unimpeded: A Public Nested Sampling Database for Cosmology

Dily Duan Yi Ong [1,2,3¶] **and Will Handley** [1,2]

1 Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, UK **2** Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge, CB3 0HE, UK **3** Newnham College, Sidgwick Avenue, Cambridge, CB3 9DF, UK ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Bayesian inference is central to modern cosmology. While parameter estimation is achievable with unnormalised posteriors traditionally obtained via MCMC methods, comprehensive model comparison and tension quantification require Bayesian evidences and normalised posteriors, which remain computationally prohibitive for many researchers. To address this, we present `unimpeded`, a publicly available Python library and data repository providing DiRAC-funded (DP192 and 264) pre-computed nested sampling and MCMC chains with their normalised posterior samples, computed using `Cobaya` (Torrado & Lewis, 2021) and the Boltzmann solver CAMB (Lewis et al., 2000; Lewis & Bridle, 2002). `unimpeded` delivers systematic analysis across a grid of eight cosmological models (including ΛCDM and seven extensions) and 39 modern cosmological datasets (comprising individual probes and their pairwise combinations). The built-in tension statistics calculator enables rapid computation of six tension quantification metrics. All chains are hosted on Zenodo with permanent access via the `unimpeded` API, analogous to the renowned Planck Legacy Archive (Dupac et al., 2015) but utilising nested sampling (Skilling, 2006) rather than traditional MCMC methods.

Motivation

With the advancement of observational cosmology, more cosmological datasets have become available and revealed tensions, such as the Hubble tension (Verde et al., 2019), σ_8 tension (Joudaki & others, 2017), and Ω_K tension (Di Valentino et al., 2020; Will Handley & Lemos, 2021). Our standard model of cosmology faces several significant challenges, giving rise to alternative models additional to the baseline ΛCDM model, aiming to better understand our universe. We need a systematic exploration across a wide range of models and datasets to accurately quantify tension when comparing and combining datasets under different cosmological models.

The Planck Legacy Archive (PLA) (Dupac et al., 2015) provides a trusted set of MCMC chains that have become a cornerstone for cosmological analysis. However, MCMC is primarily designed for parameter estimation and is inefficient at calculating the Bayesian evidence, and hence normalised posteriors, which are the key quantities for model comparison and tension quantification. Nested sampling algorithms (Feroz et al., 2009; W. J. Handley et al., 2015; Speagle, 2020) compute the evidence as a primary output, making them the superior tool. Despite their advantages, nested sampling runs are computationally expensive, creating a significant barrier for many researchers. Furthermore, the lack of a centralised, public archive for nested sampling products has hampered reproducibility and large-scale comparative studies.

unimpeded

`unimpeded` addresses these challenges directly. It provides a pip-installable tool that leverages the `anesthetic` package (W. Handley, 2019) for analysis and introduces a seamless Zenodo integration for data management. The nested sampling theory and methodology are detailed in Ong & Handley (2025). Its main features are:

1. **A Public Nested Sampling Grid:** The package provides access to a pre-computed grid of nested sampling chains for 8 cosmological models (standard ΛCDM and seven extensions), run against 39 datasets (comprising individual probes and their pairwise combinations). This saves the community significant computational resources and provides a common baseline for new analyses. Evidence and Kullback-Leibler divergence can be calculated jointly with `anesthetic` for model comparison and quantifying the constraining power of datasets and models, respectively. The scientific results from this grid are presented in Ong & Handley (in prep).
2. **Archival and Reproducibility via Zenodo:** `unimpeded` automates the process of archiving analysis products. The `DatabaseCreator` class bundles chains and metadata, uploading them to a Zenodo community to generate a permanent, citable Digital Object Identifier (DOI). The `DatabaseExplorer` class allows public user to easily download and analyse these chains, promoting open science and effortless reproducibility. Figure 1 demonstrates `unimpeded`'s visualisation capabilities using the `anesthetic` package, showing how posterior distributions (orange) are constrained relative to prior distributions (blue) for the $\Omega_k\Lambda$CDM model with Planck+CMB lensing+SDSS data.
3. **Tension Statistics Calculator:** With the nested sampling chains and the built-in tension statistics calculator, six tension quantification metrics with different characteristics are available, including the R statistic, information ratio I, suspiciousness S, Gaussian model dimensionality d_G, tension significance in units of σ, and p-value. Each of them has unique characteristics optimised for different tasks, thoroughly discussed in Ong & Handley (2025). `unimpeded` implements these statistics with the necessary F-correction to account for discarded prior volume (W. Handley & Lemos, 2019; Will Handley & Lemos, 2021). Figure 2 demonstrates the tension calculator output showing p-value derived tension significance (σ) for 31 pairwise dataset combinations across 8 cosmological models, sorted by significance to highlight the most problematic dataset pairs.

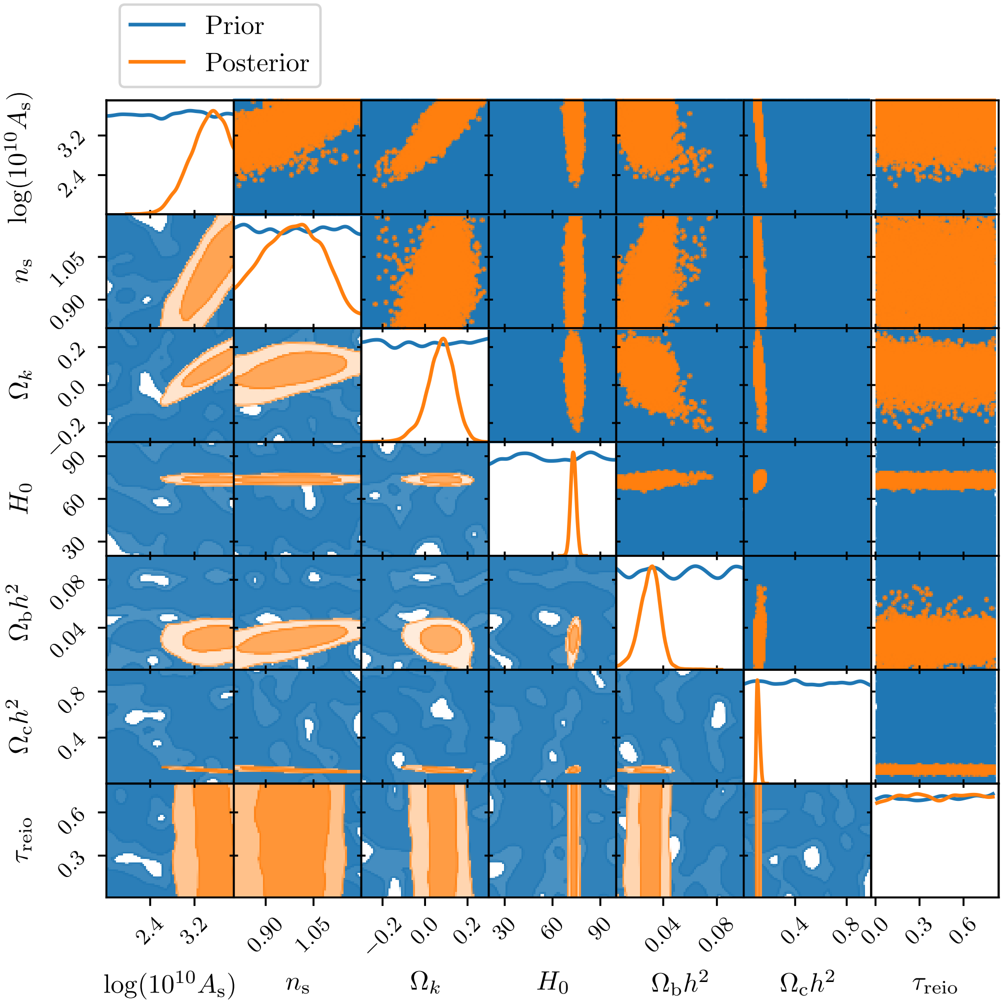

Figure 1: Corner plot visualisation showing prior (blue) and posterior (orange) distributions for the $\Omega_k\Lambda$CDM model constrained by Planck+CMB lensing+SDSS data. Diagonal panels display one-dimensional marginalised distributions, demonstrating how the data constrains parameters. Lower triangular panels show two-dimensional joint distributions with 68% and 95% confidence contours. Visualisation created using `anesthetic`.

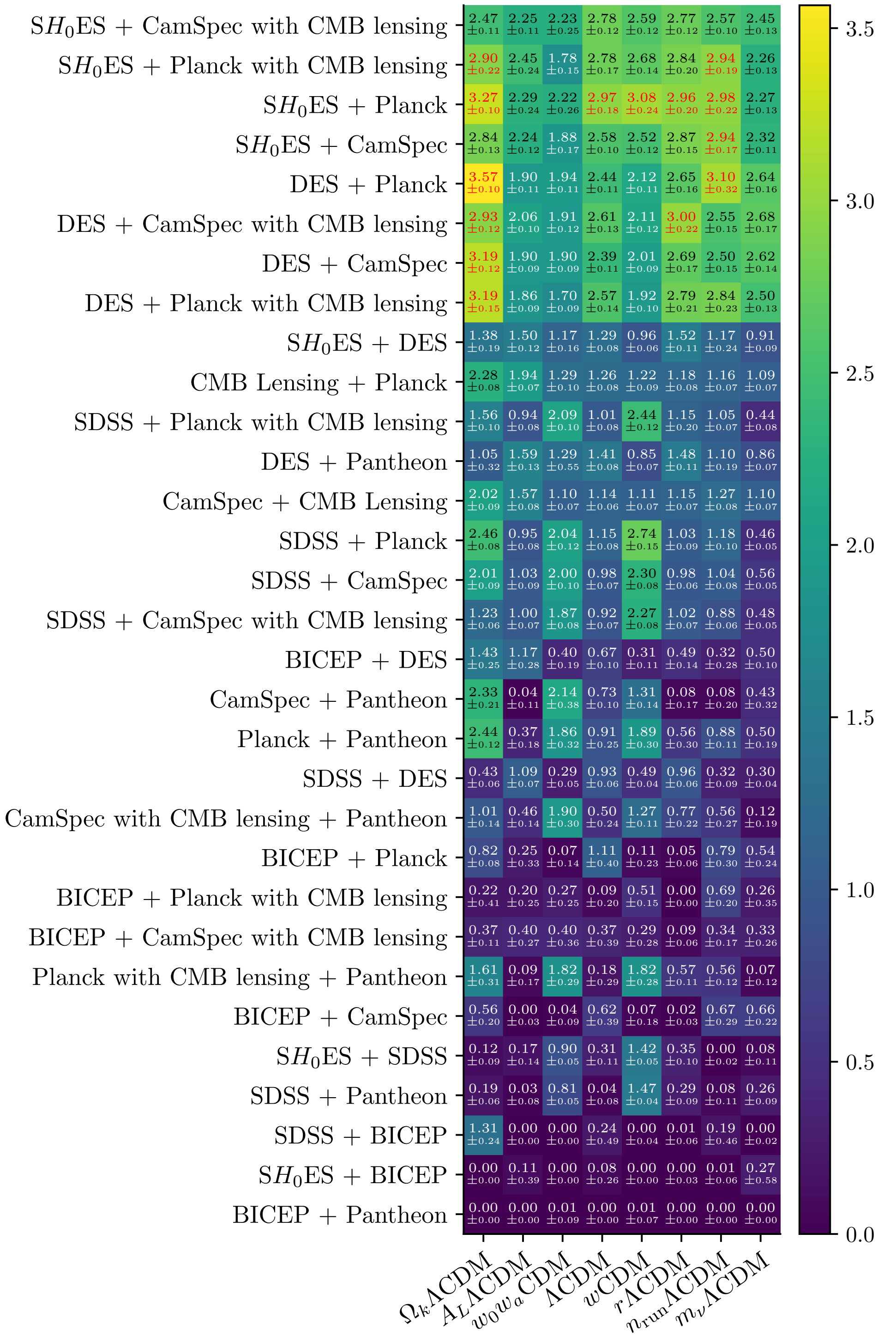

Figure 2: Tension analysis heatmap produced by `unimpeded` displaying p-value derived tension significance (σ values) for 31 pairwise dataset combinations across 8 cosmological models. Rows are sorted by significance, with the most problematic dataset pairs (highest tension) at the top. This demonstrates `unimpeded`'s capability to systematically quantify tensions and their model dependence.

While tools like `getdist` (Lewis, 2019) are excellent for MCMC analysis, and frameworks like `CosmoSIS` (Zuntz et al., 2015) or `MontePython` (Brinckmann & Lesgourgues, 2019) are used

The Journal of Open Source Software

for running simulations with samplers like Cobaya (Torrado & Lewis, 2021), unimpeded fills a unique niche. It is not a sampler but a high-level analysis and database management tool that extends the capabilities of its underlying engine, anesthetic, to create a public, reproducible, and statistically robust nested sampling resource for the cosmology community.

The package is fully documented, tested, and available for installation via the Python Package Index (PyPI).

Acknowledgements

We thank the developers of the open-source packages that this work relies upon, including anesthetic, numpy, scipy, pandas, and corner.py. This work was performed using the Cambridge Service for Data Driven Discovery (CSD3), operated by the University of Cambridge Research Computing Service, provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk).

References

Brinckmann, T., & Lesgourgues, J. (2019). MontePython 3: boosted MCMC sampler and other features. *Physics of the Dark Universe*, *24*. https://doi.org/10.1016/j.dark.2018.100260

Di Valentino, E., Melchiorri, A., & Silk, J. (2020). Planck evidence for a closed Universe and a possible crisis for cosmology. *Nature Astronomy*, *4*, 196–203. https://doi.org/10.1038/s41550-019-0906-9

Dupac, X., Arviset, C., Fernandez Barreiro, M., Lopez-Caniego, M., & Tauber, J. (2015, December). The Planck Legacy Archive. *Science Operations 2015: Science Data Management*. https://doi.org/10.5281/zenodo.34639

Feroz, F., Hobson, M. P., & Bridges, M. (2009). MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, *398*(4), 1601–1614. https://doi.org/10.1111/j.1365-2966.2009.14548.x

Handley, W. (2019). anesthetic: nested sampling visualisation. *Journal of Open Source Software*, *4*(37), 1414. https://doi.org/10.21105/joss.01414

Handley, W. J., Hobson, M. P., & Lasenby, A. N. (2015). PolyChord: nested sampling for cosmology. *Monthly Notices of the Royal Astronomical Society*, *450*(1), L61–L65. https://doi.org/10.1093/mnrasl/slv047

Handley, W., & Lemos, P. (2019). Quantifying tension: interpreting the DES evidence ratio. *Physical Review D*, *100*(4). https://doi.org/10.1103/PhysRevD.100.043504

Handley, Will, & Lemos, P. (2021). Quantifying the global parameter tensions between ACT, SPT, and Planck. *Physical Review D*, *103*(6). https://doi.org/10.1103/PhysRevD.103.063529

Joudaki, S., & others. (2017). KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, *474*(4), 4894–4924. https://doi.org/10.1093/mnras/stx2820

Lewis, A. (2019). GetDist: a Python package for analysing Monte Carlo samples. *arXiv e-Prints*. https://arxiv.org/abs/1910.13970

Lewis, A., & Bridle, S. (2002). Cosmological parameters from CMB and other data: A Monte Carlo approach. *Physical Review D*, *66*(10), 103511. https://doi.org/10.1103/PhysRevD.66.103511

Lewis, A., Challinor, A., & Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *The Astrophysical Journal*, *538*, 473–476. https://doi.org/10.1086/309179

Skilling, J. (2006). Nested sampling for general bayesian computation. *Bayesian Analysis*, *1*(4), 833–859. https://doi.org/10.1214/06-BA127

Speagle, J. S. (2020). dynesty: A dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Monthly Notices of the Royal Astronomical Society*, *493*(3), 3132–3158. https://doi.org/10.1093/mnras/staa278

Torrado, J., & Lewis, A. (2021). Cobaya: Code for Bayesian analysis of hierarchical physical models. *Journal of Cosmology and Astroparticle Physics*, *2021*(05), 057. https://doi.org/10.1088/1475-7516/2021/05/057

Verde, L., Treu, T., & Riess, A. G. (2019). Tensions between the early and late Universe. *Nature Astronomy*, *3*, 891–895. https://doi.org/10.1038/s41550-019-0902-0

Zuntz, J., Paterno, M., Jennings, E., Rudd, D., Manzotti, A., Dodelson, S., Bridle, S., Sehrish, S., & Kowalkowski, J. (2015). CosmoSIS: Modular cosmological parameter estimation. *Astronomy and Computing*, *12*, 45–59. https://doi.org/10.1016/j.ascom.2015.05.005