Neural Networks for Galactic Dust Simulations

Moutaz Haq
Supervised by: Dr. Lloyd Knox, Kevin Aylor
2018 REU, Physics Department, University of California, Davis

Abstract. B-mode polarization in the cosmic microwave background (CMB) is a signature of primordial gravitational waves. Current attempts to measure this are hampered by the presence of galactic dust which imprints an additive signal onto the CMB. Here we present partial progress toward a method of removing the unwanted dust signal, with a quantification of the uncertainty in the residuals. We show that a neural network can learn the distribution of galactic dust and produce simulations drawn from said distribution thereby virtually increasing our observation data set. We plan to train another neural network using these dust simulations as well as real images. This second neural network will act as a de-noiser to clean the effects of galactic dust from CMB observations.

Background

The discovery of the cosmic microwave background (CMB) in 1964 by radio astronomers Arno Penzias and Robert Wilson was hailed as undeniable proof of the Big Bang theory of the beginning of the Universe. Shortly after the Big Bang the Universe was filled with hot and dense plasma. The dense plasma prevented the free propagation of photons and the plasma and radiation were coupled together. After about 380,000 years of the Universe's expansion, the plasma and radiation cooled to a low enough energy where it was possible for protons to capture electrons and start forming neutral hydrogen. The neutral gas was transparent to the photons and electromagnetic radiation started to stream freely through the Universe. This event is called decoupling. The Universe is presently about 13.7 billion years old and its continued expansion since decoupling has led to redshifting of the radiation. Today, this radiation is found everywhere in the sky and is called the cosmic microwave background. The frequency spectrum of the radiation is that of a blackbody radiating at 2.7 K.

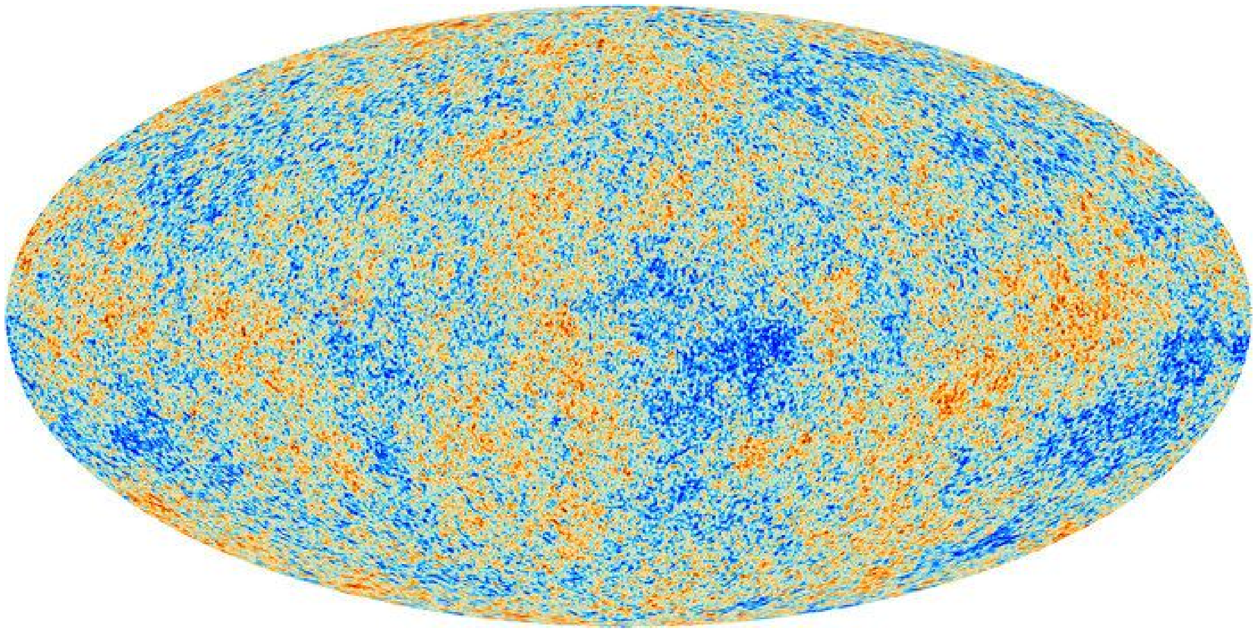

Figure 1: A temperature map of the cosmic microwave background as seen by the Planck satellite. Red areas are slightly hotter and blue areas are slightly colder than the average. Credit: ESA and the Planck Collaboration.

Although the CMB is very uniform, it does have variations depending on which part of the sky we observe. The variations can be characterized in terms of the spherical harmonics $Y_l^m(\theta, \phi)$. The CMB can be decomposed into the spherical harmonics with the coefficients a_{lm}. Summing over m and multiplying by a normalization factor results in the coefficients C_l which tell us how much angular variations corresponding to the *l*th multipole moment contribute to the CMB. Finally, $l(l+1)C_l$ is a measure of the power in the *l*th multipole moment. A plot of this versus the multipole moment is known as the power spectrum of the CMB and tells us about the variations from 2.7 K in the CMB.

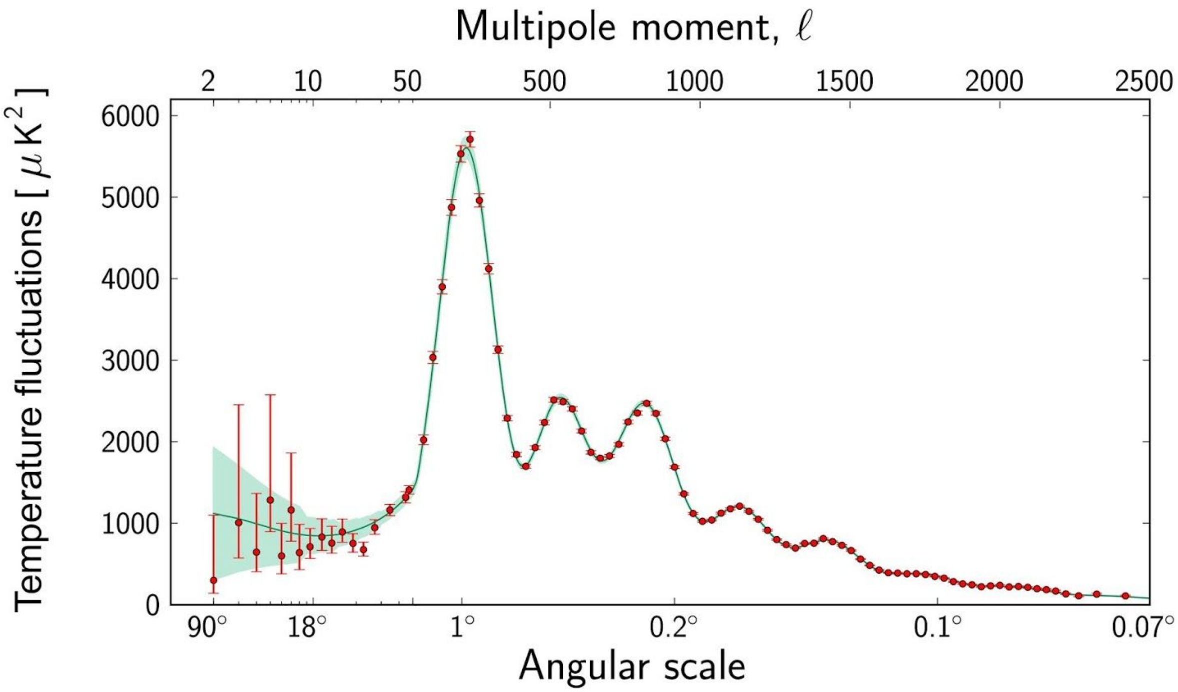

Figure 2: The CMB power spectrum. The dots indicate measurements by Planck. The solid line is the theoretical prediction by standard cosmological models. Credit: ESA and the Planck Collaboration.

The dipole moment (*l*=1) is commonly not plotted and is an artifact of our motion relative to the rest frame of the CMB. The monopole moment is simply the average temperature of the CMB (~2.7 K). Note that, ignoring the dipole moment, the variations in the CMB temperature are extremely tiny — on the order of 1 in 100,000.

The structure of the CMB gives us important information about the Universe, its structure, and its evolution. This information can be used to put constraints on cosmological models. According to inflationary models, immediately after the Big Bang the Universe underwent an extremely rapid expansion ("inflation") in which the size of the Universe increased by a factor of at least 10^{26}. This inflationary epoch ended about 10^{-33} seconds after the Big Bang. Inflation explains why the Universe is so homogeneous and isotropic and why the CMB is so uniform. Inflationary models theorize the existence of primordial gravitational waves. These are gravitational waves that started out as quantum fluctuations and were stretched to cosmic scales during inflation. Such gravitational waves would produce a certain pattern of polarization in the CMB known as B-mode polarization. However, attempts to observe B-mode polarization are hampered by an unwanted signal layered on top of the CMB: radiation from and interaction of the CMB photons with interstellar dust grains in the Milky Way galaxy.

Galactic dust grains are heated by star light in our galaxy to approximately 20 K and thermally re-radiate in a non-blackbody spectrum. Furthermore, the dust grains are preferentially aligned with the galactic magnetic field. This alignment is a result of complex interactions of the galactic magnetic field with induced currents in the dust grains since the dust grains are elongated, have angular momentum, and also move through the galactic magnetic field. The distribution of dust is random and its exact properties are not well understood. The scattering of CMB photons by the dust as well as the dust's thermal radiation contaminate the pristine CMB signal (known as foreground contamination). In fact, the scattering of CMB photons produces a foreground B-mode polarization, something which caused a problem for the BICEP2 experiment in attempts to detect primordial gravitational waves. Thus, removal of foreground contamination is key to furthering our understanding of the CMB and possibly detecting primordial gravitational waves.

Deep Convolutional Generative Adversarial Networks

We attempted to tackle this problem by employing neural networks. A neural network is a way to approximate a function through a series of linear and non-linear transformations as follows:

$$f(\mathbf{x}) \approx g_n(\mathbf{W}_n \cdots g_2(\mathbf{W}_2 g_1(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2) \cdots + \mathbf{b}_n) \qquad (1)$$

Here $f(\mathbf{x})$ is the function to be approximated. Each g_i is a non-linear transformation, each $\mathbf{W}_i$ is a linear transformation, and each $\mathbf{b}_i$ is known as a bias. Collectively $\mathbf{W}_i$ and $\mathbf{b}_i$ are the weights and biases of the neural network. The process by which the weights and biases are updated such that the output of the neural network begins to approach $f(\mathbf{x})$ is known as training.

The particular architecture we used is called a deep convolutional generative adversarial network (DCGAN). In order to use the DCGAN we split up observations of the dust into square patches of the sky measuring 20 degrees by 20 degrees. Each patch was a square image covering approximately 1% of the sky. Since cutting square patches out of the sky requires mapping from a spherical surface onto a flat surface, we limited the size of the patches to 20 degrees by 20 degrees so that each such region of the sky is well approximated by a flat surface. The patches were cut from overlapping regions of the sky such that our data set consisted of about 1000 images.

A generative adversarial network consists of two sub-networks: a generator and a discriminator. The input of the generator is a random vector and its output is an image. This image then becomes the input to the discriminator. The output of the discriminator is the probability that its input image is “real” (from the set of images of the sky) versus “fake” (not from the set of images of the sky). The discriminator is trained on both real and fake images such that it learns to discriminate between the real and fake image sets. The generator is trained such that it learns to fool the discriminator. In this way the generator learns the distribution of dust so that it can generate images that have the same statistical properties as real dust images.

In this manner we are effectively increasing our sample size of galactic dust from just one sky’s worth of observation to many. This is important because our limited sample size of one sky impedes our ability to remove foreground contamination from the CMB with any statistical significance. Finally, how can we be sure that the generator has actually learned the distribution of the dust other than just comparing generated versus real dust images by eye?

Testing the DCGAN

In order to test the efficacy of the DCGAN, we developed an image processing pipeline that ran image sets through a battery of statistics and produced certain summary statistics. The first statistic was a histogram of pixel intensities from all the pixels in an image set. The second statistic was a histogram of the power spectrum of all the images in an image set. And the third statistic was a histogram of the Minkowski functionals of all the images in an image set. The Minkowski functionals are used to characterize the topology of an object. In our case, we use the 2-dimensional Minkowski functionals of which there are three.

The power spectrum of an image was generated by a Fourier transform. It was then converted from k-space to the angular representation in terms of the multipole moments over the whole sky by multiplying each k-value by $360°/20° = 13$ since the largest Fourier mode in an image spanning 20 degrees of sky makes 13 complete cycles in a full 360 degrees.

The Minkowski functionals work by first thresholding the image such that all pixels with a value above the threshold value are set to 1 (foreground) and the remaining set to 0 (background). The first two-dimensional Minkowski functional is a count of how many pixels are in the foreground. The second Minkowski functional is a count of the perimeter around islands of foreground. The third Minkowski functional is the Euler characteristic: the number of foreground islands minus the number of background islands. In order for the Minkowski functionals to work, we log-normalized the images on $[-1, +1]$ in both the real and generated image sets using the minimum and maximum pixel values in the real image set.

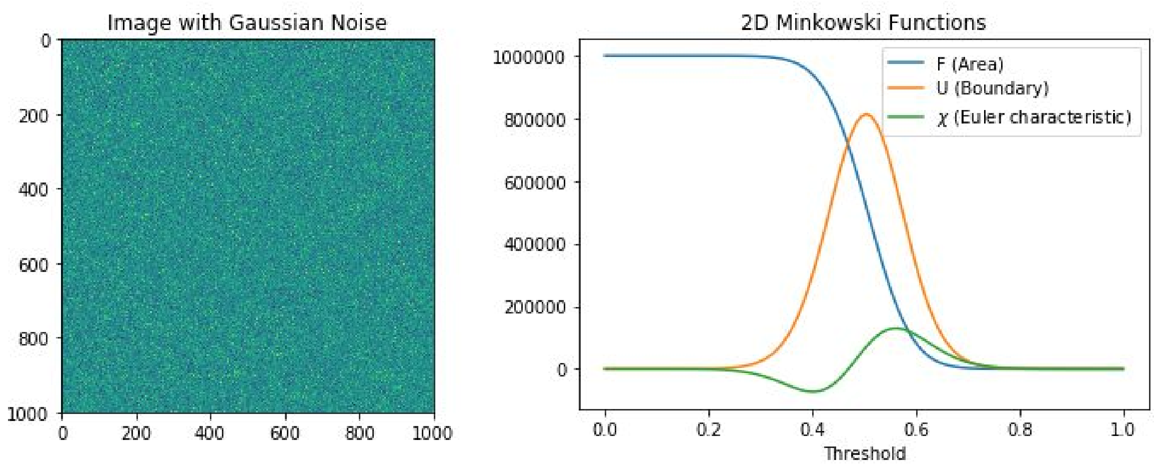

Figure 3: On the left is a 1000 by 1000 pixel image populated with random Gaussian noise. On the right are plots of the 3 Minkowski functionals for this image. The plots are characteristic of Gaussian noise.

We then compared the statistics for the set of real images against the statistics for a set of images generated by a GAN. In order to compare the statistics we used the Kolmogorov-Smirnov similarity test (KS test).

Results

Here are our results when comparing the statistics of the real image set against the statistics of those generated by one of the GANs tested, known as GAN v1.3.

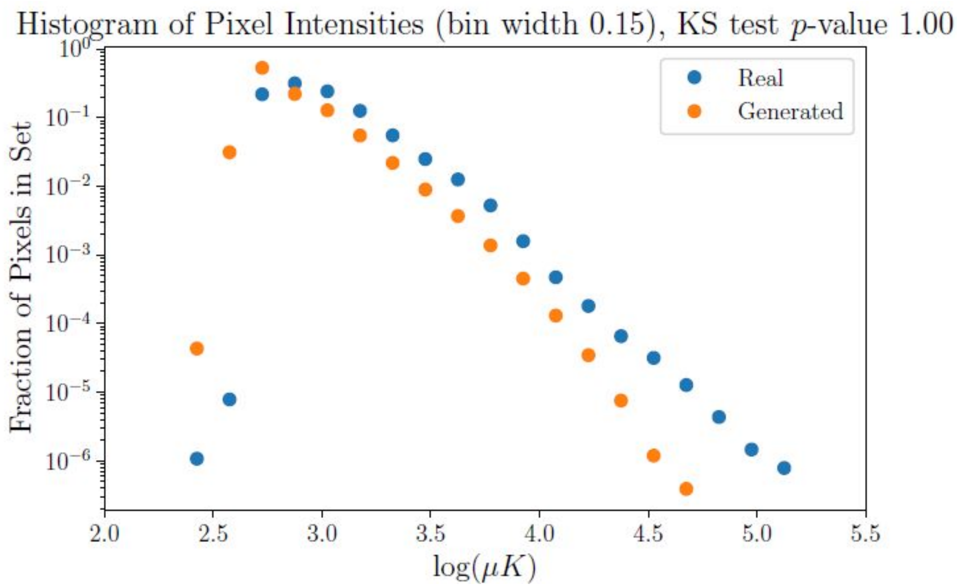

Figure 4: Histograms of pixel intensities in the real image set and a set of 1000 images generated by GAN v1.3.

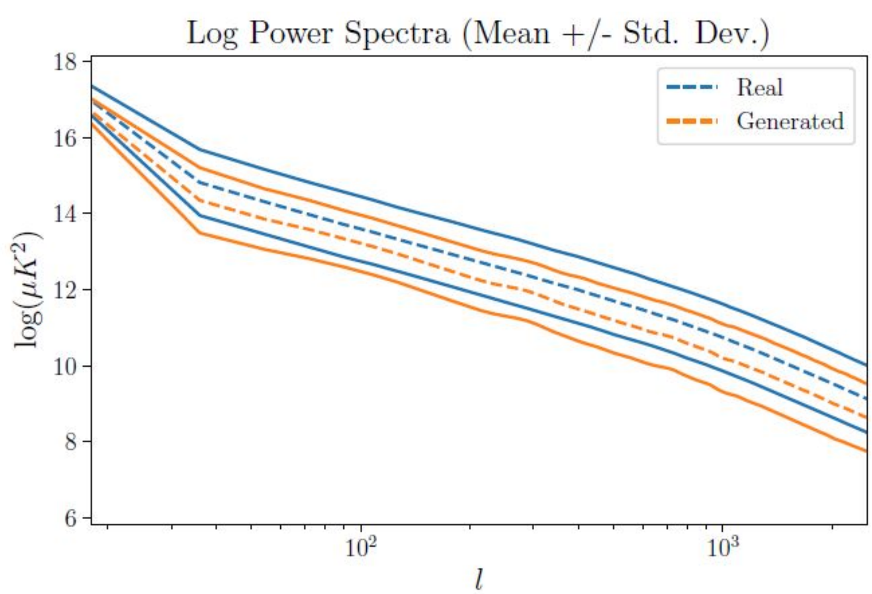

Figure 5: Log-plot of the power spectra of the images in the real set and a set of 1000 images generated by GAN v1.3. *l* is the multipole moment. Dashed lines indicate the mean power over all images in a set. Solid lines indicate 1 standard deviation from the mean.

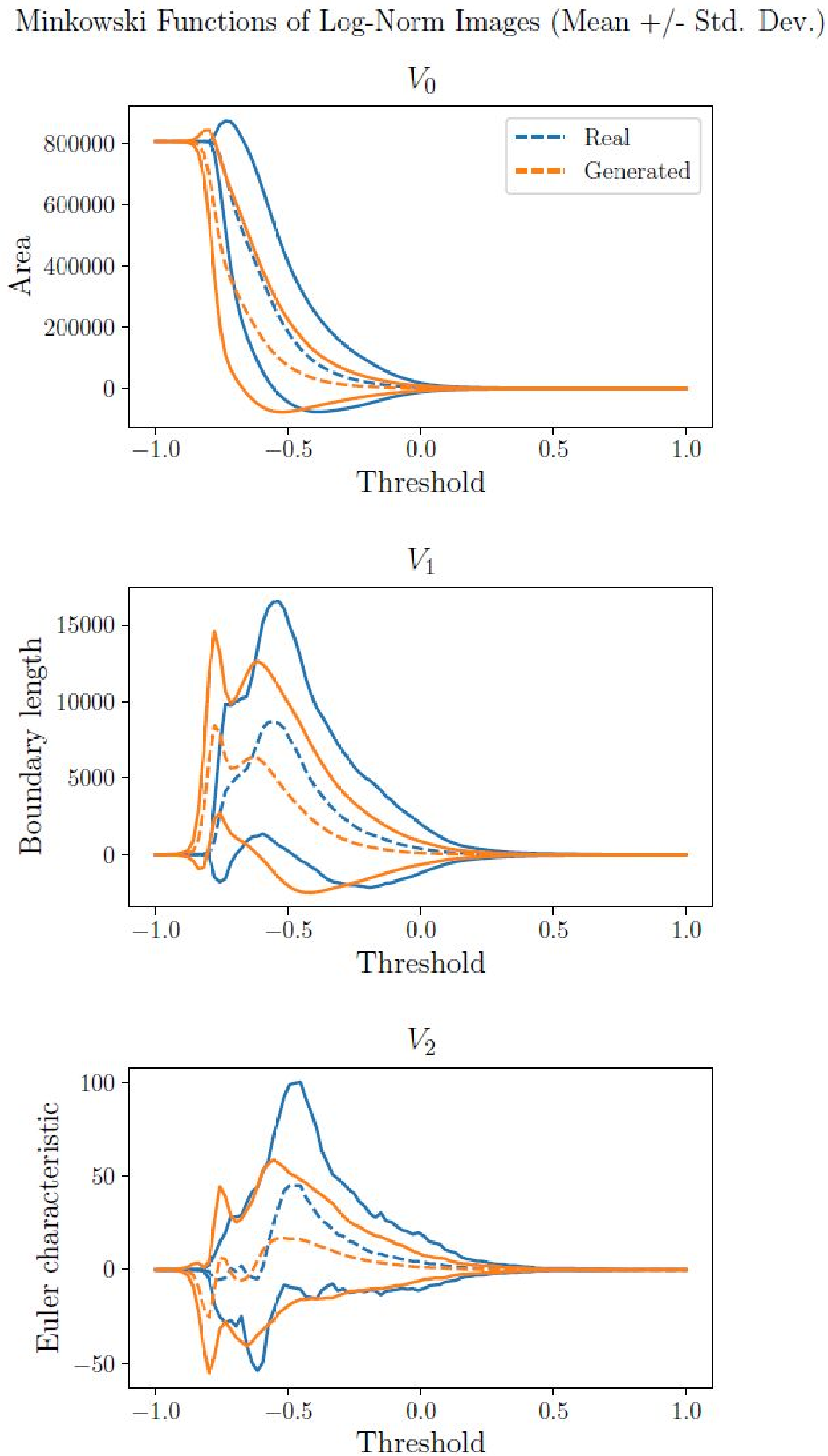

Figure 6: Plots of the 3 Minkowski functionals for images in the real set and a set of 1000 images generated by GAN v1.3. Dashed lines indicate the mean value and solid lines indicate 1 standard deviation from the mean.

Conclusion

The results showed us that GANs are capable of capturing the distribution of galactic dust. However, more work is needed to improve the GANs. Furthermore, the Kolmogorov-Smirnov similarity test is not a good one for our purposes as can be seen by the KS test *p*-value of 1.00 in figure 4.

Future Work

In future work we will train another neural network on real and DCGAN-generated dust images. We will use this neural network to de-noise the dust signal from the CMB, in a manner that allows us to estimate the statistical properties of the residual dust contamination and any noise that has been introduced.

Source Code

Source code for the image processing pipeline can be found at `https://github.com/cefarix/Dust_Simulations_Testing_Pipeline`. Source code for the GAN can be found at `https://github.com/kmaylor/K-GAN`.

References

[1] Alex Lazarian. Tracing Magnetic Fields with Aligned Grains

arXiv:0707.0858 [astro-ph]

[2] Astrophysical Foreground Maps / Planck Public Data Release 2

https://irsa.ipac.caltech.edu/data/Planck/release_2/all-sky-maps/foregrounds.html

[3] Brandon Amos. Image Completion with Deep Learning in TensorFlow.

http://bamos.github.io/2016/08/09/deep-completion

[4] A Beginner's Guide to Generative Adversarial Networks (GANs).

https://skymind.ai/wiki/generative-adversarial-network-gan

[5] CMB-S4 Science Book, First Edition

arXiv:1610.02743 [astro-ph.CO]

[6] Planck Collaboration. Planck 2018 results. XII. Galactic astrophysics using polarized dust emission.

arXiv:1807.06212 [astro-ph.GA]