TTK-20-44

Bayesian reconstruction of the inflaton's speed of sound using CMB data

Guadalupe Cañas-Herrera,[1,2,*] Jesús Torrado,[3,4,†] and Ana Achúcarro[1,5,‡]

[1] *Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands*
[2] *Leiden Observatory, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands*
[3] *Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, D-52056 Aachen, Germany*
[4] *Dept. of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[5] *Dept. of Theoretical Physics, University of the Basque Country UPV-EHU, 48080 Bilbao, Spain*

(Dated: December 9, 2020)

We update the search for features, due to transient reductions in inflaton's speed of sound, in the Cosmic Microwave Background (CMB) angular power spectrum using Planck 2018 temperature, polarization and lensing data. We develop a new methodology to test more flexible templates to reconstruct the reduction of the speed of sound based on Gaussian Processes. We formally derive a dynamical prior for the shape of the reduction using a *maximum-entropy* approach to ensure the physical conditions of the model are satisfied. The posterior allows for one or more consecutive reductions, fitting apparent features in the CMB power spectra in multipoles from a few tens to $\ell \simeq 2000$. As expected, these fits are not statistically favored with respect to the ΛCDM model. The methodology derived here allows for the inclusion of additional data sets (in particular, Large Scale Structure data), which in principle will increase the statistical significance of the reconstruction of the inflaton's speed of sound.

arXiv:2012.04640v1 [astro-ph.CO] 8 Dec 2020

I. INTRODUCTION

The standard cosmological model (ΛCDM) is currently favored by the available data. It assumes that primordial fluctuations are Gaussian and defined by an almost scale-invariant primordial power spectrum. These assumptions do not point to any particular origin, although the simplest inflationary model, canonical slow-roll single-field inflation, naturally predicts them. By contrast, other models of inflation predict deviations from the near scale-invariant spectrum in the form of features. If ever detected, they would open a new window of research in the field of primordial dynamics. See i.e: [1–4].

The study of features of primordial origin can be done within an Effective Field Theory approach. Within this scenario, features can be produced by the time dependence of primordial functions such as the slow-roll parameters or the speed of sound of the effective inflaton (the adiabatic mode). In particular, small, soft and transient reductions in the inflaton's speed of sound produce such correlated localized oscillatory features in the n-point correlation functions. In the 3-point function (or bispectrum), these localized oscillations present a distinct difference in phase between the squeezed and equilateral configurations [5].

The Planck Collaboration [6] searched for deviations of the canonical scenario in its last release of data. Nevertheless, they did not find strong evidence in the context of features in the primordial power spectrum [7]. They included, for the first time, a joint search of correlated simple features in the primordial power spectrum and in the bispectrum, also without significant results. However, the Planck Collaboration has not studied in detail different feature templates such as the above mentioned ones due to small and transient reductions of the inflaton's speed of sound. This motivates us to continue our previous study [8–11], in preparation for a future release of the Planck bispectrum likelihood or for future investigation in light of incoming Large Scale Structure surveys.

Most of the time, the study of features in both observables (primordial power spectrum and higher correlation functions) are model-dependent, both regarding their physical origin and the ansatz used. In our latest paper in this series [11], we already pointed out the need for testing more flexible feature templates to mitigate the dependence on the ansatz. Within this approach, we can test whether multiple and consecutive reductions of the inflaton's speed of sound can take place consecutively, a possibility already pointed out in the previous work [11]. Furthermore, reconstructing the inflaton's speed of sound allows us to test more complex feature templates with variable amplitude and oscillation frequency, which implies more possibilities to fit well-motivated deviations from ΛCDM beyond those that only used a pre-defined ansatz for the features.

Reconstructions at the level of the primordial power spectrum have already been attempted [12–19]. However, there is not enough constraining power in Planck's power spectrum alone to decide on a particular model for the features. Model-informed reconstructions have the advantage of increasing the constraining power by adding the information contained in higher-order correlation functions; but this is only possible if the constraints of the theoretical model are properly imposed on the reconstructed spectrum, so that it will always lead to a

* canasherrera@lorentz.leidenuniv.nl
† torrado@physik.rwth-aachen.de
‡ achucar@lorentz.leidenuniv.nl

consistent prediction. The task of imposing these physical constraints along the reconstruction is non trivial [20]. It is advisable, instead, to reconstruct the primordial dynamics directly. In our case, we reconstruct the inflaton's speed of sound: the timing, intensity and rate of its reduction. Since we are reconstructing the underlying function leading to the correlated features, it is not only guaranteed that we will obtain a consistent bispectrum feature prediction using power spectrum data alone, but we will also be able to use both data sets simultaneously to get a more stringent reconstruction once a bispectrum likelihood has been released.

In this paper, we develop a new analysis pipeline that uses Gaussian Processes (GPs), a hyper-parametric regression technique, to model the inflaton's speed of sound profile. The analytic nature of GPs makes easy to impose the constraints of the theoretical model, which involve derivatives of the reconstructed function. For a given number of nodes in the GP, we construct a prior on the hyper-parameters of the GP model (the position of the nodes and the correlation length), that maximizes entropy with respect to the bare physical constraints. In this way, we verify that nodes are not placed wherever they would lead to an unphysical reconstruction, and that the density with which the hyper-parameters are explored reproduces the measure of the physical prior [21–23].

We test our new pipeline against Planck 2018 temperature, polarization and lensing CMB angular power spectrum data, obtaining corresponding posteriors of the parameters of interest and several *maxima a posteriori*. Our results do not only reproduce our previous findings [11], but allow for combinations of multiple consecutive reductions as well as more complex shapes.

This article is organized as follows. In section II we review the theoretical framework for inflationary correlated features in the primordial power spectrum due to transient reductions in the speed of sound. In section III, we explain the methodology used to generate features in the primordial power spectrum: the parametrization for the reduction in the speed of sound (III A), the chosen priors for the the different parameters (III B) and the computational procedure (III C). In section IV, we present the results corresponding to the fitting of features using the CMB angular power spectrum. Finally, we discuss the results, draw our conclusion and show prospective work for the future in section V.

II. THEORETICAL MODEL

We follow the Effective Field Theory (EFT) of inflationary perturbations [24] to characterize the fluctuations of comoving curvature perturbations around an inflating cosmological background. It starts with an effective action for the Goldstone boson of cosmic time diffeomorphisms $\pi(t,\mathbf{x})$. This Goldstone boson is related to the comoving curvature perturbation $\mathcal{R}(t,\mathbf{x})$ through the relation $\mathcal{R} \approx -H(t)\pi(t,\mathbf{x})$, with the Hubble parameter $H(t) \equiv \dot{a}/a$, with a being the scale factor (where the dot denotes derivatives with respect to cosmic time t). The effective single field action for π up to second order is given by

$$S_2 = \int d^4x a^3 M_P^2 \epsilon_1 H^2 \left[-\frac{\dot{\pi}^2}{c_s^2} + \frac{(\partial_i \pi)^2}{a^2} \right], \tag{1}$$

where $M_P = 1/\sqrt{8\pi G}$ is the reduced Planck Mass in natural units $c = \hbar = 1$, $\epsilon_1 \equiv -\dot{H}/H^2$ is the first slow-roll parameter and $c_s \equiv c_s(t)$ is the time-dependent speed of sound.

The effective single field action up to third order, neglecting higher order slow-roll corrections ($\sim \mathcal{O}(\epsilon_1^2)$) and assuming $\dot{\pi}^3$ to be small and approximately constant reads [25],

$$S_3 = \int \Bigg[d^4x a^3 M_P^2 \epsilon_1 H^2 - 2Hsc_s^{-2}\pi\dot{\pi}^2 - (1-c_s^{-2})\dot{\pi}\left(\frac{\dot{\pi}^2}{c_s^2} - \frac{(\partial_i\pi)^2}{a^2} \right) \Bigg], \tag{2}$$

where $s \equiv s(t)$ parameterizes the change in the speed of sound $c_s(t)$ defined as

$$s \equiv \frac{\dot{c_s}(t)}{(c_s(t)H)}. \tag{3}$$

The physical details of the theory are encoded in the speed of sound c_s and in its corresponding rate of change given by s. The speed of sound c_s accounts for the effects of integrating out the heavy fields within the effective action. To get an insight of what this variable $c_s(t)$ means, we look at the particular case of an effective theory for the comoving curvature perturbation $\mathcal{R}$, when a strong turn in the inflationary trajectory in multifield space is supported by a heavy field $\mathcal{F}$ with "effective mass" $M_{\rm eff}$. In this case, the curvature perturbation $\mathcal{R}$ is kinetically coupled to the heavy field $\mathcal{F}$. This effective action is similar to the EFT of inflation equation (1), with the speed of sound c_s of the adiabatic perturbation $\mathcal{R}$ given by [26, 27]

$$c_s^{-2} = 1 + \frac{4\Omega^2}{k^2/a^2 + M_{\rm eff}}, \tag{4}$$

where Ω is the the angular velocity when there is a turn in the inflationary trajectory, inducing a momentary reduction on the speed of sound c_s [25]. The effect of this variable speed of sound c_s can be seen in the primordial power spectrum $\mathcal{P}_\mathcal{R}$, in the bispectrum $\mathcal{B}_\mathcal{R}$ and in higher-order correlation functions. In particular, transient variations of c_s produce localized oscillatory and correlated features in both $\mathcal{P}_\mathcal{R}$ and $\mathcal{B}_\mathcal{R}$ [28]. Generally, $c_s(t)$ encodes the effect of derivative interactions.

The almost scale-invariant featureless power spectrum $\mathcal{P}_{\mathcal{R}0}$, with $c_s = 1 \to u = 0$, is defined as,

$$\mathcal{P}_{\mathcal{R}0} = A_s \left(\frac{k}{k_*} \right)^{n_s - 1}, \tag{5}$$

where A_s is the scalar amplitude, k_* is the pivot scale and n_s the so-called spectral index, which depends on the slow-roll parameters as,

$$n_s \equiv 1 - 2\epsilon_1 - \epsilon_2, \tag{6}$$

where $\epsilon_2 \equiv \dot{\epsilon_1}/(\epsilon_1 H)$ is the second slow-roll parameter. Under the assumption of small, mild and transient reductions of the speed of sound c_s, the modifications in the primordial power spectrum of curvature perturbations $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$ were already calculated [5]. The quadratic action of EFT of inflation, equation (1), is divided into a free part (resembling single field inflation, with $c_s = 1$) and a small perturbation:

$$S_2 = \int \mathrm{d}^4x\, a^3 M_\mathrm{P}^2 \epsilon H^2 \left(\dot{\pi}^2 - \frac{(\partial_i \pi)^2}{a^2}\right) - \int \mathrm{d}^4x\, a^3 M_\mathrm{P}^2 \epsilon H^2 \left(\left(1 - c_s^{-2}\right)\dot{\pi}^2\right), \tag{7}$$

Transitioning from cosmic time t to the conformal time τ, so that $d\tau = dt/a(t)$, using the *in-in* formalism [29] and the following definition of the variable u,

$$u(\tau) \equiv (1 - c_s^{-2}(\tau)), \tag{8}$$

the change in the primordial power spectrum $\Delta\mathcal{P}_\mathcal{R}$ is given by the Fourier transform of the reduction in the speed of sound c_s:

$$\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}} = k \int_{-\infty}^{0} d\tau u(\tau) \sin(2k\tau). \tag{9}$$

III. METHODOLOGY

The reduction of the speed of sound and its rate of change are encoded in $u(\tau)$ and $s(\tau)$ respectively. We aim to use current cosmological data (the temperature, polarization and lensing power spectrum of Planck 2018) to estimate them given the theoretical framework presented in section II. To do that, we use *Bayesian inference*. The estimation of the joint probability distribution of a set of parameters θ of a model M given some data d, the so-called posterior $P(\theta|d, M)$, is computed using Bayes' Theorem [30]:

$$P(\theta|d, M) \propto \mathcal{L}(d|\theta, M)\Pi(\theta|M), \tag{10}$$

where $\mathcal{L}(d|\theta, M)$ is the *likelihood* (probability of observing the data d given the model M is realized with parameters θ) and $\Pi(\theta|M)$ the *prior* (probability distribution of the parameters θ given some *a priori* information). In this section, we present the methodology that we have employed to study the posterior distribution of $u(\tau)$, and hence $s(\tau)$.

A. Reconstruction model for the reduction in the speed of sound

We use Gaussian Processes (GPs) [31] as an interpolator for reconstructing the speed of sound of the inflation. We aim to reconstruct $u(\tau) \equiv (1 - c_s^{-2}(\tau))$. Since $u(\tau)$ is a negative quantity, it makes sense to reconstruct the logarithm of $-u = |u|$, to guarantee that the GP interpolator, once exponentiated, conserves sign. On the other hand, there is a choice to be made about the scale of the linear axis: whether to reconstruct $\log|u(\tau)|$ or $\log|u(\log|\tau|)|$.

Using the latter case as an example for the formulae, to reconstruct $\log|u(\log|\tau|)|$ we choose a number i of training nodes[1] $(\log|\tau_i|, \log|u_i|)$ where $\log|u_i| := \log|u(\log|\tau_i|)|$ (see Fig. 1), to which we fit a GP with kernel function

$$\kappa(\log|\tau_i|, \log|\tau_{i+1}|;\, l) = c^2 \exp\left\{-\frac{1}{2}\left(\frac{\log|\tau_i| - \log|\tau_{i+1}|}{l}\right)^2\right\}, \tag{11}$$

where c is the *output scale*, and l the *correlation length*. The output scale c plays no role in this approach, and can be fitted using *maximum likelihood* and then ignored. The correlation length will be sampled together with the position of the nodes. To compute the GPs, we use the Python package `sklearn` [32]. The *mean* of the GP is used as an interpolator for $\log|u(\log|\tau|)|$, and reads, in terms of the training nodes, as the matrix product:

$$\log|u(\log|\tau|)| = \kappa(\log|\tau|, \log|\tau_i|;\, l) \times [\kappa(\log|\tau_i|, \log|\tau_j|;\, l)]^{-1} \log|u_j|, \tag{12}$$

where the first kernel function κ is a vector of evaluations at the requested $\log|\tau|$ combined with each of the training $\log|\tau_i|$, the second one is the matrix of evaluations of κ for each pair of training nodes (i, j), and the final term is the vector of training $\log|u_j|$. Once $u(\tau)$ is generated, we calculate $s(\tau)$ from Eq. (3), which we can rewrite more conveniently as

$$\begin{aligned} s(\tau) &= \frac{1}{2}\frac{u}{1-u}\frac{\mathrm{d}\log u}{\mathrm{d}\log|\tau|} \\ &= \frac{1}{2}\frac{u}{u-1}\frac{1}{l^2}\left[(\log|\tau| - \log|\tau_i|)\,\kappa(\log|\tau|, \log|\tau_i|;\, l)\right] \times \\ &\quad [\kappa(\log|\tau_i|, \log|\tau_j|;\, l)]^{-1} \log|u_j|, \end{aligned} \tag{13}$$

where we have taken the derivative after substituting u by the mean of the GP defined in Eq. (12). Notice that this reproduces the matrix product in Eq. (12), just changing

[1] Notice that our use of GPs as interpolators does not involve machine-learning, but we are borrowing the term *training* from its literature.

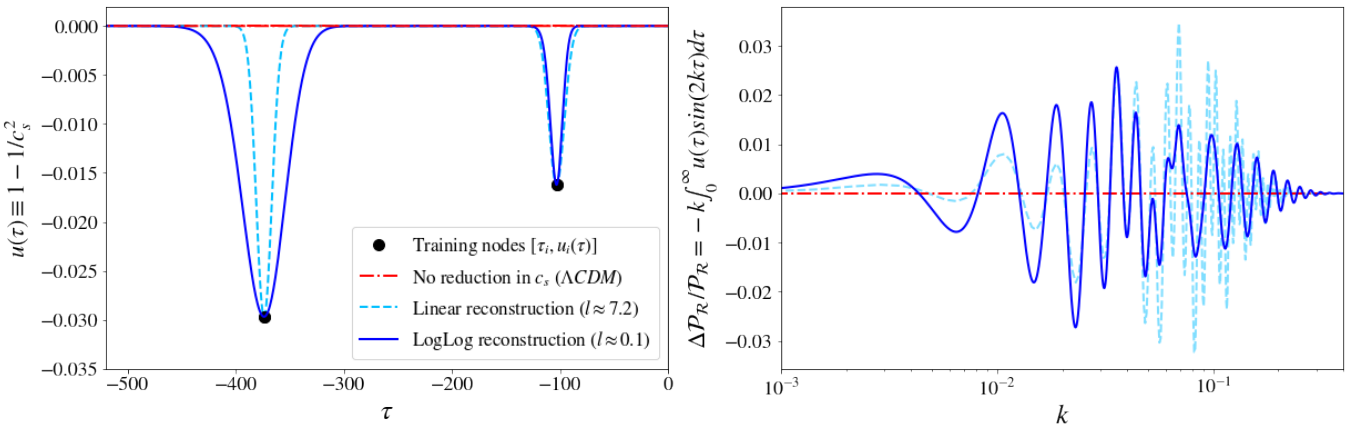

FIG. 1. **Left panel**: Reconstruction of $u(\tau)$ using a GP on $\log|u(\log|\tau|)|$ (continuous line) and on $u(\tau)$ (dashed line), both using as training nodes $(\tau, u) \in \{(-100, -0.016), (-375, -0.03)\}$, and correlation length for each model such that the width of the mode corresponding to the first node is similar (notice the difference in width of the mode of the second one). As the reconstruction is done in conformal time τ, the x-axis is always negative and $\tau = 0$ indicates the end of inflation. **Right panel**: Corresponding feature in the primordial power spectrum. Notice how, despite the similar position of the training nodes, their features look quite different: similar in the leftmost oscillations (corresponding to the node at $\tau = -100$), but very different after that, due to the broader width of the first mode.

the first vector. Finally, we compute the power spectrum feature of Eq. (9) from a fine sampling of the GP using the `FFTLog` algorithm [33, 34]. The density and limits of the $\log|\tau|$ sampling for the `FFTLog` are chosen adaptively to minimise computational costs and guarantee the accurate computation of the transform.

The most consequential difference of the choice between linear and logarithmic τ in the GP will show up whenever we have nodes separated by a distance much larger than the correlation length, appearing as isolated (log)Gaussians: in the linear case, their width in τ will be similar, whereas for the logarithmic one, the width will scale logarithmically (see Fig. 1). Looking at the first equality in Eq. (13), and seeing how s depends on the logarithmic derivative on τ, it is easy to see that the linear parameterization is going to struggle to place two or more nodes away from each other, since $s(\tau)$ will peak at highly different values in each of them, making it hard not to violate the perturbativity bounds on s (see sec. III B).

Thus for the primary results in this paper, we model $\log|u(\log|\tau|)|$. To mitigate excessive sensitivity to the prior of our results, we also perform a reconstruction in $u(\tau)$. Notice that by modelling u and not $\log|u|$ we need to deal with cases in which $u(\tau)$ goes positive, by assigning it null prior density. However, those are generally disfavored by the data (require large l compared to the distance between nodes), and the large difference between the $\log|u(\log|\tau|)|$ and the $u(\tau)$ reconstruction is useful for assessing prior sensitivity.

B. Parameters and Priors

The action described in equations (1) and (2) is perturbative in terms of $(1-1/c_s^2)$. It implies that the reduction in the speed of sound, c_s, cannot be too big ($|u| \ll 1$) and the rate of change in the reduction cannot be too fast ($|s| \ll 1$). Also, the contributions of the slow-roll corrections ϵ_1, ϵ_2 have to be smaller than those of the variable speed of sound c_s. We need to impose these conditions for all values of τ, but it is enough to restrict to the point where $u(\tau)$ and $s(\tau)$ take their maximum value ($|u|_{\max}, |s|_{\max}$). Note that imposing the perturbative limit on $|s|_{\max}$ satisfies the consistency conditions in [35–37]. In short:

$$\max(\epsilon_1, \epsilon_2) \ll \max(|u|_{\max}, |s|_{\max}) \ll 1. \qquad (14)$$

In [11] we argued that this condition could be naturally imposed by a prior Beta(5, 5) on $\max(\log_{10}|u|_{\max}, \log_{10}|s|_{\max})$ between the extremes in Eq. (14) (see left part of Fig. 2), the logarithm coming from the difference in order of magnitude between both bounds.

Contrary to [11], in this work $|u|_{\max}$ and $|s|_{\max}$ are not sampled directly. Instead, the parameter space for the feature consists of the position of the nodes $\{(\tau_i, u_i = u(\tau_i))\}$ and the correlation length l. Thus, the total number of feature parameters is $2N + 1$ for a number N of nodes. Imposing the Beta prior described above is not as simple as sampling the GP parameters from some prior, computing $|u|_{\max}$ and $|s|_{\max}$ along the reconstruction, and multiplying by the Beta density. That procedure will likely introduce undesired information in the shape of under- or overdensities in the probability

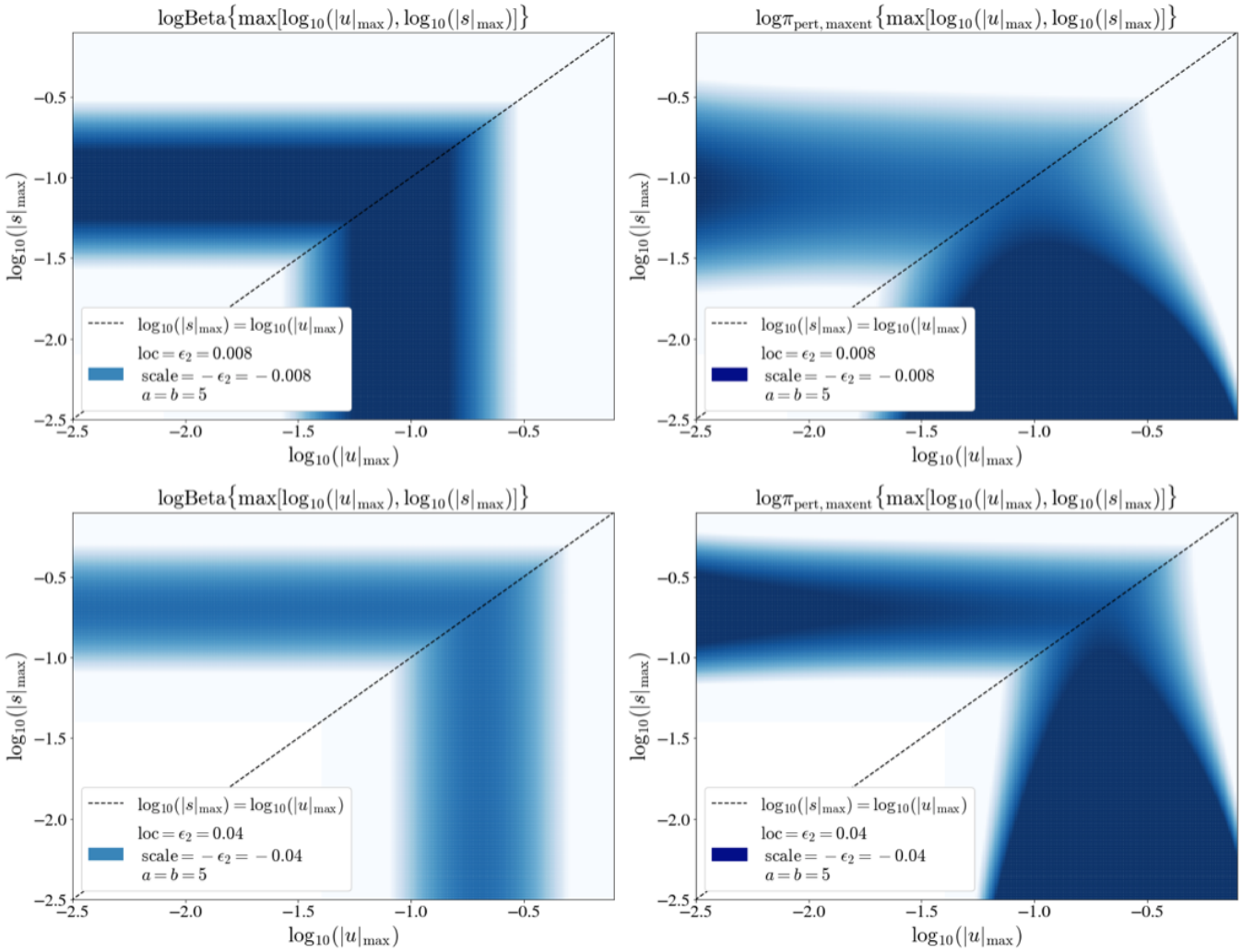

FIG. 2. A description of the Bayesian priors adopted in this study. **Left panels**: Desired Beta density distribution for the parameters $log_{10}(|u|_{\max})$ and $log_{10}(|s|_{\max})$, with $\epsilon_2 = 0.008$ and $\epsilon_2 = 0.04$, respectively. **Right panels**: probability density distribution induced by π_0 on $log_{10}(|s|_{\max})$ and $log_{10}(|u|_{\max})$ (defined as P in the denominator of Eq. (15)), for a Gaussian-Processes reconstruction in $\log|u(\log|\tau|)|$ using 2 training nodes. The initial prior $\pi_0(\tau_i, u_i, l)$ induces overdensities at low $log_{10}(|s|_{\max})$ due to this quantity being doubly-correlated to early-time position of the nodes and large correlation lengths; Eq. (15) corrects for this effect.

induced on $(|u|_{\max}, |s|_{\max})$, which would finally diverge significantly from a Beta. The correct way to proceed so that the induced probability on the physical parameters is the desired one is by constructing a distribution on the parameters of the nodes that *maximises entropy* with respect to the desired one, which can be computed, according to [21], as

$$\pi_{\text{pert, maxent}}(\tau_i, u_k, l) = \frac{\pi_0(\tau_i, u_k, l)\pi_{\text{pert}}(|u|_{\max}, |s|_{\max})}{P(|u|_{\max}, |s|_{\max} \mid \pi_0)}, \tag{15}$$

where $\pi_0(\tau_i, u_i, l)$ is some initial prior on the node parameters and $\pi_{\text{pert}}(|u|_{\max}, |s|_{\max})$ the Beta prior described above. The term in the denominator is the probability density induced by π_o on $(|u|_{\max}, |s|_{\max})$, which we compute from a Monte Carlo sample from π_0 using `PolyChord` [38, 39]. The Monte Carlo sample is fed to `GetDist` [40] to construct a density estimator. Both the Beta prior and the maxent prior can be seen in Fig. 2.

For π_0, we choose log-uniform prior distributions on the correlation length l and on the training node location (τ_i, u_i)[2]. The bounds for the time-positions τ_i are chosen so that the feature falls in the CMB window function (features running from scales $k \approx 10^{-3}$ to $k \approx 3 \times 10^{-1}$, though a larger region has been scanned as a consistency check (see Sec. IV A). The bounds for the amplitude of

[2] Note that we are sampling the training nodes and the correlation length in a logarithmic scale as we expect them to vary several orders of magnitude.

the reductions at the nodes, u_i, are chosen to generously fulfill the EFT condition in Eq. (14). Summarizing:

$$\pi_0(\log_{10}(|\tau_1|),\ldots,\log_{10}(|\tau_n|),\log_{10}(|u_1|),\ldots,\log_{10}(|u_n|),\log_{10} l) = \mathcal{U}(1.8<\log_{10}(|\tau_n|)<3.3)\times \prod_{i=n}^{2}\mathcal{U}(\log_{10}(|\tau_i|)<\log_{10}(|\tau_{i-1}|)<3.3)\times \prod_{i=1}^{n}\mathcal{U}(-4<\log_{10}(|u_i|)<0)\times\pi(l) \quad (16)$$

where $\mathcal{U}$ means a uniform distribution, and the prior on the time positions of the nodes includes sorting so that $\tau_i<\tau_{i-1}$ (i runs from $1\ldots n$). The prior on l is chosen so that it produces reasonable values of $|s|_{\max}$. For each of the two reconstructions studied here, $\log|u(\log|\tau|)|$ and $u(\tau)$, the boundaries can be chosen as[3]

$$\pi_{\log|u(\log|\tau|)|}(l) = \mathcal{U}(-2<\log_{10} l<2) \qquad \text{and}$$
$$\pi_{u(\tau)}(l) = \mathcal{U}(-2<\log_{10} l<3.3). \quad (17)$$

As an improvement on [11], in this work we do not fix the value of the slow-roll parameters ϵ_1 and ϵ_2 as bounds of the perturbativity condition Eq. (14). Instead we let the bounds of the Beta distribution run dynamically, marginalizing over the slow-roll parameters. We use as priors uniform distributions $\mathcal{U}(0.0001<\epsilon_1<0.05)$ and $\mathcal{U}(-0.06<\epsilon_2<0.06)$, which encompass the ΛCDM posterior found for them in Planck 2018 [12].

C. Data sets and sampler

To constrain the reduction of the speed of sound, we use the Planck 2018 polarized CMB and lensing data. In particular we use the product of the low multipole likelihoods `lowT` and `lowE`, the *unbinned* high-ℓ likelihood `plik_TTTEEE` and the `lensing` likelihood. We use the unbinned likelihoods because of the fast frequency of oscillations in the features, as was already pointed out in [11].

We compute the changes to the CMB power spectra C_ℓ using the Boltzmann code `CAMB` [41], modified accordingly to account for the increased sampling in k needed by the oscillatory features $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$ in the primordial power spectrum. We sample over the parameter space described in section III B, i.e. the positions of the training nodes $\{(\tau_i, u_i=u(\tau_i))\}$, the correlation length l of the GP, and the kinetic slow-roll parameters ϵ_1 and ϵ_2. We also allow for the possibility of tensor modes, as changes in the Sachs-Wolfe plateau caused by them could possibly be correlated with features at very large scales. We track as derived parameters the scalar tilt n_s, the tensor-to-scalar ratio r and the EFT parameters ($|u|_{\max}, |s|_{\max}$). We fix the rest of cosmological parameters of ΛCDM to the best fit of Planck 2018 with the present likelihoods, as well as the *nuisance* parameters of the likelihoods. This choice is justified by previous sensitivity analyses in [11], that we repeat here for the background ΛCDM parameters by exploring a broader range of τ and l than the one indicated above (see Sec. IV A).

We obtain the posterior distribution of the parameters using the sampler `PolyChord` [38, 39]. We use this nested sampler since, from previous searches, we expect the posterior distributions of $u_{\max}$ and τ_i to be multimodal. The handling of the priors, likelihoods, Boltzmann code and sampler is managed by the Bayesian framework `Cobaya` [42]. The analysis of the posterior distributions is carried out using `GetDist` [40].

We sample the posterior of two different parameterizations of the GP sound speed reconstruction: $\log|u|(\log|\tau|)$ and $u(\tau)$, in the following called simply *logarithmic* and *linear* parameterizations, respectively. We know the logarithmic parameterization is more stable numerically, as it consistently makes the reconstruction of $u(\tau)$ negative. However, we still use the runs in the linear parameterization for the purposes of assessing prior sensitivity. For the first sampling processes (up to three GPs nodes), we run `Cobaya` in parallel launching 8 MPI processes, each allowed to thread across 3 CPU cores. In the case of 4 nodes, we run `Cobaya` with 32 MPI processes, each allowed to thread across only one single CPU core. The nested sampler `PolyChord` has been run with 1000 live points (which is far above the requirements for the current number of dimensions in the parameter space) and a stopping criterion of 0.01. The computation time varies depending on the number of training nodes in the GPs: from a few days with only 1 node, up to several weeks with 4 nodes.

All the *maxima a posteriori* (MAP) presented in the next section have been obtained running `Py-BOBYQA` [43, 44] (a Python implementation of the `BOBYQA` algorithm [45], available via `Cobaya`), initialized on the relevant local maxima of the `PolyChord` samples.

IV. RESULTS

A. Consistency checks

Before presenting our results, we shortly discuss whether the assumptions made in previous sections were justified. In particular, we have tested whether we find clear posterior modes outside the (τ_i, l) prior region described section III B (the *CMB window* prior), and whether in posterior modes either in our initial prior

[3] Notice that, while smaller values of l will result in sharper reductions with too high, forbidden $|s|_{\max}$ values, larger values of l would result in small $|s|_{\max}$ values which are actually allowed as long as $|u|_{\max}$ fulfills Eq. (14). In any case, we are imposing these upper l boundaries for the main runs, since reductions with very small $|s|_{\max}$ tend not to be easily distinguishable from changes in the background cosmological model (see Sec. IV A).

or in the broader region, the assumption of no-correlation with background cosmological parameters is fulfilled.

To do that, we produced a 1-node posterior sample in the logarithmic parameterization in the enlarged prior region $0 < \log_{10}(|\tau_i|) < 4.3$ and $2 < \log_{10} l < 10$, and let the background ΛCDM parameters vary. No significant modes were found outside the original, reduced prior region. We found mild modes in the region $1.8 < \log_{10}(|\tau_i|) < 3.3$ and $2 < \log_{10} l < 3.3$, which presented some degeneracy between Ω_m, n_s and the reconstruction parameters ($\rho \approx 0.17$), due to the fact that these features can be confused with the shape of the first and second acoustic peaks (already observed in [11]). This justifies restricting ourselves to the prior described in III B, since any mode found outside of it would not be distinguishable from background cosmology.

The check for degeneracies between the c_s reconstruction parameters and the slow-roll parameters is of particular importance, since the latter determine the perturbative prior limits on the former (see Eq. 14 and Fig. 2). We have found no significant degeneracies, neither in the tests described above nor in the final runs. We have reproduced the Planck ΛCDM posterior on the slow-roll parameters in all cases (see Fig. 3).

Most of the results below have been run both in the linear and logarithmic parameterizations for the Gaussian Processes (GPs) reconstruction of the speed of sound profile $u(\tau)$. The results agree with each other, in particular, for the maxima a posteriori found at late conformal time (i.e. towards the end of inflation, with $-\tau_i$ of a few hundreds, where both parameterizations look similar). However, the logarithmic parameterization differs from the linear one when training nodes are thrown at early values of conformal time (i.e. $-\tau_i$ over 800), see Fig. 1). This is due to modes of constant width in logarithmic scale getting broader the further we go along the axis of conformal time. In the 1-node case, the linear parameterization reproduces the results in [11] (which uses a Gaussian ansatz in $u(\tau)$), whereas the logarithmic parameterization produces different 1-node posterior modes (see Sec. A 1 in appendix A).

It is worth remarking that the use of the logarithmic parametrization does not compromise the flexibility of our reconstruction of $u(\tau)$. Even though the logarithmic parameterization reconstructs naturally profiles of $u(\tau)$ with broad reductions at earlier conformal times and narrower reductions at later times (which is preferred so that $|s|_{\max}$ is not violated), narrow reductions at early conformal times can always be achieved by adding further nodes that would force the profile to return to zero. If the data and EFT conditions did allow for a narrow reduction at earlier conformal times, we would have seen it during the analysis of the posterior distributions when more than one training node was used.

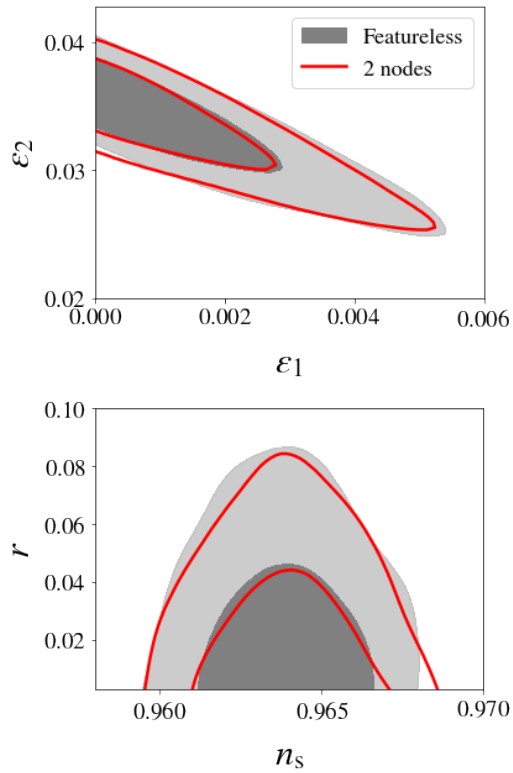

FIG. 3. Posterior distributions of the primordial parameters: the kinetic slow-roll parameters ϵ_1 and ϵ_2, and the derived spectral index n_s and tensor-to-scalar ratio r, for a 2-node reconstruction in the logarithmic parameterization (red line). The grey contours correspond to the featureless ΛCDM scenario. The correspondence between both posteriors is due to the absence of degeneracies between the c_s reconstruction parameters and the slow-roll parameters. Similar results are found for 1, 3 and 4 nodes.

B. Reconstruction of the inflaton's speed of sound profile $u(\tau)$

In this section, we present the results of the GP reconstruction in the logarithmic parameterization using the Planck data as described above, and imposing the Maximum-Entropy prior described in III B for the derived quantities $|u|_{\max}$ and $|s|_{\max}$.

When presenting our results, we use an effective $\Delta\chi^2$ obtained by using a reference value of $\chi^2 = 24611.6$ corresponding to the MAP of Planck 2018 for an equivalent combination of likelihoods, re-evaluated in our case after swapping the binned for the unbinned version of the high-ℓ likelihood. Note that this effective $\Delta\chi^2$ is not meant for model selection purposes and it is used for illustration only. As an example, a triangle plot of the posterior for the 2-nodes case can be seen in Fig. 4.

We have reconstructed the inflaton's speed of sound profile $u(\tau)$ using up to four training nodes. We have stopped there after checking that the Akaike Information Criterion (AIC) [46] has a minimum for three training

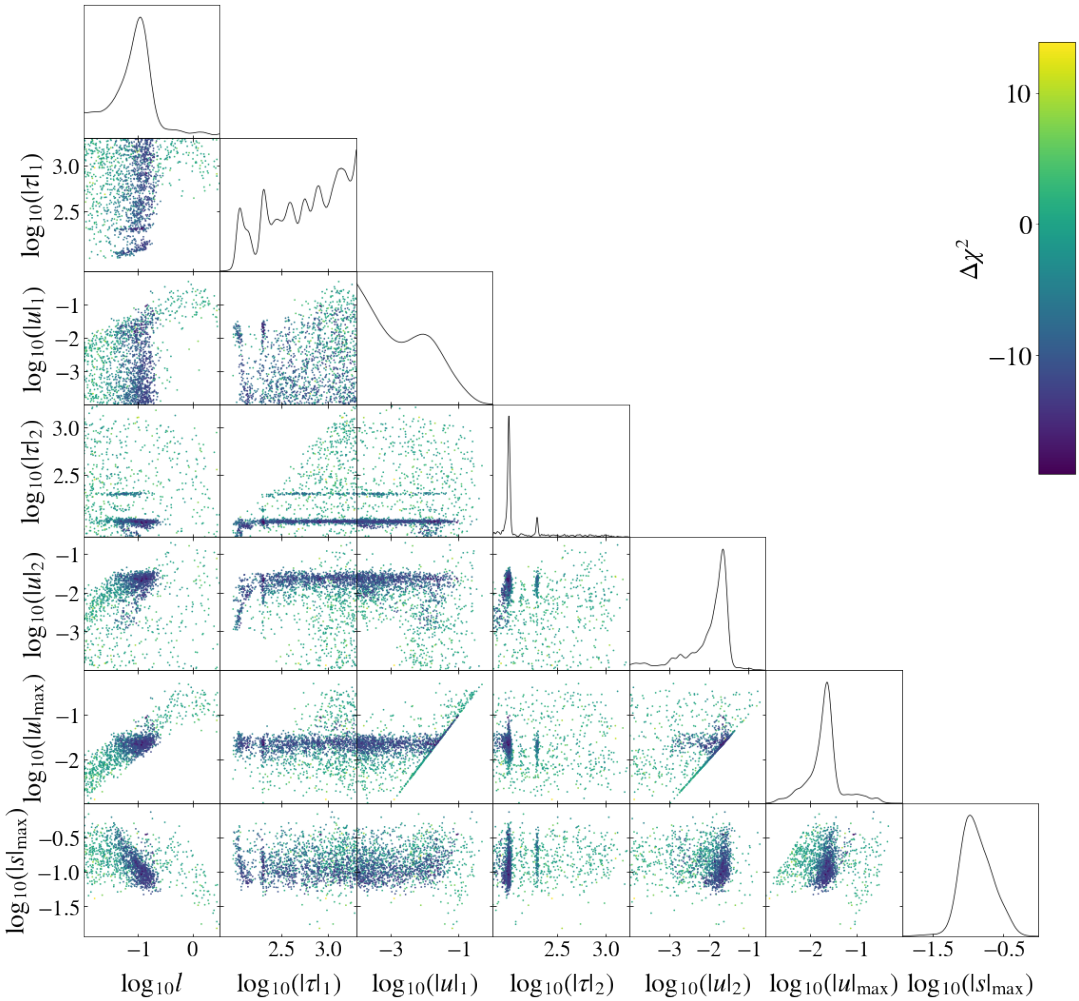

FIG. 4. Posterior distribution for the reconstruction of the speed of sound's profile $u(\tau)$ using 2 training nodes and the logarithmic parametrization shown in Fig. 1. We use $\Delta\chi^2 = \chi^2_{\rm model} - \chi^2_{\rm baseline}$ as the variable for the scatter plot's colour scale, the reference $\chi^2_{\rm baseline}$ corresponding to the MAP of the baseline ΛCDM model to the same datasets. We show the parameters of the training nodes (τ_i, u_k) and the correlation length l (described in III A; priors in III B). We also show the posteriors of the EFT parameters $(|u|_{\rm max}, |s|_{\rm max})$ (described in section II, and not sampled directly, but derived from the nodes parameters). It can be seen how longer correlation lengths (broader reductions) lead to lower values of $|s|_{\rm max}$, and vice-versa. The posterior distributions for different numbers of nodes display similar patterns. For all cases, the posterior distributions are clearly multi-modal.

nodes and stabilizes after that. The profile $u(\tau)$ shows different patterns depending on how many training nodes are used in the GP reconstruction. We have decided to classify all possible profiles $u(\tau)$ based on whether they show differentiated and non-overlapping reductions (that we denominate *dips*) or they present some kind of substructure:

- **One single dip**: usually present at either late values of conformal time (−100, −200, −400), or at earlier times (−800, ∼ −1000). Early-time dips produce features in the CMB power spectra localised in ℓ's up to the first acoustic peak, whereas features from late-time dips affect the power spectra along the full ℓ range. Similar profiles were already found in previous studies (see Sec. IV A). Details on this posterior modes can be found in appendix A 1.
- **Combination of non-overlapping reductions (2, 3 and 4 dips)**: appearing when more than one training node is used, they consist of consecutive, isolated reductions in the speed of sound.[4] Details can be found in appendices A 2, A 3 and A 4. These combinations can be classified as (for details see appendix A):

[4] Notice that the number of training nodes is not always equal to the number of dips: reconstructions with m dips found with m GP nodes usually re-appear as posterior modes in the $m+1$ GP nodes case, where one of the nodes is placed at $u_i \approx 0$.

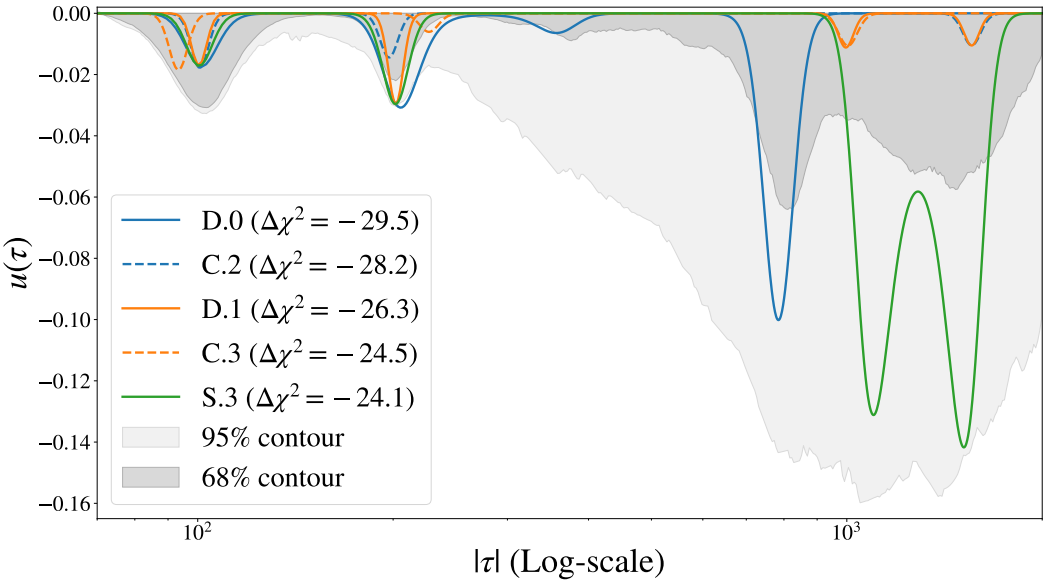

FIG. 5. Reconstruction of the inflaton's speed of sound profile $u(\tau)$ based a 4-nodes GP, where the confidence contours (68% and 95%) are shown. We are able to constrain the shape the inflaton's speed of sound more stringently at late times (up to ~ -200), whereas the confidence intervals get larger at earlier times (i.e: starting from -800). This difference in the constraining power between early and late conformal times is mostly due to early-time reductions being associated to low-multipole features where cosmic variance is largest. The best *maxima a posteriori* are also plotted on top of the confidence contours.

  - *All dips at late conformal times*: when at least two training nodes are considered, there is a preference for two of the possible dips remaining at late-time values of τ_i, combining either -100 and -200, or -400 and -200. Their effect in the CMB power spectrum overlap each other along a large range of ℓ's.
  - *Combination of early- and late-time dips:* these appear typically as a combination of features at both low ℓ's (from the early-time dips) and high ℓ's (from the late time ones), e.g. from the presence of dips both at -800 and -100.

- **Dips with substructure:** We have found some *maxima a posteriori* where the reconstructed $u(\tau)$ does not show clearly separated reductions, but a more complex profile with some degree of substructure. These substructures are presented either at early and late τ_i, trying to fit some of the characteristic features of the CMB angular power spectrum (i.e: $\ell \approx 20 - 40$ feature). The fits are presented in Appendix A, in subsection A 5.

As noticed in previous works [11, 47], we do not have a highly predictive posterior of the maximum of the rate of change of the sound speed, $|s|_{\rm max}$, whose value is mostly constrained by the prior information. By contrast, the positions of the nodes (the oscillation frequency of the features in the power spectrum) are tightly constrained within each of the multiple posterior modes, specially for nodes at late conformal time.

Using the sampling results of the profile $u(\tau)$ with four training nodes in the GPs, we have reconstructed the allowed confidence contours for $u(\tau)$ given Planck 2018 data. The result can be seen in Fig. 5. As expected from the 1-D marginalized posterior distributions, the confidence contours are narrower around $\tau_i = [-100, 200]$. These modes were found in every single reconstruction of the inflaton's speed of sound independently of the number of training nodes (and were also observed in previous studies [11]), and usually show the highest individual dip $\Delta\chi^2$ with respect to ΛCDM (since they produce features at a long range of ℓ for which Planck has low error bars). On the other hand, the confidence contours are broader for earlier conformal times $\tau_i < -400$. This is the range of τ where we have found the modes at $\tau_i = [-800, -1000]$ and some degree of substructure. In this range, the posterior distributions are not very predictive (see again Fig. 4, where the posterior peaks are small for $\tau_i < -800$), since they produce low-multipole features hidden by cosmic variance.

V. CONCLUSIONS

We have searched for features in the primordial power spectrum as given by the last release of Planck 2018 data. Following an EFT of inflation approach, we have focused our search on features coming from reductions of the sound speed of the inflaton, assuming these reductions to be small, mild and transient. These feature templates were not tested by the Planck Collaboration.

We have improved over previous studies (which used a single-reduction Gaussian ansatz) by developing a reconstruction technique for the speed of sound's profile based on Gaussian Processes. We have also marginalised over the slow roll parameters to allow for a dynamical prior. In this new pipeline, the parameters of the reconstruction (the position of the training nodes and the correlation length) are fitted to the Planck 2018 data. The physical constraints of the model are imposed on the reconstruction parameters by means of a Maximum-Entropy prior defined on the EFT quantities ($|u|_{\rm max}, |s|_{\rm max}$), which define the consistency bounds of the model. We have also tracked as derived parameters n_s and r.

This template-free reconstruction of $u(\tau)$ has allowed us to make an exhaustive search of more flexible features' templates, constrained only by EFT conditions. The analysis of the result of Bayesian parameter inference on the Planck 2018 data has demonstrated that there are many possible different and complex $u(\tau)$ profiles which are consistent with Planck's CMB power spectra. Even though none of these fits is preferred with respect to ΛCDM (their $\Delta\chi^2$'s are not significant), we have shown some promising results in terms of interesting feature templates. First, we have argued that there is a strong preference for two consecutive reductions of the speed of sound to coexist at late times around $\tau_i \approx -200$ and $\tau_i \approx -100$. Also, combinations of modes at late conformal time $\tau_i \approx -100$ and early conformal time $\tau_i \approx -800$ are also possible. Secondly, we have found certain profiles which show some degree of sub-structure at early and late conformal times. Finally, we have been able to obtain confidence contours for the $u(\tau)$ profile given the results obtained with four training nodes.

In the future, we plan to exploit this robust and novel pipeline in the search of features using new sets of data (in particular, Large Scale Structure surveys or the CMB bispectrum). The improvement of current data (for example, the polarization of the CMB) will also help to reduce the noise and, therefore, the uncertainty we have at large scales. If the noise is reduced, we could discern how realistic the reductions at earlier conformal times are. Moreover, we also consider introducing new features coming from a variable first slow-roll parameter [18] to perform a joint search of both patterns: features induced by a variable $c_s(\tau)$ and $\epsilon(\tau)$.

ACKNOWLEDGMENTS

We thank Juan Claramunt Gonzalez for useful comments concerning the mathematical notation used in the methodology. We acknowledge Santander Supercomputacion support group at the University of Cantabria who provided access to the supercomputer Altamira Supercomputer at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing simulations/analyses. GCH acknowledges support from the Delta Institute for Theoretical Physics (D-ITP consortium), a program of the Netherlands Organization for Scientific Research (NWO, OWC). JT acknowledges partial support from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement No. [616170]. AA is partially supported by the Netherlands' Organization for Scientific Research (NWO/FOM), by the Basque Government/Eusko Jaurlaritza (IT- 979-16), and by the Spanish Ministry MINECO FPA2015-64041-C2-1P (MINECO/FEDER) and PGC2018-094626-B-C21 (MCIU/AEl/FEDER,UE).

Appendix A: detailed results up to 4 nodes of $u(\tau)$ in log-log parameterization

In this appendix, we explain in detail the several *maxima a posteriori* found during the sampling runs when the profile of the inflaton's speed of sound $u(\tau)$ was reconstructed using Gaussian Processes up to 4 training nodes.

1. One dip (*denoted by "A"*)

These profiles of $u(\tau)$ show only one single reduction of the inflaton's speed of sound or *one dip*. These reductions can be found using just one training node in the Gaussian Process. The modes at late conformal time τ_i (see the first row of Fig. 6) present a well defined oscillation frequency τ_i (at -100, -200 and -400). The value of $|u|_{\rm max}$ is around 0.02 and the rate of change in the speed of sound $|s|_{\rm max} \gg |u|_{\rm max}$. These dips are exactly reproduced with the linear parametrization of the reconstruction of $u(\tau)$. They were already listed during previous searches using Planck 2013 and 2015 data [11, 48]. In particular, the mode corresponding at $\tau_i \approx -400$ was identified faintly in [11]. These modes are present in a broad multipole ℓ range ($\ell \approx [100 - 2000]$), fitting some structures in the temperature and polarization data.

Modes at early conformal time τ_i (-800, -1000) (see the second row of Fig. 6) are found, but are more poorly constrained and with worse $\Delta\chi^2$ with respect to ΛCDM. Mode A.3 at $\tau_i \approx -800$ shows similar characteristics to the modes at a late conformal time (small amplitude and

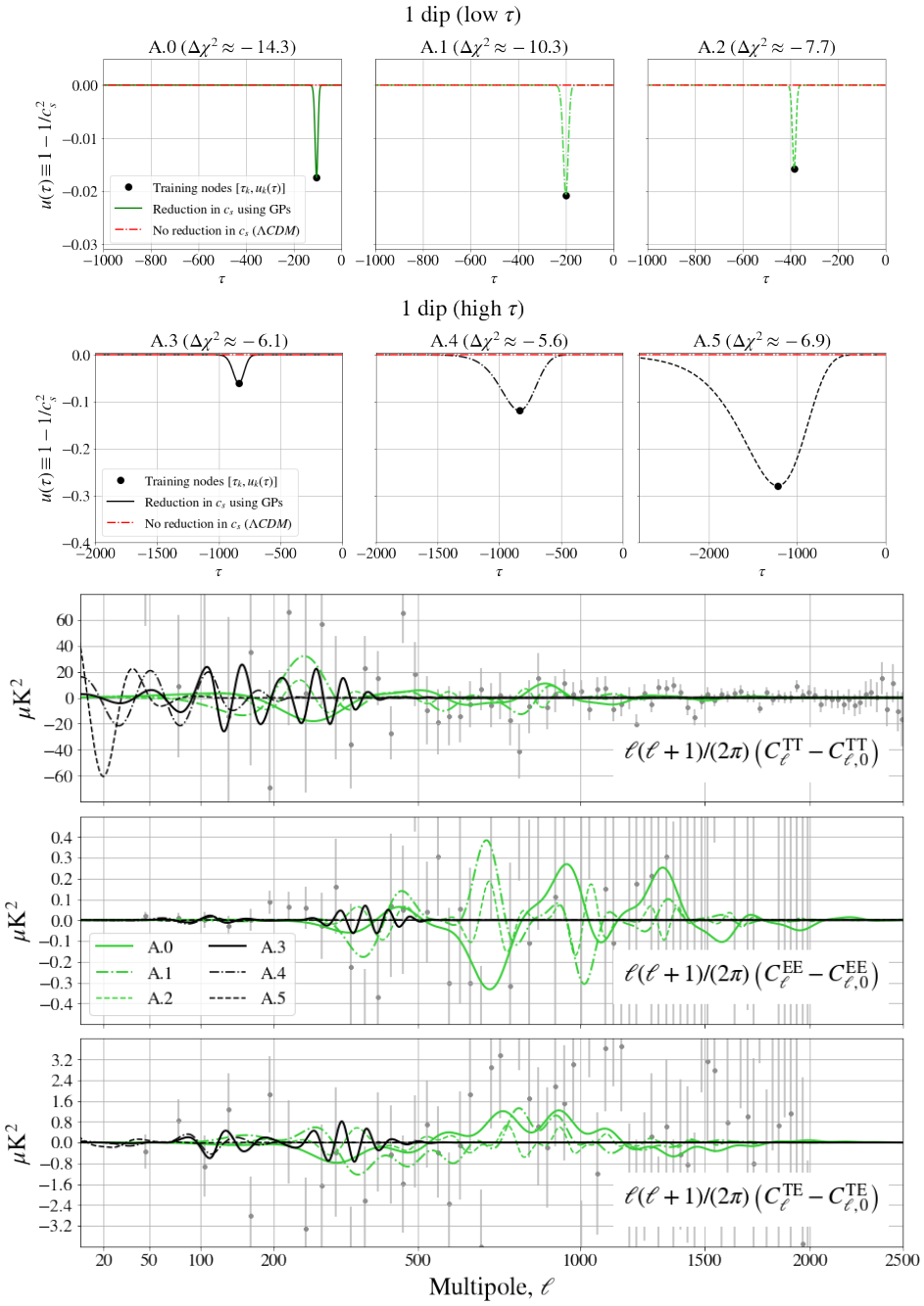

FIG. 6. One single dip, one training node. **Top:** Different profiles $u(\tau)$ for the 6 *maxima a posteriori* when only 1 training node is used (and consequently only one dip is visible). The reconstruction is done following the logarithmic parametrization explained in Sec. III A. We found a principal MAP and 5 other fits when the multimodal posterior distribution is further analysed (see, for example, Fig. 4, where other peaks in the posterior distribution are visible). **Bottom**: Differences in the CMB temperature (TT), E-polarization (EE) and cross-correlated power spectra (TE) between the MAP to the Planck 2018 data and the featureless ΛCDM baseline model for the reconstructed speed of sound profiles $u(\tau)$ A.0 - A.5 shown above. Notice how these profiles fit small deviations from ΛCDM at low and high multipoles ℓ. The same color and line-style correspondence between the $u(\tau)$ profiles and the differences in the CMB spectra has been used.

same behaviour in the EFT parameter s). It fits an apparent oscillating structure of the temperature C_ℓ at the first acoustic peak. This mode is also found with the linear parametrization and in previous searches in [11, 48]. The modes at -800 and -1000, with a larger amplitude, have $|s|_{\max} \approx |u|_{\max}$. They slightly differ from the modes found in previous studies. The main reason is that $u(\tau)$ reconstructed using the logarithmic parametrization dif-

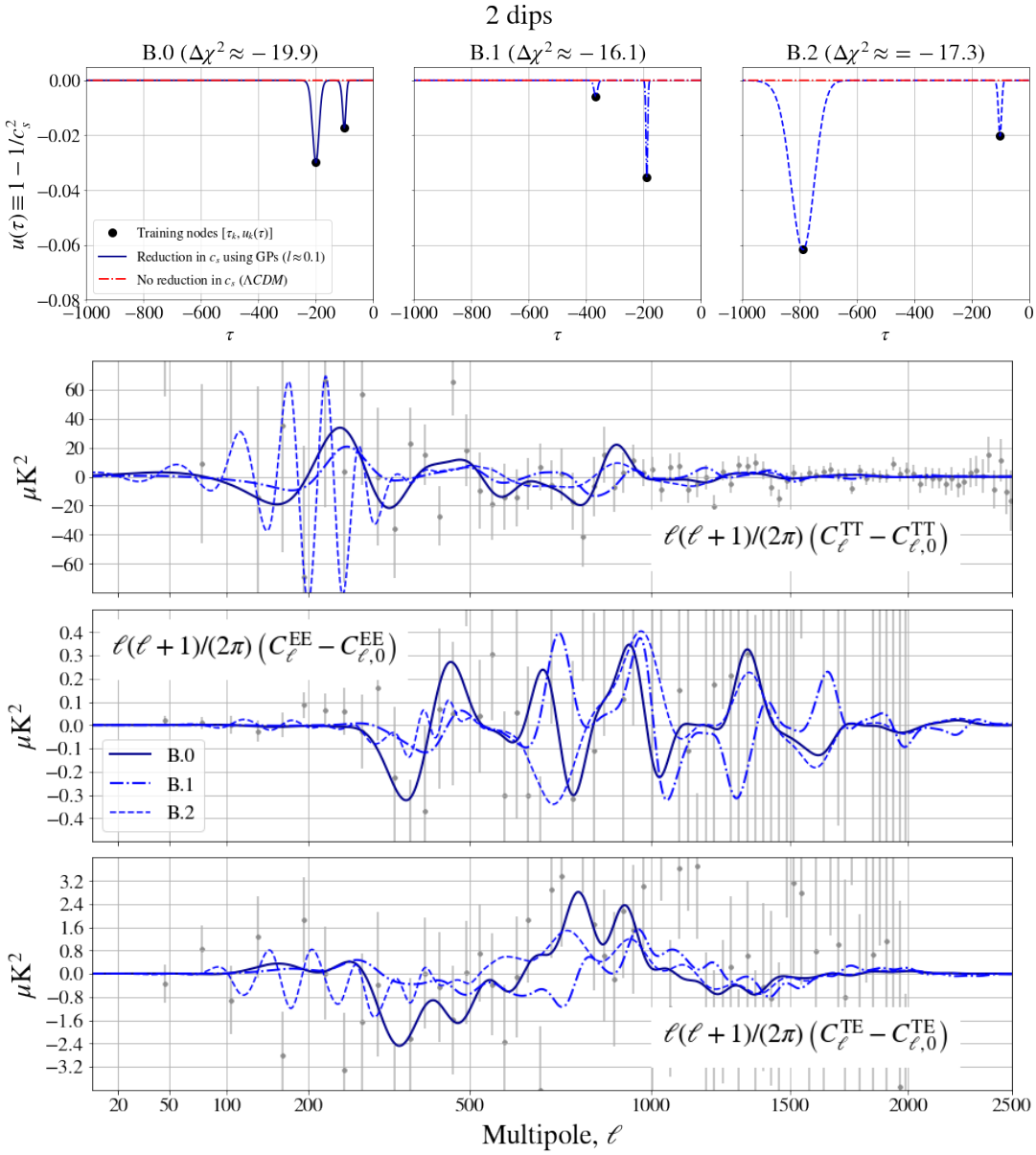

FIG. 7. 2 non-overlapping dips, 2 training nodes. **Top:** Different profiles $u(\tau)$ for the 3 *maxima a posteriori* when only 2 training nodes are used and only two clearly different dips are observed. The reconstruction is done following the logarithmic parametrization explained in Sec. III A. We found a principal best fit and 2 other fits when the multimodal posterior distribution is further studied (see, for example, Fig. 4, where other peaks in the posterior distribution are visible). **Bottom**: Differences in the CMB temperature (TT), E-polarization (EE) and cross-correlated power spectra (TE) between the best fit to the Planck 2018 data and the featureless ΛCDM baseline model for the reconstructed speed's of sound profile $u(\tau)$ shown above. Notice how these profiles fit small deviations from ΛCDM at low and high multipoles ℓ. The same color and line-style correspondence between the $u(\tau)$ profiles and the differences in the CMB spectra has been used.

fers from the linear one at high values of τ_i. The modes at -1000 try to fit the characteristic $\ell \approx 20-40$ structure of the CMB temperature angular power spectrum. As identified in [11], this kind of features impose a tighter upper limit on the scalar-to-tensor ratio r (see the tables on this section), although these are still within the analogous bounds in ΛCDM.

2. Two dips (*denoted by "B"*)

In this case, the inflaton would suffer two consecutive reductions of the speed of sound (due to, for instance, two consecutive turns in the field space). In the previous study using Planck 2015 [11], it was pointed out that, a priori, the features due to single reductions of the sound speed that do not overlap can in principle co-exist. These

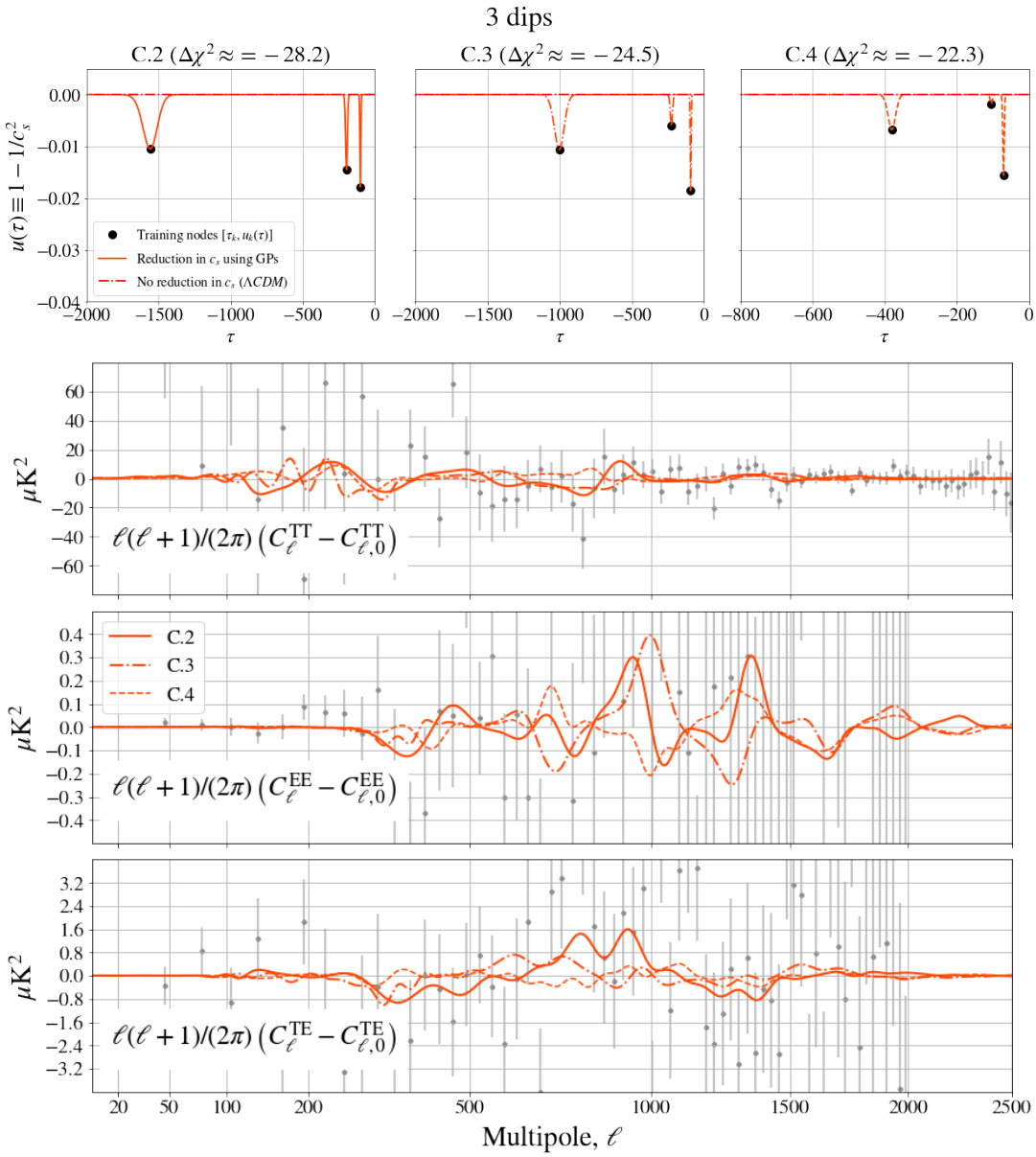

FIG. 8. 3 non-overlapping dips, 3 training nodes. **Top:** Different profiles $u(\tau)$ for the 3 *maxima a posteriori* when only 3 training nodes are used and three differentiated dips are observed. The reconstruction is done following the logarithmic parametrization explained in Sec. III A. We found 3 MAP when the corresponding multimodal posterior distribution is further studied (see, for example, Fig. 4, where other peaks in the posterior distribution are visible). **Bottom**: Differences in the CMB temperature (TT), E-polarization (EE) and cross-correlated power spectra (TE) between the best fit to the Planck 2018 data and the featureless ΛCDM baseline model for the reconstructed speed's of sound profile $u(\tau)$ shown above. Notice how these profiles fit small deviations from ΛCDM at low and high multipoles ℓ. The same color and line-style correspondence between the $u(\tau)$ profiles and the differences in the CMB spectra has been used.

combinations would be modes at early τ_i with another late mode (i.e: -1000 and -100). However, the results of the reconstruction using two training nodes show a richer picture (see Fig. 7). We have identified that the modes at -100 and -200 can result from an overlapping feature that is preferred by the data, and thus, it is the overall MAP for the 2-nodes reconstruction. The dips at -200 and -400 can also co-exist (mode B.1), with a worse $\Delta\chi^2$. These two overlapping features fit TT, TE and EE structure across a large range of ℓ. On the other hand, there is a possible combination of the modes at -800 and -100 (see the first row of Fig. 7). This $u(\tau)$ profile includes the fitting of the apparent oscillations around the first acoustic peak and small deviations across the rest of the multipole scale. All of these combinations of modes fulfill $|s|_{\rm max} \gg |u|_{\rm max}$.

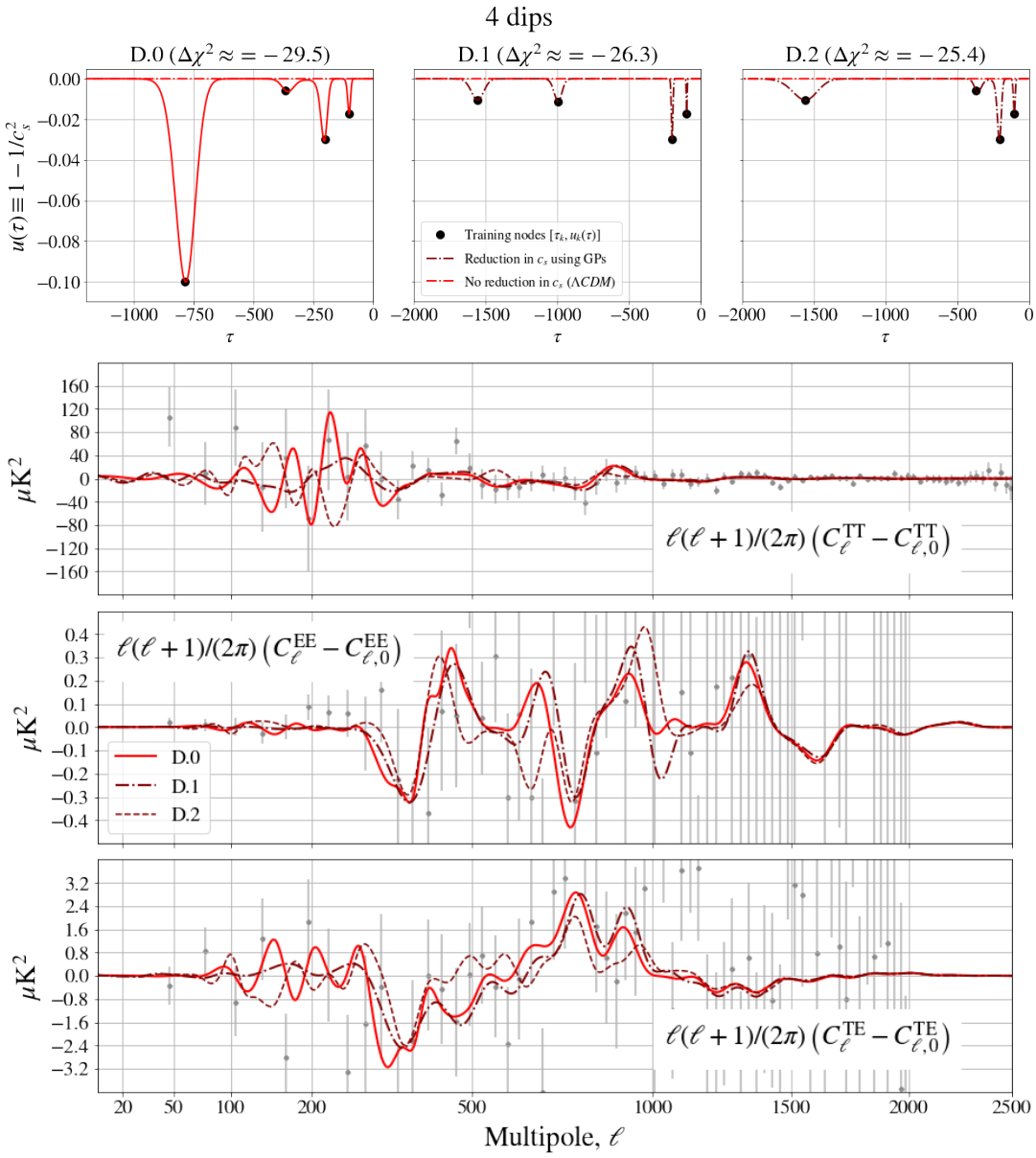

FIG. 9. 4 non-overlapping dips, 4 training nodes. **Top:** Different profiles $u(\tau)$ for the 3 *maxima a posteriori* when only 4 training nodes are used and four differentiated dips are observed. The reconstruction is done following the logarithmic parametrization explained in Sec. III A. In this case, we observe how possible reductions at $\tau \approx -100, -200, -400$ and -800 can consecutively take place. **Bottom**: Differences in the CMB temperature (TT), E-polarization (EE) and cross-correlated power spectra (TE) between the best fit to the Planck 2018 data and the featureless ΛCDM baseline model for the reconstructed speed's of sound profile $u(\tau)$ shown above. Notice how these profiles fit small deviations from ΛCDM at low and high multipoles ℓ. The same color and line-style correspondence between the $u(\tau)$ profiles and the differences in the CMB spectra has been used.

3. Three dips (*denoted by "C"*)

When the profile of the inflaton's reduction of the speed of sound $u(\tau)$ is reconstructed using 3 training nodes, we find more complex profiles. The usual dips at late τ_i around -100 and -200 combine to mild and small reductions at early conformal times around -1000 and -1500 (see modes C.2 and C.3 of Fig. A 3, respectively). The combinations that are preferred by the data are those whose earlier training node τ_i is placed around $u_k \approx -0.01$. These small dips at early τ_i were not found alone during the search using one training node (modes A shown in Fig. 6). However, these modes at early conformal times are loosely constrained and the confidence intervals are large. Overall, these profiles fits the data very similarly as the single standing modes A.0 and A.1, as well as the combination B.0. Furthermore, we have also verified that the modes A.0 and A.2 can co-exist (con-

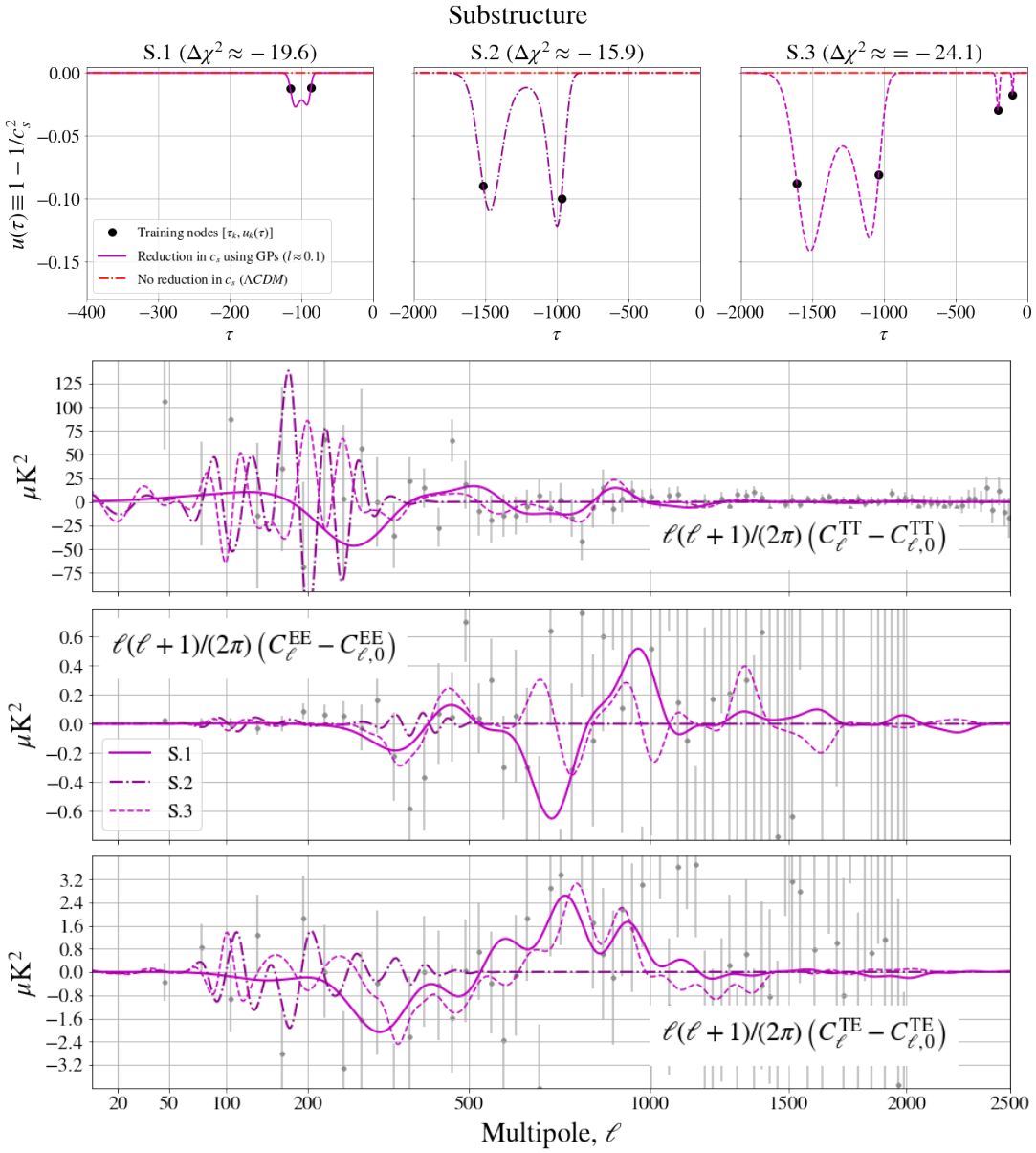

FIG. 10. Profiles with substructure (2 training nodes and 4 training nodes). **Top:** Different profiles $u(\tau)$ for 2 *maxima a posteriori* when only 2 or training nodes are used and the profiles of $u(\tau)$ show some grade of substructure. The reconstruction is done following the logarithmic parametrization explained in Sec. III A. We found 2 fits (one an earlier conformal time and a another one at late conformal time), when the corresponding multimodal posterior distributiosn are further studied (see, for example, Fig. 4, where other peaks in the posterior distribution are visible). **Bottom**: Differences in the CMB temperature (TT), E-polarization (EE) and cross-correlated power spectra (TE) between the best fit to the Planck 2018 data and the featureless ΛCDM baseline model for the reconstructed speed's of sound profile $u(\tau)$ shown above. Notice how these profiles fit small deviations from ΛCDM at low and high multipoles ℓ. The same color and line-style correspondence between the $u(\tau)$ profiles and the differences in the CMB spectra has been used.

trary to the case of reconstruction with 2 training nodes, where the combination -100 and -400 was not found) if a very small mode is added close to -100. All these profiles presented in Fig. 8 are in the limit $|s|_{\max} \gg |u|_{\max}$. Finally, the 2-dip profiles explained above are also reproduced when we run with 3 training nodes.

4. Four dips (*denoted by "D"*)

We find similar profiles for $u(\tau)$ as the ones for 3 training nodes, adding one extra dip and finding the remaining possible combination of nodes at late τ_i around -100, -200 and -400 with modes at earlier times at -800, -1000 or -1500 (see modes D.0, D.1 and D.3 of

Fig. 9). Thus, we have corroborated that the modes A.0, A.1, A.2 and A.4 can, in principle, co-exist. Still, the small modes at early $\tau_i \approx -1000, -1500$ are less likely to show up in the posterior, as they are poorly constrained given the data. It is worth mentioning that we can mostly reproduce all the different profiles found during the search using up to three nodes when four training nodes are used. In this case, one, two or three training nodes are placed in such a way the corresponding profile looks very similar to the cases A, B or C (either the training node is placed close to $u_i \approx 0$ or close to the previous training mode itself).

5. Reductions with substructure (*denoted by "S"*)

Apart from concatenations of transient reductions in the speed of sound, we have also observed some possible fits which show more complicated feature patterns according to the data. These are profiles of $u(\tau)$ that do not clearly show full dips but have some kind of substructure (see the upper row of Fig. 10, profile S.1). We find a sub-structured maxima a posteriori at a late conformal time (centred around $\tau_i \approx -100$), which resembles the mode A.0 but with two small sub-reductions. Similarly to A.0, the limit of the EFT functions is $|s|_{\max} \gg |u|_{\max}$.

Motivated by the loose constraints of the training nodes in the range of early conformal times $-800 < \tau_i < -3000$, we have launched a GPs reconstruction using two training nodes, which are restricted to remain in the range $-3500 < \tau_i < -990$, obtaining the profile S.2 of Fig. 10. This profile have the particularity that $|s|_{\max} \ll |u|_{\max}$. This profile tries to fit not only the structure of the CMB TT angular power spectrum around $\ell \approx 20 - 40$ but also the apparent structure of the first acoustic peak in the TT and TE data. To answer the question if it is possible the substructure mode S.2 to co-exist with any of the modes at late conformal time τ_i, we have relaunched the GPs reconstruction with 4 training nodes, constraining the two early ones in the range $-3500 < \tau_1 < \tau_2 < -990$. When this constraint is imposed, the later training nodes τ_3 and τ_4 are placed clearly around -100 and -200. The resulting profile (see mode S.3 in Fig. 10) can fit the CMB data at low ℓ but also in a broader range similarly to the case of the mode A.0, increasing the statistical significance $\Delta\chi^2$. In this case, the EFT limit is reverted to $|s|_{\max} \gg |u|_{\max}$ due to the narrow mode at -100.

[1] J. Chluba, J. Hamann, and S. P. Patil, International Journal of Modern Physics D **24**, 1530023 (2015), arXiv:1505.01834.

[2] X. Chen, Advances in Astronomy **2010**, 638979 (2010), arXiv:1002.1416 [astro-ph.CO].

[3] A. Slosar, X. Chen, C. Dvorkin, D. Meerburg, B. Wallisch, D. Green, and E. Silverstein, Bulletin of the AAS **51** (2019), https://baas.aas.org/pub/2020n3i098.

[4] G. A. Palma, JCAP **04**, 035 (2015), arXiv:1412.5615 [hep-th].

[5] A. Achúcarro, J.-O. Gong, G. A. Palma, and S. P. Patil, Phys. Rev. D **87**, 121301 (2013), arXiv:1211.5619 [astro-ph.CO].

[6] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, and N. Bartolo, arXiv e-prints , arXiv:1807.06209 (2018), arXiv:1807.06209 [astro-ph.CO].

[7] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, and N. Bartolo, arXiv e-prints , arXiv:1807.06211 (2018), arXiv:1807.06211 [astro-ph.CO].

[8] A. Achúcarro, V. Atal, P. Ortiz, and J. Torrado, Phys. Rev. D **89**, 103006 (2014), arXiv:1311.2552.

[9] A. Achúcarro, V. Atal, B. Hu, P. Ortiz, and J. Torrado, Phys. Rev. D **90**, 023511 (2014), arXiv:1404.7522.

[10] B. Hu and J. Torrado, Phys. Rev. D **91**, 064039 (2015), arXiv:1410.4804 [astro-ph.CO].

[11] J. Torrado, B. Hu, and A. Achúcarro, Phys. Rev. D **96**, 083515 (2017), arXiv:1611.10350.

[12] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, and et al., Astronomy and Astrophysics **594**, A20 (2016), arXiv:1502.02114.

[13] D. K. Hazra, A. Shafieloo, and T. Souradeep, Journal of Cosmology and Astroparticle Physics **2013**, 031 (2013), arXiv:1303.4143 [astro-ph.CO].

[14] P. Hunt and S. Sarkar, Journal of Cosmology and Astroparticle Physics **2014**, 025 (2014), arXiv:1308.2317 [astro-ph.CO].

[15] A. Ravenni, L. Verde, and A. J. Cuesta, Journal of Cosmology and Astroparticle Physics **2016**, 028 (2016), arXiv:1605.06637 [astro-ph.CO].

[16] A. Durakovic, P. Hunt, S. Mukherjee, S. Sarkar, and T. Souradeep, Journal of Cosmology and Astroparticle Physics **2018**, 012 (2018), arXiv:1711.08441 [astro-ph.CO].

[17] M. Ballardini, Physics of the Dark Universe **23**, 100245 (2019), arXiv:1807.05521 [astro-ph.CO].

[18] A. Durakovic, P. Hunt, S. P. Patil, and S. Sarkar, SciPost Phys. **7**, 049 (2019), arXiv:1904.00991 [astro-ph.CO].

[19] W. J. Handley, A. N. Lasenby, H. V. Peiris, and M. P. Hobson, Phys. Rev. D **100**, 103511 (2019), arXiv:1908.00906 [astro-ph.CO].

[20] S. Appleby, J.-O. Gong, D. K. Hazra, A. Shafieloo, and S. Sypsas, Physics Letters B **760**, 297 (2016), arXiv:1512.08977 [astro-ph.CO].

[21] W. Handley and M. Millea, Entropy **21**, 272 (2019), arXiv:1804.08143 [math.ST].

[22] S. Gariazzo and O. Mena, Phys. Rev. D **99**, 021301 (2019), arXiv:1812.05449 [astro-ph.CO].

[23] N. Aghanim *et al.* (Planck), (2018), arXiv:1807.06209

[astro-ph.CO].
[24] C. Cheung, A. L. Fitzpatrick, J. Kaplan, L. Senatore, and P. Creminelli, Journal of High Energy Physics **3**, 014-014 (2008), arXiv:0709.0293 [hep-th].
[25] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, JHEP **05**, 066 (2012), arXiv:1201.6342 [hep-th].
[26] A. Achúcarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, Journal of Cosmology and Astroparticle Physics **1**, 030 (2011), arXiv:1010.3693 [hep-ph].
[27] A. Achúcarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, Phys. Rev. D **84**, 043502 (2011), arXiv:1005.3848 [hep-th].
[28] A. Achúcarro, V. Atal, S. Céspedes, J.-O. Gong, G. A. Palma, and S. P. Patil, Phys. Rev. D **86**, 121301 (2012).
[29] S. Weinberg, Phys. Rev. D **72**, 043514 (2005).
[30] S. Riemer-Sorensen, "LCDM and Beyond: Cosmology Tools in Theory and in Practice: "Statistics and model selection in cosmology"," `http://icg.port.ac.uk/~jschewts/cantata/L5/Statistics_Notes.pdf` (2018), accessed: 2018-05-18.
[31] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (The MIT Press, 2005).
[32] "Scikit-learn 0.19.1 documentation: Gaussian Processes," `http://scikit-learn.org/stable/modules/gaussian_process.html` (2018), accessed: 10-06-2019.
[33] A. J. S. Hamilton, Mon. Not. Roy. Astron. Soc. **312**, 257 (2000), arXiv:astro-ph/9905191 [astro-ph].
[34] "fftlog - python wrapper FFTLog," `https://github.com/prisae/fftlog` (2016), accessed: 01-09-2019.
[35] S. Céspedes, V. Atal, and G. A. Palma, Journal of Cosmology and Astroparticle Physics **5**, 008 (2012), arXiv:1201.4848 [hep-th].
[36] P. Adshead and W. Hu, Phys. Rev. D **89**, 083531 (2014), arXiv:1402.1677 [astro-ph.CO].
[37] D. Cannone, N. Bartolo, and S. Matarrese, Phys. Rev. D **89**, 127301 (2014), arXiv:1402.2258 [astro-ph.CO].
[38] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. Roy. Astron. Soc. **450**, L61 (2015), arXiv:1502.01856.
[39] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. Roy. Astron. Soc. **453**, 4384 (2015), arXiv:1506.00171 [astro-ph.IM].
[40] A. Lewis, (2019), arXiv:1910.13970 [astro-ph.IM].
[41] A. Lewis, Phys. Rev. **D87**, 103529 (2013), arXiv:1304.4473 [astro-ph.CO].
[42] J. Torrado and A. Lewis, arXiv e-prints , arXiv:2005.05290 (2020), arXiv:2005.05290 [astro-ph.IM].
[43] C. Cartis, J. Fiala, B. Marteau, and L. Roberts, arXiv e-prints , arXiv:1804.00154 (2018), arXiv:1804.00154 [math.OC].
[44] C. Cartis, L. Roberts, and O. Sheridan-Methven, arXiv e-prints , arXiv:1812.11343 (2018), arXiv:1812.11343 [math.OC].
[45] M. J. D. Powell, Technical Report, Department of Applied Mathematics and Theoretical Physics (2009).
[46] H. Akaike, IEEE Transactions on Automatic Control **19**, 716 (1974).
[47] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. Roy. Astron. Soc. **453**, 4384 (2015), arXiv:1506.00171 [astro-ph.IM].
[48] B. Hu and J. Torrado, Phys. Rev. D **91**, 064039 (2015), arXiv:1410.4804.

Name	**A.0** ($\Delta\chi^2 = -14.3$)			**B.0** ($\Delta\chi^2 = -19.9$)			**C.2** ($\Delta\chi^2 = -28.2$)			**D.0** ($\Delta\chi^2 = -29.5$)		
Parameters	Center	lower	upper	Center	lower	upper	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-0.972	-1.997	0.300	-0.995	-1.989	0.282	-1.231	-1.211	-0.754	-0.994	-1.165	-0.917
$\log_{10}(\lvert\tau_1\rvert)$	2.014	1.959	2.299	2.305	1.993	3.299	3.192	2.834	3.297	2.896	2.397	3.100
$\log_{10}(\lvert u_1\rvert)$	-1.760	-2.687	-0.271	-1.528	-3.999	-0.593	-1.979	-1.952	-0.851	-1.101	-3.997	-0.597
$\log_{10}(\lvert\tau_2\rvert)$	-	-	-	2.002	1.845	3.145	2.295	2.237	2.395	2.564	2.285	3.301
$\log_{10}(\lvert u_2\rvert)$	-	-	-	-1.764	-3.843	-0.977	-1.840	-3.958	-1.495	-2.229	-3.987	-0.404
$\log_{10}(\lvert\tau_3\rvert)$	-	-	-	-	-	-	2.003	1.963	2.030	2.564	2.492	2.629
$\log_{10}(\lvert u_3\rvert)$	-	-	-	-	-	-	-1.749	-2.039	-1.495	-2.229	-1.854	-0.741
$\log_{10}(\lvert\tau_4\rvert)$	-	-	-	-	-	-	-	-	-	2.002	1.983	3.310
$\log_{10}(\lvert u_4\rvert)$	-	-	-	-	-	-	-	-	-	-1.764	-2.693	-0.256
$\log_{10}(\lvert u\rvert_{\max})$	-1.462	-2.694	-0.282	-1.523	-2.587	-0.500	-1.621	-1.810	-0.851	-0.998	-1.831	-0.804
$\log_{10}(\lvert s\rvert_{\max})$	-0.768	-1.442	-0.235	-0.813	-1.352	-0.312	-0.928	-1.084	-0.206	-0.266	-1.731	-0.471
ϵ_1	0.000	0.000	0.007	0.000	0.000	0.006	0.001	0.000	0.005	0.001	0.000	0.005
ϵ_2	0.036	0.021	0.039	0.036	0.025	0.040	0.035	0.027	0.037	0.032	0.025	0.038
n_s	0.964	0.960	0.968	0.963	0.959	0.968	0.963	0.961	0.967	0.963	0.963	0.968
r	0.002	0.002	0.111	0.005	0.002	0.094	0.015	0.002	0.078	0.014	0.002	0.069

TABLE I. *Maxima a posteriori* values and 68% confidence intervals of the feature and primordial parameters for the cases where $|\Delta\chi^2|$ is the largest, when a minimizer method is used. The correspond to 1, 2, 3 and 4 training nodes used in the GPs reconstruction, respectively.

Number of nodes i	**1 training node** (late conformal time)					
Modes	A.1 ($\Delta\chi^2 = -10.3$)			A.2 ($\Delta\chi^2 = -7.7$)		
Parameters	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-1.012	-1.942	-0.762	-1.445	-1.996	-0.983
$\log_{10}(\lvert\tau_1\rvert)$	2.299	2.202	2.300	2.583	2.501	2.639
$\log_{10}(\lvert u_1\rvert)$	-1.682	-2.304	-1.410	-1.804	-2.168	-0.870
$\log_{10}(\lvert u\rvert_{\max})$	-1.592	-2.313	-1.412	-1.814	-2.777	-0.282
$\log_{10}(\lvert s\rvert_{\max})$	-0.727	-0.999	-0.275	-0.617	-1.442	-0.304
ϵ_1	0.000	0.000	0.004	0.0001	0.0001	0.0043
ϵ_2	0.036	0.027	0.039	0.035	0.026	0.038
n_s	0.964	0.961	0.967	0.965	0.961	0.968
r	0.003	0.002	0.058	0.002	0.002	0.068

TABLE II. 1 node secondary *maxima a posteriori* values, at low conformal time, and 68% confidence intervals of the feature and primordial parameters when a minimizer method is used.

Number of nodes i	**1 training node** (early conformal time)								
Modes	A.3 ($\Delta\chi^2 = -6.1$)			A.4 ($\Delta\chi^2 = -5.6$)			A.5 ($\Delta\chi^2 = -6.9$)		
Parameters	Center	lower	upper	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-0.988	-1.642	-0.661	-0.488	-1.066	0.050	-0.233	-1.141	0.328
$\log_{10}(\lvert\tau_1\rvert)$	2.923	2.402	3.100	2.931	2.804	2.995	3.087	3.001	3.299
$\log_{10}(\lvert u_1\rvert)$	-1.215	-2.976	-0.807	-1.210	-1.470	-0.472	-0.554	-1.457	-0.270
$\log_{10}(\lvert u\rvert_{\max})$	-1.169	-1.880	-0.797	-1.367	-1.470	-0.474	-0.554	-1.459	-0.282
$\log_{10}(\lvert s\rvert_{\max})$	-0.473	-1.231	-0.340	-0.702	-1.371	-0.406	-0.614	-1.442	-0.329
ϵ_1	0.000	0.000	0.005	0.0001	0.0001	0.0035	0.0001	0.0001	0.0041
ϵ_2	0.036	0.026	0.040	0.037	0.028	0.037	0.036	0.026	0.038
n_s	0.964	0.960	0.966	0.963	0.961	0.967	0.963	0.961	0.967
r	0.004	0.002	0.068	0.002	0.002	0.055	0.002	0.002	0.065

TABLE III. 1 node secondary *maxima a posteriori* values and 68% confidence intervals of the feature and primordial parameters when a minimizer method is used.

Number of nodes i	**2 training nodes** (other fits)					
Modes	B.1 ($\Delta\chi^2 = -16.1$)			B.2 ($\Delta\chi^2 = -17.3$)		
Parameters	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-1.522	-1.922	0.292	-0.999	-1.462	-0.618
$\log_{10}(\lvert\tau_1\rvert)$	2.565	2.296	3.297	2.896	2.400	3.200
$\log_{10}(\lvert u_1\rvert)$	-2.230	-3.987	-0.404	-1.210	-3.997	-0.797
$\log_{10}(\lvert\tau_2\rvert)$	2.277	2.213	2.300	2.002	1.970	2.033 5
$\log_{10}(\lvert u_2\rvert)$	1.452	-3.895	-1.424	-1.693	-2.240	-1.426
$\log_{10}(\lvert u\rvert_{\max})$	-1.591	-2.299	-0.385	-1.169	-1.980	-0.803
$\log_{10}(\lvert s\rvert_{\max})$	-0.613	-1.279	-0.381	-0.468	-1.226	-0.318
ϵ_1	0.000	0.000	0.005	0.000	0.000	0.005
ϵ_2	0.037	0.026	0.037	0.036	0.026	0.039
n_s	0.963	0.961	0.967	0.964	0.960	0.968
r	0.003	0.002	0.076	0.005	0.002	0.082

TABLE IV. 2 nodes secondary *maxima a posteriori* values and 68% confidence intervals of the feature and primordial parameters, when a minimizer method is used.

Number of nodes i	**3 training nodes** (other fits)														
Modes	C.0 ($\Delta\chi^2 = -21.4$)			C.1 ($\Delta\chi^2 = -20.7$)			C.3 ($\Delta\chi^2 = -24.5$)			C.4 ($\Delta\chi^2 = -22.3$)			C.5 ($\Delta\chi^2 = -17.1$)		
parameters	Center	lower	upper	Center	lower	upper	Center	lower	upper	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-0.979	-1.634	-0.559	-1.166	-1.514	-0.646	-1.208	-1.209	-0.821	-1.297	-1.475	-0.721	-0.909	-1.443	-0.690
$\log_{10}(\lvert\tau_1\rvert)$	2.940	2.214	3,300	2.253	2.214	3.286	3.002	2.840	2.998	2.580	2.405	2.599	2.903	2.814	2.997
$\log_{10}(\lvert u_1\rvert)$	-3.205	-4.000	-0.825	-1.658	-3.989	-0.939	-1.974	-1.952	-1.288	-2.169	-3.931	-1.507	-1.205	-3.953	-0.971
$\log_{10}(\lvert\tau_2\rvert)$	2.309	1.969	3.261	2.080	1.969	2.100	2.356	2.237	2.395	2.019	2.014	2.506	2.078	1.981	2.100
$\log_{10}(\lvert u_2\rvert)$	-1.486	-3.998	-1.120	-3.487	-3.965	-1.572	-2.216	-3.724	-1.965	-2.741	-3.303	-1.560	-2.735	-3.965	-1.572
$\log_{10}(\lvert\tau_3\rvert)$	2.027	1.846	2.318	2.003	1.901	2.064	1.970	1.963	2.018	1.846	1.846	2.065	1.945	1.915	2.064
$\log_{10}(\lvert u_3\rvert)$	-1.822	-3.261	-1.415	-2.185	-2.860	-1.559	-1.733	-2.019	-1.565	-1.807	-3.261	-1.513	-2.296	-2.822	-1.629
$\log_{10}(\lvert u\rvert_{\max})$	-1.462	-2.046	-0.473	-1.633	-1.878	-0.939	-1.697	-1.697	-1.288	-1.584	-1.844	-1.409	-1.205	-1.878	-0.971
$\log_{10}(\lvert s\rvert_{\max})$	-0.768	-1.287	-0.206	-0.925	-1.216	-0.231	-0.808	-0.994	-0.516	-0.872	-1.151	-0.410	-0.547	-1.216	-0.231
ϵ_1	0.000	0.000	0.005	0.002	0.000	0.005	0.001	0.001	0.003	0.000	0.000	0.002	0.002	0.000	0.005
ϵ_2	0.036	0.024	0.039	0.033	0.025	0.039	0.035	0.031	0.036	0.036	0.031	0.039	0.033	0.025	0.038
n_s	0.963	0.960	0.968	0.964	0.960	0.968	0.963	0.962	0.964	0.964	0.961	0.965	0.963	0.961	0.966
r	0.004	0.002	0.087	0.025	0.002	0.082	0.018	0.018	0.044	0.003	0.002	0.039	0.028	0.002	0.081

TABLE V. 3 nodes secondary *maxima a posteriori* values and 68% confidence intervals of the feature and primordial parameters, when a minimizer method is used.

Number of nodes i	**4 training nodes** (other fits)					
Modes	D.1 ($\Delta\chi^2 = -26.3$)			D.2 ($\Delta\chi^2 = -25.4$)		
Parameters	Center	lower	upper	Center	lower	upper
$\log_{10} l$	-1.294857	-1.211	-0.751	-0.994857	-1.522	-0.613
$\log_{10}(\lvert\tau_1\rvert)$	3.192	2.934	3.298	3.191	2.928	3.293
$\log_{10}(\lvert u_1\rvert)$	-1.978718905	-1.952	-0.851	-1.979	-1.952	-0.851
$\log_{10}(\lvert\tau_2\rvert)$	2.99835	2.840	3.150	2.565	2.296	2.745
$\log_{10}(\lvert u_2\rvert)$	-1.95155	-3.294	-1.120	-2.230	-3.987	-0.404
$\log_{10}(\lvert\tau_3\rvert)$	2.304	2.201	2.323	2.305	2.203	2.326
$\log_{10}(\lvert u_3\rvert)$	-1.528	-2.654	-0.262	-1.528	-2.662	-0.261
$\log_{10}(\lvert\tau_4\rvert)$	2.002	1.951	2.297	2.002	1.951	2.297
$\log_{10}(\lvert u_4\rvert)$	-1.765	-2.691	-0.270	-1.763	-2.691	-0.272
$\log_{10}(\lvert u\rvert_{\max})$	-1.528	-2.223	-1.312	-1.511	-2.334	-1.422
$\log_{10}(\lvert s\rvert_{\max})$	-0.484	-0.998	-0.196	-0.845	-1.322	-0.231
ϵ_1	0.000	0.000	0.004	0.000	0.000	0.005
ϵ_2	0.035	0.024	0.040	0.036	0.027	0.036
n_s	0.963	0.961	0.966	0.964	0.960	0.967
r	0.004	0.002	0.062	0.002	0.002	0.060

TABLE VI. 4 node secondary *maxima a posteriori* values and 68% confidence intervals of the feature and primordial parameters when a minimizer method is used.

	MAP with substructure								
Modes	S.1 ($\Delta\chi^2 = -19.6$)			S.2 ($\Delta\chi^2 = -15.9$)			S.3 ($\Delta\chi^2 = -24.1$)		
Parameters	Center	lower	upper	Center	lower	upper	Center	lower	upper
$\log_{10} l$	1.228	-1.415	-0.855	-1.101	-1.999	0.499	-1.098	-1.939	0.501
$\log_{10}(\lvert\tau_1\rvert)$	2.062	1.993	2.100	3.180	2.998	3.400	3.180	2.999	3.400
$\log_{10}(\lvert u_1\rvert)$	-1.900	-3.907	-1.555	-1.048	-2.987	-0.390	-1.047	-2.950	-0.394
$\log_{10}(\lvert\tau_2\rvert)$	1.936	1.901	2.028	2.985	2.970	3.395	2.986	2.967	3.391
$\log_{10}(\lvert u_2\rvert)$	-1.937	-3.178	-1.605	-1.001	-2.973	-0.365	-1.000	-2.974	-0.361
$\log_{10}(\lvert\tau_3\rvert)$	-	-	-	-	-	-	2.299	2.202	2.300
$\log_{10}(\lvert u_3\rvert)$	-	-	-	-	-	-	-1.682	-2.304	-1.410
$\log_{10}(\lvert\tau_4\rvert)$	-	-	-	-	-	-	2.003	1.961	2.030
$\log_{10}(\lvert u_4\rvert)$	-	-	-	-	-	-	-1.749	-2.039	-1.495
$\log_{10}(\lvert u\rvert_{\max})$	-1.610	-1.875	-1.451	-0.876	-2.758	-0.355	-0.908	-2.758	-0.355
$\log_{10}(\lvert s\rvert_{\max})$	-0.799	-1.170	-0.420	-1.414	-1.775	-0.334	-0.138	-1.775	-0.334
ϵ_1	0.001	0.000	0.005	0.0002	0.0001	0.0035	0.0001	0.0000	0.0036
ϵ_2	0.034	0.028	0.037	0.036	0.029	0.038	0.034	0.024	0.037
n_s	0.964	0.960	0.968	0.965	0.962	0.967	0.964	0.961	0.967
r	0.002	0.002	0.095	0.010	0.002	0.074	0.004	0.002	0.056

TABLE VII. *Maxima a posteriori* values, for those nodes which show some degree of sub-structure, and 68% confidence intervals of the feature and primordial parameters, when a minimizer method is used.