Machine learning to improve galaxy cluster mass estimation

(Jay) Digvijay Wadekar

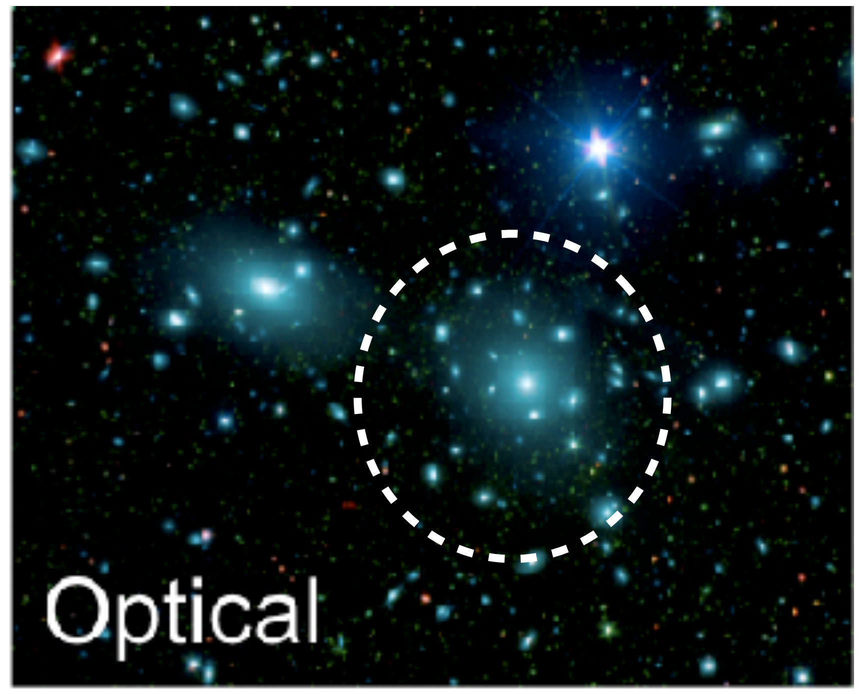
IAS

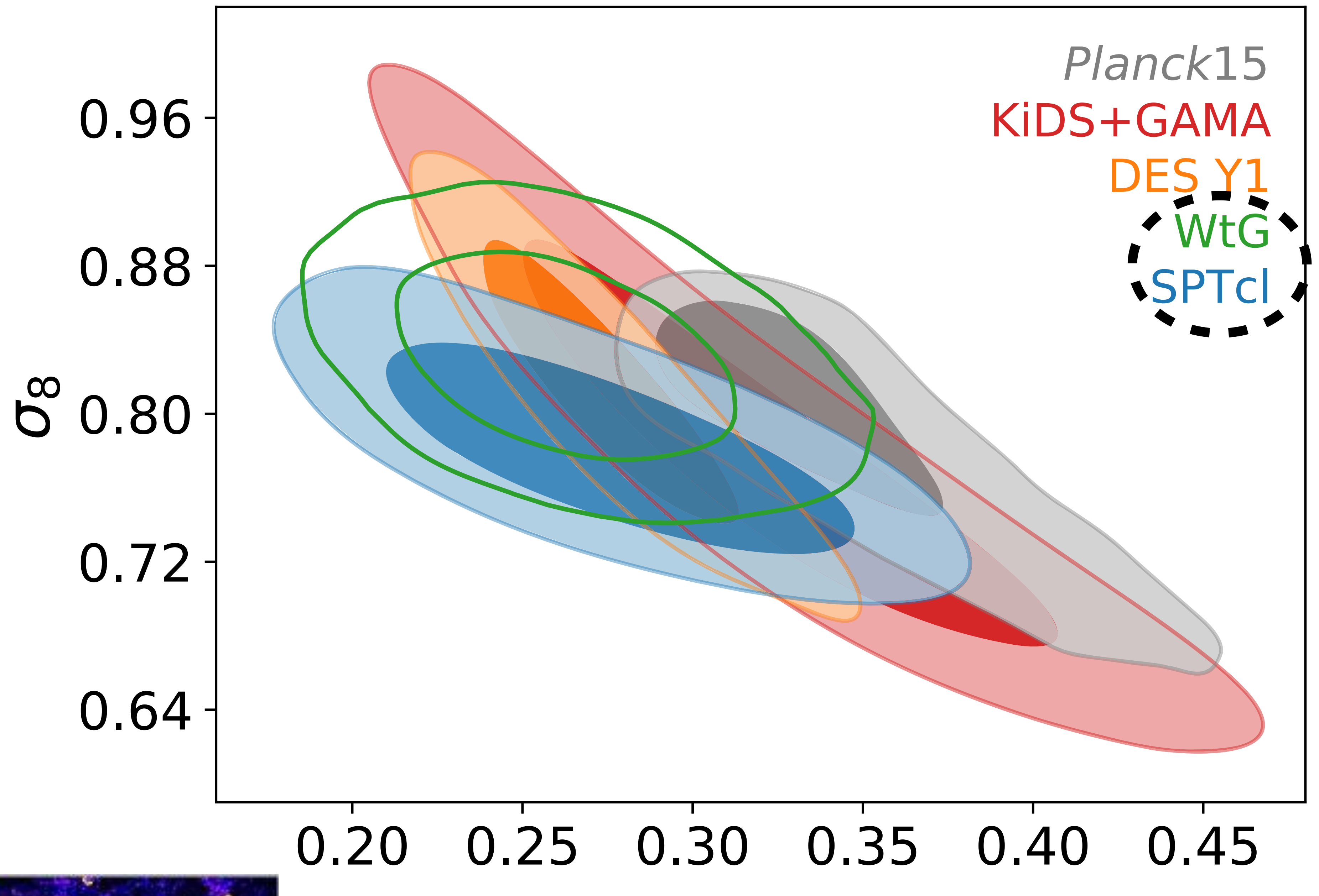
with

L. Thiele, F. Villaescusa-Navarro, C. Hill, D. Spergel, ***M. Cranmer***,
N. Battaglia, S. Ho, D. Angles-Alcazar, L. Hernquist

arXiv:2201.01305 & in prep.

Cluster mass estimation is important for cosmology

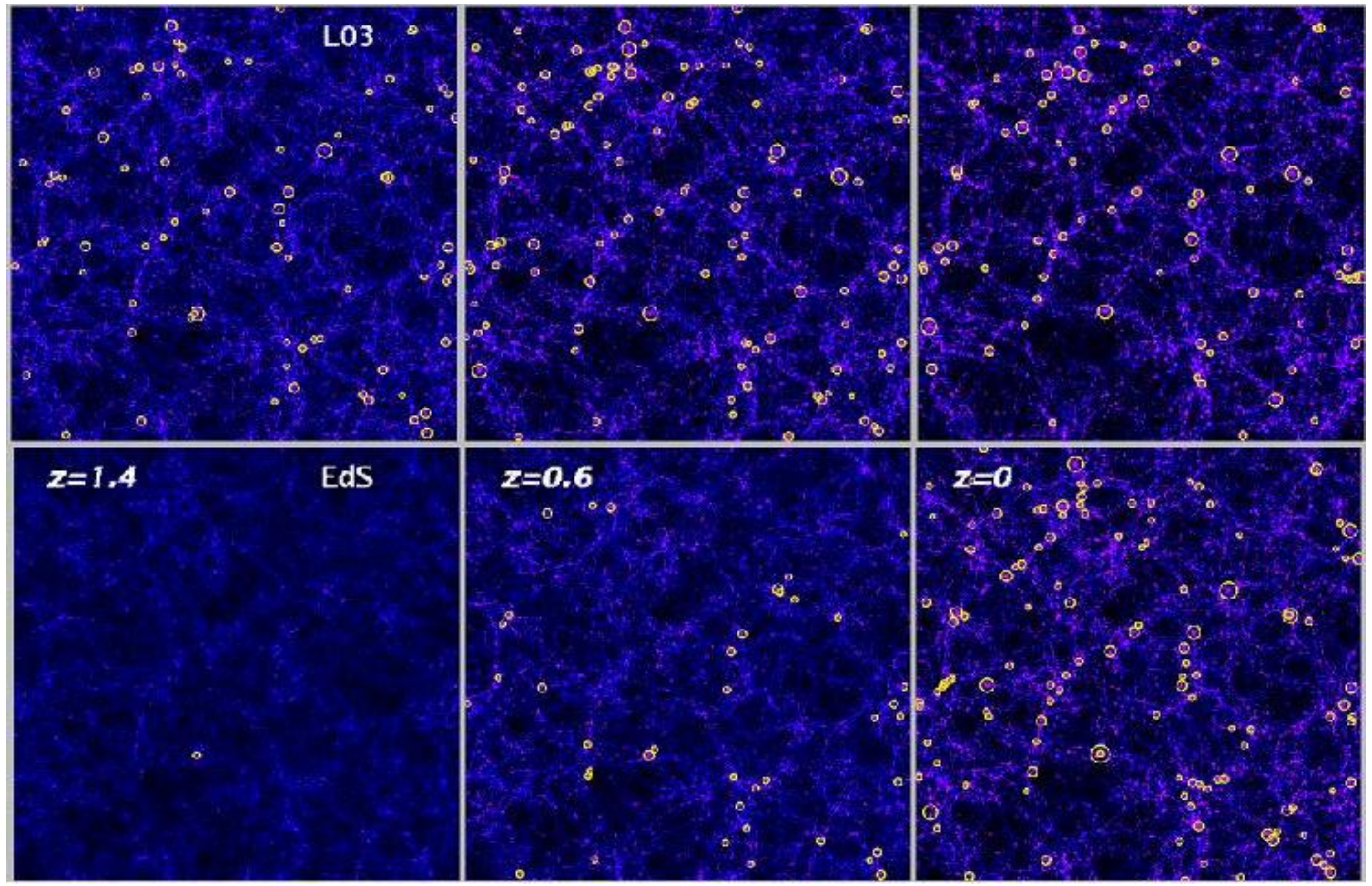




Bocquet et al. 19

ΛCDM
$\Omega_M = 0.3$
$\Lambda = 0.7$

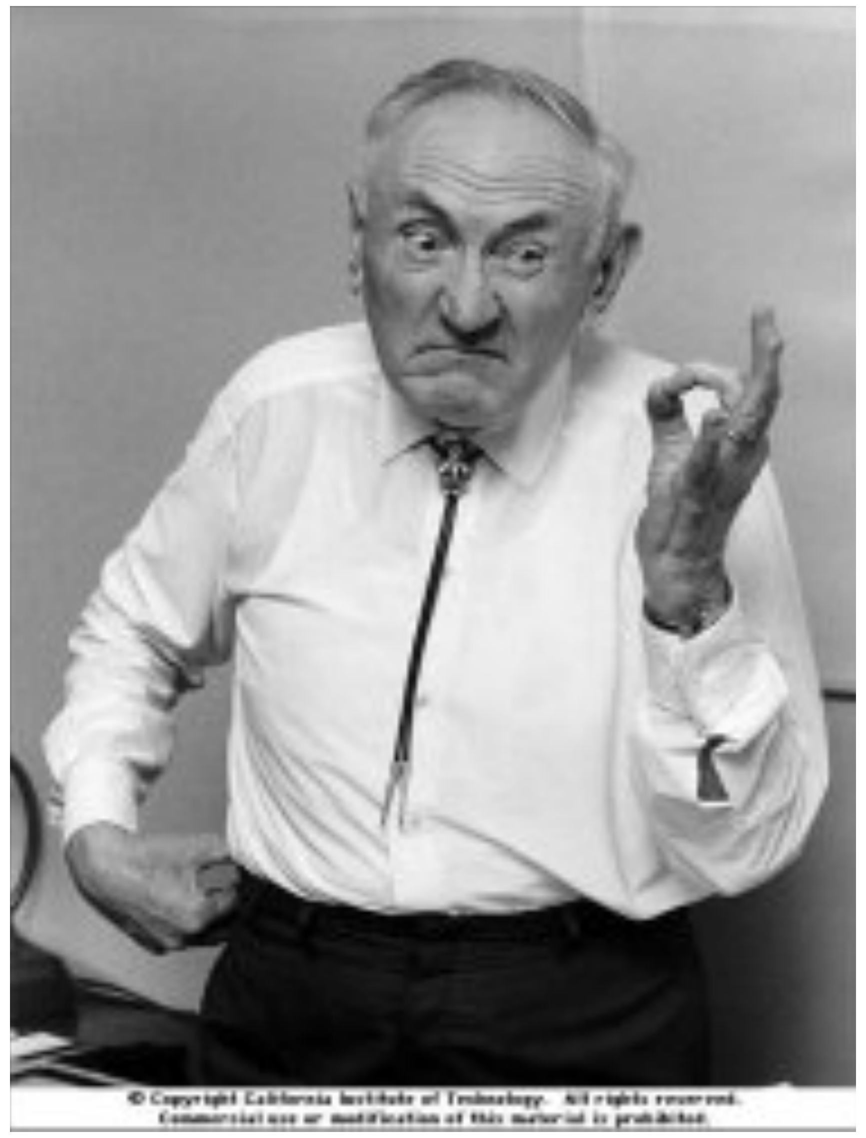
SCDM
$\Omega_M = 1$



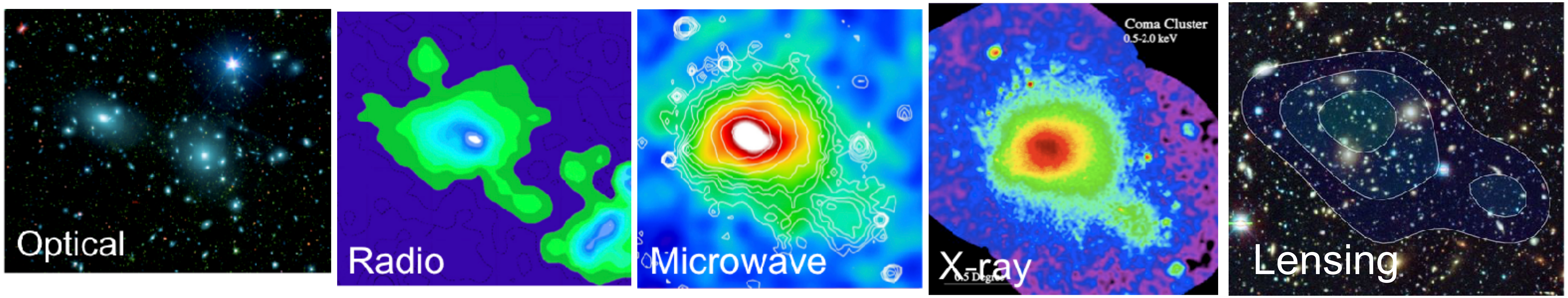
normalized to present density

Borgani & Guzzo 01

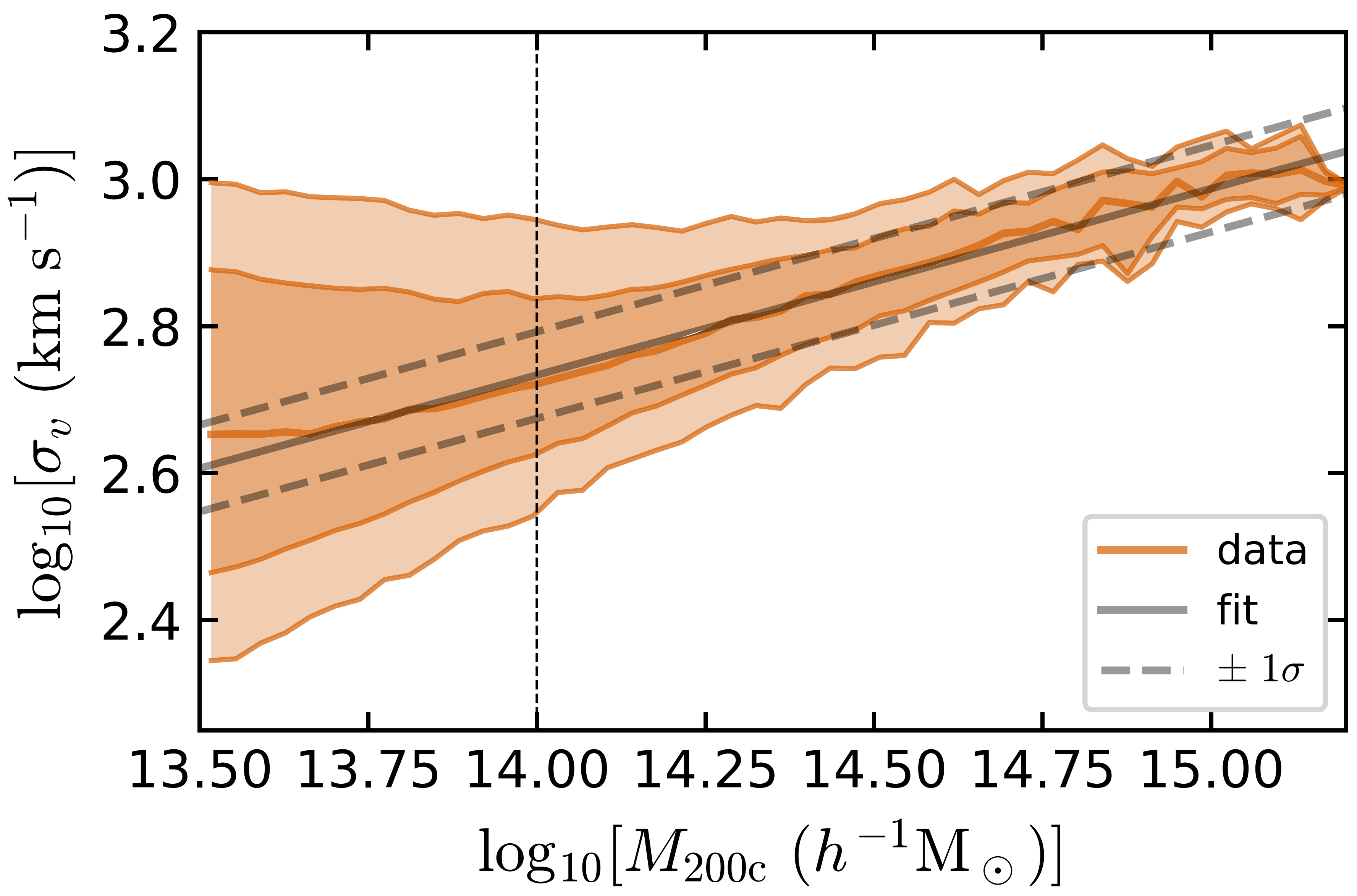
"dunkle materie"

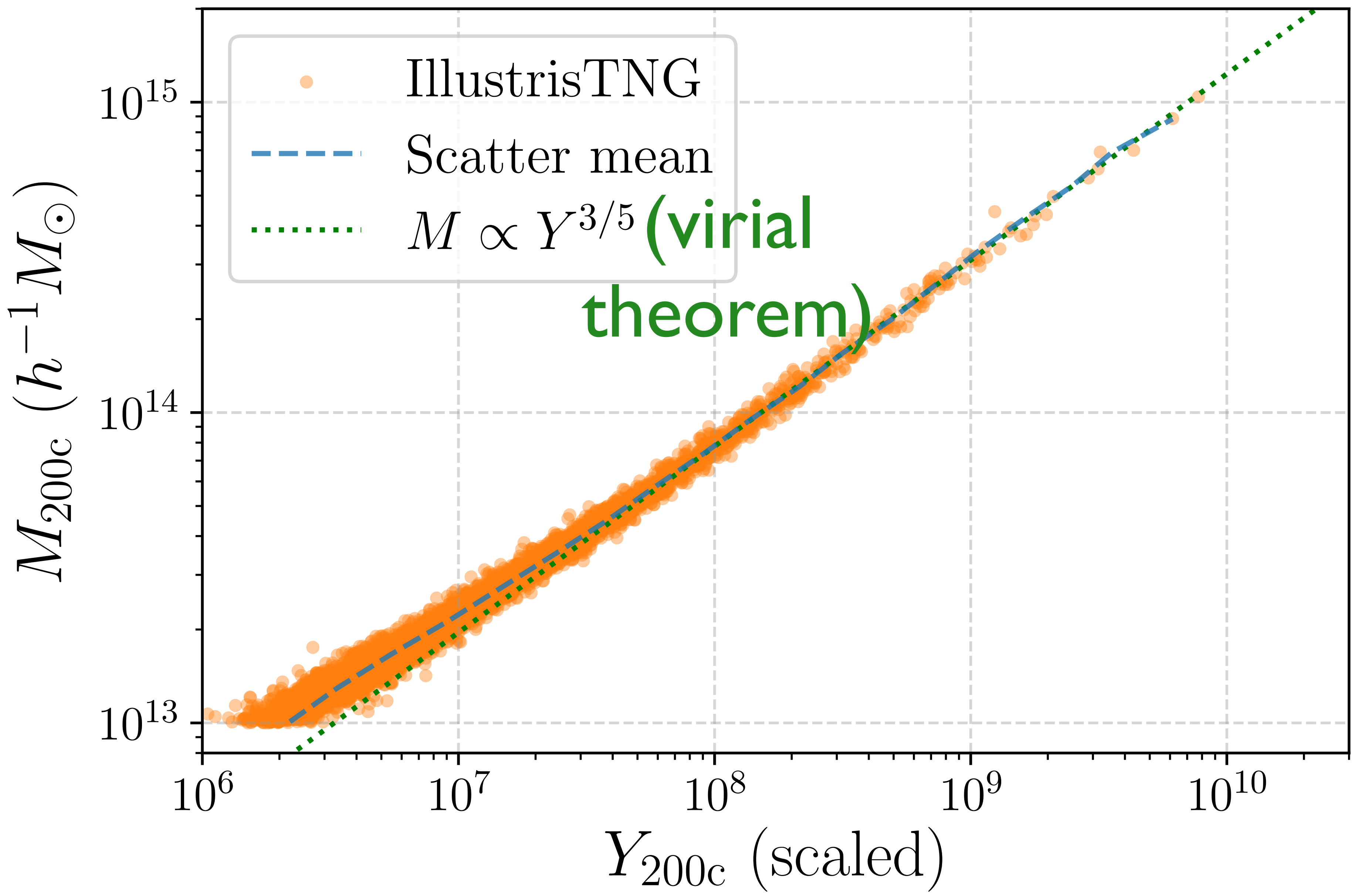


Traditional approaches for cluster mass estimation



(Velocity dispersion)



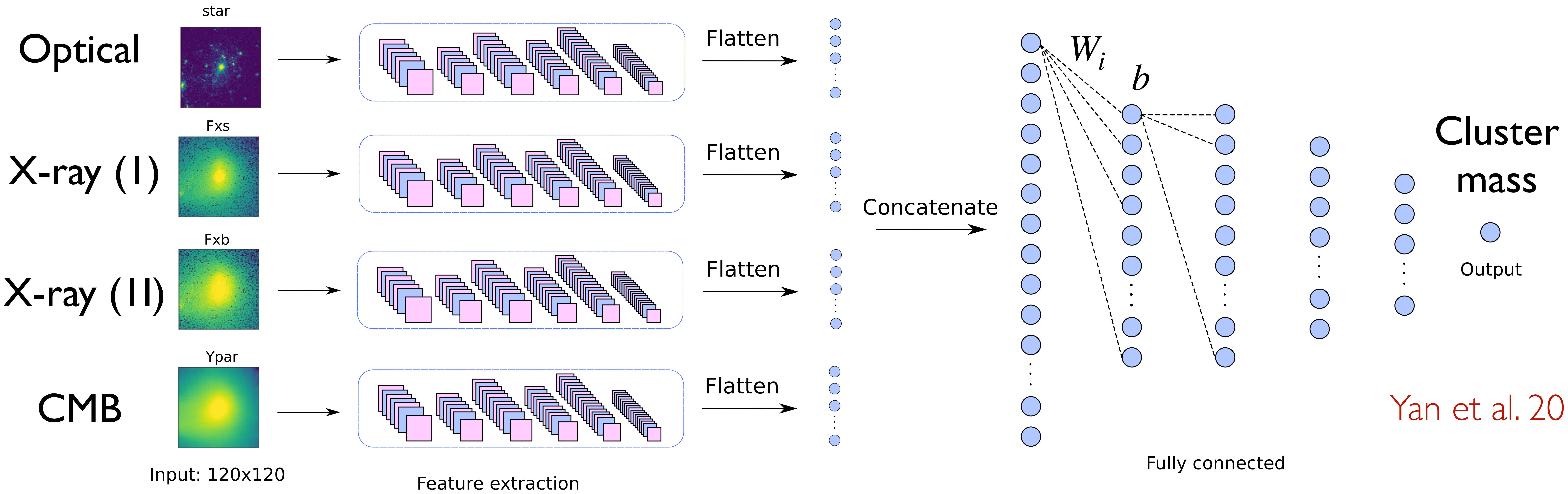


$$Y \propto \int_0^{R_{200c}} P_e(r)dV$$

$$\sim M_{\rm gas}T_{\rm gas}$$

(thermal energy of gas)

Machine learning (ML) is a potential alternative

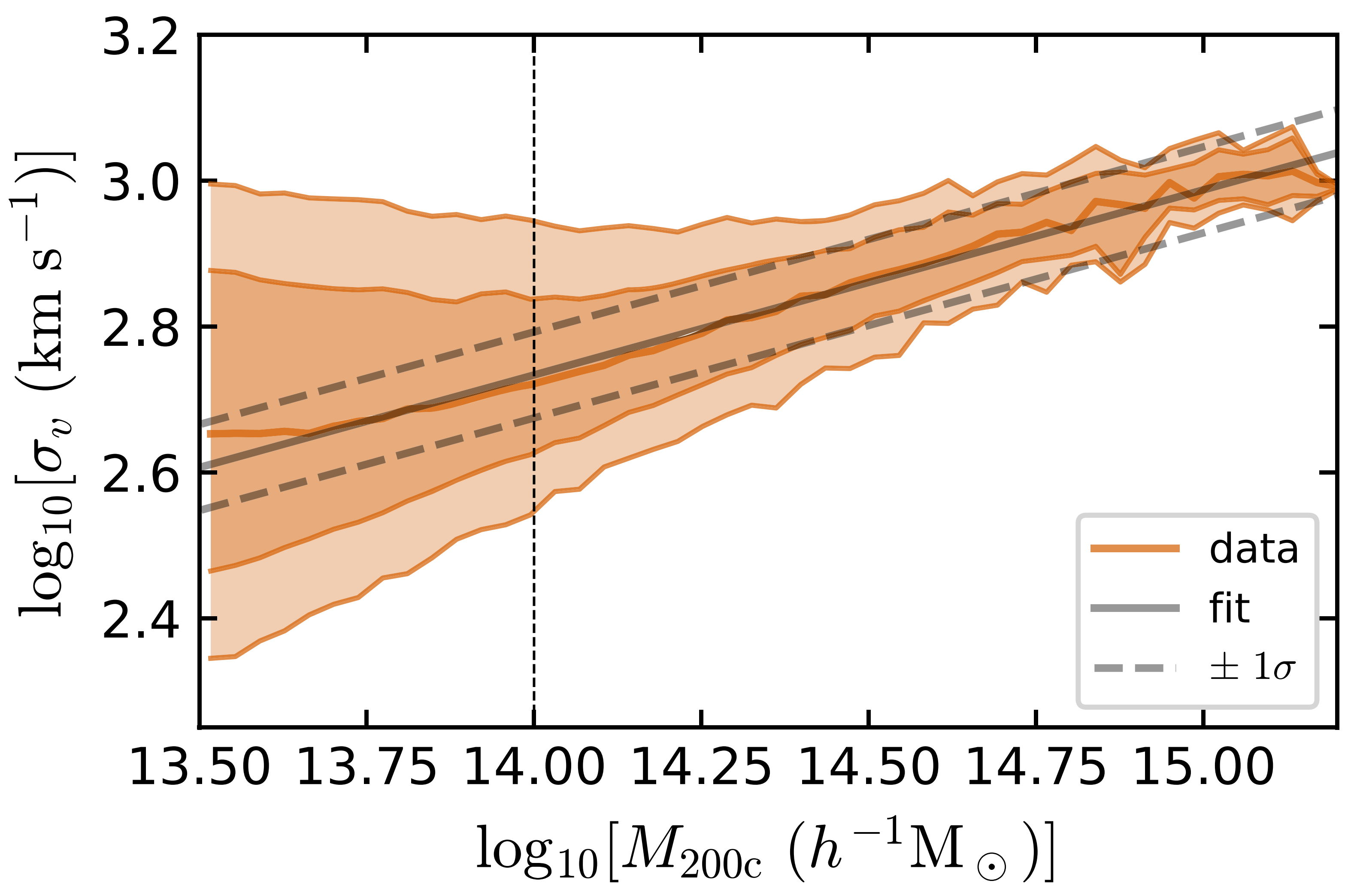


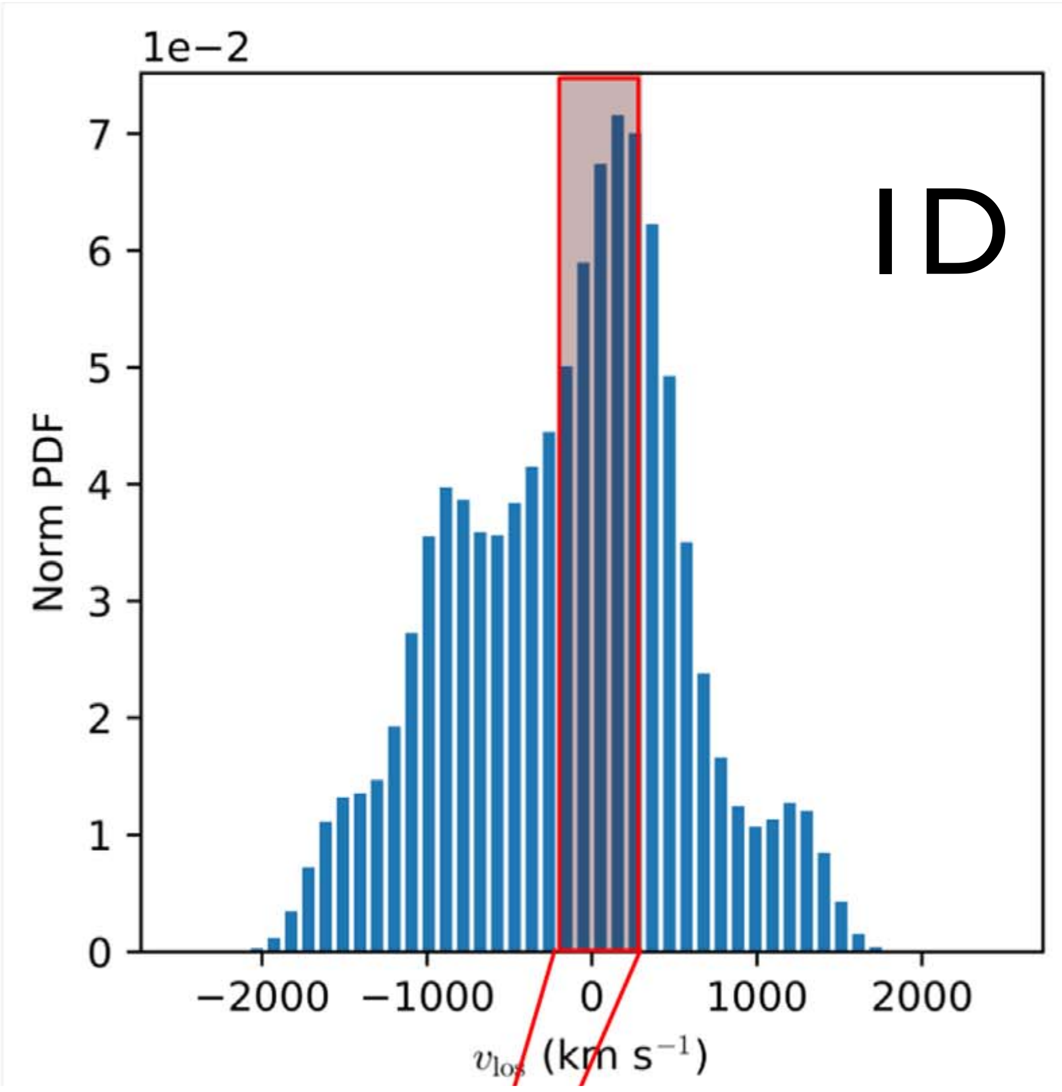
Yan et al. 20

$$x^{(2)} = \mathrm{ReLU}(\sum W_i x_i^{(1)} + b^{(2)})$$

Machine learning (ML) is a potential alternative

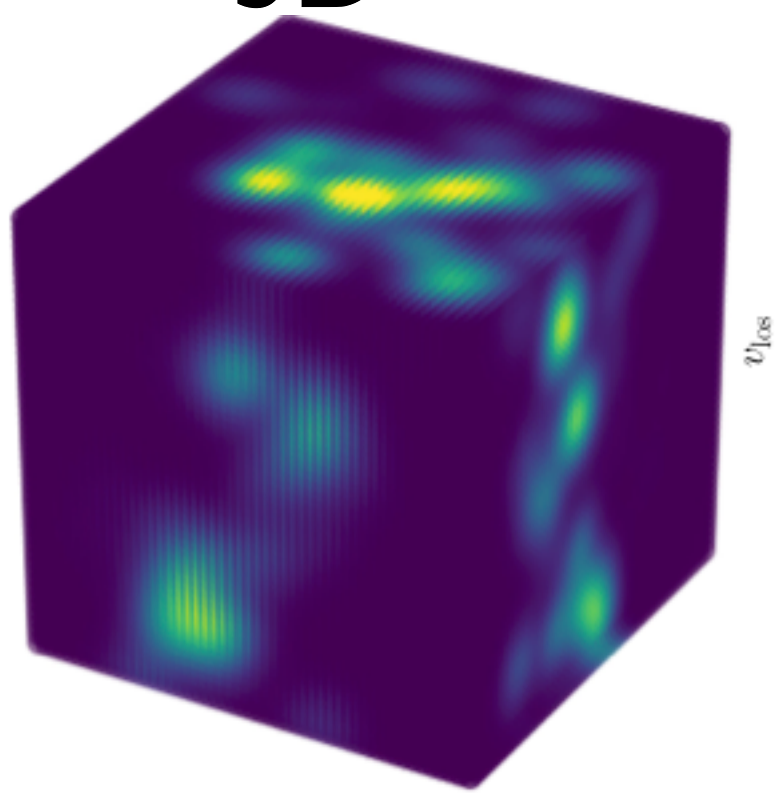
- Utilize full dataset instead of just the first order moment (σ_{velocity})



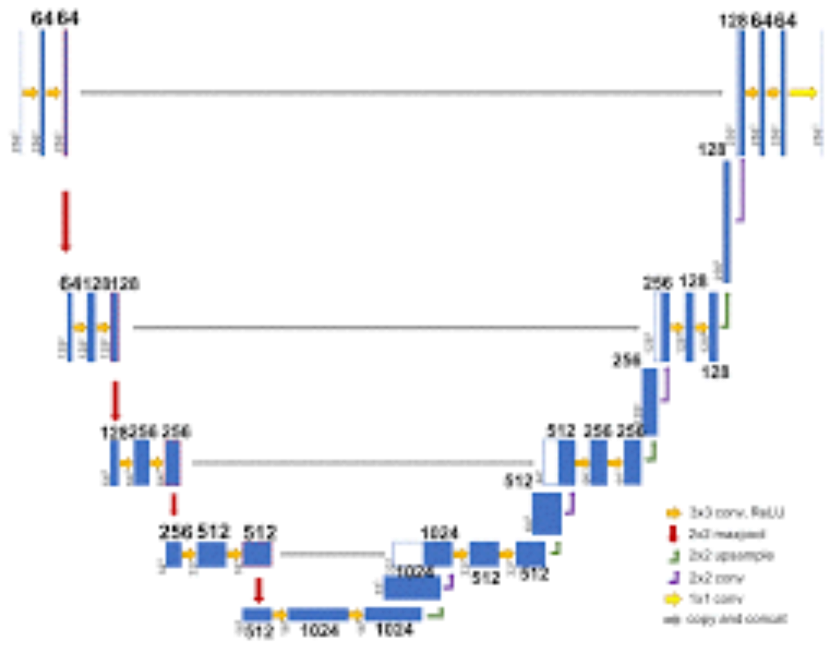


or

3D



→



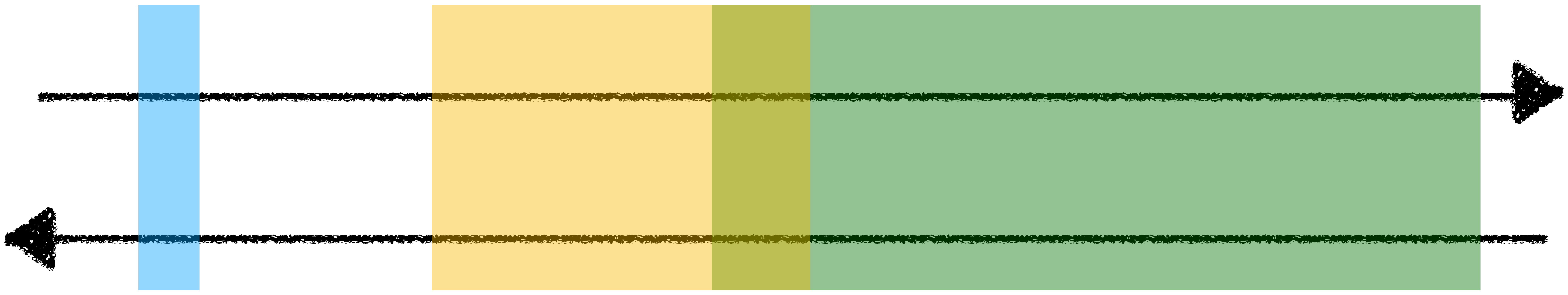
→ Cluster mass

- Ntampaka et al. 2015 (SDM)
- Ho et al. 2020, 2021 (CNN, Bayesian NN)
- Ramanah et al. 2020 (Neural flows)

Comparison of ML approaches

Power
(input dimensionality/
data size)

Generalizability/
Interpretability

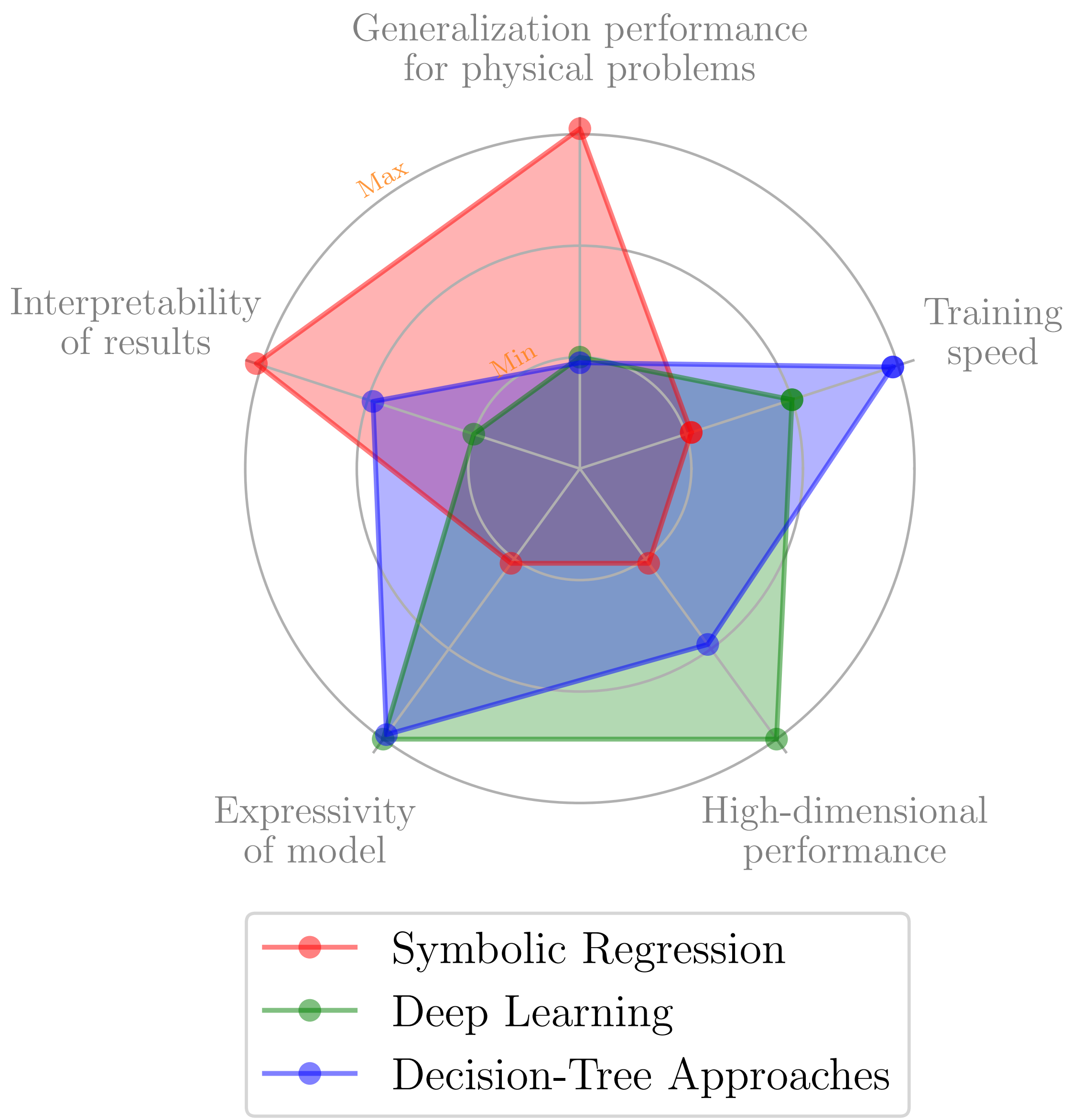


Symbolic
regression

Decision-tree
approaches
(e.g., random forests)

Deep neural
networks

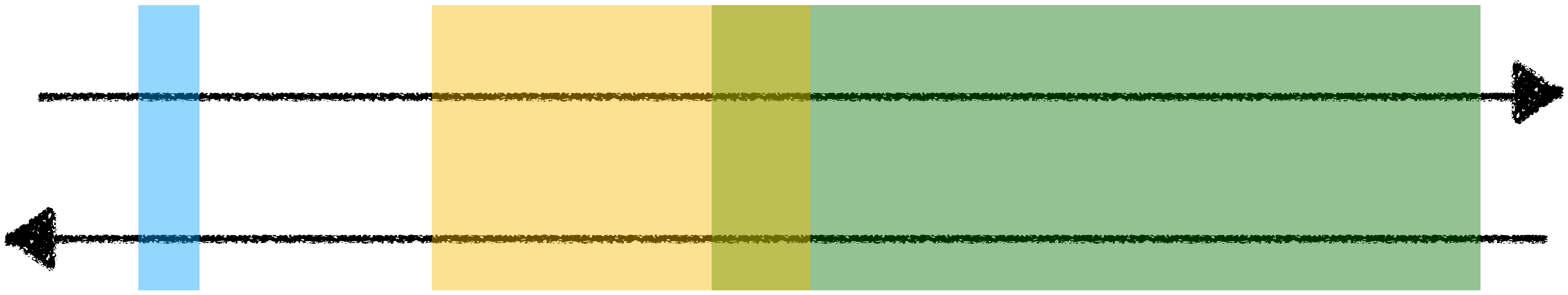
~10 parameters
~10,000 data points



Comparison of ML approaches

Power
(input dimensionality/
data size)

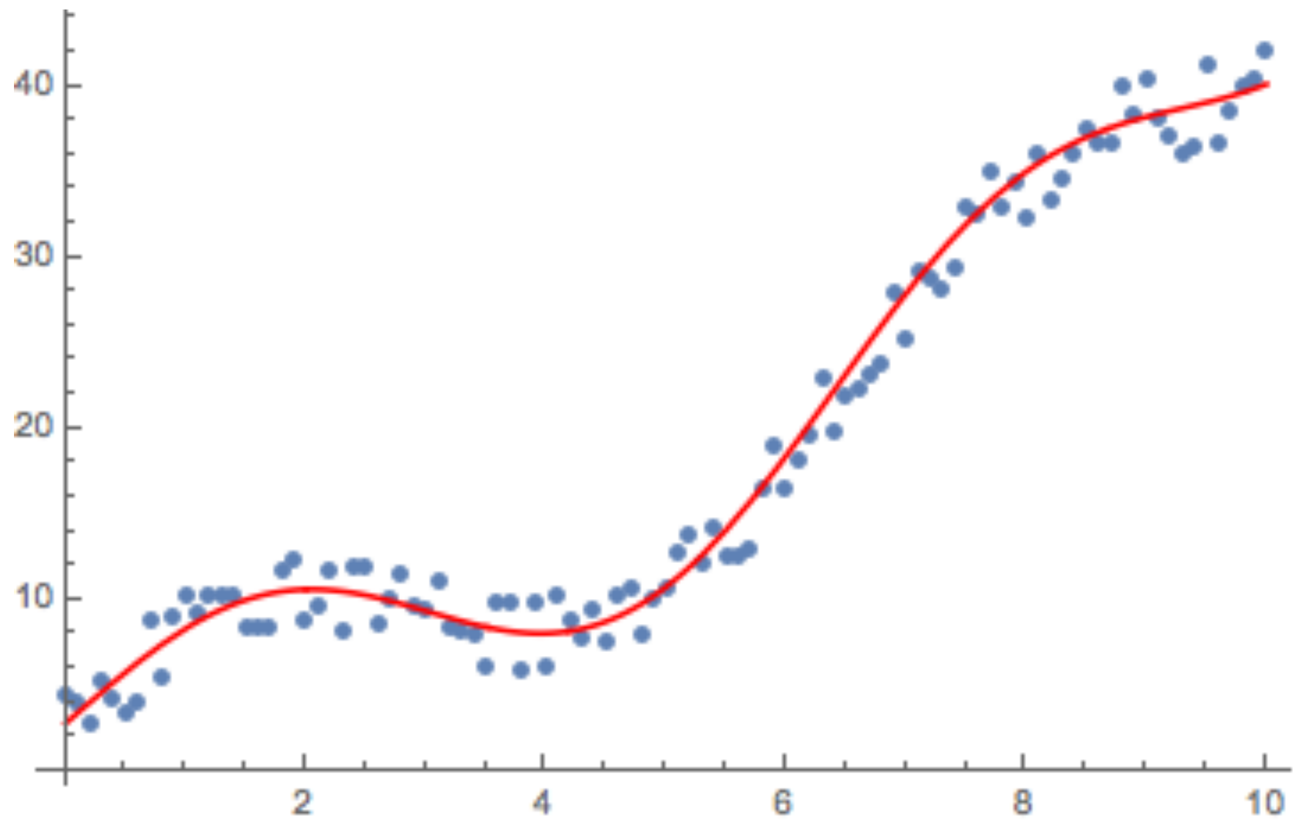
Generalizability/
Interpretability



Symbolic regression

Decision-tree approaches
(e.g., random forests)

Deep neural networks

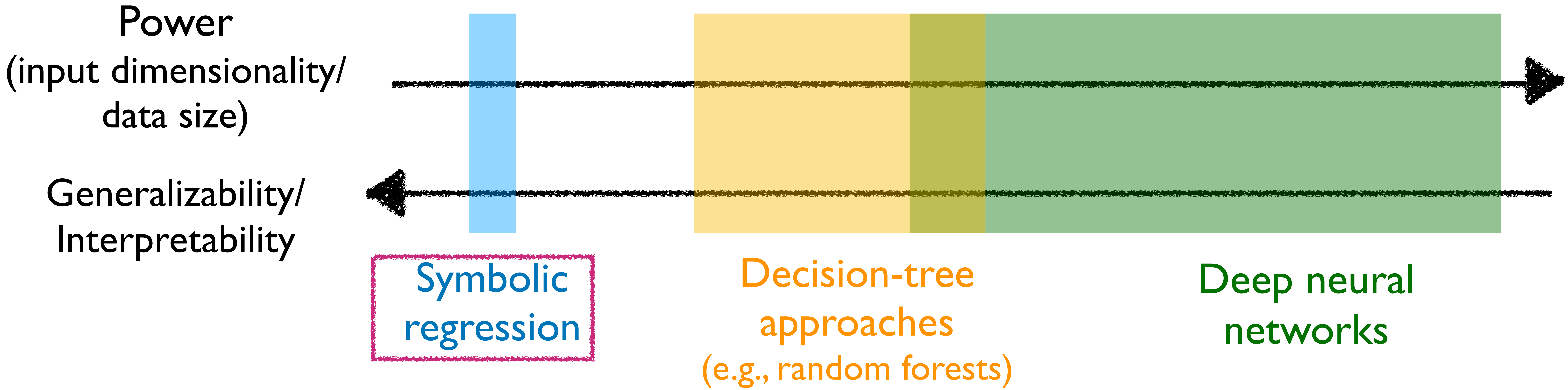


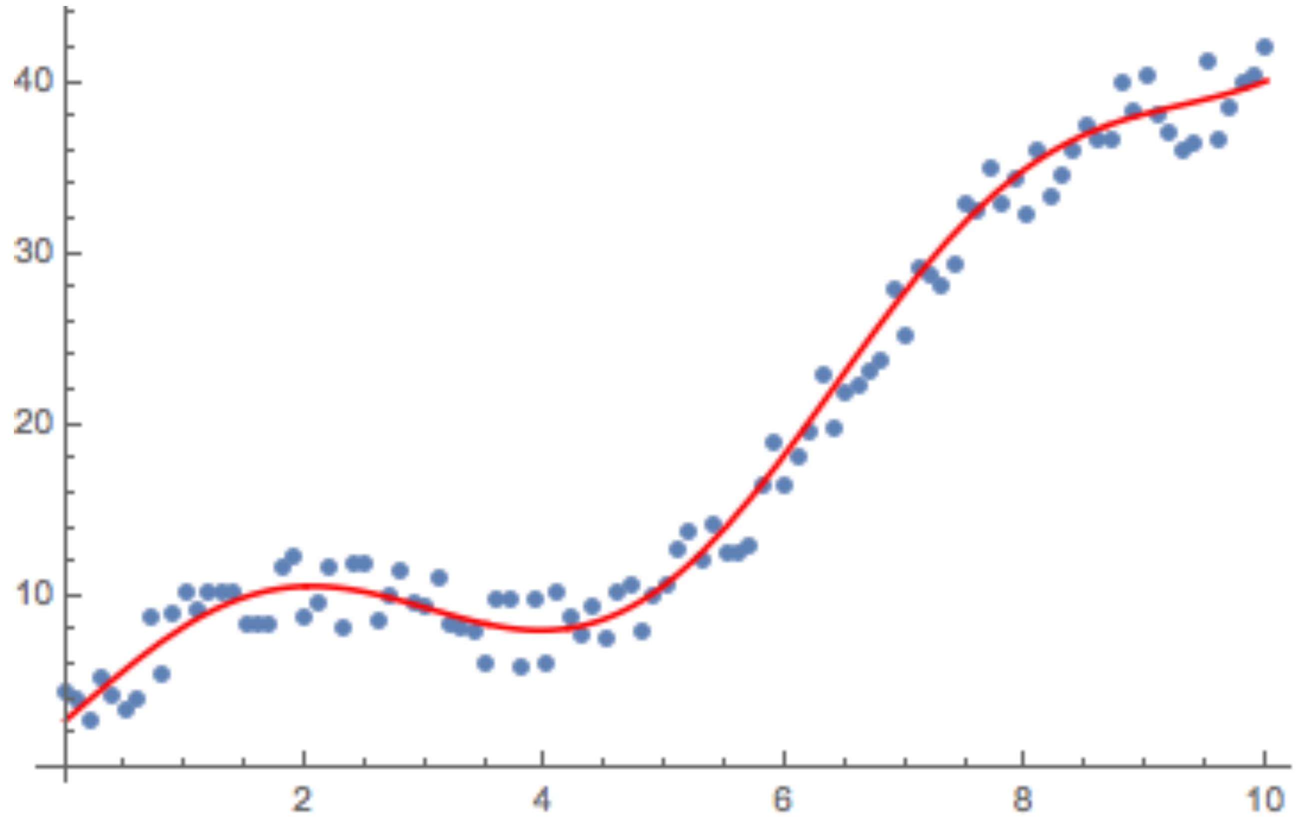
$\longrightarrow$

$\cos(x)$ ✘

$a + bx$ Close

Comparison of ML approaches





$\cos(x)$ ✘

$a + bx$ (Close)

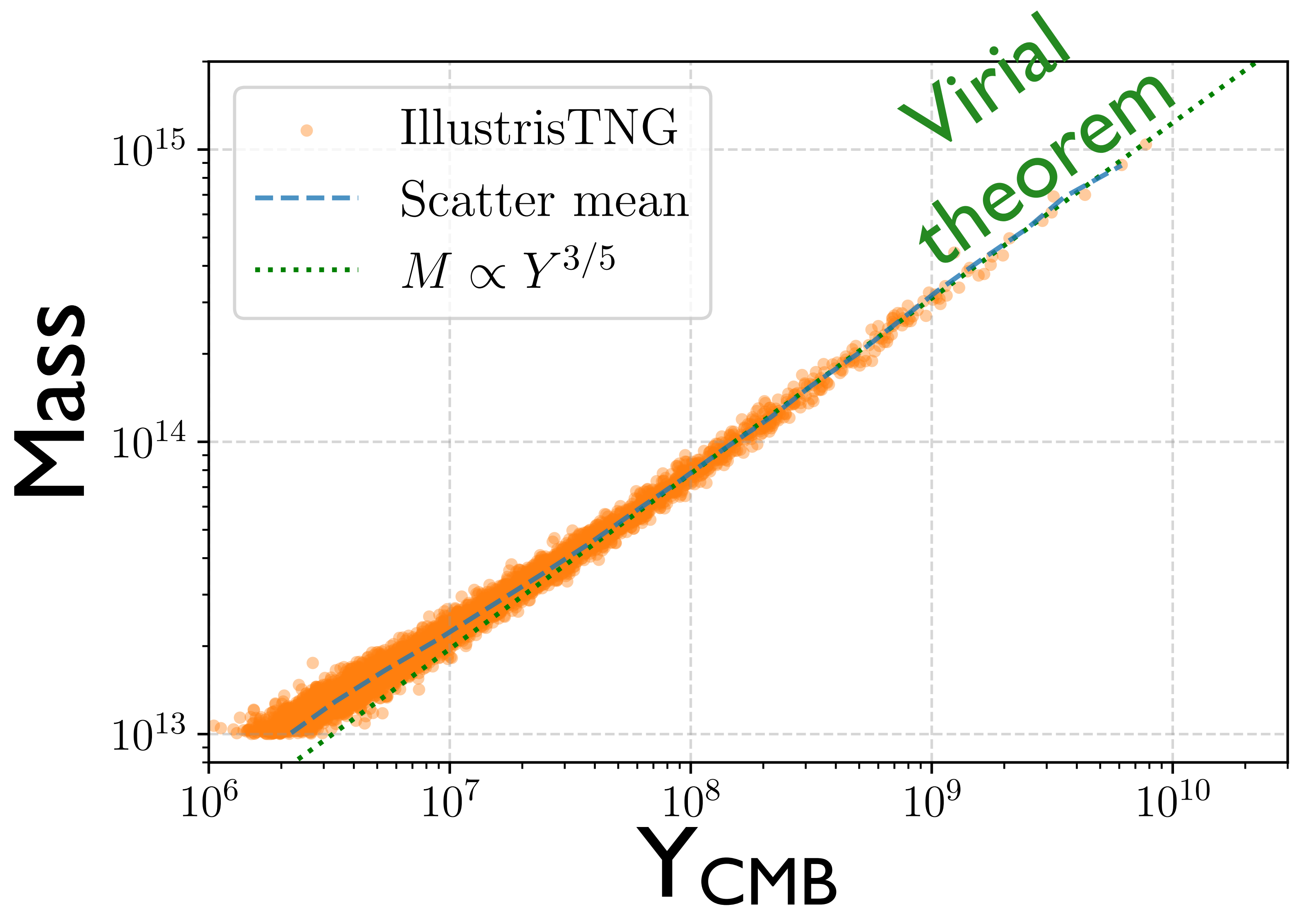
$a + bx + d\sin(x)$ (Closer!)

$a + bx + cx^2 + d\sin(x)$ ✔

PySR package:
https://github.com/MilesCranmer/PySR

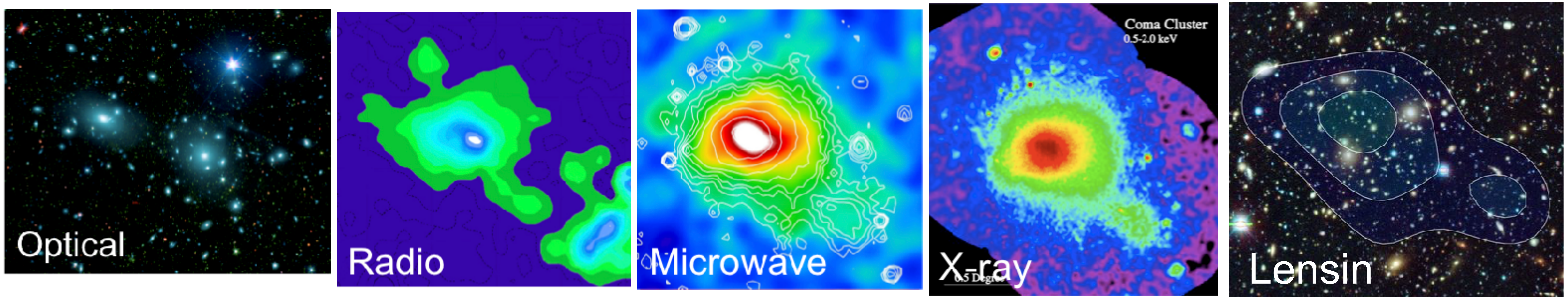
Our approach:
Symbolic regression + Random Forest

$$M_{\mathrm{cluster}} = f\left(Y_{\mathrm{CMB}}^{3/5} \right.,\ \text{observables from other surveys?} \left.\right)$$



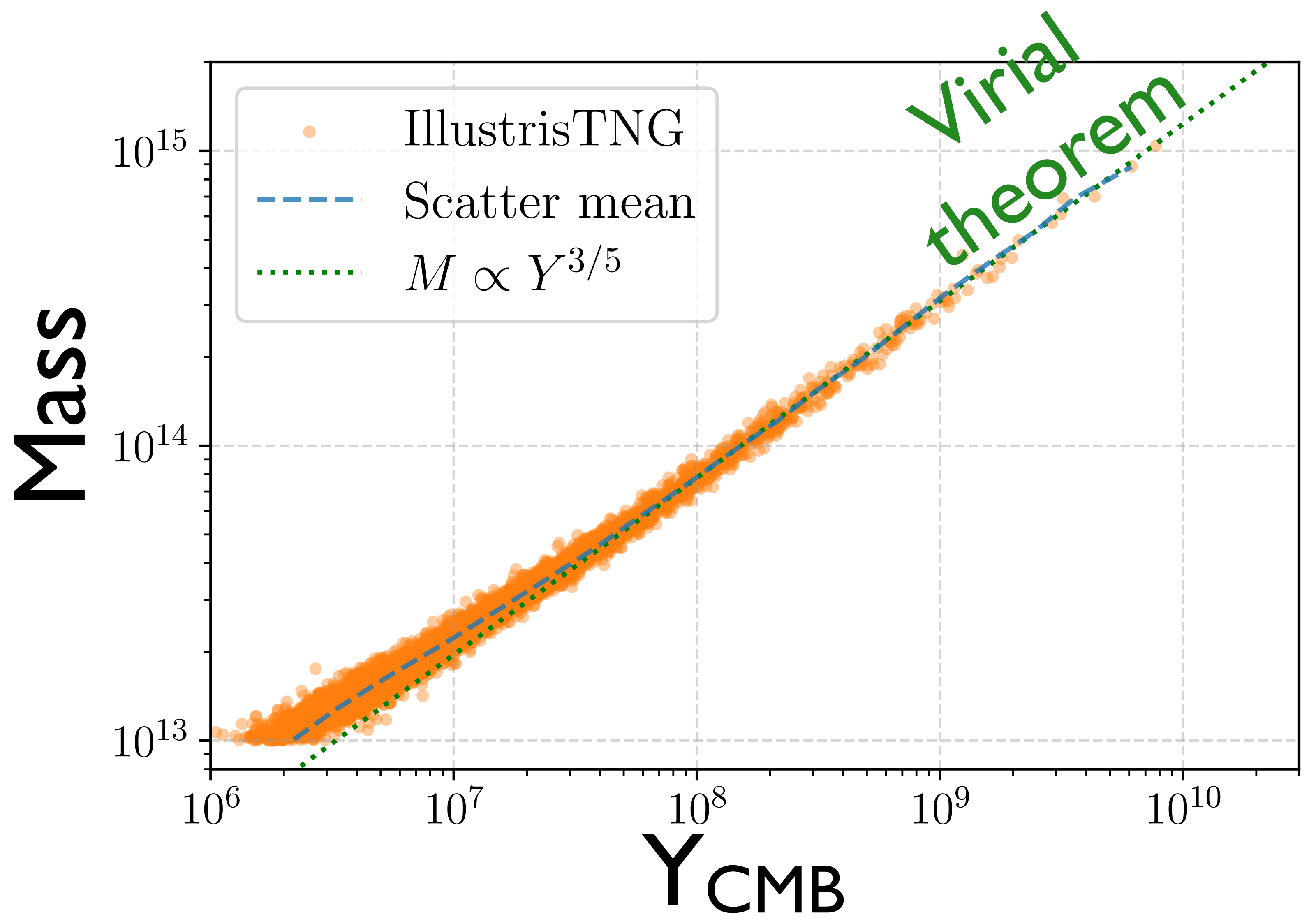
- **X-ray surveys**
  - Gas mass profile
  - Luminosity profile
  - Spectral temperature
  - Gas ellipticity
  - …….

- **Galaxy surveys**
  - Richness
  - Galaxy colors (e.g. fraction of red galaxies)
  - Stellar mass
  - …….



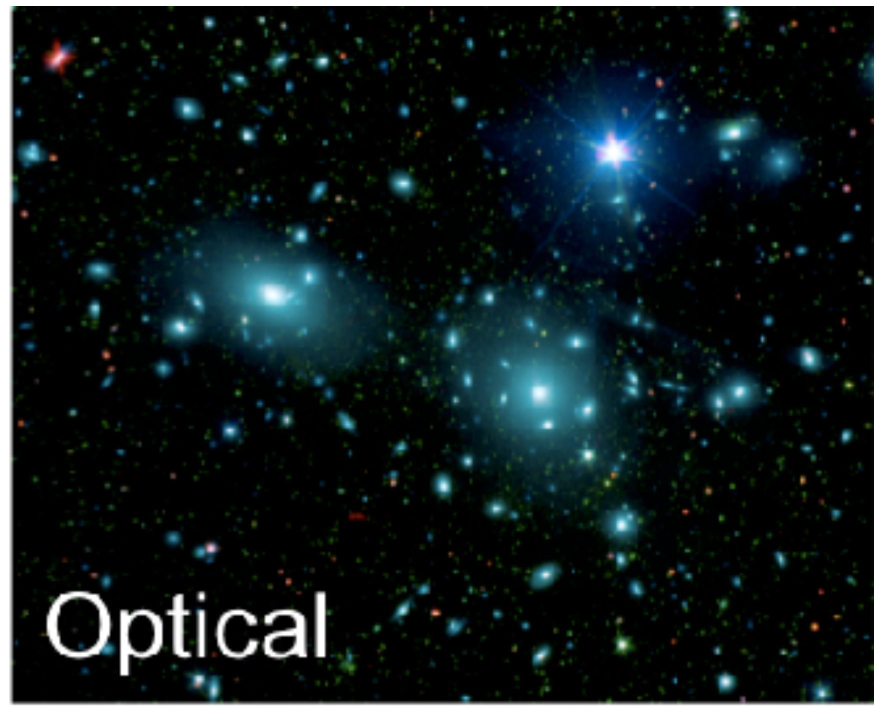
Our approach:
Symbolic regression + Random Forest

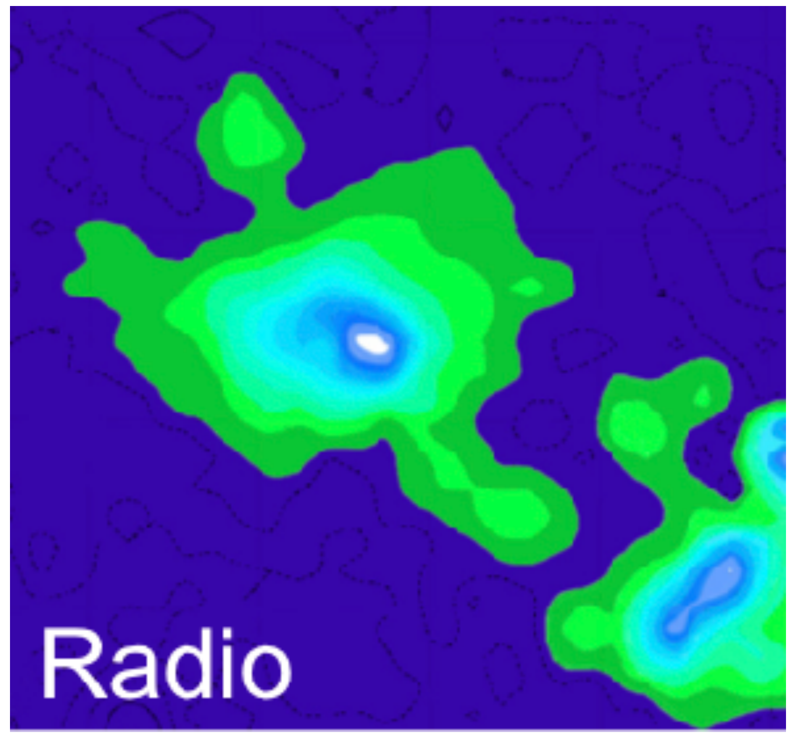
$$M_{\mathrm{cluster}} = f\left(Y_{\mathrm{CMB}}^{3/5}, \text{ observables from other surveys? } \right)$$

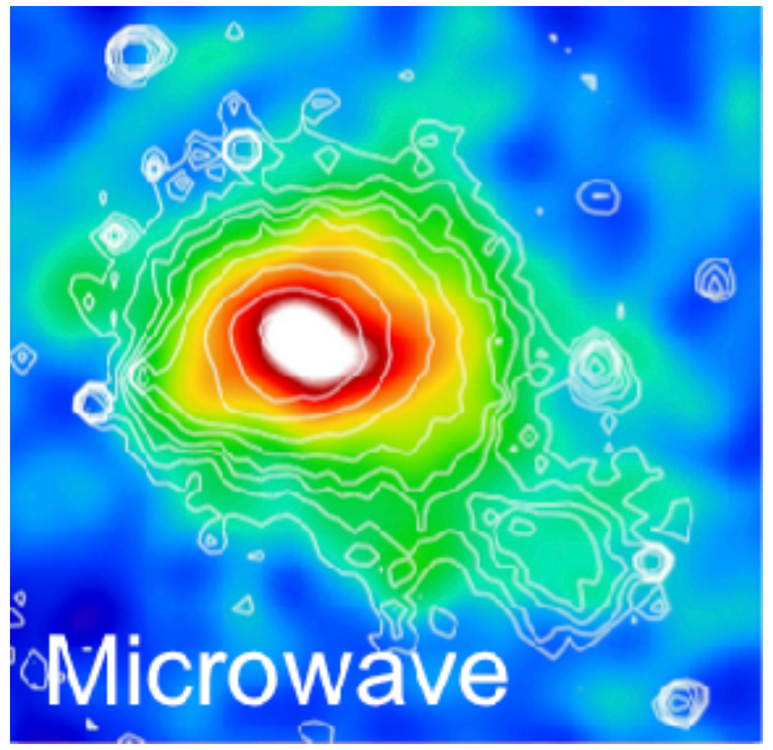


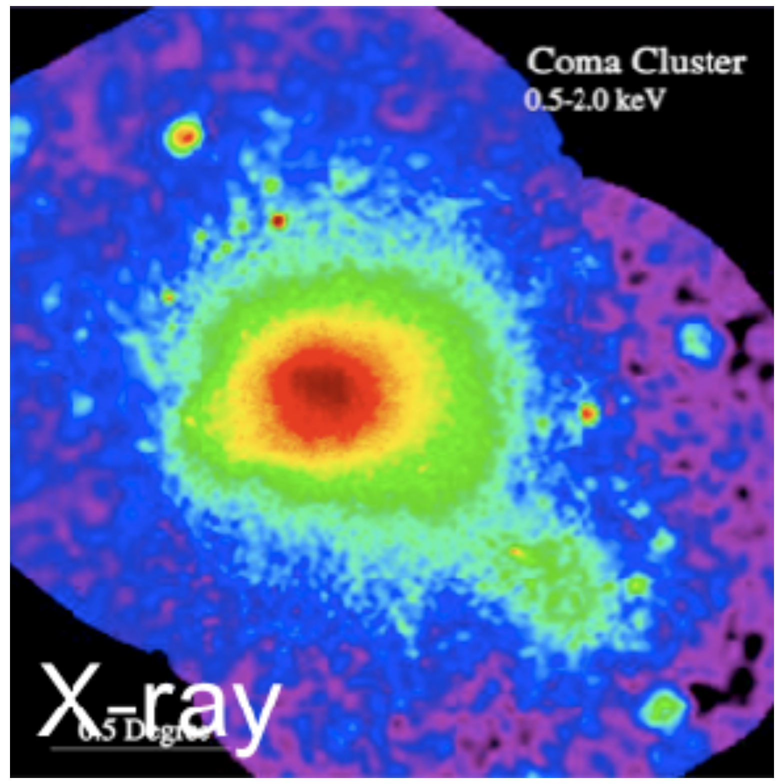
- **X-ray surveys**
  - $\boxed{\text{Gas mass profile}}$
  - Luminosity profile
  - Spectral temperature
  - Gas ellipticity
  - …….

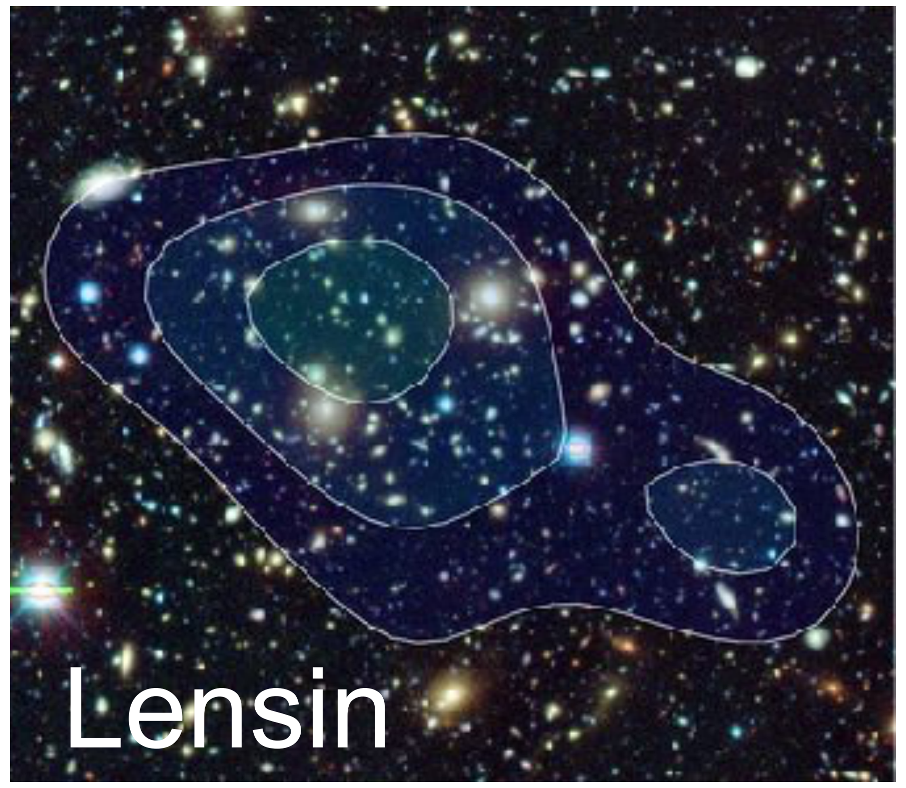
- **Galaxy surveys**
  - Richness
  - Galaxy colors (e.g. fraction of red galaxies)
  - $\boxed{\text{Stellar mass}}$
  - …….

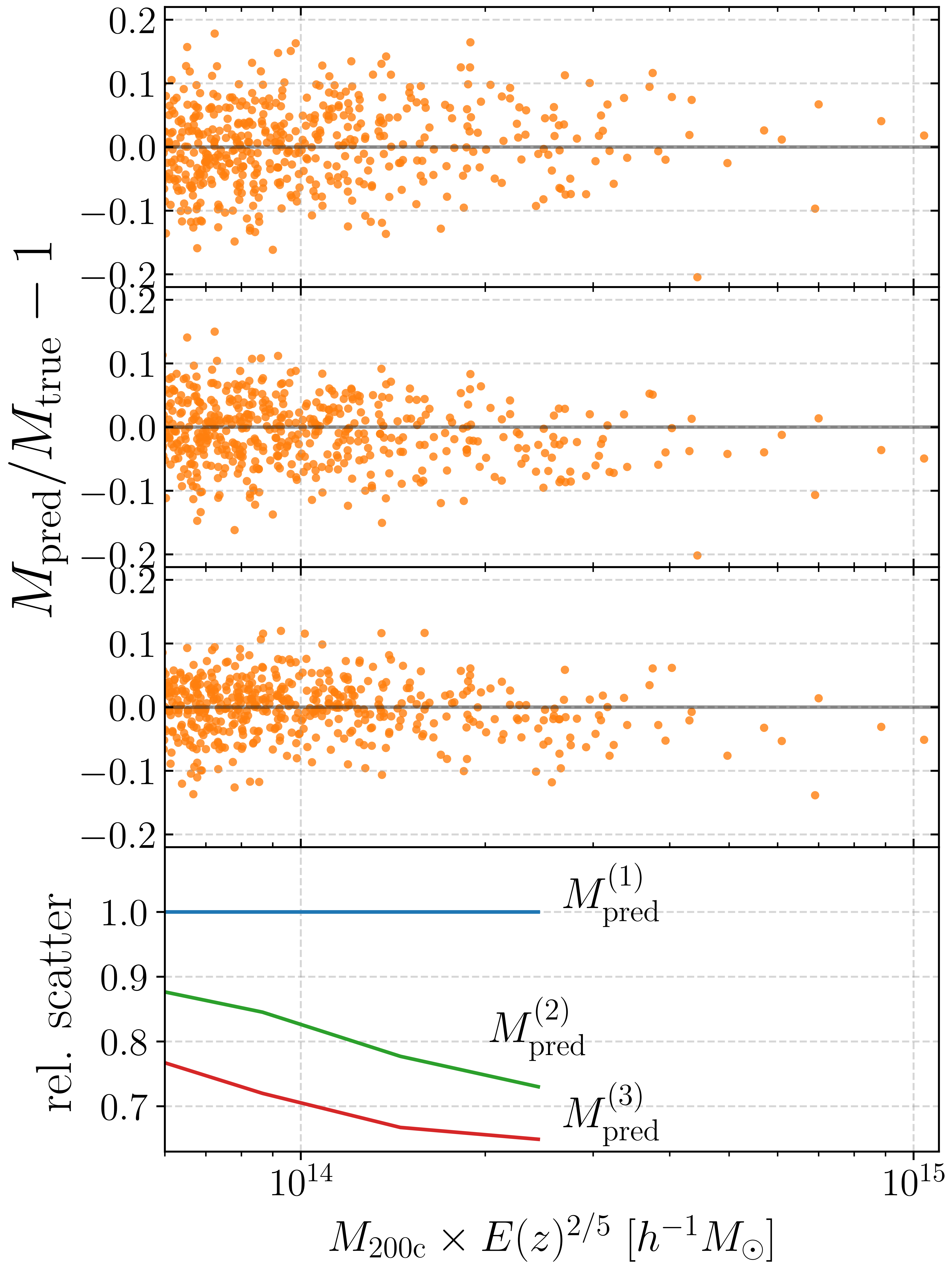












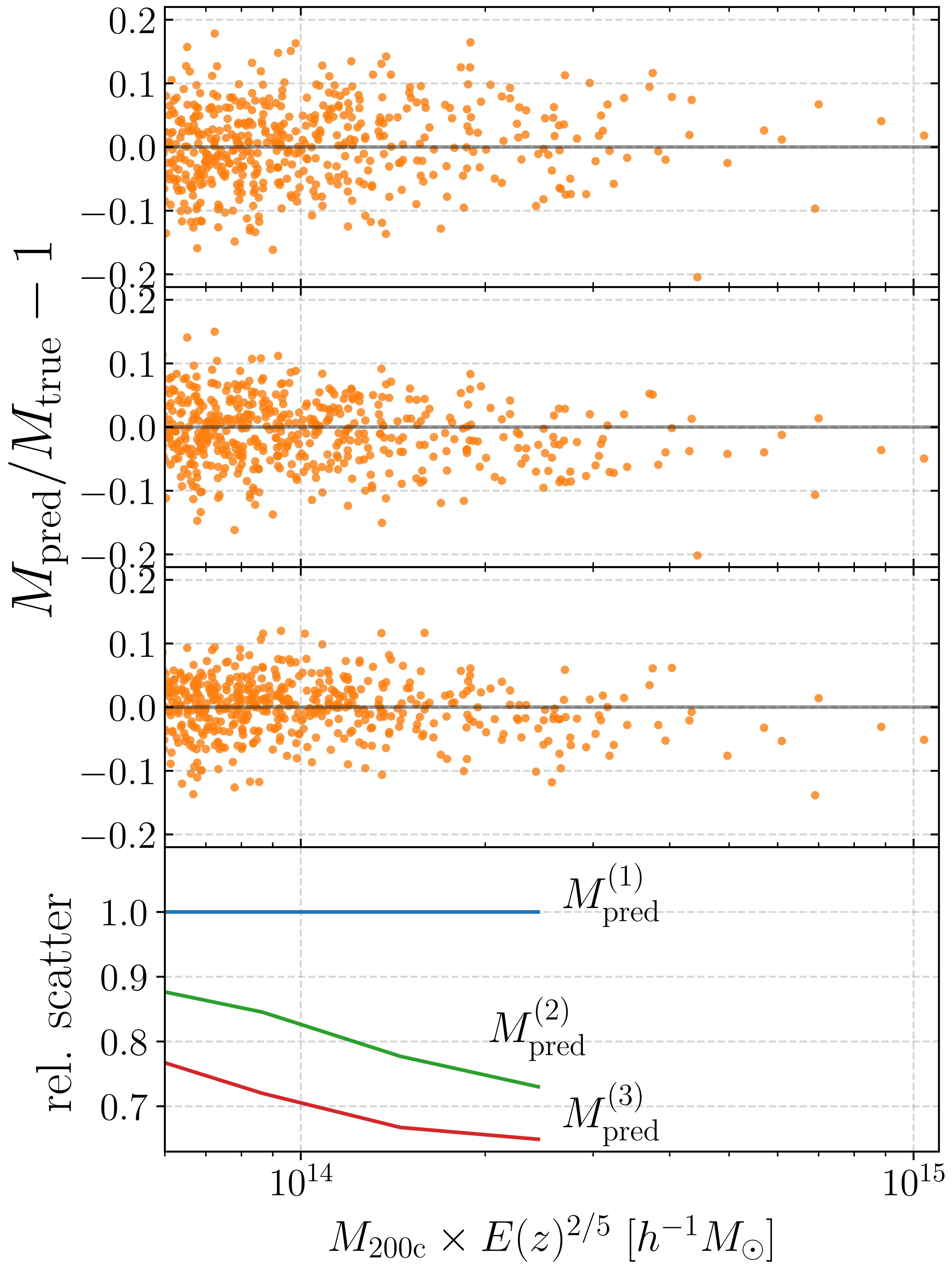
Results for IllustrisTNG

$$M^{(1)}_{\rm pred} \propto Y^{3/5}$$

$$M^{(2)}_{\rm pred} \propto Y^{3/5}\,(1 - A\,c_{\rm gas}) \qquad c_{\rm gas} \equiv \frac{M_{\rm gas}(r < R_{200c}/2)}{M_{\rm gas}(r < R_{200c})}$$

$$M^{(3)}_{\rm pred} \propto Y^{3/5} \left(\frac{B}{c_{\rm NFW}} \right)^{M_*/M_{\rm gas}}$$

Results for IllustrisTNG



$$M^{(1)}_{\rm pred} \propto Y^{3/5}$$

$$M^{(2)}_{\rm pred} \propto Y^{3/5}\,(1 - A\,c_{\rm gas}) \qquad c_{\rm gas} \equiv \frac{M_{\rm gas}(r < R_{200c}/2)}{M_{\rm gas}(r < R_{200c})}$$

$$M^{(3)}_{\rm pred} \propto Y^{3/5}\left(\frac{B}{c_{\rm NFW}}\right)^{M_*/M_{\rm gas}}$$

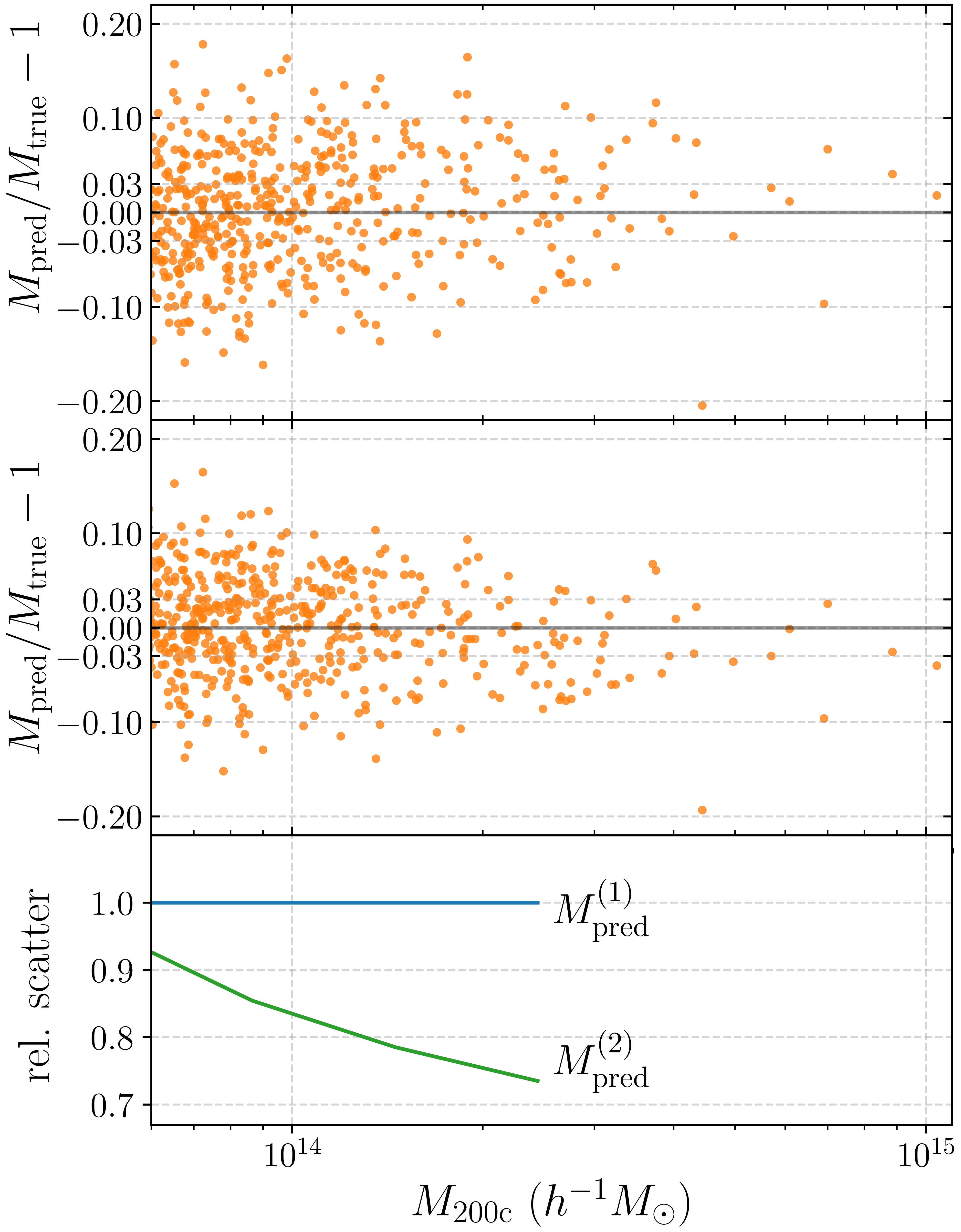
Reasons for dependence:

1. Central regions of clusters are noisier
   (conc. can be used to down-weight central regions)

2. Conversion of gas to stars reduces Y

Kravtsov et al. 06, Arnaud et al. 10

Results for IllustrisTNG



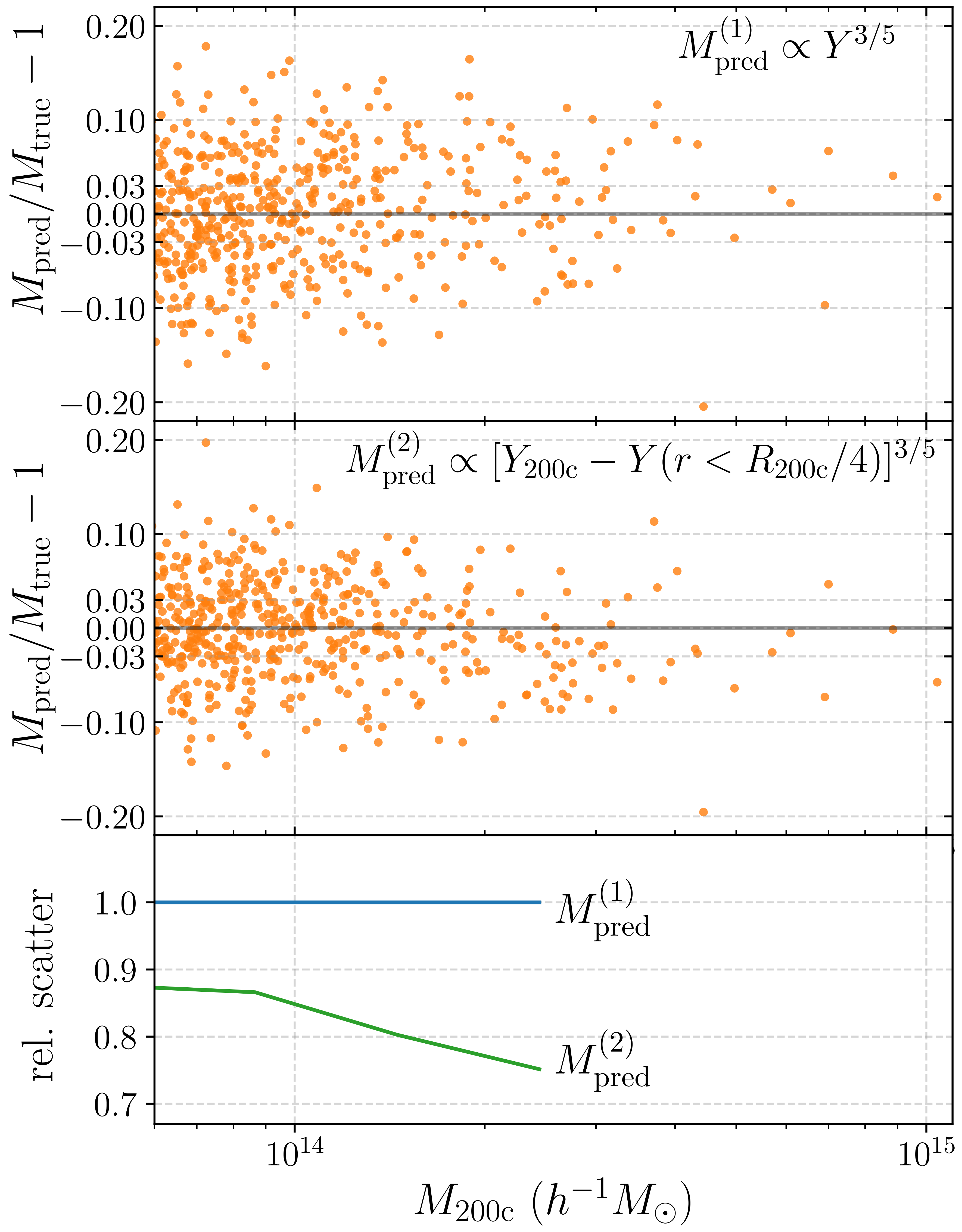
$$M^{(1)}_{\rm pred} \propto Y^{3/5}$$

$$M^{(2)}_{\rm pred} \propto Y^{3/5}\,(1 - A\,c_{\rm gas}) \qquad c_{\rm gas} \equiv \frac{M_{\rm gas}(r < R_{200c}/2)}{M_{\rm gas}(r < R_{200c})}$$

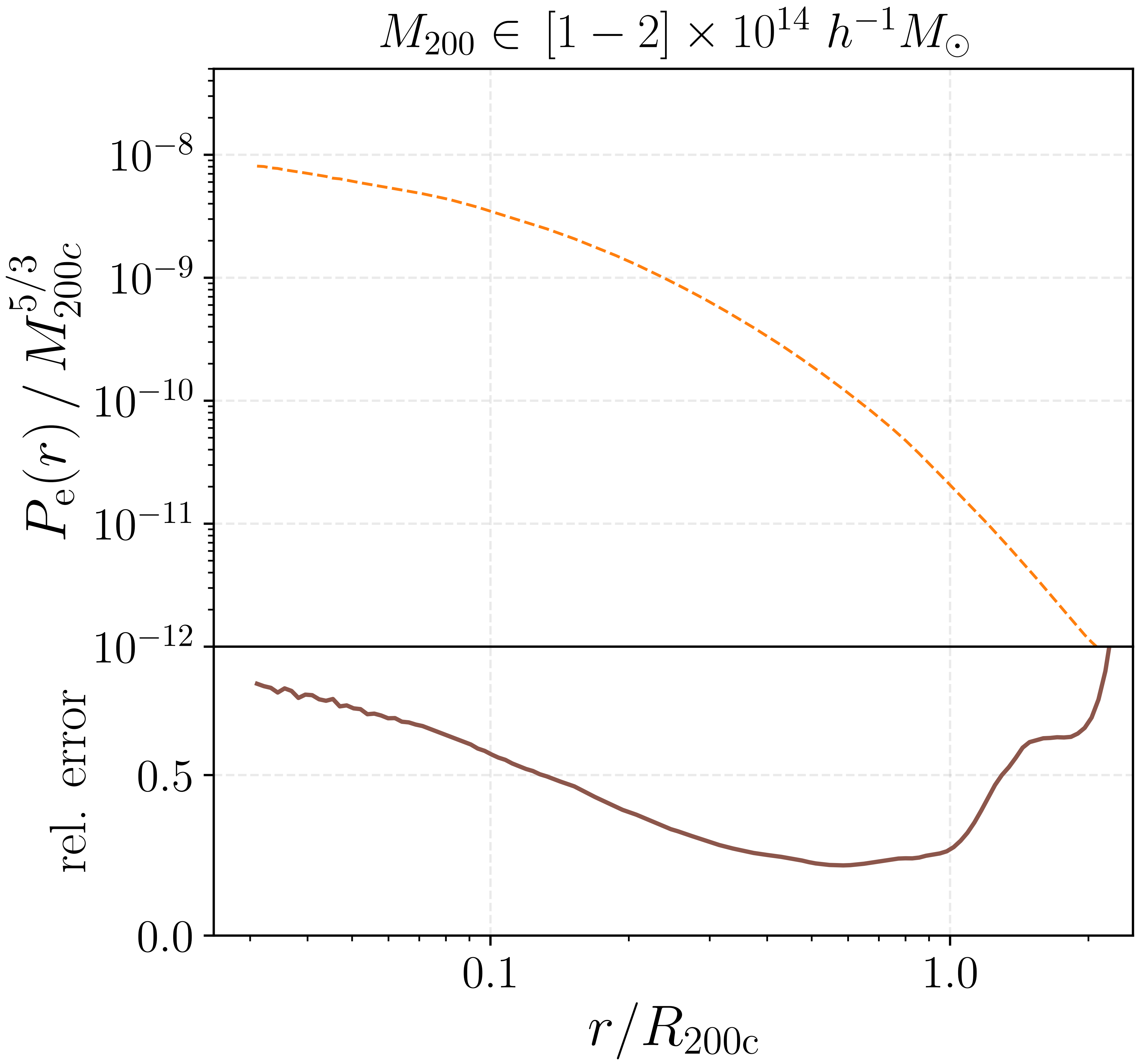
X-ray	SZ
- High resolution - Indirect probe	- Low resolution - Direct probe

Cross-checks

Excising inner cluster regions



Radial dependence of scatter

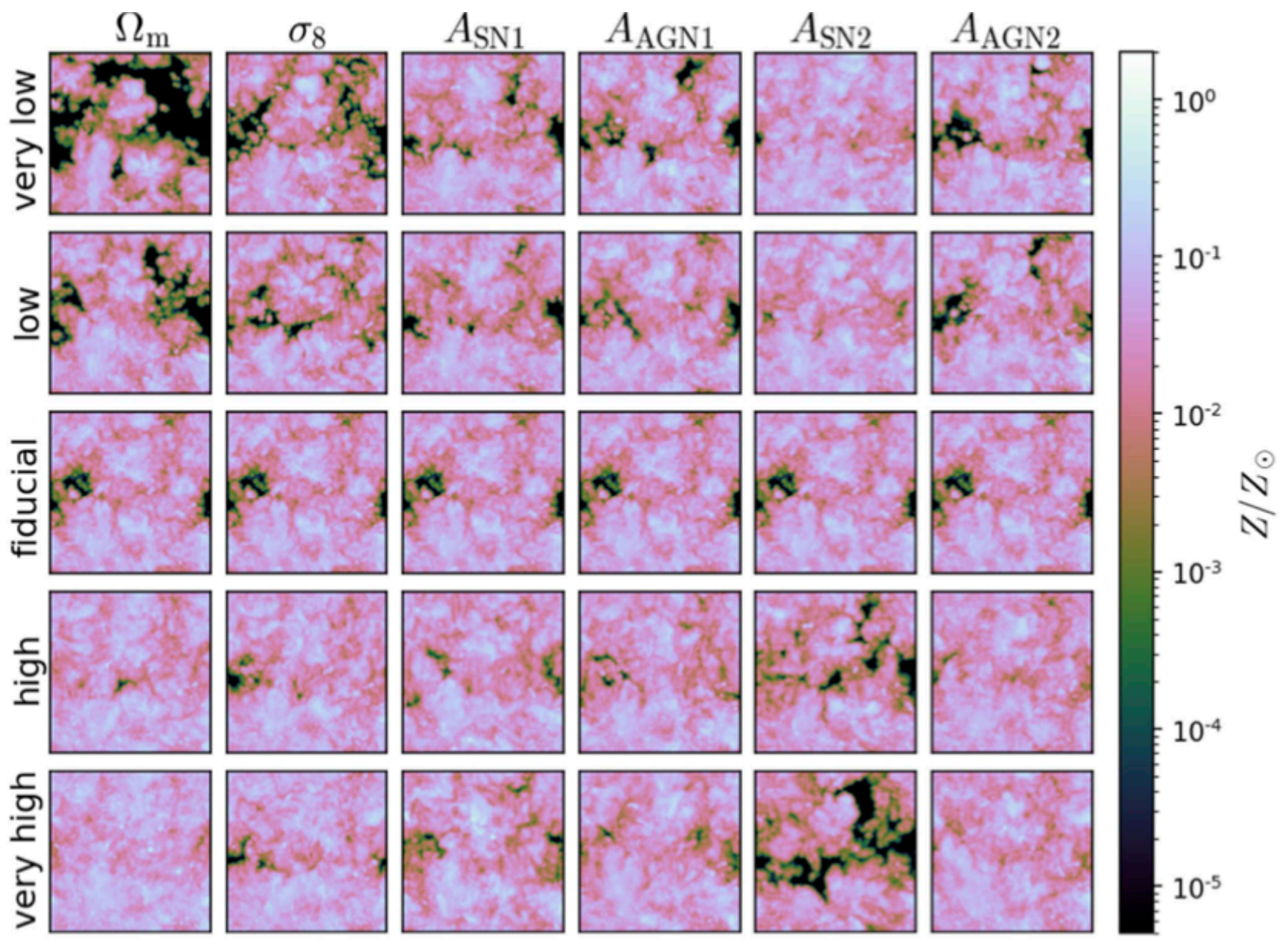


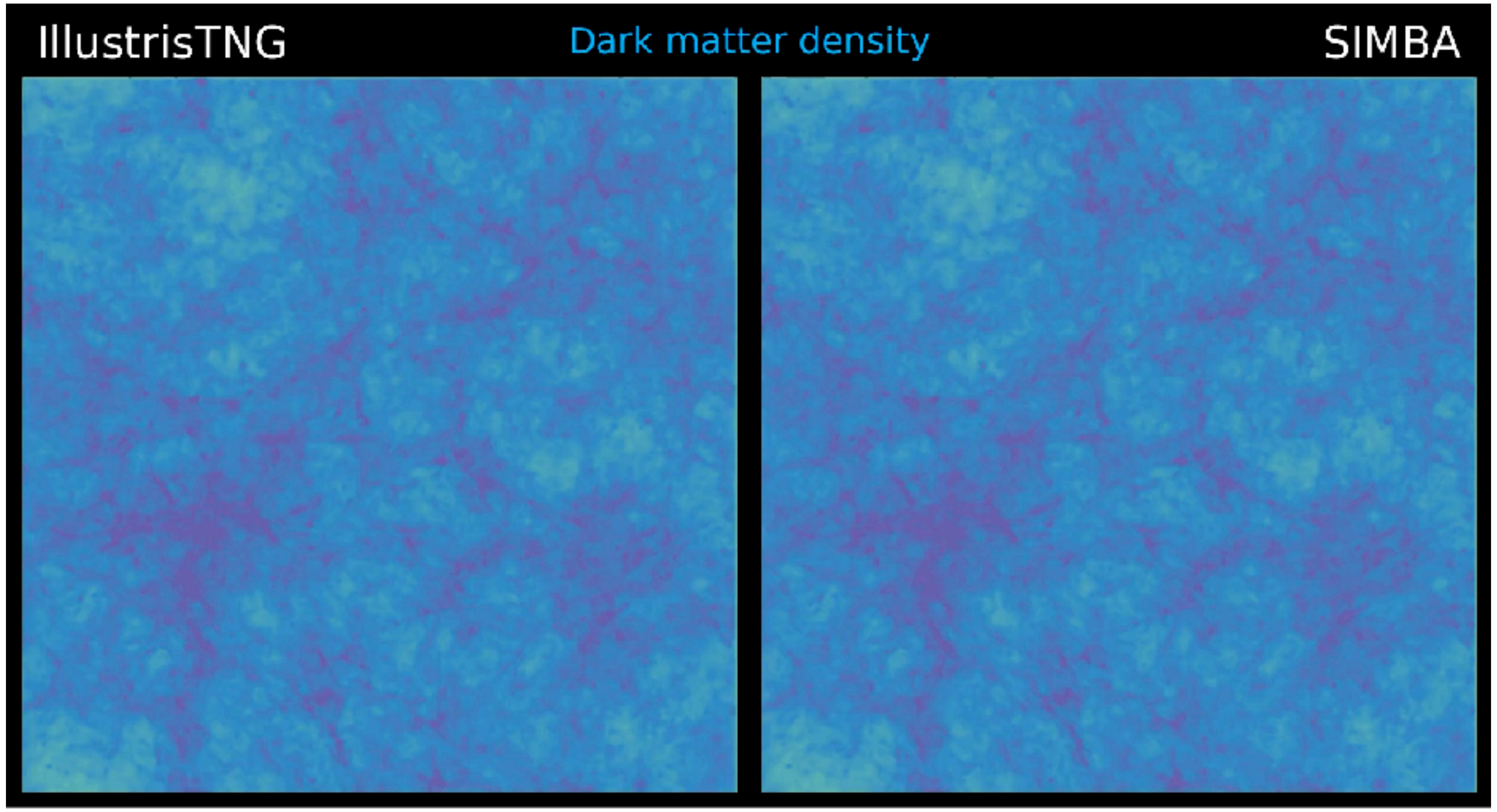
But IllustrisTNG has only one configuration of baryonic feedback and initial conditions?

Do the results hold in a more general setting?

But IllustrisTNG has only one configuration of baryonic feedback and initial conditions?

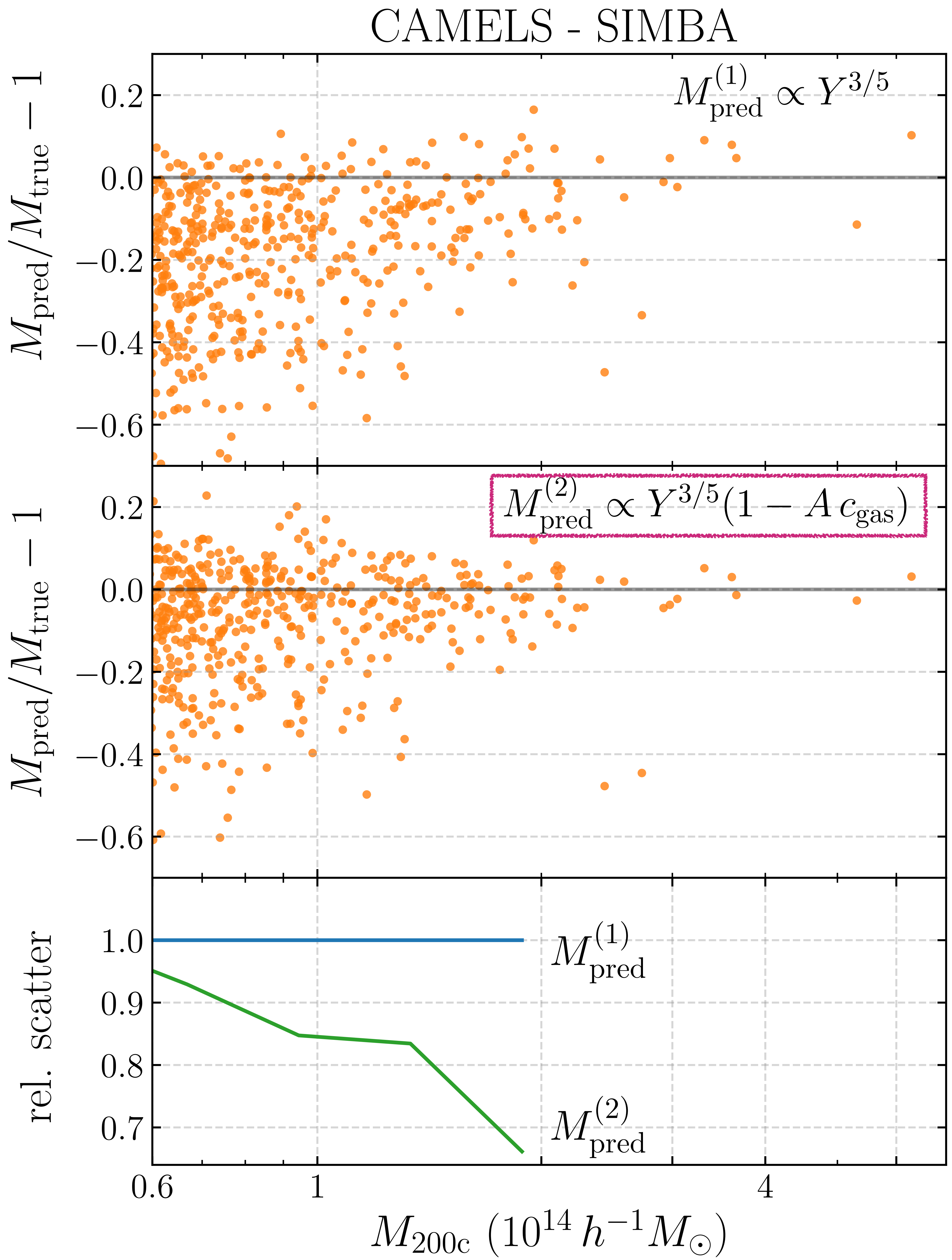
Do the results hold in a more general setting?

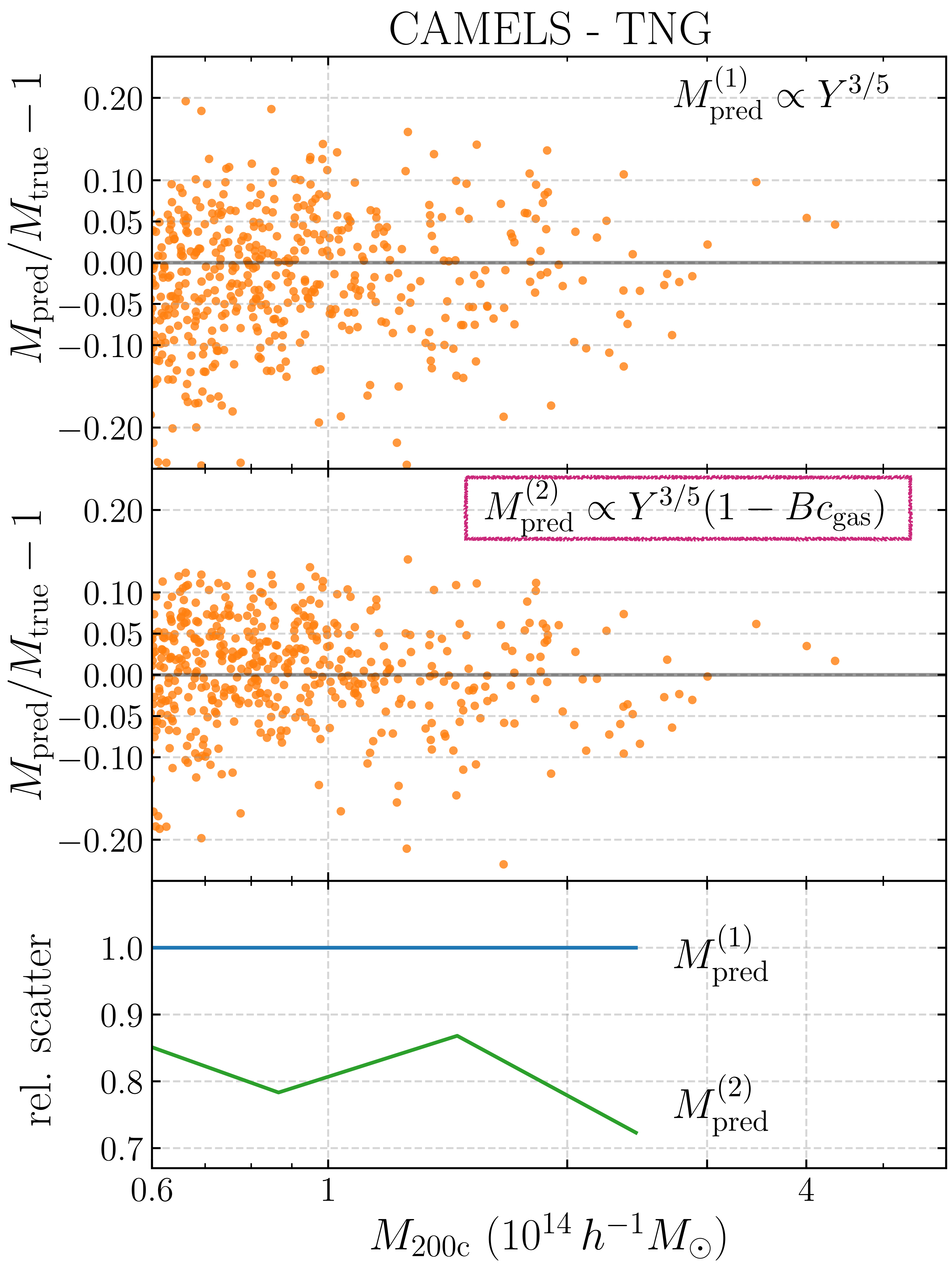




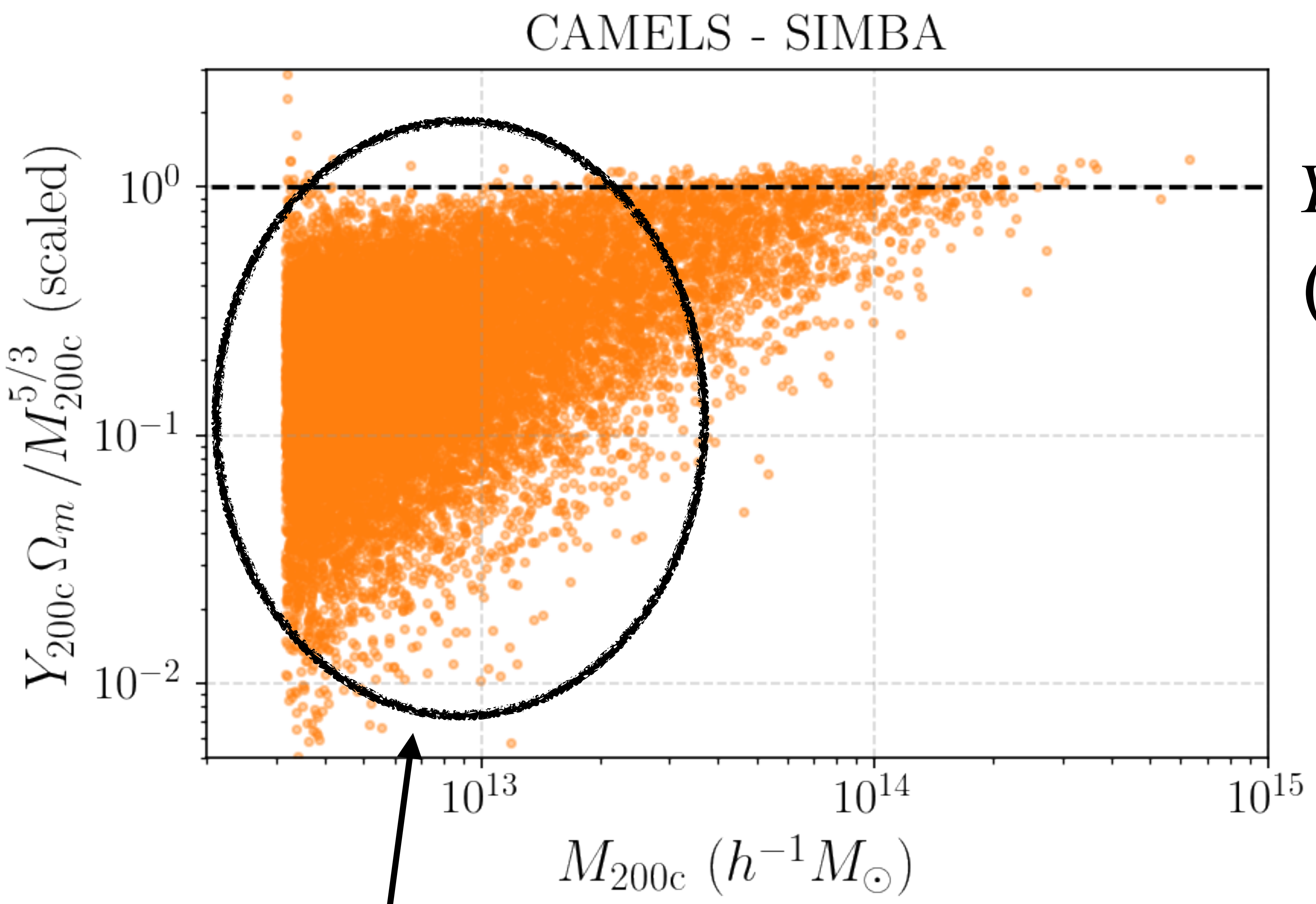
CAMELS simulations
Villaescusa-Navarro et al. 21
https://camels.readthedocs.io/

Results for CAMELS





Reducing deviation from self-similarity (pow. law)

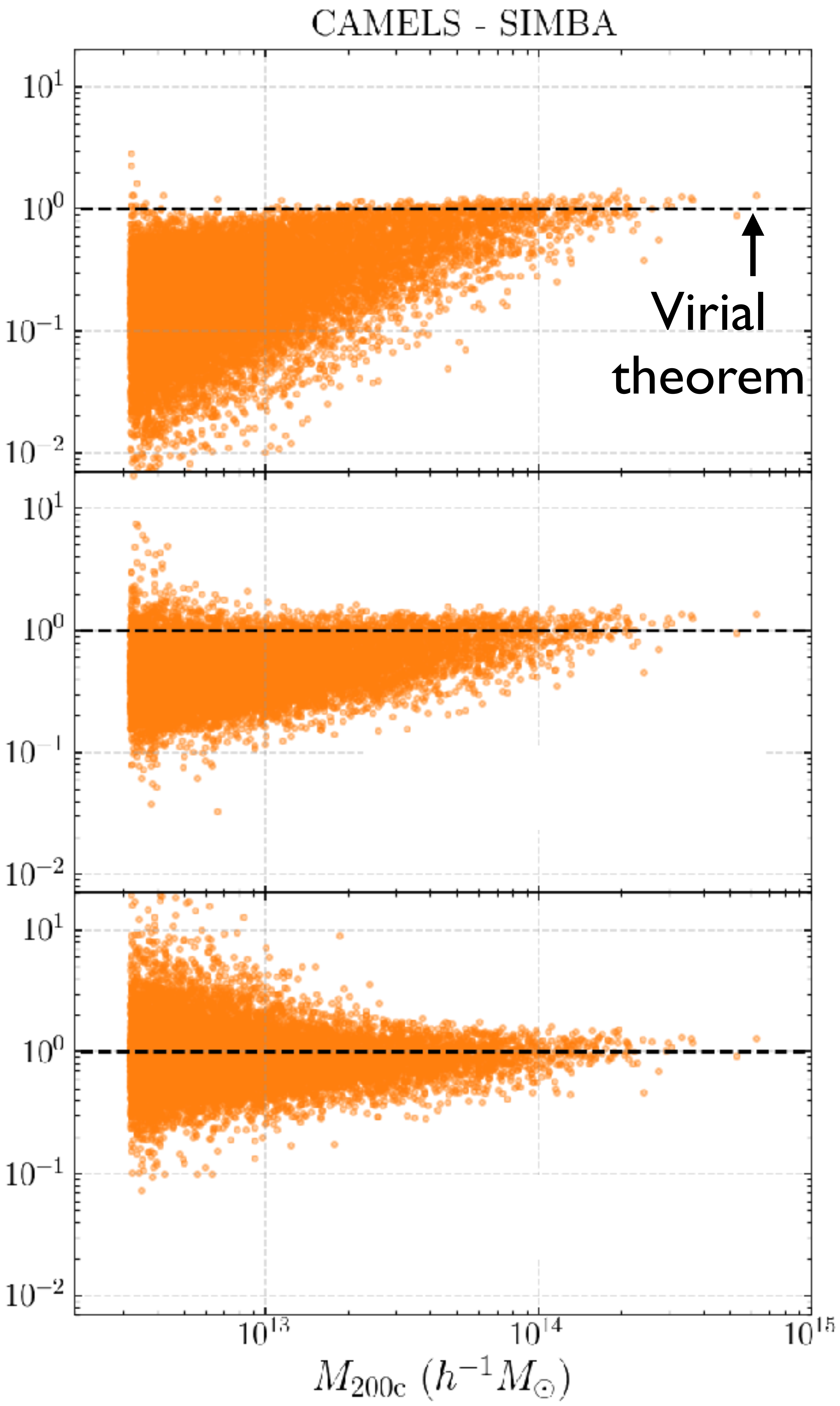


$Y \propto M^{5/3}$

(virial theorem)

Due to ejection of gas from clusters from AGN/SN feedback

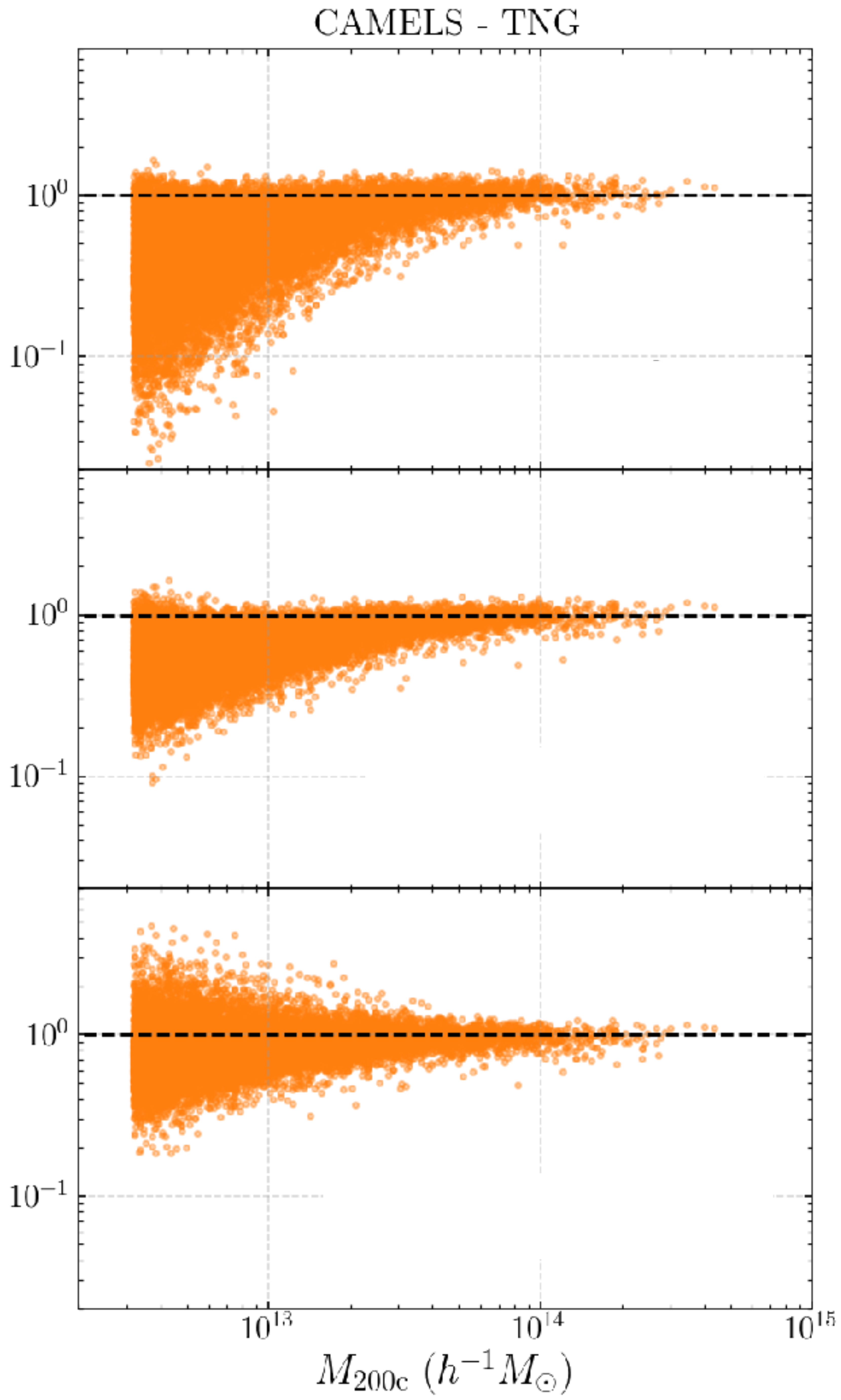
Preliminary results



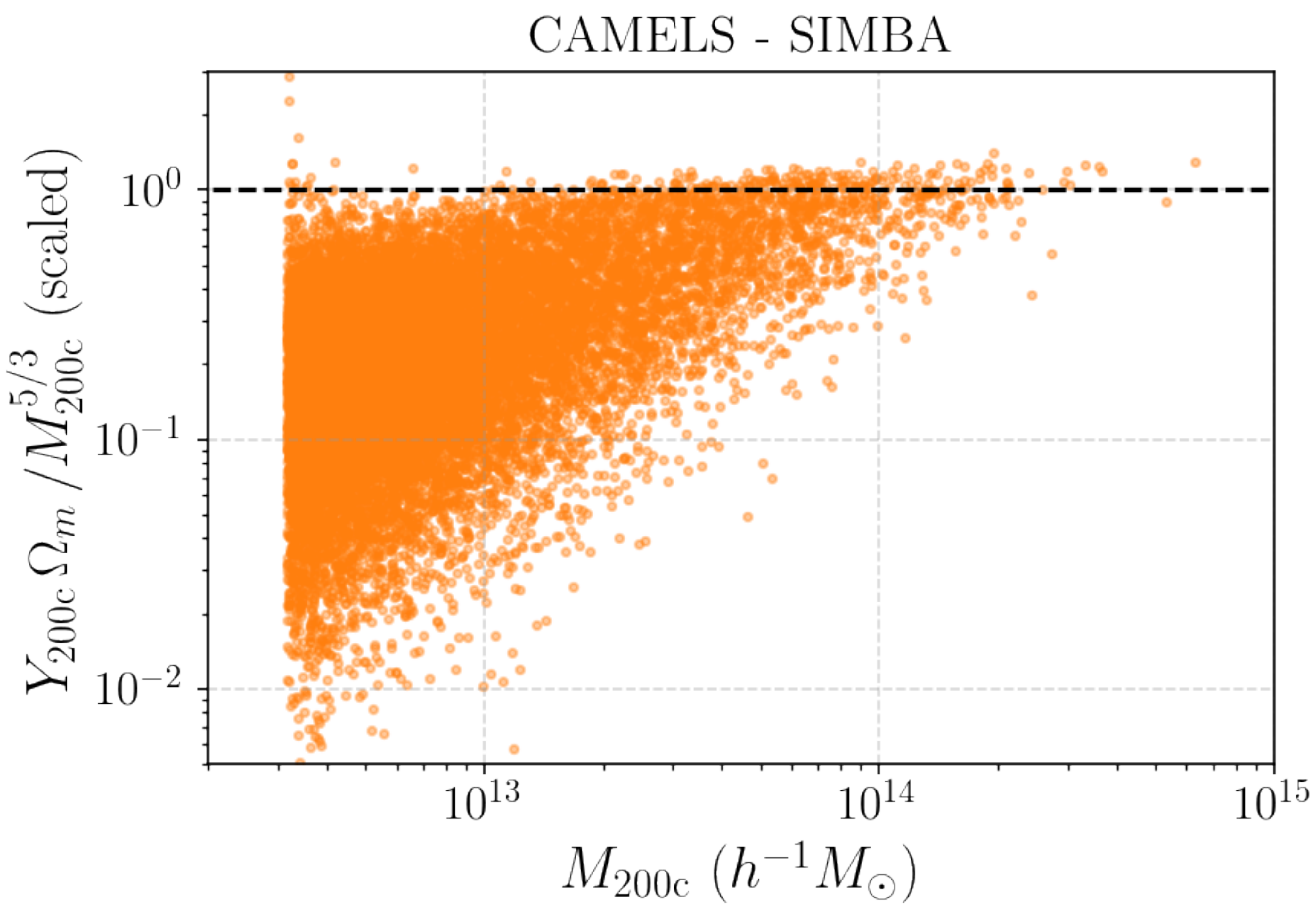
$\leftarrow \quad Y \quad \rightarrow$

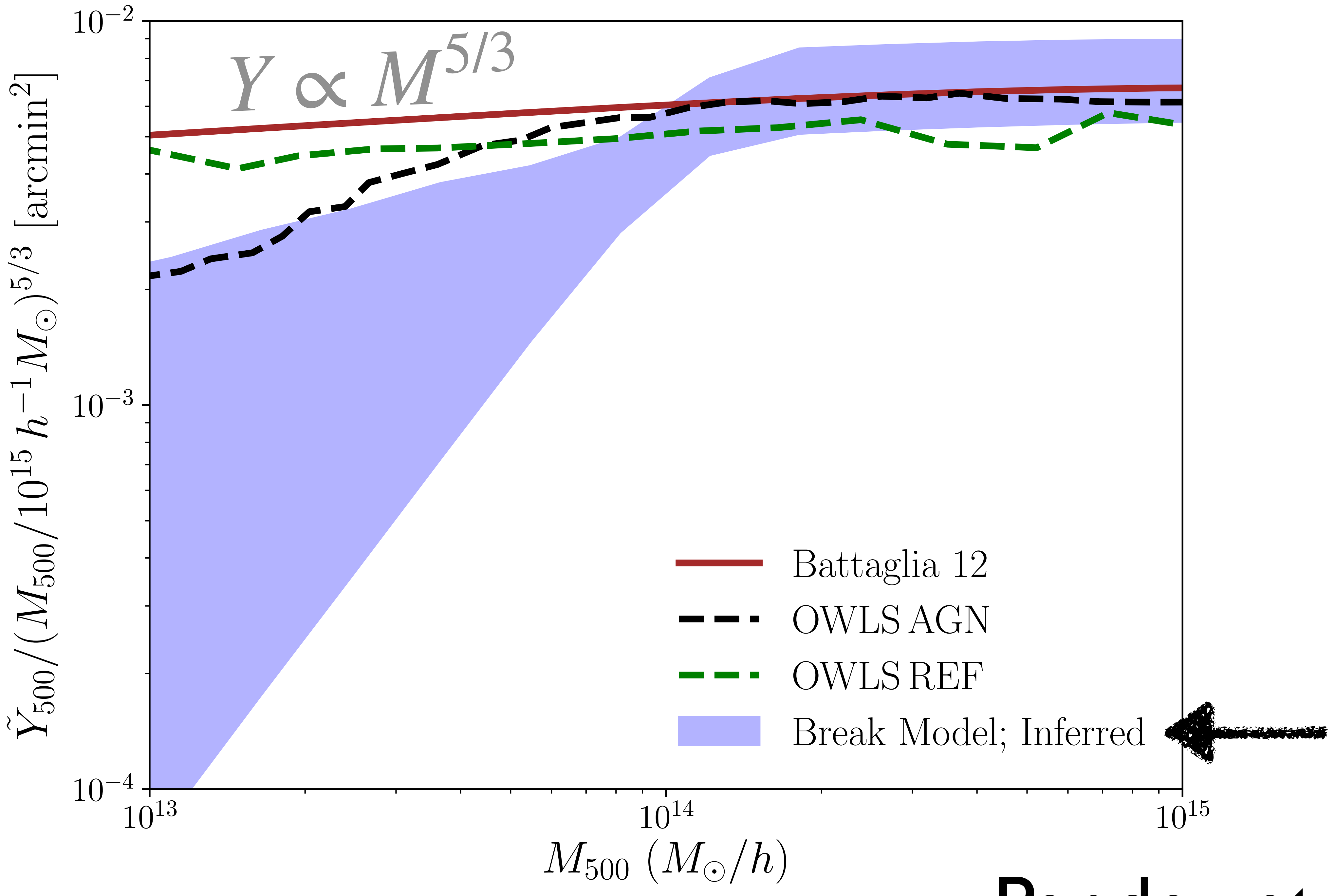
$$Y\left(1+\frac{M_{*}(r<R)}{M_{\rm gas}(r<R)}\right)$$

$$Y\left[1+\frac{M_{*}(r<R/2)}{M_{\rm gas}(r<R/2)}\right]$$



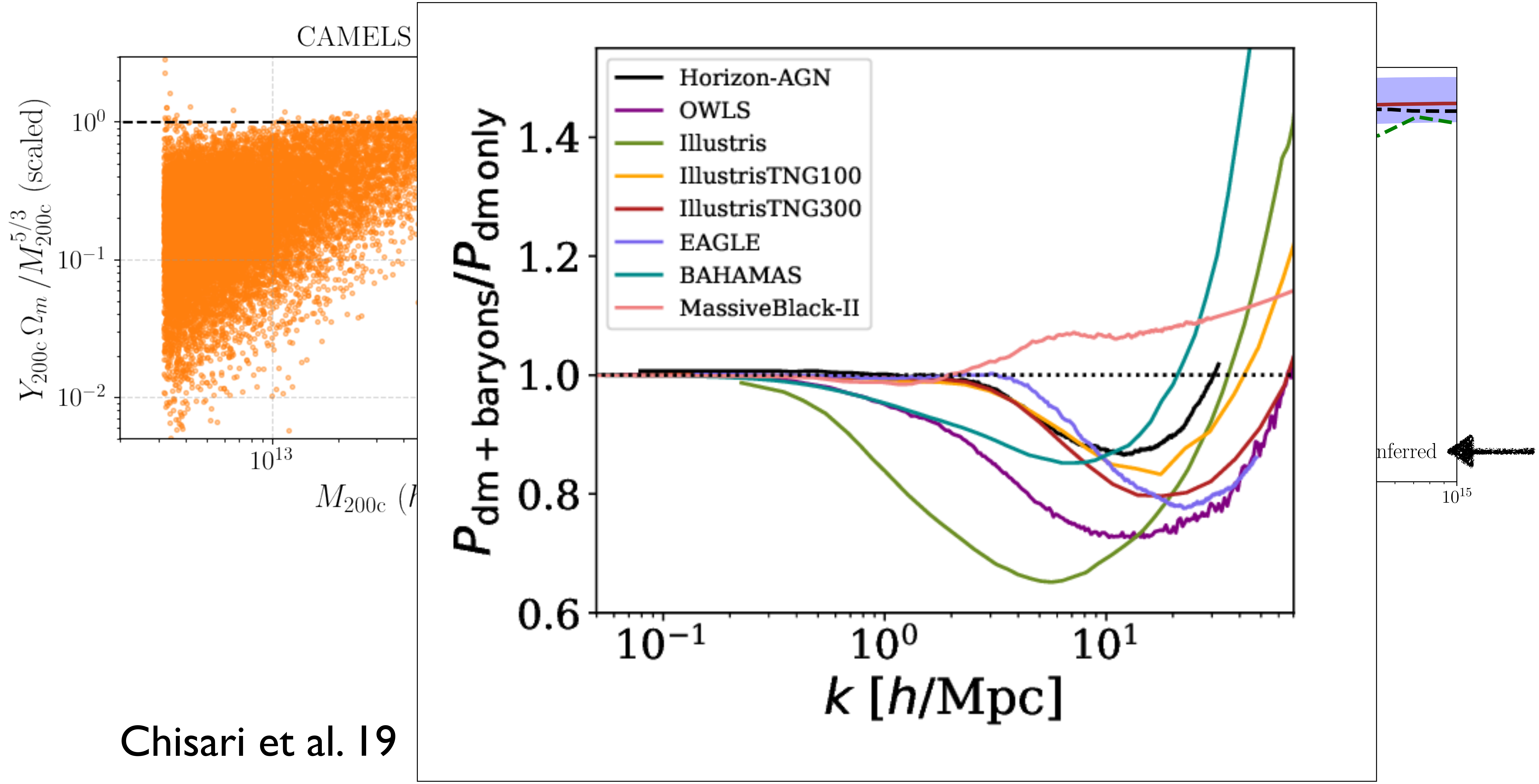
Using the Y-M measurements to constrain baryonic feedback





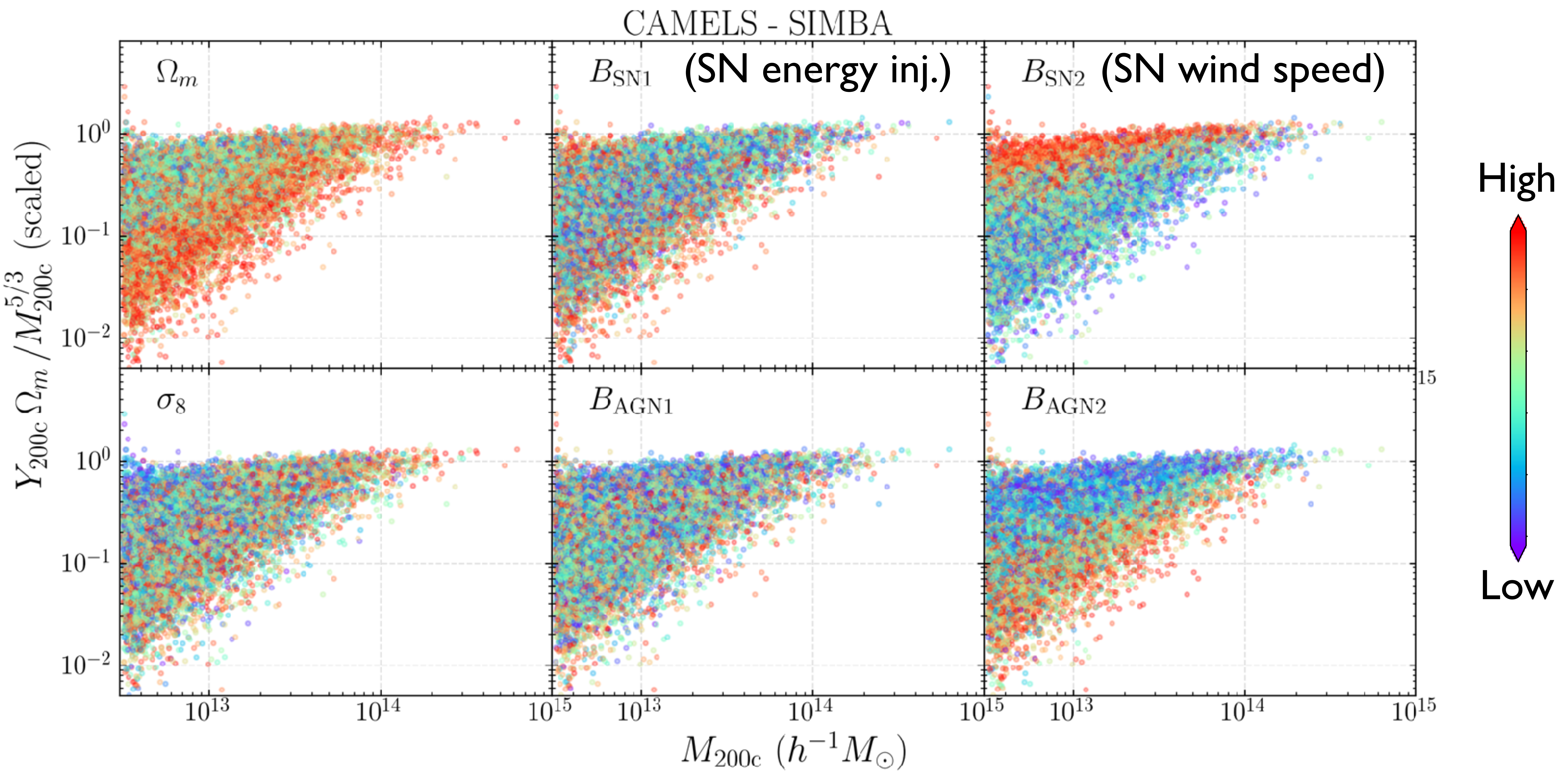
Pandey et al. 21
(ACT x DES)

Using the Y-M measurements to constrain baryonic feedback

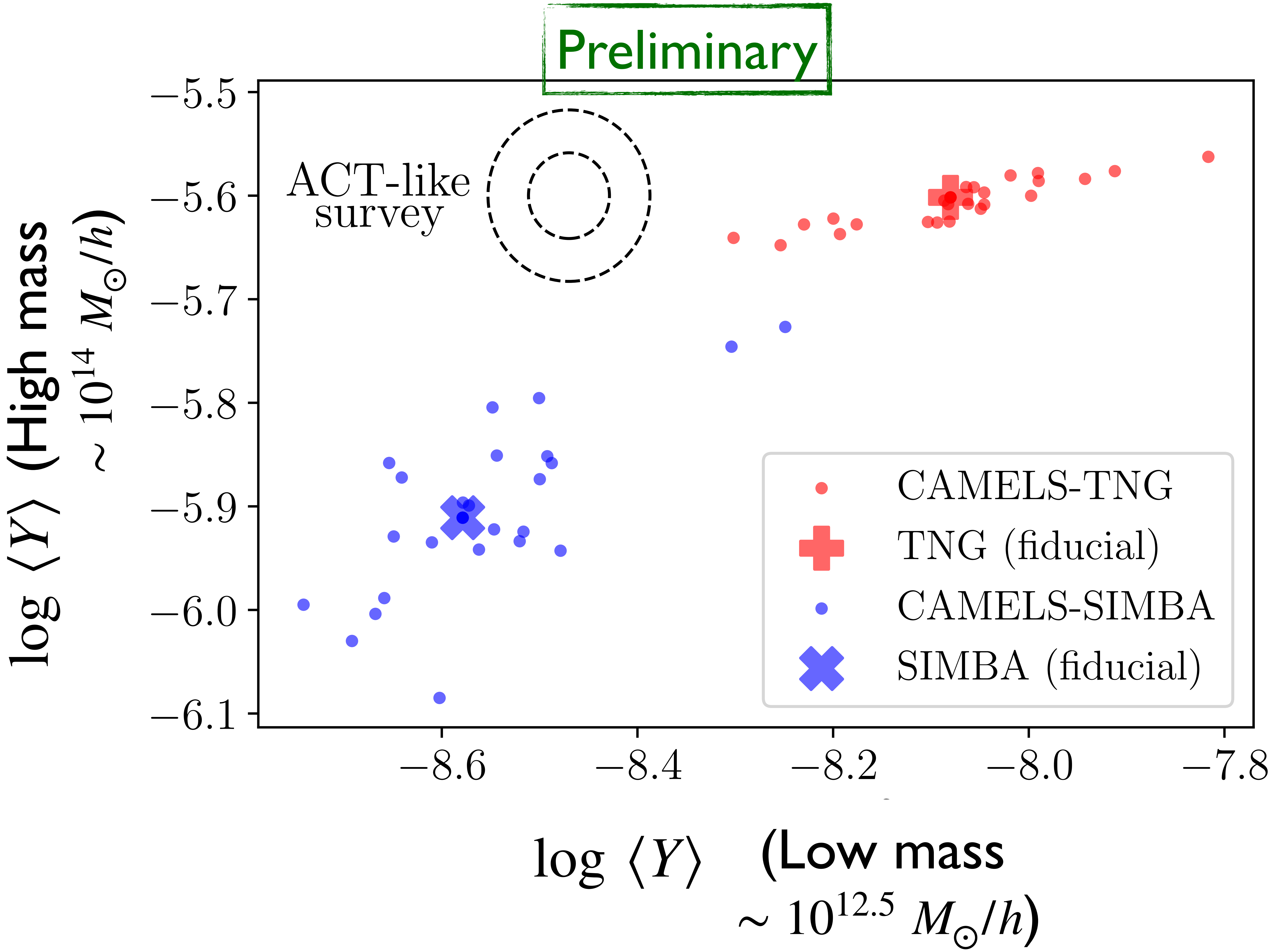


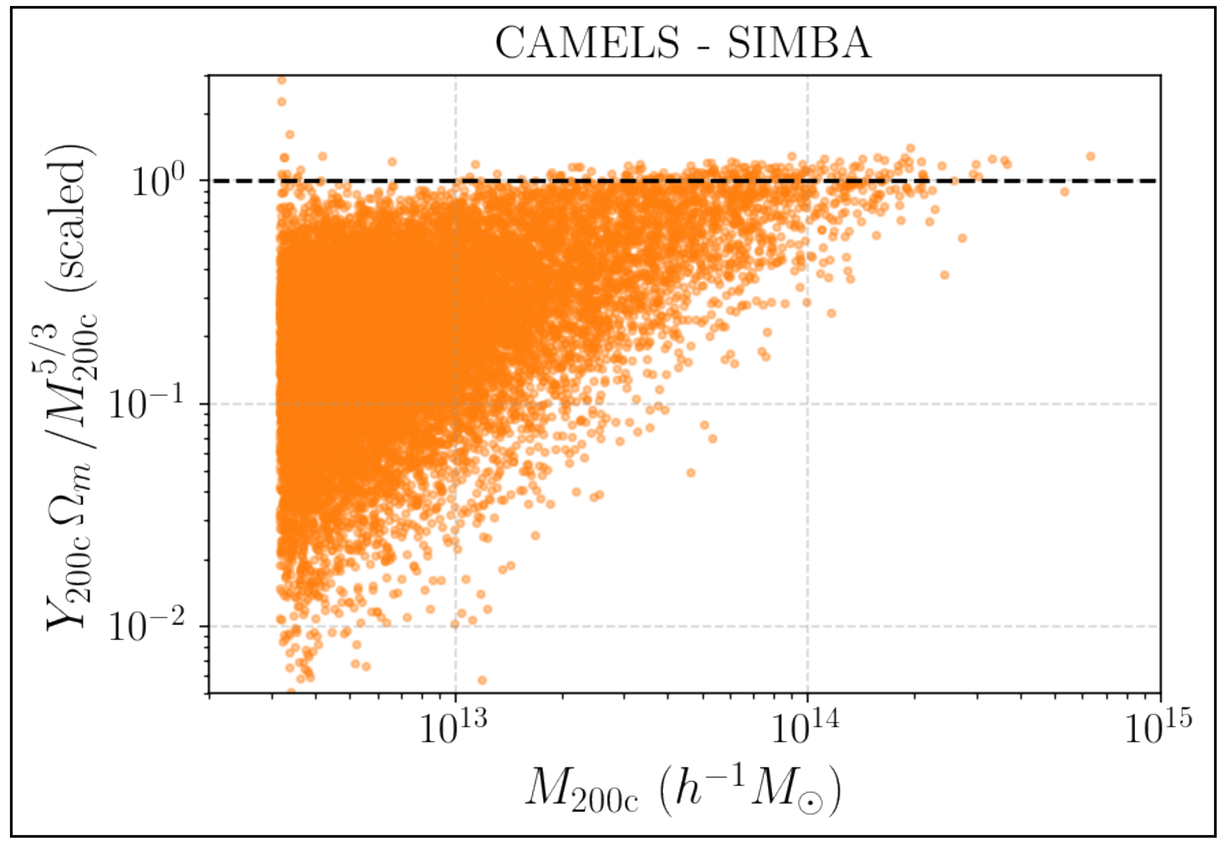
Chisari et al. 19

Y-M for CAMELS sims

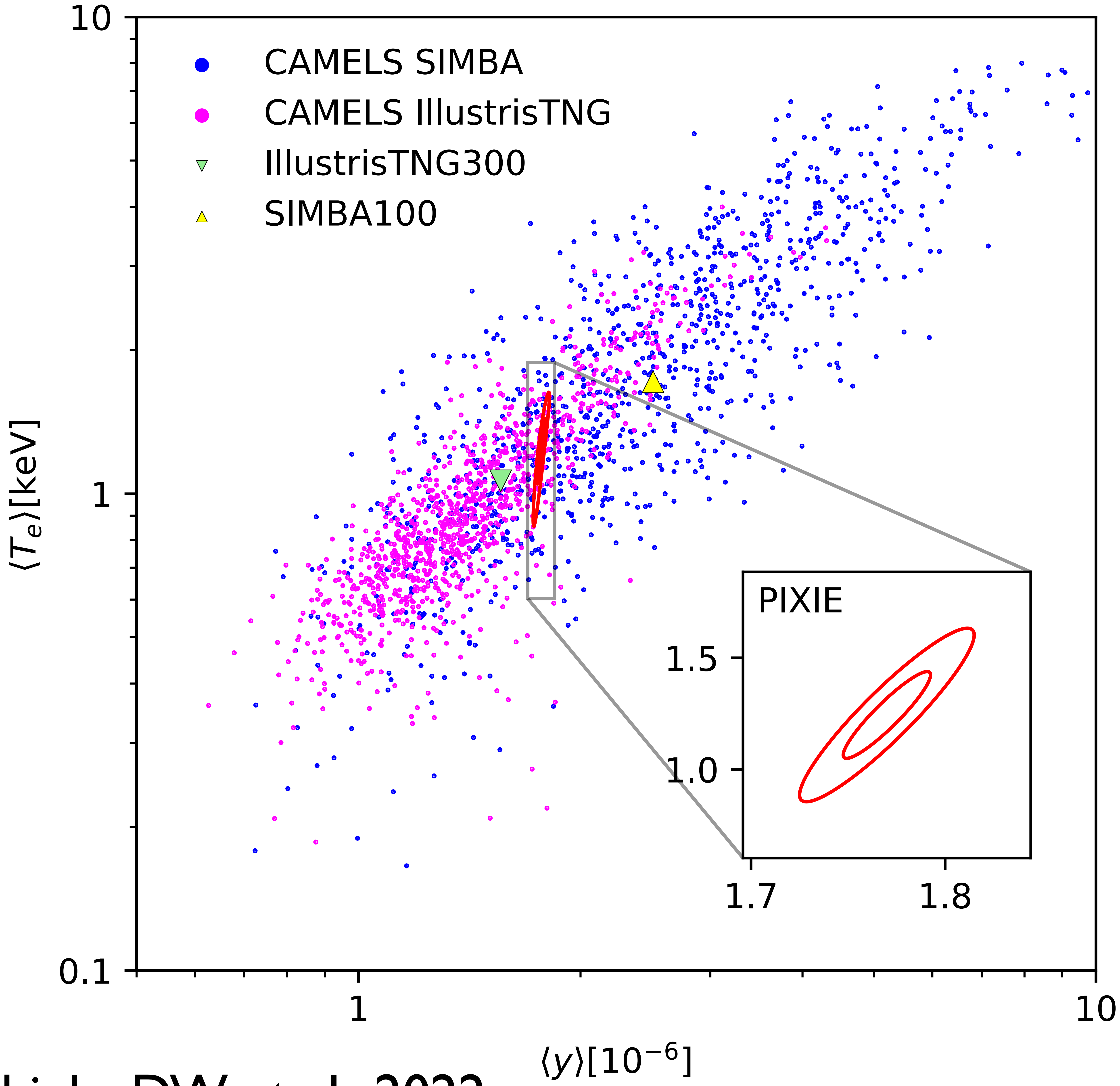


Constraints on sub-grid models



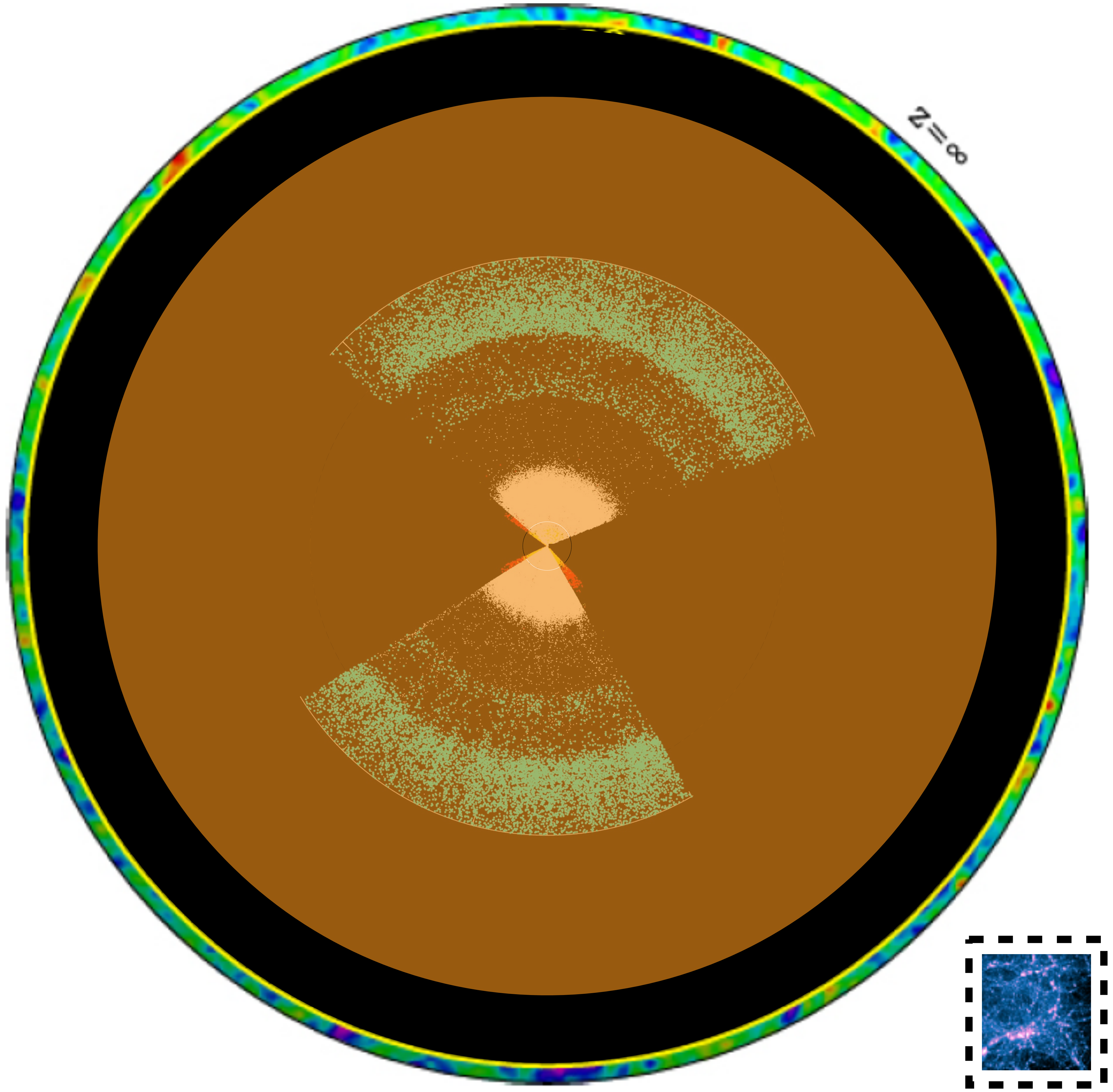


Similarly, CMB spectral distortions can also constrain baryonic feedback (% level constraints)



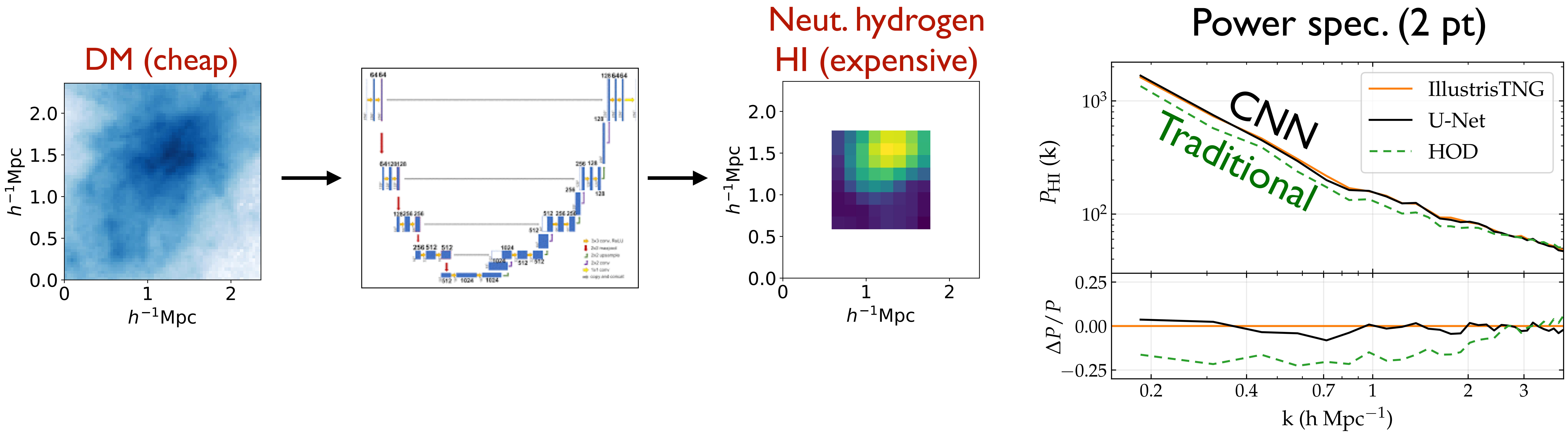
L. Thiele, DW, et al., 2022

ML for emulation of hydro simulations for future surveys



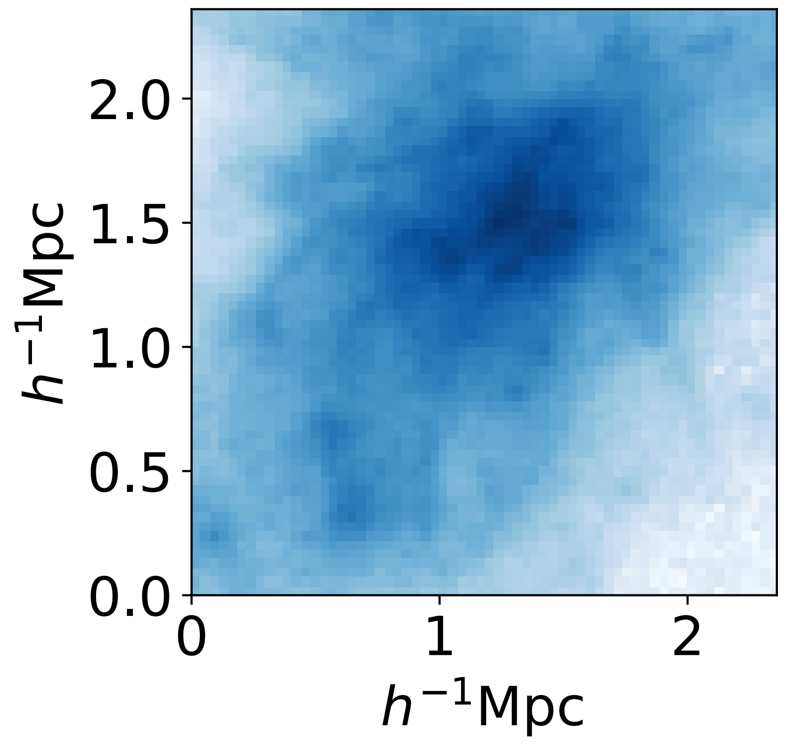
- Volume of upcoming surveys like DESI: ~$\mathcal{O}$ (10-100 Gpc^3)
- Hydro sims are expensive: ~10 million CPU hours for (0.001 Gpc^3)
- Needed to study non-linear scales where baryonic effects dominate

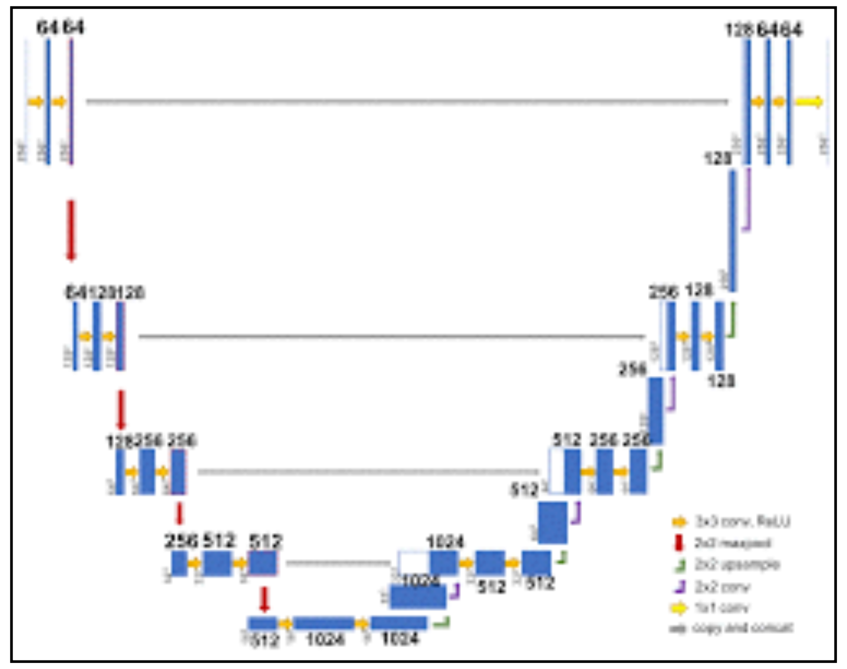
ML for emulation of hydro simulations for future surveys



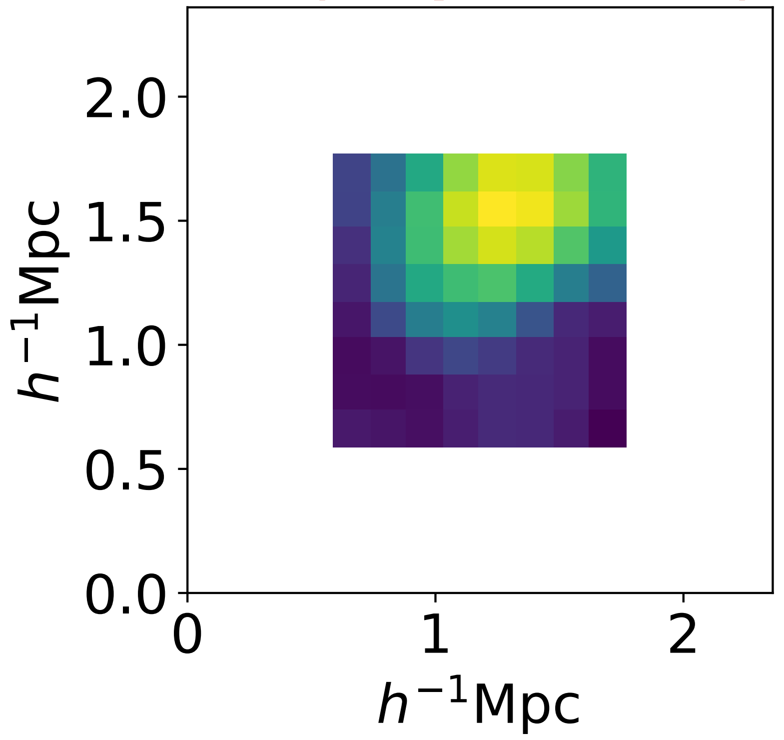
ML for emulation of hydro simulations for future surveys

DM (cheap)

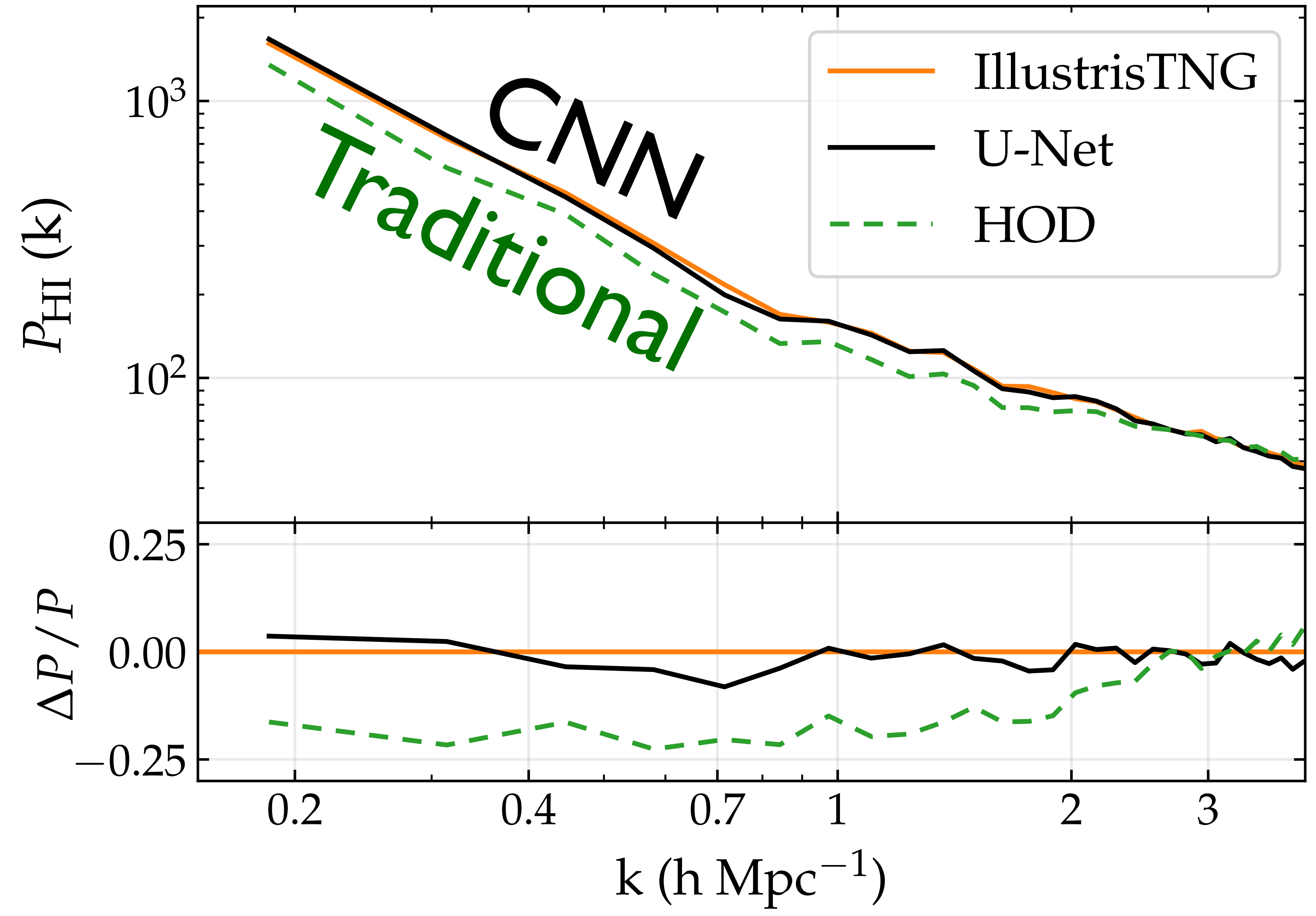




Neut. hydrogen
HI (expensive)



Power spec. (2 pt)



N-body → Galaxies

N-body → N-body + Neutrinos

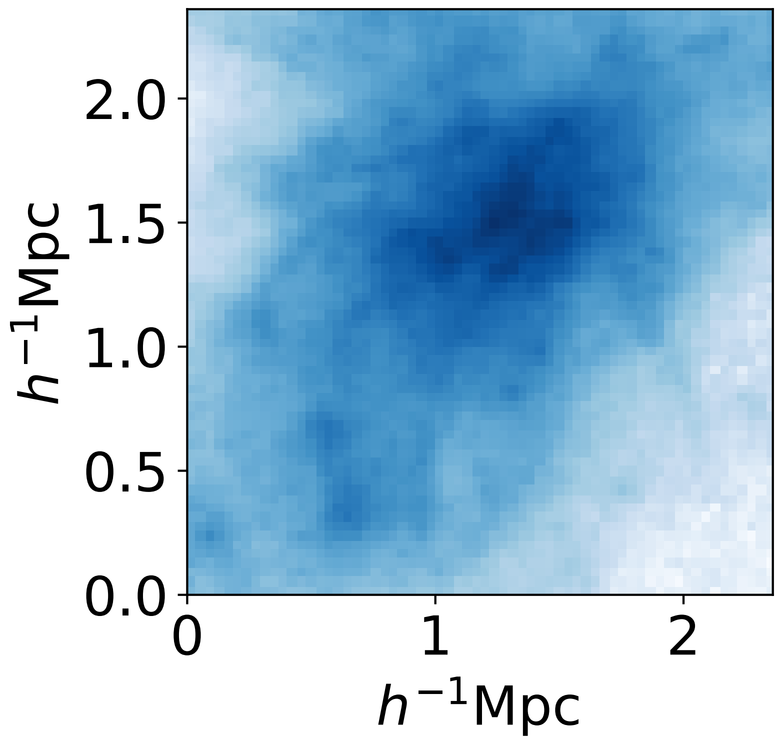
ZA (theoretical) → N-body

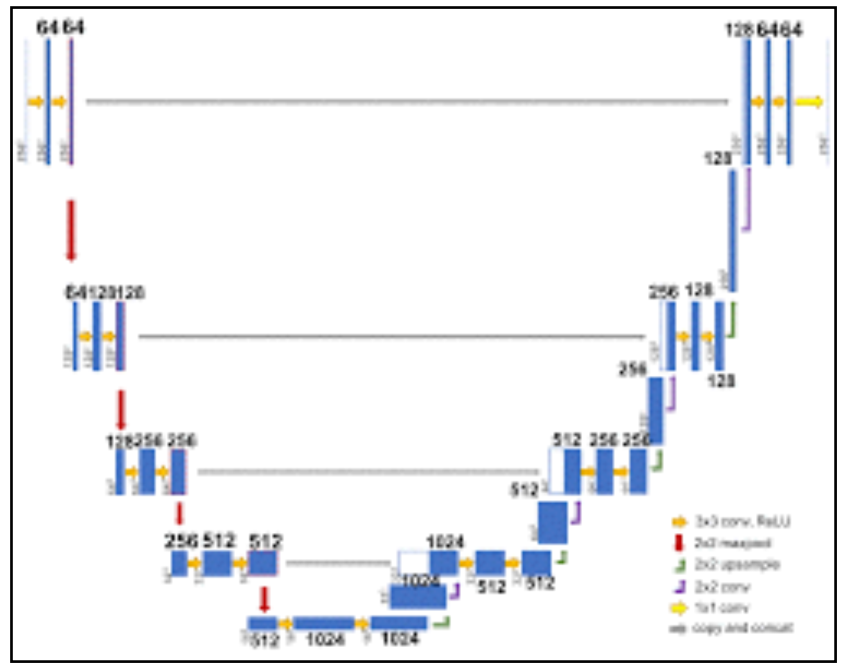
Sims/Data → Cosmo. parameters

Low res. N-body → High res. N-body

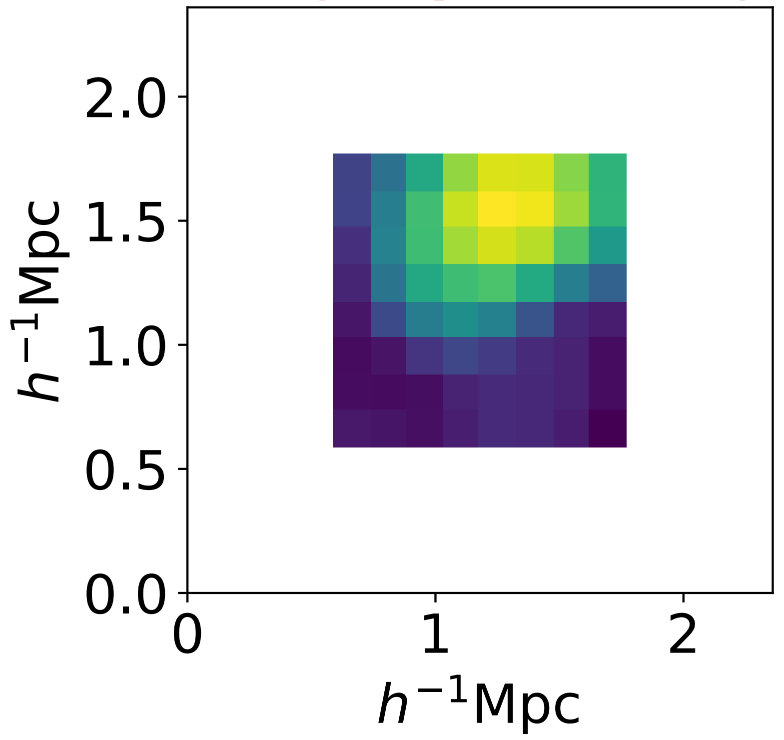
ML for emulation of hydro simulations for future surveys

DM (cheap)

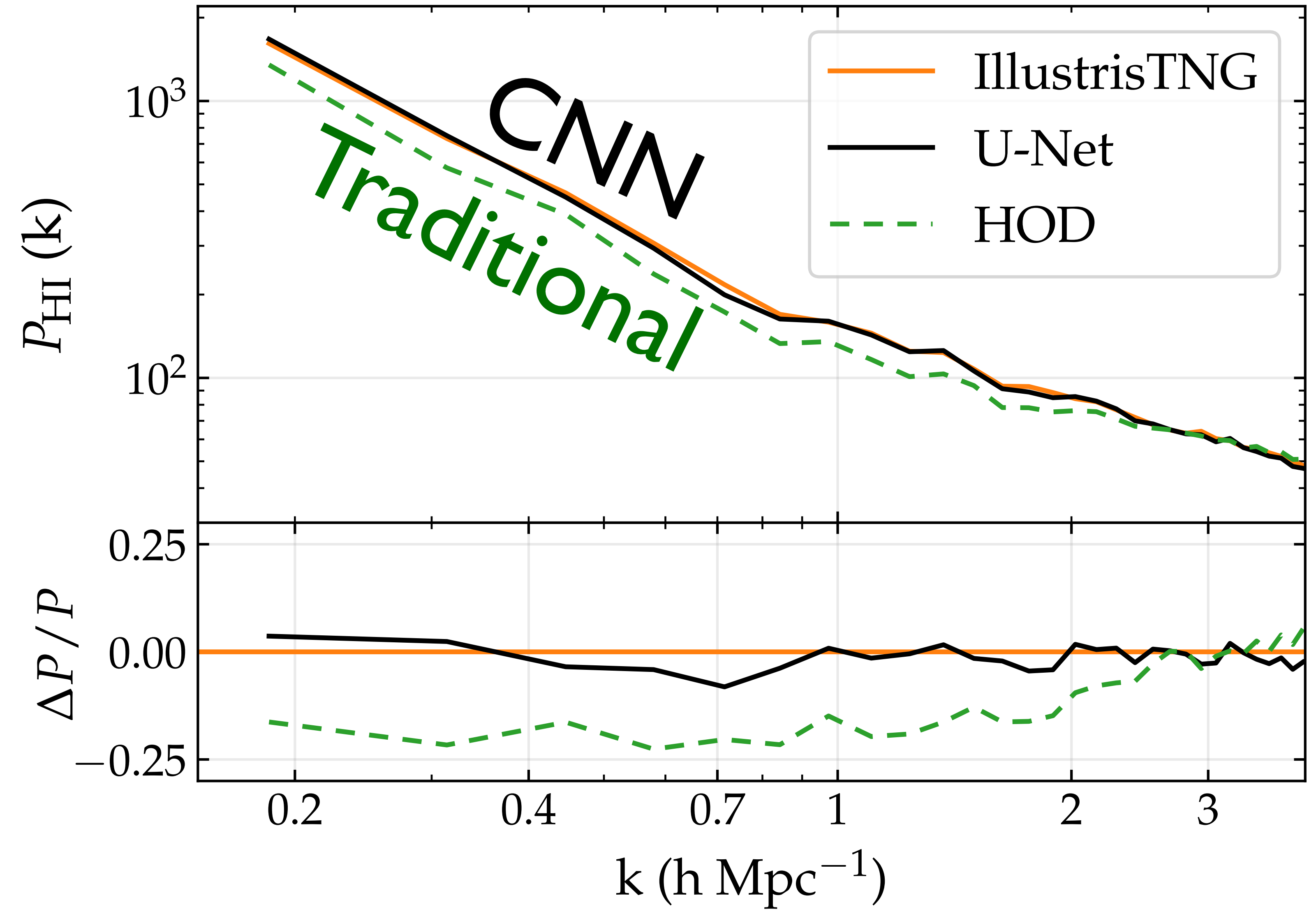




Neut. hydrogen
HI (expensive)



Power spec. (2 pt)



N-body → Galaxies

N-body → N-body + Neutrinos

ZA (theoretical) → N-body

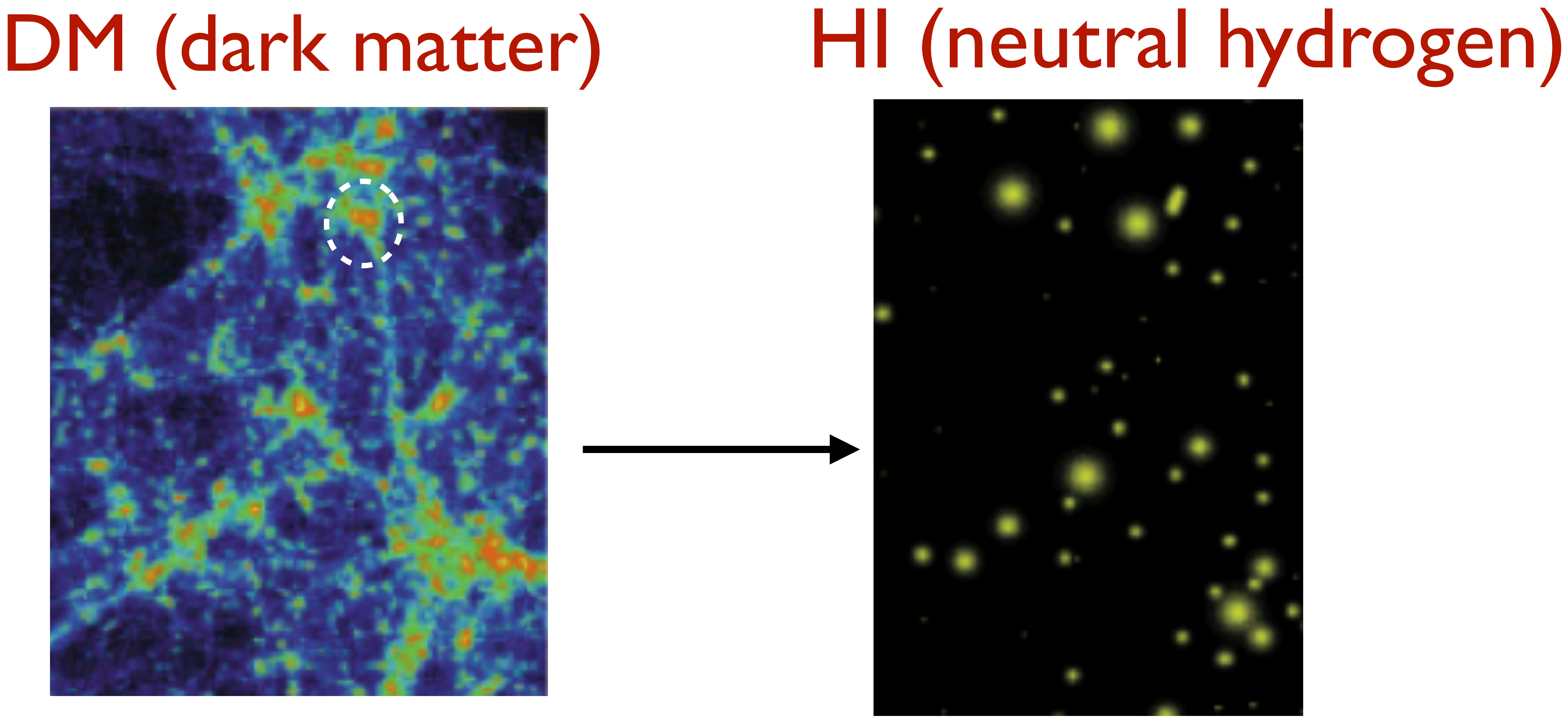
Sims/Data → Cosmo. parameters

Low res. N-body → High res. N-body

Challenges:

1. Robustness to feedback prescriptions
2. Robustness to sim resolution
3. Robustness to observational systematics

ML to model assembly/secondary bias

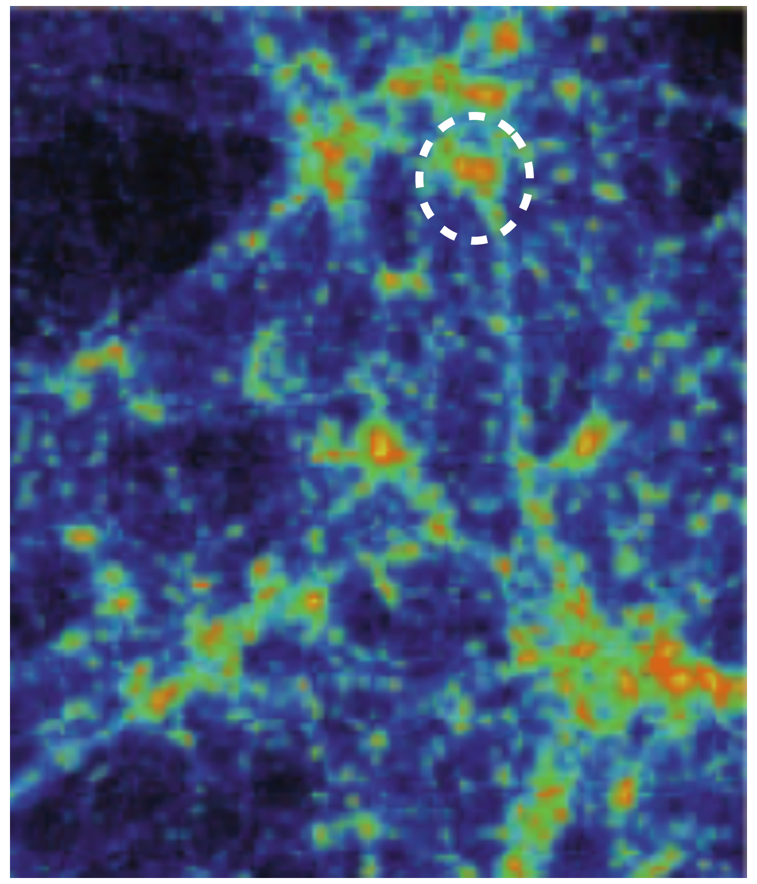


HI mass of halo = f (Halo mass only)

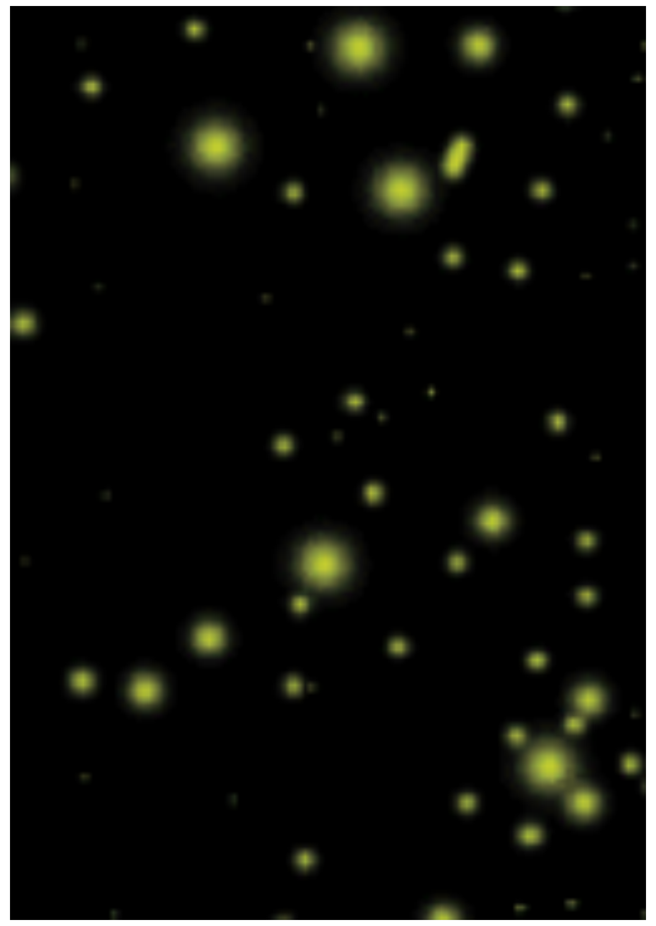
(No. of galaxies in halo)

ML to model assembly/secondary bias

DM (dark matter)



HI (neutral hydrogen)



HI mass of halo = f (Halo mass, *secondary props. ?*)

(No. of galaxies in halo)

{local env.,
conc.,
shear,….}

ML to model assembly/secondary bias

No. of galaxies in a halo $= f$ (Halo mass, *environmental shear and overdensity*)

$$N_{\rm sat}(M_h) = N_{\rm sat}^{\rm HOD}(M_h) \times (q' + A)$$

A. Delgado, DW, et al. 21

$$N_{\rm cen}(M_h) = N_{\rm cen}^{\rm HOD}(M_h) \times \left[1 + B(\delta'_{\rm env} - \overline{\delta'_{\rm env}})(1 - N_{\rm cen}^{\rm HOD})\right]$$

Neutral hydrogen content of halo $= f$ (Halo mass, *environmental shear and overdensity*)

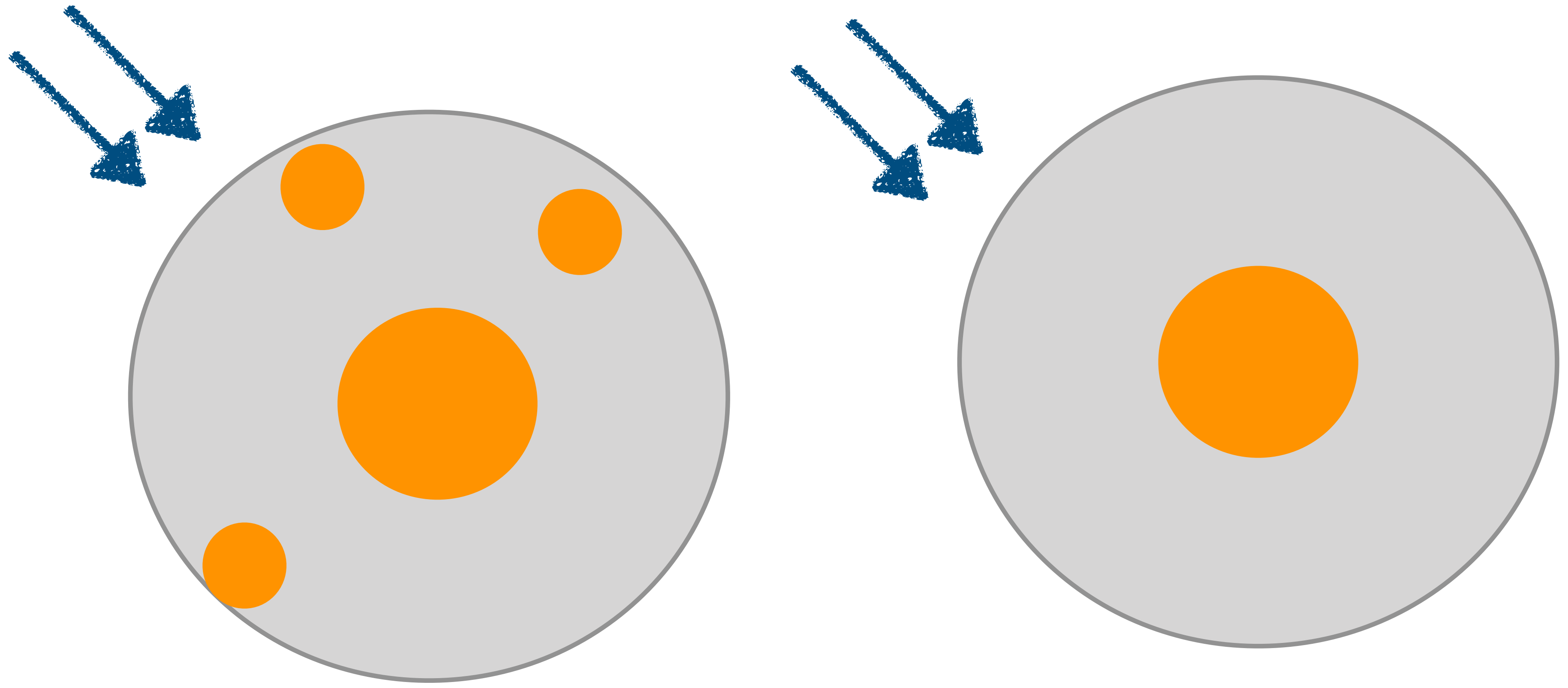
DW et al. 20

$$\frac{M_{\rm HI}}{M_{\rm HOD}} = 0.81 + 1.44\,\alpha'_{0.5}\,m_{10} - 0.57\,(\alpha'^2_{0.5}\,m_{10}^2 + \alpha'_{0.5}\,\delta'_5)$$

Why does HI content have env. dependence?

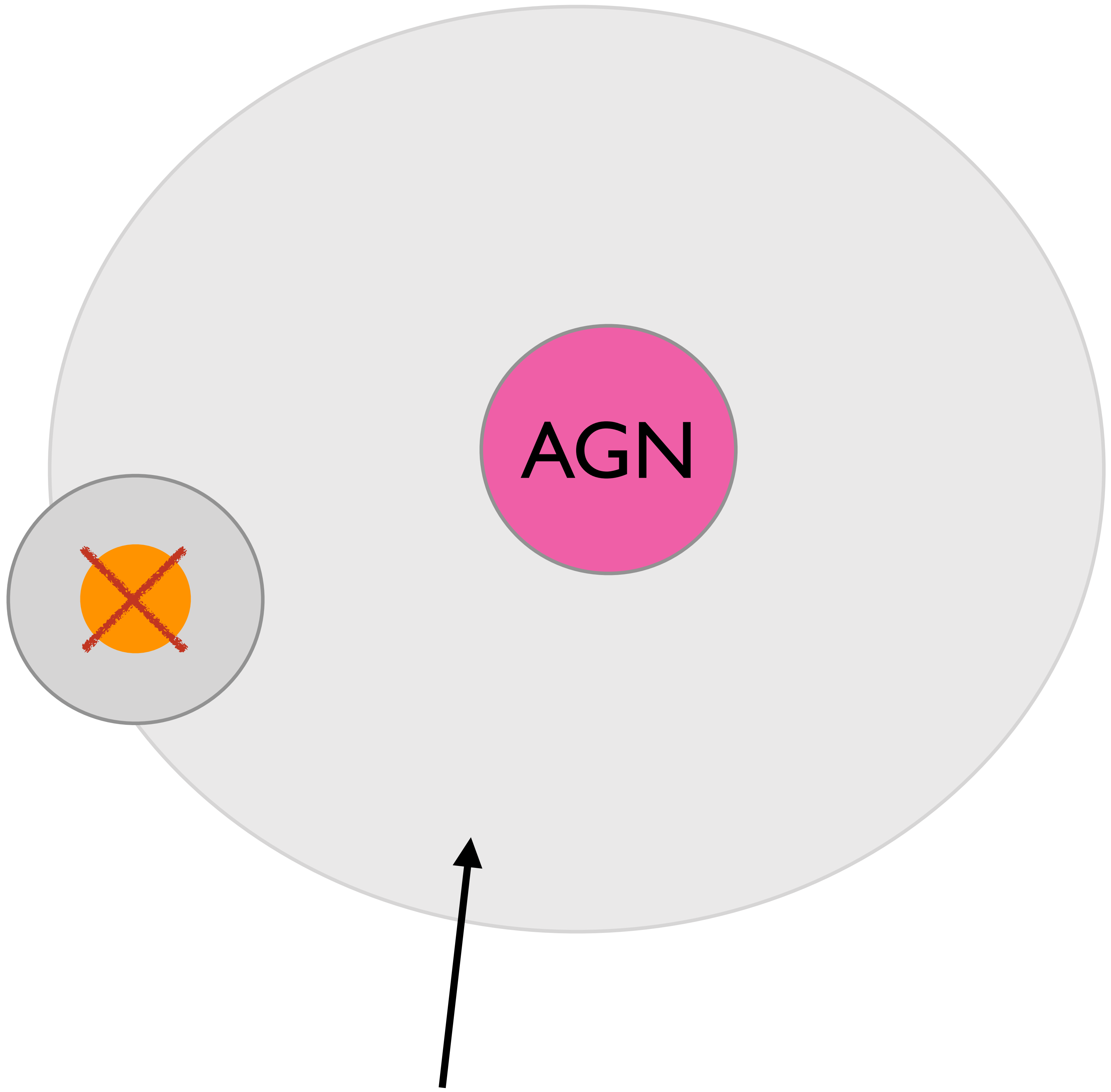
$$\frac{M_{\mathrm{HI}}}{M_{\mathrm{HOD}}} = 0.8 + 1.4\,\alpha'_{0.5} m_{10} - 0.6\,(\alpha'^{2}_{0.5} m^{2}_{10} + \alpha'_{0.5}\delta'_{5})$$

UV + X-ray background



Denser env.

→ more mergers



Denser env.: Ionized medium

(ram pressure stripping)

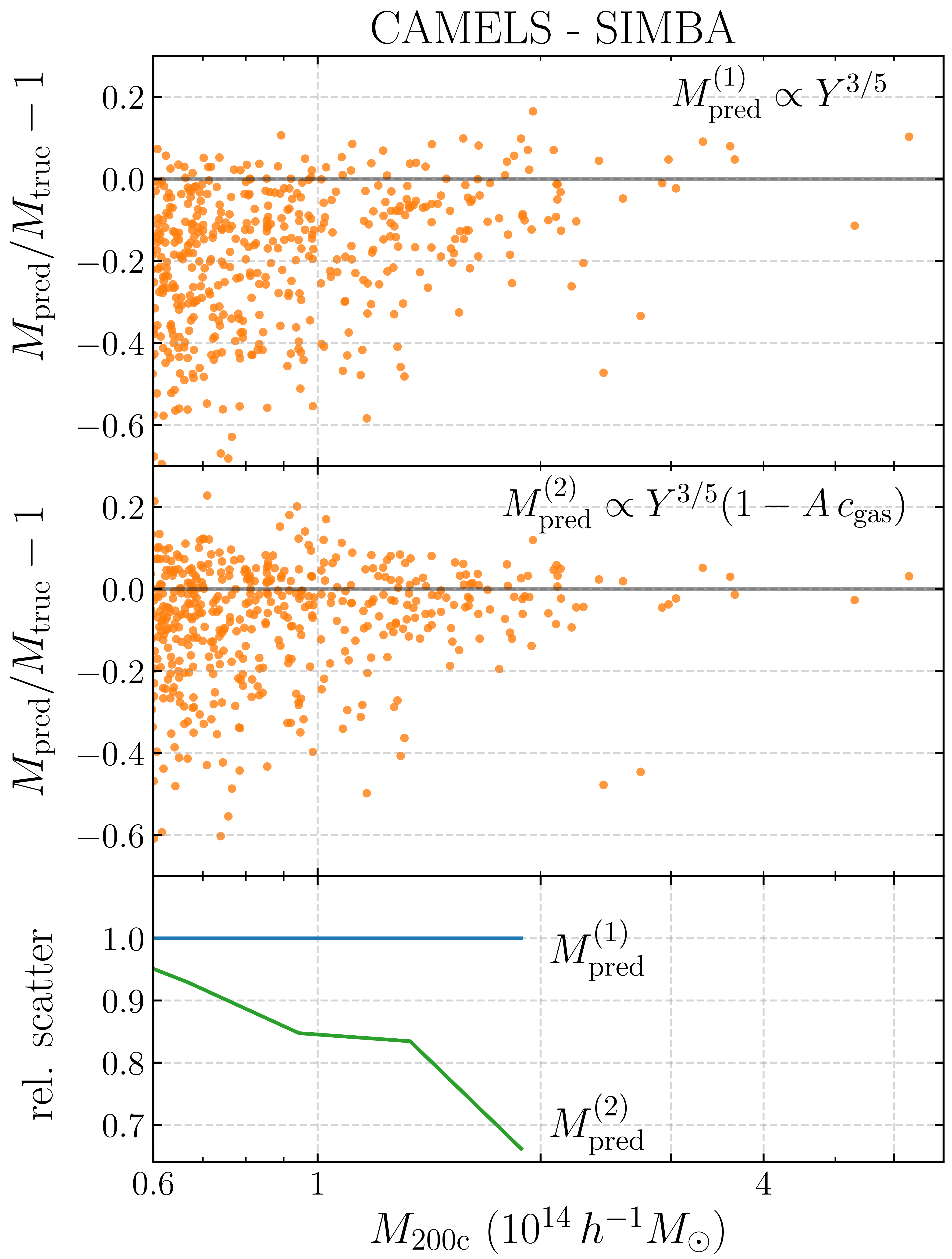
Summary

★ Symbolic regression can be used to *augment astrophysical scaling relations* and increase their precision

- Using gas conc. reduces scatter in SZ mass estimates by 20-30% for large clusters

- Including stellar to gas mass ratio reduces deviation from self-similarity by factor >2

➡Suggestions for other scaling relations?



Application to other scaling relations?

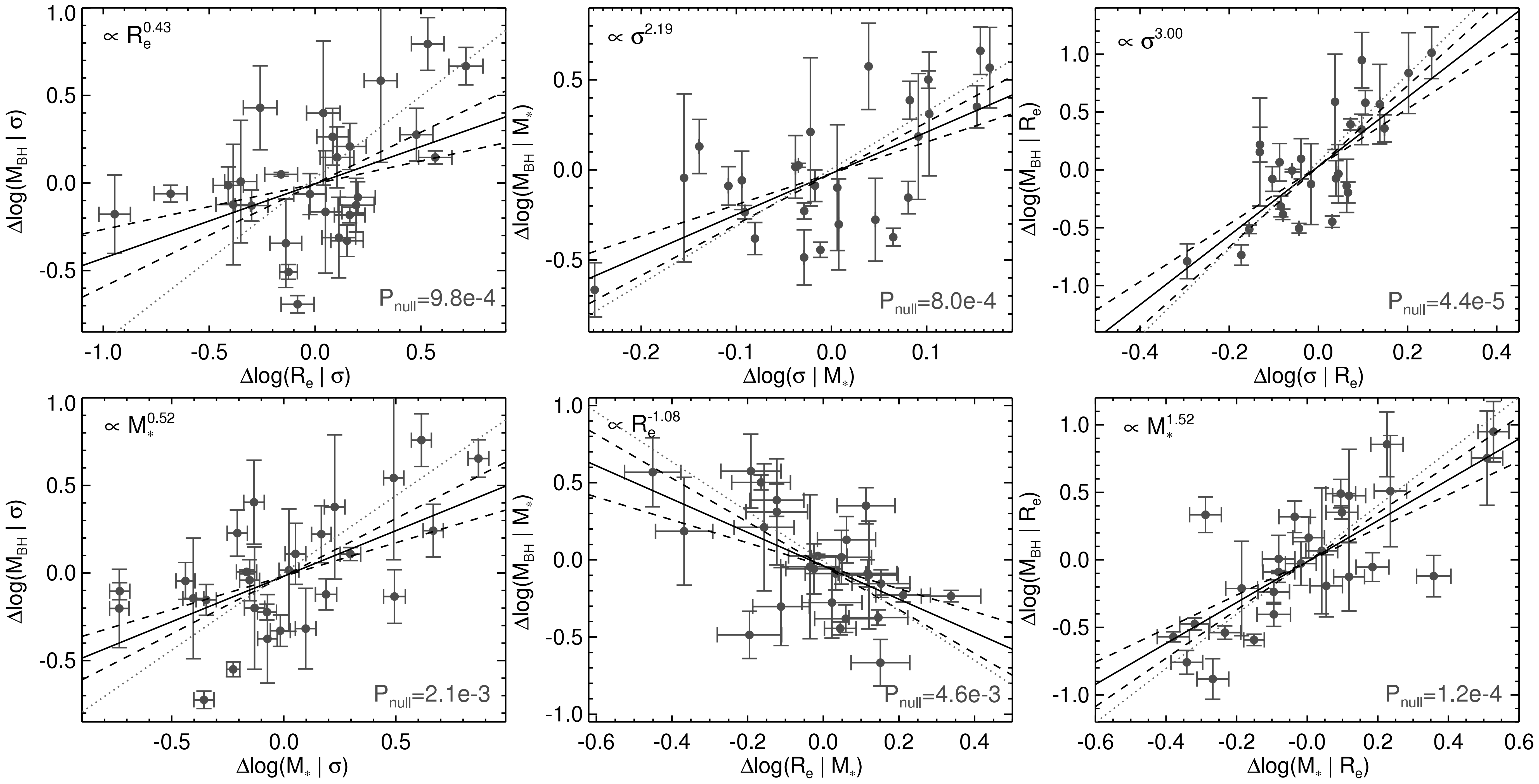
- Philips relation for supernovae

$$M_{\max}(B) = -21.726 + 2.698\,\Delta m_{15}(B)$$

- Cepheid P-L relation

$$M_{\mathrm{v}} = A(\log_{10}P - 1) - B$$

- Tully fisher relation
- Black hole-bulge mass relation
- Fundamental plane relation
- ….



Hopkins et al. 07