The multivariate halo bias in the eye of a machine

Jiaxin Han
Shanghai JiaoTong University
& Kavli IPMU

Yin Li, Yipeng Jing (SJTU), Takahiro Nishimichi, Wenting Wang and Chunyan Jiang (SHAO)

Arxiv: 1802.0917,
MNRAS

Halo Bias

- Bias: unfair selection

  - halo field is a biased sampling of the density field

$$\delta_h = b\delta_m \longrightarrow \rho_h = k\rho_m + (1-b)\bar{\rho}_h$$

  - depend on all variables that can modify the sampling: multivariate

$$b(M, c, e, j, t_f, ...)$$

  - Previous works: marginalized dependences
    - Univariate: well studied $b(M)$

$$b(M) = \int b(M, c, e, j, t_f, , ...) dP(c, e, j, t_f, ...)$$

    - Bivariate: detected “assembly bias”

$$b(M, c),\ b(M, j),\ b(M, t_f)...$$

Efficient bias measurement

- *Ensemble* estimator of bias

$$\delta_h = b\delta_m \;\Rightarrow\; \xi_{hm} = b\xi_{mm}$$

- Correlation function: average density profile

$$\xi_{hm} =< \delta_h \delta_m >$$
$$=< \delta_m | h >$$

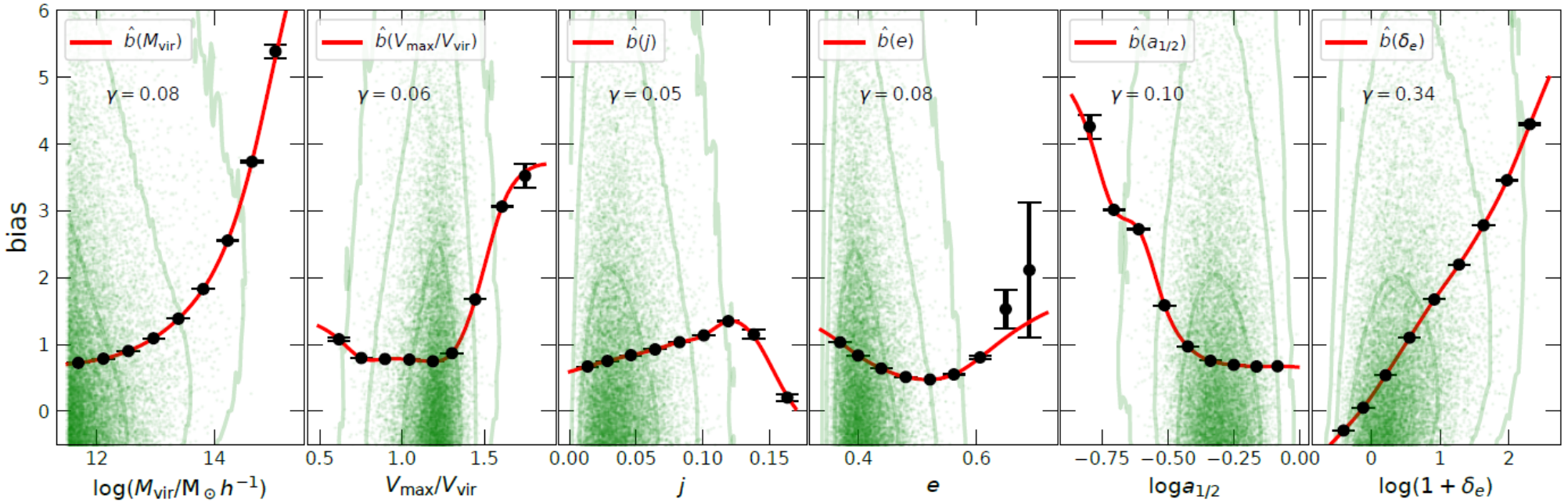
- individual bias estimator

$$\beta(r) = \delta_m(r)/\xi_{mm}(r)$$

$$\Rightarrow b =< \beta >$$

Single-variable dependence

$$b =< \beta >$$

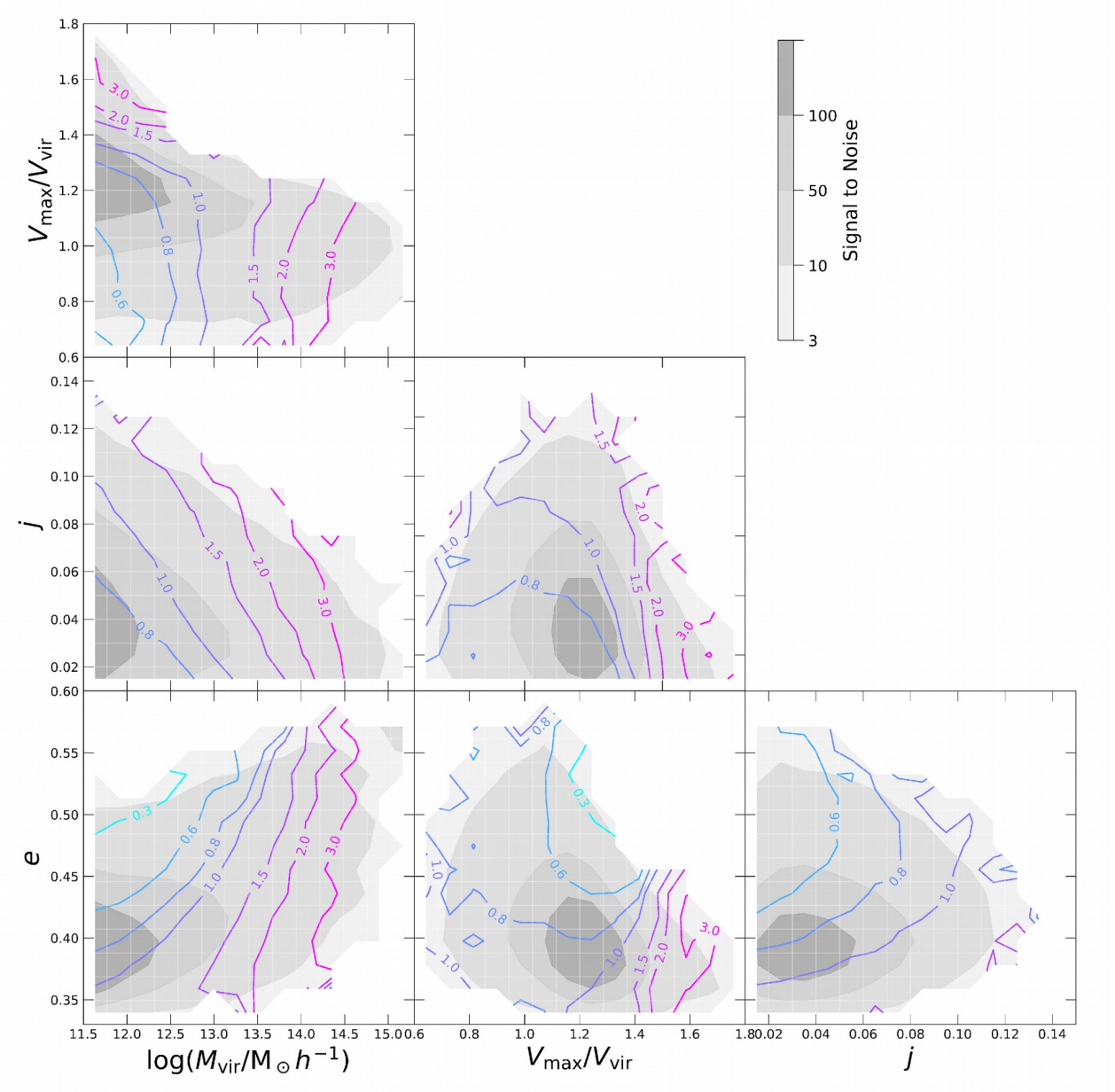


Mass | Concentration | Spin | Shape | Formation Time | Environment

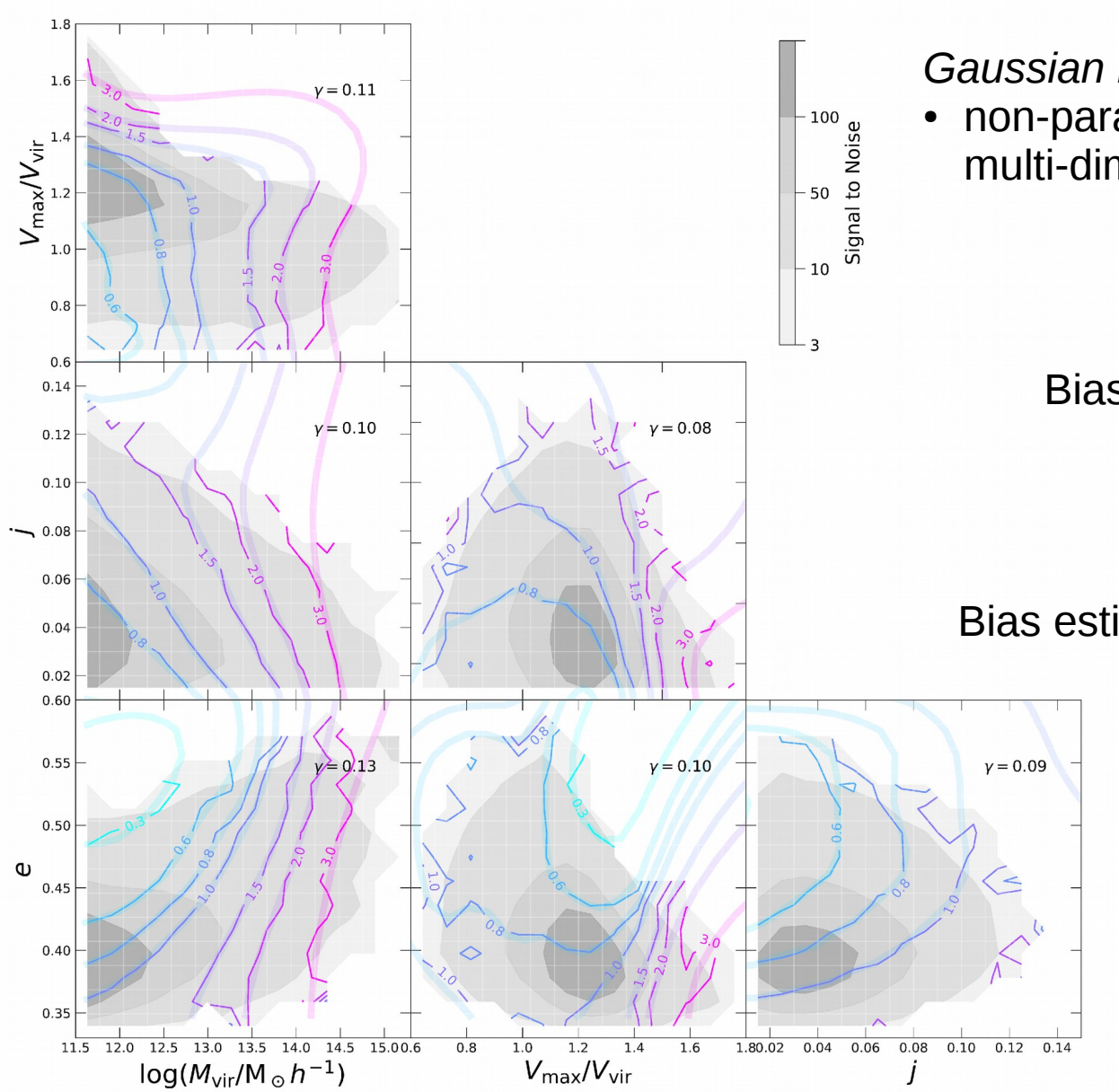
$$M_{\rm vir} > 10^{11.5} M_\odot/h \ (500 \text{ particles})$$

Data: LCDM simulation, 3072^3 particles, 600Mpc/h box

Bivariate dependence



Bivariate dependence



Gaussian Process Regression

- non-parametric fitting of multi-dimensional data

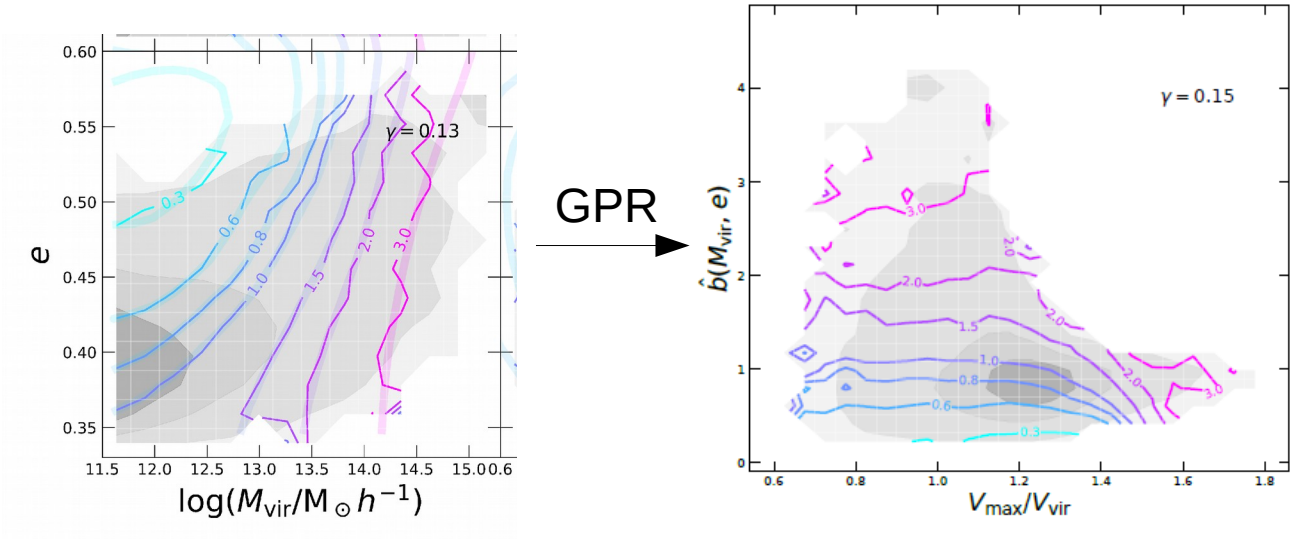
Bias maps

Gaussian Process Regression

Bias estimators $\hat{b}(x, y)$

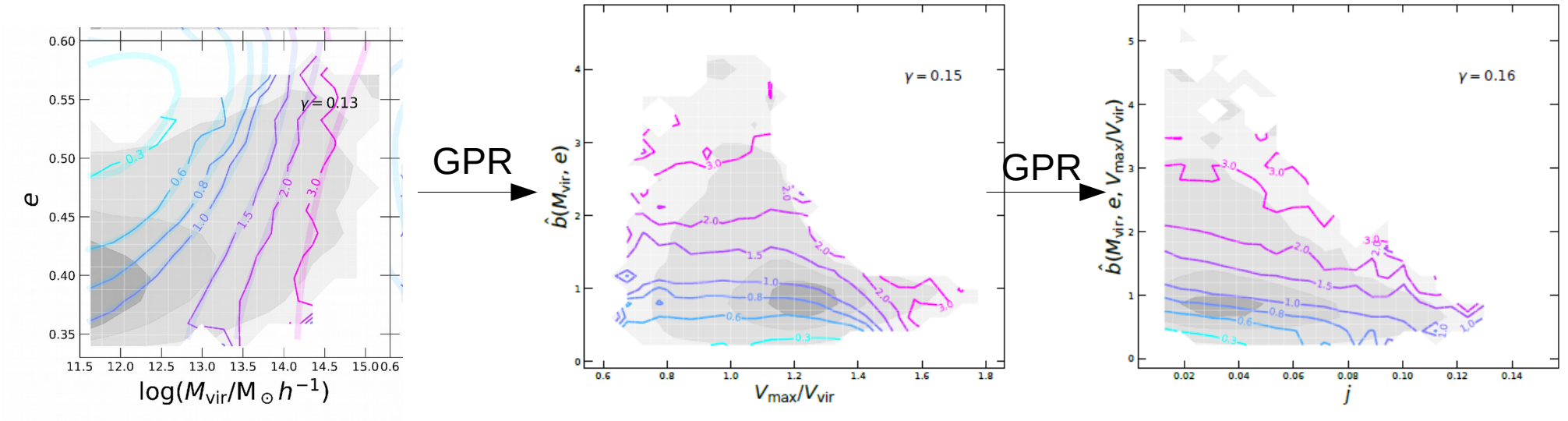
Result

Residual dependence and multivariate estimators



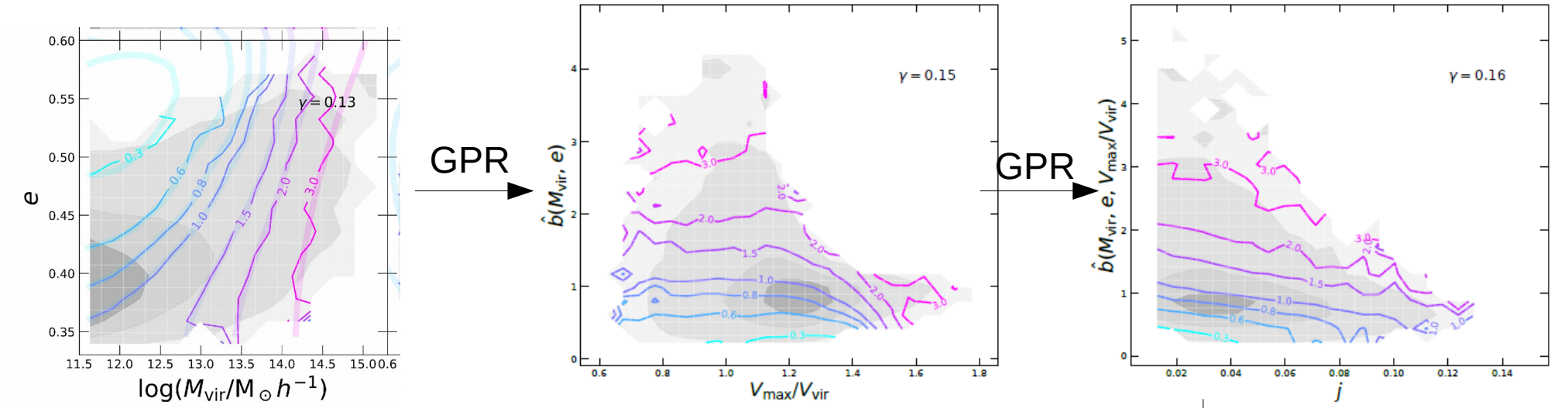
Result

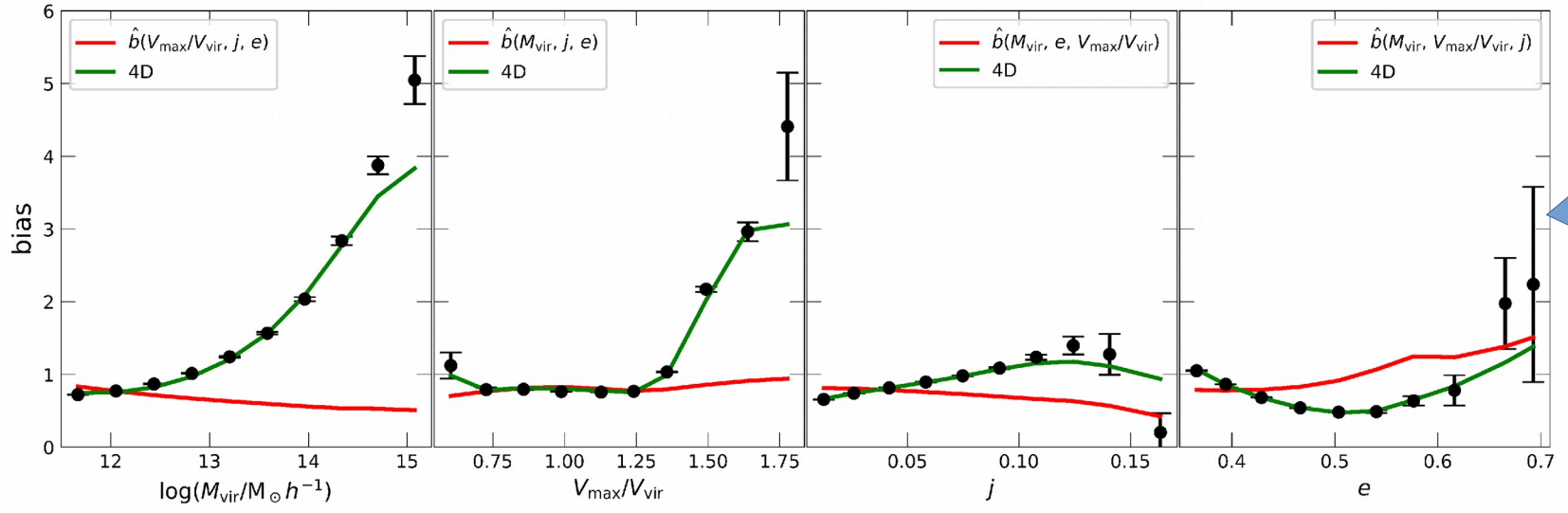
Residual dependence and multivariate estimators



Result

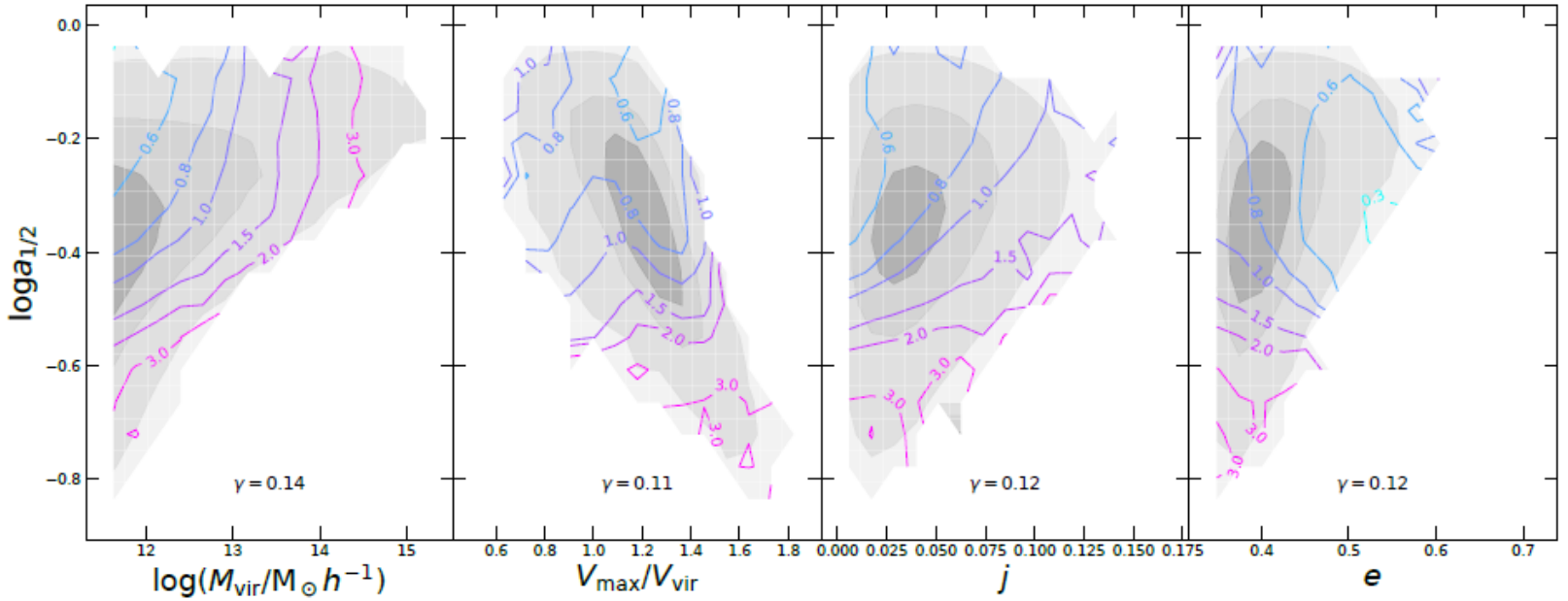
Residual dependence and multivariate estimators

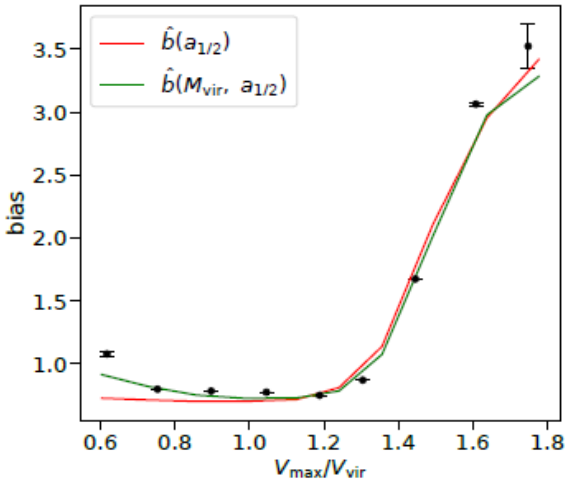


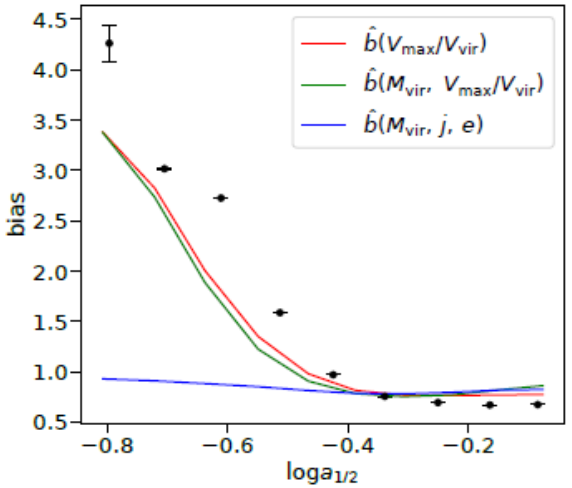


Non-redundant Variables!

Assembly History Dependence



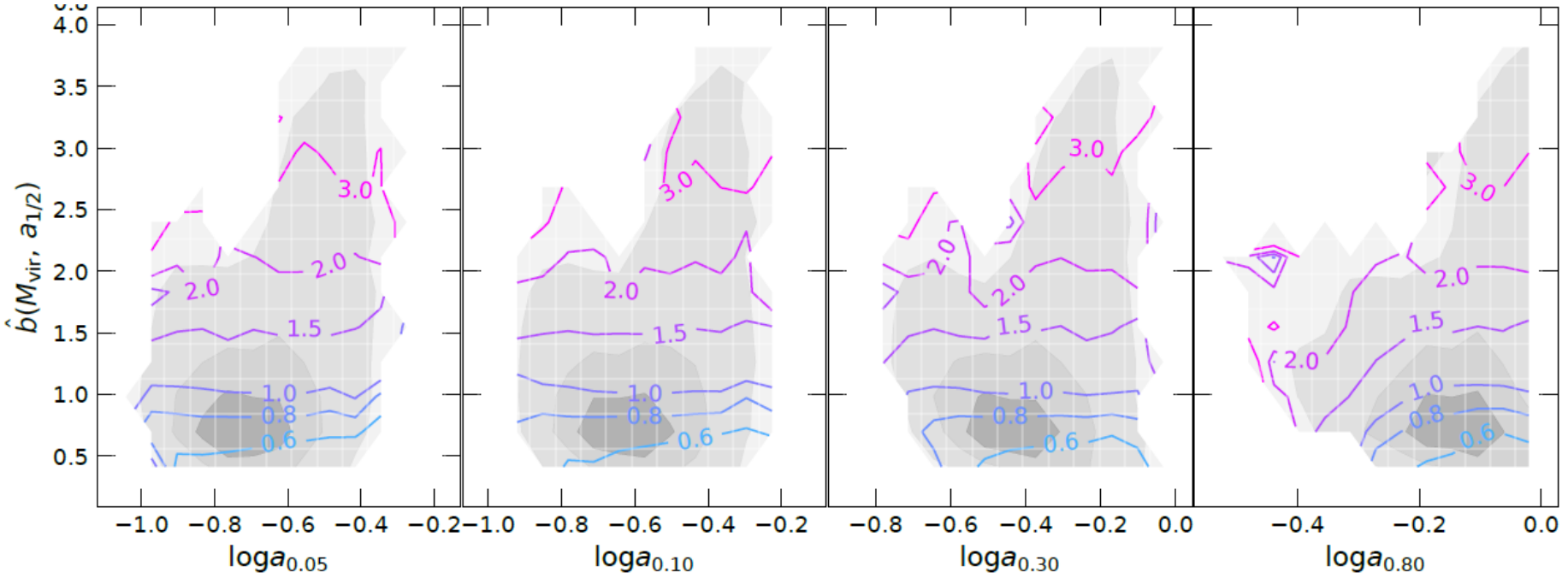




- Primary formation time dependence for early-forming haloes
- $V_{\rm max}/V_{\rm vir}$is a lossy proxy of formation time (and mass)!
- Spin and shape biases are not

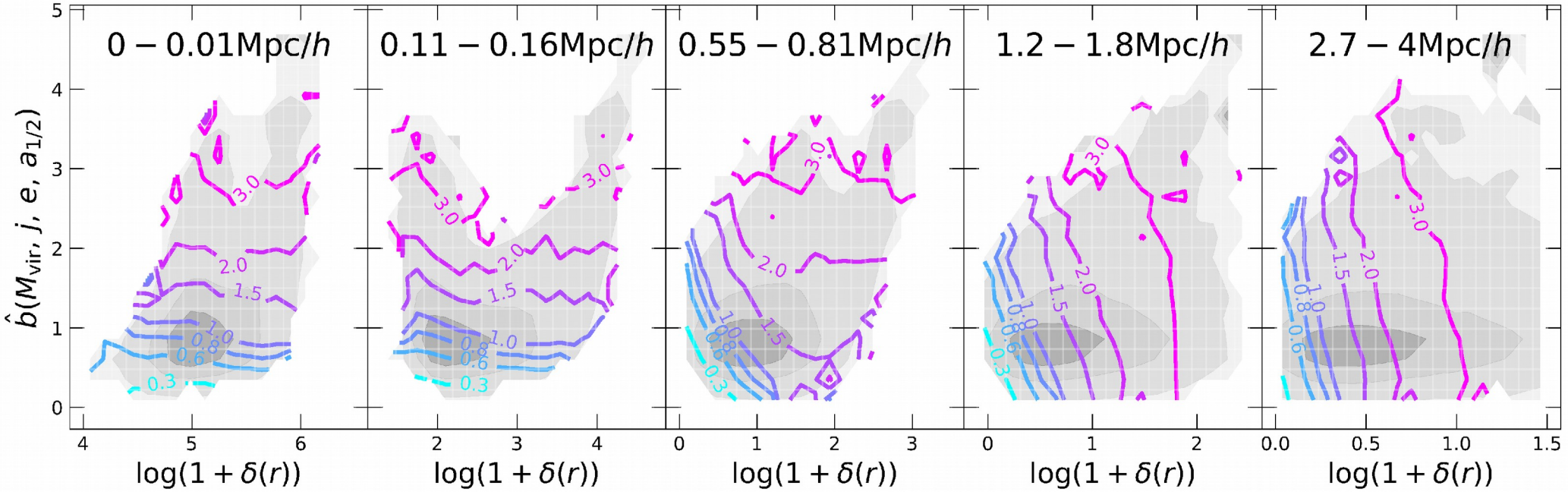
MAH dependence

Early-forming haloes: extra dependence on recent MAH



a_f : the scale factor when the halo is a fraction f of its final mass

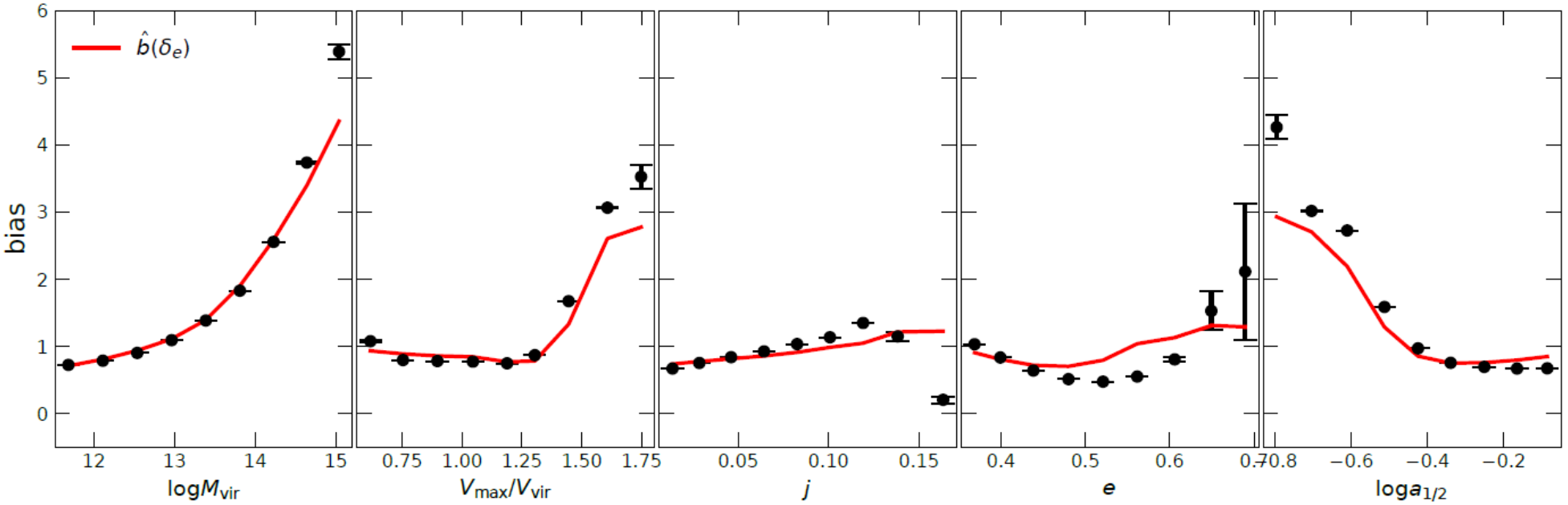
The density profile dependence

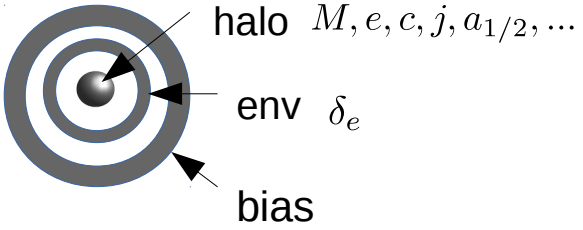


Density Profile

Environment

One environment to rule them all

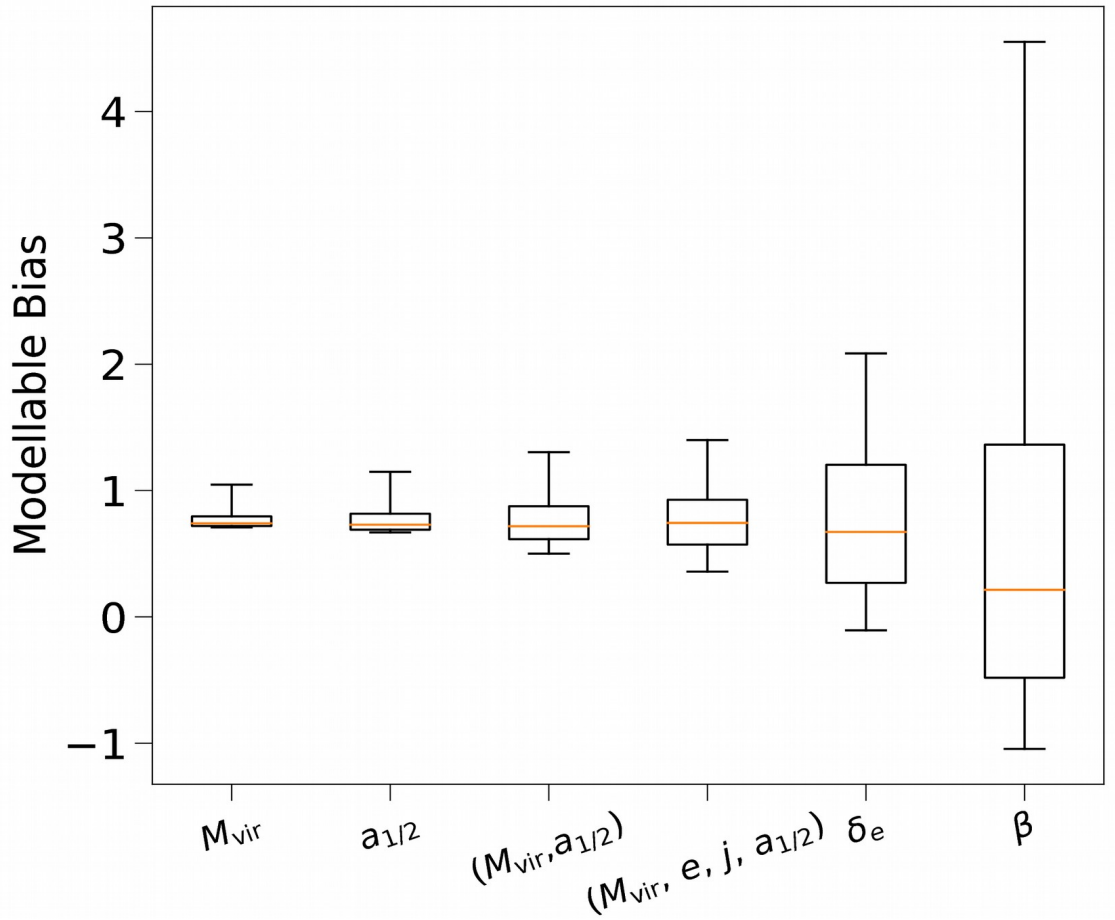




Conclusions

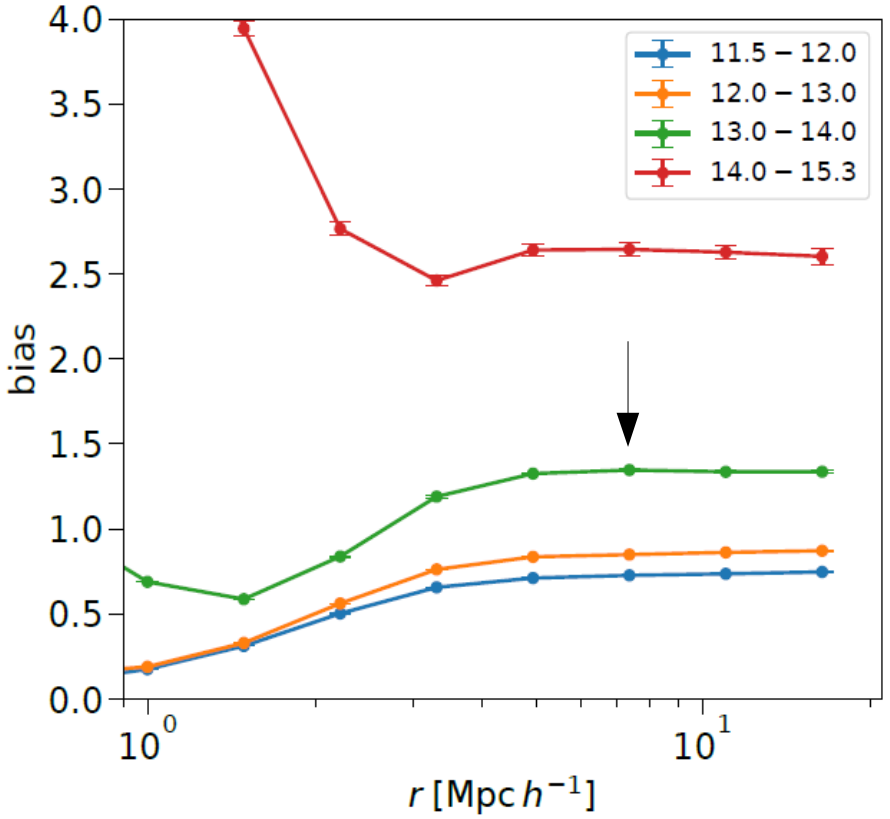
- Mass, concentration (formation time), spin, shape are non-redundant bias variables
  - Bias depends mostly on mass at high mass,
  - mostly on (recent) formation history for early-forming halos
  - Lowmass lateforming halos care more about spin and shape.
- $V_{\rm max}/V_{\rm vir}$ is a lossy proxy of formation time, while spin and shape are different
- The environment around halos captures most of the bias dependence on halo structure
  - Defined at as small as 1~2Mpc/h scale

Summary



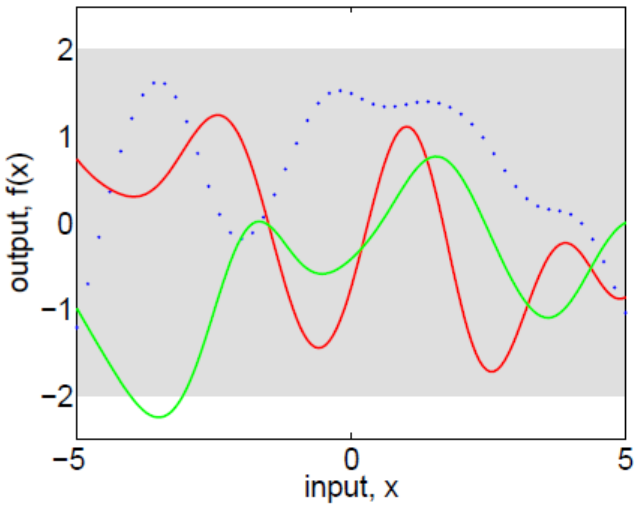
arxiv: 1802.0917

Bias scale

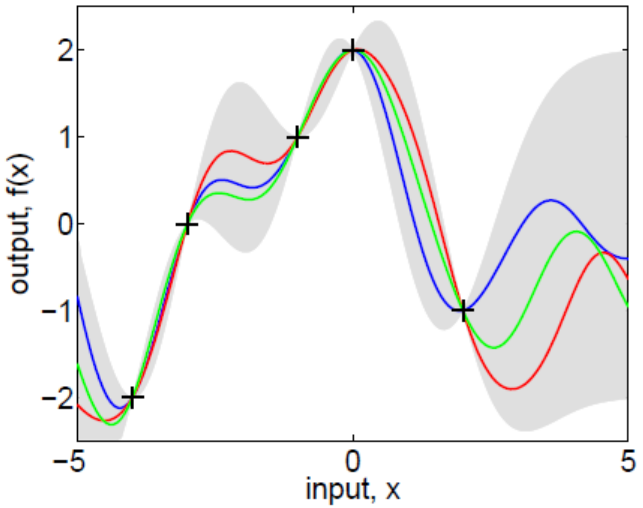

4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
bias
11.5 – 12.0
12.0 – 13.0
13.0 – 14.0
14.0 – 15.3
10^0
10^1
$r\,[\mathrm{Mpc}\,h^{-1}]$

Gaussian Process Regression

- Fitting Gaussian random field to maps—constrained Gaussian random field
  - A Gaussian random field: the joint PDF of the values at any field points are multivariate Gaussian
  - Characterized by correlation function
  - Can almost match anything→ universal fitting tool



(a), prior



(b), posterior

$$b(x) \sim \mathcal{N}\left(g(x), \sigma^2\right), \quad g(x) \sim \mathcal{GP}\left(0, K(x, x')\right) \longrightarrow \mathbb{E}[g(x')|\mathbf{b}(\mathbf{x})] = \mathbf{K}(x', \mathbf{x})[\mathbf{K}(\mathbf{x}, \mathbf{x}) + \Sigma]^{-1}\mathbf{b},$$