

NeedATool: A NEEDLET ANALYSIS TOOL FOR COSMOLOGICAL DATA PROCESSING

DAVIDE PIETROBON¹, AMEDEO BALBI^{2,3}, PAOLO CABELLA², AND KRZYSZTOF M. GORSKI^{1,4}

¹ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA;
davide.pietrobon@jpl.nasa.gov, krzysztof.m.gorski@jpl.nasa.gov

² Dipartimento di Fisica, Università di Roma “Tor Vergata,” via della Ricerca Scientifica 1, 00133 Roma, Italy;
amedeo.balbi@roma2.infn.it, paolo.cabella@roma2.infn.it

³ INFN Sezione di Roma “Tor Vergata,” via della Ricerca Scientifica 1, 00133 Roma, Italy

⁴ Warsaw University Observatory, Aleje Ujazdowskie 4, 00478 Warszawa, Poland

Received 2010 April 5; accepted 2010 August 4; published 2010 October 6

ABSTRACT

We introduce NeedATool (Needlet Analysis Tool), a software for data analysis based on needlets, a wavelet rendition which is powerful for the analysis of fields defined on a sphere. Needlets have been applied successfully to the treatment of astrophysical and cosmological observations, and in particular to the analysis of cosmic microwave background (CMB) data. Usually, such analyses are performed in real space as well as in its dual domain, the harmonic one. Both spaces have advantages and disadvantages: for example, in pixel space it is easier to deal with partial sky coverage and experimental noise; in the harmonic domain, beam treatment and comparison with theoretical predictions are more effective. During the last decade, however, wavelets have emerged as a useful tool for CMB data analysis, since they allow us to combine most of the advantages of the two spaces, one of the main reasons being their sharp localization. In this paper, we outline the analytical properties of needlets and discuss the main features of the numerical code, which should be a valuable addition to the CMB analyst’s toolbox.

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

Online-only material: color figures

1. INTRODUCTION

Over the last two decades, the detailed analysis of cosmic microwave background (CMB) radiation anisotropies has been fundamental in determining the global properties of our universe and its evolutions. The cosmological concordance model encodes the variety of processes we observe in the local universe as well as those occurring at very large scales into a few parameters. Such parameters have been measured very precisely by several CMB experiments (Mather et al. 1992; Smoot et al. 1992; de Bernardis et al. 2000; Komatsu et al. 2010), and such measurements will be further refined with the next generation of cosmological experiments. CMB data analysis is very demanding—both in terms of computational power required and sophistication of the necessary techniques—due to the complexity of the data sets and to the high degree of accuracy one wants to achieve. An overview of the techniques recently applied to astrophysical data analysis can be found in Pesenson et al. (2010). This holds in particular for cutting-edge experiments such as the ongoing *Planck* satellite.⁵

Over the last decade, wavelets (Freedman & Schneider 1998; Antoine & Vanderghyest 1999; McEwen et al. 2006, 2007; Sanz et al. 2006; Starck et al. 2006, 2009) have emerged as a very powerful tool for CMB data analysis including testing for non-Gaussianity (Vielva et al. 2004; Cabella et al. 2004), foreground subtraction (Hansen et al. 2006), point-source detection (Sanz et al. 2006), component separation (Moudén et al. 2005; Starck et al. 2006), polarization analysis (Cabella et al. 2007), and many others. The reason for such a strong interest is easily understood. As is well known, CMB models are best analyzed in the frequency domain, where the behavior at different multipoles can be investigated separately; on the

other hand, partial sky coverage and other missing observations make the evaluation of exact spherical harmonic transforms troublesome. The combination of these two features makes the time–frequency localization properties of wavelets most valuable. See Starck & Bobin (2009) for a recent review on this topic⁶ and Wiaux et al. (2008) for a numerical implementation.

Recently, a novel approach to spherical wavelets has been introduced in the statistical literature by Baldi et al. (2006), adapting tools proposed in the functional analysis literature by Narcowich et al. (2006); the first application to CMB data is due to Pietrobon et al. (2006), where needlets are used to estimate (cross-)angular power spectra in order to search for dark energy imprints on the correlation between large-scale structures and the CMB (Sachs & Wolfe 1967). Guilloix et al. (2007) investigate the effect of different window functions in needlet constructions, whereas Baldi et al. (2009a) provide further mathematical results on their behavior for partially observed sky maps. Needlets have been applied to angular power spectrum estimation in the presence of noise (Faÿ et al. 2008; Faÿ & Guilloix 2008) as well as to the estimation of the bispectrum (Lan & Marinucci 2008); the latter tool has been applied to the *WMAP*-5 data release (*WMAP*-5) by Pietrobon et al. (2009b) and Rudjord et al. (2009) to constrain the primordial non-Gaussianity parameter. The bispectrum formalism has been further exploited by Pietrobon et al. (2009a) and Rudjord et al. (2010), who addressed the sky asymmetry issue analyzing, respectively, the properties of a three-point correlation function and the primordial non-Gaussianity parameter. Cabella et al. (2010) developed the bispectrum estimator including the marginalization over the possible foreground residuals in the CMB maps, while Delabrouille et al. (2009) and Ghosh et al. (2010) produced a

⁵ [http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI\(2005\)1.pdf](http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI(2005)1.pdf)

⁶ A multiresolution package for data analysis and compression is also available from: <http://jstarck.free.fr/mresol.htm>.

foreground component separation algorithm. The analysis of directional data is described in Baldi et al. (2009b). Finally, the needlet formalism has been extended to the polarization field, as discussed by Geller & Marinucci (2010a, 2010b), Geller & Mayeli (2008, 2009), and Geller et al. (2008, 2009).

The aim of this paper is to describe a numerical code, called NeedATool (Needlet Analysis Tool), which has already been used in several of the above-mentioned analyses. We first provide a discussion of the needlet formalism in Sections 2 and 3. We then describe a viable implementation of the code in Section 4 and add our concluding remarks in Section 5.

2. NEEDLETS FRAME

Needlets enjoy several features which in general are not granted by other spherical wavelet constructions. Here, we recall some of these properties and refer to Marinucci et al. (2008) and Lan & Marinucci (2008) for a comprehensive mathematical discussion. Complementary mathematical analyses can be found in Geller & Mayeli (2007, 2008, 2009), Lan & Marinucci (2009), and Mayeli (2008). In particular, needlets have the following properties.

1. They do not rely on any tangent plane approximation (compare Sanz et al. 2006), and take advantage of the manifold structure of the sphere.
2. Being defined in harmonic space, they are computationally very convenient, and natively adapted to standard packages such as *HEALPix*⁷ (Górski et al. 2005).
3. They allow for a simple reconstruction formula (see Equation (5)), where the same needlet functions appear both in the direct and the inverse transform (see also Kerkycharian et al. 2007).
4. They are quasi-exponentially (i.e., faster than any polynomial) concentrated in pixel space (see Equation (6)).
5. They are exactly localized on a finite number of multipoles; the width of this support is explicitly known, controlling the power encoded in each multipole range (see Equation (1)).
6. Coefficients can be shown to be asymptotically uncorrelated (and hence, in the Gaussian case, independent) at any fixed angular distance, when the frequency increases.

We first recall that the spherical needlet function is defined as

$$\psi_{jk}(\hat{\gamma}) = \sqrt{\lambda_{jk}} \sum_{\ell} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} \bar{Y}_{\ell m}(\hat{\gamma}) Y_{\ell m}(\xi_{jk}), \quad (1)$$

where γ and ξ_{jk} are directions on the sphere, $Y_{\ell m}$ is a spherical harmonic function, with $\bar{Y}_{\ell m}$ identifying its complex conjugate, and $b(x)$ is a filter function defined for $x \in [1/B, B]$, on which the entire needlet construction relies. Here, we use $\{\xi_{jk}\}$ to denote a set of cubature points on the sphere, corresponding to frequency j , and λ_{jk} denotes the cubature weights. In Figure 1, the needlet profile as a function of the angle between $\hat{\gamma}$ and ξ_{jk} is shown for the choice $B = 2$, $j = 8$.

Intuitively, needlets should be viewed as a convolution of the projection operator $\sum_{m=-\ell}^{\ell} \bar{Y}_{\ell m}(\hat{\gamma}) Y_{\ell m}(\xi_{jk})$ with a suitably chosen window function $b(\cdot)$. The needlet frame construction strongly relies on the spherical harmonic decomposition which represents the mathematical environment for the derivation of the fundamental properties of needlets (Lan & Marinucci 2009); in particular, the existence of a reconstruction formula, as first

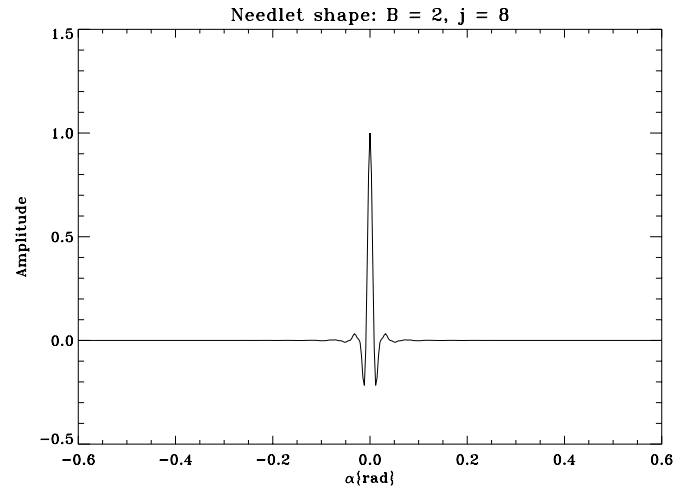


Figure 1. Needlets in pixel space. $B = 2$, $j = 8$ as a function of the angle between $\hat{\gamma}$ and ξ_{jk} .

pointed out by Freeden & Windheuser (1997) and Freeden & Schneider (1998). More details can be found in Freeden & Maier (2002). A similar approach has been followed by Starck et al. (2006, 2009), who implemented wavelets, ridgelets, and curvelets built directly on the sphere, both for scalar and spin-2 fields. All these sets of functions are examples of frames on the sphere: they are over-complete and redundant but admit a well-defined backward transformation. For this reason, frames are not only suitable for a multi-frequency analysis of a signal, like any other wavelet implementation, but also for data compression, denoising algorithms, and component separation, as described by Moudden et al. (2005) and Starck & Bobin (2009).

Besides spherical harmonic decomposition, needlet properties strongly depend on the filter function $b(\cdot)$, which controls the angular scale span covered by each needlet and ensures that needlets enjoy quasi-exponential localization properties in pixel space. Formally, we must ensure that

- i. $b^2(\cdot)$ has support in $[\frac{1}{B}, B]$ and hence $b(\frac{\ell}{B^j})$ has support in $\ell \in [B^{j-1}, B^{j+1}]$,
- ii. the function $b(\cdot)$ is infinitely differentiable in $(0, \infty)$, and
- iii. we have

$$\sum_{j=1}^{\infty} b^2\left(\frac{\ell}{B^j}\right) \equiv 1 \text{ for all } \ell > B. \quad (2)$$

It is easy to see that property (i) ensures that the needlets have bounded support in the harmonic domain, property (ii) is the crucial element in the derivation of the localization properties (Narcowich et al. 2006), and finally, property (iii) is necessary to establish the reconstruction formula (Equation (5)). Functions such as $b^2(\cdot)$ are called partitions of unity.

There are of course many possible constructions satisfying properties (i)-(iii); indeed an interesting theme of research is the derivation of optimal windows satisfying these three conditions (compare Guilloux et al. 2007), although the choice of $b(\cdot)$ is expected to exert second-order effects on the final estimates (Lan & Marinucci 2009). An explicit recipe for the construction of $b(\cdot)$ is given in Section 4. Very recently, an extended study on how needlet properties depend on the filter functions has been conducted by Scodeller et al. (2010). Interestingly, the authors explore a peculiar construction able to mimic the Spherical Mexican Hat Wavelets based on the mathematical study of Geller & Mayeli (2008, 2009).

⁷ <http://healpix.jpl.nasa.gov>

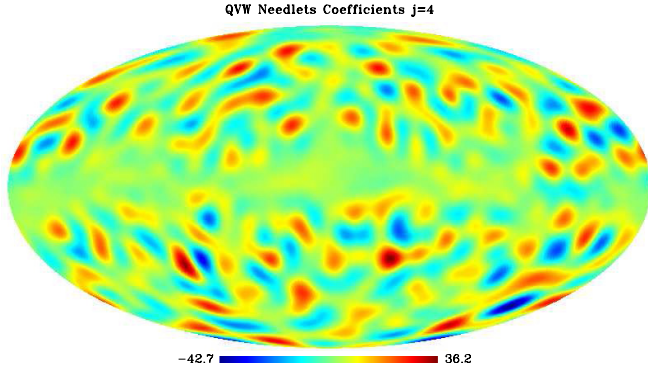


Figure 2. Needlet coefficients of the combined Q , V , and W map at the resolution $j = 4$. The B parameter is fixed to 2. Note that the anomalous bright spots found by Pietrobon et al. (2008) are clearly visible.

(A color version of this figure is available in the online journal.)

Needlet coefficients are hence given by

$$\beta_{jk} = \sqrt{\lambda_{jk}} \sum_{\ell} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\xi_{jk}). \quad (3)$$

In Figure 2, we show the needlet coefficients of the *WMAP*-5 temperature map for the specific choice $B = 2$ and $j = 4$. A remarkable aspect of this construction is that the needlet coefficients can be represented easily as a Mollweide projection in the *HEALPix* pixelization framework (Górski et al. 2005), the most widely used tool for visualization and analysis of CMB maps. This makes dealing with needlets particularly handy since it is easy to implement a needlet analysis code that exploits pre-existing *HEALPix* routines.

It is very important to stress that, although the needlets do not make up an orthonormal basis for square integrable functions on the sphere, they do represent a tight frame. In general, a tight frame on the sphere is a countable set of functions $\{e_j\}$ such that for all square integrable functions on the sphere $f \in L^2(S^2)$, we have

$$\sum_j \langle f, e_j \rangle^2 \equiv \int_{S^2} f(\hat{\gamma})^2 d\Omega,$$

so that the norm is preserved. Here, $\langle f, g \rangle$ means the scalar product or more properly the projection between the functions f and g . Of course, this norm-preserving property is shared by all orthonormal systems; however, frames do not in general make up a basis as they admit redundant elements. They can be viewed as the closest system to a basis for a given redundancy; see Hernández & Weiss (1996), Baldi et al. (2006), and Baldi et al. (2009a) for further definitions and discussion. In our framework, the norm-preserving property translates into

$$\sum_{j,k} \beta_{jk}^2 \equiv \sum_{\ell=1}^{\infty} \frac{2\ell+1}{4\pi} \hat{C}_{\ell}, \quad (4)$$

where

$$\hat{C}_{\ell} = \frac{4\pi}{2\ell+1} \sum_m |a_{\ell m}|^2$$

is the angular power spectrum of the map $T(\hat{\gamma})$. Identity (4) has indeed been verified by means of numerical simulations and

implicitly provides the correct normalization for needlets, λ_{jk} . It is basically a consequence of the peculiar partition-of-unity property of $b(\cdot)$ (Equation (2)). Equation (4) is related to a much more fundamental result, i.e., the reconstruction formula

$$T(\hat{\gamma}) \equiv \sum_{j,k} \beta_{jk} \psi_{jk}(\hat{\gamma}), \quad (5)$$

which in turn is a non-trivial consequence of the analytical properties of the $b_{\ell,j}$ functions. As mentioned before, the simple reconstruction formula of Equation (5) is typical of tight frames but does not hold in general for other wavelet systems.

The following quasi-exponential localization property of needlets is due to Narcowich et al. (2006) and motivates their name. For any $M = 1, 2, \dots$, there exists a positive constant c_M such that for any point $x \in S^2$ we have

$$|\psi_{jk}(\hat{\gamma})| \leq \frac{c_M B^j}{(1 + B^j \arccos(\langle x, \xi_{jk} \rangle))^M}. \quad (6)$$

We recall that $\arccos(\langle x, \xi_{jk} \rangle)$ is just the natural distance on the unit sphere between the points (x, ξ_{jk}) . The meaning of Equation (6) is then clear: for any fixed angular distance, the value of $\psi_{jk}(\hat{\gamma})$ goes to zero quasi-exponentially in the parameter B . This clearly establishes an excellent localization behavior in pixel space. Note that the constants c_M do depend on the form of the weight function $b(\cdot)$, and in particular on the value of the bandwidth parameter B ; typically, a better localization in multipole space (i.e., a value of B very close to unity) will entail a larger value of c_M , that is, less concentration in pixel space for any fixed j . The resulting tradeoff in the behavior over the harmonic and real spaces is expected: smaller values of B correspond to a tighter localization in harmonic space (less multipoles entering into any needlet), whereas larger values ensure a faster decay in real space (Lan & Marinucci 2009).

In Baldi et al. (2006), another relevant property of needlet coefficients was discussed, namely their asymptotic uncorrelation at any fixed angular distance, for growing frequencies j . More explicitly, at high frequency, needlet coefficients can be approximated as a sample of identically distributed and independent (under Gaussianity) coefficients. Also, in view of Equation (3), for full-sky maps and in the absence of any mask we should expect the theoretical correlation to be identically zero whenever $|j_1 - j_2| \geq 2$. This has indeed been numerically verified by Marinucci et al. (2008).

The probabilistic properties of the needlet coefficients β_{jk} have been established in Baldi et al. (2009a); in that paper, it is shown that for any two (sequence of) pixels $\xi_{jk}, \xi_{jk'}$ such that their angular distance is larger than a positive ε , for all j , we have

$$\frac{\langle \beta_{jk} \beta_{jk'} \rangle}{\sqrt{\langle \beta_{jk}^2 \rangle \langle \beta_{jk'}^2 \rangle}} \leq \frac{c_M}{(s^j \varepsilon)^{M-1}} \quad \text{for all } M = 1, 2, 3, \dots, \quad (7)$$

thus proving wavelet coefficients are asymptotically uncorrelated as $j \rightarrow \infty$ for any fixed angular distance. Equation (7) can then be seen as the statistical counterpart of Equation (6).

These properties are the basis for the large success of needlets as a CMB toolbox, in particular when dealing with masked data sets.

3. NEEDLET ESTIMATORS: TWO- AND THREE-POINT CORRELATION FUNCTIONS

Having introduced the spherical needlet frame and recalled the main properties which make needlets perform extremely well in a wide number of applications to two-dimensional fields on the sphere, we now briefly describe some important statistical techniques largely used in CMB data analysis.

3.1. (Cross-)Power Spectrum

After computing the needlet coefficients β_{jk} from a two-dimensional map (e.g., the CMB or the source count map), we can use Equation (4) to build a (cross-)correlation estimator in wavelet space, β_j , as

$$\beta_j^{\text{I}} \equiv \sum_k \frac{1}{N_{\text{pix}}(j)} \beta_{jk}^{\text{I}} \beta_{jk}^{\text{J}}, \quad (8)$$

where $N_{\text{pix}}(j)$ is the number of pixels (e.g., in the *HEALPix* scheme $N_{\text{pix}} = 12N_{\text{side}}^2$) with I and J denoting the two different maps. The theoretical prediction for β_j can be computed from the expected C_{ℓ}^{I} as

$$\beta_j^{\text{I}} = \sum_{\ell} \frac{(2\ell+1)}{4\pi} \left[b\left(\frac{\ell}{B^j}\right) \right]^2 C_{\ell}^{\text{I}}, \quad (9)$$

where we recall $C_{\ell}^{\text{I}} \equiv \langle a_{\ell m}^{\text{I}} a_{\ell m}^{\text{I}*} \rangle$ is the (cross-)angular power spectrum.

β_j provides then an unbiased estimator for the (cross-)angular power spectrum within the needlet framework. The analytic relation between β_j and C_{ℓ} underlines a few more advantages in using needlets. Indeed, it makes dealing with beam profiles and experimental window functions, which have to be taken into account when analyzing real data (see Pietrobon et al. 2006), extremely easy and straightforward. The duality that needlets embed, namely the localization both in pixel and harmonic domain, allows us also to characterize the noise properties (see Delabrouille et al. 2009 and Pietrobon et al. 2009a, 2009b for direct applications to *WMAP*-5 data).

Computing the four-point correlation function, it can be easily shown that the analytical expression for the dispersion of the estimated cross-correlation power spectrum in needlet space is

$$\Delta\beta_j^{\text{I}} = \sqrt{\sum_{\ell} \frac{(2\ell+1)}{16\pi^2} \left[b\left(\frac{\ell}{B^j}\right) \right]^4 ((C_{\ell}^{\text{I}})^2 + C_{\ell}^{\text{I}} C_{\ell}^{\text{J}})},$$

which, of course, must be taken only as an approximation when dealing with real data, where window functions, noise, and partial sky coverage have to be taken into proper account.

It is important to stress that Equation (8) generalizes into

$$\beta_{j_1 j_2}^{\text{I}} = \frac{1}{N_{\text{pix}}} \sum_k \beta_{j_1 k}^{\text{I}} \beta_{j_2 k}^{\text{J}}, \quad (10)$$

which describes the needlet coefficient covariance and has been used in Pietrobon et al. (2008) to determine the degree of anomaly of a few hot and cold spots found in the CMB temperature map.

We have shown that the needlet formalism may be suitable for the problem of angular power spectrum estimation from a CMB map (and therefore, indirectly, to the estimation of cosmological parameters). In particular, the application of needlets

to the *WMAP* three-year data led to interesting constraints on the dynamics of dark energy (Pietrobon et al. 2006) and to the measure of the difference in power between the two estimates of the power spectrum computed on the north and south CMB skies (Pietrobon et al. 2008). A detailed discussion on the application of needlets to power spectrum estimation can be found in Faÿ et al. (2008).

3.2. Needlet Bispectrum

In the previous section, we described how needlets can naturally be applied to the estimation of the two-point correlation function and how, thanks to the reconstruction formula (Equations (2) and (4)), it relates to the usual angular power spectrum. It is easy to extend the formalism to higher order correlation functions.

Here, we focus on the three-point correlation function, which plays a crucial role in CMB data analysis to detect any departure of primordial fluctuations from the Gaussian statistics, a smoking gun for non-standard inflationary models. Next we briefly review the properties of the needlet bispectrum and how it relates to the spherical harmonic bispectrum. An extensive discussion is provided in Lan & Marinucci (2008), Pietrobon et al. (2009a, 2009b), and Rudjord et al. (2009, 2010).

The needlet bispectrum is defined as follows:

$$\begin{aligned} S_{j_1 j_2 j_3} &\equiv \frac{1}{N_{\text{pix}}} \sum_k \beta_{j_1 k} \beta_{j_2 k} \beta_{j_3 k} \\ &= \sum_{\ell_1 \ell_2 \ell_3} b_{\ell_1}^{(j_1)} b_{\ell_2}^{(j_2)} b_{\ell_3}^{(j_3)} \\ &\quad \times \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \\ &\quad \times \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \hat{B}_{\ell_1 \ell_2 \ell_3}, \end{aligned} \quad (11)$$

where

$$\hat{B}_{\ell_1 \ell_2 \ell_3} \equiv \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = \sum_m a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}$$

is the estimated bispectrum, averaged over m 's. $S_{j_1 j_2 j_3}$ can be seen as a binned bispectrum, a smooth and combined component of the angular bispectrum. The bispectrum is supposed to be vanishing for a Gaussian distribution. The standard inflation mechanism (Guth 1981; Sato 1981; Linde 1982; Albrecht & Steinhardt 1982) predicts a tiny non-Gaussianity in the cosmological perturbations; this is why great effort has been spent to measure a bispectrum amplitude different from zero in the CMB data, which would provide an extraordinary handle on the early-universe physics (Smith et al. 2009; Curto et al. 2008, 2010; Komatsu et al. 2009; Natoli et al. 2009; Smidt et al. 2009). This kind of study is usually performed in terms of the nonlinear parameter f_{NL} (see, for example, Komatsu & Spergel 2001; Bartolo et al. 2004; Smith & Zaldarriaga 2006).

One of the key properties of needlets is that the sum of the squared filter functions in harmonic space, $b_{\ell j}$, is 1 (see Equation (2)). This means that even if we group multipoles and each needlet peaks at a certain multipole range the total power is conserved. Therefore, the needlet power spectrum analysis can in principle be done with any choice of parameter B , being the total power conserved, and only the correlation and localization properties affected by different widths of $b_{\ell j}$.

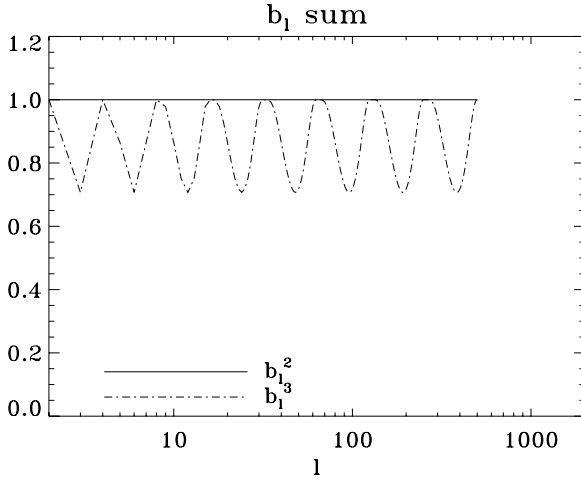


Figure 3. Solid line sum of the b_ℓ^2 ; dot-dashed line sum of the b_ℓ^3 . While the former is equal to 1 for the entire range of multipole, the latter is not.

Incidentally we note that any wavelet function defined as the difference between the square roots of the scaling functions at two different resolutions satisfies Equation (2), but the uncorrelation properties in general are not granted, since they are determined by the shape of the filter function $b_\ell(\cdot)$. See, for example, Scodeller et al. (2010) for a detailed discussion on needlet filters.

This does not hold any more when the cubic power of the filter functions contributes to the estimator used, which is indeed the case of the skewness expression and more generally of the bispectrum one. This fact is displayed in Figure 3 where we plot the sum of square and cube of the filter functions. The non-uniform sampling of the multipoles for an n -power estimator suggests that the choice of the B parameter is crucial for the analysis and must be driven by the insight on the range of multipoles to be probed.

4. NeedATool: A NUMERICAL IMPLEMENTATION OF NEEDLETS

In the previous sections, we have discussed the main needlet properties, which are indeed strictly related to the analytical properties of the filter function in harmonic space, $b(\ell/B^j)$. We now give a specific recipe for the construction of $b(\cdot)$, and describe the main features of “NeedATool” Needlet Analysis Tool—a numerical code that computes the filter functions and needlet coefficients according to the algorithm that has been applied in several cosmological analyses (Pietrobon et al. 2006, 2008, 2009a, 2009b; Cabella et al. 2010).

We give a step-by-step procedure, as implemented in NeedATool and described in Marinucci et al. (2008).

1. Construct the function

$$f(t) = \begin{cases} \exp\left(-\frac{1}{1-t^2}\right), & -1 \leq t \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

It is necessary to check that the function $f(\cdot)$ is C^∞ and compactly supported in the interval $(-1, 1)$.

2. Construct the function

$$\psi(u) = \frac{\int_{-1}^u f(t) dt}{\int_{-1}^1 f(t) dt}.$$

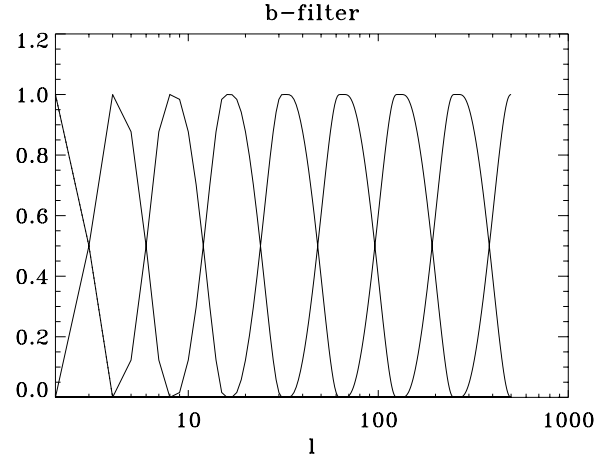


Figure 4. Filter function in ℓ -space on which the needlet construction relies. Set computed for $B = 2$.

The function $\psi(\cdot)$ is again C^∞ ; moreover it is non-decreasing and normalized so that $\psi(-1) = 0$, $\psi(1) = 1$.

3. Construct the function

$$\varphi(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq \frac{1}{B} \\ \psi\left(1 - \frac{2B}{B-1}\left(t - \frac{1}{B}\right)\right) & \text{if } \frac{1}{B} \leq t \leq 1 \\ 0 & \text{if } t > 1. \end{cases}$$

Here we are simply implementing a change of variable so that the resulting function $\varphi(\cdot)$ is constant on $(0, B^{-1})$ and monotonically decreasing to zero in the interval $(B^{-1}, 1)$. Indeed, it can be checked that

$$1 - \frac{2B}{B-1}\left(t - \frac{1}{B}\right) = \begin{cases} 1 & \text{for } t = \frac{1}{B} \\ -1 & \text{for } t = 1 \end{cases}$$

and

$$\begin{aligned} \varphi\left(\frac{1}{B}\right) &= \psi(1) = 1 \\ \varphi(1) &= \psi(-1) = 0. \end{aligned}$$

4. Construct

$$b^2(\xi) = \varphi\left(\frac{\xi}{B}\right) - \varphi(\xi).$$

The expression for $b^2(\cdot)$ is meant to ensure that the function satisfies the partition-of-unity property of Equation (2). Needless to say, for $b(\xi) = \{\varphi(\frac{\xi}{B}) - \varphi(\xi)\}^{1/2}$ we take the positive root.

Incidentally we note that property (4) is crucial in allowing for the reconstruction formula, and it is shared, although within a different setup, by the implementation described by Starck et al. (2006, 2009) who make use of the *HEALPix* software package too. In Figure 4, we show the set of filter functions in ℓ space for the choice $B = 2$. They result in a homogeneous binning in $\log \ell$, whose power adds up to 1, a crucial property on which needlet properties rely.

NeedATool⁸ computes the needlet filter functions in harmonic space and the needlet coefficients, given a set of parameters. The code is based on the publicly available *HEALPix* package⁹

⁸ http://www.fisica.uniroma2.it/~pietrobon/dp_webpage_eng.html

⁹ <http://healpix.jpl.nasa.gov/>

Table 1
Parameters Required by the Codes “synneed” and “ananeed”

Parameter	synneed	ananeed
healpix_dir	/usr/local/Healpix_2.13a	
ℓ_{\max}	500	
B	2.0	
compute_needlets	T	
mapfile	input/lcdm_map_lmax500.fits	needlet_2.00_Nj009.fits
mapnside	256	256
maskfile	input/sky_cut_1_256_ring.fits	input/sky_cut_1_256_ring.fits
bl2_root	bl2	
need_root	needlet	recmap

(Górski et al. 2005), which has to be downloaded and installed separately (version 2.10 or higher required). The software is composed of two programs “synneed” and “ananeed”; following the *HEALPix* structure, the former deconvolves a given map to obtain a needlet frame, and the latter reconstructs the original map, if a needlet’s basis is given as input. Both programs, as well as the routines necessary for NeedATool, will be fully integrated in the next *HEALPix* release. Both programs accept the same parsing file in which the fundamental parameters are provided by the user. A list of such parameters is given in Table 1. The maximum number of multipoles (ℓ_{\max}) and the B parameter are required. Given these, the codes compute the maximum j necessary to keep all the information in the map. The N_{side} of the needlet coefficients is then determined according to the relation $\ell_{\max} \leq 2N_{\text{side}}$. The filter functions $b_{\ell j}$ are computed by default, while it is possible to choose whether to compute the needlet coefficients, which is actually the most time-consuming part, by setting the keyword “compute_needlets.” It is then necessary to specify the map and its resolution. A sky mask can be applied filling the “maskfile” variable. The last two keywords set the output files. The input map of ananeed has to be the needlet output file created by synneed.

The cubature points $\mathcal{X}_j = \{\xi_{jk}\}_{k=1,2,\dots}$ are assumed to coincide with the pixelization of the unit sphere \mathbb{S}^2 provided by *HEALPix*, with N_{side} such that $l_{\max} \equiv [B^{j+1}] \leq 2N_{\text{side}}$ (with $[\cdot]$ denoting the integer part and $B > 1$). The cubature weights, λ_{jk} , are given by $1/N_{\text{pix}}$, with N_{pix} given by $12 \cdot N_{\text{side}}^2$. We then computed the β_{jk} coefficients for each k position given by the *HEALPix* scheme evaluating the projection operator, namely the product of $\sum_{\ell\ell'} Y_{\ell m} \bar{Y}_{\ell' m'}$ for each pair of pixels ξ_{jk} , by means of the *HEALPix* software package. The code is very fast and can be run on a laptop. For a low-resolution map ($N_{\text{side}} = 256$) it takes a few seconds, while it scales according to the *HEALPix* scaling laws for higher resolutions.

We present an example of the reconstruction power of the needlet frame. We produced a CMB map consistent with the *WMAP*-5 best-fit cosmological power spectrum (Komatsu et al. 2009) up to $\ell_{\max} = 500$ by using the *HEALPix* toolbox.

We processed this map through the needlet pipeline first extracting the needlet coefficients by applying *synneed*, then reconstructing the map using *ananeed*. We repeated this procedure both for the case of the entire sky and in the presence of symmetric sky cuts, 15° , 5° , 1° , and 0.1° . The results are shown in Figure 5 (left column) together with the percentage error due to the procedure (right column). The error due to the forward and backward transformation is smaller than 0.01% except for a few pixels in the case of the entire sky. The result worsens reaching 1% when a broad mask is applied, basically because the needlet coefficients at very large scale are affected by the presence of the

mask. This effect is actually expected, since needlets work well at small angular scales, namely when the filter function $b(\ell/B^j)$ groups several multipoles. The mask effect may be reduced by either fine tuning the B parameter or deconvolving the mask effect. Reconstruction error is indeed wavelet frame dependent: a better performance can be achieved when the wavelet is defined as the difference between two scaling functions at two different resolutions instead of the difference between the square root of the scaling functions. This is the implementation discussed by Starck et al. (2006) which leads to an exact reconstruction.

To give another estimate of the error due to the reconstruction, we computed the angular power spectrum of the original map, C_ℓ , as well as those of the reprocessed ones, C_ℓ^R , and compared them. The ratio is displayed in Figure 6 (left panel) for the four analyzed sky cuts applied. The full-sky reconstruction shows an excellent agreement, the difference between the power spectra being of the order of 10^{-4} ; when a mask is present, the reconstruction causes an error of few percent at very low angular scales, which decreases at small scales.

The imperfect reconstruction in the presence of missing information is expected. Therefore, it is interesting to compare the needlets’ performance to the one of the spherical harmonics implemented by *HEALPix*. We extracted the $a_{\ell m}$ from the same CMB realization for the five different sky coverages considered by using *anafast*; then we produced a CMB map out of the pseudo- $a_{\ell m}$ using *synfast*. We finally analyzed the resulting map computing its power spectrum. The difference with respect to the power spectrum computed through needlets is shown in the right panel of Figure 6. The spectra agree very well regardless of the applied mask. This confirms the high performance of needlets in reconstructing a field in the sphere. As a further figure of merit, in Figure 7 we show the percentage error between the two maps reconstructed using spherical harmonics and needlets for the broadest mask (15°). The agreement is striking: a large-scale pattern appears at the level of $<1\%$.

We further investigate the mask effect, focusing on how the presence of a sky cut couples the spherical harmonic coefficients. As is well known, in the case of partial sky coverage the spherical harmonics do not form an orthonormal basis. This can be formalized as follows:

$$T^R(\hat{\gamma}) = \sum_{\ell m} \tilde{a}_{\ell m} Y_{\ell m}(\hat{\gamma}) \quad (12)$$

$$\tilde{a}_{\ell' m'} = \int_{\mathcal{O}} T(\hat{\gamma}) \bar{Y}_{\ell' m'} d\Omega = \sum_{\ell m} a_{\ell m} \mathcal{K}_{\ell \ell' m m'} \quad (13)$$

$$\int_{\mathcal{O}} \bar{Y}_{\ell' m'}(\hat{\gamma}) Y_{\ell m}(\hat{\gamma}) W(\hat{\gamma}) d\Omega \equiv \mathcal{K}_{\ell \ell' m m'}, \quad (14)$$

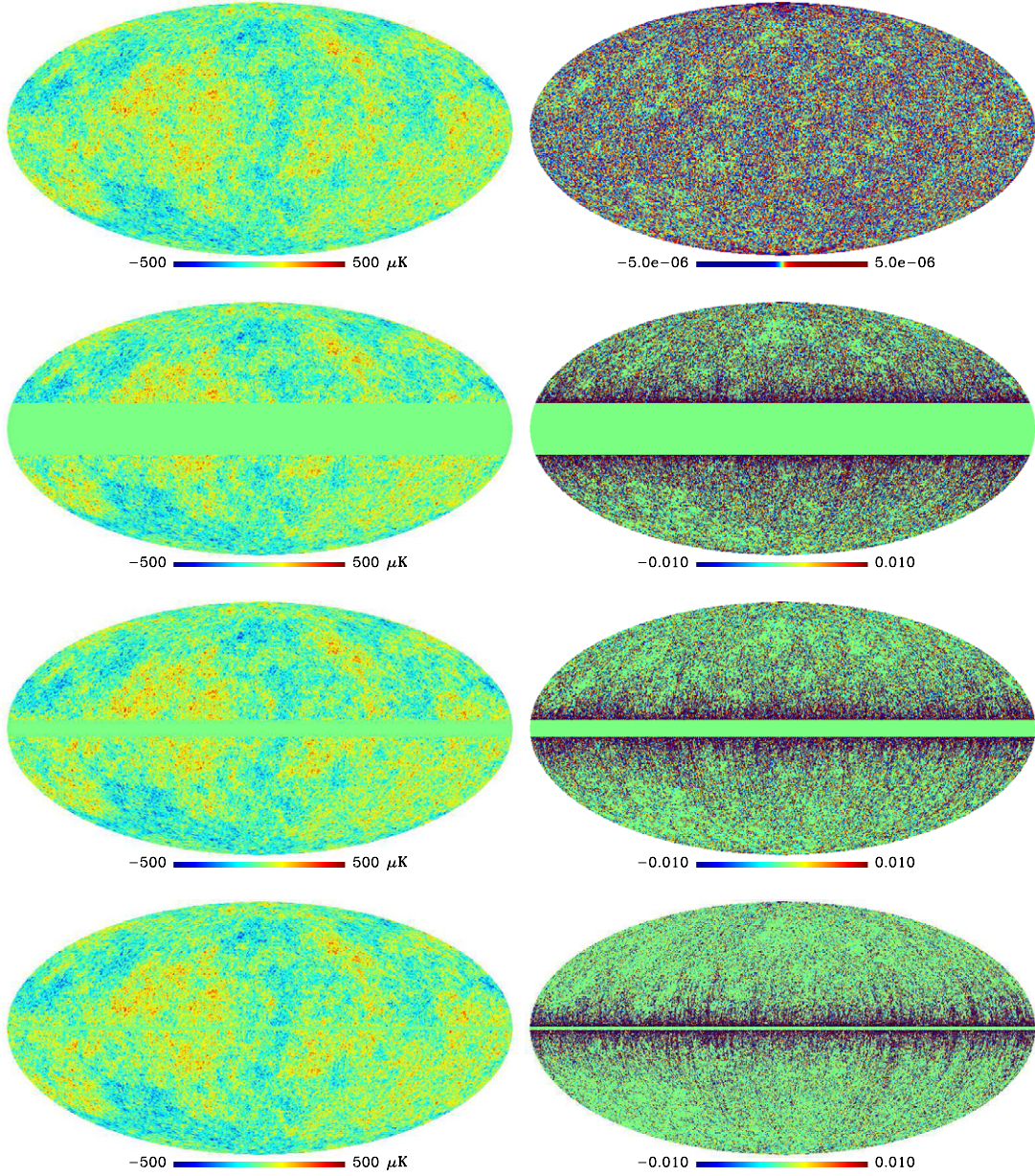


Figure 5. Left: reconstructed maps for the several sky cuts analyzed (from top to bottom full sky, 15°, 5°, and 1°); right: percentage error on the reconstructed map. (A color version of this figure is available in the online journal.)

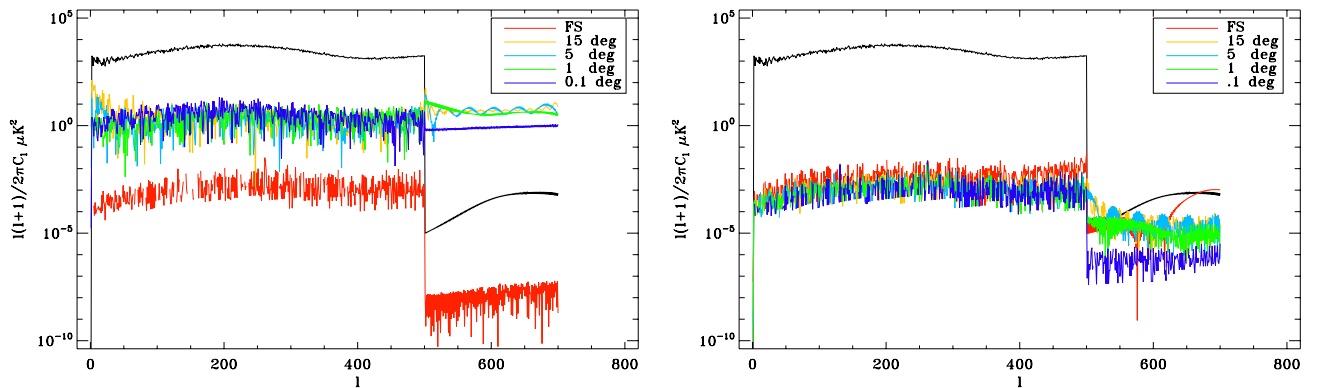


Figure 6. Left panel: absolute value of the difference between the power spectra extracted from the original map and the reconstructed one for several symmetric sky cuts: full sky (red solid line), 15° (yellow solid line), 5° (light blue solid line), 1° (green solid line), and 0.1° (blue solid line). The black solid line shows the CMB angular power spectrum for the original map. Right panel: comparison between the reconstruction using *synfast-anafast* combination and the needlets pipeline for the same sky cuts. The input angular power spectrum contains information up to $\ell = 500$, but we computed the power spectrum of the realization map up to $\ell = 700$ to check the leakage from low multipoles to high ones: this explains the edge at $\ell = 500$ in the plotted power spectra.

(A color version of this figure is available in the online journal.)

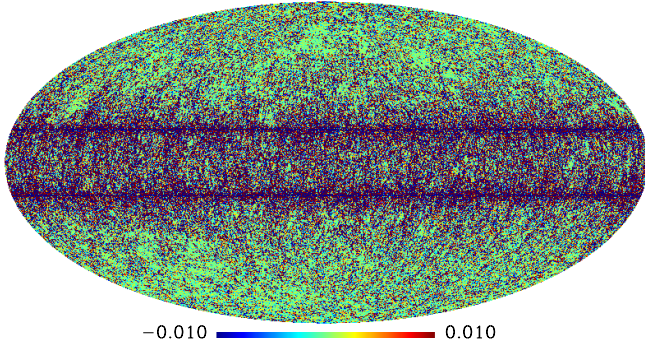


Figure 7. Percentage error between the two maps reconstructed using spherical harmonics and needlets. A symmetric sky cut of 15° is applied. The agreement is impressive.

(A color version of this figure is available in the online journal.)

where \mathcal{O} is the observed region and $\mathcal{K}_{\ell\ell'mm'}$ is the coupling matrix. In the needlet framework, this affects the β_{jk} coefficients as

$$\tilde{\beta}_{jk} = \sqrt{\lambda_{jk}} \sum_{\ell'm'} b\left(\frac{\ell'}{B^j}\right) \tilde{a}_{\ell'm'} Y_{\ell'm'}(\xi_{jk}) \quad (15)$$

$$T^R(\hat{\gamma}) = \sum_j \sum_{\ell'm'} b\left(\frac{\ell'}{B^j}\right) \tilde{a}_{\ell'm'} \sum_{\ell m} b\left(\frac{\ell}{B^j}\right) Y_{\ell m}(\hat{\gamma}) \mathcal{K}_{\ell\ell'mm'}. \quad (16)$$

As for the spherical harmonic decomposition, the map reconstructed through the needlet pipeline is computed from $\tilde{a}_{\ell m}$; in addition, the mask shows its effect indirectly on the filter functions $b(\ell/B^j)$.

To visualize this effect, in Figure 8 we plot the coupling matrix $\mathcal{K}_{\ell\ell'mm'}$ for a full-sky case (upper panel), in which it actually becomes a Kronecker's δ function, $\delta_{\ell\ell'mm'}$; a 15° symmetric sky cut (middle panel); and the convolution $\sum_j b(\ell'/B^j) \mathcal{K}_{\ell\ell'mm'} b(\ell/B^j)$, which appears in Equation (16). Since the numerical evaluation of the coupling matrix up to $\ell = 500$ is a severe computational challenge and we expect the higher multipoles to be marginally affected by a partial sky coverage, we computed Equation (14) on a smaller range of spherical harmonics in the interval $\ell \in [0, 20]$ on the *HEALPix* pixelization at $N_{\text{side}} = 16$. The presence of the mask translates into off-diagonal terms in Figure 8: each multipole is coupled to its second neighbor, whereas the coupling with the first neighbor is vanishing because of parity. The coupling is indeed not

negligible up to the fourth second neighbor. The lower panel in Figure 8 clearly depicts the properties of needlets and explains the error we obtain in the reconstruction. The correlation matrix results in a superposition of blocks, each of them corresponding to a given j : each block partially overlaps only the first neighbor, as a consequence of the compact support of the $b(\ell/B^j)$ functions. (This is shown by the white regions we observe in the plot.) Moreover, the off-diagonal elements are less powerful compared to the autocorrelation terms and the correlation extends to a lower number of multipoles. At a very large scale, however, a lack of power is present in the diagonal terms which affects the global reconstructed map. This can be seen in the right panel of Figure 5, where the error in the reconstructed map follows a peculiar pattern which is symmetric as a consequence of the sky cut.

The additional coupling due to the mask present in the needlet construction arises from the definition of the needlet itself, Equation (1), since the needlet is built from the projection operator. However, needlets provide a natural way to counterbalance this effect. At a large scale, the needlets are not sharply localized, and the coefficients also turn out to be non-vanishing in the masked region. Taking into account this leakage, and including the coefficients inside the mask in the reconstruction procedure, it is possible to reduce the effect of the coupling between needlets. This is the actual procedure implemented in the code, which leads to the results discussed in this section. The residual effect due to the presence of the mask shows up in the large-scale pattern highlighted in Figure 7.

5. CONCLUSIONS

In this paper, we have introduced NeedATool, a public software for the analysis of data sets on the sphere based on the needlet framework. The software is particularly useful for the analysis of CMB data, as shown by its successful application to, e.g., the *WMAP* data set. The needlet construction differs from other wavelet renditions due to the distinctive properties of the filter functions $b_{\ell j}$, which translate into a sharp localization in pixel space and excellent properties of non-correlation among the functions of the set. This aspect is crucial when building estimators for CMB data analysis as we discussed extensively; therefore, needlets are a very promising tool for high-accuracy cosmological experiments.

We thank Domenico Marinucci for useful discussions and Giancarlo de Gasperis and Rajib Saha for technical support.

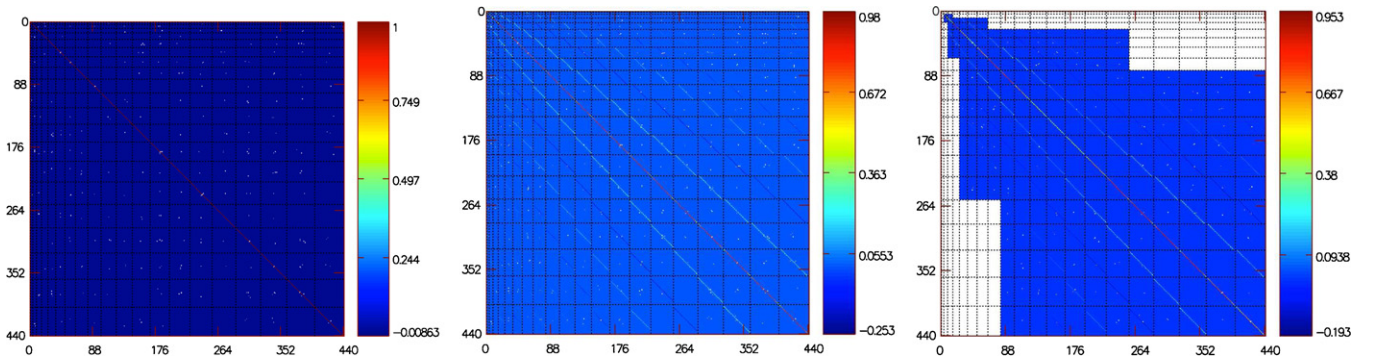


Figure 8. Coupling matrix as a function of the quantity $\ell^2 + \ell + m + 1$ for three interesting cases: full sky (left panel), 15° symmetric sky cut (middle panel), and its convolution with needlet window functions (right panel). The grid marks the multipoles.

(A color version of this figure is available in the online journal.)

REFERENCES

- Albrecht, A., & Steinhardt, P. J. 1982, *Phys. Rev. Lett.*, **48**, 1220
- Antoine, J.-P., & Vanderghenst, P. 1999, *Appl. Comput. Harmon. Anal.*, **7**, 262
- Baldi, P., Kerkycharian, G., Marinucci, D., & Picard, D. 2006, *Ann. Stat.*, **37**, 1150
- Baldi, P., Kerkycharian, G., Marinucci, D., & Picard, D. 2009a, *Bernoulli*, **15**, 438
- Baldi, P., Kerkycharian, G., Marinucci, D., & Picard, D. 2009b, *Ann. Statist.*, **37**, 3362
- Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004, *Phys. Rep.*, **402**, 103
- Cabella, P., Hansen, F., Marinucci, D., Pagano, D., & Vittorio, N. 2004, *Phys. Rev. D*, **69**, 063007
- Cabella, P., Natoli, P., & Silk, J. 2007, *Phys. Rev. D*, **76**, 123014
- Cabella, P., Pietrobon, D., Veneziani, M., Balbi, A., Crittenden, R., de Gasperis, G., Quercellini, C., & Vittorio, N. 2010, *MNRAS*, **405**, 961
- Curto, A., Macías-Pérez, J. F., Martínez-González, E., Barreiro, R. B., Santos, D., Hansen, F. K., Liguori, M., & Matarrese, S. 2008, *A&A*, **486**, 383
- Curto, A., Martínez-González, E., & Barreiro, R. B. 2010, arXiv:1007.2181
- de Bernardis, P., et al. 2000, *Nature*, **404**, 955
- Delabrouille, J., Cardoso, J.-F., Le Jeune, M., Betoule, M., Faÿ, G., & Guillaoux, F. 2009, *A&A*, **493**, 835
- Faÿ, G., & Guillaoux, F. 2008, arXiv:0807.2162
- Faÿ, G., et al. 2008, *Phys. Rev. D*, **78**, 083013
- Freedman, W., & Maier, T. 2002, *ETNA*, **14**, 40
- Freedman, W., & Schneider, F. 1998, *Inverse Probl.*, **14**, 225
- Freedman, W., & Windheuser, U. 1997, *Appl. Comput. Harmon. Anal.*, **4**, 1
- Geller, D., Hansen, F. K., Marinucci, D., Kerkycharian, G., & Picard, D. 2008, *Phys. Rev. D*, **78**, 123533
- Geller, D., Lan, X., & Marinucci, D. 2009, *Electron. J. Statist.*, **3**, 1497
- Geller, D., & Marinucci, D. 2010a, *J. Fourier Analysis & Applications*, online first (arXiv:0811.2935)
- Geller, D., & Marinucci, D. 2010b, arXiv:1006.3835
- Geller, D., & Mayeli, A. 2007, arXiv:0706.3642
- Geller, D., & Mayeli, A. 2008, arXiv:0811.4440
- Geller, D., & Mayeli, A. 2009, arXiv:0907.3164
- Ghosh, T., Delabrouille, J., Remazeilles, M., Cardoso, J., & Souradeep, T. 2010, arXiv:1006.0916
- Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, *ApJ*, **622**, 759
- Guillaoux, F., Faÿ, G., & Cardoso, J.-F. 2007, arXiv:0706.2598
- Guth, A. H. 1981, *Phys. Rev. D*, **23**, 347
- Hansen, F. K., Banday, A. J., Eriksen, H. K., Górski, K. M., & Lilje, P. B. 2006, *ApJ*, **648**, 784
- Hernández, E., & Weiss, G. 1996, *Studies in Advanced Mathematics* (Boca Raton, FL: CRC Press)
- Kerkycharian, G., Petrushev, P., Picard, D., & Willer, T. 2007, *Electron. J. Statist.*, **1**, 30
- Komatsu, E., & Spergel, D. N. 2001, *Phys. Rev. D*, **63**, 063002
- Komatsu, E., et al. 2009, *ApJS*, **180**, 330
- Komatsu, E., et al. 2010, arXiv:1001.4538
- Lan, X., & Marinucci, D. 2008, *Electron. J. Statist.*, **2**, 332
- Lan, X., & Marinucci, D. 2009, *Stochastic Processes & their Applications*, **119**, 3749
- Linde, A. D. 1982, *Phys. Lett. B*, **108**, 389
- Marinucci, D., et al. 2008, *MNRAS*, **383**, 539
- Mather, J. C., et al. 1992, in *AIP Conf. Ser.* 245, *Early Results from the Cosmic Background Explorer*, ed. H. J. Haubold & R. K. Khanna (Melville, NY: AIP), 266
- Mayeli, A. 2008, arXiv:0806.3009
- McEwen, J. D., Hobson, M. P., & Lasenby, A. N. 2006, arXiv:astro-ph/0609159
- McEwen, J. D., Vielva, P., Hobson, M. P., Martínez-González, E., & Lasenby, A. N. 2007, *MNRAS*, **376**, 1211
- Moudden, Y., Cardoso, J. F., Starck, J.-L., & Delabrouille, J. 2005, *EURASIP J. Appl. Signal Process.*, **15**, 2437
- Narcowich, F. J., Petrushev, P., & Ward, J. D. 2006, *SIAM J. Math. Anal.*, **38**, 574
- Natoli, P., et al. 2010, *MNRAS*, **400**, 1266 (arXiv:0905.4301)
- Pesenson, M. Z., Pesenson, I. Z., & McCollum, B. 2010, *Adv. Astron.* (arXiv:1003.0879)
- Pietrobon, D., Amblard, A., Balbi, A., Cabella, P., Cooray, A., & Marinucci, D. 2008, *Phys. Rev. D*, **78**, 103504
- Pietrobon, D., Balbi, A., & Marinucci, D. 2006, *Phys. Rev. D*, **74**, 043524
- Pietrobon, D., Cabella, P., Balbi, A., Crittenden, R., de Gasperis, G., & Vittorio, N. 2009a, *MNRAS*, **402**, L34
- Pietrobon, D., Cabella, P., Balbi, A., de Gasperis, G., & Vittorio, N. 2009b, *MNRAS*, **396**, 1682
- Rudjord, Ø., Hansen, F. K., Lan, X., Liguori, M., Marinucci, D., & Matarrese, S. 2009, *ApJ*, **701**, 369
- Rudjord, Ø., Hansen, F. K., Lan, X., Liguori, M., Marinucci, D., & Matarrese, S. 2010, *ApJ*, **708**, 1321
- Sachs, R. K., & Wolfe, A. M. 1967, *ApJ*, **147**, 73
- Sanz, J. L., Herranz, D., Lopez-Caniego, M., & Argüeso, F. 2006, in *Proc. 14th Eur. Signal Processing Conf. (EUSIPCO 2006)*, ed. F. Gini & E. E. Kuruoglu (arXiv:astro-ph/0609351)
- Sato, K. 1981, *MNRAS*, **195**, 467
- Scodeller, S., Rudjord, Ø., Hansen, F. K., Marinucci, D., Geller, D., & Mayeli, A. 2010, arXiv:1004.5576
- Smidt, J., Amblard, A., Serra, P., & Cooray, A. 2009, *Phys. Rev. D*, **80**, 123005
- Smith, K. M., Senatore, L., & Zaldarriaga, M. 2009, *J. Cosmol. Astropart. Phys.*, **JCAP09(2009)006**
- Smith, K. M., & Zaldarriaga, M. 2006, arXiv:astro-ph/0612571
- Smoot, G. F., et al. 1992, *ApJ*, **396**, L1
- Starck, J.-L., & Bobin, J. 2009, arXiv:0903.3383
- Starck, J.-L., Moudden, Y., Abrial, P., & Nguyen, M. 2006, *A&A*, **446**, 1191
- Starck, J.-L., Moudden, Y., & Bobin, J. 2009, *A&A*, **497**, 931
- Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, *ApJ*, **609**, 22
- Wiaux, Y., McEwen, J. D., Vanderghenst, P., & Blanc, O. 2008, *MNRAS*, **388**, 770