

PAPER • OPEN ACCESS

B-sure. Part I. Minkowski functionals as robustness test for tensor-to-scalar ratio detection from CMB observations

To cite this article: Claudio Ranucci *et al* JCAP01(2026)045

View the [article online](#) for updates and enhancements.

You may also like

- [Enhancing Morphological Measurements of the Cosmic Web with Delaunay Tessellation Field Estimation](#)
Yu Liu, Yu Yu, Pengjie Zhang et al.
- [STATISTICAL AND SYSTEMATIC ERRORS IN THE MEASUREMENT OF WEAK-LENSING MINKOWSKI FUNCTIONALS: APPLICATION TO THE CANADA–FRANCE–HAWAII LENSING SURVEY](#)
Masato Shirasaki and Naoki Yoshida
- [Bioink with cartilage-derived extracellular matrix microfibers enables spatial control of vascular capillary formation in bioprinted constructs](#)
Margo L. Terpstra, Jinyu Li, Anneloes Mensinga et al.

B -sure. Part I. Minkowski functionals as robustness test for tensor-to-scalar ratio detection from CMB observations

Claudio Ranucci<sup>,^{a,b,c,*} **Alessandro Carones**<sup>,^{a,b,c} **Léo Vacher**<sup>,^{a,b,c}
Nicoletta Krachmalnicoff^{,^{a,b,c} and **Carlo Baccigalupi**^{,^{a,b,c}}}</sup></sup></sup>

^a*International School for Advanced Studies (SISSA),*

Via Bonomea 265, Trieste 34136, Italy

^b*Istituto Nazionale di Fisica Nucleare (INFN) Sezione di Trieste,*

Via Valerio 2, Trieste 34127, Italy

^c*Institute for Fundamental Physics of the Universe (IFPU),*

Via Beirut 2, Trieste 34151, Italy

E-mail: cranucci@sissa.it, acarones@sissa.it, leo.vacher@ijclab.in2p3.fr,
nkrach@sissa.it, bacci@sissa.it

ABSTRACT: The detection of primordial B -mode polarisation of the Cosmic Microwave Background (CMB) is a major observational goal in modern Cosmology, offering a potential window into inflationary physics through the measurement of the tensor-to-scalar ratio r . However, the presence of Galactic foregrounds poses significant challenges, possibly biasing the r estimate. In this study we explore the viability of using Minkowski functionals (MFs) as a robustness test to validate a potential r detection by identifying non-Gaussian features associated with foregrounds contamination. To do so, we simulate sky maps as observed by a *LiteBIRD*-like CMB experiment, with realistic instrumental and foregrounds modelling. The CMB B -mode signal is recovered through blind component separation algorithms, and the obtained (biased) value of r is used to generate Gaussian realisation of CMB signal. Their MFs are then compared with those computed on maps contaminated by foreground residual left by component separation, looking for a detection of non-Gaussianity. Our results demonstrate that, with the experimental configuration considered here, MFs can not be reliably adopted as a robustness test of an eventual r detection, as we find that in the majority of the cases MFs are not able to raise significant warnings about the non-Gaussianity induced by the presence of foreground residuals. In the most realistic and refined scenario we adopted, the test is able to flag non-Gaussianity in $\sim 26\%$ of the simulations, meaning that there is no

*Corresponding author.

warning on the biased tensor-to-scalar ratio in $\sim 74\%$ of cases. These results suggest that more advanced statistics than MFs must be considered to look for non-Gaussian signatures of foregrounds, in order to be able to perform reliable null tests in future CMB missions.

KEYWORDS: CMBR polarisation, non-gaussianity, CMBR experiments, gravitational waves and CMBR polarization

ARXIV EPRINT: [2510.18868](#)

Contents

1	Introduction	1
2	Minkowski functionals	3
3	Methodology	5
3.1	Input simulations	7
3.2	Component separation	8
3.3	Likelihood	10
3.4	Gaussian simulations for MFs comparison	13
3.5	Compatibility analysis	15
4	Results and discussion	17
5	Conclusions	22
A	Supplementary results	23
A.1	Goodness-of-fit analysis	23
A.2	Examples of MFs in other scenarios	23
A.3	Dependence of MFs test efficiency on the foreground model complexity	24

1 Introduction

Observations of the Cosmic Microwave Background (CMB) have played a crucial role in establishing the Λ Cold Dark Matter (Λ CDM) as the standard cosmological model, providing insights into the origin of the structure, content and evolution of the Universe [1–3]. Most of the information regarding the early phases of the cosmic history has been extracted from temperature fluctuations; however, the fainter CMB polarisation field can still shed new light on the physics of the primordial Universe. A phase of exponential expansion which occurred $10^{-36} - 10^{-34}$ s after the Big Bang, called cosmic inflation [4–6], is canonically added to the standard model to explain some features of the hot Big Bang Cosmology. This expansion stretched quantum fluctuations of the metric (primordial gravitational waves, GWs) to cosmological scales, which are expected to produce a specific signature in the CMB polarisation: the primordial B -mode pattern [7–9]. Detecting primordial B -modes in CMB polarisation could enable the estimation of the amplitude of primordial gravitational waves, usually quantified by the tensor-to-scalar ratio parameter r [10], potentially providing experimental evidence for the inflationary scenario. Due to the faintness of such a signal ($> 10^3$ times weaker than temperature anisotropies), there is no direct evidence for primordial B -modes yet, and current CMB experiments have placed tight constraints on their amplitude, finding $r_{0.05} < 0.032$ [11] and $r_{0.01} < 0.028$ [12] (with a free-to-vary tensor spectral tilt) at 95% confidence when evaluated at a pivot scale of 0.05 or 0.01 Mpc^{-1} . At the same time, these experiments also observed that the power spectrum of primordial scalar perturbations, generated by inflation, is not exactly scale-independent, with the scalar spectral index

$n_s - 1 \sim -0.035$ (e.g., [2]). This measurement is compatible with several classes of inflationary models predicting r to be in the $10^{-3} - 10^{-2}$ range (see [13] and references therein).

The detection of primordial B -modes thus represents one of the main goals of future CMB missions. Their power spectrum is characterised by the presence of two different features: the reionisation bump ($\ell \lesssim 10$), associated with the scattering of CMB photons with free electrons released during cosmic reionisation, and the recombination bump ($\ell \sim 80$), which corresponds to the imprint of primordial tensor perturbations at the recombination epoch. The first can be measured only through full sky observations from space, while the second can be targeted by observing smaller regions of the sky from the ground [14–17].

Even though the only source of primordial large-scale B -modes are tensor fluctuations (at linear order), a practical measurement is complicated by several factors. The gravitational deflection of the background CMB photons by the cosmic large-scale structure creates coherent sub-degree distortions in the CMB, known as CMB lensing [18, 19]. Through this mechanism, a fraction of the parity-even E -modes is transformed into parity-odd B -modes at intermediate and small scales. Lensing B -modes have already been measured by SPTpol [20], ACTpol [21], PolarBear [22] and BICEP2/Keck [23] experiments.

Diffuse Galactic emission is significantly polarised, and in particular components such as synchrotron radiation and thermal emission from dust grains produce B -modes with a significant amplitude. These two emission mechanisms are obviously prominent around the Galactic plane, but are also clearly detectable at higher latitudes [24]. Current measurements of Galactic emission demonstrate that the Galactic B -mode signal is dominant over the cosmological signal on all scales [25–28]. At the minimum of polarised Galactic thermal dust and synchrotron, around 80 GHz, their B -mode signal represents an effective tensor-to-scalar ratio with amplitude larger than the sensitivity targeted by future CMB experiments, even in the cleanest regions of the sky [29].

Component separation methods, which exploit the different spectral energy distributions (SED) of the CMB and foregrounds to separate the different components, are thus of vital importance [30, 31]. Practical implementations of these methods must also be able to carry out this separation in the presence of instrumental noise and systematic effects [32]. Several component separation methods have been developed in the last years for CMB data analysis; here we mention two categories: 1) parametric-fitting methods [33–36], which recover the CMB signal by fitting a model of the various sky components; and 2) the “blind” methods [37, 38], whose purpose is to recover a cleaned CMB blackbody signal, without any assumption on the SED of foreground emission. Methods of the latter class, in most cases, are also referred to as minimum-variance techniques, since they reconstruct the CMB signal as the minimum variance solution from the linear combination of multi-frequency observations, thus maximally reducing the foreground contamination in the 2-point statistics [39].

The application of component separation algorithms is able to reduce most of the foregrounds impact, but residual contamination in the reconstructed CMB signal may still be comparable in amplitude to the primordial B -mode signal to be measured, thus biasing any estimate of the tensor-to-scalar ratio. In recent years, many works have tackled this problem and forecasted the quality of an r estimation with different ground-based and satellite experiments (e.g., [14, 40, 41]). In general, these works have highlighted how, if left untreated,

systematic residuals arising from a simplistic characterisation of foregrounds could bias a $r \sim 10^{-3}$ measurement by several σ , consequently leading to a false detection, with the outcome of the BICEP2 analysis [42] being a clear example. Thus, it is crucial to develop methodologies able to detect such residual contamination in the data, in order to validate any potential detection of the tensor-to-scalar ratio.

Minkowski functionals (MFs) are mathematical tools used to assess the statistics of CMB anisotropies; they have been widely explored in the mathematical literature ([43–45]), and have become popular in CMB analysis after the discussions by [46] and [47]. In the case of scalar Gaussian isotropic fields such as the CMB temperature anisotropies, the theoretical predictions of these functionals can be accurately computed. Given their low variance, any deviation from Gaussianity or statistical isotropy can be detected at high significance in a model-independent manner. MFs have been adopted to investigate non-Gaussianity and anisotropy in the CMB [48], to study the non-Gaussianity of Galactic signals [49], and also to constrain cosmological parameters [50]. They are additive for disjoint regions of the sky and invariant under rotations and translations. Moreover, these tools can be applied to blindly detect contamination by foreground and instrumental systematics in the reconstructed CMB map, given their highly non-Gaussian and anisotropic nature. The impact of such systematic and astrophysical effects on the constraining power for primordial non-Gaussianity in temperature data has also been explored by [51]. Until now, MFs have been largely applied to the scalar fields associated with the CMB polarisation, E - and B -modes ([48, 52–54]), and recently the formalism has been extended to spin- s fields [55], including their combination into the polarisation modulus P^2 [56]. The aim of this work is to explore the application of MFs as high-order statistical tools able to test the robustness of a potential detection of the tensor-to-scalar ratio r . We investigate if MFs are powerful enough to recognise a contribution to r partially or entirely sourced by residual foreground contamination, through a measurement of non-Gaussianity in the CMB maps.

The paper is organised as follows: in section 2, we introduce the MFs theoretical framework, while the full methodology adopted is described in section 3; results are presented in section 4, and we summarise our conclusions and provide future prospects in section 5.

2 Minkowski functionals

Minkowski functionals (MFs) on the sphere are now well-established tools for CMB data analysis [48]; here we recall only the basis of the theory to facilitate comparison with other works. For the complete derivation of the equations and a more detailed discussion, we refer to [55, 56] and references therein.

Given a scalar random field f observed on the unit sphere $\mathbb{S}^2$ (e.g., CMB temperature anisotropies, E -/ B -modes), the excursion set at a threshold u is defined as

$$A_u(f, \mathbb{S}^2) = \{x \in \mathbb{S}^2 : f(x) > u\}. \quad (2.1)$$

We will omit the arguments of A_u when f and its domain are clear from the context. As recalled in [47], the morphological properties of these excursion sets can be summarised

by the three MFs defined as follows:

$$\begin{aligned} V_0(A_u) &= \int_{A_u} dx, \\ V_1(A_u) &= \frac{1}{4} \int_{\partial A_u} dr, \\ V_2(A_u) &= \frac{1}{2\pi} \int_{\partial A_u} \kappa(r) dr, \end{aligned} \quad (2.2)$$

where dr denotes a line element along the boundary ∂A_u of the excursion set and $\kappa(r)$ denotes the geodesic curvature of the said boundary. Thus, V_0 represents the excursion area, V_1 is one-fourth of the boundary length (perimeter), and V_2 (genus) is associated with the number of connected regions minus the number of holes. It is notationally more convenient to replace MFs with the equivalent notion of Lipschitz-Killing curvatures, expressed as

$$\begin{aligned} \mathcal{L}_2(A_u) &= V_0(A_u) \\ \mathcal{L}_1(A_u) &= 2V_1(A_u) \\ \mathcal{L}_0(A_u) &= V_2(A_u) + \frac{1}{2\pi} V_0(A_u). \end{aligned} \quad (2.3)$$

For an isotropic Gaussian field (normalised to have unit variance), the expected value of the Lipschitz-Killing curvatures on the sphere is given by the Gaussian kinematic formula (GKF):

$$\mathbb{E}[\mathcal{L}_j(A_u)] = \sum_{k=0}^{2-j} \frac{\omega_{k+j}}{\omega_k \omega_j} \binom{k+j}{k} \cdot \rho_k(u) \cdot \mathcal{L}_{k+j}(\mathbb{S}^2) \cdot \mu^{k/2}, \quad (2.4)$$

where we defined the “flag” coefficients [57] as $\omega_j = \pi^{j/2} \left[\Gamma(\frac{j}{2} + 1) \right]^{-1}$ with $\Gamma(z)$ being the usual Gamma function, and μ as the derivative of the covariance function at the origin for the field f , computed as

$$\mu = \sum_{\ell} \frac{\ell(\ell+1)}{2} \frac{2\ell+1}{4\pi} C_{\ell}, \quad (2.5)$$

where C_{ℓ} denotes, the angular power spectrum of the field f . The quantity ρ_k represents a set of “density” functions,

$$\rho_k(u) = \frac{1}{(2\pi)^{k/2}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) H_{k-1}(u), \quad (2.6)$$

with $H_k(u)$ denoting the sequence of Hermite polynomials. With eqs. (2.3) and (2.4) we get the well-known results for the expected values of MFs, which we report normalised by area:

$$\begin{aligned} \frac{\mathbb{E}(V_0)}{4\pi} &= 1 - \Phi(u) \\ \frac{\mathbb{E}(V_1)}{4\pi} &= \frac{1}{8} \exp\left(-\frac{u^2}{2}\right) \mu^{1/2} \\ \frac{\mathbb{E}(V_2)}{4\pi} &= \frac{1}{\sqrt{(2\pi)^3}} \mu \cdot \exp\left(-\frac{u^2}{2}\right) \cdot u. \end{aligned} \quad (2.7)$$

where Φ is the cumulative function of the normal distribution. We note that the first MF depends only on u , not on the field itself as long as it is normalised to have unit variance. The above set of equations tell us that the MFs expected values, in the case of a Gaussian random field, depend on the combination of four different independent ingredients:

1. a set of universal coefficients (called “flag” coefficients);
2. the Lipschitz-Killing curvatures evaluated on the original manifold (in our case, the unitary sphere);
3. a power of the derivative of the covariance function of the field at the origin;
4. a set of “density” functions ρ_k , dependent only on the threshold at which the MFs are evaluated, and on the statistics of the field f .

We note that the presence of a mask only affects the manifold where the map is defined, therefore having an impact only on the MFs normalisation, at leading order.

A key point to notice is that, from eq. (2.4), the MFs expected value depends both on the level of Gaussianity of the field (in general, on the statistics of the field through ρ_k), and on its spectral features through the C_ℓ term in the derivative of the covariance function, in eq. (2.5). This implies that two fields will have different MFs values either due to the different statistical properties or to the different shape of the spectra. This point will be discussed in more details in the following sections.

In this work, MFs of CMB maps are evaluated with `Pynkowski`,¹ a python package developed to compute MFs and other higher order statistics of fields defined on the sphere and other manifolds. The package has been presented and validated in [55, 56], both for scalar (T , E , B , P^2) and polarisation (Q , U) maps.

Following what has been done in previous works [48, 56] in order to characterise the MFs we consider two approaches for the scale-dependent analysis of the B -mode sky maps: *i*) simultaneous exploitation of all the scales sampled in the maps (i.e., using all the multipoles available from $\ell_{\min}$ to $\ell_{\max}$) and *ii*) scales separation by using needlet bands, to isolate the contribution from different ranges of multipoles. The localisation properties of needlets (in real and harmonic space) allow a much more precise, scale-by-scale, interpretation of any possible anomalies, since non-Gaussian or statistically anisotropic signal may be relevant only in a specific range of multipoles (e.g., Galactic residuals on large scales). The needlets approach should then provide a complete view about the nature of possible non-Gaussian features detected with the MFs. Their formalism is described in more detail in section 3.2.

3 Methodology

As outlined in section 1, the goal of this work is to build a robustness test for a potential tensor-to-scalar ratio detection, using MFs as statistical tools. For reference, we describe here all the steps we follow to build this MFs-based robustness test, with the proposed

¹<https://javicarron.github.io/pynkowski/pynkowski.html>.

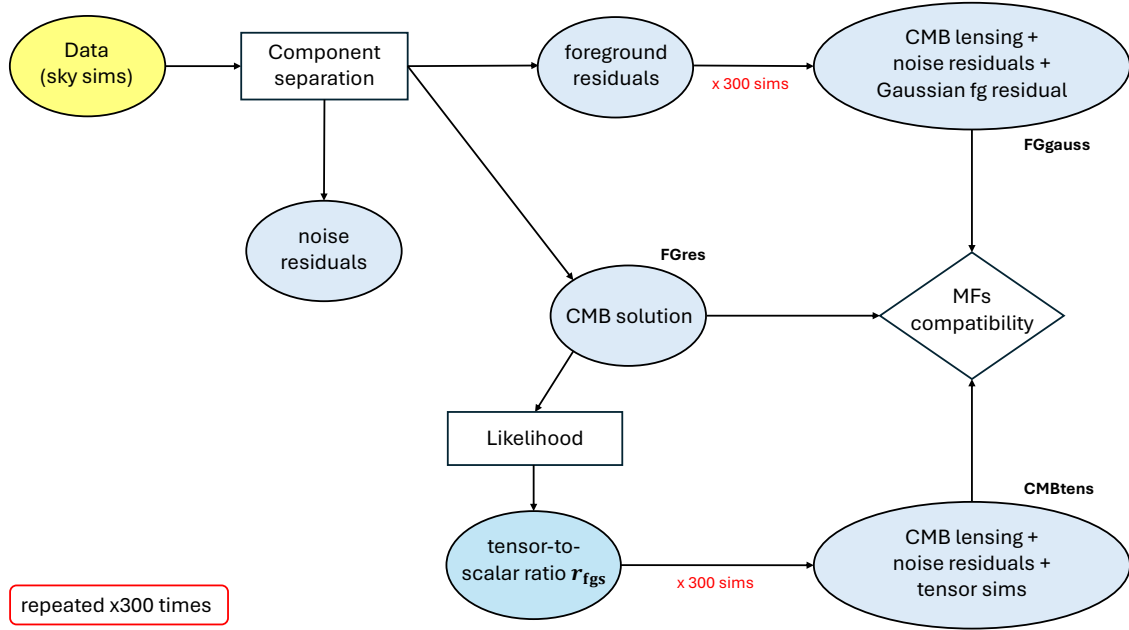


Figure 1. Overview of the steps adopted to build the MFs robustness test. Squares represent pipeline stages, while ovals indicate data products (input in yellow, intermediate in light blue). This procedure is then repeated for 300 “data” simulations.

analysis applied to the instrumental configuration of a *LiteBIRD*-like experiment.² The procedure is also illustrated in figure 1. Each one of the item is then discussed in details in the correspondent section:

1. first, we generate a set of multi-frequency *LiteBIRD*-like full-sky simulations in polarisation, (Q and U Stokes parameters), by co-adding CMB (with B -modes sourced by Gaussian realisation of lensing only, $r = 0$), instrumental noise, and different models of polarised emission from Galactic foregrounds (section 3.1). Simulations are then converted to B -mode maps through a full-sky harmonic transformation;
2. we apply component separation to these B -mode maps, obtaining three outputs: a cleaned CMB map (CMB solution), a map of noise residuals and a map of (non-Gaussian) foreground residuals (section 3.2), with the CMB solution containing the other two maps;
3. we mask in each product the expected most foreground-contaminated sky regions. We evaluate the likelihood function on the power spectrum of the cleaned CMB map, inferring the tensor-to-scalar ratio r_{fgs} , eventually biased by foregrounds (section 3.3);
4. we use the measured r_{fgs} to generate 300 Gaussian realisations of CMB B -mode polarisation maps sourced by primordial GWs only, $C_{\ell}^{BB} = r_{\text{fgs}} \cdot C_{\ell}^{r=1}$ (section 3.4.1);

² *LiteBIRD*: Lite (Light) satellite for the study of B -mode polarization and Inflation from cosmic background Radiation Detection.

5. using the power spectrum of the foreground residuals map derived from component separation (step 2), we generate a set of 300 realisations of Gaussian B -mode maps from it (section 3.4.2);

6. we compute MFs of the following sets of maps:

- **FGres**: co-addition of CMB with $r = 0$, noise residuals, non-Gaussian foreground residuals (i.e., the CMB solution returned by the component separation, step 2);
- **CMBtens**: co-addition of CMB with $r = r_{\text{fgs}}$ (step 4) and noise residuals after component separation;
- **FGgauss**: co-addition of CMB with $r = 0$, noise residuals after component separation and Gaussian foreground residuals (step 5);

all datasets include the input realisations of the lensing signal in the CMB component;

7. we run a statistical analysis to assess if MFs are able to robustly detect deviations between the cleaned CMB map and the Gaussian simulations;

This procedure is repeated for 300 realisations of the input multi-frequency maps, and applied to different combination of component separation pipeline and foregrounds model.

3.1 Input simulations

In this section we describe the sky simulations that we use as input for our pipeline (step 1), while the CMB tensor and Gaussian foregrounds simulations outlined in steps 4 and 5 are described in section 3.4.1 and 3.4.2, respectively. Maps generation, manipulation, and analysis is carried out adopting the HEALPix³ [58] pixelisation scheme by means of the healpy⁴ package [59]. For all our simulations, the pixel resolution is defined by $N_{\text{side}} = 64$, corresponding to an angular resolution of about 55 arcmin, with the maximum multipole considered being $\ell_{\text{max}} = 3N_{\text{side}} - 1 = 191$.

CMB maps are generated as Gaussian realisations from the angular power spectra of the best-fitting *Planck* 2018 parameters [2]; we produce $N = 300$ simulations where B -modes are sourced by gravitational lensing only, with no primordial signal ($r = 0$). The adopted *LiteBIRD* specifications are taken from the latest forecast of the experiment (table 13 in [14]), planned to cover 15 frequency bands with 22 effective channels between 40 and 402 GHz. All CMB maps are smoothed to a common angular resolution of $\text{FWHM}_{\text{out}} = 70.5$ arcmin, which is the largest beam considered (corresponding to the 40 GHz channel).

Noise maps are simulated by generating Gaussian realisations of white noise at the different frequency channels; noise maps are smoothed to a common resolution by considering an effective beam as

$$\text{FWHM}_{\text{eff}}(\nu) = \sqrt{\text{FWHM}_{\text{out}}^2 - \text{FWHM}(\nu)^2}. \quad (3.1)$$

To simulate Galactic foregrounds, we use the Python Sky Model (PySM)⁵ [60–62], a python package that generates full-sky simulations of Galactic foregrounds in intensity and

³<https://healpix.jpl.nasa.gov/>.

⁴<https://healpy.readthedocs.io/en/latest/>.

⁵<https://pysm3.readthedocs.io/en/latest/index.html>.

polarisation relevant for CMB experiments. Since we are focusing on polarisation, we are only interested in thermal dust (**d**) and synchrotron (**s**) emission, as they are the main contaminants of observations.

In this work we adopt three different models for synchrotron and thermal dust emission, of increasing complexity: **d0s0**, **d1s1**, and **d10s5**. The former one is a simplistic view of the sky that we use for testing and validation of our pipeline, while the others are models closer to realistic observations. We refer to the **PySM** papers and documentation for more details.

We generate Q/U full-sky simulations for each of the three models, one per frequency channels; each map is smoothed to 70.5 arcmin and then co-added to the CMB and noise simulations. We end up with sky simulations which include CMB, noise, and foregrounds, with 300 different realisations of CMB and noise for each of the three considered **PySM** models.

We then perform a full-sky harmonic transformation to convert Q/U simulations to B -mode maps, which will be our input dataset for all the subsequent analysis pipeline. We are aware that the non-locality of the $Q/U \rightarrow E/B$ transformation is creating artifacts in the final B -mode maps, due to the leakage of sky regions with very bright foregrounds radiation outside the Galactic plane. A more correct approach would be to first mask the most contaminated areas of the Q/U maps and then operate the transformation to B -modes, while correcting for the E -to- B leakage, if needed, due to observing a partial sky (see [41, 63] for more details and a MFs analysis of the problem). In the following we will not account for this issue, placing ourselves in the worst possible scenario, with no mitigation of this effect in the maps. The impact of this assumption will be discussed in the presentation of the results.

3.2 Component separation

As mentioned in section 1, in this work we adopt minimum-variance algorithms as our component separation methods to recover CMB B -modes from multi-frequency simulations of the sky. In particular, we use two implementations: Needlet Internal Linear Combination (NILC) and Multi-Clustering NILC (MC-NILC). NILC has already been applied in the past for CMB data analysis (e.g., *WMAP* [64], *Planck* [27]) and both NILC and MC-NILC will be used for current and future CMB surveys (*SO* [16, 65], *LiteBIRD* [14]).

NILC belongs to the category of blind component separation methods, since it aims at reconstructing the CMB signal without any assumptions on the foreground spectral properties. It thus represents a valuable alternative to parametric approaches, as it is not affected by spectral mis-modelling of the Galactic polarised emission, which may significantly bias the final estimate of the tensor-to-scalar ratio.

In a framework with complex foreground emission, the NILC pipeline may present some limitations or be suboptimal, since it performs simple local variance minimisation across the sky, without fully handling the local spectral variations of foregrounds. MC-NILC [38] is a NILC extension which performs component separation independently in different sky regions, where the polarised B -mode Galactic emission shows similar spectral properties. The spatial variability of dust and synchrotron spectral parameters is assessed by identifying a blind tracer of their distribution across the sky. This tracer is obtained from the simulated dataset by applying a model-independent and realistic de-noising technique [38].

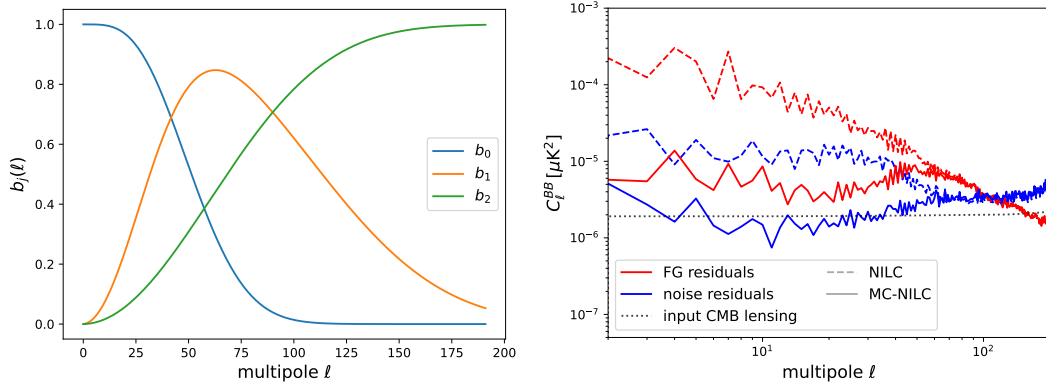


Figure 2. *Left panel:* needlelet bands configuration adopted in this work, where the first bands have been merged. *Right panel:* full-sky angular power spectra evaluated on foreground (red) and noise (blue) residuals, for NILC (dashed lines) and MC-NILC (solid lines), applied to the **d10s5** model. Black dotted line is the input CMB spectrum (lensing-only).

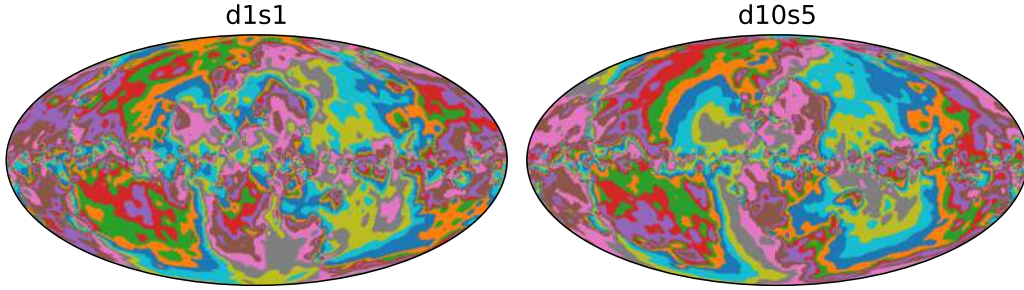


Figure 3. Sky partitions adopted within MC-NILC component separation for **d1s1** (*left*) and **d10s5** (*right*) foreground models.

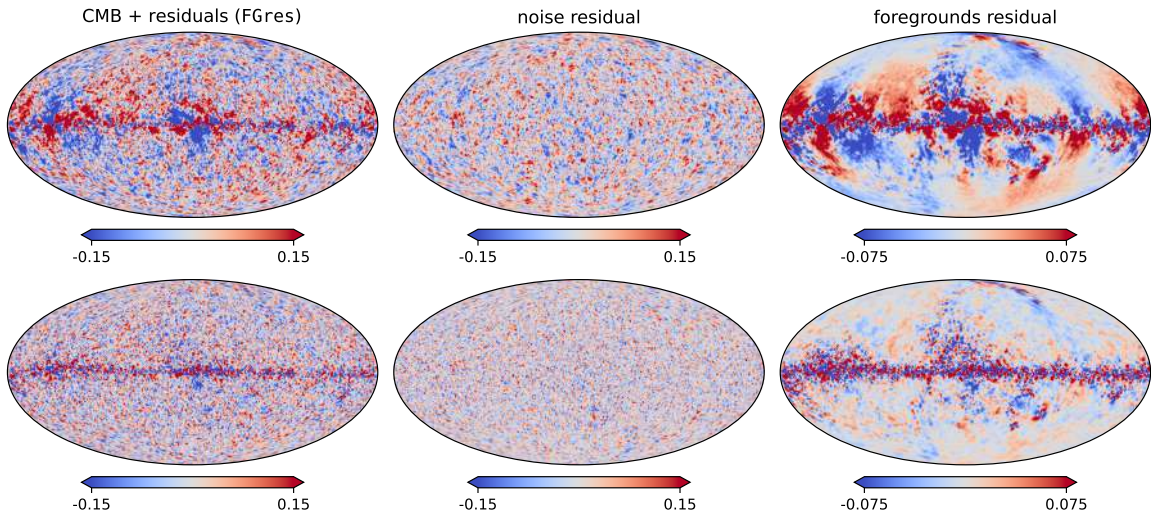


Figure 4. Component separation outputs, for one simulation. *Left column:* CMB solution. *Middle column:* noise residuals. *Right column:* foreground residuals. Outputs are shown for NILC (upper row) and MC-NILC (lower row) methods, applied on the **d10s5** model. Units are μK_{CMB} .

Both algorithms relies on the notion of needlets. Needlets are a wavelet system that guarantees simultaneous localisation in harmonic and pixel space; they have been introduced in the statistical literature in [66] and first applied to CMB data in [67]. Given any field of spin s defined on the sphere, the corresponding needlet maps β_j^s are obtained by filtering its harmonic coefficients $a_{\ell m}^s$ with a harmonic weighting function b , which isolates modes at different angular scales for each needlet scale j [68]:

$$\beta_j^s(\hat{n}) = \sum_{\ell, m} \left[a_{\ell m}^s \cdot b\left(\frac{\ell}{B^j}\right) \right] {}_sY_{\ell m}(\hat{n}) \quad (3.2)$$

where $\hat{n}$ is a direction in the sky and ${}_sY_{\ell m}$ are the spin-weighted spherical harmonics. This procedure in harmonic space is equivalent to performing a map convolution in real domain. The shape of the needlet bands is defined by the choice of the harmonic function b , whose width is set by the parameter B : lower values of B correspond to a tighter localisation in harmonic space (less multipoles entering into any needlet coefficient). We adopt the Mexican needlet construction [66] as our harmonic function, with $B = 1.3$ (figure 2, left panel). The needlet filters have been obtained through the MTNeedlet⁶ [69] package.

The adopted MC-NILC sky partitions are illustrated in figure 3, while an example of the performances of the two algorithms is reported in figures 2 (right panel) and 4. The amount of leftover contamination can be visually compared both at the map and at the power spectrum level, and the improvement achieved by MC-NILC can be clearly seen in the reduction of the noise and foreground residuals.

Usually a component separation algorithm returns a single CMB “solution” as final product, containing the cleaned CMB signal and some contamination from noise and foregrounds. In a realistic experiment, these residuals maps are not accessible. Since in this analysis we are working with simulated datasets, we have complete control over the separated sky components, as we independently generated CMB, noise, and foregrounds realisations. Component separation weights can be applied to the input noise and foregrounds maps to obtain residuals maps, useful for pipeline validation and for the characterisation of the MFs robustness test.

3.3 Likelihood

In order to infer the tensor-to-scalar ratio from the component separated maps, we adopt an inverse-Wishart distribution as our likelihood [70, 71]:

$$-\ln \mathcal{L} \left(C_\ell^{\text{obs}} | r \right) = \frac{1}{2} \sum_{\ell} (2\ell + 1) f_{\text{sky}} \left[\frac{C_\ell^{\text{obs}}}{C_\ell^{\text{th}}(r)} + \ln C_\ell^{\text{th}}(r) - \frac{2\ell - 1}{2\ell + 1} \ln C_\ell^{\text{obs}} \right], \quad (3.3)$$

where C_ℓ^{obs} is the “observed” power spectrum evaluated on the CMB solution map, containing lensing, noise and foreground residuals, and C_ℓ^{th} is the theoretical BB spectrum built as

$$C_\ell^{\text{th}}(r) = C_\ell^{\text{lensing}} + \langle C_\ell^{\text{nres}} \rangle + r \cdot C_\ell^{r=1}. \quad (3.4)$$

In the above equation, C_ℓ^{lensing} is the B -mode spectrum induced by gravitational lensing, $C_\ell^{r=1}$ is the theoretical B -mode power spectrum sourced only by tensor perturbations with

⁶<https://javicarron.github.io/mtneedlet/index.html>.

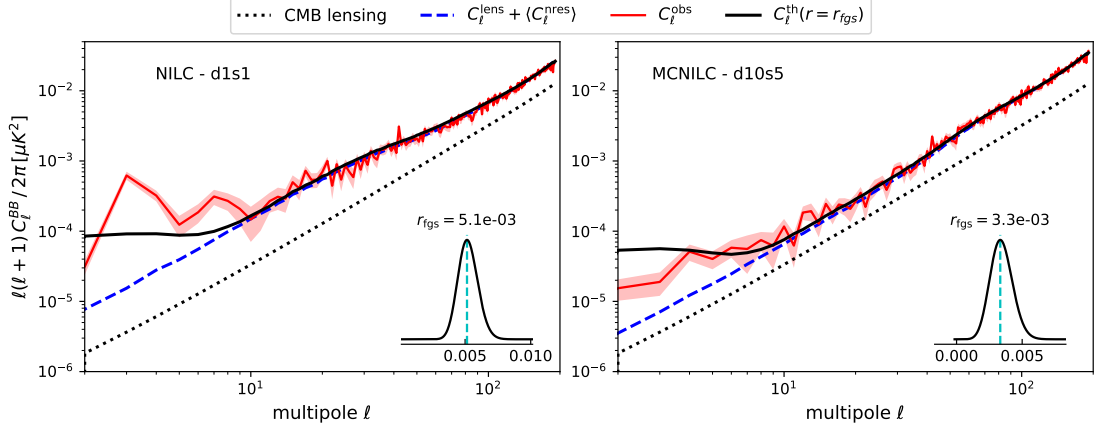


Figure 5. Output of the likelihood evaluation for a single simulation, for the NILC - d1s1 (left) and MC-NILC - d10s5 (right) scenarios, considering all the multipoles. *Main plot:* power spectra, with $C_\ell^{th}(r=0)$ in dashed blue, C_ℓ^{obs} in red solid, $C_\ell^{lensing}$ in dotted black, and best fit spectrum in black solid. The shaded area represents the standard deviation of C_ℓ^{obs} across 300 realisations. *Inset:* normalised likelihood on the tensor-to-scalar ratio, with the peak value indicated by the dashed line.

$r = 1$, and $\langle C_\ell^{nres} \rangle$ is the average of the spectra computed on all the other $N - 1 = 299$ noise residuals maps obtained from the component separation.

Angular power spectra are computed with the **anafast** routine as implemented in the **healpy** package [58, 59]. To avoid strong contamination from the Galactic plane, a simple mask⁷ is applied to the maps, retaining a sky fraction of $f_{sky} \simeq 60\%$. As it is well-known, estimating power spectra with **anafast** does not account for correlations among multipoles induced by the presence of a mask (usually referred as mode-coupling [72]). However, as discussed in [14], this approximation has a negligible impact on the power spectra estimation of foreground and noise residuals over large sky fractions, as considered in this work.

We evaluate the likelihood function over an interval of r values in the range $r \in [-1.5 \times 10^{-4}, 0.03]$, with a constant spacing of $\Delta r = 2 \times 10^{-6}$. Values of $r < 0$ are included in the analysis in order to be sensitive to negative biases on the tensor-to-scalar ratio, as our generated CMB simulations assumes $r = 0$. We take the peak value of the likelihood as our best fit tensor-to-scalar ratio r_{fgs} , where the **fgs** label indicates that the bias on the parameter is entirely due to the presence of foreground residuals, which contribute to the power observed on the large scales and mimic the behaviour of a primordial GWs signal.

The inferred value of the tensor-to-scalar ratio gathers contributions from all the angular scales included in the analysis. In general, the largest scales (low ℓ) contain most of the information about foregrounds and inflation, while the smaller scales (high ℓ) are dominated by noise and CMB lensing, having therefore a lower impact on the r measurement. To highlight the different contributions, we also evaluate the likelihood function over three multipoles range: $\Delta\ell_0 \in [2, 60]$, $\Delta\ell_1 \in [15, 175]$, $\Delta\ell_2 \in [40, 191]$. These intervals follow the scales sampled by the needlet bands used for component separation and MFs computation, including all multipoles where each of the adopted needlet bands is greater than 30%. We

⁷Available at the [Planck Legacy Archive](#).

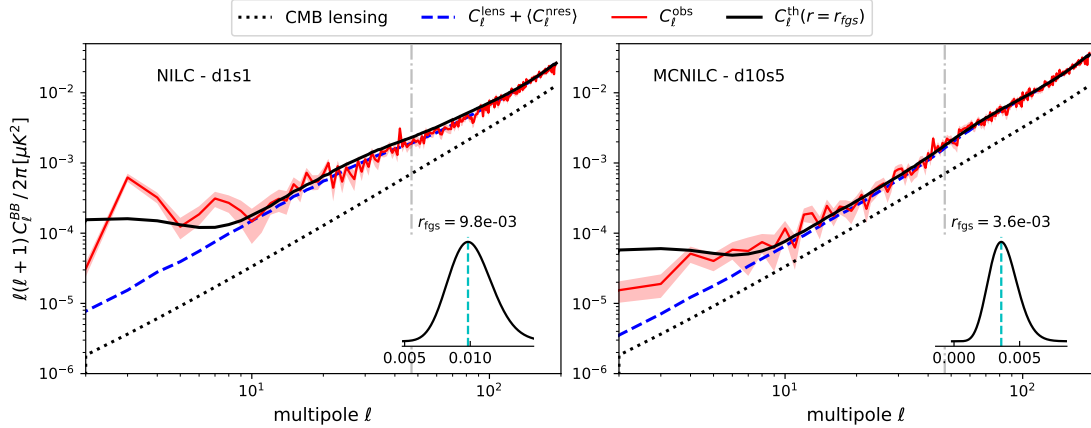


Figure 6. Same as figure 5, but considering only the largest scales in the likelihood evaluation. The maximum multipole included is indicated by the vertical line.

will focus on the first interval, as we are mainly interested on the impact of foreground residuals on the largest angular scales. An example of the likelihood inputs and outputs is provided in figure 5, where we report the output and best-fit tensor B -mode angular power spectra together with the corresponding posterior of r_{fgs} . Figure 6 instead outlines, for the same cases, the best-fit tensor spectra obtained when only the largest scales ($\Delta\ell_0$) are considered in the likelihood analysis.

Following the procedure adopted in a realistic data analysis, we evaluate the goodness-of-fit achieved to understand if we are (already) able to assess whether the bias on r is induced by the presence of foreground residuals. We note that eq. (3.3) is not a Gaussian likelihood, and thus it is not possible to define a χ^2 statistic in the usual way. However, we can still evaluate the fit quality by relying on simulations, as follows (for each simulation):

1. we use the inferred tensor-to-scalar ratio r_{fgs} to generate 300 Gaussian random realisations of CMB B -mode maps ($C_\ell^{BB} = r_{\text{fgs}} \cdot C_\ell^{r=1}$) and we co-add them with simulations of CMB lensing and noise residuals (provided by the component separation). Doing so means taking the CMB solution and replacing the foreground residuals contribution with realisations of tensor perturbations;
2. for each of these combined CMB lensing, noise, and tensor perturbation maps we compute their power spectrum and we fit for r with the same likelihood function of eq. (3.3). We take each minimum of the $-\ln\mathcal{L}$ to build a distribution which we can use to compare with the value obtained by the corresponding best fit r_{fgs} ;
3. we compute the probability-to-exceed (PTE), that is the amount of the distribution which falls above the value correspondent to the best fit r_{fgs} . If the PTE is less than 0.05 (95%), we deem the fit as not good for the considered simulation.

We repeat this analysis for all the 300 input simulations. In general, we expect the PTE distribution to be uniform in order to have a reasonable goodness-of-fit, with the percentage of incompatible simulations that should be $\sim 5\%$. The results of this analysis are shown in table 1, while the PTE distribution for the 300 simulations can be found in appendix A.

all multipoles ($\Delta\ell=2-191$)			
	d0s0	d1s1	d10s5
NILC	3.7	97.7	74.0
MC-NILC		28.0	21.7

large-scales only ($\Delta\ell_0=2-60$)			
	d0s0	d1s1	d10s5
NILC	4.0	98.0	71.0
MC-NILC		51.7	42.7

Table 1. Percentage of incompatible simulations from the goodness-of-fit analysis, considering all scales at the same time or only the largest scales. Values are reported for each scenario considered.

As it can be seen from the table, in almost all the cases we do not have a good fit quality, with PTE distributions far from being uniform. This is suggesting that the shape of the power spectra at the largest scale do not match enough for the fit to be acceptable: in a realistic scenario, we would likely get a $\text{PTE} < 0.05$, already ruling out the cosmological origin of the inferred tensor-to-scalar ratio. In this analysis we are restricting ourselves to specific foreground models, component separation pipelines, and instrumental configurations, which may not represent the most general case. To illustrate a broader scenario, we also consider a situation where no spectral mismatch is present and only higher-order statistics can reveal discrepancies between residuals and primordial cosmological signal. We will further describe this point in section 3.4.2.

3.4 Gaussian simulations for MFs comparison

The aim of the whole analysis is to compare MFs evaluated on the CMB solution (**FGres**) against those computed on two different sets of Gaussian simulation: tensor perturbations (**CMBtens**) and foreground residuals (**FGgauss**). These Gaussian simulations are co-added with CMB lensing and noise residuals maps, in order to make a comparison with the MFs of **FGres** maps. Therefore, the compared maps only differ for the foreground residuals part, which is replaced by either tensor perturbations or Gaussian foreground residuals. In this section we describe how these two sets of Gaussian simulations are generated. An example for one simulation of these two types of maps and their co-addition with CMB lensing and noise residuals is reported in figures 7 and 8.

3.4.1 Tensor perturbations (**CMBtens**)

After the likelihood evaluation on the component-separated maps, we obtain $N = 300$ values (one per simulation) of the best-fit tensor-to-scalar ratio r_{fgs} , biased by foreground residuals. Since our ultimate goal is to assess the MFs capability to discriminate between foreground contamination and a cosmological signal, we generate a set of Gaussian realisations of tensor perturbations and co-add them with simulated CMB lensing and noise residuals maps. Specifically, for each inferred r_{fgs} , we generate other 300 realisations of B -mode maps (i.e., 300×300 simulations in total) sourced only by primordial GWs ($C_\ell^{BB} = r_{\text{fgs}} \cdot C_\ell^{r=1}$), smoothed

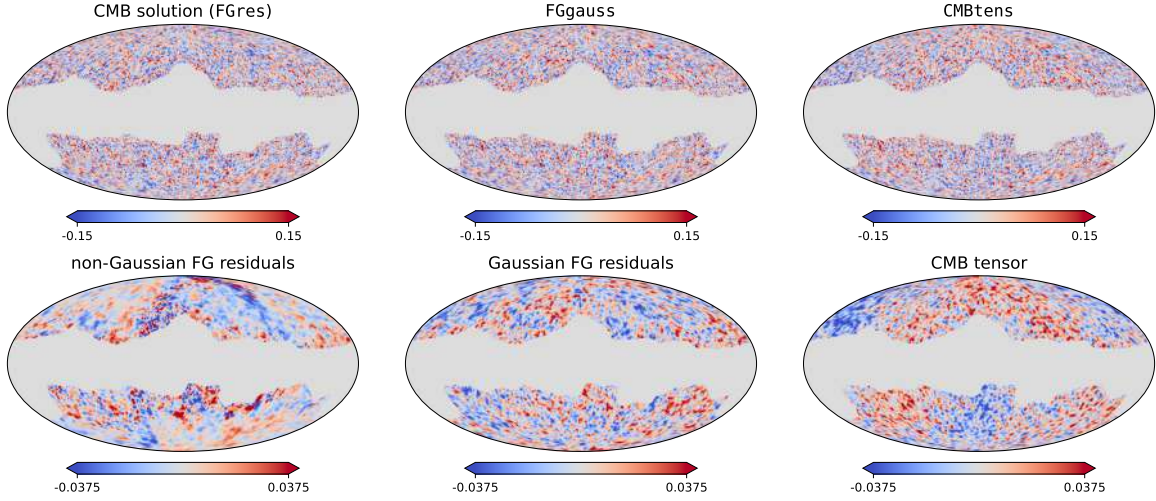


Figure 7. Example of maps of which MFs are computed, for the MC-NILC - d10s5 case, considering all the scales. *Upper row:* CMB solution (FGres) on the left, Gaussian simulations (FGgauss, CMBtens) in the middle / on the right. *Lower row:* isolated contributions, without CMB lensing and noise, with non-Gaussian (left) and Gaussian (middle) foreground residuals, and tensor perturbations (right).

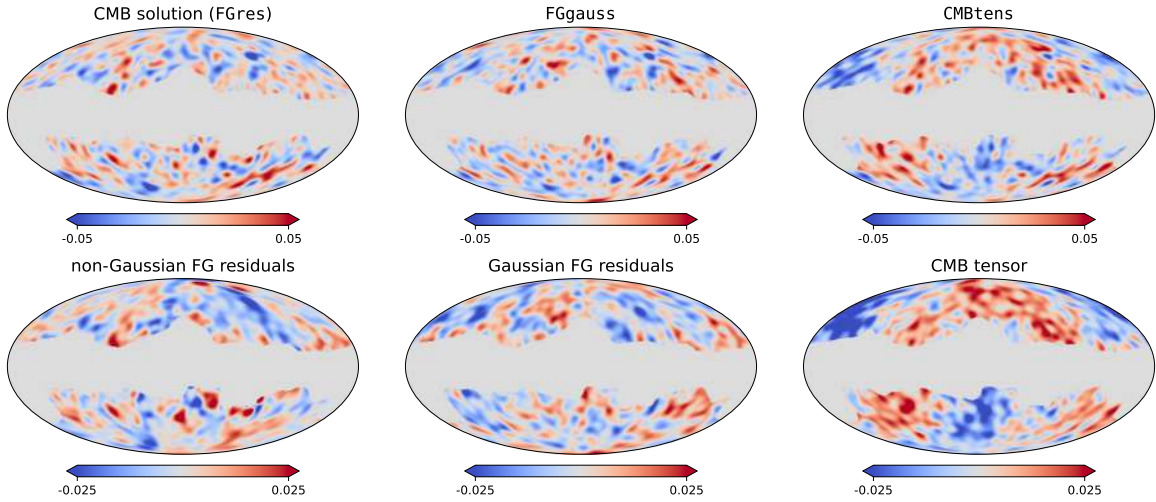


Figure 8. Same as figure 7, but in needlet space (i.e., considering separate scales), for the MC-NILC - d10s5 scenario in the b_0 needlet band (multipoles $\ell \lesssim 60$).

to the same beam resolution of the component separation outputs. If the detected tensor-to-scalar ratio is of astrophysical origin, then the CMB solution (that includes anisotropic contamination) and the tensor simulations (Gaussian) should be statistically incompatible when observed through a higher-order statistics like MFs, failing the robustness test. Vice versa, if no deviation is observed between two maps, it means that MFs are not able to detect the level of non-Gaussian contribution coming from the foreground residuals. In the latter case, the test is passed and no flag is raised. The complete statistical compatibility analysis is outlined in section 3.5.

3.4.2 Foreground residuals (FGgauss)

The noise and foreground residuals maps provided by component separation represent the level of contamination left in a cleaned CMB map by the adopted algorithm. The right column of figure 4 shows how foreground residuals display features of Galactic origin, with evident anisotropic structures over the sky. These structures are not so obvious when foreground residuals are mixed with CMB and noise (left column of figure 4). The contamination is stronger along the Galactic plane (always masked in data analysis), while at higher latitudes the foregrounds anisotropy is not strikingly evident.

In section 3.3 we fit the power spectrum of the CMB solution, measuring a biased tensor-to-scalar ratio. We find that the shapes of the power spectra — foregrounds against tensor perturbations — do not match, with a bad fit quality (table 1). This introduces a problem in the interpretation of MFs results: as outlined in section 2, their expected values depend both on the map statistics and on the spectral features through the C_ℓ term in eq. (2.4). Two maps with same statistics but different spectra will have different MFs values.

We are mainly interested in using MFs to detect deviations from the Gaussian statistics, since a difference in the power spectrum shape would already be detected by a poor goodness-of-fit in the likelihood analysis. To disentangle the two effects, we also evaluate MFs on the FGgauss maps, built by replacing the non-Gaussian foreground residuals of the FGres maps with 300 Gaussian realisation of it. These simulations are produced by taking the power spectrum of foreground residuals outside the adopted mask, and using it to generate Gaussian realisations with the `synfast` routine from `healpy`. This means that we are treating the spectrum of a single realisation as the correct underlying theoretical spectrum, with the average spectrum of the Gaussian simulations converging to the non-Gaussian one. In this way, in the comparison between FGres and FGgauss, MFs are evaluated on sky maps with the same “observed” power spectrum but different statistics, providing a critical benchmark to evaluate the MFs ability in discriminating the two fields. In a realistic experiment, this would translate to an optimal fit in the likelihood step, reproducing a situation where we obtain a biased estimate of the tensor-to-scalar ratio with a good fit quality. This approach also helps in alleviating limitations induced by some assumptions we made in our analysis, such as: *i*) the adoption of a simplified sky mask, which is sub-optimal for mitigating foregrounds contamination; *ii*) the assumption of an inverse-Wishart form for the likelihood, which does not strictly hold on a masked sky, and *iii*) the poor goodness-of-fit in the inference of the tensor-to-scalar ratio. The first two assumptions may contribute to the bias on r , and a mask or likelihood optimisation would therefore reduce the bias on the parameter.

We stress that in a realistic situation we do not have access to the component separation residuals and, consequently, to the possibility of generating this kind of simulations. Nonetheless, it is useful to explore the MFs detection power in an idealised scenario where the contribution from non-Gaussianity is isolated, in order to understand if this higher-order statistic can be a reliable tool in validating a tensor-to-scalar ratio potential detection.

3.5 Compatibility analysis

To establish the compatibility between MFs computed on different maps, we follow an approach similar to the one adopted in [48], where a thorough investigation of the statistical

properties of CMB anisotropies was conducted on real *Planck* data, comparing the MFs of a single component-separated map against those of a set of Gaussian simulations. Here we apply the same analysis, but on a simulated dataset (“data” realisation) and repeated for $N = 300$ times. We will then provide an estimation of the MFs compatibility test over this sample of simulations. For the remainder of this section, we will use “Gaussian simulations” to refer either to tensor (**CMBtens**) or Gaussian foreground residuals (**FGgauss**) simulations, as we carry out the compatibility analysis in the same way for the two sets.

MFs are evaluated for 12 thresholds u between -3 and 3 in σ units, with maps normalised by their variance. To quantify the level of deviation between MFs of different maps, χ^2 values are computed assuming a Gaussian likelihood for the MFs at every threshold:

$$\chi_s^2 = \left[y_s - \bar{y}_s^{\text{sim}} \right]^T \Sigma_s^{-1} \left[y_s - \bar{y}_s^{\text{sim}} \right], \quad s = 1, \dots, 300, \quad (3.5)$$

where y_s represents one of the three MFs ($y = v_0, v_1, v_2$) of the s -th **FGres** realisation, and $\bar{y}_s^{\text{sim}} \equiv \langle y_s^{\text{sim}} \rangle$ is the mean of the MFs computed on the 300 Gaussian simulations generated for that particular s -th “data” simulation. We recall that we have $N = 300$ realisations of **FGres** maps, and for each one of them we generate other 300 simulations of **CMBtens** or **FGgauss** maps. In the first comparison, MFs can detect deviations due to both the shape of the power spectrum and the non-Gaussianity, while in the second one the discrepancy is entirely caused by the presence of non-Gaussianity.

Correlations between different thresholds [51] are taken into account using the appropriate covariance matrix Σ_s , different for each “data” realisation. An unbiased estimator of the inverse of the covariance is given by

$$\Sigma_s^{-1} = \frac{N - d - 2}{N - 1} \mathbf{C}_s^{-1}, \quad (3.6)$$

where $N = 300$ is the number of simulations, $d = 12$ is the number of thresholds (degrees of freedom), and $\mathbf{C}_s$ is the sample covariance computed from the N Gaussian simulations (tensor or foreground residuals):

$$\mathbf{C} = C_{ij} \equiv \langle (y_i^{\text{sim}} - \bar{y}_i^{\text{sim}})(y_j^{\text{sim}} - \bar{y}_j^{\text{sim}}) \rangle, \quad (3.7)$$

with i, j referring to MFs computed at different thresholds. The pre-factor in eq. (3.6) is known as the Hartlap factor [73] and it is needed for the estimator to be unbiased.

For each “data” realisation ($\chi_{\text{data}}^2 = \chi_s^2$), we compute its probability-to-exceed (PTE), which is the amount of the χ^2 distribution of the MFs computed on Gaussian simulation exceeding the observed χ_{data}^2 realisation:

$$\text{PTE}_s = \chi_{\text{sims}}^2 > \chi_{\text{data}}^2. \quad (3.8)$$

We assign a detection (i.e., MFs are not compatible) to the s -th simulation when $\text{PTE}_s < 0.05$, for at least one of the MFs. A detection means that statistically significant deviations are observed between MFs computed on a map contaminated by non-Gaussian residuals and the mean MFs of Gaussian realisations with similar power in tensor or foreground residuals. Obtaining many detections over a large number of simulations indicates that MFs are a reliable robustness test, while a small amount of detection would suggest that MFs are not able, most of the times, to distinguish between Gaussian and non-Gaussian contributions.

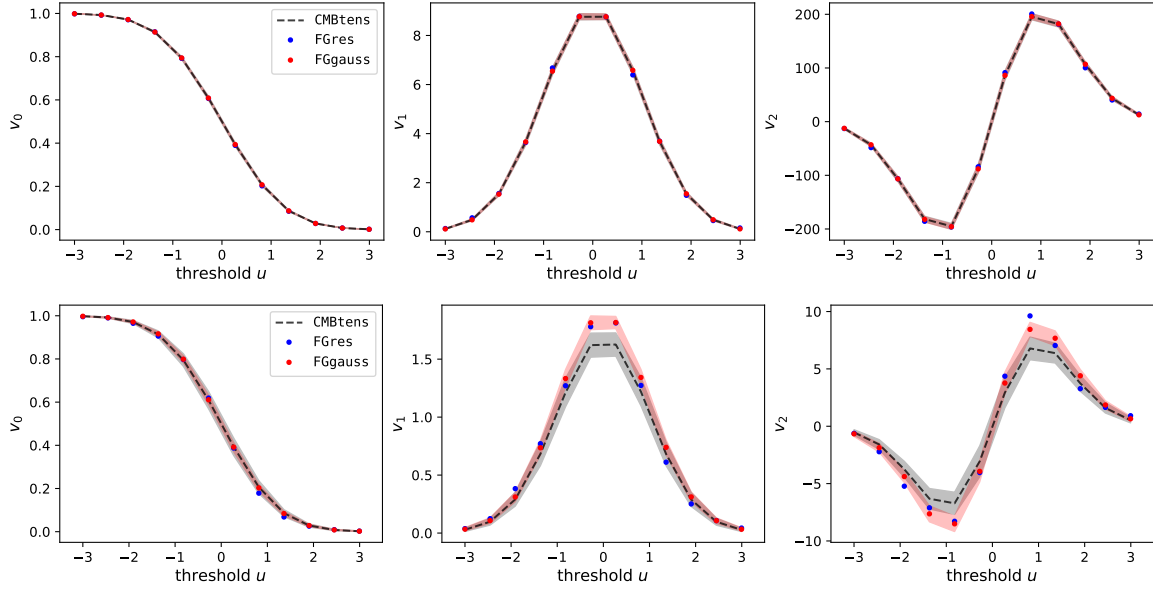


Figure 9. Minkowski functionals, for a single simulation, in the MC-NILC - **d10s5** case. MFs of the CMB solution (**FGres**) are shown with blue points; the black dashed line is the average MFs from 300 Gaussian realisations of tensors (**CMBtens**), and red points are the average MFs from 300 Gaussian realisations of foreground residuals (**FGgauss**). The correspondent gray and red shaded areas represent the standard deviation of the MFs across the simulations. *Upper row*: real space (including all scales). *Lower row*: needlet space (first band b_0 , $\ell \lesssim 60$).

We define the test efficiency $\eta(A, B)$ between two sets of data A and B as the percentage of detection we get over all the simulations,

$$\eta(A, B) = \frac{\# [\text{detections}(A, B)]}{N} = \frac{\# [\text{PTE} < 0.05]}{N}, \quad (3.9)$$

where $\#[\cdot]$ stands for the number of simulations satisfying the condition in brackets. We will report this number to quantify the detection power of the MFs test, for each combination of component separation and foregrounds model.

4 Results and discussion

In this section we report and discuss the results of this work. Tables 2 and 3 offer a comprehensive report of the results for the scenarios considered in the analysis, while figure 9 shows an example of MFs comparison for the MC-NILC - **d10s5** case as we consider it as our most realistic scenario, since **d10s5** is the model closer to observations and MC-NILC is the state-of-the-art minimum-variance component separation algorithm which will be used in future CMB experiments. In the appendix A we report examples for the NILC - **d1s1** scenario, which represents a case with more significant foreground residuals on the large scales.

We report our results in terms of the compatibility level between MFs of different maps. The efficiency of the test (percentage of incompatible simulations, eq. (3.9)) is the main summary statistic of the study, reported for all combinations of component separation technique and foregrounds modelling, in pixel and needlet space. $\eta(\text{FGres}, \text{CMBtens})$ indicates

no delensing ($A_{\text{lens}} = 1$)				
	A, B	d0s0	d1s1	d10s5
NILC	FGres, CMBtens	12.7	95.3	88.3
	FGres, FGgauss	11.0	39.3	26.7
MC-NILC	FGres, CMBtens		13.3	15.3
	FGres, FGgauss		10.3	12.7

50% delensing ($A_{\text{lens}} = 0.50$)				
	A, B	d0s0	d1s1	d10s5
NILC	FGres, CMBtens	10.3	100	95.3
	FGres, FGgauss	11.3	42.7	27.7
MC-NILC	FGres, CMBtens		14.0	20.3
	FGres, FGgauss		15.0	24.0

Table 2. Efficiency η (in %) of the MFs test on polarisation B -mode maps, in pixel space (i.e., considering all scales at the same time), both for no delensing ($A_{\text{lens}} = 1.0$) or 50% delensing ($A_{\text{lens}} = 0.5$) applied. Values are reported for each scenario considered in the analysis. $\eta(A, B)$ indicates the efficiency computed by comparing the dataset A against the reference dataset B .

the values computed by comparing MFs against those of Gaussian simulations of tensor perturbations, while $\eta(\text{FGres}, \text{FGgauss})$ refers to the comparison with Gaussian simulations of foreground residuals (both co-added with CMB lensing and noise residuals).

Figure 9 shows an example for one simulation of the computed MFs in pixel and needlet space (first band b_0), reporting MFs of the single CMB solution (blue points) along with the average ones of the CMBtens (dashed black line) and FGgauss (red points) simulations. We also show, in figure 10 and for all the 300 simulations, the MFs relative difference in units of the standard deviation of the simulations $\Delta v / \sigma = (y - \bar{y}^{\text{sim}}) / \sigma$; each line corresponds to the difference between one CMB solution realisation and the means of the two sets of Gaussian simulations, blue for CMBtens and orange for FGgauss simulations.

Concerning the real-space analysis, we find a high efficiency of the test when comparing MFs against Gaussian tensor simulations, while using NILC on the d1s1 and d10s5 models, with $\eta(\text{FGres}, \text{CMBtens}) \simeq 95\%$ and 88% respectively. This means that over 300 simulations, MFs are able to detect the discrepancy due to foregrounds contamination in 95% (or 88%) of the cases. However, this deviation is mainly driven by the large differences at the power spectrum level between the two maps: as discussed in section 2, the MFs expected values depend through eq. (2.4) both on the level of Gaussianity in the map, and on the shape of the power spectrum (included in the μ term). MFs computed on maps with mismatching power spectrum features will deviate between each other, with discrepancies sourced by both contributions. As already pointed out in section 3.4.2, the large-scale shapes of the power spectra do not match, for the two scenarios discussed in this paragraph (for comparison, the percentages of simulations providing a bad fit are $\sim 98\%$ for d1s1 and 74% for d10s5). Hence, one could already rule out the cosmological origin of the signal by evaluating the

no delensing ($A_{\text{lens}} = 1$)				
	A, B	d0s0	d1s1	d10s5
NILC	FGres, CMBtens	10.7	93.0	51.7
	FGres, FGgauss	11.7	72.7	23.3
MC-NILC	FGres, CMBtens		22.3	26.7
	FGres, FGgauss		16.0	15.0

50% delensing ($A_{\text{lens}} = 0.50$)				
	A, B	d0s0	d1s1	d10s5
NILC	FGres, CMBtens	11.7	95.7	57.0
	FGres, FGgauss	11.0	75.3	20.0
MC-NILC	FGres, CMBtens		24.3	42.0
	FGres, FGgauss		18.3	18.0

Table 3. Same as table 2, but in needlet space, for the first band b_0 (largest scales, $\ell \lesssim 60$).

goodness-of-fit, and the hypothesis would then be confirmed by the MFs test; we notice that in some cases, such as NILC - d10s5, we find that the MFs compatibility test is more powerful than the power spectrum goodness of fit.

The same conclusion can also be drawn from the lower value of the efficiency when Gaussian foreground residuals simulations are considered. We find $\eta(\text{FGres}, \text{FGgauss}) \simeq 39\%$ and $\simeq 27\%$ for d1s1 and d10s5, respectively, when NILC is applied. This amount of detection is entirely due to the different statistics in one component of the two datasets, since the corresponding power spectra are equivalent. This interpretation of the results is also confirmed by the detection efficiencies obtained when applying MC-NILC, where the large-scale features of the power spectra are more similar: the number of detections is relatively low, and close to what we find in the d0s0 scenario, where the tensor-to-scalar ratio is not biased and for which we do not expect a statistically significant detection.

In the needlet-based analysis, we compute MFs on filtered maps adopting three needlet bands which sample separate range of multipoles; here we focus on the first band b_0 which is the most relevant for our analysis ($\ell \lesssim 60$), as we expect the most relevant sources of bias to the tensor-to-scalar ratio to be in this multipole range. In this way, different contributions from different sky components should be isolated, in particular on the largest scales where foreground residuals dominate over the CMB and noise residuals components; the MFs detection power is expected to be boosted in the first needlet band as the higher multipoles noise contribution is filtered out, thus removing a Gaussian component at the map level and enhancing non-Gaussian features typical of foreground residuals. In general, we observe a slight increase in the detection number for some of the cases (e.g., for MC-NILC), while in others the efficiency remains the same or is even reduced. This is due to the fact that the tensor-to-scalar ratio used for generating the simulations is fitted by considering only the multipoles included in the needlet band, providing a better fit to the power spectrum and

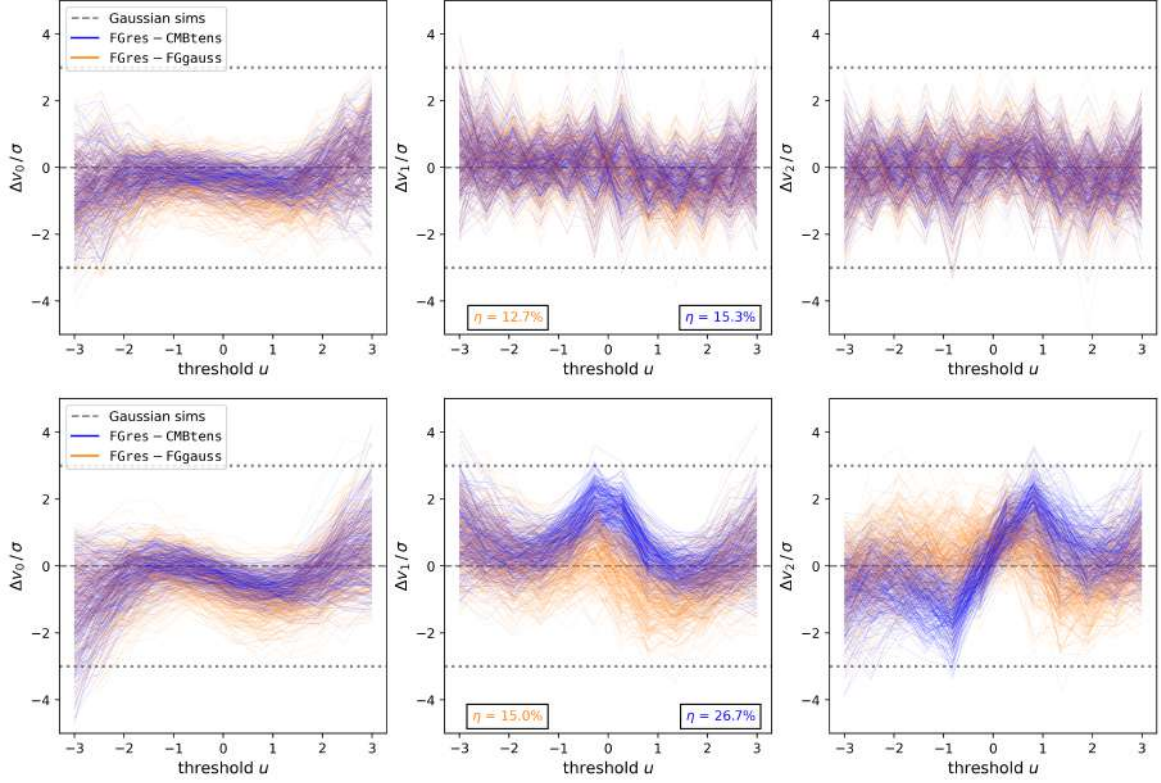


Figure 10. MFs weighted difference for the MC-NILC - d10s5 scenario, referring to FGres against CMBtens maps (blue lines) and FGres against FGgauss maps (orange lines). See text for details. Plots are shown for pixel (upper row) and needlet (band b_0 , lower row) analysis.

thus reducing deviations arising from the different spectral shapes. The detection number still remains considerably low and close to the “no detection” level typical of the d0s0 case.

We can explain why MFs can not clearly detect non-Gaussianity by looking at the maps shown in figure 7 and 8: the Gaussian component (due to CMB lensing and noise residuals) is dominating the structures of the maps, hiding the lower contribution coming from non-Gaussian foreground residuals. This suggests that the level of foreground residuals is too low with respect to the noise residuals in order to be detected by MFs. Indeed, in the NILC - d1s1 case foreground residuals are relatively larger than noise residuals with respect to other scenarios, and we register a higher number of detections.

This can be straightforwardly tested by repeating the analysis after removing CMB and noise contributions, looking at discrepancies between the non-Gaussian (lower left map of figure 7) and Gaussian (lower middle map) foreground residuals; in this case, the efficiency of the MFs test is always 100%, meaning that MFs are perfectly able to distinguish the non-Gaussianity. Once added, CMB lensing and noise residuals become the driving contribution in the MFs and tend to overwhelm the foregrounds, hindering the MFs detection power.

To better understand the relative contribution of CMB lensing and noise residuals, we repeat the same analysis while applying two independent delensing and denoising factors on the CMB and noise residuals components, respectively, after the component separation. In this way, we are removing part of the Gaussian contribution at the map level, in order

to check which is the dominant component and its impact on the ability of MFs to detect non-Gaussianity. In tables 2 and 3 we report results for a realistically attainable delensing of 50% ($A_{\text{lens}} = 0.5$), showing that in general we obtain an increase of the efficiency η as expected: we reduced one of the Gaussian contribution, and the anisotropy induced by foreground residuals is more evident to MFs. However, the detection number in the most realistic scenario (MC-NILC - d10s5) is still considerably low, being $< 30\%$. For this particular case only, we also rescale the noise residuals maps by a denoising factor of 75% ($A_{\text{noise}} = 0.25$), observing an efficiency raise to 93% and 92% for MFs compared with CMBtens and FGgauss maps, respectively, in pixel space. In the needlet-based analysis, the efficiencies reach $\sim 99\%$ in both cases. This confirms the role of noise residuals as the dominant component over CMB lensing, driving the Gaussian contribution. It also provides the noise level needed for the MFs to be considered as reliable tools for this robustness test.

It is also worth to discuss the d10s5 scenario, which at first sight provides a counter-intuitive result: in general we expect that an increase in foregrounds complexity should correspond to an increase in the number of detected deviations, with respect to the simpler d1s1 case, as the foreground residuals should be more significant. Instead, we obtain a lower amount of detections for d10s5, both in pixel- and needlet-based analysis. This can be explained by considering that in general, the application of a minimum-variance component separation pipeline to a more complex foreground model results in the increase of both foreground and noise residuals. It means that the CMB solution will contain more contamination due to foreground residuals, but it will also carry a stronger Gaussian component driven by noise residuals, which is responsible for reducing the MFs detection power (as discussed in the previous paragraph). An additional analysis on this is reported in the appendix A.

As mentioned in section 3.1, in this work we ignored the eventual artifacts introduced in B -mode maps by the harmonic transformation from the originally simulated Q/U sky maps. We decided to blindly propagate this contribution through the pipeline. In any case, its missing mitigation (before component separation) should increase the level of non-Gaussianity in the maps with respect to the reference Gaussian simulations, meaning that the number of identified false detections (i.e., the efficiency of the test η) is expected to be overestimated. However, our results suggest that even in this pessimistic scenario, MFs are not able to reliably flag a false detection of the tensor-to-scalar ratio; therefore, removing this contamination would only reduce the efficiency of the test, not changing the overall outcome.

Following the results of the comparison between MFs of FGres and FGgauss maps, we conclude that, with the obtained relative amount of foregrounds and noise residuals in the component separated maps, MFs will not be able to detect the non-cosmological nature of a potentially biased measurement of the tensor-to-scalar ratio; the level of noise residuals is masking the anisotropic non-Gaussian contribution due to foreground residuals, preventing MFs from correctly identifying the contamination. The false cosmological signal passes the MFs test in the majority of the simulations, while not providing additional information on the astrophysical nature of the contamination. Thus, MFs can not be reliably used as higher-order statistical tool for a robustness test in the event of a future tensor-to-scalar ratio detection with a *LiteBIRD*-like configuration. However, they may still be useful in scenarios where foreground residuals are relatively more dominant than noise residuals.

5 Conclusions

Future CMB experiments like the *LiteBIRD* satellite will target a detection of the primordial tensor perturbations with an overall uncertainty on the tensor-to-scalar ratio of $\delta(r=0) = 0.001$. The presence of foreground residuals left after the application of component separation algorithms could bias this measurement, providing a false detection of the primordial signal.

In this work, we explored the possibility of using Minkowski functionals (MFs) to validate an eventual detection of the tensor-to-scalar ratio, exploiting the sensitivity of this higher-order statistics to the presence of non-Gaussianity induced by foreground residuals. We applied two minimum-variance component separation algorithms (NILC and MC-NILC) to realistic sky simulation of CMB, noise, and three different models of foregrounds (`d0s0`, `d1s1`, `d10s5`) in order to obtain cleaned B -mode maps. After evaluating the likelihood function on their angular power spectrum, we recover a (biased) tensor-to-scalar ratio for each realisation. We used these inferred values of r to generate Gaussian simulations of tensor perturbations, and then we compared the MFs of these simulations with the ones evaluated on the CMB solution (the output map of the component separation), in order to check their compatibility and establish if the MFs can spot the discrepancy between the two sets of maps due to the foregrounds non-Gaussianity. Since MFs expected values depend both on the shape of the power spectrum and statistics of the considered maps, we also compare the CMB solution MFs with those computed on a set of maps that have the same CMB lensing and noise residuals, but with the non-Gaussian foreground residuals replaced by a Gaussian counterpart with the same power spectrum. In this way, the eventual deviation detected by MFs is entirely driven by non-Gaussianity.

The general outcome of the study is that MFs, with the obtained relative level of foreground and noise residuals for the considered instrumental configuration and component separation techniques, are not able to significantly detect the presence of non-Gaussianity in component-separated maps, even if this contamination is biasing the value of the inferred tensor-to-scalar ratio. This result has been observed both when using all the scales for MFs computation, or considering needlet-filtered maps (isolating the largest scales). In the majority of the cases, the test is passed and no warning is raised, meaning that in real life we would accept a false signal as a real measurement of r . This means that MFs can not be used as a reliable robustness test in the context of a map-based inference of the tensor-to-scalar ratio, at least for a *LiteBIRD*-like experiment which adopts MC-NILC as component separation algorithm. An analogous analysis could be applied to other experimental configurations, for instance to a ground-based CMB telescope like SO [16], where only a partial region of the sky is observed, and for which the MFs study could provide different results.

MFs are not the only higher-order statistical tools available for CMB studies. For instance, wavelet scattering transform (WST) [74–77] are low-variance statistical descriptor of non-Gaussian processes introduced in the field of data science, able to provide summary statistics of CMB polarisation maps. Also the bispectrum estimators [78] could represent a powerful tool for the study on non-Gaussian features in component-separated CMB maps. We will address the exploration of such tools in the context of robustness test for tensor-to-scalar measurements in a future work.

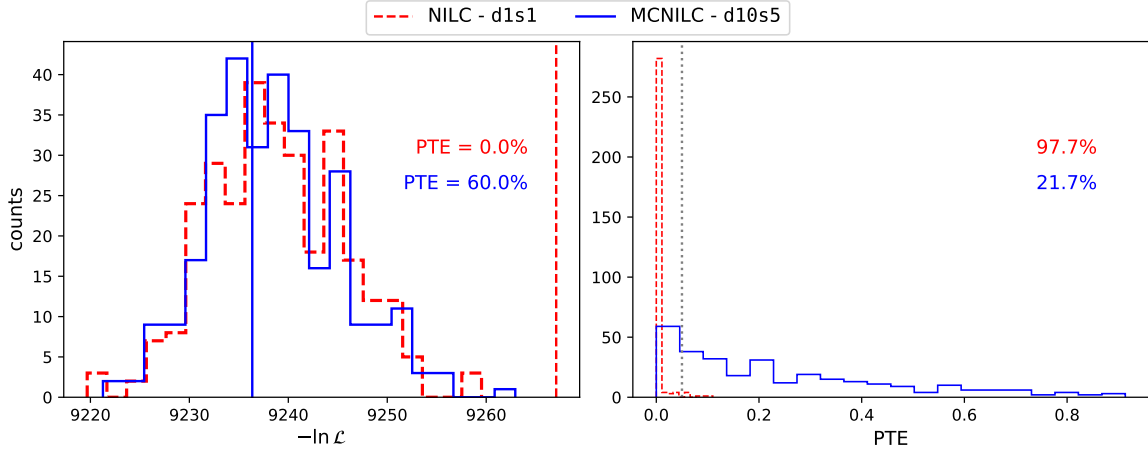


Figure 11. Results of the goodness-of-fit analysis for the NILC - **d1s1** (red dashed lines) and MC-NILC - **d10s5** (blue solid lines) cases. The left panel shows one realisation of the $-\ln \mathcal{L}$ distribution of the 300 tensor simulations, compared against the value given by the best fit on the CMB solution (vertical lines); the corresponding PTE is also indicated. The right panel reports the PTE distribution for the 300 simulations, with the percentage of incompatible simulations shown in the upper right.

Acknowledgments

We thank Anthony Banday, Luca Pagano, Giuseppe Puglisi, Giacomo Galloni, and BB for the useful comments, feedbacks, and discussions in the development of this work. The authors acknowledge partial support by the Italian Space Agency (ASI Grants No. 2020-9-HH.0 and 2016-24-H.1-2018) and the Euclid Project, as well as the RadioForegroundsPlus Project HORIZON-CL4-2023-SPACE-01, GA 101135036 and through the Project SPACE-IT-UP by the Italian Space Agency and Ministry of University and Research, Contract Number 2024-5-E.0. The authors also acknowledge partial support by the InDark and LiteBIRD Initiative of the National Institute for Nuclear Physics.

A Supplementary results

A.1 Goodness-of-fit analysis

In this section we show some examples of the goodness-of-fit analysis carried out during the inference of the tensor-to-scalar ratio (described in section 3.3). Figure 11 reports results for NILC - **d1s1** and MC-NILC **d10s5** scenarios, with the $-\ln \mathcal{L}$ distribution for one simulation on the left, and the PTE distribution on the right, along with the percentage of simulations failing the χ^2 test. For this particular realisation, the NILC - **d1s1** case fails the test ($\text{PTE} < 0.05$) while for MC-NILC - **d10s5** the test is passed ($\text{PTE} > 0.05$).

A.2 Examples of MFs in other scenarios

Here we report MFs evaluated in other scenarios considered in this work. In figure 12 we illustrate the NILC - **d1s1** case both in pixel and needlet space (largest scales), while in figure 13 we show the MFs weighted difference as presented in figure 10, for the NILC

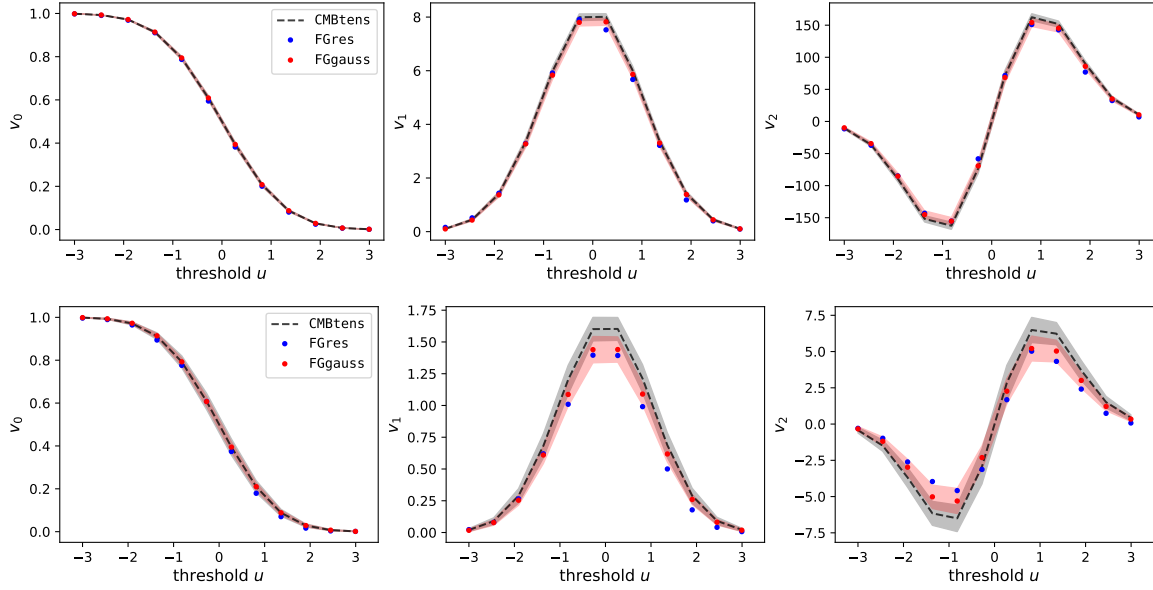


Figure 12. Same as figure 9, but for the NILC - **d1s1** scenario. *Upper row:* pixel space (including all scales). *Lower row:* needlet space (first band, $\ell \lesssim 60$).

algorithm applied on the **d0s0** and **d1s1** foreground models. The **d0s0** panels are useful to visualize a configuration where we have no pattern in the distribution of simulations, compatible with noise-like features. See the discussion in section 4 for more details.

A.3 Dependence of MFs test efficiency on the foreground model complexity

In section 4 we found a decrease in the MFs efficiency when moving from a lower (**d1s1**) to a higher (**d10s5**) foreground complexity, and we addressed it to the higher level of noise residuals present in the latter case. To explore this hypothesis, for each simulation we evaluate the relative amount of residuals by computing the ratio between the power spectra of foreground (C_ℓ^{fgs}) and noise residuals (C_ℓ^{nres}). In figure 14 we report the mean of this quantity across simulations: for the NILC algorithm (left panel) the average ratio in the **d10s5** case is slightly lower than **d1s1**, meaning that the separation between foregrounds and noise is less clear for MFs to detect, thus leading to a lower efficiency η . For the MC-NILC case the residual ratio is similar for the two models, leading to roughly the same number of detections. These considerations suggest that the MFs efficiency is not heavily dependent on the complexity of the foreground model, but rather on the relative levels of foregrounds and noise residuals, which can both increase or decrease based on the performance of the component separation.

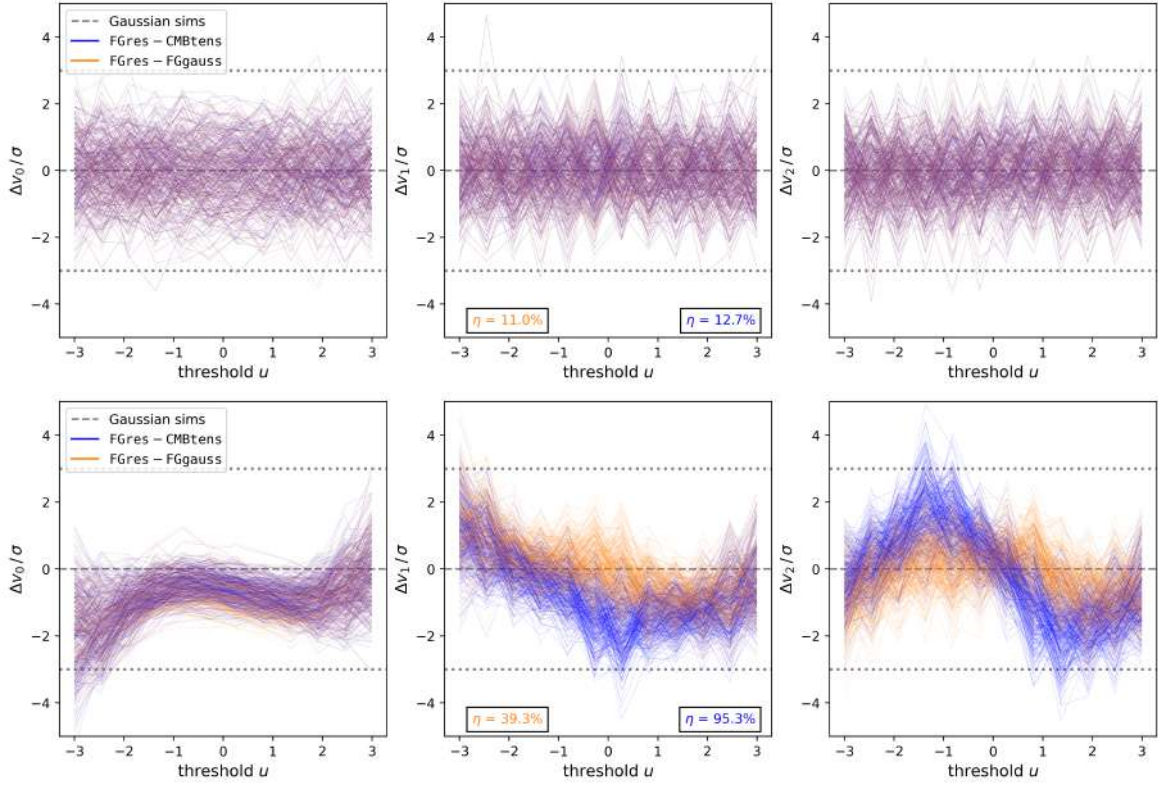


Figure 13. Same as figure 10, but for the NILC - d0s0 (upper row) and NILC - d1s1 (lower row) cases, in pixel space (including all scales).

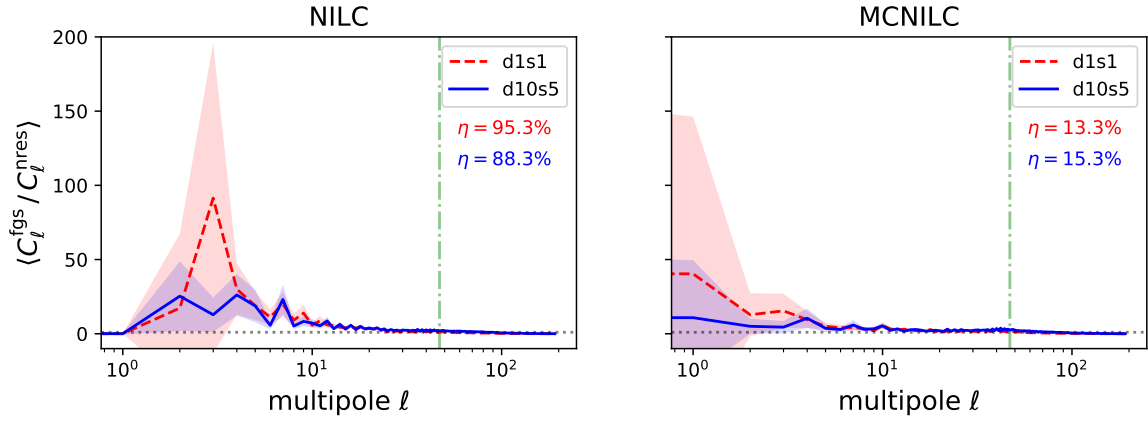


Figure 14. Average ratio between spectra of foreground and noise residuals. Dashed red lines indicate the d1s1 model, while blue solid lines are for d10s5. Shaded areas represent the standard deviation across 300 simulations. MFs test efficiencies are also reported for comparison. The vertical line indicates the smallest scale included in the first needlet band b_0 .

References

- [1] WMAP collaboration, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, *Astrophys. J. Suppl.* **208** (2013) 20 [[arXiv:1212.5225](#)] [[INSPIRE](#)].
- [2] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [3] ACT collaboration, *The Atacama Cosmology Telescope: DR4 Maps and Cosmological Parameters*, *JCAP* **12** (2020) 047 [[arXiv:2007.07288](#)] [[INSPIRE](#)].
- [4] R. Brout, F. Englert and E. Gunzig, *The Creation of the Universe as a Quantum Phenomenon*, *Annals Phys.* **115** (1978) 78 [[INSPIRE](#)].
- [5] A.A. Starobinsky, *A New Type of Isotropic Cosmological Models Without Singularity*, *Phys. Lett. B* **91** (1980) 99 [[INSPIRE](#)].
- [6] A.H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, *Phys. Rev. D* **23** (1981) 347 [[INSPIRE](#)].
- [7] M. Kamionkowski, A. Kosowsky and A. Stebbins, *A probe of primordial gravity waves and vorticity*, *Phys. Rev. Lett.* **78** (1997) 2058 [[astro-ph/9609132](#)] [[INSPIRE](#)].
- [8] W. Hu and M.J. White, *A CMB polarization primer*, *New Astron.* **2** (1997) 323 [[astro-ph/9706147](#)] [[INSPIRE](#)].
- [9] U. Seljak and M. Zaldarriaga, *Signature of Gravity Waves in the Polarization of the Microwave Background*, *Phys. Rev. Lett.* **78** (1997) 2054 [[astro-ph/9609169](#)] [[INSPIRE](#)].
- [10] PLANCK collaboration, *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10 [[arXiv:1807.06211](#)] [[INSPIRE](#)].
- [11] M. Tristram et al., *Improved limits on the tensor-to-scalar ratio using BICEP and Planck data*, *Phys. Rev. D* **105** (2022) 083524 [[arXiv:2112.07961](#)] [[INSPIRE](#)].
- [12] G. Galloni et al., *Updated constraints on amplitude and tilt of the tensor primordial spectrum*, *JCAP* **04** (2023) 062 [[arXiv:2208.00188](#)] [[INSPIRE](#)].
- [13] M. Kamionkowski and E.D. Kovetz, *The Quest for B Modes from Inflationary Gravitational Waves*, *Ann. Rev. Astron. Astrophys.* **54** (2016) 227 [[arXiv:1510.06042](#)] [[INSPIRE](#)].
- [14] LITEBIRD collaboration, *Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey*, *PTEP* **2023** (2023) 042F01 [[arXiv:2202.02773](#)] [[INSPIRE](#)].
- [15] BICEP and KECK collaborations, *Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season*, *Phys. Rev. Lett.* **127** (2021) 151301 [[arXiv:2110.00483](#)] [[INSPIRE](#)].
- [16] SIMONS OBSERVATORY collaboration, *The Simons Observatory: Science goals and forecasts*, *JCAP* **02** (2019) 056 [[arXiv:1808.07445](#)] [[INSPIRE](#)].
- [17] K. Abazajian et al., *CMB-S4 Science Case, Reference Design, and Project Plan*, [[arXiv:1907.04473](#)] [[INSPIRE](#)].
- [18] M. Zaldarriaga and U. Seljak, *Gravitational lensing effect on cosmic microwave background polarization*, *Phys. Rev. D* **58** (1998) 023003 [[astro-ph/9803150](#)] [[INSPIRE](#)].
- [19] A. Lewis and A. Challinor, *Weak gravitational lensing of the CMB*, *Phys. Rept.* **429** (2006) 1 [[astro-ph/0601594](#)] [[INSPIRE](#)].

- [20] E.J. Baxter et al., *A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope*, *Astrophys. J.* **806** (2015) 247 [[arXiv:1412.7521](#)] [[INSPIRE](#)].
- [21] B.D. Sherwin et al., *Two-season Atacama Cosmology Telescope polarimeter lensing power spectrum*, *Phys. Rev. D* **95** (2017) 123529 [[arXiv:1611.09753](#)] [[INSPIRE](#)].
- [22] POLARBEAR collaboration, *Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR experiment*, *Phys. Rev. Lett.* **113** (2014) 021301 [[arXiv:1312.6646](#)] [[INSPIRE](#)].
- [23] BICEP and KECK collaborations, *Measurements of Degree-Scale B-mode Polarization with the BICEP/Keck Experiments at South Pole*, in the proceedings of the 53rd Rencontres de Moriond on Cosmology, La Thuile, Italy, March 17–24 (2018) [[arXiv:1807.02199](#)] [[INSPIRE](#)].
- [24] R. Skalidis et al., *Local measurements of the mean interstellar polarization at high Galactic latitudes*, *Astron. Astrophys.* **616** (2018) A52 [[arXiv:1802.04305](#)] [[INSPIRE](#)].
- [25] PLANCK collaboration, *Planck 2015 results. X. Diffuse component separation: Foreground maps*, *Astron. Astrophys.* **594** (2016) A10 [[arXiv:1502.01588](#)] [[INSPIRE](#)].
- [26] PLANCK collaboration, *Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes*, *Astron. Astrophys.* **586** (2016) A133 [[arXiv:1409.5738](#)] [[INSPIRE](#)].
- [27] PLANCK collaboration, *Planck 2018 results. IV. Diffuse component separation*, *Astron. Astrophys.* **641** (2020) A4 [[arXiv:1807.06208](#)] [[INSPIRE](#)].
- [28] PLANCK collaboration, *Planck 2018 results. XI. Polarized dust foregrounds*, *Astron. Astrophys.* **641** (2020) A11 [[arXiv:1801.04945](#)] [[INSPIRE](#)].
- [29] N. Krachmalnicoff et al., *Characterization of foreground emission on degree angular scales for CMB B-mode observations. Thermal dust and synchrotron signal from Planck and WMAP data*, *Astron. Astrophys.* **588** (2016) A65 [[arXiv:1511.00532](#)] [[INSPIRE](#)].
- [30] J. Delabrouille and J.-F. Cardoso, *Diffuse source separation in CMB observations*, in the proceedings of the *International Summer School on Data Analysis in Cosmology*, Valencia, Spain, September 06–10 (2004) [[astro-ph/0702198](#)] [[INSPIRE](#)].
- [31] S.M. Leach et al., *Component separation methods for the PLANCK mission*, *Astron. Astrophys.* **491** (2008) 597 [[arXiv:0805.0269](#)] [[INSPIRE](#)].
- [32] M.H. Abitbol et al., *The Simons Observatory: gain, bandpass and polarization-angle calibration requirements for B-mode searches*, *JCAP* **05** (2021) 032 [[arXiv:2011.02449](#)] [[INSPIRE](#)].
- [33] H.K. Eriksen et al., *Joint Bayesian component separation and CMB power spectrum estimation*, *Astrophys. J.* **676** (2008) 10 [[arXiv:0709.1058](#)] [[INSPIRE](#)].
- [34] R. Stompor, S.M. Leach, F. Stivoli and C. Baccigalupi, *Maximum Likelihood algorithm for parametric component separation in CMB experiments*, *Mon. Not. Roy. Astron. Soc.* **392** (2009) 216 [[arXiv:0804.2645](#)] [[INSPIRE](#)].
- [35] LITEBIRD collaboration, *Moment expansion of polarized dust SED: A new path towards capturing the CMB B-modes with LiteBIRD*, *Astron. Astrophys.* **660** (2022) A111 [[arXiv:2111.07742](#)] [[INSPIRE](#)].
- [36] L. Vacher et al., *How bad could it be? Modelling the 3D complexity of the polarised dust signal using moment expansion*, *Astron. Astrophys.* **697** (2025) A212 [[arXiv:2411.11649](#)] [[INSPIRE](#)].

- [37] J. Delabrouille et al., *A full sky, low foreground, high resolution CMB map from WMAP*, *Astron. Astrophys.* **493** (2009) 835 [[arXiv:0807.0773](#)] [[INSPIRE](#)].
- [38] LITEBIRD collaboration, *Multiclustering needlet ILC for CMB B-mode component separation*, *Mon. Not. Roy. Astron. Soc.* **525** (2023) 3117 [[arXiv:2212.04456](#)] [[INSPIRE](#)].
- [39] M. Tegmark, A. de Oliveira-Costa and A. Hamilton, *A high resolution foreground cleaned CMB map from WMAP*, *Phys. Rev. D* **68** (2003) 123523 [[astro-ph/0302496](#)] [[INSPIRE](#)].
- [40] D. Alonso, J. Dunkley, B. Thorne and S. Næss, *Simulated forecasts for primordial B-mode searches in ground-based experiments*, *Phys. Rev. D* **95** (2017) 043504 [[arXiv:1608.00551](#)] [[INSPIRE](#)].
- [41] A. Carones, M. Migliaccio, D. Marinucci and N. Vittorio, *Analysis of Needlet Internal Linear Combination performance on B-mode data from sub-orbital experiments*, *Astron. Astrophys.* **677** (2023) A147 [[arXiv:2208.12059](#)] [[INSPIRE](#)].
- [42] BICEP2 collaboration, *Detection of B-Mode Polarization at Degree Angular Scales by BICEP2*, *Phys. Rev. Lett.* **112** (2014) 241101 [[arXiv:1403.3985](#)] [[INSPIRE](#)].
- [43] H. Tomita, *Curvature Invariants of Random Interface Generated by Gaussian Fields*, *Prog. Theor. Phys.* **76** (1986) 952.
- [44] P. Coles and J.D. Barrow, *Non-Gaussian statistics and the microwave background radiation*, *Mon. Not. Roy. Astron. Soc.* **228** (1987) 407.
- [45] J.R. Gott III et al., *Topology of Microwave Background Fluctuations: Theory*, *Astrophys. J.* **352** (1990) 1.
- [46] S. Winitzki and A. Kosowsky, *Minkowski functional description of microwave background Gaussianity*, *New Astron.* **3** (1998) 75 [[astro-ph/9710164](#)] [[INSPIRE](#)].
- [47] J. Schmalzing and K.M. Górski, *Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps*, *Mon. Not. Roy. Astron. Soc.* **297** (1998) 355 [[astro-ph/9710185](#)] [[INSPIRE](#)].
- [48] PLANCK collaboration, *Planck 2018 results. VII. Isotropy and Statistics of the CMB*, *Astron. Astrophys.* **641** (2020) A7 [[arXiv:1906.02552](#)] [[INSPIRE](#)].
- [49] F. Rahman, P. Chingangbam and T. Ghosh, *The nature of non-Gaussianity and statistical isotropy of the 408 MHz Haslam synchrotron map*, *JCAP* **07** (2021) 026 [[arXiv:2104.00419](#)] [[INSPIRE](#)].
- [50] D. Zürcher et al., *Cosmological Forecast for non-Gaussian Statistics in large-scale weak Lensing Surveys*, *JCAP* **01** (2021) 028 [[arXiv:2006.12506](#)] [[INSPIRE](#)].
- [51] A. Ducout et al., *Non Gaussianity and Minkowski Functionals: forecasts for Planck*, *Mon. Not. Roy. Astron. Soc.* **429** (2013) 2104 [[arXiv:1209.1223](#)] [[INSPIRE](#)].
- [52] V. Ganesan, P. Chingangbam, K.P. Yogendran and C. Park, *Primordial non-Gaussian signatures in CMB polarization*, *JCAP* **02** (2015) 028 [[arXiv:1411.5256](#)] [[INSPIRE](#)].
- [53] PLANCK collaboration, *Planck 2015 results. XVI. Isotropy and statistics of the CMB*, *Astron. Astrophys.* **594** (2016) A16 [[arXiv:1506.07135](#)] [[INSPIRE](#)].
- [54] L. Santos, K. Wang and W. Zhao, *Probing the statistical properties of CMB B-mode polarization through Minkowski Functionals*, *JCAP* **07** (2016) 029 [[arXiv:1510.07779](#)] [[INSPIRE](#)].
- [55] J. Carrón Duque et al., *Minkowski Functionals in $SO(3)$ for the spin-2 CMB polarisation field*, *JCAP* **01** (2024) 039 [[arXiv:2301.13191](#)] [[INSPIRE](#)].

- [56] A. Carones et al., *Minkowski functionals of CMB polarization intensity with pynkowski: theory and application to Planck and future data*, *Mon. Not. Roy. Astron. Soc.* **527** (2024) 756 [[arXiv:2211.07562](#)] [[INSPIRE](#)].
- [57] R. Adler and J. Taylor, *Random Fields and Geometry*, Springer Monographs in Mathematics, Springer New York (2009).
- [58] K.M. Górski et al., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, *Astrophys. J.* **622** (2005) 759 [[astro-ph/0409513](#)] [[INSPIRE](#)].
- [59] A. Zonca et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, *J. Open Source Softw.* **4** (2019) 1298 [[INSPIRE](#)].
- [60] B. Thorne, J. Dunkley, D. Alonso and S. Naess, *The Python Sky Model: software for simulating the Galactic microwave sky*, *Mon. Not. Roy. Astron. Soc.* **469** (2017) 2821 [[arXiv:1608.02841](#)] [[INSPIRE](#)].
- [61] A. Zonca, B. Thorne, N. Krachmalnicoff and J. Borrill, *The Python Sky Model 3 software*, *J. Open Source Softw.* **6** (2021) 3783 [[arXiv:2108.01444](#)] [[INSPIRE](#)].
- [62] PAN-EXPERIMENT GALACTIC SCIENCE GROUP collaboration, *Full-sky Models of Galactic Microwave Emission and Polarization at Sub-arcminute Scales for the Python Sky Model*, [arXiv:2502.20452](#) [[INSPIRE](#)].
- [63] L. Santos et al., *Statistical imprints of CMB B-type polarization leakage in an incomplete sky survey analysis*, *JCAP* **01** (2017) 043 [[arXiv:1612.03564](#)] [[INSPIRE](#)].
- [64] S. Basak and J. Delabrouille, *A needlet ILC analysis of WMAP 7-year data: estimation of CMB temperature map and power spectrum*, *Mon. Not. Roy. Astron. Soc.* **419** (2012) 1163 [[arXiv:1106.5383](#)] [[INSPIRE](#)].
- [65] K. Wolz et al., *The Simons Observatory: pipeline comparison and validation for large-scale B-modes*, *Astron. Astrophys.* **686** (2024) A16 [[arXiv:2302.04276](#)] [[INSPIRE](#)].
- [66] F.J. Narcowich, P. Petrushev and J.D. Ward, *Localized Tight Frames on Spheres*, *SIAM J. Math. Anal.* **38** (2006) 574.
- [67] D. Pietrobon, A. Balbi and D. Marinucci, *Integrated Sachs-Wolfe effect from the cross-correlation of WMAP 3 year and NVSS: new results and constraints on dark energy*, *Phys. Rev. D* **74** (2006) 043524 [[astro-ph/0606475](#)] [[INSPIRE](#)].
- [68] D. Geller and D. Marinucci, *Spin Wavelets on the Sphere*, [arXiv:0811.2935](#) [[INSPIRE](#)].
- [69] J.C. Duque et al., *Point Source Detection and False Discovery Rate Control on CMB Maps*, *Astron. Comput.* **28** (2019) 100310 [[arXiv:1902.06636](#)] [[INSPIRE](#)].
- [70] S. Hamimeche and A. Lewis, *Properties and use of CMB power spectrum likelihoods*, *Phys. Rev. D* **79** (2009) 083012 [[arXiv:0902.0674](#)] [[INSPIRE](#)].
- [71] M. Gerbino et al., *Likelihood methods for CMB experiments*, *Front. in Phys.* **8** (2020) 15 [[arXiv:1909.09375](#)] [[INSPIRE](#)].
- [72] A. Lewis, A. Challinor and N. Turok, *Analysis of CMB polarization on an incomplete sky*, *Phys. Rev. D* **65** (2002) 023505 [[astro-ph/0106536](#)] [[INSPIRE](#)].
- [73] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [[astro-ph/0608064](#)] [[INSPIRE](#)].

- [74] E. Allys et al., *The RWST, a comprehensive statistical description of the non-Gaussian structures in the ISM*, *Astron. Astrophys.* **629** (2019) A115 [[arXiv:1905.01372](#)] [[INSPIRE](#)].
- [75] J.-M. Delouis, E. Allys, E. Gauvrit and F. Boulanger, *Non-Gaussian modelling and statistical denoising of Planck dust polarisation full-sky maps using scattering transforms*, *Astron. Astrophys.* **668** (2022) A122 [[arXiv:2207.12527](#)] [[INSPIRE](#)].
- [76] L. Mousset et al., *Generative models of astrophysical fields with scattering transforms on the sphere*, *Astron. Astrophys.* **691** (2024) A269 [[arXiv:2407.07007](#)] [[INSPIRE](#)].
- [77] P. Campeti et al., *From few to many maps: A fast map-level emulator for extreme augmentation of cosmic microwave background systematics datasets*, *Astron. Astrophys.* **700** (2025) A136 [[arXiv:2503.11643](#)] [[INSPIRE](#)].
- [78] PLANCK collaboration, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [[arXiv:1905.05697](#)] [[INSPIRE](#)].