Non-Gaussian Statistics
&
Non-Gaussian Tensions

Marco Raveri

University of Genova

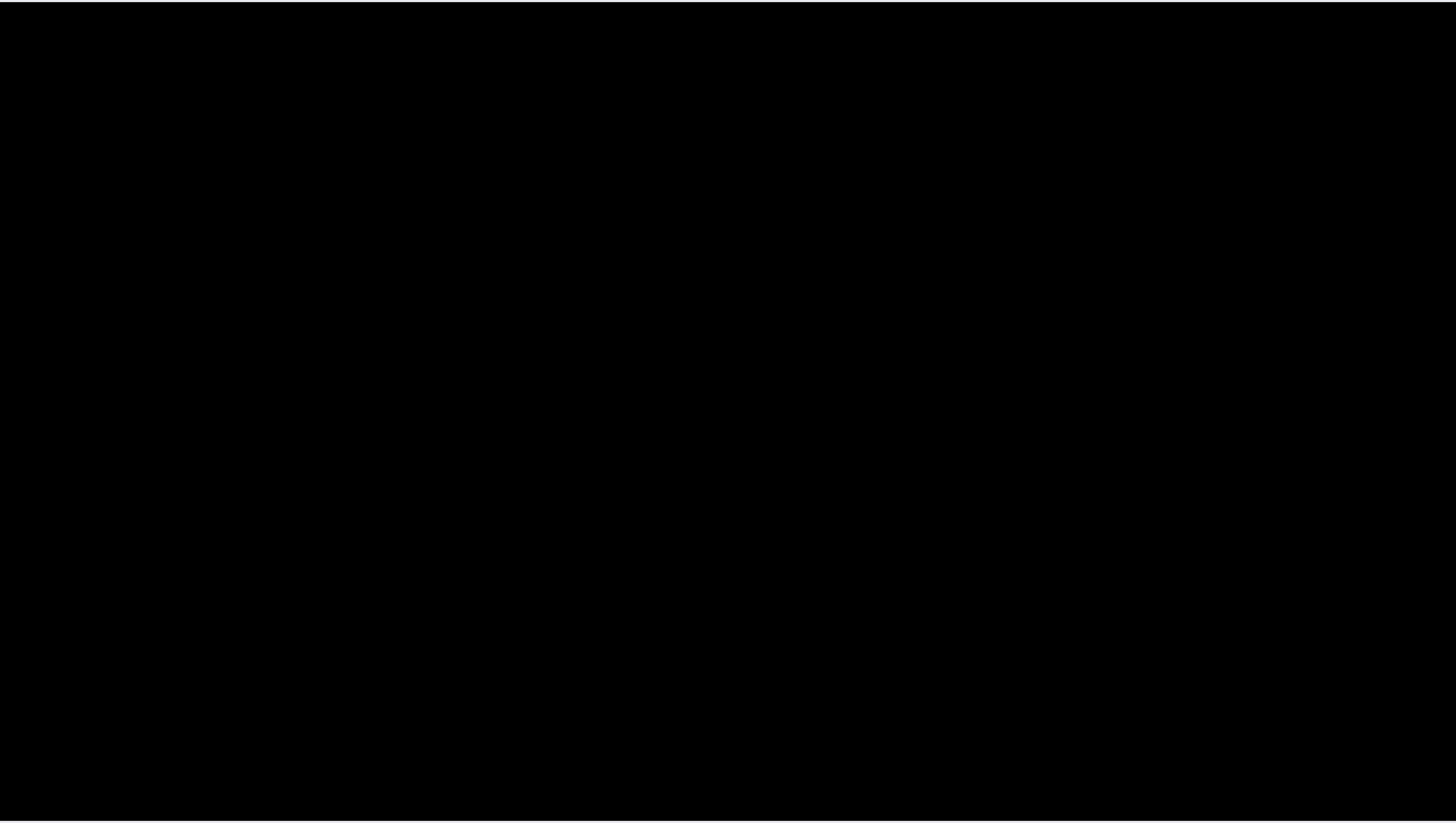
Plan

Cosmology results from non-Gaussian statistics
+
Their agreement with other probes

1. Weak gravitational lensing observables and the Dark Energy Survey
2. agreement and disagreement between different experiments
3. example use with DES Y3 data

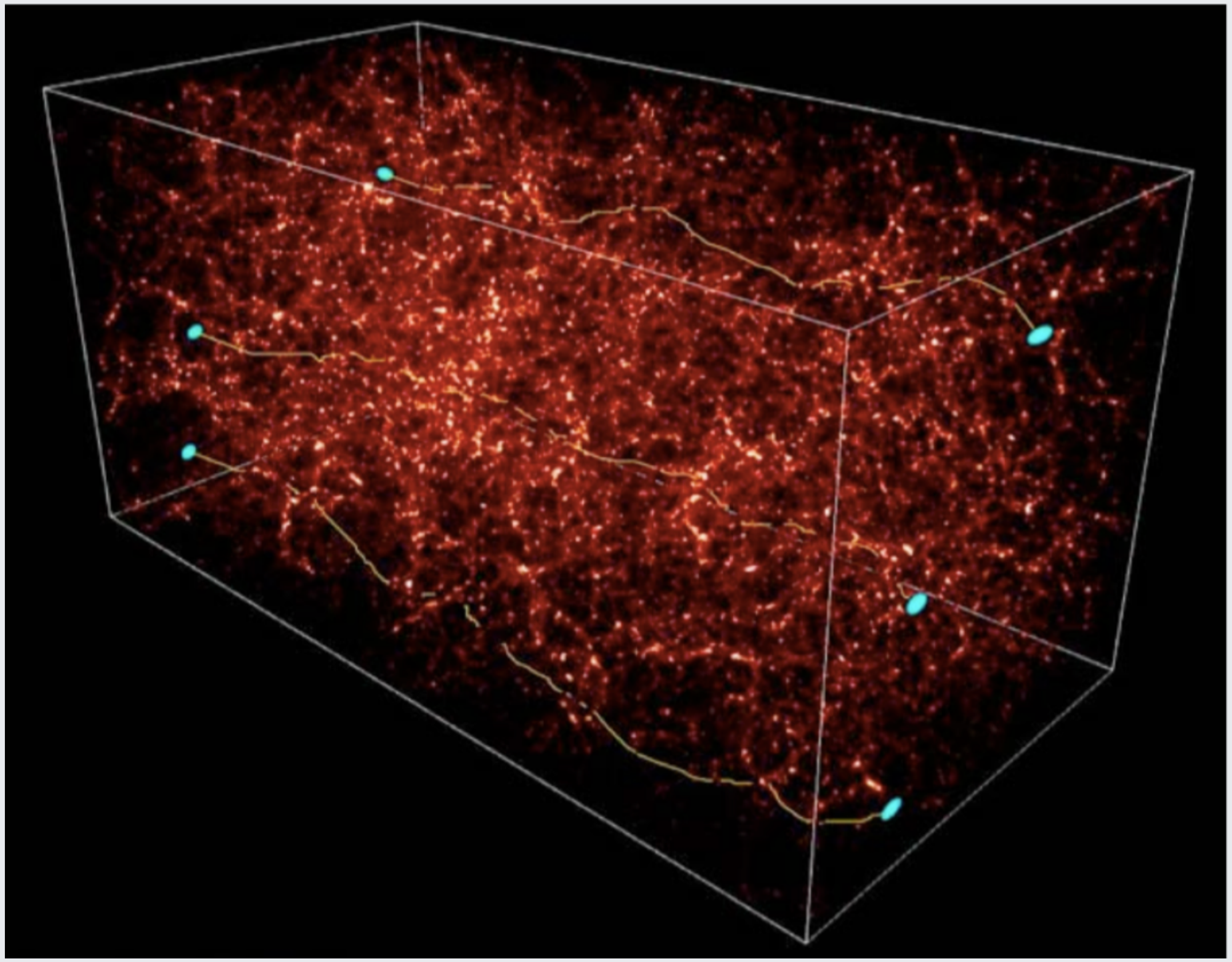
Marco Gatti
(Upenn)

Structures in the Universe



(Illustris simulation)

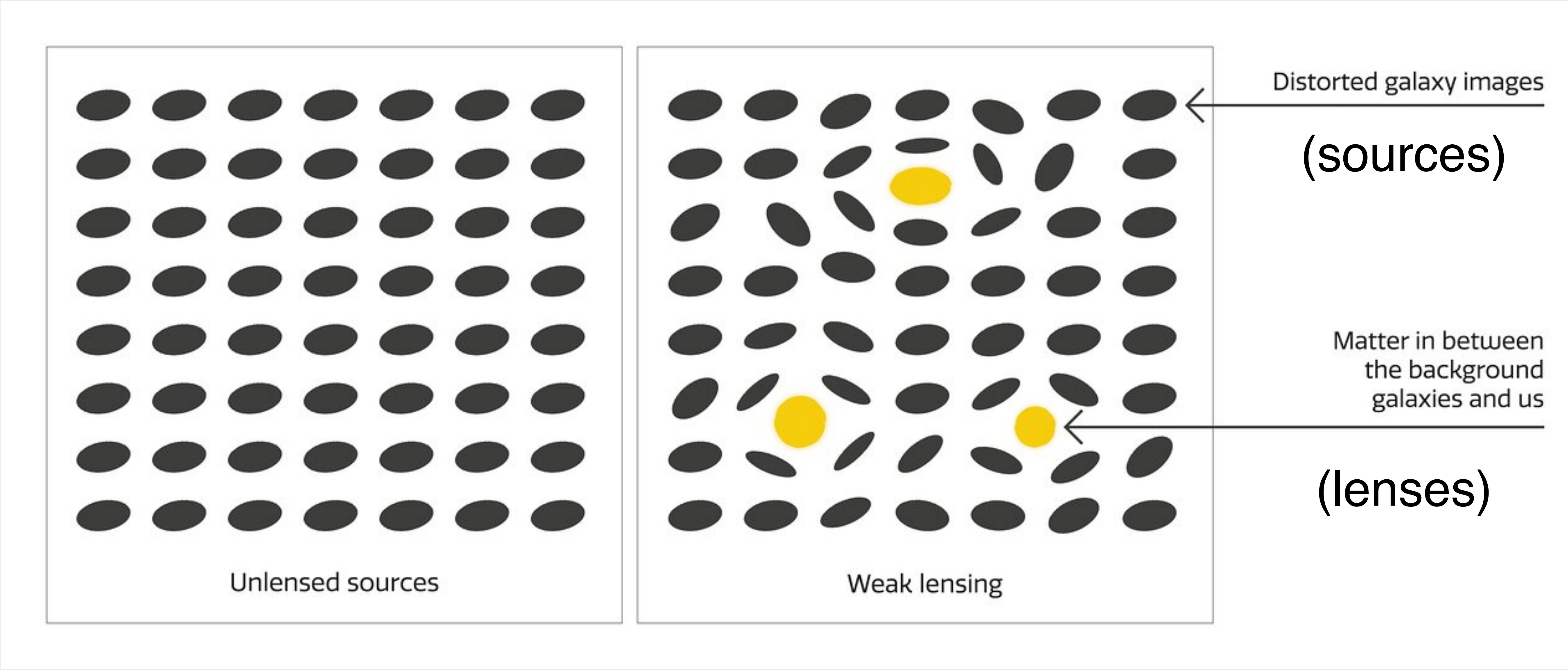
Weak Gravitational Lensing



Due to the **Large Scale Structure** of the Universe,
the path followed by the light emitted by distant galaxies will appear distorted

Gravitational lensing allows to probe the matter distribution (mostly dark)

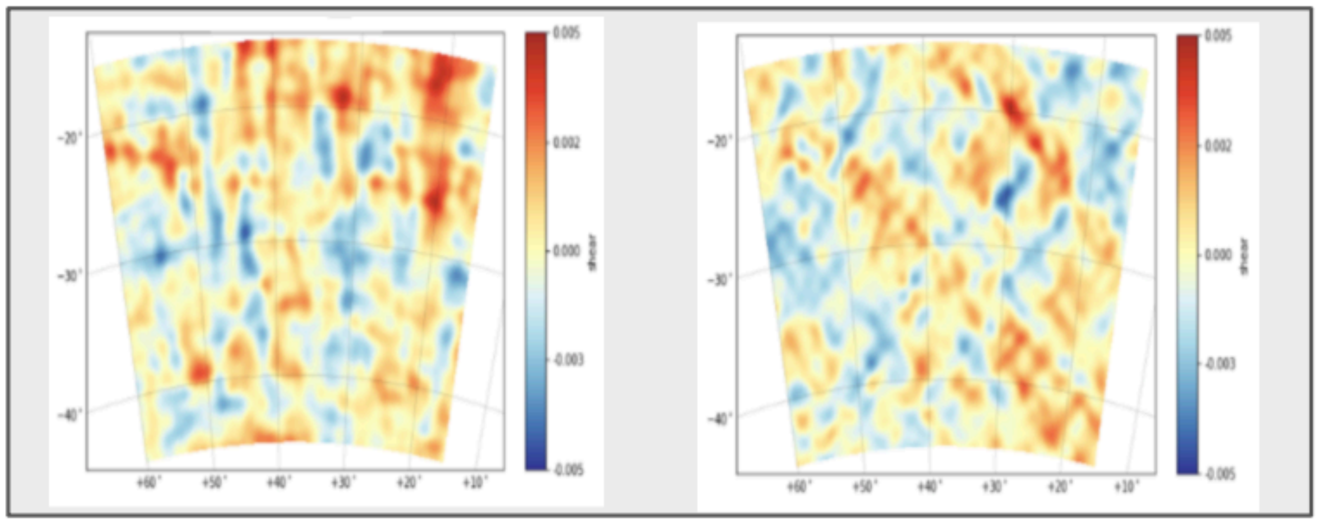
Weak Gravitational Lensing



Weak Gravitational Lensing

observable!

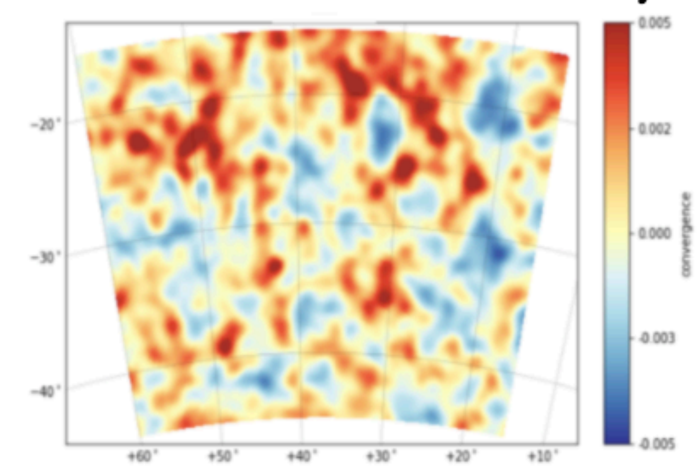
Using measured galaxies ellipticity, we can estimate the shear field (2 components)



Mass Map reconstruction

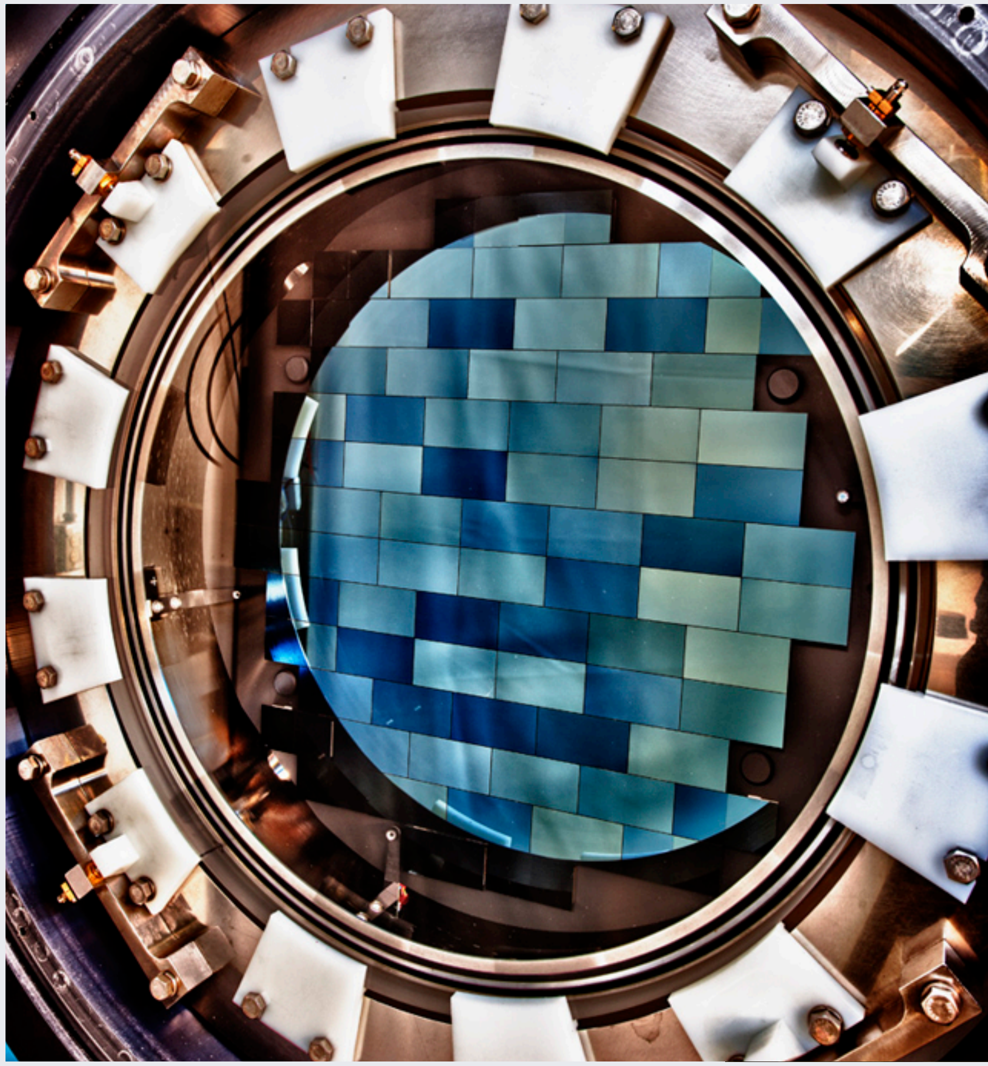
(projected) WL mass map (or convergence)

Not observable directly



(Credit: M. Gatti)

The Dark Energy Survey (DES)



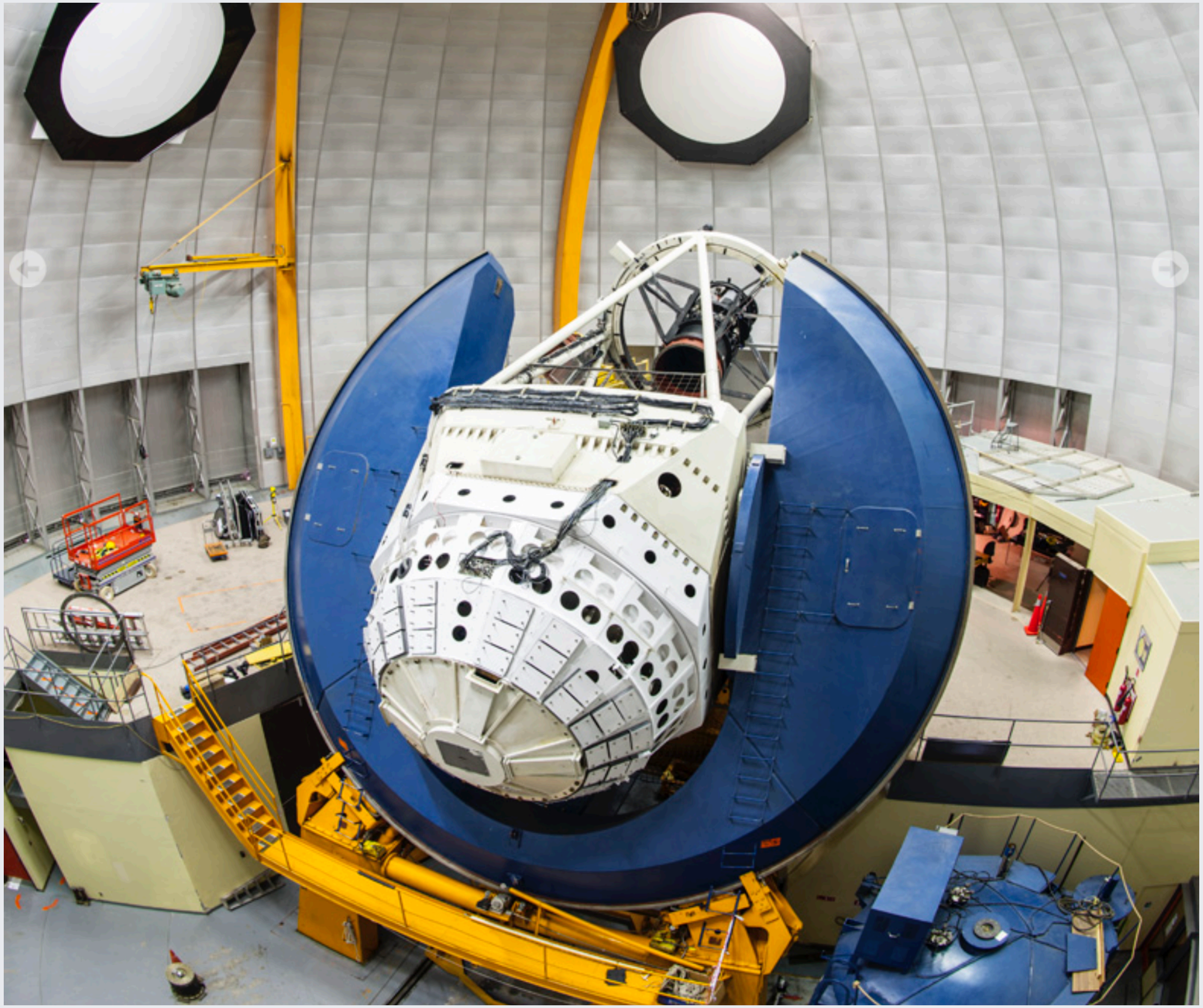
570-Megapixel digital camera, DECam, mounted on the Blanco 4-meter telescope at Cerro Tololo Inter-American Observatory (Chile).

~5k sq. deg. after 6 years of observations (one-fifth of the usable sky)

Shapes, photometric redshifts, positions of 300M galaxies

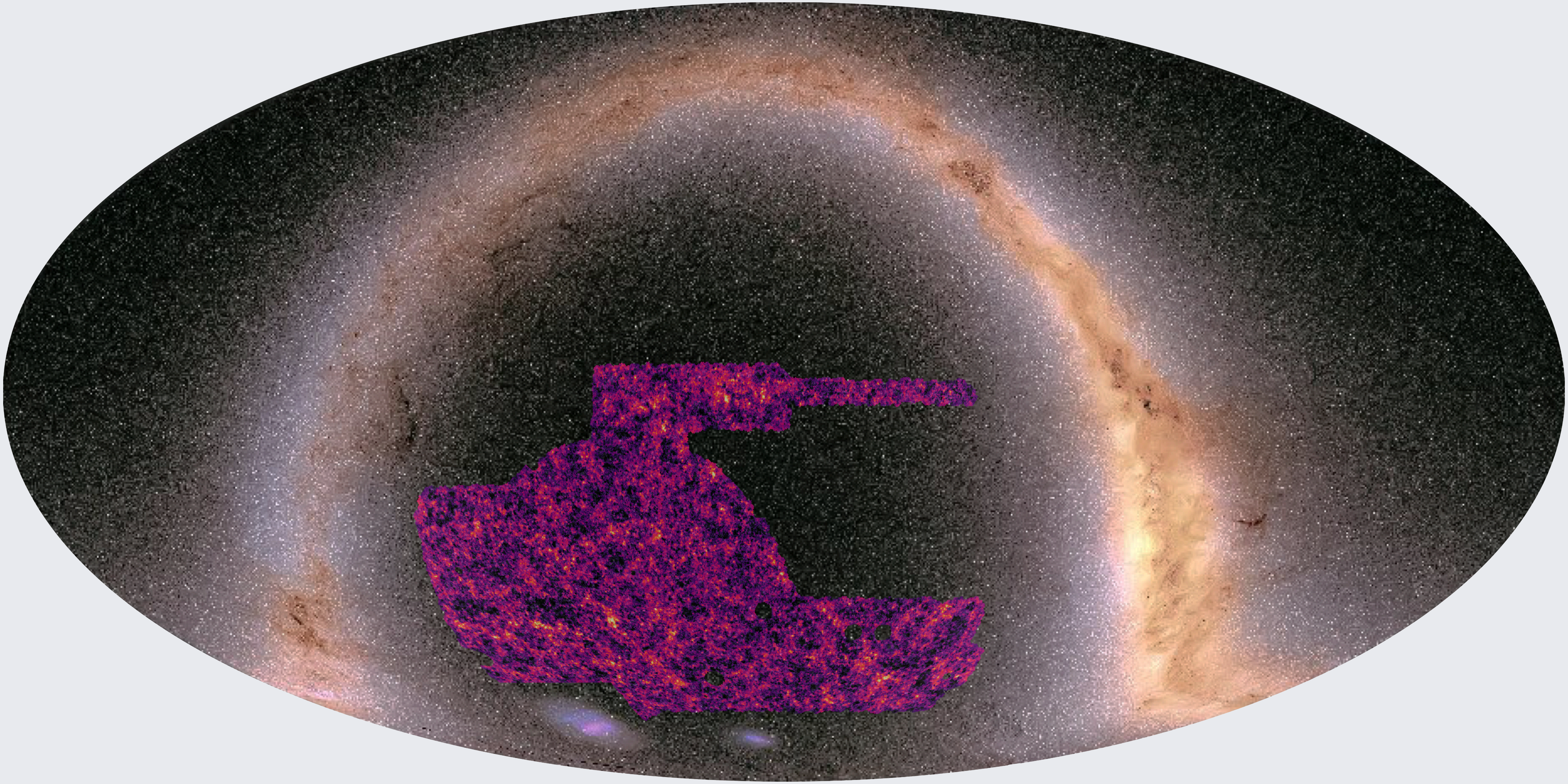
(Based on some work in DES)

The Dark Energy Survey (DES)



(Based on some work in DES)

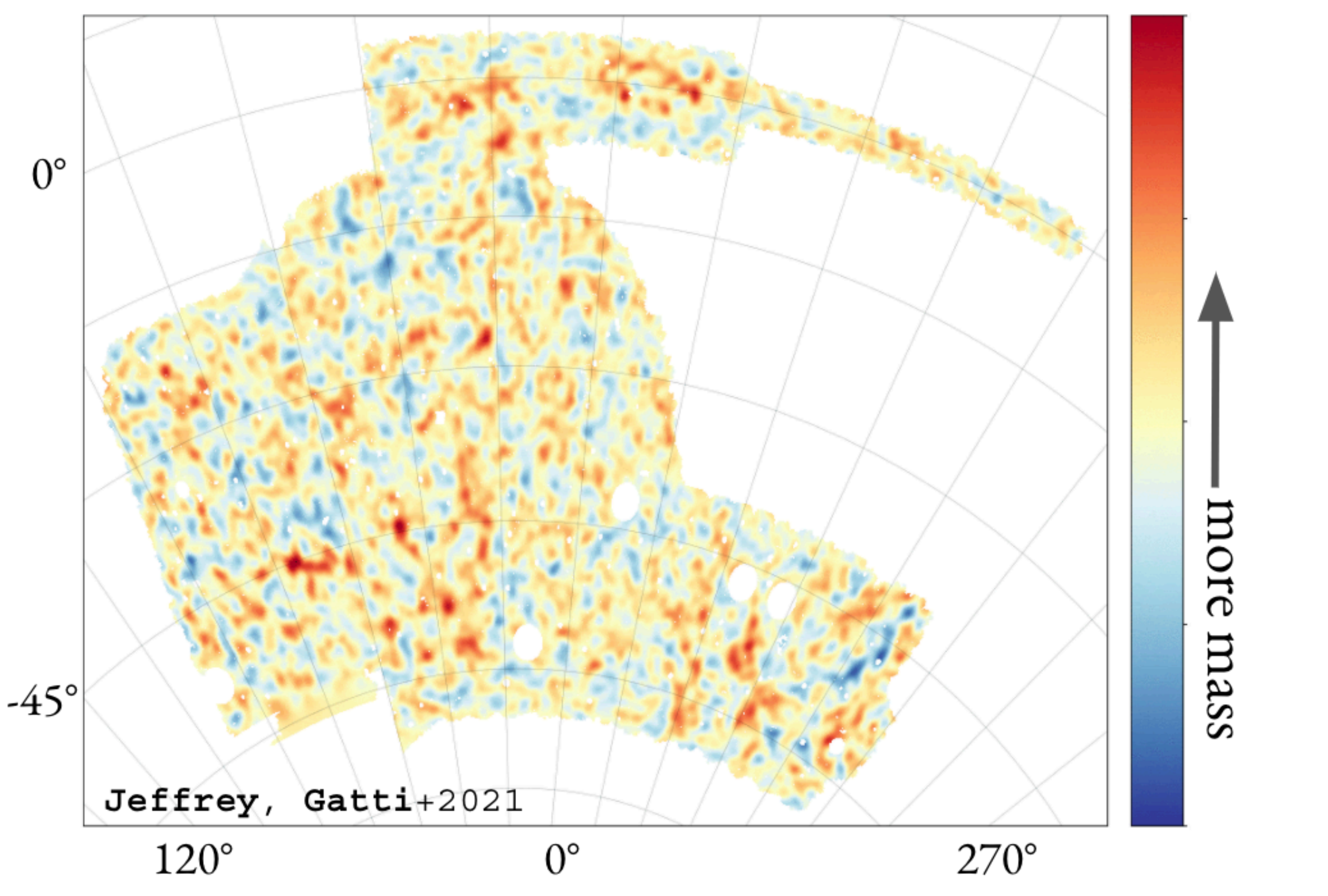
The Dark Energy Survey



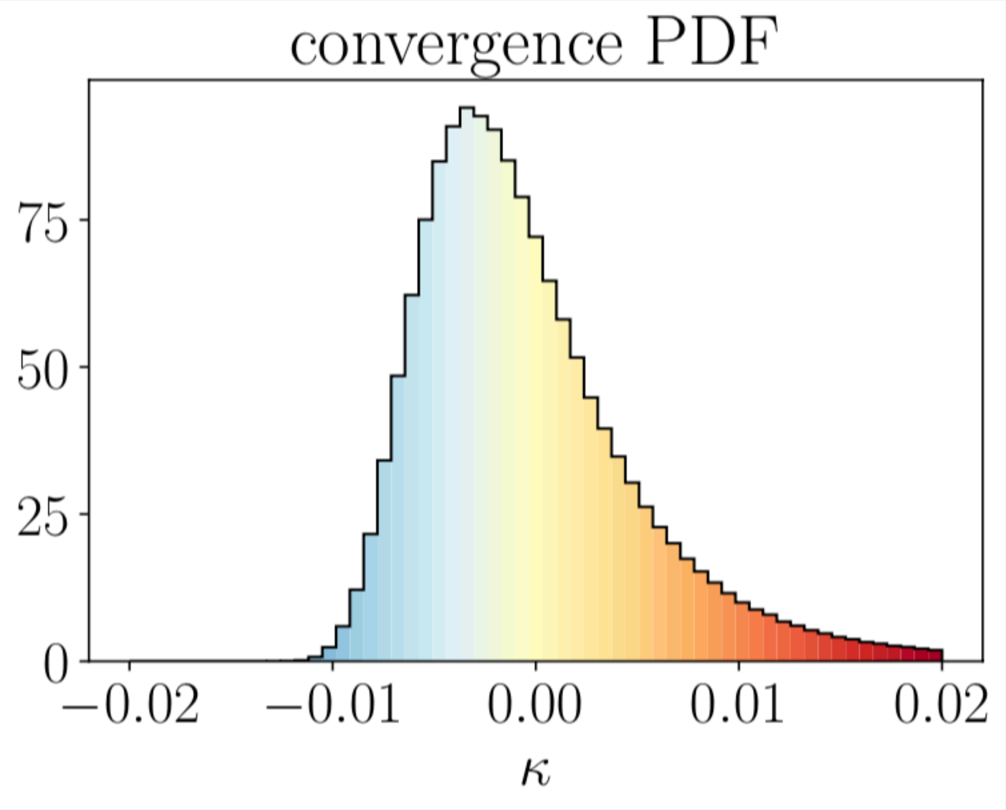
Largest map of the Dark Universe

(Based on some work in DES)

The DES mass map



5000 sq. degrees,
100 million galaxy shapes



The convergence field is **not Gaussian**

(Credit: M. Gatti)

Why looking for non-Gaussianity in LSS?

Gravitational fluid dynamics is non-linear physics

If initial conditions are Gaussian (only variance)
evolved distributions become non-Gaussian (more numbers)

↓

Extra numbers tell us how gravity clumped things together

Growing interest in weak lensing Non Gaussianity

Peaks statistics

(e.g. Kacprzak et al. 2016; Martinet et al. 2018; Peel et al. 2018; Shan et al. 2018; Ajani et al. 2020; Z¸rcher et al. 2021a,2021b..)

High order Moments

(Chang et al. 2018; Vicinanza et al. 2018; Peel et al. 2018; Gatti et al. 2020,2021…)

3pt correlation functions

(Takada & Jain 2003, 2004; Semboloni et al. 2011; Fu et al. 2014,Secco et al 2022...)

Minkowski functionals

(Kratochvil et al. 2012; Petri et al. 2015; Vicinanza et al. 2019; Parroni et al. 2020...)

Machine Learning

(Ribli et al. 2019; Fluri et al. 2018, 2019; Jeffrey et al. 2021a…)

Wavelet-based methods

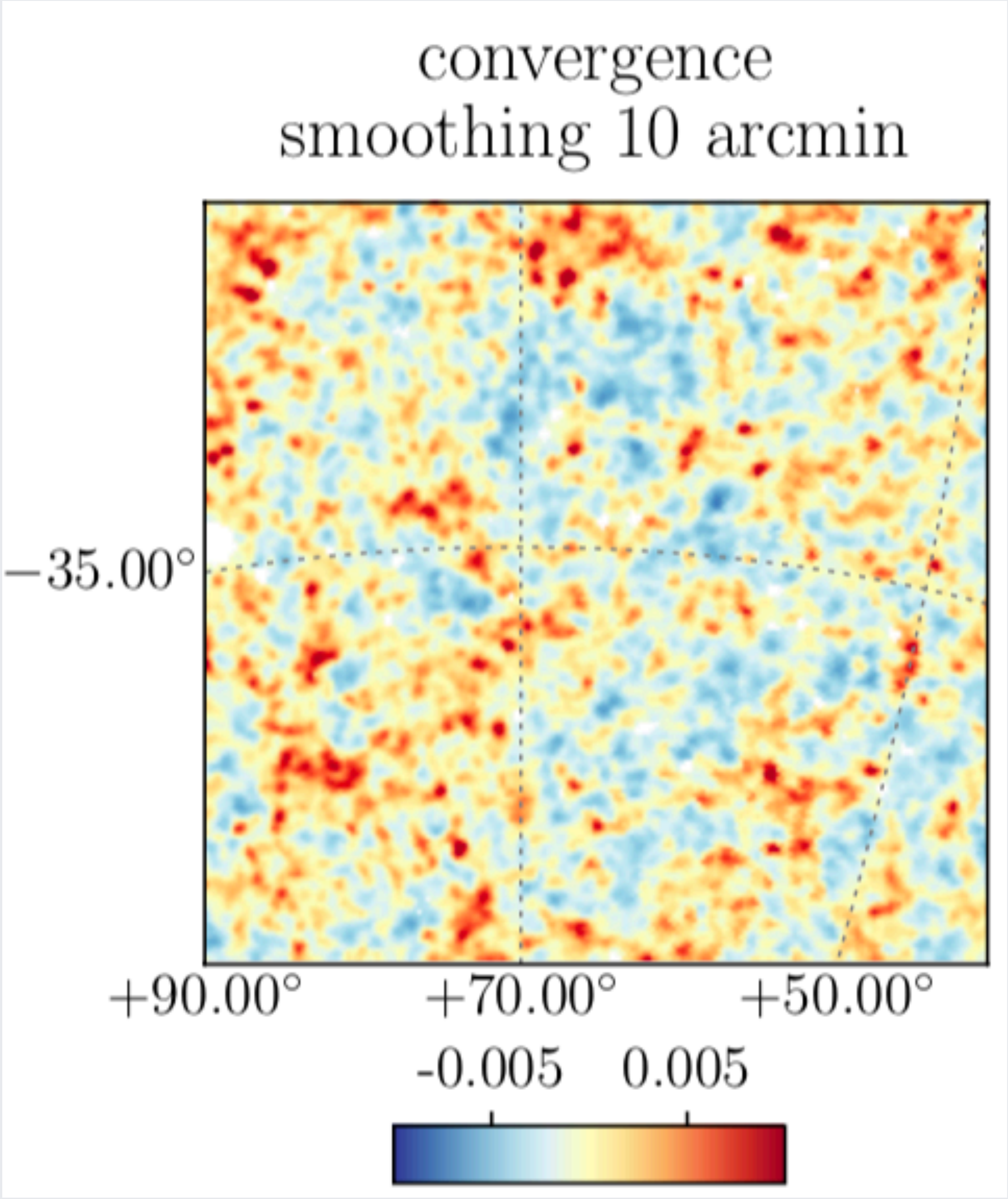
(Allys 2021, Cheng 2021, Gatti et al in prep.…)

Others (PDF,minima counts, L1-norm, k-nearest Neighbor distributions,....)

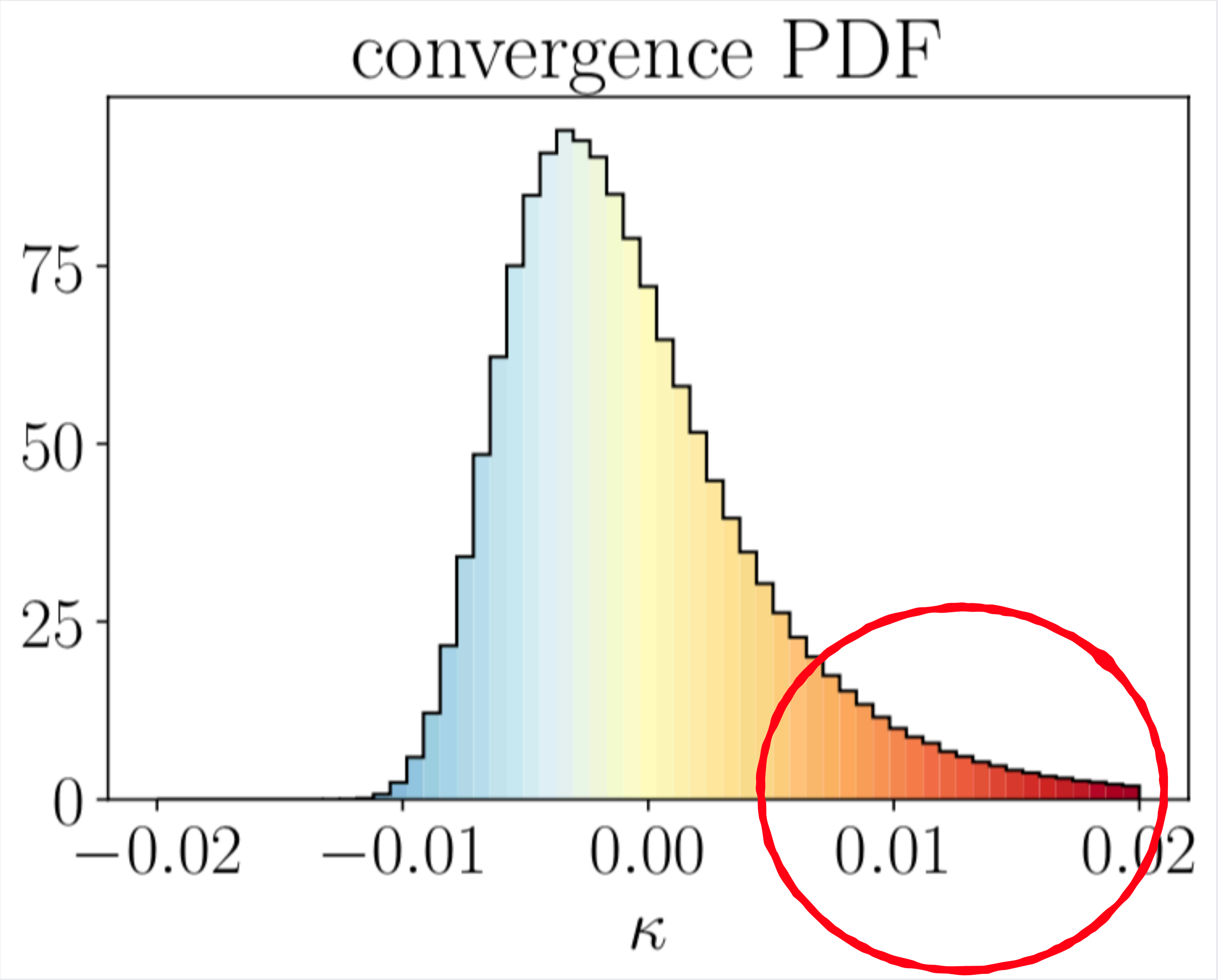
[DISCLAIMER: non-exhaustive!]

(Credit: M. Gatti)

Non-Gaussian statistics in DES: moments



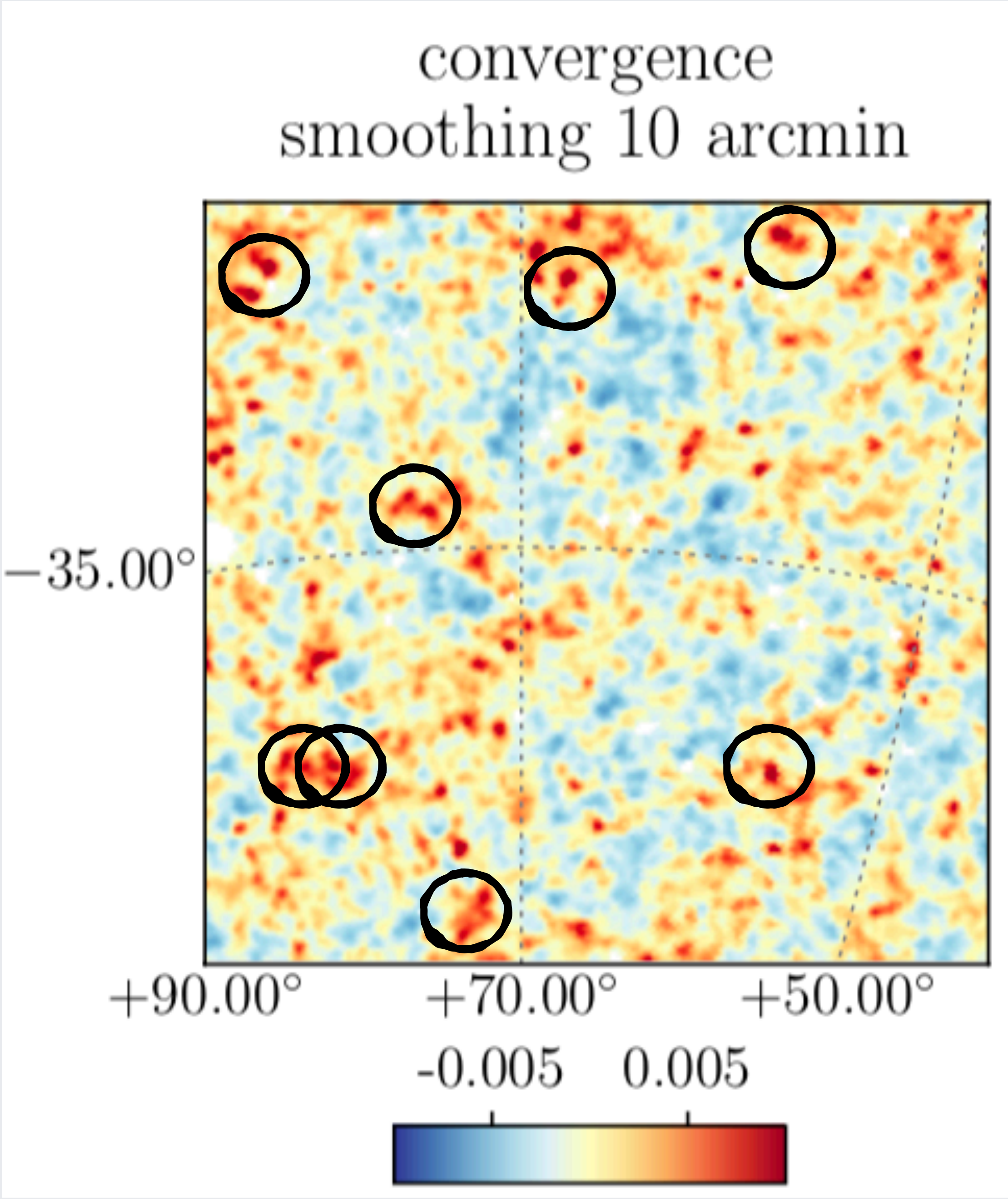
DES mass map



Observable: moments of the distribution at different smoothing scales

(DES Y3 moments analysis, Gatti+21, [2110.10141])

Non-Gaussian statistics in DES: peaks

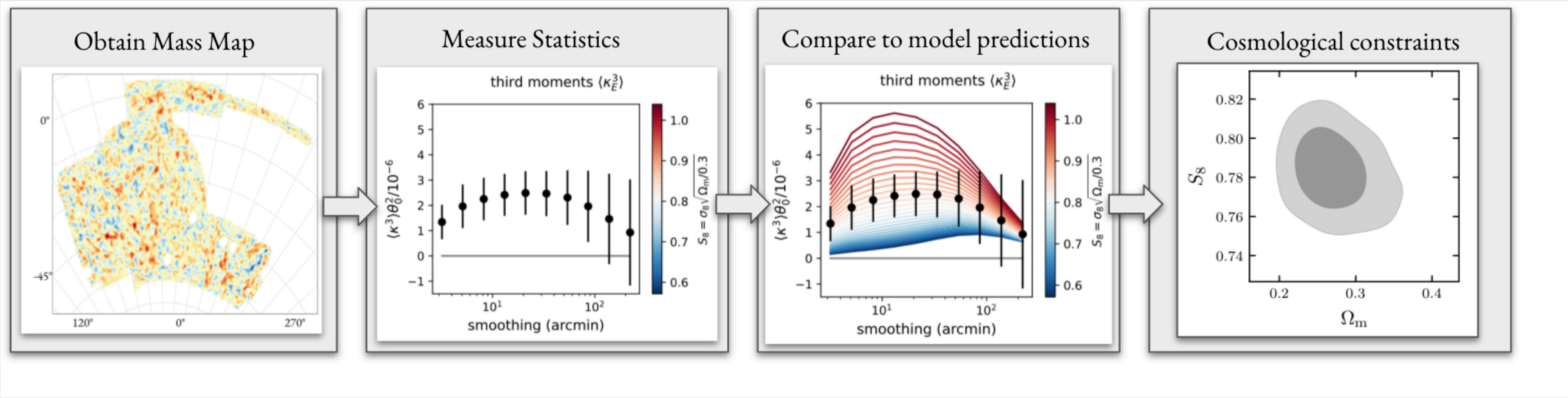


DES mass map

Observable: number of mass peaks at a given smoothing scale

(DES Y3 peaks analysis, Zuercher+22, [2110.10135])

From maps to cosmology



Two different strategies to model high order statistics

Analytical modelling

- complex to develop; not always feasible
- not computationally expensive
- adopted in the moments analysis [Gatti+21]

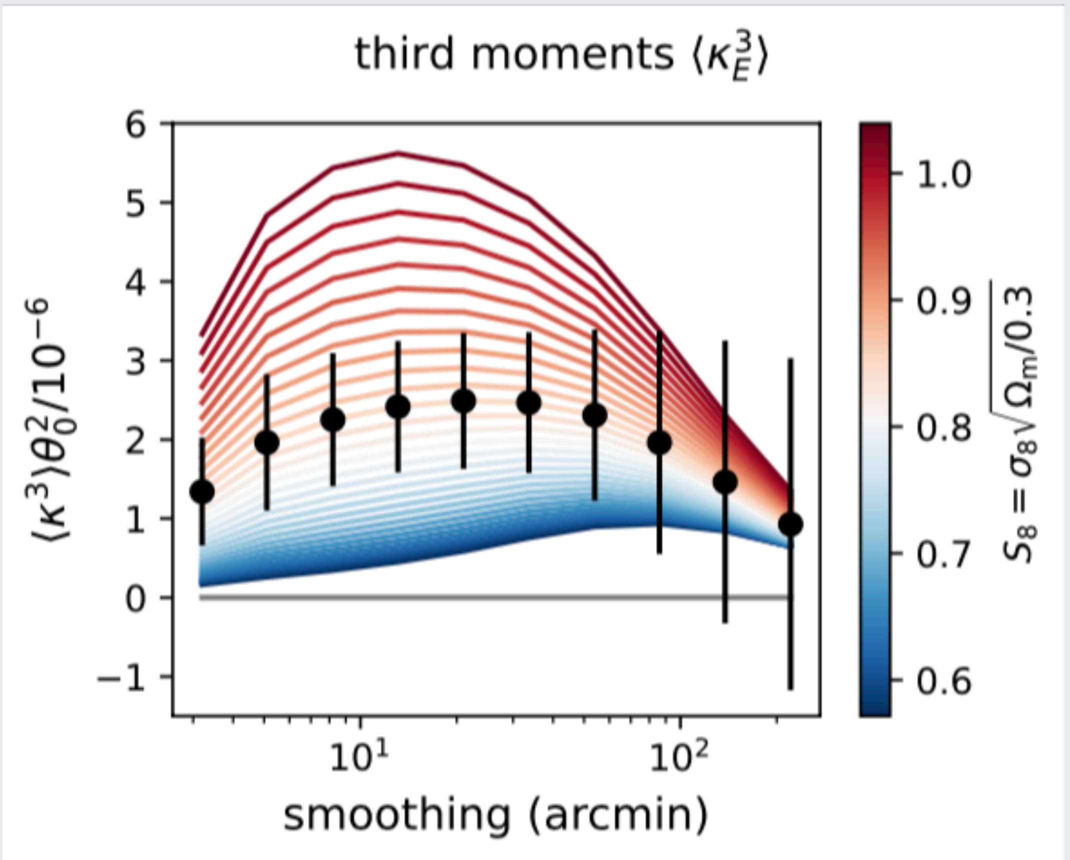
Simulation-based forward modelling

- possible for any statistic
- computationally expensive
- adopted in the peaks analysis [Zuercher+21]

(Credit: M. Gatti)

Analytical modeling: moments

The goal



The algebra

$$\langle\delta^3_{\theta_0,\mathrm{lin}}\rangle(\tau) = \frac{6}{(2\pi)^3}\int \mathrm{d}^2k_1\mathrm{d}^2k_2 W(\mathbf{k}_1,\theta_0)W(\mathbf{k}_2,\theta_0)W(\mathbf{k}_1+\mathbf{k}_2,\theta_0) \times P_{\mathrm{lin}}(\mathbf{k}_1,\tau), P_{\mathrm{lin}}(\mathbf{k}_2,\tau)F_2(\mathbf{k}_1,\mathbf{k}_2,\tau), \quad \text{(A11)}$$

$$F_2(\mathbf{k}_1,\mathbf{k}_2,\tau) = \frac{1}{2}[(1+\frac{k_1}{k_2}\cos\phi)+(1+\frac{k_2}{k_1}\cos\phi)] + [1-\mu(\tau)](\cos^2\phi-1),$$

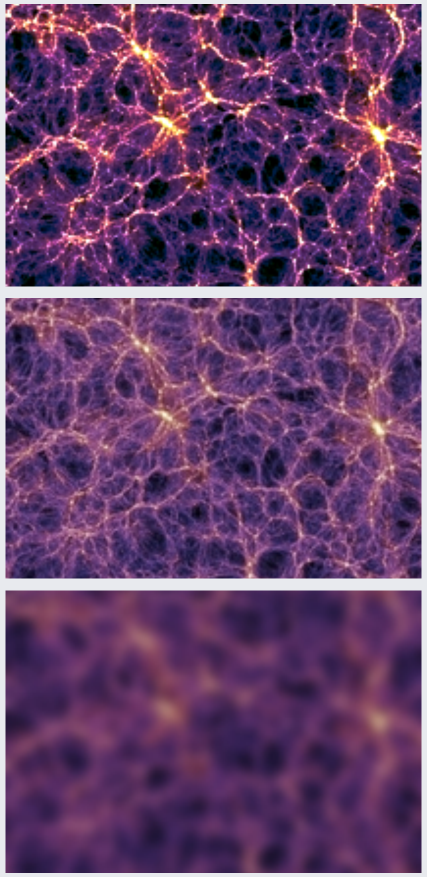
$$F_2(\mathbf{k}_1,\mathbf{k}_2,\tau) = \frac{1}{2}b_1b_2[(1+\frac{k_1}{k_2}\cos\phi)+(1+\frac{k_2}{k_1}\cos\phi)] + [1-\mu(\tau)]c_1c_2(\cos^2\phi-1) + [a_1a_2\mu(\tau)-b_1b_2+[1-\mu(\tau)]c_1c_2].$$

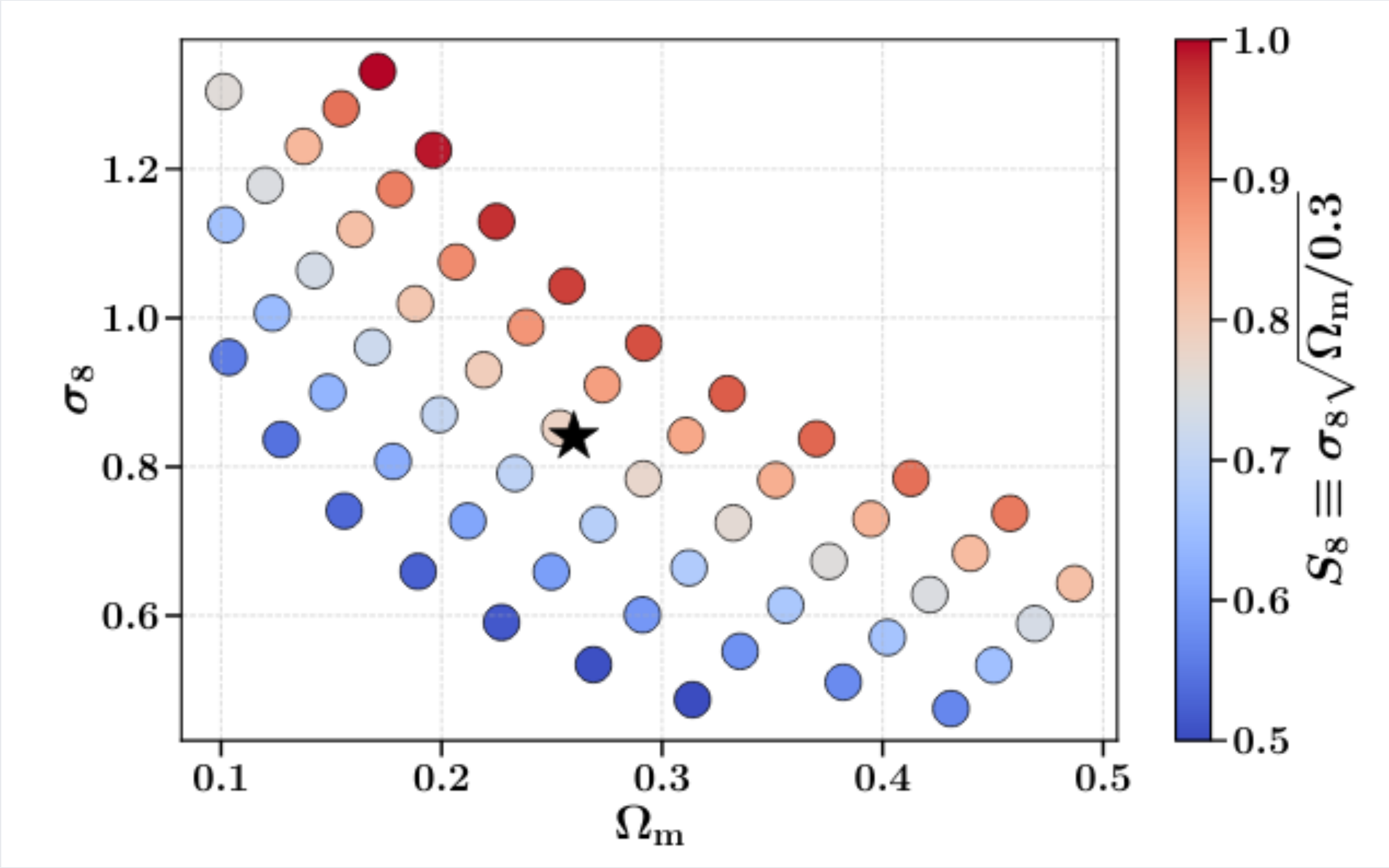
(Credit: M. Gatti)

Simulation based modeling: peaks

Simulation-based forward modeling:
lots of simulations required!

clumpiness of the
Universe



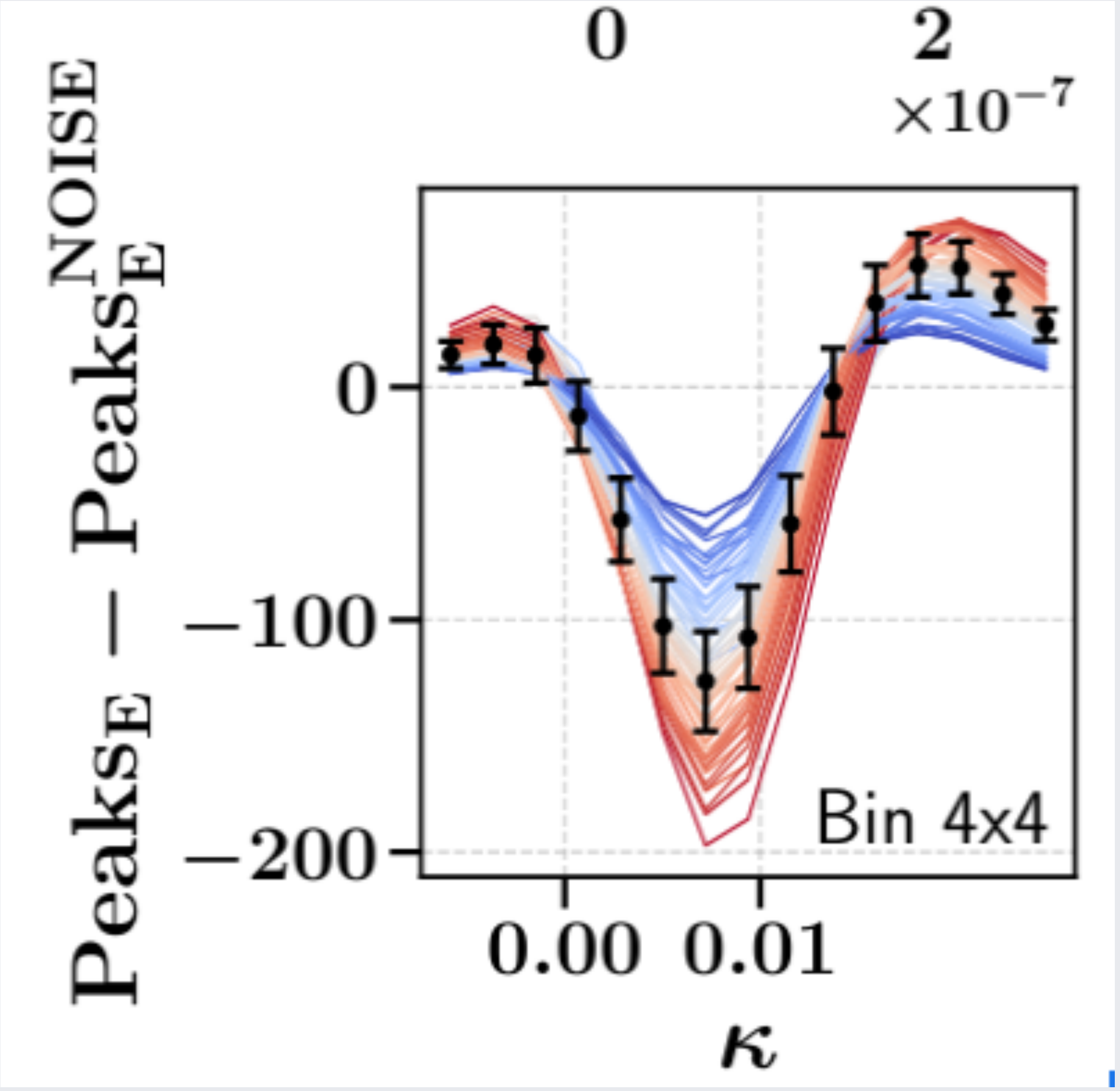


fractional matter abundance
(everything else is Dark Energy)

(Credit: D. Zuercher)

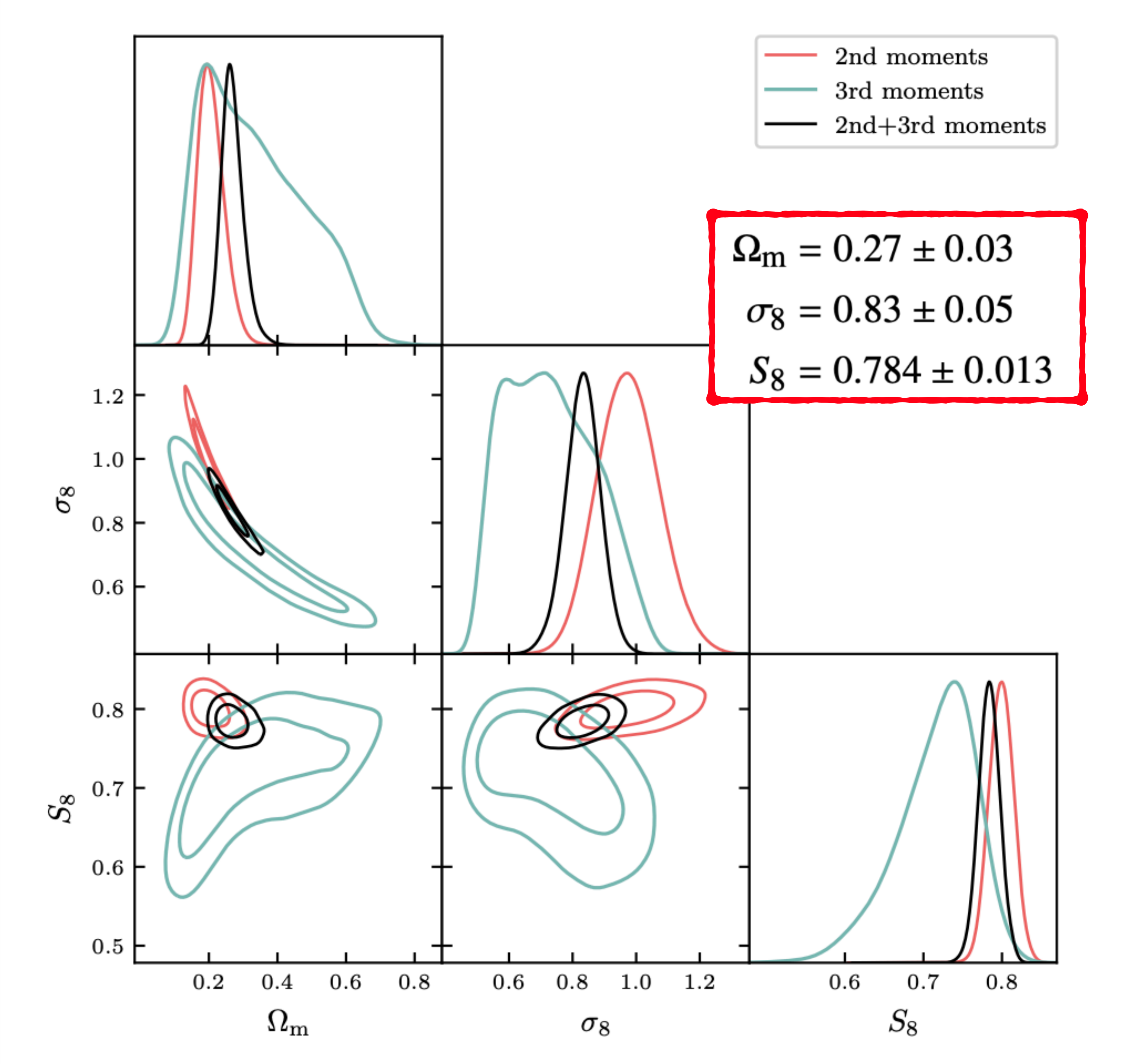
Simulation based modeling: peaks

Prediction for peak function



(Credit: D. Zuercher)

Cosmology from DES Y3 2nd+3rd moments



3rd moments probe additional non Gaussian information & break degeneracy

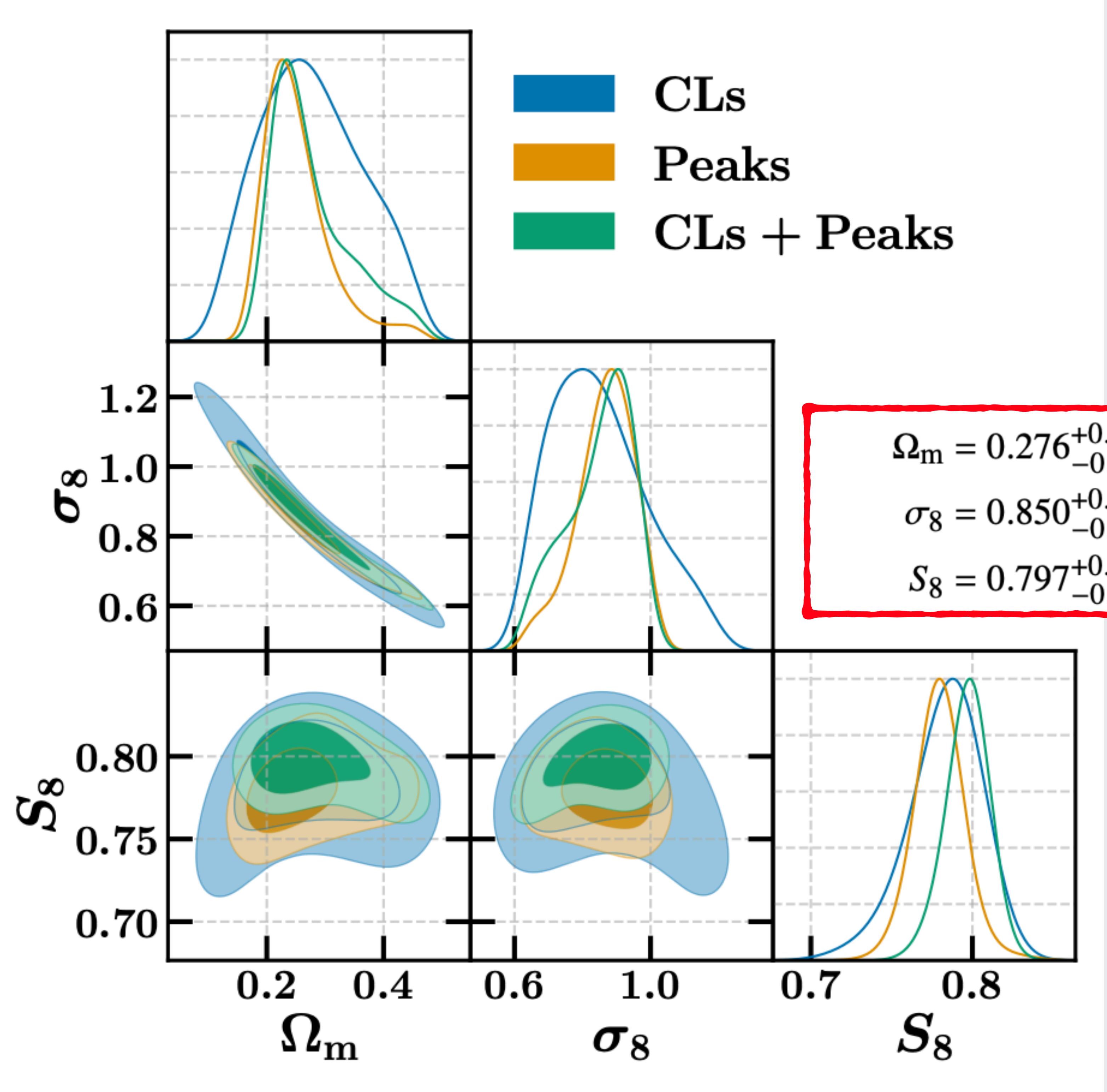
3rd moments is partially independent of second -> different impact of systematics.

3rd+2nd moments improve constraints by 30% over 2nd moments only

Most stringent constraints on S8 from a WL analysis to date!

(Credit: M. Gatti)

Cosmology from DES Y3 Power Spectra+ Peaks



$$\Omega_m = 0.276^{+0.034}_{-0.086}$$
$$\sigma_8 = 0.850^{+0.13}_{-0.068}$$
$$S_8 = 0.797^{+0.015}_{-0.013}$$

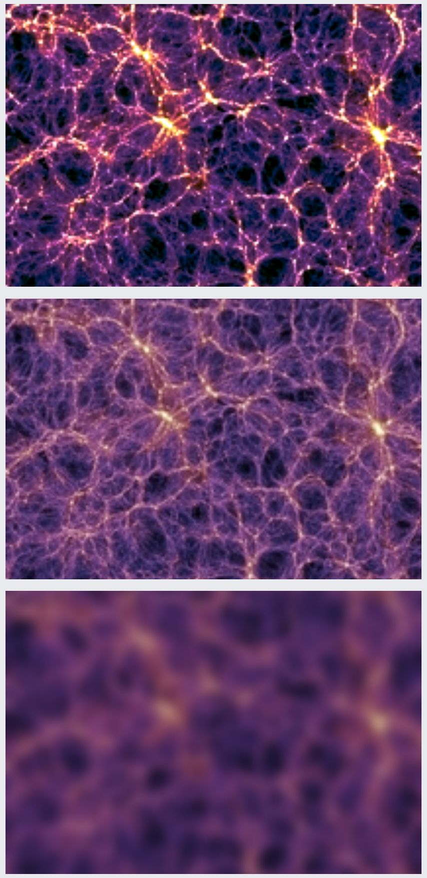
Peaks+Power Spectra(CL) improve constraints by 40% over Power Spectra only

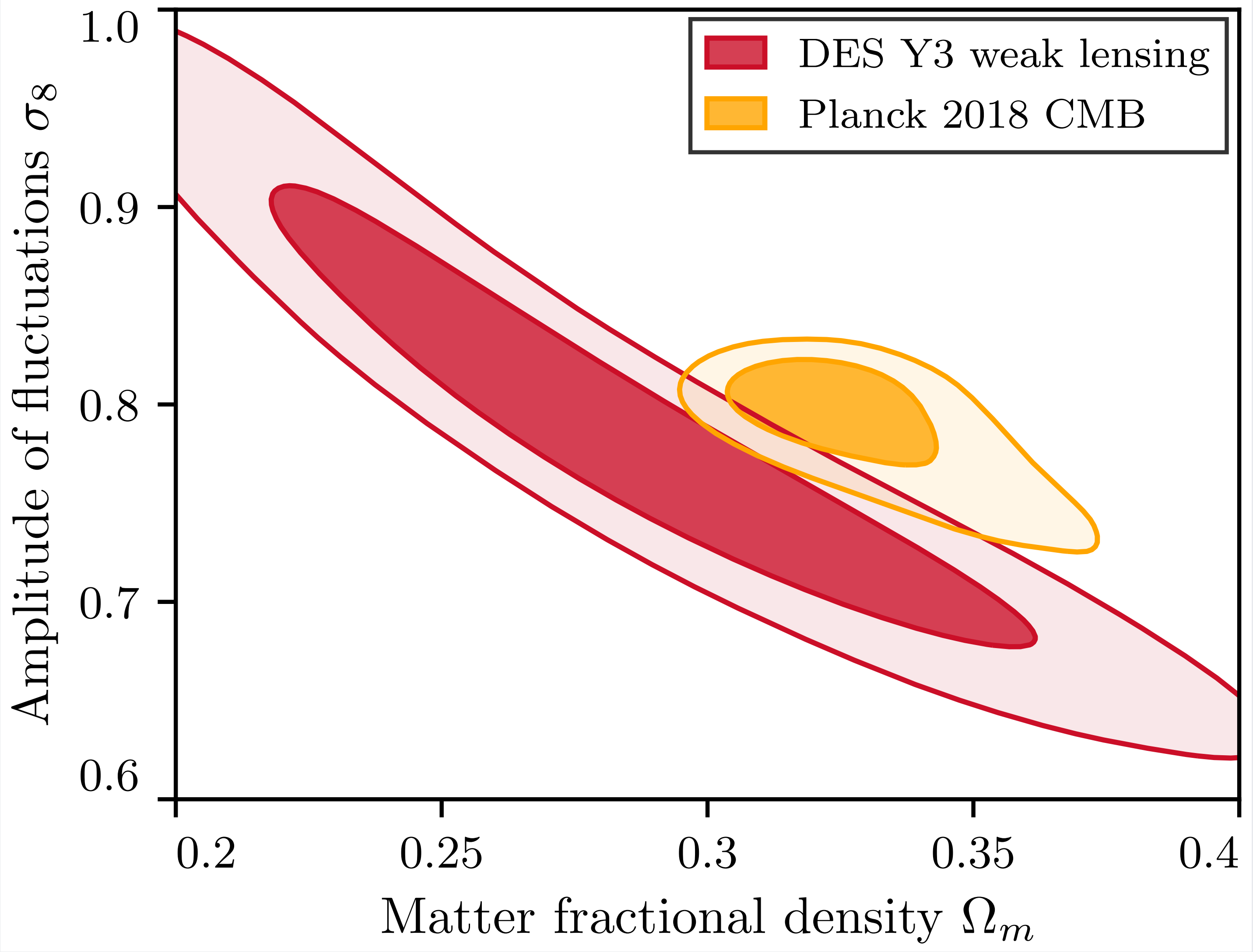
Similar constraining power on S8 of the moments analysis

Agreement or disagreement?

Seemingly easy question: do experiments agree?

clumpiness of the Universe



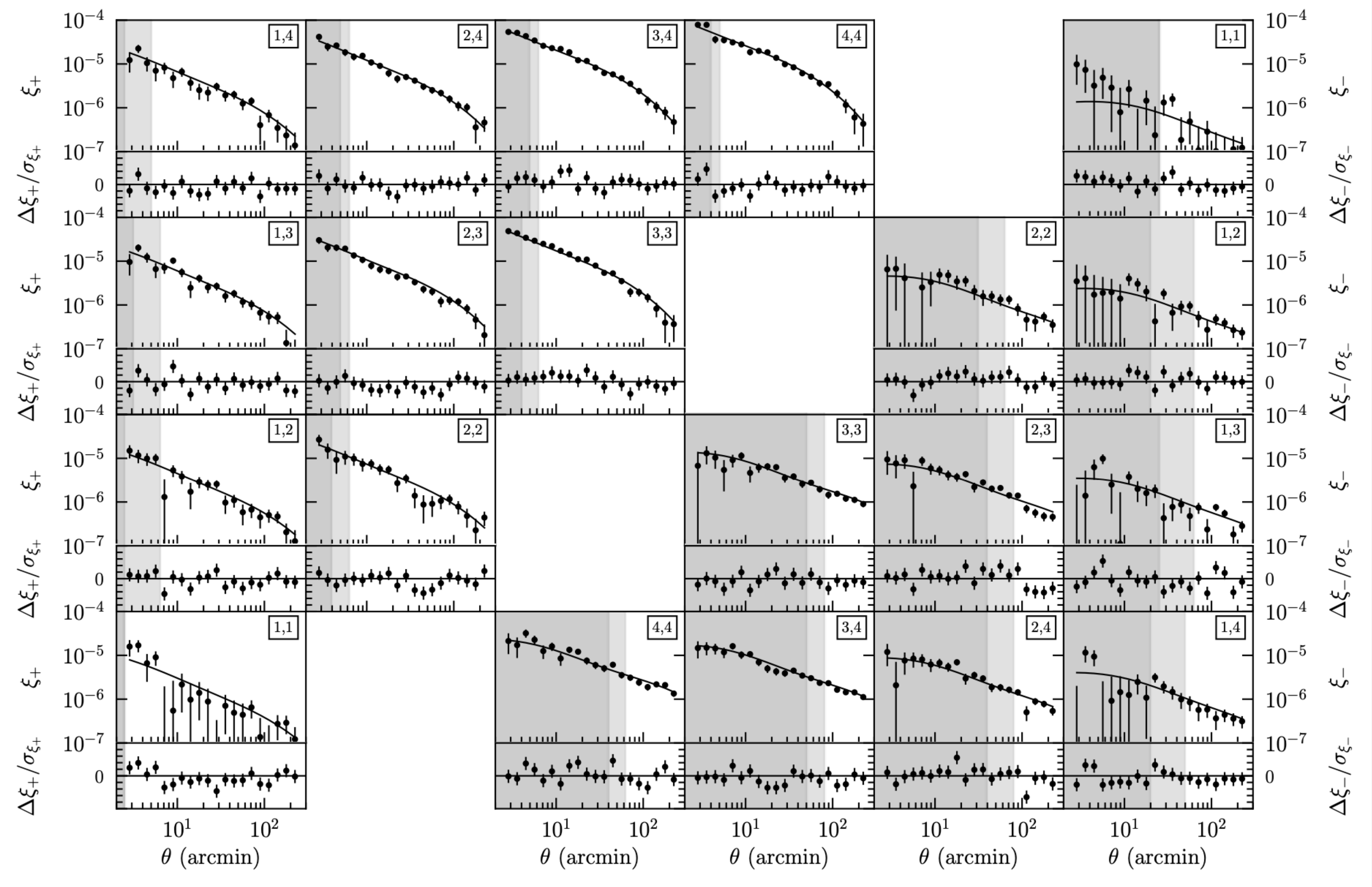


fractional matter abundance
(everything else is Dark Energy)

(Based on some work in DES)

Testing concordance: the challenges

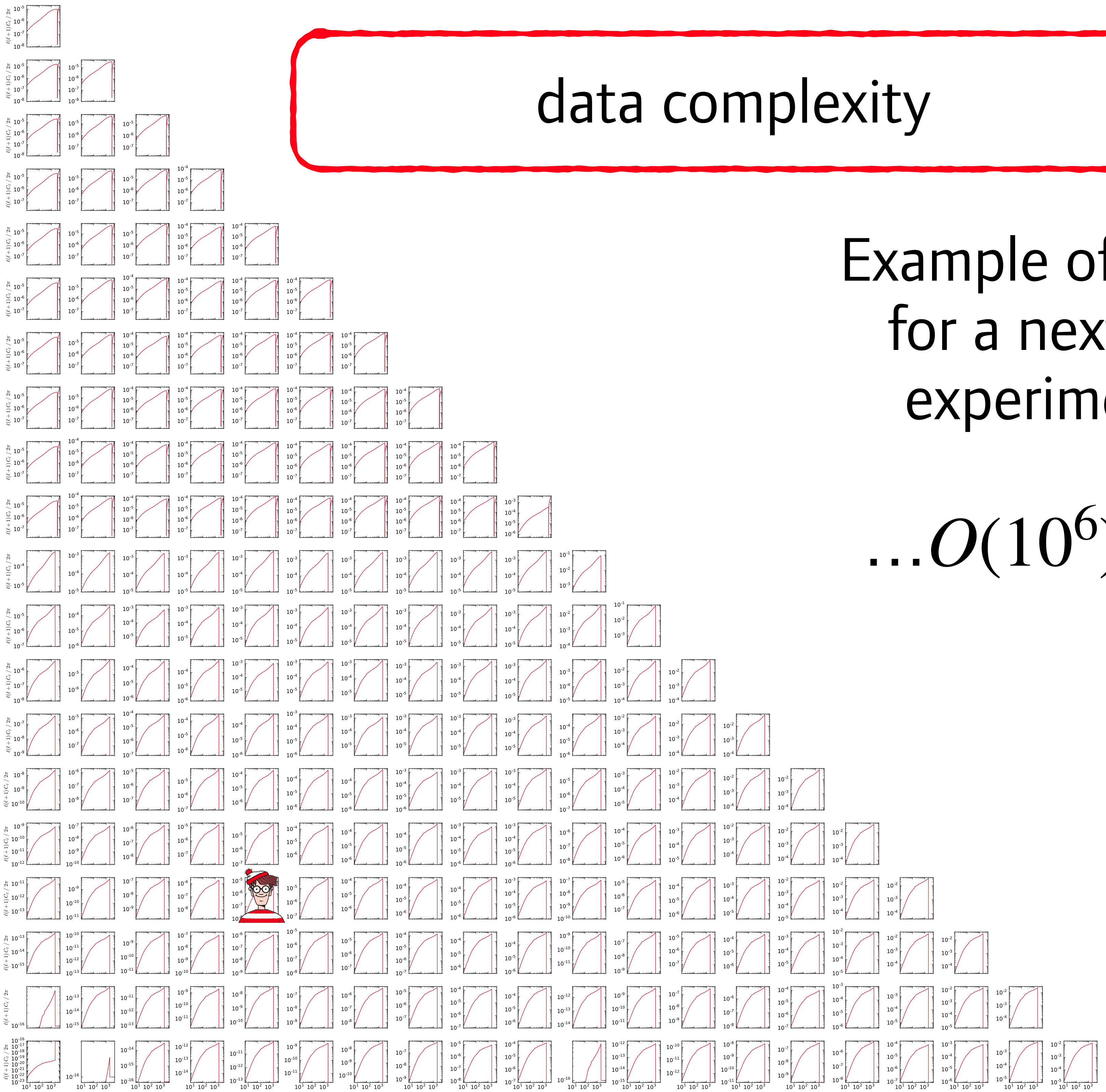
Data complexity



Dark Energy Survey

(Based on some work in DES)

Testing concordance: the challenges



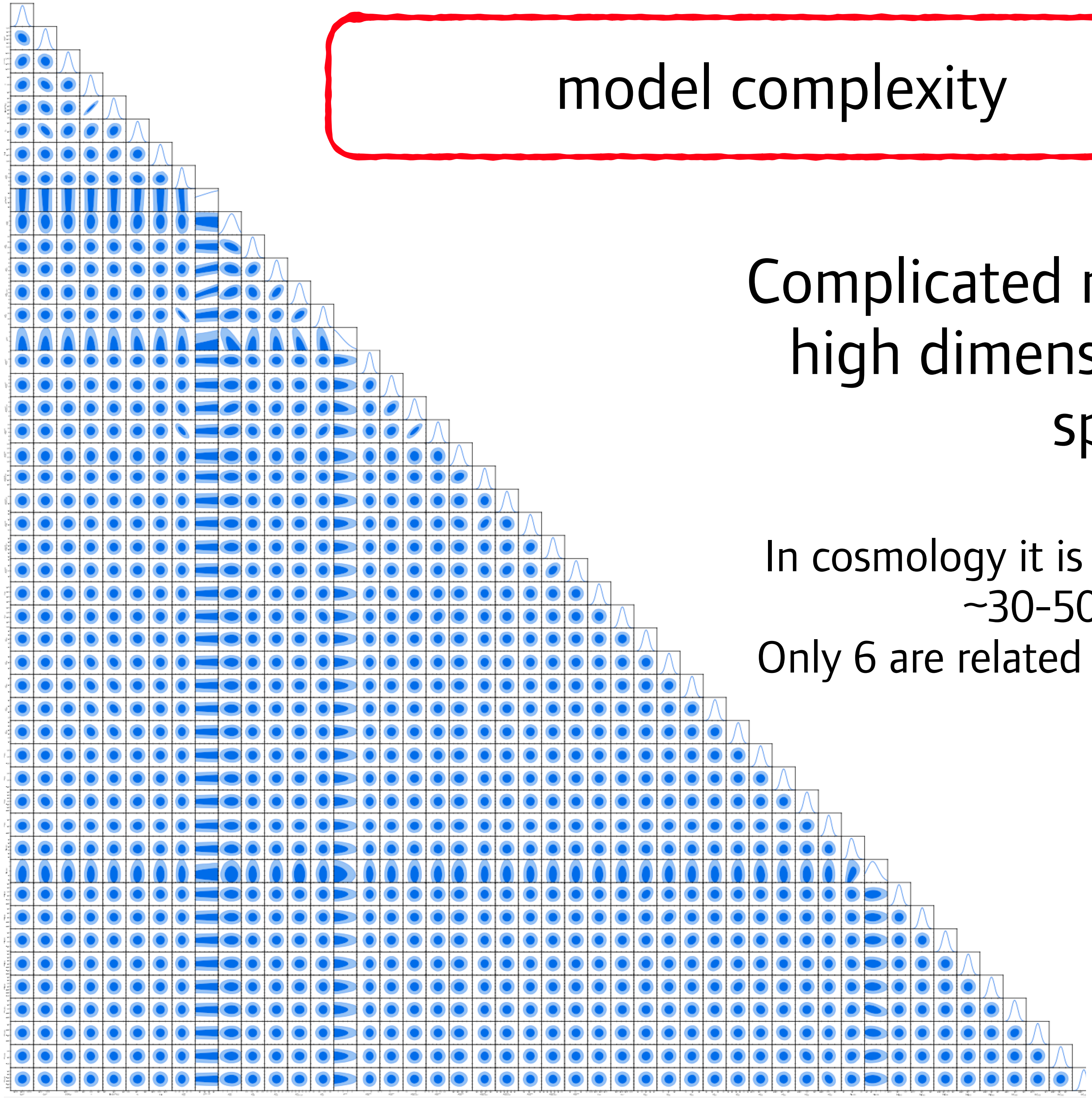
data complexity

Example of a data vector for a next generation experiment (Euclid)

...$O(10^6)$ data points!

(Based on some work in Euclid)

Testing concordance: the challenges



model complexity

Complicated models come with high dimensional parameter spaces.

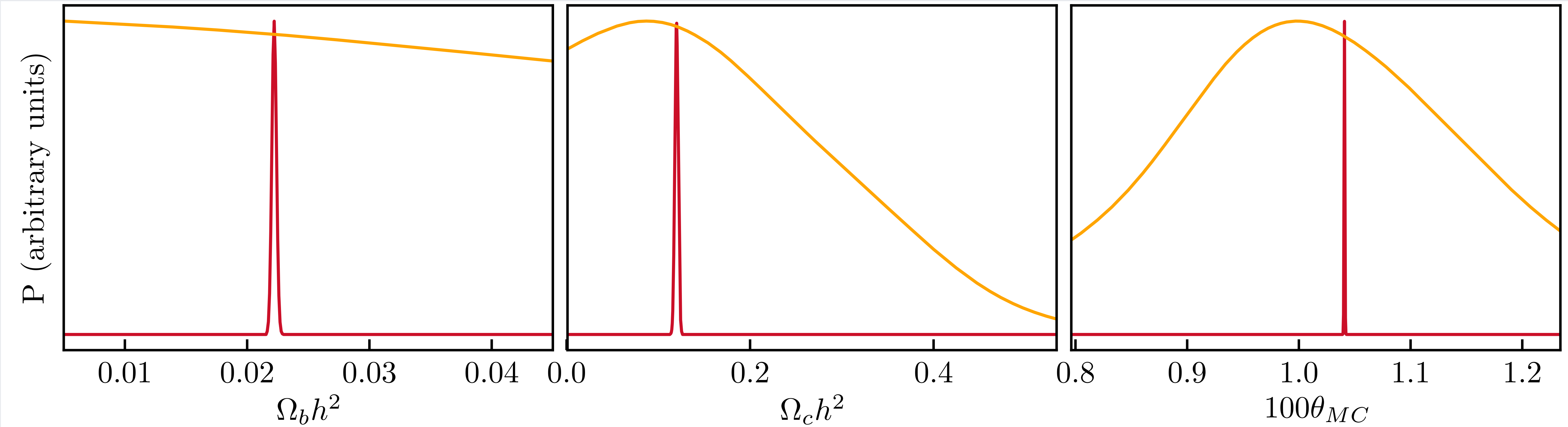
In cosmology it is customary to work with ~30-50 parameters.
Only 6 are related to fundamental physics.

(This is the DES parameter space…)

Testing concordance: the challenges

Projections of the parameter space might hide discrepancies

Do they agree?

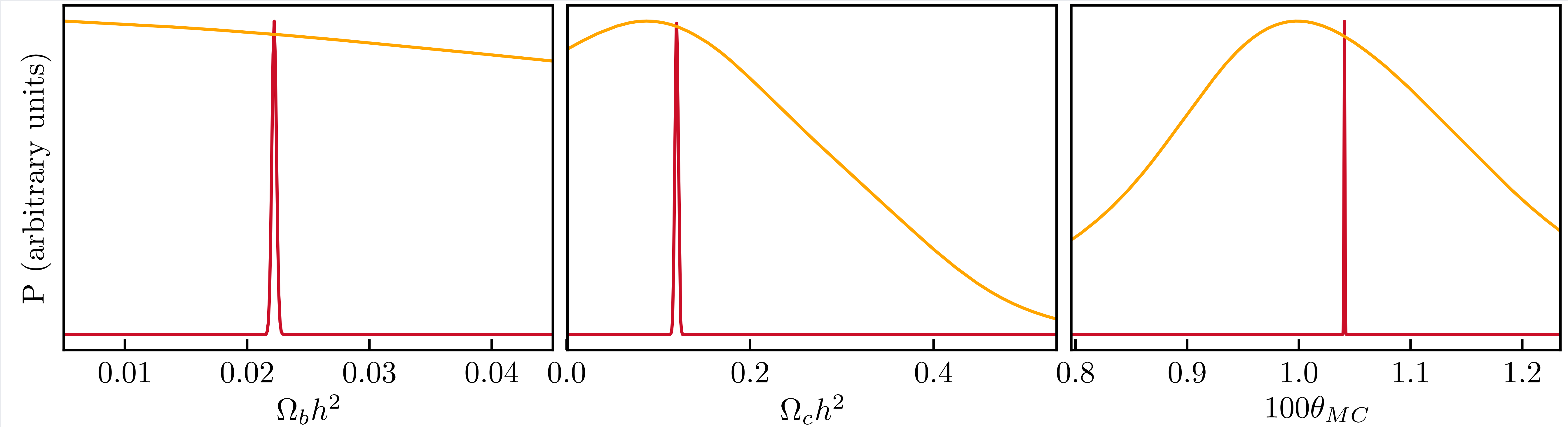


What are these data sets?

Testing concordance: the challenges

Projections of the parameter space might hide discrepancies

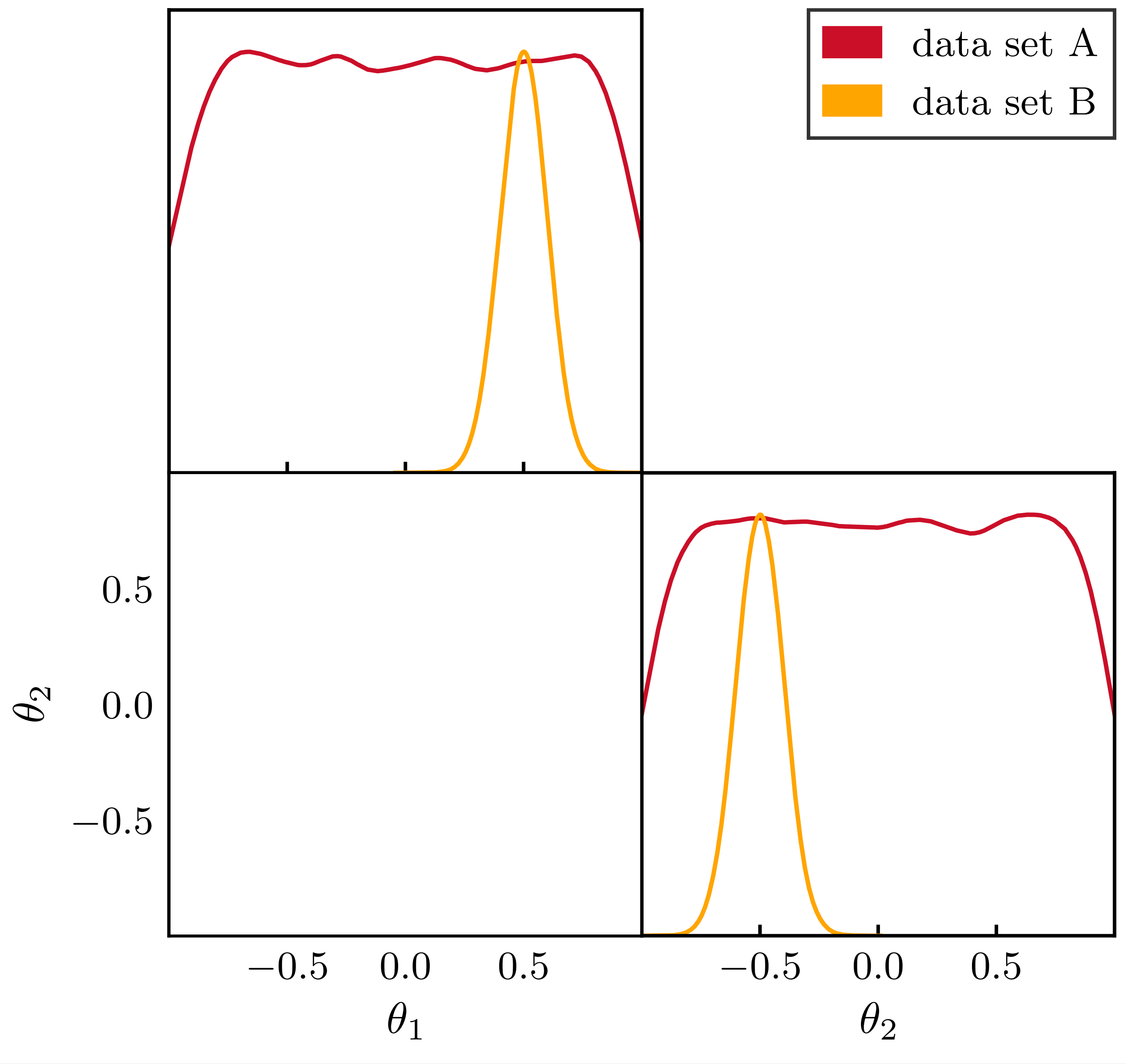
Do they agree? No, to 5 sigma



What are these data sets? Planck CMB and local Hubble constant

Testing concordance: the challenges

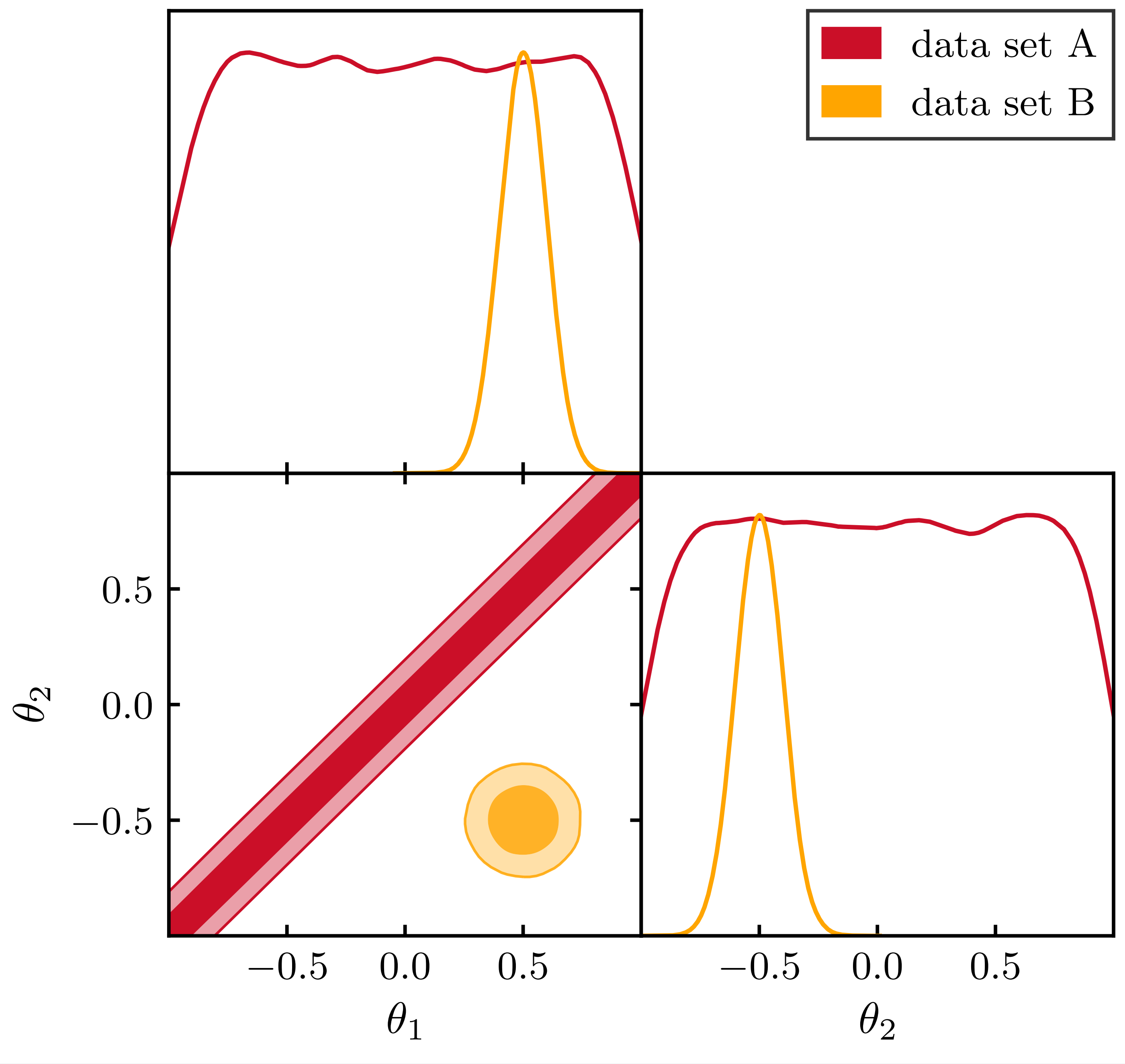
What is going on?



(made up posteriors from Lemos, MR et al arXiv:2012.09554)

Testing concordance: the challenges

We are being tricked by low dimensional projections



(made up posteriors from Lemos, MR et al arXiv:2012.09554)

Testing concordance when life is hard:

tools of the trade:

Theoretical papers:

- Raveri and Hu, "Concordance and Discordance in Cosmology" **arXiv 1806.04649**
- Raveri, Zacharegkas, Hu, "Quantifying concordance of correlated cosmological data sets" **arXiv 1912.04880**
- Raveri and Doux, "Non-Gaussian estimates of tensions in cosmological parameters" **arXiv 2105.03324**

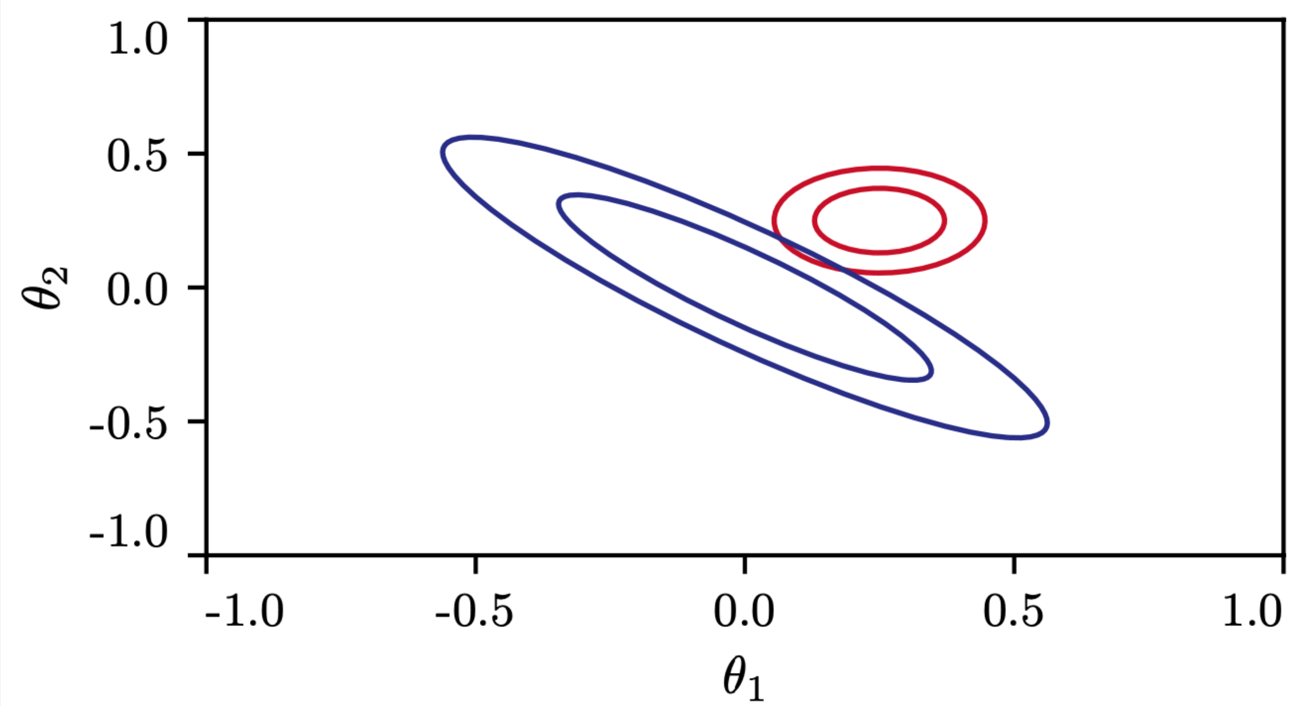
Code implementation (in Python, with several example notebooks)

```
~ pip install tensiometer
```

Used by many collaborations: DES, ACT, PTA

Parameter differences

Full dimensional distribution of differences in parameters



Start with:

$$P_1(\theta_1)P_2(\theta_2)$$

change variables:

$$P_1(\theta_1)P_2(\theta_1 - \Delta\theta)$$

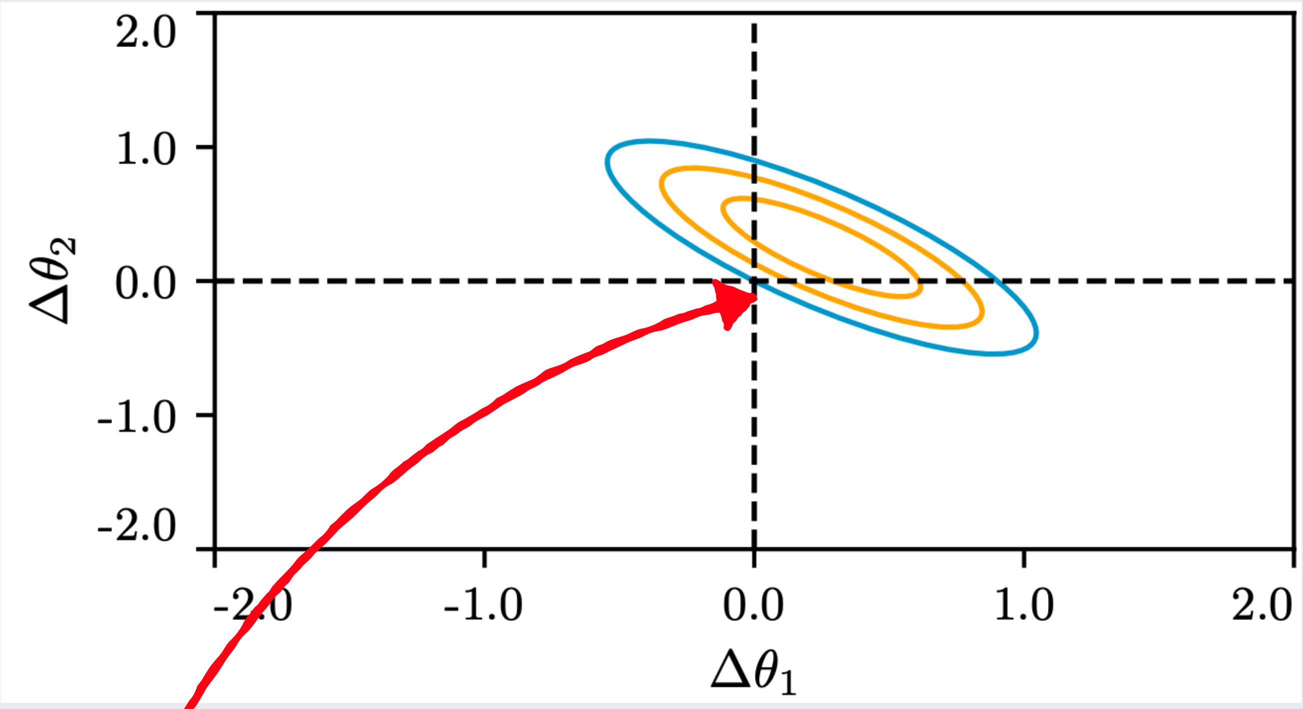
integrate out

$$P(\Delta\theta) = \int P_1(\theta_1)P_2(\theta_1 - \Delta\theta)d\theta_1$$

(Raveri and Doux arXiv:2105.03324)

Parameter differences

$$P(\Delta\theta) = \int P_1(\theta_1) P_2(\theta_1 - \Delta\theta) d\theta_1$$



Our goal:

$$\Delta = \int_{P(\Delta\theta) > P(0)} P(\Delta\theta) d\Delta\theta$$

(Raveri and Doux arXiv:2105.03324)

Parameter differences: the bad news...

$$P(\Delta\theta) = \int P_1(\theta_1) P_2(\theta_1 - \Delta\theta) d\theta_1$$

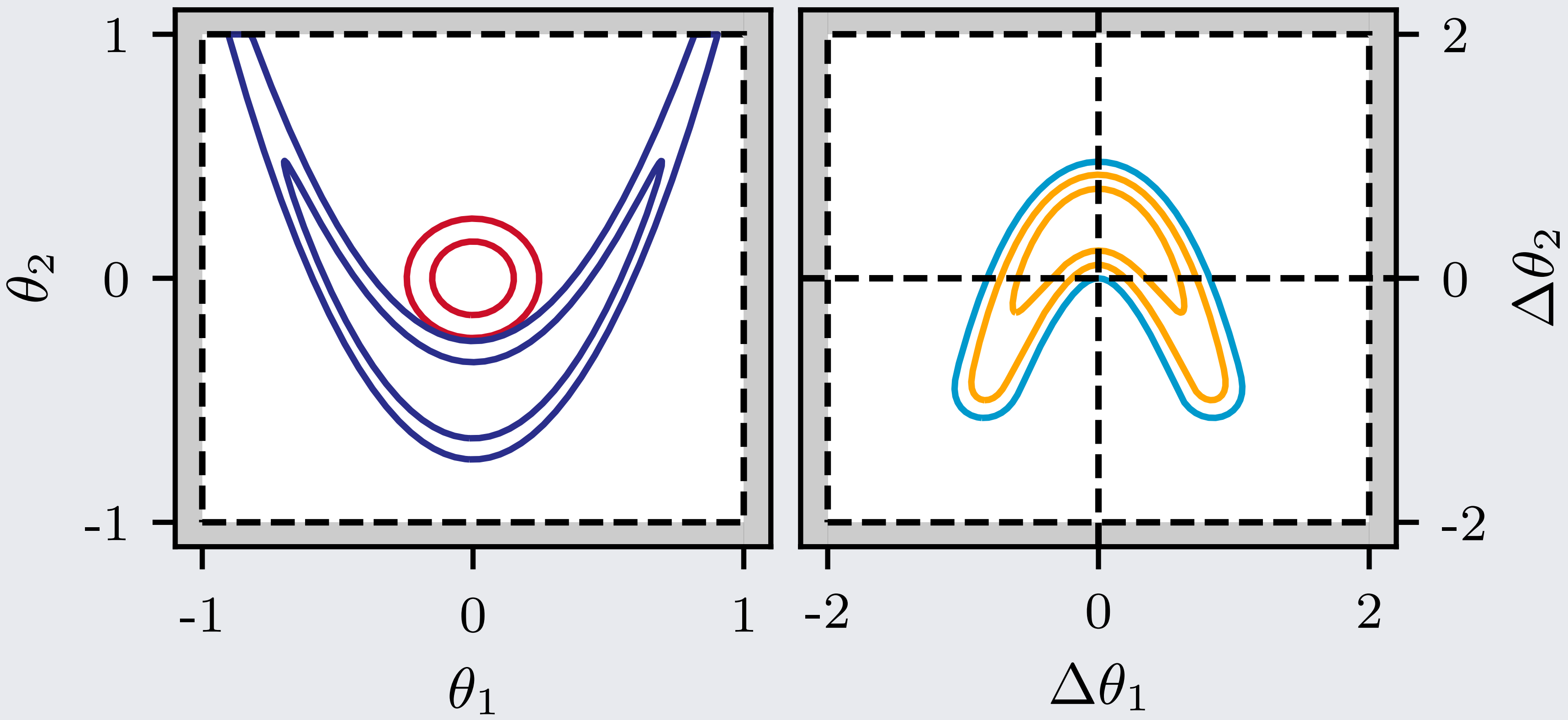
Usually very high dimensional integral

$$\Delta = \int_{P(\Delta\theta) > P(0)} P(\Delta\theta) d\Delta\theta$$

Usually very high dimensional integral

Way out assuming Gaussianity: Raveri & Hu arXiv:1806.04649

Parameter differences in practice



What to do when the distribution is **non-Gaussian**?

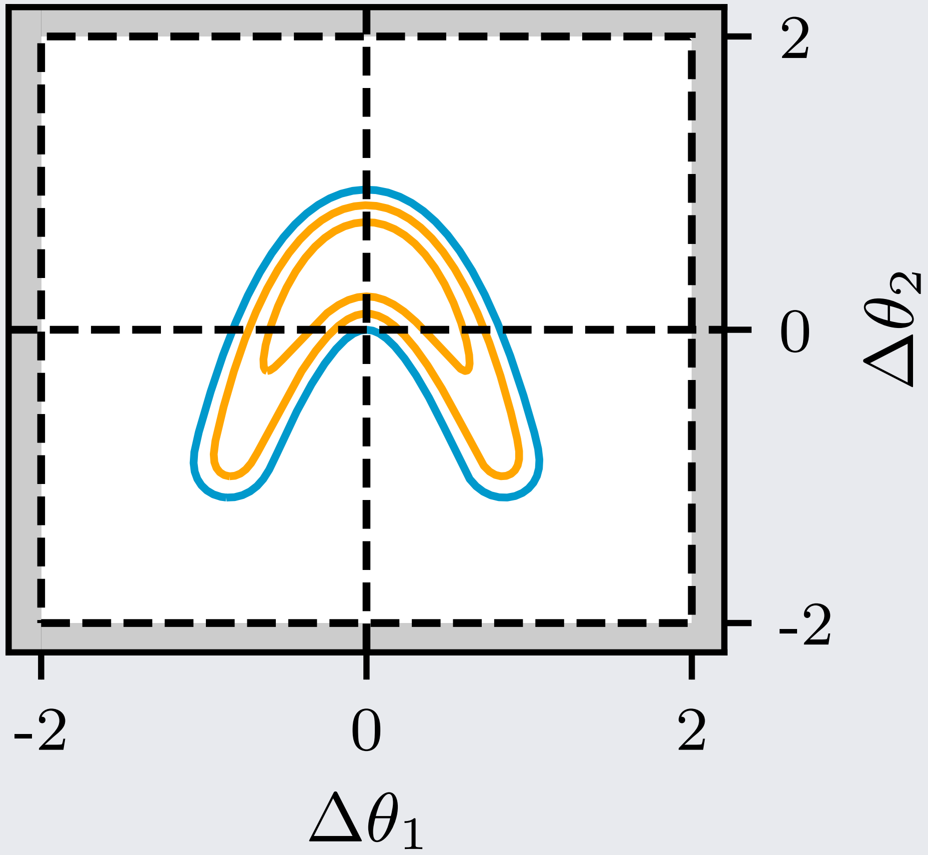
No way out of performing the integrals…

$$P(\Delta\theta) = \int P_1(\theta_1) P_2(\theta_1 - \Delta\theta) d\theta_1$$

First integral can be done with Monte Carlo

(Raveri and Doux arXiv:2105.03324)

Parameter differences in practice



$$\Delta = \int_{P(\Delta\theta) > P(0)} P(\Delta\theta) d\Delta\theta$$

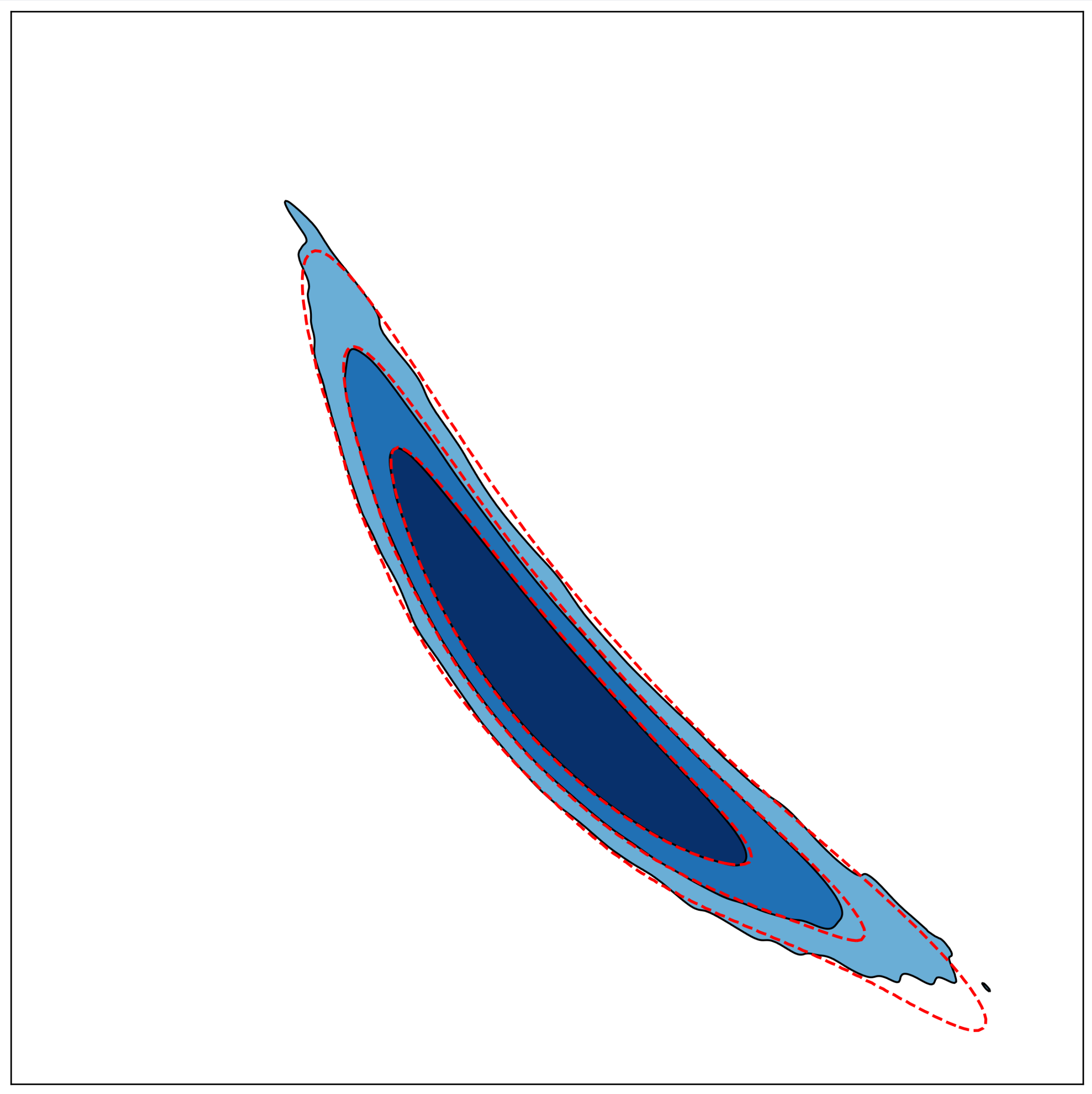
Second integral can be done with KDE but is very expensive

Naive algorithm is N^2 (not doable)
R&D arXiv:2105.03324 has the NlogN algorithm which is still very expensive (curse of dimensionality of KDEs)

(Raveri and Doux arXiv:2105.03324)

Parameter differences and machine learning

$$\Delta = \int_{P(\Delta\theta) > P(0)} P(\Delta\theta) d\Delta\theta$$



Solution:
learn the parameter difference posterior

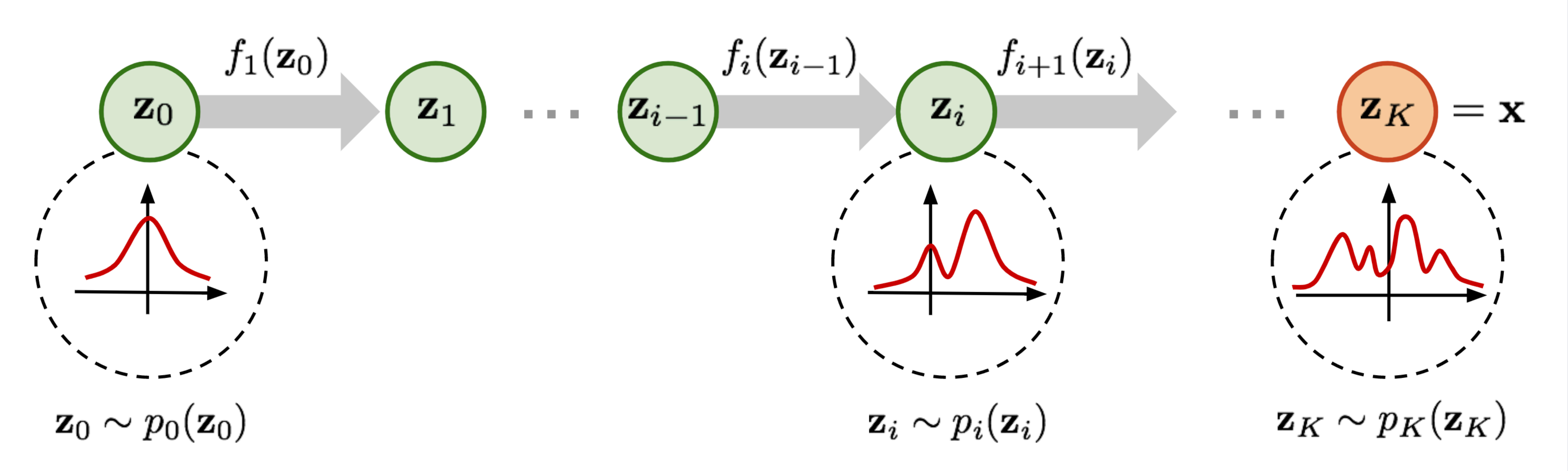
$$P(\Delta\theta)$$

…then MC integrate

(Raveri and Doux arXiv:2105.03324)

Normalizing flows

MAF normalizing flows



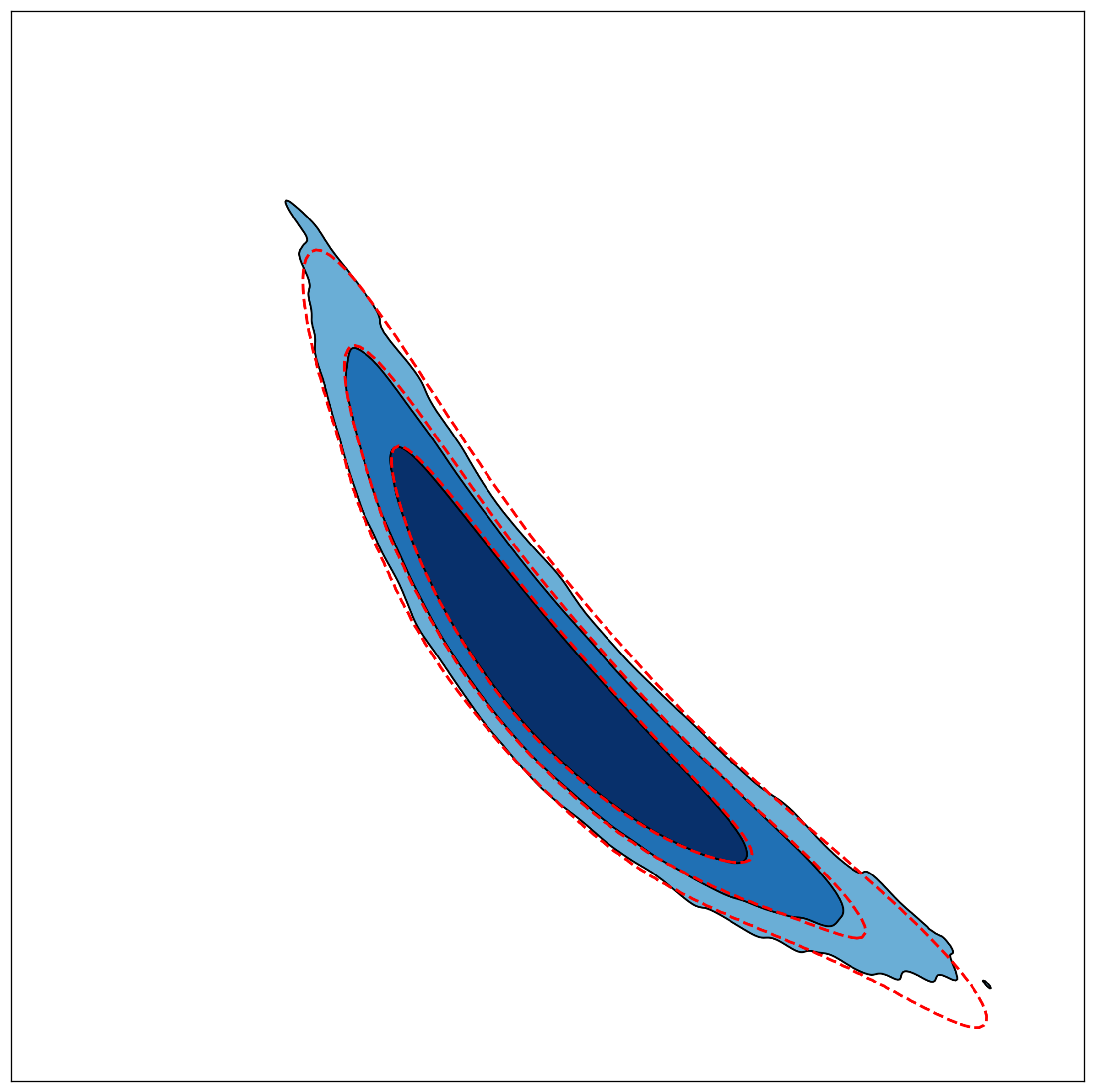
$$
\begin{aligned}
y_1 &= \mu_1 + \sigma_1 z_1 \\
y_i &= \mu(y_{1\dots i-1}) + \sigma(y_{1\dots i-1}) z_i
\end{aligned}
$$

Stacked with permutations to ensure no coordinate is unlucky

(Papamakarios, Pavlakou, Murray arXiv:1705.07057)

Normalizing flows performances

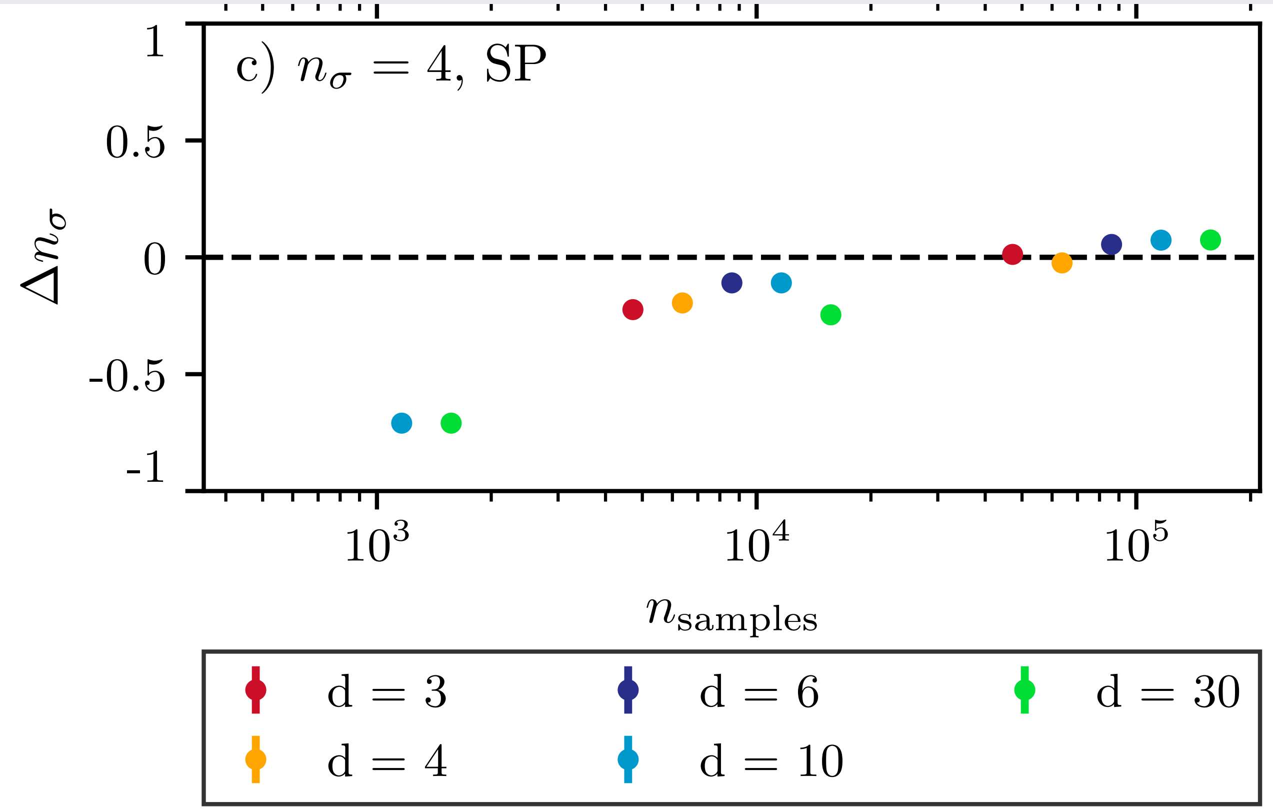
…trained PDFs are indistinguishable from real (KDE) ones…



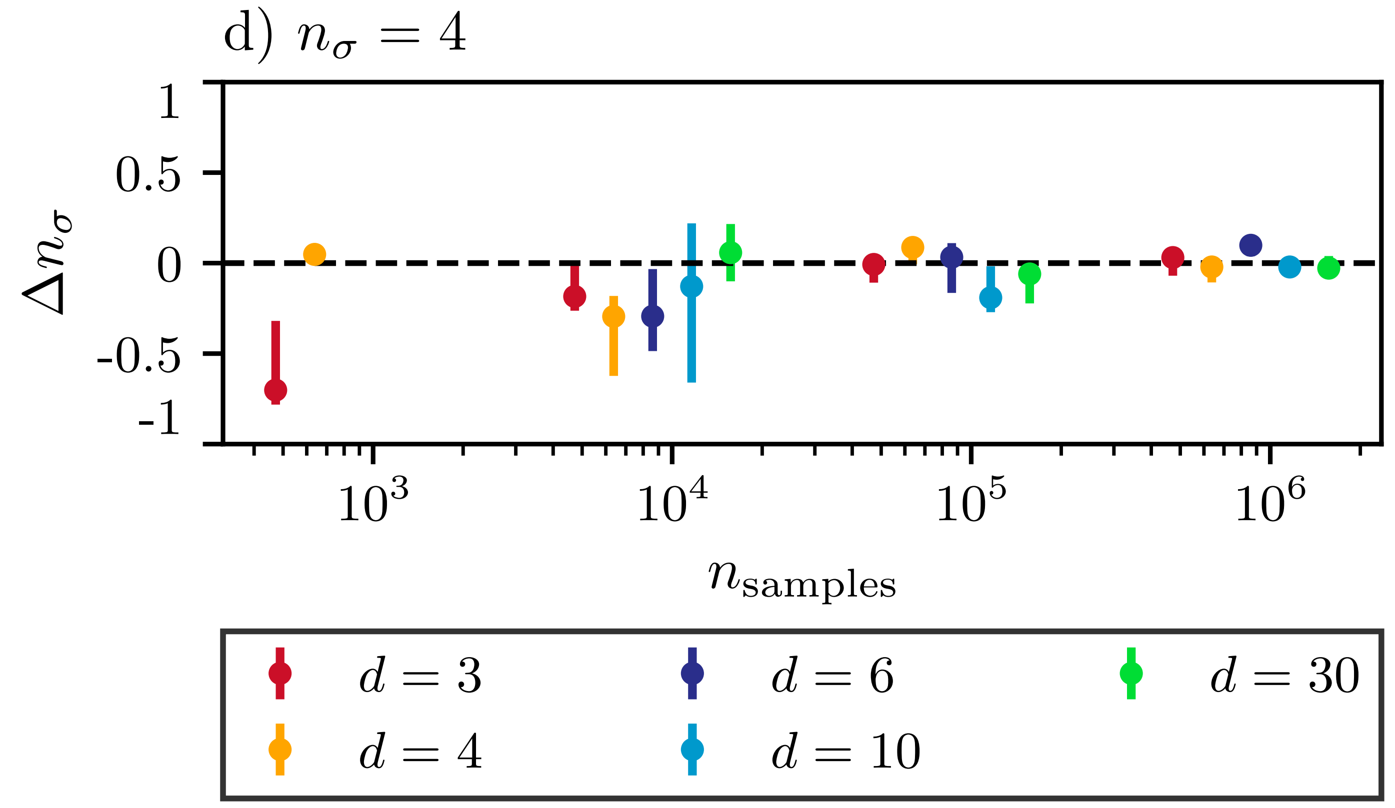
(Raveri and Doux arXiv:2105.03324)

Matches or outperforms KDE based methods in all our tests

KDE



Flow

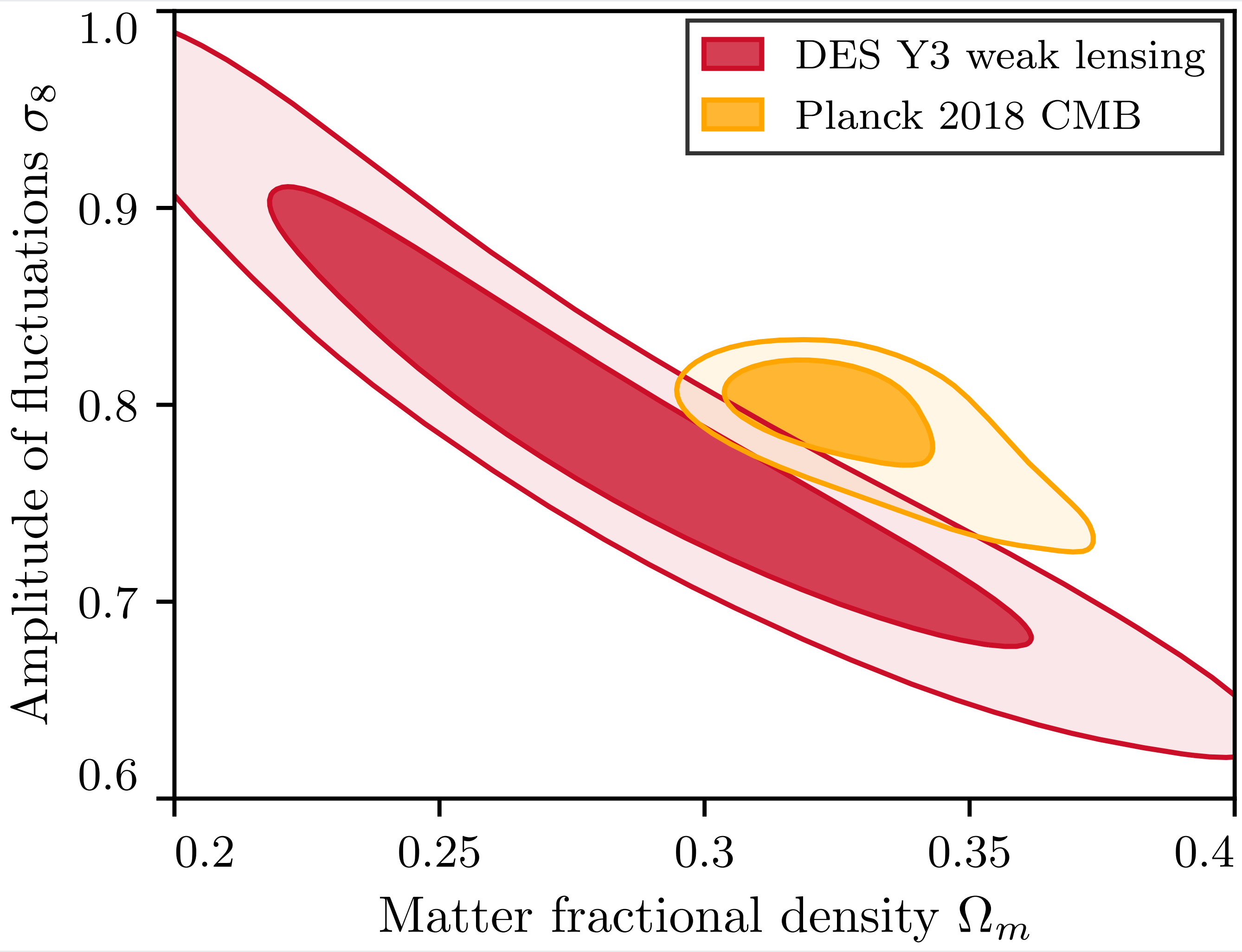


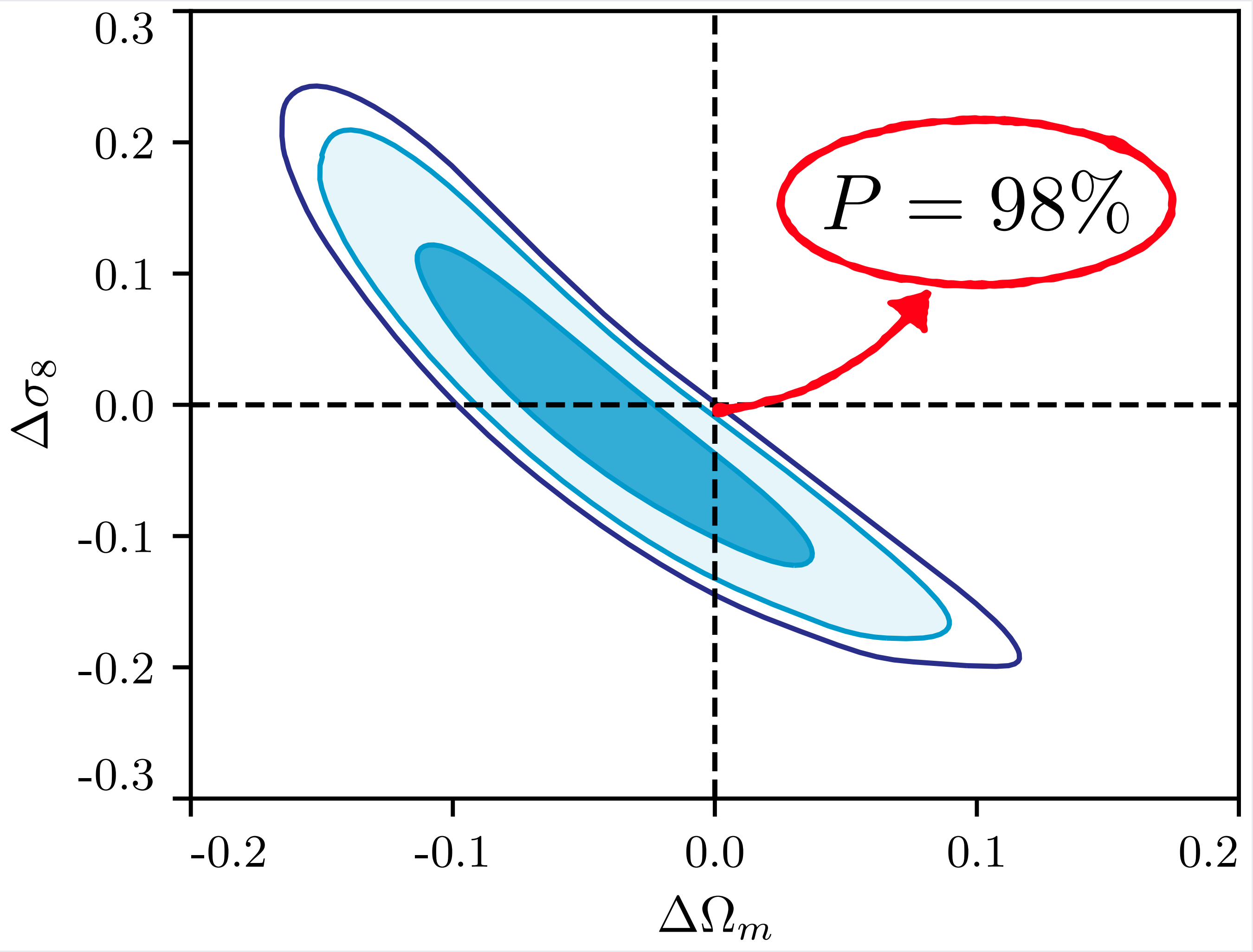
(Raveri and Doux arXiv:2105.03324)

Application to DES Y3

Original parameter space

Difference parameter space

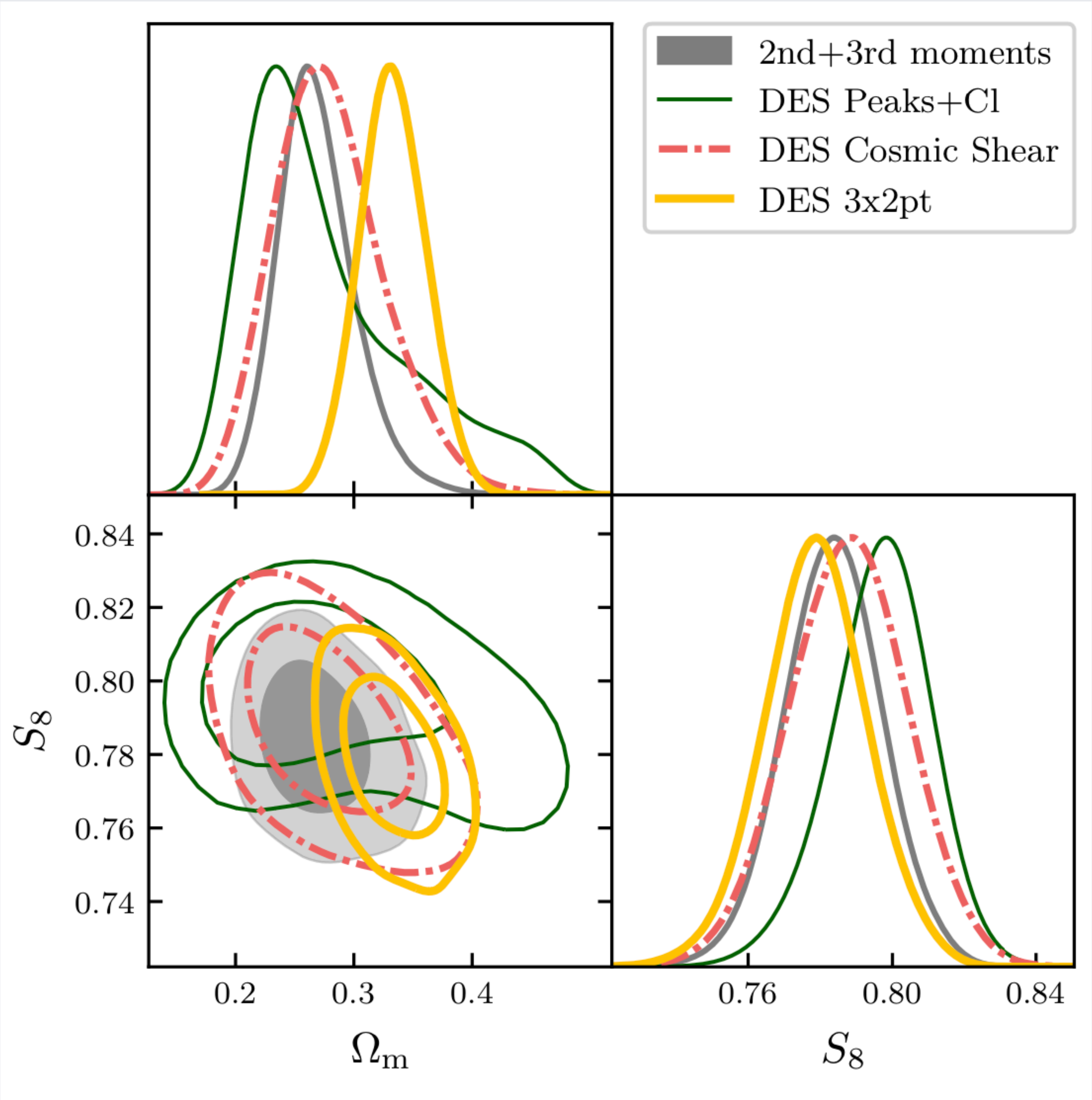




…all other parameters are hidden…

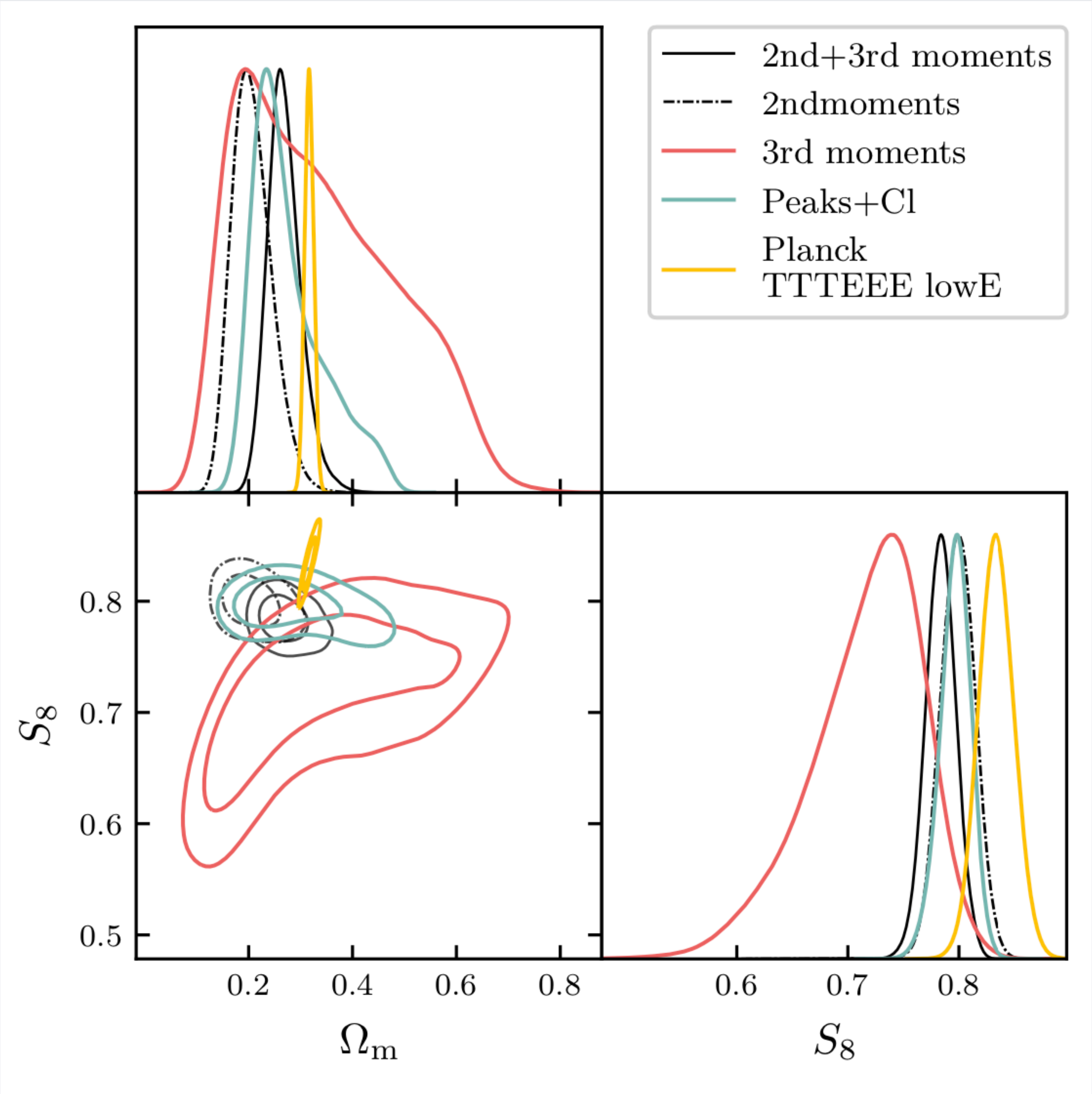
(Based on DES Y3 KP arXiv:2105.13549)

Back to non-Gaussian statistics



Broad consistency between different measurements

Back to non-Gaussian statistics



Consistency with Planck

They are consistent (<3σ), although note that 3rd moments alone shows a 2.8 tension

	Planck TTTEEE lowl lowE
2nd moments	2.7σ
3rd moments	2.8σ
2nd+3rd moments	2.2σ
Peaks+Cl	1.5σ

Outlook

* Gravity induces non-Gaussianity in cosmological observations. Rich datasets, hard to extract and model.

* Non-Gaussian statistics help with cosmological constraints. Worked example from the state-of-the-art WL survey.

* Understanding which statistic is more useful and how to best model it is the current challenge.

Outlook

* Non-Gaussianity in parameter space complicates the task of understanding when experiments agree.

* We have seen how to do it and how this is made possible by ML algorithms. De facto industry standard.

* The algorithms we use to model non-Gaussian distributions are extremely fast (and differentiable) -> more applications to come!