

Power spectrum estimation on the sphere using needlets

Frederic Guilloux^{*†}

Université Pierre et Marie Curie (UPMC), Laboratoire de Statistique Théorique et Appliquée (LSTA), France

E-mail: frederic.guilloux@upmc.fr

Needlet coefficients can be used to estimate the power spectrum of an isotropic field (such as the CMB). It enables local adaptation, both in the spatial and harmonic domains, to the signal-to-noise ratio. Moreover, it provides an efficient way to aggregate observations from multiple experiments and to estimate the power spectrum jointly from all the data sets.

Big Bang, Big Data, Big Computers

September 19-21, 2012

Laboratoire Astroparticule et Cosmologie, 10 rue A. Domon et L. Duquet, 75205 Paris 13, France

^{*}Speaker.

[†]This contribution presents some results which can be found with more details in Faÿ, Guilloux, Betoule, Cardoso, Delabrouille and Le Jeune (Phys. Rev. D, 2008) [6] and Faÿ and Guilloux (Statist. Inference for Stoch. Processes, 2011) [5].

Introduction

In the ideal case of a full-sky and infinitely resolute observation of pure, isotropic, Gaussian CMB fluctuations ΔT , the empirical power spectrum, namely, for $\ell \geq 0$, $\tilde{C}_\ell := \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_{\ell,m}^2$ – where $a_{\ell,m} := \int_{\mathbb{S}} \Delta T(\mathbf{r}) Y_{\ell,m}(\mathbf{r}) d\mathbf{r}$ – would be a sufficient statistic. It would also be the maximum-likelihood estimator of the true power spectrum $C_\ell := E(a_{\ell,m}^2)$.

However, problems occur when one wants to take into account the fact that in reality: - the spatial resolution of the instruments is limited - some experiments observe the sky only partially, others require a mask on highly contaminated regions - there are foreground to be removed, and a non-iid noise - maybe the fluctuations are not really isotropic nor Gaussian - the polarization spectrum should be computed - data come from multiple experiments and multiple detectors for each experiment, with possible calibration issues - the computational cost of the method must be reasonable - etc. The aim of this contribution is to illustrate the possible use of needlets to address some of these issues, namely :

1. the limitations due to beam smoothing and incomplete sky coverage (and/or masking),
2. the heteroscedasticity of the noise (which is still, however, supposed independent from one pixel to another),
3. and, above all, the aggregation of multiple experiments, each with its own specifications (beam, noise level (“hit map”), coverage, map resolution).

In the following, the amplitude of CMB fluctuations are considered, and they are supposed isotropic and Gaussian. The foregrounds are supposed removed (or included in the noise).

The use of needlets for some other of the above mentioned problems are addressed by other contributions (Cf contributions by Remazeilles [13] for foregrounds removal and by Marinucci [11] for Gaussianity tests). For the extension to polarization data, see [7, 3]. Other well-known methods used for classical spectral estimators can be straightforwardly adapted to the present context: for instance, computing the cross-spectra between detectors.

1. The needlet spectral estimators

1.1 Needlets

Starting from a function f on the sphere, the two following operations are equivalent :

1. Smoothing in the harmonic domain and sampling the smoothed map at some point \mathbf{r}_k :

$$f \mapsto \{a_{\ell,m}\} \mapsto \{b(\ell)a_{\ell,m}\} \mapsto \beta_k := \sum_{\ell,m} b(\ell)a_{\ell,m} Y_{\ell,m}(\mathbf{r}_k)$$

for some harmonic window function $b(\ell)$.

2. Taking the scalar product between f and an axis-symmetric function centered at \mathbf{r}_k :

$$f \mapsto \beta_k := \int_{\mathbb{S}} f(\mathbf{r}) \psi_k(\mathbf{r}) d\mathbf{r}$$

where $\psi_k(\mathbf{r}) = \sum_{\ell} b(\ell) L_{\ell}(\langle \mathbf{r}, \mathbf{r}_k \rangle)$, where L_{ℓ} is the Legendre polynomial of degree ℓ .

The first operation ensures reasonable computational cost, thanks to algorithms realizing fast Fourier transforms on the sphere.

The second one, applied for a well-chosen family $\{b_j(\ell)\}_{j \in \mathbb{N}}$ of harmonic window functions and for suitable adapted families of points $\{\mathbf{r}_{j,k}\}_{j \in \mathbb{N}, k=1:\mathbf{npix}_j}$, enables to consider the resulting coefficients $\beta_{j,k}$ as the so-called “needlet” coefficients of f on the frame $\{\psi_{j,k}\}_{j \in \mathbb{N}, k=1:\mathbf{npix}_j}$ and to deduce the properties of these coefficients $\beta_{j,k}$ from the ones of the functions $\psi_{j,k}$, called “needlets”.

In particular, it has been shown [12] that the family of harmonic window functions can be chosen in such a way that the needlets $\psi_{j,k}$ are a tight frame and possess the following localization property:

$$\forall m \in \mathbb{N}, |\psi_{j,k}(\mathbf{r})| \leq \frac{\text{cst}(m)B^j}{(1 + B^j d(\mathbf{r}, \mathbf{r}_{j,k}))^m} \quad (1.1)$$

where B is a constant and $d(\cdot, \cdot)$ is the geodesic distance on \mathbb{S} . That is, the needlet $\psi_{j,k}$ is localized at $\mathbf{r}_{j,k}$ with a tail decreasing at a “quasi-exponential” rate (polynomial of any order).

In the setting assuring (1.1), the harmonic window functions $b_j(\ell)$ are non-negative smooth functions the support of which slips to infinity with increasing length. In the following, j will be referred to the (harmonic) “scale” of the needlets.

Due to the constant in the numerator, (1.1) is useful essentially in the asymptotic of high frequencies, when the scale $j \rightarrow \infty$. The optimization of non asymptotic localization properties of the needlets by the design of the harmonic window functions was addressed in [8].

1.2 Statistical properties

In this subsection, we consider the needlet coefficients of the fluctuation field ΔT , which is supposed zero-mean and isotropic. The following properties are obvious consequences of the two equivalent constructions of needlet coefficients (the third one derives from (1.1)).

Proposition 1. *Let ΔT be a zero-mean, isotropic field on \mathbb{S} with (true) power spectrum C_ℓ . Let $b_j(\ell)$ be harmonic window functions with compact support and, for each j , let $(\mathbf{r}_{j,k}, \lambda_{j,k})_{k=1:\mathbf{npix}_j}$ in $(\mathbb{S} \times \mathbb{R}_+)^{\mathbf{npix}_j}$ be a set of points and quadrature weights assuring exact quadrature of all polynomials of degree in the support of b_j . Then*

(a) *The square of each coefficient $\beta_{j,k}$ at scale j provides an unbiased estimator of the (smoothed) power spectrum:*

$$\lambda_{j,k}^{-1} E(\beta_{j,k}^2) = C_{(j)} \quad (1.2)$$

where $C_{(j)} := \sum_{\ell \in \mathbb{N}} \frac{2\ell+1}{4\pi} b_j(\ell) C_\ell$.

(b) *The coefficients $\beta_{j,k}$ are (exactly) uncorrelated between non overlapping bands:*

$$\mathbf{Cor}(\beta_{j_1, k_1}; \beta_{j_2, k_2}) = 0 \quad (1.3)$$

for all k_1, k_2 , provided that $b_{j_1} b_{j_2} \equiv 0$.

(c) *The coefficients $\beta_{j,k}$ within a given scale j are asymptotically uncorrelated, when $j \rightarrow \infty$, provided that the corresponding points $\mathbf{r}_{j,k}$ are not too close. More precisely, there exists a function $\varphi(j) \xrightarrow{j \rightarrow \infty} 0$ such that*

$$\mathbf{Cor}(\beta_{j, k_1}; \beta_{j, k_2}) \xrightarrow{j \rightarrow \infty} 0 \quad (1.4)$$

as soon as $\frac{d(\mathbf{r}_{j, k_1}, \mathbf{r}_{j, k_2})}{\varphi(j)} \xrightarrow{j \rightarrow \infty} \infty$.

In practice,

- (a) For “reasonable” pixelization schemes, $\lambda_{j,k} \equiv \frac{1}{\text{npix}_j}$.
- (b) Each harmonic window function overlaps only with the previous one and the next one, which implies that (1.3) is valid as soon as $|j_1 - j_2| > 1$.
- (c) At a fixed scale j , we consider that $\mathbf{Cor}(\beta_{j,k_1}; \beta_{j,k_2}) \approx 0$ if k_1 and k_2 are not too close.

More sophisticated statistical properties of the needlet coefficients can be found in [1, 2].

1.3 Needlet spectral estimators (NSE)

In this subsection, we consider the needlet coefficients of a set of observation maps $\{Y_1(\mathbf{r}), \dots, Y_E(\mathbf{r})\}$ where E independent observations (indexed by e) include the same signal of interest ΔT , each with its own specifications: (inhomogeneous) noise level, coverage, beam, resolution. The different observations may come from different missions (e.g. one map from Planck and one from ACBAR) or from different detectors of the same mission. Mathematically, we consider, for $e = 1 : E$, the observations

$$Y_e(\mathbf{r}) = M_e(\mathbf{r})[(B_e * \Delta T)(\mathbf{r}) + \sigma_e(\mathbf{r})Z_e(\mathbf{r})] \quad (1.5)$$

from which we compute the needlet coefficients

$$\beta_{j,k,e} := \int_{\mathbb{S}} Y_e(\mathbf{r}) \psi_{j,k}(\mathbf{r}) d\mathbf{r}.$$

In equation (1.5) M_e is the mask (it can be any function, but in practice, for numerical integration, one will usually prefer to use a smooth function rather than a 0-1 one); B_e is the instrumental beam, supposed axis-symmetric with transfer function $B_{e,\ell}$ and transfer function at scale j $B_{e,(j)}^2 := \sum_{\ell} \frac{2\ell+1}{4\pi} (b_j(\ell))^2 B_{e,\ell}^2$; σ_e is the noise level (in general, proportional to the inverse square root of the hit map); and $Z_e(\mathbf{r})$ are standard random variables.

The needlet coefficients $\beta_{j,k,e}$ are computed only outside the mask M_e , and, even more, only for points \mathbf{r}_k which are far enough from the mask. More precisely, a spatial index k is selected only if $\int_{\mathbb{S}} (1 - W_e(\mathbf{r}))^2 \psi_{j,k}^2(\mathbf{r}) d\mathbf{r}$ is below some threshold (more discussion on the choice of the thresholds are in [8, 6, 5]).

The standard deviation of the effect of the noise on $\beta_{j,k,e}$ is

$$n_{j,k,e} = \left(\int_{\mathbb{S}} \sigma_e^2 |\psi_{j,k}|^2 \right)^{1/2}.$$

Definition 1. *The proposed estimator for the smoothed power spectrum $C_{(j)}$ defined in Proposi-*

tion 1 is:

$$\widehat{C}_{(j)}^{NSE} := \sum_{k=1}^{\mathbf{npix}_j} w_{j,k} \left(\lambda_{j,k}^{-1} \left(\widetilde{\beta}_{j,k} \right)^2 - n_{j,k}^2 \right) \quad (1.6)$$

$$\begin{aligned} \text{where: } \widetilde{\beta}_{j,k} &= \sum_{e=1}^E \omega_{j,k,e} \beta_{j,k,e} \\ \omega_{j,k,e} &\propto (B_{e,(j)}/n_{j,k,e})^2 \\ n_{j,k} &= \left(\sum_e n_{j,k,e}^{-2} \right)^{-1/2} \\ w_{j,k} &\propto (C^0 + n_{j,k}^2)^{-2}, \end{aligned}$$

C^0 being an order of magnitude of $C_{(j)}$.

As demonstrated in [6], $\widehat{C}_{(j)}^{NSE}$ is almost the maximum-likelihood estimate of $C_{(j)}$ given the $\beta_{j,k,e}$.

The intermediate quantities involved in the NSE estimator can be interpreted in the following way:

First step: aggregation of observations.

- $\widetilde{\beta}_{j,k}$ is an aggregated map computed from all the experiments which give observations around the point \mathbf{r}_k , with weights $\omega_{j,k,e}$ taking into account the local (local in space and in the harmonic domain) signal-to-noise ratio of each observation, relatively to the snr ratio of the other observations.

Figure 1 represents the weights $\omega_{j,k,e}$ for a choice of six observations. The specifications of the experiments are quite realistic, but intend only to illustrate the method.

Second step: combination of squared coefficients.

- $n_{j,k}$ is the debasing term, so that for each spatial index $k = 1 : \mathbf{npix}_j$, $\lambda_{j,k}^{-1} \left(\widetilde{\beta}_{j,k} \right)^2 - n_{j,k}^2$ is an unbiased estimator of $C_{(j)}$.

Figure 2 (left) represents these unbiased estimators. Their variance is clearly dependent on the region of the sky.

- the weights $w_{j,k}$ realize the optimal combination of these estimators, giving more importance to regions where the aggregated snr ratio is better.

These weights $w_{j,k}$ are also represented on Figure 2 (right).

1.4 Theoretical (asymptotical) result

In [5], the details of the method are discussed, so that the following asymptotic consistency theorem holds. The assumption that ΔT is Gaussian is needed only for this theorem.

Theorem 1. *If the masks and noise levels in the experiments verify some technical conditions and if the parameters of the method (harmonic window functions, thresholds) are properly chosen, then*

$$\frac{\|\widehat{C}_{(j)}^{NSE} - C_{(j)}\|^2}{C_{(j)}^2} \xrightarrow{j \rightarrow \infty} 0.$$

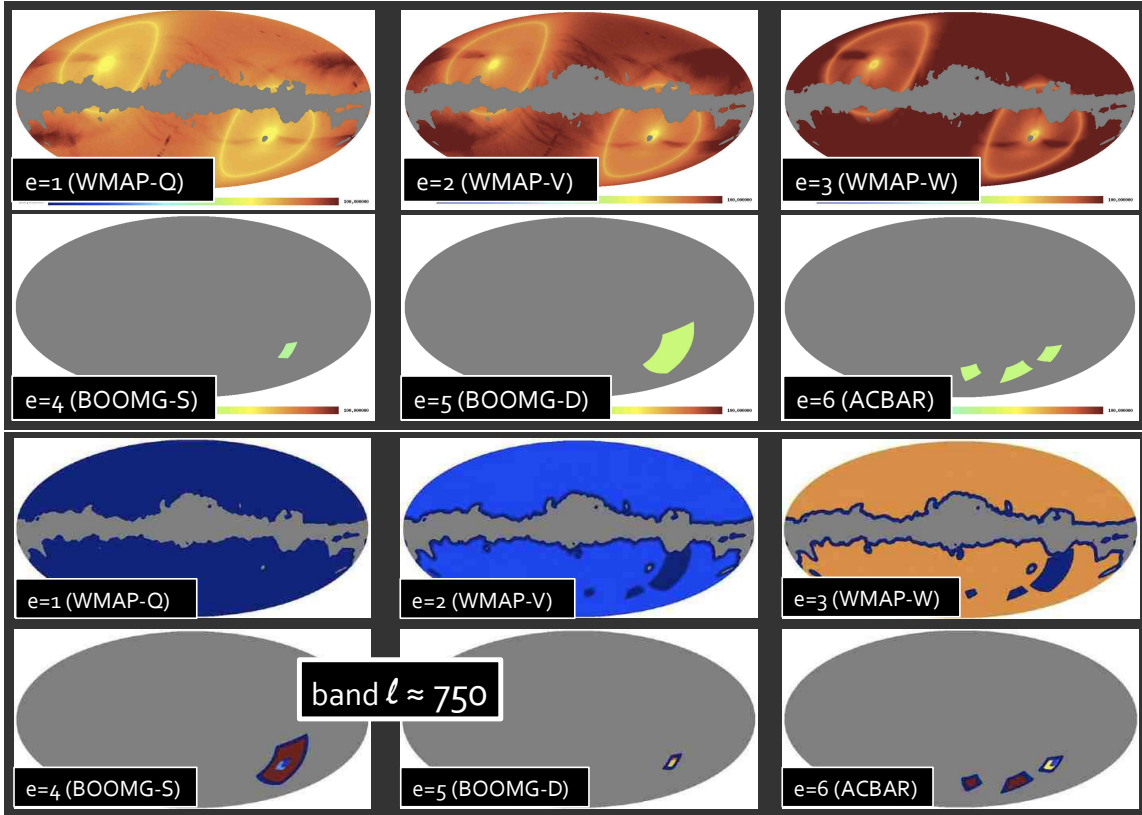


Figure 1: Top: masks M_e (in grey) and noise levels σ_e of six experiments. Bottom: corresponding aggregation weights $\omega_{j,k,e}$ for a given scale j (selecting harmonic multipoles around $\ell \simeq 750$).

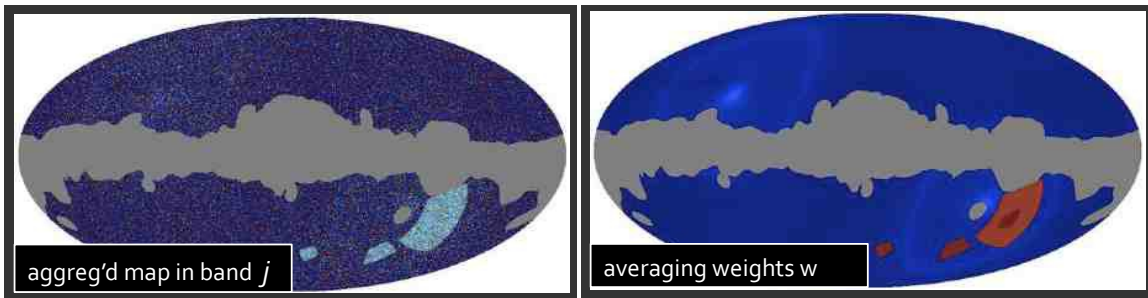


Figure 2: Left: debased aggregated map of needlet coefficients (for the same scale j as in Figure 1), i.e. the quantity into the parentheses in equation (1.6). Right: corresponding combination weights. The NSE estimator is obtained by the summation over the sphere of the product of the left and right maps.

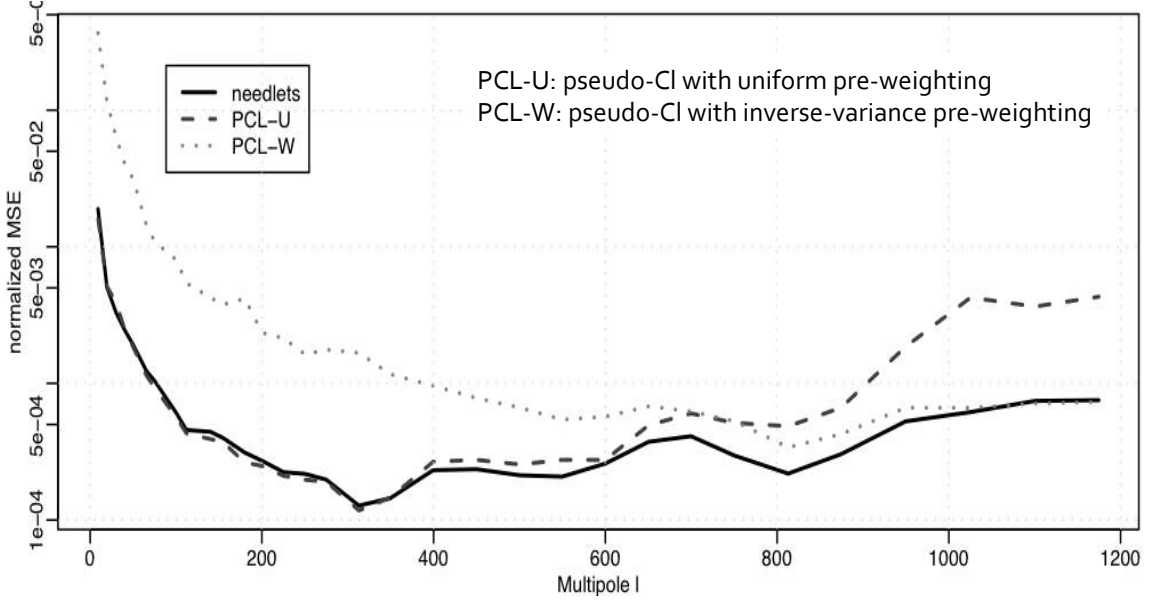


Figure 3: Normalised Mean Squared Error (i.e. $\|\widehat{C}_{(j)}^{NSE} - C_{(j)}\|^2 / C_{(j)}^2$) for three different estimators. The normalized MSE are estimated from simulated data following a model with known mask and noise level map.

2. Simulation studies

2.1 Single experiment (comparison with Pseudo- C_ℓ 's)

In a first simulation study, we consider only one single (WMAP-like) observation ($E = 1$), and we compare our estimator with the “pseudo-CI” (PCL) estimators. The latter are computed in the following way:

- For PCL-U, the empirical spectrum of the data $\widetilde{C}_\ell := \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} (\int_{\mathbb{S}} Y Y_{\ell,m})^2$ is debiased from the effect of the mask and the noise : $\widehat{C}_\ell^{PCL-U} = (\mathcal{M}_{\ell,\ell'})^{-1} \{\widetilde{C}_{\ell'} - N_{\ell'}\}$ where N_ℓ is the power spectrum of the noise and \mathcal{M} is the “coupling matrix” of the mask. Finally, $\widehat{C}_{(j)}^{PCL-U} = \sum_{\ell} \frac{2\ell+1}{4\pi} b_j(\ell) \widehat{C}_\ell^{PCL-U}$.
- For PCL-W, $1/\sigma$ -preweighting is applied to the data before computing and debiasing the empirical spectrum : $\widetilde{Y}(\mathbf{r}) = \sigma^{-1}(\mathbf{r})Y(\mathbf{r})$ and $\widehat{C}_{(j)}^{PCL-W}$ is nothing else but $\widehat{C}_{(j)}^{PCL-U}$ applied to the data \widetilde{Y} , the mask $\sigma^{-1}M$ and homogeneous noise.

More details on these procedures are given, for instance, in [9], and in [6, appendix A].

Figure 3 compares the performance of the three estimators: $\widehat{C}_{(j)}^{PCL-U}$, $\widehat{C}_{(j)}^{PCL-W}$ and $\widehat{C}_{(j)}^{NSE}$. Our estimator behave like the best of the previous ones at low ℓ 's and high ℓ 's, and improves both at medium ℓ 's. In fact, the needlet estimator realizes automatically and blindly something similar to the hybrid estimator of Efstathiou [4].

2.2 Multiple experiments (new possibility of joint estimation)

In a second simulation study, we come back to the whole set of $E = 6$ observations presented in subsection 1.3 (Figures 1-2). The details of the experiments' specifications are in [6]. In this

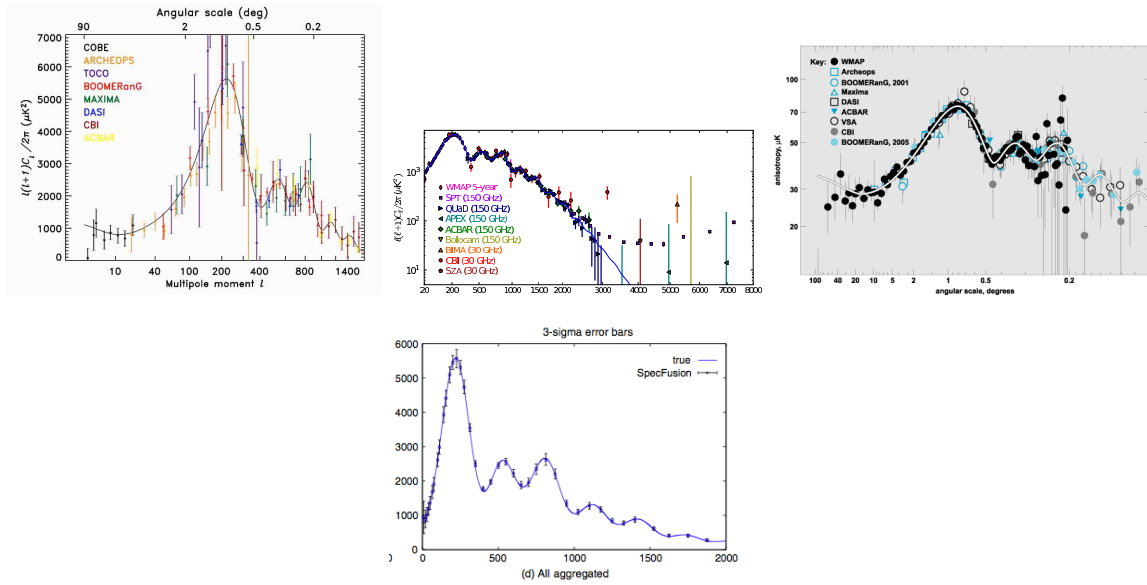


Figure 4: Top: Examples found in the literature of graphical presentations of the estimation of the power spectrum based on multiple experiments. Bottom: NSE estimation of the power spectrum from simulated multiple experiments.

context, as far as I know, there is no other method which uses all the data set to produce one estimated spectrum to which we could compare (if we except, of course, the parametric estimation of the power spectrum *via* the estimation of cosmological parameters from the experiments). The only possible comparison would be between something like the top and the bottom of Figure 4.

In order to understand how the method works, we can apply it separately to the WMAP-like observations alone; then to the BOOMERanG-like ones; and then to the ACBAR-like one. The results are shown in Figure 5 and may be compared to the result obtained by applying the method with all the six observations together (Figure 4, Bottom). At each scale, the latter is better than the estimation from the best experiment considered alone.

Figure 6 shows the correlation (estimated by monte carlo simulations) between this aggregated estimator on the one hand, and the estimators based on a single experiment on the other hand. It illustrates the fact that at low l , one could obtain the same estimation using only WMAP data, since the correlation is $\simeq 1$. As well, at high l , the same estimation is given by ACBAR only. But in a wide range of medium l 's, the aggregated estimator is significantly correlated with 2 or 3 missions, and significantly different from one of this mission taken separately. That is, one would certainly benefit from estimating the power spectrum jointly from the missions taken together.

Conclusion

Estimation of the power spectrum from needlet coefficients has some drawbacks. The principal one is that only a smoothed version of the power spectrum can be estimated, and the choice of this “binning” scheme is not independent of the optimization of the needlets’ localization – and thus, of the performance of the estimation.

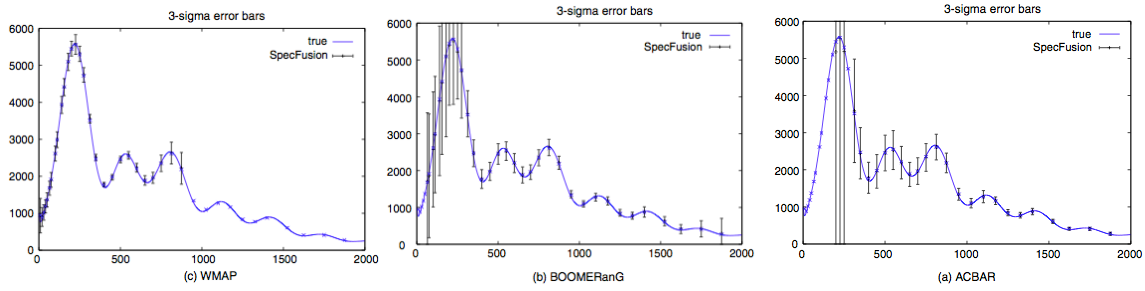


Figure 5: Estimators computed from single experiments observations. For WMAP, a resolution of $n_{\text{side}}=512$ was used, so the numerical spherical harmonic transform can be done only at low ℓ .

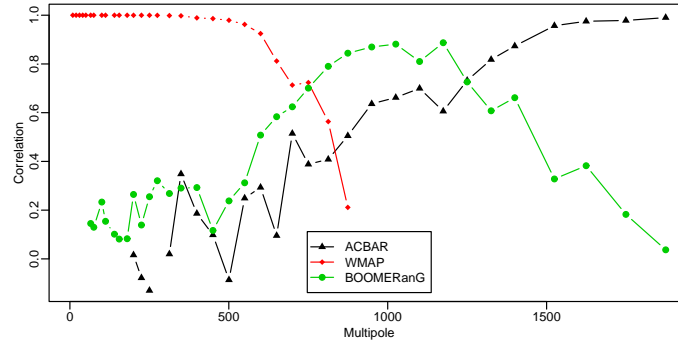


Figure 6: Correlation between estimators using only the observation from a single mission, and the estimator which aggregate the observations from all the missions.

However, the method has important advantages:

1. It fulfills L. Knox's requirements [10]: simple to understand, to implement and to interpret; computationally efficient.
2. It adapts automatically and optimally to the local signal-to-noise ratio – local both in the spatial and the harmonic domains.
3. It provides a way to aggregate complementary data sets in a very natural manner, taking automatically into account the various specifications (beams, masks, noise levels, resolutions) of the observations.

The third point is probably the most crucial. For the moment it was tested only on toy examples, but we are convinced that the method could apply to present and future real data sets. This implies that the data from all the missions are publicly available, with enough information on their specifications.

References

- [1] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard, *Asymptotics for spherical needlets*, *Ann. Statist.* 37(3):1150–1171, 2009 [arXiv: math/0606599].

- [2] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard, *Subsampling needlet coefficients on the sphere*, *Bernoulli* 15(2):438–463, 2009 [arXiv: 0706.4169].
- [3] S. Basak and J. Delabrouille, *A needlet ILC analysis of WMAP 7-year polarisation data: CMB polarisation power spectra* [arXiv: 1204.0292].
- [4] G. Efstathiou, *Myths and truths concerning estimation of power spectra: the case for a hybrid estimator*, *MNRAS* 349(2):603–626, 2004 [arXiv: astro-ph/0307515].
- [5] G. Faÿ and F. Guilloux, *Spectral estimation on the sphere with needlets: high frequency asymptotics*, *Statist. Infer. Stoch. Processes* 14(1):1–25, 2011 [arXiv: 0807.2162].
- [6] G. Faÿ, F. Guilloux, M. Betoule, J.-F. Cardoso, J. Delabrouille, and M. Le Jeune, *CMB power spectrum estimation using wavelets*, *Phys. Rev. D* 78(8):083013, 2008 [arXiv: 0807.1113].
- [7] D. Geller, F. K. Hansen, D. Marinucci, G. Kerkycharian, and D. Picard, *Spin needlets for cosmic microwave background polarization data analysis*, *Phys. Rev. D* 78(12):123533, 2008 [arXiv: 0811.2881].
- [8] F. Guilloux, G. Faÿ, and J.-F. Cardoso, *Practical wavelet design on the sphere*, *Appl. Comput. Harmon. Anal.* 26(2):143–160, 2009 [arXiv: 0706.2598].
- [9] E. Hivon, K. M. Górski, 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*, *Astrophys. J.* 567:2–17, 2002 [arXiv: astro-ph/0105302].
- [10] L. Knox, *Big overview*, (*Big Bang, Big Data, Big Computers*)
- [11] D. Marinucci, *On needlets/wavelets nonGaussianity estimators*, (*Big Bang, Big Data, Big Computers*)
- [12] F. Narcowich, P. Petrushev, and J. Ward, *Localized tight frames on spheres*, *SIAM J. Math. Anal.* 38(2):574–594, 2006.
- [13] M. Remazeilles, *Reconstruction of high-resolution SZ maps from heterogeneous datasets using needlets*, (*Big Bang, Big Data, Big Computers*)