Needlet estimation of cross-correlation between CMB lensing maps and LSS

Federico Bianchini,[a,e,d,1] **Alessandro Renzi,**[b,c] **Domenico Marinucci**[b,c]

[a]Astrophysics Sector, SISSA, Via Bonomea 265, I-34136 Trieste, Italia

[b]Dipartimento di Matematica, Universitá di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italia

[c]INFN, Sezione di Roma 2, Universitá di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italia

[d]INFN - Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italy

[e]INAF - Osservatorio Astronomico di Trieste, via Tiepolo 11, 34131, Trieste, Italy

E-mail: fbianchini@sissa.it, renzi@mat.uniroma2.it, marinucc@mat.uniroma2.it

Abstract. In this paper we develop a novel needlet-based estimator to investigate the cross-correlation between cosmic microwave background (CMB) lensing maps and large-scale structure (LSS) data. We compare this estimator with its harmonic counterpart and, in particular, we analyze the bias effects of different forms of masking. In order to address this bias, we also implement a MASTER-like technique in the needlet case. The resulting estimator turns out to have an extremely good signal-to-noise performance. Our analysis aims at expanding and optimizing the operating domains in CMB-LSS cross-correlation studies, similarly to CMB needlet data analysis. It is motivated especially by next generation experiments (such as Euclid) which will allow us to derive much tighter constraints on cosmological and astrophysical parameters through cross-correlation measurements between CMB and LSS.

---

[1]Corresponding author.

Contents

1 Introduction

One of the main puzzles of modern cosmology is the understanding of the mechanism that sources the late-time accelerated expansion of the Universe. Whether it is associated to an exotic form of energy or to some modifications of general relativity, the different scenarios can only be disentangled by probing the perturbations evolution over cosmic time. In this context, galaxy clustering and weak gravitational lensing have become promising probes not only to investigate cosmic acceleration but also the dark matter and neutrino sectors.

While the analysis of the data from the Planck satellite is approaching to an end, yielding a breakthrough in many respects for what concerns CMB studies [1], such fundamental issues have triggered the upcoming experimental efforts and in the next few years galaxy surveys such as the European Space Agency's (ESA) satellite Euclid[1] [2], the Dark Energy Spectroscopic Instrument (DESI[2]), the Large Synoptic Survey Telescope (LSST[3]) and the Wide Field Infrared Survey Telescope (WFIRST[4]), along with a plethora of ground-based high-sensitivity CMB experiments like the Simons Array[5], the South Pole Telescope (SPT-3G)[6], and the Advanced Atacama Cosmology Telescope (AdvACT)[7], will carry out observations devoted to shed light on the physics behind the dark components. In these experiments, operating and under design and construction towards the efforts of the next decade (including ground-based facilities such as the Simons Observatory[8] and CMB-S4, as well as the proposed space satellites COrE[9] e LiteBIRD[10]), the role of CMB-LSS cross correlation is double: on

---

[1]http://sci.esa.int/euclid/
[2]http://desi.lbl.gov
[3]http://www.lsst.org
[4]http://wfirst.gsfc.nasa.gov
[5]http://cosmology.ucsd.edu/simonsarray.html
[6]https://pole.uchicago.edu/spt/
[7]https://act.princeton.edu
[8]https://simonsobservatory.org
[9]http://www.core-mission.org
[10]http://litebird.jp/eng/

one side, yielding constraints on dark energy and matter through the analysis of CMB lensing by forming LSS, and on the other, to de-lens the B-modes of polarization in order to improve the constraint, or measure, of the power from primordial gravitational waves.

In particular, LSS data gathered from Euclid in the form of weak lensing and galaxy catalogues will provide an excellent tracer for the underlying gravitational potential which is responsible for the CMB lensing effect. It is then only natural to cross-correlate CMB lensing maps with LSS data to improve the constraints on dark energy models and cosmological parameters, similarly to what has been done with CMB temperature and LSS maps in order to extract faint large scale signal like the integrated Sachs-Wolfe effect (iSW), see for instance [3–8].

CMB lensing-galaxy cross-correlation measurements can be used to reconstruct the redshift evolution of the galaxy bias [9–11], the growth of structures [12] and to augment the absolute cosmic shear calibration [13]. All analyses reported to date have reconstructed the 2-point statistics either in harmonic or real space.

In this paper, we shall use instead a procedure based on needlet ideas, and we shall discuss how to improve its performance for very aggressive masks using a MASTER-like correction.

As discussed in many previous references, needlets are a form of spherical wavelets which were introduced in functional analysis and statistics by [14, 15] and have then found a number of different applications in the cosmological community over the last decade; we recall for instance [16] for a general description of the methods, [17–22] for non-Gaussianity estimation, [23–27] for foreground component separation, [22, 28, 29] for polarization data analysis, [30, 31] for extension in 3d framework and [21, 32] for trispectrum analysis.

The advantages of needlets, like those of other wavelets system, have been widely discussed in the literature; in short, they are mainly concerned with the possibility to exploit double localization properties, in the real and harmonic domain. Despite this localization in the real domain, we show here that the performance of a needlet cross-correlation estimator deteriorates badly in the presence of very aggressive sky-cuts (i.e., experiments with sky coverage much smaller than 50%). In this paper, we show how the performance of this estimator can be greatly improved by a MASTER-like correction. Thus achieving signal-to-noise figure of merits which are in some aspect superior to the corresponding results for power spectrum methods.

The plan of the paper is as follows. In section 2 we review quickly some background material on both harmonic and needlet cross-correlation analysis; we then proceed in section 3 to introduce the MASTER-like algorithm for the needlets cross-correlation estimator. Numerical evidence and some comparison on the performance of these procedures are collected in section 4, while final considerations are presented in section 5.

2 Building the cross-correlation estimators

In this section we introduce the CMB lensing-galaxy cross-correlation estimators in harmonic and needlet space. We start by briefly reviewing the theoretical framework that we exploit to model such signal and show how the weak lensing of the CMB is correlated to the large scale matter distribution. Then, starting from these concepts, we illustrate how to build an harmonic estimator of the CMB lensing-galaxy cross-correlation signal and finally we derive a needlet estimator from the harmonic one. Even though the main focus of the paper is

the measurement of the CMB lensing-galaxy cross-correlation, we recall that the estimators presented here can be applied to any scalar field on the sphere.

2.1 The weak lensing of the CMB from LSS

Gravitational lensing performs a remapping of the primordial CMB temperature and polarization anisotropies by a deflection field $\mathbf{d}(\hat{\mathbf{n}})$, so that photons coming from direction $\hat{\mathbf{n}}$ carry information about the patch of the sky in the perturbed direction $\hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}})$ (see [33]). The deflection field can be written as the gradient of a scalar potential, namely the CMB lensing potential $\phi(\hat{\mathbf{n}})$, which encodes information about the Weyl potential [11] integrated out to the last-scattering surface. Here we work in terms of the (spherical) Laplacian of the lensing potential, the CMB convergence field[12] $\kappa(\hat{\mathbf{n}}) = -\Delta_{S^2}\phi(\hat{\mathbf{n}})/2$ which describes the local (de)magnification of CMB fluctuations, while the Laplacian reads as $\Delta_{S^2} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}$. As concerns the tracer galaxies we define the projected galaxy density fluctuations as $g(\hat{\mathbf{n}}) = n(\hat{\mathbf{n}})/\bar{n} - 1$, where $n(\hat{\mathbf{n}})$ is the number of objects in a given direction, and $\bar{n}$ is the mean number of sources. In standard cosmologies both the CMB convergence and the galaxy overdensity can be written as a weighted integral of the matter overdensity $\delta(\hat{\mathbf{n}})$ along the line-of-sight (LOS):

$$X(\hat{\mathbf{n}}) = \int_0^{z_*} dz\, W^X(z)\delta(\chi(z)\hat{\mathbf{n}}, z), \tag{2.1}$$

where $X = \{\kappa, g\}$ and $W^X(z)$ is the kernel related to a given field.
The kernel W^κ quantifies the matter distribution lensing efficiency and it reads

$$W^\kappa(z) = \frac{3\Omega_m}{2c}\frac{H_0^2}{H(z)}(1+z)\chi(z)\frac{\chi_* - \chi(z)}{\chi_*}. \tag{2.2}$$

Here $H(z)$ is the Hubble factor at redshift z, $\chi(z)$ is the comoving distance to redshift z, χ_* is the comoving distance to the last scattering surface at $z_* \simeq 1090$, c is the speed of light, Ω_m and H_0 are the present-day values of matter density and Hubble parameter, respectively. Under the hypothesis that luminous matter traces the peaks of the underlying dark matter field, we write the observed projected galaxy overdensity as the sum of an intrinsic clustering term and a lensing magnification bias one, so that the galaxy kernel reads

$$W^g(z) = b(z)\frac{dN}{dz} + \mu(z). \tag{2.3}$$

The former term describes the physical clustering of the sources and is given by the product of the bias factor b with the *unit-normalized* redshift distribution of galaxies, dN/dz. The latter is related to the lensing magnification bias and it writes:

$$\mu(z) = \frac{3\Omega_m}{2c}\frac{H_0^2}{H(z)}(1+z)\chi(z)\int_z^{z_*} dz'\left(1 - \frac{\chi(z)}{\chi(z')}\right)(\alpha(z') - 1)\frac{dN}{dz'}. \tag{2.4}$$

Magnification bias is independent of the tracer bias parameter and, in the weak lensing limit, depends on the slope of the galaxy number counts α $(N(> S) \propto S^{-\alpha})$ at the flux density limit of the survey.

---

[11]Here we define the Weyl potential as $(\Psi + \Phi)/2$, half the sum of the two Bardeen potentials Ψ and Φ.
[12]This relation translates in harmonic space into $\kappa_{\ell m} = \frac{\ell(\ell+1)}{2}\phi_{\ell m}$.

At smaller angular scales ($\ell \gtrsim 20$), the Limber approximation [34] allows us to relate the theoretical two-point statistics of the CMB convergence-galaxy and galaxy-galaxy correlations to the matter power spectrum $P(k, z)$ through:

$$C_\ell^{\kappa g} = \int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} W^\kappa(z) W^g(z) P\left(\frac{\ell}{\chi(z)}, z\right);$$
$$C_\ell^{gg} = \int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} [W^g(z)]^2 P\left(\frac{\ell}{\chi(z)}, z\right). \tag{2.5}$$

We calculate the matter power spectrum using the CAMB[13] code [35], including the effect of non-linear matter clustering via the common Halofit prescription [36].

2.2 Harmonic cross-correlation estimator

Most of the cosmological observations, from CMB to galaxy surveys, provide us with data in the form of two-dimensional sky-maps.[14] The information content hidden in such maps is usually probed by means of harmonic analysis on the sphere. A popular observable that characterizes the statistical properties of a given cosmic field is the angular power spectrum C_ℓ and its reconstruction enables a direct comparison between models and data.

It is common practice to decompose the observed field $X(\hat{\mathbf{n}})$ into spherical harmonics, a frequency-space orthonormal basis for representing functions defined over the sphere, as

$$X(\hat{\mathbf{n}}) = \sum_{\ell m} x_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{2.6}$$

where the spherical harmonic coefficients are given by

$$x_{\ell m} = \int_{\mathbb{S}^2} X(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}) d\Omega. \tag{2.7}$$

For an isotropic finite variance field we have that the mean of the spherical harmonic coefficients is $\langle x_{\ell m} \rangle = 0$, while their covariance is given by $\langle x_{\ell m} x_{\ell' m'}^* \rangle = C_\ell^{XX} \delta_{\ell \ell'} \delta_{mm'}$. An unbiased estimator of the (cross-)angular power spectrum is given by (hereafter the hat $\hat{X}$ denotes estimated quantities):

$$\hat{C}_\ell^{XY} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} x_{\ell m} y_{\ell m}^*. \tag{2.8}$$

In particular, $\hat{C}_\ell$ can be shown to possess the minimal variance among the unbiased estimators (in the sense that its variance reaches the Cramèr-Rao lower bound).

Spherical harmonics are particularly appealing because they are statistically orthogonal for full-sky Gaussian-distributed sky-maps, i.e. the covariance is diagonal $\mathrm{Cov}_{\ell \ell'} \propto \delta_{\ell \ell'}$, and the power spectrum fully characterizes the behaviour of the field. However, real-world observations have to deal with a number of limitations and issues, such as the finite instrumental spatial resolution, the anisotropic noise, and asymmetric beam response. Moreover the incomplete sky coverage, motivated for example by foreground contamination or the instruments

---

[13]http://cosmologist.info/camb/

[14]This especially applies when distance information about the sources is unavailable, nevertheless the quantity of interest can always be projected on the sphere.

scanning strategy, induces a mode-coupling and a power leakage between different multipoles, as well as an overall downward shift of power [37, 38]. This makes the exact evaluation of the following spherical harmonic transform cumbersome:

$$
\begin{aligned}
\tilde{x}_{\ell m} &= \int_{\mathbb{S}^2} X(\hat{\mathbf{n}}) W(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}) d\Omega \\
&= \sum_{\ell' m'} K_{\ell m \ell' m'}[W] x_{\ell' m'},
\end{aligned}
\tag{2.9}
$$

where the kernel K, dependent on the weighting scheme $W(\hat{\mathbf{n}})$ (i.e. the mask), describes the induced mode-coupling. A common approach to obtain unbiased but slightly sub-optimal bandpower estimates is to use approximate pseudo-C_ℓ (PCL) methods us the well known MASTER (see i.e., [37, 38]). It is possible to show that the pseudo-spectrum $\tilde{C}_\ell^{XY} = (2\ell + 1)^{-1} \sum_m \tilde{x}_{\ell m} \tilde{y}_{\ell m}^*$ is related to the underlying power spectrum C_ℓ as

$$
\langle \tilde{C}_\ell^{XY} \rangle = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}^{XY},
\tag{2.10}
$$

where $M_{\ell\ell'}$ is the coupling matrix as defined in [37]. The basic idea is to invert eq. 2.10 in order to recover the underlying power spectrum, however for small sky fraction $f_{\text{sky}} = \frac{1}{4\pi} \int_{\mathbb{S}^2} W^2(\hat{\mathbf{n}}) d\Omega$, one needs to bin the pseudo-power spectrum and the coupling matrix, so that the estimator of the true cross-bandpowers $\hat{C}_L^{XY}$ writes

$$
\hat{C}_L^{XY} = \sum_{L'\ell} K_{LL'}^{-1} P_{L'\ell} \tilde{C}_\ell^{XY},
\tag{2.11}
$$

where L is the bandpower index and the binned coupling matrix can be written as

$$
K_{LL'} = \sum_{\ell\ell'} P_{L\ell} M_{\ell\ell'} B_{\ell'}^2 Q_{\ell'L'}.
\tag{2.12}
$$

Here $P_{L\ell}$ is the binning operator, $Q_{\ell L}$ is its reciprocal, and $B_{\ell'}^2$ is the pixel window function that corrects for the finite pixel size. If the true power spectrum varies slowly with respect to the coupling matrix and/or f_{sky} is large, eq. 2.10 becomes

$$
\langle \tilde{C}_\ell^{XY} \rangle \approx C_\ell^{XY} \sum_{\ell'} M_{\ell\ell'} = f_{\text{sky}} C_\ell^{XY},
\tag{2.13}
$$

which is the so-called f_{sky} approximation [39].

2.3 Needlet cross-correlation estimator

As mentioned in the introduction, some drawbacks of standard Fourier analysis on the sphere can be mitigated by the exploitation of needlet/wavelet techniques. Related advantages have already been widely discussed in the literature, see again [16, 17, 20, 30, 32, 40].

Here we simply recall that the spherical needlet system $\psi_{\{jk\}}$ can be obtained by a quadratic combination of spherical harmonics as

$$
\psi_{jk}(\hat{\mathbf{n}}) = \sqrt{\lambda_{jk}} \sum_{\ell=[B^{j-1}]}^{[B^{j+1}]} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell m}(\xi_{jk}),
\tag{2.14}
$$

where $[\cdot]$ denotes the integer part, $b(\cdot)$ is the filter function in the harmonic domain defined for $x \in [1/B, B]$, and $\{\xi_{jk}\}$ are the cubature points on the sphere corresponding to the frequency j and the location k. Since our implementation relies on the HEALPix[15] [41] pixelation scheme
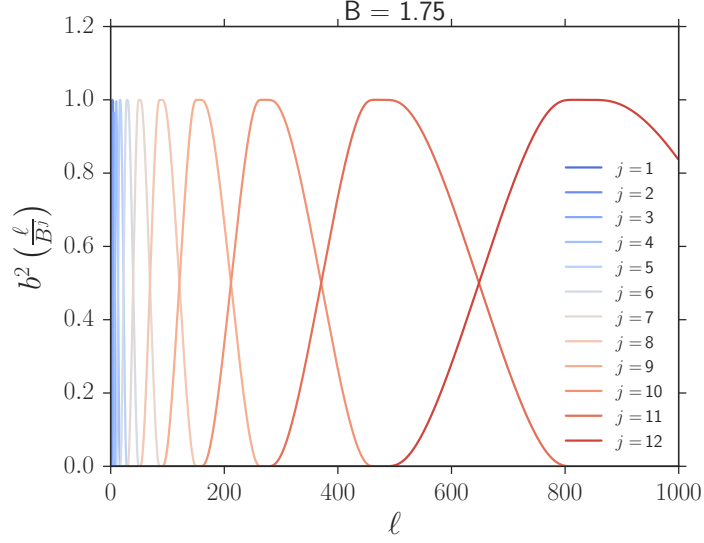
---

[15] http://healpix.jpl.nasa.gov

Figure 1. Profile of the filter function in the ℓ-space for different needlet frequencies j. The needlet width parameter is set to $B = 1.75$.

we can identify the cubature points with the pixel centers, so that the cubature weights λ_{jk} can be approximated by $4\pi/N_{\rm pix}$, where $N_{\rm pix}$ is the number of pixels for the chosen HEALPix $N_{\rm side}$ resolution and k represents the pixel number [3].

Needlets can be thought of as a convolution of the projection operator $\sum_m Y_{\ell m}^*(\hat{\mathbf{n}})Y_{\ell m}(\xi_{jk})$ with a filter function $b(\cdot)$ whose width is controlled by the only free parameter B: recipes for the construction of the function $b(\cdot)$ can be found in [16, 40, 42]. A smaller value of B corresponds to a narrower localization in ℓ-space, while a larger value translates into a more precise localization in real space. Once B is fixed, each needlet can be shown to pick up signal only from a specific range of multipoles determined by the index j: the profile of the filter function $b(\cdot)$ is shown in Fig. 1 for different frequencies. Needlet coefficients are then evaluated by projecting the centered field $X(\hat{\mathbf{n}})$ on the corresponding needlet $\psi_{jk}(\hat{\mathbf{n}})$ as

$$
\begin{aligned}
\beta_{jk} &= \int_{\mathbb{S}^2} X(\hat{\mathbf{n}})\psi_{jk}(\hat{\mathbf{n}})d\Omega \\
&= \sqrt{\lambda_{jk}} \sum_{\ell=[B^{j-1}]}^{[B^{j+1}]} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} x_{\ell m}Y_{\ell m}(\xi_{jk}).
\end{aligned}
\tag{2.15}
$$

Needlet coefficients corresponding to a given frequency j can themselves be represented as an HEALPix map. It is worth to stress that although needlets do not make up an orthonormal basis for square integrable functions on the sphere, they represent a tight frame (redundant basis) so that they allow for a simple reconstruction formula. After computing the needlet coefficients β_{jk} from the maps, we can build a spectral estimator as

$$
\hat{\beta}_j^{XY} = \frac{1}{N_{\rm pix}} \sum_k \beta_{jk}^X \beta_{jk}^Y,
\tag{2.16}
$$

and it is immediate to check that it provides an unbiased estimate of (a binned form of) the

angular power spectrum, i.e.

$$\langle \hat{\beta}_j^{XY} \rangle \equiv \beta_j^{XY} = \sum_\ell \frac{2\ell+1}{4\pi} b^2\left(\frac{\ell}{B^j}\right) C_\ell^{XY}. \tag{2.17}$$

These theoretical predictions can directly be compared to the extracted spectra, allowing for the parameter extraction process. Moreover, as noted in [3], the analytic relation between β_j and C_ℓ makes straightforward dealing with beam profiles, pixel window function, and experimental transfer functions. Note that in this paper we divide the spectral estimator 2.16 and its expected value 2.17 for a normalizing factor $\mathcal{N}$ given by

$$\mathcal{N} = \sum_\ell \frac{2\ell+1}{4\pi} b^2\left(\frac{\ell}{B^j}\right), \tag{2.18}$$

so that in the plots we show $\hat{\beta}_j^{XY} \to \hat{\beta}_j^{XY}/\mathcal{N}$.
The theoretical variance of the cross-correlation power spectrum in needlet space reads

$$(\Delta\beta_j^{XY})^2 \equiv \mathrm{Var}[\hat{\beta}_j^{XY}] = \sum_\ell \frac{2\ell+1}{16\pi^2} b^4\left(\frac{\ell}{B^j}\right)\left[(C_\ell^{XY})^2 + C_\ell^{XX}C_\ell^{YY}\right], \tag{2.19}$$

where the angular auto-spectra can be comprehensive of a noise term, i.e. $C_\ell \to C_\ell + N_\ell$, if present. Moreover, the needlets system is compactly supported in the harmonic domain and as such, for full-sky maps, the random needlets coefficients are uncorrelated by construction for $|j - j'| \geq 2$ [15].

3   MASTER algorithm for needlets

As mentioned in Sec 2.3, one of the main driver behind the development of the needlet spectral estimator is the need to overcome the issues related to Fourier analysis on the sphere in the presence of missing observation. The excellent needlets localization properties in real space represent a key feature for analyzing cosmological data on the partially observed sky, in particular it has been shown that even in the presence of masked regions the random needlet coefficients β_{jk} are asymptotically independent (over k) as $j \to \infty$ (contrary to the case of random coefficients $x_{\ell m}$) [15, 16]. However, as we shall see from simulations in the next section, the estimator defined in eq. 2.16 becomes biased for aggressive masking: here we formally study the effect of sky-cuts on the needlet power spectrum estimation.
From eq. 2.9, we find that needlet coefficients computed on a masked sky are given by

$$\tilde{\beta}_{jk} = \sqrt{\lambda_{jk}} \sum_\ell b\left(\frac{\ell}{B^j}\right) \sum_m \tilde{x}_{\ell m} Y_{\ell m}(\xi_{jk}). \tag{3.1}$$

Then, if we consider the statistic

$$\hat{\Gamma}_j^{XY} = \frac{1}{N_{\mathrm{pix}}} \sum_k \tilde{\beta}_{jk}^X \tilde{\beta}_{jk}^Y, \tag{3.2}$$

it is straightforward to see that its expectation value reads as follows

$$\langle \hat{\Gamma}_j^{XY} \rangle \equiv \Gamma_j^{XY} = \sum_{\ell m} b^2\left(\frac{\ell}{B^j}\right) \langle \tilde{x}_{\ell m} \tilde{y}_{\ell m} \rangle$$

$$= \sum_{\ell m} \sum_{\ell' m'} b^2\left(\frac{\ell}{B^j}\right) K^2_{\ell m \ell' m'}[W] C_{\ell'} \qquad (3.3)$$

$$= \sum_{\ell \ell'} \frac{2\ell + 1}{4\pi} b^2\left(\frac{\ell}{B^j}\right) M_{\ell \ell'} C_{\ell'},$$

which tells us that $\hat{\Gamma}_j^{XY}$ is an unbiased estimator for a smoothed version of the pseudo angular power spectrum $\tilde{C}_\ell$, similar to the case of $\hat{\beta}_j^{XY}$: in some sense, we can view $\hat{\Gamma}_j^{XY}$ as an estimator of the pseudo-needlet power spectrum. Using eq. 2.13, which is valid for slowly varying power spectra and/or large sky fractions, it is possible to relate the two estimators as

$$\langle \hat{\Gamma}_j^{XY} \rangle \approx f_{\text{sky}} \sum_{\ell} \frac{2\ell + 1}{4\pi} b^2\left(\frac{\ell}{B^j}\right) C_\ell$$

$$= f_{\text{sky}} \langle \hat{\beta}_j^{XY} \rangle. \qquad (3.4)$$

Before we conclude this section let us introduce a couple of remarks. We recall first that in the case of a survey with a large sky-cut, inverting the full coupling matrix become unfeasible because of singularities; hence the power spectrum can be estimated only over some subset of multipoles i.e. the power spectrum is recovered only up to some frequency windows. As discussed earlier in Sec 2.2, the choice of this frequency windows is to a good degree arbitrary; on the other hand, the needlet framework naturally provides a binning scheme which is controlled by a single width parameter B (as well as by the profile of the filter function $b(\cdot)$).

As a second difference, we note that while the PCL approach usually makes use of the *backward modelling*, where measurements are deconvolved for numerical and observational effects to match the theoretical predictions, needlets analysis is oriented towards the *forward modelling*, which turns theoretical (needlet) power spectra into pseudo-spectra that can be directly compared to the raw measurements[16] (see [43] for a discussion on forward and backward modeling). In particular, in the needlet case it is not feasible to write a closed formula such as eq. 2.11 to express the original needlet power spectrum as a function of the pseudo one, i.e. $\beta_j = \beta_j(\Gamma_j)$; however, this is not an obstacle for data analysis because the forward estimator can be used just as well to do model checks as parameter estimation.

4 Numerical evidence

In this section we describe the simulations setup exploited and the tests performed in order to compare the harmonic and needlet cross-correlation estimators.

---

[16]Note that pseudo-spectra, either in harmonic or needlet space, depend on the observational setup represented for example by the masking, the smoothing, and the apodization, while this is not the case for theoretical predictions.
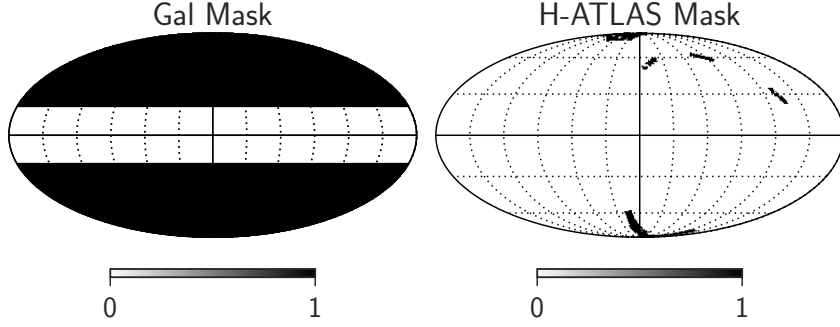
Figure 2. Masks used for the analysis. The mask with a symmetric galactic cut at $\pm 20 \deg$ ($f_{\rm sky} = 0.65$) is shown in the left part, while the H-ATLAS mask ($f_{\rm sky} = 0.013$) is shown on the right one. In both cases the black color denotes observed regions of the sky.

4.1 Simulations

We simulate a set of $N_{\rm sim} = 500$ correlated CMB convergence and galaxy density maps at an HEALPix resolution of $N_{\rm side} = 512$ (corresponding to an angular resolution of $\sim 7'.2$). For the galaxies we consider an high-z Herschel-like population with a redshift distribution as described in [11] and fix $b = 3$ for the present galaxy sample; the precise details of spectra are not fundamental since we are interested in testing the estimators. This simulations set is used in a Monte Carlo (MC) approach (i) to validate the extraction pipelines; (ii) to compute the uncertainty associated with each bin; and (iii) to quantify the degree of correlation among different needlet frequencies. A thorough description of the main steps to obtain correlated CMB lensing and galaxy maps comprising of signal and noise can be found in [9], here we simply use noise-free maps for validation purposes. Pairs of correlated signal-only Gaussian CMB convergence $\kappa_{\ell m}^S$ and galaxy density $g_{\ell m}^S$ maps are generated from the three fiducial spectra $C_\ell^{\kappa g}$, $C_\ell^{\kappa \kappa}$ and C_ℓ^{gg} [9, 44]. This is easily implemented using the synfast routine of HEALPix. In order to show the effect of masking on the reconstructed statistics, the simulated maps are masked with two different masks: we consider either a galactic mask that covers the 35% of sky ($f_{\rm sky} = 0.65$), similar to the one implemented in Planck CMB data, and a much more aggressive Herschel-Astrophysical Terahertz Large Area Survey (H-ATLAS) [45] mask with sky coverage equal only to 1.3% that comprehends the North Galactic Pole, the South Galactic Pole, and the GAMA fields. The adopted masks are shown in Fig. 2.

4.2 Results

We start by investigating the uncorrelation properties of the needlets coefficients, as a function of the width of the mask. In particular, the covariance matrix of needlet coefficients is computed by means of 500 MC simulations as

$$\mathrm{Cov}_{jj'} \equiv \mathrm{Cov}[\hat{\beta}_j, \hat{\beta}_{j'}] = \langle (\hat{\beta}_j - \langle \hat{\beta}_j \rangle_{\rm MC})(\hat{\beta}_{j'} - \langle \hat{\beta}_{j'} \rangle_{\rm MC}) \rangle_{\rm MC}. \tag{4.1}$$

The corresponding results are reported in Fig. 8 for the full sky galactic and H-ATLAS case respective. Numerical evidence is very much consistent with the theoretical expectation: in particular in the full sky and galactic mask case the correlation decrease very rapidly outside the main diagonal (where it is trivially unit, which is not reported in the table) in the case of full-sky maps. The decay is still very satisfactory when sky coverage is high, although not complete as for the galactic mask; on the other hand a very aggressive cut with sky coverage of
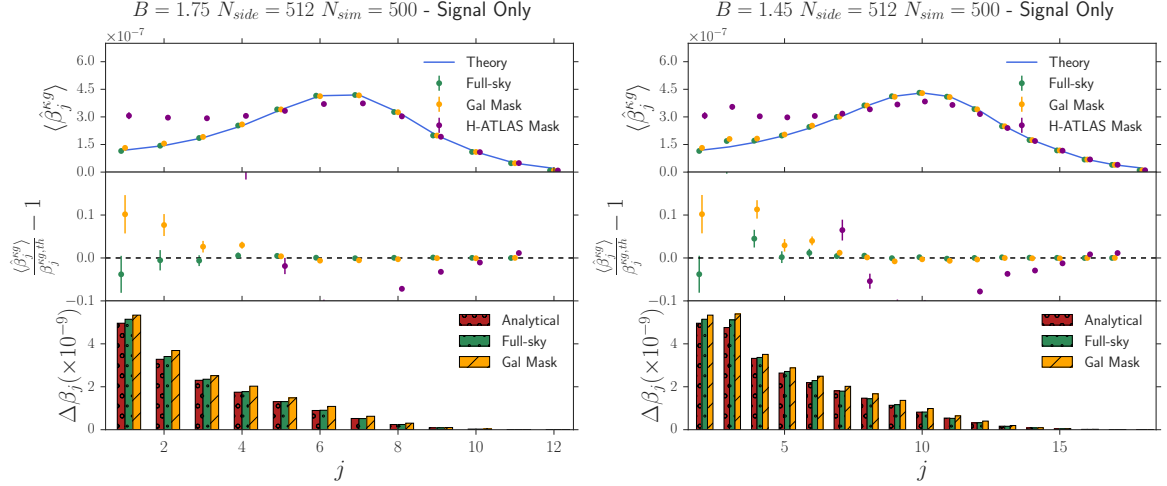
Figure 3. *Upper panel*: Recovered mean needlet cross-power spectrum between correlated CMB convergence and galaxy density maps for different masks and width parameter ($B = 1.75$ and 1.45 on the left and right parts respectively). Green, yellow and purple bandpowers represent full-sky, galactic mask (with $f_{\text{sky}} = 0.65$) and H-ATLAS mask (with $f_{\text{sky}} = 0.013$) cases respectively. Solid blue line is the generative theoretical input cross-power spectrum. Error bars shown are the diagonal components of the covariance matrices (defined in eq. 4.1), properly scaled by $\sqrt{N_{\text{sim}}}$. *Central panel*: Fractional difference between mean recovered and theoretical needlet cross- spectra for the cases shown in the upper panel. *Lower panel*: Error bars comparison for the cases shown in the upper panel. Note that the lack of power observed for $j = 12$ (or for $j = 18$ if $B = 1.45$) is due to the fact that simulated maps have been generated using spectral information up to $\ell_{\text{max}} = 2N_{\text{side}} = 1024$, while the needlet frequency $j = 12$ picks up signal in the multipole range of $458 \lesssim \ell \lesssim 1396$ ($551 \lesssim \ell \lesssim 1159$), where the power is partially missing.

1.3% deteriorate enormously the uncorrelation properties (even though $\text{Corr}_{jj'}$ is $\mathcal{O}(0.1)$ and smaller at high frequencies for $|j - j'| \geq 2$), see the bottom panel in Fig. 8. The estimated covariances are then used to derive error bars in the cross-correlation estimators reported in Fig. 3. Again, the needlet estimator is shown to perform very well in the full-sky and galactic mask cases whereas Herschel-like framework clearly requires corrections. Error bars decay rapidly for increasing frequencies as expected. For comparison, in Fig. 4 MASTER-like estimators are reported for the cross-power spectrum, while the equivalent MASTER needlet reconstruction discussed in section 3 is shown in Fig. 5, where we can see that the bias is strongly suppressed.

The most important results are collected in Fig. 6, where we report the performance of the MASTER-like corrected needlet cross correlation estimator. The figure of merit that we report is the signal-to-noise ratio (S/N), evaluated as the ratio between the analytical expected value of the estimator (numerator) and a measure of variability, which can be either the standard deviation ($\Delta\hat{\beta}_j \equiv \sqrt{\text{Cov}_{jj}}$) or the root mean square error ($\sqrt{\text{MSE}}$). The latter estimator takes in to account also the possible presence of bias, but this is so small that the two measures are largely equivalent. Clearly, an higher value of this figure of merit entails a better performance of the estimator; for comparison, the same statistics are reported also for the MASTER-like cross-correlation estimator in the standard harmonic domain.

The performance of the needlet estimator seems to be extremely satisfactory, with the figure of merit ranging from 1 to 3 for the H-ATLAS case and from 3 to 10 for the galactic case
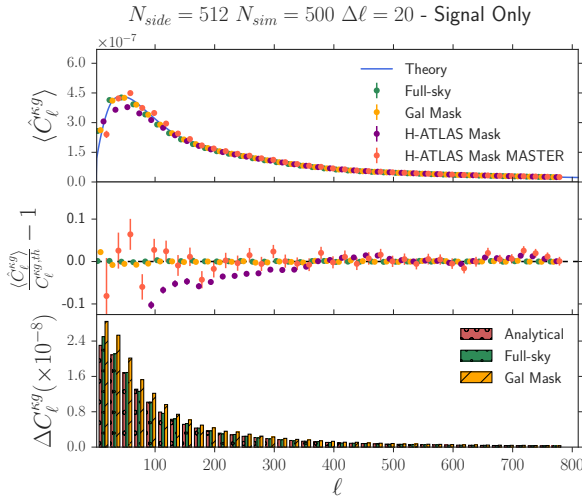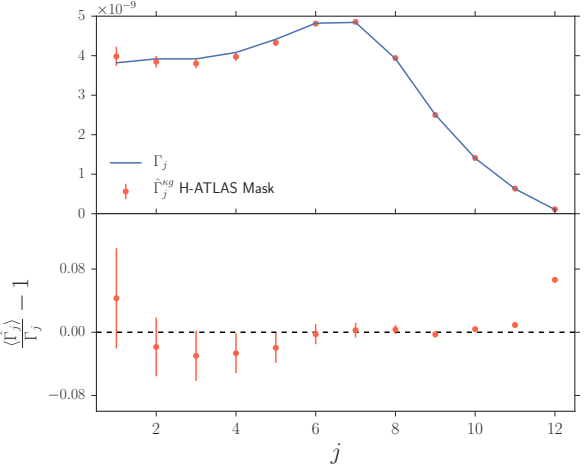
Figure 4. Same as 3 but in harmonic space.

Figure 5. Mean needlet pseudo power spectrum $\langle \hat{\Gamma}_j^{\kappa g} \rangle$ (orange circles) superimposed to the generative theoretical (pseudo) spectrum (blue line).

at the smallest frequencies $j = 3, 4, 5$ (corresponding to multipoles of the order $\ell = 6, 10, 18$ respectively). At higher frequencies, i.e. $j = 10, 11, 12$ (corresponding to central multipoles of the order $\ell = 312, 547, 957$ respectively), the figure of merit is of order 200 (30) when the galactic (H-ATLAS) mask is applied. To make a rough comparison, the figures of merit for the standard power spectrum cross-correlation estimators are in the order of 9 at $\ell = 10$, and 80 at $\ell = 800$ for the galactic mask case, while the figure of merit in the H-ATLAS scenario goes from below 1 up to roughly 6 in the same ℓ-range. To be fair, we stress that these numbers are not strictly comparable, because the bandwidths which are chosen for the standard harmonic domain estimator are constant across the multipoles domain and smaller than the equivalent needlet bandwidths, especially at high frequencies j. However, the least one can conclude from these results is that the two procedures have different advantages; in particular, we view as major assets for the needlet based algorithm the very high S/N and the natural choice of bandwidth parameters; while the advantage of the power spectrum based procedure seems the high resolution which can be achieved in multipole spaces.

As a further check we show in Fig. 7 the variance of the harmonic and needlet space estimators for the different observational setups as function of multipole ℓ and needlet frequency j, normalized to the full-sky analytical variance.

5 Conclusions

Cross-correlation analyses between independent cosmological datasets have the advantage to be potentially immune to any known (and unknown) systematics, as well as to extract signals hidden in noisy data. In this way, cross-correlation measurements can provide us with a clearer view of the large scale distribution of matter, fundamental to reconstruct the dynamics and the spatial distribution of the gravitational potential that can be then translated into constraints on cosmological parameters, breaking degeneracies with the astrophysical ones.

In this paper we begin a systematic analysis of the scientific potential associated to the expansion of the analysis domain in CMB-LSS cross-correlation studies to include the localization in the harmonic and spatial domains. In this initial application, by exploiting an
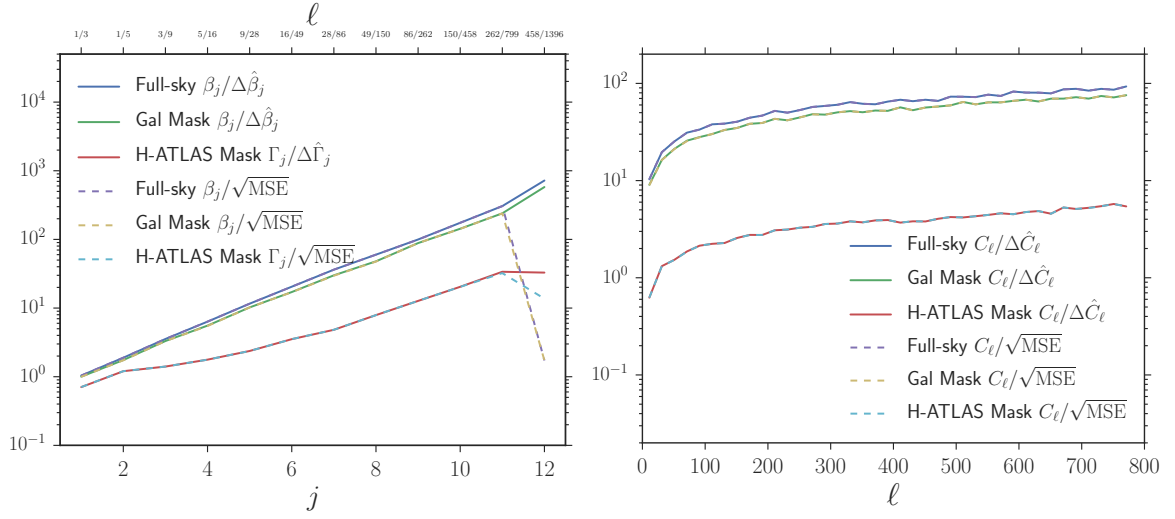
Figure 6. Different figures of merit discussed in text to assess the goodness of the estimators for the needlet (left panel) and harmonic (right panel) cases respectively. On top of the left plot we quote the multipoles that roughly correspond to a given needlet frequency j.
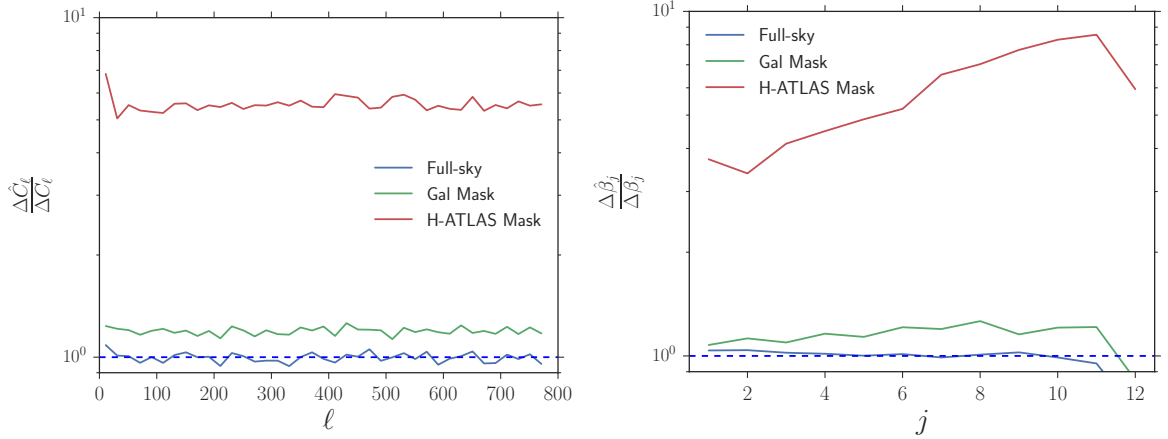

Figure 7. The variance of cross-power spectrum estimates (left panel) and needlet cross-spectra (right panel) divided by the respective analytical full-sky variance. Note that we calculate the ratio $\hat{\beta}_j/\Delta\beta_j$ for the H-ATLAS case, not the pseudo spectrum $\hat{\Gamma}_j/\Delta\Gamma_j$.

ensemble of simulations, we have shown that under the same observational configurations the needlet spectral estimator can outperform the harmonic one thanks to the excellent needlets localization properties in both pixel and frequency space, as well as their optimal window function. Moreover, we have completed an initial needlet based analysis pipeline throughout the implementation of a novel MASTER-like approach for needlet spectral reconstruction in the case of aggressive masking ($f_{\rm sky} \simeq 0.01$), reporting an higher S/N with respect to its harmonic counterpart.

Motivated by these positive indications and results, in future research we plan to explore further the trade-off between S/N and multipole localization, so as to achieve optimal bandwidth selection for a given experimental setting (such as the Euclid coverage mask). We also

aim at applying this machinery to accurate CMB maps lensed with ray-tracing techniques [46] and realistic galaxy mock catalogues based on N-body simulations by adopting, on the CMB side, the projected accuracy and sensitivity of forthcoming polarization oriented CMB probes, targeting the B-modes from cosmological gravitational waves and gravitational lensing. This work is of course preparatory for application to real data, from currently available LSS maps such as Herschel and WISExSCOS Photometric Redshift Catalogue (WISExSCOSPZ) [47] to upcoming surveys such as Euclid, LSST, DESI, and WFIRST, in order to robustly extract cosmological information from cross-correlation measurements.

Acknowledgments

References

[1] Planck Collaboration, *Planck 2015 results. I. Overview of products and scientific results*, 1502.01582.

[2] R. Laureijs, J. Amiaux, S. Arduini, J. L. Auguères, J. Brinchmann, R. Cole et al., *Euclid Definition Study Report*, 1110.3193.

[3] D. Pietrobon, A. Balbi and D. Marinucci, *Integrated Sachs-Wolfe effect from the cross correlation of WMAP 3 year and the NRAO VLA sky survey data: New results and constraints on dark energy*, *Physical Review D* **74** (aug, 2006) 043524.

[4] P. Vielva, E. Martinez-Gonzalez and M. Tucci, *Cross-correlation of the cosmic microwave background and radio galaxies in real, harmonic and wavelet spaces: detection of the integrated Sachs-Wolfe effect and dark energy constraints*, *Monthly Notices of the Royal Astronomical Society* **365** (jan, 2006) 891–901.

[5] J. D. McEwen, P. Vielva, M. P. Hobson, E. Martinez-Gonzalez and A. N. Lasenby, *Detection of the integrated Sachs-Wolfe effect and corresponding dark energy constraints made with directional spherical wavelets*, *Monthly Notices of the Royal Astronomical Society* **376** (apr, 2007) 1211–1226.

[6] D. Munshi, B. Hu, A. Renzi, A. Heavens and P. Coles, *Probing modified gravity theories with ISW and CMB lensing*, *Monthly Notices of the Royal Astronomical Society* **442** (jul, 2014) 821–837, [1403.0852].

[7] Planck Collaboration, *Planck 2013 results. XIX. The integrated Sachs-Wolfe effect*, *Astronomy & Astrophysics* **571** (nov, 2014) A19, [1303.5079].

[8] Planck Collaboration, *Planck 2015 results. XXI. The integrated Sachs-Wolfe effect*, *Astronomy & Astrophysics* (feb, 2015) 30, [1502.01595].

[9] F. Bianchini, P. Bielewicz, A. Lapi, J. Gonzalez-Nuevo, C. Baccigalupi, G. de Zotti et al., *Cross-correlation between the CMB lensing potential measured by Planck and high-z submillimiter galaxies detected by the HERSCHEL-ATLAS survey*, *The Astrophysical Journal* **802** (mar, 2015) 64, [1410.4502].

[10] R. Allison, S. N. Lindsay, B. D. Sherwin, F. de Bernardis, J. R. Bond, E. Calabrese et al., *The Atacama Cosmology Telescope: measuring radio galaxy bias through cross-correlation with*

lensing, *Monthly Notices of the Royal Astronomical Society* **451** (jun, 2015) 849–858, [1502.06456].

[11] F. Bianchini, A. Lapi, M. Calabrese, P. Bielewicz, J. Gonzalez-Nuevo, C. Baccigalupi et al., *Toward a tomographic analysis of the cross-correlation between Planck CMB lensing and h-ATLAS galaxies*, *The Astrophysical Journal* **825** (jun, 2016) 24.

[12] T. Giannantonio, P. Fosalba, R. Cawthon, Y. Omori, M. Crocce, F. Elsner et al., *CMB lensing tomography with the DES Science Verification galaxies*, *Monthly Notices of the Royal Astronomical Society* **456** (mar, 2016) 3213–3244, [1507.05551].

[13] E. J. Baxter, J. Clampitt, T. Giannantonio, S. Dodelson, B. Jain, D. Huterer et al., *Joint Measurement of Lensing-Galaxy Correlations Using SPT and DES SV Data*, eprint arXiv **1602** (feb, 2016) arXiv:1602.07384, [1602.07384].

[14] F. J. Narcowich, P. Petrushev and J. D. Ward, *Localized Tight Frames on Spheres*, *SIAM Journal on Mathematical Analysis* **38** (jan, 2006) 574–594.

[15] P. Baldi, G. Kerkyacharian, D. Marinucci and D. Picard, *Asymptotics for spherical needlets*, *The Annals of Statistics* **37** (jun, 2009) 1150–1171, [0606599v2].

[16] D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella, G. Kerkyacharian et al., *Spherical needlets for cosmic microwave background data analysis*, *Monthly Notices of the Royal Astronomical Society* **383** (jan, 2008) 539–545, [0707.0844].

[17] X. Lan and D. Marinucci, *The needlets bispectrum*, *Electronic Journal of Statistics* **2** (feb, 2008) 332–367.

[18] Ø. Rudjord, F. K. Hansen, X. Lan, M. Liguori, D. Marinucci and S. Matarrese, *An Estimate of the Primordial Non-Gaussianity Parameter fNL Using the Needlet Bispectrum from WMAP*, *The Astrophysical Journal* **701** (aug, 2009) 369–376.

[19] D. Pietrobon, P. Cabella, A. Balbi, R. Crittenden, G. De Gasperis and N. Vittorio, *Needlet bispectrum asymmetries in the WMAP 5-year data*, *Monthly Notices of the Royal Astronomical Society: Letters* **402** (feb, 2010) L34–L38.

[20] S. Donzelli, F. K. Hansen, M. Liguori, D. Marinucci and S. Matarrese, *On the linear term correction for needlet/wavelet non-Gaussianity estimators*, *The Astrophysical Journal* **755** (aug, 2012) 19, [1202.1478].

[21] D. Regan, M. Gosenca and D. Seery, *Constraining the WMAP9 bispectrum and trispectrum with needlets*, *Journal of Cosmology and Astroparticle Physics* **2015** (jan, 2015) 013–013.

[22] Planck Collaboration, *Planck 2015 results. XVII. Constraints on primordial non-Gaussianity*, *Astronomy & Astrophysics* (mar, 2016) 61, [1502.01592].

[23] J. Delabrouille, *Component separation in CMB observations*, 1006.5344.

[24] Planck Collaboration, *Planck 2013 results. XII. Diffuse component separation*, *Astronomy & Astrophysics* **571** (nov, 2014) A12, [1303.5072].

[25] Planck Collaboration, *Planck 2015 results. IX. Diffuse component separation: CMB maps*, *Astronomy & Astrophysics* (feb, 2015) , [1502.05956].

[26] K. K. Rogers, H. V. Peiris, B. Leistedt, J. D. McEwen and A. Pontzen, *SILC: a new Planck internal linear combination CMB temperature map using directional wavelets*, *Monthly Notices of the Royal Astronomical Society* **460** (aug, 2016) 3014–3028.

[27] K. K. Rogers, H. V. Peiris, B. Leistedt, J. D. McEwen and A. Pontzen, *Spin-SILC: CMB polarisation component separation with spin wavelets*, .

[28] D. Geller, F. K. Hansen, D. Marinucci, G. Kerkyacharian and D. Picard, *Spin needlets for cosmic microwave background polarization data analysis*, *Physical Review D* **78** (dec, 2008) 123533, [0811.2881].

[29] B. Leistedt, J. D. McEwen, M. Büttner, H. V. Peiris, P. Vandergheynst and Y. Wiaux, *Analysing the polarisation of the CMB with spin scale-discretised wavelets*, 1502.03120.

[30] C. Durastanti, Y. Fantaye, F. Hansen, D. Marinucci and I. Z. Pesenson, *Simple proposal for radial 3D needlets*, *Physical Review D* **90** (nov, 2014) 103532, [1408.1095].

[31] B. Leistedt, J. D. McEwen, T. D. Kitching and H. V. Peiris, *3D weak lensing with spin wavelets on the ball*, *Physical Review D* **92** (dec, 2015) 123010.

[32] A. Troja, S. Donzelli, D. Maino and D. Marinucci, *The Needlet CMB Trispectrum*, *Proceedings of the International Astronomical Union* **10** (may, 2014) 48–50.

[33] A. Lewis and A. Challinor, *Weak gravitational lensing of the CMB*, *Physics Reports* **429** (jun, 2006) 1–65, [0601594].

[34] D. N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field.*, *The Astrophysical Journal* **117** (jan, 1953) 134.

[35] A. Lewis, A. Challinor and A. Lasenby, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, *The Astrophysical Journal* **538** (aug, 2000) 473–476, [9911177].

[36] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya and M. Oguri, *Revising the Halofit Model for the Nonlinear Matter Power Spectrum*, *The Astrophysical Journal* **761** (dec, 2012) 152, [1208.2701].

[37] E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet and F. Hansen, *MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets*, *The Astrophysical Journal* **567** (mar, 2002) 2–17, [0105302].

[38] G. Efstathiou, *Myths and truths concerning estimation of power spectra: the case for a hybrid estimator*, *Monthly Notices of the Royal Astronomical Society* **349** (apr, 2004) 603–626, [0307515].

[39] E. Komatsu, B. D. Wandelt, D. N. Spergel, A. J. Banday and K. M. Gorski, *Measurement of the Cosmic Microwave Background Bispectrum on the COBE DMR Sky Maps*, *The Astrophysical Journal* **566** (feb, 2002) 19–29, [0107605].

[40] D. Marinucci and G. Peccati, *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications.* Cambridge University Press, 2011.

[41] K. M. Gorski, E. Hivon, a. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke et al., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, *The Astrophysical Journal* **622** (apr, 2005) 759–771, [0409513].

[42] J. D. McEwen, P. Vandergheynst and Y. Wiaux, *On the computation of directional scale-discretized wavelet transforms on the sphere*, in *SPIE* (D. Van De Ville, V. K. Goyal and M. Papadakis, eds.), p. 88580I, sep, 2013. DOI.

[43] J. Harnois-Déraps, T. Tröster, A. Hojjati, L. van Waerbeke, M. Asgari, A. Choi et al., *CFHTLenS and RCSLenS cross-correlation with Planck lensing detected in fourier and configuration space*, *Monthly Notices of the Royal Astronomical Society* **460** (jul, 2016) 434–457, [1603.07723].

[44] T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, S. P. Boughn, A. D. Myers et al., *Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications*, *Physical Review D* **77** (jun, 2008) 123520, [0801.4380].

[45] S. Eales, L. Dunne, D. Clements, A. Cooray, G. De Zotti, S. Dye et al., *The Herschel ATLAS*, *Publications of the Astronomical Society of the Pacific* **122** (may, 2010) 499–515, [0910.4279].

[46] M. Calabrese, C. Carbone, G. Fabbian, M. Baldi and C. Baccigalupi, *Multiple lensing of the cosmic microwave background anisotropies*, *Journal of Cosmology and Astroparticle Physics* **2015** (mar, 2015) 049–049, [1409.7680].

[47] M. Bilicki, J. A. Peacock, T. H. Jarrett, M. E. Cluver, N. Maddox, M. J. I. Brown et al., *WISE x SuperCOSMOS photometric redshift catalog: 20 million galaxies over 3pi steradians*, 1607.01182.
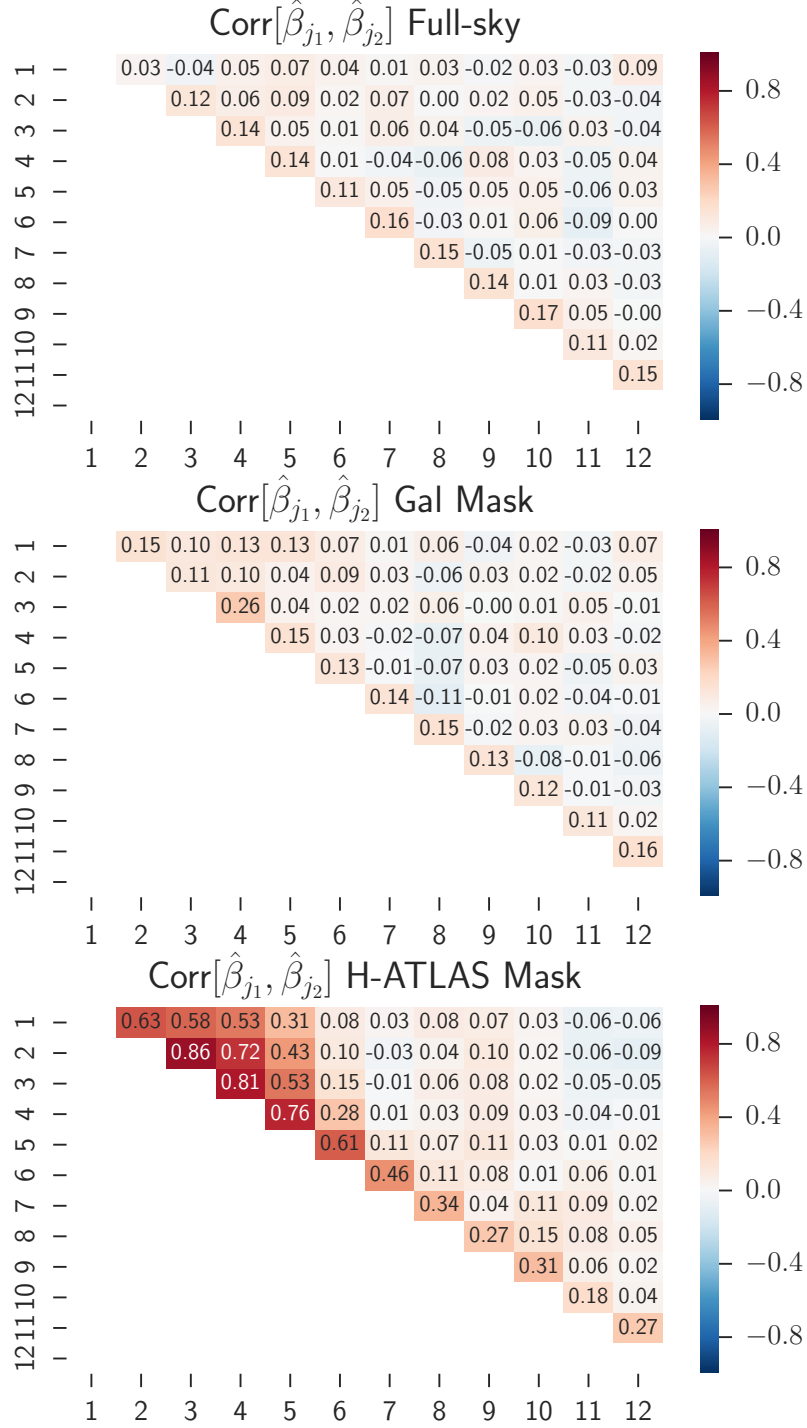
Figure 8. Cross-correlation coefficient matrices, defined as $\mathrm{Corr}_{ij} \equiv \mathrm{Cov}_{ij}/\sqrt{\mathrm{Cov}_{ii}\mathrm{Cov}_{jj}}$, about the needlet space estimator. From top to bottom we show results for the full-sky, galactic mask, and H-ATLAS mask cases respectively.