

Spatial distribution of the large-scale structure: An unsupervised search for parity violation

Samuel Hewson^{1,2}, Will Handley^{1,3,4} and Christopher G. Lester⁵

¹*Astrophysics Group, Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0EZ, United Kingdom*

²*St John's College, St John's Street, Cambridge, CB2 1TP, United Kingdom*

³*Kavli Institute for Cosmology, University of Cambridge, Cambridge, CB3 0EZ, United Kingdom*

⁴*Gonville & Caius College, Trinity Street, Cambridge, CB2 1TA, United Kingdom*

⁵*High Energy Physics Group, Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0US, United Kingdom*



(Received 24 November 2024; accepted 8 May 2025; published 17 June 2025)

We use machine learning methods to search for parity violations in the large-scale structure (LSS) of the Universe, motivated by recent claims of chirality detection using the 4-point correlation function (4PCF), which would suggest new physics during the epoch of inflation. This work seeks to reproduce these claims using methods originating from high energy collider analyses. Our machine learning methods optimize some underlying parity odd function of the data, and use it to evaluate the parity odd fraction. We demonstrate the effectiveness and suitability of these methods and then apply them to the Baryon Oscillation Spectroscopic Survey (BOSS) catalog. No strong evidence for parity violation is detected.

DOI: [10.1103/PhysRevD.111.123528](https://doi.org/10.1103/PhysRevD.111.123528)

I. INTRODUCTION

A parity transformation acts to spatially invert a system and parity has, for a long time, been a defining symmetry of the Standard Model of particle physics. Over the past 30 years, progress has been made in investigating whether parity has a role to play in cosmology. Until very recently most of that work was focused on cosmic microwave background (CMB) polarization [1–6], or on gravitational waves [7–11].

In recent years, following the proposal of Cahn *et al.* [12] the large-scale structure (LSS) of the Universe has been probed for parity violations using the 4-point correlation function (4PCF). So far these studies have used the Baryon Oscillation Spectroscopic Survey (BOSS) [13] of the Sloan Digital Sky Survey (SDSS)-III [14,15] and have reported findings of parity violation at the 8.1σ [16], 7.1σ [17], and 2.9σ -levels [18].

In three dimensions the lowest order polyhedron subject to parity asymmetry is the tetrahedron, as shown in Fig. 1. The 4PCF is, therefore, the lowest order N -point correlation function which can be probed for parity violations [12]. Tetrahedra are constructed with galaxies at each of the four vertices, and the relative positions are used to calculate the 4PCF. This method facilitates searches across groups of

four galaxies but is, by design, constrained to only search for parity violations at the four-point level. This limits the type and order of parity violations that could be detected. It is possible to consider higher N -point correlation functions, but these quickly become intractable. On top of this there could be parity violations which manifest through multiple orders of correlation, which would be undetectable.

A second drawback of the method is the apparent difficulty in quantifying the significance of the detection. Philcox [18] performed both a rank test and a χ^2 test, and compared the results to 2048 MultiDarkPATCHY simulations. These simulations are calibrated to match the BOSS two- and three-point clustering statistics [20,21], not the 4-point, which could lead to an underestimate of the uncertainty in the 4PCF calculations. Hou *et al.* [17] quantify the significance by calculating the uncertainty in the covariance matrix, this could be influenced by observational effects and instrumental systematics [17], and assumes that the 4PCF follows a Gaussian distribution [22]. In fact, recent work by Krolewski *et al.* [23] shows that correctly accounting for 8-point contributions reduces the significance of detection down to between 0 and 2.5σ .

The aim of this paper is to demonstrate that more general methods can be used to search for parity violations in the LSS. Such methods do not have to be specifically 4-point, or even N -point, but rather consider the data in spatially localized regions and search for underlying parity violations across all N -point correlations. A more specialized method will than be implemented for the 4-point case, constructed with no knowledge of a 4PCF, and

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

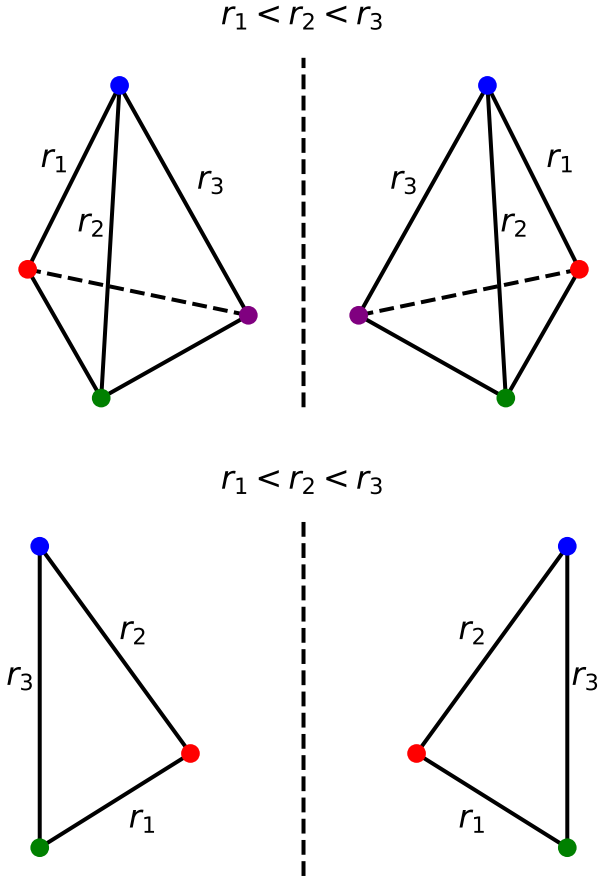


FIG. 1. Top: a tetrahedron is the lowest order polygon subject to parity violation in three dimensions, therefore the 4PCF is the lowest order N -point correlation function which can be probed for parity violations. Bottom: three points can be parity violating in two dimensions, but in three dimensions the 2D parity operation becomes equivalent to a 180° rotation. Consider rotating into the page about the reflection axis. Figure adapted from Taylor *et al.* [19].

implemented to see if the results of Hou *et al.* [17] and Philcox [18] can be reproduced.

Similar methods to the ones we applied have also been proposed and tested with CNNs and neural field scattering transforms (NSFT) [19,24]. With the recent data release from the Dark Energy Spectroscopic Instrument (DESI) [25,26] and the upcoming releases from EUCLID [27,28] and the Nancy Grace Roman Space Telescope [29,30], the methods proposed here will be useful in future exploration of the LSS.

II. THEORY AND METHODS

A. Theory of parity

In quantum mechanics the parity operator, \hat{P} is defined to spatially invert the wave function of a particle, such that a wave function $\Psi(x)$ is Parity-even if $\Psi(\hat{P}x) = \Psi(x)$ and

Parity-odd if $\Psi(\hat{P}x) = -\Psi(x)$. The importance of parity pervades throughout physics, most notably in the Standard Model of particle physics, where parity is a symmetry of both quantum electrodynamics (QED) and quantum chromodynamics (QCD) interactions, but is violated by weak interactions [31,32]. The creation of the present-day matter-antimatter asymmetry would require baryogenesis processes that violate charge and parity conservation [33,34], but since gravity is parity symmetric [35], all cosmological correlators should be too. This means that parity asymmetry in the LSS would require new physics in the epoch of inflation.

B. Method of Lester and Tombs

Current theoretical models predict parity violating mechanisms which are not observable at the Large Hadron Collider (LHC) [36], and accordingly the LHC has not been used to search for these. Parity may however be violated by unknown means, which could be detectable in LHC deposit data. Motivated by this possibility, Lester and Tombs [37] demonstrated that machine learning can be used to detect parity violating functions within datasets without any required knowledge of what is causing the violation. By training an image classification convolutional neural network (CNN) and rewarding it for detecting asymmetry, a measure of the parity of the image dataset can be determined. To demonstrate this method, Lester and Tombs [37] applied their algorithm to the MNIST data as well as more general symmetries [38]. The success of the algorithm on the MNIST data also displays the generality of the method, and its applicability outside particle physics. To further demonstrate the general success of the Lester and Tombs [37] method, a Master's project was run in 2022/3 aiming to use the method to test for parity violations in snails [39]. The population was known to be parity violating as most snail species have a strong preference for the handedness of their shells' spirals. It was successfully shown that the unsupervised learning method could detect parity violations within the images.

1. Algorithm

To quantify any parity violations existing in a sample, consider some function $g(x)$, where g represents the outputs of the CNN with inputs x . Then construct $f(x)$ to be the difference between the outputs from x , and the parity transform of x

$$f(x) = g(x) - g(\hat{P}x) \quad (1)$$

from which it is immediately clear that $f(\hat{P}x) = -f(x)$. In fact further symmetries of the system can be specified: let $S = \{S_1, S_2, \dots, S_i\}$ be the set of all the symmetries of the system. Then $f(x)$ can be defined as the sum over S_i of

the difference between the outputs from the original and parity transformed inputs as demonstrated in Eq. (2)

$$f(x) = \sum_i g(\hat{S}_i x) - g(\hat{S}_i \hat{P} x). \quad (2)$$

Equation (3) is an example of this where $f(x)$ has been constructed to obey invariance under 180° rotations, represented by \hat{R}_{180} .

$$f(x) = g(x) + g(\hat{R}_{180}x) - g(\hat{P}x) + g(\hat{R}_{180}\hat{P}x) \quad (3)$$

This creates a function for which $f(x) = f(\hat{R}_{180}x)$, enforcing 180° rotational symmetry and removing any idea of a top from the images. Accounting for symmetries in f means they do not need to be accounted for in the subnetwork g , reducing the processing required on the inputs. Further symmetries are enforced by max-pooling layers within the network.

Equation (1) will output zero for a parity even function $g(x)$, and the network can be constructed to return outputs g such that the difference of f from zero is maximized. For each batch $B = \{x_1, x_2, \dots, x_{|B|}\}$, the mean value of f is calculated.

$$\mu_B = \frac{1}{|B|} \sum_{x \in B} f(x). \quad (4)$$

The loss function is defined

$$\mathcal{L}_B = -\mu_B / \sigma_B, \quad (5)$$

where

$$\sigma_B^2 = \frac{1}{|B|} \left(\sum_{x \in B} [f(x)]^2 \right) - \mu_B^2 \quad (6)$$

is the variance of f over batches, which ensures that the network will not trivially increase the loss under $f(x) \mapsto \lambda f(x)$ for $\lambda > 1$.

After training the neural net, the function $f(x)$ can be evaluated on the testing data. If the mean value of f on that dataset differs from zero by a statistically significant amount, some parity-violating feature in the data has been detected. We compute the fraction of the outputs for which $f(x) > 0$. This positive fraction quantifies the proportion of the inputs that exist in a parity violating way, which means achiral datasets are expected to output 50%, and deviations from this imply a parity violating dataset.

It is important to note here that when searching for a violation of some symmetry then a function f can provide evidence if both:

- (1) f has the correct symmetry properties to permit the discovery, and

- (2) f yields statistically significant outputs when evaluated on independent datasets that it has not seen before and that were not used in its construction

Here we have constructed f by hand to be parity-odd, which satisfies the first requirement. The second requirement is demonstrated in Sec. IV. From these principles it follows that the most important step here is the construction of f . This method could, in principle, work without any training. The training step just acts to increase the likelihood of a discovery by optimizing, but not changing the form, of f . This also means that the net has the benefit of never creating a false positive from overtraining. If there is a positive detection, the data must inherently contain the parity violation specified by the function f , regardless of whether that f was obtained by overtraining.

III. DATASET

Following from Hou *et al.* [17] and Philcox [18], we use data release 12 (DR12) of the BOSS catalog, which contains LOWZ and CMASS samples, each being split into the North Galactic Cap (NGC) and South Galactic Cap (SGC). The target selection algorithm results in samples of primarily luminous red galaxies (LRGs), with CMASS further specified to select objects of uniform redshift [40]. This leads to a roughly mass-limited sample down to a stellar mass $M \sim 10^{11.3} h^{-1} M_\odot$ [41–44]. A redshift cutoff of $0.43 < z < 0.7$ is applied, which raises the purity of the CMASS sample by adding more LRGs [17]. For this investigation, the CMASS NGC was used, and the catalog of galaxies was treated as a field of points in three-dimensional space, a 2D projection of this is shown in Fig. 2. The BOSS dataset contains data in the form of right ascension, declination and redshift. This was converted into spatial coordinates using the fiducial cosmology specified by the BOSS collaboration [45]. Namely, a flat Λ CDM model, with matter density $\Omega_m = 0.31$, Hubble constant $h \equiv H_0 / (100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.676$, baryon density $\Omega_b h^2 = 0.022$, fluctuation amplitude $\sigma_8 = 0.8$ and spectral tilt $n_s = 0.97$. From this, the question can be asked: Is the spatial distribution of the points in the field parity violating? More specifically, is the spatial distribution of the points in the field parity violating on a given scale? The scale of the parity violation is specified to match the work of Hou *et al.* [17]. This means for the geometric analyses, groups are constructed such that points within a group are separated by a distance, ranging from $20h^{-1} \text{ Mpc}$ to $160h^{-1} \text{ Mpc}$.

To each galaxy a weight was applied

$$w = w_{\text{fkp}} w_{\text{sys}} (w_{\text{rf}} + w_{\text{cp}} - 1) \quad (7)$$

where w_{rf} is the redshift failure weight and w_{cp} the fibre collision weight; these are additive weights with default value of unity. w_{sys} is the systematic weight, which

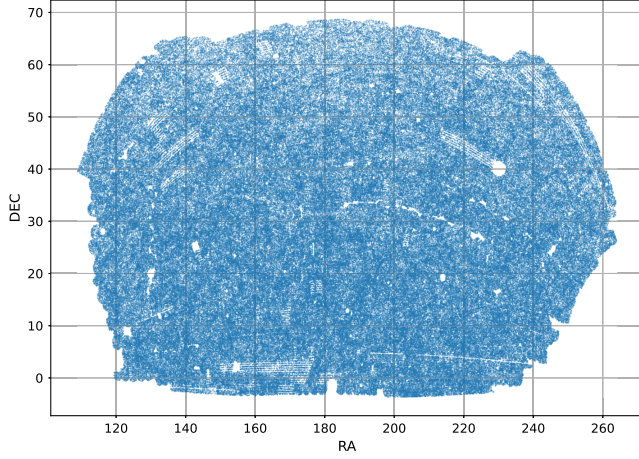


FIG. 2. The North Galactic Cap of the CMASS catalog.

accounts for stellar density and seeing [46], and the FKP weight is defined as $w_{\text{fkp}} = [1 + n(z)P_0]^{-1}$ [47], where $n(z)$ is the weighted number density as a function of redshift, and $P_0 = 10^4 [h^{-1} \text{ Mpc}]^{-3}$.

IV. VALIDATING THE METHODS

To begin, in Sec. IV A we use the CNN architecture of Lester and Tombs [37] to examine a two-dimensional projection of the data. This is followed in Sec. IV B by the addition of color channels to the images, with the aim of representing three-dimensional information via color. Finally, we investigate more geometric approaches in Sec. IV C, which have the drawback of being constrained to some order of polyhedron, much like N -point correlation functions, but consequently allow direct comparison with the 4PCF.

A. 2D convolution

We begin by using the same architecture as Lester and Tombs [37]: 2D convolutional layers, with image inputs. For our investigations, the data can be treated as a field of points, from which we constructed images to represent localized regions of the field. To ensure that this method is suitable for detecting parity violations in image data, we ran a number of tests using simulated data. These not only verify that the network is able to detect parity asymmetric data, but also help to determine a level of significance, i.e., what net fraction of the data must contain a parity violating object to result in a nonzero detection. The choice of how to represent galaxies in the images is somewhat arbitrary, and two possible examples are shown in Fig. 3. After testing multiple methods (Appendix A), we decided to use PNG images generated from scatter plots of data points. An outline of this method follows:

To begin, a field of random data points was generated, and image samples created in the following manner:

- (1) Sample a random point from the dataset, O

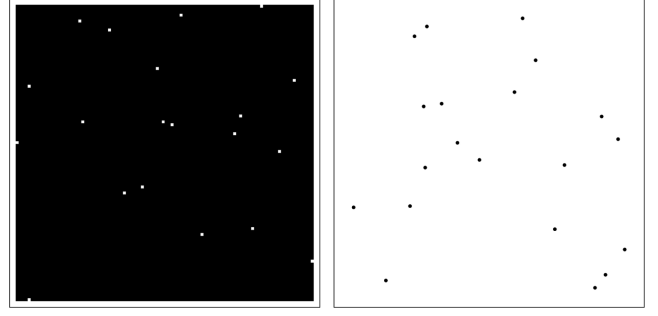


FIG. 3. Two demonstration images generated from random data. Left: an image where each galaxy has its localized pixels colored. Right: an image where each galaxy is represented by a scatter point.

- (2) Generate a square of side length l , centered on O
- (3) Rotate the square by a random angle, $\theta \in [0, 2\pi]$
- (4) Get all the data points lying within the rotated square
- (5) Create an image of the square of data points

This ensures that all the images are the same size, 64×64 pixels, following the work of James Gibbon [39]. Testing images are generated first, and then training ones with the added condition that these images may not overlap with the testing set, to avoid overfitting. An exaggerated sample is displayed in Fig. 4 to clearly show the method of sampling squares.

1. Testing method with toy data

The toy data tests were constructed on the same scale as the real images would be. As outlined in Sec. V images covering $0.5^\circ \times 0.5^\circ$ were chosen, giving an average of 22.3 galaxies per image. For the toy models we generated

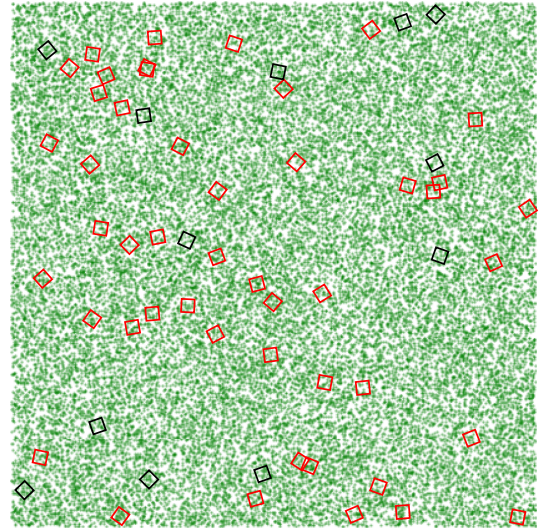


FIG. 4. A demonstration of the image sampling technique, on a background random field of 500,000 data points, with 12 testing samples shown in black and 48 training samples shown in red.

TABLE I. A summary of the crucial image based toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Description	Purpose	Detection	Control
RI	Random Generated Image Data	Check sampling method is not introducing parity violation	$49.9 \pm 0.3\%$	$49.8 \pm 0.3\%$
MNIST-RI	RI with random MNIST image overlaid in random position	Check detection of complex parity violation	$87.5 \pm 0.3\%$	$50.3 \pm 0.3\%$
F-MNIST-RI	MNIST-RI, but each MNIST image has a 50% chance of being flipped vertically before inserting	Check method of insertion is not causing spurious detection	$50.0 \pm 0.3\%$	$50.4 \pm 0.3\%$
SPR-RI-100	RI with 100 point spiral inserted in random position and random radius	Is parity violation in the same form as the data detected	$98.2 \pm 0.1\%$	$49.5 \pm 0.3\%$
F-SPR-RI-100	SPR-RI-100, but each spiral has a 50% chance of being flipped vertically before inserting	Check method of insertion is not causing spurious detection	$49.6 \pm 0.3\%$	$49.5 \pm 0.3\%$
SPR-RI-10	RI with 10 point spiral inserted in random position and random radius	Is parity violation detectable on a more realistic scale	$97.9 \pm 0.1\%$	$50.1 \pm 0.3\%$
F-SPR-RI-10	SPR-RI-10, but each spiral has a 50% chance of being flipped vertically before inserting	Check method of insertion is not causing spurious detection	$49.8 \pm 0.1\%$	$51.6 \pm 0.3\%$
SPR-RI-3	RI with 3 point spirals (triangles) inserted in random position and of random radius	Is parity violation detectable on the smallest scales	$93.5 \pm 0.2\%$	$51.1 \pm 0.3\%$
F-SPR-RI-3	SPR-RI-3, but each spiral has a 50% chance of being flipped vertically before inserting	Check method of insertion is not causing spurious detection	$51.9 \pm 0.3\%$	$49.2 \pm 0.3\%$
LH-RI-0.1	RI where every point has 2 extra points generated close to it forcing a parity odd “left-handed” triangle on the scale of 1 tenth of the average interpoint distance	Can other types of parity violation be detected	$74.6 \pm 0.3\%$	$49.6 \pm 0.3\%$
LH-RI-1	LH-RI-0.1 with extra points on the scale of the average interpoint distance	Can more realistic types of LH-RI-0.1 parity violation be detected	$67.6 \pm 0.3\%$	$49.2 \pm 0.3\%$
LH-RI-10	LH-RI-0.1 with extra points on the scale of ten times the average interpoint distance	Can more realistic types of LH-RI-0.1 parity violation be detected	$52.3 \pm 0.3\%$	$49.13 \pm 0.3\%$
SPR-RI-5-1%	RI with 5 point spiral inserted in random position and random radius in 1% of images	How sensitive is the detection	$57.6 \pm 0.3\%$	$49.8 \pm 0.3\%$
SPR-RI-5-0.1%	RI with 5 point spiral inserted in random position and random radius in 0.1% of images	How sensitive is the detection	$47.6 \pm 0.3\%$	$49.5 \pm 0.3\%$
SPR-RI-5-0.01%	RI with 5 point spiral inserted in random position and random radius in 0.01% of images	How sensitive is the detection	$48.0 \pm 0.3\%$	$48.8 \pm 0.3\%$

12,000 images, split into 9,600 training and 2,400 testing, with each image averaging 22.3 data points.

A number of toy datasets were curated in order to investigate the effectiveness of the method under different scenarios. The critical datasets are outlined with a description and purpose in Table I, with example images

displayed in Fig. 5. Each dataset has an associated control dataset, which ensures the function of the net. A control group is generated by iterating through the real group, and flipping each image with a 50% probability. For any large parity odd dataset this symmetrization process always creates an achiral dataset, which can act as a

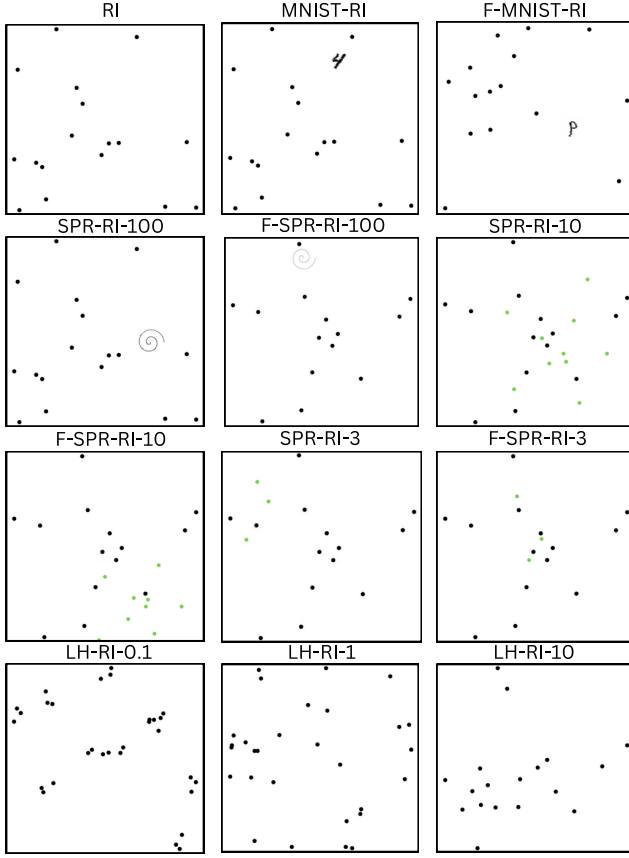


FIG. 5. A sample of toy images labeled by their data codes from Table I. For the purpose of demonstration, smaller spirals that have been overlaid have been colored green. During testing, all images were in black and white.

control.¹ Each data type was created from the same randomly generated background field, repeated over 10 different background fields to create 10 variations of each

¹A right-handed figure drawn on a flat piece of tracing paper becomes left-handed if you are allowed to turn the tracing paper over and view it from the back. The handedness of two-dimensional objects are therefore only well-defined if the surfaces on which they live are orientable or are oriented in some special way (s). The toy datasets in Table I are all intrinsically two-dimensional and so, for the reason just given, any of these datasets which violate parity only do so on account of the embedded objects being implicitly pasted the inside of the celestial sphere, centered on earth, from where they are viewed. Such geocentric parity violation is, of course, entirely unlike the sorts of large scale parity violation one might want to see in the real universe. Any large scale parity violation will (presumably) not rely on the earth being at a special place in the universe. The geocentric nature of the parity violation in these toy datasets does not prevent them being used for the purposes they were created, however. Furthermore, this geocentricity is completely removed in the parity violating datasets introduced later in Table II. These later toy datasets contain projections of intrinsically three-dimensional (parity violating) structures without any preferred orientations, and so are representative of a sort of parity violation that one might hope to see in reality.

data type. The average positive fraction over the 10 repeats is listed in Table I.

2. Summary of tests

The results of the toy data samples demonstrate that the net does not find parity violations where there are none and that it does detect parity violations where they exist on large enough scales. Crucially, it also shows that the method of sampling is not creating spurious detection. The ability to detect varying types of complex parity violation suggests that this method should be successful in finding any such violations in the real data.

These tests also give insight into the sensitivity of the method: having just 0.1% of the images contain parity violation is enough for it to be successfully detected. Perhaps the most powerful tests are SPR-RI-3 and LH-RI-10, which show that images without any visible asymmetry, can be detected by the net.

B. 2D Convolution with color

The 2D CNN is constrained by its loss of 3 dimensional information, meaning not all forms of 3D parity violation are detectable. A first attempt to resolve this is made by introducing color to the images. By normalizing the axis into the page and translating it to pixel color, it is possible to add extra information to the two-dimensional images. With this modification each image represents a voxel of the space displayed as a 2D projection, with the third dimensional information encoded in the color of a given point, this creates a color CNN (cCNN).

As in the previous case, we generate a random field of points, now in 3D space. Images are constructed following the same process outlined in Sec. IV A, with the additional step of using z values and weights to color the pixels. Under this construction, each pixel has two degrees of freedom we need to express. The first is the distance of the galaxy and the second is its weight. To represent these degrees of freedom we assign RGB values to the pixels in the following manner:

$$(r, g, b)_{jk} = \sum_{i \in jk} w_i \cdot \mathbf{c}(z_i) \quad (8)$$

$$\mathbf{c}(z) = z \cdot \mathbf{R} + (1 - z) \cdot \mathbf{B} \quad (9)$$

Where w_i is the weight of the i th galaxy in pixel jk and $\mathbf{R} = (1, 0, 0)$ and $\mathbf{B} = (0, 0, 1)$. In this way the distance correlates to the ratio between the red and blue channels, and the weight is represented by the intensity of the pixel. The background is set to $(0, 0, 0)$, so the net is looking at $\mathcal{O}(1)$ deviations from 0. A demonstration of this is visible in Fig. 6.

To account for the slight issues that arise from the fact that some of the weights are larger than 1.0 and the fact that the clipping of values to fit in the integer range $(0, 255)$ will

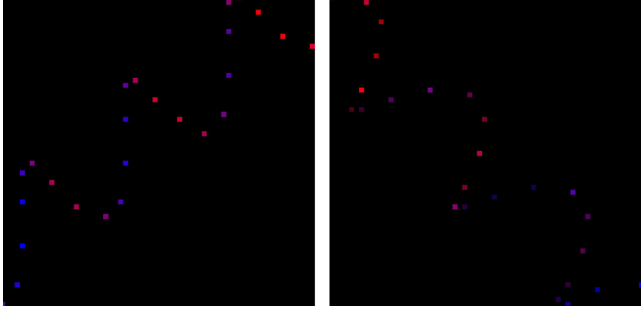


FIG. 6. Sample images containing helices. The coloration creates a parity violating dataset. Left: all galaxies have weight 1.0, and therefore are all the same intensity. Right: Galaxies have a randomly assigned weight from (0.25, 0.5, 0.75), this affects intensity.

cause some rounding issues, we also create an array dataset. This dataset contains the same 3 color channel arrays of 64×64 pixels as would be extracted from the images, but contains slightly improved accuracy values. This array can then easily be manipulated in the same way as the image data.

Under this construction we use the following identity to simplify the parity operator, to just consider an x flip

$$\hat{P} \equiv \hat{P}_x \hat{P}_y \hat{P}_z = \hat{P}_{i \in \{x,y,z\}} \hat{R} \quad (10)$$

We then enforce rotational invariance in the 2D plane, so for inputs α construct $f(\alpha)$ in the following way:

$$f(\alpha) = \sum_{\hat{R}_i} [g(\hat{R}_i \alpha) - g(\hat{R}_i \hat{P}_x \alpha)], \quad (11)$$

where $\hat{R}_i \in [\hat{R}_0, \hat{R}_{90}, \hat{R}_{180}, \hat{R}_{270}]$. For this setup then, we can use the same net as we used for the black and white images.

1. Testing method with toy data

To effectively test the functionality of this method, various datasets were constructed in three-dimensional space. The key results follow in Table II and a full list of the important test can be found in Appendix B. Because we are using the same net as for the 2D convolutions, we can directly compare the coloration scheme outlined in Sec. IV B with inputs where all galaxies are monochrome points, as show in Fig. 7. This shows significant improvement in detection, and results in detections that would not otherwise be made. These tests illustrate the success of the introduction of color channels as a method for including the third dimensional information.

C. 2D angle based method

To further the investigation into three dimensions, and facilitate testing a corollary to the 4PCF a new neural net was constructed with the same guiding principles as the CNN. Since the inputs were numeric, two options were considered: 1D convolutions or linear layers. We performed the tests on the toy data using both separately. We find that both perform equally well, but the linear layers are marginally faster, so these were chosen to be used on the real data. We begin by considering the two-dimensional case. In 2D the lowest order polyhedron subject to parity violation is the triangle. A triangle in 2D space, is defined by six degrees of freedom (d.o.f.)—three two-dimensional coordinates—to which the constraints of rotational and scale invariance are added, these reduce the six d.o.f. down to two. Two element inputs are then sufficient, and a convenient choice is the internal angles of the triangle. To make the angles subject to parity violation, they require ordering in a way that would change under the action of a parity operator. The method used is outlined as follows for groups of 3 points:

- (1) Find the mean position of the group and shift it to be the origin

TABLE II. A summary of the crucial color image based toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection; the control column shows the positive fraction for the control group and the uncolored column shows the detection for the same images colored monochrome.

Data code	Description	Purpose	Detection	Control	Uncolored
RCI	Random generated image data	Check sampling method is not introducing parity violation	$49.8 \pm 0.3\%$	$50.3 \pm 0.3\%$	$49.9 \pm 0.3\%$
HE-RCI	RCI with 1 left-handed helix inserted into the dataset for every 100 random points	Check detection of complex parity violation	$57.5 \pm 0.3\%$	$50.2 \pm 0.3\%$	$49.8 \pm 0.3\%$
F-HE-RCI	HE-RCI, but each helix has a 50% chance of being left- or right-handed	Check method of insertion is not causing spurious detection	$50.5 \pm 0.3\%$	$50.4 \pm 0.3\%$	$50.2 \pm 0.3\%$
PV-RCI	Parity violating groups of 4 points randomly oriented and distributed. Local separation \ll average interpoint distance.	Can parity violations be detected where the 2D dataset does not detect	$87.2 \pm 0.1\%$	$49.6 \pm 0.3\%$	$50.8 \pm 0.3\%$

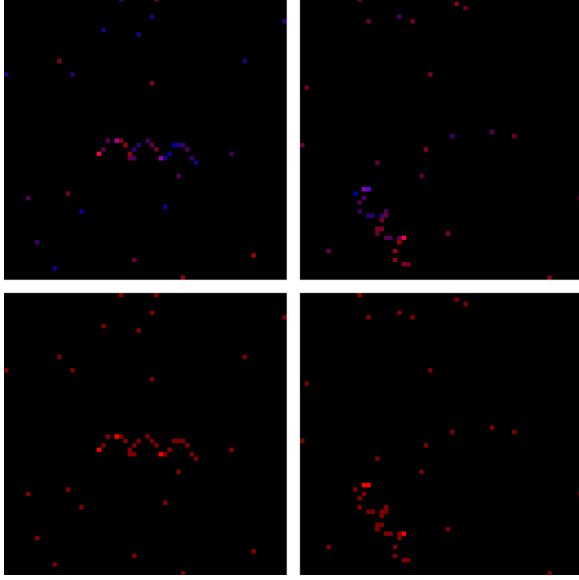


FIG. 7. Sample images from a 3D dataset containing left-handed helices and a random field of points. Top: images where points are colored by their distance into the page. The 3D information creates a parity violating dataset. Bottom: all points are the same color, the loss of information creates a dataset which is not parity violating.

- (2) Compute the distances of the points to the origin use this to order the points with the closest as the first
- (3) Reorder the second and third in an anticlockwise direction from the first
- (4) Compute the angles between the points in an anticlockwise fashion
- (5) Define \hat{P} to swap the first and third angles
- (6) Use the first two angles in every group of 3 as the net inputs

The principles of this method are displayed in Fig. 8, which illustrates why this is the appropriate parity operator. In fact this method is generally extensible to any n -gon in 2D, with the first $n - 1$ angles being the input and the parity operator inverting the order of the angles as shown in Eq. (12). Conveniently for the function of the neural net, the n th angle can be calculated from the others, as given in Eq. (13).

$$\hat{P}[\theta_1, \theta_2, \dots, \theta_{n-1}] = [\theta_n, \theta_{n-1}, \dots, \theta_2] \quad (12)$$

$$\theta_n = 2\pi - \sum_{i=1}^{n-1} \theta_i \quad (13)$$

1. Testing method with toy data

As with the CNN, our first step was to test the efficacy and sensitivity of this method. This began with testing a random field and then working through a number of toy

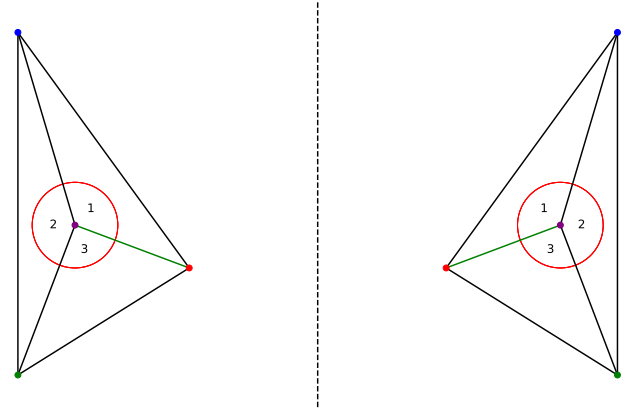


FIG. 8. A demonstration of the angle method in 2D. The line connecting the closest vertex to the centroid is shown in green. Working anticlockwise from this line, the initial angles are ordered [1,2,3], under a parity transformation this changes to [3,2,1].

data scenarios, the critical ones are outlined in Table III, and the full list can be found in Appendix B. These results crucially demonstrate that this method does not detect parity violation if not present, and can detect parity violation if it is present in at least 1% of the input data. The main drawback of this method is that for it to be successful, the input data must be sampled such that it faithfully represents the parity of the system. Conceptualize this by considering a random uniform field of two-dimensional data points. For every point in the field generate two more points, positioned relative to the first so that they form a chiral triangle, and enforce the average separation of the points in each group of three to be one tenth of the average interpoint distance. This creates a visually parity-violating dataset, as displayed in Fig. 9. If the inputs to the net are calculated from these groups of three, there is a strong parity detection, as expected. However, the real data will not have the benefit of being preselected into parity violating groups, and therefore to produce a fair test it is necessary to generate the field as before, and then sample groups of three randomly from the total field. In this scenario, it becomes almost impossible to find the parity violation, because each point in a group is being sampled from a 500,000-point field, so the chances of selecting a group that represents the chirality is negligible. This is shown in Table III with datasets PV-0.1-R2d and S-PV-0.1-R2d. This issue can be mitigated by requiring samples to be drawn on a given scale. With this additional requirement, detection of parity violation becomes significant, as long as the scale constraint matches the scale on which the parity violation exists. This idea is demonstrated in Fig. 10 and shown in Table III with dataset C-PV-0.1-R2d. Testing over different constraint scales, the data can be investigated in an analogous way to power spectrum analyses.

TABLE III. 2D toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Description	Purpose	Detection	Control
R2d	Random generated 2D point data	Check parity violation is not detected where there is none	$50.41 \pm 0.10\%$	$50.10 \pm 0.09\%$
PV-0.1-R2d	R2d with two extra points generated near each R2d point, these are generated in a way that all groups of three are of the same chirality. Groups of three are separated on the scale of one tenth the average interpoint distance	Check detection of parity violating inputs	$90.77 \pm 0.06\%$	$49.74 \pm 0.10\%$
S-PV-0.1-R2d	PV-0.1-R2d, but after generating parity violating groups all points shuffled together and new groups drawn at random	Can parity violations be detected when they are visibly present, and groups sampled randomly from the dataset	$49.70 \pm 0.10\%$	$49.68 \pm 0.10\%$
C-PV-0.1-R2d	S-PV-0.1-R2d, with a distance constraint when generating groups from the shuffled data	Can parity violations become detectable by constraining the scale	$65.10 \pm 0.04\%$	$49.85 \pm 0.04\%$

2. Summary of tests

After adding the constraint on the scale of the groups, a detection is made with a high level of significance if the constraint appropriately selects groups of data with at least $\sim 1.0\%$ containing parity violation. These tests were also extended to groups of 4 and 5 points, the full results of which can be found in Appendix C.

D. 3D angle based method

Now that the success of an angle based method in 2D has been demonstrated, this motivates exploration of a 3D equivalent. To detect parity violation, groups of four points are the smallest scale on which this method will work, and so this can act as a direct comparison to the 4PCF. As in the two-dimensional case, the critical part of this method is

defining an ordering of the angles and understanding how this ordering changes under a parity operator. There are a number of ways to define an ordering, the process used for groups of four points is outlined as follows, and displayed in Fig. 11:

- (1) Compute the mean position of the group and shift it to be the origin
- (2) Calculate the distances of the points to the origin, and order the points by distance with the closest point as the first
- (3) Use the first position as an axis, and starting from the second-closest point, work around the axis in a right-handed manner to determine the order of points 3 and 4
- (4) Compute the internal angles and order them as $[\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{34}, \theta_{24}]$

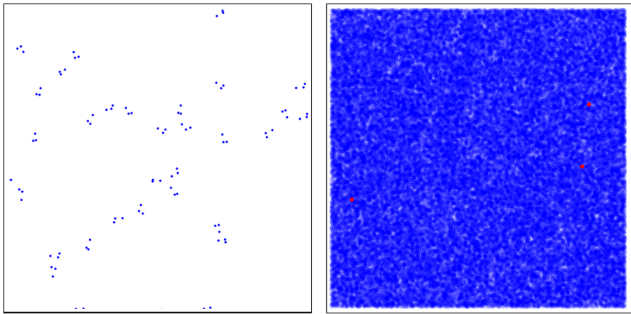


FIG. 9. An illustration of the failure of the 2 dimensional angles method to find parity violations across arbitrary scales. Left: a localized region of the field demonstrating the observable parity violation in the data. Right: the full field of points, with a randomly sampled group of 3 three shown in red. This demonstrates the impossibility of finding the parity violation when sampling across any scale: the chance of picking three points that exist in a parity violating way is negligible.

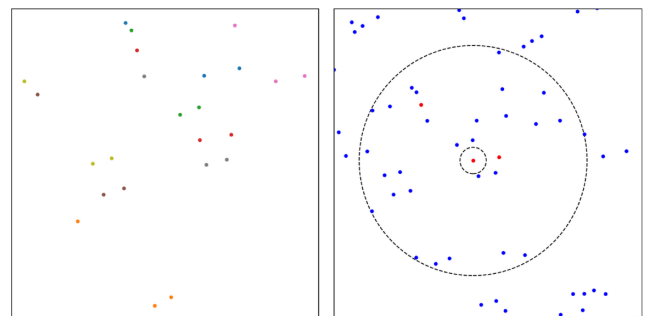


FIG. 10. By constraining the scale on which to construct a group of points, the net has a significantly improved likelihood of detection. Left: a localized region of the field of points. The field is chiral by construction, and each group of three has been colored to show the asymmetry. Right: adding a constraint on the scale over which groups can be constructed significantly increases the chances of detection.

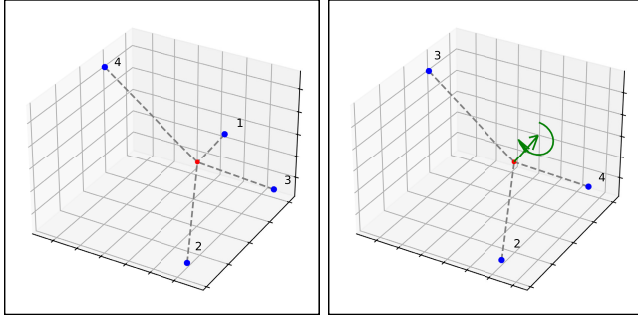


FIG. 11. An illustration of the ordering of angles in 3 dimensions. The images show a group of four points with their mean position in red. Left: the points are initially ordered by distance from the mean. Right: the ordering is finalized by starting from the second point and working in a right-handed manner about the vector describing the position of the first point. In this case points three and four are reordered.

This method fixes the positions of the first two points because that only depends on their distance to the mean. The ordering of the third and fourth points is dependent on the chirality of the group of points. This suggests an appropriate parity operator would swap positions three and four, which would reorder the angles as shown in Eq. (14).

$$\hat{P}[\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{34}, \theta_{24}] = [\theta_{12}, \theta_{14}, \theta_{13}, \theta_{24}, \theta_{34}, \theta_{23}] \quad (14)$$

To ensure this is always the appropriate parity operator, we generated 100,000 tetrahedra, and determined their orders by the process above, giving [1, 2, 3, 4]. Parity operators in three-dimensional space were defined: $\hat{P}_x, \hat{P}_y, \hat{P}_z$, flipping in the yz, zx and xy planes respectively. Combinations of these parity operators were applied to each tetrahedral group. The four combinations that do not change the parity are $[1, \hat{P}_x \hat{P}_y, \hat{P}_y \hat{P}_z, \hat{P}_z \hat{P}_x]$ and these all leave the order as [1, 2, 3, 4]. The four parity flipping combinations are $[\hat{P}_x, \hat{P}_y, \hat{P}_z, \hat{P}_x \hat{P}_y \hat{P}_z]$ and all were found to change the order of the points to [1, 2, 4, 3]. This confirmed that the parity operator given in Eq. (14) is the correct operator in this angular description. A group of 4 points in three-dimensional space has 12 degrees of freedom—4, 3D coordinates—and after accounting for rotational and scaling invariance, this reduces to 5 d.o.f. Since a tetrahedron has 6 internal angles (6C_2), the inputs are chosen to be the first 5: $[\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{34}]$. In the 2D case the final angle, which is needed for the parity transformation, can be calculated from the others, as shown in Eq. (13). In the 3D case it is not so simple. Trying to use the first five angles gives four choices for the sixth. To overcome this issue, we created a custom data class, so that every data entry has all six internal angles, the first five of which provide the basic inputs to the net,

and the sixth, which can be accessed to generate the parity transformed inputs.

1. Testing method with toy data

Once again, the significant work in creating the net was establishing its accuracy and sensitivity. Numerous tests were carried out, with the results listed in Table IV.

The main conclusions of the tests were that the method does detect significant parity violations, as long as at least 1.2% of the inputs are net parity violating. The main drawback of the method is the difficulty of ensuring the groups of four points are sampled such that if there is parity violation, it is appropriately represented. The dataset S-PV-0.1-R3 shows that some constraints are needed, and as in the 2D case, it was shown (C-PV-0.1-R3) that constraining the scale on which groups are generated is sufficient to capture the parity violation, as long as the scales are aligned. Group C-PV-1-R3 shows that parity violations are detected where the human eye would be unable to, a powerful demonstration of this method.

E. Vector input method

To finalize the search for parity violations at the 4-point level, we construct a vector method. For this net the inputs are components of the vectors describing a group of points in their mean at origin frame. To account for the scaling invariance, all inputs are normalized to make the longest vector in the group have a length of unity. The rotational invariance of the setup is accounted for in the construction of $f(x)$:

$$f(x) = \sum_{\hat{R}_i} \left[g(\hat{R}_i x) - \sum_i g(\hat{R}_i \hat{P}_i x) + \frac{1}{2} \sum_{i \neq j} g(\hat{R}_i \hat{P}_i \hat{P}_j x) - g(\hat{R}_i \hat{P}_x \hat{P}_y \hat{P}_z x) \right] \quad (15)$$

where $\hat{R}_i \in [\hat{R}_0, \hat{R}_{90}, \hat{R}_{180}, \hat{R}_{270}]$. This ensures that every orientation of each handedness is equally represented.

1. Testing method with toy data

Because this also considers 3D distributions, we used the same tests as those in Sec. IV D 1 on the 3D angle method. The results are listed in Table V and show the efficacy and suitability of this method. As with the previous geometric methods, the only way to detect parity violations is to issue a constraint on the scale of the testing groups. It is important to note that the deviations from 50% are much smaller in this case, because of the sampling over 3D, thus significance is shown using σ -levels in Sec. V.

F. Summary of sensitivities

After extensively testing each method, the net fraction of the inputs required to contain parity violating objects

TABLE IV. 3D toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Description	Purpose	Detection	Control
R3	Random generated 3D point data	Check parity violation is not detected where there is none	$49.90 \pm 0.11\%$	$50.11 \pm 0.10\%$
PV-0.1-R3	R3 with three extra points generated near each R3 point, these are generated in a way that all groups of four are of the same chirality. Groups of four are separated on the scale of one tenth the average inter point distance	Check detection of parity violating inputs	$87.59 \pm 0.11\%$	$49.69 \pm 0.07\%$
S-PV-0.1-R3	PV-0.1-R3, but after generating parity violating groups all points shuffled together and new groups drawn at random	Can parity violations be detected when they are visibly present, and groups sampled randomly from the dataset	$50.37 \pm 0.08\%$	$50.15 \pm 0.08\%$
C-PV-0.1-R3	S-PV-0.1-R3, with a distance constraint when generating groups from the shuffled data	Can parity violations become detectable by constraining the scale	$58.73 \pm 0.08\%$	$49.87 \pm 0.08\%$
C-PV-1-R3	PV-0.1-R3, but scale of parity violations now same scale as average interpoint distance. No longer visibly parity violating. Groups drawn with scale constraint	Check parity violations detectable on more realistic scales with scale constraint	$54.37 \pm 0.05\%$	$49.87 \pm 0.06\%$

in order to make a detection significant from nonzero was determined. The level for each method is outlined in Table VI. For a null detection, these are the levels above which we can conclude that parity violations are not present.

TABLE V. Results for the vector input method on the 3D toy datasets, as defined in Table XII. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Detection	Control
R3	$50.11 \pm 0.07\%$	$50.21 \pm 0.06\%$
PV-0.1-R3	$99.83 \pm 0.02\%$	$50.3 \pm 0.31\%$
S-PV-0.1-R3	$50.16 \pm 0.07\%$	$49.84 \pm 0.07\%$
C-PV-0.1-R3	$52.44 \pm 0.10\%$	$50.11 \pm 0.11\%$
C-PV-1-R3	$52.68 \pm 0.07\%$	$49.91 \pm 0.07\%$

TABLE VI. Summary of the required net fraction of the inputs containing a “parity violating” object for the network to return a fraction significantly different from random input data.

Method	Fraction to detect
2D images	$< 0.1\%$
3D images	$\sim 0.5\%$
2D angles	$\sim 1.0\%$
3D angles	$\sim 1.0\%$
3D vectors	$\sim 2.0\%$

V. GENERATING INPUTS AND RESULTS

A. Curation for 2D CNN

To generate images from the BOSS catalog the method outlined in Sec. IVA was used. 12,000 images were sampled, split into 2,400 testing and 9,600 training images, each covering a $0.5^\circ \times 0.5^\circ$ patch of the sky, a sample is shown in Fig. 12. This size was chosen to create a large enough image dataset that suitably covered the full field of galaxy data. Smaller images also offset issues with

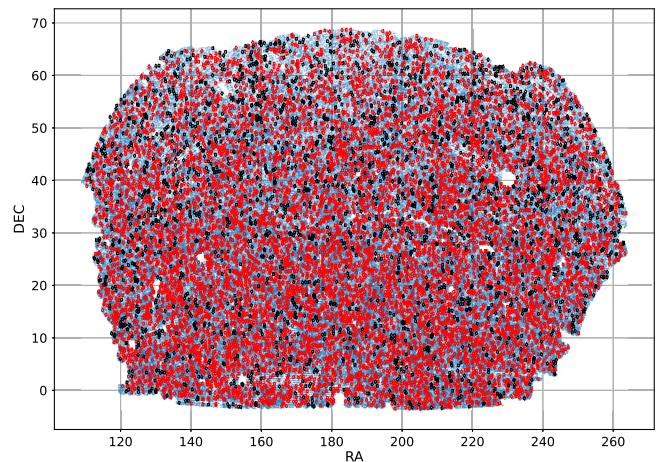


FIG. 12. A sample of images taken from CMASS. The black squares show testing images, and the red squares show training images.

TABLE VII. The average deviation from parity even of the CMASS NGC catalog for the different methods. Results for the control groups and random catalog are also shown.

Method	Detection	Control	Random
2D images	$50.4 \pm 0.3\%$	$50.1 \pm 0.3\%$	$50.8 \pm 0.3\%$
Color images	$49.7 \pm 0.3\%$	$49.4 \pm 0.3\%$	$49.3 \pm 0.3\%$
2D angles, $n = 3$	$50.15 \pm 0.09\%$	$50.02 \pm 0.09\%$	$50.35 \pm 0.10\%$
2D angles, $n = 4$	$50.05 \pm 0.06\%$	$49.85 \pm 0.06\%$	$50.20 \pm 0.06\%$
2D angles, $n = 5$	$50.30 \pm 0.06\%$	$50.27 \pm 0.06\%$	$49.65 \pm 0.06\%$
3D angles	$49.91 \pm 0.06\%$	$50.09 \pm 0.05\%$	$49.84 \pm 0.05\%$
3D vectors	$50.08 \pm 0.03\%$	$50.09 \pm 0.03\%$	$50.13 \pm 0.03\%$

approximating a spherical surface onto a flat projection and reduce overlap of images onto survey artifact regions. A full list of tests on different sizes can be found in Appendix A.

B. Curation for 2D CNN with color

To generate the 3D images, the method outlined in Sec. IV B was used. Each voxel of the space had a side length of 0.5° and a redshift depth of 0.33. Since there is no approximation of the 3D distribution, there is no constraint on the upper size of the images, and so varying image sizes were tested.

C. Curation for geometric nets

For the 2D angle method, groups were constructed from the 2D projection of the data, with points in a group constrained to be within 1° of each other. Groups were constructed on the 3, 4, and 5 point levels. In the 3D case, for both the angle and vector input methods, groups were constructed to contain 4 points, allowing for direct comparison to the 4PCF method. These groups were drawn from the CMASS catalog, with the constraint that the points must be separated by a distance in the range of $20h^{-1}$ Mpc to $160h^{-1}$ Mpc. In all cases datasets of 500,000 data groups were constructed; 400,000 for training and 100,000 for testing.

TABLE VIII. σ -level significances calculated for both (i) the difference between the CMASS data and the CMASS controls, and (ii) the difference between the CMASS data and the BOSS random catalog data. No significant parity violation is found.

Method	CMASS-control	CMASS-random
2D images	1.00	1.33
Color images	1.00	1.33
2D angles, $n = 3$	1.44	2.00
2D angles, $n = 4$	1.67	2.50
2D angles, $n = 5$	0.500	0.833
3D angles	0.00	1.00
3D vectors	0.333	1.67

D. Results

The average level of detection from bootstrapping over the testing sets and their controls for each method are given in Table VII. We also include results from running over the random catalog provided by BOSS. For every method, no significant deviation from the control or random data is detected. This is confirmed in the σ levels displayed in Table VIII. We do not find evidence of parity violation.

VI. DISCUSSION

The results suggest there is no parity violation to be found in the LSS represented in the NGC of the CMASS catalog. For the 2D projections, parity violation is not expected, because compression of a three-dimensional dataset down to two dimensions removes information about the three-dimensional distribution.

For the three-dimensional cases, the results obtained are at odds with the work of Philcox [18] and Hou *et al.* [17]. Recent analyses which account for the 8PCF bias term, have also disputed the detection of parity violation [23]. During a recent talk at the Royal Society [48], Philcox announced results that cast doubt on the parity violation detection [16,18], suggesting that sample variance could be the cause of the detection [49].

An important consideration is the “artifacts” of the dataset. The 2D projection in Fig. 2 displays notable sample artifacts, which could interfere with the analyses. In all methods, the weights applied Sec. III act to account for sampling issues, reducing the impact. For the CNN, images were constructed to further offset the impact: sampling a large enough set of small enough patches minimized overlap with artifact regions, displayed in Fig. 13. That the overall impact is negligible is perhaps best displayed by the null results in all cases. Given that no parity violation is detected, if the artifacts were interfering, it would have to be in a way that perfectly cancels with some underlying parity violation of the dataset. Since no detection is made in either 2D or 3D, the probability of this being the case is negligible. To further the rigorosity of these claims, the analyses were performed over the random catalog of the BOSS dataset. This catalog is constructed

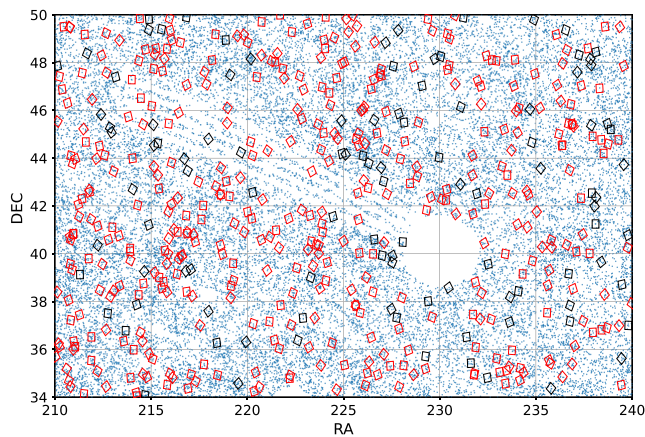


FIG. 13. The set of images generated from points near the large artifact between (210, 35) and (240, 50). Few images are affected by the presence of the artifact, and those which are affected have small overlaps into the affected region.

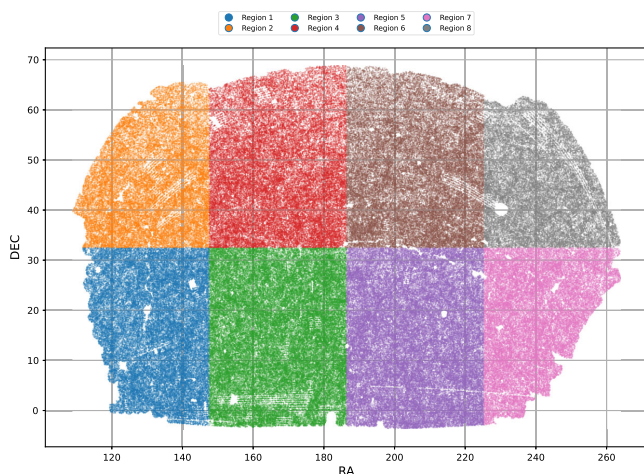


FIG. 14. The CMASS catalog split into 8 regions, each of which was investigated separately. Each region showed the same level of parity violation, as shown in Appendix D. This further suggests that the artifacts are not skewing the results.

to match the sampling geometry of the BOSS data, but is ~ 50 times the size. To resolve this size mismatch, the random data was split into 50 chunks, 10 of which were used for the analysis. The results given in Table VIII show that sampling geometry indeed introduces no parity violation. To make a final confirmation, the sky was split into 8 bins by position, displayed in Fig. 14, and inspired by Hou *et al.* [17]. Treating each of these separately, no significant difference was found between regions (Appendix D). Since every region is differently affected by artifacts, this suggests they have no noticeable impact on results.

A final important consideration is to note that the errors quoted for each dataset represent the model variance arising from our network and the resampling process. Cosmic variance from the single realization of the

Universe will contribute an extra source of uncertainty to the BOSS catalog, such that even in the limit of infinite samples of the data, some underlying nonzero baseline variance would remain.

Given the sensitivity of the methods, and the issues with searching on different scales, the CNN seems to be the most suitable method for these analyses. Its main advantage is that each image contains the full distribution for a region of space, making it the most general search method for parity violations across scales. In this work we proposed and demonstrated the effectiveness of conveying three-dimensional information through color creating a cCNN. We do not claim that this is the most suitable method, but the success of our test cases provides evidence to show the power of this method. To further this pursuit, convolutions could be extended into three dimensions; extensive further work would be required to ensure the model would correctly detect chirality. A possible implementation of this three-dimensional structure could be layers of two-dimensional images, stacked into three-dimensional voxel inputs, similar to the application in classification of medical images [50,51].

VII. CONCLUSION

We used machine learning methods to search for parity violations in the large-scale structure of the Universe. Five variations of the method were employed, two investigating the two-dimensional projection of the data, and three investigating the three-dimensional distribution. It was shown that all of these methods can detect chiral datasets down to significance fractions outlined in Table VI. Over various samples from the BOSS CMASS catalog, no parity violation was detected in any of these methods. As such, we cannot provide further support for the recent claims by Philcox *et al.* [16] and Hou *et al.* [17]. To consolidate the work outlined in this paper, the CMASS SGC and LOWZ catalogs should be explored. Uncertainties in the fiducial cosmology used (Sec. III) would alter the spatial distribution in three dimensions, and so the implications of these uncertainties should be investigated. Additionally, it would be useful to directly compare these methods to the 4PCF, and see how each performs against mock parity violating datasets. Further work should also be made in creating a three-dimensional CNN architecture to better explore local regions for parity violations across all orders in three dimensions. All of these methods should be applied to upcoming DESI, EUCLID, and Roman data releases.

ACKNOWLEDGMENTS

S. H. thanks the Cavendish Laboratory for the project opportunity. S. H. also thanks St. John's College and the University of Cambridge for funding. W. H. was supported by a Royal Society University Research Fellowship.

TABLE IX. Test results on different size image patches. On each size, tests were run on the real CMASS data and randomly generated data.

Side length	Average points	Detection	Random detection
$l = 0.3$	8.0	50.8	51.2
$l = 0.5$	22.3	50.1	50.6
$l = 0.7$	43.7	51.1	50.5
$l = 1.0$	89.2	50.1	49.3

DATA AVAILABILITY

The data that support the findings of this article are openly available [52,53].

APPENDIX A: IMAGE SIZE AND TYPE TESTS

The work carried out by James Gibbon [39] demonstrated that using different image types has no effect on the function of the net, JPEG and PNG images could both be used equally. We chose to use PNG images because space-saving was not an issue, and the higher quality would make the pngs easier to work with.

When investigating the size, the primary considerations were:

- (1) Does each image contain enough points to be able to detect parity violation on real scales. We required $3 < \langle N \rangle$, since we need at least 3 points to create a chiral object in 2D.
- (2) Are images small enough to get a large sample of images, where training and testing images do not overlap, and each set suitably represents the data.
- (3) Are images small enough to minimize effects of approximating 3D distribution to 2D plane.

These considerations led to square side lengths in the range $0.3 < l < 1.0$. A few sizes were tested on real and random data, the results of which are given in Table IX. No method had a specific advantage in testing, so we chose to use 0.5° square images because while larger point clouds could be used, 22.3 points is a reasonable level to look for parity violations while minimizing any effects of projecting onto 2D space.

APPENDIX B: IMPORTANT TEST RESULTS

The following tables list more fully the important results from the validation of the methods. Table X gives the results from the cCNN; Table XI the results from the 2D angle method; Table XII results from the 3D angle method; and Table XIII the results from the 3D vector method.

TABLE X. A summary of the important color image based toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection; the control column shows the positive fraction for the control group and the uncolored column shows the detection for the same images colored monochrome.

Data code	Description	Purpose	Detection	Control	Uncolored
RCI	Random generated image data	Check sampling method is not introducing parity violation	$49.8 \pm 0.3\%$	$50.3 \pm 0.3\%$	$49.9 \pm 0.3\%$
HE-CI	Field of left-handed helices	Check detection of complex parity violation	$97.6 \pm 0.3\%$	$50.5 \pm 0.3\%$	$62.8 \pm 0.3\%$
F-HE-CI	HE-CI, but each helix 50% chance of being left- or right-handed	Check method does not detect where there is none	$51.0 \pm 0.3\%$	$50.6 \pm 0.3\%$	$50.5 \pm 0.3\%$
HE-RCI	RCI with 1 left-handed helix inserted into the dataset for every 100 random points	Check detection of complex parity violation	$57.5 \pm 0.3\%$	$50.2 \pm 0.3\%$	$49.8 \pm 0.3\%$
F-HE-RCI	HE-RCI, but each helix has a 50% chance of being left- or right-handed	Check method of insertion is not causing spurious detection	$50.5 \pm 0.3\%$	$50.4 \pm 0.3\%$	$50.2 \pm 0.3\%$
PV-RCI	Parity violating groups of 4 points randomly oriented and distributed. Local separation \ll average interpoint distance.	Can parity violations be detected where the 2D dataset does not detect.	$87.2 \pm 0.1\%$	$49.6 \pm 0.3\%$	$50.8 \pm 0.3\%$
PV-RCI-W	PV-RCI with wider field of view, each image contains ~ 5 groups of 4.	Check detection still possible when parity violating groups have very similar pixel colors	$67.0 \pm 0.1\%$	$50.2 \pm 0.3\%$	$50.1 \pm 0.3\%$
CHE-RCI	HE-RCI, but now each helix tapers in, giving it an orientation as well as a handedness	Check detection of more complex parity violation	$58.9 \pm 0.3\%$	$49.8 \pm 0.3\%$	$49.6 \pm 0.3\%$

(Table continued)

TABLE X. (Continued)

Data code	Description	Purpose	Detection	Control	Uncolored
F-CHE-RCI	CHE-RCI, but each helix has a 50% chance of being left- or right-handed	Check method of insertion is not causing spurious detection	$50.5 \pm 0.3\%$	$50.4 \pm 0.3\%$	$50.2 \pm 0.3\%$
HE-RCI-0.1	HE-RCI with 10 times fewer helices	Check detection of complex parity violation	$54.3 \pm 0.3\%$	$50.6 \pm 0.3\%$	$49.9 \pm 0.3\%$
HE-RCI-0.01	HE-RCI with 100 times fewer helices	Check detection of complex parity violation	$52.2 \pm 0.3\%$	$50.1 \pm 0.3\%$	$49.8 \pm 0.3\%$

TABLE XI. 2D toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Description	Purpose	Detection	Control
R2d	Random generated 2D point data	Check parity violation is not detected where there is none	$50.41 \pm 0.10\%$	$50.10 \pm 0.09\%$
PV-0.1-R2d	R2d with two extra points generated near each R2d point, these are generated in a way that all groups of three are of the same chirality. Groups of three are separated on the scale of one tenth the average interpoint distance	Check detection of parity violating inputs	$90.77 \pm 0.06\%$	$49.74 \pm 0.10\%$
Fx-PV-0.1-R2d	PV-0.1-R2d with 50% of the extra points inserted with their relative x coordinate to the original point inverted	Check method of generating points is not injecting parity violation	$49.66 \pm 0.10\%$	$49.98 \pm 0.10\%$
Fy-PV-0.1-R2d	Fx-PV-0.1-R2d flipped in y instead of x	Check method of generating points is not injecting parity violation	$49.98 \pm 0.10\%$	$49.57 \pm 0.10\%$
PV-0.1-R2d-5%	PV-0.1-R2d, but only 5% of inputs generated in this way. The rest are random	Check fraction of inputs needed to detect	$52.46 \pm 0.10\%$	$50.0 \pm 0.10\%$
PV-0.1-R2d-1%	PV-0.1-R2d, but only 1% of inputs generated in this way. The rest are random	Check fraction of inputs needed to detect	$50.96 \pm 0.10\%$	$50.0 \pm 0.10\%$
S-PV-0.1-R2d	PV-0.1-R2d, but after generating parity violating groups all points shuffled together and new groups drawn at random	Can parity violations be detected when they are visibly present, and groups sampled randomly from the dataset	$49.70 \pm 0.10\%$	$49.68 \pm 0.10\%$
C-PV-0.1-R2d	S-PV-0.1-R2d, with a distance constraint when generating groups from the shuffled data	Can parity violations become detectable by constraining the scale	$65.10 \pm 0.04\%$	$49.85 \pm 0.04\%$
C-R2d	R2d with groups constructed after constraining the scale	Check action of constraining scale is not causing spurious detection	$50.34 \pm 0.09\%$	$50.12 \pm 0.10\%$
C-PV-1-R2d	PV-0.1-R2d, but scale of parity violations now same scale as average interpoint distance. No longer visibly parity violating. Groups drawn with scale constraint	Check parity violations detectable on more realistic scales with scale constraint	$55.12 \pm 0.10\%$	$49.63 \pm 0.10\%$
C-R2d-LH5	Generate groups of three from R2d with distance in range 1-5 x average interpoint distance. Remove groups of three where $\theta_{12} < \theta_{13}$. Reshuffle and drawn new groups on same scale	Can different methods of generating parity violations be detected	$52.46 \pm 0.10\%$	$50.07 \pm 0.10\%$

TABLE XII. 3D toy datasets. Each dataset is given a code and a description and has its testing purpose outlined. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Description	Purpose	Detection	Control
R3	Random generated 3d point data	Check parity violation is not detected where there is none	$49.90 \pm 0.11\%$	$50.11 \pm 0.10\%$
PV-0.1-R3	R3 with three extra points generated near each R3 point, these are generated in a way that all groups of four are of the same chirality. Groups of four are separated on the scale of one tenth the average inter point distance	Check detection of parity violating inputs	$87.59 \pm 0.11\%$	$49.69 \pm 0.07\%$
F-PV-0.1-R3d	PV-0.1-R3d with 50% of the extra points inserted with their relative x coordinate to the original point inverted	Check method of generating points is not injecting parity violation	$50.17 \pm 0.11\%$	$50.21 \pm 0.10\%$
PV-0.1-R3-5%	PV-0.1-R3, but only 5% of inputs generated in this way. The rest are random	Check fraction of inputs needed to detect	$53.67 \pm 0.07\%$	$50.01 \pm 0.08\%$
PV-0.1-R3-1%	PV-0.1-R3, but only 1% of inputs generated in this way. The rest are random	Check fraction of inputs needed to detect	$49.37 \pm 0.08\%$	$50.05 \pm 0.07\%$
S-PV-0.1-R3	PV-0.1-R3, but after generating parity violating groups all points shuffled together and new groups drawn at random	Can parity violations be detected when they are visibly present, and groups sampled randomly from the dataset	$50.37 \pm 0.08\%$	$50.15 \pm 0.08\%$
C-PV-0.1-R3	S-PV-0.1-R3, with a distance constraint when generating groups from the shuffled data	Can parity violations become detectable by constraining the scale	$58.73 \pm 0.08\%$	$49.87 \pm 0.08\%$
C-R3	R3 with groups constructed after constraining the scale	Check action of constraining scale is not causing spurious detection	$49.87 \pm 0.04\%$	$50.32 \pm 0.07\%$
C-PV-1-R3	PV-0.1-R3, but scale of parity violations now same scale as average interpoint distance. No longer visibly parity violating. Groups drawn with scale constraint	Check parity violations detectable on more realistic scales with scale constraint	$54.37 \pm 0.05\%$	$49.87 \pm 0.06\%$
C-R3-LH5	Generate groups of 4 from R3 with distance in range 1-5x average interpoint distance. Remove groups of 4 where $\theta_{13} < \theta_{14}$. Reshuffle and drawn new groups on same scale	Can different methods of generating parity violations be detected	$53.11 \pm 0.07\%$	$50.64 \pm 0.06\%$

APPENDIX C: ADDITIONAL TESTS ON NETWORKS

1. 2D angles: $n = 4, 5$

TABLE XIII. Results for the vector input method on the 3D toy datasets, as defined in Table XII. The detection column contains the positive fraction detection, and the control column shows the positive fraction for the control group for each test.

Data code	Detection	Control
R3	$50.11 \pm 0.07\%$	$50.21 \pm 0.06\%$
PV-0.1-R3	$99.83 \pm 0.02\%$	$50.3 \pm 0.31\%$
F-PV-0.1-R3d	$50.35 \pm 0.07\%$	$49.87 \pm 0.07\%$
PV-0.1-R3-5%	$53.18 \pm 0.07\%$	$49.78 \pm 0.07\%$
PV-0.1-R3-1%	$50.43 \pm 0.07\%$	$49.91 \pm 0.07\%$
S-PV-0.1-R3	$50.16 \pm 0.07\%$	$49.84 \pm 0.07\%$
C-PV-0.1-R3	$47.56 \pm 0.10\%$	$50.11 \pm 0.11\%$
C-R3	$49.98 \pm 0.11\%$	$50.09 \pm 0.10\%$
C-PV-1-R3	$47.36 \pm 0.07\%$	$49.91 \pm 0.07\%$
C-R3-LH5	$48.12 \pm 0.07\%$	$50.04 \pm 0.07\%$

As with the method using angles from groups of 3, the methods using 4 and 5 points per group were also

TABLE XIV. 2D toy datasets for $n = 4$. The datacodes are defined in Table XI.

Data code	Detection	Control
R2d	$50.22 \pm 1.0\%$	$49.87 \pm 1.0\%$
PV-0.1-R2d	$98.32 \pm 0.9\%$	$50.91 \pm 1.0\%$
PV-0.1-R2d-1%	$52.62 \pm 1.0\%$	$50.43 \pm 1.0\%$
PV-0.1-R2d-0.5%	$51.56 \pm 0.3\%$	$50.52 \pm 1.0\%$
S-PV-0.1-R2d	$50.36 \pm 0.3\%$	$50.22 \pm 1.0\%$
C-PV-0.1-R2d	$55.57 \pm 0.3\%$	$50.22 \pm 1.0\%$
C-R2d	$50.61 \pm 0.3\%$	$50.21 \pm 1.0\%$
C-PV-1-R2d	$53.69 \pm 0.3\%$	$50.32 \pm 1.0\%$
C-R2d-LH5	$51.26 \pm 0.3\%$	$50.62 \pm 1.0\%$

TABLE XV. 2D toy datasets for $n = 5$. The datacodes are defined in Table XI.

Data code	Detection	Control
R2d	$50.24 \pm 1.0\%$	$50.17 \pm 1.0\%$
PV-0.1-R2d	$99.49 \pm 0.9\%$	$50.31 \pm 1.0\%$
PV-0.1-R2d-1%	$53.32 \pm 1.0\%$	$50.38 \pm 1.0\%$
PV-0.1-R2d-0.5%	$51.76 \pm 0.3\%$	$50.21 \pm 1.0\%$
S-PV-0.1-R2d	$50.29 \pm 0.3\%$	$50.27 \pm 1.0\%$
C-PV-0.1-R2d	$54.77 \pm 0.3\%$	$50.18 \pm 1.0\%$
C-R2d	$50.51 \pm 0.3\%$	$50.43 \pm 1.0\%$
C-PV-1-R2d	$53.72 \pm 0.3\%$	$50.23 \pm 1.0\%$
C-R2d-LH5	$51.07 \pm 0.3\%$	$50.52 \pm 1.0\%$

TABLE XVI. Levels of detection in the separate regions, used to investigate the impact of the artifacts.

Data code	Detection	Control
R1	$49.96 \pm 2.0\%$	$53.43 \pm 2.0\%$
R2	$48.95 \pm 2.0\%$	$51.33 \pm 2.0\%$
R3	$50.91 \pm 2.0\%$	$50.4 \pm 2.0\%$
R4	$51.31 \pm 2.0\%$	$50.92 \pm 2.0\%$
R5	$49.65 \pm 2.0\%$	$52.43 \pm 2.0\%$
R6	$52.46 \pm 2.0\%$	$51.04 \pm 2.0\%$
R7	$48.76 \pm 2.0\%$	$52.93 \pm 2.0\%$
R8	$49.96 \pm 2.0\%$	$53.43 \pm 2.0\%$

tested. These were tested using the same tests as the $n = 3$ case. The results are given in Tables XIV and XV. The tests demonstrate that these methods are equally suitable.

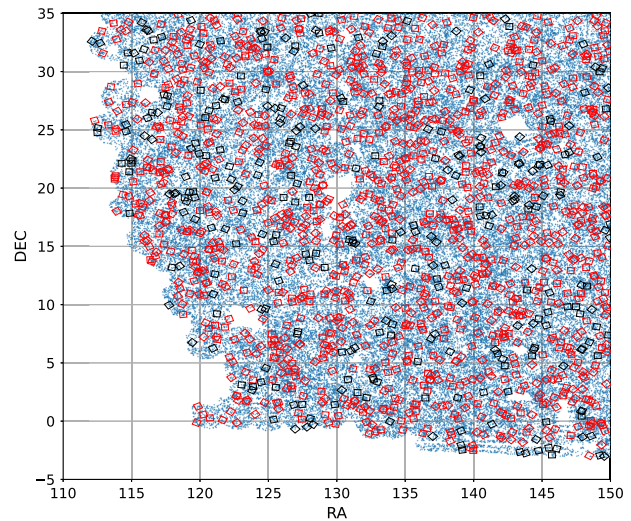


FIG. 15. The image samples drawn from the first region for localized analysis.

APPENDIX D: PATCHES OF THE SKY

In order to confirm the arguments which propose that the artifacts are having no significant impact, the CMASS NGC was split into eight regions. And image datasets compiled from each separately, the full results follow in Table XVI, where we see that no significant difference occurs between the regions. Suggesting that the artifacts are indeed having no impact. For these tests, the same size patches of the sky were used, and each region had 3,000 images generated from it. The image distribution for the first region is shown in Fig. 15.

- [1] Arthur Lue, Limin Wang, and Marc Kamionkowski, Cosmological signature of new parity-violating interactions, *Phys. Rev. Lett.* **83**, 1506 (1999).
- [2] Marc Kamionkowski and Tarun Souradeep, Odd-parity cosmic microwave background bispectrum, *Phys. Rev. D* **83**, 027301 (2011).
- [3] Maresuke Shiraishi, Parity violation in the CMB trispectrum from the scalar sector, *Phys. Rev. D* **94**, 083503 (2016).
- [4] Yuto Minami and Eiichiro Komatsu, New extraction of the cosmic birefringence from the Planck 2018 polarization data, *Phys. Rev. Lett.* **125**, 221301 (2020).
- [5] Oliver H. E. Philcox, Do the CMB temperature fluctuations conserve parity?, *arXiv:2303.12106*.
- [6] Oliver H. E. Philcox and Maresuke Shiraishi, Testing parity symmetry with the polarized cosmic microwave background, *arXiv:2308.03831*.
- [7] Shun Saito, Kiyotomo Ichiki, and Atsushi Taruya, Probing polarization states of primordial gravitational waves with cosmic microwave background anisotropies, *J. Cosmol. Astropart. Phys.* **09** (2007) 002.
- [8] Nicolás Yunes, Richard O’Shaughnessy, Benjamin J. Owen, and Stephon Alexander, Testing gravitational parity violation with coincident gravitational waves and short gamma-ray bursts, *Phys. Rev. D* **82**, 064017 (2010).
- [9] Anzhong Wang, Qiang Wu, Wen Zhao, and Tao Zhu, Polarizing primordial gravitational waves by parity violation, *Phys. Rev. D* **87**, 103512 (2013).
- [10] Giorgio Orlando, Mauro Pieroni, and Angelo Ricciardone, Measuring parity violation in the stochastic gravitational wave background with the LISA-Taiji network, *J. Cosmol. Astropart. Phys.* **03** (2021) 069.
- [11] Leah Jenks, Lyla Choi, Macarena Lagos, and Nicolás Yunes, Parameterized parity violation in gravitational wave propagation, *Phys. Rev. D* **108**, 044023 (2023).
- [12] Robert N. Cahn, Zachary Slepian, and Jiamin Hou, A test for cosmological parity violation using the 3D distribution of galaxies, *Phys. Rev. Lett.* **130**, 201002 (2023).

- [13] BOSS Collaboration, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [14] BOSS Collaboration, SDSS-III: Massive spectroscopic surveys of the distant universe, the Milky Way, and extra-solar planetary systems, *Astron. J.* **142**, 72 (2011).
- [15] BOSS Collaboration, The baryon oscillation spectroscopic survey of SDSS-III, *Astron. J.* **145**, 10 (2012).
- [16] Oliver H. E. Philcox, Jiamin Hou, and Zachary Slepian, A first detection of the connected 4-point correlation function of galaxies using the BOSS CMASS sample, [arXiv:2108.01670](https://arxiv.org/abs/2108.01670).
- [17] Jiamin Hou, Zachary Slepian, and Robert N. Cahn, Measurement of parity-odd modes in the large-scale 4-point correlation function of Sloan Digital Sky Survey baryon oscillation spectroscopic survey twelfth data release CMASS and LOWZ galaxies, *Mon. Not. R. Astron. Soc.* **522**, 5701 (2023).
- [18] Oliver H. E. Philcox, Probing parity violation with the four-point correlation function of BOSS galaxies, *Phys. Rev. D* **106**, 063501 (2022).
- [19] Peter L. Taylor, Matthew Craigie, and Yuan-Sen Ting, Unsupervised searches for cosmological parity-violation: An investigation with convolutional neural networks, *Phys. Rev. D* **109**, 083518 (2024).
- [20] BOSS Collaboration, The clustering of galaxies in the SDSS-III baryon oscillation spectroscopic survey: Modelling the clustering and halo occupation distribution of BOSS CMASS galaxies in the final data release, *Mon. Not. R. Astron. Soc.* **460**, 1173 (2016).
- [21] Francisco-Shu Kitaura *et al.*, The clustering of galaxies in the SDSS-III baryon oscillation spectroscopic survey: Mock galaxy catalogs for the BOSS final data release, *Mon. Not. R. Astron. Soc.* **456**, 4156 (2016).
- [22] Jiamin Hou, Robert N. Cahn, Oliver H. E. Philcox, and Zachary Slepian, Analytic Gaussian covariance matrices for galaxy N -point correlation functions, *Phys. Rev. D* **106**, 043515 (2022).
- [23] Alex Krolewski, Simon May, Kendrick Smith, and Hans Hopkins, No evidence for parity violation in BOSS, *J. Cosmol. Astropart. Phys.* **08** (2024) 044.
- [24] Matthew Craigie, Peter L. Taylor, Yuan-Sen Ting, Carolina Cuesta-Lazaro, Rossana Ruggeri, and Tamara M. Davis, Unsupervised searches for cosmological parity violation: Improving detection power with the neural field scattering transform, [arXiv:2405.13083](https://arxiv.org/abs/2405.13083).
- [25] DESI Collaboration, The DESI experiment part I: Science, targeting, and survey design, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036).
- [26] DESI Collaboration, The early data release of the dark energy spectroscopic instrument, (2023), <https://zenodo.org/record/7964161>.
- [27] Euclid Collaboration, Euclid definition study report, [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [28] Euclid Collaboration, Euclid preparation: VII. Forecast validation for Euclid cosmological probes, *Astron. Astrophys.* **642**, A191 (2020).
- [29] D. Spergel *et al.*, Wide-field infrared survey telescope-astronomy focused telescope assets WFIRST-AFTA 2015 report, [arXiv:1503.03757](https://arxiv.org/abs/1503.03757).
- [30] WFIRST Collaboration, The wide field infrared survey telescope: 100 Hubbles for the 2020s, [arXiv:1902.05569](https://arxiv.org/abs/1902.05569).
- [31] T. D. Lee and C. N. Yang, Question of parity conservation in weak interactions, *Phys. Rev.* **104**, 254 (1956).
- [32] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Experimental test of parity conservation in beta decay, *Phys. Rev.* **105**, 1413 (1957).
- [33] A. D. Sakharov, Violation of *CP* invariance, *C* asymmetry, and baryon asymmetry of the universe, *Pis'ma Zh. Eksp. Teor. Fiz.* **5**, 32 (1967).
- [34] Sean M. Carroll, Quintessence and the rest of the world: Suppressing long-range interactions, *Phys. Rev. Lett.* **81**, 3067 (1998).
- [35] D. Hilbert, Die grundlagen der physik. (erste mitteilung), *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.* **1915**, 395 (1915).
- [36] Christopher G. Lester and Matthias Schott, Testing non-standard sources of parity violation in jets at the LHC, trialled with CMS open data, *J. High Energy Phys.* **12** (2019) 120.
- [37] Christopher G. Lester and Rupert Tombs, Using unsupervised learning to detect broken symmetries, with relevance to searches for parity violation in nature (previously: Stressed GANs snag desserts), (2022), <https://openreview.net/forum?id=QFJ3gtbWHR>.
- [38] Rupert Tombs and Christopher G. Lester, A method to challenge symmetries in data with self-supervised learning, *J. Instrum.* **17**, P08024 (2022).
- [39] Christopher G. Lester and James Gibbon, Detection of parity violation in snails through unsupervised learning, (2023), <https://www.hep.phy.cam.ac.uk/~lester/teaching/PartIIIProjects/2023-James-Gibbon-Snails.pdf>.
- [40] BOSS Collaboration, SDSS-III Baryon Oscillation Spectroscopic Survey data release 12: Galaxy target selection and large-scale structure catalogues, *Mon. Not. R. Astron. Soc.* **455**, 1553 (2016).
- [41] BOSS Collaboration, Stellar masses of SDSS-III/BOSS galaxies at $z \sim 0.5$ and constraints to galaxy formation models, *Mon. Not. R. Astron. Soc.* **435**, 2764 (2013).
- [42] Shun Saito, Alexie Leauthaud, Andrew P. Hearin, Kevin Bundy, Andrew R. Zentner, Peter S. Behroozi, Beth A. Reid, Manodeep Sinha, Jean Coupon, Jeremy L. Tinker, Martin White, and Donald P. Schneider, Connecting massive galaxies to dark matter haloes in BOSS—I. Is galaxy colour a stochastic process in high-mass haloes?, *Mon. Not. R. Astron. Soc.* **460**, 1457 (2016).
- [43] Alexie Leauthaud, Kevin Bundy, Shun Saito, Jeremy Tinker, Claudia Maraston, Rita Tojeiro, Song Huang, Joel R. Brownstein, Donald P. Schneider, and Daniel Thomas, The stripe 82 massive galaxy project—II. Stellar mass completeness of spectroscopic galaxy samples from the Baryon Oscillation Spectroscopic Survey, *Mon. Not. R. Astron. Soc.* **457**, 4021 (2016).
- [44] Kevin Bundy, Alexie Leauthaud, Shun Saito, Claudia Maraston, David A. Wake, and Daniel Thomas, The stripe 82 massive galaxy project. III. A lack of growth among massive galaxies, *Astrophys. J.* **851**, 34 (2017).
- [45] BOSS Collaboration, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic

- Survey: Cosmological analysis of the dr12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [46] L. Anderson *et al.*, The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon acoustic oscillations in the data release 9 spectroscopic galaxy sample, *Mon. Not. R. Astron. Soc.* **427**, 3435 (2012).
- [47] Hume A. Feldman, Nick Kaiser, and John A. Peacock, Power-spectrum analysis of three-dimensional redshift surveys, *Astrophys. J.* **426**, 23 (1994).
- [48] Royal Society Events, Royal Society Events challenging the standard cosmological model, (2024), <https://royalsociety.org/science-events-and-lectures/2024/04/cosmological-model/>.
- [49] O. H. E. Philcox and J. Ereza, Could sample variance be responsible for the parity-violating signal seen in the Baryon Oscillation Spectroscopic Survey?, *Philos. T. R. Soc. A* **383**, 20240034 (2025).
- [50] Justin Ker, Satya P. Singh, Yeqi Bai, Jai Rao, Tchoyoson Lim, and Lipo Wang, Image thresholding improves 3-dimensional convolutional neural network diagnosis of different acute brain hemorrhages on computed tomography scans, *Sensors* **19**, 2167 (2019).
- [51] Satya P. Singh, Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan, and Balázs Gulyás, 3D deep learning on medical images: A review, *Sensors* **20**, 5097 (2020).
- [52] BOSS Collaboration, SDSS BOSS dr12, <https://data.sdss.org/sas/dr12/booss/lss/>.
- [53] Samuel Hewson, Unsupervised searches for parity violation (2025), <https://github.com/sh2099/parity-violation>.