A comparison of CMB angular power spectrum estimators at large scales: the TT case

D. Molinari,1,2★ A. Gruppuso,2,3 G. Polenta,4,5 C. Burigana,2,6 A. De Rosa,2 P. Natoli,2,4,6,7 F. Finelli2,3 and F. Paci8

1Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy
2INAF-IASF Bologna, Via Piero Gobetti 101, I-40129 Bologna, Italy
3INFN, Sezione di Bologna, Via Irnerio 46, I-40126 Bologna, Italy
4Agenzia Spaziale Italiana Science Data Center, c/o ESRIN, via Galileo Galilei, I-00044 Frascati, Italy
5INAF – Osservatorio Astronomico di Roma, via di Frascati 33, I-00040 Monte Porzio Catone, Italy
6Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, via Giuseppe Saragat 1, I-44100 Ferrara, Italy
7INFN, Sezione di Ferrara, via Giuseppe Saragat 1, I-44100 Ferrara, Italy
8SISSA – Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, I-34136 Trieste, Italy

ABSTRACT
In the context of cosmic microwave background (CMB) data analysis, we compare the efficiency at large scale of two angular power spectrum algorithms, implementing, respectively, the quadratic maximum likelihood (QML) estimator and the pseudo-spectrum (pseudo-$C_\ell$) estimator. By exploiting 1000 realistic Monte Carlo simulations, we find that the QML approach is markedly superior in the range $2 \leq \ell \leq 100$. At the largest angular scales, e.g. $\ell \leq 10$, the variance of the QML is almost $1/3$ ($1/2$) that of the pseudo-$C_\ell$, when we consider the WMAP kq85 (kq85 enlarged by $8^\circ$) mask, making the pseudo-spectrum estimator a very poor option. Even at multipoles $20 \leq \ell \leq 60$, where pseudo-$C_\ell$ methods are traditionally used to feed the CMB likelihood algorithms, we find an efficiency loss of about 20 per cent, when we considered the WMAP kq85 mask, and of about 15 per cent for the kq85 mask enlarged by $8^\circ$. This should be taken into account when claiming accurate results based on pseudo-$C_\ell$ methods. Some examples concerning typical large-scale estimators are provided.

Key words: methods: data analysis – methods: numerical – methods: statistical – cosmic background radiation – cosmology: observations – cosmology: theory.

1 INTRODUCTION

The pattern of the cosmic microwave background (CMB) anisotropy field can be used to probe cosmology to high precision, as shown by the Wilkinson Microwave Anisotropy Probe (WMAP) 9 yr results (Hinshaw et al. 2012) and by the very recent Planck cosmological results (see Planck Collaboration I 2013 and references therein). CMB data have given a significant contribution in setting up the Λ cold dark matter ($\Lambda$CDM) cosmological concordance model. The latter establishes a set of basic quantities for which CMB observations and other cosmological and astrophysical data sets agree:1 spatial curvature close to zero; $\sim$68.5 per cent of the cosmic density in the form of dark energy; $\sim$26.5 per cent in CDM; $\sim$5 per cent in baryonic matter; and non-perfectly scale invariant adiabatic, primordial perturbations compatible with Gaussianity (Planck Collaboration XVI 2013; Planck Collaboration XXIV 2013).

In particular, the largest scales of the temperature anisotropies map are of great interest because they directly probe the early Universe (or Inflationary Phase of the Universe; Starobinsky 1980; Guth 1981; Linde 1982; Albrecht & Steinhardt 1982). They correspond to angular scales larger than the horizon size at decoupling as observed today, i.e. $\theta > 2^\circ$ or, equivalently $\ell < \ell_{\text{dec}} \sim 90$ (see for example Page et al. 2003) with $\ell$ being the multipole order of the spherical harmonics expansion

$$\delta T(\hat{\theta}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{\theta}), \quad (1)$$

where $\delta T = T(\hat{\theta}) - T_0$ is the temperature anisotropy observed in the direction $\hat{\theta}$ relative to the CMB average temperature $T_0 \simeq 2.725$ K (Mather et al. 1999; see also Fixsen et al. 1996 for a constraint of the CMB blackbody shape), and with $a_{\ell m}$ being the coefficients of the spherical harmonics $Y_{\ell m}(\hat{\theta})$.

★E-mail: diegomolinari@inwind.it
1 See, however, Planck Collaboration XX (2013) for a possible tension concerning the $G_\Delta$ estimate from Planck CMB and galaxy clusters data.
The main contribution to the CMB anisotropies at these large scales is provided by the so-called Sachs–Wolfe effect and by a subdominant integrated Sachs–Wolfe effect (Sachs & Wolfe 1967) which is different from zero because of the recent (from a cosmological point of view) transition to an accelerated phase of the Universe (Kofman & Starobinsky 1985) likely associated with a dark energy (or cosmological constant) component. In principle, a stochastic background of primordial gravitational waves can also give a contribution to the temperature CMB anisotropies at these largest scales, depending on the tensor-to-scalar ratio, \( r \), constrained by the current data (Hinshaw et al. 2012; Planck Collaboration XVI 2013). A firm detection of the primordial gravitational waves requires CMB B modes polarization measurement at large scale (Knox & Turner 1994).

From the observational point of view, the CMB anisotropies temperature map, as observed by WMAP 9 yr, is cosmic variance dominated, i.e. cosmic variance exceeds the instrument noise, up to \( \ell = 946 \), see Bennett et al. (2013). For Planck data, this crossing happens at \( \ell \approx 1500 \) (Planck Collaboration XV 2013). Therefore, at the largest scales (\( \ell < \ell_{\text{dec}} \approx 90 \)), the effect of instrumental noise is almost negligible. Keeping this in mind, it is even more important to employ the most accurate data analysis tools.

In the current paper, we focus on the angular power spectrum (APS), which is the main observable for the diagnosis of the CMB map. The method that is capable of providing APS with no bias and with the minimum variance, as provided by the Fisher–Cramer–Rao inequality is the quadratic maximum likelihood (QML) method (Tegmark 1997; Tegmark & de Oliveira-Costa 2001). Such an optimal method has the drawback of being computationally expensive and then limited by the number of pixels. It is currently implemented and applied at low resolution (see e.g. Gruppuso et al. 2009). Several other strategies for measuring \( C_\ell \) at low resolution have been developed and applied to CMB data with excellent results. These methods include different sampling techniques such as Gibbs (Eriksen et al. 2004; Jewell, Levin & Anderson 2004; Wandelt, Larson & Lakshminarayanan 2004), adaptive importance (Benabed et al. 2009) and Hamiltonian (Taylor, Ashdown & Hobson 2008). At high multipoles (\( \ell > 30 \), Efstathiou 2004a), the so-called pseudo-\( C_\ell \) algorithms are usually preferred to others techniques. These methods, in fact, implement the estimation of power spectral densities from periodograms (Hauser & Peebles 1973). Basically, they estimate the \( C_\ell \) through the inverse harmonic transform of a masked map that is then deconvolved with geometrical kernels and corrected with a noise bias term. These techniques, such as Master (Hivon et al. 2002), cross-spectrum (Polenta et al. 2005; Saha, Jain & Souradeep 2006; Grain, Tristram & Stompor 2009), give unbiased estimates of the CMB power spectra and moreover, it has been shown that they work successfully when applied to real data at high multipoles (Kuo et al. 2004; Jones et al. 2006; Wu et al. 2007; Dunkley et al. 2009). These estimators are pretty quick and light from a computational point of view. However, it is well known that at low multipoles they are not optimal since they provide power spectra estimates with error bars larger than the minimum variance. We note that in Efstathiou (2004a), a hybrid approach has already been proposed, i.e. the QML at low and the pseudo-\( C_\ell \) at high multipoles, where a recipe for a hybrid covariance is consistently given. However, in that paper the QML is considered only up to \( \ell \approx 40 \) since it is applied to maps with a full width at half-maximum (FWHM) = 3°.

The aim of this work is to compare quantitatively the QML and pseudo-\( C_\ell \) methods under realistic assumptions focusing on the largest scales (i.e. low multipoles) where CMB anomalies are mainly located (Bennett et al. 2011; Planck Collaboration XXIII 2013). The idea is to provide the scientific community with a reference analysis comparing the heavy QML method and the light and quick pseudo-\( C_\ell \) approach in the range of multipoles \( \ell < \ell_{\text{dec}} \approx 90 \). Here, we quantitatively address this problem through realistic Monte Carlo (MC) simulations. A similar analysis has been carried out by WMAP team using the so-called \( C^{-1} \) method employed in Bennett et al. (2013) that shows an improvement with respect to the pseudo-\( C_\ell \) method mostly located at very low multipoles and in the intermediate S/N regime. Moreover, we evaluate the benefit of considering an optimal APS estimation taking into account some examples of estimators of anomalies that are commonly used at large scales to test the consistency of the observations with the \( \Lambda \)CDM model.

The paper is organized as follows. In Section 2, we describe the considered implementations of QML and pseudo-\( C_\ell \), codes, called respectively BOLPOL and cROMASTER. In Section 3, the computational requirements for both the methods are given. Section 4 is devoted to the detailed description of the comparison and corresponding results. Examples on how the previous analysis propagates to the APS based estimators, commonly used in the literature for the analysis of the large-scale anomalies, are provided in Section 4.4. Conclusions are drawn in Section 5.

2 DESCRIPTION OF THE CODES

In this section, we give a description of the two considered APS estimators. Specifically, we consider cROMASTER, as an implementation of pseudo-\( C_\ell \) method and BOLPOL as implementation of the QML estimator.

2.1 BOLPOL

In order to evaluate the APS of full sky and masked sky maps BOLPOL code adopts the QML estimator, introduced in Tegmark (1997). In this section, we describe the essence of this method. Given a CMB temperature map \( \mathbf{x} \), the QML provides estimates \( \hat{C}_\ell \) of the TT APS as

\[
\hat{C}_\ell = \left( \frac{1}{2\ell+1} \right) F_{\ell\ell} \left[ \mathbf{x}^\dagger \mathbf{E}^\nu \mathbf{x} - \text{tr}(\mathbf{N} \mathbf{E}^\nu) \right],
\]

where \( F_{\ell\ell} \) is the Fisher matrix, defined as

\[
F_{\ell\ell} = \frac{1}{2} \text{tr} \left( C^{-1} \frac{\partial C}{\partial \hat{C}_\ell} C^{-1} \frac{\partial C}{\partial \hat{C}_\ell} \right),
\]

and the \( \mathbf{E}^\nu \) matrix is given by

\[
\mathbf{E}^\nu = \frac{1}{2} C^{-1} \frac{\partial C}{\partial \hat{C}_\ell} C^{-1},
\]

with \( \mathbf{C} = \mathbf{S}(\hat{C}_\ell) + \mathbf{N} \) being the global covariance matrix (signal plus noise contribution).

Although an initial assumption for a fiducial power spectrum \( \hat{C}_\ell \) is needed in order to build the signal covariance matrix \( \mathbf{S}(\hat{C}_\ell) \), it has been proven that the QML method provides unbiased estimates of the power spectrum contained in the map regardless of the initial guess (Tegmark 1997),

\[
\langle \hat{C}_\ell \rangle = \mathcal{T}_\ell,
\]

where the average is taken over the ensemble of realizations (or, in a practical test, over MC realizations extracted from \( \mathcal{T}_\ell \)). On the other hand, the covariance matrix associated with the estimates is given by the inverse of the Fisher matrix

\[
\langle \Delta \hat{C}_\ell \Delta \hat{C}_{\ell'} \rangle = (F^{-1})_{\ell\ell'},
\]
only when the fiducial spectrum $C_\ell = \tilde C_\ell$. In this case the inverse of the Fisher matrix provides the minimum variance. According to the Cramer–Rao inequality, which sets a limit to the accuracy of an estimator, equation (6) tells us that the QML has the smallest error bars. The QML is then an optimal estimator because equations (5) and (6) are both satisfied.

For a generalization to polarization spectra see Tegmark & de Oliveira-Costa (2001).

### 2.2 cROMASTER

cROMASTER is an implementation of the pseudo-$C_\ell$ method proposed by Hivon et al. (2002), extended to allow for both auto- and cross-power spectrum estimation (see Polenta et al., 2005, for a comparison). A given sky map $x$ can be decomposed in CMB signal and noise as $x = s + n$, and we define the masked sky pseudo-spectrum as

$$\tilde C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |\tilde a_{\ell m}|^2 = \hat C_\ell^2 + \tilde N_\ell,$$

(7)

where $\hat C_\ell^2$ is the pseudo-spectrum of the sky signal, $\tilde N_\ell$ is the pseudo-spectrum of the noise present in the map, and the pseudo-$a_{\ell m}$ coefficients are computed as

$$\tilde a_{\ell m} = \int d\Omega \psi(\theta, \phi) y_{\ell m}^* (\theta, \phi),$$

(8)

where $\psi$ is the applied mask.

In order to recover the full sky power spectrum, we employ the following estimator:

$$\hat C_\ell = K_{\ell \ell}^{-1} (\tilde C_\ell - \tilde N_\ell),$$

(9)

where the mode–mode coupling kernel $K_{\ell \ell}^{-1}$ is a geometrical correction that accounts for the loss of orthonormality of the spherical harmonic functions in the cut sky (see Hivon et al. 2002, for more details). In fact, for an isotropic sky signal that is a realization of a theoretical power spectrum $C_\ell$, one can write:

$$\langle \hat C_\ell \rangle = C_\ell.$$

(10)

Hence, equation (9) defines an unbiased estimator, provided that the noise term $\tilde N_\ell$ is properly removed, which is usually done through MC simulations. However, for current generation high-sensitivity experiments, such as WMAP and Planck, the noise bias has to be known to better than 0.1 per cent accuracy, and therefore, the cross-spectrum approach is preferred (see, e.g. Bennett et al. 2013; Planck Collaboration XV 2013). In the latter case, the pseudo-$C_\ell$ are computed by combining the spherical harmonic coefficients from two or more maps with uncorrelated noise so that the ensemble average of $\tilde N_\ell$ is null, making the cross-spectrum estimator naturally unbiased.

For practical applications, i.e. involving coloured noise properties and arbitrary masks incorporating possible non-uniform weighting of the data, there is no readily available exact analytical expression for the covariance matrix of this estimator. The error bars for pseudo-$C_\ell$ methods can be roughly approximated as

$$\Delta \hat C_\ell = \sqrt{2 \over (2\ell + 1) f_{\text{sky}}^\text{eff}} (C_\ell + N_\ell),$$

(11)

where $f_{\text{sky}}^\text{eff}$ is the effective fraction of the sky used for the analysis which accounts for the weighting scheme of the pixels. For a given sky coverage, a uniform weighting scheme produces the largest $f_{\text{sky}}^\text{eff} = 1$, and is therefore the optimal choice in the signal dominated regime, which corresponds to large angular scales. On the other hand, an inverse noise weighting scheme reduces $N_\ell$ and is the best choice in the noise dominated regime, i.e. at small angular scales. A trade-off between the two weighting scheme should be applied in the intermediate regime, and an optimal pseudo-$C_\ell$ method should employ the best weighting scheme at a given multipole according to the S/N ratio at that multipole.

In order to derive an accurate estimate of the covariance matrix, the original approach proposed in Hivon et al. (2002) is to rely on signal and noise MC simulations:

$$\text{Cov} \{ \hat C_\ell, \hat C_{\ell'} \} = \langle (\hat C_\ell - \langle \hat C_\ell \rangle)(\hat C_{\ell'} - \langle \hat C_{\ell'} \rangle) \rangle_{\text{MC}}.$$

(12)

This approach has been widely used (see e.g. Jones et al. 2006; Pryke et al. 2009; Planck Collaboration II 2013), and it is also the baseline procedure for cROMASTER used for the analysis reported in this paper.

Estimating off-diagonal elements of the covariance matrix with high accuracy requires a huge number of simulations. In fact, assuming that the covariance estimate is Wishart distributed, the uncertainty on the diagonal elements simply scales as $\sqrt{(2/N_{\text{sim}})}$, while for the off-diagonal elements it scales as $\sqrt{1/N_{\text{sim}} \times \text{Cov}_{\ell \ell'} / \text{Cov}_{\ell \ell'}}$, which is significantly worse in the presence of small correlations.

Tristram et al. (2005) proposed an analytical approximation involving measured auto- and cross-power spectra which is accurate only at high multipoles and for large sky fraction. Planck power spectrum covariance is also based on an analytical approximation (Planck Collaboration XV 2013), that apart from being rather more complex, involves a fiducial model $C_\ell$ as well as assuming uncorrelated pixel noise.

### 3 COMPUTATIONAL REQUIREMENTS

BOLPOL is a fully parallel implementation of the QML method, as described in Section 2.1, written in Fortran90. Since the method works in pixel space, the computational cost rapidly increases as one considers higher resolution maps of a given sky area. The original code, that has been applied to WMAP 5 yr data in (Gruppuso et al. 2009; Paci et al. 2010), to WMAP 7 yr data in (Gruppuso et al. 2011; Paci et al. 2013), and to WMAP 9 yr data in (Gruppuso et al. 2013), performs the analysis both in temperature and polarization. Here, we consider a modification of such implementation that works only on the temperature sector and can therefore ingest higher resolution maps than the original code. The following computational requirements are referred to this version of BOLPOL which has been already applied to Planck data (Planck Collaboration XV 2013; Planck Collaboration XXIII 2013).

---

2 For the sake of completeness, we notice that cROMASTER implements also an approach for error estimation based on bootstrapping; for a given $\ell$, we generate a set of fake $\tilde a_{\ell m}$ by resampling the observed $a_{\ell m}$ with uniform probability, and these are used to compute a set of fake $\hat C_\ell$ to be fed into equation (12) to estimate the diagonal elements of the matrix. In order to account for the sky cut, only $(2\ell + 1)f_{\text{sky}}^\text{eff}$ elements of the $\tilde a_{\ell m}$ are averaged to compute the fake $\hat C_\ell$. This is a very fast procedure based only on real data, and it is especially useful for testing purposes (i.e. when having a quick look to the data or checking for instrumental systematic effects). However, it is not accurate enough for cosmological analysis especially at low multipoles and when only a small fraction of the sky is considered. For this reason, we do not employ it here.
D. Molinari et al.

Table 1. Details of the considered MC simulations. Two cases are taken into account, each of them consist of 1000 realizations from the best fit of Planck model (in fact the choice of the model is irrelevant for our purposes). First column ‘Case’ is for type of simulations. Second column ‘Res’ is for the resolution which is expressed in terms of the parameter $N_{\text{side}}$. Third column ‘Beam’ is for the adopted FWHM. Fourth column is for the level of ‘White noise’ which is given in $\mu K^2$. Fifth column ‘$N_{\text{sims}}$’ contains the number of considered simulations. Sixth column ‘Code’ specifies which code is applied for each case, where B stands for BOLPOL and C for cROMASTER. Seventh column ‘Masks’ identify the masks that have been applied to the simulated maps to perform the analysis: the WMAP kq85 masks (a) and the kq85 enlarged by $8^\circ$ (b). See the text for more details.

| Case | Res $N_{\text{side}}$ | Beam (deg) | Noise ($\mu K^2$) | $N_{\text{sims}}$ | Code | Masks |
|------|----------------------|------------|------------------|-----------------|------|-------|
| 1    | 64                   | 0.916      | 1.0              | 1000            | B, C | a, b  |
| 2    | 256                  | 0.573      | 1.0              | 1000            | C    | a, b  |

The inversion of the covariance matrix $C$ scales as the third power of the side of the matrix, i.e. $O(N^3)$ being $N$ the number of observed pixels. The number of operations is roughly driven, once the inversion of the total covariance matrix is done, by the matrix–matrix multiplications to build the operators $E_{\ell}^{\ell'}$ in equation (4) and by calculating the Fisher matrix $F_{\ell\ell'}$ given in equation (3). The memory, RAM, required to build these matrices is of the order of $O(\Delta k N^2)$ where $\Delta k$ is the range in multipoles of $C^{-1}(\partial C/\partial C^*)$ (for every $X$) that are built and kept in memory during the execution time. This implementation is memory demanding, but it has the benefit of speeding up the computations. For a temperature all-sky map of 49 152 pixels (resolution parametrized by the HEALPIX parameter $N_{\text{side}} = 64$), see Gorski et al. (2005), the total amount of RAM needed by BOLPOL is 19 TB. When we run the code into massively parallel computer clusters such as FERMI, at CINECA, it takes roughly one day using 16 384 cores to perform a MC of 1000 maps with 20 percent of the sky masked. Possible future development and optimization of the code will lead the QML method to analyse higher resolution maps, but, at the moment, the QML method can be applied only to low resolution, i.e. large angular scales.

On the other side, cROMASTER is very light and can be easily run on a common laptop since it works in harmonic space.

4 COMPARISON

The main idea of the paper is to compare the two aforementioned methods for APS extraction at the largest angular scales under realistic conditions. Our goal is to quantitatively find out what extent it is worth to use a QML method with respect to a quicker and lighter pseudo-Cl approach.

4.1 Details of the simulations

We consider several cases summarized in Table 1. The low-resolution case, namely case 1, is parametrized by $N_{\text{side}} = 64$ and corresponds to the maximum resolution that BOLPOL is currently capable to treat. The high-resolution case, i.e. case 2, with $N_{\text{side}} = 256$ is analysed by cROMASTER. Of course cROMASTER can be run also at higher resolution but this does not impact the range of multipoles where we wish to compare the two codes which essentially is $\ell < \ell_{\text{dec}} \sim 90$. The noise value of $1 \mu K^2$ at $N_{\text{side}} = 256$ is the typical variance for a Planck-like experiment (Planck Collaboration I 2013). The adopted value at large scale is still $1 \mu K^2$, that of course corresponds to a higher noise. This is done to regularize the numerics for this case. In practice, even if different, the noise levels for both cases of Table 1 are negligible and this different treatment does not impact on the analysis we perform.

Each of the cases of Table 1 is analysed with and without Galactic masks. For the latter option, we used the WMAP kq85 temperature mask (Fig. 1a), publicly available at the LAMBDA website, and an extended version of the kq85 that we call ‘kq85 + 8 deg’ (Fig. 1b) where we have extended the edges of the mask by $8^\circ$. The former mask excludes about 20 percent of the sky, the latter about 48 percent. In the full sky case and in a signal dominated regime, the two codes are algebraically equivalent. We have checked that this is indeed the case for internal consistency. Since a Galactic masking is in practice unavoidable in CMB data analysis, we do not report on this unrealistic case.

For cases 1 and 2, we analyse 1000 ‘signal plus noise’ MC simulations where the signal is randomly extracted through the synfast routine of the HEALPIX package (Gorski et al. 2005), from a ΛCDM Planck best-fitting model (Planck Collaboration XV 2013) and the noise from a Gaussian distribution with variance given by the value reported in Table 1.

The averages and variances of the APS of the two MC simulations are plotted in Fig. 2 where we considered the WMAP kq85 mask (upper panel) and the kq85 mask enlarged by $8^\circ$ (lower panel). These figures are considered as the validation of the performed extractions.\(^6\)

\(^3\)http://healpix.jpl.nasa.gov/
\(^4\)For the reader who is not familiar with HEALPix library, we remind that $N_{\text{side}}$ is implicitly defined by $N_{\text{pix}} = 12 N_{\text{side}}^2$, where $N_{\text{pix}}$ is the total number of pixels in a CMB map.
\(^5\)http://www.cineca.it/
\(^6\)http://lambda.gsfc.nasa.gov/
\(^7\)Since the codes run at different resolutions, their difference is driven by the considered smoothing, i.e. FWHM.
\(^8\)In fact, Fig. 2 might be seen as the validation of both codes and extractions at the same time. However, an extensive validation of the codes is already given in Gruppuso et al. (2009) and in Polenta et al. (2005).
Comparison of CMB TT APS estimators

4.2 Figure of Merit

In order to make a detailed comparison between the two methods, we have to define a suitable estimator. Our approach is very similar to what is proposed by Efstathiou (2004a).

For each multipole and for each realization of MC simulations, we compute the APS and build plots as in Fig. 3. In such a figure, each point $P = (x, y)$ has the abscissa $x$ given by the APS obtained with anafast in the ideal case (i.e. full sky and no noise) and the ordinate $y$ that is given by the APS estimated through the BOLPOL or cROMASTER for the cases of Table 1.

If the codes were ‘perfect’, only the diagonal of this kind of plots would be populated (see black solid line in Fig. 3). In fact, there are two clouds of points, one for BOLPOL estimates, shown in blue, and the other for the cROMASTER estimates, shown in red. The idea is to measure the dispersion of the two clouds around the solid black line. This defines our estimator aimed at the comparison of the two codes. The code that shows larger dispersion has an intrinsic larger variance in the determination of the APS. In practice, for each single multipole $\ell$, we define the variance $D^2_\ell$ as the mean of the squared distance $d_\ell$ of each point $P$ from the line $y = x$, which is the diagonal of the first quadrant of this Cartesian plane,

$$D^2_\ell \equiv \langle d_\ell^2(P_{B/C}, \text{diagonal}) \rangle,$$

where the labels $B/C$ refer to BOLPOL and cROMASTER and with $\langle \rangle$ standing for the ‘ensemble’ average. We underline that, in this way, the estimator cancels the uncertainty due to the cosmic variance that is the same for both the codes and highlight their different intrinsic variance. Taking the square root of equation (13), we obtain

$$D_{\ell}^{B/C} = \frac{\ell(\ell+1)}{2\sqrt{2\pi}} \left( \langle (C_\ell^{B/C} - C_\ell^A)^2 \rangle \right)^{1/2},$$

where $C^A_\ell$ is the APS computed with anafast in the ideal case. From equation (14), it is clear that the unit of $D_\ell^{B/C}$ is the same as the one used for the APS, that in our case is $\mu K^2$. Equation (14) is what we consider in the next section to perform the comparison.

4.3 Results

Fig. 4 shows the estimator $D_\ell$, defined in equation (14), as a function of the multipole $\ell$ for each of the cases of Table 1. This plot demonstrates that the intrinsic variance of BOLPOL is lower than the intrinsic variance of cROMASTER up to $\ell \sim 100$. The differences between the two estimators, $(D_\ell^{C} - D_\ell^{B})$ versus $\ell$ is shown in Fig. 5. This makes clear that the difference in the accuracy of the
two methods is higher at lowest multipoles and that it grows as the number of masked pixel increases. In particular, when we consider the WMAP kq85 mask (kq85 enlarged by 8°), the intrinsic dispersion introduced by the pseudo-$C_\ell$ method is at least a factor of 3 (a factor of 2) greater than that of the QML estimates for $\ell \leq 10$. In the range $20 \leq \ell \leq 60$, the QML is about 20 per cent (15 per cent) more accurate than the pseudo-$C_\ell$ method. At higher multipoles, i.e. $\ell > 100$, the larger QML intrinsic variance displayed in Fig. 4 is entirely due to the lower resolution at which BOLPOL is run with respect to cROMASTER. Note however that when the two codes are run at the same resolution, i.e. $N_{\text{side}} = 64$, the QML has always a smaller variance than cROMASTER in the commonly valid multipole domain, in as shown in Fig. 6 when we used the WMAP kq85 mask. These results show that the resolution given by $N_{\text{side}} = 64$ is enough to have an optimal APS extraction with the QML method in the range of interest ($\ell < \ell_{\text{dec}} \sim 90$) compared to pseudo-$C_\ell$ estimates performed on maps at the best resolution allowed by the observations.

11 cROMASTER results are fully reliable only up to $\ell_{\text{max}} = 2 \times N_{\text{side}} = 128$.

12 We obtained the same result when we considered the enlarged mask.

4.4 Applications

We consider here two estimators of anomalies used at large scales in CMB data analysis to illustrate the benefit of applying an optimal APS estimator. The estimator, $R$, see Kim & Naselsky (2010) and Gruppuso et al. (2011) for the TT parity analysis, is $R \equiv C_{TT}^+/C_{TT}^-$ where $C_{TT}^{+/−} \equiv 1/(\ell_{\text{tot}}^{+/−}) \sum_{\ell=2}^{\ell_{\text{max}}} \ell(\ell + 1)/(2\pi)C_{TT}^{\ell}$ and $\ell_{\text{tot}}^{+/−}$ is the total number of even (+) or odd (−) multipoles taken into account in the sum. The Variance estimator, $\sigma^2$ (e.g. Monteserin et al. 2008; Cruz et al. 2011; Gruppuso et al. 2013; Planck Collaboration XXIII 2013, and reference therein) is defined by $\sigma^2 \equiv \langle \delta T^2 \rangle \equiv \sum_{\ell=2}^{\ell_{\text{max}}} C_{TT}^{\ell} / (2\ell + 1)/(4\pi)$.

The variances of these two estimators are affected by the uncertainties on the APS. In Fig. 7, we show the ratio of the variances of $R$ obtained through the APS extracted by cROMASTER and BOLPOL for each $\ell_{\text{max}}$. In the same figure, we show also the ratio between the variances of $\sigma^2$ obtained with cROMASTER and BOLPOL. It is clear that the lower uncertainty given by BOLPOL leads to a lower variance for both the two estimators in the range of interest ($\ell \leq \ell_{\text{dec}} \simeq 90$). For the TT Parity estimator, the average gain in the efficiency of about 10 per cent when we use the WMAP kq85 mask (20 per cent for the kq85 mask enlarged by 8°) with a peak at the lowest scales
Figure 7. Ratio of the variances (cROMASTER over BOLPOL) for the Variance (solid line) and the TT Parity (dashed line) estimators as a function of the multipoles obtained from the BOLPOL and cROMASTER APS extracted in the previous section when we considered the WMAP kq85 mask (green) and the kq85 mask enlarged by 8° (magenta).

5 CONCLUSION

Our main result is given by Figs 4 and 5 where the intrinsic variance, see equation (14), of the two APS estimators, namely BOLPOL and cROMASTER, are compared under realistic conditions. We have found that the QML method is markedly preferable in the range 2 ≤ ℓ ≤ 100. Moreover, we note that the largest difference between the two codes is for the lowest multipoles: for ℓ smaller than ~20, the square root of the intrinsic variance introduced in the estimates by the pseudo-Cℓ is at least up to three times (two times) the QML one when we consider the WMAP kq85 mask (respectively, the kq85 mask enlarged by 8°). For higher multipoles (i.e. ℓ ≥ 100), we observe an opposite behaviour. This stems from the smoothing of the input maps that in turn is a consequence of the adopted resolutions. Note however that when the two codes are run at the same resolution, i.e. Nside = 64, the QML has always a smaller variance than cROMASTER in the commonly valid multipole domain, as shown in Fig. 6.

We have also analysed how the intrinsic variance of the two APS methods impacts on some typical large-scale anomaly estimators like the TT Parity estimator and Variance estimator. In conclusion, the use of BOLPOL for low-resolution map analysis will bring to tighter constraints for these kind of estimators.

Therefore, we suggest to use the QML estimator and not the pseudo-Cℓ method in order to perform accurate analyses that are based on the APS at large angular scales (at least ℓ ≤ ℓdec ≈ 90). This might be of particular interest for studying large-scale anomalies in the temperature anisotropy pattern. In a future work, we will extend this analysis to the polarization field, which is crucial to reveal the reionization imprints at large scale.

ACKNOWLEDGEMENTS

We acknowledge the use of computing facilities at NERSC (USA) and CINECA (ITALY). We acknowledge use of the HEALPIX (Gorski et al. 2005) software and analysis package for deriving the results in this paper. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA), part of the High Energy Astrophysics Science Archive Center (HEASARC). HEASARC/LAMBDA is a service of the Astrophysics Science Division at the NASA Goddard Space Flight Center. Work supported by ASI through ASI/INAF Agreement I/072/09/0 for the Planck LFI Activity of Phase E2 and by MIUR through PRIN 2009 (grant no. 2009XZ54H2).

REFERENCES

Albrecht A., Steinhardt P. J., 1982, Phys. Rev. Lett., 48, 1220
Benabed K., Cardoso J. F., Prunet S., Hivon E., 2009, MNRAS, 400, 219
Bennett C. et al., 2011, ApJS, 192, 17
Bennett C. L. et al., 2013, ApJS, 208, 20
Cruz M., Vielva P., Martnez-Gonzalez E., Barreiro R. B., 2011, MNRAS, 412, 2383
Dunkley J. et al., 2009, ApJS, 180, 306
Efstathiou G., 2004a, MNRAS, 348, 885
Efstathiou G., 2004b, MNRAS, 349, 603
Eriksen H. K. et al., 2004, ApJS, 155, 227
Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Grain J., Tristram M., Stompor R., 2009, Phys. Rev. D, 79, 123515
Gruppuso A., de Rosa A., Cabella P., Paci F., Finelli F., Natoli P., de Gasperis G., Mandolesi N., 2009, MNRAS, 400, 1
Gruppuso A., Finelli F., Natoli P., Paci F., Cabella P., de Rosa A., Mandolesi N., 2011, MNRAS, 411, 1445
Gruppuso A., Natoli P., Paci F., Finelli F., Molinari D., de Rosa A., Mandolesi N., 2013, J. Cosmol. Astropart. Phys., 07, 047
Guth A. H., 1981, Phys. Rev. D, 23, 347
Hauser M. G., Peebles P. J. E., 1973, ApJ, 185, 757
Hinshaw G. et al., 2003, ApJS, 148, 135
Hinshaw G. et al., 2013, ApJS, 208, 19
Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
Jewell J., Levin S., Anderson C. H., 2004, ApJ, 609, 1
Jones W. C. et al., 2006, ApJ, 647, 823
Kim J., Natsel’sky P., 2010, ApJ, 714, L265
Knox L., Turner M. S., 1994, Phys. Rev. Lett., 73, 3347
Kofman L., Starobinsky A. A., 1985, Sov. Astron. Lett., 11, 271
Kuo C. I. et al., 2004, ApJ, 600, 32
Linde A. D., 1982, Phys. Lett. B, 108, 389
Mathe J. C., Fixsen D. J., Shafer R. A., Mosier C., Wilkinson D. T., 1991, ApJ, 412, 2383
Monteserin C., Barreiro R. B., Vielva P., Martínez-González E., Hobson M. P., Lasenby A. N., 2008, MNRAS, 387, 209
Paci F., Gruppuso A., Finelli F., Cabella P., De Rosa A., Mandolesi N., Natoli P., 2010, MNRAS, 407, 399
Paci F., Gruppuso A., Finelli F., De Rosa A., Mandolesi N., Natoli P., 2013, MNRAS, 434, 3071
Page L. et al., 2003, ApJS, 148, 233
Planck Collaboration I, 2013, A&A, preprint (arXiv:1303.5062)
Planck Collaboration II, 2013, A&A, preprint (arXiv:1303.5063)
Planck Collaboration XV, 2013, A&A, preprint (arXiv:1303.5075)
Planck Collaboration XVI, 2013, A&A, preprint (arXiv:1303.5076)
Planck Collaboration XX, 2013, A&A, preprint (arXiv:1303.5080)
Planck Collaboration XXIII, 2013, A&A, preprint (arXiv:1303.5083)
Planck Collaboration XXIV, 2013, A&A, preprint (arXiv:1303.5084)
D. Molinari et al.

Polenta G., Marinucci D., Balbi A., de Bernardis P., Hivon E., Masi S., Natoli P., Vittorio N., 2005, J. Cosmol. Astropart. Phys., 11, 001
Pryke C. et al., 2009, ApJ, 692, 1247
Sachs R. K., Wolfe A. M., 1967, ApJ, 147, 73
Saha R., Jain P., Souradeep T., 2006, ApJ, 645, L89
Starobinsky A. A., 1980, Phys. Lett. B, 91, 99
Taylor J. F., Ashdown M. A. J., Hobson M. P., 2008, MNRAS, 389, 1284
Tegmark M., 1997, Phys. Rev. D, 55, 5895
Tegmark M., de Oliveira-Costa A., 2001, Phys. Rev. D, 64, 063001
Tristram M., Macías-Pérez J. F., Renault C., Santos D., 2005, MNRAS, 358, 833
Wandelt B. D., Larson D. L., Lakshminarayanan A., 2004, Phys. Rev. D, 70, 083511
Wu J. H. et al., 2007, ApJ, 665, 55

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