Impact of Galactic polarized emission on \(B\)-mode detection at low multipoles

G. Efstathiou,¹ S. Gratton¹⋆ and F. Paci¹,²,³,⁴

¹Kavli Institute for Cosmology, Cambridge and Institute of Astronomy, Madingley Road, Cambridge CB3 OHA
²Dipartimento di Astronomia, Università degli Studi di Bologna, via Ranzani 1, I-40127 Bologna, Italy
³INAF/IASF-BO, Istituto di Astrofisica Spaziale e Fisica Cosmica di Bologna, via Gobetti 101, I-40129 Bologna, Italy
⁴INFN, Sezione di Bologna, via Irnerio 46, I-40126 Bologna, Italy

Accepted 2009 April 28. Received 2009 April 27; in original form 2009 March 3

ABSTRACT

We use a model of polarized Galactic emission developed by the Planck collaboration to assess the impact of foregrounds on \(B\)-mode detection at low multipoles. Our main interest is in applications of noisy polarization data and in particular in assessing the feasibility of \(B\)-mode detection by Planck. This limits the complexity of foreground subtraction techniques that can be applied to the data. We analyse internal linear combination techniques and show that the offset caused by the dominant \(E\)-mode polarization pattern leads to a fundamental limit of \(r \sim 0.1\) for the tensor–scalar ratio even in the absence of instrumental noise. We devise a simple, robust, template fitting technique using multifrequency polarization maps. We show that template fitting using Planck data alone offers a feasible way of recovering primordial \(B\)-modes from dominant foreground contamination, even in the presence of noise on the data and templates. We implement and test a pixel-based scheme for computing the likelihood function of cosmological parameters at low multipoles that incorporates foreground subtraction of noisy data.

Key words: methods: data analysis – methods: statistical – cosmic microwave background – large-scale structure of Universe.

1 INTRODUCTION

In the last decade, observations of cosmic microwave background (CMB) anisotropies have provided one of the most powerful probes of cosmology. By combining CMB anisotropy data with a variety of other data, many of the key parameters that define our Universe have been determined with unprecedented precision (see e.g. Komatsu et al. 2009, and references therein). Nevertheless, many important questions remained unanswered. One of the most important is the amplitude of a ‘\(B\)-mode’ polarization signature in the CMB. Scalar perturbations generate during inflation generate purely a divergence-like \(E\)-mode polarization pattern in the CMB, whereas tensor perturbations would produce a distinctive curl-like (\(B\)-mode) polarization signature together with an \(E\)-mode pattern of roughly equal amplitude¹ (Kamionkowski, Kosowsky & Stebbins 1997; Zaldarriaga & Seljak 1997).

A detection of a primordial \(B\)-mode anisotropy would provide crucial evidence that inflation actually took place. Furthermore, a measurement of the relative amplitude of the tensor and scalar primordial power spectra (the tensor–scalar ratio \(r\), see Peiris et al. 2003, for a precise definition) would fix the energy scale of inflation via

\[
V^{1/4} \approx 3.3 \times 10^{16} r^{1/4} \text{GeV}
\]

(Lyth 1984) providing critical constraints on inflationary models (see Baumann et al. 2008, for a review). It is therefore no surprise that a number of sensitive ground-based/suborbital CMB polarization experiments are either planned or in progress. Examples include BICEP (Yoon et al. 2006), CLOVER (North et al. 2008), EBEX (Oxley et al. 2004), QUIET (Seiffert et al. 2006) and SPIDER (Crill et al. 2008). In addition, groups in Europe and the USA have considered designs for a \(B\)-mode optimized space satellite capable of probing tensor–scalar ratios \(r \lesssim 10^{-2}\) (Bok et al. 2008; de Bernardis et al. 2009).

⋆E-mail: stg20@cam.ac.uk

¹Gravitational lensing of CMB \(E\)-modes by intervening matter generates a \(B\)-mode anisotropy (see Lewis & Challinor 2006, for a review). This effect will be ignored in this paper, since we will concentrate on the detectability of tensor modes at low multipoles \(\ell \lesssim 20\), where the effects of lensing are small.
The *Planck* satellite\(^2\) is scheduled for launch in 2009 April and has polarization sensitivity in seven channels over the frequency range 30–353 GHz. As described in SPP05, the sensitivity of *Planck* limits B-mode recovery to low multipoles \(\ell \lesssim 20.\)\(^3\) Nevertheless, until a new polarization-optimized satellite is flown, *Planck* is the only experiment capable of probing these low multipoles. It is therefore important to analyse *Planck*’s performance for B-mode detection in the presence of realistic noise levels and polarized foregrounds and therefore to assess whether it can provide useful complimentary data to experiments probing higher multipoles. This is the main goal of this paper.

The problem of detecting primordial B-modes at low multipoles is unusually difficult. Unlike the detection of temperature anisotropies, Galactic foregrounds are expected to have a much larger amplitude than any putative primordial B-mode signal over the entire sky (see Section 2). Accurate foreground removal is therefore essential for B-mode detection at low multipoles. There has been an enormous amount of work on CMB foreground subtraction (see e.g. the review by Delabrouille & Cardoso 2007; Leach et al. 2008). A variety of methods have been developed, ranging from ‘blind’ techniques that make few physical assumptions concerning the Galactic foregrounds [examples include internal linear combination (ILC) (e.g. Bennett et al. 2003), independent component analysis (e.g. Hyvärinen 1999) and its fast and spectral-matching variants (Maino et al. 2002; Delabrouille, Cardoso & Patanchon 2003), ‘semiblind’ methods that make limited use of prior information on the foregrounds (such as maximum entropy; Stolyarov et al. 2002) and template matching (Bennett et al. 2003; Eriksen et al. 2004a; Slosar & Seljak 2004; Slosar, Seljak & Makarov 2004)] and parametric fitting techniques based on physical models of the foregrounds (Eriksen et al. 2006, 2008; Dunkley et al. 2008a). A number of other methods have been developed which incorporate some aspects of these techniques and use, for example, wavelet or harmonic decompositions (e.g. Tegmark, de Oliveira-Costa & Hamilton 2003; Hansen et al. 2006; Norgaard-Nielsen & Jorgensen 2008). Some of these methods provide approximations to the likelihood function for cosmological parameters (e.g. Eriksen et al. 2008) and there has also been some work (Gratton 2008) directly addressing the question of modelling the likelihood function from multifrequency maps. Almost all of these methods have been developed for temperature foreground subtraction. In contrast to the temperature data, the noise level of *Planck* polarization maps will be high. There is, therefore, limited information in polarization and hence a restriction on the complexity of polarized foreground removal algorithms that the data can support.

The problem of B-mode detection in the presence of Galactic foregrounds has been considered by a number of authors. Tucci et al. (2005) performed a Fisher matrix analysis for idealized experiments including foregrounds. Amblard, Cooray & Kaplinghat (2007) investigated harmonic ILC subtraction for B-mode detection at high multipoles (\(\ell \gtrsim 20\)) for various experimental configurations. Betoule et al. (2009) considered the application of spectral matching independent component analysis (SMICA) to perform a Fisher matrix analysis for various experiments, including *Planck*. The work most closely related to ours is the paper by Dunkley et al. (2008b), which focuses on B-mode detection with a future satellite with high signal-to-noise ratio in polarization, rather than the low signal-to-noise ratio case relevant to *Planck*. Low signal-to-noise ratio introduces additional complexity to the foreground subtraction problem, nevertheless, there are strong similarities between our approaches.

The layout of this paper is as follows. The *Planck* sky model (PSM), which is used in this paper to model polarized Galactic foregrounds, is described briefly in Section 2. This model is compared to realizations of the primordial CMB polarization signature to determine the magnitude of the foreground subtraction problem. ILC foreground subtraction is described in Section 3. We show that the ILC method is fundamentally limited for B-mode detection because of the offset associated with the dominant E-mode signal. Section 4 analyses foreground template subtraction techniques and we present a heuristic model for constructing a pixel-based polarization likelihood function. This is applied to simulations with *Planck*-like noise. Section 5 introduces a classification scheme for foreground subtraction methods based on their dominant errors. Our conclusions are summarized in Section 6.

2 THE MAGNITUDE OF THE PROBLEM

2.1 The *Planck* sky model

The PSM has been developed by the *Planck* Component Separation Working Group for use in simulations of the *Planck* mission. Summaries of the model are given by Dunkley et al. (2008b) and Leach et al. (2008) and a detailed description will be provided in a forthcoming paper by Delabrouille et al. (in preparation). The polarized foreground model used in this paper is similar to that used by Dunkley et al. (2008b). Briefly, the model includes polarization from a power-law synchrotron component with geometrical suppression factors, polarization angles and polarization fractions based on the magnetic field model of Miville-Deschenes et al. (2008). It also includes a power-law dust component based on an IRAS dust template derived by Finkbeiner, Davis & Schlegel (1999). Polarization from point sources is ignored. The resulting dust polarization fraction in this model is \(\sim 5\) per cent over most of the sky, corresponding to the ‘high’ polarization fraction used in Dunkley et al. (2008b). Preliminary indications from the BICEP experiment suggest a significantly lower polarization fraction of \(\sim 1–2\) per cent close to the Galactic plane, though this figure may be unrepresentative of regions at higher Galactic latitude where depolarization may be lower (BICEP collaboration, private communication).

\(^2\) For a description of *Planck* and its science case, see *The Scientific Programme of Planck* (2005), hereafter SPP05.

\(^3\) Although in theory it may be possible to extract some information on primordial B-modes from the noise-dominated data at higher multipoles, a number of systematic effects such as cross-polar leakage and errors in the polarizer angles are expected to dominate at higher multipoles.
Figure 1. \( Q \) (left) and \( U \) (right) maps: the upper panel shows CMB simulations. The remaining panels show the PSM at 70, 100, 143 and 217 GHz. The temperature scale (thermodynamic temperature) is in \( \mu \)K. All maps were generated at HEALPix NSIDE = 2048.

Fig. 1 shows \( Q \) and \( U \) simulations at HEALPix (Gorski et al. 2005) resolution NSIDE = 2048 for a single realization of the concordance cold dark matter model.\(^4\) The model has a tensor–scalar ratio \( r = 0.1 \). The PSM of polarized foregrounds is shown in the lower panels over

\(^4\) Throughout this paper, apart from the tensor–scalar ratio \( r \), we use the cosmological parameters derived from the 3-year Wilkinson Microwave Anisotropy Probe (WMAP) data (Spergel et al. 2007) assuming zero curvature and a single scalar spectral index. The tensor spectral index is fixed at \( n_t = 1 \).
the frequency range 70–217 GHz with the colour (grey) scales adjusted to span the range −30 to 30 μK. The structure of the foreground is fairly similar at each of these frequencies, and so we use the 217-GHz maps to define a polarization mask by simply applying a threshold to each of the $Q$ and $U$ maps. An ‘internal mask’ is then defined as the union of the two $Q$ and $U$ masks. For the tests described in this paper we use a fairly conservative mask which removes 37 per cent of the sky.

2.2 Impact of foregrounds on $B$-mode detection

Fig. 1 is not particularly useful for assessing tensor mode detection with Planck, since Planck will be noise dominated in the $B$-mode for all but the lowest multipoles. Smoothed maps, as shown in Fig. 2, are of more relevance. Here we show the maps of Fig. 1 in the regions outside the internal mask after smoothing with a Gaussian of full width at half-maximum (FWHM) 7° and repixelization to a resolution of NSIDE = 16. One can see that at this resolution (almost signal dominated for Planck) the peak-to-peak variations in the polarization maps of the primary CMB are of order $\sim 0.7$ μK. The PSM is shown for the regions outside the internal mask, but unlike Fig. 1, the temperature scales of the colour tables are set by the true maximum and minimum values of the maps. Evidently at the most sensitive Planck channel (143 GHz) foregrounds dominate over the primary CMB signal over the whole sky. In fact, to get a better feel for how accurately we need to subtract foregrounds for $B$-mode detection, Fig. 3 shows the contributions to the $Q$ and $U$ maps from the $B$-mode alone for $r = 0.1$. The rms contribution from the $B$-mode at this resolution is about a quarter of the rms of the $E$-mode, and so foreground subtraction to significantly better than 5 per cent accuracy at 143 GHz is required to detect a $B$-mode with $r = 0.1$. As we will show, this presents a formidable challenge for Planck, even with the simplified foreground model assumed in the PSM. The rms values for the PSM and for the primary CMB maps shown in Figs 2 and 3 are listed in Table 1. Power spectra for the maps shown in Fig. 2, computed using the pseudo-$C_\ell$ (PCL) estimator, are plotted in Fig. 4.

3 ILC COMPONENT SEPARATION

3.1 Summary

The internal linear combination method is very simple and well known (see e.g. Bennett et al. 2003; Eriksen et al. 2004b). Suppose we have $M$ maps $T_i(p)$ (temperature, $Q$ or $U$ polarization at each pixel $p$) at different frequencies, we find the linear combination

$$T(p) = \sum_i w_i T_i(p)$$

that minimizes the variance

$$\text{Var}(T) = \langle T^2 \rangle - \langle T \rangle^2$$

subject to the constraint that the sum over the $w_i$ is equal to unity. (Angular brackets in this section denote averages over map pixels.) Equation (3) can be minimized using a Lagrange multiplier. If the map covariance matrix

$$M_{ij} = \frac{1}{N_p} \sum_p \left[ T_i(p) - \langle T_i \rangle \right] \left[ T_j(p) - \langle T_j \rangle \right]$$

is invertible then the solution is

$$w_i = \frac{\sum_j M_{ij}^{-1}}{\sum_k M_{kj}^{-1}}$$

and we denote the corresponding map $T_{\text{ILC}}(p)$. (If $M_{ij}$ is not invertible then a family of solutions for $w_i$ exists.) As a slight variation of this method for noisy data, one might consider subtracting the noise variance from equation (4) so that we minimize primarily on the foreground residuals. In this case, the solution is identical to (5) with $M_{ij}$ replaced by

$$M_{ij} \rightarrow M_{ij} - \sigma_i^2 \delta_{ij},$$

where $\sigma_i^2$ is the noise variance of the map at frequency $i$. [Note that this variant is inconsistent with a Bayesian formulation of ILC presented in Gratton (2008).]

3.2 ILC with variable foregrounds

Let us ignore instrumental noise, but assume that the foregrounds differ with frequency. At each frequency $i$, we can write the data vector as

$$T^i = S + F^i$$

5 Note that as described in Efstathiou (2006) these PCL estimates are highly suboptimal at low multipoles, but they are perfectly adequate for illustrating the magnitude of the foreground problem. We will discuss more optimal methods in Sections 3 and 4.
Figure 2. As Fig. 1, but with maps generated at HEALPix NSIDE = 16 and a smoothing of 7° FWHM. The internal mask described in the text has been applied to the PSM.

where $S$ is the frequency independent CMB signal and $F^i$ is the foreground. (For simplicity, we assume that $S$ and $F^i$ have zero mean.) To find the ILC solution, we must extremize the quantity

$$\Sigma^2 = \sum_{ij} w_i w_j (\langle S + F^i \rangle \langle S + F^j \rangle ) + \lambda \left( \sum_i w_i - 1 \right)$$

(8)
Figure 3. Maps of the $B$-mode contributions to the primary CMB realizations ($r = 0.1$) shown in Fig. 2.

Table 1. RMS residuals outside internal mask.$^a$

| Map                | $T$ (μK) | $Q$ (μK) | $U$ (μK) |
|--------------------|----------|----------|----------|
| CMB                | 50.4     | 0.257    | 0.232    |
| CMB $B$-mode       | 2.26     | 0.062    | 0.064    |
| PSM 30 GHz         | 97.2     | 4.29     | 3.94     |
| PSM 44 GHz         | 36.0     | 1.51     | 1.37     |
| PSM 70 GHz         | 19.0     | 0.661    | 0.608    |
| PSM 100 GHz        | 22.7     | 0.720    | 0.692    |
| PSM 143 GHz        | 40.8     | 1.29     | 1.27     |
| PSM 217 GHz        | 119.2    | 3.82     | 3.76     |
| PSM 353 GHz        | 874.1    | 28.3     | 27.9     |

$^a$For maps smoothed with a Gaussian of FWHM 7$^\circ$.

with respect to the weights $w_i$ and $\lambda$. The solution is

$$w_i = \frac{(1 + \sum_{kj} X_k F^{-1}_{kj}) \sum_j F^{-1}_{ji} - \sum_k X_k F^{-1}_{ki}}{\sum_k F^{-1}_{ki} \sum_{kj} F^{-1}_{kj}}. \quad (9)$$

where

$$X_i = \langle SF^i \rangle, \quad F_{ij} = \langle F^i F^j \rangle. \quad (10)$$

and we have assumed in (9) that there is enough variation in the foregrounds with frequency that the matrix $F_{ij}$ is non-singular and therefore invertible. Notice that in the limit of zero signal, $X_i = 0$, the solution is just equation (5) with $M_{ij} = F_{ij}$. In this case the variance of the ILC solution is

$$\langle (\Delta T_{\text{ILC}})^2 \rangle = \left( \sum_{ij} F_{ij}^{-1} \right)^{-1}. \quad (11)$$

Figure 4. PCL $E$- and $B$-mode power spectrum estimates computed for the CMB simulations and foreground components of Fig. 1. The power spectra are computed for the region of the sky outside the internal mask. No instrumental noise has been added to the simulations. The blue points show the power spectrum estimates for the CMB. The red lines show the theoretical input CMB spectra. The foreground power spectra are as follows: 70 GHz (dark blue); 100 GHz (purple); 143 GHz (orange); 217 GHz (green).
This gives an indication of the best possible foreground subtraction achievable through ILC, i.e. of the limitation imposed by ‘foreground mismatch’. (We will apply the term ‘foreground mismatch’ generally to subtraction techniques to denote the residual contamination by foregrounds even in the absence of CMB signal and instrument noise.) The terms in equation (9) proportional to $X_i$ lead to an offset in the ILC solution with an amplitude proportional to $X_i$ and independent of the amplitude of the foreground. Such offsets exist even if there is not enough variation in the foregrounds to mimic the CMB. Consider the well-known case of two frequency channels with an identical foreground template:

$$\begin{align*}
T_1 &= S + \alpha F, \\
T_2 &= S + F
\end{align*}$$

(12)

(without loss of generality, since there are only two unknown functions, $S$ and $F$). The ILC solution in this case gives

$$T_{\text{ILC}} = S - \frac{\langle SF \rangle}{\langle F^2 \rangle} F,$$

(13)

independent of the parameter $\alpha$. The error in (13), which we will term ‘cross-correlation offset’ [called ‘cosmic covariance’ by Chiang, Naselsky & Coles (2008)], is independent of the amplitude of the foreground template and is irremovable, since the amplitude of the effect depends on the specific realization of the CMB signal. Equation (10) generalizes this result to an arbitrary number of channels and includes the bias caused by foreground mismatch. The cross-correlation offset averages to zero over a large number of realizations of the CMB. However, it can be significant for any single realization. Cross-correlation offset is particularly serious for $B$-mode polarization measurements. This is because we are trying to detect a $B$-mode signal that is a small fraction of the $E$-mode contribution to the $Q$ and $U$ maps (cf. Fig. 3). The cross-correlation terms $X_i$ for the $Q$ and $U$ maps are therefore dominated by the contribution from $E$-modes, and since these offset terms are fixed for a particular realization of the $E$-modes, they set a fundamental limit on the amplitude of a $B$-mode signal detectable via ILC subtraction.

This is illustrated in Fig. 5, which shows the likelihood distributions of the tensor–scalar ratio, $r$, for 10 ILC cleaned CMB realizations generated with $r = 0$. (The histogram plotted in the figure shows the peak likelihood values for 500 simulations.) In each case, we generated an NSIDE = 16 low-resolution CMB map smoothed with a Gaussian of FWHM 7° and added the PSM foregrounds at the four frequencies 70, 100, 143 and 217 GHz as plotted in Fig. 2. The internal mask of Fig. 2 was applied to each map and an ILC cleaned map was computed from the solution (5).

The pixel likelihood function is

$$\mathcal{L} \propto \frac{1}{\sqrt{|C|}} \exp \left( -\frac{1}{2} x^T C^{-1} x \right),$$

(14)

where $x$ is the $(T, Q, U)$ ILC map and $C$ is the sum of the signal ($S$) and noise ($N$) covariance matrices

$$C_{ij} = \langle x_i x_j \rangle, \quad C = S + N.$$

(15)

In the examples shown in Fig. 5, we compute the likelihood function for the $Q$ and $U$ maps allowing only the parameter $r$ to vary. (This is a very good approximation to a full likelihood analysis because $r$ is weakly correlated with other cosmological parameters.) A small diagonal noise of 0.1 μK was added to the $Q$ and $U$ maps to regularize the signal covariance matrix, which is otherwise numerically singular because the maps are overpixelized. Fig. 5 shows that the cross-correlation offset causes biases $r \sim 0.03$, but with a large spread so that biases in excess of $r \sim 0.10$ are seen. The results shown in Fig. 5 are insensitive to the number of channels used in the ILC.
Figure 6. QML estimates of the $E$- and $B$-mode polarization spectra for the realization plotted in Fig. 3 with $r = 0$. The upper panel shows the analysis of the CMB maps on the cut sky. The lower panels show the analysis of the ILC cleaned CMB maps. The error bars show the diagonal components of (20) using the theoretical input values of $r$ for each realization. The solid lines show the theoretical input spectra for $r = 0.1$.

The effects of the cross-correlation offset on the power spectra are illustrated in Fig. 6 for the specific CMB realization of Fig. 2 with $r = 0$. Here we plot the quadratic maximum likelihood (QML) estimates (Tegmark & de Oliveira-Costa 2001; Efstathiou 2006)

$$
\hat{C}_\ell = \mathcal{F}^{-1} y,
$$

where

$$
y_\ell = x_i x_j E_{ij}^r, \quad r \equiv (T, X, E, B),
$$

with matrices $E^r$ given by

$$
E^r = \frac{1}{2} C^{-1} \frac{\partial C}{\partial \ell} C^{-1}.
$$

$x$ and $C$ are the data vector and data covariance matrix defined in equation (14) and $\mathcal{F}$ is the Fisher matrix

$$
\mathcal{F}^{\ell\ell'} = \frac{1}{2} \operatorname{Tr} \left[ \frac{\partial C}{\partial \ell} C^{-1} \frac{\partial C}{\partial \ell'} C^{-1} \right].
$$

The covariance matrix of the QML estimates is given by the inverse of the Fisher matrix (19):

$$
\langle \Delta \hat{C}_\ell \Delta \hat{C}_{\ell'} \rangle = \mathcal{F}^{-1}.
$$

The upper panels of Fig. 6 show the QML estimates for the CMB maps alone computed from the cut sky. (Regularizing noise of 0.1 $\mu$K was added to the maps as in the likelihood analysis of Fig. 5.) The error bars on the points are computed from the diagonal components of the Fisher matrix (19) assuming $C_B^\theta = 0$. The lower panels show the QML power spectra for the ILC cleaned maps, illustrating the cross-correlation offset.

A significant cross-correlation offset is inherent in any ‘blind’ component separation technique. For the ILC method as presented above, the size of the offset is set by the amplitude of the $E$-mode signal with a resulting large effect on the $B$-mode. In principle, the cross-correlation offset could be reduced by first disentangling $E$ and $B$-modes and performing ILC independently on maps constructed from the two sets of modes. If the whole sky is available, the separation into $E$ and $B$-modes is exact and unambiguous. If the sky coverage is incomplete, one
could decompose the maps into almost pure $E$ and $B$-modes following the techniques described by Lewis (2003). Such a decomposition will necessarily lead to information loss since ambiguous modes must be discarded. Even if near pure $B$-modes are identified, so reducing the cross-correlation offset, we will see in the next subsection that an ILC solution can amplify the instrument noise to unacceptably high levels. As we will show in Section 4, rather than invoke some type of modal decomposition, it is possible to reduce the cross-correlation offset to negligible levels by template fitting and to avoid catastrophic amplification of instrument noise. The scheme outlined in Section 4 has the added advantage of leading to a simple model for the full polarization likelihood function.

### 3.3 ILC with noise

So far, we have ignored instrumental noise. In the presence of instrumental noise, the solution for the ILC weights is approximately (9) with $F_{ij}$ replaced by

$$\tilde{F}_{ij} = F_{ij} + \hat{\sigma}_i^2 \delta_{ij},$$

where $\hat{\sigma}_i^2$ is the noise variance of the map at frequency $i$. (This solution is exact if the noise–signal $\langle NS \rangle$ and noise–foreground $\langle NF \rangle$ covariances can be neglected.) If noise dominates in the frequency covariance matrix, the ILC solution corresponds to inverse noise-variance weighting,

$$w_i = \frac{1}{\hat{\sigma}_i^2} \left( \sum_k \frac{1}{\hat{\sigma}_k^2} \right)^{-1},$$

which has, of course, nothing to do with foregrounds. One can get a rough idea of how noise will affect the ILC solution by computing

$$\langle (\Delta T_{\text{ILC}})^2 \rangle = \left( \sum_i \tilde{F}_{ij}^{-1} \right)^{-1}$$

instead of (10). In the limit that noise dominates, the ILC residuals will simply reflect the noise level of the ILC map,

$$\hat{\sigma}_{\text{ILC}}^2 = \sum_i w_i^2 \hat{\sigma}_i^2 \approx \left( \sum_i \frac{1}{\hat{\sigma}_i^2} \right)^{-1},$$

where we have assumed that the noise at each frequency is independent.

Table 2 lists the ILC weights for the foregrounds alone (equation 5 with $F_{ij}$ replacing $M_{ij}$) computed for the pixels outside the internal mask shown in Fig. 2. The ILC residuals using these weights are listed in the last column of the table. For the $Q$ and $U$ maps the residuals are an order of magnitude lower than the rms contribution expected for a $B$-mode with $r = 0.1$. Thus, if the PSM model is correct, it is in principle possible to assign weights that subtract foregrounds to achieve a limit of $r \sim 10^{-3}$ (i.e., for the PSM, ‘foreground mismatch’ is negligible). However, three of the four weights are considerably larger than unity and will therefore amplify any instrument noise that is present (equation 24). If somehow we were given these weights, we would only able to make use of them if the instrument noise at each channel were extremely low. For example, using the instrumental sensitivities in SPP05 for Planck, these weights would lead to disastrously high noise levels of $\sim 0.6 \mu K$ for the ILC cleaned $Q$ and $U$ maps. Such a high noise level is of no use for $B$-mode detection.

The situation becomes even worse if we include the CMB in the ILC solution (lower entries in Table 2). The ILC solution listed in Table 2 is for the CMB realization of Fig. 2 with $r = 0$. The polarization power spectra for this solution are shown in the lower panel of Fig. 6. The cross-correlation offset amplifies the ILC residuals by more than an order of magnitude compared to the foreground-only solution (see the last column in Table 2) and so some foreground $B$-mode leaks into the power spectrum at low multipoles as shown in Fig. 6. This is consistent with the fundamental limit of $r \sim 0.1$ imposed by the cross-correlation offset discussed in the previous section. However, notice

### Table 2. ILC weights.

| Frequency | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $\langle (T_{\text{ILC}} - T_{\text{CMB}})^2 \rangle^{1/2}$ (μK) |
|-----------|-------|-------|-------|-------|--------------------------------------------------|
| 70 GHz    |       |       |       |       |                                                 |
| 100 GHz   |       |       |       |       |                                                 |
| 143 GHz   |       |       |       |       |                                                 |
| 217 GHz   |       |       |       |       |                                                 |

Notes. The upper table lists the ILC weights, $w_i$, computed from the foregrounds alone (i.e., equation 5 with $F_{ij}$ replacing $M_{ij}$). The column labelled $\langle (T_{\text{ILC}} - T_{\text{CMB}})^2 \rangle^{1/2}$ lists the rms residual of the ILC cleaned maps and the true CMB maps for the regions outside the internal mask shown in Fig. 2. The lower table lists the ILC weights for the combined noise-free foreground and CMB maps shown in Fig. 2.
that the weights are now even larger than for the foreground-only case and so any instrumental noise will be highly amplified in the ILC solution.

For realistic experiments, we are therefore caught between a rock and a hard place. In the presence of instrumental noise, we would like to minimize the noise when combining frequency channels (equation 22). However, this will not remove foregrounds. If we use weights that minimize the foreground residuals (which we cannot find in principle because of cross-correlation offset) we amplify the noise to unacceptable high levels.

4 TEMPLATE FITTING

The discussion of the previous section shows that a purely blind component separation method such as ILC is fundamentally limited for B-mode detection by the cross-correlation offset, even if foreground mismatch is negligible (i.e. a linear combination exists which eliminates the foregrounds to high accuracy). To reduce the cross-correlation offset, a semiblind technique is required which utilizes supplementary information on either the foregrounds or the primary CMB signal. In this section we investigate template fitting and show that this provides an acceptable method for B-mode analysis for Planck.

4.1 Summary

Let us model the data vector as

\[ x = s + F \beta + n. \]

where \( s \) is the signal, \( F \) is a template matrix, \( \beta \) is a vector of unknown parameters and \( n \) is the pixel noise vector. For example, the data vector \( x \) could be a vector of length \( 2N_p \) consisting of the \( Q \) and \( U \) maps \( x \equiv (Q, U) \), \( \beta \) could be a vector of four unknown amplitudes \( (\beta_Q^0, \beta_U^1, \beta_Q^1, \beta_U^2) \) and \( F \) a \( 2N_p \times 4 \) matrix consisting of two \( Q \) and two \( U \) foreground template maps:

\[
F = \begin{pmatrix}
F_Q^1(1) & 0 & F_Q^3(1) & 0 \\
0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
F_Q^1(N_p) & 0 & F_Q^3(N_p) & 0 \\
0 & F_U^1(1) & 0 & F_U^3(1) \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & F_U^1(N_p) & 0 & F_U^3(N_p)
\end{pmatrix}
\]

We find \( \beta \) by minimizing

\[ \chi^2 = (x - F \beta)^T C^{-1} (x - F \beta), \]

where \( C \) is the covariance matrix (15). The solution is

\[ \hat{\beta} = (F^T C^{-1} F)^{-1} (F^T C^{-1} x). \]

The minimum variance estimate of the signal vector, \( \hat{s} \), is the Wiener filtered

\[ \hat{s} = SC^{-1}(x - F \beta) \]

(see e.g. Rybicki & Press 1992). If the data vector is noise-free and contains zero foreground, we see that template matching recovers a ‘biased’ estimate of the signal,

\[ \hat{s} = s - F (F^T C^{-1} F)^{-1} F^T C^{-1} s. \]

This is the analogue of (13) for template matching (and is identical for a single foreground/template if the covariance matrix \( C \) is proportional to the identity matrix). Notice that as with equation (13) the offset is independent of the amplitude of the foreground template. Even if there is no foreground in our signal, template matching will produce a cross-correlation offset in the recovered signal that is independent of the amplitude of the foreground. As with the ILC method, the cross-correlation offset gives a fundamental irreducible limit on the amplitude of a B-mode component detectable by template matching. The critical difference with the ILC method is that the amplitude of the offset depends on the mismatch between the foreground matrices, \( F_Q(i)F_U(j) \) etc., and the appropriate submatrices \( C_{Q_Q} \) etc. of \( C \). The bigger the mismatch between the spatial structure of the foregrounds and the primordial CMB, the smaller the cross-correlation offset. The method is therefore ‘semiblind’ because it uses prior information on the signal+noise covariance matrix to determine the vector \( \beta \). As we will show below, this prior information reduces the cross-correlation offset by more than an order of magnitude compared to the ILC method which makes no assumption concerning the spatial structure of the CMB.

If the data vector is noise-free, but contains a foreground component \( F' \), \( F' \neq F \beta \) the minimum variance signal estimate will contain a foreground-mismatch term in addition to the cross-correlation offset:

\[ \hat{s} = s - F (F^T C^{-1} F)^{-1} F^T C^{-1} s + [F' - F (F^T C^{-1} F)^{-1} (F^T C^{-1} F')]. \]
Evidently, both terms must be shown to be small for $B$-mode detection.

As an illustration of the method, we have applied it to the noise-free maps illustrated in Fig. 2. We first construct inverse noise variance weighted $Q$ and $U$ maps (equation (22)) of the CMB ($r = 0$)+PSM for the four frequency channels 70, 100, 143 and 217 GHz using the SPP05 detector sensitivities. (The motivation for this will be made clear in the next section.) The resulting data vector $x$, although noise-free, is contaminated with both synchrotron and dust polarization foregrounds. We then use the 30-GHz PSM as a low-frequency foreground template and we use either the 217- or 353-GHz PSM as a high-frequency foreground template. We then compute the template subtracted data vector, $x - \mathbf{F} \beta$, and feed this into the likelihood function (14) to compute the posterior distribution of $r$. The results are shown in Fig. 7. The 30- and 217-GHz PSM templates remove foregrounds to extremely high accuracy of $r \lesssim 10^{-3}$. The cross-correlation offset term in equation (31) is negligible. The primary source of error is caused by the foreground mismatch in equation (31), though for these choices of template, the error is small. Using the 30- and 353-GHz templates, the peak of the likelihood is offset by $\sim 3 \times 10^{-3}$ as a result of foreground mismatch term at high frequencies. However, the offset is much smaller than the width of the likelihood distribution expected for Planck sensitivities and is therefore ignorable. As we will see in Section 4.3, because of Planck’s high detector noise levels, it is better to use the 353-GHz channel as a high-frequency polarization template because it has a higher foreground signal-to-noise ratio than the 217-GHz channel.

### 4.2 Parameter estimation with template marginalization

In this subsection, we will discuss the problem of parameter estimation from multiple noisy maps with foreground subtraction. The aim is to use the results of this subsection to construct an approximate, but easily calculable estimate of the likelihood function.

The data ($I$, $Q$ and $U$) per pixel $i$ at each frequency is arranged into a single column vector $\mathbf{X}$. Assume the instrument noise is described by a corresponding ‘large’ covariance matrix $\mathbf{N}$. The foregrounds are modelled by a mean map at each frequency with some spread. We denote the mean by the vector $\mathbf{F}_0$ and encode the spread via a Gaussian distribution with a covariance matrix $\mathbf{P}$. ($\mathbf{P}$ can be very general, being a ‘large’ matrix, but it can encode ‘simple’ uncertainties such as a global spectral index uncertainty, or an uncertainty about whether a particular pixel is heavily contaminated or not.)

We assume that CMB has a blackbody, spectrum, so that the CMB-induced signal from each pixel is given by $\mathbf{es}(i)$, or, for the entire data set, by $\mathbf{E}s$, where $\mathbf{s}$ is the entire CMB signal, and $\mathbf{E} \equiv \mathbf{I} \otimes \mathbf{e}$. The matrix $\mathbf{E}$ therefore duplicates the CMB signal $\mathbf{s}$ into each of the frequency bands. Finally, we assume that the foregrounds and CMB are statistically independent.

Since the CMB is assumed to be Gaussian, the CMB signal $\mathbf{s}$ is distributed as

$$
\frac{1}{\sqrt{\mid \mathbf{S} \mid}} e^{-\frac{1}{2} \mathbf{s} \mathbf{T} \mathbf{S}^{-1} \mathbf{s}},
$$

where $\mathbf{S}$ is the signal covariance matrix. The model is assumed to affect the CMB only through $\mathbf{S}$, so the probability for a model may be equated to the probability for $\mathbf{S}$. From our assumptions above, the prior on the foreground signal is taken to be

$$
\frac{1}{\sqrt{\mid \mathbf{P} \mid}} e^{-\frac{1}{2} \mathbf{F} - \mathbf{F}_0 \mathbf{T} \mathbf{P}^{-1} (\mathbf{F} - \mathbf{F}_0) \mathbf{d} \mathbf{F}}.
$$

---

**Figure 7.** Distributions of the tensor–scalar ratio for template subtracted noise-free combined frequency $Q$ and $U$ maps (frequency range 70–217 GHz), constructed as discussed in the text. Two curves are shown, one using the 30- and 217-GHz PSM maps as templates and the other using the 30- and 353-GHz PSM.
(Note we assume that $P^{-1}$ exists, though the derivation given below can easily be modified if it does not.) Noise introduces a mismatch between $X$ and $E + F$, so the probability density for the data given the signal and foreground is

$$
\frac{1}{\sqrt{|N|}} e^{-\frac{1}{2} (X - E - F)^T N^{-1} (X - E - F)} dX.
$$

(34)

To obtain the probability for a model given the data, we need to combine equations (32)–(34), multiply by a prior $p(S)$ and integrate over the signal and uncertainties in the foreground model:

$$
p(S|X) \propto p(S) \int dF \frac{1}{\sqrt{|N|}} e^{-\frac{1}{2} (X - E - F)^T N^{-1} (X - E - F)} \frac{1}{\sqrt{|P|}} e^{-\frac{1}{2} (F - F_0)^T P^{-1} (F - F_0)} \frac{1}{\sqrt{|S|}} e^{-\frac{1}{2} (X + S - F_0)^T S^{-1} (X + S - F_0)}.
$$

(35)

From now on we will drop $S$-independent factors without comment. Writing the exponent as

$$
-\frac{1}{2} (s^T F^T \begin{pmatrix} E^T N^{-1} E & E^T N^{-1} F \\ N^{-1} F & N^{-1} + P^{-1} \end{pmatrix} s + (s^T F^T \begin{pmatrix} E^T N^{-1} x \\ N^{-1} x + P^{-1} F_0^{-1} \end{pmatrix} ,
$$

(36)

performing the integral over $s$ and $F$ and using the formulae for block matrices and determinants in Press et al. (1992) the result simplifies to the compact expression

$$
\frac{p(S)}{\sqrt{|S + M|}} e^{-\frac{1}{2} (s^T (S + M)^{-1} s + y)},
$$

(37)

where we have defined $M$ via

$$
M^{-1} \equiv E^T (N + P)^{-1} E,
$$

(38)

and $Y$ is

$$
Y \equiv ME^T (N + P)^{-1} (X - F_0).
$$

(39)

Note that equation (38) has a simple interpretation. It is just the usual likelihood function (14) applied to a map $Y$ constructed from the data with a noise covariance matrix $M$ given by equation (38). Furthermore, the map (39) is basically an inverse-noise weighted map over the frequency channels (but note the ‘noise’ here is the sum of the detector noise $N$ and the foreground uncertainty $P$ and that the best-guess foreground $F_0$ is subtracted from the data at the start).

This derivation provides the justification for the following simplified model for a polarization likelihood for Planck. In Section 4.1, we demonstrated that by using two templates, foreground-mismatch could be reduced to levels that are negligible compared to the Planck instrumental noise. It is therefore a good approximation to neglect $P$ in comparison to $N$. The map $Y$ can then be approximated by subtracting the best-fitting templates (as computed in Section 4.1) from an inverse noise-variance weighted set of maps. This justifies the procedure used to construct Fig. 7.

Gratton (2008) presents a conceptually similar technique for constructing a likelihood function from noisy data in the presence of foregrounds. Given assumed priors, the scheme effectively averages over linearly weighted combinations of frequency channels. The advantage of that technique is that it self-consistently marginalizes over foregrounds, rather than selecting specific channels as templates as described above. In fact, in the noise-free limit, to the extent that (25) (with $n = 0$) is an accurate model of the data, the technique gives best-fitting weights, $w_i$, that also yield (30) as the most likely CMB map, mitigating foreground mismatch. However, with noise, the algorithm for the resulting likelihood function is considerably more complicated. A detailed comparison between the two approaches is currently in progress.

4.3 Application to simulations with noise

In this subsection, we apply the above formalism to simulations that include the Planck noise levels. The detector noise is assumed to be white and uncorrelated in each pixel. It would be straightforward to generalize this analysis to include a more realistic scanning strategy and correlated ‘destriping’ noise (all that is required is an appropriate model for the noise covariance matrices $N_i$ that includes destriping errors). However, correlated errors are expected to be small for Planck and so we ignore them in this analysis (for a detailed discussion see Efstathiou 2007). We therefore add uncorrelated white noise to the NSIDE = 2048 primary CMB+foreground maps at each of the polarization-sensitive Planck frequencies. The maps are smoothed with a Gaussian of 7’ FWHM and reconstructed at NSIDE = 16. The low-resolution noise covariance matrices are computed using the small-angle approximation (see Appendix A), which gives an excellent approximation to the true covariance matrices.

In reality, the templates will contain a primary CMB signal and so the likelihood approximation is slightly more complicated than implied in Section 4.2. As described above, the data vector $x_i$ (where the subscript $i$ denotes pixel number) is constructed as an inverse noise variance weighted sum over a set of frequency channels. The covariance matrix of this vector is written as

$$
(x_i x_j) = S_{ij} + \Phi_{ij} + N_{ij},
$$

(40)

It is unlikely that ‘external’ templates would be of any value for $B$-mode analysis.
where $S_j$, $\Phi_j$, and $N_j$ are, respectively, the primordial CMB, residual foreground and noise covariance matrices. Now construct the data vector $Y_i = x_i - F_i \beta_i$, \[ \{ \beta_i \} = \{ \beta_{(Q,U)} \} \text{ if } i = (Q,U), \] where the superscript denotes frequency. If the template subtraction removes the foregrounds, the average of (41) over noise-realizations is \[ \langle Y_i \rangle = x_i \left( 1 - \sum_k \beta_i^k \right), \] and if the coefficients $\beta$ are independent of the signal the covariance matrix $\{ Y, Y \}$ is \[ \langle Y, Y \rangle = S_{ij} \left( 1 - \sum_k \beta_i^k \right) \left( 1 - \sum_k \beta_j^k \right) + N_{ij} + N_i^j \beta_i^k \beta_j^k. \] The solution for $\beta$ is found by iteratively minimizing (27) with $C$ replaced by $\langle Y, Y \rangle$ and ignoring any weak correlation between the solution and the signal. The final data vector $Y$ and its covariance matrix (43) are then used to compute the likelihood function (15). The parameters $\beta$ are well constrained by the data and so it is a good approximation to keep them fixed at their central values. The main contribution of the $\beta$ to the error budget is via the noise term (43).

In the simulations described here, we construct the data vector $x$ from the four frequency channels 70, 100, 143, 217 GHz, since there is little additional sensitivity to primary CMB signal in the other channels. We use the 30- and 353-GHz channels as templates. The internal mask described in Section 2 is applied to all channels. The resulting noisy foreground subtracted maps are shown in Fig. 8. The upper panel of this figure shows the noise-free CMB simulations for $r = 0.1$ for the regions that lie outside the internal mask. The second panel shows the reconstruction after foreground subtraction from the noisy maps following the procedure described above. There is clearly a very good correspondence between the two sets of maps. The third panel in Fig. 8 shows the noise-free contribution of the $B$-mode to the $Q$ and $U$ maps. The foreground-subtracted reconstruction is shown in the lowest panel. Again, there is a good correspondence between the maps, but the reconstructed maps are very noisy. In fact, instrument noise dominates over foreground mismatch. A substantial component of the noise comes from the templates because the 30- and 353-GHz channels of Planck are significantly noisier than the main ‘CMB’ channels at $\sim 100$ GHz.

The likelihood functions for $r$ are shown in Fig. 9. The dotted (red) lines show the likelihood functions applied to the noise-free CMB maps (though with diagonal ‘regularizing’ noise applied, as described in Section 3) for the two simulations with $r = 0$ and 0.1 for the regions outside the internal mask. These likelihoods are close to the ‘best’ that could be achieved from a low-resolution experiment in the absence of foreground contamination. The results from our noisy foreground subtracted simulations are shown by the solid (blue) lines. The distribution for the model with $r = 0$ is peaked close to $r = 0$, so clearly residual foreground mismatch is unimportant. The increased widths of the blue curves are caused by residual instrument noise, including the noise in the template channels. All of the results described in this paper assume a nominal mission lifetime of 14 months for Planck. This ensures that every detector on Planck covers the sky twice (see SPP05). We show in a separate paper (Efstathiou & Gratton 2009) that an extended mission lifetime for Planck to four full sky surveys leads to a significant improvement in $B$-mode sensitivity.

Finally, Fig. 10 shows the QML power spectra (equations 16–20) for the foreground subtracted noisy realizations with $r = 0$ and 0.1. As discussed in Efstathiou (2006), the QML estimator eliminates mixing of $E$ and $B$-modes at low multipoles almost perfectly on a cut-sky, and so the $B$-mode spectrum for the $r = 0$ realization is indeed close to zero at low multipoles. However, this figure shows clearly that for the Planck noise levels, there is little information in the $B$-mode spectrum at multipoles $\ell \gtrsim 10$.

## 5 Comments on Foreground Removal Techniques

As summarized in the Introduction, a large number of diffuse foreground subtraction techniques have been discussed in the literature. Some of these are designed to recover ‘physical’ foregrounds, e.g. separating free–free from synchrotron emission. Other techniques are designed to tackle the problem described in this paper, i.e. the recovery of the primordial CMB anisotropy from foreground-contaminated maps. The remarks in this section apply to this latter class of techniques.

In the preceding sections, we identified two distinct forms of error, which we termed ‘cross-correlation offset’ and ‘foreground mismatch’. These two types of error provide an intuitively useful way of classifying component separation methods. Table 3 summarizes our proposed classification scheme. Any foreground subtraction technique can be placed somewhere between the blind and unblind rows of this table.

| Classification | Remarks |
|----------------|---------|
| **Blind**      |         |
| **Unblind**    |         |

Briefly, a purely blind technique such as ILC can, in theory, reduce foreground mismatch to negligible levels provided there are enough frequency bands to describe the foregrounds. (This is the case, for example, for the weights listed in the first three rows of Table 2.) However, in a purely blind technique, there is no external information to distinguish between CMB and foreground components with similar structure on the sky. The result is a cross-correlation offset that is independent of the amplitude of the foregrounds. Any purely blind technique (e.g. harmonic ILC) will show a cross-correlation offset. Unless one can isolate pure $B$-modes on a cut sky (see e.g. Lewis 2003), cross-correlation between the CMB $E$-modes and the foregrounds can produce a potentially serious cross-correlation offset.

The amplitude of the cross-correlation offset can be reduced if some additional information is provided. We have classified template matching as a ‘semiblind’ technique because it makes use of some prior information, though it is not based on a physical model of the foreground.
foregrounds. The method requires a model for the signal (primordial CMB) covariance matrix and the templates provide a model for the angular distribution of the foregrounds (spectral index variations can be taken into account by adding more templates). As shown in the previous section it is possible to reduce both cross-correlation offset and foreground mismatch to negligible levels.

A third class of technique attempts to model the foregrounds by fitting a parametric physical model (Brandt et al. 1994; Eriksen et al. 2006; Dunkley et al. 2008b). If the physical model is a correct representation of the truth, it is possible to reduce both the cross-correlation offset and foreground mismatch to negligible levels. However, this type of technique is limited, in practice, by the number of frequency channels available. The number of independent parameters describing the model must be less than or equal to the number of frequency channels. For Planck polarization, this limits the number of independent free parameters to be ≤7 (if the Q and U models are treated independently). This limits the complexity of the physical model, limiting the scope for redundancy checks. Of course, if the model is incorrect the method will be limited by foreground mismatch.

It is also useful to consider how foreground subtraction techniques are affected by instrumental noise. Instrumental noise in a purely blind technique is, in a sense, ‘uncontrollable’. For example, the ‘ideal’ weights listed in the first three rows of Table 2 remove foregrounds to high precision. However, if they were applied to noisy data they would amplify the instrumental noise to high levels (because many of the weights exceed unity). For Planck polarization, the resulting noise amplification would be unacceptable. We have shown in Section 4.2 that

Figure 8. Q and U maps: upper panel shows the noise-free CMB simulations with $r = 0.1$ for regions outside the internal mask. The second panel shows the foreground subtracted noisy reconstructions computed as described in the text. The third panel shows the $B$-mode contribution to the noise-free CMB maps. The lowest panel shows the noisy foreground subtracted reconstruction of this $B$-mode contribution.
in the template matching approach, instrumental noise is ‘controllable’ provided the templates have high signal-to-noise ratio. Instrumental noise is a major problem for model fitting techniques. As far as we are aware, nobody has yet developed a model-fitting technique that incorporates prior information on the angular variation of the spectral indices of the diffuse foregrounds (which vary slowly over the sky). As a proxy, model fitting is usually done by independently fitting parameters in a very coarsely pixelized map. This reduces the effects of instrument noise on the estimated parameters, but even then the effects of noise can be limiting. For example, in the analysis of the WMAP 5-yr polarization data at a resolution of NSIDE = 8 the synchrotron spectral index was computed at a resolution of NSIDE = 2 (48 pixels over the whole sky) in Dunkley et al. (2008a). Fairly strong (though not unreasonable) priors were imposed to find convergent solutions (e.g. the dust spectral index, which is poorly constrained by the data, was kept fixed). Nevertheless, Dunkley et al.’s results for the $E$-mode power spectrum at low multipoles compare well with those from the template cleaned maps of Gold et al. (2009). It remains to be seen whether model fitting can perform well for the more difficult problem of $B$-mode recovery for Planck. We hope to report on this in a future paper.

6 CONCLUSIONS

In this paper, we have used the PSM to assess the impact of foregrounds on $B$-mode detection by Planck at low multipoles. We have analysed the internal linear combination technique and shown that the offset caused by $E$-mode polarization pattern (cross-correlation offset) leads to a fundamental limit of $r \approx 0.1$ for the tensor–scalar ratio even in the absence of instrumental noise. This is comparable to the sensitivity limit of Planck if foregrounds are neglected. For realistic Planck instrument noise, ILC amplifies the noise of the ‘cleaned’ polarization maps to unacceptably high levels. Our results show that ‘blind’ techniques such as ILC are unsuitable for detecting primordial $B$-modes from a future low-noise ‘CMBpol’ mission.

We have analysed template fitting, using internal templates constructed from the Planck data and devised a scheme to approximate the likelihood function (14) from multifrequency maps. We have shown that this scheme works well for Planck and offers a feasible way of recovering primordial $B$-modes from dominant foreground contamination even in the presence of noise. According to the results shown in Fig. 9, Planck, after the nominal mission lifetime of 14 months, could set a useful upper limit of $r \lesssim 0.05$ if there is no primordial tensor mode and may even detect a tensor mode if $r \sim 0.1$. This provides a useful complement to ground-based/suborbital experiments which cannot probe these low multipoles ($\ell \lesssim 10$). These limits probe an interesting part of parameter space (see Efstathiou & Congchitnan 2006, for a review). For inflation with a power-law potential, $\phi^\alpha$, the scalar spectral index and tensor–scalar ratio are approximately

$$n_s \approx 1 - \frac{2 + \alpha}{2N}, \quad r \approx \frac{4\alpha}{N}, \quad \text{i.e.} \quad r \approx 8(1 - n_s) \frac{\alpha}{\alpha + 2},$$

where $N$ is the number of inflationary e-folds between the time that CMB scales crossed the ‘horizon’ and the time that inflation ends. There are indications from WMAP and CMB experiments probing higher multipoles for a small tilt of the spectral index $n_s \sim 0.97$ (Komatsu et al. 2009; Reichardt et al. 2009). If this tilt is correct, then the last of these equations suggests $r \sim 0.1$ for any $\alpha$ of order unity. For example, for $N \approx 60$ (Liddle & Leach 2003), the quadratic potential (still allowed by the data) gives $n_s \approx 0.97$, $r \approx 0.13$, within the parameter range.

The significance of this tilt depends quite sensitively on the complexity of the model (and assumed priors), for example on whether or not a tensor mode or a run in the spectral index are included in the model.
Figure 10. QML estimates of the E- and B-mode polarization spectra for the two noisy foreground subtracted simulations used to compute the likelihoods of Fig. 9. The upper panel shows the power spectra for the realization with $r = 0$ and the lower panel shows the power spectra for the realization with $r = 0.1$. The error bars show the diagonal components of (20) using the theoretical input values of $r$ for each realization. The solid lines show the theoretical input spectra for $r = 0.1$.

Table 3. Classification scheme of foreground removal techniques.

| Scheme                | Cross-correlation offset | Foreground mismatch                        |
|-----------------------|--------------------------|--------------------------------------------|
| Blind (e.g. ILC)      | Significant              | Small (given enough frequency bands)       |
| Semiblind (e.g. template fitting) | Small               | Small (given enough templates)            |
| Unblind (e.g. model fitting)     | Small (if model is correct)     | Small (if model is correct)               |

accessible to Planck. Failure to detect a B-mode at $r \gtrsim 10^{-2}$ would put pressure on ‘high-field’ ($\phi$ of order Planck scale) inflation models, in which there has been recent renewed interest (Silverstein & Westphal 2008).

The simplicity of the PSM is a source of concern. Following WMAP, there is quite a lot of information available on the polarized synchrotron emission. The low-frequency channels on Planck will provide additional information at $\lesssim 70$ GHz. Thus, there is considerable scope for redundancy checks at low frequencies, for example, by varying templates and by parameter fitting. The dust contribution to polarization is much more uncertain. Neither the level of polarization nor dust spectral index variations are well constrained by current data. We will almost certainly have to wait until Planck flies to assess whether polarized dust emission poses a serious problem for B-mode analysis. Planck is heavily reliant on the 353-GHz channel to model dust polarization, because it has the highest signal-to-noise ratio on the foreground. There is some limited scope for redundancy checks using the 217-GHz channel. However, if polarized dust emission is complex, it may not be possible to achieve a limit of $r \sim 0.1$ with Planck using template fitting or any other foreground removal technique.
ACKNOWLEDGMENTS

GE and SG thank STFC for financial support. FP thanks the Institute of Astronomy, Cambridge, for its warm hospitality during visits related to this work. The authors acknowledge the use of the PSM developed by the Component Separation Working group of the Planck Collaboration, and of the HEALPix package. We thank Anthony Challinor, Jacques Delabrouille, Jo Dunkley, Antony Lewis, Hiranya Peiris and the BICEP collaboration for useful discussions.

REFERENCES

Amblard A., Cooray A., Kaplinghat M., 2007, Phys. Rev. D, 75, 083508
Baumann D. et al., 2008, preprint (arXiv:0811.3919)
Bennett C. L. et al., 2003, ApJS, 148, 1
Betoule M., Pierpaoli E., Delabrouille J., Le Jeune M., Cardoso J.-F., 2009, A&A, submitted (arXiv:0901.1056)
Bok J. et al., 2008, preprint (arXiv:0805.4207)
Brandt W. N., Lawrence C. R., Readhead A. C. S., Pakianathan J. N., Fiola T. M., 1994, ApJ, 424, 1
Chiang L.-Y., Naselsky P. D., Coles P., 2008, preprint (arXiv:0805.4207)
Crill B. P. et al., 2008, in Oschmann J. M., de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham, p. 70102P
de Bernardis P., Bucher M., Burigana C., Piccirillo L., 2009, Exp. Astron., 23, 5
Delabrouille J., Cardoso J.-F., 2007, preprint (astro-ph/0702198)
Delabrouille J., Cardoso J.-F., Patanchon G., 2003, MNRAS, 346, 1089
Dunkley J. et al., 2008a, preprint (arXiv:0811.4280)
Dunkley J. et al., 2008b, preprint (arXiv:0811.3915)
Efstathiou G., 2006, MNRAS, 370, 343
Efstathiou G., 2007, MNRAS, 380, 1621
Efstathiou G., Congchitnan S., 2006, Prog. Theor. Phys., 163, 204
Efstathiou G., Gratton S., 2009, J. Cosmol. Astropart. Phys., submitted (arXiv:0903.0345)
Eriksen H. K. et al., 2004a, ApJS, 155, 227
Eriksen H. K., Banday A. J., Gorski K. M., Lilje P. B., 2004b, ApJ, 612, 633
Eriksen H. K. et al., 2006, ApJ, 641, 665
Eriksen H. K., Jewell J. B., Dickinson K., Banday A. J., Gorski K. M., Lawrence C. R., 2008, ApJ, 676, 10
Finkbeiner D. P., Davis M., Schlegel D., 1999, ApJ, 524, 867
Gold G. B. et al., 2009, ApJS, 180, 265
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartlemann M., 2005, ApJ, 622, 759
Gratton S., 2008, preprint (arXiv:0805.0093)
Hansen F. K., Banday A. J., Eriksen H. K., Gorski K. M., Lilje P. B., 2006, ApJ, 648, 784
Hyvärinen A., 1999, IEEE Signal Process. Lett., 6, 145
Kamionkowski M., Kosowsky A., Stebbins A., 1997, Phys. Rev. D, 55, 7368
Komatsu E. et al., 2009, ApJS, 180, 330
Leach S. M. et al., 2008, A&A, 491, 597
Lewis A., 2003, Phys. Rev. D, 68, 3509
Lewis A., Challinor A., 2006, Phys. Rev., 429, 1
Liddle A. R., Leach S. M., 2003, Phys. Rev. D, 68, 103503
Lyth D. H., 1984, Phys. Lett. B, 147, 403
Maino D. et al., 2002, MNRAS, 334, 53
Miville-Deschênes M., Ysard N., Lavabre A., Ponthieu N., Macias-Perez J. F., Aumont J., Bernard J. P., 2008, A&A, in press (arXiv:0802.3345)
Norgaard-Nielsen H. U., Jorgensen H. E., 2008, Ap&SS, 318, 195
North C. E. et al., 2008, preprint (arXiv:0805.3690)
Oxley P. et al., 2004, Proc. SPIE – Int. Soc. Opt. Eng., 5543, 320
Pérez H. et al., 2003, ApJS, 148, 213
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in Fortran 77, 2nd edn. Cambridge Univ. Press, Cambridge
Reichardt C. L. et al., 2009, ApJ, 694, 1200
Rybicki G. B., Press W. H., 1992, ApJ, 398, 169
Seiffert M. D. et al., 2006, Am. Astron. Soc. Meeting, 209, 2006
Silverstein E., Westphal A., 2008, Phys. Rev. D, 78, 6003
Slosar A., Seljak U., 2004, Phys. Rev. D, 70, 3002
Slosar A., Seljak U., Makarov A., 2004, Phys. Rev. D, 69, 123003
Spergel D. N. et al., 2007, ApJS, 170, 377
Stolyarov V., Hobson M. P., Ashdown M. A. J., Lasenby A. N., 2002, MNRAS, 336, 97
Tegmark M., de Oliveira-Costa A., 2001, Phys. Rev. D, 64, 063001
Tegmark M., de Oliveira-Costa A., Hamilton J. A., 2003, Phys. Rev. D, 68, 123523
The Planck Collaboration, 2005, in Efstathiou G., Lawrence C., Tauber J., eds, The Scientific Programme of Planck, ESA-SCI(2005). ESA, Noordwijk
Tucci M., Martínez-González E., Vielva P., Delabrouille J., 2005, MNRAS, 360, 935
Varshalovich D. A., Moskalev A. N., Khersonskii V. K., 1988, Quantum Theory of Angular Momentum. World Scientific Press, Singapore
Yoon K. W. et al., 2006, Proc. SPIE, 6275, 51
Zaldarriaga M., Seljak U., 1997, Phys. Rev. D, 55, 1830

© 2009 The Authors. Journal compilation © 2009 RAS, MNRAS 397, 1355–1373
APPENDIX A: PIXEL NOISE COVARIANCES OF DEGRADED RESOLUTION MAPS IN THE SMALL ANGLE LIMIT

A1 Temperature

Let $x_i$ denote the pixel value in the high-resolution map and $x_j$ denote the pixel value in the low-resolution map. The harmonic coefficients computed from the high-resolution map is

$$a_{lm} = \sum_j x_i \Omega_i Y_{lm}(\theta_i),$$

where $\Omega_i$ is the solid angle of a high-resolution map pixel. So, the pixel values in the degraded map are

$$X_i = \sum_{lm} a_{lm} Y_{lm}(\theta_i)f_i(\theta_i).$$

Using the addition theorem for spherical harmonics, the pixel noise covariance of the low-resolution map is

$$\langle X_i X_j \rangle = \sum_{l \leq 2} \sum_{pq} (x_i x_j) \frac{\ell(\ell + 1)}{4\pi} \frac{(2\ell + 1)}{4\pi} \Omega_i \Omega_j P_{\ell,i}(\cos \theta_i)P_{\ell,j}(\cos \theta_j)f_i(\theta_i)f_j(\theta_j).$$

This expression is time consuming to evaluate, but it simplifies significantly if the noise is diagonal $\langle x_i x_j \rangle = \delta_{ij}\sigma^2$ and if we assume small angles. In this case, for a Gaussian smoothing function, equation (A4) simplifies to

$$\langle X_i X_j \rangle \approx \frac{1}{\theta^2} \sum_p \sigma^2_p \Omega^2_p \exp \left(-\frac{\theta^2_p}{\theta^2}\right) \exp \left(-\frac{\theta^2_p}{\theta^2}\right).$$

A2 Polarization

In the case of diagonal pixel noise at high resolution, and for small angles, the polarization covariance matrices $\langle Q_i Q_j \rangle$ for the degraded resolution maps can be approximated by equation (A5). It is interesting to see why this is so. We will consider degraded resolution $Q$ maps (the analysis is identical for $U$ maps). The equivalent to (A3) in obvious notation is

$$Q_i = \frac{1}{2} \sum_{m} \sum_p q_p \left[ 2Y^2(p) 2Y(i) + -2Y^2(p) -2Y(i) \right] + iu_p \left[ 2Y^2(p) 2Y(i) - -2Y^2(p) +2Y(i) \right] \Omega_p f_i.$$  

Now the addition theorem for the tensorial harmonics is

$$\sum_{m} x_i Y_{m-m}(\theta_1, \phi_1) x_j Y_{m-m}(\theta_2, \phi_2) = (-1)^{i} \frac{2\ell + 1}{4\pi} \frac{\ell(\ell + 1)}{4\pi} \Omega_i \Omega_j \left[ 2Y^2(p) 2Y(i) + -2Y^2(p) -2Y(i) \right] \Omega_p f_i.$$  

where we use the Euler angle conventions of Varshalovich, Moskalev & Khersonskii (1988). Applying the addition theorem to (A6) gives

$$Q_i = \frac{1}{2} \left( \frac{2\ell + 1}{4\pi} \right)^{1/2} \left( q_p \left[ 2Y^2(p) 2Y(i) + -2Y^2(p) -2Y(i) \right] + iu_p \left[ 2Y^2(p) 2Y(i) - -2Y^2(p) +2Y(i) \right] \right) \Omega_p f_i.$$  

Now we can write

$$2Y_{lm} = 2\sqrt{N_i} A^e_{\ell m}(G^+ + G^-) e^{i\phi}.$$  

(see Kamionkowski et al. 1997) and so

$$Q_i = \frac{1}{2} \left( \frac{2\ell + 1}{4\pi} \right)^{1/2} \left[ 2\sqrt{N_i} A^e_{\ell m}(G^+ + G^-) \Omega_p f_i \left[ q_p \cos(2\alpha + 2\gamma) - u_p \sin(2\alpha + 2\gamma) \right] \right].$$

In the limit $\ell \to \infty$, the pre-factor in (A11), tends to

$$\left( \frac{2\ell + 1}{4\pi} \right)^{1/2} \sqrt{2N_i} A^e_{\ell m} \left( \frac{2\ell + 1}{2\pi} \right)^{1/2} \frac{1}{\ell^2}$$

and

$$G^e_{\ell m}(s) \to \frac{1}{4} \ell [J_0(s) \pm J_4(s)],$$

and so (A11) becomes

$$Q_i \approx \sum_{\ell p} \left( \frac{2\ell + 1}{4\pi} \right) \left[ J_0((2\ell + 1) \sin \beta/2) \Omega_p f_i \left[ q_p \cos(2\alpha + 2\gamma) - u_p \sin(2\alpha + 2\gamma) \right] \right].$$
Summing over $\ell$,

$$Q_i \approx \frac{\Omega_p}{\theta_p^2} \frac{1}{2\pi} \sum_{p} \left[ q_p \cos(2\alpha + 2\gamma) - u_p \sin(2\alpha + 2\gamma) \right] \exp \left( -\frac{\theta_p^2}{2\theta_i^2} \right)$$  \hspace{1cm} (A15)

for Gaussian smoothing. Thus, if the $q_p$ and $u_p$ are uncorrelated between pixels and are uncorrelated with each other, and for small angular separations, it is possible to show after some algebra that equation (A15) gives exactly the same covariance matrix as the scalar result of equation (A5):

$$\langle Q_i Q_j \rangle \approx \frac{1}{\theta_i^4} \frac{1}{2\pi^2} \sum_{p} \sigma_p^2 \Omega_p^2 \exp \left( -\frac{\theta_p^2}{2\theta_i^2} \right) \exp \left( -\frac{\theta_p^2}{2\theta_i^2} \right).$$  \hspace{1cm} (A16)

An analogous derivation applies for the covariance matrix $\langle U_i U_j \rangle$ (note that $\langle Q_i U_j \rangle \approx 0$). By comparing with numerical simulations, we find that the scalar approximation (A16) is an excellent approximation and is perfectly adequate for the smoothing scales adopted in this paper.