




# Non-Gaussian modelling and statistical denoising of Planck dust polarisation full-sky maps using scattering transforms

J.-M. Delouis<sup>1</sup> , E. Allys<sup>2</sup> , E. Gauvrit<sup>1,2</sup>, and F. Boulanger<sup>2</sup> 

<sup>1</sup> Laboratoire d’Océanographie Physique et Spatiale (LOPS), Univ. Brest, CNRS, Ifremer, IRD, 29200 Brest, France  
e-mail: [jean.marc.delouis@ifremer.fr](mailto:jean.marc.delouis@ifremer.fr)

<sup>2</sup> Laboratoire de Physique de l’École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris-Cité, 75005 Paris, France

Received 21 July 2022 / Accepted 18 October 2022

## ABSTRACT

Scattering transforms have been successfully used to describe dust polarisation for flat-sky images. This paper expands this framework to noisy observations on the sphere with the aim of obtaining denoised Stokes  $Q$  and  $U$  all-sky maps at 353 GHz, as well as a non-Gaussian model of dust polarisation, from the Planck data. To achieve this goal, we extended the computation of scattering coefficients to the HEALPix pixelation and introduced cross-statistics that allowed us to make use of half-mission maps as well as the correlation between dust temperature and polarisation. Introducing a general framework, we developed an algorithm that uses the scattering statistics to separate dust polarisation from data noise. The separation was validated on mock data before it was applied to the SRoll2Planck maps at  $N_{\text{side}} = 256$ . The validation shows that the statistics of the dust emission, including its non-Gaussian properties, are recovered until  $\ell_{\text{max}} \sim 700$ , where, at high Galactic latitudes, the dust power is weaker than that of the dust by two orders of magnitude. On scales where the dust power is weaker than one-tenth of the power of the noise, structures in the output maps have comparable statistics, but are not spatially coincident with those of the input maps. Our results on Planck data are significant milestones opening new perspectives for statistical studies of dust polarisation and for the simulation of Galactic polarised foregrounds. The Planck denoised maps are available (see [http://sroll20.ias.u-psud.fr/sroll40\\_353\\_data.html](http://sroll20.ias.u-psud.fr/sroll40_353_data.html)) together with results from our validation on mock data, which may be used to quantify uncertainties.

**Key words.** techniques: image processing – methods: statistical – submillimeter: ISM – cosmic background radiation

## 1. Introduction

The cosmic microwave background (CMB) is a prime observational probe for constraining cosmological models (Durrer 2015). Today, with the uncertainties in the CMB temperature spectrum essentially reduced to the cosmic variance (Planck Collaboration V 2020; Planck Collaboration XI 2020), the CMB experiments have shifted their focus to polarisation. In particular, accurate measurements of the tensor component ( $B$ -modes) of the polarised signal could provide direct evidence of the inflation period (Guth 1981; Linde 1982). This paramount goal of cosmology is driving the development of ambitious CMB experiments (Abazajian et al. 2016; Ade et al. 2019; LiteBIRD Collaboration 2022), but the potential detection of primordial  $B$ -modes does not only depend on increasing the signal-to-noise ratio on CMB polarisation.

The quest for CMB  $B$ -modes is also hampered by instrumental systematic effects (Planck Collaboration III 2020) and polarised foregrounds dominated by Galactic dust emission (BICEP2/Keck Array and Planck Collaborations 2015; Planck Collaboration XI 2020). In this context, the modelling of systematic effects and Galactic foregrounds must advance alongside the sensitivity of the measurements. This is a major challenge because instrumental systematics and Galactic emission are non-Gaussian signals, which in essence are difficult to model. To address this difficulty, the CMB community has been investing much effort in the development of Galactic emission

models (Delabrouille et al. 2013; Thorne et al. 2017; Vansyngel et al. 2017; Martínez-Solaache et al. 2018; Zonca et al. 2021; Hervías-Caimapo & Huffenberger 2022) combined with instruments models to produce end-to-end simulations of the data, as was done for instance for Planck (Planck Collaboration III 2020). Data simulations are essential to marginalise the inference of cosmological parameters on the nuisance signals and correct for bias (e.g. Vacher et al. 2022) and to perform likelihood-free inference methods (Planck Collaboration Int. XLVI 2016; Alsing et al. 2019; Jeffrey et al. 2022).

Non-Gaussianity is an important characteristic of Galactic foregrounds. To account for it, several authors have introduced machine-learning algorithms (Aylor et al. 2020; Krachmalnicoff & Puglisi 2021; Petroff et al. 2020; Thorne et al. 2021), but these methods need to be trained. Thereby, their use is hindered by the difficulty of building relevant training sets. Magnetohydrodynamics (MHD) simulations of the interstellar medium (Kritsuk et al. 2018; Kim et al. 2019; Pelgrims et al. 2022) are useful to develop the method, but they are far from reproducing the statistics of dust polarisation with the accuracy required for CMB component separation.

Another approach to modelling Galactic foregrounds is to rely on scattering transform statistics. These statistics were introduced in data science to discriminate non-Gaussian textures (Mallat 2012; Bruna & Mallat 2013; Cheng & Ménard 2021a), and they have since been applied to dust emission maps computed from MHD simulations (Allys et al. 2019;

Regaldo-Saint Blancard et al. 2020; Saydjari et al. 2021). Promising results have also been obtained on various astrophysical processes such as large-scale structure density field and galaxy surveys (Allys et al. 2020; Eickenberg et al. 2022; Valogiannis & Dvorkin 2022b,a), weak-lensing convergence maps (Cheng et al. 2020; Cheng & Ménard 2021b), and 21cm data of the epoch of reionisation (Greig et al. 2022). To construct these statistics, convolutions of the input image with wavelets over multiple oriented scales are combined with non-linear operators that allow efficiently characterising interactions between scales.

One notable advantage of the scattering transforms is that generative models that quantitatively reproduce the non-Gaussian structures of a given process can be constructed from a small number of realisations of this process, which can be even a single image (Bruna & Mallat 2019; Allys et al. 2020). This could allow constructing Galactic dust models directly from observational data. For this purpose, the Planck data are a key observational input. They have been used both as a template of the dust sky and to model the spectral energy distribution of dust emission (Thorne et al. 2017; Zonca et al. 2021). However, for polarisation, data noise is a severe limitation that must be circumvented. A new direction was opened by Regaldo-Saint Blancard et al. (2021), who introduced an algorithm that successfully uses scattering statistics to separate dust emission from data noise. They applied it to flat-sky Planck Stokes images at 353 GHz. Their method exploits the very different non-Gaussian structure on the sky of dust emission compared to the data noise. An iterative optimisation on sky pixels yields denoised maps and a generative model of dust polarisation.

This paper aims to extend the innovative component separation approach of Regaldo-Saint Blancard et al. (2021) to the sphere and to apply it to all-sky Planck polarisation maps. Our scientific motivation is to obtain denoised Planck Stokes maps that may be used to model the dust foreground to CMB polarisation. For the astrophysics of dust polarisation, the signal-to-noise ratio limits the range of angular scales that is accessible to study (Planck Collaboration XII 2020). Thus, our work is also a valuable contribution to statistical studies of dust polarisation at high Galactic latitudes.

While our scientific objective is the statistical denoising of dust-polarised emission, we treat this problem as the separation between two components. This differs from usual denoising algorithms that rely for instance on filtering (Wiener 1949; Zaroubi et al. 1994) or sparsity (Starck et al. 2002). In comparison to these usual algorithms, our objective is to recover a map with the correct statistics, even if it differs from the true map at the smallest scales. We also emphasise that the method we present in this paper can be similarly applied to the separation of two components of interest, and not only for denoising.

Our data objective involves several challenges. The scattering coefficients must be computed on all-sky maps in HEALPix format. The algorithm used to separate dust and data noise must accommodate variations in the dust statistics over the sky. The computing time required to compute the scattering coefficients and the map optimisation must be kept manageable. Within this framework, we developed cross-scattering coefficients to make optimal use of the available data, especially complementary half-mission maps. Cross-scattering coefficients have also been introduced by Régaldó-Saint Blancard et al. (2022) to model multi-channel dust data. In doing this, the two papers extended the common use of cross-power spectra to statistics that encode non-Gaussianity. For our project, the cross-correlation of Planck maps with independent noise

realisations facilitates the denoising. It also allows us to take into account the  $TE$  and  $TB$  correlation of dust polarisation (Planck Collaboration XI 2020).

The paper is organised as follows. Section 2 introduces the cross-scattering statistics, the algorithm, and the loss functions from a generic starting-point. Our method is validated on mock data (Sect. 3) before it is applied to the Planck maps (Sect. 4). Applications and perspectives for future work are discussed in Sect. 5. The paper results are summarised in Sect. 6. Additional figures are presented in Appendix A.

## 2. CWST and separation of the dust and noise components

We introduce our method in Sect. 2.1 before we present the cross-scattering transform on the sphere in Sect. 2.2. Next, we describe the algorithm we used to perform the component separation between dust polarisation and data noise (Sect. 2.3).

### 2.1. Introduction

We aim to characterise the statistical properties of the polarised dust emission from noisy Planck all-sky maps in HEALPix (Górski et al. 2005). We ignored the CMB, which can be either removed or neglected, and address this problem as a separation between two components: dust emission and data noise. We chose to work with SR0112Planck polarisation maps at 353 GHz (Delouis et al. 2019). We converted the all-sky Stokes  $Q$  and  $U$  maps into  $E$  and  $B$  maps and applied our method to the latter because they are independent scalars that do not depend on the chosen reference frame. This transformation being non-local, signal from the brightest areas in the Galactic plane contaminates the  $E$  and  $B$  maps at high Galactic latitude. Thus, to display the results, we transformed the denoised  $E$  and  $B$  maps back to  $Q$  and  $U$ . In doing so, we also conformed to the standard way of representing the polarised sky. In practice, we found that the use of  $E$  and  $B$  is better for the present work.

Our data-processing was the same for  $E$  and  $B$  maps. In each case, we had access to a full-mission map  $d$  and two half-missions  $d_1$  and  $d_2$ . The half-mission maps were computed from the first and second part of the mission, respectively, making their noises and time-variable instrumental systematics mostly independent. These maps can be written as the sum of the dust emission  $s$  and three different noise realisations that we call  $n$ ,  $n_1$ , and  $n_2$ . For instance,

$$d = s + n \quad (1)$$

for the full mission, and

$$d_1 = s + n_1 \quad (2)$$

for the first half mission. We assumed that the noises for the two half-missions can be considered to be statistically independent, at least for  $\ell > 30$  (see e.g. Delouis et al. 2019). Based on a simulation effort, we also assumed that we have access to a large number of realistic noise maps. The SR0112 dataset contains for instance an ensemble of 500  $(\tilde{n}, \tilde{n}_1, \tilde{n}_2)$  associated triplets of maps (Delouis et al. 2019) that can be used to simulate the noises of the full mission and of the two half-missions of the  $E$ ,  $B$ , and  $T$  maps. Finally, we also had access to an intensity map that we labelled  $T$ , which we used to characterise the statistical dependence between polarised and total intensity dust emission.

The details of the different data we used and the associated assumptions are discussed in Sects. 3 and 4.

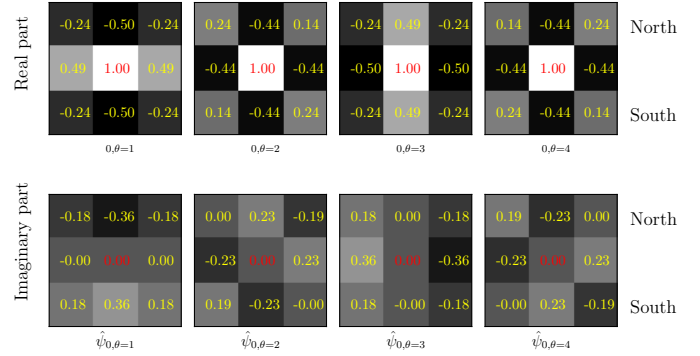
To obtain the statistical properties of the polarised dust emission, we performed a component separation of the dust and noise maps, using their different non-Gaussian characterisation as a lever arm (Regaldo-Saint Blancard et al. 2020). This component separation generates a new dust maps  $\tilde{s}$  from a gradient descent with an ensemble of statistical constraints built from cross wavelet scattering transforms (CWST) between different maps. The statistics of the polarised dust map  $s$  was then estimated from this  $\tilde{s}$  map. In the following, we call this component separation method function of cleaning using statistics, FoCUS.

## 2.2. Cross wavelet scattering transform

Scattering transforms (ST) are a recently developed type of non-Gaussian summary statistics (Mallat 2012; Bruna & Mallat 2013). The inspiration for them comes from convolutional neural networks, but they do not need any training stage to be computed (i.e. the function for computing the statistics from a set of data can be written explicitly and does not need to be learnt). They thus benefit from the low-variance efficient characterisation typical of neural networks, but give some level of interpretability through their explicit mathematical form. Several sets of ST statistics have been constructed, such as the wavelet scattering transform (WST; Allys et al. 2019; Regaldo-Saint Blancard et al. 2020) or the wavelet phase harmonics (WPH; Mallat et al. 2018; Allys et al. 2020). The construction of the ST statistics relies on two main features: scale separation (through wavelet transforms), and characterisation of the interaction between scales using non-linearity (e.g. modulus, ReLU, or phase acceleration). The CWST introduced in this paper is a new type of ST cross-statistics constructed for data defined on the sphere. It is an extension of the WST, to which it is reduced when it is used as auto-statistics.

The first building block of the CWST transform is a wavelet transform allowing fast computation directly on HEALPix data. For this purpose, we introduced a very simple multi-resolution wavelet transform defined from four  $3 \times 3$  complex kernels, which are used to compute convolutions with HEALPix maps at different resolutions. These kernels, which are called  $\tilde{\psi}_{j,\theta}$  with  $\theta$  between 1 and 4, are plotted in Fig. 1. With respect to the HEALPix conventions,  $\theta = 1$  refers to a north-south brightness oscillation associated with an east-west elongated structure. The convolution is computed by multiplying the weights of Fig. 1 to the eight neighbours of each pixel in such way that the top left value is related to the north-west HEALPix pixel. In the particular case when a pixel has only seven neighbours, the value taken for the eighth missing pixel corresponds to the value of the closest anti-clockwise pixel (e.g. north-west to replace north). With a resolution  $N_{\text{side}}$ , one pixel occurs every  $N_{\text{side}}^2$  pixels.

The different wavelets  $\psi_{j,\theta}$  are labelled on a integer scale  $j$  going from 0 to  $J - 1$  and by an integer angle  $\theta$  going from 1 to  $L$  (associated with a  $(\theta - 1) \cdot \pi/L$ ) for a total of  $J \cdot L$  wavelets. On an  $N_{\text{side}} = 256$  map, we used  $J = 8$  and  $L = 4$  (since the four  $\tilde{\psi}_{j,\theta}$  kernels define four orientations). The wavelet transform for  $j = 0$  was obtained through the convolution of the input map in HEALPix with the  $\tilde{\psi}_{0,\theta}$  kernels. The wavelet transform for  $j = 1$  was then obtained by first sub-sampling the input map by computing  $2 \times 2$  mean through the HEALPix nested indexing property, and by then computing the convolutions again with the  $\tilde{\psi}_{0,\theta}$  kernels. By repeating this process, we can compute convolutions up to scales  $j = J - 1$ .



**Fig. 1.** Representation of the four  $\tilde{\psi}_{0,\theta}$   $3 \times 3$  kernels. *First row:* real part of the kernel coefficients. *Second row:* there imaginary parts. The right labels show how the coefficients are applied to the HEALPix neighbours. In this figure, north (as defined by HEALPix) is up, west is left, and so on.

As the  $\tilde{\psi}_{0,\theta}$  kernels have a 2-pixel wavelength, the characteristic pic multipole probed by the wavelet transform at scale  $j$  is  $\ell \approx 512 \cdot 2^{-j}$ . Starting from an initial resolution of  $N_{\text{side}} = 256$  at  $j = 0$ , the resolution on which the convolution is performed at  $j > 0$  is  $N_{\text{side}} = 256 \cdot 2^{-j}$ . Moreover, the initial value of  $N_{\text{side}}$  obviously limits the number of scales that can be considered: here  $j = 7$ , which corresponds to a HEALPix map with  $N_{\text{side}} = 2$ .

We are aware of the fact that a more refined way to compute a wavelet transform on the sphere exists (see e.g. Leistedt et al. 2013; McEwen et al. 2015, 2018) from which a first implementation of scattering transform on the sphere has already been defined (McEwen et al. 2021). We have chosen to conduct this project within our scheme because it is simple to use and because of its computational efficiency for GPU-accelerated computations, especially for the implementation of new cross-statistics. We would like to transition to better wavelets in the future, however, for which we expect an improvement of the obtained results.

In the following, we use the index  $\lambda = (j, \theta)$  to describe the oriented scale associated with each  $\psi_{\lambda}$ , and we keep implicit the fact that they are defined on HEALPix maps of different  $N_{\text{side}}$ . The wavelet convolution of an image  $I$  with a  $\psi_{\lambda}$  wavelet then reads  $I \star \psi_{\lambda}(p)$ , where  $p$  is the coordinate on the sphere at the corresponding resolution.

The CWST cross-statistics are calculated on two maps  $I_a$  and  $I_b$ . Similarly to the usual WST, the CWST contains two layers of coefficients, which are characterised by one or two oriented scales  $\lambda_i$ . The whole set of statistics, called  $S(I_a, I_b)$ , is thus decomposed into the  $S_1$  coefficients at the first layer and the  $S_2$  coefficients at the second layer. When  $I_a = I_b$ , these coefficients are the usual WST coefficients (Bruna & Mallat 2013).

The coefficients at first order are called  $S_1(I_a, I_b)_{\lambda_1}$ . They are constructed from a product of convolutions of  $I_a$  and  $I_b$  at the same  $\lambda_1$  scale,

$$C_1(I_a, I_b)_{\lambda_1}^{\Re} = \text{sign}(\Re(I_a \star \psi_{\lambda_1} \cdot I_b \star \psi_{\lambda_1}^*)) \cdot \sqrt{|\Re(I_a \star \psi_{\lambda_1} \cdot I_b \star \psi_{\lambda_1}^*)|}, \quad (3)$$

and

$$C_1(I_a, I_b)_{\lambda_1}^{\Im} = \text{sign}(\Im(I_a \star \psi_{\lambda_1} \cdot I_b \star \psi_{\lambda_1}^*)) \cdot \sqrt{|\Im(I_a \star \psi_{\lambda_1} \cdot I_b \star \psi_{\lambda_1}^*)|}, \quad (4)$$



where we independently considered the real and imaginary part of the products of wavelet convolutions, and where  $*$  and  $|\cdot|$  stand for the complex conjugate (acting here on the whole  $I_b \star \psi_{\lambda_1}$  term) and absolute value, respectively. The square root allows us to recover the L1-like norm, which is useful to decrease the variance of the estimators, but introduces a bias when the correlated information between two noisy data sets is computed.

The  $S_1$  coefficients are then computed from a spatial integration,

$$S_1(I_a, I_b)_{\lambda_1} = \left\langle C_1(I_a, I_b)_{\lambda_1}^{\Re} + i \cdot C_1(I_a, I_b)_{\lambda_1}^{\Im} \right\rangle_{\text{pixels}}, \quad (5)$$

where  $i$  is the imaginary unit, and where the brackets stand for a spatial average on the sphere, which is multiplied by  $2^j$  for a uniform normalisation meaning that the  $S_1$  coefficients of white noise are constant across all scales. The  $S_1$  coefficients can also be used to compute the statistics of a single map  $I_a = I_b = I$ . In this case, the  $C_{\lambda_1}^{\Im}$  term obviously vanishes, the  $C_{\lambda_1}^{\Re}$  terms are the complex modulus of the  $I \star \psi_{\lambda_1}$  convolution, and we recover the standard WST definition  $S_1 = \langle |I \star \psi_{\lambda_1}| \rangle_{\text{pixel}}$  for  $S_1$ .

The coefficients at second order are called  $S_2(I_a, I_b)_{\lambda_1, \lambda_2}$ . They are constructed from two positive and negative terms,

$$\begin{aligned} C_{2,+}^{\Re}(I_a, I_b)_{\lambda_1, \lambda_2} &= \left| \psi_{\lambda_2} \star \text{ReLU} \left( C_1^{\Re}(I_a, I_b)_{\lambda_1} \right) \right|, \\ C_{2,-}^{\Re}(I_a, I_b)_{\lambda_1, \lambda_2} &= \left| \psi_{\lambda_2} \star \text{ReLU} \left( -C_1^{\Re}(I_a, I_b)_{\lambda_1} \right) \right|, \end{aligned} \quad (6)$$

and similarly for the imaginary terms  $C_{2,\pm}^{\Im}$ . The  $S_2$  terms are then obtained through spatial integration,

$$\begin{aligned} S_2(I_a, I_b)_{\lambda_1, \lambda_2} &= \left\langle \left( C_{2,+}^{\Re}(I_a, I_b)_{\lambda_1, \lambda_2} - C_{2,-}^{\Re}(I_a, I_b)_{\lambda_1, \lambda_2} \right) \right. \\ &\quad \left. + i \cdot \left( C_{2,+}^{\Im}(I_a, I_b)_{\lambda_1, \lambda_2} - C_{2,-}^{\Im}(I_a, I_b)_{\lambda_1, \lambda_2} \right) \right\rangle_{\text{pixels}}. \end{aligned} \quad (7)$$

For these coefficients, we also recover the standard WST definition  $S_2 = \langle |I \star \psi_{\lambda_1}| \star \psi_{\lambda_2} \rangle_{\text{pixel}}$  when  $I_a = I_b = I$ .

An algebraic sum in Eq. (7) allows us to easily identify statistical dependences between processes. When the two images are correlated or anti-correlated, we expect the  $C_1(I_a, I_b)$  to be mostly positive or negative, respectively, leading to  $S_2$  values of the same sign. On the other hand, when the two images are uncorrelated, we expect  $C_1(I_a, I_b)$  to have similar positive and negative patterns, thus leading to  $S_2$  coefficients of much lower values. In addition, using a convolution with a second wavelet allows us to quantify at which scales this correlation appears, and thus to characterise an interaction between two oriented scales.

### 2.3. Principle of the component separation

We present here the component separation method between dust polarisation and the data noise, which we call FoCUS. This method consists of generating a new dust map through a gradient-descent in pixel space under several constraints constructed from CWST statistics. In this part, we call  $u$  the dust map that is modified in the gradient descent. This map is initialised by  $d$ , and we call  $\tilde{s}$  its final value, which is the FoCUS dust map. The scientific result of this paper is this dust map, as well as its statistics. As we discuss in the validation performed in Sect. 3, while the statistics of  $\tilde{s}$  reproduce those of the unknown  $s$  dust map very well, its deterministic structures are not reproduced at the smallest scale.

The three constraints imposed on the  $u$  map are constructed from the full and half-mission maps  $d$ ,  $d_1$ ,  $d_2$ , and  $T$  temperature maps, as well as an ensemble of realisations of full and half-mission noises  $\{\tilde{n}, \tilde{n}_1, \tilde{n}_2\}$ . These constraints, which are not independent, are obtained from averages over the ensemble of noise realisations.

The first constraint is

$$S(d_1, d_2) \simeq \langle S(u + \tilde{n}_1, u + \tilde{n}_2) \rangle_{\tilde{n}}, \quad (8)$$

where the bracket designs an average over the noise realisations, here  $\{\tilde{n}_1, \tilde{n}_2\}$ . This constraint enforces the statistics of  $u$  to the statistics of  $s$  estimated from the two half-mission maps. The second constraint, which yields

$$S(d, u) \simeq \langle S(u + \tilde{n}, u) \rangle_{\tilde{n}}, \quad (9)$$

enforces the cross-statistics between  $u$  and  $d$ . We note the difference between the two constraints: the first constraint is only statistical in nature because it contains no cross-term between the denoised  $u$  map and observational data, while the second constraint includes a cross-term between  $u$  and  $d$ . This allows us to use cross-statistics between half-mission data that have independent noises for the first constraint, while we use the full-mission map, which has the smallest amount of noise, for the second one. This second term also constrains deterministic features. Similarly to a cross-spectrum computation,  $S(d, u)$  characterises the correlation between  $d$  and  $u$  (here a non-linear correlation), and hence the correlation between  $s$  and  $u$  if  $n$  and  $u$  are independent. This term thus constrains the alignment of structures in  $u$  with structures in  $s$ , which we call deterministic features. The  $C_1$  terms for instance impose that local levels of oscillation at a given scale are correlated.

Finally, the last loss constraint,

$$S(T, d) \simeq \langle S(T, u + \tilde{n}) \rangle_{\tilde{n}}, \quad (10)$$

imposes that we keep the same cross-statistics between  $T$  and  $u$  as were estimated from the  $d$  map. For this last constraint, we chose  $T$  to be the 857 GHz SRoll3 map, which was corrected for large-scale systematics present in earlier data releases (Lopez-Radcenco et al. 2021). The choice of 857 GHz instead of the 353 GHz maps avoids correlated noise between the polarisation maps  $d$  and the  $T$  map (as they are computed using data from the same detectors for intensity and polarisation). A drawback is that the 857 GHz  $T$  map includes a significant contribution from the cosmological infrared background at high Galactic latitudes, but this emission is only very weakly polarised (Feng & Holder 2020; Lagache et al. 2020). Furthermore, the non-negligible spatial variations of the spectral energy distribution of the Galactic dust induce some decorrelation between the  $T$  maps at 857 and 353 GHz (Delouis et al. 2021). An advantage of our method is that it is not hindered by these effects because we only impose that the denoised map has the same statistical dependence as the one estimated from the observational data.

We also tried to impose constraints that directly involved the recovered noise map,  $d - u$ . We considered in particular a direct constraint on its statistics, as well as on its independence of  $u$ . Neither of these constraints particularly improved the results we obtained. In particular, the independence of  $u$  and  $d - u$  is already constrained from Eq. (9). It is possible, however, that these constraints would play a non-negligible role in separating other types of astrophysical fields.

In practice, performing a gradient descent from these constraints would be computationally very costly, however, because



it would require computing an average on a large number of noise realisations (300) at each iteration. Five hundred noise realisations were available, but only 300 were used to limit memory and computing usage. We also verified that adding noise apparently does not improve the results. To avoid this, the noise-induced biases of the CWST statistics were separated on a specific term and estimated only after a certain batch of iterations. The loss term related to the first constraint is thus written

$$\text{Loss}_1 = \left\| \frac{S(d_1, d_2) - S(u, u) - B_1}{\sigma_{S(u+\tilde{n}_1, u+\tilde{n}_2)}} \right\|^2, \quad (11)$$

with

$$B_1 = \langle S(u + \tilde{n}_1, u + \tilde{n}_2) - S(u, u) \rangle_{\tilde{n}}, \quad (12)$$

where  $\|\cdot\|^2$  stands for the square Euclidean norm over the whole set of CWST statistics, and  $\sigma$  stands for the estimated standard deviation of  $\langle S(u + \tilde{n}_1, u + \tilde{n}_2) \rangle_{\tilde{n}}$ . As discussed above,  $B_1$  gives an estimate of the noise-induced bias between  $S(d_1, d_2)$  and  $S(s, s)$  that is more accurate the closer  $u$  becomes to  $s$  during the optimisation.

Similarly, the loss terms related to the second constraint yield

$$\text{Loss}_2 = \left\| \frac{S(d, u) - S(u, u) - B_2}{\sigma_{S(u+\tilde{n}, u)}} \right\|^2, \quad (13)$$

with

$$B_2 = \langle S(u + \tilde{n}, u) - S(u, u) \rangle_{\tilde{n}}, \quad (14)$$

In practice, the numerical experiments we performed showed that it was difficult to use a loss term involving a difference between two  $u$  terms, especially at the start of the optimisation where  $u_0 = d$ . To avoid this, we modified this loss term, replacing  $S(u, u)$  by  $S(d_1, d_2) - B_1$ , which proved to be much more efficient. This led to the following loss term:

$$\text{Loss}_2 = \left\| \frac{S(d, u) - S(d_1, d_2) - B_2 + B_1}{\sqrt{\sigma_{S(u+\tilde{n}, u)}^2 + \sigma_{S(u+\tilde{n}_1, u+\tilde{n}_2)}^2}} \right\|^2. \quad (15)$$

And finally the term related to the third constraint yields

$$\text{Loss}_3 = \left\| \frac{S(T, d) - S(T, u) - B_3}{\sigma_{S(T, u+\tilde{n})}} \right\|^2, \quad (16)$$

with

$$B_3 = \langle S(T, u + \tilde{n}) - S(T, u) \rangle_{\tilde{n}}. \quad (17)$$

These losses treat the Galactic signal as a homogeneous process on the sky. However, it is clear that the dust emission has a strong variation in statistical properties with Galactic latitude. To take this into account, the three losses described previously were computed from statistics estimated on different parts of sky, using five different standard Planck masks with  $f_{\text{sky}} \in [1.0, 0.73, 0.63, 0.43, 0.27]$  (Planck Collaboration Int. XXX 2016). Dust statistics are dominated by the brightest emission within the unmasked sky. Therefore, as the amplitude of the dust emission decreases steadily from the Galactic plane to the poles, the masks allow us to progressively characterise dust polarisation from bright to faint regions when  $f_{\text{sky}}$  extends from

1.0 to 0.27. Because in contrast, the noise power is quite homogeneous in Galactic latitude, this also allows us to evaluate the success of the FoCUS algorithm from a high to low signal-to-noise ratio. In practical terms, the masks are taken into account in the averages over sky pixels in Eqs. (5) and (6). Thus, the FoCUS algorithm simultaneously optimises the three loss terms for each of the five masks, that is, 15 constraints in total.

Numerically, the optimisation ran for 500 iterations between each computation of the noise-induced biases. The minimisation did not improve much, and the change in  $\tilde{s}$  was negligible after this number of iterations. This step was repeated 12 times (6000 iterations in total), at which point the modification of the estimated biases was very small. The total iteration time represents 10 h on three nodes (processor=Intel Xeon E5-2680) with 28 CPU cores, or 2 hours on three M100 GPUs. We also note that the optimisation was made on  $d - u$  rather than  $u$ , which was more stable and led to far fewer oscillations between local minima. This probably arises because the  $d - u$  contamination is close to a Gaussian random field as scales where FoCUS works, whose pixels values can be optimised in a much more independent way.

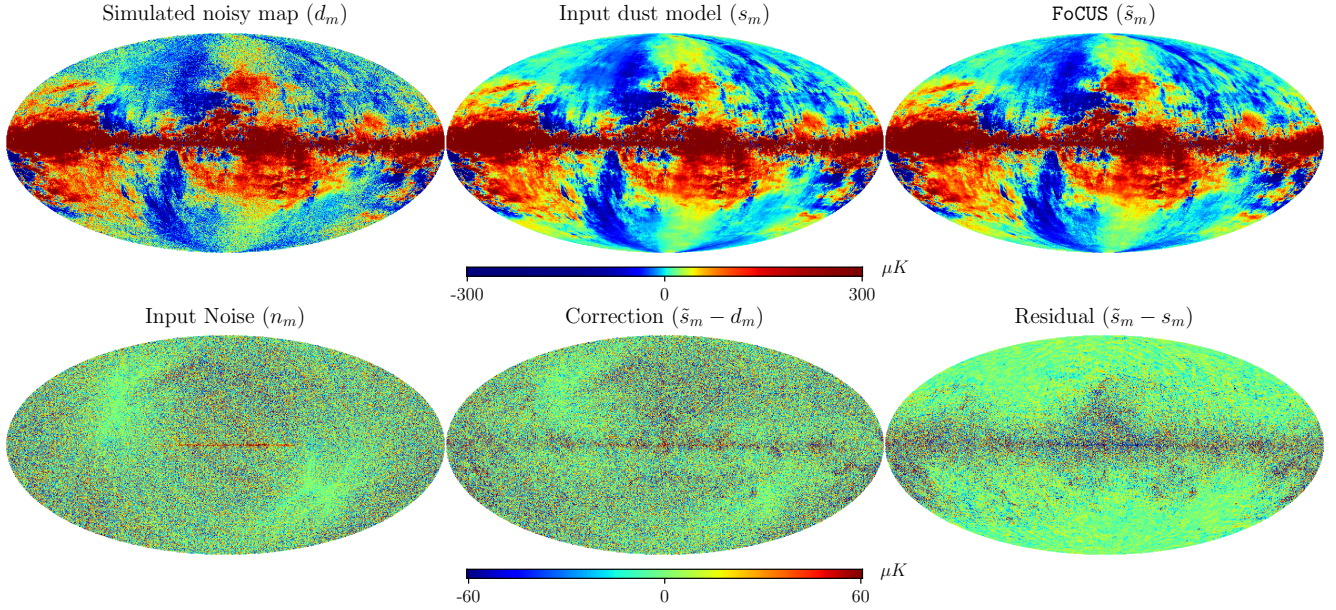
### 3. Validation of the component separation

In this section, we apply the FoCUS component separation method to mock data to assess its performance. We introduce the mock data in Sect. 3.1 and present the results of the FoCUS run in Sect. 3.2. In Sect. 3.3 we analyze the impact of each of the three terms of the loss function on the FoCUS output maps.

#### 3.1. Mock data

To build our mock data set, we combine a model of dust polarisation maps with noise simulations. We used the dust model, hereafter the Vansyngel model, that was introduced in Appendix A of Planck Collaboration III (2020). The Vansyngel model was used by both Planck Collaboration III (2020) and Delouis et al. (2019) to build end-to-end simulations of the Planck polarisation data at 353 GHz. The total intensity maps is the Planck map at 353 GHz obtained by applying the generalized needlet internal linear combination (GNILC) method of Remazeilles et al. (2011) to the 2015 release of Planck HFI maps (PR2). The Stokes  $Q$  and  $U$  maps were built from one realisation of the statistical model of Vansyngel et al. (2017). In the model of Planck Collaboration III (2020), the simulation was replaced by the Planck PR2 353 GHz maps near the Galactic plane, and the largest angular scales  $\ell < 20$  of the simulated maps were also replaced by the Planck data (see Planck Collaboration III 2020 for more details).

Away from the Galactic plane and for multipoles  $\ell > 20$ , the statistical properties of the  $Q$  and  $U$  maps were those of the Vansyngel et al. (2017) model. This model was built from a simplified description of the magnetised interstellar medium where the random component of the magnetic field is represented by Gaussian fields. The  $TE$  correlation and  $E/B$  asymmetry correlation are introduced in the model maps in spherical harmonics with random phases. However, the model  $Q$  and  $U$  maps do have some non-Gaussian characteristics that arise from the total intensity map and the modelling of the line-of-sight integration with a small number of independent emission layers. As explained in Sect. 2.3, the temperature map we used in the third loss term is the SRoll3 map at 857 GHz, even if the  $T$  map used to evaluate the  $TE/TB$  correlations is the map from the Vansyngel model



**Fig. 2.** Stokes  $Q$  maps illustrating the validation on mock data. *Top row, left to right:* one realisation of the noisy mock data, the noise-free Vansyngel model, and the result of the FoCUS algorithm. *Bottom row, left to right:* noise map used in the data simulation, the correction found by the FoCUS method, and the residual map, i.e. the difference between the Vansyngel model and the FoCUS map.

at 353 GHz. This allows us to remain consistent in testing the FoCUS algorithm and in validating the obtained result.

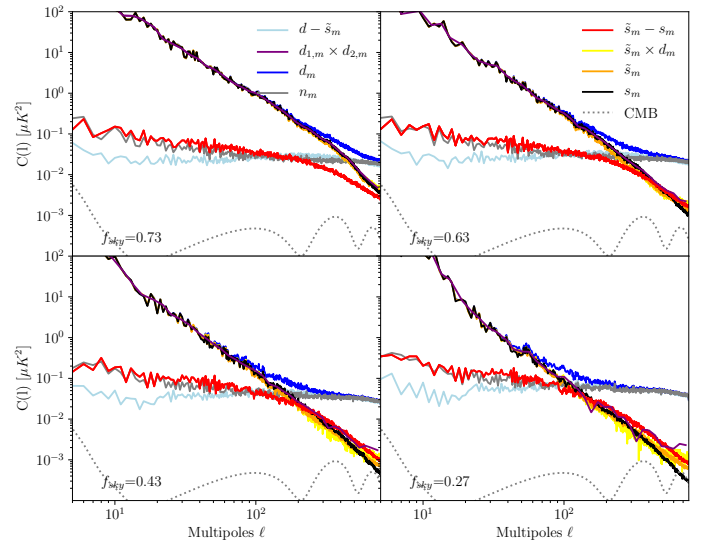
We associated the dust model  $s_m$  (here the  $m$  index is related to modeling and simulation) with noise maps from the end-to-end SRo112 dataset (Delouis et al. 2019). Ten noise realisations were added to the dust model to generate 10 mock  $d_m$  maps, and 300 additional maps, all independent, were used for the FoCUS optimisation. We applied the FoCUS method on the 10  $d_m$  and obtained 10 denoised maps  $\tilde{s}_m$ .

### 3.2. Validation of FoCUS on mock data

#### 3.2.1. FoCUS maps

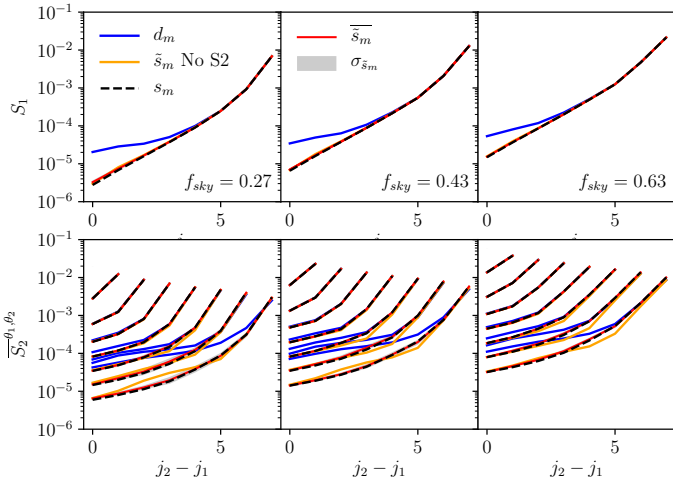
The top row of Fig. 2 presents three  $Q$  maps: from left to right, the noisy mock data ( $d_m$ ), the input dust model ( $s_m$ ), and the FoCUS output  $\tilde{s}_m$ . The bottom row shows the noise map included in  $n_m$ , the noise estimate from FoCUS  $d_m - \tilde{s}_m$ , and the residual  $\tilde{s}_m - s_m$ . The FoCUS map  $\tilde{s}_m$  is strikingly less noisy than  $d_m$ . Fig. 2 also clearly shows that the noise estimate  $d_m - \tilde{s}_m$  corresponds to what we expect: a noisy map with large-scale patterns close to the pattern of the true noise. To the eye, the residual map appears noisier where the dust emission is the brightest. Along the Galactic plane, the residuals are larger, but they represent a small fraction of the total dust signal. Away from the Galactic plane, the residuals do not correspond to leftover noise, but mainly result from small displacements of some structures between  $\tilde{s}_m$  and the true map  $s_m$ .

In Fig. 3 we present  $EE$  power spectra of the maps in Fig. 2 for four masks with  $f_{\text{sky}}$  from 0.27 to 0.73. The corresponding  $BB$  spectra are shown in Fig. A.1. In both figures, it is remarkable that the power spectra of the FoCUS output  $\tilde{s}_m$  in orange closely follow the true dust spectrum in black in the four plots, down to two orders of magnitude below the noise power for  $f_{\text{sky}} = 0.27$ . The success in reproducing deterministic dust structures is characterised by the spectra of the residual map  $\tilde{s}_m - s_m$  in red. The power of the residual map exceeds that of the dust model in black, where the  $EE$  power of dust is one-tenth of the power



**Fig. 3.** Power spectra of one realisation for the FoCUS validation on mock data. The plots show  $EE$  spectra of the dust model ( $s_m$  in black), the noisy mock data ( $d_m$  in blue), and the FoCUS output ( $\tilde{s}_m$  in orange) for four masks with  $f_{\text{sky}}$  from 0.27 to 0.73. The red curve is the power spectrum of the residual map  $\tilde{s}_m - s_m$ . The figure also presents the noise spectrum ( $n_m$ ) and the noise estimate from FoCUS ( $d_m - \tilde{s}_m$ ), as well as cross spectra between half-mission mock maps ( $d_{1,m} \times d_{2,m}$  logarithmic binned) and between  $\tilde{s}_m$  and  $d_m$ . Other mock data realisations show very consistent power spectra.

of noise. For lower signal-to-noise ratios, the amplitude of the residual spectrum is about twice that of the dust power, which indicates that structures in the FoCUS output map  $\tilde{s}_m$  are not spatially coincident with those in the input map  $s_m$ . We show below that this corresponds to a regime in which the recovered structures have the correct statistical properties, but do not reproduce the input data from a deterministic point of view. The presence of this regime has been identified in Regalado-Saint Blancard et al. (2020).



**Fig. 4.** Comparison of CWST statistics. The CWST coefficients of the FoCUS output map  $\tilde{s}_m$  (solid red line) are compared to those of the mock data  $d$  (solid blue line) and the noise-free dust model  $s_m$  (dashed black line). The thicker grey line represents the statistical variance of the FoCUS output estimated from ten simulations. The *top row* shows the coefficients  $S_1$  plotted vs scale  $j_1$ , and the *bottom row* shows the  $S_2$  coefficients averaged over  $\theta_1$  and  $\theta_2$  plotted vs the ratio of the two scales  $j_2 - j_1$ . The three columns correspond to different Galactic masks.

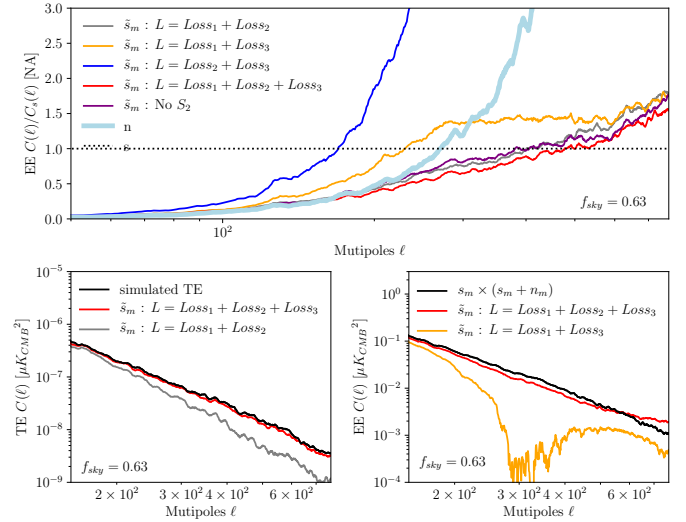
### 3.2.2. CWST statistics

One main objective of FoCUS is to derive a statistical model of dust polarisation that is unbiased by data noise from an observation. The mock data allow us to assess the success of the algorithm in this regard.

In Fig. 4 we compare the CWST coefficients of the FoCUS output map  $\tilde{s}_m$  to those of the mock data  $d_m$  and the noise-free dust model  $s_m$ . As discussed in Sect. 2.2, the CWST coefficients of a single map are the standard WST coefficients studied for instance in Allys et al. (2019). The *top row* shows the coefficients  $S_1$  averaged over  $\theta_1$  plotted versus scale  $j_1$ , and the *bottom row* shows the mean  $S_2$  coefficients averaged over  $\theta_1$  and  $\theta_2$  plotted versus the ratio of the two scales  $j_2 - j_1$ , for  $j_1 = 0$  to 6 from the bottom right to the top left.

The statistics of  $d_m$  clearly depart from those of  $s_m$ . As expected, the difference is most noticeable at small scales, as well as for the area of the sky with the lowest signal-to-noise ratio at high Galactic latitudes. For  $f_{\text{sky}} = 0.27$  and  $j_1 = 0$ , the mean  $S_2$  coefficients for  $d$  depart from those of  $s$  for all  $j_2$  but the largest scale. Extracting non-Gaussian statistics of  $s_m$  from the noisy  $d_m$  data down to scale  $j_1 = 0$  is therefore a notable challenge at high Galactic latitudes, where the power ratio between the dust signal and noise ratio is down to 1% (see Fig. 3). In this challenging context, the excellent match between coefficients for  $\tilde{s}_m$  and  $s_m$  for all masks demonstrates the remarkable success of the FoCUS algorithm in synthesising maps with the same statistics as the noise-free dust emission.

For the 353 GHz Planck data, we expect some non-negligible bias to appear on the smallest scales probed with  $N_{\text{side}} > 256$ , however, entering a regime in which even the statistical properties of the denoised map begin to differ from those of the true map. This third regime, which has also been observed in Regalado-Saint Blancard et al. (2020), is indeed expected in the limit in which the noise has a much higher level than the signal. It might be possible to use an extrapolated model of the CWST statistics of the dust to describe scales included in this regime, however.



**Fig. 5.** Power spectra of the FoCUS maps for different combinations of the three loss terms  $\text{Loss}_1$ ,  $\text{Loss}_2$ , and  $\text{Loss}_3$ . *Top plot*: ratio of the  $EE$  spectra of the residual map,  $s - \tilde{s}_m$ , and the input dust model  $s_m$ . *Bottom left plot*: compares the  $TE$  spectra obtained with or without  $\text{Loss}_3$ , in red and grey, respectively. *Bottom right plot*: compares the cross spectra between FoCUS map  $\tilde{s}_m$  and the mock data  $d$  with or without the  $\text{Loss}_2$  term in red and yellow, respectively. All of the spectra are binned in  $\ell$  bins with a width  $\Delta\ell = 10$ .

### 3.3. Impact of each loss term

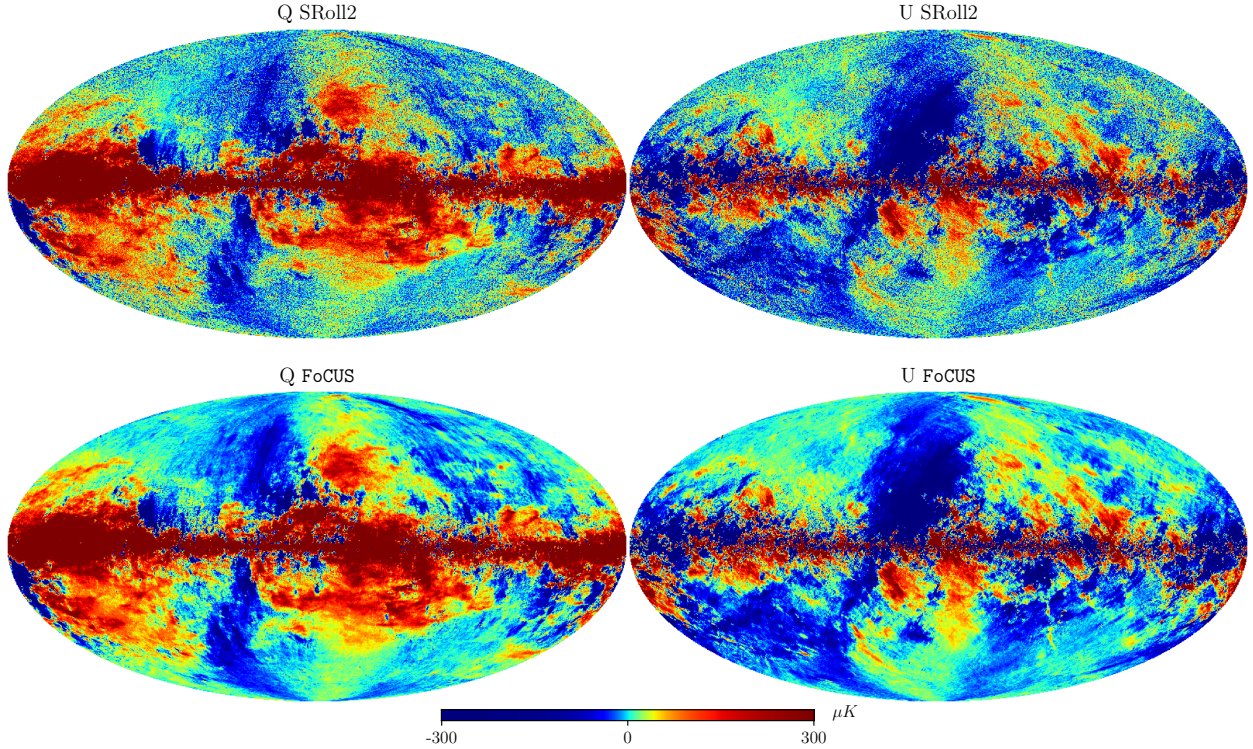
Our loss functions comprise three terms:  $\text{Loss}_1$ ,  $\text{Loss}_2$  and  $\text{Loss}_3$ . They are defined in Eqs. (11), (13), and (16). The mock data allow us to assess the impact of pairwise combinations excluding one of the loss terms for all sky masks on the FoCUS output maps.

Figure 5 includes three plots in which the input dust model  $s_m$  is drawn in black. All plots are for  $f_{\text{sky}} = 0.63$ . The top plot presents the  $EE$  spectra of the residual maps  $\tilde{s}_m - s_m$  divided by that of  $s_m$ . This ratio quantifies the ability of the FoCUS algorithm to reconstruct structures consistently with the noise-free dust model. The bottom left plot compares the  $TE$  spectrum obtained with or without the  $\text{Loss}_3$  term, in red and grey. The bottom right plot compares the cross spectra between the FoCUS map  $\tilde{s}_m$  and the mock data  $d_m$  with or without the  $\text{Loss}_2$  term in red and yellow, respectively.

In the three plots, the FoCUS output maps  $\tilde{s}_m$  obtained with the complete loss function, drawn in red, is the best result, which confirms our choice of combining the three terms of the loss function. The top plot shows that both  $\text{Loss}_1$  and  $\text{Loss}_2$  are essential for minimising the power of the residual maps. We interpret the effect of  $\text{Loss}_2$  as follows. At multipoles between  $10^2$  and  $3 \times 10^2$ , this loss allows recovering the deterministic structures, which can be characterised by the cross-statistics with the noisy-data. At higher multipoles, these cross-statistics progressively become noise dominated, and the lever-arm of this loss term to recover deterministic structures decreases.  $\text{Loss}_2$  also ensures a faster convergence of the minimisation. Furthermore, the bottom right plot shows that  $\text{Loss}_2$  is critical for matching the cross spectra  $\tilde{s}_m \times d_m$ .  $\text{Loss}_3$  has only a weak impact on the residual maps, but the bottom left plot shows that it is essential for matching the  $TE$  correlation of the dust model, which is expected because it constrains the cross statistics between dust polarisation and total intensity.

The top plot in Fig. 5 also shows the residual power for  $\tilde{s}_m$  maps obtained by discarding the  $S_2$  coefficients in the loss function in purple. A comparison the purple and red curves





**Fig. 6.** Stokes  $Q$  and  $U$  maps at 353 GHz for  $N_{\text{side}} = 256$ . *Top panels:* input Planck SRoll2 maps, and the *bottom panels* show the corresponding FoCUS maps.

shows that these coefficients only improve the residual power by a few dozen percent. However, the yellow curve in Fig. 4 demonstrates that the  $S_2$  coefficients in the loss terms are important for fully recovering the non-Gaussian properties of the dust map, as expected.

#### 4. Denoised Planck dust polarisation maps

In this section, we use the CWST and FoCUS to separate dust polarisation and data noise in Planck data. The data we used are introduced in Sect. 4.1, and the denoised Planck polarisation maps are presented in Sect. 4.2.

##### 4.1. Input data

We used the Planck Stokes  $Q$  and  $U$  maps at 353 GHz from the SRoll2 processing, which corrects for the instrumental systematics present in the Planck Legacy polarisation maps of the High Frequency Instrument<sup>1</sup> (Delouis et al. 2019). To obtain dust polarisation maps, we subtracted the CMB polarisation using  $Q$  and  $U$  maps from the SMICA component separation method (Planck Collaboration IV 2020). Over the multipole range we considered, dust polarisation dominates the CMB signal for all sky masks at 353 GHz, as shown in Fig. 3. Uncertainties on the CMB correction are thus not an issue. As explained in Sect. 2.3, the temperature map we used in the third loss is the SRoll3 map at 857 GHz.

##### 4.2. FoCUS maps

Figure 6 presents the SRoll2 Stokes  $Q$  and  $U$  maps at 353 GHz (top images) and the result of the FoCUS algorithm (bottom

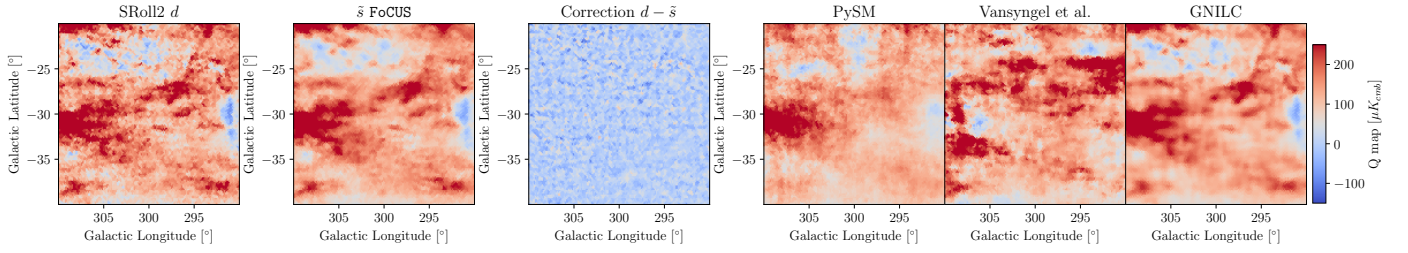
<sup>1</sup> The SRoll2 maps are available at [http://sroll20.ias.u-psud.fr/sroll22\\_files.html](http://sroll20.ias.u-psud.fr/sroll22_files.html)

images). The comparison by eye shows that noise has been subtracted without smoothing the map. This is further illustrated in Fig. 7, which zooms on one sky area to allow a detailed comparison with other dust polarisation models. Even without over-smoothing, the dust map has much less power at small scales than the noise. The denoised map is therefore expected to have a smoother texture.

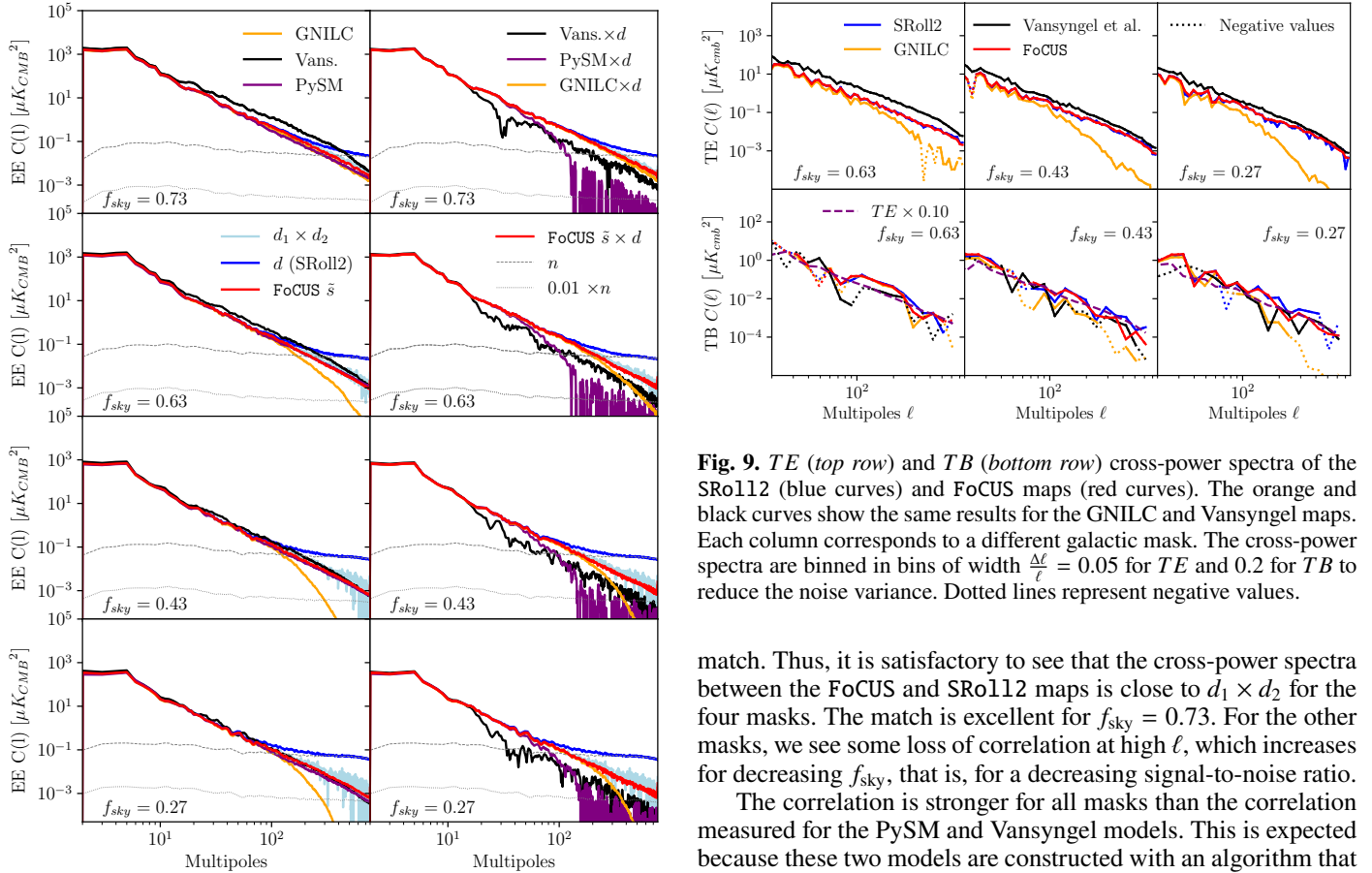
The PySM d1 model map (Thorne et al. 2017; Zonca et al. 2021) includes small-scale structure that was derived from a random Gaussian field, which appears unrealistic. The Vansyngel maps are constructed from the dust total intensity map. Their textures appear to be closer to what is statistically expected, but the small scales seem to lack elongated structures. The GNILC map shows a lack of small scales. Finally, the FoCUS map shows non-Gaussian elongated structures at all scales, which is a general characteristic of the diffuse interstellar medium. Our algorithm is able to capture this by using advanced statistical descriptors and constraints.

Power spectra of the Planck data are compared with those of the FoCUS, GNILC, PySM, and Vansyngel maps for four Galactic masks defined by their corresponding  $f_{\text{sky}}$  in Figs. 8 and A.2 for  $EE$  and  $BB$ , respectively. In each figure, the plots in the left column present power spectra and the plots in the right column show cross-spectra with the SRoll2 input maps. Each plot includes the cross spectrum between the two SRoll2 half-mission maps ( $d_1 \times d_2$ ) drawn in purple as the reference to match because it is a noise-unbiased estimate from the Planck data of the spectrum of Galactic dust emission.

In the left column of the two figures, the SRoll2 spectra show the data noise bias over an increasing range of  $\ell$  as  $f_{\text{sky}}$  decreases. It is remarkable that the power spectra of the FoCUS map are consistent with the reference for the four masks, up to multipoles where the signal power is two orders of magnitude lower than that of the noise. The difference with the GNILC method, which reduces the sky noise at the expense of



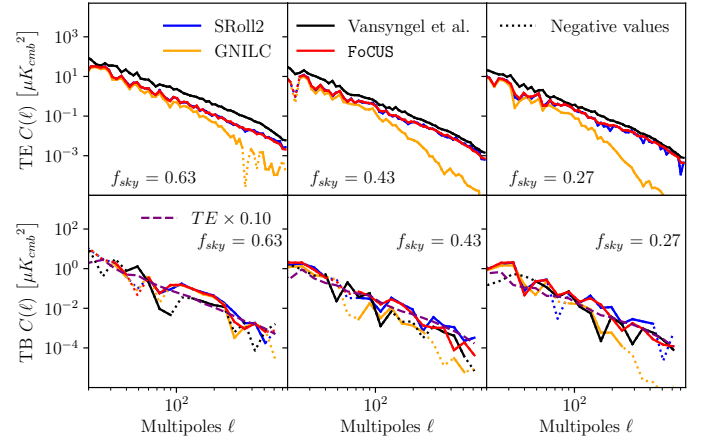
**Fig. 7.** Zoom on a sky region of the Q map. *From left to right panels:* SRoll 2.0 data, the FoCUS map, the correction computed by FoCUS to be applied to the SRoll2.0 map, the PySM d1 model, the Vansyngel et al. (2017) map, and the GNILC map.



**Fig. 8.** *EE* power spectra (left column) and cross-power spectra (right column) for Galactic masks with  $f_{\text{sky}}=0.27$  (bottom) to 0.73 (top). In each plot, the cross spectrum of the two Planck half-mission maps is drawn in purple. The blue, orange, and red curves represent the power spectra of the SRoll2, GNILC, and FoCUS maps, respectively. All the spectra are binned over ten multipoles and normalized by dividing the power with the  $f_{\text{sky}}$  value to keep the scales consistent between the plots.

small-scale smoothing, stands out in Fig. 8. The plots also show that the power spectra of the Vansyngel maps deviate somewhat from the Planck data, which is a known shortcoming of their model.

The plots in the right columns of Figs. 8 and A.2 present cross-power spectra between models and the SRoll12 maps. Our purpose is to quantify the correlation on the sky between the data and model maps, but we note that our validation of the FoCUS method has shown (see Sect. 2.2) that some correlation may arise from data noise. On these plots, like for those in the left columns, the  $d_1 \times d_2$  cross-power spectra are the references to



**Fig. 9.** *TE* (top row) and *TB* (bottom row) cross-power spectra of the SRoll12 (blue curves) and FoCUS maps (red curves). The orange and black curves show the same results for the GNILC and Vansyngel maps. Each column corresponds to a different galactic mask. The cross-power spectra are binned in bins of width  $\frac{\Delta\ell}{\ell} = 0.05$  for *TE* and 0.2 for *TB* to reduce the noise variance. Dotted lines represent negative values.

match. Thus, it is satisfactory to see that the cross-power spectra between the FoCUS and SRoll12 maps is close to  $d_1 \times d_2$  for the four masks. The match is excellent for  $f_{\text{sky}} = 0.73$ . For the other masks, we see some loss of correlation at high  $\ell$ , which increases for decreasing  $f_{\text{sky}}$ , that is, for a decreasing signal-to-noise ratio.

The correlation is stronger for all masks than the correlation measured for the PySM and Vansyngel models. This is expected because these two models are constructed with an algorithm that is not designed to preserve correlation with the input data.

#### 4.3. *TE* and *TB* correlation

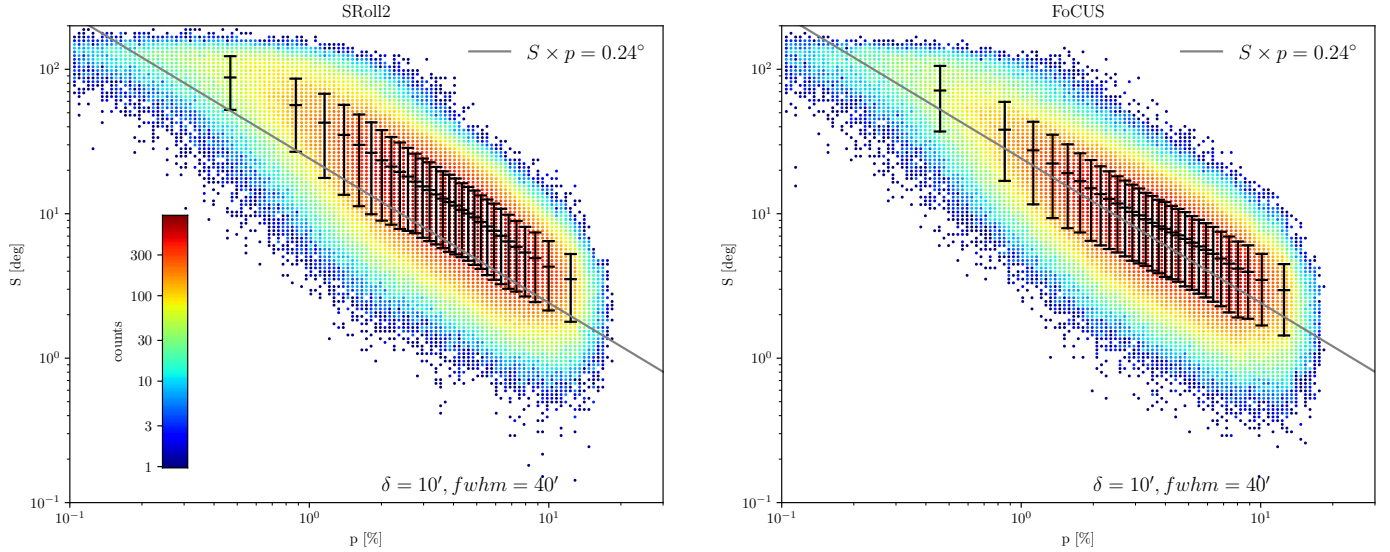
The *TE* and *TB* correlation are main statistical characteristics of dust polarised emission, determined by the analysis of Planck data (Planck Collaboration XI 2020). This important property is difficult to match without learning the model statistics from data.

In Fig. 9 the *TE* and *TB* cross-power spectra for the SRoll12 and FoCUS maps are compared. The plots show that the two sets of *TE* and *TB* cross-power spectra match and that the FoCUS algorithm reduces the variance at the highest  $\ell$ . This is a clear success of the FoCUS algorithm. Figure 9 also shows that the previous noise removal (see GNILC or Vansyngel et al. 2017 in orange or black) do not reproduce these correlations.

## 5. Projected applications and perspectives

The Planck 353 GHz polarisation maps have been extensively used for the astrophysics of dust polarisation and the modelling





**Fig. 10.** Distribution of the polarisation angle dispersion  $S$  and the polarisation fraction  $p$  values for the SRoll2 (*left panel*) and FoCUS (*right panel*) maps. The data points with error bars represent the mean value and  $1\sigma$  dispersion of  $S$  within bins of  $p$ . The data binning and representation follows that of Fig. 10 of Planck Collaboration XII (2020). The grey line shows the relation  $S \times p = 0.24^\circ$ , which is expected based on the analytical model of Planck Collaboration XII (2020) for the resolution  $FWHM = 40'$  we used.

of the Galactic dust foreground to CMB. For both topics, the FoCUS map represents a significant stepping-stone opening new prospects. We illustrate and discuss these perspectives from the points of view of astrophysics and foregrounds.

### 5.1. Astrophysics of dust polarisation

For astrophysics, the signal-to-noise ratio of the dust polarisation maps statistically conditions the range of angular scales accessible for study. The polarised emission of the diffuse ISM at high Galactic latitudes is too faint to be analysed at the full  $5'$  angular resolution of Planck (see dust power spectra in Planck Collaboration XI 2020). Most of the analysis of the dust polarised emission in Planck Collaboration XII (2020) was performed after smoothing the Planck maps to  $80'$  and even  $160'$  resolution. The FoCUS maps thus allow us to extend the range of earlier studies. As our data denoising is statistical in nature at scales at which the signal-to-noise ratio is low, the FoCUS maps are most relevant for statistical studies, in particular those that aim to statistically characterise the turbulent component of the Galactic magnetic field.

To illustrate this perspective, we use the FoCUS maps to compare the polarisation angle dispersion,  $S$ , and the polarisation fraction,  $p$ , as was done by Planck Collaboration XII (2020) with the Planck Legacy maps.  $S$ , introduced by Hildebrand et al. (2009) and Planck Collaboration Int. XIX (2015), quantifies the local non-uniformity of the polarisation angle patterns on the sky by means of the local variance of the polarisation angle map on a scale defined by a lag  $\delta$ . Regions in which the polarisation angle tends to be uniform exhibit low values of  $S$ , while regions where the polarisation patterns change on the lag-scale exhibit higher values. The polarisation fraction,  $p$ , depends on both the orientation of the mean Galactic magnetic field in the solar neighborhood and, statistically, on depolarisation resulting from changes in the magnetic field orientation along the line of sight (Planck Collaboration Int. XLIV 2016). The trend  $S \propto 1/p$  observed in the Planck data can be reproduced with the analytical model presented in Appendix A of Planck Collaboration XII (2020). The product  $S \times p$  scales with the degree of randomness

of the Galactic magnetic field: the ratio of the dispersion of the turbulent component and the strength of the mean field.

Figure 10 presents plots of  $S$  versus  $p$  for both the SRoll2 and FoCUS maps with the same presentation as in the corresponding figure (also Fig. 10) in Planck Collaboration XII (2020). Figure A.3 presents the corresponding plots for the Vansyngel model. Our maps have a resolution of  $40'$  (full width at half maximum, FWHM) to be compared with  $160'$  for the plot in Planck Collaboration XII (2020). As in Planck Collaboration Int. XIX (2015), we used a lag equal to one-fourth of the FWHM, that is,  $\delta = 10'$ . The model relation for our resolution and lag,  $S \times p = 0.24^\circ$ , is the line drawn in the two plots. Figure A.3 shows that the noise bias on the  $S$  and  $p$  relation is considerably reduced by the FoCUS algorithm. Given the excellent debiasing obtained on this validation dataset, the FoCUS plot in Fig. 10 suggests that the analytical model of (Planck Collaboration Int. XIX 2015) only provides an approximate description of the data.

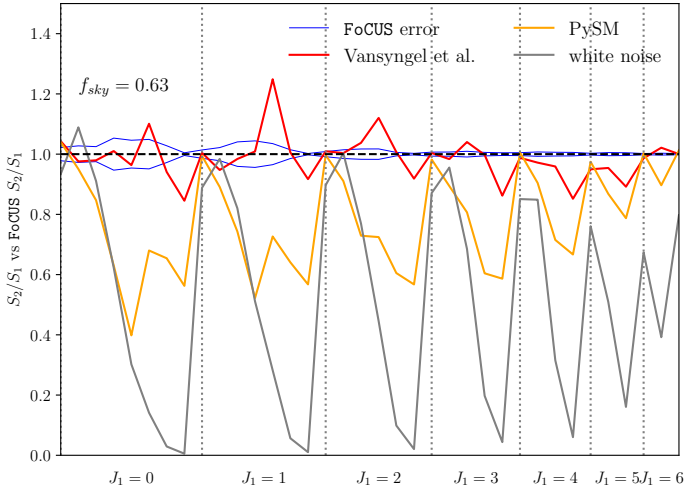
### 5.2. Galactic foreground modelling

The modelling of the dust foreground to CMB polarisation is the primary motivation of our work. The approach we followed to model dust polarisation is novel in some crucial aspects, which we summarize here.

We learned our statistical model from the Planck data using a set of summary statistics designed to perform an in-depth characterisation of non-Gaussian structures (Bruna & Mallat 2013). The FoCUS maps include non-Gaussian features that are missed by models making use of Gaussian random fields to describe foregrounds on scales at which the Planck template maps are noise dominated. This important difference is illustrated in Fig. 11, where ratios of the CWST coefficients  $S_2/S_1$ , used as a measure of non-Gaussianity, are compared between the FoCUS PySM and Vansyngel  $E$  maps.

We have applied a statistical component separation to all-sky Planck maps that allowed us to extend our statistical model of dust polarisation down to scales where the dust power is two orders of magnitude smaller than the data noise. For the sky at high Galactic latitudes that are best suited for deep CMB





**Fig. 11.** Comparison of the second-order  $S_2$  CWST coefficients used as a measure of non-Gaussianity. The  $S_2$  coefficients of  $E$  maps normalized by  $S_1$  are plotted vs scales  $J_1$  and  $J_2$ . The data are ordered by  $J_1$  values and for each  $J_1$  for increasing  $J_2$ . The plot compares  $S_2/S_1$  ratios for a Gaussian map in purple, the PySM model in yellow, and the model of Vansyngel et al. (2017) and the FoCUS map with a dashed black line. All data values are normalised by those of the FoCUS map. The two blue curves shows the uncertainty ( $\pm 1\sigma$ ) on the  $S_2/S_1$  ratio estimated from simulations.

observations, we succeed in modelling the dust polarisation up to  $N_{\text{side}} = 256$ . The resolution could be increased by modelling and extrapolating the scale dependence of the CWST break coefficients.

The CWST statistics allow us to combine different maps. We used this possibility to learn the correlation between dust total intensity and polarisation from the Planck data. Therefore, our FoCUS map reproduces the  $E/B$  asymmetry, the  $TE$ , and also  $TB$  correlations, which is a unique achievement among current dust models, as illustrated in Fig. 9. This figure also shows that denoising methods such as GNILC, which seek to minimise the difference between the denoised map and the true signal, remove power on scales that are dominated by the noise.

To proceed in the modelling of polarised Galactic foregrounds, one main objective will be to extend this analysis to multi-frequency models that take the spatial variations of the spectral energy distribution of dust polarisation into account (Ritacco et al. 2022). For this purpose, it is necessary to properly construct joint multi-channel scattering transform statistics, as well as to extend the component separation algorithm to this framework.

## 6. Conclusion

We have applied the scattering transform to Planck data in order to derive a non-Gaussian model of dust polarisation and produce denoised all-sky dust Stokes  $Q$  and  $U$  maps at 353 GHz. First, we introduced the CWST statistics that we used to characterise the non-Gaussian structure of dust polarisation. They extend the computation of scattering coefficients to the HEALPix pixelisation on the sphere and include cross-statistics that allow us to combine images. Second, we devised the FoCUS algorithm, which uses the CWST statistics to separate dust polarisation from data noise. FoCUS was validated on simulations of the Planck data before it was applied to the SRoll12 Planck maps at 353 GHz. The main results of our work are described below.

The CWST statistics and the FoCUS algorithm allow us to characterise dust polarisation down to angular scales at which

the  $EE$  dust power is two orders of magnitude smaller than that of the data noise. The FoCUS Stokes maps reproduce Planck dust polarisation power spectra estimated from cross-spectra of half-mission maps over these scales.

Our validation on mock data allowed us to compare the FoCUS output map  $\tilde{s}$  with the noise-free input map  $s$ . The spectra of the residual map  $\tilde{s} - s$  become larger than the spectra of  $s$  at scales at which the  $EE$  dust power is lower than one-tenth of the noise power. On these scales, structures in the FoCUS output maps  $\tilde{s}$  have comparable non-Gaussian statistics, estimated in terms of CWST, but are not spatially coincident with those in  $s$ .

The FoCUS Stokes maps at 353 GHz<sup>2</sup>, with a set of residual maps from our mock data analysis quantifying uncertainties, are made available to the community. The FoCUS maps open new prospects for astrophysics and the modelling of the Galactic dust foreground to CMB polarisation.

For astrophysics, the signal-to-noise ratio of the dust polarisation maps limits the range of angular scales that are accessible for study. Because our denoising of the data is statistical in nature at scales where the signal-to-noise ratio is low, the gain is in statistical studies. We illustrated this type of applications by repeating the Planck studies of the anti-correlation between the dispersion of polarisation angles and the polarisation fraction.

The FoCUS Stokes maps improve on current models of dust polarisation in two main aspects. (1) The FoCUS maps include non-Gaussian characteristics of dust polarisation, which are missed by models making use of Gaussian random fields to describe foregrounds on scales at which the Planck maps are noise dominated. (2) The CWST cross-statistics allows us to learn the correlation between dust total intensity and polarisation from the Planck data. Therefore, our FoCUS map reproduces the  $E/B$  asymmetry and the  $TE$  and  $TB$  correlations, which is a unique achievement among current dust models used to assess CMB component separation methods.

A clear path to improving our results would be to use the most recent scattering transform statistics on the sphere. Following recent works, several improvements of the scattering statistics might be made in the near future. On the one hand, more refined wavelet transforms on the sphere could be used, as discussed in Sect. 2.2. On the other hand, other successful sets of scattering statistics, which give better syntheses on regular 2D grids, could be used. For instance, the wavelet phase harmonics (Allys et al. 2020; Jeffrey et al. 2022) and their recent multi-channel extensions (Régald-Saint Blancard et al. 2022), or the more recent representations built from wavelet scattering covariances (Morel et al. 2022; Cheng et al., in prep.). However, the main challenge is to make these improvements feasible given the computational and memory costs of the FoCUS algorithm.

The scattering coefficients derived from the Planck data could also be used to generate a set of realistic synthetic non-Gaussian foreground maps. Several ways to proceed might be thought of, which would depend on the scientific objective of such a generation. This is an open issue for future works.

On a more general aspect, the CWST statistics and the FoCUS algorithm might be applied to many processes defined on the sphere, including in other areas than astrophysics, to leverage and analyse different types of correlated data. Its two main advantages are its ability to efficiently combine different datasets and the statistical constraints.

<sup>2</sup> The FoCUS maps are available at [http://sroll120.ias.u-psud.fr/sroll140\\_353\\_data.html](http://sroll120.ias.u-psud.fr/sroll140_353_data.html)

*Acknowledgements.* This work is part of the R & T Deepsee project supported by CNES. The authors acknowledge the heritage of the Planck-HFI consortium regarding data, software, knowledge. This work has been supported by the Programme National de Télédetection Spatiale (PNTS, <http://programmes.insu.cnrs.fr/pnts/>), grant n° PNTS-2020-08 and by the CNRS to the 80 Prime project AstrOcean. F.B. acknowledges support from the Agence Nationale de la Recherche (project BxB: ANR-17-CE31-0022). The authors acknowledge valuable discussions with B. Regaldo on cross-scattering coefficients. The authors also thank C. Auclair, S. Cheng, M. Eickenberg, F. Levrier, J. McEwen, S. Mallat, B. Ménard, R. Morel, and P. Richard for useful discussions.

## References

- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv e-prints [arXiv:1610.02743]
- Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, *J. Cosmol. Astropart. Phys.*, 2019, 056
- Allys, E., Levrier, F., Zhang, S., et al. 2019, *A&A*, 629, A115
- Allys, E., Marchand, T., Cardoso, J. F., et al. 2020, *Phys. Rev. D*, 102, 103506
- Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, *MNRAS*, 488, 4440
- Aylor, K., Haq, M., Knox, L., Hezaveh, Y., & Perreault-Levasseur, L. 2020, *MNRAS*, 500, 3889
- BICEP2/Keck Array and Planck Collaborations. 2015, *Phys. Rev. Lett.*, 114, 101301
- Bruna, J., & Mallat, S. 2013, *IEEE Trans. Pattern Anal. Mach. Intell.*, 35, 1872
- Bruna, J., & Mallat, S. 2019, *Math. Stat. Learn.*, 1, 257
- Cheng, S., & Ménard, B. 2021a, arXiv e-prints [arXiv:2112.01288]
- Cheng, S., & Ménard, B. 2021b, *MNRAS*, 507, 1012
- Cheng, S., Ting, Y.-S., Ménard, B., & Bruna, J. 2020, *MNRAS*, 499, 5902
- Delabrouille, J., Betoule, M., Melin, J. B., et al. 2013, *A&A*, 553, A96
- Delouis, J.-M., Pagano, L., Mottet, S., Puget, J.-L., & Vibert, L. 2019, *A&A*, 629, A38
- Delouis, J. M., Puget, J. L., & Vibert, L. 2021, *A&A*, 650, A82
- Durrer, R. 2015, *Class. Quantum Gravity*, 32, 124007
- Eickenberg, M., Allys, E., Dizgah, A. M., et al. 2022, arXiv e-prints [arXiv:2204.07646]
- Feng, C., & Holder, G. 2020, *ApJ*, 897, 140
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Greig, B., Ting, Y.-S., & Kaurov, A. A. 2022, *MNRAS*, 513, 1719
- Guth, A. H. 1981, *Phys. Rev. D*, 23, 347
- Hervías-Caimapo, C., & Huffenberger, K. M. 2022, *ApJ*, 928, 65
- Hildebrand, R. H., Kirby, L., Dotson, J. L., Houde, M., & Vaillancourt, J. E. 2009, *ApJ*, 696, 567
- Jeffrey, N., Boulanger, F., Wandelt, B. D., et al. 2022, *MNRAS*, 510, L1
- Kim, C.-G., Choi, S. K., & Flauger, R. 2019, *ApJ*, 880, 106
- Krachmalnicoff, N., & Puglisi, G. 2021, *ApJ*, 911, 42
- Kritsuk, A. G., Flauger, R., & Ustyugov, S. D. 2018, *Phys. Rev. Lett.*, 121, 021104
- Lagache, G., Béthermin, M., Montier, L., Serra, P., & Tucci, M. 2020, *A&A*, 642, A232
- Leistedt, B., McEwen, J. D., Vanderghaynst, P., & Wiaux, Y. 2013, *A&A*, 558, A128
- Linde, A. 1982, *Phys. Lett. B*, 116, 335
- LiteBIRD Collaboration (Allys, E., et al.) 2022, PTEP, accepted, [arXiv:2202.02773]
- Lopez-Radcenco, M., Delouis, J. M., & Vibert, L. 2021, *A&A*, 651, A65
- Mallat, S. 2012, *Commun. Pure Appl. Math.*, 65, 1331
- Mallat, S., Zhang, S., & Rochette, G. 2018, arXiv e-prints [arXiv:1810.12136]
- Martínez-Solaèche, G., Karakci, A., & Delabrouille, J. 2018, *MNRAS*, 476, 1310
- McEwen, J. D., Leistedt, B., Büttner, M., Peiris, H. V., & Wiaux, Y. 2015, arXiv e-prints [arXiv:1509.06749]
- McEwen, J. D., Durastanti, C., & Wiaux, Y. 2018, *Appl. Comput. Harmonic Anal.*, 44, 59
- McEwen, J. D., Wallis, C. G., & Mavor-Parker, A. N. 2021, ICLR, accepted [arXiv:2102.02828]
- Morel, R., Rochette, G., Leonarduzzi, R., Bouchaud, J.-P., & Mallat, S. 2022, arXiv e-prints [arXiv:2204.10177]
- Pelgrims, V., Ntormousi, E., & Tassis, K. 2022, *A&A*, 658, A134
- Petroff, M. A., Addison, G. E., Bennett, C. L., & Weiland, J. L. 2020, *ApJ*, 903, 104
- Planck Collaboration Int. XIX. 2015, *A&A*, 576, A104
- Planck Collaboration Int. XLIV. 2016, *A&A*, 596, A105
- Planck Collaboration Int. XLVI. 2016, *A&A*, 596, A107
- Planck Collaboration Int. XXX. 2016, *A&A*, 586, A133
- Planck Collaboration III. 2020, *A&A*, 641, A3
- Planck Collaboration IV. 2020, *A&A*, 641, A4
- Planck Collaboration V. 2020, *A&A*, 641, A5
- Planck Collaboration XI. 2020, *A&A*, 641, A11
- Planck Collaboration XII. 2020, *A&A*, 641, A12
- Regaldo-Saint Blancard, B., Levrier, F., Allys, E., Bellomi, E., & Boulanger, F. 2020, *A&A*, 642, A217
- Regaldo-Saint Blancard, B., Allys, E., Boulanger, F., Levrier, F., & Jeffrey, N. 2021, *A&A*, 649, L18
- Régaldo-Saint Blancard, B., Allys, E., Auclair, C., et al. 2022, *ApJ*, submitted [arXiv:2208.03538]
- Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, *MNRAS*, 418, 467
- Ritacco, A., Boulanger, F., Guillet, V., et al. 2022, *A&A*, submitted [arXiv:2206.07671]
- Saydjari, A. K., Portillo, S. K. N., Slepian, Z., et al. 2021, *ApJ*, 910, 122
- Starck, J.-L., Candès, E. J., & Donoho, D. L. 2002, *IEEE Trans. Image Process.*, 11, 670
- Thorne, B., Dunkley, J., Alonso, D., & Naess, S. 2017, *MNRAS*, 469, 2821
- Thorne, B., Knox, L., & Prabhu, K. 2021, *MNRAS*, 504, 2603
- Vacher, L., Aumont, J., Montier, L., et al. 2022, *A&A*, 660, A111
- Valogiannis, G., & Dvorkin, C. 2022a, *Phys. Rev. D*, 105, 103534
- Valogiannis, G., & Dvorkin, C. 2022b, *Phys. Rev. D*, 106, 103509
- Vansyngel, F., Boulanger, F., Ghosh, T., et al. 2017, *A&A*, 603, A62
- Wiener, N. 1949, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (Cambridge: MIT Press), 2
- Zaroubi, S., Hoffman, Y., Fisher, K., & Lahav, O. 1995, *ApJ*, 449, 446
- Zonca, A., Thorne, B., Krachmalnicoff, N., & Borrill, J. 2021, *J. Open Source Softw.*, 6, 3783

## Appendix A: Additional figures

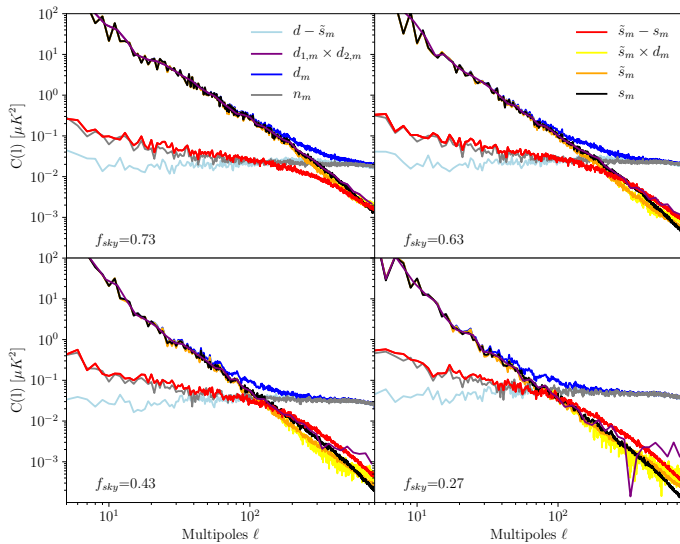
### A.1. Power spectra for $B$ -modes

Figure A.1 complements Fig. 3 by presenting  $BB$  power spectra of the FoCUS validation for one realisation of the mock data. The signal-to-noise ratio is lower for  $BB$  than  $EE$  as the  $E$  - to-  $B$  power ratio for dust emission is about 2. Furthermore, the  $TB/TE$  power ratio is about one-tenth, which decreases the impact of the loss term ( $Loss_3$ ) based on this correlation. For the mock data,  $Loss_3$  is fully ineffective because the Vansyngel model does not include the  $TB$  correlation. These effects combine to make FoCUS denoising more challenging. However, Fig. 3 shows a good consistency of the  $BB$  power spectra of the FoCUS and the input model maps for most multipoles. At very high Galactic latitudes (bottom right panel), the noisier part of the sky, the  $BB$  power spectrum of the FoCUS maps, shows a small bias compared to that of the input maps, which reflects the limitation of our method when the signal-to-noise ratio is very low ( $< 1\%$ ).

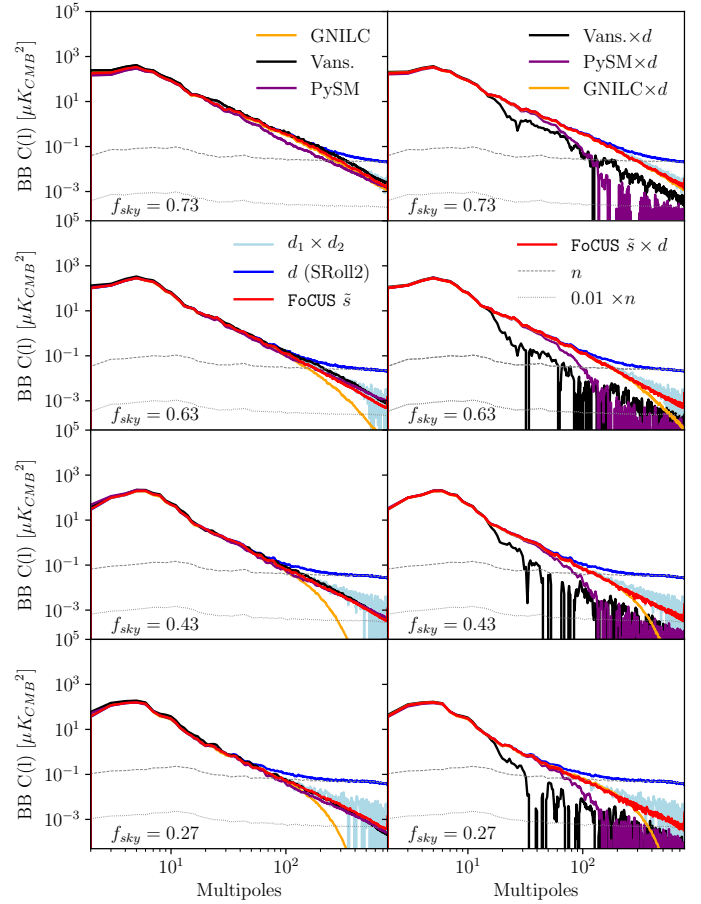
Figure A.2 shows the same set of power spectra as in Fig. 8, but for  $BB$ . These plots demonstrate that the FoCUS results are consistent for  $BB$  and  $EE$ . The fact that both the  $EE$  and  $BB$  power spectra are properly retrieved demonstrates that the FoCUS maps keep the  $EE$  to  $BB$  power asymmetry.

### A.2. $S$ and $p$ plot for the mock data

Figure A.3 complements Fig. 10 by presenting the joint distribution of  $S$  and  $p$  for the Vansyngel model. The comparison by eye of the three plots show that FoCUS considerably reduces the noise bias on the  $S - p$  relation. We note that the analytical model of Planck Collaboration XII (2020) does not match the Vansyngel model for high  $p$  values. A similar shift between the analytical



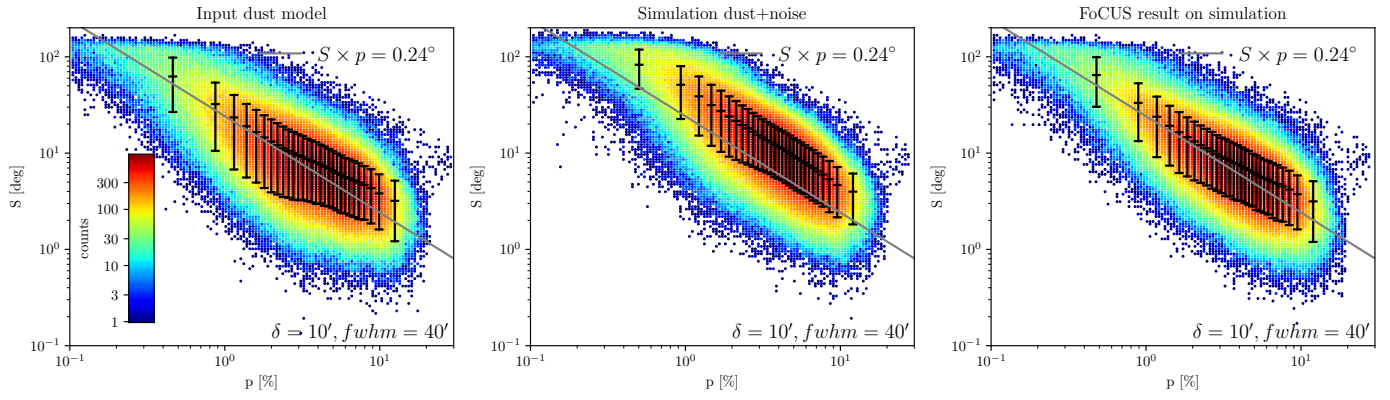
**Fig. A.1.** Power spectra of one realisation for the FoCUS validation on mock data. The plots are the same as in Fig.3, but for  $BB$  spectra.



**Fig. A.2.** Power spectra (left column) and cross-power spectra (right column) for Galactic masks with  $f_{sky} = 0.27$  (bottom) to  $0.73$  (top). The plots are the same as in Fig. 8, but for  $BB$  spectra.

model and the data is observed for the Planck maps in the right panel of Fig. 10.





**Fig. A.3.** Same plot as in Fig. 10 for the Vansyngel model and the mock data. The figure shows the joint distribution of  $S$  and  $p$  for the Vansyngel model (left panel), one realisation of the noisy mock data (middle panel), and the FoCUS denoised maps (right panel).