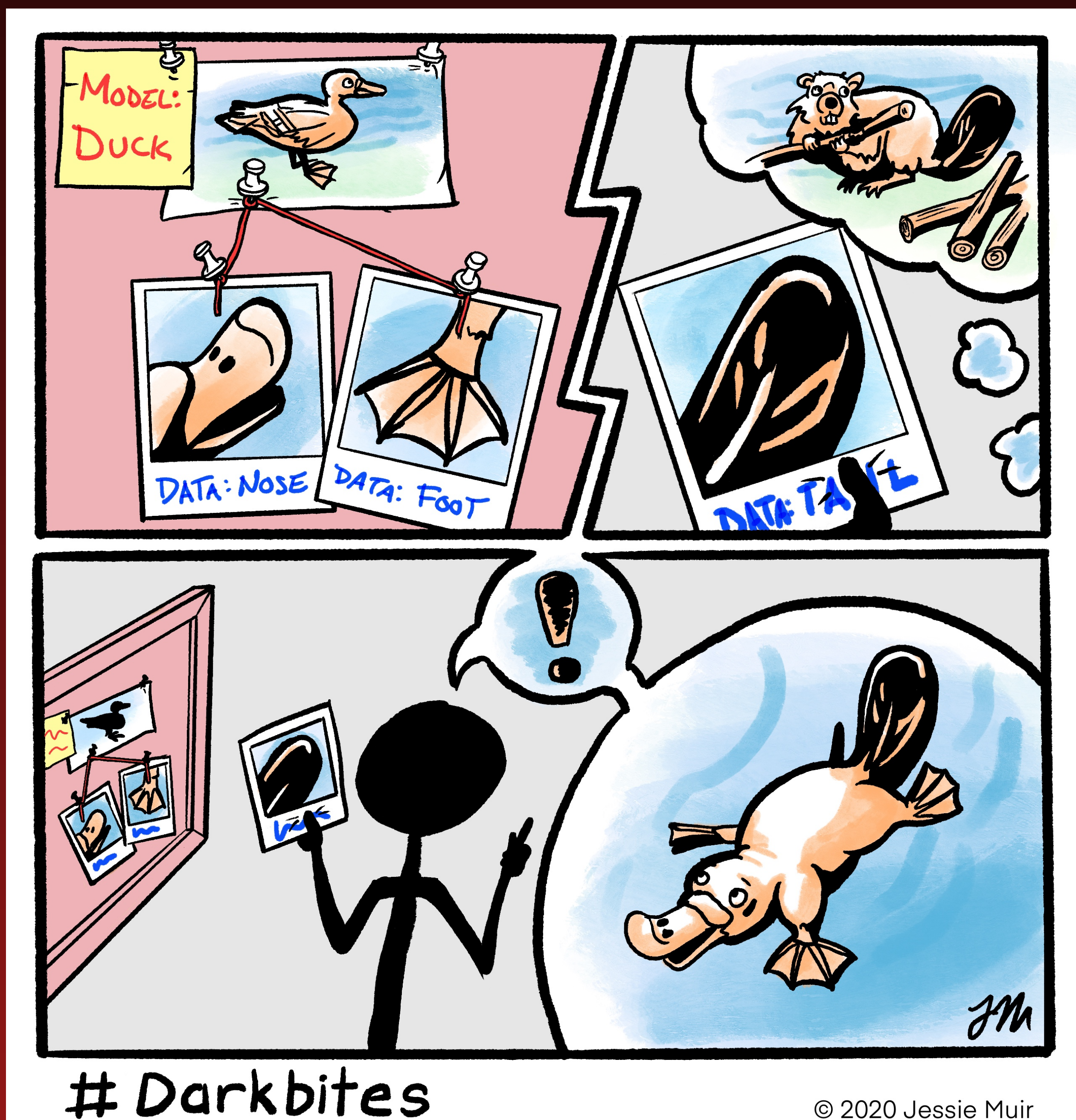
Cosmological Tensions

(and How to Find Them)



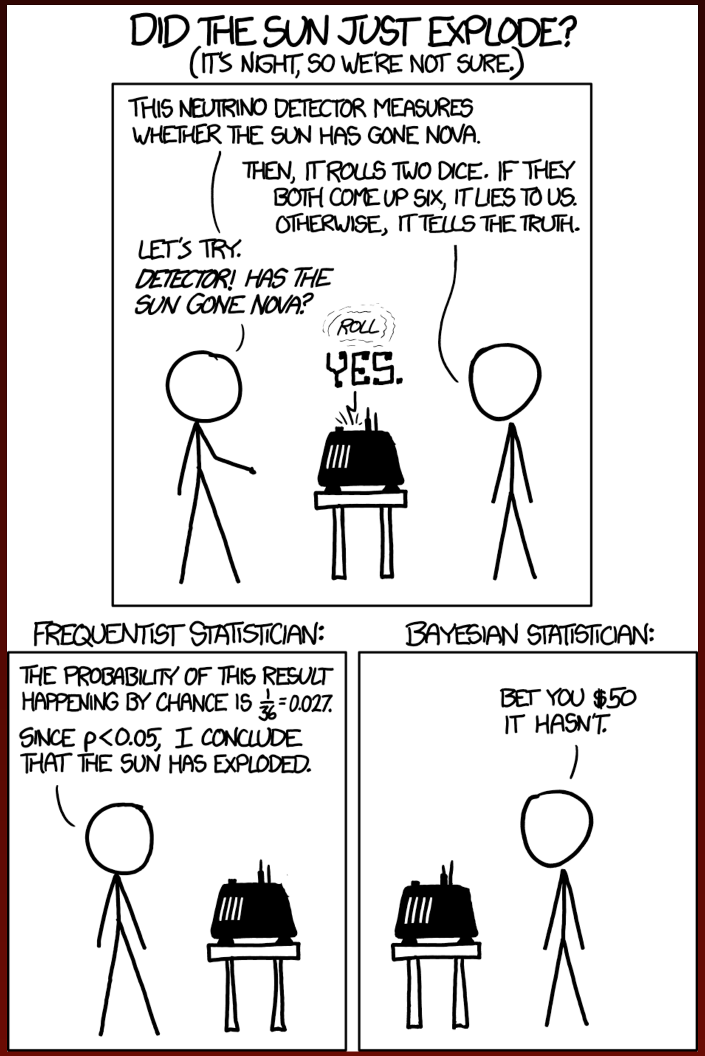
Text: Andresa Campos
@AndresaCampos

Illustration: Jessie Muir
@jlynnmuir

Pablo Lemos
University College London -
University of Sussex
07-05-2021
p.lemos@sussex.ac.uk

In collaboration with: George Efstathiou, Will Handley, Ofer Lahav, Benjamin Joachimi, Antony Lewis and others

Bayesian Statistics



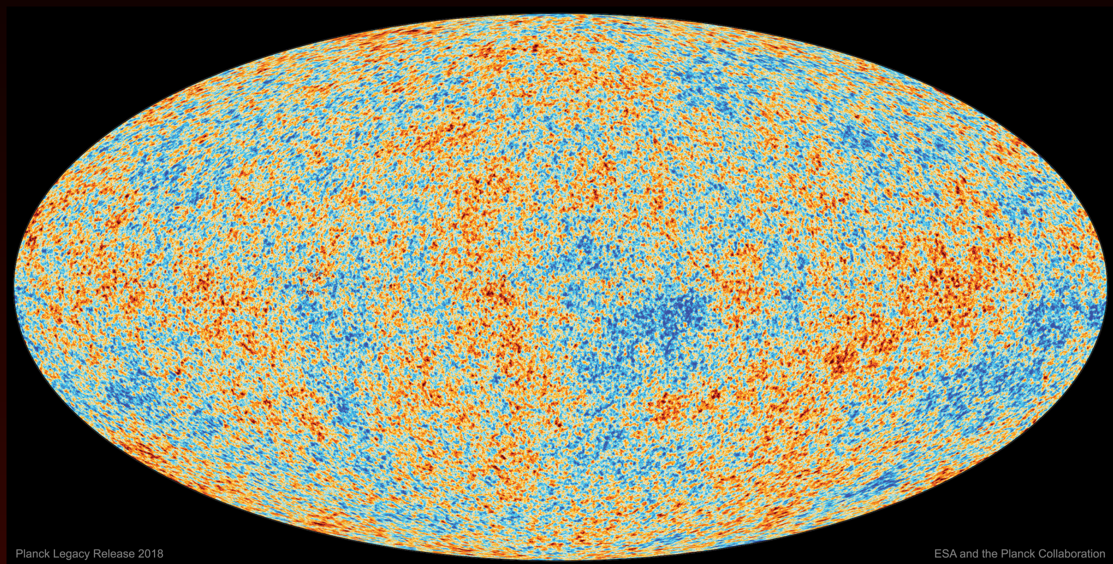
Bayesian Statistics

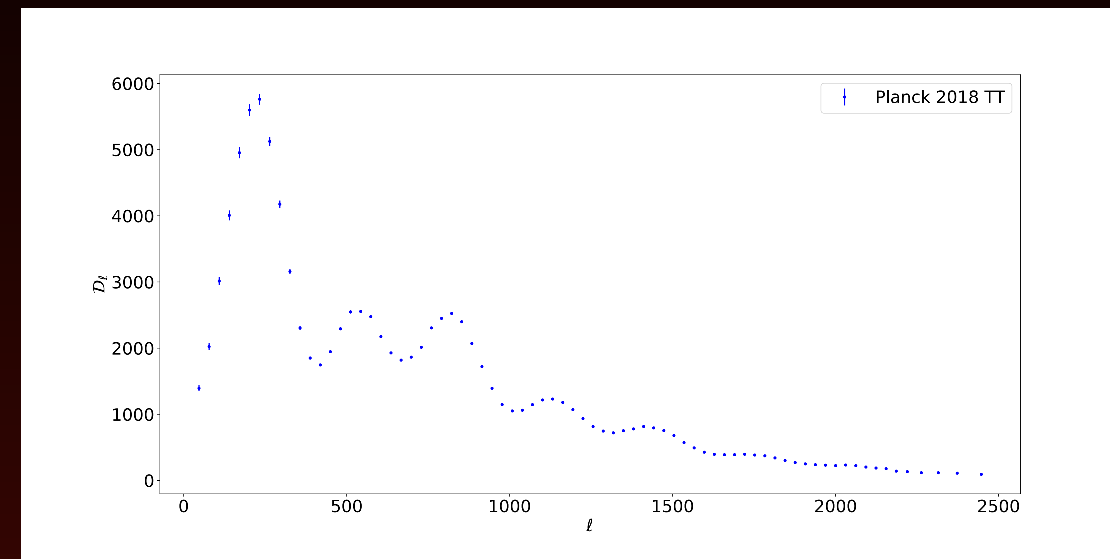
Bayes' Theorem

$$P(\theta|D,M) = \frac{P(D|\theta,M)\cdot P(\theta|M)}{P(D|M)}$$

$$\mathcal{P} = \frac{\mathcal{L}\times\Pi}{\mathcal{Z}}$$

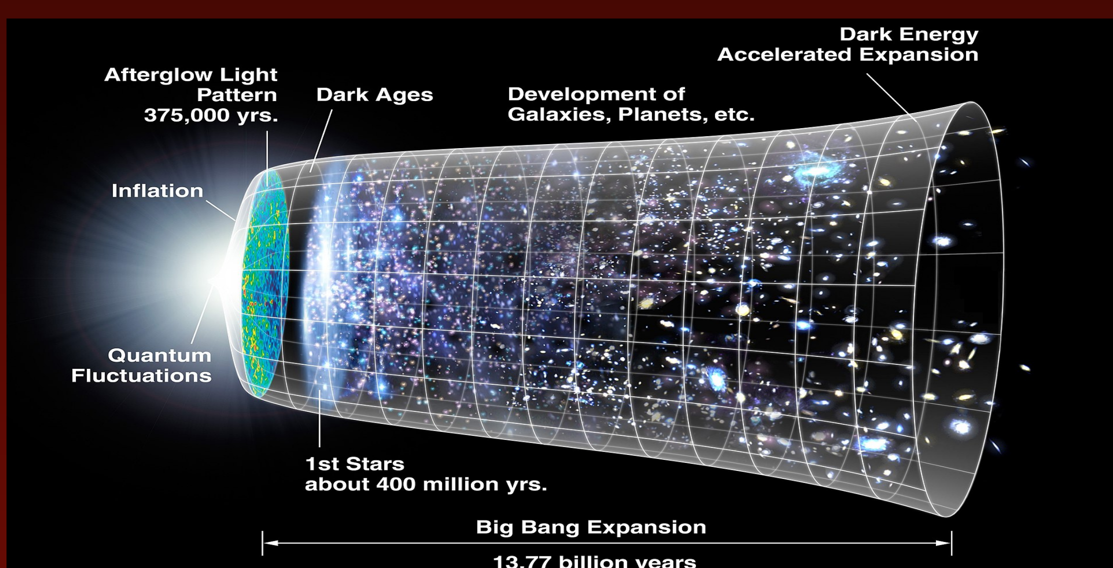
- **θ**: *Parameters*
- **D**: *Data*
- **M**: *Model*

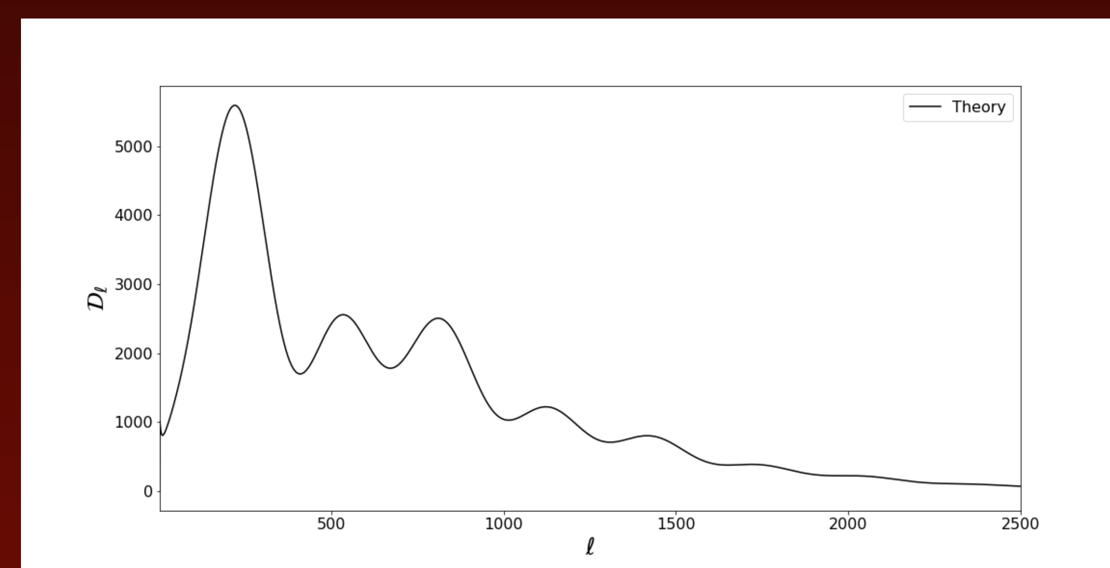




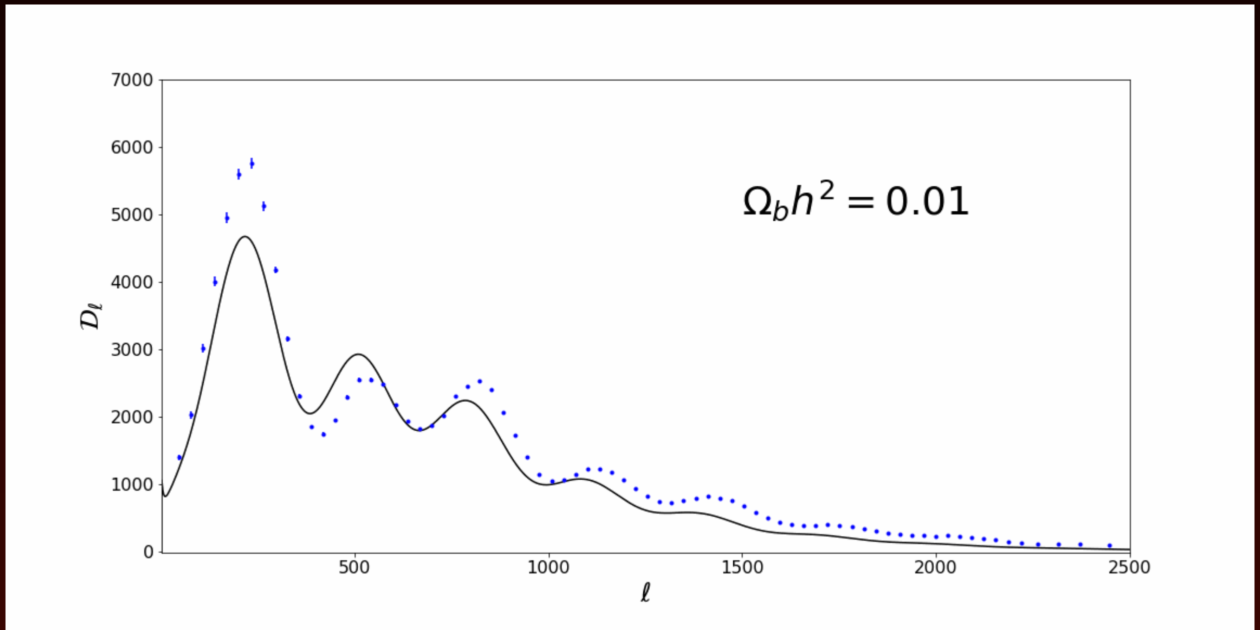
$$\mathscr{L} \equiv P(D\,|\,\theta, M)$$

$$\theta, M$$

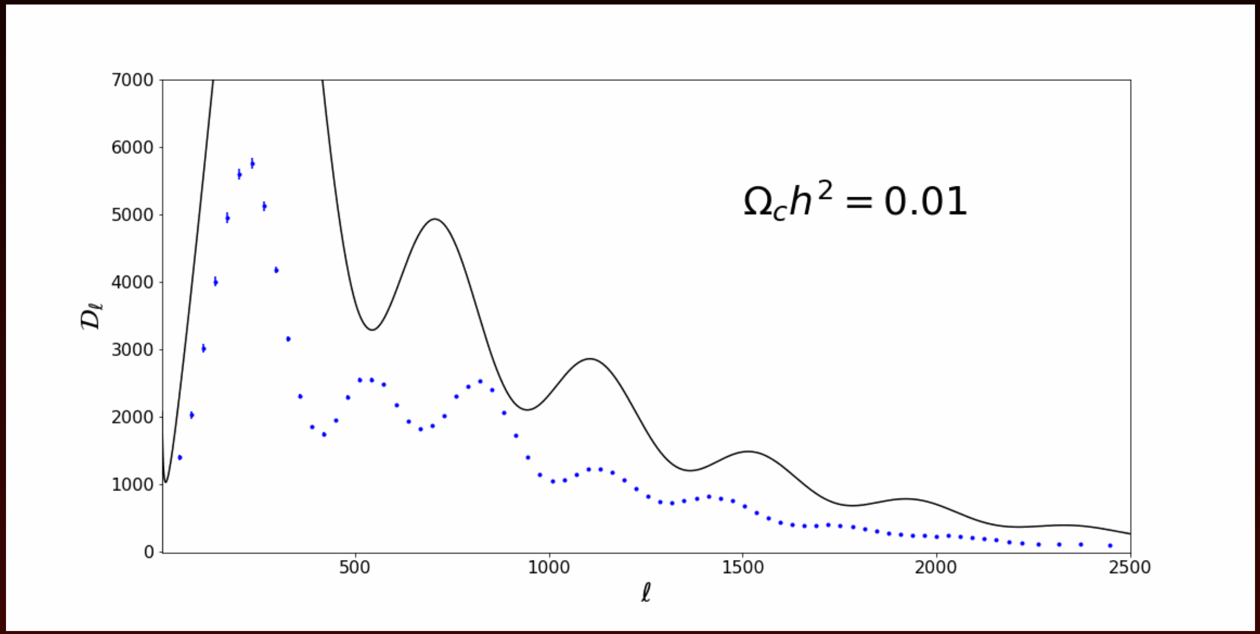




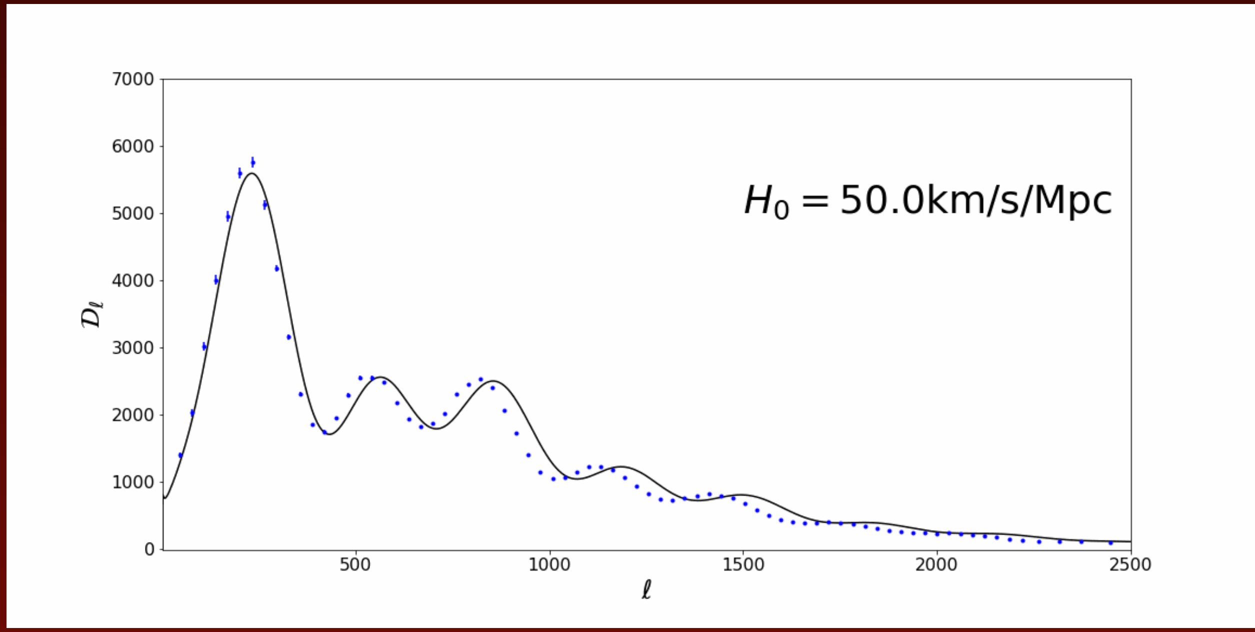
How much dark energy?



How much dark matter?



How fast does it expand?



Three types of problem

- Parameter estimation
- Model comparison
- Dataset comparison

'Tension'

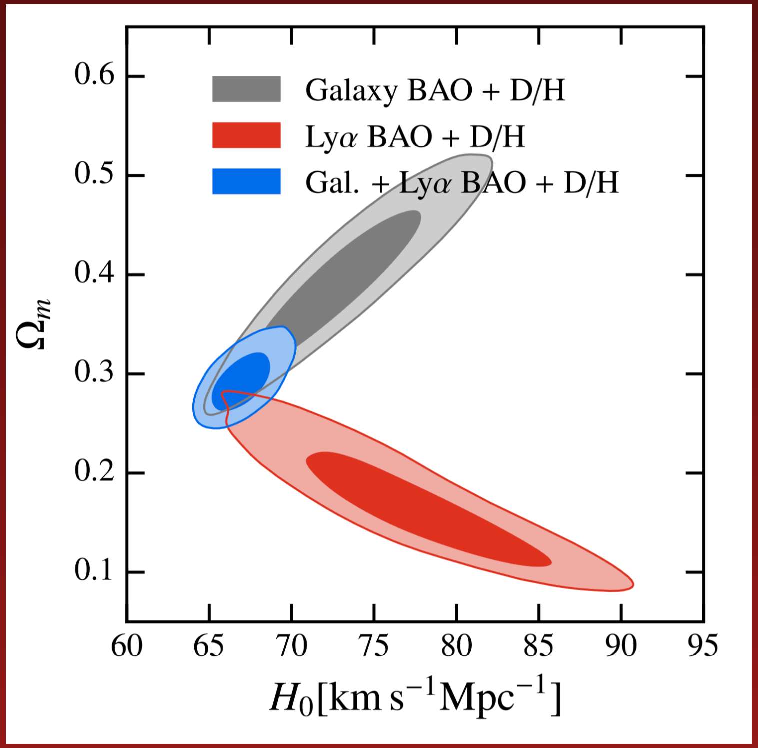
Bayesian Statistics

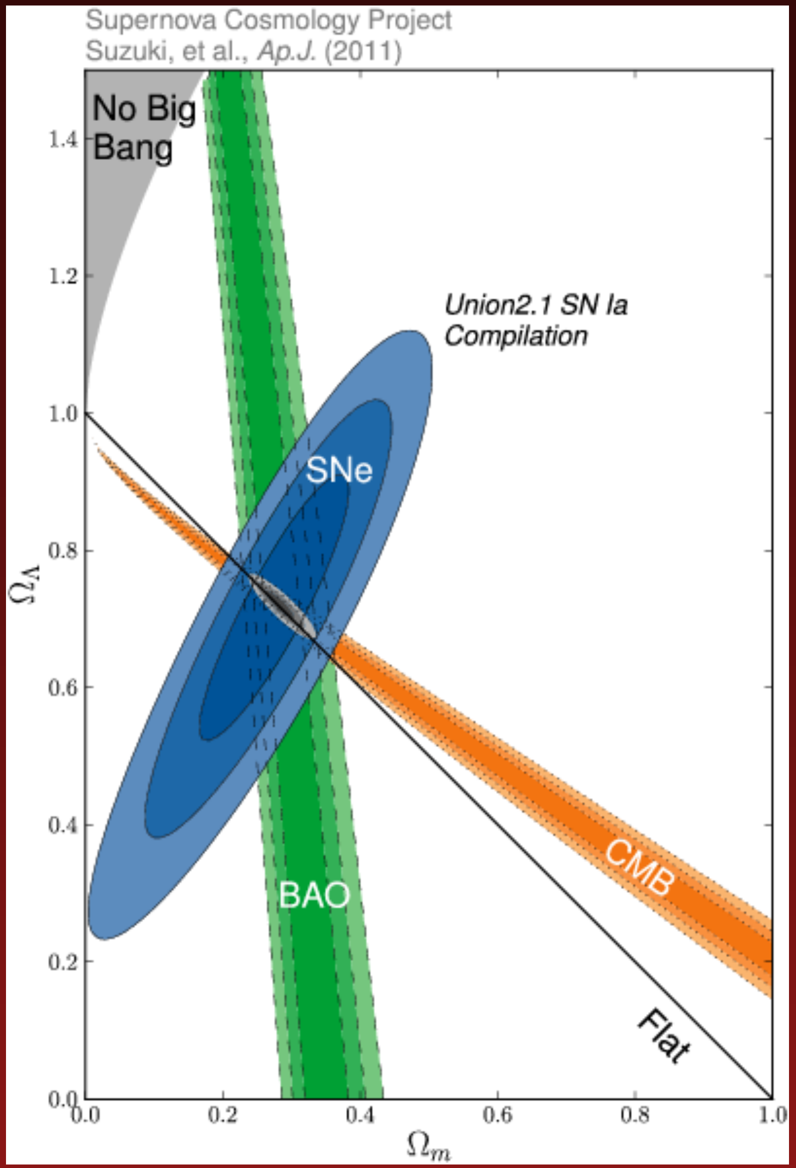
Why is tension important

- We can only combine data Sets that are **CONSISTENT**. Data set combinations are crucial to break degeneracies.

- If two data sets are in tension, there are two explanations: One (or both) data sets are wrong, or the underlying model is wrong.

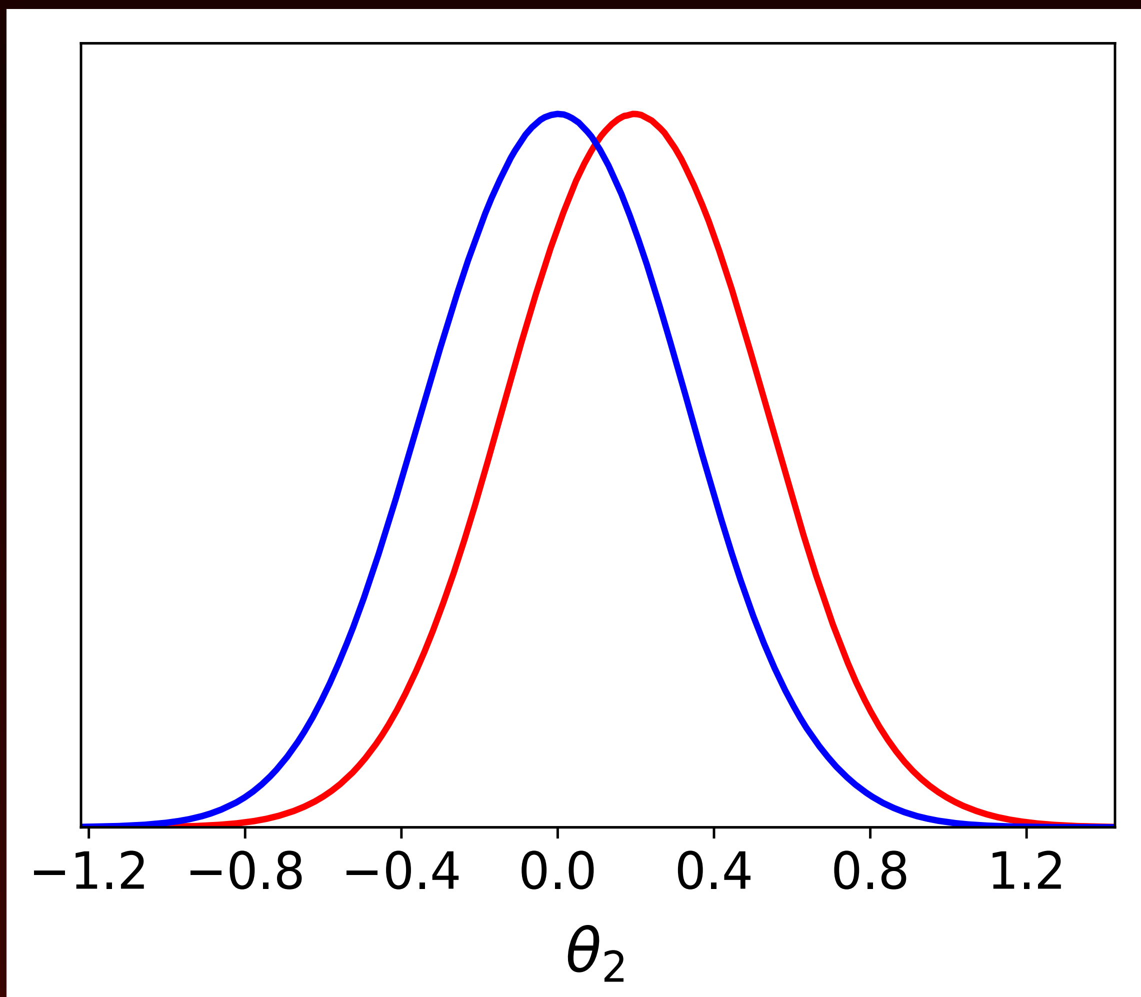
- We need a method to accurately quantify tension!



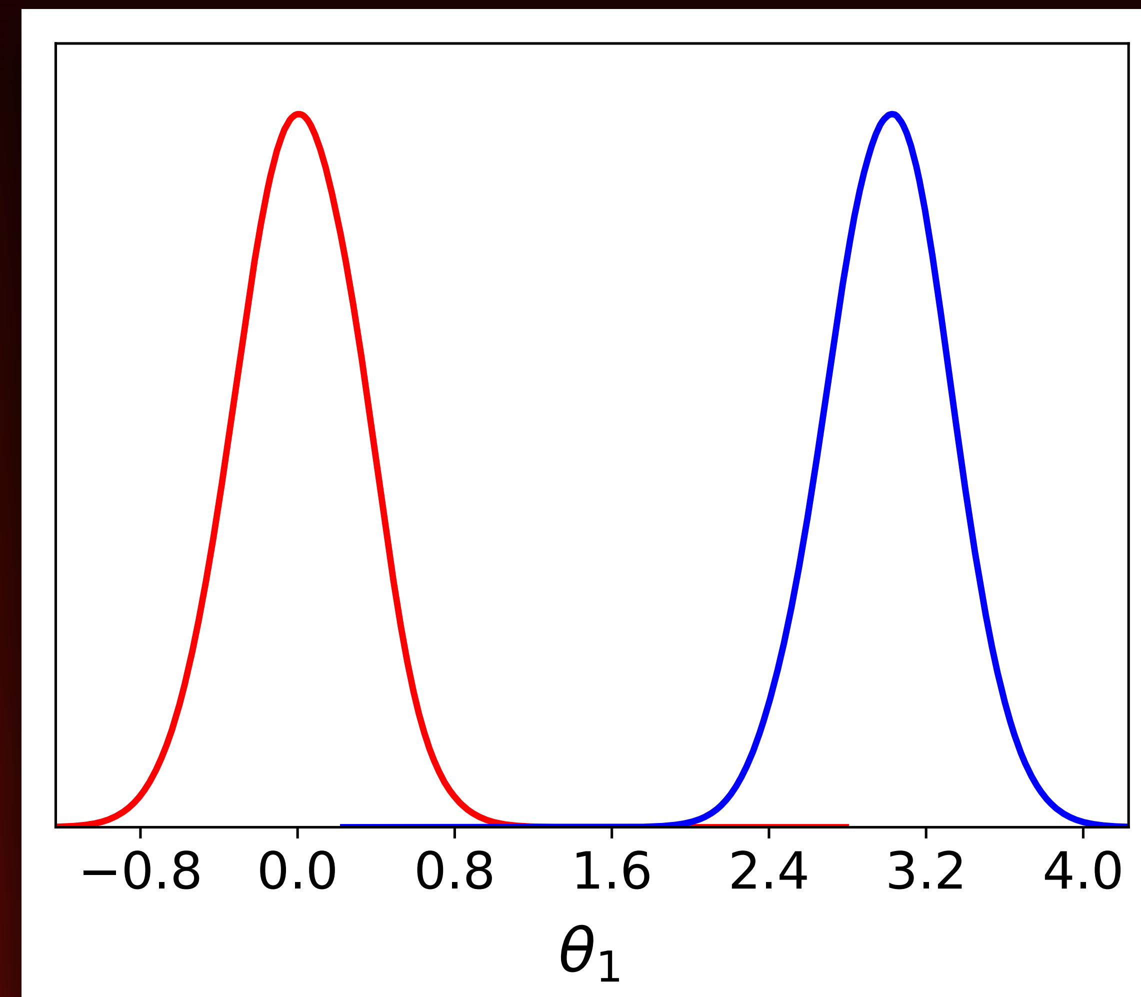


Why is Data Set Comparison Non-Trivial?

Trivial?

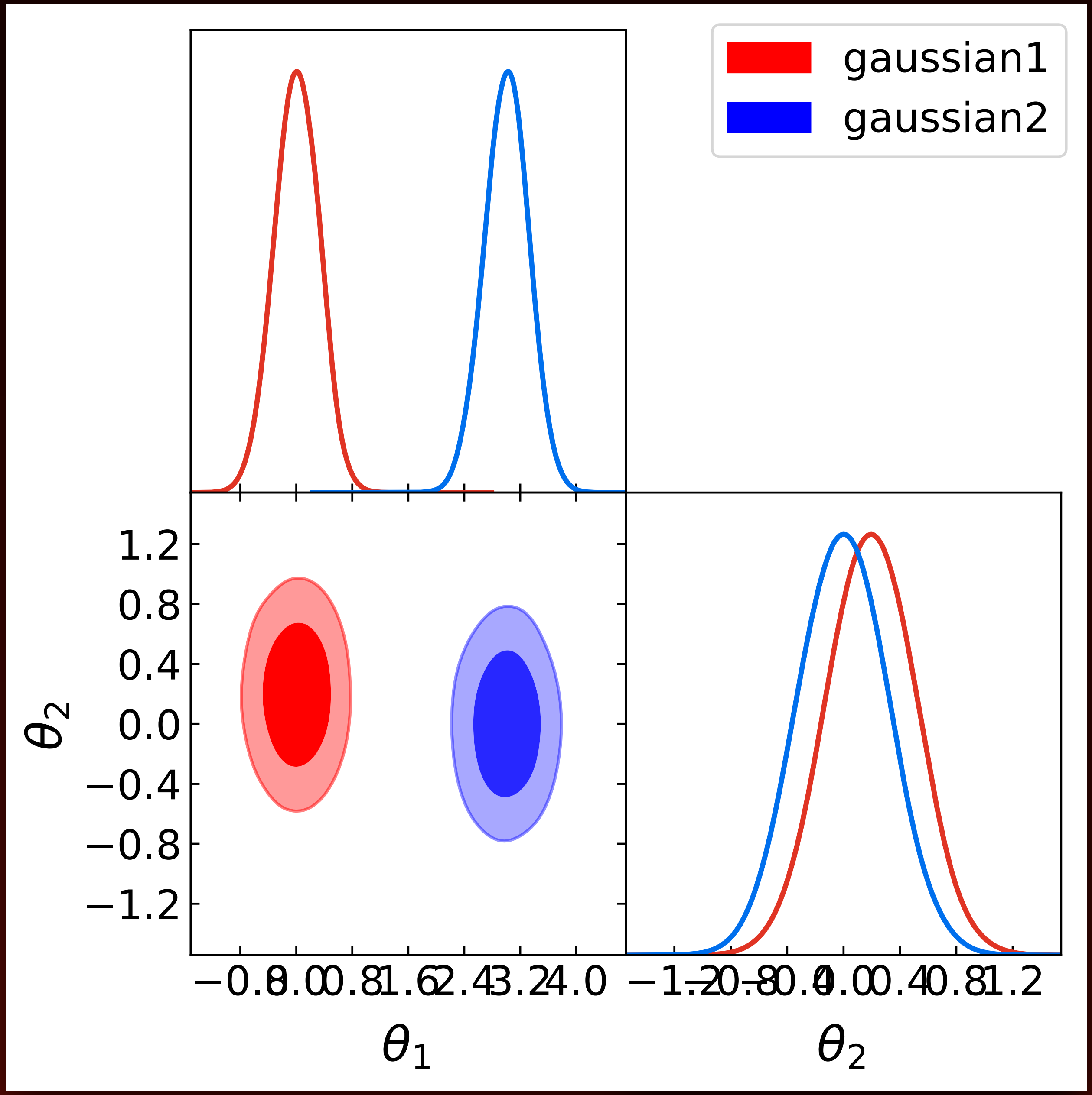


Consistent

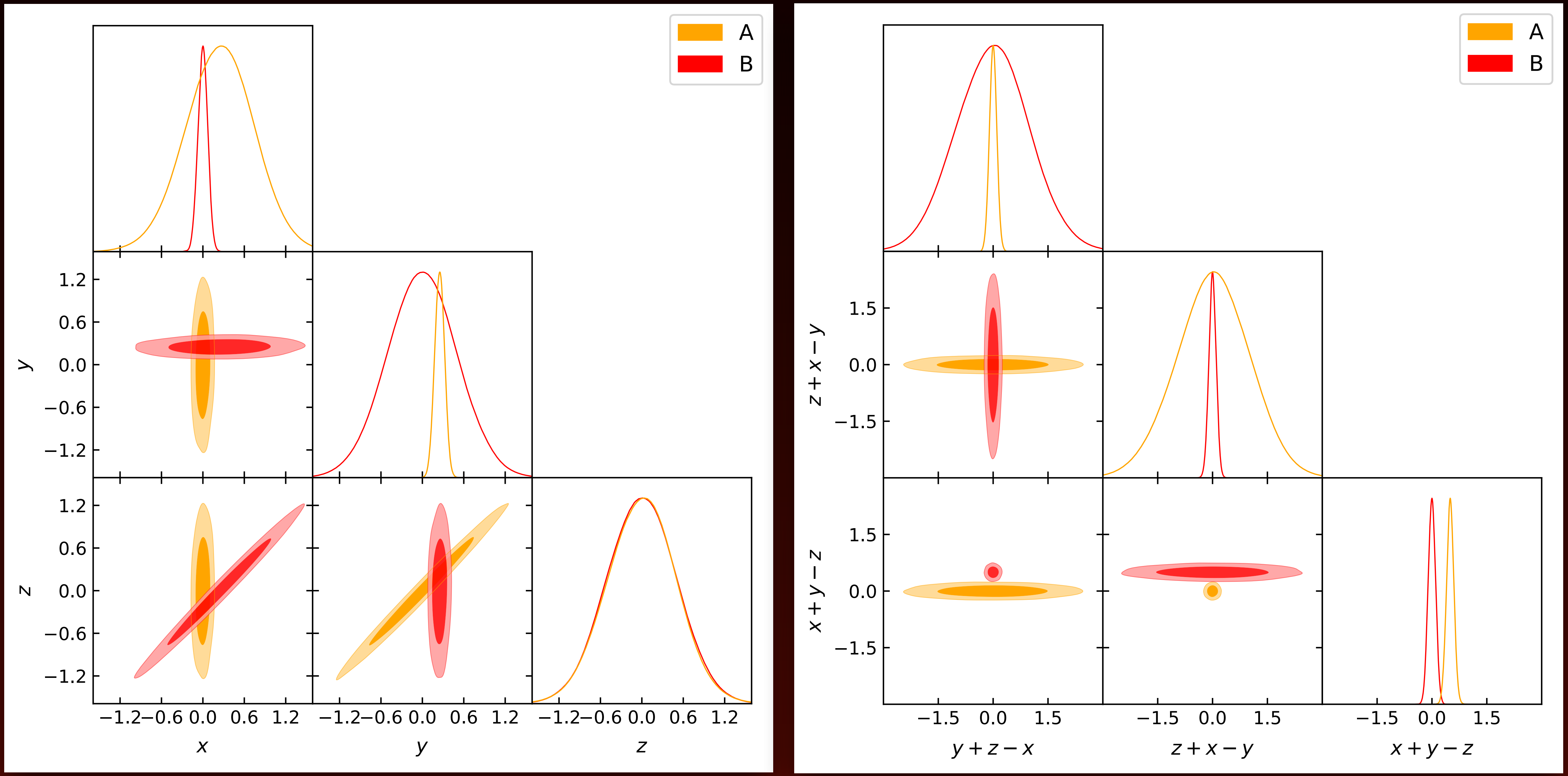


Inconsistent

Trivial?


gaussian1
gaussian2
1.2
0.8
0.4
0.0
−0.4
−0.8
−1.2
θ_2
θ_1
θ_2

Trivial?



A
B
1.2
0.6
0.0
−0.6
−1.2
y
z
−1.2 −0.6 0.0 0.6 1.2
x
−1.2 −0.6 0.0 0.6 1.2
y
−1.2 −0.6 0.0 0.6 1.2
z
A
B
1.5
0.0
−1.5
z + x − y
x + y − z
−1.5 0.0 1.5
y + z − x
−1.5 0.0 1.5
z + x − y
−1.5 0.0 1.5
x + y − z

The Bayes Ratio

Bayes Ratio

Bayesian evidence as a tool for comparing datasets

Phil Marshall
Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, USA

Nutan Rajguru
Astrophysics Group, Cavendish Laboratory, Madingley Road, Cambridge, UK

Anže Slosar
Faculty of Mathematics and Physics, University of Ljubljana, Slovenia

(Dated: February 2, 2008)

We introduce a new conservative test for quantifying the consistency of two or more datasets. The test is based on the Bayesian answer to the question, "How much more probable is it that all my data were generated from the same model system than if each dataset were generated from an independent set of model parameters?". We make explicit the connection between evidence ratios and the differences in peak chi-squared
more cheaply calculated. Calculating evidenc
data (WMAP, ACBAR, CBI, VSA), SDSS a
concordance is favoured and the tightening
justified.

Probability that both datasets come from **THE SAME** Universe

$$R \equiv \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \times \mathcal{Z}_B}$$

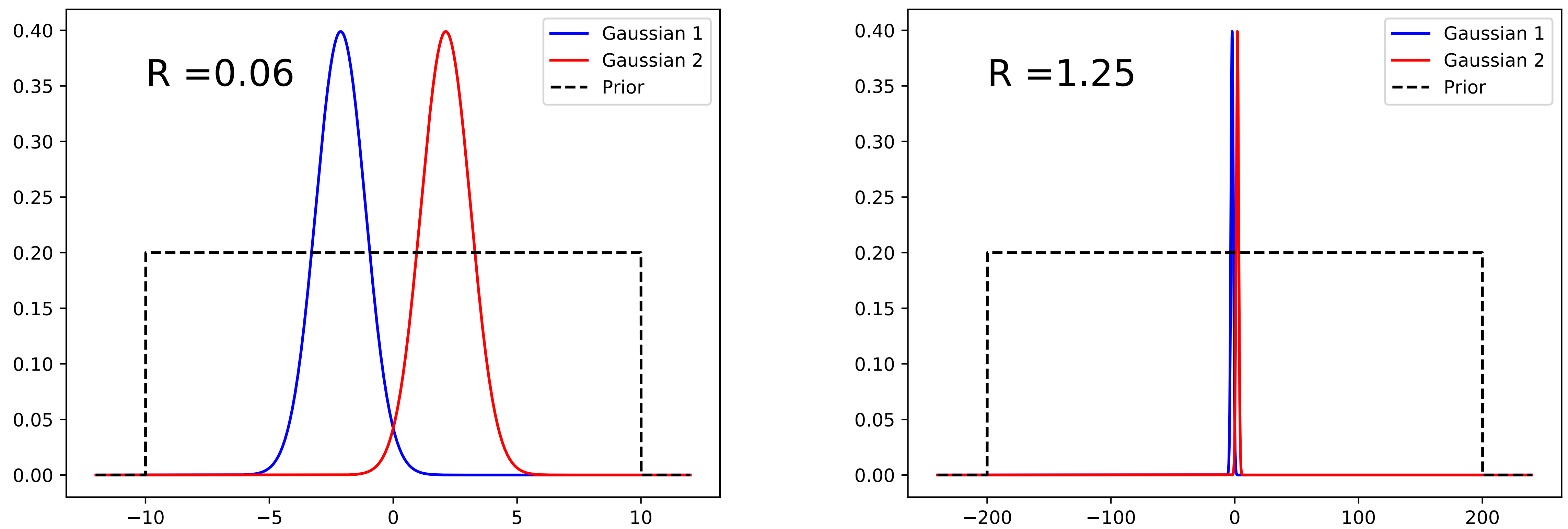
$\mathcal{Z} = \int$

Probability that both datasets come **DIFFERENT** Universes

Bayes Ratio

Toy example: 1D Gaussians:

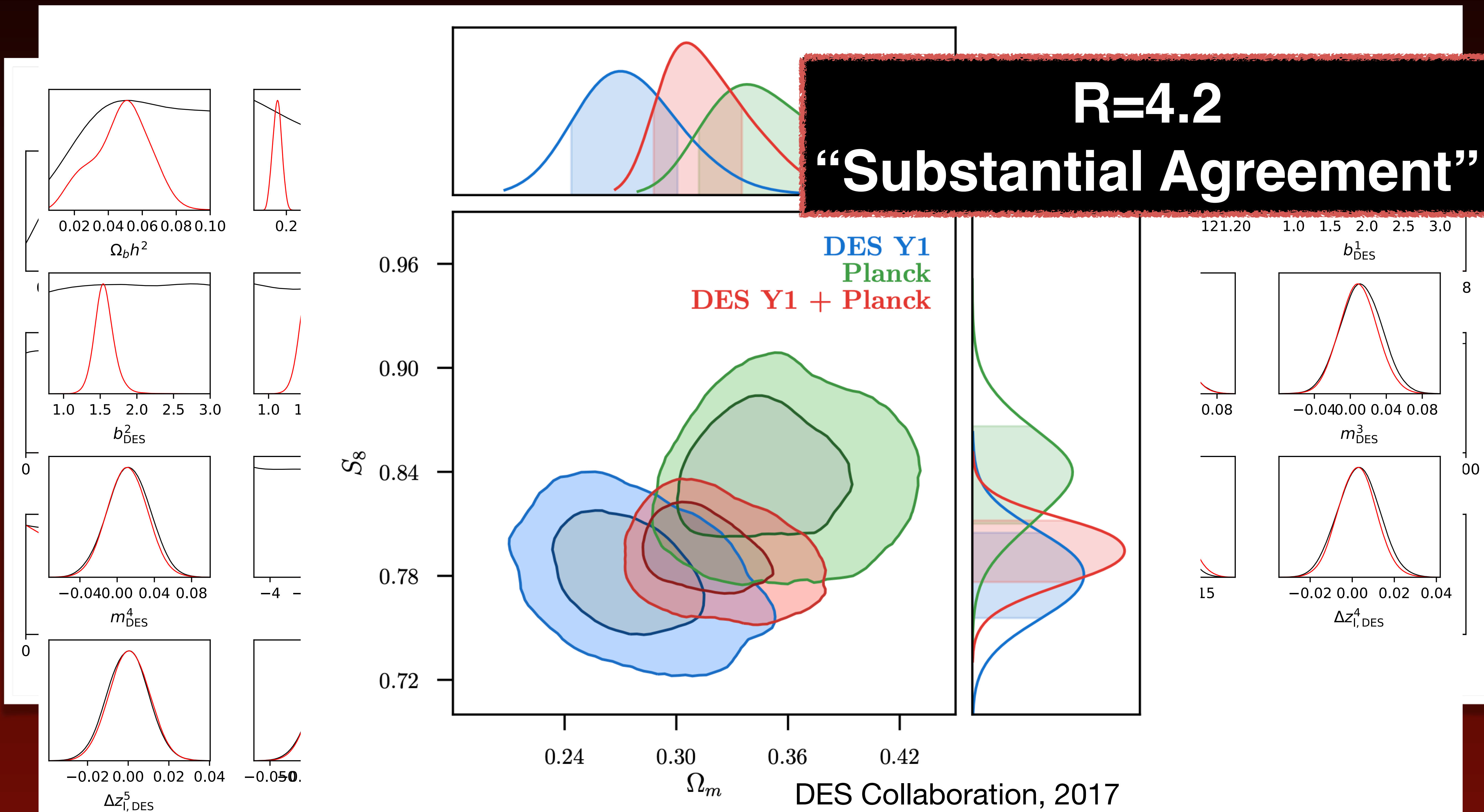
$$R \equiv \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \times \mathcal{Z}_B} \propto V_\pi$$



Prior: -10, 10 -> **Discordant** Prior: -200, 200 -> **Concordant**

Bayes Ratio

Is this a problem in Cosmology?



R=4.2

“Substantial Agreement”

DES Collaboration, 2017

The ‘Suspiciousness’

In collaboration with:
Will Handley

Proposition 1

<u>Proposition 1:</u>

If there are ***any*** *physically reasonable priors which render R significantly less than 1, then as Bayesians we should consider these datasets* ***in tension.***

Handley & PL, 2019, arXiv: 1902.04029
10.1103/PhysRevD.100.043504

Suspiciousness

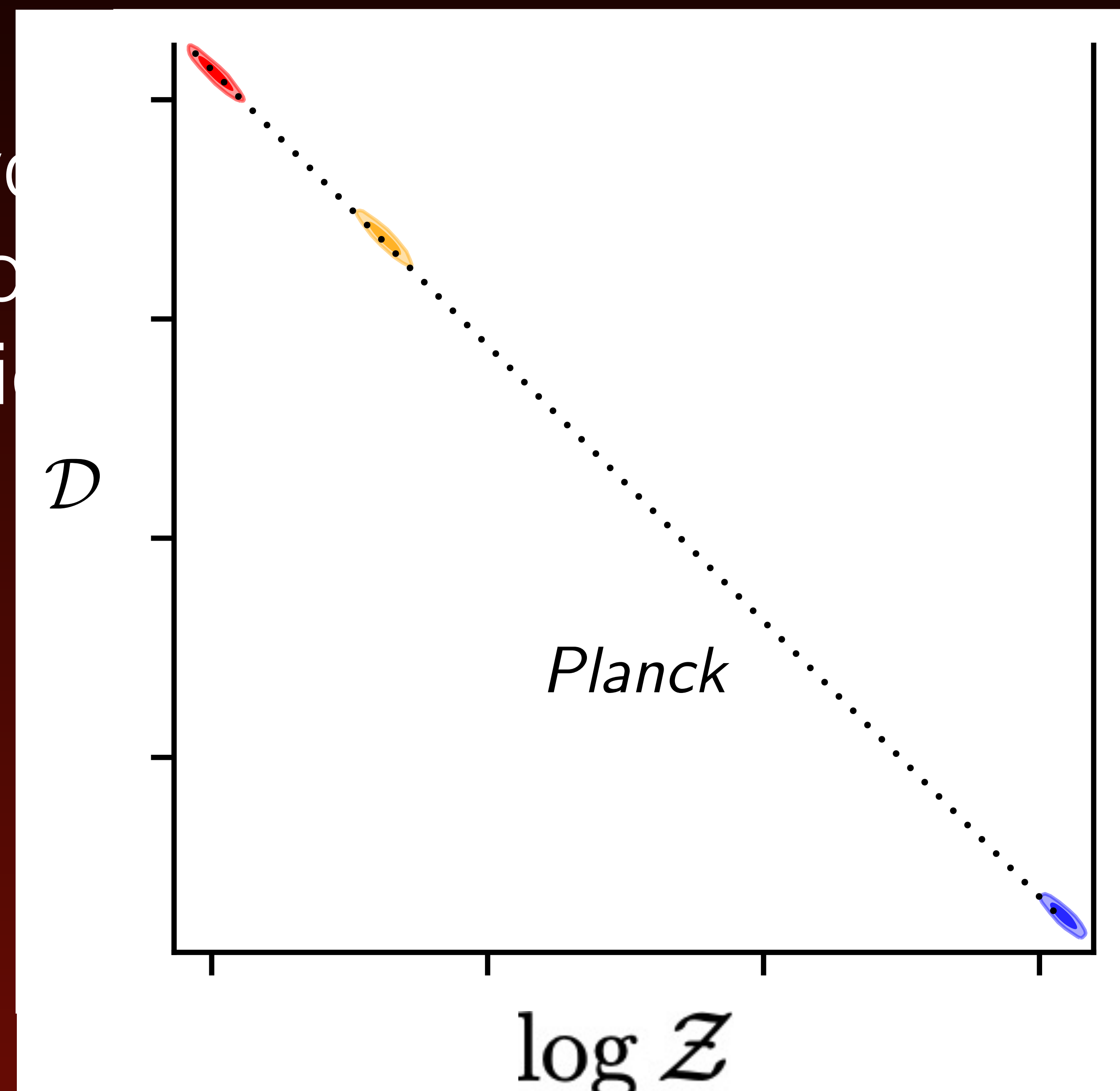
We want a method that

- Is formed of fully Bayesian quantities.
- Is independent of choice of parameterisation.
- **Has an intuitive interpretation**
- **Does not depend on prior volume**

Suspiciousness

What part of the Bayes Ratio carries the ‘prior volume dependence’?

i.e. if I double the prior vo
many possible states, so
becomes twi



Kullback-Leibler Divergence

$$\mathcal{D} \equiv \int \mathrm{d}\theta \; \mathcal{P} \log \left(\frac{\mathcal{P}}{\Pi} \right)$$

Kullback, Leibler, 1951
doi:10.1214/aoms/1177729694

Suspiciousness

So a part of the **BAYES RATIO (R)**:

$$\log R = \log \mathcal{Z}_{AB} - \log \mathcal{Z}_A - \log \mathcal{Z}_B$$

Encloses its dependence on the prior volume. We call this part the **INFORMATION (I)**:

$$\log I = \mathcal{D}_A + \mathcal{D}_B - \mathcal{D}_{AB}$$

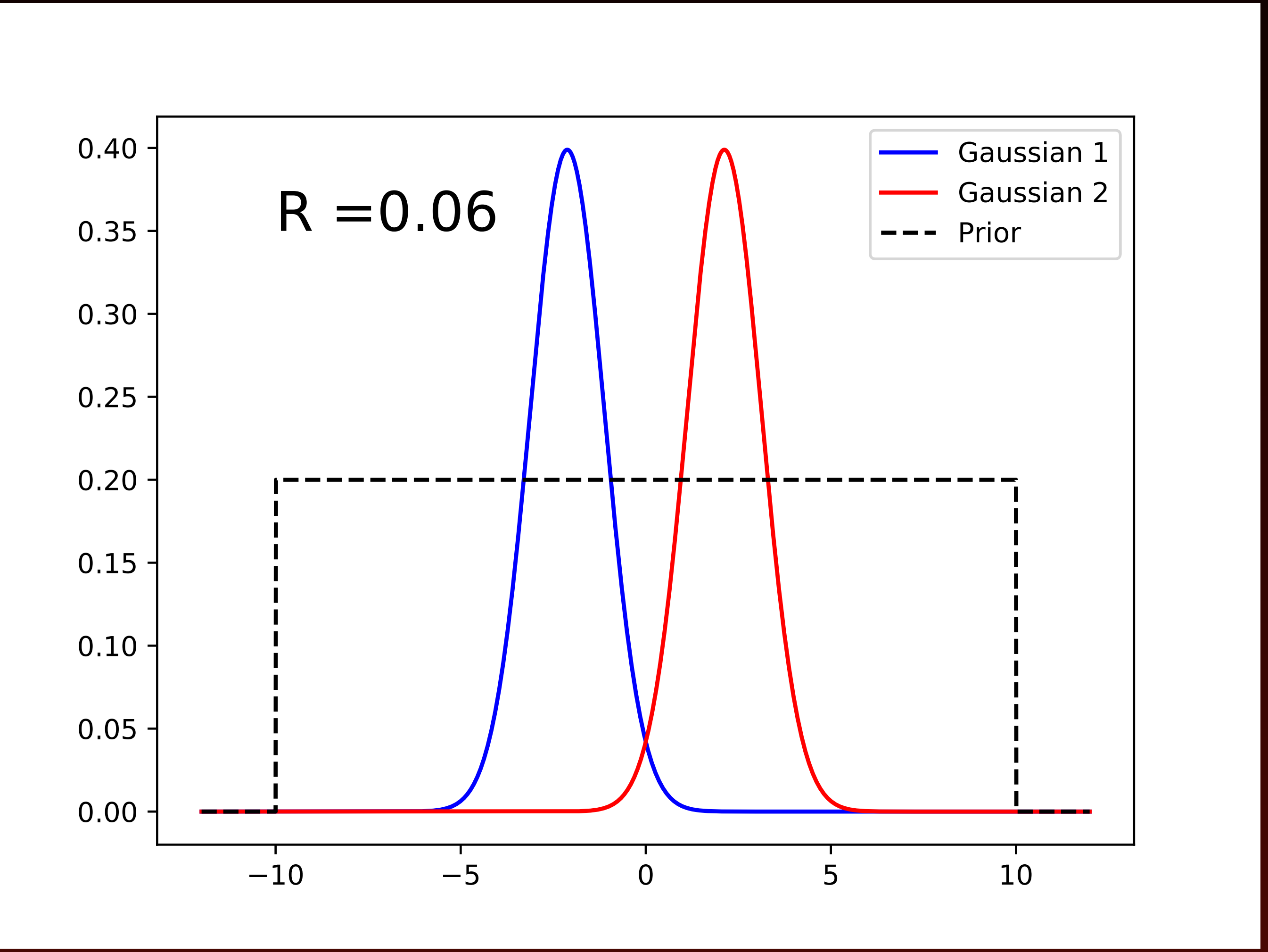
The part of R that is left, is what we call the **SUSPICIOUSNESS (S)**:

$$\log R = \log I + \log S$$

Suspiciousness

For **Gaussian Likelihoods**, the Suspiciousness follows a **chi-squared distribution**.

Therefore we can assign a **tension probability**, and interpret the result with a ‘number of sigma’ tension.

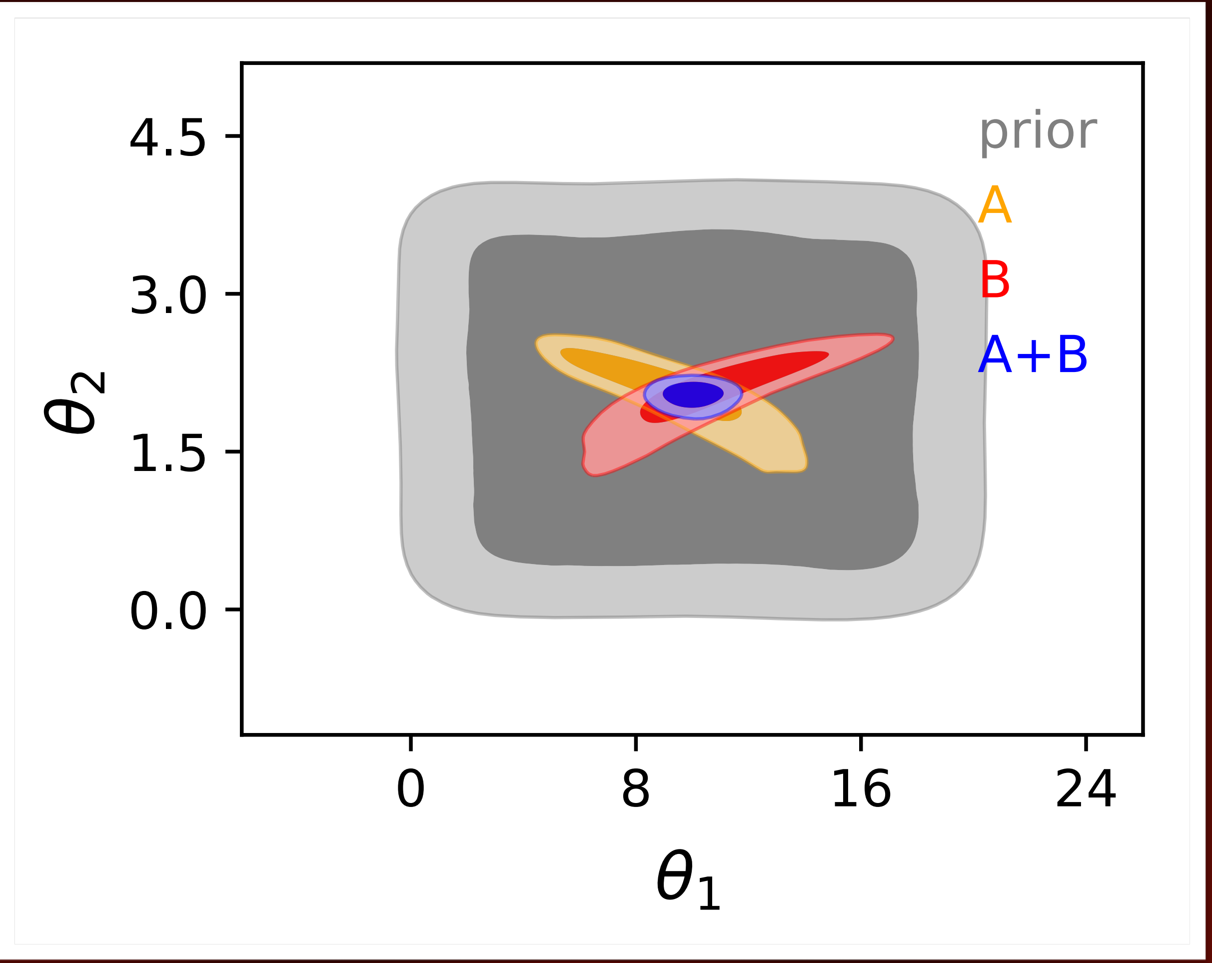


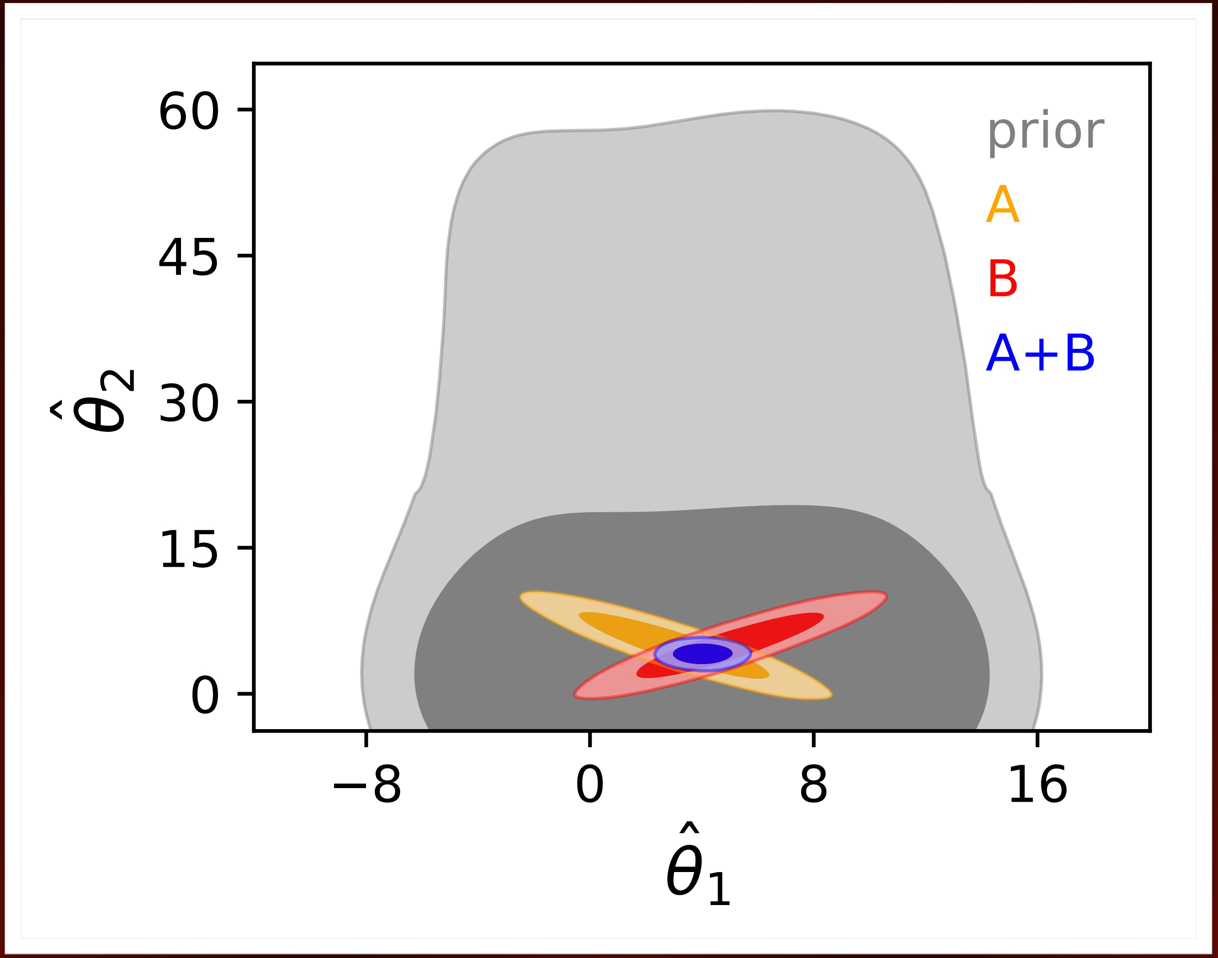
$$p = \int_{d-2\log S}^{\infty} \chi^2_d(x)\,\mathrm{d}x = \int_{d-2\log S}^{\infty} \frac{x^{d/2-1}e^{-x/2}}{2^{d/2}\Gamma(d/2)}\,\mathrm{d}x$$

Suspiciousness

What about non-Gaussian posteriors

Box Cox transformations can 'Gaussianize' the posterior, and they preserve the Suspiciousness



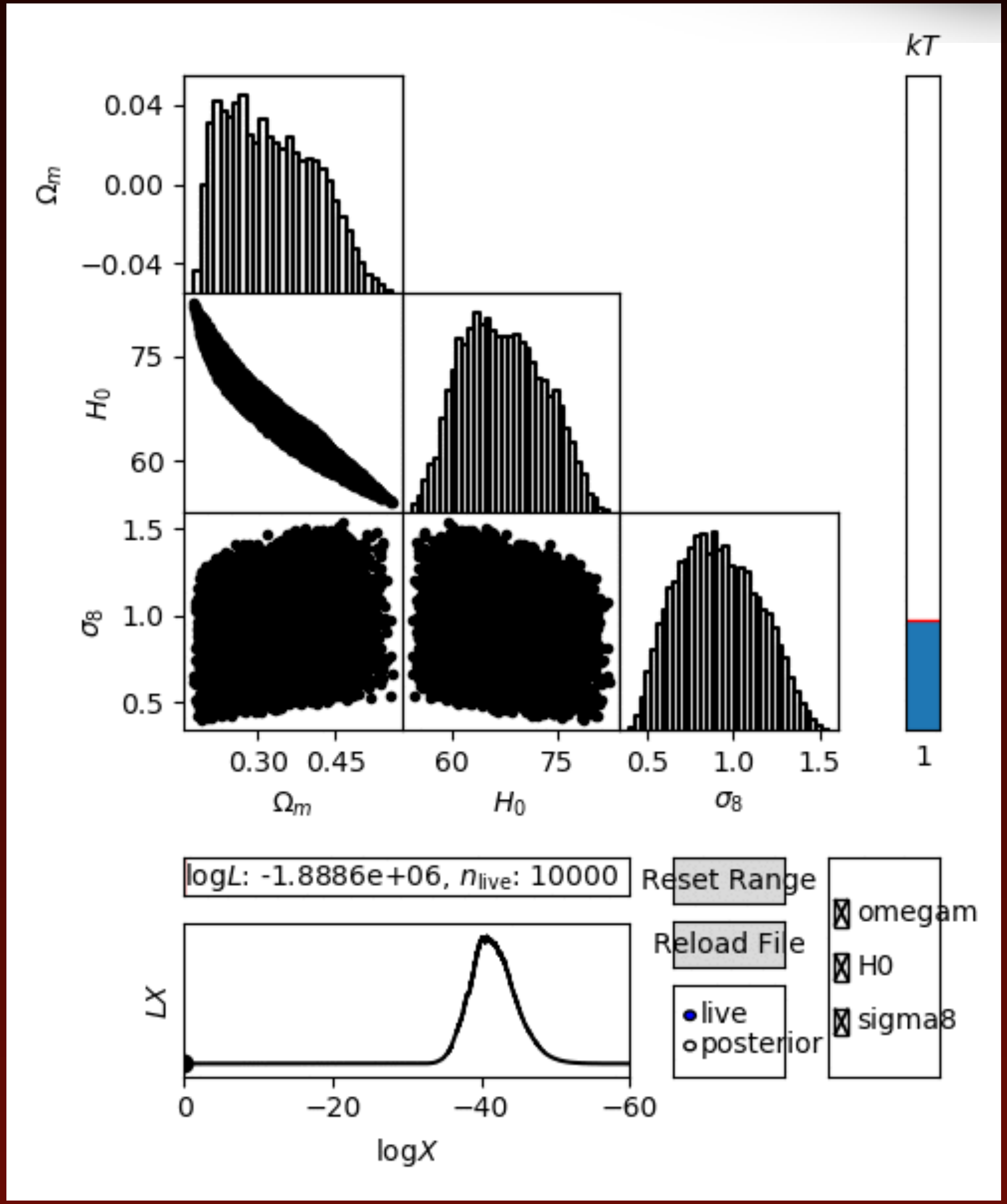


How to calculate this

Anesthetic

https://github.com/williamjameshandley/anesthetic

- Public python code
- Computation of Evidences, KL divergences, Bayesian model dimensionalities…
- Marginalised 1D and 2D plots
- Dynamic replaying of nested sampling

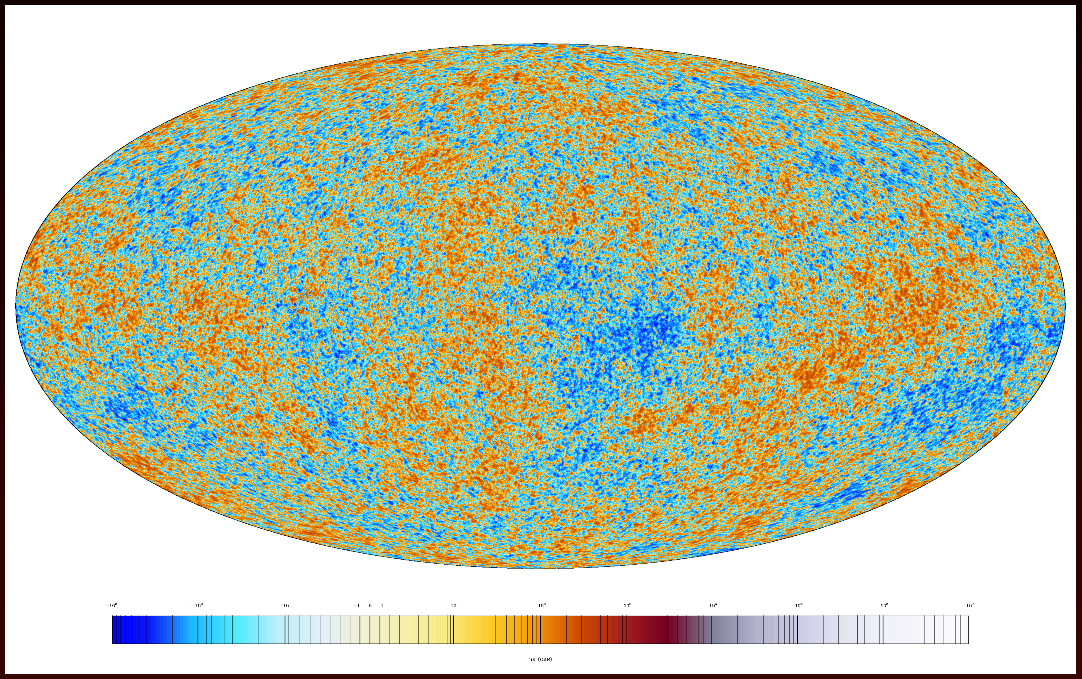


How does this work in practice?

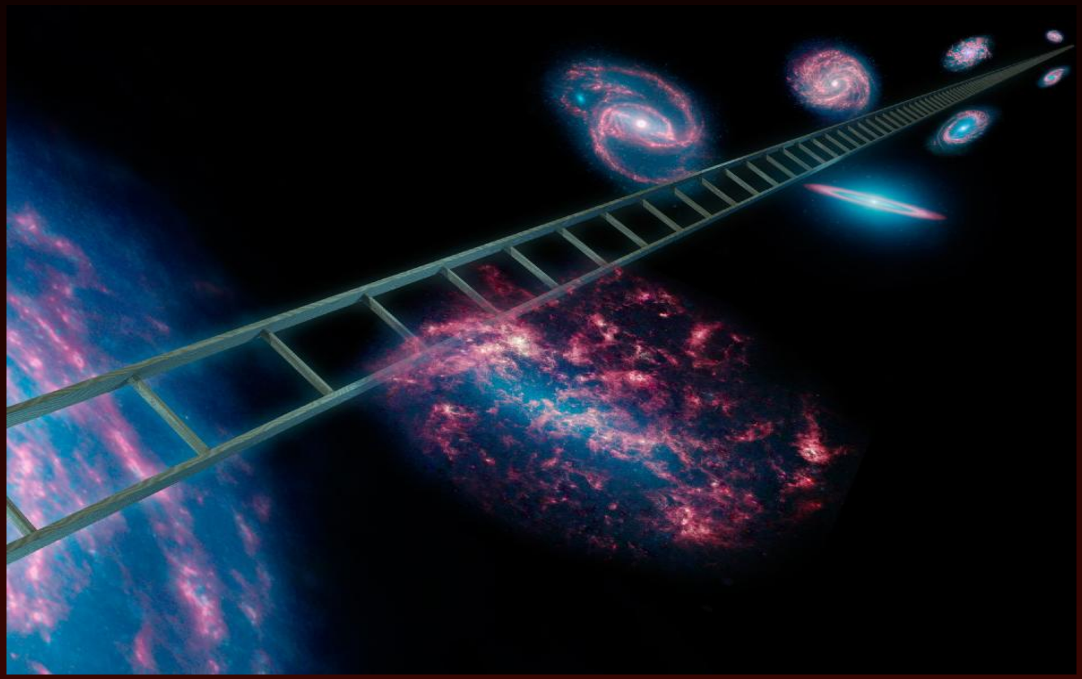
https://github.com/Pablo-Lemos/Suspiciousness-CosmoSIS.git

Application to Cosmology

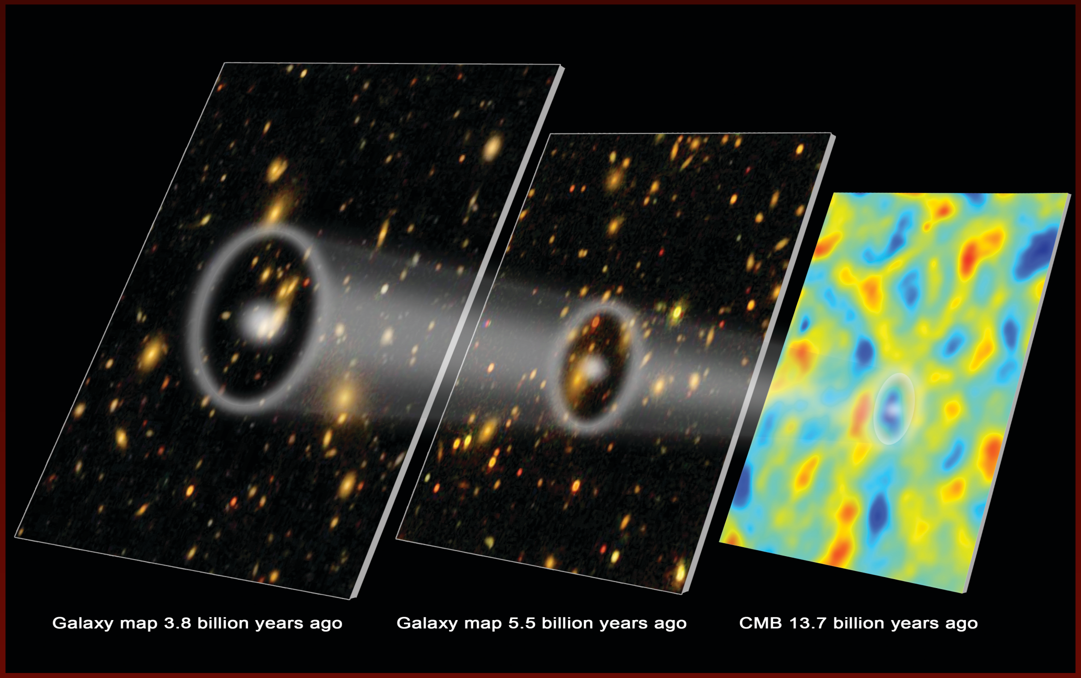
Application to Cosmology



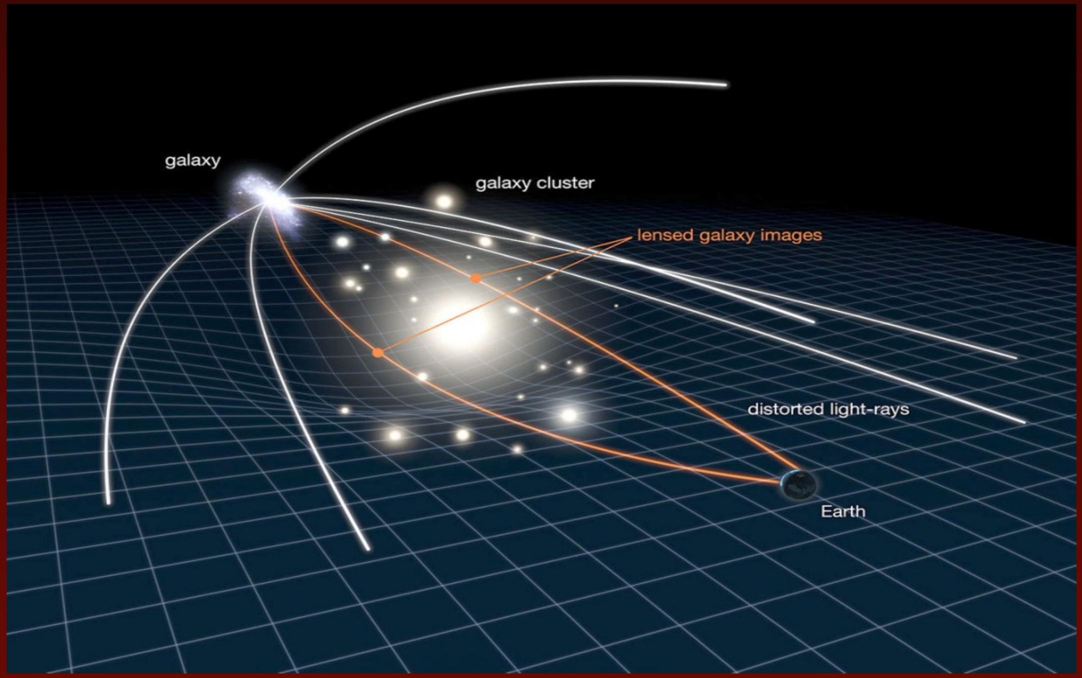
Cosmic Microwave Background (CMB) - ***PLANCK***



Cosmic Distance Ladder - ***SH0ES***



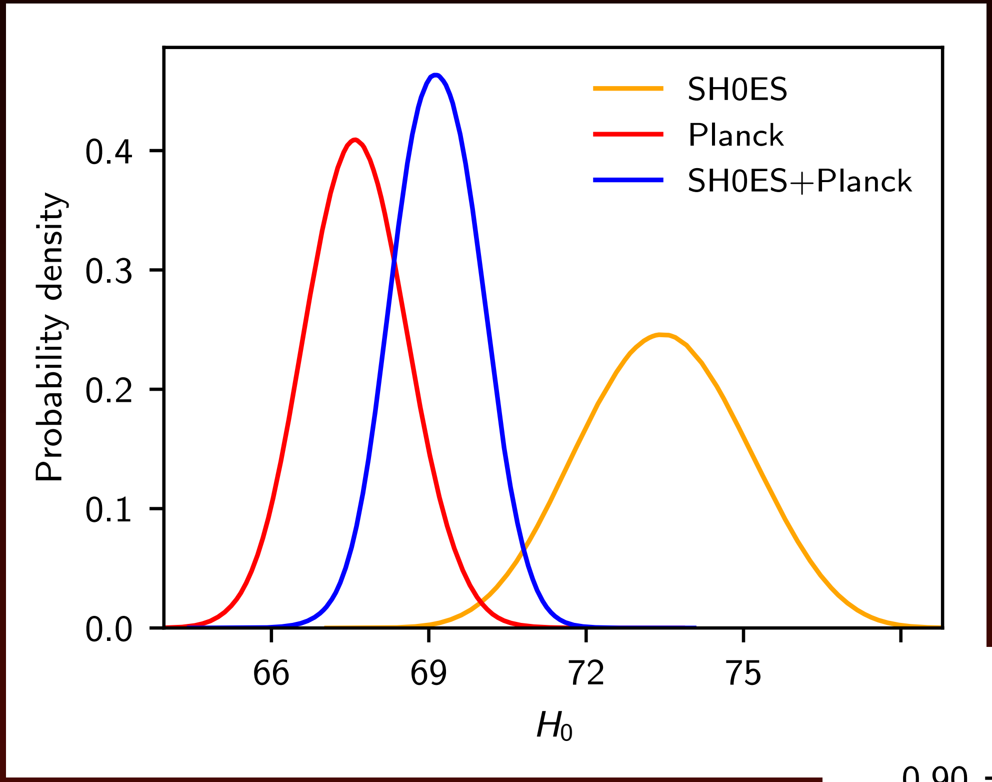
Baryon Acoustic Oscilaltions (BAO) - ***BOSS***



Weak Galaxy Lensing - ***DES***

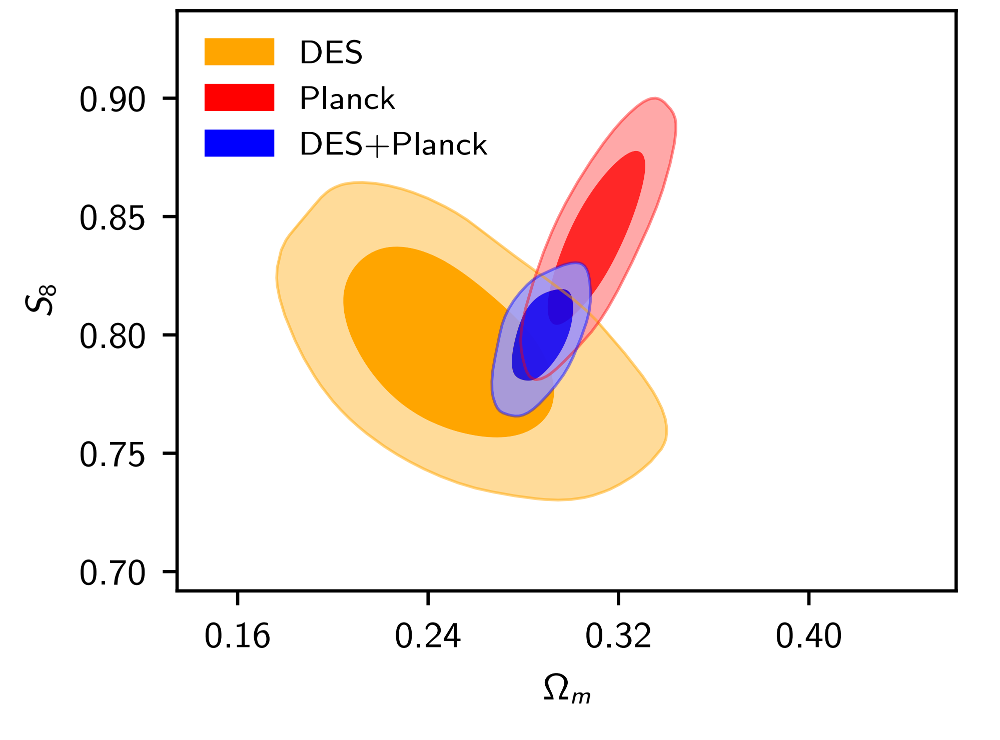
Application to Cosmology

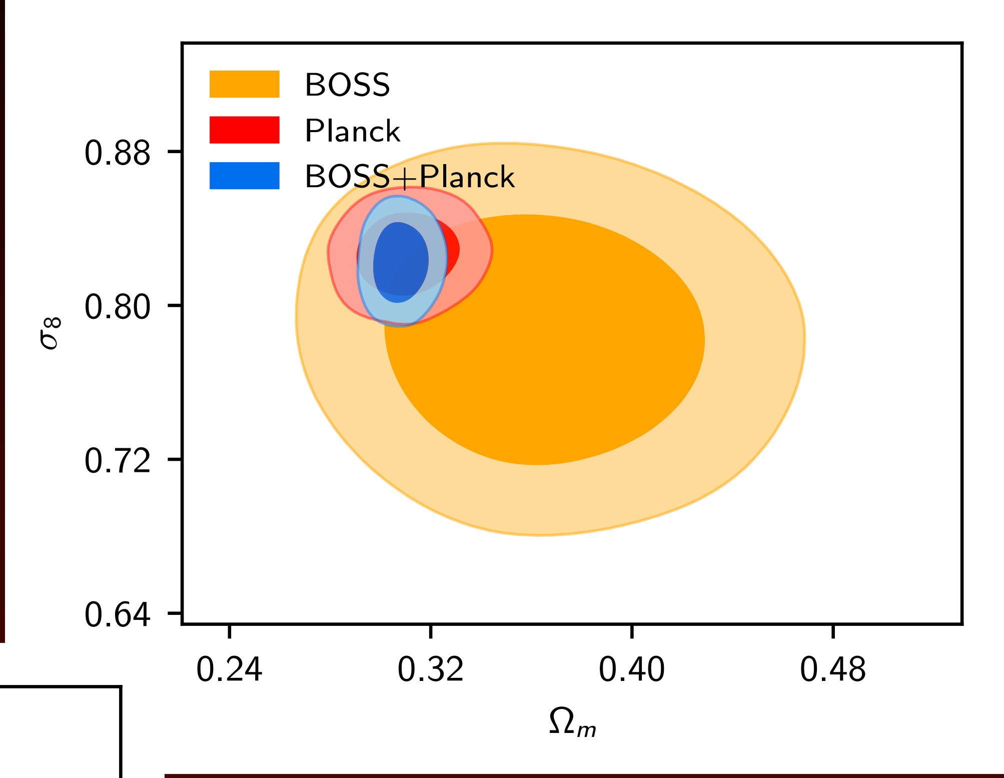
Planck vs…



SH0ES

DES

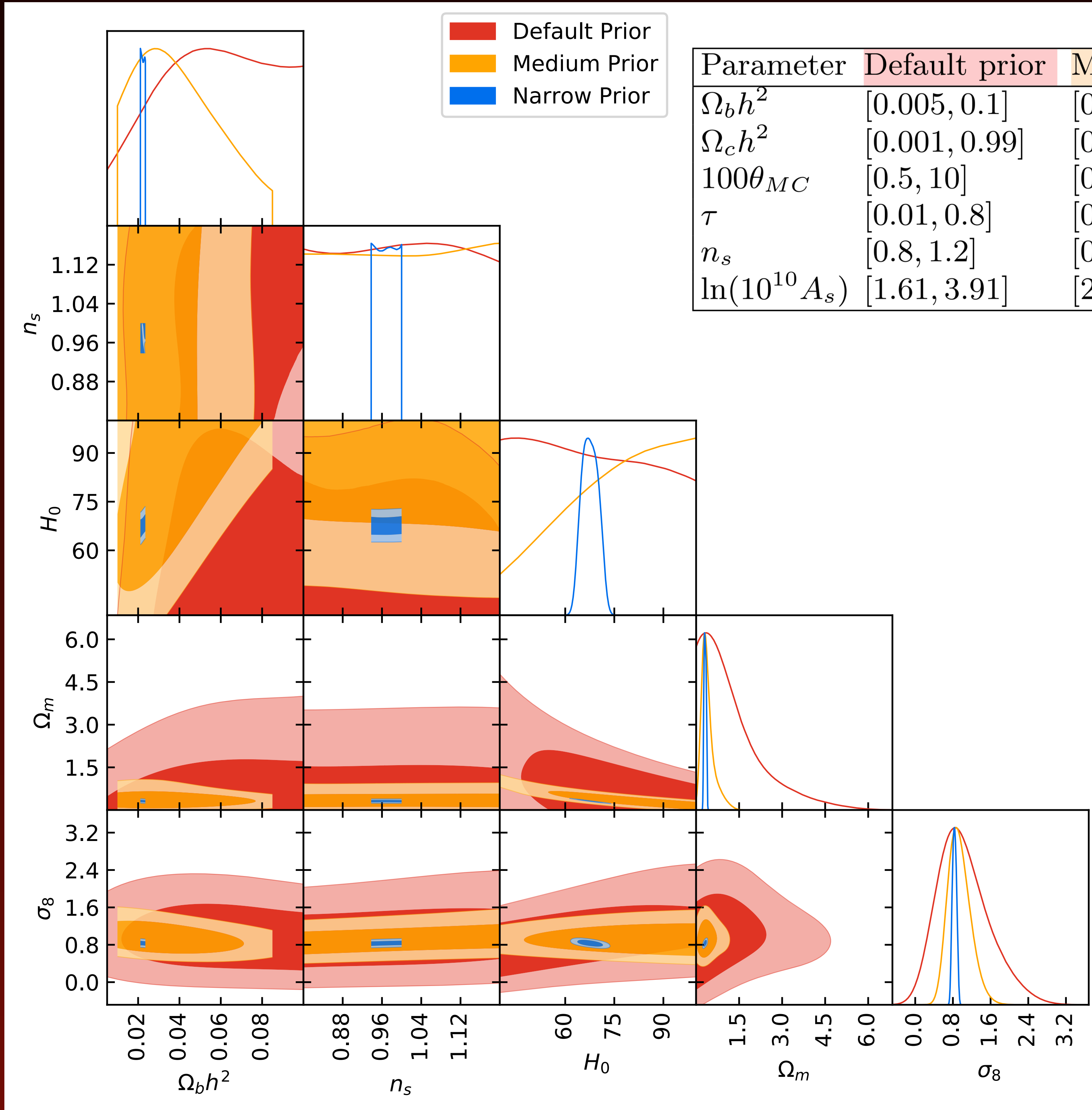




BOSS

Application to Cosmology

Three different priors



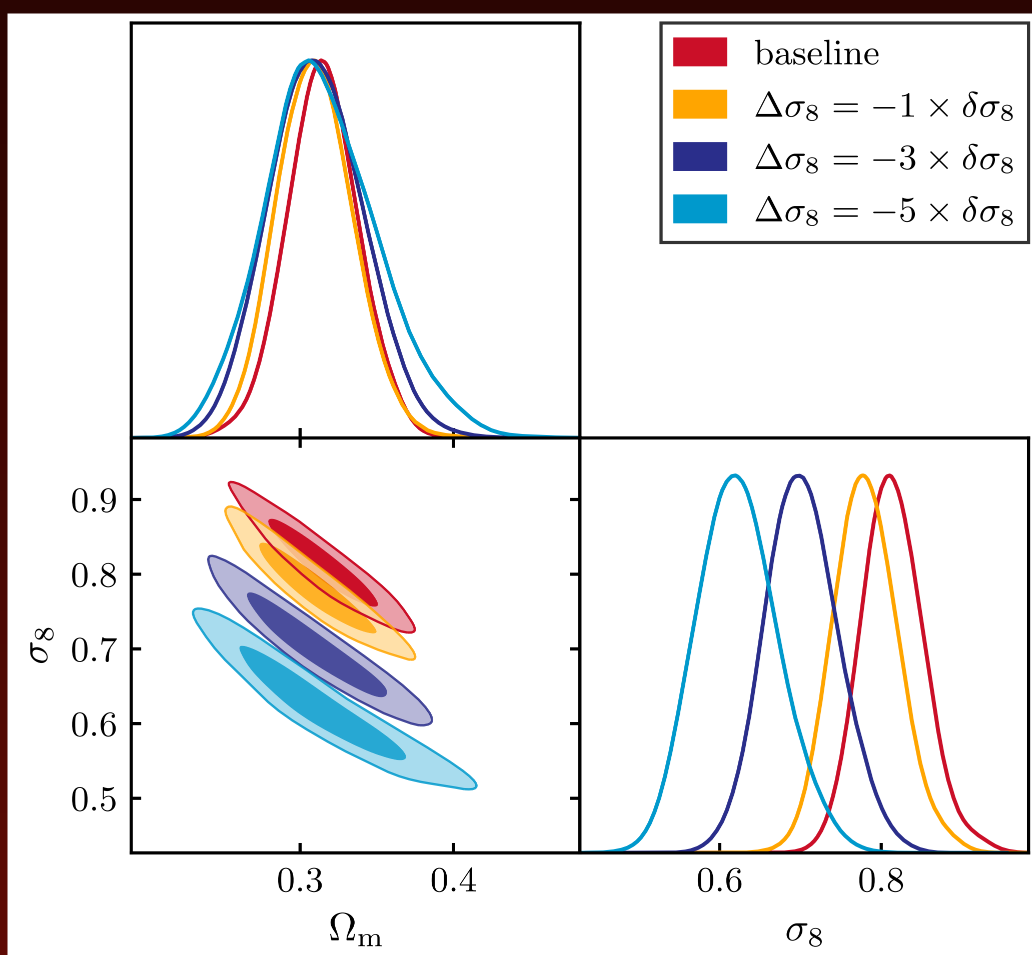
Parameter	Default prior	Medium prior	Narrow prior
$\Omega_b h^2$	$[0.005, 0.1]$	$[0.01, 0.085]$	$[0.0211, 0.0235]$
$\Omega_c h^2$	$[0.001, 0.99]$	$[0.08, 0.21]$	$[0.108, 0.131]$
$100\theta_{MC}$	$[0.5, 10]$	$[0.97, 1.5]$	$[1.038, 1.044]$
τ	$[0.01, 0.8]$	$[0.01, 0.8]$	$[0.01, 0.16]$
n_s	$[0.8, 1.2]$	$[0.8, 1.2]$	$[0.938, 1]$
$\ln(10^{10} A_s)$	$[1.61, 3.91]$	$[2.6, 3.8]$	$[2.95, 3.25]$

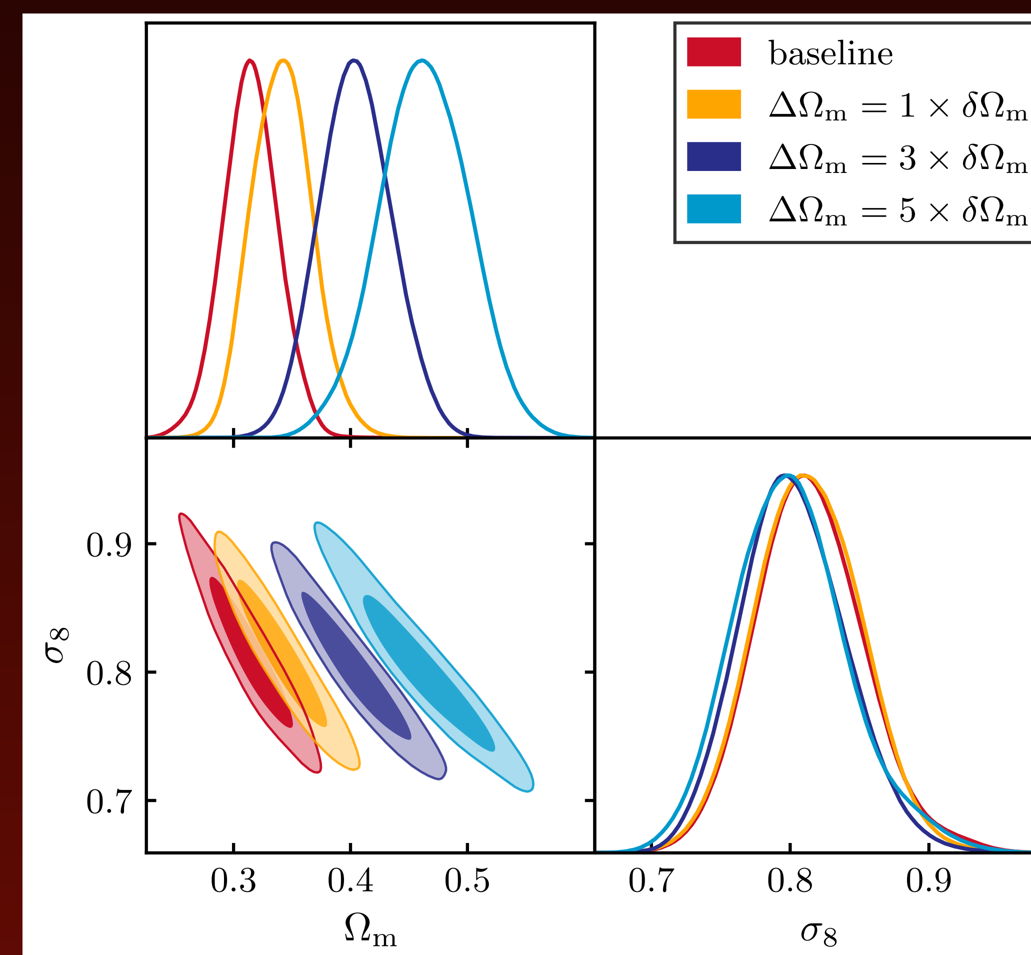
Results

Dataset	Prior	$\log R$	$\log I$	$\log S$	d	$p(\%)$
BOSS-*Planck*	default	6.30 ± 0.29	6.18 ± 0.29	0.11 ± 0.29	2.91 ± 0.51	42.66 ± 4.28
	medium	4.51 ± 0.28	4.06 ± 0.28	0.46 ± 0.28	3.30 ± 0.55	55.12 ± 4.47
	narrow	1.30 ± 0.23	0.69 ± 0.22	0.61 ± 0.22	1.67 ± 0.54	77.12 ± 14.10
DES-*Planck*	default	2.88 ± 0.35	6.15 ± 0.34	-3.28 ± 0.34	3.97 ± 0.82	3.23 ± 1.00
	medium	0.51 ± 0.34	4.00 ± 0.34	-3.49 ± 0.34	3.13 ± 0.81	2.04 ± 0.79
	narrow	-1.88 ± 0.29	0.90 ± 0.29	-2.78 ± 0.29	1.15 ± 0.77	1.44 ± 0.91
SH_0ES-*Planck*	default	-2.03 ± 0.29	1.96 ± 0.28	-3.99 ± 0.28	0.78 ± 0.52	0.25 ± 0.17
	medium	-2.50 ± 0.28	1.56 ± 0.28	-4.06 ± 0.28	1.77 ± 0.51	0.56 ± 0.24
	narrow	-2.00 ± 0.23	1.43 ± 0.23	-3.43 ± 0.23	1.92 ± 0.52	1.17 ± 0.45

Simulated DES vs Planck

We generated simulated DES data vectors, at cosmologies with a given ‘a priori’ tension with *Planck.*

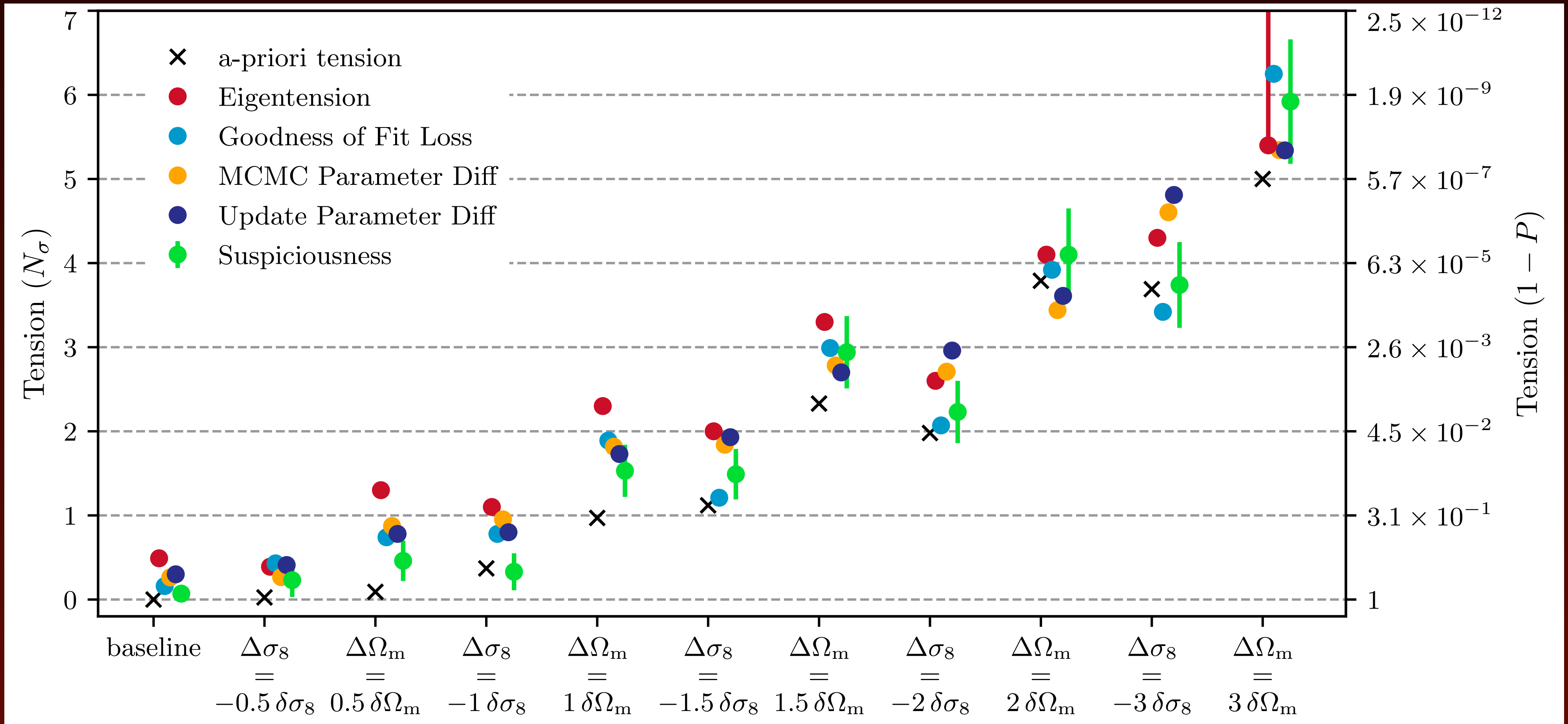




PL, Raveri & DES Collaboration;
arXiv: 2012.09554

Simulated DES vs Planck

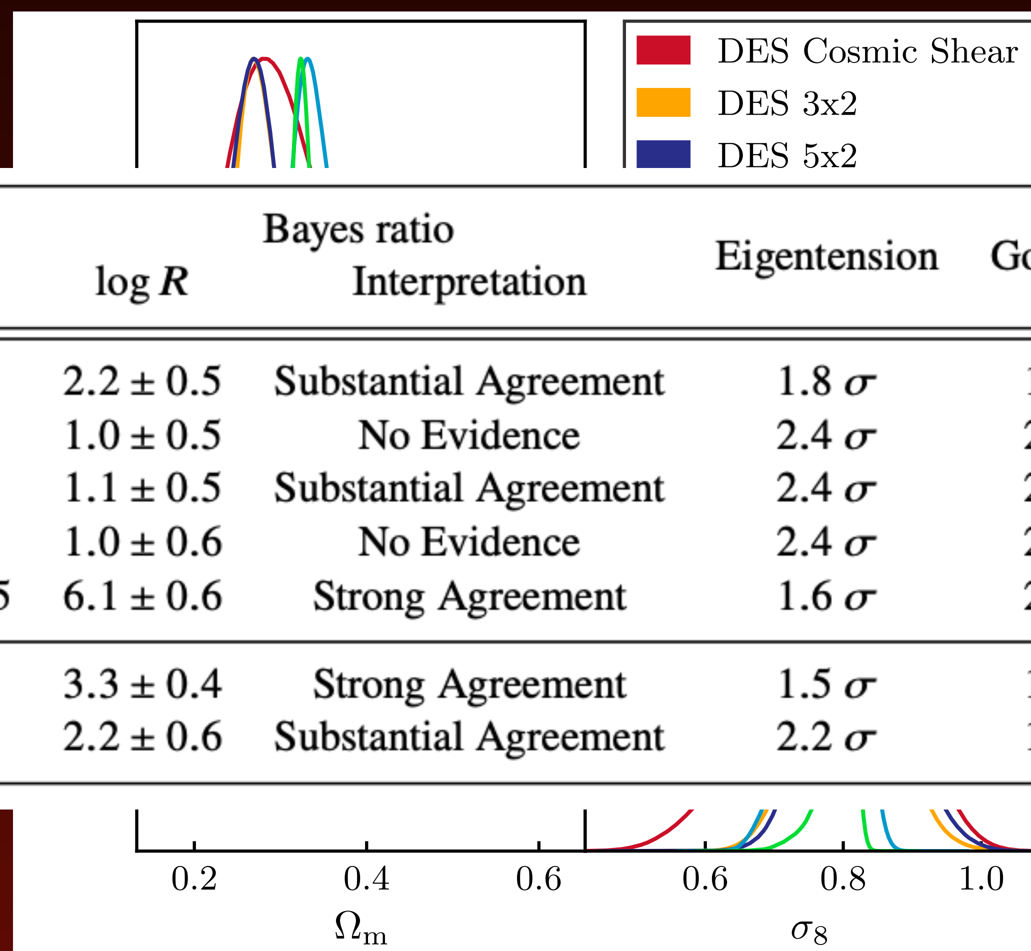
We then used the Suspiciousness, Bayes Ratio, and other statistics to quantify the tension between these simulations & *Planck.*



PL, Raveri & DES Collaboration;
arXiv: 2012.09554

DES Y1 vs Planck

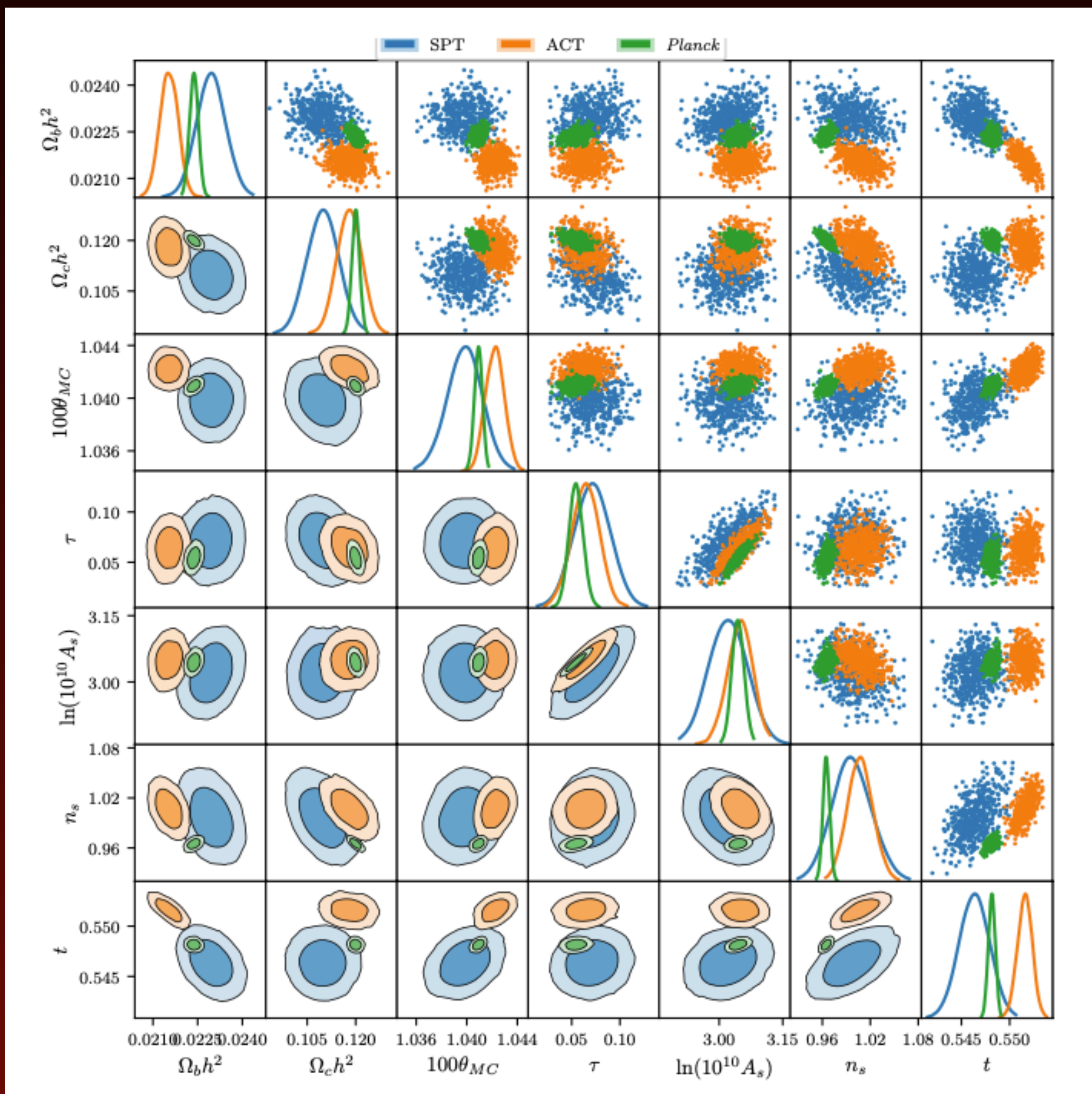
Finally, we used these metrics to recalibrate the tension between DES Y1 & *Planck*.



data set	log R	Bayes ratio Interpretation	Eigentension	GoF Loss	MCMC/Update Param Shifts	Suspiciousness
DES cosmic shear vs. *Planck* 15	2.2 ± 0.5	Substantial Agreement	$1.8\,\sigma$	$1.3\,\sigma$	$1.3/1.2\,\sigma$	$(0.7 \pm 0.4)\,\sigma$
DES 3×2pt vs. *Planck* 15	1.0 ± 0.5	No Evidence	$2.4\,\sigma$	$2.7\,\sigma$	$2.2/2.2\,\sigma$	$(2.4 \pm 0.2)\,\sigma$
DES 5×2pt vs. *Planck* 15	1.1 ± 0.5	Substantial Agreement	$2.4\,\sigma$	$2.8\,\sigma$	$2.1/2.3\,\sigma$	$(2.2 \pm 0.3)\,\sigma$
DES 5×2pt vs. *Planck* 15 + lensing	1.0 ± 0.6	No Evidence	$2.4\,\sigma$	$2.5\,\sigma$	$2.1/2.3\,\sigma$	$(2.2 \pm 0.4)\,\sigma$
DES 5×2pt + *Planck* lensing vs. *Planck* 15	6.1 ± 0.6	Strong Agreement	$1.6\,\sigma$	$2.4\,\sigma$	$1.9/2.2\,\sigma$	$(1.8 \pm 0.2)\,\sigma$
DES cosmic shear vs. *Planck* 18	3.3 ± 0.4	Strong Agreement	$1.5\,\sigma$	$1.0\,\sigma$	$1.0/1.1\,\sigma$	$(0.5 \pm 0.3)\,\sigma$
DES 3×2pt vs. *Planck* 18	2.2 ± 0.6	Substantial Agreement	$2.2\,\sigma$	$1.6\,\sigma$	$2.0/2.3\,\sigma$	$(2.4 \pm 0.2)\,\sigma$

PL, Raveri & DES Collaboration;
arXiv: 2012.09554

Application to the CMB



Dataset combination	χ^2	p	tension	$\log S$
ACT vs *Planck*	17.2	0.86%	2.63σ	-5.60
ACT vs SPT	15.4	1.77%	2.37σ	-4.68
Planck vs SPT	9.1	16.82%	1.38σ	-1.55
ACT vs *Planck*+SPT	18.4	0.52%	2.79σ	-6.22
ACT+SPT vs *Planck*	12.2	5.81%	1.90σ	-3.09
ACT+*Planck* vs SPT	10.3	11.09%	1.59σ	-2.17

Handley & PL;
arXiv: 2007.08496

Application to KiDS

KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints

Catherine Heymans[1,2⋆], Tilman Tröster[1⋆⋆], Marika Asgari[1], Chris Blake[3], Hendrik Hildebrandt[2], Benjamin Joachimi[4], Konrad Kuijken[5], Chieh-An Lin[1], Ariel G. Sánchez[6], Jan Luca van den Busch[2], Angus H. Wright[2], Alexandra Amon[7], Maciej Bilicki[8], Jelte de Jong[9], Martin Crocce[10,11], Andrej Dvornik[2], Thomas Erben[12], Maria Cristina Fortuna[5], Fedor Getman[13], Benjamin Giblin[1], Karl Glazebrook[3], Henk Hoekstra[5], Shahab Joudaki[14], Arun Kannawadi[15,5], Fabian Köhlinger[2], Chris Lidman[16], Lance Miller[14], Nicola R. Napolitano[17], David Parkinson[18], Peter Schneider[12], HuanYuan Shan[19,20], Edwin A. Valentijn[9], Gijs Verdoes Kleijn[9], and Christian Wolf[16]

Handley & Lemos (2019) propose the 'suspiciousness' statistic S that is based on the Bayes factor, R, but hardened against prior dependences. We find that the probability of observing our measured suspiciousness statistic is 0.08 ± 0.02, which corresponds to a KiDS-*Planck* tension at the level of $1.8 \pm 0.1\,\sigma$ (see Appendix G.3 for details).

The second equality follows from Bayes theorem: $P = \mathcal{L}\pi / Z$. Using this definition of $\mathcal{D}$ allows us to rephrase the suspiciousness solely in terms of the expectation values of the log-likelihoods:

$$\ln S = \langle \ln \mathcal{L}_{3\times2\mathrm{pt}+Planck} \rangle_{P_{3\times2\mathrm{pt}+Planck}} - \langle \ln \mathcal{L}_{3\times2\mathrm{pt}} \rangle_{P_{3\times2\mathrm{pt}}} - \langle \ln \mathcal{L}_{Planck} \rangle_{P_{Planck}} \,. \quad \text{(G.9)}$$

Thanks for listening!

- Cosmological ‘**Tensions**’ could be a hint of new physics, and must therefore be understood.
- Quantifying tension is therefore crucial. We propose the ‘**Suspiciousness**’ as the optimal metric of tension in Cosmology.
- The method can be extended to any other problem of assessing consistency between data sets, in astrophysics or otherwise.

p.lemos@sussex.ac.uk