

ForSE: a GAN based algorithm for extending CMB foreground models to sub-degree angular scales

NICOLETTA KRACHMALNICOFF^{1, 2, 3} AND GIUSEPPE PUGLISI^{4, 5}

¹SISSA, Via Bonomea 265, 34136 Trieste, Italy

²INFN, Via Valerio 2, 34127 Trieste, Italy

³IFPU, Via Beirut 2, 34014 Trieste, Italy

⁴Space Sciences Laboratory, University of California, Berkeley, CA 94720, USA

⁵Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

ABSTRACT

We present FORSE (Foreground Scale Extender), a novel Python package which aims at overcoming the current limitations in the simulation of diffuse Galactic radiation, in the context of Cosmic Microwave Background experiments (CMB). FORSE exploits the ability of generative adversarial neural networks (GANs) to learn and reproduce complex features present in a set of images, with the goal of simulating realistic and non-Gaussian foreground radiation at sub-degree angular scales. This is of great importance in order to estimate the foreground contamination to lensing reconstruction, de-lensing and primordial *B*-modes, for future CMB experiments. We applied this algorithm to Galactic thermal dust emission in both total intensity and polarization. Our results show how FORSE is able to generate small scale features (at 12 arc-minutes) having as input the large scale ones (80 arc-minutes). The injected structures have statistical properties, evaluated by means of the Minkowski functionals, in good agreement with those of the real sky and which show the correct amplitude scaling as a function of the angular dimension. The obtained thermal dust Stokes *Q* and *U* full sky maps are publicly available at <https://portal.nersc.gov/project/sobs/users/ForSE/>

Keywords: cosmology: cosmic background radiation — diffuse radiation — neural networks — generative adversarial networks

1. INTRODUCTION

In the last ~ 50 years, observations of the Cosmic Microwave Background (CMB) have allowed to unveil the physics and the evolution history of our Universe with an accuracy that was unimaginable when the CMB radiation was first detected in 1965 (Penzias & Wilson 1965). Thanks to the measurements of three dedicated satellites and of several ground based and balloon-borne experiments that have successfully taken data over the decades, the CMB total intensity anisotropies have been constrained with great precision. The accuracy achieved on the main cosmological parameters is today at the level of percent or sub-percent, returning a picture of our Universe in large agreement with the Λ CDM (Lambda cold dark matter) cosmological model (Planck Collaboration VI 2020; Aiola et al. 2020).

Corresponding author: Nicoletta Krachmalnicoff
nkrach@sissa.it

After the release, in 2018, of the legacy data of the Planck satellite (Planck Collaboration I 2020), CMB observations have officially entered a new era. The goal of the current dedicated ground based experiments, as well as of the future satellites, is to measure with high accuracy the anisotropies in the CMB polarized signal, and to constrain its power spectra at all the angular scales, from tens of degrees to few arc-minutes.

CMB is partially linear polarized due to scattering events of photons against free electrons in the primordial Universe and in presence of quadrupole anisotropies. The resulting polarization field, observed through the Stokes *Q* and *U* parameters, can be decomposed into a gradient (*E*-mode) and a curl (*B*-mode) component, respectively even and odd under parity inversion (Zaldarriaga & Seljak 1997; Kamionkowski et al. 1997). *E*-modes have been measured by several experiments in the past years, and their power spectrum as well as their correlation with the total intensity (*T*) field have been reconstructed up to the sub-degree angular scales (Planck Collaboration V 2020). On the other hand, it is in the

B-mode signal that potential breakthrough discoveries might be hidden.

Primordial *B*-modes are sourced by gravitational (tensor) perturbations in the early Universe, contrary to density (scalar) perturbations that can only produce the *E*-mode polarization pattern. The presence of a background of primordial gravitational waves (GWs), as a manifestation of quantum gravity, is predicted by the Inflationary paradigms, i.e. a family of theories that hypothesize a phase of accelerated expansion in the very early Universe (Guth 1981; Linde 1982; Starobinskii 1979). The detection of the imprint of these GWs in the curl component of the CMB polarized signal would not only be fundamental to shed light on Inflation, but would also have a crucial impact on several fields in fundamental Physics.

The primordial *B*-mode power spectrum is characterized by the presence of two main bumps at intermediate and large angular scales: (i) the recombination bump, at multipoles $\ell \sim 80$ (corresponding to an angular dimension of about 2° in the sky); (ii) the reionization bump, located at $\ell \lesssim 20$ (tens of degrees in the sky) and whose presence is due to a second phase of scattering events that CMB photons undertook during the reionization era. The amplitude of the primordial spectrum is parametrized by the tensor-to-scalar ratio r , whose value varies significantly among different inflationary models (Baumann 2009). The current upper limit is $r < 0.06$ (95% confidence level) coming from the combination of the BICEP2/Keck Array data with foreground observations from Planck and WMAP (BICEP2 Collaboration et al. 2018); this limit can be further shrunk to $r < 0.044$ if full sky CMB data from Planck are also considered (Tristram et al. 2020).

At smaller angular scales, the *B*-mode polarization pattern is generated as a consequence of gravitational lensing. As a matter of fact, CMB photons get deflected along their path, due to the presence of gravitational potentials generated by forming cosmological structures in the Universe. This effect causes a leak of the existing *E*-modes (generated by the primordial scalar perturbations) into the *B* ones. The resulting lensing *B*-mode power spectrum shows a broad peak, that has its maximum at $\ell \sim 1000$, corresponding to an angular scale of few arc-minutes. The amplitude of the lensing spectrum at $\ell \simeq 100$ is equivalent to the one of a primordial *B*-mode signal with $r \simeq 0.02$, and thus needs to be taken into account when targeting smaller values of r . The *B*-mode lensing signal has been first detected by the South Pole Telescope (SPTpol) in cross correlation with Cosmic Infrared Background (CIB) data (Hanson et al. 2013), and subsequently the lensing peak has been suc-

cessfully measured by the PolarBear experiment (The Polarbear Collaboration 2014), the Atacama Cosmology Telescope (Louis et al. 2017, ACTpol), and others.

In the recent years, several experiments have been built or are currently being designed with the primary goal of observing the primordial *B*-modes and constrain the value of the r parameter. Ground based experiments are focusing on the observation of the CMB polarization at intermediate and small angular scales, in order to detect the *B*-modes recombination bump. Among these experiment the Simons Observatory (SO) is currently being deployed with observations that will begin ~ 2022 from the Atacama desert (The Simons Observatory Collaboration 2019). SO will consist in three small-aperture 0.5-m telescopes (SATs) and one large-aperture 6-m telescope (LAT). The SATs will target the degree angular scales (corresponding to multipoles $30 \lesssim \ell \lesssim 500$) in about 10% of the sky, while LAT observations will focus in the arc-minute angular scales ($\ell \gtrsim 1000$) covering about 40% of the sky. The combination of the measurements of the two instruments is necessary to achieve the goal of constraining the tensor-to-scalar ratio: LAT observations at small scales will allow to reconstruct the lensing potential that causes the CMB lensing distortion, and consequently will permit to perform de-lensing of the SAT maps. With this procedure SO will reach the target sensitivity on r with $\sigma(r) \simeq 0.003$. Similarly, but on a longer time scale (2027), the Stage-IV network of ground-based observatories (CMB-S4) will combine observations at intermediate and small angular scales to target detection of primordial *B*-modes with $r \gtrsim 0.001$ (Abazajian et al. 2016). A similar target on the tensor-to-scalar ratio is forecasted for the LiteBIRD satellite, which has been selected as strategic large mission by the JAXA (Japan Aerospace Exploration Agency) space agency and will observe the CMB polarized signal at very large angular scales, with full sky coverage from the space ($\ell \lesssim 200$), with operation starting in ~ 2028 (Sugai et al. 2020).

One of the greatest challenges in the detection of primordial *B*-modes is represented by the contamination coming from Galactic foregrounds. In fact, polarized radiation from synchrotron and thermal dust emission pollutes the CMB tensor modes at the degree scales, everywhere in the sky and at every possible observation frequency (Baccigalupi 2003; Page et al. 2007; Krachmalnicoff et al. 2018, 2016; Planck Collaboration Int. XXX 2016). Several algorithm that perform the so-called *component separation*, with the goal of getting a clean CMB maps have been developed in the recent years, including blind or parametric methods, in pixel

or harmonic domain (see for example: [Planck Collaboration IX 2016; Stompor et al. 2016](#)).

At smaller angular scales the Galactic emission amplitude diminishes but it is still strong enough to significantly contaminate the CMB lensing signal. This must be taken into account as the effect of foreground on lensing reconstruction and delensing could lead to residual spurious signals preventing the detection of primordial B -modes ([Beck et al. 2020](#)).

The properties of both thermal dust and synchrotron emission are known with reasonable accuracy at large angular scales ($\gtrsim 1^\circ$) thanks to the full sky multi-frequency observations of the WMAP and Planck satellites, and these data are used to build our large scale foreground models and templates ([Planck Collaboration X 2016](#)). On the other hand, we do not have data with the sensitivity required to characterize polarized foregrounds at sub-degree resolution on large portion of the sky. Nonetheless, we know that Galactic emission have a non trivial statistics, with a non-Gaussian, non-stationary signal and a complicated morphology that reflects the complexity of the Galactic magnetic field ([Coulton & Spergel 2019](#)); we expect this to be true at all angular scales.

Being able to simulate the complexity of foregrounds, from the larger scales to the arc-minute ones, is fundamental, in order to understand their impact on both lensing and primordial B -mode reconstruction and get prepared to the challenges that we are going to face with the next generation of CMB experiments. The current lack of data represents an important limitation to achieve this goal, which is not going to be solved in the near future, as the high sensitivity/high resolution foreground observations will be carried out at the same time as the CMB one, with the coming multi-frequency observatories.

In this work, we propose to overcome the current limitation of foreground models taking advantage of Generative Adversarial Neural Networks (GANs), with the goal of simulating realistic non-Gaussian polarized Galactic emission at sub-degree angular scales. In particular, we show how GANs can be trained to understand the underlying relation between large (degree) and small (tens of arc-minutes) angular scales, when reliable data are available. The trained network can be applied to reproduce this relation and the statistical properties of small scales whenever data are not available.

The paper is organized as follows: in Section 2 we summarize the current knowledge of Galactic foreground emission in the context of CMB observations, and the status of available models with their limitations. In Section 3 we introduce GANs and we describe FORSE

(Foreground Scale Extender): the algorithm developed in the context of this work. In Section 4 we show the results obtained when the proposed method is applied to thermal dust emission in both total intensity and polarization. Our conclusions together with the description of the numerous possible future applications of our approach are summarized in Section 5.

2. GALACTIC CONTAMINATION TO CMB OBSERVATIONS

In this Section we summarize the current knowledge of Galactic emission as contaminant to CMB observations and we describe the available Galactic models that are currently being used in preparation of the future CMB experiments.

2.1. Status of observations

The microwave sky is characterized by the presence of two main Galactic emissions that generate highly linearly polarized radiation.

At low frequencies ($\lesssim 70$ GHz) synchrotron emission from cosmic ray electrons accelerating around the Galactic magnetic field represents the dominant source of radiation. Synchrotron polarized emission has been observed on the full sky by both the WMAP and Planck satellites, with the highest signal-to-noise ratio (SNR) reached at frequencies of 23 and 28.4 GHz respectively. The Stokes Q and U maps obtained from these observations have allowed to constrain the emission with sufficient sensitivity at angular scales larger than few degrees ([Planck Collaboration XXV 2016](#)). Other datasets, with partial sky coverage but with higher SNR, have been obtained from observations at lower frequencies (S-PASS [Carretti et al. \(2019\)](#); [Krachmalnicoff et al. \(2018\)](#), CBASS [Jones et al. \(2018\)](#)). Thanks to these observations we know that synchrotron emission is highly linearly polarized, with polarization fraction that can reach about 20% at intermediate and high Galactic latitudes. In the first approximation the synchrotron SED (spectral energy distribution) follows a power law $A_s(\nu) \propto \nu^{\beta_s}$ with spectral index $\beta_s \approx -3$ that presents a non negligible spatial variation ([Jew & Grumitt 2020](#); [Krachmalnicoff et al. 2018](#); [Fuskeland et al. 2014](#)). At the power spectrum level, synchrotron amplitude can be approximated with a power law as a function of multipole with $A_s(\ell) \propto \ell^{\alpha_s}$ and $\alpha_s \approx -3$, showing therefore a pretty steep decay at small angular scales ([Krachmalnicoff et al. 2018](#)).

At higher frequencies ($\gtrsim 100$ GHz) the polarized thermal emission coming from dust grains aligned with the Galactic magnetic field represents the prevalent radiation in the microwave sky. Polarization maps of thermal

dust emission have been obtained from Planck observations on the full sky at the frequency of 353 GHz with an angular resolution of about 5 arc-minutes. These maps are signal dominated everywhere in the sky at angular scales larger than $\sim 1^\circ$, while finer structures are detected with high accuracy only at low Galactic latitudes. Dust radiation in Planck data is well fitted by a single modified black body SED with a spectral index $\beta_d \approx 1.6$ and a temperature $T_d \approx 20$ K (Planck Collaboration Int. XXX 2016; Planck Collaboration XI 2020). Both these spectral parameters show variations in the sky on the degree scale (Planck Collaboration X 2016; Planck Collaboration Int. XLVIII 2016). Planck observations have also shown that there exists an asymmetry in the polarized dust radiation, with E -mode power being about twice as higher than B -modes (Planck Collaboration Int. XXX 2016; Planck Collaboration Int. XXXVIII 2016). Similarly to synchrotron radiation, the thermal dust power spectrum can be approximated at first order with a power law $A_d(\ell) \propto \ell^{\alpha_d}$, with $\alpha_d \approx -2.4$. The dust signal also show high level of non-gaussianity with non-zero polarized bispectra (the equivalent of the three point correlation function in harmonic space) that have been observed in the Planck maps (Coulton & Spergel 2019).

Polarized synchrotron and thermal dust radiation have a high degree of spatial correlation, due to the same underlying Galactic magnetic field. In particular, the joint analysis of Planck and WMAP data allowed to measure the correlation coefficient ρ which shows a progressive decay as a function of multipole, with $\rho \approx 0.5$ at $\ell \approx 10$ and approaching values compatible with zero at the degree angular scales (Planck Collaboration XI 2020; Choi & Page 2015).

2.2. Available foreground models

Given the brightness of polarized Galactic emission and its complexity having reliable models of foregrounds is fundamental to be able to optimize and test the capability of component separation algorithm to retrieve clean CMB maps. In fact, in the recent years a mismodeling of thermal dust emission has already led to a false detection of primordial B -modes from the observations of the BICEP2 experiment (BICEP2 collaboration I 2014; Planck Collaboration Int. XXX 2016; BICEP2/Keck Array and Planck Collaborations 2015).

Current models are largely based on foreground templates obtained from the datasets described in the previous section. These models are included in dedicated packages that allow to simulate the microwave sky in the different frequency channels (e.g. the Python Sky

Model, PySM¹, Thorne et al. (2017); the Planck Sky Model, PSM, Delabrouille et al. (2013)). As it has been already stressed, these templates have a close match to the real sky at angular scales $\gtrsim 1^\circ$ but cannot be used to simulate smaller structures, since at those scales observations are largely contaminated by instrumental noise. This limitation is usually overcome by extrapolating foreground power spectra at higher multipoles considering a simple power law models. Gaussian realizations of these spectra are then generated and combined with the large scale templates in order to obtain foreground maps with the inclusion of small scale structures. Although this procedure can guarantee to have, at least at the first order, foreground simulations with the correct amplitude at all scales, it obviously represents quite a rough simplification from a statistical point of view. In fact, the higher order statistics of foregrounds must be taken into account to estimate its impact on both lensing reconstruction and on possible detection of primordial non-Gaussianity. New, more sophisticated Galactic models are therefore necessary in order to achieve this goal.

A second approach to model the Galactic emission is based on numerical magnetohydrodynamic (MHD) simulations. In this case realistic magnetized interstellar medium structure is simulated taking into account turbulence driven by star-formation and supernova feedbacks. Synthetic foreground maps can be generated through MHD simulations, and it has been shown that properties of real Galactic emission can be retrieved, as for example the asymmetry in the amplitude of polarized E -modes compared to the B ones observed in thermal dust radiation (see for example Kim et al. (2019) and references therein). However with this approach is difficult to simulate maps with large scale structures corresponding to the real foreground morphology and for this reason MHD simulations are currently little used in CMB studies.

Vansyngel et al. (2018) proposed a hybrid approach, combining magnetic field informations with the latest Planck observations. They aimed at characterizing the small scales of thermal dust polarization by relating them to the underlying structure of the Galactic magnetic field, modeled as a uniform field plus a random turbulent component. The parameters of the magnetic field model are constrained so that they could reproduce the dust EE, BB and TE power spectra observed

¹ The PySM package, its documentation and a description of the implemented sky models is available here <https://pysm3.readthedocs.io/en/latest/>

in Planck maps. Their simulated maps are produced at an angular resolution of 0.5° .

3. THE FORSE ALGORITHM

In this work, we propose a new approach to generate simulated maps of the polarized Galactic emission at sub-degree angular scales. The package that we developed, called FORSE (Foreground Scale Extender), is based on the use of Neural Networks (NNs) to generate realistic small scale structures in foreground maps².

NNs have started to be widely used in Astrophysics and Cosmology, with several applications also in the CMB field: from parameter estimation (Krachmalnicoff & Tomasi 2019) to lensing reconstruction (Caldeira et al. 2019), foreground characterization and component separation (Petroff et al. 2020; Farsian et al. 2020).

In this section we summarize how GANs, the family of NNs used in this project, work and we describe the architecture used for our application.

3.1. Generative Adversarial Networks (GANs)

The GAN approach was firstly proposed by Goodfellow et al. (2014) and since then it has become more and more popular in several contexts where a generative framework can be applied, e.g. in natural image processing as image inpainting, super-resolution, image de-noising. To date, there are several GANs proposed in the literature and the major differences can mostly be found in the network architecture implementations (e.g. see Deep-Convolutional GAN (DCGAN) Radford et al. 2016, Super-Resolution GAN Ledig et al. 2016, Contextual-Attention Yu et al. 2018)

The general idea of GAN framework relies on an *adversarial game* between two NNs: a *Generator*, G , and a *Discriminator*, D . Practically, G generates new data instances that must match the general feature of a given data sets (the training set); D evaluates the authenticity of the samples generated by G , distinguishing whether each sample belongs or not to the actual training set. D returns a probability value between 0 and 1, with 1 (0) labeling an authentic (mock) image. The goal of D is thus to label all the generated images as “mock”, whereas the goal of G is to improve the quality of the generated images in order to fool D .

The whole process of training a GAN therefore consists of iterating between two feedback loops: G being

² The word “forse” means “maybe” in Italian. With this name we emphasize the fact that the small scale structures generated by our algorithm are not expected to reproduce the true morphology of foregrounds but just to predict realistic features with the correct statistical properties.

in a feedback loop with D , and D in loop with the authentic set of images. Further details about our GAN implementation can be found in Section 3.2.

During the past few years, GAN have been adopted on a wide range of applications in cosmology. Kodi Ramanah et al. (2020) and Mustafa et al. (2017) used GANs in order to generate high-quality samples of respectively N-body simulations and weak lensing convergence maps. Aylor et al. (2019) generated thermal dust intensity emission maps with an effective non-Gaussian statistical model and trained a DCGAN architecture from the observations inferred by the *Planck* satellite. Puglisi & Bai (2020) showed that GANs can be efficiently applied to Galactic foreground unpolarized and polarized emission in order to generate emission at small scales and in-paint the maps on regions smaller than $\lesssim 0.5$ degrees.

3.2. The implemented GAN

The problem that we want to address can be summarized as follows. Let M be a map encoding foreground emission at a given angular resolution. M can be expressed as the sum of two maps encoding large and small scale structures, respectively M_{LS} and M_{SS} :

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

We can also make the assumption that the small scale features are actually modulated by the large ones, with $M_{SS} = M_{LS} \cdot m_{ss}$. In this way we have:

$$\tilde{m}_{ss} \equiv m_{ss} + 1 = \frac{M}{M_{LS}}. \quad (2)$$

The goal of this work is to use GANs to generate a map of realistic small scale structures (\tilde{m}_{ss}) given the real large scale ones.

3.2.1. Architecture

The FORSE package is based on a modified version of the DCGAN described in Radford et al. (2016). In its original version the *Generator* (G) of a DCGAN takes as input a vector of random numbers and, by reshaping it and applying a series of convolutional layers, generates a mock image. However, our application is different, as both the input and the output of G are images.

The GAN architecture implemented in this work is sketched in Figure 1. G takes as input images of 320×320 pixels which are maps (patches) of the foreground sky where only the large scale features are present (M_{LS}). A series of three convolutional layers is then applied with the dimension of the kernel being 5×5 . In the first layer 64 filters are used, and no stride is applied, leading to an image cube with dimension of

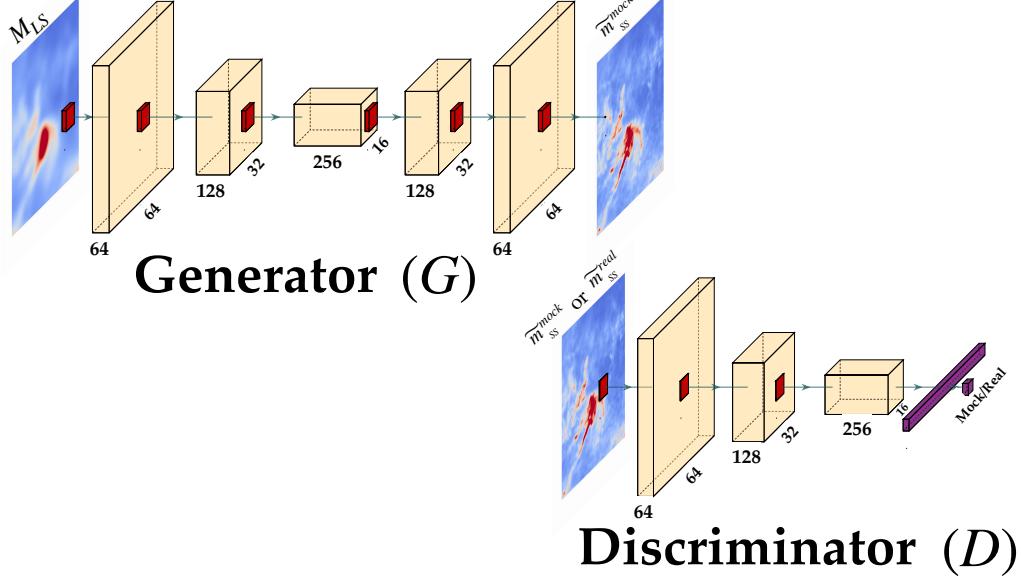


Figure 1. Sketch of the GAN architecture developed for this work. The *Generator* (G) takes as input images of 320×320 pixels, which represent patches of the sky where only large scale structures are present (M_{LS}). The *Discriminator* (D) network takes as input real or mock small scale structures and classify them. The red squares represent the convolutional kernels (with dimension 5×5). A detailed description of the different layers is given in Section 3.2.1.

$320 \times 320 \times 64$ pixels. In the following convolutional layers a stride of 2 is used and the number of filters is doubled, generating therefore cubes of $160 \times 160 \times 128$ and $80 \times 80 \times 256$ pixels after the second and third convolution, respectively. After each convolution, a *LeakyRelu* activation function is applied with slope equal to 0.2, and a *BatchNormalization* layer is added.

The decoding part of G is symmetric to the encoding one; upsampling layers are combined with convolutional ones in order to restore the output dimension of the image with 320×320 pixels (see Figure 1). The output layer is activated with a *tanh* function.

The *Discriminator* (D), also shown in Figure 1, takes as input \tilde{m}_{ss} images of 320×320 pixels. After three convolutional layers (which are analogous to the encoding part of G) the resulting cube is flattened into a 1-D vector which is then densely connected to the output unit, activated though a *sigmoid* function.

The GAN architecture and the training procedure (described in the following section) has been developed using the *Keras*³ python library with a *tensorflow* backend.

3.2.2. Training procedure

In all our applications we trained our GAN by minimizing a *binary cross-entropy* loss function applying a

³ See <https://keras.io/api/> also for a thorough description of the layers and functions used in our GAN.

stochastic gradient descent. As suggested in Radford et al. (2016), we used the *Adam* optimizer, with a learning rate of 0.0002 and a momentum term $\beta_1 = 0.5$.

The training was done considering mini-batches of $N_b = 16$ images with the following two steps:

1. N_b patches of the real large scale structure (M_{LS}) were fed to G which generates the corresponding \tilde{m}_{ss}^{mock} and passes them to D . In this phase G was trained by maximizing the probability that \tilde{m}_{ss}^{mock} are classified as real images (class 1) by D . The back-propagation is done trough the whole GAN ($G+D$) but only the weights of G are optimized.
2. In the second step $N_b/2$ \tilde{m}_{ss}^{mock} and $N_b/2$ \tilde{m}_{ss}^{real} images were passed to D which was trained to classify them as mock (class 0) and real (1), respectively.

As a pre-step, all the M_{LS} and the \tilde{m}_{ss}^{real} patches were normalized, in order to have pixel values in the range $[-1, 1]$.

4. APPLICATION TO THERMAL DUST EMISSION

As described in Section 2, thermal dust emission represents one of the strongest contaminant to CMB observations at frequency around 150 GHz, and shows high level of non-Gaussianity in both total intensity and polarization. We therefore choose to apply our algorithm to thermal dust maps, with the goal of obtaining sim-

ulations with realistic and non-trivial small scale structures at sub-degree angular resolution. In this Section, we describe the methodology used as well as the results obtained.

Existing templates of thermal dust emission on the full sky have been obtained by applying component separation algorithms to Planck maps. In this work, we used the ones obtained from the GNILC (Generalized Needlet Internal Linear Combination) method at 353 GHz, which have the advantage of not being contaminated by the CIB radiation in total intensity (for a detailed description of the data and the method applied to get this thermal dust template we refer to [Planck Collaboration Int. XLVIII 2016](#)). The GNILC dust maps have an angular resolution that varies in the sky and depends on the SNR of the Planck high frequency maps: in total intensity the effective beam FWHM (full width half maximum) ranges from 5 to about 22 arc-minutes, while in polarization it varies in the interval 5–80 arc-minutes ([Planck Collaboration IV 2020](#))⁴. Both polarization and intensity maps follow the HEALPix pixelization scheme ([Górski et al. 2005](#)) at $N_{side} = 2048$, i.e. ~ 1.7 arc-minute pixel resolution.

4.1. Total intensity

We firstly tested the ability of FORSE of generating realistic small scale features in total intensity (quantified by the Stokes I parameter). To do so, we needed a set of patches for which the real small scale structures have been observed. As stated above, the GNILC template in total intensity has a variable angular resolution, which is equal to 5 arc-minutes in the regions close to the Galactic plane thus making them suitable to be used to train the GAN. In order to further select only those regions less contaminated by noise, we built a mask from the Planck HFI 353 GHz I map. In particular, we took into account only that part of the sky where the SNR in the 353 GHz map at the full angular resolution (about 5 arc-minutes) is above a threshold of 8. We smoothed the obtained mask with a Gaussian beam with FWHM of 1° , in order to regularize its borders. We also cut out the inner part of the Galactic plane (with $b < 10^\circ$) which, due to its peculiarity, is not suitable to be used as part of the training set. The resulting mask is shown in Figure 2 and covers a fraction of the sky of about 23%.

In order to train the GAN, both M_{LS} and \tilde{m}_{ss}^{real} images were needed. The M_{LS} patches were generated as tiles of $20^\circ \times 20^\circ$ and 320×320 pixels from the GNILC

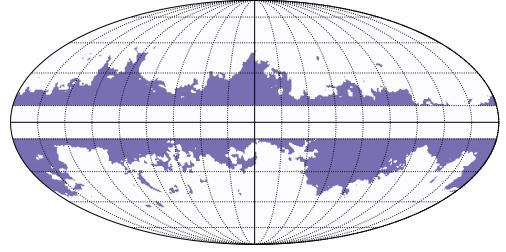


Figure 2. Sky mask adopted in this work (see Sect.4.1 for details). Images used to perform the GAN training have been selected within the purple region.

I template, smoothed at the angular resolution of 80 arc-minutes. On the other hand, \tilde{m}_{ss}^{real} patches were obtained from the same map smoothed at 12 arc-minutes angular resolution and divided by M_{LS} (see definition in Eq. 2). All the tiles were taken from the region of the sky within the mask in Figure 2, for a total of 350 pair of images. We trained our GAN on NVIDIA V100 GPUs, with the procedure outlined in Section 3.2.2, for about 10^5 iterations, saving both the G and D weights every 1000 steps. As described below, we obtained the best performance after 37000 iterations, and therefore, in the following, we report the results corresponding to this GAN configuration.

Figure 3 reports the input and output of G after training, as well as the comparison with the real small scale features, for three different sky location. Images are normalized in the range $[-1, 1]$. These results show the capability of FORSE of generating small scale features with a non-trivial statistics, starting only from the large scale ones. We stress again that, with this approach, we do not expect to generate small scale structures that reproduce the morphology of the real ones, but to retrieve the correct statistics. It is however remarkable how the network is able to reconstruct the main small scale features (e.g. peaks and troughs), implying that during training it has learned the underlying relation between large and small scales.

In Figure 4 we show a comparison of the generated small scale structures with the real ones, from a statistical point of view. In order to do so we made use of the Minkowski functionals \mathcal{V}_0 , \mathcal{V}_1 , \mathcal{V}_2 , as defined in [Mantz et al. \(2008\)](#), which are sensitive to the presence of non-Gaussian structures. The three functionals characterize the area, the perimeter and the connectivity of the features respectively. We computed \mathcal{V}_0 , \mathcal{V}_1 , \mathcal{V}_2 as a function of the threshold ρ in the range $[-1, 1]$, for all the 350 patches and for \tilde{m}_{ss}^{real} and \tilde{m}_{ss}^{mock} ; in Figure 4, for each functional, we plotted the mean and 1σ deviation. The two distributions present a remarkable agreement,

⁴ The GNILC maps are available on the Planck Legacy Archive website <http://pla.esac.esa.int/pla/#maps>

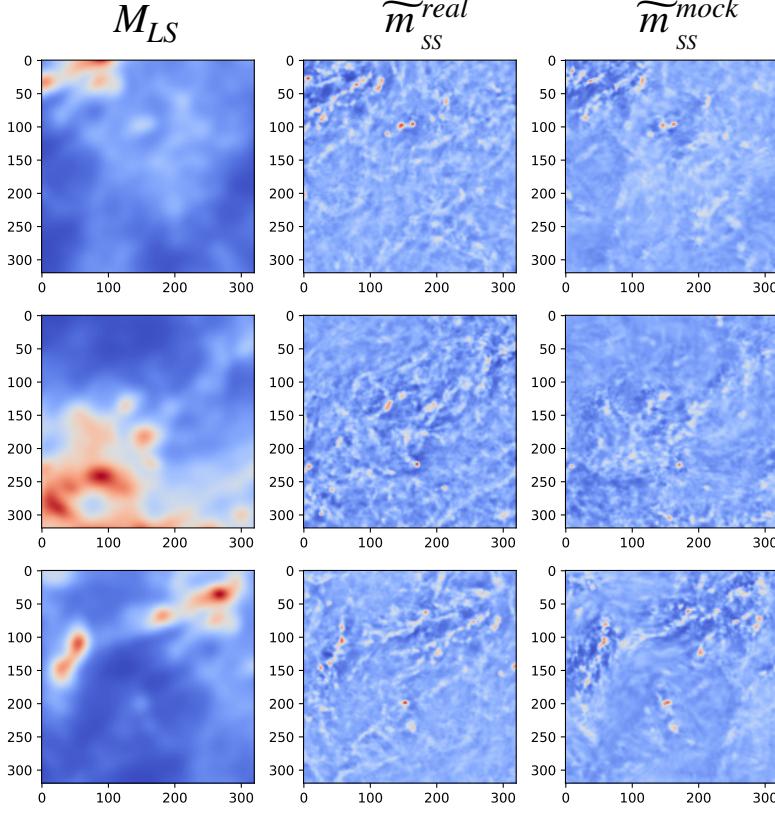


Figure 3. Input (left column) and output (right column) of FORSE when applied to total intensity dust map for three different sky regions. The middle column shows the real small scale features (12 arc-minutes) present in the sky. Maps are rescaled to unphysical units ranging in the same interval $[-1, 1]$.

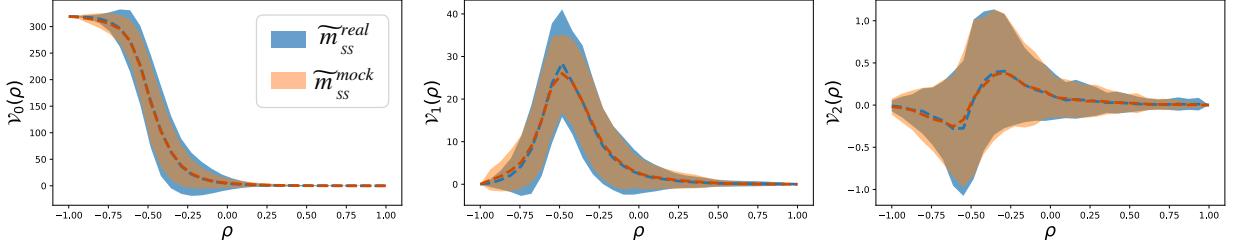


Figure 4. Minkowski functionals (\mathcal{V}_0 , \mathcal{V}_1 , \mathcal{V}_2) as a function of the threshold ρ , for \tilde{m}_{ss}^{real} (blue) \tilde{m}_{ss}^{mock} (orange) for the case of total intensity dust maps. The functionals are computed for the 350 patches used to train the GAN and we report the mean (dashed lines) and 1σ deviation (shaded areas) of the distributions. The high level of superposition between the curves demonstrates the ability of FORSE to reproduce the correct statistical properties of small scale foregrounds.

with superposition at the level of 76% (\mathcal{V}_0), 84% (\mathcal{V}_1) and 91% (\mathcal{V}_2). We have computed these values for all the different GAN configurations saved during training, and chose the one that led to the best agreement (reached after 37000 iterations). This results clearly show how the approach developed with FORSE allows to generate

small scale feature on foreground maps that match the statistical properties of the real ones.

Once the small scale features are generated by the GAN they needed to be normalized back in order to restore physical units. In this case, where we tested the feasibility of the approach in total intensity, the normal-

ization to physical units is trivial, as, for each considered patch, we know the amplitude of the real small scales structures and we could therefore use this information. In practice, we rescaled each \tilde{m}_{ss}^{mock} patch in order to match its mean and standard deviation with the ones of the corresponding \tilde{m}_{ss}^{real} . Once physical units were restored, we combined the large and small scale patches and got the final image as $M = M_{LS} \cdot \tilde{m}_{ss}$ (see relation 2).

Figure 5 shows our final results. In the left panel, we report the large scale and combined images (large and small scales) for the same three patches shown in Figure 3, as an example. We report also their angular power spectra (right panel) computed with the NaMaster code (Alonso et al. 2019). The comparison of the combined M^{mock} and M^{real} images and spectra not only shows that FORSE generates small scale structures with realistic morphology and statistical properties, but also that the correct amplitude scaling as a function of the angular scale is recovered after the normalization to physical units.

4.2. Polarization

After having tested our approach in total intensity, we could apply it to polarized emission. Although being more challenging, this represents the most interesting application of FORSE, as current CMB experiments are focusing on the observations of the polarized cosmological signal (see Section 1) and we suffer from a significant lack of foreground data.

The GNILC polarized thermal dust template has an angular resolution below 12 arc-minutes (the target resolution of FORSE) only in about 9% of the sky and mainly in the inner Galactic plane region ($|b| < 10^\circ$). Given this, we could not apply the procedure used in total intensity for Stokes Q and U maps, as we did not have enough data to perform training. To overcome this limitation, we made the assumption that small scale structures in polarization follow the same statistics as the ones in total intensity. This represents a reasonable assumption to the first order; in fact, we can presume that the dust grains population producing the polarized emission is the same as the one responsible in emitting the unpolarized signal. Additionally, we already know that thermal dust radiation has similar two point correlation functions (power spectra) in polarization and total intensity (Planck Collaboration XI 2020).

With this assumption, we proceeded by training separately two different GANs (both with the architecture described in Section 3.2), for Stokes Q and U patches respectively. In particular, we used as input to the GANs, images of the large scale structures $M_{LS}^{Q(U)}$ taken from

the GNILC Stokes Q (U) map, at the angular resolution of 80 arc-minutes. The *Generator* G was trained to generate small scale structures ($\tilde{m}_{ss}^{mock,Q(U)}$) at the angular resolution of 12 arc-minutes, starting from the large scale ones. These were then compared by the *Discriminator* D with the real small scale structures in total intensity ($\tilde{m}_{ss}^{real,I}$) and the GAN weights were thus optimized in order to make $\tilde{m}_{ss}^{mock,Q(U)}$ indistinguishable from $\tilde{m}_{ss}^{real,I}$. We applied the same training procedure used for total intensity (see Section 3.2.2).

Since our goal was to obtain Q and U full sky maps, we generated the $M_{LS}^{Q(U)}$ tiles in such a way that they cover the full celestial sphere, accounting for a total of 174 images for each of the two Stokes parameter (the description of the procedure adopted to make the projections from the celestial sphere to the flat patches and back is reported in Appendix A). In order to get the best possible result, we optimized the GAN's weights directly on these images (after having normalized each of them in the range $[-1, 1]$) and compared the G output with the $\tilde{m}_{ss}^{real,I}$ from the set of images used in the total intensity case. As before, and for both Q and U optimization, we trained the GAN for about 10^5 iterations, saving weights every 1000 steps. We reached the best performance (in terms of Minkowski functionals, see Figure 7) after 83000 and 88000 iterations, for Stokes Q and U maps respectively.

Figure 6 shows the results after training, for one patch and for both Stokes Q and U maps. The small scale structures generated by the GAN are compared with the ones from the Gaussian extrapolation of the dust power maps, obtained as follows:

1. polarization (E and B -modes) power spectra were computed from the GNILC Q and U full sky maps at the angular resolution of 80 arc-minutes and fitted with a power law in the range of multipoles $20 - 100$;
2. the fitted spectra were multiplied by a function $f_\ell = 1 - B_\ell^2$ (where B_ℓ represents the beam window function with FWHM = $80'$) in order to remove the large angular scales. A pair of Q and U maps was obtained as a Gaussian realization of them;
3. the obtained maps were multiplied by the GNILC Q and U maps at 80 arc-minutes in order to modulate the small scale structures, and then multiplied by a factor that restore the correct amplitude of the power spectra;
4. lastly the obtained small scale structure maps are summed to the GNILC large scale ones.

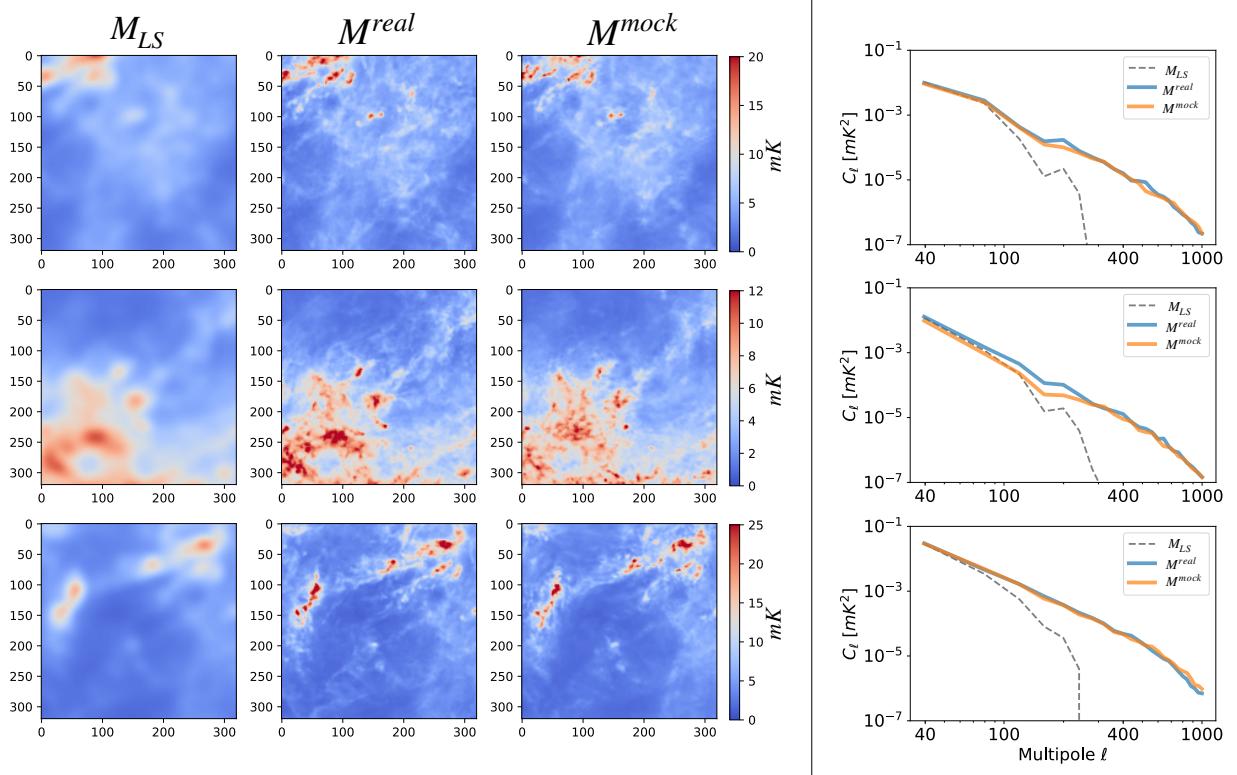


Figure 5. Left panel: large scale (M_{LS}) and combined ($M^{real/mock}$) images for the three different patches shown in Figure 3 as obtained after restoration of physical units. Right panel: angular power spectra of the images shown on the left.

This procedure generates maps which satisfy relation 2, with Gaussian small scale structures \tilde{m}_{ss}^{gauss} , and whose power spectra follow the power law extrapolation of the real large scale features. We stress that a similar procedure is also implemented in the publicly available foreground simulation packages (see Section 2.2).

Figure 7 shows the Minkowski functionals of the GAN-generated small scale structures for Q and U maps. The functionals are compared with the ones of the real structures in total intensity (which represent our target distribution), as well as with the ones computed from the Gaussian small scale features. The distributions of the mock structures are in good agreement with the real total intensity ones, with a superposition of the three functionals (\mathcal{V}_0 , \mathcal{V}_1 , \mathcal{V}_2) at the level of (79%, 85%, 89%) and (87%, 84%, 87%) for Q and U maps respectively. On the contrary the distributions are very different from the Gaussian ones, showing once again how FORSE is able to generate non-trivial structures on maps.

As for the total intensity case, we needed to normalize back the $\tilde{m}_{ss}^{mock,Q(U)}$ images to physical units. Since in polarization we do not have the information on the amplitude of the real small scale structure, we performed

the rescaling by means of Gaussian extrapolations described above. Although these simulations present different statistical properties, they predict, at a first order, the correct amplitude of the small scale structures (i.e. the extrapolation is specifically built in such a way that the spectra have the correct scaling). In practice, the $\tilde{m}_{ss}^{mock,Q(U)}$ maps were rescaled in order to have the same mean and standard deviation as the corresponding Gaussian small scales structures in physical units. A second multiplication factor was applied in order to minimize the distance between the power spectra of the $\tilde{m}_{ss}^{mock,Q(U)}$ and $\tilde{m}_{ss}^{gauss,Q(U)}$ at $\ell > 160$.

Figure 8 shows the final results, with the combination of real large scales and mock small ones for Stokes Q and U and polarized intensity $P = \sqrt{Q^2 + U^2}$, together with the E and B -modes power spectra. As it can be seen the GAN is able to produce non-Gaussian structures that after the normalization in physical units have the correct amplitude and scaling. However we notice that, given the fact that we trained the network to reproduce the statistics of total intensity features, it was unable to retrieve the thermal dust asymmetry of E and B modes. This limitation might be overcome by adopt-

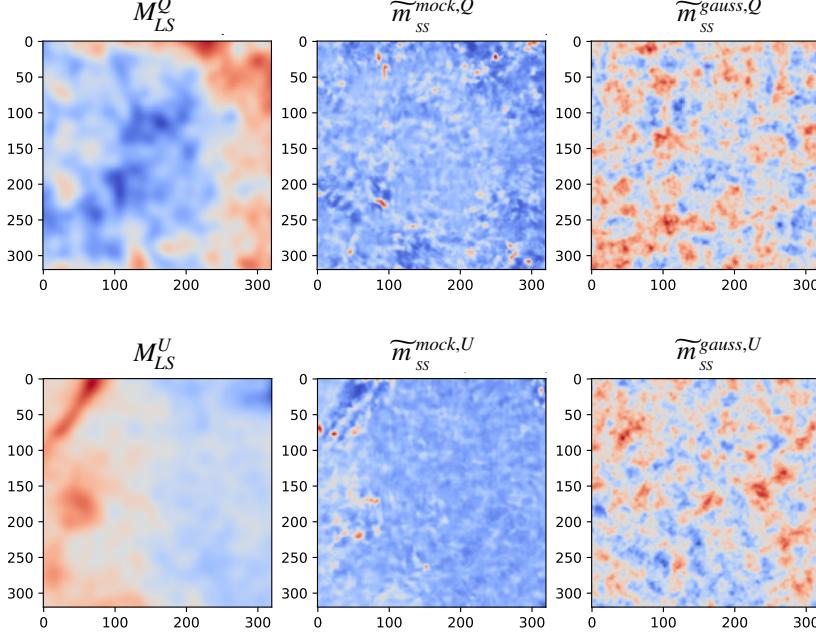


Figure 6. Input (left column) and output (middle column) of FORSE when applied to Stokes Q and U maps (upper and lower rows respectively). The small scale structures generated by the GAN are compared with Gaussian ones obtained as described in Section 4.2. Maps are rescaled to unphysical units ranging in the interval $[-1, 1]$.

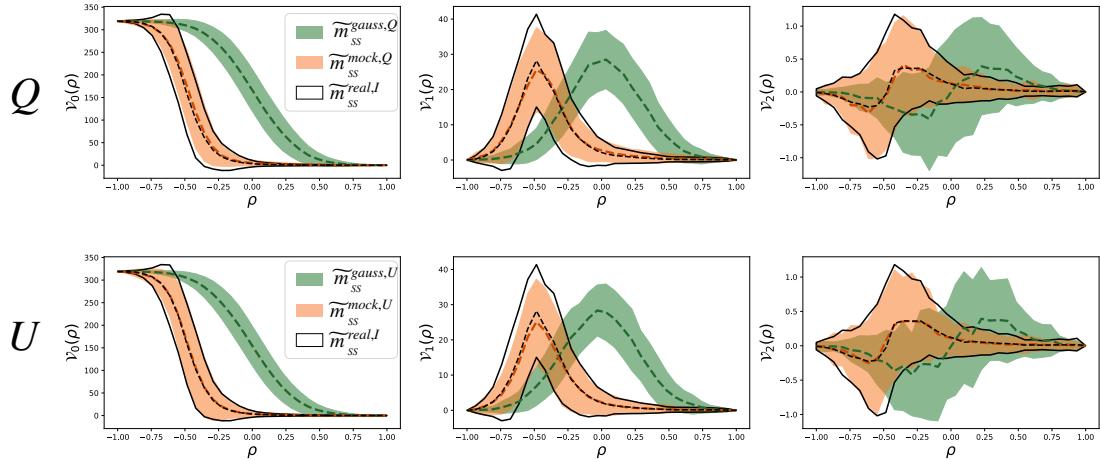


Figure 7. Minkowski functionals for real small scales structures in total intensity (black lines), features generated by the GAN (orange) and Gaussian ones (green). The functionals are computed for the 174 patches used to train the GAN and we report the mean (dashed lines) and 1σ deviation (shaded areas) of the distributions.

ing two different training sets obtained from polarized intensity and polarization angle patches, we outlook to explore this in the future, once high resolution polarization data will become publicly available.

4.2.1. Full sky maps

Since we wanted to obtain polarization full sky maps of dust emission with the inclusion of GAN generated

small scale structures, we needed to reproject the obtained square patches into a full sky HEALPix map.

The projection into the spherical domain is complicated by the fact that there is no large scale continuity in two consecutive patches between the high-resolution predictions performed by FORSE. This is due to the fact that when the neural network generates the small scale features for a given patch, it is essentially *blind* to

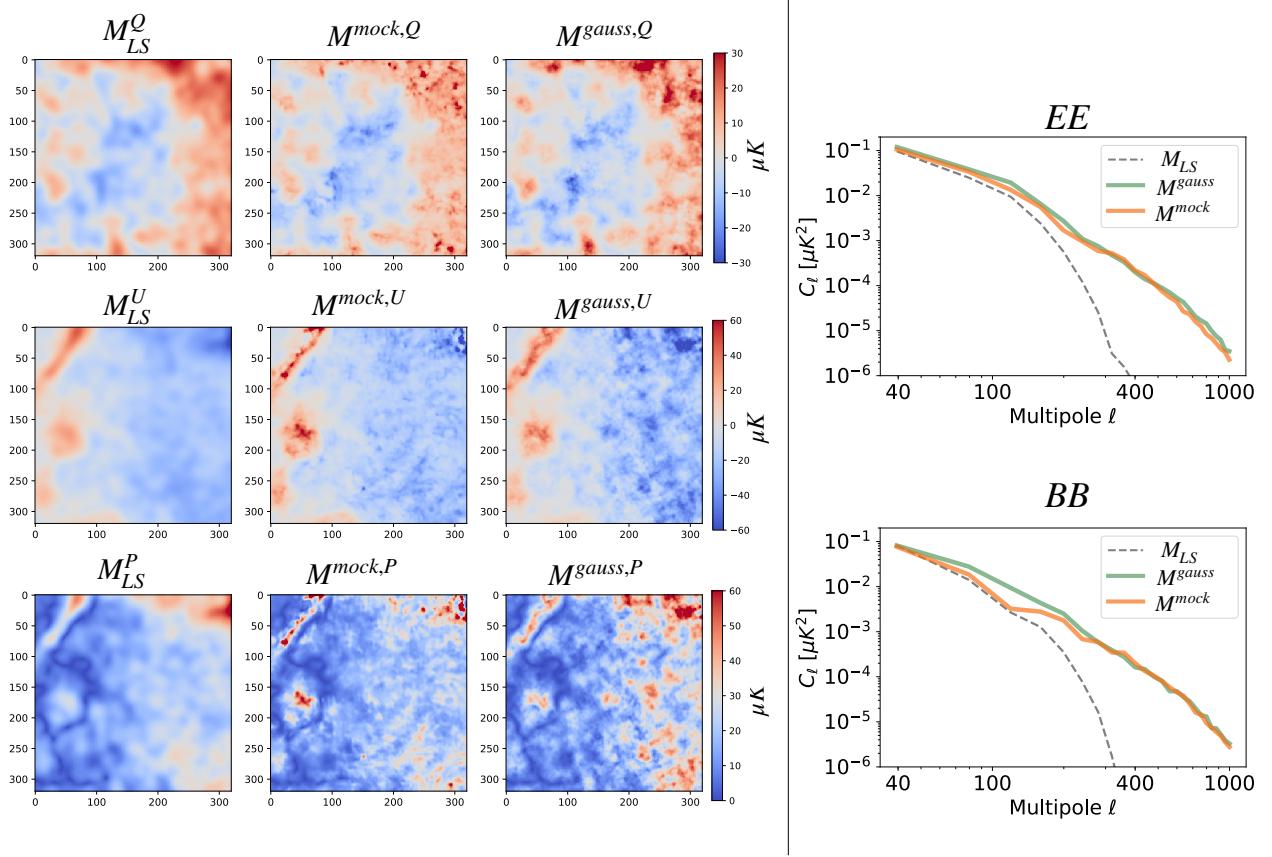


Figure 8. Left panel: large scales (M_{LS} , right column) and combined images with small scales features obtained from the GAN (M^{mock} , middle column) or as Gaussian realizations (M^{gauss} , right column), for Q (upper row), U (middle row) and P (lower row). The sky region is the same as the one shown in Figure 6. Right panel: comparison of the E (upper plot) and B -mode (lower plot) power spectra.

the structures encoded at the edges of its nearest neighbors, and, as a result, we observe a feature mismatch at the borders. In order to reduce the border effects, we (i) applied a cosine apodized 2D window function (see right panel of Fig. 12) to the square patches before performing the full-sky projection and (ii) we averaged the patches in the overlapping area with weights given by the apodization window. The details on how the projection is performed are outlined in Appendix A.

In Figure 9 we show the final reprojected HEALPix Q and U maps at angular resolution of 12 arc-minutes (with $N_{side} = 2048$) compared with the input GNILC ones at 80 arc-minutes. We notice that the injection of small scales do not globally change the large scale features and, by looking at the difference between maps (third row in Figure 9), we see that no border effects are visible. However, the application of an apodization window function and the averaging in the reprojection procedure could causes an extra-smoothing of features especially where the patch overlap is large (i.e. close to the Galactic poles). To quantify this effect we proceeded

as follows: we resampled back the obtained full sky maps into square patches and computed the ratio (r_ℓ^p) between the polarization power spectra of each patch before and after the reprojection. Any smoothing effect should therefore be seen as a ratio $r_\ell^p > 1$. Since the effect depends on the overlap of adjacent patches, which is a function of the latitude at which the tiles are centered (see left panel of Figure 12), we averaged r_ℓ^p over all the patches located at the absolute Galactic latitude. Results are reported in Figure 10 for the ratio between EE spectra (similar results are obtained for BB). As it can be seen the loss of power is limited, below a factor 2, at all multipoles and for all the patches centered at Galactic latitudes $|b_c| \leq 72^\circ$ while, as expected, is more critical at the Galactic poles.

Finally, Figure 11 shows the polarization power spectra computed from the final full sky Q and U maps compared with the ones of the input large scale maps. As expected the NN overall effect into the EE and BB power spectra results into a further extrapolation of the spectra up to $\ell \sim 1000$, where we can notice the drop in power

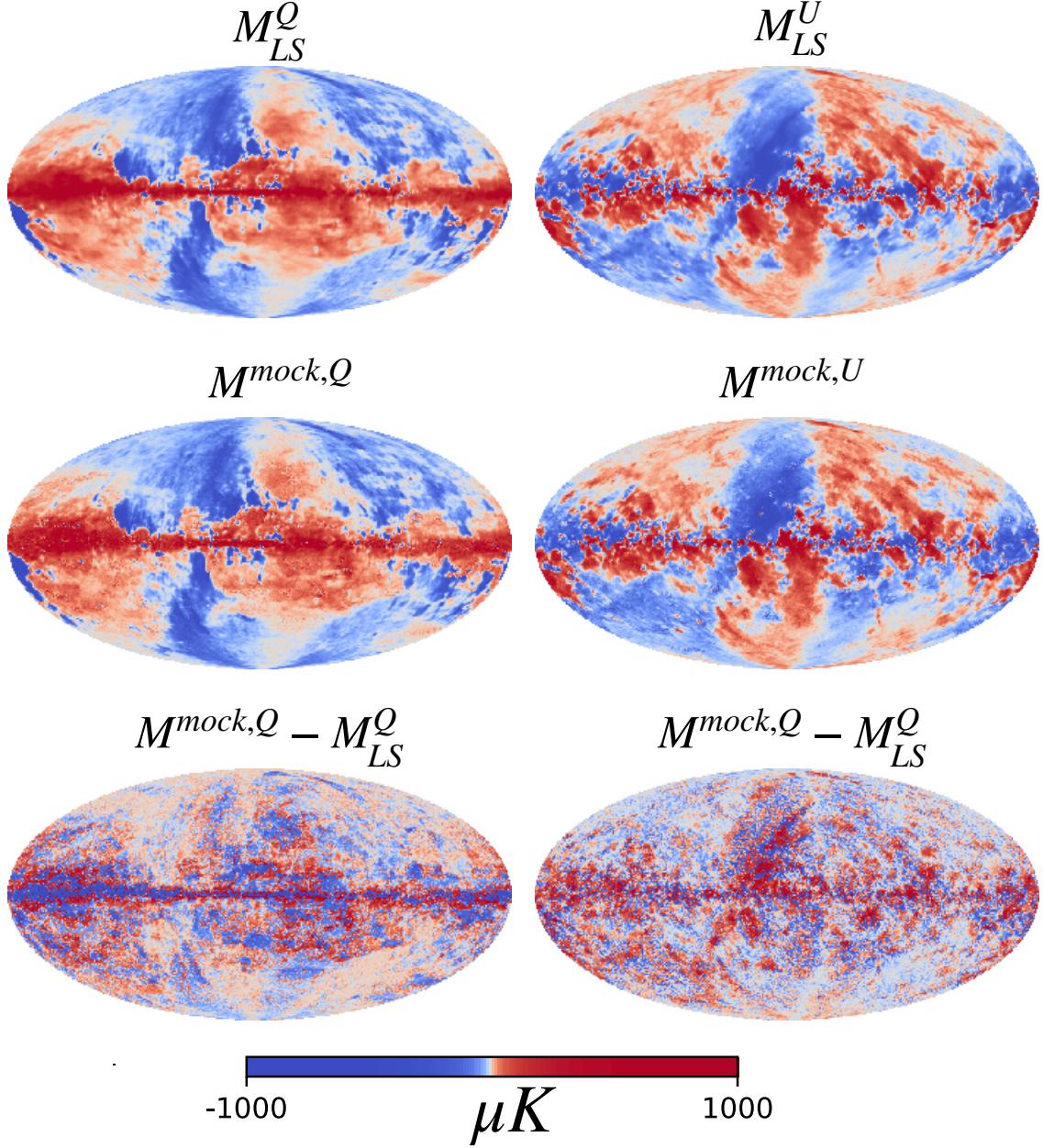


Figure 9. Top row: full sky polarization maps (left: Stokes Q , right: Stokes U) for the GNILC template at 80 arc-minutes angular resolution. These maps are the input to our algorithm. Middle row: maps with small scale features, up to 12 arc-minutes, generated by FORSE . Bottom row: difference between the two maps. Notice the residuals mostly encode smaller angular scales as expected.

due to the resolution scale. No additional loss of power due to the reprojection is visible at the full sky level. We further compared the power spectra of the input low resolution Q and U maps with the ones estimated from the GAN maps by estimating the index of the power law as $C_\ell \propto \ell^{-\alpha}$ and we performed the fit on a different multipole range, i.e. $\ell \leq 100$ for the former and $\ell \leq 800$ for the latter. The spectral indices estimates from the large scale maps are $\alpha_{EE} = 2.48$ and $\alpha_{BB} = 2.46$, whereas for the NN maps we get 2.54 and 2.38 respectively for

EE and BB; indicating that the NN does not induce any pivoting scale in the polarization power spectra at sub-degree angular scales.

However, The E/B asymmetry holds up only to the scales where the dust polarization has been probed by latest Planck measurements. As already stressed, this is somewhat expected since no polarization small scale data, that encode this characteristic, are given as training features to the GAN. Once high resolution data will be made publicly available from the current and future

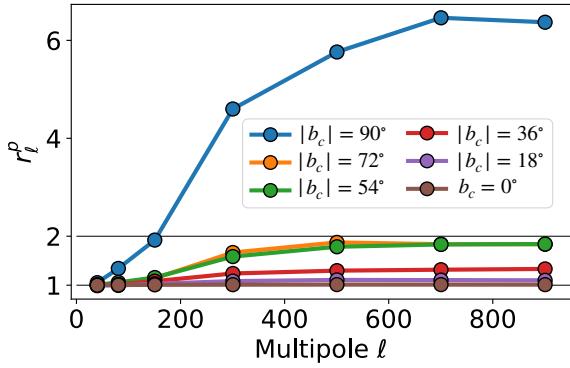


Figure 10. Ratio between E-mode power spectra before and after reprojection averaged over patches located at the same absolute Galactic latitude. Smoothing effects due to the full sky reprojection appear as $r_\ell^p > 1$.

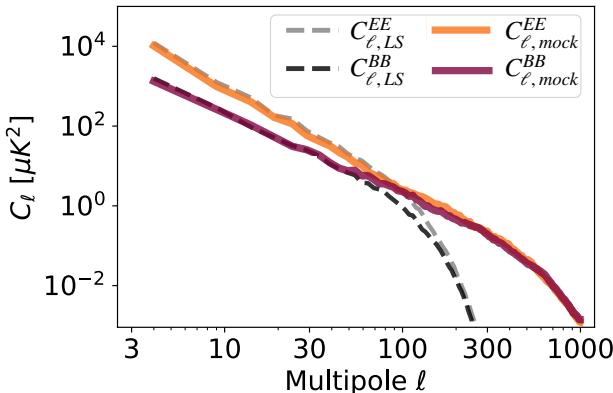


Figure 11. Polarization power spectra of the low resolution GNILC maps at 80 arc-minutes (dashed lines) and the final maps obtained with FORSE at 12 arc-minutes (solid lines).

high frequency experiments, even in partial regions of the sky, we plan to update the training weights with those data sets to improve fidelity in the generation of dust polarization features.

5. DISCUSSION AND CONCLUSIONS

In this work we have presented a novel approach, based on the use of GANs, to simulate realistic non-Gaussian sub-degree foreground emission. This represents an important task, as in current foreground simulations small scale structures are usually injected as Gaussian random realizations of the expected power spectra. This simplification of the statistical properties of foregrounds makes it difficult to estimate their impact, especially on gravitational lensing reconstruction and de-lensing. Having the possibility to rely on realistic foreground simulations at all the angular scales would thus be extremely important in the preparation of the next generation of CMB experiments. To our knowledge, this

is the first time that GAN methodology is used to generate polarization maps of astrophysical emission and to extend it to smaller angular scales.

We applied our algorithm, named FORSE, to thermal dust radiation considering both total intensity and polarized signal. In all the cases, we trained the GAN to generate small scale structures, at 12 arc-minutes, starting from large ones, at 80 arc-minutes. We applied the method to square sky patches with physical dimension of $20^\circ \times 20^\circ$ and resolution of 320×320 pixels.

In our first test in total intensity, we could make use of available observations from the Planck satellite (at 353 GHz) of the small scale structure of thermal dust in regions close to the Galactic plane. We therefore trained the network to produce sub-degree features that mimic the statistics of the real ones. Results (shown in Figures 3, 4 and 5) demonstrate how FORSE is able to achieve this goal. In particular, we characterized the obtained structures by mean of the three Minkowski functionals (\mathcal{V}_0 , \mathcal{V}_1 , \mathcal{V}_2) which are sensitive to the presence of non-Gaussianity. The agreement between the distributions computed from real and mock structures is remarkably good (see Figure 4). We also computed the angular power spectra of the final images, showing how the amplitude of the generated features follow the correct scaling as a function of the angular scale.

We applied our approach also to polarized emission. Since in polarization there are not publicly available high resolution data, we made the assumption that small scale features on Stokes Q and U maps follow the same statistics as the one in total intensity. We trained the networks accordingly, having as target distribution the one of small scale I structures. The results for polarization maps are reported in Figures 6, 7 and 8. Again, we find a very good agreement between the statistical properties of the generated structures on Q and U maps and the target total intensity ones, with distributions of the Minkowski functionals significantly different from those of a Gaussian field. Moreover, polarization power spectra show the correct scaling as a function of multipoles.

We reprojected the flat patches into the Celestial sphere, obtaining Stokes Q and U full sky maps of the thermal dust emission at sub-degree angular resolution which are shown in Figure 9. Maps are made publicly available online⁵, they are stored as a fits file and we follow the same conventions as in the GNILC dust map (i.e. μK_{CMB} , Galactic coordinates, same pixelization).

⁵ Maps can be downloaded from <https://portal.nersc.gov/project/sobs/users/ForSE/>

The reprojection on the sphere causes an extra smoothing of the generated feature, which is particularly important in the regions close to the Galactic poles (Fig. 10). Furthermore, due the fact that our GAN could not been trained with polarization data, the EE and BB power spectra of full sky maps reproduce the scaling observed at large angular scales but they do not preserve the E/B asymmetry that have been observed at low multipole at smaller angular scales.

As in this paper we presented the method and its first application to thermal dust emission, we plan to study the impact of the predicted non-Gaussian foreground on lensing reconstruction, de-lensing and primordial B -mode detection in an upcoming work. We also foresee further developments of the approach presented in this work, with several possible applications. In particular, it could be used to further extend the angular resolution of foreground models at even smaller scales. This could be done in different ways. One is obviously by making use of new data that will become publicly available in the coming years, which will include also the observations of small scale structures in polarization, allowing to overcome the assumptions we have made in our work. The second way would be to train and apply the GAN predictions iteratively, in a sort of “fractal” approach, by considering the foreground statistical properties as scale invariant. One could further think of training a

NN to learn and reproduce the relation between large and small structures as well as their statistical properties obtained from MHD simulations and then apply it to low-resolution data. Finally, although we have applied FORSE specifically to thermal dust emission, as one of the strongest contaminant to CMB observations, nothing prevents to adopt it on other kind of emissions, as long as a sufficient amount of data exists.

ACKNOWLEDGMENTS

NK acknowledges support from the ASI-Cosmos and ASI-LiteBIRD programs, from the INDARK and LiteBIRD INFN initiatives and from the “Leonardo da Vinci” grant (MIUR-2019). NK also thanks the Simons Foundation for hosting her at the Flatiron Institute (New York, USA) in fall 2019, and the whole group of Cosmology x Data Science at CCA (Center for Computational Astrophysics) for the extremely useful discussions at the beginning of this work. The authors thank Prof. Carlo Baccigalupi for valuable comments. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

REFERENCES

- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv e-prints, arXiv:1610.02743.
<https://arxiv.org/abs/1610.02743>
- Aiola, S., Calabrese, E., Maurin, L., et al. 2020, arXiv e-prints, arXiv:2007.07288.
<https://arxiv.org/abs/2007.07288>
- Alonso, D., Sanchez, J., Slosar, A., & LSST Dark Energy Science Collaboration. 2019, MNRAS, 484, 4127, doi: [10.1093/mnras/stz093](https://doi.org/10.1093/mnras/stz093)
- Aylor, K., Haq, M., Knox, L., Hezaveh, Y., & Perreault-Levasseur, L. 2019.
<https://arxiv.org/abs/1909.06467>
- Baccigalupi, C. 2003, NewAR, 47, 1127, doi: [10.1016/j.newar.2003.09.038](https://doi.org/10.1016/j.newar.2003.09.038)
- Baumann, D. 2009, arXiv e-prints, arXiv:0907.5424.
<https://arxiv.org/abs/0907.5424>
- Beck, D., Errard, J., & Stompor, R. 2020, JCAP, 2020, 030, doi: [10.1088/1475-7516/2020/06/030](https://doi.org/10.1088/1475-7516/2020/06/030)
- BICEP2 Collaboration, Keck Array Collaboration, Ade, P. A. R., et al. 2018, PhRvL, 121, 221301, doi: [10.1103/PhysRevLett.121.221301](https://doi.org/10.1103/PhysRevLett.121.221301)
- BICEP2 collaboration I. 2014, Physical Review Letters, 112, 241101, doi: [10.1103/PhysRevLett.112.241101](https://doi.org/10.1103/PhysRevLett.112.241101)
- BICEP2/Keck Array and Planck Collaborations. 2015, Physical Review Letters, 114, 101301, doi: [10.1103/PhysRevLett.114.101301](https://doi.org/10.1103/PhysRevLett.114.101301)
- Caldeira, J., Wu, W. L. K., Nord, B., et al. 2019, Astronomy and Computing, 28, 100307, doi: [10.1016/j.ascom.2019.100307](https://doi.org/10.1016/j.ascom.2019.100307)
- Carretti, E., Havercorn, M., Staveley-Smith, L., et al. 2019, MNRAS, 489, 2330, doi: [10.1093/mnras/stz806](https://doi.org/10.1093/mnras/stz806)
- Choi, S. K., & Page, L. A. 2015, arXiv: 1509.05934
- Coulton, W. R., & Spergel, D. N. 2019, JCAP, 2019, 056, doi: [10.1088/1475-7516/2019/10/056](https://doi.org/10.1088/1475-7516/2019/10/056)
- Delabrouille, J., Betoule, M., Melin, J. B., et al. 2013, A&A, 553, A96, doi: [10.1051/0004-6361/201220019](https://doi.org/10.1051/0004-6361/201220019)
- Farsian, F., Krachmalnicoff, N., & Baccigalupi, C. 2020, JCAP, 2020, 017, doi: [10.1088/1475-7516/2020/07/017](https://doi.org/10.1088/1475-7516/2020/07/017)
- Fuskeland, U., Wehus, I. K., Eriksen, H. K., & Næss, S. K. 2014, ApJ, 790, 104, doi: [10.1088/0004-637X/790/2/104](https://doi.org/10.1088/0004-637X/790/2/104)
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, <https://arxiv.org/abs/1406.2661>

- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759, doi: [10.1086/427976](https://doi.org/10.1086/427976)
- Guth, A. H. 1981, PhRvD, 23, 347, doi: [10.1103/PhysRevD.23.347](https://doi.org/10.1103/PhysRevD.23.347)
- Hanson, D., Hoover, S., Crites, A., et al. 2013, PhRvL, 111, 141301, doi: [10.1103/PhysRevLett.111.141301](https://doi.org/10.1103/PhysRevLett.111.141301)
- Jew, L., & Grumitt, R. D. P. 2020, MNRAS, 495, 578, doi: [10.1093/mnras/staa1233](https://doi.org/10.1093/mnras/staa1233)
- Jones, M. E., Taylor, A. C., Aich, M., et al. 2018, MNRAS, 480, 3224, doi: [10.1093/mnras/sty1956](https://doi.org/10.1093/mnras/sty1956)
- Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, PhRvD, 55, 7368, doi: [10.1103/PhysRevD.55.7368](https://doi.org/10.1103/PhysRevD.55.7368)
- Kim, C.-G., Choi, S. K., & Flauger, R. 2019, ApJ, 880, 106, doi: [10.3847/1538-4357/ab29f2](https://doi.org/10.3847/1538-4357/ab29f2)
- Kodi Ramanah, D., Charnock, T., Villaescusa-Navarro, F., & Wandelt, B. D. 2020, MNRAS, 495, 4227, doi: [10.1093/mnras/staa1428](https://doi.org/10.1093/mnras/staa1428)
- Krachmalnicoff, N., Baccigalupi, C., Aumont, J., Bersanelli, M., & Mennella, A. 2016, A&A, 588, A65, doi: [10.1051/0004-6361/201527678](https://doi.org/10.1051/0004-6361/201527678)
- Krachmalnicoff, N., & Tomasi, M. 2019, A&A, 628, A129, doi: [10.1051/0004-6361/201935211](https://doi.org/10.1051/0004-6361/201935211)
- Krachmalnicoff, N., Carretti, E., Baccigalupi, C., et al. 2018, A&A, 618, A166, doi: [10.1051/0004-6361/201832768](https://doi.org/10.1051/0004-6361/201832768)
- Ledig, C., Theis, L., Huszar, F., et al. 2016. <https://arxiv.org/abs/1609.04802>
- Linde, A. D. 1982, Physics Letters B, 108, 389, doi: [10.1016/0370-2693\(82\)91219-9](https://doi.org/10.1016/0370-2693(82)91219-9)
- Louis, T., Grace, E., Hasselfield, M., et al. 2017, JCAP, 6, 031, doi: [10.1088/1475-7516/2017/06/031](https://doi.org/10.1088/1475-7516/2017/06/031)
- Mantz, H., Jacobs, K., & Mecke, K. 2008, J. Stat. Mech. P12015 J. Stat. Mech, 12, doi: [10.1088/1742-5468/2008/12/P12015](https://doi.org/10.1088/1742-5468/2008/12/P12015)
- Mustafa, M., Bard, D., Bhimji, W., et al. 2017, CosmoGAN: creating high-fidelity weak lensing convergence maps using Generative Adversarial Networks, Tech. rep.
- Page, L., Hinshaw, G., Komatsu, E., et al. 2007, ApJS, 170, 335, doi: [10.1086/513699](https://doi.org/10.1086/513699)
- Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419, doi: [10.1086/148307](https://doi.org/10.1086/148307)
- Petroff, M. A., Addison, G. E., Bennett, C. L., & Weiland, J. L. 2020, arXiv e-prints, arXiv:2004.11507. <https://arxiv.org/abs/2004.11507>
- Planck Collaboration I. 2020, A&A, 641, A1, doi: [10.1051/0004-6361/201833880](https://doi.org/10.1051/0004-6361/201833880)
- Planck Collaboration Int. XLVIII. 2016, A&A, 596, A109, doi: [10.1051/0004-6361/201629022](https://doi.org/10.1051/0004-6361/201629022)
- Planck Collaboration Int. XXX. 2016, A&A, 586, A133, doi: [10.1051/0004-6361/201425034](https://doi.org/10.1051/0004-6361/201425034)
- Planck Collaboration Int. XXXVIII. 2016, A&A, 586, A141, doi: [10.1051/0004-6361/201526506](https://doi.org/10.1051/0004-6361/201526506)
- Planck Collaboration IV. 2020, A&A, 641, A4, doi: [10.1051/0004-6361/201833881](https://doi.org/10.1051/0004-6361/201833881)
- Planck Collaboration IX. 2016, A&A, 594, A9, doi: [10.1051/0004-6361/201525936](https://doi.org/10.1051/0004-6361/201525936)
- Planck Collaboration V. 2020, A&A, 641, A5, doi: [10.1051/0004-6361/201936386](https://doi.org/10.1051/0004-6361/201936386)
- Planck Collaboration VI. 2020, A&A, 641, A6, doi: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910)
- Planck Collaboration X. 2016, A&A, 594, A10, doi: [10.1051/0004-6361/201525967](https://doi.org/10.1051/0004-6361/201525967)
- Planck Collaboration XI. 2020, A&A, 641, A11, doi: [10.1051/0004-6361/201832618](https://doi.org/10.1051/0004-6361/201832618)
- Planck Collaboration XXV. 2016, A&A, 594, A25, doi: [10.1051/0004-6361/201526803](https://doi.org/10.1051/0004-6361/201526803)
- Puglisi, G., & Bai, X. 2020. <https://arxiv.org/abs/2003.13691>
- Radford, A., Metz, L., & Chintala, S. 2016, in 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings, ed. Y. Bengio & Y. LeCun. [http://arxiv.org/abs/1511.06434](https://arxiv.org/abs/1511.06434)
- Starobinskii, A. A. 1979, Soviet Journal of Experimental and Theoretical Physics Letters, 30, 682
- Stompor, R., Errard, J., & Poletti, D. 2016, PhRvD, 94, 083526, doi: [10.1103/PhysRevD.94.083526](https://doi.org/10.1103/PhysRevD.94.083526)
- Sugai, H., Ade, P. A. R., Akiba, Y., et al. 2020, Journal of Low Temperature Physics, 199, 1107, doi: [10.1007/s10909-019-02329-w](https://doi.org/10.1007/s10909-019-02329-w)
- The Polarbear Collaboration. 2014, ApJ, 794, 171
- The Simons Observatory Collaboration. 2019, JCAP, 2019, 056, doi: [10.1088/1475-7516/2019/02/056](https://doi.org/10.1088/1475-7516/2019/02/056)
- Thorne, B., Dunkley, J., Alonso, D., & Næss, S. 2017, MNRAS, 469, 2821, doi: [10.1093/mnras/stx949](https://doi.org/10.1093/mnras/stx949)
- Tristram, M., Banday, A. J., Górska, K. M., et al. 2020, arXiv e-prints, arXiv:2010.01139. <https://arxiv.org/abs/2010.01139>
- Vansyngel, F., Boulanger, F., Ghosh, T., et al. 2018, Astronomy & Astrophysics, 618, C4, doi: [10.1051/0004-6361/201629992e](https://doi.org/10.1051/0004-6361/201629992e)
- Yu, J., Lin, Z., Yang, J., et al. 2018, arXiv preprint arXiv:1801.07892
- Zaldarriaga, M., & Seljak, U. 1997, PhRvD, 55, 1830, doi: [10.1103/PhysRevD.55.1830](https://doi.org/10.1103/PhysRevD.55.1830)

APPENDIX

A. FROM HEALPIX TO FLAT PATCHES AND BACK

In order to build the training datasets we needed to extract square images from the GNILC maps, which are projected on the sphere considering the HEALPix tessellation scheme (Górski et al. 2005). In particular, the flat images correspond to patches with angular dimension of $20^\circ \times 20^\circ$ in the sky and resolution of 3.75 arc-minutes (320×320 pixels). We chose the pixel dimension in the flat reprojection to be larger than the HEALPix one (which is 1.72 arc-minutes for a pixelization at $N_{side} = 2048$); in this way, to each pixel in the flat domain corresponds about 2 pixels in the HEALPix one and we further reduced the possibility of having missing pixels in the flat projection.

To perform the projections we made use of `astropy`⁶ and the `reproject`⁷ python packages.

From HEALPix to flat: for the polarization case, we needed to get flat images that cover the whole Celestial sphere, in order to generate small scale features everywhere in the sky. In particular, we considered 174 locations, corresponding to the center of our $20^\circ \times 20^\circ$ patches. The patch centers were taken from a grid, with number of patches in Galactic longitude changing as a function of Galactic latitude. In particular, the grid encodes 10 locations in Galactic latitude equally spaced by 18° , with each patch overlapping by 2° . The size of the grid adaptively changes in Galactic longitude: namely 7 patches are centered at $b = \pm 90^\circ$ and oriented at equally spaced longitudes, 10 patches at $b = \pm 72^\circ$ and 20 patches for the other latitude centroids. This choice was motivated by finding a trade off between having a full and homogeneous coverage of the sphere and the necessity to include some degree of overlap between patches in order to minimize border effects (see Section 4.2.1).

In the left panel of Figure 12 we show for each pixel in the HEALPix map the number of flat patches that includes it. The degree of overlap increases with the Galactic latitude, reaching its maximum at the Galactic Poles.

From flat to HEALPix: as described in Section 4.2.1 we needed to apply an apodization to the small scale patches generated by FORSE before reprojecting them into the sphere, in order to reduce the border effects. In particular, we applied a *cosine* apodized 2D window function, W_{apo} , defined as :

$$W_{apo}(x) = \begin{cases} \frac{1}{2} (1 - \cos(\pi x)) & \text{if } x < 1 \\ 1 & \text{otherwise} \end{cases}$$

where $x \equiv \sqrt{(1 - \cos \eta)/(1 - \cos \eta_{apo})}$, η being the angular separation between the pixels and the edge of the image, η_{apo} the apodization length chosen to be the same as the overlap scale (i.e. 2°). The resulting apodization window is shown in the right panel of Figure 12. The patches were then combined in the overlapping area as a weighted average with weights given by W_{apo} into a single HEALPix map.

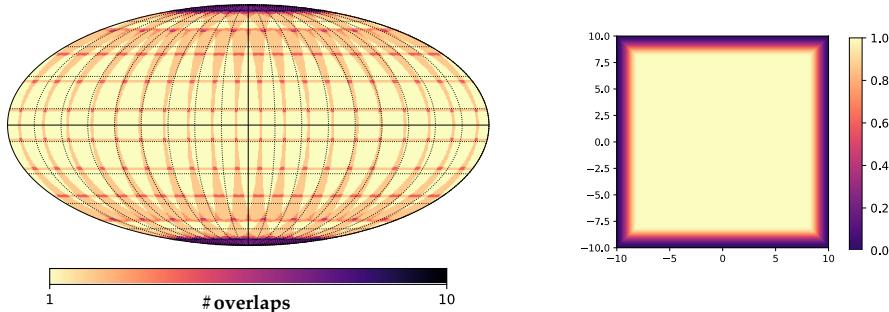


Figure 12. Left panel: map of overlaps given the scheme used to make the projections. Right panel: apodization window adopted to project the square patches into Celestial sphere with the HEALPix gridding.

⁶ <https://www.astropy.org/>

⁷ <https://reproject.readthedocs.io/en/stable/index.html>.