Quantifying Radio Source Morphology

Lachlan J. Barnes,[1] Andrew M. Hopkins,[1] Lawrence Rudnick,[2] Heinz Andernach,[3] Michael Cowley,[4,5] Nikhel Gupta,[6] Ray P. Norris,[7,8] Stanislav S. Shabala,[9] and Tayyaba Zafar[1]

[1]School of Mathematical and Physical Sciences, 12 Wally's Walk, Macquarie University, NSW 2109, Australia
[2]Minnesota Institute for Astrophysics, 116 Church St. SE, Minneapolis, MN 55455, USA
[3]Depto. de Astronomía, DCNE, Univ. de Guanajuato, Callejón de Jalisco s/n, Guanajuato, CP 36023, GTO, Mexico
[4]School of Chemistry & Physics, Faculty of Science, Queensland University of Technology, Brisbane, QLD 4000, Australia
[5]University of Southern Queensland, Centre for Astrophysics, West Street, Toowoomba QLD 4350, Australia
[6]CSIRO Space & Astronomy, Bentley, WA, Australia
[7]ATNF, CSIRO Space & Astronomy, P.O. Box 76, Epping, NSW 1710, Australia
[8]Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
[9]School of Natural Sciences, University of Tasmania, Private Bag 37, Hobart, TAS 7001, Australia
Author for correspondence: Lachlan Barnes, Email: lachlan.barnes3@hdr.mq.edu.au.

Abstract

The advent of next-generation telescope facilities brings with it an unprecedented amount of data, and the demand for effective tools to process and classify this information has become increasingly important. This work proposes a novel approach to quantify the radio galaxy morphology, through the development of a series of algorithmic metrics that can quantitatively describe the structure of radio source, and can be applied to radio images in an automatic way. These metrics are intuitive in nature and are inspired by the intrinsic structural differences observed between the existing Fanaroff-Riley (FR) morphology types. The metrics are defined in categories of asymmetry, blurriness, concentration, disorder, and elongation (*ABCDE*/single-lobe metrics), as well as the asymmetry and angle between lobes (source metrics). We apply these metrics to a sample of 480 sources from the Evolutionary Map of the Universe Pilot Survey (EMU-PS) and 72 well resolved extensively studied sources from An Atlas of DRAGNs, a subset of the revised Third Cambridge Catalogue of Radio Sources (3CRR). We find that these metrics are relatively robust to resolution changes, independent of each other, and measure fundamentally different structural components of radio galaxy lobes. These metrics work particularly well for sources with reasonable signal-to-noise and well separated lobes. We also find that we can recover the original FR classification using probabilistic combinations of our metrics, highlighting the usefulness of our approach for future large data sets from radio sky surveys.

Keywords: radio continuum: galaxies, galaxies: jets, galaxies: structure

1. Introduction

1.1 Radio Galaxies and Morphology

Supermassive black holes (SMBH) at the core of galaxies, surrounded by an accretion disk of hot gas (e.g., Wilson & Colbert, 1995; Best & Heckman, 2012; Kormendy & Ho, 2013; Morganti, 2019) are referred to as active galactic nuclei (AGN). AGN sources with radio emission in the form of symmetric jets and lobes are known as DRAGNs (double radio sources associated with a galactic nucleus), a term coined by Leahy (1993). DRAGNs (also referred to as 'radio galaxies' in this work) exhibit a wide range of structural features and have been classified into distinct morphological types (e.g., Miley, 1980).

Fanaroff & Riley (1974) first classified DRAGNs into two categories, based on the ratio of a/b (sometimes referred to as the FR ratio, e.g., Brand et al., 2023). Here a represents the distance between brightest spots in opposite lobes, and b is the total extent of the radio emission. Objects with $a/b < 0.5$ were classified as class 1 (FRI) and those with $a/b > 0.5$ classified as class 2 (FRII). FRI and FRII type sources appear structurally different. FRI sources tend to have initially collimated jets that decelerate and flare into turbulent, edge-dimmed lobes (e.g., Laing & Bridle, 2014; Hardcastle & Croston, 2020; Saikia,

2022). FRII sources are believed to have highly collimated, relativistic jets that remain narrow over large distances and terminate in bright hotspots at the edges of the lobes, making them edge-brightened (e.g., Shabala & Alexander, 2009; Proctor, 2011; Saripalli et al., 2012; Rossi et al., 2020; Ndung'u et al., 2024).

The difference between structural features of the two FR types is likely due to both jet power and the radio source environment density (e.g., Best, 2009; Krause et al., 2012; Saripalli et al., 2012; Kapińska et al., 2017; Yates-Jones et al., 2023). FRI sources tend to be hosted by massive ellipticals in galaxy clusters (i.e., a high density environment), whereas FRII sources tend to be hosted by lower mass galaxies, and are often found in less dense environments (Morganti, 2019; Shabala et al., 2020). Due to the higher environmental density and lower power, the jets of FRI sources decelerate and dissipate on tens of kpc scales (e.g., Laing & Bridle, 2002; Laing & Bridle, 2014), whereas those of FRII sources can remain relativistic for hundreds of kpc (e.g., Mingo et al., 2019; Rossi et al., 2020; Gordon et al., 2023). While the original classification indicated that FRII sources tend to have higher luminosities than FRI sources, later work at higher frequencies found there was significant

overlap in luminosity for the two morphologies (e.g., Best, 2009; Gupta et al., 2024b). In addition to luminosity overlap, FRI and FRII sources can also share structural features, making it difficult to precisely determine the morphology of DRAGNs. Low flux density (faint) or sources at different redshifts and environments, can influence the observed morphology leading to hybrid sources (HyMoRS; Kapińska et al., 2017), with one lobe showing FRI-like morphology and the other FRII-like (Best, 2009; Bastien et al., 2021). Further, Fanaroff & Riley (1974) exclude the central source (core) in their classification. Some authors do not, and this can lead to a core-dominant FRII source being misclassified as an FRI (Norris et al., 2025).

The development of next-generation radio telescope facilities, driven by the vision of the Square Kilometre Array (SKA) since the early 1990s (Carilli & Rawlings, 2004), has led to further advancements in radio telescope technology (such as improved sensitivity, resolution, and survey efficiency). The development of the SKA Pathfinders and their large-scale surveys has resulted in an unprecedented influx of radio data. This has created a growing demand for efficient and effective tools to process and classify radio sources. However, the wide range of morphologies present in radio galaxies makes this a significant challenge for automatic classification.

1.2 Machine Learning

Machine learning (ML) techniques for a wide variety of classification problems have been gradually incorporated into astronomical applications (Fayyad et al., 1995). For example, ML has been used for learning and classifying the morphology of different radio objects (e.g., Ma et al., 2019; Wu et al., 2019; Galvin et al., 2020; Gupta et al., 2022, 2023; Alam et al., 2025), simulating radio sources (e.g., Bastien et al., 2021; Andrianomena & Tang, 2024) and the development of radio source finding tools (e.g., Riggi et al., 2023). While ML is extremely capable at handling vast amounts of data, the process of training and testing in many ML approaches is computationally intensive (see Ball & Brunner, 2010), and the results can be subject to overfitting and generalisation issues (Tolley, 2024). Further, many of the existing ML models are trained on radio images from a selection of surveys with varying resolution and signal-to-noise which may not be well-matched to the datasets they are later applied to (see Ndung'u et al., 2023; Riggi et al., 2023). The lack of large annotated training datasets remains one of the greatest challenges in assessing and improving the overall performance of ML image classification algorithms (Becker et al., 2021). Automating the detection and annotation of radio sources is a separate challenge, although recent work (Gupta et al., 2023, 2024a) introduces a novel ML tool which appears to reliably address this aspect of the problem.

1.3 Feature Extraction

Algorithmic feature extraction from images can be faster and more computationally efficient, when compared to many ML approaches, and has been used for image classification across many disciplines for several decades (e.g., Xu et al., 2023). Within astronomy, it has been shown to be a robust approach

for classifying the morphology of optical galaxies (e.g., Kent, 1985; Abraham et al., 1994; Bershady et al., 2000). For example, the development of the CAS parameters (Concentration, Asymmetry, and Smoothness, Conselice et al., 2000; Conselice, 2003) and the Gini and M20 parameters (Lotz et al., 2004) have been used extensively to distinguish between early- and late-type galaxies, and have been modified in recent years to be more effective (see Ferrari et al., 2015). A number of different tools have also been developed to apply algorithmic feature extraction for the purpose of classifying classes of objects from images. Some examples are WND-CHARM (Weighted Neighbour Distances using a Compound Hierarchy of Algorithms Representing Morphology, Orlov et al., 2008; Shamir, 2009) and COSFIRE filters (Combination Of Shifted FIlter REsponses, Azzopardi & Petkov, 2013; Ndung'u et al., 2024).

While it is evident that feature extraction methods are powerful tools in the classification and analysis of galaxy morphology at optical wavelengths, their application to radio astronomy has only begun in recent years. Feature extraction techniques have been used to measure the coarse-grained complexity of complex and analogous radio sources (e.g., Segal et al., 2019, 2023), and directly analyse radio galaxy features such as FR classification, hotspot brightness (e.g., Becker & Grobler, 2019), size, eccentricity, orientation, and symmetry of a radio galaxy (e.g., Javaherian et al., 2023). Other works have applied existing feature based metrics to radio galaxies. For example, Sadeghi et al. (2021) incorporated rotation, translation, and scale invariant image moments, based on Zernike polynomials (Teague, 1980) and Ntwaetsile & Geach (2021) used Haralick features (Haralick et al., 1973) to characterise the morphology of FRI and FRII type sources. Brand et al. (2023) extracted the FR ratio, if a galaxy is bent or not, the number of bright spots present in the galaxy, and the ratio of the size of bright spots to the total size of the galaxy to aid the training of a CNN in the classification of radio galaxy morphology. Although still incorporating aspects of ML approaches for radio source morphology classification, these studies highlight the potential of algorithmic feature extraction techniques to classify and quantify radio source morphology.

The motivation for this work stems from the need for efficient and scalable methods to classify complex extended radio sources. Algorithmic feature extraction approaches have demonstrated strong potential, and are likely to serve as a valuable method for automating radio galaxy classification. We propose a novel approach to the classification of such sources through the development of a series of algorithmic metrics that can be applied to any radio image. These metrics are designed to quantify key structural features of radio galaxy lobes, such as the concentration of flux or the asymmetry. We seek to establish a foundation for a robust classification framework that can accommodate future radio surveys and higher-resolution observational datasets. Our approach is adaptable to evolving interpretations of radio galaxy morphology. Unlike ML models, which require retraining when new types of sources emerge, our metrics remain agnostic to specific classification schemes and can be applied to any radio galaxy image in an automatic way.

2. Data and Pre-processing

2.1 EMU and RadioGalaxyNET

The Evolutionary Map of the Universe (EMU; Norris et al., 2011; Hopkins et al., 2025) is a wide-field radio continuum survey, using the ASKAP telescope. The primary goal of EMU is to make a deep (~ 900 MHz) radio continuum survey of the entire Southern sky at a resolution of $\sim 15''$ and sensitivity of ~ 25 μJy/beam. It is anticipated that EMU will detect and catalogue tens of millions sources, many of which will be extended and complex. A key motivation for this work is the analysis of large numbers of radio sources, for which we start by drawing from EMU data. In preparation for the full EMU survey (Norris et al., 2011), the EMU Pilot Survey (EMU-PS; Norris et al., 2021) was conducted to test the survey strategy and data processing pipeline. Phase 1 of the EMU-PS was limited to a contiguous region of 270 deg^2 within the Dark Energy Survey region (DES; Abbott et al., 2018), lying at a declination $< -30°$ and Galactic latitude of $> +20°$, with a sensitivity of 25–35 μJy/beam at 944 MHz.

RadioGalaxyNET (Gupta et al., 2024a) is developed from a curated dataset of DRAGNs classified by professional astronomers, designed for training machine learning models in radio morphology recognition. Because of the manual nature of identifying the original dataset, the DRAGNs used will tend to favour higher signal-to-noise systems, although this is not a fixed or quantitative threshold. It simply arises since very low surface brightness objects are harder to visually identify. The full RadioGalaxyNET dataset contains 2800 EMU-PS $15' \times 15'$ cutouts in three channels (unprocessed radio, processed radio, and infrared), resulting in a total of 4155 annotated sources, since some cutouts contain multiple sources. The annotations consist of radio morphological class information, bounding box details for associated components of each radio galaxy, radio galaxy segmentation masks, and the host galaxy positions from the infrared images. This work primarily analyses a subset of the sources from the RadioGalaxyNET data.

In order to ensure a source is sufficiently resolved and contains enough observable structure for our analysis, we define a minimum area threshold of the radio cutouts of 20 times the ASKAP synthesised beam size (1 beam $\approx 13'' \times 11''$; Norris et al., 2021). The minimum area for each cutout is therefore 2860 arcsec^2 or ~ 715 pixels (1 pixel ≈ 2 arcsec in the EMU-PS mosaic). Below this value the analysis of the radio lobes can become compromised by the limited number of beams (resolution elements). To produce the cutout images from the EMU-PS, we first extracted the coordinates of the bounding boxes from the RadioGalaxyNET data. These coordinates were transformed to match the orientation of the EMU-PS mosaic, from which we directly produce the cutouts of the individual radio galaxies. There are 595 sources from RadioGalaxyNET that met the minimum size requirements, with an average cutout size of ~ 3600 pixels. The cutouts were made on the 'unsmooth' EMU-PS mosaic, in order to maximise the resolution in different regions.

2.2 3CRR and An Atlas of DRAGNs

The Third Cambridge Catalogue of Radio Sources (3C; Edge et al., 1959; Archer et al., 1959) is a radio catalogue originally observed at 159 MHz, then revised at 178 MHz (3CR; Bennett & Simth, 1962; Bennett, 1962). The 3CR catalogue contained 328 sources and was one of the first comprehensive lists of extragalactic radio sources. However, the original survey had relatively limited angular resolution, leading to source blending and other biases (Véron, 1977). To improve upon this, a further revision of 173 bright radio sources at 178 MHz was conducted (3CRR; Laing et al., 1983). While the initial 3CR selection suffered from resolution limitations, 3CRR sources have since been observed in much greater detail and resolution. The 3CRR catalogue contains all extragalactic radio sources in the Northern Hemisphere with a 178 MHz flux density greater than 10.9 Jy.

An Atlas of DRAGNs (Leahy et al., 2013) is a subset of the 3CRR catalogue, consisting of the nearest 85 DRAGNs. The Atlas of DRAGNs website[a] contains high-quality images of the DRAGNs compiled from different studies. The images in the atlas have high resolution, and sufficient sensitivity and spatial frequency coverage to clearly and accurately image the faintest regions of a radio source. Most images are at frequencies near 1.5 GHz, although the frequencies range from 0.3 to 8.4 GHz in the atlas. The difference in resolution between EMU-PS and 3CRR sources from this atlas is highlighted in Figure 1. There is also a dedicated page providing supplementary information about each radio source, containing FR classification, redshift, radio power, and linear size. Given the well-documented nature of these sources, we use this atlas as a reference dataset in this work.

Example EMU source   Example 3CRR source

158"x158"   256"x256"
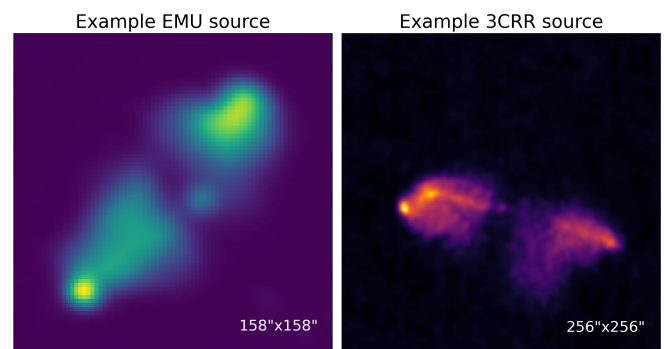
Figure 1. Comparison of two double lobed radio sources to highlight the resolution between datasets. As observed in J202644.4–552227 (left) and 3C285 (right), both sources show similar structural features. Angular size of each image shown in the bottom right.

The cutouts of the 3CRR sources, as well as the corresponding supplementary information, were directly downloaded from the Atlas of DRAGNs site. We do not use the full sample of 85 sources in this work. Through manual inspection, it was found that two sources had corrupt FITS files, two sources were not centred in the image, two were extremely bent (see Appendix 1), and seven contained excessive structural com-

[a]https://www.jb.man.ac.uk/atlas/

plexity too intricate to be processed reliably in an automatic way (see Table 1). Our reference dataset therefore consists of 72 3CRR sources (17 FRI and 55 FRII).

Table 1. List of sources and reasons for omission from analysis.

Reason for Omission	Sources
Corrupt FITS	3C236, NGC6251
Not centred	3C35, 3C465
Extremely bent	3C83.1B, 3C264
Complicated structure	3C16, 3C48, 3C84, 3C274, 3C293, 3C314.1, 3C274

Table 2. Summary of the metrics presented in this work. All metrics are unitless except for the angle between lobes, which is measured in degrees.

Metric	Summary
Asymmetry	Multiple metrics. Shape asymmetry, flux distribution asymmetry, shape and flux distribution asymmetry between both lobes of a given source.
Blurriness	Measure of how blurred the edge of a radio lobe is.
Concentration	Concentration of flux within a lobe.
Disorder	Measure of how complicated the shape of the lobe is.
Elongation	Measure of how stretched or compact a radio lobe is.
Angle Between Lobes	Measure of how bent a radio source is.

3. Development of Quantitative Metrics

As discussed above in Section 1, the morphology of a radio source does not always fit into well defined, rigid categories (such as FRI or FRII for example). Further, an object's classification can become ambiguous if it contains characteristics of different categories (e.g., HyMoRS) or when its observed structure varies with resolution. Recently, Rudnick (2021) proposed the concept of assigning radio sources descriptive criteria–based #*tags* instead of putting sources into such categories. Such #*tags* must have well defined criteria, and are ideally quantitative.

Of particular relevance to this work are structural #*tags*, which contain information about the number of jets, the peak flux separation, FR-morphology, shape, or symmetry, among other properties for a given radio source. In this work, we define a series of metrics that can be used to characterise and quantify the structure of a radio source and be incorporated into the #*tag* system. These metrics include asymmetry, blurriness, concentration, disorder, and elongation (referred to as the *ABCDE* metrics), which are applied to a single lobe of a DRAGN system. We also present metrics that are applied to the whole DRAGN system, the asymmetry and angle between lobes (ABL). A brief summary of each of the metrics is provided in Table 2. As the shape of a radio lobe is directly related to the astrophysics of the host AGN, we selected metrics designed to quantitatively capture both morphological structure and flux distribution of a lobe. It is important to note that while many of the following metrics are not new, the innovation in this work lies in their application to quantifying the structure of extended, complex radio sources.

To calculate each of the above metrics, all radio galaxy source images (from both the EMU-PS and 3CRR catalogue) are divided into two halves, with the expectation that each half will contain one lobe of the galaxy. For rectangular images, we cut the image perpendicular to and half-way along the major axis. For square images, we simply cut the x-axis in half. This is assuming the source is both centred and symmetrically distributed in the image. From a visual inspection of both EMU and 3CRR datasets, the vast majority of galaxies were centred in the radio image. Through manual inspection, we found 3 sources that were extremely bent with both lobes in one half of the cutout (See Appendix 1). These sources were dropped

from our analysis. We then identify the brightest pixel in each lobe, ensuring a minimum distance of three pixels from the border of each half. Further, for all metrics except the angle between lobes and asymmetry, we use a minimum threshold of 10% of the peak flux value to ensure lobe isolation and minimise biases from high noise pixels, setting pixel values below this threshold to NaN (Not a Number). For metrics that involve the summation of pixels, the NaNs are excluded.

3.1 Pre-processing

Many of the EMU-PS sources contain a very bright core. As the metric calculations are based on the brightest pixel, this can influence the results. We therefore implement a core-removal process. We first apply a minimum flux threshold of 150 µJy, corresponding to roughly 5σ of the EMU-PS imaging (σ = 25-30 µJy/beam, Norris et al., 2021). Pixels with intensities less than this threshold are are set to zero to remove noise for the purposes of core identification only. Following the work outlined in Hancock et al. (2012), to highlight potential edges and features in an image we generate a curvature map by convolving the thresholded image with the following Laplacian kernel:

$$L_{xy}^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} . \tag{1}$$

The resulting curvature map emphasises regions with sharp intensity transitions, aiding in the accurate localisation of a core. We then generate a binary map with labelled features, calculate their sizes, and retain the largest components. If an image contains two lobes and a core (i.e., number of components > 2), a circular mask is applied around the local maximum nearest to the infrared coordinates of the host galaxy, which is expected to be coincident with the core. This process is fully automatic and demonstrated in Figure 2. Through an initial visual inspection of the 592 EMU sources, we estimate that 186 (30%) sources contain a dominant core, i.e., the central component was brighter than the two lobes. We then applied the above core-removal process to the 592 images and, after another visual inspection, the core of 102 sources was successfully removed. However, this process is not perfect

(see Appendix 1). We found that the number of components detected was sometimes incorrect, leading sources that did not have a dominant core to accidentally have a lobe masked. Sources with a dominant core remaining after the attempted masking (84) or an inadvertently masked lobe (28) were subsequently dropped from our analysis, leading to a final EMU-PS dataset consisting of 480 sources.
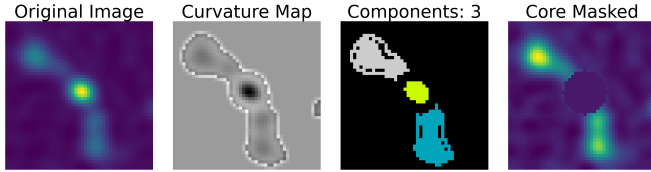
Figure 2. Visualisation of the core masking process. From left to right is the original EMU-PS radio image, the curvature map, the labelled components, and the final image with the core removed.

3.2 Asymmetry

The asymmetry of a galaxy has been defined in a number of ways in optical astronomy (e.g., Schade et al., 1995; Abraham et al., 1996a,b; Conselice, 1997; Brinchmann et al., 1998). These methods provide a framework for morphological classification that can also be applied to radio sources. We adopt the approach to asymmetry as described in Conselice et al. (2000) and Conselice (2003), which compares a source image (I), with itself rotated by an angle of $180°$ (I_{Rot}). We define the centre of rotation as the brightest pixel in a given lobe. We produce the difference image between the original and rotated image, and take the absolute value of the sum over all pixels, η, and normalising by the total flux of the original lobe, I_{Tot}, to produce a shape asymmetry metric (A_S). The absolute value of the difference of two images will produce twice the flux for a perfectly asymmetric source (i.e., each bright pixel has no corresponding bright pixel after the image is rotated), and we therefore include the factor of $1/2$ to ensure A_S is always between 0 and 1:

$$A_S = \frac{1}{2} \sum_{i,j}^{\eta} \frac{|I - I_{Rot}|}{I_{Tot}} . \qquad (2)$$

We also include a metric to describe the flux distribution across a lobe, defining a ratio asymmetry (A_R), which is the absolute value of the ratio of the original and rotated lobe, summed over all pixels, η (Equation 3). For a completely uniform (or symmetric) source, each pixel ratio would be 1, and the sum of the ratio image will be number of pixels. We therefore normalise by the number of pixels, meaning a perfectly uniform source will produce a minimum A_R value of 1:

$$A_R = \frac{1}{\eta} \sum_{i,j}^{\eta} \left| \frac{I}{I_{Rot}} \right| . \qquad (3)$$

Examples of the images of A_S and A_R (before summing the pixels) are shown in Figure 3.
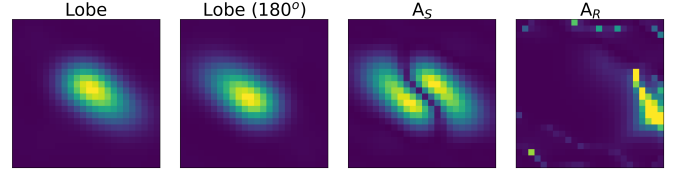
Figure 3. Visualisation of both asymmetry calculations. From left to right: original lobe image, its rotated counterpart, the residual image after taking the absolute value of the difference between the original and rotated image, and the residual image after taking the ratio of the original and rotated images.

We can define analogous calculations to A_S and A_R to compare how asymmetric the two lobes (I_1 and I_2) of a radio galaxy are to each other. We denote these new parameters with the subscript p to represent the comparison of the pair of lobes in the DRAGN system. We can define two shape asymmetry metrics, the first being the total asymmetry ($A_{S,p}$, Equation 4) which is weighted by the relative flux in each lobe. This takes a similar form to the A_S calculation, however we instead normalise by the sum of the total flux in both lobes. To compare only how symmetric shapes of the lobes are, without the relative flux weighting, we must normalise each lobe by the total flux, I_{Tot} before computing the difference ($A'_{S,p}$, Equation 5). The ratio asymmetry between lobes, $A_{R,p}$ (Equation 6), is determined in the same way as A_R, using the absolute value of the ratio of one lobe with the rotated other, normalised by the number of pixels, η. As before, we can also normalise by the total flux before computing the ratio ($A'_{R,p}$, Equation 7). As with the single-lobe calculations, $A_{S,p}$ and $A'_{S,p}$ both have values between 0 and 1, and $A_{R,p}$ and $A'_{R,p}$ are always greater than 1.

$$A_{S,p} = \frac{1}{2} \sum_{i,j}^{\eta} \frac{|I_1 - I_{2,Rot}|}{I_{1,Tot} + I_{2,Tot}} , \qquad (4)$$

$$A'_{S,p} = \frac{1}{2} \sum_{i,j}^{\eta} \left| \frac{I_1}{I_{1,Tot}} - \frac{I_2}{I_{2,Tot}} \right| , \qquad (5)$$

$$A_{R,p} = \frac{1}{\eta} \sum_{i,j}^{\eta} \left| \frac{I_1}{I_{2,rot}} \right| , \qquad (6)$$

$$A'_{R,p} = \frac{1}{\eta} \sum_{i,j}^{\eta} \left| \frac{I_1/I_{1,Tot}}{I_{2,rot}/I_{2,Tot}} \right| . \qquad (7)$$

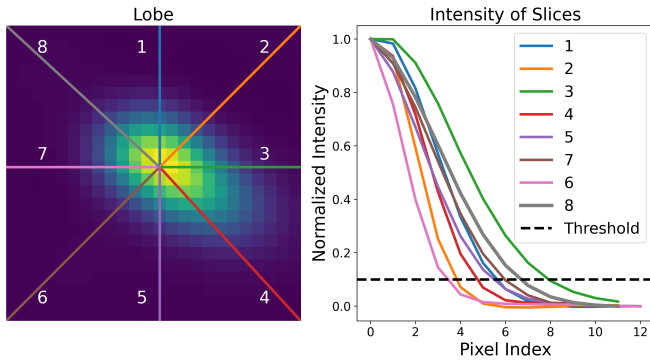## 3.3 Blurriness

We calculate how blurry (or diffuse) radio lobes are by measuring the rate at which the flux of a lobe decreases from the hotspot radially outwards. We first define eight cardinal directions, taking slices of the pixel intensities radially outwards from the brightest pixel (i.e., the hotspot). To calculate the rate at which the flux decreases, we use the `scipy.optimize` `curve_fit` function to fit a straight line to each of the eight

slices and extract the gradient (Figure 4). While we acknowledge that the pixel intensity is not linear, this approach provides a robust first-order estimate of blurriness, as more complex models (e.g., exponential or power law fits) introduce additional parameters that may not be necessary for comparative classification. The gradient of pixel intensities (normalised by the peak brightness) in the whole slice alone is not enough to determine how blurred or sharp a lobe is. We therefore also measure the gradient of the three outermost pixels in each slice. We chose to use 3 pixels as this is the minimum number of points we can have to fit a gradient, and limit the number of pixels from within the lobe. We average over all eight directions for both the full slice and outermost pixels ($\bar{g}_{full}$ and $\bar{g}_{outer}$, respectively). To estimate the how blurred a lobe is, we then take the ratio of $\bar{g}_{full}$ and $\bar{g}_{outer}$:

$$B = \frac{\bar{g}_{full}}{\bar{g}_{outer}} . \tag{8}$$

Small B values (low blurriness) are produced when the average edge gradient of flux is steeper than the average gradient of the whole lobe. This could be indicative of a relatively constant brightness distribution across the lobe with a sharp drop to the background, possibly produced by a lobe expanding into a dense environment, such as within a galaxy cluster. Conversely, higher B values (high blurriness) may indicate more diffuse lobe structure and potentially a less dense environment.

Figure 4. Visualisation of the blurriness calculation, highlighting the 8 cardinal directions and corresponding normalised intensity slices measured away from the brightest pixel of a lobe.
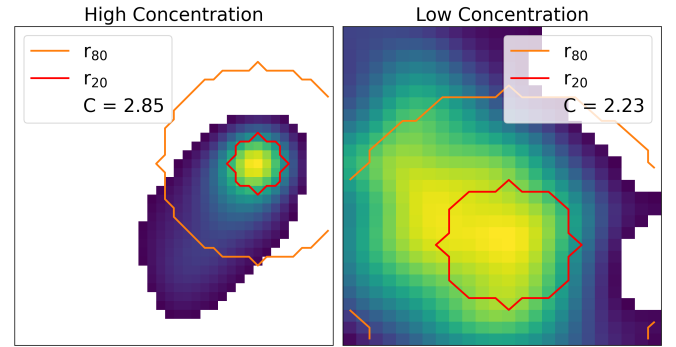
3.4 Concentration

The concentration metric describes how the flux is distributed within a lobe. Our approach to concentration is analogous to the approach described in Abraham et al. (1994), with concentration defined as the ratio of flux between inner and outer isophotes of normalised radii. We define concentration, C, as the ratio of radii containing 80% and 20% of the flux in a lobe, centred on the brightest pixel:

$$C = \frac{r_{80}}{r_{20}} . \tag{9}$$

To calculate the radius of each isophote, we sum the number of pixels until 20% and 80% of the total lobe flux is achieved.

This sum is treated as an area, α, and converted into a circular radius, r, using the relationship $r = \sqrt{\alpha/\pi}$. Here, C has a minimum value of 2 for a completely uniform source. Larger values of C mean the flux is more concentrated within the lobe, and can be indicative of a hotspot. A comparison showing high and low concentration lobes is presented in Figure 5.

Figure 5. Visualisation of the concentration metric comparing lobes with high and low concentration (left and right, respectively). Radii corresponding to 20% and 80% of the lobe flux shown in red and orange, respectively.

3.5 Disorder

We include a metric to quantify how perturbed a radio galaxy lobe is. Perturbations in the lobe structure can be caused by differing jet power, or from dynamic forces within the host galaxy environment. We refer to this metric as 'disorder', D, and following the definition of complexity by Watson (2012), define it as the ratio of the perimeter of a lobe squared and its area multiplied by 4π:

$$D = \frac{\text{Perimeter}^2}{4\pi\text{Area}} . \tag{10}$$

To determine the perimeter, we measure the length of a fitted contour around pixels above the minimum threshold (10% of the peak brightness). To calculate the area of the lobe, we simply sum the number of non-NaN pixels above the minimum threshold. The D metric therefore has a minimum value of 1 for a perfectly circular source. Figure 6 highlights the difference between lobes with high and low disorder.
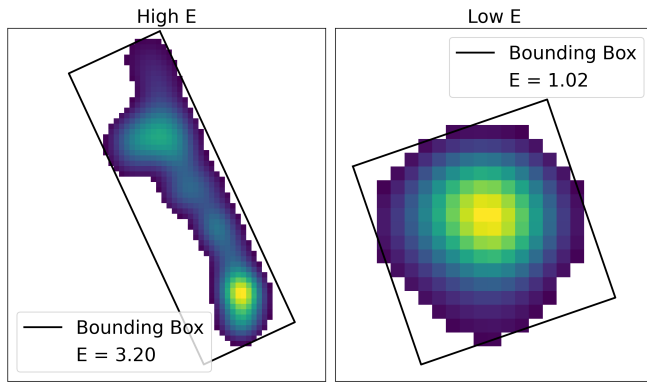
3.6 Elongation

We also measure the elongation of a lobe for a radio source, which can be related to the power of a radio jet as more powerful radio jets can produce more extended structures. We define it as the ratio of the major axis and the minor axis of a bounding box fitted to a radio lobe:

$$E = \frac{\text{Major axis}}{\text{Minor axis}} . \tag{11}$$

This metric will have a minimum value of 1 for a square or circularly shaped object. We perform a principle component analysis (PCA) on the pixels above the minimum threshold to orient and fit the bounding box, where we can then extract the lengths of the major and minor axes of the box. This is shown in Figure 7.
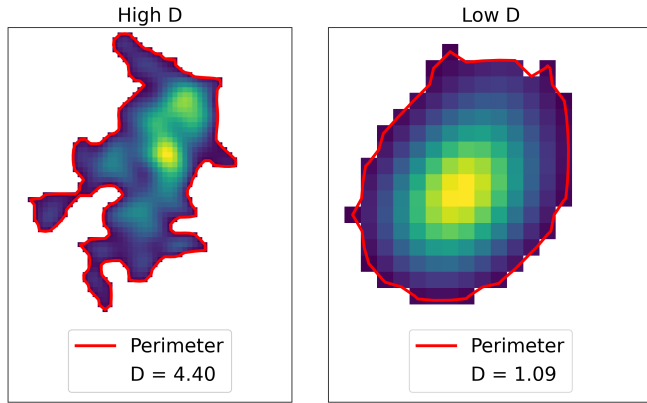
Figure 6. Visualisation of the disorder metric comparing lobes with high and low disorder (left and right, respectively). Perimeter calculated by the length of the red contour.

Figure 7. Visualisation of the elongation metric comparing lobes with high and low elongation (left and right, respectively). The bounding box of each lobe is shown in black.

3.7 Angle Between Lobes

We include a metric that estimates how bent a radio source is, the angle between lobes (ABL). The ABL can serve as a proxy for environmental effects, as bent radio galaxies are predominantly found in dense and dynamic environments such as galaxy clusters. To calculate the ABL, we treat the brightest pixels in each lobe as points in space, L_1 and L_2, where we can then determine vectors to the core of the radio galaxy, C (Equations 11a and 11b). For the location of the core, we prioritise using the infrared coordinates of the host galaxy when available, otherwise we simply use the geometric centre of original radio galaxy cutout. Once the vectors have been determined, we compute the dot product (Equation 12) and magnitude of the vectors (Equations 12a and 12b) to extract the angle between lobes (Equation 13), which is then converted into degrees. This calculation is illustrated in Figure 8. We will refer to ABL, $A_{S,p}$, $A'_{S,p}$, $A_{R,p}$, and $A'_{R,p}$ as source metrics hereafter.

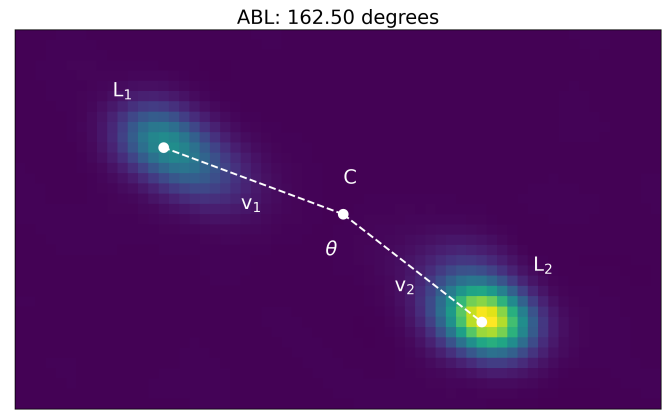$$\mathbf{v}_1 = \mathbf{L}_1 - \mathbf{C} = (x_1 - x_c, y_1 - y_c), \tag{11a}$$

$$\mathbf{v}_2 = \mathbf{L}_2 - \mathbf{C} = (x_2 - x_c, y_2 - y_c), \tag{11b}$$

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = (x_1 - x_c)(x_2 - x_c) + (y_1 - y_c)(y_2 - y_c), \tag{12}$$

$$|\mathbf{v}_1| = \sqrt{(x_1 - x_c)^2 + (y_1 - y_c)^2}, \tag{12a}$$

$$|\mathbf{v}_2| = \sqrt{(x_2 - x_c)^2 + (y_2 - y_c)^2}, \tag{12b}$$

$$\cos(\theta) = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|} \Rightarrow \theta = \arccos\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|}\right). \tag{13}$$


Figure 8. Example calculation of the ABL for a radio source from the EMU-PS. Positions of the brightest pixels in each lobe and the location of the host galaxy shown by the white dots and labelled L_1, L_2 and C, respectively. Vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ shown by the dashed lines. Angle θ indicated by the arc between both vectors.

3.8 Resolution Dependence of Metrics

It is well known that metrics used in quantifying shapes, such as those presented in this work, are likely to be sensitive to the resolution of the images analysed. The shape of unresolved or marginally resolved sources (i.e., components only one or two beams wide) will be determined by the shape of the telescope synthesised beam, which is usually an elliptical Gaussian (Boyce et al., 2023). To understand the effect resolution has on each of our metrics, we convolve high resolution images with 3 different 2D Gaussian kernels, each with increasing standard deviation ($\sigma = 3$, 6, and 9) to produce images with progressively worse resolution. The convolution is performed in real space using direct kernel application. As the σ increases, we begin to lose more and more structure until the lobes become unresolved (Figure 9). We chose to convolve the 3CRR images for convenience, but the following analysis should be independent of the input images.

We apply each metric to the convolved images and compare them with the original values in Figure 10. We can see that the ABL calculation (Figure 10k) is the most robust against resolution differences. The scattered points in this metric result from location of the brightest pixel changing, often moving closer to the centre of a lobe after convolving. We see that convolving tends to produce smaller values for shape
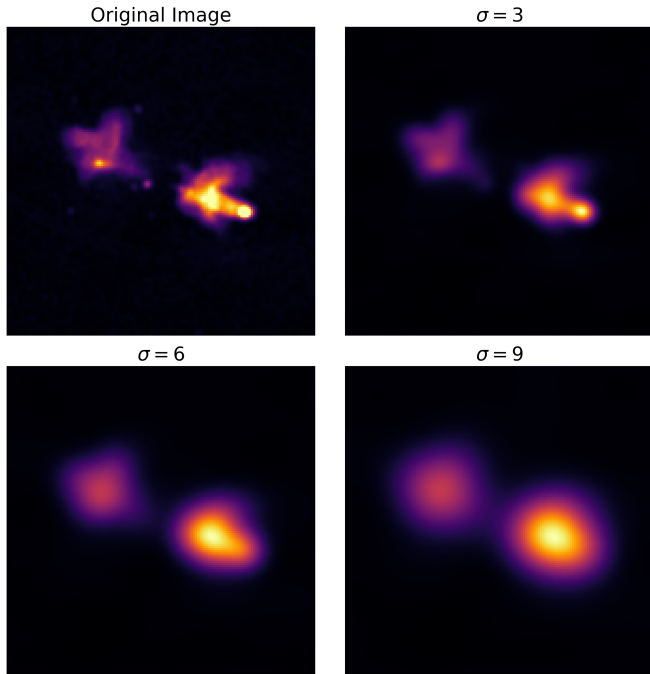
Figure 9. Effect of convolving radio image of 4C73.08 (top left) with 2D Gaussian kernels of increasing σ in units of pixels.

asymmetry measures (A_S, $A_{S,p}$ and $A'_{S,p}$). This is an expected result, as the flux in the lobes becomes more evenly distributed and circular in shape, resulting in lower asymmetry. This appears in Figure 10c, g, and h, where we can see smaller asymmetry values with increasing σ.

The other metrics show less intuitive trends. The process of convolving reduces the brightness of the peak pixel, spreading that flux over the surrounding pixels. For the metrics that use threshold masking (all apart from the ABL and asymmetry metrics), using 10% of a lower peak brightness means more pixels lie above the threshold. With B for example, more dim pixels around the edge of the lobe are retained. These pixels would produce a smaller average outer gradient making the overall ratio a larger number, moving the points above the 1 : 1 line (Figure 10a). This effect would be increased as sigma gets larger, which is why the σ = 9 points sit higher than the σ = 6 and σ = 3 points. For concentration, this may increase r_{80} at a faster rate than the r_{20}, making the overall concentration measurement higher.

We see for disorder and elongation that the convolved images produce fairly stable values, that then decrease gradually with increasing σ. This again is an expected result as the shape of the lobe becomes more circular with convolving, therefore making both disorder and elongation calculations approach unity. For disorder, the increase in the number of pixels will increase both the perimeter and area, however the perimeter^2 dependence will cause the disorder calculation to become much larger. For elongation, we expect the value to become closer to unity after convolving. It is possible that the increase in the number of dim pixels included some of the jet structure that was previously too dim, therefore increasing the overall

elongation value. We have similar expectations with A_R, in that as we convolve the distribution of flux is smoothed, so A_R should approach unity. As before with elongation, the inclusion of jet structure that was previously too dim, before convolving, can again increase the values of A_R.

The above analysis shows that the metrics can be highly resolution dependent, as evidenced in Figure 10, with many metrics showing high scatter when comparing original and convolved values. Higher resolution will enable the metrics to be more robust, as there is more structural information available to be extracted.

4. Results and Analysis

In this section, we present the results of applying the *ABCDE*, and source metrics to both the EMU-PS and 3CRR datasets to characterise and quantify the structure of the radio lobes. In Tables 3 and 4, we include a summary of the computed metric values for EMU and 3CRR sources, respectively. This data is available upon request. The name, host position, flux, and angular size for the EMU sources were obtained directly from RadioGalaxyNET (Gupta et al., 2024a). The name, flux, angular size, redshift and FR classification for the 3CRR sources were obtained from the Atlas of DRAGNs site (Leahy et al., 2013), with the host coordinates cross-matched from the Vizier catalogue of 3CRR sources (Laing et al., 2003). Note, all metric values are expressed in $\log_{10}$, except for ABL. These results form the basis for the subsequent analyses, where we explore correlations between metrics and if our metrics can effectively characterise the structure and morphology of radio galaxies.

4.1 Single-Lobe Metrics

Figure 11 shows pairwise combinations of the single-lobe calculations (*ABCDE*), with the normalised kernel density estimate (KDE) for each metric in the diagonals. For most of the metrics, the EMU and 3CRR sources occupy similar areas of parameter space, indicating the minimum image area of 20 ASKAP beams for the EMU sources allows them to be sufficiently resolved for the metric results to be comparable with the 3CRR data. While some metric combinations show slight correlations (e.g., B and C) or more complicated relationships (e.g., A_S and A_R) we do see high scatter in each subplot, indicating that these metrics my be independent. We calculated the Spearman rank correlation coefficient (ρ, Siegel & Castellan, 1988; Conover, 1999) on each of the unique metric combinations from Figure 11. Values of ρ ≈ 0 correspond to little to no correlation, with ρ ≈ ±1 correspond to a strong positive or negative correlation, respectively. We found that most of the metric combinations had –0.054 < ρ < 0.27, with the exception of A_S-A_R (ρ = 0.72) and *D-E* (ρ = 0.57). This indicates that the metrics are, in fact, largely independent of each other and therefore measure fundamentally different structural features of a radio galaxy lobe.

Since the FRI/II dichotomy has been a very productive scheme for uncovering the physics of radio galaxies, we will explore to what extent these metrics can reproduce this dichotomy. Both FR-types occupy similar regions in metric
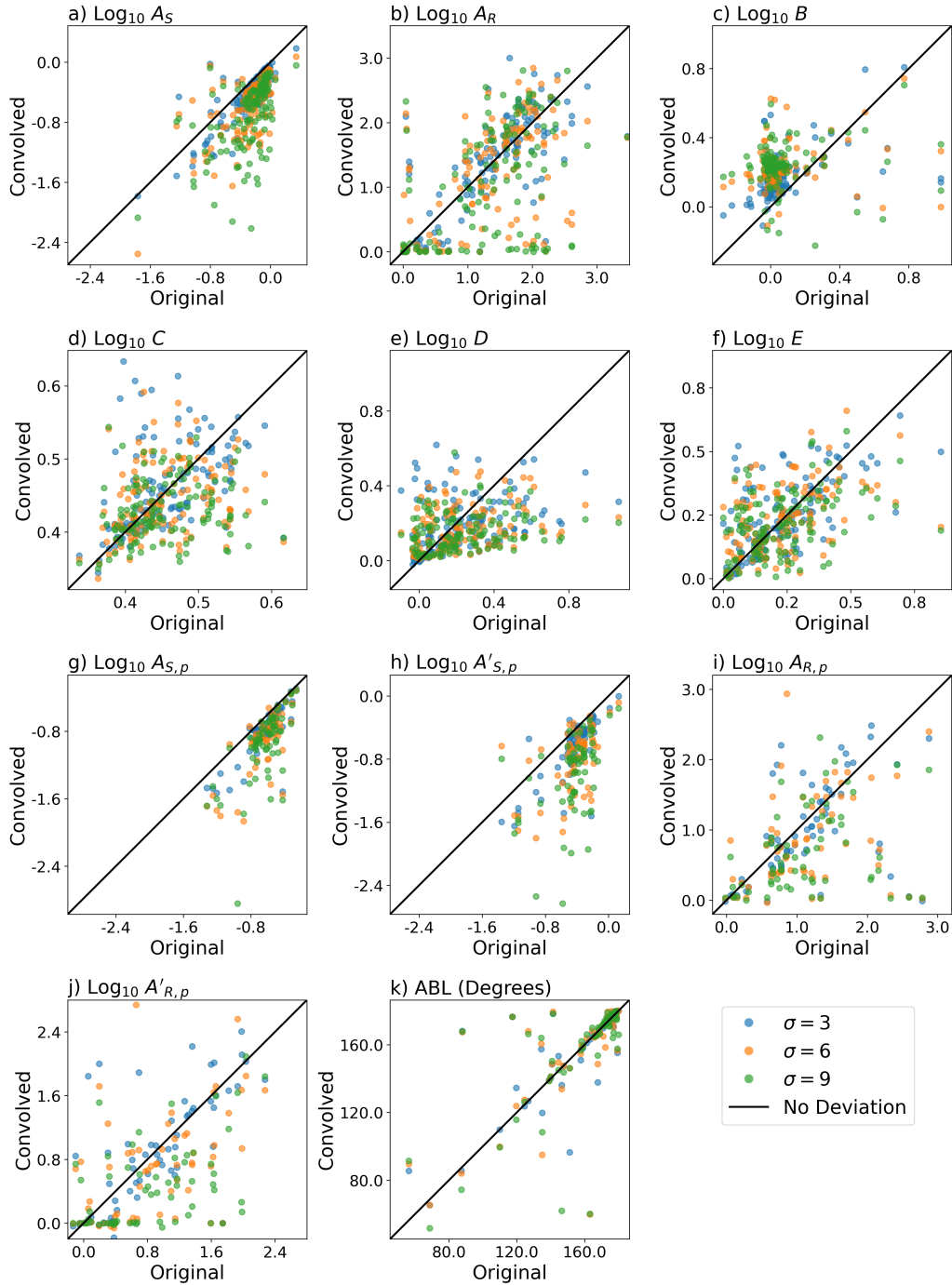
Figure 10. Comparison of the metric outputs of original 3CRR images and convolved versions. Blue, orange and green points correspond to images convolved with Gaussian kernels with σ = 3, 6, and 9 respectively. Black 1:1 line indicates no deviation in the metrics after convolving. Panels **a)-e)** have two points per source, one for each lobe.

space for the majority of our metrics, with the exception A_S, A_R where we see stronger separation between FRI and FRII sources. FRI sources also tend to have a broad range of values, when compared to the FRII sources, particularly in the B, C, and E metrics. This reinforces the need for additional metrics to completely characterise the shape of radio lobes. We explore correlations between metrics and FR distributions in more detail in subsequent sections.

4.1.1 Asymmetry

Evident in both the scatter and KDE panels in Figure 11, the FRI and FRII sources are found in different areas of the asymmetry parameter space. We see that for both A_S and A_R that the FRI sources tend be less asymmetric than FRII sources. While the EMU sources share the same region of parameter space as the FRIs there is evidence of a bimodal distribution, highlighted in the KDE of the A_R metric. There is a highly
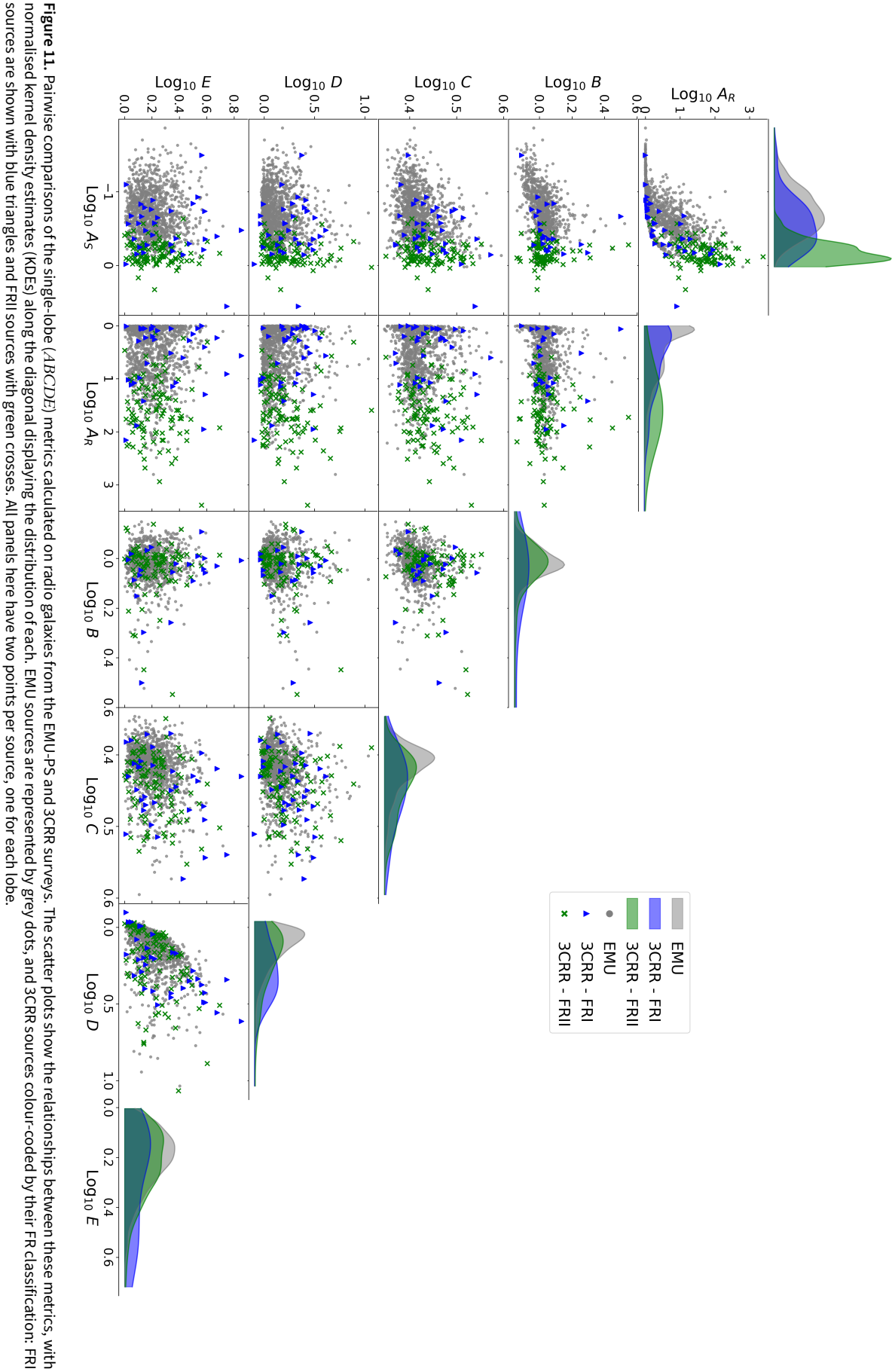
Figure 11. Pairwise comparisons of the single-lobe ($ABCDE$) metrics calculated on radio galaxies from the EMU-PS and 3CRR surveys. The scatter plots show the relationships between these metrics, with normalised kernel density estimates (KDEs) along the diagonal displaying the distribution of each. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses. All panels here have two points per source, one for each lobe.

Table 3. Top 10 rows of $ABCDE$, and source metric calculations for the EMU-PS sources.

Name	RA (°)	Dec (°)	Flux (Jy)	Ang Size (")	ABL (°)	A_{S1}	A_{R1}	A_{S2}	A_{R2}	A_{Sp}	A'_{Sp}	A_{Rp}	A'_{Rp}	$B1$	$B2$	$C1$	$C2$	$D1$	$D2$	$E1$	$E2$
J200753.8−502330	301.98	−50.39	0.00	109.58	168.14	−0.35	1.40	−0.50	0.78	−0.50	−0.41	1.22	0.70	0.04	0.03	0.45	0.43	0.71	0.32	0.22	0.45
J200825.1−525526	302.11	−52.92	0.00	153.56	176.56	−0.56	0.63	−0.50	0.76	−0.69	−0.87	0.77	0.18	−0.02	0.03	0.41	0.41	0.21	0.24	0.39	0.28
J200828.6−530013	302.12	−53.00	0.08	104.78	162.74	−0.60	1.00	−0.69	0.59	−1.27	−0.97	0.41	0.17	0.04	0.03	0.39	0.37	0.11	0.08	0.34	0.28
J200920.8−494020	302.33	−49.67	0.10	107.25	159.24	−0.59	0.58	−0.65	0.72	−0.82	−0.72	0.32	0.08	−0.03	0.00	0.40	0.40	0.04	−0.00	0.25	0.14
J200947.6−592758	302.45	−59.47	0.00	90.75	117.14	−0.97	0.02	−0.68	0.58	−1.05	−1.07	0.16	0.02	NaN	NaN	0.39	0.36	0.31	0.30	0.25	0.19
J201017.6−522855	302.57	−52.49	0.00	115.48	173.50	−0.18	1.13	−0.45	0.87	−0.50	−0.24	1.09	0.94	0.11	−0.05	0.44	0.45	0.41	0.01	0.22	0.13
J201051.4−525836	302.72	−52.98	0.03	137.84	165.68	−0.51	0.85	−0.58	0.59	−0.80	−0.52	0.41	0.32	−0.00	0.04	0.44	0.39	0.31	0.01	0.47	0.09
J201052.7−582128	302.72	−58.36	0.01	105.75	164.57	−0.36	0.82	−0.22	0.57	−0.59	−0.47	0.79	0.80	0.04	0.06	0.46	0.45	0.13	0.55	0.20	0.02
J201055.3−492051	302.73	−49.34	0.02	158.69	162.90	−0.57	0.68	−0.54	0.32	−1.15	−0.94	0.03	−0.04	0.03	0.01	0.41	0.43	0.01	0.03	0.14	0.17
J201107.5−525706	302.78	−52.95	0.03	82.40	178.51	−0.63	1.22	−1.01	0.01	−0.93	−1.01	0.20	−0.01	0.22	−0.06	0.41	0.43	0.16	0.06	0.19	0.12

Table 4. Top 10 rows of $ABCDE$, and source metric calculations for the 3CRR sources.

Name	RA (°)	Dec (°)	Flux (Jy)	Ang Size (")	ABL (°)	A_{S1}	A_{R1}	A_{S2}	A_{R2}	A_{Sp}	A'_{Sp}	A_{Rp}	A'_{Rp}	$B1$	$B2$	$C1$	$C2$	$D1$	$D2$	$E1$	$E2$	z	FR Class
4C12.03	2.47	12.73	10.90	232.00	179.36	−0.04	1.69	−0.71	0.11	−0.55	−0.38	0.58	1.98	−0.12	0.68	0.43	0.44	0.29	0.57	0.41	0.23	0.16	II
3C19	10.23	33.17	13.20	6.80	173.08	−0.07	1.40	−0.08	1.30	−0.58	−0.31	1.13	0.91	0.02	0.08	0.42	0.49	0.03	0.22	0.10	0.15	0.48	II
3C20	10.79	52.06	46.80	53.10	174.63	−0.07	1.88	−0.14	1.95	−0.70	−0.43	0.97	0.80	0.06	0.07	0.47	0.42	0.34	0.01	0.08	0.13	0.17	II
3C28	13.96	26.41	17.80	45.60	157.69	−0.44	1.47	−0.12	2.18	−0.62	−0.31	2.33	0.67	−0.11	−0.14	0.41	0.40	0.30	0.21	0.02	0.21	0.20	II
3C31	16.85	32.41	18.30	2700.00	87.27	−0.17	2.02	−0.80	0.07	−0.46	−0.64	0.72	0.25	0.04	NaN	0.47	0.42	0.49	0.21	0.58	0.10	0.02	I
3C33	17.22	13.34	59.30	257.00	177.31	−0.44	1.67	−0.20	2.08	−0.54	−0.43	0.78	0.34	−0.04	−0.01	0.41	0.42	0.02	0.04	0.07	0.28	0.06	II
3C33.1	17.43	73.20	14.20	238.70	172.26	−0.13	2.39	−0.09	2.54	−0.59	−0.45	1.09	1.93	0.02	0.09	0.46	0.48	0.28	0.66	0.26	0.37	0.18	II
3C42	22.13	29.05	13.10	31.00	165.69	−0.15	2.24	−0.31	1.65	−0.74	−0.43	0.75	0.62	0.03	−0.02	0.43	0.41	0.02	−0.01	0.14	0.05	0.40	II
3C46	23.87	37.90	11.10	164.00	173.00	−0.13	1.74	−0.15	1.35	−0.51	−0.19	1.69	1.60	−0.01	0.09	0.48	0.49	0.18	0.20	0.41	0.32	0.44	II
3C47	24.10	20.96	28.80	77.50	178.38	−0.13	2.32	−0.29	2.03	−0.70	−0.44	0.63	0.42	NaN	0.02	0.42	0.46	0.19	0.13	0.39	0.28	0.42	II

concentrated population at very low A_R, and another more scattered population at moderate to high A_R. The initial comparison of EMU and 3CRR sources indicates that the low A_R values coincide with the FRIs and the moderate to high A_R values with FRIIs. We explore the A_S-A_R parameter space further in Figure 12.
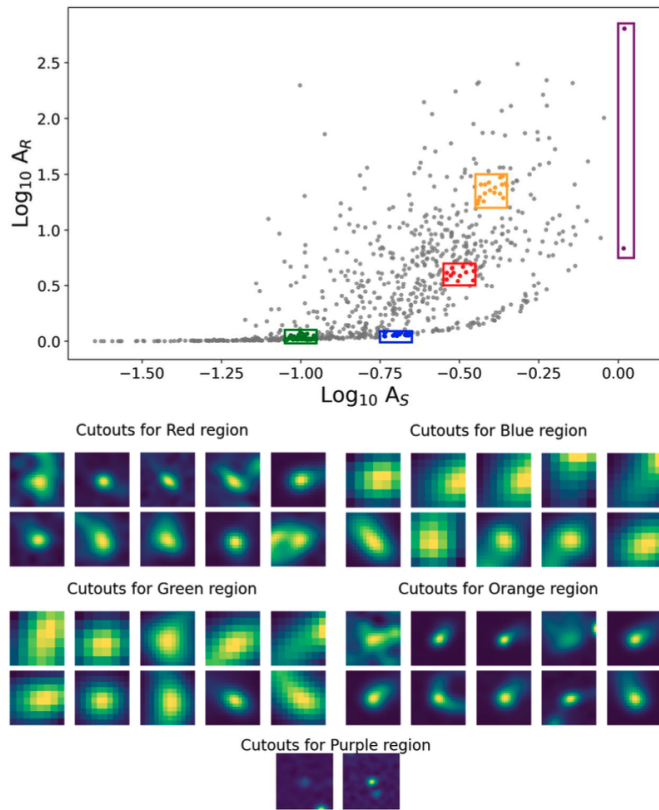
Figure 12. Enlarged version of the A_S-A_R sub-panel from Figure 11 showing only EMU data. In this figure different regions of this parameter space are highlighted by different colours. Panels below are lobe cutouts for each highlighted region.

The concentrated tail at low A_R values (e.g., the blue and green regions) may be produced when the brightest pixel is near the edge of the lobe cutout, therefore limiting the size of the cutout and causing the overall flux distribution to appear more uniform. This will result in lower values for the A_R metric, forming the tail with $\log_{10}(A_R)$ values close to 0. We can relate this to the abundance of FRI sources present in the tail. FRI sources are often referred to as core-brightened, with the peak brightness occurring in the jets near the core. The code splits an image of a radio galaxy in half under the assumption that there will be a radio lobe in each half. As the FRI hotspots tend to be closer to the core, the brightest pixel for an FRI source would therefore be near the border of the lobe cutout, limiting the number of pixels in the asymmetry calculations. The tail is therefore likely an artefact of both the flux distribution of the sources and of how we isolate a lobe in our algorithm. For objects found in the more scattered section of this parameter space (e.g., the red and orange regions), the asymmetry metric appears to be working as expected, with

the scatter resulting from actual variance in lobe structure in different sources.

In the purple region, there are two measurements with $\log_{10}(A_S) > 0$, which should not occur for a normally distributed source. We can see in the lobe cutouts for this region of parameter space that both contain two separate circular sources, with one near the centre and the other offset. From visual inspection, the central sources appear to have low rotational asymmetry, and will not contribute much to the asymmetry value. For the off-centre sources however, the rotation by $180°$ will produce two bright spots after taking the absolute value, corresponding to a high A_S value. We should also note that both of these sources have relatively low overall brightness, and, we can see some of the noise fluctuations in surrounding pixels.

We conducted simulations to investigate the features observed in Figure 12, particularly what produces the concentrated tail and what causes sources to have $\log_{10}(A_S) > 0$. To simulate a perfectly symmetric lobe (red points), with the only asymmetry coming from noise fluctuations we first populate a 1D array with values ranging between 1×10^{-5} and 1, referred to as array x. We then generate a second 1D array, y, which is identical to array x except for the addition of random noise, normally distributed around zero with an amplitude of 0.1:

$$y = x + \text{noise} ,$$

where we then calculate A_S and A_R by substituting x for lobe 1 and y for lobe 2 into Equations 2 and 3, respectively.

To simulate an asymmetric lobe (blue points), we generated 15 different 25×25 Gaussian profiles, with ranging intensities, σ, skew, and rotation (Figure 13). We loop over each of the profiles, systematically adding noise until we have the same number of data points as our EMU-PS data (480), repeating the asymmetry calculations. The noise added in this loop is normally distributed around zero, with amplitude ranging from 0.015 to 0.9.
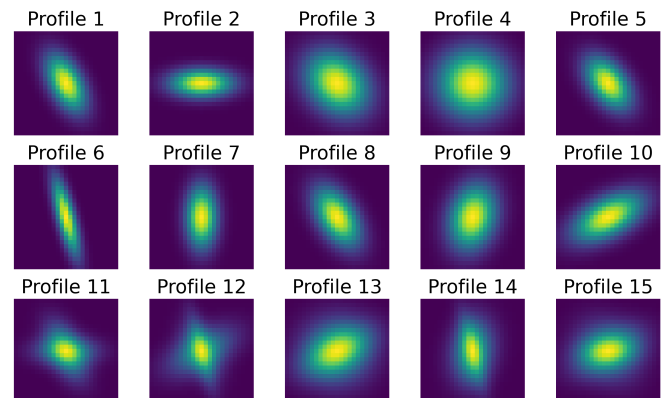
Figure 13. Collection of the 15 different 2D Gaussian profiles used in the simulations.

Finally for the noise–only simulation (green points), we randomly populate a 15×15 array with Gaussian noise matching the noise distribution of the EMU-PS, i.e., centred on zero with σ = 23 μJy/beam (Norris et al., 2021).

The asymmetry calculations on the simulated data (Figure 14) revealed that the noise-only points explicitly exist at $\log_{10}(A_S) > 0$. This is likely due to normalising by the total flux. The sum of an array populated by Gaussian noise distributed around zero approaches 0 as we increase the number of points in the array. This means that for the noise-only points (i.e., sources with poor signal-to-noise), we normalise by a very small number, leading to $\log_{10}(A_S)$ values greater than 0. In this work we wanted to investigate how the metrics perform on a variety of different sources, and we do not include any minimum signal to noise cutoff.
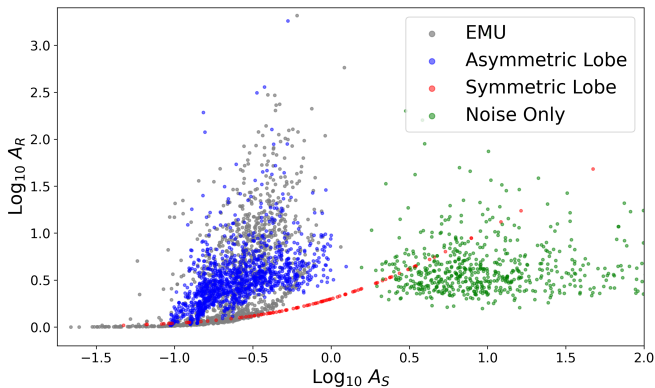
Figure 14. Enlarged version of the A_S-A_R sub-panel from Figure 11 with only EMU data, comparing how the asymmetry parameters characterise different simulated data. Grey, blue, red, and green correspond to data points from the EMU-PS, simulated asymmetric lobes, simulated symmetric lobes, and Gaussian distributed noise, respectively.

While the initial asymmetry results (both A_S and A_R) taken at face value suggest that FRII sources tend to be more asymmetric than FRIs, they need to be interpreted with caution. The way A_S and A_R are defined, using the brightest pixel as a centre of rotation, the tendency is for FRI sources to have fewer pixels, and a more symmetrical structure within the cutout for each lobe. As discussed earlier, a limited number of pixels (i.e., if the bright spots of a lobe are near the edge of a lobe cutout) can produce a low asymmetry value. Conversely, FRII lobes tend to have hotspots, and hence a brightest pixel, positioned such that more of the actual lobe structure is included in the measurement, leading to higher values of asymmetry. This is not to say that these measures are wrong, but it is important to interpret them given how they are defined, reflecting the degree of asymmetry only within the region sampled.

4.1.2 Blurriness and Concentration

Figure 15 reveals a weak positive correlation between the B and C metrics ($\rho = 0.27$) for the EMU and for the 3CRR data ($\rho = 0.24$). Small B values can be indicative of an edge-brightened source, possibly identifying lobes that are compressed by the environment. Conversely larger B values indicate a rapid flux drop from the brightest pixel to the outer parts of the lobe (i.e., higher concentration of flux), corresponding to a edge-dimmed source, that could exist in a lower density environment. These implications can be tested by investigating the number density of surrounding galaxies (e.g., Brough

et al., 2013) for sources with different B and C values. We can also see the FRI and FRII sources span very similar ranges for both B and C, (although FRIIs may have slightly higher concentration values than FRIs). When developing B and C, it was anticipated that the hotspots would result in higher concentration (and blurriness) values for FRII sources, potentially showing higher separation between FRI and FRII types in this metric space. These findings further highlight the limitations of the binary FR classification. Quantitative metrics, such as B and C, provide more nuanced descriptions of radio galaxy morphology. This reinforces the need to move away from the binary FR classification system and instead use quantitative descriptive metrics for a given source.
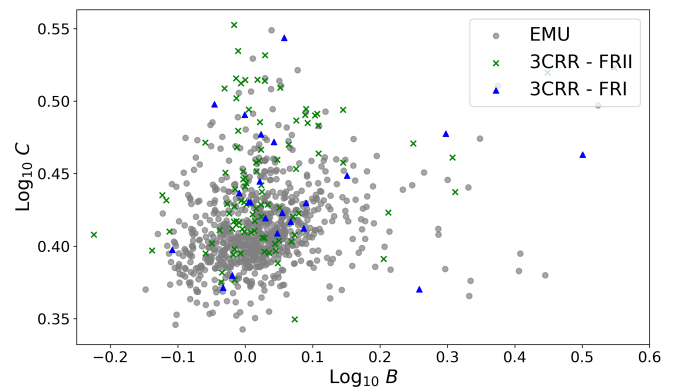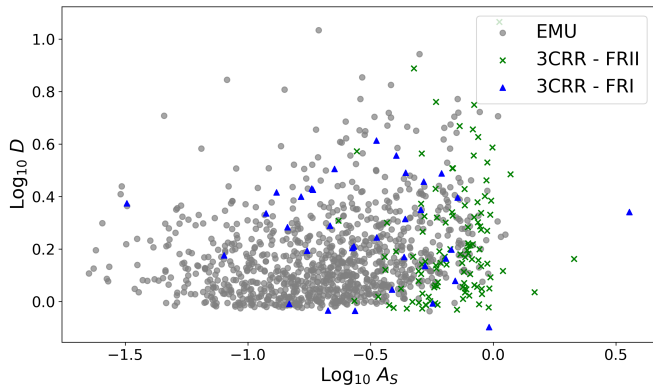

Figure 15. Enlarged version of the B-C sub-panel from Figure 11. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses.
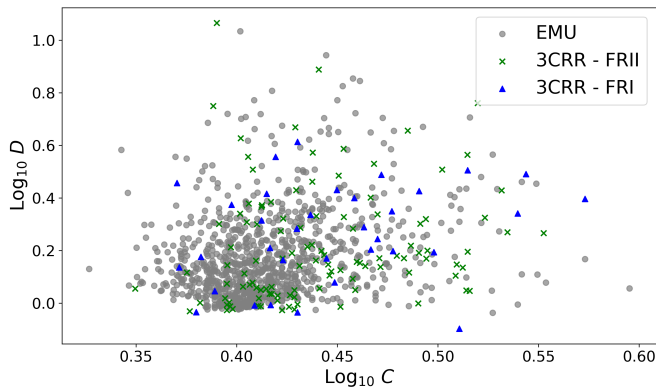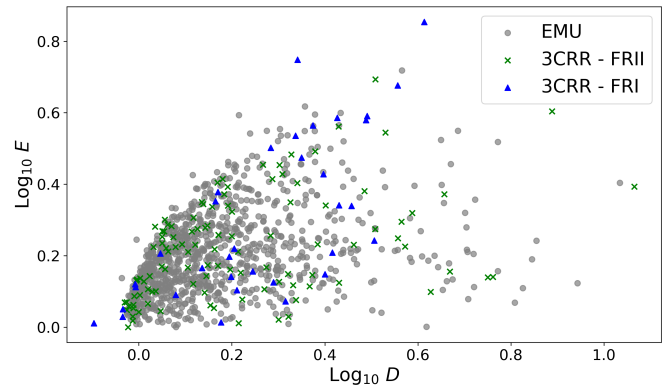
4.1.3 Disorder and Elongation

We explore connections between A_S-D and C-D in Figures 16 and 17, respectively. We find that A_S and D are very weakly correlated ($\rho = 0.16$ and 0.077 for EMU and 3CRR, respectively), and that there are no sources with both high disorder and low asymmetry. Similarly, we find that C and D are weakly correlated ($\rho = 0.27$ for both EMU and 3CRR), and there are no sources with both high disorder and low concentration. Highly disordered structures tend to coincide with more asymmetric or diffuse features, making it less likely for a source to simultaneously have low asymmetry or high concentration. From the KDE distributions in Figure 11, while the FRI sources appear to have slightly larger values of D than FRII sources, this distinction is not clear. This could be due to the lower jet power and brightness of the FRI sources than those FRII sources. The lobe structure may therefore be more sensitive to noise fluctuations, increasing the overall disorder of a lobe. The lack of FR separation in C and D suggest that either these metrics may not be sufficiently sensitive to hotspots and jet power, or that other factors, such as variations in lobe brightness distribution and environmental effects, play a dominant role in determining concentration and disorder.

We find D and E are moderately correlated ($\rho = 0.57$ and 0.51 for EMU and 3CRR sources, respectively). The compar-

Figure 16. Enlarged version of the A_S-D sub-panel from Figure 11. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses.

Figure 18. Enlarged version of the D-E sub-panel from Figure 11. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses.

Figure 17. Enlarged version of the C-D sub-panel from Figure 11. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses.

ison between the D and E metrics in Figure 18 highlights a general logarithmic envelope, with no sources exhibiting both high elongation and low disorder. Highly elongated jets or lobes will be more susceptible to environmental interactions, turbulence, and instabilities, all of which can contribute to increased disorder. There is also no clear separation between FR types in this parameter space. This suggests that disorder and elongation alone are not primary distinguishing factors in the classical FR classification. Instead, these metrics may be more reflective of environmental influences and source dynamics rather than intrinsic differences between FR types.

4.2 Source Metrics

As before with the single-lobe calculations, Figure 19 shows the pairwise combinations of the source metrics (ABL, $A_{S,p}$, $A'_{S,p}$, $A_{R,p}$, and $A'_{R,p}$), again with the normalised KDE along the diagonal. We can again see that the EMU and 3CRR sources occupy the same areas of parameter space, with differentiation between FRI and FRII sources in each of the asymmetry metrics. FRIIs tend to be more asymmetric than the FRIs. This is likely due to the different structure around the hotspots in each

of the radio lobes produced from the higher power jets than in FRI sources. Further, the brightest parts of a FRI source tend to be near the core, where the jets are still highly collimated. This, in conjunction with the limited number of resolution elements for some sources (see Section 4.1.1), results in lower asymmetry for the FRI sources. The features observed when comparing the shape and ratio asymmetries for the source asymmetry calculations are notably similar to those seen in the single-lobe asymmetry results. We expect both lobes of a radio source to have similar shape and structure when produced from jets of the same power and in the same environment. For the small number of cases where the environment is not consistent for both lobes (e.g., if the galaxy is on the edge of a cluster), than we expect to see higher asymmetry between lobes. We do however see some variation when comparing the flux-weighted (A_S and A_R) and normalised versions (A'_S and A'_R) of the asymmetry metrics. Looking at the KDEs for each, we can see that the flux-weighted calculations show a wider spread of values than the normalised ones. This is likely a direct result of the difference in brightness between the two lobes. This difference could be from jets with differing power or a projection effect with the emission of the closer lobe being brighter. This analysis supports the findings of Rodman et al. (2019), who studied the relationship between length asymmetries of radio lobe and the environment properties of the host AGN. They found that the FRI type sources tended to have smaller asymmetries than FRII sources, and that there was a stronger anti-correlation between length ratio and ambient galaxy density than with lobe luminosity and environment. However, we acknowledge that this study had a limited sample size (6 FRI, 16 FRII, 1 hybrid), with limited positive results for FRIs. Our findings are also consistent with Yates–Jones et al. (2021).

4.2.1 Angle Between Lobes

Figure 20 shows that the majority of EMU and FRII sources show little to no bending. There is a distinct separation between FR types at an ABL of roughly 160°, with FRIs tend-
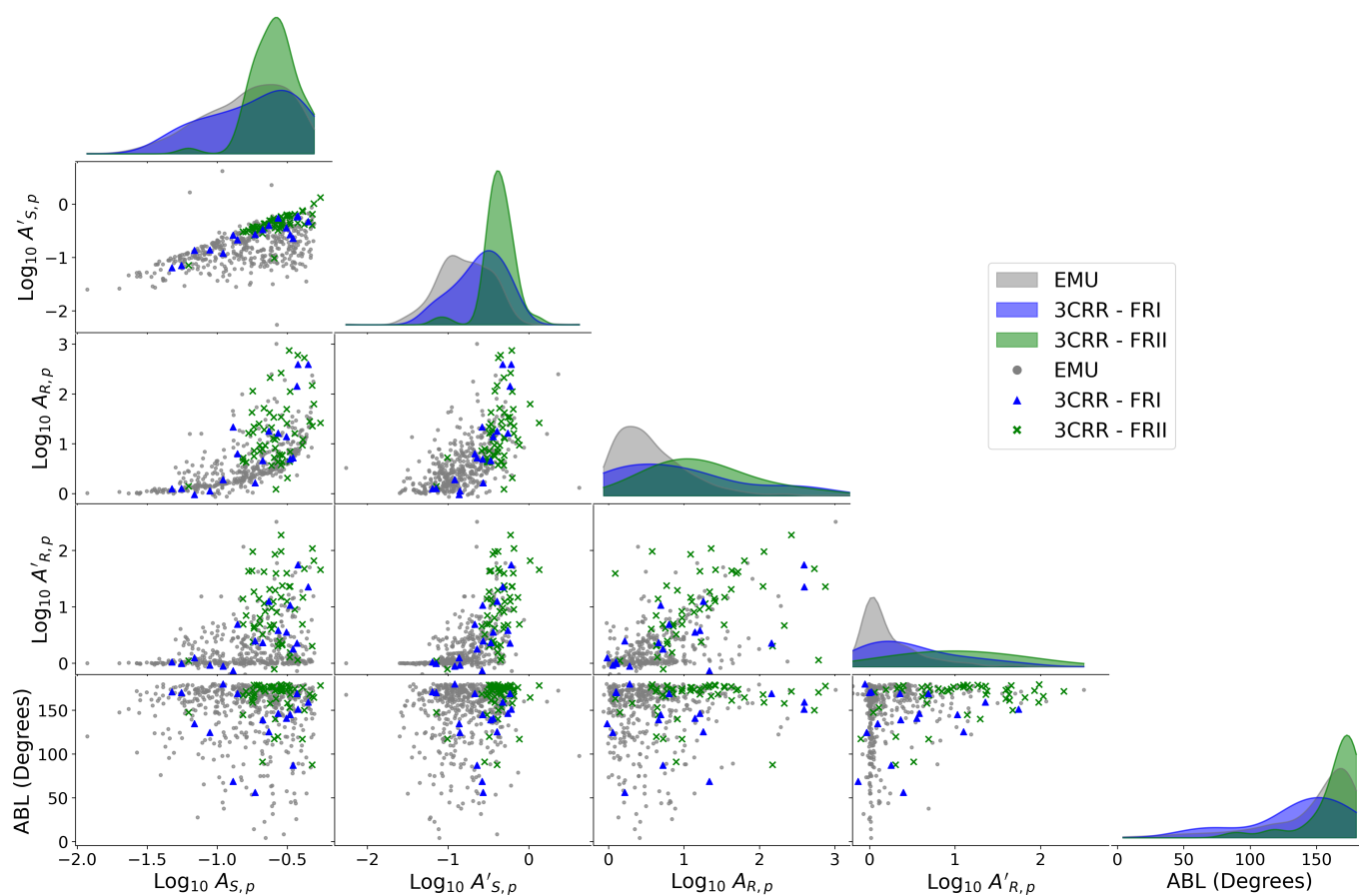
Figure 19. Pairwise comparisons of the source metrics (ABL, $A_{S,p}$, $A'_{S,p}$, $A_{R,p}$, $A'_{R,p}$) calculated on radio galaxies from the EMU-PS and 3CRR surveys. The scatter plots show the relationships between these metrics, with normalised KDEs along the diagonal displaying the distribution of each. EMU sources are represented by grey dots, and 3CRR sources colour-coded by their FR classification: FRI sources are shown with blue triangles and FRII sources with green crosses.

ing to be more bent than FRIIs. Specifically, we find that 53.21% of EMU sources and 76.36% of 3CRR FRII sources have an ABL > 150°, with only 29.41% of FRI sources with an ABL > 150°. This is to be expected as FRIs have lower power and are often found in higher density environments (like galaxy clusters) than FRIIs. In addition to the lower jet power, FRIs also tend to have slower jet speed. The jets are then more likely to be slowed down through entrainment (e.g., Laing & Bridle, 2014), therefore making them easier to bend (Banfield et al., 2015) than FRII sources.
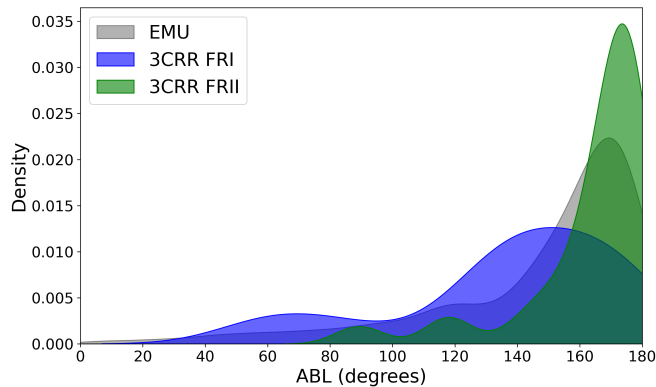
Figure 20. Enlarged version of the ABL KDE sub-panel from Figure 11. EMU sources are represented in grey, and 3CRR sources colour-coded by their FR classification: FRI sources are shown in blue and FRII sources in green.

4.3 Data Clustering in Parameter Space

To confirm that the metrics presented in this work are connected with physically meaningful aspects of radio galaxy morphology, we used Gaussian Mixture Modelling (GMM; VanderPlas, 2016), via the `sklearn.mixture GaussianMixture` package. GMM models the data as a probabilistic mixture of Gaussian distributions, where each cluster is characterised by a mean and covariance structure rather than a strict threshold in any single metric (see e.g., Bishop, 2006). This allows the clustering process to account for both individual contributions and correlations between metrics, meaning that the separation of clusters emerges from their joint distribution rather than from any one dominant parameter. In this way, GMM can identify if there are multiple populations within a dataset. Before applying GMM, we normalised all metrics to a standard scale spanning a range of −1 to 1 with a mean of 0, to ensure that each feature contributes equally to the clustering process.

In order to compare against FRI and FRII populations, we specified the GMM to produce two clusters, treating the *ABCDE* and source metrics as features. We did GMM clustering with the single-lobe metrics (*ABCDE*) and source metrics separately. This was because the source measurements contain half as many data points as the single-lobe measurements (i.e., one measurement per source rather than one per lobe). Doubling the source metrics to match the number of single-lobe measurements would artificially inflate their influence in the clustering process, making subsequent interpretation more ambiguous.

Apart from a mean vector and a covariance matrix, GMM also returns a mixture coefficient (weight) that reflects the prior probability of a data point belonging to that component. These mixture weights are not associated with individual features but instead represent the relative prevalence of each Gaussian component in the data. To explore how each feature (or metric in this case) contributes to the separation between clusters, we examined the differences between means of each feature across the two clusters. This provides a simple estimate of which features vary most between clusters, offering insight into their relative contribution to the clustering structure. Although this is not a formal measure of feature importance, it allows us to visualise which metrics show the strongest differences between clusters. This is explored for the single-lobe metrics in Figure 21. We can see that the metrics A_S, A_R, B, and D have the highest influence in creating the clusters, with C and E showing little to no contribution. For the source metrics (Figure 22), metrics $A_{S,p}$, $A'_{S,p}$ and ABL, have the highest influence in creating the clusters.
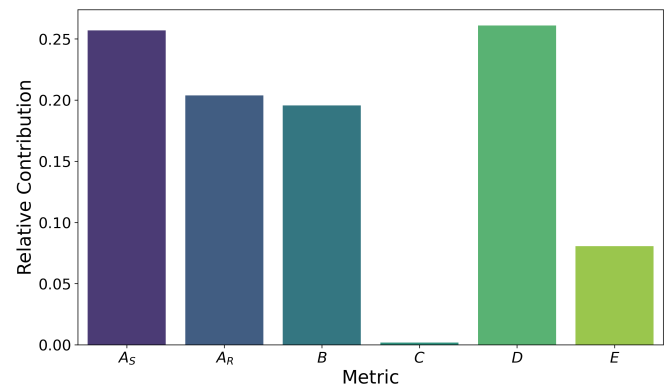
Figure 21. Comparison of the relative contributions of single-lobe metrics in distinguishing between GMM clusters.
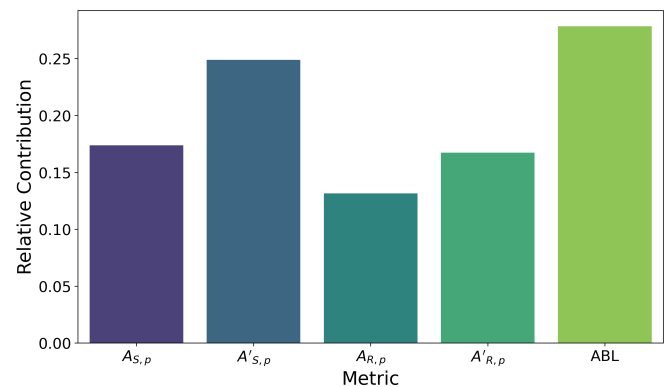
Figure 22. Comparison of the relative contributions of source metrics in distinguishing between GMM clusters.

Figure 23 compares the distributions of the GMM clusters with the FR distributions of the single-lobe metrics from Figure 11. The GMM clusters match closely with the FR distributions (cluster 1 with FRI, cluster 2 with FRII) in our metric space. The cluster and FR distributions are particularly close

in the A_R, and D metrics, reflecting the relative importance of these metrics in the clustering process. The clusters made with the source metrics also align closely with FR distributions (Figure 24), particularly in $A_{R,p}$ and $A'_{R,p}$. This suggests that while the GMM relies more heavily on $A_{S,p}$ and $A'_{S,p}$ to separate the data probabilistically, the actual FR classification differences manifest more strongly in the $A_{R,p}$ and $A'_{R,p}$ distributions. The results highlighted in Figures 23 and 24 confirm that both the *ABCDE* and source metrics are physically motivated and are not arbitrary shape analysis calculations.

To assess the performance of the clustering in distinguishing between FR morphologies, we compute completeness and reliability for each class. We define completeness as the fraction of sources of a given FR type that are correctly assigned to the corresponding cluster. Specifically for FRI sources, completeness is defined as the number of FRI sources in the FRI-predominant cluster divided by the total number of FRI sources. Reliability quantifies how pure a given cluster is with respect to an FR type. We define it as the fraction of sources in the FRI-predominant cluster that are actually classified as FRI. The same definitions apply to FRII sources. A high completeness value indicates that most FRI (or FRII) sources are correctly assigned, while a high reliability value suggests that the cluster is largely free from contamination by the other FR type. These values, for both single-lobe and source metrics, are compiled in Table 5. On average, the clustering is more successful at grouping the FRII sources together, likely due to larger number of FRIIs in our sample.

Table 5. Completeness and reliability of GMM clusters mapped to FR classifications, using the single-lobe (*ABCDE*) and source metrics (ABL, $A_{S,p}$, $A'_{S,p}$, $A_{R,p}$, $A'_{R,p}$).

Metric Type	Completeness (%)		Reliability (%)	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Single-Lobe	70.6	78.2	50.0	89.6
Source	58.8	89.1	62.5	87.5

To further quantify the similarity between distributions of the FR types with their respective clusters, we perform the Kolmogorov–Smirnov (KS, Kolmogorov, 1933; Smirnov, 1948) test. Specifically we use the two-sample KS test, which takes the null hypothesis that two distributions come from the same population. If the p-value of the KS test is small ($p < 0.05$), we can conclude that the two distributions were sampled from different underlying populations. We used the KS test to compare FRI with cluster 1 and FRII with cluster 2, compiling the resulting p-values in Table 7. We can see that for all metrics $p \gg 0.05$, consistent with the null hypothesis, that both are drawn from the same population. This further reinforces that the sources in cluster 1 are consistent with the properties of FRIs, and those in cluster 2 with FRIIs.

While we do not expect a perfect 1:1 correlation between the GMM clusters and the FR classifications, the observed correspondence between them in this simple comparison strongly reinforces the value of quantifying radio galaxy structure

Table 6. Kolmogorov–Smirnov test results comparing FRI sources with Cluster 1 and FRII sources with Cluster 2 for both the single-lobe metrics (*ABCDE*), and source metrics.

| Metric	FRI vs Cluster 1	FRII vs Cluster 2
A_S	$p = 3.840 \times 10^{-1}$	$p = 1.000$
A_R	$p = 1.160 \times 10^{-1}$	$p = 1.000$
B	$p = 8.280 \times 10^{-1}$	$p = 7.020 \times 10^{-1}$
C	$p = 9.130 \times 10^{-1}$	$p = 1.000$
D	$p = 5.410 \times 10^{-1}$	$p = 4.240 \times 10^{-1}$
E	$p = 9.190 \times 10^{-1}$	$p = 1.000$
$A_{S,p}$	$p = 9.420 \times 10^{-1}$	$p = 1.000$
$A'_{S,p}$	$p = 9.070 \times 10^{-1}$	$p = 1.000$
$A_{R,p}$	$p = 6.640 \times 10^{-1}$	$p = 9.990 \times 10^{-1}$
$A'_{R,p}$	$p = 3.650 \times 10^{-1}$	$p = 9.920 \times 10^{-1}$
ABL	$p = 2.790 \times 10^{-1}$	$p = 9.520 \times 10^{-1}$

through the *ABCDE* and source metrics. The fact that these metrics can probabilistically separate sources into distinct clusters that correspond to known FR types (to some degree) demonstrates the utility of these metrics as a flexible tool for classifying and quantifying radio galaxy morphology.

Table 7. Kolmogorov–Smirnov test results comparing FRI sources with Cluster 1 and FRII sources with Cluster 2 for both the single-lobe metrics (*ABCDE*), and source metrics.

Metric	FRI vs Cluster 1	FRII vs Cluster 2
A_S	$p = 3.840 \times 10^{-1}$	$p = 1.000$
A_R	$p = 1.160 \times 10^{-1}$	$p = 1.000$
B	$p = 8.280 \times 10^{-1}$	$p = 7.020 \times 10^{-1}$
C	$p = 9.130 \times 10^{-1}$	$p = 1.000$
D	$p = 5.410 \times 10^{-1}$	$p = 4.240 \times 10^{-1}$
E	$p = 9.190 \times 10^{-1}$	$p = 1.000$
$A_{S,p}$	$p = 9.420 \times 10^{-1}$	$p = 1.000$
$A'_{S,p}$	$p = 9.070 \times 10^{-1}$	$p = 1.000$
$A_{R,p}$	$p = 6.640 \times 10^{-1}$	$p = 9.990 \times 10^{-1}$
$A'_{R,p}$	$p = 3.650 \times 10^{-1}$	$p = 9.920 \times 10^{-1}$
ABL	$p = 2.790 \times 10^{-1}$	$p = 9.520 \times 10^{-1}$

4.4 Computational Efficiency

To estimate how computationally efficient our algorithm is, we timed how long it took to compute both the *ABCDE* and source metrics for all images in the EMU and 3CRR datasets. These computations were performed on an Apple MacBook Pro with an M4 chip and 16 GB of RAM, running in a Python notebook. We repeated this process 10 times for both datasets, then took the average. We obtained an average computation time of 10.56 seconds for the 480 EMU images, averaging 0.022 seconds per image. For the 72 3CRR images, we found an average computation time of 5.02 seconds, averaging 0.070 seconds per image. The average number of pixels in the EMU images was 3, 596.88, whereas for the higher-resolution 3CRR
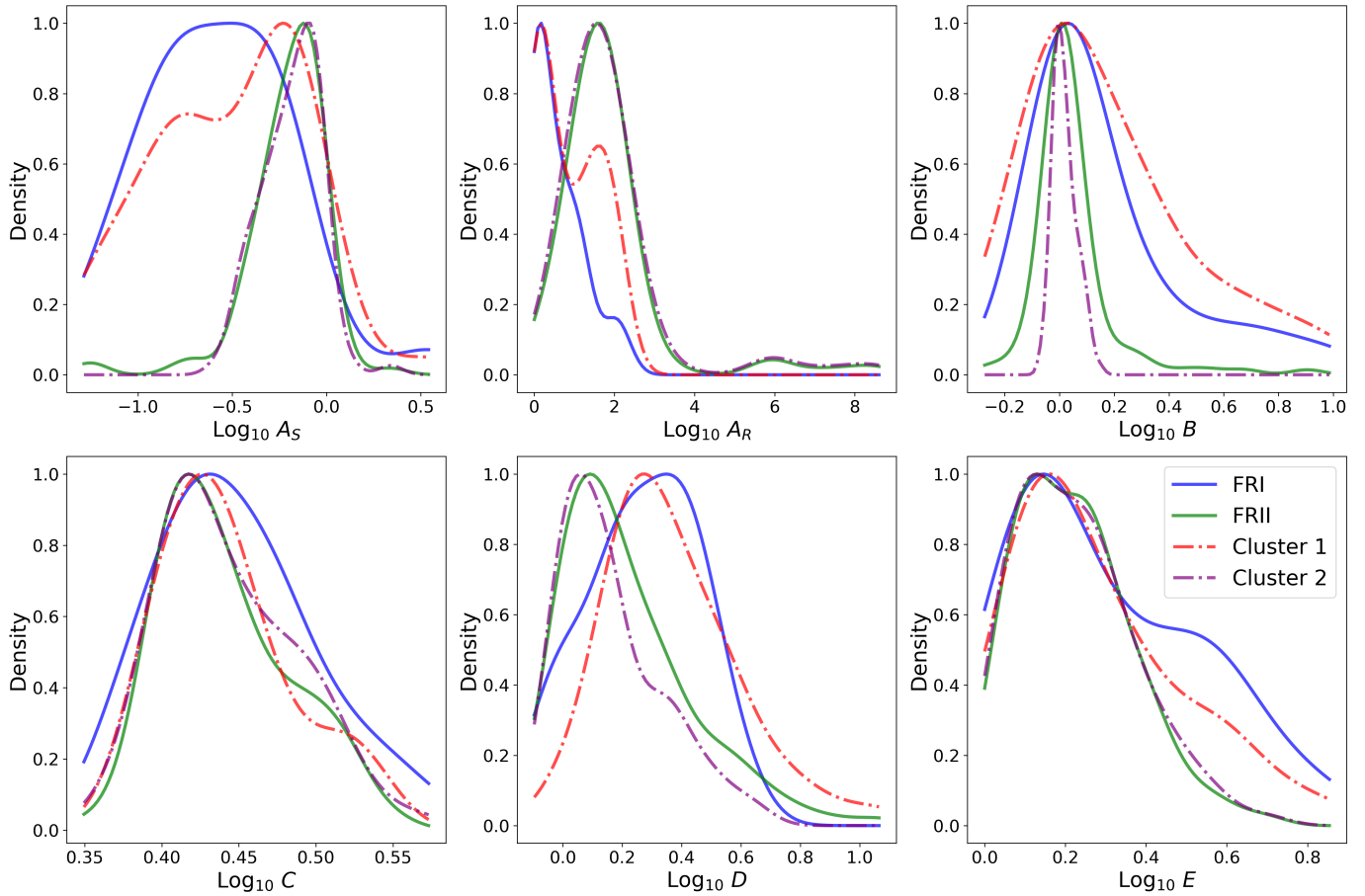
Figure 23. Comparison of the distributions of the GMM clusters with the FR distributions of the single-lobe metrics from Figure 11. Blue and green lines correspond to the distributions of FRI and II sources respectively. Red and purple dashed lines correspond to GMM clusters 1 and 2 respectively. Note how for A_R and D, the distributions are quite different for FRIs and FRIIs, and are well-matched to the Cluster 1 and 2 distributions, respectively.

images, the average number of pixels was $115,729.85$. The difference in computation time between EMU and 3CRR is therefore likely dependent on the number of pixels in the image. This demonstrates that the algorithm performs better than linear, as a roughly $30\times$ increase in the number of pixels resulted in only a $\sim 3\times$ increase in computation time on average. This efficiency highlights the algorithm's scalability, making it feasible for large-scale surveys like EMU. Further optimisations, such as parallelising computations or leveraging GPU acceleration, may improve performance even further.

5. Conclusions

We have introduced a series of quantitative metrics to describe the structure of a single radio galaxy lobe, and of the whole DRAGN. These include two asymmetry calculations, blurriness, concentration, disorder, and elongation ($ABCDE$), asymmetry and angle between lobes (source metrics). We developed this series of metrics metrics to be applied to large radio datasets, such as those that will be produced by EMU, and surveys from MeerKAT, LOFAR, SKA (in future), in an automatic and more efficient way than current machine learning approaches. We applied our metrics to two datasets: 480 sources from the EMU-PS and 72 sources from the 3CRR catalogue. We propose that

these metrics offer the prospect of deeper insights into both the structure and astrophysical processes shaping the radio galaxy lobes than the earlier Fanaroff-Riley morphology classification system (illustrative examples are shown in Figure 25). These metrics can also be incorporated into the #*tag* system proposed by Rudnick (2021). While the current set of metrics presented in this work provides an intuitive and effective starting point for quantifying complex radio structures, it is by no means exhaustive.

We found that the our metrics are largely independent of each other with Spearman rank correlation coefficient of $-0.01 < \rho < 0.3$ for all metric combinations except A_S-A_P and D-E, which had a ρ of 0.72 and 0.57, respectively. While the quantitative values for some of these metrics will change with the resolution (see Section 3.8), their qualitative performance remains consistent. The minimum image area of 20 beams for the EMU sources resulted in them being sufficiently resolved for the metric results to be comparable with the 3CRR data, evidenced by the metric results of both datasets occupying the same areas of parameter space. Further, a source with higher surface brightness will be less susceptible to noise biases and therefore should produce more robust results. We therefore acknowledge that this series of metrics will tend to have better
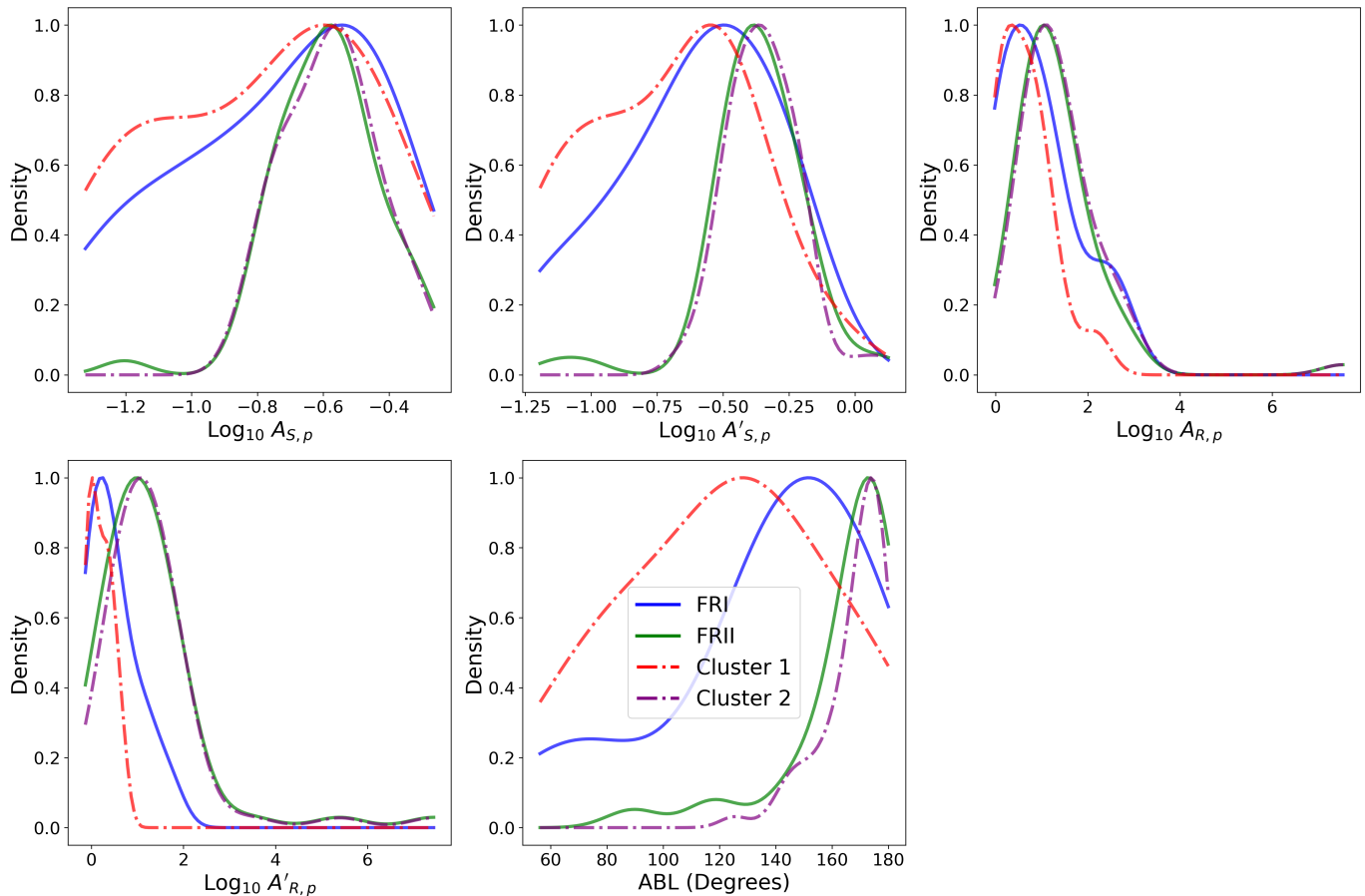
Figure 24. Comparison of the distributions of the GMM clusters with the FR distributions of the source metrics from Figure 11. Blue and green lines correspond to the distributions of FRI and II sources respectively. Red and purple dashed lines correspond to GMM clusters 1 and 2 respectively. Again, note how for $A'_{S,p}$, $A'_{R,p}$ and ABL, the difference in distributions between FRIs and FRIIs, that are well approximated by the distributions of Clusters 1 and 2, respectively.

performance on well resolved sources with high signal-to-noise, on average. In both single-lobe and source calculations, we find that the FRII types tend to have higher asymmetry than the FRIs, likely due to the differences in radio power and the structural features around hotspots. This could also be due to a potential selection effect. For FRI sources, a lack of sensitivity will mean that only the jets will be seen, which is expected to be symmetric before the jets becomes disrupted, whereas the edge-brightened nature of FRII sources allow us to see more complex structure of the lobe. Additionally, the flux-weighted asymmetry metrics in the source calculations show more scatter when compared to the normalised versions, indicating some brightness differences between the lobes.

We performed clustering analysis on the 3CRR data using Gaussian mixture modelling (GMM) with the single-lobe *ABCDE* metrics and source metrics separately. We find the distributions of the GMM clusters align with the distributions of the FR classifications in our parameter space, reinforcing the physical relevance of these metrics. While we do not expect a perfect one-to-one correspondence between the clusters and FR types, the observed overlap highlights the effectiveness of this quantitative approach. Notably, the clustering was more successful in grouping FRII sources, likely due to the larger

sample size in our data.

We find that ABL is a relatively important source metric for radio galaxy morphology. The ABL measurements show a clear distinction between FR types, with FRI sources being significantly more bent than FRII sources. This is consistent with the expectation that FRI sources, which are typically found in higher-density environments and often have lower jet power, experience greater interactions with the surrounding medium. While ABL does not directly fit into the single-lobe *ABCDE* framework, it plays a significant role in characterising source morphology. Additionally, our clustering analysis shows that ABL, along with $A_{S,p}$ and $A'_{S,p}$, has a strong influence on how clusters are assigned through GMM. ABL is therefore a key parameter in defining the source structure, highlighting the need to consider both single-lobe and source metrics when analysing radio galaxy morphology.

This series of metrics can be particularly valuable in the era of large-scale surveys, where millions of sources will need to be analysed in an automated and efficient manner. By capturing key aspects of radio lobe structure, our metrics can reveal critical information about AGN and the astrophysical processes driving jet formation, particle acceleration, and interactions between the radio source and the surrounding environment.
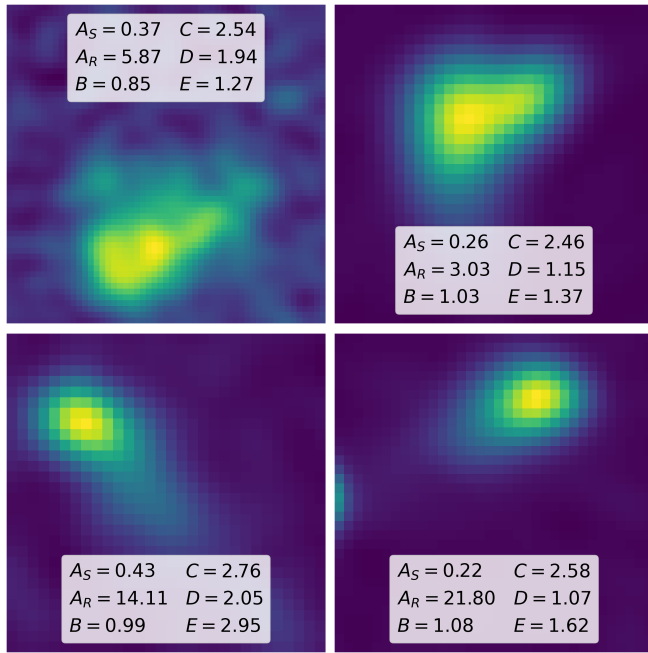
Figure 25. Multi-panel comparison of single lobes from different sources in the EMU-PS. Visual differences in each lobe structure are highlighted by different values of the *ABCDE* metrics in the legend.

Further, the metrics presented in this work are agnostic to the changing emphases in radio galaxy interpretation. They do not depend on specific morphological labels, such as FRI or FRII. As a result, they can be used to efficiently characterise large surveys without being constrained by predefined classifications. Once characterised, different combinations of metrics can then be applied to identify new and interesting types of sources.

6. Future work

Future work will focus on refining the implementation of the metrics, to reduce the effect of sources with limited sampling. This could be done by implementing a region-based criteria for lobe detection, instead of using the brightest pixel, by detecting extended regions of high flux and using contour-based methods to locate the true lobes. This could enable, for example, the modification of the definitions of asymmetry metrics to span the full extent of a lobe plus jet, providing a more astrophysically intuitive measurement. This work does not explicitly exclude low surface brightness sources in order to investigate how the metrics perform on a range of source types. Further investigation, perhaps through binning sources by brightness, can provide a more robust understanding of how the metrics perform with varying source brightness. The inclusion of new metrics into our current framework may also improve the ability to quantify complex radio structures, offering the prospect of deeper insights about radio galaxy morphology. Applying these metrics to a dataset with a larger sample size, such as MiraBest[b] (Porter & Scaife, 2023) is also a crucial step in the testing of the metrics.

[b]https://doi.org/10.5281/zenodo.4288837

A key direction will be further exploring the relationship between the the metrics presented in this work and the underlying astrophysical processes. Linking these metrics to characteristics such as AGN properties, galaxy number density, and host galaxy properties (such as mass, colour, optical morphology, among others), will enable a more comprehensive understanding of the forces shaping radio galaxies.

Multi-frequency analysis of radio sources will be crucial in future studies. By examining how the metric values vary across different radio frequencies, we can gain a clearer picture of how various physical processes, such as different emission mechanisms, environmental impacts, and jet composition and its dynamics, manifest at different frequencies. This will not only enhance our understanding of radio galaxies and their classifications but also provide valuable insights into the broader mechanisms driving galaxy evolution.

References

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18

Abraham, R. G., Tanvir, N. R., Santiago, B. X., et al. 1996a, MNRAS, 279, L47

Abraham, R. G., Valdes, F., Yee, H. K. C., & van den Bergh, S. 1994, ApJ, 432, 75

Abraham, R. G., van den Bergh, S., Glazebrook, K., et al. 1996b, ApJS, 107, 1

Alam, A., Pimbblet, K., & Gordon, Y. 2025, PASA, 42, e013

Andrianomena, S., & Tang, H. 2024, JCAP, 2024, 034

Archer, S., Baldwin, J. E., Edge, D. O., et al. 1959, URSI Symp. 1: Paris Symposium on Radio Astronomy, 9, 487

Azzopardi, G., & Petkov, N. 2013, IEEE TPAMI, 35, 490

Ball, N. M., & Brunner, R. J. 2010, IJMPD, 19, 1049

Banfield, J. K., Wong, O. I., Willett, K. W., et al. 2015, MNRAS, 453, 2326

Bastien, D. J., Scaife, A. M. M., Tang, H., Bowles, M., & Porter, F. 2021, MNRAS, 503, 3351

Becker, B., & Grobler, T. 2019, in 2019 International Multidisciplinary Information Technology and Engineering Conference, 1–8

Becker, B., Vaccari, M., Prescott, M., & Grobler, T. 2021, MNRAS, 503, 1828

Bennett, A. S. 1962, Mem. R. Astron. Soc., 68, 163

Bennett, A. S., & Simth, F. G. 1962, MNRAS, 125, 75

Bershady, M. A., Jangren, A., & Conselice, C. J. 2000, AJ, 119, 2645

Best, P. N. 2009, Astron. Nachr., 330, 184

Best, P. N., & Heckman, T. M. 2012, MNRAS, 421, 1569

Bishop, C. M. 2006, Pattern Recognition and Machine Learning, 1st edn. (Springer New York, NY)

Boyce, M. M., Hopkins, A. M., Riggi, S., et al. 2023, PASA, 40, e028

Brand, K., Grobler, T. L., Kleynhans, W., et al. 2023, MNRAS, 522, 292

Brinchmann, J., Abraham, R., Schade, D., et al. 1998, ApJ, 499, 112

Brough, S., Croom, S., Sharp, R., et al. 2013, MNRAS, 435, 2903

Carilli, C. L., & Rawlings, S. 2004, New Astron. Rev., 48, 979

Conover, W. J. 1999, Practical nonparametric statistics, Vol. 350 (John Wiley & Sons)

Conselice, C. J. 1997, PASP, 109, 1251

Conselice, C. J. 2003, ApJS, 147, 1

Conselice, C. J., Bershady, M. A., & Jangren, A. 2000, ApJ, 529, 886

Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E., & Archer, S. 1959, Mem. R. Astron. Soc., 68, 37

Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P

Fayyad, U. M., Smyth, P., Weir, N., & Djorgovski, S. 1995, J. Intell. Inf. Syst., 4, 7

Ferrari, F., Carvalho, R. R. d., & Trevisan, M. 2015, ApJ, 814, 55

Galvin, T. J., Huynh, M. T., Norris, R. P., et al. 2020, MNRAS, 497, 2730

Gordon, Y. A., Rudnick, L., Andernach, H., et al. 2023, ApJS, 267, 37

Gupta, N., Hayder, Z., Norris, R. P., Huynh, M., & Petersson, L. 2024a, PASA, 41, e001

Gupta, N., Huynh, M., Norris, R. P., et al. 2022, PASA, 39, e051

Gupta, N., Hayder, Z., Norris, R. P., et al. 2023, PASA, 40, e044

Gupta, N., Norris, R. P., Hayder, Z., et al. 2024b, PASA, 41, e027

Hancock, P. J., Murphy, T., Gaensler, B. M., Hopkins, A., & Curran, J. R. 2012, MNRAS, 422, 1812

Haralick, R. M., Shanmugam, K., & Dinstein, I. 1973, IEEE Trans. Syst. Man. Cybern., SMC–3, 610

Hardcastle, M. J., & Croston, J. H. 2020, New A Rev., 88, 101539

Hopkins, A. M., Kapinska, A., Marvil, J., et al. 2025, Publications of the Astronomical Society of Australia, 1–32

Javaherian, M., Miraghaei, H., & Moradpour, H. 2023, AJ, 166, 124

Kapińska, A. D., Terentev, I., Wong, O. I., et al. 2017, AJ, 154, 253

Kent, S. M. 1985, ApJS, 59, 115

Kolmogorov, A. 1933, Giorn Dell'inst Ital Degli Att, 4, 89

Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511

Krause, M., Alexander, P., Riley, J., & Hopton, D. 2012, MNRAS, 427, 3196

Laing, R. A., & Bridle, A. H. 2002, MNRAS, 336, 1161

Laing, R. A., & Bridle, A. H. 2014, MNRAS, 437, 3405

Laing, R. A., Riley, J. M., & Longair, M. S. 1983, MNRAS, 204, 151

Laing, R. A., Riley, J. M., & Longair, M. S. 2003, VizieR Online Data Catalog: Bright radio sources at 178 MHz (3CRR) (Laing+ 1983), VizieR On-line Data Catalog: J/MNRAS/204/151. Originally published in: 1983MN-RAS.204..151L

Leahy, J. P. 1993, Jets in Extragalactic Radio Sources, 421, 1

Leahy, J. P., Bridle, A. H., & Strom, R. G. 2013, An Atlas of DRAGNs

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163

Ma, Z., Xu, H., Zhu, J., et al. 2019, ApJS, 240, 34

Miley, G. 1980, Annu. Rev. Astron., 18, 165

Mingo, B., Croston, J. H., Hardcastle, M. J., et al. 2019, MNRAS, 488, 2701

Morganti, R. 2019, Proc. Int. Astron. Union, 15, 229

Ndung'u, S., Grobler, T., Wijnholds, S. J., Karastoyanova, D., & Azzopardi, G. 2023, New Astron. Rev., 97, 101685

Ndung'u, S., Grobler, T., Wijnholds, S. J., Karastoyanova, D., & Azzopardi, G. 2024, MNRAS, 530, 783

Norris, R. P., Yew, M., Crawford, E., et al. 2025, PASA, in preparation

Norris, R. P., Hopkins, A. M., Afonso, J., et al. 2011, PASA, 28, 215

Norris, R. P., Marvil, J., Collier, J. D., et al. 2021, PASA, 38, e046

Ntwaetsile, K., & Geach, J. E. 2021, MNRAS, 502, 3417

Orlov, N., Shamir, L., Macura, T., et al. 2008, Pattern Recognit. Lett., 29, 1684

Porter, F. A. M., & Scaife, A. M. M. 2023, RASTI, 2, 293

Proctor, D. D. 2011, ApJS, 194, 31

Riggi, S., Magro, D., Sortino, R., et al. 2023, Astron. Comput., 42, 100682

Rodman, P. E., Turner, R. J., Shabala, S. S., et al. 2019, MNRAS, 482, 5625

Rossi, P., Bodo, G., Massaglia, S., & Capetti, A. 2020, A&A, 642, A69

Rudnick, L. 2021, Galaxies, 9, 85

Sadeghi, M., Javaherian, M., & Miraghaei, H. 2021, AJ, 161, 94

Saikia, D. J. 2022, J. Astrophys. Astron., 43, 97

Saripalli, L., Subrahmanyan, R., Thorat, K., et al. 2012, ApJS, 199, 27

Schade, D., Lilly, S. J., Crampton, D., et al. 1995, ApJ, 451, L1

Segal, G., Parkinson, D., Norris, R. P., & Swan, J. 2019, PASP, 131, 108007

Segal, G., Parkinson, D., Norris, R., et al. 2023, MNRAS, 521, 1429

Shabala, S., & Alexander, P. 2009, ApJ, 699, 525

Shabala, S. S., Jurlin, N., Morganti, R., et al. 2020, MNRAS, 496, 1706

Shamir, L. 2009, MNRAS, 399, 1367

Siegel, S., & Castellan, N. 1988, Nonparametric Statistics for the Behavioral Sciences, McGraw–Hill international editions statistics series (McGraw–Hill)

Smirnov, N. 1948, Ann. Math. Stat., 19, 279

Teague, M. R. 1980, JOSA, 70, 920

Tolley, E. 2024, URSI RSL, 5, doi:10.46620/23–0028

VanderPlas, J. 2016, Python Data Science Handbook: Essential Tools for Working with Data, 1st edn. (O'Reilly Media, Inc.)

Véron, P. 1977, A&AS, 30, 131

Watson, A. 2012, Mathematica, 14, doi:10.3888/tmj.14–5

Wilson, A. S., & Colbert, E. J. M. 1995, ApJ, 438, 62

Wu, C., Wong, O. I., Rudnick, L., et al. 2019, MNRAS, 482, 1211

Xu, Z., Ahmad, S., Liao, Z., Xu, X., & Xiang, Z. 2023, J. Intell. Syst., 32, doi:10.1515/jisys-2023-0111

Yates-Jones, P. M., Shabala, S. S., & Krause, M. G. H. 2021, MNRAS, 508, 5239

Yates-Jones, P. M., Shabala, S. S., Power, C., et al. 2023, PASA, 40, e014

Appendix 1. Troublesome Images

The metric algorithm described above works on a series of assumptions. First, we require that an image is centred in the cutout, so that there is a lobe in each half (cutting the image perpendicular to and half–way along the major axis). This is not true for all sources, for example, those that are extremely bent. For these objects both lobes are in the same half of the cutout (e.g., Figure 26). For such sources, rather than our metrics quantifying some properties of a lobe, we would instead have some measurement of the core and perhaps jets, and either missing both lobes or possibly incorporating one of them. While the metric calculation would return some measurement, it would not be the measurement intended. Accordingly, through visual inspection, extremely bent sources were omitted from our analysis.

Second, we ideally want sources that are not core-dominated, so that the core is less luminous than the lobes. This is because our algorithm uses the brightest pixel as the base for many of the metric operations (A_S, A_R, B, C, for example). As mentioned in Section 3.1, the core removal process we implemented did not always work perfectly. For example, there were cases where an erroneous number of components were detected. If the number of components was under–estimated
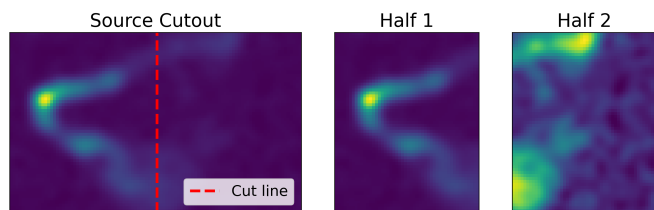
Figure 26. Example of an extremely bent source where both lobes are in one side of the cutout. Red dashed line indicates the midpoint of the major axis of the source image.

the core would not be masked out, and if it was over–estimated a lobe could be masked out (see Figure 27 for examples of each case). Both scenarios again would lead to the metric measurements not reflecting the actual lobe structure. Such sources were excluded from our analysis.
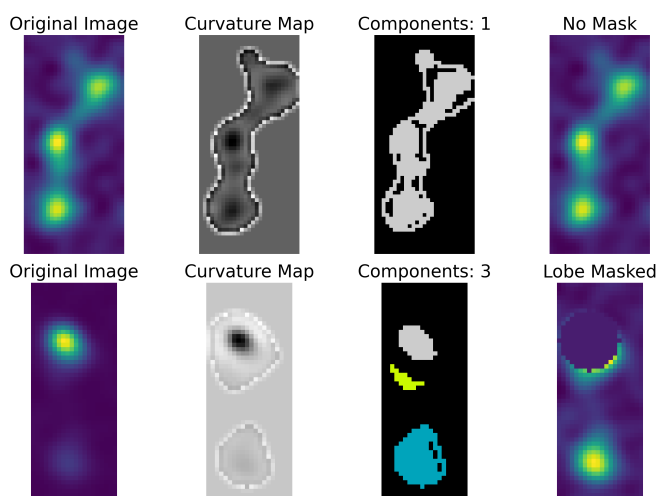

Figure 27. Illustrative examples of when the core-removal process fails. Top row: number of source components was underestimated, and no mask was applied. Bottom row: number of source components was overestimated, and a lobe was masked.

Appendix 2. Metric differences between lobes

To quantify how the metrics differ between lobes, we simply take the absolute difference of the $\log_{10}$ values the metric values for each lobe:

$$\Delta x = |log_{10}(x_1) - log_{10}(x_2)|, \tag{14}$$

where x_1 and x_2 are the metric values of lobe 1 and 2 respectively. In Figure 28 we show the range of differences between each lobe, across each of the $ABCDE$ metrics. We calculate the percentage of sources with $\Delta x < 0.3$ (Table 8). We can see for both EMU and 3CRR sources, the majority of lobe measurements are within a factor of two of each other for each of the $ABCDE$ metrics. The A_R metric appears to show the greatest degree of difference suggesting that it could be the most sensitive to variations between lobes.
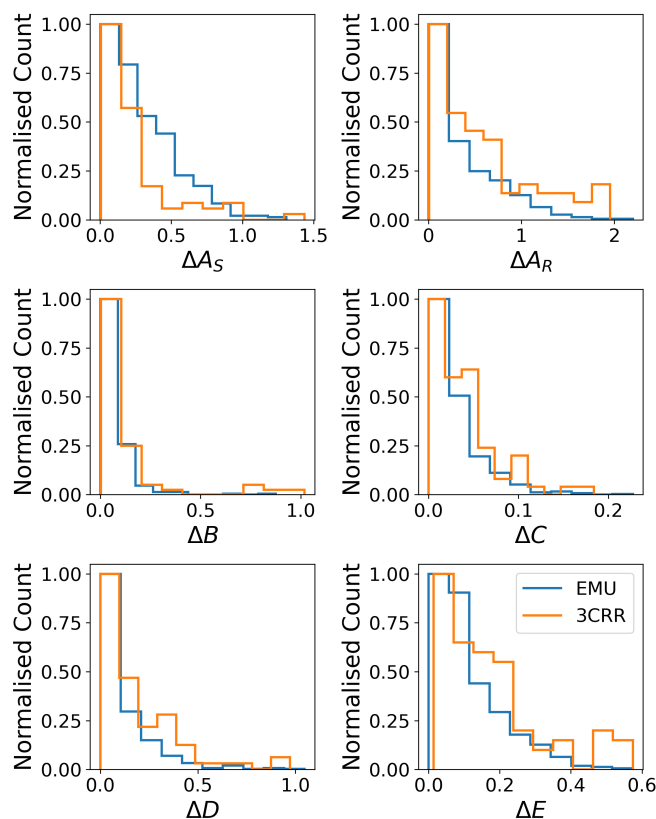
Figure 28. Normalised histograms of absolute differences (Δx) between single lobe metrics for EMU (blue) and 3CRR (orange) sources.

Table 8. Percentage of sources with $\Delta x < 0.3$ for both EMU and 3CRR datasets.

Δx	EMU (%)	3CRR (%)
ΔA_S	59.9	76.4
ΔA_R	58.2	40.3
ΔB	97.6	91.2
ΔC	100.0	100.0
ΔD	90.2	75.0
ΔE	93.5	83.3