Non-Gaussianity and Minkowski functionals: forecasts for Planck

A. Ducout,1⋆ F. R. Bouchet,1 S. Colombi,1 D. Pogosyan2 and S. Prunet1

1Institut d’Astrophysique de Paris, CNRS & UPMC, 98 bis boulevard Arago, F-75014 Paris, France
2Department of Physics, University of Alberta, 11322-89 Avenue, Edmonton, Alberta T6G 2G7, Canada

Accepted 2012 November 23. Received 2012 November 19; in original form 2012 September 4

ABSTRACT

We study Minkowski functionals as probes of primordial non-Gaussianity in the cosmic microwave background, specifically for the estimate of the primordial ‘local’ bi-spectrum parameter $f_{NL}$, with instrumental parameters which should be appropriate for the Planck experiment. We use a maximum likelihood approach, which we couple with various filtering methods and test thoroughly for convergence. We included the effect of inhomogeneous noise as well as astrophysical biases induced by point sources and by the contamination from the Galaxy. We find that when Wiener filtered maps are used (rather than simply smoothed with Gaussian), the expected error on the measurement of $f_{NL}$ should be as small as $\Delta f_{NL} \simeq 10$ when combining the three channels at 100, 143 and 217 GHz in the Planck extended mission setup. This result is fairly insensitive to the non-homogeneous nature of the noise, at least for realistic hitmaps expected from Planck. We then estimate the bias induced on the measurement of $f_{NL}$ by point sources in those three channels. With the appropriate masking of the bright sources, this bias can be reduced to a negligible level in the 100 and 143 GHz channels. It remains significant in the 217 GHz channel, but can be corrected for. The Galactic foreground biases are quite important and present a complex dependence on sky coverage: making them negligible will depend strongly on the quality of the component separation methods.

Key words: methods: numerical – methods: statistical – cosmic background radiation.

1 INTRODUCTION

The Planck satellite was launched in 2009 May (Tauber et al. 2010; Planck Collaboration et al. 2011a) to observe the microwave sky and particularly the cosmic microwave background (CMB). It will provide major pieces of information about the early evolution of the Universe and the origin of structures thanks to its unprecedented combination of resolution, sensitivity and spectral coverage. One of the important objectives of this mission is to bring new constraints on inflation theory and primordial non-Gaussianities (NGs). In the simplest inflationary models based on a single slowly rolling scalar field, primordial fluctuations should be only weakly non-Gaussian (Maldacena 2003; Bartolo, Matarrese & Riotto 2006). Yet in more general models a much higher level of NG is expected. We can cite models with multiple scalar fields, features in the inflaton potential, non-adiabatic fluctuations, non-standard kinetic terms, warm inflation, deviations from Bunch–Davies vacuum (review in Bartolo et al. 2004) or completely different mechanisms, for example topological defects such as cosmic strings (Kaiser & Stebbins 1984). NG is often parametrized in a phenomenological way by the non-linear ‘local’ coupling parameters $f_{NL}, g_{NL}, \ldots$ which appear in the perturbative development of the primordial curvature perturbation (Komatsu & Spergel 2001; Okamoto & Hu 2002),

$$\Phi(x) = \phi_L(x) + f_{NL} (\phi_L^2(x) - \langle \phi_L^2(x) \rangle) + g_{NL} \phi_L^4(x) + \cdots ,$$

(1)

where $\phi_L(x)$ is the linear Gaussian part of the Bardeen curvature. We will focus here on constraining the first parameter, $f_{NL}$.

The latest results from Wilkinson Microwave Anisotropy Probe (WMAP) (Komatsu et al. 2011) with 7 years of data provide the constraint

$$f_{NL} = 32 \pm 21$$

(2)

at the 68 per cent confidence level. With Planck, we expect the error on $f_{NL}$ to be at best of the order of 2.5 from both temperature and polarization (Sefusatti et al. 2009; Yadav & Wandelt 2010) or of the order of 5 for the temperature only (Komatsu & Spergel 2001). These constraints on $f_{NL}$ were obtained with bi-spectrum measurements (Komatsu, Spergel & Wandelt 2005; Yadav, Komatsu & Wandelt 2007; Yadav et al. 2008). The CMB bi-spectrum is the harmonic transform of the three-point correlation function and it has been shown that this is theoretically an optimal estimator for $f_{NL}$ as it saturates the Cramér–Rao inequality for a weak NG (Babich 2005).

However, alternative statistics to the bi-spectrum have been also developed, which can at least serve as checks and diagnoses of

E-mail: ducout@iap.fr

© 2012 The Authors

Published by Oxford University Press on behalf of the Royal Astronomical Society
the results obtained from the bi-spectrum. One of the main reasons to study various probes of NG is indeed that they are affected differently by different systematics and contaminants such as inhomogeneous noise and foreground residuals induced by our Galaxy and point sources, as well as secondary anisotropies (e.g., Aghanim, Majumdar & Silk 2008, and references therein) such as integrated Sachs–Wolfe effect, Sunyaev–Zeldovich effect (Sunyaev & Zeldovich 1972) and weak gravitational lensing, which can all contribute to NG of the CMB in a non-trivial way. In particular, the most serious bias on $f_{NL}$ from secondary anisotropies arises from the coupling between weak gravitational lensing and integrated Sachs–Wolfe effect (see Goldberg & Spergel 1999; Serra & Cooray 2008; Hanson et al. 2009; Mangilli & Verde 2009; Munshi & Heavens 2010), which adds complexity to the analyses.

Here we focus on a complete set of topological tools, the Minkowski functionals (MFs) introduced in cosmology by Mecke, Buchert & Wagner (1994). MFs describe the morphological features of random fields over excursion sets, i.e. regions where the field exceeds some threshold level $v$. These well-known probes of primordial NGs (Schmalzing & Buchert 1997; Schmalzing & Gorski 1998; Winitzki & Kosowsky 1998) have been widely used in two and three dimensions, for instance on WMAP CMB data (Hikage, Komatsu & Matsubara 2006; Hikage et al. 2008) and on the Sloan Digital Sky Survey galaxy catalogue (Park et al. 2005; Hikage et al. 2006).

For CMB studies, MFs provide a nice complement to the bi-spectrum for several reasons. First, at variance with the bi-spectrum, which evolves in harmonic (or Fourier) space, MFs are defined in real space, which makes a robust implementation for MFs in practice much easier than for the bi-spectrum. Secondly, MFs are sensitive to the full hierarchy of higher-order correlations, instead of third order only, and can provide additional information on all the non-linear coupling parameters beyond the sole $f_{NL}$. Because they are measured on excursions of the density field smoothed with an isotropic window, MFs only probe angular averages of higher order statistics, leaving out the angular dependences, at variance with the bi-spectrum. This explains why MFs are less optimal than the bi-spectrum in terms of disentangling between various models of weak NG. However, the very nature of these angular averages considerably eases statistical analysis and reduces the number of parameters while performing maximum likelihood analysis. As a result, even though suboptimal, MFs represent powerful tools of investigation of NG. For instance, the first limit on the primordial NG in the isocurvature perturbation was measured with MFs, even though theoretical predictions were computed on the bi-spectrum (Hikage et al. 2009). Note finally that the constraint on $f_{NL}$ obtained from the analysis of WMAP3 with MFs reads

$$f_{NL} = 11 \pm 40$$

at the 68 per cent confidence level (Hikage et al. 2008), a rather competitive result, after all, compared to the bi-spectrum constraints from WMAP7 (equation 2).

In this paper, we investigate constraints on $f_{NL}$ obtained with MFs from a practical point of view, in order to ease at best future application of MFs to Planck data. After detailing the MF estimators and the Bayesian method we implemented, we review the effects of the main systematics that can affect/bias the results. These systematics can be of instrumental nature – inhomogeneous noise, beams – or of astrophysical nature: the observed microwave sky contains foreground emission from the Milky Way and from extragalactic sources.

Our Galaxy is indeed a strong source of contamination. It is generally accounted for by masking and/or by using component separation methods and by assuming that the final residual bias on $f_{NL}$ is negligible compared to error bars (Komatsu et al. 2003; Hikage et al. 2008). Nevertheless it has been shown that component separation leaves Galactic features in CMB maps (Chiang et al. 2003). Indeed, the various methods of component separation (Leach et al. 2008) need to be controlled at the new level of accuracy on $f_{NL}$ expected for Planck. Another approach is to use foreground templates and to marginalize over them (Komatsu et al. 2002, 2011). Here, we chose to analyze the behaviour of MFs when using masks and a naive model of component separation. For that, we used foreground templates provided by the Planck Sky model (PSM) code (Delabrouille et al. 2012).

Point sources are mostly unresolved galaxies (at least in low-density regions of Galactic emission), some emitting a signal sufficiently high to be detected individually, and the others forming a diffuse unresolved background. Point sources consist first of radio galaxies, active galactic nuclei which emit in radio frequencies with synchrotron process. They can also be dusty starburst galaxies which are observed via the thermal emission of their dust heated by the ultra-violet emission of young stars. Current instruments are not able to detect all these galaxies individually and the integrated emission of all the faint galaxies forms a diffuse background, the cosmic infrared background (CIB) which has been recently studied and observed (Puget et al. 1996; Lagache & Puget 2000; Planck Collaboration et al. 2011c). The strong point sources are accounted for by masking them but the fainter ones can still induce biases in NG studies. The point source bi-spectrum has already been measured (Komatsu et al. 2011) and its effect in the estimates of $f_{NL}$ evaluated, for bi-spectrum measurements (Babich & Pierpaoli 2008; Serra & Cooray 2008; Lacasa et al. 2012).

In addition to these contributions, as mentioned above, CMB contains secondary anisotropies imprinted between the surface of last scattering and present time. While we leave the treatment of these secondary anisotropies for future work, we shall treat in detail the contamination from our Galaxy and point sources.

This paper is organized as follows. In Section 2, we review the Bayesian method we use to optimize the constraint on $f_{NL}$ from MFs. In Section 3, we study the effects of pixelization and filtering. Section 4 deals with the effect of inhomogeneous noise. In Sections 5 and 6, we study astrophysical sources of systematic effects: first point sources, then Galactic foregrounds, Section 7 summarizes and discusses the results. For completeness, some additional technical details can be found in Appendix A, which discusses analytic predictions for MFs in the weakly non-Gaussian regime, Appendix B, which details with the algorithm used to measure MFs, Appendix C, with a convergence study of our $\chi^2$, and Appendix D, which provides some details on the Planck simulation experiments performed in this work.

The pixelization scheme adopted in this paper is the same as in Planck processed maps, namely HEALPix1 (Górski et al. 2005).

2 METHOD

In this section, we set notations by first recalling basics of MFs and their measurement inside an excursion of varying height (Section 2.1). Then we detail the Bayesian approach we use to estimate $f_{NL}$ (Section 2.2) by comparing the measurement $\hat{f}$ of a functional or

---

1 Available at http://healpix.jpl.nasa.gov
a combination of functionals to its expected value. Within a set of reasonable simplifying assumptions, in particular weak NG, this approach reduces to a simple $\chi^2$ test (Section 2.2.1). The numerical estimate of this $\chi^2$ requires the accurate calculation of the covariance matrix of our estimator $\hat{y}$ from a large set of Gaussian simulations, as discussed in Section 2.2.2. Additional details concerning the Gaussian simulations set up as well as the method used to test the convergence of the $\chi^2$ are given in Appendix C. Once we have a numerically robust $\chi^2$ test, we study its sensitivity to important parameters of the problem, such as the number of bins used to explore the excursion and the range of the excursion levels, as well as the choice of the functionals and/or their combination (Section 2.3).

2.1 Minkowski functionals

For a two-dimensional field $\delta$ of zero mean and of variance $\sigma^2$, defined on the sphere, and smoothed with a window of typical size $\ell$ (to be defined later), we consider an excursion set of height $v = \delta/\sigma$, i.e. the set of points where the field exceeds the threshold $v$. In what follows, we shall study the topological properties of the excursion with four quantities, denoted by $V_k$ ($k = 0, 1, 2, 3$). The first three ones correspond to MFs: $V_0$ is the fractional Area of the regions above the threshold, $V_1$ is the Perimeter of these regions and $V_2$ is the Genus, defined as the total number of connected components of the excursion above the threshold minus the total number of connected components under the threshold. The fourth one, $V_3$ that will be called the Number of clusters, also noted $N_{\text{clusters}}$, is just the number of connected regions above the threshold for positive thresholds and reversely for negative thresholds (these regions are also known as Betti numbers or as cold/hot spots; Chingangbam et al. 2012). Analytic formulae (theoretical expectation values) for the quantities $V_k$ are known for Gaussian and weakly non-Gaussian fields and are summarized in Appendix A. An interesting property of these functionals is that the pure spectral dependence can be factorized out:

$$V_k(v) = A_k v_k(v),$$

where the amplitude $A_k$ is determined by the shape of the power spectrum of the field fluctuations. The renormalized functionals $v_k$ depend only on the non-Gaussian corrections, i.e. on the behaviour of correlation functions beyond second order. The analysis of $v_k$ allows us to focus on NG, which is the goal of this paper.

In what follows, we denote by $\tilde{X}$ an estimator of the quantity $X$. We calculate $\tilde{v}_k = \tilde{V}_k/A_k$ on a pixelized map by proceeding as described in Appendix B to estimate $\tilde{V}_k$. The quantity $\tilde{A}_k$ itself is obtained by direct measurement of $\sigma_0$ and $\sigma_1$ on each map (see Appendix A for definitions).

2.2 Testing non-Gaussianity with a $\chi^2$

To provide constraints on NG, we adopt a Bayesian approach similar to that of Hikage et al. (2008), comparing measurement of normalized functionals $v_k$ on the data map to theoretical predictions of these functionals for a given model of NG. Here the predictions are obtained from the average of measurements of $v_k$ on a large number of non-Gaussian simulations. We will focus now on the non-linear coupling parameter ‘local’ $f_{\text{NL}}$ but the method can be applied to other types of NGs.

To perform the measurements, we consider an ensemble of $n_{\text{bins}}$ threshold values $v_i$ ranging from $-v_{\text{max}}$ to $+v_{\text{max}}$, defining a vector $\tilde{v} = [\tilde{v}(v_1), \ldots, \tilde{v}(v_{n_{\text{bins}}})]$ for each of the functionals. The statistics under consideration will then be a vector of $n = n_{\text{functionals}} \times n_{\text{bins}}$ elements, $\tilde{y} = \tilde{v}$ if only one functional is used in the analysis, or $\tilde{y} = [\tilde{v}_1, \tilde{v}_2, \ldots]$ if a combination of $n_{\text{functionals}} > 1$ functionals is considered.

The Bayes formula is written as

$$P(f_{\text{NL}} | \tilde{y}) = \frac{P(\tilde{y} | f_{\text{NL}})P(f_{\text{NL}})}{\int P(\tilde{y} | f_{\text{NL}})P(f_{\text{NL}}) \, df_{\text{NL}}}.$$  

In what follows, we shall take a flat prior for $f_{\text{NL}}$, with $P(f_{\text{NL}})$ a constant, while the evidence $\int P(\tilde{y} | f_{\text{NL}})P(f_{\text{NL}}) \, df_{\text{NL}}$ will be just considered as a normalization.

2.2.1 Construction of a $\chi^2$ test

We assume that the likelihood $P(\tilde{y} | f_{\text{NL}})$ is a Gaussian, which allows us to define a simple $\chi^2$ test for $f_{\text{NL}}$. This approximation is expected to be good in the regime where the fluctuations of each component $v_i$ of the vector $\tilde{y}$ remain small compared to its ensemble average $\langle v_i \rangle$, when one considers a large number of realizations of the random field. Keeping these fluctuations small sets practical constraints on (i) the value of $v_{\text{max}}$, which should not be too extreme to avoid sensitivity to rare events, (ii) the smoothing scale $\ell$, which should be small enough to probe a sufficient number of independent harmonic modes on the sky. Because we consider only small values of $\ell$ in most of what follows our main concern is the choice of $v_{\text{max}}$. We shall restrict to the range $v_{\text{max}} < 5$ (see also the discussion in Gott et al. 1990).

With all these simplifications, the posterior becomes

$$P(f_{\text{NL}} | \tilde{y}) \propto \exp \left[ -\frac{\chi^2(\tilde{y}, f_{\text{NL}})}{2} \right]$$

with

$$\chi^2(\tilde{y}, f_{\text{NL}}) \equiv \left[ \tilde{y} - \langle y(f_{\text{NL}}) \rangle \right] C^{-1} \left[ \tilde{y} - \langle y(f_{\text{NL}}) \rangle \right]$$

where $\langle y(f_{\text{NL}}) \rangle$ is the model under test. It can be derived analytically in the weakly non-Gaussian regime (see Hikage et al. 2008, and Appendix A) or by direct measurements on simulations. Since we want to take into account complex additional contributions such as inhomogeneous noise and masks, we use here the more flexible second approach with non-Gaussian simulations provided by Elsner & Wandelt (2009). So $\langle y(f_{\text{NL}}) \rangle \equiv \langle \langle y(f_{\text{NL}}) \rangle \rangle$ is the mean of $y$ measured over $m_{\text{NG}}$ reference maps with a given level of NG $f_{\text{NL}}$. In practice, we shall take

$$m_{\text{NG}} = 200.$$  

Since NG is weak (Komatsu et al. 2011), one can compute the covariance matrix $C$ in a Monte Carlo fashion relying on simulations of Gaussian maps of the CMB (to which we add, when relevant, beam effects, noise, masks). In that case, the posterior probability distribution function $P(f_{\text{NL}} | \tilde{y})$ is expected to be very close to a Gaussian, as illustrated by Fig. 1 for the Perimeter. Indeed, in the weakly non-Gaussian regime, $\langle y(f_{\text{NL}}) \rangle$ is, at first order, a linear function of $f_{\text{NL}}$ (see e.g. Appendix A).

To estimate the most likely posterior value of $f_{\text{NL}}$ and an uncertainty, instead of trying to find iteratively the maximum of $P(f_{\text{NL}} | \tilde{y})$ followed by an estimate of the local curvature using e.g. the Fisher
information matrix to estimate the error, we prefer, for simplicity and robustness, to compute directly the average and the variance of the posterior distribution, \( P(f_{NL}|\hat{y}) \). This latter is estimated using a number of equally spaced bins, \( f_{NL} = f_i \{ i = 1, \ldots, n_{\text{max}} \} \), with for the purpose of this paper \( n_{\text{max}} = 21 \). From this approximation of the posterior, the average and the variance are directly estimated numerically via the simple formulæ

\[
\hat{f}_{NL} = \frac{\sum_i f_i P_i}{\sum_i P_i},
\]

\[
(\Delta f_{NL})^2 = \frac{\sum_i (f_i - \hat{f}_{NL})^2 P_i}{\sum_i P_i},
\]

where \( P_i \) is proportional to \( P(f_{NL} = f_i|\hat{y}) \).

Finally, one might consider performing a number of realizations of the data \( \hat{y} \) – also called in this paper ‘test’ maps – to have a more accurate description of the ‘typical’ expected posterior probability. However such a forecasting is not obvious, as we have to define a frequentist average over Bayesian quantities, to predict the typical value expected for \( f_{NL} \) and \( \Delta f_{NL} \). Our choice, equivalent to Fisher information forecast in the pure Gaussian case, consists in estimating \( f_{NL} \) and \( \Delta f_{NL} \) from the following averages:

\[
\langle \hat{f}_{NL} \rangle = \frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} f_{NL,i},
\]

\[
\langle \Delta f_{NL}^2 \rangle = \frac{1}{m_{\text{test}} - 1} \sum_{i=1}^{m_{\text{test}}} \Delta f_{NL,i}^2,
\]

where \( \hat{f}_{NL,i} \) and \( \Delta f_{NL,i} \) are obtained from the \( \chi^2 \) analysis of ‘data’ map number \( i \) using equations (10) and (11). In what follows, we shall always take (for each individual value of \( f_{NL} \) considered in the test maps)

\[
m_{\text{test}} = 200.
\]

### Table 1.

| \( v_{\text{max}} \) | \( \Delta v \) | \( n_{\text{bins}} \) | \( m \) |
|----------------------|-----------------|-----------------|-----|
| 5                    | 1               | 11              | \( 7.0 \times 10^3 \) |
| 3.5                  | 0.7             | 26              | \( 8.0 \times 10^3 \) |
| 2                    | 0.4             | 51              | \( 9.0 \times 10^3 \) |
| 1                    | 0.14            | 101             | \( 10.0 \times 10^3 \) |
| 0.7                  | 0.04            |                 | \( 15.0 \times 10^3 \) |

2.2.2 Convergence issues

The covariance matrix is estimated from the average over \( m \) Gaussian simulations, on each of which a \( \chi^2 \) vector of functionals is measured:

\[ C_{ij} = \langle (\chi^2_{ij} - \bar{\chi}^2_{ij}) (\chi^2_{ji} - \bar{\chi}^2_{ji}) \rangle, \]

with \( \bar{\chi}^2_{ij} \equiv \langle \chi^2_{ij} \rangle \).

The calculation of the \( \chi^2 \) requires the inversion of the covariance matrix, an uneasy task, because \( C \) can be ill-conditioned, especially if \( m \) is not large enough. The number of simulations required to have a good estimate of the \( \chi^2(\hat{y}, f_{NL}) \) function indeed depends on the number of bins, \( n_{\text{bins}} \), on the functionals under consideration, their number, \( n_{\text{functionals}} \), and on the choice of \( v_{\text{max}} \). The way we estimate \( m \) is exposed in detail in Appendix C. The results of our analyses are summarized in Table 1, where the minimum values of \( m \) required for having a better than two per cent accuracy on the \( \chi^2 \) obtained from the perimeter are displayed for various realistic set-ups in terms of the number of bins and of the excursion range. Similar orders of magnitude are expected for other functionals or when combining a set of functionals, as explicitly checked for the combination \( V_1 + V_2 \), \( n_{\text{bins}} = 26 \) and \( v_{\text{max}} = 3.5 \), where the corresponding value found in Table 1 remains unchanged within 5 per cent. This latter property comes from the fact that in practice, two different functionals are much less correlated than the same functional estimated at two successive bins of the excursion: combining functionals doubles the dimensions of matrix \( C \) but does not make it significantly more degenerate. In the subsequent calculations performed in this paper, we shall take

\[ m = 10,000, \]

a number of simulations sufficient for a good estimate of the \( \chi^2 \) for \( n_{\text{bins}} \lesssim 50 \) and \( v_{\text{max}} \gtrsim 3.5 \).

2.3 Sensitivity of the estimator

In this section we discuss the sensitivity of our \( \chi^2 \) estimator, that is the uncertainty on the estimated \( f_{NL} \), with respect to the number of bins, \( n_{\text{bins}} \), the excursion range, \( v_{\text{max}} \), and the set of functionals used, whether it is the Area, the Perimeter, the Genus, the Number
Table 2. Investigation of the best combination for \((v_{\text{max}}, r_{\text{bins}})\). This table gives the uncertainty in \(f_{\text{NL}}\) for different values of the number of bins \(n = r_{\text{bins}}\) and the threshold \(v_{\text{max}}\). The calculations are performed for the Perimeter, which is the most sensitive to \(f_{\text{NL}}\). The results would be nearly the same for the Genus and \(N_{\text{clusters}}\), while the Area is quite insensitive to \(n\) and \(v_{\text{max}}\) for the range of values tested here. From this table it is fairly easy to conclude that the combination \((v_{\text{max}}, r_{\text{bins}}) = (3.5, 26)\) is close to optimal. In particular it is not necessary in practice to go beyond \(r_{\text{bins}} \geq 26\). Note for completeness that the error bars computed here correspond to the specific set up of Appendix C (see also the caption of the next table).

| \(v_{\text{max}}\) | \(n = 11\) | \(n = 26\) | \(n = 51\) | \(n = 101\) |
|-------------------|----------|----------|----------|----------|
| 5                 | 29       | 25.5     | 25       | 25       |
| 3.5               | 27       | 25       | 25       | 25       |
| 2                 | 37.5     | 37       | 37       | 36.5     |

Table 3. Sensitivity of the estimator: error bars on the measurement of \(f_{\text{NL}}\) for each functional, and for various combinations. The calculations are performed assuming \(n = n_{\text{functional}} \times r_{\text{bins}}\) with \(r_{\text{bins}} = 26\) and \(v_{\text{max}} = 3.5\). Note for completeness that the error bars computed here correspond to the specific set up of Appendix C (see also Appendix D): extended mission for the three channels and noise filtering with a Gaussian window of FWHM \(\sigma_{\text{FWHM}} = 10\) arcmin and HEALPix resolution parameter \(N_{\text{side}} = 2048\).

| Functionals \(V_k\) | \(\langle \Delta f_{\text{NL}}^2 \rangle^{1/2}\) |
|---------------------|-------------------|
| \(V_0\)             | 44                 |
| \(V_1\)             | 25                 |
| \(V_2\)             | 25                 |
| \(V_3\)             | 28                 |
| \(V_1 + V_2\)       | 20                 |
| \(V_1 + V_2 + V_3\) | 19                 |
| \(V_1 + V_2 + V_3 + V_0\) | 18.5              |

of clusters, or an arbitrary combination of any of them. Our simulation setup is the same as in the previous paragraph (and detailed in Appendix C). We checked, in this full sky configuration with homogeneous noise, that our estimator is unbiased, \(\langle \hat{f}_{\text{NL}} \rangle = f_{\text{NL}}\) (within numerical limits set by the finiteness of \(m_{\text{NG}}\) and \(m_{\text{test}}\)).

The results of our analyses are displayed in Tables 2 and 3. They show that the combination \((v_{\text{max}}, r_{\text{bins}}) \simeq (3.5, 26)\) (17) is fairly optimal and shall represent our choice in the subsequent analyses. Note that it is not important to notice that assuming the covariance matrix to be diagonal, as done for instance in Eriksen et al. (2004) and Gott et al. (1990), is not a good approximation in the case of \(f_{\text{NL}}\) and decreases by more than a factor of 2 the constraining power of the \(\chi^2\).

The comparison between various functionals leads to the following ranking in term of discriminative power:

Perimeter \(\gtrsim\) Genus \(\gtrsim N_{\text{clusters}} \gg\) Area. (18)

While the Perimeter, the Genus and the Number of cluster present the same order of sensitivity, the Area is about twice less discriminative than them. Most of the information on \(f_{\text{NL}}\) can be extracted by a combined analysis of the perimeter and the genus, \(V_1 + V_2\), with a little improvement when taking into account the number of clusters, \(V_1 + V_2 + V_3\), but the area does not carry significant pieces of additional information compared to the three others.

3 FILTERING AND SMOOTHING

The measurement of MFs requires smoothing of temperature maps in order to remove the contribution of the noise. This smoothing is performed at various scales on a pixelized map in order to extract all the statistical information available. First, one has to deal with discreteness effects brought by the pixelization. In Section 3.1, we show that for practical measurements, it is not needed to have a large value of HEALPix resolution parameter \(N_{\text{side}}\) to extract all the relevant information even if this means a pixel size comparable to the size of the smoothing window: this is because we explicitly account for these pixelization effects in the model. In Section 3.2, we consider Gaussian smoothing and show that most information on \(f_{\text{NL}}\) obtained from MFs is present at the smallest possible scales, i.e. at scales comparable to the size of the beam of the instrument. Note, however, that this result stands for local \(f_{\text{NL}}\) and might vary for different types of NG. Finally, Section 3.3 deals with the Wiener filtering for the field, its first and second derivatives. We show that the results obtained with a simple combination of Wiener filters set much better constraints on \(f_{\text{NL}}\) than Gaussian smoothing at various scales.

In what follows, Gaussian smoothing scale will be expressed in units of the full width at half-maximum (FWHM), \(\sigma_{\text{FWHM}}\) (see Appendix D2 for more details).

3.1 Pixelization effects: choice of \(N_{\text{side}}\)

In principle, the pixel size should be small compared to the smoothing kernel size in order to avoid systematic errors introduced by the discrete nature of the pixelization (see e.g. Colombi, Pogosyan & Souradeep 2000; Novikov, Colombi & Doré 2006). We checked, in practice, that no bias is introduced on the measurement of \(f_{\text{NL}}\) even when the pixel size becomes comparable to the smoothing window size, because the defects of the pixelization are present as well in the model. However, introducing larger pixels is similar to coarse graining and decreases the effective level of statistical richness, which in turn increases the uncertainty on \(f_{\text{NL}}\). Furthermore, a too large pixel size would simply make the additional filtering operation inoperative: instead, what we would get in that case is simply the dominant part of filtering to be a convolution with a top-hat of the pixel shape.

Table 4 gives the error obtained on \(f_{\text{NL}}\) when using the combination of all four functionals, for various values of HEALPix resolution parameter \(N_{\text{side}}\), different levels of noise and a Gaussian smoothing with \(\sigma_{\text{FWHM}} = 5\) arcmin. For Planck purpose, we also find that the improvement brought by the \(N_{\text{side}} = 2048\) resolution is tiny, as expected for higher level of noise, and this table shows that it is not needed in practice to go beyond \(N_{\text{side}} = 1024\), which is rather handy computationally speaking. From now on, unless specified otherwise, we shall assume \(N_{\text{side}} = 1024\) (19) for all subsequent analyses.

3.2 Gaussian smoothing

Smoothing with a Gaussian kernel depends only on the angular scale, \(\sigma_{\text{FWHM}}\). \textit{A priori}, each type of NG is characterized by a specific scale range where the sensitivity of the estimator of \(f_{\text{NL}}\) is the best and we have to stay aware that we are limited here to only one particular case of NG, although quite typical.
We tested different smoothing scales, including even very small scales \((\theta^S_{\text{FWHM}} = 5 \text{ arcmin})\), smaller than size of the beam \((\theta^B_{\text{FWHM}} = 7.2 \text{ arcmin in the case of the combined channels})\), and just above the size of the pixel, which might look awkward, but in fact does not introduce any significant bias on the measurement of \(f_{\text{nl}}\), as already argued in Section 3.1 when discussing about pixelization effects.

One issue about smoothing at large scales is that it reduces the number of independent modes available on the sky. This in turns reduces the quality of the measurement of the MFs in the tails (large values of \(|v|\)) and can make the likelihood function non-Gaussian, as discussed in the beginning of Section 2.2.1, particularly if \(v_{\text{max}}\) is too large. For instance, with \(v_{\text{max}} = 4\), the Gaussian assumption for the likelihood is legitimate for \(\theta^S_{\text{FWHM}} = 5 \text{ arcmin} \) but not for \(\theta^S_{\text{FWHM}} = 40 \text{ arcmin}\). Here we checked that our \(\chi^2\) analysis was still valid and well converged at smoothing scales as large as \(\theta^S_{\text{FWHM}} = 40 \text{ arcmin}\) for our default choice for the parameters, \(v_{\text{max}} = 3.5, \theta_{\text{max}} = 26\) and \(m = 10,000\).

Table 5 summarizes the results of our analyses in the case where the combination of all functionals is used at various scales or various combinations of scales. Note that when combining four scales, we have a large number of entries in the data vector, \(n = n_{\text{mins}} \times n_{\text{functionals}} \times n_{\text{scales}} \simeq 400\), but we checked that this did not affect the convergence of the calculation of the covariance matrix with \(m = 10,000\). The results of Table 5 show that there is no significant statistical information for \(\theta^S_{\text{FWHM}} \gtrsim 20 \text{ arcmin}\) Most of

\[
\begin{array}{cccc}
\text{Noise} & N_{\text{side}} = 512 & N_{\text{side}} = 1024 & N_{\text{side}} = 2048 \\
\ell_{\text{max}} & 1000 & 2000 & 3500 \\
0.33 \mu\text{K.deg} & 19.5 & 16.5 & 16 \\
0.5 \mu\text{K.deg} & 20 & 16.5 & 16.5 \\
0.7 \mu\text{K.deg} & 20 & 17 & 16.5 \\
1.1 \mu\text{K.deg} & 20 & 18.5 & 18 \\
\end{array}
\]

3.3 Wiener filters

Now we consider Wiener filtering, which is an optimal way of recovering the signal of a data map \(D = S + N\) in case both the underlying signal \(S\) and the noise \(N\) are Gaussian. The Wiener filter is written, in harmonic space, as

\[
W_M = \frac{C_\ell}{C_\ell + N_\ell},
\]

where \(C_\ell\) is the forecasted power spectrum of the signal we want to analyse and \(N_\ell\) is the known power spectrum of the noise. Here, signal will refer to the dominant, Gaussian, cosmological part of the CMB. Obviously, this is sub-optimal, since we are after the non-Gaussian part of the CMB, while all the rest should be considered as noise. However, using a Wiener filter designed for extracting the non-Gaussian signal would require a stronger prior, making the measurement of \(f_{\text{nl}}\) potentially more accurate but only for a restricted class of NGs.

Similarly, one can define in harmonic space a ‘derivative’ Wiener operator,

\[
W_{D1} = \sqrt{\ell(\ell + 1)} \frac{C_\ell}{C_\ell + N_\ell},
\]

and a ‘second derivative’ Wiener filter optimal for recovering the Laplacian of the signal, \(\Delta S\),

\[
W_{D2} = \ell(\ell + 1) \frac{C_\ell}{C_\ell + N_\ell}.
\]

In practice, we use a smoothed version of the quantity \(C_\ell/(C_\ell + N_\ell)\), a ‘Wiener-like’ function used for component separation method (CMB removal) in Planck HFI Core Team et al. (2011). Before applying the Wiener filters we also correct the map for the Gaussian beam corresponding to the channel configuration. The three resulting filters, \(W_M, W_{D1}\) and \(W_{D2}\), are represented in Fig. 2.

Table 5. Sensitivity of the estimator of \(f_{\text{nl}}\) versus Gaussian smoothing scale \(\theta^S_{\text{FWHM}}\) when the combination of all functionals, \(V_0 + V_1 + V_2 + V_3\), is used. The forecasted quantity \((\tilde{\chi}^2_{\text{nl}})^{1/2}\) is given for the combined 100, 143 and 217 GHz channels in the extended Planck mission (see Appendix D).

| \(\theta^S_{\text{FWHM}}\) | \((\tilde{\chi}^2_{\text{nl}})^{1/2}\) |
|-----------------|-----------------|
| 5'              | 16.5            |
| 10'             | 20              |
| 20'             | 26.5            |
| 40'             | 34              |
| 5' + 10'        | 14              |
| 5' + 10' + 20'  | 13.5            |
| 5' + 10' + 20' + 40' | 13.5        |
A. Ducout et al.

Table 6. Sensitivity $\langle \Delta f_{NL} \rangle ^{1/2}$ when Wiener filters are applied to the data maps prior to MF measurements. It is calculated in the framework of the extended mission for the three channels.

| Functional | $W_M$ | $W_{D1}$ | $W_{D2}$ | $W_M + W_{D1}$ | $W_M + W_{D2}$ | $W_M + W_{D1} + W_{D2}$ |
|------------|-------|----------|----------|----------------|----------------|-------------------------|
| $V_0$      | 51    | 54       | 58       | 33             | 35             | 33                      |
| $V_1$      | 14    | 53       | 56       | 9              | 13             | 8.3                     |
| $V_2$      | 21    | 55       | 57       | 16             | 18             | 16                      |
| $V_3$      | 20    | 54       | 57       | 16             | 19             | 15.6                    |
| $V_1 + V_2$| 13    | 52       | 55       | 8.6            | 12.5           | 8                       |
| $V_1 + V_2 + V_3$ | 12.5 | 50       | 54       | 8              | 12             | 7.6                     |
| $V_0 + V_1 + V_2 + V_3$ | 12.4 | 49       | 52       | 7.6            | 12             | 7                       |

Table 6 summarizes the results of our analyses using them. It shows, not surprisingly, that $W_M$ alone does much better than Gaussian smoothing (Table 5) and a significant improvement is obtained when combining $W_M$ with $W_{D1}$ resulting in an overall reduction of the error on $f_{NL}$ of 30 per cent compared to the best results obtained with Gaussian smoothing. On the other hand, $W_{D2}$ does not significantly improve the result but this is somewhat expectable: for a stationary and isotropic random field, there is no correlation between the field and its first derivatives, but there is a strong correlation between the field and its second derivatives.

4 INHOMOGENEOUS NOISE

Due to the scanning strategy and the orientations of the different horns and bolometers (Dupac & Tauber 2005), the distribution of the noise in the raw sky maps is correlated, anisotropic and inhomogeneous. In what follows, we treat effects of inhomogeneity and anisotropy of the noise by relying on hitmaps (generated with the software MADAM of Keihänen, Kurki-Suonio & Poutanen 2005), in which each pixel value represents the number of times the pixel has been observed by the satellite. Our modelling of the noise in each pixel of the map $i$ then can be written as

$$\text{noise}(i) = \sigma_{\text{isotropic noise}} \times \mathcal{N}(0, 1) \times \sqrt{\frac{\text{hitmap}}{\text{hitmap}(i)}}. \quad (23)$$

Note thus that since correlations are neglected, anisotropy of the noise is modelled only partly, through the anisotropy of the hitmap. Even more realistic analyses would take into account the effects of correlations in the noise, which will be addressed in future work. Figs 3 and 4 show a hitmap and the corresponding noise map for the 143 GHz channel in a simulation of Planck signal (which used the characteristics of the instrument described in Planck HFI Core Team et al. 2011).

To analyse the impact of various levels of realism in modelling of the noise on the determination of $f_{NL}$, we consider two configurations: one where inhomogeneous noise is included in all the steps of the $\chi^2$ analysis, and one where only each of the $m_{\text{rand}} = 200$ test maps has its own realization of inhomogeneous noise – relying on the same hitmap but with different random seeds. We test here the null hypothesis $f_{NL} = 0$ but we checked that the conclusions do not change significantly for other values of $f_{NL}$. All the simulations are performed as detailed in Appendix D for full sky surveys with no foregrounds, $N_{\text{side}} = 2048$, and a Gaussian smoothing with $\theta_{N_{\text{side}}} = 5 \text{arcmin}$ to filter out the noise, which corresponds to a rather (almost the most) pessimistic case in terms of inhomogeneous noise.

The results of our analyses are summarized in Tables 7 and 8 for four levels of noise which are likely to bracket the actual sensitivity of Planck. To understand the results displayed in the tables, Figs 5 and 6 compare the effect of neglecting inhomogeneous noise to the presence of a ‘true’ $f_{NL}$ on the functionals for the nominal mission in the 143 GHz channel.

The Area functional, $V_0$, seems fairly insensitive to the effect of the inhomogeneity of the noise, which in turn makes the determination of $f_{NL}$ from the Area quite robust to that respect. This is not very surprising: the area is proportional to the cumulated one point distribution function (PDF). The presence of inhomogeneous noise locally induces a convolution of this PDF with a Gaussian of varying width depending on the value of the number of hits in the map. The effect of this convolution is negligible when the rms

2 The one obtained in the 143 GHz channel, to be specific.
Effect of neglecting the presence of inhomogeneous noise when estimating $f_{NL}$ and the error bar $\Delta f_{NL}$. Three settings are considered, as detailed below. For the reference maps and the test maps, we introduce inhomogeneous noise. These simple statements are confirmed by the examination of Fig. 5 shows that other functionals are rather sensitive to the effect of inhomogeneous noise, which is also expected. Indeed, we can guess that the presence of inhomogeneous noise augments the contrasts between cold and hot spots compared to the homogeneous case, hence building up a signal in $V_1$, $V_2$ and $V_3$. However, this signal is also contained in the parameter $A_1$ in equation (4) which tends to compensate for the effect on $V_1$. Hence, the normalized functionals, $v_k$, appear to be less affected by inhomogeneous noise than the ‘true’ functionals, $V_k$, as illustrated by Fig. 6. There is still some rather significant residual signal roughly proportional to the leading, Gaussian order of $v_k$, at least for the small smoothing scale considered here, as if the normalization factor $A_k$ was still slightly away from its actual value. This non-trivial effect might be due to deviation from homogeneity and isotropy of the field, induced by the noise. However, concerning $f_{NL}$ estimates, the parity of the black curves in the different panels of Fig. 6 is opposite to that of the curves corresponding to a true primordial $f_{NL}$ (green and red curves): we do not expect in that case the presence of inhomogeneous noise to introduce any bias on the measurement of $f_{NL}$. These simple statements are confirmed by the examination of Table 7. On the other hand, the presence of inhomogeneous noise makes the uncertainty on $f_{NL}$ slightly larger, increasing with the average level of noise, as shown in Table 8. Fortunately, for the three combined channels in extended Planck mission, the effects of the inhomogeneity of the noise become nearly negligible and can be fairly ignored, as we shall do from now on.

Table 7. Effect of neglecting the presence of inhomogeneous noise when estimating $f_{NL}$ and the error bar $\Delta f_{NL}$. Three settings are considered, as detailed below. The table is otherwise similar to Table 7.

| Noise ($\mu$K deg) | 1.1 | 0.7 | 0.5 | 0.33 |
|-------------------|-----|-----|-----|-----|
| Functional | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ |
| $V_0$ | 0.09 | 38.6 | 1 | 38.6 | 0.9 | 38.6 | 0.9 | 38.6 |
| $V_1$ | −9 | 24 | −4 | 23.6 | −19 | 22.9 | −0.8 | 21.5 |
| $V_2$ | −2 | 23 | −0.4 | 22 | 0.2 | 22 | −0.1 | 22 |
| $V_3$ | −1.9 | 25 | −0.2 | 24.7 | 0.5 | 24.6 | −0.3 | 24.5 |
| $V_0 + V_1 + V_2 + V_3$ | −4.4 | 18.8 | −3.8 | 17.9 | −2.3 | 17 | −0.8 | 16.7 |

$\sigma_{\text{noise}}$ of the local noise is small compared to the rms $\sigma_0$ of the field under consideration. \(^3\) Here this is the case: the additional Gaussian smoothing with $\theta_{\text{FWHM}}^N = 5$ arcmin reduces the typical local rms of the smoothed noise map to $\sigma_{\text{noise}} \lesssim 4 \times 10^{-6}$ whatever the channel considered, to be compared to $\sigma_0 \sim 4 \times 10^{-5}$.

The examination of Fig. 5 shows that other functionals are rather sensitive to the effect of inhomogeneous noise, which is also expected. Indeed, we can guess that the presence of inhomogeneous noise augments the contrasts between cold and hot spots compared to the homogeneous case, hence building up a signal in $V_1$, $V_2$ and $V_3$. However, this signal is also contained in the parameter $A_1$ in equation (4) which tends to compensate for the effect on $V_1$. Hence, the normalized functionals, $v_k$, appear to be less affected by inhomogeneous noise than the ‘true’ functionals, $V_k$, as illustrated by Fig. 6. There is still some rather significant residual signal roughly proportional to the leading, Gaussian order of $v_k$, at least for the small smoothing scale considered here, as if the normalization factor $A_k$ was still slightly away from its actual value. This non-trivial effect might be due to deviation from homogeneity and isotropy of the field, induced by the noise. However, concerning $f_{NL}$ estimates, the parity of the black curves in the different panels of Fig. 6 is opposite to that of the curves corresponding to a true primordial $f_{NL}$ (green and red curves): we do not expect in that case the presence of inhomogeneous noise to introduce any bias on the measurement of $f_{NL}$. These simple statements are confirmed by the examination of Table 7. On the other hand, the presence of inhomogeneous noise makes the uncertainty on $f_{NL}$ slightly larger, increasing with the average level of noise, as shown in Table 8. Fortunately, for the three combined channels in extended Planck mission, the effects of the inhomogeneity of the noise become nearly negligible and can be fairly ignored, as we shall do from now on.

Table 8. Effect of inhomogeneous noise when estimating $f_{NL}$ and the error bar $\Delta f_{NL}$. Three settings are considered, as detailed below. The table is otherwise similar to Table 7.

| Noise ($\mu$K deg) | 1.1 | 0.7 | 0.5 | 0.33 |
|-------------------|-----|-----|-----|-----|
| Functional | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ | $(\hat{f}_{NL})$ | $(\Delta f_{NL}^2)^{1/2}$ |
| Configuration 1\(^a\) | 0.19 | 17.8 | −0.05 | 16.5 | 0.044 | 16.2 | 0.07 | 16 |
| Configuration 2\(^b\) | −4.4 | 18.8 | −3.8 | 17.9 | −2.3 | 17 | −0.8 | 16.7 |
| Configuration 3\(^c\) | 0.002 | 25.5 | −0.44 | 23 | −0.15 | 19.9 | −0.15 | 17 |

\(^a\)For the reference maps and the test maps, we introduce isotropic noise.
\(^b\)For the reference maps, we introduce isotropic noise and for test maps we introduce inhomogeneous noise.
\(^c\)For the reference maps and the test maps, we introduce inhomogeneous noise.

5 FOREGROUNDS I: POINT SOURCES

‘Point sources’ refer to the (large) number of radio and infrared (IR) galaxies that are detectable at the CMB frequencies. These galaxies are in general not all resolved by CMB all-sky experiments like WMAP or Planck, even if the brightest objects are detected individually. The faint ones contribute to an inhomogeneous sky background. In this section, we test for the first time the effect of these sources on the estimate of $f_{NL}$ with MFs. In previous studies relying on MF analyses (e.g. Hikage et al. 2008; Natoli et al. 2010), the effect of point sources was indeed supposed to be completely subtracted off by an appropriate masking or simply negligible compared to error bars.

This section is divided into two parts: Section 5.1 describes in details the simulation pipeline we used to perform our analyses while Section 5.2 discusses the results of our investigations.

5.1 Method

To perform our simulations of the data, we need to compute accurately the contribution of point sources. To do so, we use the PSM code described in Delabrouille et al. (2012).\(^7\) As reviewed in Section 5.1.1, the point source contribution depends significantly on frequency. As a result, our simulations and analyses will separately consider the three cosmological channels, 100, 143 and 217 GHz, in the extended mission configuration for the noise level (see Appendix D). In particular each channel will have different masking treatment for the brightest point sources. Since masks represent a crucial part of the treatment of point sources, we discuss about them in Section 5.1.2. Other technical details about our simulations are provided in Section 5.1.3.

\(^3\) Note that this argument would be valid as well for a non-Gaussian noise.

\(^4\) http://www.apc.univ-paris7.fr/~delabrou/PSM/psm.html
Figure 5. Effect of inhomogeneous noise on the ‘raw’ functionals $V_k$ for the nominal mission in the 143 GHz channel and a Gaussian smoothing with $\theta_{5000}^2 = 5$ arcmin. The relative difference between the $V_k$ measured in different types of maps and the Gaussian limit $V_k^G$ is displayed as a function of $\nu$. Each panel corresponds to an individual functional. The curves represented on each panel are calculated by the average over 200 realizations, with $f_{NL} = 0$ and inhomogeneous noise for the black thick curve and with primordial NG ($f_{NL} \neq 0$) and homogeneous noise for the green and red curves, while $V_k^G$ was computed alike with homogeneous noise. The grey shade around the thick black curve represents the uncertainty induced by the fact that $m_{\text{test}} = 200$.

5.1.1 Point sources simulations

The PSM code (Delabrouille et al. 2012) is specifically designed to simulate all relevant sky emissions at Planck frequencies, including secondaries and foreground emission, as they were known before the launch of Planck. In this paragraph we describe only the part of the PSM that deals with point sources, the rest of the simulation pipeline being detailed in Section 5.1.3.

First, we use the PSM code to add radio sources to our simulations: these are mostly active galactic nuclei (AGN) observed via their synchrotron emission. The PSM relies on numerous surveys of those radio sources at frequencies ranging from 0.85 to 4.85 GHz to model this emission. In regions not observed by surveys or with shallower observations, sources are copied from other regions, until a coverage down to at least 20 mJy at 5 GHz is achieved over the full sky. Then flux densities are extrapolated at all frequencies by using a power-law approximation for the spectra, of the form $S_\nu \propto \nu^{-\alpha}$. For the spectral index $\alpha$ estimates, sources are classified into steep or flat spectrum class and $\alpha$ is drawn from a Gaussian distribution with mean and variance corresponding to its class (Ricci et al. 2006) and matching WMAP data (Bennett et al. 2003) in several frequency ranges. Besides, WMAP sources are accounted for in the simulations. Source counts at 5 and 20 GHz are found to be consistent with the model of Toffolati et al. (1998) and an updated version of the model of de Zotti et al. (2005). These sources are known to contribute essentially at low frequencies, from 30 to 90 GHz but they have been detected at higher frequencies up to 217 GHz (Planck Collaboration et al. 2011a,b).

We note that radio sources are nearly Poisson distributed on the sky, so they essentially contribute to the CMB as a shot noise.

Secondly, we add IR sources to the simulations. Indeed, at high frequencies (above 150 GHz), a thermal emission arising from dust heated by the UV emission of young stars also contributes. In addition to normal stars surrounded by a disc, numerous starburst galaxies which form stars at extreme rates contribute to this thermal emission. In the PSM, sources are taken from the IRAS point source catalogue (PSC) (Beichman et al. 1988) and the faint source catalogue (FSC) (Moshir, Kopman & Conrow 1992). The flux densities are extrapolated to Planck frequencies by adopting a model with modified blackbody spectra and the gaps in sky coverage are filled up using the same procedure as for the radio sources.
Finally, we add to the simulations the CIB which is possibly the dominating component. Distant starburst galaxies are not all detected individually and the cumulated emission from the fainter ones form a diffuse background of anisotropies. The PSM adopts the count model of Lapi et al. (2006), which is consistent with Submillimetre Common-User Bolometer Array (SCUBA) and Max Planck Millimeter Bolometer (MAMBO) surveys. Sources are clustered following model 2 in Negrello et al. (2004) and the spatial distribution follows the method of González-Nuevo, Toffolatti & Argüeso (2005), then flux densities are extrapolated to all frequencies. The CIB power spectra of the simulation agree sufficiently well with those measured by Planck (Planck Collaboration et al. 2011c) and Atacama Cosmology Telescope (ACT; Dunkley et al. 2011) for our forecast studies.

The main difference between the distribution of radio and IR sources on the sky is that IR sources, being observed either as individual entities or as a background, are clustered in their host dark-matter haloes. So their power spectrum is not flat as for the radio sources.

5.1.2 Masks

Brightest point sources can be detected individually and can be masked properly when their flux density is beyond a chosen detection threshold when compared to the level $\sigma = \sigma_{\text{noise}}$ of the underlying noise in the CMB data. Here, we create three sets of point source masks corresponding to three flux density cuts (referred simply as flux cuts). The choice of these different flux cuts has been mainly derived from the Early Release Compact Source Catalogue (ERCSC) published by the Planck Collaboration (Planck Collaboration et al. 2011b). The first set of masks refers to sources beyond a $10\sigma$ level in the three bands 100, 143 and 217 GHz, corresponding to respective flux density thresholds at 0.5, 0.33, 0.28 Jy as chosen in the ERCSC. The second set of masks concerns sources beyond the $5\sigma$ level, which corresponds to the threshold choice in the cleanest parts of the sky of the ERCSC. The third set of masks corresponds to the $3\sigma$ level, which is not mentioned in the ERCSC, but we use it because we believe it represents a more appropriate set of masks for cosmological purposes. Indeed, in the ERCSC, the goal was to set flux density thresholds to have a sufficiently good signal-to-noise ratio for reliable analysis of the point sources properties. Our goal here is to remove the contribution from the point sources, which requires a much less stringent criterion on the quality of their detection. Furthermore, the ERCSC signal-to-noise ratio level does not match that of the nominal mission and by no mean that of the extended mission.

Each mask associated with an individual point source is a disc of radius three times the FWHM of the beam of the instrument in...
the channel under consideration. When adding up the contributions of all the sources, a certain fraction $1 - f_{\text{sky}}$ of the sky is masked, as illustrated by Fig. 7. The value of the sky fraction which is then used, $f_{\text{sky}}$, ranges from $f_{\text{sky}} = 0.90$ for the 100 GHz channel up to $f_{\text{sky}} = 0.99$ for the 217 GHz channel. These differences come from two factors on which the construction of masks depends: the beam width and the number of point sources detected beyond the threshold of interest. These two parameters decrease when passing from 100 to 217 GHz.

5.1.3 Simulation pipeline

To test the effects of point sources in the estimation of $f_{\text{NL}}$ or more specifically the approximation of neglecting their presence, we add their contribution only to the ‘data’ (test) maps, $\hat{y}$ in equation (8). We simulate $m_{\text{test}} = 200$ of these tests maps with $f_{\text{NL}}^{\text{prim}} = \{-10, 0, 10, 50\}$, where $f_{\text{NL}}^{\text{prim}}$ stands for the ‘primordial’ $f_{\text{NL}}$ (to be contrasted later with other contributions to the effective $f_{\text{NL}}$ arising from biases induced by unaccounted point sources). The $m = 10,000$ Gaussian maps used to compute the covariance matrix $C$ as well as the $m_{\text{NG}} = 200$ non-Gaussian maps used to calculate the model prediction $\tilde{y}(f_{\text{NL}})$ in equation (8) neglect this contribution, but have exactly the same treatment otherwise, including sky coverage and instrumental noise as detailed below. This way, our analysis will be able to confirm if appropriate masking is enough to render the effects of point sources negligible on the measurement of $f_{\text{NL}}$.

The details of our simulation pipeline now follow. CMB maps are created first with the beam corresponding to each channel frequency $f$ (see Appendix D), and are supplemented with point sources only for the test maps convolved with the same beam and with the noise corresponding to each channel $f$ for the extended mission. Next, point source masks are applied to the maps. These masks depend on the channel $f$ so the beam width $\theta_{\text{FWHM}}^b(f)$ is a parameter – and on the chosen flux cut $\beta$. The punched holes are filled by diffusive inpainting. Then, the maps are smoothed with a Gaussian window of size $\theta_{\text{FWHM}}^b = 5$ arcmin. Finally, a Galactic mask is applied to the maps, here corresponding to a valid fraction of the sky $f_{\text{sky}} = 0.80$. The procedure to construct the Galactic mask will be described in Section 6. The complete pipeline is summarized as follows:

$$\text{map} = \text{CMB} \left( f_{\text{NL}} \right) * \text{beam} \left[ \theta_{\text{FWHM}}^b(f) \right]$$

$$+ \text{foreground of sources} * \text{beam} \left[ \theta_{\text{FWHM}}^b(f) \right]$$

$$+ \text{noise} (f) + \text{point sources mask} (\beta, f) \text{ inpainted}$$

$$+ \text{smoothing} \left( \theta_{\text{FWHM}}^b = 5 \text{ arcmin} \right)$$

$$+ \text{Galactic mask} (f_{\text{sky}}).$$

(24)

We checked that the results derived in this section are qualitatively the same for other values of $\theta_{\text{FWHM}}^b$ and for the Wiener filters studied in Section 3.3. Of course, a quantitative calculation of the biases induced by point sources will require a new analyse each time a new filter is considered.

5.2 Results

Table 9 shows the estimate on $f_{\text{NL}}$ obtained for different channels as a function of source masking level and primordial NG, $f_{\text{NL}}^{\text{prim}}$. Again, a frequentist average of the posterior averages is performed over 200 test maps realizations, and is noted $\langle f_{\text{NL}} \rangle$. Note that while point sources can introduce a significant bias on the estimate of $f_{\text{NL}}^{\text{prim}}$, they do not change significantly the error bars, $\Delta f_{\text{NL}}$, that depend linearly on the inverse square root of sky coverage given the overall level of noise (Table 14). Therefore, error bars on the measured $f_{\text{NL}}$ are not mentioned further in this section, for simplicity. We now discuss the results obtained in Table 9, starting with the 100 and 143 GHz channels, dominated by radio sources (Section 5.2.1) and finishing with the 217 GHz channel, where one has to account for the additional IR source contribution (Section 5.2.2).

5.2.1 100 and 143 GHz: effect of radio sources

In the first two bands, 100 and 143 GHz, faint point sources are composed mainly of radio sources, as can be seen in the ERCSC. Radio sources are not clustered: their power spectrum is known to be flat so they act as a positive, uncorrelated noise. To understand the effect of such a noise, we study in details the configuration of a minimal mask in the 143 GHz channel as illustrated for each functional by Fig. 8.

As a positive uncorrelated noise, radio sources do not significantly affect the Area MF. Indeed, the positive nature of such a noise is subtracted out when computing the density contrast $\delta$ of the map and, as discussed in Section 4, the convolution effect of a zero average distribution on the Area is negligible as long as its variance is small compared to the variance of the signal.

The presence of point sources brings an excess of positive clusters; it also slightly shifts the curve $V(v)$ to smaller values of $v$ as an effect of subtracting the average from the temperature map when computing the density contrast $\delta$. As can be easily noticed on the Perimeter panel of Fig. 8. These effects induce an overall positive bias for the Genus and the Perimeter, while $N_{\text{clusters}}$ presents an excess for positive thresholds but a deficit (of under-dense regions) for negative thresholds. These biases remain after renormalization, i.e. when passing from $V_1$ to $v_1$, except for the perimeter, where division by the factor $A_1$ inverted the bias. Indeed, the presence of point sources increases the ratio $\sigma_1/\sigma_0$, hence the measured value of $A_1$, $k > 0$ (see Appendix A).

Table 10 shows the corresponding bias on the measurement of $f_{\text{NL}}$ in the null hypothesis $f_{\text{NL}}^{\text{prim}} = 0$. The results of this table can partly be inferred intuitively from the examination of Fig. 8: small bias for the Area, positive bias for the Genus and $N_{\text{clusters}}$, negative bias for the Perimeter, resulting in an overall positive bias for the...
combination of all functionals. The examination of other masking levels confirms the results of this analysis: the effect of point sources in the 100 and 143 GHz is a positive bias on the measured \( f_{\text{NL}} \), which, in addition, does not depend on the primordial level of NG,

\[
\hat{f}_{\text{NL}} = f_{\text{NL}} + f_{\text{bias}},
\]

and decreases, as expected, when more point sources are excluded by the masks, as illustrated by Table 11. In particular, the bias induced by point sources becomes nearly negligible compared to expected error bars (\( \Delta f_{\text{NL}} \lesssim 10 \), Table 14) when masks are set up at the 3σ level.

5.2.2 217 GHz: effect of radio and IR sources

In the 217 GHz band, in addition to radio sources, an IR background contributes to the faint point sources, which results in new biases on the measurement of \( f_{\text{NL}} \), as Table 9 shows.

The contribution from unclustered radio sources is a decreasing function of frequency and masking. It should act, as in the 100 and 143 GHz channels, as a positive bias on the measurement of \( f_{\text{NL}} \) that does not depend on the value of \( f_{\text{NL}} \) (see equation 25).

On the other hand, IR sources are clustered and form mainly a diffuse, unresolved background, which cannot be dealt with masks. They induce a bias on the measurement of \( f_{\text{NL}} \) which appears to depend on the value of \( f_{\text{NL}} \) as can be inferred from Table 9. To analyse this bias more in depth, we concentrate on a configuration where the contribution of radio sources is masked out as much as possible and nearly negligible, with the 3σ flux cut setting for the masks. Fig. 9 displays the functionals obtained in two cases, \( f_{\text{NL}} = 0 \) and \( f_{\text{NL}} = 50 \). It is interesting to notice that the curve obtained for \( f_{\text{NL}} = 50 \) is nearly exactly the sum of the curve for \( f_{\text{NL}} = 50 \) with no point source contribution and the curve for \( f_{\text{NL}} = 0 \) with point sources. Unfortunately, this quite general\(^6\) linear property does not translate in a simple way in terms of bias on the measured \( f_{\text{NL}} \), when performing the \( \chi^2 \) analysis, as illustrated by Tables 12 and 13. What we find, instead, is the bias due to IR sources to be roughly proportional to \( f_{\text{NL}} \), when one considers the combination of all functionals to perform the measurements.

Our final approximation for the total bias in the 217 GHz channel is therefore

\[
\hat{f}_{\text{NL}} @217\text{GHz} = f_{\text{NL}} + f_{\text{bias}} + f_{\text{bias, radio}} + f_{\text{bias, IR}}
\]

(26)

with \( f_{\text{bias, radio}} \) depending only of the masking level, as in Section 5.2.1. Here, this bias grows from negligible for the 3σ masks to \( f_{\text{bias, radio}} \approx 5 \) for the 10σ masks (right-column part of Table 9 with \( f_{\text{NL}} = 0 \)). The other term, \( f_{\text{bias, IR}} \), does not depend on masking but is approximately proportional to \( f_{\text{NL}} \) for moderate values of \( f_{\text{NL}} \):

\[
|f_{\text{bias, IR}}| \lesssim \frac{f_{\text{NL}}}{2}, \quad |f_{\text{NL}}| \lesssim 50.
\]

(27)

Note that with Planck extended mission signal-to-noise ratio, one expects to be able to characterize more accurately the IR background. It might then be possible to account for it in a better way, by e.g. including it in the model itself instead of ignoring it due to the lack of precise knowledge.

6 FOREGROUNDS II: GALACTIC RESIDUALS AND GALACTIC MASK

Galactic signals are a major issue for cosmological studies of the CMB and have to be accurately assessed. It is usually treated in two ways: masking alone or in conjunction with component separation. Here, we test the effects of the Galactic residuals left from these methods on the estimation of \( f_{\text{NL}} \) with MFs by using simulations of the Foreground emission, masks constructed from these simulations for different sky coverages and a naive model of component separation quality.

The philosophy adopted here is similar to that in the previous section: we do not include in the estimator model the templates of Galactic foregrounds. Instead, we analyse the biases on \( f_{\text{NL}} \) induced by neglecting the presence of Galactic emission. Such biases are expected to be small if one restricts to regions far from the Galactic plane or/and if proper component separation has been performed prior to the measurements.

---

\(^6\) It also applies to radio sources (Section 5.2.1).

---

**Table 9.** Estimates of \( f_{\text{NL}} \) in the presence of point sources. To create the test maps, we used the procedure described in equation (24) with a noise at the level of the extended mission in each band. The analyses are performed for combined MFs, \( V_0 + V_1 + V_2 + V_3 \).

| Flux cut (detection level in the ERCSC) | \( f = 100 \text{ GHz} \) | \( f = 143 \text{ GHz} \) | \( f = 217 \text{ GHz} \) |
|----------------------------------------|-------------------------|-------------------------|-------------------------|
| \( f_{\text{NL}} = 0 \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) |
| 10σ | 14 | 10 | 5 |
| 5σ | 8 | 4 | 0.4 |
| 3σ | 3 | 1.4 | 0.6 |
| \( f_{\text{NL}} = -10 \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) |
| 10σ | 2 | −0.2 | −10 |
| 5σ | −1.5 | −6 | −15 |
| 3σ | −6.5 | −9 | −16 |
| \( f_{\text{NL}} = 10 \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) |
| 10σ | 25 | 21 | 21 |
| 5σ | 19 | 14 | 16 |
| 3σ | 13 | 12 | 15.5 |
| \( f_{\text{NL}} = 50 \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) | \( \langle \hat{f}_{\text{NL}} \rangle \) |
| 10σ | 64 | 61 | 72 |
| 5σ | 57 | 54 | 69 |
| 3σ | 52 | 51.7 | 69 |
Figure 8. Effect of the presence of point sources on the measurement of MFs in the 143 GHz channel. Here, only the brightest point sources, with an ERCSC signal-to-noise ratio larger than $\beta = 10\sigma$, are masked out. The configuration is the same as in Fig. 6: normalized functionals $v_k$ are plotted as functions of the threshold $\nu$, after subtracting the Gaussian limit prediction, $v_k^G$. The black thick curve with the grey shade corresponds to $f_{NL} = 0$ but with point sources, while the two other ones correspond to the expected curves in the presence of primordial $f_{NL}$ as shown inside each panel.

Table 10. Bias on the measurement of $f_{NL}$ introduced by point sources at 143 GHz in case of weak masking at the $10\sigma$ detection level in the ERCSC. Each column corresponds to using a specific functional in the $\chi^2$ analysis, or, for the last one, the combination of all functionals. Here, the null hypothesis $f_{NL}^{\text{prim}} = 0$ is tested, but in practice, the effective bias does not depend on the value of $f_{NL}^{\text{prim}}$ (equation 25).

| $f_{NL}^{\text{prim}} = 0$ | $V_0$ | $V_1$ | $V_2$ | $V_3$ | All |
|---------------------------|-------|-------|-------|-------|-----|
| $\langle \hat{f}_{NL} \rangle$ | 0.8   | -8    | 14    | 2.3   | 10  |

This section is divided into four parts. The first one, Section 6.1, details our simulation settings. The second one, Section 6.2, looks at the statistical uncertainty expected on the measured $f_{NL}$ from MFs as a function of sky coverage. The third one, Section 6.3, examines, as functions of sky coverage, the biases brought by Galactic foregrounds, if these latter would not be removed at all, and the expected improvements on such biases brought by component separation.

Table 11. Bias on the measurement of $f_{NL}$ introduced by point sources at 100 and 143 GHz as a function of masking level, expressed here in terms of ERCSC signal-to-noise ratio threshold, assuming that the combination of all functionals is used to measure $f_{NL}$. This bias, modelled by equation (25), does not depend on the actual level of primordial NG, $f_{NL}^{\text{prim}}$, and can therefore be easily corrected for. Here, as detailed in Section 5.1.3, the bias is obtained in the following configuration for the extended mission: $N_{\text{side}} = 1024$, $\ell_{\text{max}} = 2000$, Gaussian smoothing with $\theta_{\text{FWHM}} = 5$ arcmin, $n_{\text{bins}} = 26$ and $\nu_{\text{max}} = 3.5$. With a different set up, a new estimate of the bias would be needed, but this is an easy exercise.

| Flux cut | $f_{NL}^{\text{bias}}$ @ 100 GHz | $f_{NL}^{\text{bias}}$ @ 143 GHz |
|----------|-----------------------------|-----------------------------|
| $10\sigma$ | 14                         | 10                          |
| $5\sigma$  | 8                          | 4                           |
| $3\sigma$  | 3                          | 1.5                         |

6.1 Method

To perform our simulations, we need a way to generate realistic maps of the Galactic emission, which is reviewed in Section 6.1.1.
To do so, we use again the PSM (Delabrouille et al. 2012). Section 6.1.2 explains how masks are generated in order to exclude regions where Galactic emission is too strong. Section 6.1.3 details the very simple method we decided to use to assess the effect of component separation quality. Other technical details of our simulations are provided in Section 6.1.4.

6.1.1 Simulations of the foreground emission: the different components

We use again the PSM to construct the Galactic emission. The code uses template maps interpolated at the desired frequencies. The detailed modelling of each component is described in Delabrouille et al. (2012). We detail hereafter the list of components we simulated in our map of the Foreground emission.

The diffuse Galactic emission arises from the interstellar medium (ISM) in the Milky Way. The ISM is composed of different phases, from cold molecular clouds to hot ionized regions, of magnetic fields and cosmic rays. The intensity of the corresponding emission depends on Galactic latitude. It is stronger near the centre of the Galaxy and decreases at lower/higher latitudes following approximately a cosecant law $1/\sin(|b|)$. More precisely, we usually classify the different ISM components according to their physical emission processes:

(i) synchrotron radiation is emitted by relativistic cosmic rays spiralling in the Galactic magnetic field. Its intensity depends on the
cosmic ray density and on the magnetic field strength perpendicular to the line of sight;

(ii) free–free emission originates from the ionized medium in the ISM, as a result of the interaction of free electrons with positively charged nuclei. It comes principally from star-forming regions in the Galactic plane;

(iii) there is also a thermal emission coming from dust grains heated by stars, which is the dominant contribution above 70 GHz;

(iv) another type of emission has been found at microwave frequencies which is probably due to small spinning dust particles ('Anomalous emission');

(v) At these frequencies, there are also molecular lines emerging from molecular clouds, particularly those of $^{12}$CO at 100 and 217 GHz.

For the last contribution, i.e. CO lines, templates and models of the emission are mostly unknown at present time and we choose not to model it and to simulate only the 143 GHz part of the Galactic foreground. However, concerning noise level, our analyses correspond to a combination of the 100, 143 and 217 GHz channels in the framework of the Planck extended mission. The specific CO contribution, even if sub-dominant compared to the other physical processes, will be studied when Planck templates of CO will be available.

The resulting foreground map is represented in Fig. 10, with two colour scales enhancing different aspects of this emission.

### 6.1.2 Galactic mask

To create Galactic masks, we consider the fraction $f_{\text{sky}}$ of the sky we aim to keep, which sets up an intensity threshold for the Galactic foreground above which the corresponding region of the sky is masked out. Once the mask function $M$ is set up, which is equal to one for valid pixels and zero for excluded pixels, convolution of this function is performed with a Gaussian kernel of size $\theta_{\text{FWHM}}^S = 5^\circ$ to obtain a smoothed version $M_{\text{smoothed}}$. New masks with smooth boundaries are extracted from this map by selecting pixels with $M_{\text{smoothed}} \geq M_{\text{thresh}}$ and excluding the others, where the value of $M_{\text{thresh}}$ is tuned to get the correct value of $f_{\text{sky}}$. An example of masks constructed that way is given in Fig. 11.

### 6.1.3 Component separation quality

In order to assess simply the impact of residuals from the component separation, i.e. impact of its quality, we model component separation results by simply adding to the CMB the map of Galactic foregrounds multiplied by a scaling factor $\alpha$, where $\sigma_{\text{CMB@HL}}$ and $\sigma_{\text{FG@HL}}$ are, respectively, the standard deviations of the CMB and foreground maps at high latitudes, i.e. measured in pixels outside the mask with $f_{\text{sky}} = 0.80$. So for a 'quality factor' $\alpha = 1$, the normalized contributions of Galactic foreground and CMB at high latitudes will be the same. The value $\alpha = 0.77$ corresponds to the initial level of foreground emission, so it is equivalent to no component separation (top panel of Fig. 12). Realistic values of $\alpha$, when examining maps obtained from actual component separation methods, rather appear to range typically between $\alpha = 0.01$ (bottom panel of Fig. 12) and $\alpha = 0.05$.

Obviously, our modelling of Galactic residuals after component separation is very rough, but it should be sufficient to assess what should be the value of $f_{\text{sky}}$ for making the biases on $f_{\text{NL}}$ induced by these residuals negligible. A better modelling of the residuals would require detailed examination of the results obtained from actual component separation. Furthermore, a new analysis would be required each time a new component separation method is considered: this is far beyond the scope of this paper.
or Wiener filters discussed in Section 3.3. Finally, the Galactic mask calculated as in Section 6.1.2 is added, for a given $f_{\text{sky}}$. Note that at variance with the point source analysis, the Galactic mask is not inpainted, due to its rather large size: the pixels inside the mask are just ignored by the MFs code. The complete pipeline is summarized as follows:

$$\text{map} = \text{CMB}(f_{\text{prim}}) \ast \text{beam} \left[ \theta^b_{\text{FWHM}}(f) \right] + \alpha \times \frac{\sigma_{\text{CMB@HL}}}{\sigma_{\text{FG@HL}}} \times \text{Gal.FG}(f) \ast \text{beam} \left[ \theta^b_{\text{FWHM}}(f) \right] + \text{noise}(f) + \text{smoothing} \left[ \theta^S_{\text{FWHM}} \right] \text{ or Wiener filtering} + \text{Galactic mask}(f_{\text{sky}}).$$

(28)

6.2 Sensitivity versus sky coverage

First, we test the sensitivity of our $\chi^2$ estimator to sky coverage, using maps of the CMB without Galactic foreground, i.e. $\alpha = 0$ and just looking at the resulting error bars, $\Delta f_{\text{NL}}$. Table 14 shows $\langle \Delta f_{\text{NL}}^2 \rangle^{1/2}$ as a function of $f_{\text{sky}}$, obtained from the combination of all four functionals for various Gaussian smoothing scales $\theta^S_{\text{FWHM}}$ and different Wiener filterings. As expected, $\langle \Delta f_{\text{NL}}^2 \rangle^{1/2}$ is approximately a linear function of $1/\sqrt{f_{\text{sky}}}$. Note that these results do not change significantly in the presence of Galactic foregrounds as long as $f_{\text{sky}} \lesssim 0.80$.

6.3 Effect of Galactic foreground

In the following, we study the actual effects of the Galactic foregrounds, which present a rather complex behaviour as a function of sky coverage, $f_{\text{sky}}$. Our analyses restrict to Gaussian smoothing and put aside Wiener filtering. From the qualitative point of view, the latter indeed leads to results similar to what is obtained with the smallest Gaussian smoothing scale. This section is divided into two parts. Section 6.3.1 assumes no component separation and analyses in detail the biases brought on the measurement of $f_{\text{NL}}$ by the Galactic foregrounds, while Section 6.3.3 examines the biases as functions of component separation quality factor, $\alpha$.

6.3.1 Two behaviours, two components

We now examine what kind of biases the Galactic foregrounds induce on $f_{\text{NL}}$. Table 15 shows our forecast for $\langle f_{\text{NL}} \rangle$ as a function of $f_{\text{sky}}$ and $f_{\text{prim}}$, at the smallest Gaussian smoothing scale, $\theta^S_{\text{FWHM}} = 5$ arcmin. Remember that Galactic foregrounds are completely present, without any attempt to remove them with component separation.

Numerous interesting results can be extracted from Table 15.

(a) First, close to the Galactic plane, so for $f_{\text{sky}}$ close to unity, a very important signal, that we denote by $f_{\text{Gal-plane}}$, overrides the primordial one. Indeed, when $f_{\text{sky}} = 0.90$, $(\tilde{f}_{\text{NL}})$ does not depend much on the value of $f_{\text{prim}}$ and can be approximated as follows:

$$\langle f_{\text{NL}} \rangle = f_{\text{Gal-plane}} = f_{\text{Gal-plane, const}} + \frac{f_{\text{prim}}}{10}, \quad f_{\text{sky}} = 0.9 \quad (29)$$

with $f_{\text{Gal-plane, const}} \simeq -22$ for the $V_1 + V_2 + V_3$ combination and $f_{\text{Gal-plane, const}} \simeq -26$ for the $V_0 + V_1 + V_2 + V_3$ combination.

(b) At higher latitudes, a second type of signal less powerful than $f_{\text{Gal-plane}}$ appears. This new contribution, denoted by $f_{\text{Gal-high}}(f_{\text{sky}})$,
Table 15. Effect of Galactic foregrounds on the measurement of $f_{NL}$ as a function of sky coverage $f_{sky}$ and level of primordial NG $f_{NL}^{prim}$. The results displayed in this table assume Gaussian smoothing with $\theta_{FWHM}^S = 5$ arcmin in the simulations used to perform the calculations (Section 6.1.4).

| $\alpha = 0.77$ | $f_{NL}^{prim} = 0$ | $f_{NL}^{prim} = -10$ | $f_{NL}^{prim} = 10$ | $f_{NL}^{prim} = 50$ |
|---------------|----------------|-----------------|----------------|----------------|
| $f_{sky} = 0.90$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ |
| $f_{NL}^{Gal}$ for $V_1 + V_2 + V_3$ | $-22$ | $-23$ | $-21$ | $-15$ |
| $f_{NL}^{Gal}$ for $V_0 + V_1 + V_2 + V_3$ | $-26$ | $-27$ | $-25$ | $-21$ |
| $f_{sky} = 0.80$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ |
| $V_0$ | $-61$ | $-63$ | $-58$ | $-48$ |
| $V_1$ | $-42$ | $-52$ | $-32$ | $14$ |
| $V_2$ | $-12$ | $-22$ | $-2$ | $34$ |
| $V_3$ | $-58$ | $-64$ | $-51$ | $19$ |
| $V_1 + V_2 + V_3$ | $0.8$ | $-12$ | $10.6$ | $53$ |
| $V_0 + V_1 + V_2 + V_3$ | $-5$ | $-16.2$ | $6.5$ | $50$ |
| $f_{sky} = 0.70$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ |
| $V_0$ | $-15$ | $-20$ | $-12$ | $4.5$ |
| $V_1$ | $-5$ | $-15$ | $5$ | $45$ |
| $V_2$ | $8$ | $-1.3$ | $17$ | $51$ |
| $V_3$ | $-22$ | $-30$ | $-13$ | $23$ |
| $V_1 + V_2 + V_3$ | $17$ | $6$ | $27$ | $63$ |
| $V_0 + V_1 + V_2 + V_3$ | $16$ | $5$ | $26$ | $63$ |
| $f_{sky} = 0.60$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ |
| $V_1 + V_2 + V_3$ | $13$ | $4$ | $23$ | $60$ |
| $V_0 + V_1 + V_2 + V_3$ | $12$ | $3$ | $22$ | $60$ |
| $f_{sky} = 0.40$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ | $\langle \hat{f}_{NL} \rangle$ |
| $V_1 + V_2 + V_3$ | $9$ | $-0.1$ | $18.5$ | $56$ |
| $V_0 + V_1 + V_2 + V_3$ | $9$ | $-0.16$ | $18.5$ | $56$ |

Table 16. Galactic bias at high latitudes, $f_{NL}^{Gal-high}$, as a function of sky coverage, $f_{sky}$, as explained in the main text. The results displayed in this table assume Gaussian smoothing with $\theta_{FWHM}^S = 5$ arcmin in the simulations used to perform the calculations (Section 6.1.4).

| $f_{sky}$ | 0.80 | 0.70 | 0.60 | 0.40 |
|-----------|------------|------------|------------|------------|
| $f_{NL}^{Gal-high}$ for $V_1 + V_2 + V_3$ | 22 | 17 | 13 | 9 |
| $f_{NL}^{Gal-high}$ for $V_0 + V_1 + V_2 + V_3$ | 21 | 16 | 12 | 9 |

Table 17. Galactic foreground bias as a function of smoothing scale $\theta_{FWHM}^S$ for various sky coverages, $f_{sky}$, in the null hypothesis $f_{NL}^{prim} = 0$. The measured quantity $\langle \hat{f}_{NL} \rangle$ is given for the combination $V_0 + V_1 + V_2 + V_3$. Because our $\chi^2$ estimates are made unreliable at large smoothing scales when sky coverage is small, there are three values missing in the table.

| $\theta_{FWHM}^S$ | $f_{sky} = 0.80$ | $f_{sky} = 0.70$ | $f_{sky} = 0.60$ | $f_{sky} = 0.40$ |
|------------------|-----------------|-----------------|-----------------|-----------------|
| 5 arcmin | $-5$ | 16 | 12 | 9 |
| 10 arcmin | 12 | 25 | 17 | $-$ |
| 20 arcmin | 32 | 39 | 22 | $-$ |
| 40 arcmin | 37 | 36 | 25 | $-$ |

6.3.2 Smoothing

To understand more deeply the effects of the two types of Galactic foregrounds we found above, we examine what happens when the smoothing scale is varied. Table 17 gives $\langle \hat{f}_{NL} \rangle$ as a function of $\theta_{FWHM}^S$ for various sky coverages. Obviously, one can infer from this table that the positive bias increases with smoothing scale, suggesting that the Galactic signal at high latitude dominates at large scales, while the Galactic plane signal is present with its negative bias only at small scales. This result, in addition to the analyses performed in previous paragraphs, show that MFs remain helpful in understanding and isolating the different biases induced by Galactic foregrounds, similarly as for the bi-spectrum. This demonstrates again the discriminative power of MFs.
most of the information on NG is contained at the smallest possible scales, of the order of the beam size.

(vi) Point source foregrounds introduce a bias on the measurement of $f_{NL}$ that can be estimated and thus corrected for accurately.\(^7\)\(^8\) Note that with appropriate masking of the brightest sources followed by a simple in-painting procedure, this bias becomes negligible except at the 217 GHz frequency.

(vii) Galactic foregrounds introduce a complex bias on the measurement of $f_{NL}$ that depends on the fraction of sky covered, $f_{sky}$, or in other words, depends on how much the most luminous part of the Galaxy has been masked out. This bias can be again corrected for and is coincidentally negligible when the combination of Perimeter, Genus and number of clusters is used and $f_{sky} = 0.8$. With appropriate component separation the bias due to Galactic foregrounds should become negligible compared to the error bars on $f_{NL}$, at least for $f_{sky} < 0.8$.

(viii) With all the effects described above under control, we expect to be able to measure $f_{NL}$ using MFs with an error of the order of $\Delta f_{NL} \lesssim 0.1$.

These are excellent news overall, since it suggests that MFs should be capable of putting interesting constraints on $f_{NL}$, even in view of the performance of fully optimal estimators for that purpose. It is reasonable to assume that they should also provide non-trivial constraints on other types of primordial NG even if there are no prediction yet at the time of the analysis which would allow building dedicated and somewhat more stringent indicators. Finally, although rather detailed, our analyses could be improved by lifting the following limitations before practical applications are considered.

(i) Our model for testing component separation quality is rather naive and two major issues remain to be addressed before drawing definite conclusions: (a) component separation methods do not move Galactic components the same way in each part of the sky nor each component: the description of Galactic residuals in terms of a contribution simply proportional to them has to be improved; (b) here, Galactic components are ‘added’ everywhere: component separation can subtract CMB signal too, an effect that we did not consider here.

In fact, the analysis of component separation method quality needs a specific study for each method at use (Leach et al. 2008).

(ii) In this study we did not consider the use of foreground templates in the reference maps (Komatsu et al. 2002, 2011) to assess for the residuals. The use of templates, if they are reliable, is expected to correct for the biases introduced by the foregrounds. We can thus consider the results of our analyses as pessimistic in that respect.

(iii) We did not characterize biases induced by secondary anisotropies: they could be important, even dominant, as a bispectrum is created from the covariance between weak lensing and Sunyaev–Zeldovich effect or Integrated Sachs–Wolfe (ISW) effect. Indeed, previous studies (Goldberg & Spergel 1999; Serra & Cooray 2008; Hanson et al. 2009) have warned about these spurious signals and a future work about their effects on MFs is planned.

(iv) Our analyses neglected correlations in the noise. In a future work we shall refine them when reliable simulations of correlated noise are available.

7 And the number of clusters, that we call an MF here to simplify the presentation.

8 In particular by exploiting the nearly linear nature of MFs with respect to various independent non-Gaussian contributions.
In this paper, we have not used explicitly the analytic formulations of MFs for our NG studies, as was done for example, in Hikage et al. (2006, 2008). Actually, using the ‘Skewness parameters’ (Matsubara 2003) to study biases induced by secondaries, Galactic foreground and point sources would allow us to use the numerous bi-spectra studies in an effective way. On the other hand, the advantage of our Monte Carlo method is that it can be generalized to any statistics, for instance the skeleton length in the excursion (Novikov et al. 2006), or any type of NG, e.g. that induced by cosmic strings (e.g., Bouchet, Bennett & Stebbins 1988; Ringeval 2010, and references therein).

ACKNOWLEDGMENTS

We thank G. Roudier for providing us the hitmaps, J.-F. Cardoso for useful discussions about component separation and G. Castex for advices on point source masks.

REFERENCES

Aghanim N., Majumdar S., Silk J., 2008, Rep. Progress Phys., 71, 066902
Babich D., 2005, Phys. Rev. D., 72, 043003
Babich D., Pierpaoli E., 2008, Phys. Rev. D., 77, 123011
Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103
Bartolo N., Matarrese S., Riotto A., 2006, J. Cosmol. Astropart. Phys., 6, 24
Beichman C. A., Neugebauer G., Habing H. J., Clegg P. E., Chester T. J., eds, 1988, Infrared Astronomical Satellite (IRAS) Catalogs and Atlases. Explanatory Supplement, Vol. 1. Library of Congress Cataloging-in-Publication Data, Superintendent of Documents, US Government Printing Office, Washington D.C.
Bennett C. L. et al., 2003, ApJS, 148, 97
Bouchet F. R., Bennett D. P., Stebbins A., 1988, Nat, 335, 410
Chiang L.-Y., Naselsky P. D., Verkhodanov O. V., Way M. J., 2003, ApJ, 590, L65
Chingangbam P., Park C., Yogendran K. P., van de Weygaert R., 2012, ApJ, 755, 122
Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 2774
Delabrouille J. et al., 2012, preprint (arXiv:1207.3675)
de Zotti G., Ricci R., Mesa D., Silva L., Mazzotta P., Toffolatti L., González-Nuevo J., 2005, A&A, 431, 893
Dunkley J. et al., 2011, ApJ, 739, 52
Dupac X., Tauber J., 2005, A&A, 430, 363
Elsner F., Wandelt B. D., 2009, ApJS, 184, 264
Eriksen H. K., Novikov D. I., Lilje P. B., Banday A. J., Winitzki S., Kosowsky A., 1998, New Astron., 3, 75
Gonzalez-Nuevo J., Lesgourgues J., Toffolatti L., 2005, ApJ, 621, 1
Gott J. R. III, Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Gott J. R. III, Park C., Juszkiewicz R., Bies W. E., Bennett D. P., Bouchet F. R., Stebbins A., 1990, ApJ, 352, 1
Hanson D., Smith K. M., Challinor A., Liguori M., 2009, Phys. Rev. D., 80, 083004
Hikage C., Komatsu E., Matsubara T., 2006, ApJ, 653, 11
Hikage C., Matsubara T., Coles P., Liguori M., Hansen F. K., Matarrese S., 2008, MNRAS, 389, 1439
Hikage C., Koyama K., Matsubara T., Takahashi T., Yamaguchi M., 2009, MNRAS, 398, 2188
Kaiserman N., Kuroki-Sunoi H., Poutanen T., 2005, MNRAS, 360, 390
Komatsu E., Pogosian D., Shandarin S., 1990, MNRAS, 242, 200
Kofman L., Pogosian D., Shandarin S. F., Melott A. L., 1992, ApJ, 393, 437
Komatsu E. et al., 2003, ApJS, 148, 119
Komatsu E., Spergel D. N., 2001, Phys. Rev. D., 63, 063002
Komatsu E., Wandelt B. D., Spergel D. N., Banday A. J., Górski K. M., 2002, ApJ, 566, 19
Komatsu E., Spergel D. N., Wandelt B. D., 2005, ApJ, 634, 14
Komatsu E. et al., 2011, ApJS, 192, 18
Lacasa F., Aghanim N., Kunz M., Frommert M., 2012, MNRAS, 421, 1982
Lagache G., Puget J. L., 2000, A&A, 355, 17
Lapi A., Shankar F., Mao J., Granato G. L., Silva L., De Zotti G., Danese L., 2006, ApJ, 650, 42
Leach S. M. et al., 2008, A&A, 491, 597
Maldacena J., 2003, J. High Energy Phys., 5, 13
Mandolesi N. et al., 2010, A&A, 520, A3
Mangilli A., Verde L., 2009, Phys. Rev. D., 80, 123007
Matsubara T., 2003, ApJ, 584, 1
Matsubara T., 2010, Phys. Rev. D., 81, 083505
Mecke K. R., Buchert T., Wagner H., 1994, A&A, 288, 697
Moshir M., Kopman G., Conrow T. A. O., 1992, in Moshir M., Kopman G., Conrow T. A. O., eds, IRAS Faint Source Survey, Explanatory Supplement Version 2. Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA
Munshi D., Heavens A., 2010, MNRAS, 401, 2406
Natoli P. et al., 2010, MNRAS, 408, 1658
Negrello M., Magliocchetti M., Moscardini L., De Zotti G., Granato G. L., Silva L., 2004, MNRAS, 352, 493
Novikov D., Colombi S., Doré O., 2006, MNRAS, 366, 1201
Okamoto T., Hu W., 2002, Phys. Rev. D., 66, 063008
Park C. et al., 2005, ApJ, 633, 11
Planck Collaboration et al., 2011a, A&A, 536, A1
Planck Collaboration et al., 2011b, A&A, 536, A7
Planck Collaboration et al., 2011c, A&A, 536, A18
Planck HFI Core Team et al., 2011, A&A, 536, A6
Pogosyan D., 1989, The 2D Numerical Code for the Large Scale Structure of the Universe in the Adhesion Model, Preprint A-7, Section of Physics and Astronomy, Estonian Academy of Sciences. Tallinn, Estonia
Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5
Ricci R., Prandoni I., Gruppioni C., Sault R. J., de Zotti G., 2006, A&A, 445, 465
Ringeval C., 2010, Adv. Astron., 2010
Schmalzing J., Buchert T., 1997, ApJ, 482, L1
Schmalzing J., Gorski K. M., 1998, MNRAS, 297, 355
Sefusatti E., Liguori M., Yadav A. P. S., Jackson M. G., Pajer E., 2009, J. Cosmol. Astropart. Phys., 12, 22
Serra P., Cooray A., 2008, Phys. Rev. D., 77, 107305
Sunyaev R. A., Zeldovich Y. B., 1972, Comment. Astrophys. Space Phys., 4, 173
Tauber J. A. et al., 2010, A&A, 520, A1
Toffolatti L., Argueso Gomez F., de Zotti G., Mazzei P., Franceschini A., Danese L., Burigana C., 1998, MNRAS, 297, 17
Vanmarcke E., 1983, Random Fields. The MIT Press, Cambridge, MA, p. 372
Winitzki S., Kosowsky A., 1998, New Astron., 3, 75
Yadav A. P. S., Komatsu E., Wandelt B. D., 2007, ApJ, 664, 680
Yadav A. P. S., Komatsu E., Wandelt B. D., Liguori M., Hansen F. K., Matarrese S., 2008, ApJ, 678, 578
Yadav A. P. S., Wandelt B. D., 2010, Adv. Astron., 2010

APPENDIX A: MINKOWSKI FUNCTIONALS: DEFINITIONS AND THEORY

For a field \( f(x) \) of zero average and variance \( \sigma^2 \) defined on the two-dimensional sphere \( S^2 \), an overdense excursion set is defined as

\[
\Sigma \equiv \{ x \in S^2 | f(x) > \nu \sigma \}. \tag{A1}
\]

The boundary of the excursion is

\[
\partial \Sigma \equiv \{ x \in S^2 | f(x) = \nu \sigma \}. \tag{A2}
\]
Then the three MFs on the sphere are

Area : \( V_0(v) = \frac{1}{4\pi} \int_{\Sigma} d\Omega, \) \hspace{1cm} (A3)

Perimeter : \( V_1(v) = \frac{1}{4\pi} \int_{\partial \Sigma} d\ell, \) \hspace{1cm} (A4)

Genus : \( V_2(v) = \frac{1}{4\pi} \frac{1}{2\pi} \int_{\partial \Sigma} \kappa d\ell, \) \hspace{1cm} (A5)

where \( d\Omega \) and \( d\ell \) are, respectively, elements of solid angles (surface) and of angle (distance), and \( \kappa \) is the geodesic curvature. Note that the Genus can be also expressed as the number of components in the excursion minus the number of holes in the excursion.

The fourth functional we use in this paper, \( V_3(v) \), is defined, for \( v > 0 \), as the number of components in the excursion. Symmetrically, for \( v < 0 \), it is the number of underdense components (or the number of components in the excursion \( \{x \in \mathbb{S}^2 \mid f(x) < \nu \sigma_0 \} \)).

In the Gaussian limit, the functionals can be expressed in the following way (see e.g. Vanmarcke 1983; Matsubara 2010):

\( V_k(v) = A_k v_k(v), \) \hspace{1cm} (A6)

with

\( v_k(v) = \exp(-v^2/2) H_{k-1}(v), \quad k \leq 2 \) \hspace{1cm} (A7)

\( v_3(v) = \frac{e^{-v^2}}{\text{erfc}(v/\sqrt{2})} \) \hspace{1cm} (A8)

and

\( H_0(v) = e^{-v^2/2} \left( -\frac{d}{dv} \right)^n e^{-v^2/2}. \) \hspace{1cm} (A9)

The amplitude \( A_1 \) depends only on the shape of the power spectrum \( C_\ell \):

\( A_k = \frac{1}{(2\pi)^{k+1}/2} \alpha_1 \frac{\sigma_1}{2\sigma_0} \left( \frac{\sigma_1}{\sqrt{2} \sigma_0} \right)^k, \quad k \leq 2 \) \hspace{1cm} (A10)

\( A_3 = \frac{2}{\pi} \left( \frac{\sigma_1}{\sqrt{2} \sigma_0} \right)^2 \) \hspace{1cm} (A11)

where \( \alpha_1 \equiv \pi^{k/2}/\Gamma(k/2 + 1) \), which gives \( \alpha_0 = 1, \alpha_1 = 2, \alpha_2 = \pi \) and \( \sigma_0 \) and \( \sigma_1 \) are, respectively, the rms of the field and its first derivatives.

In the weakly non-Gaussian regime, one can write, at leading order (see e.g. Matsubara 2010)

\( v_k \simeq v_k^{(0)} + v_k^{(1)} \sigma_0, \) \hspace{1cm} (A12)

where \( v_k^{(0)} \) is given by equations (A7) and (A8), while the first order non-Gaussian correction writes

\( v_k^{(1)}(v) = \frac{S}{6} H_{k+2}(v) - \frac{kS}{4} H_k(v) - \frac{k(k-1)S}{4} H_{k-2}(v), \) \hspace{1cm} (A13)

with

\( H_0(v) = e^{-v^2/2} \int_v^\infty e^{-\tau^2/2} d\tau = e^{v^2/2} \sqrt{\frac{\pi}{2}} \text{erfc}(\frac{v}{\sqrt{2}}). \) \hspace{1cm} (A14)

\(^9\) A component is a connected subset of the excursion.

and where

\( S = \frac{\langle f^4 \rangle}{\sigma_0^4}, \) \hspace{1cm} (A15)

\( S_L = \frac{\langle f^2 \nabla^2 f \rangle}{\sigma_0^4 \sigma_L^2}, \) \hspace{1cm} (A16)

\( S_H = \frac{2 \langle |\nabla f|^2 \rangle}{\sigma_L^4} \) \hspace{1cm} (A17)

are skewness parameters (Matsubara 2003). Each of them is a weighted integral of the bi-spectrum and is thus directly proportional to \( f_{3\ell} \).

**APPENDIX B: ALGORITHM USED TO COMPUTE MINKOWSKI FUNCTIONALS**

The code we developed is available by simple e-mail request. The numerical technique we use for measuring MFs on HEALPix maps consists, for each threshold value of the temperature, of two steps.

In the first step of our algorithm, an on-grid cluster analysis is performed in order to define the connected ensembles of pixels that have temperature values above the threshold. The output of this step is a map of integers (flags) that assign a negative ‘outside’ flag for pixels below the threshold and a positive flag that corresponds to the cluster number of each connected region above the threshold. Optionally, the boundary pixels (defined either within high excursion regions or just outside them) can be marked. The implementation follows closely the CND_REG2D procedure developed first for the Adhesion model (Pogosyan 1989; Kofman, Pogosyan & Shandarin 1990; Kofman et al. 1992).

The on-grid cluster analysis then proceeds as follows. At the onset, all pixels are considered as ‘unseen’. In the outer loop, the code scans the map, checking the temperature values until the first pixel exceeding the threshold is encountered. Along the way, the checked pixels that have been found to be below the threshold are marked with the ‘seen, below’ flag. The new pixel above the threshold is assigned the flag that corresponds to the cluster order number (zero for the first found). Its neighbours are investigated. The ones below the threshold are marked as ‘seen, below’, the ones above the threshold are given the same cluster number flag and are put on to stack for further analysis. Next, the pixel from a stack is drawn, and its neighbours are checked in the same way, with the ones above the threshold further added to the stack. This inner loop proceeds until the stack is exhausted which signifies that all connected pixels belonging to the first cluster are found. The control is reverted to the outer scanner that proceeds checking the pixels, skipping over those already analysed, until the next previously unseen pixel above the threshold is found. This pixel acquires the next cluster order number and then the inner, stack-based, loop finds all the pixels connected to it, and so on.

This code is fast and linear in the number of pixels since the neighbours of each pixel are analysed only once. Optionally, the boundary pixels can be marked differently from pixels inner to the regions. Different decisions about what constitutes connected pixels can be easily implemented. In the current implementation, we consider all pixels that have at least a common vertex belonging to the same cluster (as we shall see below this has to be taken into account when considering the Euler characteristic of the total excursion set). The immediate byproducts of the cluster analysis...
are the volume (area) of all connected regions above the threshold, which constitutes the first MF, and their number $N_{\text{cluster}}$.\(^{13}\)

The other two MFs of the 2D excursion sets, the Euler characteristics and the perimeter of the set, are computed in the second step of our algorithm, using the just obtained clustering information.

The Euler characteristic of each individual cluster, and, after summation over all the clusters, of the whole excursion set above the threshold, can be obtained in one pass by analysing the pixel grid vertices on the cluster boundaries. Gauss–Bonnet theorem links Euler characteristics of the region to the integration of the curvature of its boundary. However, since topological properties are invariant under any continuous transformation of the boundary, the need to explicitly evaluate the boundary curvature is eliminated by considering the curvature to be accumulated just in the outside vertices of the boundary pixels. Thus one only needs to assign the appropriate curvature weights to the grid vertices and sum over them.

Similar procedure on the Cartesian grid has been described in Gay, Pichon & Pogosyan (2012). A 3D version of the code is also readily available.

Clearly, the vertex contribution is determined solely by the temperature distribution of the pixels that form this vertex – which are below and which are above the threshold. Necessary weights can be boot-strapped by considering the elementary situations. Most of the grid vertices in HEALPix pixellization are regular, being formed by four adjacent pixels.

(a) Consider a single pixel cluster above the threshold. Its Euler characteristic is $\chi = 1$. It has four boundary vertices, all four equivalent, formed by one adjacent pixel above the threshold and three below. They should contribute equally, thus we assign the weight 1/4 to any vertex of this type.

(b) Consider next a two pixels cluster, formed by pixels having a common side. The boundary has six vertices, four corner ones of the type (a) and two new ones, which are formed by two side-by-side adjacent pixels above the threshold and two below it. The cluster has $\chi = 1$, thus new vertices must contribute a weight equal to 0.

(c) Consider a three pixels cluster that forms a corner. It has five vertices of type (a), which weights add up to 5/4, two vertices of type (b) and one new vertex that is formed by three pixels above the threshold and one below it. For the total $\chi = 1$ this vertex must contribute the weight $-1/4$.

(d) The last possibility arises when we consider the boundary vertex formed by two pixels above the threshold that are touching just at the vertex. Its weight depends whether we consider the clusters to be linked through the vertex or disjoint. In the former case, which corresponds to our choice in cluster analysis, the weight is $1 - 6/4 = -1/2$, while in the latter case it is $2 - 6/4 = +1/2$.

However, for statistical analysis of the total Euler characteristic of the excursion sets, both choices, while exact for the corresponding clustering decision, will lead to biased results. Indeed, one may argue that in a situation when the discretized field has high and low pixels mixing at a vertex, it is equally probable that the regions below the threshold or above it connect through this vertex. Assigning the weight zero in this case will reflect such a 'statistical' consideration and this is the choice that we make in our statistical analyses.

(e) The vertex that is formed by four pixels above the threshold is not a boundary vertex and its weight is zero.

The HEALPix grid on a sphere has eight special vertices that are formed only by three pixels. Their weights need to be defined separately. Consider elementary clusters that have one of such vertices.

(f) A one-pixel cluster with a special vertex has three regular vertices of type (a). Thus, a special vertex with one pixel above the threshold and two below has the weight 1/4.

(g) A two pixels cluster with a special vertex on a side shows that a special vertex formed by two pixels above the threshold and one below has the weight 0.

(h) A three pixels cluster with a special vertex inside has only three exterior regular vertices of type (a). Thus, a special vertex formed by all three pixels above the threshold, albeit being an inner one, has a weight 1/4. These eight special vertices are providing the Euler characteristic of the entire spherical manifold, $\chi = 2$, when all pixel values lie above the threshold.

Once the weights are defined, the computation of the Euler characteristic of individual clusters and the entire excursion set is a simple one loop over vertices to determine their type from the temperature value at the adjacent pixels and add the weights based on the index of pixels that are above the threshold. Our method does not involve any differentiation or integration of the field. Moreover, if only the total $\chi$ is needed, the cluster analysis step can be omitted. The code works in the presence of masks of arbitrary complexity, treating the field as defined on a manifold that itself has a non-trivial Euler characteristic.

The measurement of the perimeter of the excursion set is somewhat more complex. Our technique is again based on a scan over the boundary grid vertices, for each of which we linearly interpolate the field based on the values in the four (or three) adjacent pixels. Interpolation is least-squares, linear, performed in the plane tangent to the sphere at the vertex. Interpolation requires first derivatives of the field which are precomputed by standard HEALpix routines and stored. After finding the best-fitting plane for the local field behaviour, we compute the intersection of this plane with the polygon formed by the pixel centres around the vertex. The length of the intersection is then added to the perimeter of the cluster the vertex belongs to. We found that linear interpolation is sufficient given the level of accuracy offered by the pixellization and there is no advantage to go for a higher order interpolation in comparison with increasing the pixellization level.

The code we developed along these lines is available by simple e-mail request.

**APPENDIX C: INVERSION OF THE COVARIANCE MATRIX: A CONVERGENCE STUDY OF THE $\chi^2$**

The inversion of matrix $C$ in equation (7) is performed using Singular Value Decomposition (SVD): $C = U\Sigma U^T$ or

$$C_{ij} = \sum_{k=1}^{n} U_{ik} d_{ik} U_{jk}, \quad (C1)$$

hence

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{1}{d_{ik}} U_{ik} U_{jk} [\hat{y}_i - \bar{y}_i(f_{NL})] [\hat{y}_j - \bar{y}_j(f_{NL})]. \quad (C2)$$

where $U$ is an orthogonal matrix and $D$ a diagonal matrix such that its diagonal terms $d_{ii}$, the singular values, are ranked in descending order. Here, these latter are expected to be positive, since $C$ is positive-definite by construction. When there is a large contrast

\(^{10}\)To compute $N_{\text{cluster}}$ for negative thresholds, one just multiplies the map by $-1$, and repeats the procedure.
between the singular values, \( r \equiv d_1/d_n \gg 1 \), the calculation of the inverse of \( C \) becomes an ill-conditioned problem if \( C \) is not computed with sufficient accuracy. In our calculations, the accuracy is mainly controlled by the number \( m \) of simulations used to estimate \( C \), but it also depends on other parameters such as the number of bins, the value of \( v_{\text{max}} \), the smoothing scale \( \ell \), the presence of masks, etc. In particular, the ratio \( r \) increases when the space between successive bins in the excursion is reduced, because they become more correlated. A sufficiently accurate calculation of \( C \) can be costly. For instance, it takes about 3 min at present time to make MFs on a HEALPix map with \( N_{\text{side}} = 2048 \), \( v_{\text{max}} = 3.5 \), \( n_{\text{bins}} = 26 \) and a Gaussian smoothing of \( \theta^S_{\text{FWHM}} = 10 \) arcmin, an operation that can become prohibitive if repeated many times; this is why it is worth investigating in details what should be a reasonable value of \( m \) to have a few percent error on the estimate of the \( \chi^2 \).

In this appendix, instead of studying the accuracy in the determination of the inverse of \( C \), we analyse directly the convergence of the \( \chi^2 \) as a function of \( m \) for \( f_{\text{NL}} = 0 \). We focus on the perimeter functional, \( \hat{\chi} = \hat{\phi}_1 \), but similar results would be obtained for other functionals. To perform the Gaussian simulations for computing the covariance matrix, we shall use the following typical configuration: the HEALPix resolution is \( N_{\text{side}} = 2048 \) and the power spectrum is calculated as detailed in Appendix D. The maps are convolved with a Gaussian beam of size \( \theta^b_{\text{FWHM}} = 7.1 \) arcmin, then supplemented with a white noise of 0.33 \( \mu \text{K.deg} \), and finally smoothed with a Gaussian smoothing kernel with \( \theta^S_{\text{FWHM}} = 10 \) arcmin. Our concept of infinity corresponds to \( m = 20000 \) simulations for estimating the covariance matrix in the ‘ideal’ case.

Here, we want to be able to estimate \( m \) for various choices of the parameters which are the most influential in the convergence of the covariance matrix, namely the number of bins, \( n \) and the excursion range, \( v_{\text{max}} \). We already computed a covariance matrix \( C_{20k} \) for a rather large value of \( m = 20000 \). However, the optimal value of \( m \) remains unknown and might be larger than 20 000 in some cases. Now, to make a convergence study, we should realize a large number of trials, each of them with a subset of \( m \) realizations, in order to estimate the error in the estimate of \( C \) with \( m \) simulations. To avoid that procedure to be prohibitive, instead of generating random maps and measuring the genus on each of them, we use the property that the random vectors \( \hat{\chi} \equiv \hat{y}^{(0)} - \hat{y}^{(i)} \) are nearly Gaussianly distributed to simulate directly and rapidly a Gaussian distribution of \( m \) vectors \( \hat{\chi} \) of zero average and covariance \( C_{20k} \). To do that, we use the following standard procedure. We draw 

\[
\hat{\chi}_N = N(0, 1) \text{ of size } m, \tag{C3}
\]

then the vector 

\[
\hat{\chi} = U(\sqrt{D} \hat{\chi}_N) \tag{C4}
\]

verifies, as required, 

\[
\langle \hat{\chi} \hat{\chi}^T \rangle = UDUT = C_{20k}. \tag{C5}
\]

The point is to see whether using \( m \) realizations, \( \hat{\chi}^{(i)} \), \( i = 1, \ldots, m \), of vector \( \hat{\chi} \) to estimate the covariance matrix is sufficient to compute the \( \chi^2 \) accurately enough. For one data realization with \( f_{\text{NL}} = 0 \), 

\[
\hat{\chi} = \hat{\phi}_1(0) + \hat{\phi}_2, \tag{C6}
\]

where \( \hat{\chi}^{(i)} \) is generated randomly the same way as just above, the estimator of the \( \chi^2 \) writes 

\[
\chi^2_{\text{MC}}(m) = [\hat{\chi} - \hat{\chi}(f_{\text{NL}})]^T C_{\text{MC}}^{-1}(m) [\hat{\chi} - \hat{\chi}(f_{\text{NL}})], \tag{C7}
\]

with the covariance matrix estimated from \( m \) realizations, 

\[
C_{\text{MC}}(m) = \frac{1}{m} \sum_{i=1}^{m} \chi^{(i)}(\hat{\chi}^{(i)})^T, \tag{C8}
\]

while the ‘true’ \( \chi^2 \) should write 

\[
\chi^2_{\text{true}} = [\hat{\chi} - \hat{\chi}(f_{\text{NL}})]^T C^{-1}_{\text{true}} [\hat{\chi} - \hat{\chi}(f_{\text{NL}})], \tag{C9}
\]

where we approximate here \( C^{-1}_{\text{true}} \) by the inverse of \( C_{20k} \). The relative difference between the exact \( \chi^2 \) and the one driven from \( m \) simulations writes 

\[
\epsilon_m(\hat{\chi}) = \frac{\chi^2_{\text{MC}}(m) - \chi^2_{\text{true}}}{\chi^2_{\text{true}}}, \tag{C10}
\]

where \( \hat{\chi} \equiv (\hat{\chi}^{(0)}, \hat{\chi}^{(1)}, \ldots, \hat{\chi}^{(m)}) \). To assess the convergence, we need the second moment of \( \epsilon_m \) over many realizations \( M \) of \( \hat{\chi} \) (this corresponds to \( M \times (m + 1) \times n \) random numbers in total) to be small 

\[
\langle (\epsilon_m(\hat{\chi}))^2 \rangle \leq \epsilon^2. \tag{C11}
\]

We performed this exercise for \( \epsilon = 2 \) per cent, with \( M = 1000 \) and for the following set of values of \( m = 1000, i \in \{1, \ldots, 20\} \). The results of our analyses are summarized in Table 1.

**APPENDIX D: SIMULATING PLANCK DATA**

**D1 Observations with Planck**

Planck is the third and latest generation of space observatory designed to observe the anisotropies of the CMB over the entire sky, at a high sensitivity and angular resolution. It observes the sky in nine frequency bands: the Low Frequency Instrument (LFI; Mandolesi et al. 2010) covers the 30, 44 and 70 GHz bands with amplifiers cooled to 20 K; the High Frequency Instrument (HFI; Planck HFI Core Team et al. 2011) covers the 100, 143, 217, 353, 545 and 857 GHz bands with bolometers cooled to 0.1 K.

In order to assess the characteristics of Planck which are particularly relevant for the specific purposes of this paper, i.e. the noise level and the resolution of the CMB map analysed, we consider the expected performances of the three HFI channels at 100, 143 and 217 GHz, and we adopt two durations for the mission: 15
months and 30 months which we shall refer to in the following as the nominal and extended mission.

The specifications of the three channels that are studied here are gathered in Table D1. The beams are assumed to be isotropic (‘circular’) Gaussian. The beam size in Table D1 is given in terms of ‘FWHM’ scale as detailed in Section D2.

D2 Simulations

The simulation procedure used in this article can be summarized as follows:

\[
\text{map} = \text{CMB}(a_{\ell m}, f_{NL}) \ast \text{beam}\langle \theta_{\text{FWHM}}(f) \rangle \\
+ \text{Galactic Foreground} \ast \text{beam}\langle \theta_{\text{FWHM}}(f) \rangle \\
+ \text{Sources Foreground} \ast \text{beam}\langle \theta_{\text{FWHM}}(f) \rangle \\
+ \text{noise}(f) + \text{Point Sources Mask inpainted}(\beta, f) \\
+ \text{smoothing} \langle \theta_{\text{FWHM}}(f) \rangle \text{ or Wiener filtering} \\
+ \text{Galactic Mask}(f_{\text{sky}}),
\]

where \( f \) is the considered frequency. We mainly describe here how the first line of this equation is dealt with, while other aspects are examined in various sections of the article.

To generate the Gaussian part of the CMB, we use harmonic coefficients \( a_{\ell m} \) derived from standard cold dark matter (CDM) cosmology, with the best parameters obtained from WMAP7+BAO+H\(_0\) (Komatsu et al. 2011): \( \Omega_{\Lambda} = 0.728, \Omega_{m}h^2 = 0.1123, \Omega_{b}h^2 = 0.0226, H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}, n_s = 0.967, \tau = 0.085 \) and \( \Delta_{\text{L}}^2(k_0) = 2.42 \times 10^{-9} \).

The maps are generated using HEALPix pixelization (G"orski et al. 2005). Most of the calculations performed in that paper use \( N_{\text{side}} = 1024 \) and assume a truncation of the harmonic modes at \( \ell = \ell_{\text{max}} = 2000 \), but we also examined other values of \( N_{\text{side}} \) and \( \ell_{\text{max}} \) as shown in Table 4.

For the non-Gaussian part of the CMB, we used updated versions of simulations of harmonic coefficients provided by Elsner & Wandelt (2009) which allow us to choose \( N_{\text{side}} \) and \( \ell_{\text{max}} \) up to \( \ell = 3500 \).

Convolution, in particular with the Gaussian window, was performed in harmonic space using HEALPix package. In our conventions, the Gaussian kernel writes, in harmonic space,

\[
W_\ell = \exp \left[ -\ell(\ell + 1) \frac{\theta_{\text{FWHM}}^2}{2} \right],
\]

where \( \theta_{\text{FWHM}} \) is the Gaussian size. Here we will use instead the corresponding FWHM: \( \theta_{\text{FWHM}} = \theta_{\text{FWHM}} \times 2.35 \).

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