

FORSE+: Simulating non-Gaussian CMB foregrounds at 3 arcmin in a stochastic way based on a generative adversarial network

Jian Yao^{1,2,3} , Nicoletta Krachmalnicoff^{1,2,3} , Marianna Foschi⁴ ,
Giuseppe Puglisi^{5,6,7} , and Carlo Baccigalupi^{1,2,3} 

¹ SISSA, Via Bonomea 265, 34136 Trieste, Italy
e-mail: jyao@sissa.it

² INFN, Via Valerio 2, 34127 Trieste, Italy

³ IFPU, Via Beirut 2, 34014 Trieste, Italy

⁴ Instituto de Astrofísica de Andalucía, Glorieta de la Astronomía s/n, 18008 Granada, Spain

⁵ Dipartimento di Fisica e Astronomia, Università degli Studi di Catania, Via S. Sofia, 64, 95123 Catania, Italy

⁶ INFN – Sezione di Catania, Via S. Sofia 64, 95123 Catania, Italy

⁷ INAF – Osservatorio Astrofisico di Catania, Via S. Sofia 78, 95123 Catania, Italy

Received 1 March 2024 / Accepted 4 April 2024

ABSTRACT

We present FORSE+, a Python package that produces non-Gaussian diffuse Galactic thermal dust emission maps at arcminute angular scales and that has the capacity to generate random realizations of small scales. This represents an extension of the FORSE (Foreground Scale Extender) package, which was recently proposed to simulate non-Gaussian small scales of thermal dust emission using generative adversarial networks (GANs). With the input of the large-scale polarization maps from observations, FORSE+ has been trained to produce realistic polarized small scales at 3' following the statistical properties, mainly the non-Gaussianity, of observed intensity small scales, which are evaluated through Minkowski functionals. Furthermore, by adding different realizations of random components to the large-scale foregrounds, we show that FORSE+ is able to generate small scales in a stochastic way. In both cases, the output small scales have a similar level of non-Gaussianity compared with real observations and correct amplitude scaling as a power law. These realistic new maps will be useful, in the future, to understand the impact of non-Gaussian foregrounds on the measurements of the cosmic microwave background (CMB) signal, particularly on the lensing reconstruction, de-lensing, and the detection of cosmological gravitational waves in CMB polarization *B*-modes.

Key words. methods: data analysis – cosmic background radiation – diffuse radiation

1. Introduction

The cosmic microwave background (CMB) contains valuable information about the initial conditions in the early Universe and physical processes that happened during its propagation from the last scattering surface to our observation (Hu & Dodelson 2002). In addition to the great success of observations and interpretation of the CMB total intensity signal, mainly sourced by scalar perturbations (Hinshaw et al. 2013; Planck Collaboration VI 2020), the focus of experimental efforts has now been transferred to the observations of the CMB polarization signal. In particular, the quest for the curl or *B*-mode component in the CMB polarized signal has drawn the most attention nowadays, as it can be a unique probe of primordial gravitational waves (PGWs, i.e., tensor-mode cosmological perturbations; Zaldarriaga & Seljak 1997; Kamionkowski et al. 1997; Kamionkowski & Kovetz 2016). These tensor perturbations are predicted by the families of inflationary theories, which describe a history of accelerated expansion in the very early Universe (Guth 1981). Great efforts are being made with many ongoing and future telescopes, including ground-based ones, such as PolarBear¹ (POLARBEAR Collaboration 2022),

SPT² (SPT Collaboration 2023), ACT³ (Madhavacheril et al. 2024), SO⁴ (Ade et al. 2019), CMB-S4⁵ (Abazajian et al. 2016), BICEP⁶ (The BICEP/Keck Collaboration 2022), and AliCPT⁷ (Li et al. 2018), balloon-borne ones such as SPIDER⁸ (Spider Collaboration 2022), and telescopes on satellites, like LiteBIRD⁹ (LiteBIRD Collaboration 2023). The tightest constraint on the amplitude of the PGW, in terms of the tensor-to-scalar ratio, $r \equiv A_T/A_S$, is $r < 0.032$ (Tristram et al. 2022), at a 95% confidence level.

While the scalar perturbations do not contribute to the CMB *B*-modes and only source the parity-even *E*-modes, the gravitational lensing effect from the forming cosmological structures along the propagation of CMB photons, with a typical deflection angle of a few arcmins, converts a small portion of *E*-modes into *B*-ones. These lensing-induced *B*-modes domi-

² South Pole Telescope.

³ Atacama Cosmology Telescope.

⁴ Simons Observatory.

⁵ CMB-Stage-IV.

⁶ Background Imaging of Cosmic Extragalactic Polarization.

⁷ Ali CMB Polarization Telescope.

⁸ Suborbital Polarimeter for Inflation Dust and the Epoch of Reionization.

⁹ Lite (Light) satellite for the studies of *B*-mode polarization and Inflation from the cosmic background Radiation Detection.

¹ Polarization of the Background Radiation experiment.

nate over the PGWs at arcminute angular scales (Hu 2000; Planck Collaboration VIII 2020). Some inflationary models also predict nearly Gaussian statistics of the CMB anisotropies with small deviations known as primordial non-Gaussianity (PNG), which can be used to further constrain the vast space of inflationary theories that go beyond a simple single field slow roll scenario (Bartolo et al. 2004; Planck Collaboration IX 2020). However, the lensing effect will also induce non-Gaussianity into the distribution of CMB photons by distorting the paths and coupling different multipole bins of power spectra.

Despite being a contaminant to PGWs and PNG, CMB lensing carries cosmological information on processes occurring during structure formation, on neutrino masses (Allison et al. 2015), and Dark Energy (Acquaviva & Baccigalupi 2006). Effective techniques capable of reconstructing the lensing field have been developed, mainly relying on detecting the non-Gaussianity presented in the observed CMB maps (Hu & Okamoto 2002; Maniyar et al. 2021). The latest detection of lensing potential comes from the ACT experiment, which achieves a 43σ detection of the CMB lensing signal (Qu et al. 2024).

The primary source of contamination to CMB B -modes arises from astrophysical processes causing diffuse emissions within our own Galaxy. The two main polarized components are represented by the thermal dust and synchrotron emission, dominating at high ($\gtrsim 70$ GHz) and low ($\lesssim 70$ GHz) frequencies, respectively, exceeding the CMB B -mode signal in any frequency and any position on the sky (Krachmalnicoff et al. 2016, 2018). They are caused by dust grains heated by starlight, exhibiting a quasi-thermal emission at about 18 K, and cosmic ray electrons spiraling around the lines of the Galactic magnetic field, respectively. The class of data analysis algorithms capable of extracting the cosmological B -mode anisotropies out of a multi-frequency dataset is known as component separation. Several implementations exist, exploiting the different properties that CMB and foregrounds have, such as frequency dependency and spatial distribution. Depending on the assumptions made about the foregrounds, these methods can generally be classified into blind (Yao et al. 2018; Santos et al. 2021; Zhang et al. 2019; Delabrouille et al. 2009) and parametric ones (Stompor et al. 2009; Planck Collaboration IV 2020; BeyondPlanck Collaboration 2023).

Unlike the CMB, diffuse foregrounds generically exhibit a large degree of non-Gaussianity. This comes primarily from large-scale observations (Coulton & Spergel 2019; Allys et al. 2019) and is expected to also be true at smaller scales, since dust grains are highly locally distributed and the magnetic field in the diffuse interstellar medium is highly turbulent. The foreground non-Gaussianity introduces coupling between different modes in the angular power spectra, which induces non-diagonal terms in the covariance matrices of the observed signal. On the other hand, covariance matrices are usually treated as diagonal under the assumption that all the sky components behave like Gaussian random fields. Additionally, approaches relying on power spectra will inherently ignore the non-Gaussianity, as the two-point functions are insufficient in capturing its presence. Therefore, the existence of non-Gaussianity in the Galactic foregrounds will potentially induce systematic errors in primordial B -mode detection.

Abril-Cabezas et al. (2024) claim that the constraints on r are robust against the non-Gaussianity in the polarized dust at the power spectra level, but they only focus on $30 \leq \ell \leq 300$, the angular scales at which the PGWs cause a maximum B -mode contribution known as recombination bump. At smaller scales

where lensing dominates, non-Gaussian foreground will possibly cause bias to lensing reconstruction (Beck et al. 2020) and de-lensing. Moreover, non-Gaussianity also affects PNG measurements (Cabella et al. 2010).

The main reason for neglecting non-Gaussianity is that observations of foregrounds at arcminute scales are lacking, and so is our capability of generating realizations of simulations of their patterns. Current observations of polarized foregrounds are limited to degree scales over a large sky fraction and can only reach a higher resolution at sub-degree or arcminute scales for portions of the sky (Planck Collaboration XI 2020; Bernardi et al. 2004; Remazeilles et al. 2015).

The way to simulate foregrounds relies on data and phenomenology. Foreground observations at a given frequency are used as templates and extrapolated to other frequency bands by making assumptions about their spectral energy distributions (SEDs). These templates are mostly based on the observations by the *Planck* (Planck Collaboration IV 2020) and *Wilkinson Microwave Anisotropy Probe* (WMAP) satellite observations (Remazeilles et al. 2015; Bennett et al. 2013), which characterize the signal down to a limited angular resolution (around 1°). There exist several widely used packages that simulate Galactic foreground maps, such as the Python Sky Model (PySM, Thorne et al. 2017) and *Planck* Sky Model (Delabrouille et al. 2013). The latest version of PySM3 package (Zonca et al. 2021; PanEx Collaboration, in prep.) uses foreground templates from the latest available observations and extrapolates them to arcminute angular scales. These packages usually fit a power law to the observed large-scale power spectra of foregrounds and extrapolate to small scales with Gaussian realizations of them. Although commonly used, these packages are not able to produce foreground small scales with significant non-Gaussianity comparable to the observed large scales.

Clark & Hensley (2019) provides a data-driven framework to construct three-dimensional Stokes parameter maps of thermal dust emission in position-position-velocity space using only neutral hydrogen data on the basis that HI is strongly correlated with the dust in the diffuse interstellar medium (Lenz et al. 2017). By integrating over the velocity space, they obtained the polarized dust emission maps over the full sky at a resolution of $16.2'$, corresponding to that of HI data, which are in good agreement with the *Planck* observed 353 GHz dust maps in terms of several physical properties such as the polarization fraction and power spectra, although they do not have a thorough discussion about the non-Gaussianity contained in their maps.

Hervías-Caimapo & Huppenberger (2022) uses a large number of filaments in the distribution of the thermal dust grains together with the large-scale template from data to reproduce the main features of the *Planck* 353 GHz map. These include the power spectrum slopes of intensity and polarization maps, the ratios between EE, BB, and TE power spectra, and the level of non-Gaussianity in the total intensity map, which can be controlled by the density of filaments. When focusing on scales of arcminutes or tens of arcminutes, $\ell = 300\text{--}1200$, however, the *Planck* 353 GHz total intensity map has more non-Gaussianity with respect to the generated small scales from the model based on filaments.

Diffuse foreground models are also generated by exploiting magnetohydrodynamic (MHD) simulations. They model the physical processes of the interstellar medium, such as heating and cooling of the gas and the interaction between dust grains and the turbulent magnetic field. The thermal dust emission can be obtained by integrating the emission of dust grains along the line of sight (Padoan et al. 2001; Planck Collaboration Int. XX

2015; Kim et al. 2019). The simulated maps are thus non-Gaussian because of the MHD processes and can reach small scales according to the resolution of the MHD simulation itself. However, the MHD simulations can only reproduce the statistical properties of the Galaxy, and thus fail to generate the specific morphology of Galactic foreground emission. Another important limiting factor of MHD simulations is that they are computationally expensive, especially in achieving a high resolution.

An additional technique consists of simulating the foreground by exploiting innovative algorithms capable of modeling the emission pattern at a high resolution on the basis of the observed one at moderate and low angular scales. The FORSE model introduced by Krachmalnicoff & Puglisi (2021, hereafter KP2021) is able to generate small scales, up to $12'$, starting from the low-resolution polarized dust observations. It utilizes a generative adversarial network (GAN), which is trained to inject small-scale features with statistical properties like the ones observed at a high resolution in the intensity maps.

In this study, we further extend the FORSE algorithm to: (i) produce dust polarization maps at $3'$ resolution, which is vital for checking the stability of lensing reconstruction and de-lensing algorithms in CMB observations, (ii) stochastically generate multiple realizations of small-scale features, which are most important in estimating the uncertainty associated with foreground variations.

Additional uses of deep generative models for dust simulations exist in the literature. For example, Aylor et al. (2021) use GANs to generate simulated total intensity maps from the observed *Planck* generalized needlet internal linear combination (GNILC) map at 353 GHz. Other common generative models, such as variational autoencoders (Thorne et al. 2021) or diffusion models (Heurtel-Depeiges et al. 2023), are also used.

The outline of the paper is as follows. In Sect. 2, we summarize the first version of the FORSE algorithm and present the extension and methodology used to achieve the new version operating at arcminute scale. In Sect. 3, we describe the pre-processing, training, and post-processing of the new model and also present our validation procedure. In Sect. 4, full-sky maps are presented and compared with maps from the latest version of PySM3 package. Conclusions are given in Sect. 5.

2. The ForSE+ model

In this section, we briefly review the basic structure and assumptions of the FORSE algorithm presented in KP2021. We then introduce our new version of the code, FORSE+, which allows one to generate maps of the thermal dust emission with non-Gaussian structures at $3'$ angular resolution, in both a deterministic and a stochastic way.

2.1. Review of the ForSE model: From $80'$ to $12'$

As has already been mentioned, the FORSE model, introduced and validated in KP2021, is based on GANs and allows one to produce non-Gaussian full-sky maps of polarized thermal dust emission at an angular resolution of $12'$ from low-resolution *Planck* observations at $80'$.

Generative adversarial networks are a particular family of networks (Goodfellow et al. 2014) whose characteristic feature is to be composed of two sub-networks called Generator (G) and Discriminator (D), which are trained to compete against each other. In practice, the goal of G is to produce new images that are compared by D with a set of real images that are the training set. Once the training of the two sub-networks is done in an adversar-

ial way, G is able to produce images that have the same statistical properties as those belonging to the training set, in such a way that mock and real images are no longer distinguishable by D .

In the FORSE implementation, the input to G are images at a low resolution ($80'$) of the thermal dust emission observed by the *Planck* satellite and processed through the GNILC method¹⁰ (Planck Collaboration IV 2020). The output of G are small-scale features at $12'$, which are then compared by D with real observations in total intensity at the same angular resolution. We note that different training is performed for total intensity and Stokes Q and U maps, but always having as the training set images of small scales $\tilde{m}_{12'}^{I,20^\circ}$ in total intensity. KP2021 therefore assumed that thermal dust statistical properties of small scales in polarization are the same as for Stokes I maps. We also rely on that assumption in this work.

In KP2021 and in this work, the following definition of “small scales” is used. Let M_{TOT} be a foreground map at a given angular resolution, which can be seen as the sum of two maps containing, respectively, only large- and small-scale structures:

$$M_{\text{TOT}} = M_{\text{LS}} + M_{\text{SS}}. \quad (1)$$

Assuming that the map encoding small-scale features, M_{SS} , is modulated by the large scales, M_{LS} – that is, $M_{\text{SS}} = M_{\text{LS}} \cdot m_{\text{ss}}$ – we have

$$M_{\text{TOT}} = M_{\text{LS}} + M_{\text{LS}} \cdot m_{\text{ss}} = M_{\text{LS}} \cdot \tilde{m}_{\text{ss}}, \quad (2)$$

where $\tilde{m}_{\text{ss}} = m_{\text{ss}} + 1$ represents the small-scale map generated by network G , having as an input M_{LS} .

Although there exist neural networks designed to work on the sphere (Krachmalnicoff & Tomasi 2019), our GAN deals with flat two-dimensional images. For this reason, we had to project the input maps onto flat patches, and project output patches back onto the sphere, after the application of the trained GAN model. We used the same projection strategy as that described in the appendix of KP2021, which divides the GNILC thermal dust Stokes Q and U maps at $80'$ with $N_{\text{side}} = 2048$ ¹¹ into 174 square patches that have 320×320 pixels and a physical side length of 20° . We note that this projection on flat patches and reprojection on the sphere can introduce distortions in the final full-sky map. In order to mitigate this effect, the flat patches overlap each other. We estimate that the final level of distortion induced by our procedure is always less than 7% of the signal (around 2% on average).

From now on we will use $\tilde{m}_{z'}^{X,y^\circ}$ and $M_{z'}^{X,y^\circ}$ to refer to the small-scale (output of the network) and large-scale (input of the network) patches, respectively, where $X = I/Q/U$ defines the Stokes parameters, y° the physical dimension of the patch in degrees, and z' its angular resolution in arcminutes.

The GAN architecture, training procedure, input maps, and output validation of the FORSE model are fully described in KP2021.

2.2. ForSE+: Producing non-Gaussian dust maps at $3'$

In this work, we implement FORSE+, an updated version of the FORSE code and its training procedure, with two objectives:

¹⁰ http://pla.esac.esa.int/pla/aio/product-action?MAP_MAP_ID=COM_CompMap_IQU-thermaldust-gnilc-unires_2048_R3.00.fits

¹¹ Presented in the Hierarchical Equal Area Latitude Pixels (HEALPix; Górski et al. 2005) format.

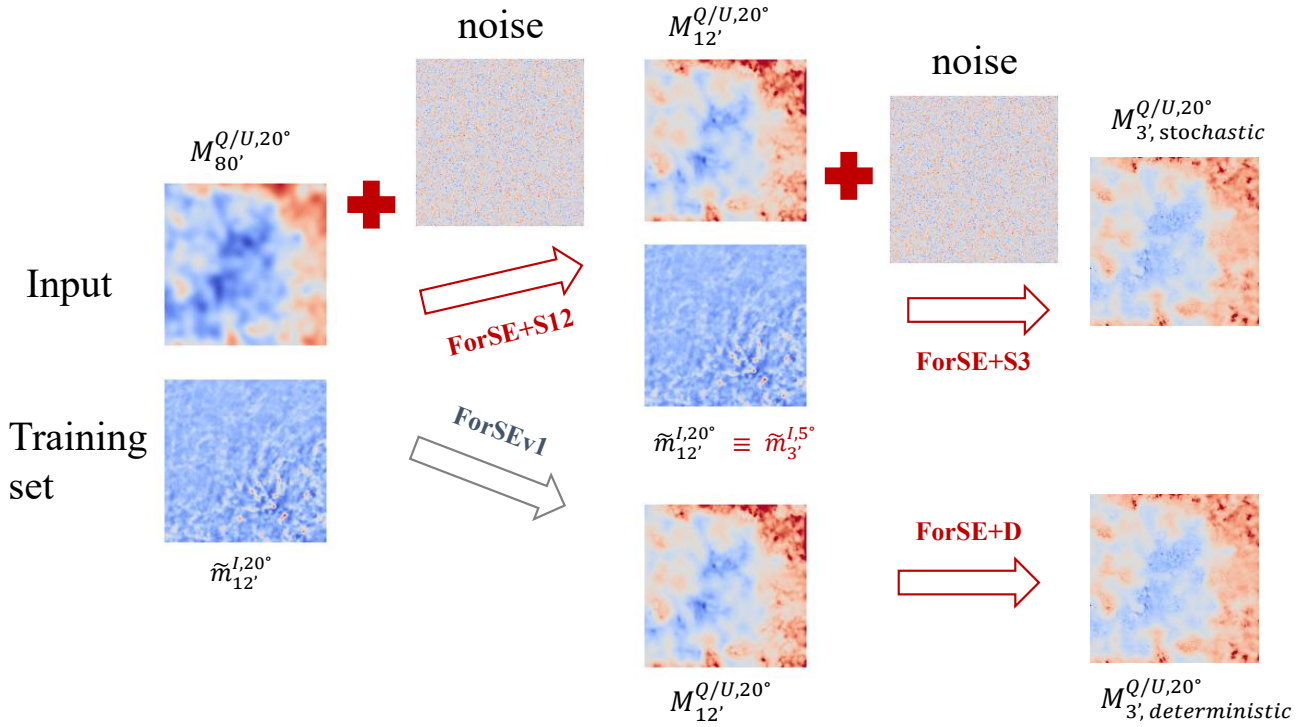


Fig. 1. Structures in FORSE (KP2021) and FORSE+ (this work). FORSE is designed to achieve a 12' resolution from input large scales at 80'. We have three kinds of newly trained models here: FORSE+S12 and FORSE+S3 to produce stochastic maps at 12' and 3', and FORSE+D to generate a deterministic map at 3'. The output, $M_{12'}^{Q/U,20^\circ}$, from FORSE and FORSE+S12 are the input to FORSE+D and FORSE+S3 to get deterministic and stochastic small scales at 3', respectively.

1. To allow the generation of full-sky polarization maps with non-Gaussian features at an enhanced resolution of 3'.
2. Starting from the same low-resolution maps, to generate multiple realizations of stochastic small scales that still have the correct non-Gaussian statistical properties.

In the sections below, we explain the assumptions and methodology used to achieve these two goals. Figure 1 sketches the input and output of the new FORSE+ models and in Table 1 we summarize the new models trained in this paper.

In order to incorporate these extensions, we used the same GAN architecture as for the FORSE model, re-implemented using the new Tensorflow¹² framework (version 2.6.0), and we performed additional fine-tuning steps of some hyperparameters of the networks to improve our results.

2.2.1. Scale invariance assumption

The first goal of our implementation of FORSE+ is to generate a map of polarized dust emission with non-Gaussian features at arcminute angular scales. As we anticipated, this is crucial to understand the possible impact of non-Gaussianity in the extraction of the CMB *B*-mode lensing signals, which peak at these scales. In order to achieve this, we need to find a way to overcome the current limitation in the observational data at our disposal. As a matter of fact, in order to train our GANs we can only rely on a training set composed of 350 patches, with dimensions of $20^\circ \times 20^\circ$ and 320×320 pixels, at an angular resolution of 12', taken from the total intensity GNILC *Planck* map at 353 GHz, described in KP2021. No other observations of thermal dust emission at a higher angular resolution in a portion of the sky large enough to be used as training set are available.

The idea to circumvent this restriction and still be able to reach a resolution higher than 12', even in the absence of a proper training set, is to make a scale-invariance assumption, applying our GAN model in an iterative way. In practice, our set of training squared patches with dimensions of $20^\circ \times 20^\circ$ and a resolution of 12' can equally be treated as having dimensions of $5^\circ \times 5^\circ$ and resolution of 3', since the network does not have a sense of physical units and is only sensitive to the ratio between the dimension of the patch and its angular resolution (i.e., $\tilde{m}_{12'}^{I,20^\circ} \equiv \tilde{m}_{3'}^{I,5^\circ}$). In this way, the same dataset used to train the first version of FORSE in KP2021 can be used to train a new GAN model. This model takes the output of the first iteration of the code as input and generates non-Gaussian scales at 3'. As has been mentioned, this implies that we are assuming scale invariance for the statistical properties of dust emission; that is, scales at 12' have the same properties as those at 3'. This assumption is justified by the fact that current observations of the dust polarization power spectra shows that they can, at the first order, be approximated as a power law as a function of the angular scales (Planck Collaboration XI 2020).

Additional pre- and post-processing of the input patches (including upsampling, smoothing, and the sub-division of patches) is needed to train the GAN in the correct way and to be able to restore full-sky maps, as is described in Sect. 3.2.1.

2.2.2. Stochasticity

Our second goal is to produce different realizations of the non-Gaussian small-scale structures. This is important to estimate the variance of the signal we are simulating as well as the correlation among different angular scales. The way we achieved this was

¹² <https://www.tensorflow.org>

Table 1. Summary of three newly trained models in this paper and the first version of the model proposed in KP2021.

Model	Description (output)	Input
FORSE	KP2021 version to simulate deterministic thermal dust emission maps at 12'	Planck GNILC maps at 80'
FORSE+S12	New model to simulate stochastic thermal dust emission at 12'	Planck GNILC maps at 80' plus random component
FORSE+D	New model to simulate deterministic thermal dust emission maps at 3'	Maps at 20', smoothed from FORSE 12' maps
FORSE+S3	New model to simulate stochastic thermal dust emission maps at 3'	Maps at 20' smoothed from FORSE+S12 plus random component

by simply adding a random component to the large-scale maps that are the input of our GANs. We then trained new models on these signal+random component maps, always using the 350 $\tilde{m}_{12'}^{I,20^\circ}$ patches as the training set.

The random component that we considered was simply generated as a random realization from a uniform distribution in the range $[-1, 1]$. Since our input maps were always re-scaled to have pixel values ranging in the same interval to be compatible with the input of our neural networks, we had a signal-to-noise ratio of ~ 1 . We refer to this as “stochastic training;” in other words, a random component was added to the input signal, FORSE+S, as opposed to the deterministic case (FORSE+D), in which we did not add any random component to the input maps.

3. Pre-processing, training, and post-processing of FORSE+

In this section, we describe the training details of FORSE+S and FORSE+D, including the pre-processing and post-processing steps, and present the results on flat-squared patches, including maps and power spectra, before reprojecting them to full-sky maps.

3.1. FORSE+S to 12'

We first describe how we generate stochastic maps with non-Gaussian features at a resolution of 12'. We also applied a similar procedure to construct maps at 3', as is described in Sect. 3.2. For clarity, we will call FORSE+S12 the model that generates stochastic maps at 12' and FORSE+S3 the one that goes up to 3'.

As was mentioned above, we injected stochasticity into our generative process by simply adding a random component to the low-resolution maps that are the input of our GAN. By doing so, FORSE+S is able to generate non-Gaussian small-scale features that still partially depend on the real observed large-scale structures but will have a different morphology as we change the realization of the random component in the input dataset.

3.1.1. Training and post-processing

As in KP2021, we trained two models for Q and U maps separately. The inputs to the generator, G , were the 174 $M_{80'}^{Q/U,20^\circ}$ patches that together cover the full sky, with dust signal plus an additional random component. The training set, which encodes the target statistical distribution of small scales, was the 350 $\tilde{m}_{12'}^{I,20^\circ}$ maps. The weights of the neural works were updated for 2×10^5 epochs and saved every 500 steps. Since we do not want

to generalize the usage of the trained model – the trained model only needs to predict the output from the training data – it is not a problem if there is an over-fitting during training. Therefore, there was no consideration of a separate validation set during the training process.

Following KP2021, we used the Minkowski functionals (MFs), as implemented in Mantz et al. (2008), as the criteria with which to choose the best epoch. For two-dimensional fields, three kinds of MFs can be built: $\mathcal{V}_0$, $\mathcal{V}_1$, and $\mathcal{V}_2$. They fully describe the statistical properties of the field and represent the covered area ($\mathcal{V}_0$), the boundary length ($\mathcal{V}_1$), and the number difference of connected domains and holes of the image's feature ($\mathcal{V}_2$), as a function of the threshold, ρ (Hadwiger 1957).

During the GAN training process, the goal is to produce small-scale feature maps with statistical properties as close as possible to the ones of the training set. We quantified the level of agreement by calculating the overlapping fractions between the distributions of MFs of the maps in the training set ($\tilde{m}_{12'}^{I,20^\circ}$) and those of the generated ones ($\tilde{m}_{12'}^{Q/U,20^\circ}$). The distributions of MFs are indicated by the $\pm 1\sigma$ variation among the training set or generated maps and the overlapping fractions were computed as the ratio between the intersection area and the total area spanned by the two distributions. In practice, we calculated the MFs overlapping for each saved epoch of G and selected as the best model the one with the highest score. In doing so, we computed the MFs for the output maps, $\tilde{m}_{12'}^{Q/U,20^\circ}$, by applying G to maps with the realization of a random component different from the one used for training.

The best models are obtained after 5500 epochs and 6000 epochs for Q and U , respectively. Their MFs are shown in Fig. 2 and compared with the ones from the training set. The overlap among the distributions is at a level of 50%–60%, comparable with the one obtained in Krachmalnicoff & Puglisi (2023, that includes corrections to KP2021). In comparison, Gaussian small scales have MFs with obviously different shapes from these two sets of maps, as is illustrated in Fig. 7 of KP2021.

We also computed the overlapping fraction of the MFs by considering 100 different realizations of the small-scale maps (obtained by changing the random component in the input). The mean (standard variation) values of the overlapping fraction for $\mathcal{V}_0$, $\mathcal{V}_1$, and $\mathcal{V}_2$ are 59.1% (1.5%), 62.9% (0.6%), and 55.8% (0.7%), respectively, for Stokes Q maps, and 58.8% (2%), 56.3% (1.3%), and 45.5% (1.1%), respectively, for U maps. These numbers show the robustness of FORSE+S in generating stochastic small scales with non-Gaussian high-order statistics.

Since in the training procedure both the input maps and the training sets are normalized in the range of $[-1, 1]$, the output maps also have pixel values in this range, and therefore need to be normalized to restore physical units. We achieved this by

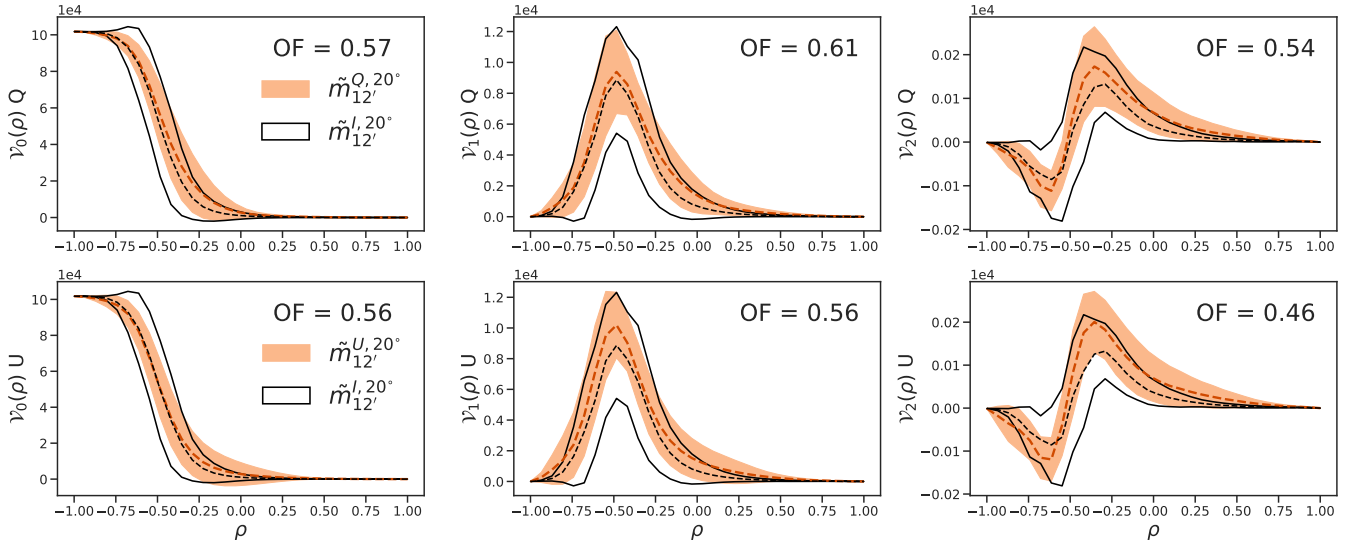


Fig. 2. MFs of small scales at 12' produced by FORSE+S, compared with the ones from the intensity maps in the training set. The overlapping fractions (OFs) of each pair of MFs are also shown. The dashed line represents the mean over the set of different patches and the shaded area is their standard deviation. The distribution of Q maps is shown in the upper panel, and U in the lower panel.

following the procedure of KP2021, hence ensuring that, for each patch and for both Q and U , the amplitude of the power spectrum of the produced small scales matches the extrapolation at higher multipoles of the power spectrum of the low-resolution input maps at 80'.

3.1.2. Results of FORSE+S at 12'

Figure 3 shows $M_{12'}^{Q/U, 20^\circ}$ patches with two different realizations of small scales at 12' from FORSE+S12, after the normalization mentioned above, in the second and third columns, compared with the deterministic results from FORSE in the first column. Both the differences at small scales and consistency at large scales between the outputs from FORSE and FORSE+S12 show the capability of the trained model to produce small-scale features in the map space.

To further validate these maps, we calculated the second-order statistics – the power spectrum – using the Namaster package (Alonso et al. 2019)¹³. In Fig. 4, we show the EE and BB power spectra from the QU squared patches of low-resolution maps, the output from FORSE, and 100 realizations of FORSE+S12 in black, green, and gray, respectively. The mean values of power spectra from 100 stochastic realizations are also shown in red. The output maps from FORSE+S12 are consistent with the deterministic output maps in terms of the power spectra, as was expected.

3.2. FORSE+D and FORSE+S3 to 3'

We describe now the procedures that we followed to generate maps at the resolution of 3', by using our GAN model iteratively both in the case of FORSE+D and for FORSE+S3. The whole procedures, including several pre- and post-processing steps, are sketched in Fig. 5 and described in the following (see also Foschi 2021).

3.2.1. Pre-processing for the training of ForSE+D

As was mentioned above, we reached the resolution of 3' by assuming scale invariance in the statistics of the thermal dust emission. We exploited the output of the first GAN model as the input to a second generative step by using the same training set and considering it to be composed of 350 patches with dimensions of $5^\circ \times 5^\circ$ at a resolution of 3', as was explained in Sect. 2.2.1. Therefore, since the first iteration of FORSE allows one to go from maps with dimensions of $20^\circ \times 20^\circ$ and a resolution of 80' to maps at a resolution of 12', the second iteration can produce $5^\circ \times 5^\circ$ maps at a resolution of 3'. In order to preserve the proportions among all the relevant quantities (i.e., patch dimension, resolution of input, and resolution of output), the input patches for the second iteration should have dimensions of $5^\circ \times 5^\circ$ and a resolution of 20'. We can obtain those patches by smoothing and subdividing the output of FORSE.

In practice, we pre-processed each of the 174 $M_{12'}^{Q/U, 20^\circ}$ maps with 320×320 pixels obtained from the first iteration in the following way:

1. We upsampled each map with 320×320 pixels to 1280×1280 by repeating each pixel four times.
2. We smoothed these maps with a Gaussian kernel in order to reach a resolution of 20'.
3. We subdivided each of these large squared patches into $5^\circ \times 5^\circ$ patches, each one containing 320×320 pixels. We divided each large patch into 49 small ones, with a large overlap among them, made of 160 pixels on each side. This overlap is needed in order to avoid border effects when the composition and reprojection on the sphere is performed.
4. At the end of this procedure, we have a set of $174 \times 49 = 8526$ patches for Q and U , which is the total amount of patches covering the full sky and represents the $M_{20'}^{Q/U, 5^\circ}$ that will be used as the input to the second GAN iteration.

We applied the same pre-processing to the output of FORSE and FORSE+S12 in order to produce maps at 3' in both the deterministic case and the stochastic one. In the stochastic case, we added an additional random component, as is described in Sect. 2.2.2.

¹³ https://namaster.readthedocs.io/en/latest/source/sample_flat.html

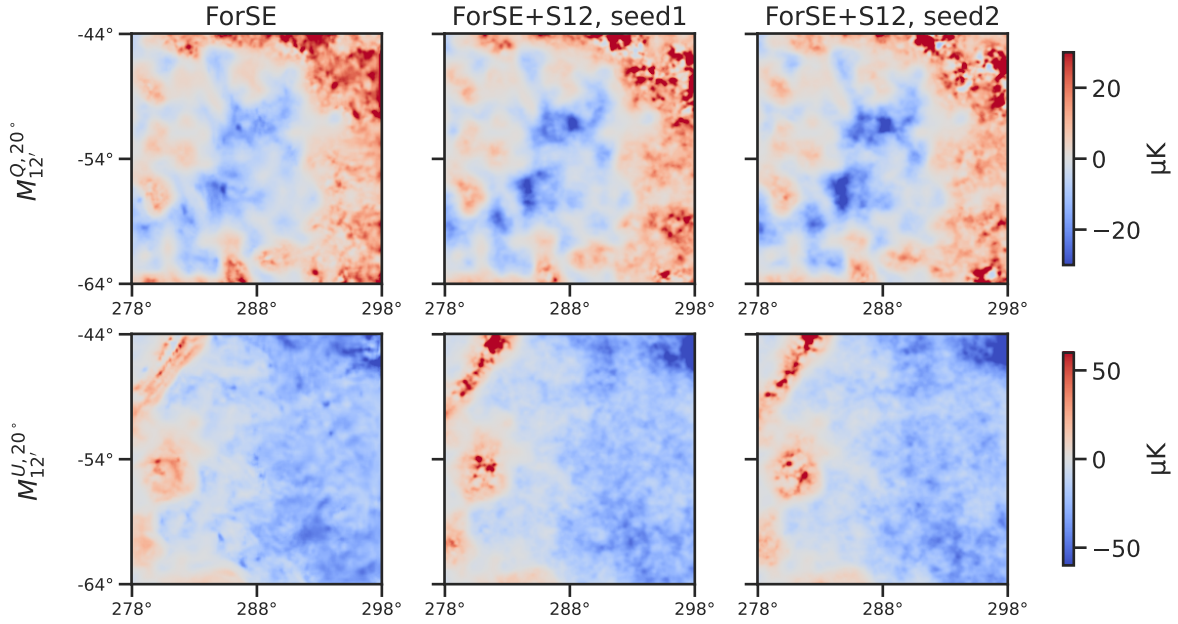


Fig. 3. Maps of $20^\circ \times 20^\circ$ patches at $12'$. From left to right are the deterministic map from FORSE and two stochastic realizations from FORSE+S12. Throughout this paper, all the maps are shown in Galactic coordinates.

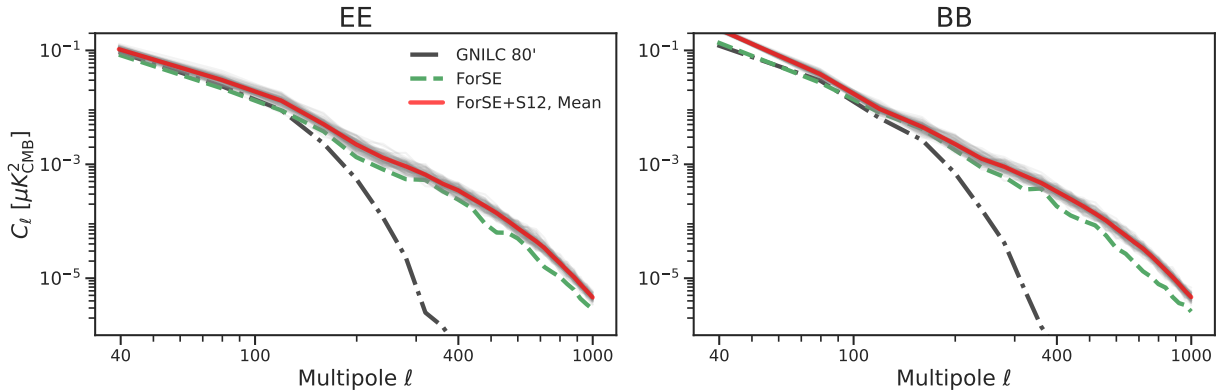


Fig. 4. EE and BB power spectra of the squared patches at $12'$ shown in Fig. 3. Dash-dotted black lines are the power spectra for GNILC $80'$ patches and green lines are for the deterministic $12'$ patches. Gray lines show the power spectra of all the 100 realizations from FORSE+S12 and red lines are the means.

3.2.2. Training and post-processing of ForSE+D

Once the pre-processing steps described above are performed, the obtained 8526 patches can be used to train a new GAN model. However, the time cost is basically linear with the number of patches. We note that the 8526 patches have repeating pixels among them (see the steps above to get 8526 patches); thus, they are actually not independent. In order to save computational time, we fed as input to the generator, G , only a subset of 696 patches, randomly selected from the total 8526 patches. Once the network was trained, we applied it to the remaining patches.

In Fig. 6 we show the MFs of the generated small scales at $3'$. The overlap with the target distribution is at a level of 70–80% in the deterministic case. In the stochastic case (not shown), on the other hand, the overlap ranges between 60% and 70%, showing therefore that we are also able to generate non-Gaussian small-scale features at this higher resolution.

The output of the GAN models are patches, $\tilde{m}_{3'}^{Q/U,5^\circ}$, containing the small-scale features that, as in the first iteration, have pixel values in the range of $[-1, 1]$. We therefore normalized them in physical units before multiplying them with the large-scale maps at the resolution of $20'$, obtaining 8526 $M_{3'}^{Q/U,20^\circ}$ maps.

We recall that each of these patches comes from a subdivision of the larger maps with physical dimensions of $20^\circ \times 20^\circ$ (as is described in Sect. 3.2.1) into 49 sub-patches with a large overlap among each other. Therefore, before reprojecting them on the sphere to obtain the full-sky maps, we needed to recombine them. In order to avoid border effects, we did this by using $\cos^2$ apodization window functions, as is shown in Fig. 7: each sub-patch was weighted with the corresponding window function, then they were summed together to form the 174 $M_{3'}^{Q/U,20^\circ}$ maps that could then be reprojected on the sphere.

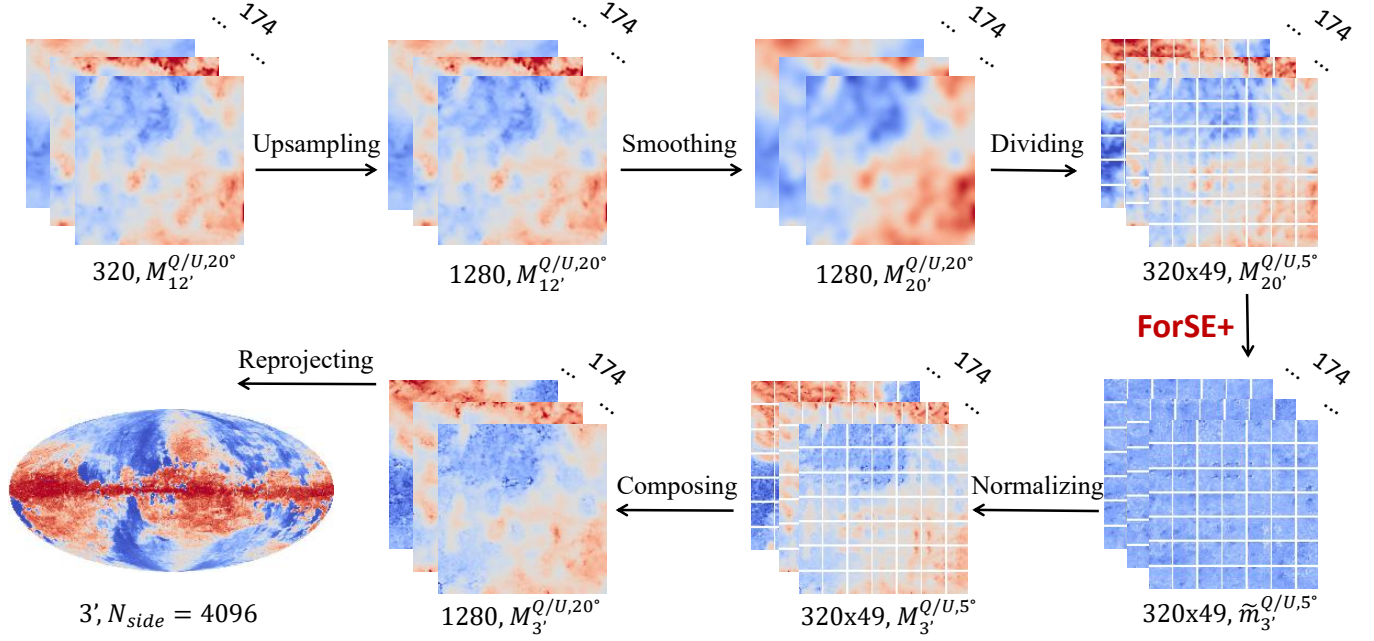


Fig. 5. Pipeline to prepare the training data for FORSE+D (upper row, from left to right) and post-processing to transform the output patches from FORSE+D into the full-sky map at 3' in the HEALPix format (lower row, from right to left). 174 in the upper right corner of each tile means the number of images. The first number beneath the images indicates the number of pixels on each side of the image and the second expression has the same subscript and superscript as that described in Sect. 2.1. We note that after the “Dividing” and before the “Composing” steps, each original image was divided into 49 sub-patches, with a side length of 5°. At the end, all the flat patches were reprojected onto a HEALPix map with N_{side} equal to 4096.

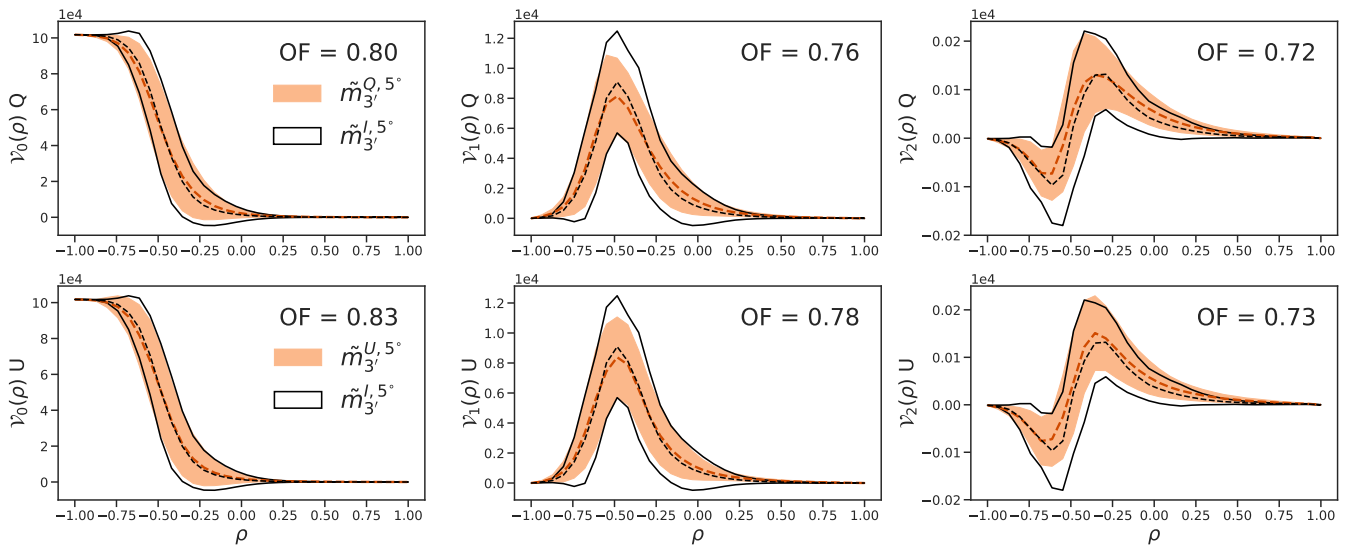


Fig. 6. MFs overlapping between deterministic 3' small scales and intensity small scales.

3.2.3. Results of FORSE+D and FORSE+S3 at 3'

In Fig. 8, we show the comparisons of a selected patch at 80', 12' from FORSE and maps at 3' from FORSE+D, from left to right. It is clear that FORSE+D can inject small-scale structures, following the modulation of the large-scale emission.

The power spectra, computed on the same patch, are shown in Fig. 9. As can be seen, the amplitude at low ℓ matches the one from the low-resolution GNILC maps, and the generated small scales have power at higher multipoles, with no breaks in the power spectrum that follows a power law, as was expected.

FORSE+S3 was utilized to generate stochastic small scales at 3', as is mentioned in Sect. 2.2.2. Two realizations of small scales for the Q map in the range of $[-1, 1]$ with a side length of 5° and centered on $[288^\circ, -61.5^\circ]$ (i.e., at the position of the dashed red box in the upper right plot of Fig. 8) are shown in the upper panel of Fig. 10, and in the lower panel we show the normalized maps combined with the large scales. The differences between the two realizations should be noted and both of them trace the large-scale features of the input maps. We also show the power spectra of 100 realizations in Fig. 9, where we also plot the mean, validating the results of stochasticity on the power spectrum level.

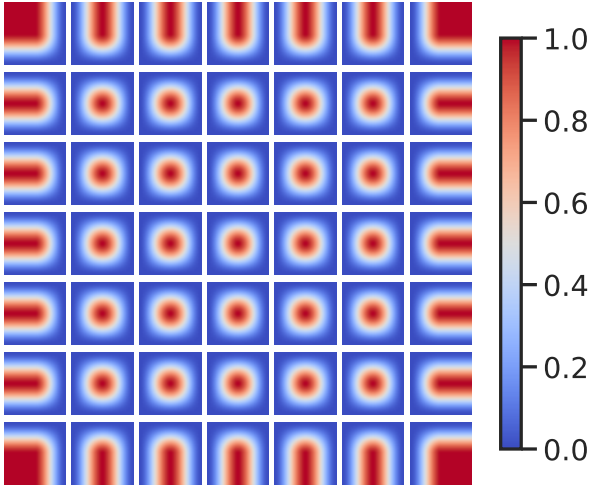


Fig. 7. Apodized window function of each sub-patch at different positions to compose 49 sub-patches with a side length of 5° into a patch with a length of 20° .

4. Validations of full-sky maps from FORSE+

In this section, we show the maps and power spectra after reprojecting the flat patches back to the sphere, following the algorithm in the appendix of KP2021. Before showing the results, we introduce a further step to adjust the E-to-B ratio in the simulated maps to match the observations at large scales.

4.1. Fine-tuning of the E-to-B ratio to match observations

We recall that, due to the limitation of observational data in polarization, we used the statistical properties of the total intensity small scales to be the ground truth of both Q and U during the training process. This implies that the output high-resolution map will have the same power for E and B modes. However, high-frequency observations of the Planck satellite have shown the existence of an asymmetry between the measured power of thermal dust emission in the E and B modes, with $A_{BB}/A_{EE} \sim 0.5$ over the multipole range of $40 \leq \ell \leq 600$ (Planck Collaboration XI 2020). Therefore, in order to match the real observations, we applied a fine-tuning to the full-sky maps, obtained after all the steps mentioned above.

We first transformed QU maps into the harmonic space to obtain the $a_{\ell m}^E$ and $a_{\ell m}^B$ coefficients. Then we applied a factor of $\sqrt{0.5}$ to the $a_{\ell m}^B$ to decrease its power, since $a_{\ell m}^E$ and $a_{\ell m}^B$ have the same variance as was just mentioned. Finally, we transformed the tuned $a_{\ell m}^B$ back to the map space to get maps that match the observed ratio of 0.5 at the power spectra level. We note that the large scales in the output maps of FORSE/FORSE+ remain to be the observed ones, so the tuning was applied only to the small scales; that is, $a_{\ell m}^B$ belonging to $\ell > 135$ ¹⁴ and the transition between large and small scales was smoothed with a sigmoid function to avoid discontinuities in the final power spectra.

The E-to-B ratios of the power spectra calculated for different fractions of sky are shown in Fig. 11. The red lines show the ratios after the tuning steps, which were indeed corrected to ~ 2 for the injected small scales.

4.2. Full-sky maps

We are now ready to present the full-sky maps and specifically, we show the FORSE+D maps at $3'$, at $N_{\text{side}} = 4096$, in Fig. 12, compared with the input GNILC ones at $80'$. The difference between the two highlights the presence of small-scale features in the FORSE+D maps. When zooming into a patch that consists of two $M_{3'}^{Q/U, 20^\circ}$ patches we also find no clear edge effects. These results show that border effects from reprojection are negligible.

In Fig. 13, the power spectra of the maps on the sphere at different resolutions are shown for different sky fractions. As was expected, our maps at $3'$ can further extrapolate the power up to $\ell \sim 3600$, which corresponds to the drop scale of $3'$. We also note the absence of discontinuities at the transition from large scales to small scales ($\ell > 200$), implying that the small scales are produced with a similar pattern as the one at large scales, attributed to the normalization step to rescale the generated small scales in the range of $[-1, 1]$ to the correct amplitude, as was mentioned in Sect. 3.2.2. In order to make a comparison with the latest d9 dust model from PySM3¹⁵ (PanEx Collaboration, in prep.), we also show the power spectra of d9 maps at 353 GHz in blue. Although the latter was produced with entirely different methods, the power spectra of FORSE+D maps at $3'$ are impressively close to those of d9 maps, even for different sky fractions.

We also calculated the power spectra of 100 realizations of full-sky maps at $3'$, generated with FORSE+S3, with a 80% sky mask, up to $\ell_{\text{max}} = 4096$ and with a bin width of $\Delta\ell = 160$. The covariance matrix from these 100 realizations is shown in Fig. 14. The correlation among multipoles at small scales is further evidence that the small scales at $\ell > 800$ were synthesized in a non-Gaussian way. In fact, if the small-scale features were produced with Gaussian properties, we would not observe any non-diagonal correlation, which is verified from our experiment to calculate the covariance matrix of a purely Gaussian field. We devote the next section to further deepening the non-Gaussian properties of our maps.

4.3. Non-Gaussianity measured on the sphere

In previous sections, we considered the MFs in order to characterize non-Gaussianity for flat patches. Here, we expand the analysis to the full-sky maps, by exploiting the algorithms described in Grewal et al. (2022), in order to calculate MFs for spherical maps in the HEALPix format. We focus on the small scales injected by FORSE and FORSE+D and here we used the multipole range $[200, 2048]$ to band-pass filter the raw maps and applied a 80% sky mask to ignore the Galactic plane.

In Fig. 15 we show the results of the MFs statistics applied to the sphere, confirming our previous claims. Differences in the FORSE+D map are visible with respect to the case of Gaussian maps in dashed gray for all three kinds of MFs. The Gaussian maps were generated from a random realization from the power spectra of the FORSE+D Q map. We also show the results of the latest PySM3 d9 dust maps in blue, which exhibits a deviation from Gaussianity, although being similar to the modulated Gaussian maps in orange, whose non-Gaussianity is supposed to derive only from the modulation of large scales. We further checked that the results are robust for the 40% sky mask and for $\ell_{\text{min}} = 500$ and 1000.

¹⁴ Corresponding to $80'$: $\ell = \pi/\theta = \pi/(80/60/180 * \pi) = 135$.

¹⁵ <https://github.com/galsci/pysm>

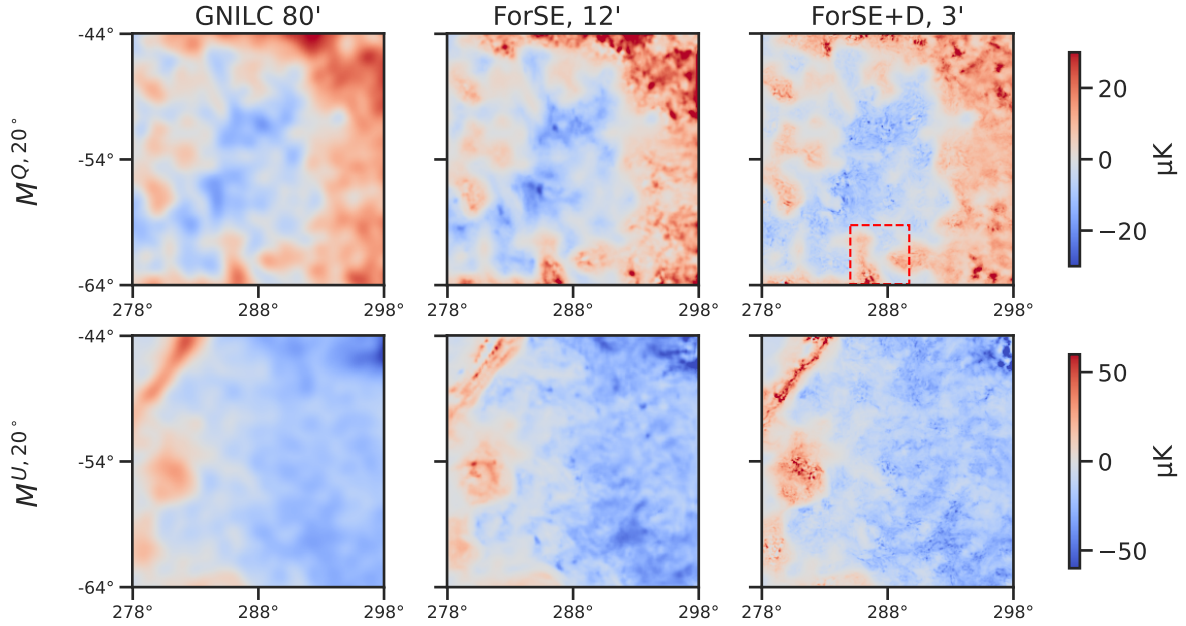


Fig. 8. $20^\circ \times 20^\circ$ patches of three different resolutions. From left to right are GNILC maps at $80'$, FORSE maps at $12'$, and FORSE+D maps at $3'$.

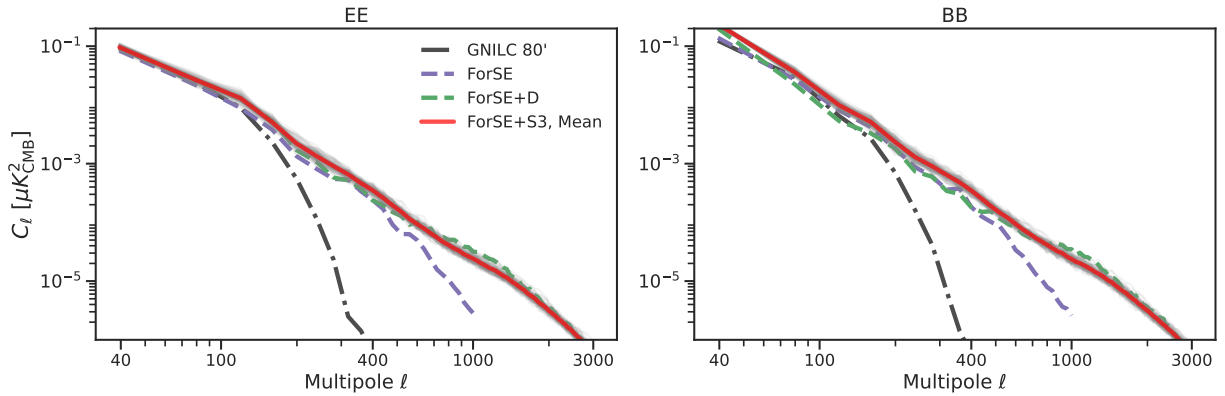


Fig. 9. EE and BB power spectra of the squared patches at $3'$ shown in Fig. 8. Dash-dotted black lines are the power spectra for GNILC $80'$ patches and purple lines are for the deterministic $12'$ patches. Green lines depict the behavior of FORSE+D maps, while gray lines show the power spectra of all the 100 realizations from FORSE+S3 and red lines are the means.

5. Discussion

In this paper, we extend the ability of the FORSE model proposed in KP2021 based on the GAN technique to simulate non-Gaussian stochastic small scales of polarized thermal dust emission up to $3'$. We have trained three new models, which are summarized in Table 1, together with the one trained in KP2021.

Based on the results obtained in KP2021, our first test was to bring stochasticity into our models by adding a random component into the input and to train a new network called FORSE+S12 so that it can generate different maps with different seeds. Minkowski functionals were used to quantify the level of non-Gaussianity in the maps and in Fig. 2 we demonstrate that the polarized thermal dust small scales have a similar level of non-Gaussianity to that in the intensity small scales. Different realizations of maps at $12'$ for a specific patch and the corresponding power spectra are shown in Fig. 4, which indeed have expected variations at small scales and the correct amplitude scaling with multipoles for power spectra.

We then considered the case of a $3'$ angular scale, where lensing B -modes have the strongest signal, and we studied the deter-

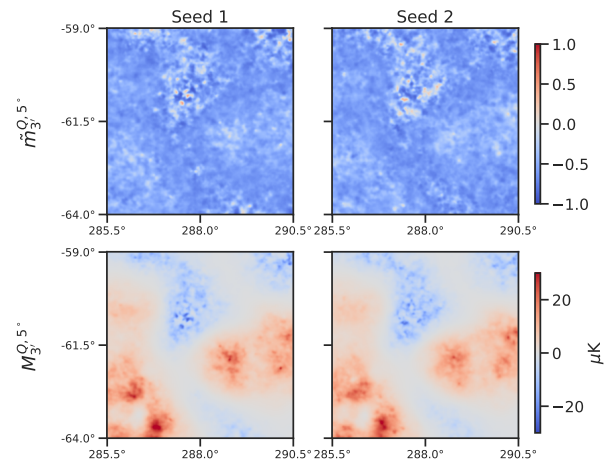


Fig. 10. Upper panel: two stochastic realizations of Q maps, $\tilde{m}_{3'}^{Q,5^\circ}$, from FORSE+S3. Lower panel: $M_{3'}^{Q,5^\circ}$ maps of the upper panel after the normalization step. This patch is at the position of the dashed red box in the upper right plot of Fig. 8.

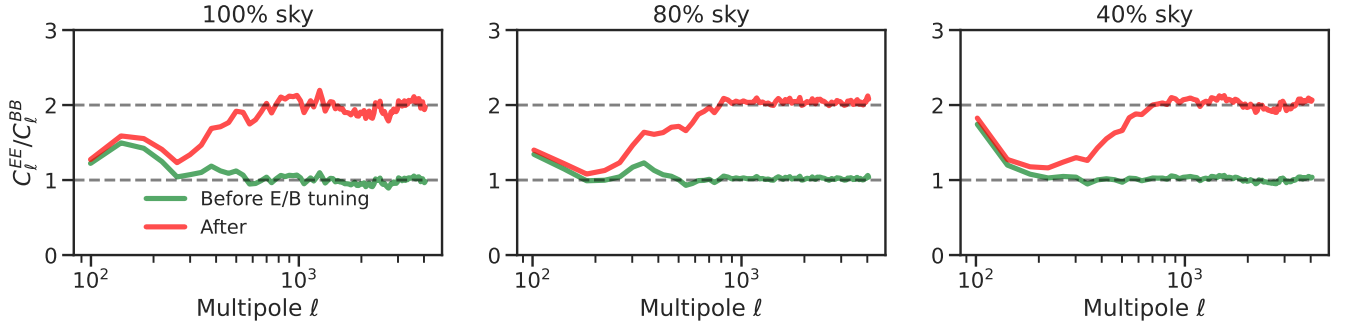


Fig. 11. E-to-B ratios of non-tuned and tuned maps in green and red lines, for different sky masks.

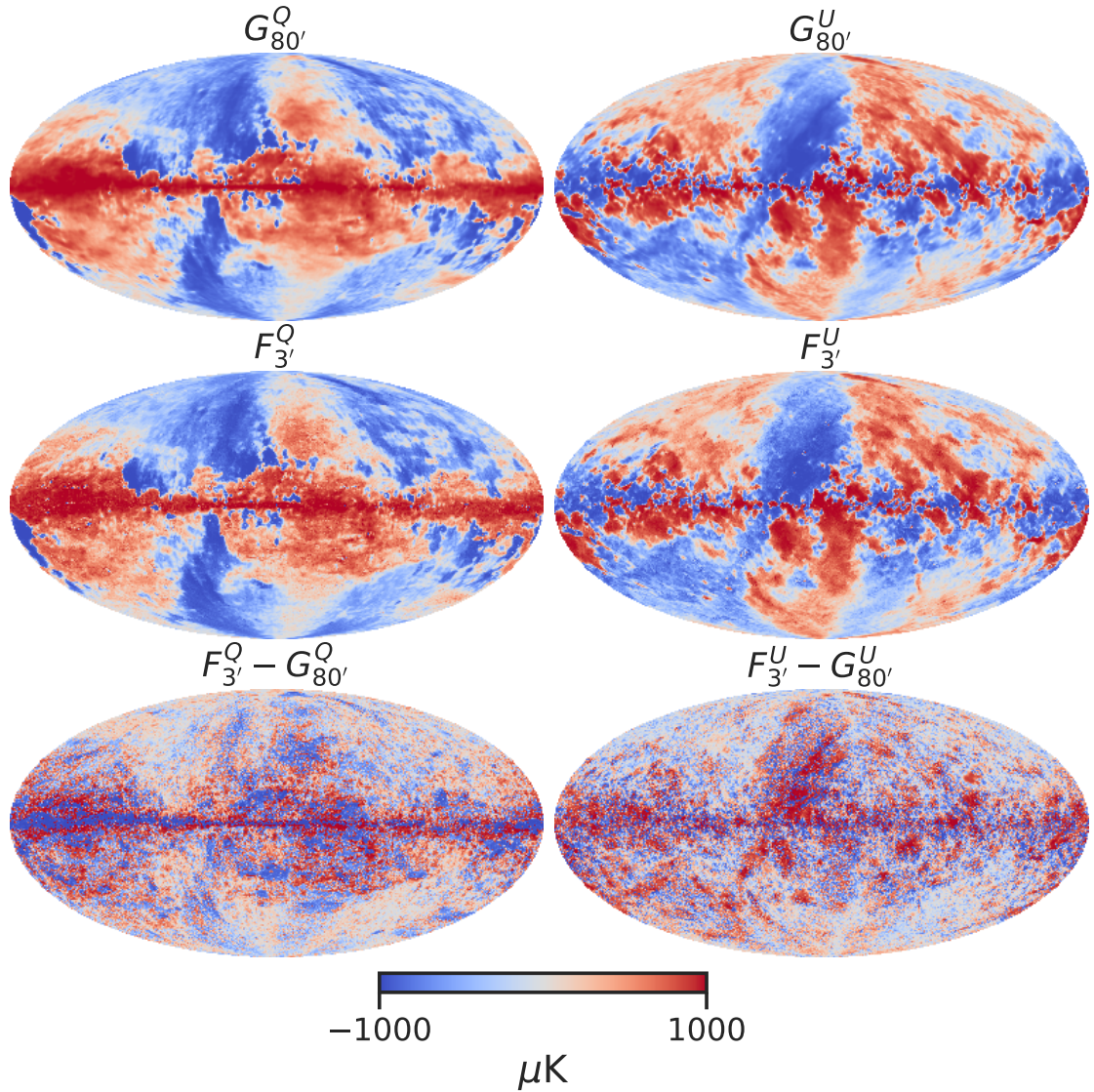


Fig. 12. Top panel: full-sky polarization maps (left: Stokes Q , right: Stokes U) for the GNILC template at a $80'$ angular resolution. These maps are the input to our algorithm. Middle panel: maps with small-scale features, up to $3'$, generated by FORSE+D. Bottom panel: the difference between the two maps. The residuals mostly encode smaller angular scales, as was expected.

ministic case first, FORSE+D. By relying on the scale invariance assumption discussed in Sect. 2.2.1 and the pre- and post-processing steps outlined in Fig. 5, we trained the model to generate maps at $3'$ out of those at $12'$, which are the output from FORSE. MFs distributions in Fig. 6 represent a verification

of the training process. Maps and power spectra of a selected patch of maps at $3'$ are shown in Fig. 8, where the $3'$ power is present as expected. FORSE+S3 was finally trained in order to generate stochastic small scales at $3'$. Two realizations of small scales are shown in Fig. 10, with their power spectra shown

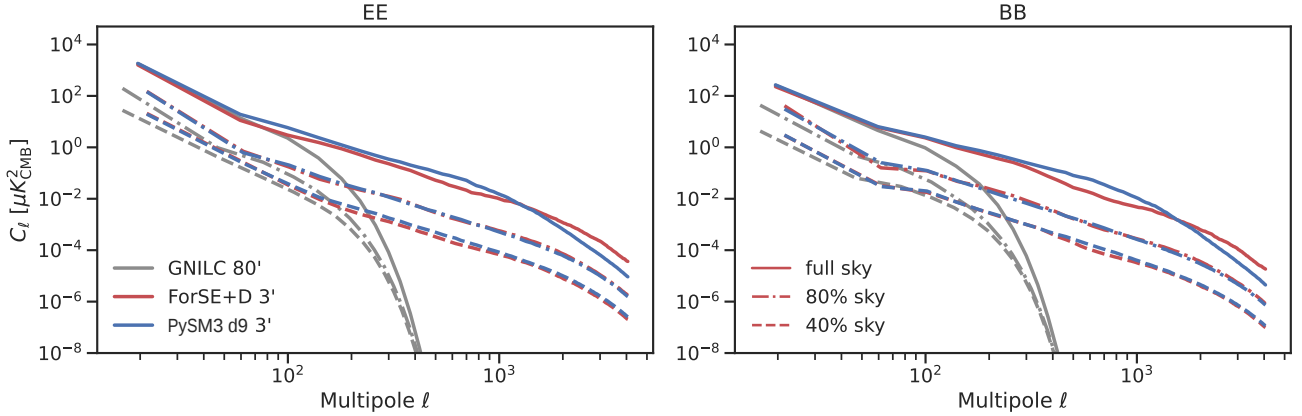


Fig. 13. EE and BB Power spectra of GNILC 80', deterministic 3' in gray and red, with $f_{\text{sky}} = 1, 0.8$, and 0.4 in solid, dash-dotted, and dashed lines, respectively. In order to make a comparison, we also show the power spectra of PySM3 d9 dust map at 353 GHz in blue.

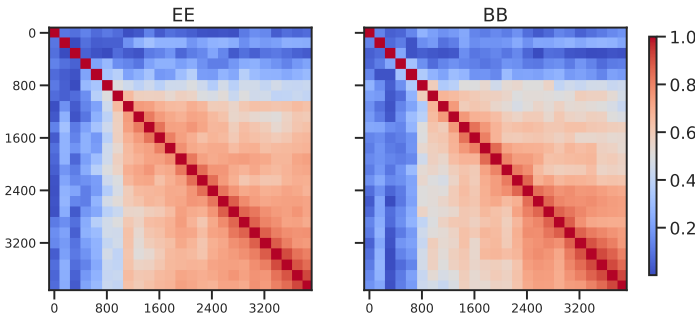


Fig. 14. Covariance matrices of maps at 3' on the sphere with an 80% sky mask, shown in absolute values. Calculated for power spectra of an ℓ -bin range of $[40, 4096]$ with a bin width of 160. Non-diagonal terms are normalized with the diagonal values, i.e., $R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$.

in the right panel of Fig. 8, showing the expected multipole scaling.

After obtaining the small scales on flat patches, we reprojected them onto the celestial sphere in order to get full-sky maps in the HEALPix format, with $N_{\text{side}} = 2048$ for maps at 12' and 4096 for maps at 3'. We show the Mollview projection of the deterministic maps at 3' in Fig. 12 and compare them with the observed *Planck* GNILC maps, validating both the effectiveness of the injected small scales and the reprojection process. The power spectra of deterministic maps at 80', 12', and 3' with 100%, 80%, and 40% sky masks are plotted in Fig. 13, indicating that the small scales generated preserve the same anisotropies as the low-resolution observations.

We further obtain the covariance matrix of power spectra up to $\ell_{\text{max}} \sim 4000$ from 100 realizations of maps at 3' in Fig. 14, which has a strong correlation between different multipoles at small scales, and thus highlights the non-Gaussianity in our maps. We note that for a Gaussian field, the variances along the diagonal line are smaller than the variances we obtained here and the off-diagonal correlation is zero. We repeated the calculation of the covariance matrix for the PySM3 d11 model, which is also designed to generate multiple realizations of small scales of thermal dust emission (PanEx Collaboration, in prep.), and find that off-diagonal elements are also close to zero, meaning that the level of non-Gaussianity of the injected small scales on the full sky is small.

Finally, we measured the level of non-Gaussianity using MFs in the simulated full-sky maps in Fig. 15, and compared it with the small scales simulated within the Gaussian assumption. The difference between the MFs obtained from the two kinds of maps is another clear indication of the non-Gaussian small-scale component generated with FORSE+.

Once the steps above are achieved, we can use the synthetic dust Stokes Q and U maps at 353 GHz at 3' resolution as a template and appropriately scale across different frequencies by taking the observed dust SEDs. The series of maps can then serve as a simulation suite to obtain the scatter of component separation against variations in the foreground realization. Moreover, high-resolution maps are essential in order to validate the lensing reconstruction methods, tested so far on foreground models where the arcminute structure is either absent or Gaussian (Lonappan et al. 2024), or based on models of the Galactic magnetic fields in order to reproduce a non-Gaussian pattern (Beck et al. 2020). In a forthcoming work, we will investigate if any variation of the results shown in Beck et al. (2020) occurs when considering the foreground maps exhibiting non-Gaussianity generated through GANs in this paper.

With the new version of PySM3 under development, we capitalize on integrating our maps into a new model of the latest PySM packages for public use. Maps will be shared upon request to the authors. The code to generate maps has also been made publicly available¹⁶. Finally, we report here future improvements in the FORSE algorithm. First, more observations are needed. In fact, the assumptions made in Sect. 2.2 are somewhat a compromise due to the lack of observation at the required resolution. The models will get more reliable as more observations become available. Second, by considering the network architecture, the loss function accounting for the non-Gaussianity of the targets produced by the GAN may be considered, as in the current implementation this process is not automatized. A way to quantify the level of non-Gaussianity with a formalism that is differentiable with respect to the input pixels would be desirable, as then we could construct the loss function of the network in order to include the information of non-Gaussianity, which will effectively guide the generator. The last point we want to mention is that the operation of adding noise is to some extent like the training process of diffusion models in deep learning, which is more natural than what is done in this work, so if we

¹⁶ <https://github.com/yaojian95/ForSEplus>

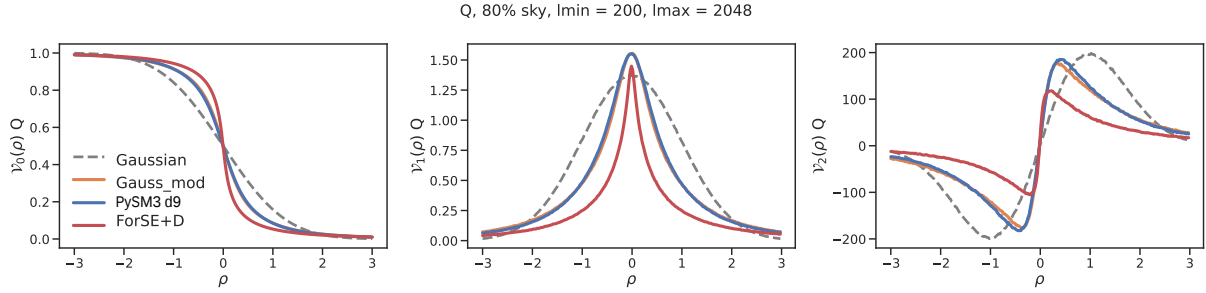


Fig. 15. Three kinds of MFs for the FORSE+D Q full-sky map at $3'$ with an 80% sky mask (red). The raw map is band-pass filtered to keep small scales only, corresponding to the multipole range [200, 2048]. The dashed gray lines show the MFs of the Gaussian maps. Results for PySM3 d9 and modulated Gaussian maps are also shown in blue and orange, respectively.

turn to use the network of diffusion model, it may have a better performance.

Acknowledgements. The authors acknowledge partial support by the Italian Space Agency LiteBIRD Project (ASI Grants No. 2020-9-HH.0 and 2016-24-H.1-2018), as well as the InDark and LiteBIRD Initiative of the National Institute for Nuclear Physics, and the RadioForegroundsPlus Project (HORIZON-CL4-2023-SPACE-01, GA 101135036). G.P. acknowledges support from Italian Research Center on High Performance Computing Big Data and Quantum Computing (ICSC), project funded by European Union – NextGenerationEU – and National Recovery and Resilience Plan (NRRP) – Mission 4 Component 2 within the activities of Spoke 3 (Astrophysics and Cosmos Observations). J.Y. thanks Yun Zheng for useful discussions. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a US Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. Software used: Astropy (Astropy Collaboration 2013, 2018), PySM (Zonca et al. 2021; Thorne et al. 2017), NumPy (Harris et al. 2020), Namaster (Alonso et al. 2019), numba (Lam et al. 2015), Tensorflow (Abadi et al. 2015), reproject (Robitaille et al. 2023), and HEALPix (Zonca et al. 2019; Górski et al. 2005).

References

- Abadi, M., Agarwal, A., Barham, P., et al. 2015, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, <https://www.tensorflow.org>
- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv e-prints, [arXiv:1610.02743]
- Abri-Cabezas, I., Hervias-Caimapo, C., von Hausegger, S., Sherwin, B. D., & Alonso, D. 2024, *MNRAS*, **527**, 5751
- Acquaviva, V., & Baccigalupi, C. 2006, *Phys. Rev. D*, **74**, 103510
- Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, *JCAP*, **2019**, 056
- Allison, R., Caucal, P., Calabrese, E., Dunkley, J., & Louis, T. 2015, *Phys. Rev. D*, **92**, 123535
- Allys, E., Levrier, F., Zhang, S., et al. 2019, *A&A*, **629**, A115
- Alonso, D., Sanchez, J., Slosar, A., & LSST Dark Energy Science Collaboration 2019, *MNRAS*, **484**, A127
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
- Aylor, K., Haq, M., Knox, L., Hezaveh, Y., & Perreault-Levasseur, L. 2021, *MNRAS*, **500**, 3889
- Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004, *Phys. Rep.*, **402**, 103
- Beck, D., Errard, J., & Stompor, R. 2020, *JCAP*, **2020**, 030
- Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, *ApJS*, **208**, 20
- Bernardi, G., Carretti, E., Fabbri, R., et al. 2004, *MNRAS*, **351**, 436
- BeyondPlanck Collaboration (Andersen, K. J., et al.) 2023, *A&A*, **675**, A1
- Cabella, P., Pietrobon, D., Veneziani, M., et al. 2010, *MNRAS*, **405**, 961
- Clark, S. E., & Hensley, B. S. 2019, *ApJ*, **887**, 136
- Coulton, W. R., & Spergel, D. N. 2019, *JCAP*, **2019**, 056
- Delabrouille, J., Cardoso, J. F., Le Jeune, M., et al. 2009, *A&A*, **493**, 835
- Delabrouille, J., Betoule, M., Melin, J. B., et al. 2013, *A&A*, **553**, A96
- Foschi, M. 2021, Master's Thesis, University of Trento
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv e-prints [arXiv:1406.2661]
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Grewal, N., Zuntz, J., Tröster, T., & Amon, A. 2022, *Open J. Astrophys.*, **5**, 13
- Guth, A. H. 1981, *Phys. Rev. D*, **23**, 347
- Hadwiger, H. 1957, *Vorlesungen ueber Inhalt, Oberfläche und Isoperimetrie, Die Grundlehren der mathematischen Wissenschaften* (Springer)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
- Hervias-Caimapo, C., & Huppenberger, K. M. 2022, *ApJ*, **928**, 65
- Heurtel-Depeiges, D., Burkhart, B., Ohana, R., & Régalo-Saint Blancard, B. 2023, arXiv e-prints [arXiv:2310.16285]
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hu, W. 2000, *Phys. Rev. D*, **62**, 043007
- Hu, W., & Dodelson, S. 2002, *ARA&A*, **40**, 171
- Hu, W., & Okamoto, T. 2002, *ApJ*, **574**, 566
- Kamionkowski, M., & Kovetz, E. D. 2016, *ARA&A*, **54**, 227
- Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, *Phys. Rev. D*, **55**, 7368
- Kim, C.-G., Choi, S. K., & Flauger, R. 2019, *ApJ*, **880**, 106
- Krachmalnicoff, N., & Puglisi, G. 2021, *ApJ*, **911**, 42
- Krachmalnicoff, N., & Puglisi, G. 2023, *ApJ*, **947**, 93
- Krachmalnicoff, N., & Tomasi, M. 2019, *A&A*, **628**, A129
- Krachmalnicoff, N., Baccigalupi, C., Aumont, J., Bersanelli, M., & Mennella, A. 2016, *A&A*, **588**, A65
- Krachmalnicoff, N., Carretti, E., Baccigalupi, C., et al. 2018, *A&A*, **618**, A166
- Lam, S. K., Pitrou, A., & Seibert, S. 2015, *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, 7
- Lenz, D., Hensley, B. S., & Doré, O. 2017, *ApJ*, **846**, 38
- Li, H., Li, S.-Y., Liu, Y., et al. 2018, *Natl. Sci. Rev.*, **6**, 145
- LiteBIRD Collaboration (Allys, E., et al.) 2023, *Progr. Theor. Exp. Phys.*, **2023**, 042F01
- Lonappan, A. I., Namikawa, T., Piccirilli, G., et al. 2024, *JCAP*, **2024**, 009
- Madhavacheril, M. S., Qu, F. J., Sherwin, B. D., et al. 2024, *ApJ*, **962**, 113
- Maniyar, A. S., Ali-Haïmoud, Y., Carron, J., Lewis, A., & Madhavacheril, M. S. 2021, *Phys. Rev. D*, **103**, 083524
- Mantz, H., Jacobs, K., & Mecke, K. 2008, *J. Stat. Mech.: Theory Exp.*, **2008**, 12015
- Padoan, P., Goodman, A., Draine, B. T., et al. 2001, *ApJ*, **559**, 1005
- Planck Collaboration IV. 2020, *A&A*, **641**, A4
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Planck Collaboration VIII. 2020, *A&A*, **641**, A8
- Planck Collaboration IX. 2020, *A&A*, **641**, A9
- Planck Collaboration XI. 2020, *A&A*, **641**, A11
- Planck Collaboration Int. XX. 2015, *A&A*, **576**, A105
- POLARBEAR Collaboration 2022, *ApJ*, **931**, 101
- Qu, F. J., Sherwin, B. D., Madhavacheril, M. S., et al. 2024, *ApJ*, **962**, 112
- Remazeilles, M., Dickinson, C., Banday, A. J., Bigot-Sazy, M. A., & Ghosh, T. 2015, *MNRAS*, **451**, 4311
- Robitaille, T., Ginsburg, A., Mumford, S., et al. 2023, <https://doi.org/10.5281/zenodo.7584411>
- Santos, L., Yao, J., Zhang, L., et al. 2021, *A&A*, **650**, A65
- Spider Collaboration 2022, *ApJ*, **927**, 174
- SPT Collaboration 2023, *Phys. Rev. D*, **108**, 023510
- Stompor, R., Leach, S., Stivoli, F., & Baccigalupi, C. 2009, *MNRAS*, **392**, 216
- The BICEP/Keck Collaboration 2022, *ApJ*, **927**, 77
- Thorne, B., Dunkley, J., Alonso, D., & Naess, S. 2017, *MNRAS*, **469**, 2821
- Thorne, B., Knox, L., & Prabhu, K. 2021, *MNRAS*, **504**, 2603
- Tristram, M., Banday, A. J., Górski, K. M., et al. 2022, *Phys. Rev. D*, **105**, 083524
- Yao, J., Zhang, L., Zhao, Y., et al. 2018, *ApJS*, **239**, 36
- Zaldarriaga, M., & Seljak, U. 1997, *Phys. Rev. D*, **55**, 1830
- Zhang, P., Zhang, J., & Zhang, L. 2019, *MNRAS*, **484**, 1616
- Zonca, A., Singer, L., Lenz, D., et al. 2019, *J. Open Source Softw.*, **4**, 1298
- Zonca, A., Thorne, B., Krachmalnicoff, N., & Borrill, J. 2021, *J. Open Source Softw.*, **6**, 3783