Secondary non-Gaussianity and cross-correlation analysis

Dipak Munshi, 1,2⋆ Alan Heavens, 1 Asantha Cooray3 and Patrick Valageas4

1Scottish Universities Physics Alliance (SUPA), Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ
2School of Physics and Astronomy, Cardiff University, CF24 3AA
3Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
4Institut de Physique Théorique, CEA Saclay 91191,Gif-sur-Yvette, France

Accepted 2011 February 28. Received 2011 February 9; in original form 2010 August 30

ABSTRACT

We develop optimized estimators of two sorts of power spectra for fields defined on the sky, in the presence of partial sky coverage. The first is the cross-power spectrum of two fields on the sky; the second is the skew spectrum of three fields. The cross-power spectrum of the cosmic microwave background (CMB) sky with tracers of large-scale structure is useful as it provides valuable information on cosmological parameters. Numerous recent studies have proved the usefulness of cross-correlating the CMB sky with external data sets, which probes the integrated Sachs–Wolfe (ISW) effect at large angular scales and the Sunyaev–Zel’dovich (SZ) effect from hot gas in clusters at small angular scales. The skew spectrum, recently introduced by Munshi & Heavens, is a powerful statistic, as it is optimized to study particular forms of non-Gaussianity, such as may arise in the early Universe, but in addition, it retains information on the nature of non-Gaussianity. As such, it allows a robust statistical analysis, where contributions from primordial and contaminating non-Gaussianity can be estimated. In this paper we develop the mathematical formalism for the skew spectrum of three different fields. When applied to the CMB, this allows us to explore the contamination of the skew spectrum by secondary sources of CMB fluctuations, in the case where the foreground contamination and the primary signal are not independent. After developing the analytical model we use them to study specific cases of cosmological interest which include cross-correlating the CMB with various large-scale tracers to probe the ISW and SZ effects for cross-spectral analysis. Next we use the formalism to study the signal-to-noise ratio for the detection of the weak lensing of the CMB by cross-correlating it with different tracers, as well as point sources for CMB experiments such as Planck.

Key words: methods: analytical – methods: numerical – methods: statistical – cosmic background radiation – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Observations of the cosmic microwave background (CMB) and large-scale structure carry complementary cosmological information. While all-sky CMB observations such as NASA’s WMAP1 and ESA’s current Planck2 experiments primarily probe the distribution of matter and radiation at redshift \( z = 1300 \), large-scale surveys tend to give us a window at low redshift \( z \sim 0 \). The main advantage of cross-correlating such independent data sets lies in the fact that it is possible to highlight signals which may not be otherwise detected in individual data sets independently. Earlier studies in this direction include, for example, Peiris & Spergel (2000) who carried out a detailed error forecast of such cross-correlation analyses for cosmological parameters. Clearly, for these tracers to be effective in constraining cosmology, they should be as numerous as possible to reduce the Poisson noise and the survey should cover as large a fraction of the sky as possible to reduce sample variance.

*E-mail: Dipak.Munshi@astro.cf.ac.uk
1 http://map.gsfc.nasa.gov/
2 http://www.rssd.esa.int/index.php?project=Planck

© 2011 The Authors
Monthly Notices of the Royal Astronomical Society © 2011 RAS
Various authors have used different external data sets with specific astrophysical tracers to trace the large-scale structure (LSS), with one of the main motivations being to detect the ISW effect as predicted for ΛCDM cosmology. Earlier studies in this direction include Fosabala & Gaztanaga (2004); Fosabala et al. (2003) who cross-correlated the SDSS Data Release 1 galaxies with the first-year full-sky WMAP data. Nolta et al. (2004) cross-correlated the NVSS radio source catalogue with first-year full-sky WMAP data. Scranton et al. (2003) correlated the Sloan Digital Sky Survey against the WMAP data. Boughn & Crittenden (2005, 2004, 2005) used two tracers of the LSS: the HEAO1 A2 full-sky hard X-ray map and the NVSS large area radio galaxy survey. A maximum likelihood fit to both data sets yields a detection of an ISW amplitude at a level consistent with what is predicted by the ΛCDM cosmology. Most of these studies detected ISW effect at a level of 2σ–3σ although the error analysis models and the statistic used were sometimes completely different [see Ho et al. 2008 for tomographic studies involving ISW and Hirata et al. (2008) for weak lensing detection]. The ISW effect remains one of the most direct and quantitative measures of the dark energy available to us today. Future all-sky missions such as Planck will provide an excellent possibility to extend these studies to a higher confidence regime. While the above studies are mainly focused on large angular scales, where the ISW effect plays an important role, at small angular scales, the presence of clusters and probably the associated filamentary network in which they reside can also affect CMB maps through the Sunyaev–Zel’dovich effect (Sunyaev & Zel’dovich 1980) as well as through the X-ray maps via bremsstrahlung. Cross-correlation analysis of the diffuse soft X-ray background maps of ROSAT with WMAP first-year data were performed by Diego, Silk & Silwa (2003, a, b). This study was motivated by the fact that hot gas in clusters can be more easily detected by cross-correlating X-ray and CMB maps. Although no evidence was found of this effect, it opens the possibility of detecting such an effect in future high-resolution CMB maps. All these act as a motivation for development of a generic techniques to cross-correlate high-resolution CMB maps with other maps from LSS surveys. In this paper we focus on cross-correlating two or three different data sets, but the challenges are similar to those arising from a single data set. For example, the estimation of the power spectrum from a single high-resolution map poses a formidable numerical problem in terms of computational requirements. Typically two different methods are followed. The first is the non-linear maximum likelihood method, or its quadratic variant, which can be applied to smoothed degraded maps, as it is not possible to directly invert a full pixel covariance matrix (Tegmark 1997). To circumvent this problem, a pseudo-\( C_l \) (PCL) values technique was invented (Hivon et al. 2002) which is unbiased though remains suboptimal. More recently, Efstathiou (2004, 2006) has shown how to optimize these estimators which can then be used to analyse high-resolution maps in a very fast and accurate way. We generalize the PCL-based approach here to compute the cross-correlation of different data sets. The method developed here is completely general and can be applied to an arbitrary number of data sets. For example, our formalism can analyse the degree of cross-correlation among various CMB surveys observing the same region of the sky with different noise levels and survey strategies.

For near-Gaussian fields, the two-point analysis from any cosmological survey provides the bulk of the cosmological information. However, going one step further, at the level of the three-point correlation, the detection of departure from Gaussianity in the CMB can probe primary non-Gaussianity (e.g. Munshi & Heavens 2010) as well as the mode-coupling effects due to secondaries. The possibility of further improving the detection of primordial non-Gaussianity with CMB maps, given the current hints with WMAP data (Yadav & Wandelt 2008; Smith, Senatore & Zaldarriaga 2009), provides further motivation in this direction (Afshrodi 2004; Afshrodi, Loh & Strauss 2004; Alishahiha, Silverstein & Tong 2004; Arkani-Hamed et al. 2004; Boughn & Crittenden 2005; Chen et al. 2007a; Chen, Easther & Lim 2007b). One of the prominent contributions to the secondary non-Gaussianity is the coupling of weak lensing and sources of secondary contributions such as Sunyaev–Zel’dovich (SZ) (Goldberg & Spergel 1999; Cooray & Hu 2000). Although weak lensing produces a characteristic signature in the CMB angular power spectrum, its detection has proved to be difficult internally from the CMB power spectrum alone. The non-Gaussianity imprinted by lensing into the primordial CMB remains below the detection level of current experiments, although with Planck the situation is likely to improve. The difficulty originates mainly from the fact that such detections are linked to the four-point statistics of the lensing potential. However, cross-correlating CMB data with external tracers means lensing signals can be probed at the level of the mixed bispectrum. After the first unsuccessful attempt to cross-correlate WMAP against SDSS, recent efforts by Smith, Zahn & Dore (2007) have found a clear signal of weak lensing of the CMB, by cross-correlating WMAP against NVSS. Their work also underlines the link between three-point statistics estimators and the estimators for weak lensing effects on CMB.

The study of non-Gaussianity is primarily focused on the bispectrum as it saturates the Cramér–Rao bound (Babich 2005); however, in practice it is difficult to probe the entire configuration dependence in the harmonic space from noisy data. The cumulant correlators or multipoint correlators collapsed to probe the two-point statistic. These were introduced in the context of analysing galaxy clustering by Szapudi & Szalay (1999), and were later found to be useful for analysing projected surveys such as APM (Munshi, Melott & Coles 2000). Being two-point statistics, they can be analysed in the multipole space by defining an associated power spectrum. Recent studies by Cooray (2006) and Cooray, Li & Melchiorri (2008) have shown its wider applicability including, e.g. in 21-cm studies. However, the multspectrum elements defined in multipole space are difficult to estimate directly from the data because of their complicated response to partial sky coverage and inhomogeneous noise, as well as associated high redundancy in the information content. However, such issues are well understood in the context of power-spectrum analysis. Borrowing from previous results (Hivon et al. 2002; Yadav & Wandelt 2008; Smith et al. 2009), in this paper we show how the cross-power spectrum and the skew spectrum can be studied in real data in an optimal way. We concentrate on two effects: the cross-correlation power spectrum, which is recovered by cross-correlating two different (but possibly correlated) data sets, focusing on weak lensing effects on the CMB, and secondly the contributions to the skew spectrum from foreground effects. The relation of such cross-power spectrum estimators with higher-order multspsctra such as the bispectrum is also discussed in the context of methods known as pseudo-\( C_l \) values and quadratic estimators. We derive the error covariance matrices and discuss their validity in the signal- and noise-dominated regimes and comment on their relationship to the Fisher matrix. The layout of the paper is as follows: in Section 2 we use the...
formalism based on pseudo-$C_l$ analysis for power spectra to study the cross-correlation power spectrum of different data sets. While we keep the analysis completely general, it is specialized for the case of near-all-sky analysis and use it to compute the signal-to-noise ratio and the covariance of estimated $C_l$ values for various tracers with Planck-type all-sky experiments. Possibilities of using various weights which can make the pseudo-$C_l$ approach near optimal in limiting cases of the signal-dominated regime or the noise-dominated regime are also discussed.

In Section 3 we continue our discussion on Pseudo-$C_l$ but generalize it to the analysis of the skew spectrum. Such an estimator can handle the partial sky coverage and noise in a very straightforward way, but in general it is suboptimal. In the high-$l$ regime where mode–mode coupling can be modelled by decreasing the density of modes by the fraction of the sky observed, $f_{\text{sky}}$, one can make such an estimator nearly optimal using a suitable weighting. After a very brief introduction to various physical effects in Section 4, which introduces mode–mode coupling which leads to CMB bispectra, we move on to develop a crude but fast estimator for the skew spectrum in Section 5. Section 6 is devoted to developing the mixed bispectrum analysis in a fully optimal way by introducing inverse covariance weighting of the data. We analyse both one-point and two-point collapsed bispectral analyses. The one-point estimator or the mixed skewness is introduced – being a one-point estimator it compresses all the available information in a bispectrum to a single number. Next, we introduce the mixed skew spectrum which compresses various components of a bispectrum to a power spectrum in an optimum way. Finally in Section 7, the general formalism of bispectral analysis is used for specific cases of interest.

2 GENERALIZED PSEUDO-$C_l$ ESTIMATOR FOR CROSS-CORRELATION ANALYSIS: GENERALIZATION TO ARBITRARY DATA SETS

Typically a maximum likelihood estimator (or its quadratic version) or a PCL estimator is used for estimation of power spectra. On the one hand quadratic maximum likelihood (QML) estimators are slow but have the advantage of optimal weighting of data. On the other hand, the PCL estimator is fast but suboptimal. There has been progress in combining the speed of PCL estimator with the optimality of the QML estimator by introducing appropriate weights in pixel space depending on the noise (Efstathiou 2004, 2006). In this section we consider pseudo-$C_l$ estimation of the cross-power spectrum of two fields, $\Phi_1(\hat{\Omega}), \Phi_2(\hat{\Omega})$, with partial sky coverage. Estimators are constructed from the spherical harmonic transforms $a_{lm}^{XY}$ over the partial sky, where the tilde indicates that the fields are assumed to take zero value in unobserved regions.

The $a_{lm}^{XY}$ are related to the true all-sky spherical harmonics $a_{lm}^{XY}$ by a linear transformation, via a transformation matrix $K_{lm'lm}$, which we will compute, and where possible direct inversion to get the $a_{lm}$ is much faster than maximum likelihood analysis. Using a suitable choice of the weighting function for the data, this estimation can also be made nearly optimal. In practice, some averaging of the power spectra in band powers may be necessary to regularize the inversion.

2.1 Estimator for $C_{\ell}^{XY}$

We will start by defining the harmonic coefficients of all-sky maps $\Phi(\hat{\Omega})$; it is useful to define the all-sky harmonics using the transform, $a_{lm} = \int \Phi(\hat{\Omega}) Y_{lm}(\hat{\Omega}) \, d\Omega$. In most observational situations, we will only have part of the sky covered. The mask will be denoted by a weight $w(\hat{\Omega})$, equal to zero or one in simple cases. Typically, maps are pixelized using a pixelization scheme. The transformation from the pixel-space discrete maps $\Phi_{lm}^{X,Y}(\hat{\Omega}_i)$ to the harmonic domain for maps $X$ or $Y$ is obtained by transforming the weighted field over the whole sky:

$$a_{lm}^{X,Y} = \sum_{\text{pixels}} \Phi_{lm}^{X,Y}(\hat{\Omega}_i) w_{lm}^{X,Y}(\hat{\Omega}_i) Y_{lm}(\hat{\Omega}_i). \tag{1}$$

No summation over repeated indices is assumed. The superscripts label two different fields, $Y_{lm}(\hat{\Omega}_i)$ represents the spherical harmonics and $\Omega_i$ is the pixel area (which we will assume independent of the pixel position). The same expression in the continuous limit can be written as

$$a_{lm}^{X,Y} = \int \Phi(\hat{\Omega}) \Phi_{lm}^{X,Y}(\hat{\Omega}) Y_{lm}(\hat{\Omega}) \, d\Omega \equiv \sum_{lm'} a_{lm'}^{X,Y} K_{lm'm}. \tag{2}$$

Expanding the weighting function on a spherical basis one can write down the coupling matrix $K_{lm'l'm}$, which encodes all the information regarding mode–mode coupling due to partial sky coverage as [see e.g. Hivon et al. (2002) for a detailed derivation]

$$K_{lm'l'm} = \int w(\hat{\Omega}) Y_{lm'}^{*} (\hat{\Omega}) Y_{l'm} (\hat{\Omega}) d\hat{\Omega}$$

$$= \sum_{l'm} (-1)^{m+l'} \tilde{w}_{lm} \left( \frac{(2l+1)(2l'+1)}{4\pi} \right)^{1/2} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -m_1 & m_2 & m_3 \end{pmatrix}, \tag{3}$$

where $\tilde{w}_{lm}$ is the transform of the window or (arbitrary) weighting function. The matrices represent $3j$ symbols. The quantum numbers $l$ and $m$ need to satisfy certain conditions for the Wigner $3j$ functions to have non-vanishing values. Adopting the notion of Efstathiou (2004), we write the pseudo-$C_l$ in terms of the underlying true cross-power spectrum $C_{\ell}$.

Start by defining the PCL estimators from the direct harmonic transforms:

$$C_{\ell}^{XX} = \frac{1}{2(2\ell+1)} \sum_m |a_{lm}^{X}|^2, \quad C_{\ell}^{XY} = \frac{1}{2(2\ell+1)} \sum_m \Re \left( \bar{a}_{lm}^{X} a_{lm}^{Y} \right). \tag{4}$$

© 2011 The Authors, MNRAS 414, 3173–3197

Monthly Notices of the Royal Astronomical Society. © 2011 RAS.
where $\alpha = XX$, $YY$ or $XY$, and we can estimate the unbiased estimators for various power spectra as

$$\hat{C}_\ell = (M^{-1})^W \hat{C}^\alpha.$$  

In practical cases $M_W$ will be close to singular, and some band-power averaging may be necessary to give an invertible mixing matrix.

The mode-mixing matrix $M$ can be expressed in terms of 3$j$ symbols as (Hivon et al. 2002)

$$M^\alpha_{\ell_1 \ell_2} = (2\ell_2 + 1) \sum_\ell \frac{(2\ell + 1)}{4\pi} W^\alpha_\ell (\ell \quad \ell_2 \quad \ell_1)^2,$$

and $\hat{W}^\alpha_\ell = \frac{1}{\ell_2 + 1} \sum_w |w_{\ell m}|^2$ is the power spectrum associated with the mask. Note that the transformation matrices $M^{XX}$ depend on the power spectrum of the weighting function $w^{XX}_\ell$, whereas the matrix $M^{XY}$ for cross-power spectra is determined by the cross-power spectra $\hat{W}^XY$, with two weighting functions. Independent of the choice of weights, the estimators $\hat{C}_\ell^\alpha$ remain unbiased.

We have assumed that the noise in different surveys is independent. If the noise in two surveys is correlated then any noise contribution needs to be subtracted after treatment for the mask. A Monte Carlo approach is typically employed in such a procedure.

### 2.2 Covariances of pseudo-$C_\ell$ values

The covariances of the pseudo-cross-power spectra estimates can be computed analytically for arbitrary sky coverage and a non-uniform Gaussian noise distribution. The deviation of the estimated $C_\ell$ from the ensemble average $\langle \hat{C}_\ell \rangle$ is denoted by $\delta C_\ell$:

$$\delta \hat{C}_\ell^\alpha = \hat{C}_\ell^\alpha - \langle \hat{C}_\ell^\alpha \rangle, \quad \alpha \in \{XX, YY, XY\}. $$

As before, we will be considering near-all-sky surveys. The construction of estimators does not depend on this assumption but computation of covariance of estimators depend crucially on near-orthogonality of cut-sky modes. Our results depend on the assumption that the noise in different surveys is uncorrelated and Gaussian, characterized by a pixel variance of $[\sigma^2(\hat{Q}_{\text{pixel}})]^\alpha$.

However, the noise distribution can be inhomogeneous. We are concerned here with the computation of the covariance of estimated $C_\ell$ values jointly from two different fields and their cross-correlation. We begin by defining the covariance matrix:

$$\langle \delta \hat{C}_\ell^\alpha \delta \hat{C}_\ell^{\beta} \rangle = \langle \hat{C}_\ell^\alpha \hat{C}_\ell^{\beta} \rangle - \langle \hat{C}_\ell^\alpha \rangle \langle \hat{C}_\ell^{\beta} \rangle, \quad \alpha, \beta \in \{XX, YY, XY\}. $$

The covariance of $C_\ell$ values from individual fields can be expressed as follows. These results extend the results obtained for individual surveys (Efstathiou 2004) (the results for single fields can be recovered as a special case by identifying $\ell = 0$):

$$\langle \delta \hat{c}_\ell^{XX} \delta \hat{c}_\ell^{YY} \rangle = \frac{1}{4\pi} \sum_\ell \left\{ 2C^{XX}_\ell C^{YY}_\ell \sum_M \text{Re} \left\{ [w^X w^Y]_{LM} [w^X w^Y]_{LM} \right\} + \sqrt{C^{XX}_\ell C^{YY}_\ell} \sum_M \text{Re} \left\{ (w^X w^X)_{LM} (w^2 \sigma^2)^{YY}_{LM} \right\} \right\} (L \quad \ell \quad \ell')^2.$$  

Here we have introduced the following notations:

$$[w^X w^X]_{LM} = \int d\Omega [w^X(\Omega)]^2 Y^*_{LM}(\Omega),$$

$$[w^X w^Y]_{LM} = \int d\Omega [w^X(\Omega)] w^Y(\Omega) Y^*_{LM}(\Omega),$$

$$[w^2 \sigma^2]^{XX}_{LM} = \int d\Omega [w^X(\Omega)]^2 \sigma^2_X(\Omega) Y^*_{LM}(\Omega),$$

$$[w^2 \sigma^2]^{YY}_{LM} = \int d\Omega [w^X(\Omega)]^2 \sigma^2_Y(\Omega) Y^*_{LM}(\Omega).$$

As before, the noise contribution to the individual surveys $[\Phi^X(\Omega) = \Phi^2_X(\Omega) + \Phi^2_X(\Omega)]$ is characterized in pixel space by the noise maps $[\Phi^2_X(\Omega) \Phi^2_X(\Omega)] = [\sigma^2_X(\Omega)]^2 \delta(\Omega - \Omega')$. The noise spectrum associated with the noise maps has variance $C^\alpha = \Omega_p \sigma^2(\hat{Q})$. Here $\Omega_p = 4\pi/N_{\text{pix}}$ is the individual pixel area assumed to be independent of the pixel position. The number of pixel required to cover the entire sky is $N_{\text{pix}}$. The harmonic decomposition for noise maps, masked noise maps etc. is displayed for survey $X$. Similar expressions hold for the second survey and the cross-terms for product of two surveys are also defined in an analogous manner.

For our derivation we start from equation (1) and for simplification we have used the orthogonality relationship for spherical harmonics. These therefore are approximate results valid in the limiting case of all-sky experiments. The total scatter or error depends mainly on two different contributions. The terms involving power spectra of the mask are directly linked with cosmic variance contributions. These terms dominate at large angular scales. The terms that contribute at small angular scales are determined by the noise in individual data sets. There
are additional terms which are mixed terms and depend on both the mask and the noise, and which may play a dominant role on intermediate angular scales.

Extending the above results to the covariance of \( \mathcal{C}_{L}^{XY} \) for the cross-power-spectrum can be expressed as

\[
\langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XY} \rangle = \frac{1}{4\pi} \sum_{L} \left\{ \sqrt{C_{L}^{XX} C_{L'}^{XY} C_{L'}^{YY}} \sum_{M} \text{Re} \left\{ [w^L w^X]_{LM} [w^Y w^Y]_{LM} \right\} + C_{L}^{XY} \sum_{M} \text{Re} \left\{ [w^L w^X]_{LM} [w^X w^Y]_{LM} \right\} \right. \\
+ \sqrt{C_{L}^{XX} C_{L'}^{XY}} \sum_{M} \text{Re} \left\{ [w^2 \sigma^2]_{LM} (w^X w^Y)_{LM} \right\}
+ \sqrt{C_{L'}^{YY} C_{L'}^{YY}} \sum_{M} \text{Re} \left\{ (w^Y w^Y)_{LM} (w^2 \sigma^2)_{LM} \right\} \\
+ \sum_{M} \text{Re} \left\{ (w^2 \sigma^2)_{LM} (w^2 \sigma^2)_{LM} \right\} \left( \begin{array}{ccc} L & \ell & \ell' \\ 0 & 0 & 0 \end{array} \right)^2.
\]

(15)

Note that the noise contributions are the same for both equations (10) and (15). This is related to the fact that we have assumed the noise to be uncorrelated. As we are focusing on joint estimation from two data sets, the error bars on \( \mathcal{C}_{L}^{XY} \) depend also on the individual noise spectra. In our derivation we have assumed that all the three power spectra are being estimated from the data simultaneously:

\[
\langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{YY} \rangle = \frac{1}{2\pi} \sum_{L} \left\{ \sqrt{C_{L}^{XY} C_{L'}^{XY} C_{L'}^{YY}} \sum_{M} \text{Re} \left\{ [w^L w^X]_{LM} [w^Y w^Y]_{LM} \right\} \right. \\
+ \sqrt{C_{L}^{XY} C_{L'}^{XY}} \sum_{M} \text{Re} \left\{ (w^2 \sigma^2)_{LM} (w^X w^Y)_{LM} \right\} \left( \begin{array}{ccc} L & \ell & \ell' \\ 0 & 0 & 0 \end{array} \right)^2.
\]

(16)

and similarly for \( \langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XX} \rangle \).

Finally the error covariances associated with deconvolved estimators \( \hat{C}_L \) can be expressed in terms of that of the convolved estimators \( \mathcal{C}_L \) as follows:

\[
\langle \hat{\delta \mathcal{C}}_{L}^{XY} \hat{\delta \mathcal{C}}_{L'}^{XY} \rangle = \sum_{\ell} \sum_{\ell'} |M^{-1}|_{\ell \ell'} \langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XY} \rangle |M^{-1}|_{\ell' \ell'}, \quad \alpha, \beta, \gamma, \delta \in \{XX, YY, XY\}.
\]

(17)

The above results represent generic matrix multiplications and no specific symmetry is assumed. The matrix \( M_{\ell \ell'}^{ab} = M_{\ell'}^{ab} \delta_{ab} \) consists only of block diagonal entries. In deriving these results it is assumed that the coverage of the sky is near complete. This will mean that the windows associated with the various couplings are sharper than any features in the power spectra. The shape of the mask and the noise covariance properties are quite general at this stage. If the \( C_L \) values of individual data sets are known from independent estimations then cross-spectra deconvolution of the cross-spectra \( C_L^{XY} \) can be simply written as

\[
\langle L \rangle_{L'} \langle \hat{\delta \mathcal{C}}_{L}^{XY} \hat{\delta \mathcal{C}}_{L'}^{XY} \rangle = \sum_{\ell} [M^{XY}]_{\ell \ell'} \langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XY} \rangle [M^{XY}]_{\ell' \ell'}.
\]

(18)

In the limiting situation when the survey area covers almost the entire sky, these equations take a much simpler form which are in common use in the literature. If \( f_{\text{sky}} \) is the fraction of the sky covered then one can write

\[
\langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XY} \rangle = \frac{1}{(2l + 1) f_{\text{sky}}} \left( C_{L}^{XX} C_{L'}^{XY} + (C_{L'}^{XY})^2 \right) \delta_{\ell \ell'}, \quad \langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{XY} \rangle = \frac{2}{(2l + 1) f_{\text{sky}}} \left( C_{L'}^{XY} \right)^2 \delta_{\ell \ell'}.
\]

(19)

\[
\langle \delta \mathcal{C}_{L}^{XY} \delta \mathcal{C}_{L'}^{YY} \rangle = \frac{2}{(2l + 1) f_{\text{sky}}} \left( C_{L'}^{YY} \right)^2 \delta_{\ell \ell'}.
\]

(20)

The autospectra \( C_L^{XX} \) and \( C_L^{YY} \) in these expressions take contributions from both signal and noise \( C_L \) values, i.e. \( C_L = C_L^S + C_L^N / \sigma_L^2 \). However, the noise contribution is missing in the cross-spectra \( C_L^{XY} \) as we have assumed the noise in these two surveys to be uncorrelated.

While the results derived above are completely general, we will focus on a few special case studies. We consider cross-correlation of the integrated Sachs–Wolfe (ISW) effect with specific LSS tracers. This is a valuable cosmological tool, e.g. to constrain the dark energy equation of state parameter. The specific cosmological parameters that we took for our analysis are \( h = 0.65, \Omega_m = 0.3, \Omega_L = 0.65, \Omega_b = 0.05, \sigma_8 = 0.86, n_s = 1, y_{\text{ISW}} = 0.25 \). Modelling of clustering statistics of haloes which is crucial for the determination of SZ statistics is discussed extensively in the literature; see e.g. Komatsu & Kitayama (1999). We list below the power spectra associated with the ISW effect \( C_L^{ISW} \), the tracer fields \( \mathcal{C}_L \) and their cross-correlation \( C_L^{ISW-X} \) in terms of the matter power spectrum \( P_{ab}(k) \):

\[
C_L^{ISW} = \frac{2}{\pi} \int k^2 dk P_{ab}(k) \left[ I_0^{SW}(k) \right]^2, \quad C_L^X = \frac{2}{\pi} \int k^2 dk P_{ab}(k) \left[ I_0^X(k) \right]^2, \quad C_L^{ISW-X} = \frac{2}{\pi} \int k^2 dk P_{ab}(k) \left[ I_0^{SW}(k) \right] \left[ I_0^X(k) \right].
\]

(21)

The integrands \( I_0(k) \) depend on the window functions which in turn depend on cosmological parameters. For ISW we have

\[
I_0^{SW}(k) = \int_0^\infty dr W^{SW}(r, k) j_0(kr), \quad W^{SW}(r, k) = -3\Omega_M \left( h_0 \right) F, \quad \delta(k, r) = G(r) \delta(k, 0), \quad F(r) = \frac{G(r)}{a(r)}.
\]

(22)

Here \( \Omega_M \) is cosmological density parameter, \( h_0 \) is Hubble constant, \( k \) is the wave vector, \( G \) is the growth factor, \( a \) is the scalefactor of the Universe and \( r \) is conformal distance or equivalently look-back time from observer at redshift \( z = 0 \). See Cooray (2002) and Komatsu & Kitayama (1999) for a more detailed discussion on the modelling of the SZ effect based on halo models. The window functions for the
Figure 1. The power-spectra $C_l$ values are plotted for various combinations of fields. In the left-hand panel we show the power spectrum corresponding to an LSS tracer such as radio sources from NVSS and that of CMB ISW. The cross-power spectrum associated with the two is also depicted. In the right-hand panel we show SZ and ISW power spectrum and their cross-correlation. The CMB power spectrum is also plotted in both panels. SZ curves include only the part that is correlated with the large-scale density field.

Figure 2. Left: the variance for estimated $C_l$ corresponding to ISW, SZ and their cross-correlation. Right: the ISW versus local tracers (NVSS type) analysis. The number density of galaxies for an NVSS-type survey was taken to be $\bar{N} = 7 \times 10^8 \text{sr}^{-1}$. For the CMB a Planck-type experiment was assumed. Results plotted are for all-sky survey (i.e. $f_{\text{sky}} = 1$). For a near-all-sky survey, the variances will scale as $f_{\text{sky}}^{-1}$.

NVSS-type galaxies as well as the SZ surveys depend on modelling the redshift distributions of the tracers. This is typically modelled using a prescription from the halo model of structure formation; we follow Cooray (2002).

Using results in equation (20) the signal-to-noise ratio of the ISW and tracer cross-correlation can now be written in terms of the $C_l$ values defined in equation (21):

$$S/N = f_{\text{sky}} \sum_l (2l + 1) \left\{ \frac{(C_l^{\text{ISW}} - X_l)^2}{(C_l^{\text{ISW}})^2 + (C_l^{\text{ISW}} + N_l^{\text{ISW}})^2 (C_l^X + N_l^X)} \right\}.$$  \hspace{1cm} (23)

The individual power spectra include noise contributions. The noise $N_l^{\text{ISW}}$ is the detector noise in the relevant CMB experiment. It also includes the theoretical CMB power spectra which acts as a noise for the detection of cross-spectra, whereas in LSS surveys it is the shot noise $N_l^X = 1/\bar{N}$. The $C_l$ values for this study are plotted in Fig. 1. The individual variance in power spectra are plotted in Fig. 2. The cumulative signal-to-noise ratio is plotted in Fig. 3. Fig. 1 plots the estimation error for various $C_l$ values, i.e. $\langle [\delta C_l^{\text{ISW}}]^2 \rangle$, $\langle [\delta C_l^X]^2 \rangle$ and $\langle [\delta C_l^{\text{ISW}} - X_l]^2 \rangle$.

The PCL approach is not optimal. It does not optimally weight the noisy data. This means the variance associated with a PCL approach is higher compared to an optimal estimator. Optimal estimator incorporates an inverse variance weighting. It was shown however that the PCL approach can be improved by incorporating weights. In the noise-dominated high-$l$ regime the inverse covariance weighting is essentially reproduced by inverse noise weighting. The low-$l$ regime is better analysed with a QML approach (Efstathiou 2004). In the intermediate regime various fiducial weighting schemes can be introduced in the PCL approach to reduce the scatter in the estimates. These estimates at various $l$ values are then combined in an optimal way. Such an hybrid estimator can also be introduced in the context of cross-correlational analysis. The results developed here are valid for arbitrary weighting schemes and will be useful for the development of such a technique. The weights, for a near-optimal estimator for cross-correlation, will depend on the noise in both surveys. These methods can potentially combine the speed of the PCL estimators and the optimality of QML estimators.
3 THE PSEUDO-$\mathcal{C}_l$ ESTIMATOR FOR MIXED BISPECTRUM ANALYSIS: GENERALIZATION TO ARBITRARY SKY COVERAGE AND MULTIPLE FIELDS

The statistics of temperature fluctuations in the sky are very nearly Gaussian, but small departures from Gaussianity can put constraints on early-universe scenarios. Secondary non-Gaussianity on the other hand can provide valuable information to distinguish structure formation scenarios, and when used with constraints from the power spectrum it can be a very valuable tool. However, estimation of the bispectrum for each triplet of harmonics modes can be difficult to perform numerically. Munshi & Heavens (2010) introduced a limited data compression method for three-point functions which reduces the data to a single function (the skew spectrum), and which can be made optimal for estimating a bispectrum form. In the same spirit, we define a pseudo-skew spectrum for three arbitrary fields defined on a cut sky, and show how it is related to the skew spectrum on the uncut sky. The PCL-based approach described here is not optimal; however, it can be made optimal with a suitable choice of weights.

Let us assume that we have three fields which are defined over the observed sky. The product of two of these fields as $\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})$ has an associated mask, which we denote by $w_X(\hat{\Omega})$ and which is a product of two masks associated with the individual fields. Analogously, the third field $\Phi^Z(\hat{\Omega})$ is observed with a mask $w_Z(\hat{\Omega})$.

From the harmonic transforms of these fields we study the skew spectrum and express it in terms of the mixed bispectra of fields $\Phi^X(\hat{\Omega})$, $\Phi^Y(\hat{\Omega})$ and $\Phi^Z(\hat{\Omega})$, $B_{ljm}^{XY}$. We develop this generally, but the results we derive will be useful for the study of primordial non-Gaussianity. Here we consider a single field (the CMB), but it is a field with contributions from various components, and the skew spectrum contains terms from various triplets of different (or repeated) fields.

3.1 Estimator for $\hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z}$

The skew spectrum we will use is simply the cross-spectrum associated with the product field $\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})$, constructed from fields $\Phi^X(\hat{\Omega})$ and $\Phi^Y(\hat{\Omega})$, against $\Phi^Z(\hat{\Omega})$. The fundamental quantity we consider is the cross-skew-spectrum $\hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z}$, which we define below in equation (24). Assuming that the noise in each field is independent, it will not contribute to the cross-skew-spectrum, but it will be implicit in the harmonic coefficients below.

### 3.1.1 All-sky analysis

We start by introducing the cross-skew-spectrum $\hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z}$, associated with the product map $\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})$ and the field $\Phi^Z(\hat{\Omega})$. In the absence of sky cuts and instrumental noise, we can write

$$\hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z} = \frac{1}{2l+1} \sum_a \text{Re} [a_{lm}^{XY} a_{lm}^{ZY}^*],$$

where $a_{lm}^{XY}$ is the spherical harmonic transform of the product $\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})$, and $a_{lm}^{ZY}$ is a standard:

$$a_{lm}^{XY} = \int d^3\Omega \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega}); \quad a_{lm}^{ZY} = \int d^3\Omega \Phi^Z(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega}).$$

Assuming homogeneity and isotropy, the correlation function $C_{lXY,Z}(\hat{\Omega}, \hat{\Omega}')$ of $\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})$ and $\Phi^Z(\hat{\Omega})$ can be written in terms of $\hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z}$:

$$C_{lXY,Z}(\hat{\Omega}, \hat{\Omega}') = \langle \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})\Phi^Z(\hat{\Omega}') \rangle = \sum_{l_{m1},l_{m2}} \langle a_{l_{m1}m}^{XY} a_{l_{m2}m}^{ZY} \rangle Y_{lm1}(\hat{\Omega})Y_{lm2}(\hat{\Omega}') = \frac{1}{4\pi} \sum_l (2l+1) p_l |\cos(\hat{\Omega} \cdot \hat{\Omega}')}| \hat{\mathcal{C}}_{lXY,Z}^{X,Y,Z}. \quad (26)$$

© 2011 The Authors, MNRAS 414, 3173–3197

Monthly Notices of the Royal Astronomical Society. © 2011 RAS
Here \( P_{l} (\hat{\Omega} \cdot \hat{\Omega}') \) is a Legendre polynomial of order \( l \). The three-point correlation function in the harmonic domain can similarly be used to introduce the mixed bispectrum \( B_{l_{1}l_{2}l_{3}}^{XYZ} \) for the related fields. Assuming statistical isotropy,

\[
(a_{l_{1}m_{1}}^{X}a_{l_{2}m_{2}}^{Y}a_{l_{3}m_{3}}^{Z}) = \left( \begin{array}{ccc} l_{1} & l_{2} & l_{3} \\ m_{1} & m_{2} & m_{3} \end{array} \right) B_{l_{1}l_{2}l_{3}}^{XYZ}.
\]

(27)

Our aim is to compute the cross-correlation power spectra of the product fields \( \Phi^X(\Omega)\Phi^Y(\Omega) \) and \( \Phi^Z(\Omega) \). Using a harmonic decomposition we can relate the multipoles \( a_{lm}^{XY} \) with multipoles \( a_{lm_{1}}^{X} \) and \( a_{lm_{2}}^{Y} \):

\[
a_{lm}^{XY} = \int d\hat{\Omega} Y_{lm}^{*}(\hat{\Omega}) \Phi^X(\hat{\Omega}) \Phi^Y(\hat{\Omega}) = \sum_{l_{1}m_{1}l_{2}m_{2}} a_{l_{1}m_{1}}^{X}a_{l_{2}m_{2}}^{Y} \int d\hat{\Omega} Y_{lm_{1}}^{*}(\hat{\Omega}) Y_{l_{2}m_{2}}(\hat{\Omega})
\]

\[
= \sum_{l_{1}m_{1}l_{2}m_{2}} \sum_{l_{1}m_{1}l_{2}m_{2}} (-1)^{m} a_{l_{1}m_{1}}^{X} a_{l_{2}m_{2}}^{Y} \left( \frac{(2l_{1}+1)(2l_{2}+1)(2l+1)}{4\pi} \begin{pmatrix} l & l_{1} & l_{2} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l_{1} & l_{2} \\ -m & m_{2} & m_{2} \end{pmatrix} \right).
\]

(28)

We have used equation (C4) in deriving the last step. Contracting with the multipole of the remaining field \( a_{lm_{3}}^{Z} \) we can see that it directly probes the mixed bispectrum associated with these three different fields (Cooray 2001a):

\[
\langle a_{lm}^{XY}a_{lm_{3}}^{Z} \rangle = C_{l}^{XYZ} \delta_{ll_{3}} \delta_{m_{3}m_{2}} ;
\]

\[
C_{l}^{XY,Z} = \sum_{l_{1}l_{2}} B_{l_{1}l_{2}}^{XYZ} \frac{(2l_{1}+1)(2l_{2}+1)}{4\pi(2l+1)} \begin{pmatrix} l & l_{1} & l_{2} \\ 0 & 0 & 0 \end{pmatrix}.
\]

(29)

Since the bispectrum is determined by triangular configuration in the multipole space \((l_{1}, l_{2}, l)\), the power spectrum \( C_{l} \) defined above captures information about all the possible triangular configurations when one of its sides is fixed at length \( l \). However, this data compression is not done optimally as it does not weight the contributions from each bispectrum component with their inverse variance. The error covariance matrix can be computed exactly and depends on higher order moments of signal and noise, as well as their cross-correlations.

### 3.1.2 Partial sky coverage

It is possible to extend the above result to take into account partial sky coverage. Assuming that the composite map \( \Phi^X(\Omega)\Phi^Y(\Omega) \) is masked with arbitrary mask \( w_{A}(\Omega) \) and that the map \( \Phi^Z(\Omega) \) is masked with \( w_{B}(\Omega) \), we can write the cut-sky multipoles \( \tilde{a}_{lm}^{XY} \) and \( \tilde{a}_{lm}^{Z} \) in terms of their all-sky counterparts as well as the multipoles associated with the mask multipoles as follows:

\[
\tilde{a}_{lm}^{XY} = \int \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega}) w_{A}(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega}) d\hat{\Omega} ;
\]

\[
\tilde{a}_{lm}^{Z} = \int \Phi^Z(\hat{\Omega}) w_{B}(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega}) d\hat{\Omega}.
\]

(30)

On further simplification using spherical decomposition of individual fields, we can express the composite harmonic coefficient \( \tilde{a}_{lm}^{XY} \) in terms of the individual ones:

\[
\tilde{a}_{lm}^{XY} = \sum_{l_{1}m_{1}l_{2}m_{2}} a_{l_{1}m_{1}}^{X}a_{l_{2}m_{2}}^{Y} w_{A_{l_{1}m_{1}l_{2}m_{2}}} \int d\hat{\Omega} Y_{l_{1}m_{1}}^{*}(\hat{\Omega}) Y_{l_{2}m_{2}}(\hat{\Omega}) Y_{lm_{1}}(\hat{\Omega}) Y_{lm_{2}}^{*}(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega})
\]

\[
\tilde{a}_{lm}^{Z} = \sum_{l_{1}m_{1}l_{2}m_{2}} a_{l_{1}m_{1}}^{Y} w_{B_{l_{1}m_{1}l_{2}m_{2}}} \int d\hat{\Omega} Y_{l_{1}m_{1}}^{*}(\hat{\Omega}) Y_{l_{2}m_{2}}(\hat{\Omega}) Y_{lm_{1}}^{*}(\hat{\Omega}) Y_{lm}^{*}(\hat{\Omega}).
\]

(31)

For partial sky coverage one can obtain with a tedious but straightforward algebra:

\[
\tilde{C}_{l}^{XY,Z} = \frac{1}{2l+1} \sum_{m} \Re \{ \langle [\tilde{a}_{lm}^{XY}]* [\tilde{a}_{lm}^{Z}] \rangle \}
\]

\[
= \sum_{l_{1}l_{2}} \sum_{l_{1}l_{2}} \left( \frac{2(l_{1}+1)}{4\pi} \begin{pmatrix} l & l_{1} & l_{2} \\ 0 & 0 & 0 \end{pmatrix} \right)^{2} |w_{A_{l_{1}m_{1}l_{2}m_{2}}}|^{2} B_{l_{1}l_{2}}^{XYZ} \frac{(2l_{1}+1)(2l_{2}+1)}{(2l+1)4\pi} \begin{pmatrix} l_{1} & l_{2} & l' \\ 0 & 0 & 0 \end{pmatrix} = M_{l} C_{l}^{XY,Z}.
\]

(32)

Here \( B_{l_{1}l_{2}}^{XYZ} \) is the angle-averaged bispectrum introduced above in equation (27). We have assumed that \( w_{A} = w_{B} \), i.e. we take the mask associated with the product map \( \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega}) \) to be the same as the mask associated with the third map \( \Phi^Z(\hat{\Omega}) \). In case we have two different masks we need to replace \( |w_{l}|^{2} \) above by \( \sum_{m} \Re w_{l_{1}m_{1}l_{2}m_{2}} w_{l_{1}m_{1}l_{2}m_{2}}^{*} \). In addition to mask each experimental set-up involves finite size of the beams. The effect of beam smoothing can be incorporated by suitably multiplying the harmonics with a beam window function (assumed circular) \( b_{l} \), i.e. \( a_{lm} \rightarrow b_{lm} a_{lm} \). The effect of pixelization can also be incorporated in a similar way. However, typically generation of map is done using overpixelization at scale much smaller than beam smoothing to minimize the loss of information.

This is one of the important results in this paper. It shows how the pseudo-skew spectrum is related to the all-sky skew spectrum, and is a computationally efficient way to estimate the latter. By suitable choice of weight functions it can be made optimal. It is valid for a completely general mask where \( w_{l} \) represents the spherical transform of the mask. The mixing matrix \( M_{l} \) that appears above in equation (32) is the same as the one defined in equation (7).

The transformation matrix \( M_{l} \) used here is the same as that we introduced for the recovery of cross-power spectra. The power spectra associated with the mask \( w_{l} \), plays the same role in construction of \( M_{l} \). For simplicity we have assumed that different data sets have the same mask, but it is trivial to generalize for two different masks. A detailed comparison of the level of suboptimality will be made with numerical
4 CMB SECONDARY BISPECTRUM: REVIEW OF EXISTING MODELS

The formalism developed so far is quite general and can handle mixed bispectra of different kinds. The main goal was to relate the skew spectrum with the corresponding mixed bispectrum. To make concrete predictions we need to consider a specific form for the bispectrum. Following Spergel & Goldberg (1999), Goldberg & Spergel (1999), Cooray & Hu (2000) and Verde & Spergel (2002) we expand the observed temperature anisotropy $\delta T(\hat{\Omega})$ in terms of the primary anisotropy $\delta T_p$, and due to lensing of primary, $\delta T_L$, and the other secondaries from coupling LSS, $\delta T_S$:

$$\delta T(\hat{\Omega}) = \delta T_p(\hat{\Omega}) + \delta T_L(\hat{\Omega}) + \delta T_S(\hat{\Omega}).$$

(33)

Expanding the respective terms in spherical harmonics, we can write

$$\delta T_p(\hat{\Omega}) \equiv \sum_{lm} a_{lm} Y_{lm}(\hat{\Omega}); \quad \delta T_L(\hat{\Omega}) \equiv \sum_{lm} a_{lm} \nabla \Psi(\hat{\Omega}) \cdot \nabla Y_{lm}(\hat{\Omega}); \quad \langle \Psi(\hat{\Omega}) \delta T_S(\hat{\Omega}) \rangle = \sum_i \frac{2l_i + 1}{4\pi} \beta_i P_i(\hat{\Omega} \cdot \hat{\Omega}).$$

(34)

The secondary bispectrum for the CMB then takes contributions from many products of the $P, L, S$ terms [$\Psi(\hat{\Omega})$ is the projected lensing potential]. For example, one term arises from the products of $\delta T_p \delta T_L \delta T_S$:

$$B_{ilj}^{P,L,S} = \sum_{m_1 m_2 m_3} \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right) \int \langle \delta T_p(\hat{\Omega}_1) \delta T_L(\hat{\Omega}_2) \delta T_S(\hat{\Omega}_3) \rangle Y^*_{l_1 m_1}(\hat{\Omega}_1) Y^*_{l_2 m_2}(\hat{\Omega}_2) Y^*_{l_3 m_3}(\hat{\Omega}_3) d\Omega_1 d\Omega_2 d\Omega_3$$

$$= \sum_{m_1 m_2 m_3} \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right) \langle \delta T_p \rangle_{l_1 m_1} \langle \delta T_L \rangle_{l_2 m_2} \langle \delta T_S \rangle_{l_3 m_3}.$$ 

(35)

It is possible to invert the relation using isotropy of the background cosmology:

$$\langle \delta T_p \rangle_{l_1 m_1} \langle \delta T_L \rangle_{l_2 m_2} \langle \delta T_S \rangle_{l_3 m_3} = \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right) B_{ilj}^{P,L,S}.$$ 

(36)

Explicit calculations, detailed in Goldberg & Spergel (1999) and Cooray & Hu (2000), found the mixed bispectrum to be of the following form:

$$B_{ilj}^{P,L,S} = -\left( \beta_i C_i \frac{l_2(l_2+1) - l_1(l_1+1) - l_3(l_3+1)}{2} + \text{cyc.perm.} \right) \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right) = b_{ilj} I_{ilj},$$

(37)

where $\beta_i$ represents lensing-ISW or lensing-SZ cross-correlation power spectra, $C_i$ is the unlensed CMB angular power spectrum and we have defined the reduced bispectrum $b_{ilj}$ (Bartolo, Matarrese & Riotto 2006), and the additional geometrical factor, which originates from the integral involving three spherical harmonics over the entire sky:

$$I_{ilj} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right).$$

(38)

The cross-correlation power spectra appearing in the above expression denotes the coupling of lensing with a specific form of secondary anisotropy (see e.g. Cooray & Hu 2000). The bispectrum contains all the information at the three-point correlation function level and can be reduced to one-point skewness or the two-point collapsed correlation function or the associated power spectra, the skew spectrum. Here we have considered the secondaries of the CMB; however, the analysis holds if external tracers such as the radio galaxy surveys NVSS or 21-cm observations are used instead. In the next few subsections we will discuss the problem of estimation of the skew spectrum in a nearly optimal way. This will lead to a discussion of development of optimal techniques in the subsequent sections. We will also tackle the problem of joint estimation of several bispectra and associated estimation errors. As is known and will be discussed in the following sections that the estimation of CMB bispectra is similar and related to the case of lensing reconstruction of the CMB sky (Smith et al. 2007).

The total scatter or error depends mainly on two different contributions. The factors involving power spectra of the mask are directly linked with cosmic variance contributions. These terms dominate at large angular scales. The terms which contribute at small angular scales are determined by noise $\sigma^2$ in individual data sets. There are additional terms which are mixed terms and depend on both mask and noise. The mixed terms play a dominant role in intermediate angular scales.

5 ESTIMATION OF SKEW SPECTRA: GENERALIZATION TO THE CASE OF MIXED SECONDARY SKEW SPECTRA

Estimation of the primordial bispectrum is a major goal of all CMB missions. In Munshi & Heavens (2010) we designed estimators that can tackle this problem in the presence of a mask (partial sky coverage) and noise. However, any such measurements will also get contributions from secondary sources which act as contaminants in such measurements. These secondary bispectra are of course interesting in themselves as...
they can provide a valuable tool to study LSS formation scenarios. In this section we will consider a variety of contaminants to the primordial bispectrum. We design estimators which can optimally estimate the skewness and skew spectra associated with these secondary bispectra. The formalism can also provide the level of contamination from secondary sources (including lensing) in the estimation of primordial bispectrum. The secondaries can be categorized by the form of the bispectrum induced. The correlation between the ISW effect with lensing, the Rees–Sciama effect with lensing and the SZ effect with lensing – all have similar bispectra which depend on the CMB power spectrum. We consider these first. Other contributors involve line-of-sight integrations, such as the Ostriker–Vishniac effect correlated with the SZ effect (thermal or kinetic).

The problem of estimation of the secondary skew spectrum is very similar to that of the primary CMB bispectrum (see e.g. Babich & Zaldarriaga 2004; Babich, Creminelli & Zaldarriaga 2004; Babich & Pierpaoli 2008). There has been a recent surge in activity in this area, driven by the claim of detection of non-Gaussianity in the WMAP data release (see e.g. Yadav, Komatsu & Wandelt 2007; Yadav & Wandelt 2008; Yadav et al. 2008). Starting with Heavens (1998) different techniques were developed which introduce various weighting schemes in the harmonic domain to make the method optimal, i.e. saturates the Cramer–Rao bound (Babich 2005). See also Creminelli, Senatore & Zaldarriaga (2007) and Smith et al. (2009). Maps are constructed by weighting the observed CMB sky with \( l \)-dependent weights obtained from inflationary theoretical models. These weighted maps are then used to compute one-point quantities which are the generalization of skewness and can be termed mixed skewness. These mixed skewness measures are useful estimators of the \( f_{NL} \) parameters. A more general treatment is provided in Smith & Zaldarriaga (2006) and Smith et al. (2007, 2009), who took into account mode–mode coupling in an exact way with the use of a proper inverse-covariance weighting of harmonic modes.

The recent work by Munshi & Heavens (2010) has improved the situation by focusing directly on the skew spectrum. Their technique does not compress all the available information in the bispectrum into a single number but provides a power spectrum that depends on the harmonic wavenumber \( l \). This method has the advantage of being able to separate various contributions as they will have different dependence on \( l \), thus allowing an assessment of whether any non-Gaussianity is primordial or not. In this section we compute the contaminating secondary bispectra contributions from lensing-secondary coupling. In Section 5.1 we will consider the bispectra which do not involve line-of-sight integration. These include bispectra linked with the cross-correlation of lensing and secondaries, i.e. ISW, RS and SZ. Next in Section 5.2 we consider the bispectra that involve Ostriker–Vishniac effect that give rise to the line-of-sight integration. Finally in Section 5.3 we quote the general result for skew-spectrum for any specific model of bispectra that can be approximated as a product of several factors (factorizable model).

### 5.1 Bispectra without line-of-sight integration involving ISW-lensing, RS-lensing and SZ-lensing

The study of the bispectrum related to secondary anisotropies (see Cooray & Seth 2002) is arguably as important as that generated by the primary anisotropy. Primary non-Gaussianity in simpler inflationary models is vanishingly small (Salopek & Bond 1990, 1991; Falk et al. 1993; Gangui et al. 1994; Wang & Kamionkowski 2001; Acquaviva et al. 2003; Maldacena 2003); see Bartolo et al. (2006) and references therein for more details. However, variants of simple inflationary models such as multiple scalar fields (Linde & Mukhanov 1997; Lyth, Ungarelli & Wands 2003), features in the inflationary potential, non-adiabatic fluctuations, non-standard kinetic terms, warm inflation (Gupta et al. 2002; Moss & Xiong 2007), or deviations from Bunch–Davies vacuum can all lead to a much higher level of non-Gaussianity.

Early observational work on the bispectrum from COBE (Komatsu et al. 2002) and MAXIMA (Santos et al. 2003) was followed by a much more accurate analysis with WMAP (Cremielli 2003; Komatsu et al. 2003; Corasaniti, Giannantonio & Melchiorri 2005; Creminelli et al. 2007; Spergel et al. 2007). The primary bispectrum encodes information about inflationary dynamics and hence can constrain various inflationary scenarios, whereas the secondary bispectrum will provide valuable information regarding the low-redshift universe and constrain structure formation scenarios. These bispectra are generated because of the cross-correlation effect of lensing due to various intervening materials and the secondary anisotropy such as the SZ effect due to inverse Compton scattering of CMB photons from hot gas in the intervening clusters. The decay of the peculiar gravitational potential along the line of sight in a \( \Lambda \)CDM cosmology introduces an additional fluctuation in the CMB sky which is correlated with lensing generated by the potential. The cross-correlation generates a non-zero bispectrum. Likewise the second-order corrections to the linear potential too can generate a non-zero bispectrum which can be probed using the same approach as the resulting bispectrum has the same form:

\[
 b_{l_{1}l_{2}l_{3}} = -\frac{1}{2} \left( l_{2}(l_{2}+1) - l_{1}(l_{1}+1) - l_{3}(l_{3}+1) \right) C_{l_{1}l_{2}l_{3}}^{S} \beta_{l_{1}} + \text{cyc.perm.}, \tag{39}
\]

The function \( \beta_{l} \) takes different form for ISW-lensing, RS-lensing or SZ-lensing; see Goldberg & Spergel (1999) for derivation.

The power spectrum \( C_{l_{1}}^{S} \) is the \textit{unlensed} target angular spectrum of the primordial CMB anisotropy (only signal contribution is included). We have introduced the superscript \( S \) to distinguish it from the \( C_{l_{1}} \) values that appear as the denominators which take contributions from the instrumental noise (\( C_{l_{1}}^{S} = C_{l_{1}}^{N} + b_{l_{1}}^{2} \)), where \( C_{l_{1}}^{N} \) denotes the instrumental noise and \( b_{l_{1}} \) is the instrumental beam.

The construction of our estimator follows exactly the same procedure as that of the estimator for primordial non-Gaussianity. We define associated fields \( A', B', C' \) on the celestial sphere from the original harmonics with appropriate weights that depend on the CMB power spectra and lensing-secondary cross-spectra. For SZ we only have 2D fields; however, for ISW we need to construct 3D fields. These fields are next used for constructing an unbiased near-optimal estimator \( C_{l_{1}}^{S,1} \).
We define the following harmonics: $A_{lm}^{(i)}$, $B_{lm}^{(i)}$ and $C_{lm}^{(i)}$. These are constructed from the harmonics of the observed CMB sky. These harmonics will be next used to construct an optimum estimator for the corresponding power spectrum associated with the bispectrum (skew-spectrum):

$$
A_{lm}^{(1)} = \frac{a_{lm}}{C_l} \tilde{C}_l^S, \quad B_{lm}^{(1)} = l(l+1) \frac{a_{lm}}{C_l} \tilde{C}_l^T, \quad C_{lm}^{(1)} = \frac{a_{lm}}{C_l} \tilde{C}_l^T 
$$

$$
A_{lm}^{(2)} = -l(l+1) \frac{a_{lm}}{C_l} \tilde{C}_l^S, \quad B_{lm}^{(2)} = \frac{a_{lm}}{C_l} \tilde{C}_l^T, \quad C_{lm}^{(2)} = \frac{a_{lm}}{C_l} \tilde{C}_l^T 
$$

$$
A_{lm}^{(3)} = \frac{a_{lm}}{C_l} \tilde{C}_l^S, \quad B_{lm}^{(3)} = \frac{a_{lm}}{C_l} \tilde{C}_l^T, \quad C_{lm}^{(3)} = l(l+1) \frac{a_{lm}}{C_l} \tilde{C}_l^T 
$$

(40)

The corresponding fields that we construct are $A^{(i)}(\hat{\Omega}) \equiv \sum_l Y_{lm}(\hat{\Omega}) A_{lm}^{(i)}$ and in an analogous manner $B^{(i)}$ and $C^{(i)}$. The optimized skew spectrum in the presence of all-sky coverage and homogeneous noise can now be written as

$$
C_{lm}^{2,1} = \frac{1}{2l+1} \sum_m \sum_i \text{Real} \left\{ [A^{(i)}(\hat{\Omega}) B^{(i)}(\hat{\Omega})]_m C^{(i)*}_m (\hat{\Omega})_m \right\} + \text{cyc.perm.} 
$$

(41)

The cyclic terms that are considered here are constructed from the corresponding terms in the expression for the reduced bispectrum discussed above equation (39). The above estimator can be seen as a scalar product of the theoretical bispectra and the estimated bispectra with inverse variance weighting. In the absence of spherical symmetry, due to the presence of sky cuts or because of inhomogeneous noise, the optimized estimator is more complicated, and we refer to Munshi & Heavens (2010) for details:

$$
\tilde{C}_l^{2,1} = \frac{1}{f_{sk}} \sum \left[ C_l^{AB,C} - C_l^{A,B,C} - C_l^{B,A,C} - C_l^{(AB),C} \right] + \text{cyc.perm.} 
$$

(42)

The terms without averaging such as $\tilde{C}_l^{AB,C}$ are direct estimates from the observed partial sky with inhomogeneous noise. The Monte Carlo corrections such as $C_l^{A,B,C}$ are constructed by cross-correlating the product of the observed map $A$ and a Monte Carlo map $B$ with a Monte Carlo map $C$, and then taking an ensemble average over many realizations. The denominator $f_{sk}$, which represents the fraction of the sky covered, is introduced to correct for the effect of partial sky coverage. For completeness, the skewness associated with this form of bispectrum can be expressed as a weighted sum of the corresponding $C_l$ values:

$$
\tilde{S} = \sum_i (2l+1) \tilde{C}_l^{2,1} = \sum_{l_1 l_2} \frac{\tilde{B}_{l_1 l_2} B_{l_1 l_2}}{C_l^{l_1} C_l^{l_2}} 
$$

(43)

The $\tilde{B}_{l_1 l_2}$ is the estimated bispectrum and $B_{l_1 l_2}$ is the target model used to generate the weights. Constructing such weighted maps can clearly be seen as a way to construct a matched-filter estimator for the detection of non-Gaussianity. It is optimally weighted by the inverse cosmic variance and achieves maximum response when the observed non-Gaussianity matches with a specific theoretical input. The skew spectrum also allows for the analysis of more than one specific type of non-Gaussianity from the same data – allowing a joint analysis to determine cross-contamination from various contributions.

### 5.2 Bispectra involving line-of-sight integration: the Ostriker–Vishniac effect and its correlation with other secondary anisotropies

Another set of secondary bispectra involving any of the Ostriker–Vishniac (see e.g. Jaffe & Kamionkowski 1998) effect, the SZ thermal effect or the kinetic SZ effect (Cooray 2001b; Castro 2004) or a combination of these has the following form of reduced bispectrum (Cooray & Hu 2000), which involves a line-of-sight integration along $r$:

$$
b_{l_1 l_2 r} = \int dr [f_l(r) g_l(r) + \text{cyc.perm.}] 
$$

(44)

The additional five terms can be constructed from cyclic permutations of the first term. The construction of weighted maps follow the same principle with the use of kernels $f_l(r)$ and $g_l(r)$ that are associated with any of the scattering secondaries that involve a line-of-sight integration. We do not write down the exact expressions for $f_l(r)$ and $g_l(r)$; these can be found in, e.g. Cooray & Hu (2000).

The weights for integrated effects depend on the radial direction. The associated fields that we construct will be three-dimensional to take into account the radial dependence. Defining,

$$
A_{lm}(r) = \frac{a_{lm}}{C_l} f_l(r), \quad B_{lm}(r) = \frac{a_{lm}}{C_l} g_l(r), \quad C_{lm}(r) = \frac{a_{lm}}{C_l}, 
$$

(45)

where $A(\hat{\Omega}, r) = \sum_l A_{lm}(r) Y_{lm}(\hat{\Omega})$ and $B = \sum_l B_{lm}(r) Y_{lm}(\hat{\Omega})$ are fields constructed from the generic functions represented by $f_l$ and $g_l$. Following Munshi & Heavens (2010) we construct

$$
\tilde{C}_l^{2,1}(r) = \frac{1}{2l+1} \int dr d\Omega \sum_m \text{Real} \left\{ (AB)_{lm} C^*_m \right\} + \text{cyc.perm.} = \sum_{l_1 l_2} \frac{\tilde{B}_{l_1 l_2} B_{l_1 l_2}}{C_l^{l_1} C_l^{l_2}} 
$$

(46)

and from this compute the skew-spectrum by carrying out the line-of-sight integration:

$$
\tilde{C}_l^{2,1} = \int dr \tilde{C}_l^{2,1}(r). 
$$

(47)
This is the generalization of the all-sky estimator of the skew spectrum of Munshi & Heavens (2010), but for three distinct fields with a bispectrum given by line-of-sight integration. The linear terms reduce the scatter thereby improving the optimality. They are important in the absence of spherical symmetry which is destroyed because of galactic or point-source mask or non-uniform noise. However, the analysis presented here does not take into account the mode–mode coupling in an accurate way, and uses a $f_{\text{sky}}$ proxy which can only be improved by constructing a completely optimal estimator. We will develop such an estimator in later sections.

In the next section we consider the contamination of the primary skew spectrum by secondary non-Gaussianity from point sources.

5.3 General expression

From the examples above, it is clear that following the ansatz proposed by Smith & Zaldarriaga (2006) the reduced bispectrum $B_{\text{red}}^{XYZ}$ can be decomposed into a sum of terms that are factorizable in $l_1$, $l_2$ and $l_3$ in such a way that the resulting expression can be expressed as

$$ b_{l_1l_2l_3}^{XYZ} = \frac{1}{6} \sum_{i} N_{\text{fact}}^i a_i^l b_i^l c_i^l + \text{cyc.perm.} \tag{48} $$

Here $N_{\text{fact}}^i$ is the total number of terms. The formalism developed in this section can deal with such a generic reduced bispectrum. The procedure involve construction of fields such as $A^0$, $B^0$ and $C^0$ using the weights $a$, $b$ and $c$ (not necessarily of a specific form), a skew-spectrum can always be constructed by a similar manipulation as outlined above. In certain situations the $C_{l}^{2,1}(r)$ might actually also have radial dependence, in which case a line-of-sight integration needs to be performed to match observations and the reduced factors $A^0$ etc. need to be replaced with their functional analogue, e.g. $A^0(r)$:

$$ \hat{C}_{l}^{2,1} = \sum_{l' l''} \sum_{ij} \left( \frac{2l' + 1)(2l'' + 1)}{4\pi} \right) \left( \begin{array}{ccc} l & l' & L \\ 0 & 0 & 0 \end{array} \right) \frac{1}{C_{l}^{XX}} \frac{1}{C_{l'}^{YY}} \frac{1}{C_{l''}^{ZZ}} \left[ A_l B_{l'} C_{l''}^{i} + \text{cyc.perm.} \right] \left[ A_{l'} B_{l''} C_{l}^{i} + \text{cyc.perm.} \right]. \tag{49} $$

Previous results can be seen as special examples of this generalized expression. The formalism developed here for the analysis and estimation of secondary bispectra depends on finding appropriate functional forms for $A_l$, $B_l$ and $C_l$ that can accurately describe a secondary bispectrum. All the specific cases that are of particular interest to us can be decomposed into the above factorizable form. In certain situations the forms $A_l$ can also be a function of the radial distance $r$ along the line of sight. In such situations, the estimator which is a function of $r$ is eventually integrated out to obtain the final estimate.

5.4 Cross-contamination from point sources and primary non-Gaussianity

The bispectra associated with unclustered point sources is modelled as $B_{ps}^{111} = \text{const}$. The constant depends on the flux limit. More complicated modelling, which incorporates certain aspects of halo models, can be used for better accuracy (see e.g. Serra & Cooray 2008):

$$ S = \sum_{l_1 l_2 l_3} \frac{B_{ps}^{111} B_{sec}^{111}}{C_{l_1}^{1} C_{l_2}^{1} C_{l_3}^{1}}. \tag{50} $$

Similarly, given a model of primary non-Gaussianity, one can construct a theoretical model for the computation of $B_{prim}^{111}$ [see Munshi & Heavens (2010) for more about various models and the construction of optimal estimators]. However, the formalism developed here allows us to estimate the level of cross-contamination from secondaries when estimating a specific model of primary non-Gaussianity. If we denote a specific type of primordial non-Gaussianity by $B_{prim}$ and secondary by $B_{sec}$ then the cross-contamination of skewness can be expressed as

$$ S = \sum_{l_1 l_2 l_3} \frac{B_{prim}^{111} B_{sec}^{111}}{C_{l_1}^{1} C_{l_2}^{1} C_{l_3}^{1}}. \tag{51} $$

Similar results hold for the skew spectrum. A more general treatment based on Fisher analysis of multiple bispectra is presented in the subsequent sections.

In addition to various sources mentioned above, second-order corrections to the gravitational potential through gravitational instability too can also act as a source of secondary non-Gaussianity (Munshi, Souradeep & Starobinsky 1995). However, there are no available factorizable approximations as in equation (49) for such non-Gaussianity.

A general comment about the type of estimator that we have considered in this section is in order. The estimator we considered here is virtually optimal in all practical situations where the survey nearly covers the entire sky. However, for a more accurate treatment of noise at high $l$ as well as for an accurate treatment of mode–mode coupling at low-$l$, a robust formalism that works directly with the field harmonics is required. While such a treatment is optimal, it can be rather cumbersome to implement numerically and it can also be computationally expensive. Nevertheless, various factorization schemes exist which can reduce the cost of computation drastically. We will consider the optimal estimators in the next section.

For construction of secondary bispectra the cosmological parameters that we have considered are the same as the ones we have considered in a previous section. The halo model was employed to compute $\beta_l$. The results are plotted in Fig. 4. The resulting skew spectrum is plotted in Figs 5 and 6.
Figure 4. The cross-spectra $\beta_l$ (sometimes also denoted by $b_l$ and not to be confused with the beam) introduced in equation (37) required for the construction of the bispectra is plotted for ISW cross-lensing (middle panel), for SZ cross-lensing (left-hand panel) and for PS cross-lensing (right-hand panel) bispectrum as a function of $l$. See text for details.

Figure 5. The skew-spectra $C_{21}^l$ is plotted as a function of harmonics $l$. Various panels correspond to different choice of mixed bispectra; point source (PS) cross-lensing (right-hand panel), SZ cross-lensing (middle panel) and ISW cross-lensing (left-hand panel). An all-sky coverage $f_{\text{sky}} = 1$ is assumed.

Figure 6. The mixed skewness $S$ is plotted as a function of harmonics $l$ for various choice of mixed bispectra. We plot mixed skewness associated with ISW cross-lensing (right-hand panel), SZ cross-lensing (middle panel) and ISW cross-lensing (right-hand panel) as a function of $l$. As before we have assume all-sky coverage $f_{\text{sky}} = 1$.

6 OPTIMIZED ANALYSIS OF MIXED BISPECTRA: GENERALIZATION OF A PURE ONE-POINT ESTIMATOR TO MIXED ONE- AND TWO-POINT ESTIMATORS

Starting from Babich (2005) a complete analysis of the bispectrum in the presence of partial sky coverage and inhomogeneous noise was developed (see also Babich 2005; Creminelli et al. 2006; Yadav et al. 2008). A specific form for a bispectrum estimator was introduced which is both unbiased and optimal. This was further developed and used by Smith et al. (2007) for lensing reconstruction and by Smith & Zaldarriaga (2006) for general bispectrum analysis and for one-point estimator for $f_{\text{NL}}$. The analysis depends on finding suitable inverse cosmic variance weighting $C^{-1}$ of modes. It deals with mode–mode coupling in an exact way. In a recent work Munshi & Heavens (2010) further extended this analysis by incorporating two-point statistics or the skew spectrum which we have already introduced above. We generalize their results in this work for the case of mixed bispectra for both one-point and two-point studies involving three-way correlations. The analytical results presented here are being kept as much general as possible. However, in the next sections we specialize them to individual cases of lensing reconstruction and the mixed bispectrum associated with lensing and the SZ effect as concrete examples.

The estimator presented here is accurate and can work with arbitrary scan strategies. The implementation, however, can be computationally demanding. This may restrict the $l_{\text{max}}$ to which such a method can be implemented in realistic situations. Even then such estimators can be used at lower $l_{\text{max}}$ to judge the accuracy of other approximate methods. We will consider the estimation of both the skew-spectrum and skewness in this section. The construction of a skewness estimator in the harmonic space is presented in the next subsection. The estimators...
work with harmonics defined over the observed sky. In each case a functional \( \hat{Q} \) is defined, which takes observed harmonics \( \hat{X}, \hat{Y}, \hat{Z} \) as input. Similarly derivatives of \( Q \) with reference to \( \hat{X}, \hat{Y}, \hat{Z} \) are also constructed. These are then used for the construction of the optimized estimator. The estimator also depends on the auto- and cross-covariance of these data vectors. To keep our method very general we consider three different data vectors. These will be important for considering mixed bispectrum involving external tracers that are often used for the study of CMB lensing at the level of bispectrum (internal detection of lensing is only at the level of trispectrum) or bispectrum that can be estimated jointly from the frequency-cleaned SZ maps and CMB maps. These specific cases of the general result will be considered in Sections 7.1 and 7.2.

Another large class of problems that can be studied using the formalism here is that of studying re-ionization using mixed bispectrum involving CMB temperature \((T)\), electric \(E\) (or magnetic \(B\)) polarization and a large-scale tracer field \(S\) (Cooray 2004a,b). The construction of an optimized estimator using the harmonics \( E_{\text{in}} \) (or \( B_{\text{lm}} \), \( \delta T_{\text{in}} \) and \( S_{\text{in}} \) of these fields follows exactly the same procedure outlined below. A detailed analysis will be presented elsewhere.

### 6.1 One-point estimator: mixed skewness \((\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})\Phi^Z(\hat{\Omega}))\)

We are interested in constructing an optimal and unbiased estimator for the estimation of mixed skewness \( \langle \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})\Phi^Z(\hat{\Omega}) \rangle \). The fields \( \Phi^X(\hat{\Omega}) = \sum_{lm} a_{lm}^X Q_{lm}(\hat{\Omega}) \) and similarly for \( \Phi^Y \) and \( \Phi^Z \), are defined over the entire sky, though observed with a mask and non-uniform noise coverage. The non-uniform coverage imprints a mode-mode coupling \( C_{XX,l_{in}m_{in}l_{im}m_{im}}^{1} \) in the observed multipoles of a given field in the harmonic space \( (a_{im}^X, a_{im}^Y) \). For the construction of the optimal estimator it will be useful to define \( \tilde{a}_{lm}^X \) as: \( \tilde{a}_{lm}^X = \sum_{lm}^{1} C_{XX,l_{in}m_{in}l_{im}m_{im}}^{1} a_{lm}^X \). Here \( \tilde{a}_{lm}^X \) represents the harmonics of the data in pixel (real) space \( \Phi^X \) with an inverse covariance weighting. Next we need to deal with the covariance matrix of the modes \( a_{lm}^X \) in terms of that of \( \tilde{a}_{lm}^X \). The autocovariance matrix for \( a_{lm}^X, C_{XX} \), and that of \( \tilde{a}_{lm}^X, \tilde{C}_{XX} \) are related by the following expression: \( \tilde{C}_{XX} = \sum_{lm}^{1} (\tilde{a}_{lm}^X) (\tilde{a}_{lm}^X)^{\dagger} = [C_{XX}]^{1}_{lm,lm} a_{lm}^X a_{lm}^{X\dagger} \). Similarly, the cross-covariance for two different fields \( \tilde{a}^{X} \) and \( \tilde{a}^{Y} \) with inverse variance weighting, in harmonic space can be written as \( \tilde{C}_{XY} = \langle \tilde{a}_{lm}^{X}\tilde{a}_{lm}^{Y}\rangle = [C_{XX}C_{XY}]^{1}_{lm,lm} \).

The estimator that we construct will be based on functions \( Q[\hat{X}, \hat{Y}, \hat{Z}] \) which depends on the input fields, and its derivatives with reference to the fields, e.g. \( \partial Q[\hat{X}, \hat{Y}, \hat{Z}] / \partial \hat{Z}_{lm} \). The derivatives are themselves a map with harmonics described by the free indices \( lm \), and are constructed out of two other maps. The function \( \hat{Q} \) on the other hand is an ordinary number which depends on all the three input functions and lacks free indices:

\[
\hat{Q}[\hat{X}, \hat{Y}, \hat{Z}] = \sum_{lm}^{1} \sum_{l_{in}m_{in}l_{im}m_{im}}^{1} B^{XYZ}_{l_{in}m_{in}l_{im}m_{im}} (l_{in} \ l_{im} \ m_{in} \ m_{im}) \tilde{a}_{lm}^{X\dagger} \tilde{a}_{lm}^{Y} \tilde{a}_{lm}^{Z}.
\]

(52)

Similar expressions hold for other fields such as \( \tilde{a}_{lm}^{X} \), \( \tilde{a}_{lm}^{Y} \). The pre-factor \( \frac{1}{6} \) originates from symmetry considerations. Introducing a more compact notation \( x_{i} \), where \( x_{1} = a_{lm}^{X} \), \( x_{2} = a_{lm}^{Y} \), \( x_{3} = a_{lm}^{Z} \), we can write the one-point estimator for the mixed skewness as

\[
E[\tilde{x}_{i}] = \frac{1}{F} \left\{ Q[\tilde{x}_{i}] - \sum_{i} \langle \tilde{x}_{i} \rangle \langle \partial \tilde{Q}[\tilde{x}_{i}] \rangle \right\}.
\]

(53)

This is the main result of the paper, generalizing work by Smith & Zaldarriaga (2006) to mixed fields. The estimator is similar to that introduced by Munshi & Heavens (2010). There are two contributions. The first term is cubic in the input maps (data vectors) whereas the second term is linear. The second term reflects lack of spherical symmetry due to the presence of an observational mask and an asymmetric noise. Such realistic effects are incorporated in Monte Carlo (MC) realizations. The ensemble averaging \( \langle \tilde{Q}[\tilde{x}_{i}] \rangle \) in the linear terms represents Monte Carlo averaging using simulated non-Gaussian maps. The associated Fisher matrix (a scalar in this case) can be written in terms of the functions \( Q[\tilde{x}_{i}] \), its derivative and the cross-covariance matrices involving different fields. The computation of the Fisher matrix too involves MC realizations with a detailed modelling of noise and sky coverage:

\[
F = \frac{1}{36} \sum_{ij} \sum_{lm_{in}lm_{in}l_{im}lm_{im}} \left\{ \langle \delta_{lm_{in}} \delta_{lm_{in}} \partial_{lm_{in}} \partial_{lm_{in}} Q[\tilde{x}_{i}] \rangle - \langle \delta_{lm_{in}} \delta_{lm_{in}} \partial_{lm_{in}} \partial_{lm_{in}} \rangle \langle \partial_{lm_{in}} \partial_{lm_{in}} Q[\tilde{x}_{i}] \rangle \right\}.
\]

(54)

The Fisher matrix defined here is the inverse of the variance of the estimator \( E \), i.e. \( F^{-1} = \langle (\delta E)^{2} \rangle \) and is a number.

Here we have used the shorthand notation for \( \langle a_{lm}^{X} a_{lm}^{Y} \rangle = \sum_{lm}^{1} C_{XX}^{1}_{lm,lm} \). In case of joint estimation of different bispectra from the same data sets, we can extend the above discussion and write

\[
E[\tilde{x}_{i}] = F^{-1} \left\{ Q[\tilde{x}_{i}] - \sum_{i} \langle \tilde{x}_{i} \rangle \langle \partial \tilde{Q}[\tilde{x}_{i}] \rangle \right\}.
\]

(55)

Here the Fisher matrix \( F_{\alpha\beta}^{-1} = \langle \delta E_{\alpha} \delta E_{\beta} \rangle \) encodes the inverse estimator covariance for two different estimators \( E_{\alpha} \) and \( E_{\beta} \) related to two different mixed bispectra \( B^{X} \) and \( B^{Y} \), that are recovered using the same data sets \( x_{i} \). These bispectra can be different types of secondaries or can also include specific models of primaries. The derivation of the Fisher matrix considers all the possible couplings of the cut-sky harmonics. The term with self-couplings is subtracted using the second term of equation (54). The self-coupling terms are labelled as the \( \beta \)
terms in Munshi & Heavens (2010) and the terms that survive are analogous to the $\alpha$ terms.

$$F_{\alpha\alpha} = \frac{1}{36} \sum_{l_1 l_2 l_3} B_{l_1 l_2 l_3} \left( \begin{array}{ccc} l_1 & l_2 & l_3 \\ l_4 & l_5 & l_6 \end{array} \right) \left( \begin{array}{ccc} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \end{array} \right) \left\{ \bar{C}^{XX} l_{i1} m_{i4} \bar{C}^{YY} l_{i2} l_{i5} \bar{C}^{ZZ} l_{i3} m_{i6} + \text{cyc. perm.} \right\}. \tag{56}$$

The cyclic permutations here represent five additional terms that reflect permutations of the superscripts $X, Y, Z$ along with associated subscripts $l_i$. In general these terms will depend both on the autocovariance matrices and on the cross-covariance among data sets.

The off-diagonal terms of the Fisher matrix $F_{\alpha\beta}$ for the joint estimation of multiple bispectra $\alpha \neq \beta$ from identical data sets can be computed in an analogous manner. The individual cases which we will discuss next can be recovered as various limiting cases from this master result. The Fisher matrix (which is a number in this particular case) for the mixed bispectrum in case of all-sky coverage and constant variance noise can be expressed as

$$F_{\alpha\alpha} = \frac{1}{36} \sum_{l_1 l_2 l_3} B_{l_1 l_2 l_3} B_{l_1 l_2 l_3} \left[ \frac{1}{C_{l_1}^{XX} C_{l_1}^{YY} C_{l_1}^{ZZ}} \delta_{l_1 l_2} \delta_{l_1 l_3} + \text{cyc. perm.} \right]. \tag{57}$$

In addition to the mixed bispectra, the Fisher matrix depends on individual power spectra and the cross-spectra of the data sets involved. The covariance matrix includes both signal and noise. The term that is depicted here depends only on the autospectra of individual data sets. There will be additional terms resulting from the cyclic permutations of the first term. These will depend on the cross-spectra of data sets involved in addition to the autospectra. However, only the first term survives if we assume all the cross-correlations to be weak when computing the error bars from a Fisher analysis. Next, we will consider the special situations where one or more data sets are identical which can be of practical importance. The mixed skewness estimators that we consider here will be generalized to mixed skew-spectra subsequently.

### 6.1.1 Special case (A): $\Phi^2(\hat{\Omega}) = \Phi^Y(\hat{\Omega}), \langle \Phi^X(\hat{\Omega}) \Phi^Y(\hat{\Theta}) \rangle$

In certain practical situations we will encounter cases where two of the three fields are identical, e.g. as we will discuss in a later section, such a situation arises in the context of cross-correlating frequency-cleaned SZ maps against CMB sky. The corresponding Fisher matrix can be recovered by simply setting $\Phi^Z(\hat{\Omega}) = \Phi^Y(\hat{\Theta})$:

$$F_{\alpha\alpha} = \frac{1}{36} \sum_{l_1 l_2 l_3} B_{l_1 l_2 l_3} B_{l_1 l_2 l_3} \left[ \frac{1}{C_{l_1}^{XX} C_{l_1}^{YY} C_{l_1}^{ZZ}} + 4 \left( \frac{C_{l_1}^{XY}}{C_{l_1}^{XX} C_{l_1}^{YY}} \right) \left( \frac{C_{l_1}^{XY}}{C_{l_1}^{XX} C_{l_1}^{YY}} \right) \right], \quad \alpha = XXX. \tag{58}$$

The second term can be ignored while computing the Fisher elements, if the correlation between the two different data sets is weak or $C_{l_1}^{XX} C_{l_1}^{YY} \gg (C_{l_1}^{XY})^2$.

### 6.1.2 Special case (B): $\Phi^2(\hat{\Omega}) = \Phi^Y(\hat{\Theta}) = \Phi^X(\hat{\Omega}), \langle \Phi^X(\hat{\Omega}) \rangle$

Finally, if we identify all the three fields, i.e. $\Phi^Z(\hat{\Omega}) = \Phi^X(\hat{\Theta}) = \Phi^Y(\hat{\Theta})$, we recover an estimator for the skewness (primary or secondary). This is perhaps the most important estimator from observational point of view. It has been used by various groups mainly for the estimation of primordial non-Gaussianity:

$$F_{\alpha\alpha} = \frac{1}{6} \sum_{l_1 m_1} B_{l_1 l_1 l_1} B_{l_1 l_1 l_1} \left[ C^{-1} \right]_{l_1 m_1, l_1 m_1} \left[ C^{-1} \right]_{l_2 m_2, l_3 m_3} \left[ C^{-1} \right]_{l_1 m_3, l_2 m_6}, \quad \alpha = XXX. \tag{59}$$

In the limit of all-sky coverage and constant variance noise, the Fisher error corresponding to the estimator reduces to

$$F_{\alpha\alpha} = \frac{1}{6} \sum_{l_1 l_2 l_3} B_{l_1 l_1 l_1} B_{l_1 l_1 l_1} \left[ C_{l_1}^{XX} C_{l_1}^{YY} C_{l_1}^{ZZ} \right]^{1/2}, \quad \alpha = XXX. \tag{60}$$

For high $l$ a scaling $f_{\text{sky}}^{-1/2}$ is sufficient to describe the effect of partial sky coverage on the error covariance matrix. In the case of bispectrum involving a single field, it is possible to replace the sum over all $l$ described above with a sum that only covers the upper triangular elements of the matrix $B_{l_1 l_1 l_1}$. In case of mixed spectra, $B_{l_1 l_1 l_1}$, the entire range of $l$ needs to be considered for each choice of triplet $XYZ$. We have considered the autocovariance of these estimators, i.e. $\alpha = \beta$ but computing cross-covariances among various estimators $\alpha \neq \beta$ can be done in an analogous manner which will be important for joint estimation of multiple bispectra from identical data sets. Though the one-point estimator compresses all the available information into a single number, it loses sensitivity to the origin of the non-Gaussianity, but it has an advantage of having a maximum signal-to-noise ratio. Being a real-space object the one-point estimators can be estimated in a straightforward way in the pixel space. Note that at the third order the one-point object does not have any Gaussian contribution from the source or from the noise (assumed Gaussian).
6.2 Two-point estimators: mixed skew spectrum \( \langle \Phi^X(\tilde{\Omega})\Phi^Y(\tilde{\Omega}')\Phi^Z(\tilde{\Omega}) \rangle \)

The error-analysis for the power spectra associated with mixed bispectra can be carried out along the same line.

We begin by constructing the functionals \( Q_l \) and its derivative with reference to various input fields. We use these to construct an optimum and unbiased estimator to correlate the field \( \Phi^X(\tilde{\Omega}) \) with the product of two such fields \( \Phi^Y(\tilde{\Omega})\Phi^Z(\tilde{\Omega}) \). We consider the most general possible case of the skew spectrum associated with the mixed bispectrum \( B^{XYZ} \). In contrast to the case of one-point mixed skewness, the functionals \( Q_l \), which are cubic in input fields, are no-longer scalars but depend on the free (not summed over) index \( L \). The skew-spectrum \( \tilde{E}_L \) for a given harmonic \( L \) thus estimated using \( Q_l \) and its derivatives \( \tilde{\alpha}^X_{lm}, \tilde{\alpha}^Y_{lm}, \tilde{\alpha}^Z_{lm} \) are optimum. We will also show here that the various skewnesses defined in the previous sections are a weighted sum of the corresponding skew-spectrum:

\[
\tilde{Q}_L[\tilde{X}, \tilde{Y}, \tilde{Z}] = \frac{1}{6} \sum_m \tilde{\alpha}^X_{LM,m} \sum_{l',m'} B^{XYZ}_{Ll'm'} \left( \begin{array}{ccc} L & l' & l'' \\ M & m' & m'' \end{array} \right) \tilde{a}^Z_{l'm'} \tilde{a}^X_{l''m''},
\]

(61)

\[
\tilde{\alpha}^X_{lm} \tilde{Q}_L[\tilde{Y}, \tilde{Z}] = \delta_{ll'} \frac{1}{6} \sum_{m'} B^{XYZ}_{Ll'm'} \left( \begin{array}{ccc} L & l & l'' \\ M & m & m'' \end{array} \right) \tilde{a}^Z_{l'm'} \tilde{a}^Y_{l'm''},
\]

(62)

\[
\tilde{\alpha}^X_{lm} \tilde{Q}_L[\tilde{X}, \tilde{Z}] = \frac{1}{6} \sum_m \tilde{a}^X_{lm} \sum_{l',m'} B^{XYZ}_{Ll'm'} \left( \begin{array}{ccc} L & l & l' \\ M & m & m' \end{array} \right) \tilde{a}^Z_{l'm}' \tilde{a}^X_{l'm''}.
\]

(63)

While \( Q_l \) is a number (cubic function of input maps) for a given \( L \), the derivatives are maps that are quadratic in the input maps. The derivatives will be important in constructing the linear terms that are important in reducing the variance of the estimator in an absence of spherical symmetry, which is the case in the presence of inhomogeneous noise or partial sky coverage. The derivative with reference to harmonics \( \tilde{a}^X_{lm} \) is different from the one with reference to \( a^X_{lm} \) and \( a^X_{lm} \); this is related to the fact that we are constructing an estimator that cross-correlates the data \( \Phi^X(\tilde{\Omega}) \) with the product field \( \Phi^Y(\Phi^Z)(\tilde{\Omega}) \). The index \( L \), associated with the field \( \Phi^X(\tilde{\Omega}) \), is a free index and is not summed over. Whereas the indices \( lm \) and \( l'm'' \) are summed over.

Using these expressions we can write down the optimized skew-spectrum estimator as

\[
E^{XYZ}_{Ll'm'}[x_i] = [F^{-1}]^{XYZ}_{Ll'm'} \left\{ Q_L[\tilde{x}_i] - \sum_j \tilde{x}_j [\tilde{\alpha}^{l_j}_{l'm'j} Q_L[\tilde{x}_j]]_{LM,j} \right\}.
\]

(64)

The index \( i \) is not summed over and represents all input harmonics \( \tilde{a}^X_{lm}, \tilde{a}^Y_{lm}, \tilde{a}^Z_{lm} \). The normalization is the same as the inverse Fisher matrix \( F^{-1}_{LL'} \). \( Q_L[\tilde{x}_i] \) can be expressed in terms of the covariance matrices among different data sets, \( [C^{-1}_{lm_1,lm_2}]_{ij} \). It also depends on the various derivatives \( \tilde{\alpha}^{l_j}_{l'm'j} \). The averaging involves Monte Carlo simulations of maps that need to be generated assuming a specific correlations structure and a target model for the non-Gaussianity:

\[
[F]_{LL'} = \sum_{l_j} \sum_{m_j} \{ [\tilde{\alpha}^{l_j}_{l'm'j} Q_L[\tilde{x}_j]] [C^{-1}_{lm_1,lm_2}]_{ij} \} \{ [\tilde{\alpha}^{l_j}_{l'm'j} Q_L[\tilde{x}_j]] [C^{-1}_{lm_1,lm_2}]_{ij} \} = \{ [\tilde{\alpha}^{l_j}_{l'm'j} Q_L[\tilde{x}_j]] [C^{-1}_{lm_1,lm_2}]_{ij} \}.
\]

(65)

A detailed derivation of the Fisher matrix here follows a similar technique as used in Munshi & Heavens (2010). The first term in the above expression when worked out will have contributions from terms that represent coupling of modes that appear in the same 3 symbols. These self-coupling terms are called \( \beta \) terms. The beta contributions from the first term will be exactly cancelled by the second term. The remaining surviving terms are called \( \alpha \) terms. These terms signify coupling of modes that appear in different 3 symbols. The \( \alpha \) terms can be further classified in two different categories. Those which involve coupling of free–free terms represented by the appearance of the \( C^{XYZ}_{LM,LM'} \) matrix. They were represented as PP terms in Munshi & Heavens (2010). There are two PP contributions. The first involves only autospectra of various data sets. The second one also includes cross-spectra of \( \Phi^X, \Phi^Y, \Phi^Z \) as well as all three autospectra. The other set of terms are called QQ contributions and involve coupling of a free mode and a summed-over mode – denoted by the covariances \( C^{XYZ}_{LM,LM'} \). There are four such QQ terms all of which depend on a specific cross-covariance and two autocovariance matrices. The difference between PP and QQ contributions do not arise in the Fisher analysis of skewness estimation as all modes are summed over. The Fisher matrix can be finally expressed as

\[
F^{QQ}_{LL'} = \frac{1}{36} \sum_{LM} \sum_{L'M'} \sum_{l_1,m_1} \sum_{l_2,m_2} B^{XYZ}_{Ll_1m_1} B^{XYZ}_{L'l_2m_2} \left( \begin{array}{ccc} L & l_1 & l_1' \\ M & m_1 & m_1' \end{array} \right) \left( \begin{array}{ccc} L' & l_2 & l_2' \\ M' & m_2 & m_2' \end{array} \right) \times \left\{ \left[C^{XX}_{LM,LM'} \right] \left[C^{YY}_{l_1m_1,l_1'm_1'} \right] \left[C^{ZZ}_{l_1m_1,l_1'm_1'} \right] + \left[C^{YY}_{LM,LM'} \right] \left[C^{ZZ}_{l_1m_1,l_1'm_1'} \right] \left[C^{XX}_{l_1m_1,l_1'm_1'} \right] \right\} \times \left\{ \left[C^{YY}_{l_2m_2,l_2'm_2'} \right] \left[C^{ZZ}_{l_2m_2,l_2'm_2'} \right] \left[C^{XX}_{l_2m_2,l_2'm_2'} \right] \right\} \times \left\{ \left[C^{YY}_{l_1m_1,l_1'm_1'} \right] \left[C^{ZZ}_{l_1m_1,l_1'm_1'} \right] \left[C^{XX}_{l_1m_1,l_1'm_1'} \right] \right\}, \quad \alpha = X \otimes YZ.
\]

© 2011 The Authors, MNRAS 414, 3173–3197
Monthly Notices of the Royal Astronomical Society © 2011 RAS
Downloaded from https://academic.oup.com/mnras/article-abstract/414/4/3173/994424 on 29 July 2018
In the case of near-all-sky experiments, the off-diagonal elements of the Fisher matrix are expected to be relatively smaller. The diagonal elements as before can be scaled by \( f_{sk} \) (the fraction of the sky covered by a near-all-sky experiment). The covariance matrices can now be expressed only as a function of the related \( C_l \) values the auto- and cross-correlation power spectra:

\[
F_{LL'}^{uu} = \frac{1}{36} \left\{ \delta_{LL'} \sum_{l,l'} \left[ B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \frac{1}{C_Y^{l'}} + B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \left( \frac{C_Y^{l}}{C_Y^{l'}} \right) \right] \right. \\
\left. + \sum_l B_{Ll}^{XY} B_{L'l'}^{XY} \left( \frac{C_X^{l}}{C_{XX}^{l}} \right) \left( \frac{C_Y^{l}}{C_{YY}^{l}} \right) + B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \left( \frac{C_Y^{l}}{C_{YY}^{l}} \right) \right\}, \quad \alpha = X \otimes YZ.
\]

The error estimates for the other two power spectra \( E_l^{X,Y} \) and \( E_l^{Y,X} \) can also be computed analogously. All three of these estimators will carry complementary information about the mixed bispectra \( B^{XYZ} \). In deriving these expressions, the all-sky limits of \( \hat{C}^{XY}_{\ell_{in},\ell_{in}'} = (1/C_X^{XY}) \delta_{\ell_{in}'} \delta_{\ell_{in}'} \) were used along with the fact that we can write \( C_{\ell_{in},\ell_{in}'} = (C_X^{XY} C_Y^{XY}) \delta_{\ell_{in}'} \delta_{\ell_{in}'} \) for the all-sky case. In our notation \( C_X^{XY} = C_X^{YY} \) as they denote the same cross-spectra. If the cross-correlation among two or more fields vanish, or two or all three fields are identified, the above expression simplifies considerably.

For a joint estimation of several bispectra from the same data, one can write the following expression:

\[
\hat{E}_{\ell}^\alpha[\tilde{x}] = \sum_{L} \sum_{\ell} \left[ \left( N^{-1}\right)^{\alpha}_{LL'} \right] Q_{\ell}^\beta[\tilde{x}] - \sum_{i=1,2,3} \left[ \tilde{x}_i \right] \left( \delta_{\ell_{in}} Q_{\ell}^\beta[\tilde{x}] \right)_{MC}.
\]

Here the indices \( \alpha \) and \( \beta \) correspond to different bispectra \( B_{L,l}^{XY} \) or \( B_{L,l}^{XY} \) which can be jointly estimated using the corresponding skew-spectra \( E_{\ell}^{XY} \) or \( E_{\ell}^{XY} \). Below we consider two special cases for the skew spectra that we have considered so far. The expressions for \( F_{LL'}^{uu} \) and \( N_{LL'}^{\alpha} \) can be obtained simply by replacing the product \( B_{L,l}^{XY} B_{L,l'}^{XY} \) by \( [B_{L,l}^{XY}]^2[B_{L,l'}^{XY}]^2 \). In certain situations when accurate noise modelling is difficult or unlikely, an approximate proxy for \( C^{-1} \) is used in the form of a regularization matrix \( R \) which acts as a smoothing of the data. The resulting data vector \( \tilde{x}^R = \tilde{x} \) is now used for developing an unbiased but suboptimal estimator by replacing \( \tilde{x}_i \) with \( \tilde{x}^R \).

Next we will consider specific cases of interest where two or more fields are identical.

**6.2.1 Special case (A):** \( \Phi^X(\hat{\Omega}) = \Phi^Y(\hat{\Omega}) \langle \Phi^X(\hat{\Omega}) \Phi^Y(\hat{\Omega}) \rangle^X \)

The estimator in this case corresponds to \( E_{\ell}^{X,Y} \). It represents the cross-correlation of a squared field \( \Phi^Y(\hat{\Omega}) \) against another field \( \Phi^X(\hat{\Omega}) \):

\[
F_{LL'}^{uu} = \frac{1}{36} \left\{ 2 \delta_{LL'} \sum_{l,l'} B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \frac{1}{C_Y^{l'}} + \sum_{l} B_{Ll}^{XY} B_{L'l'}^{XY} \left( \frac{C_X^{l}}{C_{XX}^{l}} \right) \left( \frac{C_Y^{l}}{C_{YY}^{l}} \right) \right\}, \quad \alpha = X \otimes YY.
\]

Cross-terms involving \( C_i^{XY} \) contribute to the off-diagonal elements of the Fisher matrix. In case \( \vert C_i^{XY} \vert^2 \ll C_i^{XX} C_i^{YY} \) these terms can be ignored. The expression simplifies a lot if we simply consider only the first term, which depends on the mixed bispectra and the power spectra of individual data sets.

**6.2.2 Special case (B):** \( \Phi^X(\hat{\Omega}) = \Phi^Y(\hat{\Omega}) = \Phi^X(\hat{\Omega}), \langle \Phi^X(\hat{\Omega}) \Phi^X(\hat{\Omega}) \rangle^X \)

If we identify all the three different fields, we recover the Fisher matrix of skew-spectrum or the bispectrum related to a bispectrum. This was introduced by Munshi & Heavens (2010) in the context of primordial bispectrum studies. It can be used to differentiate primordial contributions from secondary contributions or any unknown contribution from incomplete foreground removal or from point-source contamination:

\[
F_{LL'}^{uu} = \frac{1}{36} \left\{ 2 \delta_{LL'} \sum_{l,l'} B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \frac{1}{C_Y^{l'}} + \sum_{l} B_{Ll}^{XY} B_{L'l'}^{XY} \frac{1}{C_X^{l}} \frac{1}{C_Y^{l'}} \right\}, \quad \alpha = X \otimes XX.
\]

The first term contains two identical PP contributions and the next term signifies the four identical QQ contributions.

This particular case is of most importance for the analysis of the primordial bispectrum. For every specific model for primordial non-Gaussianity this estimator can provide an optimal estimate as well as cross-contamination from various secondary sources. It is also useful to study the cross-talk among various models of primordial non-Gaussianity.
6.2.3 Special case (C): $\Phi^X(\hat{\Omega}) = \Phi^X(\hat{\Omega})$, $\langle \Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})\Phi^X(\hat{\Omega}) \rangle$

The estimator we consider here is $E_{L}^{XY}$ which cannot be derived by the simple identification of superscripts. We define the new functions related to the estimators $Q_L$, $\tilde{\alpha}_{lm}^X \hat{Q}_L$ and $\tilde{\alpha}_{lm}^X \hat{Q}_L$ according to the prescription given above:

$$\hat{Q}_L[\tilde{X}, \tilde{Y}] = \frac{1}{6} \sum_M \tilde{\alpha}_{LM}^X \sum_{l,m',m''} B_{Lm' m''}^{XY} \left( L \quad l' \quad l'' \right) \tilde{\alpha}_{l'm'n'}^X \tilde{\alpha}_{l'm'n'}^Y,$$

$$\tilde{\alpha}_{lm}^X \hat{Q}_L[\tilde{X}, \tilde{Y}] = \frac{1}{6} \sum_M \tilde{\alpha}_{LM}^X \sum_{l,m',m''} B_{Lm' m''}^{XY} \left( L \quad l' \quad l'' \right) \tilde{\alpha}_{l'm'n'}^X,$$

$$\tilde{\alpha}_{lm}^X \hat{Q}_L[\tilde{X}, \tilde{Y}] = \frac{1}{6} \sum_M \tilde{\alpha}_{LM}^X \sum_{l,m',m''} B_{Lm' m''}^{XY} \left( L \quad l' \quad l'' \right) \tilde{\alpha}_{l'm'n'}^Y + \delta_{lm} \delta_{Mm} \sum_{l',m'} B_{Ll' m'}^{XY} \left( L \quad l \quad l'' \right) \tilde{\alpha}_{l'm'n'}^X \tilde{\alpha}_{l'm'n'}^Y. \quad (71)$$

The main difference from the previous cases that we have considered so far originates from the fact that the harmonic modes of the field $\Phi^X$ is associated with both free indices as well as the indices that are summed over. The estimator in this case takes the usual form

$$\hat{E}_L[\tilde{x}_L] = \sum_{L'} \left[ F^{-1}_{L L'} \left\{ Q_{L'}[\tilde{x}_L] - \sum_{i=1,2} [\tilde{x}_L, \tilde{x}_M] Q_{L'}[\tilde{x}_M] \right\} \right]. \quad (72)$$

But, the associated Fisher matrix can be expressed in terms of the mixed bispectra and the related power spectra:

$$F_{LL'}^{XY} = \frac{1}{36} \left\{ B_{LL'}^{XY} \left( \frac{C_{LX}^{XX}}{C_{LX}^{XY}} \frac{1}{C_{LX}^{TT}} \frac{1}{C_{LX}^{YY}} \right) + \left\{ B_{LL'}^{XY} B_{LL'}^{XY} \left( \frac{C_{LX}^{XY}}{C_{LX}^{XX}} \frac{1}{C_{LX}^{YT}} \frac{1}{C_{LX}^{YY}} \right) \right\} \right\}$$

$$+ \sum_l \left\{ 2B_{LL}^{XY} B_{LL'}^{XY} \left( \frac{C_{LX}^{XX}}{C_{LX}^{XY}} \frac{1}{C_{LX}^{YT}} \frac{1}{C_{LX}^{YY}} \right) + \frac{1}{C_{LX}^{XX}} \right\} B_{LL'}^{XY} B_{LL'}^{XY} \left( \frac{C_{LX}^{XY}}{C_{LX}^{XX}} \frac{1}{C_{LX}^{YT}} \frac{1}{C_{LX}^{YY}} \right), \quad \alpha = XY \otimes X. \quad (73)$$

The terms that contain a double summation over the indices $l, l'$ are the PP terms described before. They contribute to the diagonal elements of the Fisher matrix in the all-sky limit. The QQ terms contribute to both diagonal and off-diagonal entries. The Fisher matrix in each case for the corresponding estimator can be recovered by summing over the free indices, i.e. $F = \sum_{L'} F_{LL'}$. The one-point estimator will have better signal-to-noise ratio but they are less likely to differentiate among models or detect contamination from an unknown source.

If we ignore correlation between individual data sets, the Fisher matrix can be further simplified. In many previous cases the off-diagonal elements vanished if we ignore the cross-correlation among two data sets. This is not the situation here. This is related to the fact that we are considering cross-correlation between a field $\Phi^X(\hat{\Omega})$ and its product $[\Phi^X(\hat{\Omega})\Phi^Y(\hat{\Omega})]$ with another field. The results that we have considered in this section will be useful for analysing the individual results that we discuss next.

The formalism developed here is quite generic and does not depend on detailed modelling of the target bispectrum. As long as an accurate factorizable model can be found, an optimum estimator can be developed following these principles. In our studies we have only focused here on the scalar fields and their higher order cross-correlation statistics. Generalization to higher order statistics involving higher spin objects such as polarization will be reported elsewhere (Munshi et al., in preparation).

7 SPECIFIC EXAMPLES

The discussion so far has been completely general. We specialize now for a few practical cases of cosmological importance. These correspond to the study of mixed bispectra associated with lensing-induced correlation of secondaries and CMB as well as frequency-cleained SZ catalogues against the CMB sky.

7.1 Lensing reconstruction

7.1.1 One-point estimator

Various estimators associated with lensing reconstruction were introduced by different authors (e.g. Hu 2000; Hu & Okamoto 2002). It was recently studied by Smith et al. (2007) and was used to probe the effect of lensing in CMB by cross-correlating with external data sets such as NVSS survey against WMAP observations:

$$s_{\text{lens}} = \frac{1}{2N} \sum_{l,m} B_{l_1 l_2 l_3}^{dd} \left( \frac{1}{m_1} \frac{1}{m_2} \frac{1}{m_3} \right) \tilde{\delta}_{m_1} \tilde{\delta}_{m_2} \tilde{\delta}_{m_3} \left[ C_{l_1 l_2 l_3}^{\delta T} \tilde{\delta}_{l_1} \tilde{\delta}_{l_2} \tilde{\delta}_{l_3} \right]. \quad (74)$$

Here $\delta_{lm} = \delta T_{lm}$ is the shorthand for CMB temperature multipole. This result is obtained by writing the reconstructed lensing potential in terms of the CMB harmonics and cross-correlating it with low-redshift large-scale tracers such as galaxy surveys (Smith et al. 2007). The bispectrum $B_{l_1 l_2 l_3}^{dd}$ depends, in addition to the $C_l$ values of the CMB multipole, on the cross-correlation between the CMB sky $\delta(\hat{\Omega})$ and the
low-redshift tracer field $\psi(\Omega)$. The reduced bispectrum of interest $b^{\delta\psi}_{\ell_1\ell_2\ell_3}$ and the related form factor $f_{\ell_1\ell_2\ell_3}$ can be written as

\[ b^{\delta\psi}_{\ell_1\ell_2\ell_3} = \{ f_{\ell_1\ell_2\ell_3}C_{\ell_1}^{\delta} + f_{\ell_2\ell_1\ell_3}C_{\ell_2}^{\delta} \} C_{\ell_3}^{\psi}, \]

\[ f_{\ell_1\ell_2\ell_3} = \frac{1}{2} (l_2(l_2 + 1) + l_3(l_3 + 1) - l_1(l_1 + 1)). \]

The multipole $\delta_{\ell_1m_1}$ and $\delta_{\ell_2m_2}$ are associated with the CMB sky and $\psi_{\ell_3m_3}$ is the multipole associated with the LSS tracer at low redshift and hence correlates with the lensing potential (e.g. NVSS survey). The above estimator directly probes the cross-correlation between the lensing potential harmonics $\phi_m$ constructed from temperature harmonics $\delta_m$ and the harmonics of the tracers $\psi_m$. It is interesting to notice that the estimator constructed lacks the term that signifies a correlation between $\delta(\Omega)$ and $\psi(\Omega)$ through the coupling $C^{\delta\psi}$, though the bispectrum itself depends directly on the cross-power spectra. Using the results derived before, we can write the Fisher matrix associated with this estimator as

\[ F = \frac{1}{2} \sum_{\ell_m, \ell_m'} B^{\delta\psi}_{\ell_1\ell_2\ell_3}B^{\delta\psi}_{\ell_1\ell_2',\ell_3'} \left[ \frac{\partial}{\partial \psi_{\ell_m}} \sum_{\ell_m\ell_m'} Q_{\ell_m\ell_m'} \langle \delta_{\ell_m\ell_m'} \rangle \right] \]

As before the inverse Fisher matrix $F^{-1}$ plays the role of a normalization constant. Next, we will generalize these estimators to the case of skew spectrum.

### 7.1.2 Estimators for the skew spectrum

Instead of the one-point estimator described above, we compute the two-point estimator or the skew spectrum as follows:

\[ E_L[\tilde{\delta}, \tilde{\psi}] = \sum_{L'} \left\{ F^{-1}[L,L'] \left[ Q_L[\tilde{\delta}, \tilde{\psi}] - \sum_{\ell_m} \tilde{\psi}_{\ell_m} \langle \delta_{\ell_m\ell_m'} \rangle \right] \right\}. \]

The corresponding expressions for the functions $Q_L[\tilde{\phi}]$ and $\tilde{\phi}_{\ell_m\ell_m'}$ are given by

\[ Q_L = \sum_{L,M} \tilde{\psi}_{L,M} \delta_{m,m'} \left( \frac{L + 1}{L} \right) \left( \frac{m}{m} \right) \left( \frac{m'}{m'} \right), \]

\[ \tilde{\phi}_{\ell_m\ell_m'} = \langle \delta_{\ell_m\ell_m'} \rangle \sum_{l_m' m_m'} B^{\delta\phi}_{\ell_1\ell_2\ell_3} \left( \frac{L + 1}{L} \right) \left( \frac{m}{m} \right) \left( \frac{m'}{m'} \right). \]

The corresponding Fisher matrices that turn out to be diagonal can be written as

\[ F_{LL'} = \sum_{L,M} \frac{1}{2} B^{\delta\psi}_{\ell_1\ell_2\ell_3}B^{\delta\psi}_{\ell_1\ell_2',\ell_3'} \left[ C^{\delta\psi}_{L,M},C^{\delta\psi}_{L,M'} \right], \]

which finally leads us to

\[ F_{LL'} = \frac{1}{2} \sum_{l_m, l_m'} B^{\delta\psi}_{\ell_1\ell_2\ell_3}B^{\delta\psi}_{\ell_1\ell_2',\ell_3'} \left( \frac{1}{C_L} \right) \left( \frac{1}{C_{L'}} \right). \]

A comparison with the previous estimator shows that the presence of off-diagonal entries in the Fisher matrix is related to a direct correlation between data sets. Since we have ignored cross-correlation between the tracer field $\psi$ and the CMB, the Fisher matrix $F_{LL'}$ we recover is entirely diagonal. The power spectra of individual surveys, i.e. $C^{\delta\psi}_L$ or $C^{\delta\psi}_L$, includes noise contributions too. As before we recover $F = \sum_{LL'} F_{LL'}$.

It is possible to work with CMB sky internally, without external data sets (such as NVSS or similar galaxy surveys), to probe weak lensing, e.g. the power spectrum of the lensing potential itself is related to a four-point statistics of the temperature, which makes it noise dominated. The use of external tracers such as galaxy surveys can reduce the problem to three-point level thus lowering the need on sensitivity of the instrument. The discussion above can have direct relevance for use of other tracers such as the one with neutral hydrogen observations (Zahn & Zaldarriaga 2006).

### 7.2 SZ–CMB$^2$ mixed bispectrum

The secondary bispectrum caused by the SZ effect is one of the most pronounced secondary bispectrum among many others (Goldberg & Spergel 1999; Spergel & Goldberg 1999; Cooray & Hu 2000; Cooray 2001a,b, 2002). Following Cooray, Hu & Tegmark (2000) we study if frequency-cleaned maps of all-sky CMB and SZ maps can also be used to construct power spectra associated with the mixed bispectrum with a signal-to-noise ratio that can be detectable with the ongoing CMB experiments. It probes the mode-coupling effects generated by the correlation involved in gravitational lensing angular deflections in CMB and the SZ effects due to large-scale pressure fluctuations. As before the estimator can be constructed from CMB $\tilde{a}_{\ell_m}$ and SZ $\tilde{b}_{\ell_m}$ multipoles. There is a possibility of constructing the correlating the product map $\tilde{\Omega}(\Omega)\tilde{\Omega}(\Omega)$ with $\tilde{a}(\Omega)$ as well as $\tilde{b}(\Omega)$ and $\tilde{s}(\Omega)^2$. In terms of the suboptimal estimators introduced before they correspond to $C^{\delta\psi}_L$ and $C^{\psi\psi}_L$.
respectively. In the second case, analysis is exactly the same as that of lensing reconstruction discussed before. However, in the first case the optimal estimator is expressed as follows:

$$E_{\tilde{L},\alpha}^{\tilde{L},\alpha} = \sum_{\tilde{L}} \langle F^{-1} \rangle_{\tilde{L},\alpha} \left\{ Q_{L}^{\tilde{L}}(\tilde{a}, \tilde{s}) - \sum_{\tilde{l}} \bar{a}_{\tilde{l}} \langle \partial_{\tilde{l}}^{\alpha} Q_{L}^{\tilde{L}}(\tilde{a}, \tilde{s}) \rangle - \sum_{\tilde{l}} \bar{a}_{\tilde{l}} \langle \partial_{\tilde{l}}^{\alpha} Q_{L}^{\tilde{L}}(\tilde{a}, \tilde{s}) \rangle \right\}.$$

(82)

The above estimator is the same as equation (68); with corresponding $Q_{L}^{\tilde{L}}(\tilde{a}, \tilde{s})$ function and its derivatives defined in equation (63). The mixed bispectrum of CMB$^2$-SZ is known to be exactly the same as that of the bispectrum we considered in the lensing reconstruction. This is true not only of the SZ-lensing bispectrum but also of other lensing-induced correlation-related bispectra. The only difference is in the different $b_{\tilde{l}}$ values involved. We quote the result for the Fisher matrix elements below:

$$F_{\tilde{L},\tilde{L}} = \frac{1}{36} \left\{ \sum_{\tilde{l}} \left[ B_{\tilde{L},\tilde{L}}^{\tilde{l}} \left( \frac{1}{C_{L}^{\tilde{l}}} \right) + \sum_{\tilde{l}} \left[ 4 \bar{A}_{\tilde{L},\tilde{L}}^{\tilde{l}} \left( \frac{C_{L}^{\tilde{l}}}{C_{\tilde{L}}^{\tilde{l}}} \right) \left( \frac{C_{L}^{\tilde{l}}}{C_{\tilde{L}}^{\tilde{l}}} \right) \right] \right\}.$$

(83)

This is an application of the case considered in equation (69). For the one-point mixed skewness associated with this power spectra, the related Fisher error is simply given by the sum over all the elements. $F = \sum_{\tilde{L},\tilde{L}} F_{\tilde{L},\tilde{L}}$. The cross-correlations $C_{\tilde{l}}^{\tilde{l}}$ between the two maps $s(\tilde{\Omega})$ and $a(\tilde{\Omega})$ introduces the off-diagonal elements in the Fisher matrix even for the case of all-sky coverage and homogeneous noise. Ignoring the correlations we can recover the Fisher matrix elements derived for the case of lensing reconstruction.

In addition to considering the cross-correlation of $s(\tilde{\Omega})$ and $a(\tilde{\Omega})$ as discussed above, the other estimator of non-Gaussianity that we can consider is the cross-correlation of product field $s(\tilde{\Omega})a(\tilde{\Omega})$, which is the same as the estimator defined in equation (72), with the relevant $Q$ term and its derivative given by equation (71). The associated Fisher matrix is given by equation (73). The one-point estimator recovered from both of these degenerate estimators will be the same.

8 CONCLUSIONS

Extending the previous work (e.g. Hivon et al. 2002) for the estimation of power spectra from correlated data sets, we show how PCL-based approaches can be used for the estimation of cross-correlation power spectra from multiple cosmological surveys through a joint analysis. Analytical results were derived under very general conditions using an arbitrary mask as well as arbitrary noise properties. We also kept the weighting of the data completely general. Our analytical results also include a systematic analysis of the covariance of various deconvolved $C_l$ values characterizing auto- and cross-power spectra from a joint analysis. While PCL-based approaches are known to be unbiased, they are not in general optimal. However, they can be made to act in a near-optimal way by the introduction of weights in different regimes corresponding to signal or noise domination. These studies will be useful in analysing simulated as well as real survey data either in projection or in 3D. We specialize these expressions to recover well-known $f_{sky}$ approximation used in the literature for the error analysis. Using a halo model inspired approach, we compute the expected cross-correlation signal in cross-correlating an NVSS-type survey with the CMB sky through the ISW effect. We also study the cross-correlation between the frequency-cleaned SZ surveys against the ISW effect. The cross-correlation study also provides the covariances among different estimated $C_l$ values and the signal-to-noise ratio of detection for a specific survey. However, we would like to stress that the formalism developed here is more powerful and can tackle many issues in analysing realistic surveys. A detailed study using simulations will be presented elsewhere.

The analysis of the bispectrum is one step beyond the power spectrum and provides additional cosmological information. The primary motivation to date has been to put constraints on early-universe scenarios; however, secondary bispectra can play a significant role in enhancing our understanding of LSS formation scenarios. The secondary bispectrum is mainly related to mode-coupling by secondary effects and lensing. We study various statistics that can directly handle realistic data sets. Extending previous work by Munshi & Heavens (2010), we take into account multiple correlated fields which are used for constructing a mixed skewness at one-point level as well as constructing a skew spectrum at the level of two-point. A very general framework was developed for the study of bispectrum from correlated fields in an unbiased and optimized way. We introduce the inverse covariance weighting and specialize our results for the analysis of bispectrum originating from lensing-secondary correlations. A simple-minded approach that handles the noise and partial sky coverage in a nearly optimal way using Monte Carlo techniques is also discussed. We also develop an approach based on PCL to study the skew spectrum. This approach, whilst suboptimal, can handle the noise and partial sky coverage directly. It is also possible to use weights to make it near optimal in the limit of high $l$, and can be useful mainly because of its speed of handling MC realizations. In its most general form, the estimator $E_L$ (equation 64) for the skew spectrum of mixed fields includes the effect of partial sky coverage and inhomogeneous noise, and provides a compact function that can be compared with theoretical models to identify the source of the correlations between the fields. The associated Fisher matrix (equation 66) allows the statistical analysis of the $E_L$ estimates, allowing the estimation of the relative contributions from different physical processes.

For specific examples we have focused on probing the secondary non-Gaussianity with Planck-type all-sky experiments and surveys such as NVSS. The signal-to-noise ratios for cross-correlation studies involving lensing potential and secondaries such as SZ and ISW would allow detection with Planck. However, to differentiate among various effects one would need to go beyond the cumulative signal-to-noise ratio estimates, but the statistics that we have introduced here will be useful as diagnostic tools.

There has been a lot of work by a number of authors to detect correlations between the WMAP CMB and LSSs, which typically conclude with a constraint on the dark energy (accelerating universe). Analysis of secondary bispectrum has also been attempted. However, at this point, consistent simulations that can correctly take into account the correlation between CMB and LSS, and the impact of LSS on the various observables, still remain to be developed. Though a patchwork of simulations is getting ready, we still lack suitable simulations that can be
used both for cross-correlational analysis and for the entire range of bispectrum analysis at the moment. Our approach can be invaluable in quantifying the accuracy of such consistency check and eventually in putting constraints on cosmology using real high-resolution data. We have not taken into account the errors or residuals from foreground removals. Some of the foreground contaminations may well be correlated to various LSS tracers. These issues and how a PCL-based approach can tackle them will be dealt with elsewhere.

A few comments on the computational costs of various estimators are in order. The reduced bispectrum $b_{l_1l_2l_3}$ is an object with $O(l_{\text{max}}^3)$ associated degrees of freedom. A brute force evaluation of an optimal estimator for skew-spectrum has an associated cost of $O(l_{\text{max}}^5)$. However, the use of factorization ansatz (see e.g. equation (48)) for the reduced bispectrum brings down the cost to $O(N_{\text{side}}l_{\text{max}}^3)$. Using such a factorization it is possible to implement the optimal estimator for WMAP resolution and beyond. The nearly optimal estimator discussed in Section 5 is much faster and limited only by the speed of spherical harmonic transform and scales roughly as $l_{\text{max}}^3$. For bispectra that require construction of weights that depend on the line-of-sight integration, additional computational costs due to the computation of quadrature are nominal. The PCL-based approach is computationally very cheap and can be used for testing pipelines for sanity checks before sophisticated optimal or near-optimal estimators are employed. It also allows the computation of scatter in various estimates and their covariances as discussed in Section 3. For a comparison in terms of the number of pixels in a map in the HEALPix\(^3\) format, the resolution is specified by $l_{\text{side}}$ which can be probed is typically taken to be $l_{\text{max}} = 2N_{\text{side}}$. The formalism developed in this paper has been used recently in a follow-up paper (Calabrese et al. 2010), where the theory of mixed skew-spectrum of the CMB anisotropies, optimized for the detection of the secondary bispectrum, generated by the correlation of the CMB lensing potential and the ISW effect, or the SZ effect, was used on the WMAP 5-yr data.

9 ACKNOWLEDGMENTS

DM was supported by an STFC rolling grant at Royal Observatory at Edinburgh, Institute for Astronomy, when this work was initiated. DM also acknowledges support from STFC standard grant ST/G002231/1 at School of Physics and Astronomy at Cardiff University where this work was completed. It is a pleasure to acknowledge useful exchanges with Jacques Delabrouille, Matthias Bartelmann, Peter Coles and Patricio Vielva Martinez. AC acknowledges support from NSF CAREER AST-0645427. We would like to thank our anonymous referee for constructive comments.

REFERENCES

Acquaviva V., Bartolo N., Matterese S., Riotto A., 2003, Nuclear Phys. B, 667, 119
Afshordi N., 2004, Phys. Rev. D, 70, 083536
Afshordi N., Loh Y., Strauss M. A., 2004, Phys. Rev. D, 69, 083524
Alishahiha M., Silverstein E., Tong T., 2004, Phys. Rev. D, 70, 123505
Arkani-Hamed N., Creminelli P., Mukohyama S., Zaldarriaga M., 2004, J. Cosmol. Astropart. Phys., 04, 001
Babich D., 2005, Phys. Rev. D, 72, 043003
Babich D., Pierpaoli E., 2008, Phys. Rev. D, 77, 123011
Babich D., Zaldarriaga M., 2004, Phys. Rev. D, 70, 083005
Babich D., Creminelli P., Zaldarriaga M., 2004, J. Cosmol. Astropart. Phys., 08, 009
Bartolo N., Matterese S., Riotto A., 2006, J. Cosmol. Astropart. Phys., 06, 024
Boughn S., Crittenden R., 2004, Nat, 427, 45
Boughn S., Crittenden R., 2005, MNRAS, 360, 1013
Boughn S., Crichton R., 2005, New Astron. Rev., 49, 75
Calabrese, Smidt J., Ambland A., Cooray A., Melchiorri A., Serra P., Heavens A., Munshi D., 2010, Phys. Rev. D, 81, 043529
Castro P., 2004, Phys. Rev. D, 67, 044039 (erratum D70, 049902)
Chen G., Szapudi I., 2006, ApJ, 647, L87
Chen X., Huang M., Kachru S., Shiu G., 2007a, J. Cosmol. Astropart. Phys., 01, 002
Chen X., Easther R., Lim E. A., 2007b, J. Cosmol. Astropart. Phys., 06, 023
Cooray A., 2001a, Phys. Rev. D, 64, 043516
Cooray A., 2001b, Phys Rev. D, 64, 063514
Cooray A., 2002, Phys. Rev. D, 62, 103506
Cooray A., 2004a, Phys. Rev. D, 70, 063509
Cooray A., 2004b, Phys. Rev. D, 70, 023508
Cooray A., 2006, Phys. Rev. Lett., 97, 261301
Cooray A. R., Hu W., 2000, ApJ, 534, 533
Cooray A., Seth R., 2002, Phys. Rep., 372, 1
Cooray A., Hu W., Tegmark M., 2000, ApJ, 540, 1
Cooray A., Li C., Melchiorri A., 2008, Phys. Rev. D, 77, 103506
Corasaniti P. S., Giannantonio T., Melchiorri A., 2005, Phys. Rev. D, 72, 023514
Creminelli P., 2003, J. Cosmol. Astropart. Phys., 10, 003

\(^3\) http://healpix.jpl.nasa.gov/

© 2011 The Authors, MNRAS 414, 3173–3197
Monthly Notices of the Royal Astronomical Society. © 2011 RAS
In this appendix we detail the derivation of the error covariance terms for the ordinary cross-spectrum quoted in the paper. We will decompose the entire error covariance matrix into three terms. These three different terms represent signal–signal, noise–noise and signal–noise correlations which can play an important role. The entire derivation will depend on the assumption that the survey covers nearly all of the sky and that the noise we consider is Gaussian:

\[ \delta C_l = \sum_m \delta C_{lm}^s + \sum_n \delta C_{ln}^t \]

and their cross \( \delta C_l^s \delta C_l^t \). We will derive the expressions for these individual covariances in terms of the signal \( C_l \) values and the noise covariance. At intermediate values the cross terms which represent the signal–noise correlations can play an important role. The entire derivation will depend on the assumption that the survey covers nearly all of the sky and that the noise we consider is Gaussian:

\[ \tilde{\delta} C_l = \delta C_l^s + \delta C_l^t \]

The total covariance \( \langle \delta C_l \delta C_l' \rangle \) can be expressed in terms of \( \langle \delta C_l^s \delta C_l^t \rangle \), \( \langle \delta C_l^s \delta C_l^N \rangle \) and their cross \( \langle \delta C_l^s \delta C_l^N \rangle \). We will derive the expressions for these individual covariances in terms of the signal \( C_l \) values and the noise covariance. Clearly, we can write \( \langle \delta C_l \delta C_l' \rangle = \langle \delta C_l^s \delta C_l^s \rangle + \langle \delta C_l^s \delta C_l^t \rangle + \langle \delta C_l^t \delta C_l^t \rangle + \langle \delta C_l^t \delta C_l^N \rangle + \langle \delta C_l^N \delta C_l^N \rangle \).
\( \langle \delta C^0_l \delta C^0_l \rangle + \langle \delta C^0_l \delta C^1_l \rangle + \langle \delta C^1_l \delta C^0_l \rangle \). The first two terms will be referred to as the signal–signal and noise–noise covariances while the last two terms are cross or mixed terms which take contribution from both signal and noise.

A1 Signal–signal

We will start from the definition of \( \tilde{C}_l \) and calculate the covariance \( \langle \delta \tilde{C}_l \delta \tilde{C}_l \rangle \). Our results are valid for arbitrary mask. However, in our derivation we will be using completeness of spherical harmonic basis in which we expand all fields. This means the results are valid for the near-all-sky coverage:

\[
\langle \delta \tilde{C}_l \delta \tilde{C}_l \rangle = \frac{1}{2l + 1} \frac{1}{2l' + 1} \sum_{m} \sum_{m'} \langle \delta a_l^m \delta a_{l'}^{m'} \rangle ;
\]

\[
= [d] \frac{1}{2l + 1} \frac{1}{2l' + 1} \sum_{l,m} \sum_{l',m'} \sum_{l''} \sum_{m''} K_{lm,lm'} K_{l'm',l''m''} \hat{C}_{l'm'} \hat{C}_{l''m''} \delta a_{lm} \delta a_{l'm'} \delta a_{l''m''} .
\]

\( \langle \delta C_l \delta C_l \rangle = \langle \delta \tilde{C}_l \delta \tilde{C}_l \rangle - \langle \tilde{C}_l \rangle \langle \tilde{C}_l \rangle = 2 \frac{1}{2l + 1} \frac{1}{2l' + 1} \sum_{l,m} \sum_{l',m'} K_{lm,lm'} K_{l'm',l''m''} \hat{C}_{l'm'} \hat{C}_{l''m''} C_l C_l .
\]

The \( C_l \) on the right-hand side represents all-sky underlying the theoretical target \( \tilde{C}_l \) values. The pre-factor 2 represents the number of possible permutation of the angular harmonics which avoids self-coupling. Next, using the definition of the mode-coupling matrix equation (3), we will evaluate the following sum involving the mode-coupling matrix \( K \):

\[
\sum_{l,m} K_{lm,lm'} K_{l'm',l''m''} = \sum_{l,m} \int \langle \tilde{C}_l \rangle Y^*_{lm}(\Omega) \langle \tilde{C}_l \rangle Y_{lm'}(\Omega) d\Omega = \int \langle \tilde{C}_l \rangle \overline{\langle \tilde{C}_l \rangle} d\Omega
\]

\[
= \int \delta \Omega \int \delta \Omega' \langle \tilde{C}_l \rangle \overline{\langle \tilde{C}_l \rangle} Y_{lm}^*(\Omega) Y_{lm'}(\Omega') d\Omega'd\Omega' = \int \langle \tilde{C}_l \rangle \overline{\langle \tilde{C}_l \rangle} d\Omega
\]

\[
= \sum_{LM} (-1)^M \langle \delta C^0_L \delta C^0_L \rangle \left( \begin{array}{ccc} l & l' & L \\ -m & m' & M \end{array} \right) \left( \begin{array}{ccc} l & l' & L' \\ 0 & 0 & 0 \end{array} \right) .
\]

The second step of the above derivation relies on the fact that the survey covers nearly all of the sky, so that we can use the orthogonality and completeness of the spherical basis (equation (B1)). The last step involves expanding the squared mask \( \langle \tilde{C}_l \rangle \overline{\langle \tilde{C}_l \rangle} = \sum_{lm} \langle w^2 \rangle \langle \tilde{C}_l \rangle \overline{\langle \tilde{C}_l \rangle} \) in the harmonic basis and expressing the integral involving the product of three spherical harmonic functions in terms of Wigner 3j symbols (Gaunt Integral), i.e. equation (C4). The sum over the product of the other two coupling matrices can also be expressed by introducing a different dummy variable \( L' \), which will be summed over:

\[
\sum_{l,m} K_{lm,lm'} K_{l'm',l''m''} = \sum_{L,M} \langle \delta C^0_L \delta C^0_L \rangle \left( \begin{array}{ccc} l & l' & L \\ -m & m' & M \end{array} \right) \left( \begin{array}{ccc} l & l' & L' \\ 0 & 0 & 0 \end{array} \right) .
\]

In further simplification we will assume the mode-coupling matrix \( K \) to be narrow as it represents a near-all-sky coverage. This will allow us to replace \( C_l \) and \( \hat{C}_l \) by \( \langle \tilde{C}_l \rangle \) and to move them out of the summation. After some straightforward but lengthy algebra, which involves the orthogonality property of the Wigner 3j symbols (equation B2), we can finally write

\[
\langle \delta \tilde{C}_l \delta \tilde{C}_l \rangle = 2 \langle \tilde{C}_l \rangle \frac{1}{3\pi} \sum_{L,M} \langle \delta C^0_L \delta C^0_L \rangle \left( \begin{array}{ccc} l & l' & L \\ 0 & 0 & 0 \end{array} \right) ^2 .
\]

The factor of 2 results from the two different possibilities of coupling various \( a_{lm} \). We have used the orthogonality properties of the Wigner 3j symbols. Here the results are for a single survey.

If we consider two different surveys characterized by the harmonics \( a^X_{lm} \) and \( a^Y_{lm} \) and try to estimate the error covariance for the estimator, \( \langle \delta \tilde{C}^X_{l} \delta \tilde{C}^Y_{l} \rangle \) will involve a four-point correlation in harmonic space \( \langle \delta a^X_{lm} \delta a^X_{lm'} \delta a^Y_{lm''} \delta a^Y_{lm'''} \rangle \) using Wick’s theorem and avoiding a self-coupling term (which needs to be subtracted) we will get the signal–signal contribution with the pre-factor \( 2 C^X_{l} C^Y_{l} \) as given in equation (16). Similarly, the estimator \( \langle \delta \tilde{C}^X_{l} \delta \tilde{C}^X_{l} \rangle \) will have terms with the pre-factors \( C^X_{l} C^X_{l} \) and \( C^X_{l} C^Y_{l} \). The full form of these terms will take into account the integrals involving respective windows for each survey which can be derived as outlined above (equation 15). Finally the signal–signal contribution for the covariance \( \langle \delta \tilde{C}^X_{l} \delta \tilde{C}^Y_{l} \rangle \) will have two identical terms with the pre-factor \( C^Y_{l} C^Y_{l} \); see equation (10). The multiplicative factors for these pre-factors can be computed following the procedure outlined above.

A2 Noise–noise

Next we consider the derivation for the contribution from the noise. We will assume an uncorrelated Gaussian noise. However, the formalism developed here is robust to deal with non-uniform noise:

\[
\langle \delta \tilde{C}^N_{l} \delta \tilde{C}^N_{l} \rangle = \frac{1}{2l + 1} \frac{1}{2l' + 1} \sum_{m} \sum_{m'} \langle \tilde{R}_{lm} \tilde{R}^*_{lm'} \rangle \langle \tilde{R}_{lm} \tilde{R}^*_{lm'} \rangle .
\]
Using the definition of the noise harmonics \( \tilde{n}_{lm} = \int d\Omega \Phi^N(\Omega)w(\Omega)Y_{lm}^*(\Omega) \), the covariance of the noise PCL can be expressed in terms of the mask and the noise characteristics. The noise is assumed to be uncorrelated and Gaussian but non-uniform, \( \langle \Phi^N(\Omega)\Phi^N(\Omega') \rangle = \sigma^2(\Omega)\delta_{\Omega\Omega'}(\Omega - \Omega') \):

\[
\langle \tilde{n}_{lm}\tilde{n}_{lm'}^* \rangle = \int d\Omega \int d\Omega'\langle \Phi^N(\Omega)\Phi^N(\Omega') \rangle w(\Omega)w(\Omega')Y_{lm}(\Omega)Y_{lm'}^*(\Omega') = \int [\sigma^2(\Omega)w^2(\Omega)]Y_{lm}(\Omega)Y_{lm'}^*(\Omega') d\Omega \tag{A11}
\]

\[
= \sum_{LM} [\sigma^2 w^2]_{LM} \int d\Omega Y_{LM}(\Omega)Y_{lm}(\Omega)Y_{lm'}^*(\Omega') 
= \sum_{LM} [w^2\sigma^2]_{LM} \sqrt{\frac{(2l+1)(2l'+1)(2L+1)}{4\pi}} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}. \tag{A12}
\]

Using equation (A10) and the orthogonality property of Wigner 3j functions, we can finally arrive at the following expression

\[
\langle \delta\tilde{C}^N_l \delta\tilde{C}^N_{l'} \rangle = \frac{1}{2\pi} \sum_{LM} [w^2\sigma^2]_{LM} \frac{1}{4\pi} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{A13}
\]

The noise contribution to the covariance matrix depends on the power spectrum of a squared noise map which can be constructed from the original map. The required mask is the squared original mask.

When we are considering the noise contribution to the covariance \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \), we need to consider \( \langle \Phi^N_l \Phi^N_{l'} \Omega \rangle \). The required couplings will all involve cross-correlations of noise in two different surveys which we have assumed to vanish identically \( \langle \Phi^N_l(\Omega)\Phi^N_{l'}(\Omega') \rangle = 0 \), hence there is no noise-only term in equation (16). For the same reason, the only noise contribution to the covariance \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \) will be from the terms involving \( \langle \Phi^N_l(\Omega)\Phi^N_{l'}(\Omega') \rangle \) (equation (10)). On the other hand, the \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \) too has an identical contribution of \( \langle \Phi^N_l(\Omega)\Phi^N_{l'}(\Omega') \rangle \) (equation (15)). This term can be computed using the technique outlined above.

### A3 Signal–noise

In case of the cross-terms we follow the same analysis. There will be two averages, one involving signal multipoles and the other involving noise multipoles. The expressions for both of them are derived above. The final result is obtained on simplifications that involves orthogonality of Wigner 3j symbols:

\[
\langle \delta\tilde{C}^S_l \delta\tilde{C}^S_{l'} \rangle = \frac{1}{2\pi} \sum_{LM} [w^2\sigma^2]_{LM} \frac{1}{4\pi} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}. \tag{A15}
\]

Using the expressions derived above for \( \langle \tilde{N}_{lm}\tilde{N}_{lm'}^* \rangle \) and \( \langle \tilde{n}_{lm}\tilde{n}_{lm'}^* \rangle \) jointly,

\[
\langle \tilde{N}_{lm}\tilde{N}_{lm'}^* \rangle = \frac{1}{2\pi} \sum_{LM} [w^2\sigma^2]_{LM} \frac{1}{4\pi} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}^2 \tag{A16}
\]

Using the definition of the noise harmonics \( \tilde{n}_{lm} = \int d\Omega \Phi^N(\Omega)w(\Omega)Y_{lm}^*(\Omega) \), we can derive the final result. The derivation follows the steps similar to those outlined in the previous subsections, using the orthogonality property of the 3j symbols (equation C2). The final expression is

\[
\langle \delta\tilde{C}^S_l \delta\tilde{C}^N_{l'} \rangle = \frac{1}{2\pi} \sqrt{C_l C_{l'}} \sum_{LM} [w^2(\Omega)]_{LM} [w^2(\Omega)]_{LM} \frac{1}{4\pi} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{A17}
\]

When computing the signal–noise cross-terms for the error covariance in two different surveys, it is easy to see, due to the lack of correlation among the noise in different surveys, that both the covariances \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \) and \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \) will have two different cross-terms, one involving signal component from survey X characterized by \( \sqrt{C_l C_{l'}}C_l^{XX}C_{l'}^{YY} \) and noise component from Y and the other involving the opposite combination, i.e. noise component from X and signal from survey Y characterized by \( \sqrt{C_l C_{l'}}C_l^{YY}Y_{l'}^{XX} \). Finally the covariance \( \langle \delta\tilde{C}^{XX}_l \delta\tilde{C}^{YY}_{l'} \rangle \) will have only one cross term that involves a signal component from the cross-correlation of survey X and Y characterized by \( \sqrt{C_l C_{l'}}C_l^{XX}C_{l'}^{YY} \) and a noise component from survey Y.

### APPENDIX B: SPHERICAL HARMONICS

The completeness relationship for the spherical harmonics is given by

\[
\sum_{lm} Y_{lm}(\hat{\Omega})Y_{lm}^*(\hat{\Omega}) = \delta_{2D}(\hat{\Omega} - \hat{\Omega'}). \tag{B1}
\]
The orthogonality relationship is as follows:
\[
\int d\Omega \ Y_{lm}(\hat{\Omega}) Y_{l'm'}^{*}(\hat{\Omega}) = \delta^K_{ll'} \delta^K_{mm'}. \tag{B2}
\]

**APPENDIX C: 3j SYMBOLS**

The following properties of 3j symbols were used to simplify various expressions:

\[
\sum_{l_3,m_3} (2l_3 + 1) \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l \\ m_1' & m_2' & m \end{pmatrix} = \delta^K_{m_1 m_1'} \delta^K_{m_2 m_2'}, \tag{C1}
\]

\[
\sum_{m_1m_2} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3' \\ m_1 & m_2 & m_3' \end{pmatrix} = \frac{\delta^K_{l_1 l_2} \delta^K_{m_1 m_3}}{2l_3 + 1}, \tag{C2}
\]

\[
(-1)^m \begin{pmatrix} l & l & l' \\ m & -m & 0 \end{pmatrix} = \frac{(-1)^l}{\sqrt{(2l + 1)}} \delta^K_{l' 0}. \tag{C3}
\]

\[
\int d\Omega Y_{lm}(\hat{\Omega}) Y_{l'm'}^{*}(\hat{\Omega}) Y_{LM}(\hat{\Omega}) = \sqrt{\frac{(2l + 1)(2l' + 1)(2L + 1)}{4\pi}} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} l & l' & L \\ 0 & 0 & 0 \end{pmatrix}. \tag{C4}
\]

This paper has been typeset from a TeX/LaTeX file prepared by the author.