

Connection between SDSS galaxies and ELUCID subhalos in the eye of machine learning

Xiaoju Xu

Shanghai Jiao Tong University

Galaxy-halo connection

➤ Empirical models

HOD, CLF, SHAM,

assume galaxy-halo connection, model observables and fit to observation

model dependent

➤ Simulations

SAM, hydrodynamic

galaxy formation physics: gas cooling, star formation, feedback,

slow, not 100% the same with observation

➤ Machine learning

learn the galaxy-halo (subhalo) relation in SAM and hydro

depend on SAM and hydro

Galaxy-halo connection

➤ Empirical models

HOD, CLF, SHAM,

assume galaxy-halo connection, model observables and fit to observation

model dependent

➤ Simulations

SAM, hydrodynamic

galaxy formation physics: gas cooling, star formation, feedback,

slow, not 100% the same with observation

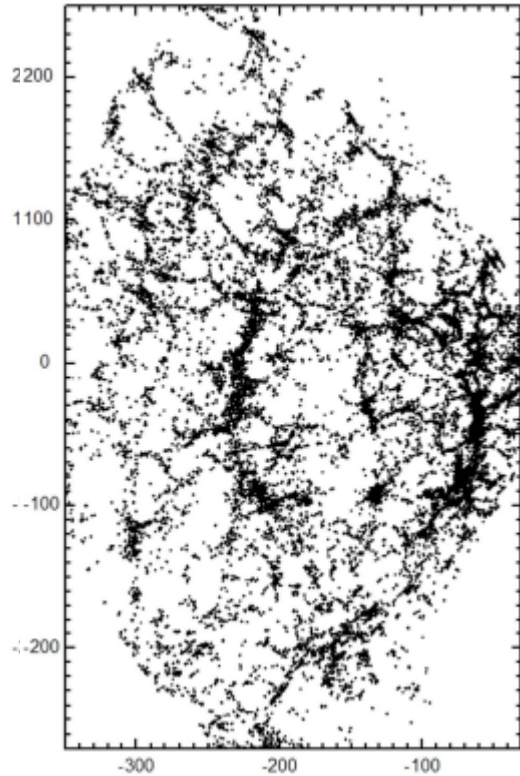
➤ Machine learning

learn the galaxy-halo (subhalo) relation in SAM and hydro

depend on SAM and hydro

use observed galaxies and simulated halos (subhalos) ?

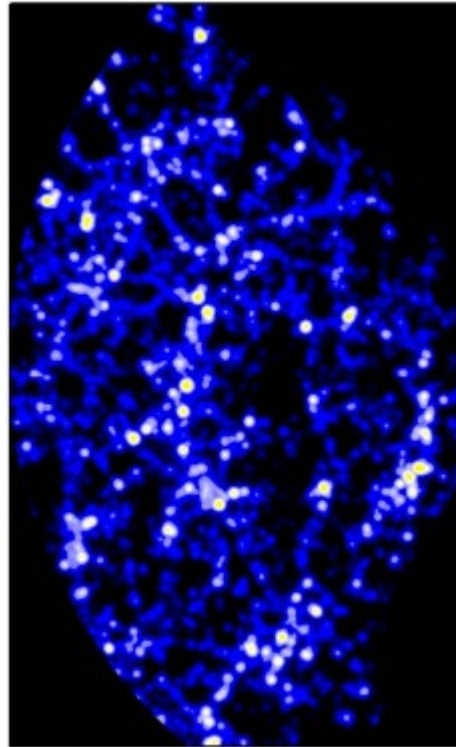
ELUCID constrained simulation



SDSS galaxy

Find underlying matter

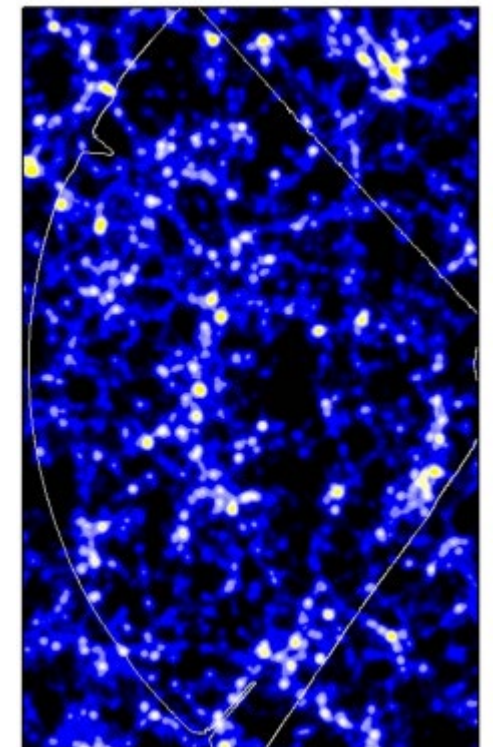
Group finder +
density
reconstruction



Constructed field

Find initial field and run

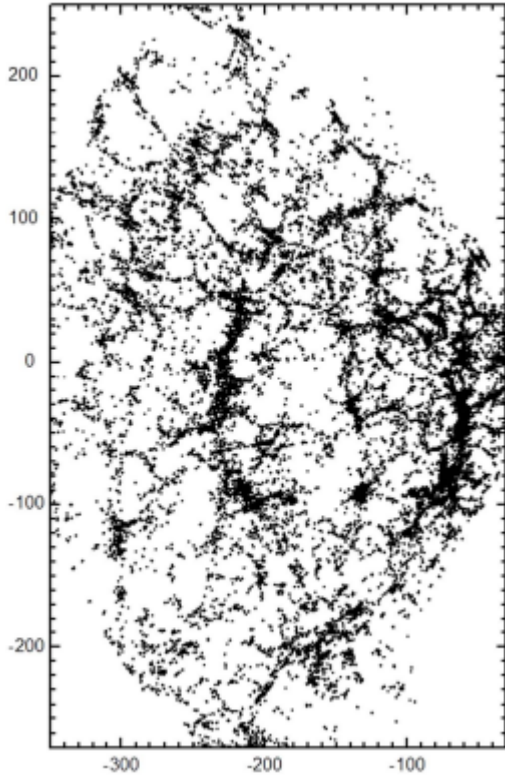
HMC+PM



ELUCID

(Yang2007, Wang2009, Wang2014, Wang2016)

SDSS-ELUCID matching



SDSS galaxy

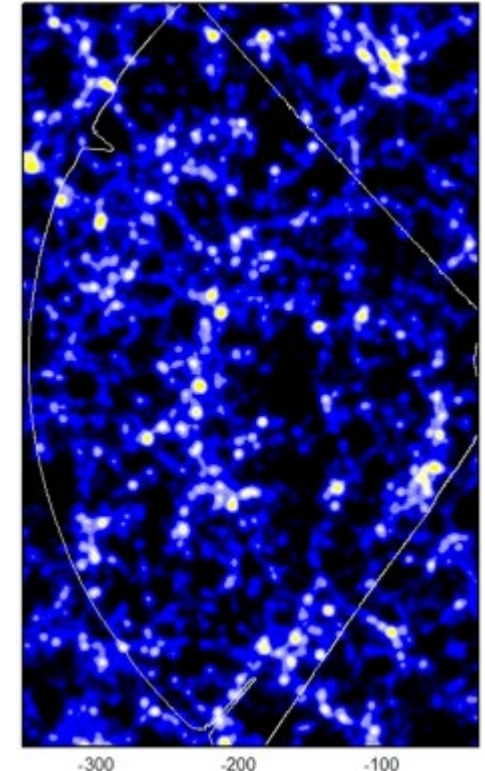
neighborhood abundance matching (Yang2018)

$$P(r_p, \pi, M_{\text{sh}}) = M_{\text{sh}} \exp\left(-\frac{r_p^2}{2r_{\text{off}}^2}\right) \exp\left(-\frac{\pi^2}{2v_{\text{off}}^2}\right).$$

galaxy properties

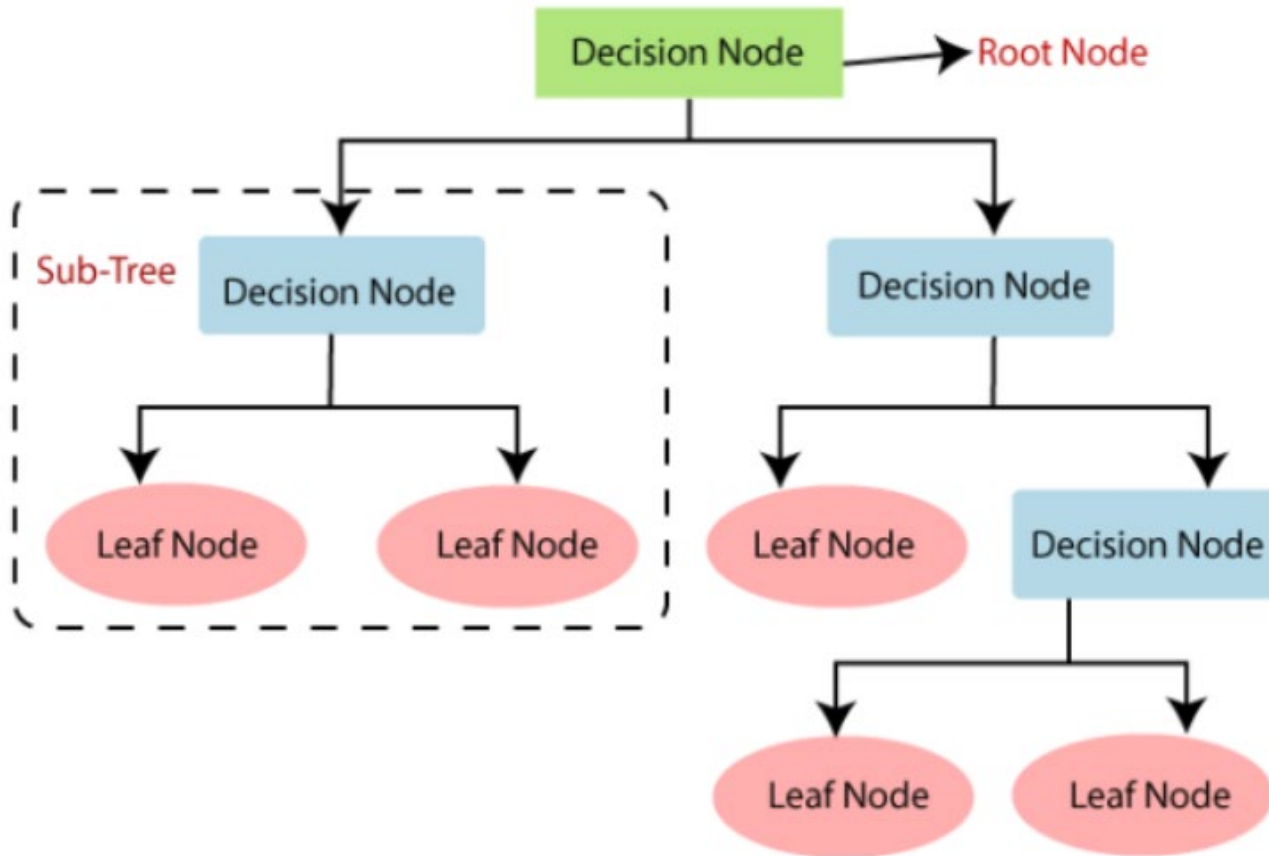
subhalo properties

Mag	Mstar	SFR	Msub	Zform	Vmax



ELUCID subhalo

Learn the relation between SDSS galaxies and ELUCID subhalos with Random Forest



Input: subhalo properties

mass indicators: $M_{\text{sub}}, M_{\text{peak}}, V_{\text{max}} \dots$

assembly: $z_{0.5}, z_{\text{acc}}, z_{\text{mpeak}} \dots$

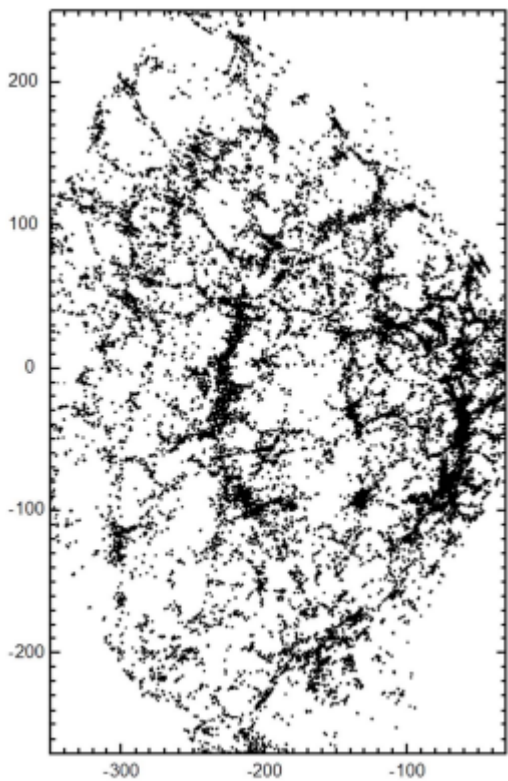
merger: $N_{\text{merg}}, z_{\text{first}}, z_{\text{last}}$

environment: density, web

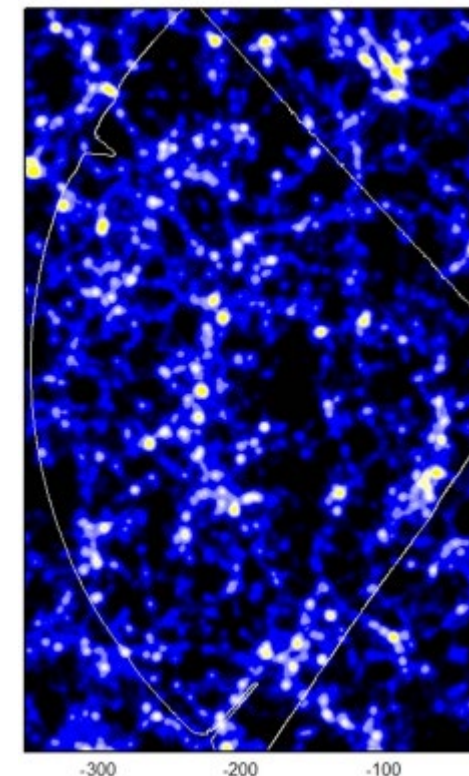
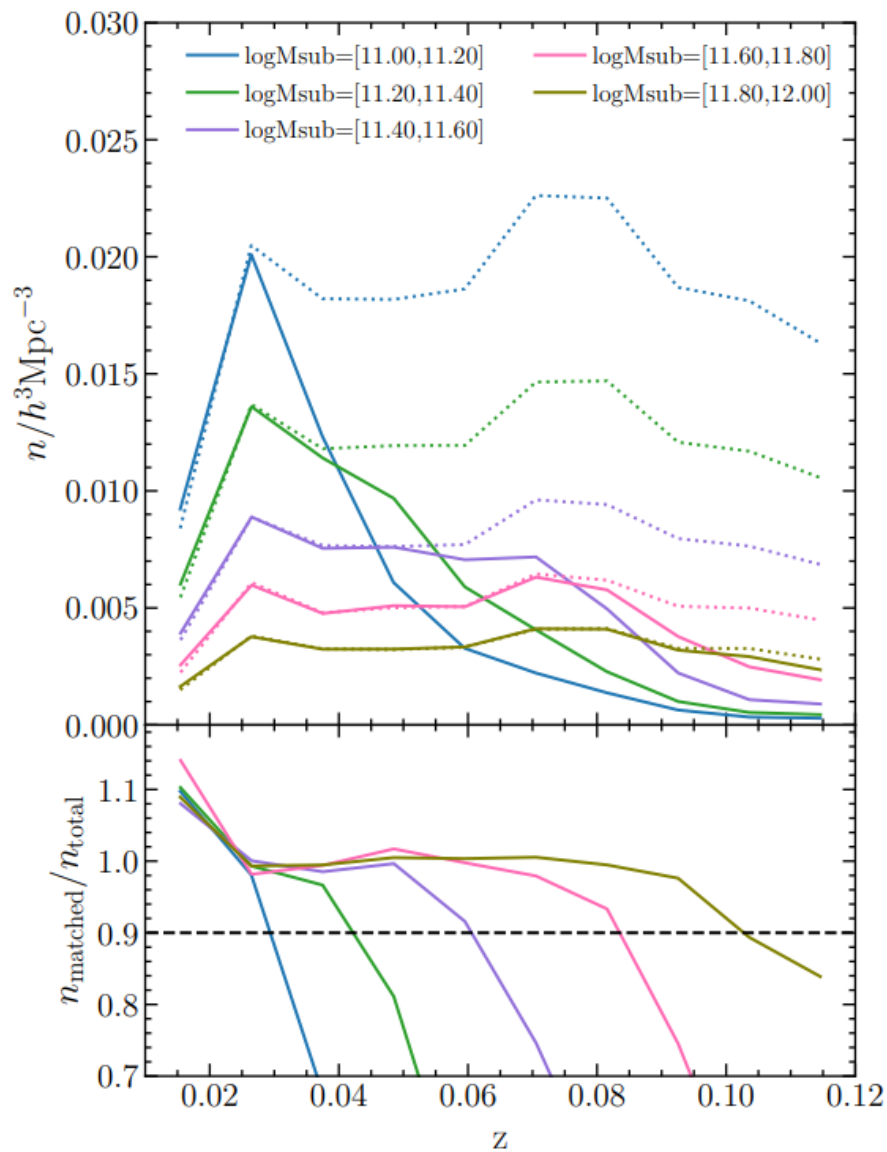
Output: galaxy properties

$M_r, g-r, M_{\text{star}}, \text{SFR} \dots$

Sample selection



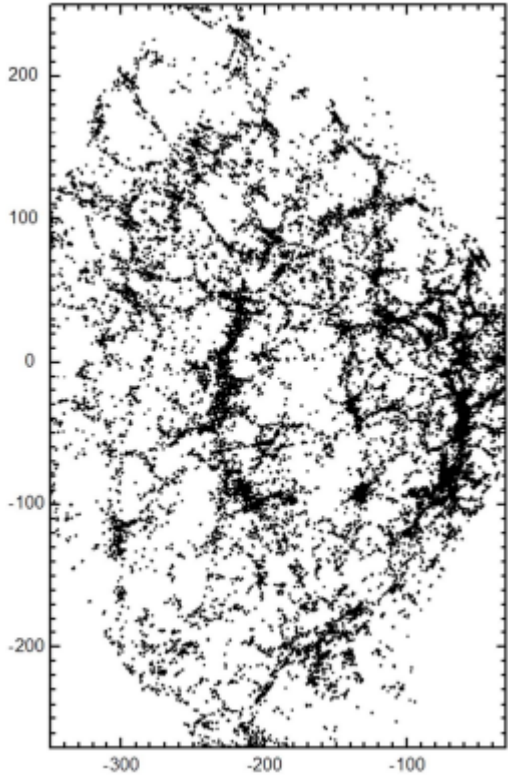
SDSS galaxy



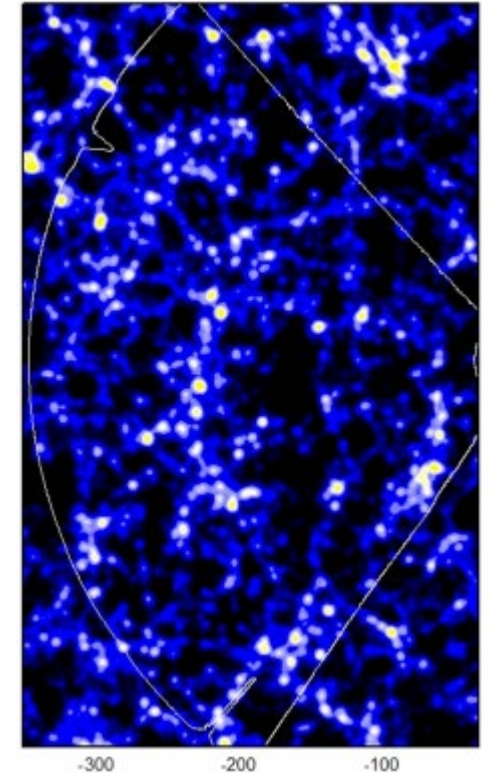
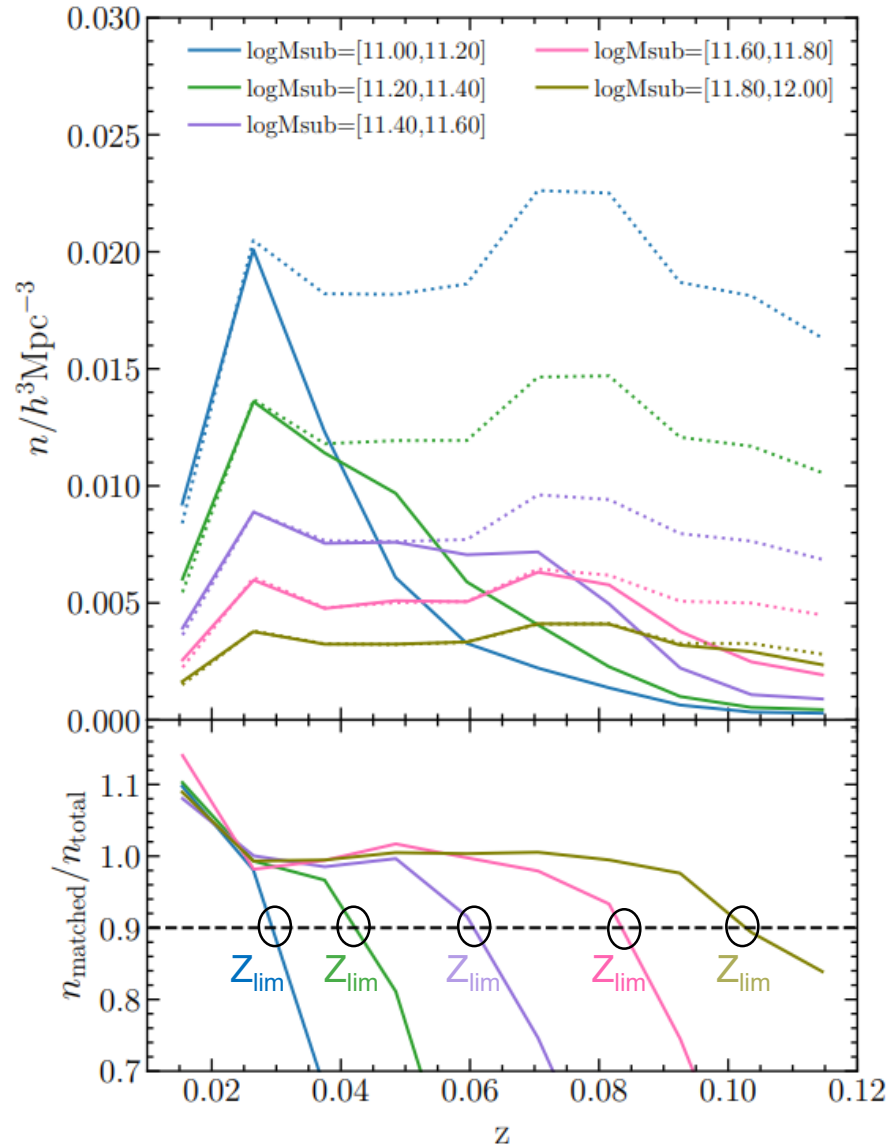
ELUCID subhalo

Msub complete:

Sample selection



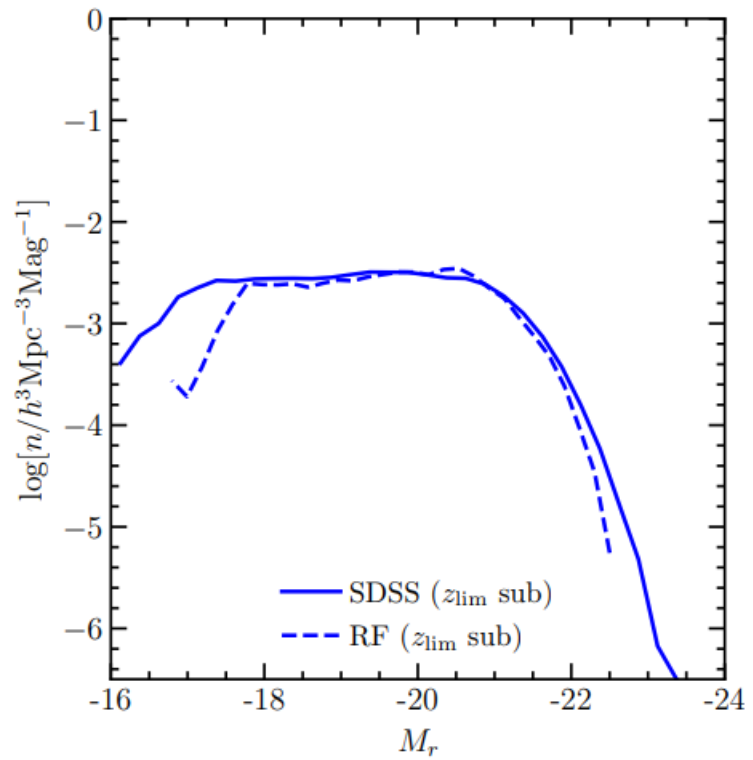
SDSS galaxy



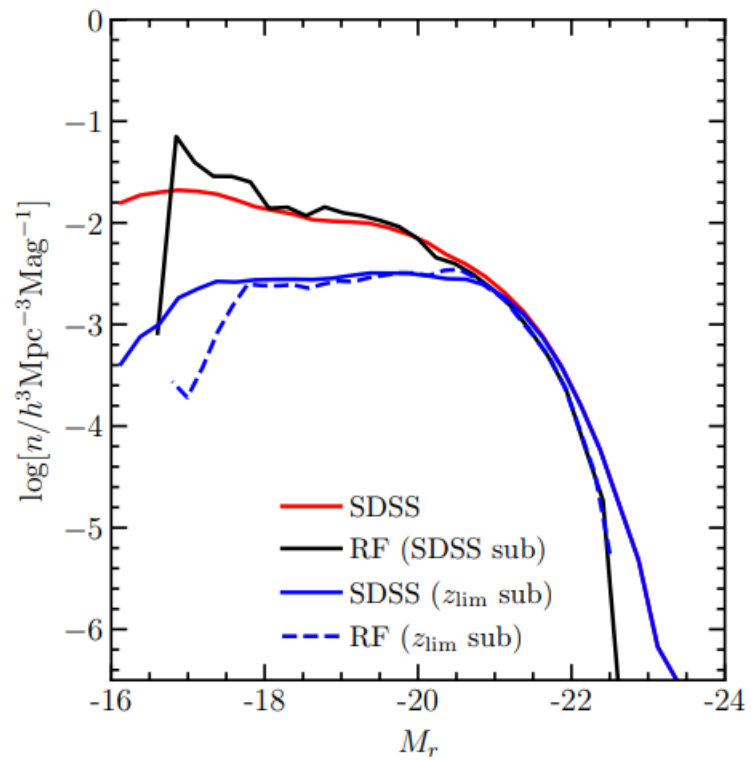
ELUCID subhalo

Msub complete:
Select galaxies in SDSS- ELUCID
matched catalog with $z < z_{\text{lim}}$

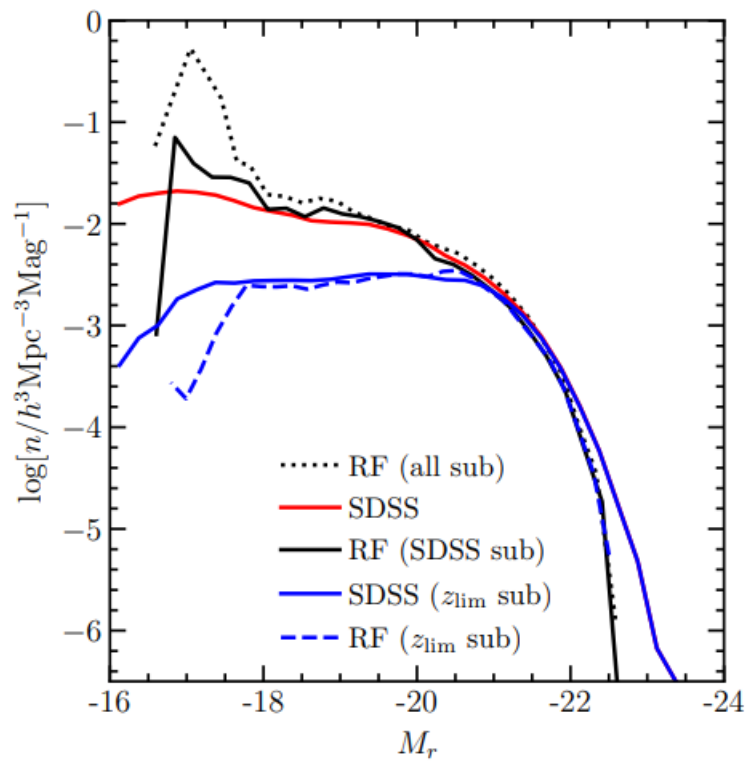
Mr prediction



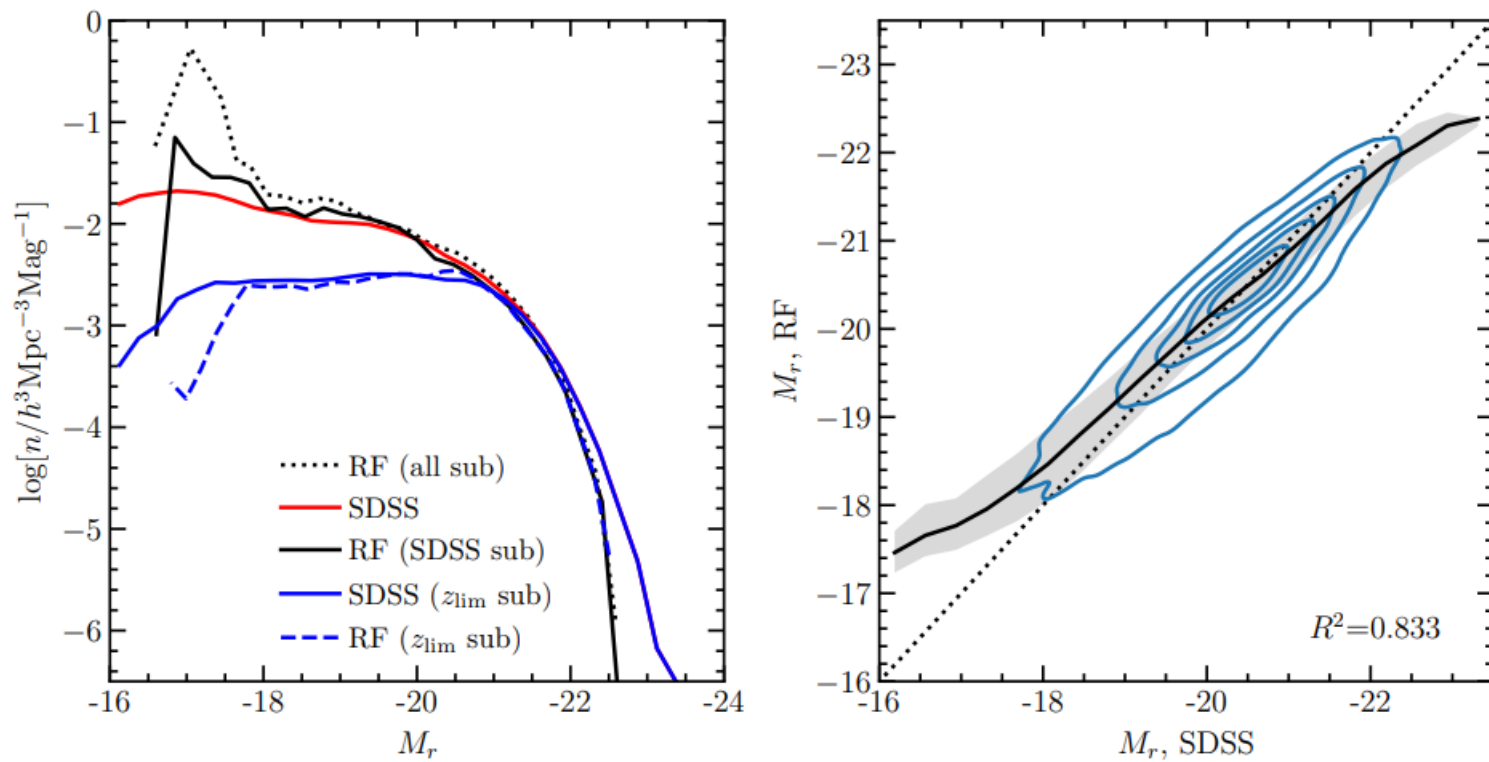
Mr prediction



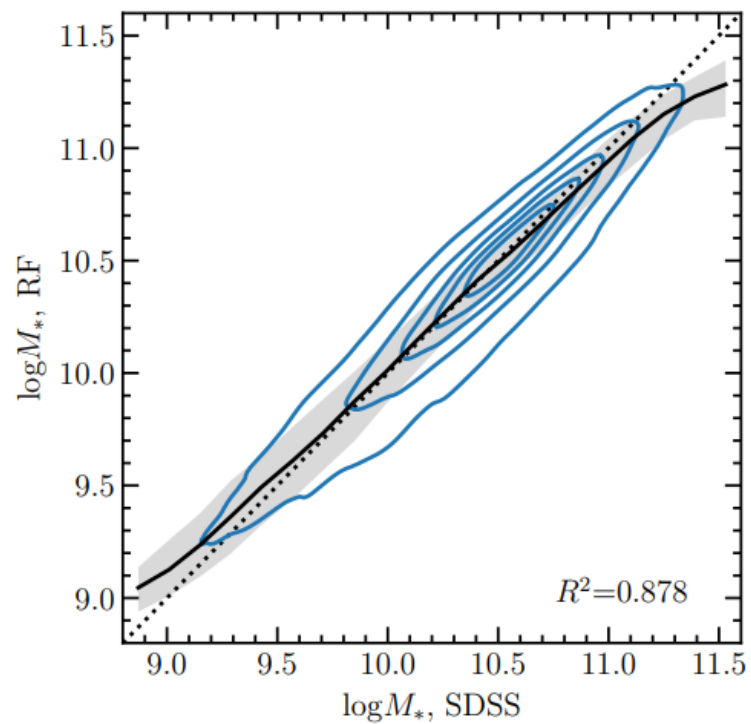
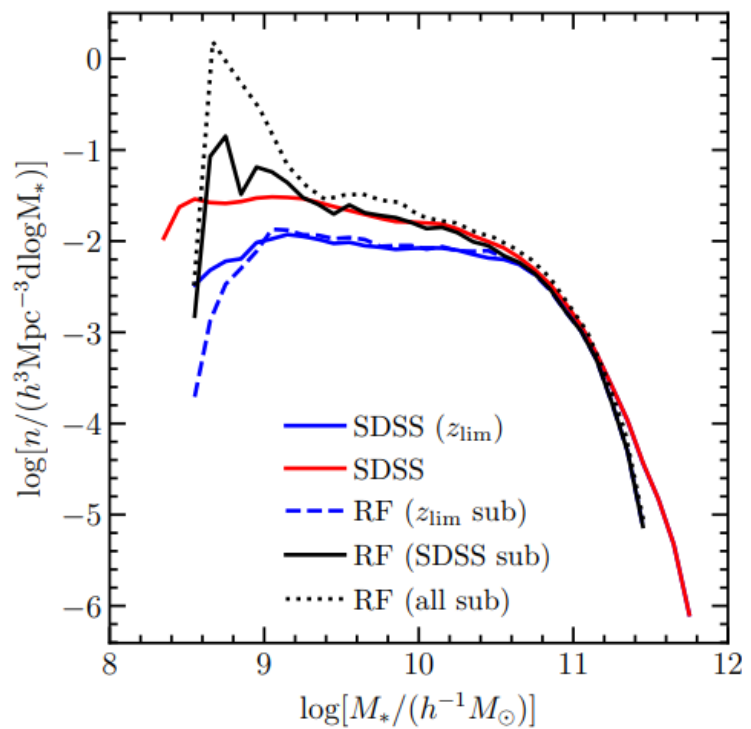
Mr prediction



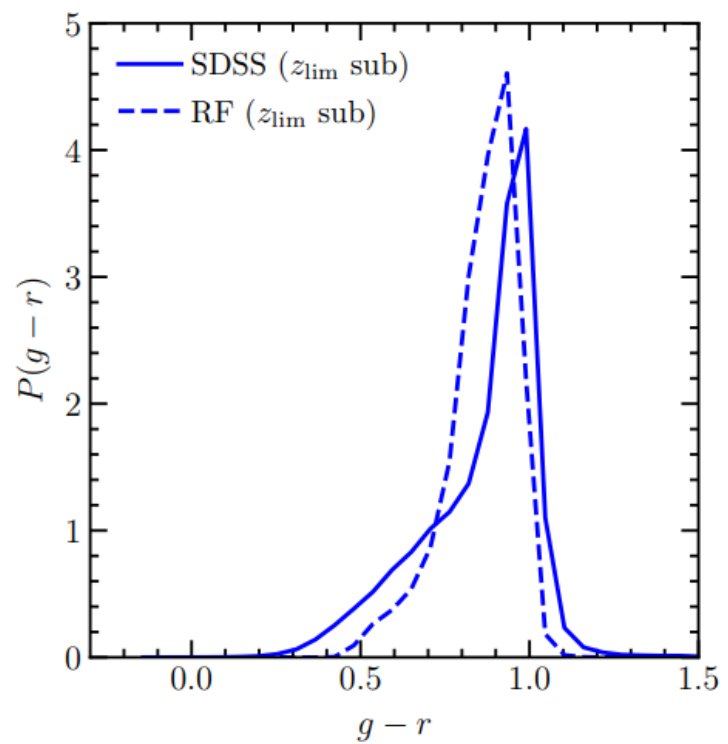
Mr prediction



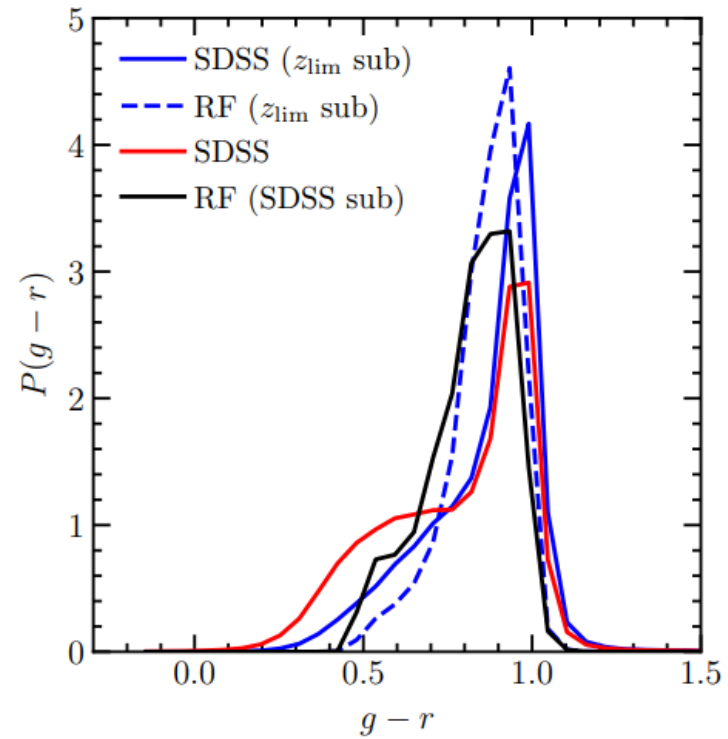
Stellar mass prediction



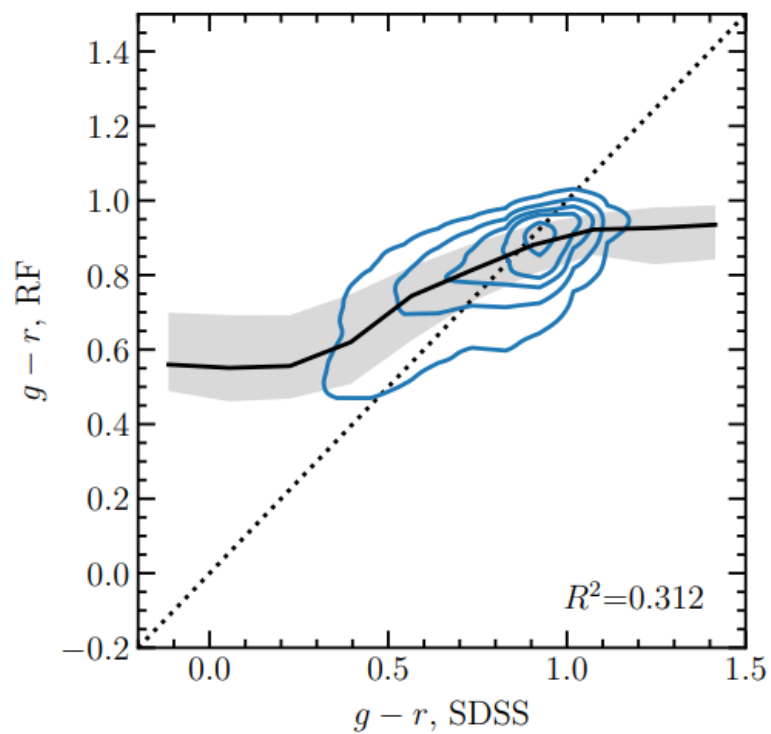
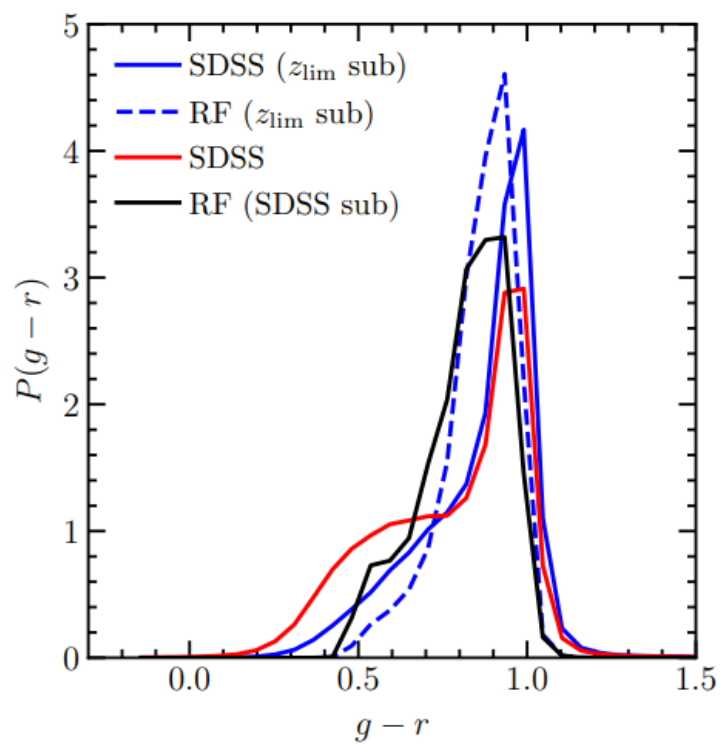
Color prediction



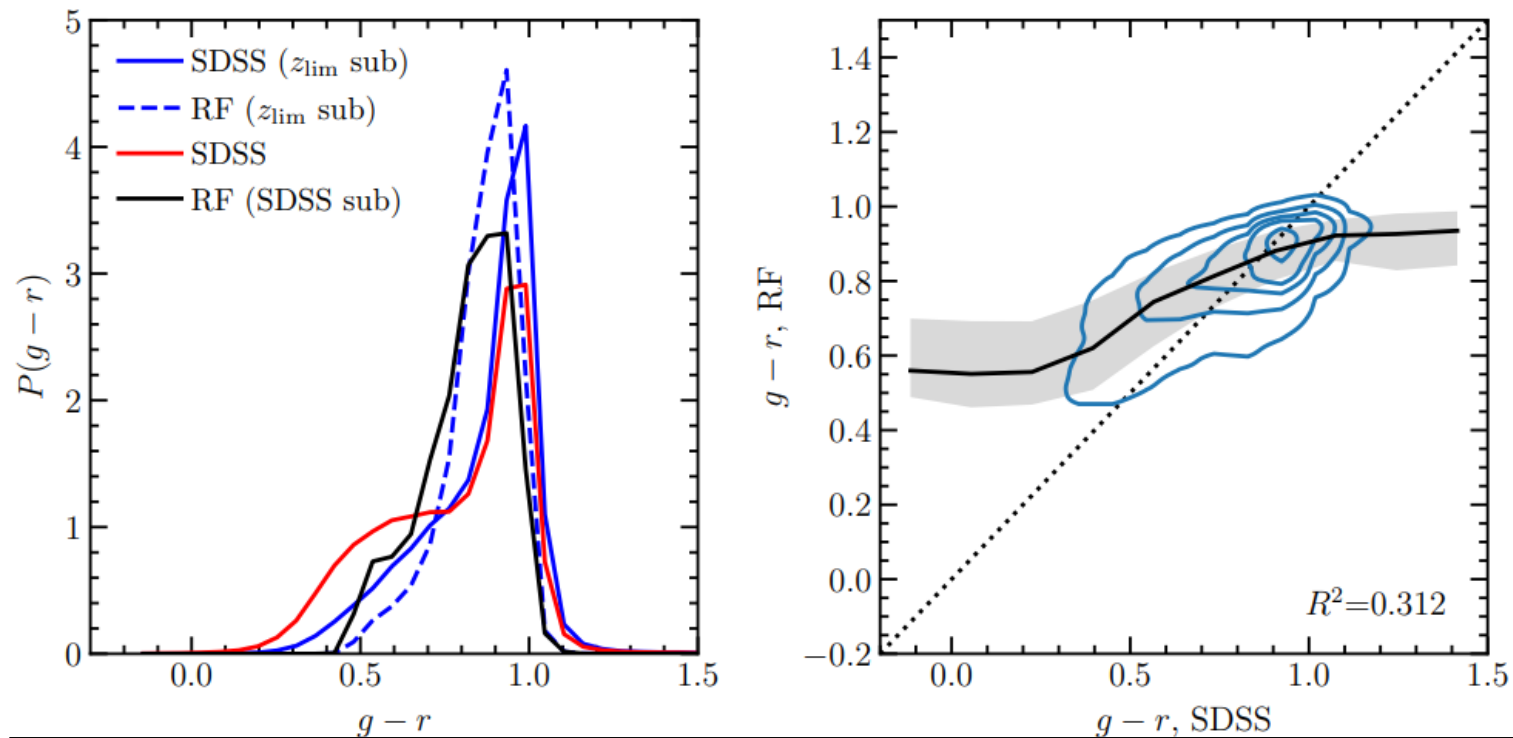
Color prediction



Color prediction



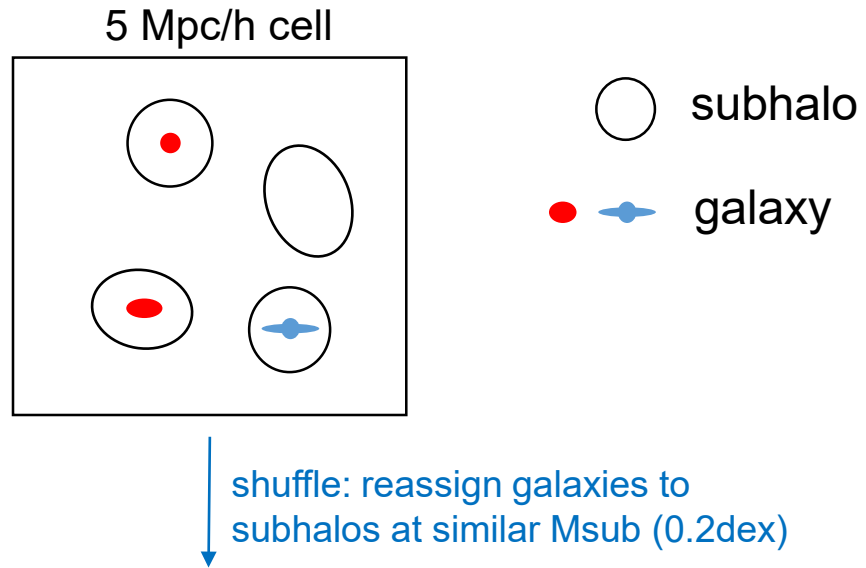
Color prediction



Low prediction score:

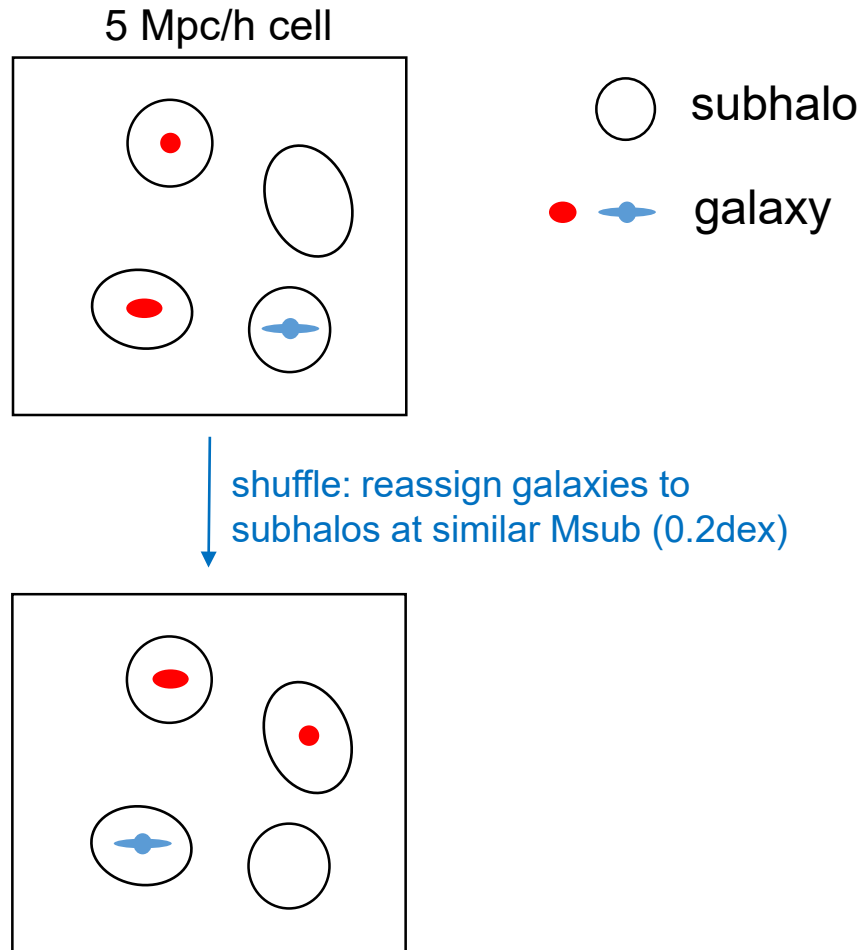
- Low connection between galaxy color and subhalo properties
- Noises in the training data: mismatch between galaxy and subhalo

Mismatch effect with SAM



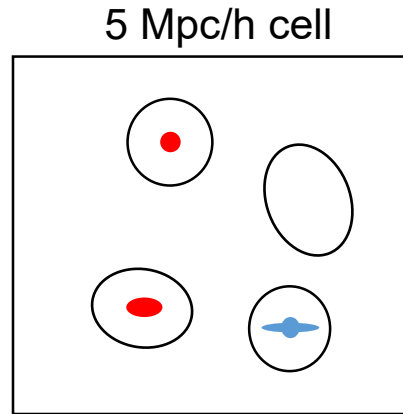
- L_{galaxies} (Luo2016) SAM implemented on ELUCID
- Same subhalos as the SDSS-ELUCID sample

Mismatch effect with SAM



- L_{galaxies} (Luo2016) SAM implemented on ELUCID
- Same subhalos as the SDSS-ELUCID sample

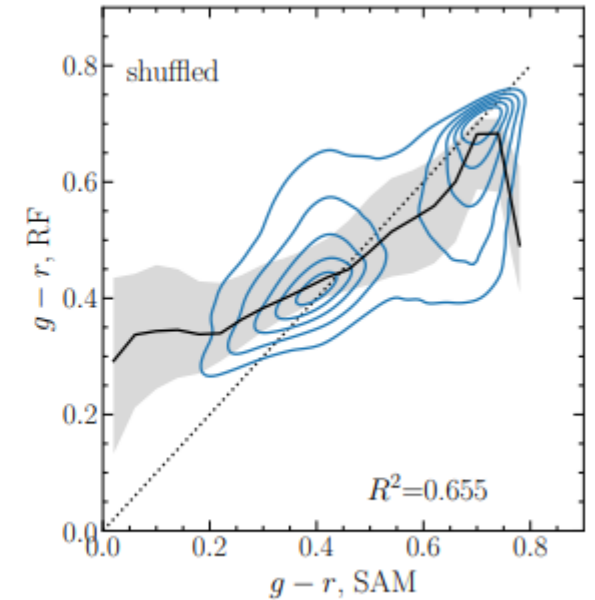
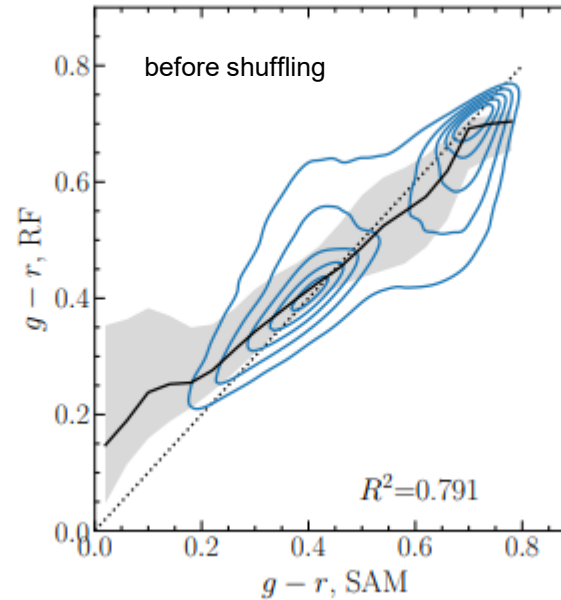
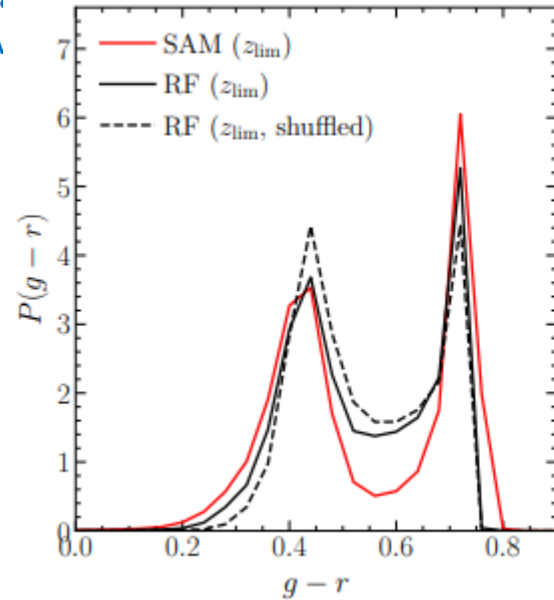
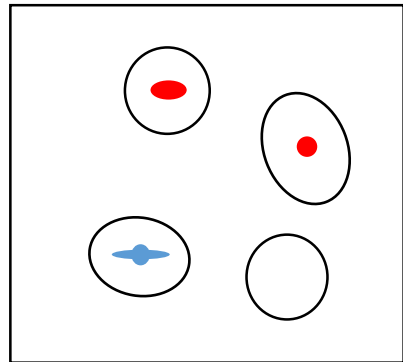
Mismatch effect with SAM



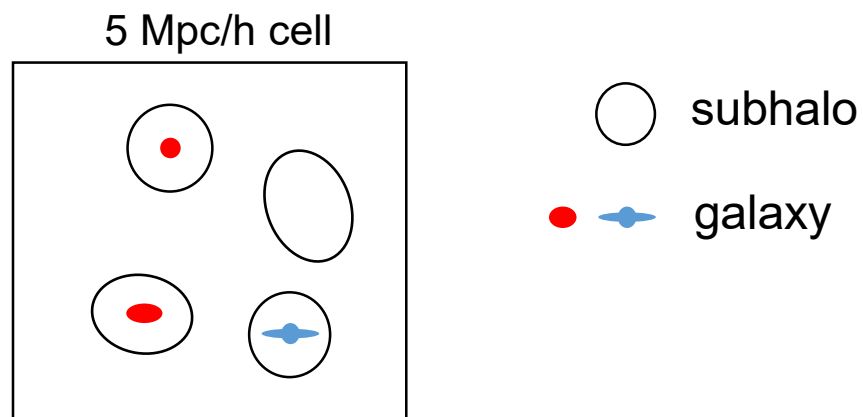
○ subhalo
● galaxy

- Lgalaxies (Luo2016) SAM implemented on ELUCID
- Same subhalos as the SDSS-ELUCID sample

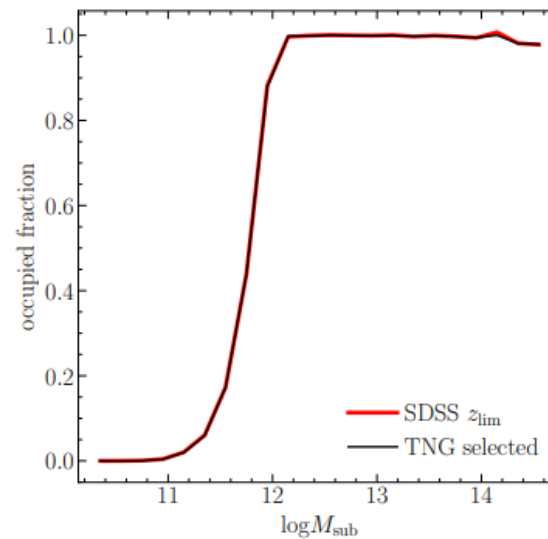
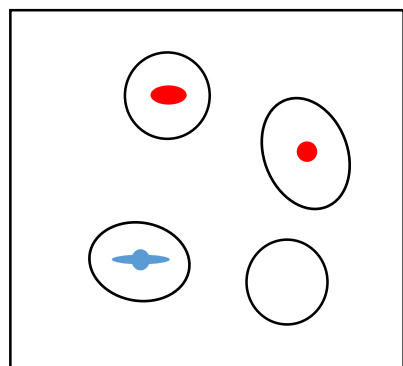
shuffle: reassign gal;
subhalos at similar M



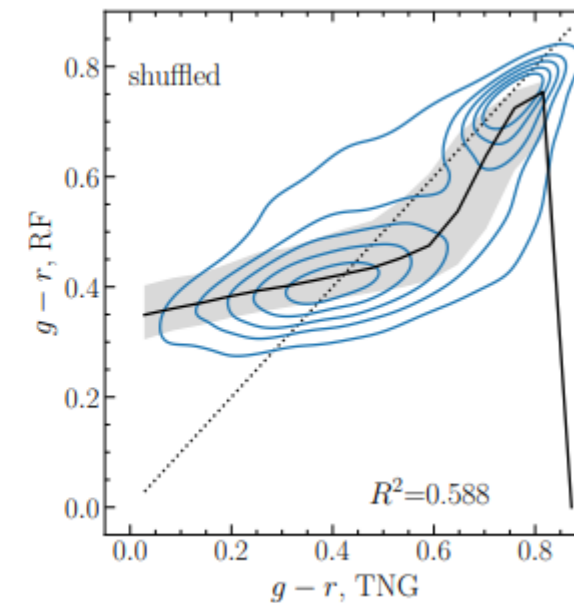
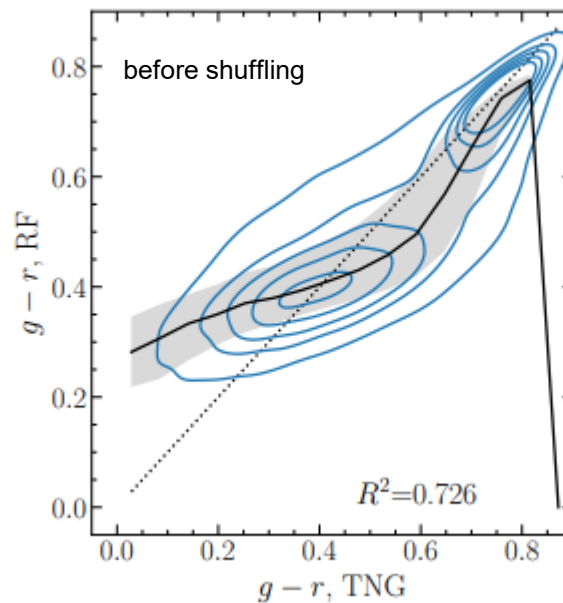
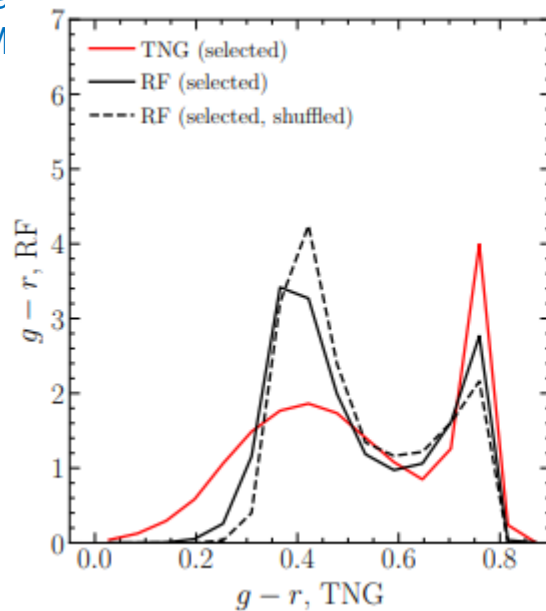
Mismatch effect with TNG



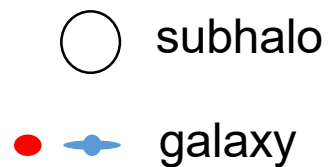
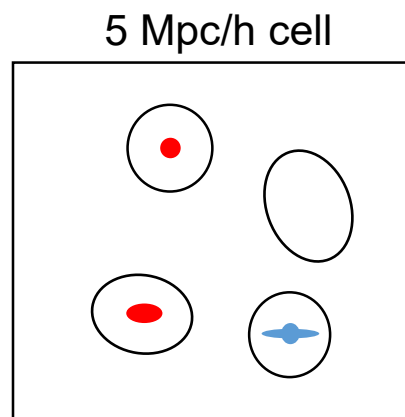
shuffle: reassign galaxy subhalos at similar M



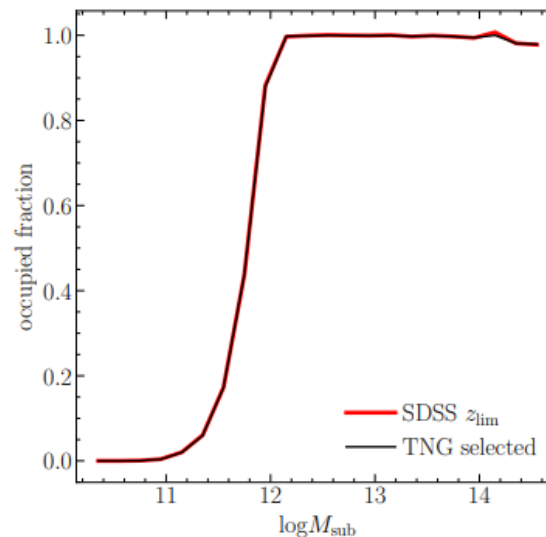
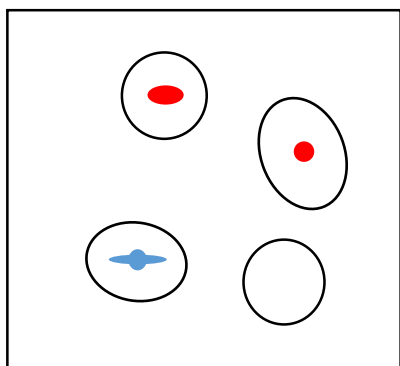
select similar subhalo population in TNG



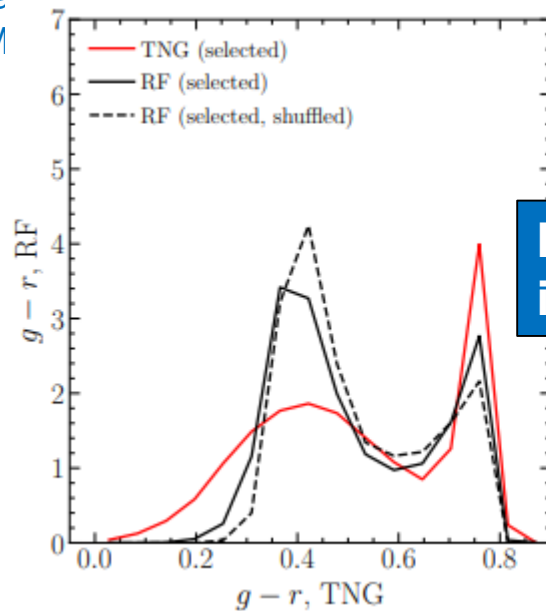
Mismatch effect with TNG



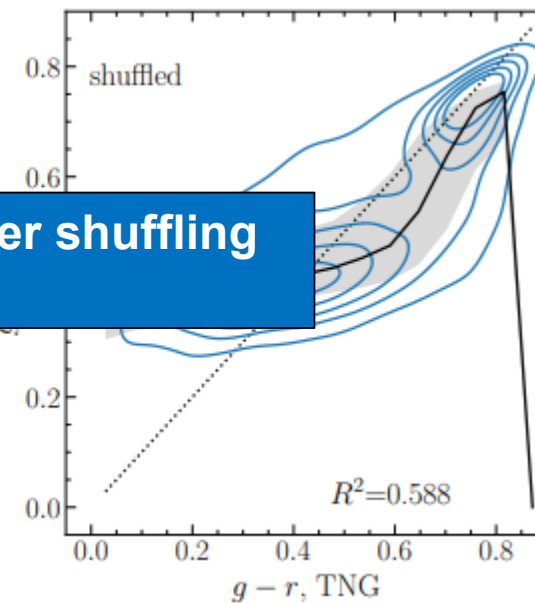
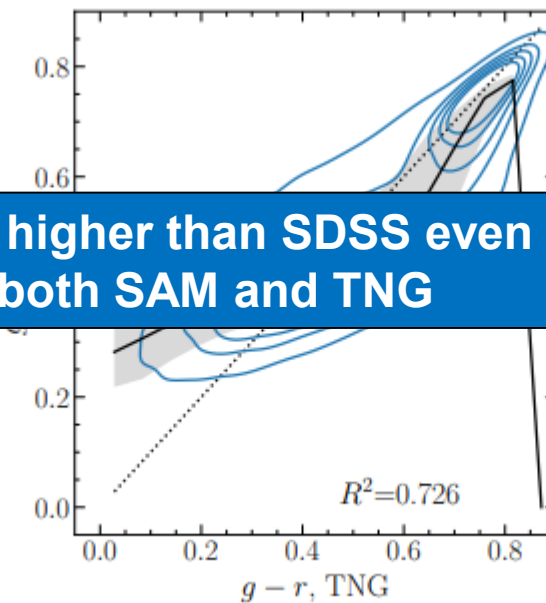
shuffle: reassign galaxy subhalos at similar M



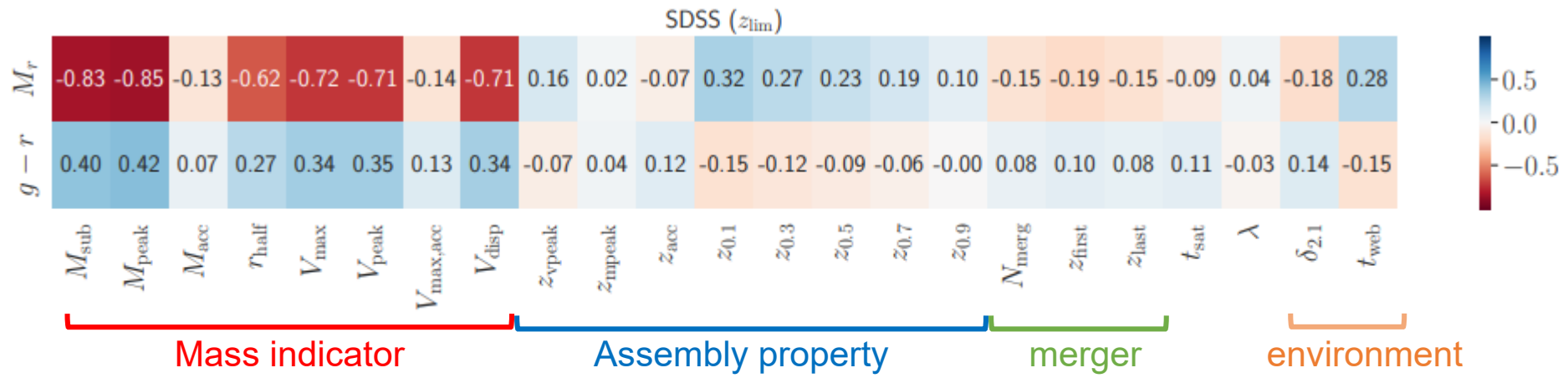
select similar subhalo population in TNG



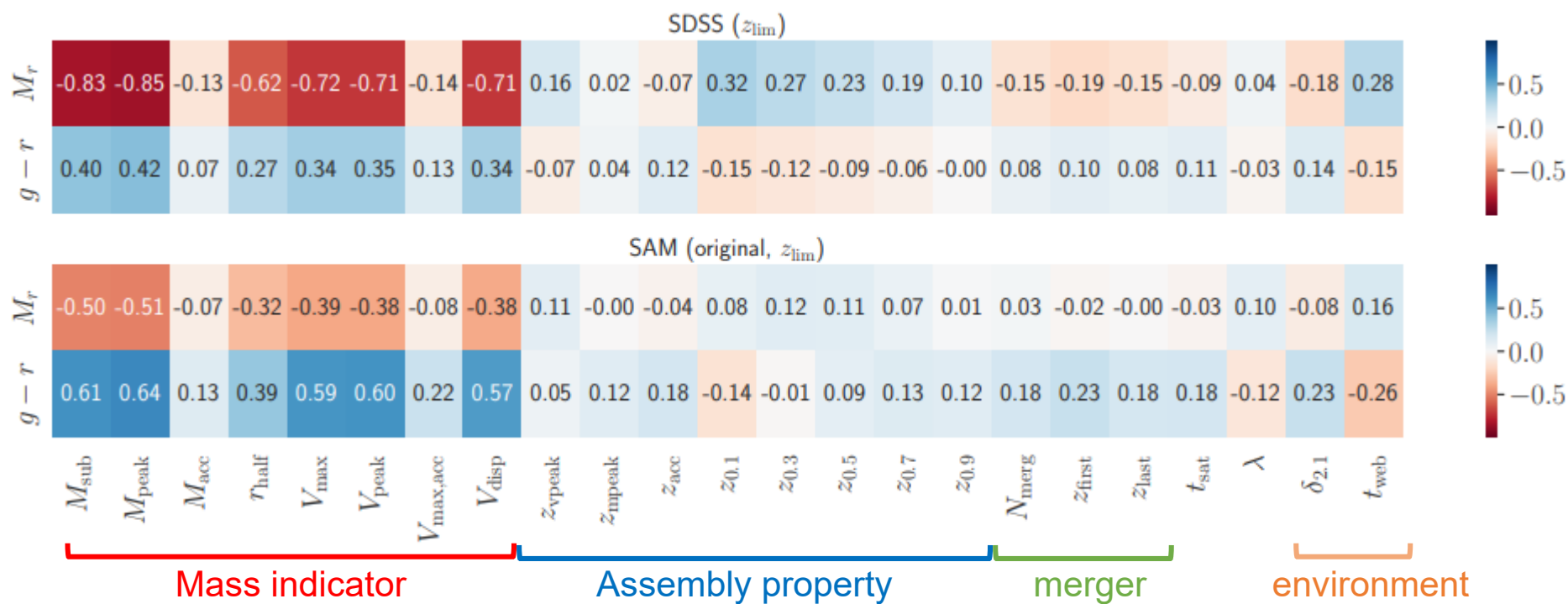
R^2 higher than SDSS even after shuffling in both SAM and TNG



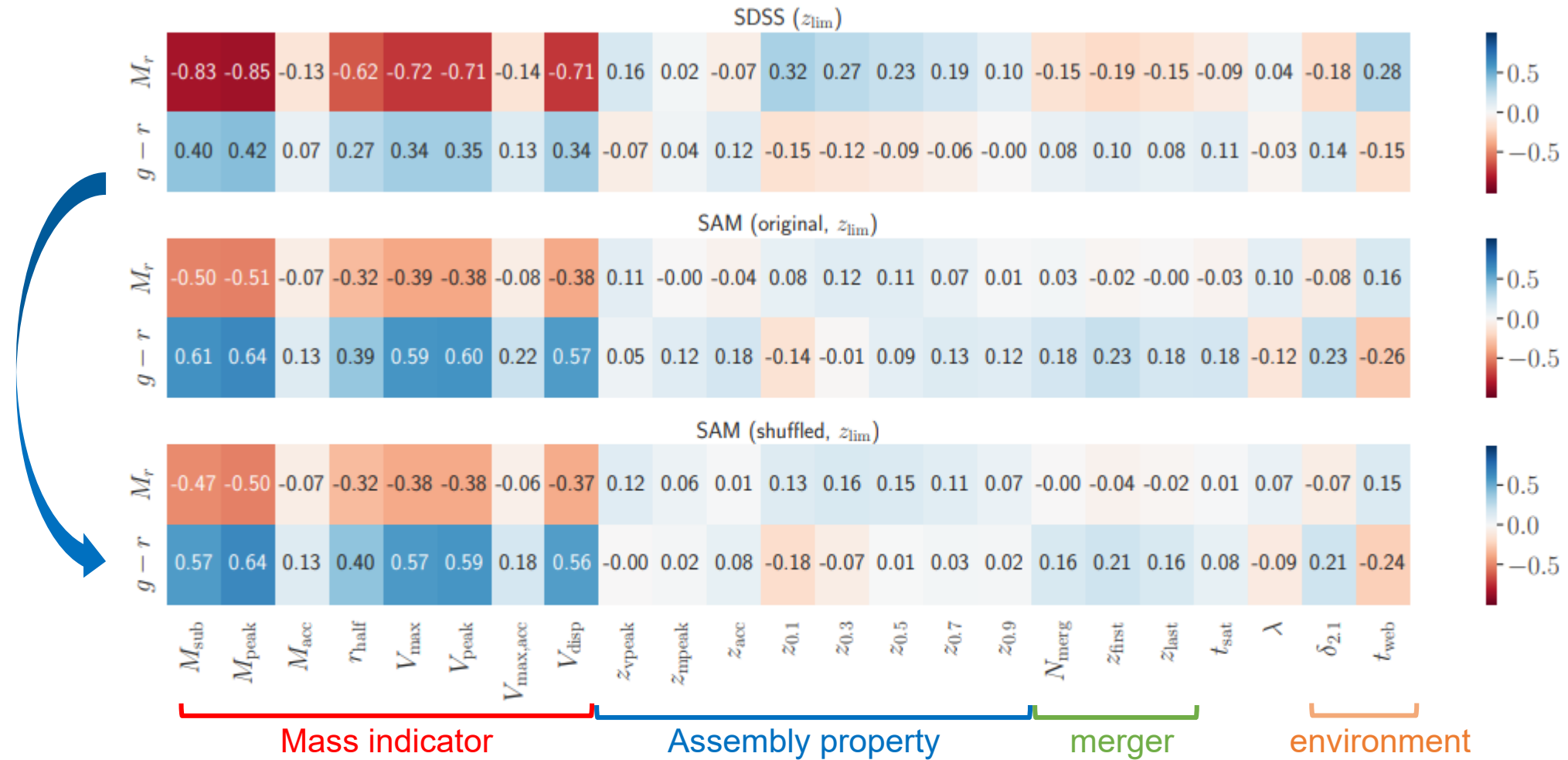
Pearson correlation coefficients



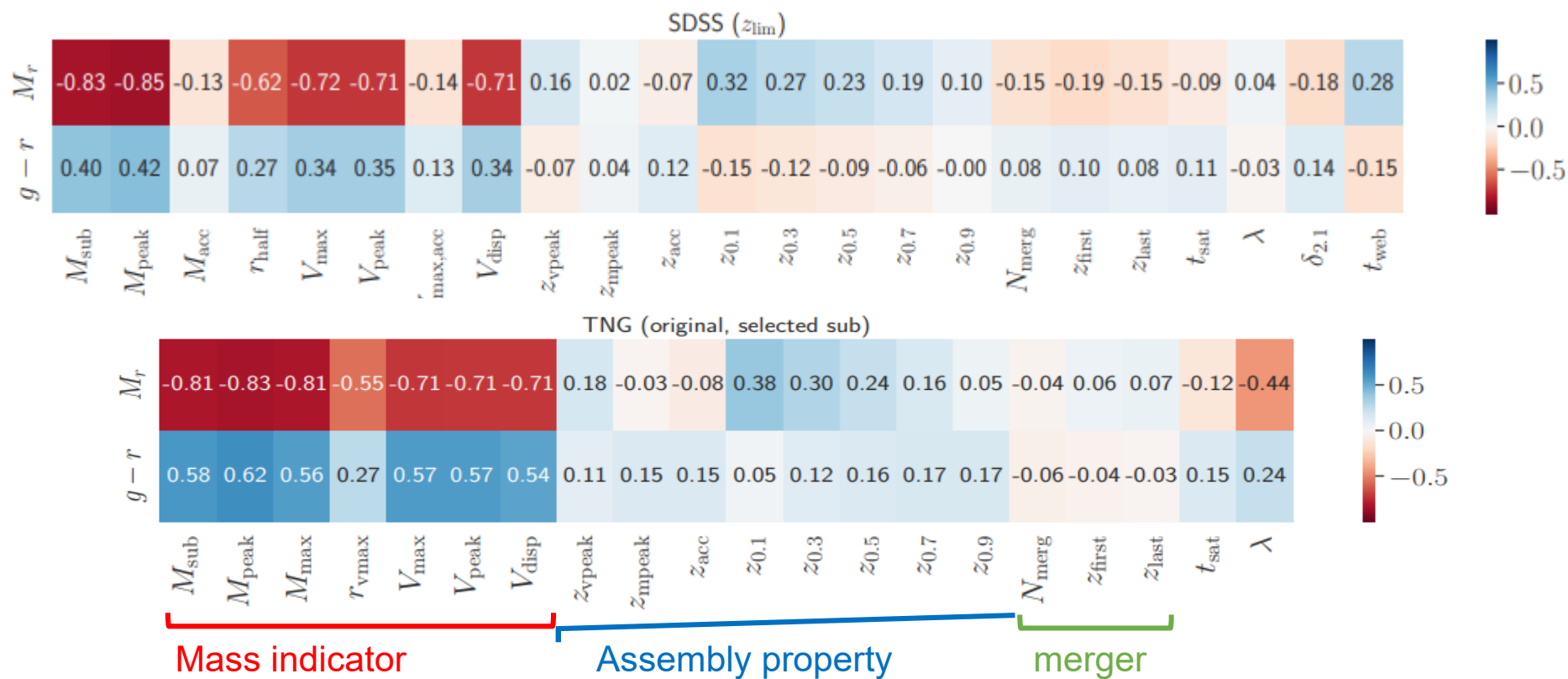
Pearson correlation coefficients: with SAM



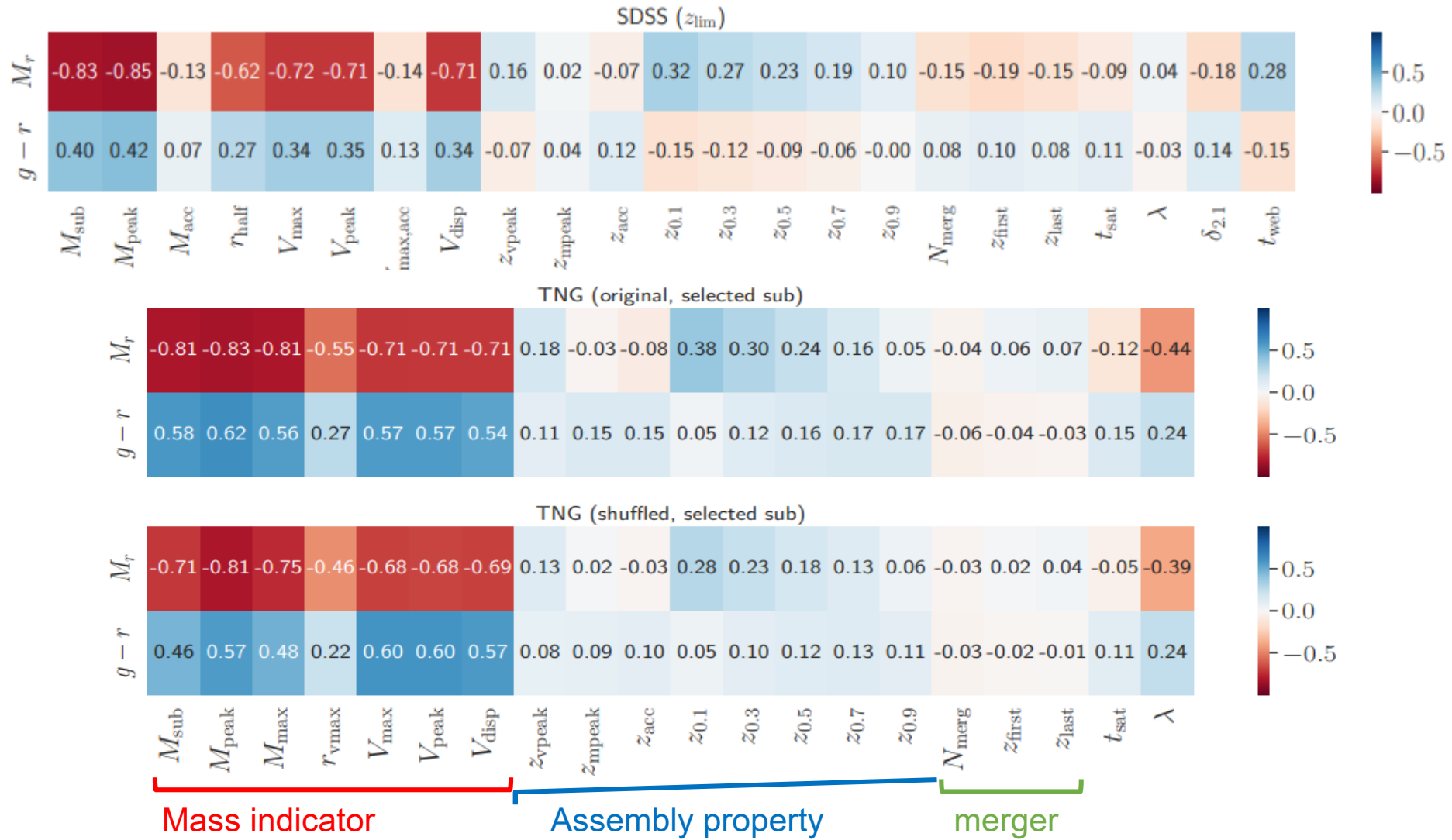
Pearson correlation coefficients: with SAM



Pearson correlation coefficients: with TNG



Pearson correlation coefficients: with TNG



Summary

- Study galaxy-subhalo relation in SDSS-ELUCID matched catalog
- Machine learning can predict M_r , but not color
- The mismatch may be a reason for the low performance of color prediction
- Mismatch effect has small impact on both SAM and TNG
- Connection between SDSS color and subhalo properties is weaker than SAM and TNG

