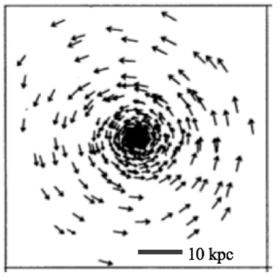
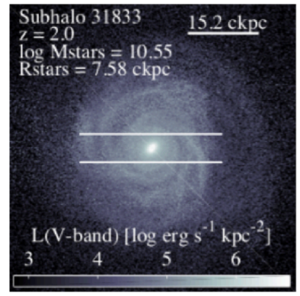
Stellar Property Statistics of Massive Halos: Common Kernel Shapes from Multiple Cosmo-Hydro Simulations

Evrard+ 1994

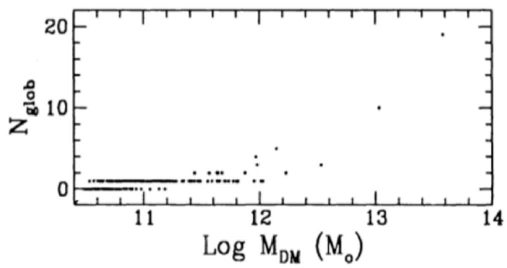


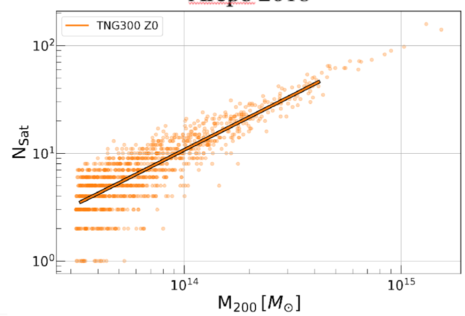
P3MSPH 1994



Pillepich+ 2018

Arepo 2018



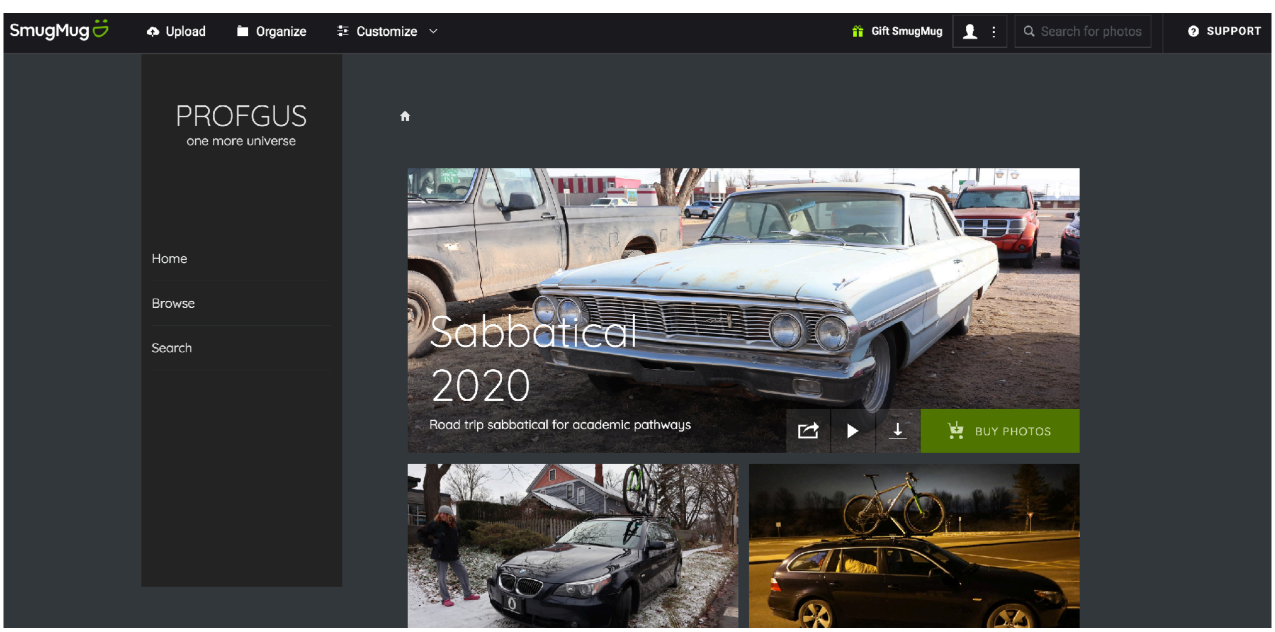


August (Gus) Evrard

Arthur F. Thurnau Professor
Departments of Physics and Astronomy
Leinweber Center for Theoretical Physics (LCTP)
University of Michigan

Road trip sabbatical!

https://profgus.smugmug.com/Sabbatical2020/



27-28 Jan 2020 - KIPAC + LBL Seminars – Gus Evrard, U Michigan

27-28 Jan 2020 - KIPAC + LBL Seminars – Gus Evrard, U Michigan

27-28 Jan 2020 - KIPAC + LBL Seminars – Gus Evrard, U Michigan

27-28 Jan 2020 - KIPAC + LBL Seminars – Gus Evrard, U Michigan

Quick overview of key findings

Verification/Consistency test

D. Anbajagane, A.E. Evrard, A. Farahi, D.J. Barnes,
K.Dolag, I.G. McCarthy, D.Nelson, A. Pillepich , arXiv:2001.02283

Dhayaa Anbajagane
Undergraduate, U Michigan

Arya Farahi
MIDAS Postdoc
U Michigan

Three cosmological simulations with AGN feedback:

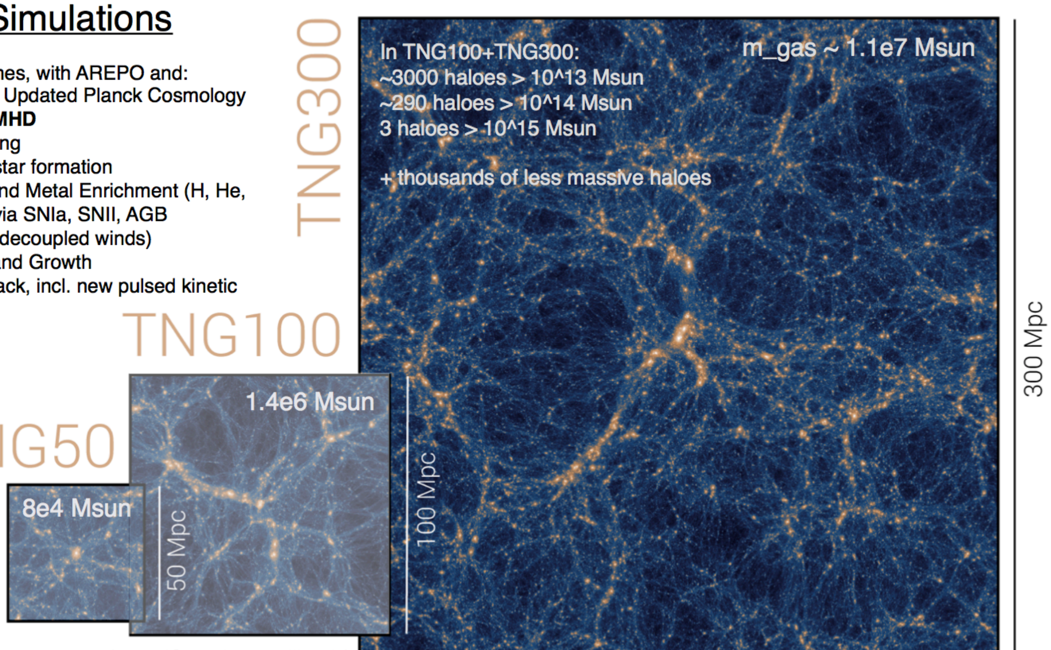
- IllustrisTNG (Arepo) : **Pillepich**+ 2018
- BAHAMAS + MACSIS (SPH-Gadget) : **McCarthy**+ (2017), **Barnes**+(2017)
- Magneticum (SPH-Gadget-X) : Ragagnin, **Dolag**+ (2019)

Large samples of halos with $M_{200}>10^{13.5}$ M_{Sun}

Simulation	N_{sam}	$\log_{10}(M_{20})$	L (Mpc)	Ω_m	Ω_b	f_b	σ_8
BM	9368	15.6	596	0.3175	0.049	0.154	0.834
MGTM	4202	14.9	500	0.2726	0.0456	0.167	0.809
TNG300-1	1130	14.6	302.6	0.3089	0.0486	0.157	0.8159

The TNG Simulations

Three flagship volumes, with AREPO and:
- Cosmological ICs: Updated Planck Cosmology
- GRAVITY+ **ideal MHD**
- Gas Cooling/Heating
- Threshold-based star formation
- Stellar Evolution and Metal Enrichment (H, He, Fe, C, Mg, …Eu) via SNIa, SNII, AGB
- Stellar Feedback (decoupled winds)
- Black Hole Seed and Growth
- Black Hole Feedback, incl. new pulsed kinetic **BH-driven winds**



(same ICs and res as Illustris)

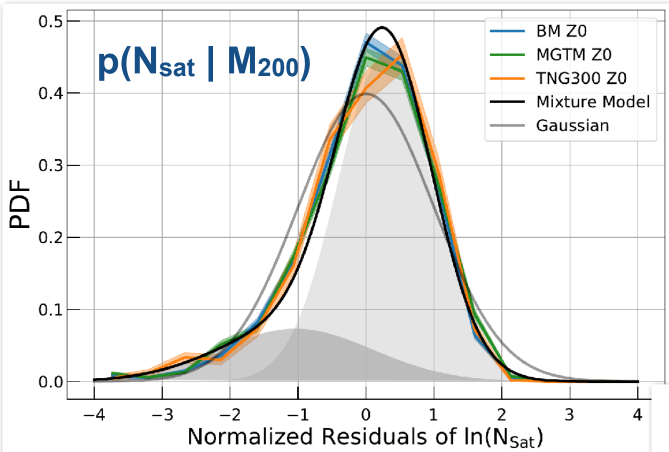
TNG Model:
Weinberger:2017
Pillepich:2018a

TNG100+TNG300:
Springel et al. 2018
Pillepich et al. 2018b
Naiman et al. 2018
Marinacci et al. 2018
Nelson et al. 2018

TNG50:
Nelson et al. 2019b
Pillepich et al. 2019

IllustrisTNG in one slide — Annalisa Pillepich for Gus Evrard

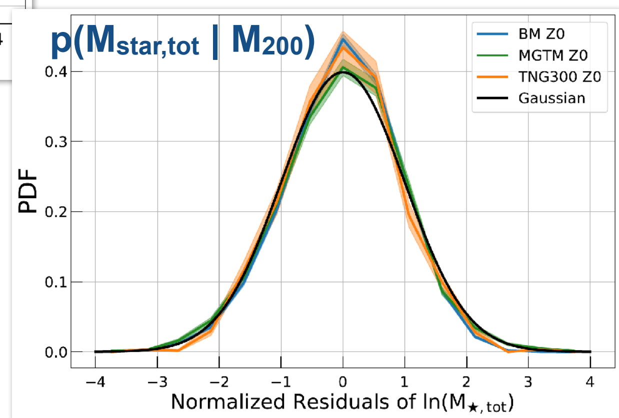
Main result: Consistent kernel shapes



N_{sat} = # satellite galaxies < R_{200} with M_{star} > 1e10 M_{sun}

Gaussian mixture (skew-normal)

Skewness driven by early forming halos



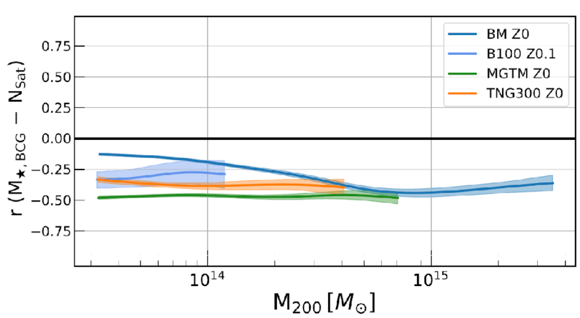
$M_{star,tot}$ = total stellar mass within R_{200}

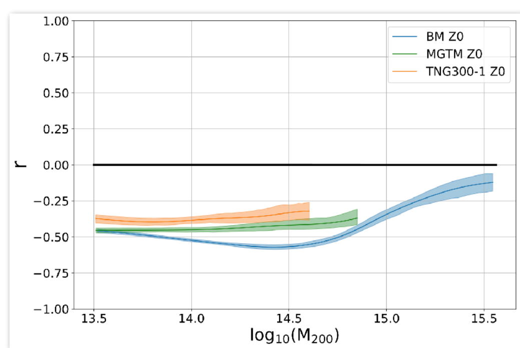
Nearly Gaussian PDF (in log M_{star})

Reflects central limit theorem?

Bonus: halo mass-conditioned correlations

correlation coefficients of properties (< R_{200}) @ z=0





$M_{star,BCG}$ and N_{sat} | M_{halo}

$M_{star,tot}$ and M_{gas} | M_{halo} @z=0

Why multi-simulation studies?

- **Verification exercise** for solutions of complex, non-linear models

  Do independent teams produce consistent statistical outcomes for large-scale structure and galaxy statistics?

- Estimate **theoretical uncertainties** in the non-linear regime

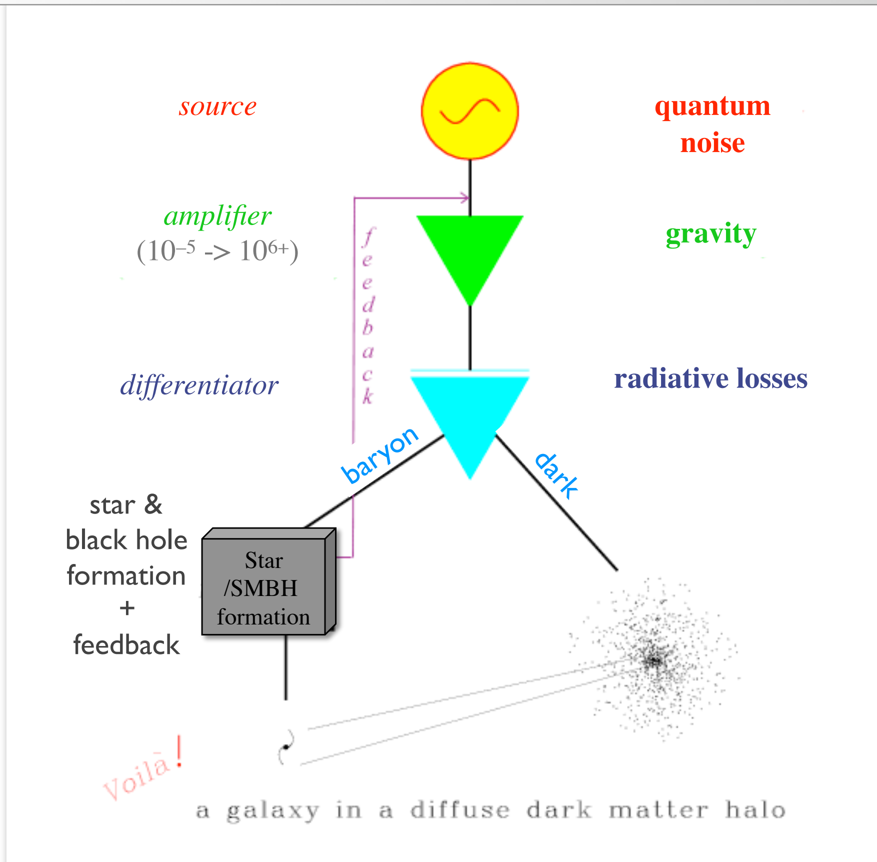
  Can we build more robust likelihoods for interpreting survey data?

- Establish high quality **reference models** (data compression)

  e.g., What is the form of halo multi-property (**S**) kernel, Pr(**S** | M, z)?

Galaxy formation: astrophysics in a cosmological context

cosmic engineering: a flowchart for galaxy formation



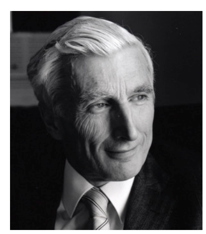
basic theoretical picture has been in place for nearly half a century

Mon. Not. R. astr. Soc. (1978) **183**, 341–358

Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering

S. D. M. White and M. J. Rees *Institute of Astronomy, Madingley Road, Cambridge*

Received 1977 September 26



Summary. We suggest that most of the material in the Universe condensed at an early epoch into small 'dark' objects. Irrespective of their nature, these objects must subsequently have undergone hierarchical clustering, whose present scale we infer from the large-scale distribution of galaxies. As each stage of the hierarchy forms and collapses, relaxation effects wipe out its substructure, leading to a self-similar distribution of bound masses of the type discussed by Press & Schechter. The entire luminous content of galaxies, however, results from the cooling and fragmentation of residual gas within the transient potential wells provided by the dark matter. Every galaxy thus forms as a concentrated luminous core embedded in an extensive dark halo. The observed sizes of galaxies and their survival through later stages of the hierarchy seem inexplicable without invoking substantial dissipation; this dissipation allows the galaxies to become sufficiently concentrated to survive the disruption of their halos in groups and clusters of galaxies. We propose a specific model in which $\Omega \simeq 0.2$, the dark matter makes up 80 per cent of the total mass, and half the residual gas has been converted into luminous galaxies by the present time. This model is consistent with the inferred proportions of dark matter, luminous matter and gas in rich clusters, with the observed luminosity density of the Universe and with the observed radii of galaxies; further, it predicts the characteristic luminosities of bright galaxies and can give a luminosity function of the observed shape.

what's involved in the "full physics" of galaxy formation? *Answer*: **A LOT!**

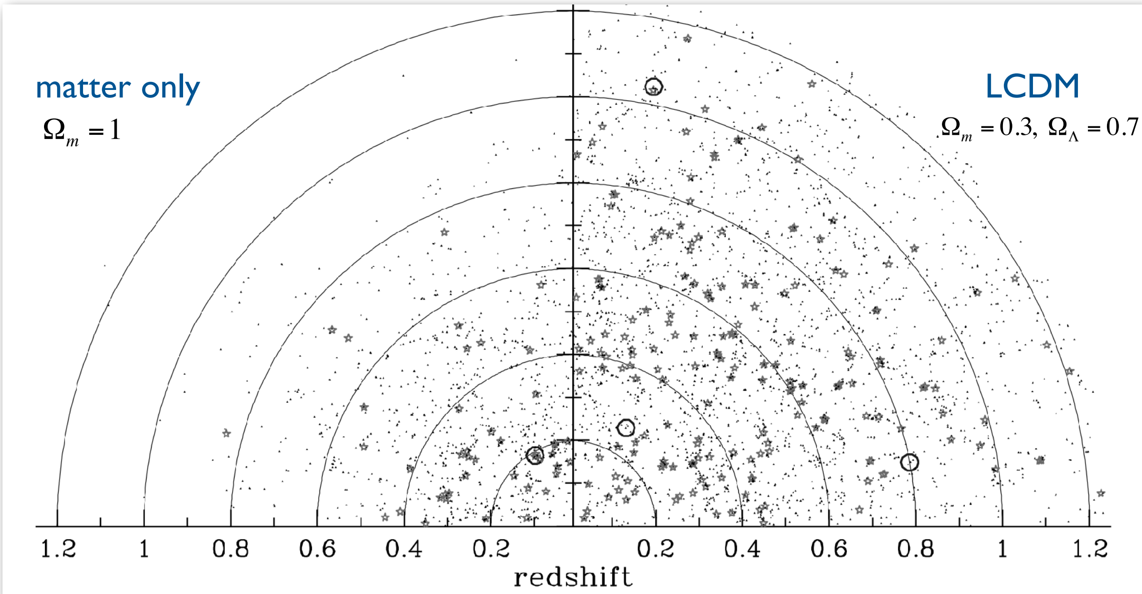
- gravity and background cosmology
  - solved with mature N-body simulation methods
- gas dynamics of (thermal) baryonic plasma
  - solved with (magneto-)hydrodynamic methods including treatment of shocks
  - require wide dynamic range and solvers that capture/resolve turbulence
- thermal conduction
  - can erase T gradients; sensitive to B-field structure
- interaction of baryons with a radiation field
  - optically thin gas in weak radiation field may **radiatively cool** gas
  - optically thick or strong radiation fields may heat/move gas (chemistry)
- star formation on resolved ("mesoscopic", 100pc ~1 kpc) scales
  - requires rules for converting gas into stars (and seed black holes)
  - requires an Initial Mass Function (IMF) and properties of evolved stellar populations
- supernova / supermassive black holes (SMBH) feedback
  - blast waves and jets drive winds that can unbind gas in small galaxies
  - metal production; turbulence; regulating cooling in cores of massive halos; cosmic rays
- molecular chemistry, dust
  - cools to low T, affects radiative opacity, …

Cluster cosmology
+ Massive Halo population models

counting experiment version of cluster cosmology

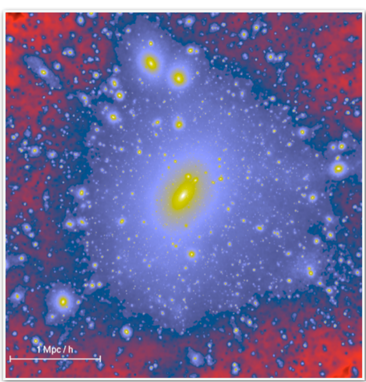
Evrard et al. (Virgo Consortium) 2002

- massive halos in lightcone outputs of Hubble Volume simulations (DM only)
  - symbols: log10(M) > 14 (dots), 14.5(stars), 15(circles)
  - virtual observer at bottom center of graphic (z=0)
  - P(k) amplitudes chosen to yield same "near field" population (z<0.2)



Clusters and Halos occupy **fundamentally different spaces**

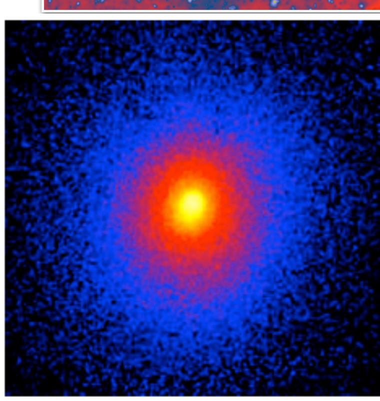
simulation

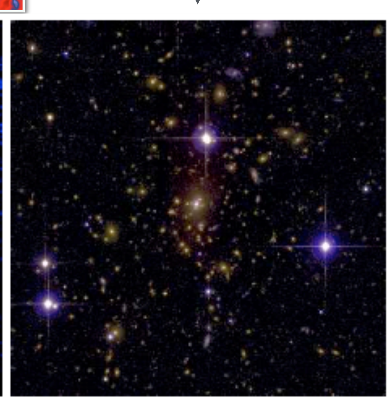


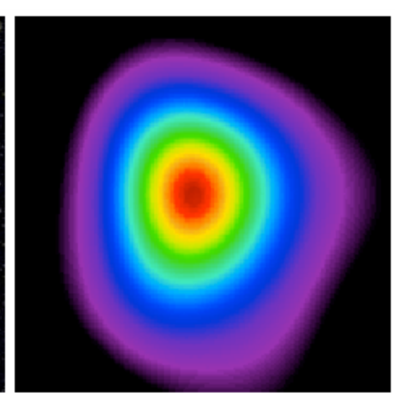
← dark matter halo in a N-body simulation
6D phase space

cluster Abell 1835 in multiple wavebands
2+1 sky + color/redshift space

Allen, Evrard + Mantz 2011







X-ray optical mm (SZ)

Property kernels map halo mass to observable properties, **S**

$$\mathbf{s} \equiv \ln \mathbf{S} \qquad \mu \equiv \ln M$$

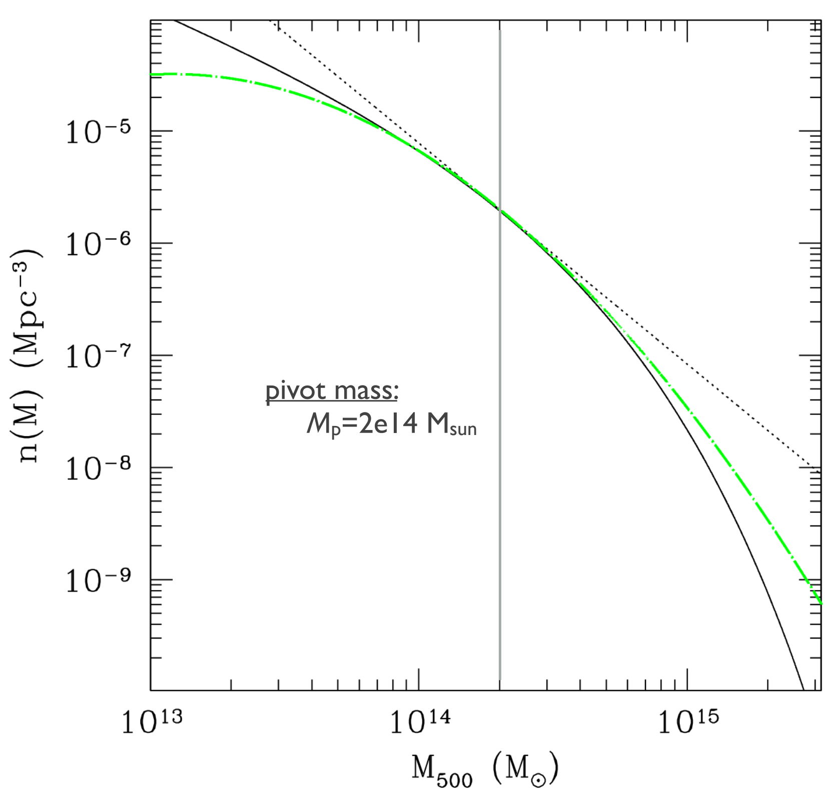
$$n(\mathbf{s}, z) = \int d\mu \; n(\mu, z) \; p(\mathbf{s} \,|\, \mu, z)$$

Mass function
Space density of halos
sensitive to cosmology

Property kernel
aka, Mass-Observable Relation (MOR)
sensitive to astrophysics

mass function shape: locally polynomial in log space

Evrard+ 2014



$$\log(n_1(M)) = \log(A) - \beta_1 \log(M/M_p) - \beta_2/2\,(\log(M/M_p))^2$$

Tinker MF (z=0.2)
~local **quadratic**

convolution with a
log-normal property kernel
is analytic

Simple form for the multi-property space density of halos

Evrard+ 2014

scaling relation $\langle \mathbf{s}|\mu\rangle = \boldsymbol{\pi} + \boldsymbol{\alpha}\mu$ log-means (vector)

$C_{ab} = \langle (s_a - \langle s_a|\mu\rangle)(s_b - \langle s_b|\mu\rangle)\rangle$ covariance of log-observables

joint, multi-observable space density

μ = ln *M*
s_i = ln S_i

$$n_1(\mathbf{s}) = A_1' \exp\left[-\frac{1}{2}\left((\mathbf{s}-\boldsymbol{\pi})^T \mathbf{C}^{-1}(\mathbf{s}-\boldsymbol{\pi}) - \frac{\langle\mu|\mathbf{s}\rangle_1^2}{\sigma^2_{\mu|\mathbf{s},1}}\right)\right]$$

$$A_1' = A\,\sigma_{\mu|\mathbf{s},1}\,\left((2\pi)^{N-1}|\mathbf{C}|\right)^{-1/2}$$

$$\langle\mu|\mathbf{s}\rangle_1 = \left[\boldsymbol{\alpha}^T \mathbf{C}^{-1}(\mathbf{s}-\boldsymbol{\pi}) - \beta_1\right]\sigma^2_{\mu|\mathbf{s},1}$$ **mean** selected mass

$$\sigma^2_{\mu|\mathbf{s},1} = (\boldsymbol{\alpha}^T \mathbf{C}^{-1}\boldsymbol{\alpha})^{-1}.$$ **variance** in selected mass

scaling behavior of property-selected samples

Evrard+ 2014

log-normal form of property-conditioned samples, $\Pr(s_b \mid s_a)$

$$\langle s_b|s_a\rangle_1 = \pi_b + \alpha_b[\,\langle\mu|s_a\rangle_1 + \beta_1\, r_{ab}\,\sigma_{\mu|a,1}\,\sigma_{\mu|b,1}\,]$$ **mean** s_b for s_a selection

$$\sigma^2_{b|a,1} = \alpha_b^2\,[\,\sigma^2_{\mu|a,1} + \sigma^2_{\mu|b,1} - 2r_{ab}\,\sigma_{\mu|a,1}\sigma_{\mu|b,1}\,]$$ **variance** ”

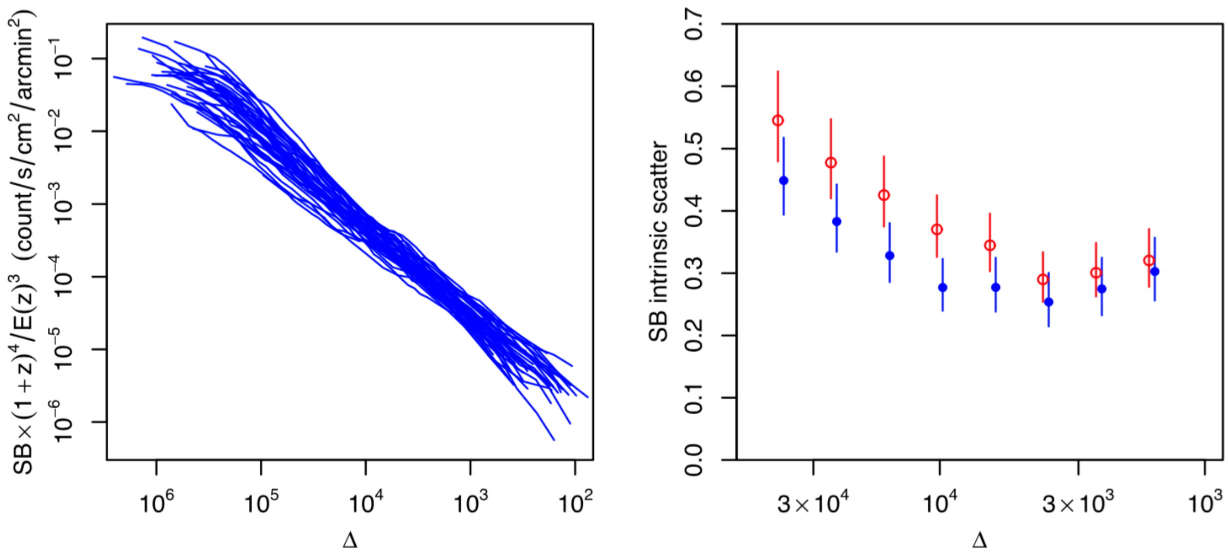
μ=ln M
s_i=ln S_i

These are what we actually measure in survey samples!

self-similarity in Tx-selected sample of relaxed clusters

Mantz et al, 2014, 2016

sample of 40 ***dynamically relaxed***, hot (kT>5 keV) clusters spanning 0 < z < 1

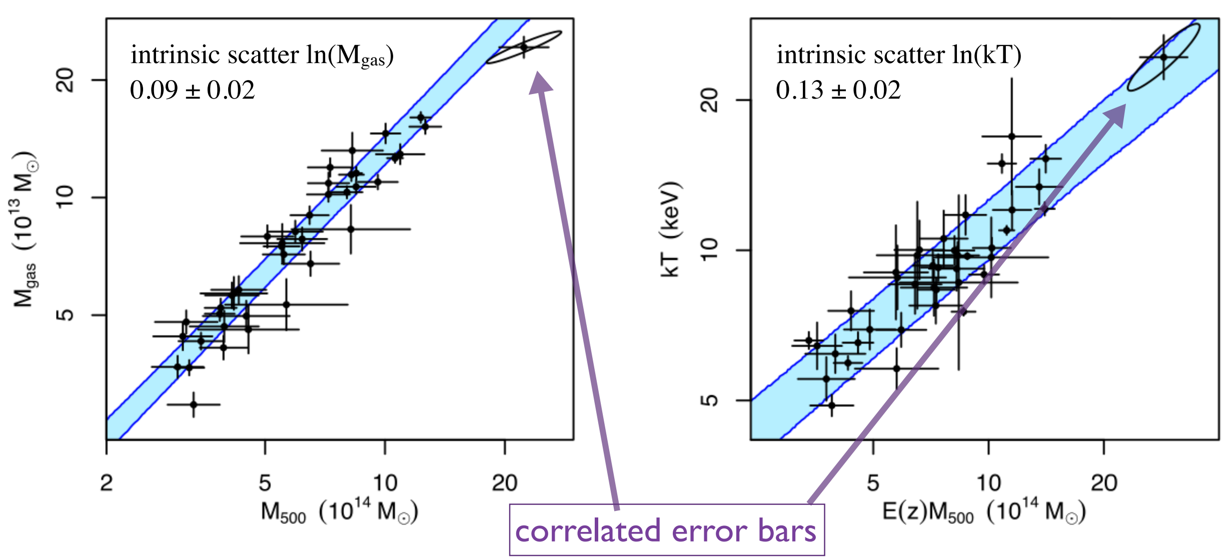


Scalings analysis of relaxed clusters

Mantz et al, 2014, 2016

– scaling of Mgas and Tx are consistent with self-similar expectations

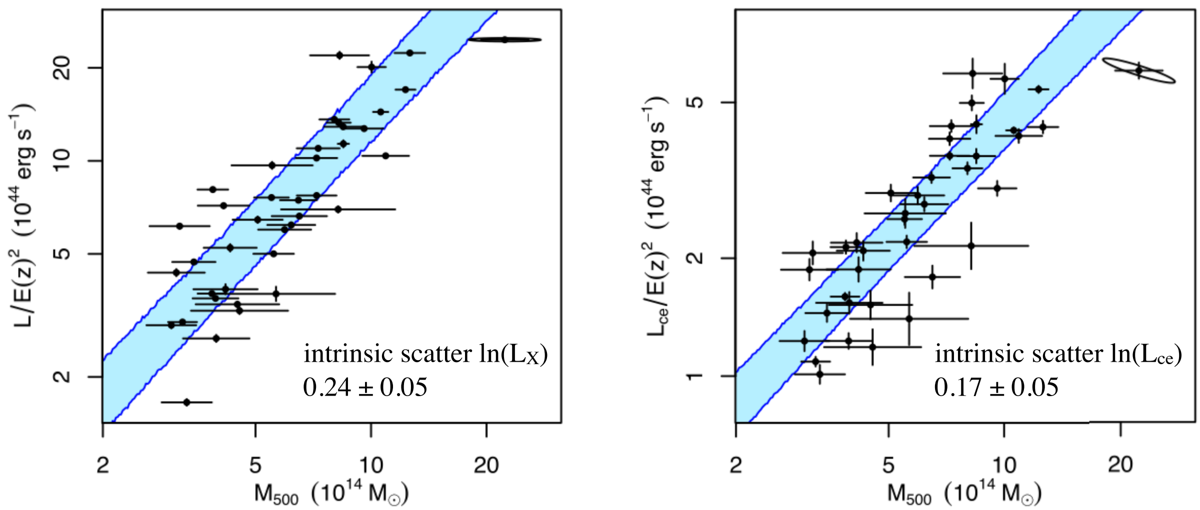
– total masses, M500, are X-ray hydrostatic estimates (~OK for relaxed systems?)



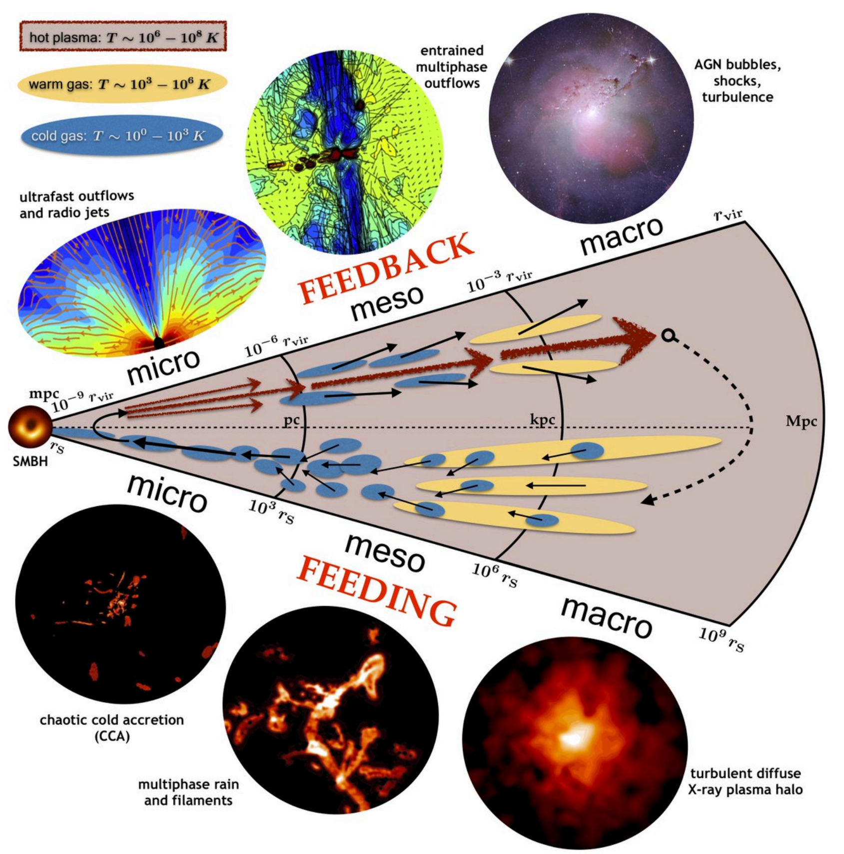
MOR analysis of relaxed clusters

Mantz et al, 2016

– Lx looks better when you put your thumb over the core



Cluster core physics: precipitation/chaotic, cold accretion driven by SMBH

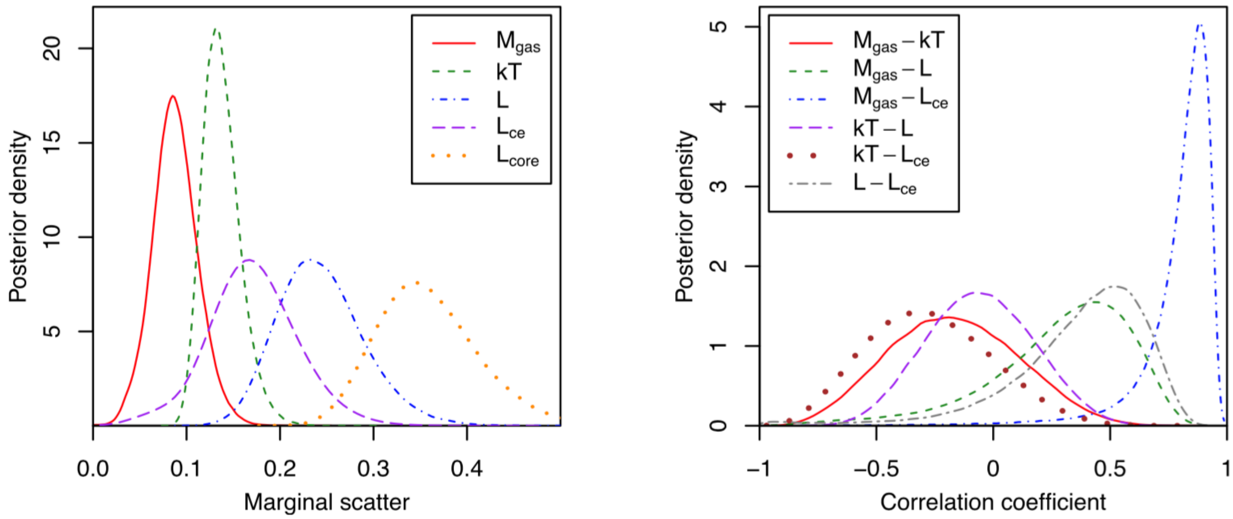


Max Gaspari
(via Twitter)

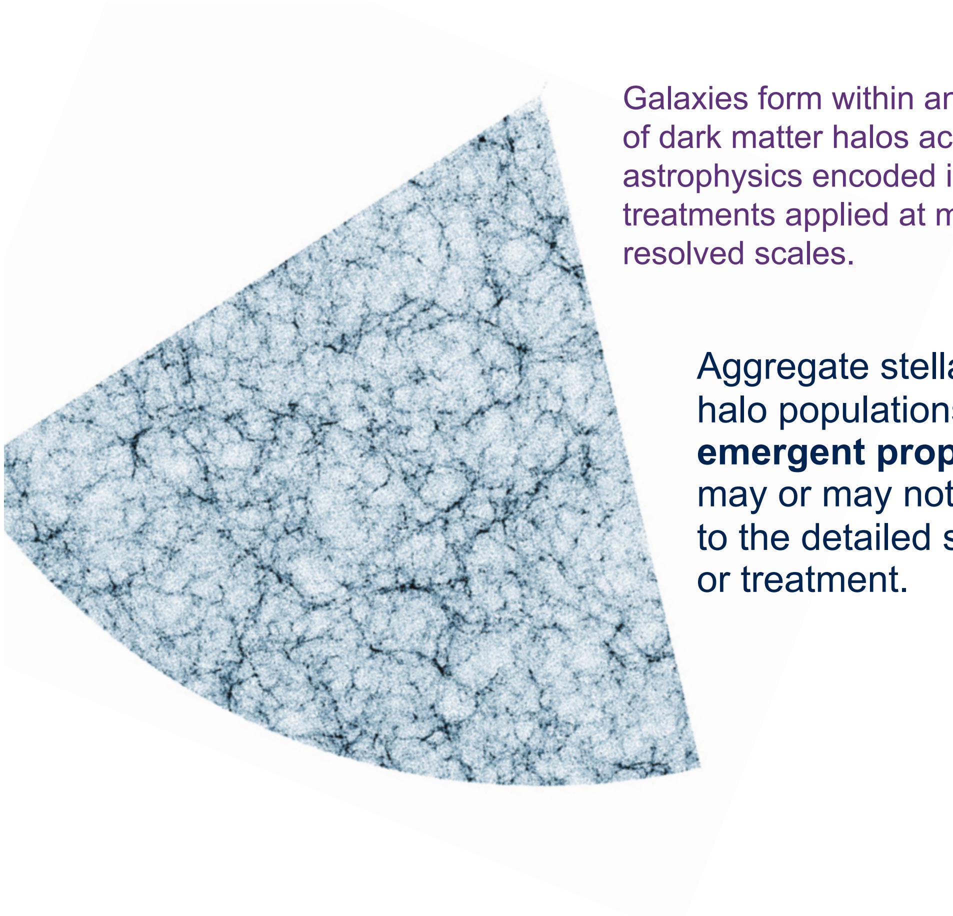
Scaling analysis of relaxed clusters

Mantz et al, 2016

– intrinsic scatter in $\ln(S_i)|M_{500}$; $\boldsymbol{S} = \{ M_{gas}, kT, L, L_{ce}, L_{core} \}$; gas mass is best
– first estimates of property covariance at fixed total mass



Local linear regression (LLR) and property kernels

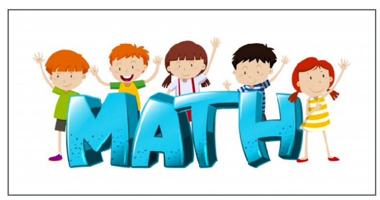


Galaxies form within an evolving web of dark matter halos according to astrophysics encoded in “sub-grid” treatments applied at minimally resolved scales.

Aggregate stellar statistics of halo populations are thus **emergent properties** that may or may not be sensitive to the detailed sub-grid scale or treatment.

Kernel shape via Local Linear Regression (LLR)

- “**Local**” scaling with halo mass of a property, *S*a, by **mass-weighted linear regression**
- Marginalize over mass yields kernel shape



Population Model mean: (local intercept; local slope)

$$\langle s_a \mid M, z \rangle = \pi_a(M_c, z) + \alpha_a(M_c, z) \ln(M/M_c),$$

Local linear regression: minimize weighted dev’s

$$\epsilon_a^2(M,z) = \sum_{i=1}^{n} w_i^2 \left(s_{a,i} - \alpha_a(M,z)\mu_i - \pi_a(M,z)\right)^2,$$

Mass-dependent weight

$$w_i = \frac{1}{\sqrt{2\pi}\sigma_{\rm LLR}} \exp\left\{-\frac{\mu_i^2}{2\sigma_{\rm LLR}^2}\right\},$$

width in log mass 0.2 dex (=0.46 in ln M)

Local deviations

$$\delta s_{a,i} \equiv s_{a,i} - \widetilde{\pi}_a(M_i, z).$$

Local normalized deviations

$$\widehat{\delta s}_{a,i} \equiv \frac{\delta s_{a,i}}{\widetilde{\sigma}_a} = \frac{s_{a,i} - \widetilde{\pi}_a(M_i, z)}{\widetilde{\sigma}_a}$$

Simpler version (no math)…

- Art by chainsaw

- Art by paintbrush

Simulation characteristics

Table 1. Simulation characteristics and $z = 0$ halo population sizes. Empirical sources for tuning sub-grid parameters are given in the last column and consist of the Galaxy Stellar Mass Function (GSMF), SuperMassive Black Hole scaling (SMBH), Metallicity scaling (Metals), and cluster hot gas mass fraction $< R_{500c}$ (CL $f_{\rm gas}$). All assume a flat ΛCDM cosmology, so $\Omega_\Lambda = 1 - \Omega_m$, and their cosmic mean baryon fraction is $f_b \equiv \Omega_b/\Omega_m$. The MGTM output is actually $z = 0.03$ and B100 (used only for testing resolution) is $z = 0.12$ while the other two are exactly $z = 0$. See text for references.

Simulation$^\diamond$	L [Mpc]*	Ω_m	f_b	σ_8	$\epsilon_{DM}^{z=0}$ [kpc]	$m_\star$ [M$_\odot$]$^\spadesuit$	$\log_{10}(M_{20})^\heartsuit$	$N_{\rm sam}$†	Calibration
BM	596	0.3175	0.154	0.834	5.96	1.2×10^9	15.6	9430	GSMF, CL $f_{\rm gas}$
B100	143	0.2793	0.166	0.821	2.86	1.4×10^8	14.1	96	GSMF, CL $f_{\rm gas}$
MGTM	500	0.2726	0.167	0.809	5.33	5.0×10^7	14.9	4207	SMBH, Metals, CL $f_{\rm gas}$
TNG300	303	0.3089	0.157	0.8159	1.48	1.1×10^7	14.6	1146	See Pillepich et al. (2018a)

$^\diamond$ See text for description of acronyms.

* Comoving simulation cube length except for MACSIS (subset of the BM data), which subsamples a 3.2 Gpc cubic volume.

$^\spadesuit$ Initial stellar particle mass.

$^\heartsuit$ Upper limit of LLR regression at $z = 0$, the 20th most massive halo mass, in M$_\odot$.

† Number of halos with total mass, $M_{200c} > 10^{13.5} M_\odot$. The number above $10^{13.8}$ M$_\odot$ for BM is ≈ 4400.

Our focus is on halos hosting groups and clusters of galaxies

$$M_{200c} > 10^{13.5} M_\odot.$$

Halos and their stellar properties

Table 3. Property components of LLR regression vector, **S**.

Symbol	Quantity
$N_{\rm sat}$	Count of non-central galaxies within R_{200c} †
$M_{\star,\rm tot}$	Total stellar mass within R_{200c} ($M_{\odot}$).
$M_{\star,\rm BCG}$	Central galaxy stellar mass within 100 kpc ($M_{\odot}$).

† Stellar mass-limited, $M_{\star} > \left(\frac{f_{\rm b,\,Sim}}{f_{\rm b,\,MGTM}}\right) \times 10^{10}\, M_{\odot}$.

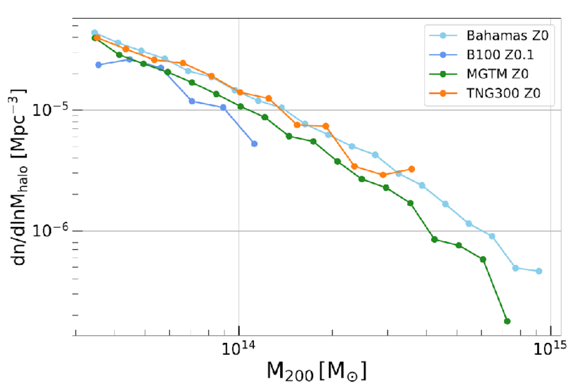
Table 2. Property definitions.

Quantity	Definition
Halo center	Minimum gravitational potential*
Halo total mass, M_{200c}	All species within R_{200c} sphere ♡
Halo stellar mass, $M_{\star,\rm tot}$	All stellar particles within R_{200c}
Galaxy center	Most bound particle†
Galaxy stellar mass, $M_{\star}$	Gravitationally-bound stellar particles

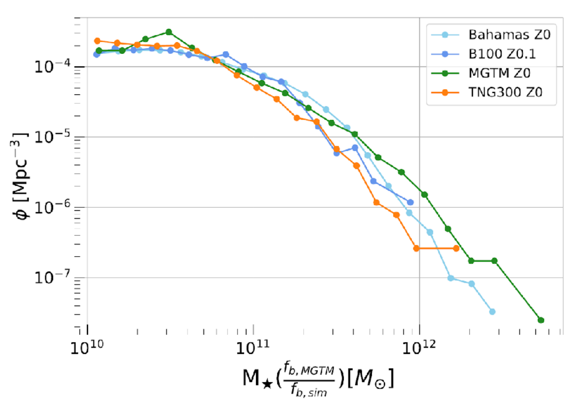
* FOF links of 0.2 (BM, TNG) and 0.16 (MGTM) mean separation.
♡ TNG halo masses use only FOF particle set.
† Of any species (BM, TNG) or collisionless only (MGTM).

"Mass functions"

Halo mass function

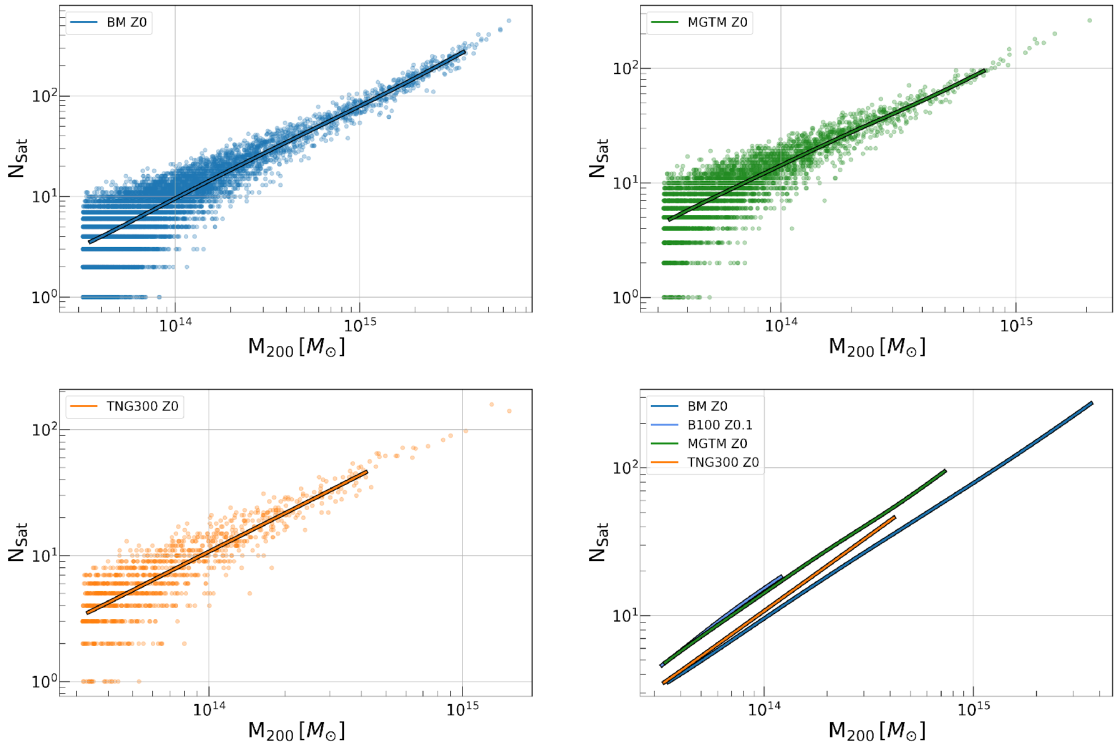


Galaxy stellar mass function
(high mass halos only)

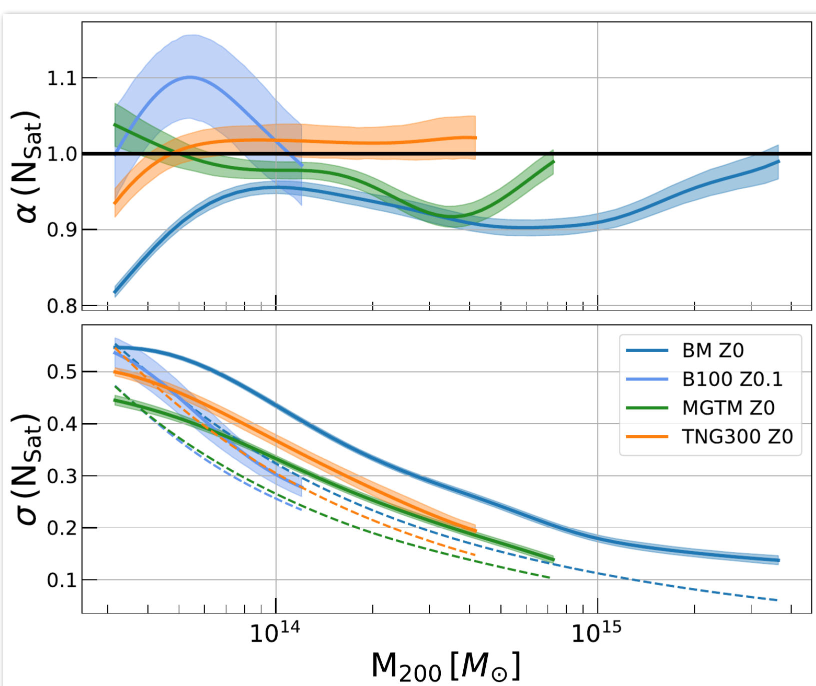


Satellite HOD statistics

N_{sat} = # satellite galaxies with M_{star}>1e10 M_{sun} and R<R_{200}



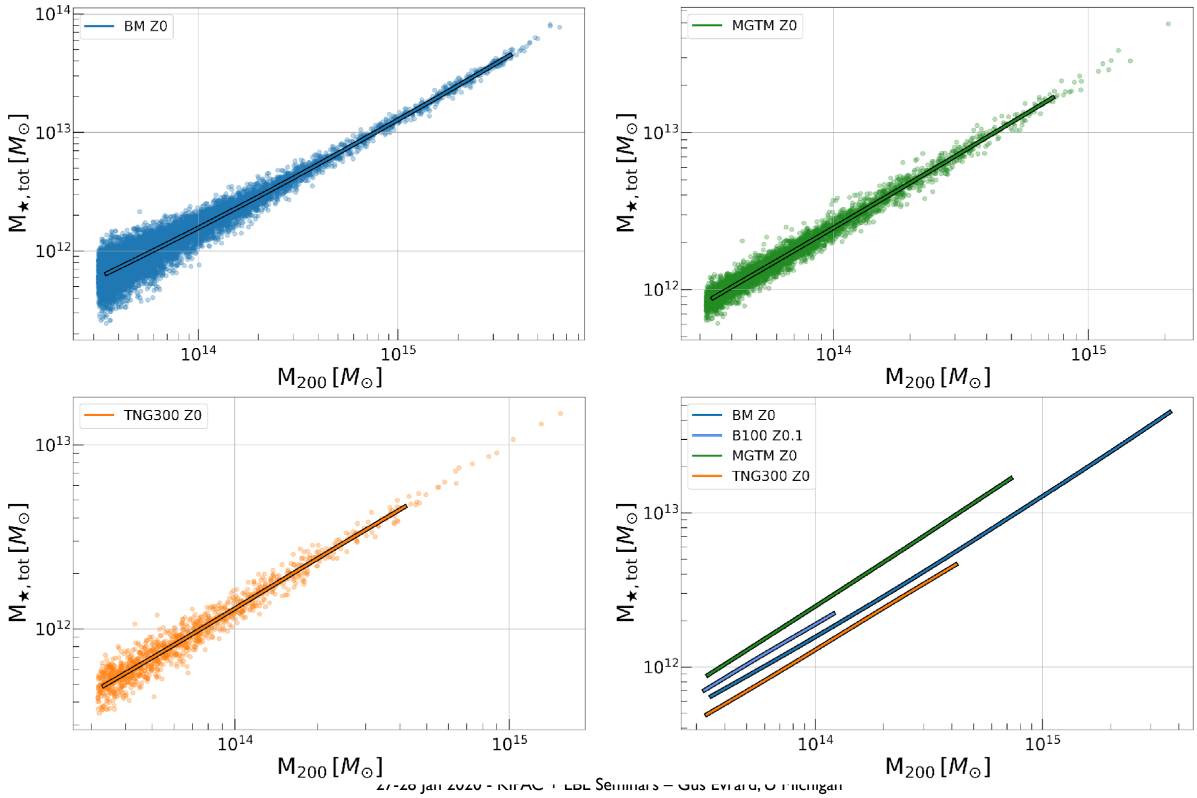
N_{sat} (<R_{200}) mass-dependent slope+scatter @z=0



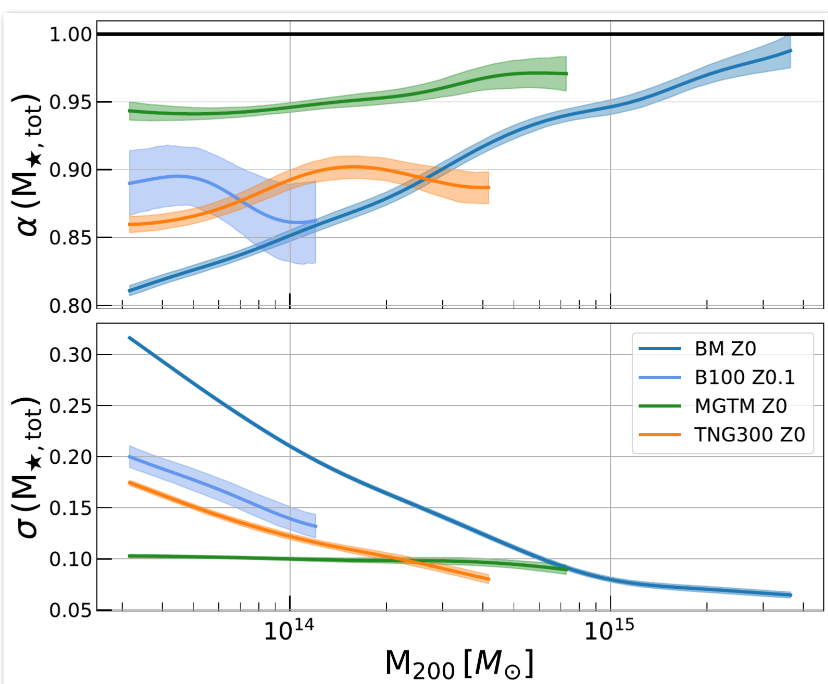
- Statistical (bootstrap) uncertainties smaller than systematic differences
- Absolute differences are at ~10% level
- All models see the scatter in lnN_{sat} declining with mass, slightly super-Poisson in amplitude (numerical resolution?)

Total stellar mass statistics

$M_{star,tot}$ is total stellar mass within R_{200}



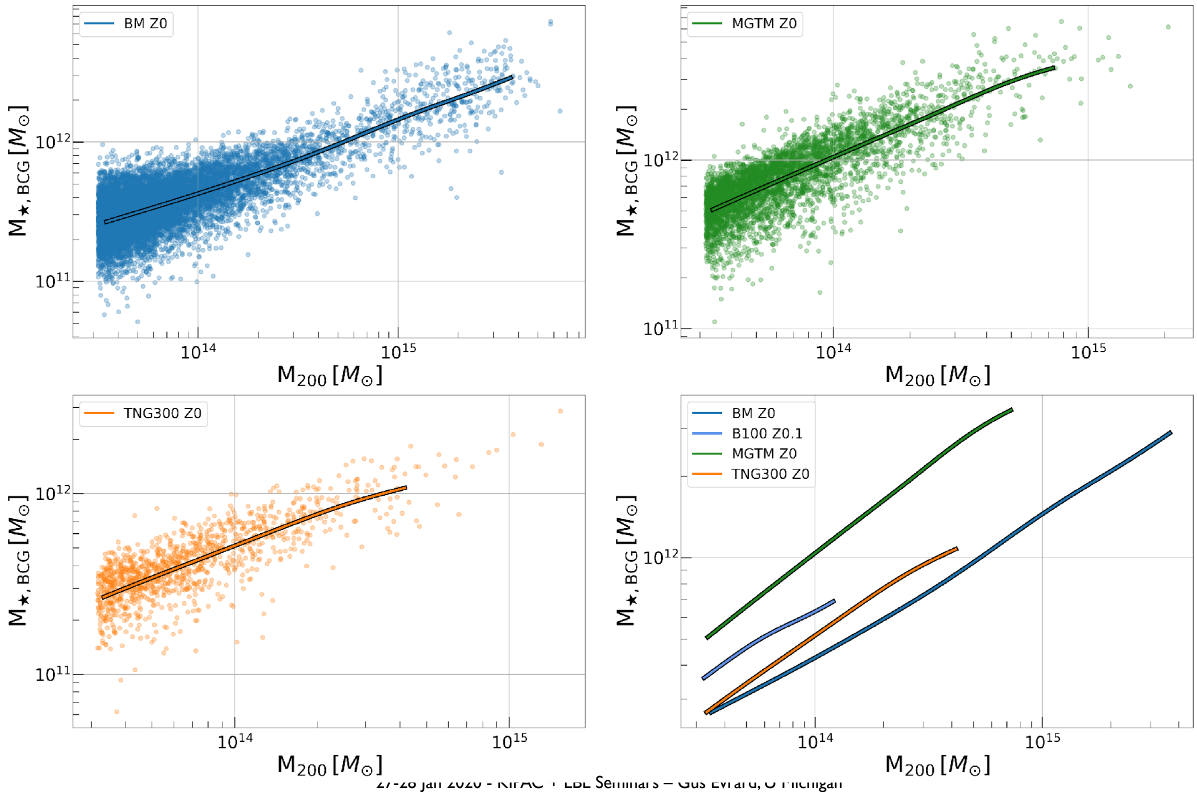
M_{star} ($<R_{200}$) slope+scatter @z=0



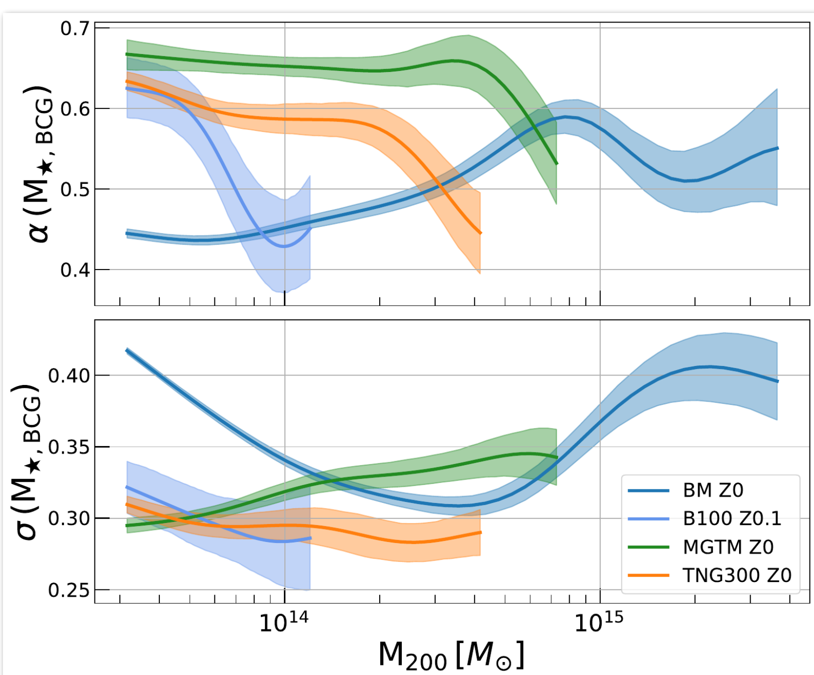
- Consistent sub-linear slopes
- Behaviors vary (MGTM=constant)
- BM scatter enhanced by poor stellar mass resolution (~12 particles at 1e10 limit)
- Scatter below 10% expected for M_{halo}>2e14

BCG stellar mass statistics

$M_{star,BCG}$ is total stellar mass within 100 kpc

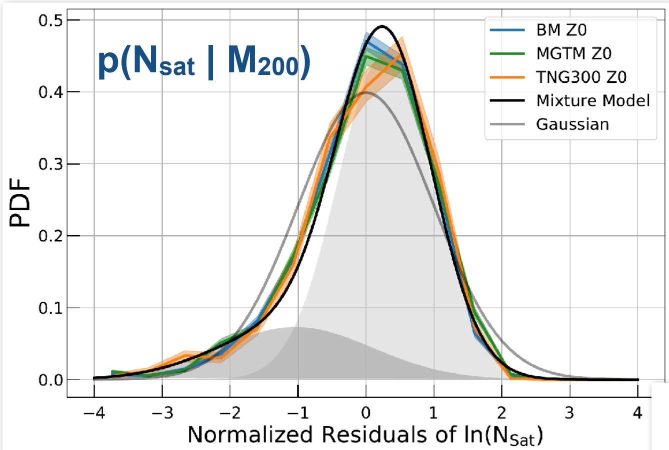


$M_{star,BCG}$ (<100 kpc) slope+scatter @z=0

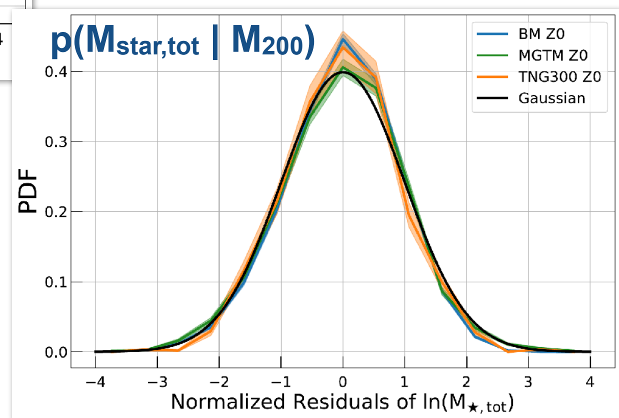


- Highly sub-linear slopes
- 30±5 % scatter near 10^{14} Msun

Consistent kernel shapes!

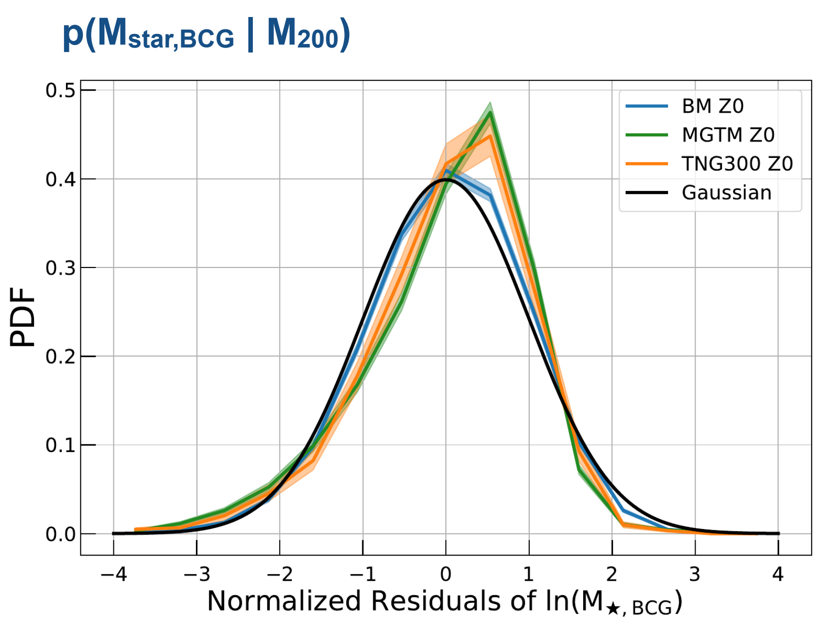


Skew-normal / Gaussian mixture PDF



Nearly Gaussian PDF (in log M_{star})

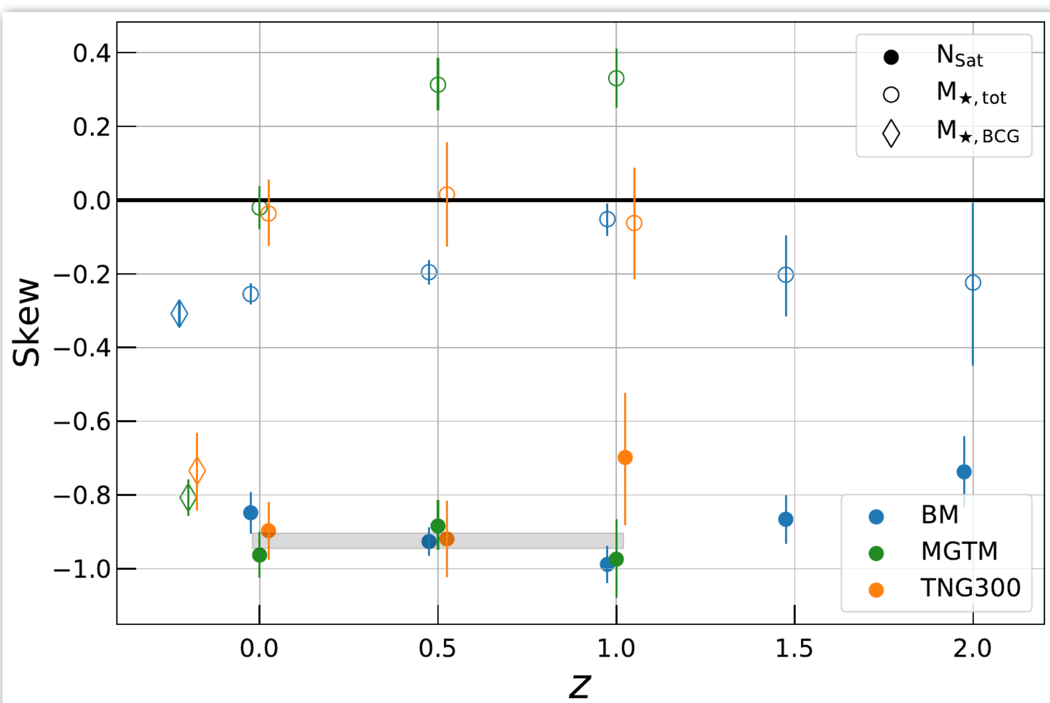
BCG kernel @z=0: probably skew?



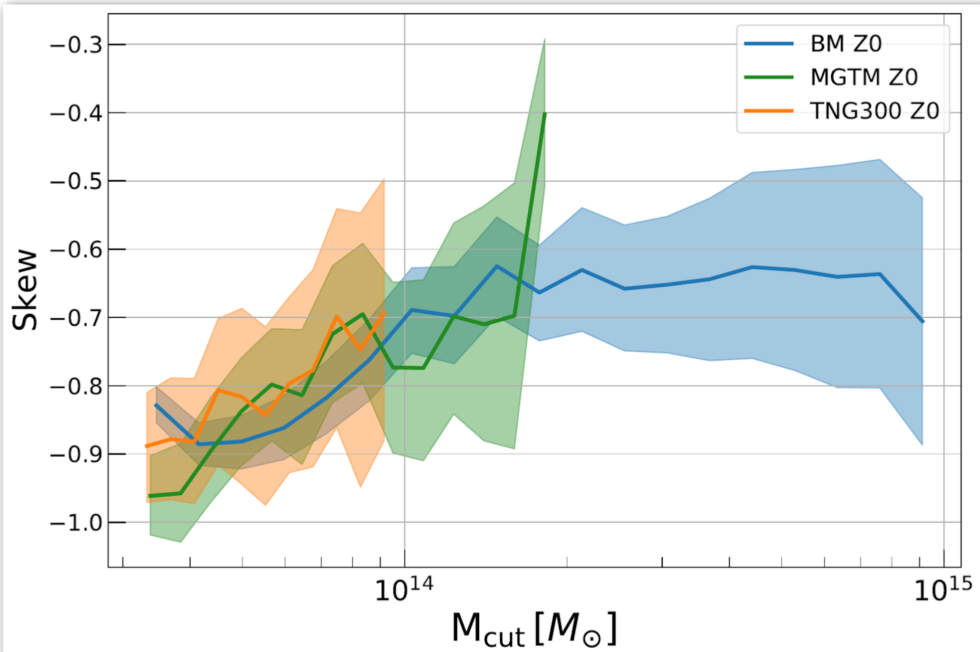
N_{sat} skewness and Gaussian mixture model kernel fit

Skewness

$$\gamma = \frac{\left\langle \left(x - \langle x \rangle\right)^3 \right\rangle}{\sigma_x^3}$$



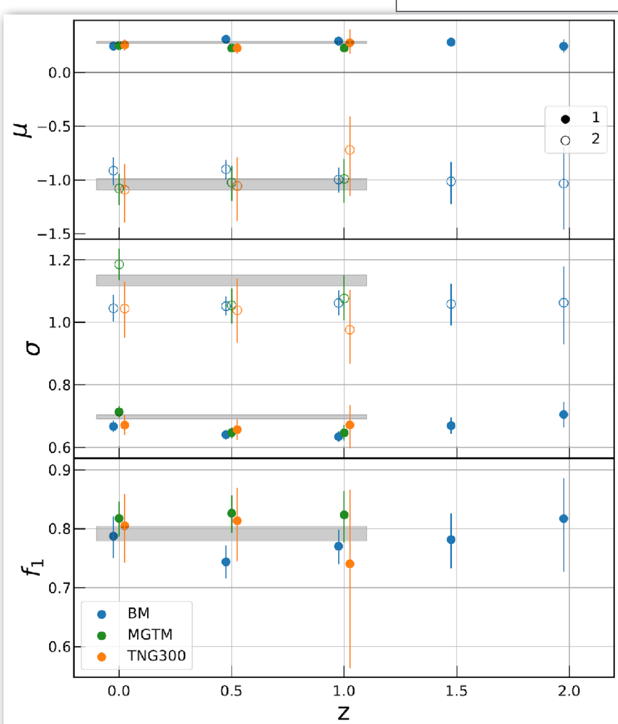
N_{sat} skew: weak mass dependence



Halo mass cutoff

N_{sat} kernel: Gaussian mixture fit

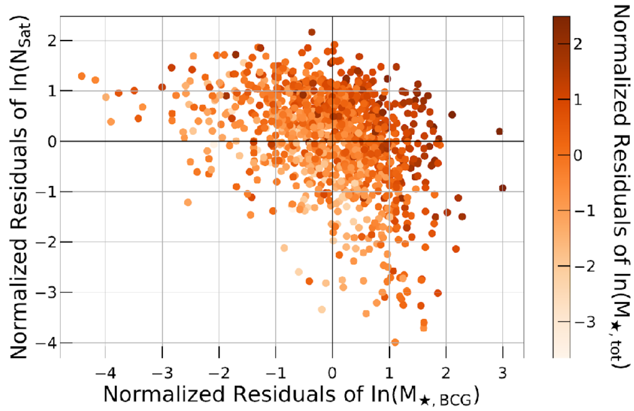
$$\Pr(x) = f_1\, G(x - \mu_1, \sigma_1) + (1 - f_1)\, G(x - \mu_2, \sigma_2),$$



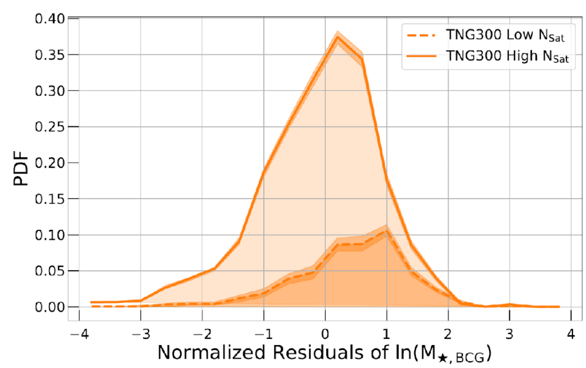
Parameter	Value
γ	-0.91 ± 0.02
f_1	0.79 ± 0.01
μ_1	0.28 ± 0.01
μ_2	-1.04 ± 0.05
σ_1	0.68 ± 0.01
σ_2	1.13 ± 0.02

- 80% of population is contained in a Gaussian component with variance roughly half (0.68)^2 that of the full population

Gaussian mixture application



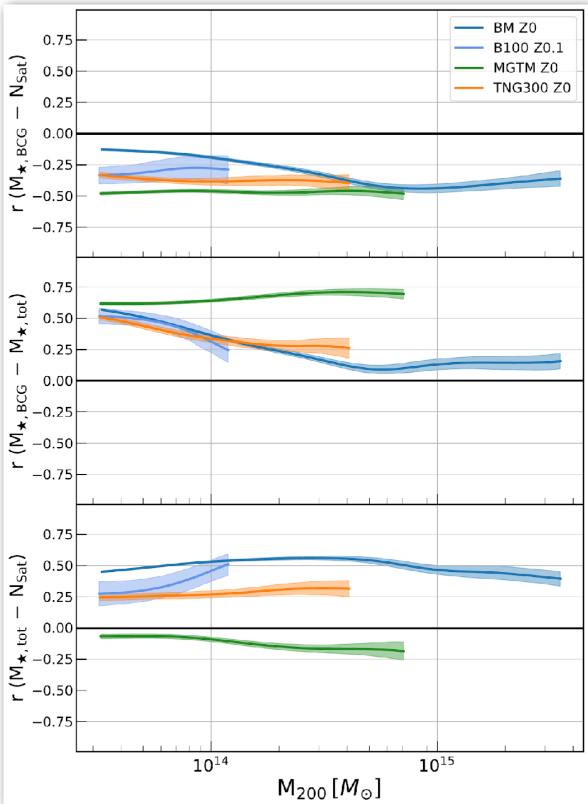
Important sampling the two components yields shaded PDFs in BCG stellar mass



Stellar property correlations

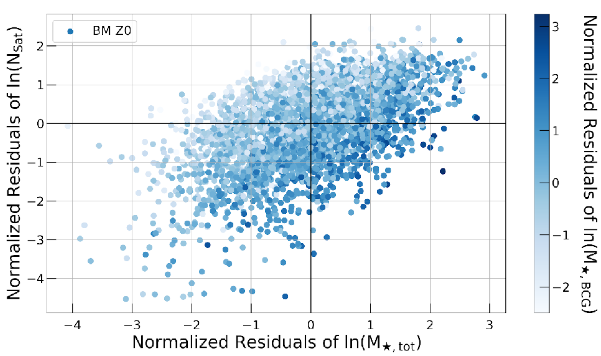
caveat: non-Gaussian shapes…

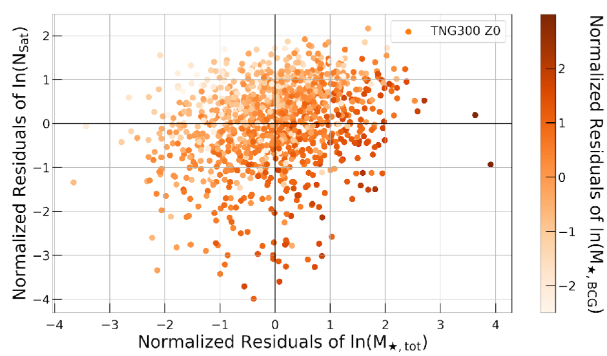
LLR correlation coefficients @ fixed M_{halo}

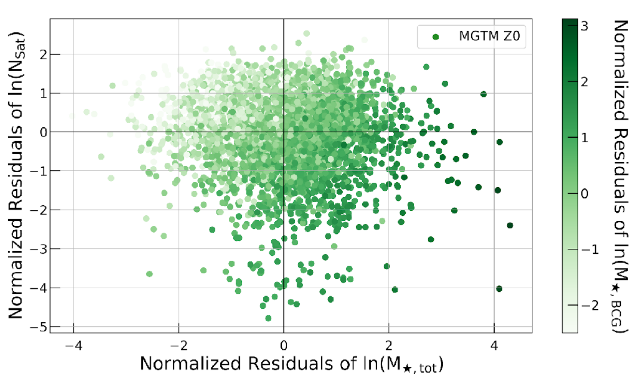


- Anti-correlation between BCG stellar mass and N_{sat} count
- Positive correlation of BCG stellar mass and total halo stellar mass (MGMT strongest)
- Mixed results for total stellar mass and N_{sat} count (MGMT is outlier)

Normalized mass-conditioned residuals

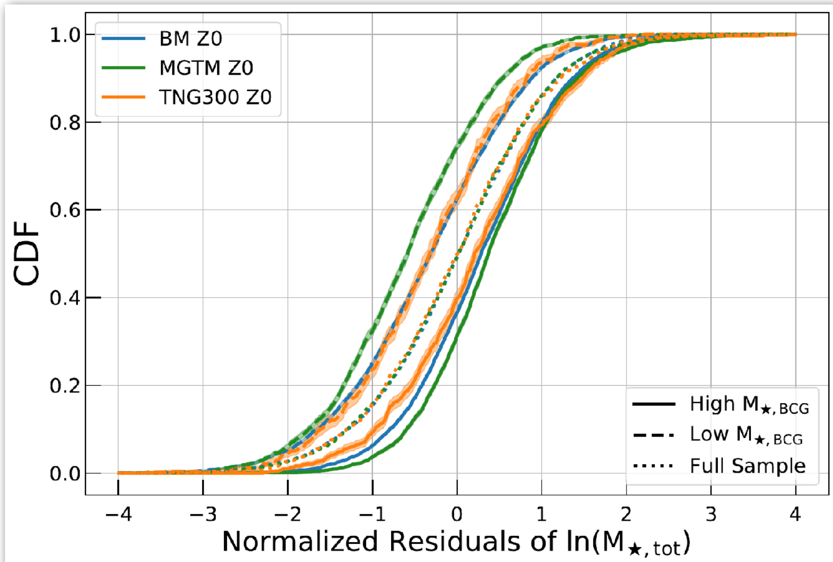






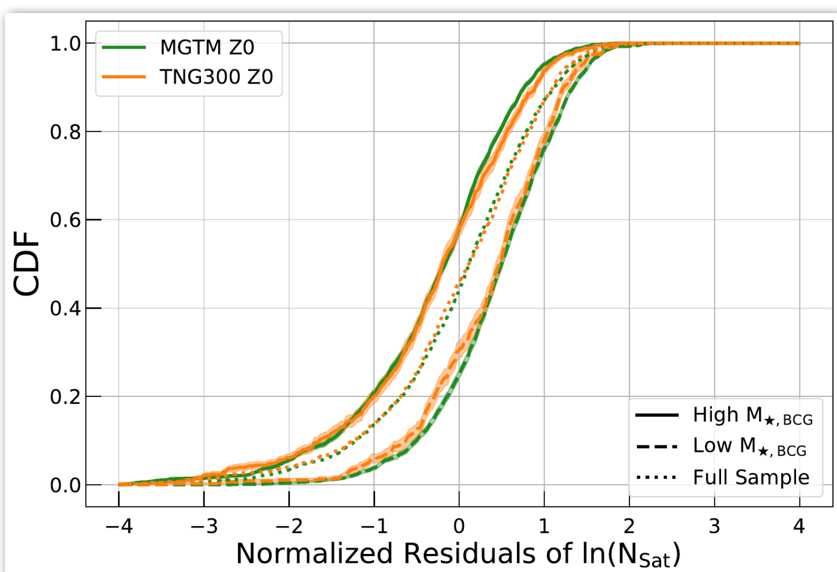
$M_{star,tot}$ PDF split by BCG stellar mass @z=0

- BCG stellar mass relates to formation epoch; early = more massive
- More massive than average BCG stellar mass => higher $M_{star,tot}$

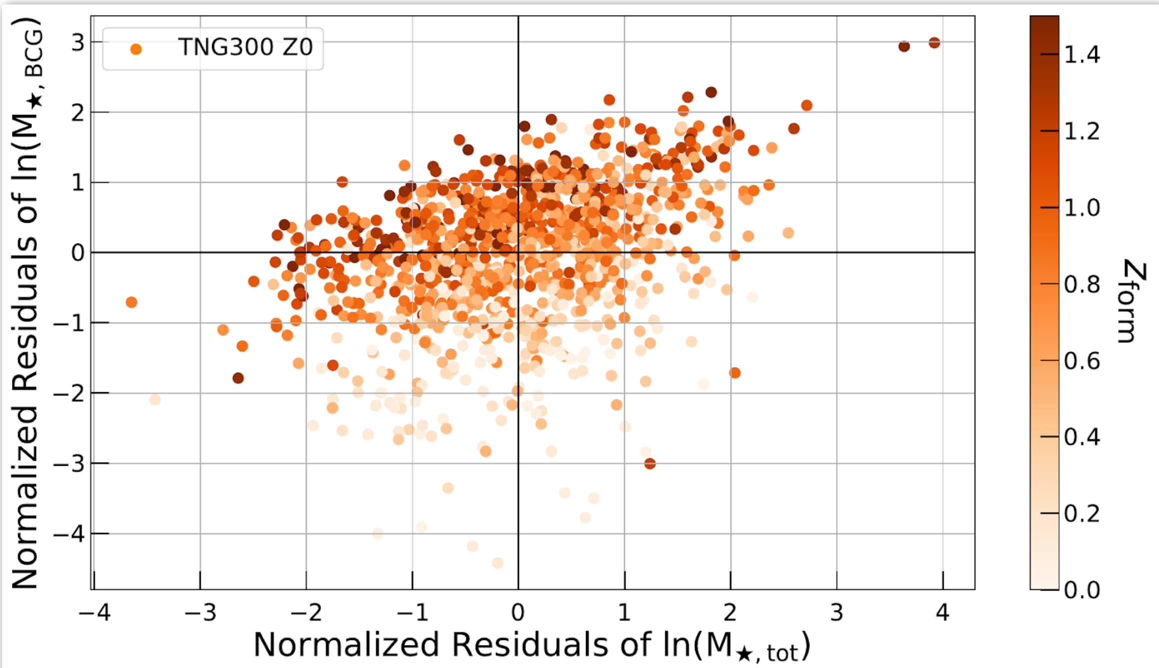


N_{sat} PDF split by BCG stellar mass @z=0

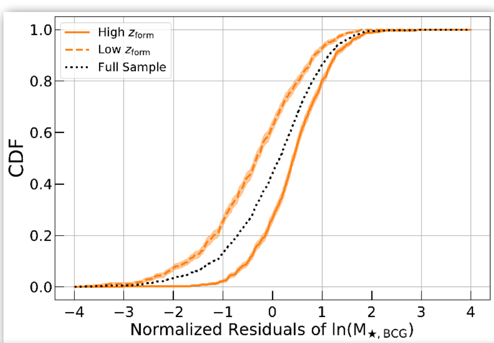
- More massive than average BCG stellar mass => lower N_{sat}



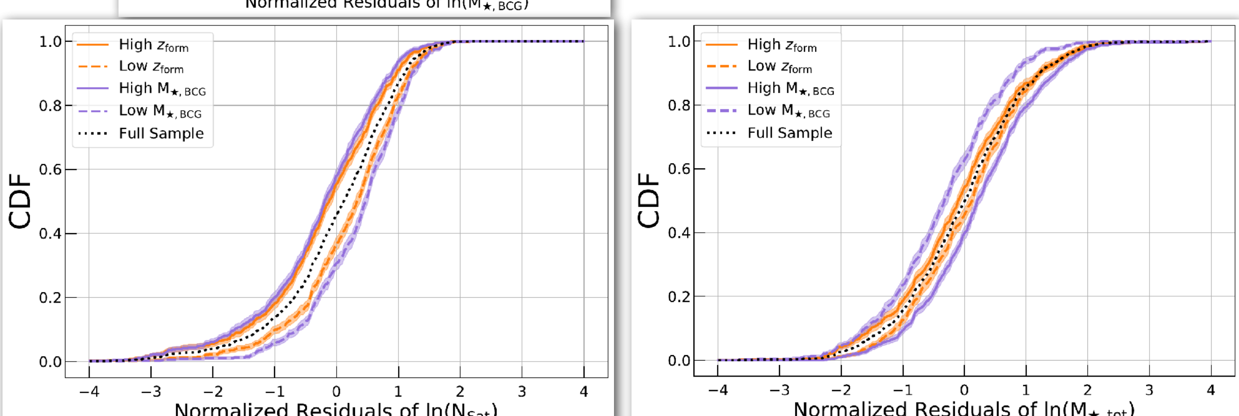
Halo formation epoch (TNG only)



Halo formation epoch (TNG only)



- BCG stellar mass and Nsat sensitive to formation epoch (with opposite signs)
- A halo's total stellar mass is insensitive to formation epoch but correlates with BCG stellar mass

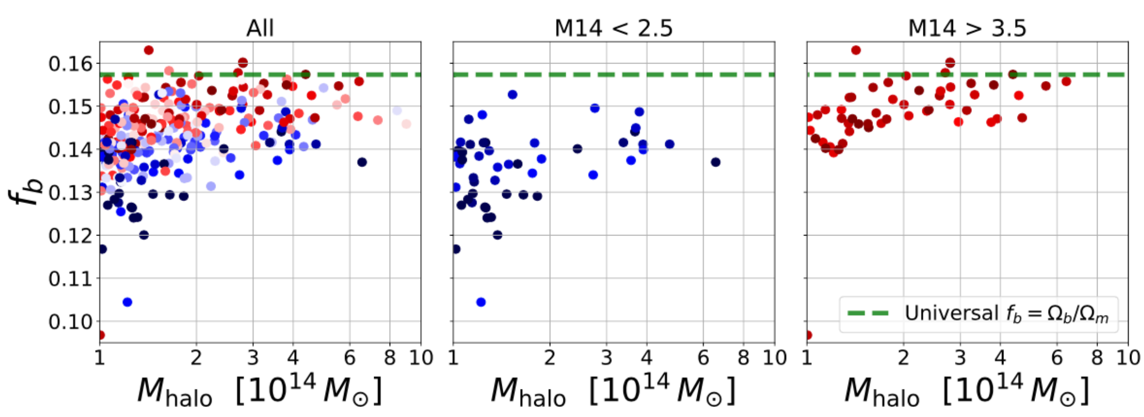


27-28 Jan 2020 - KIPAC + LBL Seminars – Gus Evrard, U Michigan

What about the hot gas?

TNG: magnitude gap and baryon fraction

Farahi, Ho + Trac, 2001.05639



Summary of findings

- **Stellar property statistics** of massive halos: **robust** features
  - **PDF shapes** (3D properties within R_{200})

    $p(\ln N_{sat} \mid M, z)$ is skewed negative, fit by Gaussian mixture model

    $p(\ln M_{star,tot} \mid M, z)$ & $p(\ln M_{gas} \mid M, z)$ closer to Gaussian
  - **Correlations** (mass-conditioned, recall non-Guassianity…)

    BCG stellar mass and N_{sat} mildly anti-correlated

    BCG stellar mass and total halo stellar mass mildly correlated

    N_{sat} and total halo stellar mass + for BM/TNG but – for MGTM

    (MGTM halos more dominated by BCG)

* Left to do…

– widen set of properties (galaxy velocities, hot gas measures)

– extend to directly observational measures (projection, mis-centering)

– test for shifts using halo mass proxy {e.g., M_{gas}, T_x} in survey samples

Thoughts on cyberinfrastructure

6.3 Cyberinfrastructure improvements

Our study has been greatly enabled by the availability of full data releases of the IllustrisTNG simulations (Nelson et al. 2019), and a partial public release of Magneticum Pathfinder (Ragagnin et al. 2017). The democratization of the data, through the availability of uniform, catalog-level simulation products to the public, is key to permitting more in-depth analyses of halo population statistics derived from multiple cosmological simulations.

Looking even further, reproducible computational science benefits from having open community access to the specific simulation methods, both production and analysis code bases, used in simulation studies (*e.g.*, Stodden et al. 2016). The scale of simulation data volumes makes it difficult to move high-resolution data to a central location, but a future in which distributed, containerized analysis environments (Raddick et al. 2019) operate using improved discoverability standards (Languignon et al. 2017) could greatly simplify and empower verification studies of the type we perform here.

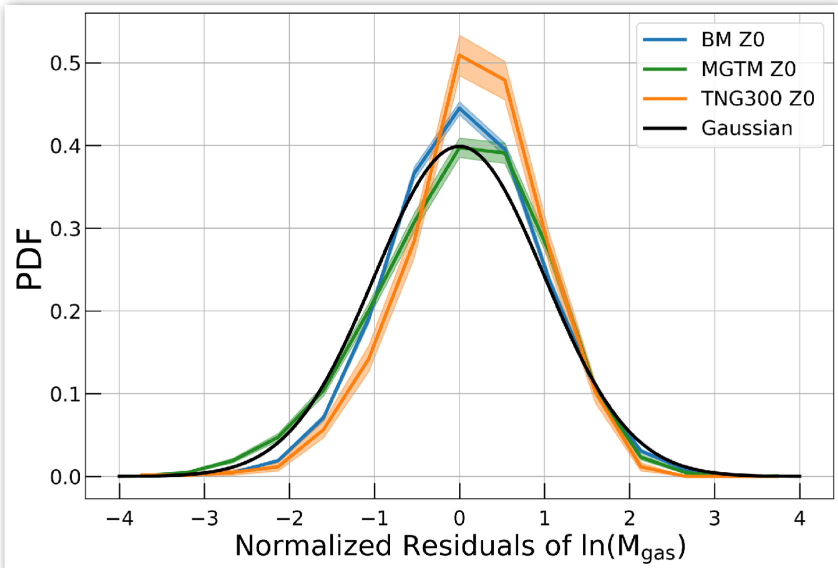
Dhayaa Anbajagane
Undergraduate, U Michigan

Thanks to Dhayaa

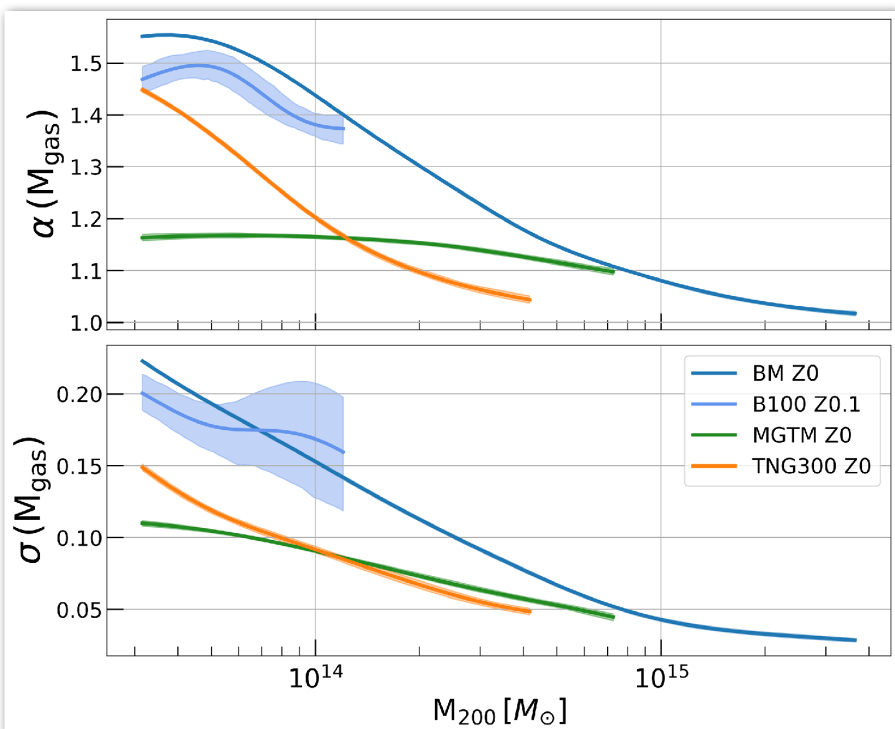
The End

M_{gas} = hot gas mass < R_{200} @ z=0

Mildly skewed log-normal



M_{gas} = hot gas mass < R_{200} @ z=0



Scatter in total halo mass at fixed gas mass