Performance of XFast\text{er} likelihood in real CMB experiments

G. Rocha,\textsuperscript{1,2} C. R. Contaldi,\textsuperscript{3} L. P. L. Colombo,\textsuperscript{4} J. R. Bond,\textsuperscript{5} K. M. Górski,\textsuperscript{1,6} and C. R. Lawrence\textsuperscript{1}

\textsuperscript{1}Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena CA 91109, U. S. A.
\textsuperscript{2}Department of Physics, California Institute of Technology, Pasadena, 91125, U. S. A.
\textsuperscript{3}Department of Physics, Imperial College, University of London, South Kensington Campus, London SW7 2AZ, U. K.
\textsuperscript{4}Department of Physics and Astronomy, University of Southern California, University Park Campus, Los Angeles CA 90089, U. S. A.
\textsuperscript{5}The Canadian Institute for Theoretical Astrophysics, CITA, University of Toronto, 60 St. George Street, Toronto, Ontario, M5S 3H8, Canada
\textsuperscript{6}Warsaw University Observatory, Aleje Ujazdowskie 4, 00478 Warszawa, Poland

31 August 2010

ABSTRACT

We assess the strengths and weaknesses of several likelihood formalisms, including the \textit{XFaster} likelihood. We compare the performance of the \textit{XFaster} likelihood to that of the Offset Lognormal Bandpower likelihood on simulated data for the Planck satellite. Parameters estimated with these two likelihoods are in good agreement. The advantages of the \textit{XFaster} likelihood can therefore be realized without compromising performance.

Key words: Cosmology: observations – methods: data analysis – cosmic microwave background

1 INTRODUCTION

The temperature and polarization anisotropies of the Cosmic Microwave Background (CMB) contain a wealth of cosmological information. In extracting this information from measurements of the CMB, the likelihood function \( L(C_\ell) = P(\text{data}|C_\ell) \), where \( C_\ell \) is the theoretical spectrum for some cosmological model, plays an important role. For Gaussian fluctuations, \( L(C_\ell) \) is given by a Multivariate Gaussian of the observed data (section 2). For low resolution data (\( \ell \lesssim 10^2 \), defined more precisely later) in the usual spherical harmonic expansion, it is computationally feasible to evaluate this directly. For high resolution data, it is not, and computationally tractable approximations are required. Fortunately, the large number of independent samples of the universe at high resolution, coupled with a much more nearly diagonal covariance matrix, mean that approximations exist that are accurate as well as fast. The subject of this paper is the performance of one such high \( \ell \) method, the “\textit{XFaster} likelihood.”

A successful high \( \ell \) method must account correctly for correlations induced in the angular power spectra by partial sky coverage, non-uniform noise, and non-zero beamwidths, as well as the temperature and polarization cross power. A number of approaches for high-\( \ell \) have been proposed. For temperature alone, these include the Gaussian, Offset Lognormal, Equal Variance (Bond, Jaffe, & Knox 2000), WMAP (Verde et al. 2003), and SCR (Smith, Challinor, & Rocha 2006) likelihood approximations. For temperature and polarization together, these include the Offset Lognormal Bandpower, Hamimeche and Lewis (Hamimeche & Lewis 2008), and \textit{XFaster} (Rocha et al. 2009) likelihoods.

This paper is organized as follows. In Section 2 we describe the Multivariate Gaussian likelihood both in pixel and harmonic space, give an overview of existing high \( \ell \) approximations (Section 2.2) and their limitations. An account of their performances applied to simulated Planck data has been given in (Rocha et al. 2009). We further test the performance of \textit{XFaster} likelihood implemented in a new, modified version of the publicly available software \textit{CosmoMC} code\textsuperscript{1} (Lewis and Bridle (2002)) by applying it to estimation of cosmological parameters, and compare it to the Offset Lognormal Bandpower likelihood. We further test our approach to tackle the asymmetry of the beams by comparing parameters estimated for Planck simulated data convolved with a symmetric and an asymmetric beam using both likelihood approaches.

\textsuperscript{1}http://cosmologist.info/cosmomc/
2 LIKELIHOOD FOR A GAUSSIAN SKY

Pixel temperature fluctuations $T(\hat{n})$ (Stokes $I$) on the celestial sphere can be expanded in terms of spherical harmonics $Y_{\ell m}$ as

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}),$$

with coefficients $a_{\ell m}$. Polarization fluctuations (Stokes $Q$ and $U$) can be expanded in spin-2 spherical harmonics, $2 Y_{\ell m}$, with $E$ and $B$ (gradient- or curl-type) polarization coefficients

$$(Q \pm iU)(\hat{n}) = \sum_{\ell m} (a_{\ell m}^E \pm ia_{\ell m}^B) Y_{\ell m}(\hat{n}).$$

If the CMB is a Gaussian isotropic field, then the probability of a measurement of the sky given a model is described by a Multivariate Gaussian of the observed data:

$$L(d|p) = \frac{1}{2\pi^{N/2}|C|^{1/2}} \exp \left(-\frac{1}{2}d^{\top}C^{-1}d\right)$$

where $C$ is the covariance of the data $d$, and $p$ is the set of model parameters. $C(p) = S(p) + N$, where $S$ is the sky signal and $N$ is the noise. Since measurements of the sky are pixelated, the above likelihood is often called the "pixel-based likelihood" when estimated in the pixel domain. It can be evaluated directly if the number of pixels is not too large, but for a full-sky experiment such as Planck with 5' pixels it is impossible with current computers.

2.1 Exact likelihood in harmonic space

If the CMB is Gaussian, its statistical properties are represented fully by the underlying power spectrum $C_{\ell}$. The multipole harmonic coefficients, $a_{\ell m}^X$ (where $X$ is $T, E$, or $B$), on different scales are independent of one another, and we can write

$$\left\langle (a_{\ell m}^X)^* a_{\ell' m'}^X \right\rangle = \delta_{\ell \ell'} \delta_{m m'} C_{\ell}^{XX'}.$$

The $a_{\ell m}^X$ are complex. Under the assumption of Gaussianity, their real and imaginary parts are independent and Gaussian distributed, hence their phases are random. Since the CMB is real, they must satisfy $a_{\ell m}^* = (-1)^m a_{\ell -m}^*$. This means that $a_{\ell 0}$ is real.

For $m = 0$ we have

$$P(a_{\ell 0}|C_{\ell}) da_{\ell 0} = \frac{1}{(2\pi)^{3/2}|C_{\ell}|^{1/2}} \exp \left(-\frac{1}{2}a_{\ell 0}^{\top}C_{\ell}^{-1}a_{\ell 0}\right).$$

For $m \neq 0$, we have for the real part of the $a_{\ell m}^X$

$$P(\Re\{a_{\ell m}\}|C_{\ell}) da\Re\{a_{\ell m}\} = \frac{1}{(2\pi)^{3/2}|C_{\ell}|^{1/2}} \exp \left(-\frac{1}{2}\Re\{a_{\ell m}\}C_{\ell}^{-1}\Re\{a_{\ell m}\}\right),$$

where

$$a_{\ell m} = \begin{pmatrix} a_{\ell m}^T \\ a_{\ell m}^E \\ a_{\ell m}^B \end{pmatrix}$$

and

$$C_{\ell} = \begin{pmatrix} C_{\ell}^{TT} & C_{\ell}^{TE} & 0 \\ C_{\ell}^{ET} & C_{\ell}^{EE} & 0 \\ 0 & 0 & C_{\ell}^{BB} \end{pmatrix},$$

Similarly for the imaginary part.

Combining together all the values of $m$, we find, for a particular $\ell$:

$$P(\hat{C}_{\ell}|C_{\ell}) \propto |\hat{C}_{\ell}|^{\frac{2\ell + 1}{2}} |C_{\ell}|^{-\frac{2\ell + 1}{2}} \exp \left\{ -\frac{2\ell + 1}{2} \text{Tr} \left( \hat{C}_{\ell} C_{\ell}^{-1} \right) \right\}.$$

where

$$\hat{C}_{\ell}^{XX'} = \sum_{m=-\ell}^{\ell} (a_{\ell m}^X)^* a_{\ell m}^{X'} \frac{2\ell + 1}{2},$$

and the normalisation is independent of $C_{\ell}$ and $\hat{C}_{\ell}$. In data analysis, the measured power spectrum $\hat{C}_{\ell}$ is a fixed quantity, hence the dependence of the likelihood on this value is generally dropped. In this case, up to a constant, we can write the log-likelihood as

$$-2 \ln P(\hat{C}_{\ell}|C_{\ell}) = (2\ell + 1) \left( \ln |C_{\ell}| + \text{Tr} \left( \hat{C}_{\ell} C_{\ell}^{-1} \right) \right),$$

i.e., the inverse Wishart distribution. If we consider only one measurement, e.g., one $T$-mode or $B$-mode, we can write the likelihood function
\[ -2 \ln P(\hat{C}/C) = (2\ell + 1) \left( \ln \left( \frac{C}{\hat{C}} \right) + \frac{\hat{C}}{C} \right), \]  

i.e., the inverse Gamma distribution, where \( C \) is the theoretical value of \( C^\ell \) (or \( C^B \)) and \( \hat{C} \) is the measured value.

We can write an exact expression for the likelihood function for our measured power spectrum \( \hat{C} \) given the true underlying power spectrum \( C \), which is a function of cosmological parameters. Since this likelihood is usually considered in the context of data analysis, it is common to regard the measured \( \hat{C} \) as fixed quantities, and to write the likelihood as

\[
\ln P(\hat{C}/C) = \sum_{\ell} \left( \frac{2\ell + 1}{2} \right) \left( \ln \left( \frac{C}{\hat{C}} \right) + \frac{\hat{C}}{C} \right) + \text{constant},
\]

where the constant depends on \( \hat{C} \). For fixed \( \hat{C} \), this function peaks at \( C = \hat{C} \). However if we wish to consider the likelihood as a function of \( \hat{C} \) then it is necessary to include the \( \hat{C}-\)dependence of the likelihood, in which case it should be written as

\[
\ln P(\hat{C}/C) = \sum_{\ell} \left( \frac{2\ell - 1}{2} \right) \ln \hat{C} - \frac{2\ell + 1}{2} \left( \ln C + \frac{\hat{C}}{C} \right) + \text{constant}.
\]

For a fixed underlying power spectrum \( C \), this function peaks at \( \hat{C} = \frac{2\ell + 1}{2\ell + 3} C \).

This is adequate for a full-sky experiment with an infinitely narrow beam and no instrumental noise. For a partial or “cut” sky it is necessary to account for the correlations between the \( \hat{C} \) that are introduced. In addition, real experiments always have non-uniform noise, and must estimate bandpowers \( \hat{C}_B \) rather than individual \( \hat{C}_\ell \). We need to find an appropriate likelihood function that includes the correct correlations and accounts properly for noise.

2.2 **Approximating the likelihood at high-\( \ell \)**

Although the signal and the noise can be assumed Gaussian, the distribution of the band powers is non-Gaussian. This is the so-called 'cosmic bias'—the distribution is skewed towards higher \( C_\ell \) values. This effect is most noticeable at low \( \ell \) where cosmic variance of the signal dominates the error bars. As we will see, the full likelihood includes the cosmic bias whereas the Gaussian approximation of the Fisher matrix does not. The Joint likelihood for temperature and polarization carries an extra complication in that one has to find an approximation that properly accounts for the temperature and polarization cross power. We will start by considering the current approximations derived for temperature alone, followed by an account on existing attempts to extend it to a joint likelihood for temperature and polarization.

2.2.1 **Temperature only**

**Gaussian Likelihood**—The first level of approximation is to use a likelihood that is Gaussian in the \( \hat{C} \) (\cite{Bond:1996}, i.e.,

\[
P(\hat{C}/C) \propto \exp \left\{ -\frac{1}{2} (\hat{C} - C)^T S^{-1} (\hat{C} - C) \right\},
\]

where \( C \) is a vector of \( C_\ell \) values (and similarly \( \hat{C} \)) and \( S^{-1} \) is the inverse signal covariance matrix. However this likelihood function is well-known to be biased, (see e.g., \cite{Bond:1996, Smith:2000}). This Gaussian likelihood can be implemented in two ways. The version discussed in \cite{Bond:1996} considers the case where the signal covariance matrix is derived from the measured power spectrum, rather than the theoretical power spectrum. This results in an overestimation of the errors if the measured power spectrum has fluctuated upwards, and hence upward fluctuations are given less weight, leading to an overall downward bias. The covariance matrix can instead be computed from the theoretical power spectrum, as used in \cite{Verde:2003}, leading to an overestimation of the amplitude.

**Offset Lognormal Likelihood**—A better approximation to the likelihood is the Offset Lognormal approximation (\cite{Bond:1996}), given by

\[
PLN(\hat{C}/C) \propto \exp \left\{ -\frac{1}{2} (\hat{z} - z)^T M(\hat{z} - z) \right\},
\]

where \( z = \ln(C + x_\ell) \) and the matrix \( M \) is related to the inverse covariance matrix by

\[
M_{\ell\ell'} = (C + x_\ell) S_{\ell\ell'}^{-1} (C' + x_\ell')
\]

(The offset factors \( x_\ell \) are simply a function of the noise and beam of the experiment.) This likelihood function is still slightly biased, but in the opposite direction to that of the Gaussian likelihood.

To some extent the Offset Lognormal approximation addresses this issue, in that the transformation of variables from \( C_\ell \) to \( Z_\ell = \ln(C + x_\ell) \) has a constant curvature matrix. Hence the uncertainties on the \( Z_\ell \) do not depend on the \( Z_\ell \) estimated, avoiding the cosmic bias.
To proceed, assume a normal distribution in this new variable $Z_t$ instead of $C_t$. The likelihood is now closer to the exact one. Nevertheless, there is still a small bias opposite to that introduced by the Gaussian approximation in $C_t$.

**WMAP Likelihood**—Taking advantage of the fact that the bias for the Offsets Lognormal likelihood is opposite to that introduced by the Gaussian likelihood, the WMAP team defined a likelihood that is a weighted combination of the two [Verde et al. 2003]:

$$
\ln P_{\text{WMAP}}(\hat{C}|C) = \frac{1}{3} \ln P_{\text{Gauss}}(\hat{C}|C) + \frac{2}{3} \ln P_{\text{LN}}(\hat{C}|C)
$$

(18)

This likelihood is a significantly better approximation for the case of a Gaussian CMB. However, the fact that a likelihood function is accurate in the absence of correlations (when the probability of a measured power spectrum is purely a function of the input power spectrum rather than having any additional dependence on the cosmology) is not a guarantee that it will perform well when applied to a non-Gaussian sky. For instance, [Smith, Challinor, & Rocha 2006] have shown that the WMAP likelihood gives significant biases in the dark energy parameter, $w$, when considering the lensed B-mode power spectrum.

**Equal Variance Likelihood**—The equal variance likelihood proposed by [Bond, Jaffe, & Knox 2000] is given by:

$$
\ln L = -\frac{1}{2} G \left[ e^{-(z-b)} - (1 - (z-b)) \right]
$$

(19)

with

$$
z = \ln \left( q_b + q_b^N \right)
$$

(20)

and

$$
G = \left[ e^{-\sigma_z} - (1 - \sigma_z) \right]^{-1}, \sigma_z = \frac{\sqrt{F_{bb}}}{(q_b + q_b^N)}
$$

(21)

The noise offset $q_b^N$ is estimated using the equation of the maximum likelihood solution for the $q_b$, replacing the observed map with the average of the noise Monte Carlo simulation power spectra $\langle \tilde{N}_\ell \rangle$.

**SCR Likelihood**—[Smith, Challinor, & Rocha 2006] developed a new likelihood to tackle the non-Gaussianity of the lensed sky for studies of the B-mode power spectrum; however, this likelihood has not been extended to the joint probability distribution for all modes. By considering the curvature (with respect to the measured $\hat{C}_\ell$) of the exact log-likelihood expression for a Gaussian sky (equation 14) at its peak, and also the third derivative, a new likelihood can be derived which is Gaussian in $x_\ell = (\hat{C}_\ell)^{1/3}$, where both the second and third derivatives with respect to $x_\ell$ take the correct values at the peak of the likelihood.

The SCR likelihood approximates the normalised distribution $P(\hat{C}_\ell|\theta)$ as Gaussian in some function of the $\hat{C}_\ell$, and takes the form

$$
\ln P(\hat{C}_\ell|\theta) \approx \ln A - \frac{1}{2} \sum_{\ell \ell'} M_{\ell \ell'}^{-1} (\hat{x}_\ell - \mu_\ell)(\hat{x}_{\ell'} - \mu_{\ell'}),
$$

(22)

where

$$
\hat{x}_\ell = C_\ell^{1/3}
$$

(23)

$$
\mu_\ell = \left( \frac{2\ell - 1}{2\ell + 1} C_\ell \right)^{1/3},
$$

(24)

and

$$
M_{\ell \ell'}^{-1} = 3C_\ell^{2/3}\left( \frac{2\ell - 1}{2\ell + 1} \right)^{1/6} S_{\ell \ell'}^{-1} 3C_{\ell'}^{2/3}\left( \frac{2\ell' - 1}{2\ell' + 1} \right)^{1/6}.
$$

(25)

Here $S_{\ell \ell'}$ is the covariance matrix of the measured $\hat{C}_\ell$ at parameters $\theta$. The normalisation is

$$A^{-1} \propto (\det M_{\ell \ell'})^{1/2} \prod_\ell \mu_\ell,
$$

(26)

which can be approximated by $A \propto \prod_\ell 1/C_\ell$.

Applying this expression to a Gaussian simulation gives results almost indistinguishable from the exact likelihood expression, as shown in [Smith, Challinor, & Rocha 2006]. Ignoring the $(2\ell - 1)/(2\ell + 1)$ factors, however, is a bad approximation at low $\ell$, since it ignores the fact that, for a fixed $C_\ell$, the peak of the likelihood is slightly below $C_\ell$. This ends up translating to an underestimation of the theoretical power spectrum. The performance of the Gaussian, WMAP, and two versions of the SCR likelihoods on full-sky lensed simulations was compared in [Smith, Challinor, & Rocha 2006]. The WMAP likelihood function gives very different results to the new SCR likelihood function, and shows a significant bias in the dark energy parameter, $w$. 

2.2.2 Temperature + polarization

We cannot merely extend the above approximations to build a joint likelihood for temperature and polarization. For instance, assume that we approximate the likelihood for $TE$ as a Gaussian. The mode and variance of the Gaussian distribution for $TE$ depend on $TT$ and $EE$. Hence it is not enough to consider the joint likelihood as a product of independent $TT$, $EE$, and $TE$ likelihoods. Indeed the trick is to find a way of coupling these components reliably. Furthermore, given that the $TE$ power spectrum is at times negative, we cannot build a joint likelihood as a product of independent Offset Lognormal likelihoods. As a quick fix, one could try the following:

**Offset Lognormal Bandpower Likelihood**—This likelihood is a joint likelihood for temperature and polarization built as a Gaussian for $TT$ and Offset Lognormal for $TT$, $EE$, and $BB$. However, this approximation does not properly take into account temperature and polarization cross power.

The following two likelihoods, the Hamimeche and Lewis and XFastest likelihoods, do attempt to take temperature and polarization cross power into account properly.

**Hamimeche and Lewis (HL) Likelihood**—The Hamimeche & Lewis (2008) likelihood generalizes the full-sky exact likelihood given by Equation (11) to the cut sky by considering a quadratic expression of the form

$$
\ln L(C_\ell(\hat{C}_\ell)) = -\frac{1}{2} \sum_\ell g(2\ell + 1) \left( \frac{C^{obs}_\ell}{\hat{C}_\ell + \langle \hat{N}_\ell \rangle} \right) + \ln \left( \hat{C}_\ell + \langle \hat{N}_\ell \rangle \right),
$$

where $g(x) = \text{sgn}(x - 1)\sqrt{2(x - \ln x - 1)}$ and $[g(D_\ell)]_{ij} = g(D_{\ell,ij})\delta_{ij}$. This approximation involves a fiducial model so that the covariance can be pre-computed. It is assumed that the matrix of estimators $\hat{C}_\ell$ is positive-definite, although this assumption may break down for some estimators at low-$\ell$.

**XFastest Likelihood**—The XFastest likelihood, introduced in Contaldi et al. (2009) and Rocha et al. (2009), takes the following form for temperature alone:

$$
\ln L = -\frac{1}{2} \sum_\ell g(2\ell + 1) \left( \frac{C^{obs}_\ell}{\hat{C}_\ell + \langle \hat{N}_\ell \rangle} \right) + \ln \left( \hat{C}_\ell + \langle \hat{N}_\ell \rangle \right),
$$

where a tilde connotes a quantity estimated on the cut-sky. The power spectrum is parameterized through a set of deviations $q_\ell$ from a template full-sky spectrum $C^{(S)}_\ell$.

$$
\hat{C}_\ell = \sum_{\ell'} K_{\ell\ell'} B_{\ell\ell'} C^{(S)}_{\ell'} q_\ell,
$$

where $K_{\ell\ell'}$ is the coupling matrix due to the cut sky observations, $B_\ell$ expresses the effect of a finite beam, and $F_\ell$ is a transfer or filter function accounting for the effect of pre-filtering the data in both time and spatial domains.

Extending to polarization, we have

$$
\ln L = -\frac{1}{2} \sum_\ell g(2\ell + 1) \left( \text{Tr} \left( \hat{D}_\ell^{obs} \left( \hat{D}_\ell + \langle \hat{N}_\ell \rangle \right)^{-1} \right) + \ln \left( \hat{D}_\ell + \langle \hat{N}_\ell \rangle \right) \right),
$$

where the matrix $C$ is block diagonal: $C \rightarrow \text{diag}(\hat{D}_{\ell_{\min}}, \hat{D}_{\ell_{\min}+1}, \ldots, \hat{D}_{\ell_{\max}})$, with each multipole’s covariance given by the $3 \times 3$ matrix

$$
\hat{D}_\ell = \begin{pmatrix}
C^{TT}_\ell & C^{TE}_\ell & C^{TB}_\ell \\
C^{ET}_\ell & C^{EE}_\ell & C^{EB}_\ell \\
C^{BT}_\ell & C^{BE}_\ell & C^{BB}_\ell
\end{pmatrix}.
$$

This likelihood follows intuitively from the full-sky exact likelihood, the Inverse Wishart distribution, as given by Equation (11).

2.3 The likelihood at low-$\ell$

The pixel-based Multivariate Gaussian likelihood given by Equation (3) can be computed up to $\ell \approx 30$, 40, and is adequate for comparison with XFastest as shown in Rocha et al. (2009). (Faster methods (see summary by Ashdown et al. 2010) have been developed, but are not necessary here.) We use an implementation known as Blike, described as follows.

A CMB map can be written as an ordered vector $d = (T_{i_1}, T_{i_2}, \ldots, T_{i_7}, Q_{j_1}, Q_{j_2}, \ldots Q_{j_7}, U_{j_1}, U_{j_2}, \ldots U_{j_7})$, comprising all pixels with valid observations. In general $n_T \neq n_P$ and the sets of indexes of temperature and polarization measurements will be different. Assuming that both CMB and noise fluctuations in each pixel are Gaussian-distributed with zero mean, the likelihood for $d$ has the functional form given in equation (1), where the covariance matrix has a block structure:

$$
C = \begin{pmatrix}
<TT>_{(n_T \times n_T)} & <TQ>_{(n_T \times n_P)} & <TU>_{(n_T \times n_P)} \\
<QT>_{(n_P \times n_T)} & <QQ>_{(n_P \times n_P)} & <QU>_{(n_P \times n_P)} \\
<UT>_{(n_P \times n_T)} & <UQ>_{(n_P \times n_P)} & <UU>_{(n_P \times n_P)}
\end{pmatrix}
$$

(32)
Correlations between, e.g., temperature measurements in pixels $i_1$ and $i_2$ can be written as:

$$\langle T_{i_1} T_{i_2} \rangle = \sum_{\ell=2}^{\ell_{\text{max}}} \frac{2\ell + 1}{4\pi} \hat{C}_{\ell} P_{\ell}(\theta_{i_1 i_2}) + N_{i_1 i_2},$$  \hspace{1cm} (33)$$

$P_{\ell}(x)$ are the ordinary Legendre polynomials, and $\theta_{i_1 i_2}$ is the angle between the centers of pixels $i_1$ and $i_2$. Notice that the $\{C_{\ell}\}$ include the contribution of the beam and pixel window, i.e., $\{C_{\ell}\} = \{C_{\ell}^{\text{th}}\} b_{\ell}^2 w_{\ell}^2$, where $\{C_{\ell}^{\text{th}}\}$ is the theory power spectrum and $b_{\ell}$ and $w_{\ell}$ are the harmonic transform of the beam and window functions respectively. For uncorrelated noise, $N_{i j} = n_i^2 \delta_{i j}$, where $n_i$ is the rms noise in pixel $i$. In general, $N$ is a dense matrix. Similar expressions hold for correlations involving $Q$ and $U$ (see e.g.,\cite{tegmark2001}). The choice of $\ell_{\text{max}}$ in Equation (33) depends on several factors, including the smoothing beam, the signal-to-noise ratio, and the pixelization scheme.

3 COMPARING LIKELIHOODS

3.1 Simulations

To compare the performance of $\text{XFaster}$ with other likelihood functions, we use simulations developed within the Planck CTP working group. Planck (\textit{Planck Blue Book} 2005, Tauber et al. 2009) is a full-sky experiment covering frequencies from 30 to 857 GHz with beams ranging in size from 33$'$ to 5$'$. A full description of the simulations is given in\cite{ashdown2013, rocha2009a}. Practical considerations of computational resources having to do with the size of the time-ordered data (TOD), the number of pixels in the maps, and the number of multipoles that had to be calculated, dictated the choice of the 70 GHz channel for the simulations. Higher frequency channels have higher angular resolution and sensitivity, and will extend to smaller angular scales with reduced error bars. Recent increases in computational capability make it possible now to generate thousands of Monte Carlo simulations at the higher frequencies as well. Results will be presented in a future publication\cite{rocha2010a, ashdown2010}.

The 70 GHz simulations used here include the CMB, realistic detector noise, and noise induced by temperature fluctuations of the 20-K hydrogen sorption cooler. To mimic the sensitivity of combination of channels, as would be used for separation of foregrounds with real data, the white noise level was taken to be lower than that expected for the 70 GHz channel alone. The white noise per sample was 2025.8 $\mu$K and the $1/f$ noise power spectrum had knee frequency 0.05 Hz and slope $-1.7$. The input sky signal used to generate the “observed map” was the CMB map derived from the Planck CMB reference sky available\footnote{http://lambda.gsfc.nasa.gov/product/map/dr1/lcdm.cfm} which uses the WMAP 1-year $a_{100}$ up to $\ell = 70$ to generate the CMB. In other words, the large scale structure of the observed map is a WMAP-constrained realization.

The TOD were generated using modules of the Planck simulator pipeline, \textsc{Levels} (Reinecke et al. 2005). Maps were made from the simulated time-ordered data (TOD) using the destriping code Springtide (Poutanen 2003, Ashdown et al. 2007a, b, 2009, Ashdown 2009b). Where a sky cut was applied, it was made at the boundary where the total intensity of the diffuse foregrounds and point sources exceeded twice the CMB sigma. Pixels missing due to the scanning strategy were masked. The beams of the detectors have FWHMs of 13$'$ for planck optical system.

One hundred Monte Carlo signal simulations were generated from the best fit WMAP + CBI + ACBAR $\Lambda$CDM power spectrum\footnote{available in http://lambda.gsfc.nasa.gov/product/map/dr1/lcdm.cfm} with BB mode power set to zero. For the asymmetric beam case, both signal and noise simulations were generated in the time domain.

The “low-$\ell$ dataset” of the Phase 2 simulations was generated directly at $N_{\text{side}} = 16$. The procedure adopted ensured consistency between the low- and high-$\ell$ datasets used to test the $\text{XFaster}$ power spectrum and likelihood estimator; however, the low-$\ell$ dataset lacks the artifacts connected to smoothing and degradation of higher resolution maps that will be present in the final Planck maps.

As pointed out in\cite{rocha2009a}, since the large scale structure of the observed map is derived from real observations, i.e., a WMAP constrained realization, it is not necessarily consistent with the best-fit spectrum at low multipoles. This discrepancy is evident when comparing the cosmological parameters estimated with $\text{XFaster}$ power spectrum and likelihood and the Offset Lognormal likelihood with the theoretical best fit parameters. The Monte Carlo simulations, however, are realizations of the first year WMAP+CBI+ACBAR best fit $\Lambda$CDM power spectrum for Phases 2a and 2b, so such discrepancy is no longer present. Parameters estimated from these Monte Carlo simulations maps are shown to be close to the WMAP best fit parameters.

3.2 Comparisons

\cite{rocha2009a} (Figs. 14 and 15) showed that for $TT$ and $TE$, all of the likelihood approximations described above except the Gaussian

\textcopyright 2005\textsuperscript{2}
likelihood converge to the same form at multipoles above 10 or 20. For $EE$, however, the Gaussian and the Lognormal likelihood differs from the rest up to high $\ell \simeq 10$. Rocha et al. (2009) (Fig. 17) also showed by comparing XFaster with the pixel-based likelihood code (BFlike) that a transition from low-$\ell$ to high-$\ell$ codes was appropriate for Planck in the range $\ell_{\text{trans}} = 30–40$

Here we assess the performance of the XFaster likelihood by comparing cosmological parameters obtained with the XFaster and the Offset Lognormal Bandpower (i.e., Offset Lognormal likelihood for $TT$, $EE$, $BB$, and Gaussian for $TE$) likelihoods. XFaster computes the likelihood of a model passed to it by a modified version of the publicly available CosmoMC code (Lewis and Bridle 2002) for cosmological parameter Markov Chain Monte Carlo estimations. There is no need for window functions or the band power spectrum itself. The inputs are the raw pseudo-$C_\ell$ of the observations plus the kernel and transfer function required by XFaster to relate the cut-sky pseudo-$C_\ell$ to the full-sky $C_\ell$.

For the Offset Lognormal Bandpower likelihood, window functions are required that properly account for band power spectrum correlations. We used two different window functions, top hat (box) and Fisher-weighted (Rocha et al. 2009), written $\text{BF}_{\text{like}}$ code such as C $\ell$, $K_{\ell}$, and $\tau_0$; the spectral index of the initial density perturbation, $\omega_f = 0.05 \pm 0.02$; and $\omega_\Lambda = 0.71$, respectively; the ratio of the sound horizon to the angular diameter distance at decoupling, $\theta_\text{s}$; the spectral index of the initial fluctuation spectrum, $n_s$; the overall normalization of the spectrum $\log[10^{10} A_\text{v}]$ at $k = 0.05 \text{ Mpc}^{-1}$ ($A_s$); the Hubble constant $H_0$; and the reionization redshift $z_{\text{re}}$. We treat $\tau$ in two different ways. Rocha et al. (2009) showed that “high $\ell$” codes could be used to determine parameters from the “observed map” quite well if $\tau$ was fixed in the fit to the value of the input model. This works because $\tau$ is constrained primarily by data at $\ell < 30$. We indicate when $\tau$ is held fixed.

Figure 1 from Rocha et al. (2009), shows one-dimensional marginalised parameter distributions from XFaster, for three cases: 1) the observed map; the ensemble average of Monte Carlo simulations; and the observed map, but holding $\tau$ fixed at its input value and using the XFaster likelihood only for $\ell > 30$. The input parameters are recovered quite well from the ensemble average. The red lines show that if $\tau$ is fixed, high-$\ell$ codes such as XFaster can ignore the low multipoles that they are not designed to calculate, and still recover the other input parameters quite well.

Figure 2 shows parameter distributions from the Offset Lognormal Bandpower likelihood for a Fisher-weighted ($\mathcal{F}_{bb}$) window function, for both the observed map and for the ensemble average of Monte Carlo simulations. The input parameters are recovered from the ensemble average simulated data, but not from the observed map, particularly $A_s$. This is not a surprise. As described in § 3.1, the observed map is fixed by WMAP for $2 \leq \ell \leq 70$. It is therefore affected by low-$\ell$ anomalies arising from any cause, including the details of WMAP

\[ X_{\text{faster}} = \frac{\ell + 1}{2}\sum_{\ell} f_{\ell} \]

which is used to calculate the expectation values for the deviations $q_\ell$(when a shape model, $C_\ell^S$ is considered), or bandpowers $C_b$ (when $C_\ell^S$ is assumed to be flat).

\[ \langle q_\ell \rangle = \frac{\mathcal{I}[W^b_\ell C_\ell]}{\mathcal{I}[W^b_\ell C_\ell^S]} \langle C_b \rangle = \frac{\mathcal{I}[W^b_\ell C_\ell]}{\mathcal{I}[W^b_\ell]} \]

where $W^b_\ell$ is the band power window function, with $C_\ell^S = \ell(\ell+1)C_\ell^{S}/2\pi$. We define normalized window functions

\[ \mathcal{I}[W^b_\ell C_\ell^S] = 1, \]

using the fact that

\[ \langle (C_b^{\text{obs}} - \bar{N}_\ell) \rangle \rightarrow \bar{C}_\ell \]

Extending to polarization:

\[ W^b_\ell = \frac{4\pi}{(2\ell + 1)} \sum_{\ell'} F_{bb} \sum_{\ell} g(2\ell' + 1) \frac{\bar{C}_{\ell'}^{(S)}}{C_{\ell'}^{(S)} + (\bar{N}_{\ell'})^2} K_{\ell'\ell} F_{\ell} B_{\ell}. \]

\[ W^b_\ell = \frac{4\pi}{(2\ell + 1)} \sum_{\ell'} F_{bb} \sum_{\ell} g(2\ell' + 1) \text{Tr} [W_{\ell'\ell'} K_{\ell'}] \]

where $W_{bb} = \bar{D}_{\ell}^{-1} \frac{\partial}{\partial \omega_b} \bar{D}_{\ell}^{-1}$, and $K_{\ell}$ gives the cut-sky response to the individual full-sky multipoles,

\[ K_{\ell} = \left( \begin{array}{c} K_{\ell,TT} F_{\ell,TT}^T B_{\ell}^2 \\ K_{\ell,EE} F_{\ell,EE} B_{\ell}^2 \\ + K_{\ell,TT} F_{\ell,EE} B_{\ell}^2 \\ K_{\ell,EE} F_{\ell,BB} B_{\ell}^2 \\ + K_{\ell,EE} F_{\ell,BB} B_{\ell}^2 \end{array} \right). \]
processing. The WMAP best-fit parameters, on the other hand, are obtained with heavy marginalization of the low-$\ell$ points by foregrounds, and are therefore little affected by the low-$\ell$ anomalies. Since the Monte Carlo simulations are realizations of the WMAP best-fit model parameters, we expect no systematic bias from the ensemble of simulations, but significant offsets from the parameters derived from the observed map, as confirmed.

Figure 5 shows the effect of fixing $\tau$ in the case of the Offset Lognormal Bandpower likelihood for a top-hat window function for the observed map. Green dashed lines are computed from all multipoles, with $\tau$ free to vary, while for the blue solid lines use only $\ell > 30$, with $\tau$ fixed at the input value. As with XFaster in Figure 4 fixing $\tau$ and ignoring low multipoles gives good results with the other parameters.

Figures 2 and 3 show parameter distributions from the Offset Lognormal Bandpower likelihood in the symmetric and asymmetric beam cases, for the observed power spectrum and the ensemble average of the Monte Carlo simulations, respectively. Parameters for the two cases are roughly consistent with each other. They can be compared to the equivalent distributions from the Offset Lognormal Bandpower likelihood for the ensemble average power spectrum of the Monte Carlo simulations.

Table 1 is the same for the Offset Lognormal Bandpower likelihood and the XFaster likelihood for the ensemble average power spectrum of the Monte Carlo simulations. Our aim is solely to check how the parameter uncertainties for Planck from both likelihoods compare to those of WMAP and to our Fisher predictions. The uncertainties are 2–3 times better than those for WMAP except for $\tau$ estimated with the Offset Lognormal Bandpower likelihood, for which the uncertainty on $\tau$ is 0.015 compared to $\sim 0.007$ for XFaster, 0.017 for WMAP, and 0.004 for our Fisher predictions (Rocha et al. 2004). Table I.

Table 1. Parameter estimates and uncertainties from the Offset Lognormal Bandpower and XFaster likelihoods for Planck simulations at 70 GHz, compared to WMAP (Dunkley et al. 2009) and Fisher uncertainties (Rocha et al. 2004). Estimates are for the ensemble average of 100 Monte Carlo simulations. Last column displays input parameter values $\pm$ Fisher uncertainties for reference.

| Parameter | WMAP | Offset Lognormal Bandpower | XFaster | Fisher |
|-----------|------|-----------------------------|---------|--------|
| $\tau$    | 0.087 ± 0.017 | 0.099 ± 0.015 | 0.1105 ± 0.0064 | 0.1103 ± 0.004 (4%) |
| $n_s$     | 0.963 ± 0.015 | 0.965 ± 0.008 | 0.9621 ± 0.0133 | 0.9582 ± 0.004 (4%) |
| $\omega_b$| 0.02273 ± 0.015 | 0.0229 ± 0.0003 | 0.0225 ± 0.00042 | 0.02238 ± 0.00018 (0.8%) |

4 CONCLUSIONS

Parameters estimated with the XFaster and Offset Lognormal Bandpower likelihoods agree well. As the XFaster likelihood is estimated for individual multipoles, a large number of Monte Carlo simulations is required for accurate estimates of low-$\ell$ correlations. If only a small number of Monte Carlo simulations (such as the 100 used in this study), binning of the band power spectrum estimated with XFaster used along with the Offset Lognormal Bandpower likelihood helps to partially correct these correlations. For a large number of Monte Carlo simulations this is unnecessary. There XFaster likelihood, however, has at least three advantages. First, the Offset Lognormal likelihood does not properly take into account the temperature-polarization cross power. This is likely to become evident with a larger number of simulations, or at lower noise levels, such as anticipated with the Planck HFI 143 GHz channel. We are investigating this further; results will be presented in (Rocha et al. 2010). Second, the Offset Lognormal Bandpower likelihood requires calculation of a window function. Third,
the XFaster likelihood can go straight from maps to parameters (via its raw pseudo-$C_\ell$), bypassing the band power spectrum estimation step. These advantages make XFaster a adequate procedure to estimate cosmological parameters from Planck data in the high multipole regime. As a bonus, XFaster performs reasonably well for moderately low multipoles as well. Although hybridization with a likelihood code able to handle fully the challenges of multipoles less than, say, 40, will be necessary for the best estimates parameters, XFaster could be used alone where accuracy can be traded for speed.
Figure 2. Marginalised parameter distributions from the Offset Lognormal Bandpower likelihood (Offset Lognormal for $TT$, $EE$, $BB$, Gaussian for $TE$), for a Fisher-weighted ($F_{bb}$) window function. Blue solid lines are from the observed map. Red dashed lines are from the ensemble average of Monte Carlo simulations. Input parameter values are marked by vertical lines.

5 ACKNOWLEDGMENTS

The work reported in this paper was partially done within the CTP Working Group of the PLANCK Consortia. PLANCK is a mission of the European Space Agency. GR would like to thank useful discussions with Jeff Jewell. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. This work has made use of the HEALPix package (Gorski et al. 2005), and of the PLANCK satellite simulation package, LevelS, (Reinecke et al. 2005), which is assembled by the Max Planck Institute for Astrophysics PLANCK Analysis Centre (MPAC). The PLANCK Project in the US is supported by the NASA Science Mission Directorate. The research described in this paper was partially carried out at the Jet propulsion Laboratory, California Institute of Technology, under a contract with NASA. Copyright 2009. All rights reserved.
Figure 3. Marginalised parameter distributions from the Offset Lognormal Bandpower likelihood for a top-hat window function, for the observed map. Green dashed lines are computed from all multipoles, with $\tau$ free to vary, while for the blue solid lines use only $\ell > 30$, with $\tau$ fixed at the input value. Input parameter values are marked by vertical lines.

REFERENCES
Ashdown M. A. J., Baccigalupi C., Balbi A., Bartlett J. G., et al., A&A, 467, 761, 2007a
Ashdown M., et al., A&A, 471, 361, 2007b
Ashdown M. A. J., Baccigalupi C., Balbi A., Bartlett J. G., et al., A&A, 493, 753, 2009
Ashdown M. A. J., 2009b, in preparation
Ashdown M. A. J., Baccigalupi C., Balbi A., Bartlett J. G., et al., 2010, in preparation
Bond J. R., Jaffe A. H., Knox L., 2000, ApJ, 533, 19
Contaldi C. R., Bond J. R., Crill, Hivon E., et al., 2009, in preparation
Dunkley J. et al., Astrophys. J. Suppl. 180 (2009) 306
Figure 4. Comparison of symmetric (red dashed lines) and asymmetric (blue solid lines) beams in marginalised parameter distributions from the Offset Lognormal Bandpower likelihood, for the observed map. $\tau$ is not fixed.

Fendt W. A., Wandelt B. D., Astrophys. J. 654 (2006) 2
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., et al., ApJ, 622, 759-771, 2005
Hamimeche S., Lewis A., 2008, astro-ph/0801055
Lewis A. and Bridle S., 2002, Phys. Rev. D 66, 103511
Planck Collaboration, Planck Blue Book, arXiv:astro-ph/0604069
Poutanen T., de Gasperis G., Hivon E., Kurki-Suonio H., et al., A&A, 449, 1311, 2006
Reinecke M., Dolag K., Hell R., Bartelmann M. and Enßlin T., 2005
Rocha G., Trotta R., Martins C. J. A. P., Melchiorri A., Avelino P. P., Bean R. and Viana P. T., 2004, MNRAS, 352, 20R
Rocha G., Contaldi C., Bond D., Górski K. M., 2009, submitted to MNRAS
Rocha G., Contaldi C.R., Bond J.R., Górski K.M., 2010, in preparation
Figure 5. Same as Figure 4 but for the ensemble average of the Monte Carlo simulations.

Smith S., Challinor A., Rocha G., 2006, Phys. Rev. D37, 023517
Tauber J., Mandolesi N., Puget J. L., Bersanelli M., et al., Planck Pre-Launch Status: The Planck Mission, 2009, submitted to A&A
Tegmark M., & de Oliveira-Costa A., 2001, Phys.Rev. D64, 063001
Verde L., Peiris H. V., Spergel D. N., Nolta M. R., et al., 2003, ApJS, 148, 195
Figure 6. Comparison of different window functions with the Offset Lognormal Bandpower likelihood. Parameter distributions are shown for the observed map for top-hat (green dashed lines) and a Fisher ($F_{bb}$) (dark blue solid lines) window functions, for the asymmetric beam case. Input model parameter values are marked by vertical black lines. Most parameters improve with $F_{bb}$-windows, while uncertainties are unaffected except in the case of $r$. The 95% upper limit on $r$ is higher by 15% when using $F_{bb}$ windows.
Figure 7. Parameter constraints from the Offset Lognormal Bandpower likelihood with a top-hat window function, computed for the observed power spectrum (black dashed lines) and for the ensemble average of 100 Monte Carlo simulations (blue solid lines).
Figure 8. Same as Figure 7 for a Fisher ($\mathcal{F}_{\text{ba}}$) window function, computed for the observed power spectrum (blue solid lines) and for the ensemble average of 100 Monte Carlo simulations (red dashed lines).