PHYSICAL REVIEW D **104,** 023502 (2021)

Statistics of nonpolarized points in the CMB polarization maps

Jaan Kasak,[*] James Creswell,[†] and Pavel Naselsky[‡]
Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark

Hao Liu[§]
School of Physics and Material Science, Anhui University, 111 Jiulong Road, Hefei, Anhui 230601, China

(Received 27 January 2021; accepted 4 May 2021; published 2 July 2021)

The nonpolarized points (NPP) of the Q and U Stokes parameters of the CMB can be classified according to the geometry of the polarization field. We describe a procedure to identify these points in the pixelized sky and present the shape of the polarization angles in the vicinity of NPPs. We design a test of Gaussianity using the Kullback-Leibler divergence. We show that the total number density of nonpolarized points of the E and B families is closely related to the presence of lensing and the tensor-to-scalar ratio r. We further show that in the absence of lensing, the total number of NPPs of all types does not depend on r, while the lensing effect removes this degeneracy. This analysis is applied to the CMB maps from the 2018 Planck release. We show that there is a general consistency of SMICA and NILC maps compared to a reference set of Gaussian simulations. The strongest discrepancies are found in the Commander (with corresponding p value 0.07) and SEVEM ($p = 0.01$) maps.

DOI: 10.1103/PhysRevD.104.023502

I. INTRODUCTION

Verifying the Gaussianity and statistical isotropy of the CMB polarization is an important test of the Lambda-CDM model as well as inflation and other early Universe physics. In this direction, taking under consideration forthcoming CMB experiments, statistical characterization of the B mode of polarization is especially important [1–3]. Much effort has been made towards accurate determination of power spectra, and in addition, other estimators and techniques have been developed to characterize the polarization fluctuations. These include Minkowski functionals, peak analysis, power asymmetry, parity asymmetry, single- and multidimensional moments, etc. (see, for review, [4,5]). Since the anomalies of the CMB up to now do not have theoretical bases, it seems to be very important to investigate all possible estimators of non-Gaussianity and statistical anisotropy in order to split the effects of systematics and component separation (foregrounds) from primordial sources.

In this direction, we would like to draw attention to the statistical peculiarities of the Q and U components of the Stokes "vector," related to the nonpolarized points (NPP) in the intensity map $I = \sqrt{Q^2 + U^2}$ and the corresponding features of the polarization angle $\tan(2\Psi) = \frac{U}{Q}$. In principal, the Stokes parameters Q and U comprise a continuous spin-2 valued function on the sphere. Because of their continuity, there will be well-defined contours along which $Q = 0$ and similarly, contours along which $U = 0$ (so-called contours of percolation [6]). At the points where these contours intersect, Q and U are simultaneously 0, and the total sky signal is unpolarized [7]. At the same time, these NPP manifest themselves as points of absolute minimum (zeros) of the polarization intensity, which for Gaussian Q and U has a Rayleigh distribution. Thus, there is a connection between the statistics of NPP and the statistics of minima of the non-Gaussian field of polarization intensity. Note that the existence of these points is a natural product for a correlated pseudovector field as points of connection of the domains with a different polarization, similar to the domains in ferromagnetics. The number density of these points, their morphology, and their relative concentration will reveal some peculiarities of the morphology of the polarization due to different systematic effects, instrumental noise, and residuals of the component separation technique.

In [7], these nonpolarized points for random Gaussian isotropic fields were investigated. See also [8]. Several results from this work can be summarized as follows:

(i) The total number density of NPP is a constant depending only on the correlation radius of the signal.
(ii) NPP can be classified into one of three kinds (knots, foci, and saddles) depending on the local geometry of the field.

---

[*] jaan.kasak@abzu.ai
[†] james.creswell@nbi.ku.dk
[‡] naselsky@nbi.dk
[§] ustc_liuhao@163.com

(iii) The ratios of number densities of these different kinds of NPP are fixed, independent of the power spectrum and depend only on the assumption of Gaussianity.

The third result motivates a simple test of Gaussianity, based on counting and classifying NPP and calculating the corresponding ratios of their counts. In any finite sky area, random sampling error permits some departure of the actual counts from the theoretical number densities. Therefore, we compare the variation of the ratios across the sky with that of Gaussian simulations.

One additional feature is the properties of E and B modes of polarization, widely used in the search for cosmological gravitational waves. In [9] it was shown that Q and U can be decomposed into (Q_E, U_E) and (Q_B, U_B) families, which generate the corresponding E and B modes: $(Q_E, U_E) \rightarrow E$ and $(Q_B, U_B) \rightarrow B$. In this transform, we remain in the domain of the Stokes parameters, and the statistics of NPP apply to each.

In Sec. II, we review the theory of NPP in the polarization field and derive the main analytical results. In Sec. III, we summarize the procedure for measuring NPP in pixelized data. The theory of NPP in the E and B modes is discussed in Sec. IV, with emphasis on the effect and detectability of lensing by this estimator. In Sec. V, the statistics of the singular point ratios are explored, and Gaussian simulations are compared to the Planck 2018 CMB maps [10]. A brief conclusion is given in Sec. VI.

II. THEORY OF NONPOLARIZED POINTS

Inflation predicts Gaussian statistics for the CMB, including its polarization. These assumptions allow us to also calculate predictions for the densities of nonpolarized points in the CMB. Unfortunately, the Planck CMB products (SMICA, Commander, NILC, and SEVEM) are contaminated by foreground residuals and instrumental noise. These factors will affect the statistics of NPP compared to a pure cosmological signal. At the same time, these peculiarities of statistics can be used for determination of their morphology and amplitudes, and they can be masked out by more complicated filters designed in the domain of polarization intensity, e.g., [11]. We will discuss these anomalies in the next sections of the article, and now we turn to the theory for Gaussian signals, following [7]. Note that this model is exact for statistically isotropic Gaussian noise added to the CMB.

A. Statistics of Gaussian Q and U and their derivatives

We start with the basis for our description of the CMB polarization field from [7] for total Q and U without separation into E and B families. We assume that Q and U are random Gaussian variables with possibly nonzero means and define the Stokes vector,

$$\boldsymbol{P} = \begin{pmatrix} Q \\ U \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \mu_Q \\ \mu_U \end{pmatrix}, \tag{1}$$

where μ_Q and μ_U are the means of Q and U. The covariance matrix of Q and U is

$$\boldsymbol{\Sigma}_{QU} = \begin{pmatrix} \sigma_{QQ} & \sigma_{QU} \\ \sigma_{QU} & \sigma_{UU} \end{pmatrix}. \tag{2}$$

Since only second-order correlations exist, the statistics of the Stokes parameters are described by a Gaussian distribution. Using Eqs. (1) and (2), we write the joint probability density for Q and U,

$$\mathcal{P}(Q,U) = \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{QU})^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\boldsymbol{P}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}_{QU}^{-1}(\boldsymbol{P}-\boldsymbol{\mu})\right]. \tag{3}$$

We now make our first assumptions that Q and U have the same variance, namely the spectral parameter σ_0^2. We also take them to be completely uncorrelated, $\sigma_{QU} = 0$. The mean of random variables with a Gaussian distribution does not affect the shape of the PDF, only its location in the space of random variables, so we take $\boldsymbol{\mu} = 0$, and move to the polarization vector $\boldsymbol{P} \rightarrow \boldsymbol{P} + \boldsymbol{\mu}$, such that $\boldsymbol{P} - \boldsymbol{\mu} \rightarrow \boldsymbol{P} + \boldsymbol{\mu} - \boldsymbol{\mu} = \boldsymbol{P}$. With these assumptions, Eq. (3) simplifies to

$$\mathcal{P}(Q,U) = \frac{1}{2\pi\sigma_0^2} \exp\left(-\frac{1}{2}\frac{I^2}{\sigma_0^2}\right), \tag{4}$$

where $I^2 = ||\boldsymbol{P}||^2 = Q^2 + U^2$, and σ_0^2 is the variance of Q and U. For the purposes of analyzing NPP, we also want to describe the statistics of the field derivatives. We define the Jacobian,

$$\boldsymbol{J} = \begin{pmatrix} Q_\theta & Q_\varphi \\ U_\theta & U_\varphi \end{pmatrix}, \tag{5}$$

where

$$Q_\theta = \frac{\partial Q}{\partial \theta}, \qquad U_\theta = \frac{\partial U}{\partial \theta}, \qquad Q_\varphi = \frac{\partial Q}{\partial \varphi}, \qquad U_\varphi = \frac{\partial U}{\partial \varphi}. \tag{6}$$

In constructing a probability density for the derivatives, similar assumptions are made as for the field components themselves. Namely, that they are uncorrelated and have a variance tied to spectral parameters, $\sigma_{Q_iQ_i} = \sigma_{U_iU_i} = \sigma_1^2/2$. All other variances are then zero.

B. Classification of nonpolarized points

The trace and determinant of the Jacobian determine the classification of each singular point into either a focus, knot, or saddle. The conditions for foci, knots, and saddles are, respectively, given by

$$\det(\boldsymbol{J}) > \left[\frac{1}{2}\mathrm{Tr}(\boldsymbol{J})\right]^2 > 0 \quad \text{(foci)},$$
$$\left[\frac{1}{2}\mathrm{Tr}(\boldsymbol{J})\right]^2 > \det(\boldsymbol{J}) > 0 \quad \text{(knots)},$$
$$\det(\boldsymbol{J}) < 0 \quad \text{(saddles)}. \tag{7}$$

Trajectories of the polarization vector for each kind of NPP can be plotted around the singular point by solving the system of equations defined by the Jacobian in Eq. (5) acting on a polarization state. Examples are depicted in Fig. 1.

We next turn to discussion of the statistics of these NPP, based on [7]. The number density of different NPP is given by [7]

$$N_f = \frac{\sqrt{2}}{16\pi r_c^2} \quad N_k = \frac{1}{16\pi r_c^2}(2-\sqrt{2}) \quad N_s = \frac{1}{8\pi r_c^2}. \tag{8}$$

These are the densities of foci, knots, and saddles, respectively. The radius of correlation is introduced, defined as $r_c = \sigma_0/\sigma_1$, where σ_0^2 is the variance of Q and U and $\sigma_1^2/2$ is the variance of the derivatives, as above. Summing these densities gives the total density of nonpolarized points,

$$N_{\text{total}} = N_f + N_k + N_s = \frac{1}{4\pi r_c^2}. \tag{9}$$

Of note is that the ratios of these densities are parameter independent. We write them in terms of the focus density,

$$\frac{N_f}{N_k} = \sqrt{2}+1 \qquad \frac{N_f}{N_s} = \frac{\sqrt{2}}{2}. \tag{10}$$

Following the arguments here, a statistical ensemble of the pure Gaussian CMB skies should be one in which these ratios hold, and measuring these ratios in data allows us to describe the statistics of a CMB. However, for a single realization of the CMB sky with some finite number of NPP, the actual detected ratios are random variables. Thus, our goal is to extend the analysis of [7] in order to determine the uncertainties of these ratios.

Assuming that the underlying Q and U data are Gaussian random variables, we can derive the distribution followed by the ratios. Let $\hat{\sigma}_0^2$ be the sample variance of Q and U. It follows a chi-squared distribution. $\hat{\sigma}_1^2$ similarly is chi-squared distributed. The correlation length, $r_c = \sigma_0/\sigma_1$, is therefore the ratio of two chi-squared-distributed variables, which means it should follow (with suitable normalizations) an F distribution. We illustrate the distribution of the correlation length r_c in Fig. 2 for a statistical ensemble of 10^6 Gaussian realizations and compare the result with the probability density function of an F distribution, given by

$$F(x,\alpha,\beta) = Ax^{\frac{\alpha}{2}-1}\left(1+\frac{\alpha}{\beta}x\right)^{-\frac{\alpha+\beta}{2}}, \tag{11}$$

where A is a normalization constant and α and β are parameters of the distribution.

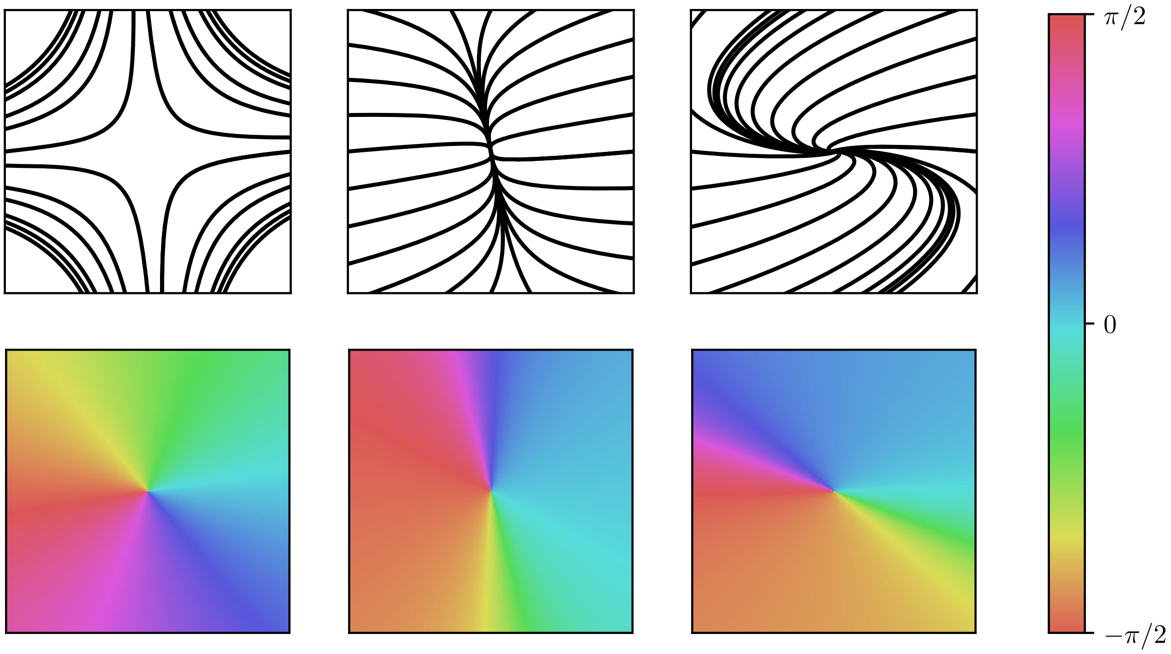

FIG. 1. Samples of NPP for Gaussian Q and U in flat (θ,φ) space. The first column features a saddle, the second a knot, and the last a focus. The second row is an image of the polarization angle Ψ linearly interpolated around the singular point, based on the first-order derivatives of Q and U evaluated at the singular point itself and defining its classification.

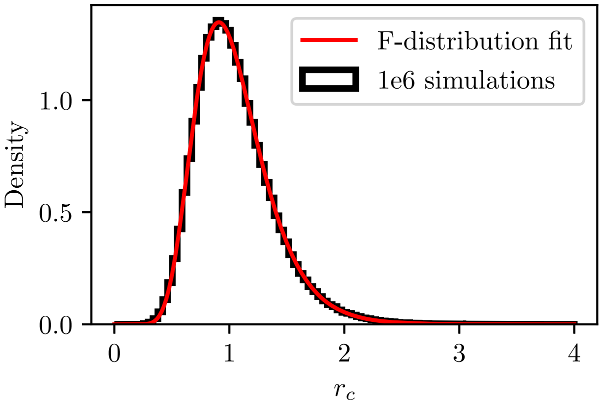

FIG. 2. The correlation length r_c follows an F distribution. Shown is the distribution of r_c from 10^6 Gaussian simulations compared to an F-distribution fit with parameters $\alpha = \beta \approx 42$. At high resolution, the F distribution reduces to a Cauchy distribution.

The analysis of NPP for Gaussian Q and U needs to be completed by description of the properties of polarization angle in the vicinity of these points. The behavior of Ψ, presented in the lower row of Fig. 1, is driven by the following equation:

$$\frac{d}{d\Theta}\tan(2\Psi) = \frac{\det(\boldsymbol{J})}{Q^2}, \tag{12}$$

where Θ is the polar angle in the coordinate system centered on the nonpolarized point. From this equation, we can explain part of the behavior observed in Fig. 1: the sign of the determinant controls whether the polarization angle increases or decreases clockwise around the NPP. For saddles, Ψ decreases clockwise; for knots and foci, Ψ increases.

III. DETECTING NONPOLARIZED POINTS IN INTENSITY MAPS

Nonpolarized points were defined as the places where both Q and U are zero. For the practical analysis of NPP, we need to implement some filtration for the Planck data in order to restore approximate continuity of the discrete pixelized data. We will use a standard Gaussian filter with a FWHM of 0.5 deg. Since these data are very noisy, to some extent, the Gaussian filter will remove high multipoles from the analysis and effectively decrease the power of the noise. Solving for NPP in the Q, U field configuration would be complicated, as the extremal points can be minima, maxima, or saddles. So we instead look to the polarization intensity, that is bounded from below by 0, where the singular points should be. Floating point inaccuracies and the resolution of available data further complicate the search, the true NPP will not appear as perfect zeroes in the data. As such, any local minimum in the polarization intensity with a sufficiently small value could be a candidate for a singular point. To deal with this, we introduce the parameter ε with the same units as the polarization intensity. This parameter is chosen according to the correlation scale of the map. We then select contiguous subthreshold basins, within which both Q and U separately change sign, as candidates for NPP. Local minima pixels within each basin are determined using a recursive search, and duplicate minima that are nearly adjacent are filtered out. The remainder are considered NPP and used for calculating ratios inside the region in question. Therefore, the conditions for a NPP at a particular pixel $\boldsymbol{n}_0$ are

1. (1) The Stokes parameters satisfy $|Q|, |U| < \varepsilon$ for an appropriately chosen value of ε.
2. (2) The polarization intensity is a local minimum, meaning $I(\boldsymbol{n}_0) < I(\boldsymbol{n}_i)$ for all pixels $\boldsymbol{n}_i$ adjacent to $\boldsymbol{n}_0$.
3. (3) Both positive and negative values of Q and U are found in the connected sub-ε region including $\boldsymbol{n}_0$.
4. (4) There are no other candidates also meeting conditions 1–3 with $I < I(\boldsymbol{n}_0)$ within an angular distance $\alpha_{\max}$ of $\boldsymbol{n}_0$.

ε and $\alpha_{\max}$ are parameters of the search method. ε is chosen such that the number of detected nonpolarized points matches the theoretical expectation of $1/r_c^2$. This value is found via a bisection search. $\alpha_{\max}$ is a suitable linear scaling of r_c, which also acts as the average angular distance between two NPPs. Once the pixels corresponding to each NPP are determined, they can be classified by simply evaluating the Jacobian at each pixel and applying the rules described above in Sec. II. The code that performs this detection of singular points is available at [12], and an illustrative example of NPP detection and classification is shown in Fig. 3.

It is convenient to count NPP in a target sky that has been divided into subareas. In any finite sky area, the actual ratios of counts will deviate from the Gaussian prediction. By dividing the sky into different areas and calculating the ratio within each, we can effectively run many "simulations" in one sky and compare the resulting variation to that of Gaussian simulations. This results in an estimator that is a more informative than calculating the full-sky ratio alone.

Simulations for these results were created in CAMB using parameters derived from the Plik likelihood [13].

For data serialized in the HEALPix [14] format, a HEALPix pixel at some fixed N_{side} is the natural choice. What remains is the choice of N_{side}. For statistical purposes, it should be large enough that a distribution of evaluated ratios from each of the subareas is available for analysis, yet not too large that clustering algorithms take up a lot of memory and CPU time.

In this work, we settled on choosing $N_{\text{side}} = 8$ and 16 for the mother pixel, preserving the pixelization for the sky map $N_{\text{side}} = 2048$. This choice of overpixelization

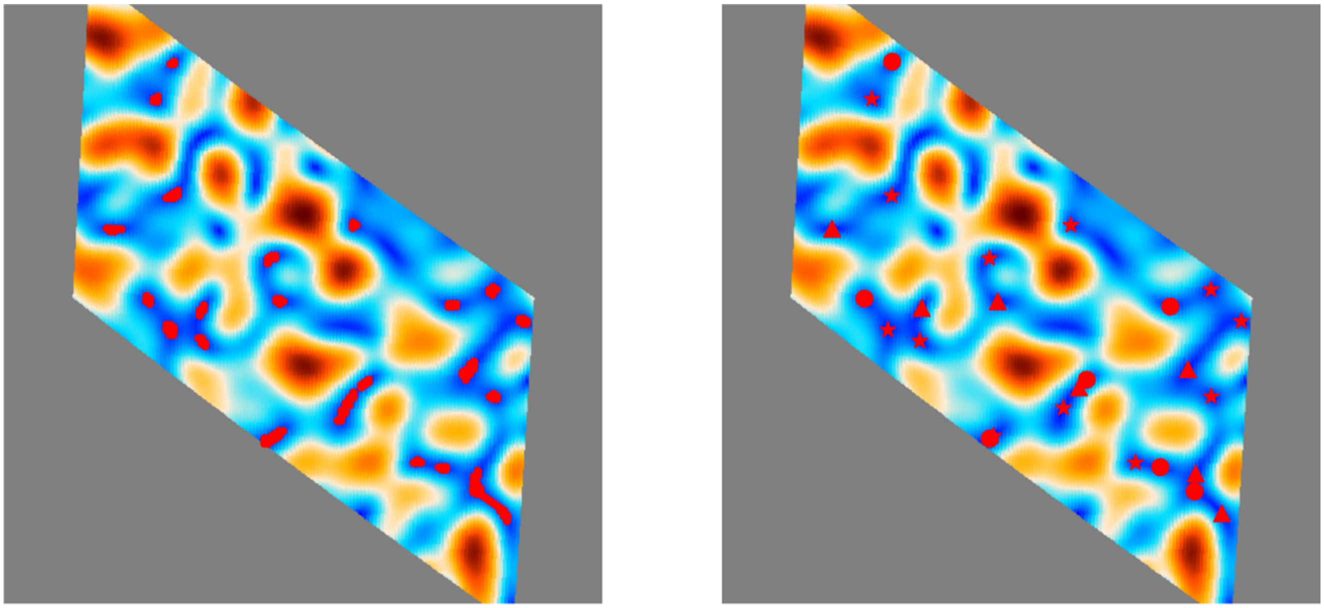

FIG. 3. Finding and classifying NPP in a sample mother pixel with $N_{\mathrm{side}} = 16$ from Gaussian simulations. In the left panel are highlighted sufficiently cold areas in the polarization intensity that also contain crossing points in both Q and U. In the right panel are the NPP that are kept for analysis of ratios. Circles are foci, triangles are knots, and stars are saddles. In the right panel, there are 7 foci, 7 knots, and 12 saddles. The ratios of this subarea are $N_f/N_k = 1$ and $N_f/N_s = 7/12$.

determines the "mother pixels" within each of which NPP are counted and ratios calculated to derive the final distribution functions. In practice, however, the detection algorithm is run on the full sky maps, and afterwards the points are assigned to their mother pixels. This means that the algorithm is not at risk of edge effects.

Areas of the sky obscured by the strongest foreground sources can be excluded from the analysis. In this work, we choose to downgrade Galactic masks to the mother pixel resolution using an area-preserving method. Beginning with the $N_{\mathrm{side}} = 2048$ Planck polarization confidence masks released with the component separation results [10], a mother pixel is masked or unmasked based on whether the number of masked subpixels exceeds a threshold, which is chosen such that the total masked sky area is the same in the downgraded mask. It should be noted that this area-preserving method does not necessarily maintain the original meaning of the mask. It is possible to adjust the threshold to make the mask more conservative or less conservative. In general, the method is local and can be applied to any particular sky area, including regions suspected to be contaminated by foreground residuals, which might be interesting objects of analysis. However, in this work, we adopt the area-preserving method as a compromise.

The basic results, including the masks in use throughout the rest of the paper, are illustrated in Fig. 4. Generally, no obvious patterns in the sky distribution of the ratios is visible in these plots. Mother pixels within which the ratios vary significantly from the Gaussian expectation are visible but apparently randomly distributed on the sky.

IV. E/B-FAMILY DECOMPOSITION AND THE TOTAL NUMBER DENSITY OF NONPOLARIZED POINTS

As we have pointed out above, the decomposition of the Stokes parameters into E and B families provides a direct link to the corresponding E and B modes of polarization (see, for details, [9]). Since all these transforms are based on liner combination of Q and U parameters, both E and B families will preserve the statistical properties of Q and U. Thus, Gaussian Q and U will generate Gaussian Q_E, U_E and Q_B, U_B, and all the criteria for NPP, presented in Sec. III, will be fully applicable for the E and B families.

We begin with the number density of NPP. In Fig. 5, we show the number of counts for NPP for E and B families, taken from random realisations of best fit Planck 2018 ΛCDM model with tensor-to-scalar ratio $r = 0.05$ and with lensing of E mode included. We used smoothing of the simulated maps with Gaussian kernel with FWHM 0.5 deg.

As it is seen from Fig. 5, the distribution of total number of singular points for unseparated signal and the corresponding E/B families are approximately Gaussian with a small separation of the means for unseparated and E family, and a strong difference between the E and B families. Note a very interesting effect, that the point of maximum of NPP for the E family is slightly bigger than for the unseparated signal. This phenomenon can be understood in terms of the correlation radius r_c from Eq. (9).

For unseparated signal (Q, U), the correlation radius is given by the ratio $r_c = \sigma_0/\sigma_1$, where

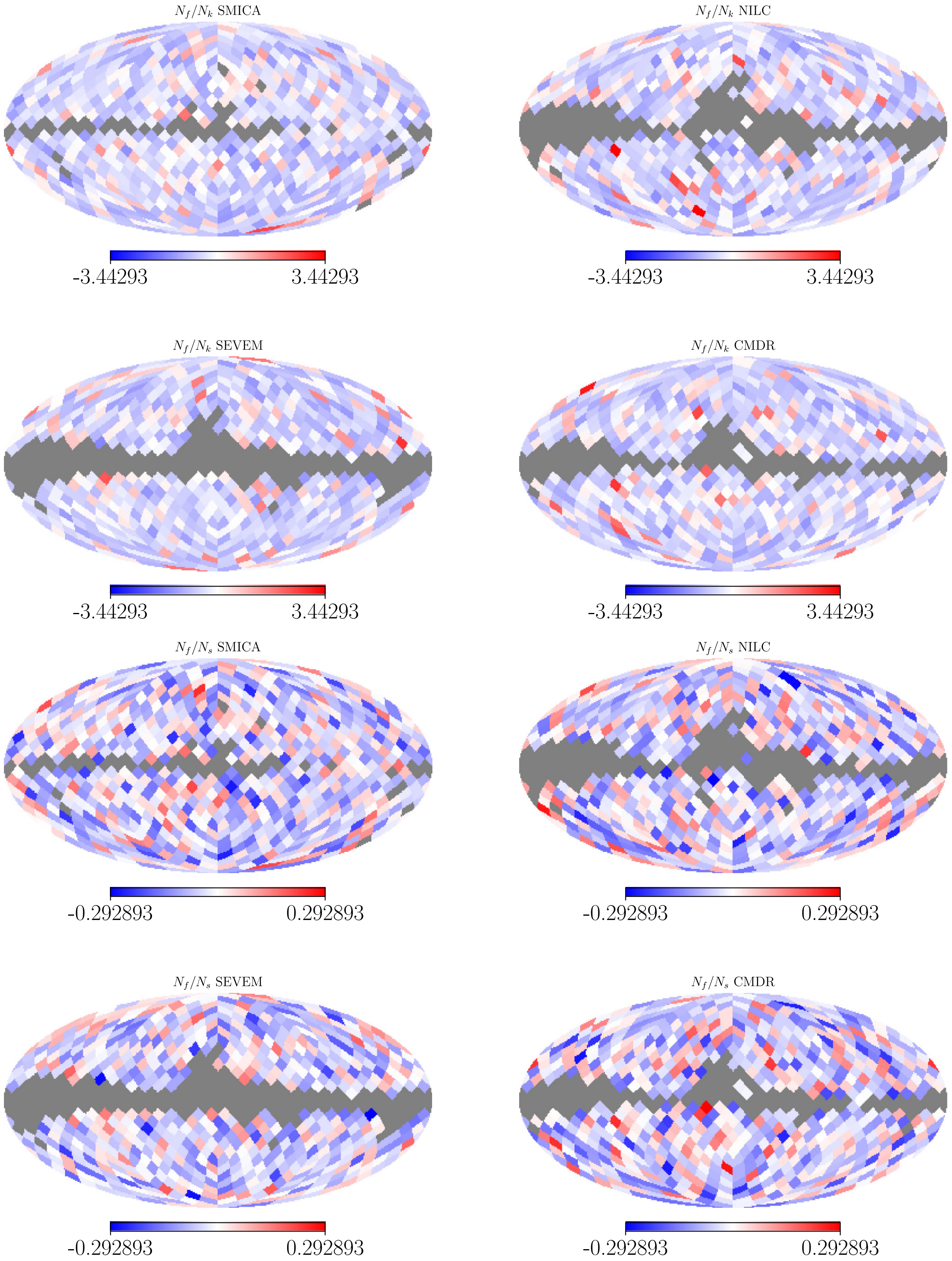

FIG. 4. The ratios N_f/N_k (upper four panels) and N_f/N_s (lower four panels) computed in subareas of $N_{\mathrm{side}} = 8$ using the procedure described in Sec. III, expressed as absolute differences from the theoretical ratios $N_f/N_k = \sqrt{2} + 1$ and $N_f/N_s = \sqrt{2}/2$. The four Planck CMB products SMICA, NILC, SEVEM, and Commander are shown. The area-preserving masking has been applied as described in the text; masked mother pixels are shown in gray.

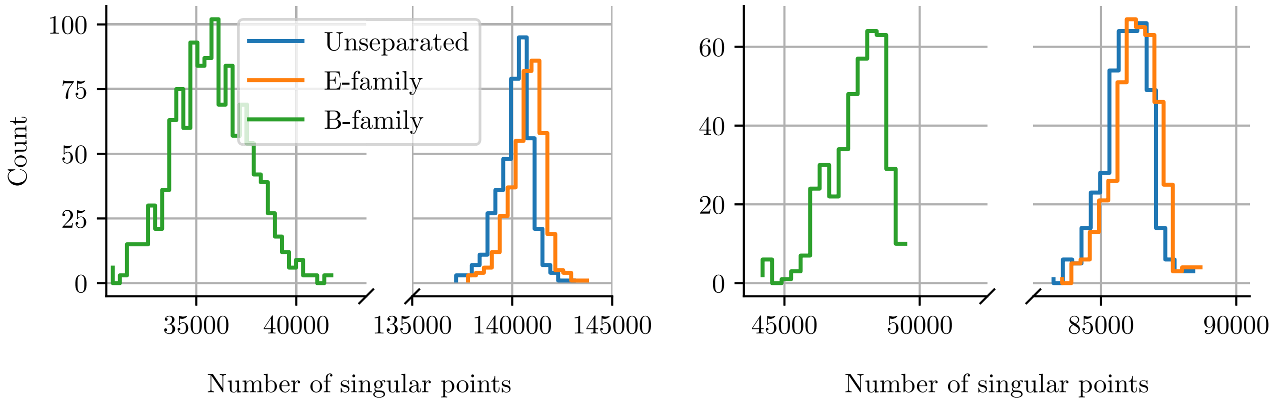

FIG. 5. Range of the number of nonpolarized points in simulations with $r = 0.05$ and 0.5-deg smoothing (see Sec. IV for details). The blue lines are for the full unseparated polarization map, and the orange lines are for the E family only. The B family is shown with green lines in the left-offset panels. The left pair of panels are based on simulations without lensing, and the right pair of panels include lensing.

$$\Delta^2 = 2\sigma_0^2 = \langle (Q_E + Q_B)^2 + (U_E + U_B)^2 \rangle = \Delta_E^2 + \Delta_B^2, \tag{13}$$

$$\sigma_1^2 = \gamma_E^2 + \gamma_B^2, \tag{14}$$

and $\Delta_E^2 = \langle Q_E^2 \rangle + \langle U_E^2 \rangle$ and $\Delta_B^2 = \langle Q_B^2 \rangle + \langle U_B^2 \rangle$ are the corresponding variances for amplitudes of E/B families, and γ_E^2, γ_B^2 are the variances for the first derivatives: $\gamma_E^2 = 4(\langle (Q_E')^2 \rangle + \langle (U_E')^2 \rangle)$ and $\gamma_B^2 = 4(\langle (Q_B')^2 \rangle + \langle (U_B')^2 \rangle)$. So, for the whole sky the total number of unseparated NPP can be represented as follows:

$$\begin{aligned} N_{\mathrm{tot}} &= \frac{\gamma_E^2 + \gamma_B^2}{\Delta_E^2 + \Delta_B^2} \simeq N_{\mathrm{tot}}^E \left(1 + \frac{\gamma_B^2}{\gamma_E^2} - \frac{\Delta_B^2}{\Delta_E^2}\right) \\ &= N_{\mathrm{tot}}^E \left[1 - \frac{\Delta_B^2}{\Delta_E^2}\left(1 - \frac{N_{\mathrm{tot}}^B}{N_{\mathrm{tot}}^E}\right)\right]. \end{aligned} \tag{15}$$

Thus, for $\frac{N_{\mathrm{tot}}^B}{N_{\mathrm{tot}}^E} < 1$ the total number of unseparated NPP is less than for the E family, which one can see reflected in the Gaussian simulations presented in Fig. 5. Also, from Eq. (15), we can see that the size of the deviation will grow according to Δ_B^2, and therefore, also according to r, the tensor-to-scalar ratio.

At the end of this section, we would like to discuss the dependence of the total number of NPP for the B family on the tensor-to-scalar ratio r, based on analysis of pure Gaussian simulations. We will exploit the fact that for the primordial CMB signal without lensing of the E mode, the total number of NPP N_{tot}^B depends only on the ratio γ_B^2/Δ_B^2, where both numerator and denominator are proportional to r. Thus, N_{tot}^B should not depend on r without incorporation of the lensing effect.

From a theoretical point of view, we may expect that with weak lensing of the E mode, the number of NPP will be practically the same as without lensing for all realistic values of the tensor-to-scalar ratio r. For the B family, the dependency N_{tot}^B on r is nontrivial due to the following reasons.

Firstly, it should critically depend on the balance between theoretical (without smoothing by the antenna beam or any others filters) correlation radius r_{cmb} and the effective scale of smoothing Θ_{sm}. If $\Theta_{\mathrm{sm}} \geq r_{\mathrm{cmb}}$, then the γ_B^2 parameter is still proportional to r for the CMB B family, and it is affected by the contribution of the lensed E component,

$$\gamma_B^2 \simeq \gamma_{B,\mathrm{cmb}}^2(r) + \gamma_{B,\mathrm{lens}}^2, \tag{16}$$

where the first term corresponds to the cosmological B mode with tensor-to-scalar ratio r, and the second one is for the lensing effect (almost independent of r). The same representation is valid for the variances,

$$\Delta_B^2 \simeq \Delta_{B,\mathrm{cmb}}^2(r) + \Delta_{B,\mathrm{lens}}^2. \tag{17}$$

Taking into account that $\gamma_{B,\mathrm{cmb}}^2(r)$ and $\Delta_{B,\mathrm{cmb}}^2(r)$ both are proportional to r, we can represent these terms in the following way:

$$\Delta_{B,\mathrm{cmb}}^2(r) = \frac{r}{r_*} C, \qquad \gamma_{B,cmb}^2(r) = \frac{r}{r_*} D, \tag{18}$$

where r_* is some arbitrary normalization parameter, and C and D correspond to $\Delta_{B,\mathrm{cmb}}^2(r)|_{r=r_*}$ and $\gamma_{B,\mathrm{cmb}}^2(r)|_{r=r_*}$.

Secondly, the total number of NPP N_{tot}^B critically depends on the ratios between the CMB and the lensing terms in Eqs. (16)–(17). Namely,

$$N_{\mathrm{tot}}^B(r) = \frac{\frac{r}{r_*} D + \gamma_{B,\mathrm{lens}}^2}{\frac{r}{r_*} C + \Delta_{B,\mathrm{lens}}^2}, \tag{19}$$

whence

$$N^B_{\rm tot}(r)\underset{r\to 0}{\longrightarrow}\frac{\gamma^2_{B,\rm lens}}{\Delta^2_{B,lens}}=N_{B,\rm lens} \tag{20}$$

in the small-r limit, and

$$N^B_{\rm tot}(r)\underset{r\to\infty}{\longrightarrow}\frac{D}{C}=N_{B,\rm cmb} \tag{21}$$

when r is large. Thus, the entire range of variation of $N^B_{\rm tot}(r)$ is reduced to a transition between two constants. The details of this transition can be traced as follows. Let us assume that the CMB tail of Eq. (19) dominates over the lensed part: $\frac{r}{r_*}D\gg\gamma^2_{B,\rm lens}$ and $\frac{r}{r_*}C\gg\Delta^2_{B,\rm lens}$. In this case, the total number of NPP is given by the following asymptotic:

$$\frac{N^B_{\rm tot}(r)}{N_{B,\rm cmb}}\simeq 1-\frac{\Delta^2_{B,\rm lens}}{C}\left(\frac{r_*}{r}\right)\left(1-\frac{N_{B,\rm lens}}{N_{B,\rm cmb}}\right). \tag{22}$$

Here $N_{B,\rm lens}/N_{B,\rm cmb}\ll 1$ and the total number of NPP is slightly smaller than for pure CMB.

From a practical point of view, of greatest interest is the asymptotics $N_{B,\rm lens}/N_{B,\rm cmb}\gg 1$, when the structure of the polarized B signal is completely determined by the lensing effect of E mode. In the absence of delensing, this effect will dominate at $r\ll 0.01$, which is a target for the forthcoming CMB experiments. Although delensing will change the nature of this asymptotic, note that delensing is currently possible in the power spectrum domain but not in the map domain where the NPP analysis naturally works. In the future, NPP analysis might be useful to test the quality of delensing methods.

For $N_{B,\rm lens}/N_{B,\rm cmb}\gg 1$, we get

$$\frac{N^B_{\rm tot}(r)}{N_{B,\rm lens}}\simeq 1-\left(\frac{r}{r_*}\right)\frac{C}{\Delta^2_{B,\rm lens}}\left(1-\frac{N_{B,\rm cmb}}{N_{B,\rm lens}}\right). \tag{23}$$

Confirmation of the asymptotics in Eq. (23) can be seen in Fig. 6, where we give results of numerical simulations for total numbers of NPP for the unseparated signal, E and B families with and without lensing. To generate an ensemble of Gaussian realizations, we used the following parameters from Planck 2018 date release in standard notation [15]: $H_0=67.5$ km/sMpc, $\Omega_b h^2=0.022$, $\Omega_c h^2=0.122$, $A_s{=}210^{-9}$, $\tau=0.06$, and $n_s=0.965$.

V. NONPOLARIZED POINTS IN THE 2018 PLANCK MAPS

A. Contamination of the 2018 Planck maps

Both the theory and computational methods for counting nonpolarized points can now be applied to the latest Planck release. We also get a baseline for these methods by comparison with many Gaussian realizations from a best-fit power spectrum with $r=0.05$. With the method of NPP, we can measure how well each Planck map matches the Gaussian expectation. The Kullback-Leibler entropy [16] was also used as an indicator for how well the Planck maps matched the Gaussian realizations.

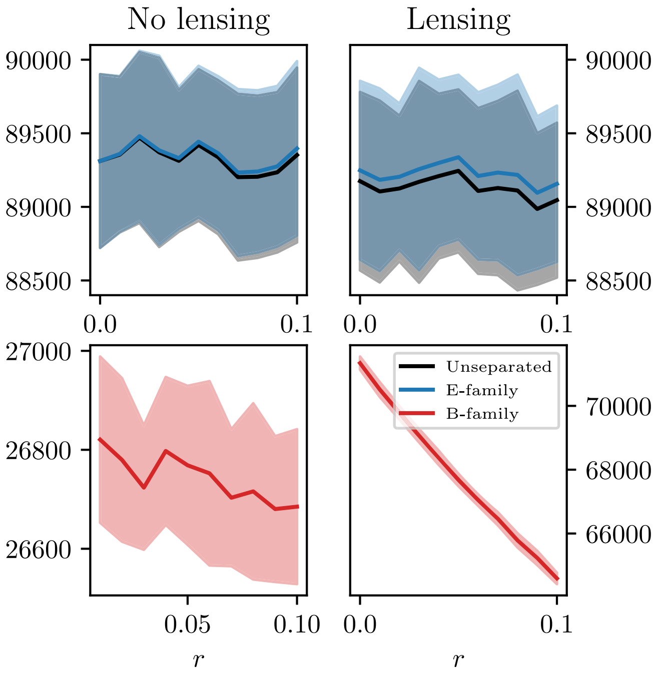

FIG. 6. The results of numerical simulation for total numbers of NPP for unseparated signal (gray), E (blue), and B families (red) with and without lensing. The left column corresponds to the unlensed case and the right column is for lensed. The width of the contours corresponds to 68% confidence level. The solid lines correspond to the mean over realizations.

An important feature of Planck's polarization data is a relatively small signal-to-noise ratio (SNR) for the E mode (SNR ∼ 1) and absence of detection for the primordial B mode. In addition, when analyzing the Planck data, one cannot ignore the effects of systematics, which can interfere with the instrumental noise and the foreground residuals. For illustration of the morphology of these noncosmological components of the derived Planck 2018 CMB products in Fig. 7, we show the intensity of the difference between SMICA and Commander maps, defined as

$$\Delta I=\sqrt{(Q_{\rm SMICA}-Q_{\rm Comm})^2+(U_{\rm SMICA}-U_{\rm Comm})^2} \tag{24}$$

Both these maps were taken with maximal resolution $N_{\rm side}=2048$. In the map of difference SMICA-Commander, the cosmological signal is removed, and this map clearly illustrates the interference between the instrumental noise and scan strategy patterns. Note that reduction of angular resolution down to $N_{\rm side}=128$ decreases the amplitude of the noise but preserves the morphology of contaminants.

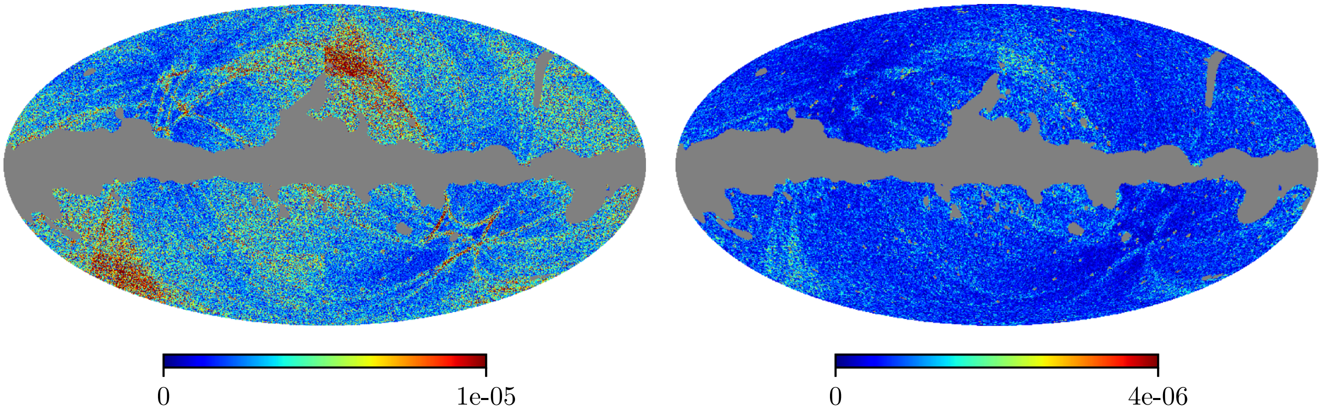

FIG. 7. The maps of differences between SMICA and Commander, i.e., $\sqrt{(Q_{\mathrm{SMICA}} - Q_{\mathrm{Comm}})^2 + (U_{\mathrm{SMICA}} - U_{\mathrm{Comm}})^2}$, for $N_{\mathrm{side}} = 2048$ (left panel) and $N_{\mathrm{side}} = 128$ (right panel).

Thus, in reality, applying NPP analysis to the Planck CMB products, we are not directly testing Gaussianity of the primordial CMB alone but also potential contamination of the component separation procedure. As a simple model, we can write the measured polarization field as

$$Q_{\mathrm{pl}} = Q_{\mathrm{cmb}} + Q_R + Q_{\mathrm{noise}} = Q_{\mathrm{cmb}} + Q_N,$$
$$U_{\mathrm{pl}} = U_{\mathrm{cmb}} + U_R + U_{\mathrm{noise}} = U_{\mathrm{cmb}} + U_N, \qquad (25)$$

where Q_{cmb}, U_{cmb} is the primordial signal, Q_R, U_R are the residuals of systematics and foregrounds, and $Q_{\mathrm{noise}}, U_{\mathrm{noise}}$ are noise. The NPP we measure are those where $Q_{\mathrm{pl}} = 0, U_{\mathrm{pl}} = 0$, i.e., $Q_{\mathrm{cmb}} = -Q_N, U_{\mathrm{cmb}} = -U_N$. Any nonzero residuals or noise will therefore perturb the location of the NPP compared to the primordial signal. Also, note that if the residuals are correlated with the CMB component or with the noise component, this can influence the location and density of NPP in the Planck maps. As we have pointed out in Sec. II, the effect of existence of NPP is not local (see Fig. 1). In the vicinity of each CMB NPP, the morphology of the signal can be represented as follows:

$$Q_{\mathrm{pl}} \simeq Q'_x x + Q'_y y + Q_N(x, y),$$
$$U_{\mathrm{pl}} \simeq U'_x x + U'_y y + U_N(x, y), \qquad (26)$$

where x and y correspond to the θ and ϕ coordinates around NPP, centered at $x = 0$, $y = 0$, and Q'_x, Q'_y, U'_x, U'_y are the corresponding derivatives at that point.

At the point $x = 0$, $y = 0$, the CMB component vanishes, while $Q_N(0,0) \neq 0, U_N(0,0) \neq 0$. Thus, at CMB NPP the structure of the polarization for Planck noisy maps will be destroyed by the noise. However, linear behavior of the CMB signal still will be preserved in the area around NPP, when the terms linear in x, y will dominate over the noise component. Qualitatively, the corresponding scale can be estimated as follows. The linear terms for Q_{pl}, U_{pl} in Eq. (26) are of the order of

$$A \sim \sigma_1 \eta \simeq \sigma_0 \frac{\eta}{r_c}, \qquad (27)$$

where η stands for the x or y coordinates. The noise term in Eq. (26) is about σ_N, where σ_N is rms of noise. Thus $A > B$ leads to the following constraint on η:

$$\eta \geq r_c \frac{\sigma_N}{\sigma_0} \sim r_c (\mathrm{SNR})^{-1}. \qquad (28)$$

Thus, if signal-to-noise ratio SNR $\gg 1$, the structure of the signal around CMB NPP is very well detectable for

$$r_c (\mathrm{SNR})^{-1} \leq \eta \leq n r_c, \qquad (29)$$

where n is 1 to 3, and marginally, it can be seen even for SNR ~ 1 to 2.

In Fig. 8, we show some $2.5° \times 2.5°$ patch of the Planck 2018 Commander polarization map, decomposed into E and B families and smoothed by a Gaussian filter with FWHM of 0.5 deg. The maps of intensities for the E and B families indicate the position of the NPP as points of minima, while the line segments show the polarization angles. The two bottom maps illustrate the morphology of the maps of polarization angles of the E and B families in detail. From the maps of intensity (the top row), one can find that all zones colored by deep blue correspond to the NPP. The bottom row maps show the anomalies of the polarization angle in the vicinity of these points in full agreement with theoretical expectations (for comparison, see Fig. 1). Taking into account that $2.5° \times 2.5°$ maps in Fig. 8 correspond to the fraction of the sky $f_{\mathrm{sky}} \simeq 1.5 \times 10^{-4}$, from Fig. 5, one can find that the number of NPP in Fig. 8 is in agreement with theoretical expectations as well.

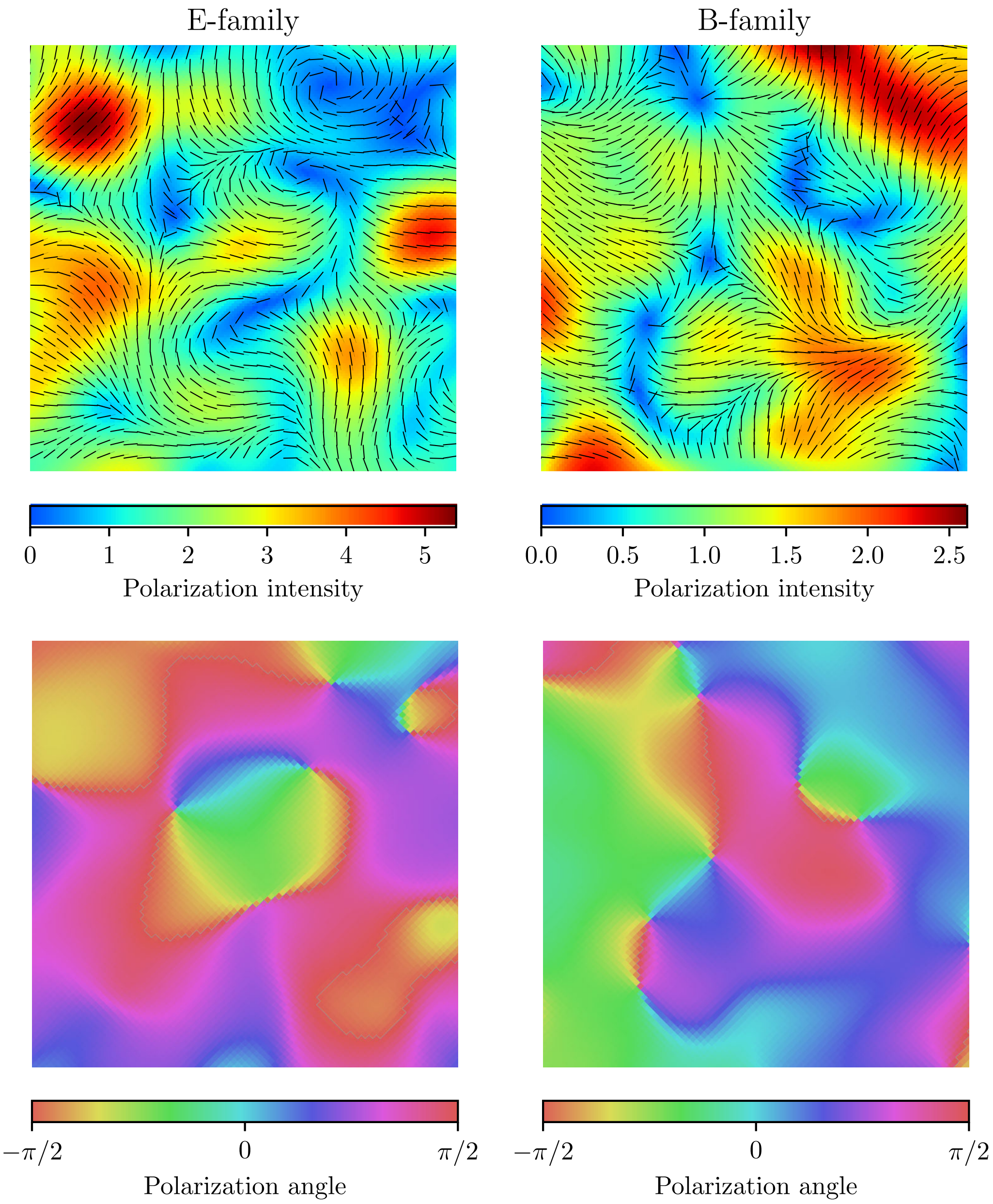

FIG. 8. Top row: E and B families of the Commander polarization maps for Q and U. The background of the maps corresponds to the intensity for each family (in μK). Orientation of the Stokes vector is indicated by black lines. The Commander map has angular resolution $N_{\mathrm{side}} = 2048$ with Gaussian smoothing by 0.5°. The bottom row shows the polarization angle for the same maps, correspondingly. The nonpolarized points are clearly visible in this representation.

B. Ratio distributions for the Planck maps

In this section, we will address the problem of ratios for knots, foci, and saddles [see Eq. (10)] for the Planck 2018 SMICA, Commander, NILC, and SEVEM maps. An important feature of this test is that it depends only on statistical properties of the signal (CMB plus noise), and if noise is the dominating part of the signal, it will reveal important information about itself. For evaluation of statistical significance of the ratio test, it is possible to use any simulations: for example, Planck FFP 10 or pure Gaussian.

For the ratio test, all Planck maps and Gaussian simulations of the sky were smoothed by Gaussian filter with FWHM 0.5 deg. and divided into subareas (mother pixels) of HEALPix pixels at N_{side} 8 and 16. In each subarea outside the common mask, all NPP were counted, and ratios were evaluated. These were used to construct density histograms (see Fig. 9) with the aim of comparing them to Gaussian simulations. An important choice in creating histograms is a choice of the number of bins.

The simple Sturges [17] estimator of $\log_2(n) + 1$ was used to provide a suitable number of bins for analyzing the

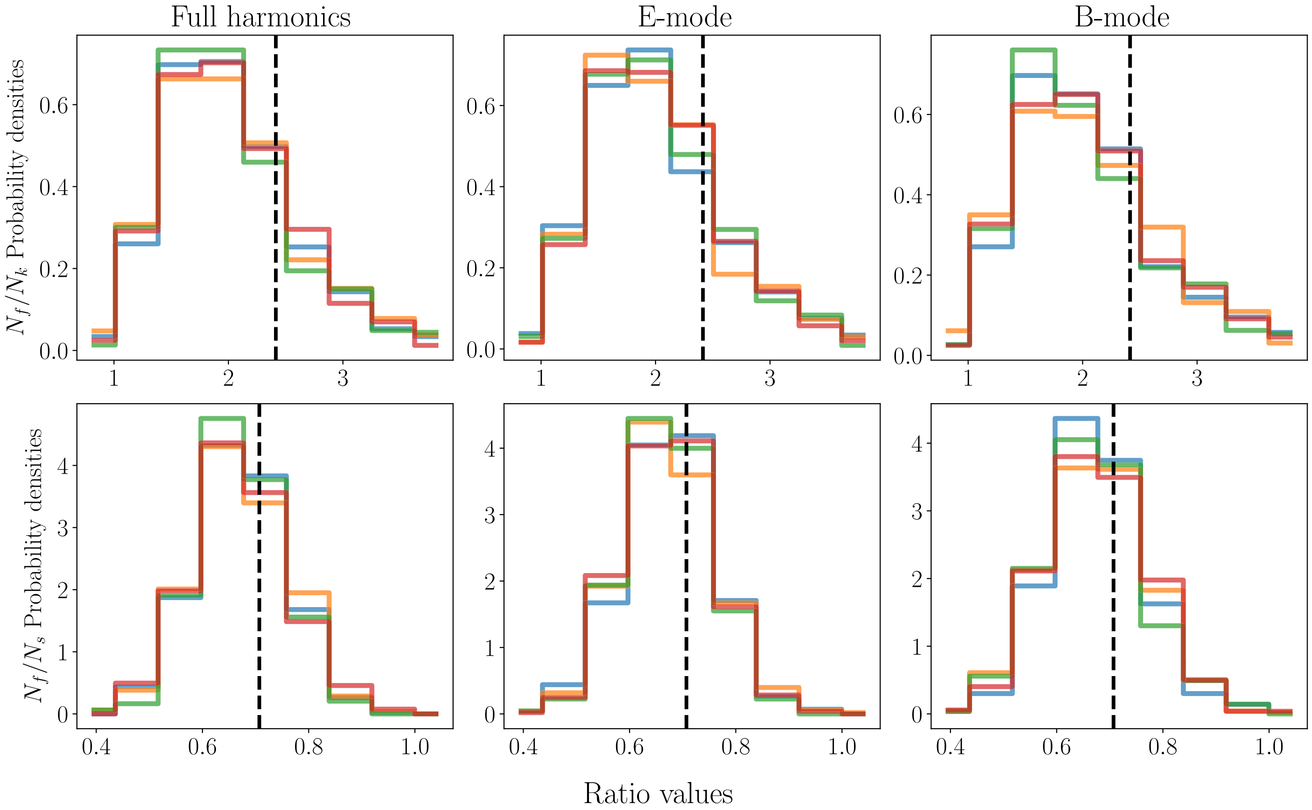

FIG. 9. Ratio distributions for skies decomposed into separate sets of E/B families. From left to right each column corresponds to full harmonics, the E mode, and the B mode. The top row corresponds to the f/k ratio, and the bottom to the f/s ratio. Each color corresponds to one of the Planck maps. Blue is SMICA, yellow is NILC, green is SEVEM, and red is Commander.

data, where n is the number of data points available. The Doane estimator [18] was also tried to compensate for possible skewing in the data, but it gave the same results as Sturges, so the latter was used. Each subarea contributes one data point for analysis, and given the values of N_{side} we were working at, we had at most either 768 ratio values for $N_{\text{side}} = 8$ or 3072 ratio values for $N_{\text{side}} = 16$. Based on this, the Sturges estimator provides 9 to 11 as the range for suitable numbers of bins. Bin edges were chosen accordingly to encompass every realization of the sky under analysis. A cutoff value for the ratio value was used to ignore statistical outliers, subareas with an unlikely small number of NPP.

Constructing these ratio distributions show that they are close to the theoretical expectations for a Gaussian random process. Precise properties of each distribution are given in Table I.

C. Comparison with Gaussian distributions

As an empirical test for the Planck maps, the Kullback-Leibler entropy between each of the Planck maps and a background Gaussian was calculated. The KL entropy is a measure of the divergence between a sample probability distribution P and a background distribution Q. For discrete distributions, it is defined as

$$\Theta_{KL}(P|Q) = \sum_i P_i \ln\left(\frac{P_i}{Q_i}\right). \tag{30}$$

Hence, the closer each P_i is to each Q_i, the smaller the value of $\Theta_{KL}(P|Q)$ will be and the better the match between the two distributions. As another way to compare the Planck distribution functions to the distribution functions from the Gaussian simulations, a Kolmogorov-Smirnov test was performed. The Kolmogorov-Smirnov estimator between two samples is defined as the maximum

TABLE I. Statistical information for all Planck maps and a ratio distribution from many Gaussian simulations. All data are calculated from the density histograms in Fig. 9, at the same fidelity of nine bins. Given are the expectation value (ex) and standard deviation (stdev). The theoretical ratio values to as many significant digits are 2.4142 and 0.7071 for the f/k and f/s ratios, respectively.

Map name	ex (f/k)	stdev (f/k)	ex (f/s)	stdev (f/s)
Gaussian	2.0349	0.3240	0.6763	0.0071
SMICA	2.0149	0.3357	0.6716	0.0076
NILC	2.0143	0.3756	0.6719	0.0084
SEVEM	1.9934	0.3377	0.6702	0.0067
CMDR	2.0073	0.3190	0.6712	0.0087

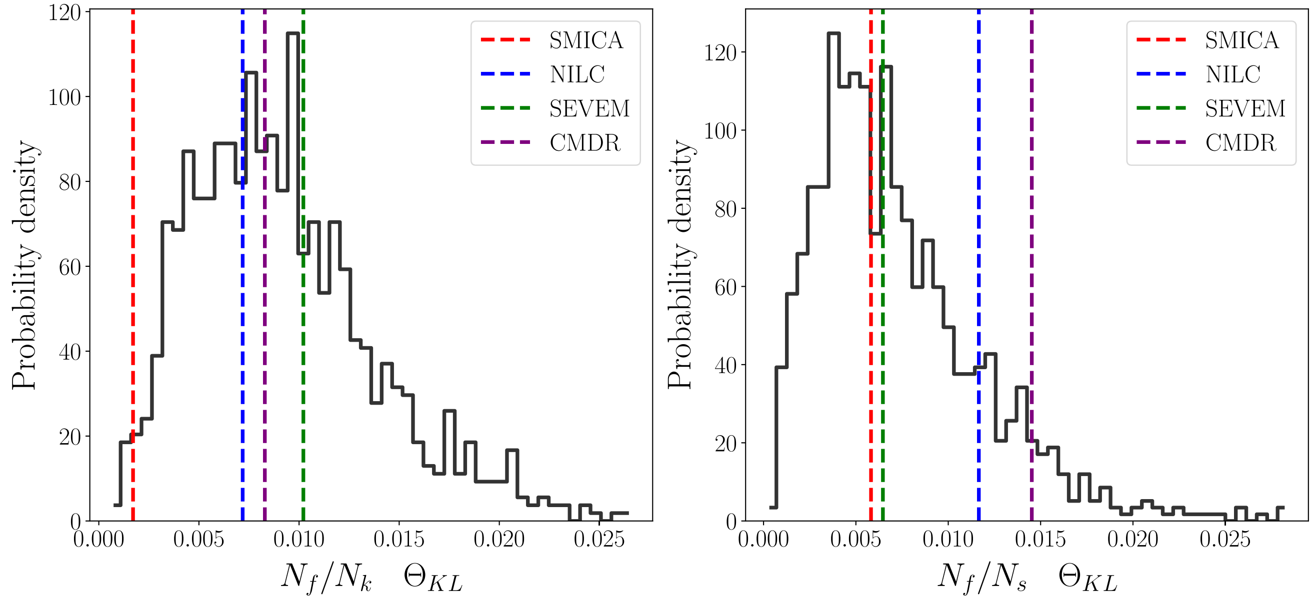

FIG. 10. The vertical dotted lines are KL-entropy values of the Planck ratio distributions against an average background from Gaussian simulations. The black line is the distribution of the KL-entropy values of 1035 unique pairings of many Gaussian realizations. These allow us to calculate p values for finding entropies of the Planck maps at least that large.

distance between their cumulative distribution functions [19]. This quantity follows the Kolmogorov distribution depending on the size of the samples, from which each KS score can be converted to a p value can be calculated. The bigger the KS score, the greater inferred difference between the two underlying distribution functions.

The KL entropy is a logarithmic estimator based on exponential probability densities that have positive values in the whole domain. Due to this, discrete bins that end up having no data points for some realizations of the sky contribute infinities to the entropy sum. To deal with this, we keep track of the set of bins where a zero has occurred and exclude those when calculating the KL entropy. Other methods like minimally smoothing the distributions and doping each bin with an extra data point were considered, but they all performed the same. As such, dropping the zero bins was the preferred method. This issue does not apply to the KS estimator, which, being based on the cumulative distribution function, is unaffected by null bins.

TABLE II. Data from all Planck maps. p values give the probability of finding a KL entropy or KS statistic at least that large.

KL	f/k KL entropy	f/k p value	f/s KL entropy	f/s p value
SMICA	0.0017	0.9884	0.0058	0.5044
NILC	0.0072	0.6135	0.0117	0.1560
CMDR	0.0083	0.5130	0.0145	0.0713
SEVEM	0.0102	0.3324	0.0065	0.4661

KS	f/k KS score	f/k p value	f/s KS score	f/s p value
SMICA	0.0400	0.2076	0.0459	0.1017
NILC	0.0380	0.3248	0.0466	0.1317
CMDR	0.0424	0.1859	0.0506	0.0686
SEVEM	0.0662	0.0095	0.0518	0.0747

The background Gaussian is created by averaging the distributions from 46 unique Gaussian simulations. To give meaning to the entropies calculated in that way, a background distribution for them was also realized. The 46 Gaussian distributions allow us to match up 1035 unique pairs and calculate the KL entropy of each pairing. A density distribution of the background entropy values provides the opportunity of finding a p value for each measured entropy in the Planck maps. These results are visualized in Fig. 10 and Table II. Table II also shows the results of the KS test.

VI. CONCLUSIONS

The nonpolarized points in the CMB polarization field can be classified according to the local geometry of the field, which is converged in the points of local minima of the intensity and anomalies of the polarization angle. Gaussianity predicts that the different types of NPP should occur in certain fixed ratios. By measuring the variation of these ratios in different subareas of the sky, one can construct a sensitive test of non-Gaussianity and statistical anisotropy, complementary to [4,20].

We separate the Q and U Stokes parameters into the E and B families and applied the NPP theory to these families. Our analysis revealed the regularities of their statistics depending on the amplitude of the tensor-to-scalar ratio r and lensing effect. We have shown that in the absence of lensing, the total number of NPPs of all types does not depend on the tensor-to-scalar ratio, while the lensing effect removes this degeneracy. Thus, the analysis of NPP is a powerful tool for investigating the manifestation of the effect of lensing for the next generation of CMB experiments devoted to detect or constrain cosmological models with $r \leq 10^{-2}$. The same test can be used for detection of E/B leakage corrections and study of the performance of delensing methods.

We have applied theory of NPP to the Planck 2018 CMB polarization maps and showed, that there is general consistency of SMICA and NILC compared to a reference set of Gaussian simulations according to the KL entropy. The strongest discrepancies are found in Commander ($p \approx 0.07$) and NILC ($p \approx 0.15$). Although these discrepancies are not highly significant, they may point to the presence of foreground, and component separation residuals in these two maps compared to SMICA and SEVEM.

The KS nonparametric test for equality of distributions can also be used. Overall, the KS test also reveals a general consistency with the KL entropy, except for SEVEM, whose f/k ratios are peculiar at the $p < 0.01$ level. Similarly, the KS test suggests disagreement of the SEVEM f/s ratios at a modest significance of around $p = 0.07$, compared to no result ($p = 0.47$) from the KL entropy. In general, we expect the KS estimator to be more sensitive to local deviations, or deviations in the tails of the distribution functions, compared to the KL entropy. Departures of this kind might therefore exist in the SEVEM ratios. With the upcoming era of high-precision, ground-based and space-based CMB polarization observations, methods exploiting NPP statistics will be useful for characterizing and testing the statistical properties of derived maps.

ACKNOWLEDGMENTS

The HEALPix pixelization [14] scheme was heavily used in this work, and we thank them for their contributions to the field. This research was partially funded by the Villum Fonden through the Deep Space project.

[1] U. Seljak and M. Zaldarriaga, Phys. Rev. Lett. **78**, 2054 (1997).

[2] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997).

[3] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Phys. Rev. D **55**, 7368 (1997).

[4] Y. Akrami *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A7 (2020).

[5] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A5 (2020).

[6] P. D. Naselsky and D. I. Novikov, Astrophys. J. Lett. **444**, L1 (1995).

[7] P. D. Naselsky and D. I. Novikov, Astrophys. J. **507**, 31 (1998).

[8] A. D. Dolgov, A. G. Doroshkevich, D. I. Novikov, and I. D. Novikov, JETP Lett. **69**, 427 (1999).

[9] H. Liu, J. Creswell, and P. Naselsky, J. Cosmol. Astropart. Phys. 05 (2018) 059.

[10] Y. Akrami *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A4 (2020).

[11] H. Liu, J. Creswell, S. von Hausegger, and P. Naselsky, J. Cosmol. Astropart. Phys. 05 (**2019**) 003.

[12] J. Kasak, Code repository for singular point analysis, https://gitlab.com/aramend/libpol (2020).

[13] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A3 (2020).

[14] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, Astrophys. J. **622**, 759 (2005).

[15] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A6 (2020).

[16] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).

[17] H. A. Sturges, J. Am. Stat. Assoc. **21**, 65 (1926).

[18] D. P. Doane, Am. Stat. **30**, 181 (1976).

[19] N. Smirnov, Moscow Univ. Math. Bull. **2**, 3 (1939).

[20] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A16 (2016).