Improving astrophysical scaling relations with interpretable ML

(Jay) Digvijay Wadekar

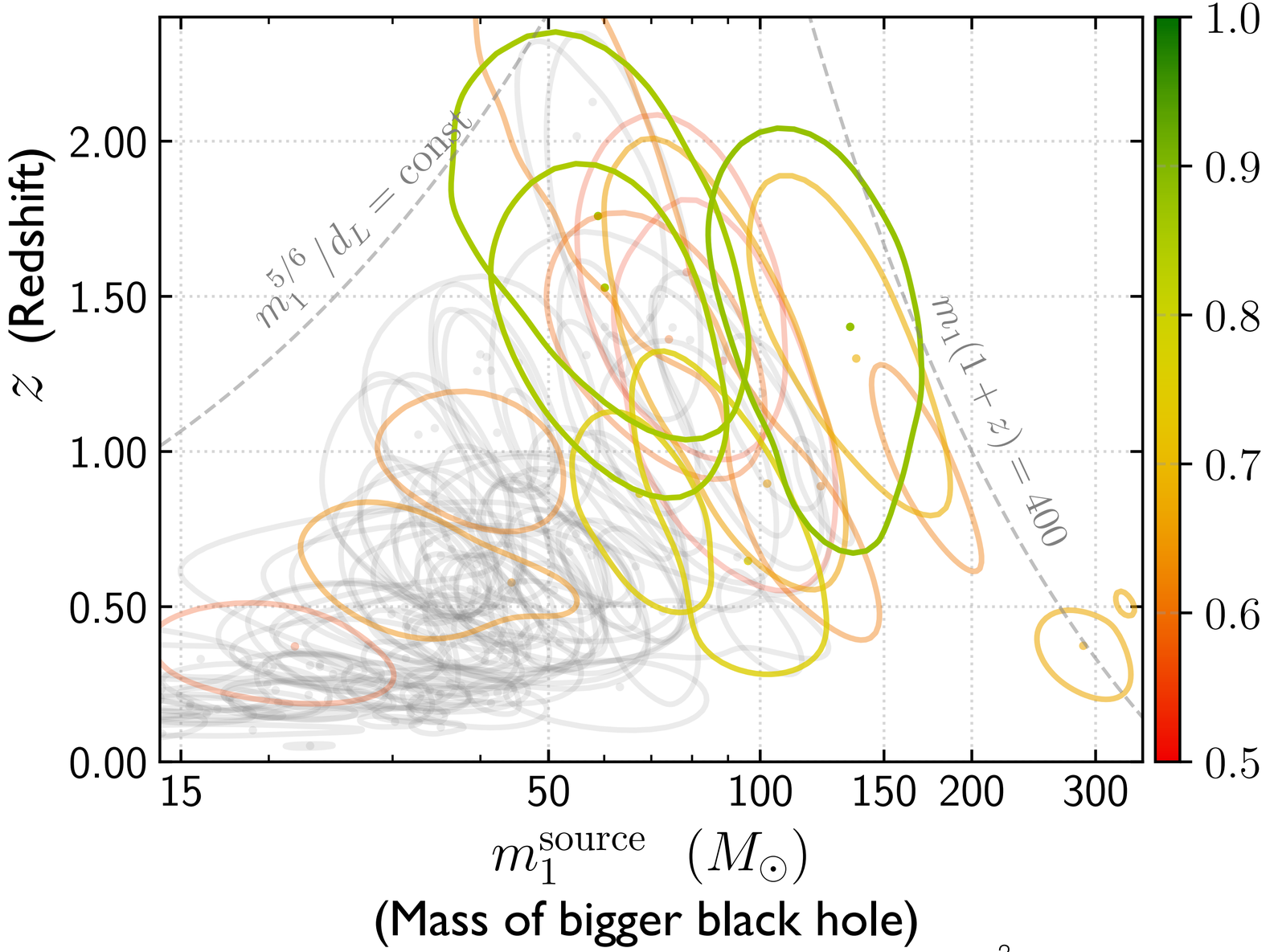
Johns Hopkins U

with CCA cosmology group

(L. Thiele, J.C. Hill, F. Villaescusa-Navarro, D. Spergel, M. Cranmer, S. Pandey, N. Battaglia, S. Ho, D. Nagai, L. Hernquist)

arXiv:2201.01305 & 2209.02075 & in prep

Gravitational wave data-analysis (IAS pipeline)



Gray: Older detections (LIGO-Virgo catalog)

Colored: New detections from our new search

Probability of being of astrophysical origin

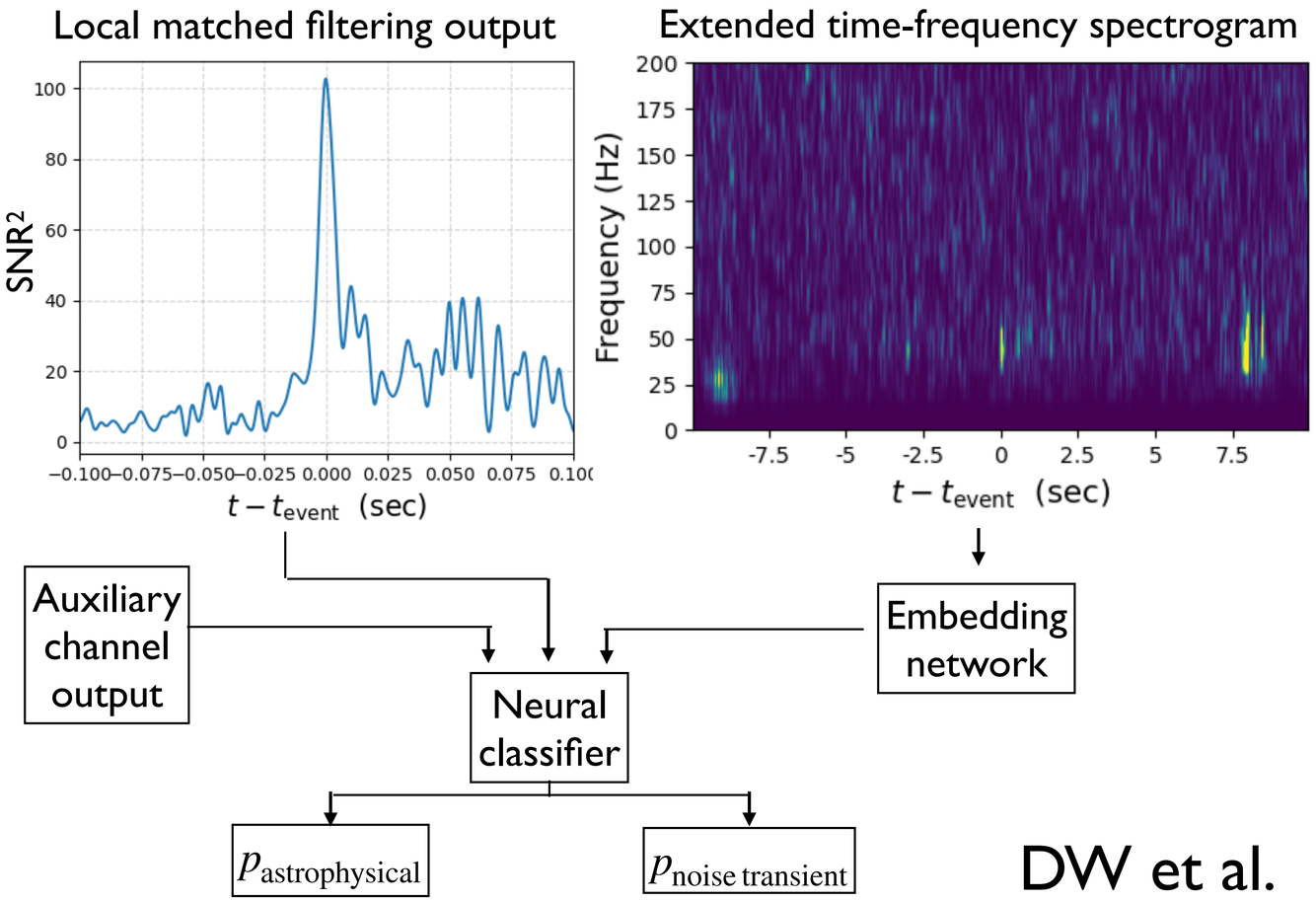
DW et al. 2312.06631

Enhancing a GW search pipeline with ML

1. Autoencoders for generating efficient template banks

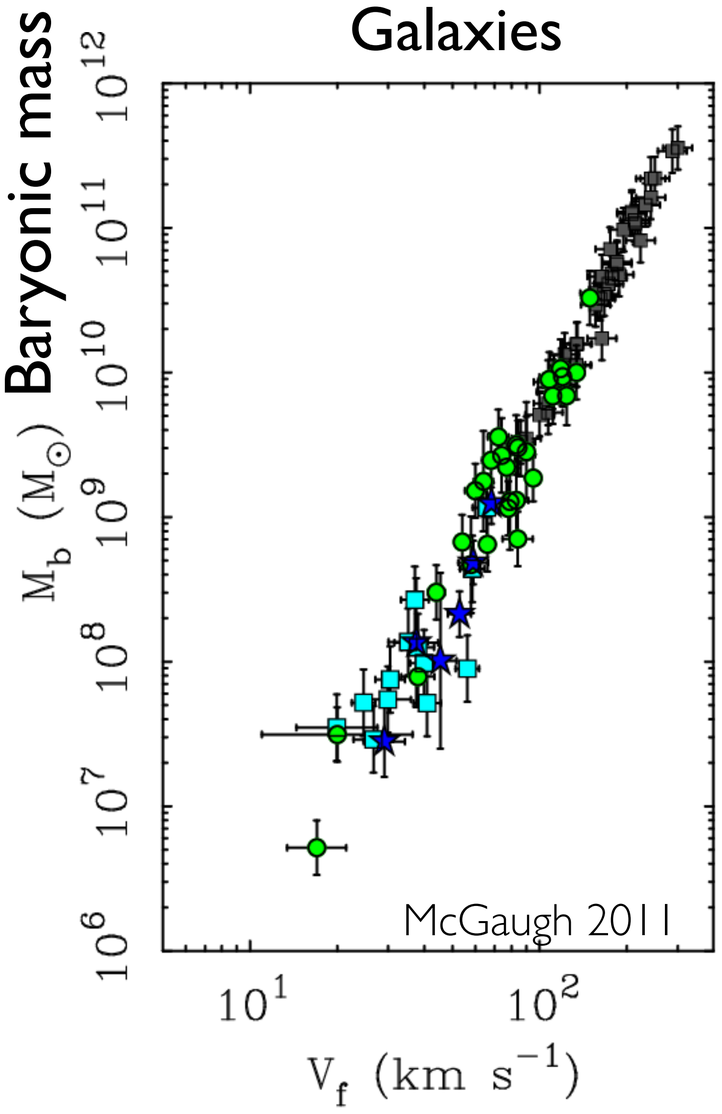
2. Marginalized detection statistic using normalizing flows

3. CNN + Attention for glitch mitigation



DW et al. in prep

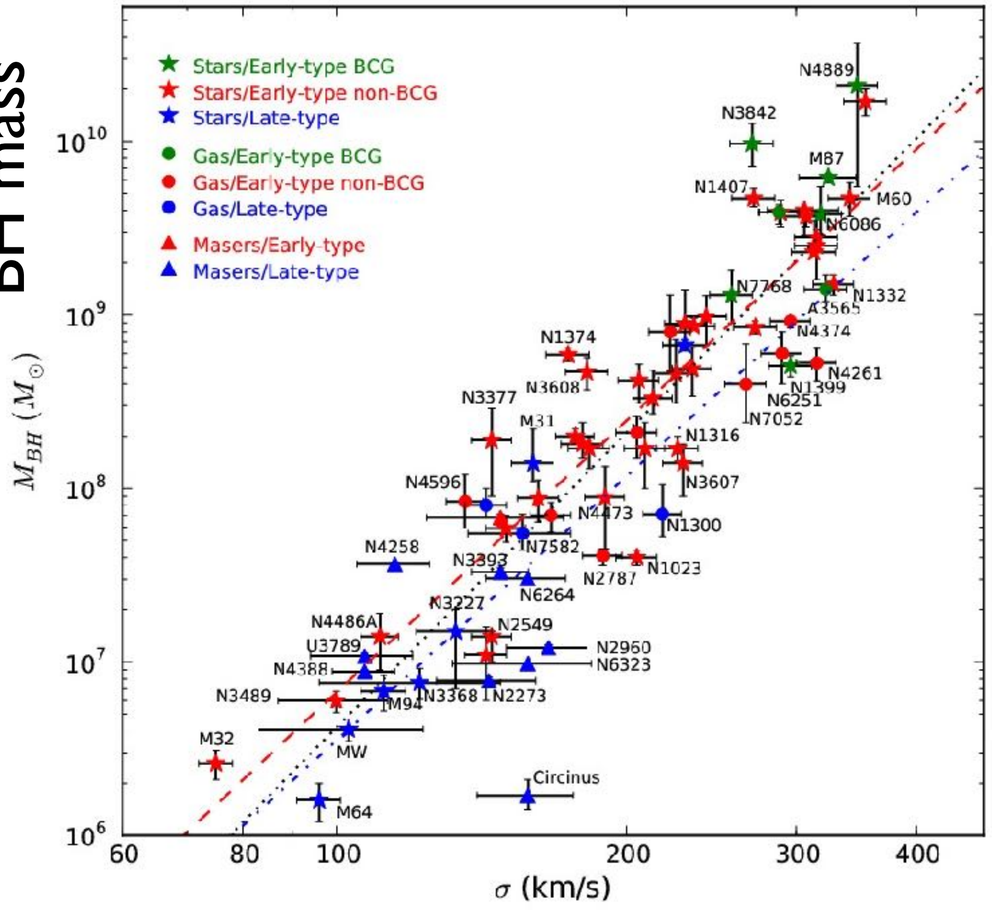
What are scaling relations?



- Low-scatter relations between properties of complex astrophysical systems
- Often found empirically in observational/ simulation data
- Often found by fitting power laws to 2D data

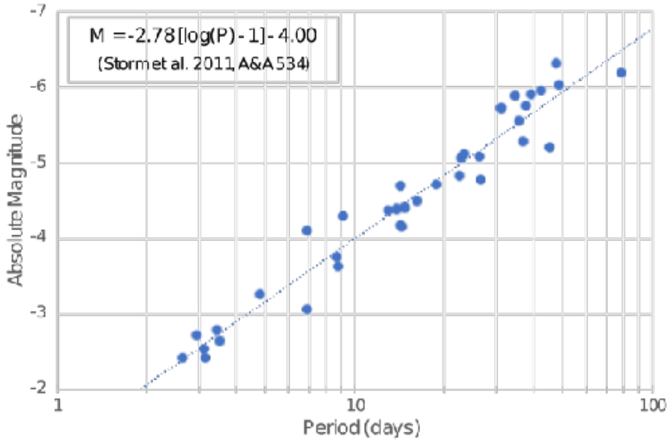
A few popular scaling relations

BH mass

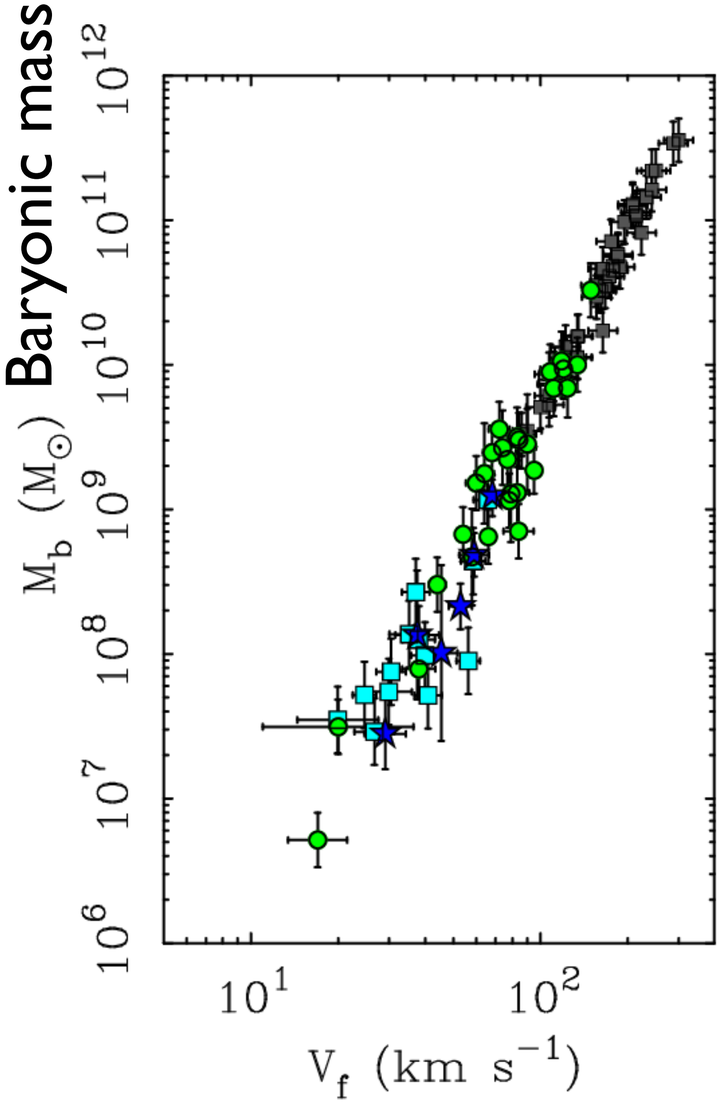


Bulge vel. dispersion

Cepheid period-luminosity relation

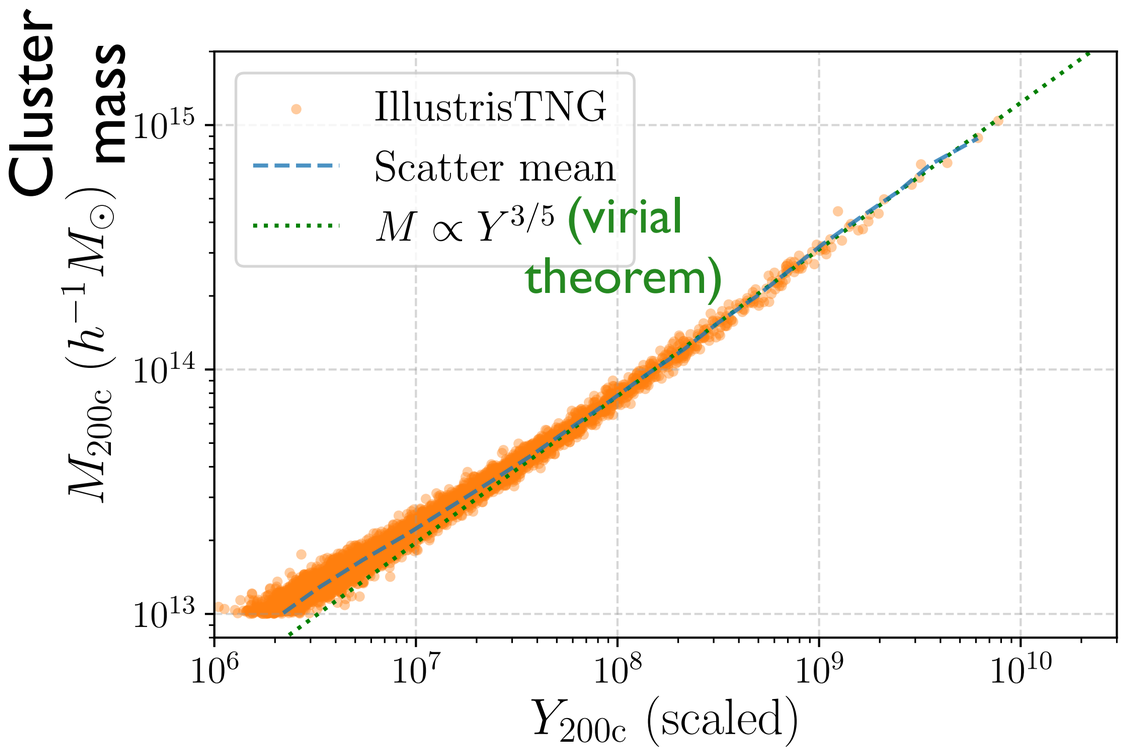


Baryonic mass



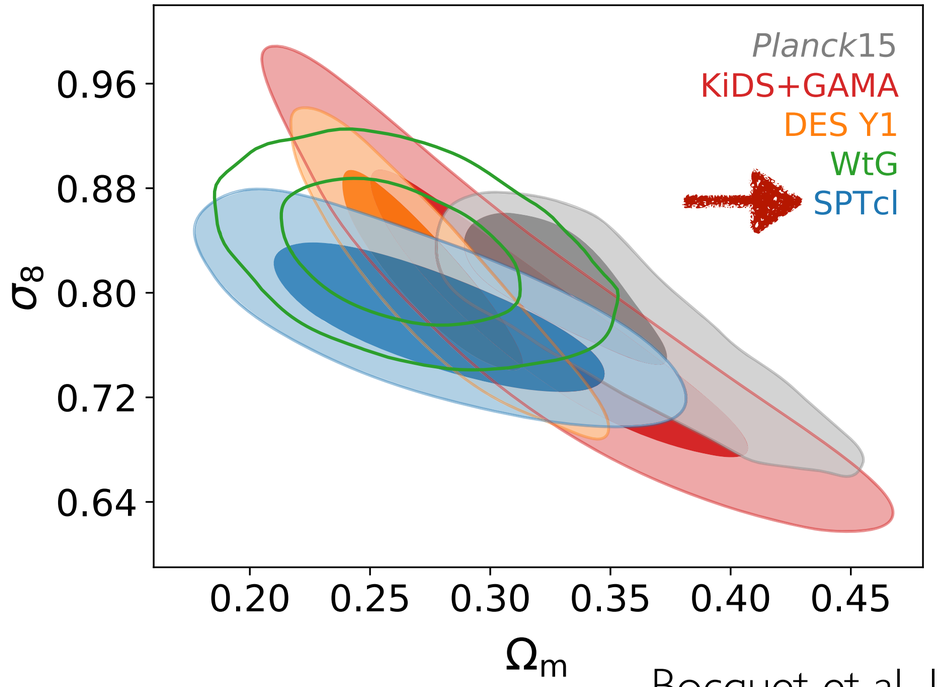
Rotation velocity

Y-M relation is important for cluster cosmology



$$Y_{\rm CMB} \propto M_{\rm gas} T_{\rm gas}$$

(thermal energy of ionized gas)



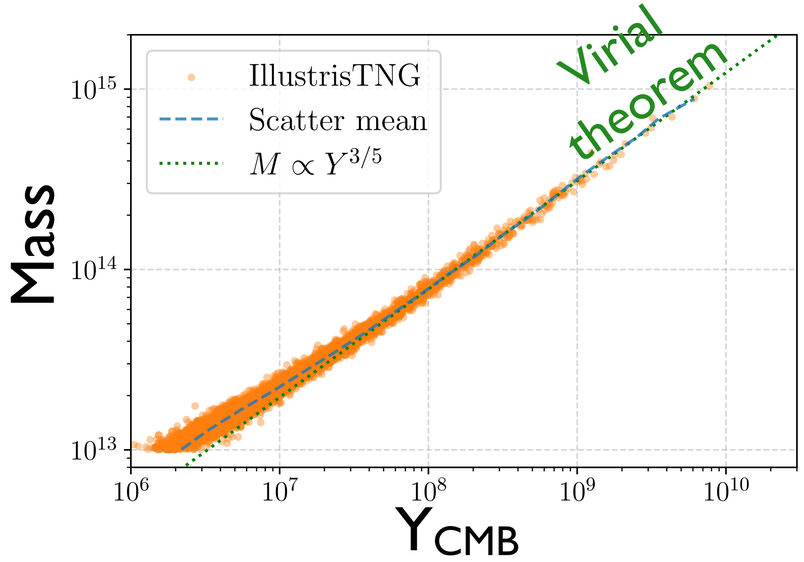
Bocquet et al. 19

Most scaling relations were obtained
by fitting curves (mostly power laws) to 2D data

Can we find relations with
better performance in higher dim. spaces
with ML tools?

Problem statement

$$M_{\rm cluster} = f\left(Y^{3/5}_{\rm CMB} \text{ , other observables?? } \right)$$

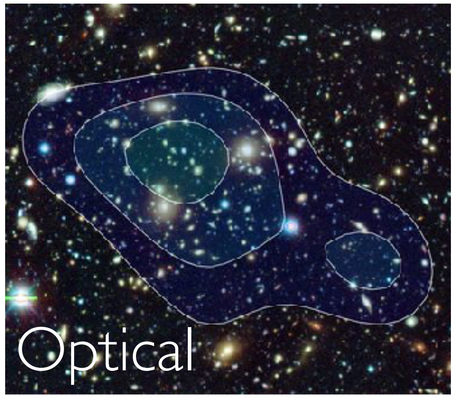


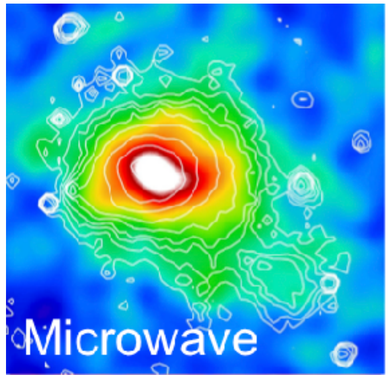
- X-ray/CMB surveys

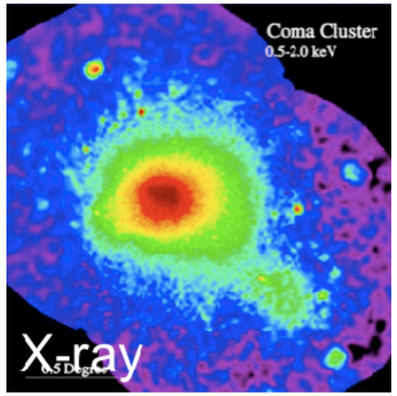
- Gas mass/pressure profile
- Luminosity profile
- Spectral temperature
- Gas ellipticity
- ……

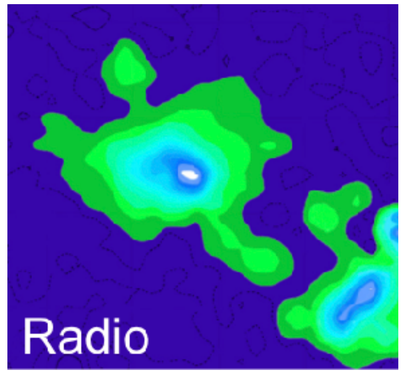
- Galaxy surveys

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

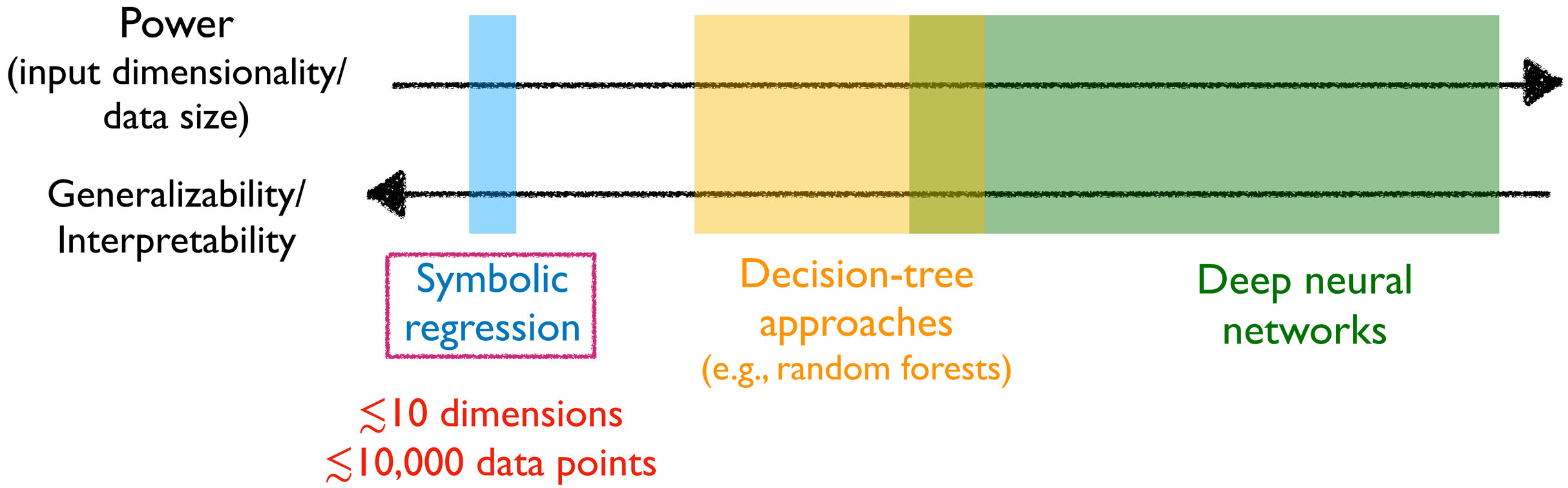








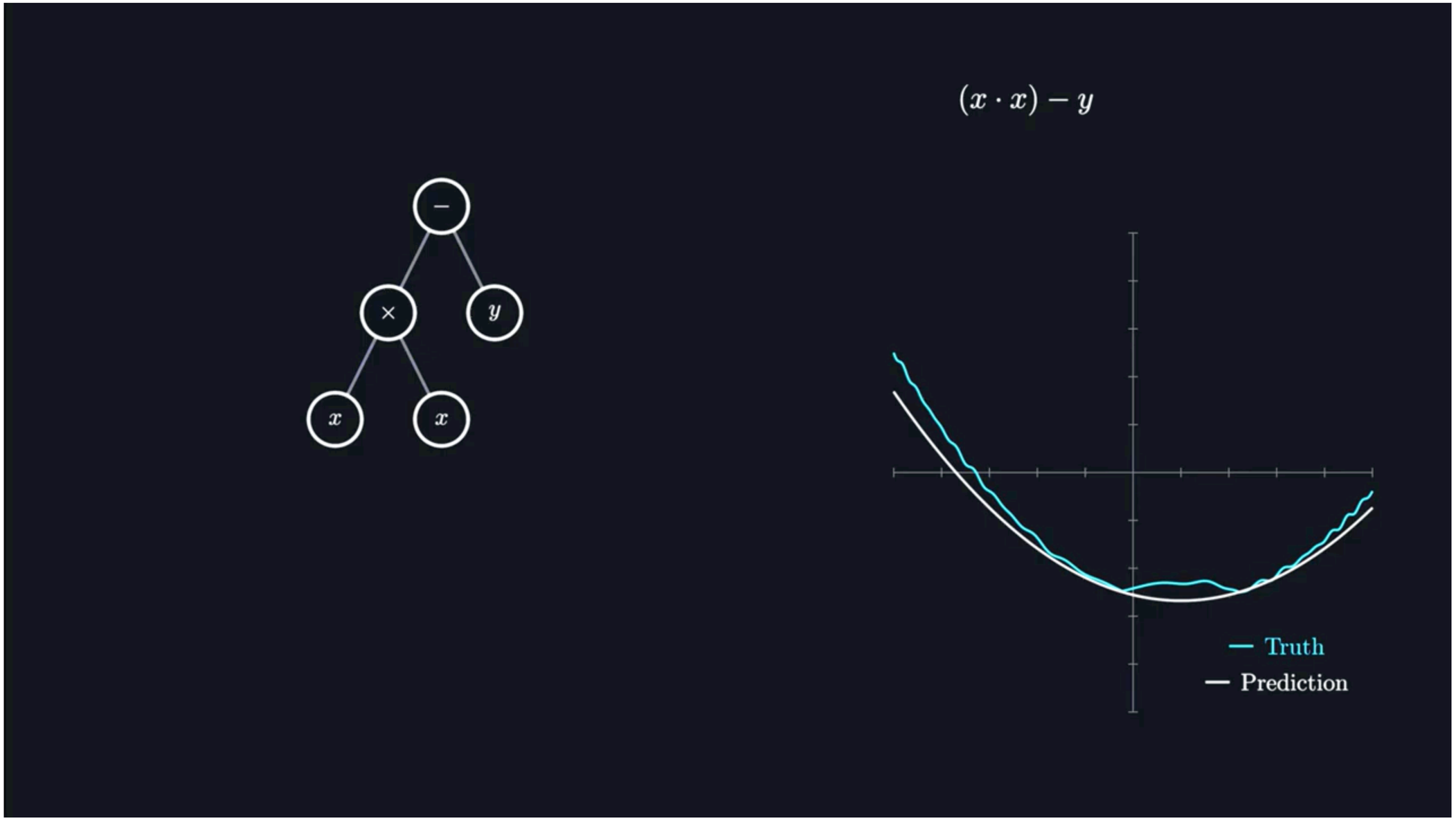
ML tools could be of help

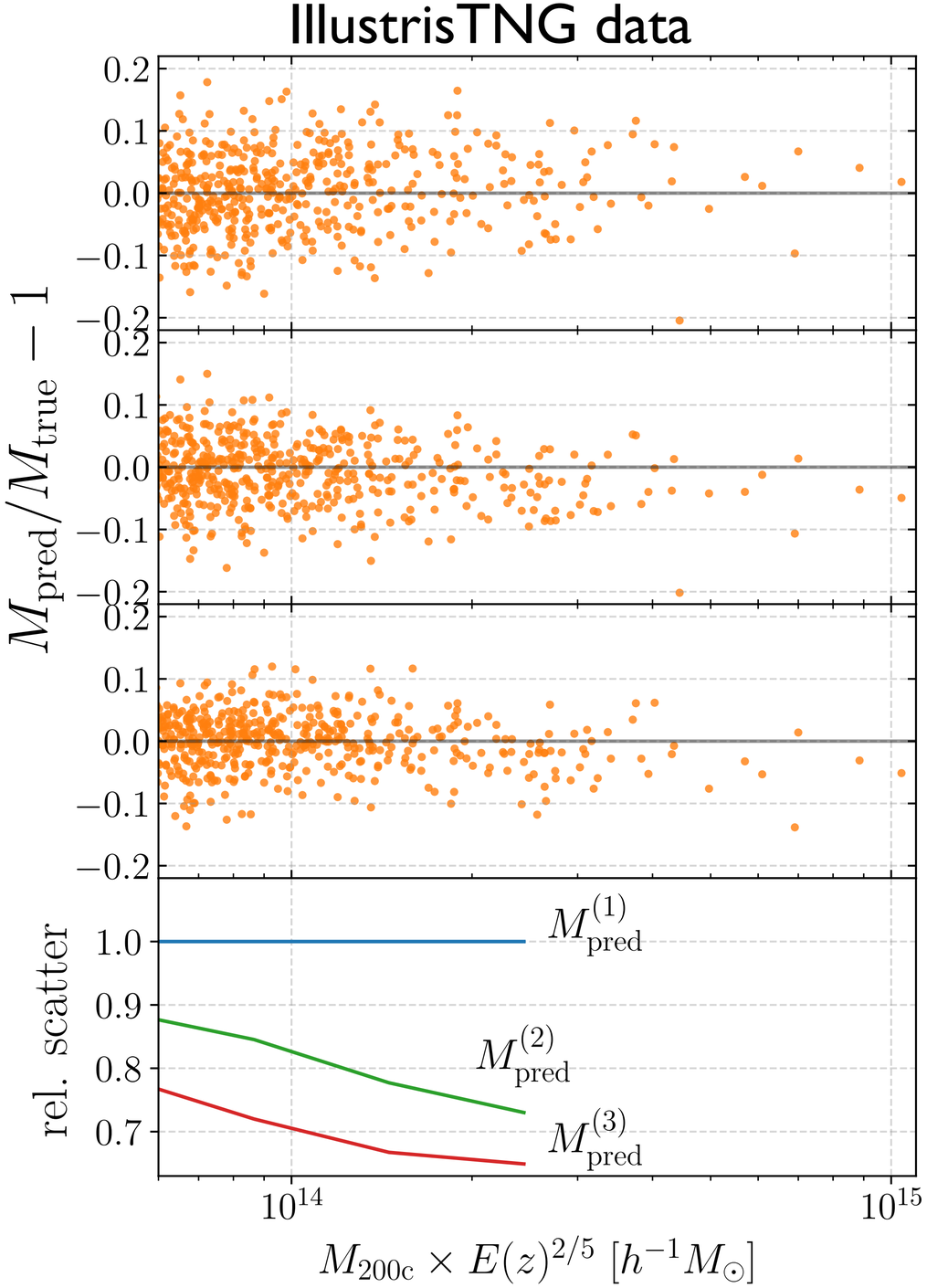


Symbolic regression

(PySR: Cranmer et al. 23)

https://youtu.be/3ezic5Usyxs?si=_ApC5KczchQUPUYg&t=15

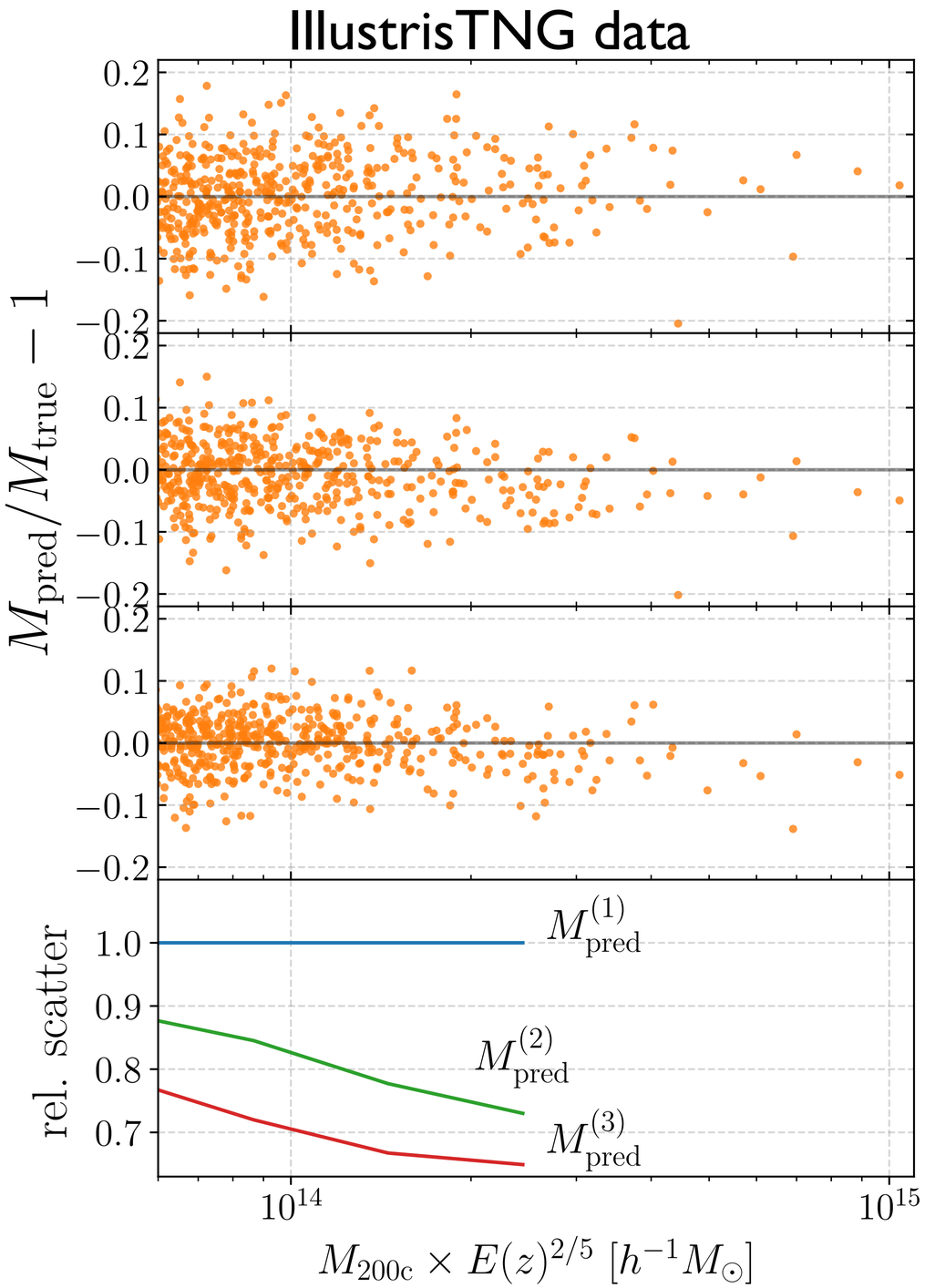




$$M^{(1)}_{\rm pred} \propto Y^{3/5}$$

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

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



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

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

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

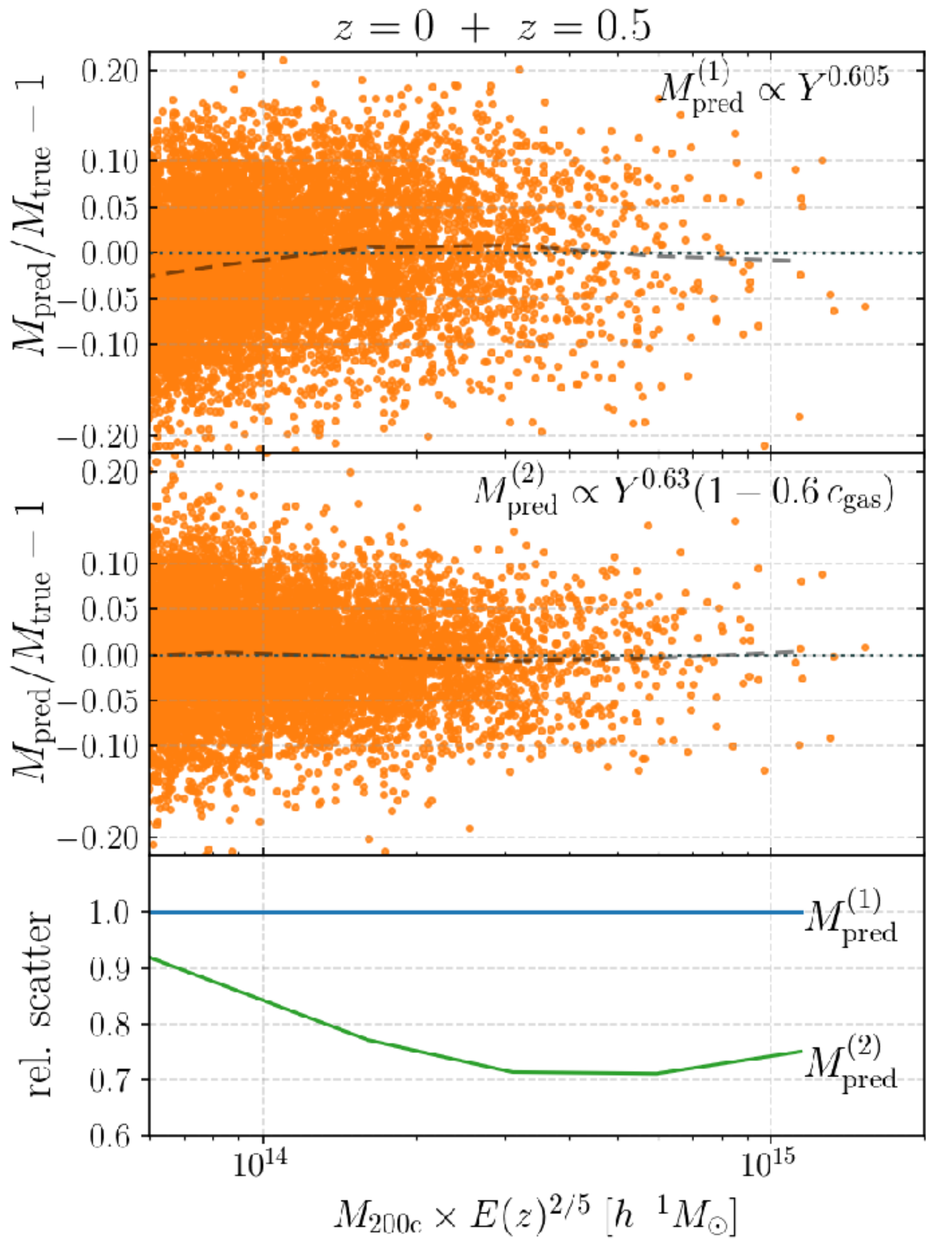
Reasons for dependence:

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

2. Conc. of gas is related to non-thermal pressure

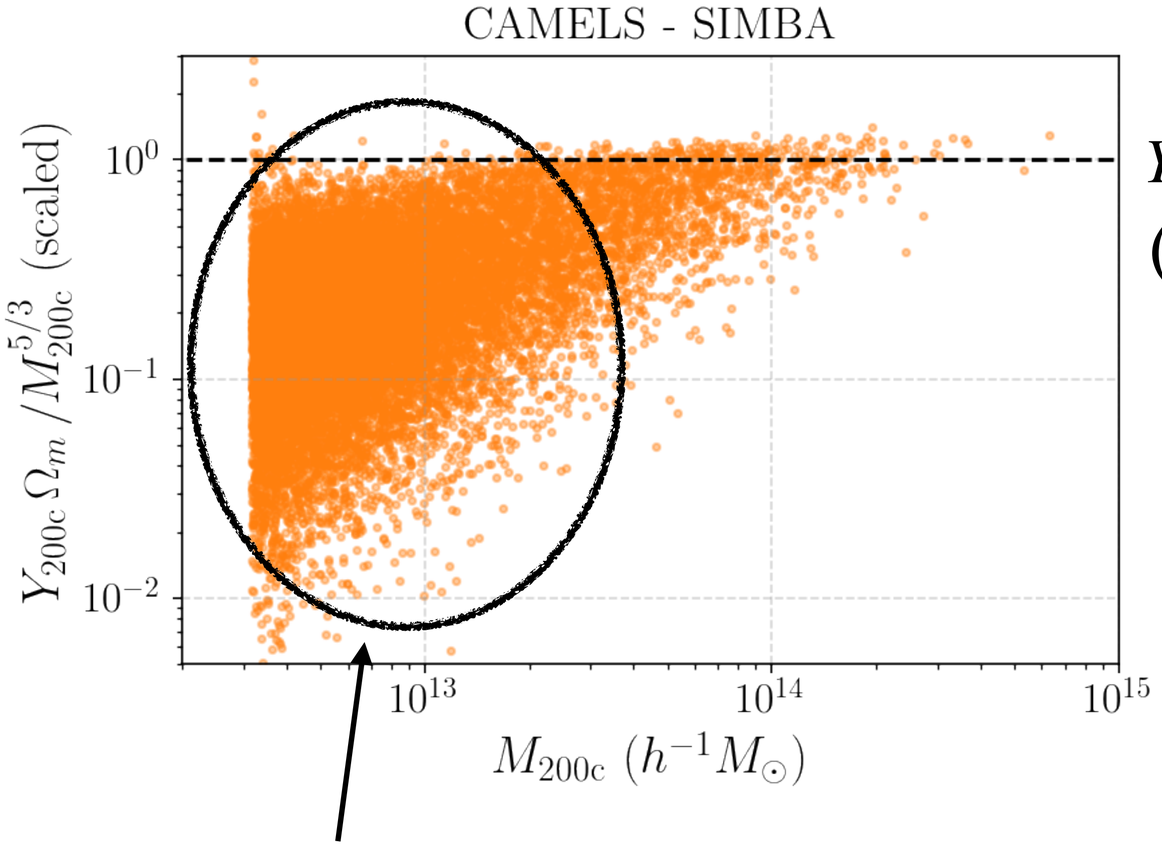
Kravtsov et al. 06,
Arnaud et al. 10

Excellent generalization to high masses!!



Wadekar et al.,
in prep
(using MTNG dataset)

Part II : Reducing deviation from self-similarity (pow. law)



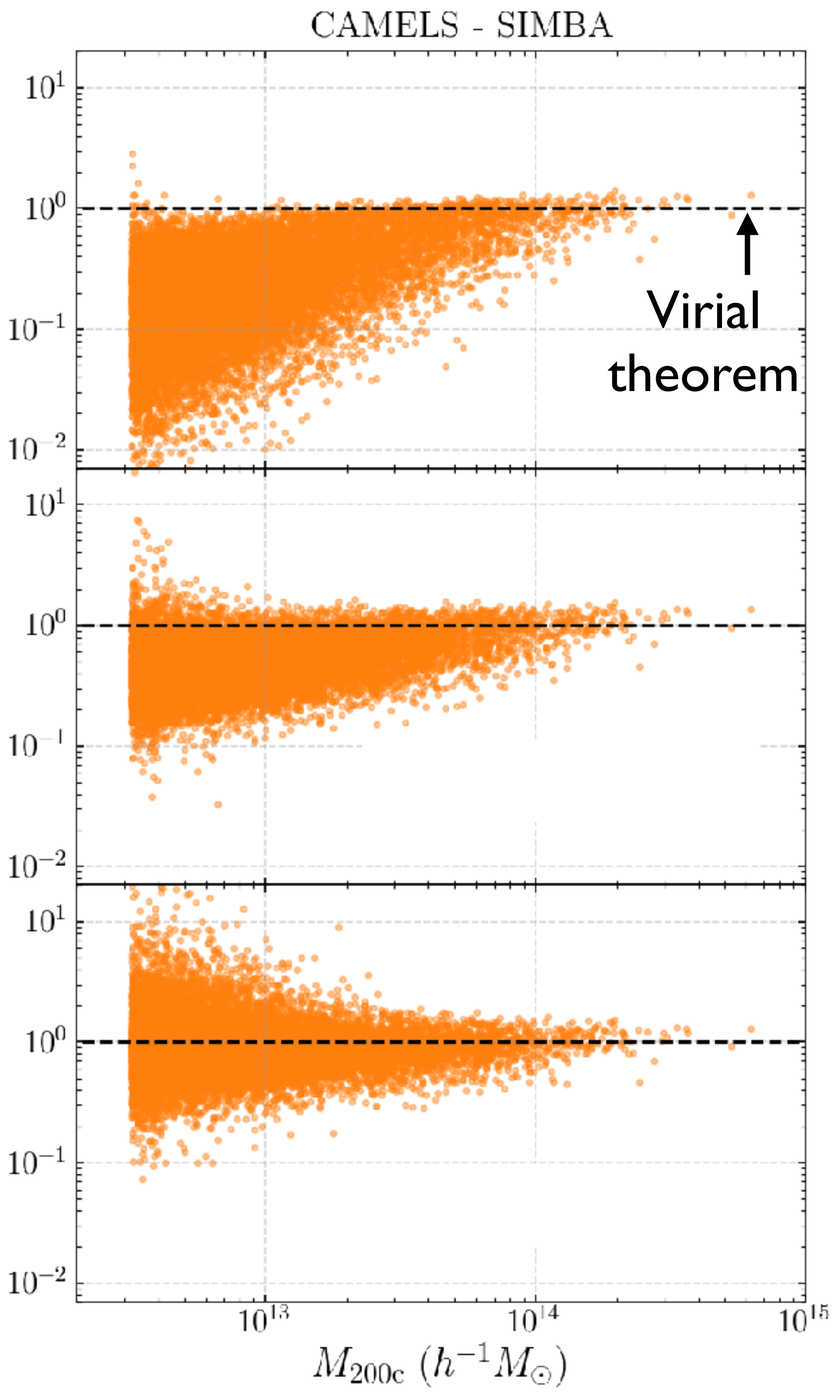
$Y \propto M^{5/3}$

(virial theorem)

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

Ref. to Paco and Arif's talks for intro to CAMELS and SIMBA

Results

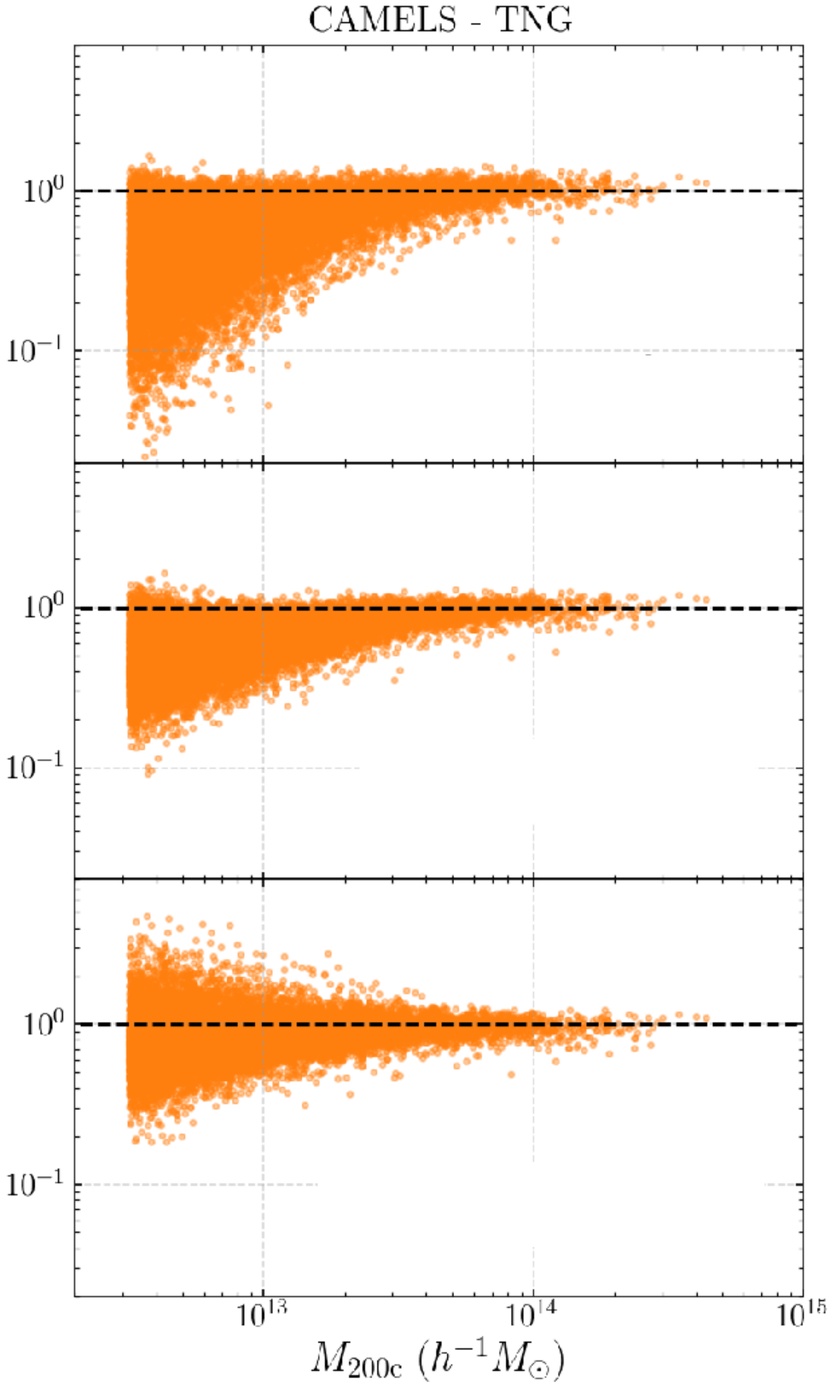


Y

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

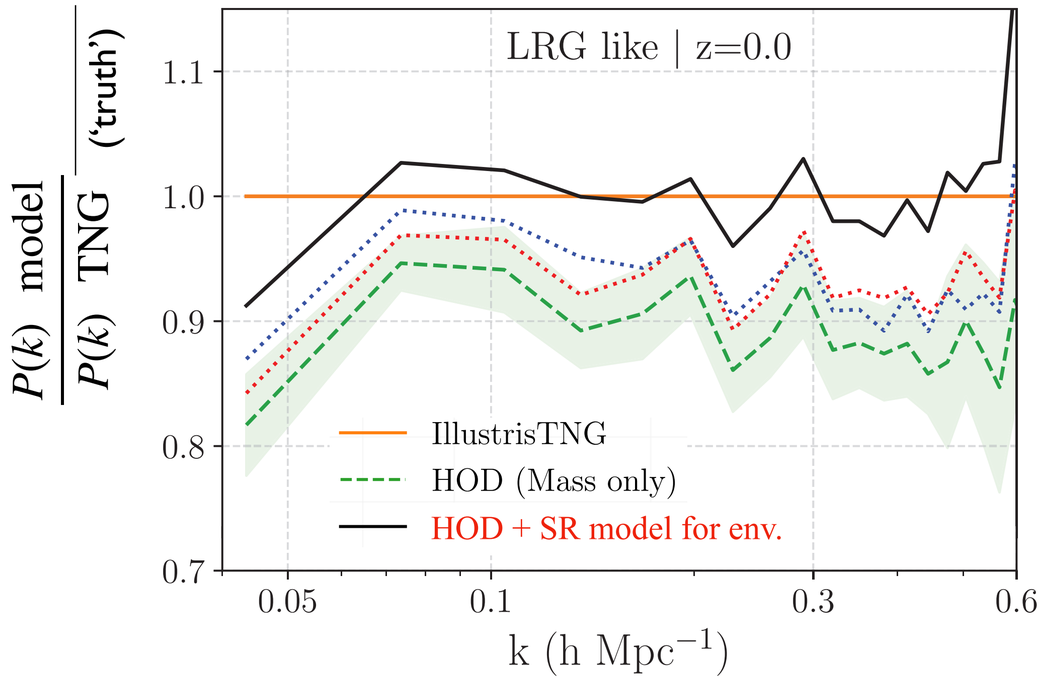
$$\sim (M_*+M_{\rm gas})\,T_{\rm gas}$$

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

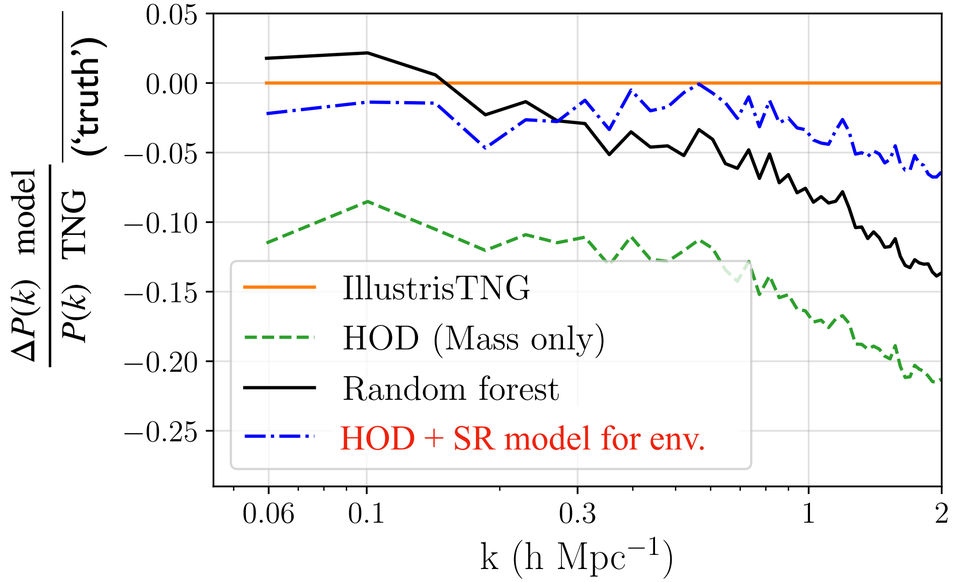


Other applications: galaxy-halo connection

Galaxies

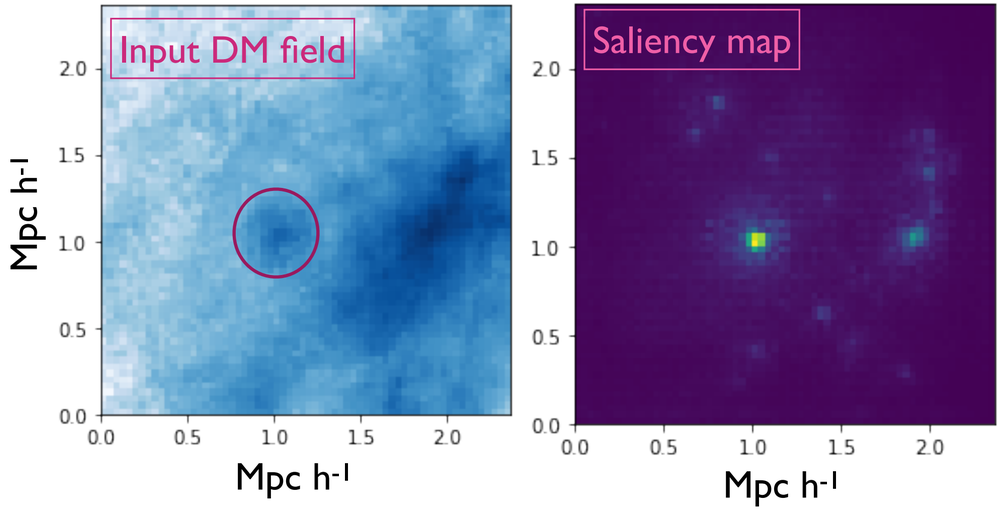


21cm



Delgado, DW et al. 22

DW et al. 20



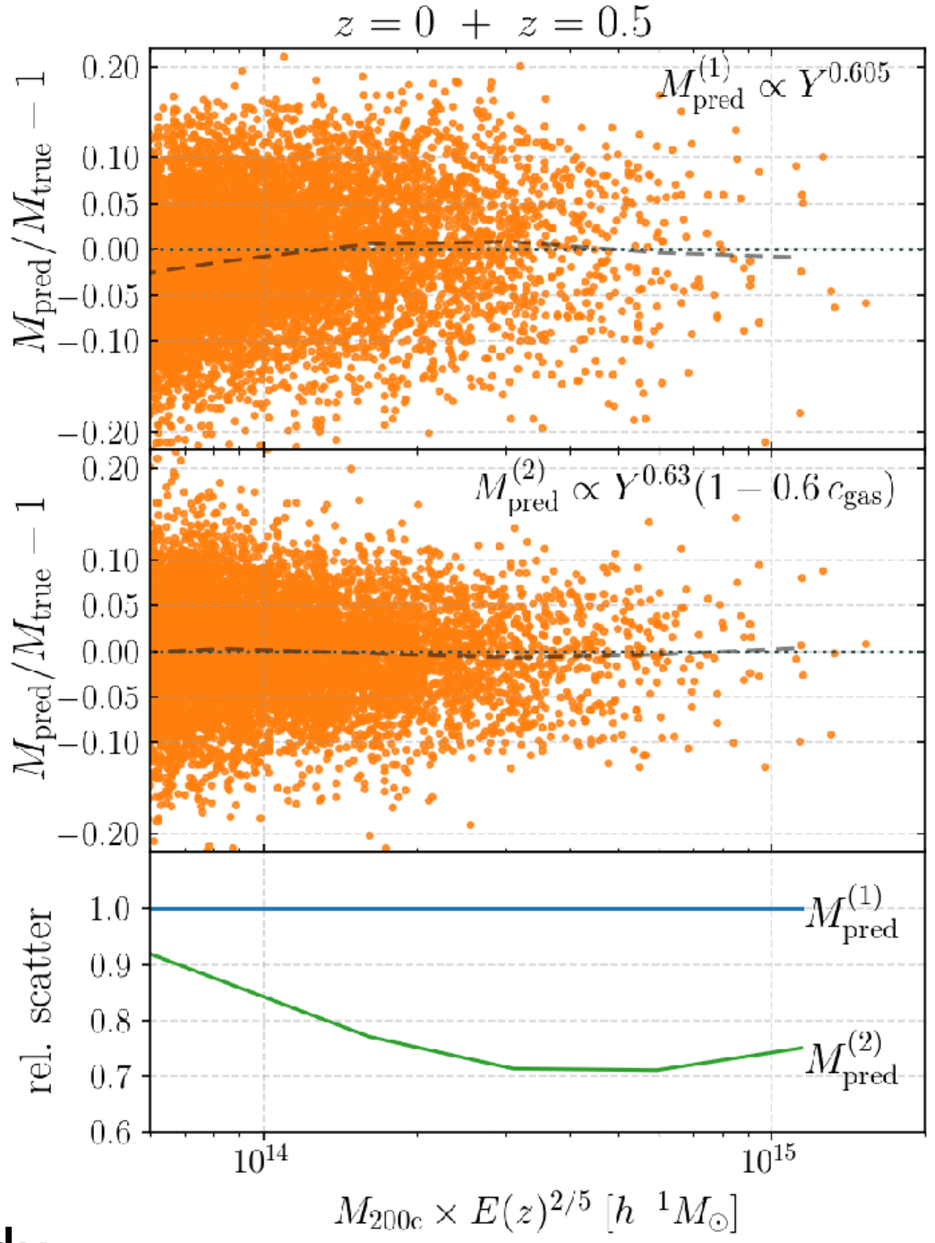
Summary

★ Interpretable ML tools like symbolic regression can be used to robustly improve astrophysical scaling relations

- 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

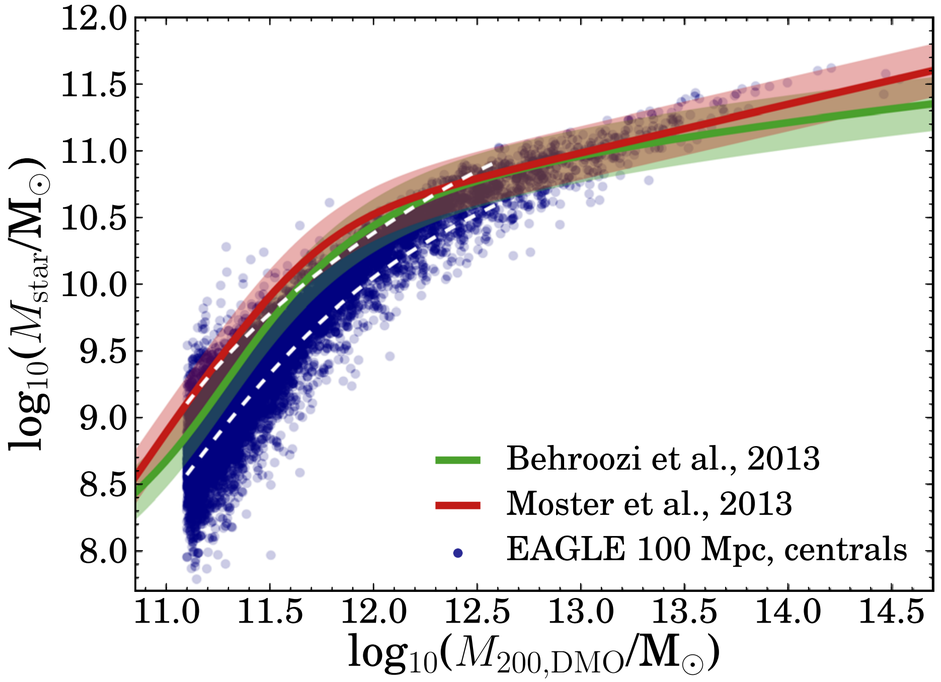
➡Suggestions for other scaling relations?



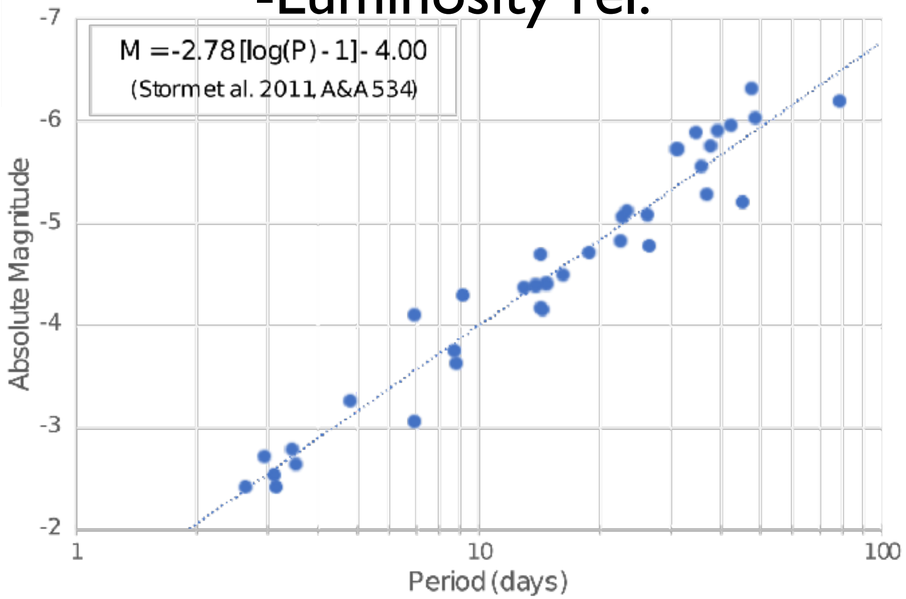
jayw@ias.edu

Application to other scaling relations?

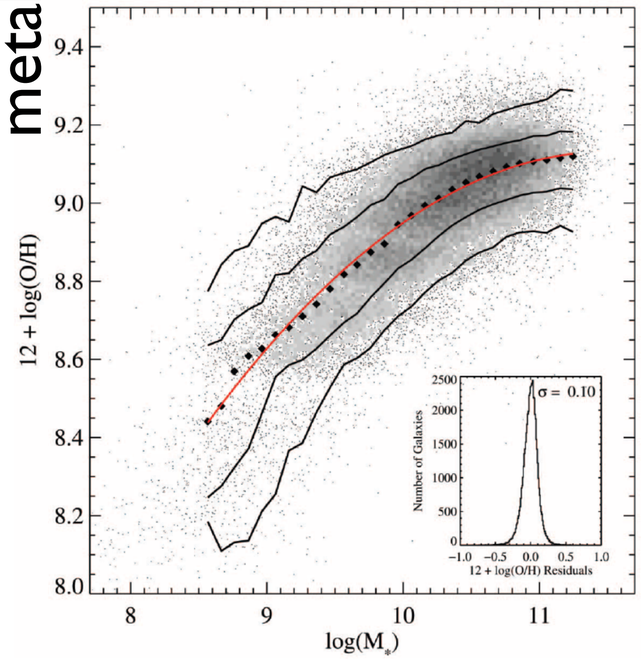
Stellar-halo mass relation



Cepheid Period -Luminosity rel.



metallicity



Stellar mass

- Black hole - bulge mass
- Neutron star prop - GW freq.

Choosing equations with symbolic reg. (Pareto curve)