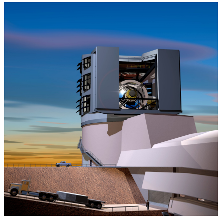
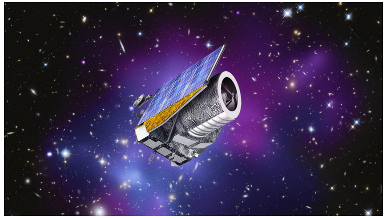
Comparing Galaxy Formation Models using the Bias Expansion

Mahlet Shiferaw
Advisors: Risa Wechsler, Nickolas Kokron

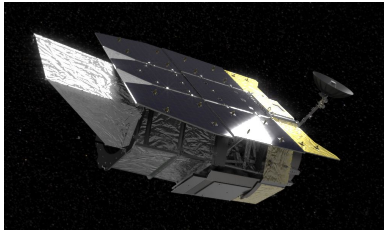
KIPAC
KAVLI INSTITUTE FOR PARTICLE ASTROPHYSICS & COSMOLOGY



Spherex

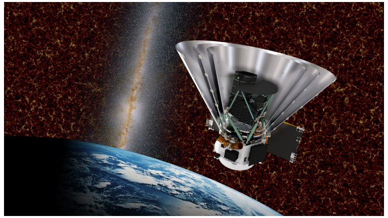


Euclid

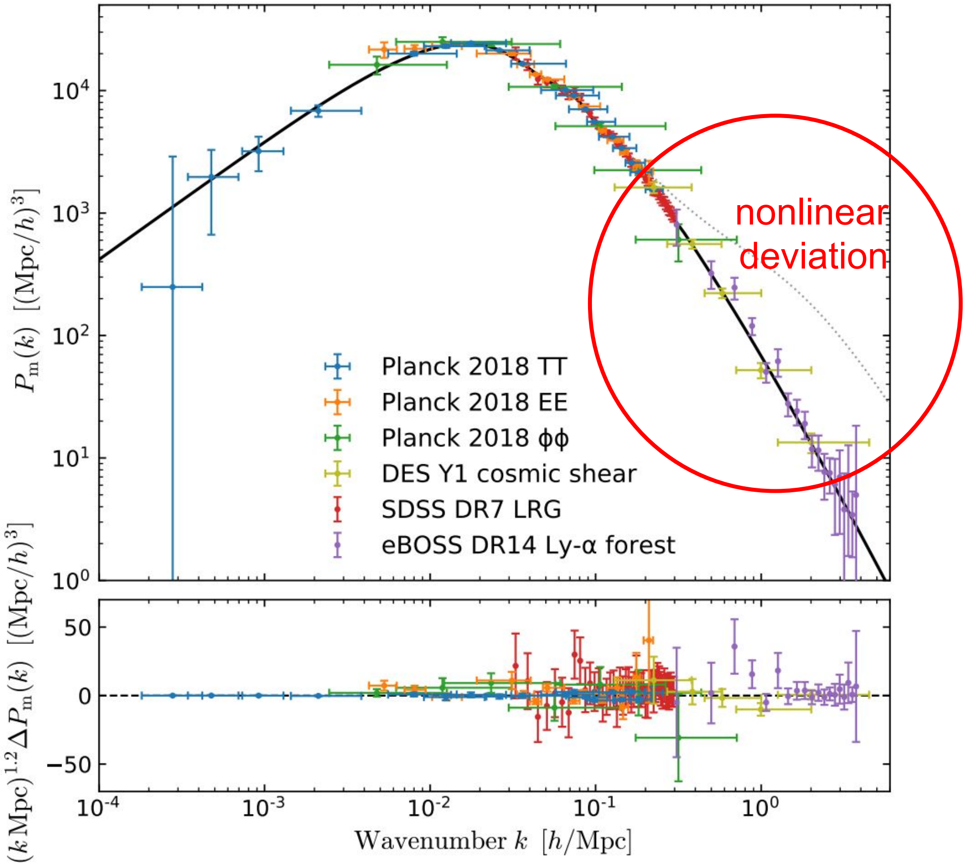


Roman Space Telescope

DESI



Rubin/LSST



We will have a lot of data! (especially at small scales)

Chabanier et al 2019

Galaxy Bias

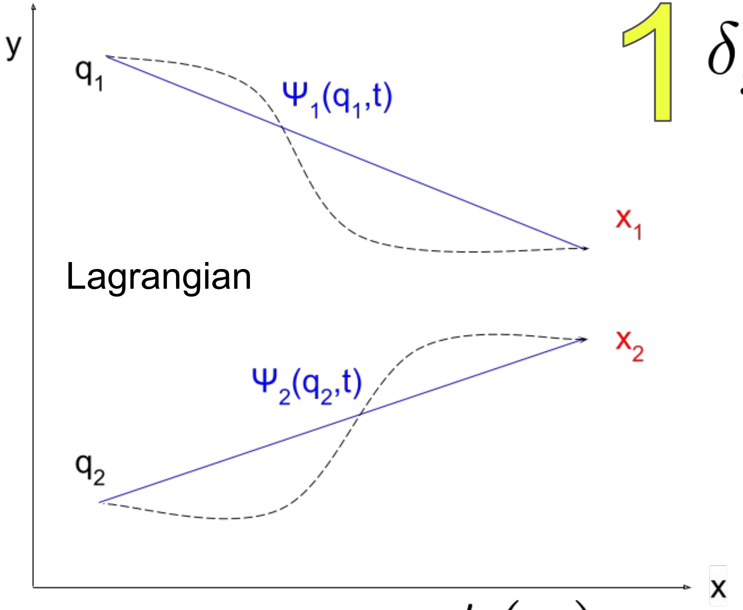
1 $$\delta_g(\boldsymbol{x},\tau) = \sum_O b_O(\tau) O(\boldsymbol{x},\tau)$$

Perturbation Theory

1 $$\delta_g(\boldsymbol{x}, \tau) = \sum_O b_O(\tau) O(\boldsymbol{x}, \tau)$$

Perturbation Theory

2 Lagrangian

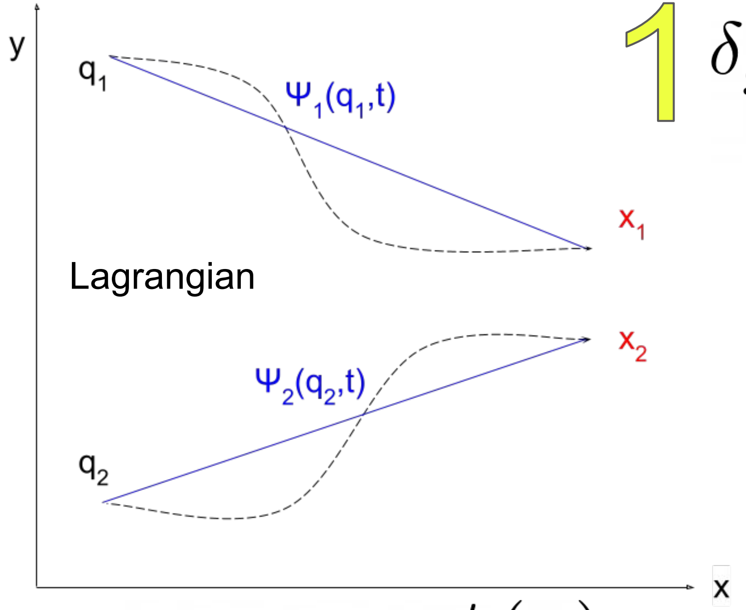


$$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\psi}(\boldsymbol{q})$$

1 $$\delta_g(\boldsymbol{x},\tau) = \sum_O b_O(\tau) O(\boldsymbol{x},\tau)$$

Perturbation Theory

2 Lagrangian



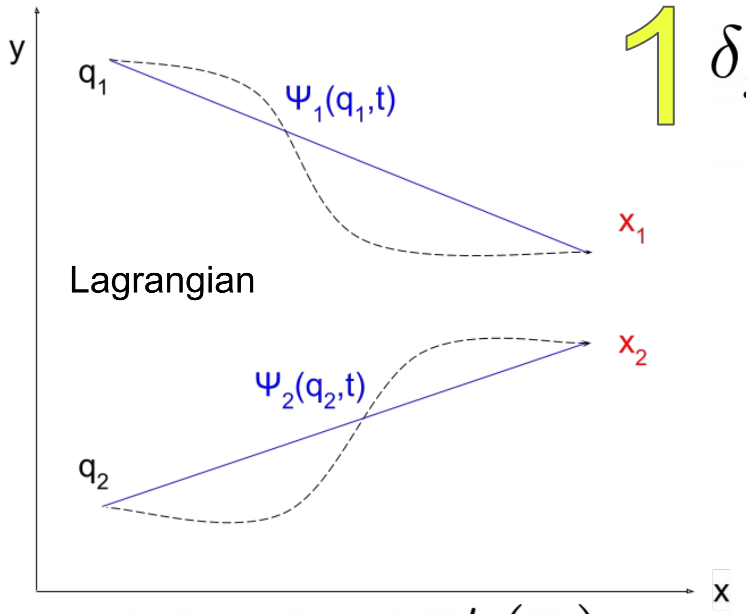
$$\boldsymbol{x} = \boldsymbol{q} + \psi(\boldsymbol{q})$$

Hybrid Effective Field Theory (HEFT)

3 $$\delta_g(\boldsymbol{x}) = \delta_m(\boldsymbol{x}) + b_1 \mathcal{O}_\delta(\boldsymbol{x}) + b_{\nabla^2}\mathcal{O}_{\nabla^2\delta}(\boldsymbol{x}) + b_2\mathcal{O}_{\delta^2}(\boldsymbol{x}) + b_{s^2}\mathcal{O}_{s^2}(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

1 $$\delta_g(\boldsymbol{x},\tau) = \sum_O b_O(\tau) O(\boldsymbol{x},\tau)$$

Perturbation Theory

2 

$$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\psi}(\boldsymbol{q})$$

4 $$S = \left\langle [\epsilon(\boldsymbol{x})]^2_{k_{\max}} \right\rangle$$

Loss function

Hybrid Effective Field Theory (HEFT)

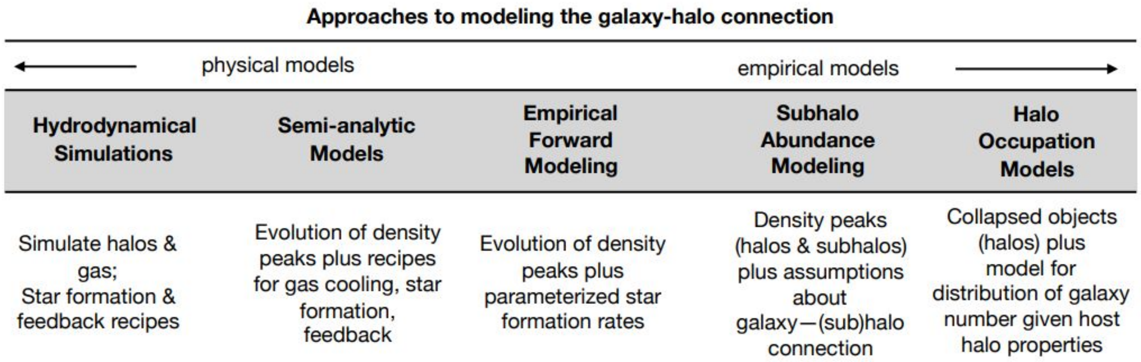
3 $$\delta_g(\boldsymbol{x}) = \delta_m(\boldsymbol{x}) + b_1 \mathcal{O}_\delta(\boldsymbol{x}) + b_{\nabla^2} \mathcal{O}_{\nabla^2\delta}(\boldsymbol{x}) + b_2 \mathcal{O}_{\delta^2}(\boldsymbol{x}) + b_{s^2} \mathcal{O}_{s^2}(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

Approaches to modeling the galaxy-halo connection

← physical models | empirical models →

Hydrodynamical Simulations	Semi-analytic Models	Empirical Forward Modeling	Subhalo Abundance Modeling	Halo Occupation Models
Simulate halos & gas; Star formation & feedback recipes	Evolution of density peaks plus recipes for gas cooling, star formation, feedback	Evolution of density peaks plus parameterized star formation rates	Density peaks (halos & subhalos) plus assumptions about galaxy—(sub)halo connection	Collapsed objects (halos) plus model for distribution of galaxy number given host halo properties

Wechsler + Tinker 2018



Our Approach: We examine quenched and star-forming galaxy samples from models with disparate approaches using a perturbative bias expansion.

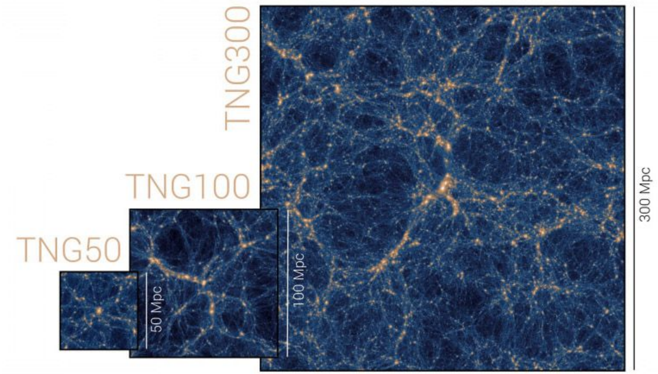
1. Enables a direct comparison between distinct modeling approaches
2. See how uncertainties in galaxy-halo connection impact bias parameters
3. Allows us to inform priors on these bias parameters for future surveys

Wechsler + Tinker 2018

Initial Conditions: 2500^3 dark matter particles, L_{box} = 205 Mpc/h
z=0 (for now)

Models

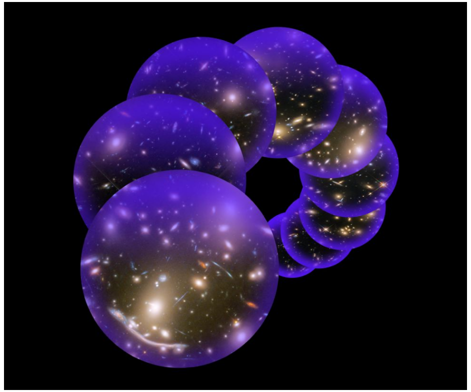
Nelson et al 2021



IllustrisTNG Project

Full, hydrodynamical (more physical) model

Behroozi et al 2019



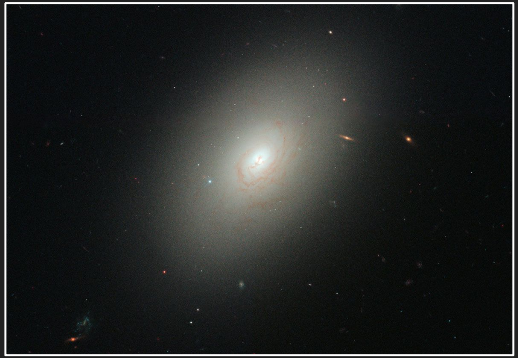
UniverseMachine (UM)

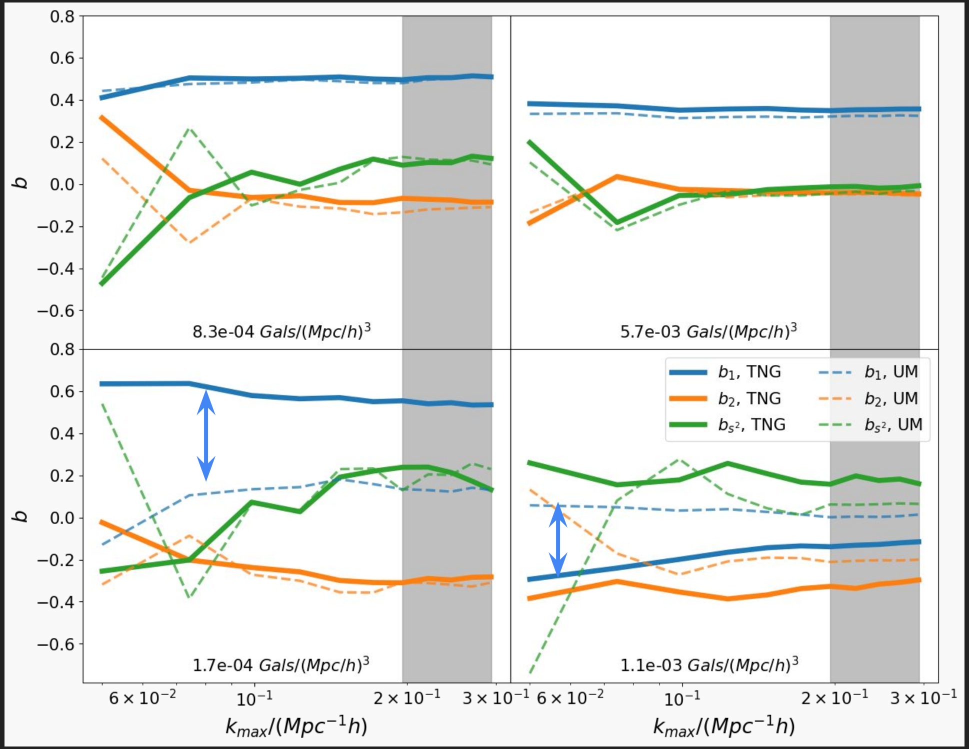
Semi-analytic (more empirical) model

← physical models			empirical models →	
Hydrodynamical Simulations	**Semi-analytic Models**	**Empirical Forward Modeling**	**Subhalo Abundance Modeling**	**Halo Occupation Models**

Preliminary Results

Bias as a function of k_{max}:





Bias parameters have significant disagreement between UM & TNG in star-forming samples!

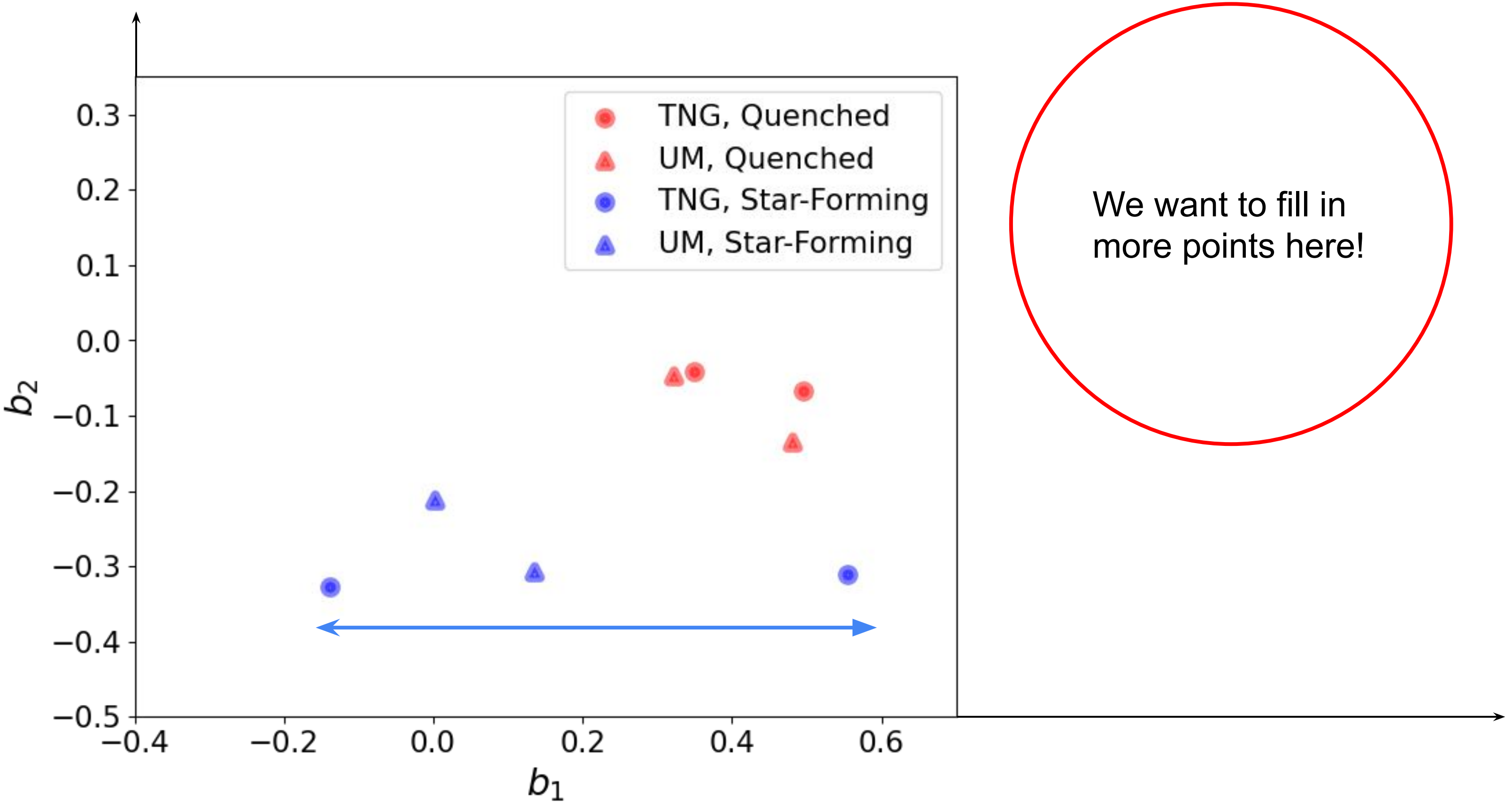
Conclusions

1. Using the language of bias, we have successfully created a framework for a **1:1 comparison between models** with disparate approaches
2. The linear bias relation between the galaxy and matter distributions **agree for for UM and TNG in quenched samples**, but **differ in star-forming samples**
3. This suggests that we need **a wider prior for star-forming galaxies,** but can put **tighter constraints on quenched galaxies**

Next steps:

- Fill in parameter space at higher z (0.5-1.5)
- Fill in priors with intermediate number densities

A closer look at k_{max}=0.2 h/Mpc

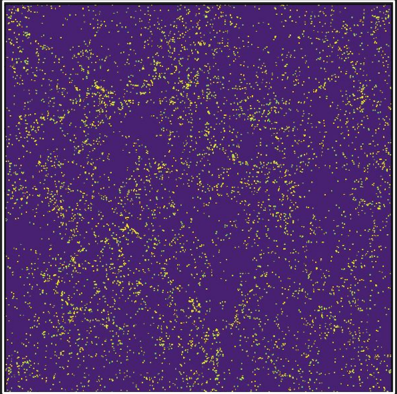


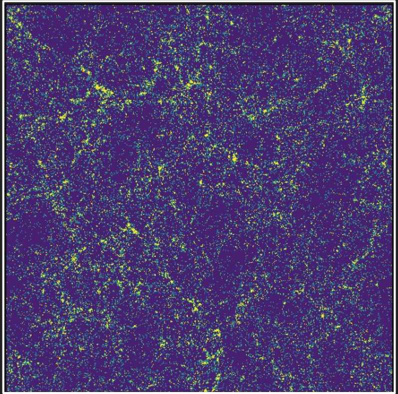
Galaxy Samples

Low density, High stellar mass | **High density, Low stellar mass**

Quenched

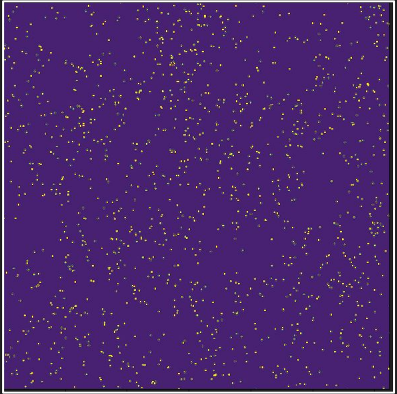
NASA, ESA, R.M. Crockett et al

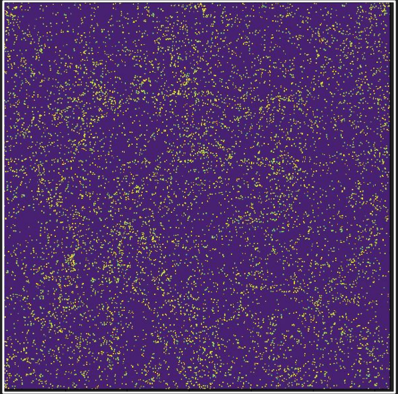




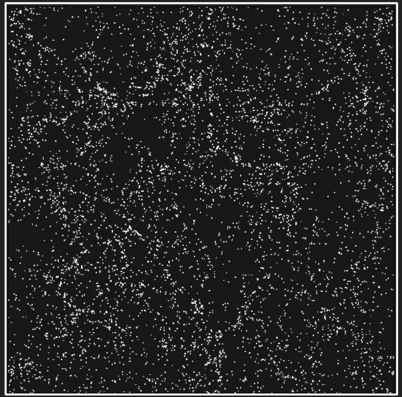
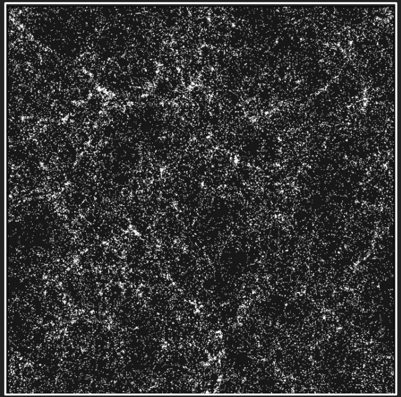
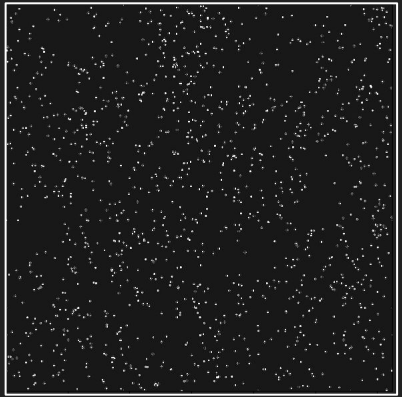
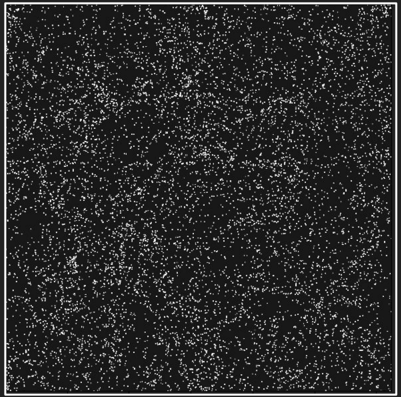
Star-forming

Credit: ESA/Hubble

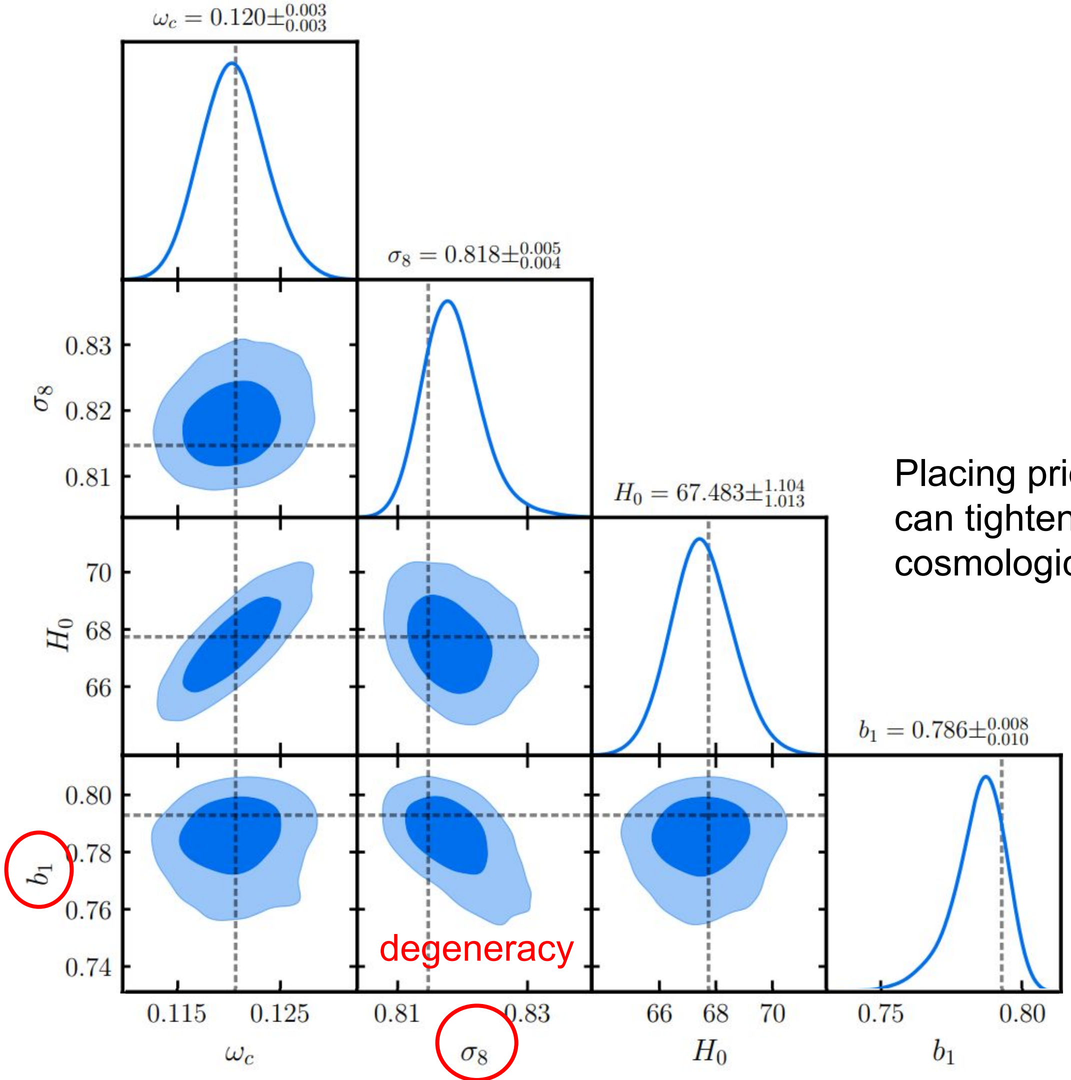




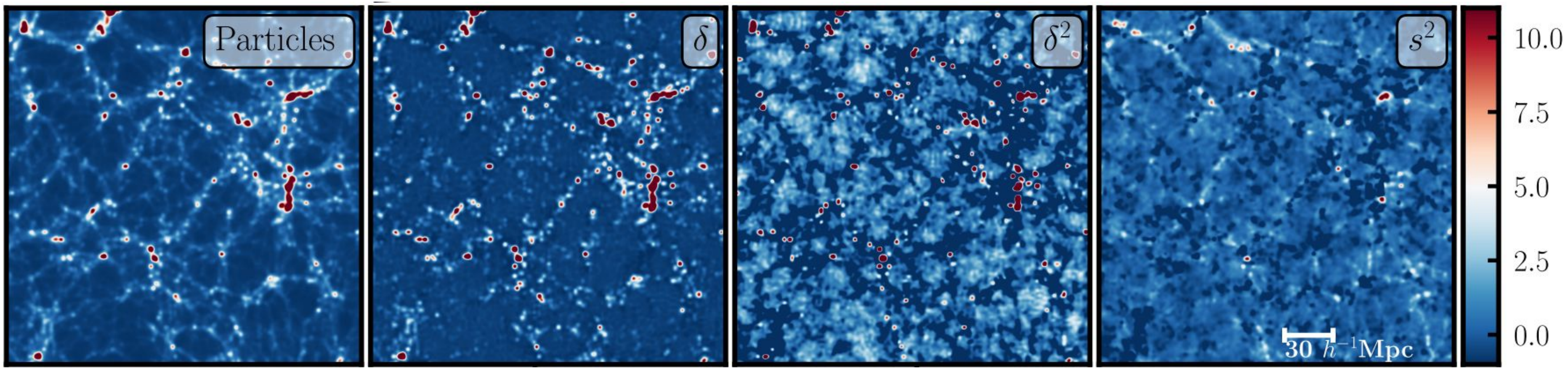
Galaxy Samples

		Low density, High stellar mass	High density, Low stellar mass
Quenched	 NASA, ESA, R.M. Crockett et al		
Star-forming	 Credit: ESA/Hubble		

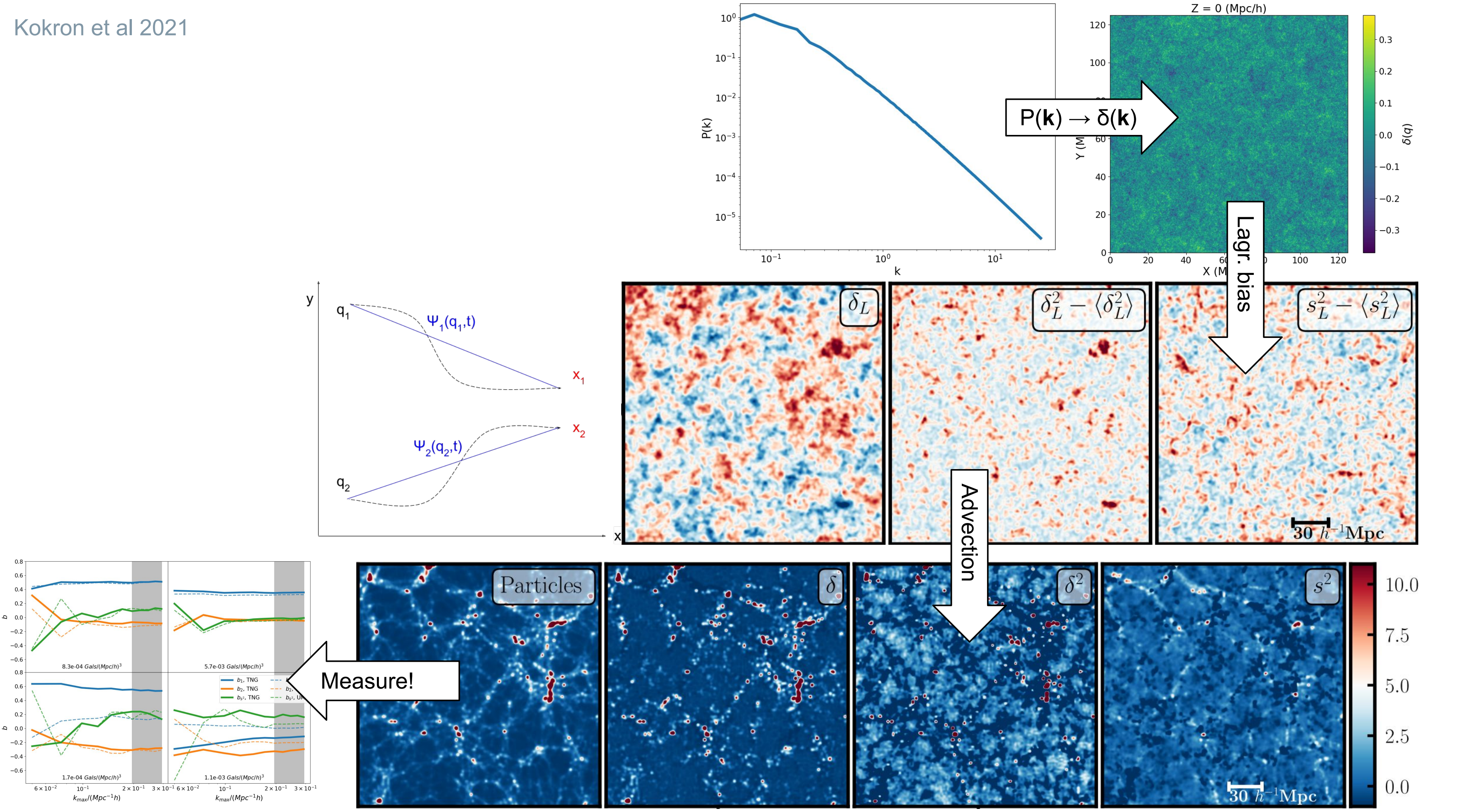
Kokron et al 2021



Placing priors on bias parameters can tighten constraints on cosmological parameters!


Particles
δ
δ^2
s^2
30 h^{-1}Mpc
10.0
7.5
5.0
2.5
0.0

Kokron et al 2021



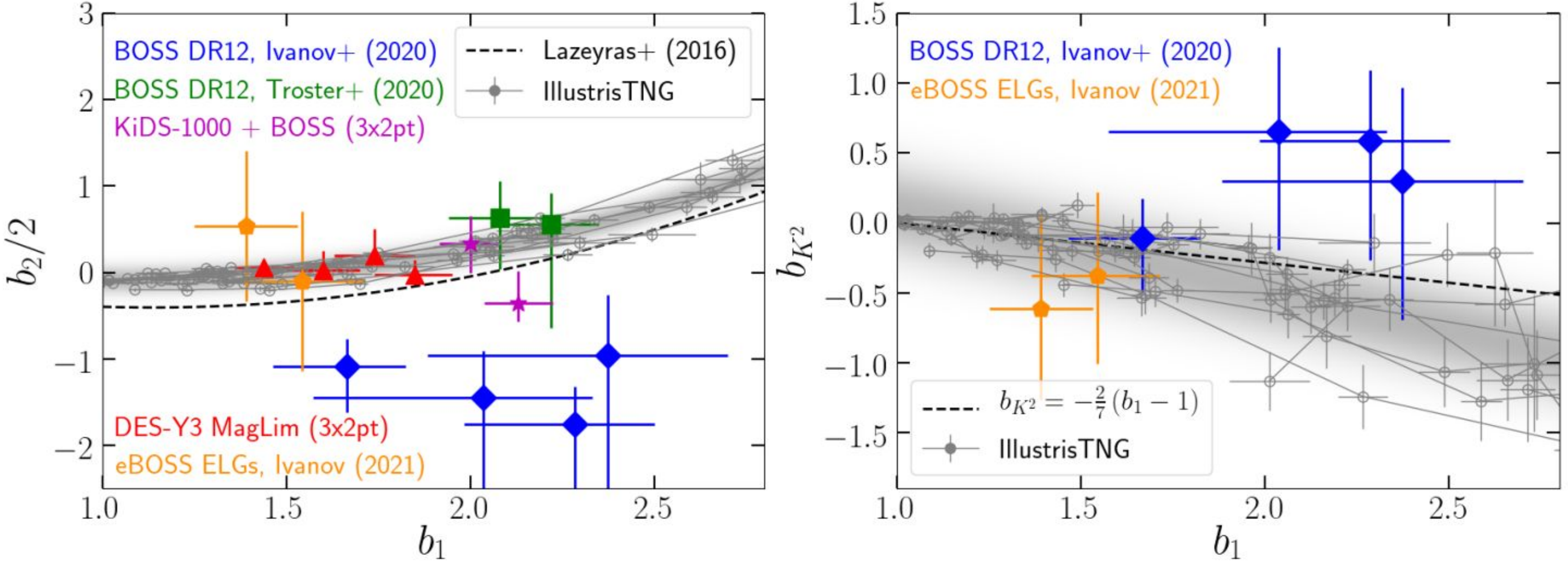


Figure 7. Comparison between the $b_2(b_1)$ and $b_{K^2}(b_1)$ relations obtained in this paper for IllustrisTNG galaxies and recent estimates from real galaxy data analyses. The grey data points show the result for all of the simulated galaxy samples in Figs. 4 and 5 put together, without distinguishing by redshift or selection criterion. The grey shaded map shows a Gaussian prior based on the IllustrisTNG data points (cf. Eqs. (3.8) and (3.9)). The colored data points show the relations inferred from recent cosmological inference analyses of observed galaxies: blue diamonds from Ref. [86], green squares from Ref. [87], magenta stars from Ref. [88], red triangles from Ref. [90], and orange pentagons from Ref. [89]. All error bars are 68% confidence limits.

Expand perturbatively…

1) To first order: $$\delta_g(\boldsymbol{x}) = \underline{b_1}\delta_m(\boldsymbol{x})$$

2) Taylor expand: $$\delta_g(\boldsymbol{x}, \tau) = \sum_O \underline{b_O(\tau)} O(\boldsymbol{x}, \tau)$$

3) Hybrid expansion: $$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\psi}(\boldsymbol{q})$$

$$\delta_g(\boldsymbol{x}) = \delta_m(\boldsymbol{x}) + \underline{b_1}\mathcal{O}_\delta(\boldsymbol{x}) + \underline{b_{\nabla^2}}\mathcal{O}_{\nabla^2\delta}(\boldsymbol{x}) + \underline{b_2}\mathcal{O}_{\delta^2}(\boldsymbol{x}) + \underline{b_{s^2}}\mathcal{O}_{s^2}(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

Measuring the bias parameters via maximum likelihood

1) Define the stochasticity field:

$$\epsilon(\boldsymbol{x}) = \delta_g(\boldsymbol{x}) - \left(\delta_m(\boldsymbol{x}) + \sum_O b_O(\tau) O(\boldsymbol{x}, \tau)\right)$$

2) Take the Fourier Transform and define its mean/variance:

$$\langle \epsilon(\boldsymbol{k}) \rangle = 0$$

$$\left\langle \epsilon(\boldsymbol{k}) \epsilon\left(\boldsymbol{k}'\right) \right\rangle \equiv P_{\mathrm{err}}(|\boldsymbol{k}|)$$

3) Find the bias parameters that **minimize this loss function** (up to some k_{max}):

$$S = \left\langle [\epsilon(\boldsymbol{x})]^2_{k_{\mathrm{max}}} \right\rangle$$

$$P\Big(\delta_h(\boldsymbol{k}) - \delta_{h,\mathrm{det}}(\boldsymbol{k})\Big) = \int d\varepsilon_m(\boldsymbol{k})\; P\Big(\delta_h(\boldsymbol{k}) - \delta_{h,\mathrm{det}}(\boldsymbol{k}) - b_1\varepsilon_m(\boldsymbol{k}),\; \varepsilon_m(\boldsymbol{k})\Big)$$
$$= (2\pi)^{-1/2}\left|P^{\varepsilon}_{hh}(k) + 2b_1 P^{\varepsilon}_{hm}(k) + b_1^2 P^{\varepsilon}_{mm}\right|^{-1/2}$$
$$\times \exp\left[-\frac{1}{2}\frac{|\delta_h(\boldsymbol{k}) - \delta_{h,\mathrm{det}}(\boldsymbol{k})|^2}{P^{\varepsilon}_{hh}(k) + 2b_1 P^{\varepsilon}_{hm}(k) + b_1^2 P^{\varepsilon}_{mm}}\right]. \tag{3.4}$$

$$\epsilon(\boldsymbol{x}) = \delta_g(\boldsymbol{x}) - \left(\delta_m(\boldsymbol{x}) + \sum_O b_O(\tau) O(\boldsymbol{x}, \tau)\right)$$

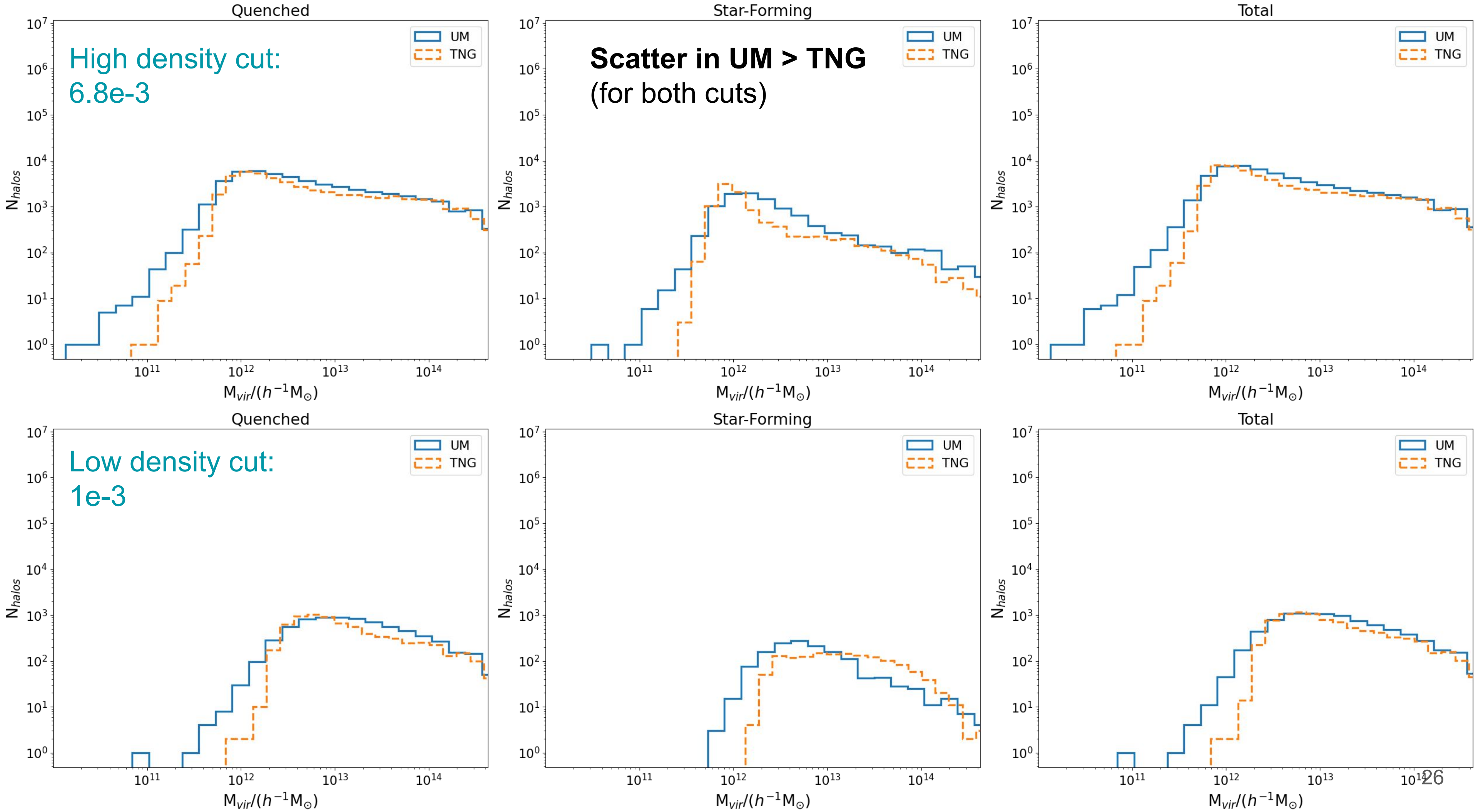
$$S = \left\langle [\epsilon(\boldsymbol{x})]^2_{k_{\max}} \right\rangle$$

$$S \approx \int_{|\mathbf{k}|<k_{\max}} \frac{d^3k}{(2\pi)^3} \left\| \delta_h(\mathbf{k}) - \delta_m(\mathbf{k}) - \sum_i b_i \mathcal{O}_i(\mathbf{k}) \right\|^2 . \qquad (14)$$

Galaxy Samples

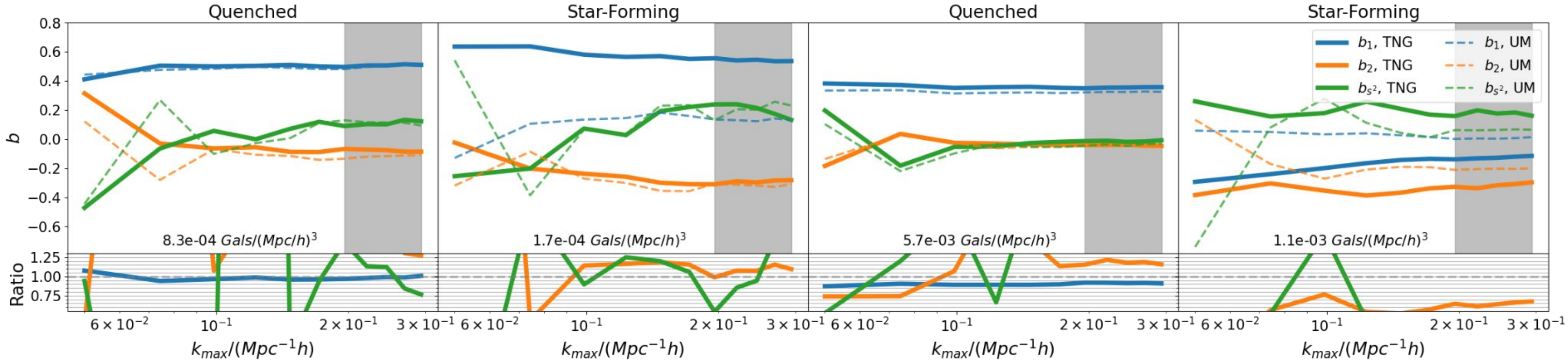
1. Set the total number density with a **stellar mass cut**
2. Split the galaxies into **quenched** and **star-forming** samples with a **specific star formation rate cut**

	Low density (Gals/(Mpc/h)^3)	High density (Gals/(Mpc/h)^3)
Quenched	8.3e-4	5.7e-3
Star-forming	1.7e-4	1.1e-3
Total	1e-3	6.8e-3



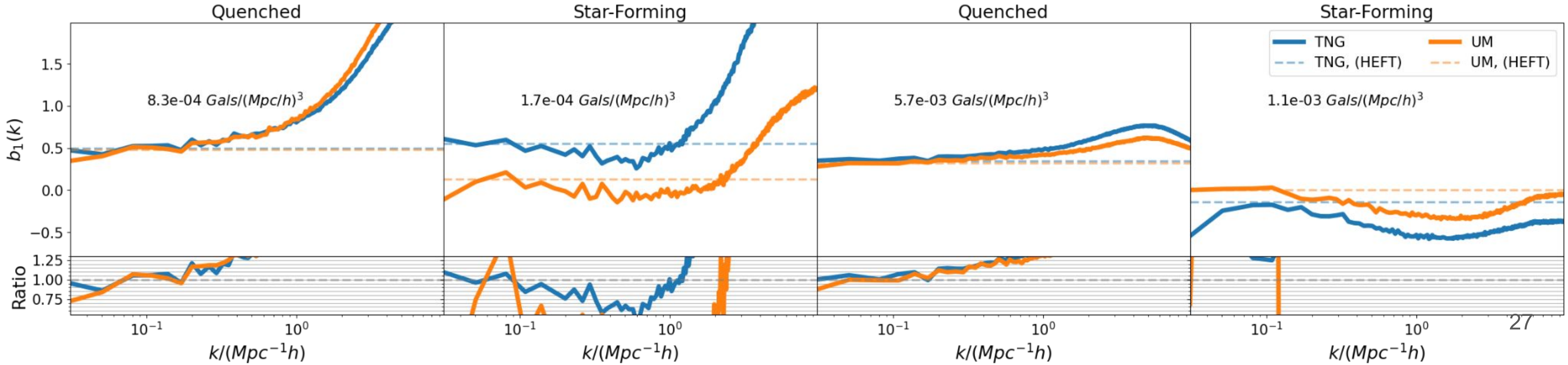
Quenched
Star-Forming
Total
High density cut:
6.8e-3
Scatter in UM > TNG
(for both cuts)
Low density cut:
1e-3
UM
TNG
N_{halos}
$M_{vir}/(h^{-1}M_{\odot})$

Bias parameters as a function of k_{max}:

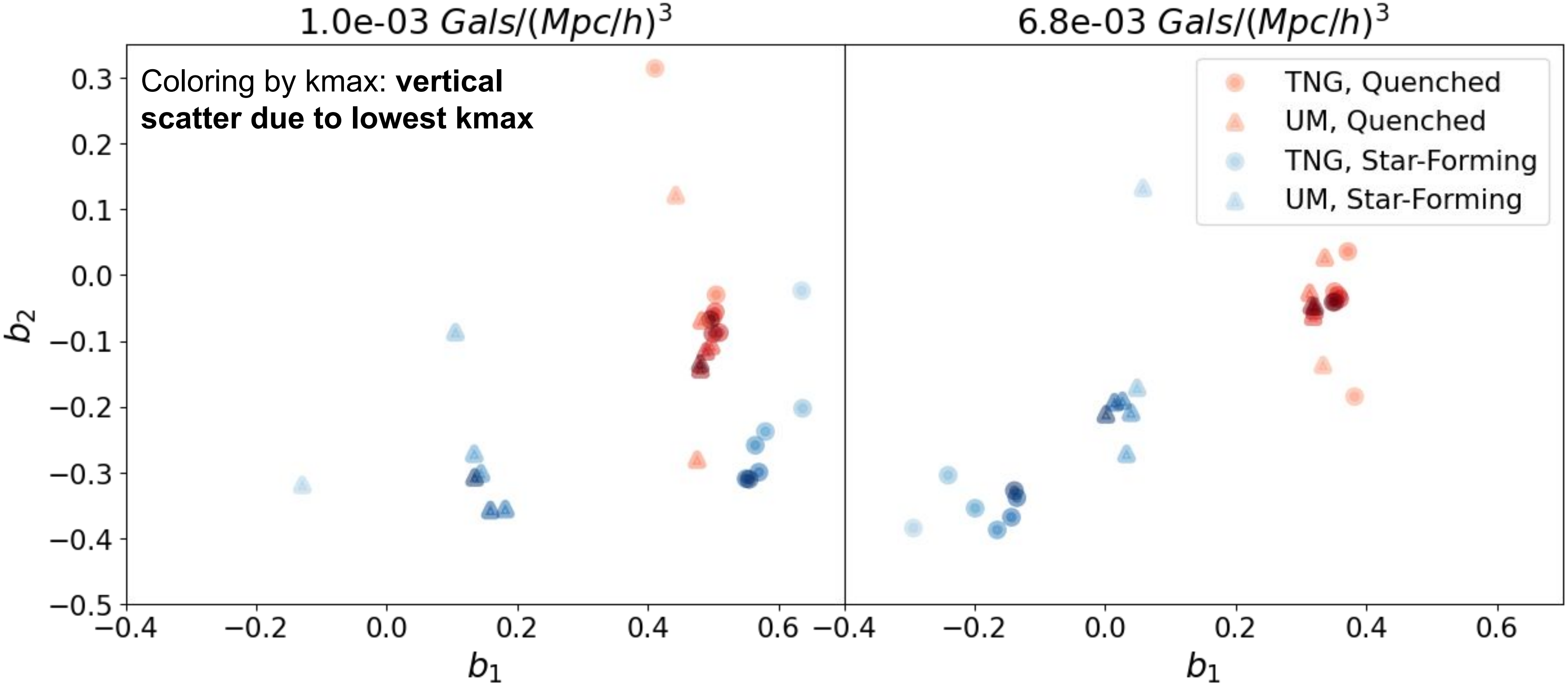


Sanity check with linear theory at large scales:

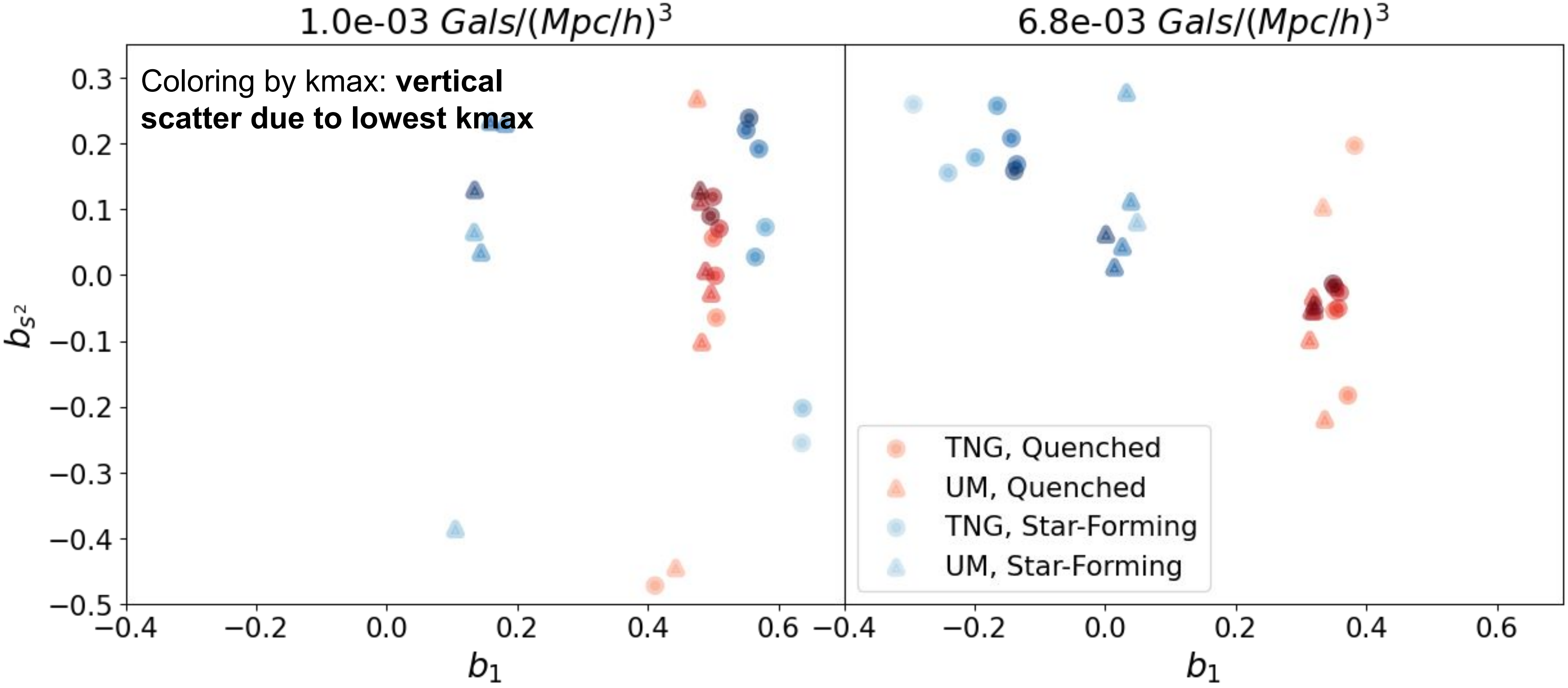
$$\langle \delta_g(\boldsymbol{k})\delta_m(\boldsymbol{k}')\rangle = b_1 \langle \delta_m(\boldsymbol{k})\delta_m(\boldsymbol{k}')\rangle$$



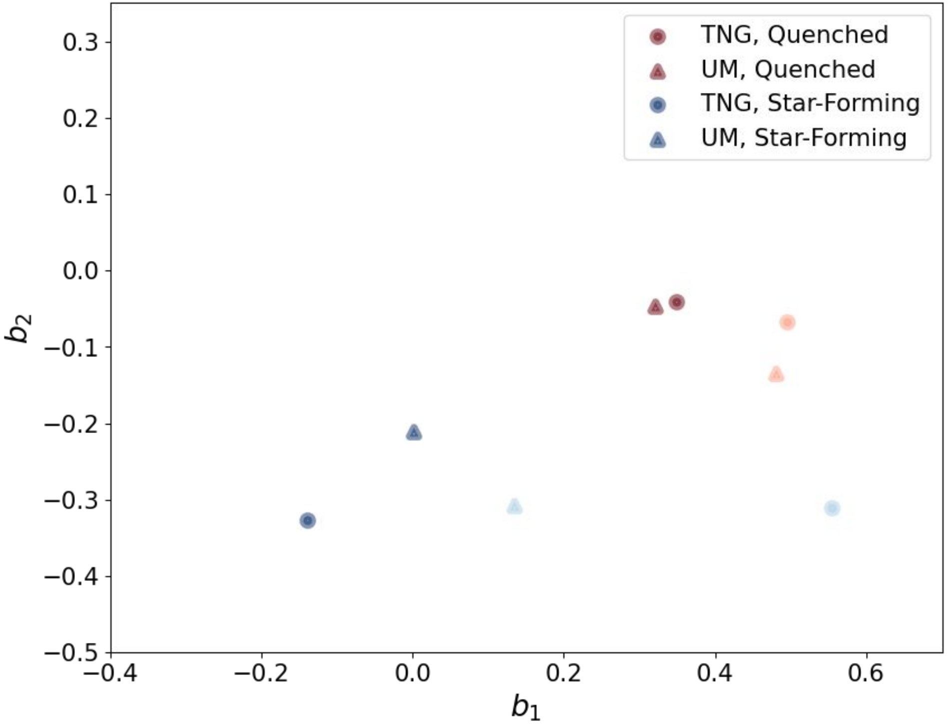
b1 vs b2 in large scale regime (kmax<0.2 h/Mpc)



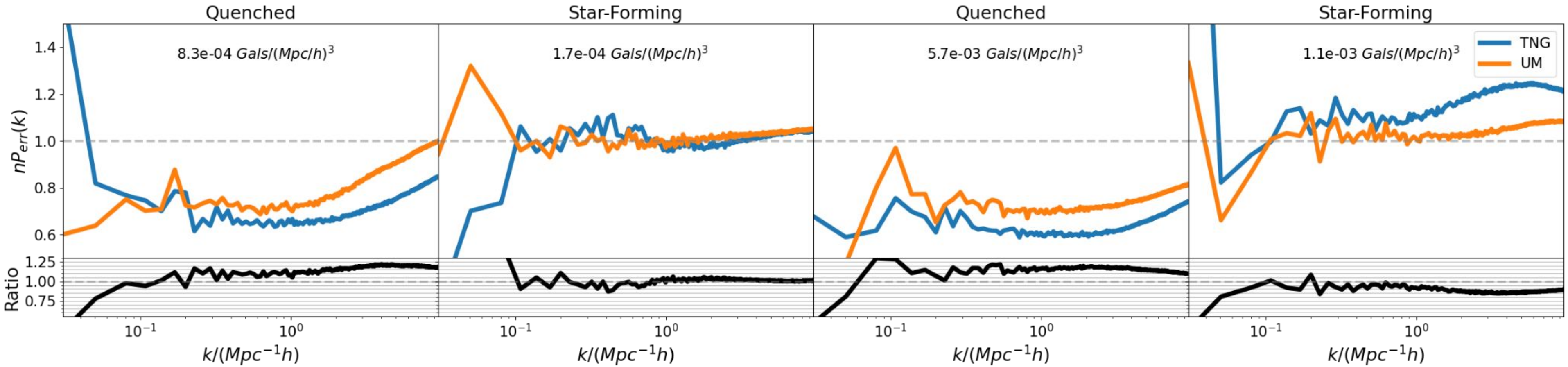
b1 vs bs2 in large scale regime (kmax<0.2 h/Mpc)



A closer look at k_{max}=0.2 h/Mpc



Comparing stochasticity power spectrum at z =0:



Comparing HEFT/galaxy power spectrum at z=0: