Model-independent forecasts of CMB angular power spectra for the Planck mission

Amir Aghamousa[*]

Department of Physics, University of Pune, Pune 411 007 India and Centre for Modeling and Simulation, University of Pune, Pune 411 007, India

Mihir Arjunwadkar[†]

National Centre for Radio Astrophysics, University of Pune Campus, Pune 411 007, India and Centre for Modeling and Simulation, University of Pune, Pune 411 007, India

Tarun Souradeep[‡]

Inter-University Centre for Astronomy and Astrophysics, University of Pune Campus, Pune 411 007, India

The Planck mission, designed for making measurements of the cosmic microwave background (CMB) radiation with unprecedented accuracy and angular resolution, is expected to release its entire data in the near future. In this paper, we provide model-independent forecasts for the TT, EE, and TE angular power spectra for the Planck mission using synthetic data based on the best-fit Lambda cold dark matter (ΛCDM) model. The nonparametric function estimation methodology we use here is based on the agnostic viewpoint of allowing the data to speak for themselves rather than letting the models decide what is inferred from the data. Our analysis indicates that the three Planck angular power spectra will be determined sufficiently well for $2 \leq l \lesssim l_{\max}$, where $l_{\max} = 2500(TT)$, $1377(EE)$, and $1727(TE)$ respectively. A key signature of reionization, namely, a bump at low values of l, is evident in our forecasts for the EE and TE power spectra. Nonparametric confidence bands in the phase shift (ϕ_m) versus acoustic scale (l_A) plane, corresponding to the first eight peaks in the TT power spectrum, show a confluence region for $300 \lesssim l_A \lesssim 305$ which is in good agreement with the estimate $l_A = 300$ based on the best-fit ΛCDM model. From our results, we expect that the final Planck data should lead to accurate model-independent estimates of CMB angular power spectra using our nonparametric regression formalism.

I. INTRODUCTION

Observations of the microwave sky reveal that the temperature of the cosmic microwave background (CMB) radiation is not exactly the same in all directions. These small fluctuations in temperature are imprinted on the entire sky, implying that the CMB is anisotropic. These primordial anisotropies were first discovered in 1992 by the cosmic background explorer (COBE). This was followed by a remarkable series of ground-based and balloon-borne experiments, and more recently by the Wilkinson microwave anisotropy probe (WMAP). These fluctuations are believed to have been generated within 10^{-35} seconds of the big bang. CMB anisotropies are therefore a rich source of information about the early universe, and have revolutionized the way we understand our Universe. A study of CMB anisotropies also helps in probing the fundamental physics at energy scales much higher in magnitude compared to those accessible to

particle accelerators. CMB anisotropies are sensitive to classical cosmology parameters such as expansion rate, curvature, cosmological constant, matter content, radiation content, and baryon fraction, and provide insights for modeling structure formation in the Universe [1]. For example, measurements of the CMB anisotropies with ever-increasing precision have made it possible to establish a standard cosmological model that asserts that the Universe is spatially nearly flat [2].

The CMB contains an additional wealth of information about the Universe through its (linear) polarization. The CMB has acquired linear polarization through Thomson scattering during either decoupling or reionization, and sourced by the quadrupole anisotropy in the radiation distribution at that time [3,4]. The dependence of CMB polarization on cosmological parameters differs from that of temperature anisotropies. As such, it helps break degeneracies and constrain cosmological parameters better. Accurate measurements of CMB polarization will therefore enable us to affirm the validity and consistency of different cosmological models [5–7].

The Planck mission [8] is a space-based full-sky probe for third-generation CMB experiments designed to make

[*]amir@cms.unipune.ac.in; a.aghamousa@gmail.com
[†]mihir@ncra.tifr.res.in
[‡]tarun@iucaa.ernet.in

extensive measurements of CMB polarization. Indeed, one of the main objectives of this mission is to measure primordial fluctuations in the CMB with an accuracy prescribed by the fundamental astrophysical limits, through improvements in sensitivity and angular resolution, and through better control over noise and confounding foregrounds. The higher angular resolution of Planck implies that higher-order peaks in the CMB angular spectra can be determined with better precision, which in turn translates to determination of cosmological parameters (such as baryon and dark matter densities) with improvement in statistical precision by an order of magnitude. Furthermore, the Planck mission is designed to make extensive measurements of the E-mode polarization spectrum over multipoles up to $l \approx 1500$ with unprecedented precision, together with good control over polarized foreground noise. These measurements are therefore expected to provide insights into the physics of the early universe, epoch of recombination, structure formation, allowable modes of primordial fluctuations (adiabatic versus isocurvature modes), reionization history of the Universe, etc., and help in establishing constraints on the primordial power spectrum. The Planck mission will also help constrain the fundamental physics at high energies which are impossible to probe through terrestrial experiments [8].

In our previous work [9], we estimated the CMB TT power spectrum for four phased WMAP data releases using a nonparametric function estimation methodology [10,11]. This methodology does not impose any specific form or model on the power spectrum. It determines the best fit by optimizing a measure of smoothness that depends only on characteristics of the data. This ensures that the fit and the subsequent analysis is approximately model independent for sufficiently large data sizes. Further, this methodology quantifies uncertainties in the fit in the form of a high-dimensional ellipsoidal confidence set. This confidence set is centered at the fit, and captures the true but unknown power spectrum with a prespecified probability. This confidence set is the prime inferential object of this methodology, and allows addressing complex inferential questions about the true but unknown power spectrum meaningfully within a unified framework.

The Planck mission recently released a partial TT data set, and is expected to release its final and entire CMB temperature and E-mode polarization data sets in the near future. In this paper, we therefore attempt to make model-independent forecasts for the three CMB power spectra using synthetic Planck-like data conforming to the specifications and characteristics of the Planck mission. These synthetic data are based on the assumption that the best-fit ΛCDM model [12] is the true model of the Universe.

In what follows, Sec. II briefly describes the three CMB angular power spectra studied here. This is followed by a description (Sec. III) of the synthetic data used in this work, our results (Sec. IV), and conclusions (Sec. V).

II. TEMPERATURE AND POLARIZATION POWER SPECTRA

The polarization of radiation is commonly characterized in terms of the Stokes parameters I, Q, U and V [13]. It is well known that the CMB cannot have circular polarization, so we only consider linear polarization. The parameters Q and U which describe linear polarization constitute a rank-2 symmetric trace-free tensor $\mathcal{P}_{ab}$. Any two-dimensional symmetric tensor can be represented in the form of two scalar fields. Such factorization is unique, and is similar to the decomposition of a vector field into a gradient and a divergence-free component [14]. Conventionally, $\mathcal{P}_{ab}$ is therefore expressed in terms of its electric (i.e., gradient) and magnetic (i.e., curl) components, P_E and P_B respectively.

For the CMB, the P_E and P_B scalars are defined on a sphere, and expanded in terms of spherical harmonics as

$$P_E(\theta, \phi) = \sum_{l \geq 2} \sum_{|m| \leq l} \sqrt{\frac{(l-2)!}{(l+2)!}} a_{lm}^E Y_{lm}(\theta, \phi), \quad (1)$$

$$P_B(\theta, \phi) = \sum_{l \geq 2} \sum_{|m| \leq l} \sqrt{\frac{(l-2)!}{(l+2)!}} a_{lm}^B Y_{lm}(\theta, \phi), \quad (2)$$

which define the E- and B-mode multipoles a_{lm}^E and a_{lm}^B respectively. Any CMB measurement can be decomposed into three maps (T, E and B respectively). Hence, a total of six angular power spectra (TT, EE, BB, TE, TB, and EB) can be obtained from these three components [15]. These six power spectra are defined by expanding the T, E and B maps in terms of spherical harmonics, resulting into the following correlation structure:

$$\langle a_{lm}^{Y*} a_{l'm'}^{Y'} \rangle = C_l^{YY'} \delta_{ll'} \delta_{mm'}, \quad (3)$$

where Y, Y' are E, B or T. In the absence of parity violation and assuming Gaussian fluctuations, temperature and polarization anisotropies of the CMB are described by the C_l^{TT}, C_l^{EE}, C_l^{BB}, C_l^{TE} power spectra completely [4,5,14]. Since B-mode polarization is not expected to be detected by the Planck mission accurately [8], we focus only on the C_l^{TT}, C_l^{EE} and C_l^{TE} power spectra.

The standard deviation of the C_l^{TT}, C_l^{EE} and C_l^{TE} power spectra is approximately given by [8,16]

$$\Delta C_l^{TT} \simeq \frac{2}{(2l+1)f_{\text{sky}}} (C_l^{TT} + \omega_T^{-1} W_l^{-2})^2 \quad (4)$$

$$\Delta C_l^{EE} \simeq \frac{2}{(2l+1)f_{\text{sky}}} (C_l^{EE} + \omega_P^{-1} W_l^{-2})^2 \quad (5)$$

$$\Delta C_l^{TE} \simeq \frac{2}{(2l+1)f_{\text{sky}}} ((C_l^{EE} + \omega_P^{-1} W_l^{-2})$$
$$\times (C_l^{TT} + \omega_T^{-1} W_l^{-2}) + (C_l^{TE})^2), \qquad (6)$$

where $\omega_T \equiv (\sigma_{p,T}\theta_{\text{FWHM}})^{-2}$ and $\omega_P \equiv (\sigma_{p,P}\theta_{\text{FWHM}})^{-2}$ are the weights per solid angle for temperature and polarization, and f_{sky} is the fraction of observed sky in the experiment. The $\sigma_{p,T}$ and $\sigma_{p,P}$ are noise standard deviations per resolution element ($\theta_{\text{FWHM}} \times \theta_{\text{FWHM}}$). The window function for a Gaussian beam is

$$W_l = \exp\left(-\frac{l(l+1)}{2l_{\text{beam}}^2}\right), \qquad (7)$$

where $l_{\text{beam}} = \sqrt{8\ln 2}/\theta_{\text{FWHM}}$.

The importance of decomposing the $\mathcal{P}_{ab}$ tensor in terms of the E and B modes comes from the fact that linear scalar perturbations do not generate any B-mode polarization [4]. The tensor mode contributes to both E as well as B modes, whereas the vector mode contributes only to the B-mode polarization. Therefore, the E part of the decomposition stems from the scalar and tensor modes, whereas the B part originates in the vector and tensor modes [4,8,14].

Cosmological relevance of these polarization modes is as follows. In contrast to the temperature anisotropies which originate in photon density fluctuations at the last scattering, the E mode follows the velocity of the cosmological plasma at decoupling. The E mode therefore contains greater information about cosmological parameters such as baryon and cold dark matter densities [6,14]. Polarization power spectra have an oscillatory structure which is analogous to that of the TT power spectrum. For example, peaks in the EE power spectrum are out of phase with those in the TT power spectrum due to anisotropy generated at the last scattering. The TE power spectrum, which has a higher amplitude compared to the EE power spectrum, is a measure of the correlations (positive or negative) between density and velocity fluctuations [14,17]. Further, the phase difference between acoustic peaks in the TT, EE and TE power spectra can be used as a model-independent check for the physics of acoustic oscillations [14].

Adiabatic and isocurvature perturbations also have different effects on the phase of the CMB polarization spectra. Predicted polarization power spectra for isocurvature perturbations show out-of-phase peaks and dips compared to those for adiabatic perturbations. As a result, power spectra from isocurvature perturbations appear to be l-shifted versions of those for adiabatic perturbations [18]. A meaningful estimation of the polarization power spectra can therefore be used to determine which of the two scenarios is closer to truth.

An important cosmological phenomenon that affects the polarization spectra is reionization. Reionization of the Universe started when the first generation of stars started producing a flux of photons. The resulting free electrons started rescattering the CMB radiation. Although only a small fraction of CMB photons got scattered this way during the reionization era, the imprints of reionization are expected to be seen as distortions in the polarization power spectra at large angular scales of the order of 10 deg. The height and location of the reionization bump [19–21] expected at low multipoles ($l \lesssim 20$) has information related to total optical depth and the reionization epoch redshift [6,20]. Although the precision of reionization bump detection is limited by cosmic variance at low l [22], constraining it will help understand the reionization history better. It will also help break degeneracies between several cosmological parameters by constraining the optical depth better [6,7].

III. SYNTHETIC DATA FOR THE PLANCK MISSION

For forecasting CMB angular power spectra for the Planck mission, we generate the synthetic Planck-like data using the FuturCMB code [23]. FuturCMB generates a simulated angular power spectrum using a user-provided theoretical power spectrum C_l^{true} (which is assumed to be the true spectrum) for frequency channels representing Planck measurements, and generates the corresponding noise power spectrum N_l conforming to the Planck characteristics. This is done by generating a random realization of the spherical harmonic coefficients a_{lm}, assumed to be Gaussian random variables with mean zero and variance

$$\text{Var}(a_{lm}) = C_l^{\text{true}} + N_l. \qquad (8)$$

For C_l^{true}, we use spectra generated using CAMB [24] for the best-fit ΛCDM cosmological parameters obtained from the WMAP 7-yr data [12]. We also limit FuturCMB to $l \leq 2500$, a range that corresponds to the three Planck frequency channels (100, 142 and 217 GHz). N_l is the noise power spectrum given by

$$N_l = \omega^{-1} W_l^{-2}, \qquad (9)$$

where ω is ω_T and ω_P for temperature and polarization respectively, and W_l is the window function for a Gaussian beam [Eq. (7)]. The noise in the TE power spectrum is taken to be zero because noise contributions from different maps are uncorrelated [23]. FuturCMB then calculates the power spectra data C_l^{map} as

$$C_l^{\text{map}} = \frac{1}{2(l+1)} \sum_{m=-l}^{+l} |a_{lm}|^2. \qquad (10)$$

Equation (10) is an unbiased estimator of $\text{Var}(a_{lm})$ [Eq. (8)]. Its expected value is therefore equal to $C_l^{\text{true}} + N_l$. The noise spectra are expected to dominate over the true spectrum for sufficiently high values of l. This is seen in Figs. 1–3, where
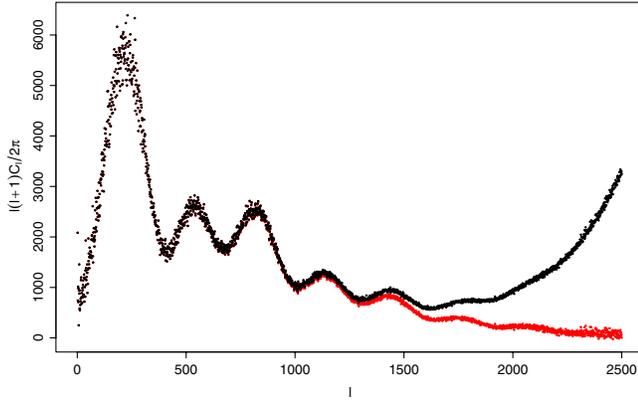
FIG. 1 (color online). A realization of simulated TT power spectrum data for the Planck mission, generated using FuturCMB [23]. Black points: data including noise; red points: simulated data after subtracting noise.
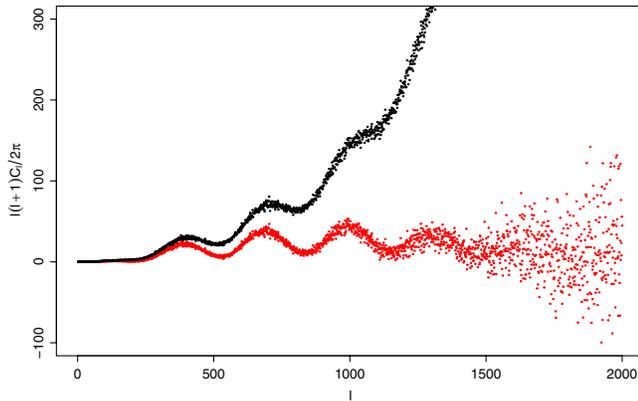
FIG. 2 (color online). A realization of simulated EE power spectrum data for the Planck mission, generated using FuturCMB [23]. Black points: data including noise; red points: simulated data after subtracting noise.
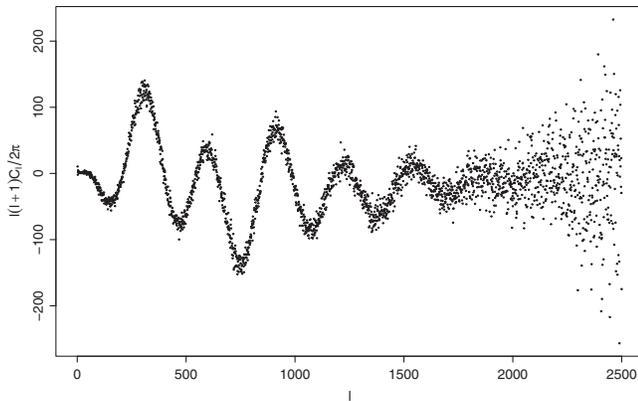
FIG. 3. A realization of simulated TE power spectrum data for the Planck mission, generated using FuturCMB [23]. FuturCMB assumes the noise to be zero for the TE data.

the black points represent the FuturCMB output C_l^{map} of the TT, EE and TE spectra: The upward trends in the tail of TT and EE spectra are the result of noise dominating the data at high ls. The TE power spectrum, on the other hand, does not show any such upward trend because the noise in the TE spectrum is assumed zero. To obtain synthetic TT and EE data, we therefore subtract the corresponding noise spectra from the C_l^{map} output of FutureCMB (Figs. 1–2, red points). The covariance matrix of the simulated angular power spectra is taken to be a diagonal matrix with diagonal elements defined in Eqs. (4), (5) and (6) for TT, EE, and TE respectively.

IV. RESULTS AND DISCUSSION

A. Nonparametric fits to synthetic Planck data

For estimating the power spectra from synthetic Planck data, we use the nonparametric regression method described in [9]. While our synthetic data are generated under the assumption that the ΛCDM model as estimated from the WMAP 7-yr data is the true model of the Universe, this formalism for nonparametric regression and inference itself does not make any assumptions about the shape of the true regression function underlying the data; it is asymptotically model independent. Further, while the correlations in our synthetic data are zero by construction, this regression formalism is capable of incorporating a nondiagonal covariance matrix describing correlated noise.

Under this formalism, the nonparametric fit can be characterized by its effective degrees of freedom (EDoF), which can be thought of as the equivalent of the number of parameters in a parametric regression problem. Using this methodology, we obtain nonparametric fits to the synthetic TT, EE and TE data by appropriately constraining the EDoF of the fits.

Figures 4, 5 and 6 show nonparametric fits (red curves) to the TT, EE, and TE data respectively, which are in good agreement with the underlying ΛCDM spectra (C_l^{true}, black curves) used to generate the synthetic data. This shows that this nonparametric regression methodology, which does not assume any specific form of the true (but generally unknown) regression function, can recover the underlying true spectrum with high accuracy especially where noise levels are not too high. Additional details about these nonparametric fits can be found in Appendices A and B.

B. How well will the final Planck spectra be determined by data alone?

To see how noise in the data affects local uncertainties in a fitted spectrum, we compute approximate 95% confidence intervals for each fitted C_l using 5000 randomly sampled spectra from the corresponding confidence set. The ratio of this confidence interval to the absolute value of the fitted $|C_l|$, assumed to be nonzero, is a relative measure
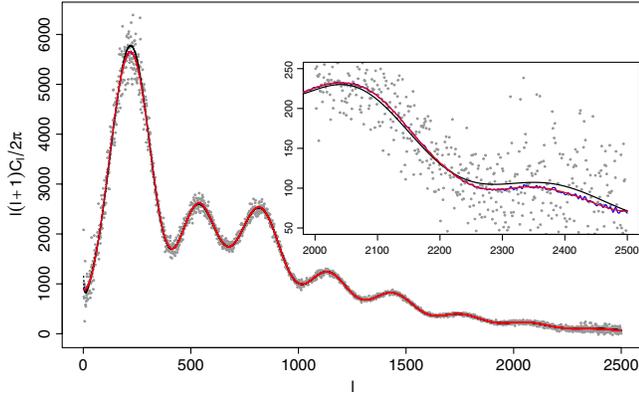
FIG. 4 (color online).   *TT* nonparametric fits. Blue: full-freedom fit (EDoF ≈ 72), red: restricted-freedom fit (EDoF = 27), black: best-fit ΛCDM spectrum, grey: simulated data realization.
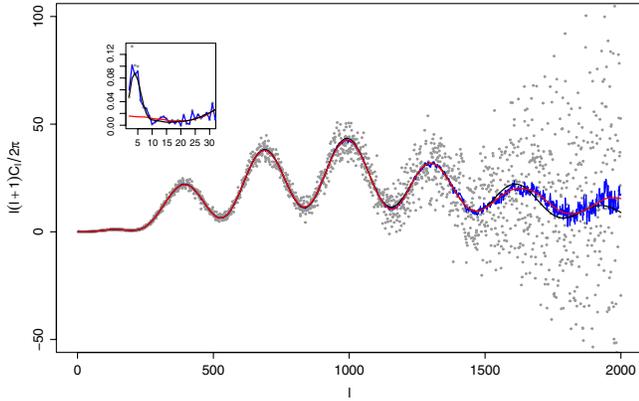
FIG. 5 (color online).   *EE* nonparametric fits. Blue: full-freedom fit (EDoF ≈ 190), red: restricted-freedom fit (EDoF = 24), black: best-fit ΛCDM spectrum, grey: simulated data realization.
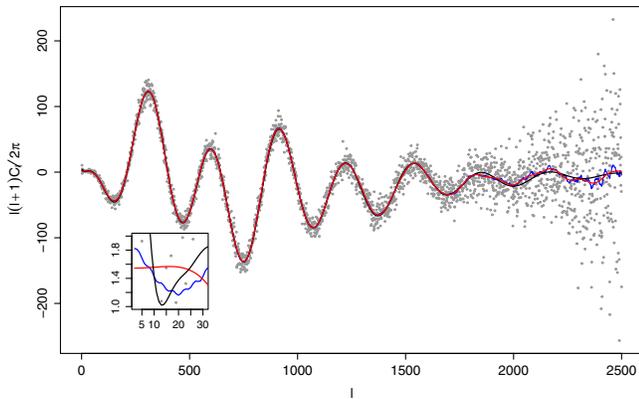
FIG. 6 (color online).   *TE* nonparametric fits. Blue: full-freedom fit (EDoF ≈ 95), red: restricted-freedom fit (EDoF = 40), black: best-fit ΛCDM spectrum, grey: simulated data realization.
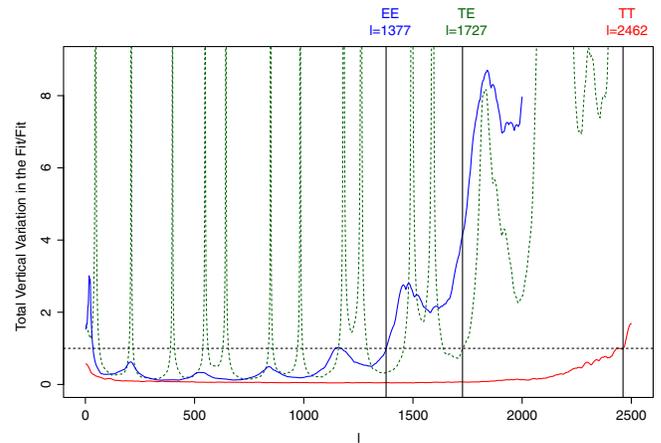
FIG. 7 (color online).   The results of a probe of the confidence sets for the *TT* (red), *EE* (blue), and *TE* (green) nonparametric restricted-freedom fits to the synthetic Planck data, to determine how well the fits are expected to be determined by the data alone. The quantity plotted for each data realization is the total vertical variation at each l within the respective 95% (2σ) confidence set, divided by the absolute value of the fit (assumed nonzero): Values ≪ 1 indicate that the fit is tightly determined by the data, whereas values ≳1 indicate that the data contain very little information about the height of the angular power spectrum at that l. Disregarding the low-l region for the *EE* fit, and spikes for the *TE* fit (which arise from nearly zero fitted C_l values), the marked vertical lines indicate the approximate l value at which each curve rises above 1.

of how well each fitted C_l is determined [9,11]: A value ≪ 1 implies that the fit is well determined by the data, and a value ≳1 implies that the data contain very little information about height of the power spectrum at that l. In Fig. 7, we plot this relative error (95%) for all three spectra as a function of the multipole index l. We see that, by this criterion, the Planck power spectra are expected to be well determined up to $l \approx 2462(TT)$, $1377(EE)$ and $1727(TE)$. Since the *TE* fit oscillates around zero (Fig. 6), this quantity takes very high values at l's where the *TE* spectrum has a nearly zero value. This results in multiple spikes in Fig. 7 (green dash curve), but this does not imply that the fit is ill determined at these l's. Ignoring these spikes, we see that the relative error in the *TE* fit is below unity up to $l \approx 1727$, which indicates the range over which this fit is expected to be well determined by data alone.

C. Uncertainties on the locations and heights of peaks and dips

Locations and heights of peaks and dips in the CMB power spectra are governed by cosmological parameters. Uncertainties in the location and height of a peak or a dip in a fitted spectrum can thus help assess uncertainties in the values of related parameter. Following the procedure outlined in [9], we sampled the 95% confidence set of each fitted spectrum uniformly to generate spectral variations

while ensuring that at least 5000 of these are acceptable (see Appendix B for details). Figures 8, 9, 10, and show the results of this exercise, together with tabulated values in Tables I, II, and III respectively. The box around a peak or a dip represents the largest horizontal and vertical variations in the scatter; these represent the 95% confidence intervals on the location and height of a peak or a dip.

For the TT fit, these boxes around peaks and dips are tiny (Fig. 8), which is a reflection of the accuracy of the Planck TT data. Such precise determination of peaks and dips will clearly lead to more robust estimates of related

FIG. 10 (color online). 95% confidence boxes the locations and heights of peaks and dips in the TE fit. Black curve is the restricted-freedom monotone fit to the synthetic Planck TE data (grey points). The number of acceptable spectral variations sampled from the 95% confidence set is 5000. These uncertainties are tabulated in Table III.
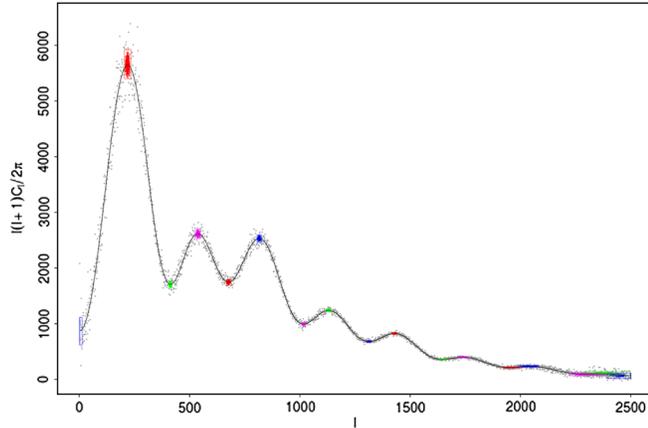
FIG. 8 (color online). 95% confidence boxes the locations and heights of peaks and dips in the TT fit. Black curve is the restricted-freedom monotone fit to the synthetic Planck TT data (grey points). The number of acceptable spectral variations sampled from the 95% confidence set is 5000. These uncertainties are tabulated in Table I.
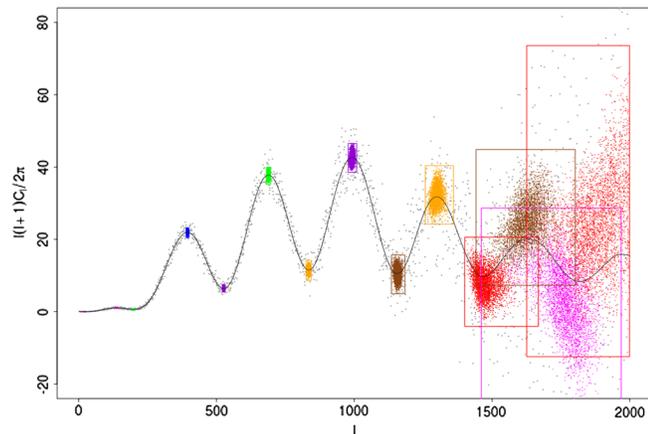
FIG. 9 (color online). 95% confidence boxes the locations and heights of peaks and dips in the EE fit. Black curve is the restricted-freedom monotone fit to the synthetic Planck EE data (grey points). The number of acceptable spectral variations sampled from the 95% confidence set is 5000. These uncertainties are tabulated in Table II.

cosmological parameters than what is currently available. The theoretical TT power spectrum (Fig. 4, black curve) shows a small upturn at low l. This upturn is primarily the result of the integrated Sachs-Wolf effect. In Fig. 8, this upturn corresponds to the first dip in the spectral variations sampled from the 95% confidence set.

Peaks and dips in the EE power spectrum show reasonably low uncertainties up to $l \approx 1200$ (Fig. 9). Beyond this, the data contain high levels of noise, and therefore all uncertainties become much larger. This is in agreement with the behavior of the EE curve in Fig. 7 (blue curve). We also expect a small bump in the EE power spectrum which is related to the epoch of reionization. This bump is indeed seen in full-freedom nonparametric EE fit in Fig. 5. The restricted-freedom fit, however, does not recover this bump except for a weak peak at low multipoles ($l = 20$). We believe this to be more a limitation of methodology. The restricted-freedom fit may not be appropriate at very low multipoles since features at low multipole will not be captured in a limited cosine expansion. The 95% uncertainties of locations and heights of peaks and dips are reported in Table II. The uncertainty boxes on peaks and dips in the TE fit (Fig. 10) are reasonably small until $l \approx 1800$, again in agreement with the result depicted in Fig. 7 (green dashed curve). The somewhat peculiar uncertainty boxes on the last three peaks in the TE fit are due to the fact that there are two spurious peaks at $l = 1920$ and $l = 2070$ in the restricted-freedom fit (see Appendix A) which are a result of the high noise levels at high ls. At low multipoles, the TE fit also shows a bump (Fig. 6) which is related to the epoch of reionization. Although a similar bump in the EE fit is known to be more informative [8], a determination of this peak in the TE spectrum should also lead to useful information about reionization.

TABLE I.　95% Confidence interval on several features of TT angular power spectrum.

Peak location	Peak height	Dip location	Dip height
l_1:(204, 234)	h_1:(5401.543, 5925.611)	$l_{1+\frac{1}{2}}$:(406, 421)	$h_{1+\frac{1}{2}}$:(1638.521, 1769.197)
l_2:(525, 549)	h_2:(2527.018, 2691.862)	$l_{2+\frac{1}{2}}$:(666, 688)	$h_{1+\frac{1}{2}}$:(1689.753, 1793.608)
l_3:(805, 827)	h_3:(2465.483, 2596.654)	$l_{3+\frac{1}{2}}$:(1009, 1028)	$h_{1+\frac{1}{2}}$:(969.1924, 1019.3422)
l_4:(1120, 1141)	h_4:(1215.164, 1267.128)	$l_{4+\frac{1}{2}}$:(1301, 1325)	$h_{1+\frac{1}{2}}$:(663.6103, 695.9041)
l_5:(1418, 1439)	h_5:(806.5287, 845.0737)	$l_{5+\frac{1}{2}}$:(1630, 1666)	$h_{1+\frac{1}{2}}$:(348.6765, 370.1998)
l_6:(1714, 1761)	h_6:(387.0516, 412.0869)	$l_{6+\frac{1}{2}}$:(1922, 2007)	$h_{1+\frac{1}{2}}$:(199.8328, 225.3753)
l_7:(1989, 2085)	h_7:(219.8173, 251.4042)	$l_{7+\frac{1}{2}}$:(2220, 2442)	$h_{1+\frac{1}{2}}$:(64.88648, 112.57429)
l_8:(2282, 2498)	h_8:(71.91322, 141.91509)	$l_{8+\frac{1}{2}}$:(2394, 2500)	$h_{1+\frac{1}{2}}$:(15.27434, 111.02498)

TABLE II.　95% Confidence interval on several features of EE angular power spectrum.

Peak location	Peak height	Dip location	Dip height
l_1:(7, 26)	h_1:(−0.0025, 0.0235)	$l_{1+\frac{1}{2}}$:(17, 27)	$h_{1+\frac{1}{2}}$:(−0.0029, 0.0181)
l_2:(133, 144)	h_2:(0.9234, 1.2542)	$l_{2+\frac{1}{2}}$:(193, 205)	$h_{2+\frac{1}{2}}$:(0.4620, 0.8600)
l_3:(390, 401)	h_3:(20.4654, 23.2528)	$l_{3+\frac{1}{2}}$:(521, 532)	$h_{3+\frac{1}{2}}$:(5.4424, 7.5745)
l_4:(681, 697)	h_4:(35.2102, 39.9650)	$l_{4+\frac{1}{2}}$:(825, 844)	$h_{4+\frac{1}{2}}$:(8.6629, 14.2180)
l_5:(978, 1009)	h_5:(38.5544, 46.5531)	$l_{5+\frac{1}{2}}$:(1135, 1184)	$h_{5+\frac{1}{2}}$:(4.9459, 15.6909)
l_6:(1258, 1360)	h_6:(24.2445, 40.3917)	$l_{6+\frac{1}{2}}$:(1401, 1668)	$h_{6+\frac{1}{2}}$:(−4.0435, 20.6010)
l_7:(1442, 1802)	h_7:(7.3139, 44.8427)	$l_{7+\frac{1}{2}}$:(1461, 1969)	$h_{7+\frac{1}{2}}$:(−33.4637, 28.6720)
l_8:(1626, 1999)	h_8:(−12.4126, 73.5849)	$l_{8+\frac{1}{2}}$:(1755, 2000)	$h_{8+\frac{1}{2}}$:(−46.4440, 65.6556)

TABLE III.　95% Confidence interval on several features of TE angular power spectrum.

Peak location	Peak height	Dip location	Dip height
l_1:(3, 38)	h_1:(0.6092, 2.5822)	$l_{1+\frac{1}{2}}$:(142, 158)	$h_{1+\frac{1}{2}}$:(−49.8581, −37.7906)
l_2:(301, 316)	h_2:(113.1289, 136.2571)	$l_{2+\frac{1}{2}}$:(461, 477)	$h_{2+\frac{1}{2}}$:(−83.9852, −68.6596)
l_3:(588, 605)	h_3:(26.2370, 44.0873)	$l_{3+\frac{1}{2}}$:(741, 758)	$h_{3+\frac{1}{2}}$:(−145.3450, −124.9381)
l_4:(903, 923)	h_4:(56.4278, 76.1789)	$l_{4+\frac{1}{2}}$:(1061, 1086)	$h_{4+\frac{1}{2}}$:(−94.7879, −76.2354)
l_5:(1199, 1243)	h_5:(4.2335, 23.1510)	$l_{5+\frac{1}{2}}$:(1353, 1400)	$h_{5+\frac{1}{2}}$:(−76.0768, −55.1005)
l_6:(1511, 1579)	h_6:(1.9257, 25.1787)	$l_{6+\frac{1}{2}}$:(1650, 1750)	$h_{6+\frac{1}{2}}$:(−49.6416, −25.6796)
l_7:(1772, 1954)	h_7:(−14.4294, 13.7468)	$l_{7+\frac{1}{2}}$:(1822, 2081)	$h_{7+\frac{1}{2}}$:(−42.0202, −2.8288)
l_8:(1884, 2208)	h_8:(−20.8813, 37.7615)	$l_{8+\frac{1}{2}}$:(1959, 2282)	$h_{8+\frac{1}{2}}$:(−50.6428, 11.4636)
l_9:(2023, 2354)	h_9:(−21.1805, 40.5083)	$l_{9+\frac{1}{2}}$:(2073, 2419)	$h_{9+\frac{1}{2}}$:(−83.8615, 10.2379)
l_10:(2155, 2499)	h_10:(−48.9460, 75.1592)	$l_{10+\frac{1}{2}}$:(2257, 2500)	$h_{10+\frac{1}{2}}$:(−116.2237, 60.3390)

D. Are the acoustic peaks in the TT and EE spectra out of phase with respected to each other?

From the fundamental physics of the CMB anisotropies, we expect the acoustic peaks in the TT and EE power spectra to be out phase with respect to each other. One way of establishing this is by considering the ratio of peak locations in the EE fit to the corresponding ones in the TT fit, which should be $(m+0.5)/m$ for the mth peak [18]. In Fig. 11, we depict the 95% confidence intervals on peak locations in the TT power spectrum (blue) against the corresponding intervals for the EE spectrum (red). Also plotted are the peak location pairs corresponding to the best-fit ΛCDM model (black dots), and points (green) based on the approximate expectation $(m+0.5)/m$. Since the CMB power spectrum is a two-dimensional projection of three-dimensional acoustic oscillation, the analytic approximation (green dots) relating periodicity three-dimensional wave number to a corresponding periodicity in (two-dimensional wave number) multipoles, l works only at large l. Not surprisingly, the deviation from $(m+0.5)/m$ behavior is also seen in the theoretical prediction. What is remarkable and relevant is that the data recovers theoretical prediction of ΛCDM model with adiabatic initial conditions for perturbations so well.

E. An estimate of the acoustic scale parameter l_A

Table I lists the 95% confidence intervals on peak and dip locations and heights for the TT power spectrum fit. As a way of illustrating the role of these uncertainties in the estimation of cosmological parameters, we consider, for TT power spectrum, the following relationship $l_m = l_A(m - \phi_m)$ [25,26] between the location l_m of the mth peak, the
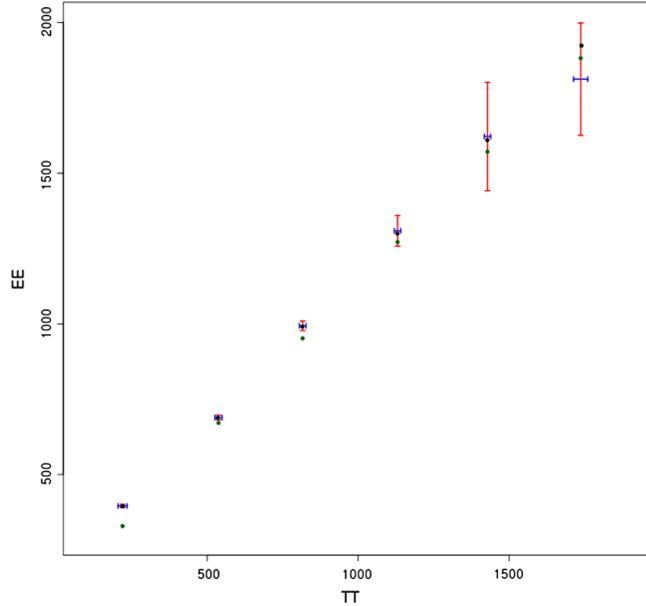
FIG. 11 (color online). The 95% confidence intervals on peak locations in the *EE* fit (red) plotted against the corresponding confidence intervals in the *TT* fit (blue). Also plotted are the peak location pairs corresponding to the best-fit ΛCDM model (black dots), and points (green) based on the theoretical expectation $(m + 0.5)/m$. All the plotted quantities are by and large consistent with each other, indicating that the expected behavior of out-of-phase peak locations is indeed vindicated by the data.

acoustic scale l_A, and the phase shift parameter ϕ_m. In this formulation, the physically meaningful range of values for ϕ_m is $|\phi_m| < 1$. If we substitute the end-points of the 95% confidence interval for the *m*th peak location, then this relationship results into hyperbolic confidence bands in the l_A-ϕ_m plane (Fig. 12). The intersection of these bands (for the first 8 peaks in the *TT* fit) determine an estimated confidence interval for the acoustic scale $300 \leq l_A \leq 305$
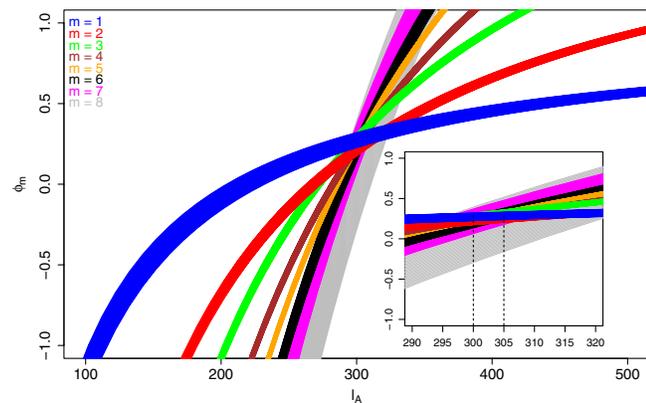
FIG. 12 (color online). Confidence "bands" for the acoustic scale l_A and the shift ϕ_m for the *m*th peak, as derived from the 95% confidence intervals on the first eight peak locations of estimated *TT* power spectrum.
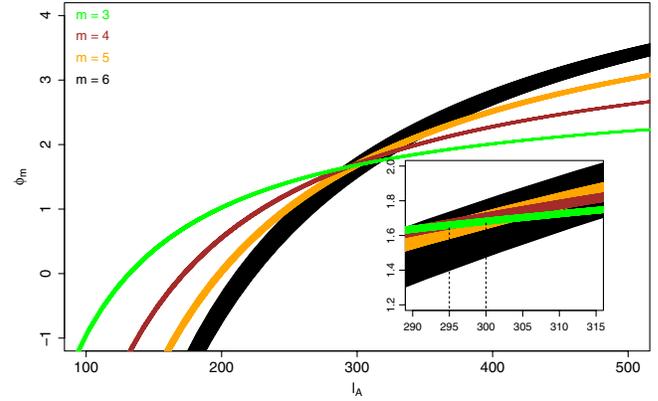
FIG. 13 (color online). Confidence "bands" for the acoustic scale l_A and the shift ϕ_m for the *m*th peak, as derived from the 95% confidence intervals on the first four acoustic peak locations of estimated *EE* power spectrum.

which is in agreement with the reported value $l_A = 300$ by [27]. Similarly, Fig. 13 depicts ϕ_m versus l_A for first 4 acoustic peaks in the *EE* fit. The intersection of bands estimate acoustic scale in the range of $295 \leq l_A \leq 300$. The corresponding ϕ values in both plots ($\phi \approx 0.25$ for *TT* and $\phi \approx 1.65$ for *EE*) indicate *TT* and *EE* power spectra are out of phase. In comparison with our previous estimate $200 \leq l_A \leq 400$ [9] based on the WMAP 7-yr data, the estimate from Planck is expected to improve remarkably. Furthermore, any additional information about the phase shifts ϕ_m can lead to an even more refined estimate for acoustic scale.

V. CONCLUSION

While a partial *TT* power spectrum data set has been released by the Planck mission, the final *TT* and polarization data are yet to be released. In this paper, we have therefore attempted to gleam at what the final *TT*, *EE* and *TE* power spectrum data from the Planck mission might yield when they are released.

For the present analysis, we have used synthetic Planck-like data sets based on the best-fit ΛCDM model and conforming to the characteristics of the Planck mission. We have analyzed these synthetic data using a nonparametric regression and inference methodology which yields data-driven, nearly model-independent estimates and inferences for power spectra and cosmological parameters. This methodology is agnostic in the sense that it does not assume any particular cosmological model to be the truth, and hence allows the data to speak for themselves as much as possible.

Our nonparametric *TT* power spectrum fit has all the peaks resolved well up to $l \leq 2500$, and we see that the *EE* and *TE* power spectra are reasonably well determined as well. We expect the locations and heights of peak and dip in the power spectra to be determined quite precisely, leading

to stronger constraints on cosmological parameters and derived quantities such as the acoustic scale l_A. Our results further lead us to expect that the out-of-phase peaks in the TT and EE spectra will reconfirm the fundamental physics of acoustic oscillations far better.

We therefore conclude that the final Planck data, when released, should lead to a far better understanding of the reionization history of the Universe and help resolve the issue of adiabatic versus isocurvature perturbations through much tighter constraints on the power spectra and cosmological parameters, even with an agnostic, model-independent approach to estimation and inference.

APPENDIX A: FULL-FREEDOM AND RESTRICTED-FREEDOM NONPARAMETRIC FITS

The full-freedom fits are obtained by minimizing the risk function subject to monotonicity constraints on the shrinkage parameters [9]. Such full-freedom fits can be quite oscillatory especially where the noise levels in the data are high. Despite this fit being a reasonable fit (in the sense that it captures the essential trend in the data well), all cosmological models expect far smoother angular power spectra. To account for this, we minimize the risk again, restricting the EDoF of the fit to a value less than that for the full-freedom fit. We continue reducing the EDoF in this way until we obtain an acceptably smooth fit. We call this the *restricted-freedom fit*.

The full-freedom fit for the TT power spectrum is almost smooth, but with tiny wiggles (Fig. 4, blue curve), and corresponds to EDoF ≈ 72. We achieve a numerically smooth fit for the TT power spectrum at EDoF ≈ 27 (Fig. 4, red curve). We see that this restricted-freedom fit follows the full-freedom fit almost exactly, but without the additional wiggles which resulted from noise in the data. For the EE power spectrum, the full-freedom fit has EDoF ≈ 189 (Fig. 5, blue curve). This fit is quite wiggly at high multipoles, as can be expected from high noise levels in data there. The Planck proposal [8] also expects the EE power spectrum to have high noise at high multipoles, and hence unable to resolve peaks beyond $l \lesssim 1000$. Therefore, we have excluded the EE fit beyond $l = 2000$ (Fig. 5) from further analysis. The restricted-freedom smoother version

of the EE fit corresponds to EDoF $= 23.79$ (Fig. 5, red curve). We see again that the smooth fit follows the full-freedom fit, but averages over the wiggles. For the simulated TE data, we obtain the full-freedom fit at EDoF ≈ 95 (Fig. 6, blue curve). The peaks are resolved well up to $l \approx 2000$. At higher multipoles, we see a wiggly fit with false peaks due to high levels of noise. A smoother restricted-freedom fit is obtained at EDoF $= 40$ (Fig. 6, red curve). Despite this additional smoothing, because of the high noise at high multipoles, we see two false bumps at $l = 1920, 2070$, and a false peak at $l = 2304$.

APPENDIX B: PROBING THE CONFIDENCE SETS

A high-dimensional confidence set (for a prespecified level of confidence, typically, 95%) around a fit is the prime inferential object for this nonparametric methodology [9]. By probing the confidence set, we can determine uncertainties on specific feature of a fit, validate different cosmological models against the data, etc. For finding uncertainties on the location and height of a peak or a dips (Figs. 8, 9, and 10) we use the same methodology as was used in [9]: We uniformly sample the confidence set of each smooth fit, and record the peaks and dips of acceptable spectra thus sampled. The most extreme variations in the height or location define a confidence interval on the respective estimates (represented in the figures as boxes). Compared to [9], we change a few criteria for accepting such sampled spectra to account for the greater angular resolution of the Planck data: For the TT and TE fits, we select spectra with exactly 8 and 10 peaks respectively. In the case of the TE fit, this criterion includes a peak at low multipoles ($l \leq 50$) and a false peak at $l = 2304$. For the EE fit, we sample spectra with 8 peaks for $l \leq 2000$, together with the additional condition that there must be a peak at low multipoles ($l \leq 50$). The last peak in the EE restricted-freedom fit consists of two tiny but close bumps due to noise in the data. Such fine structure is not expected here on theoretical grounds. Therefore, in case we find a sampled spectrum with two tiny peaks around the location of the eighth peak in the fit, but with separation less than 10 multipoles, we consider them as a single peak and record their average location and height.

[1] W. Hu, Lect. Notes Phys. **470,** 207 (1996).

[2] G. F. Smoot, Rev. Mod. Phys. **79,** 1349 (2007).

[3] M. J. Rees, Astrophys. J., Lett. **153,** L1 (1968).

[4] W. Hu and M. White, New Astron. **2,** 323 (1997).

[5] M. Zaldarriaga, in *Measuring and Modeling the Universe*, Volume Two: Carnegie Observatories Astrophysics Series,

edited by Wendy Freeman (Cambridge University Press, Cambridge, England, 2004), p. 309.

[6] M. Zaldarriaga, D. N. Spergel, and U. Seljak, Astrophys. J. **488,** 1 (1997).

[7] D. J. Eisenstein, W. Hu, and M. Tegmark, Astrophys. J. **518,** 2 (1999).

[8] Planck Collaboration, arXiv:astro-ph/0604069.

[9] A. Aghamousa, M. Arjunwadkar, and T. Souradeep, Astrophys. J. **745**, 114 (2012).

[10] R. Beran, J. Am. Stat. Assoc. **95**, 155 (2000).

[11] C. R. Genovese, C. J. Miller, R. C. Nichol, M. Arjunwadkar, and L. Wasserman, Stat. Sci. **19**, 308 (2004).

[12] D. Larson *et al.*, Astrophys. J. Suppl. Ser. **192**, 16 (2011).

[13] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998), 3rd ed.

[14] A. Challinor, Lect. Notes Phys. **653**, 71 (2004).

[15] A. de Oliveira-Costa, *Astronomical Polarimetry: Current Status and Future Directions*, edited by A. Adamson, C. Aspin, C. Davis, and T. Fujiyoshi, Astronomical Society of the Pacific Conference Series Vol. 343 (ASP, San Francisco, 2005), p. 485.

[16] L. Knox, Phys. Rev. D **52**, 4307 (1995).

[17] D. Scott and G. Smoot (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[18] J. L. Sievers *et al.*, Astrophys. J. **660**, 976 (2007).

[19] M. Zaldarriaga, Phys. Rev. D **55**, 1822 (1997).

[20] M. Kaplinghat, M. Chu, Z. Haiman, G. P. Holder, L. Knox, and C. Skordis, Astrophys. J. **583**, 24 (2003).

[21] G. P. Holder, Z. Haiman, M. Kaplinghat, and L. Knox, Astrophys. J. **595**, 13 (2003).

[22] W. Hu and G. P. Holder, Phys. Rev. D **68**, 023001 (2003).

[23] L. Perotto, J. Lesgourgues, S. Hannestad, H. Tu, and Y. Y. Y. Wong, J. Cosmol. Astropart. Phys. 10 (**2006**) 013.

[24] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).

[25] W. Hu, M. Fukugita, M. Zaldarriaga, and M. Tegmark, Astrophys. J. **549**, 669 (2001).

[26] M. Doran and M. Lilley, Mon. Not. R. Astron. Soc. **330**, 965 (2002).

[27] L. Page *et al.*, Astrophys. J. Suppl. Ser. **148**, 233 (2003).