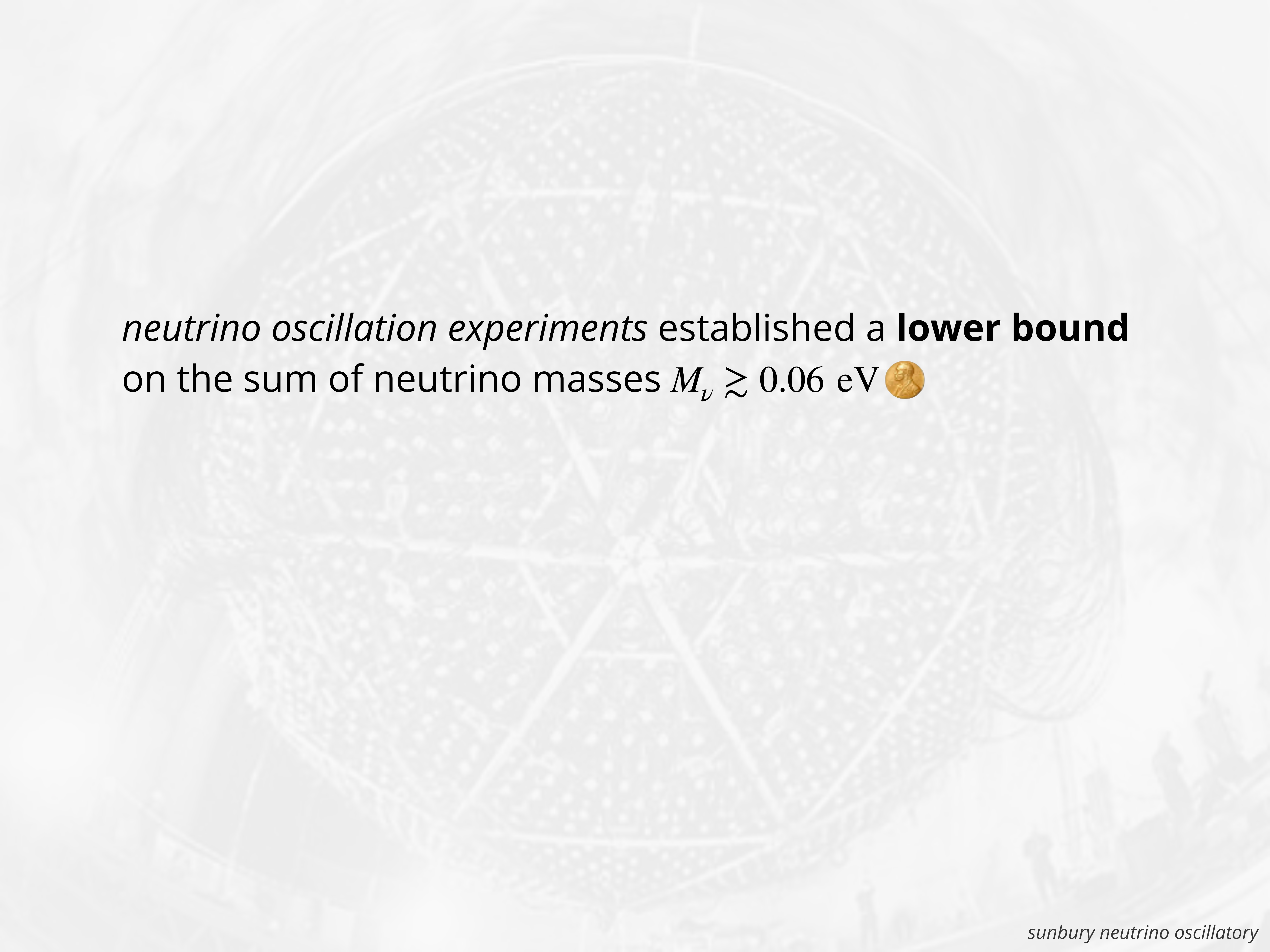
constraining M_ν with *the bispectrum*

Changhoon Hahn

PRINCETON UNIVERSITY

arXiv:1909.05273
arXiv:1909.11107
arXiv:2012.02200

Feb 17, 2021 — APEC Seminar IPMU



neutrino oscillation experiments established a **lower bound** on the sum of neutrino masses $M_\nu \gtrsim 0.06\ \mathrm{eV}$

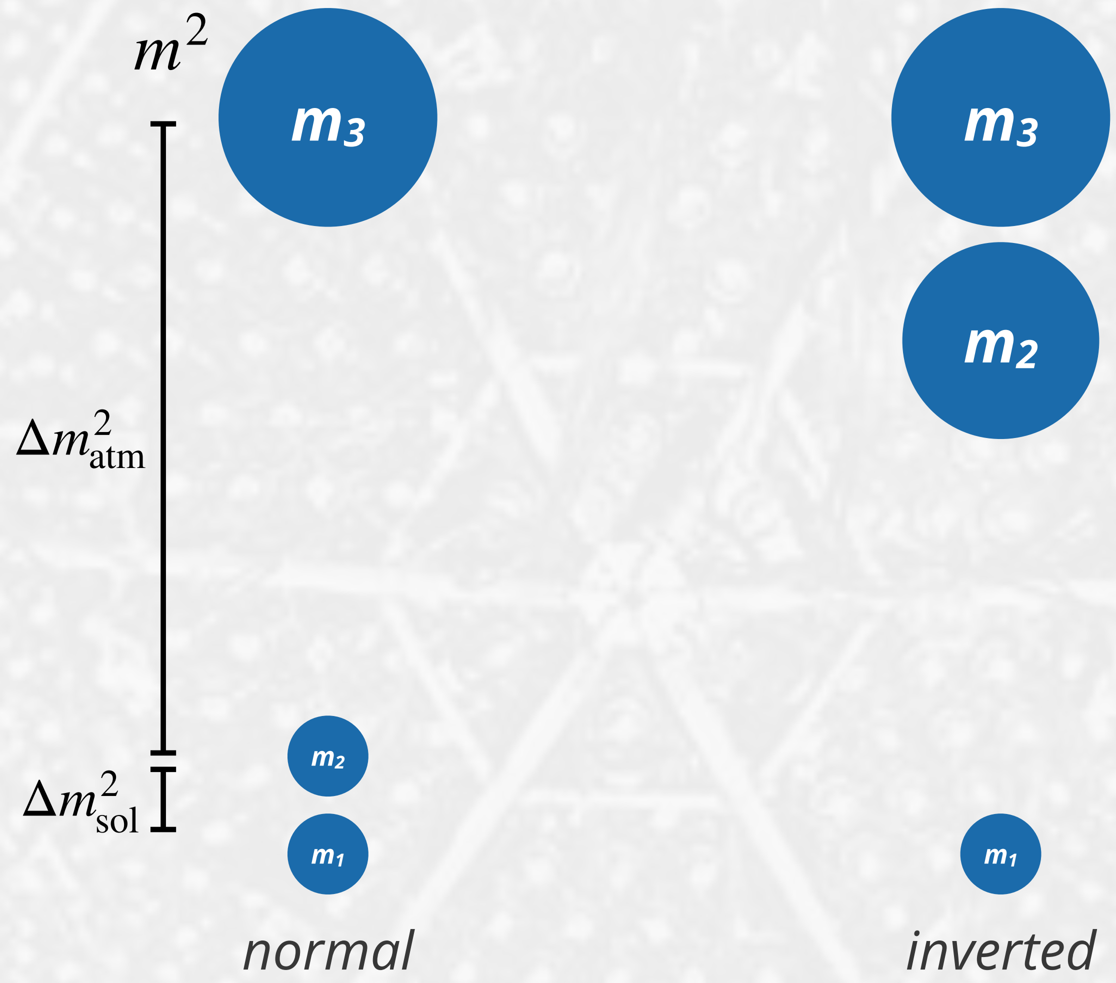
sunbury neutrino oscillatory

neutrino oscillation experiments established a **lower bound** on the sum of neutrino masses $M_\nu \gtrsim 0.06\ \mathrm{eV}$

but they don't measure M_ν

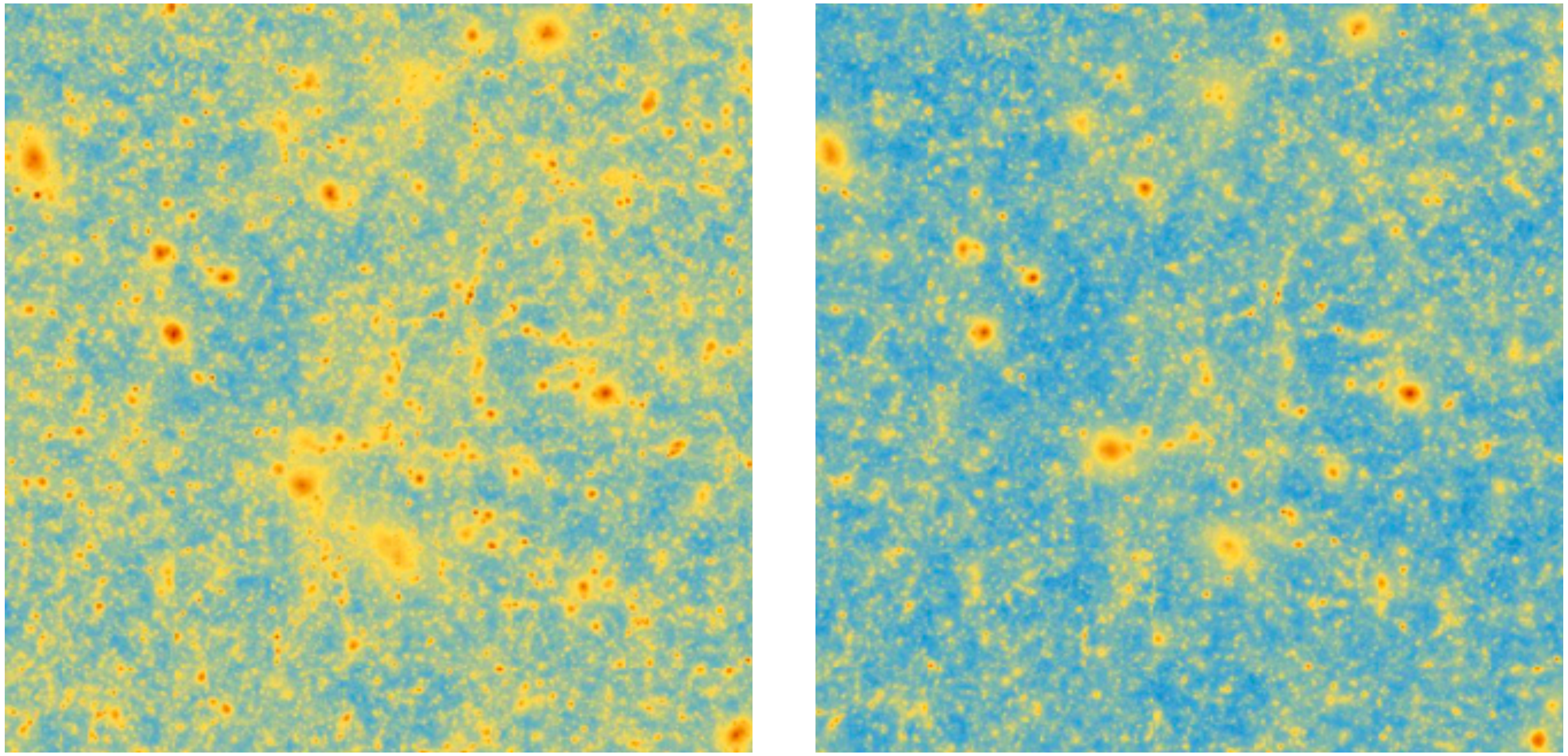
sunbury neutrino oscillatory

precise M_ν reveal physics **beyond the Standard Model** (*e.g.* distinguish between *"normal" and "inverted" hierarchies)*



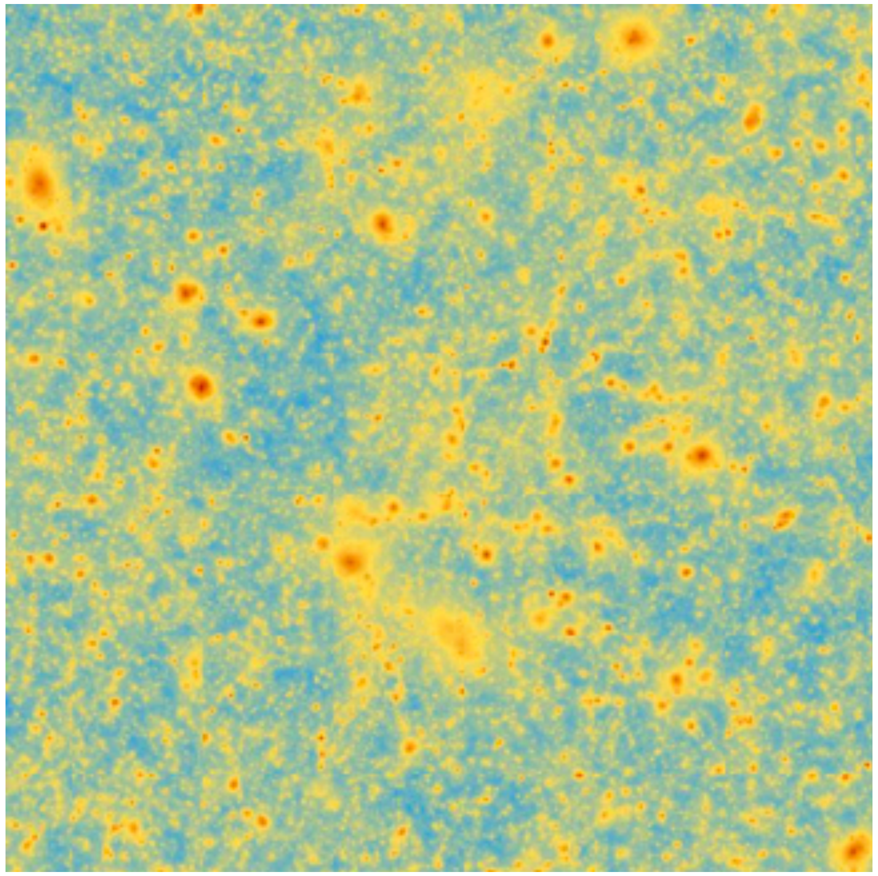
upcoming laboratory experiments are *not sensitive to* M_ν

$M_\nu > 0$ eV **suppresses the growth of structure** on *small scales* below their free-streaming scale

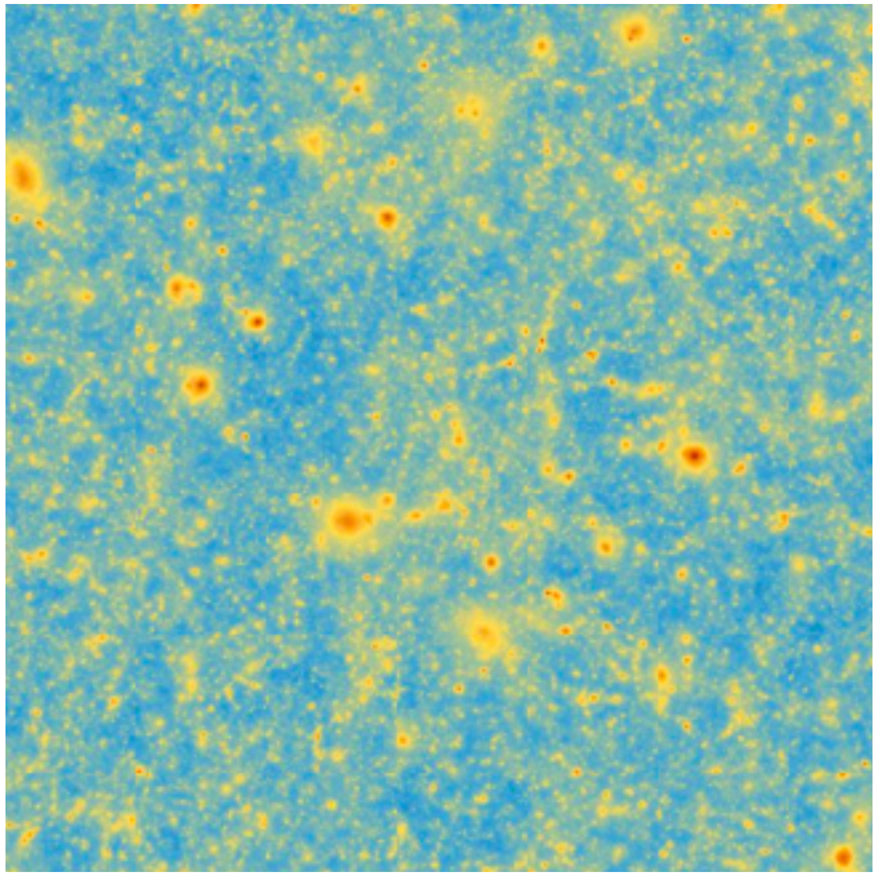


credit: BAHAMAS

$M_\nu > 0$ eV **suppresses the growth of structure** on *small scales* below their free-streaming scale



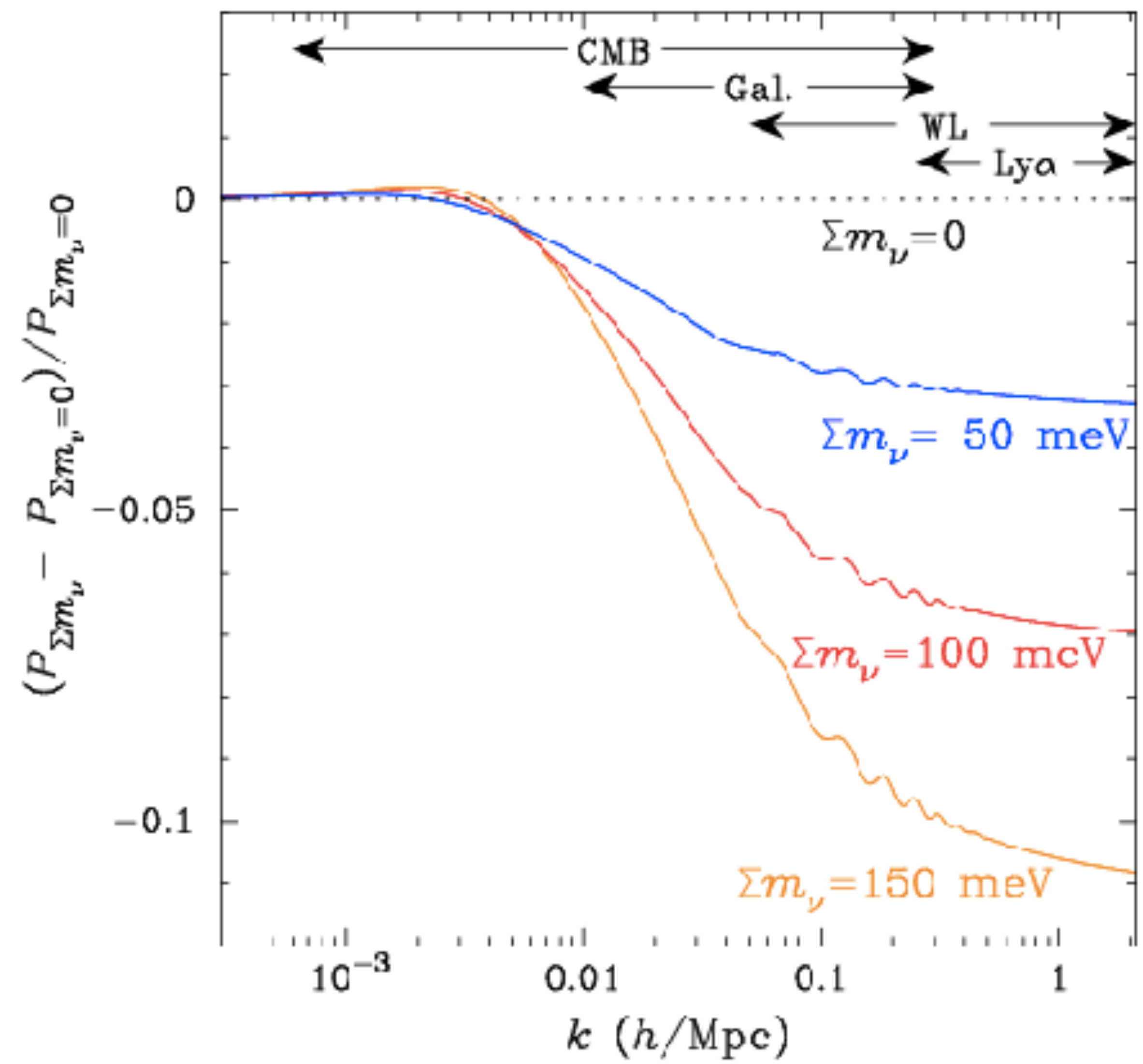
$M_\nu = 0$ eV



$M_\nu > 0$ eV

credit: BAHAMAS

$M_\nu > 0$ eV **suppresses the growth of structure** on *small scales* below their free-streaming scale



credit: Lesgourgues+(2013)

best constraints currently come from cosmology: **CMB + LSS**

$M_\nu < 0.13\ \mathrm{eV}$ (*95%, Planck TT+lowE+lensing+BAO*)

Planck Collaboration (2018)

best constraints currently come from cosmology: **CMB + LSS**

$M_\nu < 0.13$ eV (*95%, Planck TT*+*lowE*+*lensing*+*BAO*)

KATRIN 1.1eV upper limit (Aker+2019)

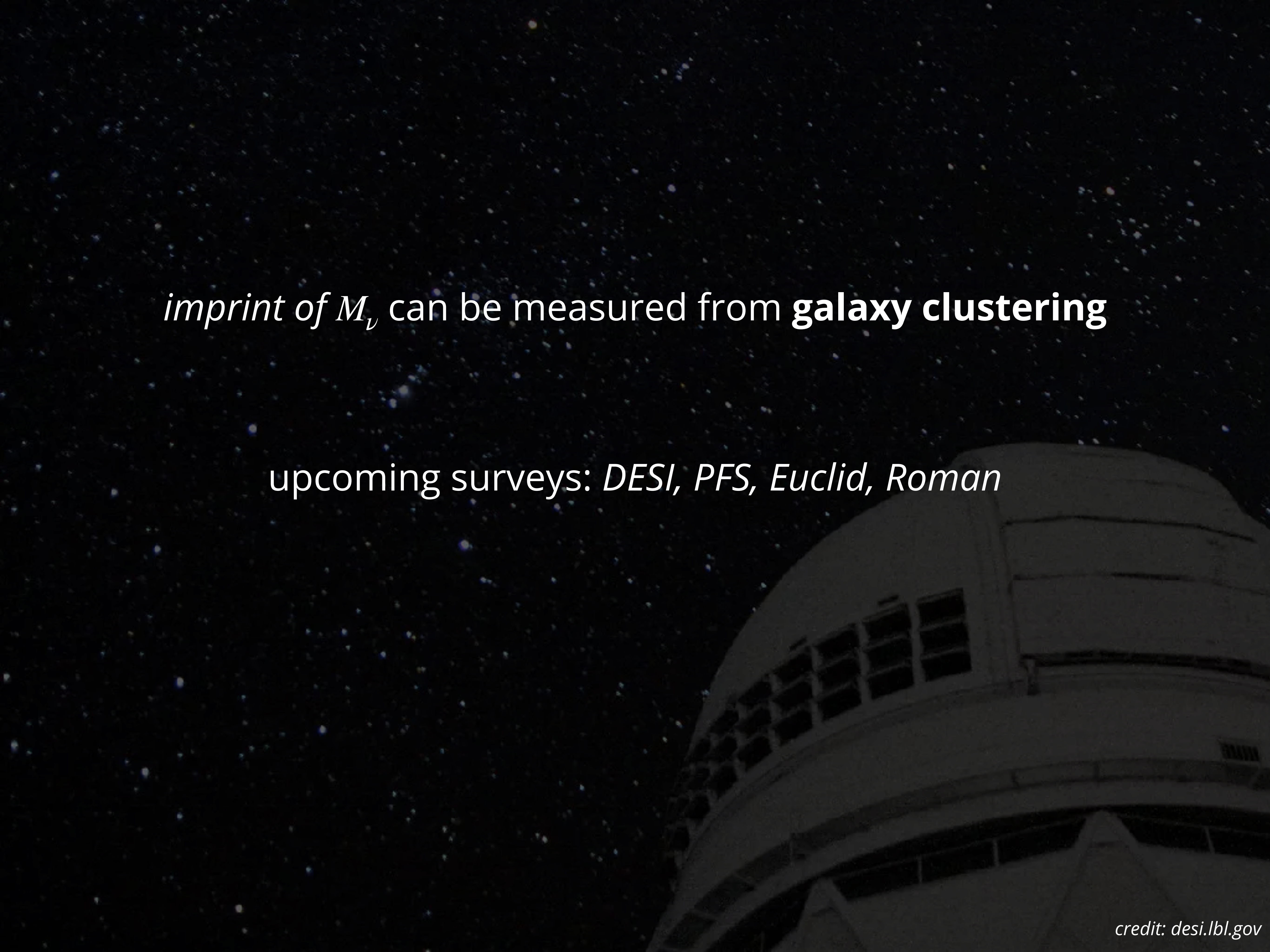
CMB measures $A_s e^{-2\tau}$ and thus **heavily rely on τ constraints**

Allison et al. (2015), Mishra-Sharma et al. (2018)

CMB measures $A_s e^{-2\tau}$ and thus **heavily rely on τ constraints**

but upcoming ground-based experiments (*CMB-S4*)
will not directly constrain τ … *LiteBIRD?, LiteCOrE?*

Allison et al. (2015), Mishra-Sharma et al. (2018)

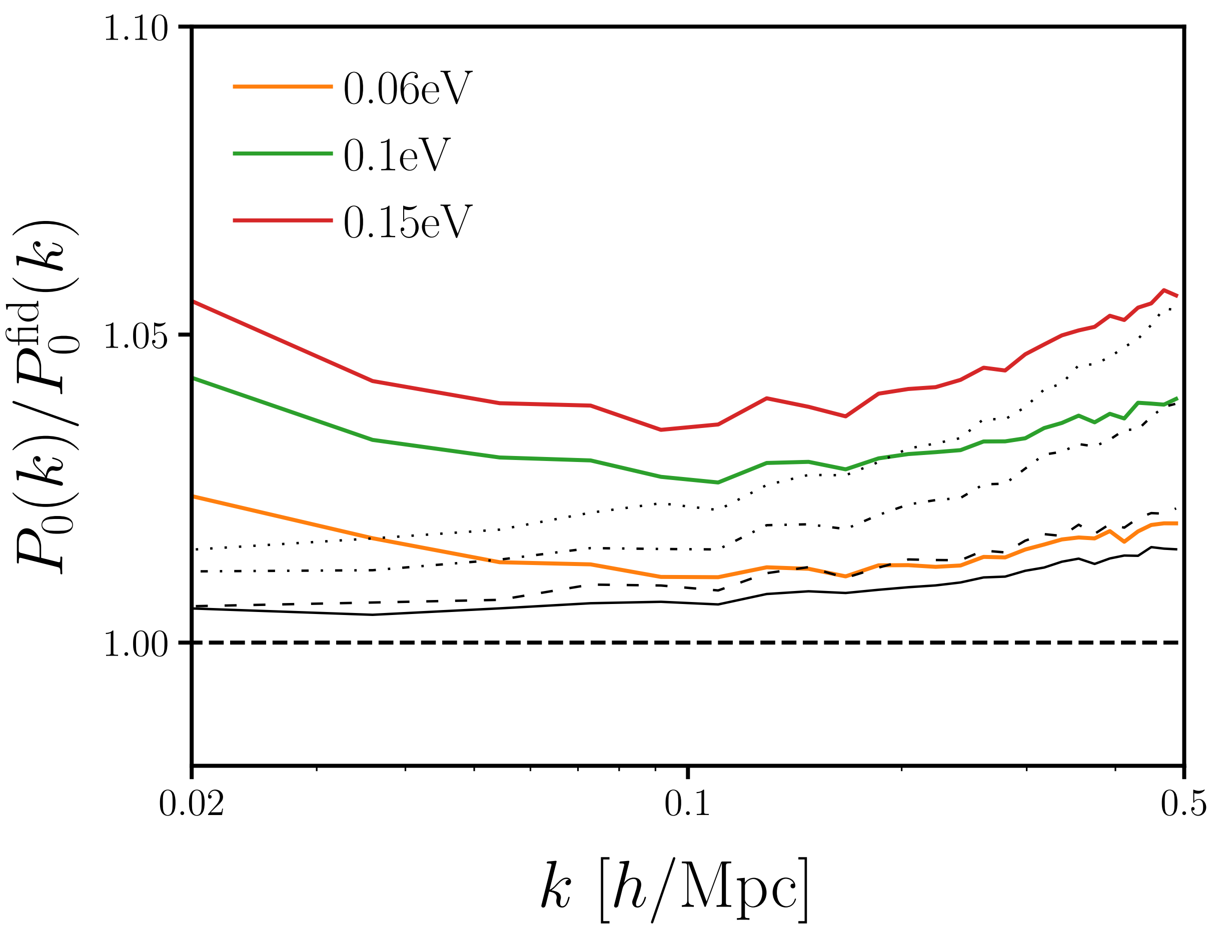


imprint of M_ν can be measured from **galaxy clustering**

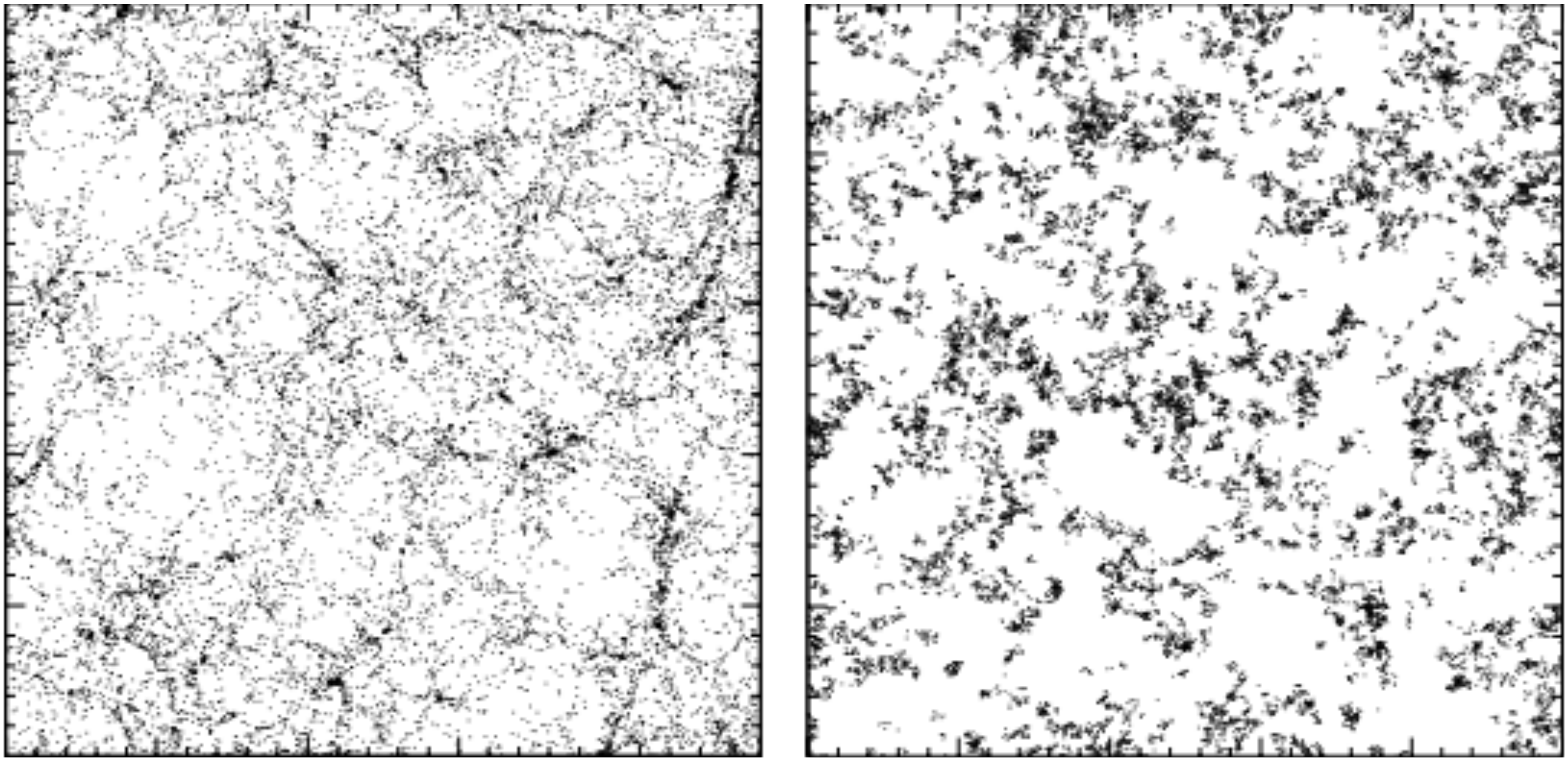
upcoming surveys: *DESI, PFS, Euclid, Roman*

credit: desi.lbl.gov

$M_\nu - \sigma_8$ **degeneracy is a major limitation** for the redshift-space *power spectrum*

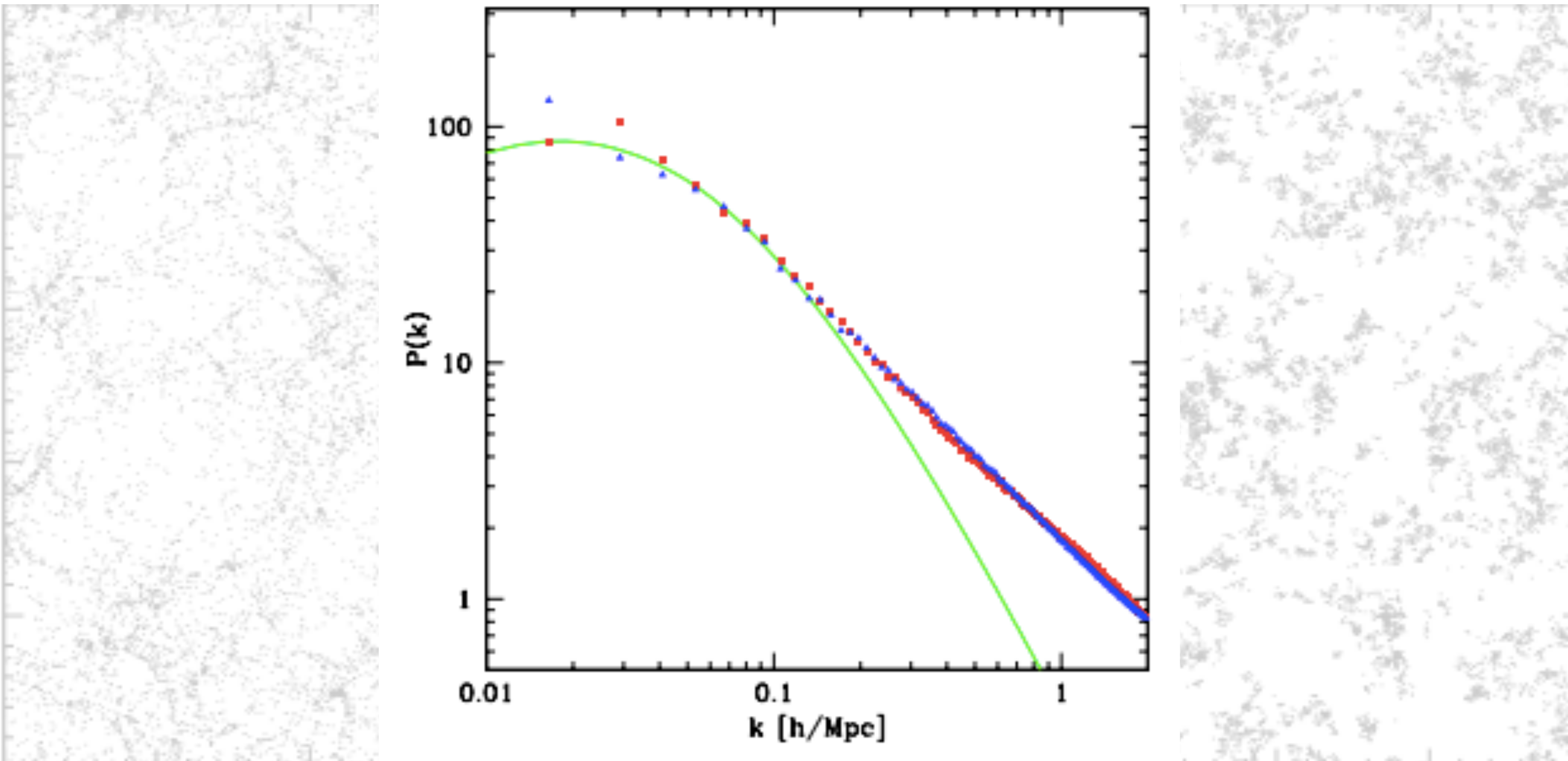


HADES simulation (Villaescusa-Navarro+2018) information in the *nonlinear regime* cascades from the power spectrum to **higher-order statistics** these two distributions have the ***same power spectrum***



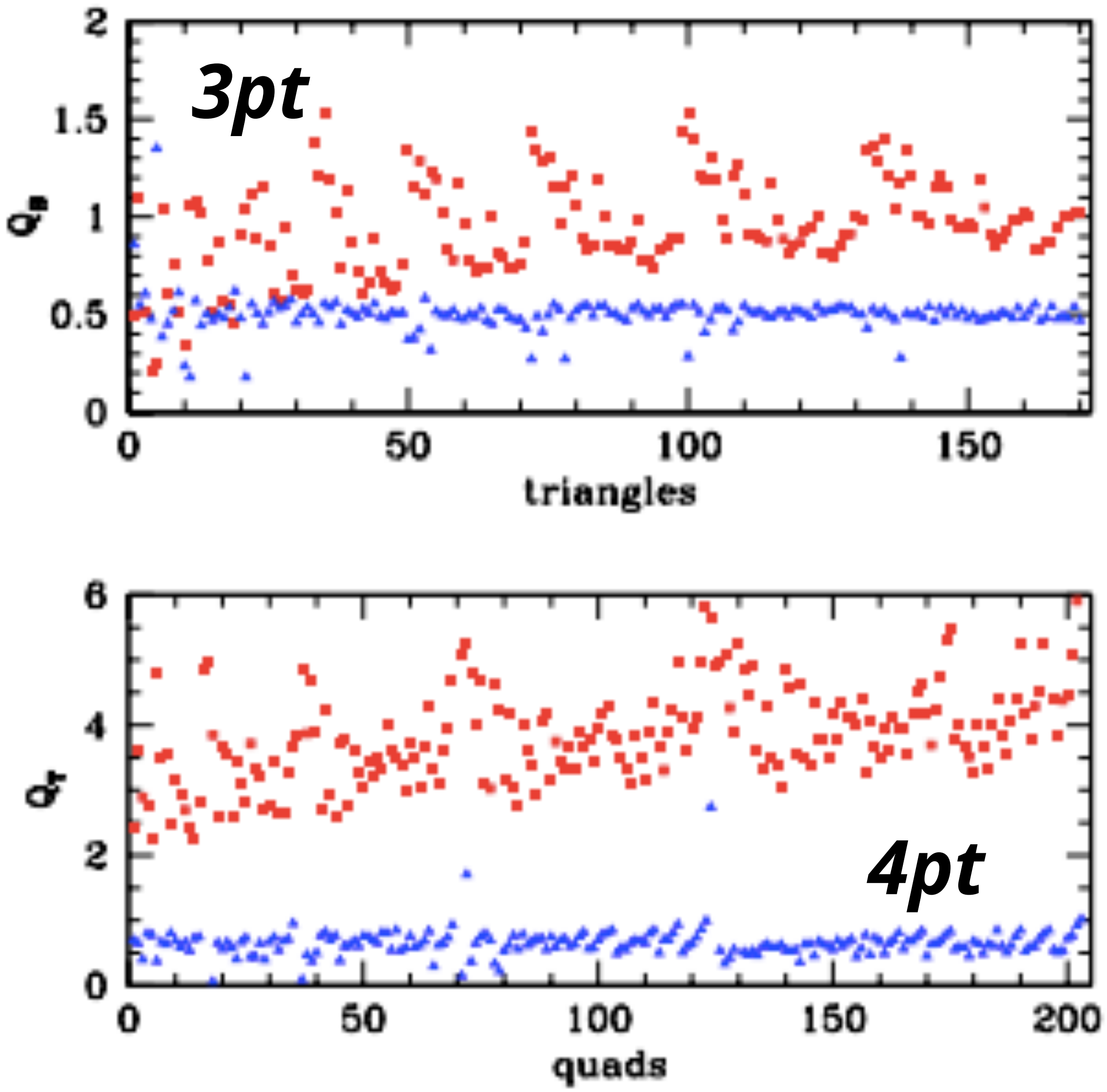
Sefusatti+(2005)

these two distributions have the ***same power spectrum***



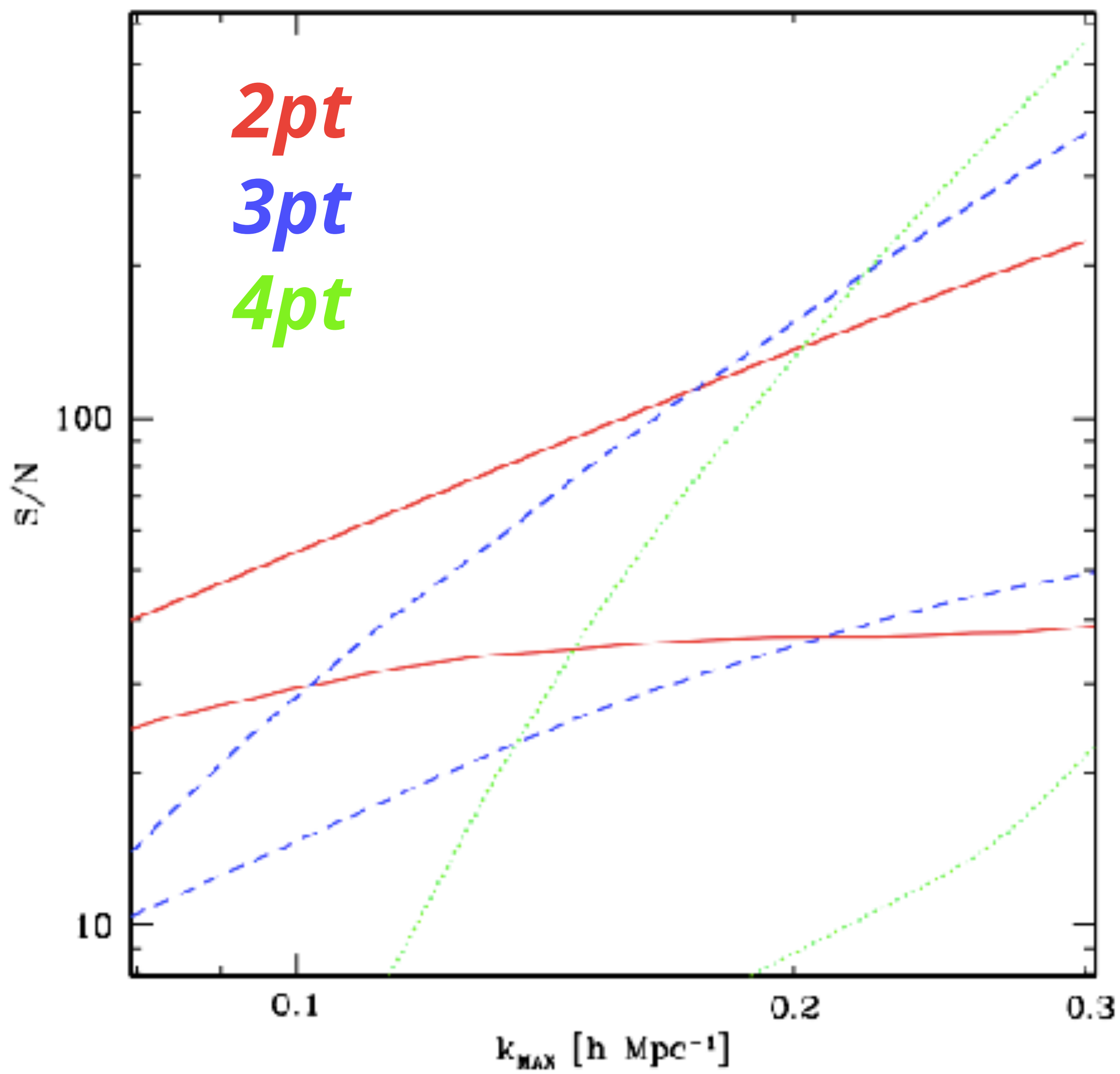
Sefusatti+(2005)

but very ***different higher-order statistics***



Sefusatti+(2005)

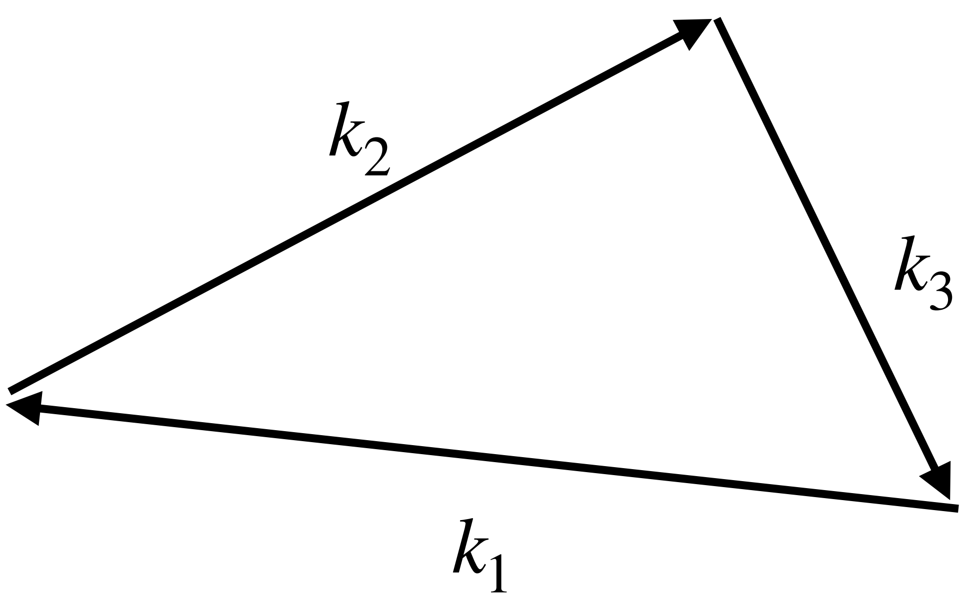
higher-order statistics have **higher signal-to-noise** on small scales



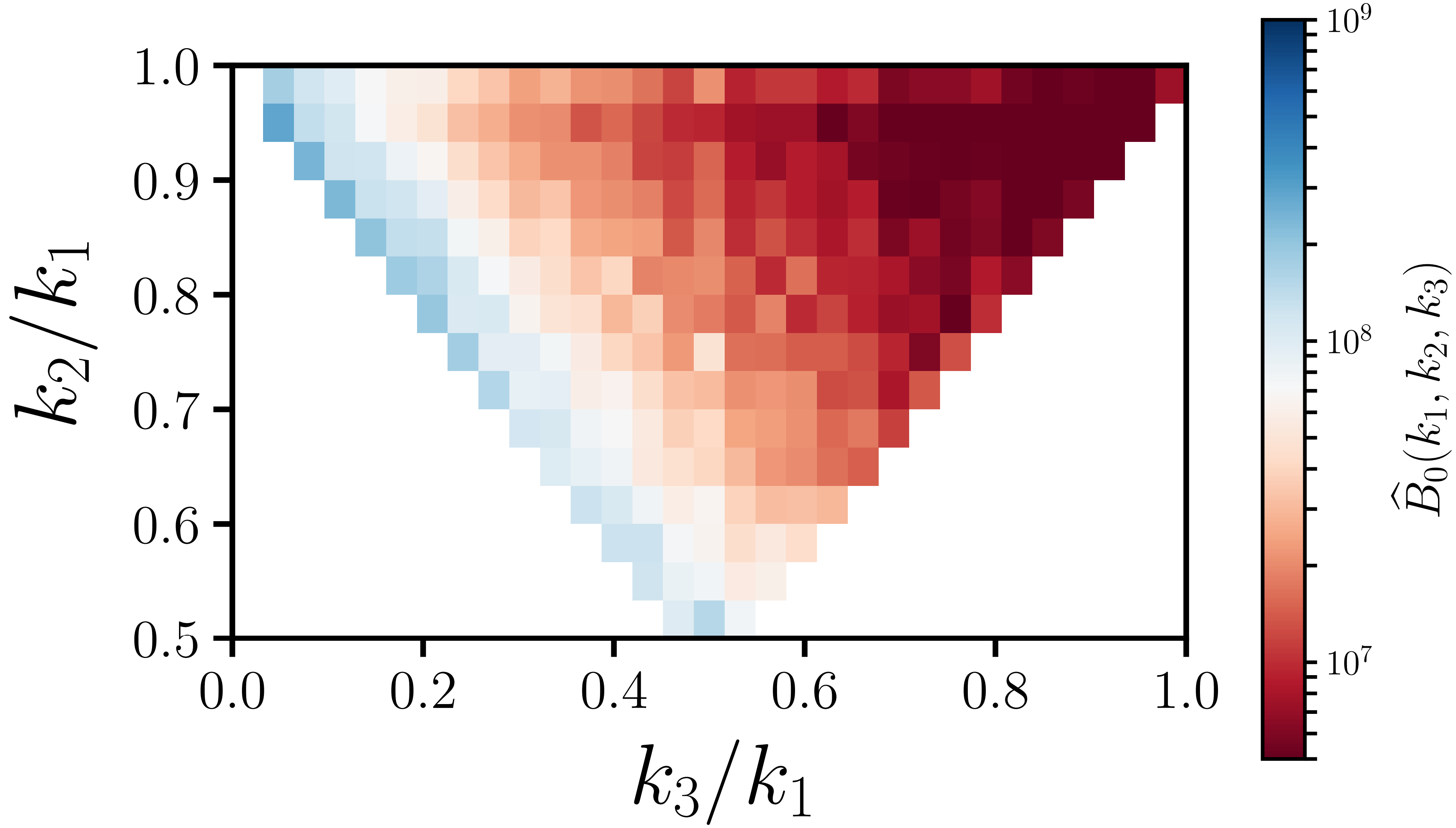
Sefusatti+(2005)

introduction to ***the bispectrum***

$$\langle\delta(\mathbf{k_1}),\delta(\mathbf{k_2}),\delta(\mathbf{k_3})\rangle = \delta_D(\mathbf{k_1},\mathbf{k_2},\mathbf{k_3})B(\mathbf{k_1},\mathbf{k_2},\mathbf{k_3})$$

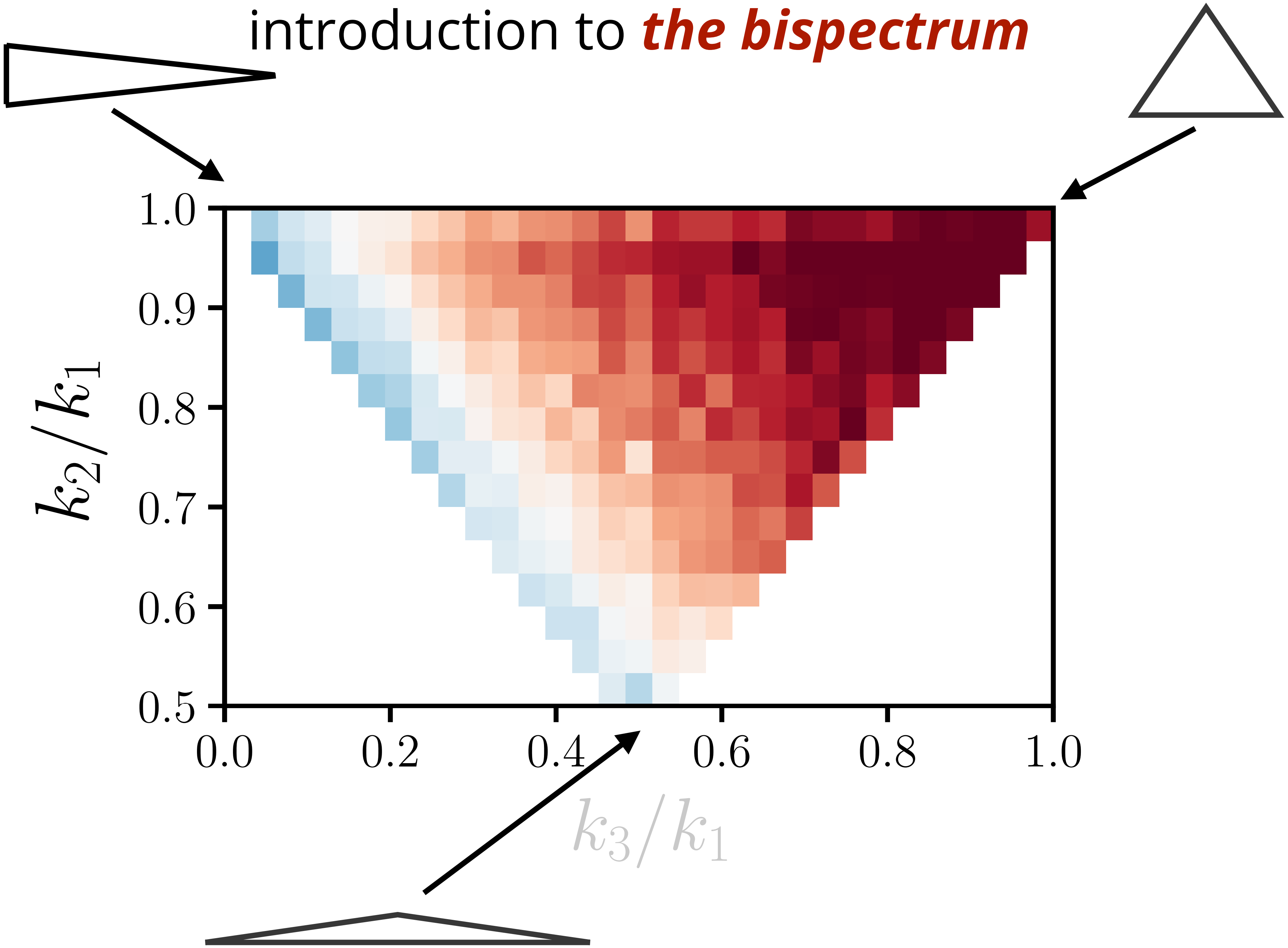


introduction to ***the bispectrum***



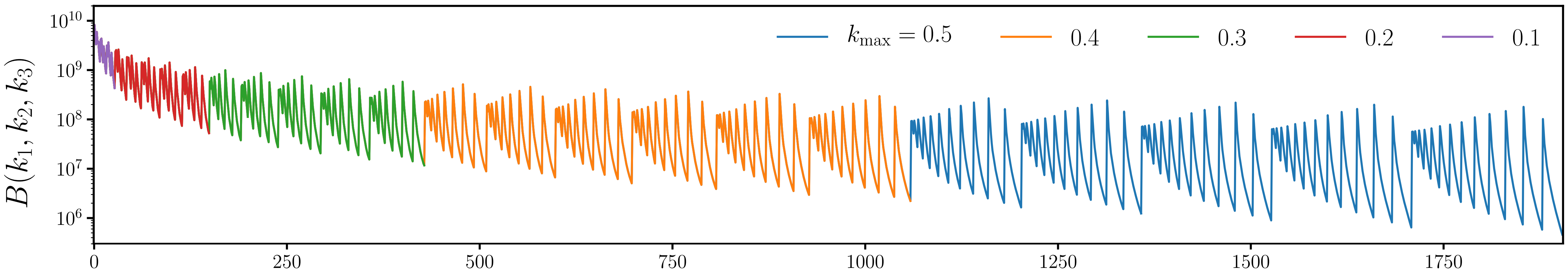
github.com/changhoonhahn/pySpectrum

introduction to ***the bispectrum***



github.com/changhoonhahn/pySpectrum

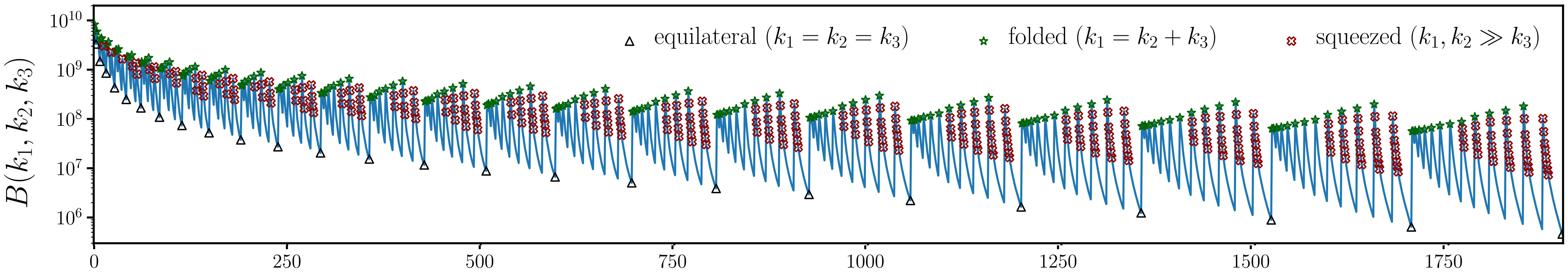
introduction to ***the bispectrum***



$k_1 \geq k_2 \geq k_3$

github.com/changhoonhahn/pySpectrum

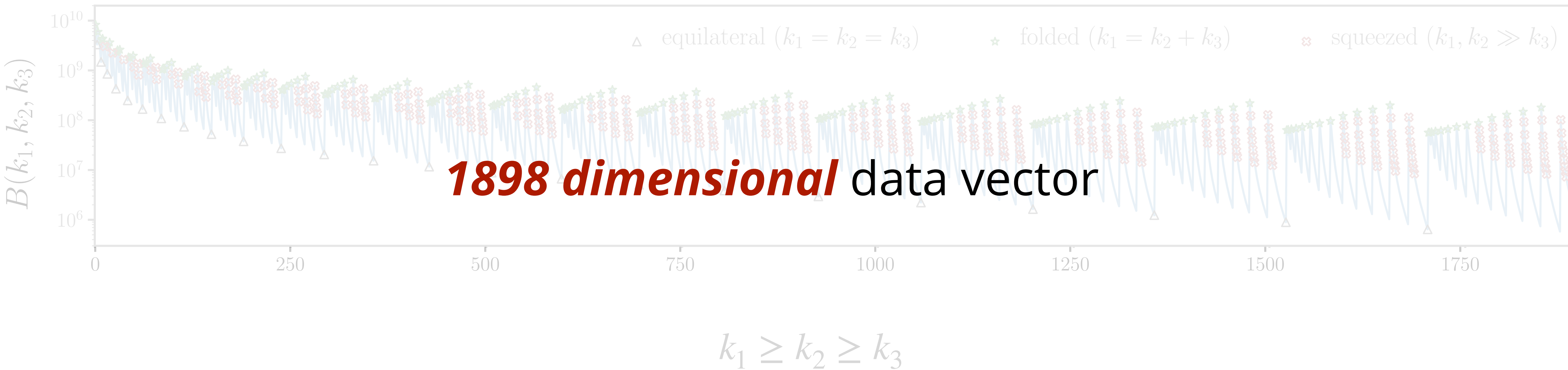
introduction to ***the bispectrum***



$k_1 \geq k_2 \geq k_3$

github.com/changhoonhahn/pySpectrum

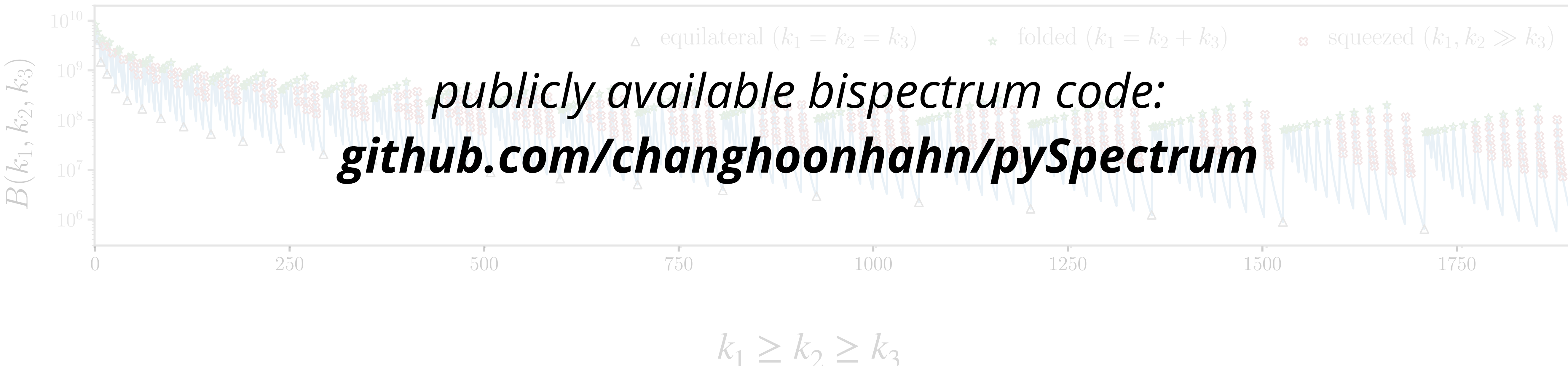
introduction to ***the bispectrum***



1898 dimensional data vector

github.com/changhoonhahn/pySpectrum

introduction to ***the bispectrum***



publicly available bispectrum code:

github.com/changhoonhahn/pySpectrum

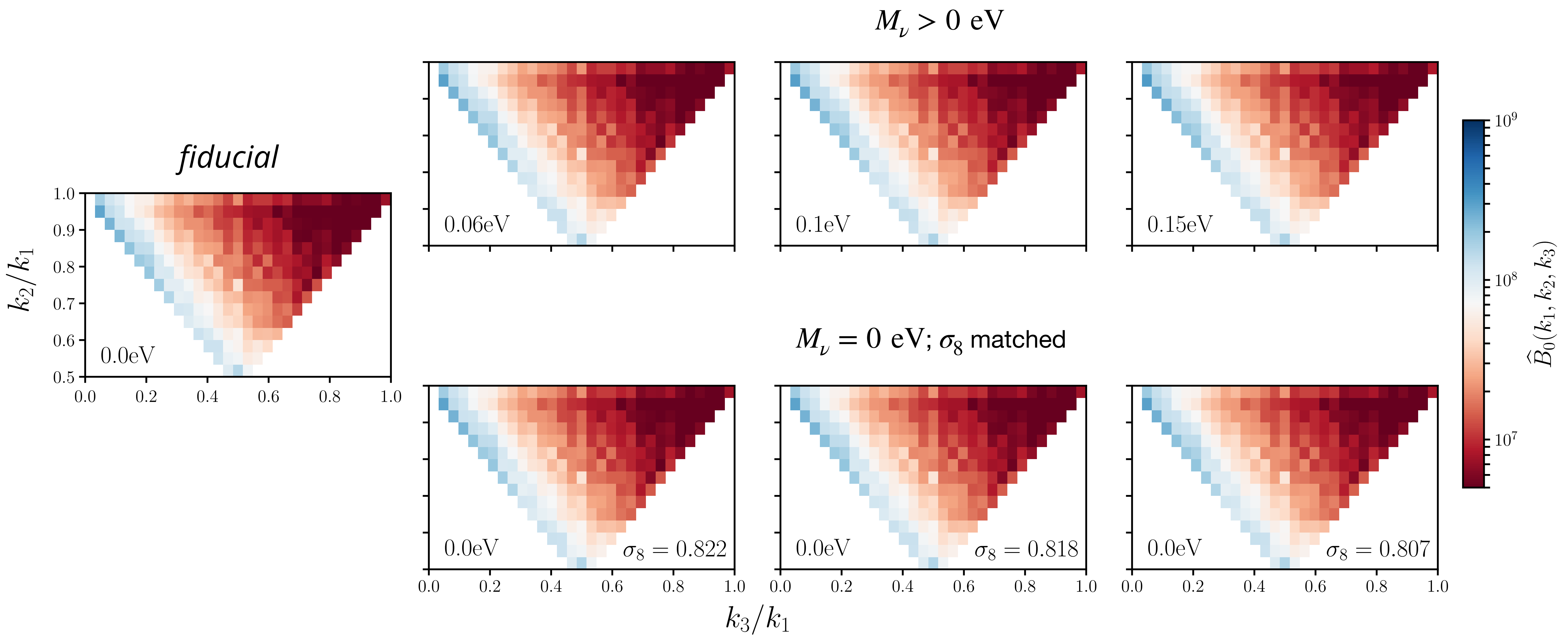
does ***the bispectrum improve*** M_ν ***constraints***?

can the bispectrum help *break* $M_\nu - \sigma_8$ *degeneracy?*

Name	M_ν	Ω_m	Ω_b	h	n_s	σ_8^m	σ_8^c	ICs	realizations
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.833	0.833	Zel'dovich	100
	0.06	0.3175	0.049	0.6711	0.9624	0.819	0.822	Zel'dovich	100
$M_\nu > 0$ eV	0.10	0.3175	0.049	0.6711	0.9624	0.809	0.815	Zel'dovich	100
	0.15	0.3175	0.049	0.6711	0.9624	0.798	0.806	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.822	0.822	Zel'dovich	100
$M_\nu = 0$ eV	0.0	0.3175	0.049	0.6711	0.9624	0.818	0.818	Zel'dovich	100
σ_8 matched	0.0	0.3175	0.049	0.6711	0.9624	0.807	0.807	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.798	0.798	Zel'dovich	100

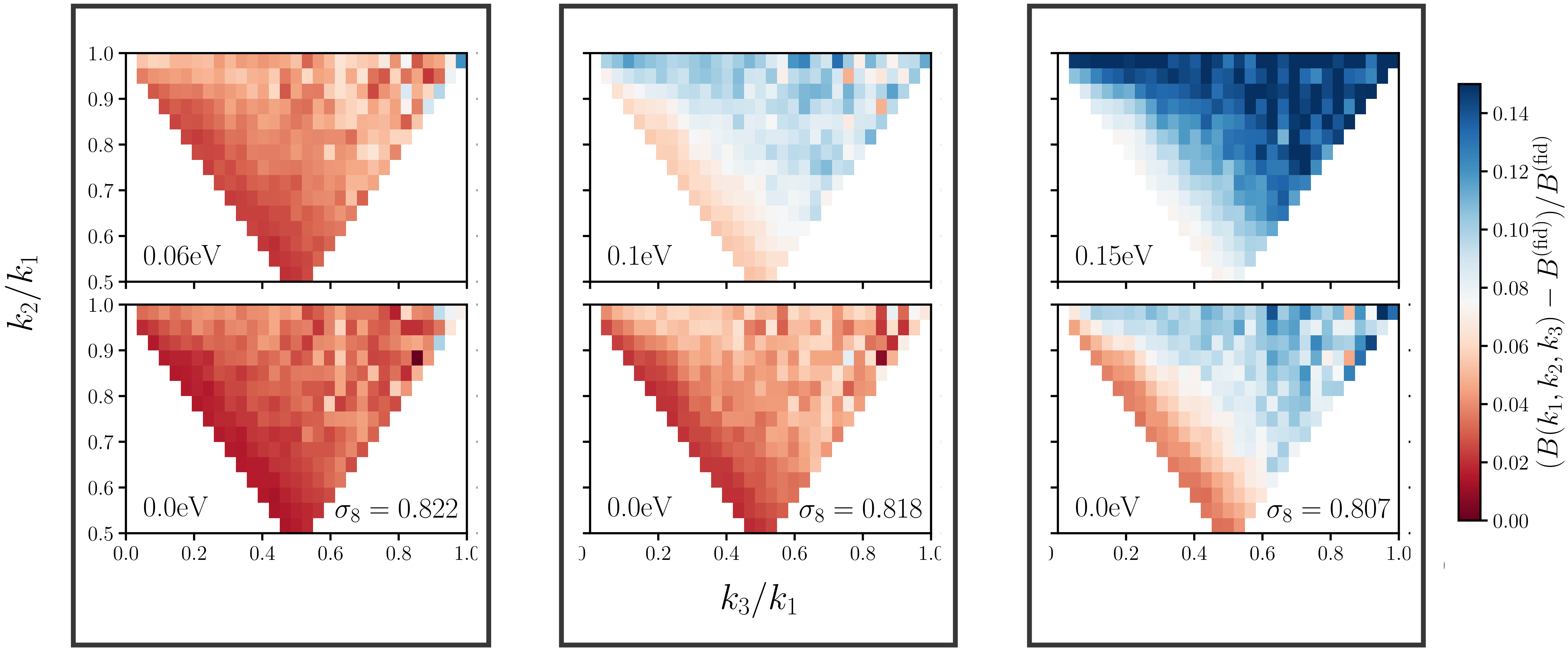
HADES simulation suite (Villaescusa-Navarro+2018)
franciscovillaescusa.github.io/hades.html

can the bispectrum help *break $M_\nu - \sigma_8$ degeneracy?*



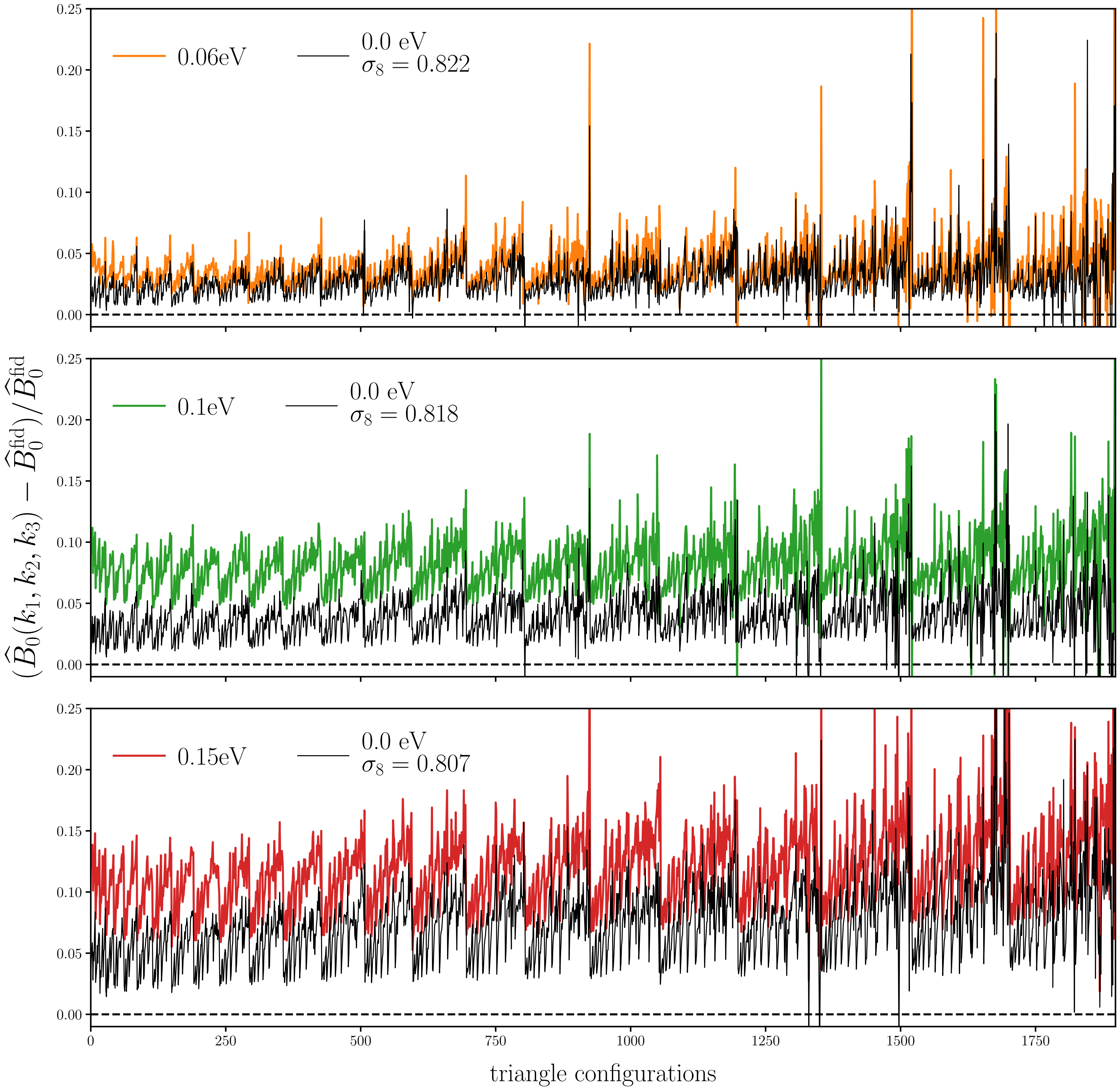
Hahn+(2020)

yes! M_ν imprint on the bispectrum is different than σ_8



Hahn+(2020)

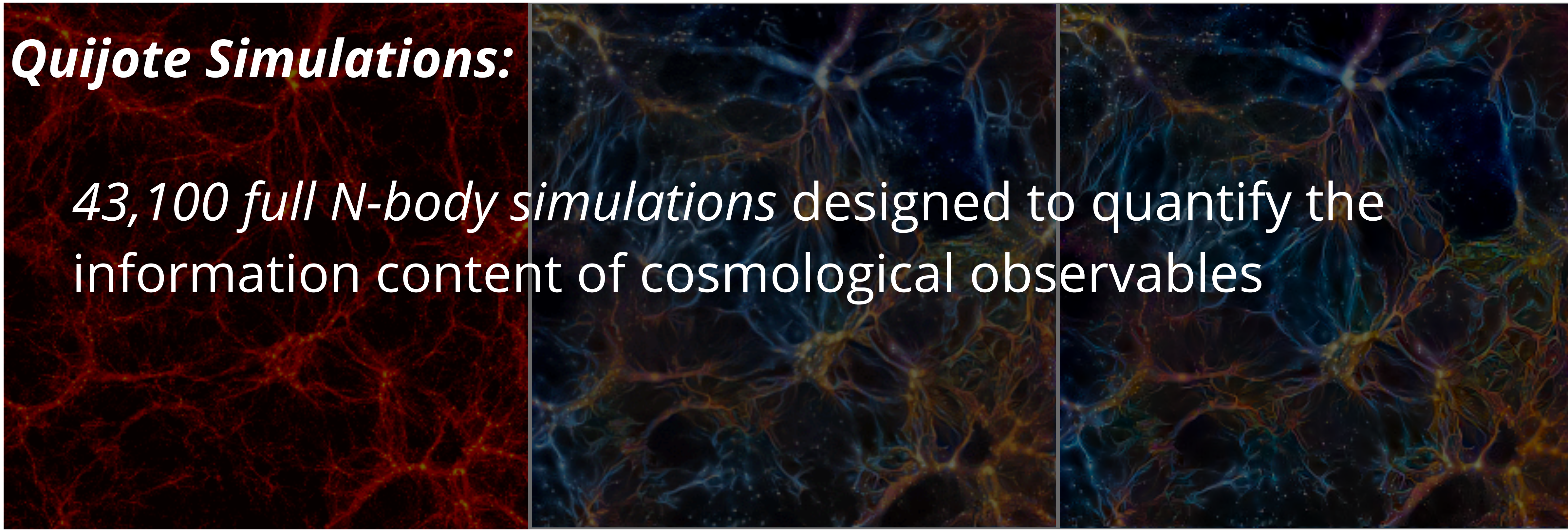
yes! M_ν imprint on the bispectrum is different than σ_8



Hahn+(2020)

does ***the bispectrum improve M_ν constraints***? **YES** does ***the bispectrum improve*** M_ν ***constraints***? **YES**

what's the **total information content** of *the bispectrum*?



Quijote Simulations:

43,100 full N-body simulations designed to quantify the information content of cosmological observables

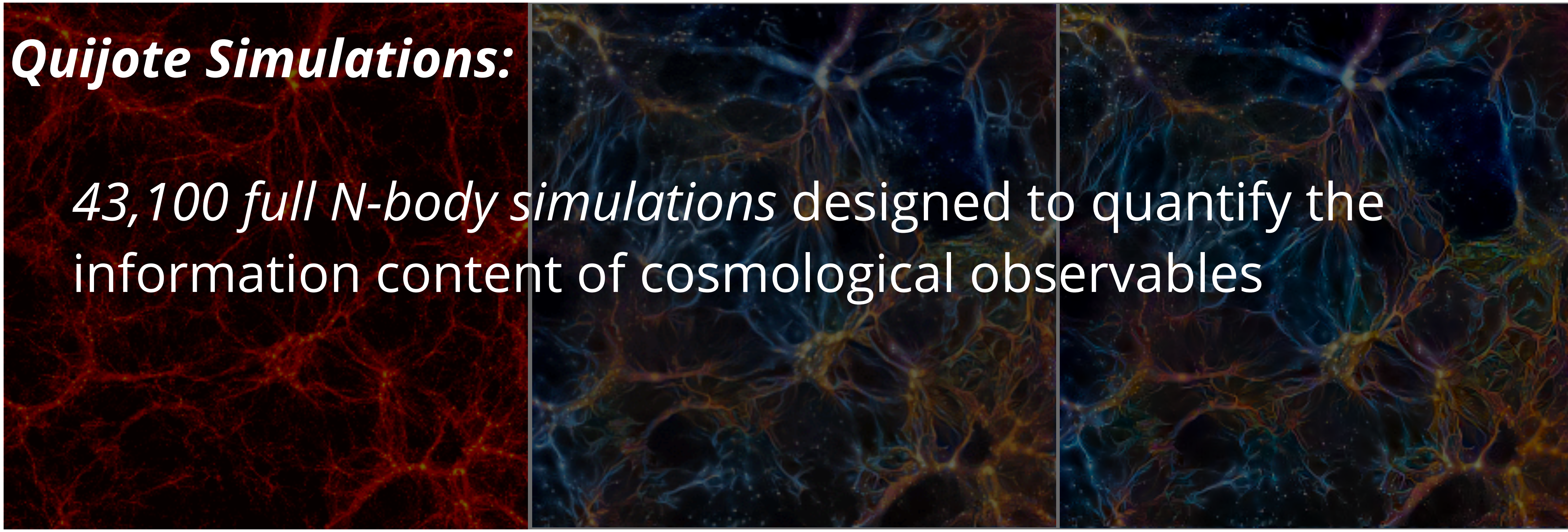
Features

- Simulations run with the TreePM code Gadget-III
- More than 35 Million CPU hours
- Boxes of 1 Gpc/h. Combined total volume of 43100 (Gpc/h)^3
- 17100 simulations for a fiducial Planck cosmology
- Between 500 and 1000 simulations/cosmology for 17 different cosmologies
- 11000 simulations in different latin-hypercubes
- More than 8.5 trillions of particles at a single redshift
- Billions of halos and voids identified
- Full snapshots at redshifts 0, 0.5, 1, 2, 3 and 127 (initial conditions)
- More than 200000 halo catalogues
- More than 200000 void catalogues
- More than 1 million power spectra
- More than 1 million bispectra
- More than 1 million correlation functions
- More than 1 million marked power spectra
- More than 1 million probability distribution functions
- More than 1 Petabyte of data publicly available

all publicly available

github.com/franciscovillaescusa/Quijote-simulations

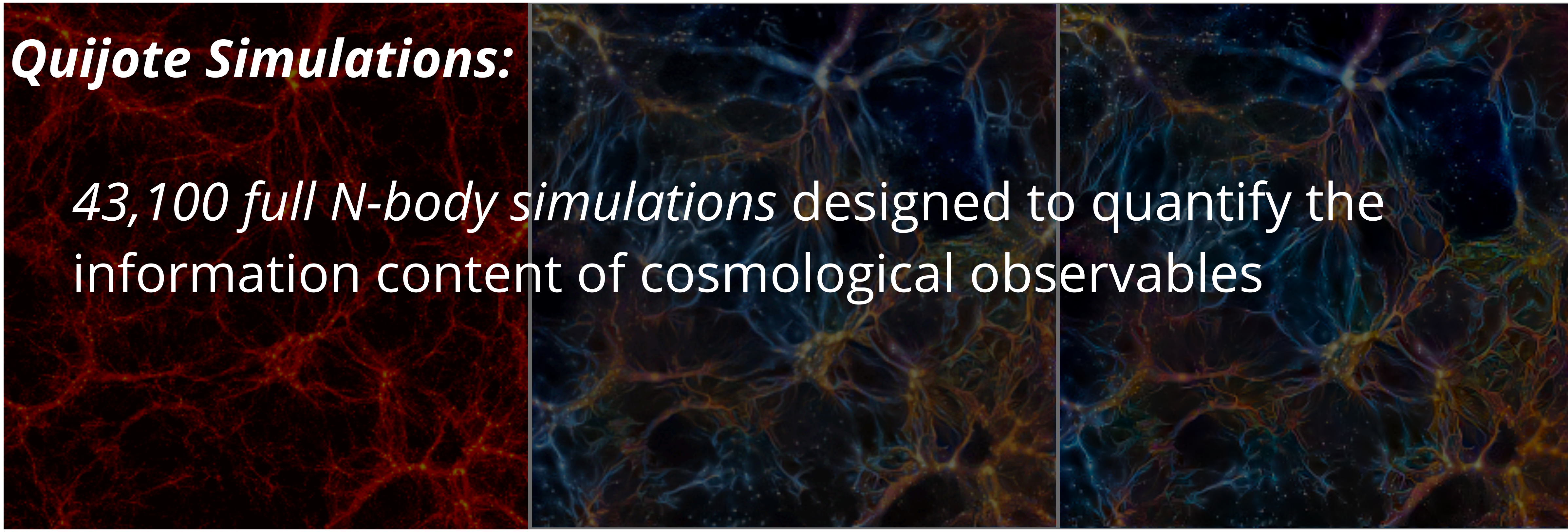
Villaesucsa-Navarro, CHH+(2019)



Quijote Simulations:

43,100 full N-body simulations designed to quantify the information content of cosmological observables

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$



Quijote Simulations:

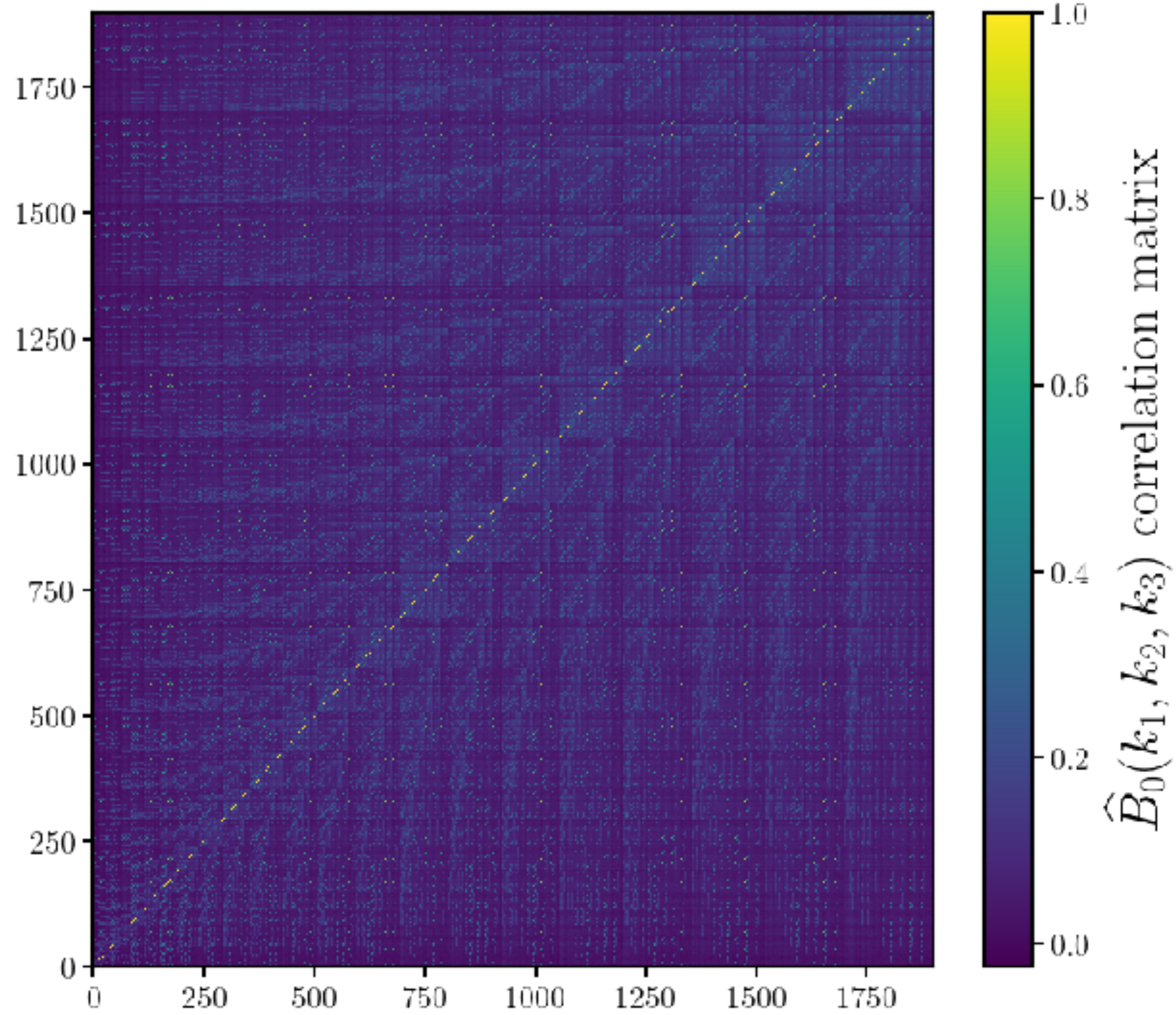
43,100 full N-body simulations designed to quantify the information content of cosmological observables

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\boxed{\mathbf{C}^{-1}}\left(\frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

difficult to accurately estimate the **1898x1898 covariance matrix**

previous work used fewer triangles, PT estimates, ignored non-Gaussian term, etc

$\mathbf{C}$ estimated using ***15,000 N-body simulations***!



Hahn+(2020)

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \boxed{\frac{\partial B}{\partial \theta_j}} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

Hahn+(2020)

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \boxed{\frac{\partial B}{\partial \theta_j}} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

fiducial $M_\nu = 0.0\mathrm{eV}$ $\Omega_m = 0.3175$ $\Omega_b = 0.049$ $h = 0.6711$ $n_s = 0.9624$ $\sigma_8 = 0.834$

Hahn+(2020)

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \boxed{\frac{\partial B}{\partial \theta_j}} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

$\boxed{\Omega_m^+ = 0.3275}$

fiducial $M_\nu = 0.0\mathrm{eV}$ $\Omega_m = 0.3175$ $\Omega_b = 0.049$ $h = 0.6711$ $n_s = 0.9624$ $\sigma_8 = 0.834$

$\boxed{\Omega_m^- = 0.3075}$

Hahn+(2020)

$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \boxed{\frac{\partial B}{\partial \theta_j}} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

	$M_\nu^{+++} = 0.4$					
	$M_\nu^{++} = 0.2$					
	$M_\nu^{+} = 0.1$	$\Omega_m^+ = 0.3275$	$\Omega_b^+ = 0.051$	$h^+ = 0.6911$	$n_s^+ = 0.9824$	$\sigma_8^+ = 0.849$
fiducial	$M_\nu = 0.0\mathrm{eV}$	$\Omega_m = 0.3175$	$\Omega_b = 0.049$	$h = 0.6711$	$n_s = 0.9624$	$\sigma_8 = 0.834$
		$\Omega_m^- = 0.3075$	$\Omega_b^- = 0.047$	$h^- = 0.6511$	$n_s^- = 0.9424$	$\sigma_8^- = 0.819$

$$\frac{\partial B}{\partial \theta} \approx \frac{B(\theta^+) - B(\theta^-)}{\theta^+ - \theta^-}$$

Hahn+(2020)

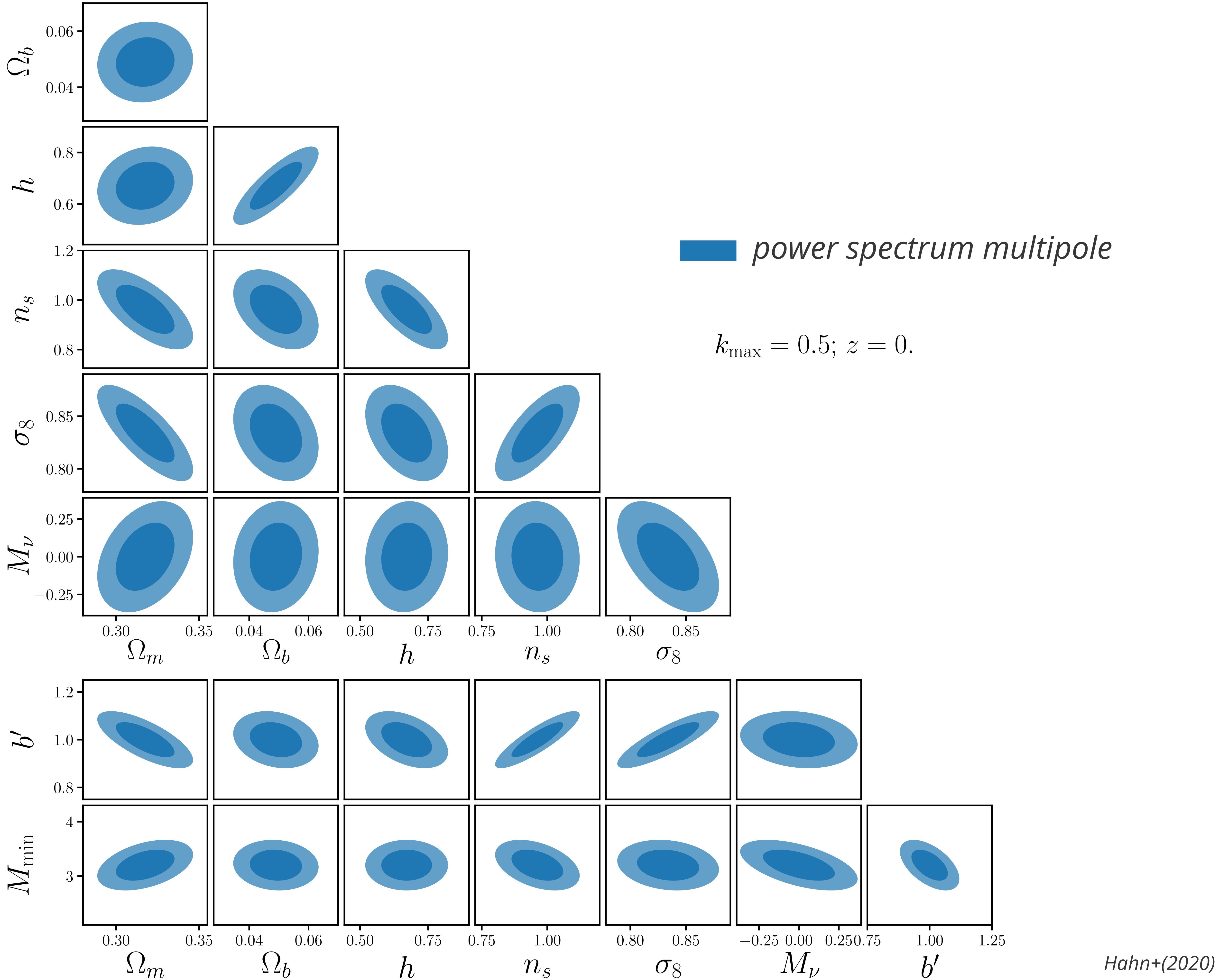
$$F_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\left(\frac{\partial B}{\partial \theta_i}^T \boxed{\frac{\partial B}{\partial \theta_j}} + \frac{\partial B}{\partial \theta_i}^T \frac{\partial B}{\partial \theta_j}\right)\right]$$

	$M_\nu^{+++} = 0.4$					
	$M_\nu^{++} = 0.2$					
	$M_\nu^{+} = 0.1$	$\Omega_m^{+} = 0.3275$	$\Omega_b^{+} = 0.051$	$h^{+} = 0.6911$	$n_s^{+} = 0.9824$	$\sigma_8^{+} = 0.849$
fiducial	$M_\nu = 0.0\mathrm{eV}$	$\Omega_m = 0.3175$	$\Omega_b = 0.049$	$h = 0.6711$	$n_s = 0.9624$	$\sigma_8 = 0.834$
		$\Omega_m^{-} = 0.3075$	$\Omega_b^{-} = 0.047$	$h^{-} = 0.6511$	$n_s^{-} = 0.9424$	$\sigma_8^{-} = 0.819$

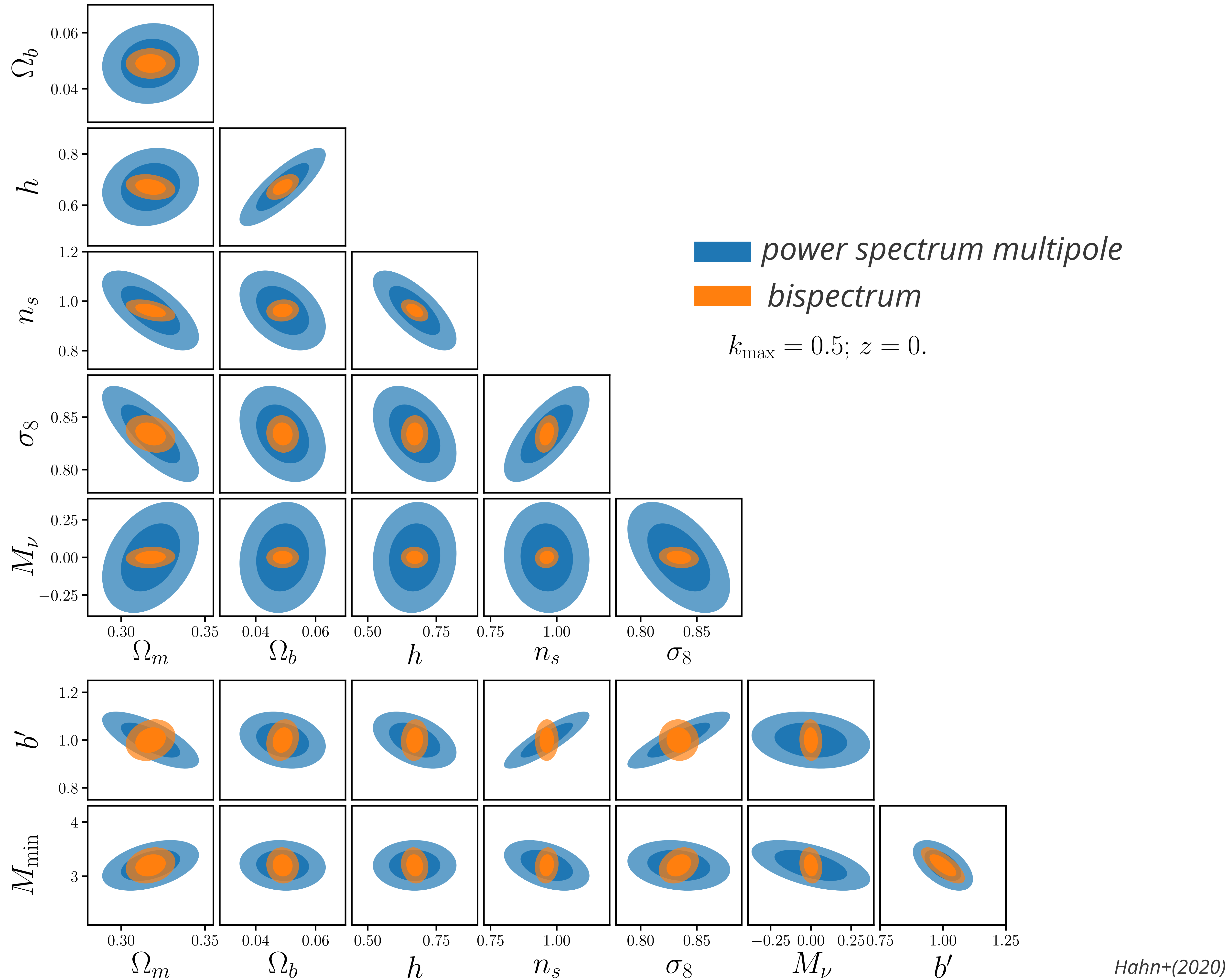
$$\frac{\partial B}{\partial \theta} \approx \frac{B(\theta^+) - B(\theta^-)}{\theta^+ - \theta^-}$$

each box is a different cosmology with ***500 N-body*** simulations

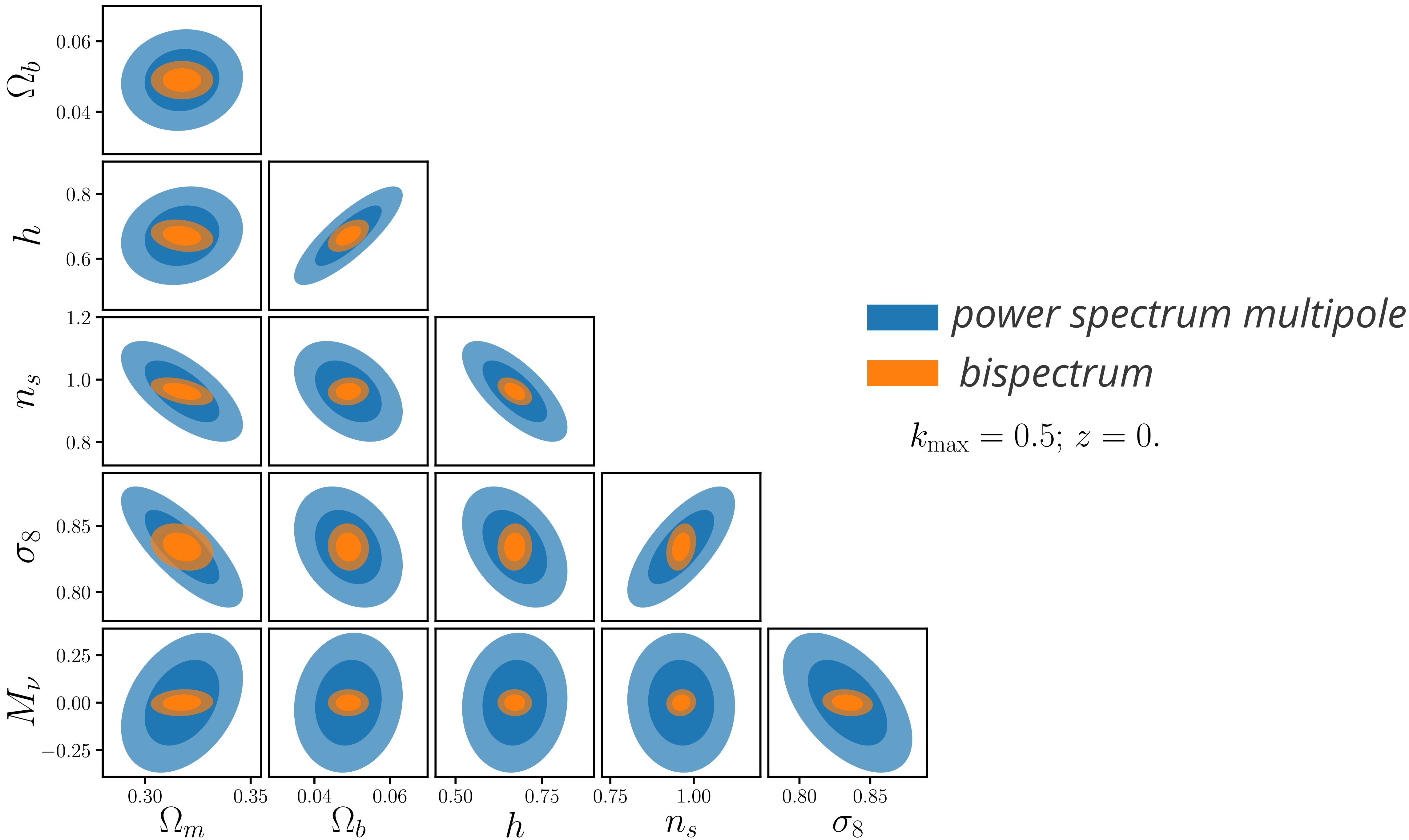
Hahn+(2020)



power spectrum multipole
$k_{\rm max} = 0.5;\ z = 0.$
Ω_b
h
n_s
σ_8
M_ν
b'
$M_{\rm min}$
Ω_m
Hahn+(2020)

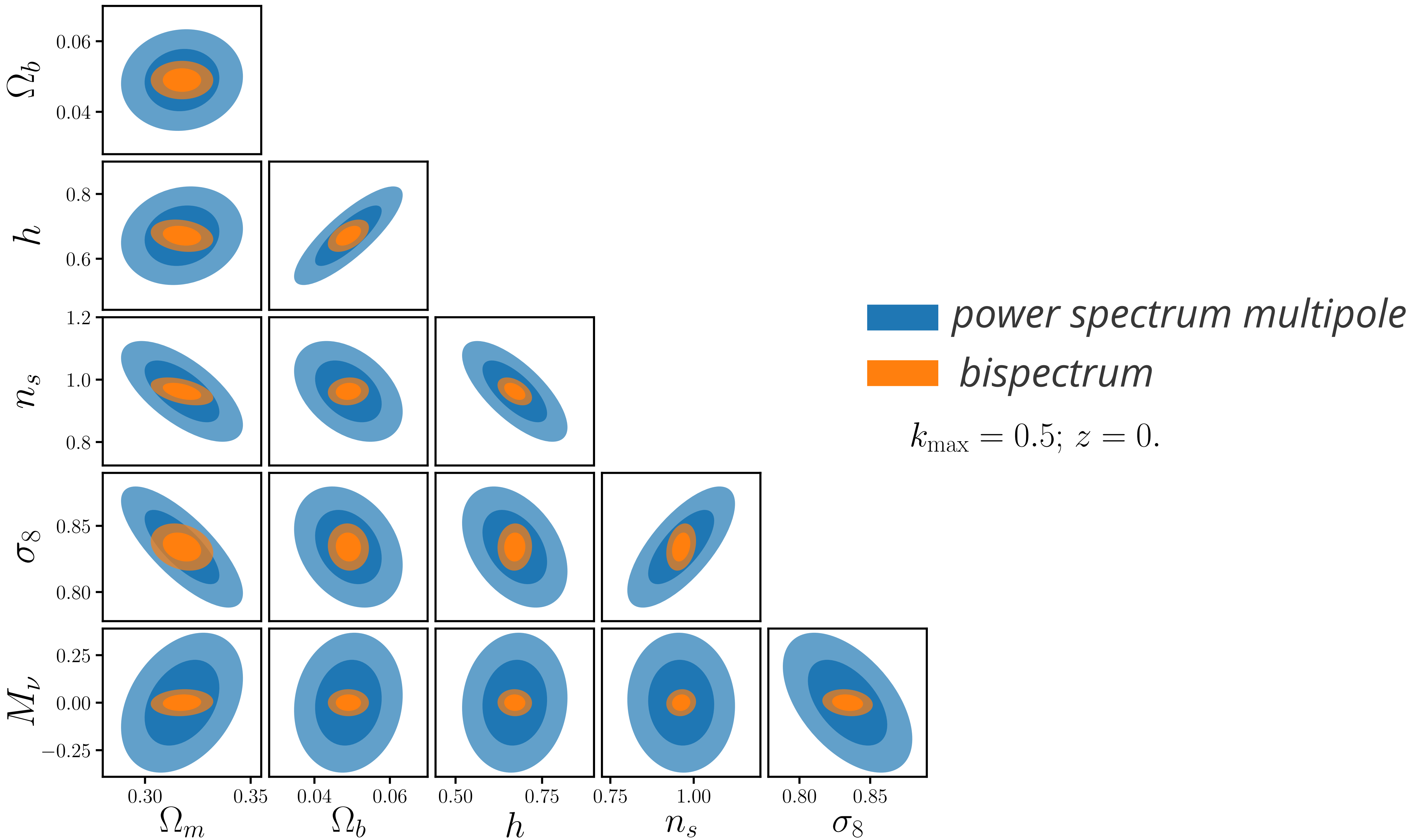


power spectrum multipole
bispectrum
$k_{\max} = 0.5;\ z = 0.$
Ω_b
h
n_s
σ_8
M_ν
b'
$M_{\min}$
Ω_m
0.06
0.04
0.8
0.6
1.2
1.0
0.8
0.85
0.80
0.25
0.00
−0.25
1.2
1.0
0.8
4
3
0.30
0.35
0.04
0.06
0.50
0.75
0.75
1.00
0.80
0.85
−0.25
0.00
0.25
0.75
1.00
1.25
Hahn+(2020)



$\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$ constraints $\{1.9,\ 2.6,\ 3.1,\ 3.6,\ 2.6\times\}$ **tighter**

Hahn+(2020)



power spectrum
$\sigma_{M_\nu} = \pm 0.297$

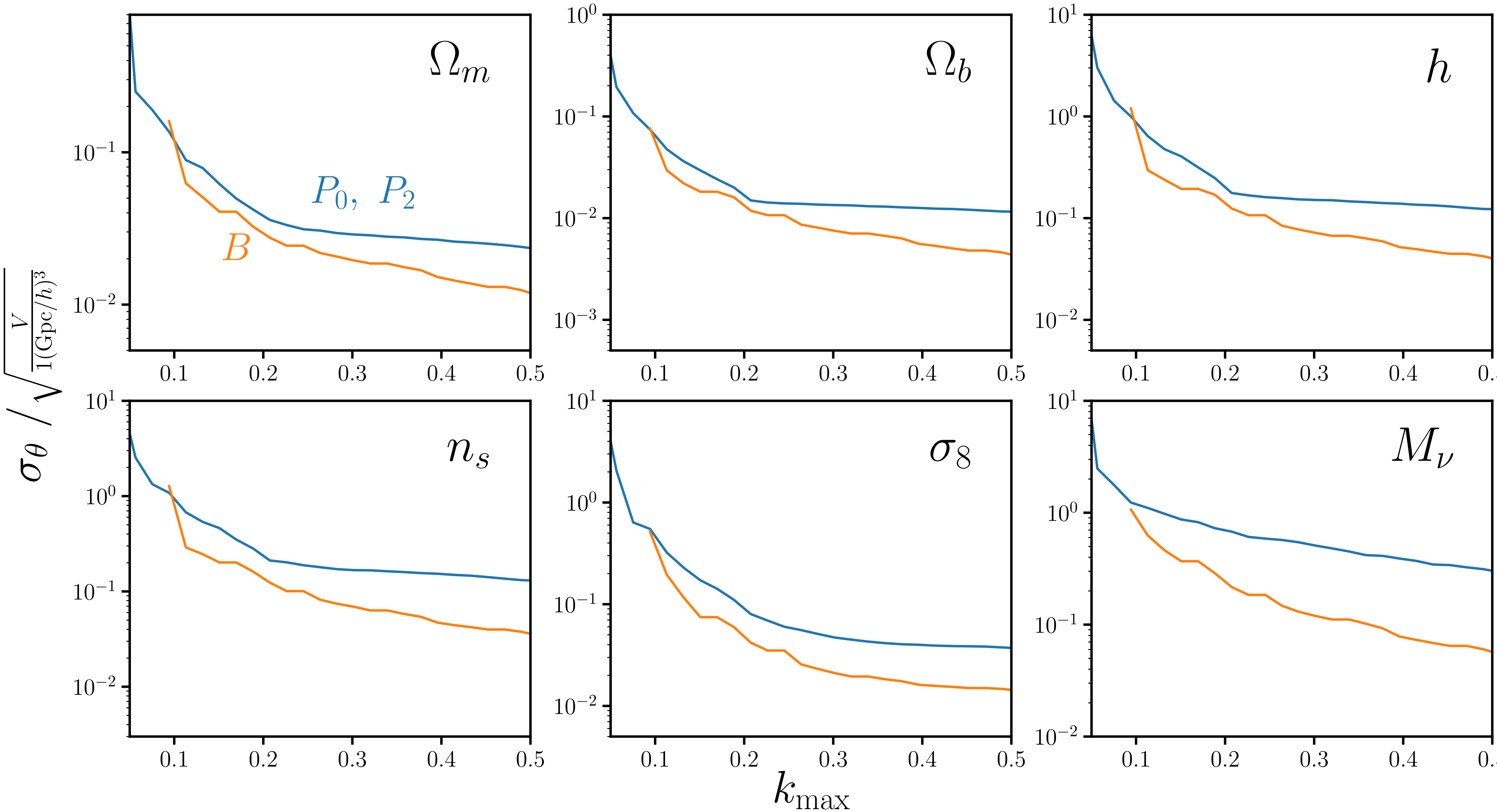
→

bispectrum
$\sigma_{M_\nu} = \pm 0.057$

>5x tighter M_ν constraints

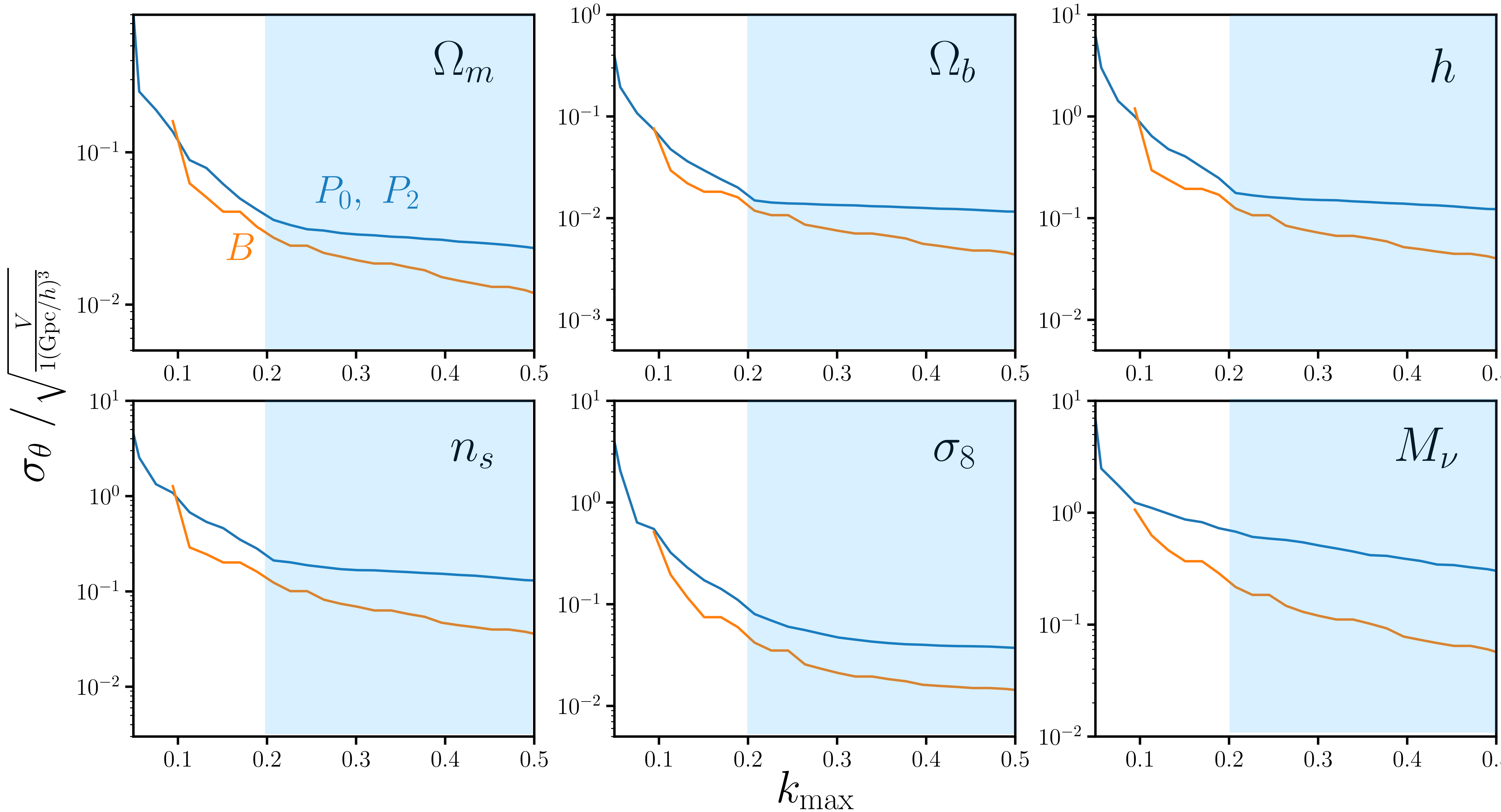
Hahn+(2020)

bispectrum improves parameter constraints **even at lower** $k_{\max}$



Hahn+(2020)

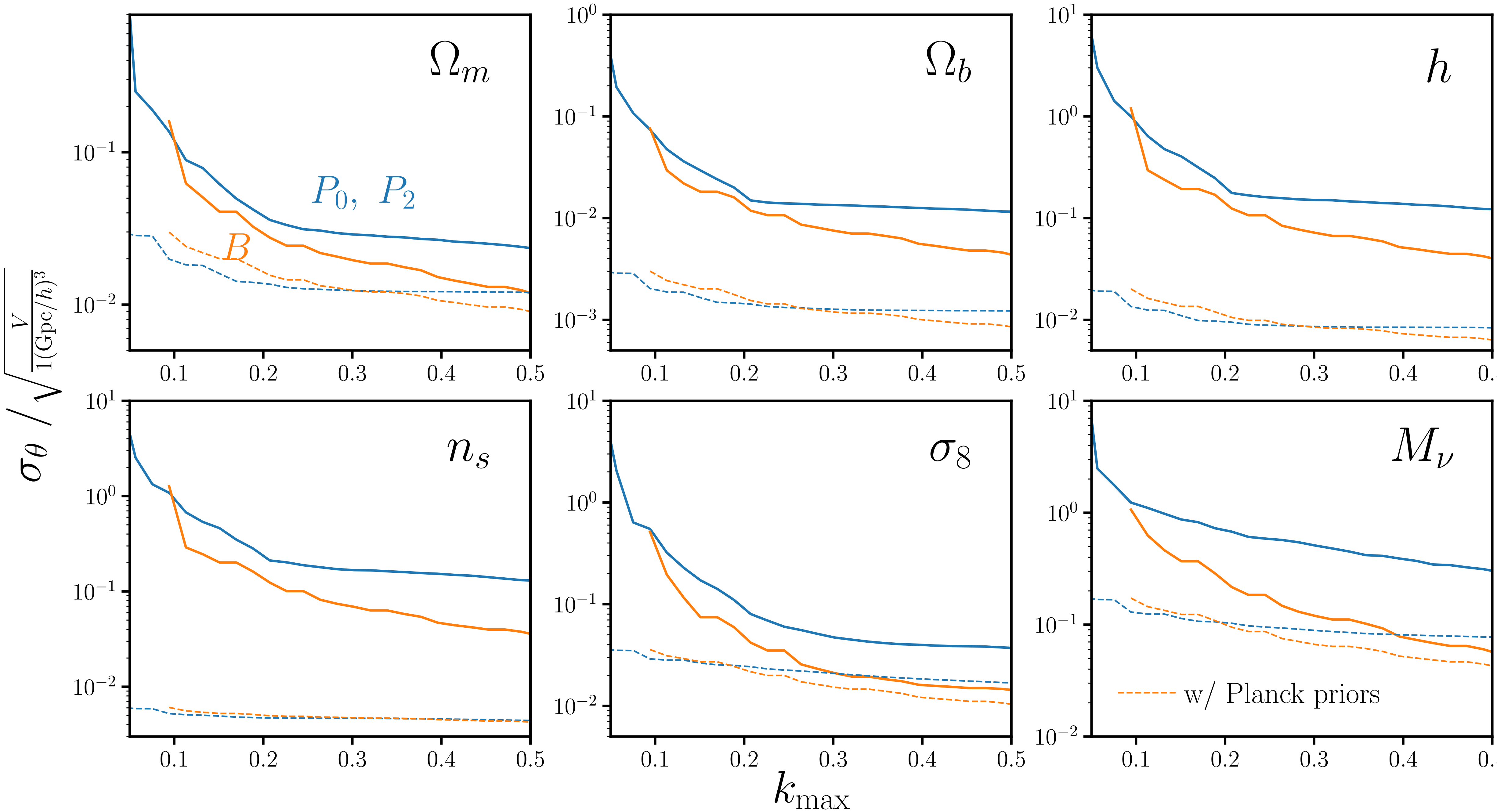
bispectrum improves parameter constraints **even at lower** $k_{\rm max}$



information content of P_ℓ *saturates beyond* $k_{\rm max} > 0.2\ h/{\rm Mpc}$

Hahn+(2020)

bispectrum improves parameter constraints **even at lower** $k_{\max}$



and with ***Planck priors***

Hahn+(2020)

the redshift-space bispectrum breaks parameter degeneracies and ***significantly improves*** parameter constraints

$1.9,\ 2.6,\ 3.1,\ 3.6,\ 2.6\times$ *tighter than* P_ℓ *for* $\Omega_m, \Omega_b, h, n_s, \sigma_8$

$5\times$ *tighter* M_ν *constraints*

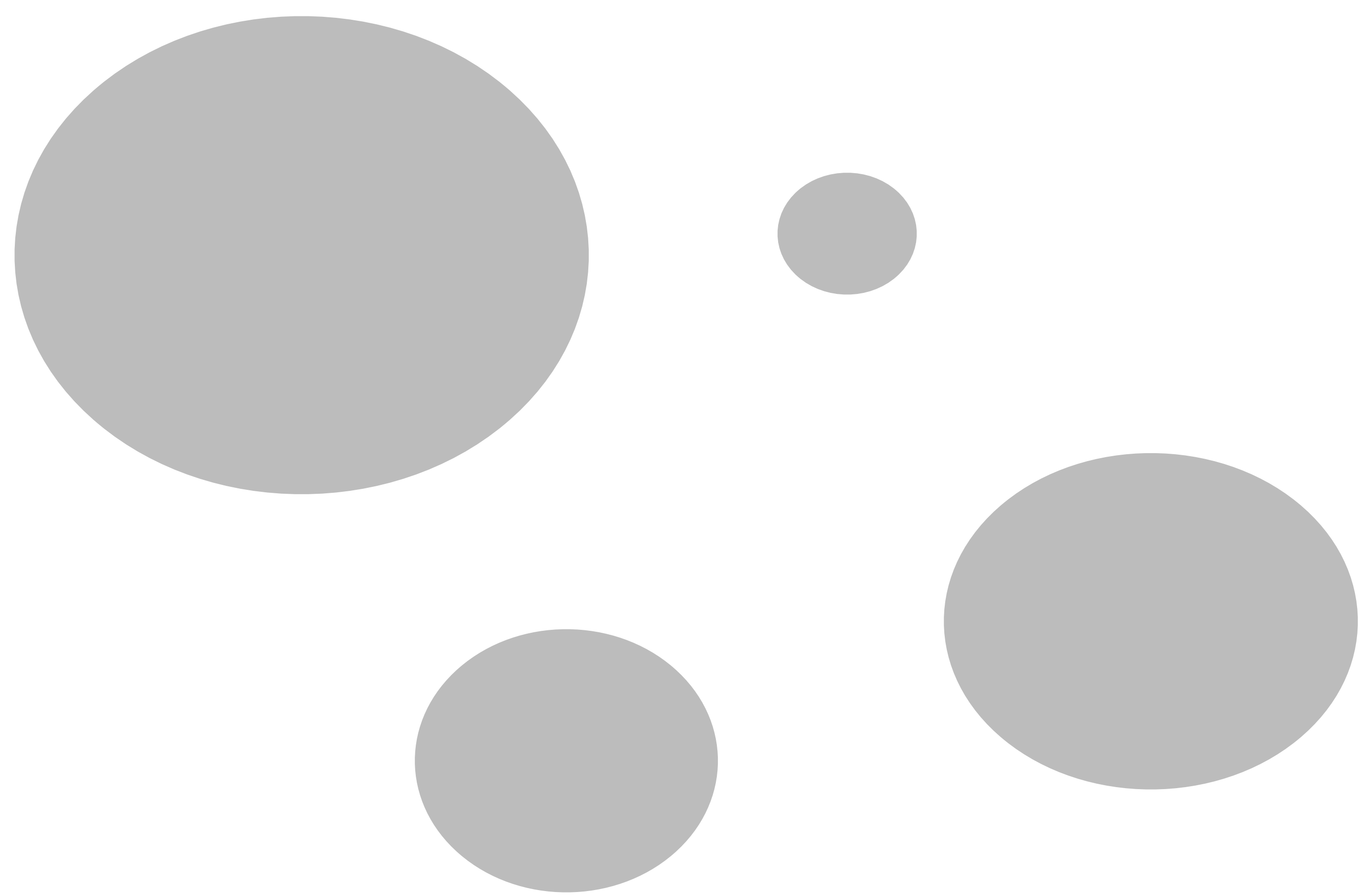
halo
*
the redshift-space bispectrum breaks parameter degeneracies and ***significantly improves*** parameter constraints

$1.9,\ 2.6,\ 3.1,\ 3.6,\ 2.6\times$ *tighter than* P_ℓ *for* $\Omega_m, \Omega_b, h, n_s, \sigma_8$

$5\times$ *tighter* M_ν *constraints* *marginalizing over* $b_1, b_2, \gamma_2, M_{\rm min}, A_{\rm SN}, B_{\rm SN}$ **does not impact** the improvement *marginalizing over* $b_1, b_2, \gamma_2, M_{\mathrm{min}}, A_{\mathrm{SN}}, B_{\mathrm{SN}}$ **does not impact** the improvement but we want to include a **more complete galaxy bias** model

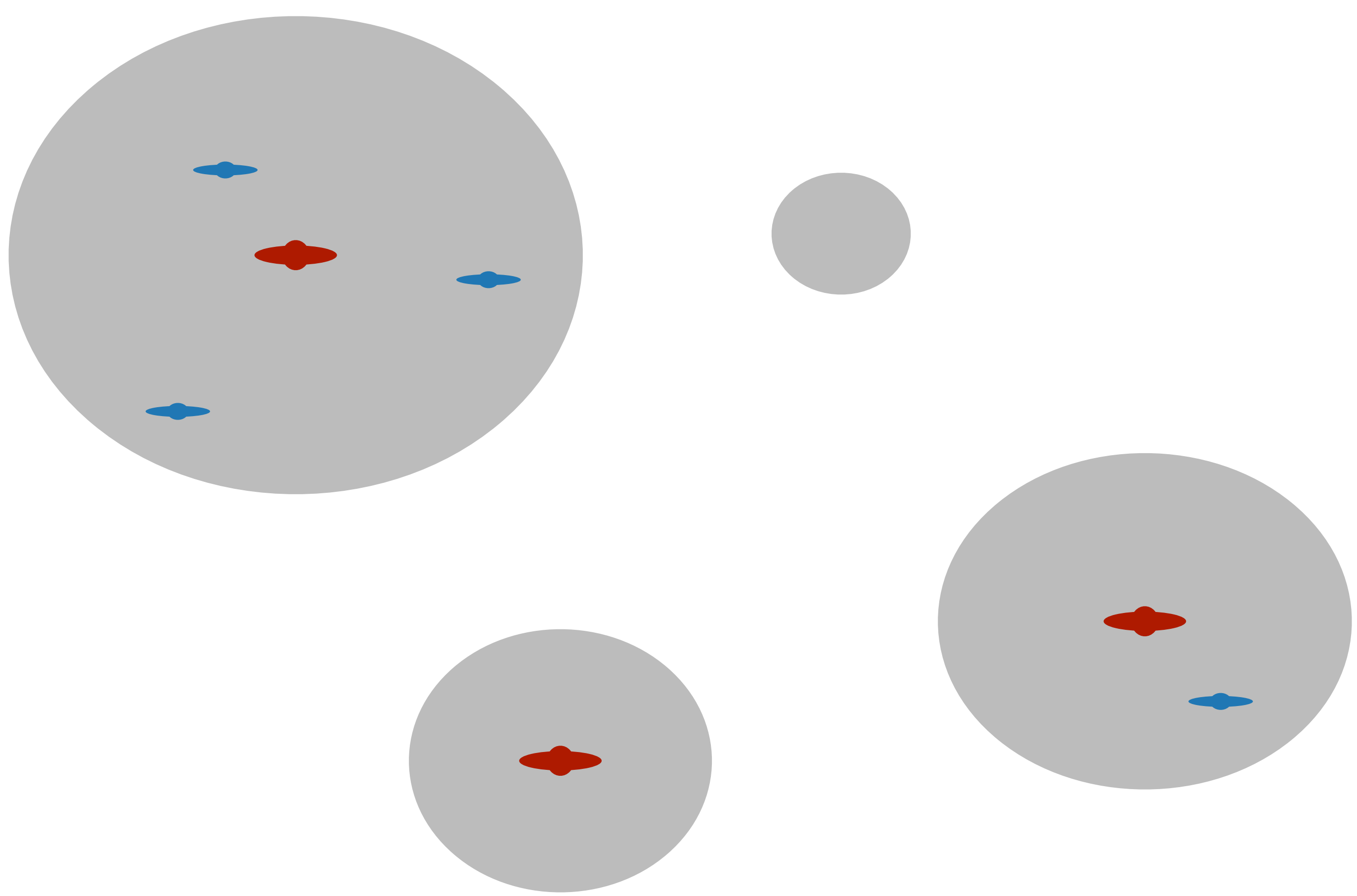
Halo Occupation Distribution (HOD) models populate halos from the *probability of a given halo hosting N galaxies*

$$N_{\mathrm{gal}} \sim P(N \,|\, M_h)$$



Halo Occupation Distribution (HOD) models populate halos from the *probability of a given halo hosting N galaxies*

$$N_{\rm gal} \sim P(N \,|\, M_h)$$



Halo Occupation Distribution (HOD) models populate halos from the *probability of a given halo hosting N galaxies*

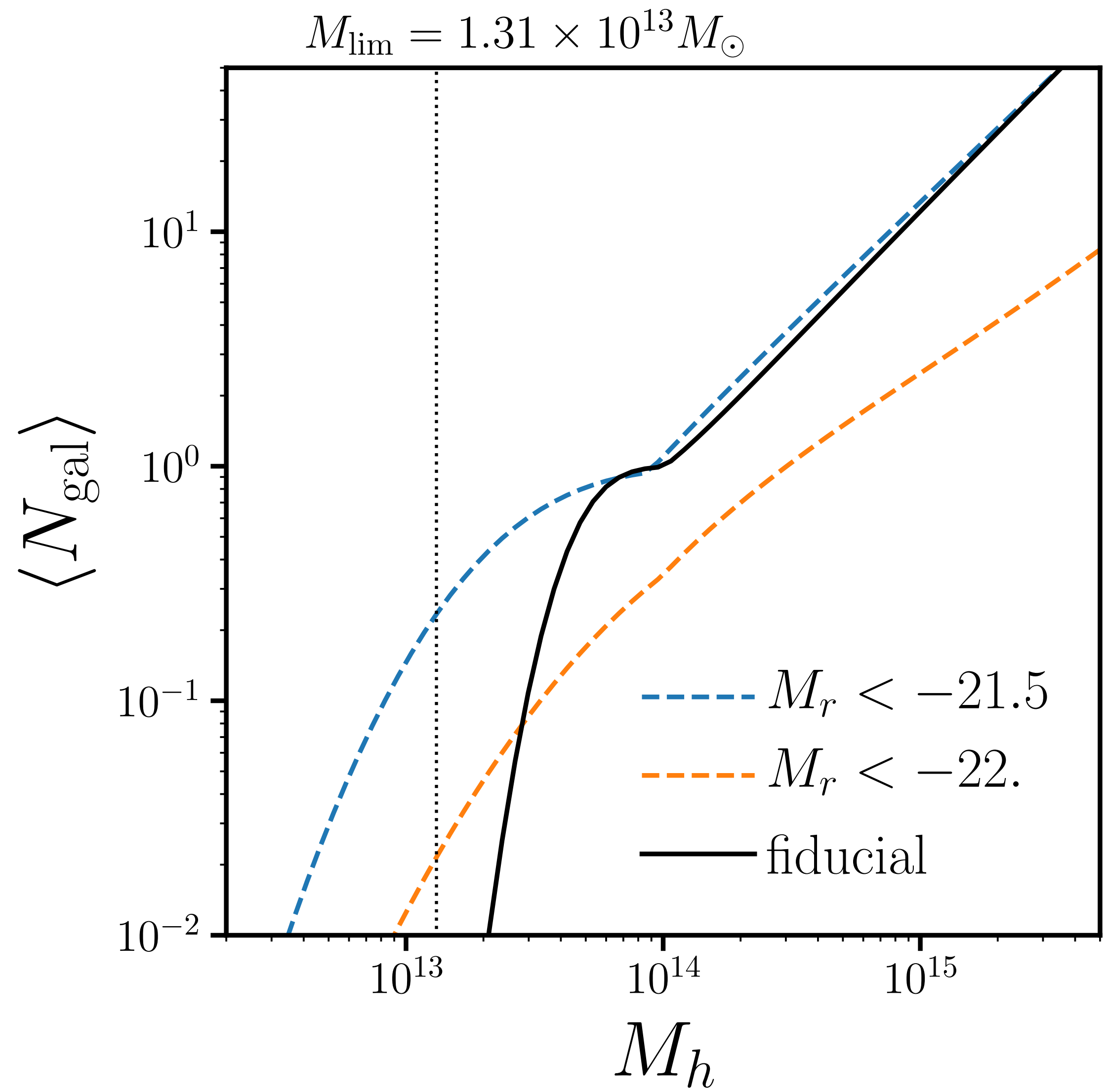
$$N_{\rm gal} \sim P(N\,|\,M_h)$$

$$\langle N_{\rm gal}\rangle = \langle N_{\rm cen}\rangle + \langle N_{\rm sat}\rangle$$

$$\langle N_{\rm cen}\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M_h - \log M_{\rm min}}{\sigma_{\log M}}\right)\right] \qquad \langle N_{\rm sat}\rangle = \langle N_{\rm cen}\rangle\left(\frac{M_h - M_0}{M_1}\right)^{\alpha}$$

standard HOD model (Zheng+2007)

fiducial HOD based on best-fit HOD parameters for the *SDSS* $M_r < -21.5$ *and* $M_r < -22$ *samples*



Zheng+(2007)

molino

changhoonhahn.github.io/molino/

15,000 mocks at fiducial HOD

+ (*500 N-body*) x (*5* HOD) x (*24* param)

75,000 galaxy catalogs

molino

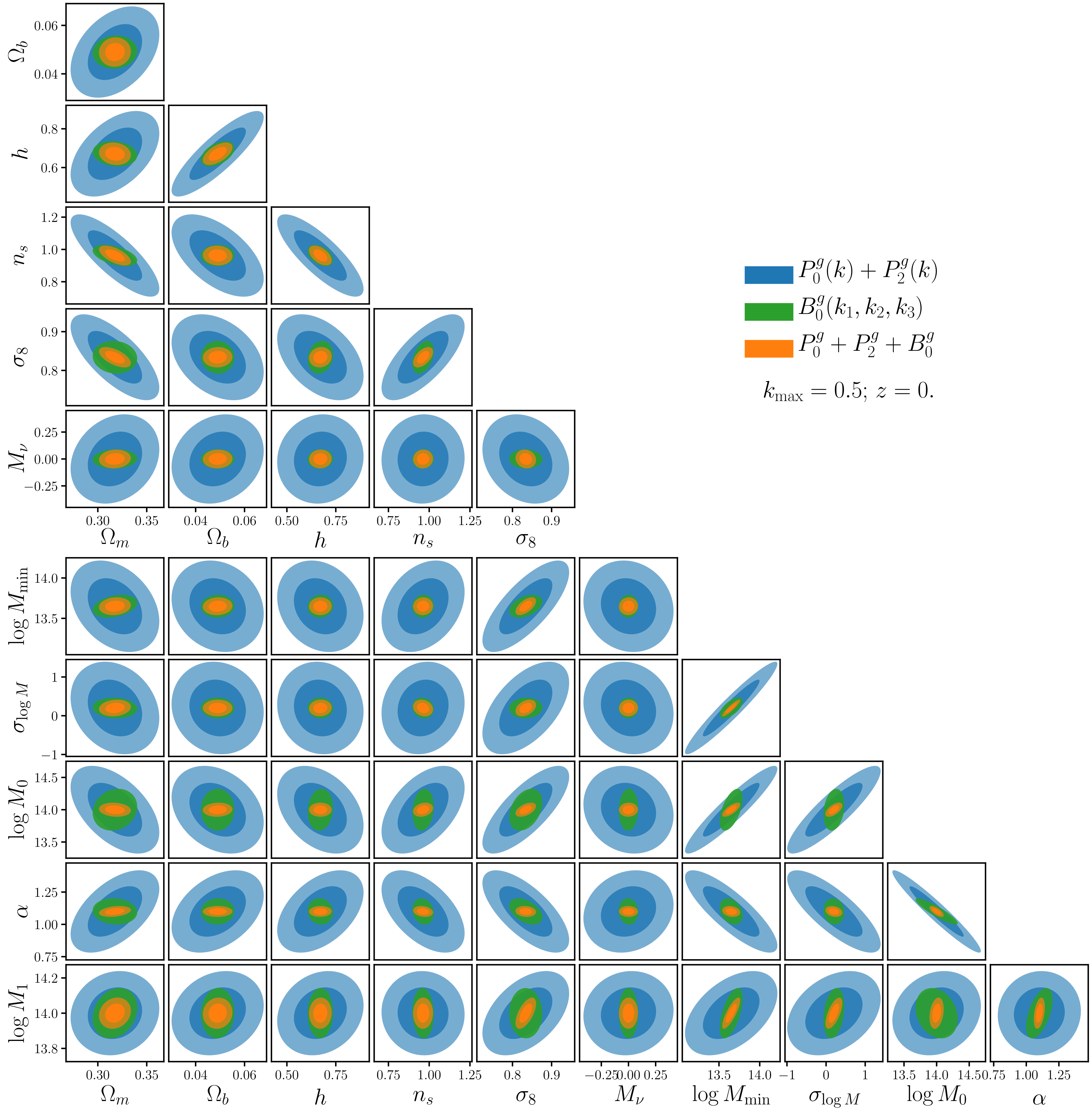
changhoonhahn.github.io/molino/

75,000 galaxy catalogs to quantify the information content of any galaxy cosmological observable

6 cosmological parameters + ***5 HOD parameters***

$M_\nu^{+++} = 0.4$										
$M_\nu^{++} = 0.2$										
$M_\nu^{+} = 0.1$	$\Omega_m^+ = 0.3275$	$\Omega_b^+ = 0.051$	$h^+ = 0.6911$	$n_s^+ = 0.9824$	$\sigma_8^+ = 0.849$	$M_{\rm min}^+ = 13.7$	$\sigma_{\log M}^+ = 0.22$	$\log M_0^+ = 14.2$	$\alpha^+ = 1.3$	$\log M_1^+ = 14.2$
$M_\nu = 0.0\text{eV}$	$\Omega_m = 0.3175$	$\Omega_b = 0.049$	$h = 0.6711$	$n_s = 0.9624$	$\sigma_8 = 0.834$	$M_{\rm min} = 13.65$	$\sigma_{\log M} = 0.2$	$\log M_0 = 14$	$\alpha = 1.1$	$\log M_1 = 14$
	$\Omega_m^- = 0.3075$	$\Omega_b^- = 0.047$	$h^- = 0.6511$	$n_s^- = 0.9424$	$\sigma_8^- = 0.819$	$M_{\rm min}^- = 13.6$	$\sigma_{\log M}^- = 0.18$	$\log M_0^+ = 13.8$	$\alpha^- = 0.9$	$\log M_1^- = 13.8$

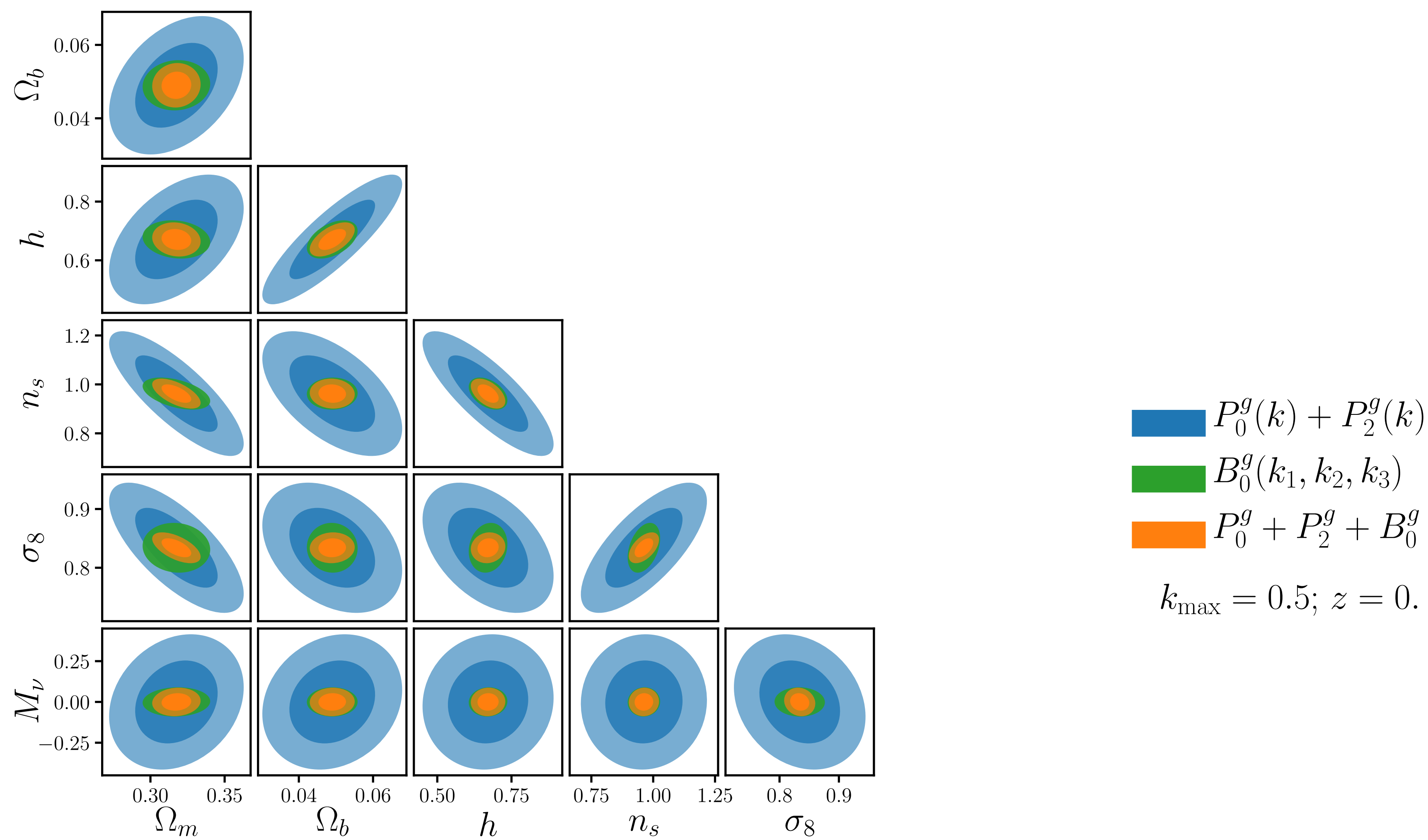
with molino, we can quantify the total information content of ***the redshift-space galaxy bispectrum***



$P_0^g(k) + P_2^g(k)$
$B_0^g(k_1, k_2, k_3)$
$P_0^g + P_2^g + B_0^g$
$k_{\rm max} = 0.5$; $z = 0$.
Ω_m
Ω_b
h
n_s
σ_8
M_ν
$\log M_{\rm min}$
$\sigma_{\log M}$
$\log M_0$
α
$\log M_1$

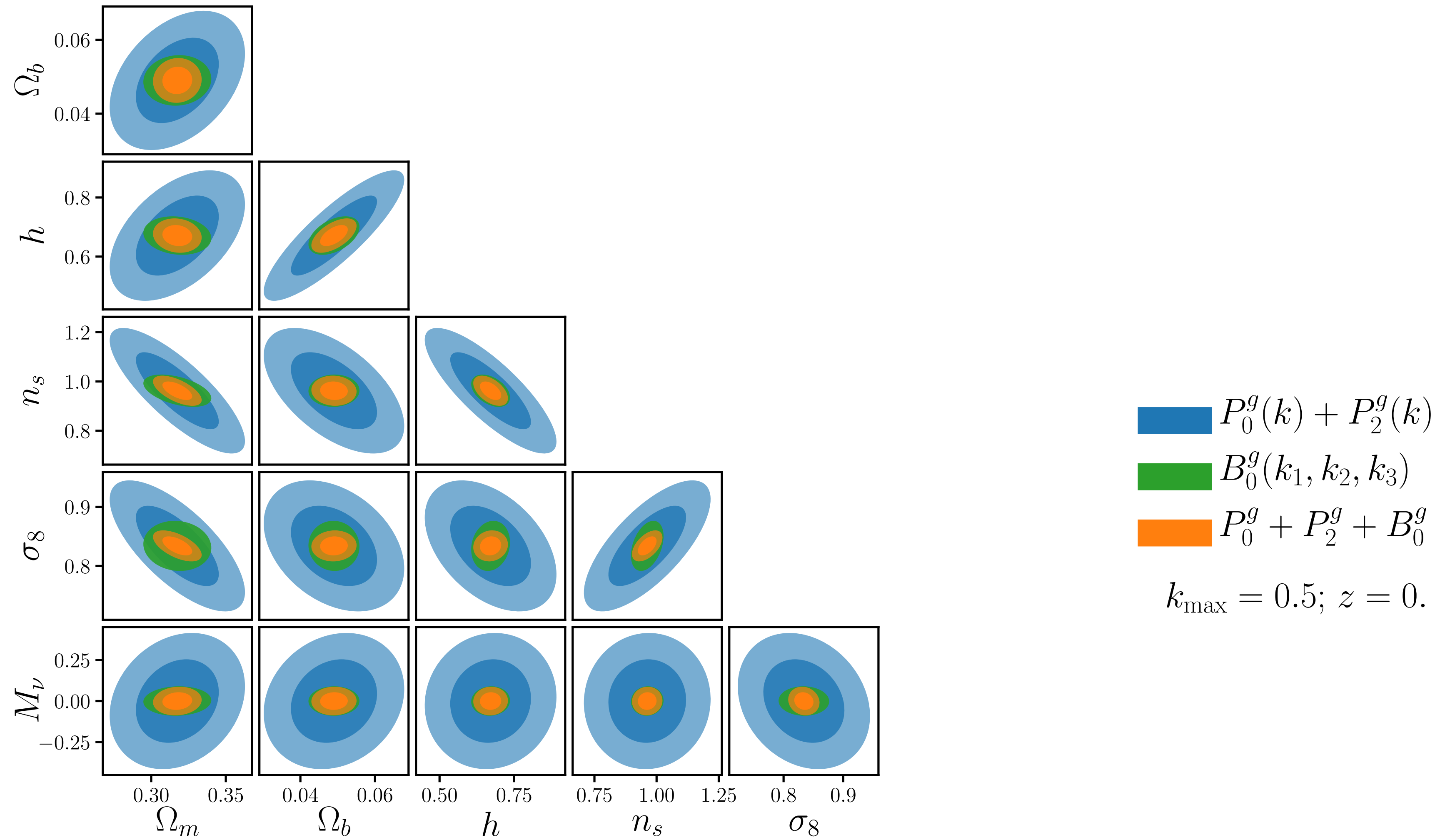
Hahn+(2021)

joint galaxy $P_\ell + B$ constraints on Ω_m, Ω_b, h, n_s, σ_8 are **$\gtrsim 3\times$ tighter** than the galaxy P_ℓ alone



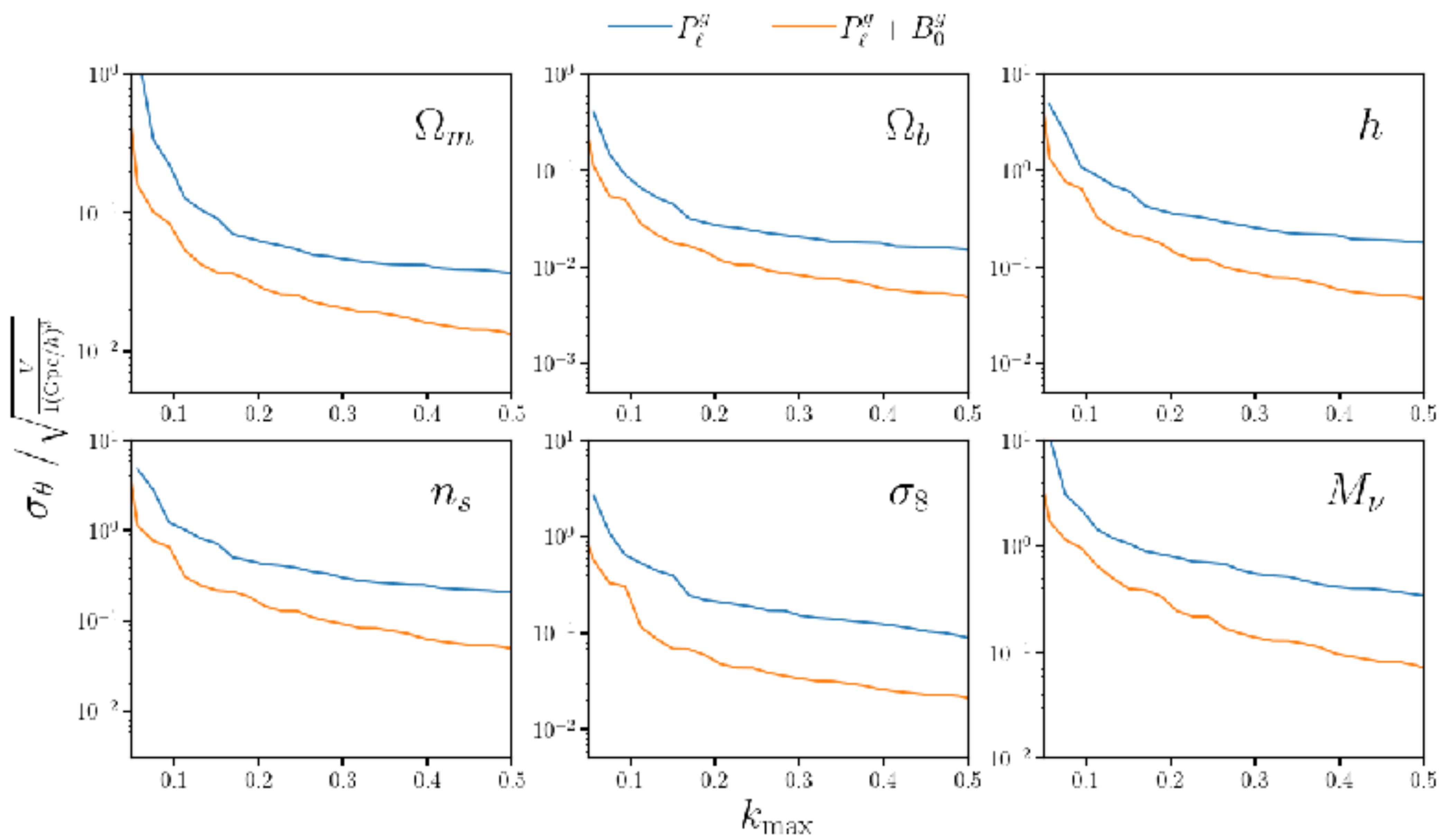
Hahn+(2021)

joint galaxy $P_\ell + B$ constraints on Ω_m, Ω_b, h, n_s, σ_8 are $\gtrsim 3\times$ **tighter** than the galaxy P_ℓ alone; *M_ν constraint* is $\gtrsim 5\times$ **tighter**



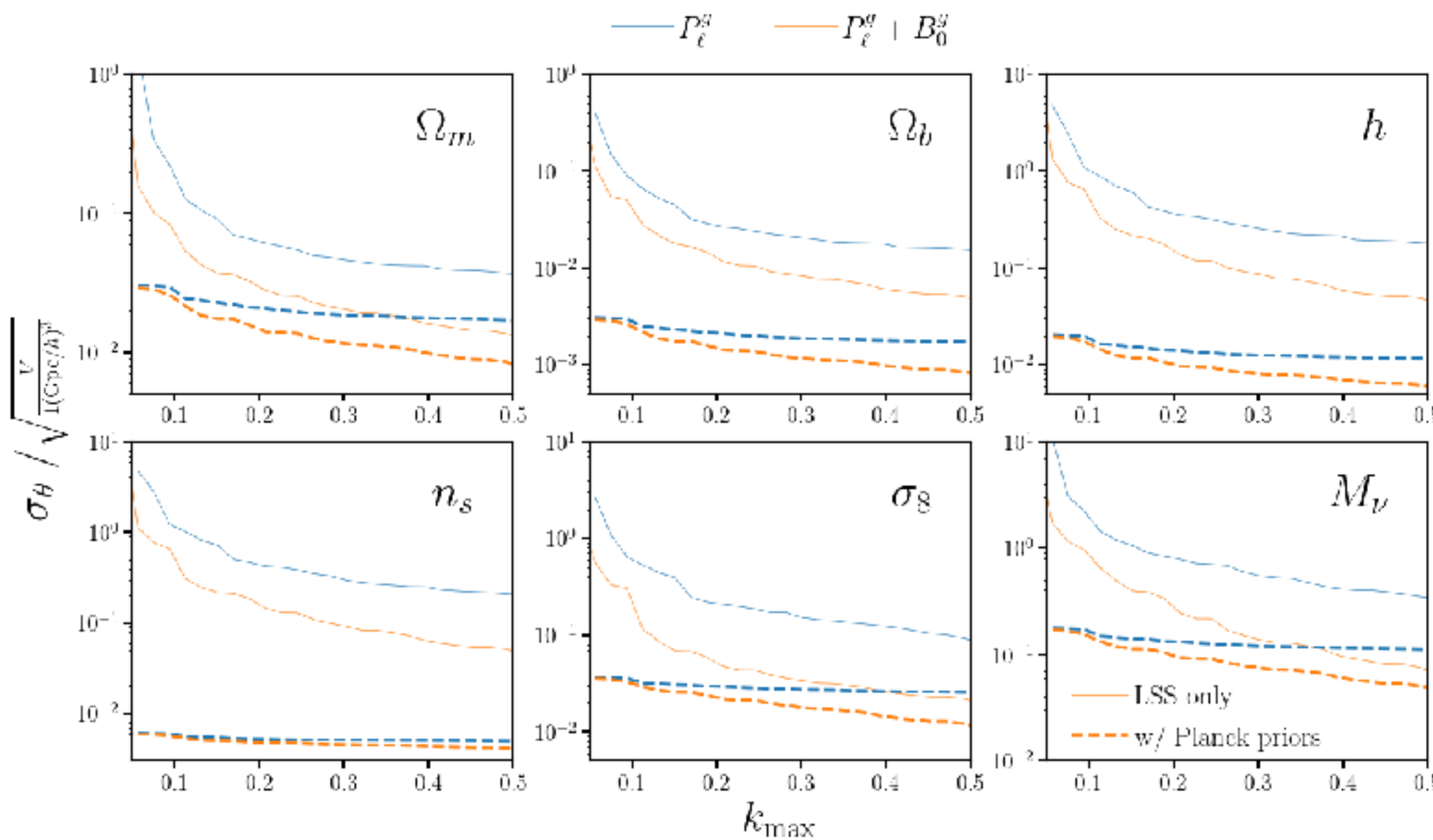
Hahn+(2021)

significant **constraining power** on small *nonlinear scales*



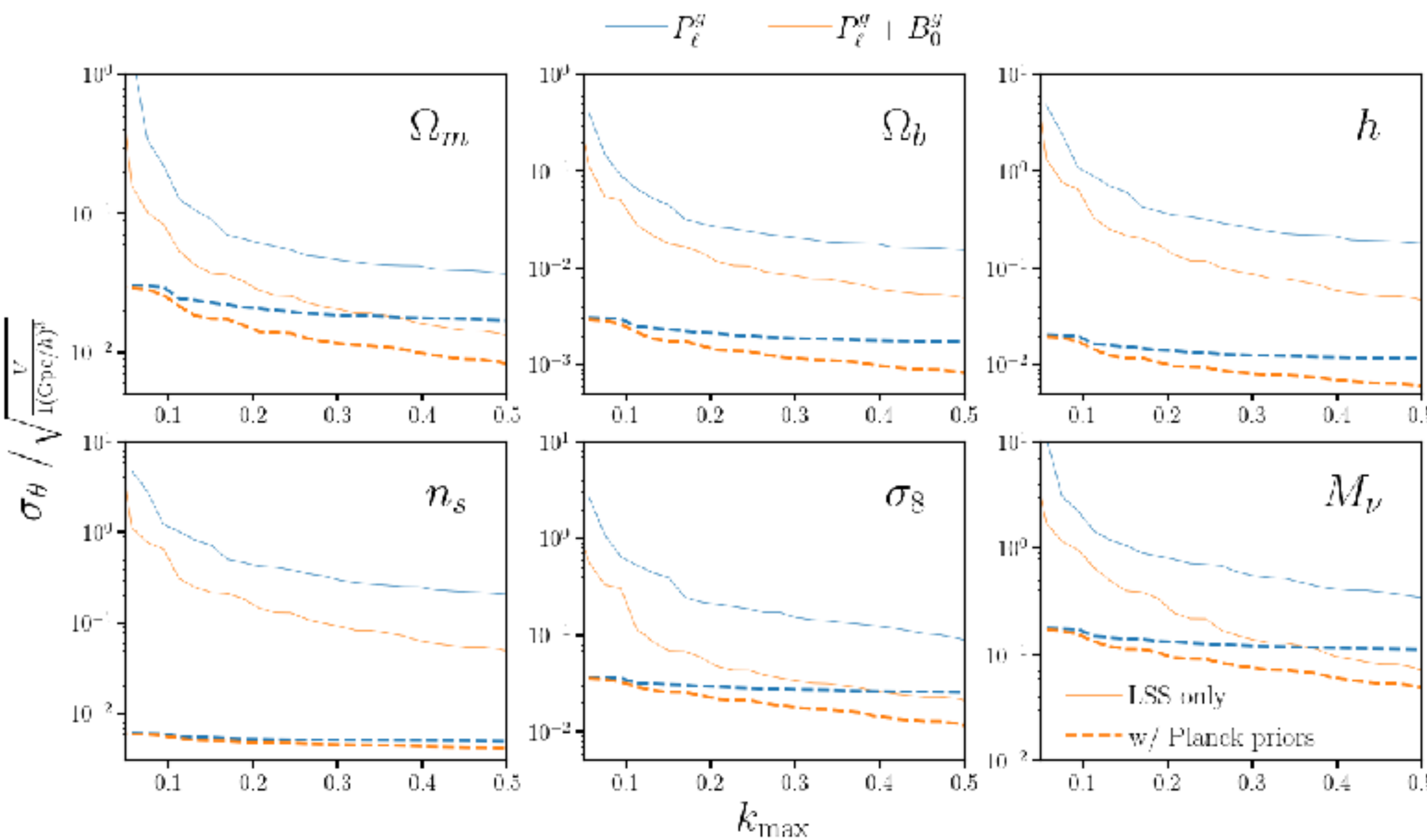
Hahn+(2021)

$\gtrsim 2\times$ **tighter** parameter constraints even *with Planck priors*



Hahn+(2021)

information content of P_ℓ **saturates** $k_{\rm max} \gtrsim 0.1\ h/{\rm Mpc}$



Hahn+(2021)

the **bispectrum** significant improves parameter constraints over P_ℓ — *even after marginalizing over HOD parameters*

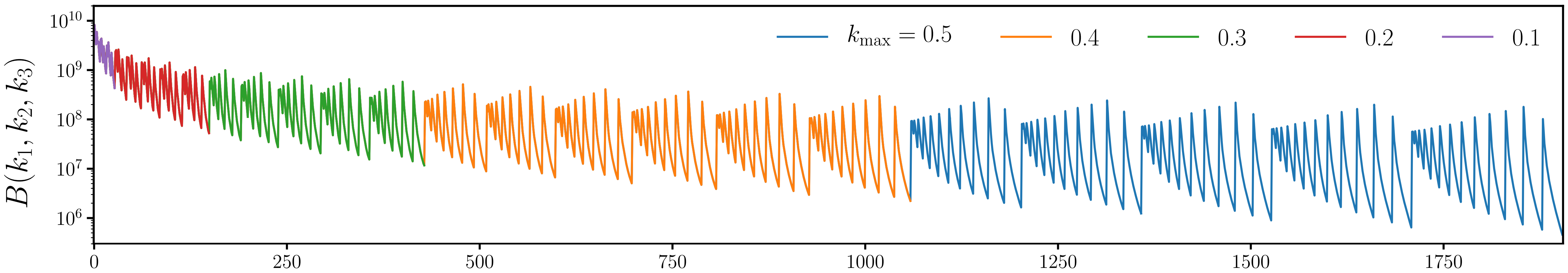
there's *significant constraining power* in the **nonlinear regime** that can be exploited the bispectrum is *very promising* but there are **some challenges**

$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*

theoretical limitations on nonlinear scales

systematic effects

$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ is **1898 dimensional**

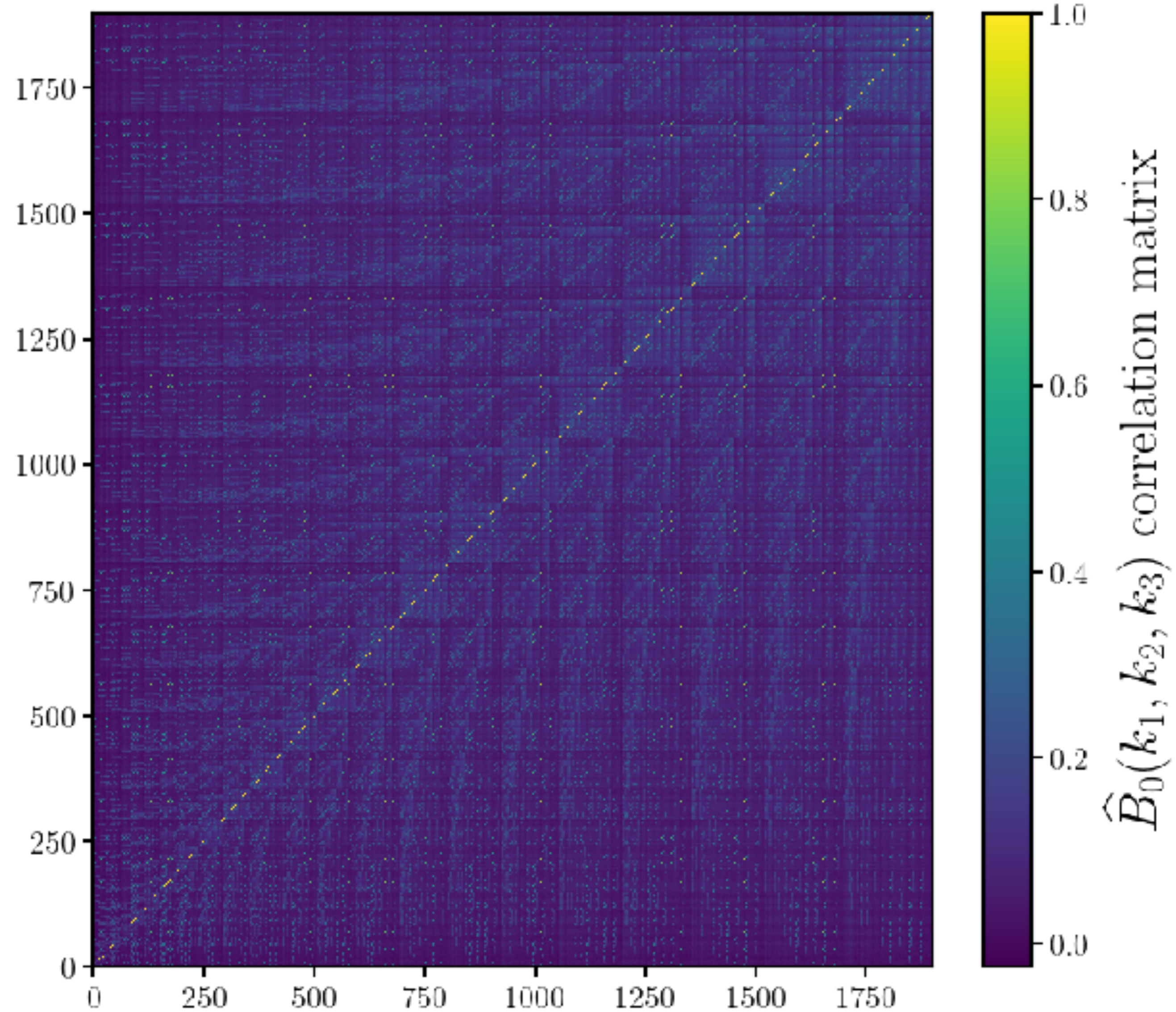


$k_1 \geq k_2 \geq k_3$

github.com/changhoonhahn/pySpectrum

we used *15,000* mocks to estimate $\mathbf{C}$ but for current analyses

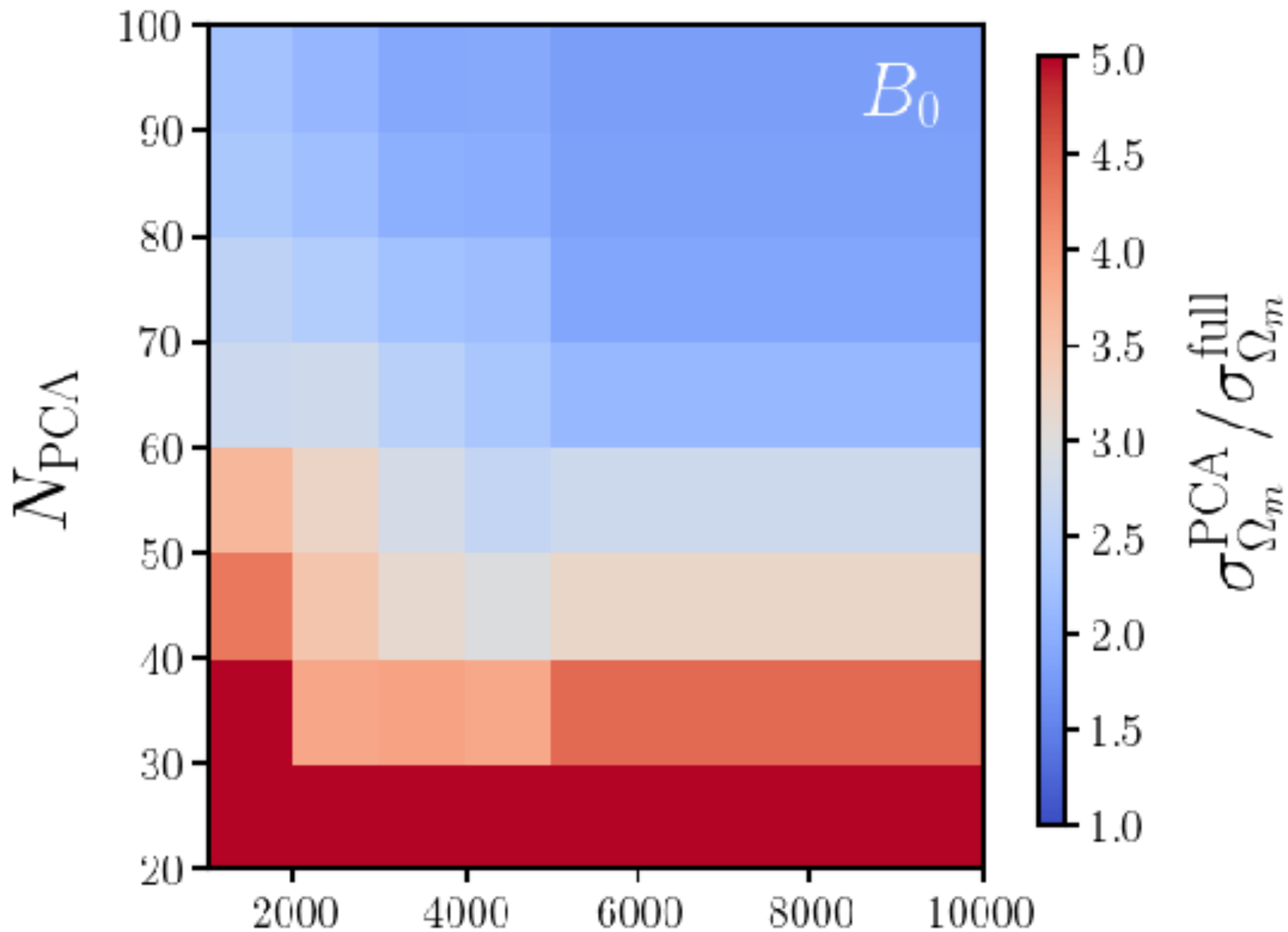
$N_{\rm mock} \sim 2000$



Hahn+(2020)

if use a **compressed bispectrum** how much constraining power do we retain?

with **PCA** data compression we lose a lot of the constraining power we gain from the bispectrum



N_{mock} used to
estimate the compression

Hahn+ in prep.

we can do better than PCA compression by **maximizing the Fisher information** (KL/MOPED/score)

linear combination of the bispectrum: $cB = \mathbf{b}^T B$

that maximizes $F_{ij} = \frac{1}{2}\left(\frac{\mathbf{b}^T\mathbf{C}_{,i}\mathbf{b}}{\mathbf{b}^T\mathbf{C}\mathbf{b}}\right)\left(\frac{\mathbf{b}^T\mathbf{C}_{,j}\mathbf{b}}{\mathbf{b}^T\mathbf{C}\mathbf{b}}\right) + \frac{(\mathbf{b}^T B_{,i})(\mathbf{b}^T B_{,j})}{\mathbf{b}^T\mathbf{C}\mathbf{b}}$

$$cB = \left(\frac{\partial B}{\partial \theta_i}\mathbf{C}^{-1}\right)B$$

**Tegmark+(1997), Heavens+(2000), Alsing+(2018), Gualdi+(2018, 2019)*

we can do better than PCA compression by **maximizing the Fisher information** (KL/MOPED/score)

linear combination of the bispectrum: $cB = \mathbf{b}^T B$

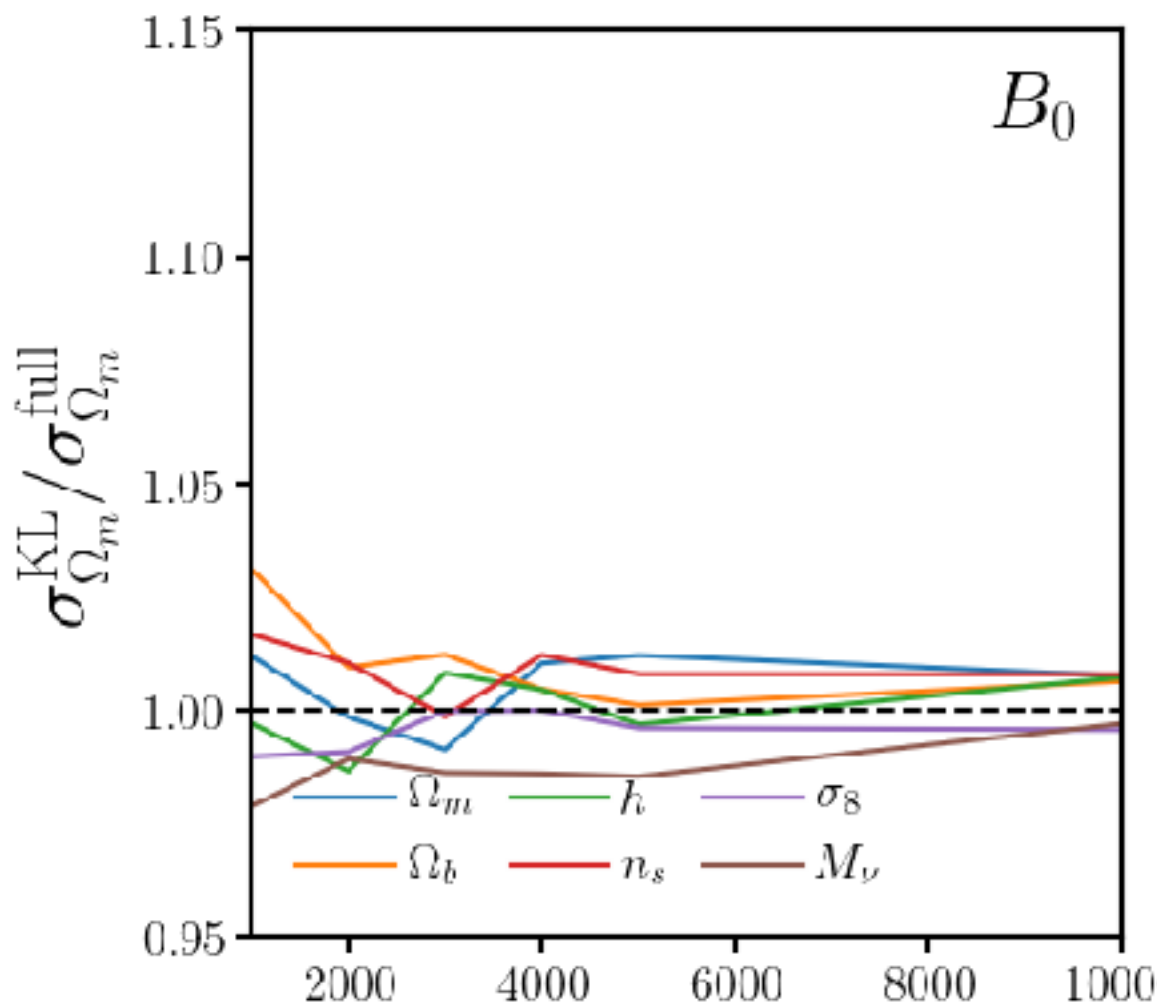
that maximizes $$F_{ij} = \frac{1}{2}\left(\frac{\mathbf{b}^T\mathbf{C}_{,i}\mathbf{b}}{\mathbf{b}^T\mathbf{C}\mathbf{b}}\right)\left(\frac{\mathbf{b}^T\mathbf{C}_{,j}\mathbf{b}}{\mathbf{b}^T\mathbf{C}\mathbf{b}}\right) + \frac{(\mathbf{b}^T B_{,i})(\mathbf{b}^T B_{,j})}{\mathbf{b}^T\mathbf{C}\mathbf{b}}$$

$$cB = \left(\frac{\partial B}{\partial \theta_i}\mathbf{C}^{-1}\right) B$$

1898 dimensions $\longrightarrow$ **6 dimensions** (# of θ_{cosmo})

**Tegmark+(1997), Heavens+(2000), Alsing+(2018), Gualdi+(2018, 2019)*

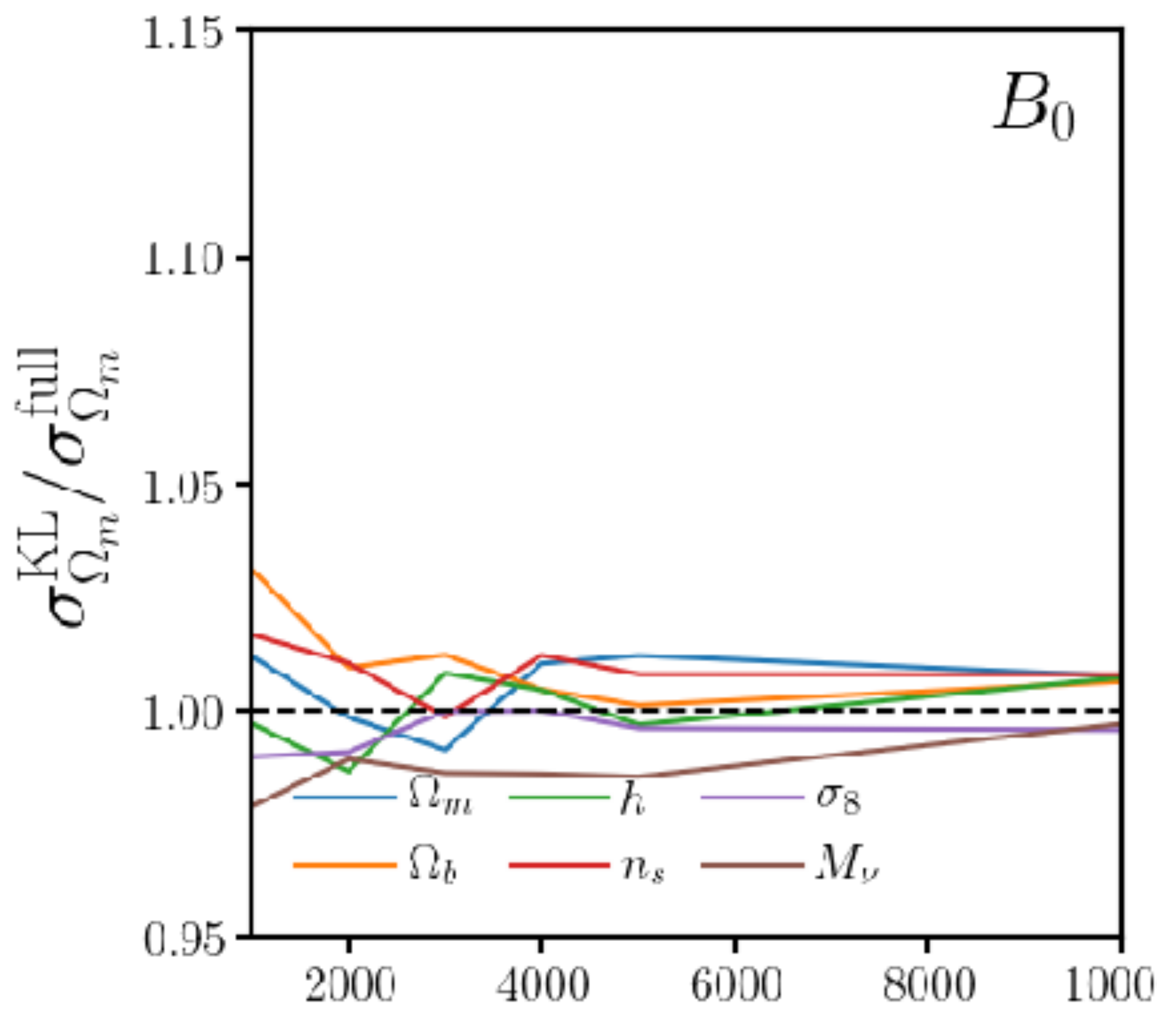
we can do better than PCA compression by **maximizing the Fisher information** (KL/MOPED/score)



N_{mock} used to
estimate the compression

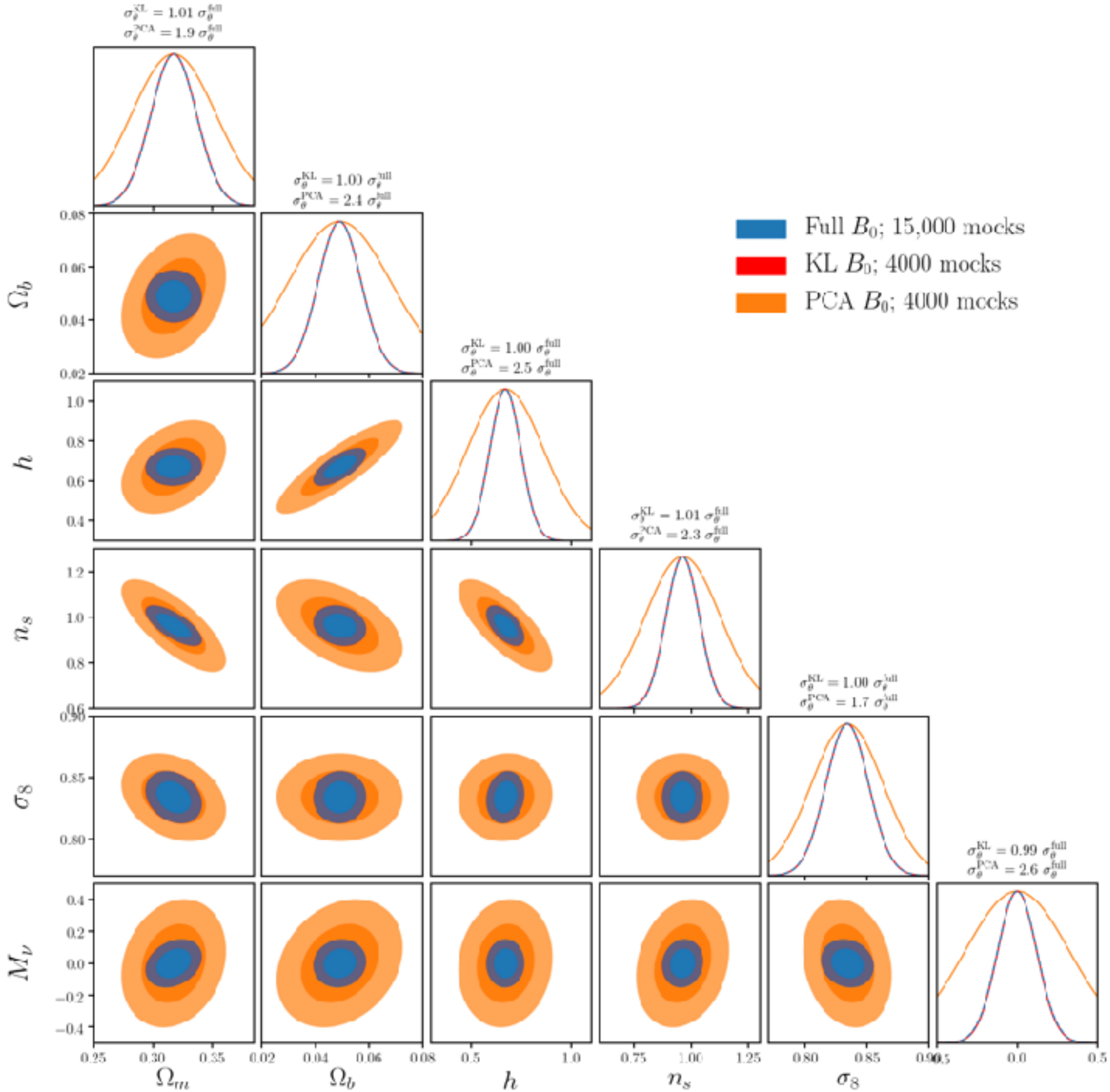
Hahn+ in prep.

even with $N_{\rm mock} < 4000$ we can extract *most* of the constraining power of the bispectrum with KL compression



$N_{\rm mock}$ used to estimate the compression

Hahn+ in prep.



Full B_0; 15,000 mocks
KL B_0; 4000 mocks
PCA B_0; 4000 mocks
$\sigma_\theta^{\rm KL} = 1.01\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 1.9\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm KL} = 1.00\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 2.4\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm KL} = 1.00\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 2.5\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm KL} = 1.01\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 2.3\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm KL} = 1.00\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 1.7\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm KL} = 0.99\ \sigma_\theta^{\rm full}$
$\sigma_\theta^{\rm PCA} = 2.6\ \sigma_\theta^{\rm full}$
Ω_m
Ω_b
h
n_s
σ_8
M_ν

Hahn+ in prep.

$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*

theoretical limitations on nonlinear scales

systematic effects

we can dramatically reduce the dimensionality of B while exploiting most of its constraining power with ***clever data compression***

~~$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*~~

theoretical limitations on nonlinear scales

systematic effects

in standard cosmological parameter inference, we infer the posteriors — probability of θ_{cosmo} given some measurement x

$$p(\theta \mid x) \propto p(x \mid \theta)\, p(\theta)$$

in standard cosmological parameter inference, we infer the posteriors — probability of θ_{cosmo} given some measurement x

$$p(\theta \mid x) \propto \underbrace{p(x \mid \theta)}_{\textbf{\textit{likelihood}}} \overbrace{p(\theta)}^{\textbf{\textit{prior}}}$$

in standard cosmological parameter inference, we infer the posteriors — probability of θ_{cosmo} given some measurement x

$$p(\theta \mid x) \propto \boxed{p(x \mid \theta)}\, p(\theta)$$

$$p(x|\theta) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \mathbf{C}}} \exp\left[\frac{1}{2}(x - \text{model}(\theta))^T \mathbf{C}^{-1} (x - \text{model}(\theta))\right]$$

which we sample using some analytic likelihood with Monte Carlo method

standard cosmological parameter inference **assumes:**

$$p(x\,|\,\theta) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\mathbf{C}}}\exp\left[\frac{1}{2}(x-\mathrm{model}(\theta))^T\mathbf{C}^{-1}(x-\mathrm{model}(\theta))\right]$$

a Gaussian likelihood

standard cosmological parameter inference **assumes:**

$$p(x\,|\,\theta) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\mathbf{C}}}\exp\left[\frac{1}{2}(x - \text{model}(\theta))^T\mathbf{C}^{-1}(x - \text{model}(\theta))\right]$$

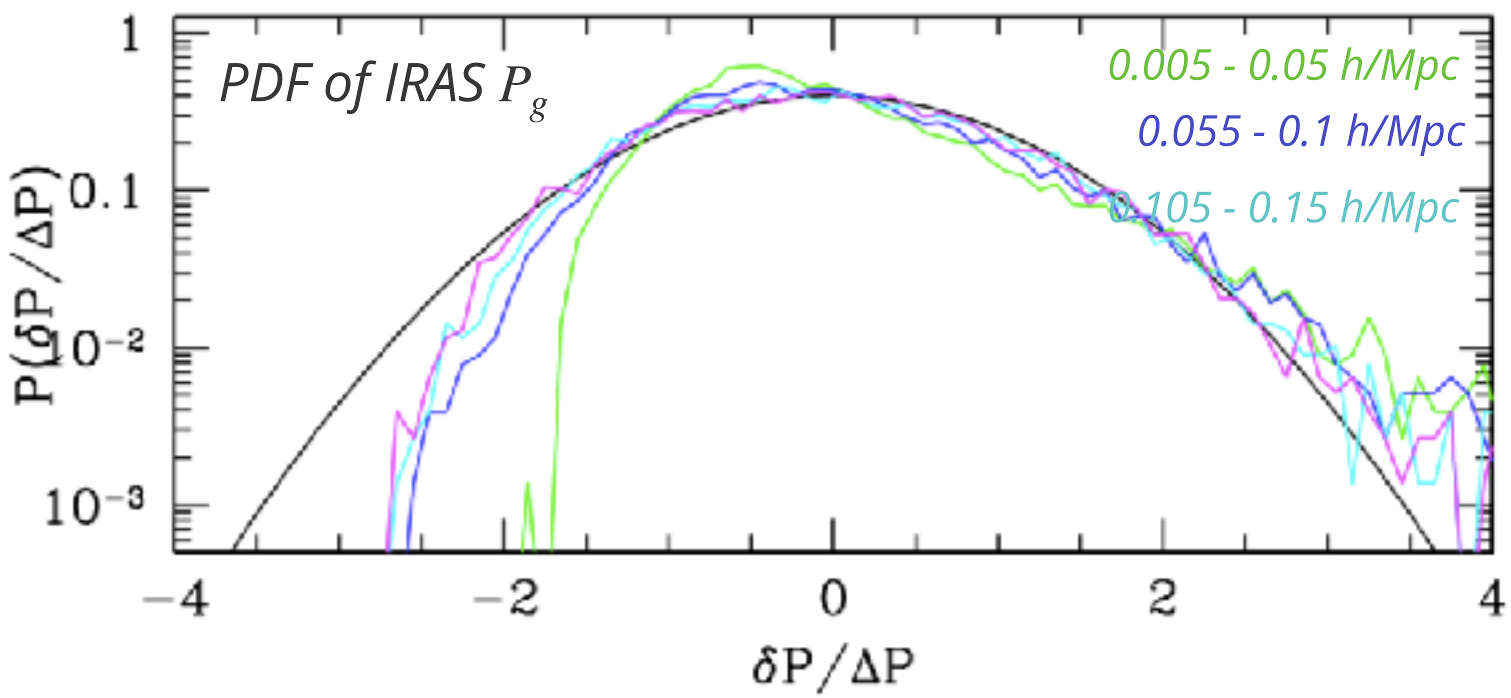
a Gaussian likelihood, *some perturbation theory model*

standard cosmological parameter inference **assumes:**

$$p(x\,|\,\theta) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\mathbf{C}}}\exp\left[\frac{1}{2}(x - \mathrm{model}(\theta))^T\mathbf{C}^{-1}(x - \mathrm{model}(\theta))\right]$$

a Gaussian likelihood, some perturbation theory model, *systematics corrected in the measurement*

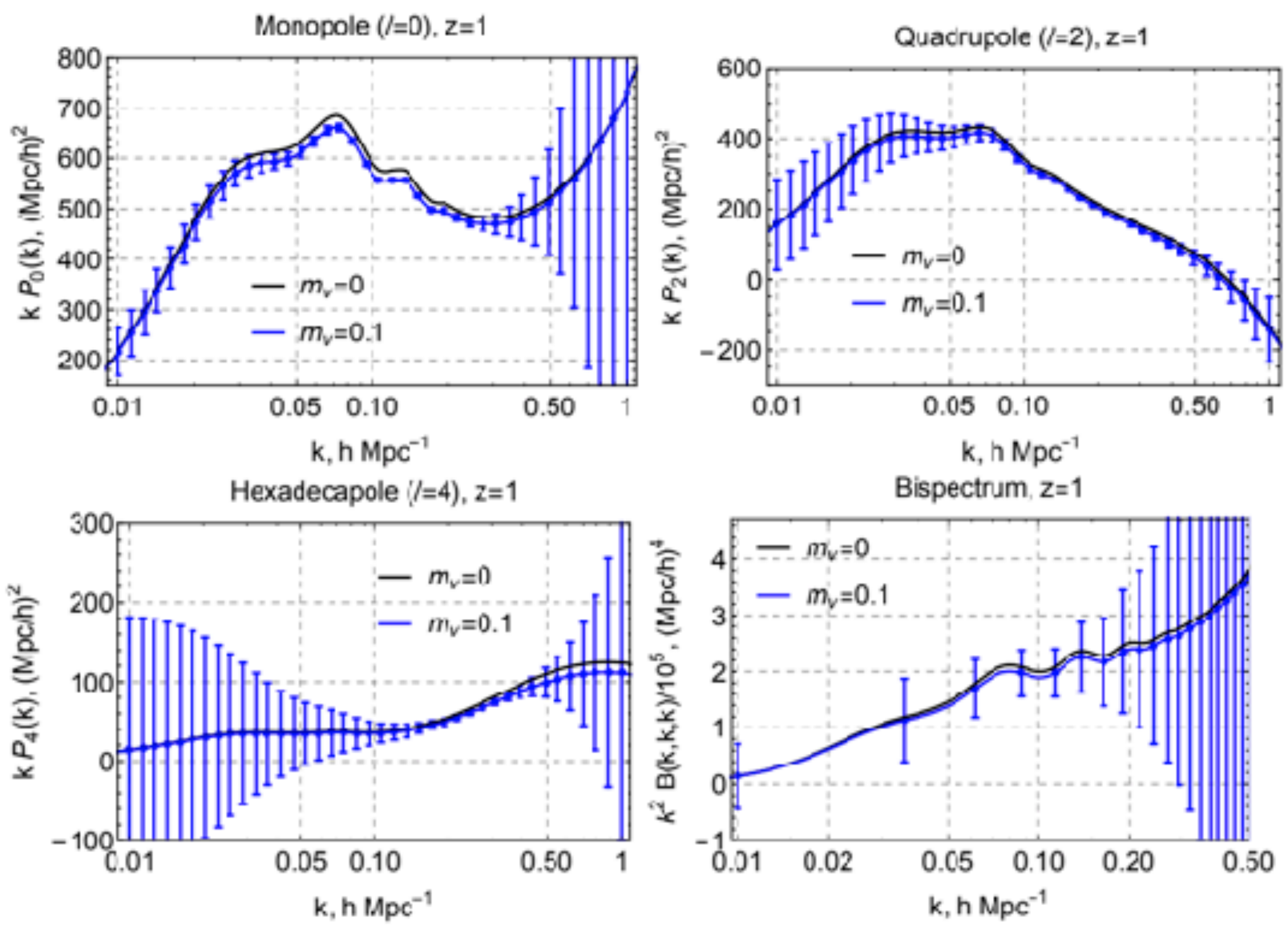
Gaussian likelihood, PT model, systematics



Gaussian assumption relies on **central limit theorem** — breaks down on low S/N regimes

Scoccimarro (2000)

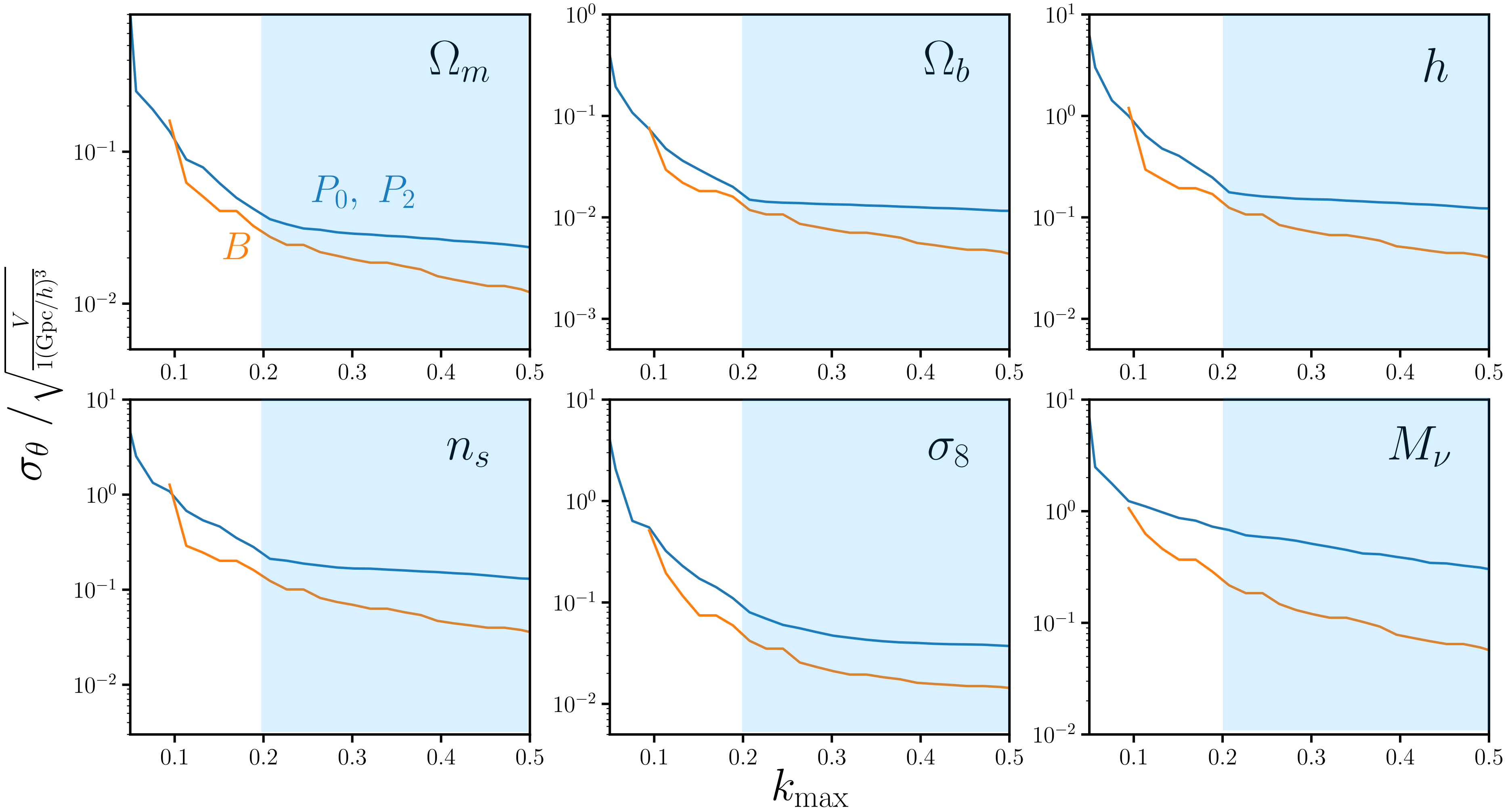
~~Gaussian likelihood~~, **PT model**, systematics



are limited to the *linear regime*

Chudaykin & Ivanov (2019)

~~Gaussian likelihood~~, **PT model**, systematics

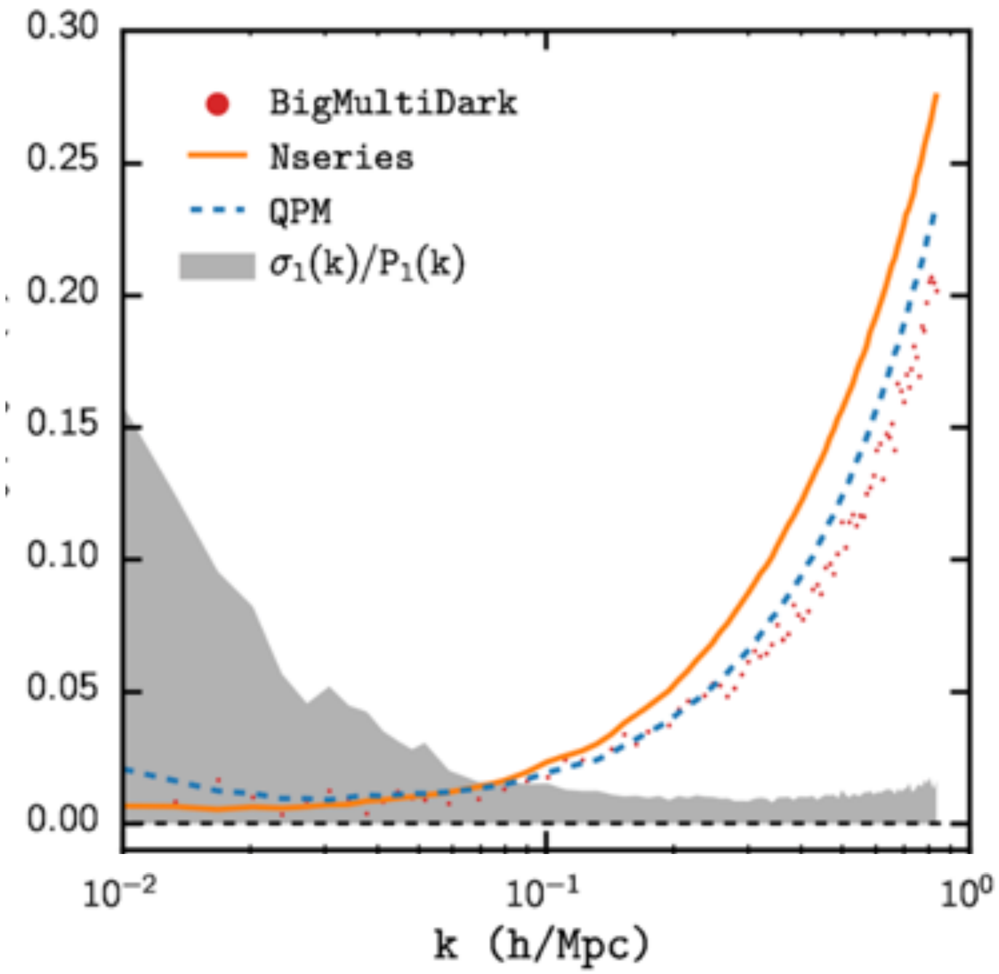


are limited to the *linear regime* — can't exploit the *constraining power on nonlinear regimes*

Chudaykin & Ivanov (2019)

~~Gaussian likelihood, PT model~~, **systematics**

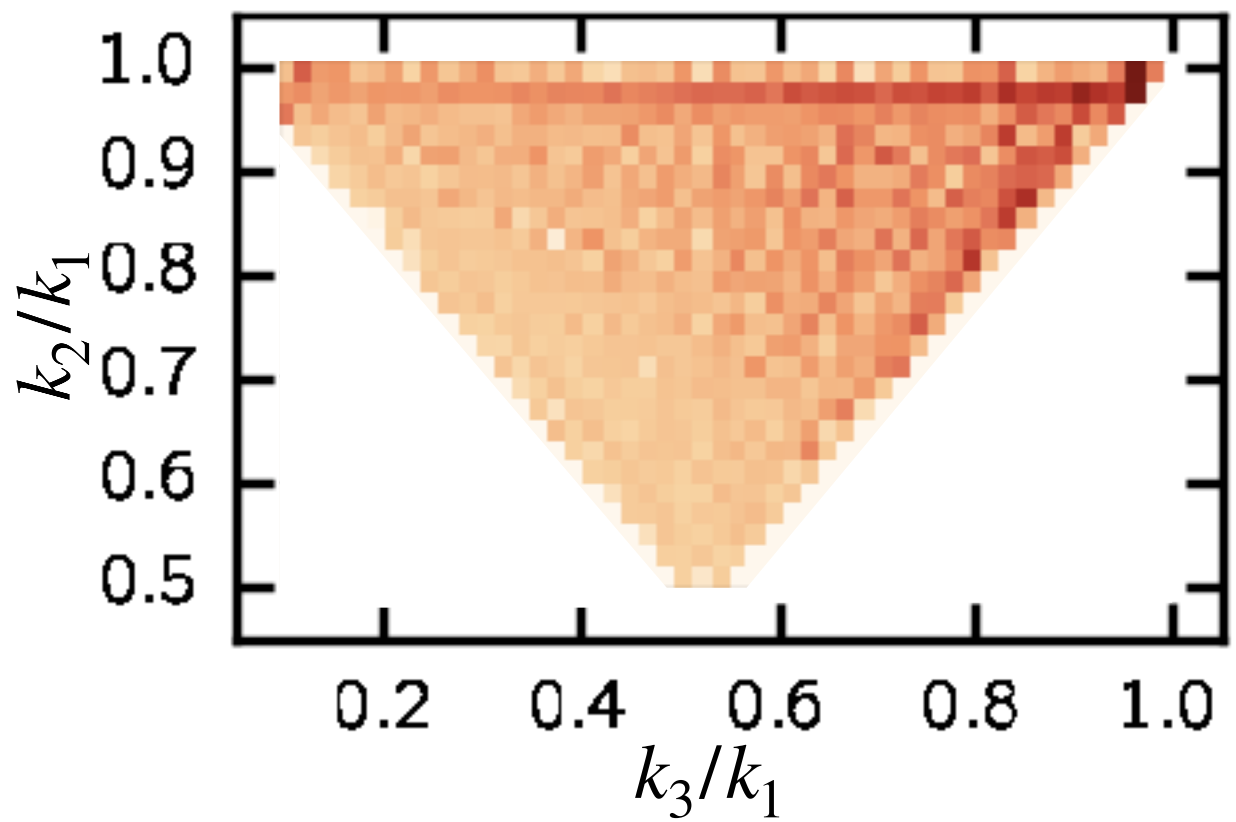
normalized P_0 residual



e.g. fiber collisions were difficult to correct in BOSS P_ℓ

Hahn+(2017)

~~Gaussian likelihood, PT model~~, **systematics**



e.g. fiber collisions were difficult to correct in BOSS P_ℓ — even ***harder for the bispectrum***

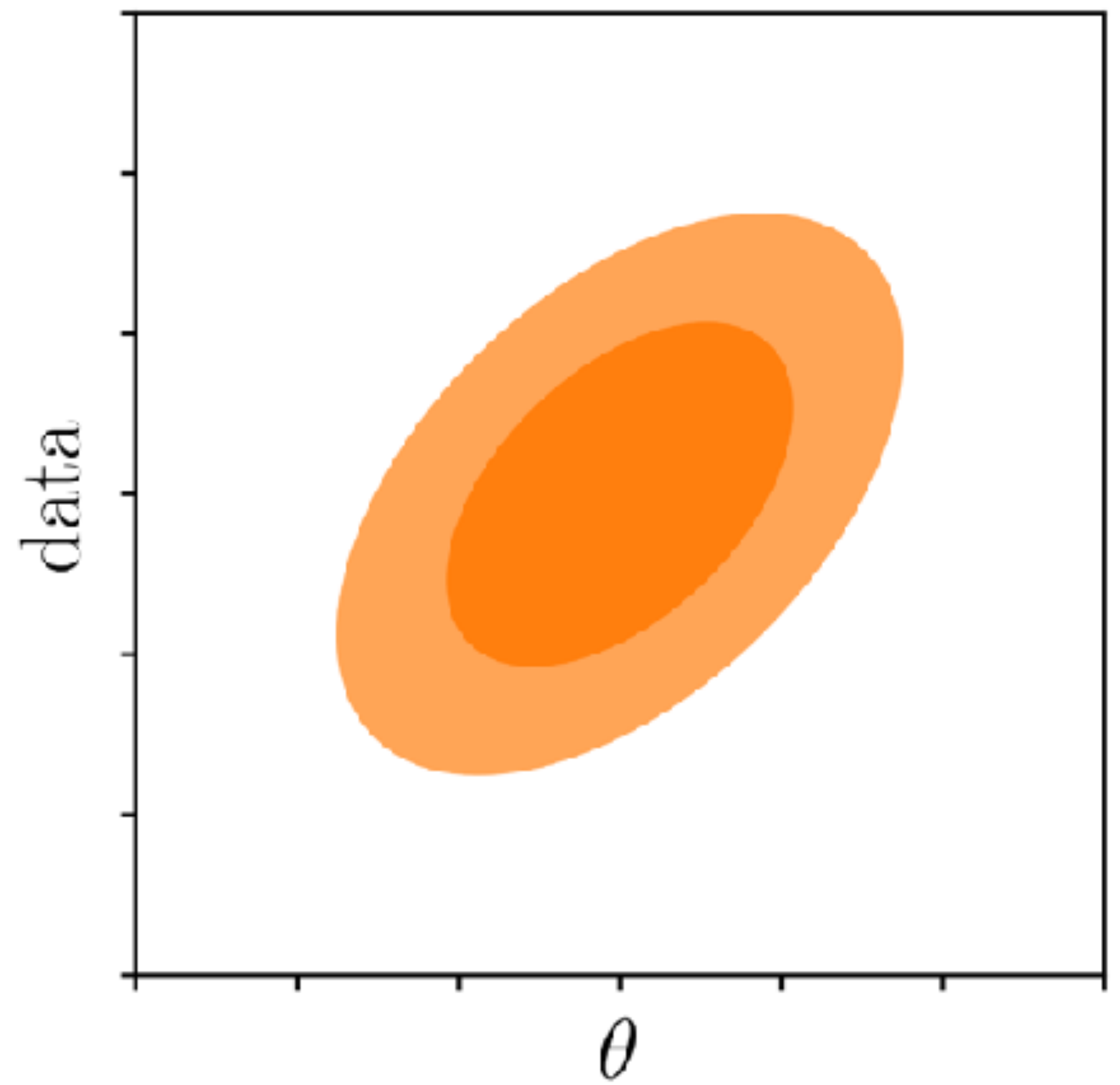
Hahn+(2017)

likelihood-free inference provides an alternative parameter inference framework that

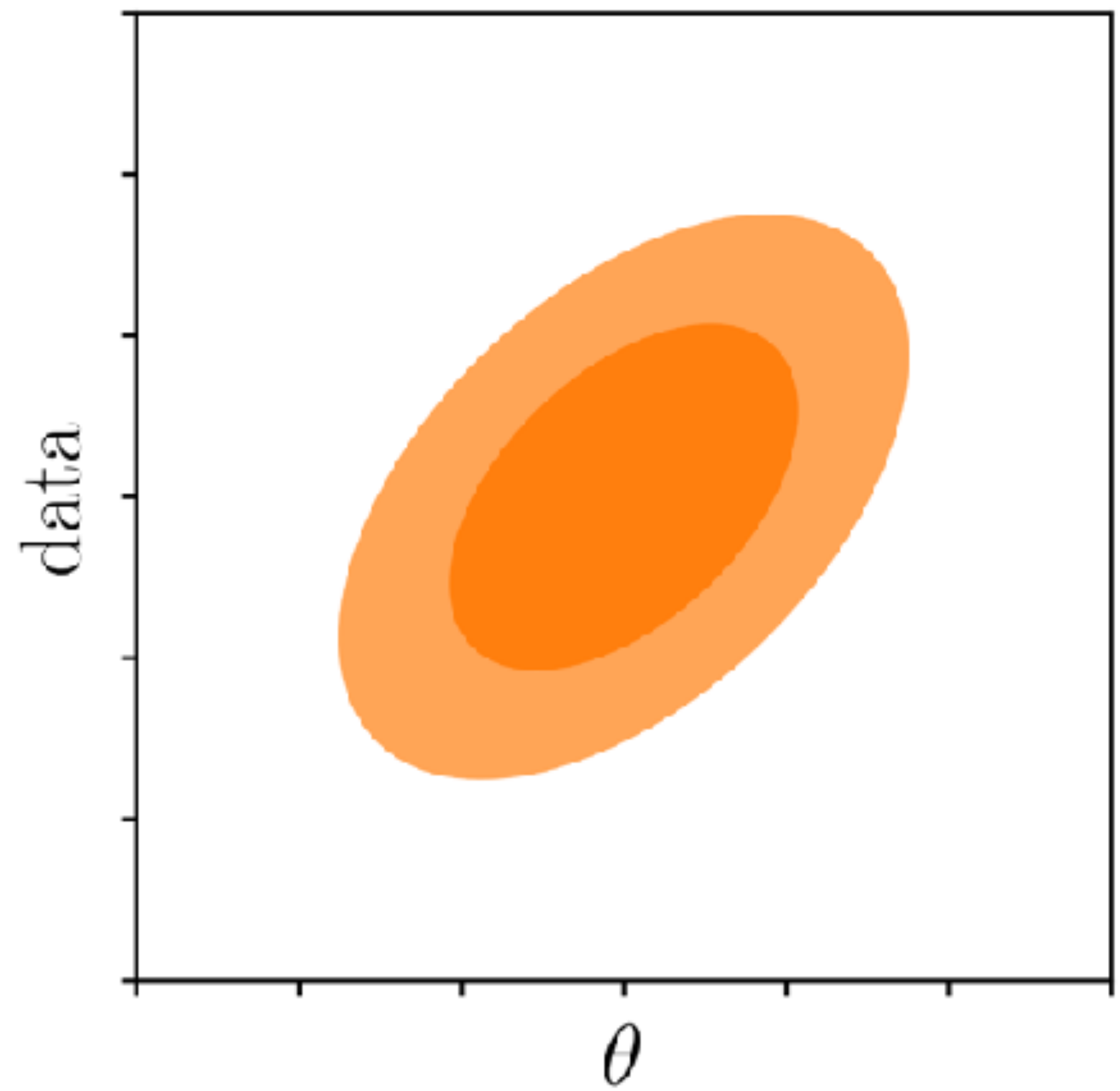
makes no assumptions on the likelihood

extends to nonlinear regimes

systematics can be included in the forward model consider the joint distribution $p(\text{data}, \theta)$ for 1D data and 1D θ

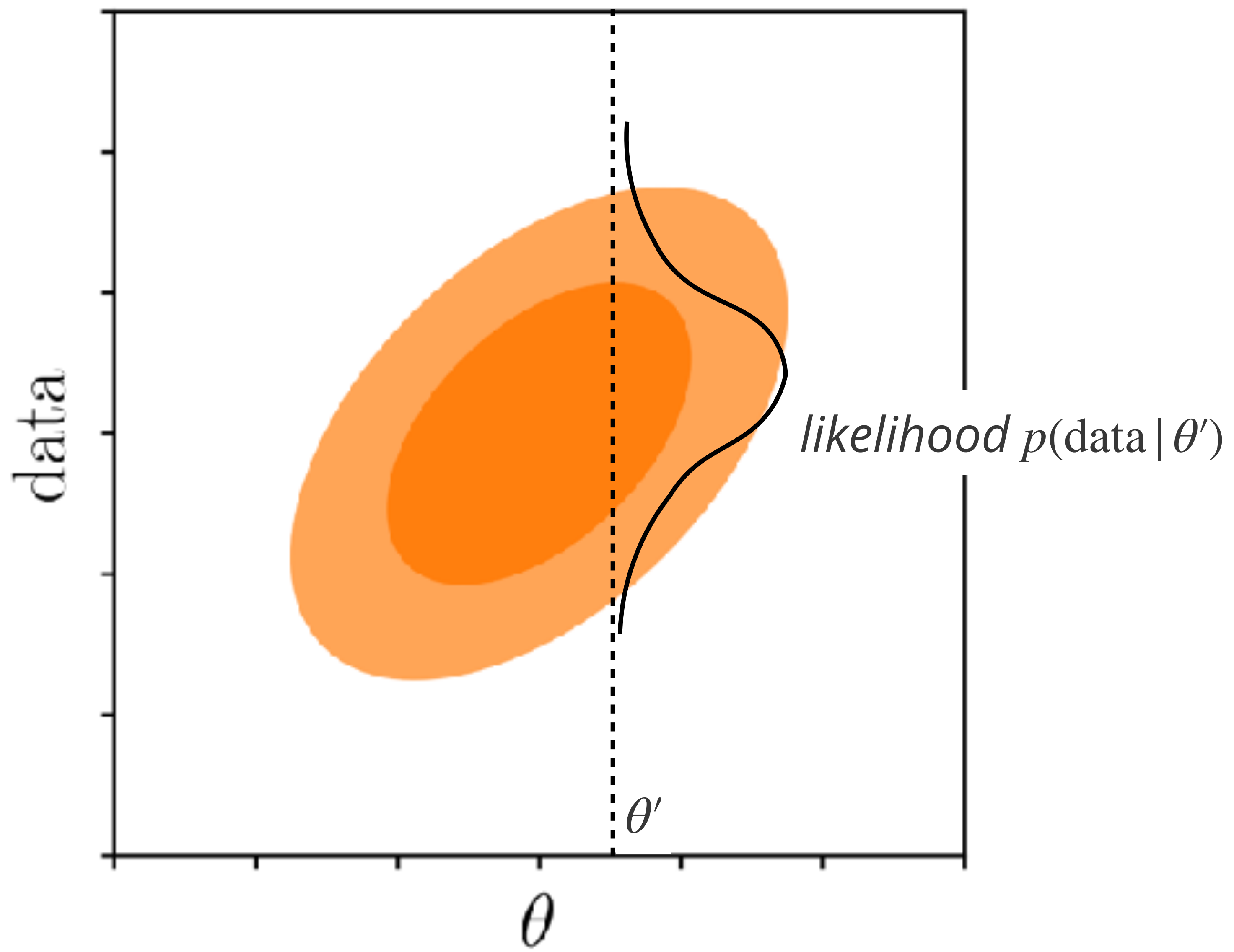


from $p(\mathrm{data}, \theta)$ we can get both the posterior and likelihood:

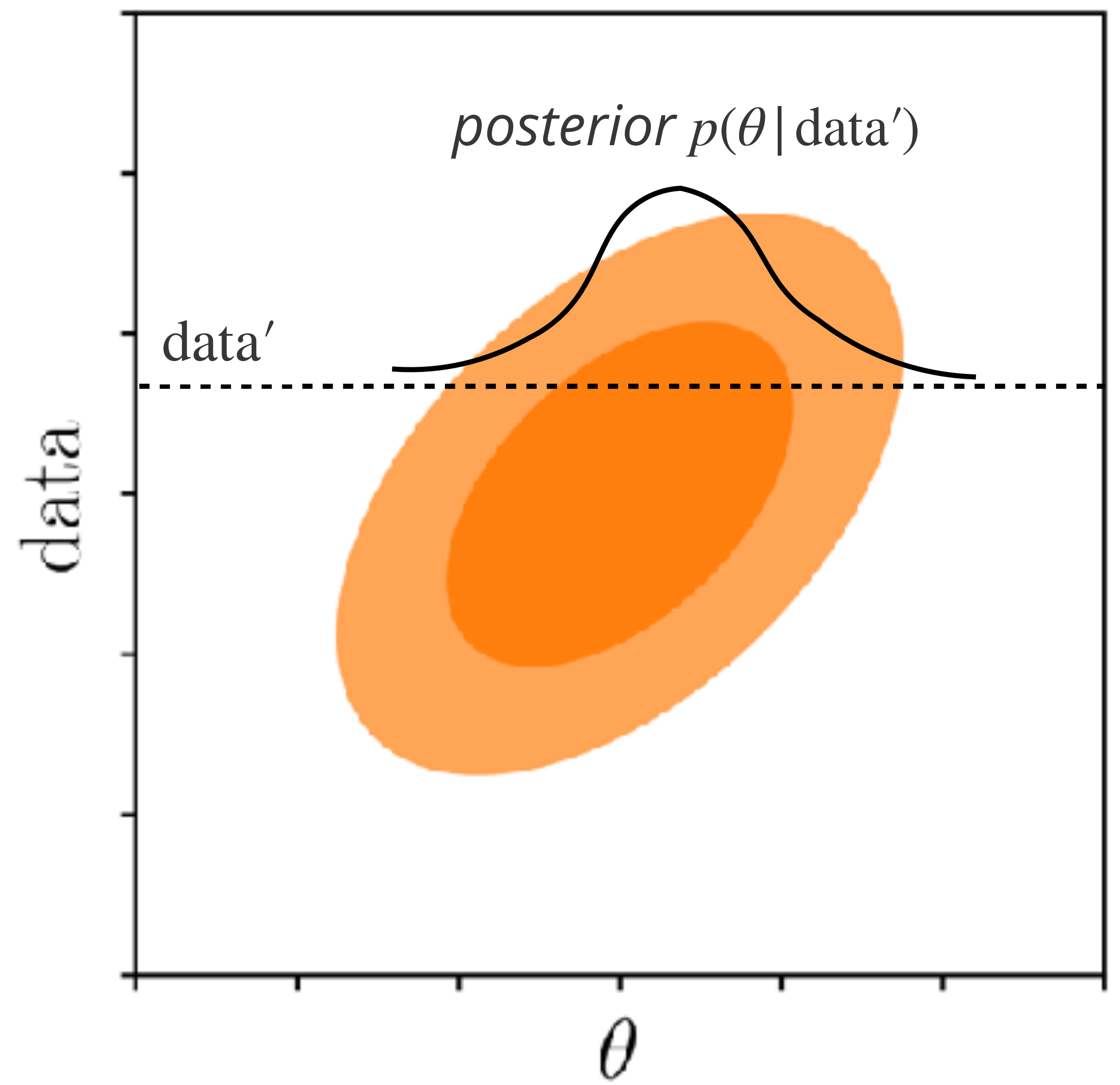


from $p(\mathrm{data}, \theta)$ we can get both the posterior and likelihood:

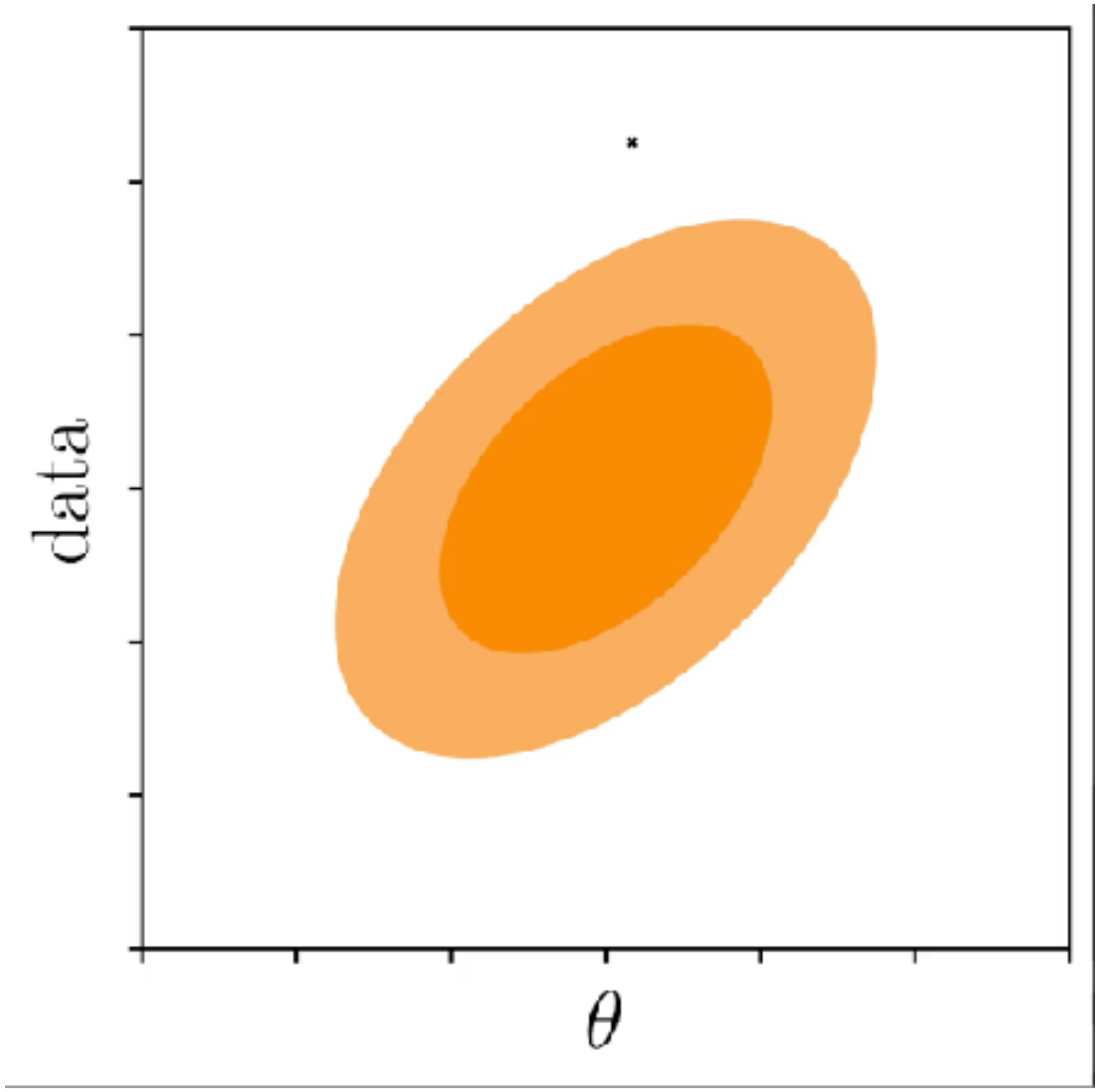
$p(\mathrm{data}, \theta)$ *conditioned on* θ' = **likelihood**



from $p(\mathrm{data}, \theta)$ we can get both the posterior and likelihood:
$p(\mathrm{data}, \theta)$ conditioned on data' = **posterior**

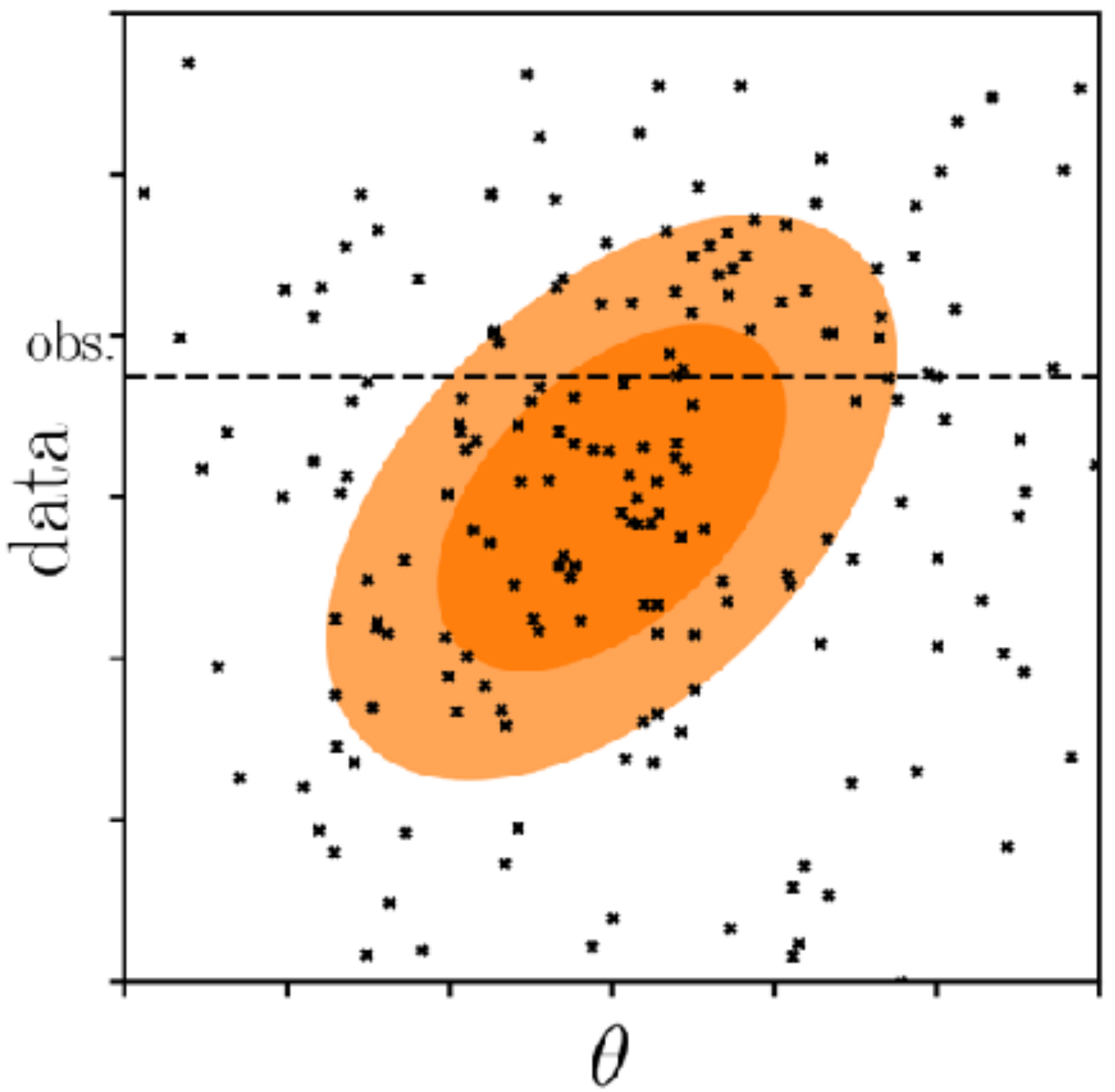


likelihood-free inference are methods that use *forward modeled simulations* to estimate the posterior one way to do likelihood-free inference is by *brute force* — **approximate bayesian computation**



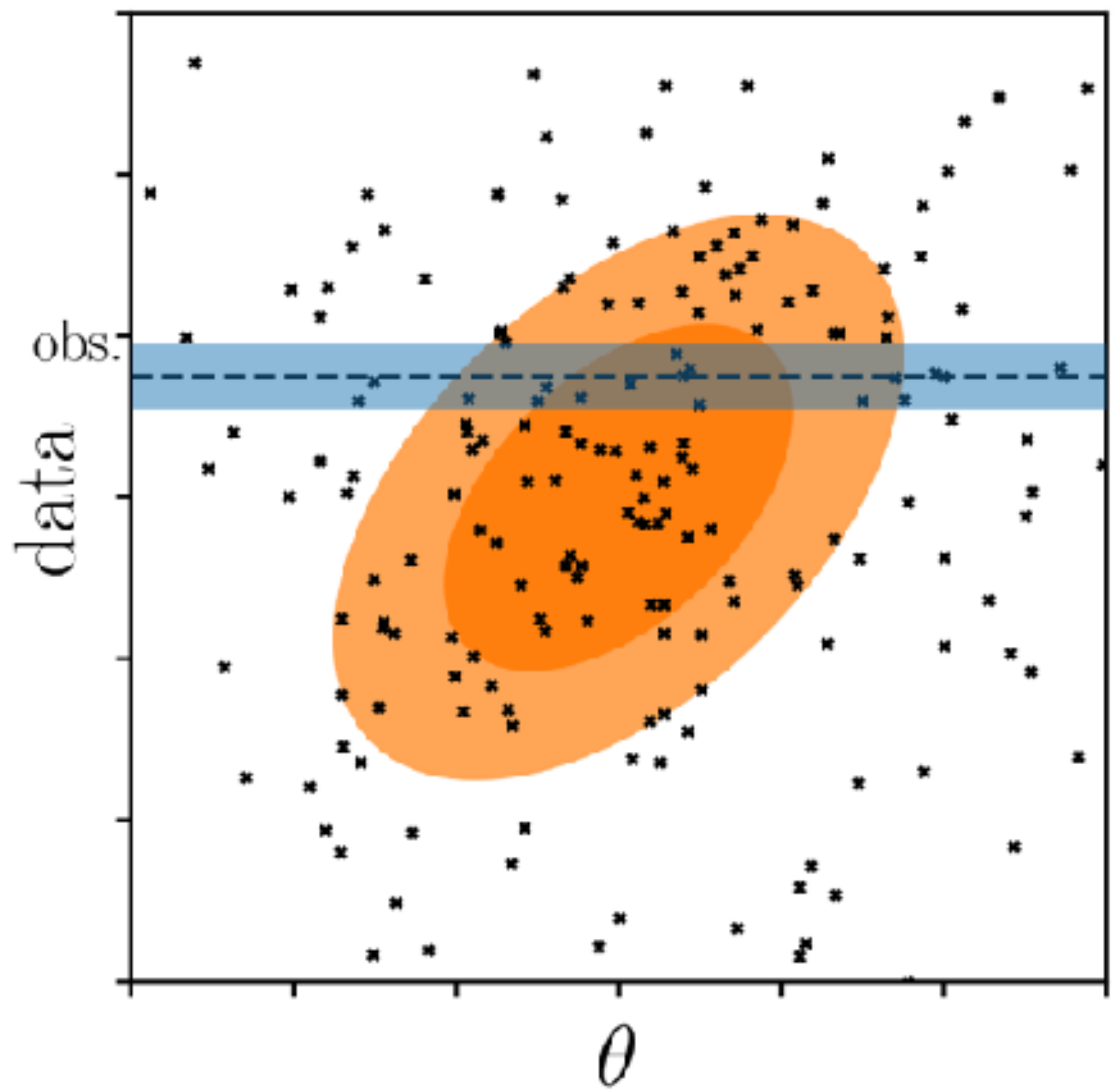
sample θ' from the prior $p(\theta)$ and run the simulation on every θ'

one way to do likelihood-free inference is by *brute force* —
approximate bayesian computation



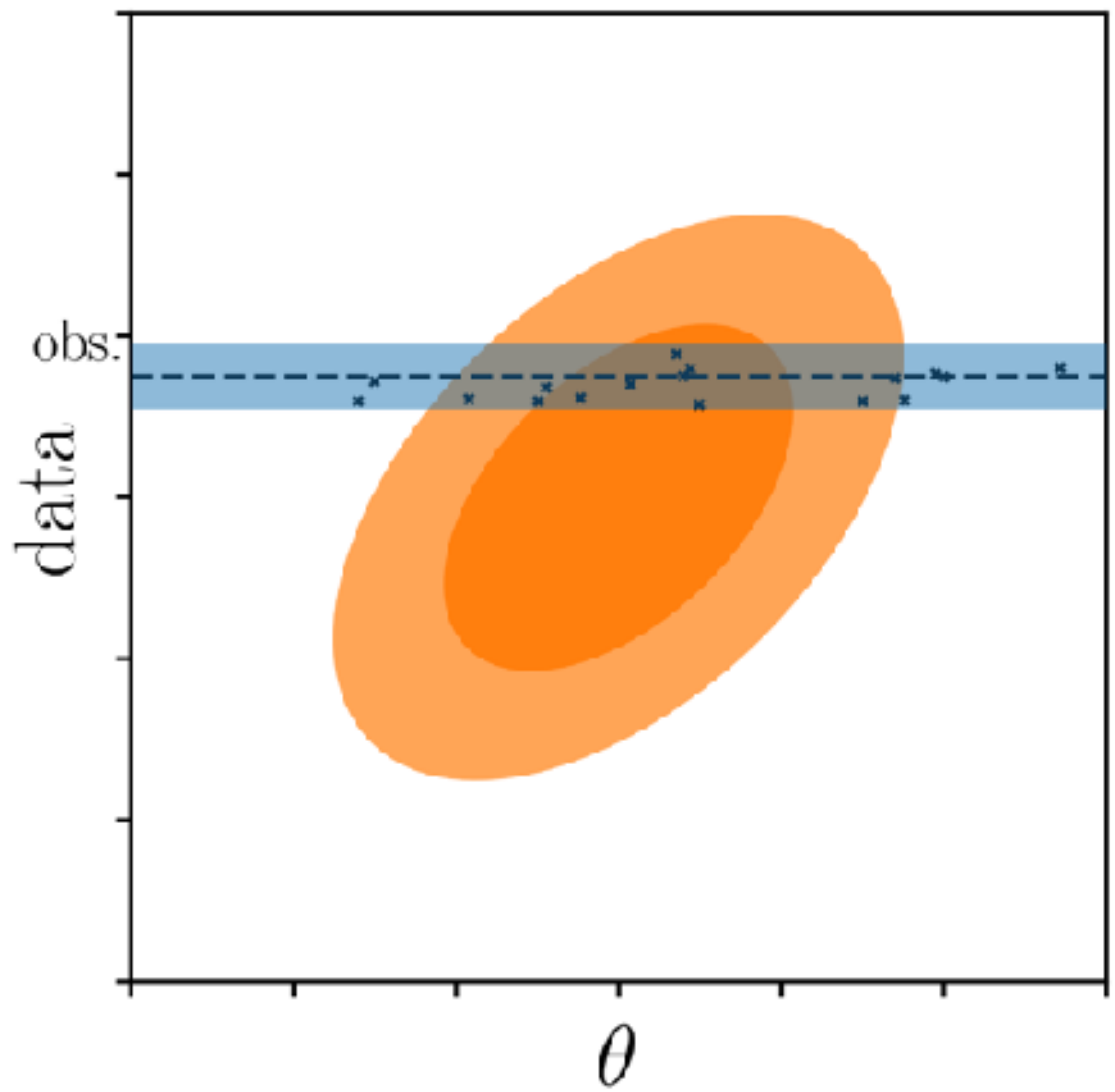
sample θ' from the prior $p(\theta)$ and run the simulation on every θ'

one way to do likelihood-free inference is by *brute force* —
approximate bayesian computation

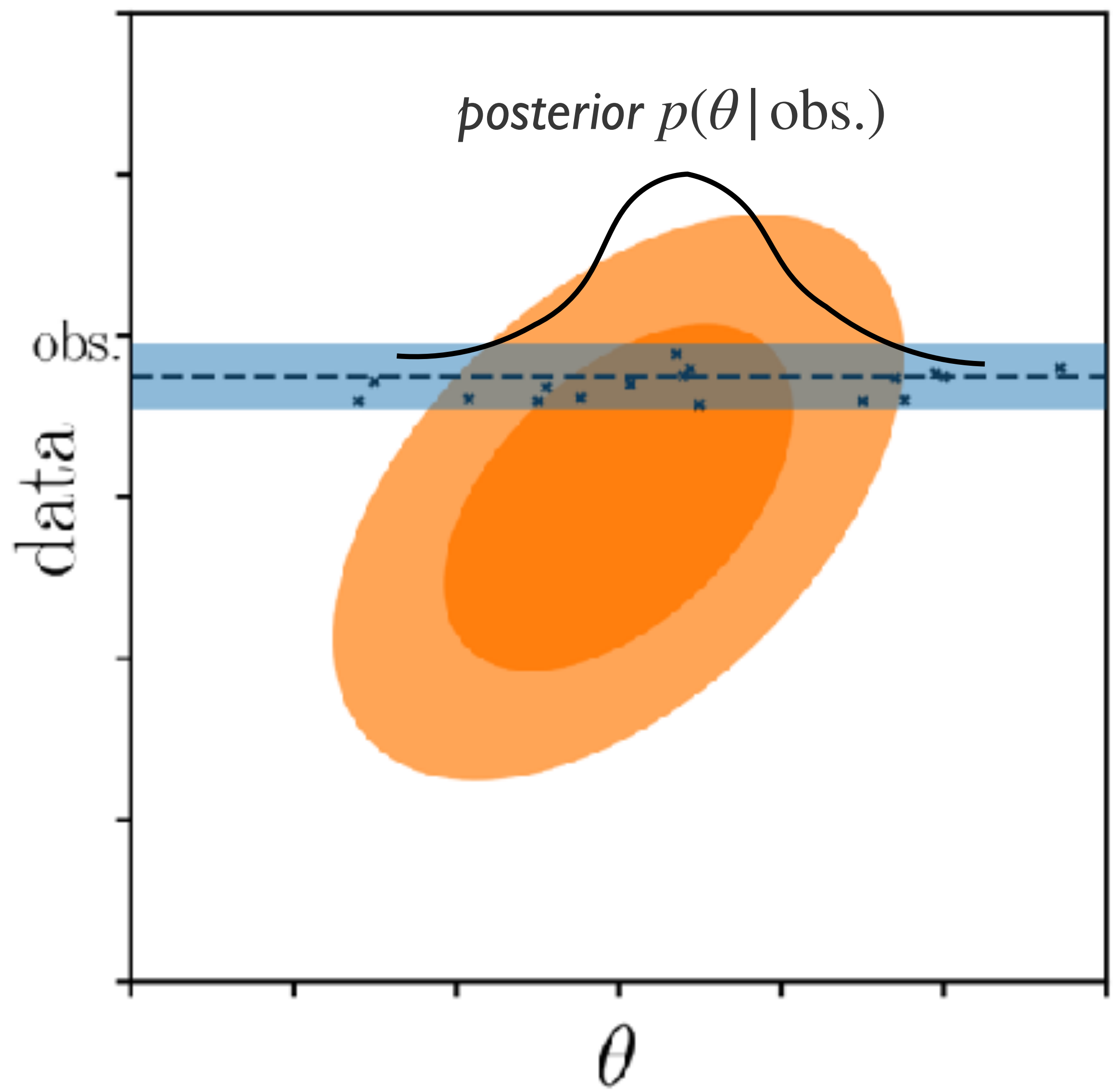


only keep simulations *close* to the observations

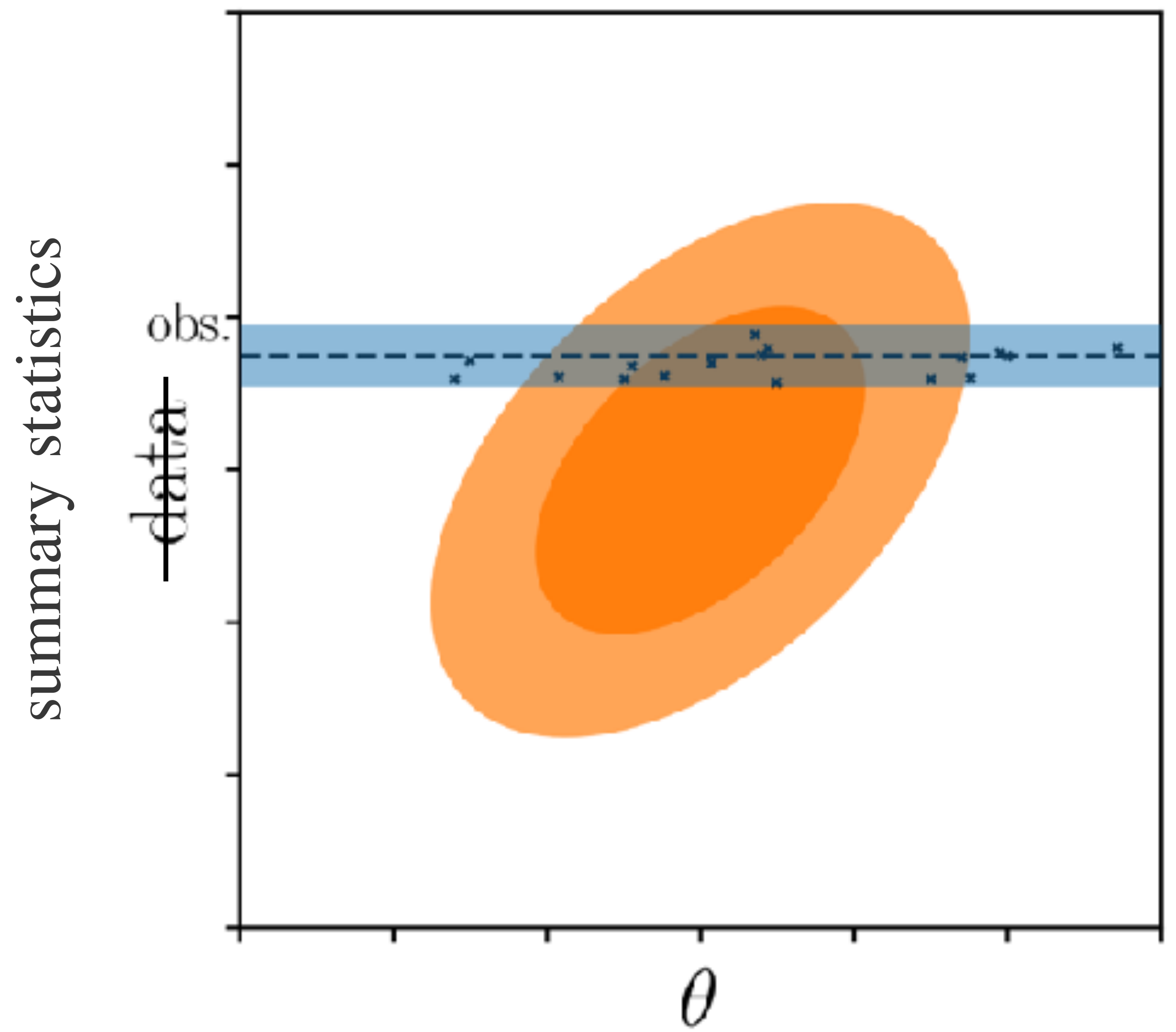
one way to do likelihood-free inference is by *brute force* —
approximate bayesian computation



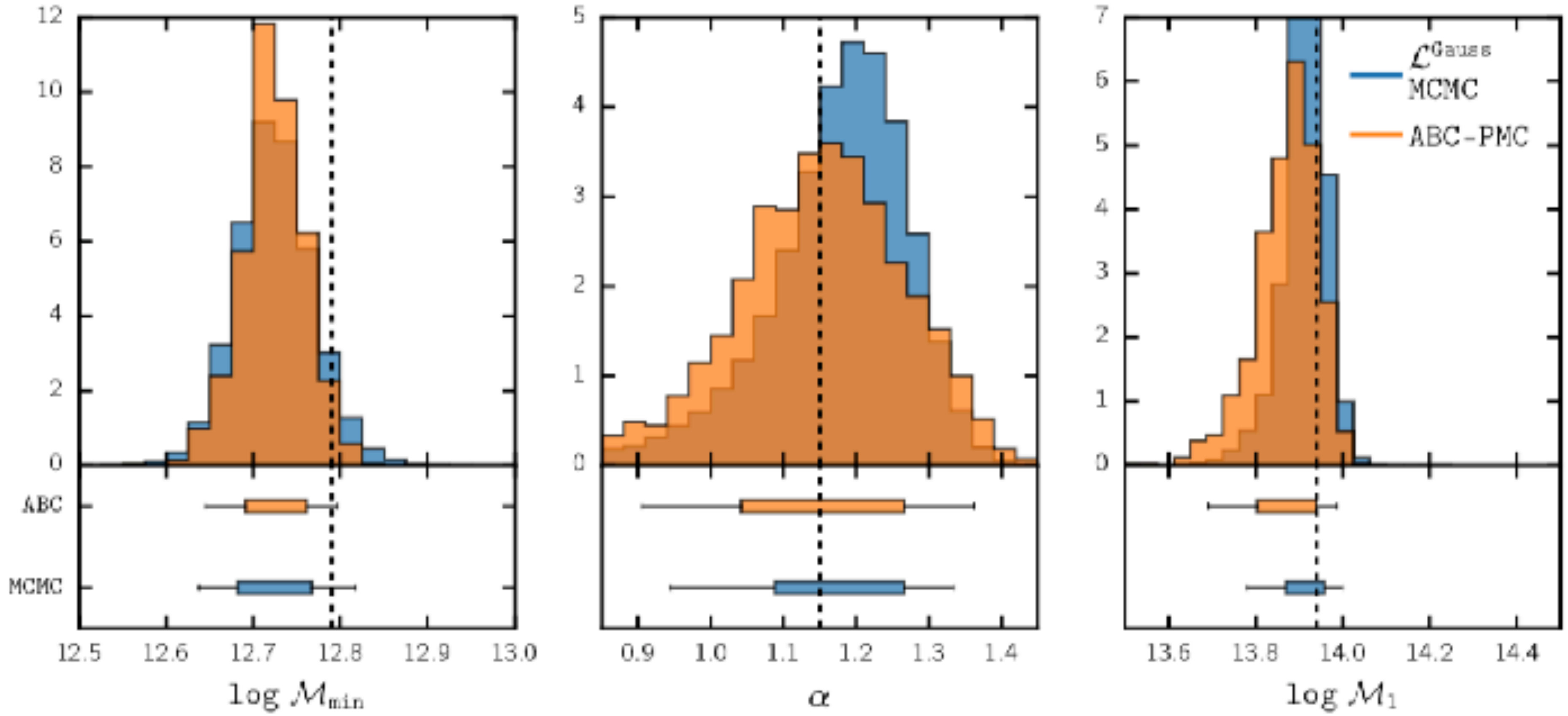
only keep simulations *close* to the observations one way to do likelihood-free inference is by *brute force* — **approximate bayesian computation**



the posterior can be approximated from these simulations more efficiently ABC using **summary statistics** (*e.g.* ξ, w_p, P_ℓ, B) and better sampling (*e.g. population monte carlo*)

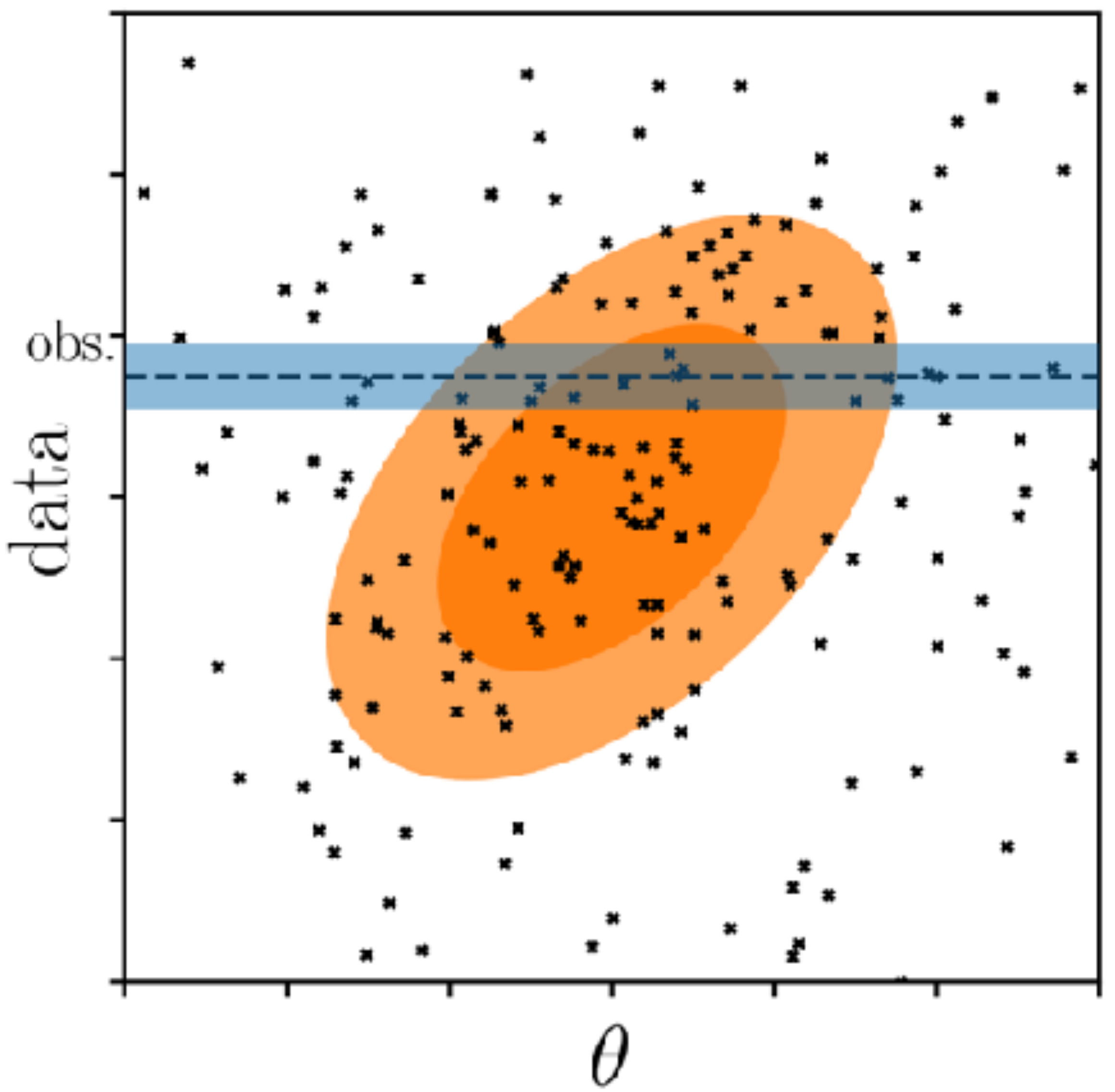


Hahn+(2017) — HOD analysis using **ABC-PMC** with 2PCF and *group multiplicity function*



Hahn+(2017b)

ABC throws out most of the simulations



each simulation is *a sample drawn from* $p(\mathrm{data}, \theta)$

each simulation is *a sample drawn from* $p(\text{data}, \theta)$, estimating $p(\text{data}, \theta)$ is a **density estimation** problem

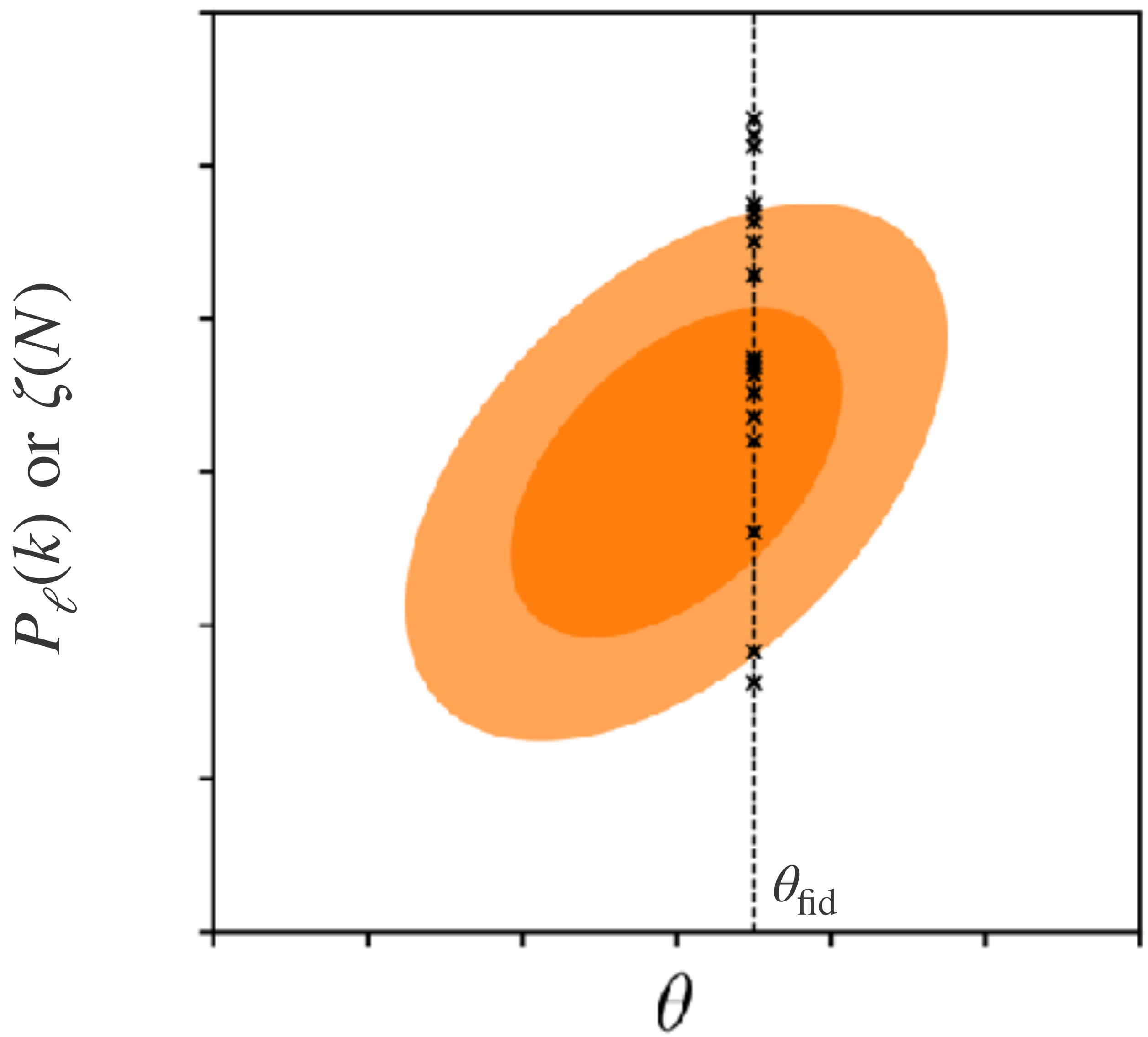
e.g. kernel density estimation, Gaussian mixture models, normalizing flows

Hahn (2019b) — can we solve the density estimation problem *using simulations for estimating the covariance matrix?*

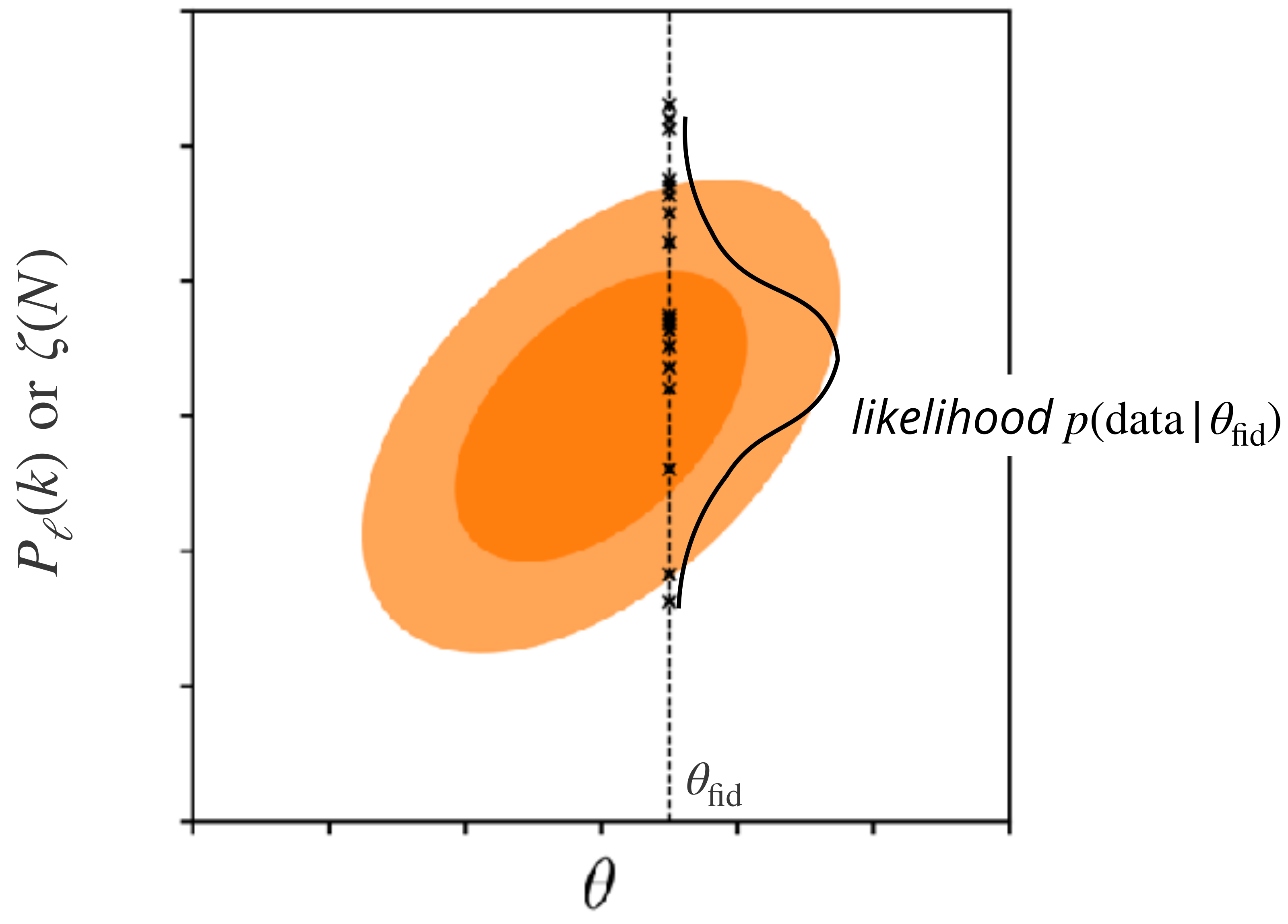
Hahn (2019b) — can we solve the density estimation problem *using simulations for estimating the covariance matrix?*

Beutler+(2017) RSD analysis using BOSS DR12 $P_\ell(k)$ — 4000 mocks

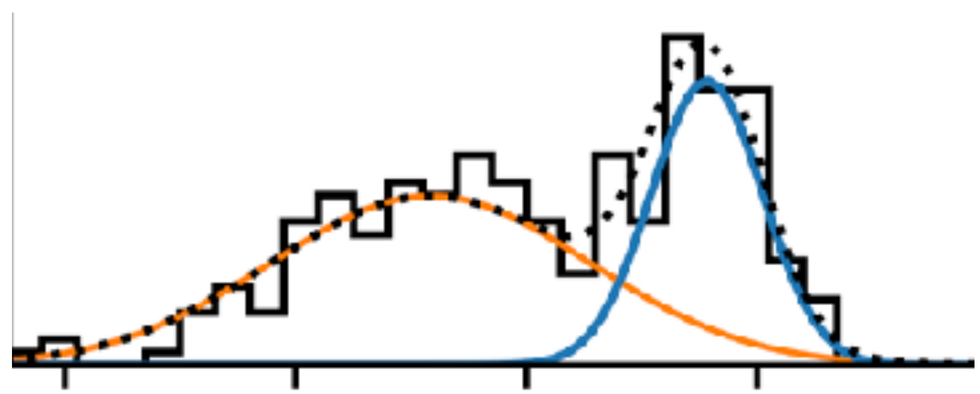
Sinha+(2017) HOD analysis using SDSS DR7 $\zeta(N)$ — 20,000 mocks in both analyses we have **1000s** $p(\mathrm{data}, \theta)$ **samples** at θ_{fid}



in both analyses we have **1000s** $p(\mathrm{data}, \theta)$ **samples** at θ_{fid} so we can use the mocks to directly estimate the $p(\mathrm{data}\,|\,\theta)$ *distribution*



in Sinha+(2017) we have 20,000 mocks for an 8D $\zeta(N)$ likelihood distribution — *Gaussian Mixture Model*

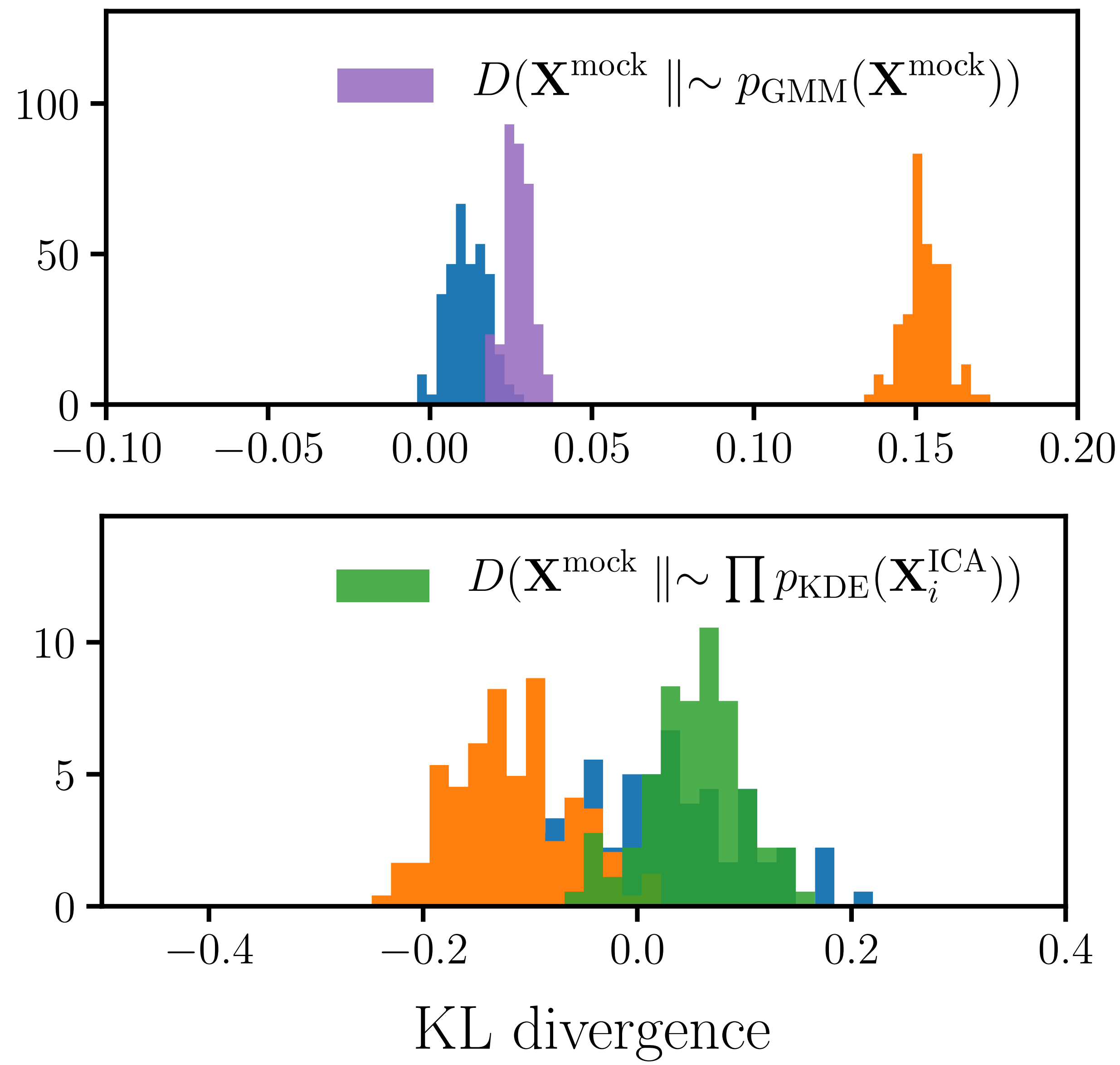


$$p(\zeta \,|\, \theta) = \sum_{i=1}^{k} \pi_i \; \mathcal{N}(x; \theta_i)$$

in Beutler+(2017) we have 4,000 mocks for an 37D $P_\ell(k)$ likelihood distribution — *independent component analysis*

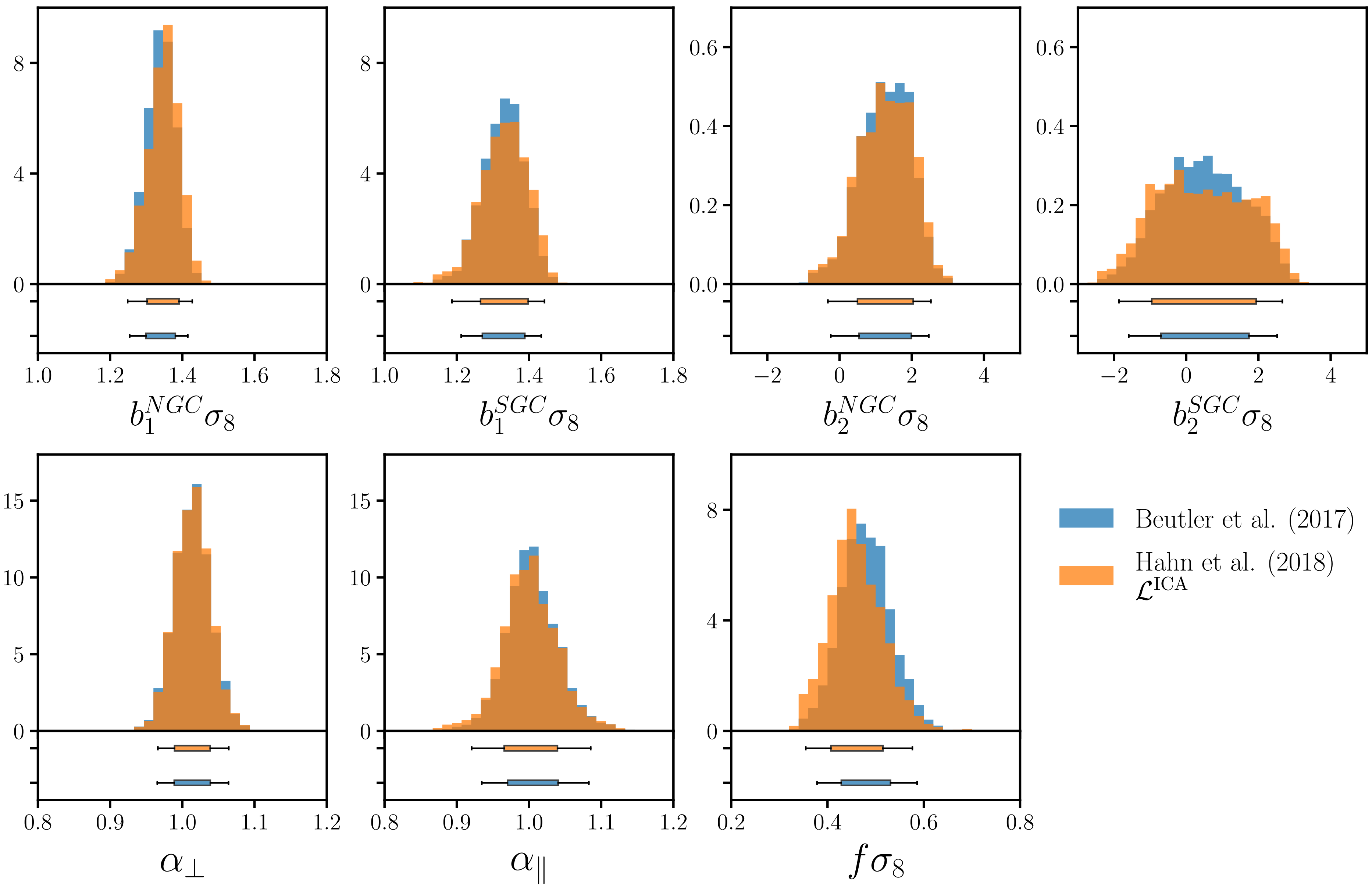
$$p(P_\ell | \theta) = \prod_{n=1}^{37} p_{x_n^{\rm IC}}(x)$$

37D distribution $\longrightarrow$ *37 independent 1D distributions* density estimation $\zeta(N)$ and $P_\ell(k)$ likelihoods are **more accurate** than the standard likelihoods



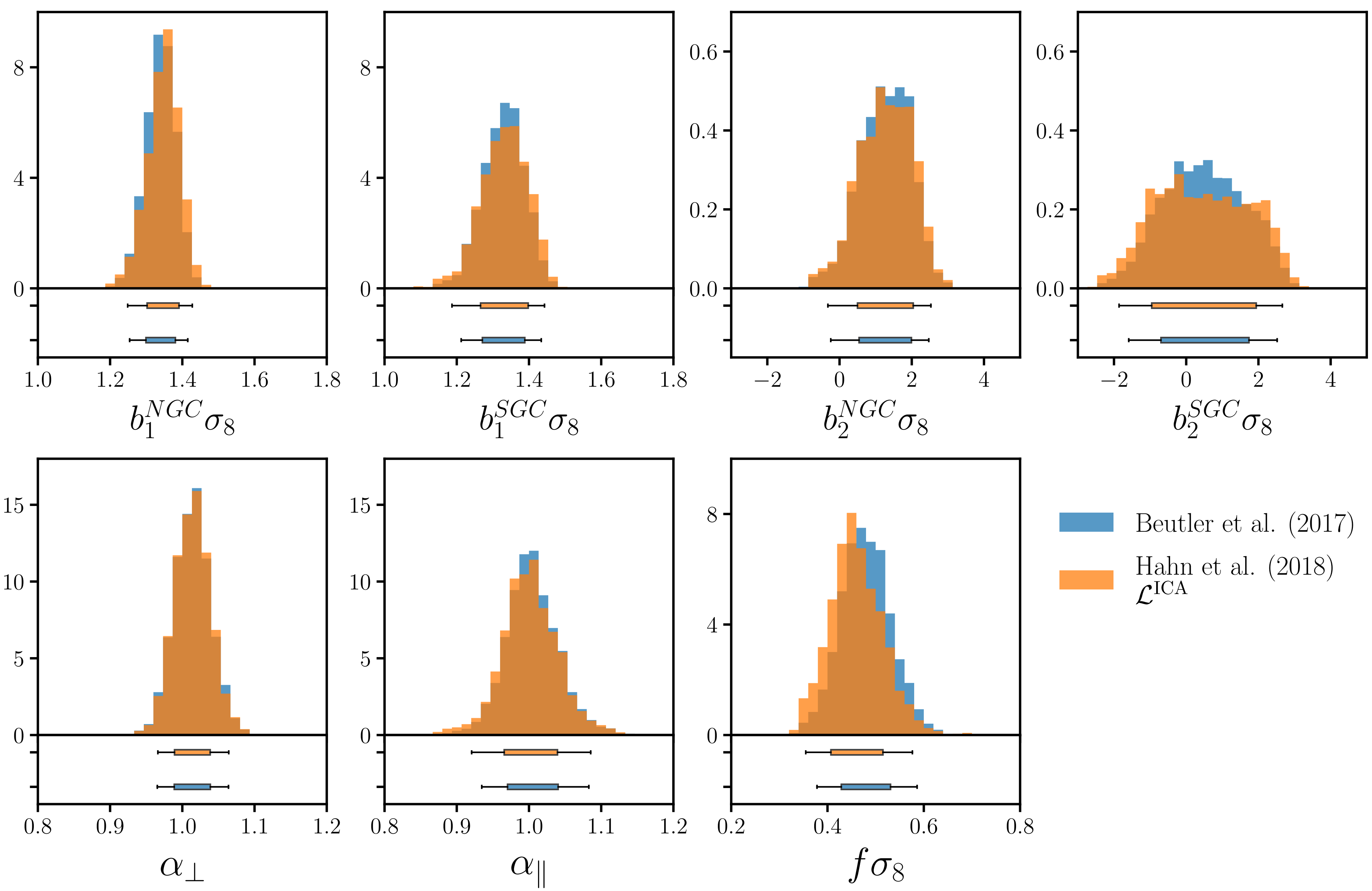
Hahn+(2019b)

non-Gaussianity of the $P_\ell(k)$ likelihood impacts the Beutler+(2017) posteriors — 0.44σ **shift in $f\sigma_8$ constraints**



Hahn+(2019b)

likelihood-free inference is possible with ~1000s of simulations!



Hahn+(2019b)

still only the ***tip of the LFI iceberg!***

DELFI with NDE (Alsing+2019)
ABC with Conditional Density Estimation (Izbicki+2018)
Sequential Neural Posterior Estimation (Lueckmann+2019)
Bayesian Optimization LFI (Gutmann & Corannder 2016)
Inference Aware Neural Optimization (de Castro & Dorigo 2018)
...

~~$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*~~

theoretical limitations on nonlinear scales

systematic effects

~~$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*~~

~~*theoretical limitations on nonlinear scales*~~

systematic effects

LFI is *fully simulation-based* so we can **probe nonlinear scales with *N* body simulations**

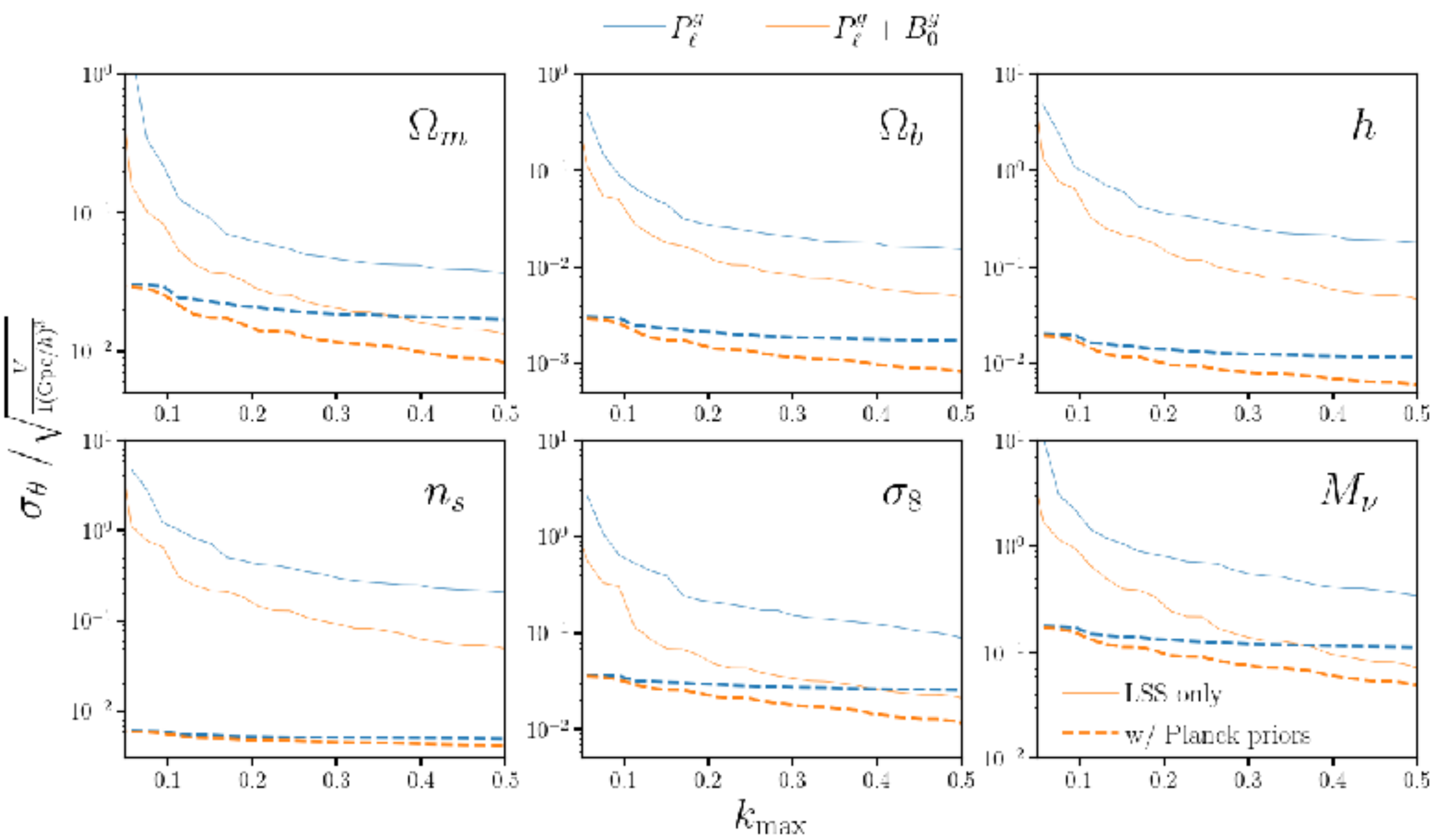
~~$B(k_1, k_2, k_3 < 0.5\ h/\mathrm{Mpc})$ *is 1898 dimensional*~~

~~*theoretical limitations on nonlinear scales*~~

~~*systematic effects*~~

LFI is *fully simulation-based* so we can probe nonlinear scales with *N* body simulations and **forward model systematics** next: B BOSS reanalysis with *data compression* and *likelihood-free inference* to constrain M_ν

next: B BOSS reanalysis with *data compression* and *likelihood-free inference* to constrain M_ν **then to DESI and PFS** all the forecasts were for a $1\ h^{-3}Gpc^3$ volume



Hahn+(2021)

DESI forecasts

BGS — $V_{\mathrm{BGS}} = 3.44\ (h^{-1}\mathrm{Gpc})^3$	σ_{M_ν}
$P_\ell\ \ [k_{\max} = 0.5]$	$0.18\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.038\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \mathit{Planck}$	$0.029\,\mathrm{eV}$

DESI — $V_{\mathrm{DESI}} \sim 50\ (h^{-1}\mathrm{Gpc})^3$	
$P_\ell\ \ [k_{\max} = 0.5]$	$0.047\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.010\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \mathit{Planck}$	$0.0095\,\mathrm{eV}$
*$\mathit{Planck}\ +\ \mathrm{DESI}$	$\sim 0.02\,\mathrm{eV}$

DESI forecasts

BGS — $V_{\mathrm{BGS}} = 3.44\ (h^{-1}\mathrm{Gpc})^3$	σ_{M_ν}
$P_\ell\ \ [k_{\max} = 0.5]$	$0.18\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.038\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \textit{Planck}$	$0.029\,\mathrm{eV}$

DESI — $V_{\mathrm{DESI}} \sim 50\ (h^{-1}\mathrm{Gpc})^3$	
$P_\ell\ \ [k_{\max} = 0.5]$	$0.047\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.010\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \textit{Planck}$	$0.0095\,\mathrm{eV}$
*$\textit{Planck}\ +\ \mathrm{DESI}$	$\sim 0.02\,\mathrm{eV}$

$> 4\sigma$ distinction between

normal (*0.057* eV) or inverted (*0.097* eV) hierarchies

DESI forecasts

BGS — $V_{\mathrm{BGS}} = 3.44\ (h^{-1}\mathrm{Gpc})^3$	σ_{M_ν}
$P_\ell\ \ [k_{\max} = 0.5]$	$0.18\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.038\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \textit{Planck}$	$0.029\,\mathrm{eV}$

DESI — $V_{\mathrm{DESI}} \sim 50\ (h^{-1}\mathrm{Gpc})^3$	
$P_\ell\ \ [k_{\max} = 0.5]$	$0.047\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]$	$0.010\,\mathrm{eV}$
$P_\ell + B_0\ \ [k_{\max} = 0.5]\ +\ \textit{Planck}$	$0.0095\,\mathrm{eV}$
*$\textit{Planck}\ +\ \mathrm{DESI}$	$\sim 0.02\,\mathrm{eV}$

DESI will have much *higher number densities* than our forecast the **Molino Suite** — *75,000 publicly available* galaxy mocks to quantify information content of any observable!

the ***galaxy bispectrum*** significantly improves parameter constraints

$\gtrsim \mathbf{3} \times$ *tighter* Ω_m, Ω_b, h, n_s, σ_8 ; $\gtrsim \mathbf{5} \times$ *tighter* M_ν

there's significant constraining power in the **nonlinear regime**

with **likelihood-free inference** we can exploit the *full constraining power of the bispectrum* to constrain M_ν

credit: desi.lbl.gov