

CMB + BAO + SN

Bayesian Model Comparison and Dataset Consistency

Lukas T. Hergt

Laboratoire de Physique des 2 infinis — Irène Joliot-Curie (IJCLab)

IN2P3, CNRS

<lukas.hergt@ijclab.in2p3.fr>

Synergistic Power of Combined Cosmological Observables



Objective: Put various tensions into perspective

Models

- Baseline: Λ CDM
 - Cold Dark Matter with cosmological constant Λ
 - flat ($\Omega_K = 0$)
 - 1 massive neutrino $m_\nu = 0.06$ eV
- Neutrinos
 Λ CDM + $\sum m_\nu$
- Curvature
 Λ CDM + Ω_K
- Constant Dark Energy
 w CDM
- Evolving Dark Energy
 $w_0 w_a$ CDM

Data/Likelihoods

- CMB and CMB lensing

- **Planck** PR3 & PR4:

low- ℓ TT	low- ℓ EE	high- ℓ $TTTEEE$
Commander	SimAll (PR3)	Plik (PR3)
Commander	SimAll (PR3)	CamSpec (PR4)
Commander	Lollipop (PR4)	Hillipop (PR4)

- Baryon Acoustic Oscillations (BAO)
 - **SDSS DR12**
 - **SDSS DR16**
 - **DESI DR2**
- Supernovae (SN)
 - **Pantheon⁺**
 - **DES y5**

The 3 Pillars of Bayesian Inference

Parameter Estimation

What do the data tell us about the parameters of a model?
(e.g. the size or age of a Λ CDM universe)

Bayes' Theorem:

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

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

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

$$\propto \mathcal{L}(\theta) \times \Pi(\theta)$$

Model Comparison

How much does the data support one model compared to another?

(e.g. cosmology with a cosmological constant vs with a dynamic dark energy)

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(D|M_1) \times P(M_1)}{P(D|M_2) \times P(M_2)}$$

$$= \frac{\mathcal{Z}_{M_1}}{\mathcal{Z}_{M_2}} \times \frac{P(M_1)}{P(M_2)}$$

$$\text{posterior odds} = \text{evidence ratio} \times \text{prior odds}$$

Occam's razor:

$$\log \mathcal{Z} = \langle \log \mathcal{L} \rangle_{\mathcal{P}} - \mathcal{D}_{\text{KL}}$$

$$\text{log-evidence} = \text{fit} - \text{Occam penalty}$$

Dataset Consistency

Do different datasets A and B make consistent predictions for the same model?

(e.g. CMB vs BAO data)

H_0 : datasets described by the same parameters

H_1 : datasets described by distinct parameters

$$\mathcal{R} = \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \mathcal{Z}_B},$$

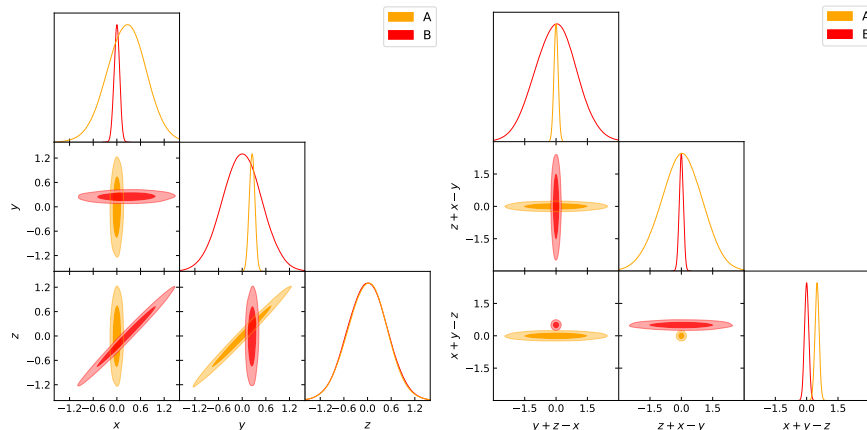
$$\mathcal{R} = \log \mathcal{S} - \mathcal{I},$$

$$\log \mathcal{S} = \langle \log \mathcal{L}_{AB} \rangle_{\mathcal{P}_{AB}}$$

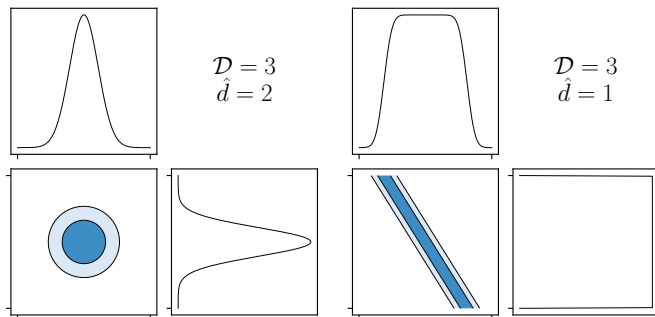
$$- \langle \log \mathcal{L}_A \rangle_{\mathcal{P}_A}$$

$$- \langle \log \mathcal{L}_B \rangle_{\mathcal{P}_B}$$

$$d - 2 \log \mathcal{S} \sim \chi_d^2$$

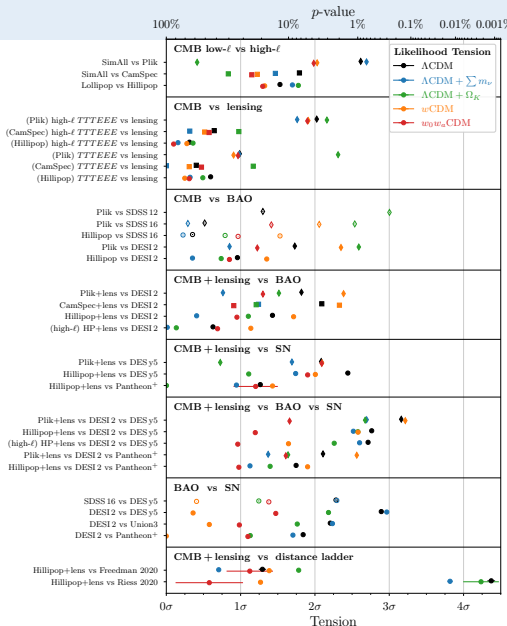
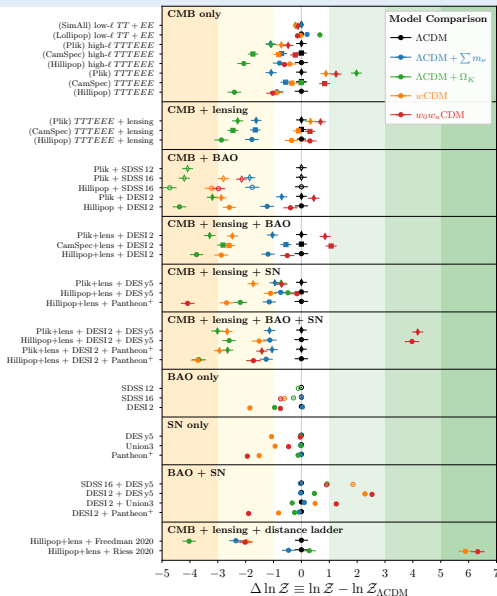


- Projections can be misleading. → Can miss tensions on marginalised parameters.
- Evidence ratio $\mathcal{R}$ and likelihood ratio $\mathcal{S}$ take **full dimensionality** into account, (estimated by the Gaussian Model Dimensionality $\hat{d}$).

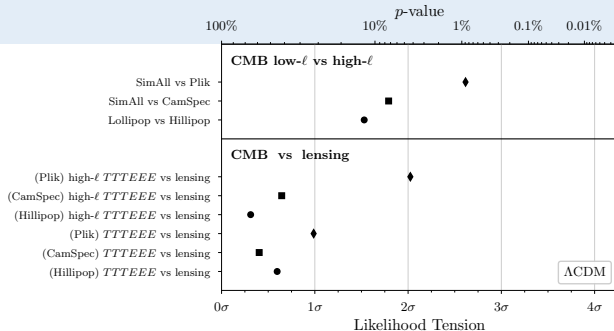


- **Same** number of sampling parameters.
- **Same** Kullback–Leibler divergence $\mathcal{D} = \left\langle \ln \frac{\mathcal{P}}{\Pi} \right\rangle_{\mathcal{P}}$ (i.e. same prior to posterior compression).
- **Different** dimensionality $\frac{\hat{d}}{2} = \langle (\ln \mathcal{L})^2 \rangle - \langle \ln \mathcal{L} \rangle^2$ (i.e. different number of *constrained* parameters).

Slightly Overwhelming Overview

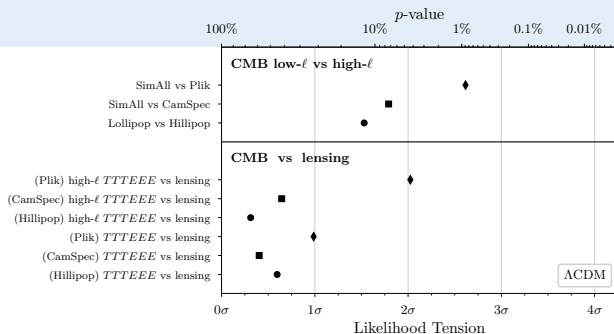


CMB + lensing



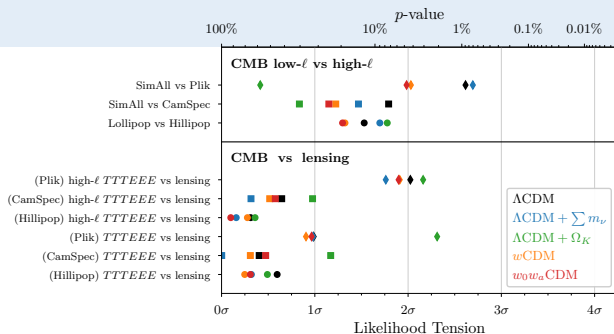
CMB: internal consistency

- Tensions reduce from
Plik $\rightarrow$ CamSpec $\rightarrow$ Hillipop.
 - low- ℓ vs high- ℓ
 - CMB vs lensing



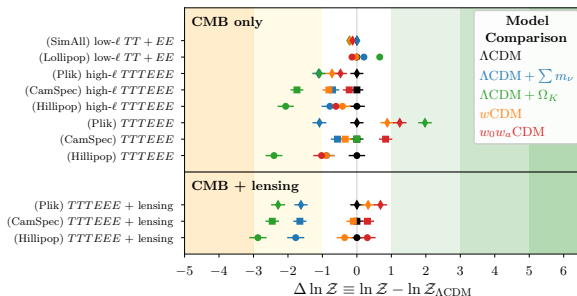
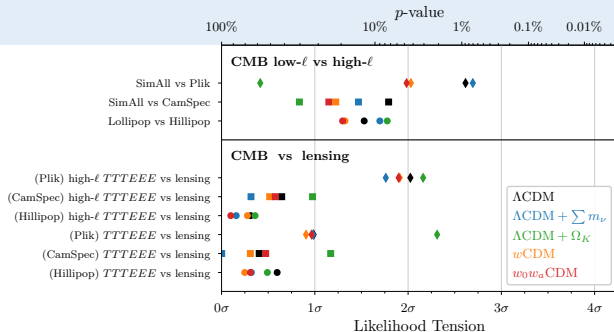
CMB: internal consistency

- Tensions reduce from
Plik $\rightarrow$ CamSpec $\rightarrow$ Hillipop.
 - low- ℓ vs high- ℓ
 - CMB vs lensing



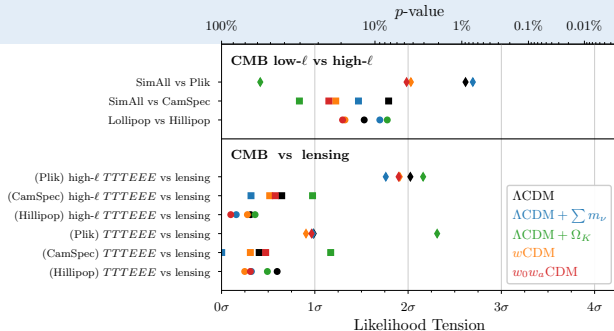
CMB: internal consistency

- Tensions reduce from
Plik $\rightarrow$ CamSpec $\rightarrow$ Hillipop.
 - low- ℓ vs high- ℓ
 - CMB vs lensing
- No strong preference for any Λ CDM extension.

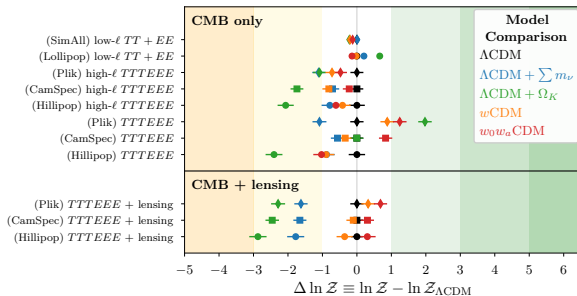


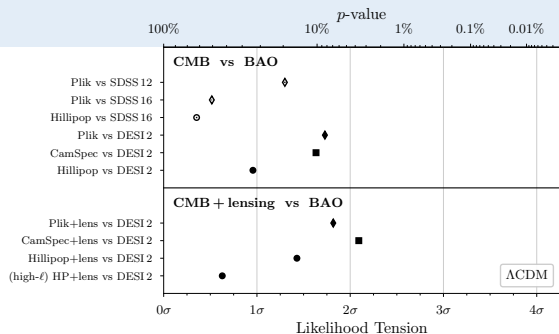
CMB: internal consistency

- Tensions reduce from
Plik $\rightarrow$ CamSpec $\rightarrow$ Hillipop.
 - low- ℓ vs high- ℓ
 - CMB vs lensing
- No strong preference for any Λ CDM extension.



- Curvature Tension** with Plik:
 - Slight tension between CMB and lensing for Λ CDM + Ω_K .
 - Slight preference for Λ CDM + Ω_K (curvature).
 - Related to A_{lens} anomaly.



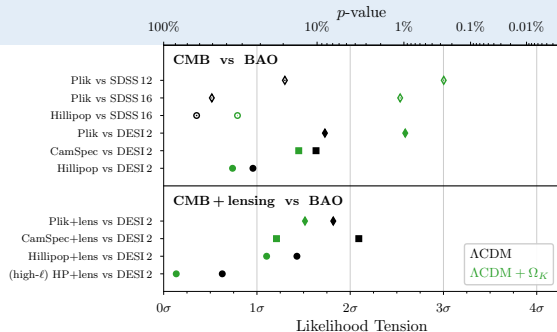


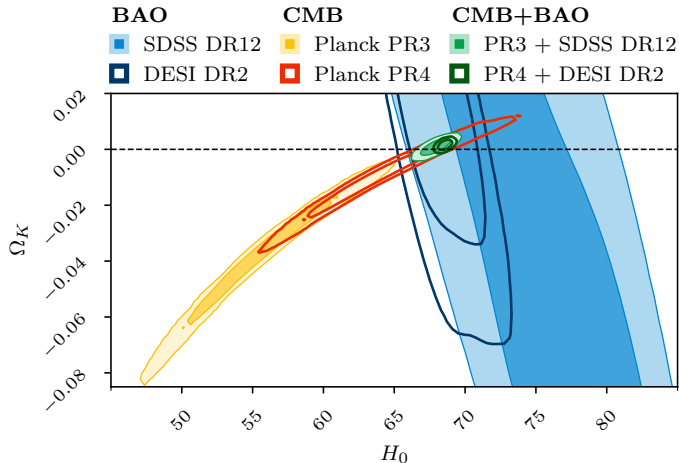
CMB + BAO

• Curvature Tension with Plik

(Λ CDM + Ω_K):

- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillipop.





• Curvature Tension with Plik

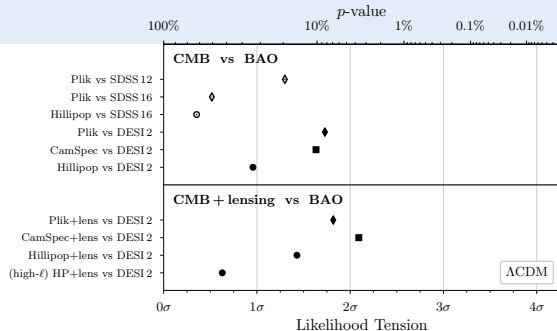
(Λ CDM + Ω_K):

- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillipop.

• Tensions reduce from Plik $\rightarrow$ Hillipop.

• BAO update from DESI DR2:

- Hints of tension with CamSpec + lensing.
- Goes away with Hillipop + lensing.



• Curvature Tension with Plik

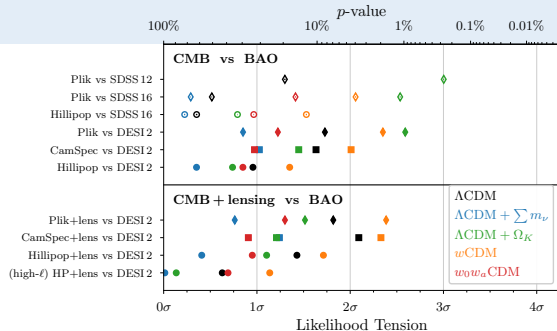
(Λ CDM + Ω_K):

- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillipop.

• Tensions reduce from Plik $\rightarrow$ Hillipop.

• BAO update from DESI DR2:

- Hints of tension with CamSpec + lensing.
- Goes away with Hillipop + lensing.



Curvature Tension with Plik

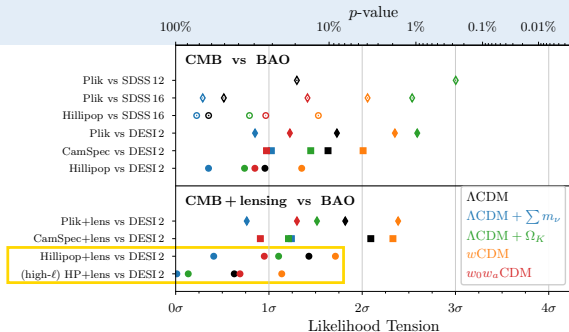
(Λ CDM + Ω_K):

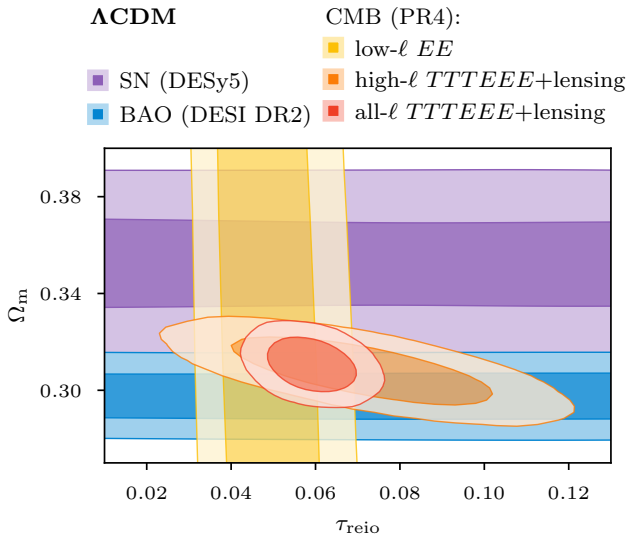
- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillipop.

Tensions reduce from Plik $\rightarrow$ Hillipop.

BAO update from DESI DR2:

- Hints of tension with CamSpec + lensing.
- Goes away with Hillipop + lensing.
- Removing low- ℓ further reduces tension.
 $\Rightarrow$ “ τ -Tension”





• Curvature Tension with Plik

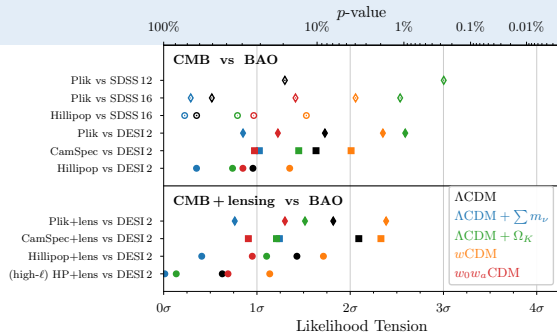
(Λ CDM + Ω_K):

- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillpop.

• Tensions reduce from Plik $\rightarrow$ Hillpop.

• BAO update from DESI DR2:

- Hints of tension with CamSpec + lensing.
- Goes away with Hillpop + lensing.
- Removing low- ℓ further reduces tension.
 $\Rightarrow$ “ τ -Tension”



Curvature Tension with Plik

(Λ CDM + Ω_K):

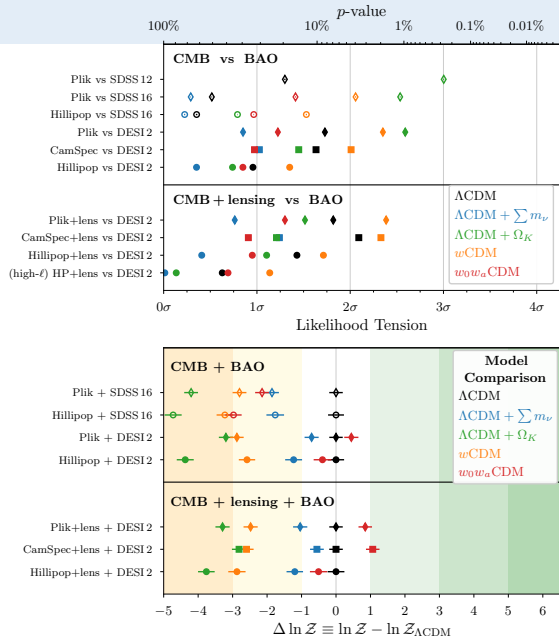
- 2.3σ tension between Plik and lensing.
- 3σ tension between Plik and SDSS 12.
- Reduces to 2.5σ with SDSS 16 or DESI 2.
- Goes away when switching to CamSpec or Hillipop.

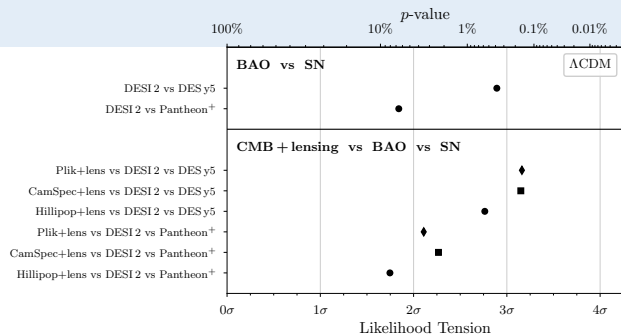
Tensions reduce from Plik $\rightarrow$ Hillipop.

BAO update from DESI DR2:

- Hints of tension with CamSpec + lensing.
- Goes away with Hillipop + lensing.
- Removing low- ℓ further reduces tension.
 $\Rightarrow$ “ τ -Tension”

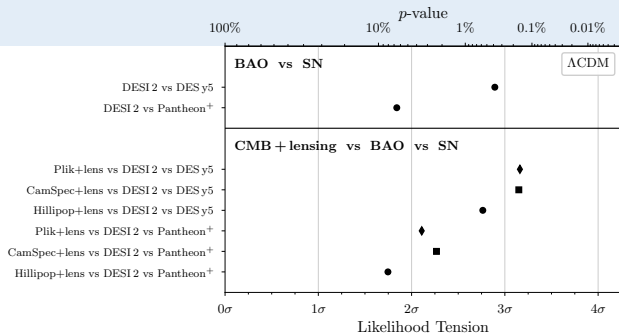
No preference for extensions.

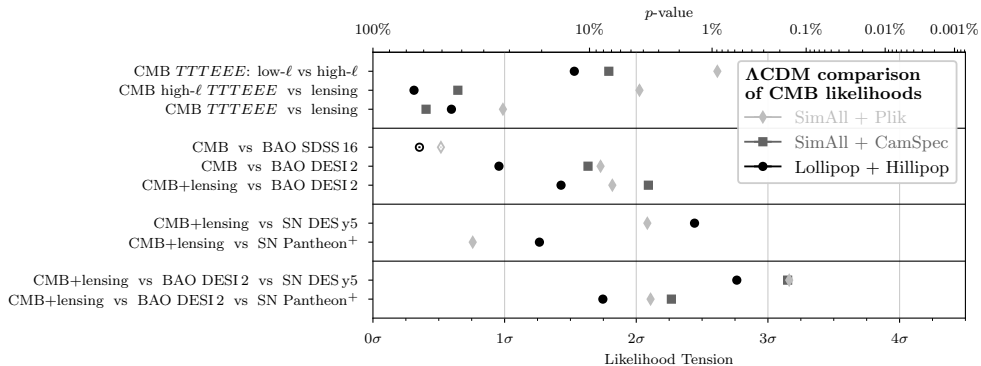


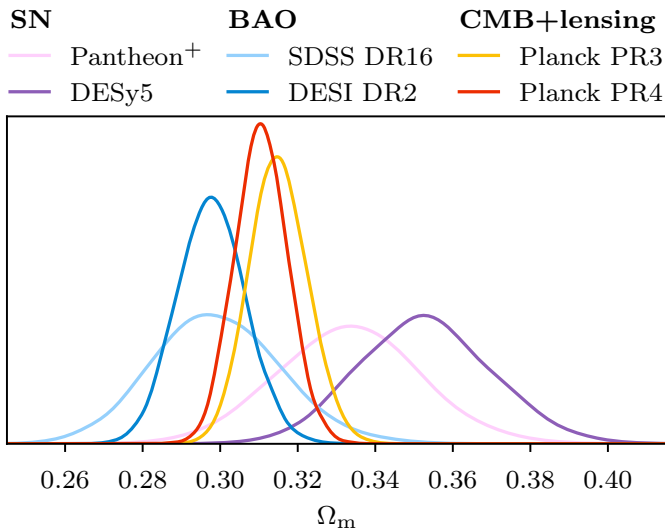


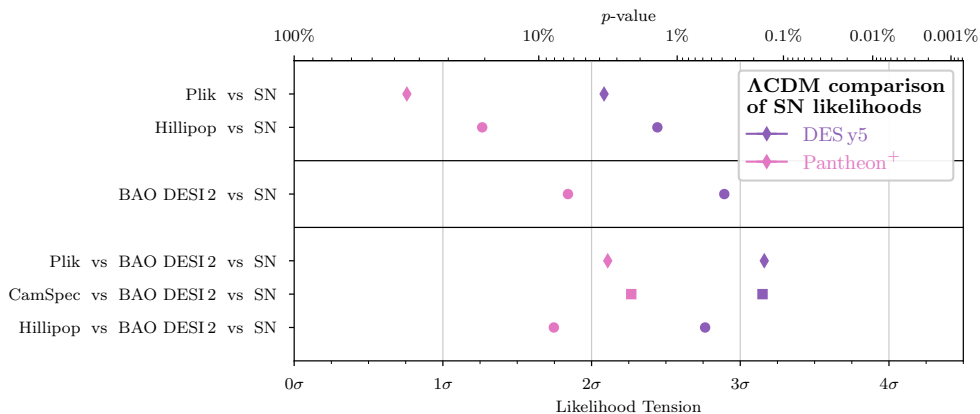
CMB + BAO + SN

- $\sim 3\sigma$ tension with DES y5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DES y5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).

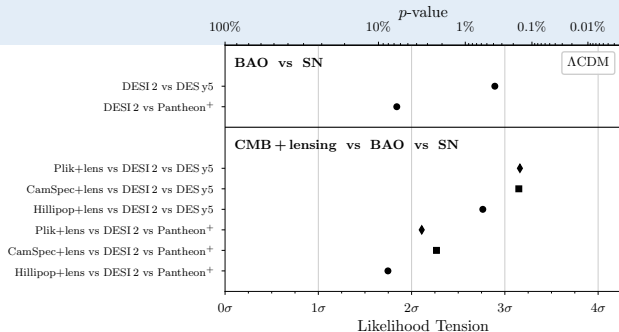




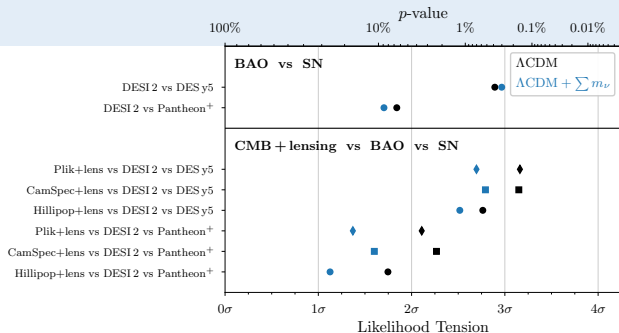


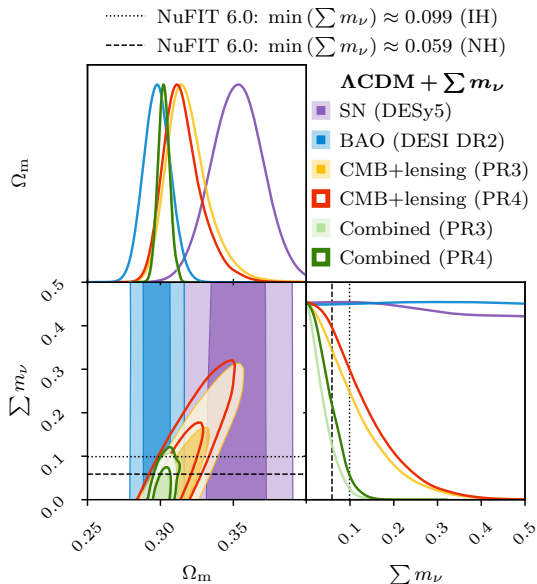


- $\sim 3\sigma$ tension with DES y5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DES y5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).

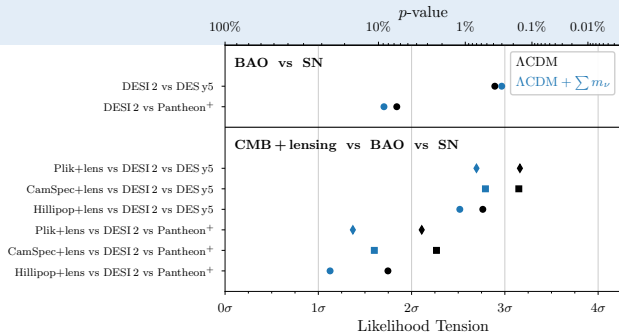


- $\sim 3\sigma$ tension with DESy5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DESy5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).
- Tensions reduce from Λ CDM $\rightarrow \Lambda$ CDM + $\sum m_\nu$ (~ 0.25 to 0.5σ).

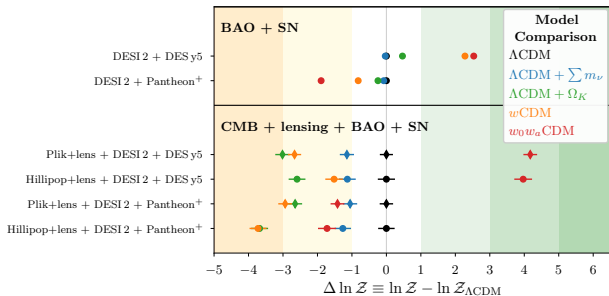
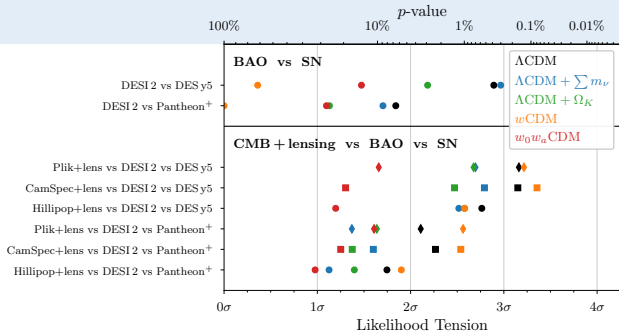


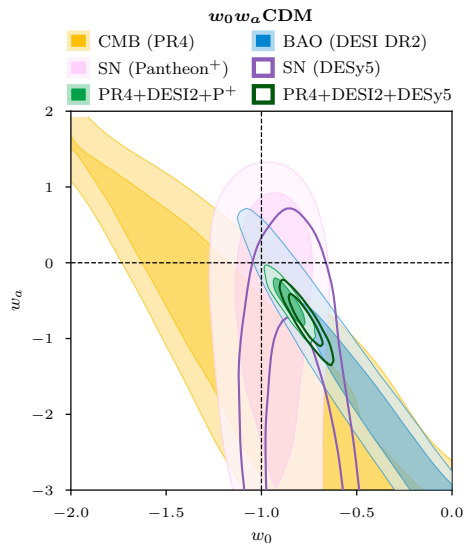


- $\sim 3\sigma$ tension with DESy5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DESy5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).
- Tensions reduce from Λ CDM $\rightarrow \Lambda$ CDM + $\sum m_\nu$ (~ 0.25 to 0.5σ).

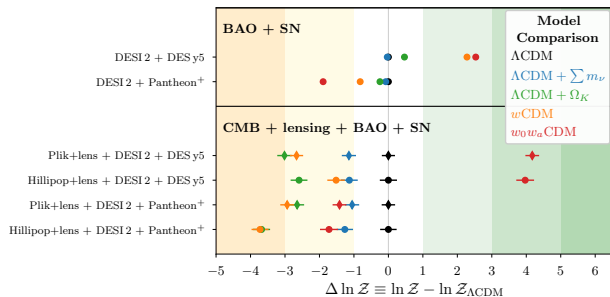
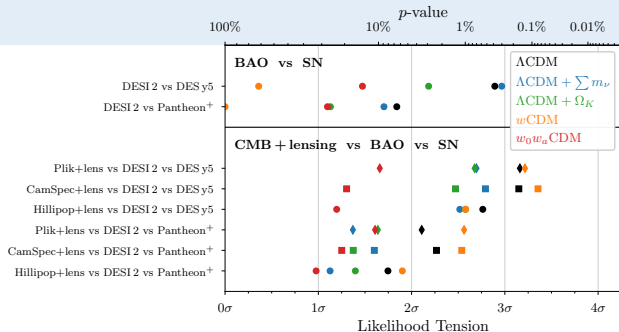


- $\sim 3\sigma$ tension with DES y5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DES y5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).
- Tensions reduce from Λ CDM $\rightarrow \Lambda$ CDM + $\sum m_\nu$ (~ 0.25 to 0.5σ).
- Tensions reduce from Λ CDM $\rightarrow w$ CDM or w_0w_a CDM (~ 1 to 2σ).
- Preference for w CDM and w_0w_a CDM:
 - Preference only in combination with DES y5.
 - No preference with Pantheon⁺.





- $\sim 3\sigma$ tension with DES y5 data.
- Tensions reduce across models:
 - from Plik $\rightarrow$ Hillipop ($\sim 0.5\sigma$).
 - from DES y5 $\rightarrow$ Pantheon⁺ ($\sim 1\sigma$).
- Tensions reduce from Λ CDM $\rightarrow \Lambda$ CDM + $\sum m_\nu$ (~ 0.25 to 0.5σ).
- Tensions reduce from Λ CDM $\rightarrow w$ CDM or $w_0 w_a$ CDM (~ 1 to 2σ).
- Preference for w CDM and $w_0 w_a$ CDM:
 - Preference only in combination with DES y5.
 - No preference with Pantheon⁺.



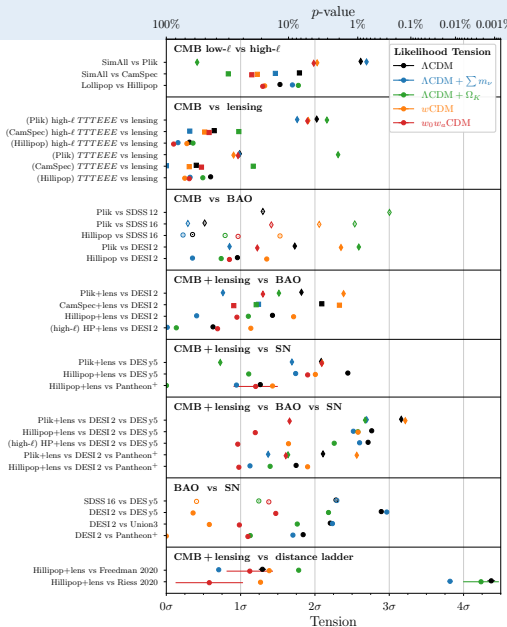
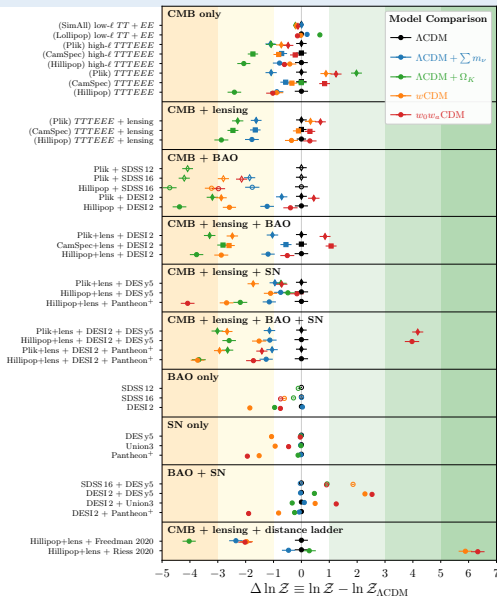
Conclusion

- Tensions vary with choice of CMB likelihood:
 - 0.5σ to 1.5σ .
 - Trend: tensions reduce from Plik $\rightarrow$ CamSpec $\rightarrow$ Hillipop.
 - Note: tensions might tighten for ACT & SPT.
- Tensions and Model preference vary with choice of SN likelihood:
 - $\sim -1\sigma$ going from DESy5 $\rightarrow$ Pantheon⁺.
 - Model comparison switches from “strongly preferred” to “disfavoured”.
- Λ CDM still going strong for majority of data combinations.
- Curvature tension?
 - Model preference for Λ CDM + Ω_K .
 - Tension between Plik and lensing, or Plik and BAO for Λ CDM + Ω_K .
 - Both model preference and tension go away with CamSpec and Hillipop.
- τ tension?
 - Exclusion of low- ℓ EE reduces tension between CMB and BAO.
 - Irrelevant when combining CMB, BAO and SN.
- Preference for $w_0 w_a$ CDM?
 - Only specifically for Plik + DESI + DESy5.

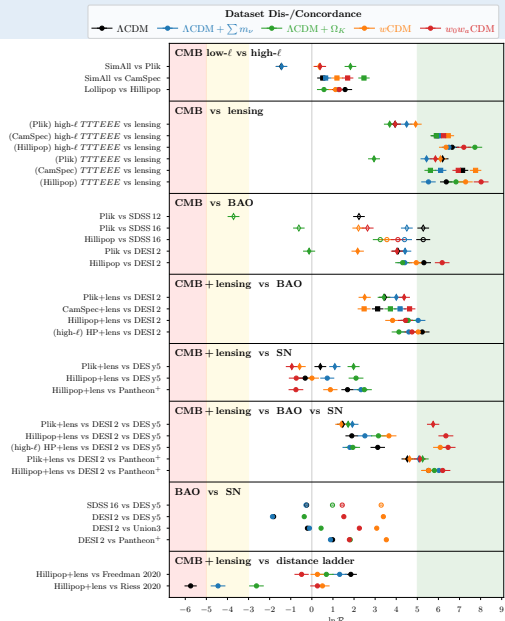
Should we drop Λ CDM and make $w_0 w_a$ CDM our new Cosmic Standard Model?

- “Strong claims need strong evidence.”
- Singling out specific dataset/likelihood combinations not enough.
- Evidence needs to be sufficiently strong independent of dataset/likelihood choice.

Slightly Overwhelming Overview



Slightly Overwhelming Overview



The 3 Pillars of Bayesian Inference

Parameter Estimation

What do the data tell us about the parameters of a model?
(e.g. the size or age of a Λ CDM universe)

Bayes' Theorem:

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

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

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

$$\propto \mathcal{L}(\theta) \times \Pi(\theta)$$

Model Comparison

How much does the data support one model compared to another?

(e.g. cosmology with a cosmological constant vs with a dynamic dark energy)

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(D|M_1) \times P(M_1)}{P(D|M_2) \times P(M_2)}$$

$$= \frac{\mathcal{Z}_{M_1}}{\mathcal{Z}_{M_2}} \times \frac{P(M_1)}{P(M_2)}$$

$$\text{posterior odds} = \text{evidence ratio} \times \text{prior odds}$$

Occam's razor:

$$\log \mathcal{Z} = \langle \log \mathcal{L} \rangle_{\mathcal{P}} - \mathcal{D}_{\text{KL}}$$

$$\text{log-evidence} = \text{fit} - \text{Occam penalty}$$

Dataset Consistency

Do different datasets A and B make consistent predictions for the same model?

(e.g. CMB vs BAO data)

H_0 : datasets described by the same parameters

H_1 : datasets described by distinct parameters

$$\mathcal{R} = \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \mathcal{Z}_B},$$

$$\mathcal{R} = \log \mathcal{S} - \mathcal{I},$$

$$\log \mathcal{S} = \langle \log \mathcal{L}_{AB} \rangle_{\mathcal{P}_{AB}}$$

$$- \langle \log \mathcal{L}_A \rangle_{\mathcal{P}_A}$$

$$- \langle \log \mathcal{L}_B \rangle_{\mathcal{P}_B}$$

$$d - 2 \log \mathcal{S} \sim \chi_d^2$$

- Bayesian inference quantifies Occam's Razor:
 - *"Frustra fit per plura quod potest fieri per pauciora"* — William of Occam
(*"It is futile to do with more things that which can be done with fewer"*)
 - *"Accept the simplest explanation that fits the data."* — David MacKay
- Consider the Evidence $\mathcal{Z} \equiv P(D|M)$:
 - Parameter estimation $\rightarrow$ normalisation constant
 - Model comparison $\rightarrow$ critical update factor from model prior to model posterior
- Properties of the evidence: rearrange Bayes' theorem for parameter estimation

$$\mathcal{P}(\theta) = \frac{\mathcal{L}(\theta)\Pi(\theta)}{\mathcal{Z}} \quad \Rightarrow \quad \log \mathcal{Z} = \log \mathcal{L}(\theta) - \log \frac{\mathcal{P}(\theta)}{\Pi(\theta)}$$

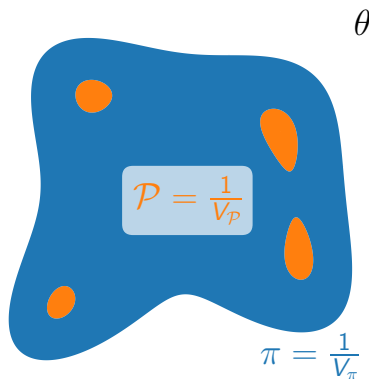
- Evidence is composed of a "goodness of fit" term and "Occam Penalty"
- RHS true for all θ . Take max likelihood point θ_{ML} :
 - Be more Bayesian and take posterior average to get the "Occam equation"

$$\log \mathcal{Z} = \log \mathcal{L}_{\max} - \log \frac{\mathcal{P}(\theta_{ML})}{\Pi(\theta_{ML})}$$

$$\log \mathcal{Z} = \langle \log \mathcal{L} \rangle_{\mathcal{P}} - \mathcal{D}_{KL}$$

- Natural regularisation which penalises models with too many parameters.

Prior and posterior volumes



- The KL divergence between **prior** Π and **posterior** $\mathcal{P}$ is defined as:

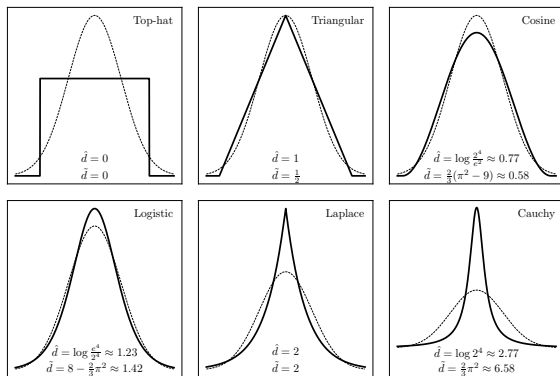
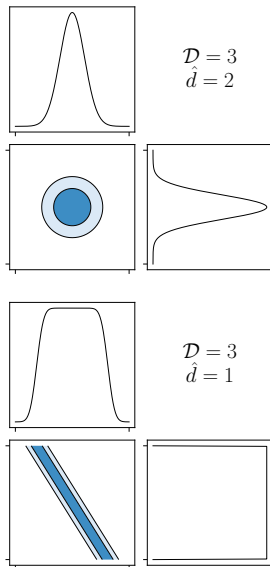
$$\mathcal{D}_{\text{KL}} = \left\langle \log \frac{\mathcal{P}}{\Pi} \right\rangle_{\mathcal{P}} = \int \mathcal{P}(\theta) \log \frac{\mathcal{P}(\theta)}{\Pi(\theta)} d\theta.$$

- Whilst not a distance, $\mathcal{D} = 0$ when $\mathcal{P} = \Pi$.
- Occurs in the context of machine learning as an objective function for training functions.
- In Bayesian inference it can be understood as a log-ratio of “volumes”:

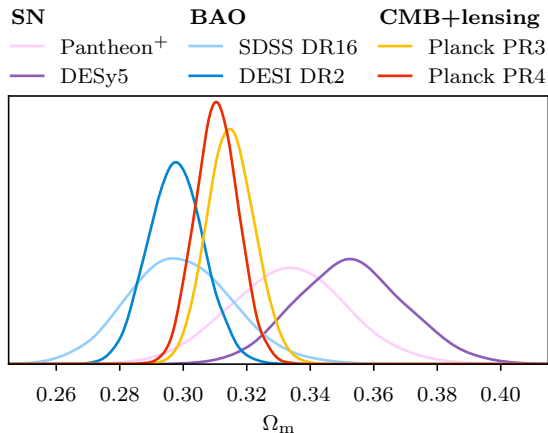
$$\mathcal{D}_{\text{KL}} \approx \log \frac{V_{\Pi}}{V_{\mathcal{P}}}.$$

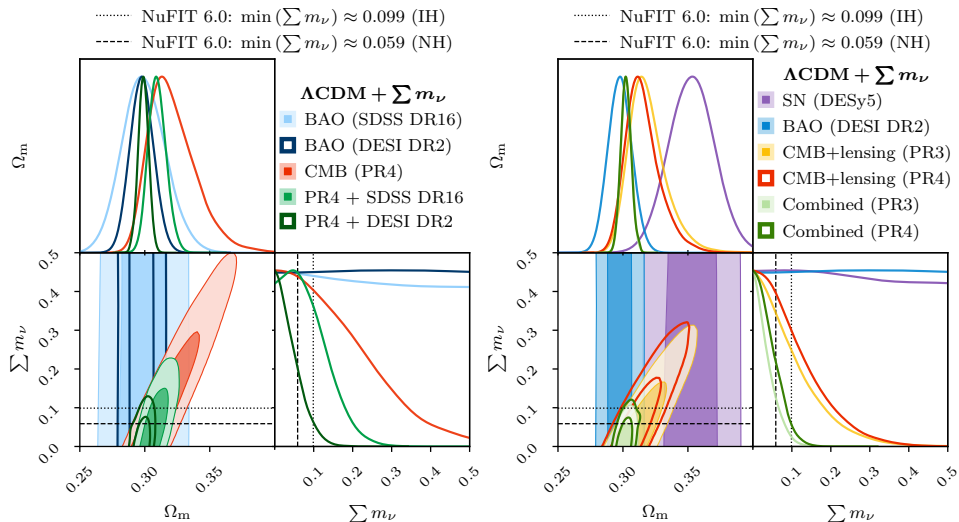
(this is exact for top-hat distributions).

Bayesian Model Dimensionality



Handley & Lemos (2019)





Constant and Evolving Dark Energy

