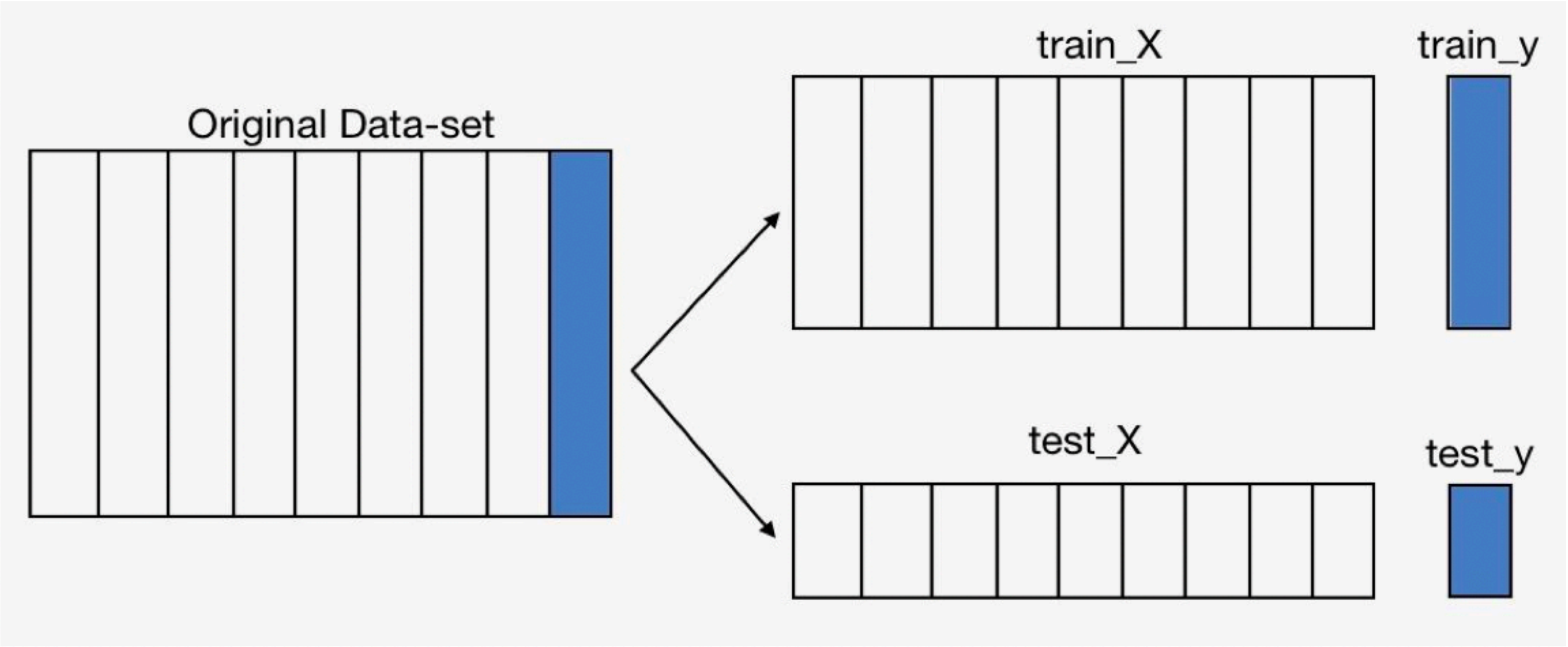
With the informations encoded in the Latent vector we are able to reconstruct the SZ map

Finding the way in the Forest



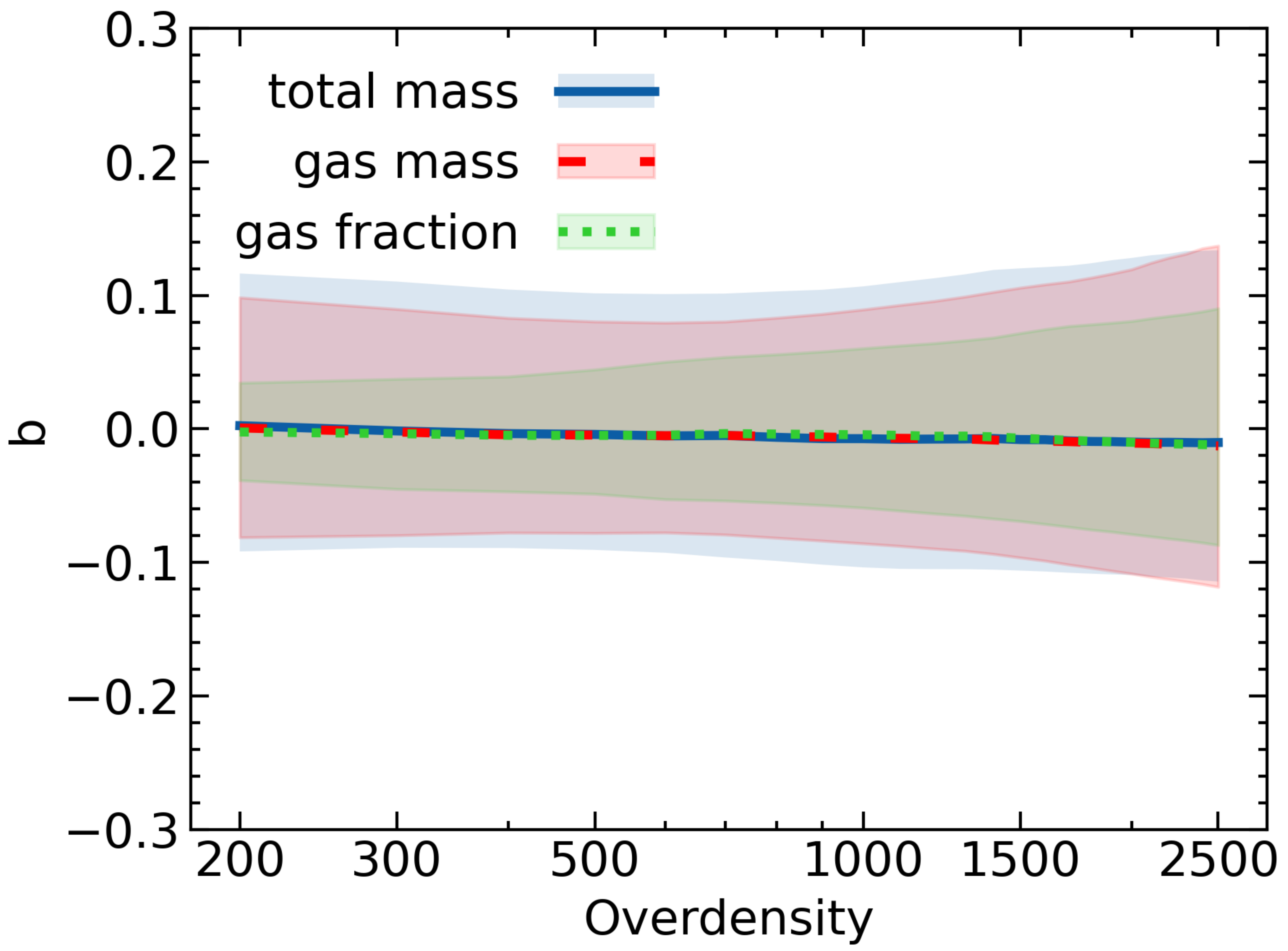


A RANDOM FOREST regressor is used to correlate the informations encoded in the latent vector with the total and gas mass profiles

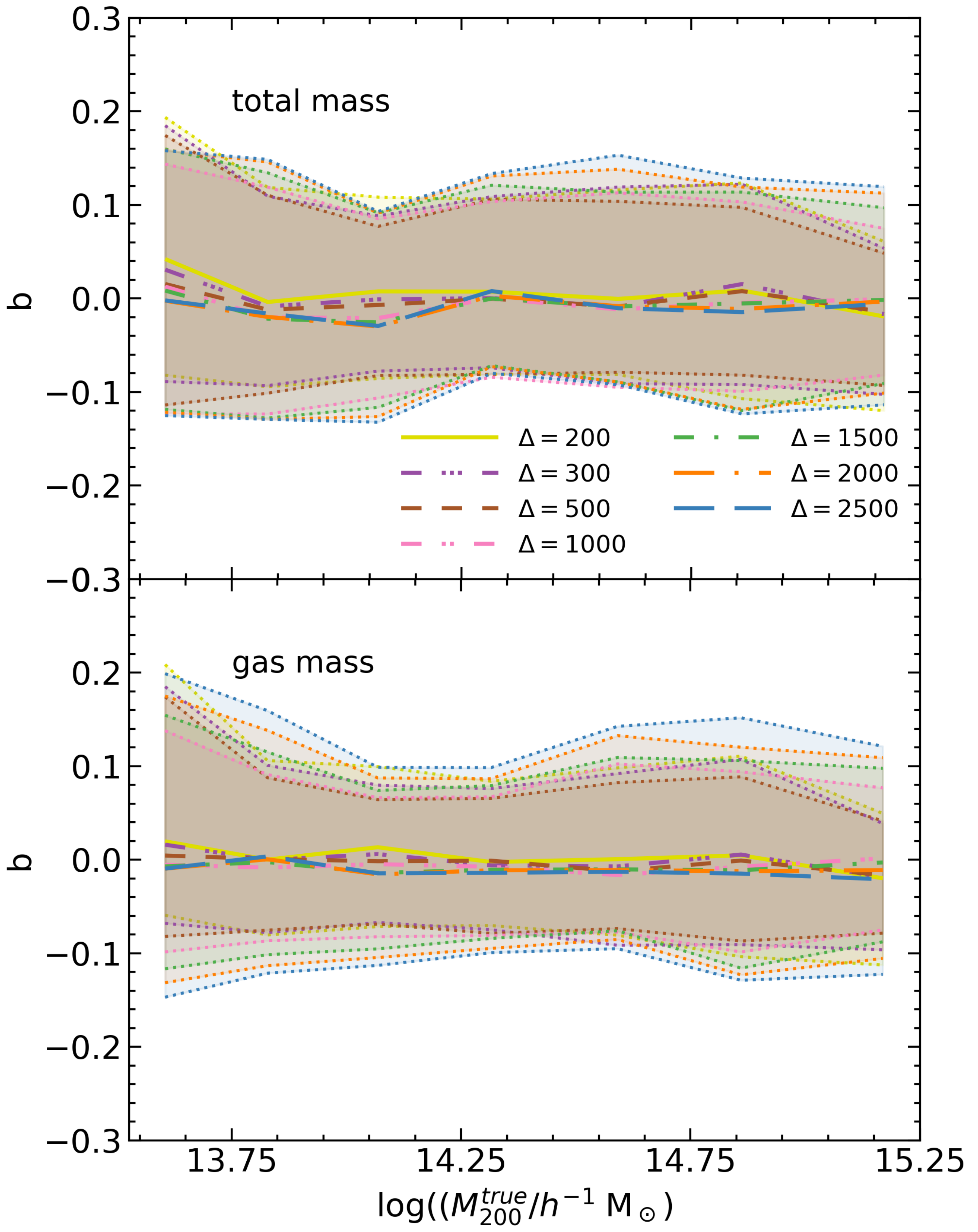


The training of the network was done dividing the dataset in the standard 80%-20% Train and Test sets

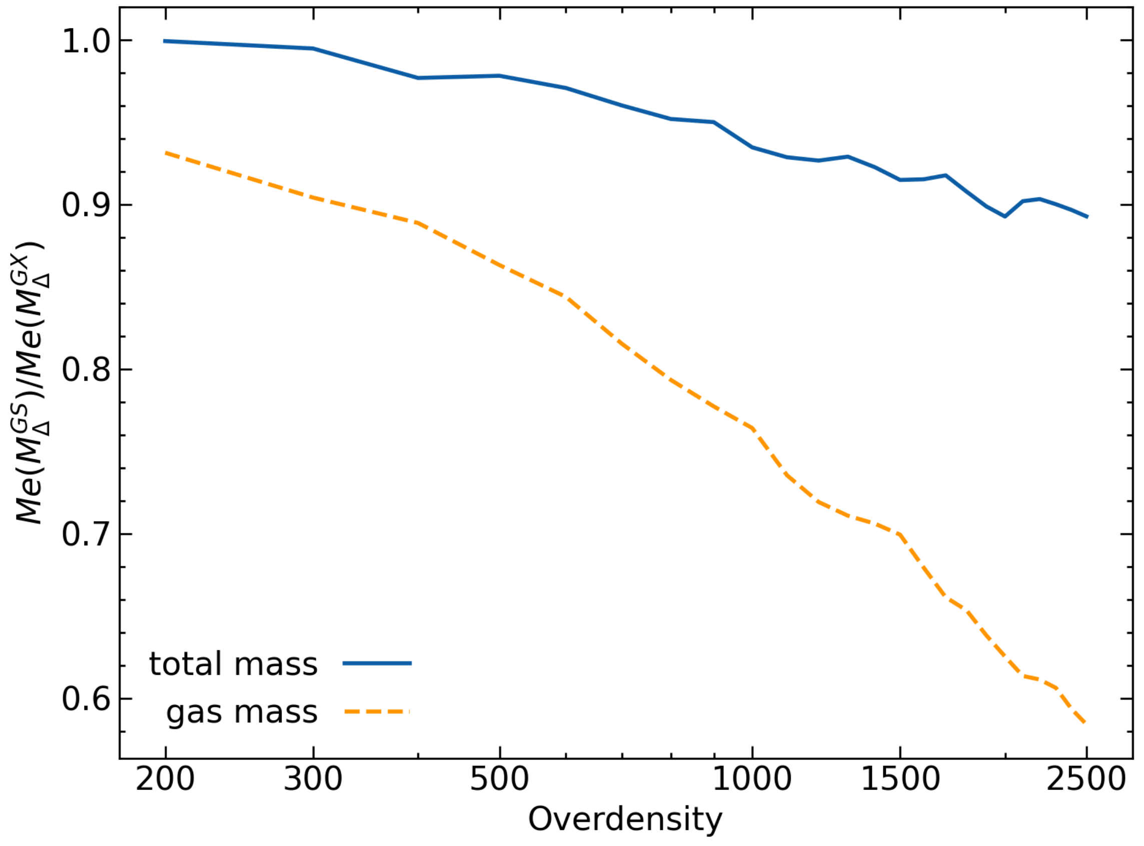
Predicted vs true profiles



- Scatter ~10% along the whole profile
- Unbiased profiles
  - Total mass
  - Gas mass
  - Gas fraction

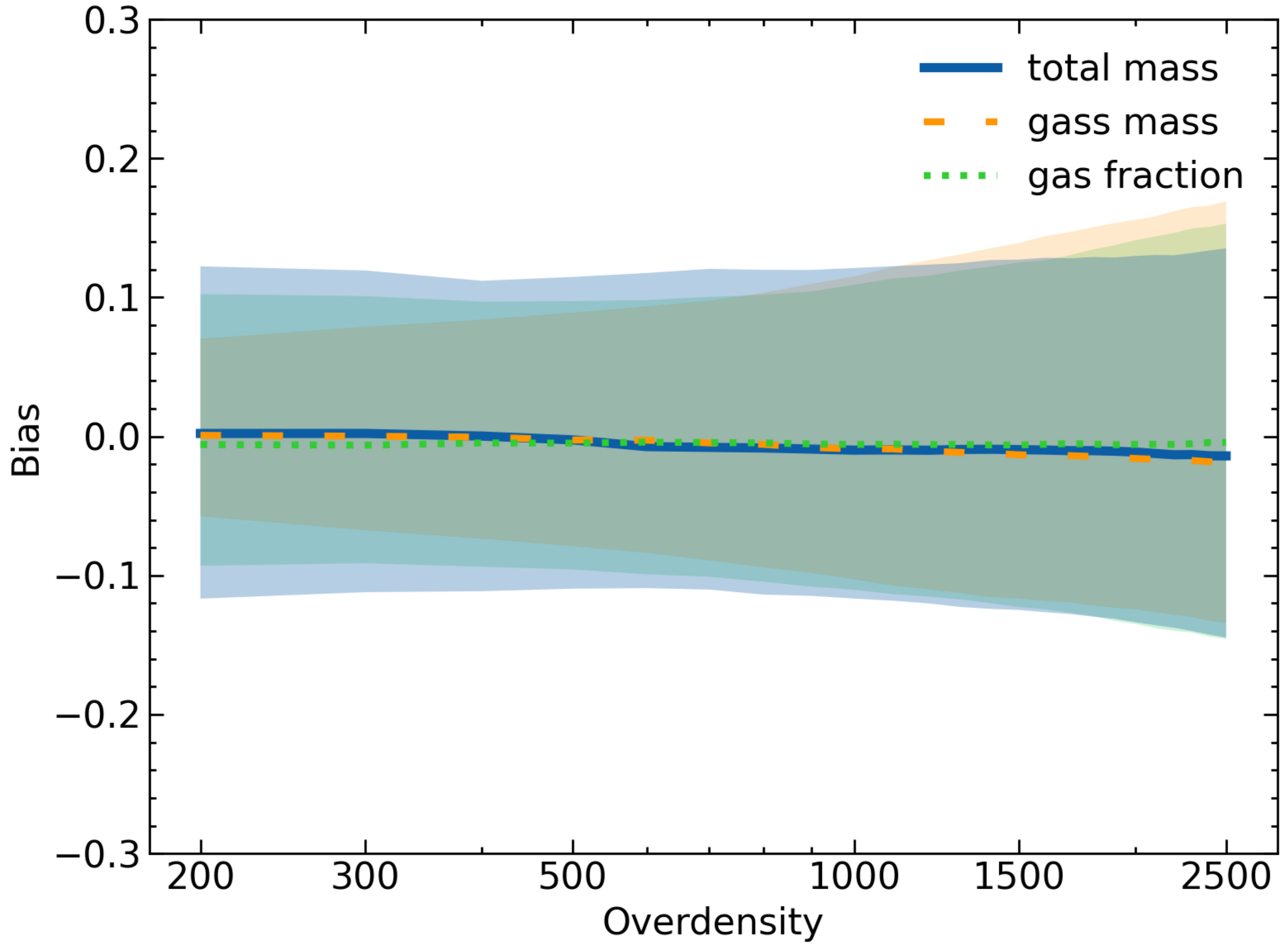


Predicted vs true profiles

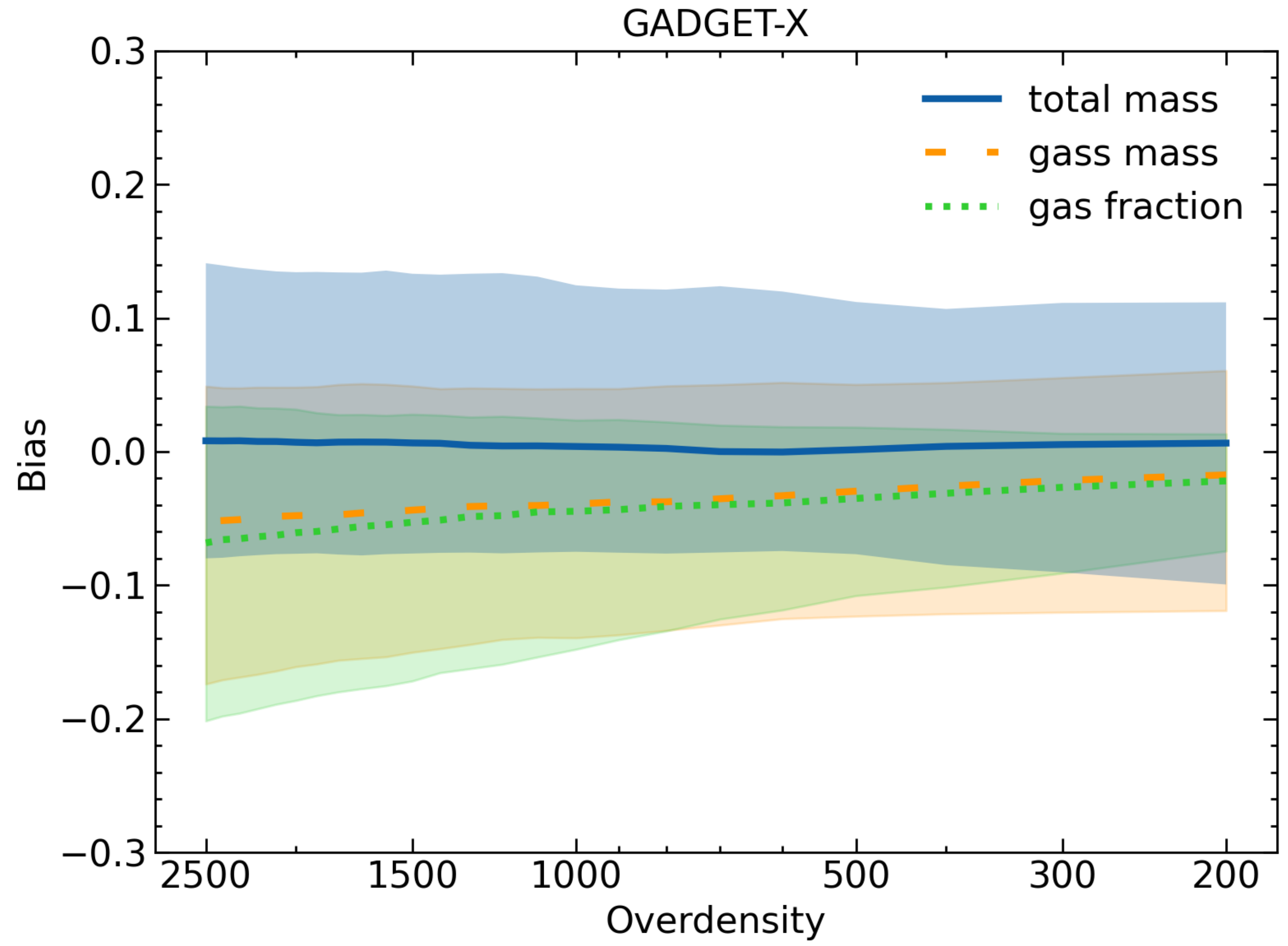


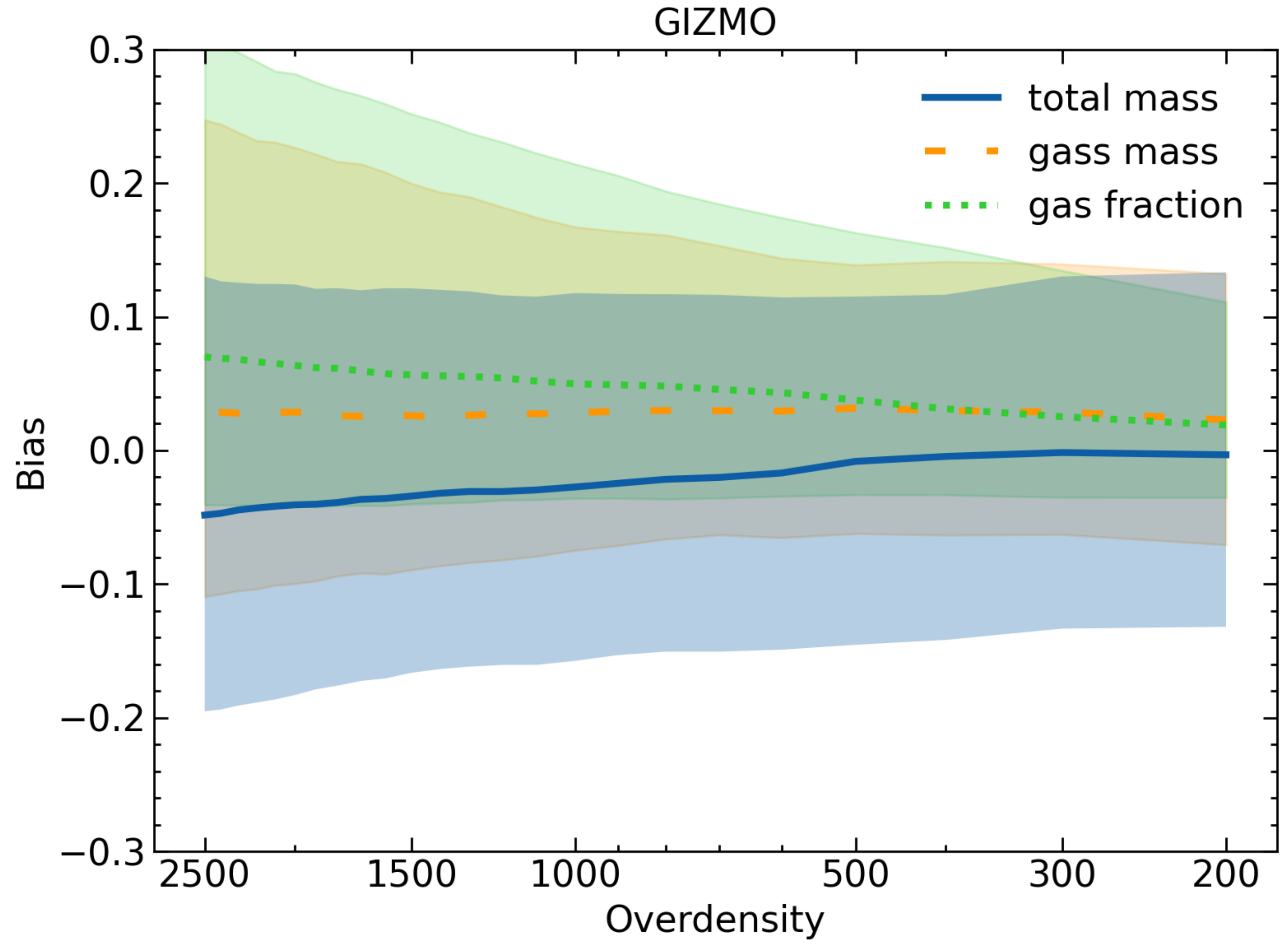
GADGET-X and GIZMO-SIMBA profiles are quite different especially in the cluster core (Talk by W. Cui)

TRAINING:GADGET-X + GIZMO-SIMBA



- Unbiased mass reconstruction by testing on a mixture of GADGET-X and GIZMO_SIMABA clusters
- Bias < 5% when tested on single Gadget-X or GIZMO-SIMBA samples

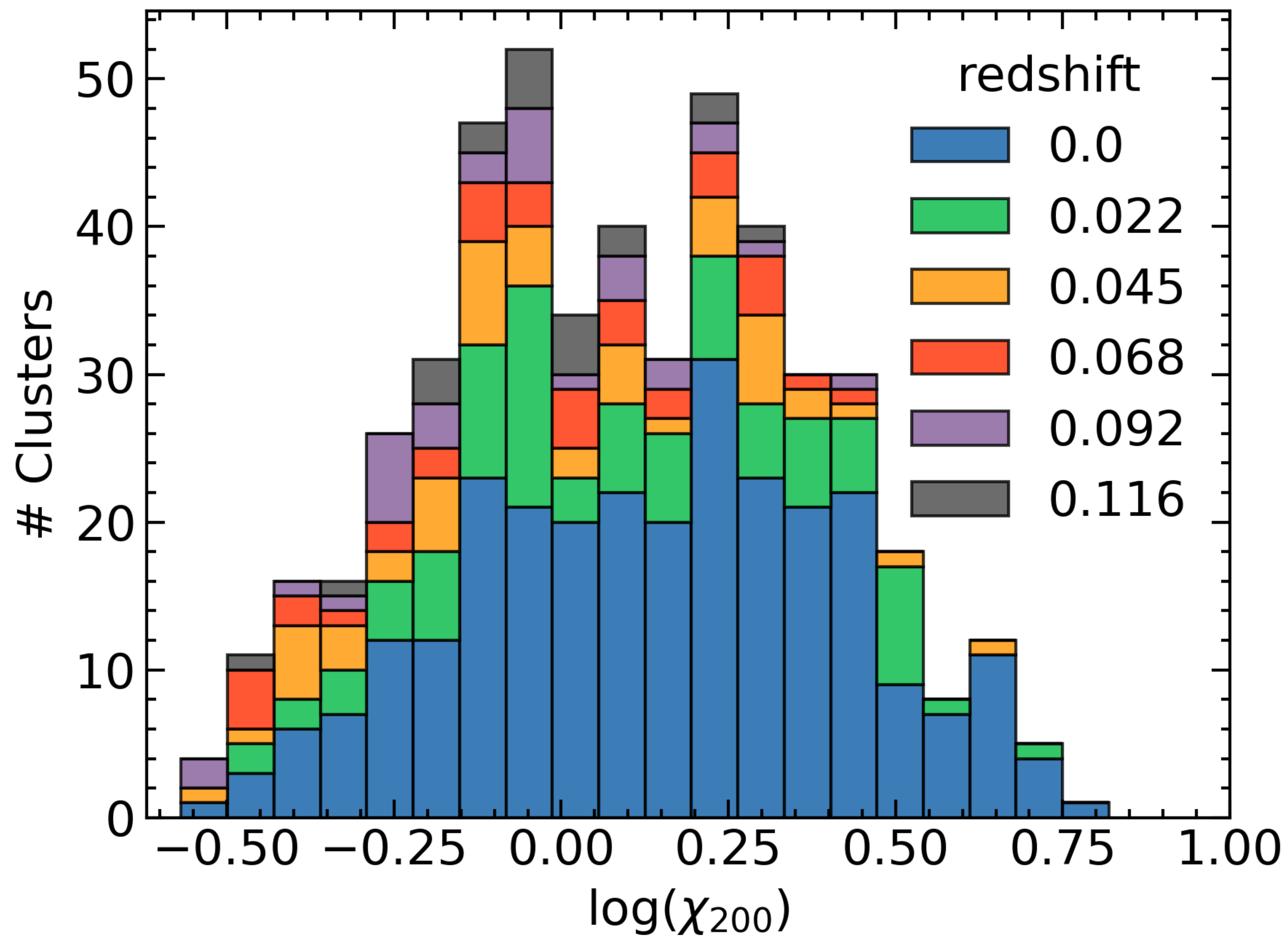


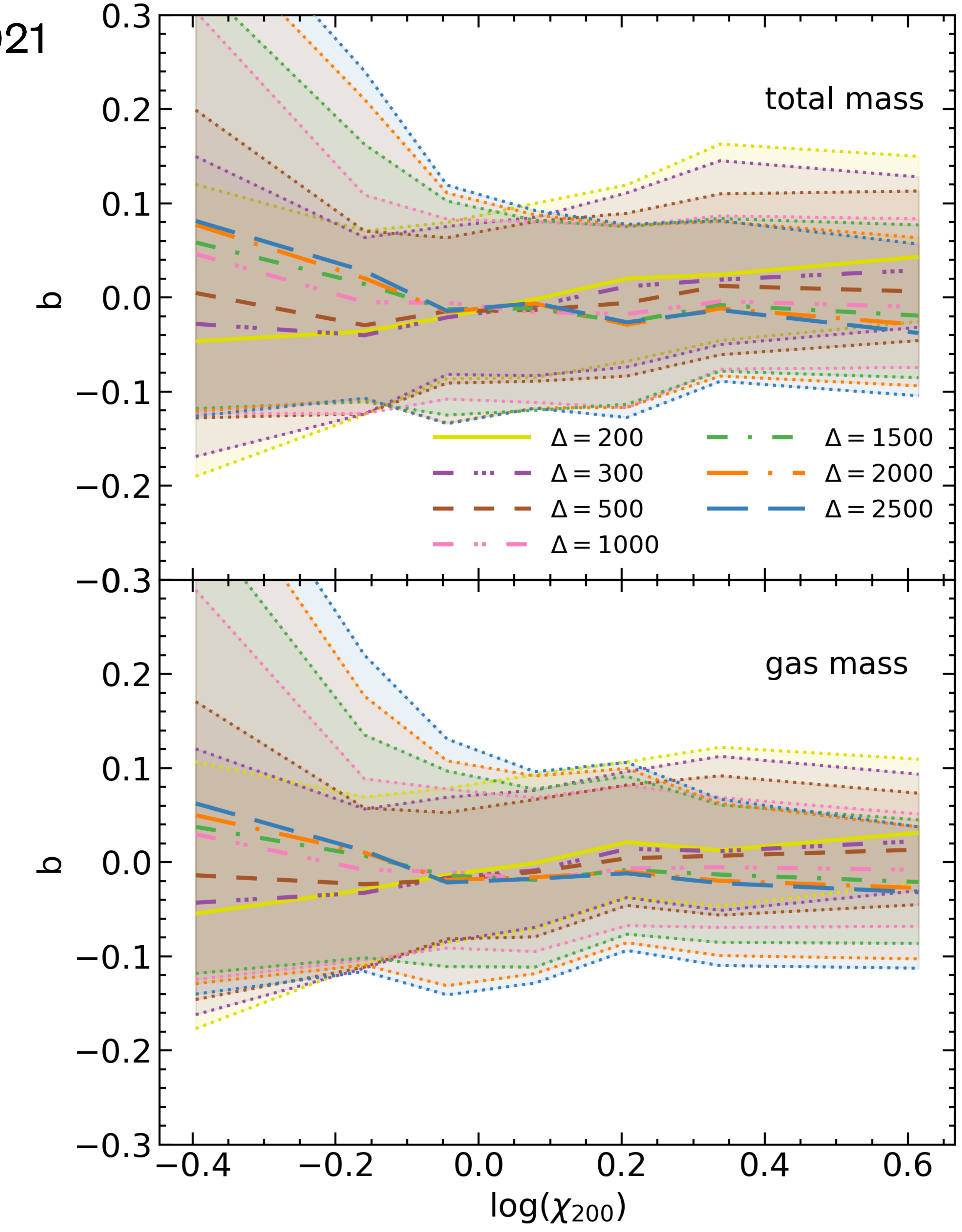


Predicted vs true profiles: dynamical state dependence

The dynamical state is defined using the χ parameter as defined in De Luca el al. 2021

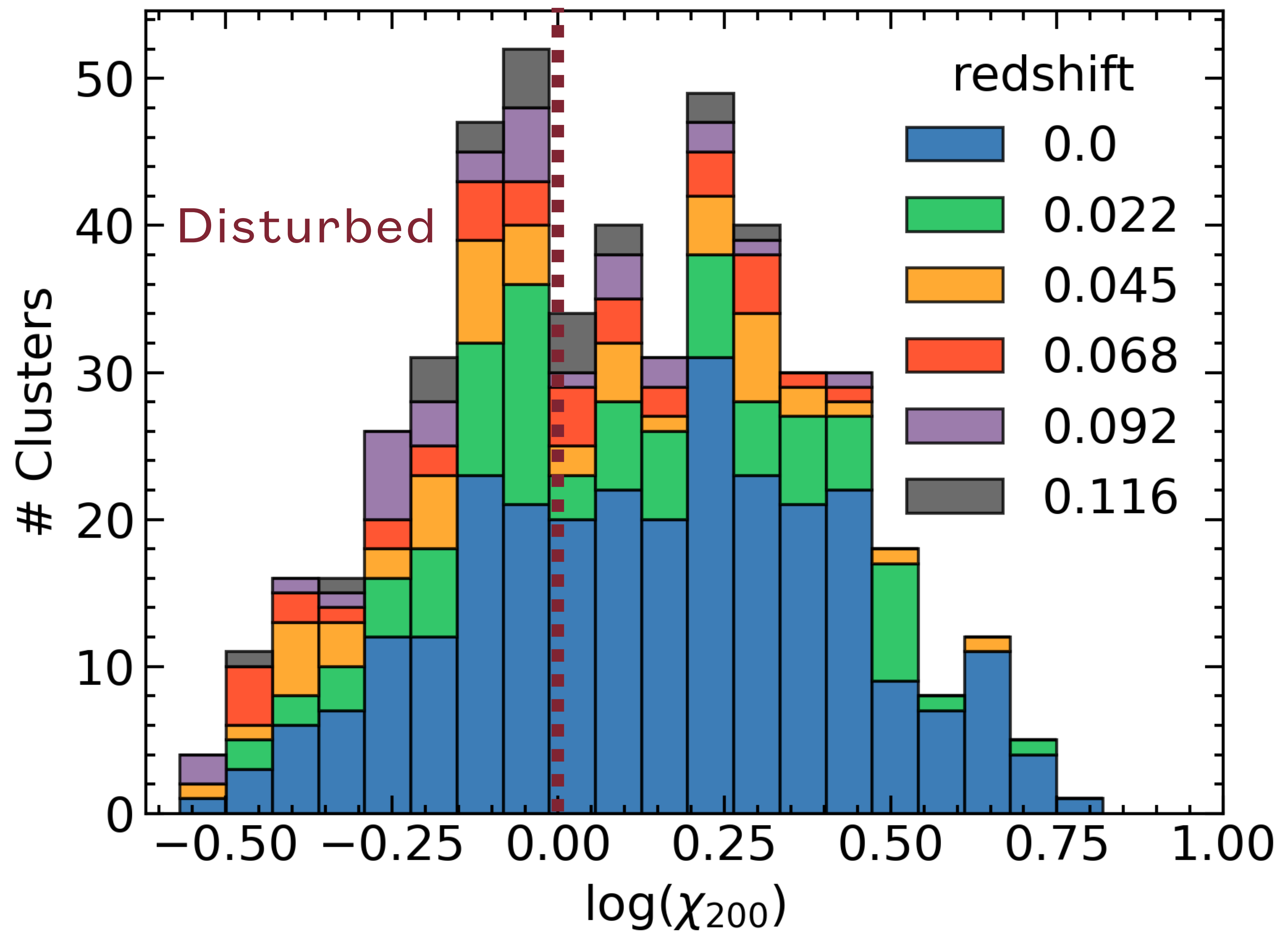
$$\chi_{200} = \sqrt{\frac{2}{\left(\frac{f_s}{0.1}\right)^2 + \left(\frac{\Delta_r}{0.1}\right)^2}}$$





Predicted vs true profiles: dynamical state dependence

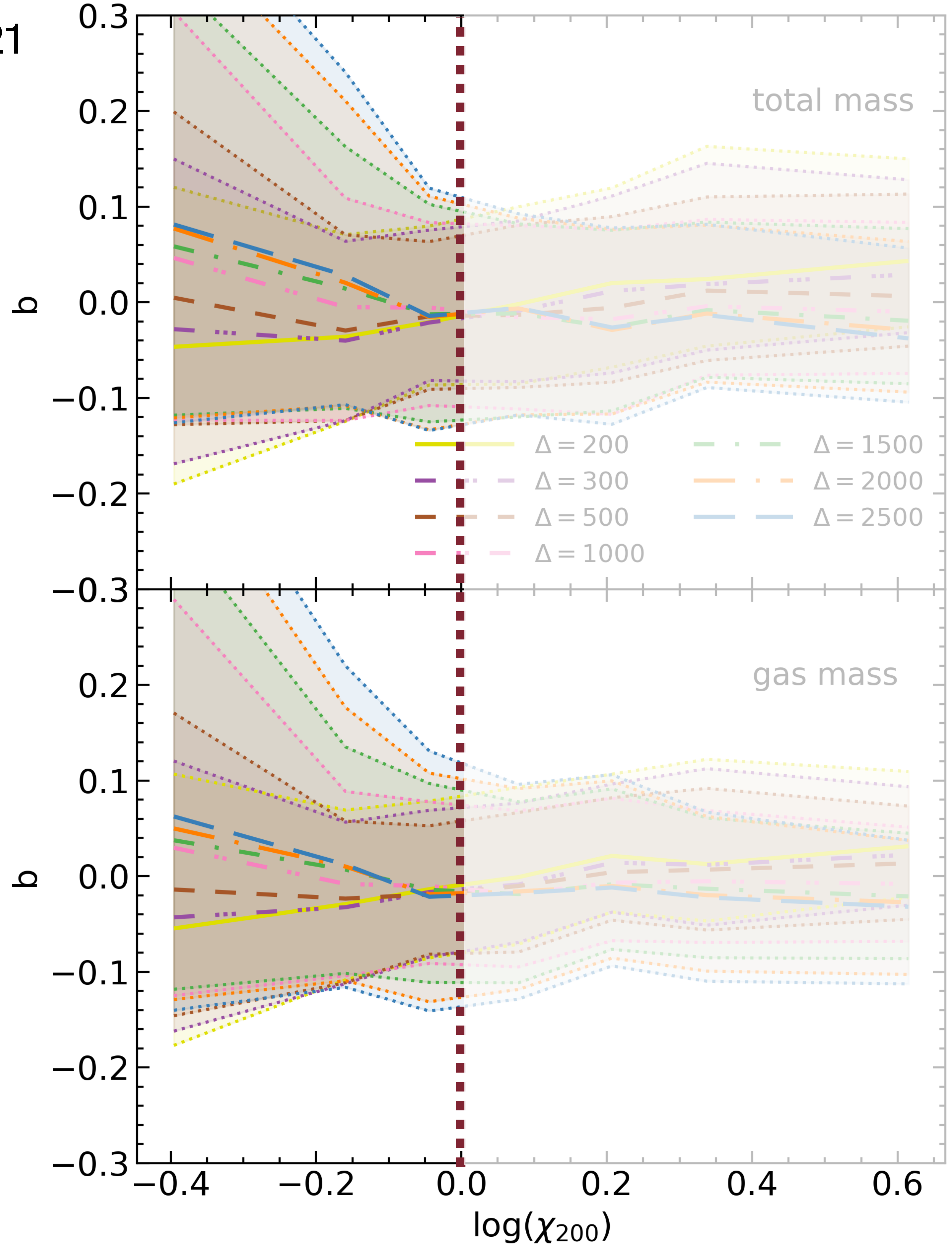
The dynamical state is defined using the χ parameter as defined in De Luca el al. 2021



$$\chi_{200} = \sqrt{\frac{2}{\left(\frac{f_s}{0.1}\right)^2 + \left(\frac{\Delta_r}{0.1}\right)^2}}$$

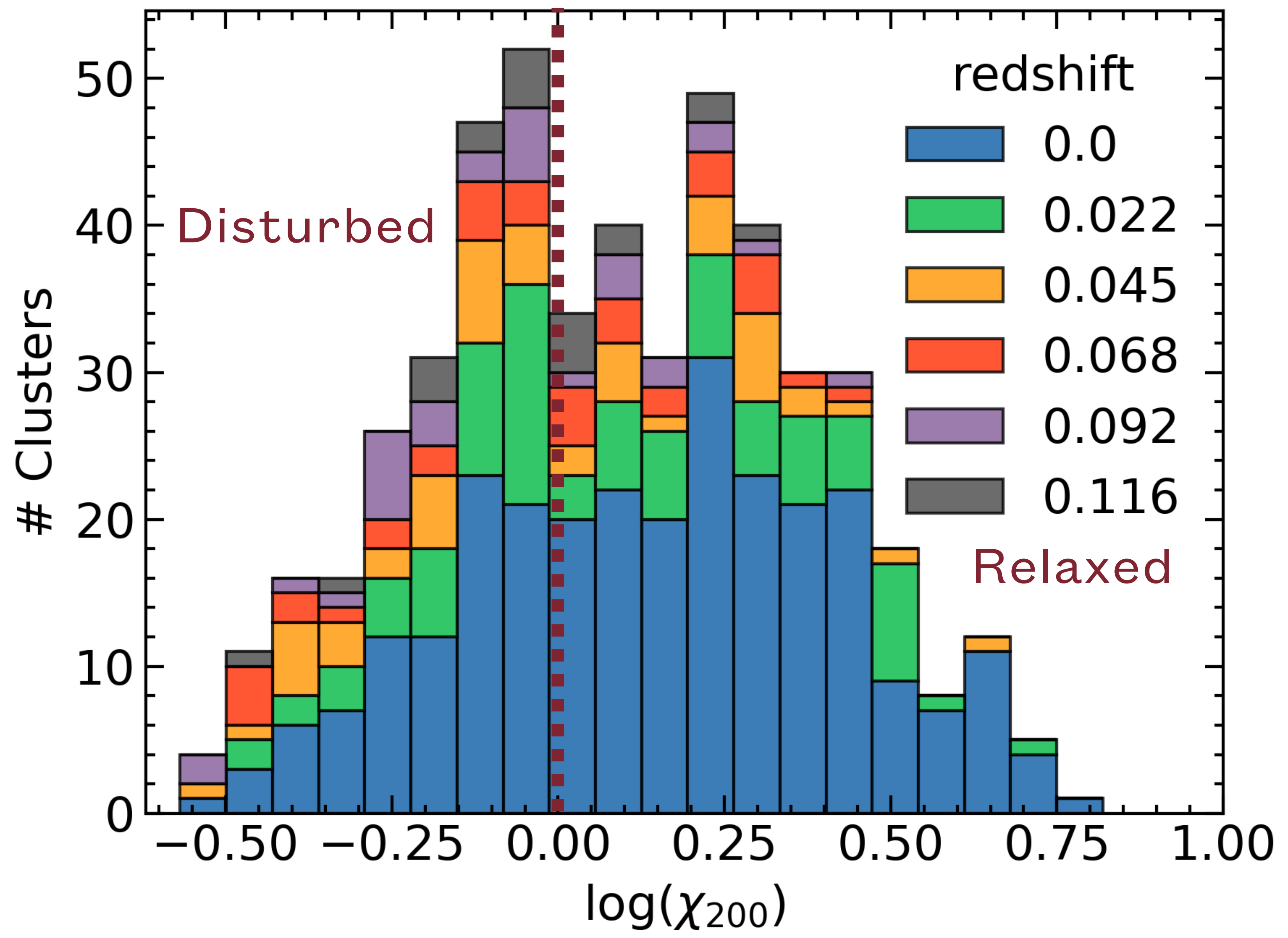
DISTURBED

- Underestimated mass in the outskirts
- Overestimated mass in the core



Predicted vs true profiles: dynamical state dependence

The dynamical state is defined using the χ parameter as defined in De Luca el al. 2021



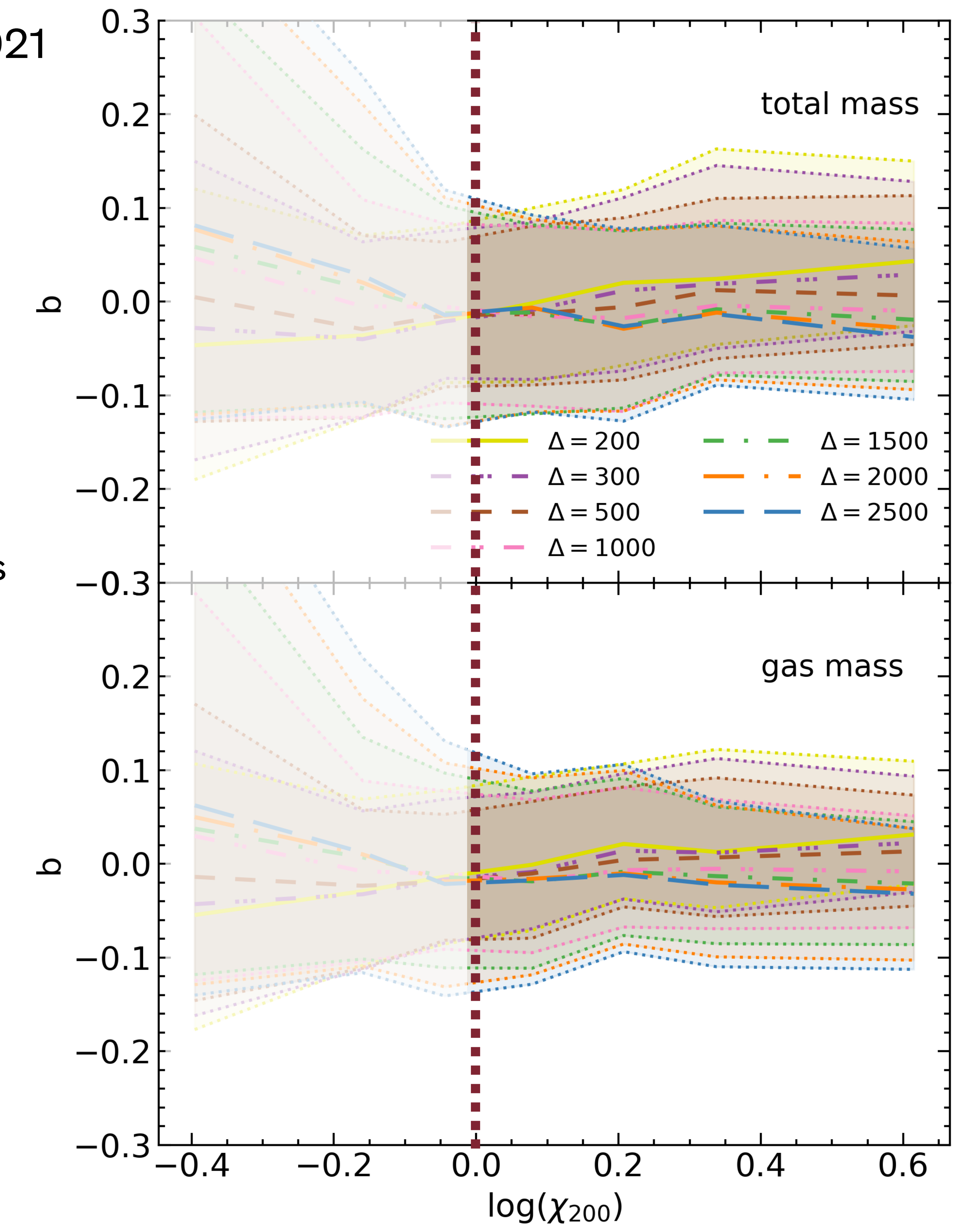
$$\chi_{200} = \sqrt{\frac{2}{\left(\frac{f_s}{0.1}\right)^2 + \left(\frac{\Delta_r}{0.1}\right)^2}}$$

DISTURBED

- Underestimated mass in the outskirts
- Overestimated mass in the core

RELAXED

- Overestimated mass in the outskirts
- Underestimated mass in the core

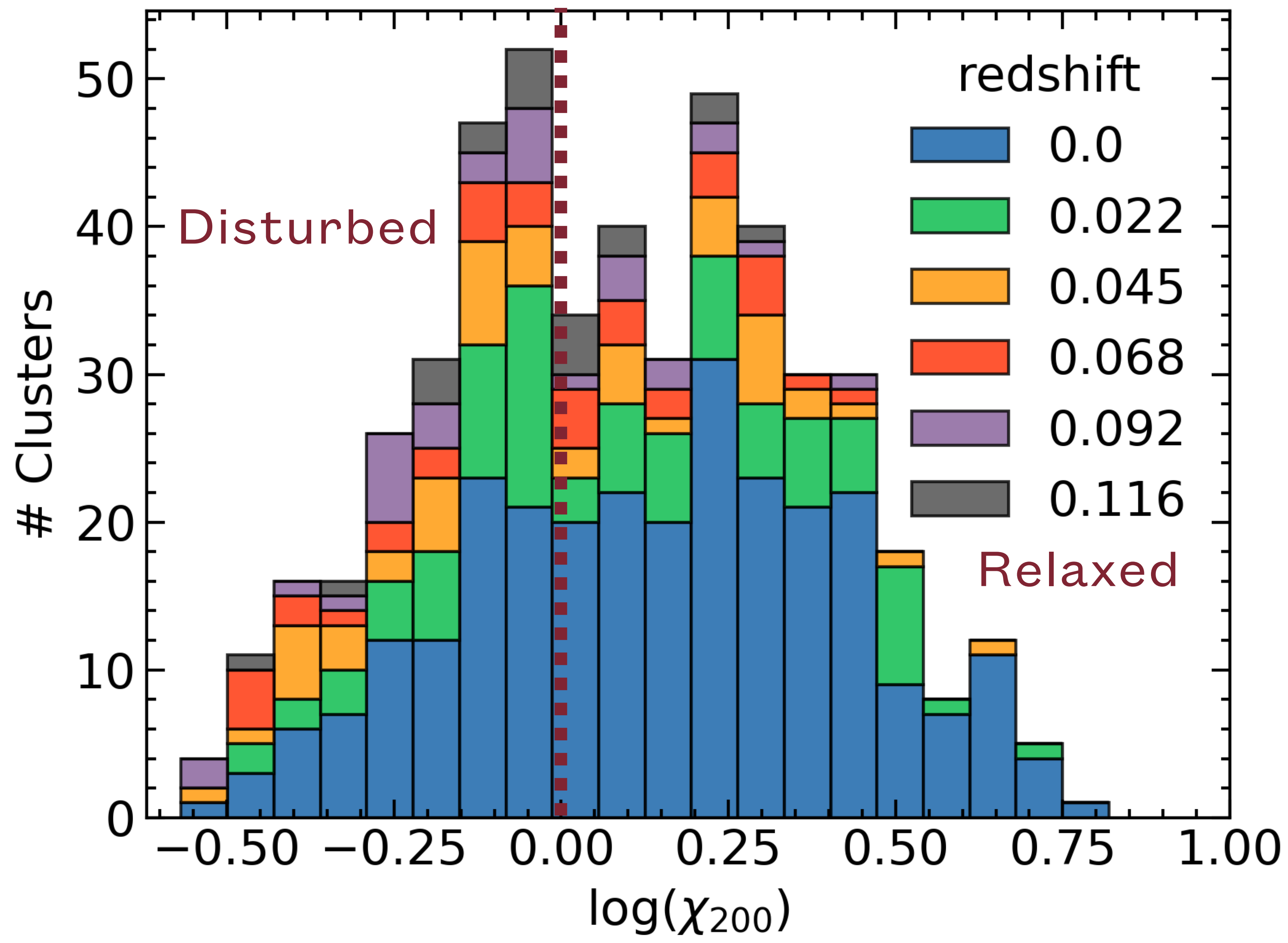


No dependence with the dynamical state at Δ=500

Predicted vs true profiles: dynamical state dependence

The dynamical state is defined using the χ parameter as defined in De Luca el al. 2021

$$\chi_{200} = \sqrt{\frac{2}{\left(\frac{f_s}{0.1}\right)^2 + \left(\frac{\Delta_r}{0.1}\right)^2}}$$

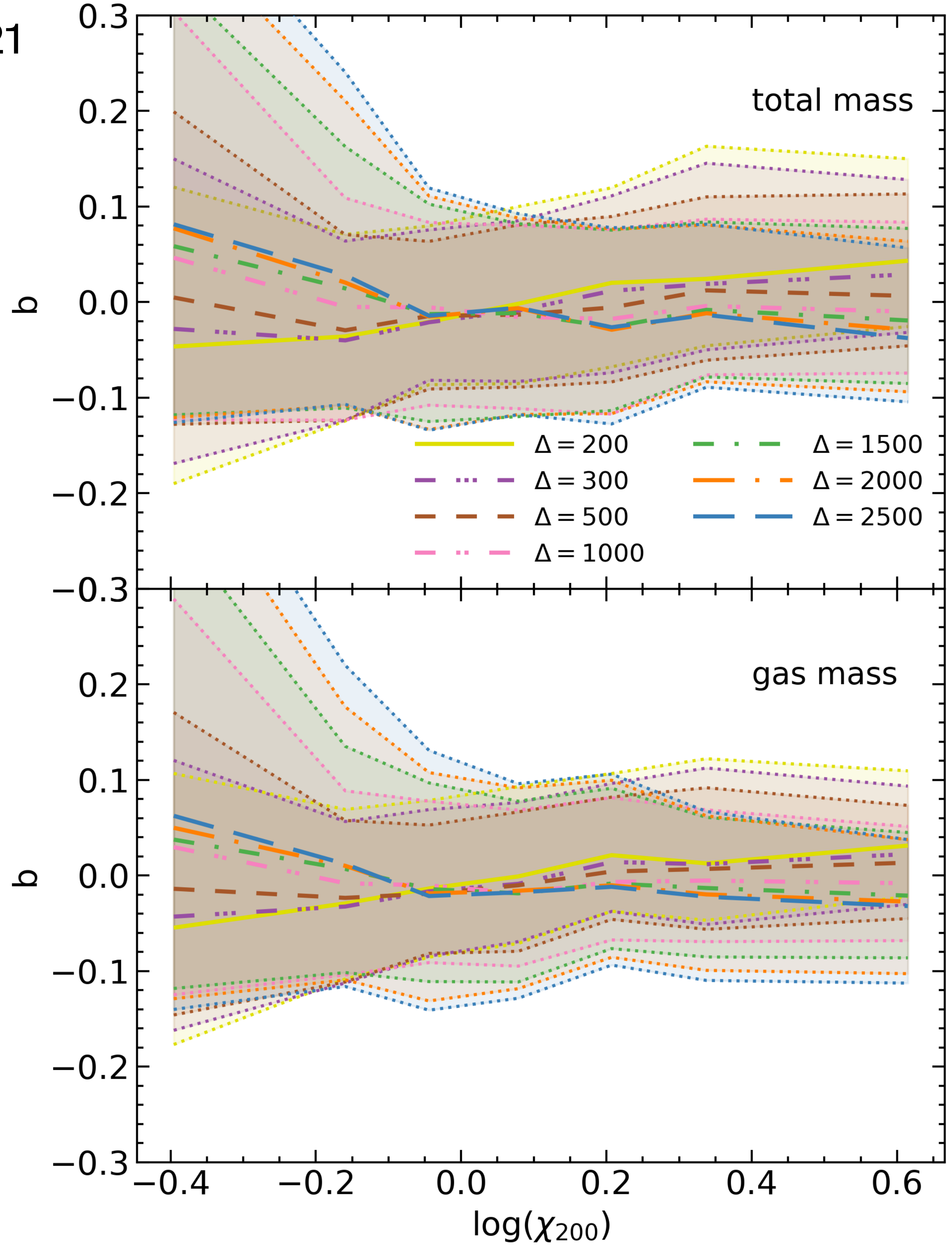


DISTURBED

- Underestimated mass in the outskirts
- Overestimated mass in the core

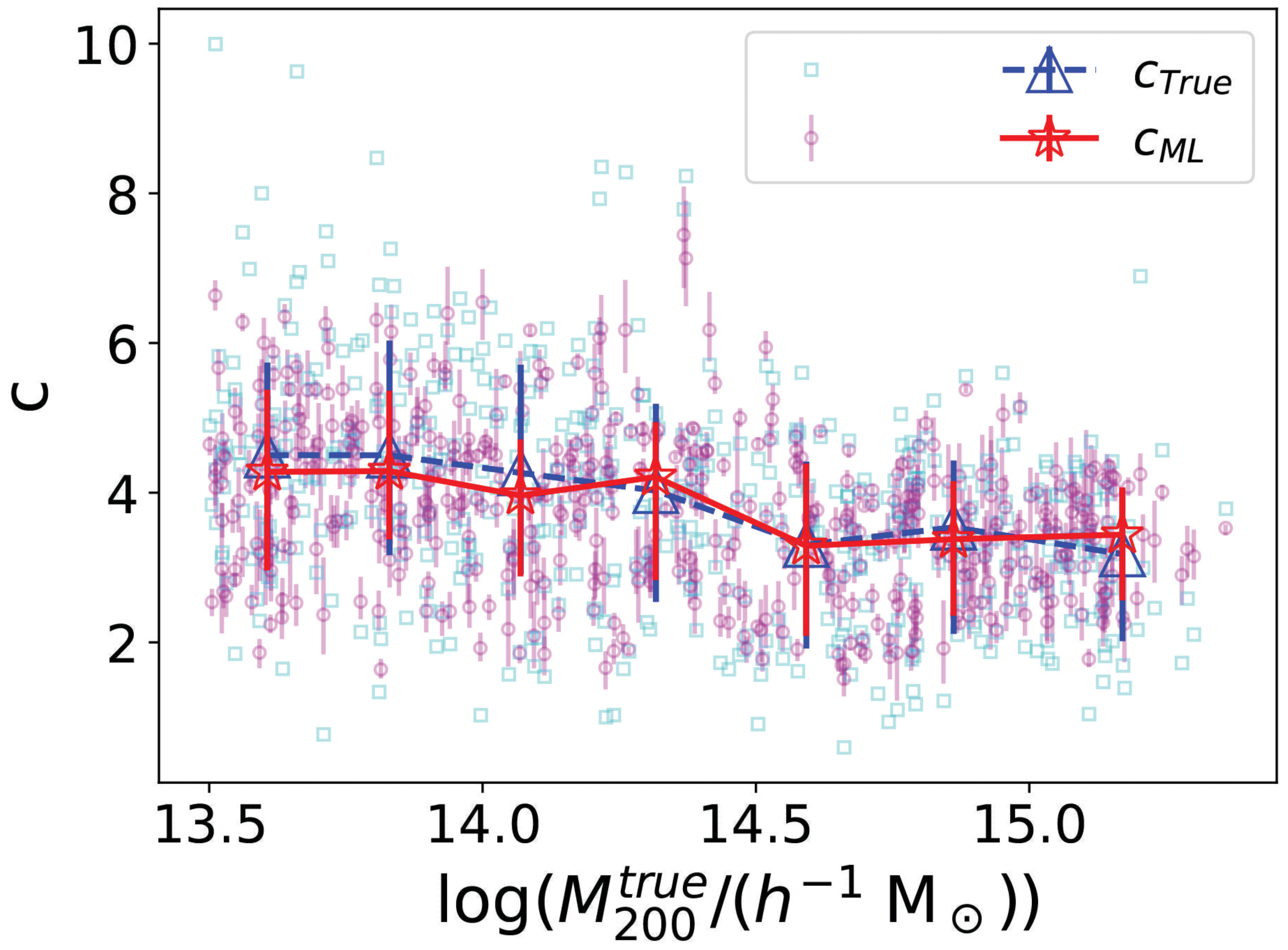
RELAXED

- Overestimated mass in the outskirts
- Underestimated mass in the core



No dependence with the dynamical state at Δ=500

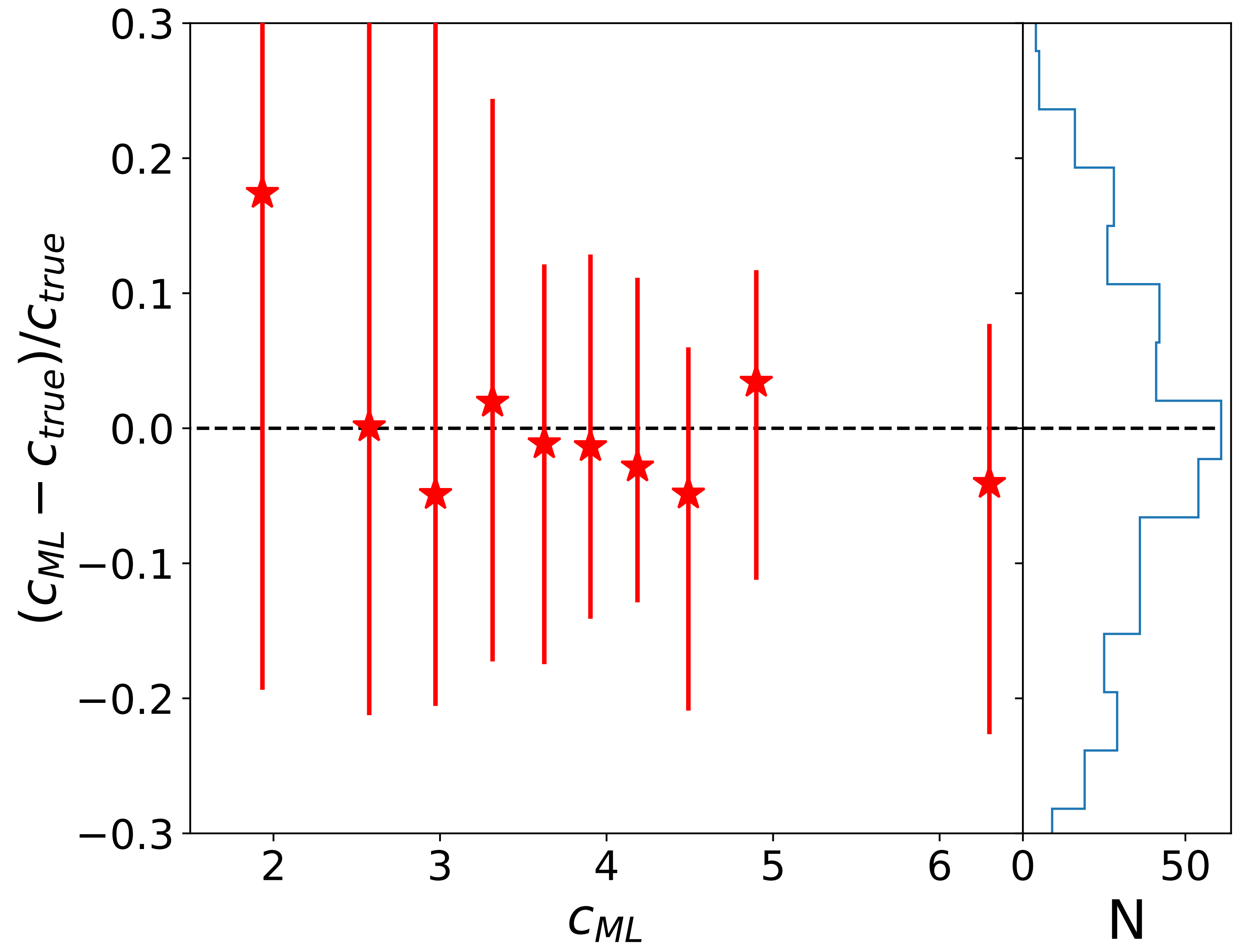
Concentration



NFW profile

$$\rho(r) = \frac{\rho_0}{\frac{r}{r_s}\left(1 + \frac{r}{r_s}\right)}$$

$$c_{ML} = R_{200}/r_s$$



- Median predicted concentration is indistinguishable from the true one
- ML concentration as less scatter than the true one (less projection effects)

- Bias between c_true and c_ML compatible with zero for intermediate and high concentrations
- Hint of underestimation of c for high concentration systems
- ML overestimate the concentration (the mass) of low c systems (Mayor mergers, disturbed clusters)

ML vs Hydrostatic Equilibrium

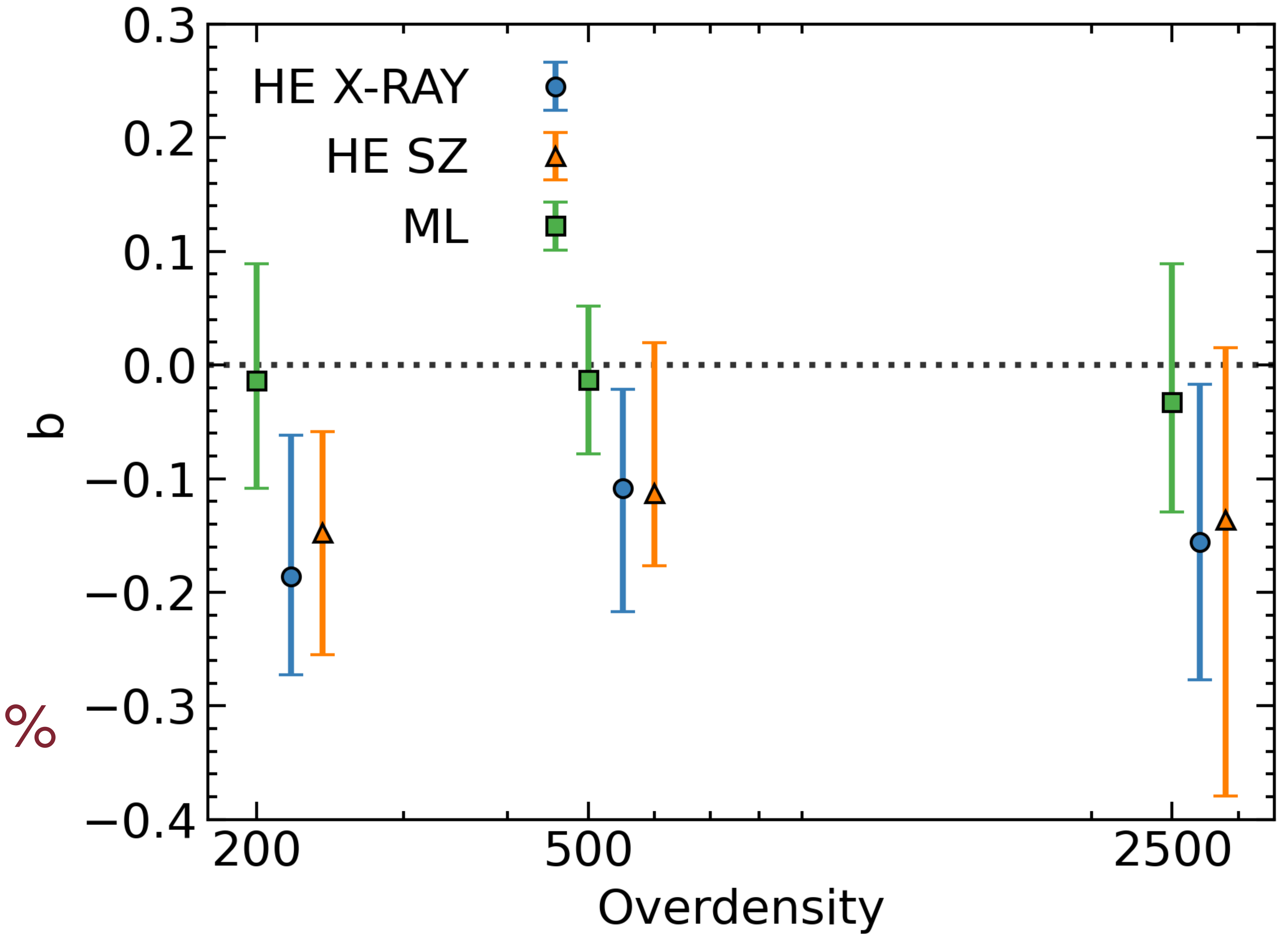
Comparison with mass estimated with HE model (*Gianfagna et al.2021)* on a subsample of the 52 most massive clusters

HE method

- well known hydrostatic bias between 10% and 20%
- Scatter around 20%

ML method

- Unbiased
- Scatter around 10%



Conclusions

Machine learning models applied for the first time to infer mass radial profiles of galaxy clusters from Sunyaev Zel'dovich mock images extracted from The Three Hundred hydrodynamical simulation.

- Unbiased reconstruction of total and gas mass radial profiles with a scatter of ~10%
- Results of the method do not depend on clusters mass
- Dependence by the dynamical state as a function of the overdensity.
- the scatter increases in unrelaxed clusters due to projection effects.
- Unbiased Gas fraction inferred with a scatter between 5% and 10%
- The concentration parameter is unbiased with a scatter between 10 and 20 per cent for cML > 2.
- ML-predicted c–M relation is in agreement with the true one.
- Better accuracy and precision than the Hydrostatic equilibrium method

Ongoing work

- ➡ Extend this approach to infer other ICM radial profiles:
  - gas temperature
  - Gas pressure
- ➡ Application on mock SZ observations:
  - noise
  - instrumental effects
  - limited angular resolution.
- ➡ Apply the model on real Compton-y parameter maps