







A Novel Method for Estimating the Ambient Medium Density Around Distant Radio Sources from Their Observed Radio Spectra

Anna Wójtowicz¹ , Łukasz Stawarz¹ , Jerzy Machalski¹ , and Luisa Ostorero^{2,3} 

¹ Astronomical Observatory of the Jagiellonian University, ul. Orła 171, 30-244 Kraków, Poland; awojtowicz@oa.uj.edu.pl

² Dipartimento di Fisica—Università degli Studi di Torino, Via P. Giuria 1, I-10125 Torino, Italy

³ Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy

Received 2020 November 21; revised 2021 June 22; accepted 2021 June 28; published 2021 November 30

Abstract

The dynamical evolution and radiative properties of luminous radio galaxies and quasars of the FR II type, are well understood. As a result, through the use of detailed modeling of the observed radio emission of such sources, one can estimate various physical parameters of the systems, including the density of the ambient medium into which the radio structure evolves. This, however, requires rather comprehensive observational information, i.e., sampling the broadband radio continua of the targets at several frequencies, and imaging their radio structures with high resolution. Such observations are, on the other hand, not always available, especially for high-redshift objects. Here, we analyze the best-fit values of the source physical parameters, derived from extensive modeling of the largest currently available sample of FR II radio sources, for which good-quality multiwavelength radio flux measurements could be collected. In the analyzed data set, we notice a significant and nonobvious correlation between the spectral index of the nonthermal radio emission continuum, and density of the ambient medium. We derive the corresponding correlation parameters, and quantify the intrinsic scatter by means of Bayesian analysis. We propose that the discovered correlation could be used as a cosmological tool to estimate the density of ambient medium for large samples of distant radio galaxies. Our method does not require any detailed modeling of individual sources, and relies on limited observational information, namely, the slope of the radio continuum between the rest-frame frequencies 0.4 and 5 GHz, possibly combined with the total linear size of the radio structure.

Unified Astronomy Thesaurus concepts: [Intergalactic medium \(813\)](#); [Active galactic nuclei \(16\)](#); [Fanaroff-Riley radio galaxies \(526\)](#); [Bayesian statistics \(1900\)](#)

Supporting material: machine-readable table

1. Introduction

High-power radio galaxies are believed to constitute a parent population of radio-loud quasars (e.g., Barthel 1989). Their large-scale radio structures are typically characterized by the edge-brightened morphology of the *classical-double* type, consisting of a pair of well-collimated jets terminating in bright, distinct hotspots, and surrounded by extended lobes that dominate the radiative outputs of systems at radio wavelengths (hereafter “FR II” sources, following Fanaroff & Riley 1974). There is an ongoing debate as to whether this large-scale morphology is determined solely by the kinetic power of the jets (and so, whether it depends only on the total radio power of a source), or also by the type of host galaxy (e.g., Ledlow & Owen 1996; Koziel-Wierzbowska & Stasińska 2011; Wing & Blanton 2011; Capetti et al. 2017; Mingo et al. 2019).

Furthermore, high-power radio galaxies at lower redshifts seem to avoid dense large-scale environments, especially when compared with local low-power radio galaxies that are often found in the centers of rich clusters (e.g., Zirbel 1997; Harvanek & Stocke 2002; Wing & Blanton 2011, and references therein). For example, a careful examination of the NRAO⁴ Very Large Array (VLA) imaging survey of Abell clusters at northern declinations (Owen et al. 1992; Owen & Ledlow 1997), when restricted to the systems with richness $R \geq 0$ and redshifts $z < 0.25$, returns only two FR IIs hosted by

the *brightest cluster galaxies* (Stawarz et al. 2014; Hagino et al. 2015; see also Cheung et al. 2019). In general, i.e., regardless of the classification of the host as a Brightest Cluster Galaxy, there still is a significant difference in the environment of galaxies displaying different radio morphologies. For example, very recently Croston et al. (2019) demonstrated that, among radio galaxies at $z < 0.4$, FR IIs are characterized by a lower cluster association fraction, and systematically poorer environments.

Whether the difference in the large-scale environment of low- and high-power radio galaxies in the local universe is significant or not, and whether this difference persists also in the higher- z universe, both remain open questions (see in this context, e.g., Hill & Lilly 1991; Belsole et al. 2007; Antognini et al. 2012; Massaro et al. 2019, 2020). Part of the problem here is that the X-ray observations of a sufficient depth and quality to allow one to estimate—via the detection of a thermal (free-free and line) emission component—the temperature and spatial distribution of the hot gaseous fraction of the intergalactic medium (IGM), are currently still rather sparse, and restricted to the nearest and/or the brightest systems (see Belsole et al. 2007; Ineson et al. 2017).

In this paper, we explore the possibility of using high-power FR II radio galaxies as cosmological probes of their environment, and in particular, of using basic observed properties of their radio emission continua to estimate the gas density of the ambient medium into which their radio structures evolve.⁵ Note

⁴ The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

⁵ We note that very recently Turner et al. (2020) proposed a method that through Bayesian inference provides PDFs for the most likely redshift of radio galaxies based on the radio-frequency imaging and photometry.

in this context, that since FR IIs are expected to reside in the external regions of galaxy clusters, as mentioned in previous paragraphs, an estimate of the surrounding gas density would likely yield an estimate of the density of a galaxy group that is falling into the cluster. The external regions of galaxy clusters are indeed rather inhomogeneous, in terms of both galaxy density and gas density (e.g., Rines et al. 2002; Coe et al. 2012; Geller et al. 2014; Yu et al. 2016; Haines et al. 2018; Sohn et al. 2019). A typical elliptical host of an FR II source (e.g., Sikora 2007; Zheng et al. 2020, and references therein) is likely to be the brightest galaxy member of such a small group (see Lin et al. 2010) that is captured by a cluster.

In principle, the general idea that FR IIs could be used in this or another way as probes of the distant universe may not sound surprising or novel. Indeed, the dynamical evolution of such sources is rather well understood (see Begelman & Cioffi 1989; Kaiser & Alexander 1997; Kawakatu & Kino 2006, and the next section) as is the production of nonthermal emission in their expanding radio lobes (e.g., Kaiser et al. 1997; Stawarz et al. 2008).

In the framework of this relatively simple scenario, the total kinetic power of the jets, along with properties of the ambient medium (density and pressure profiles), are key factors determining the volume, expansion velocities, and total energy content of the lobes, and hence, the spectral evolution of the lobes' nonthermal emission. As a result, through detailed modeling of the observed radio emission continua of FR IIs, using approximate prescriptions that combine dynamical evolution of large-scale structures with the particle and radiative transfer equations, one could hope to constrain the main physical parameters of the model, such as the ambient gas density. This is, however, a rather time-consuming procedure, requiring, in addition, comprehensive observational information, and as such could hardly be used when dealing with large samples of sources at large cosmological distances emerging from new-generation massive radio surveys, such as the Low-Frequency Array (LOFAR) Two-Meter Sky Survey survey (Hardcastle et al. 2019).

Here, we propose an alternative approach that enables us to avoid detailed modeling of individual sources, and is based on limited observational information. This method instead exploits a nonobvious correlation emerging between the best-fit values of radio spectral index, determined between two rest-frame frequencies, and the density of the ambient medium—hereafter referred to as *the model data*—derived from detailed dynamical-radiative modeling of a large sample of FR IIs with available multiwavelength radio flux-density measurements.

This paper is organized as follows: In Section 2, we describe the dynamical-radiative model used to derive the model data. In Section 3, we perform a correlation analysis on the model data, and the regression analysis by means of the Bayesian approach. In Section 4, we verify our findings by examining good-quality high-angular resolution X-ray data available for some of the sources from our sample. Final conclusions are given in Section 5.

2. Dynamical Modeling

In this paper, we explore the parameter space emerging from modeling the largest sample of FR II radio sources selected from various radio catalogs and surveys, namely, the Cambridge Catalogs of Radio Sources (3C–6C–5C), the Bologna Sky Surveys (B2–B3), and the Green Bank Sky Surveys

(GB–GB2), as described extensively in Machalski et al. (2021, ApJS, in press). Keeping in mind that our dynamical-radiative modeling (see below) deals exclusively with the extended lobes of classical doubles, and requires rather good-quality radio photometry and imaging in a wider frequency range, the following criteria have been applied when selecting the targets: (i) We included only the lobe-dominated sources previously identified as FR IIs, i.e., those classical doubles with the core/jet emission accounting for <5% of the total emission at radio frequencies; in other words, all the core-dominated sources present in the considered radio catalogs and surveys, even if classified as FR IIs, have been excluded from the analysis. (ii) We included only the sources that are resolved on the available radio maps sufficiently well, so that the projected linear sizes and axial ratios of the radio lobes (assuming cylindrical geometries) could be measured. (iii) We considered only the sources with a robust optical identification of the host galaxy, with measured spectroscopic redshift. (iv) We required well-characterized radio spectra of the sources, with flux densities measured at least at five observing frequencies, covering the frequency range between 74 MHz and 5 GHz, so that the two-point spectral index between the rest-frame frequencies 0.4 and 5 GHz could be determined.

There were no limits assigned for the redshift range, radio luminosities, or linear sizes of the sources included in the final sample. The redshifts of the selected objects span the range from 0.03 up to 4.41, with a mean value of 0.77.

The preliminary sample analyzed here consists of 271 targets, including 188 sources selected from the 3CRR, 6CE, and 7CRS in the compilation of Grimes et al. (2004), and 69 low-luminosity B2 and B3 sources from de Ruiter et al. (1986) and Fanti et al. (1986, 1987), supplemented with GB/GB2 middle-luminosity sources from Machalski (1998). Most of these sources (>80%) are identified as radio galaxies, with only a minority (<20%) classified spectroscopically as quasars.

For the modeling procedure, we use the DYNAGE algorithm of Machalski et al. (2007). This code has already been applied and verified several times to individual double radio sources belonging to various classes and types, such as giant radio galaxies, double-double radio galaxies, X-shaped radio galaxies, and luminous radio quasars (e.g., Machalski et al. 2010, 2011, 2016; Bhatta et al. 2018; O'Sullivan et al. 2019). It is based on the Kaiser & Alexander (1997) analytical model for the dynamical evolution of classical doubles, augmented by the prescription for energy evolution of the lobes' nonthermal emission by Kaiser et al. (1997).

In particular, the Kaiser & Alexander (1997) model describes a self-similar expansion of a radio cocoon powered over the time τ by a pair of jets with the total kinetic luminosity Q , and expanding into the hot gaseous environment whose density is modeled with a β profile

$$\rho(r) = \rho_0 \times (r/a_0)^{-\beta}, \quad (1)$$

where $r > a_0$ is the radial distance from the center of the host galaxy. The entire bulk kinetic energy carried away by the jets from the radio core is converted, at the termination shocks (*hotspots*), into the internal energy of a magnetized plasma, which subsequently inflates the radio cocoon, forming, in this way, extended radio lobes, and providing the pressure support for the cocoon's sideways expansion. At the shock front, the jets' particles attain a nonthermal energy distribution, approximated in the model by a single power law with the energy

index s , so that the corresponding *injection* spectral index of the hotspots' synchrotron emission (in the optically thin segment of the spectrum) is $\alpha_{\text{inj}} = (s - 1)/2$. Further downstream the shock, i.e., deep within the cocoon, particles lose their energies due to radiative and adiabatic cooling; as a result, the radio continuum, representing the integrated radiative output of the entire lobes, steepens at higher frequencies to $\alpha > \alpha_{\text{inj}}$.

The numerical code DYNAGE has been developed to solve the inverse problem, i.e., to determine the four main free parameters of the model, namely, the jet kinetic power Q , the jet lifetime τ , the injection index α_{inj} , and the central density of the gaseous environment ρ_0 , by fitting the model to a given set of observables, including the linear size of the lobes, their volume, as well as the slope and normalization of the observed radio continuum. This fitting procedure yields a unique set of best-fit parameters. Obviously, the model adopts several crude approximations. For example, we assume the jet inclinations as either 70° or 90° for most of the radio galaxies in the sample, while in the case of quasars we accommodate even smaller values, all as guided by the lobe/counter-lobe radio appearance in individual targets (surface brightness asymmetry, etc.). Also, we consider a universal slope for the ambient gas density profile $\beta = 3/2$. We emphasize, in this context, that the β model for the IGM density distribution, assumes in general a hydrostatic equilibrium for the isothermal hot gas and galaxies (King 1962; Cavaliere & Fusco-Femiano 1976); when applied to galaxy clusters' X-ray surface brightness profiles, the resulting best-fit value for the β parameter (as defined in Equation (1) above) are, indeed, typically within the range between 1.2 and 2.5 (see, e.g., the recent analysis by Käfer et al. 2019). We also note that, in the framework of the Kaiser & Alexander (1997) model, in which the radio structure is assumed to evolve in a power-law external density profile, a_0 and ρ_0 are not independent parameters, and therefore the model fit returns in fact the best-fit value for the characteristic quantity $\rho_0 a_0^\beta$ rather than ρ_0 alone.

Most importantly, the model treats α_{inj} as a *free* parameter. This parameter should, in fact, be considered as an *effective* injection spectral index, i.e., the injection index averaged over a broader spectral range and over the lifetime of a source, since the particle energy spectrum formed at the termination shocks may evolve with time, and may be far more complex than a single power law (see the discussion in Machalski et al. 2007). The integrated spectral index of the lobes' radio emission, α , within a given frequency range, depends obviously on α_{inj} , but also on the other model parameters that determine the energy evolution of the radiating particles within the lobes, $\alpha = f(\alpha_{\text{inj}}, Q, \tau, \rho_0)$. In other words, the spectral curvature of the lobes' radio continuum at a given moment τ , encodes the information on the injection index, as well as on the general evolutionary history of a source.

Within this framework note that, in the Kaiser & Alexander (1997) model, the linear size of the source, D , scales with the source lifetime as $D \propto (Q/\rho_0 a_0^\beta)^{1/(5-\beta)} \tau^{3/(5-\beta)}$, so that with the adopted $\beta = 3/2$ the ambient medium density at distances $r \simeq D/2$

$$\log \rho \propto -5 \log D + \log Q + 3 \log \tau. \quad (2)$$

On the right-hand side of the above relation, only the source linear size D may be estimated directly from observations, while the jet power Q and the jet lifetime τ can only be determined through a careful modeling by using algorithms

such as DYNAGE. Yet, as mentioned in the previous paragraph, those unknown parameters determine, at the same time, the observable curvature of the aging radio continuum of the lobes. It is thus reasonable to assume that the $\rho = f(D, Q, \tau)$ scaling is equivalent to the relation $\rho = f(D, \alpha, \alpha_{\text{inj}})$. If this is indeed the case, the modeling of a large sample of sources would in principle enable to *calibrate* such a relation, obtaining at the end a simple correlation of the parameter we seek, $\log \rho$, solely with the observable $\log D$ and α . The additional dependence on α_{inj} would then manifest as a correlation spread (a spread that, on the other hand, should not be substantial as long as α and α_{inj} are tightly correlated; see Appendix A for additional details).

In the following section, we demonstrate that in the parameter space emerging from the DYNAGE modeling of the compiled sample of FR II radio galaxies not only does a well-defined correlation plane ($\log \rho, \log D, \alpha$) exist, but in fact there is a much simpler correlation $\log \rho \propto \alpha$, albeit with a lower significance. By means of a Bayesian approach, we quantify the intrinsic spread of those correlations, and show that this spread decreases when $\log \rho_0$ (or rather $\log \rho_0 a_0^\beta$) is used in the regression analysis instead of $\log \rho$.

The full sample of 271 targets, along with the corresponding model data, is presented in Appendix B. The statistical analysis given in the following Section 3, is however performed on a slightly smaller sample of 253 sources, formed after excluding all the objects with linear sizes $D > 1$ Mpc. Indeed, we argue that for such giant radio sources, our simplified evolutionary scenario of a self-similar lobes' expansion in a $\beta = 3/2$ ambient density atmosphere, may not be really justified anymore, since at such large $\mathcal{O}(1 \text{ Mpc})$ distances from the galaxy group centers, hot gas may be far from the hydrostatic equilibrium condition (e.g., Walker et al. 2019, and references therein). In this context, see also Appendix C and the discussion therein.

We finally note that, although the spectral index between the two chosen rest-frame frequencies could be determined directly from the observations of the selected targets, in our model data we use instead the values derived from the modeled spectra between the emitted frequencies 0.4 and 5 GHz.

3. Statistical Analysis

3.1. Correlation Plane

Let us assume a linear relation $y = y(X)$ between the *response* variable $y \equiv \log \rho$, and the *prediction* variables $X \equiv (\log D, \alpha)$,

$$\log \rho = a + b \log D + c \alpha + \epsilon, \quad (3)$$

where $\theta \equiv (a, b, c)$ is the set of the coefficients of the linear relation and $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is the error term with a normal distribution and variance σ . Our training data set \mathcal{D} is the DYNAGE matrix $(\log \rho_i, \log D_i, \alpha_i)$, where $i = 1, 2, \dots, N$ with $N = 253$, the source linear size, D , is expressed in kiloparsecs (so that $D < 10^3$), and the ambient medium density, ρ , is given in the units of $10^{-28} \text{ g cm}^{-3}$ (so that, for example, the value $\log \rho \simeq 0$ corresponds roughly to the gas number density $n \simeq \rho/m_p \sim 10^{-4} \text{ cm}^{-3}$).

The data set displays a strong, statistically significant correlation, for which the p -values are $< 2.2 \times 10^{-16}$ for both Pearson's product-moment correlation test (returning the coefficient $\rho = -0.785$) and Kendall's rank correlation

Table 1
Results of the Correlation Analysis and Median-fit Parameters Following from the Bayesian Regression Analysis

Data Set (1)	Correlation Analysis		Model (4)	Bayesian Regression Analysis			
	Pearson (2)	Kendall (3)		\tilde{a} (5)	\tilde{b} (6)	\tilde{c} (7)	$\tilde{\sigma}$ (8)
$(\log \rho_i, \log D_i, \alpha_i)$	$\rho = -0.785$ $p < 2.2e - 16$	$\tau = -0.593$ $p < 2.2e - 16$	$\log \rho \sim$ $\mathcal{N}(a + b \log D + c \alpha, \sigma)$	0.28 ± 0.23	-1.01 ± 0.07	3.25 ± 0.22	0.40 ± 0.02
$(\log \rho_i, \alpha_i)$	$\rho = 0.459$ $p = 1.4e - 14$	$\tau = 0.328$ $p = 7.62e - 15$	$\log \rho \sim$ $\mathcal{N}(a + b \alpha; \sigma)$	-1.51 ± 0.29	2.65 ± 0.33		0.58 ± 0.03
$(\log(\rho_0 a_0^{3/2})_i, \log D_i, \alpha_i)$	$\rho = 0.757$ $p < 2.2e - 16$	$\tau = 0.520$ $p < 2.2e - 16$	$\log \rho_0 a_0^{3/2} \sim$ $\mathcal{N}(a + b \log D + c \alpha, \sigma)$	-3.97 ± 0.22	0.50 ± 0.06	3.26 ± 0.23	0.39 ± 0.02
$(\log(\rho_0 a_0^{3/2})_i, \alpha_i)$	$\rho = 0.675$ $p < 2.2e - 16$	$\tau = 0.452$ $p < 2.2e - 16$	$\log \rho_0 a_0^{3/2} \sim$ $\mathcal{N}(a + b \alpha; \sigma)$	$-3.09_{-0.23}^{+0.24}$	$3.55_{-0.27}^{+0.26}$		0.44 ± 0.02

Note. Column (1): data set; Column (2): Pearson's product-moment correlation coefficient ρ and the corresponding p -value; Column (3): Kendall's rank correlation coefficient τ and the corresponding p -value; Column (4): statistical model for the Bayesian regression analysis; Columns (5–7): coefficients of the linear relation; Column (8): intrinsic spread related to the unspecified variables.

(returning $\tau = -0.593$), as summarized in Table 1. However, as the relation expressed in Equation (3) follows from a much more complex model, which was briefly described in Section 2, there may be embedded dependence on additional variables. We thus decided to find the linear relation coefficients θ using the Bayesian approach to constrain the credible *distributions* of the model correlation parameters. This approach enables us to quantify the intrinsic scatter arising from the dependence on variables not taken into account in the correlation analysis.

In this approach, our statistical model \mathcal{M} states that the response variable is sampled from the normal distribution $y \sim \mathcal{N}(\bar{y}; \sigma)$ with mean

$$\overline{\log \rho} = a + b \log D + c \alpha \quad (4)$$

and variance σ representing the intrinsic spread related to the unspecified variables. With this approach, we aim to determine, the posterior probability density function (PDF) of the model parameters $\theta = (a, b, c; \sigma)$ via the Bayes' theorem

$$P(\theta|\mathcal{D}, \mathcal{M}) = \frac{P(\mathcal{D}|\theta, \mathcal{M}) P(\theta|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}, \quad (5)$$

where $P(\mathcal{D}|\theta, \mathcal{M})$ is the likelihood, i.e., the probability of measuring the set of data \mathcal{D} given particular values of the model parameters θ , $P(\theta|\mathcal{M})$ is the prior, i.e., the probability distribution for the parameters in the model, and finally

$$P(\mathcal{D}|\mathcal{M}) = \int P(\mathcal{D}|\theta, \mathcal{M}) P(\theta|\mathcal{M}) d\theta \quad (6)$$

is the model evidence.

We perform the analysis following closely Ostorero et al. (2017), in particular using the code APEMoST,⁶ (Gruberbauer et al. 2009) with 2×10^6 Markov chain Monte Carlo iterations and 20 chains to ensure a sufficiently complete sampling of the parameter space. We adopt independent, non-informative uniform priors on (a, b, c) , with the parameter space boundaries set to $[-10, 10]$. For the intrinsic spread σ , on the other hand, which by definition is always a positive number, we assume the

distribution function that describes a variate with mean r/μ , and variance r/μ^2 , namely,

$$P(\sigma|\mathcal{M}) = \frac{\mu^r}{\Gamma(r)} x^{r-1} \exp(-\mu x), \quad (7)$$

where $x = 1/\sigma$, and $\Gamma(r)$ is the Euler gamma function; in our calculations, we set $r = \mu = 10^{-5}$ and the variability interval boundaries $[0.01, 1000]$. The random number generator was set with bash command `GSL_RNG_TYPE="taus"` and the initial seed of the random number generator was set with `GSL_RNG_SEED=$RANDOM`.

The results of the Bayesian regression analysis as described above, are summarized in Table 1, and presented in Figure 1. In order to better visualize the significance of the correlation, as well as spread in the data around the correlation plane, in Figure 2, we plot the data points from the set \mathcal{D} along the projected plane with the median-fit correlation parameters' values $(\tilde{a}, \tilde{b}, \tilde{c})$ and $\pm \tilde{\sigma}$ deviations (see Table 1 for the parameter values). We see that for any pair of values of $\log D$ and α , the density of the ambient medium $\log \rho$ falls within a factor $\simeq 0.80$ from the mean relation at the 68% confidence level.

3.2. Simplified Linear Relation

Since it is not always possible to measure the projected linear sizes of distant radio galaxies, we also investigate whether spectral information alone, i.e., the slope of the observed radio emission continuum between two selected frequencies, is sufficient to estimate the density of the ambient medium on intergalactic scales. In our *reduced* training data set $(\log \rho_i, \alpha_i)$, we still see a statistically significant correlation between $\log \rho$ and α , for which Pearson's product-moment correlation test returns $\rho = 0.459$, and Kendall's rank correlation gives $\tau = 0.328$; the p -values are $< 10^{-14}$ in both tests.

We analyze this correlation further using a Bayesian approach, in the same manner as outlined in Section 3.1, assuming in particular that our response variable $\log \rho$ is sampled from the normal distribution $\log \rho \sim \mathcal{N}(a + b \alpha; \sigma)$, and seeking the posterior PDF of the model parameters

⁶ Automated Parameter Estimation and Model Selection Toolkit; <http://apemost.sourceforge.net/> 2011 February.

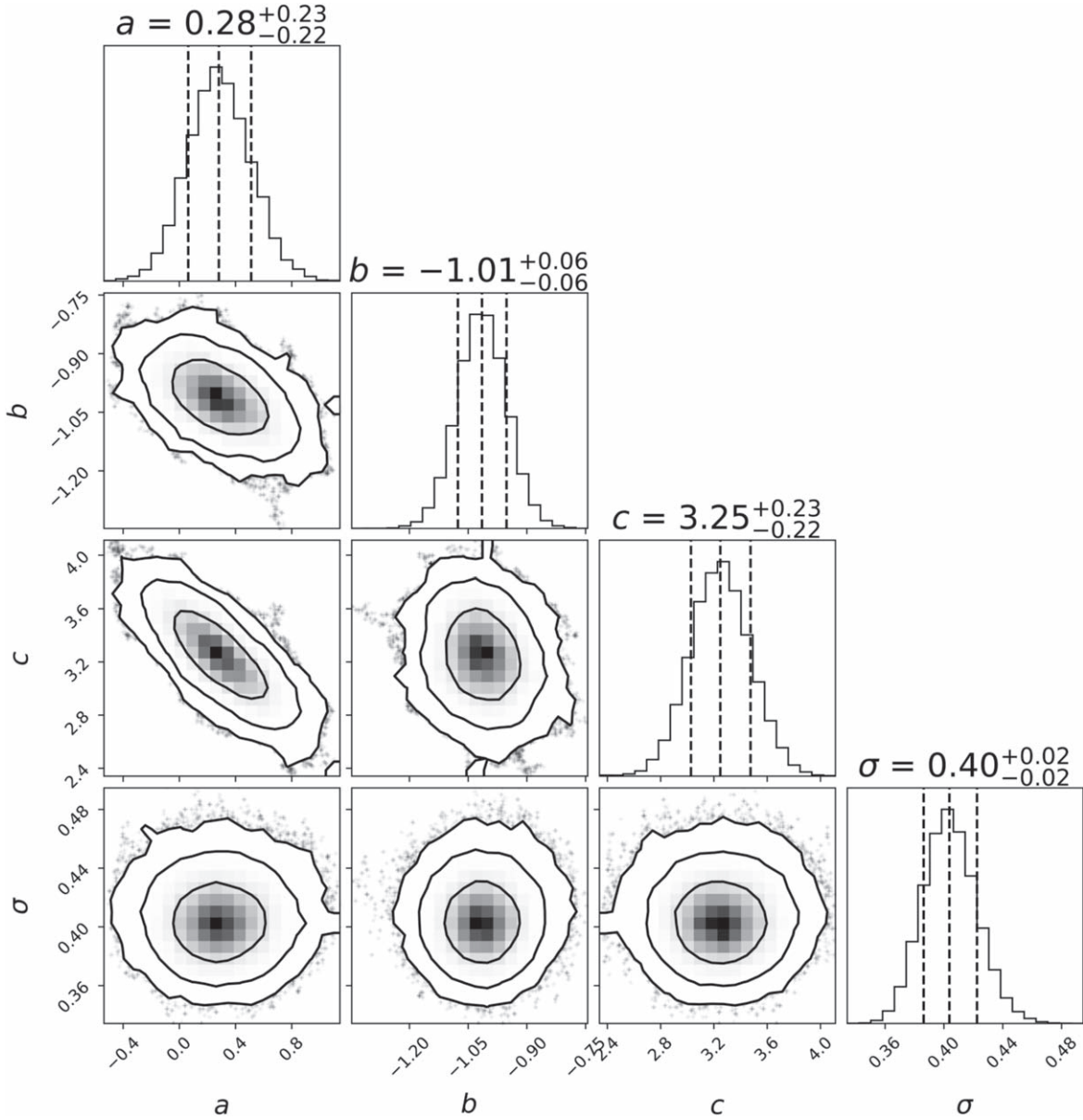


Figure 1. Results of the Bayesian analysis applied to the data set $(\log \rho, \log D_i, \alpha_i)$, with $\log \rho$ as the dependent variable. The panels show the marginalized PDFs of the model parameters $\theta = (a, b, c; \sigma)$. Contours correspond to the 68.3%, 95.4%, and 99.7% confidence levels, respectively. The median values and their marginalized 1σ uncertainties are denoted by dashed lines on the histograms showing the posterior probability distributions of the model parameters.

$\theta = (a, b; \sigma)$. The results of the analysis are summarized in Table 1, and can be visualized in Figures 3 and 4. For any given value of the spectral index α , the density of the ambient medium, $\log \rho$, falls within a factor $\simeq 1.16$ from the mean relation at the 68% confidence level.

3.3. Alternative Relations

The relatively large spread in the simplified $\log \rho \propto \alpha$ correlation, is predominantly due to the fact that the variable $\log \rho$ describes the density of the gas just ahead of the jets' termination shocks, i.e., at the distances $r \sim D/2$ from the centers of the systems; with the assumed density profile $\rho(r) \propto r^{-3/2}$ (Equation (1)), this variable has to be therefore a rather strong function of the source linear size D , which itself varies rather substantially between the objects included in our sample (in particular, from several kiloparsecs up to 1 Mpc).

Any correlation not explicitly involving the parameter $\log D$, must therefore be characterized by a large scatter.

In order to minimize this dependence, and in this way, to narrow the spread in the simplified linear relation involving solely the radio spectral information, we replace the response variable $\log \rho$ with $\log \rho_0 a_0^{3/2}$. The units for the latter are 10^8 cgs, so that, for example, with $a_0 = 10$ kpc, the $\log \rho_0 a_0^{3/2} = 0$ value corresponds to the gas number density $n_0 \simeq \rho_0/m_p \simeq 10^{-2} \text{ cm}^{-3}$.

First, however, we analyze the modified correlation plane, by forming an *alternative* training data set $(\log[\rho_0 a_0^{3/2}]_i, \log D_i, \alpha_i)$. We repeat the correlation analysis and the Bayesian regression analysis for this data set, the results of which are summarized in Table 1. As given, the change in the response variable does not dramatically alter the correlation significance, or the spread inferred from the Bayesian analysis, as long as both prediction variables $\log D$ and α are included: the

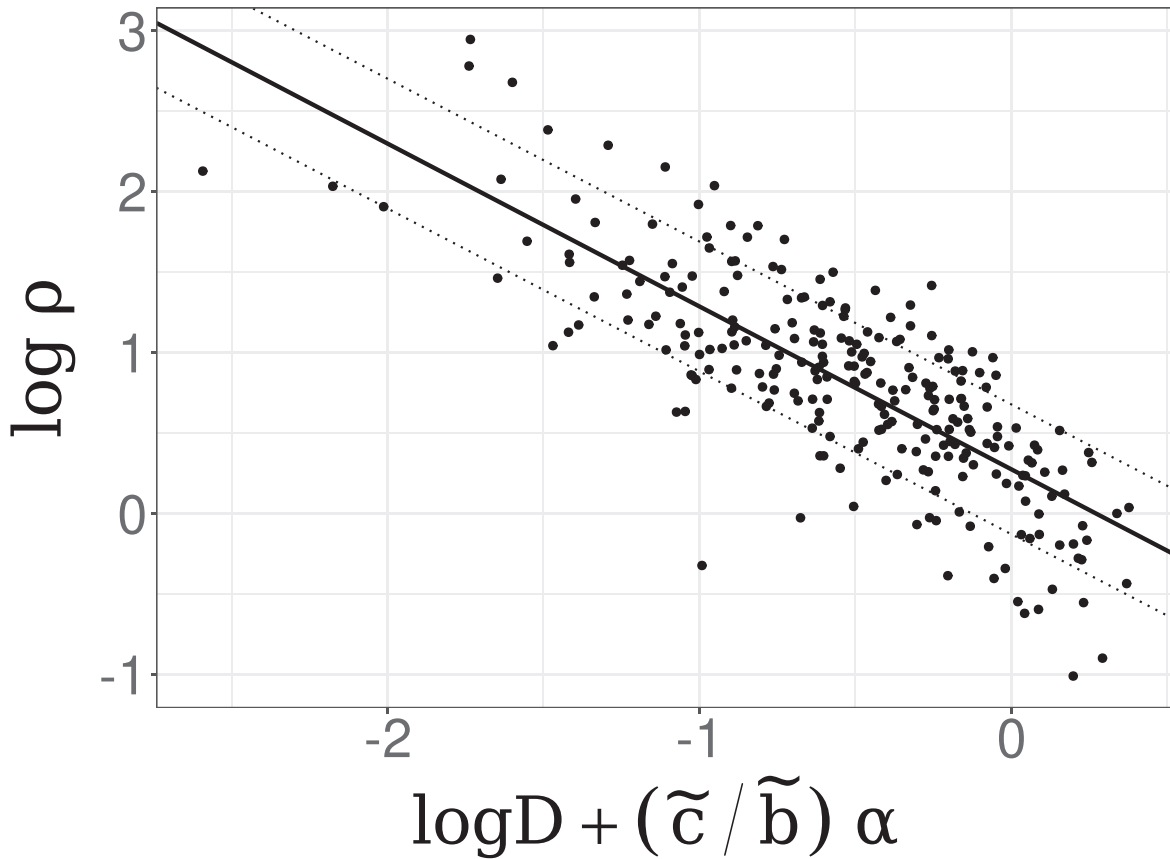


Figure 2. Data points from the set $(\log \rho_i, \log D_i, \alpha_i)$ plotted on the projected plane. The solid line corresponds to the regression line $\log \rho = \tilde{a} + \tilde{b} \log D + \tilde{c} \alpha$, while the dashed lines indicate the $\pm \tilde{\sigma}$ deviation of the relation; see Table 1 for the median-fit regression parameters' values $(\tilde{a}, \tilde{b}, \tilde{c}; \tilde{\sigma})$.

overall correlation is similarly very significant (p -values $< 2.2 \times 10^{-16}$ for both tests, with the Pearson's test yielding $\rho = 0.757$, and Kendall's rank correlation yielding $\tau = 0.520$), while the response variable $\log \rho_0 a_0^{3/2}$ falls within a factor $\simeq 0.78$ from the mean relation at the 68% confidence level for any given pair $(\log D, \alpha)$.

Finally, we analyze the *alternative reduced* training data set $(\log[\rho_0 a_0^{3/2}]_i, \alpha_i)$, in which we do see an improved correlation between $\log \rho_0 a_0^{3/2}$ and α (compared to that seen between $\log \rho$ and α), with the Pearson's product-moment correlation test yielding $\rho = 0.675$ and the Kendall's rank correlation yielding $\tau = 0.452$ with p -values $< 2.2 \times 10^{-16}$ in both tests. We repeat the Bayesian regression analysis also for this data set; the results are summarized in Table 1, and can be visualized in Figures 5 and 6. This time, for every value of the spectral index α , the *central* density of the ambient medium $\log \rho_0 a_0^{3/2}$ falls within a factor $\simeq 0.88$ from the mean relation at the 68% confidence level.

4. Discussion

As mentioned in the Introduction, the direct method of estimating the density of the hot gaseous fraction of the IGM, relies on detailed modeling of good-quality X-ray measurements of the extended, thermal emission component surrounding radio structures. Such data are currently available only for a small fraction of the nearby and/or brightest sources subjected to deep exposures with either the Chandra or XMM-Newton X-ray telescopes.

In this context, first we refer to the results presented in Belsole et al. (2007), who analyzed the available Chandra and XMM-Newton data for powerful FR IIs selected from the Third Cambridge Catalog of Radio Sources, Revised Edition (3CRR; Laing et al. 1983), within the redshift range $0.45 < z < 1.0$, and with the 178 MHz power spectral density between $\sim 7 \times 10^{26}$ and $\sim 10^{28}$ W Hz $^{-1}$ sr $^{-1}$ (see Figure 10 in Appendix B for the distribution of the 1.4 GHz power with redshift for the sources included in our sample).

Assuming the X-ray surface brightness profiles, corresponding to the IGM emission component, in a form $\Sigma_X \propto [1 + (r/r_c)^2]^{-3\tilde{\beta}+0.5}$, Belsole et al. (2007) were able to constrain the parameters $\tilde{\beta}$ and r_c for 10 out of the selected 20 targets. We note that with such a parameterization of the surface brightness, the corresponding gas density profile reduces to the one given in our Equation (1) for $\beta = 3\tilde{\beta}$ and $r_c = a_0 \ll r$. Nine of those 10 sources overlap with our sample, namely, 3C 200, 3C 207, 3C 220.1, 3C 254, 3C 265, 3C 292, 3C 295, 3C 330, and 3C 427.1. For these, based on the best-fit values provided in Belsole et al. (2007) for the central gas number density n_0 and the core radius r_c , we therefore calculate the parameters $\log \rho_0 r_c^{3/2}$ with $\rho_0 = m_p n_0$, which can be confronted directly with the regression lines following from our Bayesian regression analysis, keeping in mind that in our modeling a_0 and ρ_0 are not independent parameters, as discussed in Section 3.3. The results of this comparison are presented in Figure 7.

As shown, the central gas densities emerging from the observational constraints by Belsole et al. (2007) are within 3σ

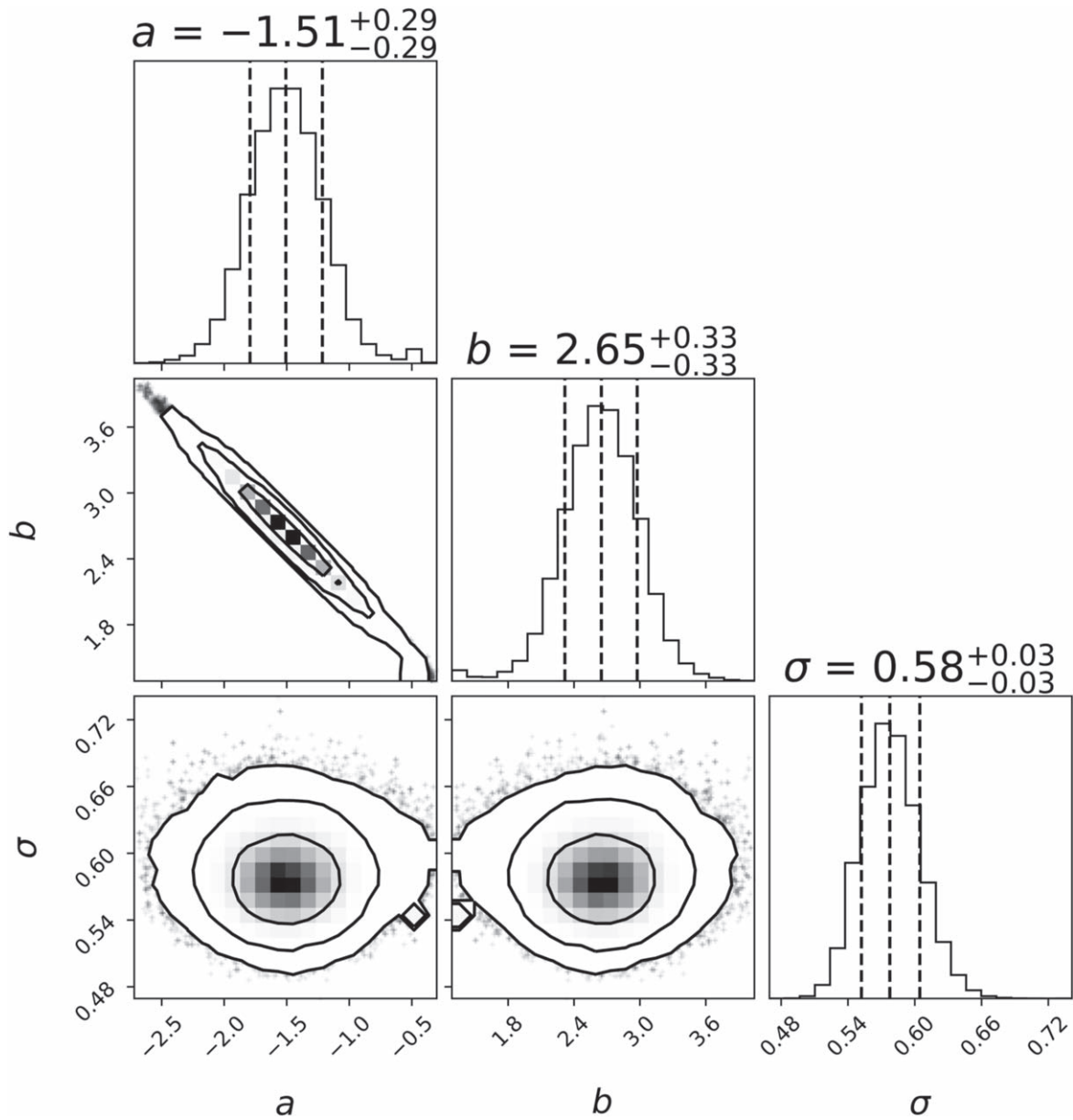


Figure 3. Results of the Bayesian analysis applied to the data set $(\log \rho_i, \alpha_i)$, with $\log \rho$ as the dependent variable. The panels show the marginalized PDFs of the model parameters $\theta = (a, b, \sigma)$. Contours correspond to the 68.3%, 95.4%, and 99.7% confidence levels, respectively. The median values and their marginalized 1σ uncertainties are denoted by dashed lines on the histograms showing the posterior probability distributions of the model parameters.

from our best-fit relations for all the sources but 3C 295 (given the errors). These “X-ray” values are, nonetheless, in all the cases systematically larger than the ones implied by our modeling and regression analysis. The reasons for such a discrepancy can be, at least partially, the difference between the universal β parameter assumed in the DYNAGE modeling, and the $\tilde{\beta}$ values following from the X-ray data fitting: indeed, even though the best-fit values of $\tilde{\beta}$ provided in Belsole et al. (2007) are in many cases consistent within the errors with the universal $\tilde{\beta} = 0.5$, they are nonetheless in the range $0.5 \lesssim \tilde{\beta} \lesssim 1.0$, implying steeper slopes of the ambient density profiles than the one assumed in the DYNAGE modeling.

We also observed that, for all the overlapping sources except 3C 295, the observational constraints on r_c are particularly uncertain, with the relative errors exceeding 50%. This is due to the extended wings of the point spread functions of the bright

unresolved X-ray cores of the targets which, even with the superb (\sim arcsec) angular resolution of the Chandra ACIS imaging instrument, precludes any precise determination of the diameter of a central plateau in their hot gaseous atmospheres (which is typically of the order of a few/several arcsecs, or less). Importantly, the source 3C 295, for which the observational constraints on r_c are relatively tight, i.e., $r_c \simeq 3.4 \pm 0.25''$, is also the one characterized by the largest discrepancy between the two values of the central gas density. This particular object is, on the other hand, the smallest radio galaxy in this overlapping subsample, with the linear size of its radio structure $D \simeq 34$ kpc; and since $D/2 \lesssim r_c \simeq 18.8 \pm 1.4$ kpc, we conclude that the DYNAGE model assumption, regarding a self-similar evolution of the radio structure in a power-law ambient density profile, is hardly fulfilled for the target. This signals that the method proposed here may not apply to compact radio galaxies.

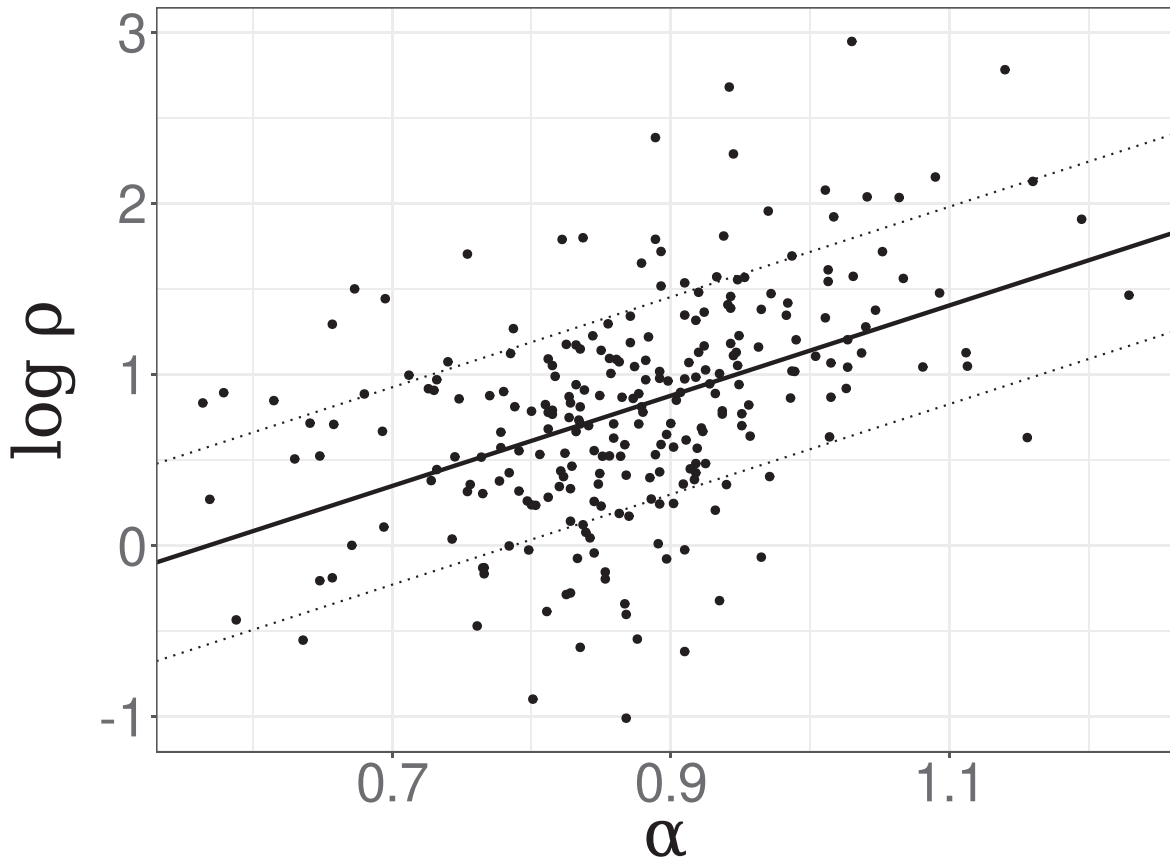


Figure 4. Data points from the set $(\log \rho_i, \alpha_i)$ plotted with the linear regression line $\log \rho = \bar{a} + \bar{b} \alpha$ (solid line), with $\pm \bar{\sigma}$ deviations (dashed lines); see Table 1 for the median-fit regression parameters' values $(\bar{a}, \bar{b}; \bar{\sigma})$.

Meanwhile, for the majority of sources for which the X-ray surface brightness profiles cannot be constrained at all (or cannot be constrained precisely) from the X-ray imaging data, but the X-ray emission of the IGM can still be seen in the total radiative outputs of the systems, one can consider the total IGM-related X-ray luminosity as a proxy for the central gas density. Indeed, considering solely the free-free continuum component of the thermal IGM emission, $\beta = 3/2$ gas density profile, and uniform gas temperature within the entire emission volume, this luminosity reads as

$$\begin{aligned}
 L_X &\propto \int \rho^2 dV \propto \int_0^{a_0} \rho_0^2 r^2 dr \\
 &+ \int_{a_0}^{r_t} [\rho_0 a_0^{3/2}]^2 r^{-1} dr \\
 &\sim [\rho_0 a_0^{3/2}]^2 \times \left(\frac{1}{3} + \ln \frac{r_t}{a_0} \right), \quad (8)
 \end{aligned}$$

meaning an approximate square dependence of L_X on $\rho_0 a_0^{3/2}$, modulo the slowly varying factor $(\frac{1}{3} + \ln r_t/a_0)$ for the termination radius $r_t \gg a_0$.

Keeping the above relation in mind thereafter we refer to the results presented in Ineson et al. (2017), who studied a sample of FR II radio galaxies with available good X-ray coverage, selected from various flux-limited, spectroscopically complete, low-frequency radio surveys. The sample is representative for radio-loud active galaxies at redshifts $z \sim 0.1$ and $z \sim 0.5$, and contains the objects with the 151 MHz power spectral density between $\sim 0.2 \times 10^{26}$ and $0.4 \times 10^{28} \text{ W Hz}^{-1} \text{ sr}^{-1}$. For these

objects, Ineson et al. (2017) estimated the IGM-related total X-ray luminosities (within the r_{500} radii), L_{IGM} , which are plotted in Figure 8 for the objects overlapping with our list against the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$, or the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha, \sigma)$ values, following from our Bayesian regression analysis.

There is a positive, statistically significant correlation between L_{IGM} and $\log \rho_0 a_0^{3/2}$. In particular, for the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$ model the Pearson's product-moment correlation test yields $\rho = 0.613$ and p -value $\simeq 0.012$, and the Kendall's rank correlation yields $\tau = 0.517$ and p -value $\simeq 0.005$. Meanwhile, the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha, \sigma)$ model seems to work equally well in this respect, with the corresponding parameters $\rho = 0.716$ and p -value $\simeq 0.002$ for the Pearson's test, and $\tau = 0.60$ and p -value $\simeq 0.001$ for the Kendall's rank correlation.

5. Conclusions

Our analysis of the data emerging from the extensive DYNAGE modeling of a large sample of FR II radio sources, reveals a positive and statistically significant correlation between the density of the ambient medium at scales comparable to the linear sizes of the sources' radio structures, $\log \rho$, and the spectral index between the emitted frequencies 0.4 and 5 GHz, α . Moreover, the intrinsic scatter in the correlation—which may be interpreted as a combination of the uncertainty resulting from the DYNAGE model assumptions, and the dependence on additional hidden parameters—can be further minimized by either including the additional variable, in

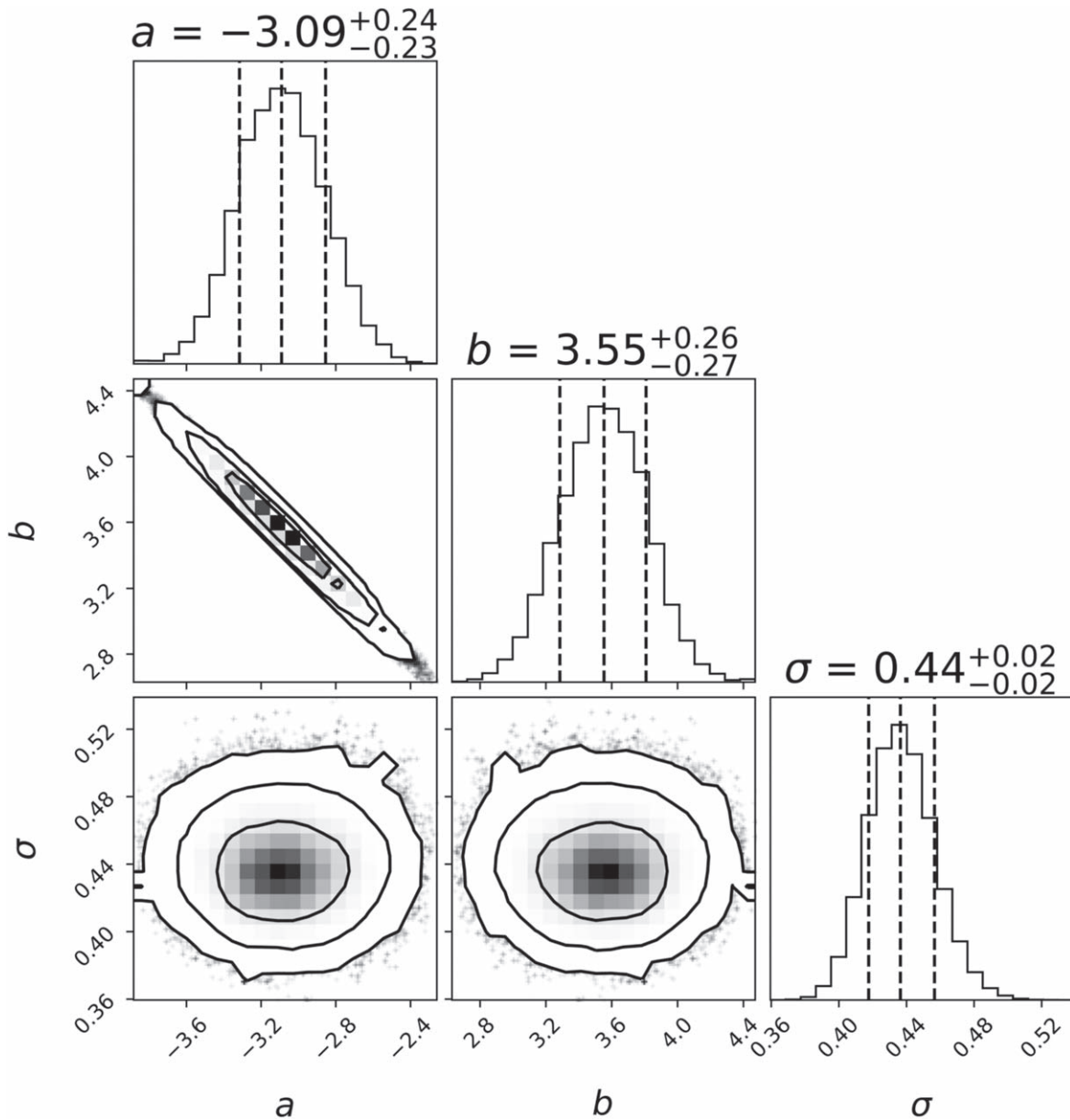


Figure 5. Results of the Bayesian analysis applied to the data set $(\log[\rho_0 a_0^{3/2}], \alpha_i)$, with $\log \rho_0$ as the dependent variable. The panels show the marginalized PDFs of the model parameters $\theta = (a, b; \sigma)$. Contours correspond to the 68.3%, 95.4%, and 99.7% confidence levels, respectively. The median values and their marginalized 1σ uncertainties are denoted by dashed lines on the histograms showing the probability distributions of the model parameters.

particular the observed projected size of the radio structure $\log D$, and/or using the central gas density $\log \rho_0$ as a dependent variable instead of $\log \rho_0 a_0^{3/2}$ (for the assumed $\beta = 3/2$ density profile); in either of these cases, the statistical model allows us to predict the density of the ambient medium with an accuracy of about 0.3 dex.

The fact that there is *some positive* correlation between the environment density of FR IIs and the slope of the integrated radio emission of their extended lobes, should not, in fact, be that surprising, keeping in mind the work by O’Sullivan et al. (2015), who reported on the absence of radio galaxies with a large degree of linear polarization (at 1.4 GHz) and steep radio spectra. When combined with the analysis of the neighboring galaxy counts, this observational finding indicates that the decreased polarization degree correlates with the richness of the local environment (as expected if the depolarization is

predominantly due to the Faraday effect related to the surrounding magnetoionic medium), and hence that the spectral index is indeed positively correlated with the environment density (see the discussion in Section 3.4.4. therein).

There are several important implications of the emerging dependence, each deserving an in-depth discussion. One is related to the physics of hotspots in FR II radio galaxies, which constitute spectacular manifestations of mildly relativistic shocks formed when extremely low-density but relativistic and highly magnetized outflows (jets), collide with high-density but nonrelativistic gas of the IGM (e.g., Meisenheimer et al. 1989), and in particular, to the issue of particle acceleration taking place at and around the fronts of such shocks. Namely, keeping in mind the $\log \rho - \alpha$ positive relation on one hand, and the $\alpha - \alpha_{\text{inj}}$ tight and also positive correlation on the other hand, one may ask how exactly could the increasing density of the ambient medium decrease the

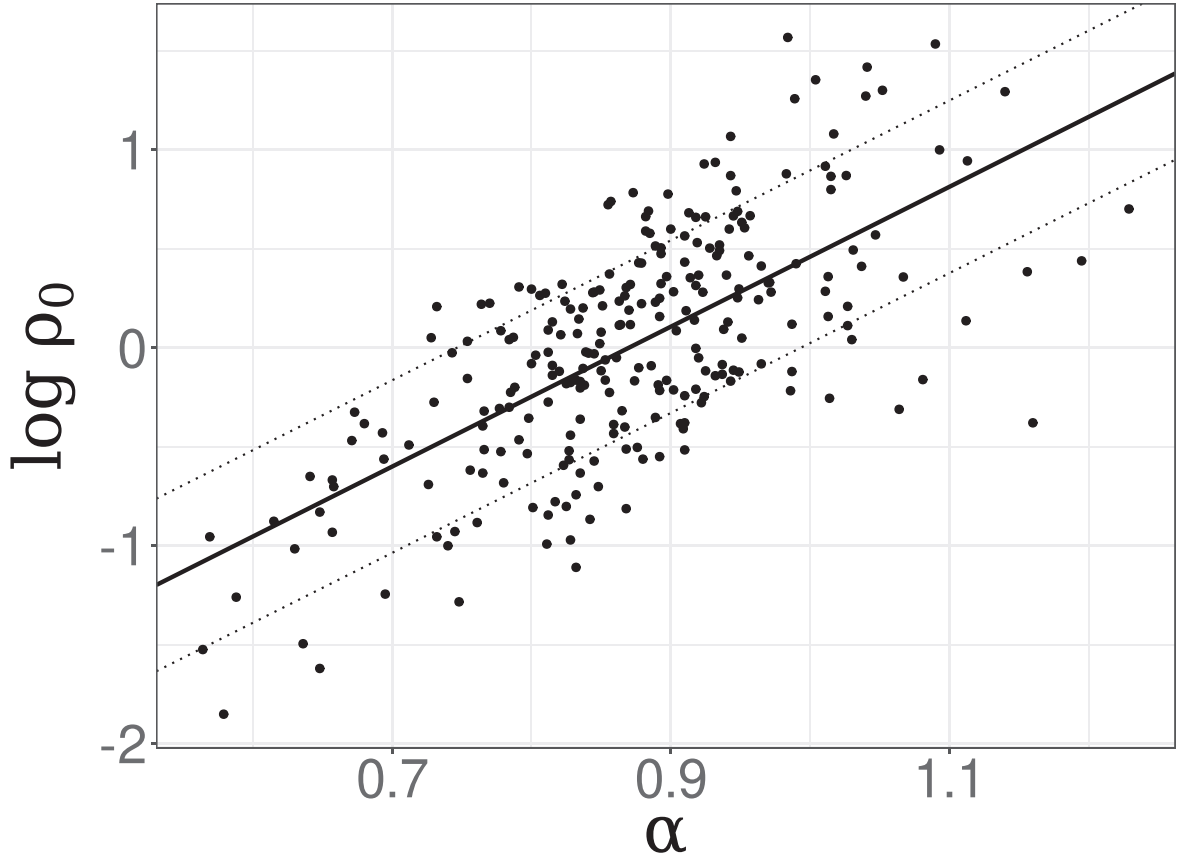


Figure 6. Data points from the set $(\log[\rho_0 a_0^{3/2}]_i, \alpha_i)$ plotted with the linear regression line $\log[\rho_0 a_0^{3/2}] = \tilde{a} + \tilde{b} \alpha$ (solid line), with $\pm\tilde{\sigma}$ deviations (dashed lines); see Table 1 for the median-fit regression parameters' values $(\tilde{a}, \tilde{b}; \tilde{\sigma})$.

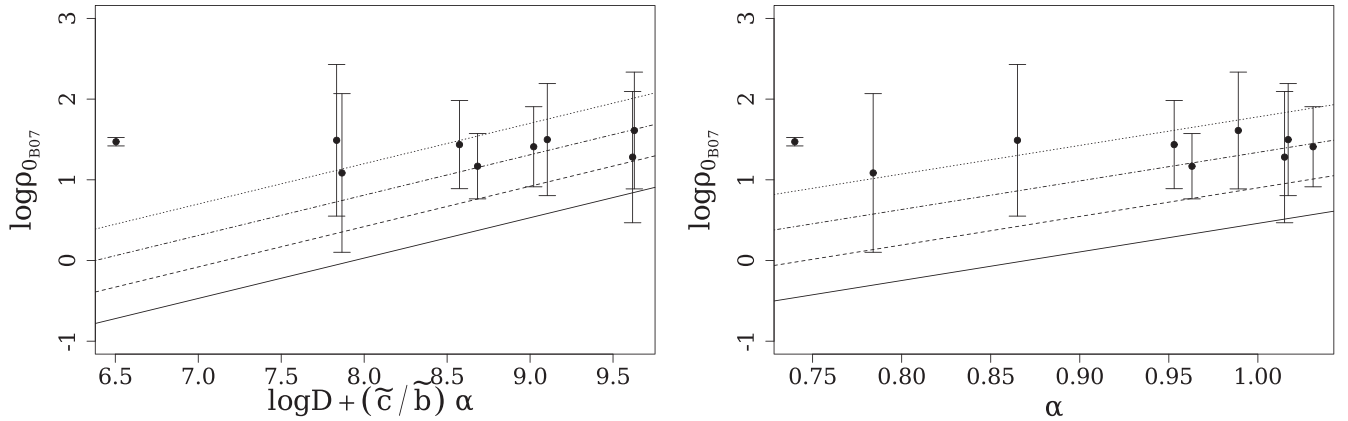


Figure 7. Comparison between the parameter $\log \rho_0 r_c^{3/2}$ [$10^8 \text{ g cm}^{-3/2}$ units], calculated based on the best-fit values for n_0 and r_c obtained by Belsole et al. (2007) from the X-ray data analysis, and the regression (solid) lines following from our Bayesian analysis assuming the model $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$, or the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha, \sigma)$ (left and right panels, respectively; see Section 3.3 and Table 1). Dashed, dotted-dashed, and dotted lines correspond to $+1\sigma$, $+2\sigma$, and $+3\sigma$ intervals from the best-fit relations, respectively.

efficiency of accelerating jet electrons at the termination shocks to ultrarelativistic energies.

The other emerging ramification of this study is the general problem of high-redshift radio galaxies (for a review see McCarthy 1993; Miley & De Breuck 2008), which are known to display steeper radio emission continua than their low-redshift counterparts (e.g., Klamer et al. 2006). Those topics will be addressed in future publications. Here, instead, we only report on the newly recognized correlations, and on how our findings seem to be supported by the analysis of good-quality

and high-angular resolution X-ray data available for some of the targets. In particular, this comparison signals that the main limitation of our model is related to the assumption of a self-similar evolution of a radio structure in a power-law ambient medium density $\rho \propto r^{-3/2}$. As a result, the method proposed here should be taken with caution when dealing with the most compact targets, i.e., those with $D \sim \mathcal{O}(10 \text{ kpc})$ and smaller, and also with giants, for which $D \sim \mathcal{O}(1 \text{ Mpc})$.

With proper multiwavelength support, however, one could possibly improve the emerging statistical correlations, for the

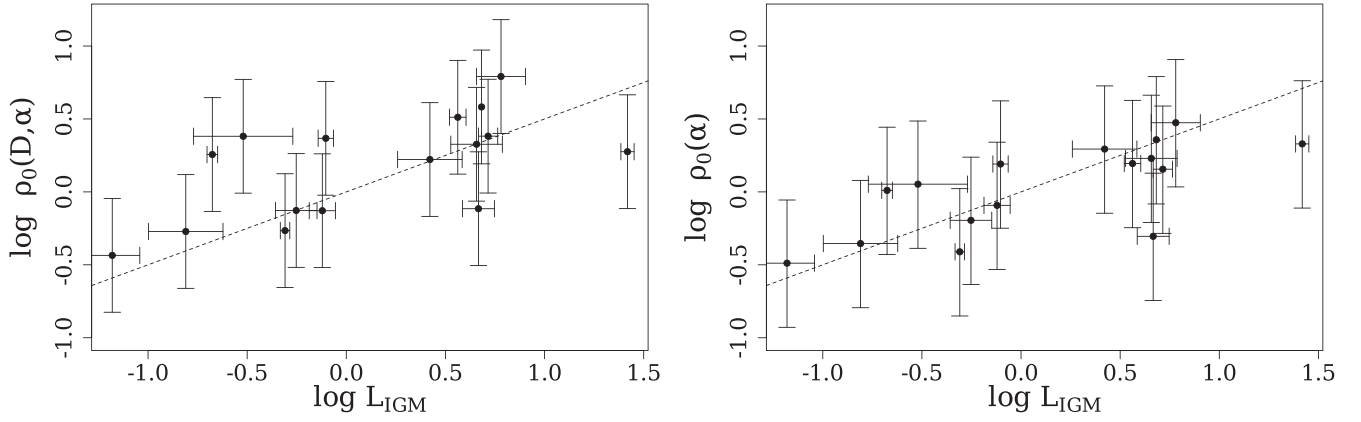


Figure 8. Comparison between the IGM-related total X-ray luminosities (within the r_{500} radii), in the units of 10^{43} erg s^{-1} , as estimated by Ineson et al. (2017), and the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$, or the $\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha, \sigma)$ values (left and right panel, respectively), following from our Bayesian regression analysis, as discussed in Section 3.3 (see also Table 1). The dashed lines represent the $y = \frac{1}{2}x$ scaling, that can be expected based on Equation (8), included to guide the eye.

purpose of using them as well-established cosmological tools when dealing with large samples of objects emerging from wide-area and high-sensitivity radio surveys (such as the one carried out by LOFAR, or the Square Kilometer Array, or the VLA Sky Survey; (e.g., Agudo et al. 2015; Hardcastle et al. 2019; Lacy et al. 2020, respectively) combined with the modern and next-generation optical photometric and spectroscopic surveys (like the Sloan Digital Sky Survey, or the forthcoming Large Synoptic Survey Telescope; Ivezić et al. 2019; Ahumada et al. 2020, respectively), and the new-generation X-ray surveys (in particular, extended ROentgen Survey with an Imaging Telescope Array (eROSITA) on board the Spectrum-Roentgen-Gamma mission; e.g., Predehl et al. 2021).

We acknowledge stimulating discussions on Bayesian analysis with A. Diaferio. The authors also acknowledge C.C. Cheung and S.P. O’Sullivan for the discussions and helpful comments on the manuscript.

A.W. and Ł.S. were supported by the Polish National Science Centre grant 2016/22/E/ST9/00061. J.M. was partly supported by the National Science Centre, Poland grant UMO 2018/29/B/ST9/01793. L.O. acknowledges partial support from the Italian Ministry of Education, University and

Research (MIUR) under the Departments of Excellence grant L.232/2016, and from the INFN grant InDark.

Appendix A The Spectral Index Correlation

As discussed in Section 2, based on the $\rho = f(D, \alpha, \alpha_{\text{inj}})$ relation expected in the framework of the DYNAGE modeling, we anticipate a simpler scaling $\rho = f(D, \alpha)$ involving only these independent variables D and α , which are accessible for direct observations. This anticipation is justified, since in our DYNAGE training data set we see a strong, positive correlation (p -value $< 2 \times 10^{-16}$) between the spectral index α for the two given emitted frequencies 0.4 and 5 GHz, and the injection spectral index α_{inj} , with $\rho = 0.73$ for the Pearson’s product-moment correlation test, and $\tau = 0.51$ for the Kendall’s rank correlation. The Bayesian analysis performed in the same manner as outlined in Section 3, assuming the model $\alpha_{\text{inj}} \sim \mathcal{N}(a + b \alpha; \sigma)$ for the dependent variable α_{inj} , yields the correlation parameter median-fit values $a = 0.294 \pm 0.015$, $b = 0.303 \pm 0.017$, and the intrinsic scatter $\sigma = 0.0322^{+0.0015}_{-0.0014}$; this correlation, overlaid with the data points, is displayed in Figure 9.

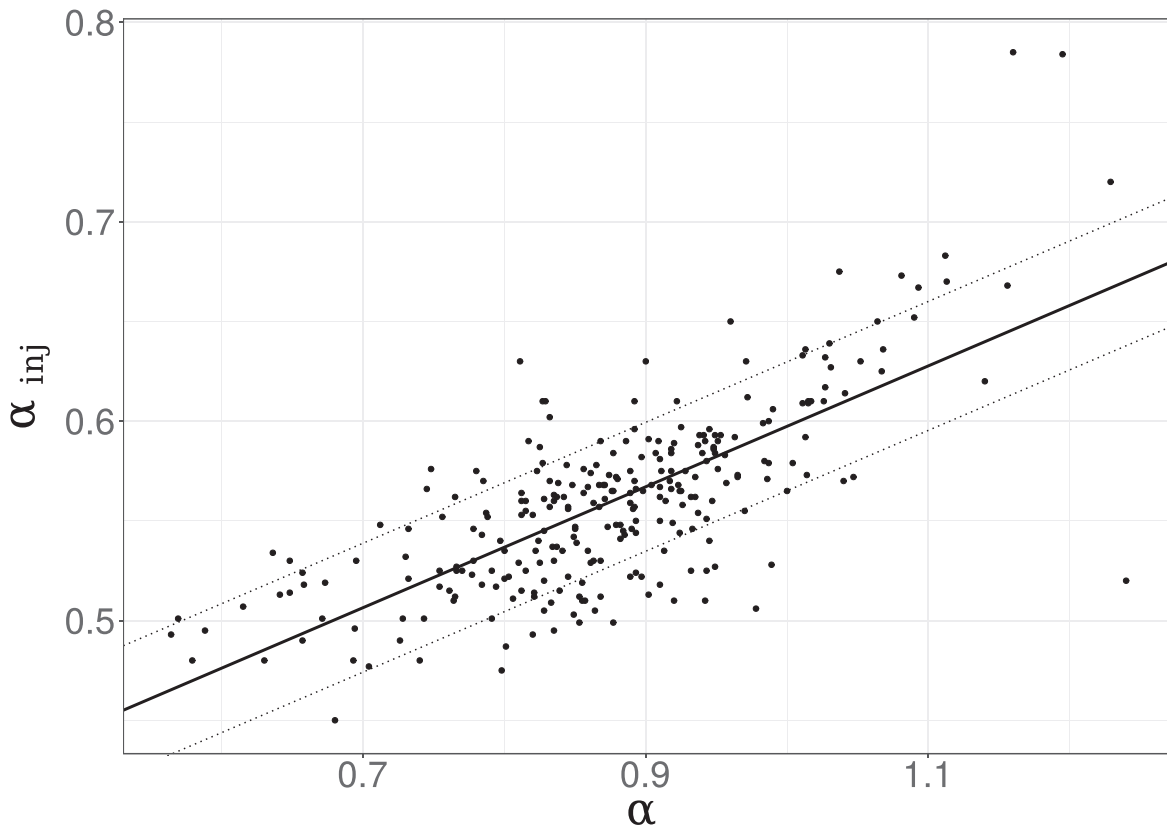


Figure 9. Data points from our DYNAGE training data set plotted along the correlation $\alpha_{\text{inj}} = \bar{a} + \bar{b} \alpha \pm \sigma$, for the median-fit correlation parameters' values $\bar{a} = 0.294$, $\bar{b} = 0.303$, and $\bar{\sigma} = 0.0322$.

Appendix B The Sample and the Model Data

Table 2 presents the sample of 271 FR II sources analyzed in this paper, with their basic observed properties (power spectral density at 1.4 GHz, total linear size, and aspect ratio of the radio lobes), as well as the corresponding DYNAGE model data (including the jet lifetime, total jet kinetic power, density

of the gaseous atmosphere, injection spectral index, and the integrated spectral index between the emitted frequencies 0.4 and 5 GHz). In Figure 10, we plot the distribution of the 1.4 GHz power density with redshifts for the sources included in the sample. For the final revised sample, along with an in-depth discussion on the model fitting procedure, see Machalski et al. (2021, ApJS, accepted).

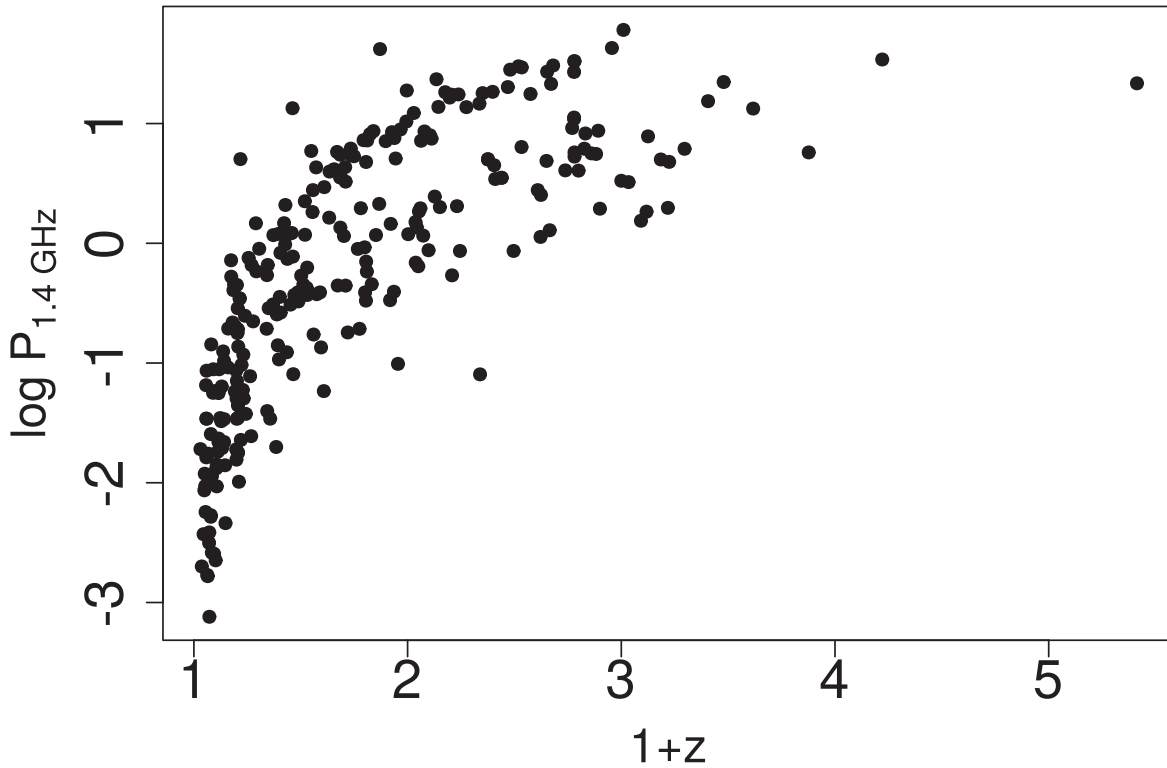


Figure 10. The distribution of the observed power spectral density at 1.4 GHz, in units of $10^{26} \text{ W Hz}^{-1} \text{ sr}^{-1}$, with redshift, for the sample of 271 FR II sources analyzed in this paper.

Table 2
The Sample of the Selected 271 FR II Sources with the Corresponding DYNAGE Model Data

Name	(1 + z)	$P_{1.4 \text{ GHz}}$ ($10^{26} \text{ W Hz}^{-1} \text{ sr}^{-1}$)	D (kpc)	AR	τ (Myr)	Q ($10^{45} \text{ erg s}^{-1}$)	$\rho_0 a_0^{3/2}$ ($10^8 \text{ g cm}^{-3/2}$)	ρ ($10^{-28} \text{ g cm}^{-3}$)	α_{inj}	α
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
3C6.1	1.84	8.65	199	3	4	14.55	0.292	1.827	0.54	0.797
3C9	3.01	60.53	119	3	1.97	89.74	1.295	16	0.617	1.027
3C13	2.351	17.95	238	3	6.2	28.05	1.12	5.033	0.59	0.951
3C14	2.469	20.18	205	3	3.22	41.4	0.39	2.295	0.59	0.909
3C16	1.405	0.832	594	3.3	30	3.524	2.14	2.538	0.63	0.971

Note. Column (1): name of the source; Column (2): spectroscopic redshift; Column (3): observed power spectral density at 1.4 GHz; Column (4): total linear size of the radio structure; Column (5): aspect ratio of the radio lobes; Column (6): jet lifetime; Column (7): total jet kinetic power; Column (8): central density of the gaseous atmosphere; Column (9): gas density at distances $D/2$ from the center of the host galaxy; Column (10): injection spectral index; Column (11): integrated spectral index of the lobes' radio emission, derived from the model spectra between the emitted frequencies 0.4 and 5 GHz. Table 2 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

(This table is available in its entirety in machine-readable form.)

Appendix C Giant Radio Sources in the Sample

In Figure 11, we present the studied correlations between the model parameters for the entire sample of 271 sources (our *full sample*), with triangles representing the most compact sources with linear sizes $D < 30 \text{ kpc}$, black dots representing intermediate objects with $30 \text{ kpc} < D < 1 \text{ Mpc}$, and squares denoting giant radio galaxies with $D > 1 \text{ Mpc}$. As follows, giant radio galaxies evolve in a substantially sparser ambient medium when compared to smaller sources, as in fact expected. Moreover, there is a gradual progression toward lower values of ρ from the most compact objects to giant radio galaxies (see both upper panels in Figure 11); there is no such a trend when the central gas density ρ_0 is used instead (see the two lower

panels in the figure), and this assures us that the DYNAGE algorithm captures correctly the dependencies between the model parameters. On the other hand, we also see that giant radio galaxies form an outlier population with respect to the correlations followed by all the other (smaller) sources.

Here, it is important to note that assuming the normality of the noise in a Bayesian regression analysis, the conflicting sources of information may contaminate the inference. This results in an undesirable effect—the posterior will concentrate in an area in between the main population of data points and outliers with a scaling large enough to incorporate them all, which leads to erroneous predictions and also can enlarge the spread. This effect can be mitigated if a population of outlying sources can be dropped from the sample. Generally, in order to minimize the

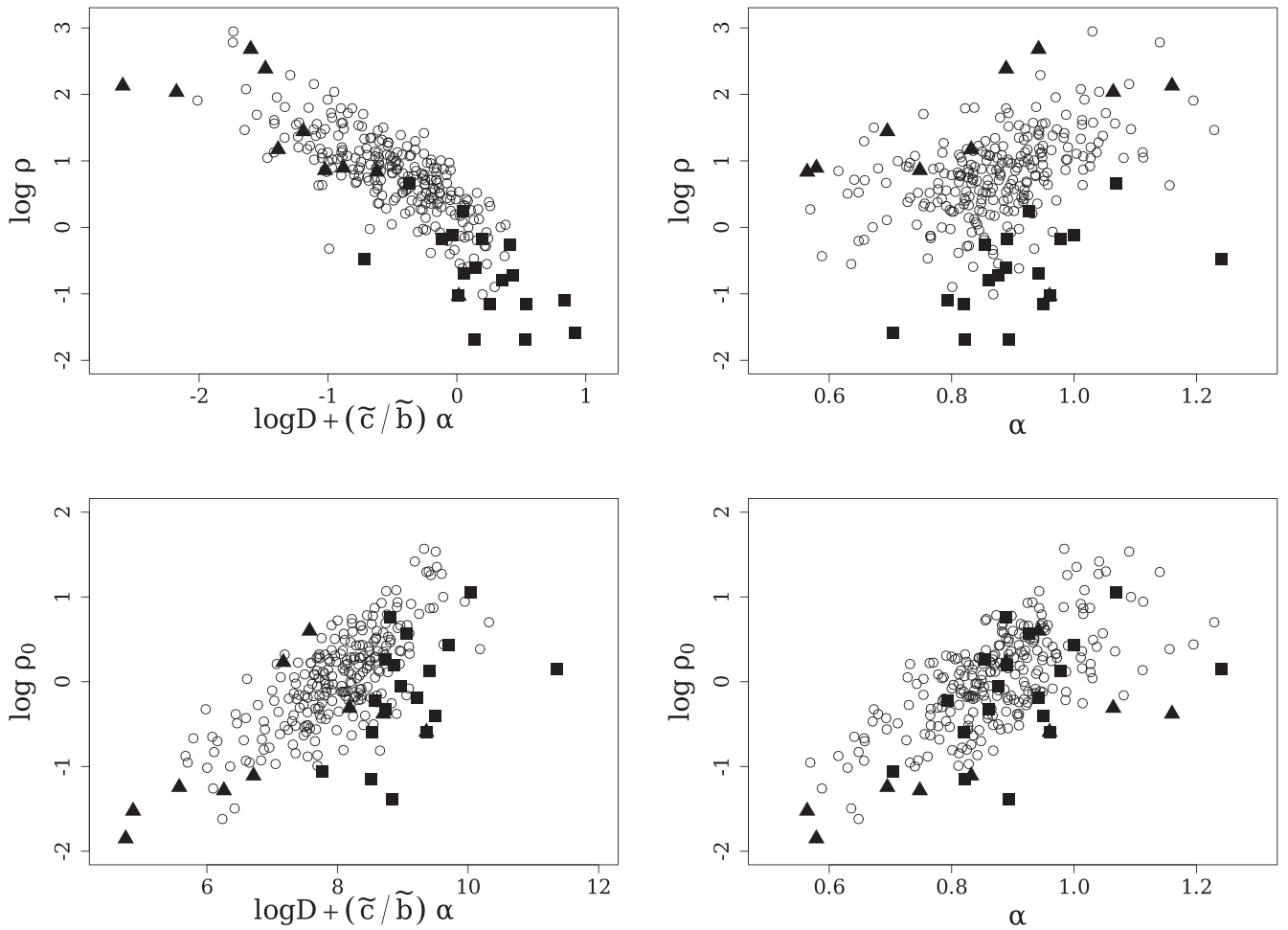


Figure 11. Correlations between the model parameters for the entire sample of 271 sources (our full sample), with triangles representing the most compact sources with linear sizes $D < 30$ kpc, empty dots representing intermediate objects with $30 \text{ kpc} < D < 1$ Mpc, and squares denoting giant radio galaxies with $D > 1$ Mpc.

influence of the outliers on our prediction, the modification of model is recommended, for instance, replacing it with Student-t distribution that can efficiently identify and reject an outliers from the analysis, for more details see Martinez-Cantin et al. (2017).

Indeed, for the full sample, the residual density distributions (after the regression fit is made), deviate from normal. We quantify this finding by means of the Shapiro–Wilk and the Anderson–Darling normality tests, which are summarized in Table 3. As follows, in all the cases but $\rho_0(\alpha)$, for the full

sample the distributions of the residuals vary from a Gaussian distribution on 95% level of confidence (p -values < 0.05). Meanwhile, for the $D < 1$ Mpc subsample, the p -values are all > 0.05 , meaning that the distributions of residuals are in this case consistent with normal.

We note also the double-peaked residual distribution of the full sample, which seems to be smoothed when extracting the giant radio sources, seems to confirm our hypothesis about the subpopulations of sources in our database.

Table 3
Normality Tests for the Residual Distributions

Dataset (1)	Model (2)	SW Statistics (3)	AD Statistics (4)	Residual Distribution (5)
full sample	$\log \rho \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$	$W = 0.98$ $p = 0.0014$	$A = 0.78$ $p = 0.04$	
$D < 1$ Mpc	$\log \rho \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$	$W = 0.99$ $p = 0.11$	$A = 0.27$ $p = 0.67$	
full sample	$\log \rho \sim \mathcal{N}(a + b \alpha; \sigma)$	$W = 0.97$ $p = 1.4e - 5$	$A = 2.45$ $p = 3.3e - 6$	
$D < 1$ Mpc	$\log \rho \sim \mathcal{N}(a + b \alpha; \sigma)$	$W = 0.99$ $p = 0.37$	$A = 0.44$ $p = 0.29$	
full sample	$\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$	$W = 0.99$ $p = 0.01$	$A = 0.67$ $p = 0.08$	
$D < 1$ Mpc	$\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \log D + c \alpha, \sigma)$	$W = 0.99$ $p = 0.38$	$A = 0.27$ $p = 0.68$	
full sample	$\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha; \sigma)$	$W = 0.99$ $p = 0.10$	$A = 0.75$ $p = 0.05$	
$D < 1$ Mpc	$\log \rho_0 a_0^{3/2} \sim \mathcal{N}(a + b \alpha; \sigma)$	$W = 0.99$ $p = 0.15$	$A = 0.75$ $p = 0.05$	

Notes. Column (1): sample analyzed—*full sample* refers to 271 sources listed in Appendix B, while “ $D < 1$ Mpc” refers to the subsample of 253 sources with linear sizes $D < 1$ Mpc; Column (2): statistical model considered; Column (3): the value of the Shapiro–Wilk statistic and the corresponding p -value; Column (4): the value of the Anderson–Darling statistic with the corresponding p -value; Column (5): the residual density distributions resulting from fitting the regression line to the given data set, x -axis range was set to $(-3, 3)$, and y -axis $(0, 1)$ for all plots.

ORCID iDs

Anna Wójtowicz  <https://orcid.org/0000-0001-5666-8665>
 Łukasz Stawarz  <https://orcid.org/0000-0001-8294-9479>
 Jerzy Machalski  <https://orcid.org/0000-0001-7506-1180>
 Luisa Ostorero  <https://orcid.org/0000-0003-3983-5980>

References

- Agudo, I., Boettcher, M., Falcke, H. D. E., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 93
- Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, *ApJS*, 249, 3
- Antognini, J., Bird, J., & Martini, P. 2012, *ApJ*, 756, 116
- Barthel, P. D. 1989, *ApJ*, 336, 606
- Begelman, M. C., & Cioffi, D. F. 1989, *ApJ*, 345, L21
- Belsole, E., Worrall, D. M., Hardcastle, M. J., et al. 2007, *MNRAS*, 381, 1109
- Belsole, E., Worrall, D. M., Hardcastle, M. J., & Croston, J. H. 2007, *MNRAS*, 381, 1109
- Bhatta, G., Stawarz, Ł., Markowitz, A., et al. 2018, *ApJ*, 866, 132
- Capetti, A., Massaro, F., & Baldi, R. D. 2017, *A&A*, 601, A81
- Cavaliere, A., & Fusco-Femiano, R. 1976, *A&A*, 49, 137
- Cheung, C. C., Giacintucci, S., & Clarke, T. E. 2019, *ApJ*, 875, 108
- Coe, D., Umetsu, K., Zitrin, A., et al. 2012, *ApJ*, 757, 22
- Croston, J. H., Hardcastle, M. J., Mingo, B., et al. 2019, *A&A*, 622, A10
- de Ruiter, H. R., Parma, P., Fanti, C., et al. 1986, *A&AS*, 65, 111
- Fanaroff, B. L., & Riley, J. M. 1974, *MNRAS*, 167, 31P
- Fanti, C., Fanti, R., de Ruiter, H. R., et al. 1986, *A&AS*, 65, 145
- Fanti, C., Fanti, R., de Ruiter, H. R., et al. 1987, *A&AS*, 69, 57
- Geller, M. J., Hwang, H. S., Diaferio, A., et al. 2014, *ApJ*, 783, 52
- Grimes, J. A., Rawlings, S., & Willott, C. J. 2004, *MNRAS*, 349, 503
- Gruberbauer, M., Kallinger, T., Weiss, W. W., et al. 2009, *A&A*, 506, 1043
- Hagino, K., Stawarz, Ł., Siemiginowska, A., et al. 2015, *ApJ*, 805, 101
- Haines, C. P., Finoguenov, A., Smith, G. P., et al. 2018, *MNRAS*, 477, 4931
- Hardcastle, M. J., Williams, W. L., Best, P. N., et al. 2019, *A&A*, 622, A12
- Harvanek, M., & Stocke, J. T. 2002, *AJ*, 124, 1239
- Hill, G. J., & Lilly, S. J. 1991, *ApJ*, 367, 1
- Ineson, J., Croston, J. H., Hardcastle, M. J., et al. 2017, *MNRAS*, 467, 1586
- Käfer, F., Finoguenov, A., Eckert, D., et al. 2019, *A&A*, 628, A43
- Ivezic, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111
- Kaiser, C. R., & Alexander, P. 1997, *MNRAS*, 286, 215
- Kaiser, C. R., Dennett-Thorpe, J., & Alexander, P. 1997, *MNRAS*, 292, 723
- Kawakatu, N., & Kino, M. 2006, *MNRAS*, 370, 1513
- King, I. 1962, *AJ*, 67, 471
- Klamer, I. J., Ekers, R. D., Bryant, J. J., et al. 2006, *MNRAS*, 371, 852
- Kozieł-Wierzbowska, D., & Stasińska, G. 2011, *MNRAS*, 415, 1013
- Lacy, M., Baum, S. A., Chandler, C. J., et al. 2020, *PASP*, 132, 035001
- Laing, R. A., Riley, J. M., & Longair, M. S. 1983, *MNRAS*, 204, 151
- Ledlow, M. J., & Owen, F. N. 1996, *AJ*, 112, 9
- Lin, Y.-T., Shen, Y., Strauss, M. A., et al. 2010, *ApJ*, 723, 1119
- Machalski, J. 1998, *A&AS*, 128, 153
- Machalski, J., Chyży, K. T., Stawarz, Ł., et al. 2007, *A&A*, 462, 43
- Machalski, J., Jamrozy, M., & Konar, C. 2010, *A&A*, 510, A84
- Machalski, J., Jamrozy, M., Stawarz, Ł., et al. 2011, *ApJ*, 740, 58
- Machalski, J., Jamrozy, M., Stawarz, Ł., et al. 2016, *A&A*, 595, A46
- Martinez-Cantin, R., Tee, K., & McCourt, M. 2017, arXiv:1712.04567
- Massaro, F., Álvarez-Crespo, N., Capetti, A., et al. 2019, *ApJS*, 240, 20
- Massaro, F., Capetti, A., Paggi, A., et al. 2020, *ApJS*, 247, 71
- McCarthy, P. J. 1993, *ARA&A*, 31, 639
- Meisenheimer, K., Roser, H.-J., Hiltner, P. R., et al. 1989, *A&A*, 219, 63
- Miley, G., & De Breuck, C. 2008, *A&ARv*, 15, 67
- Mingo, B., Croston, J. H., Hardcastle, M. J., et al. 2019, *MNRAS*, 488, 2701
- Ostorero, L., Morganti, R., Diaferio, A., et al. 2017, *ApJ*, 849, 34
- O'Sullivan, S. P., Gaensler, B. M., Lara-López, M. A., et al. 2015, *ApJ*, 806, 83
- O'Sullivan, S. P., Machalski, J., Van Eck, C. L., et al. 2019, *A&A*, 622, A16
- Owen, F. N., & Ledlow, M. J. 1997, *ApJS*, 108, 41
- Owen, F. N., White, R. A., & Burns, J. O. 1992, *ApJS*, 80, 501
- Predehl, P., Andriutschke, R., Arefiev, V., et al. 2021, *A&A*, 647, A1
- Rines, K., Geller, M. J., Diaferio, A., et al. 2002, *AJ*, 124, 1266
- Sikora, M., Stawarz, Ł., & Lasota, J.-P. 2007, *ApJ*, 658, 815
- Sohn, J., Geller, M. J., Walker, S. A., et al. 2019, *ApJ*, 871, 129
- Stawarz, Ł., Ostorero, L., Begelman, M. C., et al. 2008, *ApJ*, 680, 911
- Stawarz, Ł., Szostek, A., & Cheung, C. C. 2014, *ApJ*, 794, 164
- Turner, R. J., Drouart, G., Seymour, N., et al. 2020, *MNRAS*, 499, 3660
- Walker, S., Simionescu, A., Nagai, D., et al. 2019, *SSRv*, 215, 7
- Wing, J. D., & Blanton, E. L. 2011, *AJ*, 141, 88
- Yu, H., Diaferio, A., Agulli, I., et al. 2016, *ApJ*, 831, 156
- Zheng, X. C., Röttgering, H. J. A., Best, P. N., et al. 2020, *A&A*, 644, A12
- Zirbel, E. L. 1997, *ApJ*, 476, 489