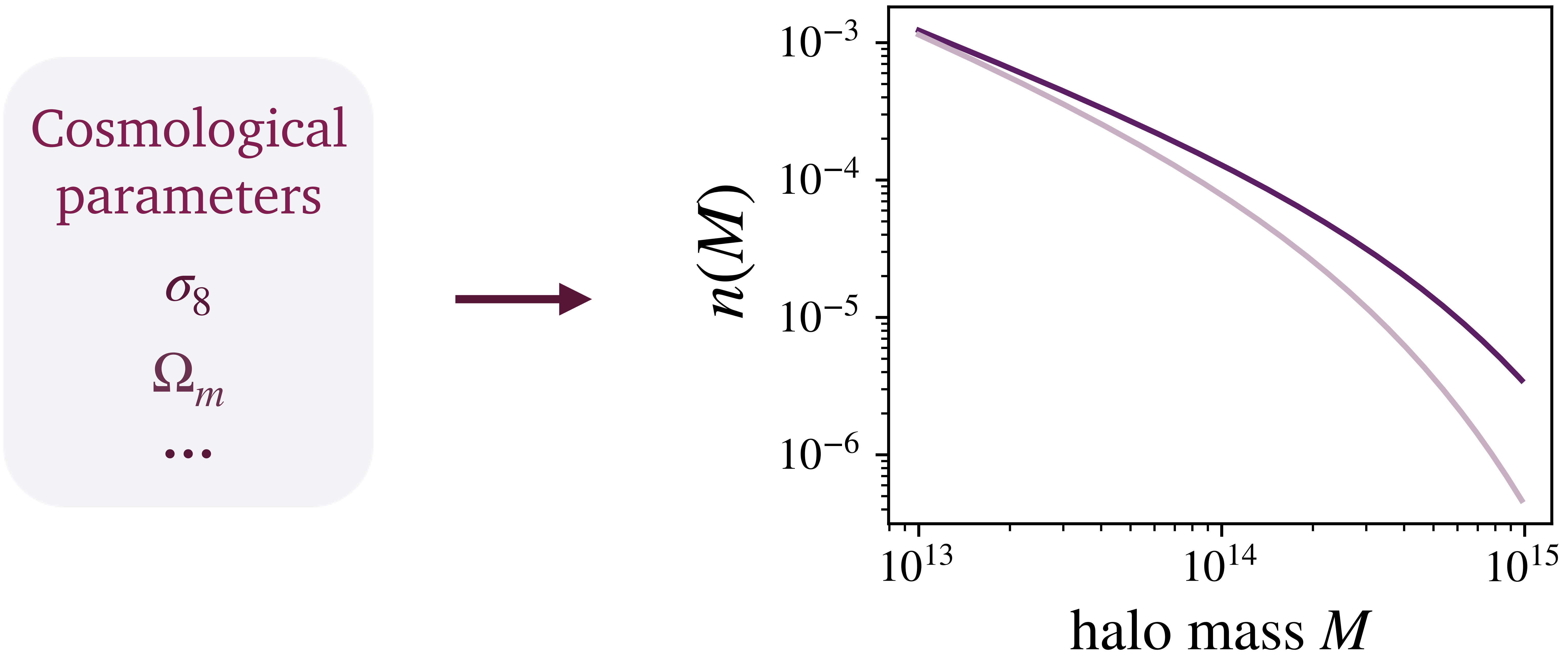
Explaining Halo Abundance with Interpretable Deep Learning

Lillian (Ningyuan) Guo
PhD Student, UCL

with Luisa Lucie-Smith (MPA), Hiranya Peiris (Cambridge), and Andrew Pontzen (UCL)

ningyuan.guo.20@ucl.ac.uk

Why do halo mass functions matter?



- Require accuracy at 1% or better to constrain cosmological parameters using galaxy cluster counts from forthcoming surveys, e.g. Rubin or Euclid (McClintock et al. 2019, Sartoris et al. 2016, Euclid Collaboration: Castro et al. 2023).

Universal halo mass function

Cosmological parameters → $P(k)$ → $\sigma(M)$ → $n(M)$

Linear theory

Calculate variance of overdensity in spheres of mass M

Cosmology-independent map, e.g. Press-Schechter, Tinker, etc.

- A simple, physically motivated halo mass function model.

$$n(M) = -\frac{\rho_c}{M}\,\Omega_m \frac{d\log\sigma}{d\log M}\,f(\sigma)$$

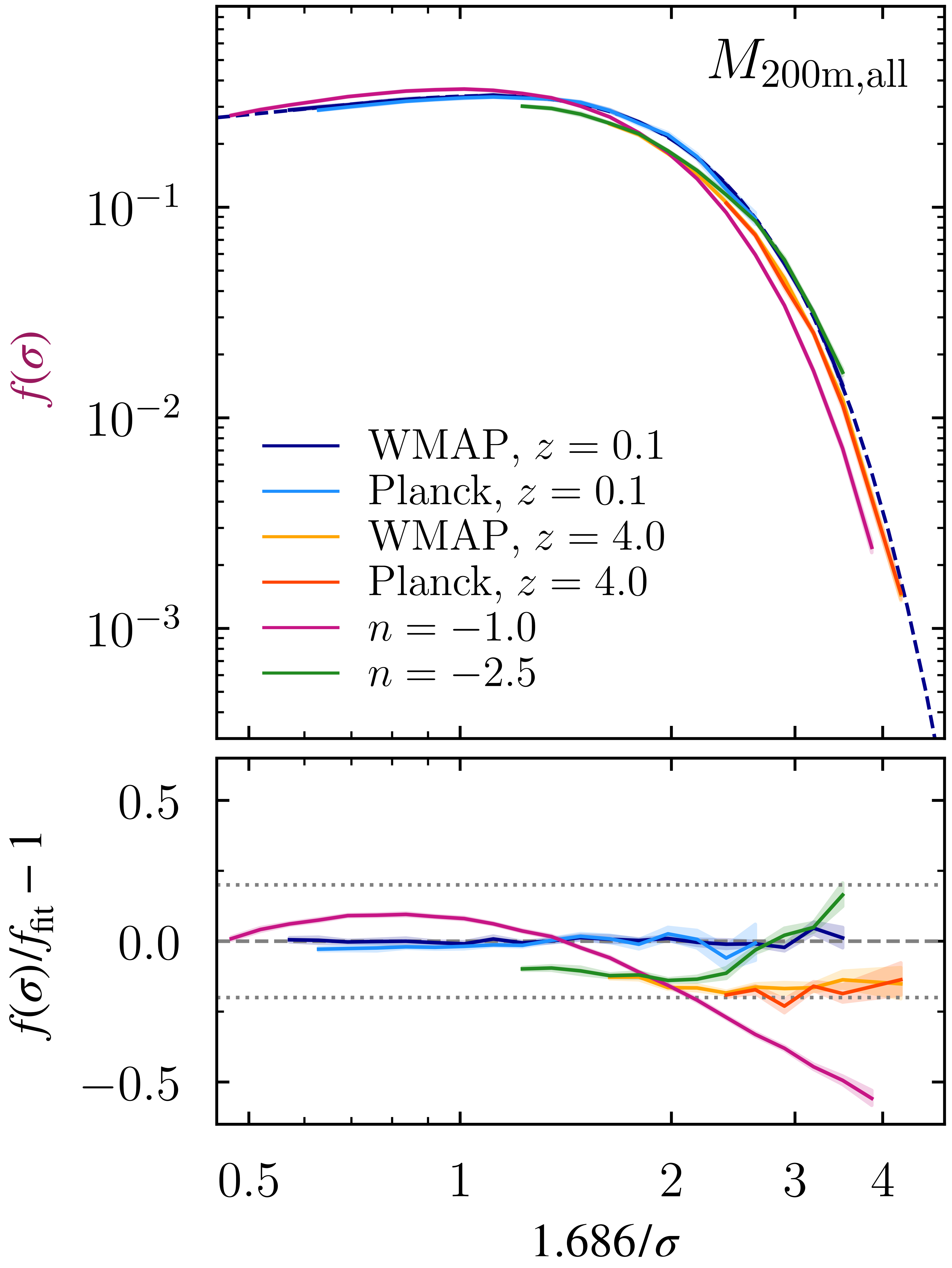
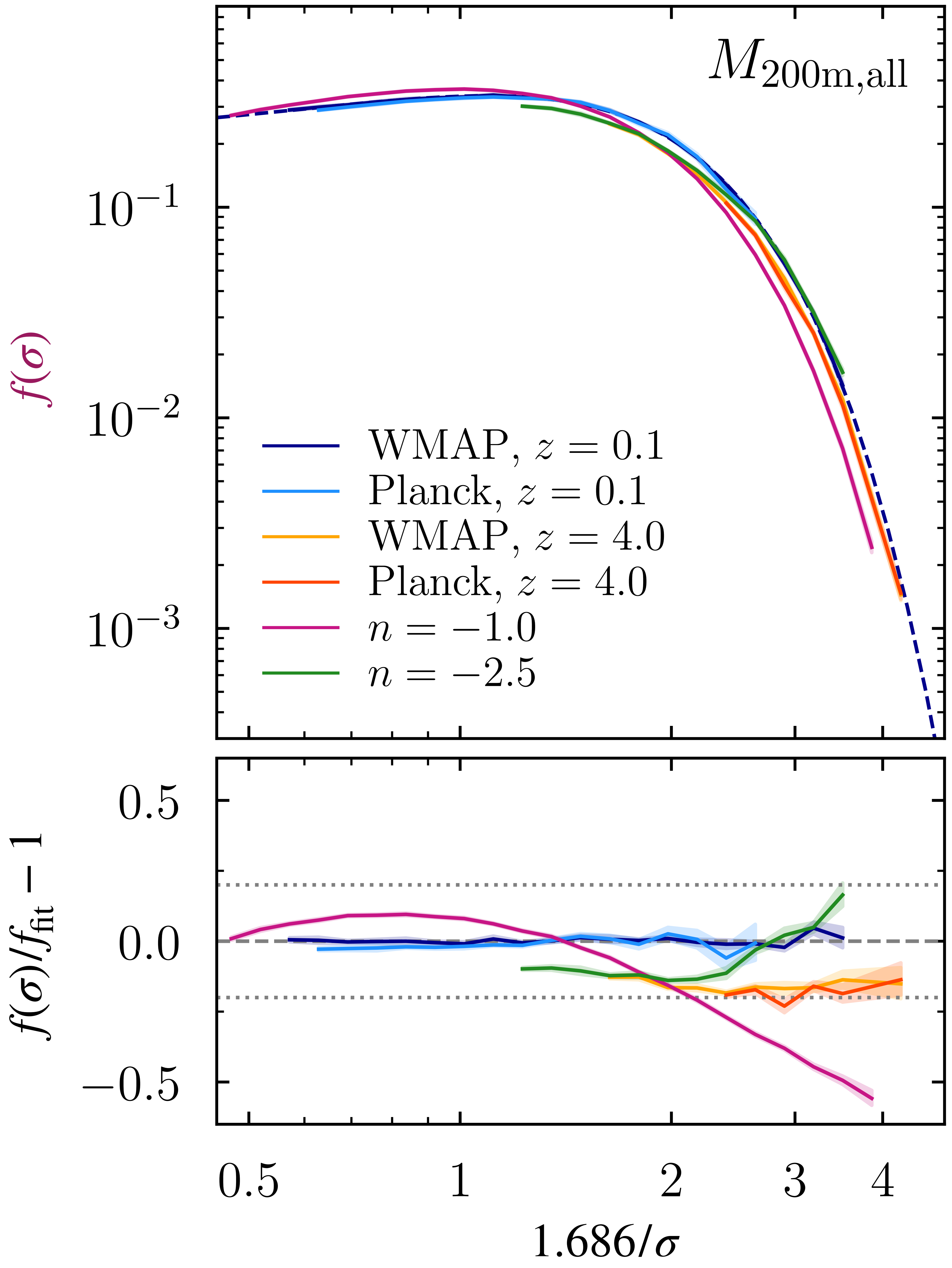


Figure: Diemer 2020

$$n(M) = -\frac{\rho_c}{M}\ \Omega_m \frac{d\log\sigma}{d\log M}\ f(\sigma)$$

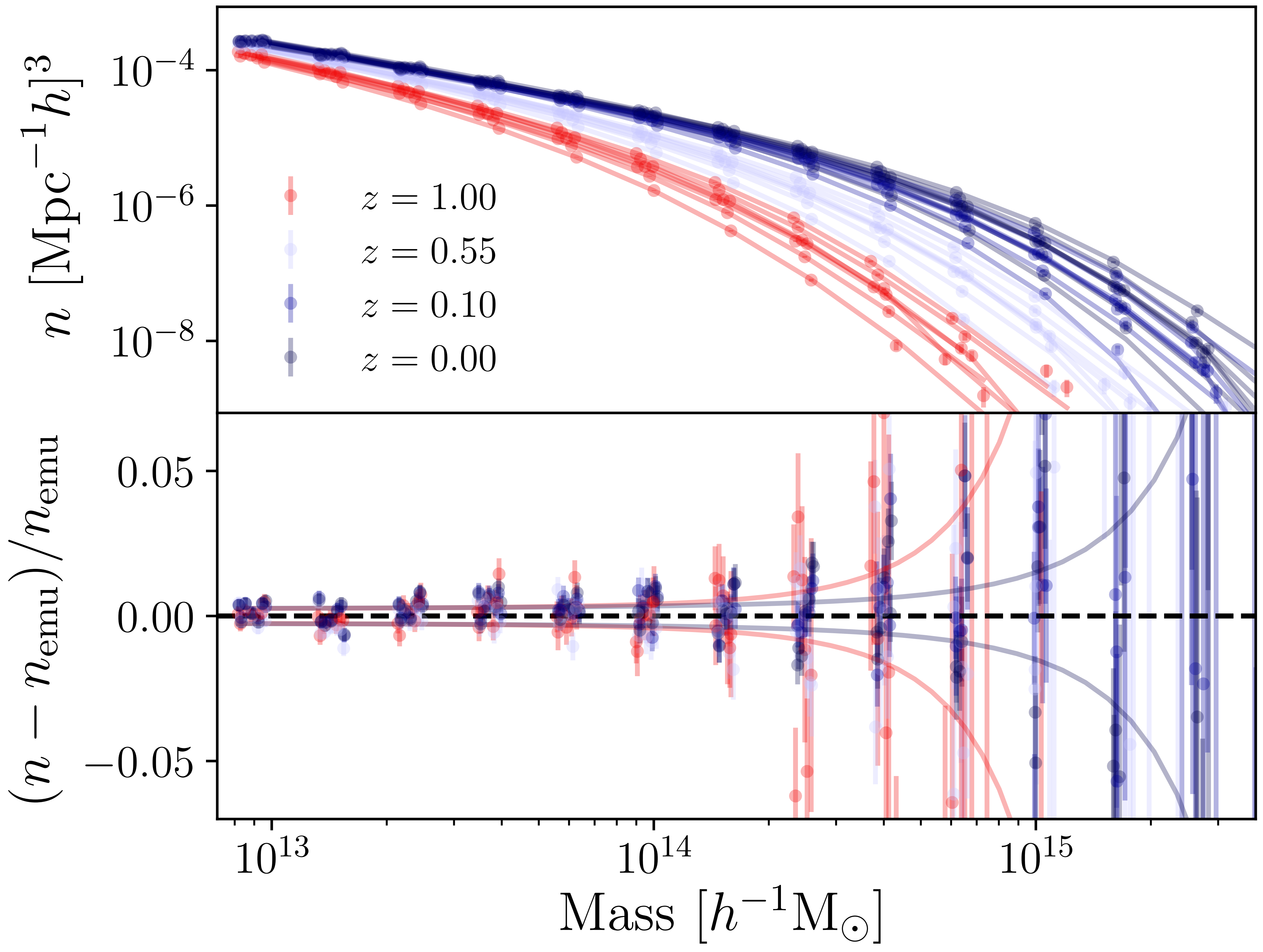
- Numerical simulations show the mass function is not universal at the percent level required for precision cosmology.

Figure: Diemer 2020



Beyond universality: emulators?

AEMULUS emulator (McClintock et al. 2019)



Advantage:

- Fit numerical simulations.
- Provide error estimates.

Disadvantage:

- Cannot generalise outside the parameter space it is trained on.

Beyond universality: what physical factors matter?

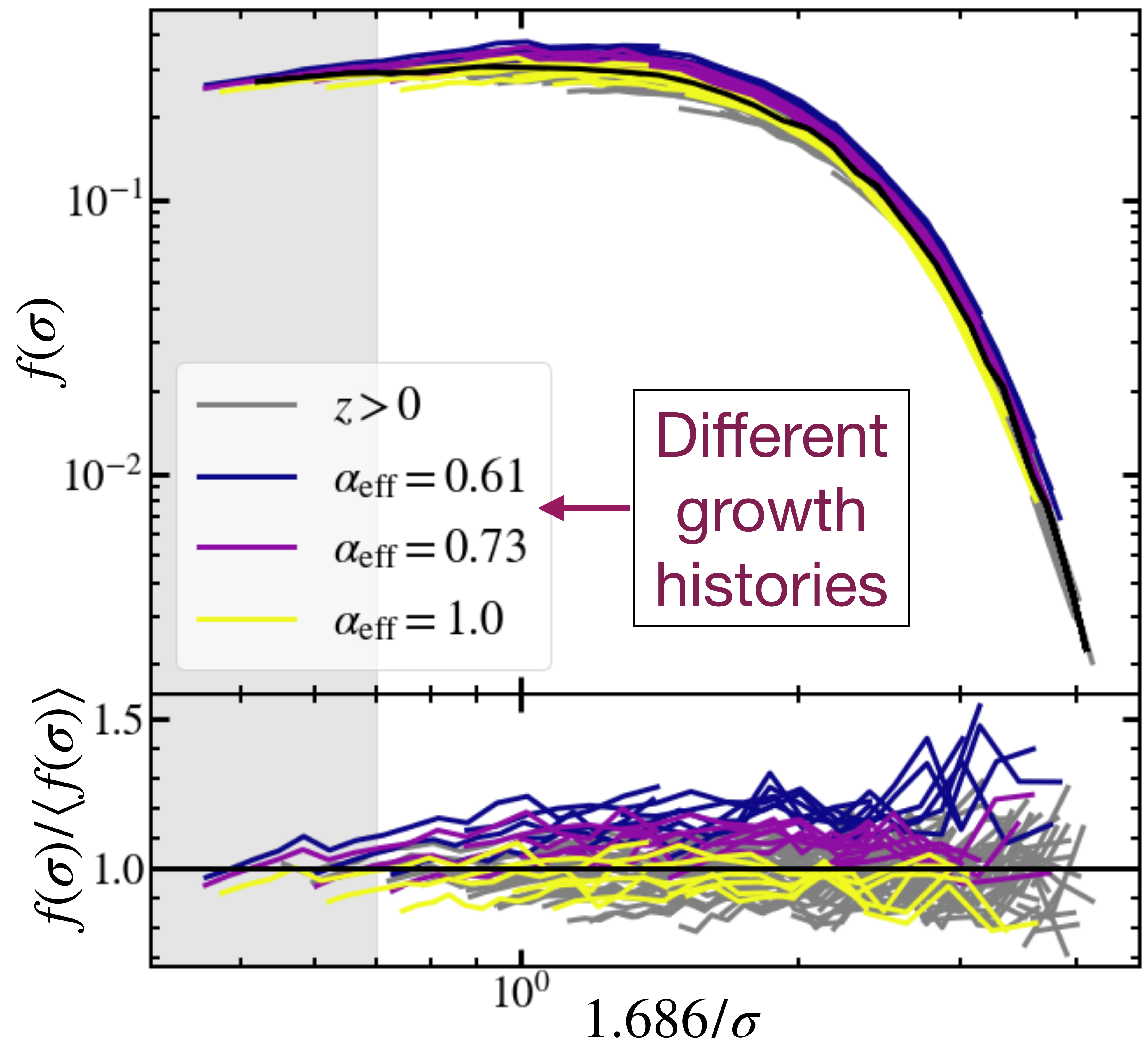
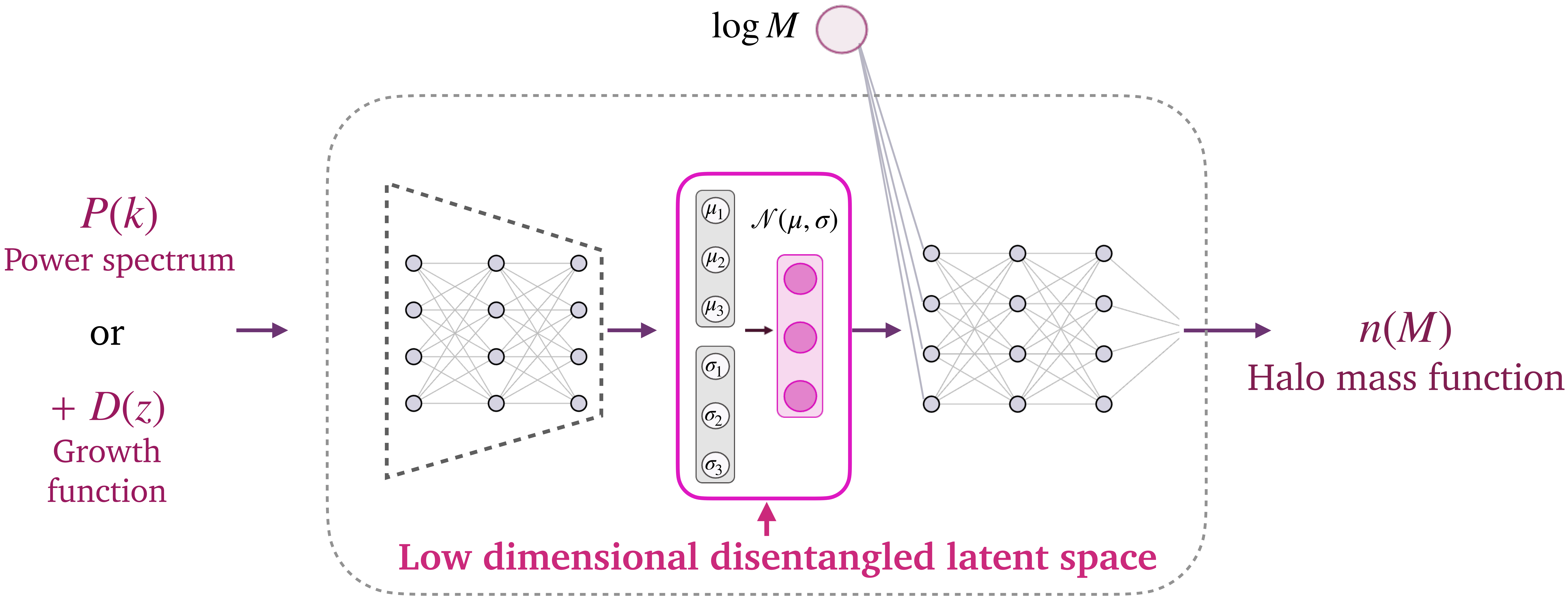


Figure: Ondaro-Mallea et al. 2022

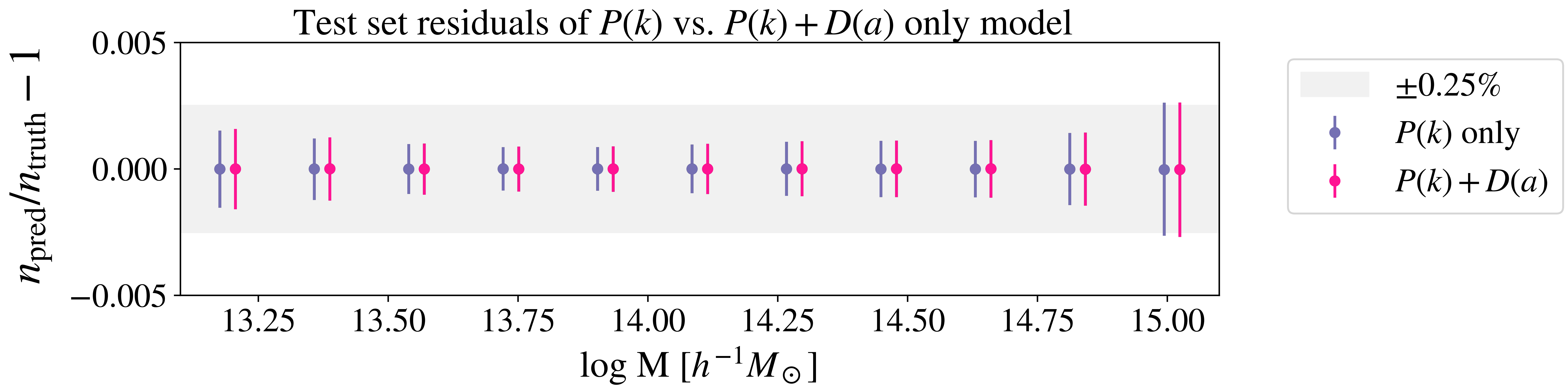
- Understanding what drives non-universality will help us build more generalisable accurate mass functions.
- E.g. is information from growth function important? (Courtin et al. 2011, Ondaro-Mallea et al. 2022)

Finding factors governing the halo mass function



Interpretable variational encoder:
Lucie-Smith et al. 2022, Iten et al. 2020.

Is growth function needed?



Ground truth mass functions at $z = 0$ from AEMULUS (wCDM cosmology).

- Growth function is not needed in addition to $P(k)$.
- Is the model learning growth information from $P(k)$?

Interpreting latent variables

- What do latent variables capture?
- Interpret using mutual information, an information theoretic metric of non-linear correlation.

$$\mathrm{MI}\left(x, y\right) = \int_x \int_y p\left(x, y\right) \ln\left[\frac{p\left(x, y\right)}{p\left(x\right) p\left(y\right)}\right] dxdy$$

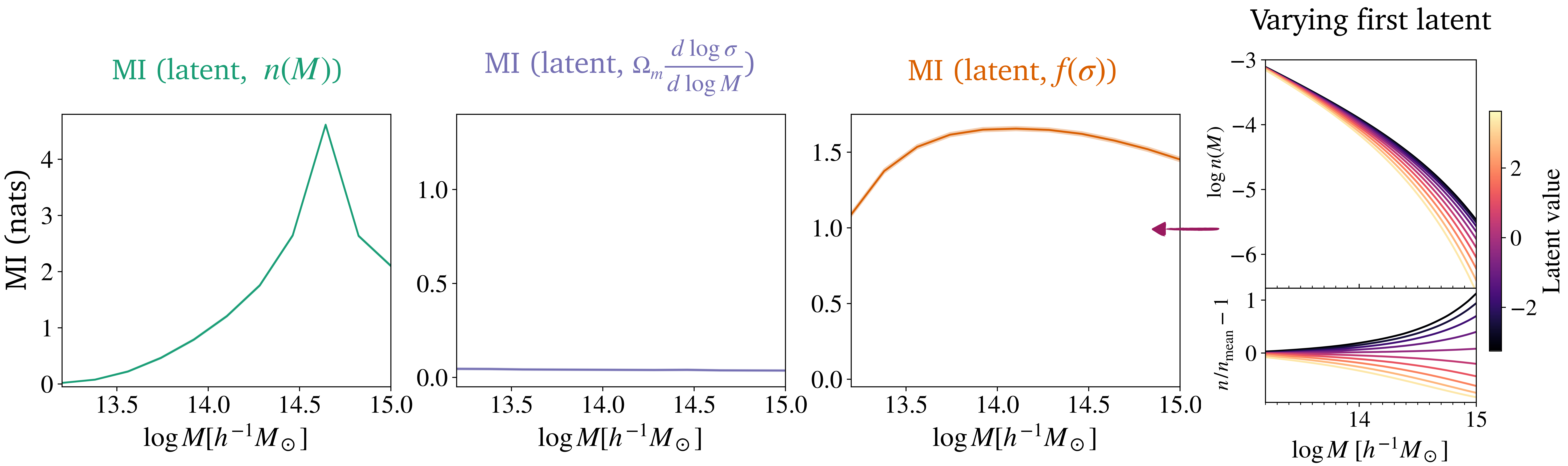
Interpreting latent variables: First latent

$$n(M) = -\frac{\rho_c}{M}\,\Omega_m \frac{d\log\sigma}{d\log M}\; f(\sigma)$$

σ(M): captures most of the cosmological dependence

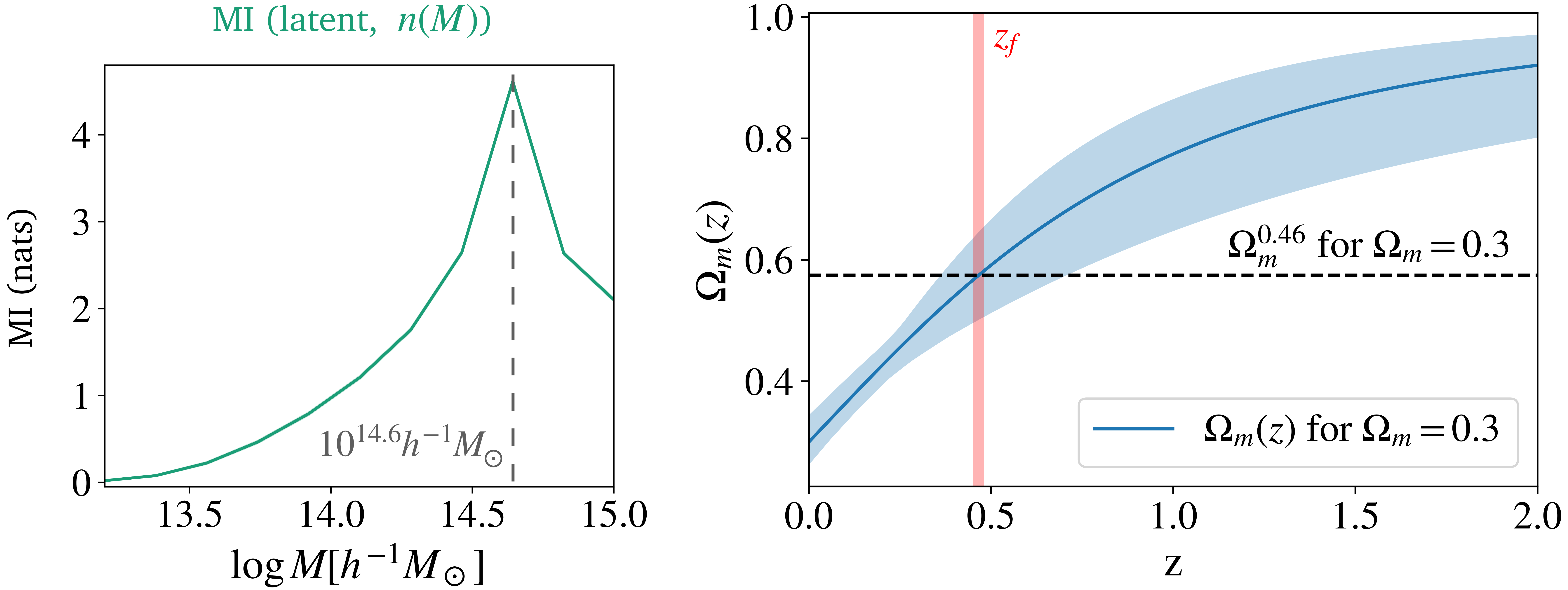
Maps between $f(\sigma)$ and $n(M)$

Mass function shape



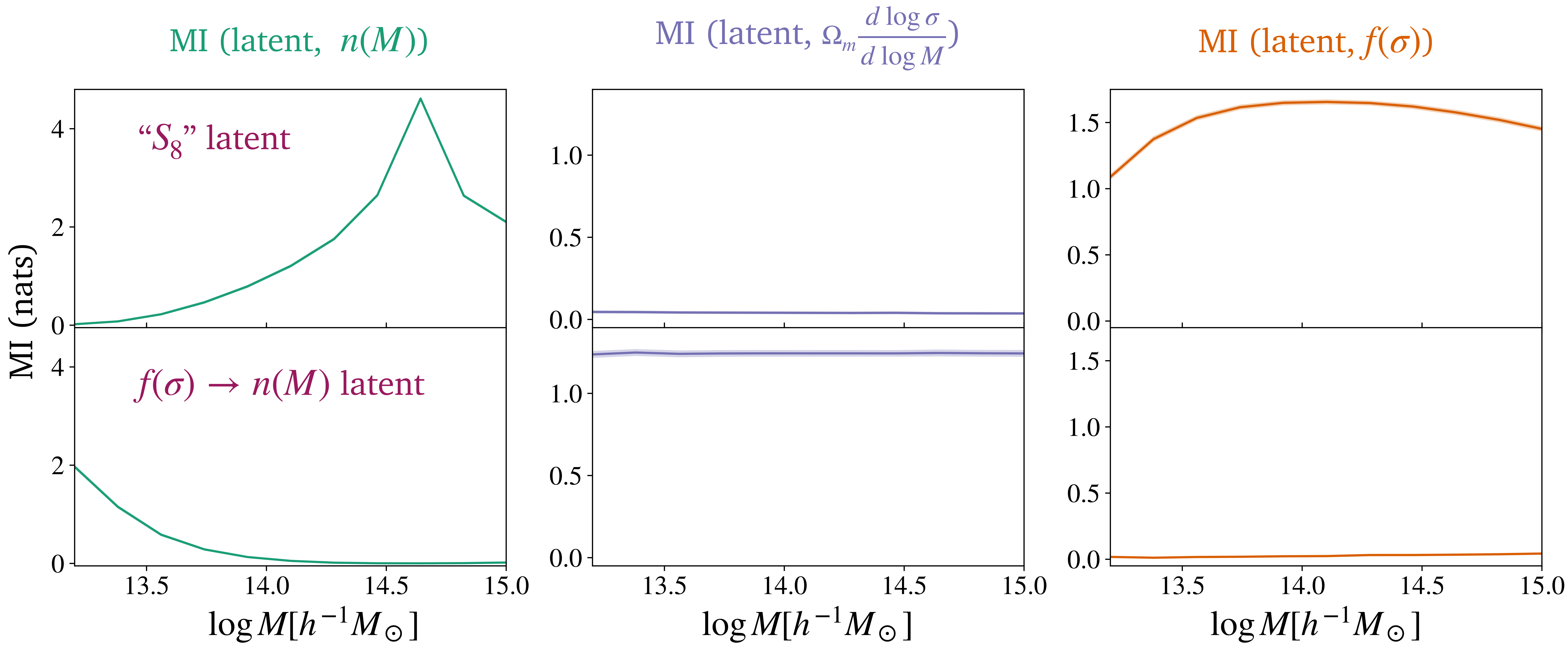
First latent: "S₈"

- Well-approximated by $\Omega_m^{0.46}\sigma_8$, which can be predicted from the formation history of $M = 10^{14.6}h^{-1}M_\odot$ halos and their mass fluctuation variance.

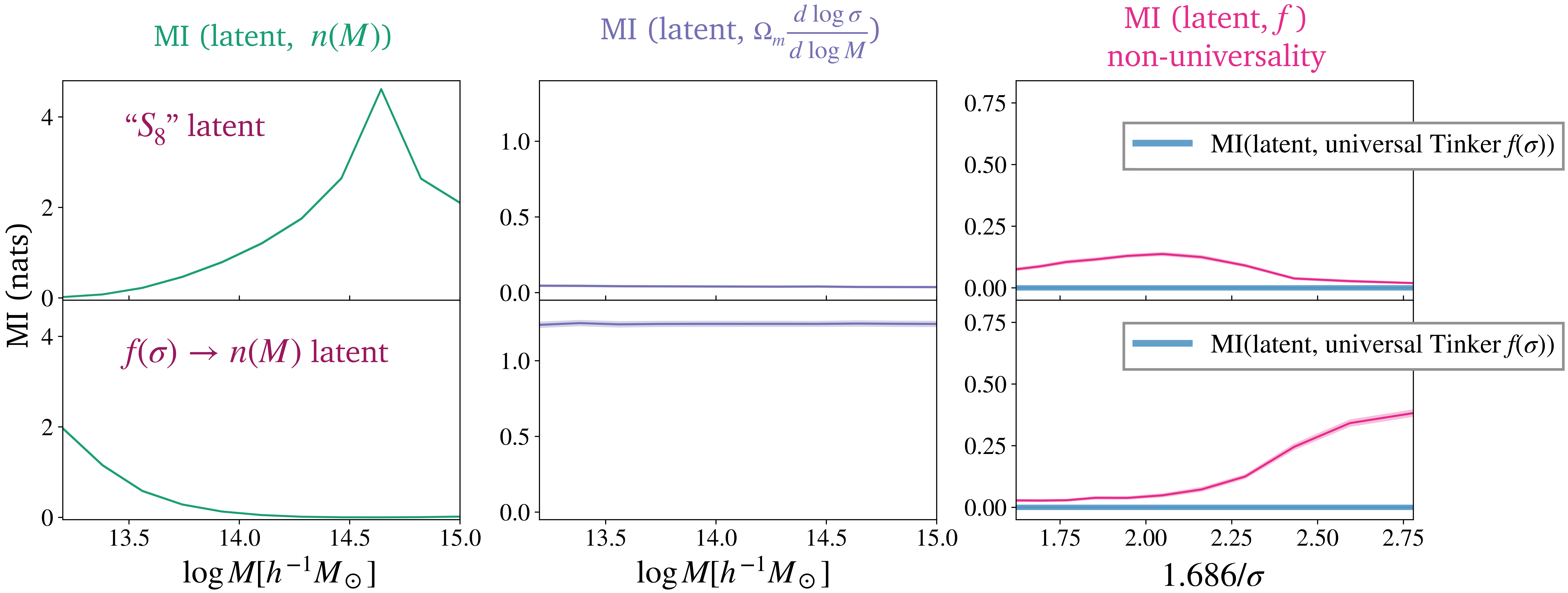


Half-mass z_f estimated using Correa et al. 2015

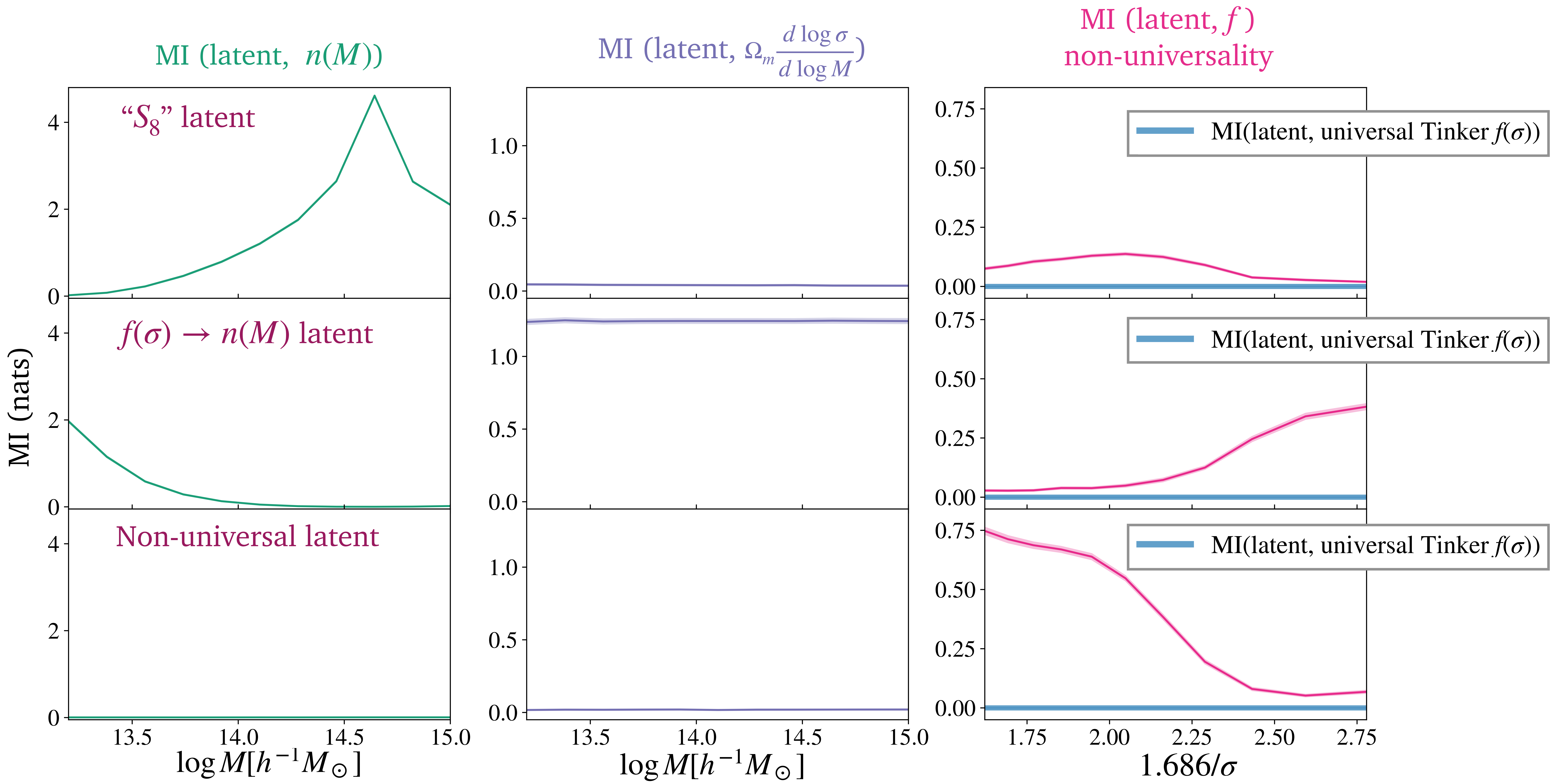
$$n(M) = -\frac{\rho_c}{M}\,\Omega_m \frac{d\log\sigma}{d\log M}\, f(\sigma)$$



Information beyond universality



Universal = Tinker. Non-universal = beyond Tinker.



MI (latent, $n(M)$)
MI (latent, $\Omega_m \frac{d\log\sigma}{d\log M}$)
MI (latent, f)
non-universality
"S_8" latent
$f(\sigma) \rightarrow n(M)$ latent
Non-universal latent
MI (nats)
MI(latent, universal Tinker $f(\sigma)$)
$\log M[h^{-1}M_\odot]$
$1.686/\sigma$

Summary

- Using deep learning to understand the origin of non-universality in the halo mass function.
- Found that growth function provides no additional information to $P(k)$ for predicting the mass function at $z = 0$.
- Using mutual information to interpret the latent variables to understand their physical relevance to structure formation.
- The low dimensional latent space can be sampled over to train accurate emulators using fewer simulations.

Thank you for your attention!

ningyuan.guo.20@ucl.ac.uk

Extra slides

Interpretable latent representation

- Latent representation is interpretable if it is disentangled: each physically relevant factor is captured by a different, independent latent variable.

- Train network to reduce loss (Higgins et al. 2017)

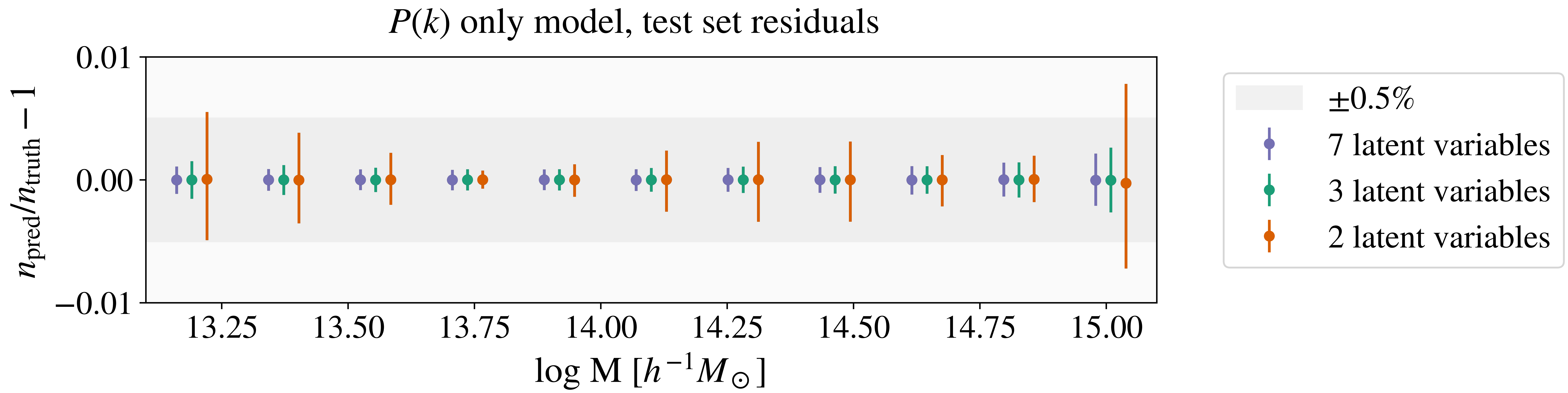
$$\mathscr{L} = \text{MSE} + \beta \cdot \mathscr{D}_{\text{KL}}(q(\mathbf{z}\,|\,x)||p(\mathbf{z}))$$

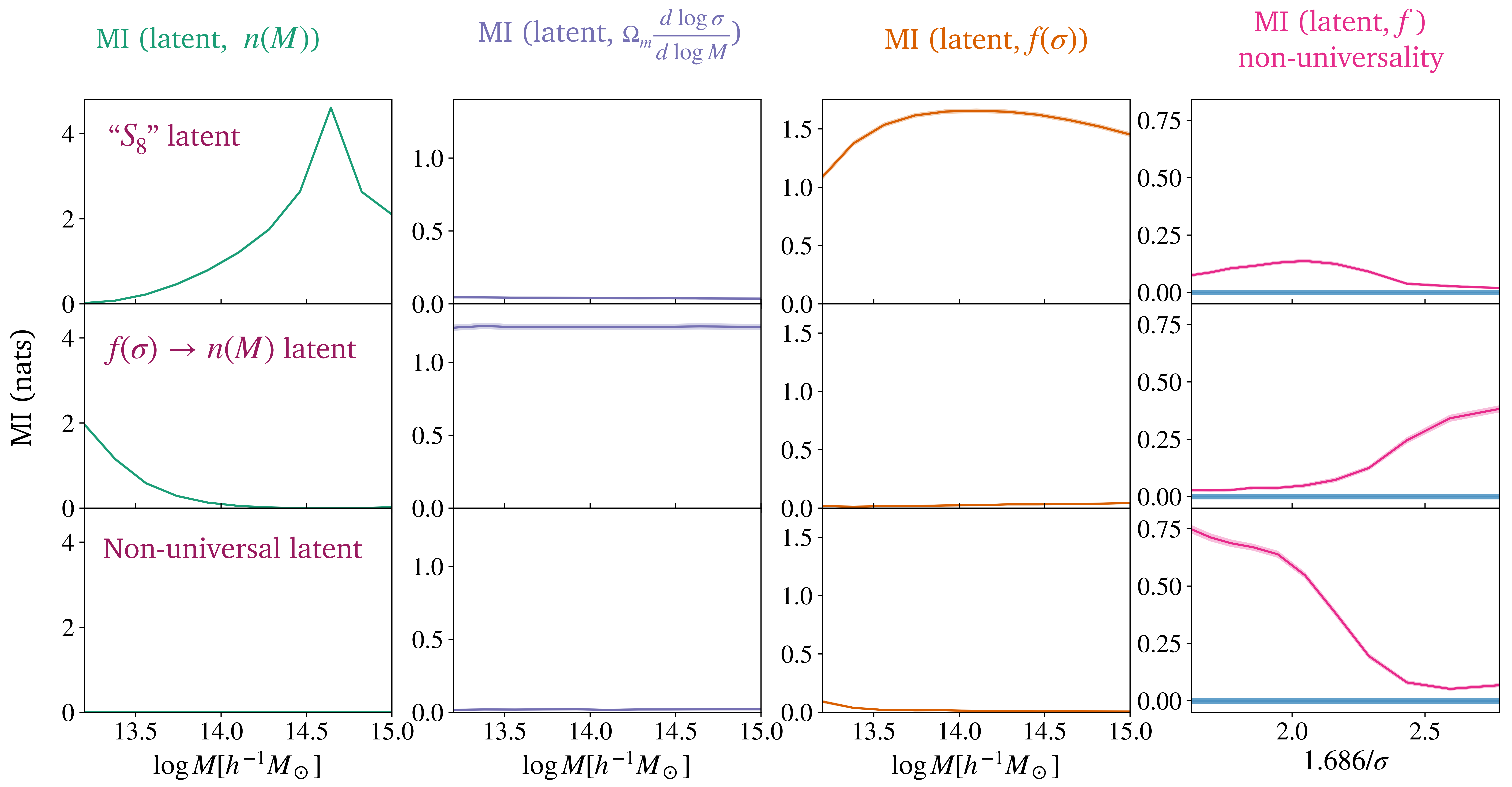
Penalise inaccurate predictions

$$\text{MSE} = \frac{1}{N}\sum_{N}\left(\log\frac{dn_{\text{pred}}}{d\log M} - \log\frac{dn_{\text{truth}}}{d\log M}\right)^2$$

Encourage latent distribution learnt, $q(\mathbf{z}\,|\,x)$, to resemble independent unit Gaussians $p(\mathbf{z})$

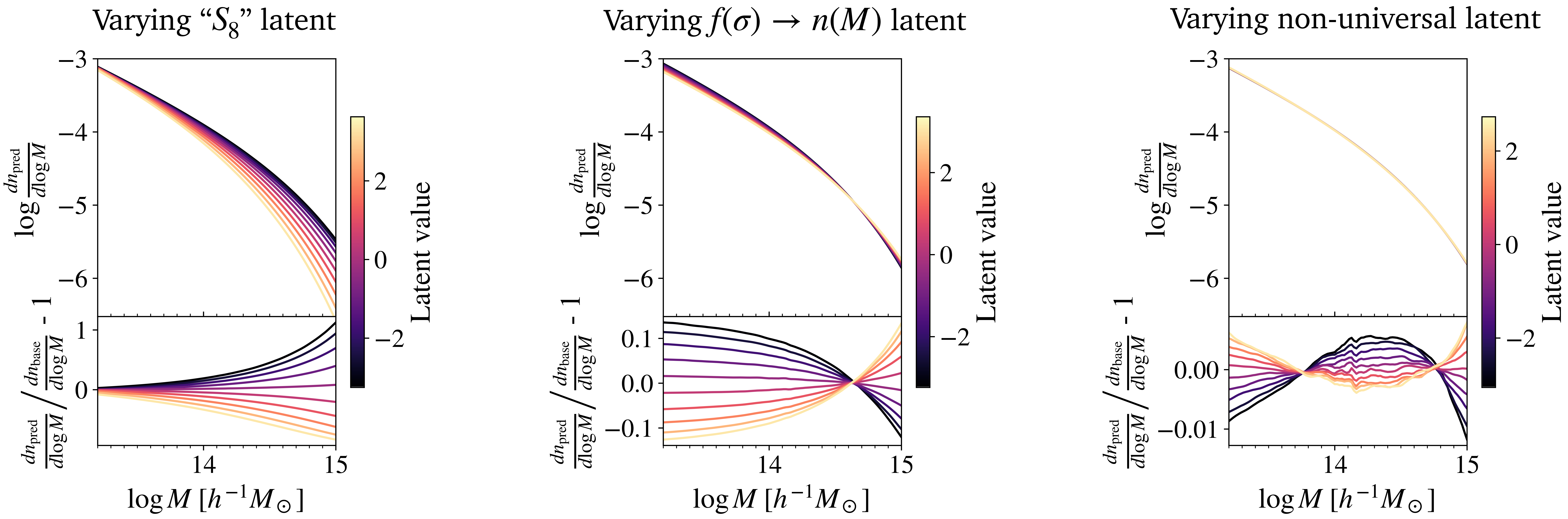
Number of latent variables required



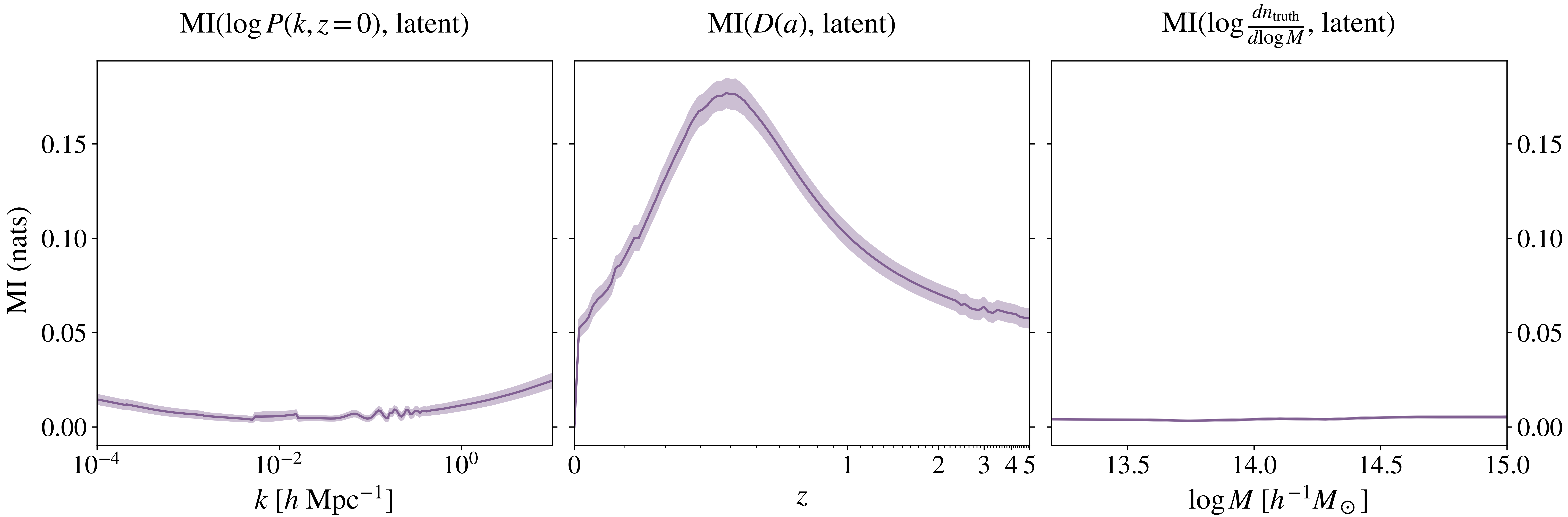


MI (latent, $n(M)$)
MI (latent, $\Omega_m \frac{d\log\sigma}{d\log M}$)
MI (latent, $f(\sigma)$)
MI (latent, f)
non-universality
"S_8" latent
$f(\sigma) \rightarrow n(M)$ latent
Non-universal latent
MI (nats)
$\log M[h^{-1}M_\odot]$
$1.686/\sigma$

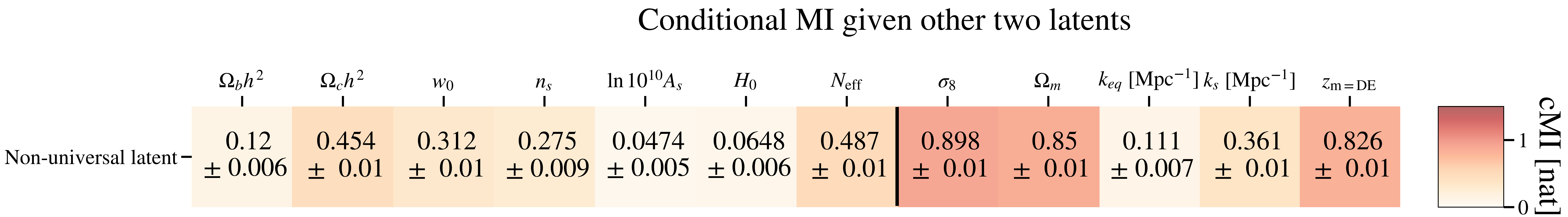
Latent traversals



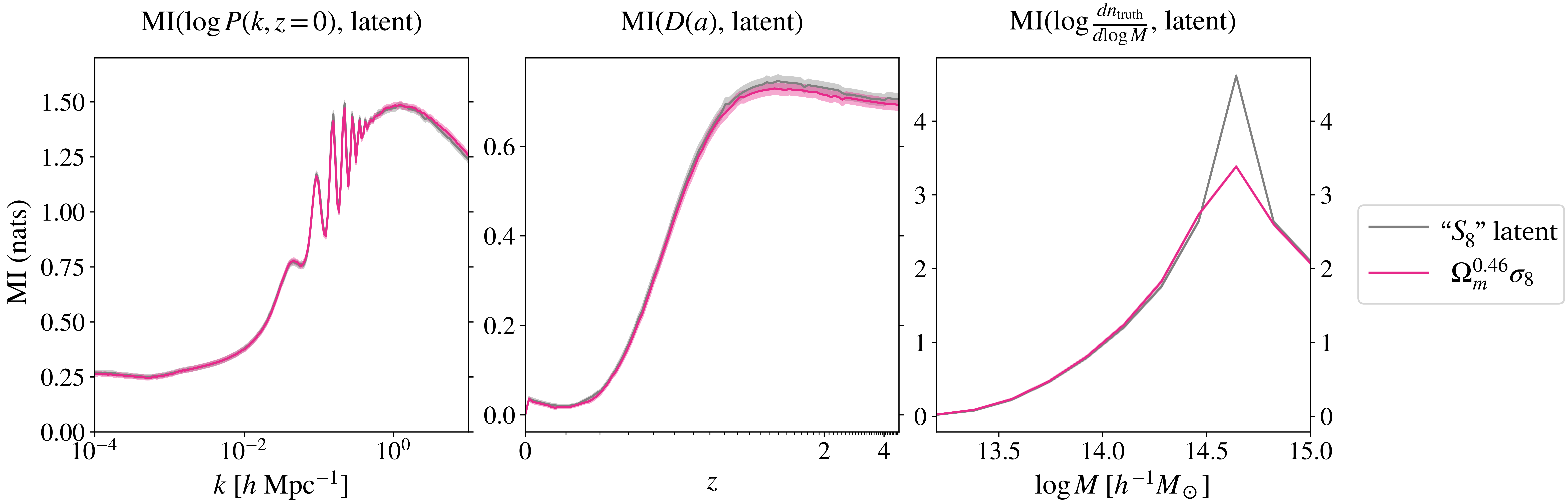
Non-universal latent



Non-universal latent

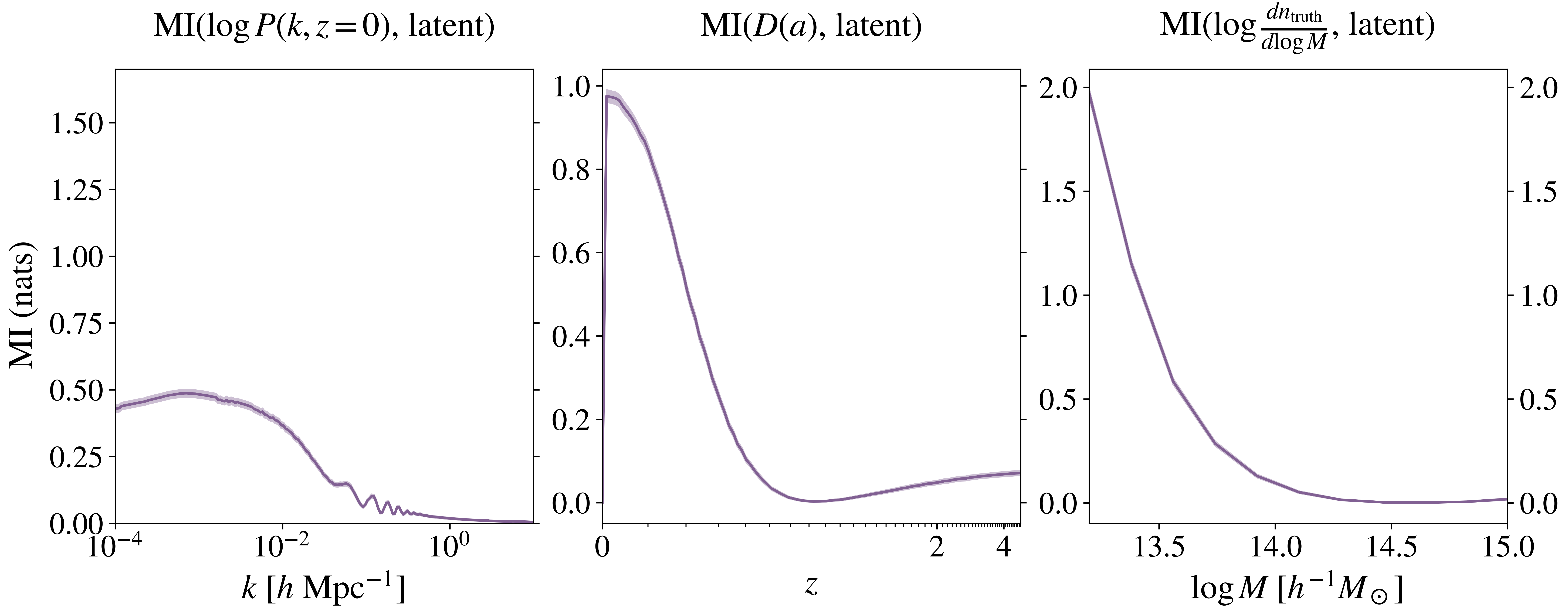


“S8” latent



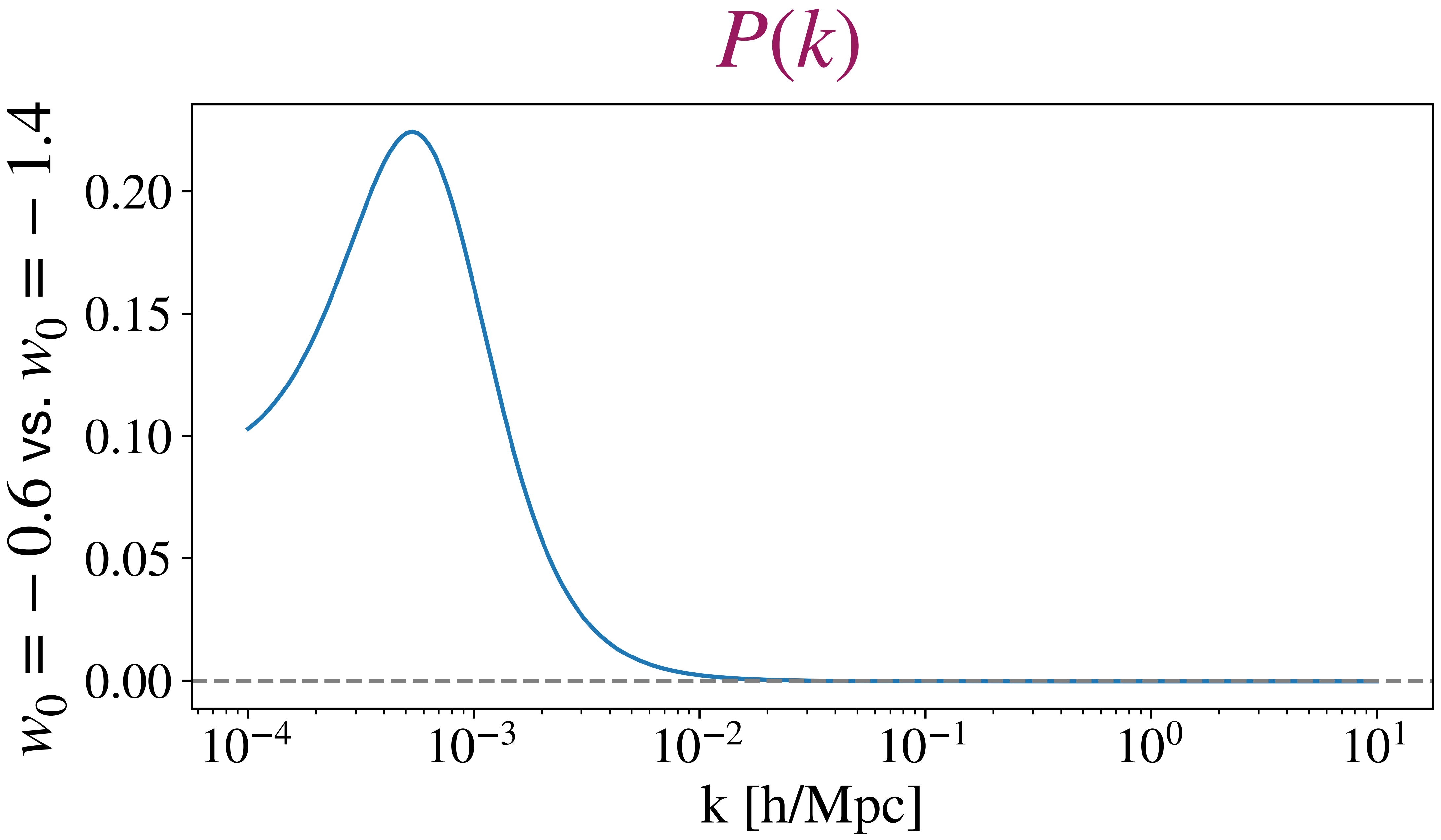
MI(log $P(k, z=0)$, latent)
MI($D(a)$, latent)
MI(log $\frac{dn_{\rm truth}}{d\log M}$, latent)
MI (nats)
1.50
1.25
1.00
0.75
0.50
0.25
0.00
10^{-4}
10^{-2}
10^{0}
k [h Mpc^{-1}]
0.6
0.4
0.2
0.0
0
2
4
z
4
3
2
1
0
13.5
14.0
14.5
15.0
$\log M$ [$h^{-1} M_\odot$]
“S_8” latent
$\Omega_m^{0.46}\sigma_8$

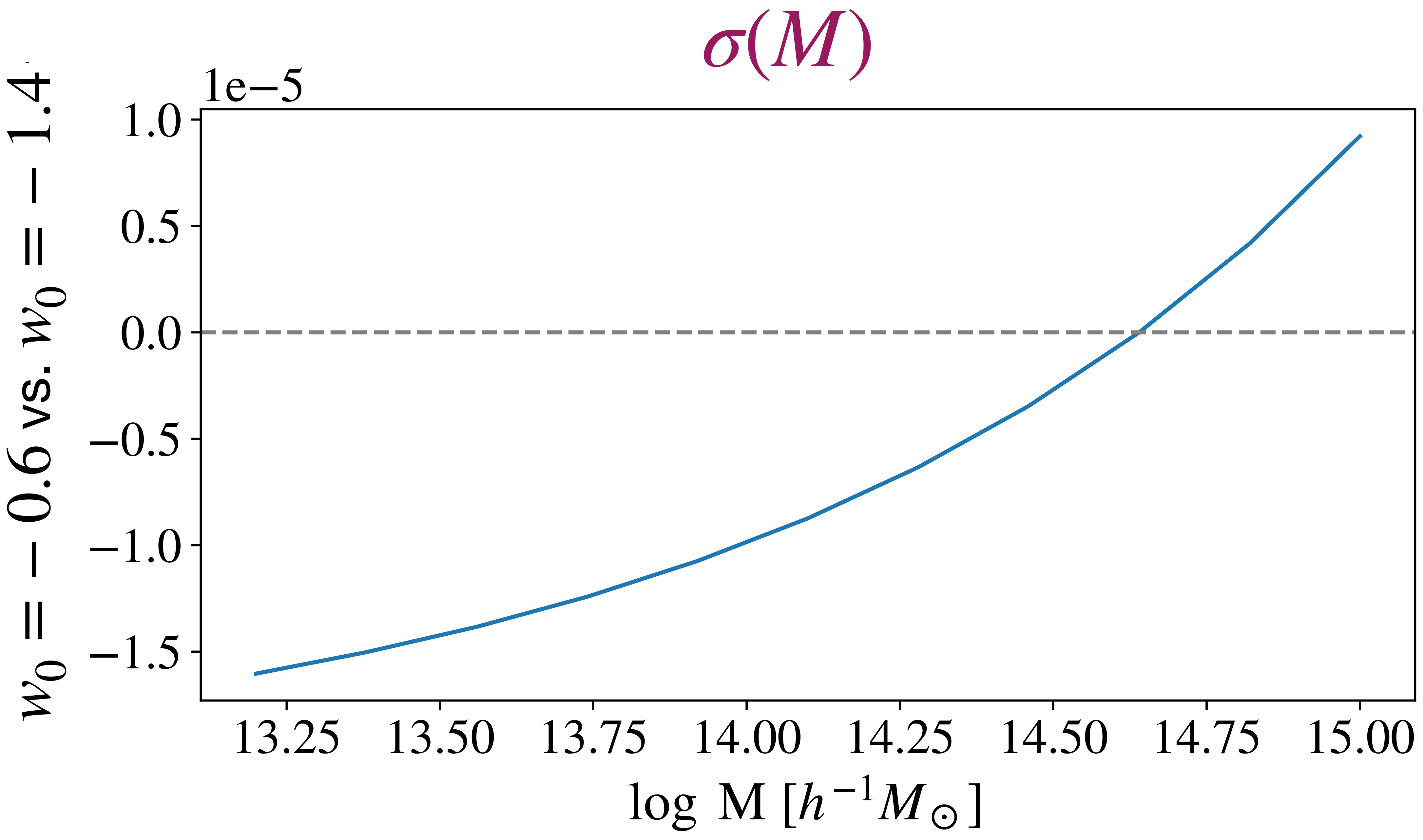
$f(\sigma) \rightarrow n(M)$ latent



Information on w_0 in power spectrum

Fractional difference after accounting for normalisation





"S8" latent

MI with cosmological parameters varied | MI with physical quantities

	$\Omega_b h^2$	$\Omega_c h^2$	w_0	n_s	$\ln 10^{10} A_s$	H_0	$N_{\rm eff}$	σ_8	Ω_m	k_{eq} [Mpc^{-1}]	k_s [Mpc^{-1}]	$z_{\rm m=DE}$
"S_8" latent	0.129 ± 0.007	0.714 ± 0.01	0.451 ± 0.009	0.0722 ± 0.005	0.104 ± 0.008	0.232 ± 0.008	0.0533 ± 0.006	1.14 ± 0.01	0.059 ± 0.005	1.1 ± 0.01	0.0249 ± 0.003	0.625 ± 0.01

MI [nat] (colour scale 0–1)

	$\Omega_b h^2$	$\Omega_c h^2$	w_0	n_s	$\ln 10^{10} A_s$	H_0	$N_{\rm eff}$	σ_8	Ω_m	k_{eq} [Mpc^{-1}]	k_s [Mpc^{-1}]	$z_{\rm m=DE}$
"S_8" latent	0.194 ± 0.01	0.128 ± 0.007	0.295 ± 0.01	0.0861 ± 0.008	0.158 ± 0.009	0.0881 ± 0.008	0.11 ± 0.008			0.0368 ± 0.004	0.0627 ± 0.007	0.36 ± 0.01

cMI [nat] (colour scale 0–1)