OPTIMAL BINNING OF THE PRIMORDIAL POWER SPECTRUM

Paniez Paykari and Andrew H. Jaffe
Imperial College London, Astrophysics Group, Blackett Laboratory, Imperial College, London SW7 2AZ, UK; p.paykari06@ic.ac.uk, a.jaffe@ic.ac.uk
Received 2009 January 22; accepted 2010 January 6; published 2010 February 3

ABSTRACT

The primordial power spectrum describes the initial perturbations in the universe, which eventually grew into the large-scale structure we observe today, and thereby provides an indirect probe of inflation or other structure-formation mechanisms. In this paper, we will investigate the best scales the primordial power spectrum can be probed with in accordance with the knowledge about other cosmological parameters such as \( \Omega_b, \Omega_\gamma, \Omega_{\Lambda}, h, \) and \( \tau. \) The aim is to find the most informative way of measuring the primordial power spectrum at different length scales, using different types of surveys and the information they provide for the desired cosmological parameters. We will find the optimal binning of the primordial power spectrum for this purpose by making use of the Fisher matrix formalism. To investigate the correlations between the cosmological parameters, mentioned above, and a set of primordial power spectrum bins, we make use of principal component analysis and the Hermitian square root matrix formalism. To investigate the correlations between the cosmological parameters, normalised above, and a set of primordial power spectrum bins, we make use of principal component analysis and the Hermitian square root matrix formalism. To investigate the correlations between the cosmological parameters, normalised above, and a set of primordial power spectrum bins, we make use of principal component analysis and the Hermitian square root matrix formalism. To investigate the correlations between the cosmological parameters, normalised above, and a set of primordial power spectrum bins, we make use of principal component analysis and the Hermitian square root matrix formalism.

Key words: cosmic background radiation – cosmological parameters – early universe – large-scale structure of universe

Online-only material: color figures

1. INTRODUCTION

The primordial power spectrum encodes the physics of the early universe and its measurement is one of the key research areas in modern cosmology. There are many proposed models that try to describe the early universe, out of which the theory of inflation (Guth 1981; Linde 1982) is currently the most favored. The simplest models of inflation predict almost purely adiabatic primordial perturbations with a nearly scale-invariant power spectrum (PS; i.e., \( P(k) \propto k \))—the so-called Harrison–Zeldovich spectrum. Indeed this form of power spectrum fits the current observations very well. However, there are various models for the generation of the perturbations with deviations from the perfectly scale-invariant power spectrum. The simplest are the slow-roll inflationary models which describe the deviations through a minimal scale dependence ("running") of the power law index of the power spectrum. Other models generating deviations from scale-invariance include, for example, multiple scalar fields during inflation (Bridle et al. 2003; Contaldi et al. 2003), multiple inflation, and various phenomenological models resulting in features such as an exponential cutoff on large scales or a power law with superimposed harmonic wiggles (due to features in the inflation potential for example)—refer to Verde & Peiris (2008) and references therein. Determining the primordial power spectrum will therefore give us a better insight into the conditions of the early universe and help us choose from the many proposed models of the early universe.

The drawback, however, is that we cannot measure the primordial power spectrum directly, and our path to its measurement is through different experimental techniques such as cosmic microwave background (CMB) measurements and large-scale structure (LSS) surveys. The outcome of such surveys is a convolution of the primordial power spectrum and (the square of) a transfer function which, in turn, depends on the cosmological parameters (which we will collectively call \( \theta_i \)). Here, we list some examples.

1. For galaxy surveys, the observable PS is related to the primordial PS through the matter PS, \( P_{\delta}(k) \), as

\[
P_b(k) = b^2(k)P_\delta(k) \simeq b^2(k)2\pi^2kT^2(k)\Delta^2_\zeta(k),
\]

where \( \Delta^2_\zeta \) is the primordial PS and \( T(k) \) is the transfer function and \( b(k) \) is the bias.

2. For CMB surveys, the angular PS is

\[
C_\ell = 4\pi \int_0^\infty d \ln k \Delta^2_\zeta(k) \Delta^2_\zeta(k),
\]

where \( \ell \) is the angular wavenumber on the sky (corresponding to an angular scale on sky via \( \ell \sim 180^\circ/\theta \)) and \( \Delta_\ell(k) \) is the angular transfer function of the radiation anisotropies.

We define the primordial curvature power spectrum as \( \Delta^2_\zeta(k) = (k^3/2\pi)P_\zeta(k) = A(k/0.05)^n_{s-1} \), where \( A \) is the amplitude and \( n_s \) is the spectral index. The notation refers to the gauge-invariant curvature perturbation \( \zeta \) (Bardeen 1980). Other types of power spectra, such as the weak lensing and peculiar velocity power spectra, have similar forms; they depend on the cosmological parameters, through a transfer function, as well as the primordial PS. As these power spectra are jointly sensitive to the primordial PS and \( \theta_i \), there is an induced statistical degeneracy between them.

To recover the continuous primordial power spectrum, we need to deconvolve it from discrete data such as the CMB power spectrum \( C_\ell \) or the band powers of the LSS power spectrum. This deconvolution, along with our lack of knowledge of the cosmological parameters that determine the transfer function, induces a correlation between those parameters and those that determine the primordial power spectrum, which limits our ability to recover the primordial power spectrum completely. In some cases, such as measurements of the CMB, even a perfect survey cannot recover the primordial power spectrum completely (Hu & Okamoto 2004). However, the
transfer function of different techniques varies in the scales and parameters that they measure and this means the induced degeneracy between the parameters is different for different surveys. Therefore, one survey can help fill the gaps of others and together they can make significant improvements. Hence, combining surveys will enable us to improve, for example, the resolution of our measurements of the primordial power spectrum.

One method to recover the primordial power spectrum is to measure its amplitude in a series of bins. The more bins with high signal-to-noise ratio (S/N), the more accurately the power spectrum can be reconstructed. The aim of this work is to find an optimal binning for the primordial power spectrum based on the knowledge (or, better, lack of knowledge) of other cosmological parameters from some specific surveys, such as the Sloan Digital Sky Survey (SDSS; Gunn & Weinberg 1995) and Planck (Delabrouille et al. 1998). We then want to quantify the correlations between these carefully chosen bins. Therefore, the aim is to explore, as the data improve, what new information can be learnt about the primordial PS and what exactly needs to be improved to better constrain the primordial PS. The motivation is to test the assumptions about the initial conditions besides getting better constraints on parameters based on the same set of assumptions. Acknowledging the degeneracy between the cosmological parameters and the primordial PS, we want to investigate which scales the primordial PS can best be probed with in future experiments.

One common method for error estimation is to use a Fisher matrix analysis. The Fisher matrix is generally used to determine the sensitivity of a particular survey to a set of parameters and has been largely used for forecasting and optimization. It is defined as the ensemble average of the curvature of a function $F$ (i.e., it is the average of the curvature over many realizations of signal and noise):

$$F_{\alpha\beta} = \langle \beta \rangle = \frac{\partial^2 \ln L}{\partial \theta_\alpha \partial \theta_\beta},$$  

(3)

where $L$ is the likelihood function. Its inverse is an approximation of the covariance matrix of the parameters, by analogy with a Gaussian distribution in the parameter space.

The Cramer–Rao inequality1 states that the smallest error measured for $\theta_\alpha$ by any unbiased estimator (such as the maximum likelihood) is $1/\sqrt{F_{\alpha\alpha}}$ and $\sqrt{(F^{-1})_{\alpha\alpha}}$ for non-marginalized and marginalized1 one-sigma errors, respectively.

We further note, as in all uses of the Fisher matrix, that any results thus obtained must be taken with the caveat that these relations only map onto realistic error bars in the case of a Gaussian distribution, usually most appropriate in the limit of high S/N and/or relatively small scales, so that the conditions of the central limit theorem apply. As long as we do not find extremely degenerate parameter directions, we expect that our results will certainly be indicative of a full analysis, using simulations and techniques such as Bayesian Experimental Design (Trotta 2007; Parkinson et al. 2007; P. Paykari & A. H. Jaffe 2010, in preparation).

The Fisher matrix for CMB surveys is given by

$$F_{\ell\ell} = f_{sky} \frac{2\ell + 1}{2} \delta_{\ell\ell} \left[ C_\ell + w^{-1} e^{2\sigma^2} \right]^{-2},$$  

(4)

where $C_\ell$ is the angular PS, $w$ is the weight defined as $(\Delta \Omega \sigma^2)^{-1}$ with $\Delta \Omega$ being the real space pixel size and $\sigma^2$ being the noise per pixel, $e^{-\sigma^2}$ is the window function1 for a Gaussian beam (where $\sigma = \theta_{WHM}/\sqrt{8\ln 2}$), and $f_{sky}$ is the fraction of the sky observed. The factor $f_{sky}(2 + 1)$ gives the number of independent modes at a given wavenumber; the term proportional to $C_\ell$ is the sample (or cosmic) variance contribution, and the $w^{-1} e^{2\sigma^2}$ term is the noise contribution. Note that the diagonal form for the matrix implies diagonal (uncorrelated) errors on the $C_\ell$’s.

The $F_{\ell\ell}$ gives the errors on the $C_\ell$’s. Therefore, to find the errors on other parameters, we use the Jacobian

$$F_{\alpha\beta} = \sum_{\ell\ell} F_{\ell\ell} \frac{\partial C_\ell}{\partial \theta_\alpha} \frac{\partial C_\ell}{\partial \theta_\beta},$$  

(5)

where $\theta_\alpha$ and $\theta_\beta$ are different parameters.

For a volume-limited galaxy survey, the Fisher matrix (Tegmark 1997) is

$$F_{nn'} = \frac{k_v^2 \Delta k V}{(2\pi)^3 (P_n + 1/\bar{n})^2},$$  

(6)

where $V$ is the total volume of the survey, $\bar{n}$ is the number density of the survey $(N_{tot}/V)$, $P_n$ is the galaxy PS in each $k_v$ bin, and $\Delta k$ is the bin width. Similar to the CMB power spectrum case, $k_v^2 \Delta k V$ counts the number of independent modes, $P_n$ gives the sample variance, and $1/\bar{n}$ the noise variance due to Poisson counting errors. This, again, gives us the errors on the galaxy PS and we use the Jacobian to get the errors on other parameters

$$F_{\alpha\beta} = \sum_{nn'} F_{nn'} \frac{\partial P(k_n)}{\partial \theta_\alpha} \frac{\partial P(k_n)}{\partial \theta_\beta}.$$  

(7)

Fisher matrices for different surveys can easily be combined by simple summation: $F = F_{\text{Galaxy}} + F_{\text{CMB}}$; they are proportional to the log of the likelihood function and this is equivalent to the multiplication of independent likelihoods to combine them. Equivalently, we can think of them as the weights (inverse noise variance) of the experiments, which add for a Gaussian distribution. The nonzero correlation between the parameters in the covariance matrix makes interpreting the errors somewhat more difficult than the uncorrelated case. We will discuss various methods for decorrelating the power spectra and cosmological parameters.

2. METHOD

The aim is to investigate the primordial PS in a “non-parametric” way (we use quotations to remind the reader that “non-parametric” merely means that we use a

1 It should be noted that the Cramer–Rao inequality is a statement about the so-called “Frequentist” confidence intervals and is not strictly applicable to “Bayesian” errors.

2 Integration of the joint probability over other parameters.

3 This damps power on larger $\ell$’s; as we get closer to the resolution limit of the survey $C_\ell$’s start to correlate.

4 Note that this equation only applies to linear regime, as non-linearities impose non-Gaussianities.
very general model, potentially with a very large number of parameters). For this purpose, we define the primordial PS as a series of top-hat bins:

$$\Delta^2(k) = \sum_B w_B(k) Q_B,$$

where $Q_B$ is the power in each bin $B$ and $w_B = 1$ if $k \in B$ and 0 otherwise. The cosmological parameters under investigation are (and of the form) $\Omega_m$, $\Omega_b$, $\Omega_\Lambda$, $h$, and $\tau$. We will choose a geometrically flat CDM model with adiabatic perturbations and the WMAP5 (Dunkley et al. 2009) values for the parameters: $\Omega_m = 0.214 \pm 0.027$, $\Omega_b = 0.044 \pm 0.003$, $\Omega_\Lambda = 0.742 \pm 0.03$, $\tau = 0.087 \pm 0.017$, and $h = 0.719 \pm 0.0265$, where $H_0 = 100 h \text{ km}^{-1} \text{ Mpc}^{-1}$. $\Omega_b = 0$ was chosen, as massive neutrinos introduce some difficulties in the Fisher matrix analysis (Eisenstein et al. 1999) and therefore were ignored for now. The CMBfast software (Seljak & Zaldarriaga 1996) was used to generate the matter and CMB power spectra. The surveys chosen for this initial investigation are the projected results from the SDSS Bright Red Galaxies (BRGs) sample and the Planck Surveyor CMB Power Spectrum.6

3. APPLICATION TO SURVEYS

3.1. Galaxy Surveys—SDSS (BRG)

A galaxy PS is related to the matter PS via a parameter called bias—Equation (1). For the BRG sample of SDSS, this is assumed linear and scale-independent with the form $P_g(k) = b^2 P_m(k)$, where the bias is $b \geq 2.0$ (Mann et al. 1998; Scherrer & Weinberg 1998; Hütsi 2006; Seljak & Warren 2004). The survey specifications for the BRG sample are $\bar{n} = 10^5/V$ and $V = (1 \text{ h}^{-1} \text{ Gpc})^3$ (Gunn & Weinberg 1995).

For the $\theta_i$, the derivatives in the Jacobian were obtained numerically using the Taylor expansion

$$P(\theta_i) = P(\theta_0) + \left( \frac{\partial P}{\partial \theta_i} \right) \Delta(\theta_i).$$

The width and direction of the step are quite important here. A two-sided derivative was chosen so that the derivative is centered on the default value $\theta_0$, with a step size of $\Delta(\theta_i)/2$ on each side. This is accurate to first order in $\Delta(\theta_i)$ (a one-sided derivative would be at a slightly shifted place of $\theta_i + \Delta(\theta_i)/2$, and is only accurate to second order; Eisenstein et al. 1999). The width of the step should be small enough to give accurate results and yet big enough to avoid numerical difficulties. This was taken to be a 5% variation, therefore a 2.5% width on each side. Other studies have shown that this turns out to be the best step size, giving the most accurate results (Eisenstein et al. 1999).

For the primordial PS bins, the derivative is proportional to the matter transfer function

$$\frac{\partial P_g(k_a)}{\partial \Delta^2(k_a)} = 4 \times 2\pi^2 k T^2(\theta) \delta_{mn},$$

where $n$ and $n'$ refer to the bins. The $k$-range for SDSS is $0.006 \lesssim k/(h \text{ Mpc}^{-1}) \lesssim 0.1$. The minimum value for the wavenumber, $k_{min}$, is obtained from the largest scale of the survey—$(2\pi/V^{1/3})$. Its maximum value, $k_{max}$, is chosen to avoid non-linearities. Simulations of a very similar flat model (Meiksin et al. 1998) suggested a $k_{max}$ of $0.1 h \text{ Mpc}^{-1}$. This is also very close to the scale at which departures from linear theory was seen by Percival & White (2009).

The derivatives in the Jacobian need to be averaged into bins. In Section 4, we will explain the criteria for choosing the widths and locations of the bins.

3.2. CMB Surveys—Planck

One thing to note in the case of CMB power spectra is that the output of CMBfast is of the form $C_{\ell} = (\ell(\ell+1)/2\pi)C_{\ell}$, so the CMB Fisher matrix, Equation (4), is multiplied by this traditional factor $\ell(\ell+1)/2\pi$. The specifications for Planck High Frequency Instrument (HFI; we use the $v = 100 \text{ GHz}$ channel and conservatively assume that other frequencies are used for foreground cleaning) are $\theta_{FWHM} = 10.7 = 0.003115 \text{ radians}$, $\sigma_{pix} = 1.7 \times 10^{-6}$, $w^{-1} = 0.28 \times 10^{-15}$ (Delabrouille et al. 1998). The derivatives in the Jacobian were again obtained numerically by the Taylor expansion

$$C_{\ell}(\theta_i) = C_{\ell}(\theta_0) + \left( \frac{\partial C_{\ell}}{\partial \theta_i} \right) \Delta(\theta_i).$$

The same arguments as in the SDSS case applies for the width and direction of the step here. In the case of the primordial PS bins, the derivative becomes

$$\frac{\partial C_{\ell}}{\partial \Delta^2(k)} = 2(\ell+1) \int_{k_{min}}^{k_{max}} dk |\Delta(k)|^2.$$ 

This needs to be averaged into $k$ bins, as discussed in Section 4. The chosen $k$-range for Planck is $0.0001 \lesssim k/(h \text{ Mpc}^{-1}) \lesssim 0.1$, where $k_{min}$ was obtained from $k_{min} = \ell_{min}/d_A = 2/d_A$, where $d_A$ is the angular-diameter distance to the surface of last scattering obtained to be $\sim 14 \text{ Gpc}$ (Dunkley et al. 2009).

3.3. SDSS and Planck

As explained above, to combine data from different surveys, we can add the Fisher matrices obtained for each of them. We expect to see an improvement on the errors of both the bins and cosmological parameters. Equivalently, this will enable us to have narrower bins without sacrificing the S/N per bin.

4. OPTIMAL BINNING

As explained before, a set of primordial PS bins are part of our parameter space. In this section, we will explain how these bins are chosen. We construct the bins to have the same contribution to the Fisher matrix; that is, they each have the same S/N. We take the signal in each bin to be the amplitude of the primordial PS in that bin and the noise to be given by the inverse of the square root of the diagonal elements of the Fisher matrix. For this, we construct a signal vector, $S$, which contains the amplitude of the primordial PS for all the bins and the values of the cosmological parameters. We weight our Fisher matrix by this vector

$$F_{\alpha\beta} = S_\alpha S_\beta.$$

To obtain $\Delta(k)$, CMBfast needed to be altered to give the radiation transfer functions at all $\ell$’s. Then, for each $\ell$, this was interpolated in $k$. 

---

5 These are bright galaxies, which means the survey will be quite deep, with $z \sim 0.25-0.5$. Also, these trace the elliptical galaxies, which are thought to be better tracers of mass at this redshift range.

6 http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI%282005%291_V2.pdf
where there is no Einstein summation. This now gives us a \((S/N)^2\) matrix, where the square root of its diagonal elements are the \(S/N\) for the bins, and the weighted errors for \(\theta_i\). It is worth emphasizing that it is this \((S/N)^2\) Fisher matrix that will be diagonalized later on.

Our algorithm will result in more bins where the \(S/N\) is greater, sampling more finely where the signal is strongest (we will see this explicitly in our discussion of the CMB power spectrum in Section 6.1.2, which has considerable structure and therefore varying \(S/N\)). There are circumstances in which we might instead want to place bins by hand if we are looking for specific features (but of course we will always be limited by the \(S/N\) of our measurements).

For SDSS, we start with the maximum number of bins possible in our \(k\)-range, set by the usual properties of the Fourier transform. These imply that the scale of the survey not only determines \(k_{\text{min}}\), but also limits our resolution: \(k_{\text{min}} = (\Delta k)_{\text{min}} = (2\pi/V^{1/3})\); narrower bins would become highly correlated. Therefore, we set up a series of bins with this minimum width in our \(k\)-range. We then construct a Fisher matrix for this set of bins (and \(\theta_i\)) and weight it by the signal vector, \(S\), for this set. With this binning adopted, the \(S/N\) values range from 3.7 in the first bin to 35.1 in the last bin. Knowing that the bin widths chosen are the minimum possible and that increasing bin widths will increase the \(S/N\) value, we conclude that the bin with the maximum \(S/N\) cannot be changed and hence we make other bins wide enough to reach the \(S/N\) in this bin. To obtain this “optimal” binning, we add small bins until their \(S/N\) is within 15\% of the maximum \(S/N\):

\[
\frac{\text{Max}(S/N) - (S/N)_i}{\text{Max}(S/N)} = 0.15, \tag{14}
\]

where \(i\) refers to the bins. This finally gives us eight bins with their \(S/N\) ranging from 30–35.

For Planck, the bins are obtained so that their \(S/N\) matches that of the SDSS bins. The reason for applying this criteria to Planck is to allow for a fair comparison between the results from SDSS and Planck. This criteria gives us a total of 23 bins for Planck.

In the case of the combined Planck and SDSS, we require only that the \(S/N\) of the bins are equal within 50\%. This now gives us the optimal resolution of the primordial PS we can achieve from SDSS and Planck. We have a total of 48 bins with \(S/N\) being in the vicinity of \(\sim 20.0\) and, therefore, still comparable to the \(S/N\) values in the other cases.

An alternative, and perhaps more appropriate, way to determine the binning would be to take the marginalized errors as the noise. This would be obtained by inverting the Fisher matrix in each iteration to get the covariance matrix, which gives the marginalized covariances of the bins and \(\theta_i\). We would then take the sub-block of this covariance matrix that refers to the bins only, and invert it to get a marginalized Fisher matrix for the bins. We would then feed this Fisher matrix into Equation (13). However, this method could not be implemented because the SDSS Fisher matrix is not invertible; the SDSS Fisher matrix is not a positive definite matrix because it is asked to estimate too many parameters. There are a total of \(n\) data points \((n\) galaxy PS bins\) and we are asking these to predict \(n + m\) parameters \((n\) primordial PS bins and \(m\theta_i\)). Also, note that whichever of the methods presented uses the correlated errors as the noise. We now discuss the decorrelation of the parameters.

5. DECORRELATING THE PARAMETERS

5.1. Principal Component Analysis

Having obtained a set of bins, and therefore parameters we wish to estimate (or, in this case, forecast), we of course find that for these realistic experiments the parameter errors are often highly correlated (i.e., the Fisher matrix has significant off-diagonal components). To overcome this problem, we make use of principal component analysis (PCA) to obtain an orthogonal basis (onto which the original parameters will be projected); the covariance(Fisher) matrix is a symmetric \(n \times n\) matrix and therefore can be diagonalized using its eigenvectors. This has the form \(\mathbf{C} = \mathbf{E}^T \Lambda \mathbf{E}\), where \(\mathbf{C}\) is the covariance matrix, \(\mathbf{E}\) is an orthogonal matrix with the eigenvectors of \(\mathbf{C}\) as its rows and \(\Lambda\) is the diagonal matrix with the eigenvalues of \(\mathbf{C}\) as its diagonal elements\(^8\). This constructs a new set of variables \(\mathbf{X}\) that are orthogonal to each other and are a linear combination of the original parameters \(\mathbf{O}\), through the eigenvectors

\[
\mathbf{X} = \mathbf{E} \mathbf{O}. \tag{15}
\]

The \(X_i\) are called the principal components of the experiment and are ordered so that \(X_1\) has the smallest eigenvalue and \(X_2\) the largest. In this construction, the eigenvalues are the variances of the new parameters, so \(X_1\) and \(X_2\) are the best- and worst-measured components, respectively. The eigenvectors have been normalized so that \(\sum_j e_{ij}^2 = 1\), where \(e_{ij}\) are the elements of \(E_i\). We list some properties of PCA below.

1. The main point of PCA is to assess the degeneracies (correlations) among the parameters that are not resolved by the experiments, be they fundamental as from cosmic variance or due to the noise and coverage of the experiment. In our case, it will especially help us to see the correlation among the bins of the primordial PS, and between the bins and the cosmological parameters.

2. The eigenvalues obtained measure the performance of the experiment—a larger number of small eigenvalues means a better experiment. Another measure of the performance of the experiments is to see how they mix physically independent parameters such as, say, \(n_s\), the spectral index, and \(\Omega_b\). This sort of mixture may be improved by improving the experiment’s noise properties or increasing its area or volume.\(^9\)

All the above points may be summed up to conclude that in a perfect setting (if we could resolve cosmic variance and the geometrical degeneracy) we would expect a one-to-one relation between the old and the new parameters—the Fisher matrix would be diagonal. Each of the original parameters would contribute to one and only one of the new parameters, with zero contribution from the others.

Note that the principal components obtained are not strictly unique and depend on the form of the variables (e.g., whether we use \(\Omega_b\) or \(\log \Omega_b\)), as well as where they are evaluated.

\(^8\) It is common to construct the covariance matrix for PCA. However, the Fisher matrix can be used instead; the eigenvectors stay the same, but eigenvalues are reciprocals.

\(^9\) However, the so-called “geometrical degeneracy” (Zaldarriaga et al. 1997; Efstathiou & Bond 1999) cannot be improved by improving the experiments; two models with same primordial PS, the same matter content, and the same comoving distance to the surface of last scattering produce identical CMB PS.
5.2. Hermitian Square Root

Another approach to remove the correlations between the uncertainties is to use the Hermitian square root of the Fisher matrix as a linear transformation on the parameter space (Bond et al. 1998; Hamilton 1997a, 1997b). This transformation is obtained by

$$F^{1/2} = E^T A^{1/2} E,$$

(16)

where, like before, $E$ is the eigenvector matrix and $A$ is a diagonal matrix containing the eigenvalues. It has the property $F = F^{1/2} F^{1/2} = (F^{1/2})^T F^{1/2}$ and therefore the condition $(F^{-1/2}) F (F^{-1/2}) = (F^{-1/2})^T F (F^{-1/2}) = \text{diag}$ is satisfied. Unlike PCA, $F^{1/2}$ does not give us an orthogonal basis and instead can be thought of as giving "window functions" for the primordial PS resulting in uncorrelated parameters (in the Gaussian limit). We define a window matrix by

$$H_{nm} = \frac{(F^{1/2})_{nm}}{\sum_n (F^{1/2})_{nm}},$$

(17)

which satisfies the normalization condition $\sum_n H_{nm} = 1$. Hence, the windowed PS is defined as

$$\tilde{P}_m = \sum_n H_{nm} P(k_n),$$

(18)

where $P(k_n)$ is the original primordial PS. Note that this windowed PS is constructed for a visual presentation and understanding of the underlying correlations (indeed, it can be manifestly unphysical if, as we will see, the window function is negative). Again, in a perfect setting—a diagonal Fisher matrix—we would expect this windowed PS to be equal to the primordial PS (i.e., with each window function comprising a single bin).

We obtain this window matrix for the marginalized Fisher matrix of the bins and hence it can only be applied to the Fisher matrices of Planck and the combination of Planck and SDSS, which are invertible.

6. RESULTS

6.1. PCA

The principal components, $X_i$, obtained for SDSS, Planck, and their combination are shown as color-coded matrix plots; $X_i$ are shown from left to right with increasing errors (which is equal to $1/\Lambda_i$, as the eigenvalues are constructed for the Fisher matrix). The original parameters, $O_i$, are shown vertically starting with the bins on the bottom to $\theta_i$ on the top. For the bins, the vertical width of the box is an indication of the bin width. The last five principal components can be ignored as they are not measured—refer to Table 1 and the text for more details. At the bottom, we show the color plot indicating different levels of contribution to the principal components.

(A color version of this figure is available in the online journal.)

![Figure 1. Principal components of SDSS with no priors on $\theta_i$. $X_i$ are shown from left to right with increasing errors ($\pm 1/\sqrt{N}$). Original parameters, $O_i$, are shown vertically starting with the bins on the bottom to $\theta_i$ on the top. For the bins, the vertical width of the box is an indication of the bin width. The last five principal components can be ignored as they are not measured—refer to Table 1 and the text for more details. At the bottom, we show the color plot indicating different levels of contribution to the principal components.](image)

parameters (or eight different combinations of the parameters, i.e., $X_i$) can be measured.

The best-measured principal component, $X_1$, has contributions from the cosmological parameters ($\theta_i$) only, with $h$ dominating. The fact that there is more than one cosmological parameter contributing to this principal component means that SDSS can only measure a (linear) combination of them—evincing a degeneracy between these parameters. $X_2$ measures a combination of the bins and $\theta_i$, which suggests a degeneracy between the highlighted bins, $h$ and $\Omega$. Other principal components, $X_3\text{--}X_8$, measure the bins only, with no contribution from $\theta_i$ at all, and the correlation among the bins is between neighboring ones only. Intuitively, one expects more correlation between the bins and the cosmological parameters. The errors for the bins are related to the matter transfer function—Equation (10). Therefore, one might expect that a change in $\theta_i$ would induce a change in the matter transfer function and hence a correlation between bins and $\theta_i$. However, consider Figure 2 showing the individual derivatives that contribute to the Jacobian. The derivatives with respect to $\theta_i$ are of similar magnitudes (apart from $\tau$ which was multiplied by 200 for presentation). However, the derivative with respect to the primordial PS bins is rescaled by $10^{-8}$ to fit in the same range as the rest of the derivatives. This suggests that perhaps the changes in $\theta_i$ are not large enough in this setting to have a significant effect on the matter transfer function and therefore the correlation is not that significant to show effects in the PCA.

![Figure 2. Individual derivatives that contribute to the Jacobian. The derivatives with respect to $\theta_i$ are of similar magnitudes (apart from $\tau$ which was multiplied by 200 for presentation). However, the derivative with respect to the primordial PS bins is rescaled by $10^{-8}$ to fit in the same range as the rest of the derivatives. This suggests that perhaps the changes in $\theta_i$ are not large enough in this setting to have a significant effect on the matter transfer function and therefore the correlation is not that significant to show effects in the PCA.](image)
Note that the correlation between the bins shows the limits to what we can learn about the primordial PS. This correlation arises due to our lack of knowledge of the cosmological parameters. If we knew the parameters perfectly, we would have what is shown in Figure 3 (which is, in fact, the Fisher matrix itself). Generally, it seems that SDSS measures cosmological parameters better than the primordial PS and considering the primordial PS, measures small scales better than large.

We also investigated what improvements we would see given better—realistic—knowledge of the cosmological parameters. Hence, WMAP5 priors (Dunkley et al. 2009) were added to constrain the \( \theta_i \) in the Fisher matrix, by adding the inverse variance of each parameter to the Fisher matrix, i.e., ignoring the WMAP5 correlations. The result is shown in Figure 4. Note that the errors on the principal components have reduced and now all \( X_i \) can be measured well—Table 1. Some of the degeneracies between the cosmological parameters have been broken. For example, \( \Omega_b \) and \( \tau \) dominate completely in \( X_{10} \) and \( X_{12} \), respectively, with no contribution from any other parameter. This is expected as WMAP5 does a good job measuring these cosmological parameters. With respect to the primordial spectrum, these improved constraints on cosmological parameters have only helped to measure linear combinations of the bins better and have not been able to break the degeneracy between them.

6.1.2. Planck

For Planck there are a total of 23 bins and, with the five \( \theta_i \), we have 28 principal components, shown in Figure 5. They

|                  | SDSS      | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 |
|------------------|-----------|----|----|----|----|----|----|----|----|----|-----|-----|-----|
| No priors        | 0.0038    | 0.0196 | 0.0288 | 0.0331 | 0.0349 | 3E5 | NaN | NaN | ... | ... | ... | ... |
| No \( \theta_i \) | 0.0282    | 0.0299 | 0.0317 | 0.0340 | 0.0359 | ... | ... | ... | ... | ... | ... | ... |
| WMAP5 priors     | 0.0037    | 0.0189 | 0.0287 | 0.0330 | 0.0344 | 0.0423 | 0.1233 | 0.5871 | ... | ... | ... | ... |

|                  | Planck    | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 |
|------------------|-----------|----|----|----|----|----|----|----|----|----|-----|-----|-----|
| PCA-No priors    | 0.0004    | 0.0022 | 0.0094 | 0.0209 | 0.0300 | 0.0402 | 0.0516 | 0.0716 | 0.2576 | ... | ... | ... |
| PCA-No \( \theta_i \) | 0.0110    | 0.0149 | 0.0181 | 0.0224 | 0.0333 | 0.0422 | 0.0572 | ... | ... | ... | ... | ... |
| PCA-Margin.      | 0.0181    | 0.0210 | 0.0250 | 0.0297 | 0.0384 | 0.0514 | 0.0679 | ... | ... | ... | ... | ... |
| Hermitian Sqrt   | 0.0165    | 0.0184 | 0.0188 | 0.0190 | 0.0218 | 0.0496 | 0.1164 | ... | ... | ... | ... | ... |

|                  | Planck and SDSS | X1 | X2 | X3 | X4 | X5 | X6 | X10 | X11 | X12 | X20 | X21 | X22 |
|------------------|------------------|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|
| PCA-No priors    | 0.0004            | 0.0020 | 0.0073 | 0.0353 | 0.0469 | 0.0552 | 0.0643 | 0.1135 | 0.2068 | ... | ... | ... |
| PCA-No \( \theta_i \) | 0.0561            | 0.0568 | 0.0575 | 0.0624 | 0.0729 | 0.0891 | 0.1003 | ... | ... | ... | ... | ... |
| PCA-Margin.      | 0.0227            | 0.0288 | 0.0304 | 0.0408 | 0.0514 | 0.0578 | 0.0734 | ... | ... | ... | ... | ... |
| Hermitian Sqrt   | 0.0233            | 0.0234 | 0.0235 | 0.0243 | 0.0384 | 0.0736 | 0.5616 | ... | ... | ... | ... | ... |

Table 1: Errors for Different Sets for SDSS, Planck, and Combination of Planck and SDSS
are all measured well and better than SDSS—Table 1. The reflection of the acoustic peaks of $C_\ell$'s on the bin sizes can clearly be seen; those corresponding to the peaks are measured with better resolution (see Figure 6 to see a pictorial version of the contributions to the Jacobian. The summation over $\ell$ in the Fisher matrix gives the oscillatory feature seen in $k$ space.\)

Just like SDSS, Planck measures the cosmological parameters better than the primordial PS and overall, gives smaller errors and smaller correlations between them than SDSS. This is not surprising as we already know Planck does a good job measuring the cosmological parameters; it measures $\Omega_b$ and $h$ very well, with only slight correlation with other cosmological parameters.

The rest of principal components contain the highly correlated bins only, with no single large contribution from any of them. Intuitively, one might expect the correlation to be between neighboring bins only. The reason for the longer range correlation lies in the form of the radiation transfer function; for each $\ell$, this transfer function spans a $k$-range around $k=\ell/d_A$, where $d_A$ is the angular-diameter distance to the last-scattering surface. This is due to the projection of a three-dimensional universe onto a two-dimensional sphere around us. Equation (12) shows the contribution to the Jacobian for the Fisher matrix analysis. For each $\ell$, this derivative integrates the radiation transfer function over the $k$-range of the bins. This would be reflected as correlation between neighboring bins. However, remember that in the Fisher matrix analysis the $\ell$'s are summed over (Equation (5)) and this induces correlation between all bins; Figure 6 shows a pictorial version of Equation (12), weighted by the primordial PS. Note how each $\ell$ spans a range of $k$. The summation over all $\ell$'s means that, for example, the bin with $\ell = 400$ dominating

has contributions from all $\ell$’s from 100 to 500, with each spanning a different range of $k$. This induces correlation between bins of all scales.

This correlation between small and large scales might even be worse when there is a degeneracy between the measured cosmological parameters. For example, consider an experiment with which we attempt to measure two parameters, such as the spectral index $n_s$ and $\Omega_b$, where $n_s$ is dominant on large scales and $\Omega_b$ is dominant on small scales—Figure 7. The degeneracy between these parameters could induce a degeneracy between large and small scale bins.

Figure 8 shows the principal components for the bins with no $\theta_i$—i.e., assuming cosmological parameters are known perfectly. Since Planck measures the cosmological parameters very well, we might expect little change. Indeed, not much is changed. Note, however, that the smallest error for this set is still larger than the smallest error for the set including $\theta_i$. This is because the $\theta_i$ are individually measured better than the primordial PS bins and hence they reduce the errors; instead, comparing the largest errors of both sets shows the improvements. Despite the smaller errors for this set, the correlation between the bins is not significantly improved.

We now consider the correlation between the bins for the marginalized Fisher matrix (that is, marginalized over the other cosmological parameters, $\theta_i$). This is obtained by inverting the parent Fisher matrix to get a covariance matrix, giving the marginalized errors for all the parameters (in the Gaussian limit). We can further take the sub-block of this matrix which holds the errors for the bins and invert it to obtain a marginalized Fisher matrix for the bins alone. The principal components for this Fisher matrix are shown in Figure 9. The first thing to note is that bins contribute more significantly to some of the principal
components. In particular, there are some mid-scale bins which seem to be measured well. For example, consider $X_{19} - X_{22}$; the marginalization has uncorrelated some mid-scale bins from the rest of the bins.

Another interesting result is that very large and very small scales never really dominate in the principal components with large errors. They only contribute to them at levels of $\lesssim 0.01$. Recall that $X_i$ with large errors are the most highly correlated and therefore the fact that mid-scale bins do not contribute to
these principal components means that they are measured quite well.

In conclusion, Planck will largely decorrelate the primordial PS from the $\theta_i$ (and therefore the transfer function) but cannot completely uncorrelate the bins themselves.

6.1.3. Planck and SDSS

The results are shown in Figure 10. Combining surveys has clearly helped to improve the resolution of the primordial PS. The data now support a total of 48 bins in the same $k$-range. Again the cosmological parameters are measured better than the primordial PS and there is also smaller correlation between the cosmological parameters compared to the previous cases. There is also smaller correlation between the bins themselves. Features of both SDSS and Planck can clearly be seen here. For example, acoustic oscillations in the $C_\ell$'s still influence the bin sizes and resolution of the primordial PS. It also seems like small scales are measured better than large scales, which is a feature seen in the SDSS case.

Figure 11 shows the results for the marginalized Fisher matrix of the bins for SDSS and Planck combined. Compare it to Figure 10; not much change can be seen.

6.2. Hermitian Square Root of Fisher Matrix

Figure 12 shows the window functions for Planck derived from the Hermitian square root decorrelation. Note that only the magnitude of the components of $H_m$’s are important and not their sign. However, it is worth mentioning that for the non-marginalized Fisher matrix (both for Planck and its combination with SDSS), the window functions have only positive values. Therefore, the lack of knowledge of the cosmological parameters (and the induced correlation between the bins) induces non-physical negative values into the window functions. The window functions, $H_m$’s, are plotted in the order of increasing errors, so that $H_1$ is the best- and $H_{23}$ is the worst-measured vector. This shows that correlation is only between neighboring bins and, that bins on small scales are measured better than the ones on large scales (just like the PCA case).

(A color version of this figure is available in the online journal.)
lowest and highest errors, respectively. Planck cannot decorrelate the bins completely and some correlations between neighboring bins can be seen. In addition, large scales (i.e., bins in the range of $k \sim 0.02$–0.04 $h$ Mpc$^{-1}$) have a large contribution to the $H_m$, compared to the bins on smaller scales. These window functions clearly show the influence of cosmic variance. Compare this to Figure 13, where we diagonalized the marginalized Fisher matrix through its eigenvectors (this is exactly Figure 9 plotted in this form for easier comparison). In the PCA case, the correlations seem not to be only between neighboring bins, but between bins of all scales, which is not seen in this case. Also, the compactness seen here (i.e., more of a traditional window-function feature) is not seen in the PCA case; there is no particular scale that contributes significantly to the principal components.

Figure 14 shows the windowed PS for Planck. It is plotted so that each $\tilde{P}_m$ is placed at the $k_n$ from which it receives the largest contribution. The vertical error bars shown are $\Delta^2(k_n)(HF^{-1}H^T)$, where $\Delta^2(k_n)$ are the amplitude of the primordial PS in the bins and $(HF^{-1}H^T)$ are the errors propagated through the $H_m$ distribution. The horizontal error bars are the half-width at half-maximum in each direction of the main peak of each $H_m$. The original primordial PS is plotted for comparison. Remember that $\tilde{P}_m$ is not a physical PS per se. However, the observed differences from the original PS arise due to the induced correlations between the bins. In a perfect setting, where there are no correlations between bins, we expect $\tilde{P}_m = \Delta^2(k)$. Note that the main feature of this plot is that vertical errors, unlike those for the original primordial PS, are not correlated. The correlation between the errors has been transferred to overlaps between the window functions—as shown in Figure 12.

Figures 15 and 16 show the same set of results for combination of Planck and SDSS. Again, large scales contribute to $H_m$ with the largest errors. There is less correlation between neighboring bins, so that each $\tilde{P}_m$ is placed at the $k_n$ from which it receives the largest contribution. The vertical error bars shown are $\Delta^2(k_n)(HF^{-1}H^T)$, where $\Delta^2(k_n)$ are the amplitude of the primordial PS in the bins and $(HF^{-1}H^T)$ are the errors propagated through the $H_m$ distribution. The horizontal error bars are the half-width at half-maximum in each direction of the main peak of each $H_m$. The original primordial PS is plotted for comparison. Remember that $\tilde{P}_m$ is not a physical PS per se. However, the observed differences from the original PS arise due to the induced correlations between the bins. In a perfect setting, where there are no correlations between bins, we expect $\tilde{P}_m = \Delta^2(k)$. Note that the main feature of this plot is that vertical errors, unlike those for the original primordial PS, are not correlated. The correlation between the errors has been transferred to overlaps between the window functions—as shown in Figure 12.
Planck & SDSS

Figure 16. Windowed PS obtained from Planck and SDSS. There is a better match between the two PS on small scales compared to Planck on its own. However, the differences on large scales still remain.

(A color version of this figure is available in the online journal.)

Figure 17. Showing Figure 11—plotted for easier comparison with Figure 15. (A color version of this figure is available in the online journal.)

bins compared to the Planck case. Also, note that bins in this case are narrower and therefore correlation between neighboring bins still means correlation between a narrower range of \( k \). Compare Figure 15 to Figure 17 (same as Figure 11). Again, the PCA case has wider effective bins, more than observed for Planck on its own. Figure 15 indicates that bins in the vicinity of \( k \sim 0.02-0.025 \, h \text{Mpc}^{-1} \) contribute very strongly to the \( H_m \) compared to other bins, in particular the last window function, \( H_{\text{MS}} \). This effect is shown in \( \hat{P}_m \), with \( \hat{P}_{11} \) having a very large amplitude—Figure 16.

7. CONCLUSIONS

The primordial PS holds precious information about the physics of the early universe and constraining it is one of the key goals of modern cosmology. However, the induced degeneracy between the cosmological parameters determining the matter/radiation transfer functions and the primordial PS limit our ability to recover the primordial PS, even from a perfect survey, especially in the case of CMB measurements (Hu & Okamoto 2004). Different surveys probe different scales with different accuracies and might not be able to constrain the primordial PS to a desired resolution on their own. In combination, however, they make significant improvements. In this paper, we have investigated these limits and improvements for CMB and LSS surveys, exemplified by Planck and the SDSS BRG sample. For this purpose, we have assumed a non-parametric form for the primordial PS and have constructed a parameter space containing a set of carefully chosen bins of the primordial PS along with a set of cosmological parameters. We constructed a Fisher matrix for this parameter space for the two different surveys separately and combined. By diagonalizing these Fisher matrices, via two different methods of eigenvector decomposition (PCA) and the Hermitian square root, we have investigated the induced correlation between the primordial PS bins and the cosmological parameters.

In the PCA case, we conclude that SDSS and Planck together measure the cosmological parameters to a better extent, and even break the degeneracy between them. More importantly for our purposes, they can increase the obtainable resolution of the primordial PS by a factor of 2 and can also condense the correlation between bins to be only among neighboring ones. On the whole, these experiments combined will constrain small scales better than large scales.

By the use of the Hermitian square root of the Fisher matrix, we managed to divert the correlation among the marginalized errors of the bins to the correlation between the bins themselves. In this case, the combination of SDSS and Planck helped to decrease the level of correlation between neighboring bins, but also, because it has helped to increase the resolution of the bins, correlation between neighboring bins means correlation between a smaller range of \( k \).

Clearly combining the two surveys will constrain the primordial PS better than current measurements, and better than each experiment on its own. Obviously, further surveys of other phenomena related to the evolution of fluctuations, such as Ly-\( \alpha \) (e.g., SDSS Ly\( \alpha \) PS), weak lensing (e.g., Euclid), peculiar velocities (e.g., Cluster Imaging Experiment, CIX), etc. will help further measure the primordial PS.

We especially thank Carlo Contaldi for his great suggestions. We thank Dmitri Novikov, George Bendo, Daniel Mortlock, and Gavin Nicholson for their great help. This work was supported by STFC.

REFERENCES

Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882
Bond, J. R., Jaffe, A. H., & Knox, L. 1998, Phys. Rev. D, 57, 2117
Bridle, S. L., Lewis, A. M., Weller, J., & Efstathiou, G. 2003, MNRAS, 342, L72
Contaldi, C. R., Peloso, M., Kofman, L., & Linde, A. 2003, J. Cosmol. Astropart. Phys., 7, 2
Delabrouille, J., Puget, J., Gispert, R., & Lamarre, J. 1998, arXiv: astro-ph/9810478
Dunkley, J., et al. 2009, ApJS, 180, 306
Efstathiou, G., & Bond, J. R. 1999, MNRAS, 304, 75
Eisenstein, D. J., Hu, W., & Tegmark, M. 1999, ApJ, 518, 2
Gunn, J., & Weinberg, D. 1995, in Proc. 35th Herstmonceux Conf., Wide Field Spectroscopy and the Distant Universe, ed. S. J. Maddox & A. Aragon-Salamanca (Singapore: World Scientific), 3
Guth, A. H. 1981, Phys. Rev. D, 23, 347
Hamilton, A. J. S. 1997a, MNRAS, 289, 285
Hamilton, A. J. S. 1997b, MNRAS, 289, 295
Hu, W., & Okamoto, T. 2004, Phys. Rev. D, 69, 043004
Hutsi, G. 2006, A&A, 449, 891
Linde, A. D. 1982, Phys. Lett. B, 108, 389
Mann, R. G., Peacock, J. A., & Heavens, A. F. 1998, MNRAS, 293, 209
Meiksin, A., White, M., & Peacock, J. A. 1998, in Abstracts of the 19th Texas Symp. on Relativistic Astrophysics and Cosmology, held in Paris, France, 1998 December 14–18, meeting abstract, ed. J. Paul, T. Montmerle, & E. Aubourg (Gif-sur-Yvette: CEA Saclay)
Parkinson, D., Blake, C., Kunz, M., Bassett, B. A., Nichol, R. C., & Glazebrook, K. 2007, MNRAS, 377, 185
Percival, W. J., & White, M. 2009, MNRAS, 393, 297
Scherrer, R. J., & Weinberg, D. H. 1998, ApJ, 504, 607
Seljak, U., & Warren, M. S. 2004, MNRAS, 355, 129
Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437
Tegmark, M. 1997, Phys. Rev. Lett., 79, 3806
Trotta, R. 2007, MNRAS, 378, 819
Verde, L., & Peiris, H. 2008, J. Cosmol. Astropart. Phys., 7, 9
Zaldarriaga, M., Spergel, D. N., & Seljak, U. 1997, ApJ, 488, 1