Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **421,** 369–380 (2012) doi:10.1111/j.1365-2966.2011.20311.x

*WMAP*7 constraints on oscillations in the primordial power spectrum

P. Daniel Meerburg,[1,2⋆] Ralph A. M. J. Wijers[1,2⋆] and Jan Pieter van der Schaar[1,3⋆]

[1]*Gravitation and AstroParticle Physics Amsterdam, University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands*
[2]*Astronomical Institute 'Anton Pannekoek', University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands*
[3]*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands*

Accepted 2011 December 1. Received 2011 November 30; in original form 2011 October 10

ABSTRACT

We use the 7-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*7) data to place constraints on oscillations supplementing an almost scale-invariant primordial power spectrum. Such oscillations are predicted by a variety of models, some of which amount to assuming that there is some non-trivial choice of the vacuum state at the onset of inflation. In this paper, we will explore data-driven constraints on two distinct models of initial state modifications. In both models, the frequency, phase and amplitude are degrees of freedom of the theory for which the theoretical bounds are rather weak: both the amplitude and frequency have allowed values ranging over several orders of magnitude. This requires many computationally expensive evaluations of the model cosmic microwave background (CMB) spectra and their goodness of fit, even in a Markov chain Monte Carlo (MCMC), normally the most efficient fitting method for such a problem. To search more efficiently, we first run a densely-spaced grid, with only three varying parameters: the frequency, the amplitude and the baryon density. We obtain the optimal frequency and run an MCMC at the best-fitting frequency, randomly varying all other relevant parameters. To reduce the computational time of each power spectrum computation, we adjust both comoving momentum integration and spline interpolation (in l) as a function of frequency and amplitude of the primordial power spectrum. Applying this to the *WMAP*7 data allows us to improve existing constraints on the presence of oscillations. We confirm earlier findings that certain frequencies can improve the fitting over a model without oscillations. For those frequencies we compute the posterior probability, allowing us to put some constraints on the primordial parameter space of both models.

Key words: methods: data analysis – cosmic background radiation – inflation – cosmology: theory – early Universe.

1 INTRODUCTION

The observed statistical distribution of temperature fluctuations in the cosmic microwave background (CMB) is believed to be largely determined by the physics in the very early Universe. These CMB fluctuations were sourced by quantum fluctuations during an epoch of accelerated expansion early on in the history of the Universe, known as inflation. They then induce curvature perturbations in the geometry of space–time, which are preserved after horizon crossing during inflation. Once they re-enter the horizon at some later time, they couple to radiation and matter, becoming responsible for the observed statistical distribution of the CMB temperature fluctuations and the large-scale structure (LSS) of matter in the Universe, respectively. Within the six-parameter Λ cold dark matter (ΛCDM), these initial conditions are described by only two parameters: the amplitude of the primordial power spectrum of scalar perturbations, and the tilt n_s describing the scale dependence of the power spectrum to first order. By investigating the CMB power spectrum, one is therefore able to probe high energy physics in the early Universe. Given the large number of models describing the physics of the early Universe (see e.g. Chen 2010b for a recent overview), these two parameters (within the six-parameter ΛCDM) are not sufficient to be able to discriminate between various proposed models. Additional degrees of freedom, derived from the statistical analysis of the late-time distribution of temperature (CMB) or density (LSS) fluctuations, could potentially be used to break the degeneracy between various models. Possible extensions to the six-parameter model include tensor degrees of freedom (gravitational waves), higher order corrections to the scalar *and* tensor power spectra, and deviations from Gaussianity measured through high-order correlation functions. The ultimate goal, of course, is to formulate a theoretical prediction of what these degrees of freedom should be, and constrain these from the observational data, just as we constrain the two parameters describing the tilt and the amplitude of the primordial

⋆E-mail: meerburg@princeton.edu (PDM); r.a.m.j.wijers@uva.nl (RAMJW); j.p.vanderschaar@uva.nl (JPvdS)

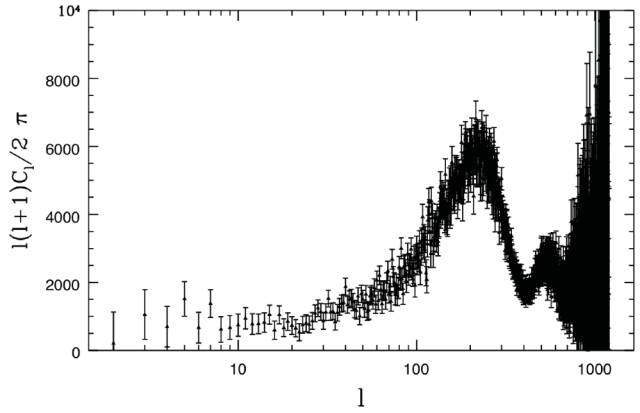

Figure 1. *WMAP*7 data with added errors from measurement and cosmic variance. The error bars are derived from the diagonal terms in the Fisher matrix. The multipole moments are slightly coupled (inducing off-diagonal elements), so a correct treatment of errors requires use of the entire Fisher matrix, which is done when calling the likelihood *WMAP* code.

power spectrum of scalar perturbations. The additional constrained parameters test the uniqueness of a proposed model and thus contribute to the understanding of the physics governing the early, inflationary Universe.

In this paper, we will consider modifications to the primordial power spectrum. In particular, we will search for evidence of oscillations in the almost scale-invariant spectrum. The motivation to search for these modifications is provided by a rapidly increasing number of theoretical models expecting such features in the primordial correlation statistics: e.g. in the power spectrum (Martin & Brandenberger 2001; Easther et al. 2001, 2003; Danielsson 2002; Greene et al. 2004; Schalm, Shiu & van der Schaar 2004, 2005; Easther, Kinney & Peiris 2005b; Cai, Hu & Zhang 2010; Chen 2010a; Flauger et al. 2010; Jackson & Schalm 2010, 2011; Achucarro et al. 2011; Chen 2011) and bispectrum (Chen et al. 2007; Chen, Easther & Lim 2008; Meerburg, van der Schaar & Corasaniti 2009; Chen 2010a; Meerburg, van der Schaar & Jackson 2010; Flauger & Pajer 2011; Meerburg & van der Schaar 2011). In this paper we will only consider oscillations in the power spectrum. Fergusson et al. (2010) have attempted to constrain features in the bispectrum and Meerburg (2010) has proposed a method to effectively search for features in the CMB bispectrum; recently LSS data have been proposed to search for these types of bispectra (Cyr-Racine & Schmidt 2011). The frequency, amplitude and possibly the phase and first-order scale dependence of these features are determined by the detailed physics of inflation. Detecting or constraining these parameters would help us determine the precise physics of inflation.

This is not the first time features in the power spectrum will be explored. Notably, Easther, Kinney & Peiris (2005a), Easther et al. (2005b), Martin & Ringeval (2004, 2005), Covi et al. (2006), Hamann et al. (2007), Pahud, Kamionkowski & Liddle (2009) and Flauger et al. (2010) have all investigated possible oscillations in the primordial power spectrum. Although there has been no clear detection, the data certainly allow small oscillations, as can be seen from Fig. 1. Previous analyses were done on 1-, 3- and 5-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data (as well as Sloan Digital Sky Survey data). In this paper, we will aim at extending and improving the analysis using the latest *WMAP* release, *WMAP*7 (Komatsu et al. 2011). One of the key hurdles in searching for oscillations in the data is the frequency of the oscillations. The frequency of the oscillations in the primordial power spectrum is a free parameter which spans several orders of magnitude for many of the proposed models. When probing the joint likelihood of our cosmological model, the large range in frequencies results in a number of issues (Groeneboom & Elgaroy 2008).

Foremost, it requires a high resolution in 'sample space', i.e. looking for the best-fitting parameter to the data requires a large number of computations of the CMB power spectrum from the primordial power spectrum (using e.g. CAMB, Lewis, Challinor & Lasenby 2000 or CMBFAST, Seljak & Zaldarriaga 1996). This necessitates an efficient computing scheme. Unfortunately, the computation of the late-time power spectrum from the primordial one involves a convolution of the transfer function $\Delta_l(k)$ with the primordial power spectrum $P(k)$. The appearance of rapid oscillations in both the transfer functions and the primordial spectrum requires smaller and more frequent steps in the integration in comoving momentum space k, increasing the required computational time for each run significantly.

Secondly, the late-time spectrum must be computed for each angular scale l. Usually, given the smoothness of the primordial spectrum, it suffices to do a spline interpolation on a number of C_l. This reduces the computational time, since the computation of the transfer functions[1] is the most time-consuming part. With the addition of oscillations on top of the smooth primordial spectrum, the number of C_l necessary to obtain an accurate fit of the interpolated C_l will depend on the frequency of the primordial oscillations. As this frequency increases, at some point *all* l will need to be considered in order to resolve the superimposed oscillations. Computing all l requires us to compute all transfer functions, which again increases the computational time. Lastly, there will often be a number of frequencies that tend to improve the fit within the large range we explore, rather than just a single one. For example, if a frequency ω is a good fit to the data, there is a fair chance that 2ω will be a good fit as well. This is a major issue, as a multidimensional parameter space is most effectively searched through a Markov chain Monte Carlo (MCMC), which is a random process. Therefore, when the frequency is not fixed, an MCMC approach to the best fit is not efficient as the likelihood is spiked and the random nature of the MCMC chain will lead to frequent tunnelling of one local maximum to another within the multidimensional joint likelihood. Flauger et al. (2010) showed a way to circumvent this issue by first taking a large number of samples fixed on a grid. Obviously, a grid does not allow us to vary all parameters within our cosmological model, as the number of samples grows quickly with the dimensionality of parameter space. Instead, the priority lies in varying the initial conditions, i.e. the oscillatory component of the primordial power spectrum. Once the best fit has been determined (within the prior frequency range, and the grid resolution), one can run MCMCs with a fixed best-fitting frequency determined through the grid. The joint likelihood is expected to no longer contain local maxima and an MCMC should converge quickly. This eventually allows us to put constraints on the amplitude and perhaps the first-order scale dependence of the amplitude of the oscillatory feature.

Although there are a large number of different features and oscillations predicted by a variety of models, in this paper we will focus on only two, distinct, theoretically motivated modified primordial power spectra. We will introduce these modified power spectra in Section 2. In Section 3 we explain how to optimize the search by making the numerical computation of the power

[1] These can generally not be precomputed, as they depend on all relevant parameters of the theory, which are varied when searching for the best fit.

spectrum frequency-dependent. To find a best-fitting frequency before we apply the MCMC, we use grid sampling. We report our findings in Section 4. Once we have established a best-fitting value of the frequency, we run an MCMC with that best-fitting value for two models. The results are discussed in Section 5 and we compare these to theoretically derived constraints on primordial parameter space. We draw conclusions in Section 7.

2 TWO MODELS OF INITIAL-STATE MODIFICATIONS

Although the standard Bunch Davies (BD) vacuum state during inflation is an excellent fit to the currently available CMB data, theoretical considerations have questioned its validity and uniqueness over the last decade (see Easther et al. 2001, 2003, 2005b; Martin & Brandenberger 2001; Danielsson 2002; Greene et al. 2004; Schalm et al. 2004, 2005, and references therein). The main reason for casting doubt on the BD assumption is the fact that the temperature fluctuations in the CMB ultimately find their origin in quantum fluctuations in a vacuum state just at the onset of inflation. Predicting the spectrum of CMB temperature fluctuations requires determining this initial vacuum state for quantum modes at extraordinarily high momentum scales, far beyond any high-energy cut-off scale, where a perturbative quantum field theory description is no longer expected to be accurate. From a theoretical point of view, it is not understood why the time-dependent inflationary background would transfer this unknown high-energy physics to a regular quantum field theory description in the BD state at later times. In wait for a more complete understanding, this vacuum state ambiguity has motivated phenomenological approaches, usually relying on low-energy effective field theory expectations, that have suggested the presence of high-energy corrections to the BD state. A typical prediction of these models is the appearance of (small) oscillations on top of the standard primordial power spectrum of inflationary perturbations.

Besides constraining these characteristic oscillations in general, an additional goal of this work is to study to what extent the currently available CMB power spectrum data can distinguish between two rather general classes of models that have been proposed to describe initial-state modifications. The first class of models is known as the boundary effective field theory (BEFT) approach to determine initial-state modifications (Greene et al. 2004; Schalm et al. 2004). In this proposal, one fixes an initial *time* where one calculates corrections to the usual BD initial condition using a low-energy effective boundary Lagrangian. The result is an explicitly scale-dependent Bogolyubov parameter β_k describing the modification with respect to the usual BD vacuum state, giving rise to strongly scale-dependent oscillatory corrections to the primordial power spectrum.

The second class of models is known as the new physics hypersurface (NPH) approach to initial-state modifications (Danielsson 2002; Easther et al. 2002). In the NPH scenario, one traces every momentum mode back to some large *physical scale* of new physics M and imposes, on a rather ad hoc basis, the standard flat space vacuum state (corresponding to positive frequency modes only), mode by mode, resulting in a k-independent Bogolyubov parameter β_k. A small departure from scale invariance only arises after taking into account the slow-roll evolution of the Hubble parameter, which affects the amplitude of the oscillatory corrections described by $\beta_k \propto H/M$. A recent effort by Jackson & Schalm (2011) to compute the low-energy vacuum state effects of a massive scalar field in an inflationary background carries the important prospect to ground the NPH proposal in a more solid effective field theory description. The authors showed that integrating out an arbitrary massive scalar field can indeed affect the vacuum, resulting in rather similar corrections to the inflationary power spectrum as in the original NPH scenario. For our purposes, the power spectrum predicted by Jackson & Schalm (2011) will be used to set up a general phenomenological parametrization of the almost scale-invariant NPH class of models.

These two classes of models describing initial-state modifications, NPH and BEFT serve as a nice benchmark to study how constraining and distinguishing the most recent CMB power spectrum data is when considering superimposed oscillations. The data analysis is in fact best performed using a primordial power spectrum parametrization of the oscillating corrections involving a set of independent parameters. The general parametrization of the primordial power spectrum for the two classes of models that we will make use of is (Greene et al. 2004)

$$P(k) = P_0(k)\,[1 + \beta k A_0 \sin(2A_1 k + \phi)] \tag{1}$$

for the BEFT scenario, while for the new Jackson–Schalm scenario (Jackson & Schalm 2011) (from here on NPH)

$$P(k) = P_0^*\,[a_0 + a_1 \ln k/k_* + (a_2 + a_3 \ln k/k_*) \sin(a_4 \ln k/k_* + \zeta)]\,. \tag{2}$$

In equation (1), β is a parameter determined by the details of the physics in the ultraviolet, and is naturally expected to be $\mathcal{O}(1)$, while A_0 and A_1 are (related) parameters predicted by the BEFT method. $P_0(k)$ is the power spectrum from canonical slow-roll inflation in a BD state. This contains a possible tilt n_s. In equation (2), k_* is a pivot scale for normalization, which is set to $k_* = 0.002\,\mathrm{Mpc}^{-1}$, P_0^* is the scale-independent power spectrum, and a possible tilt is not included. The reason for this difference is that the oscillatory correction in the NPH scenario can have a tilt a_3 that differs from the tilt $a_1 = n_s - 1$. For the BEFT, we will be interested in possibly constraining A_0, A_1 and ϕ, while for the NPH model we will consider a_2, a_3, a_4 and ζ. The parameters a_0 and a_1 will be considered zero-order contributions to the power spectrum and are constrained through n_s and P_*. As already briefly alluded to, we will constrain each parameter independently and apply theoretical dependencies between those parameters only afterwards. For example, A_1 in the BEFT model is related to the amplitude A_0. The reason not to implement this directly is that separating these parameters results in a much 'cleaner' sampling, because the frequency causes the likelihood to be highly irregular. We will apply the theoretically predicted relations between the amplitude and the frequency once the posterior distribution has been determined.

For more specific details and theoretical background on the NPH and BEFT models, we refer to Greene et al. (2004) and Jackson & Schalm (2011), and Danielsson (2002), respectively.

3 CODE OPTIMIZATION

We have modified the CAMB code (Lewis et al. 2000) and the publicly available COSMOMC package (Lewis & Bridle 2002) in order to efficiently search for oscillations. The modification is built upon work done by Martin & Ringeval (2004, 2005). The late-time power spectrum is basically a convolution of the transfer function $\Delta_l(k)$ and the primordial power spectrum $P(k)$,

$$C_l \propto \int_0^\infty k^2 \mathrm{d}k\, P(k) \Delta_l^2(k). \tag{3}$$

The transfer function contains all the physics that governs the evolution of the Universe and describes an interplay between radiation

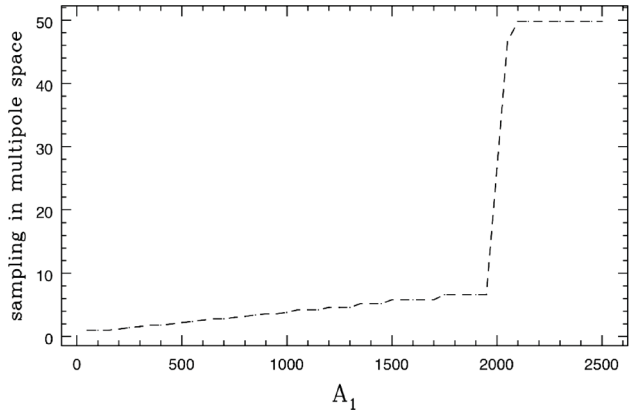

Figure 2. Sampling increase for splining in multipole space as a function of A_1 for the BEFT model. We fixed $A_0 = 10$, its largest possible value. As the frequency A_1 increases, it requires an increasing number of neighbouring l in the interpolation to maintain the same accuracy (which was set by *WMAP* best-fitting standards). Around $A_1 \sim 2000$, the sampling factor increases beyond 50, which implies that all l values are required (i.e. splining is not sufficient and one needs to compute the full transfer matrix).

pressure and gravitational collapse. This causes the transfer function to be highly oscillatory. Consequently, to gain sufficient accuracy, one has to numerically integrate over k-space with an adequate number of steps. The transfer functions are not only a function of comoving momentum but also depend on the angular scale l. To increase the speed of the code, one usually does not compute the complete transfer matrix, but rather limited number of l, and spline interpolates between them.

If the primordial power spectrum is smooth, both the integral over comoving momentum and the spline interpolation in l can be done with limited samples. Once we allow for (rapid) oscillations in the primordial power spectrum, the number of samples needs to be increased.

Numerically, the most time-consuming application is the number of transfer functions we will have to compute. Therefore, the fewer transfer functions we compute, the faster we can determine the angular power spectrum C_l. Obviously, if the number of oscillations in our primordial spectrum is large, we will need many transfer functions to resolve those using a spline interpolation. If the wavelength of the primordial signal is $\delta l \sim 1$, no spline interpolation will resolve the primordial signal.

3.1 Frequency-dependent power spectrum computation

We have modified CAMB and CosmoMC in order to optimally compute the power spectrum given a primordial frequency. Fixing all parameters to the best-fitting *WMAP*7 values (Komatsu et al. 2011), except for the frequency in the primordial power spectrum, we investigated the convergence in the late-time power spectrum when we increased the sampling in k- and l-spaces. As a null hypothesis, with 'optimal' accuracy, we took the primordial spectrum without oscillations. By plotting the frequency against the sample increase required to retain the same accuracy, we obtained an estimate of the optimal sampling for every frequency. As an example, in Fig. 2 we show the increase in l-sampling required to obtain an accurate computation of the power spectrum as a function of the BEFT frequency A_1. Only for relatively low frequencies $A_1 < 2000$ and $a_4 < 60$ can we use splines in l-space to determine the CMB power spectrum. At higher frequencies, we need to compute the full transfer matrix in order to resolve oscillations.[2] For the sampling in k-space, the optimal sampling was derived when computing all transfer functions. We find that over the range of frequencies we will probe in this paper, we need to approximately double the resolution in k-space when doing the integration.

4 GRID SAMPLING

Given the number of free parameters describing the initial conditions (at least A_s, n_s and the frequency and amplitude of the oscillating correction), many computations of the power spectrum are required to obtain the best-fitting (distribution). A commonly used approach is to apply an MCMC, randomly pick values of a set of parameters, and based on the likelihood of the computed run, reject or accept this point to be part of our parameter probability estimate. This approach is highly efficient, as we only compute a limited number of points in the multidimensional likelihood to determine the posterior distribution. Once we add oscillations to the primordial power spectrum, an MCMC method becomes less efficient, as the likelihood is expected to become irregular in the coordinate of the oscillation (the frequency) and there are many local maxima in the likelihood function. At some point, we will leave this local maximum and end up in another. Therefore, constraints on the initial conditions are hard to recover since the MCMC will constantly move to different local maxima in likelihood space.[3] Flauger et al. (2010) proposed a different approach. Instead of an MCMC over all parameter space, they first considered a grid over a limited number of parameters. The parameters to vary in the grid should be those that determine the primordial conditions (e.g. the frequency). They also identified the baryon density Ω_b to be degenerate with the amplitude and frequency of the oscillations, as it influences the height of the first peak. Finally, they showed that oscillations in axion monodromy models of inflation are a better fit primarily to the first peak in the angular power spectrum, and hence could mimic some of the effects produced by the baryon energy density Ω_b. The grid therefore samples at least three parameters (amplitude, frequency and baryon density). The other parameters are fixed to the *WMAP*7 (Komatsu et al. 2011) best-fitting values. As such, the null hypothesis (no oscillations) is the fit to beat. The advantage of the grid is that one probes the likelihood completely, although under fixed conditions for most of the ΛCDM parameters. Under the assumption that these other parameters are not (strongly) correlated with the varying parameters, we should be able to determine the absolute maximum in the likelihood rather than a local one. Once this local maximum has been determined, we can perform an MCMC with the frequency set to the best-fitting value, and allow all other parameters to vary. This should probe the likelihood of e.g. the amplitude more efficiently.

We ran three grids, one for the BEFT model and two for the NPH model. While in this paper we have opted to avoid details about the exact theoretical prediction of the parameters in each model, we decided to consider two grids for the NPH model focusing on two frequency regimes a_4. This is motivated by the fact that the NPH

[2] The sampling is from a set of predetermined l (e.g. $l = \{1, 2, 3, \ldots, 100, 150, \ldots, 1000, 1100, \ldots\}$). The number of points used in the interpolation is increased within this set of predetermined l until this set is unable to achieve the same accuracy as one would obtain by computing all l. This set is based on efficient computation of a smooth primordial spectrum.

[3] Constraints on other parameters unrelated to the initial conditions can be recovered with sufficiently large samples.

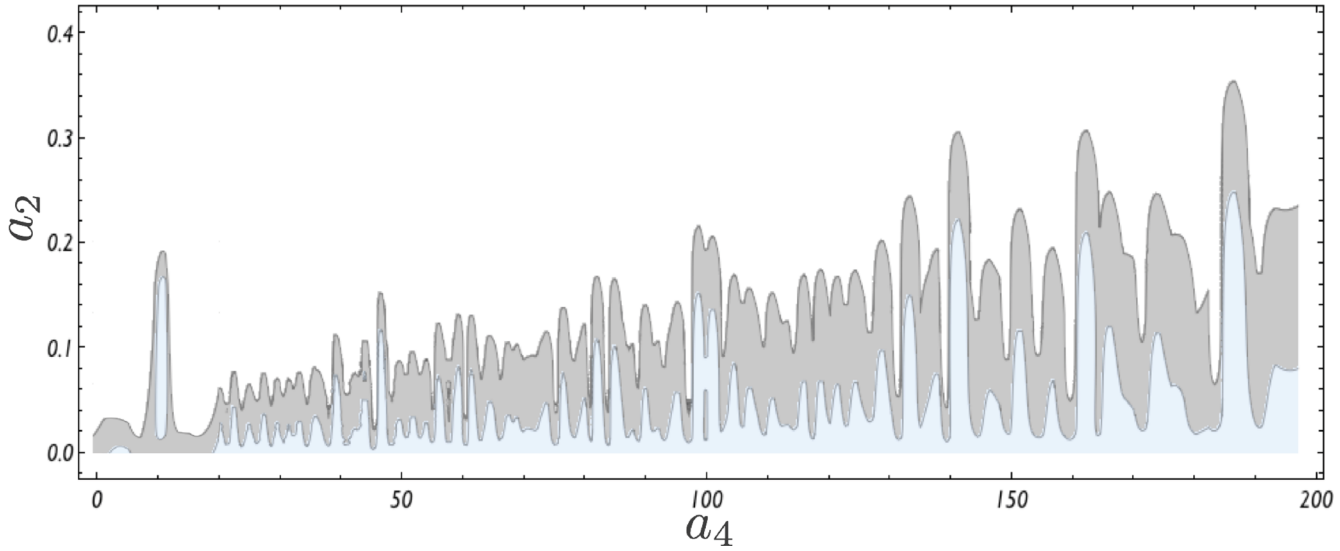

Figure 3. NPH. The 68 and 95 per cent confidence levels for amplitude a_2 versus the frequency a_4 for the low-frequency grid $1 \leq a_4 \leq 200$ marginalized over the baryon density. There are many local peaks in the marginalized likelihood. The two most likely grid points are ($a_4 = 46.5$, $a_2 = 0.056$) and ($a_4 = 98.41$, $a_2 = 0.147$). Towards higher frequencies, larger amplitudes are allowed by the data but do not necessarily represent significant improvements of the overall fit. Around $a_4 = 10$, there exist a range of frequencies that map some of the features in the slope of the first peak.

model is not expected to have a significant number of oscillations, as the theoretically predicted frequency is slow-roll suppressed, i.e. $a_4 \propto \epsilon$. We have therefore investigated low frequencies ($1 \leq a_4 \leq 200$) with 200 log-spaced samples, and a high-frequency regime ($100 \leq a_4 \leq 1000$) separated into 500 logarithmically spaced samples. The amplitude a_2 runs from 0 to 0.4 in 120 and 200 equidistant steps, respectively. For the NPH grid $a_3 = \zeta = 0$. For the low-frequency grid, we set $0.021 \leq \Omega_b h^2 \leq 0.026$ in 10 equidistant steps, resulting in a total of 240 000 grid points, while for the high-frequency regime we considered $0.02 \leq \Omega_b h^2 \leq 0.027$ in 16 equidistant steps, with a total of 1.6 million grid points.

In Figs 3 and 4 we show the confidence contours for grids obtained for the NPH model of the frequency versus the amplitude in the low- and high-frequency regimes, marginalized over the baryon energy density Ω_b. Peaks are areas in which the fit is best for non-zero values of the amplitude of the modification, while valleys represent frequencies which are not a good fit to the data and the best fit is no modification. For example, one can consider a likely frequency (peak) and plot the probability of the amplitude for that frequency to find that the most likely value for the amplitude is non-zero. In Fig. 5 we show the joint likelihood contour plot for the amplitude a_2 and the baryon density $\Omega_b h^2$ marginalized over the frequency a_4.

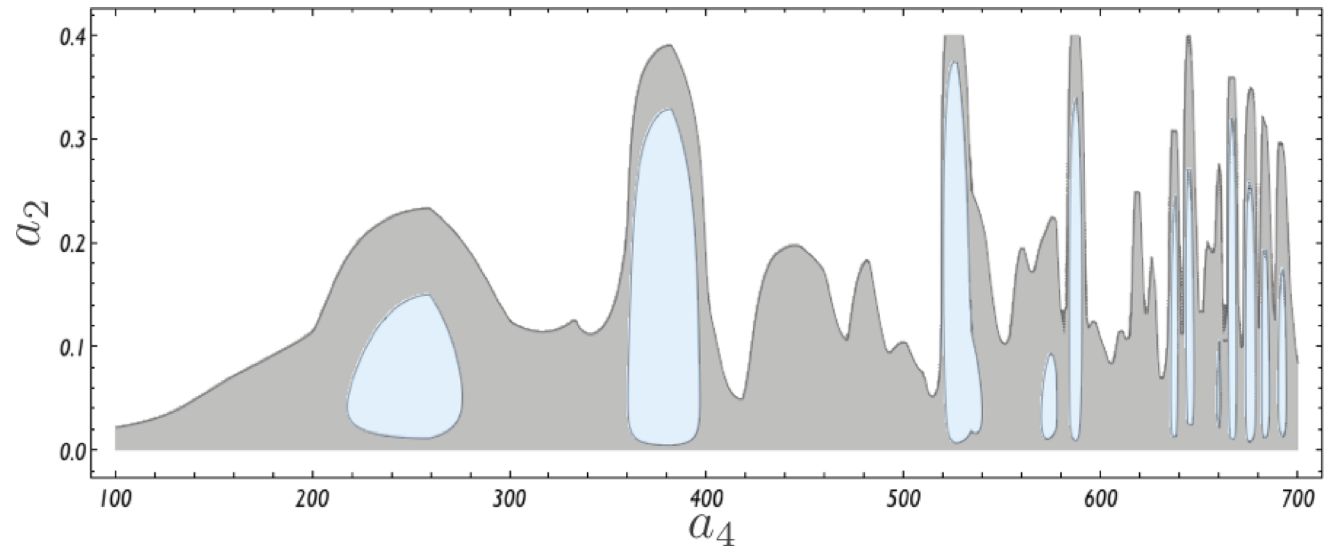

Figure 4. NPH. The 68 and 95 per cent confidence levels for a part of the high-frequency grid. Note that the absence of a match between the low- and high-frequency regimes is artificial. In the high-frequency grid, there are only six samples for $100 \leq a_4 \leq 400$. Additionally, a 100 per cent likelihood was assumed within one grid. For very high frequencies, the wavelength of the oscillations is too small to be resolved at large angular scale ($l < 200$). Primordial oscillations are resolved only in the second peak and beyond. Very high frequencies can produce glitches in the large-scale wing of the first peak in the angular power spectrum. For some specific frequencies and amplitudes, these glitches can match outliers in the wing of the first peak of the observed angular power spectrum resulting in $\Delta\chi^2 \simeq 12$ for $a_4 \sim 998$.

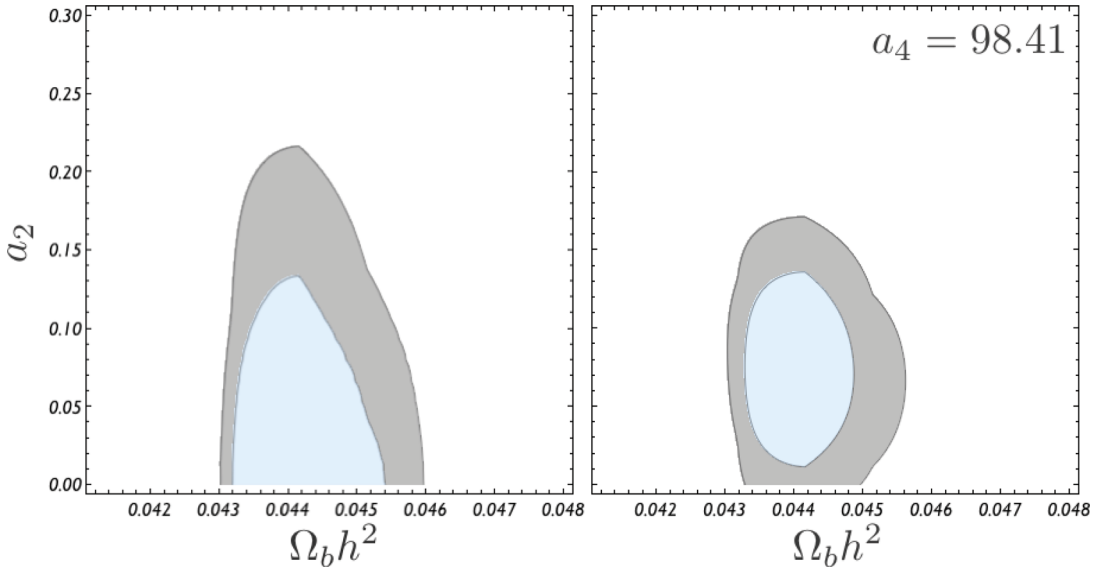

Figure 5. NPH. The 68 and 95 per cent confidence levels of $\Omega_b h^2$ versus a_2. The left-hand panel shows the joint likelihood after marginalizing over the frequency a_4 (low-frequency grid), while the right-hand panel shows the joint likelihood for the best-fitting frequency $a_4 = 98.41$.

We find that $a_2 = 0$ is the most likely value in the low-frequency grid. We also derived the joint likelihood for one of the best-fitting frequencies ($a_4 = 98.74$) to show that the best-fitting point has a non-zero amplitude a_2 with almost 95 per cent confidence level (CL).

For the high-frequency regime, we show the effect of the highest frequencies on the marginalized amplitude a_2 in Fig. 6. Note that the confidence levels are determined assuming that the grid contains all possible values the frequency could have. In other words, one would hope that for either extremely large or small frequencies the likelihood of the fit would go to zero. The problem is that it does not, and therefore we can only determine the confidence levels within a prior determined parameter domain. We have partly motivated this domain on theoretical arguments. Observationally, data are the limiting factor.

It should be obvious that many different frequencies represent good fits. As we had foreseen, this complicates running a large MCMC for *all* relevant cosmological parameters. The best-fitting point we find for the NPH model ($\Delta\chi^2 \sim 12$) centres around very high frequencies, with $a_4 = 980$ and $a_2 = 0.39$. In fact for this frequency, the best-fitting amplitude probably lies outside the domain of $0 \leq a_2 \leq 0.4$. This implies a relatively large number of primordial oscillations. The angular power spectrum for such high frequencies has most of its oscillations damped, since this frequency is only resolved at scales beyond the first peak. Martin & Ringeval (2004) showed that for these scales the amplitude will be suppressed, similar to how the overall power is damped. Given the large measurement error at these scales, we would not expect the fit to improve that much. It turns out that at large angular scales, the barely resolved high-frequency causes glitches in the large-scale wing of the

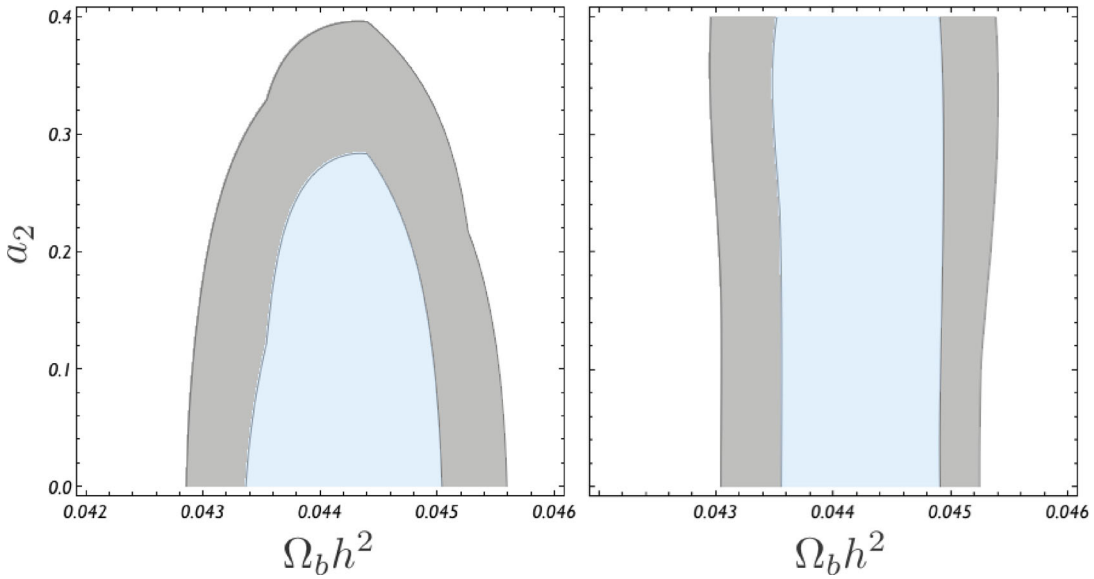

Figure 6. NPH. The 68 and 95 per cent confidence levels of $\Omega_b h^2$ versus a_2. The left-hand panel shows the joint likelihood after marginalizing over the frequency a_4 for the high-frequency grid up to $a_4 \leq 700$ (63 frequencies). If one includes all frequencies in the marginalization, the constraint on a_2 disappears (right-hand panel).

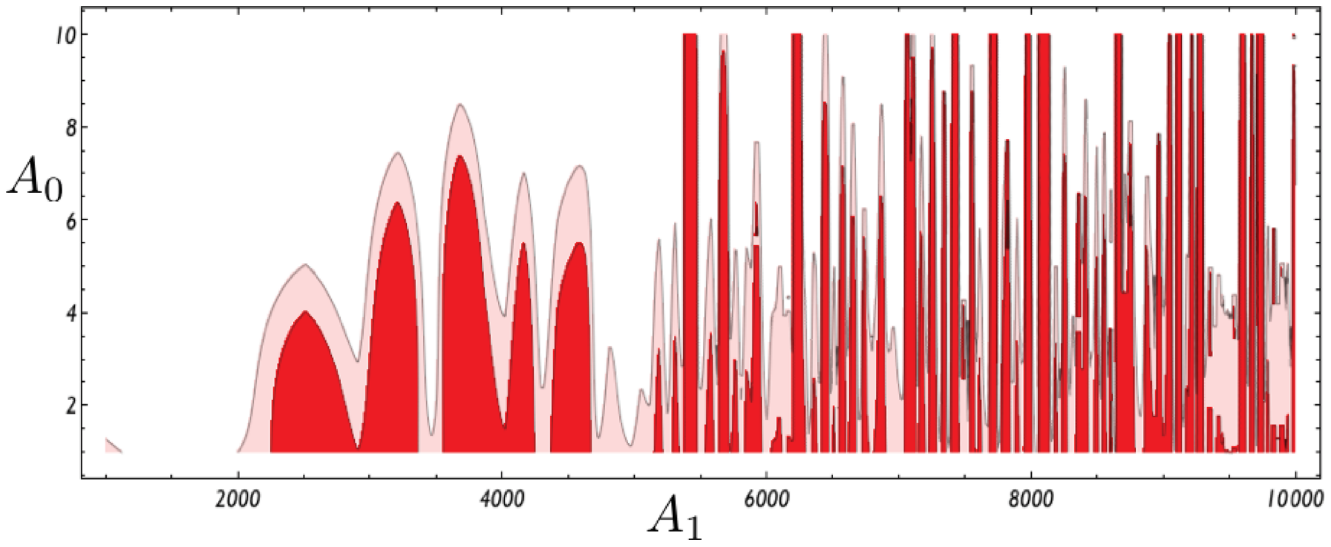

Figure 7. BEFT. The 68 and 95 per cent confidence levels for amplitude A_0 versus the frequency A_1 marginalized over the baryon density. Again, there are many local maxima in the likelihood. The cut-off in amplitude A_0 is set by theoretical constraints. If we would allow larger amplitudes ($A_0 > 10$), the likelihood contours would likely move towards higher A. The most likely grid point is ($A_1 = 7708$, $A_1 = 10$).

first peak. The *WMAP* data contain outliers in the large-scale (small l) wing of the first peak (attempts have been made to understand this in terms of a feature in the primordial spectrum; see e.g. work by Covi et al. 2006 and recently Dvorkin & Hu 2011 performed a principal component analysis to search for such a feature); these are fitted for very specific unresolved frequencies of the primordial power spectrum. These glitches are expected as the angular power spectrum will no longer be able to sample all oscillations appearing in the primordial power spectrum. We have not been able to get rid of these glitches by increasing the overall accuracy of the numerics of the code.

In the low-frequency regime, we obtain two frequencies that give an improvement of $\Delta\chi^2 \sim 6$ compared to no oscillations. In the next section, we will discuss the results from an MCMC run for one of these frequencies, with a varying phase and amplitude. For the NPH model, the results can be summarized as follows: for frequencies $1 < a_4 < 200$, $a_2 < 0.13$ at 68 per cent and $a_2 < 0.21$ at 95 per cent confidence. For frequencies up to $a_4 = 700$, $a_2 < 0.29$ at 68 per cent and $a_2 < 0.39$ at 95 per cent, while for higher frequencies the amplitude is not constrained within the bounds of the grid. The most likely value within the low-frequency grid is $a_4 = 46.5$ with an amplitude of $a_2 \sim 0.056$.

For the BEFT model, there is a constraint set by BEFT validity on the value of A_0. A BEFT approach to initial-state modifications is only valid if the physical momentum is smaller than the cut-off scale at the boundary, i.e. $k/a_0 < M$. The maximum value of k is set by the smallest observable scale in the CMB, $k_{\rm max} = \mathcal{O}(0.1)$ Mpc^{-1} and therefore we deduce $A_0 \leq 10$ from equation (1). The number of samples in A_0 is set to 120 equidistant values between 1 and 10. For BEFT, the frequency is not constrained by slow-roll and is proportional to M/H; the scale of new physics is divided by the Hubble scale. Consequently, the effective frequency can be quite high. We sample 700 logarithmically spaced steps between 10^3 and 10^4, making up a total of 1 344 000 grid points.

The likelihood confidence contours for the BEFT model are different from those of the NPH model (Fig. 7). Most importantly, the resulting contour does not have a vanishing amplitude, which was assumed as a prior based on the theoretical form of A_0. The best-fitting point in the grid is given by $A_1 = 7708$, with $A_0 = 10$ corresponding to an improvement of $\Delta\chi^2 \simeq 13$. The corresponding late time power spectrum is shown in Fig. 8.

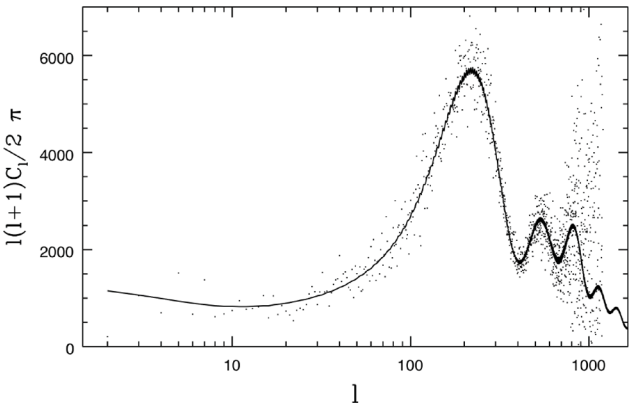

Figure 8. The best-fitting angular power spectrum derived from the grid for the BEFT model with $A_0 = 10$ and $A_1 = 7708$, resulting in $\Delta\chi^2 \simeq 13$ relative to no oscillations.

5 MCMC AND MODEL CONSTRAINTS

We analysed the *WMAP* data using Monte Carlo Markov sampling with a fixed frequency derived from the grids in the previous section for both BEFT and NPH models. We set a Gelman–Rubin (GR) criterion of $R - 1 < 0.01$ (Gelman & Rubin 1992). The GR diagnostic R relies on parallel chains to test whether they all converge to the same posterior distribution by considering the variance of the parameters in each chain compared to the variance of the same parameters over all parallel chains. Convergence is diagnosed once the chains have 'forgotten' their initial values, and the output from all chains has become indistinguishable ($R - 1 = 0$). For the NPH model, we ran eight parallel MCMC with $a_4 \sim 45.9$, close to the best-fitting point in the low-frequency grid. In addition to a_2, we allowed both a_3 and ζ to vary. We assumed flat priors with $-0.6 \leq a_2 \leq 0.6$, $-0.5 \leq a_3 \leq 0.5$ and $0 \leq \zeta \leq \pi$. We did not reach our desired GR criterion with the first chains, which is mostly entirely due to the slow convergence of the amplitude P_*. After we ran eight

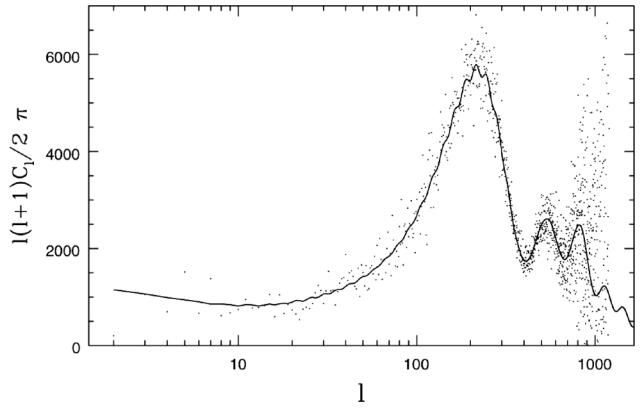

Figure 9. Best fit for the NPH model from the MCMC chain, with a fixed frequency $a_4 \sim 46$. The improvement of the fit compared to no oscillations is $\Delta\chi^2 \sim 12$. From the grid we found that an equally good fit is for a frequency of $a_4 \sim 98$, about double this frequency. We find that the best-fitting amplitude is rather large, $a_2 \sim 0.14$, with the 68 per cent level still allowing zero amplitude (Fig. 11). The amplitude is similar to the best-fitting amplitude of the axion monodromy model derived by Flauger et al. (2010). However, this best fit was at a translated frequency of about 150.

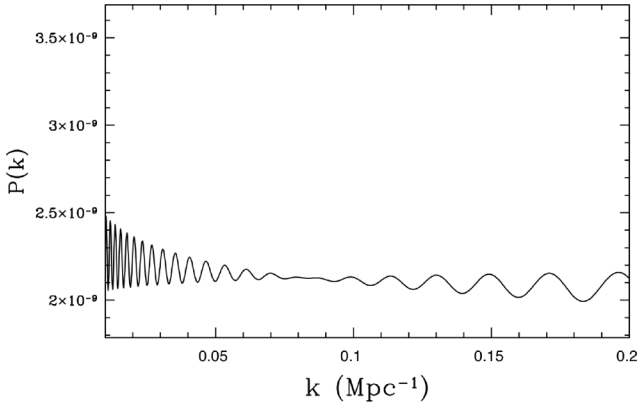

Figure 10. The primordial spectrum from the NPH model with $a_4 \sim 46$, $a_2 = 0.14$, $a_3 = -0.04$ and $\phi = 1.5$.

chains, we derived a proposal covariance matrix. We used this covariance matrix to speed up the convergence. With the addition of a covariance proposal, we easily obtained $R - 1 < 0.003$ over four chains. The result for some of the marginalized and joint likelihoods is shown in Fig. 11 for the NPH model. Note the high correlation between a_2, the amplitude of the oscillation, and a_3, the scale dependence of the oscillatory correction. We found no evidence for a strong correlation between a_2 and a_3 and the other parameters, which proves that a grid sampling is a very good first estimate of the best-fitting values of both parameters. We found that ζ is weakly correlated with both the amplitude P_* and the dark matter energy content of the Universe, explaining the relatively slow convergence of these distributions.

From the MCMC for the NPH scenario, we deduce marginalized best-fitting values $a_2 = 0.15 \pm 0.17$, $a_4 = 0.04 \pm 0.04$ and $\zeta = 1.6 \pm 0.7$ with a fixed frequency of $a_4 = 45.9$ and a best-fitting point with $\Delta\chi^2 \simeq 12$. Therefore, we conclude that at the 1σ level, a_2 is consistent with zero. In Figs 9 and 10 we show the corresponding primordial and late time power spectra respectively.

To relate these to constraints on the NPH model, one must investigate the relation between a_2, a_3 and a_4 as derived by Jackson & Schalm (2011). The exact form of these parameters is presented in equation (31) of their paper.[4] We derive a relation between a_2, a_3 and a_4:

$$\frac{a_2}{a_3} \simeq -\frac{1}{a_4}\frac{M}{H}\left[5/2 + \ln M/H\right]^{-1}. \tag{4}$$

Here, we took the lower limit $\Lambda = M$, where we consider $M/H > 10^2$. The observational limits were derived for $a_4 \sim 46$, which puts a bound on $M/H > 10^3$ for $\epsilon < 0.01$ (slow-roll). This results in a theoretical constraint $a_2/a_3 < -2.3$. We can derive a similar constraint for the upper limit $\Lambda = (1/2)(H + M^2/H)$ or $\Lambda/H \sim 1/2(M/H)^2$. It turns out that a_4 in this limit can only have a negative sign. For a sine, this means the amplitude picks up a minus sign. We find $M/H > 10^3$ and

$$\frac{a_2}{a_3} \simeq \frac{1}{a_4}\frac{M}{H}(2 - \ln M/2H)\left[5/2 + \ln\frac{1}{2}(M/H)^2\right]^{-1}, \tag{5}$$

which together results in $a_2/a_3 > 5.86$.

In Fig. 12 we show the joint likelihood together with the theoretically allowed values of a_2 and a_3 for a given frequency a_4. If this frequency is a valid signal, a minimal improvement of the confidence contours could exclude this oscillation to be due to an NPH altogether.

For the BEFT model, the bound on the frequency is not that stringent and we simply chose the best-fitting frequency of $A_1 = 7708$ obtained from the grid. Note that this high frequency requires us to compute all l in order to resolve the primordial oscillation. This significantly increases computation time for each power spectrum. From the MCMC we would like to derive proper distributions, which could be an issue if we put priors on A_0 that exclude values that have non-zero probability. Therefore, we allow $0 \leq A_0 \leq 25$ and $-\pi \leq \phi \leq \pi$.

The eight parallel MCMC for the BEFT model resulted in an improvement of $\Delta\chi^2 \sim 16$, with all parameters satisfying the GR diagnostic, except A_0 and ϕ. We ran six additional parallel chains with an estimated covariance matrix from the first eight chains. The GR diagnostic was reduced for both A_0 and ϕ to $R - 1 < 0.01$.

Again, we find no indication that A_0 and ϕ are strongly correlated with any of the other parameters. In fact, A_0 almost seems independent of the other parameters. This is reflected in the fast convergence of the distribution of this parameter, which reached $R - 1 < 0.2$ after running the first eight chains (which is remarkable given the length of the chains). Again, the phase ϕ is weakly correlated with all the energy densities, leading to a relatively slow convergence of these distributions, even after estimating the covariance matrix. This correlation, however, is even less profound here than in the NPH model. These parameters reach $R - 1 < 0.02$ within the limited running time (six chains, each 120 h on eight core 2.2-GHz CPUs, resulting in chains of approximately 45 000 samples) of the chains.

For a comparison with theoretical bounds, we derive a relation between A_0 and A_1 as $A_1 \simeq A_0 \times (M/H)$ (see Greene et al. 2004, table 1). For a fixed frequency of $A_1 \sim 7708$ and assuming that the true distribution is a Gaussian, we derive $7.0 \leq A_0 \leq 15.2$ at 68 per cent confidence, which is almost 3σ away from zero. From this value, we derive that $500 \leq M/H \leq 10^3$ at 68 per cent confidence, which lies within theoretical predictions. Based on the best-fitting

[4] We derived these constraints in the assumption $\epsilon_1/\epsilon_2 \sim 1$ and $H_* \sim H$ due to a weak scale dependence (Jackson & Schalm 2011).

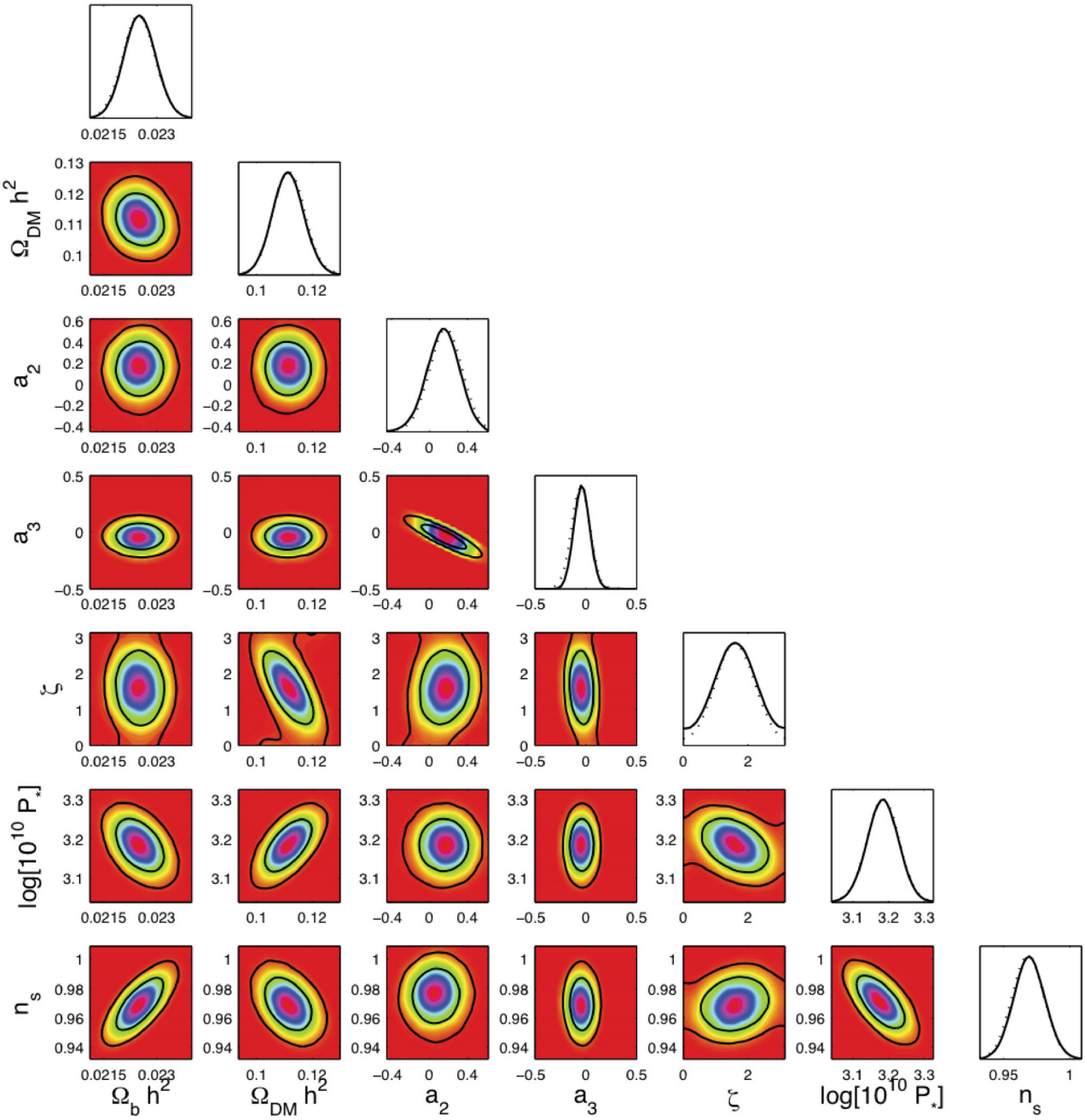

Figure 11. Marginalized and joint likelihoods for several parameters in the new NPH scenario. The parameters a_2 and a_3 show strong correlation.

amplitude $A_0 \sim 13$ and the average likelihood, we consequently obtain a slightly different constraint of $450 \leq M/H \leq 850$ at 68 per cent confidence. We argued that in principle A_0 is constrained to be smaller than 10 in the BEFT model. We set the coupling constant $\beta = 1$ but in fact β is $\mathcal{O}(1)$, which allows for small increase of the total amplitude, βA_0, possibly matching the best-fitting amplitude. Unlike the best-fitting oscillation from the NPH model, this correction is inconsistent with zero at the 3σ level.

6 EFFECTS ON OTHER PARAMETERS

Although there is little to no correlation between most of the parameters from the oscillating component and the other parameters within ΛCDM, the cross-correlation between ΛCDM parameters is affected differently for BEFT and NPH (the best-fitting values however are similar). For example, although the confidence levels of n_s are hardly affected by the presence of oscillations, in the NPH model, a weak correlation between the phase and the energy densities causes the probability distribution in Ω_{DM} to broaden, i.e. the correlation between the phase and the energy densities in the NPH scenario causes the uncertainty in those parameters to increase. This effect is shown in Fig. 13.

7 CONCLUSIONS

We used the latest *WMAP* data to constrain oscillations on top of an almost scale-invariant primordial power spectrum. We argued that the primary difficulty in constraining these models with the data is

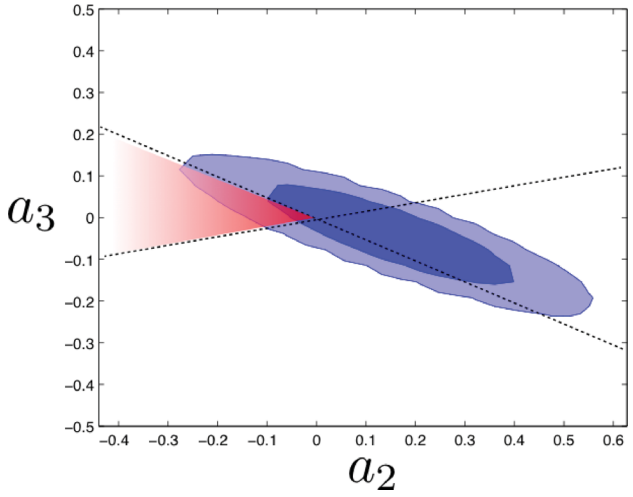

Figure 12. Joint likelihood for a_2 and a_3 together with the theoretical constraints relating a_2 to a_3 showing 68 and 95 per cent confidence levels. The region (shaded, red) allowed within the NPH model constitutes a small part of the observationally allowed region. It shows that a marginal improvement of the confidence contours could exclude NPH as the source of this oscillation. In addition, both parameters are perfectly consistent with zero.

related to the irregular likelihood function of the frequency of these oscillations. There are many equally good fits and these fits are at discrete frequencies. For all other parameters, one tries to constrain in a ΛCDM model of the Universe, such degeneracy usually does not exist and it is quite sufficient to scan parameter space using an MCMC. In order to avoid the random jumps from one maximum in the likelihood to another, it is preferable to keep the frequency fixed. However, the frequency is one of the key parameters being constrained, and it would make no sense to fix it prior to analysing the data. In order to estimate the frequency before the MCMC, we applied a grid search for the best fit. This grid quickly becomes incalculable for a large number of parameters (which is the primary reason to run an MCMC instead) and we only varied the parameters characterizing the modification of the power spectrum as well as the baryon density, which previously had been identified to be correlated with the parameters of the oscillatory correction. Once the best fit had been established, we performed an MCMC with the frequency fixed to its best-fitting value.

We found that the addition of oscillations on top of the smooth power spectrum can improve the overall fit. An improvement up to $\Delta\chi^2 \sim 16$ was found once we applied an MCMC with a predetermined best-fitting value of the frequency, which resulted in a best-fitting amplitude almost 3σ away from zero. We did not find significant correlation between the oscillatory parameters *besides* the frequency and the parameters in ΛCDM, confirming that grid sampling is a fairly robust first estimate of the best-fitting frequency. It must be noted, however, that we did not vary the frequency and the phase at the same time. The reason to leave out the phase in the grid was driven by computational costs and assessed through earlier findings where no strong correlations were reported. We do believe that there should be some correlation between the phase and the frequency and it would be interesting to vary both within a grid, possibly in favour of the baryon density. We will leave this for a future analysis.

In the NPH scenario, very high frequencies should lead to unobservable amplitudes. For completeness, we analysed such high frequencies. We found that for some specific frequencies we could obtain an improvement in the goodness of fit up to $\Delta\chi^2 \sim 12$ compared to no oscillations. Analysing the angular power spectrum with these extreme primordial frequencies revealed that these frequencies are in fact unresolved at large angular scales (small l), where the wavelength of the oscillating primordial power spectrum is smaller than the angular distance between subsequent l. This leads to small glitches in the slope of the first peak of the angular power spectrum, which fit some of the observed outliers in the slope of the first peak in the observed power spectrum. As such, high frequencies, although unresolved, could account for some of the large-scale effects we observe in the data, and improve the overall fit. We conclude, however, that these oscillations cannot be caused by NPH modifications and must be due to a different model, possibly an axion monodromy inflation-type model (Flauger et al. 2010).

Although in the grid for the low-frequency domain $1 \leq a_4 \leq 200$ of the NPH model we could only found one frequency with an improvement of $\Delta\chi^2 \simeq 6$, we ran an MCMC near this frequency to obtain an improvement of $\Delta\chi^2 \simeq 12$ corresponding to a

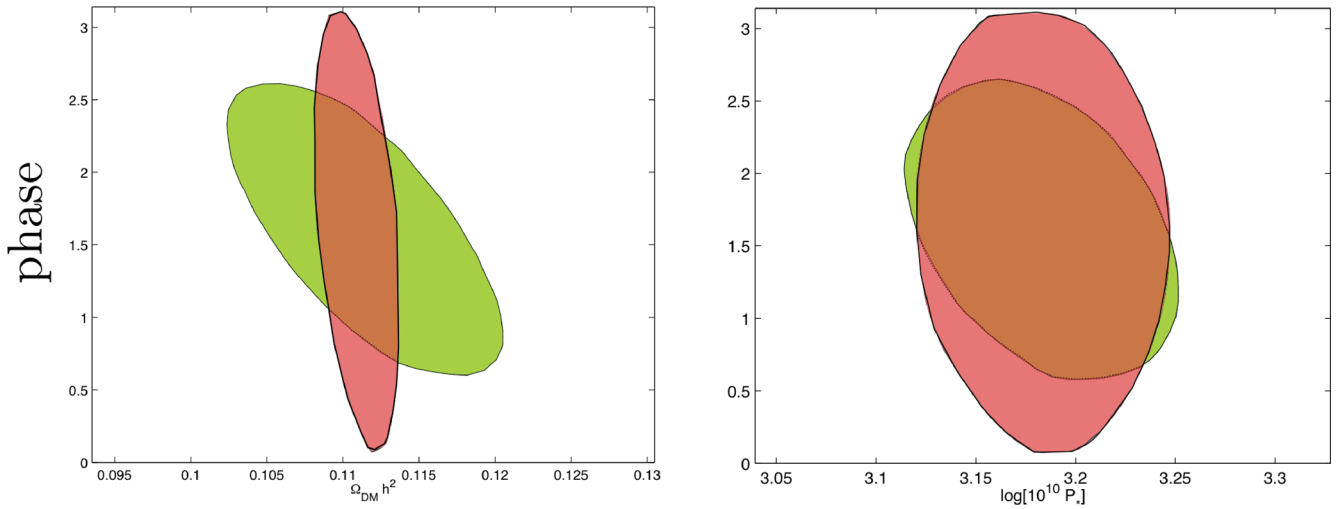

Figure 13. The 68 per cent confidence contours between the phase ζ (NPH, green) and ϕ (BEFT, red) and the dark matter energy density and the amplitude. For sake of comparison, we shifted the central value of ϕ upwards, while contour levels were maintained. Both examples show how the correlation between ζ and P_* and Ω_{DM} affects the uncertainty in these parameters. It does seem that overall uncertainty is conserved, as the error in ϕ is larger than the error in ζ.

best-fitting amplitude of $a_2 = 0.16$. This is close to the best-fitting amplitude found by Flauger et al. (2010) with an improvement of $\Delta\chi^2 \simeq 11$ (*WMAP*5). Their improvement appeared at a (translated) frequency of $a_4 \sim 150$. It is interesting to note that there appear to be subsequent improvements at low frequencies close to 50, 100 (*WMAP*7) and 150 (*WMAP*5), which could be an additional hint we might be looking at an oscillation as opposed to noise, because the improvement appears at equidistant intervals in frequency.

Using an MCMC we have been able to put constraints on several primordial parameters. In particular, we derived constraints on the amplitude a_2, the tilt a_3 and the phase ζ of the oscillatory correction in the NPH model. We used these observational constraints to test the NPH model by implementing the relation between various parameters as predicted by the NPH model. This shows that the signal found at this frequency could originate from NPH modifications to the primordial power spectrum. However, an improvement of the confidence levels would exclude this possibility.

For the BEFT model, we ran one grid and we found that even before varying all parameters in an MCMC we could achieve an improvement of $\Delta\chi^2 \simeq 13$. This improvement corresponded to a frequency $A_1 = 7708$ and an amplitude $A_0 = 10$. We ran an MCMC around this frequency and found that the theoretical bound on A_0 does not probe the observationally best-fitting point. In order to recover the full distribution of A_0, we set an upper limit on $A_0 < 25$ when running the MCMC. We found a best-fitting improvement of $\Delta\chi^2 \sim 16$ compared to no oscillations with a phase of $\phi = -0.21$ and an amplitude of $a_2 = 12.25$. This best-fitting value of A_0 is theoretically not expected by a BEFT modification of the primordial power spectrum. The amplitude of the oscillatory correction is also proportional to a coupling constant β which we set to 1. Once we relax this assumption, larger amplitudes are theoretically possible through βA_0. These results enabled us to put a constraint on the ratio M/H of $5 \times 10^2 \lesssim M/H \lesssim 10^3$, which is a very realistic possibility of this parameter ratio.

It is hard to assess the significance of the improved fits, although serious attempts have been made to quantify any anomalies in the CMB by Bennet et al. (2011), showing that there is no compelling evidence for any of the proclaimed anomalies, such as localized features and multipole alignment. In the special case of oscillations, we could think of two possibilities: investigate simulated CMB data without oscillations with realistic beam and noise, and investigate simulated CMB data *with* oscillations and realistic beam and noise. For example, if we can find oscillations in simulated data without oscillations, this would suggest oscillations can well be mimicked by noise. Complementarily, generating realistic CMB data including primordial oscillations, can we recover these oscillations by analysing the data? In particular, it would be interesting to determine whether there exists a threshold amplitude for oscillations to be recovered. Preferably, we should generate a large number of maps in order to see if we can assess a probability of recovering oscillations from the data. A similar proposal and subsequent analysis was performed by Hamann, Shafieloo & Souradeep (2010) using the a deconvolution method on *WMAP* 5-year data. We will report our findings in future work.

For now, we conclude that a primordial power spectrum with no oscillations is consistent with the data for most amplitude and frequency of the primordial signal. For some frequencies, a non-zero amplitude of the oscillatory corrections to the power spectrum seems to be preferred by the data. These signals can be investigated and used to constrain primordial parameter space, possibly signifying some of the detailed physics driving inflation.

ACKNOWLEDGMENTS

We would like to thank Mark Jackson for very helpful discussions and for reading the proof manuscript. We would also like to thank Christophe Ringeval, who provided some of the essential code which was used throughout this paper. PDM would like to thank Yuri Cavecchi and Marcin Zielinski for helping setting up the code on the LISA super cluster. PDM was supported by the Netherlands Organization for Scientific Research (NWO), NWO-toptalent grant 021.001.040. The research of JPvdS is financially supported by Foundation of Fundamental Research of Matter (FOM) research grant 06PR2510.

REFERENCES

Achucarro A., Gong J. O., Hardeman S., Palma G. A., Patil S. P., 2011, J. Cosmol. Astropart. Phys., 1101, 030
Bennett C. L. et al., 2011, ApJS, 192, 17
Cai R. G., Hu B., Zhang H. B., 2010, J. Cosmol. Astropart. Phys., 1008, 025
Chen X., 2010a, J. Cosmol. Astropart. Phys., 1012, 003
Chen X., 2010b, Adv. Astron., 2010, 638979
Chen X., 2011, preprint (arXiv:1104.1323)
Chen X., Huang M. X., Kachru S., Shiu G., 2007, J. Cosmol. Astropart. Phys., 0701, 002
Chen X., Easther R., Lim E. A., 2008, J. Cosmol. Astropart. Phys., 0804, 010
Covi L., Hamann J., Melchiorri A., Slosar A., Sorbera I., 2006, Phys. Rev. D, 74, 083509
Cyr-Racine F. Y., Schmidt F., 2011, Phys. Rev. D, 84, 083505
Danielsson U. H., 2002, Phys. Rev. D, 66, 023511
Dvorkin C., Hu W., 2011, Phys. Rev. D, 84, 063515
Easther R., Greene B. R., Kinney W. H., Shiu G., 2001, Phys. Rev. D, 64, 103502
Easther R., Greene B. R., Kinney W. H., Shiu G., 2002, Phys. Rev. D, 66, 023518
Easther R., Greene B. R., Kinney W. H., Shiu G., 2003, Phys. Rev. D, 67, 063508
Easther R., Kinney W. H., Peiris H., 2005a, J. Cosmol. Astropart. Phys., 0505, 009
Easther R., Kinney W. H., Peiris H., 2005b, J. Cosmol. Astropart. Phys., 0508, 001
Fergusson J. R., Liguori M., Shellard E. P. S., 2010, preprint (arXiv:1006.1642)
Flauger R., Pajer E., 2011, J. Cosmol. Astropart. Phys., 1101, 017
Flauger R., McAllister L., Pajer E., Westphal A., Xu G., 2010, J. Cosmol. Astropart. Phys., 1006, 009
Gelman A., Rubin D., 1992, Stat. Sci., 7, 457
Greene B. R., Schalm K., van der Schaar J. P., Shiu G., 2004, in Chen P., Bloom E., Madejski G., Patrosian V., eds, Proc. 22nd Texas Symp. on Relativistic Astrophysics at Stanford University, Stanford, CA. p. 1
Groeneboom N. E., Elgaroy O., 2008, Phys. Rev. D, 77, 043522
Hamann J., Covi L., Melchiorri A., Slosar A., 2007, Phys. Rev. D, 76, 023503
Hamann J., Shafieloo A., Souradeep T., 2010, J. Cosmol. Astropart. Phys., 1004, 010
Jackson M. G., Schalm K., 2010, preprint (arXiv:1007.0185)
Jackson M. G., Schalm K., 2011, preprint (arXiv:1104.0887)
Komatsu E., et al., 2011, ApJS, 192, 18
Lewis A., Bridle S., 2002, Phys. Rev. D, 63, 103511
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Martin J., Brandenberger R. H., 2001, Phys. Rev. D, 63, 123501
Martin J., Ringeval C., 2004, Phys. Rev. D, 69, 083515
Martin J., Ringeval C., 2005, J. Cosmol. Astropart. Phys., 0501, 007
Meerburg P. D., 2010, Phys. Rev. D, 82, 063517
Meerburg P. D., van der Schaar J. P., 2011, Phys. Rev. D, 83, 043520
Meerburg P. D., van der Schaar J. P., Corasaniti P. S., 2009, J. Cosmol. Astropart. Phys., 0905, 018

Meerburg P. D., van der Schaar J. P., Jackson M. G., 2010, J. Cosmol. Astropart. Phys., 1002, 001
Pahud C., Kamionkowski M., Liddle A. R., 2009, Phys. Rev. D, 79, 083503
Schalm K., Shiu G., van der Schaar J. P., 2004, J. High Energy Phys., 04, 076
Schalm K., Shiu G., van der Schaar J. P., 2005, AIP Conf. Proc., 743, 362
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437

This paper has been typeset from a TEX/LATEX file prepared by the author.