`unimpeded`: A Public Grid of Nested Sampling Chains for Cosmological Model Comparison and Tension Analysis

Dily Duan Yi Ong[1,2] **and Will Handley**[1,2]

[1]Kavli Institute for Cosmology, University of Cambridge,
Madingley Road, Cambridge, CB3 0HA, U.K.
[2]Cavendish Laboratory, University of Cambridge,
J.J. Thomson Avenue, Cambridge, CB3 0HE, U.K.

E-mail: dlo26@cam.ac.uk

Abstract. Bayesian inference is central to modern cosmology, yet comprehensive model comparison and tension quantification remain computationally prohibitive for many researchers. To address this, we release `unimpeded`, a publicly available Python library and data repository providing pre-computed nested sampling and MCMC chains. We apply this resource to conduct a systematic analysis across a grid of eight cosmological models, including ΛCDM and seven extensions, and 39 datasets, including individual probes and their pairwise combinations. Our model comparison reveals that whilst individual datasets show varied preferences for model extensions, the base ΛCDM model is most frequently preferred in combined analyses, with the general trend suggesting that evidence for new physics is diluted when probes are combined. Using five complementary statistics, we quantify tensions, finding the most significant to be between DES and Planck ($\sigma = 3.57$) and SH0ES and Planck ($\sigma = 3.27$) within ΛCDM. We characterise the S_8 tension as high-dimensional ($d_G = 6.62$) and resolvable in extended models, whereas the Hubble tension is low-dimensional and persists across the model space. Caution should be exercised when combining datasets in tension. The `unimpeded` data products, hosted on Zenodo, provide a powerful resource for reproducible cosmological analysis and underscore the robustness of the ΛCDM model against the current compendium of data.

ArXiv ePrint: [INSERT ARXIV NUMBER IF AVAILABLE]

---

[1]Corresponding author.

Contents

1 Introduction

Bayesian methods of inference are widely used in modern cosmology in parameter estimation, model comparison and tension quantification. Parameter estimation refers to the process of determining the values of cosmological parameters, which describe the properties of a model using observed data. Model comparison refers to the evaluation and selection between different cosmological models and tension quantification is the measurement and study of discrepancies between different observed datasets, which are predicted to be in agreement theoretically by a cosmological model. The last two of these have gained more prominence in recent times due to disparities that have emerged within the context of the concordance model regarding the estimated value of the Hubble constant H_0 [1] using Cosmic Microwave Background (CMB) and Supernovae data (commonly referred to as the Hubble tension), the clustering σ_8 [2] using CMB and weak lensing, and the curvature Ω_K [3, 4] using CMB and lensing/BAO, and between CMB datasets. These tensions, may excitingly, point towards physics beyond the standard Λ Cold Dark Matter (ΛCDM) concordance model.

     Parameter estimation has commonly been performed using the Markov chain Monte Carlo (MCMC) methods, which are effective for exploring the posterior distributions of model parameters given a set of data and a model. The Planck Legacy Archive (PLA) [5] has been an invaluable community resource, providing MCMC chains for a grid of models and datasets, primarily facilitating parameter estimation. However, MCMC methods are not suitable for calculating Bayesian evidence, which is essential for model comparison and tension quantification. Nested sampling [6–9] has emerged as a powerful alternative, specifically tailored for model comparison and tension quantification. It is a Monte Carlo sampling technique used to efficiently compute the evidence and concurrently generate samples from the posterior distribution as a by-product, and hence, enabling parameter estimation without extra expense. However, the computational cost of nested sampling is still significant, especially

when considering the large parameter spaces and complex likelihoods involved in modern cosmological analyses.

This paper introduces `unimpeded`[1], a publicly available pip-installable Python library and associated data repository. The primary aim of `unimpeded` is to provide an analogous grid to the PLA but utilising nested sampling chains, thereby enabling robust model comparison and tension quantification alongside parameter estimation. This initiative directly supports the goals of our DiRAC-funded projects (DP192 and 264), which seek a systematic examination of model-dataset combinations to uncover patterns that might illuminate the path towards resolving current cosmological puzzles or identifying a successor to ΛCDM. The `unimpeded` grid is designed to incorporate a broad variety of modern datasets and expanding as new data and models become relevant. All associated data products, including nested sampling and MCMC chains, are made publicly and permanently available on Zenodo[2].

This paper is structured as follows. In Section 2, we review the theoretical foundations of Bayesian inference and the three pillars of Bayesian cosmological analysis, namely parameter estimation, model comparison and tension quantification. We also discuss the concordance ΛCDM model and other models. We then detail our specific methodological approach with nested sampling in Section 3. The `unimpeded` library and its core functionalities are introduced in Section 4. We apply this framework to cosmological data, presenting our main findings in Section 5, presenting wide grids of model comparison and tension statistics.

2 Theory

2.1 Nomenclature

The Universe provides a single laboratory for physics, but with experimental settings we can only observe, not control. Our goal is to utilises observables to rigorously test predictive cosmological models, seeking either to quantify how likely they are to be true given the data, and to improve their parameter constraints. Bayesian inference provides a framework for systematically updating our beliefs about model and their parameters in light of new data.

A predictive model $\mathcal{M}$ contains a set of variable parameters θ, with some observed dataset D. D is typically a collection of measurements or observations, such as the Cosmic Microwave Background (CMB), baryon acoustic oscillations (BAO), supernovae data, weak lensing data and gravitational waves. Cosmological models $\mathcal{M}$ are theoretical frameworks that describe the physical properties and evolution of the Universe, typically expressed by a metric with a set of cosmological parameters $\theta_{\mathcal{M}}$. $\theta_{\mathcal{M}}$ has the subscript $\mathcal{M}$ to indicate that the parameters are specific to the model $\mathcal{M}$. For example, the concordance ΛCDM model has 6 parameters: the Hubble constant H_0, the baryon density $\Omega_b h^2$, the cold dark matter density $\Omega_c h^2$, the scalar spectral index n_s, the amplitude of primordial scalar perturbations A_s and the reionisation optical depth τ_{reio}. When only one model is considered, we can drop the subscript $\mathcal{M}$ and write θ instead of $\theta_{\mathcal{M}}$.

2.1.1 Bayesian Inference

Before analysing any data, we can express our beliefs about the parameters θ for a specific model $\mathcal{M}$, termed the prior,

$$P(\theta|\mathcal{M}) \equiv \pi(\theta). \tag{2.1}$$

---

[1]The `unimpeded` library and its source code are available at https://github.com/handley-lab/unimpeded.
[2]https://zenodo.org/

Common choices for the prior include uniform or log-uniform distributions over a range of theoretically and physically allowed values or Gaussian distributions centred around expected values. After considering the observed data D, we can update our beliefs about the parameters θ, termed the posterior,

$$\mathrm{P}(\theta|D,\mathcal{M}) \equiv \mathcal{P}(\theta). \tag{2.2}$$

Both the prior and posterior are probability density functions (PDFs), which integrate to 1 over all possible values of θ.

The likelihood function describes how probable the observed data D is given a specific set of cosmological parameter values θ and a specific model $\mathcal{M}$,

$$\mathrm{P}(D|\theta,\mathcal{M}) \equiv \mathcal{L}(\theta). \tag{2.3}$$

While $\mathrm{P}(D|\theta,\mathcal{M})$ treats D as the variable and θ and $\mathcal{M}$ as fixed parameters, $\mathcal{L}(\theta)$ treats θ as the variable. This equivalence represents a shift from prediction to inference, which is fundamental in cosmology as we only have one universe to observe. We cannot create new, independent universes to measure, and therefore only have one D. $\mathrm{P}(D|\theta,\mathcal{M})$ is therefore repurposed into a function of θ, which we can test with a wide range of θ values and find the maximised $\mathcal{L}(\theta)$ that explains our single cosmic observation. Likelihood is not a PDF like prior and posterior, and it does not necessarily integrate to 1 over θ space.

The evidence, or the marginal likelihood [10], is the probability of observing D given M, derived from the likelihood by integrating over all parameters and weighted by the prior,

$$\mathrm{P}(D|\mathcal{M}) \equiv \mathcal{Z} = \int \mathrm{P}(D|\theta,\mathcal{M})\mathrm{P}(\theta|\mathcal{M})d\theta. \tag{2.4}$$

Dropping the model dependence, we have:

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta. \tag{2.5}$$

It can be intuitively understood as a "prior-weighted average likelihood". Mathematically, it is the normalising constant that ensures the posterior integrates to unity. The evidence is usually ignored during parameter estimation (see Section 2.2) but plays a crucial role in model comparison (see Section 2.3).

2.1.2 Kullback-Leibler Divergence

The Kullback-Leibler (KL) divergence, $\mathcal{D}_{\mathrm{KL}}$, quantifies the information gain, or compression, between the prior distribution $\pi(\theta)$ and the posterior distribution $\mathcal{P}(\theta)$ [11]. It has been widely used by cosmologists [12–22] and is defined as the average of the Shannon Information, $\mathcal{I}(\theta)$, over the posterior:

$$\mathcal{I}(\theta) = \log\frac{\mathcal{P}(\theta)}{\pi(\theta)}, \tag{2.6}$$

$$\mathcal{D}_{\mathrm{KL}} = \int \mathcal{P}(\theta)\log\frac{\mathcal{P}(\theta)}{\pi(\theta)}\,d\theta = \left\langle\log\frac{\mathcal{P}}{\pi}\right\rangle_{\mathcal{P}} = \langle\mathcal{I}\rangle_{\mathcal{P}} \approx \log\left(\frac{V_\pi}{V_P}\right). \tag{2.7}$$

A higher $\mathcal{D}_{\mathrm{KL}}$ indicates a larger information gain when moving from the prior to the posterior and is consequently a useful measure of the constraining power of the data. $\mathcal{D}_{\mathrm{KL}}$ can be understood as approximately the logarithm of the ratio of the prior volume, V_π, to the posterior volume, $V_\mathcal{P}$. This relationship is exact in the case of uniform ("top-hat") prior and

posterior distributions, but remains highly accurate when a broad prior is used, where the prior is 'locally flat" around the posterior peak.

$\mathcal{D}_{\mathrm{KL}}$ is a strong function of the prior, and it inherits the property of being additive for independent parameters from the Shannon Information. A key practical consideration is that its calculation requires a properly normalised posterior distribution, which in turn requires knowledge of the Bayesian evidence, $\mathcal{Z}$. Consequently, this quantity is not attainable with common MCMC sampling techniques, which typically generate samples from an unnormalised posterior. To compute it, more computationally intensive algorithms such as nested sampling are necessary.

The KL divergence can be related directly to the Bayesian evidence via the expression [23]:

$$\log \mathcal{Z} = \langle \log \mathcal{L} \rangle_{\mathcal{P}} - \mathcal{D}_{\mathrm{KL}}, \tag{2.8}$$

where $\langle \log \mathcal{L} \rangle_{\mathcal{P}}$ is the posterior average of the log-likelihood. This relation, sometimes referred to as the Occam's Razor equation [24], illustrates how the evidence naturally implements Occam's Razor by penalising unnecessary model complexity. As discussed further in Section 2.3, the penalty factor between competing cosmological models can be approximated using the difference in their respective $\mathcal{D}_{\mathrm{KL}}$. Using the nested sampling chains from `unimpeded`, we compute $\mathcal{D}_{\mathrm{KL}}$ for every model-dataset combination in our grid, enabling a systematic comparison of the constraining power of different datasets and models. These results are presented in Section 5.4.

2.2 Parameter estimation

The goal of parameter estimation is to determine the posterior probability distribution of the parameters, $\mathrm{P}(\theta|D, \mathcal{M})$. Combining the prior, likelihood and evidence from Section 2.1, $\mathrm{P}(\theta|D, \mathcal{M})$ is stated by the Bayes' theorem as:

$$\mathrm{P}(\theta|D, \mathcal{M}) = \frac{\mathrm{P}(D|\theta, \mathcal{M})\mathrm{P}(\theta|\mathcal{M})}{\mathrm{P}(D|\mathcal{M})}, \tag{2.9}$$

$$\mathcal{P}(\theta) = \frac{\mathcal{L}(\theta) \times \pi(\theta)}{\mathcal{Z}}. \tag{2.10}$$

It describes how our initial beliefs $\pi(\theta)$ about θ are updated in light of the observed data D under the assumed model $\mathcal{M}$. MCMC algorithms are typically used for exploring the posterior, particularly in high-dimensional parameter spaces [25]. However, the evidence cannot be obtained due to technical reasons, so MCMC methods give unnormalised posterior $\mathrm{P}(\theta|D, \mathcal{M}) \propto \mathcal{L}(\theta) \times \pi(\theta)$. However, specialised algorithms like nested sampling, as implemented in codes such as `PolyChord` [26, 27], can also efficiently generate posterior samples while simultaneously calculating the evidence. The `unimpeded` library provides access to both nested sampling and MCMC chains for all model-dataset combinations, enabling comprehensive parameter estimation analysis. We present parameter constraints for key cosmological parameters across our grid in Section 5.2.

2.3 Model comparison

Model comparison addresses the question of how much the data D support each competing models $\{\mathcal{M}_1, \mathcal{M}_2, \cdots\}$, where each model $\mathcal{M}_i$ has its own set of parameters $\theta_{\mathcal{M}_i}$. The goal is to compute the posterior $\mathrm{P}(\mathcal{M}_i|D)$ of a model $\mathcal{M}_i$ is true given the data D, which can be

used to rank and select models. From the Bayes' theorem, we have:

$$P(\mathcal{M}_i|D) = \frac{P(D|\mathcal{M}_i)P(\mathcal{M}_i)}{P(D)}, \tag{2.11}$$

$$= \frac{\mathcal{Z}_i \pi_i}{\sum_j \mathcal{Z}_j \pi_j}. \tag{2.12}$$

Evidence is the inner product of the likelihood function and the prior function over the parameter space, it can also be viewed as the prior-weighted average likelihood. A model's evidence is maximised when its most predictive region of parameter space, i.e. where the prior is highest, coincides with the region of highest likelihood. A model is penalised, however, in two key scenarios: first, a direct conflict where the data favour parameter values the prior deemed unlikely; and second, a penalty for complexity, where a wide prior dilutes the evidence by spreading its predictive power too thinly. Since the prior must integrates to unity, a broader prior implies a lower height, reducing the evidence integral even if the data are well-fit within that space. The evidence thus naturally rewards models that are both predictive and simple, the latter naturally and quantitatively implements Occam's Razor[3].

A common approach is to make no prior assumption. Models have uniform prior $\pi = P(\mathcal{M}_i) = $ constant, i.e. $\pi_i = \pi_j$, and $P(\mathcal{M}_i|D)$ simplifies to the ratio of just the evidence of model $\mathcal{M}_i$ to the sum of evidences of all models under comparison,

$$P(\mathcal{M}_i|D) = \frac{\mathcal{Z}_i}{\sum_j \mathcal{Z}_j}. \tag{2.13}$$

While the Bayes Factor is widely used for comparing two models, the approach in Equation (2.13) provides the advantages of yielding the normalised posterior probability for each model, which is not limited to pairwise comparisons and provides an intuitive ranking among the entire set of competing models. We therefore adopt this method for the analysis in Section 5.3. Using the evidence values computed from nested sampling in `unimpeded`, we calculate model probabilities for all eight cosmological models across all datasets, revealing which model extensions are preferred by individual probes and their combinations.

2.4 Tension Quantification

Tension quantification assesses the statistical consistency between different datasets, say D_A and D_B, when interpreted under a common underlying model $\mathcal{M}$. In a Bayesian context, several metrics can be employed to diagnose and quantify the degree of agreement or disagreement. The following sections describe five such statistics utilised in this work, drawing on established methods from the literature [7, 24, 28]. For readers interested in alternative approaches to quantifying tensions, the DES collaboration paper [28] provides a comprehensive comparison of different tension metrics and their applications to cosmological data. The pre-computed chains from `unimpeded` enable application of any preferred tension metric beyond those presented here. We apply these five complementary statistics to quantify tensions across all pairwise dataset combinations in our grid, with comprehensive results presented in Section 5.5.

---

[3] Among competing hypotheses, the one with the fewest assumptions should be selected.

2.4.1 Combining likelihoods

To perform a joint analysis of two statistically independent datasets, D_A and D_B, their likelihoods are combined multiplicatively: $\mathcal{L}_{AB} = \mathcal{L}_A \mathcal{L}_B$. The posteriors and evidences for the individual and joint datasets are defined as:

$$\mathcal{P}_A = \frac{\mathcal{L}_A \pi_A}{\mathcal{Z}_A}, \quad \mathcal{P}_B = \frac{\mathcal{L}_B \pi_B}{\mathcal{Z}_B}, \quad \mathcal{P}_{AB} = \frac{\mathcal{L}_A \mathcal{L}_B \pi_{AB}}{\mathcal{Z}_{AB}}. \tag{2.14}$$

$$\mathcal{Z}_A = \int \mathcal{L}_A \pi_A \, d\theta, \quad \mathcal{Z}_B = \int \mathcal{L}_B \pi_B \, d\theta,$$

$$\mathcal{Z}_{AB} = \int \mathcal{L}_A \mathcal{L}_B \pi_{AB} \, d\theta. \tag{2.15}$$

Here, π_A, π_B, and π_{AB} denote the prior distributions for the individual and joint analyses. In this work, we assume that the priors agree on the shared parameters [29], such that $\pi_A(\theta_{\text{shared}}) = \pi_B(\theta_{\text{shared}}) = \pi_{AB}(\theta_{\text{shared}})$, while nuisance parameters unique to each dataset retain their respective priors. θ is taken to be the complete set for the joint analysis, including all cosmological parameters and any nuisance parameters unique to each dataset.

2.4.2 The R statistics

The R statistic quantifying the consistency between two datasets, denoted by subscripts A and B, within a shared underlying model $\mathcal{M}$ [30]. It is defined through a series of equivalent expressions that relate the evidences and conditional probabilities of the datasets:

$$R = \frac{\mathcal{Z}_{AB}}{\mathcal{Z}_A \mathcal{Z}_B} = \frac{\mathrm{P}(D_A, D_B)}{\mathrm{P}(D_A)\mathrm{P}(D_B)} = \frac{\mathrm{P}(D_A|D_B)}{\mathrm{P}(D_A)} = \frac{\mathrm{P}(D_B|D_A)}{\mathrm{P}(D_B)}. \tag{2.16}$$

R provides a direct measure of inter-dataset consistency, interpreted with respect to unity. If $R \gg 1$, knowledge of one dataset has strengthened our confidence in the other by a factor of R, indicating concordance. If $R \ll 1$, the datasets are inconsistent. The introduction of the second dataset diminishes our confidence in the first under the assumed model, prompting a re-evaluation of the shared model or the datasets themselves. While this establishes a clear framework for consistency, it is crucial to remember that the magnitude of R does not represent an absolute degree of tension, as its value is always conditional on the chosen model and prior.

The R statistic satisfies several desirable properties for a tension metric: it is dimensionally consistent, symmetric with respect to the datasets ($R_{AB} = R_{BA}$), invariant under reparametrisation, and constructed from fundamental Bayesian quantities. However, a crucial property of R is its strong dependence on the prior probability distribution, $\pi(\theta)$ [8]. This dependency can be made explicit by rewriting R in terms of the posteriors $\mathcal{P}_A$ and $\mathcal{P}_B$:

$$\begin{aligned}
R &= \frac{1}{\mathcal{Z}_A \mathcal{Z}_B} \int \mathcal{L}_A(\theta) \mathcal{L}_B(\theta) \pi(\theta) d\theta \\
&= \int \frac{\mathcal{L}_A(\theta)\pi(\theta)}{\mathcal{Z}_A} \frac{\mathcal{L}_B(\theta)\pi(\theta)}{\mathcal{Z}_B} \frac{1}{\pi(\theta)} d\theta \\
&= \int \frac{\mathcal{P}_A(\theta)\mathcal{P}_B(\theta)}{\pi(\theta)} d\theta \\
&= \left\langle \frac{\mathcal{P}_B}{\pi} \right\rangle_{\mathcal{P}_A} = \left\langle \frac{\mathcal{P}_A}{\pi} \right\rangle_{\mathcal{P}_B},
\end{aligned} \tag{2.17}$$

where we have assumed the datasets are independent. The final line shows that R can be thought of as the posterior average of the ratio of one posterior to the shared prior, averaged over the other posterior.

One should note that reducing the width of the prior on shared, constrained parameters will reduce the value of R, thereby increasing the apparent tension between the datasets. This behaviour is opposite to the prior's effect on evidence alone, where narrower priors typically increase the evidence. This creates an attractive balance: one cannot arbitrarily tune priors to increase evidence for a model without simultaneously making it more susceptible to tension if the datasets are not in perfect agreement. While this prior dependence is a feature of a coherent Bayesian analysis, it means that the interpretation of a single R value requires care. If R indicates discordance, this conclusion is robust, since the prior volume effect typically acts to increase R and mask tension. However, if R indicates agreement, one must consider if this is merely the result of an overly wide prior.

2.4.3 The Information Ratio

The information ratio, I, is defined in terms of the Kullback-Leibler divergences $\mathcal{D}_{\mathrm{KL}}$ from the individual data (D_A, D_B) and joint data (D_{AB}) analyses [8]:

$$\log I = \mathcal{D}_{\mathrm{KL}}^A + \mathcal{D}_{\mathrm{KL}}^B - \mathcal{D}_{\mathrm{KL}}^{AB}. \tag{2.18}$$

To understand the behaviour of I, we can employ the volumetric approximation of Equation (2.7), $\mathcal{D}_{\mathrm{KL}} \approx \log(V_\pi) - \log(V_{\mathcal{P}})$, as discussed in Section 2.1.2. Substituting this into the Equation (2.18) yields:

$$\log I \approx \log(V_\pi) - \log(V_{\mathcal{P}}^A) - \log(V_{\mathcal{P}}^B) + \log(V_{\mathcal{P}}^{AB}). \tag{2.19}$$

In a Bayesian interpretation of probability, a highly improbable event is a highly surprising one. I quantifies this "surprise" of agreement, i.e. different datasets making the same predictions. Considering the $\log(V_\pi)$ term, a larger prior volume V_π signifies greater initial uncertainty, making the subsequent agreement of two constraining datasets (with small posterior volumes $V_{\mathcal{P}}^A$ and $V_{\mathcal{P}}^B$) a more surprising outcome, which results in a larger value of $\log I$ and mathematically encodes this greater degree of surprise. In addition, a more constraining dataset results in a smaller $V_{\mathcal{P}}$, making $\log(V_{\mathcal{P}})$ a larger negative number and thus $-\log(V_{\mathcal{P}})$ a larger positive number. Consequently, the terms $-\log(V_{\mathcal{P}}^A)$ and $-\log(V_{\mathcal{P}}^B)$ increase the value of $\log I$. This is intuitive: if two highly constraining posteriors (tiny $V_{\mathcal{P}}$) end up agreeing, it is far more surprising than if two vague, less constraining posteriors (large $V_{\mathcal{P}}$) happen to agree. Conversely, if the datasets are in tension, their posteriors barely overlap, causing the joint posterior volume $V_{\mathcal{P}}^{AB}$ to become extremely small. This makes $\log(V_{\mathcal{P}}^{AB})$ a large negative number, which in turn significantly decreases the value of $\log I$. Therefore, a very low or negative $\log I$ is a strong signal of dataset disagreement.

2.4.4 Suspiciousness

While the information ratio quantifies the surprise of agreement, the suspiciousness S quantifies the statistical conflict between the likelihoods $\mathcal{L}$ of the two datasets. It is defined by both prior-dependent information ratio I and the R statistics [8], and has been applied to quantify tensions in various cosmological contexts [2, 7]:

$$S = \frac{R}{I}, \quad \log S = \log R - \log I. \tag{2.20}$$

Since R and I transform similarly under prior volume alterations, S is largely unaffected by changing the prior widths, as long as this change does not significantly alter the posterior. However, this prior-independence comes at the cost of the direct probabilistic interpretation inherent in R, requiring more care to calibrate its scale for significance. Substituting the definitions of $\log R$ and $\log I$ from Equations (2.8) and (2.16) and Equation (2.18) into Equation (2.20), $\log S$ can be expressed directly in terms of posterior-averaged log-likelihoods:

$$\log S = \langle \log \mathcal{L}_{AB} \rangle_{\mathcal{P}_{AB}} - \langle \log \mathcal{L}_A \rangle_{\mathcal{P}_A} - \langle \log \mathcal{L}_B \rangle_{\mathcal{P}_B}. \tag{2.21}$$

When likelihoods $\mathcal{L}_A$ and $\mathcal{L}_B$ are in strong agreement, $\log S$ is zero or positive. Conversely, if the likelihoods are in tension, $\log S$ becomes negative, with larger negative values indicating stronger tension. $\log S$ can also be calibrated into a tension probability, p, and an equivalent significance in Gaussian standard deviations, σ (see Section 2.4.6).

2.4.5 Bayesian Model Dimensionality

While the Kullback-Leibler divergence, $\mathcal{D}_{\mathrm{KL}}$, discussed in Section 2.1.2 provides a single value for the total information gain, it marginalises out any information about individual parameters. It cannot tell us how many parameters are being constrained by the data, nor what each parameter is constraining. For instance, a strong, correlated constraint between two parameters can yield the same $\mathcal{D}_{\mathrm{KL}}$ as two well-constrained but independent parameters (visually demonstrated in [24]). In high-dimensional cosmological analyses, where corner plots [31] show only marginalised views and can hide complex degeneracies, a metric is needed to quantify the effective number of constrained parameters. The Bayesian Model Dimensionality, d, was introduced to fulfil this role [24] and is defined as:

$$\begin{aligned}
\frac{d}{2} &= \int \mathcal{P}(\theta) \left(\log \frac{\mathcal{P}(\theta)}{\pi(\theta)} - \mathcal{D}_{\mathrm{KL}} \right)^2 d\theta \\
&= \left\langle \left(\log \frac{\mathcal{P}}{\pi} \right)^2 \right\rangle_{\mathcal{P}} - \left\langle \log \frac{\mathcal{P}}{\pi} \right\rangle_{\mathcal{P}}^2 \\
&= \mathrm{var}(\mathcal{I})_{\mathcal{P}} \\
&= \langle (\log \mathcal{L})^2 \rangle_{\mathcal{P}} - \langle \log \mathcal{L} \rangle_{\mathcal{P}}^2,
\end{aligned} \tag{2.22}$$

where $\mathcal{I} = \log \frac{\mathcal{P}}{\pi}$ is the Shannon Information mentioned in Equation (2.6). d is the variance of $\mathcal{I}$ over posterior, and hence, a higher-order statistic than the KL divergence. The Bayesian Model Dimensionality possesses several important properties. Crucially, it is only weakly prior-dependent, as the evidence contributions required to normalise the posterior and prior in the $\mathcal{I}$ term in Equation (2.22) effectively cancel out. Furthermore, like the KL divergence, it is additive for independent parameters and invariant under a change of variables. When combining datasets, the number of parameters that are constrained becomes:

$$d_{A \cap B} = d_A + d_B - d_{AB}. \tag{2.23}$$

2.4.6 Tension Probability and Significance

Suspiciousness discussed in Section 2.4.4 can be calibrated into a more intuitive tension probability, p, and an equivalent significance expressed in Gaussian standard deviations, σ. This calibration relies on the approximation that in the case of a Gaussian likelihood, the quantity $d - 2 \log S$ follows a χ^2 distribution. The number of degrees of freedom is given

by the Bayesian model dimensionality, d, as discussed in Section 2.4.5. p represents the probability that a level of discordance at least as large as the one observed could arise by chance. It is calculated using the survival function of the χ_d^2 distribution:

$$p = \int_{d-2\log S}^{\infty} \chi_d^2(x)\,\mathrm{d}x = \int_{d-2\log S}^{\infty} \frac{x^{d/2-1}e^{-x/2}}{2^{d/2}\Gamma(d/2)}\,\mathrm{d}x. \tag{2.24}$$

This p-value can then be converted into an equivalent significance on a Gaussian scale, σ, using the inverse complementary error function (Erfc^{-1}):

$$\sigma = \sqrt{2}\,\mathrm{Erfc}^{-1}(p). \tag{2.25}$$

Following standard conventions, if $p \lesssim 0.05$ (corresponding to $\sigma \gtrsim 2$), the datasets are considered to be in moderate tension, while $p \lesssim 0.003$ ($\sigma \gtrsim 3$) corresponds to strong tension.

2.4.7 The Look Elsewhere Effect

The Look Elsewhere Effect (LEE) arises when multiple statistical tests are performed, increasing the probability of finding a seemingly significant result purely by chance. This effect is particularly relevant to our analysis, where we systematically evaluate tension statistics for $N = 248$ distinct model-dataset combinations (see Section 5). Without accounting for multiple comparisons, the likelihood of encountering at least one false positive becomes substantial.

Rather than applying a Bonferroni correction to each individual p-value (which would change as the grid expands), we instead adopt a significance threshold that naturally accounts for the look elsewhere effect. Under the null hypothesis of no genuine tension, p-values are uniformly distributed between 0 and 1. Therefore, if we perform $N = 248$ independent tests, we expect exactly one result to have $p \leq 1/N$ purely by chance. This provides a natural threshold:

$$\sigma_{\text{threshold}} = \sqrt{2}\,\mathrm{Erfc}^{-1}\left(\frac{1}{N}\right). \tag{2.26}$$

For our grid with $N = 248$, this gives $\sigma_{\text{threshold}} \approx 2.88$. This threshold is not arbitrary, it represents the significance level at which we would expect only one false positive across all 248 tests if there were no genuine tensions. Any result highlighted above this threshold is more extreme than what we would expect from random fluctuations alone. This approach has the advantage that the threshold itself reflects the scope of the analysis, while individual p-values and σ values remain interpretable independently of the grid size.

2.4.8 Model-Weighted Average Tension Statistics

To evaluate the overall tension between datasets across our entire model space, we employ a model-weighted average for each tension statistic. This approach provides a single, summary ranking of datasets that accounts for the fact that some models are better supported by the data than others, and therefore, their tension statistics should have a heavier weighting under the Bayesian framework. The tension statistic for each model is weighted by its posterior probability, $\mathrm{P}(\mathcal{M}_i|D)$, calculated by Equation (2.12). For a generic tension metric between datasets D_A and D_B, this average is computed as:

$$\langle\mathrm{Statistic}(D_A, D_B)\rangle_{\mathcal{M}} = \sum_i \mathrm{P}(\mathcal{M}_i|D) \times \mathrm{Statistic}(\mathcal{M}_i, D_A, D_B). \tag{2.27}$$

For example, the model-weighted Kullback-Leibler divergence is:

$$\langle \mathcal{D}_{\mathrm{KL}} \rangle_{\mathrm{P}(\mathcal{M})} = \sum_i \mathrm{P}(\mathcal{M}_i | D) \times \mathcal{D}_{\mathrm{KL}}(\mathcal{M}_i). \tag{2.28}$$

Similarly, the model-weighted tension significance is:

$$\langle \sigma \rangle_{\mathrm{P}(\mathcal{M})} = \sum_i \mathrm{P}(\mathcal{M}_i | D) \times \sigma(\mathcal{M}_i). \tag{2.29}$$

These model-weighted statistics are presented in the heatmaps in Sections 5.3 and 5.5.

2.5 Cosmological Models

We consider a comprehensive set of cosmological models extending the standard ΛCDM paradigm. All models are described within the framework of the Friedmann-Lemaître-Robertson-Walker (FLRW) metric with general relativity. The background expansion is governed by the Friedmann equation, and initial conditions are set by nearly scale-invariant, adiabatic Gaussian scalar perturbations [32]. All models are implemented using the Cobaya framework [33, 34], which interfaces with the CAMB Boltzmann code [35] to compute theoretical predictions for the observables. Each model is elaborated in detail in the following subsubsections.

2.5.1 Baseline: ΛCDM

The baseline ΛCDM model describes a spatially flat universe with a cosmological constant (dark energy equation of state $w = -1$), cold dark matter, and a power-law spectrum of adiabatic scalar perturbations. The cold dark matter paradigm was established by [36], while the cosmological constant component emerged from the discovery of cosmic acceleration through Type Ia supernovae observations [37, 38]. The model has been extensively validated by cosmic microwave background measurements [32]. The model is characterized by six fundamental parameters: the baryon density parameter $\omega_b = \Omega_b h^2$, the cold dark matter density parameter $\omega_c = \Omega_c h^2$, the angular scale of the sound horizon at recombination θ_* (often parameterised as $100\theta_{MC}$), the reionisation optical depth τ, the scalar spectral index n_s, and the amplitude of scalar perturbations $\ln(10^{10} A_s)$.

The Hubble parameter evolves as:

$$H^2(a) = H_0^2 \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_\Lambda \right], \tag{2.30}$$

where a is the scale factor, Ω_r is the radiation density parameter, Ω_m is the total matter density parameter, and Ω_Λ is the cosmological constant density parameter. The flatness constraint imposes $\Omega_r + \Omega_m + \Omega_\Lambda = 1$.

The primordial scalar power spectrum is:

$$\mathcal{P}_s(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1}, \tag{2.31}$$

where $k_0 = 0.05\,\mathrm{Mpc}^{-1}$ is the pivot scale. The baseline assumes three standard neutrinos with $N_{\mathrm{eff}} = 3.046$ and minimal neutrino mass $\Sigma m_\nu = 0.06\,\mathrm{eV}$. The helium abundance is computed consistently with Big Bang nucleosynthesis.

Free parameters: ω_b, ω_c, θ_*, τ, n_s, $\ln(10^{10} A_s)$.

2.5.2 Varying curvature: $\Omega_k \Lambda$CDM

This extension allows for spatial curvature by freeing the curvature density parameter Ω_k [32]. The Friedmann equation becomes:

$$H^2(a) = H_0^2 \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda \right], \qquad (2.32)$$

with the constraint $\Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda = 1$. Positive Ω_k corresponds to an open universe (negative spatial curvature), while negative Ω_k describes a closed universe (positive spatial curvature). Non-zero curvature alters the universe's spatial geometry, modifying the evolution of metric perturbations and photon geodesics. Boltzmann codes account for this by solving the perturbation equations on a curved background, which causes a characteristic angular shift in the CMB power spectrum's acoustic peaks and modifies the late-time integrated Sachs-Wolfe effect.

Free parameters: ΛCDM parameters $+ \Omega_k$.

2.5.3 Constant dark energy equation of state: wCDM

This model generalizes the cosmological constant to a dark energy fluid with constant equation of state $w = p_{\mathrm{DE}}/\rho_{\mathrm{DE}}$ [39]. The dark energy density evolves as $\rho_{\mathrm{DE}}(a) \propto a^{-3(1+w)}$, modifying the Friedmann equation to:

$$H^2(a) = H_0^2 \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_{\mathrm{DE}} a^{-3(1+w)} \right]. \qquad (2.33)$$

The model assumes spatial flatness. Within the Parameterised Post-Friedmann (PPF) framework, dark energy is treated as a perfect fluid with a sound speed typically set to unity ($c_s^2 = 1$). Consequently, dark energy perturbations are negligible on sub-horizon scales, influencing structure growth primarily through the modified background expansion and the late-time integrated Sachs-Wolfe (ISW) effect on the CMB.

Free parameters: ΛCDM parameters $+ w$.

2.5.4 Time-varying dark energy: $w_0 w_a$CDM

This model allows for time-varying dark energy using the Chevallier-Polarski-Linder (CPL) parameterisation [40, 41]:

$$w(a) = w_0 + w_a(1 - a), \qquad (2.34)$$

where w_0 is the present-day equation of state and w_a characterizes its time evolution. The dark energy density evolves as:

$$\rho_{\mathrm{DE}}(a) = \rho_{\mathrm{DE},0} \, a^{-3(1+w_0+w_a)} \exp[-3w_a(1-a)]. \qquad (2.35)$$

The model assumes spatial flatness. Similar to wCDM, this model uses the PPF formalism where dark energy is a smooth component with $c_s^2 = 1$, preventing it from clustering. The time-varying equation of state produces a more complex background evolution, altering the growth history of matter perturbations and creating a distinct signature in the late-time ISW effect compared to a constant w.

Free parameters: ΛCDM parameters $+ w_0 + w_a$.

2.5.5 Varying lensing amplitude: $A_L\Lambda$CDM

This model introduces a phenomenological parameter A_L that scales the lensing potential power spectrum, allowing for deviations from the standard lensing predictions [32]. The parameter modifies the lensed CMB power spectra by scaling the lensing potential correlations:

$$C_\ell^{\text{lensed}} = C_\ell^{\text{unlensed}} + A_L \Delta C_\ell^{\text{lensing}}, \tag{2.36}$$

where $\Delta C_\ell^{\text{lensing}}$ represents the correction due to gravitational lensing. The standard ΛCDM prediction corresponds to $A_L = 1$. Values $A_L > 1$ indicate enhanced lensing effects, while $A_L < 1$ suggest reduced lensing. This extension was motivated by the Planck collaboration's observation of a preference for $A_L > 1$ in the CMB temperature data, providing a way to test the consistency of gravitational lensing predictions. This parameter does not alter the physical evolution of perturbations but acts as a phenomenological scaling of the gravitational lensing potential. In Boltzmann codes, the calculated lensing potential power spectrum is multiplied by A_L, which directly modifies the smoothing of the CMB acoustic peaks and the amplitude of the lensing-induced B-mode spectrum.

Free parameters: ΛCDM parameters $+ A_L$.

2.5.6 Varying neutrino masses: $m_\nu\Lambda$CDM

This model extends the standard ΛCDM framework by allowing the sum of the three active neutrino masses, Σm_ν, to vary as a free parameter. The effective number of relativistic species is held fixed at the standard value, $N_{\text{eff}} = 3.046$ [42]. The contribution of massive neutrinos to the cosmic energy budget today is:

$$\Omega_\nu h^2 = \frac{\Sigma m_\nu}{93.14\,\text{eV}}. \tag{2.37}$$

Massive neutrinos act as hot dark matter, suppressing the growth of structure below their free-streaming length due to their large thermal velocities. Boltzmann codes solve the full neutrino Boltzmann equation to model this effect, which manifests as a distinct, scale-dependent suppression in the matter power spectrum and subtly alters the CMB via the early ISW effect.

Free parameters: ΛCDM parameters $+ \Sigma m_\nu$.

2.5.7 Running spectral index: $n_{\text{run}}\Lambda$CDM

This model extends the primordial power spectrum beyond a simple power law by allowing the scalar spectral index n_s to vary with scale k. This scale dependence is parameterised by the "running of the spectral index," defined as $n_{\text{run}} \equiv dn_s/d\ln k$ [32, 43]. While the baseline ΛCDM model assumes $n_{\text{run}} = 0$, a non-zero running is a generic prediction of many inflationary models. In the context of single-field slow-roll inflation, the running is a second-order effect in the slow-roll parameters and is predicted to be very small. A detection of a significant non-zero running would therefore challenge the simplest inflationary scenarios and provide crucial insights into the shape of the inflaton potential or point towards more complex physics in the early universe.

The primordial scalar power spectrum is modified to include a logarithmic scale-dependent term in the exponent:

$$\mathcal{P}_s(k) = A_s \left(\frac{k}{k_0}\right)^{n_s - 1 + \frac{1}{2} n_{\text{run}} \ln\left(\frac{k}{k_0}\right)}, \tag{2.38}$$

where n_s is the spectral index and n_{run} is the running, both evaluated at the pivot scale $k_0 = 0.05\,\mathrm{Mpc}^{-1}$. This form arises from a first-order Taylor expansion of the spectral index $n_s(k)$ in $\ln k$.

     Free parameters: ΛCDM parameters $+\ n_{\mathrm{run}}$.

2.5.8   Primordial gravitational waves: $r\Lambda$CDM

This model includes primordial tensor perturbations (gravitational waves) characterized by the tensor-to-scalar ratio r at the pivot scale, providing a key test of inflationary theory [44, 45]. The spectrum of relic gravitational waves from inflation was first calculated by [46]. The tensor power spectrum is:

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_p} \right)^{n_t}, \tag{2.39}$$

where $r = A_t/A_s$ is evaluated at a chosen pivot scale (typically $k_p = 0.002\,\mathrm{Mpc}^{-1}$ for tensor modes). The tensor spectral index n_t is often constrained by the inflationary consistency relation $n_t = -r/8$.

     Free parameters: ΛCDM parameters $+\ r$ (and optionally n_t if not fixed by consistency relation).

3   Methodology

3.1   Nested Sampling

Nested sampling, introduced by Skilling (2004) [6, 47], is a Monte Carlo method designed for Bayesian computation. It is particularly powerful for calculating the Bayesian evidence (or marginal likelihood), a key quantity for model comparison and tension quantification, while simultaneously producing posterior samples for parameter estimation. The algorithm transforms the multi-dimensional evidence integral into a one-dimensional integral over prior volume, which is then solved numerically.

     Nested sampling is generally considered the "ground truth" method for evidence calculation, representing the reference standard against which other approaches are compared. While alternative methods exist that aim to calculate evidence more efficiently, such as harmonic mean estimators [48] and MC evidence [49], these typically require validation against nested sampling results to establish their accuracy. The pre-computed grid of nested sampling chains provided by `unimpeded` therefore serves not only as a resource for cosmological model comparison and tension quantification, but also as a reference dataset for assessing the performance of alternative evidence estimation techniques. Recent comprehensive reviews of nested sampling methodology and applications can be found in [50, 51].

     We use the publicly available `PolyChord` sampler [26, 27], which provides a robust and efficient implementation of nested sampling well suited to the high-dimensional parameter spaces typical of modern cosmology. This section outlines the core methodology of the nested sampling algorithm.

3.1.1   Generating Samples and Increasing Likelihood

The fundamental principle of nested sampling is to explore the parameter space θ by iteratively moving through nested contours of constant likelihood $\mathcal{L}_i$. The algorithm begins by drawing n_0 initial "live points" from the prior distribution $\pi(\theta)$. At each iteration i, the point with the lowest likelihood, $\mathcal{L}_i$, among the current set of live points is identified. By

some criterion we then choose whether or not to remove this from the live set and add it to a collection of "dead points". By a different criterion it is then replaced with new points drawn from the prior $\pi(\theta)$, subject to the hard constraint that their likelihood $\mathcal{L}(\theta)$ must be greater than $\mathcal{L}_i$. This ensures that the likelihoods of the dead points, $\{\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \dots\}$, form a monotonically increasing sequence.

3.1.2 Prior Volume Contraction

This iterative deletion of live points with the lowest likelihood systematically contracts the region of parameter space, leading to the peak(s) of the posterior. The prior mass $X(\mathcal{L})$ is the measure of the fraction of the prior mass[4] contained within an iso-likelihood contour $\mathcal{L}(\theta) = \mathcal{L}$. It is calculated by integrating the element of prior mass $dX = \pi(\theta)d\theta$ covering all likelihood values greater than $\mathcal{L}$[6]:

$$X(\mathcal{L}) = \int_{\mathcal{L}(\theta) > \mathcal{L}} \pi(\theta)d\theta. \tag{3.1}$$

By construction described in Section 3.1.1, the algorithm generates a sequence of increasing likelihoods $\mathcal{L}_1 < \mathcal{L}_2 < \cdots < \mathcal{L}_i$ corresponding to a sequence of shrinking prior volumes $X_i = X(\mathcal{L}_i)$, where $X_1 > X_2 > \cdots > X_i$. At each iteration i, the removal of the point with likelihood $\mathcal{L}_i$ corresponds to shrinking the prior volume from X_{i-1} to X_i, so the ratio of the volumes $t_i = X_i/X_{i-1}$. The initial prior volume is $X_0 = 1$ (the entire prior).

More formally, the shrinkage of the prior volume is a stochastic process with distribution $P(t_i) = n_i t_i^{n_i-1}$, where n_i is the live point count at iteration i. The expected logarithm of the prior volume at iteration i is given by the sum over the live point count n_k at each preceding iteration k [52]:

$$\langle \log X_i \rangle = -\sum_{k=1}^{i} \frac{1}{n_k}. \tag{3.2}$$

In the simplified case of a constant number of live points, $n_k = n_{\text{live}}$ for all k, this sum reduces to $\langle \log X_i \rangle = -i/n_{\text{live}}$. This leads to the well-known exponential approximation for the prior volume, $\langle X_i \rangle \approx e^{-i/n_{\text{live}}}$.

This exponential compression allows the algorithm to efficiently traverse the parameter space from the broad prior towards the narrow, high-likelihood regions where the posterior mass is concentrated. However, as we discuss in Section 3.1.7, Equation (3.2) provides the rigorous framework necessary for analysing runs where the number of live points varies. Each dead point is associated with a specific set of parameters θ_i, likelihood $\mathcal{L}_i$ and an estimated prior volume X_i, enabling the reconstruction of the evidence integral, discussed in Section 3.1.3.

3.1.3 Evidence Estimation

The primary strength of nested sampling is its ability to directly calculate the Bayesian evidence, $\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta$. Instead of integrating the likelihood over all possible parameters, which is computationally prohibitive in high dimensions, this integral can be reformulated in terms of the prior volume X. Since $\mathcal{L}$ can be expressed as an inverse function of its enclosed

---

[4]In the context of nested sampling, "prior volume" and "prior mass" are used interchangeably to refer to the same fundamental concept.

prior volume, $\mathcal{L}(X)$, the evidence integral can be rewritten as a one-dimensional integral from $X = 0$ to $X = 1$ [6]:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)dX. \tag{3.3}$$

The nested sampling algorithm provides a discrete sequence of points $(\mathcal{L}_i, X_i)$ that allows for a numerical approximation of this integral. Using a simple quadrature scheme, the evidence can be estimated as a weighted sum over the discarded "dead" points:

$$\mathcal{Z} \approx \sum_{i \in \text{dead}} w_i \mathcal{L}_i, \tag{3.4}$$

where $\mathcal{L}_i$ is the likelihood of the i-th discarded point and w_i is the associated prior volume, or weight. This weight represents the prior mass contained within the shell between successive likelihood contours, $w_i = X_{i-1} - X_i$. In practice, we approximate $X_i \approx e^{-i/n_{\text{live}}}$ and can thereby estimate the weights for the summation. The evidence can then be used in model comparison and tension quantification.

3.1.4 Importance Weight and Posterior Estimation

In addition to evidence calculation, the collection of live and dead points can be used to derive posterior inferences, and hence, for parameter estimation. Each "dead point" θ_i is associated with a likelihood $\mathcal{L}_i$ and a prior mass weight, w_i, which represents the element of prior mass of the shell in which point θ_i was sampled. The importance weight, or the posterior, p_i for each dead point is its contribution to the evidence $\mathcal{Z}_i = w_i \mathcal{L}_i$, normalised by the total evidence $\mathcal{Z}$ from Equation (3.4):

$$p_i = \frac{w_i \mathcal{L}_i}{\mathcal{Z}}. \tag{3.5}$$

The posterior expectation value for a function of the parameters, $f(\theta)$, is the weighted sum:

$$\langle f(\theta) \rangle \approx \sum_{j \in \{\text{dead}\}} p_j f(\theta_j). \tag{3.6}$$

This allows for the construction of marginalised posterior distributions, credibility intervals, and other standard Bayesian parameter summaries.

3.1.5 Algorithm Termination and Stopping Criterion

As iterations repeat, prior mass weights w_i monotonically decrease, and the likelihoods $\mathcal{L}_i$ monotonically increase. Live points are therefore concentrated in regions of high likelihood, and are associated with tiny prior mass. The nested sampling algorithm is terminated at the i-th iteration when the remaining posterior mass is some small fraction of the currently calculated evidence:

$$Z_{\text{live}} \approx \langle \mathcal{L}_{\text{live}} \rangle X_i, \tag{3.7}$$

where $\langle \mathcal{L}_{\text{live}} \rangle$ is the average likelihood of the current live points. By this stage, the estimated remaining evidence from the live points is a negligible fraction of the evidence accumulated thus far. A common stopping criterion is to halt the process when the expected future contribution to the evidence is smaller than a user-defined tolerance ϵ:

$$Z_{\text{live}} < \epsilon Z_{\text{dead}}. \tag{3.8}$$

The stopping criterion ensures that the final evidence estimate and posterior samples are robust and that computational effort is not wasted on regions of the parameter space with insignificant posterior mass.

3.1.6 Evidence Correction for Unphysical Parameter Space

The standard nested sampling algorithm, as outlined in the preceding sections, implicitly assumes that the entire prior volume is accessible and yields a non-zero likelihood. In practice, many cosmological models possess parameter spaces with regions that are "unphysical." These are regions where the model violates fundamental physical constraints, such as predicting a negative age for the Universe, failing to converge during numerical evolution, or producing spectra with unphysical features. In our analysis pipeline, these unphysical points are assigned a minimal log-likelihood value, effectively a numerical log-zero. We therefore partition the parameter space $\theta \in \Omega$ into two disjoint subspaces: the "physical" subspace Ω_{phys}, where the likelihood $\mathcal{L}(D|\theta, \mathcal{M}) > 0$, and the "unphysical" subspace Ω_{unphys}, where $\mathcal{L}(D|\theta, \mathcal{M}) \leq 0$. The process of generating samples from the prior $\pi(\theta|\mathcal{M})$ typically begins by drawing a point from a unit hypercube, a D-dimensional space $[0,1]^D$, where D is the number of model parameters $\theta_{\mathcal{M}}$, with each coordinate sampled uniformly as $(u_1, u_2, \ldots, u_D)$ where $u_i \in [0,1]$, which is then transformed into the physical parameter space via the prior transformation. The physical parameter space corresponds to the actual parameter ranges (e.g., $H_0 \in [60, 80]$ km s^{-1} Mpc^{-1}, $\Omega_{\mathrm{b}} h^2 \in [0.019, 0.026]$), obtained by applying the appropriate prior type (uniform, Gaussian, log-uniform, etc.) to each unit hypercube coordinate. This transformed point may fall into either Ω_{phys} or Ω_{unphys}. Since only points in Ω_{phys} will enter the nested sampling algorithm and contribute to the evidence integral, it is crucial to account for the fraction of the prior volume that is inaccessible due to unphysicality. This accessible volume fraction can be estimated via rejection sampling:

$$V_{\mathrm{phys}} \approx \frac{n_{\mathrm{prior}}}{n_{\mathrm{total}}} = \frac{\# \text{ of points in } \Omega_{\mathrm{phys}}}{\# \text{ of points in } \Omega_{\mathrm{phys}} + \# \text{ of points in } \Omega_{\mathrm{unphys}}}. \tag{3.9}$$

The corrected, true evidence is therefore:

$$\mathcal{Z}_{\mathrm{true}} = \mathcal{Z}_{\mathrm{raw}} \times \left(\frac{n_{\mathrm{prior}}}{n_{\mathrm{total}}} \right). \tag{3.10}$$

In logarithmic form, which is used for all computations, the correction is additive:

$$\log(\mathcal{Z}_{\mathrm{true}}) = \log(\mathcal{Z}_{\mathrm{raw}}) + \log \left(\frac{n_{\mathrm{prior}}}{n_{\mathrm{total}}} \right). \tag{3.11}$$

It is important to note that the correction factor is $n_{\mathrm{prior}}/n_{\mathrm{total}}$, not $n_{\mathrm{live}}/n_{\mathrm{total}}$. This is because the volume shrinkage during the initial compression phase (where n_{prior} is reduced to n_{live}) is already correctly tracked by `PolyChord`'s `update_evidence()` routine at each nested sampling step. The ratio $n_{\mathrm{prior}}/n_{\mathrm{total}}$ accounts only for the fraction of parameter space that is physical versus unphysical, as determined by the initial rejection sampling.

3.1.7 Dynamic Nested Sampling & Synchronous Parallel Sampling

Our statistical analysis is performed using the dynamic nested sampling framework implemented in `PolyChord` [26, 27]. While this framework supports adaptive live point allocation [53], we use a constant target number of live points during the main sampling phase,

leveraging the framework's efficient synchronous parallelisation across HPC cores and integrated termination scheme. Our runs proceed through three distinct phases, as illustrated in Figure 1.

Initial Compression. The process begins with an initial set of $n_{\text{prior}} \approx 10,000$ live points sampled directly from the prior and verified to yield physical solutions (i.e., $\mathcal{L}(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M}) > 0$). The compression phase consists of the first $n_{\text{prior}} - n_{\text{live}} \approx 9,000$ iterations of the nested sampling algorithm, during which the lowest-likelihood points are sequentially deleted without replacement, reducing the live point count from n_{prior} to the target value $n_{\text{live}} \approx 1,000$, which will enter the main nested sampling stage. During this phase, the number of active points n_k in Equation (3.2) decreases from n_{prior} down to n_{live}. Each deleted point contributes to the evidence integral with its appropriate prior volume weight $w_i = X_{i-1} - X_i$, where the prior volume shrinks according to the decreasing live point count. This phase efficiently accumulates evidence from the vast, low-likelihood regions of the prior volume.

Synchronous Parallel Sampling. During the main sampling phase, the live point count oscillates above the target value $n_{\text{live}} \approx 1,000$, corresponding to $n_k \approx n_{\text{live}}$ in Equation (3.2). These oscillations, clearly visible in Figure 1, are a characteristic feature of `PolyChord`'s synchronous parallelisation scheme. In each iteration, a batch of the lowest-likelihood points (equal to the number of parallel cores, in our case 760) is discarded, and the same number of new points are generated simultaneously. This synchronous approach, where all cores must wait for the slowest likelihood evaluation in the batch to complete, is crucial for preventing statistical bias. An asynchronous approach would preferentially sample regions of parameter space with faster likelihood evaluations (e.g., flat universes over curved ones), leading to an incorrect posterior.

Final Deletion. The run terminates with a final phase where all remaining live points are systematically removed one by one. In this stage, the active point count n_k decreases from n_{live} down to 1. This process is the dynamic framework's integrated termination procedure, which replaces the separate calculation of the remaining evidence $Z_{\text{live}} \approx \langle \mathcal{L} \rangle_{\text{live}} X_i$ found in earlier nested sampling implementations [27]. The total number of iterations required is related to the Kullback–Leibler divergence between the prior and the posterior. However, a strict precision criterion can cause nested sampling to continue beyond the point of maximum information gain from prior to posterior (as reflected by $\mathcal{D}_{\text{KL}}$). For example, in Figure 1, the iteration count ratio for Pantheon between $w_0 w_a$CDM and ΛCDM is approximately 1:2.66, whilst the corresponding $\mathcal{D}_{\text{KL}}$ ratio (shown in Figure 6) is only 1:14. This discrepancy arises because the precision criterion requires nested sampling to carry on sampling even after the bulk of information gain has been extracted, ensuring convergence to a high-precision evidence estimate.

Prior Volume in a Three-Phase Run. The varying number of live points during the compression and deletion phases means the simple exponential approximation for prior volume does not hold throughout the entire run. The rigorous relationship between iteration number and prior volume is given by Equation (3.2), which correctly accounts for the changing live point count n_k throughout all three phases of the run [52]. This provides a precise mapping between the iteration number (x-axis of Figure 1) and the expected log-prior volume being explored.

Figure 1 demonstrates two key features: (1) for the same dataset but different models (e.g., ΛCDM and $w_0 w_a$CDM for `Pantheon`, shown in orange and green), the iteration count increases only slightly for the more complex model; (2) for the same model but different
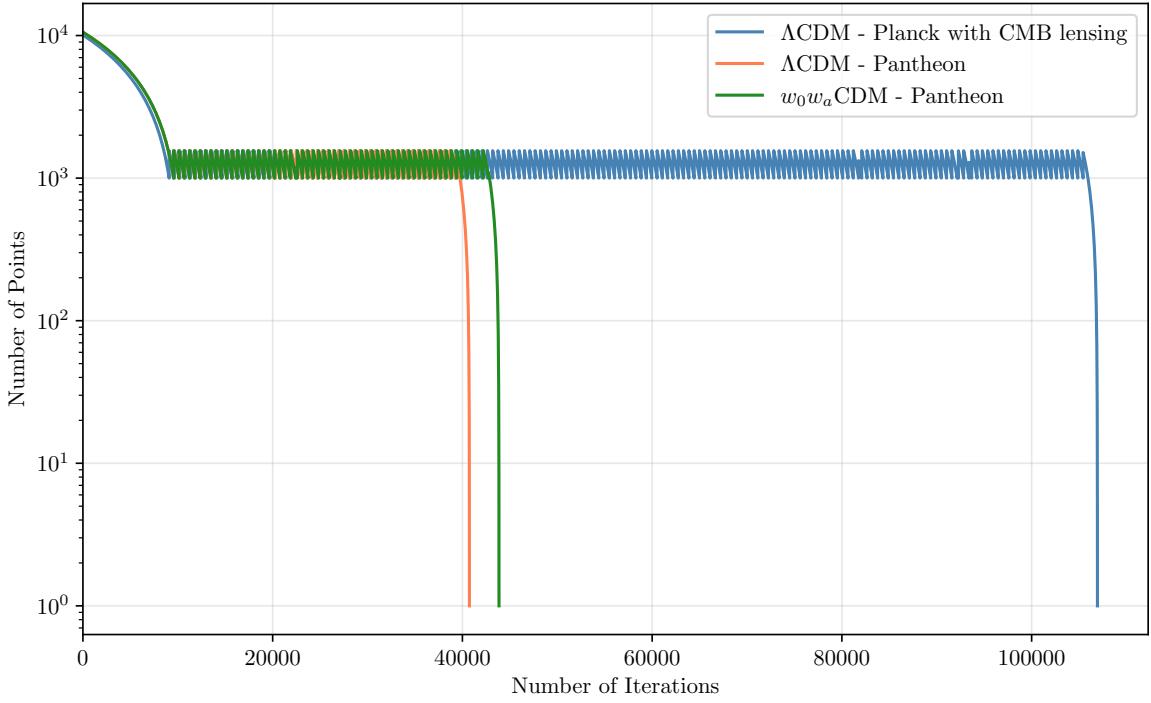
Figure 1: Evolution of the live point count throughout nested sampling runs for ΛCDM and $w_0 w_a$CDM models with Planck with CMB lensing and Pantheon datasets. Three distinct phases are visible: (1) initial compression where the first $n_{\mathrm{prior}} - n_{\mathrm{live}} \approx 9{,}000$ iterations sequentially delete the lowest-likelihood points without replacement, reducing the live point count from $n_{\mathrm{prior}} \approx 10^4$ to the target value $n_{\mathrm{live}} \approx 10^3$, which will enter the main nested sampling stage; (2) main sampling phase where the live point count oscillates above n_{live} due to synchronous parallel processing with 760 cores; (3) final deletion phase where the live point count decreases from n_{live} to zero as the remaining points are systematically removed one by one. The iteration number i (x-axis) maps to the compressed log-prior volume via $\langle \log X_i \rangle = -\sum_{k=1}^{i} 1/n_k$, where n_k is the live point count at iteration k, as shown by the y-axis (see Section 3.1.7 for details) [52]. The different termination points reflect the different Kullback–Leibler divergences between prior and posterior for each model-dataset combination. More complex models like $w_0 w_a$CDM (green) require slightly more iterations than simpler models like ΛCDM (orange), but this effect is not as dominant as the variation across different datasets, with Planck with CMB lensing (blue) requiring substantially more iterations than Pantheon (orange, green).

datasets (e.g., ΛCDM with `Planck+lensing` vs. `Pantheon`, shown in blue and orange), the iteration count varies significantly, with `Planck`'s larger parameter space requiring substantially longer run times. This behaviour is also illustrated in Figures 6, 7, and 8, where the KL divergence values remain similar across models (across rows) but vary greatly across datasets (across columns).

3.2 Cosmological Datasets

We analyse a comprehensive set of cosmological observations spanning multiple redshift ranges and probing different physical phenomena. CMB observations probe the early uni-

verse at $z \approx 1100$, while late-universe probes including baryon acoustic oscillations, Type Ia supernovae, and weak gravitational lensing constrain the expansion history and structure formation across cosmic time. To constrain the model parameters, we perform MCMC and nested sampling runs using the Cobaya framework [34], which interfaces with the PolyChord sampler [26, 27] and the CAMB Boltzmann code [35]. A full list of the likelihood packages used for this analysis is provided in Table 1.

3.2.1 Planck

This dataset comprises high-precision measurements of CMB temperature and polarisation anisotropies from the surface of last scattering ($z \approx 1100$) using the *Plik* high-ℓ likelihood [55, 67]. Our analysis utilises four likelihood components: low-ℓ temperature and E-mode polarisation (`planck_2018_lowl.TT`, `planck_2018_lowl.EE`) covering $\ell = 2$–29, the high-ℓ TTTEEE likelihood (`planck_2018_highl_plik.TTTEEE`) spanning $\ell = 30$–2508, and a Sunyaev-Zel'dovich (SZ) foreground prior (`planck_2018_highl_plik.SZ`). Temperature fluctuations arise from acoustic oscillations in the primordial photon-baryon plasma [68, 69], whilst polarisation E-modes trace Thomson scattering during the recombination and reionisation epochs. This dataset provides strong constraints on the fundamental cosmological parameters: the baryon density $\Omega_b h^2$, cold dark matter density $\Omega_c h^2$, Hubble parameter H_0, primordial amplitude A_s, spectral index n_s, and reionisation optical depth τ.

3.2.2 Planck with CMB lensing

This dataset combines the Planck 2018 CMB measurements with the CMB lensing reconstruction (`planck_2018_lensing.clik`) [55, 56, 67]. The lensing likelihood is added to the four baseline Planck components (low-ℓ TT and EE, high-ℓ TTTEEE, and SZ foreground). This combination provides enhanced constraints by breaking geometric degeneracies and improving measurements of the matter density Ω_m and the clustering amplitude σ_8 through an independent probe of large-scale structure growth.

3.2.3 CamSpec

This dataset represents an alternative high-ℓ analysis of Planck 2018 data, using the CamSpec 2021 likelihood (`planck_2018_highl_CamSpec2021.TTTEEE`) combined with the same low-ℓ likelihoods as the baseline Planck analysis [57]. CamSpec employs distinct foreground modelling and power spectrum estimation compared to the official *Plik* pipeline, including different approaches to dust cleaning and the treatment of systematics. It provides an independent systematic cross-check, which is particularly valuable for assessing the robustness of cosmological parameter constraints to the choice of analysis methodology.

3.2.4 CamSpec with CMB lensing

This dataset combines the CamSpec CMB analysis (low-ℓ TT and EE plus high-ℓ CamSpec2021 TTTEEE) with the CMB lensing reconstruction (`planck_2018_lensing.clik`) [56, 57]. It provides an independent systematic cross-check with enhanced parameter constraints from lensing, and is particularly valuable for assessing whether tensions in Ω_m and σ_8 persist across different CMB analysis pipelines.

3.2.5 CMB Lensing

This dataset is the standalone measurement of the lensing potential power spectrum, derived from the gravitational deflection of CMB photons by intervening large-scale structure [56].

Dataset	Likelihood
Cosmic Microwave Background	
Planck [54, 55]	`planck_2018_lowl.TT`
	`planck_2018_lowl.EE`
	`planck_2018_highl_plik.TTTEEE`
	`planck_2018_highl_plik.SZ`
Planck with CMB lensing [54–56]	`planck_2018_lowl.TT`
	`planck_2018_lowl.EE`
	`planck_2018_highl_plik.TTTEEE`
	`planck_2018_highl_plik.SZ`
	`planck_2018_lensing.clik`
CamSpec [57]	`planck_2018_lowl.TT`
	`planck_2018_lowl.EE`
	`planck_2018_highl_CamSpec2021.TTTEEE`
CamSpec with CMB lensing [56, 57]	`planck_2018_lowl.TT`
	`planck_2018_lowl.EE`
	`planck_2018_highl_CamSpec2021.TTTEEE`
	`planck_2018_lensing.clik`
CMB Lensing [56]	`planck_2018_lensing.clik`
BICEP [58, 59]	`bicep_keck_2018`
Baryon Acoustic Oscillations	
SDSS [60–63]	`bao.sixdf_2011_bao`
	`bao.sdss_dr7_mgs`
	`bao.sdss_dr16_baoplus_lrg`
	`bao.sdss_dr16_baoplus_elg`
	`bao.sdss_dr16_baoplus_qso`
	`bao.sdss_dr16_baoplus_lyauto`
	`bao.sdss_dr16_baoplus_lyxqso`
Type Ia Supernovae	
SH$_0$ES [64, 65]	`H0.riess2020Mb`
	`sn.pantheon`
Pantheon [65]	`sn.pantheon`
Weak Lensing	
DES [66]	`des_y1.joint`

Table 1: Cosmological datasets and their corresponding likelihood components used in the analysis. Datasets are grouped by observational type with references to the actual data packages and implementation repositories used. Likelihood names correspond to those used by `Cobaya`.

The lensing reconstruction (`planck_2018_lensing.clik`) uses quadratic estimators [70] on Planck temperature and polarisation maps to extract the lensing convergence signal. It probes the matter distribution and structure growth over cosmic history (primarily $z \sim 0.5$–5), breaking geometric degeneracies and enhancing constraints on Ω_m, σ_8, and the sum of neutrino masses $\sum m_\nu$. We run CMB lensing alone to enable tension quantification analysis on the rest of the CMB data with and without lensing.

3.2.6 BICEP

This dataset consists of degree-scale B-mode polarisation measurements from the BICEP/Keck Array 2018 data release (`bicep_keck_2018`), which searches for primordial gravitational waves from inflation [58, 59]. Observations at 95, 150, and 220 GHz from the South Pole target the cleanest sky region with the lowest Galactic foreground contamination. B-modes can originate from tensor perturbations (inflationary gravitational waves) or from the weak gravitational lensing of E-modes, which acts as a foreground in this search. The BK18-only constraints on the tensor-to-scalar ratio are $r_{0.05} < 0.06$ (95% CL), directly probing the inflationary energy scale through the relation $V^{1/4} \propto r^{1/4}$ [5].

3.2.7 SDSS

This dataset is a compilation of baryon acoustic oscillation (BAO) measurements from seven independent surveys spanning $z = 0.1$ to $z > 2$ [60–63]. We combine measurements from: 6dFGS (`bao.sixdf_2011_bao`, $z = 0.106$), SDSS DR7 MGS (`bao.sdss_dr7_mgs`, $z = 0.15$), and five eBOSS DR16 tracers covering $z = 0.698$–2.33 using luminous red galaxies (LRG), emission line galaxies (ELG), quasars (QSO), plus the Lyman-α forest auto-correlation and its cross-correlation with quasars at $z > 2$. The BAO feature represents the imprint of primordial sound waves at recombination [71–73], with a characteristic sound horizon scale of $r_{\rm drag} \approx 147$ Mpc for Planck-like ΛCDM cosmologies [67]. This provides a standard ruler that measures both the angular diameter distance $D_A(z)$ and the Hubble parameter $H(z)$ as functions of redshift, thereby constraining Ω_m, H_0, and dark energy dynamics.

3.2.8 SH$_0$ES

This dataset provides a Gaussian prior on the Type Ia supernova absolute magnitude M_b, derived from the local distance ladder [64, 65, 74]. We implement this through the `H0.riess2020Mb` likelihood, which sets $M_b = -19.253 \pm 0.027$ mag based on HST observations of Cepheid variable stars in SN Ia host galaxies with three geometric anchors (Milky Way parallaxes, the Large Magellanic Cloud, and the NGC 4258 water maser). This M_b prior is used alongside the Pantheon SN Ia dataset (`sn.pantheon` with `use_abs_mag: true`) to derive an H_0 value of 73.2 ± 1.3 km s^{-1} Mpc^{-1}, anchoring the local expansion rate. This result creates a tension of approximately 4σ with the CMB-inferred value of $H_0 \approx 67$ km s^{-1} Mpc^{-1} within the ΛCDM model, motivating searches for new physics or systematic effects.

3.2.9 Pantheon

The Pantheon sample is a compilation of 1048 spectroscopically confirmed Type Ia supernovae from Pan-STARRS1 (PS1), SDSS, SNLS, low-z surveys, and HST (`sn.pantheon`) [65, 74]. It covers the redshift range $0.01 < z < 2.3$ with standardised peak magnitudes corrected for light-curve shape and colour using the SALT2 fitter. The luminosity distance-redshift relation $d_L(z)$ directly probes the expansion history $H(z) = H_0 E(z)$ through the integral $d_L(z) = c(1+z) \int_0^z dz'/H(z')$, providing evidence for cosmic acceleration [37, 38] at $z \sim 0.5$

and constraining Ω_m and the dark energy equation of state w. When analysed without an external H_0 calibration, this dataset constrains the degenerate product $H_0 \times M_b$ rather than absolute distances.

3.2.10 DES

This dataset is the Dark Energy Survey Year 1 (DES Y1) "3×2pt" analysis (`des_y1.joint`) [66, 75], which combines three two-point correlation functions: cosmic shear (the weak lensing auto-correlation of background galaxies at $z \sim 0.2$–1.3), galaxy clustering (the angular auto-correlation of foreground lens galaxies in five tomographic redshift bins), and galaxy-galaxy lensing (the cross-correlation between lens positions and source shears). This joint analysis, spanning 1321 deg^2, probes both the expansion history through geometric effects and structure growth through gravitational lensing. It primarily constrains the matter density Ω_m and the clustering amplitude through the parameter combination $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$. The DES Y1 3×2pt analysis finds $S_8 = 0.773 \pm 0.026$, showing a mild tension of approximately 2σ with the higher value of $S_8 \approx 0.83$ inferred from Planck.

3.3 Tools

This work utilises a modified version of `Cobaya 3.5.2`[5] [33, 34] for sampling and modelling framework, which interfaces likelihoods from different datasets with the Boltzmann code `CAMB 1.4.2.1` [35, 76, 77]. The specific likelihoods used for each dataset are listed in Table 1. `PolyChord 1.22.1`[6] was used as a nested sampling tool to explore parameter spaces and generate posterior samples with 1000 live points. Subsequent analysis, including plots and tension statistic computation, was performed using `anesthetic`[7] and `unimpeded`.

4 `unimpeded` in action

`unimpeded` is a Python-based tool designed to streamline access to pre-computed cosmological chains and facilitate Bayesian analyses. This section outlines its installation, available data, and basic usage.

4.1 Installation

The Python library `unimpeded` is publicly available on GitHub. To ensure a clean installation and avoid conflicts with other packages, we highly recommend creating and activating a dedicated Python virtual environment before proceeding.

```
python -m venv venv
source venv/bin/activate
```

The simplest method is to install the latest stable release from the Python Package Index (PyPI) using `pip`:

```
pip install unimpeded
```

Alternatively, for users interested in modifying the source code or contributing to development, an editable version can be installed directly from the GitHub repository:

---

[5]https://github.com/AdamOrmondroyd/cobaya
[6]https://github.com/PolyChord/PolyChordLite
[7]https://github.com/handley-lab/anesthetic

```
git clone https://github.com/handley-lab/unimpeded
cd unimpeded
pip install -e .
```

The full source code, along with further documentation and examples, is hosted at the GitHub repository: https://github.com/handley-lab/unimpeded. If user wish to use unimpeded's tension statistics calculator (see Section 4.4), chain analysis functionality and visualisation tools (see Section 4.5), anesthetic [78] needs to be installed in the same virtual environment venv where unimpeded is installed. This can be done via pip:

```
pip install anesthetic
```

Alternatively, an editable version of anesthetic can be installed from the GitHub repository:

```
git clone https://github.com/handley-lab/anesthetic.git
cd anesthetic
pip install -e .
```

4.2   Available models and datasets

unimpeded provides access to a growing grid of both nested sampling chains and MCMC chains generated using Cobaya [34], with nested sampling performed by PolyChord [26, 27]. unimpeded currently covers 8 cosmological models, detailed in Section 2.5, and their prior ranges are summarised in Table 2. unimpeded currently presents 10 datasets and their pairwise combinations, detailed in Section 3.2. The likelihood(s) used by Cobaya for each dataset's nested sampling and MCMC runs are listed in Table 1. These chains are stored on Zenodo in csv format and are accessible directly through the unimpeded API (see Section 4.3), or from the Zenodo website.

4.3   Loading chains and information

The primary interface for accessing the pre-computed results is the DatabaseExplorer class in unimpeded.database. It provides a programmatic workflow for downloading nested sampling and MCMC chains and their associated metadata. This example demonstrates the standard user workflow in python. The process begins by instantiating the DatabaseExplorer, which lists the available content through its .models and .datasets attributes. Subsequently, both nested sampling (ns) and MCMC (mcmc) chains, along with their corresponding metadata files, are downloadable for a selected model-dataset combination. The code shows an example of downloading the nested sampling ('ns') chains for the $\Omega_k\Lambda$CDM model ('klcdm) constrained by the DES and CamSpec with CMB lensing joint-dataset (''des_y1.joint+planck_2018_CamSpec''). Correspondence between cosmological models and datasets and their unimpeded input strings are provided in Table 3 and Table 4, respectively. The call to dbe.download_samples returns a samples object containing the full posterior samples and prior samples, including their parameter values and importance weights. Complementarily, dbe.download_info retrieves the info object, which is a yaml file containing the complete run settings used by Cobaya and PolyChord for the analysis. samples and are immediately ready for analysis with tools like anesthetic [78].

Model	Parameter	Prior range	Definition
ΛCDM	H_0	[20, 100]	Hubble constant
	τ_{reio}	[0.01, 0.8]	Optical depth to reionization
	$\Omega_b h^2$	[0.005, 0.1]	Baryon density parameter
	$\Omega_c h^2$	[0.001, 0.99]	Cold dark matter density parameter
	$\log(10^{10} A_s)$	[1.61, 3.91]	Amplitude of scalar perturbations
	n_s	[0.8, 1.2]	Scalar spectral index
$\Omega_k\Lambda$CDM	Ω_k	[-0.3, 0.3]	Curvature density parameter (varying curvature)
wCDM	w	[-3, -0.333]	Constant dark energy equation of state
$w_0 w_a$CDM	w_0	[-3, 1]	Present-day dark energy equation of state
	w_a	[-3, 2]	Dark energy equation of state evolution (CPL parameterisation)
$m_\nu\Lambda$CDM	Σm_ν	[0.06, 2]	Sum of neutrino masses (eV)
$A_L\Lambda$CDM	A_L	[0, 10]	Lensing amplitude parameter
$n_{\mathrm{run}}\Lambda$CDM	n_{run}	[-1, 1]	Running of spectral index ($dn_s/d\ln k$)
$r\Lambda$CDM	r	[0, 3]	Scalar-to-tensor ratio

Table 2: Cosmological parameters for the models analysed in this work. The baseline ΛCDM model contains six fundamental parameters, with extensions adding additional parameters to test specific physical hypotheses. Prior ranges are specified based on theoretical constraints and observational bounds.

```
from unimpeded.database import DatabaseExplorer

# Initialise DatabaseExplorer
dbe = DatabaseExplorer()

# Get a list of currently available models and datasets
models_list = dbe.models
datasets_list = dbe.datasets

# Choose model, dataset and sampling method
method = 'ns' # 'ns' for nested sampling, 'mcmc' for MCMC
model = ''klcdm'' # from models_list
dataset = ''des_y1.joint+planck_2018_CamSpec'' # from datasets_list

# Download samples chain
samples = dbe.download_samples(method, model, dataset)

# Download Cobaya and PolyChord run settings
info = dbe.download_info(method, model, dataset)
```

4.4 Tension Statistics Calculator

To perform a tension analysis between two datasets, D_A and D_B, one must first run three separate nested sampling analyses to obtain the chains for: (1) D_A alone, (2) D_B alone, and (3) the joint dataset D_{AB}. These full nested sampling runs across a collection of models and

Model	unimpeded **Input**
ΛCDM	``lcdm``
$\Omega_k\Lambda$CDM	``klcdm``
wCDM	``wlcdm``
$w_0 w_a$CDM	``walcdm``
$A_L\Lambda$CDM	``Alcdm``
$m_\nu\Lambda$CDM	``mlcdm``
$n_{\rm run}\Lambda$CDM	``nrunlcdm``
$r\Lambda$CDM	``rlcdm``

Table 3: Correspondence between cosmological models described in Section 2.5 and their `unimpeded` input strings.

Dataset	unimpeded **Input**
Planck	``planck_2018_plik_nolens``
Planck with CMB lensing	``planck_2018_plik``
CamSpec	``planck_2018_CamSpec_nolens``
CamSpec with CMB lensing	``planck_2018_CamSpec``
CMB Lensing	``planck_2018_lensing``
BICEP	``bicep_keck_2018``
SDSS	``bao.sdss_dr16``
SH$_0$ES	``H0.riess2020Mb``
Pantheon	``sn.pantheon``
DES	``des_y1.joint``

Table 4: Correspondence between cosmological datasets described in Table 1 and their `unimpeded` input strings.

datasets took months to complete on a high performance computer, but `unimpeded` enables users to access these chains in seconds on a laptop, with only 2 lines of code demonstrated in this minimal working example. Please note that this functionality requires `anesthetic` to be installed in the same Python environment as `unimpeded` (see Section 4.1).

```
from unimpeded.tension import tension_calculator

tension_samples = tension_calculator(method='ns',
                                      model='lcdm',
                                      datasetA='planck_2018_CamSpec',
                                      datasetB='des_y1.joint',
                                      nsamples=1000)
```

The output of the `tension_calculator()` is an `anesthetic.samples.Samples` data structure containing the values for the tension statistic detailed in Section 2.4, which directly correspond to the theoretical quantities defined previously:

- **R statistics**: The function calculates 'logR' as $\log \mathcal{Z}_{AB} - \log \mathcal{Z}_A - \log \mathcal{Z}_B$, matching the definition of the logarithmic R statistic (Section 2.4.2).

- **Information Ratio**: 'I' is computed as $\mathcal{D}_{\rm KL}^A + \mathcal{D}_{\rm KL}^B - \mathcal{D}_{\rm KL}^{AB}$, as defined in Section 2.4.3.

- **Suspiciousness**: 'logS' is calculated as $\langle \log \mathcal{L} \rangle_{\mathcal{P}_{AB}} - \langle \log \mathcal{L} \rangle_{\mathcal{P}_A} - \langle \log \mathcal{L} \rangle_{\mathcal{P}_B}$, corresponding to the practical computational form of suspiciousness from Equation (2.21).

- **Bayesian Model Dimensionality**: The dimensionality of the shared parameter space, 'd_G', is computed as $d_A + d_B - d_{AB}$, as defined in Section 2.4.5.

- **Tension Probability and Significance**: The function uses 'd_G' and 'logS' to compute the p-value ('p') and its equivalent Gaussian significance ('tension') in units of σ, as described in Section 2.4.6.

This automated calculation provides a consistent and reproducible method for applying the full suite of Bayesian tension metrics across the large grid of datasets and models provided by `unimpeded`.

4.5 Analysing chains with `anesthetic`

The nested sampling chains generated by `unimpeded` are readily processed and analysed using the `anesthetic` package [78]. This package provides both quantitative statistical measures and powerful visualisation tools. A key function is `NestedSamples.stats()`, which computes summary statistics essential for cosmological inference, including the evidence `logZ`, which is essential for model comparison (see Section 2.3) and `logL_P` (the posterior-averaged log-likelihood $\langle \ln \mathcal{L} \rangle_P$). In addition to these quantitative diagnostics, `anesthetic` can be used to generate corner plots to visualise the one- and two-dimensional marginalised distributions. This functionality is particularly useful as it can plot the distributions for both the posterior and the prior samples, allowing for a direct visual assessment of the information gain for each parameter.

4.6 Future functions

Future work on `unimpeded` will focus on expanding its capabilities in three primary directions. First, we plan to develop machine learning emulators for both likelihoods and full posterior distributions. Trained on the extensive set of nested sampling chains generated by `unimpeded`, these emulators will facilitate extremely rapid parameter estimation and model exploration. Furthermore, this infrastructure will enable the application of simulation-based inference (SBI) methodologies, which are essential for cosmological analyses where the likelihood function is intractable. Second, the pre-computed grid of cosmological background and perturbation quantities will be expanded. This crucial update will incorporate the latest astronomical data from current and future surveys, including but not limited to ACT, Pantheon+, DESI, Euclid, and LISA, ensuring that `unimpeded` remains relevant for modern cosmology. Finally, we will implement importance sampling. This feature will provide a computationally inexpensive method for re-weighting existing posterior samples to account for different model assumptions, thereby significantly accelerating the process of updating cosmological constraints.

5 Results

5.1 Public Release of Chains via `unimpeded`

We have publically released a library of nested sampling and MCMC chains, across 8 cosmological models, 10 datasets and 31 pairwise dataset combinations, as detailed in section 4.2.

These chains are accessible via the `unimpeded` Python package and are stored on Zenodo, ensuring permanent public access and citable DOIs for specific data releases. This fulfills a key objective of our DiRAC-funded projects (DP192 and 264) by providing a community resource analogous to, but extending the capabilities of, the Planck Legacy Archive.

5.2 Parameter Estimation

The `unimpeded` framework enables efficient parameter estimation by providing direct access to pre-computed full nested sampling chains and MCMC chains. This allows users to by-pass the computationally intensive step of generating these chains themselves, facilitating rapid and robust cosmological inference. We demonstrate this capability by constraining the parameters of the $\Omega_k\Lambda$CDM model using a combination of SH0ES and DES Year 1 data[8]. Figure 2 illustrates how the posterior (orange) for certain parameters are significantly more constrained compared to the broad prior (blue). The diagonal plots show the marginalised posterior probability distribution for that specific parameter. For parameters well-constrained by these datasets, such as $\log(10^{10}A_{\rm s})$, $n_{\rm s}$, and Ω_k, the posteriors appear as narrow peaks demonstrating substantial information gain. However, for parameters that these particular datasets do not strongly constrain, such as $\tau_{\rm reio}$, the posteriors remain broad and similar to the priors. This visualisation was created using `anesthetic`.

5.3 Model Comparison

The Bayesian evidence values computed via nested sampling from `unimpeded` form the basis for rigorous model comparison. Here, we present a systematic comparison of eight cosmological models using both individual and combined datasets, as outlined in Section 4.2. Since we used uniform priors for the set of competing models, $P(\mathcal{M}_i) = $ constant, the posterior probability of a model given the data D, $P(\mathcal{M}_i|D) = \mathcal{Z}_i/\sum_j \mathcal{Z}_j$ (Equation (2.13)), provides a self-contained, normalised probability distribution over the models, allowing for a direct and intuitive ranking of their relative support from the data. Since the posterior probabilities can span many orders of magnitude, we present the natural logarithm of $\log P(\mathcal{M}_i|D)$, where higher values (i.e., less negative) indicate stronger evidence in favour of a given model.

The results of our model comparison are summarised in Figures 3 to 5. Figure 3 presents a heatmap of $\log P(\mathcal{M}_i|D)$ for each of the eight models tested against 10 individual datasets. The colour scale indicates the level of support, with bluer colours corresponding to higher $P(\mathcal{M}_i|D)$ and redder colour indicating that a model is more disfavoured relative to the others. To structure the visualisation, the datasets (y-axis) are sorted by descending constraining power (model-posterior-weighted average $\langle \mathcal{D}_{\rm KL} \rangle$), whilst the models ($x$-axis) are sorted by their descending $\mathcal{D}_{\rm KL}$ values from the Planck with CMB lensing dataset. Figures 4 and 5 show a similar analysis but for various combinations of datasets, designed to leverage their complementary constraining power.

One should note that in Equation (2.13), the sum of evidences is taken over all models being compared for a specific dataset or combination of datasets. Therefore, the numerical values of $\log P(\mathcal{M}_i|D)$ are only comparable horizontally across models for a fixed dataset, and not vertically across datasets for a fixed model. To enable recovery of the raw log-evidence values, the final column (in yellow) of each heatmap displays the normalising factor $\log\left(\sum_j \mathcal{Z}_j\right)$, which is the logarithm of the denominator of Equation (2.13). The raw log-

---

[8]The specific `unimpeded` input for this dataset combination is `H0.riess2020Mb+des_y1.joint` (see Table 1).
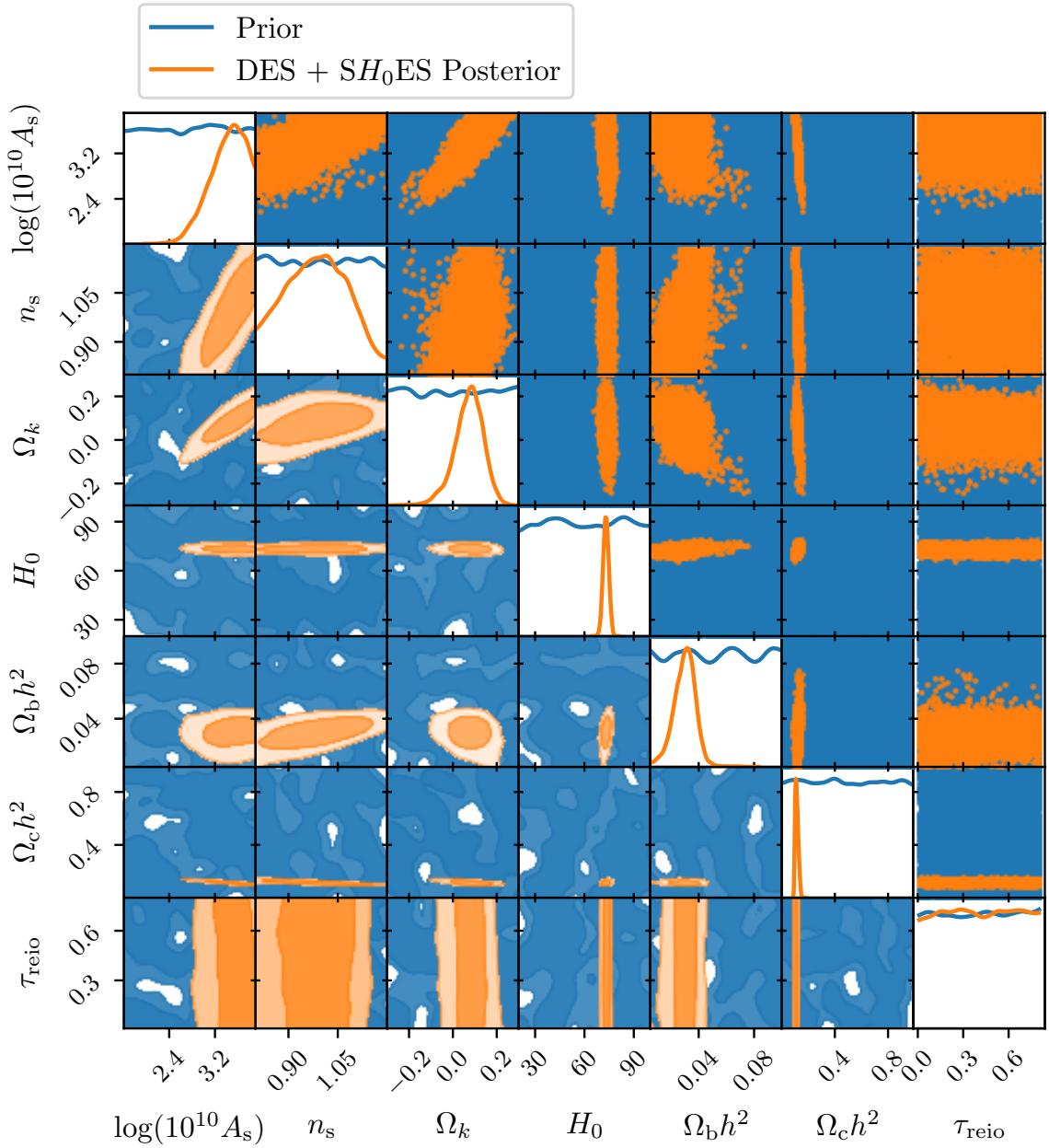
Figure 2: Corner plot showing posterior distributions for the $\Omega_k\Lambda\text{CDM}$ cosmological model constrained by Planck with CMB lensing + SDSS data. The diagonal panels show the one-dimensional marginalised prior (blue) and Planck with CMB lensing + SDSS posterior (orange) distributions, demonstrating the constraining power of the observational data. The lower triangular panels display the two-dimensional joint posterior and prior, where the inner (darker blue and darker orange) and outer (lighter blue and lighter orange) contours correspond to the 68% (1σ) and 95% (2σ) credible regions, respectively. The upper triangular panels show scatter plots of samples drawn from the posterior, visually representing parameter correlations. The posterior volume (orange) is much smaller than the prior volume (blue). This corner plot was created using `anesthetic`.

evidence for any specific model-dataset combination can be recovered by multiplying the probability value in that cell by the normalising factor of that row.

The model preference exhibits a dependence on the specific dataset being considered. As shown in Figure 3, an analysis of individual datasets reveals a diversity in the preferred cosmological model. No single model is universally favoured. Instead, different probes indicate a weak preference for different extensions to the base ΛCDM model. For instance, the SDSS dataset weakly prefers the $A_L\Lambda$CDM model, whilst DES weakly prefers a non-flat universe ($\Omega_k\Lambda$CDM). The Planck primary and CamSpec datasets both weakly prefer the wCDM model, characterised by a constant but non-standard dark energy equation of state. Other datasets show weak preferences for a running spectral index ($n_{\rm run}\Lambda$CDM for Pantheon) or massive neutrinos ($m_\nu\Lambda$CDM for SH0ES), whilst BICEP weakly prefers the base ΛCDM model itself. Notably, the ΛCDM model, though not always exhibiting the highest $\log {\rm P}(\mathcal{M}_i|D)$, emerges as the most consistently well-performing model amongst the eight models considered.

This picture changes when datasets are combined, as illustrated in Figures 4 and 5. In the combined analyses, the base ΛCDM model is most often the preferred scenario. We emphasise that this analysis involves comparing model performance horizontally within each dataset row; due to differences in data normalisation, a vertical comparison of log-evidence values across different datasets for a fixed model is not meaningful.

5.4 Constraining Power of Models and Datasets

The Kullback-Leibler divergence $\mathcal{D}_{\rm KL}$ quantifies the information gain from prior to posterior after taking into account the data, providing a measure of how much the data constrains each model-dataset combination (see Section 2.1.2 for details in theory). Figure 6 presents a heatmap of $\mathcal{D}_{\rm KL}$ for each of the eight models tested against 10 individual datasets. Figures 7 and 8 show a similar analysis but for various combinations of datasets. Higher values of $\mathcal{D}_{\rm KL}$ indicate that the dataset provides stronger constraints on the model parameters, representing greater information gain from the prior to the posterior. The datasets (y-axis) are sorted in descending order by the model posterior ${\rm P}(\mathcal{M}_i|D)$-weighted average $\mathcal{D}_{\rm KL}$ (Equation (2.28)), thereby ranking each dataset by its constraining power, weighted by the posterior probability of each model. Similarly, the models along the x-axis are sorted in descending order according to their $\mathcal{D}_{\rm KL}$ values for the Planck with CMB lensing dataset.

5.5 Tension Quantification

To systematically quantify the consistency between the various cosmological datasets employed in this work, we utilised the tension statistics calculator implemented in the `unimpeded` package and the nested sampling chains it offers. We performed a comprehensive tension analysis across 31 pairwise dataset combinations for each of the 8 cosmological models under consideration. For each pair, we computed the tension statistics discussed in section 2.4, including p-value significance σ (Figure 9), Bayesian Model Dimensionality (d_G) (Figure 10), Information Ratio (Figure 11), R statistic (Figure 12), and Suspiciousness (Figure 13), providing a comprehensive view of the statistical agreement between datasets. For each of the tension statistics, the dataset (y-axis) is ranked ascendingly or descendingly using its model posterior ${\rm P}(\mathcal{M}_i|D)$ weighted average, as stated by Equation (2.27).

The results of this extensive analysis are summarised in Figure 9, which presents the tension significance expressed as the equivalent Gaussian sigma (σ) of the p-value (see Section 2.4 for the theory and equations). A crucial feature of this representation is that the

	wCDM	$w_0 w_a$CDM	ΛCDM	$\Omega_k\Lambda$CDM	$A_L\Lambda$CDM	$m_\nu\Lambda$CDM	$n_{\rm run}\Lambda$CDM	$r\Lambda$CDM	$\sum_j \log\mathcal{Z}_j$
Planck	-0.97 ± 0.14	-1.50 ± 0.19	-3.11 ± 0.21	-1.26 ± 0.17	-3.07 ± 0.21	-5.44 ± 0.21	-7.07 ± 0.22	-6.67 ± 0.22	-1438.67 ± 0.10
Planck with CMB lensing	-0.84 ± 0.13	-0.98 ± 0.15	-1.89 ± 0.20	-4.46 ± 0.23	-4.74 ± 0.23	-4.93 ± 0.23	-6.19 ± 0.22	-6.35 ± 0.23	-1443.91 ± 0.12
CamSpec with CMB lensing	-0.90 ± 0.14	-1.08 ± 0.15	-1.65 ± 0.17	-3.72 ± 0.22	-4.62 ± 0.21	-4.73 ± 0.21	-6.25 ± 0.22	-5.11 ± 0.21	-5321.37 ± 0.10
CamSpec	-0.97 ± 0.13	-1.23 ± 0.15	-1.97 ± 0.19	-2.08 ± 0.19	-3.14 ± 0.20	-5.02 ± 0.22	-6.25 ± 0.20	-5.77 ± 0.21	-5316.35 ± 0.10
DES	-2.99 ± 0.14	-3.38 ± 0.15	-1.84 ± 0.12	-1.35 ± 0.11	-1.80 ± 0.13	-2.22 ± 0.14	-2.94 ± 0.14	-1.82 ± 0.12	-277.51 ± 0.06
BICEP	-1.98 ± 0.08	-2.23 ± 0.09	-1.70 ± 0.08	-1.81 ± 0.08	-1.74 ± 0.07	-1.96 ± 0.08	-2.46 ± 0.09	-5.77 ± 0.11	-280.12 ± 0.03
SH_0ES	-3.33 ± 0.09	-4.06 ± 0.09	-1.83 ± 0.08	-2.10 ± 0.08	-1.95 ± 0.08	-1.60 ± 0.07	-1.91 ± 0.08	-1.78 ± 0.08	-527.39 ± 0.03
CMB Lensing	-2.10 ± 0.07	-2.39 ± 0.08	-1.63 ± 0.07	-2.23 ± 0.08	-2.05 ± 0.07	-1.81 ± 0.07	-3.73 ± 0.09	-1.81 ± 0.07	-11.15 ± 0.03
SDSS	-1.93 ± 0.07	-3.22 ± 0.08	-1.76 ± 0.06	-2.92 ± 0.08	-1.74 ± 0.06	-2.40 ± 0.07	-1.84 ± 0.06	-1.82 ± 0.06	-16.45 ± 0.03
Pantheon	-3.32 ± 0.05	-3.99 ± 0.06	-1.83 ± 0.04	-2.08 ± 0.04	-1.78 ± 0.04	-1.86 ± 0.04	-1.76 ± 0.04	-1.82 ± 0.04	-519.52 ± 0.02

Figure 3: Heatmap of the log-posterior model probabilities, $\log \mathrm{P}(\mathcal{M}_i|D)$, for each cosmological model (x-axis) evaluated against individual dataset (y-axis). Bluer colours indicate stronger statistical support for a model given the data. Comparison should only be made horizontally across models for a fixed dataset, as the sum of evidences in Equation (2.13) is taken over all models for that specific dataset. The final column (in yellow) shows the normalising factor $\log\left(\sum_j \mathcal{Z}_j\right)$, the logarithm of the denominator of Equation (2.13). The raw log-evidence for any model-dataset combination can be recovered by multiplying the probability value in that cell by the normalising factor of that row. The results show that while different datasets favour different model extensions, the base ΛCDM model emerges as the most consistently well-performing model across all individual datasets (overall blue).

	wCDM	w_0w_aCDM	ΛCDM	$\Omega_k\Lambda$CDM	$A_L\Lambda$CDM	$m_\nu\Lambda$CDM	$n_{run}\Lambda$CDM	$r\Lambda$CDM	$\sum_j \log \mathcal{Z}_j$
DES + Planck	-0.65 ±0.13	-1.12 ±0.19	-2.67 ±0.25	-7.03 ±0.27	-2.72 ±0.25	-6.46 ±0.26	-8.21 ±0.26	-7.97 ±0.27	-1714.48 ±0.14
DES + Planck with CMB lensing	-0.39 ±0.10	-1.26 ±0.22	-4.65 ±0.27	-8.10 ±0.28	-4.13 ±0.28	-7.70 ±0.28	-8.68 ±0.28	-8.10 ±0.28	-1718.58 ±0.17
DES + CamSpec with CMB lensing	-0.58 ±0.12	-0.97 ±0.17	-3.41 ±0.25	-6.47 ±0.26	-4.34 ±0.25	-6.48 ±0.27	-7.24 ±0.27	-6.61 ±0.27	-5596.10 ±0.15
DES + CamSpec	-0.46 ±0.09	-1.65 ±0.22	-3.13 ±0.25	-5.82 ±0.26	-2.19 ±0.25	-6.19 ±0.27	-7.25 ±0.27	-6.51 ±0.27	-5591.36 ±0.15
BICEP + Planck	-1.62 ±0.20	-1.36 ±0.18	-3.31 ±0.22	-0.83 ±0.14	-2.92 ±0.23	-6.14 ±0.25	-7.37 ±0.24	-7.03 ±0.24	-1716.79 ±0.12
SH_0ES + Planck	-3.72 ±0.25	-1.69 ±0.21	-2.06 ±0.22	-4.57 ±0.25	-0.46 ±0.09	-4.98 ±0.24	-6.66 ±0.25	-5.63 ±0.25	-1967.21 ±0.13
BICEP + Planck with CMB lensing	-0.28 ±0.06	-2.03 ±0.24	-2.40 ±0.24	-5.01 ±0.27	-5.52 ±0.26	-6.42 ±0.27	-7.10 ±0.28	-6.86 ±0.27	-1722.16 ±0.17
SH_0ES + Planck with CMB lensing	-2.18 ±0.22	-1.33 ±0.18	-0.78 ±0.13	-3.61 ±0.22	-2.54 ±0.21	-3.94 ±0.24	-5.30 ±0.24	-3.94 ±0.23	-1973.02 ±0.11
BICEP + CamSpec with CMB lensing	-0.55 ±0.10	-1.44 ±0.19	-1.92 ±0.21	-4.34 ±0.24	-5.31 ±0.24	-5.48 ±0.23	-6.36 ±0.24	-6.35 ±0.24	-5599.52 ±0.13
BICEP + CamSpec	-1.36 ±0.17	-1.09 ±0.15	-1.87 ±0.19	-1.65 ±0.18	-3.25 ±0.22	-5.51 ±0.23	-6.89 ±0.23	-6.26 ±0.23	-5594.55 ±0.10
SH_0ES + CamSpec	-2.49 ±0.21	-1.61 ±0.18	-1.16 ±0.16	-3.04 ±0.22	-1.16 ±0.16	-3.96 ±0.22	-5.73 ±0.22	-5.09 ±0.22	-5844.75 ±0.10
SH_0ES + CamSpec with CMB lensing	-2.24 ±0.21	-1.35 ±0.17	-0.75 ±0.12	-3.39 ±0.23	-2.65 ±0.22	-3.80 ±0.22	-5.36 ±0.23	-4.12 ±0.23	-5849.94 ±0.11
Planck + Pantheon	-4.28 ±0.24	-4.80 ±0.24	-0.93 ±0.16	-4.59 ±0.24	-0.62 ±0.12	-4.27 ±0.23	-5.75 ±0.24	-4.79 ±0.24	-1957.85 ±0.13
SDSS + Planck	-3.84 ±0.23	-3.19 ±0.23	-0.67 ±0.12	-5.64 ±0.23	-0.95 ±0.16	-4.23 ±0.22	-5.53 ±0.23	-5.46 ±0.23	-1450.65 ±0.12
SDSS + Planck with CMB lensing	-2.09 ±0.21	-3.05 ±0.24	-0.32 ±0.06	-4.71 ±0.24	-3.42 ±0.24	-3.37 ±0.24	-4.63 ±0.23	-4.38 ±0.23	-1456.32 ±0.14
CMB Lensing + Planck	-0.84 ±0.13	-0.98 ±0.15	-1.89 ±0.20	-4.46 ±0.23	-4.74 ±0.23	-4.93 ±0.23	-6.19 ±0.22	-6.35 ±0.23	-1443.91 ±0.12

Color bar label: $\log P(\mathcal{M}_i|D)$

Figure 4: Same as Figure 3, but for combinations of datasets. The final column (in yellow) shows the normalising factor $\log\left(\sum_j \mathcal{Z}_j\right)$, the logarithm of the denominator of Equation (2.13), allowing recovery of raw log-evidence values by multiplying the probability in each cell by the normalising factor of that row. The combination of multiple probes sharpens the model comparison, further strengthening the preference for ΛCDM and increasing the degree to which extended models are disfavoured. Part 1 of combined datasets.
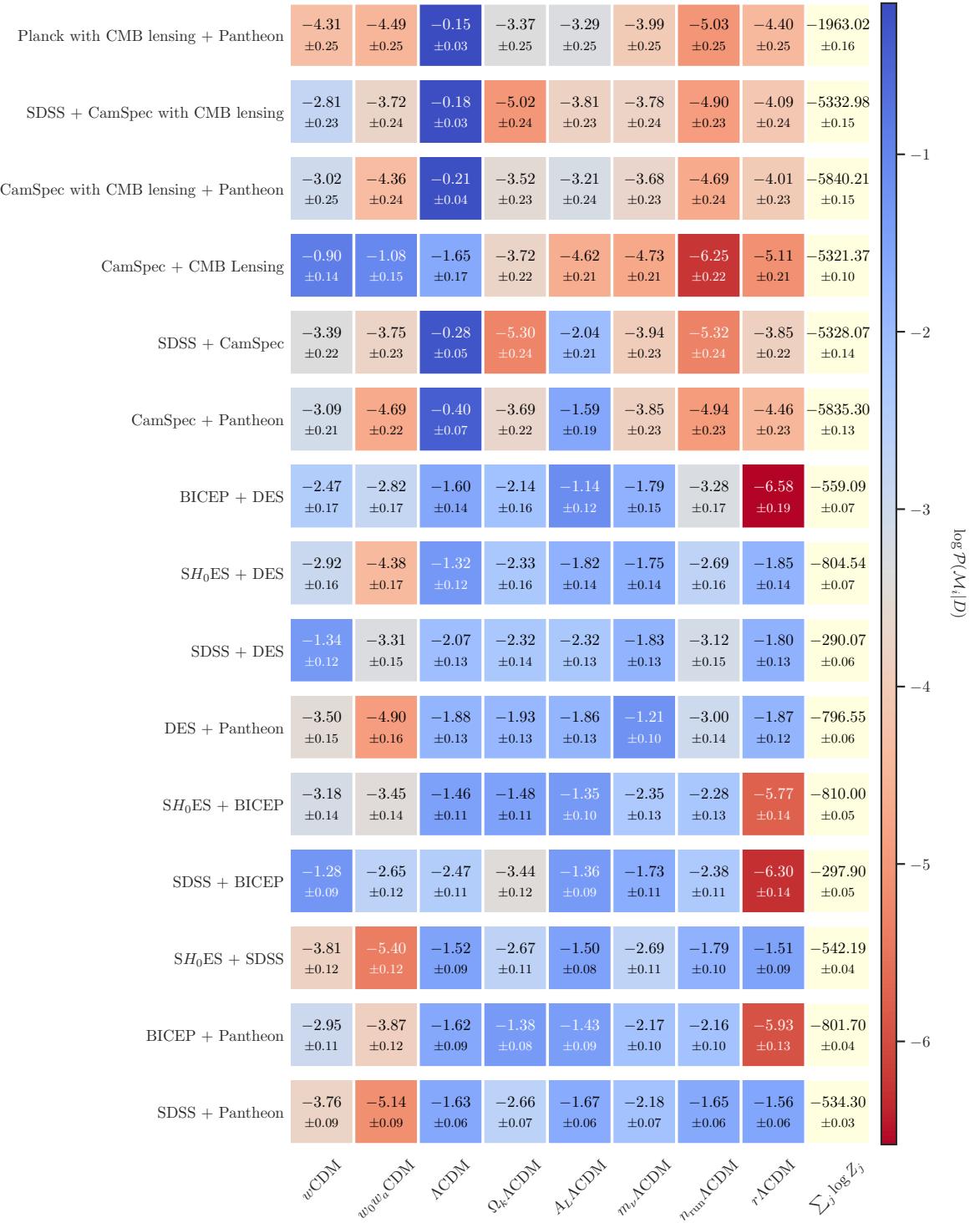
Figure 5: Same as Figure 3, but for combinations of datasets. The final column (in yellow) shows the normalising factor $\log\left(\sum_j \mathcal{Z}_j\right)$, the logarithm of the denominator of Equation (2.13), allowing recovery of raw log-evidence values by multiplying the probability in each cell by the normalising factor of that row. The combination of multiple probes sharpens the model comparison, further strengthening the preference for ΛCDM and increasing the degree to which extended models are disfavoured. Part 2 of combined datasets.

	$n_{run}\Lambda$CDM	$A_L\Lambda$CDM	$r\Lambda$CDM	$\Omega_k\Lambda$CDM	$m_\nu\Lambda$CDM	w_0w_aCDM	wCDM	ΛCDM
Planck	52.69 ±0.20	52.55 ±0.19	51.79 ±0.19	51.48 ±0.19	51.03 ±0.19	49.10 ±0.19	48.55 ±0.19	49.16 ±0.19
Planck with CMB lensing	52.48 ±0.19	52.47 ±0.20	52.23 ±0.20	52.12 ±0.20	51.16 ±0.19	49.31 ±0.18	49.15 ±0.18	48.49 ±0.18
CamSpec with CMB lensing	48.98 ±0.19	48.43 ±0.19	47.70 ±0.18	47.52 ±0.19	47.61 ±0.19	45.73 ±0.18	45.44 ±0.17	44.88 ±0.18
CamSpec	48.59 ±0.19	47.43 ±0.19	47.69 ±0.19	47.05 ±0.18	47.45 ±0.18	45.11 ±0.18	44.81 ±0.18	44.62 ±0.18
DES	24.31 ±0.13	23.60 ±0.13	23.61 ±0.13	23.62 ±0.13	23.93 ±0.13	25.37 ±0.13	25.19 ±0.14	23.48 ±0.13
BICEP	11.30 ±0.08	10.52 ±0.08	14.21 ±0.09	10.69 ±0.08	10.76 ±0.08	11.04 ±0.08	10.78 ±0.08	10.57 ±0.08
SH_0ES	10.41 ±0.08	10.44 ±0.08	10.33 ±0.08	10.37 ±0.08	10.13 ±0.08	12.05 ±0.09	11.38 ±0.09	10.35 ±0.08
CMB Lensing	11.00 ±0.08	9.26 ±0.07	9.14 ±0.07	9.70 ±0.07	8.89 ±0.07	9.73 ±0.08	9.37 ±0.07	9.00 ±0.08
SDSS	8.31 ±0.06	8.26 ±0.06	8.31 ±0.06	9.23 ±0.06	8.99 ±0.06	11.18 ±0.08	10.55 ±0.07	8.23 ±0.06
Pantheon	3.43 ±0.04	3.42 ±0.04	3.46 ±0.04	3.46 ±0.04	3.50 ±0.04	5.12 ±0.05	4.51 ±0.05	3.46 ±0.04

Figure 6: This heatmap illustrates the Kullback-Leibler divergence ($\mathcal{D}_{KL}$) for each dataset (y-axis) and model (x-axis) combination, with higher values (bluer colours) indicating a greater overall constraint. Datasets are sorted vertically by their model-posterior-weighted average $\langle \mathcal{D}_{KL} \rangle_{P(\mathcal{M})}$ (Equation (2.28)), while models are sorted horizontally by their $\mathcal{D}_{KL}$ from the Planck with CMB lensing dataset. A prominent feature is the strong vertical gradient, showing that $\mathcal{D}_{KL}$ varies significantly among datasets but remains relatively constant across models for a given dataset. This indicates that the information gain is predominantly determined by the statistical power of the observational probe, with more constraining, information-rich datasets naturally yielding higher $\mathcal{D}_{KL}$ values.
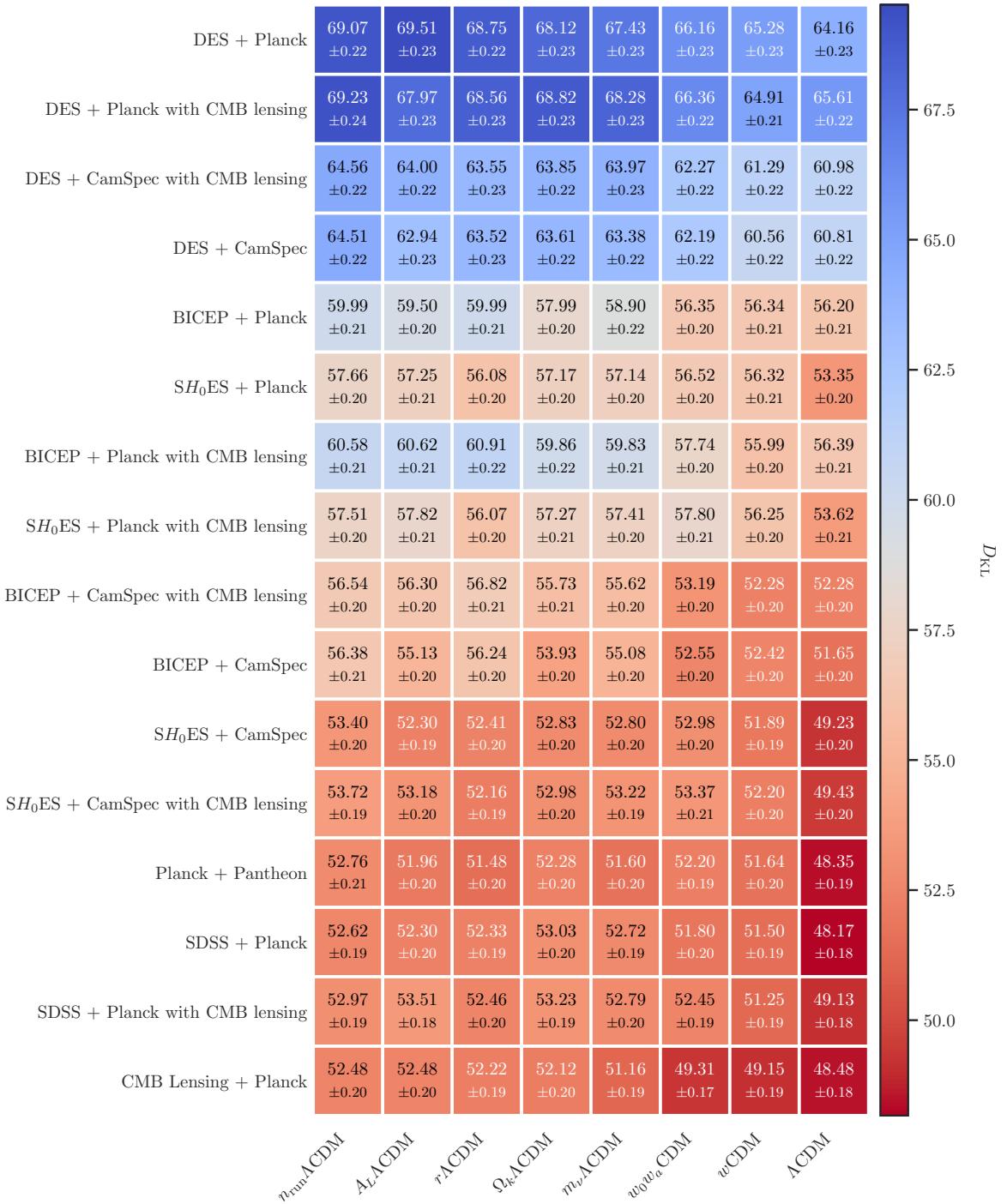
Figure 7: Same as Figure 6, but for combinations of datasets (y-axis). Combined datasets yield substantially higher $\mathcal{D}_{\mathrm{KL}}$ values compared to individual datasets, reflecting the enhanced constraining power from multiple complementary observational probes. The strong vertical gradient persists, with $\mathcal{D}_{\mathrm{KL}}$ varying significantly among dataset combinations but remaining relatively constant across models for a given combination. Part 1 of combined datasets.

	$n_{run}\Lambda\mathrm{CDM}$	$A_L\Lambda\mathrm{CDM}$	$r\Lambda\mathrm{CDM}$	$\Omega_k\Lambda\mathrm{CDM}$	$m_\nu\Lambda\mathrm{CDM}$	$w_0w_a\mathrm{CDM}$	$w\mathrm{CDM}$	$\Lambda\mathrm{CDM}$
Planck with CMB lensing + Pantheon	53.01 ±0.19	52.81 ±0.20	51.92 ±0.20	51.43 ±0.20	52.16 ±0.18	52.61 ±0.19	52.34 ±0.20	48.28 ±0.18
SDSS + CamSpec with CMB lensing	49.20 ±0.18	49.09 ±0.19	48.02 ±0.18	49.42 ±0.19	49.05 ±0.18	48.99 ±0.18	47.75 ±0.18	44.97 ±0.18
CamSpec with CMB lensing + Pantheon	48.70 ±0.18	48.48 ±0.18	47.76 ±0.18	47.95 ±0.19	48.02 ±0.19	48.51 ±0.19	47.54 ±0.18	44.70 ±0.17
CamSpec + CMB Lensing	48.98 ±0.18	48.44 ±0.18	47.70 ±0.19	47.53 ±0.18	47.61 ±0.18	45.75 ±0.18	45.44 ±0.18	44.88 ±0.18
SDSS + CamSpec	49.16 ±0.17	47.93 ±0.19	47.27 ±0.18	49.22 ±0.19	48.78 ±0.19	48.46 ±0.18	47.66 ±0.19	44.43 ±0.18
CamSpec + Pantheon	48.70 ±0.18	47.44 ±0.19	47.88 ±0.19	47.37 ±0.19	47.65 ±0.19	48.35 ±0.18	47.10 ±0.19	44.34 ±0.18
BICEP + DES	35.32 ±0.16	33.10 ±0.16	38.75 ±0.17	34.27 ±0.16	34.30 ±0.16	35.59 ±0.16	35.73 ±0.17	33.88 ±0.17
SH_0ES + DES	32.06 ±0.15	30.79 ±0.15	30.81 ±0.15	32.03 ±0.15	31.62 ±0.15	33.68 ±0.16	32.54 ±0.16	30.31 ±0.15
SDSS + DES	27.82 ±0.14	26.39 ±0.14	25.99 ±0.13	27.74 ±0.14	27.41 ±0.13	30.54 ±0.14	29.07 ±0.15	26.10 ±0.13
DES + Pantheon	25.47 ±0.13	23.80 ±0.13	23.94 ±0.13	25.13 ±0.13	24.11 ±0.13	27.48 ±0.14	26.42 ±0.14	23.73 ±0.13
SH_0ES + BICEP	22.32 ±0.12	21.18 ±0.12	25.65 ±0.14	21.59 ±0.12	22.17 ±0.12	22.83 ±0.13	22.70 ±0.13	21.46 ±0.13
SDSS + BICEP	19.07 ±0.10	18.31 ±0.11	23.02 ±0.12	19.85 ±0.11	18.83 ±0.11	21.37 ±0.11	20.50 ±0.11	19.18 ±0.11
SH_0ES + SDSS	15.62 ±0.09	15.37 ±0.09	15.37 ±0.09	16.28 ±0.10	16.68 ±0.10	19.70 ±0.11	17.97 ±0.10	15.34 ±0.09
BICEP + Pantheon	14.92 ±0.10	14.16 ±0.09	18.58 ±0.11	14.22 ±0.10	14.98 ±0.10	16.07 ±0.10	15.35 ±0.10	14.48 ±0.10
SDSS + Pantheon	8.63 ±0.06	8.62 ±0.06	8.52 ±0.06	9.39 ±0.06	9.27 ±0.07	12.71 ±0.08	10.99 ±0.08	8.57 ±0.06

D_{KL}

Figure 8: Same as Figure 6, but for combinations of datasets (y-axis). Combined datasets yield substantially higher $\mathcal{D}_{\mathrm{KL}}$ values compared to individual datasets, reflecting the enhanced constraining power from multiple complementary observational probes. The strong vertical gradient persists, with $\mathcal{D}_{\mathrm{KL}}$ varying significantly among dataset combinations but remaining relatively constant across models for a given combination. Part 2 of combined datasets.

numerical σ values are directly comparable both across rows and down columns, unlike the model comparison heatmap in Section 5.3.

Each of the five heatmaps is sorted to bring the most concerning dataset combinations to the top, providing an immediate visual guide to potential tensions. For the p-value significance in Figure 9, rows are sorted in descending order of their average σ across all models. The subsequent heatmaps for the $\log R$, $\log I$, and $\log S$ statistics (Figures 11 to 13) are sorted in ascending order to place the most negative values—indicating strong tension—at the top. The Bayesian model dimensionality in Figure 10 is also sorted in ascending order. We employ red highlighting to flag values that cross specific thresholds of concern. For the p-value, we highlight $\sigma > 2.88$, a threshold calculated by Equation (2.26) in Section 2.4.7 that accounts for the look-elsewhere effect across our 248 analyses (8 models $\times$ 31 dataset pairs). This threshold is not arbitrary, if there were no genuine tensions, we would expect exactly one result to reach $\sigma = 2.88$ purely by chance. Figure 9 shows 14 dataset-model combinations with $\sigma > 2.88$, significantly more than the single false positive expected under the null hypothesis. Rather than correcting individual p-values or σ values (which would change as the grid expands), we apply this threshold, ensuring that σ values remain directly interpretable independently of the grid size. For the other statistics, red flags indicate $\log R < 0$ and $\log S < 0$, signalling dataset inconsistency and direct likelihood conflict, respectively.

An analysis of the p-value significance in Figure 9 immediately identifies the well-known tensions in cosmology. The comparisons of DES vs Planck ($\sigma = 3.57$ in ΛCDM) and SH0ES vs Planck ($\sigma = 3.27$ in ΛCDM) exhibit the highest significance, exceeding our $\sigma > 2.88$ threshold. Other comparisons involving these datasets, such as DES vs CamSpec and SH0ES vs CamSpec, also show tension in ΛCDM ($\sigma = 3.19$ and $\sigma = 2.84$, respectively). The results demonstrate model dependence; for instance, the DES vs Planck tension is alleviated in all model extensions (e.g., dropping to $\sigma = 1.90$ in wCDM), whereas the SH0ES vs Planck tension remains above 2σ in most models, only falling below this mark for the $w_0 w_a$CDM and $\Omega_k\Lambda$CDM models. This suggests that the physics introduced in the extended models is more effective at resolving the S_8 tension than the Hubble tension. We note that the DES vs Planck tension is expected to relax when we extend the grid to include the DES Year 3 (Y3) data release.

The suite of five statistics provides a far more nuanced picture than the p-value alone, revealing crucial differences in the nature of these tensions. The rankings of the most problematic dataset pairs are broadly consistent across the p-value, Information Ratio (Figure 11), and Suspiciousness (Figure 13) heatmaps. For example, DES vs Planck and SH0ES vs Planck comparisons populate the top rows of all three, showing highly negative values for $\log I$ and $\log S$ (e.g., for DES vs Planck in ΛCDM, $\log I = -3.36$ and $\log S = -4.67$), confirming a genuine conflict between their likelihoods and minimal posterior overlap. However, a stark disagreement emerges when comparing these to the R statistic (Figure 12). For SH0ES vs Planck, while $\log S$ is strongly negative (-4.19 in ΛCDM), $\log R$ is positive ($+1.19$), indicating concordance. This discrepancy arises because the Suspiciousness is prior-independent, whereas the R statistic is not. The positive $\log R$ signifies that despite the likelihood conflict, the combined posterior is still substantially more constraining than the prior, a common feature in high-dimensional parameter spaces. This highlights the value of using the prior-independent Suspiciousness to isolate direct data conflict.

A multi-metric analysis allows a deeper physical interpretation of the tensions. The Hubble tension (SH0ES vs CMB comparisons) is characterised by high σ, negative $\log I$ and $\log S$, but a low Bayesian dimensionality (e.g., $d_G = 1.36$ for SH0ES vs Planck in ΛCDM,

see Figure 10). This confirms that the conflict is sharp but concentrated in a very small number of parameter dimensions, principally H_0. In stark contrast, the S_8 tension (DES vs CMB comparisons) appears as a more systemic disagreement. For DES vs Planck in ΛCDM, not only are σ, $\log I$ and $\log S$ all indicative of tension, but the dimensionality is very high ($d_G = 6.62$). This indicates that the datasets disagree across a wide range of parameter directions, representing a more fundamental inconsistency within the ΛCDM framework. The fact that this high-dimensional tension is largely resolved in extended models reinforces the interpretation that it may be a signature of new physics.

In summary, this comprehensive five-statistic analysis provides a detailed and robust characterisation of the consistency landscape. We find that relying on a single metric like the p-value can be misleading. The combined view confirms that the DES vs CMB and SH0ES vs CMB tensions are the most significant statistical conflicts in the data, but their natures are profoundly different. The S_8 tension is a high-dimensional problem that is effectively resolved by allowing for new physical degrees of freedom, whereas the Hubble tension is a sharp, low-dimensional conflict that persists across models and is only flagged as a severe issue by prior-independent metrics like Suspiciousness. This nuanced understanding, gained by synthesising information from multiple complementary statistics, is crucial for guiding future model building and determining which dataset combinations can be reliably used for joint cosmological analyses. Caution should be exercised when combining datasets in tension. Conversely, pairs at the bottom of the rankings, such as BICEP vs Pantheon, show excellent agreement across all five metrics ($\sigma \approx 0$, $\log R > 0$) and can be combined with confidence.

Our findings are consistent with the curvature tension analysis of [3], which reported similar moderate tensions between Planck 2018 and CMB lensing ($\sigma = 2.49 \pm 0.07$) and between Planck 2018 and BAO ($\sigma = 3.03 \pm 0.06$) in the context of curved $\Omega_K\Lambda$CDM cosmologies. However, our model comparison results show lower Bayes factors (1.85 log units for $\Omega_k\Lambda$CDM vs ΛCDM compared to 4 log units in that work), which can be attributed to the deliberately wider priors adopted in our analysis using the Cobaya defaults. These wider priors provide greater flexibility for importance reweighting if tighter priors are desired in future analyses. Whilst that work focused specifically on the curvature parameter Ω_K, our systematic analysis across eight model extensions and 31 dataset pairs provides a broader view of the tension landscape, demonstrating that the methodology is robust and the tensions persist across multiple cosmological frameworks.

Dataset pair	$\Omega_k\Lambda$CDM	$A_L\Lambda$CDM	w_0w_aCDM	ΛCDM	wCDM	$r\Lambda$CDM	$n_{\rm run}\Lambda$CDM	$m_\nu\Lambda$CDM
SH_0ES vs CamSpec with CMB lensing	2.47 ±0.11	2.25 ±0.11	2.23 ±0.25	2.78 ±0.12	2.59 ±0.12	2.77 ±0.12	2.57 ±0.10	2.45 ±0.13
SH_0ES vs Planck with CMB lensing	2.90 ±0.22	2.45 ±0.24	1.78 ±0.15	2.78 ±0.17	2.68 ±0.14	2.84 ±0.20	2.94 ±0.19	2.26 ±0.13
SH_0ES vs Planck	3.27 ±0.10	2.29 ±0.24	2.22 ±0.26	2.97 ±0.18	3.08 ±0.24	2.96 ±0.20	2.98 ±0.22	2.27 ±0.13
SH_0ES vs CamSpec	2.84 ±0.13	2.24 ±0.12	1.88 ±0.17	2.58 ±0.10	2.52 ±0.12	2.87 ±0.17	2.94 ±0.17	2.32 ±0.11
DES vs Planck	3.57 ±0.10	1.90 ±0.11	1.94 ±0.11	2.44 ±0.11	2.12 ±0.11	2.65 ±0.16	3.10 ±0.32	2.64 ±0.16
DES vs CamSpec with CMB lensing	2.93 ±0.12	2.06 ±0.10	1.91 ±0.12	2.61 ±0.13	2.11 ±0.12	3.00 ±0.22	2.55 ±0.15	2.68 ±0.17
DES vs CamSpec	3.19 ±0.12	1.90 ±0.09	1.90 ±0.09	2.39 ±0.11	2.01 ±0.09	2.69 ±0.17	2.50 ±0.15	2.62 ±0.14
DES vs Planck with CMB lensing	3.19 ±0.15	1.86 ±0.09	1.70 ±0.09	2.57 ±0.14	1.92 ±0.10	2.79 ±0.21	2.84 ±0.23	2.50 ±0.13
SH_0ES vs DES	1.38 ±0.19	1.50 ±0.12	1.17 ±0.16	1.29 ±0.08	0.96 ±0.06	1.52 ±0.11	1.17 ±0.24	0.91 ±0.09
CMB Lensing vs Planck	2.28 ±0.08	1.94 ±0.07	1.29 ±0.10	1.26 ±0.08	1.22 ±0.09	1.18 ±0.08	1.16 ±0.07	1.09 ±0.07
SDSS vs Planck with CMB lensing	1.56 ±0.10	0.94 ±0.08	2.09 ±0.10	1.01 ±0.08	2.44 ±0.12	1.15 ±0.20	1.05 ±0.07	0.44 ±0.08
DES vs Pantheon	1.05 ±0.32	1.59 ±0.13	1.29 ±0.55	1.41 ±0.08	0.85 ±0.07	1.48 ±0.11	1.10 ±0.19	0.86 ±0.07
CamSpec vs CMB Lensing	2.02 ±0.09	1.57 ±0.08	1.10 ±0.07	1.14 ±0.06	1.11 ±0.07	1.15 ±0.07	1.27 ±0.08	1.10 ±0.07
SDSS vs Planck	2.46 ±0.08	0.95 ±0.08	2.04 ±0.12	1.15 ±0.08	2.74 ±0.15	1.03 ±0.09	1.18 ±0.10	0.46 ±0.05
SDSS vs CamSpec	2.01 ±0.09	1.03 ±0.09	2.00 ±0.10	0.98 ±0.07	2.30 ±0.08	0.98 ±0.06	1.04 ±0.08	0.56 ±0.05
SDSS vs CamSpec with CMB lensing	1.23 ±0.06	1.00 ±0.07	1.87 ±0.08	0.92 ±0.07	2.27 ±0.08	1.02 ±0.07	0.88 ±0.06	0.48 ±0.05
BICEP vs DES	1.43 ±0.25	1.17 ±0.28	0.40 ±0.19	0.67 ±0.10	0.31 ±0.11	0.49 ±0.14	0.32 ±0.28	0.50 ±0.10
CamSpec vs Pantheon	2.33 ±0.21	0.04 ±0.11	2.14 ±0.38	0.73 ±0.10	1.31 ±0.14	0.08 ±0.17	0.08 ±0.20	0.43 ±0.32
Planck vs Pantheon	2.44 ±0.12	0.37 ±0.18	1.86 ±0.32	0.91 ±0.25	1.89 ±0.30	0.56 ±0.30	0.88 ±0.11	0.50 ±0.19
SDSS vs DES	0.43 ±0.06	1.09 ±0.07	0.29 ±0.05	0.93 ±0.06	0.49 ±0.04	0.96 ±0.06	0.32 ±0.09	0.30 ±0.04
CamSpec with CMB lensing vs Pantheon	1.01 ±0.14	0.46 ±0.14	1.90 ±0.30	0.50 ±0.24	1.27 ±0.11	0.77 ±0.22	0.56 ±0.27	0.12 ±0.19
BICEP vs Planck	0.82 ±0.08	0.25 ±0.33	0.07 ±0.14	1.11 ±0.40	0.11 ±0.23	0.05 ±0.06	0.79 ±0.30	0.54 ±0.24
BICEP vs Planck with CMB lensing	0.22 ±0.41	0.20 ±0.25	0.27 ±0.25	0.09 ±0.20	0.51 ±0.15	0.00 ±0.00	0.69 ±0.20	0.26 ±0.35
BICEP vs CamSpec with CMB lensing	0.37 ±0.11	0.40 ±0.27	0.40 ±0.36	0.37 ±0.39	0.29 ±0.28	0.09 ±0.06	0.34 ±0.17	0.33 ±0.26
Planck with CMB lensing vs Pantheon	1.61 ±0.31	0.09 ±0.17	1.82 ±0.29	0.18 ±0.29	1.82 ±0.28	0.57 ±0.11	0.56 ±0.12	0.07 ±0.12
BICEP vs CamSpec	0.56 ±0.20	0.00 ±0.03	0.04 ±0.09	0.62 ±0.39	0.07 ±0.18	0.02 ±0.03	0.67 ±0.29	0.66 ±0.22
SH_0ES vs SDSS	0.12 ±0.09	0.17 ±0.14	0.90 ±0.05	0.31 ±0.11	1.42 ±0.05	0.35 ±0.10	0.02 ±0.02	0.08 ±0.11
SDSS vs Pantheon	0.19 ±0.06	0.03 ±0.08	0.81 ±0.05	0.04 ±0.08	1.47 ±0.04	0.29 ±0.09	0.08 ±0.11	0.26 ±0.09
SDSS vs BICEP	1.31 ±0.24	0.00 ±0.00	0.00 ±0.00	0.24 ±0.49	0.00 ±0.04	0.01 ±0.06	0.19 ±0.46	0.00 ±0.02
SH_0ES vs BICEP	0.00 ±0.00	0.11 ±0.39	0.00 ±0.00	0.08 ±0.26	0.00 ±0.00	0.00 ±0.03	0.01 ±0.06	0.27 ±0.58
BICEP vs Pantheon	0.00 ±0.00	0.00 ±0.00	0.01 ±0.09	0.00 ±0.00	0.01 ±0.07	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00

Significance (p-value σ)

Figure 9: A heatmap quantifying the tension between 31 pairwise dataset combinations (y-axis) across 8 cosmological models (x-axis). The tension is expressed as the significance in equivalent Gaussian standard deviations (σ), derived from the p-value, allowing for direct comparison across the grid. The dataset pairs are sorted vertically in descending order of their average tension across all models (Equation (2.29)), placing the most discordant combinations at the top. Values with $\sigma > 2.88$, highlighted in red, exceed the significance threshold that accounts for the look-elsewhere effect across all 248 analyses performed. This threshold is defined such that if no genuine tensions existed, only one false positive would be expected by chance (see Section 2.4.7). We observe 14 such instances.
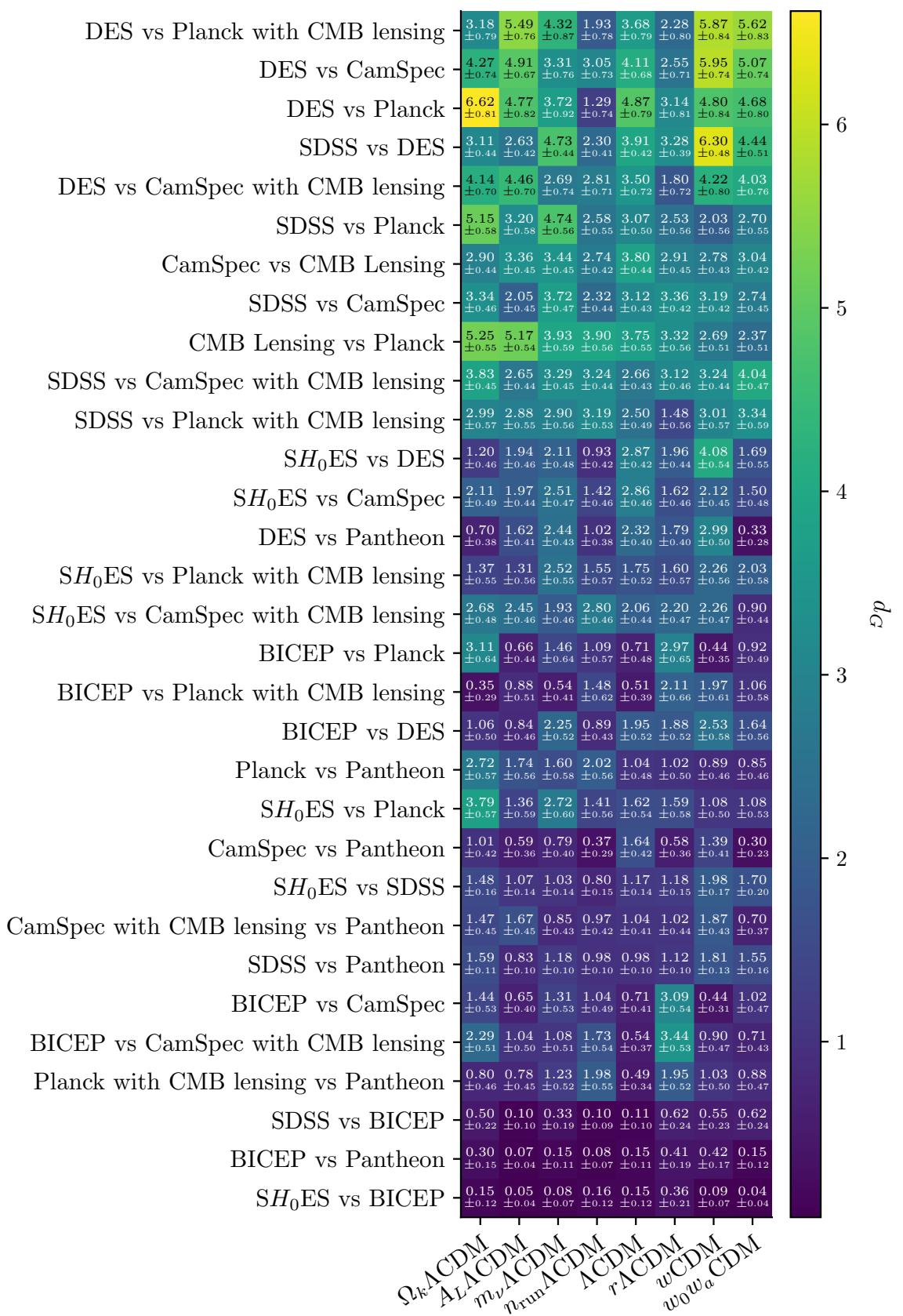
Figure 10: A heatmap quantifying the Bayesian Model Dimensionality (d_G) for 31 pairwise dataset combinations (y-axis) across 8 cosmological models (x-axis). d_G measures the effective number of constrained parameters in the shared parameter space of two datasets (see Section 2.4.5), allowing for direct comparison across the grid. The dataset pairs are sorted vertically in ascending order of their average dimensionality across all models. This metric distinguishes between sharp, low-dimensional conflicts and broader, systemic disagreements.

A heatmap of the Information Ratio (I) with $\log\mathcal{I}$ color scale. Rows are dataset pairs; columns are cosmological models ($n_{\rm run}\Lambda$CDM, ΛCDM, w_0w_aCDM, $A_L\Lambda$CDM, $\Omega_k\Lambda$CDM, wCDM, $m_\nu\Lambda$CDM, $r\Lambda$CDM).

| Dataset pair	$n_{\rm run}\Lambda$CDM	ΛCDM	w_0w_aCDM	$A_L\Lambda$CDM	$\Omega_k\Lambda$CDM	wCDM	$m_\nu\Lambda$CDM	$r\Lambda$CDM
SH_0ES vs BICEP	-0.59 ±0.17	-0.52 ±0.17	0.40 ±0.23	-0.11 ±0.15	-0.51 ±0.17	-0.43 ±0.16	-1.20 ±0.14	-1.10 ±0.19
BICEP vs Pantheon	-0.13 ±0.13	-0.43 ±0.13	0.12 ±0.14	-0.17 ±0.15	-0.07 ±0.13	-0.05 ±0.14	-0.69 ±0.13	-0.91 ±0.15
BICEP vs DES	0.29 ±0.23	0.16 ±0.22	0.82 ±0.23	1.03 ±0.22	0.04 ±0.23	0.24 ±0.23	0.39 ±0.22	-0.92 ±0.24
SDSS vs BICEP	0.58 ±0.15	-0.31 ±0.12	0.85 ±0.16	0.54 ±0.14	0.07 ±0.16	0.84 ±0.16	0.93 ±0.14	-0.51 ±0.17
DES vs Pantheon	2.25 ±0.19	3.21 ±0.19	3.06 ±0.21	3.24 ±0.19	1.96 ±0.19	3.29 ±0.20	3.32 ±0.19	3.15 ±0.19
SH_0ES vs DES	2.65 ±0.22	3.51 ±0.22	3.72 ±0.22	3.25 ±0.21	1.95 ±0.22	4.02 ±0.23	2.44 ±0.21	3.13 ±0.22
SDSS vs Pantheon	3.11 ±0.10	3.12 ±0.10	3.59 ±0.12	3.07 ±0.10	3.30 ±0.10	4.06 ±0.12	3.22 ±0.10	3.25 ±0.09
SH_0ES vs SDSS	3.11 ±0.14	3.24 ±0.14	3.53 ±0.15	3.34 ±0.14	3.33 ±0.14	3.96 ±0.15	2.45 ±0.14	3.26 ±0.13
BICEP vs CamSpec	3.51 ±0.28	3.53 ±0.26	3.60 ±0.29	2.83 ±0.28	3.82 ±0.28	3.18 ±0.27	3.11 ±0.29	5.64 ±0.30
CamSpec with CMB lensing vs Pantheon	3.70 ±0.26	3.63 ±0.26	2.35 ±0.26	3.38 ±0.27	3.01 ±0.28	2.42 ±0.26	3.09 ±0.27	3.40 ±0.26
CamSpec vs Pantheon	3.34 ±0.25	3.74 ±0.26	1.92 ±0.25	3.41 ±0.27	3.15 ±0.26	2.21 ±0.26	3.31 ±0.26	3.27 ±0.27
Planck with CMB lensing vs Pantheon	2.89 ±0.27	3.68 ±0.25	1.82 ±0.27	3.08 ±0.29	4.14 ±0.28	1.32 ±0.28	2.51 ±0.27	3.75 ±0.26
BICEP vs Planck with CMB lensing	3.20 ±0.31	2.70 ±0.30	2.63 ±0.29	2.37 ±0.30	2.98 ±0.29	3.95 ±0.28	2.09 ±0.31	5.53 ±0.31
BICEP vs CamSpec with CMB lensing	3.73 ±0.30	3.18 ±0.27	3.59 ±0.28	2.67 ±0.29	2.51 ±0.29	3.93 ±0.29	2.75 ±0.29	5.09 ±0.30
BICEP vs Planck	4.01 ±0.29	3.53 ±0.30	3.79 ±0.29	3.57 ±0.28	4.17 ±0.30	3.02 ±0.28	2.88 ±0.30	6.02 ±0.31
Planck vs Pantheon	3.37 ±0.28	4.27 ±0.26	2.05 ±0.28	4.01 ±0.27	2.66 ±0.28	1.42 ±0.26	2.94 ±0.28	3.77 ±0.28
SH_0ES vs Planck with CMB lensing	5.37 ±0.29	5.20 ±0.29	3.55 ±0.30	5.09 ±0.30	5.21 ±0.30	4.28 ±0.29	3.88 ±0.29	6.50 ±0.29
SH_0ES vs CamSpec	5.63 ±0.29	5.72 ±0.26	4.19 ±0.28	5.57 ±0.27	4.60 ±0.27	4.32 ±0.28	4.80 ±0.29	5.61 ±0.28
SH_0ES vs CamSpec with CMB lensing	5.66 ±0.28	5.79 ±0.28	4.40 ±0.28	5.70 ±0.28	4.92 ±0.27	4.61 ±0.28	4.51 ±0.28	5.87 ±0.28
SH_0ES vs Planck	5.45 ±0.29	6.17 ±0.29	4.65 ±0.30	5.72 ±0.29	4.69 ±0.30	3.61 ±0.29	4.01 ±0.30	6.01 ±0.30
SDSS vs DES	4.81 ±0.20	5.61 ±0.19	6.02 ±0.22	5.47 ±0.19	5.10 ±0.20	6.67 ±0.20	5.52 ±0.20	5.93 ±0.20
SDSS vs Planck with CMB lensing	7.82 ±0.29	7.57 ±0.27	8.04 ±0.29	7.23 ±0.28	8.12 ±0.29	8.43 ±0.27	7.34 ±0.30	8.08 ±0.29
SDSS vs CamSpec with CMB lensing	8.09 ±0.26	8.15 ±0.27	7.92 ±0.26	7.62 ±0.27	7.33 ±0.26	8.23 ±0.26	7.54 ±0.28	7.98 ±0.27
DES vs Planck	7.92 ±0.34	8.48 ±0.31	8.33 ±0.33	6.64 ±0.33	7.00 ±0.32	8.47 ±0.33	7.55 ±0.34	6.65 ±0.34
SDSS vs CamSpec	7.74 ±0.27	8.41 ±0.25	7.84 ±0.28	7.76 ±0.27	7.09 ±0.27	7.70 ±0.26	7.68 ±0.27	8.72 ±0.27
SDSS vs Planck	8.40 ±0.28	9.22 ±0.28	8.51 ±0.29	8.51 ±0.28	7.68 ±0.27	7.58 ±0.28	7.31 ±0.28	7.78 ±0.28
CamSpec vs CMB Lensing	10.61 ±0.27	8.73 ±0.26	9.12 ±0.27	8.24 ±0.27	9.24 ±0.26	8.75 ±0.27	8.74 ±0.27	9.12 ±0.27
DES vs CamSpec	8.40 ±0.32	7.25 ±0.31	8.29 ±0.32	8.10 ±0.31	7.06 ±0.32	9.44 ±0.30	7.98 ±0.31	7.77 ±0.32
DES vs Planck with CMB lensing	7.56 ±0.33	6.36 ±0.33	8.32 ±0.33	8.10 ±0.32	6.92 ±0.34	9.42 ±0.32	6.81 ±0.33	7.30 ±0.33
DES vs CamSpec with CMB lensing	8.73 ±0.32	7.38 ±0.31	8.86 ±0.32	8.03 ±0.32	7.27 ±0.31	9.33 ±0.32	7.56 ±0.32	7.74 ±0.32
CMB Lensing vs Planck	11.22 ±0.28	9.66 ±0.27	9.53 ±0.28	9.32 ±0.29	9.07 ±0.29	8.76 ±0.28	8.74 ±0.31	8.69 ±0.29

Figure 11: A heatmap quantifying the tension using the Information Ratio (I) for 31 pairwise dataset combinations (y-axis) across 8 cosmological models (x-axis). I quantifies tension by comparing the $\mathcal{D}_{\rm KL}$ of the combined posterior relative to the individual posteriors (see Section 2.4.3), allowing for direct comparison across the grid. The dataset pairs are sorted vertically in ascending order of their average I across all models, placing the combinations with the most negative I values, and thus the strongest tension, at the top. A negative I ($I < 0$) signifies that the volume of the combined posterior is substantially smaller than would be expected from statistically consistent datasets, pointing to minimal overlap between their individual parameter constraints. This metric therefore provides an intuitive, volume-based measure of statistical surprise.

Figure 12 — Heatmap of $\log \mathcal{R}$ (logarithmic R statistic) for 31 pairwise dataset combinations across 8 cosmological models.

Dataset pair	$n_{run}\Lambda$CDM	w_0w_aCDM	ΛCDM	$m_\nu\Lambda$CDM	wCDM	$r\Lambda$CDM	$A_L\Lambda$CDM	$\Omega_k\Lambda$CDM
SH_0ES vs BICEP	-0.34	0.57	-0.35	-1.12	-1.11	-0.69	0.02 ±0.15	0.02 ±0.17
BICEP vs Pantheon	0.12 ±0.14	0.35 ±0.14	-0.70	-0.36	0.30 ±0.14	-1.40	0.16 ±0.14	0.46 ±0.14
BICEP vs DES	0.67 ±0.24	1.33 ±0.23	0.47 ±0.23	0.92 ±0.22	1.04 ±0.23	-0.44 ±0.25	0.95 ±0.23	-0.42 ±0.25
SDSS vs BICEP	0.69 ±0.14	1.48 ±0.17	-0.31	1.33 ±0.15	1.33 ±0.17	-0.63	0.90 ±0.15	-0.65
SH_0ES vs Planck with CMB lensing	1.08 ±0.28	1.99 ±0.30	1.23 ±0.29	0.89 ±0.30	0.24 ±0.29	2.48 ±0.29	2.42 ±0.31	1.24 ±0.29
SH_0ES vs CamSpec with CMB lensing	1.64 ±0.28	2.62 ±0.28	1.56 ±0.28	1.33 ±0.29	0.81 ±0.29	1.58 ±0.29	2.76 ±0.28	1.26 ±0.28
SH_0ES vs CamSpec	1.45 ±0.29	2.67 ±0.29	1.64 ±0.27	1.66 ±0.30	0.82 ±0.29	1.45 ±0.29	2.94 ±0.28	0.15 ±0.28
SH_0ES vs DES	2.54 ±0.22	3.41 ±0.23	2.70 ±0.22	2.44 ±0.22	3.75 ±0.23	2.11 ±0.23	2.28 ±0.22	1.47 ±0.23
DES vs Pantheon	2.18 ±0.20	3.03 ±0.20	2.28 ±0.19	3.35 ±0.19	3.31 ±0.21	2.26 ±0.20	2.21 ±0.19	1.99 ±0.19
SH_0ES vs Planck	1.19 ±0.29	2.75 ±0.30	1.74 ±0.29	0.90 ±0.30	-0.58	1.65 ±0.30	3.41 ±0.30	-2.33 ±0.30
SDSS vs Pantheon	3.61 ±0.11	3.74 ±0.13	3.63 ±0.11	3.75 ±0.11	3.15 ±0.12	3.74 ±0.10	3.52 ±0.11	4.01 ±0.11
SH_0ES vs SDSS	3.63 ±0.15	3.56 ±0.16	3.73 ±0.15	2.98 ±0.15	3.10 ±0.16	3.75 ±0.14	3.85 ±0.15	4.02 ±0.16
CamSpec vs Pantheon	3.67 ±0.26	1.17 ±0.25	3.98 ±0.27	3.60 ±0.27	1.76 ±0.26	3.71 ±0.27	3.92 ±0.27	1.03 ±0.26
CamSpec with CMB lensing vs Pantheon	3.98 ±0.26	1.41 ±0.27	3.96 ±0.26	3.59 ±0.27	1.88 ±0.27	3.59 ±0.27	3.87 ±0.28	2.95 ±0.29
Planck with CMB lensing vs Pantheon	3.33 ±0.28	0.88 ±0.27	3.99 ±0.25	3.23 ±0.27	0.26 ±0.28	4.17 ±0.27	3.61 ±0.29	3.59 ±0.27
BICEP vs CamSpec	3.74 ±0.29	4.31 ±0.29	3.73 ±0.27	3.38 ±0.29	3.57 ±0.27	7.18 ±0.31	3.57 ±0.28	4.17 ±0.28
BICEP vs Planck	4.17 ±0.29	4.41 ±0.30	3.54 ±0.30	3.24 ±0.31	3.39 ±0.29	7.43 ±0.31	3.93 ±0.29	4.24 ±0.31
BICEP vs CamSpec with CMB lensing	4.33 ±0.30	3.87 ±0.28	3.43 ±0.27	3.17 ±0.29	4.32 ±0.29	6.50 ±0.30	3.04 ±0.29	3.19 ±0.29
BICEP vs Planck with CMB lensing	3.44 ±0.31	3.09 ±0.29	3.13 ±0.30	2.39 ±0.30	4.41 ±0.28	7.16 ±0.32	2.83 ±0.30	3.24 ±0.31
Planck vs Pantheon	3.41 ±0.29	1.08 ±0.28	4.36 ±0.26	3.35 ±0.28	0.34 ±0.27	4.04 ±0.29	4.57 ±0.28	-0.92 ±0.25
DES vs Planck	3.53 ±0.33	5.45 ±0.34	3.97 ±0.32	2.88 ±0.34	5.03 ±0.33	2.22 ±0.33	3.86 ±0.33	-2.71 ±0.33
DES vs CamSpec	4.46 ±0.32	5.44 ±0.33	3.18 ±0.30	3.54 ±0.31	6.03 ±0.31	3.54 ±0.32	5.27 ±0.32	0.10 ±0.32
DES vs CamSpec with CMB lensing	4.76 ±0.33	6.28 ±0.31	2.88 ±0.31	3.28 ±0.32	6.12 ±0.32	3.09 ±0.31	4.87 ±0.32	1.37 ±0.32
DES vs Planck with CMB lensing	3.34 ±0.33	5.92 ±0.32	1.93 ±0.33	2.32 ±0.33	6.30 ±0.33	2.94 ±0.33	5.24 ±0.33	0.59 ±0.33
SDSS vs DES	5.56 ±0.21	7.18 ±0.22	5.42 ±0.20	6.70 ±0.21	7.48 ±0.21	5.72 ±0.20	5.10 ±0.19	5.82 ±0.21
SDSS vs Planck with CMB lensing	7.45 ±0.29	5.19 ±0.29	7.35 ±0.28	8.02 ±0.30	4.70 ±0.28	7.85 ±0.29	7.09 ±0.29	6.71 ±0.30
SDSS vs CamSpec with CMB lensing	8.04 ±0.27	5.41 ±0.27	8.06 ±0.27	8.20 ±0.28	4.88 ±0.27	7.67 ±0.27	7.39 ±0.27	6.48 ±0.27
SDSS vs CamSpec	7.50 ±0.28	5.44 ±0.28	8.18 ±0.26	8.22 ±0.29	4.27 ±0.27	8.46 ±0.27	7.58 ±0.27	4.43 ±0.28
SDSS vs Planck	7.87 ±0.28	6.02 ±0.29	8.67 ±0.28	8.10 ±0.28	3.51 ±0.29	7.53 ±0.29	8.34 ±0.28	2.99 ±0.28
CamSpec vs CMB Lensing	9.87 ±0.28	8.67 ±0.28	8.09 ±0.27	8.24 ±0.28	8.31 ±0.27	8.60 ±0.27	6.70 ±0.28	6.72 ±0.27
CMB Lensing vs Planck	10.51 ±0.29	8.84 ±0.28	8.76 ±0.27	8.21 ±0.31	8.13 ±0.28	8.04 ±0.30	6.30 ±0.29	4.95 ±0.29

Figure 12: A heatmap quantifying inter-dataset consistency using the logarithmic R statistic ($\log R$) for 31 pairwise dataset combinations (y-axis) across 8 cosmological models (x-axis). The R statistic is a prior-dependent measure of consistency that compares the joint Bayesian evidence of two datasets to the product of their individual evidences (see Section 2.4.2), and is interpreted relative to unity. The dataset pairs are sorted vertically in ascending order of their average $\log R$ across all models, placing the most inconsistent combinations at the top. Values of $R > 1$ ($\log R > 0$) indicate concordance, where each dataset strengthens the probability of the other. Conversely, values with $\log R < 0$ ($R < 1$), highlighted in red, indicate inconsistency, signifying that the joint probability of the data is lower than would be expected if the datasets were independent under the assumed model.
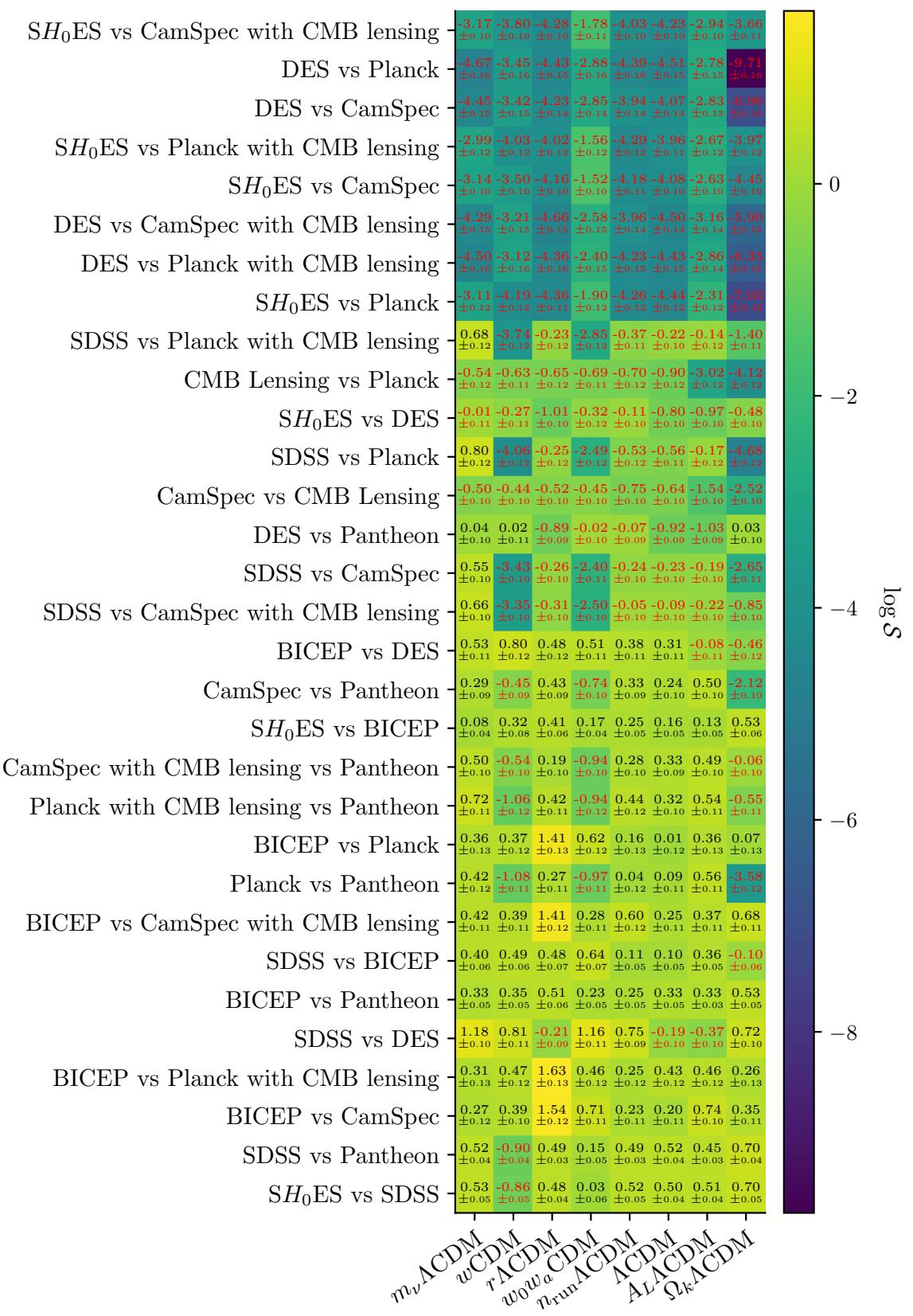
Figure 13: A heatmap quantifying tension using the logarithmic Suspiciousness ($\log S$) for 31 pairwise dataset combinations (y-axis) across 8 cosmological models (x-axis). S is a prior-independent metric that quantifies the statistical conflict between the likelihoods of two datasets (see Section 2.4.4), allowing for direct comparison across the grid. The dataset pairs are sorted vertically in ascending order of their average $\log S$ (Equation (2.27)) across all models, placing the combinations with the most negative $\log S$ values, and thus the strongest tensions, at the top. Values with $\log S \geq 0$ indicate agreement, while values with $\log S < 0$, highlighted in red, indicate tension, with more negative values signifying a stronger conflict between the datasets.

6 Conclusions

In this work, we have introduced `unimpeded`, a comprehensive and publicly available resource for Bayesian cosmological analysis. We have performed a systematic nested sampling analysis of eight cosmological models, from the base ΛCDM paradigm to seven well-motivated extensions, constrained by a suite of 39 individual and combined datasets. The primary data product of this analysis is an extensive repository of MCMC and nested sampling chains, hosted on Zenodo, which we provide to the community to facilitate reproducible and extensible cosmological research. The use of deliberately wide priors ensures that these chains are a versatile resource, suitable for importance reweighting and a wide range of future studies.

Our analysis yields two principal scientific conclusions. First, through a comprehensive model comparison, we find that whilst individual datasets show varied preferences for model extensions, the base ΛCDM model is most frequently preferred in combined analyses, with the general trend suggesting that evidence for new physics is diluted when probes are combined. This reinforces the predictive power and economy of the standard cosmological model. Second, by employing five complementary tension statistics, we systematically quantified the discordances between key datasets. We find the most significant tensions to be between SH0ES and Planck ($\sigma = 3.27$) and between DES(Y1) and Planck ($\sigma = 3.57$), within ΛCDM. Our multi-metric approach reveals that these tensions have profoundly different natures: the S_8 tension between DES and Planck is a high-dimensional disagreement ($d_G = 6.62$) that is mildly alleviated in models with a varying dark energy equation of state, whereas the Hubble tension between SH0ES and Planck is a sharp, low-dimensional conflict ($d_G = 1.36$) that persists across almost all model extensions considered.

The `unimpeded` resource provides a powerful platform for future investigations. The upgrade to DES Year 3 data is expected to clarify the status of the S_8 tension, and our framework provides the ideal foundation for a rapid and consistent analysis of this and other forthcoming datasets. Caution should be exercised when combining datasets in tension. By providing a standardised and accessible suite of Bayesian analysis products, we hope to accelerate progress in understanding the remaining tensions within the cosmological landscape and to robustly test the limits of the ΛCDM model.

Acknowledgments

References

[1] L. Verde, T. Treu and A.G. Riess, *Tensions between the early and late universe*, *Nature Astronomy* **3** (2019) 891 [1907.10625].

[2] S. Joudaki, C. Blake, A. Johnson, A. Amon, M. Asgari, A. Choi et al., *KiDS-450: testing extensions to the standard cosmological model*, *Mon. Not. Roy. Astron. Soc.* **471** (2017) 1259 [1701.05000].

[3] W. Handley, *Curvature tension: Evidence for a closed Universe*, *Phys. Rev. D* **103** (2021) L041301 [`1908.09139`].

[4] E. Di Valentino, A. Melchiorri and J. Silk, *Planck evidence for a closed Universe and a possible crisis for cosmology*, *Nature Astronomy* **4** (2020) 196 [`1911.02087`].

[5] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [`1807.06209`].

[6] J. Skilling, *Nested sampling for general bayesian computation*, *Bayesian Analysis* **1** (2006) 833.

[7] W. Handley and P. Lemos, *Quantifying the global parameter tensions between ACT, SPT, and Planck*, *Phys. Rev. D* **103** (2021) 063529 [`2007.08496`].

[8] W. Handley and P. Lemos, *Quantifying tensions in cosmological parameters: Interpreting the DES evidence ratio*, *Phys. Rev. D* **100** (2019) 043504 [`1902.04029`].

[9] P. Lemos, F. Köhlinger, W. Handley, B. Joachimi, L. Whiteway and O. Lahav, *Quantifying Suspiciousness within correlated data sets*, *Monthly Notices of the Royal Astronomical Society* **496** (2020) 4647 [`1910.07820`].

[10] R. Trotta, *Bayes in the sky: Bayesian inference and model selection in cosmology*, *Contemporary Physics* **49** (2008) 71 [`0803.4089`].

[11] S. Kullback and R.A. Leibler, *On information and sufficiency*, *The annals of mathematical statistics* **22** (1951) 79.

[12] S. Seehars, A. Amara, A. Refregier, A. Paranjape and J. Akeret, *Information gains from cosmic microwave background experiments*, *Phys. Rev. D* **90** (2014) 023533 [`1402.3593`].

[13] A. Nicola, A. Amara and A. Refregier, *Consistency tests in cosmology using relative entropy*, *Journal of Cosmology and Astroparticle Physics* **2019** (2019) 011 [`1809.07333`].

[14] A. Hosoya, T. Buchert and M. Morita, *Information Entropy in Cosmology*, *Physical Review Letters* **92** (2004) 141302 [`gr-qc/0402076`].

[15] L. Verde, P. Protopapas and R. Jimenez, *Planck and the local Universe: Quantifying the tension*, *Physics of the Dark Universe* **2** (2013) 166 [`1306.6766`].

[16] S. Seehars, S. Grandis, A. Amara and A. Refregier, *Quantifying concordance in cosmology*, *Phys. Rev. D* **93** (2016) 103507 [`1510.08483`].

[17] S. Grandis, S. Seehars, A. Refregier, A. Amara and A. Nicola, *Information gains from cosmological probes*, *Journal of Cosmology and Astroparticle Physics* **2016** (2016) 034 [`1510.06422`].

[18] M. Raveri, M. Martinelli, G. Zhao and Y. Wang, *Information Gain in Cosmology: From the Discovery of Expansion to Future Surveys*, *arXiv e-prints* (2016) arXiv:1606.06273 [`1606.06273`].

[19] S. Hee, W.J. Handley, M.P. Hobson and A.N. Lasenby, *Bayesian model selection without evidences: application to the dark energy equation-of-state*, *Monthly Notices of the Royal Astronomical Society* **455** (2016) 2461 [`1506.09024`].

[20] S. Grandis, D. Rapetti, A. Saro, J.J. Mohr and J.P. Dietrich, *Quantifying tensions between CMB and distance data sets in models with free curvature or lensing amplitude*, *Monthly Notices of the Royal Astronomical Society* **463** (2016) 1416 [`1604.06463`].

[21] G.-B. Zhao, M. Raveri, L. Pogosian, Y. Wang, R.G. Crittenden, W.J. Handley et al., *Dynamical dark energy in light of the latest observations*, *Nature Astronomy* **1** (2017) 627 [`1701.08165`].

[22] A. Nicola, A. Amara and A. Refregier, *Integrated cosmological probes: concordance quantified*, *Journal of Cosmology and Astroparticle Physics* **2017** (2017) 045 [`1706.06593`].

[23] L.T. Hergt, W.J. Handley, M.P. Hobson and A.N. Lasenby, *Bayesian evidence for the tensor-to-scalar ratio r and neutrino masses m_ν: Effects of uniform vs logarithmic priors*, *Physical Review D* **103** (2021) 123511 [`2102.11511`].

[24] W. Handley and P. Lemos, *Quantifying dimensionality: Bayesian cosmological model complexities*, *Phys. Rev. D* **100** (2019) 023512 [`1903.06682`].

[25] A. Lewis and S. Bridle, *Cosmological parameters from CMB and other data: A Monte Carlo approach*, *Phys. Rev. D* **66** (2002) 103511 [`astro-ph/0205436`].

[26] W.J. Handley, M.P. Hobson and A.N. Lasenby, *PolyChord: nested sampling for cosmology*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) L61 [`1502.01856`].

[27] W.J. Handley, M.P. Hobson and A.N. Lasenby, *PolyChord: next-generation nested sampling*, *Mon. Not. Roy. Astron. Soc.* **453** (2015) 4384 [`1506.00171`].

[28] P. Lemos, M. Raveri, A. Campos et al., *Assessing tension metrics with dark energy survey and planck data*, *Monthly Notices of the Royal Astronomical Society* **505** (2021) 6179 [`2012.09554`].

[29] H. Bevins, W. Handley, P. Lemos, P. Sims, E. de Lera Acedo and A. Fialkov, *Marginal bayesian statistics using masked autoregressive flows and kernel density estimators with examples in cosmology*, *Phys. Sci. Forum* **5** (2022) 1 [`2207.11457`].

[30] P.J. Marshall, N. Rajguru and A. Slosar, *Bayesian evidence as a tool for comparing datasets*, *Phys. Rev. D* **73** (2006) 063526 [`astro-ph/0512075`].

[31] D. Foreman-Mackey, *corner.py: Scatterplot matrices in Python*, *The Journal of Open Source Software* **1** (2016) 24.

[32] Planck Collaboration, P.A.R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown et al., *Planck 2013 results. XVI. Cosmological parameters*, *Astron. Astrophys.* **571** (2014) A16 [`1303.5076`].

[33] J. Torrado and A. Lewis, "Cobaya: Bayesian analysis in cosmology." Astrophysics Source Code Library, record ascl:1910.019, Oct., 2019.

[34] J. Torrado and A. Lewis, *Cobaya: Code for Bayesian analysis of hierarchical models*, *JCAP* **2021** (2021) 057 [`2005.05290`].

[35] A. Lewis, A. Challinor and A. Lasenby, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, *The Astrophysical Journal* **538** (2000) 473 [`astro-ph/9911177`].

[36] G.R. Blumenthal, S.M. Faber, J.R. Primack and M.J. Rees, *Formation of galaxies and large-scale structure with cold dark matter*, *Nature* **311** (1984) 517.

[37] A.G. Riess, A.V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P.M. Garnavich et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *The Astrophysical Journal* **116** (1998) 1009 [`astro-ph/9805201`].

[38] S. Perlmutter, G. Aldering, G. Goldhaber, R.A. Knop, P. Nugent, P.G. Castro et al., *Measurements of ω and λ from 42 high-redshift supernovae*, *The Astrophysical Journal* **517** (1999) 565 [`astro-ph/9812133`].

[39] M.S. Turner and M. White, *Cdm models with a cosmological constant*, *Phys. Rev. D* **56** (1997) R4439 [`astro-ph/9701138`].

[40] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, *Int. J. Mod. Phys. D* **10** (2001) 213 [`gr-qc/0009008`].

[41] E.V. Linder, *Exploring the expansion history of the universe*, *Physical Review Letters* **90** (2003) 091301 [`astro-ph/0208512`].

[42] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti and P.D. Serpico, *Relic neutrino decoupling including flavour oscillations*, *Nucl. Phys. B* **729** (2005) 221 [`hep-ph/0506164`].

[43] A. Kosowsky and M.S. Turner, *Cbr anisotropy and the running of the scalar spectral index*, *Phys. Rev. D* **52** (1995) R1739 [astro-ph/9504071].

[44] A.H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, *Phys. Rev. D* **23** (1981) 347.

[45] A.A. Starobinsky, *A new type of isotropic cosmological models without singularity*, *Phys. Lett. B* **91** (1980) 99.

[46] A.A. Starobinsky, *Spectrum of relict gravitational radiation and the early state of the universe*, *JETP Lett.* **30** (1979) 682.

[47] J. Skilling, *Nested sampling*, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, R. Fischer, R. Preuss and U. von Toussaint, eds., vol. 735 of *Amer. Inst. Phys. Conference Proc.*, (Garching, Germany), pp. 395–405, 2004, DOI.

[48] D. Piras, A. Polanska, A. Spurio Mancini, M.A. Price and J.D. McEwen, *The future of cosmological likelihood-based inference: accelerated high-dimensional parameter estimation and model comparison*, *Open Journal of Astrophysics* **7** (2024) [2405.12965].

[49] A. Heavens, Y. Fantaye, A. Mootoovaloo, H. Eggers, Z. Hosenie, S. Kroon et al., *Marginal likelihoods from monte carlo markov chains*, *arXiv e-prints* (2017) [1704.03472].

[50] J. Buchner, *Nested sampling methods*, *Statistics Surveys* **17** (2023) 169 [2101.09675].

[51] G. Ashton, C. Talbot, R. Prix, G. Mastrantonio, L. Faire, H.S. Ly et al., *Nested sampling for gravitational-wave astronomy*, *Nature Reviews Methods Primers* **2** (2022) 39 [2204.02906].

[52] Z. Hu, A. Baryshnikov and W. Handley, *aeons: approximating the end of nested sampling*, *arXiv e-prints* (2023) [2312.00294].

[53] E. Higson, W. Handley, M. Hobson and A. Lasenby, *Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation*, *Stat. Comput.* **29** (2019) 891 [1704.03459].

[54] Planck Collaboration, N. Aghanim et al., *Planck 2018 results. V. CMB power spectra and likelihoods*, *Astron. Astrophys.* **641** (2020) A5 [1907.12875].

[55] K. Benabed, "Planck likelihood code (clik)." https://github.com/benabed/clik, 2023.

[56] Planck Collaboration, *Planck 2018 results. viii. gravitational lensing*, *Astron. Astrophys.* **641** (2020) A8 [1807.06210].

[57] G. Efstathiou and S. Gratton, *A Detailed Description of the CamSpec Likelihood Pipeline and a Reanalysis of the Planck High Frequency Maps*, *arXiv e-prints* (2019) arXiv:1910.00483 [1910.00483].

[58] BICEP/Keck Collaboration, *Improved constraints on primordial gravitational waves using planck, wmap, and bicep/keck observations through the 2018 observing season*, *Phys. Rev. Lett.* **127** (2021) 151301 [2110.00483].

[59] BICEP/Keck Collaboration, "BICEP/Keck BK18 Data Release." http://bicepkeck.org/BK18_datarelease/, 2021.

[60] F. Beutler, C. Blake, M. Colless et al., *The 6df galaxy survey: baryon acoustic oscillations and the local hubble constant*, *Mon. Not. R. Astron. Soc.* **416** (2011) 3017 [1106.3366].

[61] A.J. Ross, L. Samushia, C. Howlett et al., *The clustering of the sdss dr7 main galaxy sample – i. a 4 per cent distance measure at z = 0.15*, *Mon. Not. R. Astron. Soc.* **449** (2015) 835 [1409.3242].

[62] S. Alam, M. Aubert, S. Avila et al., *Completed sdss-iv extended baryon oscillation spectroscopic survey: Cosmological implications from two decades of spectroscopic surveys at the apache point observatory*, *Phys. Rev. D* **103** (2021) 083533 [2007.08991].

[63] Cobaya Development Team, "BAO data for Cobaya."
https://github.com/CobayaSampler/bao_data, 2023.

[64] A.G. Riess, S. Casertano, W. Yuan, L.M. Macri, D. Scolnic et al., *Cosmic distances calibrated to 1% precision with gaia edr3 parallaxes and hubble space telescope photometry of 75 milky way cepheids confirm tension with λcdm*, *Astrophys. J. Lett.* **908** (2021) L6 [2012.08534].

[65] D.M. Scolnic, D.O. Jones, A. Rest et al., *The complete light-curve sample of spectroscopically confirmed sne ia from pan-starrs1 and cosmological constraints from the combined pantheon sample*, *Astrophys. J.* **859** (2018) 101 [1710.00845].

[66] DES Collaboration, *Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **98** (2018) 043526 [1708.01530].

[67] Planck Collaboration, *Planck 2018 results. vi. cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].

[68] P.J.E. Peebles and J.T. Yu, *Primeval adiabatic perturbation in an expanding universe*, *The Astrophysical Journal* **162** (1970) 815.

[69] W. Hu and M. White, *A cmb polarization primer*, *New Astron.* **2** (1997) 323 [astro-ph/9706147].

[70] W. Hu and T. Okamoto, *Mass reconstruction with cosmic microwave background polarization*, *The Astrophysical Journal* **574** (2002) 566 [astro-ph/0111606].

[71] D.J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *The Astrophysical Journal* **496** (1998) 605 [astro-ph/9709112].

[72] D.J. Eisenstein, I. Zehavi, D.W. Hogg, R. Scoccimarro, M.R. Blanton, R.C. Nichol et al., *Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies*, *The Astrophysical Journal* **633** (2005) 560 [astro-ph/0501171].

[73] S. Cole, W.J. Percival, J.A. Peacock, P. Norberg, C.M. Baugh, C.S. Frenk et al., *The 2df galaxy redshift survey: Power-spectrum analysis of the final dataset and cosmological implications*, *Monthly Notices of the Royal Astronomical Society* **362** (2005) 505 [astro-ph/0501174].

[74] Cobaya Development Team, "Supernova data for Cobaya."
https://github.com/CobayaSampler/sn_data, 2023.

[75] M. Kilbinger, *Cosmology with cosmic shear observations: a review*, *Rep. Prog. Phys.* **78** (2015) 086901 [1411.0115].

[76] C. Howlett, A. Lewis, A. Hall and A. Challinor, *CMB power spectrum parameter degeneracies in the era of precision cosmology*, *Journal of Cosmology and Astroparticle Physics* **2012** (2012) 027 [1201.3654].

[77] A.J. Mead, C. Heymans, L. Lombriser, J.A. Peacock, O.I. Steele and H.A. Winther, *Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces*, *Monthly Notices of the Royal Astronomical Society* **459** (2016) 1468–1488.

[78] W. Handley, *anesthetic: nested sampling visualisation*, *J. Open Source Softw.* **4** (2019) 1414 [1901.03279].