The Impact of Galaxy Formation on Galaxy Biasing, and Informative Priors for Primordial non-Gaussianity Constraints

Lucia A. Perez

With Anne Moore, Shy Genel, Elisabeth Krause, Rachel Somerville

Flatiron Research Fellow at the CCA
Future Faculty in the Physical Sciences Postdoctoral Fellow at Princeton University

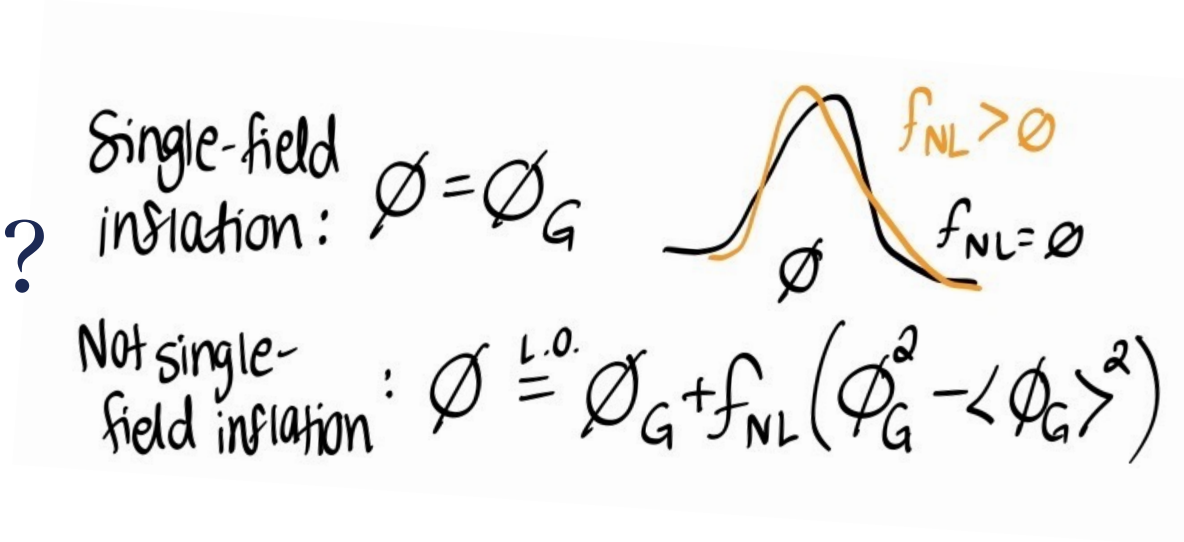
arxiv.org/abs/2602.04987

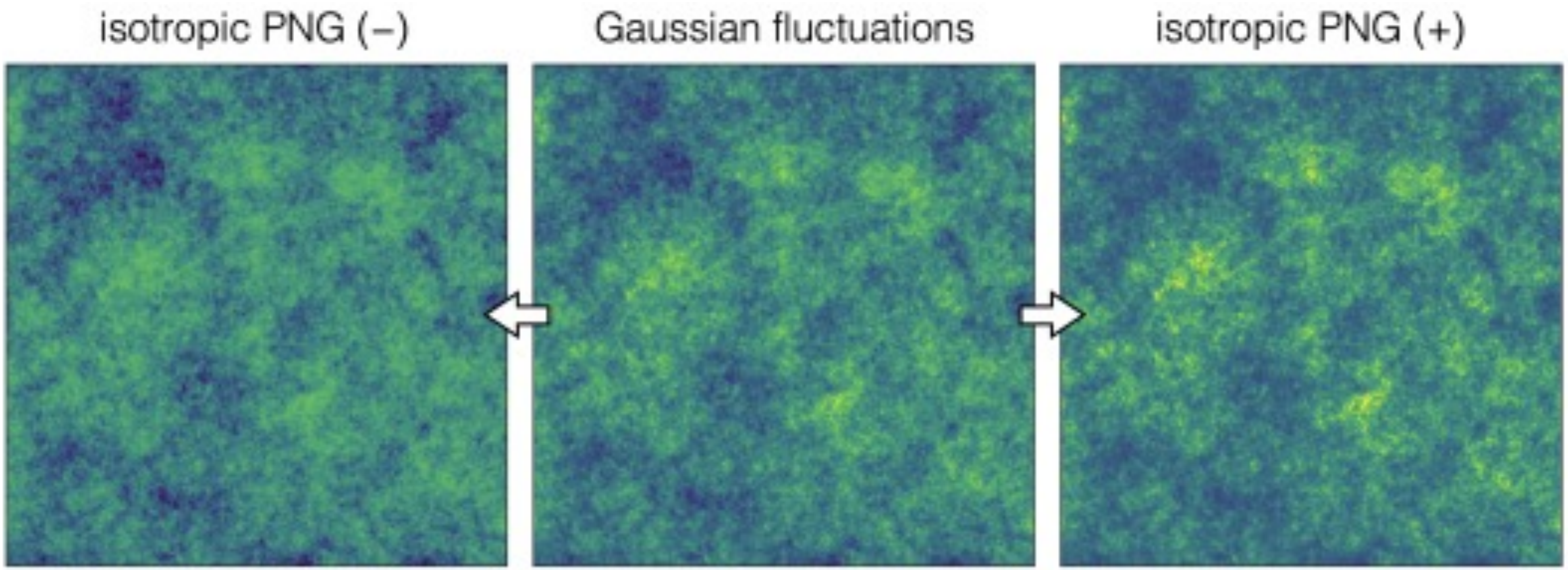
FLATIRON INSTITUTE

April 21st, 2026 | BCCP/Cosmology Seminar

better galaxies = better cosmology

How non-Gaussian were the primordial energy fluctuations? How did inflation occur?





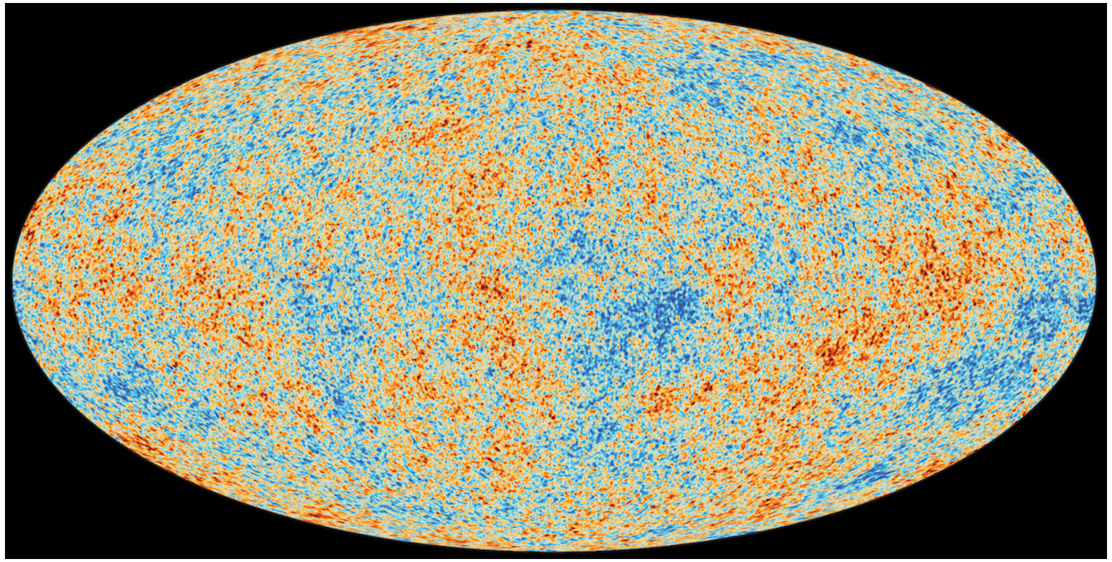
https://www.ipmu.jp/en/20231221-Non-Gaussianity , based off Kurita & Takada 2023

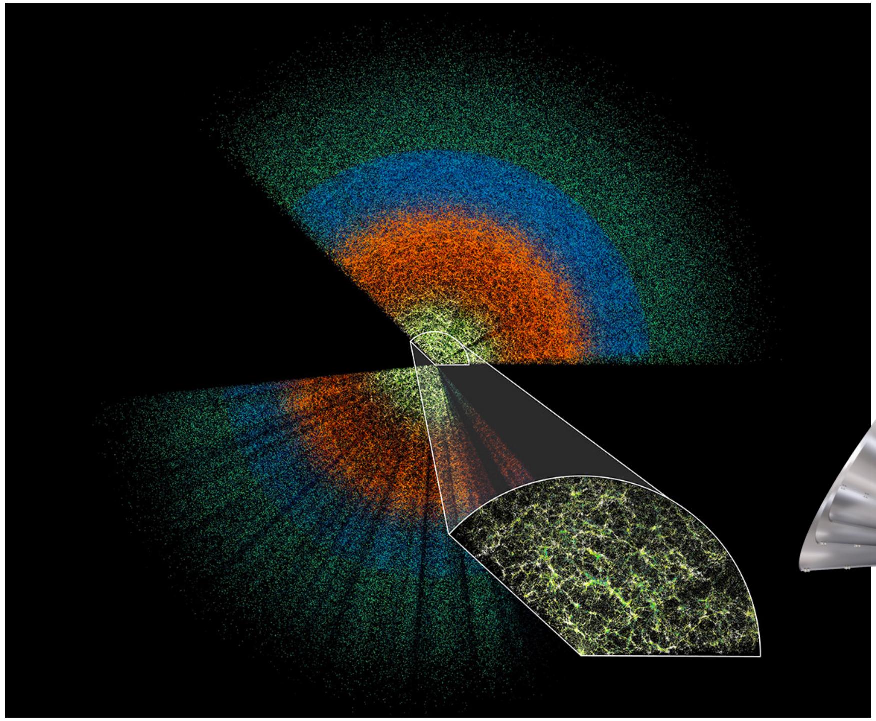
f_{NL} measures local primordial non-Gaussianity

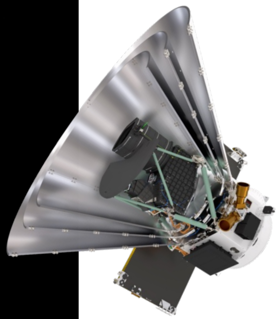
Measuring non-zero f_{NL} helps constrain the process of inflation

$f_{NL} = -0.9 \pm 5.1$ *(68%)*

but I can't go past k ~0.05 h^{-1} Mpc and I'm nearing my cosmic variance limit ☹



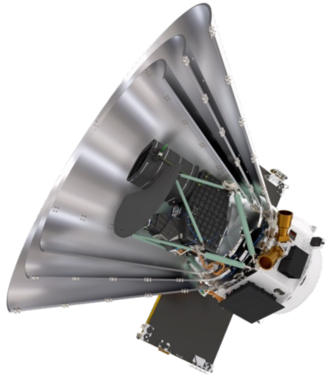




That's okay—I'll use galaxy clustering past k > 0.3 h^{-1} Mpc and reduce that error bar to within ±1!

f_{NL} is **completely degenerate** with galaxy bias parameters in analyses of galaxy clustering!

FLATIRON INSTITUTE



$$\delta_g(\vec{k},z) \overset{\text{L.O.}}{=} b_1(z)\,\delta_m(\vec{k},z) + b_\phi(z)\, f_{NL}^{loc}\, \phi(\vec{k},z) + \varepsilon(z)$$

galaxy density contrast

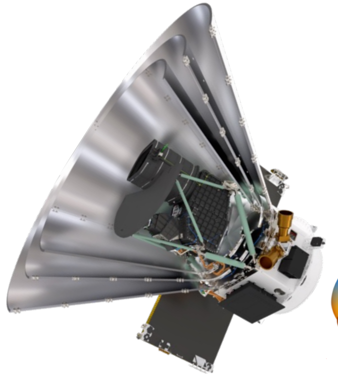
matter density contrast

primordial gravitational potential

random noise

f_{NL} is **completely degenerate** with galaxy bias parameters in analyses of galaxy clustering!

FLATIRON INSTITUTE



$$\delta_g(\vec{k},z) \overset{\text{L.O.}}{=} b_1(z)\,\delta_m(\vec{k},z) + b_\phi(z)\,f_{NL}^{loc}\,\phi(\vec{k},z) + \varepsilon(z)$$

galaxy density contrast — matter density contrast — primordial gravitational potential — random noise

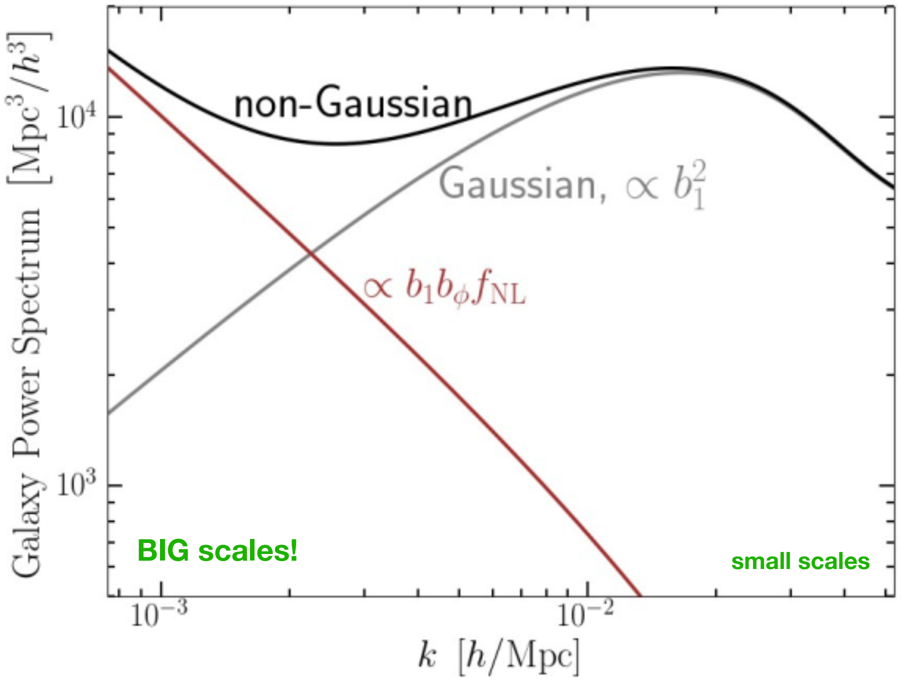
$$\hookrightarrow \delta_g(\vec{k},z) \equiv \frac{n_g(\vec{k},z)}{\bar{n}_g(\vec{k},z)} - 1\,;\quad \frac{\text{local}}{\text{mean}} \text{ comoving galaxy densities}$$

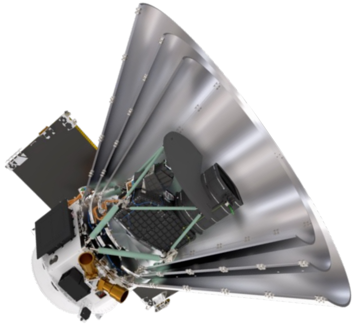
f_{NL} is **completely degenerate** with galaxy bias parameters in analyses of galaxy clustering!

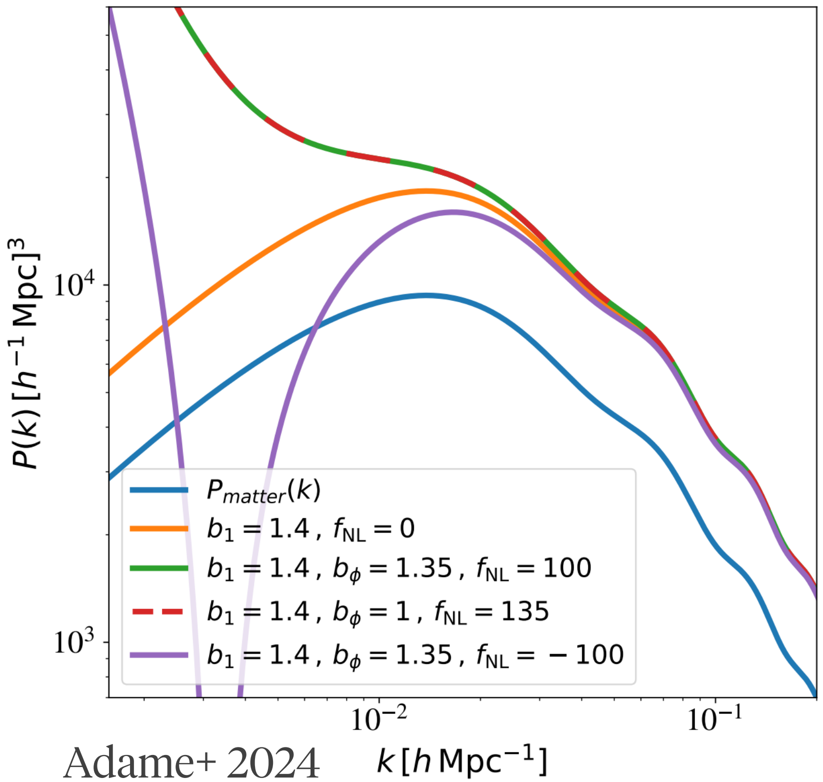
FLATIRON INSTITUTE

$$P_{gg}(\vec{k},z) \overset{L.O.}{=} P_{mm}(\vec{k},z)\left\{b_1^2 + \frac{2 b_1 b_\phi f_{NL}}{\mathcal{M}(\vec{k},z)} + \frac{b_\phi^2 f_{NL}^2}{\mathcal{M}(\vec{k},z)^2}\right\} + P_{\varepsilon\varepsilon}$$

galaxy power spectrum — matter power spectrum







Adame+ 2024

Can we get around the degeneracy of b_1 b_ϕ f_{NL} ?

FLATIRON INSTITUTE

Let's try what works for halos:

"universality relation"

$$b_\phi(z) = 2\delta_c(\ b_1(z) - 1\ )$$

Can we get around the degeneracy of b_1 b_ϕ f_{NL} ?

FLATIRON INSTITUTE

Let's try what works for halos:
"universality relation"

$$b_\phi(z) = 2\delta_c(\ b_1(z) - 1\ )$$

Issue 1: not perfect for simulated halos

Issue 2: priors on b_1 and b_ϕ are ***supremely*** important for your final f_{NL} constraint

Can we get around the degeneracy of b_1 b_ϕ f_{NL} ?

FLATIRON INSTITUTE

Let's try what works for halos:

"universality relation"

$$b_\phi(z) = 2\delta_c(\ b_1(z) - 1\ )$$

Issue 2 : priors on b_1 and b_ϕ are ***supremely*** important for your final f_{NL} constraint

Can we actually constrain $f_{\rm NL}$ using the scale-dependent bias effect?

An illustration of the impact of galaxy bias uncertainties using the BOSS DR12 galaxy power spectrum

Alexandre Barreira[a,b]

[a]Excellence Cluster ORIGINS, Boltzmannstraße 2, 85748 Garching, Germany
[b]Ludwig-Maximilians-Universität, Schellingstraße 4, 80799 München, Germany

Same BOSS galaxy data, different priors for b_ϕ and b_1

Can we get around the degeneracy of b_1 b_ϕ f_{NL} ?

FLATIRON INSTITUTE

Let's try what works for halos:

"universality relation"

$$b_\phi(z) = 2\delta_c(\ b_1(z) - 1\ )$$

Issue 2 : priors on b_1 and b_ϕ are ***supremely*** important for your final f_{NL} constraint

Can we actually constrain $f_{\rm NL}$ using the scale-dependent bias effect?

An illustration of the impact of galaxy bias uncertainties using the BOSS DR12 galaxy power spectrum

Alexandre Barreira[a,b]

[a]Excellence Cluster ORIGINS, Boltzmannstraße 2, 85748 Garching, Germany
[b]Ludwig-Maximilians-Universität, Schellingstraße 4, 80799 München, Germany

Same BOSS galaxy data, different priors for b_ϕ and b_1

↓

$f_{NL} = 16 \pm 16$ ***or***

$f_{NL} = 230 \pm 226$

Can we get around the degeneracy of b_1 b_ϕ f_{NL} ?

FLATIRON INSTITUTE

Let's try what works for halos:
"universality relation"

$$b_\phi(z) = 2\delta_c(\ b_1(z) - 1\ )$$

Issue 2 : priors on b_1 and b_ϕ are ***supremely*** important for your final f_{NL} constraint

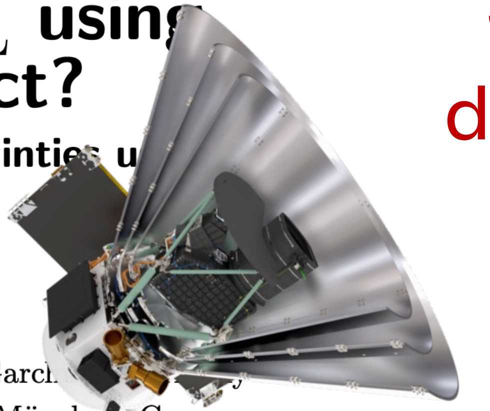
Can we actually constrain $f_{\rm NL}$ using the scale-dependent bias effect?

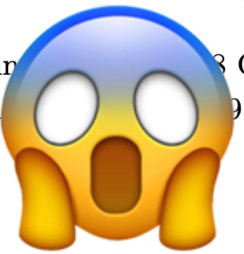
An illustration of the impact of galaxy bias uncertainties u BOSS DR12 galaxy power spectrum

Alexandre Barreira[a,b]

[a]Excellence Cluster ORIGINS, Boltzmann ... Garch...

[b]Ludwig-Maximilians-Universität, Schelli... München, Germany





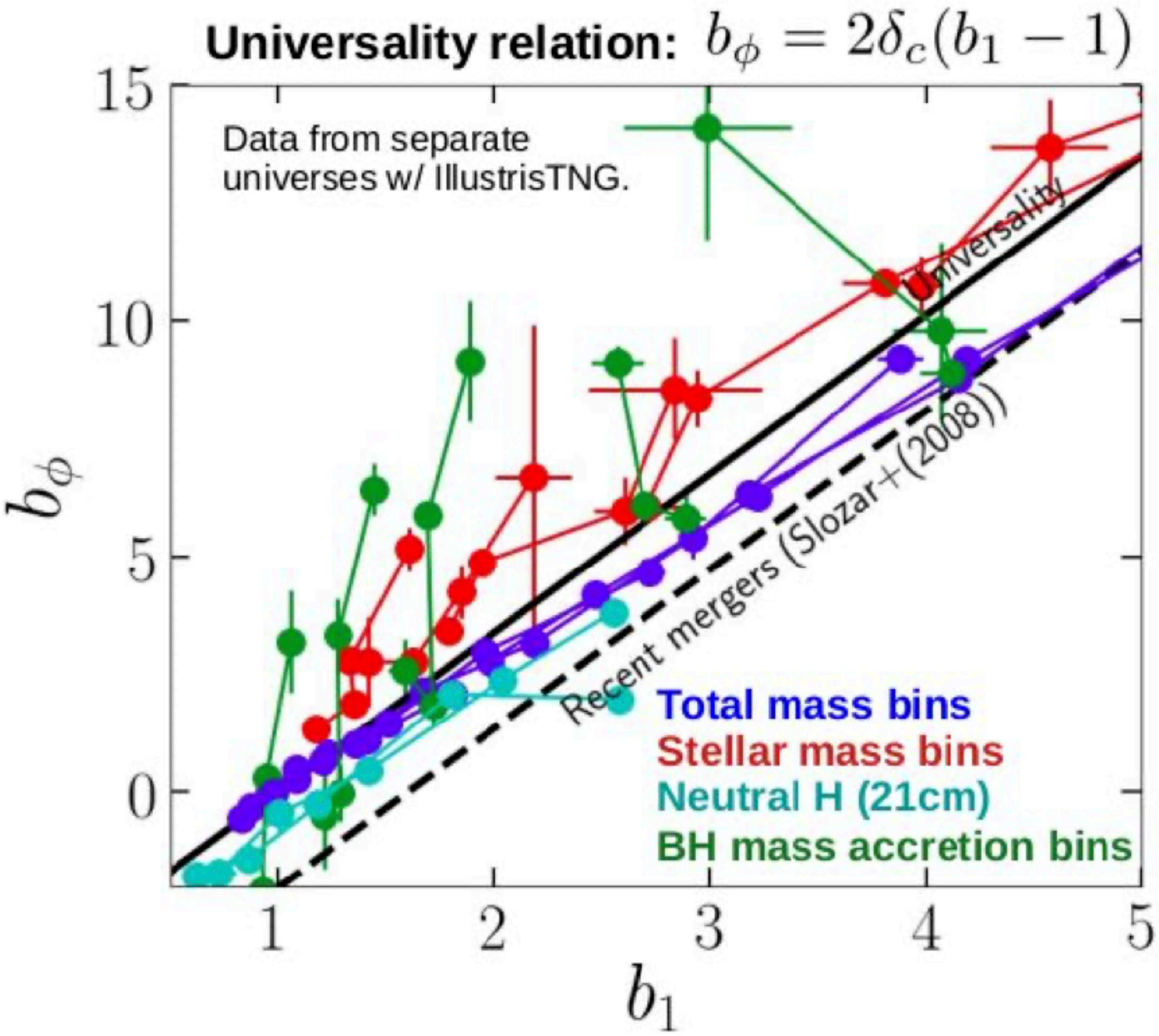
Same BOSS galaxy data, different priors for b_ϕ and b_1

↓

$f_{NL} = 16 \pm 16$ ***or***

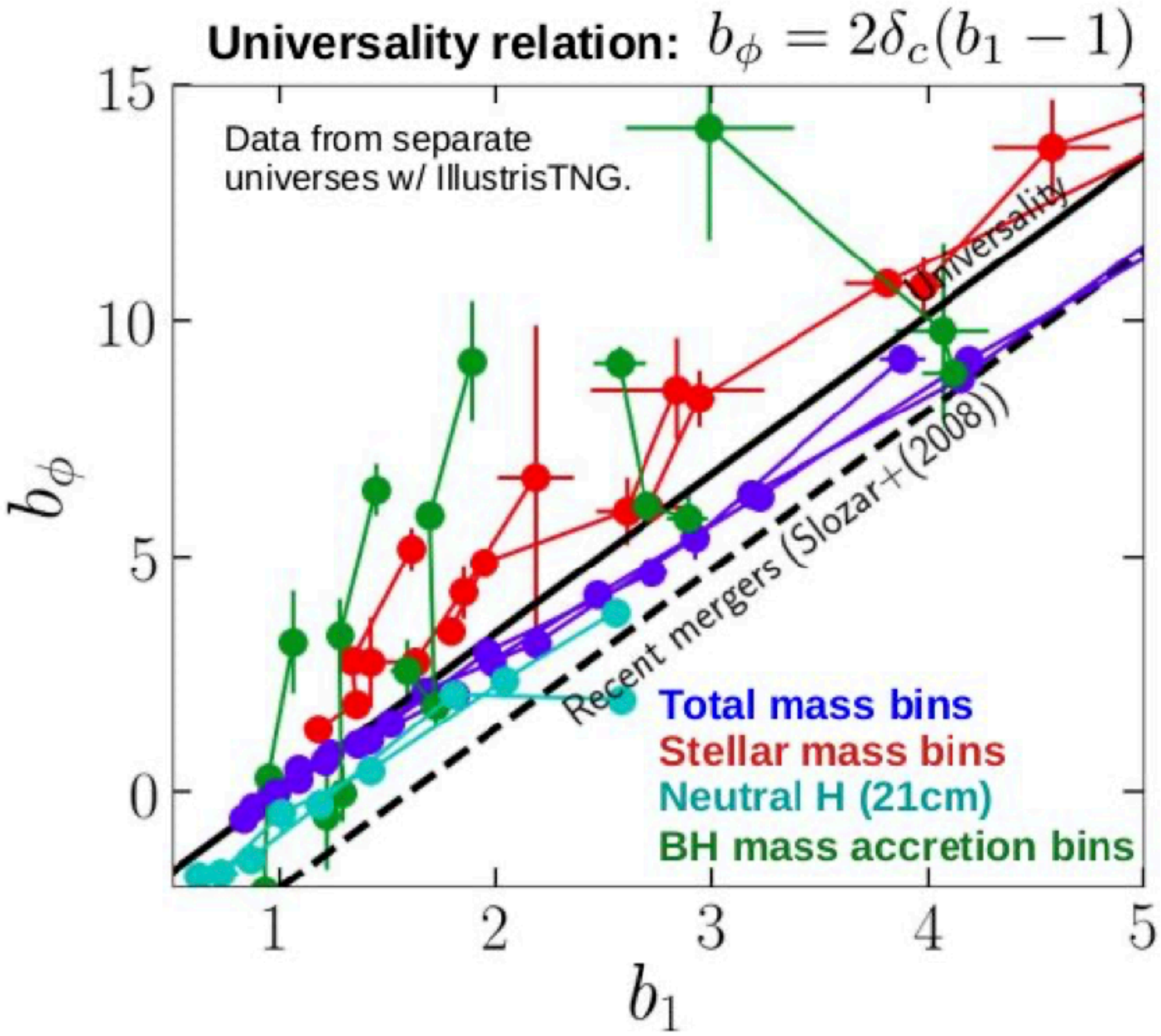
$f_{NL} = 230 \pm 226$

Can galaxy simulations help?



Barreira++ arXiv:2006.09368 , arXiv:2107.06887 , arXiv:2112.03253

Can galaxy simulations help?

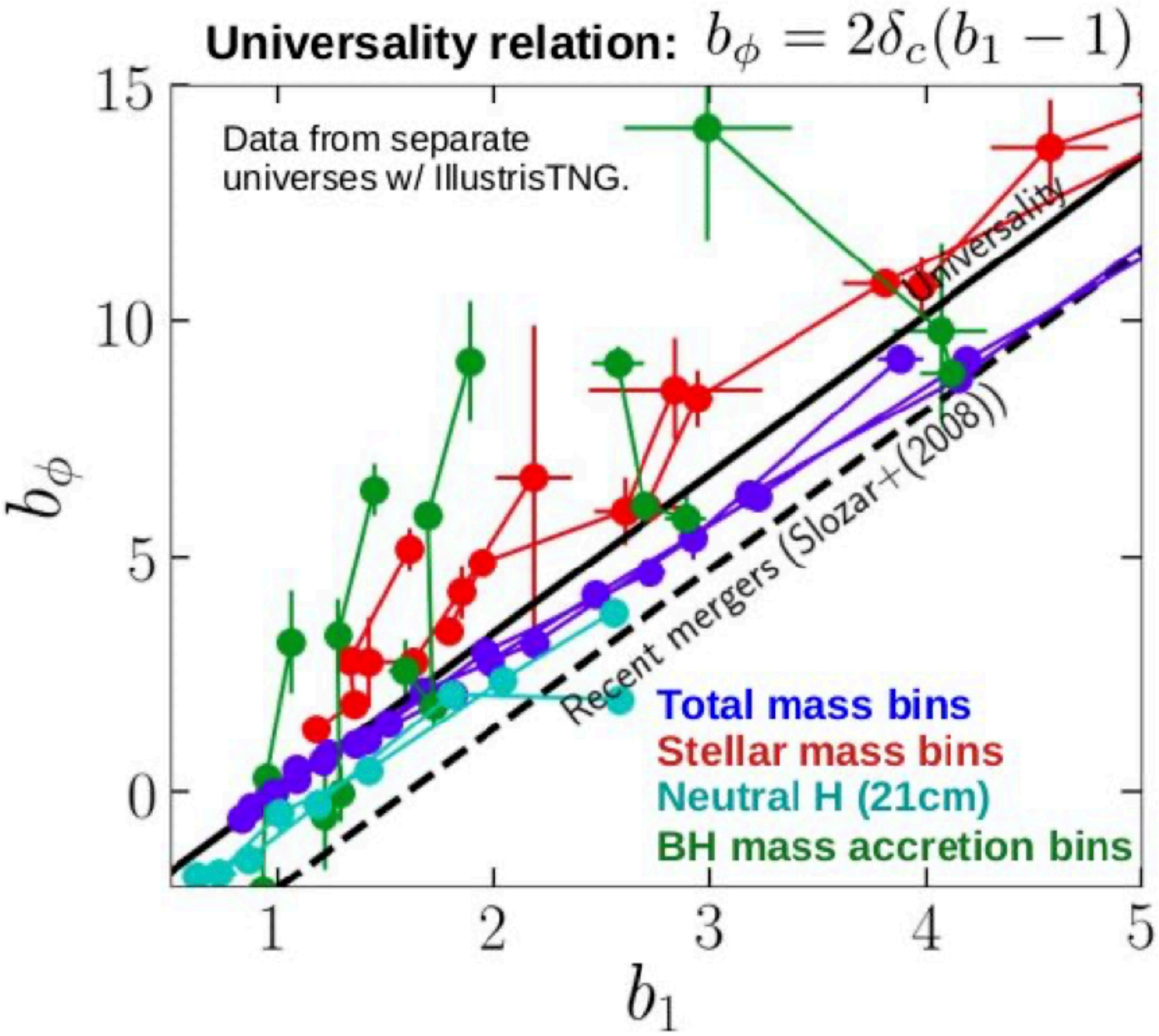


What about different parameters within the same galaxy model?

What about other galaxy models?

Barreira++ arXiv:2006.09368 , arXiv:2107.06887 , arXiv:2112.03253

Can galaxy simulations help?



What about different parameters within the same galaxy model?

What about other galaxy models?

How much better could our cosmology be if we quantify our ignorance around galaxy formation?

Barreira++ arXiv:2006.09368 , arXiv:2107.06887 , arXiv:2112.03253

Let's pin down b_ϕ with the Santa Cruz Semi-Analytic Model for Galaxy Formation

FLATIRON INSTITUTE

Separate Universe simulations to measure b_ϕ

$(205\ h\text{-}1\ cMpc)^3$, $N=1280^3$, same initial conditions, different σ_8

+

Several dozen unique runs of the SC-SAM

Vary 3 parameters for SN & AGN feedback (see Perez, Genel, + 2023)

Separate Universes method for measuring galaxy bias

FLATIRON INSTITUTE

$$\delta_g(\vec{k},z) \overset{\text{L.O.}}{=} b_1(z)\,\delta_m(\vec{k},z) + b_\phi(z)\, f_{NL}^{loc}\, \phi(\vec{k},z)$$

$$\hookrightarrow \delta_g(\vec{k},z) \equiv \frac{n_g(\vec{k},z)}{\bar{n}_g(\vec{k},z)} - 1; \quad \frac{\text{local}}{\text{mean}} \text{ comoving galaxy densities}$$

Separate Universes method for measuring galaxy bias

$$\delta_g(\vec{k},z) \overset{\text{L.O.}}{=} b_1(z)\,\delta_m(\vec{k},z) + b_\phi(z)\, f_{NL}^{loc}\, \phi(\vec{k},z)$$

$$\hookrightarrow \delta_g(\vec{k},z) \equiv \frac{n_g(\vec{k},z)}{\bar{n}_g(\vec{k},z)} - 1; \quad \frac{\text{local}}{\text{mean}} \text{ comoving galaxy densities}$$

The local structure formation response to long-wavelength perturbations is <u>equivalent</u> to structure formation in a different cosmology

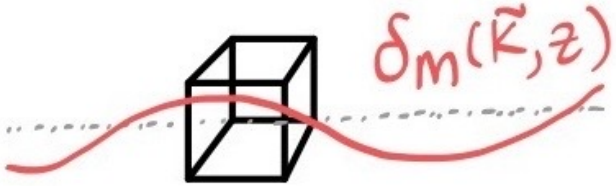
Separate Universes method for measuring galaxy bias

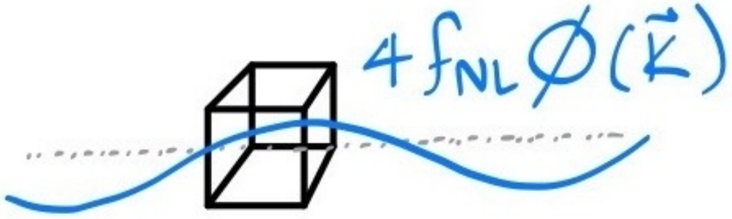
FLATIRON INSTITUTE

$$\delta_g(\vec{k},z) \overset{L.O.}{=} b_1(z)\,\delta_m(\vec{k},z) + b_\phi(z)\, f_{NL}^{loc}\, \phi(\vec{k},z)$$

$$\hookrightarrow \delta_g(\vec{k},z) \equiv \frac{n_g(\vec{k},z)}{\bar{n}_g(\vec{k},z)} - 1; \quad \frac{\text{local}}{\text{mean}} \text{ comoving galaxy densities}$$

$$b_1 \equiv \frac{d\ln(n_g(z))}{d\delta_L(z)} \qquad b_\phi \equiv 4\frac{d\ln(n_g(z))}{d\delta A_s}$$





The local structure formation response to long-wavelength perturbations is <u>equivalent</u> to structure formation in a different cosmology

Bias → the response of galaxy density to a different cosmology!

a.k.a. "Separate Universes"

(from the peak background split)

Let's pin down b_ϕ with the Santa Cruz Semi-Analytic Model for Galaxy Formation

Separate Universe simulations to measure b_ϕ

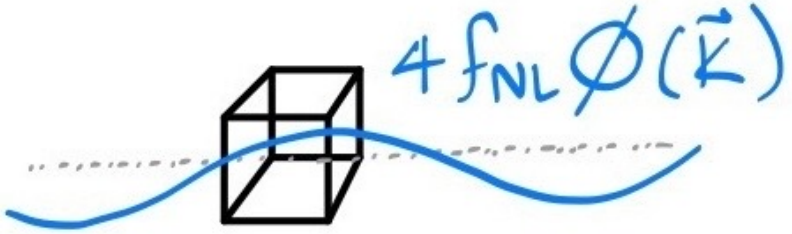
$(205\ h\text{-}1\ \mathrm{cMpc})^3$, $N=1280^3$, same initial conditions, different σ_8

+

Several dozen unique runs of the SC-SAM

Vary 3 parameters for SN & AGN feedback (see Perez, Genel, + 2023)

$$b_\phi \equiv 4\,\frac{d\ln(n_g(z))}{d\,\delta A_s} \rightarrow 4 f_{NL}\,\phi_L$$



b_ϕ → how galaxy counts/density respond to changing the primordial power spectrum (A_s or σ_8)

Let's pin down b_ϕ with the Santa Cruz Semi-Analytic Model for Galaxy Formation

FLATIRON INSTITUTE

Separate Universe simulations to measure b_ϕ

$(205\ h^{-1}\ \mathrm{cMpc})^3$, $N=1280^3$, same initial conditions, different σ_8

+

Several dozen unique runs of the SC-SAM

Vary 3 parameters for SN & AGN feedback (see Perez, Genel, + 2023)

A_{SN1} x, $+A_{SN2}$

$$\dot{m}_{\mathrm{out}} = \epsilon_{\mathrm{SN}} \left(\frac{V_0}{V_{\mathrm{c}}}\right)^{\alpha_{\mathrm{rh}}} \dot{m}_*,$$

⬆ mass outflow rate due to SN & massive stars

⬇ mass ejected by an AGN in radio jets

A_{AGN1} x

$$\dot{m}_{\mathrm{radio}} = \kappa_{\mathrm{radio}} \left[\frac{kT}{\Lambda(T, Z_{\mathrm{h}})}\right] \left(\frac{M_{\mathrm{BH}}}{10^8\ \mathrm{M}_\odot}\right)$$

Let's pin down b_ϕ with the Santa Cruz Semi-Analytic Model for Galaxy Formation

FLATIRON INSTITUTE

Separate Universe simulations to measure b_ϕ

$(205\ h^{-1}\ cMpc)^3$, $N=1280^3$, same initial conditions, different σ_8

+

Several dozen unique runs of the SC-SAM

Vary 3 parameters for SN & AGN feedback (see Perez, Genel, + 2023)

→

Understand how b_ϕ and b_ϕ (b_1) varies with galaxy selection across different parametrizations of one galaxy formation model

Let's pin down b_ϕ with the Santa Cruz Semi-Analytic Model for Galaxy Formation

FLATIRON INSTITUTE

Separate Universe simulations to measure b_ϕ

$(205\ h^{-1}\ \mathrm{cMpc})^3$, $N=1280^3$, same initial conditions, different σ_8

\+

Several dozen unique runs of the SC-SAM

Vary 3 parameters for SN & AGN feedback (see Perez, Genel, + 2023)

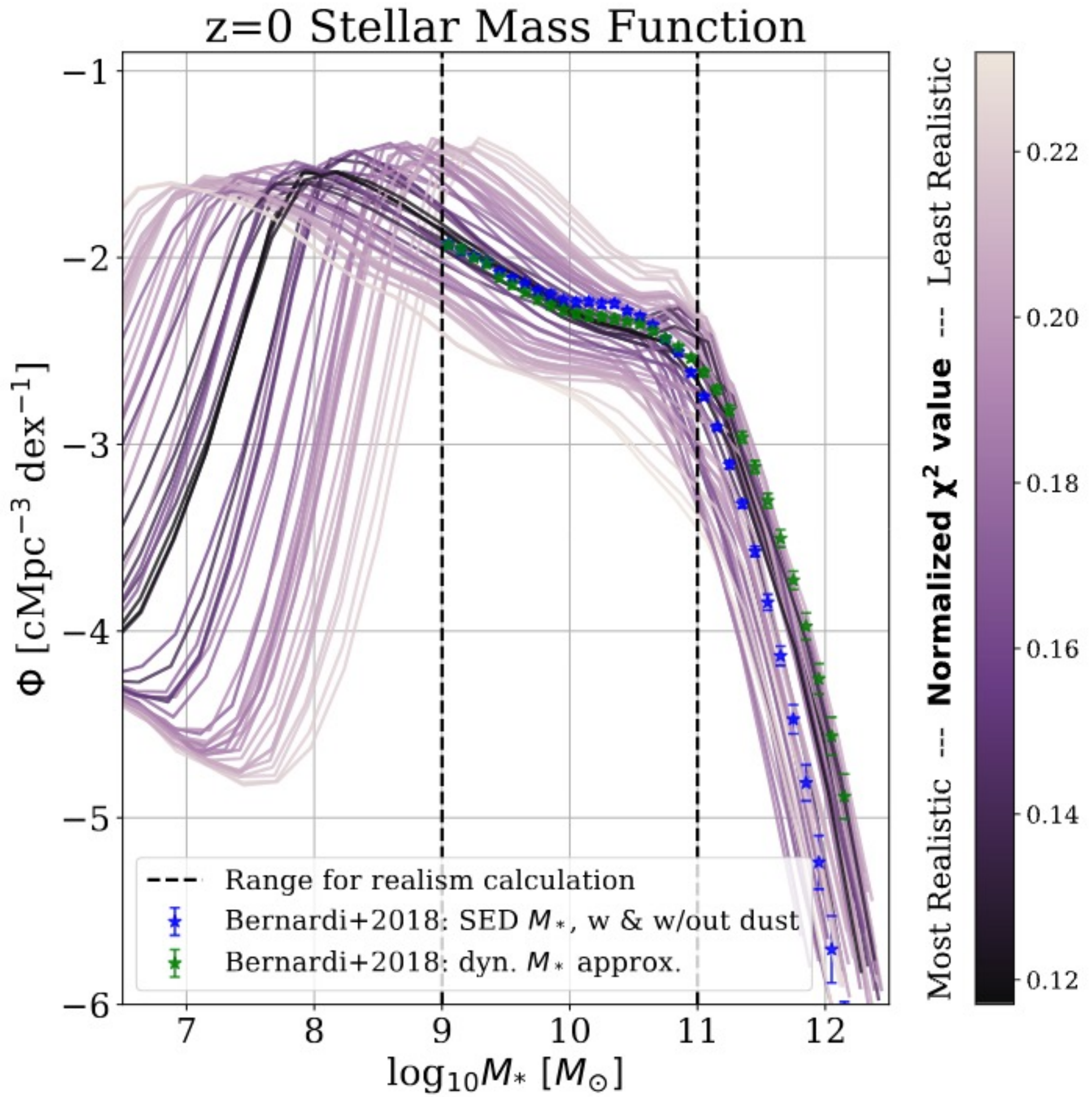
→

Understand how b_ϕ and b_ϕ (b_1) varies with galaxy selection across different parametrizations of one galaxy formation model

↑

Do we agree with the IllustrisTNG findings, and the few others measured for SAMs? Let's get physically-motivated b_ϕ priors for f_{NL} analyses?

How “realistic” is a given SAM parametrization?



Here, we compare to z=0 stellar mass function between $9 < \log_{10}(M_{star}) < 11$

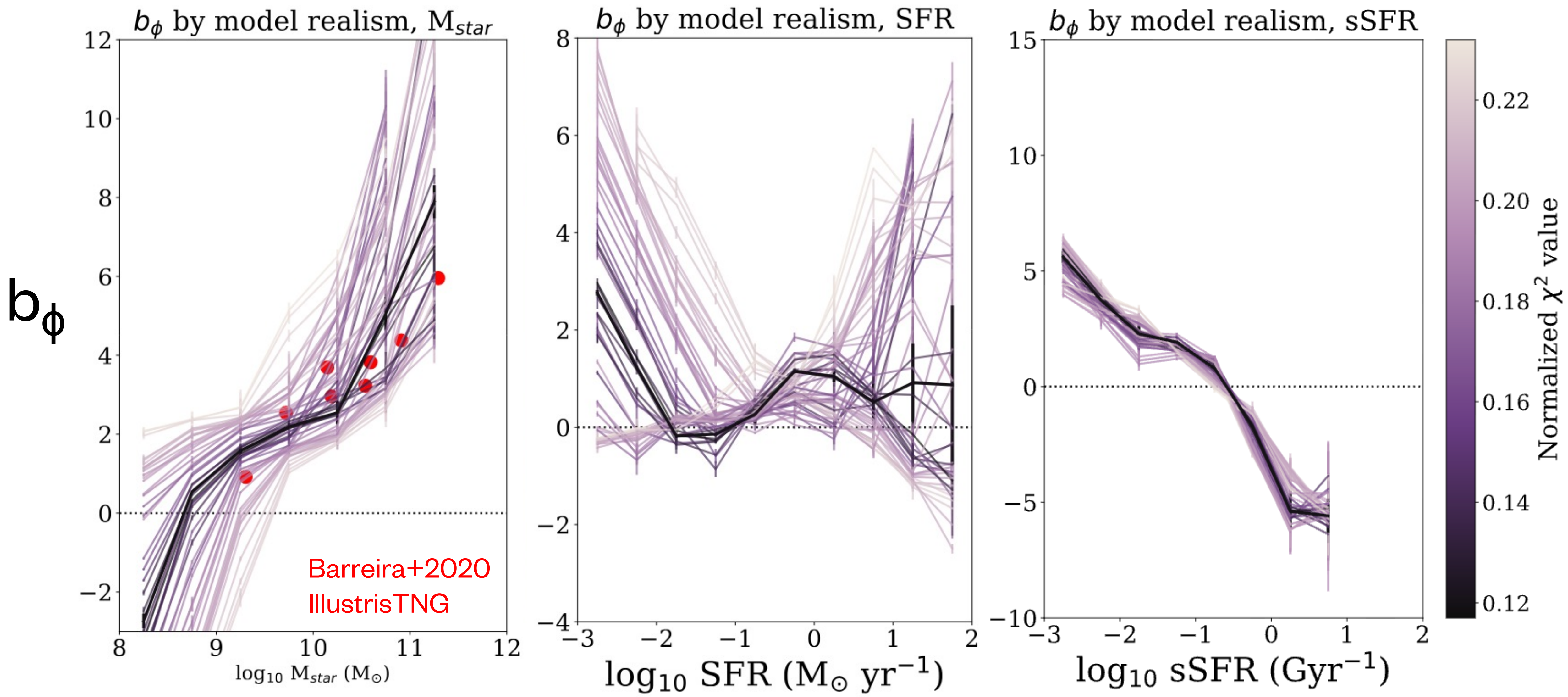
Light- to-stellar mass is hard science

Black → beige is most to least ‘realistic’

Even extreme parameters can yield realistic SMFs in unique combinations!

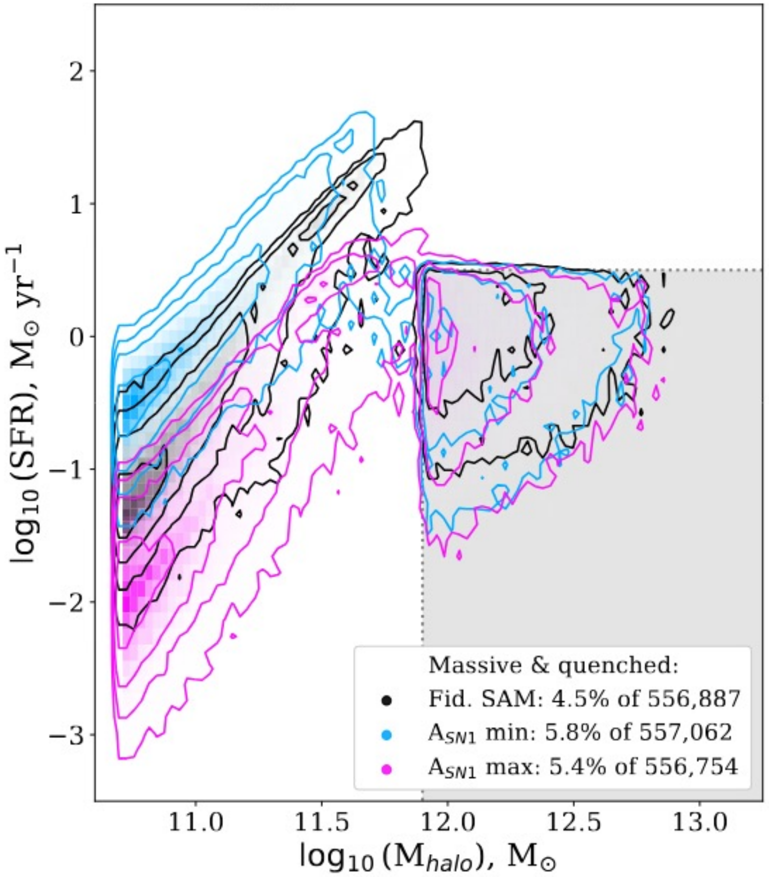
b_ϕ different galaxy selections, z=1

b_ϕ → how galaxy counts/density respond to changing the primordial power spectrum (A_s or σ_8)

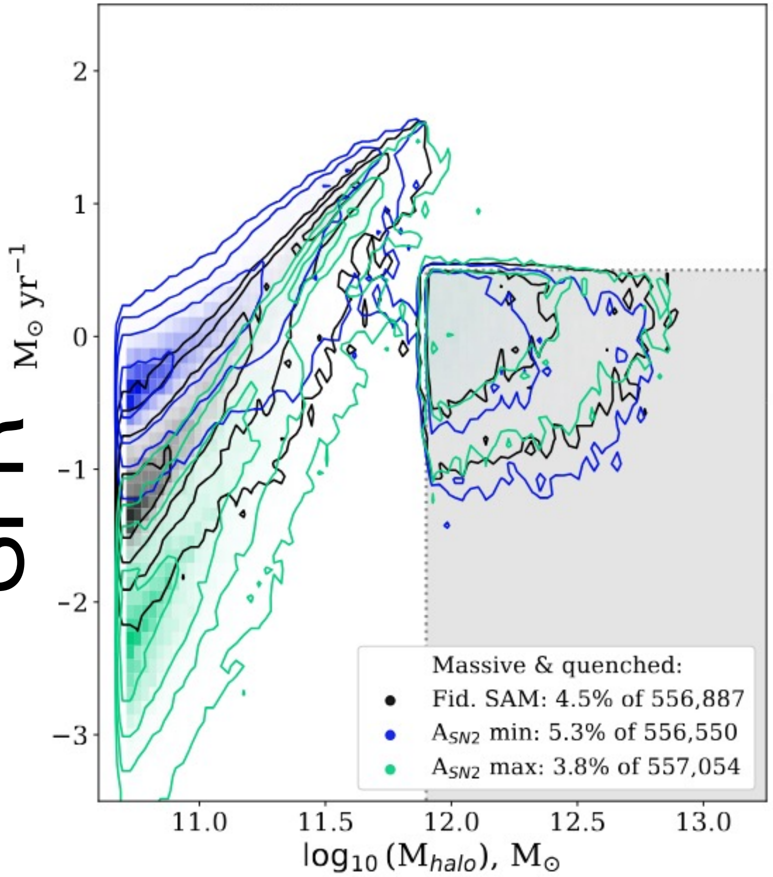


Why? The galaxy-halo connection!

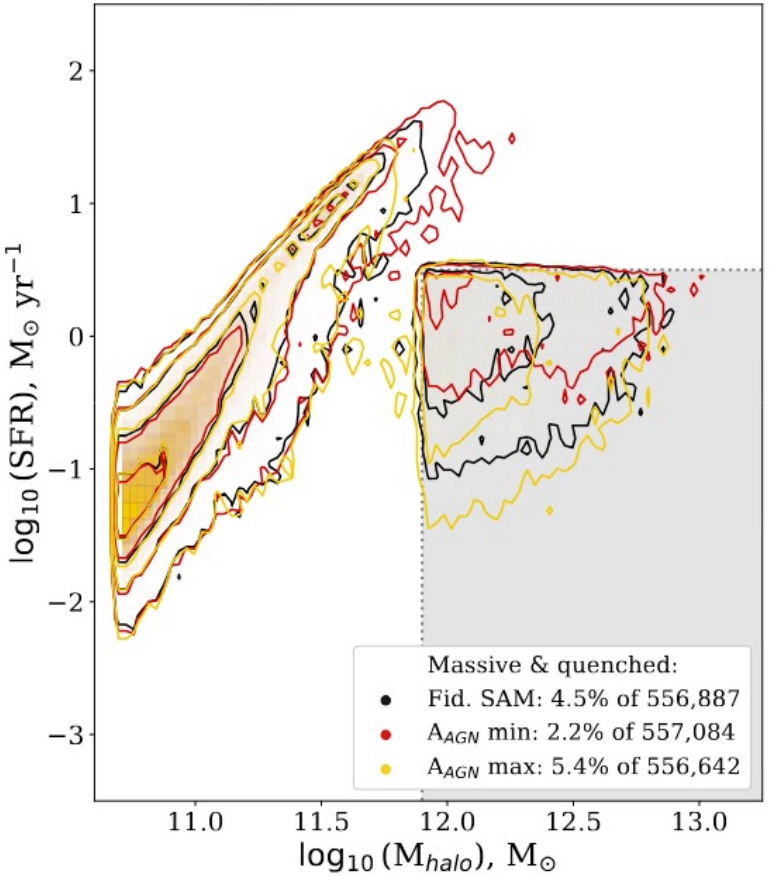
FLATIRON INSTITUTE



SFR



Halo Mass



$$\dot{m}_{\rm out} = \underbrace{A_{SN1} \times \epsilon_{\rm SN}}\left(\frac{V_0}{V_c}\right)^{\alpha_{\rm rh} + A_{SN2}} \dot{m}_*,$$

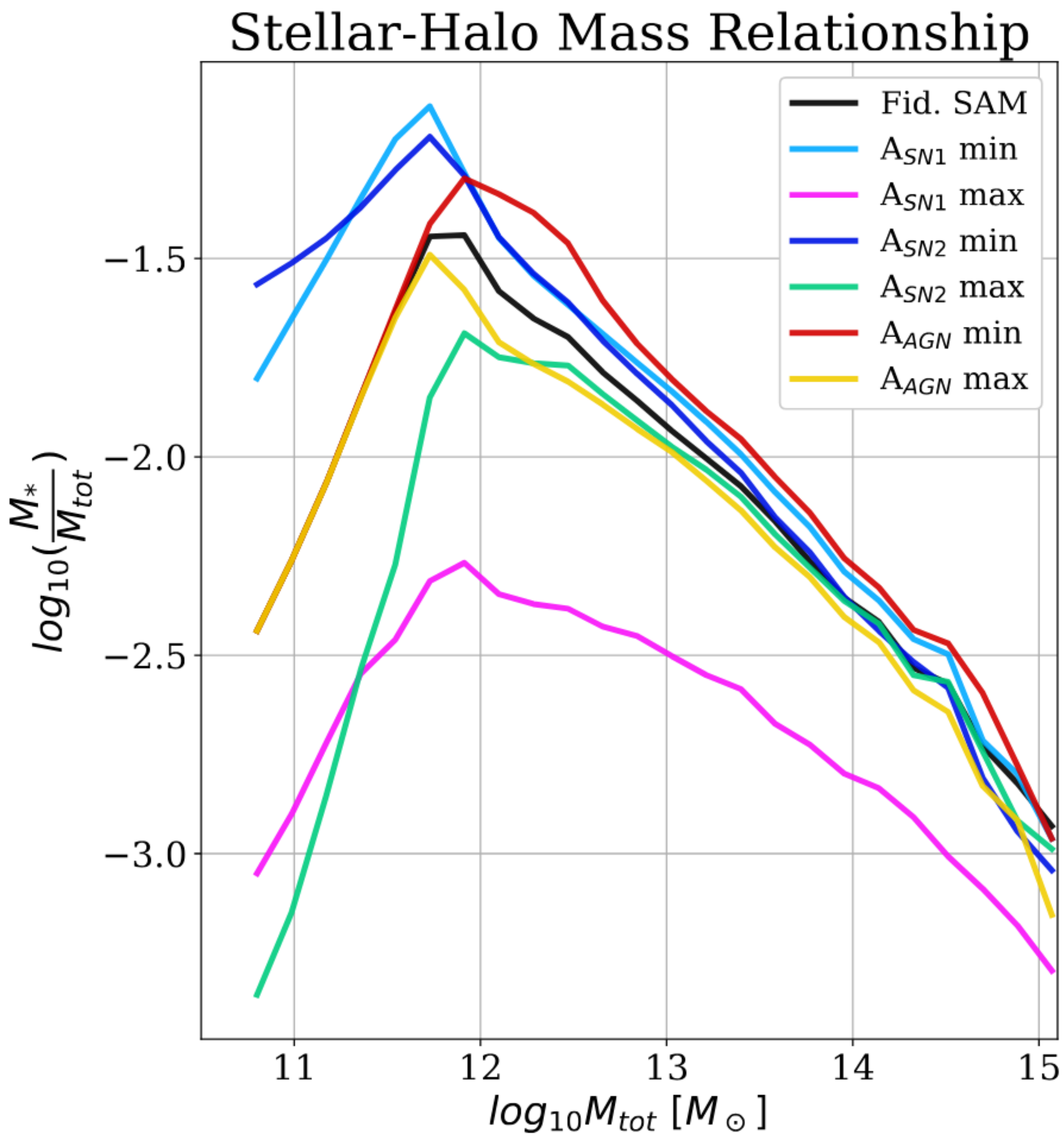
⬆ mass outflow rate due to SN & massive stars

⬇ mass ejected by an AGN in radio jets

$$\dot{m}_{\rm radio} = A_{AGN1} \times \kappa_{\rm radio}\left[\frac{kT}{\Lambda(T, Z_{\rm h})}\right]\left(\frac{M_{\rm BH}}{10^8\,{\rm M}_\odot}\right)$$

Why? The galaxy-halo connection!

FLATIRON INSTITUTE



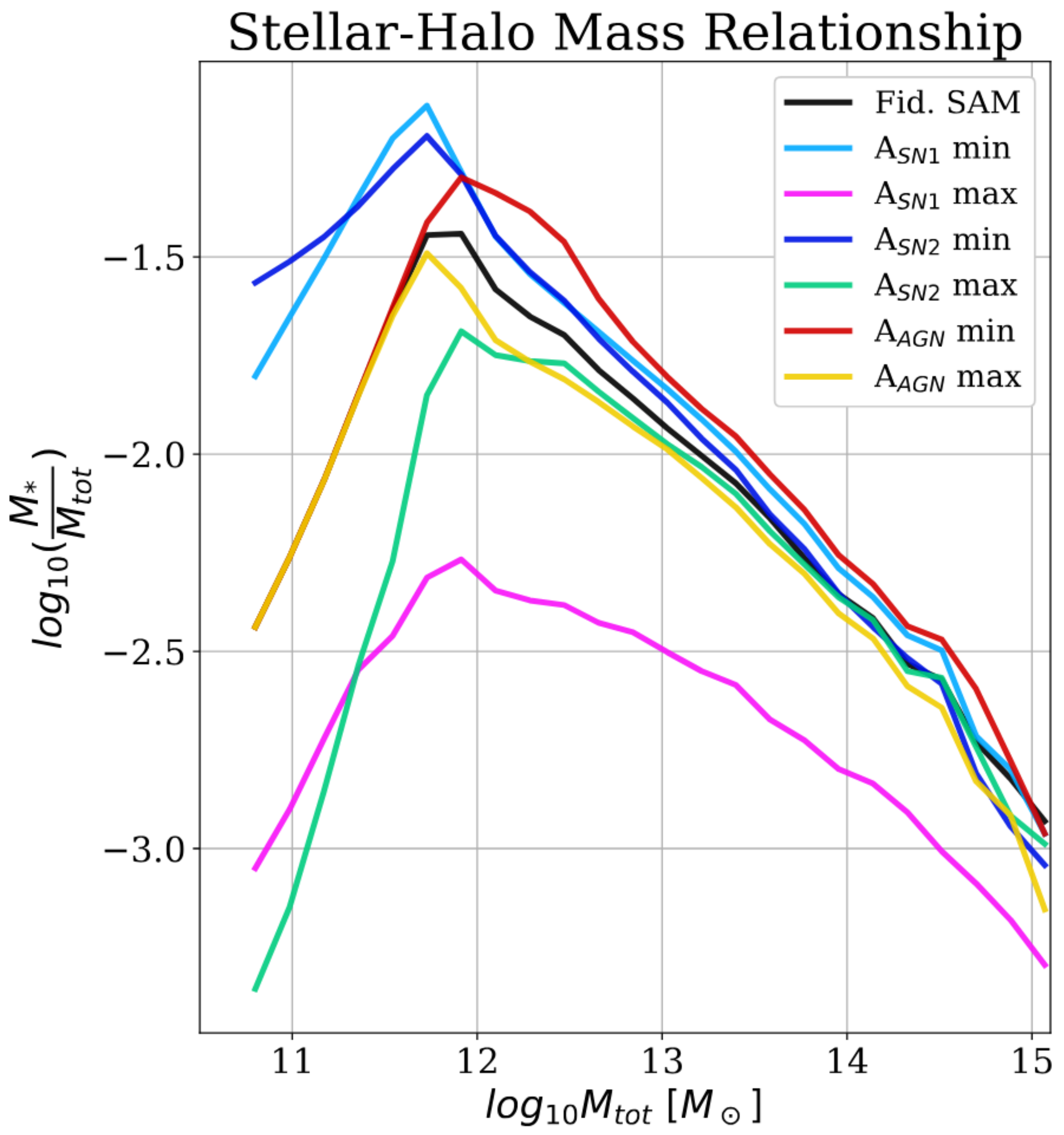
$$\dot{m}_{\rm out} = A_{SN1} \times \epsilon_{\rm SN} \left(\frac{V_0}{V_{\rm c}}\right)^{\alpha_{\rm rh} + A_{SN2}} \dot{m}_*,$$

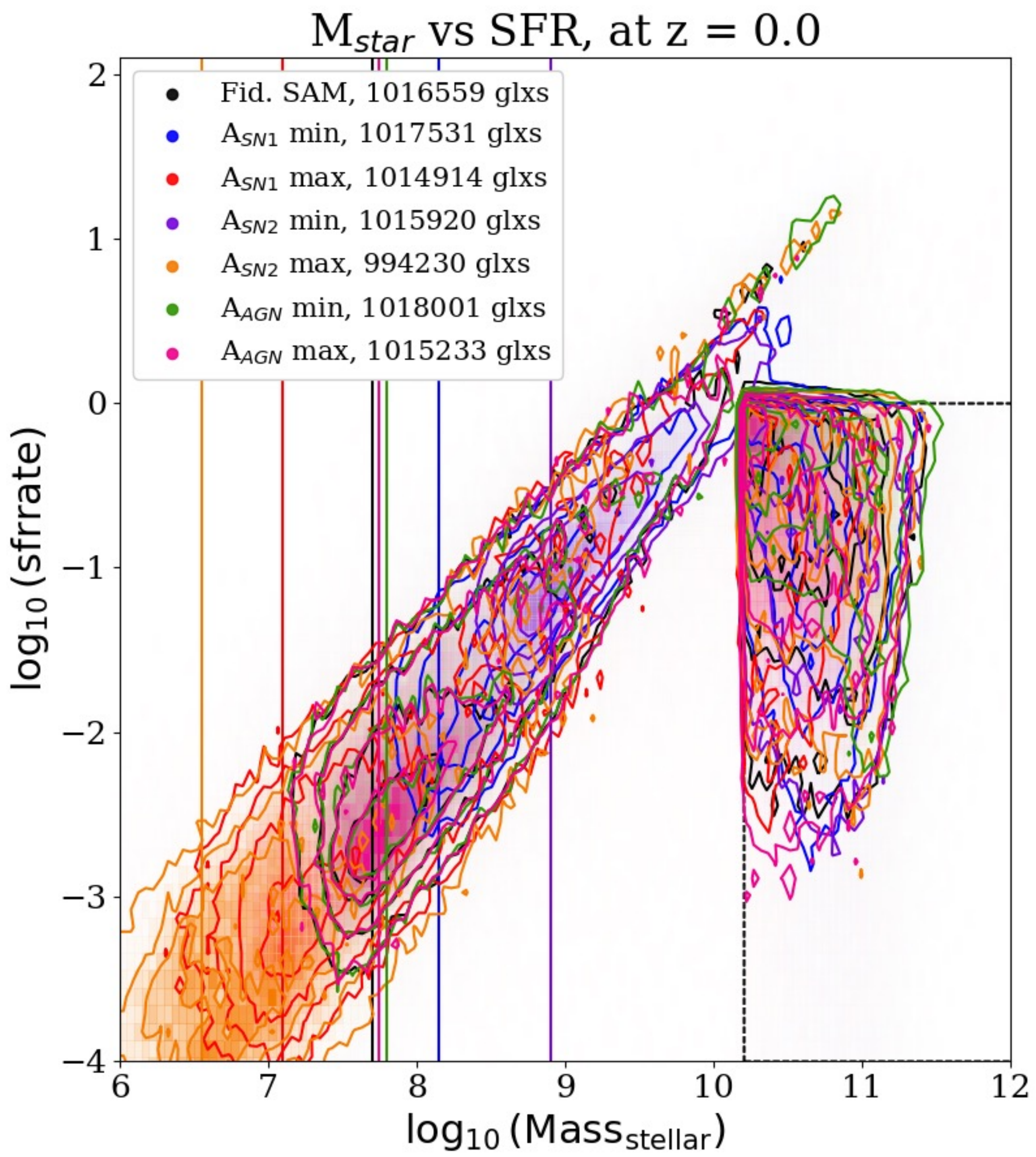
↑ mass outflow rate due to SN & massive stars

↓ mass ejected by an AGN in radio jets

$$\dot{m}_{\rm radio} = A_{AGN1} \times \kappa_{\rm radio} \left[\frac{kT}{\Lambda(T, Z_{\rm h})}\right] \left(\frac{M_{\rm BH}}{10^8\, {\rm M}_\odot}\right)$$

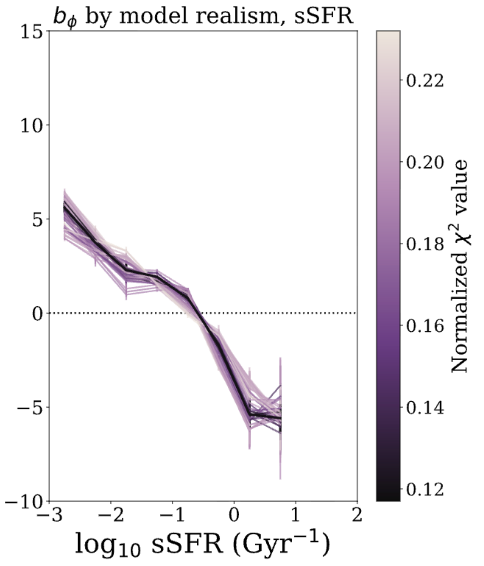
Why? The galaxy-halo connection!

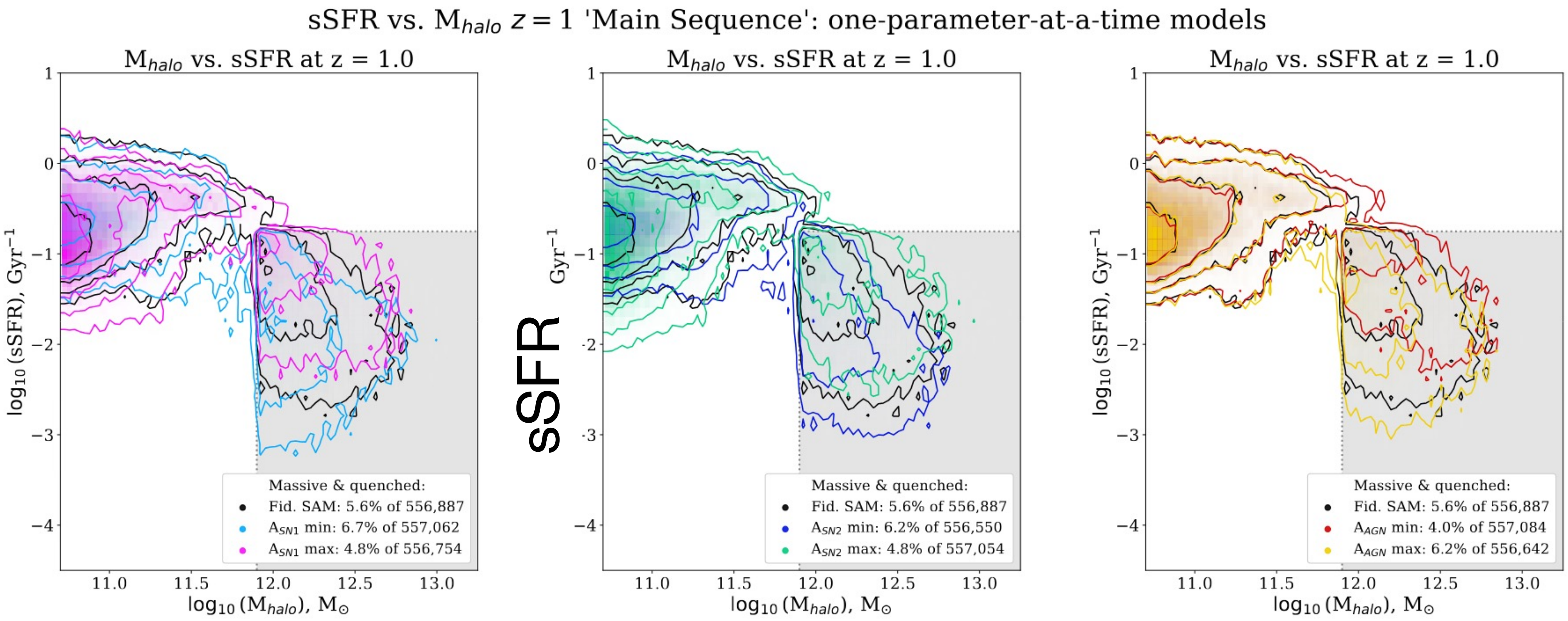




Why? The galaxy-halo connection!

Regulated sSFR- M_{halo} makes b_ϕ(sSFR) standard



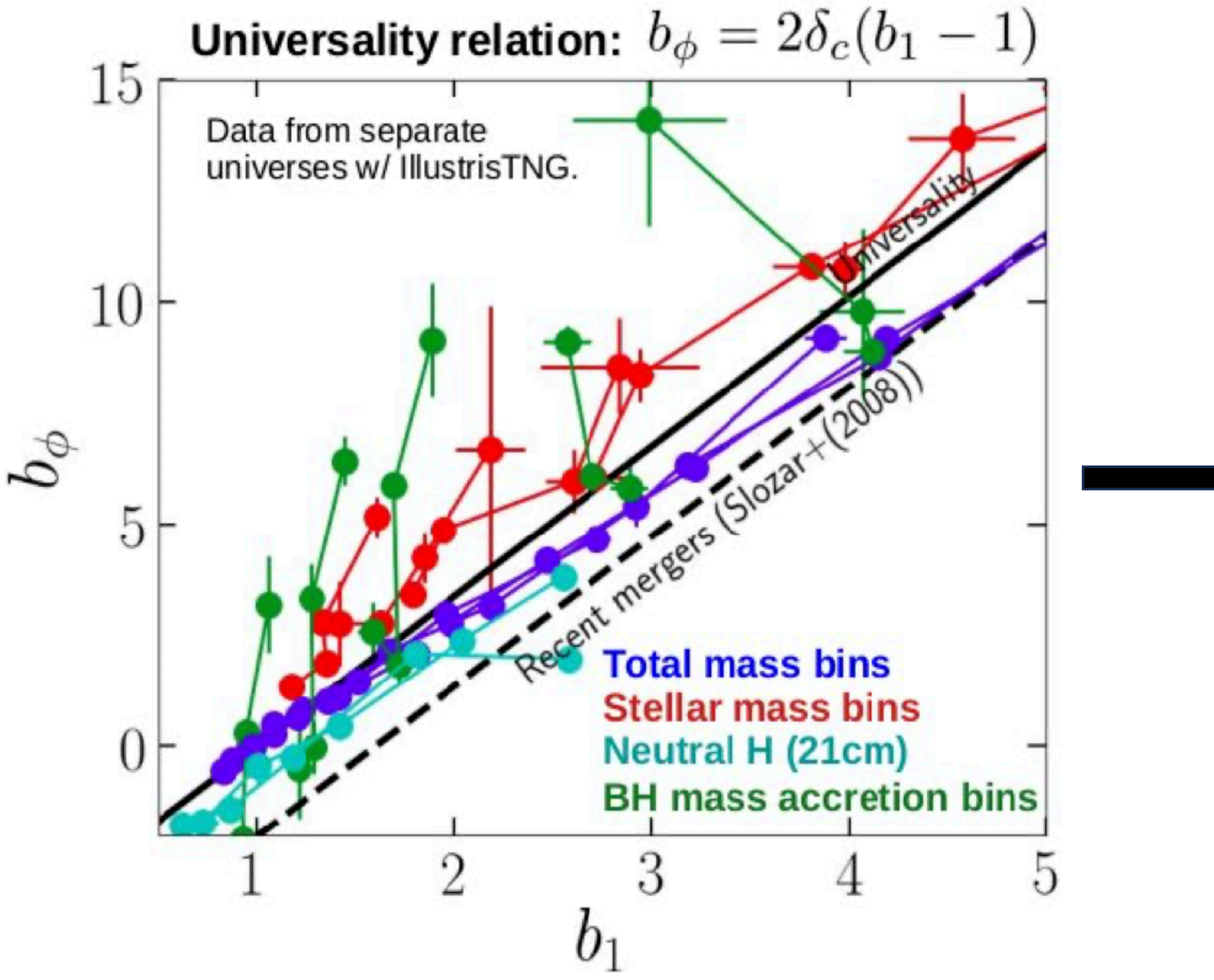


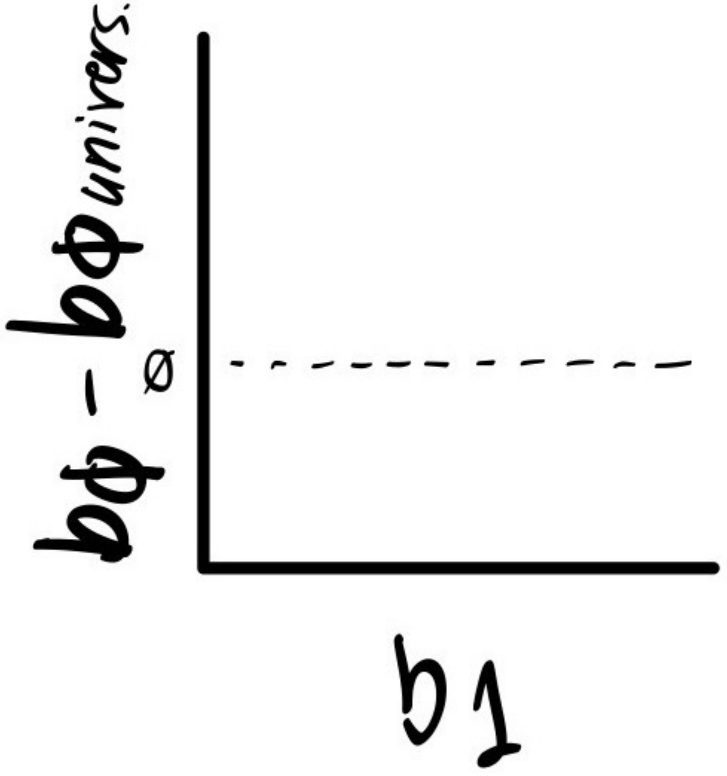
sSFR

Halo Mass

(The SC-SAM is correctly regulating star formation even as the parameters go wild)

(b_ϕ - $b_{\phi,\text{ universality}}$) vs. b_1





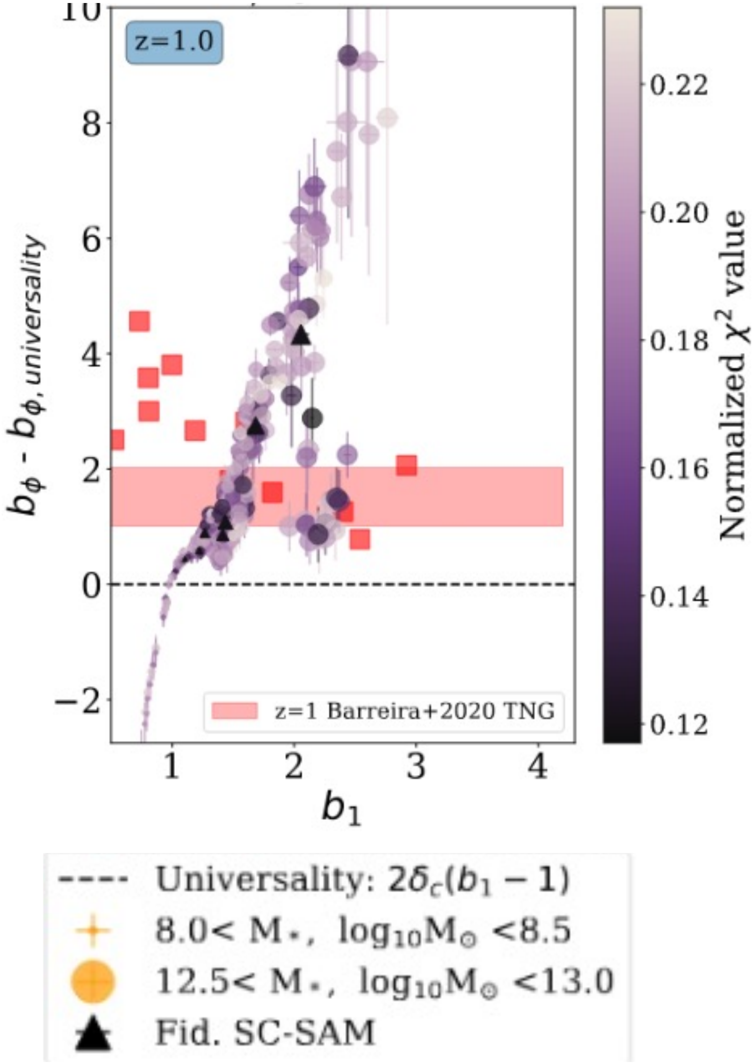
If universality describes well, points will clump near 0!

Barreira++ arXiv:2006.09368 , arXiv:2107.06887, arXiv:2112.03253

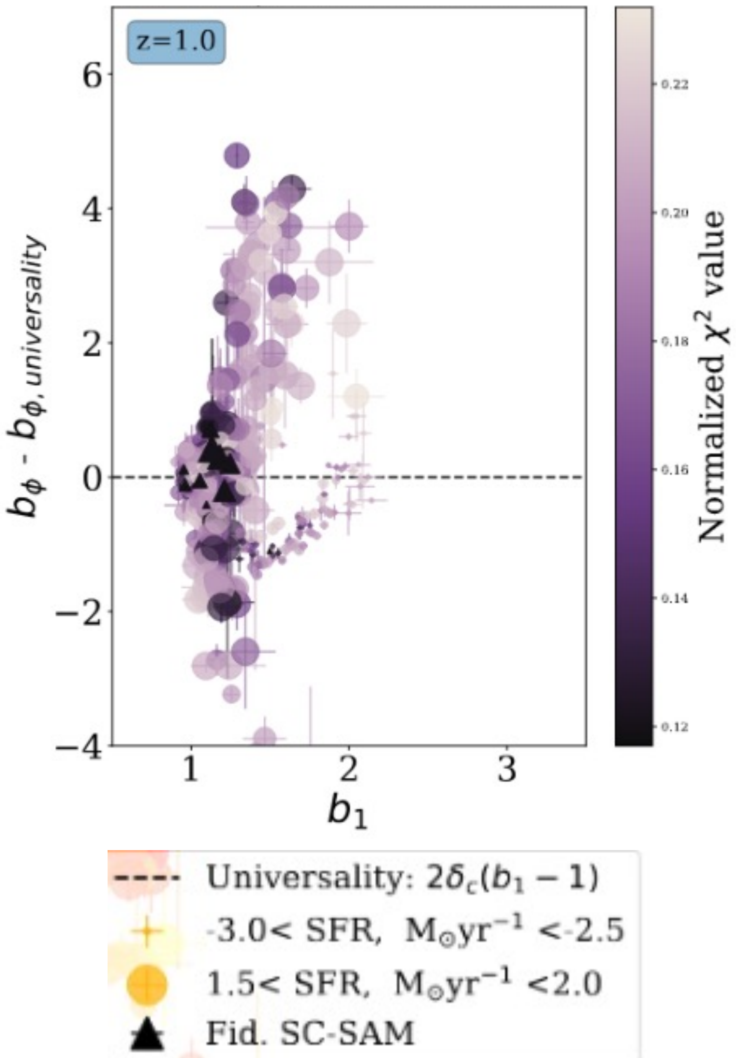
b_ϕ vs. b_1 for M_{star}, SFR, sFR

FLATIRON INSTITUTE

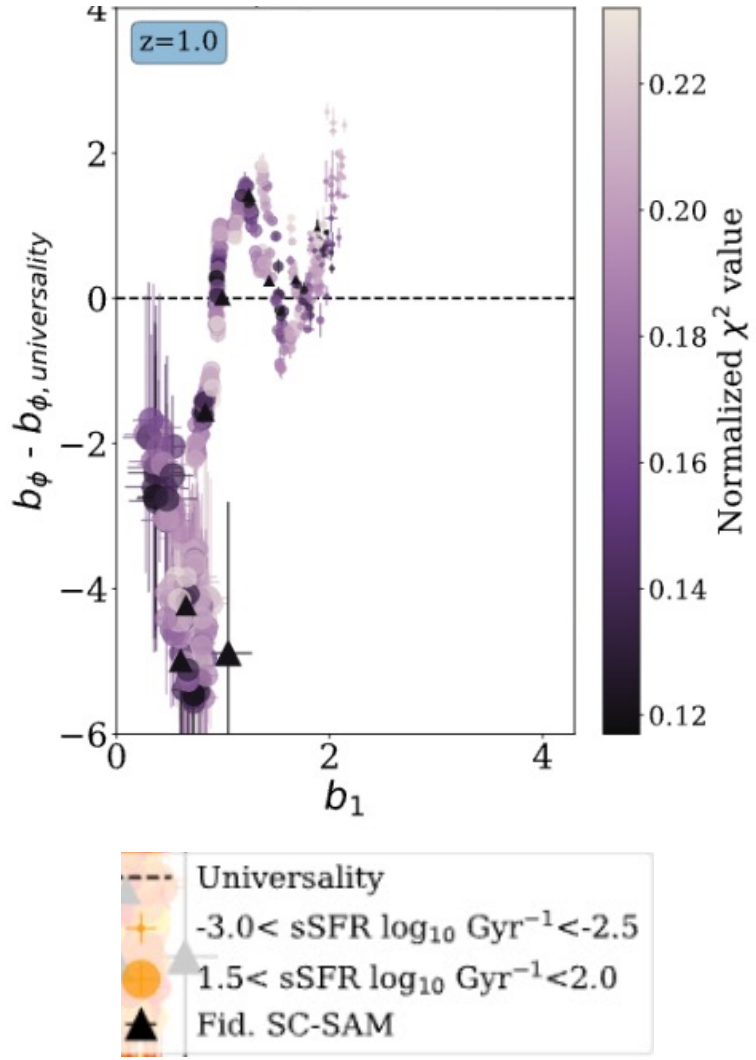
Stellar Mass selections



Instan. SFR selections



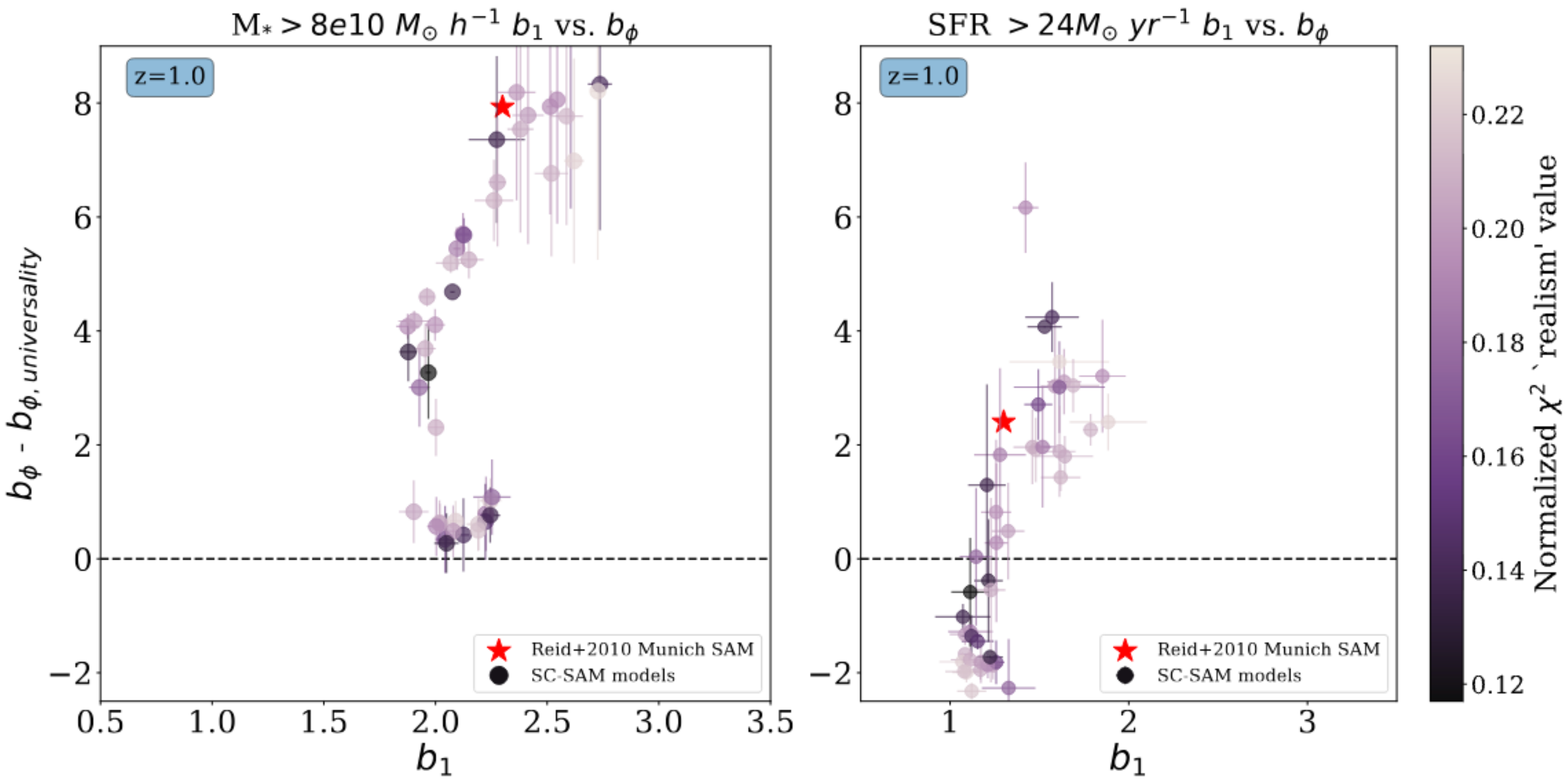
Specific SFR selections



What about other galaxy models?

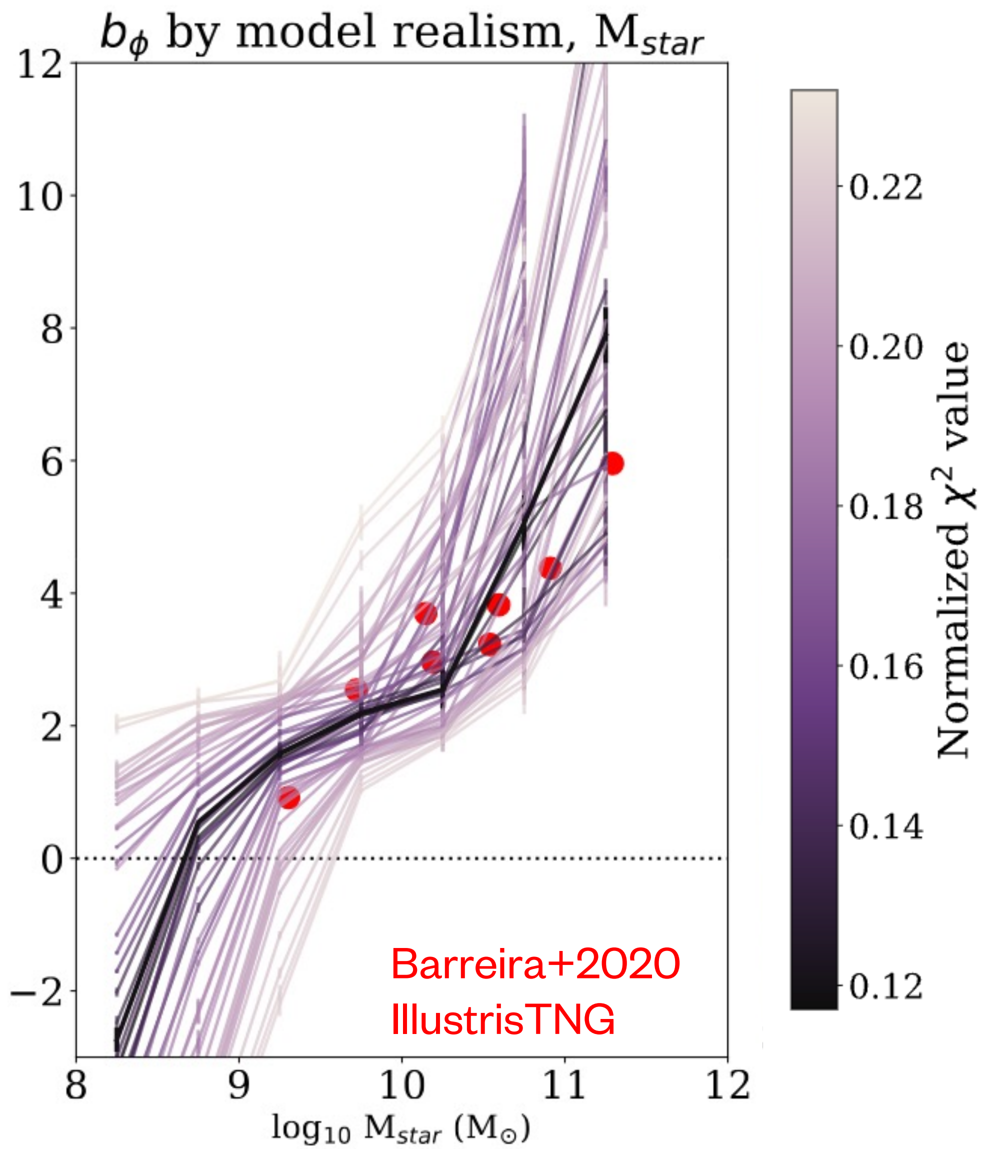
FLATIRON INSTITUTE

High mass galaxies   Strongly star forming galaxies



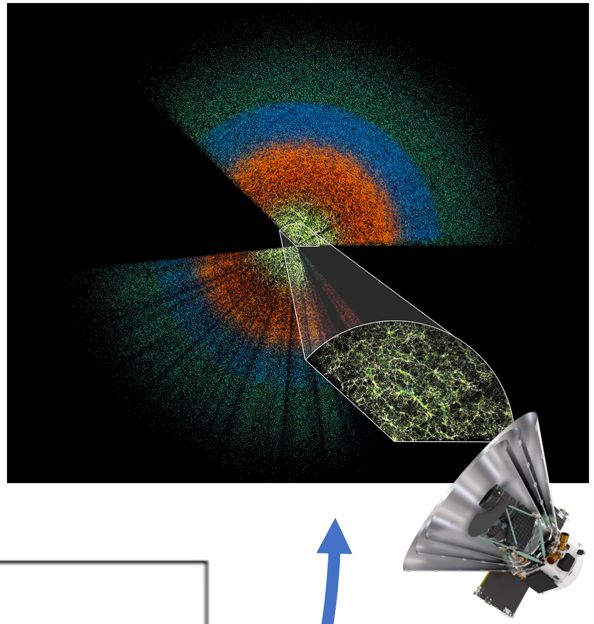
Reid et al. (2010) measure non-Gaussian assembly bias in the Munich SAM

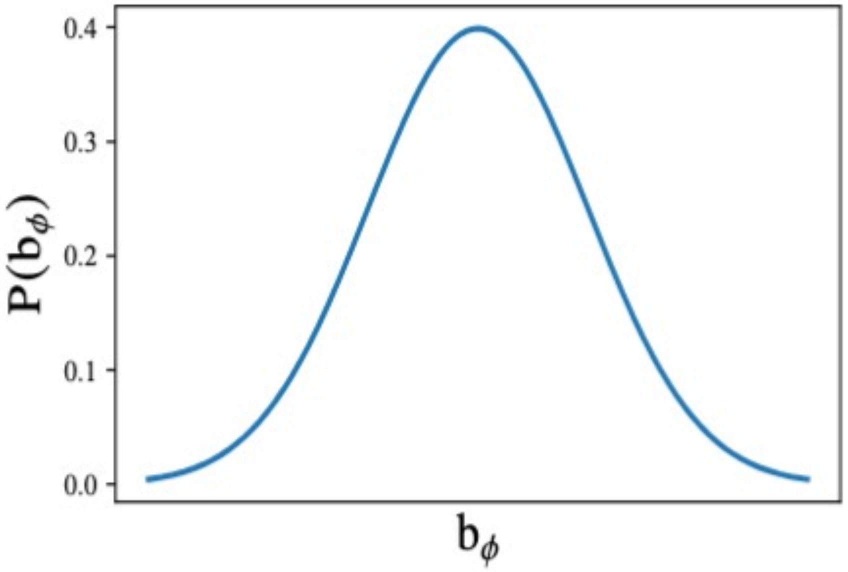
Next: get useful priors for f_{NL} studies



b_ϕ

But not all those models are realistic! The prior should reflect that.



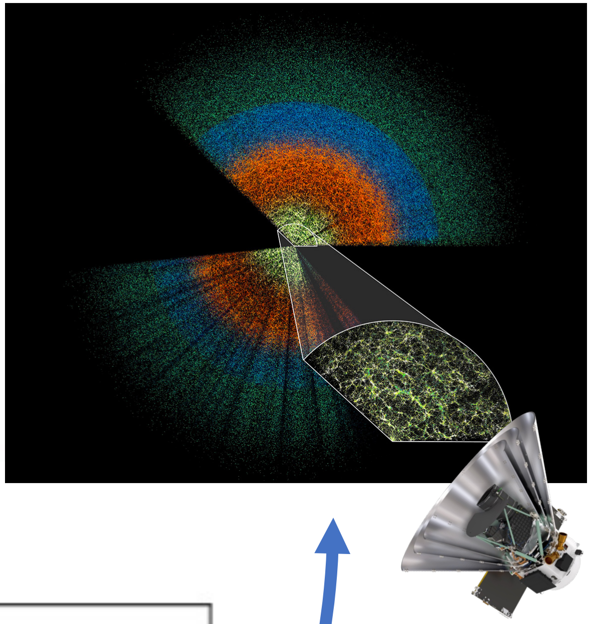


Galaxies selected by e.g. high stellar mass or SFR

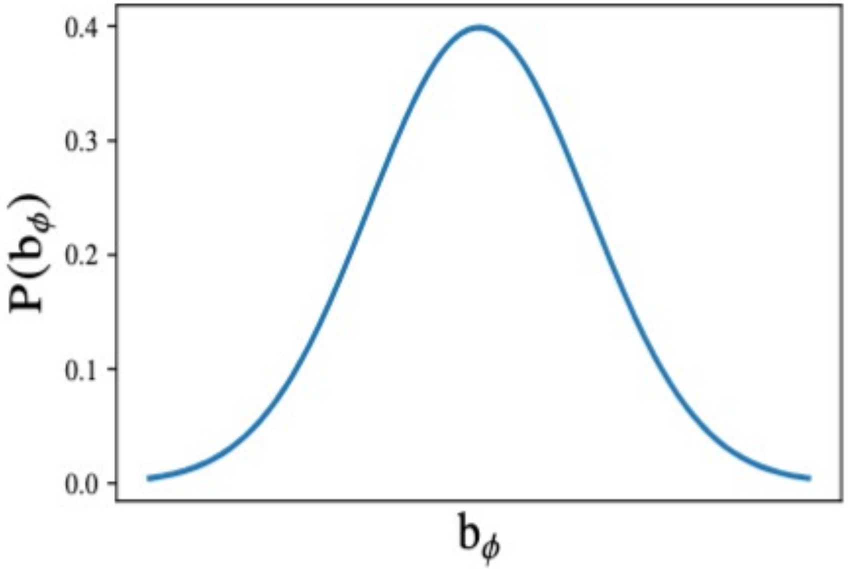
Next: get useful priors for f_{NL} studies

$$p(b_\phi | \mathrm{Obs}) = \int d\Theta_{GF}\; p(b_\phi | \Theta_{GF})\; p(\Theta_{GF} | \mathrm{Obs})$$

- prior on b_ϕ, for some selection, given a comparison observation
- galaxy formation parameters
- b_ϕ for some set of params
- realism of a set of params conditioned on some observable



Galaxies selected by e.g. high stellar mass or SFR



Led by Annie Moore, U of AZ PhD candidate

Next: get useful priors for f_{NL} studies:

FLATIRON INSTITUTE

$$p(b_\phi \mid \mathrm{Obs}) = \int d\Theta_{GF}\; p(b_\phi \mid \Theta_{GF})\; p(\Theta_{GF} \mid \mathrm{Obs})$$

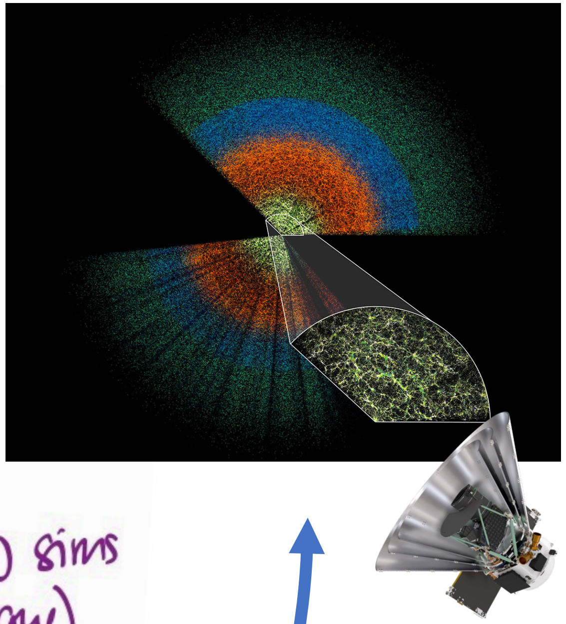
prior on b_ϕ, for some selection, given a comparison observation

galaxy formation parameters

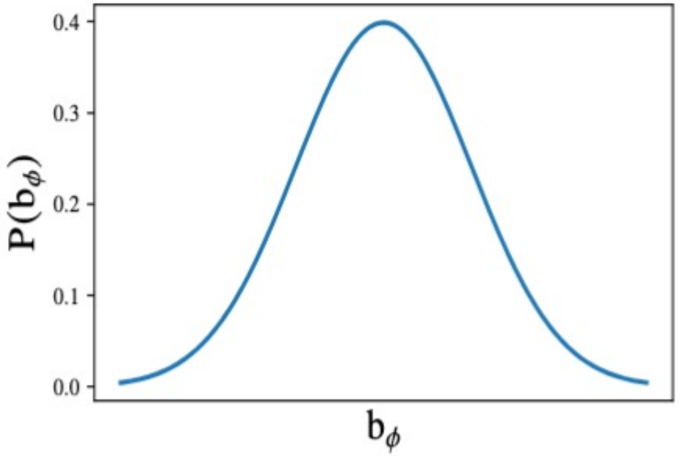
b_ϕ for some set of params
⇓
use separate universe SC-SAM simulations + b_ϕ's

realism of a set of params conditioned on some observable
⇓
use full CAMELS-SAM $L=100\ h^{-1}$Mpc suite of 1000 sims (marginalize over cosmology)

b_ϕ prior marginalized over galaxy formation



Led by Annie Moore, U of AZ PhD candidate



Galaxies selected by e.g. high stellar mass or SFR

Next: get useful priors for f_{NL} studies:

$$p(b_\phi|\mathrm{Obs}) = \int d\Theta_{GF}\, p(b_\phi|\Theta_{GF})\, p(\Theta_{GF}|\mathrm{Obs})$$

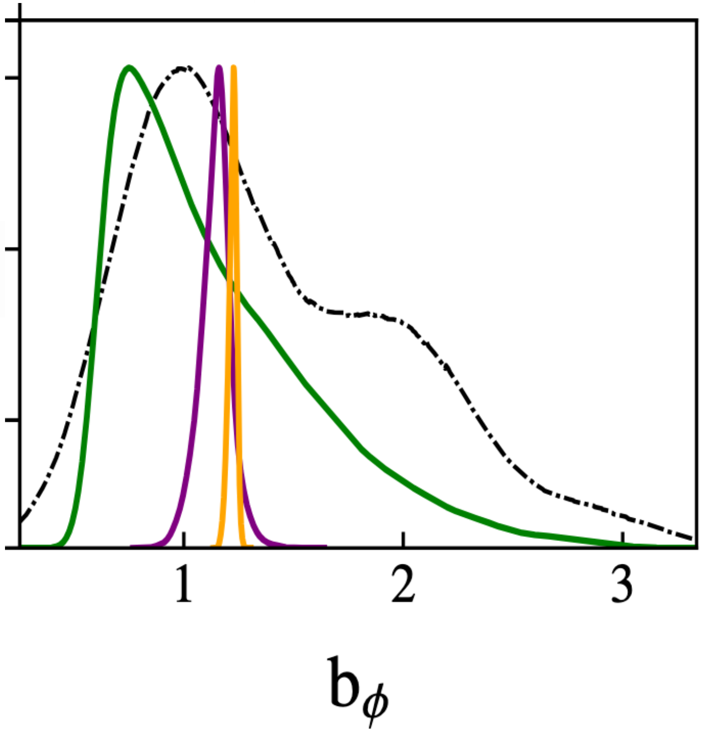
FLATIRON INSTITUTE

DARK ENERGY SPECTROSCOPIC INSTRUMENT

Preliminary results: this leads to significant narrowing in the b_ϕ prior, $\sigma_{b\phi}$ reduced by ~90%!

But it depends on the observable, limited by the fewer Sep. Univ. SC-SAMs

DESI ELG-like selection in stellar mass and specific SFR



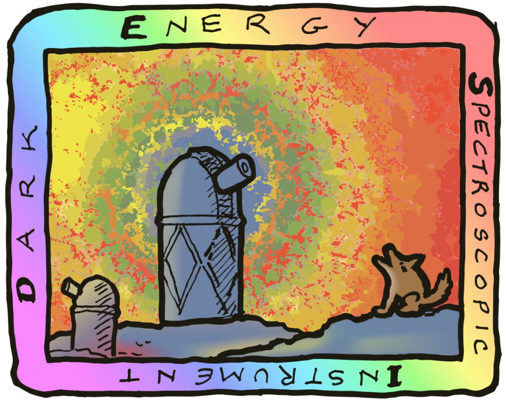
Led by Annie Moore, U of AZ PhD candidate

b_ϕ Priors Conditioned on Observed Data

Metric	$\langle b_\phi \rangle_{SU}$	$\langle b_\phi \rangle_{\chi^2\{Obs\}}$			$p(b_\phi\|\{Obs\})$		
		SMF	SHMR	$M_* - Z_*$	SMF	SHMR†	$M_* - Z_*$
$\langle b_\phi \rangle$	1.43	1.20	1.41	1.39	1.15	1.25	1.15
σ_{b_ϕ}	0.69	0.48	0.58	0.62	0.08	0.02	0.45

Next: get useful priors for f_{NL} studies:

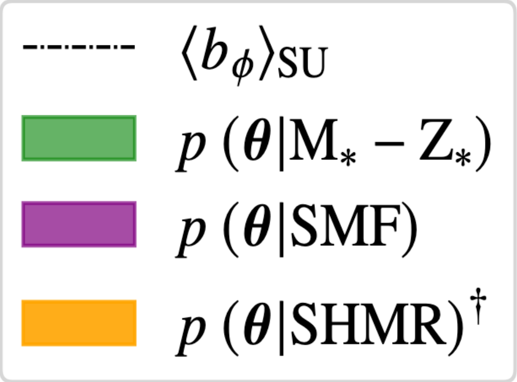
$$p(b_\phi | \mathrm{Obs}) = \int d\Theta_{GF}\, p(b_\phi | \Theta_{GF})\, p(\Theta_{GF} | \mathrm{Obs})$$

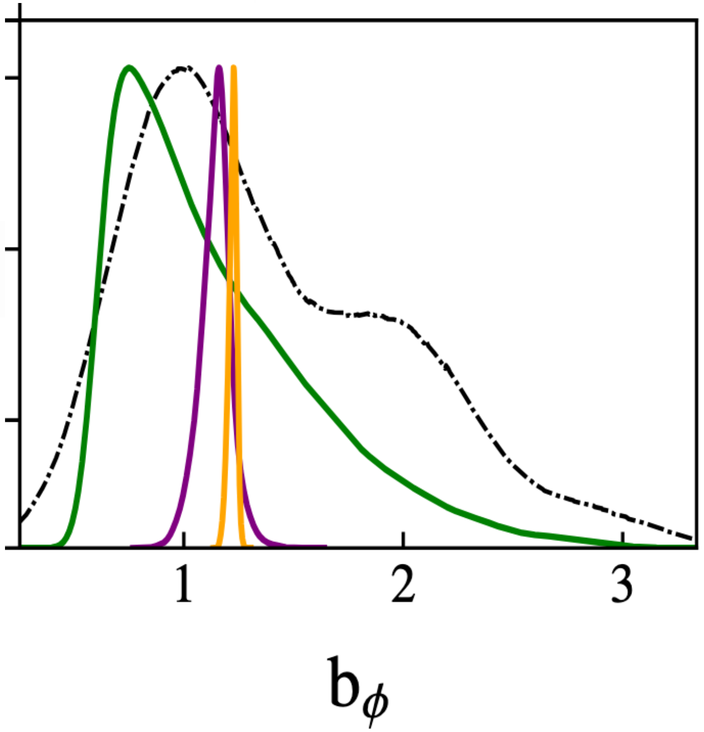


Preliminary results: this leads to significant narrowing in the b_ϕ prior, $\sigma_{b\phi}$ reduced by ~90%!

2026arXiv260421790M

DESI ELG-like selection in stellar mass and specific SFR





Led by Annie Moore, U of AZ PhD candidate

b_ϕ Priors Conditioned on Observed Data

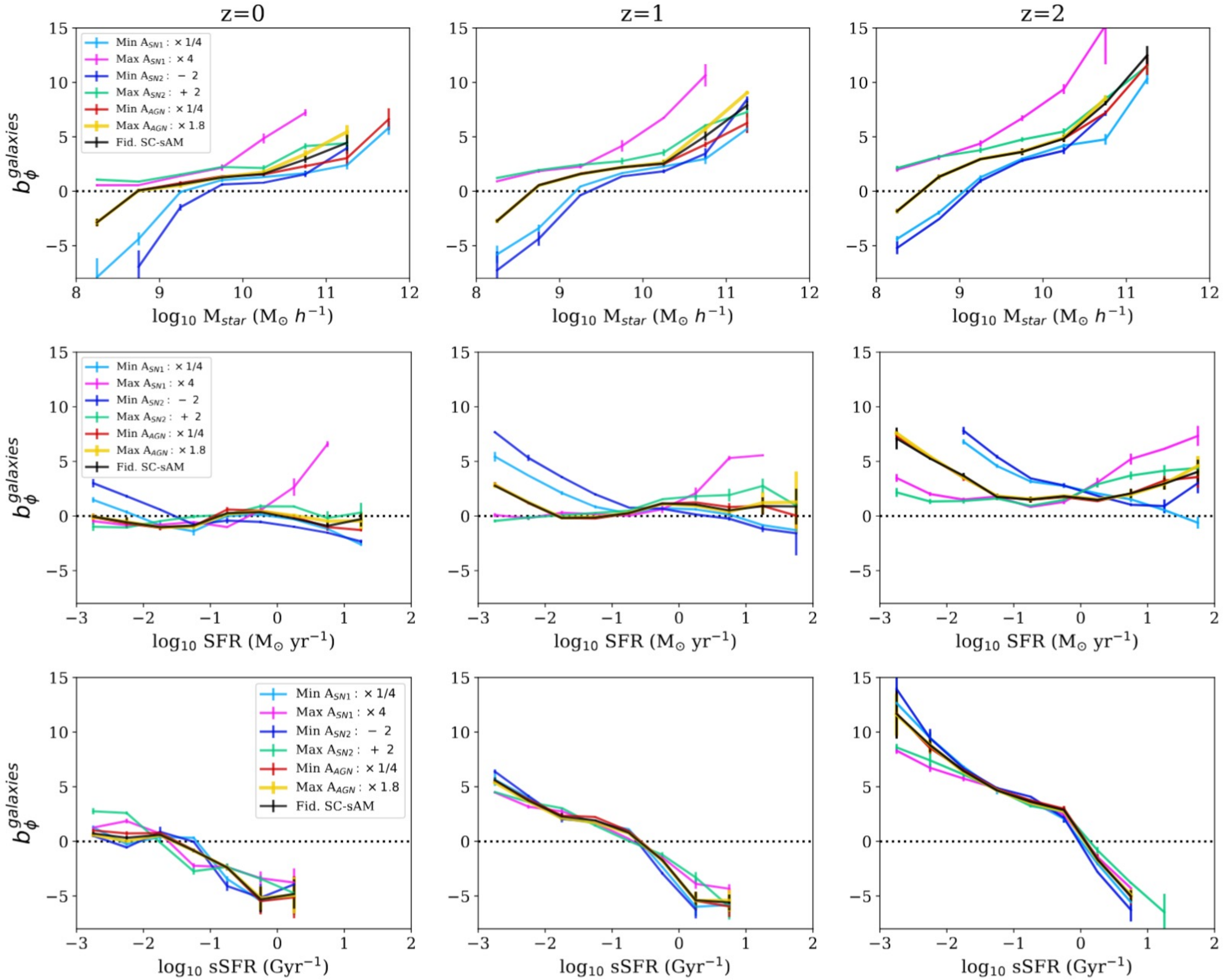
Metric	$\langle b_\phi \rangle_{SU}$	$\langle b_\phi \rangle_{\chi^2\{Obs\}}$			$p(b_\phi \| \{Obs\})$		
		SMF	SHMR	$M_* - Z_*$	SMF	SHMR†	$M_* - Z_*$
$\langle b_\phi \rangle$	1.43	1.20	1.41	1.39	1.15	1.25	1.15
σ_{b_ϕ}	0.69	0.48	0.58	0.62	0.08	0.02	0.45

Summary & Takeaways: arxiv.org/abs/2602.04987

FLATIRON INSTITUTE

- f_{NL} will not be well constrained until we better understand how galaxy formation affects galaxy bias parameters!
- With separate universe simulations and the Santa Cruz SAM, we measure non-Gaussian galaxy bias b_ϕ across several dozen parametrizations varying SN and AGN feedback
- All selections strongly diverge from the universality relation:
  - Our stellar mass findings aligns well with previous IllustrisTNG b_ϕ measurements, but diverge a bit in $b_\phi(b_1)$
  - SFR selections yield chaotic b_ϕ and $b_\phi(b_1)$
  - sSFR seems robust to galaxy formation parameterizations, likely due to regulated star formation in the SC-SAM
- We agree with the $b_\phi\,(b_1)$ measured in the Munich SAM—but we need more galaxy formation models!
- We can get much narrower priors for b_ϕ by conditioning on how realistic a particular SC-SAM parametrization is given some observable (Moore, LAP, & Krause in final prep)

FLATIRON INSTITUTE



Redshift evolution of b_ϕ? Effect of individual parameters?