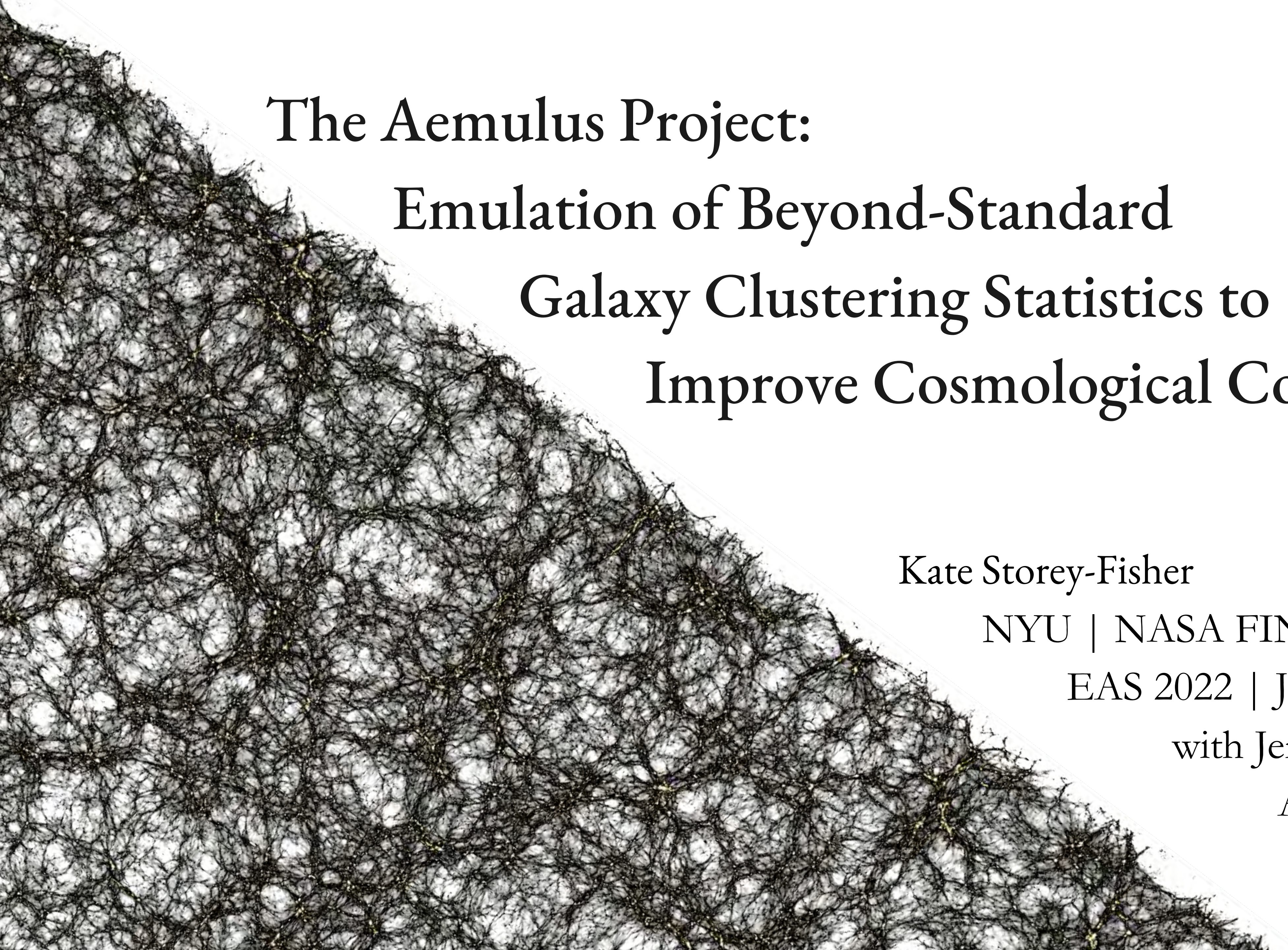


The Aemulus Project: Emulation of Beyond-Standard Galaxy Clustering Statistics to Improve Cosmological Constraints

Kate Storey-Fisher

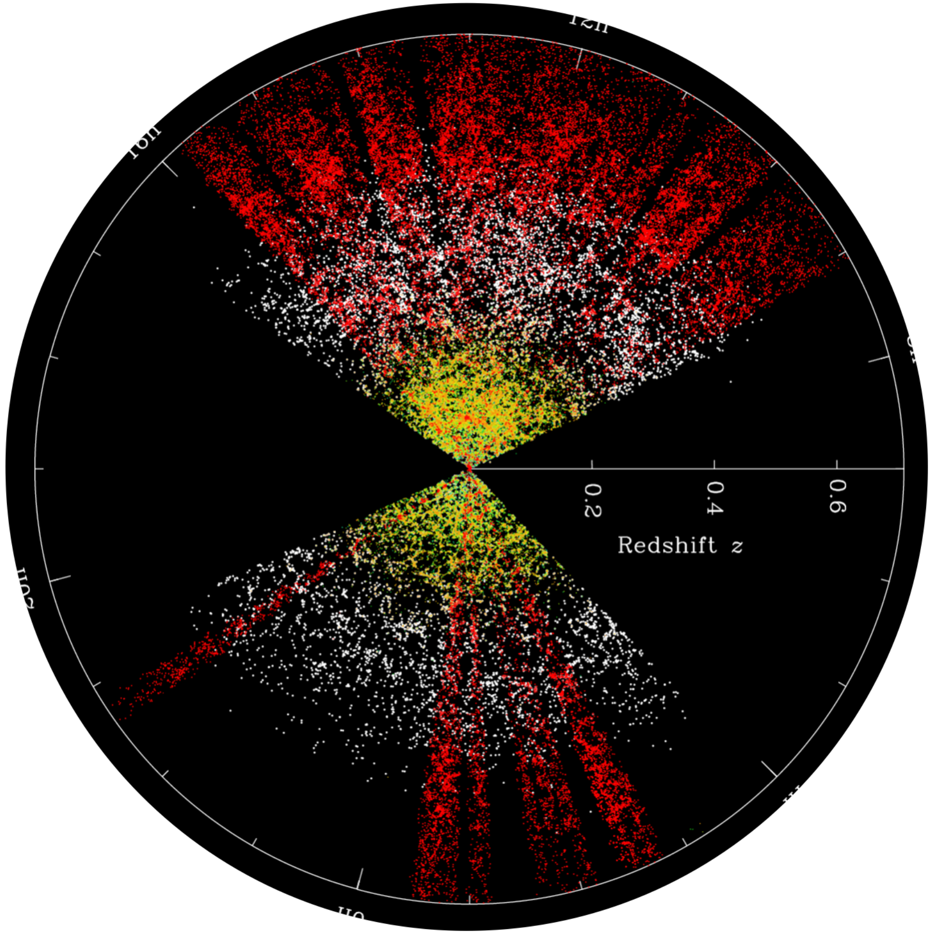
NYU | NASA FINESST

EAS 2022 | July 1, 2022

with Jeremy Tinker & the Aemulus Collaboration

Large-scale structure encodes the expansion history of the universe.

SDSS + BOSS Luminous Red Galaxies

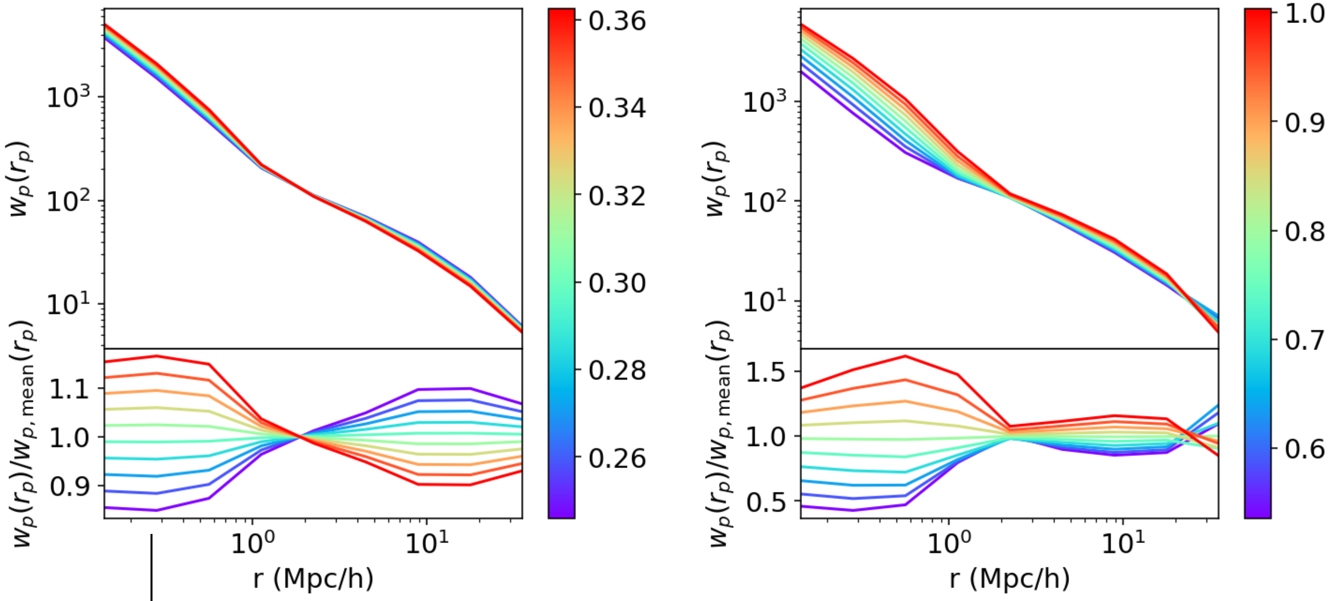


Michael Blanton / SDSS III

Summary statistics contain information about the cosmological growth of structure parameters.

Ω_m, matter density

σ_8, amplitude of fluctuations



The nonlinear regime is difficult & expensive to model!

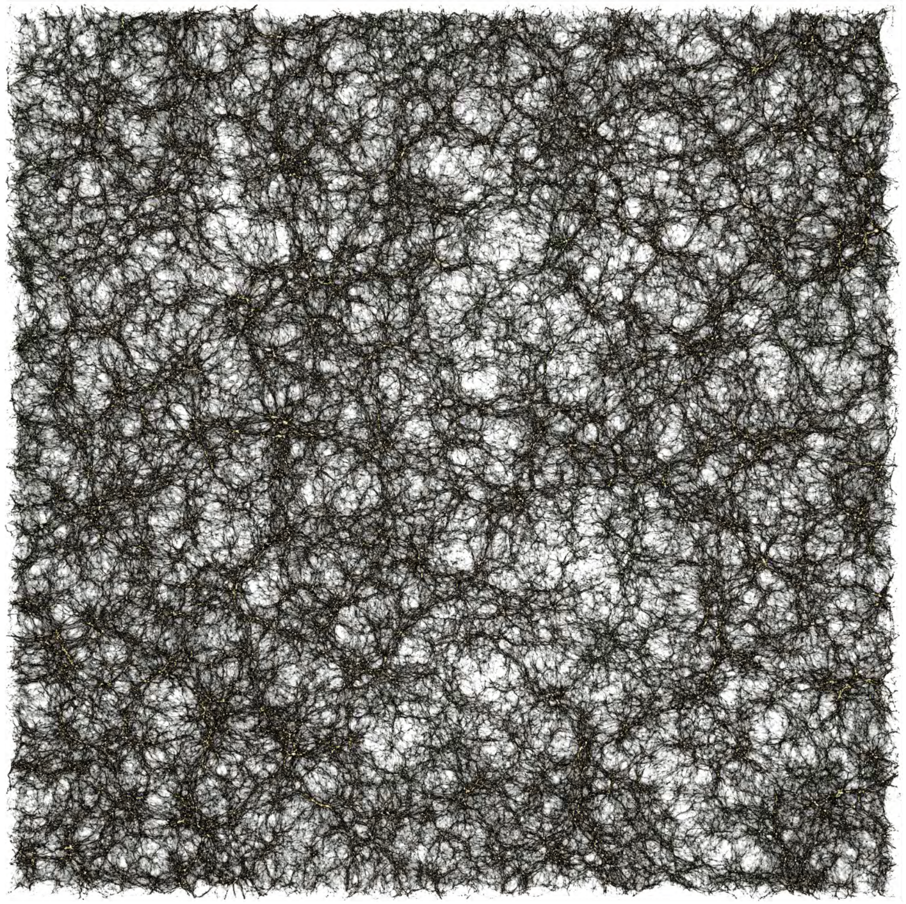
The Aemulus Project

Jeremy Tinker, Risa Wechsler, Joe DeRose, Zhongxu Zhai, Arka Banerjee, Tom McClintock, et al

Goal: Extract more cosmological information from small-scale galaxy clustering.

Problem: Cosmological simulations are expensive; can't compare observations to all possible models.

The Aemulus approach: Using a sparse suite of high-resolution N-body simulations, train an *emulator* on summary statistics to achieve %-level predictions, and apply emulator to infer parameters of observed data.



Slice of Aemulus N-body simulation

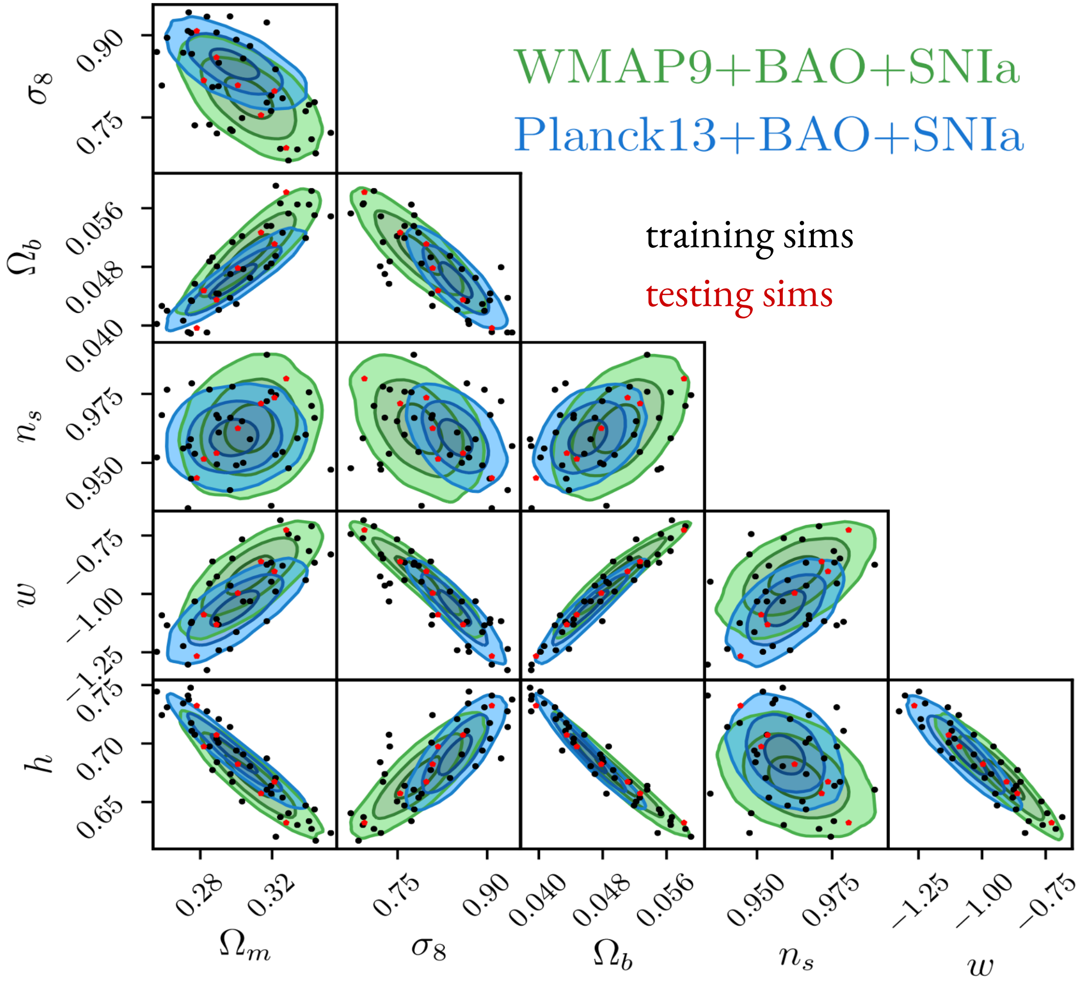
Aemulus I: DeRose+2018 (1804.05865)

Other Aemulus Papers:
II: McClintock+2018
III: Zhai+2018
IV: McClintock+2019
V: Zhai+2022

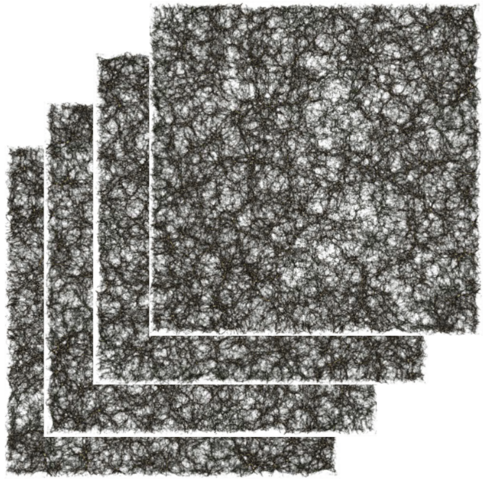
The Aemulus Simulations

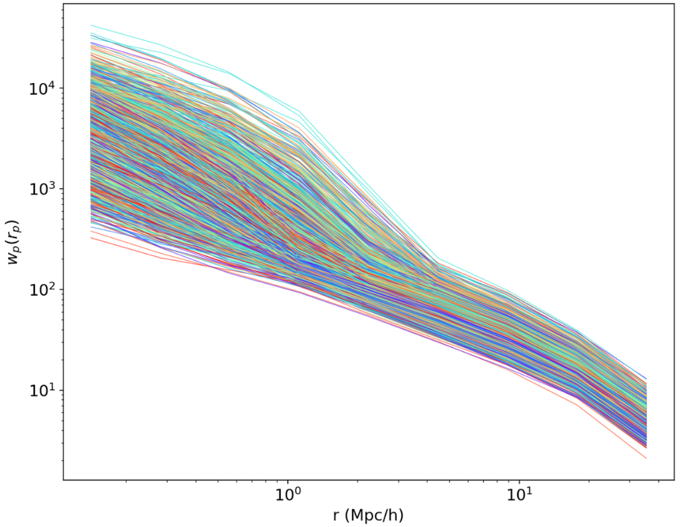
Set of simulations that span the expected parameter space of a **7D cosmological model** & an **11D HOD model** (halo occupation distribution), including **3 assembly bias parameters**

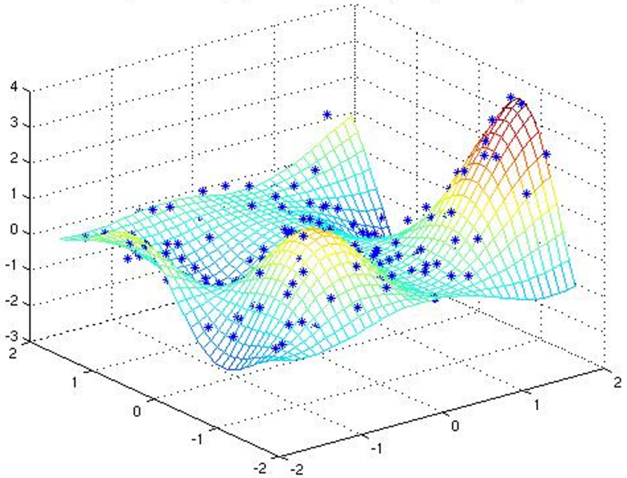
- Mass resolution = 3.5 x 10^{10} $h^{-1}M_{sun}$, L = 1.05 h^{-1}Gpc
- **Training set: 4000 mock catalogs** (40 cosmologies, 100 unique HODs each)
- **Test set: 700 mock catalogs** (7 cosmologies, 5 realizations each, 100 HODs each)

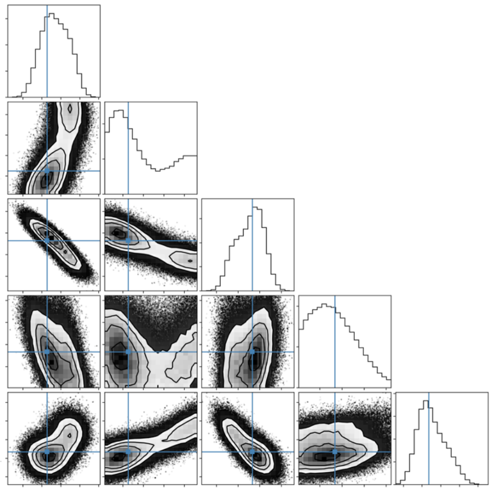


The Emulation Approach







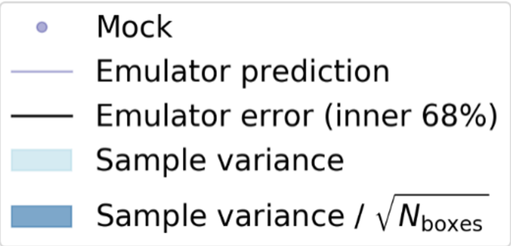


Generate **simulations** that span a prior space of cosmological and galaxy bias parameters.

Compute summary statistics on these to build up a **training set** for our model.

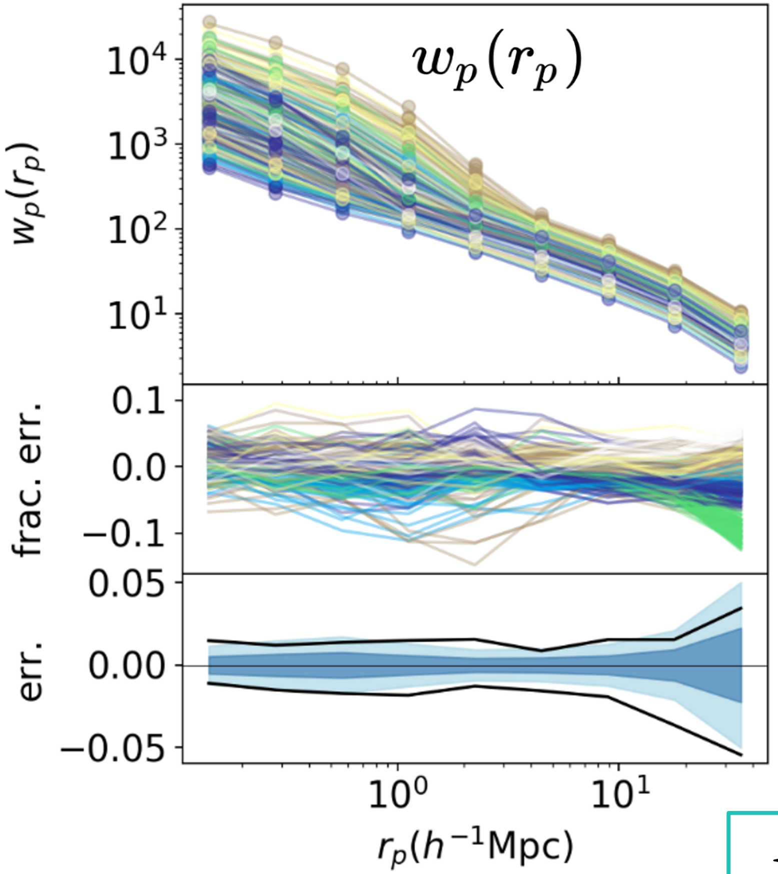
Train a **Gaussian process** (GP) to learn the correlation between the parameters and statistics.

Use **Markov Chain Monte Carlo** with the trained GP to recover test simulation parameters.

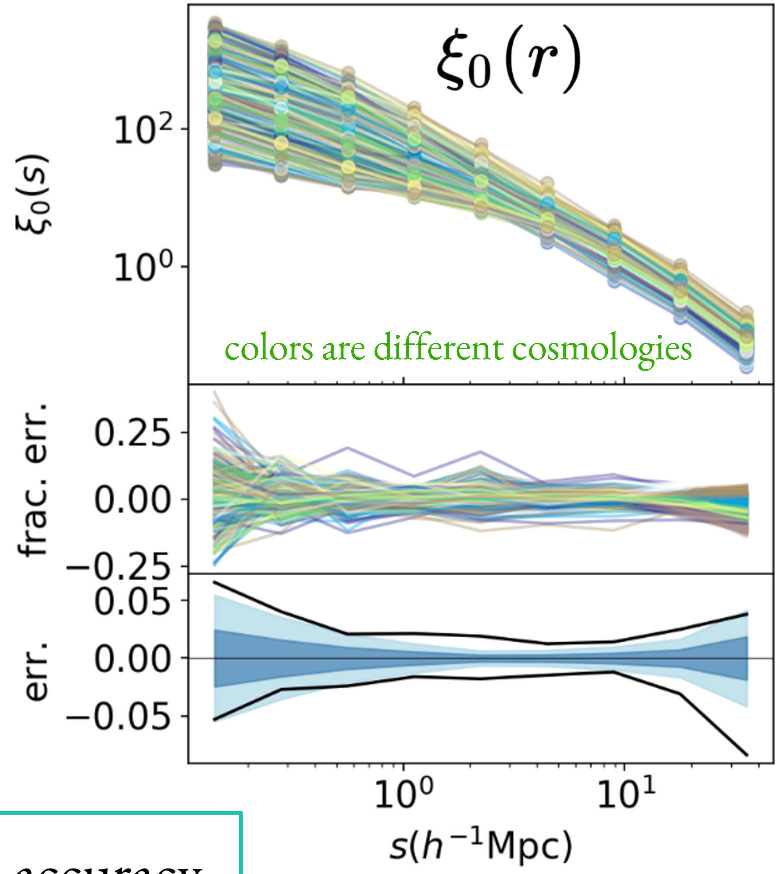


Emulation of Standard Clustering Statistics

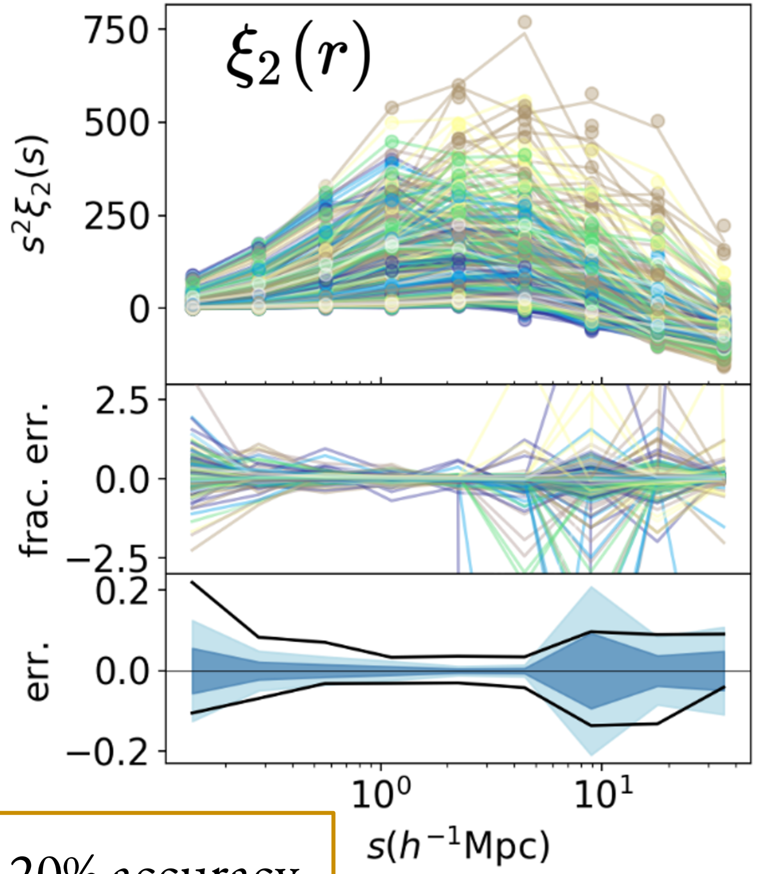
Projected correlation function:



Monopole of the correlation function:



Quadrupole of the correlation function:

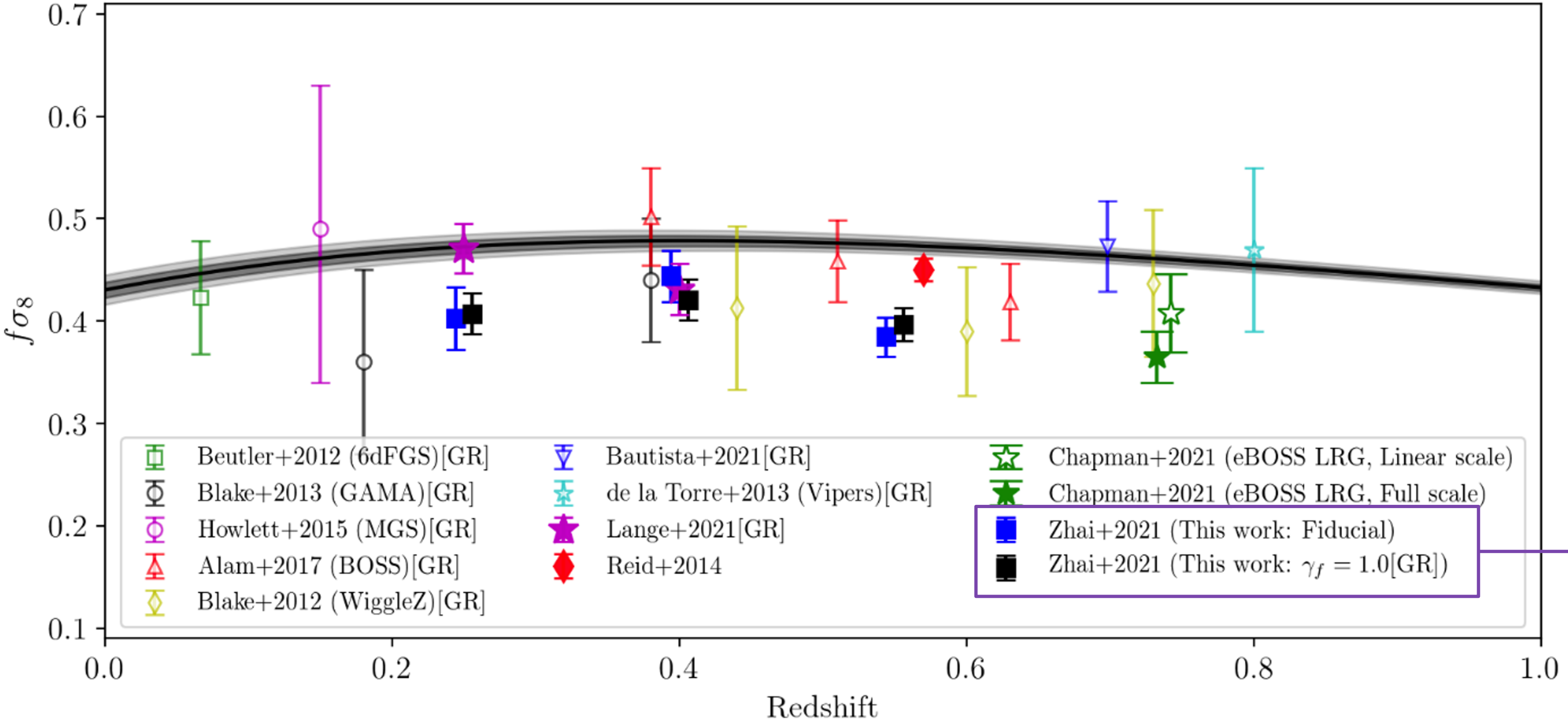


1-5% accuracy

5-20% accuracy

KSF+ (in prep.)

Measurement of the growth of structure parameter $f\sigma_8$



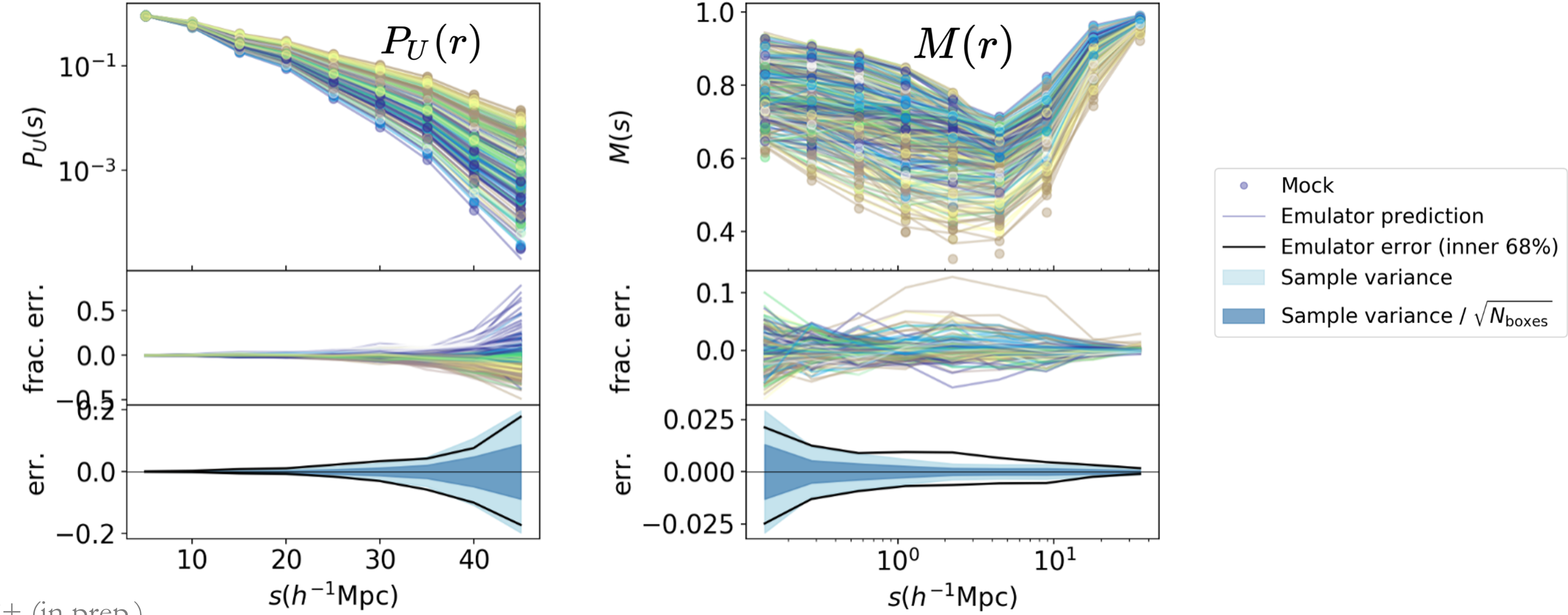
Measurements with the previous Aemulus emulators using these standard statistics

Zhai+2022 (incl. **KSF)**
([2203.08999](2203.08999))

Beyond Standard Clustering Statistics

Underdensity probability function:
% of randomly placed spheres of radius r with number density < threshold

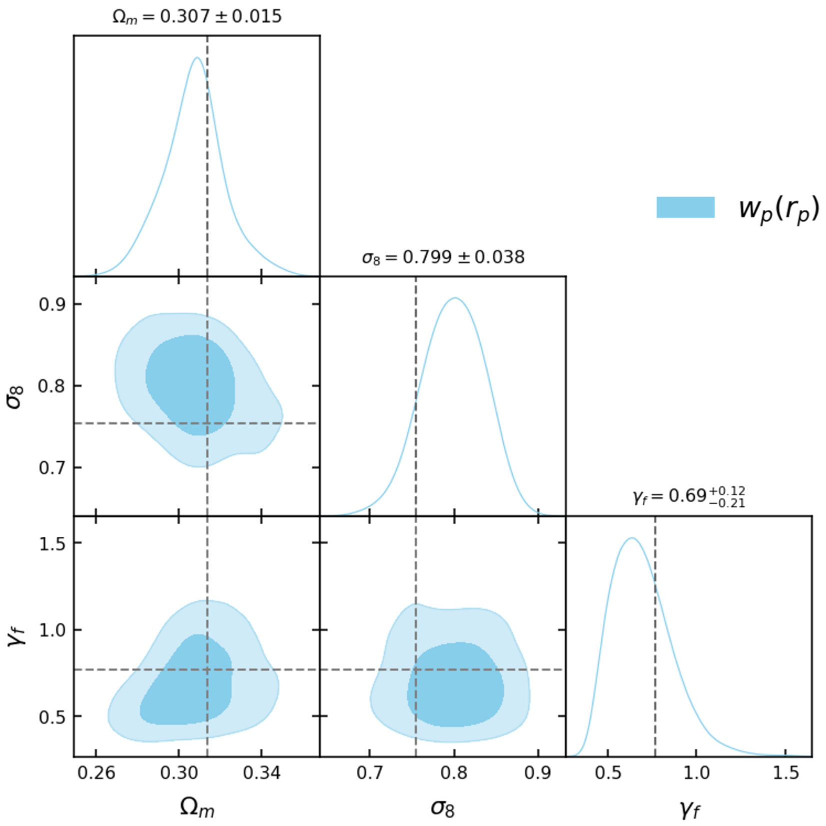
Marked correlation function:
the monopole with galaxy pairs weighted by the local density



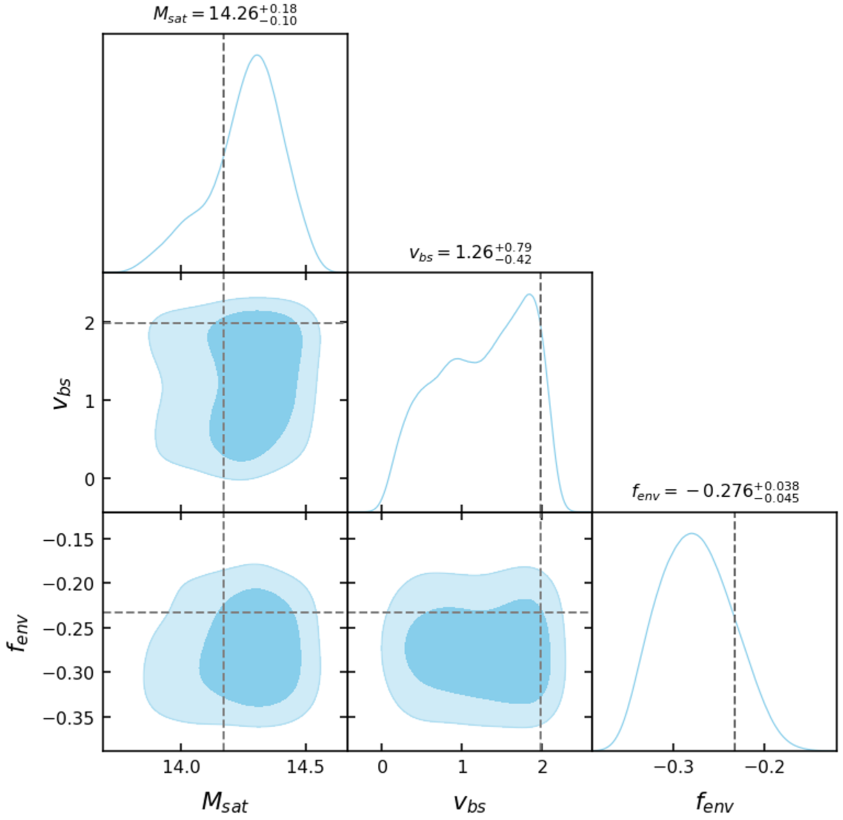
KSF+ (in prep.)

Including beyond-standard statistics improves precision & accuracy.

Cosmological parameters



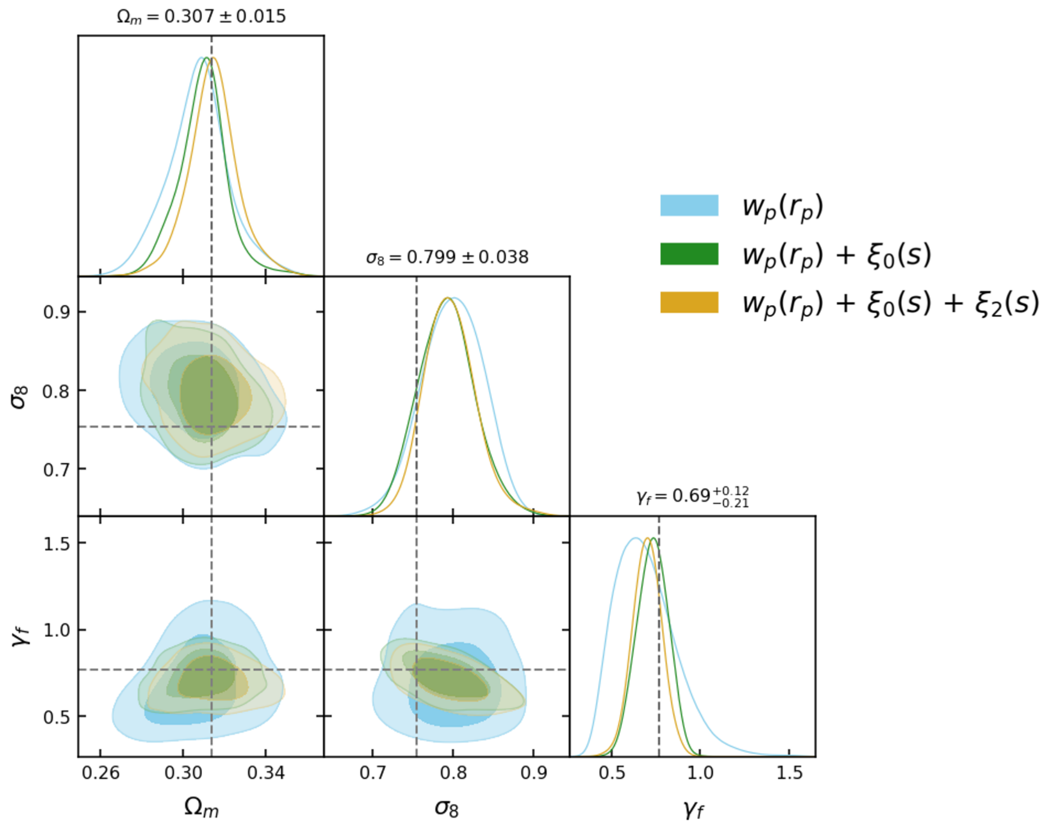
Galaxy bias parameters



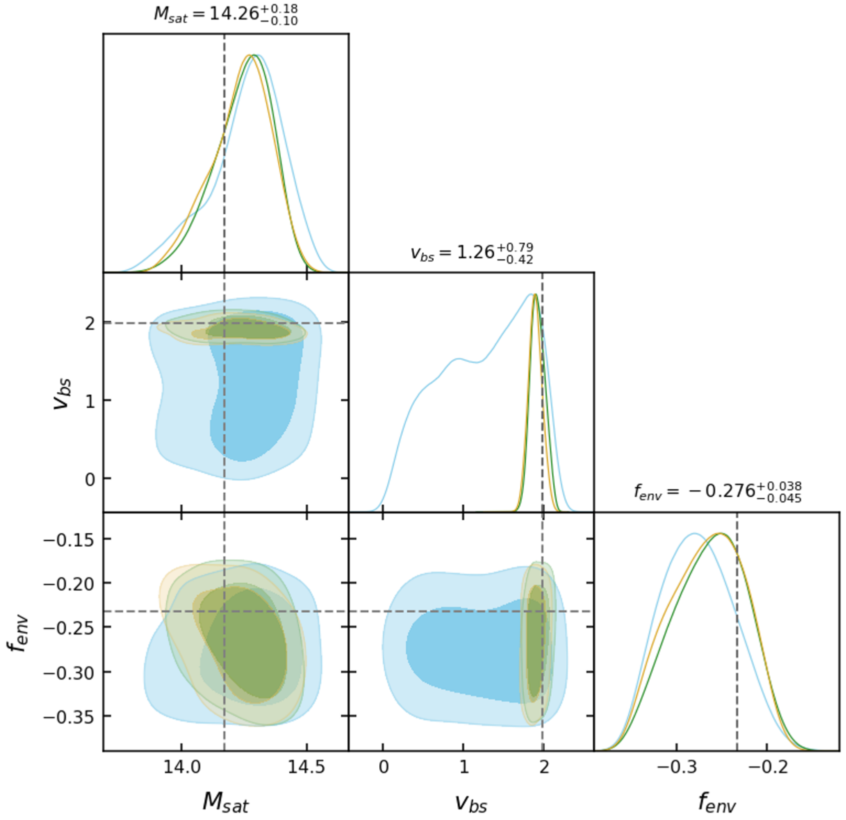
KSF+ (in prep.)

Including beyond-standard statistics improves precision & accuracy.

Cosmological parameters



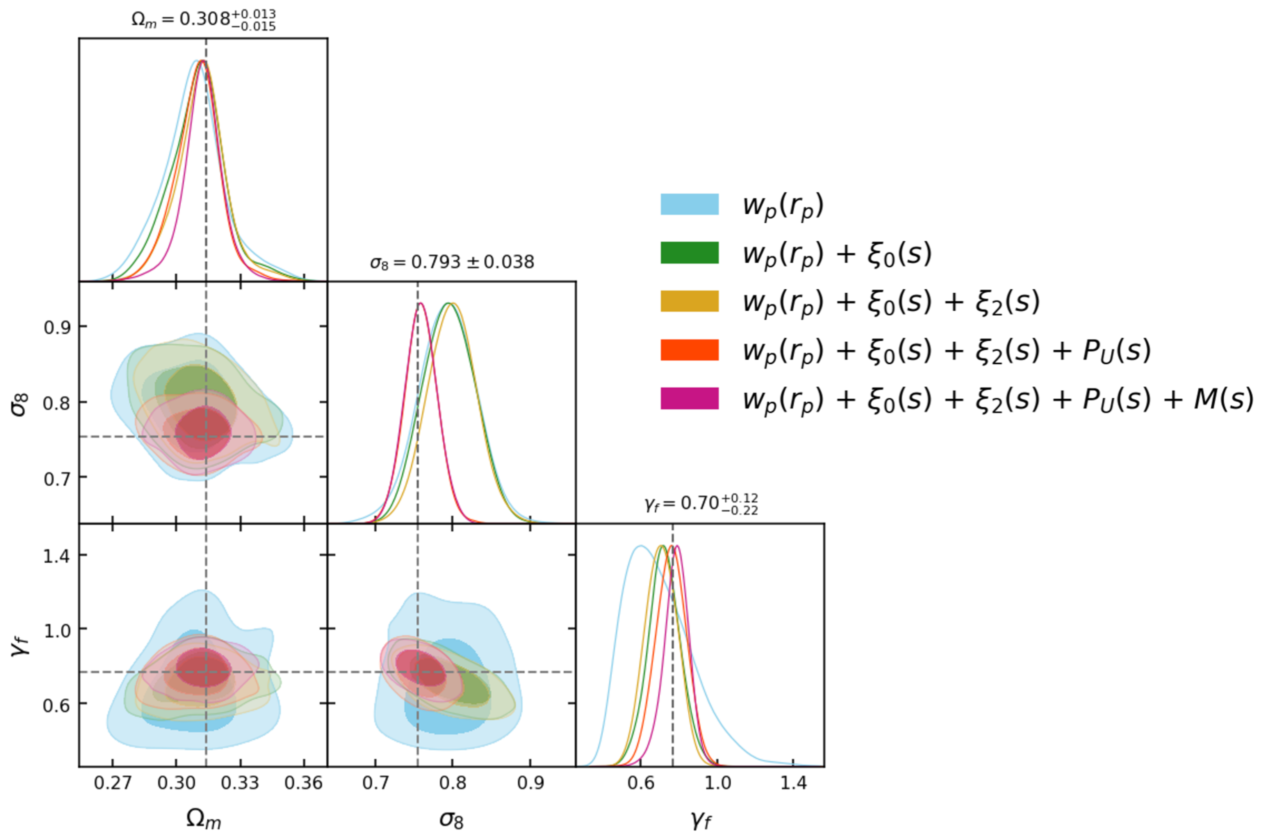
Galaxy bias parameters



KSF+ (in prep.)

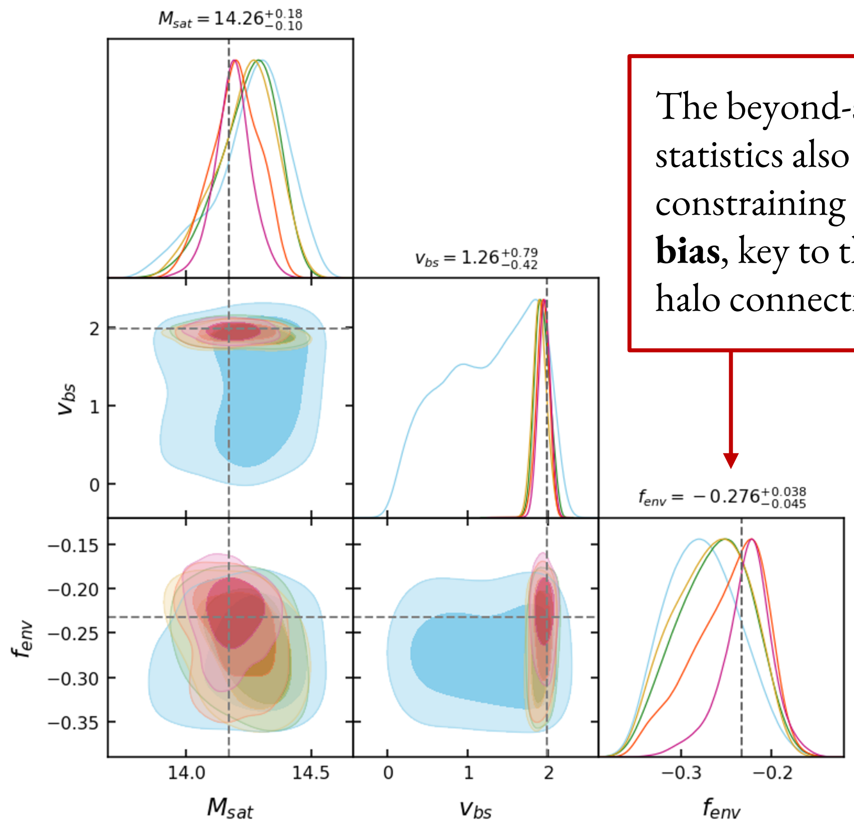
Including beyond-standard statistics improves precision & accuracy.

Cosmological parameters



Significant improvement on
σ_8 & γ_f from $P_U(r)$ and $M(r)$

Galaxy bias parameters

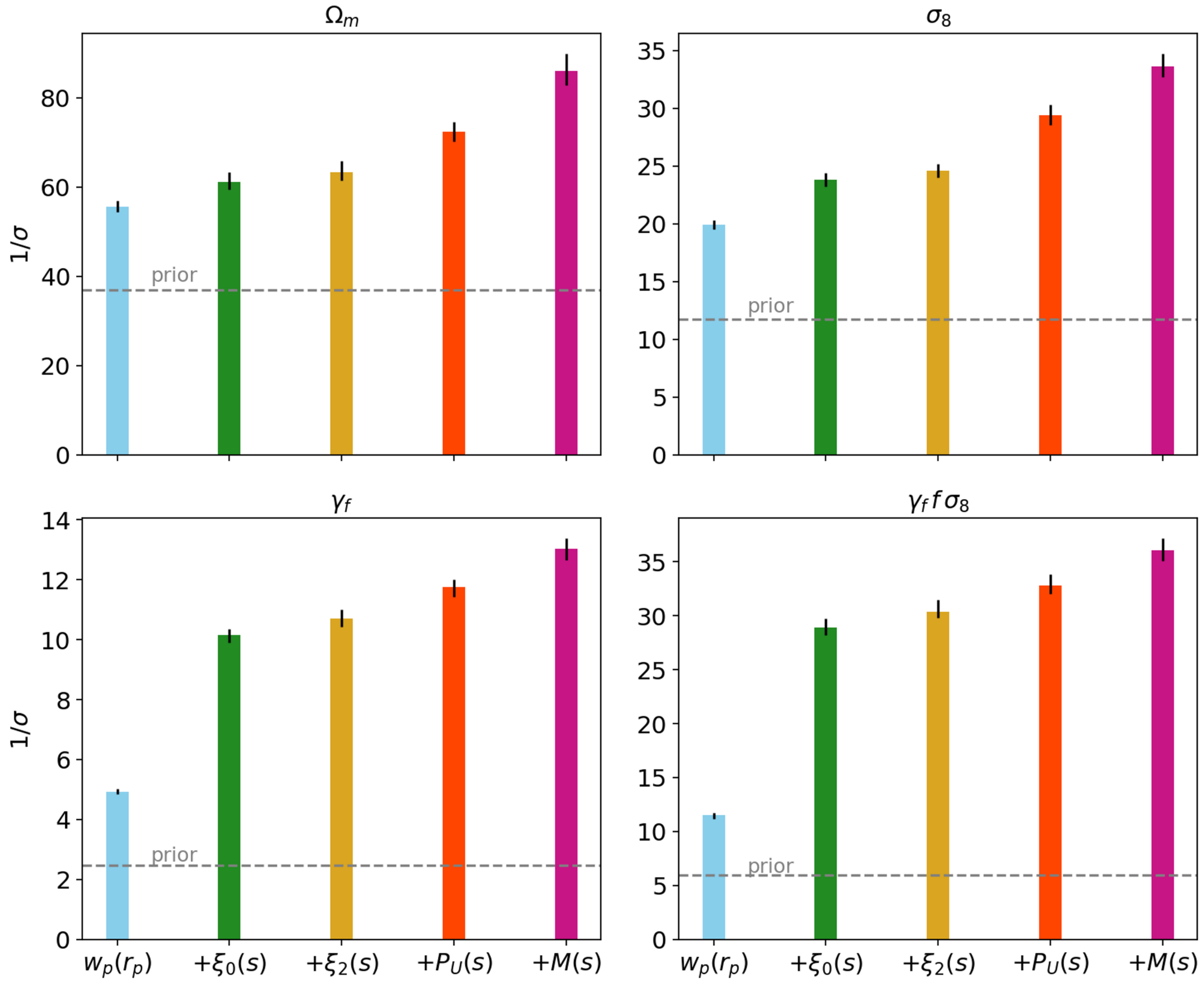


The beyond-standard statistics also aid in constraining **assembly bias**, key to the galaxy-halo connection.

Significant improvement on
M_{sat} from $P_U(r)$ and f_{env} from $M(r)$

KSF+ (in prep.)

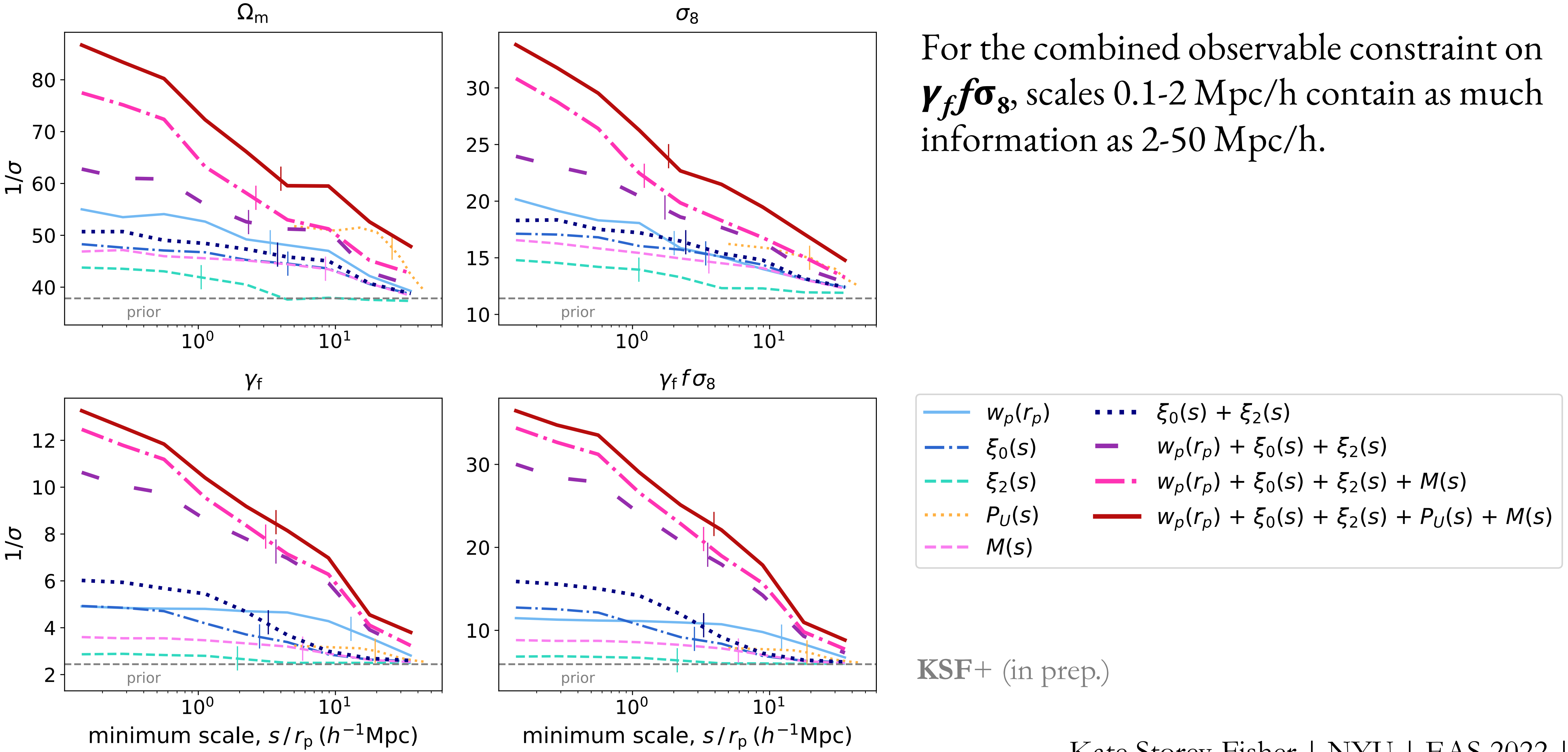
Including beyond-standard statistics improves precision & accuracy.



With the addition of the underdensity probability function and marked correlation function, we increase the constraining power on **σ_8 by 33%, Ω_m by 25%**, and **$\gamma_f f \sigma_8$ by 15%**.

KSF+ (in prep.)

Small scales add significant information to cosmological constraints.



For the combined observable constraint on $\boldsymbol{\gamma_f f \sigma_8}$, scales 0.1-2 Mpc/h contain as much information as 2-50 Mpc/h.

KSF+ (in prep.)

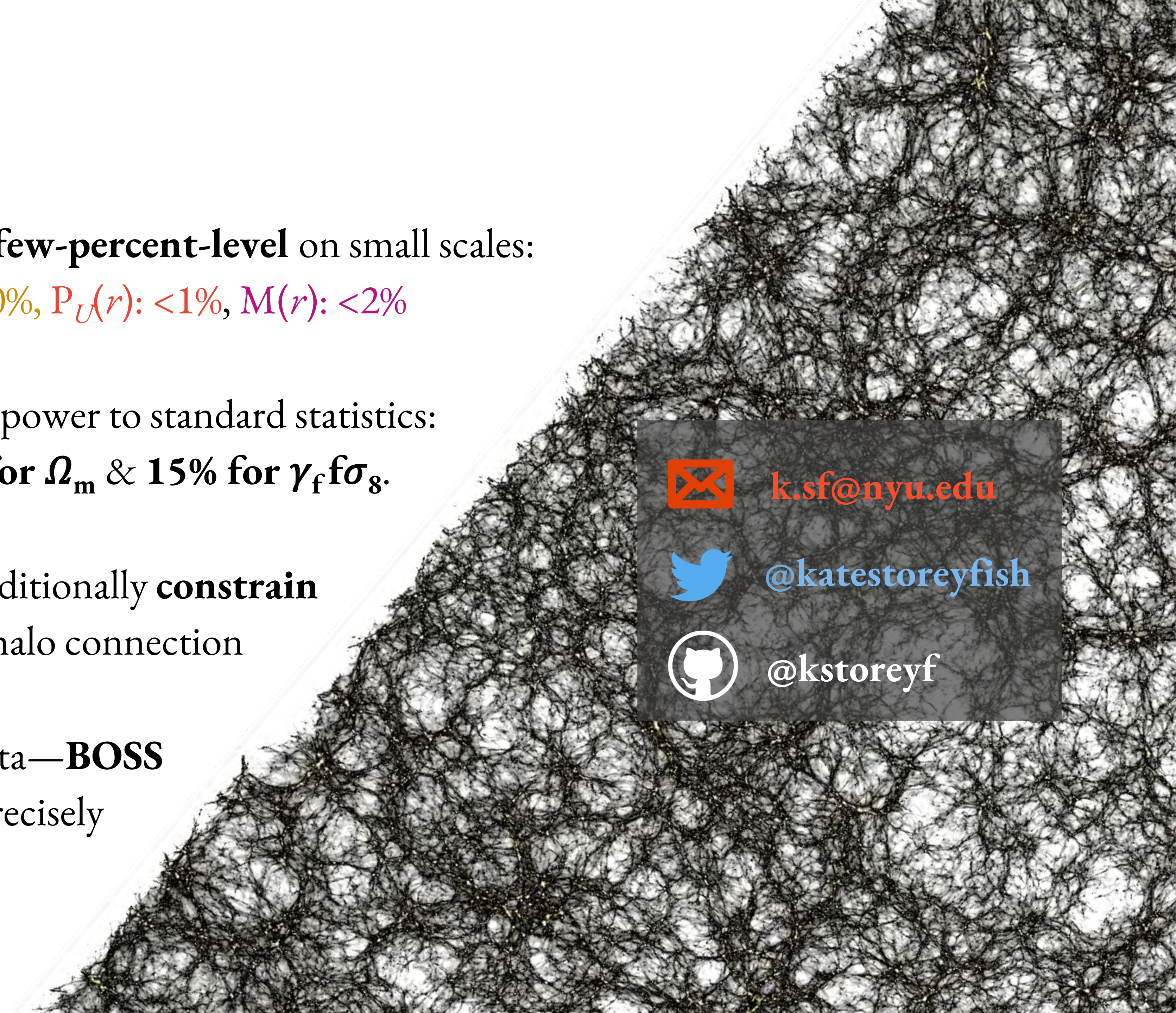
Summary

- Our emulators can predict at the **few-percent-level** on small scales: $w_p(r_p)$: ~2%, $\xi_0(r)$: <5%, $\xi_2(r)$: ~10%, $P_{\mathcal{U}}(r)$: <1%, $M(r)$: <2%

- $P_{\mathcal{U}}(r)$ and $M(r)$: add constraining power to standard statistics: precision ↑ by **33% for σ_8, 25% for Ω_m** & **15% for $\gamma_f f\sigma_8$**.

- The beyond-standard statistics additionally **constrain assembly bias**, key to the galaxy-halo connection

- Next: apply these emulators to data—**BOSS** LRG sample / **DESI**—to more precisely measure cosmological parameters



k.sf@nyu.edu
@katestoreyfish
@kstoreyf

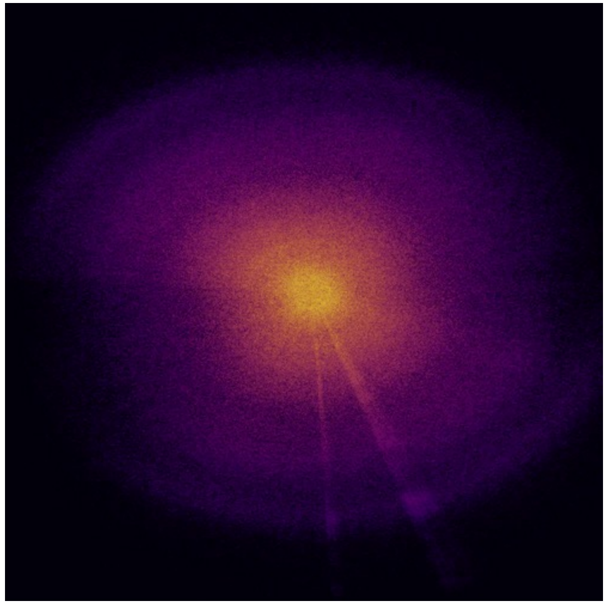
Also talk to me about…

large-scale structure!

galaxies!

statistics & machine learning!

- Symmetry-preserving methods for connecting galaxies to halos (EAS S11 Talk)
- A generalized correlation function estimator & measuring inhomogeneity (2105.02434)
- Anomaly detection in galaxy images with generative models (2105.02434)
- The Gaia quasar sample for large-scale structure cosmology



- Postdoc opportunities :)

k.sf@nyu.edu

@katestoreyfish

@kstoreyf