Beyond Fisher Forecasting for Cosmology

Joseph Ryan,1,2 Brandon Stevenson,1 Cynthia Trendafilova,1,2 and Joel Meyers1

1Department of Physics, Southern Methodist University, 3215 Daniel Ave, Dallas, Texas 75205, USA
2Center for AstroPhysical Surveys, National Center for Supercomputing Applications, Urbana, IL, 61801, USA

(Dated: May 10, 2023)

The planning and design of future experiments rely heavily on forecasting to assess the potential scientific value provided by a hypothetical set of measurements. The Fisher information matrix, due to its convenient properties and low computational cost, provides an especially useful forecasting tool. However, the Fisher matrix only provides a reasonable approximation to the true likelihood when data are nearly Gaussian distributed and observables have nearly linear dependence on the parameters of interest. Also, Fisher forecasting techniques alone cannot be used to assess their own validity. Thorough sampling of the exact or mock likelihood can definitively determine whether a Fisher forecast is valid, though such sampling is often prohibitively expensive. We propose a simple test, based on the Derivative Approximation for Likelihoods (DALI) technique, to determine whether the Fisher matrix provides a good approximation to the exact likelihood. We show that the Fisher matrix becomes a poor approximation to the true likelihood in regions where two-dimensional slices of level surfaces of the DALI approximation to the likelihood differ from two-dimensional slices of level surfaces of the Fisher approximation to the likelihood. We demonstrate that our method accurately predicts situations in which the Fisher approximation deviates from the true likelihood for various cosmological models and several data combinations, with only a modest increase in computational cost compared to standard Fisher forecasts.

I. INTRODUCTION

Parameter forecasts play an essential role in the planning, design, and development of future experiments. For any given model and hypothetical measurement uncertainties, forecasting allows a determination of the precision with which any model parameter can be measured by an experiment before data are collected. The Fisher information matrix [1–4] provides a particularly useful and efficient tool for forecasting.

The Fisher matrix is straightforward to compute and requires only $O(N)$ likelihood evaluations for a model with $N$ parameters. The inverse of the Fisher matrix provides a lower bound on the variances and covariances of any unbiased estimators of model parameters, allowing for the analytic determination of marginalized posterior variances and constraint contours. Due largely to these virtues, Fisher forecasting has become ubiquitous in cosmology [5–7].

On the other hand, Fisher forecasting suffers from a number of limitations. The Fisher matrix only provides an exact description when the posterior is a Gaussian function of the model parameters. This condition is not met unless the data are Gaussian distributed and the observables depend linearly on the model parameters. Due to the assumption of Gaussianity, the Fisher matrix always predicts symmetric, elliptical constraint contours. When data are abundant and model parameters are well constrained, the Fisher approximation often provides a good description. However, significant parameter degeneracy or weakly constraining data can lead to cases where Gaussianity provides a poor approximation; the Fisher matrix will fail to produce accurate forecasts in these circumstances.

Forecasting techniques that do not assume a Gaussian posterior include grid sampling [8], Markov Chain Monte Carlo (MCMC) methods [9–12], and nested sampling [13–16]. These methods provide a more robust determination of non-Gaussian posteriors and parameter constraints, but they are much more computationally expensive than Fisher forecasts, requiring many evaluations of the likelihood function for each model and experimental configuration for which a forecast is desired.

An alternative is provided by the Derivative Approximation for Likelihoods (DALI) technique [17, 18], which can be thought of as a non-Gaussian generalization of Fisher matrix methods [19]. The DALI method is much more efficient than the sampling methods mentioned here, requiring only $O(N^2)$ likelihood evaluations when forecasting constraints on $N$ parameters. The DALI method allows for some non-Gaussian aspects of the posterior to be captured, and is capable of producing curved and asymmetric constraint contours (see Figure 1 for an example of curved constraint contours). However, the DALI method generally requires numerical integration to calculate marginalized constraints, in contrast to Fisher forecasts for which marginalization can be achieved analytically. This numerical integration can be problematic because the forecasted constraints will depend on the details of the numerical integration, including sensitivity to the convergence criterion and the choice of priors on the parameters. Numerical integration also adds computation time; while it is typically a modest computational cost for models with few parameters, it can be a significant burden for models with many parameters. Furthermore, as we will discuss below, the DALI method is not...
always guaranteed to improve upon Fisher forecasts. In some cases the DALI method provides estimates which deviate from the true likelihood more than the Fisher approximation.

A challenge of utilizing Fisher forecasts is that the Fisher matrix does not contain the information necessary to determine whether the Gaussian approximation inherent in the technique is justified. One therefore needs to employ a technique that goes beyond the Fisher approximation in order to determine its validity. This could be accomplished with techniques like MCMC forecasts, but this would become prohibitively expensive to carry out for every model and every experimental configuration where a forecast is desired. One of the main goals of this paper is to demonstrate a means by which we can leverage the DALI method to quickly check the validity of Fisher forecasts, without needing to do the numerical integration required to produce a full set of marginalized constraint contours in the DALI method.

Previous studies have demonstrated limitations of Fisher forecasting and identified some sources for its inaccuracies in certain situations [20, 26]. We rely here on a method of testing the validity of the Fisher technique that should apply quite generally. Our approach utilizes the DALI method to determine when the Fisher approximation fails to accurately match the true likelihood. Various schemes for utilizing the DALI methods to more precisely approximate weakly non-Gaussian likelihoods have also been developed [27, 28].

This paper is organized as follows. In Section II we review the DALI formalism [17, 18]. In Section III we develop a test to determine whether Fisher forecasts are expected to receive non-trivial corrections, based on the shapes of two-dimensional cross-sections of the posterior estimated with the DALI method. We then go on to demonstrate the method using cosmological forecasts using expansion history observations, the cosmic microwave background (CMB) power spectrum, and primordial abundance measurements. We describe the details of the forecasts in Section IV and present the results in Section V. We comment on some limitations of the DALI method in Section VI and conclude in Section VII.

II. DALI FORMALISM

The starting point for both the Fisher and DALI forecasting methods is a Taylor series expansion of the log-likelihood

\[ \mathcal{L} = -\log P \]

in the model parameters \( p^\alpha \) about the best-fit point defined by \( p^\alpha_{\text{fid}} \), where “log” refers to the natural logarithm. Here, \( P \) is the posterior, and \( \alpha \) is an index that runs from 1 to \( N \), where \( N \) is the number of parameters. The fiducial best-fit point is typically chosen based on knowledge of the fit of the model to some existing set of observational data. The Fisher approximation treats this expansion to the lowest non-trivial order, resulting in an approximate log-likelihood which is second-order in the difference between model parameters \( p^\alpha \) and their fiducial best-fit values \( p^\alpha_{\text{fid}} \) (since first derivatives of the likelihood vanish at the best-fit point). The posterior in the Fisher approximation is then a Gaussian, which can be written as

\[ P \approx N_0 \exp \left[ -\frac{1}{2} F_{\alpha\beta} \Delta p^\alpha \Delta p^\beta \right] , \]

where repeated indices are summed, \( N_0 \) is a normalization constant, \( \Delta p^\alpha := p^\alpha - p^\alpha_{\text{fid}} \), and

\[ F_{\alpha\beta} := \mu_{\alpha\beta} M \mu_{\beta\alpha} . \]

In Eq. (2), \( \mu \) is a vector of the observables predicted by a given model evaluated at the fiducial values of the model’s parameters, and \( M := C^{-1} \) is the inverse of the covariance matrix of the data. We assume for simplicity that the covariance matrix does not depend on the model parameters. The vectors \( \mu_{\alpha} \) are contracted with \( M \) in the data space, so that the components \( F_{\alpha\beta} \) of the Fisher matrix can also be written as \( F_{\alpha\beta} = \mu^\alpha_{\alpha} M_{\alpha\beta} \mu_{\beta\beta} \), where Latin indices run from 1 to \( n \), \( n \) being the number of data points. There is no difference, in this paper, between tensor components with upper and lower indices.

The Derivative Approximation for Likelihoods (DALI) method treats the Taylor series expansion of the posterior at higher order, thereby accounting for non-Gaussian parameter dependence in the posterior [17]. A straightforward expansion in terms of \( \Delta p^\alpha \) leads to divergent approximations of the posterior, but the series can be re-arranged in order of derivatives of the observables \( \mu \) such that the posterior is normalizable and positive-definite at every order. The “doublet-DALI” approximation of the posterior is the lowest order correction to the Fisher ap-
proximation, and it is given by

\[
P = N_1 \exp \left[ -\frac{1}{2} F_{\alpha\beta} \Delta p^\alpha \Delta p^\beta - \frac{1}{2} G_{\alpha\beta\gamma} \Delta p^\alpha \Delta p^\beta \Delta p^\gamma - \frac{1}{8} H_{\alpha\beta\gamma\delta} \Delta p^\alpha \Delta p^\beta \Delta p^\gamma \Delta p^\delta \right] + O(3),
\]

where

\[
G_{\alpha\beta\gamma} := \mu_{\alpha\beta\gamma} M \mu_{\gamma},
\]

and

\[
H_{\alpha\beta\gamma\delta} := \mu_{\alpha\beta\gamma\delta} M \mu_{\gamma\delta},
\]

and the notation \(O(3)\) refers to terms higher than second order in the derivatives of \(\mu\). The DALI expansion can be carried out to any order, but in this paper we focus on the Fisher and the doublet-DALI approximations, since the doublet-DALI approximation suffices for testing the validity of the Fisher approximation, as will be described later.

In order to test the validity of the Fisher approximation, one just needs to visually examine the extent to which the level curves of Eq. (6) deviate from those of Eq. (7). To assess the validity of the Fisher approximation. In Sec. V we show the results of this test.

As an optional supplement to the cross-section test, a quantitative measure that can be used to indicate a difference between the Fisher and DALI approximations to the likelihood hypersurfaces is the extrinsic curvature of the level curves. To compute the extrinsic curvature as a function of model parameters, we define the constraint equation

\[
\Phi (p^\alpha) := \mathcal{L} (p^\alpha) - C = 0,
\]

and a unit normal vector

\[
n_\alpha := \frac{\Phi_\alpha}{\sqrt{g_{\alpha\beta} \Phi_\alpha \Phi_\beta}},
\]

such that the extrinsic curvature, in two dimensions, is

\[
K = n^\alpha \cdot n = \Phi_{\alpha\beta} \Phi_{\alpha+} - \Phi_{\alpha\beta} \Phi_{\alpha+} \Phi_{\alpha\beta} \Phi_{\alpha+}.
\]

If \(K < 0\) at any points on a two-dimensional cross-section of the parameter space, then the two-dimensional likelihood level curves will be concave at those points. Since the Fisher-approximated level curves are convex everywhere, concavity of the DALI-approximated level curves signals a region in which the two approximations differ (though they can differ for other reasons). Identifying regions in which the DALI-approximated level curves exhibit concavity can aid in finding cases in which the Fisher approximation breaks down, especially since regions of concavity can be detected without the necessity of making plots. We show the level curves and points where the DALI level curves are concave in Figs. 2–9.

Different approaches to the geometry of likelihood contours and its applications have been discussed elsewhere [20, 32].

### III. FISHER VALIDITY TEST

The Fisher approximation can break down if the observables predicted by a given model depend on the model’s parameters in a non-linear fashion. This leads to non-Gaussian correlations among the model parameters represented by terms containing second- or higher-order derivatives of \(\mu\). These non-Gaussian correlations manifest themselves in the form of non-ellipsoidal (often banana-shaped) constraint contours when the posterior is plotted in parameter space. Ellipsoidal constraint contours are convex everywhere, while non-ellipsoidal constraint contours can be concave in some regions of the parameter space. We can therefore leverage the DALI method, which is sensitive to non-linear parameter dependence of the posterior, to signal the breakdown of the Fisher approximation.

Qualitatively, this can be achieved by choosing any two of the \(N\) parameters \(\{p^\alpha\}\), and assigning the remaining \(N - 2\) parameters their fiducial values. This is equivalent to taking a two-dimensional cross-section of the \((N - 1)\)-dimensional constant-\(P\) hypersurface, on which one can plot the level curves of the functions

\[
C = \frac{1}{2} F_{\alpha\beta} \Delta p^\alpha \Delta p^\beta
\]

and

\[
C = \frac{1}{2} F_{\alpha\beta} \Delta p^\alpha \Delta p^\beta + \frac{1}{2} G_{\alpha\beta\gamma} \Delta p^\alpha \Delta p^\beta \Delta p^\gamma + \frac{1}{8} H_{\alpha\beta\gamma\delta} \Delta p^\alpha \Delta p^\beta \Delta p^\gamma \Delta p^\delta.
\]

for various values of \(C\). Because Eq. (7) describes the level curves of the DALI approximation, and Eq. (6) describes the level curves of the Fisher approximation, one just needs to visually examine the extent to which the level curves of Eq. (7) deviate from those of Eq. (6) to assess the validity of the Fisher approximation. In Sec. V we show the results of this test.

The sets of observables that we study fall into three categories: measurements of the expansion history in the range \(z \leq 2.73\), CMB observations, and measurements of primordial light element abundances. To obtain forecast constraints, we consider predictions of the expansion history data from three models (2-parameter flat \(\Lambda\)CDM, 3-parameter non-flat \(\Lambda\)CDM, and 3-parameter flat \(w\)CDM). We study an 8-parameter flat \(\Lambda\)CDM model with the CMB data, and we examine 3-parameter models of Big Bang Nucleosynthesis (BBN) with the primordial abundance observations.

### IV. DATA AND FORECASTS

The sets of observables that we study fall into three categories: measurements of the expansion history in the range \(z \leq 2.73\), CMB observations, and measurements of primordial light element abundances. To obtain forecast constraints, we consider predictions of the expansion history data from three models (2-parameter flat \(\Lambda\)CDM, 3-parameter non-flat \(\Lambda\)CDM, and 3-parameter flat \(w\)CDM). We study an 8-parameter flat \(\Lambda\)CDM model with the CMB data, and we examine 3-parameter models of Big Bang Nucleosynthesis (BBN) with the primordial abundance observations.
The reader should keep in mind that our aim is not to use the DALI formalism to obtain new constraints or forecasts on cosmological model parameters. Our models and data should be seen merely as a proving ground for our proposed Fisher validity test. We only intend to show some examples of what can be done with the formalism, in addition to analyzing how and why the DALI approximation breaks down.

### A. Expansion history data

The expansion history measurements we use consist of a set of cosmic chronometer and standard ruler data. The cosmic chronometer \( H(z) \) data, collected from [33, 39], comprise a set of uncorrelated measurements of the Hubble parameter as a function of redshift, and are compiled in [10].

The standard ruler data can be broken down into two subsets: baryon acoustic oscillation (BAO) and quasar angular size (QSO). The BAO subset consists of eleven distance and Hubble parameter measurements, collected from Refs. [41–45] and compiled in Table 1 of Ref. [46]. Unlike the \( H(z) \) data, subsets of these BAO data are correlated; we use the covariance matrices given by Eq. (20) in Ref. [47] and Eq. (8) of Ref. [46] to account for the correlations (see references for further details).

The QSO subset comprises measurements of the angular sizes of 120 intermediate-luminosity quasars, from knowledge of the intrinsic linear size \( l_m \) of the quasars within the sample, and measurements of their angular diameter distances at a given redshift. These measurements were taken from Ref. [48]. In contrast to the recent analysis of Ref. [49], we do not vary \( l_m \) as a free parameter. We choose instead to follow the practice of Refs. [49, 47] in setting \( l_m = 11.03 \) pc while ignoring the relatively small uncertainty of 0.25 pc.

We fit three simple models to the data using the Fisher and DALI approximations, and compare these fits to those obtained from the exact (i.e. not approximated using the DALI method) likelihoods. The first, and simplest, model is the 2-parameter \( \Lambda \)CDM model, with an expansion history characterized by the Hubble parameter

\[
H(z) = H_0 \sqrt{\Omega_m (1 + z)^3 + (1 - \Omega_m) (1 + z)^{3(1+w_X)}}.
\]

This model has two parameters: the Hubble constant, \( H_0 \), and the non-relativistic mass density in units of the critical mass density, \( \Omega_m \). Throughout this article, “\( \Lambda \)CDM” stands for “\( \Lambda \) Cold Dark Matter”, where \( \Lambda \) refers to the cosmological constant, and cold dark matter is the dominant contributor to the nonrelativistic mass density. We also consider the non-flat \( \Lambda \)CDM model, with the Hubble parameter

\[
H(z) = H_0 \sqrt{\Omega_m (1 + z)^3 + \Omega_k (1 + z)^2 + \Omega_{\Lambda}}.
\]

This model adds one free parameter compared to 2-parameter \( \Lambda \)CDM: the mass density of the cosmological constant \( \Lambda \) in units of the critical mass density, \( \Omega_{\Lambda} \). In contrast to 2-parameter \( \Lambda \)CDM, the 3-parameter \( \Lambda \)CDM model can have curved spatial hypersurfaces, with the amount of curvature quantified by the curvature energy density parameter in units of the critical mass density, \( \Omega_k := 1 - \Omega_m - \Omega_{\Lambda} \). Finally, we consider a flat \( w \)CDM model in which the dark energy is represented by a perfect fluid with equation of state \( p_X = w_X \), and a Hubble parameter given by

\[
H(z) = H_0 \sqrt{\Omega_m (1 + z)^3 + (1 - \Omega_m) (1 + z)^{3(1+w_X)}}.
\]

This model is a simple generalization of flat \( \Lambda \)CDM, which allows for simple time variation of the dark energy density through the factor \( (1 + z)^{3(1+w_X)} \).1 Here the free parameters are \( H_0 \), \( \Omega_m \), and \( w_X \).

To estimate the exact likelihood corresponding to the expansion history data, we use Cobaya [50, 51] to sample over the exact likelihoods and the DALI-approximated likelihoods, using Cobaya’s MCMC sampler (adapted from CosmoMC [10, 12]). For the expansion history data we used a convergence criterion of either \( R - 1 = 10^{-5} \) or \( R - 1 = 5 \times 10^{-6} \), depending on the model and data combination.

### B. CMB forecasts

The unlensed CMB, lensing deflection, and lensed CMB power spectra used in our forecasts are computed with the CAMB Boltzmann code [52, 53]. Lensing reconstruction noise is calculated with CLASS_delens [4], according to the implementation detailed in [54], and Fisher and DALI forecasts are calculated using FisherLens.

The two CMB experimental noise configurations that we consider are given in Table [4] and we assume Gaussian noise with temperature power spectrum of the form

\[
N_{\ell}^{TT} = \Delta_T^2 \exp \left( \frac{\ell(\ell+1)}{8 \log 2} \theta_{\text{FWHM}}^2 \right). \tag{14}
\]

Here \( \Delta_T \) is the instrumental noise in \( \mu \)K-rad, \( \theta_{\text{FWHM}} \) is the full-width at half-maximum beam size in radians, and we assume fully polarized detectors such that \( N_{\ell}^{EE} = N_{\ell}^{BB} = 2N_{\ell}^{TT} \).

The elements of the Fisher matrix are given by

\[
F_{\alpha\beta}^{\text{CMB}} = \sum_{\ell_1, \ell_2} \sum_{XY,Z} \frac{\partial\ell_{\alpha\beta}^{XY}}{\partial p^X} \left[ \text{Cov}_{\ell_1, \ell_2}^{XY, WZ} \right]^{-1} \frac{\partial\ell_{\alpha\beta}^{WZ}}{\partial p^\beta}, \tag{15}
\]

1 Because the model has flat spatial hypersurfaces, the dark energy density in units of the critical mass density must obey the constraint \( \Omega_X = 1 - \Omega_m \).
where \( p^\alpha \) are the cosmological parameters of interest, and \( XY, WZ \) run over \( TT, TE, EE, \) and \( dd \). The lensing deflection spectrum is \( C_{\ell}^{dd} = \ell (\ell + 1) C_{\ell}^{\phi \phi} \). The elements of the DALI tensors needed for the doublet-DALI approximation are given by

\[
G_{\alpha \beta \gamma \delta}^{CMB} = \sum_{\ell_1, \ell_2} \sum_{XY, WZ} \frac{\partial C_{\ell_1}^{XY}}{\partial p^\alpha \partial p^\beta} \left[ C_{\ell_1}^{XY, WZ} \right]^{-1} \frac{\partial C_{\ell_2}^{WZ}}{\partial p^\gamma \partial p^\delta},
\]

and

\[
H_{\alpha \beta \gamma \delta}^{CMB} = \sum_{\ell_1, \ell_2} \sum_{XY, WZ} \frac{\partial C_{\ell_1}^{XY}}{\partial p^\alpha \partial p^\beta} \left[ C_{\ell_1}^{XY, WZ} \right]^{-1} \frac{\partial C_{\ell_2}^{WZ}}{\partial p^\gamma \partial p^\delta}. \tag{16}
\]

For \( TT \) spectra, the sums over \( \ell \) go from \( \ell_{\min} = 30 \) to \( \ell_{\max} = 3000 \), and when performing lensing reconstruction we take \( \ell_{\max} = 3000 \) for temperature; for polarization spectra, we take \( \ell_{\max} = 5000 \). For the lensing spectrum, we include in the sums the modes from \( \ell_{\min} = 2 \) to \( \ell_{\max} = 5000 \). We assume a sky fraction of \( f_{\text{sky}} = 0.6 \), and we include a prior on \( \tau \) with \( \sigma_\tau = 0.007 \). The cosmological parameters, fiducial values, and step sizes for numerical derivatives used in our forecasts are given in Table II. We specify the primordial Helium abundance such that it is consistent with predictions from BBN.

Gravitational lensing of the primary CMB changes the statistics of the temperature and polarization anisotropies, coupling modes of different \( \ell \) and leading to non-Gaussian off-diagonal contributions to the power spectrum covariance matrix. Neglecting these contributions when performing parameter forecasts will lead to overly optimistic estimates of parameter constraints, since one is effectively reducing the number of independent \( \ell \) modes [55]. In this work, we do not include the lensing-induced non-Gaussian covariances in our forecasts, in order to facilitate consistent comparison with the results of our mock likelihood analysis, which assumes Gaussian statistics for the CMB anisotropies. Neglecting the lensing-induced non-Gaussian covariances leads to slightly tighter constraint contours when forecasting using lensed CMB power spectra, but we do not expect it to significantly impact their shape.

In addition to information from CMB forecasts, we also include information from Baryon Acoustic Oscillation (BAO) experiments. The elements of the Fisher information matrix are given by

\[
F_{\alpha \beta}^{BAO} = \sum_a \frac{1}{\sigma_{f,a}^2} \frac{\partial f_a}{\partial p^\alpha} \frac{\partial f_a}{\partial p^\beta}, \tag{18}
\]

where \( f_a \equiv f(z_a) = r_s/d_V(z_a) \); \( r_s \) is the sound horizon at photon-baryon decoupling, and \( d_V \) is the volume distance to source galaxies at redshifts \( z_a \). Similarly to the CMB case, the relevant BAO DALI tensors are given by

\[
G_{\alpha \beta}^{BAO} = \sum_a \frac{1}{\sigma_{f,a}^2} \frac{\partial f_a}{\partial p^\alpha} \frac{\partial f_a}{\partial p^\beta}, \tag{19}
\]

and

\[
H_{\alpha \beta}^{BAO} = \sum_a \frac{1}{\sigma_{f,a}^2} \frac{\partial f_a}{\partial p^\alpha \partial p^\beta} \frac{\partial f_a}{\partial p^\gamma}. \tag{20}
\]

The forecasted fractional errors in \( f_a \) for the DESI experiment are given in Table III. We estimate the exact likelihood corresponding to our CMB experiments defined above, we use Cobaya to sample over the posterior. Lensed CMB power spectra and lensing spectra are computed using the CAMB cosmological Boltzmann code [52, 53]. Our likelihood code is based on the mock CMB likelihood from MontePython [11].

### Table I. Noise levels for CMB experiments used in forecasts. Experiment A corresponds approximately to the beam size and white noise level expected from Simons Observatory [55], and Experiment B corresponds to that of CMB-S4 [56, 57].

| Label    | \( \Delta \phi \) (\( \mu \text{K}-\text{arcmin} \)) | \( \theta_{\text{FWHM}} \) (arcmin) |
|----------|-------------------|------------------|
| Experiment A | 5                 | 1.4              |
| Experiment B | 1                 | 1.4              |

### Table II. Fiducial cosmological parameters and step sizes for numerical derivatives used in forecasts, reproduced from [58].

| Parameter | Fiducial Value | Step Size |
|-----------|----------------|-----------|
| \( \Omega_c h^2 \) | 0.1197          | 0.0030    |
| \( \Omega_b h^2 \) | 0.0222          | 8.0 \times 10^{-4} |
| \( \theta_\text{s} \) | 0.010409        | 5.0 \times 10^{-5} |
| \( \tau \) | 0.060           | 0.020     |
| \( A_s \) | 2.196 \times 10^{-9} | 0.1 \times 10^{-9} |
| \( n_s \) | 0.9655          | 0.010     |
| \( N_{\text{eff}} \) | 3.046           | 0.080     |
| \( m_\nu \) (meV) | 60               | 20        |

### Table III. Expected fractional uncertainties on \( r_s/d_V \) from forecast DESI data, reproduced from [58].

| Redshift | \( \frac{\delta (r_s/d_V)}{(r_s/d_V)} \) (%) | Redshift | \( \frac{\delta (r_s/d_V)}{(r_s/d_V)} \) (%) |
|----------|-------------------|----------|-------------------|
| 0.15     | 1.89              | 0.10     | 0.59              |
| 0.25     | 1.26              | 1.15     | 0.60              |
| 0.35     | 0.98              | 1.25     | 0.57              |
| 0.45     | 0.80              | 1.35     | 0.66              |
| 0.55     | 0.68              | 1.45     | 0.75              |
| 0.65     | 0.60              | 1.55     | 0.95              |
| 0.75     | 0.52              | 1.65     | 1.48              |
| 0.85     | 0.51              | 1.75     | 2.28              |
| 0.95     | 0.56              | 1.85     | 3.03              |
We define the log-likelihood function as 

\[ \mathcal{L}(\{C_\ell\}|\{\hat{C}_\ell\}) = -\frac{1}{2} \sum_\ell (2\ell + 1) f_{\text{sky}} \times \left\{ \text{Tr}[\hat{C}_\ell C_\ell^{-1}] - \ln|\hat{C}_\ell C_\ell^{-1}| - k \right\}, \]

where \( k \) counts the number of correlated Gaussian fields included in the forecast. \( \hat{C}_\ell \) corresponds to simulated CMB power spectra calculated according to our fiducial values and representing our mock experimental data. \( C_\ell \) are functions of the cosmological parameters, sampled over by Cobaya. When using temperature, \( E \)-mode polarization, and lensing information, we have \( k = 3 \), and the \( C_\ell \) become \( 3 \times 3 \) matrices given by

\[ C_\ell = \begin{bmatrix} C_{\ell TT}^{TT} + N_{\ell TT}^{TT} & C_{\ell EE}^{TE} & 0 \\ C_{\ell EE}^{TE} & C_{\ell EE}^{EE} + N_{\ell EE}^{EE} & 0 \\ 0 & 0 & C_{\ell dd}^{dd} + N_{\ell dd}^{dd} \end{bmatrix}. \]  

The \( \hat{C}_\ell \) matrices have the same functional form in terms of their respective CMB anisotropy and noise power spectra. The sky fraction and fiducial parameter values are the same as those used in the Fisher and DALI forecasts, and we include a Gaussian prior on \( \sigma_z = 0.007 \).

We use the same \( \ell \) ranges in all forecasts, with the \( C_\ell \) and \( \hat{C}_\ell \) matrices taking a different form depending on the value of \( \ell \) in the sum in Eq. (21). For \( 2 \leq \ell < 30 \),

\[ C_\ell = [C_{\ell dd}^{dd} + N_{\ell dd}^{dd}] \cdot \]

For \( 30 \leq \ell \leq 3000 \), the matrices take the form given in Eq. (22). Finally, for \( 3000 < \ell \leq 5000 \), we want to exclude \( TT \) information, while keeping \( TE \) and \( EE \) information. We accomplish this by setting the \( TT \) entry to be the same in both \( C_\ell \) and \( \hat{C}_\ell \), in both cases using the simulated CMB power spectrum representing mock data, so that the difference does not contribute to the likelihood.

For BAO data, our log-likelihood function is defined as

\[
\mathcal{L} \left( \left\{ \frac{r_s}{d_V} \right\} \left\{ \frac{\dot{r}_s}{d_V} \right\} \right) = -\frac{1}{2} \sum_z \left( \frac{r_s}{d_V} - \frac{\dot{r}_s}{d_V} \right)^2 \left( \sigma \left( \frac{r_s}{d_V} \right) \right)^{-2}.
\]

From CAMB, we get the angular diameter distance, \( d_A \), the Hubble rate, \( H \), and the radius of the sound horizon at the drag epoch, \( r_s(z_d) \), and we calculate \( r_s/d_V \) with

\[
d_V(z) = \left[ cz(1 + z)^2 d_A^2(z) H^{-1}(z) \right]^{1/3}
\]

at each of the redshifts specified in Table III [58]. Whether our mock likelihood forecasts include BAO or only CMB data, we use a convergence criterion of \( R - 1 = 0.01 \).

C. Primordial abundance measurements

We utilize observations of the light element abundances to place constraints on the baryon-to-photon ratio, \( \eta_0 \equiv \eta \times 10^{10} \), the light relic density, \( N_\nu \), and the neutron lifetime, \( \tau_n \). We focus on observations of the primordial abundance of deuterium, \( D/H \), and the primordial mass fraction of helium-4, \( Y_p \).

To predict the primordial abundances for a given set of cosmological parameters, we use the publicly available PArthENoPE software\(^6\) [60] [62]. We utilize the standard settings for PArthENoPE discussed in [62] with a 9-nucleide nuclear network, with nuclear rates fixed at their best-fit values. Our choice to fix nuclear rates will give errors that are slightly tighter than other published results. We believe this choice is justified because, as mentioned in Sec. IV, our goal is not to present fully robust constraints, but to test the validity of the Fisher approximation.

We found that numerical second derivatives of the primordial abundances predicted by PArthENoPE were not sufficiently stable for use in our analysis. In order to calculate the derivatives used in the Fisher matrix and DALI tensors, we therefore used the semi-analytic parametrization of the light element abundances from Ref. [63]

\[
Y_p = 0.24695 \left( \frac{\eta_0}{6.123} \right)^{0.039} \left( \frac{N_\nu}{3.0} \right)^{0.163} \left( \frac{\tau_n}{879.4 \text{ s}} \right)^{0.729},
\]

and

\[
\frac{D}{H} = 2.493 \times 10^{-5} \times \left( \frac{\eta_0}{6.123} \right)^{-1.634} \left( \frac{N_\nu}{3.0} \right)^{0.405} \left( \frac{\tau_n}{879.4 \text{ s}} \right)^{0.418}.
\]

The observational errors on the primordial abundances are assumed to be Gaussian and are taken to be \( \sigma(Y_p) = 0.0040 \) from Ref. [63] and \( \sigma(D/H) = 0.030 \times 10^{-5} \) from Ref. [65]. We consider two different uncertainties for the neutron lifetime discussed below. The exact likelihood follows a form analogous to Eq. (24). We used Cobaya to sample over model parameters for both the DALI approximation and the exact likelihood, using PArthENoPE to compute the primordial abundances at each sampled point for the exact likelihood. A convergence criterion of \( R - 1 = 0.005 \) was used for both the DALI and exact likelihood sampling.

V. CROSS-SECTION TEST RESULTS

Here we apply the cross-section test described in Sec. III to the models and data sets described in Sec. IV.

\textsuperscript{6}http://parthenope.na.infn.it/
A. Expansion history data

We present our key findings in Figures 2-6. For both the flat $\Lambda$CDM and flat $w$CDM models, the Fisher approximation and DALI approximation to the likelihood level curves show significant differences. This can be seen very clearly in Figure 2, which shows the results of the cross-section test applied to the flat $\Lambda$CDM model with QSO data. In the top right panel, the 1- and 2-$\sigma$ confidence contours derived from the Fisher approximation are shown as purple dotted curves, the 1- and 2-$\sigma$ confidence contours derived from the DALI approximation are shown as green dashed curves, and the red dots represent points in the parameter space at which contours of constant likelihood are concave, $K < 0$. As discussed in Sec. III if the confidence contours in a region of parameter space have negative extrinsic curvature, then the confidence contours are concave there. From the marginalized contours shown in the triangle plot, we can see that the exact likelihood and the DALI approximation are concave, and that the concavity points in the direction of increasing $H_0$, and slightly in the direction of increasing $\Omega_m$, matching what we see in the upper panel of Figure 2. A similar conclusion can be drawn from Figure 3, which shows the results of the cross-section test and MCMC sampling applied to the flat $\Lambda$CDM model with $H(z)$ data.
H(z) data. In this case, because the H(z) data give tighter constraints on the cosmological model parameters, the difference between the exact and DALI contours is less pronounced. Nevertheless, the cross-section test indicates that the Fisher approximation is not fully valid in parts of the parameter space, and comparison with the exact likelihood confirms the breakdown of the Fisher approximation. In contrast, Figure 4 shows a case in which the Fisher approximation works well, at least out to 2σ. In this figure, the upper right panel displays the results of the cross-section test applied to the flat ΛCDM model along with a triangle plot showing the marginalized contours obtained from MCMC sampling of the same model and data. From the triangle plot, we can see no difference between the exact likelihood or the DALI approximation, and the confidence contours of both align closely with the confidence contours of the Fisher approximation. From the upper panel of Figure 4, we can see that while there is negative extrinsic curvature in some regions of the parameter space, these regions do not overlap with the 2σ confidence contours, and the Fisher approximation and DALI approximation predict similar level curves of the likelihood. The Fisher approximation is therefore adequate within 2σ of the best-fit point.

Figure 5 shows an application of the cross-section test to the 3-parameter flat wCDM model, along with a triangle plot showing the marginalized contours obtained from MCMC sampling. The panels in the upper right triangle of this figure show 2-dimensional cross-sections of the Fisher and DALI approximations to the flat wCDM likelihood. In each of these upper right panels, the model parameter not shown is held fixed at its best-fit value. In these cross-section panels, as in Figures 4, purple dotted curves represent the likelihood contours derived from the Fisher approximation, and green dashed curves represent the likelihood contours derived from the DALI approximation. Red dots indicate regions of the parameter space in which K < 0. The main difference between the flat wCDM and the flat ΛCDM results is that, because we apply the cross-section test to slices of the (approximate) likelihood for flat wCDM, it is not as easy to compare these results to what we would expect from an inspection of the marginalized constraint contours shown in the lower left panels of Figure 5. Nevertheless, we can see from the lower left panels that the Fisher approximation provides a poor approximation of the exact likelihood, and while the DALI approximation is not significantly better, it does roughly capture the shape of the exact likelihood. The concavity of the DALI contours is evident in the Ωm - H0 and Ωm - w panels of Figure 5, and the output of the curvature calculation reflects this.

Figure 6 shows the marginalized and cross-sectional constraint contours for a flat wCDM model, though here...
the model is fitted to BAO data only. The un marginalized cross-sections from the DALI and Fisher approximations agree well to 2σ, except for the $H_0 - \nu$ cross-section, which shows noticeable disagreement between DALI and Fisher even at 1σ. This example demonstrates that one should be wary of the Fisher approximation when the DALI approximation to the likelihood differs from the Fisher approximation in the vicinity of the best-fit point, even if only for a subset of parameters.

B. CMB forecasts

Figure 7 shows results for CMB forecasts for Experiment B defined in Table I using the methods described in Section IV B. The cross-section with the most pronounced deviation from an ellipsoidal shape is the $\Omega_c h^2 - m_\nu$ cross-section, shown in the top-right panel. This is consistent with a general pattern we can see by examining Figure 7. Specifically, we find that there is noticeable disagreement between the Fisher forecast and the mock likelihood forecast for the two-dimensional marginalized parameter constraints involving $m_\nu$ and $\Omega_c h^2$. We find one isolated point of negative extrinsic curvature in the $A_s - \tau$ cross-section, but it is located far from the 2σ region of interest, and the Fisher and DALI cross-sections match closely. Accordingly, we find that the Fisher-approximated marginalized contour for $A_s$ and $\tau$ matches well with that obtained from the mock likelihood.

There is one more source of discrepancy between the Fisher and mock likelihood forecasts that results not from the breakdown of the Fisher approximation, but rather from the marginalization procedure. Marginalized Fisher constraints are computed by simply inverting the Fisher matrix, which implicitly treats all parameters as if they take values from $-\infty$ to $+\infty$. However, sampling the mock likelihood is only possible in regions where parameters take physically meaningful values. This makes little difference for parameters whose fiducial values are many standard deviations away from the non-physical regime. However, in the case considered here, $m_\nu$ is physically restricted to be non-negative, and the fiducial value of 60 meV lies only about 1σ from 0. As a result, marginalized constraints obtained with the mock likelihood (and the DALI approximation) show deviation from the Fisher forecast for parameters that are degenerate with $m_\nu$. This causes some non-ellipticity of the marginalized contours involving $A_s$ and $\tau$ on the low end of each of these parameters. One could choose to also marginalize the Fisher forecasts by sampling the Fisher-approximated likelihood, restricting to the same parameter range. We tested this procedure and found in that case the difference in the low end of the marginalized constraints on $A_s$ and $\tau$ between the Fisher forecast and the mock likelihood disappeared, though the disagreement on the high end of these parameters (and on $m_\nu$) resulting from concavity of the likelihood contours remained.

In Table IV we present the upper and lower 68% confidence intervals for $m_\nu$ for several sets of CMB forecasts. One might be surprised by the fourth line, corresponding to the results presented in Figure 7 which shows that the Fisher forecast produces larger errors than the DALI and mock likelihood forecasts. This behavior can be understood in terms of the non-ellipticity discussed above. (Similar behavior can be seen in Figure 5.) In Table IV we show two additional forecasts involving Experiment B, where the Fisher forecasted errors closely match those from the DALI and mock likelihood treatments. Firstly, we show the effect of including BAO constraints in the forecast, which tightens $m_\nu$ contours and results in similar constraints predicted by all three forecasting methods. One can also change the fiducial value of $m_\nu$, for example to 200 meV, which shifts the constraints away from the $m_\nu = 0$ cutoff and the contours regain their elliptical, Gaussian shapes, with constraints once again being similar across each of the three methods.

C. Primordial abundance measurements

The results of the analysis with primordial abundance measurements, detailed in Section IV C, are shown in Figures 8 and 9. The plots use the same observational uncertainties on D/H and $Y_p$, but different uncertainties on the neutron lifetime $\tau_n$.

Figure 8 shows the results with $\sigma(\tau_n) = 0.6$ s, which is the current global average \[66\]. The concavity of the cross-section of the DALI approximation to the likelihood shown in the $N_\nu - \eta_{10}$ panel shows that the Fisher approximation is near the edge of its applicability, especially as one moves toward the tails of the distribution. This is reflected in the marginalized likelihood contours, which show that the 1σ constraints agree well among the Fisher, DALI, and exact likelihood treatments, though there are slight differences among the 2σ contours. Figure 9 shows the results of the same analysis but with a much weaker constraint on the neutron lifetime of $\sigma(\tau_n) = 30$ s. All parameters are more weakly constrained in this case, and the deviation of the Fisher approximation from the exact likelihood becomes more prominent. These cases show concavity in the DALI-approximated likelihood cross-sections, suggesting that the Fisher approximation may not be valid in some parts of parameter space. However, one can see that the deviation from ellipticity is small, and the level curves of the Fisher and DALI approximations do not significantly differ. One would therefore expect that the Fisher approximation provides a good approximation to the exact likelihood, which we can see is borne out in the comparison of the marginalized constraint contours at least out to 2σ.
FIG. 7. Constraint contours for CMB Experiment B, from the three methods discussed in the text. The fiducial value of $m_\nu$ is taken to be 60 meV, and BAO data is not included. In the top-right panel, one can see the negative extrinsic curvature present in the $\Omega_c h^2 - m_\nu$ cross-section. This is consistent with the lower-left panel, where one can see that the constraint contours from the DALI and mock likelihood predictions are curved, and they do not agree well with the Fisher forecast. Although they do not show any points of negative extrinsic curvature in the likelihood cross-sections, the other $m_\nu$ marginalized contours also show disagreement between the results from the Fisher method, and the DALI and mock likelihood methods, for the reasons discussed in the text. Neutrino masses are reported in meV.

VI. LIMITATIONS OF DALI APPROXIMATION

We have demonstrated how two-dimensional slices of the DALI approximation to the likelihood level curves can be used to detect where the Fisher approximation breaks down. One might reasonably wonder whether Fisher forecasts should simply be replaced with the DALI approximation for all parameter forecasts, given that our
| Configuration   | Fiducial $m_{\nu}$ | Fisher   | DALI          | Mock Likelihood |
|-----------------|-------------------|----------|---------------|-----------------|
| Exp A           | 60                | 60 ± 60  | $66^{+27}_{-55}$ | $70^{+30}_{-57}$ |
| Exp A w/BAO     | 60                | 60 ± 27  | $59^{+45}_{-38}$ | $60^{+25}_{-24}$ |
| Exp A           | 200               | 200 ± 41 | $200^{+30}_{-25}$ | $200^{+28}_{-24}$ |
| Exp B           | 60                | 60 ± 54  | $64^{+49}_{-53}$ | $65^{+29}_{-24}$ |
| Exp B w/BAO     | 60                | 60 ± 25  | $60^{+43}_{-38}$ | $60^{+25}_{-24}$ |
| Exp B           | 200               | 200 ± 32 | $200^{+45}_{-39}$ | $200^{+28}_{-24}$ |

TABLE IV. Forecasted upper and lower 68% confidence intervals for $m_{\nu}$ in meV derived using the Fisher matrix, the DALI technique, and a mock likelihood. Fisher gives a larger interval than DALI or the mock likelihood when the fiducial value of $m_{\nu}$ is 60 meV, as discussed in more detail in the text. The three forecasting techniques yield similar confidence intervals when one sets the fiducial value of $m_{\nu}$ to be more than 1σ away from 0 meV, or when BAO data is included to further tighten forecast constraints.

![Likelihood cross-sections](image1.png)

![Likelihood cross-sections](image2.png)

FIG. 8. Parameter constraints from primordial abundance measurements assuming $\sigma(\tau_n) = 0.6$ s in accordance with the current global average [66]. Concavity of the DALI contour in the top middle panel suggests that the Fisher approximation may be suspect over some range of parameter space. The marginalized 2-σ contours of the exact likelihood shown in the bottom panels show a very slight deviation from the Fisher approximation.

![Marginalized likelihood contours](image3.png)

FIG. 9. Same as Figure 8 but with $\sigma(\tau_n) = 30$ s. One can see that as observational constraints are weakened, the deviation from the Fisher approximation becomes more prominent.

### A. Numerical / convergence issues

Obtaining marginalized constraints in the DALI formalism requires numerical integration of the approximated likelihood. Compared to evaluation of the exact likelihood, the DALI formalism can offer a significant reduction in computation time, since it requires many fewer evaluations of the likelihood. However, this numerical integration, even if it is faster, introduces some undesirable sensitivity to numerical details. For example, MCMC techniques often rely upon a convergence criterion related...
TABLE V. Variances of parameters of flat ΛCDM model fit-
ted to BAO data, computed with the exact likelihood, the
dal approximation of the likelihood, and the Fisher ap-
proximation of the likelihood. We used a convergence crite-
on of $R - 1 = 5 \times 10^{-6}$. Across all four runs, the variances
computed with the exact likelihood are fairly stable, while
the variances computed with the DALI approximation exhibit
larger-amplitude fluctuations.

| Run | Parameter | Exact | DALI |
|-----|-----------|-------|------|
| 1   | $H_0$     | 1.08  | 1.49 |
|     | $\Omega_m$| 0.000287 | 0.000879 |
| 2   | $H_0$     | 1.08  | 1.10 |
|     | $\Omega_m$| 0.000287 | 0.000283 |
| 3   | $H_0$     | 1.10  | 1.23 |
|     | $\Omega_m$| 0.000288 | 0.000548 |
| 4   | $H_0$     | 1.11  | 1.10 |
|     | $\Omega_m$| 0.000289 | 0.000414 |
| Fisher | $H_0$ | 1.00  |     |
|       | $\Omega_m$| 0.000270 |     |

One might hope that summary statistics, like the vari-
ance of means among Markov chains to the mean of the
variance of samples within each chain [67]. The choice
of this convergence criterion can often impact the results
that will be found in a given case. One can typically
produce more stable and repeatable results at the cost
of extra computational time by requiring a tighter con-
vergence criterion, though this can become quite costly
when one requires robust predictions more than a few $\sigma$
away from the best-fit point.

One might hope that summary statistics, like the vari-
ance of a given marginalized parameter constraint, might
be mostly insensitive to numerical details. However, we
find that this is not reliably the case in our DALI fore-
casts. From Figure 4 we can see that for the flat ΛCDM
model with BAO data, both the DALI and the Fisher ap-
proximations appear to agree with the constraint con-
tours of the exact likelihood (at least to within $2\sigma$). If
we calculate the variances of the model parameters using
the exact, Fisher, and DALI likelihoods, however, we find
less agreement than what seems to be implied by the fig-
ure. Table V shows the results of these calculations. We
computed the variances of the model parameters, using
both the exact likelihood and the DALI approximation,
in several runs. We found that while the variances of
the parameters computed using the exact likelihood ex-
hibited relatively little variation, the variances computed
using the DALI approximation can show as much as a
61% deviation from the variance averaged across all runs
(in the case of $\Omega_m$ in run 1).

Another sensitivity to numerical details arises in the
treatment of priors. Numerical integration requires that
some choices be made about the range over which to
sample parameters. One will generally find different
marginalized constraints for different choices of priors,
and it is not always obvious which choice should be
made. For example, when model parameters are subject
to physical constraints (such as for parameters that are
only physically meaningful when greater than or equal
to zero), it might seem natural to place this restriction
on the numerical integration. However, this treatment of
priors will lead to marginalized constraints that should
not be compared directly to what is obtained by inverting
the Fisher matrix (which formally integrates all param-
eters over all real numbers). This discussion does not
imply that anything is wrong with any particular choice
of priors used when carrying out numerical integration,
but the fact that the results will depend on that choice
makes the resulting forecasts less transparent.

B. Toy Models

Though it may be possible to overcome the numerical
issues discussed in the previous subsection, we now turn
to cases in which the DALI approximation to the likeli-
hood is worse than the Fisher approximation. We focus
on two toy models, both of which depend on a single pa-
rameter $a$. In the first model, the observable $x$ takes the
form $x = a^2$, and in the second model, $x$ takes the form
$x = (1 + a^2)^{1/4}$. The key difference between these two
models is that the first exhibits super-linear (quadratic)
dependence on $a$, while the second exhibits sub-linear
dependence on $a$.

We tested the behavior of both the Fisher and DALI
approximations against that of the exact likelihood for
each model by plotting $\log(P)$ versus $a$, where $P$ is the
posterior. These plots are shown in Figures 10 and 11.

From Figure 10 we can see that the DALI approximation
matches the exact likelihood perfectly when the given
model has quadratic parameter dependence. The Fisher
approximation, as expected, only works well in the vicin-
ity of one of the critical points of $\log(P)$. Figure 11 by
contrast, shows a very different picture. For our sub-
linear model, both the Fisher and DALI approximations
describe the exact likelihood well near one of the critical
points of $\log(P)$, but they both break down far from the
critical point. Furthermore, the DALI approximation to
the likelihood exhibits a spurious peak at large values of
$a$ that does not appear in the Fisher approximation. We
can see that the DALI approximation to the likelihood
is strictly worse than the Fisher approximation for that
region of parameter space in this model.

VII. CONCLUSION

We have developed a scheme to test the validity of the
Fisher approximation that is based on the DALI formal-

---

7 For a discussion of sub-linear parameter dependence in the con-
text of Fisher forecasts see, e.g. Ref. [68].
We applied our cross-section test to three cosmological data sets: expansion history data, forecasted CMB observations, and primordial light element abundance measurements. Cases that exhibited concavity of likelihood contours and significant differences between the cross sections of the DALI-approximated and Fisher-approximated likelihood level curves typically showed significant deviation of the Fisher approximation from the true likelihood, suggesting that Fisher forecasts were insufficient in these cases. While we focused here on cosmological forecasting, the techniques we developed are applicable to any statistical forecast where Fisher forecasting can be employed.

In the cases where the Fisher formalism breaks down, we found that the DALI forecasts sometimes provide a better match to the true likelihood. However, we do not generally recommend that DALI forecasts should replace Fisher forecasts. We discussed how numerical issues can complicate interpretation of DALI forecasts and also how DALI forecasts in some cases perform worse than the corresponding Fisher forecasts. We therefore recommend that the cross-section test described here be used as a means to check the validity of Fisher forecasts, and that one should instead rely on sampling of the true likelihood in cases where the Fisher formalism breaks down.

Fisher forecasts are by no means perfect and are built on assumptions that we often know to be violated at some level by real observations. However, the efficiency with which the Fisher matrix can be computed and the simplicity of its interpretation ensure that the Fisher formalism will continue to hold a prominent role in forecasting. The test we developed here provides a means to test the validity of Fisher forecasts in any given scenario, and its application will help to discriminate the cases where we should be reasonably confident in Fisher forecasts from those in which a more complete treatment is required.

ACKNOWLEDGMENTS

The authors would like to thank Nicola Bellomo, Connor Hayward, Bohua Li, and Paul Shapiro for helpful discussions. This work is supported in part by the US Department of Energy under Grant DE-SC0010129. Computational resources for this research were provided by SMU’s Center for Research Computing.

1. R. A. Fisher, “A mathematical Examination of the Methods of determining the Accuracy of Observation by the Mean Error, and by the Mean Square Error,” MNRAS 80 (June, 1920) 758–770.
2. R. A. Fisher, “On the Mathematical Foundations of Theoretical Statistics,” Phil. Trans. Roy. Soc. Lond. A 222 (1922) 309–368.
3. R. A. Fisher, “Theory of Statistical Estimation,” Proceedings of the Cambridge Philosophical Society 22.
4. R. A. Fisher, “The logic of inductive inference,” Journal of the Royal Statistical Society 98 no. 1, (1935) 39–82. http://www.jstor.org/stable/29842435.
5. G. Jungman, M. Kamionkowski, A. Kosowsky, and D. N. Spergel, “Cosmological parameter determination with microwave background maps,” Phys. Rev. D 54 (1996) 1332–1344. arXiv:astro-ph/9512139.
[63] T.-H. Yeh, K. A. Olive, and B. D. Fields, “The impact of new $d(p, \gamma)3$ rates on Big Bang Nucleosynthesis,” *JCAP* 03 (2021) 046, arXiv:2011.13874 [astro-ph.CO].

[64] E. Aver, K. A. Olive, and E. D. Skillman, “The effects of He I $\lambda10830$ on helium abundance determinations,” *JCAP* 07 (2015) 011, arXiv:1503.08146 [astro-ph.CO].

[65] R. J. Cooke, M. Pettini, and C. C. Steidel, “One Percent Determination of the Primordial Deuterium Abundance,” *Astrophys. J.* 855 no. 2, (2018) 102, arXiv:1710.11129 [astro-ph.CO].

[66] Particle Data Group Collaboration, P. A. Zyla et al., “Review of Particle Physics,” *PTEP* 2020 no. 8, (2020) 083C01.

[67] A. Gelman and D. B. Rubin, “Inference from Iterative Simulation Using Multiple Sequences,” *Statistical Science* 7 (Jan., 1992) 457–472.

[68] N. Lee and Y. Ali-Haimoud, “Probing small-scale baryon and dark matter isocurvature perturbations with cosmic microwave background anisotropies,” *Phys. Rev. D* 104 no. 10, (2021) 103509, arXiv:2108.07798 [astro-ph.CO].