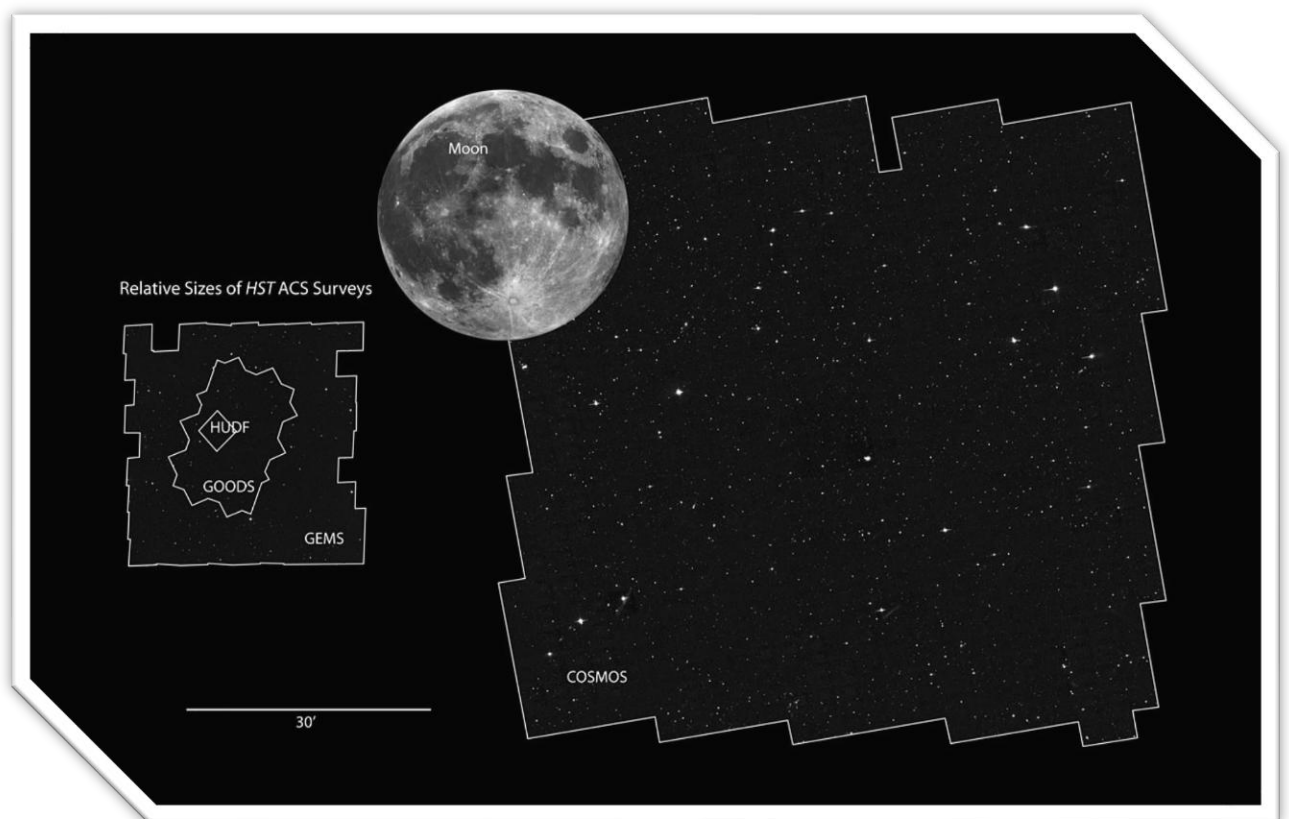


UNIVERSITY OF THESSALY
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Master's Thesis

Constraining the Radio Spectral Energy Distribution of Active and Star-forming Galaxies in the COSMOS field

LOUKAS DOUKAS



Submitted in partial fulfillment of the requirements for the Master's Degree in Applied
Physics

LAMIA 2025

© 2025 Loukas Doukas

The approval of this thesis by the Department of Physics of the Faculty of Sciences at the University of Thessaly does not imply acceptance of the author's views (Law 5343/32, Article 202, Paragraph 2).

Εγκρίθηκε από τα Μέλη της Τριμελούς Εξεταστικής Επιτροπής:

Πρώτος Εξεταστής Φίλιππος Σοφός
(Επιβλέπων) Επίκουρος Καθηγητής, Τμήμα Φυσικής, Πανεπιστήμιο Θεσσαλίας

Δεύτερος Εξεταστής Ελένη Βαρδουλάκη
 Αστροφυσικός - Εθνικό Αστεροσκοπείο Αθηνών (Εκπρόσωπος της
 Διεθνούς Αστρονομικής Ένωσης για την Εκπαίδευση (NAEC-GR))

Τρίτος Εξεταστής Θεόδωρος Καρακασίδης
 Καθηγητής, Τμήμα Φυσικής, Πανεπιστήμιο Θεσσαλίας

It was approved by the members of the three-member examination committee:

First Examiner Filippos Sofos
(Supervisor) Assistant Professor, Department of Physics, School of Sciences,
 University of Thessaly

Second Examiner Eleni Vardoulaki
 Astrophysicist - National Observatory of Athens (Representative of the
 International Astronomical Union for Education (NAEC-GR))

Third Examiner Theodoros Karakasidis
 Professor, Department of Physics, School of Sciences, University of
 Thessaly

Πρόλογος

Πρώτα απ' όλα, θέλω να ευχαριστήσω την επιβλέπουσα της διπλωματικής εργασίας μου, **Δρ. Ελένη Βαρδουλάκη**, αστροφυσικό με διδακτορικό από το Πανεπιστήμιο της Οξφόρδης και ειδίκευση στους ενεργούς γαλαξιακούς πυρήνες και στη μελέτη τους με τη χρήση δεδομένων από το ευρύ φάσμα της ηλεκτρομαγνητικής ακτινοβολίας, για την πολύτιμη καθοδήγηση, την αδιάκοπη υποστήριξη και τη συνεργασία μας στο παρόν εγχείρημα, το οποίο φέραμε εις πέρας με επιτυχία. Επίσης, ευχαριστώ θερμά τον συνεπιβλέποντα της μεταπτυχιακής μου διατριβής, **Επίκουρο Καθηγητή κ. Φίλιππο Σοφό**, με γνωστικό αντικείμενο την Υπολογιστική Φυσική Συμπυκνωμένης Ύλης, για τη βοήθεια, την εποικοδομητική του συμβολή και τις εύστοχες επιστημονικές παρατηρήσεις του κατά τη διάρκεια της εργασίας μου. Θερμές ευχαριστίες εκφράζω και προς τον **Καθηγητή κ. Θεόδωρο Καρακασίδη**, Πρόεδρο του Τμήματος Φυσικής του Πανεπιστημίου Θεσσαλίας, με γνωστικό αντικείμενο την Εφαρμοσμένη Φυσική, ο οποίος συμμετείχε ως τρίτος εξεταστής στην τριμελή εξεταστική επιτροπή της διπλωματικής μου εργασίας, για την προσεκτική ανάγνωση του κειμένου και για τις πολύτιμες υποδείξεις του. Ιδιαίτερη μνεία οφείλω στον **Dr. Krešimir Tisanić**, ο οποίος εκπόνησε το διδακτορικό του στο Ruđer Bošković Institute (University of Zagreb), για τη γενναιόδωρη διάθεση των δεδομένων και καταλόγων από τις δημοσιεύσεις του, που αποτέλεσαν τη βάση για την προσπάθειά μας να αναπαράγουμε τα γραφήματα των εργασιών του σχετικά με τους AGN και τους SFG. Θα ήθελα, επίσης, να ευχαριστήσω τους φίλους μου για την ηθική υποστήριξη και την ενθάρρυνσή τους καθ' όλη τη διάρκεια αυτής της απαιτητικής αλλά και δημιουργικής πορείας. Τέλος, ευχαριστώ την οικογένειά μου για την κατανόησή που επέδειξε, ιδιαίτερα κατά τη διάρκεια των τελευταίων μηνών της προσπάθειάς μου αλλά και για την ολόψυχη αγάπη και υποστήριξή τους όλα αυτά τα χρόνια. Αφιερώνω αυτή την εργασία στην πρώτη μου ξαδέρφη *Νάντια*.

Λουκάς Δούκας

Prologue

First and foremost, I would like to express my sincere gratitude to my thesis supervisor, **Dr. Eleni Vardoulaki**, an astrophysicist with a PhD from the University of Oxford and expertise in active galactic nuclei and their study using data across the electromagnetic spectrum, for her invaluable guidance, unwavering support, and our fruitful collaboration on this project, which we successfully brought to completion. I also warmly thank the co-supervisor of my master's thesis, **Assistant Professor Dr. Philippos Sofos**, whose academic field is Computational Condensed Matter Physics, for his assistance, constructive input, and insightful scientific comments throughout the course of this work. I extend my sincere appreciation to **Professor Dr. Theodoros Karakasidis**, Chair of the Department of Physics at the University of Thessaly, whose academic expertise lies in Applied Physics, and who served as the third examiner in the three-member evaluation committee of my thesis, for his careful reading of the manuscript and his valuable suggestions. I would like to make a special acknowledgment to **Dr. Krešimir Tisanić**, who completed his PhD at the Ruđer Bošković Institute (University of Zagreb), for generously providing the data and catalogs from his publications, which formed the foundation of our effort to reproduce the plots presented in his work on AGN and SFG populations. I would also like to thank, my friends for their moral support and encouragement throughout this demanding yet creatively fulfilling journey. Finally, I thank my family for the understanding they showed, especially during the final months of my effort, and for their wholehearted love and support throughout all these years. I dedicate this thesis to my first cousin, *Nantia*.

Loukas Doukas

Μελέτη της Κατανομής Φασματικής Ενέργειας Ραδιοκυμάτων των Ενεργών και Αστρογόνων Γαλαξιών στο πεδίο COSMOS

Λουκάς Δούκας

Πανεπιστήμιο Θεσσαλίας, Τμήμα Φυσικής, 2025

Επιβλέπων Καθηγητής: Φίλιππος Σοφός, [Επίκουρος Καθηγητής](#)

Συνεπιβλέπουσα: Δρ. Ελένη Βαρδουλάκη, [Αστροφυσικός ΕΑΑ](#)

Περίληψη

Η παρούσα εργασία μελετά τη μέση κατανομή φασματικής ενέργειας ραδιοκυμάτων των ενεργών (active galactic nuclei-AGN) και αστρογόνων γαλαξιών (star-forming galaxies SFG) στο πεδίο COSMOS, χρησιμοποιώντας δεδομένα πολλαπλών συχνοτήτων (LOFAR 144MHz (median rms: 100 μ Jy/beam, PI: Vardoulaki; project DDT19_002), GMRT 610MHz (median rms: 39 μ Jy/beam, PI: H. R. Klöckner; project 11HRK01), GMRT 325 MHz (median rms: 97 μ Jy/beam, PI: S. Croft; project 07SCB01), VLA-COSMOS 1.4 GHz Joint Project (median rms: 12 μ Jy/beam, (Schinnerer, et al., 2010)) και VLA-COSMOS 3 GHz Large Project (median rms: 2.3 μ Jy/beam, (Smolcic, et al., 2017))). Σκοπός μας είναι να αξιοποιηθούν προηγούμενες μελέτες και να αναπροσαρμοστεί η κλίση του μέσου ραδιοφασματικού δείκτη (α) που υπολογίζουν οι Tisacic et al. (2019, 2020), κάνοντας χρήση των νέων παρατηρήσεων στο πεδίο COSMOS με το International LOFAR Telescope (ILT). Η μεθοδολογία περιλαμβάνει τη δημιουργία καταλόγου για το πεδίο COSMOS στα 144MHz με τρόπο που μας δίνει τη δυνατότητα να μετρήσουμε με ακρίβεια την πυκνότητα ροής στα ραδιοκύματα των σημειακών και των εκτενών πηγών αλλά και να υπολογίσουμε τη θέση και το σχήμα τους σε δισδιάστατη προβολή στον ουρανό. Οι πληθυσμοί που ελέγχουμε είναι ενεργοί γαλαξιακοί πυρήνες στους οποίους η εκπομπή στα ραδιοκύματα υπερβαίνει την τιμή που μπορεί να παραχθεί μόνο από αστρογένεση (radio excess AGN - RxAGN). Επίσης επιλέγουμε συγκεκριμένους πληθυσμούς ραδιογαλαξιών που παρουσιάζουν ασθενείς γραμμές εκπομπής στο οπτικό φάσμα (low excitation radio galaxies-LERGs) με χαρακτηριστικά $0 \leq z \leq 4$ & $24 \leq \log L(1.4\text{GHz}) \leq 26$. Για τους αστρογόνους γαλαξίες επιλέξαμε αυτούς με υψηλό ρυθμό αστρογένεσης (highly star formation rate HSFR) με $\text{SFR} > 100$ solar masses/yr και $0.3 \leq z \leq 4$. Για την σωστή και απευθείας σύγκριση με τις υπάρχουσες έρευνες των Tisacic et al. ακολουθούμε την ίδια μεθοδολογία. Κανονικοποιούμε τις τιμές της πυκνότητας ροής στη συχνότητα $\nu_b = 4.0$ GHz και $\nu_b = 4.3$ GHz για τα AGN και SFG αντιστοίχως, και προσαρμόζουμε ένα μοντέλο της μορφής $S \propto \nu^{-\alpha}$, ώστε να ταιριάζει στα δεδομένα αριστερά (μικρότερες συχνότητες) και δεξιά (μεγαλύτερες συχνότητες) του σημείου κανονικοποίησης. Επιβεβαιώσαμε έτσι, πως το μοντέλο που

προσαρμόζεται καλύτερα στα δεδομένα είναι δύο συναρτήσεις της μορφής $S \propto \nu^{-\alpha}$ (broken power law) με τιμές για τον ραδιοφασματικό δείκτη για τα RxAGN $\alpha_1 = 0.42$ πριν τα 4GHz και $\alpha_2 = 1.16$ μετά τα 4GHz, ενώ για τα HSFG είναι $\alpha_1 = 0.58$ πριν τα 4.3GHz και $\alpha_2 = 0.94$ μετά τα 4.3GHz. Η προσθήκη των δεδομένων στα 144 MHz βελτιώνει σημαντικά τον προσδιορισμό της συμπεριφοράς του ραδιοφασματικού δείκτη στις χαμηλές συχνότητες, επιτρέποντας τον εντοπισμό φαινομένων όπως η καμπυλότητα (curvature) του SED και η πιθανή παρουσία απορρόφησης λόγω self ή free-free absorption. Οι συγκεκριμένες τιμές πιθανόν να διαφοροποιηθούν, όταν χρησιμοποιήσουμε και τα δεδομένα των 60 MHz για το πεδίο COSMOS ή/και αν μέσω τεχνικών VLBI διορθώσουμε το beam από επίμηκες σε κυκλικό.

Λέξεις Κλειδιά: *Radio SED, AGN, Star-forming galaxies, COSMOS field, LOFAR, Broken power-law.*

Constraining the Radio Spectral Energy Distribution of Active and Star-forming Galaxies in the COSMOS field

Loukas Doukas

University of Thessaly, Department of Physics, 2025

Supervisor: Filippas Sofos, Assistant Professor

Co-supervisor: Dr. Eleni Vardoulaki, Astrophysicist NOA

Abstract

This study investigates the average radio spectral energy distribution (RSED) of active galactic nuclei (AGN) and star-forming galaxies (SFG) in the COSMOS field, utilizing multi-frequency data (*LOFAR 144MHz (median rms: 100 μ Jy/beam, PI: Vardoulaki; project DDT19_002) , GMRT 610MHz (median rms: 39 μ Jy/beam, PI: H. R. Klöckner; project 11HRK01), GMRT 325 MHz (median rms: 97 μ Jy/beam, PI: S. Croft; project 07SCB01) , VLA-COSMOS 1.4 GHz Joint Project (median rms: 12 μ Jy/beam, (Schinnerer, et al., 2010)) και VLA-COSMOS 3 GHz Large Project (median rms: 2.3 μ Jy/beam, (Smolcic, et al., 2017)). Our aim is to build upon previous studies and refine the slope of the mean radio spectral index (α) derived by Tisanić et al. (2019, 2020), leveraging new observations in the COSMOS field with the International LOFAR Telescope (ILT). The methodology includes the construction of a 144 MHz source catalog for the COSMOS field, enabling accurate measurements of the flux density, position, and morphology of both compact and extended sources in two-dimensional sky projection. The populations examined include radio-excess AGN (RxAGN), where the radio emission exceeds the level expected from star formation alone, as well as specific classes of radio galaxies exhibiting weak optical emission lines (low-excitation radio galaxies, LERGs) within $0 \leq z \leq 4$ & $24 \leq \log L(1.4\text{GHz}) \leq 26$. For star-forming galaxies, we selected those with high star formation rates (HSFR; $\text{SFR} > 100$ solar masses/yr) and $0.3 \leq z \leq 4$. For a consistent and direct comparison with the results of Tisanić et al., we adopt the same methodology. The flux densities are normalized at $\nu_b = 4.0$ GHz for AGN and $\nu_b = 4.3$ GHz for SFGs, and a model of the form $S \propto \nu^{-\alpha}$ is fitted to the data on either side of the normalization frequency. We find that the most suitable model is $S \propto \nu^{-\alpha}$, a broken power law, with spectral indices for RxAGN of $\alpha_1 = 0.42$ below 4 GHz and $\alpha_2 = 1.16$ above 4 GHz, and for HSFG of $\alpha_1 = 0.58$ below 4.3 GHz and $\alpha_2 = 0.94$ above 4.3 GHz. The inclusion of the 144 MHz data significantly improves the characterization of the spectral index behavior at low frequencies, enabling the identification of features such as SED curvature and possible self- or free-free absorption effects. These values may be further*

refined with the use of 60 MHz data for the COSMOS field and/or with VLBI techniques to correct the beam from elongated to circular.

Keywords: *Radio SED, AGN, Star-forming galaxies, COSMOS field, LOFAR, Broken power-law.*

Chapter 1: Introduction.....	1
1.1 Overview of the Research Topic.....	1
1.2 Importance and Motivation.....	2
1.3 Historical Background, Review of SED Literature, and Research Contribution.....	3
1.3.1. Historical Background	3
1.3.2. Literature Review on Radio Spectral Energy Distributions.....	4
1.3.3. Research Contribution.....	12
1.4 Theoretical Framework.....	13
1.4.1 COSMOS field.....	13
1.4.2 Physical Models of Radio Emission	15
1.4.3 Active Galactic Nuclei (AGN).....	21
1.4.4 Star-Forming Galaxies (SFGs).....	22
1.4.5 Giant Radio Galaxies (GRGs).....	24
1.4.6 Underlying Principles and Equations.....	25
1.5 Thesis Outline and Cosmology	28
Chapter 2: Data and Methods	30
2.1 LOFAR 144 MHz Observations in COSMOS.....	30
2.2 Ancillary Radio Datasets in COSMOS: VLA (3 GHz & 1.4 GHz), GMRT (610/325 MHz)	30
2.2.1 The VLA-COSMOS 3 GHz Large Project	31
2.2.2 VLA-COSMOS 1.4 GHz Large Project.....	32
2.2.3 GMRT 610/325 MHz Surveys of COSMOS	32
2.3 Understanding the datasets: Counterpart, PyBDSF, and SuperDeblended.....	34
2.3.1 Catalogue Counterpart	34
2.3.2 Catalogue PyBDSF	36
2.3.3 Catalogue SuperDeblended.....	38
2.4 Catalogue LOFAR-masks generation	39
2.4.1 Source detection, flux measurement, and centroid determination	39
2.4.2 Flux extraction, beam correction, and catalogue preparation	41
2.4.3 Identification and deblending of the two GRGs in the COSMOS field.....	42
2.5 Correction, error analysis and matching process of radio source measurements.....	45
2.5.1 Correction of flux densities through identification and removal of secondary blended sources	45
2.5.2 Error estimation in position and flux density, SNR and RMS measurements	47
2.5.3 Matching process cat masks & deblended, counterpart and pybdsf	49
2.6 Final Catalogue/dataset.....	52
2.7 Data Visualization and Classification	53
Chapter 3: Analysis & Results.....	57
3.1 Description of the Sample.....	57

3.1.1 Subsample for AGNs	57
3.1.2 Subsample for SFGs.....	58
3.2 Normalization Procedures for Radio SED Construction	59
3.2.1 General Normalization Procedure for both Subsamples	59
3.2.2 AGN Subsample: Normalization Procedure (Specific Implementation)	62
3.2.3 SFG Subsample: Normalization Procedure (Specific Implementation)	62
3.3 Data Representation	63
3.3.1 AGN - RSED	64
3.3.2 SFG - RSED.....	67
3.3.3 SFG Individual Cases.....	70
3.4 Numerical Results.....	71
3.4.1 Quantitative Analysis of the AGN Sample	71
3.4.2 Quantitative Analysis of the SFG Sample	72
Chapter 4: Discussion	73
4.1 Interpretation and Scientific Assessment of Results.....	73
4.1.1 AGN Radio SEDs: Best-Fit Approaches and Results.....	73
4.1.2 SFG Radio SEDs: Best-Fit Approaches and Results	75
4.2 Limitations and Challenges.....	77
4.3 Future Work and Prospects	78
Chapter 5: Conclusion & Outlook	80
5.1 Summary of Key Findings and Major Conclusions.....	80
5.2 Implications and Significance.....	81
5.3 Outlook for Future Research.....	82
Chapter 6: References.....	82
6.1 Online Reference Links	86
Chapter 7: Appendices.....	87
7.1 List of figures.....	87
7.2 List of tables.....	89
7.3 MATLAB Pseudo-Codes.....	89

Glossary of Acronyms

ACS: Advanced Camera for Surveys

AGN: Active Galactic Nucleus

ASKAP: Australian Square Kilometre Array Pathfinder

AUI: Associated Universities, Inc.

AURA: Association of Universities for Research in Astronomy

BMAJ: Beam Major Axis

BMIN: Beam Minor Axis

BPL: Broken Power-Law

CD: Coordinate Description matrix

COSMOS: Cosmic Evolution Survey

CROTA: Coordinate Rotation angle

CRPIX: Coordinate Reference Pixel

CRVAL: Coordinate Reference Value

DDT: Director's Discretionary Time

DEC: Declination

ELAIS: European Large-Area ISO Survey

EMU: Evolutionary Map of the Universe

ESA: European Space Agency

FIR: Far-Infrared

FIRST: Faint Images of the Radio Sky at Twenty cm

FITS: Flexible Image Transport System

FR: Fanaroff-Riley

FWHM: Full Width at Half Maximum

GMRT: Giant Metrewave Radio Telescope

GOODS: Great Observatories Origins Deep Survey

GRG: Giant Radio Galaxy

HBA: High Band Antenna

HERG: High-Excitation Radio Galaxies

HLAGN: High-Luminosity AGN

HSC: Hyper Suprime-Cam

HSFG: Highly Star-Forming Galaxy

HSFR: High star formation rates

HST: Hubble Space Telescope

ICM: Intracluster Medium

IGM: Intergalactic Medium

ILT: International LOFAR Telescope

IR: Infrared

IRRC: Infrared–Radio Correlation

ISM: Interstellar Medium

JWST: James Webb Space Telescope

Jy: Jansky

LERG: Low-Excitation Radio Galaxy

LOFAR: Low Frequency Array

LR: Likelihood Ratio

MATLAB: MATrix LABoratory

MCMC: Markov Chain Monte Carlo

MIGHTEE: MeerKAT International GHz Tiered Extragalactic Exploration

MIPS: Multiband Imaging Photometer for Spitzer

MIR: Mid-Infrared

MLAGN: Moderate-Luminosity AGN

MSMF: Multi-Scale Multi-Frequency

NASA: National Aeronautics and Space Administration

NGC: New General Catalogue

NIR: Near-Infrared

NRAO: National Radio Astronomy Observatory

NVSS: NRAO VLA Sky Survey

ODR: Orthogonal Distance Regression

PACS: Photodetector Array Camera and Spectrometer

PI: Principal Investigator

PL: Power-Law

PSF: Point Spread Function

PyBDSF: Python Blob Detection and Source Finder

RA: Right Ascension

RFI: Radio Frequency Interference

RIT: Rochester Institute of Technology

RMS: Root Mean Square

RSED: Radio Spectral Energy Distribution

SCUBA-2: Submillimeter Common-User Bolometer Array 2

SED: Spectral Energy Distribution

SFG: Star-Forming Galaxy

SFR: Star Formation Rate

SKA: Square Kilometre Array

SNR: Signal-to-Noise Ratio

SPL: Single Power-Law

SSA: Synchrotron Self-Absorption

STScI: Space Telescope Science Institute

TGSS: TIFR GMRT Sky Survey

TOPCAT: Tool for OPerations on Catalogues And Tables

ULIRG: UltraLuminous InfraRed Galaxies

UMD: University of Maryland

UV: Ultraviolet

VIMOS: Visible Multi-Object Spectrograph

VLA: Very Large Array

VLT: Very Large Telescope

WCS: World Coordinate System

Chapter 1: Introduction

1.1 Overview of the Research Topic

The focus of this research is the **radio spectral energy distribution (RSED)** of active and star-forming galaxies in the COSMOS field. The COSMOS field is an extensively studied region of the sky that benefits from multi-wavelength observations, ranging from X-rays to radio frequencies, providing an unparalleled dataset for studying galaxy evolution across cosmic time. By analyzing the radio SEDs of galaxies in this field, the study aims to explore how emission at radio wavelengths correlates with physical processes, such as star formation and black hole activity. The radio SED, primarily driven by synchrotron and thermal free-free emission, serves as a vital tool for disentangling contributions from star formation and active galactic nuclei (AGN) at different frequency ranges, especially below 10 GHz. This separation is critical for deriving accurate physical properties, such as star formation rates and AGN feedback efficiency, offering insights into the life cycle of galaxies (Scoville, et al., 2007); (Tisanic, et al., 2020)).

In this context, the value of the average **radio spectral index (α)** plays a critical role in determining key galaxy properties and in interpreting results from radio surveys. Even small deviations in α can significantly affect the derived radio luminosities and star-formation rates, impacting our understanding of galaxy evolution on a cosmological scale. While many past surveys have assumed a fixed spectral index (typically $\alpha = 0.7$) for all sources, recent studies demonstrate that this simplification can introduce systematic biases in statistical analyses (Tisanic, et al., 2020).

The importance of studying radio SEDs lies in their ability to reveal how galaxies interact with their environments and regulate their star formation over time. AGN, for example, influence their surroundings through mechanical feedback, often quenching star formation in massive galaxies (Delvecchio, et al., 2017). On the other hand, star-forming galaxies exhibit radio emission tightly correlated with recent star formation, providing a dust-free probe of these processes, which makes radio emission a complementary tool to infrared and optical observations in tracing star formation activity (Condon, 1992); (Tisanic, et al., 2019)). Understanding these phenomena in the COSMOS field allows us to observe these mechanisms in action across a wide range of galaxy populations and redshifts, making it possible to track their evolution over billions of years (Tisanic, et al., 2020). Furthermore, the results of this research will contribute to improving models of galaxy evolution and preparing for the capabilities of next-generation radio observatories like the Square Kilometer Array (SKA) (Vardoulaki, et al., 2021).

The central scientific question addressed by this study is: ***How can the inclusion of new LOFAR radio data from the COSMOS field refine the average radio spectral energy distribution (RSED) of active and star-forming galaxies, and how does this affect the inferred spectral slope and its dependence on redshift and physical properties?*** In this research, we adopt a statistical analysis approach rather than focusing on individual case studies. This methodology allows us to leverage the comprehensive multi-wavelength coverage available in the COSMOS field to derive robust average properties of galaxy populations (Tisanic, et al., 2020); (Delvecchio, et al., 2017)). Specifically, we aim to disentangle the radio emission due to AGN from that arising from star formation, investigating how these contributions vary with galaxy properties such as redshift, stellar mass, and star formation rate (Scoville, et al., 2007); (Condon, Cotton, & Broderick, 2002)). A critical component of this analysis is constraining the spectral index, which characterizes the slope of the radio SED, and identifying any deviations from simple or broken power-law models. For example, Tisanic et al. (2019) found curvature in the average radio SED of star-forming galaxies at low frequencies, likely due to synchrotron aging, while in a separate study (Tisanic et al. 2020) deviations in AGN were interpreted as evidence of mixed radio-mode activity or environmental effects.

These deviations may signal processes such as synchrotron aging or changes in the relative contributions of thermal and non-thermal emission (Tisanic, et al., 2019); (Vardoulaki, et al., 2021)). By statistically linking these findings to the broader properties of galaxy populations, the study provides a comprehensive view of how star formation and AGN activity co-evolve across cosmic time and varying environments (Delvecchio, et al., 2017).

1.2 Importance and Motivation

Understanding the radio spectral energy distribution (RSED) is essential for uncovering the astrophysical processes that shape galaxies over cosmic time. Unlike optical or infrared diagnostics, which are significantly affected by dust attenuation and obscuration, the radio SED offers a dust-independent view of galactic activity, enabling the study of obscured star formation and AGN feedback even in the most heavily dust-enshrouded environments. By analyzing the SED, it is possible to distinguish between synchrotron and free-free emission, two key mechanisms that provide insights into the evolution of cosmic structures such as AGN and star-forming galaxies (SFGs). For AGN, deviations from simple power-law models are not just spectral features, but carry crucial information about the underlying physical mechanisms. These deviations, often observed in AGN-dominated radio SEDs, help constrain jet lifetimes, energy loss mechanisms, and environmental density, providing insights into synchrotron aging, self-absorption, and the influence of the surrounding medium (Tisanic, et al., 2020). These features offer critical information about the growth of supermassive black holes, their feedback processes, and their role in shaping the radio luminosity of galaxies throughout the universe (Delvecchio, et al., 2017).

The radio SED allows for robust estimation of star formation rates (SFRs), offering an unimpeded view of star-forming activity even in dusty environments (Condon, Radio Emission from Normal Galaxies, 1992). The tight infrared-radio correlation enables robust modeling of SEDs, helping to identify the contribution of thermal free-free emission and synchrotron radiation from supernova remnants (Tisanic, et al., 2019). Accurate characterization of the SED is also crucial for applying K-corrections, ensuring reliable measurements of intrinsic galaxy properties across different redshifts (Tisanic, et al., 2020). By examining the flattening of the radio spectrum at higher frequencies — caused by increasing thermal free-free emission — and the steepening at lower frequencies due to diffuse synchrotron components, the SED of star-forming galaxies offers valuable insight into the interplay between star formation and radio emission processes.

Distinguishing AGN from SFGs based on their SEDs relies on the fundamentally different origins of their radio emission, which produce distinct spectral signatures that can be identified and modeled accordingly. In active galaxies, radio waves are generated by synchrotron radiation from relativistic jets or compact cores powered by supermassive black holes. This produces distinctive SED features, including flat or inverted spectra at low frequencies due to self-absorption and steep slopes at higher frequencies caused by synchrotron aging (Vardoulaki, et al., 2021). Such characteristics allow AGN to be identified through their deviations from the SEDs of SFG (Tisanic, et al., 2020).

In contrast, the radio emission in SFGs primarily arises from massive star formation processes, such as synchrotron radiation from supernova remnants and thermal free-free emission from ionized gas in HII regions (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002). The SEDs of SFGs typically show a steep slope at lower frequencies and a flatter component at higher frequencies due to the thermal contribution (Tisanic, et al., 2019). By analyzing these features, it becomes possible to separate the AGN-related radio emission from the star-forming processes, even in galaxies where both mechanisms coexist. Correlations like the infrared-radio relationship further assist in identifying excess radio emission, a hallmark of active nuclei activity (Delvecchio, et al., 2017).

1.3 Historical Background, Review of SED Literature, and Research Contribution

1.3.1. Historical Background

The history of radio astronomy began in the early 20th century with a series of pioneering discoveries that opened a new window onto the universe. The first detection of cosmic radio waves by **Karl Jansky** in 1932 marked the birth of radio astronomy, as he identified radio emission from the Milky Way while investigating sources of interference in transatlantic telecommunications. In 1937, Grote Reber advanced this field by designing and building the earliest radio telescope specifically intended for astronomical research and conducted systematic sky surveys, confirming that celestial objects could emit significant radio waves. These early observations laid the foundation for understanding that radio emission originates from both **thermal** and **non-thermal** processes, ultimately leading to the identification of synchrotron radiation as a dominant emission mechanism.

By the 1950s, the advent of interferometry and high-resolution radio imaging allowed astronomers to associate radio sources with optical counterparts. The work of **Martin Ryle** and **Antony Hewish**, pioneers in the development of radio interferometry, enabled the precise localization of extragalactic radio sources, leading to the discovery that some of the most luminous radio emitters were quasars—compact yet highly energetic objects powered by accretion onto supermassive black holes (SMBHs). As a result, radio data emerged as a key approach for examining AGN and their significance in the evolutionary processes of galaxies. Around the same time, radio studies of nearby galaxies, particularly by **Allan Sandage** and **Thomas Matthews**, revealed that synchrotron radiation was a key tracer of star formation, as relativistic electrons produced in supernova remnants were found to be a primary source of radio emission in star-forming galaxies.

A major milestone in the historical development of radio astronomy was the discovery of the **infrared-radio correlation (IRRC)**. First identified by van der Kruit (1971) and later quantified by Helou et al. (1985), the IRRC established a tight empirical relationship between the total infrared and radio luminosities of star-forming galaxies. This correlation revealed that both infrared and radio emission trace recent star formation, and provided a strong justification for using radio continuum as a dust-unbiased SFR indicator (Condon, 1992). The IRRC underpins much of the current methodology in extragalactic radio astronomy, especially in high-redshift and dusty galaxy studies.

As observational capabilities improved, theoretical models of radio emission were developed to describe the observed spectra. The introduction of **power-law models** allowed astronomers to quantify radio sources based on their α , which is defined through the relation:

$$S_\nu \propto \nu^{-\alpha} \quad (1.1)$$

in which S_ν refers to the flux density at the specific frequency ν and α stands for the spectral index. This frequency ν refers to the **rest-frame frequency**, after correcting for the cosmological redshift z . The spectral index provides critical insights into the physical conditions of the emitting source (Condon, Radio Emission from Normal Galaxies, 1992). For pure synchrotron emission, its value typically ranges between **0.5 and 1.0**, depending on the energy distribution of the relativistic electrons. In compact radio sources, self-absorption can flatten the spectrum at lower frequencies, while in extended lobes of radio galaxies, synchrotron aging steepens the radio spectrum at higher frequencies becomes steeper as relativistic electrons gradually lose energy.

While power-law models provided a first-order approximation of radio SEDs, they were found to be insufficient in explaining spectral deviations observed in many sources. This led to the development of **broken power-law models** and hybrid models incorporating multiple emission mechanisms,

including thermal free-free emission from HII regions (Tisanic, et al., 2019 & 2020). These refinements allowed for a more accurate interpretation of radio observations, particularly in studies where AGN and star formation processes coexist within the same galaxy. Crucially, the availability of multi-frequency data across a wide spectral range enables accurate modeling of radio SEDs, as it allows for the identification of spectral curvature, breaks, and the presence of multiple emission mechanisms.

The evolution of radio astronomy has been driven by advancements in observational facilities, particularly in the development of large radio interferometers. Several major radio telescopes have significantly advanced our understanding of galaxy evolution, including:

- o **Very Large Array (VLA,1980):** One of the most powerful interferometers, providing deep, high-resolution radio surveys across multiple frequencies, crucial for identifying AGN and star-forming galaxies.
- o **Giant Metrewave Radio Telescope (GMRT,2001):** Covers intermediate-frequency observations (~150-1500 MHz), facilitating deep surveys of the radio sky and contributing to the discovery of high-redshift radio galaxies.
- o **Low-Frequency Array (LOFAR,2012):** Extends radio observations to very low frequencies (~30-240 MHz), enabling studies of synchrotron self-absorption and spectral turnover.
- o **MeerKAT (2018):** A highly sensitive radio array in South Africa, operating primarily in the 1–3.5 GHz range ((L-band (900–1670 MHz) and S-band (1.75–3.5 GHz)), serving as a precursor to SKA and significantly improving our ability to detect faint extragalactic radio sources. Major surveys like *MIGHTEE* (MeerKAT International GHz Tiered Extragalactic Exploration) leverage MeerKAT's capabilities to provide deep, wide-field observations of fields such as COSMOS, enabling statistical studies of galaxy evolution and AGN activity.
- o **Square Kilometer Array (SKA,2028):** The upcoming radio observatory is set to transform our knowledge of radio sources, offering unmatched sensitivity and resolution (with a frequency range of approximately 0.5 GHz to 25–50 GHz), enabling the identification of faint sources at high redshifts and facilitating the exploration of the universe's earliest galaxy formation periods.

These advances in radio astronomy have fundamentally deepened our insight into the energetic processes governing galaxy evolution. By enabling precise measurements across a broad spectral range, modern radio observations have allowed us to disentangle the complex interplay between synchrotron emission from relativistic electrons and thermal emission from ionized gas. As new facilities further push the frontiers of sensitivity and frequency coverage, we stand to uncover even fainter and more distant populations of radio sources, ultimately refining our understanding of the formation and transformation of galaxies throughout cosmic history.

1.3.2. Literature Review on Radio Spectral Energy Distributions

1.3.2.1 Historical Development of SED and RSED

The study of galaxies across the electromagnetic spectrum has long relied on constructing their Spectral Energy Distributions (SEDs) to probe their physical properties. Early observations in the radio regime (e.g., (Helou, Soifer, & Rowan-Robinson, 1985); (Condon, Radio Emission from Normal Galaxies, 1992)) revealed the existence of both thermal and non-thermal components in the radio emission of galaxies, highlighting the importance of star formation and AGN activity in shaping the observed spectra (see Figure 1). Further advancing the field, (Helou, Soifer, & Rowan-Robinson, 1985) demonstrated that the far-infrared and radio outputs of star-forming galaxies are closely correlated, establishing a crucial observational benchmark for studies of star formation across diverse galaxy populations. This empirical relation revealed a deep connection between the thermal emission from dust

heated by young stars and the non-thermal synchrotron emission arising from supernova remnants, reinforcing the use of radio observations as a reliable tracer of star formation activity.

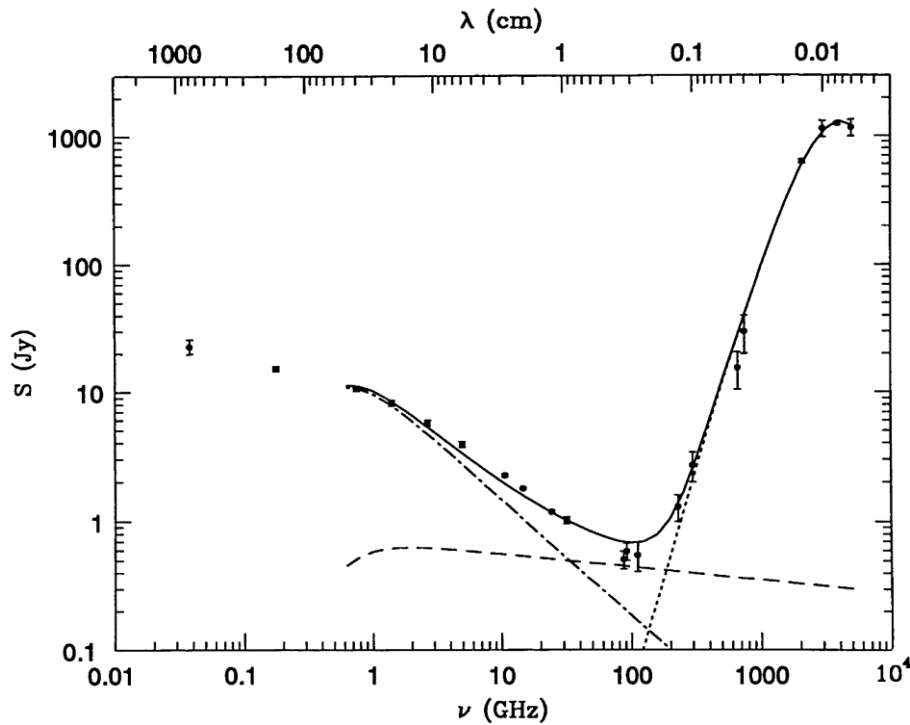


Figure 1: Observed radio to far-infrared (FIR) spectrum of the starburst galaxy M82 (Klein et al. 1988, Carlstrom & Kronberg 1991), decomposed into synchrotron (dot-dash), free-free (dashed), and thermal dust (dotted) emission components. Below ~ 1 GHz, free-free and synchrotron radiation are suppressed due to opacity in HII regions. Free-free emission peaks in the 30–200 GHz range, while thermal dust emission from ~ 45 K grains, with opacity $\propto \nu^{1.5}$, dominates at $\nu \gtrsim 200$ GHz. The plot shows flux density (Jy) versus frequency (GHz, bottom axis) and wavelength (cm, top axis) (Condon, *Radio Emission from Normal Galaxies*, 1992).

The development of deep radio surveys, such as those presented in the VLA-COSMOS projects (Schinnerer, et al., 2010); (Smolcic, et al., 2017)), allowed the systematic construction of RSEDs for large samples of galaxies. These efforts provided the first statistical insights into how radio emission correlates with galaxy properties such as stellar mass, SFR, and redshift.

A significant step forward was the introduction of high-sensitivity radio observations at multiple frequencies, enabling the direct measurement of the radio spectral index for individual sources (e.g., (Delhaize, et al., 2017); (Smolcic, et al., 2017)). Additionally, spectral ageing, caused by energy losses of cosmic ray electrons through synchrotron and inverse Compton mechanisms, results in the radio spectral index becoming steeper at higher frequencies. This approach moved beyond the assumption of a fixed spectral slope and permitted the exploration of spectral curvature and deviations associated with aging electron populations (Blundell & Rawlings, *The Spectra and Energies of Classical Double Radio Lobes*, 2000); (Blundell & Rawlings, *Spectral ageing: a new age perspective*, 2002)).

The refinement of RSED studies was further propelled by the VLA-COSMOS 3 GHz Large Project (Smolcic, et al., 2017), which combined deep and wide radio data to probe the cosmic evolution of both SFGs and AGN up to high redshifts. In parallel, advanced techniques, such as the PyBDSF software (Mohan & Rafferty, 2015), improved the detection and characterization of faint radio sources, crucial for accurate RSED measurements.

More recently, works such as those by Tisanić et al. (2019, 2020) systematically constructed average RSEDs for populations of highly star-forming galaxies and AGN separately, emphasizing the role of the radio spectrum as a tracer of underlying physical processes. These studies established the

RSED as a powerful tool not only for diagnosing star formation and AGN activity but also for investigating their coevolution with host galaxy properties across cosmic time.

Nevertheless, radio-based SEDs alone lack sensitivity to certain key galaxy properties, such as dust-obscured star formation and stellar mass, highlighting the importance of complementary multiwavelength data for a complete characterization of galaxy evolution. In summary, the historical development of SEDs and, more specifically, RSEDs reflects the broader advancement of radio astronomy, from early discoveries of emission mechanisms to the modern era of large-scale, multi-frequency surveys that enable a comprehensive understanding of galaxy evolution through radio observations.

1.3.2.2 Current understanding of RSEDs for AGN and SFG (with *Tisanić* diagrams annotation)

The characterization of the RSED has emerged as a crucial tool for studying the physical processes governing galaxy evolution. In contrast to traditional studies that focused on single-frequency radio observations, modern approaches rely on multi-frequency, rest-frame analyses that allow the direct tracing of synchrotron aging, active galactic nucleus activity, and star formation processes across cosmic time. Recent advances in large-area, deep radio surveys, such as the VLA-COSMOS 3 GHz Large Project, have enabled the construction of statistically robust average RSEDs for distinct galaxy populations, providing new insights into their evolutionary pathways.

Particularly important is the realization that a simple single power-law description of the radio SED may not suffice, especially for galaxies hosting AGN or undergoing intense star formation. Spectral curvature, flattening, or steepening at different frequencies reflects the interplay of various mechanisms, including cosmic ray electron energy losses, magnetic field variations, and environmental factors. The latest studies, including those by Tisanić et al. (2019, 2020), have employed survival analysis techniques to incorporate non-detections (upper limits) and to derive representative RSEDs separately for AGN and HSEFs, thus setting a new benchmark in the field.

RSED of AGN: Insights from (Tisanić, et al., 2020) - The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of active galactic nuclei

In their study of radio-excess AGN (RxAGN) selected from the VLA-COSMOS 3 GHz Large Project, Tisanić et al. (2020) constructed the average RSED by normalizing individual radio flux densities and binning them in rest-frame frequency space. The sample spans a redshift range of $0 \leq z \leq 4$ and radio luminosities between $(\log)L_{1.4 \text{ GHz}} = 24\text{--}26 \text{ W/Hz}$.

AGN observed in radio surveys can generally be divided into two main classes based on their optical emission-line properties and accretion modes: High-Excitation Radio Galaxies (HERGs) and Low-Excitation Radio Galaxies (LERGs). HERGs, also known as radiative-mode or quasar-mode AGN, exhibit strong optical emission lines and are powered by efficient, radiatively dominant accretion of cold gas, often at rates exceeding $\sim 1\%$ of the Eddington limit. In contrast, LERGs (also referred to as jet-mode or kinetic-mode AGN) show weak or absent emission lines and are powered by radiatively inefficient accretion of hot gas, producing most of their energy output in the form of radio jets rather than radiation. LERGs are preferentially hosted by massive, early-type galaxies with old stellar populations and low star formation rates. In this study, we focus exclusively on the LERG population, as they constitute the dominant class of radio AGN in our sample and are most relevant for tracing the interplay between radio-mode feedback and galaxy evolution in the contemporary universe (Heckman & Best, 2014).

Two models were fitted to the data: a single power-law (PL) and a broken power-law (BPL) with a fixed break frequency at 4.0 GHz (rest-frame). The single power-law fit yields a spectral index of $\alpha = 0.64 \pm 0.07$, indicative of a relatively steep radio spectrum dominated by optically thin synchrotron

emission (see Figure 2). However, the broken power-law model provides a better fit, suggesting a spectral flattening below the break frequency and steepening above it. Specifically, the spectral indices were found to be $\alpha_1 = 0.28 \pm 0.03$ below 4.0 GHz and $\alpha_2 = 1.16 \pm 0.04$ above 4.0 GHz, highlighting significant spectral curvature associated with aging electron populations and possible transitions between different emission mechanisms. This behavior is consistent with the effects of synchrotron cooling and energy losses, where lower frequencies are dominated by optically thin synchrotron emission, while higher frequencies reflect the impact of electron aging and possibly additional radiative losses.

K. Tisanic et al.: Radio SED of AGN

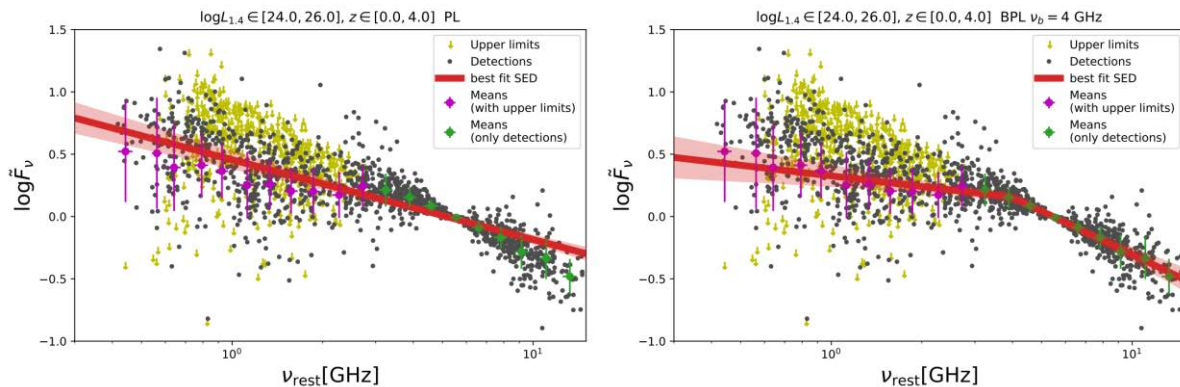


Figure 2: Radio SED of radio-excess AGN (RxAGN) as a function of rest-frame frequency (ν_{rest}) for sources with $\log L_{1.4} \in [24.0, 26.0]$ and redshift $z \in [0, 4]$, based on the VLA-COSMOS 3 GHz Large Project (Tisanic, et al., 2020). The left panel shows the best-fit single power-law model ($\alpha = 0.64 \pm 0.07$), while the right panel shows the best-fit broken power-law model ($\alpha_1 = 0.28 \pm 0.03$ below 4.0 GHz and $\alpha_2 = 1.16 \pm 0.04$ above 4.0 GHz), both overplotted with red lines and shaded confidence intervals. Gray points represent detections, yellow arrows upper limits, green squares means fluxes (detections only), and magenta squares means including upper limits using survival analysis.

The survival analysis applied in their work, which incorporates both detections and upper limits, ensures that the derived average RSED accurately represents the underlying population, accounting for observational biases at low flux densities. This result underscores the necessity of adopting more complex models than a simple power-law when interpreting the radio SEDs of AGN populations. Figure 2 illustrates that the broken power-law better captures the curvature of the average RxAGN radio SED, highlighting the importance of more complex models in describing the underlying population.

RSED of Highly Star-Forming Galaxies: Insights from (Tisanic, et al., 2019) - The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of highly star-forming galaxies

Tisanić et al. (2019) has already analyze the RSED to a sample of HSFGs, characterized by their intense infrared-derived star formation rates (see Figure 3). Similarly to the AGN analysis, individual flux densities were normalized, and the average RSED was constructed across a broad range of rest-frame frequencies.

Star-forming galaxies encompass a broad range of systems with varying rates of stellar mass assembly. Typical SFGs in the local universe exhibit moderate SFRs, generally below a few solar masses per year, and often follow the so-called star-forming "main sequence," a tight correlation between stellar mass and SFR. In contrast, HSFGs represent the upper extreme of this distribution, with SFRs often exceeding $10\text{--}100 \text{ M}_{\odot} \text{ yr}^{-1}$. These galaxies are characterized by intense infrared emission due to dust-reprocessed starlight, elevated gas fractions, and often more compact morphologies. The physical conditions in HSFGs, including higher gas densities, pressures, and magnetic field strengths, can significantly influence their radio spectral properties, with enhanced contributions from free-free emission and synchrotron radiation. The distinction between SFGs and HSFGs is thus primarily quantitative, but it also reflects underlying differences in the interstellar medium conditions and star

formation processes. In our analysis, we specifically focus on the HSFG population to investigate how extreme star formation environments shape the observed radio spectral energy distributions.

A broken power-law model was found to best describe the observed trend. The spectral index at low frequencies (below the break) is relatively flat, reflecting a mixture of thermal (free-free) and non-thermal (synchrotron) emission, while at higher frequencies, the spectrum steepens due to synchrotron aging and reduced thermal contribution. Although specific numerical values for the spectral indices were detailed in figure caption, the general behavior reflects the known spectral characteristics of star-forming galaxies, with flatter spectra at low frequencies and steeper ones at high frequencies.

Notably, Tisanić et al. employed Kaplan–Meier survival analysis to properly handle upper limits, providing robust estimates of mean flux densities and their uncertainties. The results illustrate that while star-forming galaxies can be modeled with broken power-laws, the degree of spectral curvature is less extreme than in AGN, consistent with the dominance of younger, freshly accelerated electron populations in SFGs.

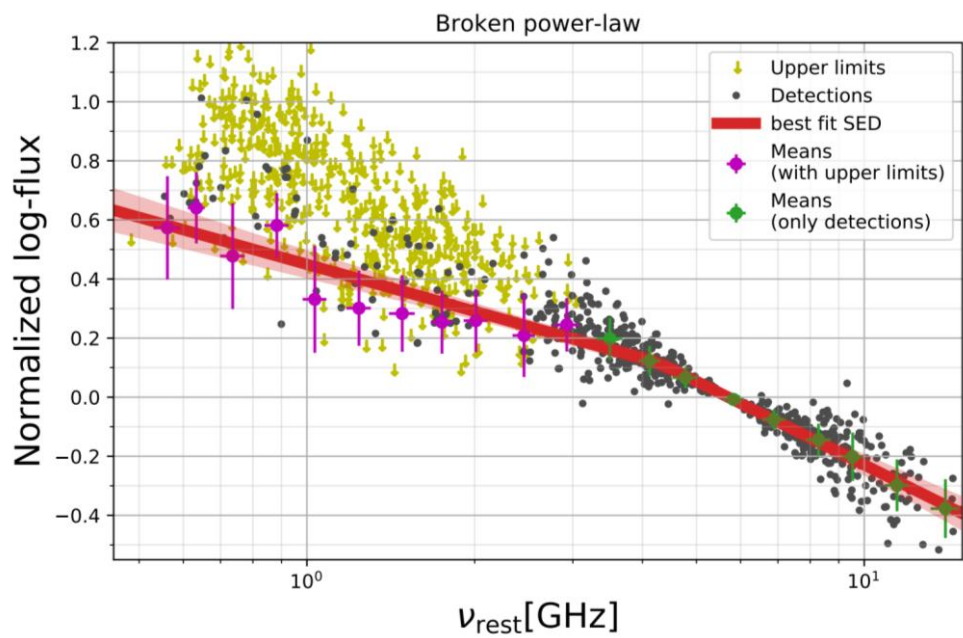


Figure 3: Radio SED of highly star-forming galaxies (HSFGs) as a function of rest-frame frequency (ν_{rest}), constructed from VLA and GMRT data in the COSMOS field (Tisanić, et al., 2019). The red line shows the best-fit broken power-law model with a break frequency at 4.3 ± 0.6 GHz, spectral index $\alpha_1 = 0.51 \pm 0.04$ below the break and $\alpha_2 = 0.98 \pm 0.07$ above. Gray points represent detections, yellow arrows upper limits, green diamonds means fluxes (detections only), and magenta squares means including upper limits using Kaplan–Meier survival analysis.

The radio SED of HSFGs is best described by a broken power-law model, reflecting the interplay between non-thermal synchrotron and thermal free-free emission at low frequencies, and synchrotron aging at higher frequencies. Quantitatively, the best-fitting broken power-law model for the HSFG sample indicates a spectral index of $\alpha_1 = 0.51 \pm 0.04$ below a break frequency of $\nu_b = 4.3 \pm 0.6$ GHz, and a steeper index of $\alpha_2 = 0.98 \pm 0.07$ above this frequency. The flattening at low frequencies is attributed to a combination of optically thin synchrotron emission and a non-negligible contribution from thermal free-free processes, whereas the steepening at higher frequencies reflects the energy losses of cosmic ray electrons via synchrotron and inverse Compton mechanisms. The relatively moderate spectral curvature compared to AGN suggests that the radio emission in HSFGs is dominated by ongoing, recent star formation rather than by aging electron populations associated with long-lived AGN activity.

1.3.2.3 Correlation of the Radio SED with Star Formation Rate, Redshift, and Stellar Mass

The RSED encapsulates crucial information about the physical conditions within galaxies, and its shape and slope can be closely linked to fundamental galaxy properties such as the SFR, stellar mass, and cosmic epoch (redshift).

Studies such as Tisanić et al. (2019, 2020) have demonstrated that galaxies with intense star formation activity, characterized by high infrared-derived SFRs, exhibit RSEDs with relatively flatter low-frequency spectral indices compared to AGN-dominated sources. This flattening is primarily attributed to the combined contribution of synchrotron and thermal free-free emission, where the thermal component becomes more significant in galaxies with higher SFRs. Consequently, the presence of substantial star formation not only elevates the overall radio luminosity but also modifies the spectral shape, particularly at frequencies below a few GHz.

The observed RSED is also significantly influenced by redshift evolution. As shown in Smolčić et al. (2017) and Delhaize et al. (2017), the radio spectral index tends to evolve mildly with increasing redshift, potentially reflecting changes in the dominant emission mechanisms or in the cosmic ray electron populations over cosmic time. In high-redshift galaxies, environmental conditions such as increased magnetic field strengths or enhanced inverse Compton losses against the cosmic microwave background can further steepen the radio spectra, particularly at higher rest-frame frequencies.

Regarding stellar mass, Darvish et al. (2016, 2017) highlighted that massive galaxies tend to host more evolved AGN activity, which can imprint characteristic signatures on their RSEDs. In AGN-dominated systems, spectral steepening at higher frequencies, as identified by Tisanić et al. (2020), may be more pronounced in more massive hosts, where older electron populations dominate the radio emission. Overall, the RSED serves not only as a diagnostic of the dominant physical processes within galaxies but also as a sensitive tracer of their star formation activity, evolutionary state, and mass build-up across cosmic time.

1.3.2.4 Major Multiwavelength Surveys Enabling Radio SED Studies in the COSMOS Field

The construction of robust RSEDs requires datasets that combine broad frequency coverage with reliable multiwavelength information and accurate redshifts. The COSMOS field is uniquely suited for such analyses, as it brings together some of the deepest and most homogeneous surveys across the radio, optical, near-infrared, and mid-infrared regimes. These complementary datasets ensure precise source identification and allow for consistent measurements of radio properties in a well-characterized extragalactic environment. The following major surveys and data products constitute the foundation for scientific research in the COSMOS field:

- **HST-COSMOS** (Hubble Space Telescope COSMOS Survey, (Scoville, et al., 2007)): Launched in 2003, HST-COSMOS provided high-resolution optical imaging (primarily in the I-band) across the entire 2 deg² COSMOS field. This unprecedented dataset enabled detailed morphological classification, structural measurements of galaxies, and the identification of optical counterparts to sources detected at other wavelengths.
- **Subaru Suprime-Cam COSMOS Survey** (Taniguchi et al., 2007): Using the wide-field Suprime-Cam on the Subaru telescope, this survey delivered deep optical images in multiple bands (B, V, r', i', z') covering the COSMOS field. The high photometric quality and depth of these images are crucial for measuring accurate photometric redshifts, detecting faint galaxies, and studying galaxy colors and stellar populations.
- **Spitzer-COSMOS** (S-COSMOS, Sanders et al., 2007): Conducted with the Spitzer Space Telescope, the S-COSMOS survey provided mid-infrared imaging at several wavelengths (3.6,

4.5, 5.8, 8.0, and 24 μm). These data are essential for investigating dust emission, obscured star formation, and the contribution of AGN in galaxies across the field.

- **zCOSMOS** (Lilly, et al., 2007): Initiated in 2005, zCOSMOS is a large spectroscopic redshift survey targeting galaxies in the COSMOS field using the VLT. It offers highly accurate redshift measurements for tens of thousands of galaxies, enabling reliable distance estimates, studies of galaxy environments, and calibration of photometric redshifts.
- **UltraVISTA Survey** (McCracken et al., 2012): Beginning in 2009, UltraVISTA is an ultra-deep near-infrared imaging survey carried out with VISTA. It delivers exceptionally deep images in the Y, J, H, and Ks bands, which are key for deriving galaxy stellar masses, rest-frame colors, and high-quality photometric redshifts, especially at higher redshifts ($z > 1$).
- **VLA-COSMOS 3 GHz Large Project** (Smolcic, et al., 2017): Completed in the mid-2010s, this survey used the VLA to map the entire COSMOS field at 3 GHz with high sensitivity (rms $\sim 2.3 \mu\text{Jy/beam}$). The resulting radio catalog enables studies of star formation, AGN activity, and radio spectral properties across large and representative galaxy samples.
- **LOFAR Two-Metre Sky Survey** (LoTSS, (Shimwell, et al., 2022)): The LoTSS project provides wide-area, low-frequency (120–168 MHz) radio observations of the northern sky, including the COSMOS field. Its deep radio maps at 144 MHz are vital for detecting synchrotron emission from both AGN and star-forming galaxies, facilitating the study of low-frequency radio properties and source populations. LoTSS data form the primary radio dataset for this thesis, enabling a systematic investigation of faint and extended radio sources in COSMOS. Advanced calibration techniques ensure reliable measurements even in challenging regions, while the exceptional depth and sky coverage of LoTSS are illustrated in Figures 4 and 5

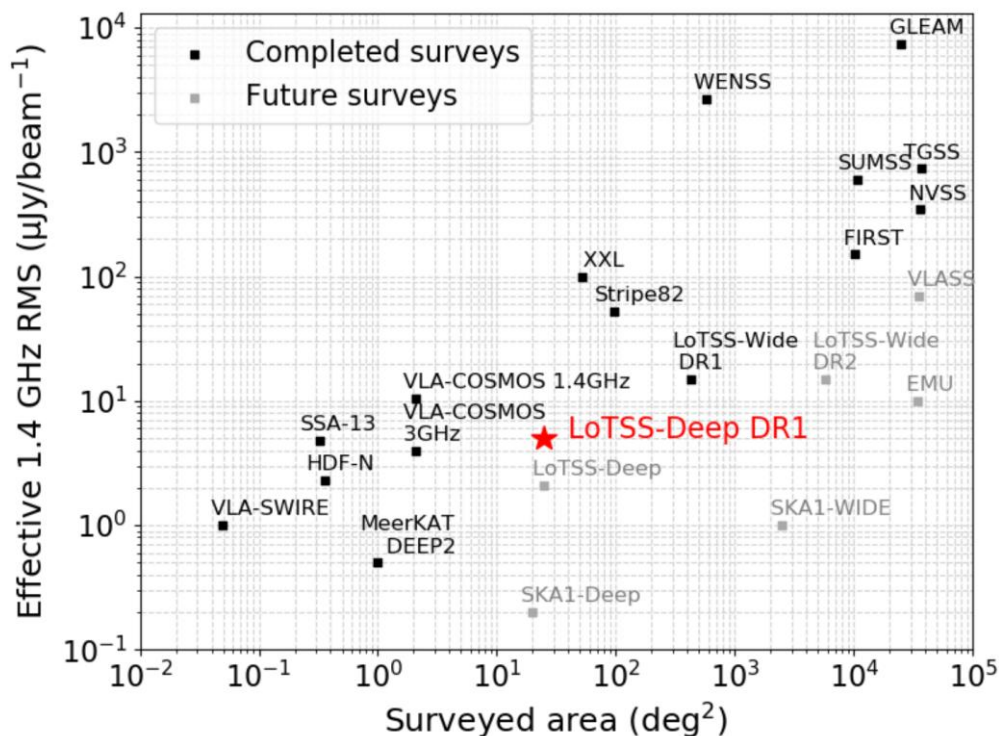


Figure 4: Comparison of the effective noise levels (converted to 1.4 GHz assuming a spectral index of $\alpha = 0.7$) versus surveyed area for existing and upcoming radio continuum surveys. The LoTSS-Deep Fields (red cross) demonstrate a unique combination of high sensitivity (down to $\sim 30 \mu\text{Jy beam}^{-1}$) and wide sky coverage, significantly outperforming earlier low-frequency surveys such as TGSS and WENSS. While most earlier wide-area surveys lacked either sensitivity or angular resolution, LoTSS, with its ~ 6 arcsec resolution and sophisticated direction-dependent calibration (Tasse, et al., 2021), enables robust detection of both compact and extended radio sources. The COSMOS field analyzed in this work lies within the LoTSS-DR2 RA-1 region and benefits from comparable depth, though it is not separately indicated in figure.

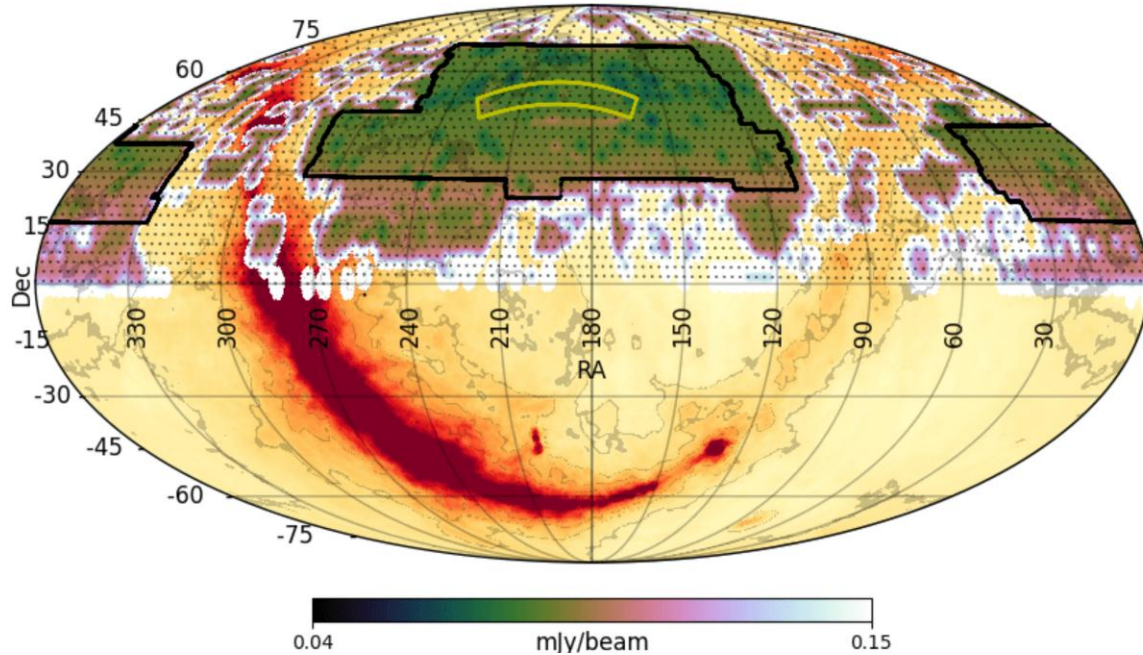


Figure 5: Status of the LOFAR Two-metre Sky Survey (**LoTSS**) observations as of April 2021, showing the approximate current sensitivity coverage overlaid on the Haslam et al. (1982) 408 MHz all-sky map. Colours range from yellow to deep red, representing the brightness temperature at 408 MHz, with contours highlighting Galactic emission. The black and yellow outlines correspond to the LoTSS-DR2 and DR1 footprints, respectively, while the black dots denote the 3168 planned LoTSS grid pointings. At the time of this map, 1623 pointings had been fully observed (including all from DR2), 154 were partially completed, and 1391 remained unobserved. The green region indicates areas with the highest sensitivity (rms ~ 0.04 mJy/beam), achieved through advanced direction-dependent calibration. This figure highlights the rapid expansion and technical maturity of LoTSS, which aims to deliver high-resolution, deep low-frequency radio imaging across the northern sky. (Shimwell, et al., 2022)

Looking ahead, future radio surveys and observatories such as the Square Kilometre Array (SKA) and its precursors (e.g., MeerKAT and ASKAP) are expected to dramatically enhance our ability to study extragalactic radio sources. Their unprecedented sensitivity and broad frequency coverage will open new parameter space for the detection and characterization of faint and distant galaxies, probing galaxy evolution and the interplay of physical processes across cosmic time. The comprehensive multiwavelength legacy established by efforts ranging from HST-COSMOS to LoTSS-DR2 provides an essential foundation for these next-generation studies, ensuring continuity and maximizing scientific return as radio astronomy enters a new era.

1.3.2.5 **Limitations and Uncertainties in the Use of RSEDs**

While the construction and analysis of RSEDs provide valuable insights into the physical processes shaping galaxy evolution, several limitations and uncertainties must be acknowledged. Understanding these caveats is crucial not only for the correct interpretation of current results but also for identifying directions for future improvements. Below, we present the major observational and methodological challenges that must be considered when working with radio datasets:

❖ Sensitivity limitations and upper limits

The sensitivity of current radio surveys, although impressive, imposes detection thresholds that lead to incomplete sampling, especially at high redshifts and low luminosities. The treatment of non-detections through survival analysis partially mitigates this issue, but uncertainties remain, particularly in the low-flux regime.

❖ Spectral curvature and frequency coverage

The accurate characterization of spectral curvature requires observations spanning a wide frequency range. Gaps in frequency coverage, especially at very low (sub-GHz) and high (>10 GHz) rest-frame frequencies, can lead to biases in the derived spectral indices and hinder the detection of spectral features such as turnovers or breaks.

❖ Impact of thermal emission

At frequencies above a few GHz, the contribution of thermal free-free emission becomes increasingly significant. Separating the thermal and non-thermal components of the radio emission often relies on assumptions or models, introducing additional uncertainties into the interpretation of the RSED shape.

❖ Environmental and evolutionary effects

The radio properties of galaxies are influenced by their environment (e.g., cluster versus field galaxies) and their evolutionary state (e.g., age of electron populations). Accounting for these factors requires careful sample selection and large, statistically robust datasets, which are not always available.

❖ Uncertainties in rest-frame frequency estimation

The determination of rest-frame frequencies depends critically on accurate redshift measurements. Although spectroscopic redshifts from surveys like zCOSMOS provide high reliability, for large samples relying on photometric redshifts, systematic uncertainties can propagate into the derived RSEDs.

Despite these limitations, ongoing and upcoming radio surveys—combined with advancements in analysis techniques, such as improved stacking methods and machine learning approaches—promise to significantly refine the construction and interpretation of RSEDs. Upcoming studies utilizing next-generation instruments like the Square Kilometre Array (SKA), will provide broader frequency coverage and deeper sensitivities, enabling a more detailed and physically grounded understanding of the radio emission processes across diverse galaxy populations and cosmic epochs. Therefore, while current RSED studies offer a powerful framework for investigating galaxy evolution, they must be interpreted in light of their observational and methodological constraints, with continuous efforts directed toward their improvement and expansion.

1.3.3. Research Contribution

A significant advance in the study of radio spectral energy distributions (SEDs) for extragalactic sources has been made possible through the integration of new low-frequency observations, particularly those provided by the LOFAR (Low Frequency Array) surveys, into established multifrequency analyses. Until recently, most studies of radio SEDs—especially within the context of the COSMOS field—were based primarily on data from the VLA and GMRT, which offered detailed coverage across the 0.325–3 GHz regime but lacked the angular resolution at the lowest frequencies needed to fully characterize the radio continuum of both AGN and SFGs.

This research builds directly on the methodological framework established by Tisanic et al. (2019, 2020), who demonstrated that the simple power-law model, often adopted for the average radio SED, fails to account for the observed spectral curvature and deviations, particularly in the presence of changes in the spectral slope. Their findings, based on robust statistical approaches such as survival

analysis and binned stacking, established that a BPL model provides a more physically motivated and accurate description of the radio SEDs for both highly star-forming galaxies and radio-excess AGN, capturing the complex interplay between synchrotron, thermal, and environmental processes (see 1.3.2.2).

The present study advances this methodology by extending the spectral coverage of radio SEDs to even lower frequencies, incorporating for the first time the 144 MHz regime. The addition of these data enables the development of a new and more precise fitting model, as the inclusion of points at 144 MHz allows for a more robust determination of the spectral slope at low frequencies. This leads to a more accurate estimate of the low-frequency spectral index, while the high-frequency part of the spectrum remains essentially unchanged. Advanced statistical techniques, including survival analysis, are employed to ensure that upper limits are properly accounted for, resulting in a representative and unbiased characterization of the average SED across the full frequency range.

The broader implication of this approach is that the refined description of the SED at the lowest frequencies provides new insight into the physical mechanisms governing radio emission in galaxies—particularly those affecting the low-frequency spectral slope, such as synchrotron aging, absorption processes, and the interplay of different energy-loss mechanisms. As even lower-frequency observations (for example, at 60 MHz) become available in the future and are integrated into this framework, it is expected that our understanding of radio SEDs and their spectral indices will deepen further, providing a progressively more comprehensive understanding of the physical mechanisms driving galaxy evolution throughout cosmic time.

1.4 Theoretical Framework

1.4.1 COSMOS field

The Cosmic Evolution Survey (COSMOS) field represents a groundbreaking initiative in observational cosmology, focusing on a **2-square-degree** equatorial region in the constellation Sextans (centered at Right Ascension **10h 00m 28.6s** and DEclination **+02° 12' 21"** (J2000)). This region was carefully selected for its lack of interference from nearby stars or gas clouds, allowing unobstructed views of deep space (Scoville, et al., 2007). Covering an area approximately 16 times larger than the full moon, the COSMOS field enables the study of galaxies, AGN, dark matter, and large-scale structures over a vast range of redshifts. The field's location close to the celestial equator (Dec +2°) enables access from both northern and southern hemisphere observatories. However, from higher northern latitudes, COSMOS never rises more than ~30 degrees above the horizon and is not visible year-round, which can limit observing conditions during certain months. By capturing light across multiple wavelengths—from X-rays to radio waves—this survey provides an unparalleled dataset, essential for exploring how the fundamental properties and environments of galaxies have evolved over 75% of the universe's history (Darvish, et al., 2016). This region is illustrated in Figure 6, which shows the COSMOS Hubble ACS Field.

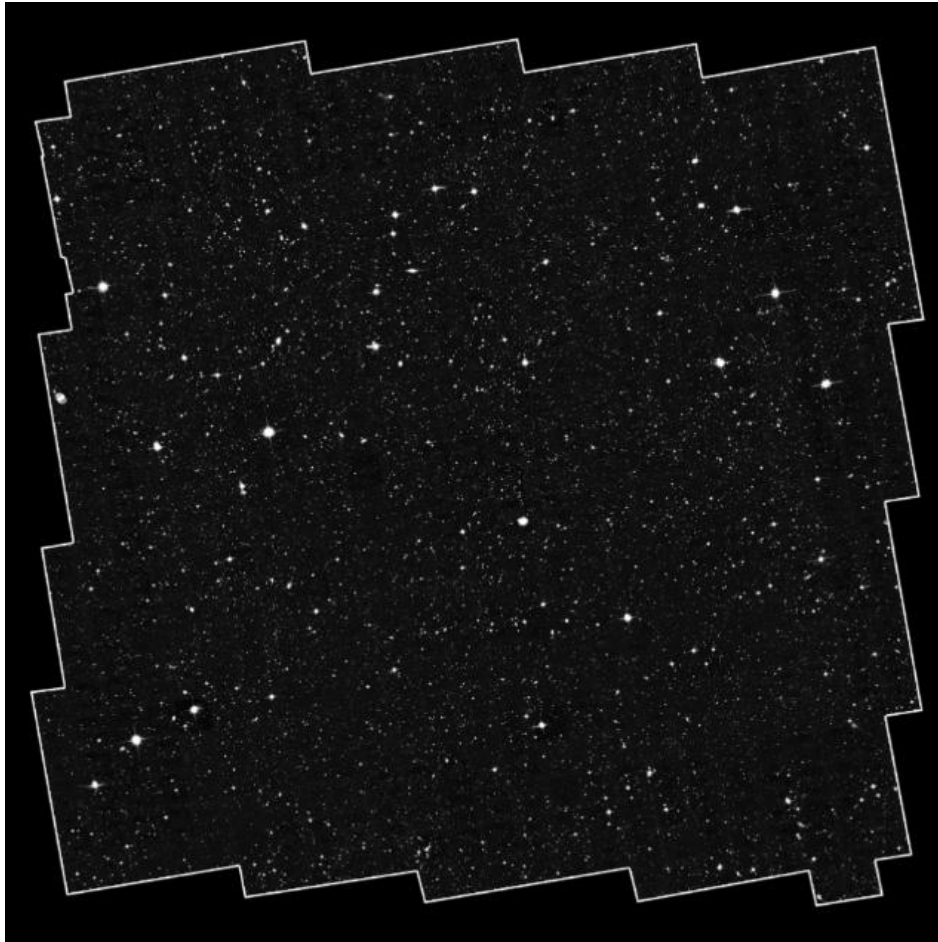


Figure 6: The COSMOS Hubble ACS Field (NASA, ESA, Anton M. Koekemoer (STScI), Nick Scoville (Caltech))

The COSMOS project was initiated to tackle several of the most significant challenges in cosmology, such as the interplay between galaxy formation, AGN activity, and large-scale structures (Scoville, et al., 2007). The survey area was strategically chosen to maximize accessibility for ground-based and space-based observatories worldwide. With contributions from over 100 scientists and major observatories—including the Hubble Space Telescope, the Subaru Telescope, and the VLA radio telescope—the project combines high-resolution imaging with deep spectroscopic and multiwavelength datasets (Tisanic, et al., 2020) (Tisanic, et al., 2019). The depth and breadth of the COSMOS field make it uniquely suited for reducing cosmic variance and detecting faint objects at high redshifts, enabling researchers to probe the physical and environmental factors shaping galaxies and their evolution (Delvecchio, et al., 2017).

Among the numerous sub-projects within COSMOS, the VLA-COSMOS 3 GHz Large Project plays a central role in tracing radio continuum emission at high resolution, making it instrumental for this study. Spanning 384 hours of observations, this project delivers deep radio maps with unprecedented sensitivity, enabling the detection of faint AGN and star-forming galaxies up to redshifts of $z \sim 5$ (Smolcic, et al., 2017). By combining this data with observations at other wavelengths, the VLA-COSMOS project offers critical insights into the distribution and evolution of radio-emitting sources. It also facilitates the distinction between AGN and star-forming galaxies, contributing to the broader goals of understanding galaxy assembly and star formation in diverse cosmic environments (Delvecchio, et al., 2017). Figure 7 presents two examples of radio galaxies observed in the COSMOS field, each exhibiting extended lobes and prominent core emission.

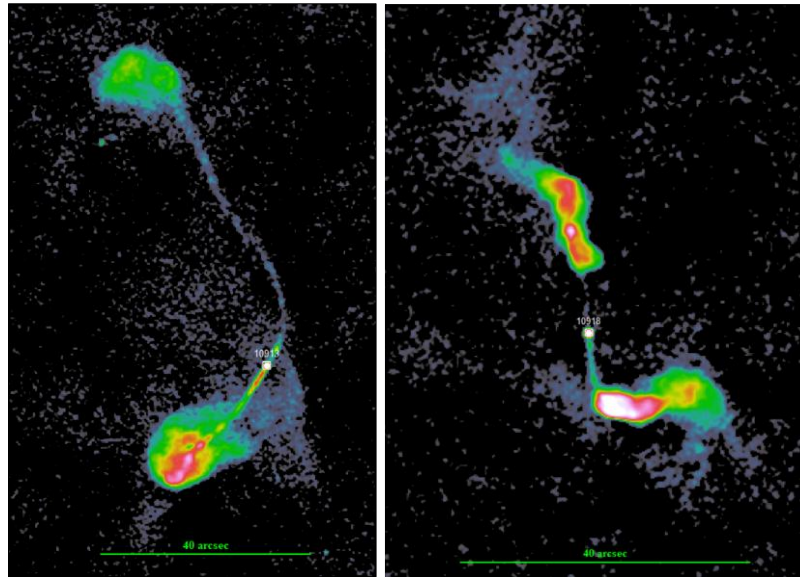


Figure 7: Postage-stamp cutouts from the VLA-COSMOS 3 GHz continuum mosaic, imaged using the MSMF (Multi-Scale Multi-Frequency) algorithm. The images showcase two examples of radio galaxies with extended lobes and core emission (left ID_VLA:10913 right ID_VLA:10918). The color scale represents the radio flux density in Jy/beam, highlighting the morphology and intensity structure of both compact and extended radio sources within the COSMOS field.

This equatorial region hosts a staggering diversity of cosmic objects, with over 2 million galaxies identified, spanning redshifts from $z \sim 0.1$ to $z \sim 6$. Among them are approximately 2000 X-ray-detected AGN, 3000 ULIRGs (UltraLuminous InfraRed Galaxies), and tens of thousands of star-forming galaxies (Darvish, et al., 2016). The large-scale structures in the field include massive dark matter halos and galaxy clusters, with masses reaching up to $\sim 10^{14} M_{\odot}$ (Toni, Maturi, Finoguenov, Moscardini, & Castignani, 2024). Observations also reveal complex environments where merging and interactions trigger intense star formation or AGN activity (Scoville, et al., 2007).

At radio wavelengths, synchrotron emission dominates, originating from relativistic electrons spiraling in magnetic fields, a hallmark of AGN jets and supernova remnants (Tisanic, et al., 2020) (Tisanic, et al., 2019). Thermal free-free emission contributes at higher frequencies, tracing ionized gas in star-forming regions. The sensitivity of $\sim 2.3 \mu\text{Jy}$ at 3 GHz allows the detection of faint sources, providing insights into starburst activity and AGN fueling mechanisms (Delhaize, et al., 2017). Combined gravitational lensing studies map the distribution of dark matter, showing the interplay between visible and invisible components across cosmic scales (Darvish, et al., 2016).

These observations and their physical implications make this dataset a cornerstone for studying galaxy evolution, large-scale structure, and the co-evolution of dark matter, AGN, and star formation through cosmic time. The region's depth and coverage provide an unparalleled laboratory for testing models of structure formation and feedback processes in the universe (Vardoulaki, et al., 2021).

1.4.2 Physical Models of Radio Emission

The study of radio emission from astrophysical sources provides crucial insights into the physical processes governing galaxies and their environments. Two key mechanisms—synchrotron radiation and free-free emission—dominate the radio regime, each tracing different components of the interstellar and intergalactic medium (ISM, IGM). Relativistic electrons spiraling within magnetic fields generate synchrotron radiation, while free-free emission is emitted by ionized gas present in regions with active star formation. In the following sections, we explore the fundamental physics,

spectral characteristics, and astrophysical applications of these mechanisms, highlighting their role in understanding galaxy evolution and the dynamics of cosmic structures.

1.4.2.1 Synchrotron Radiation

Synchrotron radiation is one of the most significant mechanisms of radio emission in astrophysical sources, such as AGN and SFGs. This process arises from the interaction of relativistic electrons, which move at velocities close to the speed of light, with magnetic fields. As the electrons spiral around the magnetic field lines, they accelerate, emitting electromagnetic radiation that spans a broad spectrum. This mechanism is pivotal for understanding the physical conditions within galaxies, such as magnetic field strengths and the energy distribution of charged particles (Rybicki & Lightman, 1979) emitted radiation is highly polarized and anisotropic, concentrated in a narrow cone along the direction of motion of the relativistic electron due to relativistic beaming effects. The *power radiated per electron* is proportional to both the magnetic field strength B and the energy of the particle. This relationship is described by the equation:

$$P = \frac{2}{3} \frac{e^4 B^2 \sin^2 \theta}{m_e^2 c^3} \quad (1.2)$$

where e refers to the charge of the electron, m_e to its mass, c to the speed of light, and θ is the angle between the electron's velocity vector and the direction of the magnetic field.

The spectrum of synchrotron radiation arises from a population of relativistic electrons with a range of energy. If the energy distribution of these electrons follows a power law,

$$N(E) \propto E^{-p} \quad (1.3)$$

where E is the energy and p are the power-law index, the resulting flux density of the emitted radiation at a given frequency ν is also a power law:

$$S_\nu \propto \nu^{-\alpha} \quad (1.4), \alpha = \frac{p-1}{2} \quad (1.5).$$

Here, α is the spectral index, which characterizes the slope of the emission spectrum. Typical values of α for synchrotron radiation in astrophysical contexts range from 0.5 to 1.0, depending on the physical conditions of the emitting region (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002).

At lower frequencies, synchrotron self-absorption can occur. This phenomenon arises when the medium becomes optically thick to its own radiation, causing the spectrum to flatten below a critical frequency ν_a . In this regime, the flux density follows:

$$S_\nu \propto \nu^{5/2} \quad (1.6).$$

For higher frequencies, where the medium becomes optically thin ($\nu > \nu_a$), the flux density returns to the power-law behavior, $S_\nu \propto \nu^{-\alpha}$ (Rybicki & Lightman, 1979).

In galactic environments, deviations from a simple power-law synchrotron spectrum are commonly observed. These arise due to energy losses of relativistic electrons (so-called "synchrotron aging") and the temporal evolution of the supernova rate, which injects cosmic ray electrons into the ISM. The resulting spectral energy distribution can exhibit a "break" or a curvature, with the slope of the spectrum steepening at higher frequencies. Such features are particularly informative for

constraining the age and energy distribution of the electron population, as well as the strength and topology of the galactic magnetic field.

The synchrotron radiation mechanism provides invaluable insights into the physical processes of galaxies and other cosmic environments. The spectral steepening observed at high frequencies is a direct consequence of synchrotron cooling, as relativistic electrons lose energy over time due to radiation. This steepening, characterized by a deviation from a simple power-law spectrum, offers a diagnostic tool for understanding the age and energy distribution of the electron population. For instance, extended regions such as jets and lobes in AGN show steep spectra due to aging electrons, while compact core-dominated sources exhibit flatter spectra (Blundell & Rawlings, *The Spectra and Energies of Classical Double Radio Lobes*, 2000).

Furthermore, the spectral index serves as a powerful parameter for probing the magnetic field strengths and electron energy distributions in star-forming galaxies and AGN. Combined with multiwavelength observations, synchrotron radiation allows researchers to disentangle the contributions of thermal and non-thermal processes, offering a comprehensive view of galaxy evolution across cosmic time (Murphy E. J., et al., 2018). This makes synchrotron radiation an essential tracer in radio surveys such as the VLA-COSMOS 3 GHz Project, where distinguishing between AGN-related and star-formation-related emission is critical.

Jets and Lobes in AGN

Synchrotron radiation is a dominant mechanism in the extended jets and lobes of AGN, which transport energy from the central SMBH to distances far beyond the host galaxy. Observations across a wide frequency range, from gigahertz (VLA) to megahertz (LOFAR), have enabled detailed studies of the morphology and spectral evolution of AGN jets and lobes, revealing the spectral properties of the emission.

For instance, in radio galaxies such as Cygnus A and Hercules A, synchrotron radiation from relativistic jets and lobes display steep spectral indices ($\alpha \sim 1$), indicative of electron cooling through synchrotron losses as they propagate outward (Blundell & Rawlings, *The Spectra and Energies of Classical Double Radio Lobes*, 2000). Moreover, these observations show spectral steepening at higher frequencies due to the aging of relativistic electrons, a feature that constrains the energy distribution of the electron population and the magnetic field strength (Rybicki & Lightman, 1979). In contrast to the extended lobes, the compact cores of AGN typically exhibit flatter or even inverted spectra, associated with self-absorbed synchrotron emission from the innermost jet regions.

High-resolution radio studies of AGN lobes often reveal morphological and spectral **asymmetries** in jets and lobes, which often reflect interactions with the intergalactic medium. These observations provide critical information about the feedback processes that AGN exert on their surrounding environment, regulating galaxy evolution (Santoro, Tadhunter, Baron, Morganti, & Holt, 2020).

Magnetic Fields in Star-Forming Regions

Synchrotron radiation also serves as an invaluable tool for probing the magnetic fields in star-forming regions. In galaxies such as M82 (see Figure 8), which exhibit intense starburst activity, radio observations reveal the synchrotron emission from relativistic electrons accelerated in supernova remnants. Observations with the VLA at GHz frequencies have shown strong correlations between the radio and infrared emission, providing evidence that synchrotron radiation traces the supernova rate, and by extension, the star formation rate (Sargent, et al., 2010).

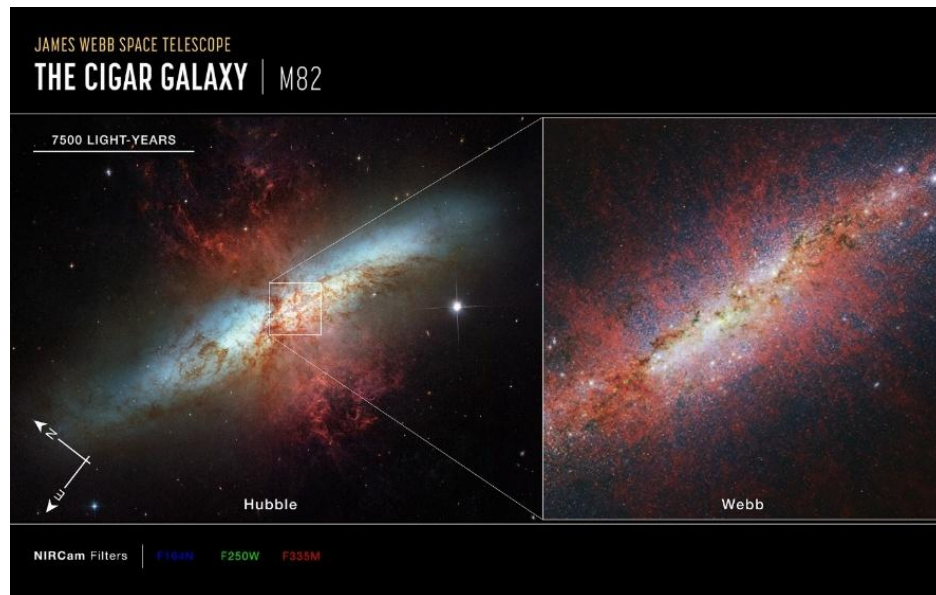


Figure 8: Multi-wavelength view of the starburst galaxy M82 (the “Cigar Galaxy”), combining Hubble (left) and James Webb (right) observations. M82 is a prototypical starburst galaxy, hosting intense star-forming activity, supernova explosions, and strong magnetic fields. JWST reveals intricate structures of ionized gas, dust lanes, and embedded star-forming regions. In such galaxies, synchrotron radiation from relativistic electrons traces magnetic field structures, while thermal free-free emission highlights regions of massive star formation. (NASA, ESA, CSA, STScI, Alberto Bolatto (UMD))

Additionally, low-frequency observations with LOFAR have been used to measure the spectral curvature caused by synchrotron self-absorption in dense star-forming regions. This spectral curvature helps estimate the magnetic field strength, as well as the energy density of cosmic rays, shedding light on the physical conditions that regulate star formation, including the influence of magnetic pressure on gas collapse and the confinement of cosmic rays.

In addition to synchrotron processes, thermal free-free emission also arises in star-forming regions, offering complementary diagnostics of ionized gas and recent massive star formation.

1.4.2.2 Free-free emission

Free-free emission, also known as thermal Bremsstrahlung, is a fundamental mechanism of electromagnetic radiation in ionized astrophysical plasmas. It results from the deflection of free electrons by ions in plasma, leading to the production of continuous electromagnetic radiation. This emission mechanism plays a crucial role in astrophysical environments dominated by ionized gas, such as HII regions, supernova remnants, and galaxy clusters. The process is especially important in the radio regime, where its spectral characteristics provide unique insights into the physical conditions of astrophysical plasmas (Rybicki & Lightman, 1979).

Physical Principles

At the core of free-free emission lies the Coulomb interaction between free electrons and ions. When an electron passes near an ion, it experiences the Coulomb force, which alters its velocity and trajectory. This deceleration, or “braking,” causes the electron to emit radiation. Unlike synchrotron radiation, where the emission is directional, free-free radiation is isotropic, with its intensity dependent on the temperature and density of the plasma.

The total power emitted by a single electron during a close encounter with an ion is proportional to the square of the ion's charge (Z^2) and inversely proportional to the electron's mass (m_e). This dependence on Z^2 explains why regions with highly ionized gases, such as those containing heavier

elements, produce stronger free-free emissions. For a plasma with an electron temperature T , the emissivity (power per unit volume) of free-free radiation at frequency ν is given by:

$$\varepsilon_\nu \propto n_e n_i z^2 T^{-1/2} e^{\frac{-h\nu}{kT}} g_\nu \quad (1.7)$$

where:

- n_e and n_i are the number densities of electrons and ions,
- T is the electron temperature,
- g_ν is the Gaunt factor, a quantum mechanical correction accounting for the interaction's details at different frequencies,
- h is Planck's constant, and
- k is Boltzmann's constant.

The exponential term $e^{\frac{-h\nu}{kT}}$ indicates that the emission decreases rapidly at frequencies much higher than the thermal energy scale (kT), restricting the strongest free-free emission to radio and submillimeter wavelengths. This property is particularly significant for studies of star-forming galaxies, as it allows the use of free-free radiation to trace ionized gas without interference from dust attenuation (Kennicutt & Evans, 2012). This makes free-free emission a valuable complement to synchrotron diagnostics, particularly in identifying compact HII regions and quantifying recent massive star formation in galaxies observed at GHz frequencies.

Spectral Characteristics and Dependence on Physical Conditions

The spectrum of free-free emission is distinctively flat in the radio regime, with the flux density scaling approximately as:

$$S_\nu \propto \nu^{-0.1} \quad (1.8) .$$

This weak frequency dependence makes it easily distinguishable from the steep spectra of synchrotron radiation. The nearly flat spectrum arises because the emission at lower frequencies is offset by increased optical depth due to free-free absorption, while higher-frequency emission is limited by the thermal energy of the electrons.

At low frequencies ($\nu < 1$ GHz), free-free absorption can dominate in dense regions, resulting in a turnover in the spectrum. This is particularly important for compact HII regions or dense starburst galaxies, where the high density of ionized gas can cause the optical depth to exceed unity. By analyzing this turnover, it is possible to estimate the electron density and the emission measure ($EM = n_e^2 L$, where L is the path length of the ionized gas).

Free-free emission also depends on the plasma temperature, scaling inversely with $T^{1/2}$. This makes it a useful probe for estimating the temperature of ionized regions in galaxies and clusters. For example, in galaxy clusters, free-free emission contributes significantly to the X-ray spectrum, while in HII regions, it provides a critical tool for measuring the star formation rate (Rybicki & Lightman, 1979).

Relevance to Mixed Sources

In astrophysical systems like star-forming galaxies, free-free emission is rarely isolated. It often coexists with non-thermal synchrotron radiation, particularly in regions where both massive star formation and supernova activity occur. For example, in galaxies such as M82, free-free emission dominates at higher radio frequencies (> 100 GHz), while synchrotron radiation contributes significantly

at lower frequencies (<100 GHz). This division arises because free-free emission has a flatter spectrum compared to the steep power-law decline of synchrotron radiation.

To model such mixed sources, broken power-law models are commonly employed. These models describe the observed spectrum as the sum of thermal and non-thermal components:

$$S_\nu = S_{\nu,thermal} + S_{\nu,non-thermal} \quad (1.9),$$

where $S_{\nu,thermal} \propto \nu^{-0.1}$ and $S_{\nu,non-thermal} \propto \nu^{-\alpha}$, with $\alpha \sim 0.7$ for typical synchrotron radiation (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002). By fitting such models to multi-frequency data, it becomes possible to disentangle the contributions of free-free and synchrotron emission, providing a detailed view of the physical processes occurring in these environments. This is particularly crucial when analyzing galaxies hosting both AGN and star formation activity, where a single-component model may lead to biased estimates of SFR or AGN luminosity.

Astrophysical Applications

The study of ionized gas in SFGs and active galactic nuclei relies heavily on the detection of thermal radiation mechanisms like free-free emission. This process serves as a vital dust-independent tracer of SFRs, with ionized hydrogen in HII regions radiating strongly at radio wavelengths. By isolating the thermal emission, astronomers can disentangle it from non-thermal contributions, enabling a clearer characterization of galaxies with mixed emission sources. Observations conducted with instruments such as the VLA and LOFAR have revealed the essential role of free-free emission in starburst galaxies, linking it directly to massive star formation, supernova activity, and the dynamics of ionized gas.

Free-free emission has been extensively studied in the COSMOS field using high-resolution radio observations, enabling the characterization of ionized gas in both SFGs and regions hosting AGN. For example, the GMRT project at 610MHz & 325MHz, which covers the 2-square-degree COSMOS field, has provided unparalleled sensitivity ($\sim 39 \mu\text{Jy/beam}$ and $\sim 97 \mu\text{Jy/beam}$) to detect faint thermal emission from ionized regions in galaxies. These observations have revealed the importance of free-free emission as a dust-independent tracer of SFRs, particularly in galaxies heavily obscured by dust, where traditional tracers like H α or UV emission fail.

One remarkable finding in the COSMOS field involves the study of the infrared-radio correlation (IRRC) in star-forming galaxies. Observations combining deep radio data from the VLA with infrared data from the Spitzer and Herschel space telescopes have demonstrated that free-free emission contributes significantly to the radio flux at higher frequencies (>10GHz), particularly in compact starburst galaxies. This makes high-frequency observations with next-generation instruments like the ngVLA critical for refining thermal SFR diagnostics and extending studies of dust-obscured star formation to higher redshifts. By isolating the free-free component, researchers were able to refine SFR calibrations, even in galaxies where synchrotron radiation dominates at lower frequencies (Delhaize, et al., 2017). These findings highlight the importance of multi-frequency observations in disentangling the contributions of thermal and non-thermal processes to the radio spectrum.

The COSMOS field has also provided critical insights into the connection between free-free emission and large-scale structures. Using spectroscopic redshift data from the zCOSMOS project, researchers have mapped the distribution of star-forming galaxies within cosmic filaments and clusters, demonstrating that regions with higher galaxy densities exhibit enhanced free-free emission due to increased ionized gas content. This phenomenon is particularly evident in galaxy groups at redshifts $z \sim 1$, where feedback from AGN and mergers trigger bursts of star formation, significantly boosting the free-free emission detected at radio wavelengths (Darvish, et al., 2017). The ability to trace these interactions in the COSMOS field underscores the unique value of combining high-resolution radio observations with spectroscopic surveys.

1.4.3 Active Galactic Nuclei (AGN)

Active Galactic Nuclei (AGN) represent one of the most fascinating and critical components in the study of cosmic evolution. These energetic regions at the centers of galaxies are powered by accretion onto **SMBH**, which can release immense amounts of energy across the electromagnetic spectrum. Depending on their dominant emission mechanisms, AGN are broadly divided into two classes: radiative-mode AGN, powered by efficient accretion and emitting strongly across the spectrum, and jet-mode AGN, dominated by mechanical feedback through powerful radio jets (Heckman & Best, 2014; Tisanic et al., 2020). They are crucial to galaxy evolution, since their energetic feedback controls star formation and influences the characteristics of galaxies and their surroundings over cosmic time (Kormendy & Ho, 2013); (Heckman & Best, 2014)). Furthermore, AGN act as probes of large-scale structure and cosmic reionization, providing insights into the interplay between black hole growth and the formation of galaxies.

The radio emission from AGN primarily arises through synchrotron radiation, a process in which relativistic electrons spiral in the presence of magnetic fields. In the radio regime, AGN are further classified according to their large-scale radio morphology into Fanaroff-Riley type I (FR I) and type II (FR II) sources (Fanaroff & Riley, 1974). FR I sources show brightest emission close to the core, while FR II sources are characterized by prominent hotspots at the edges of their radio lobes, reflecting different jet powers and environmental interactions. This emission is characteristic of AGN jets and lobes, which transport energy from the central engine to the outer regions of the host galaxy and beyond. Synchrotron radiation produces a broad, continuous spectrum often described by a power-law function, where the spectral index depends on the distribution of electron energies. At lower frequencies, synchrotron self-absorption becomes significant, causing a flattening or even inversion of the spectrum. This occurs when the medium becomes optically thick, as photons are reabsorbed by the same population of relativistic electrons that emitted them (Rybicki & Lightman, 1979). This is illustrated in Figure 9, which shows a radio image of the iconic FR II galaxy Cygnus A, with its prominent jets and lobes.

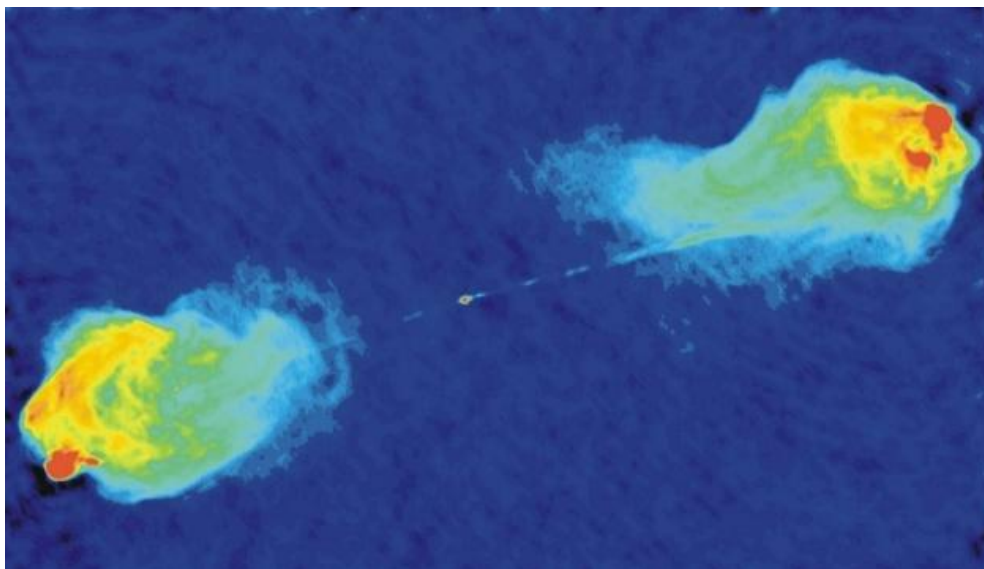


Figure 9: Radio galaxy Cygnus A. (Image courtesy of NRAO/AUI; R. Perley, C. Carilli & J. Dreher)

While simple power-law models are often used to describe AGN radio emissions, deviations from this behavior can provide crucial insights into the physical condition of the source. At higher frequencies, spectral steepening is commonly observed, attributed to synchrotron cooling as electrons lose energy over time. Such deviations are influenced by factors including the size of the emitting region, magnetic field strength, and the electron energy distribution. Additionally, the redshift of the

source plays a role, as it shifts the observed spectrum and alters the apparent steepening or flattening. For compact sources, such as those dominated by core emission, self-absorption effects are more pronounced, while extended sources, like jets and lobes, exhibit steeper spectra due to aging electrons and energy losses (Blundell & Rawlings, Spectral ageing: a new age perspective, 2002); (Vardoulaki, et al., 2021)).

The study of these spectral features provides a powerful diagnostic for understanding the evolution and physical properties of AGN. By comparing observed spectral indices across redshifts and environments, researchers can infer changes in accretion modes, jet activity, and feedback processes. Furthermore, disentangling the contributions of AGN from star-forming processes in galaxies is essential for accurately characterizing their impact on cosmic star formation. Particularly since AGN can contaminate traditional SFR tracers such as infrared or UV luminosity.

The COSMOS field plays a pivotal role in advancing our understanding of AGN, providing an extensive dataset for studying their properties and evolution in diverse. This comprehensive coverage enables the detailed characterization of AGN and their host galaxies, offering insights into their physical properties, such as black hole accretion rates, stellar masses, and morphologies, as well as their large-scale environments (Scoville, et al., 2007); (Koekemoer, et al., 2007)). By combining radio observations data, researchers can disentangle emission mechanisms, identify obscured AGN missed at shorter wavelengths, and study the infrared-radio correlation to distinguish AGN from star-forming galaxies. Additionally, the COSMOS field's extensive spectroscopic follow-up (e.g., zCOSMOS) provides precise redshifts and emission-line diagnostics, further enabling the characterization of AGN accretion properties and feedback processes.

The large-scale structures within the COSMOS field, such as galaxy clusters and dark matter halos, provide critical context for understanding how AGN are influenced by and contribute to their environments. For example, AGN in dense environments may exhibit enhanced jet activity or altered accretion modes due to interactions within clusters or mergers, phenomena that are well-sampled in the COSMOS field's extensive catalog of over 2000 X-ray-detected AGN and numerous radio-detected sources. The combination of spatially resolved imaging and statistical samples in COSMOS enables studies of the co-evolution of AGN and galaxies in unprecedented detail, solidifying its role as a cornerstone for understanding the interplay between black hole activity, star formation, and large-scale structure formation over the history of the universe.

1.4.4 Star-Forming Galaxies (SFGs)

Star-forming galaxies (SFGs) are vital for understanding the mechanisms governing the formation and evolution of galaxies across cosmic time. These systems are defined by their active star formation, which is driven by the availability of cold molecular gas that serves as the raw material for stellar birth. SFGs span a range of environments, from quiescent disks in the local universe to intensely starbursting systems at high redshifts. They function as natural laboratories for studying processes such as the conversion of gas into stars, the effects of feedback from massive stars and supernovae, and the regulation of star formation by environmental factors. Their study is indispensable for piecing together the star formation history of the universe and understanding how galaxies assemble their stellar mass over billions of years (Kennicutt & Evans, 2012).

A key component of radio emission in SFGs is synchrotron radiation, which originates from relativistic electrons accelerated in the magnetic fields of supernova remnants. When these electrons spiral in magnetic fields, they emit non-thermal radiation that dominates the low-frequency part of the radio spectrum (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002). The spectral index of this emission, typically around 0.7, reflects the energy distribution of the

relativistic electrons and the magnetic field strength in the galaxy. Synchrotron radiation serves as an indirect tracer of the supernova rate, which is directly linked to the formation of massive stars (Condon, Radio Emission from Normal Galaxies, 1992). This makes it a critical diagnostic for understanding the end stages of stellar evolution and the interaction of supernovae with the ISM (Murphy E. , et al., 2011).

At higher frequencies, the radio spectrum of SFGs is increasingly influenced by thermal free-free emission. This emission arises from the ionized gas in HII regions surrounding young, massive stars (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002). Unlike synchrotron radiation, which exhibits a steep spectrum, free-free emission has a flat spectral slope ($\alpha \approx 0.1$), making it distinguishable in the RSED. The strength of free-free emission is directly proportional to the number of ionizing photons produced by young stars, making it an important tool for estimating the SFR in galaxies (Vardoulaki, et al., 2021). As highlighted by Tisanić et al. (2019), **this separation of synchrotron and free-free components in the radio SED enables more accurate SFR determinations and provides valuable diagnostics of the physical conditions within the ISM of SFGs**. This makes free-free emission a reliable tracer of massive star formation in dense, dust-enshrouded environments, such as compact starburst galaxies and submillimeter galaxies, where traditional optical or IR tracers often fail (Murphy E. , et al., 2011).

The SFR in SFGs is intricately linked to their radio emission, particularly through the infrared-radio correlation (IRRC). This empirical relationship (eq 1.10) connects the thermal infrared emission, which traces dust heated by young stars, with the non-thermal synchrotron radiation in the radio (Delhaize, et al., 2017). The correlation is remarkably tight across a wide range of redshifts and galaxy types, underscoring the physical connection between star formation and radio emission processes (Smolcic, et al., 2017). The IRRC provides a robust method for estimating SFRs in galaxies, particularly at high redshifts, where traditional tracers like H α or UV emission are challenging to measure due to dust attenuation (Magnelli, et al., 2015). Recent studies indicate a decreasing IR-to-radio luminosity ratio (q_{TIR}) with increasing redshift, which may reflect enhanced magnetic field strengths or changes in cosmic ray propagation at earlier epochs (Delhaize, et al., 2017).

The total infrared-to-radio luminosity ratio q_{TIR} , is commonly defined as

The radio spectral energy distribution of SFGs reflects the balance between synchrotron and free-free emission. Deviations from the standard IRRC or variations in the spectral index of synchrotron emission can reveal the physical conditions within a galaxy. For instance, galaxies with intense star formation or interactions may exhibit flatter synchrotron spectra due to increased cosmic ray production or enhanced magnetic fields (Condon, Cotton, & Broderick, Radio Sources and Star Formation in the Local Universe, 2002). Similarly, an excess of free-free emission may indicate regions of concentrated massive star formation (Murphy E. , et al., 2011). By analyzing the detailed radio SED of SFGs, researchers can probe the processes driving star formation, the effects of stellar feedback, and the interplay between gas, stars, and magnetic fields in these systems (Delhaize, et al., 2017).

The COSMOS field provides a unique opportunity to investigate SFGs and their development, due to its comprehensive multiwavelength data, ranging from ultraviolet to radio, covering two square degrees (Scoville, et al., 2007). This breadth of data enables researchers to investigate the infrared-radio correlation (IRRC) and accurately derive star formation rates (SFRs) in galaxies, even at high redshift and in dusty environments (Delhaize, et al., 2017); (Magnelli, et al., 2015)). By combining radio and infrared observations, COSMOS allows for detailed separation of synchrotron and free-free emission, offering valuable insights into the physical conditions within star-forming regions (Condon, Radio Emission from Normal Galaxies, 1992). Large-scale spectroscopic campaigns, such as zCOSMOS, further enrich the dataset with precise redshift measurements, facilitating three-dimensional mapping of galaxy environments and the study of how cosmic structure influences star formation activity (Lilly, et al., 2007); (Darvish, et al., 2017)).

The combination of radio, infrared, optical, and spectroscopic data in COSMOS supports high-precision SFR measurements while providing the means to probe feedback processes, gas accretion, and environmental effects that regulate galaxy growth over billions of years (Delhaize, et al., 2017). This depth has led to the identification of rare SFG populations, such as extreme starbursts or merging systems, which often show excess radio emission—highlighting the impact of energetic feedback (Murphy E. , et al., 2011); (Smolcic, et al., 2017)). Overall, this field enables researchers to study SFGs both as individual galaxies and within the context of larger cosmic structures, significantly advancing our understanding of the mechanisms shaping star formation and galaxy assembly across cosmic time (Scoville, et al., 2007); (Smolcic, et al., 2017)).

In summary, the COSMOS field serves as a crucial resource for studying star-forming galaxies, offering a comprehensive, multiwavelength dataset that enables precise characterization of their physical properties and star formation processes. The field's combination of radio, infrared, and spectroscopic data provides a unique opportunity to probe the interplay between synchrotron and free-free emission, trace the infrared-radio correlation across cosmic time, and investigate the influence of large-scale structures on star formation. Through these observations, advanced our understanding of how star-forming galaxies evolve and interact with their environments, offering critical insights into the processes shaping galaxy assembly over billions of years (Scoville, et al., 2007); (Smolcic, et al., 2017)).

1.4.5 Giant Radio Galaxies (GRGs)

Giant Radio Galaxies (GRGs) are extragalactic radio sources with projected linear sizes exceeding 0.7–1 Mpc, which places them at the extreme end of the radio galaxy size distribution (Machalski, Jamroz, & Zola, 2001); (Delhaize, et al., 2021). GRGs are powered by central AGN that launch relativistic jets, producing large-scale radio lobes through the continuous injection of synchrotron-emitting particles into the intergalactic medium (IGM). The large sizes of GRGs are typically attributed to a combination of long-lived jet activity and evolution in relatively low-density environments, allowing radio structures to propagate with minimal disruption over timescales of several tens to hundreds of Myr.

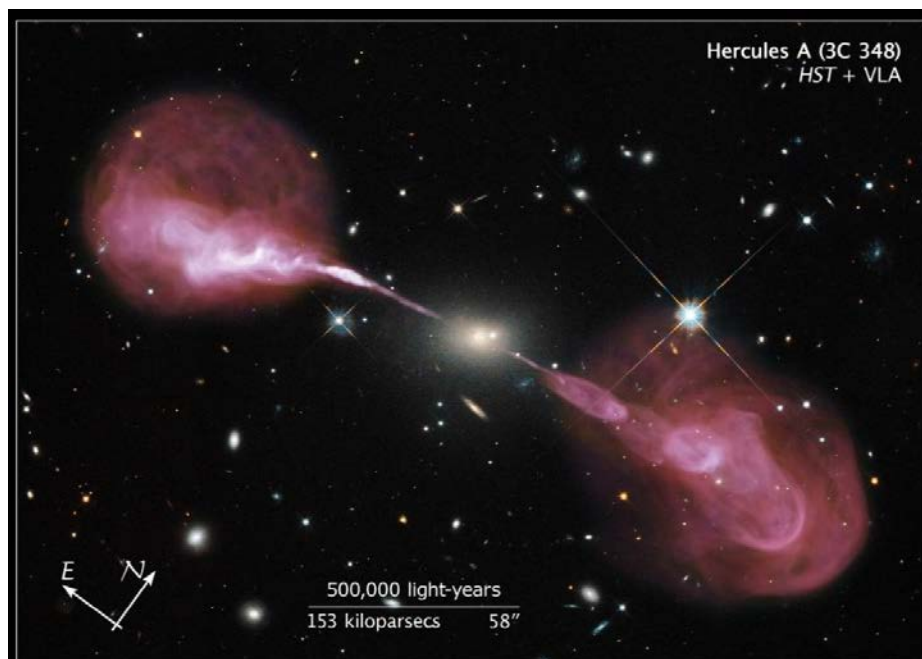


Figure 10: Hubble Space Telescope view of the giant elliptical galaxy Hercules A (3C 348). The pink regions show radio jets, revealed by the Very Large Array (VLA), protruding from the galaxy's center and extending over 1.5 million light-years.

*The scale bar represents 500,000 light-years (153 kpc), and the compass arrow shows the orientation of north and east.
(Image credit: NASA, ESA, S. Baum and C. O'Dea [RIT], R. Perley [NRAO], and the Hubble Heritage Team [STScI/AURA])*

Morphologically, most GRGs exhibit double-lobed (FRII or hybrid FRI/II) radio structures, as illustrated in *Figure 10* by the example of **Hercules A (3C 348)**, where the radio emission (in pink, VLA) extends well beyond the host elliptical galaxy (HST optical). The lobes in such systems are characterized by steep-spectrum synchrotron emission and often display hotspots at their outer edges, although in several cases, surface brightness dimming or environmental asymmetries may cause deviations from this idealized structure.

The host galaxies of GRGs are almost exclusively massive ellipticals, typically with low rates of ongoing star formation and weak or undetectable radiative AGN signatures in the optical, infrared, or X-ray regimes. The radio jets are generally associated with low-excitation (LERG) AGN, consistent with low-efficiency accretion from hot gas in the surrounding medium (Machalski, Jamrozy, & Zola, 2001). The observed radio spectra of GRGs are generally steep ($\alpha \lesssim 0.7$), reflecting aging electron populations and energy losses over their extended lifetimes.

GRGs are rare in current samples, with several hundred to $\sim 1,000$ known examples, depending on selection criteria. Their rarity is primarily a result of observational biases, as the lobes of older, larger sources exhibit low surface brightness and are easily missed in surveys lacking adequate sensitivity to diffuse emission, particularly at frequencies above 1 GHz. Recent deep, low-frequency surveys, such as those conducted with LOFAR and MeerKAT, have revealed new GRG candidates, suggesting that their space density may be substantially higher than previously thought (Delhaize, et al., 2021).

The study of GRGs is important for several reasons. They provide constraints on the duty cycle and lifetime of radio-loud AGN, allow direct measurements of the density and pressure of the IGM on megaparsec scales, and offer insight into the interaction between radio jets and their large-scale environment. Moreover, the existence of such large radio structures requires that the jets remain stable and collimated over long timescales, placing limits on the variability and fueling processes of the AGN.

1.4.6 Underlying Principles and Equations

The study of radio emissions from galaxies and AGN requires a precise mathematical framework to interpret observational data and extract meaningful physical properties. Several key equations form the foundation of this analysis, allowing us to convert observed quantities into intrinsic properties of cosmic sources. This section introduces the fundamental relations used in radio astronomy, focusing on the derivation and application of these equations in the context of galaxy evolution studies.

We begin by examining radio luminosity, which quantifies the intrinsic power output of a source at a given frequency and is directly related to the observed flux density and distance. Next, we explore the infrared-radio correlation (IRRC), an empirical relation linking total infrared and radio luminosities, which serves as a powerful tool for estimating SFRs in galaxies. Finally, we introduce the concept of the K-correction, a necessary adjustment for comparing radio observations across different redshifts by shifting measurements to the rest frame of the source. Throughout this section, we connect these principles to the COSMOS field, demonstrating their application in large-scale surveys of star-forming galaxies and AGN.

1.4.6.1 Distinguishing Luminosity, Brightness and Radio luminosity

In astrophysics, the concepts of **luminosity** and **brightness** are often used interchangeably in casual discussions, but they represent fundamentally different properties of an astronomical source.

Luminosity (L) refers to the intrinsic total energy emitted per unit time by an astronomical object across all wavelengths or within a specific frequency range. It is an inherent property of the source and remains independent of the observer's location. Luminosity is typically measured in watts (W) or erg/s and serves as a crucial parameter for characterizing astrophysical processes, from star formation to AGN activity (Rybicki & Lightman, 1979).

On the other hand, brightness (or flux density, S_ν) represents the apparent energy received by an observer per unit area and per unit frequency or wavelength. Brightness depends on the observer's distance from the source, following the inverse square law: as the distance increases, the received flux decreases. This distinction is crucial, as radio observations measure flux density, which must be converted to luminosity to accurately study the intrinsic properties of galaxies and AGN (Condon, Cotton, & Broderick, 2002).

Radio luminosity (L_ν) quantifies the total power emitted by an astronomical source at a specific radio frequency. Unlike the observed flux density (S_ν), which depends on the distance between the source and the observer, radio luminosity is an intrinsic property of the source and provides a direct measure of its total energy output in the radio regime. It is typically expressed in W/Hz and is a fundamental parameter in the study of physical processes in galaxies, including synchrotron emission from relativistic electrons and thermal free-free radiation from ionized gas (Condon, 1992).

1.4.6.2 Radio Luminosity and Its Relation to Observed Flux

In radio astronomy, the **radio luminosity** L_ν of a source at a given frequency ν quantifies its intrinsic power output per unit frequency interval. It is a fundamental quantity for characterizing astrophysical sources such as galaxies and AGN. To connect observations with physical properties, one must relate the **observed flux density** S_ν to the intrinsic luminosity via the **luminosity distance** D_L . This relationship is governed by the inverse-square law and the geometry of isotropic emission.

The correct relation between radio luminosity and flux density is:

$$L_\nu = 4\pi D_L^2 S_\nu \quad (1.11)$$

where:

- L_ν is the specific luminosity in units of W/Hz,
- D_L is the luminosity distance to the source (typically in meters or parsecs),
- S_ν is the observed flux density at frequency ν , usually in units of $\text{W} \cdot \text{m}^{-2} \cdot \text{Hz}^{-1}$ or Jy.

This equation allows researchers to infer the total energy output of a radio source, corrected for the dilution of flux with distance. It is crucial for studying synchrotron emission from star-forming galaxies and AGN jets, as well as free-free radiation from ionized regions.

For clarity, the inverse relationship can be expressed as:

$$S_\nu = \frac{L_\nu}{4\pi D_L^2} \quad (1.12)$$

This form illustrates how flux density decreases with distance, following the expected $1/D^2$ behavior of isotropic radiation.

When dealing with cosmological sources, it is essential to also apply a **K-correction** to account for redshift-induced effects, which is discussed in the next section.

Here, L_ν is measured in **W/Hz**, providing a direct way to estimate the intrinsic power output of a source in the radio domain. This equation is critical in studying galaxies and AGN, as it allows researchers to infer the total energy output from observed flux measurements, correcting for the distance effect (Condon, Radio Emission from Normal Galaxies, 1992).

In the COSMOS field, deep radio surveys such as the **VLA-COSMOS 3 GHz Large Project** utilize this relation to estimate the radio luminosities of star-forming galaxies and AGN up to high redshifts. These estimates help trace cosmic star formation history and AGN activity over billions of years, providing insights into the evolution of galaxies in different environments (Smolcic, et al., 2017).

1.4.6.3 Linking Infrared and Radio Emission: IRRC and Redshift Corrections

A key concept in linking radio and infrared observations is the infrared–radio correlation (IRRC), an empirical relation that connects the total infrared luminosity of a galaxy, which traces dust-reprocessed starlight, with its radio luminosity, primarily arising from synchrotron emission. This relation is quantitatively expressed through the logarithmic infrared-radio ratio, q , defined as:

$$q_{TIR} = \log_{10} \left(\frac{L_{TIR}}{3.75 \times 10^{12} \text{W}} \right) - \log_{10} (L_{1.4\text{GHz}} / \text{WHz}^{-1}) \quad (1.10)$$

where L_{TIR} is the total infrared luminosity (in watts), $L_{1.4\text{GHz}}$ is the monochromatic radio luminosity at 1.4 GHz (in W Hz^{-1}), and $3.75 \times 10^{12} \text{W}$ is the reference normalization corresponding to a frequency of $80\mu\text{m}$ (Delhaize, et al., 2017); (Helou, Soifer, & Rowan-Robinson, 1985)). The IRRC is remarkably tight for star-forming galaxies across a wide range of redshifts, indicating a robust connection between star formation and radio synchrotron emission. For instance, compact AGN with flat or inverted spectra can show significant excess at radio frequencies relative to their infrared output, deviating from the standard IRRC trend.

The use of 1.4 GHz as the reference frequency in the IRRC is motivated by multiple factors. This frequency lies in a regime where synchrotron emission from supernova remnants dominates the radio output of star-forming galaxies, minimizing contamination from thermal (free-free) processes. Additionally, many large-scale radio surveys such as NVSS, FIRST, and VLA-COSMOS have been conducted at 1.4 GHz, facilitating consistent and comparable measurements across studies. Observations at lower frequencies are more susceptible to synchrotron self-absorption, which can lead to systematic biases, particularly in compact sources. As such, 1.4 GHz has become a stable and widely adopted benchmark for characterizing radio emission associated with star formation. Such deviations are commonly interpreted as indicators of AGN dominance, particularly in compact or radio-excess sources.

To ensure consistency when comparing radio fluxes and luminosities across different redshifts, it is common practice to apply a K-correction. This adjustment accounts for the redshift-induced shift in the observed frequency and enables astronomers to derive rest-frame quantities that better reflect the intrinsic emission of the source. The K-corrected luminosity is given by the expression:

$$L_{\nu(\text{rest})} = 4\pi D_L^2 S_{\nu(\text{obs})} (1+z)^{(\alpha-1)} \quad (1.14)$$

where α is the spectral index of the source, and z is the redshift.

The application of K-corrections is particularly important when constructing average spectral energy distributions or analyzing the evolution of star-forming galaxies over cosmic time (Tisanic, et al., 2019); (Smolcic, et al., 2017)).

Both the IRRC and the K-correction are integral to the interpretation of radio observations in the COSMOS field. They enable the disentanglement of AGN and star formation activity and allow for reliable estimates of physical properties across cosmic time, forming a crucial part of the methodology in this study.

The application of the equations and concepts discussed above is particularly powerful in the context of the COSMOS field, where extensive multiwavelength coverage and high-resolution radio data provide an ideal testbed for physically motivated analyses. By employing these tools—such as flux-to-luminosity conversions, infrared–radio correlation diagnostics, and redshift-based corrections—it becomes possible to extract intrinsic galaxy properties with high fidelity. In conclusion, the integration of these radio diagnostics within the COSMOS framework offers a powerful pathway toward disentangling the interplay between star formation, AGN activity, and large-scale structure evolution over cosmic time.

1.5 Thesis Outline and Cosmology

Following the introductory chapter, which presents the theoretical framework, relevant physical models, and the cosmological context of the study, the thesis is structured into six core chapters.

Chapter 2 outlines the methodology, including the generation of LOFAR-based radio masks, flux extraction procedures, and the identification of GRGs within the COSMOS field. It also provides a detailed overview of the external datasets (Counterpart, PyBDSF, SuperDeblended), and explains the cross-matching strategy, error analysis, and flux correction methodology applied to construct the final radio catalogue.

Chapter 3 focuses on the selection and normalization of AGN and SFG subsamples, and presents the construction of the RSEDs, including the visualization of binned data and the extraction of numerical results.

Chapter 4 provides the astrophysical interpretation of the RSEDs, comparing them with previous studies, highlighting their implications, and discussing the limitations and prospects for future research.

Chapter 5 summarizes the key findings and major conclusions, assesses the significance of the results, and outlines directions for future investigation.

Chapter 6 includes the complete list of references cited throughout the thesis.

Finally, **Chapter 7** contains supplementary material, including figures, tables, and the MATLAB code implementations used in the analysis.

In this study, we adopt the standard convention for the spectral index α , which characterizes how the observed flux density of a source varies with frequency. For synchrotron radiation, which is prevalent in star-forming galaxies and AGN, the spectral index generally ranges from about 0.5 to 1.0, indicative of the underlying energy distribution of relativistic electrons. Steeper indices ($\alpha > 1$) indicate rapid electron cooling at high frequencies, while flatter spectra ($\alpha < 0.5$) may result from thermal contributions or synchrotron self-absorption.

To convert observed quantities into intrinsic physical parameters, such as luminosity or comoving distance, a cosmological framework must be defined. In this study, we assume a spatially flat, low-density Universe dominated by dark energy, consistent with the standard Λ CDM (Lambda Cold Dark Matter) model. The specific cosmological parameters adopted are:

- **$H_0=70 \text{ km s}^{-1} \text{ Mpc}^{-1}$** : This is the Hubble constant, which sets the current expansion rate of the Universe. It relates the recessional velocity of a galaxy to its distance and is crucial for converting redshift measurements into physical distances. A value of 70 is commonly used in extragalactic studies and lies within the range of current observational constraints from missions such as WMAP and Planck.
- **$\Omega_M=0.3$** : This parameter represents the matter density of the Universe, normalized to the critical density. It includes both baryonic (ordinary) and dark matter. A value of 0.3 implies that matter accounts for roughly 30% of the total energy density of the Universe.
- **$\Omega_\Lambda=0.7$** : This denotes the contribution of dark energy (associated with the cosmological constant Λ) to the energy density of the Universe. Together with Ω_M , it satisfies the flatness condition $\Omega_M + \Omega_\Lambda = 1$, implying a geometrically flat Universe. Dark energy drives the accelerated expansion observed in the current epoch and plays a dominant role in cosmic evolution at low redshifts.

These cosmological parameters are essential for computing key quantities such as luminosity distances, comoving volumes, and the evolution of source populations over cosmic time. Their consistent application ensures that results derived in this study—such as radio luminosities, spectral energy distributions, and galaxy number densities—are physically meaningful and comparable with other works in the literature.

Chapter 2: Data and Methods

2.1 LOFAR 144 MHz Observations in COSMOS

For this study, we utilize deep radio observations of the COSMOS field at 144 MHz, conducted with the **LOFAR High Band Antenna** (HBA, Tim Shimwell & Cyril Tasse) array (PI: Vardoulaki, project DDT19_002). The total integration time amounts to 48 hours, acquired as twelve independent 4-hour observing blocks. This scheduling was necessitated by the low elevation of the COSMOS field (declination $< +30^\circ$), which restricts continuous observations to a maximum of approximately 4 hours per session.

All data analyzed in this work were obtained exclusively with the Dutch LOFAR stations, resulting in a synthesized beam of approximately 10–15 arcseconds. However, due to the low elevation of the COSMOS field, the beam is notably elongated (major axis up to $15''$), which should be considered when comparing source morphologies and positional uncertainties. The angular resolution is well matched to the available GMRT surveys and is suitable for robust cross-matching with higher-frequency radio catalogs. A Director's Discretionary Time (DDT) proposal has also been approved to re-observe COSMOS using both the Dutch and international LOFAR stations, which will ultimately deliver even higher angular resolution and improved image fidelity.

Data processing was carried out using the LoTSS 6" pipeline (Shimwell & Hardcastle, priv. comm.; Vardoulaki et al. in prep.), which applies facet-based, direction-dependent calibration to correct for ionospheric and instrumental beam effects. The resulting mosaics achieve a median rms sensitivity of approximately $100 \mu\text{Jy/beam}$ in the central region, although noise increases toward the edges due to primary beam attenuation and local calibration differences.

Source extraction and cataloguing were performed with PyBDSF, a tool optimized for the challenges of low-frequency LOFAR data. The resulting catalog contains over 5,000 radio sources above a conservative 5σ threshold, with astrometric alignment referenced to the VLA-COSMOS frame. The typical positional accuracy for strong sources is better than 0.5 arcseconds. Cross-matching procedures were applied to minimize spurious associations and sidelobe contamination.

A key strength of the LOFAR COSMOS survey is its synergy with the extensive ancillary data in COSMOS, enabling secure multiwavelength identifications and robust population studies of AGN and star-forming galaxies at low radio frequencies. The survey's depth and frequency coverage are especially powerful for probing steep-spectrum synchrotron emission, remnant AGN activity, and diffuse cluster sources.

The completeness and reliability of the source catalog were assessed via injection-and-recovery simulations, as described in the LOFAR Deep Fields release documentation. Overall, the LOFAR 144 MHz imaging of COSMOS represents a significant advance in the low-frequency radio characterization of this extensively studied extragalactic field.

2.2 Ancillary Radio Datasets in COSMOS: VLA (3 GHz & 1.4 GHz), GMRT (610/325 MHz)

The COSMOS field is uniquely equipped with a wealth of ancillary radio datasets spanning more than two orders of magnitude in frequency, from LOFAR at 144 MHz to VLA at 3 GHz and beyond.

This unprecedented combination of sensitive, homogeneous, and wide-area radio surveys enables comprehensive studies of both AGN and star-forming galaxies, facilitating robust population analyses, spectral characterization, and multi-frequency investigations. The synergy of these ancillary datasets with the extensive multiwavelength coverage in COSMOS allows for secure identifications, detailed spectral modeling, and statistically powerful studies of galaxy evolution across cosmic time.

2.2.1 The VLA-COSMOS 3 GHz Large Project

The **VLA-COSMOS 3 GHz Large Project** stands as one of the most ambitious radio continuum surveys to date, leveraging the advanced capabilities of the Karl G. Jansky Very Large Array (VLA) to achieve both remarkable depth and high spatial resolution across the full 2 square degree COSMOS field. Conducted between 2012 and 2014, the survey consists of 384 hours of on-source integration time in A and C configurations, yielding a uniform and sensitive mosaic at a central frequency of 3 GHz (10 cm).

The final 3 GHz mosaic achieves a median rms sensitivity of $2.3 \mu\text{Jy}/\text{beam}$ in the central regions, with local noise rising only modestly toward the edges of the field. The synthesized beam size is 0.75 arcseconds FWHM, providing sub-arcsecond angular resolution that is well-matched to the high-quality multiwavelength data available in COSMOS. This combination allows for robust identification, morphological classification, and cross-matching of radio sources with their optical, infrared, and X-ray counterparts.

A total of 10,830 radio sources are catalogued above a conservative 5σ detection threshold. Of these, approximately 8,696 sources fall within regions unmasked by bright optical or near-infrared artifacts, making them ideally suited for multiwavelength analysis (Smolcic, et al., 2017). The entire 2 deg² COSMOS field is covered by the survey, although the effective unmasked area is about 1.77 deg².

One of the hallmarks of the VLA 3 GHz survey is its combination of sensitivity and resolution. With a point source detection limit reaching down to $\sim 11.5 \mu\text{Jy}$ (5σ), the survey is sensitive to both faint star-forming galaxies at moderate to high redshift and to radio-quiet or radio-intermediate AGN. The high angular resolution enables the study of both compact and extended morphologies, facilitating the separation of core and jet components in radio AGN and the identification of resolved star-forming disks.

Calibration was performed using standard VLA procedures, with the flux scale tied to Perley & Butler (2013) and phase referencing performed regularly to ensure astrometric accuracy. The typical positional uncertainty is on the order of 0.05–0.1 arcsec for faint sources, improving for brighter detections.

Data products include both the full primary beam-corrected mosaic and several multi-resolution catalogs. The primary catalog is based on the PyBDSF source extraction algorithm, which is optimized for both compact and extended emission. Each catalog entry includes the integrated and peak flux densities, positional uncertainties, deconvolved sizes, and local noise estimates. The survey team also produced ancillary products such as residual maps and completeness simulations to quantify the fidelity and reliability of detections.

The VLA-COSMOS 3 GHz Large Project is distinguished by its unprecedented synergy with the extensive ancillary data available in COSMOS, including high-resolution optical/NIR imaging from HST and UltraVISTA, mid-IR data from Spitzer, far-IR/submm from Herschel, and deep Chandra/XMM-Newton X-ray data. This comprehensive coverage allows for nearly complete identification of radio counterparts, robust redshift estimation (spectroscopic and photometric), and detailed host galaxy characterization (Delvecchio, et al., 2017).

2.2.2 VLA-COSMOS 1.4 GHz Large Project

The **VLA-COSMOS 1.4 GHz Large Project** represents a cornerstone in the radio characterization of the COSMOS field and served as the primary high-resolution radio imaging prior to the advent of the 3 GHz survey. This survey provides deep 1.4 GHz (20 cm) imaging over the full 2 deg² COSMOS area, enabling statistical studies of both faint and bright radio populations with unprecedented spatial resolution for its time (Schinnerer, et al., 2010).

Observations were conducted with the VLA in both A and C configurations, accumulating more than 240 hours of integration between 2003 and 2006. The final mosaic achieves a typical synthesized beam of 1.5" × 1.4", matched to the best optical and infrared datasets available in the field. The median rms sensitivity reaches approximately 10–12 μJy/beam in the central regions, with the noise increasing gradually toward the edges due to primary beam attenuation and the survey geometry.

The main 1.4 GHz source catalog, as published in Schinnerer, et al., 2010, contains 2,864 radio sources detected above a 5σ threshold. The catalog provides accurate measurements of integrated and peak flux densities, astrometric positions, and deconvolved source sizes. The astrometric precision is of order ~0.15 arcsec for faint sources, benefiting from phase-referencing to calibrator sources and cross-checks with the optical/IR reference frame.

This legacy survey played a critical role in establishing the radio properties of the COSMOS galaxy population, providing the foundation for multi-frequency radio spectral analyses, luminosity functions, and source classification schemes in the field. The 1.4 GHz data have been widely used as the radio anchor for cross-matching with multiwavelength catalogs and for comparison with other deep radio surveys.

A key advantage of the 1.4 GHz survey is the balance between sensitivity and sky coverage, allowing for robust statistical studies of both AGN and star-forming galaxies over a wide redshift range. The field's extensive ancillary data—including HST, Subaru, UltraVISTA, Spitzer, and Herschel—enabled nearly complete identification and photometric redshift estimation for radio sources (Laigle, et al., 2016).

2.2.3 GMRT 610/325 MHz Surveys of COSMOS

The **Giant Metrewave Radio Telescope** (GMRT) surveys at 325 MHz and 610 MHz provide vital low-frequency coverage of the COSMOS field, bridging the spectral gap between traditional VLA surveys and the new generation of ultra-low-frequency observations with LOFAR. These datasets are vital for mapping radio continuum emission from both AGN and SFGs, particularly at frequencies that are sensitive to older synchrotron populations and steep-spectrum radio sources.

GMRT 610 MHz COSMOS Survey

The GMRT 610 MHz survey of COSMOS, originally described in Schinnerer et al. (2010) and re-processed by Tisanić et al. (2019), covers approximately 1 deg² centered on the COSMOS field (PI: H. R. Klöckner; project 11HRK01). The survey achieves a synthesized beam of ~6 arcseconds and reaches a median rms sensitivity of ~40 μJy/beam in the deepest regions (see Figure 11). The observations were performed using multiple pointings and mosaicked to ensure uniform noise and resolution characteristics across the central field.

Source extraction and cataloging were carried out using PyBDSF, yielding a catalog of 3,826 radio sources above a 5σ detection threshold ($\sim 300 \mu\text{Jy}$). The catalog provides positions, integrated and peak flux densities, local noise estimates, and deconvolved source sizes for each detection. Astrometric calibration was performed using cross-matching with the VLA-COSMOS 3 GHz catalog, resulting in a typical positional accuracy better than ~ 0.5 arcsec for bright sources.

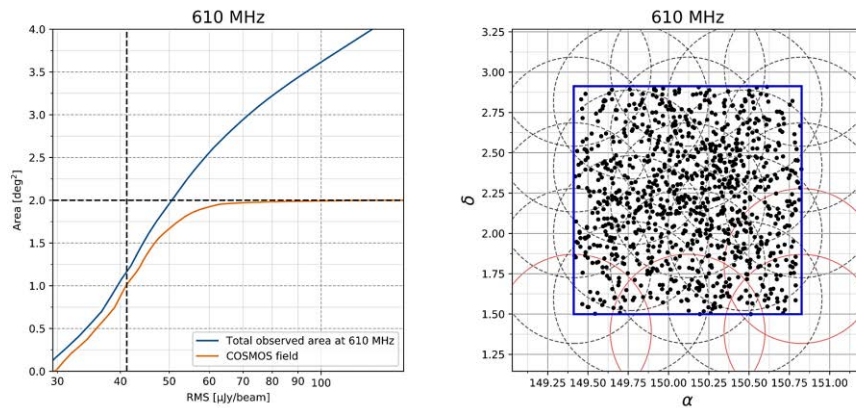


Figure 11: Left: The blue curve displays the visibility function for the complete 610 MHz mosaic, while the orange curve corresponds to the central 2 deg^2 COSMOS field. The black dashed line marks the median RMS measured within the COSMOS region. Right: Positions of 610 MHz radio sources (black dots) within the COSMOS field (blue box) (Tisanic, et al., 2019).

GMRT 325 MHz COSMOS Survey

Complementary observations at 325 MHz further extend the frequency baseline of the COSMOS radio coverage. The 325 MHz survey covers a similar $\sim 1 \text{ deg}^2$ area, with a synthesized beam of ~ 10 arcseconds and a median rms sensitivity of $97 \mu\text{Jy/beam}$ in the central field (see Figure 12, PI: S. Croft; project 07SCB01). Source extraction procedures were identical to those applied at 610 MHz, producing a catalog of 3,183 sources above a 5σ detection threshold ($\sim 500 \mu\text{Jy}$).

As with the 610 MHz data, astrometric alignment was verified and refined via cross-matching with higher-frequency VLA catalogs. The positional accuracy for bright sources is typically better than 1 arcsec. The GMRT catalogs include both compact and extended sources, providing valuable input for studies of radio spectral indices, especially for sources not detected at higher frequencies.

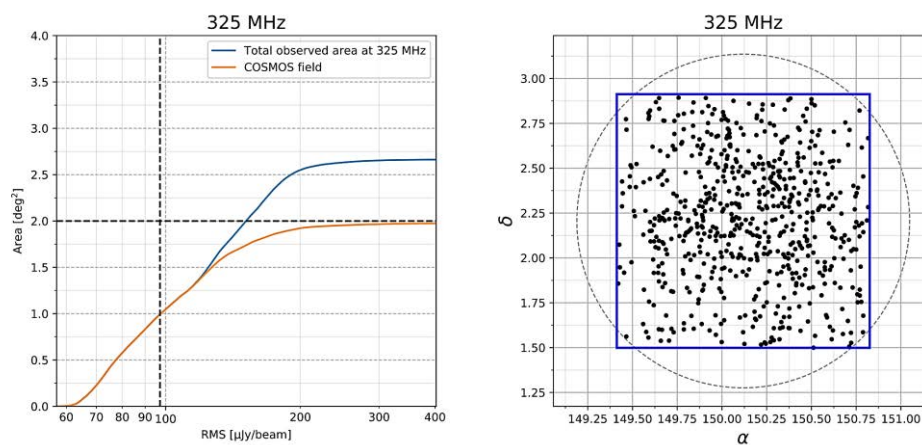


Figure 12: Left: The blue line represents the visibility function for the full 325 MHz mosaic, while the orange line corresponds to the central 2 deg^2 COSMOS area. The median RMS within the COSMOS field is indicated by the black dashed line. Right: Positions of 325 MHz radio sources (black dots) within the COSMOS field (blue box) (Tisanic, et al., 2019).

Survey Strategy, Calibration, and Data Products

Both GMRT surveys employed rigorous data reduction protocols to mitigate the effects of radio frequency interference (RFI), direction-dependent calibration errors, and primary beam attenuation. The images were corrected for beam effects, and flux density scales were anchored to standard calibrators. Detailed completeness and reliability analyses were conducted using Monte Carlo simulations (see Tisanić et al. 2019, 2020).

2.3 Understanding the datasets: Counterpart, PyBDSF, and SuperDeblended

A comprehensive understanding of the multiwavelength radio population in COSMOS requires the integration of several key catalogues, each constructed with distinct methodologies and scientific priorities. The Counterpart catalogue provides reliable multiwavelength associations and classifications, while the PyBDSF catalogue offers robust radio source extraction optimized for high-resolution continuum data. In parallel, the SuperDeblended catalogue incorporates sophisticated deblending techniques to recover accurate flux densities for sources in confused regions. Together, these datasets deliver complementary insights into source identification, photometric measurements, and the characterization of both AGN and SFGs. The following overview summarizes the purpose, structure, and scientific value of each catalogue, clarifying their individual roles and how they are integrated within the present analysis.

2.3.1 Catalogue Counterpart

The Counterpart catalogue is a critical component of the multi-wavelength identification process within the VLA-COSMOS 3 GHz Large Project (Smolcic, et al., 2017). In this study, we utilize the version of the counterpart catalogue released in January 2017 as part of the official data products of the VLA-COSMOS 3 GHz Large Project. This version, distributed under the filename `VLA_3GHz_counterpart_array_20170210_paper_smolcic_et_al.fits`, includes associations between radio sources and optical/NIR counterparts based on likelihood-ratio techniques, along with photometric redshifts and multiwavelength AGN/SFG classifications. Its primary purpose is to establish robust associations between the radio sources detected at 3 GHz and their corresponding counterparts in optical and near-infrared (NIR) surveys. These identifications are essential for assigning photometric and spectroscopic redshifts, galaxy properties such as stellar mass and SFR, and ultimately for distinguishing between AGN and SFGs.

As visualized in the TOPCAT column viewer, the catalogue contains 27 columns describing various properties of the radio sources and their associated counterparts. The most important fields include:

- **Astrometry:** RA and DEC for both the radio (RA_VLA_J2000, DEC_VLA_J2000) and optical counterparts (RA_CPT_J2000, DEC_CPT_J2000), along with their separation (SEP_VLA_CPT).
- **Redshift:** Photometric redshift (Z_BEST), essential for rest-frame calculations and luminosity functions.
- **Fluxes:** Total integrated 3 GHz flux (FLUX_INT_3GHz), and additional entries for radio flux at 10 cm and 21 cm, which are used to infer the radio spectral index (for sources with multi-frequency detections).
- **Star formation metrics:** Infrared luminosities (L_TIR_SF) and SFR estimates using various calibrations (e.g. SFR_KENN98).

- **AGN/SFG flags:** Multiple flags indicate the classification of each object (e.g., Xray_AGN, MIR_AGN, SED_AGN, HLAGN, MLAGN, Quiescent_MLAGN, Clean_SFG), following methods outlined in Delvecchio et al. (2017) and Novak et al. (2018).
- **Radio excess flag:** A boolean flag indicating whether the source shows excess radio emission compared to its SFR, a common AGN indicator.

By combining these parameters, it becomes possible to conduct statistically robust analyses of galaxy populations with respect to redshift, stellar mass, and radio luminosity, which in turn enables the classification of radio sources by physical type and the development of luminosity functions across cosmic time.

The construction of the catalogue relied on a likelihood-ratio (LR) technique to statistically associate radio sources with optical/NIR detections. This method, described in detail in Smolčić et al. (2017), balances the positional offset, magnitude distribution of true counterparts, and surface density of background sources to determine the most probable counterpart for each radio source. Only associations above a certain LR threshold were retained, ensuring high completeness and reliability.

The matching process incorporated the **COSMOS2015** catalogue (Laigle, et al., 2016) as its base for photometry and redshifts. This catalogue provides deep multi-band photometry (over 30 bands), enabling precise photometric redshift estimation with median accuracies of $\Delta z/(1+z) \approx 0.007$, making it ideal for identifying the optical counterparts of faint radio sources out to $z \sim 6$.

In total, out of the 10,830 radio sources in the full 3 GHz COSMOS catalogue, 7,729 sources were successfully assigned a counterpart via this method (Delvecchio, et al., 2017). These associations form the basis for all population studies of AGN and SFGs in COSMOS.

The Counterpart catalogue underpins a wide range of scientific investigations. It enables:

- **Separation of AGN from SFGs:** Using multiwavelength diagnostics, sources can be classified into radiative-mode (HLAGN), jet-mode (MLAGN), or purely star-forming galaxies, following the methodology in (Delvecchio, et al., 2017) and (Smolcic, et al., 2017).
- **Radio spectral studies:** With multiple frequencies included (e.g., 1.4 GHz, 3 GHz), spectral indices can be derived for sources with sufficient coverage, an essential component of this thesis.
- **Environmental studies:** The availability of redshifts and classifications allows for clustering analysis, group membership identification (Vardoulaki, Gozaliasl, Finoguenov, Novak, & Khosroshahi, 2025), and studies of galaxy evolution in different environments.
- **Calibration of selection effects:** By having a well-defined matching procedure and completeness corrections, the catalogue allows for volume-limited and bias-corrected statistical studies (e.g., (Smolcic, et al., 2017)).

Accordingly, the Counterpart catalogue underpins multiwavelength and statistical studies focused on radio sources in the COSMOS field. It's creation marks a significant step forward in bridging radio astronomy with optical and IR surveys to enable physically meaningful classifications and evolutionary tracking across redshift.

2.3.2 Catalogue PyBDSF

The second major catalogue used in this work is constructed via **PyBDSF** (Python Blob Detection and Source Finder), an automated tool designed to identify and parametrise radio sources in astronomical images (Mohan & Rafferty, 2015; Tasse, et al., 2021). PyBDSF is widely used in radio astronomy for processing interferometric continuum maps, and is the standard source extraction tool adopted in major radio surveys such as **LoTSS**, **VLA-COSMOS**, and **EMU**.

In this project, PyBDSF was applied to the full-band 3 GHz VLA-COSMOS continuum mosaic. The final PyBDSF catalogue used in this work is stored in the file `image_fullband.int.restored_3GHz-Default-3GHzarea-SupplyMaps-withFlagBeam.srl.fits`, which contains 1984 radio sources (each fitted with one or two Gaussian components) and 46 columns of parameters. These detections represent a purely radio-based identification, prior to any cross-matching with ancillary data. The resulting catalogue is essential for quantifying fluxes and morphologies directly from the radio maps and serves as the basis for constructing the initial source list each fitted with one or more Gaussian components.

PyBDSF identifies sources through the following multi-step process:

1. **Noise Characterization:** First, the software estimates the background noise (RMS) and mean maps across the image. These are critical for detecting sources above a statistically meaningful threshold (commonly 3σ or 5σ).
2. **Island Detection:** Pixels above a specified threshold (`thresh_pix`) are grouped into “islands” if they are connected to surrounding pixels above a lower threshold (`thresh_isl`). These islands are considered candidate source regions.
3. **Gaussian Fitting:** Each island is modeled as a combination of one or more 2D Gaussian components, optimally fitting the emission structure. The best-fit parameters are then reported per source.
4. **Residual and Model Maps:** The residual image (original minus fitted models) and the model image are generated to assess the quality of the fit and completeness of source recovery.

This process involves the decomposition of complex radio emission into simpler Gaussian components, accompanied by the construction of noise and model maps that describe the background characteristics and fitted source structure, respectively. These outputs are essential for assessing the quality of the fit and ensuring the reliability of the source extraction.

The PyBDSF catalogue contains 46 columns, including:

- **Source position:** RA, DEC coordinates of the source centroid, along with their uncertainties (`E_RA`, `E_DEC`), and the boundaries of the fitted Gaussian components (`RA_max`, `DEC_max`, `RA_min`, `DEC_min`).
- **Fluxes:** Total integrated and peak flux densities at 144 MHz (`Total_flux`, `Peak_flux`), as well as their uncertainties (`E_Total_flux`, `E_Peak_flux`). It also includes the total flux integrated over the source island (`Isl_total_flux`, `E_Isl_total_flux`).
- **Morphological parameters:** Major and minor axes and position angle (`Maj`, `Min`, `PA`), with associated errors (`E_Maj`, `E_Min`, `E_PA`). The catalogue also provides observed (pre-deconvolution) values (`Maj_img_plane`, `Min_img_plane`, `PA_img_plane`) and their uncertainties (`E_Maj_img_plane`, `E_Min_img_plane`, `E_PA_img_plane`).
- **Deconvolution parameters:** Parameters representing the intrinsic (deconvolved) source size and orientation after beam removal (`DC_Maj`, `DC_Min`, `DC_PA`), and their uncertainties (`E_DC_Maj`, `E_DC_Min`, `E_DC_PA`).

- **Residual and noise statistics:** Local rms noise estimate within the island (Isl_rms), residual rms and mean maps (Resid_Isl_rms, Resid_Isl_mean) computed after model subtraction.
- **Quality flags:** Beam flag (Flag_beam) indicating whether the source is unresolved, and a classification code (Code) for further analysis or filtering.

These parameters allow not only accurate photometry but also a morphological characterization of each radio source (e.g. compact vs extended, resolved vs unresolved).

For deep fields with a high source density—such as COSMOS or ELAIS-N1—standard PyBDSF runs can suffer from elevated rms noise estimates due to the presence of numerous sources. This can bias source detection toward brighter sources, missing faint but genuine detections. To address this, the following **double-pass PyBDSF approach** was implemented, based on the methodology used in the **ELAIS-N1 Deep Fields catalogue (LoTSS-Deep DR2)**:

- **First pass:** PyBDSF is run on the original image to identify and subtract bright sources, creating a residual image.
- **Residual-based rms map:** A new, more accurate rms noise map is derived from the residual image, reducing background overestimation.
- **Second pass:** PyBDSF is rerun using the improved rms map as input, enabling deeper and more reliable detection of faint sources near the noise level.
- **Outcome:** This method increases source completeness—especially in high-density fields—while maintaining a low false detection rate (only ~2–3% for $4 < \text{SNR} < 5$; see LoTSS DR2, Sect. 2).
- **Documentation:** The process follows the detailed recipe described in *Making the ELAIS-N1 DR2 PyBDSF Catalogue* (2021), demonstrating ~30% improved depth over the standard 5σ approach.

This approach substantially increases source completeness at low signal-to-noise ratio (SNR) thresholds (e.g. from 5σ down to 4σ), without significantly inflating the false detection rate (only ~2–3% for $4 < \text{SNR} < 5$; see also discussion in ELAIS-N1 DR2 documentation).

The PyBDSF catalogue plays a foundational role in:

- **Quantifying intrinsic radio properties** (e.g. total and peak flux densities, angular sizes),
- **Spectral index measurements**, when cross-matched with lower-frequency observations (e.g. LOFAR 150 MHz),
- **Morphological classification**, which is essential for identifying FR-type sources or distinguishing jets/lobes from compact AGN cores,
- **Input to cross-matching procedures** with optical/IR catalogues (e.g. via likelihood ratio methods),
- **Supporting environmental studies** by measuring source sizes, luminosities, and orientations in relation to host galaxies and group dynamics (e.g. (Vardoulaki, Gozaliasl, Finoguenov, Novak, & Khosroshahi, 2025)).

Furthermore, the radio-only detection offered by PyBDSF is complementary to the **Counterpart catalogue**, which relies on positional matching. As such, PyBDSF can detect sources lacking reliable optical/IR counterparts—important for identifying obscured or high-redshift AGN.

2.3.3 Catalogue SuperDeblended

The **SuperDeblended catalogue** represents one of the most complete and accurate datasets of FIR-to-radio photometry in the COSMOS field to date. In this study, we utilize the version provided in the SuperDeblend-LOFAR-FIR-radio-catalog.fits. file, which contains 201,468 sources and is based on the COSMOS implementation by Jin et al. (2023). It is the product of a sophisticated photometric extraction technique designed to overcome the severe source confusion inherent in far-infrared and submillimeter imaging, especially in deep extragalactic surveys like COSMOS. This technique, developed by Liu et al. (2018) ["Super-deblended" Dust Emission in Galaxies: I. The GOODS-North Catalog and the Nature of Faint Submillimeter Galaxies] and adapted for COSMOS by (Jin, et al., 2018), allows for high-fidelity flux measurements of dusty star-forming galaxies, even in the most crowded regions of the sky. The full methodology and updated implementation for the COSMOS field are detailed in Jin et al. (2023), which expands the SuperDeblended framework to the full COSMOS area using Herschel, SCUBA2, and VLA priors.

The need for “super-deblending” arises from the large point spread functions (PSFs) of FIR/(sub)mm instruments like Herschel/SPIRE and SCUBA2, where typical beam sizes can range from 10 to 36 arcseconds. These scales are far larger than typical galaxy separations, resulting in blended flux contributions from multiple sources in a single beam. Traditional blind source extraction methods often fail to resolve these sources or accurately assign their fluxes.

The SuperDeblended approach circumvents this problem by employing **prior-based photometry**, where source positions are fixed based on high-resolution imaging (from radio or near-infrared data), and photometric fitting is performed on the lower-resolution FIR/(sub)mm images. Crucially, the prior list includes not only sources detected in MIPS 24 μ m or VLA radio maps but also a mass-selected sample from the UltraVISTA Ks-band catalogue, achieving a highly complete prior population of more than **150.000 galaxies**.

Photometric measurements were performed using **PSF fitting** (with *galfit*), and the fluxes were calibrated using extensive **Monte Carlo simulations**. This calibration accounts for flux biases, residual crowding, and uncertainties introduced by faint-source subtraction. Each band was fitted separately, and the final photometry was iteratively refined through **SED fitting** to optimize source selection at each wavelength step.

The version of the SuperDeblended catalogue used in this project contains **201,468 sources**, with measurements and derived properties across the infrared and radio regimes. Key columns include:

- **Positional Information:** RA, Dec of each source.
- **Redshifts:** Photometric redshift (z_{phot}) and spectroscopic redshift (z_{spec}), where available.
- **Signal-to-Noise Ratio and Redshift-corrected FIR Luminosities:** SNR_IR, z_{FIR} , SFR_IR, and their associated errors
- **Fluxes at Multiple Radio Frequencies:** f10cm, f20cm, f23cm, f50cm, along with simulation-derived uncertainties (dfXcm).

These parameters allow a detailed reconstruction of the radio spectral energy distributions (SEDs) for each source, which is particularly relevant to this thesis.

The SuperDeblended pipeline is characterized by:

- **FIR/(sub)mm flux prediction via SED fitting** using a four-component model (stellar, AGN torus, dust continuum, and radio power-law).
- **Selection of priors for each band** based on predicted flux and local image depth, ensuring optimal source density (~ 1 per beam).

- **Faint-source subtraction** for priors below the detection threshold but with confident flux estimates.
- **Residual source fitting** to account for sources detected in residual images that were not part of the initial prior list.
- **Band-by-band corrections** using simulations, particularly for crowding and background contamination.

The method accounts for challenges specific to the COSMOS field, such as its shallower PACS and MIPS data relative to GOODS-N. (Jin, et al., 2018) adapted the L18 GOODS-N pipeline by adding a **stellar-mass selected prior layer**, ensuring that most FIR-detectable galaxies were included, even if they lacked 24 μ m or radio detections.

The SuperDeblended catalogue is essential for any study involving:

- **Dusty star-forming galaxies** across redshifts $0 < z < 7$
- **Accurate SFR measurements** from dust emission, particularly at high redshift where UV-based SFRs are unreliable
- **Spectral index derivations**, by combining LOFAR, GMRT, and VLA fluxes
- **Population statistics**, such as luminosity functions, number counts, and redshift distributions of IR-selected galaxies
- **Identification of extreme starbursts and radio-loud AGN candidates** in the FIR regime

In this thesis, the SuperDeblended catalogue serves as the principal reference for retrieving flux measurements across multiple radio frequencies and for extract redshift (spectroscopic or photometric). Owing to its exceptional photometric depth and accuracy, it allows for the detailed reconstruction of SED for a large number of sources. The combination of reliable FIR luminosities with matched radio fluxes enables robust determinations of the infrared–radio correlation, facilitating studies of its possible evolution with redshift and galaxy type.

Furthermore, by integrating flux information from instruments operating at 144 MHz, 325 MHz, 610 MHz, 1.4 GHz, and 3 GHz, the catalogue provides the necessary foundation for analyzing spectral curvature, quantifying radio excess, and disentangling contributions from star-forming and AGN-dominated galaxies. These capabilities make the SuperDeblended catalogue an indispensable tool for understanding the physical emission mechanisms and evolutionary pathways of radio sources in the COSMOS field.

2.4 Catalogue LOFAR-masks generation

2.4.1 Source detection, flux measurement, and centroid determination

In the COSMOS field, we are dealing with a vast number of radio sources, which means that their detection heavily depends on the method used to identify them. In our study, we implement an original approach that is based on a thresholding technique. Through this, we create a set of binary masks by thresholding the image, isolating regions of significant emission, from which we extract information—we are able to construct our own final catalogue.

More specifically, the map we are working with is the `image_fullband.int.restored_3GHz.fits`, and using the Algorithm 1 script in MATLAB, we set our detection threshold at 3σ . In our case, the σ corresponds to 150×10^{-6} Jy/beam in flux density per beamunits, thus our threshold is set to 450×10^{-6} Jy/beam. The code is instructed to assign the value 1 (white) to regions above this threshold in the new

map, and the value 0 (black) to regions below it. In this way, we generate a new binary map, which visually appears in black and white but has successfully isolated all values in the original map that exceed the defined threshold.

This process is performed on a per-pixel basis, since individual pixels contain the flux density information we are interested in. Accordingly, each detected source corresponds either to a single pixel or to a “island” of adjacent pixels above the threshold, with its corresponding pixel count and a unique source ID. Simultaneously, we extract the centroid of each radio source and construct a catalogue that includes this information.

The centroid is defined as the pixel within the source region that has the highest flux value among all its constituent pixels. This centroid is initially expressed in image coordinates (x, y), but will later be converted into Right Ascension (RA) and Declination (Dec) values.

Subsequently, by using the Algorithm 2, and relying on the necessary header information, the centroid coordinates are transformed into RA and Dec values. These are the fundamental quantities used in radio astronomy to describe the position of a radio source on the sky. This transformation relies on the World Coordinate System (WCS) parameters contained in the FITS file header, which define how to convert image pixel coordinates into celestial coordinates.

Specifically, the following WCS parameters are used:

- ✓ CRPIX1, CRPIX2: the reference pixel coordinates (i.e., the reference point in the image array),
- ✓ CRVAL1, CRVAL2: the corresponding celestial coordinates (RA, Dec) of the reference pixel,
- ✓ CDELT1, CDELT2: the pixel scale in degrees (deg/pixel) along the x and y axes.

The transformation from pixel coordinates (x, y) to celestial coordinates is performed according to the general form of the WCS transformation:

$$RA = (x - CRPIX1) * CDELT1 + CRVAL1 \quad (2.1)$$

$$Dec = (y - CRPIX2) * CDELT2 + CRVAL2 \quad (2.2)$$

where (x, y) are the pixel coordinates, and the CRPIX, CDELT, and CRVAL values are obtained from the FITS file header.

In cases where the FITS header includes image rotation (i.e., a non-zero value in CROTA2), the code is adapted to incorporate a rotation matrix using the CD matrix (parameters: CD1_1, CD1_2, CD2_1, CD2_2). This approach is essential for datasets where the observation field is rotated, and in such cases, the use of CD matrix parameters takes precedence over the standard CDELT/CROTA values. The transformation in this case follows the form:

$$RA = CRVAL1 + CD1_1 * dx + CD1_2 * dy \quad (2.3)$$

$$Dec = CRVAL2 + CD2_1 * dx + CD2_2 * dy \quad (2.4)$$

where $dx = x - CRPIX1$, $dy = y - CRPIX2$.

The calculation of the flux (flux_sum) for each region is performed by summing the pixel values within that region. The position of each region (centroid) is defined based on the coordinates of the pixel with the maximum flux inside the region, instead of calculating the average RA and Dec. The results for each region are recorded in a data table and stored in a file, following an initial column-based structure.

2.4.2 Flux extraction, beam correction, and catalogue preparation

After the creation of the binary masks and the initial construction of a source list containing the radio sources that exceed our 3σ detection threshold, the next step is data cleaning and flux correction.

More specifically, for each radio source identified through the masking process—consisting of one or more adjacent pixels—the code retrieves the pixel values, sums them for each region, and calculates the total flux of the source in units of Jy/beam. These values represent the sum of pixel flux densities, as each pixel in the radio map contains flux density information expressed in Jy/beam. As a sanity check, we select source ID 700 from our catalogue and manually verify that the summed pixel values yield a total of 0.014586503413739 Jy/beam, confirming the correct execution of the script.

Before converting the flux values from Jy/beam to Jy, we need to estimate how many pixels correspond to one beam area. This quantity is known as pixels per beam and is calculated by dividing the beam area (in arcsec²) by the area of a single pixel. According to the FITS header, each pixel has a side length of 1.5 arcsec, resulting in an area of:

$$1.5 \text{ arcsec} \times 1.5 \text{ arcsec} = 2.25 \text{ arcsec}^2.$$

Thus, the number of pixels per beam is:

$$\text{Pixels per Beam} = A_{\text{beam}} / 2.25 \quad (2.5)$$

To compute the beam area A_{beam} , we consider the elliptical shape of the LOFAR synthesized beam. Based on the FITS header, the beam has full width at half maximum (FWHM) values of $\theta_1 = 6 \text{ arcsec}$ (minor axis, BMIN) and $\theta_2 = 15 \text{ arcsec}$ (major axis, BMAJ). The area of the beam in arcsec² is calculated using the following equation for a 2D Gaussian beam:

$$A_{\text{beam}} = \frac{\pi * \theta_1 * \theta_2}{4 * \ln(2)} \quad (2.6)$$

where θ_1 and θ_2 are the FWHM values along the minor and major axes, respectively, and the division by $\ln(2)$ arises from the integration of the 2D Gaussian function and ensures consistency with standard radio astronomy definitions of beam area.

Once the number of pixels per beam has been computed, we convert the measured flux densities (in Jy/beam) into total flux values (in Jy) for each radio source using the following relation:

$$\text{Flux (Jy)} = \text{Region Flux (Jy/beam)} / \text{Pixels per Beam} \quad (2.7)$$

which yields total flux values that are normalized with respect to the beam and are suitable for physical interpretation and comparison across sources (Algorithm 3).

In parallel, we classify each source as either **resolved** or **unresolved**. This is done by comparing the physical area of the source to the size of the beam. The source area is first converted from pixel units to arcsec² by multiplying the number of pixels by 2.25 arcsec². Then, we subtract the beam area from this value:

$$\text{Corrected Area} = \text{Source Area (arcsec}^2\text{)} - A_{\text{beam}} \quad (2.8).$$

This subtraction is applied exclusively to the source area and serves the purpose of classification. If the corrected area is positive, the source extends beyond the beam and is considered resolved. If the corrected area is negative, the beam encompasses the source, and it is thus considered unresolved.

As mentioned earlier, the elliptical beam influences the morphology of sources in the map, often giving them an oval appearance due to the asymmetry of the beam axes. Furthermore, the relatively low

elevation of LOFAR during the COSMOS observations contributed to a poorer angular resolution, reinforcing the importance of correcting for the beam.

Before proceeding with the cross-matching between our sources and external catalogs, it is essential to first complete the construction of our own final catalog, specifically incorporating the treatment of the two GRGs described above. In addition, we must perform key corrections and calculations, including the estimation of the local RMS and signal-to-noise ratio (SNR), as well as the correction of the measured flux densities based on information from the deblended catalog. Only after these steps have been completed can we reliably proceed with source associations and classification.

2.4.3 Identification and deblending of the two GRGs in the COSMOS field

In this section, we present the identification and detailed deblending procedure applied to the two recently discovered Giant Radio Galaxies (GRG-1 and GRG-2) in the COSMOS field. Both GRGs were detected as part of the MIGHTEE survey and exhibit linear sizes of 2.4 and 2.0 Mpc, respectively (Delhaize, et al., 2021). Due to their exceptional size and low surface brightness, a dedicated methodology was required to accurately measure their total flux densities and associate them with their host galaxies. Here, we outline the region definition, source association, and flux extraction steps carried out for each GRG, providing the basis for their inclusion in the final radio source catalog (Algorithm 8). *For a detailed overview of the general properties and astrophysical significance of Giant Radio Galaxies, the reader is referred to Section 1.4.5.*

2.4.3.1 GRG-1

The first detected GRG, designated **MGTC J095959.63+024608.6** (hereafter **GRG-1**), is a striking example of a Giant Radio Galaxy with a physical size of **2.42 Mpc** and a spectroscopic redshift of **$z = 0.1656$** . GRG-1 exhibits two extended lobes and a central core, displaying a morphological characteristic of FR II-type radio galaxies, although its total radio power lies below the classical FRI/FRII luminosity divide (Delhaize, et al., 2021). This is illustrated in Figure 13, which presents multi-frequency images of GRG-1, revealing its double-lobed morphology and central core.

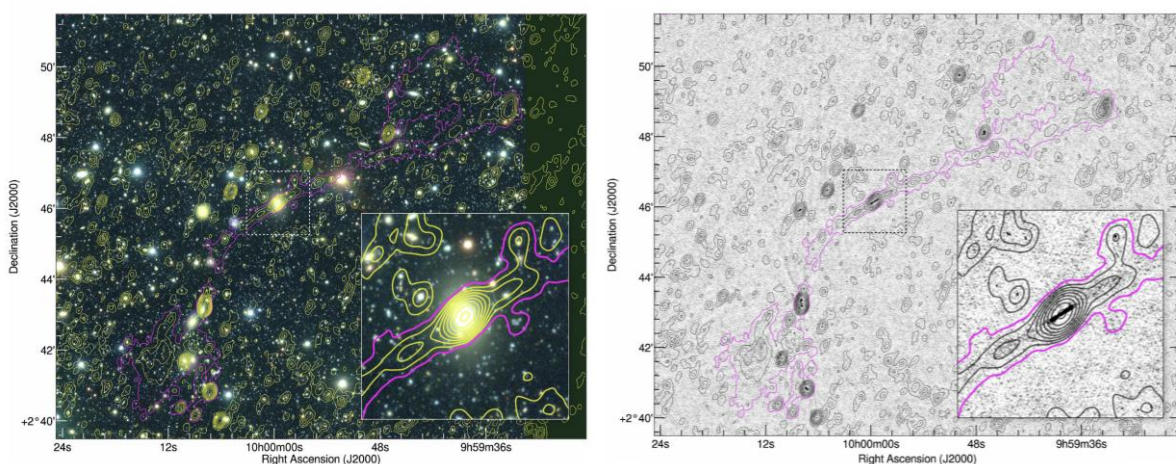


Figure 13: Composite multi-frequency images of **GRG1**. Left: **GRG1** as observed in MIGHTEE (yellow contours), overlaid on a composite optical image from the HSC *g*, *r*, and *i* bands. Radio contours increase in fixed steps up to high dynamic range, with the lowest level at $7 \mu\text{Jy beam}^{-1}$ shown in magenta to trace the full spatial extent of the GRG. The inset highlights the core region and the host galaxy. Right: **GRG1** as seen in the VLA 3 GHz data (greyscale background), with MIGHTEE contours overlaid in black (same levels as on the left). The hotspot in the northern lobe is weakly visible at 3 GHz. The inset shows a zoom-in of the central region, revealing a double-lobed/jetted morphology.

The host galaxy is a massive elliptical galaxy, with a stellar mass of approximately $\sim 2.6 \times 10^{11} M_{\odot}$, a low star formation rate ($\sim 2 M_{\odot}/\text{yr}$), and no indications of radiatively efficient AGN activity, either in the infrared or X-ray regime. The location of GRG-1 within a small galaxy group—confirmed through the detection of diffuse X-ray emission—supports scenarios in which interactions with the intergalactic and intracluster medium (IGM, ICM) shape the morphology and dynamics of the lobes.

A significant characteristic of GRG-1 is the detection of bending in the southern jet, indicating possible interactions with an anisotropic environment or varying density conditions. Additionally, the northern lobe exhibits peripheral diffuse emission, oriented perpendicular to the main jet axis, possibly as a result of local gas flows within the IGM.

The spectral index of the core is relatively flat ($\alpha \approx 0.51$), suggesting contributions from synchrotron self-absorption at the base of the jet. The total radio power at 1.2 GHz is approximately $2.4 \times 10^{24} \text{ W Hz}^{-1}$, a value lower than the typical average for known GRGs, placing GRG-1 within a previously unexplored region of the radio power–linear size (P–D) diagram.

In the COSMOS field, and based on LOFAR detections, GRG-1 is not optically visible, likely due to the limited spatial resolution of the radio telescope. As a result, its identification was achieved indirectly rather than directly. Specifically, the approach followed involved using the MIGHTEE_Continuum_DR1_COSMOS_5p2arcsec and 8p9arcsec to extract region files around the location of GRG-1 (ds9 region), convert them into the appropriate coordinate system (WCS – fk5), and then determine through custom code which mask IDs fall within the GRG. Finally, we considered the mask that lies closest to the actual center of the giant radio galaxy as representative (see Figure 14).

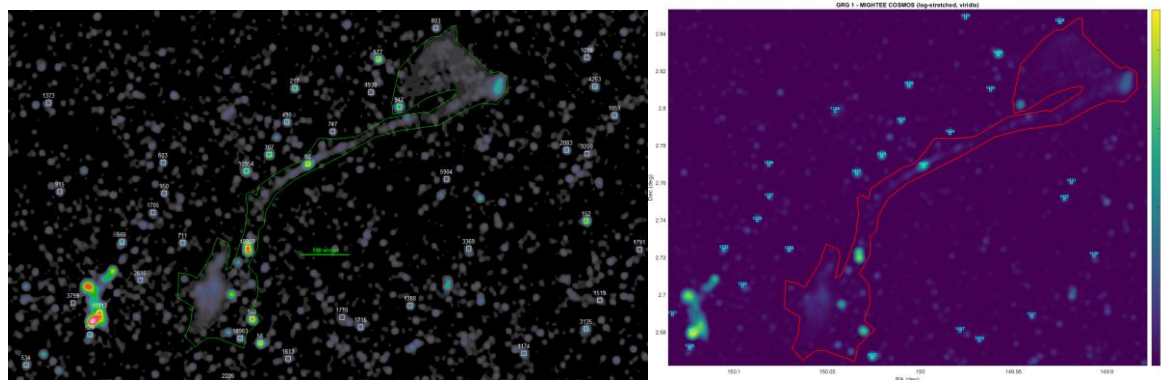


Figure 14: GRG-1. Left: The GRG-1 radio galaxy as visualized in SAOImage DS9, with the polygonal region manually defined to encompass its extended radio structure. Right: The same region plotted using MATLAB on the log-stretched, viridis-scaled MIGHTEE-COSMOS 3 GHz image. Overlaid in red are the final polygon boundaries used to calculate the total flux assigned to the central source ($ID_{VLA} = 80$, $ID_{MASKS} = 1447$). Cyan markers denote the positions and IDs of the radio sources.

As observed in Figure 14, the central source of GRG-1 coincides with the source $ID = 1447$ in the catalog constructed based on the masks. We define this source in MATLAB as the central one, which will be assigned the total flux, calculated as the sum of all radio sources located within the region we have defined. The sources found within this region, including the central source, are those with the following IDs: 1385, 1386, 1389, 1391, 1397, 1402, 1407, 1408, 1410, 1412, 1414, 1447, 1459, 1499, 1502, 1507, 1563. However, we do not subtract the flux of sources 1408, 1410, 1412, and 1502, as these are located at different redshifts and are not physically associated with the GRG-1 structure. Therefore, the total flux is computed by summing only the sources within the region that are assumed to be physically connected to the central component.

In conclusion, GRG-1 appears to represent an advanced evolutionary system indicating long-term activity in a low-density environment. It offers an excellent laboratory for studying the interaction of

radio sources with the cosmic medium, as well as the long-term evolution of AGNs over cosmological timescales.

2.4.3.2 GRG 2

The second object, **MGTC J100016.84+015133.0** (hereafter **GRG-2**), is another remarkable Giant Radio Galaxy, identified in the MIGHTEE survey within the COSMOS field. GRG-2 exhibits a projected linear size of **2.04 Mpc** and a spectroscopic redshift of **$z = 0.3363$** , confirming its status among the largest known radio galaxies. Notably, both lobes and a possible hotspot in the northern lobe are clearly detected for the first time in MIGHTEE, while previous high-resolution VLA observations revealed only the compact radio core, with the extended emission resolved out due to surface brightness limitations. The core is located at $RA = 10^h00^m16.84^s$, $Dec = +01^d51^m33.0^s$, coinciding with an optically luminous elliptical host galaxy (see Figure 15).

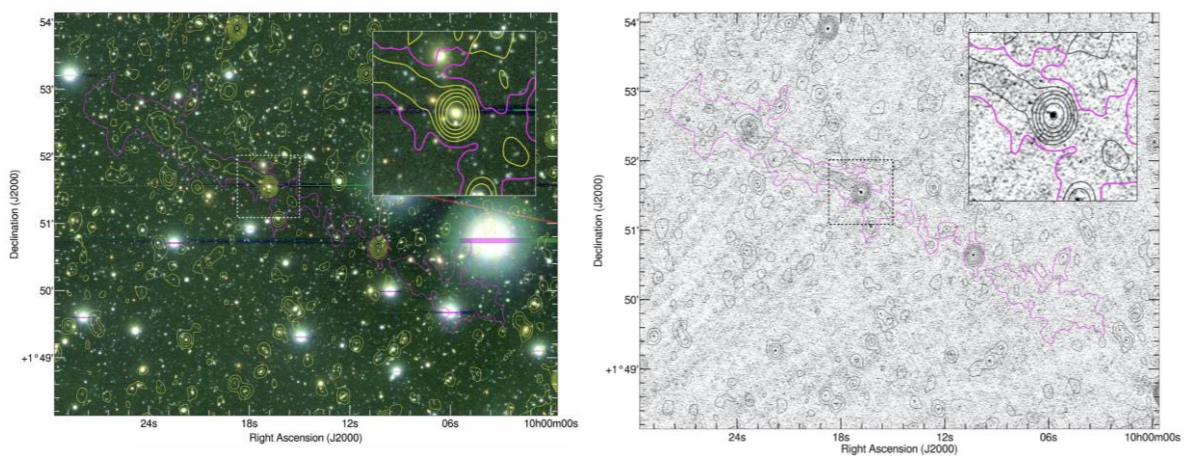


Figure 15: Composite multi-frequency images of **GRG-2**. Left: GRG2 as seen in MIGHTEE (yellow contours), overlaid on an optical composite from the HSC g, r, and i bands. Radio contours are shown in increasing steps, with the lowest contour at $5 \mu Jy beam^{-1}$ (highlighted in magenta) marking the full extent of the radio source. The inset shows an enlarged view of the core region. Right: GRG-2 as detected in the VLA 3 GHz data (greyscale background) with black MIGHTEE contours overlaid (same levels as in the left panel). A compact radio core is clearly seen at 3 GHz, and faint extended emission is visible in the northern lobe. Together, these views confirm the double-lobed structure of GRG-2 and provide evidence for a central AGN, highlighting the symmetry and scale of the source. (Delhaize, et al., 2021)

The host galaxy, also elliptical, exhibits a high stellar mass ($\sim 6.3 \times 10^{10} M_{\odot}$), a very low star formation rate ($\sim 0.5 M_{\odot}/yr$), and no signs of strong AGN activity in either the infrared or X-ray bands. The multiwavelength analysis of its spectral energy distribution (SED) did not reveal any significant contribution from radiative AGN mechanisms, confirming that GRG-2 is another example of a Low-Excitation Radio Galaxy (LERG).

Despite its symmetric morphology, a brightness asymmetry is observed between the two lobes: the northern lobe appears brighter, while the southern lobe is more diffuse. This asymmetry may indicate relativistic beaming effects due to the inclination of the system with respect to the line of sight. The spectral index of the core is nearly flat ($\alpha \approx +0.29$), a characteristic typical of an active central jet.

The total radio power at 1.2 GHz is approximately $\sim 1.3 \times 10^{24} W Hz^{-1}$, placing GRG-2 below the average power of known GRGs. Like GRG-1, GRG-2 occupies a previously unexplored region of the power–linear size (P–D) diagram, consistent with the expected evolution of low-power radio sources in low-density environments.

GRG-2 is hosted within a small galaxy group, where it constitutes the most massive member—supporting the idea that the surrounding environment plays a crucial role in shaping and sustaining large-scale radio structures.

In the COSMOS field, GRG-2 is not optically detected, similar to GRG-1. Consequently, an indirect approach is required to identify and estimate its flux density, by creating a new region file in DS9 and manually outlining the source area as precisely as possible.

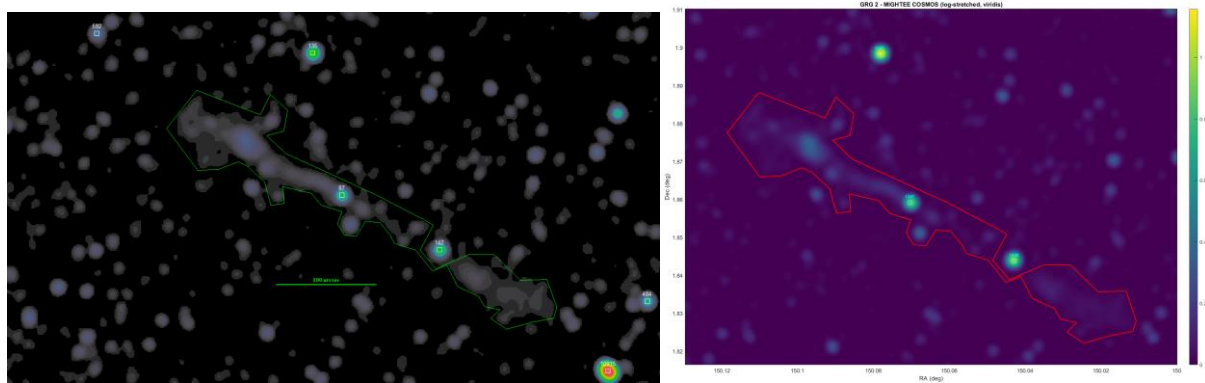


Figure 16: GRG-2. Left: The GRG-2 radio galaxy as visualized in SAOImage DS9, with the polygonal region manually defined to encompass its extended radio structure. Right: The same region plotted using MATLAB on the log-stretched, viridis-scaled MIGHTEE-COSMOS 3 GHz image. Overlaid in red are the final polygon boundaries used to calculate the total flux assigned to the central source (ID_VLA=87, ID_MASKS=1367). Cyan markers indicate the positions and IDs of the radio sources.

As shown in Figure 16, GRG-2 also contains several sources from the masks within its defined region. Specifically, based on the mask catalog IDs, these sources are: 1338, 1339, **1367**, 1371, and 1383. The central source in this case is identified as ID = 1367. Following the same procedure as for GRG-1, we summed the individual flux densities of all associated sources and assigned the total flux to this central component.

It is worth noting a specific detail here. In the final catalog we are constructing, and more precisely in the *ResolvedFlag* column, the central IDs corresponding to **GRG-1** and **GRG-2** (i.e., **1447** and **1367** respectively) will be labeled with their names rather than the standard “Resolved” or “Unresolved” designation, so that they can be distinguished from the other catalog entries. These two objects represent special cases that required dedicated analysis and interpretation.

2.5 Correction, error analysis and matching process of radio source measurements

2.5.1 Correction of flux densities through identification and removal of secondary blended sources

After establishing the mask-based source detection methodology in the previous section, we proceed with the correction of the flux density originally measured for each radio source. This correction is performed with respect to the SuperDeblended sources found within the radio islands, that is, within the defined masks. The purpose of this correction is to remove any flux contribution that does not originate from the primary source we are observing, but rather from unrelated background sources detected independently in the SuperDeblended catalog. These background sources, although located within the same region, are physically distinct and must not be attributed to the emission of the source of interest.

This procedure will also be repeated using the Counterpart and PyBDSF catalogs, in order to identify which sources from each catalog fall inside the radio islands (see Figure 17). The full knowledge of all secondary sources present within each region is essential for accurately characterizing the emission properties of the primary sources.

The process began by converting the celestial coordinates (RA and Dec) of all sources into pixel coordinates using the WCS information of the LOFAR image, carefully correcting for the inversion of the Y-axis present in the FITS file. Each white mask is defined by a specific set of pixel coordinates corresponding to the emission area of a detected source. By matching the pixel coordinates of catalog sources against the pixels of each mask, we determined whether a given source is located within a mask.

For every mask, we recorded the number of sources that were found inside and their corresponding IDs. This matching was performed separately for each catalog following exactly the same method, allowing us to establish, for every detected radio source, whether there were nearby secondary sources that could potentially contaminate the flux measurement. This information provided the basis for the subsequent flux correction, where the contribution of these secondary sources was accounted for to recover the intrinsic properties of the primary emission (see Algorithm 5).

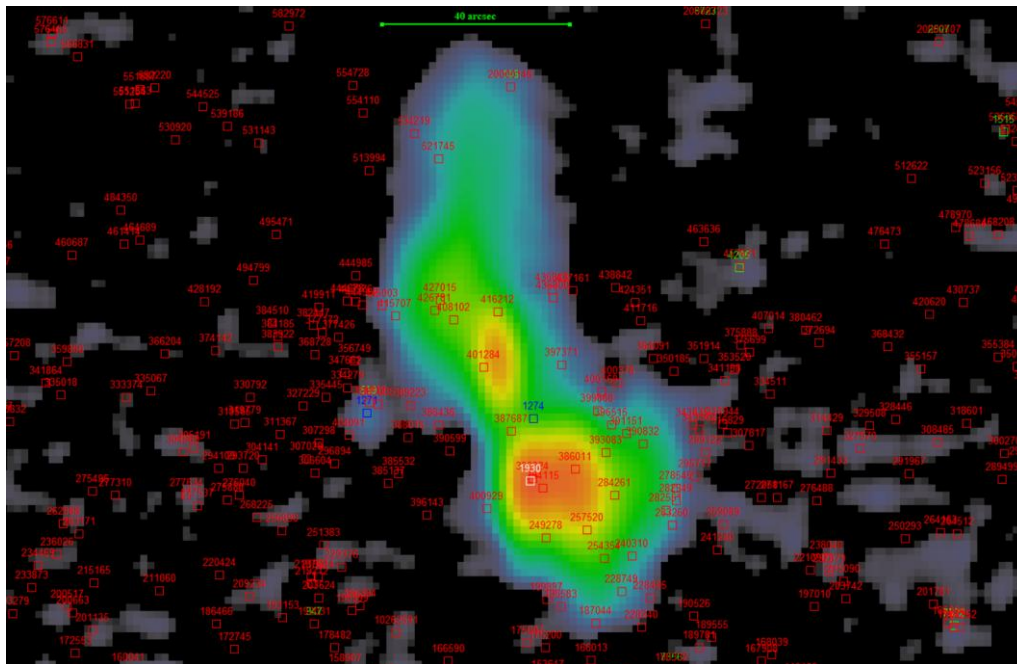


Figure 17: The white box marks the centroid of the **masks** based on the maximum pixel value (ID:1930), the red boxes represent the **SuperDeblended** catalog (multiple sources are found within the radio island, with the most likely central one being the source with ID: 387687, also supported by the counterpart), the green boxes represent the **Counterpart** catalog (ID:10918), and the blue box corresponds to the **PyBDSF** catalog (ID:1274).

After the matching process, we were able to identify the number and IDs of sources from the external catalogs that fall within each of the approximately 2500 masks we created. Naturally, the number of matches varies depending on the size and morphology of each radio source. In some cases, only a single SuperDeblended source is located within the mask, while in others multiple sources may be present. Regardless of the number, the contribution of these additional sources must be considered, as it can significantly affect the total flux density measured within the region. Therefore, it is necessary to remove their contribution in order to obtain a corrected flux density that truly reflects the emission of the primary source.

However, this leads to a fundamental question: should we subtract the flux densities of all the matched SuperDeblended sources within each mask, or should we retain one (or possibly more)? The adopted strategy is to subtract the flux densities of all secondary sources identified within a mask, keeping only the central source, which is considered to represent the true emission of the primary radio galaxy. A second important issue then arises: how do we define which source should be regarded as the central one?

To identify the central source within each region, a two-step strategy was employed. First, we investigated whether any of the catalog sources matched within a mask were associated with a known FR-type radio galaxy. These FR sources originate from a refined subset of the broader Counterpart catalog, specifically selecting sources that have been morphologically classified as Fanaroff-Riley type I or II (FRI/FR II) radio galaxies, or other powerful radio AGN. Thus, the FR source catalog can be considered as a filtered version of the Counterpart catalog, retaining only prominent and physically significant radio sources, based on their large-scale radio structure. The inclusion of FR information is crucial because these sources are known to dominate the radio emission in their environment, and they typically represent the main component of extended radio galaxies.

If an FR source was detected within a region, we selected as the central source the SuperDeblended entry that lay closest to the FR source's position. To determine the proximity, we calculated the angular distance between the positions of the SuperDeblended sources and the FR source using a simple Euclidean approximation:

$$distance = \sqrt{(\Delta RA)^2 + (\Delta DEC)^2} \quad (2.9)$$

where ΔRA and ΔDEC are the differences in right ascension and declination between the two sources, respectively. Although a more sophisticated calculation accounting for spherical geometry could be used, given the very small separations involved (typically a few arcseconds), this linear approximation is entirely sufficient and accurate for our purposes.

In cases where no FR source was identified within a region, the central source was selected based on the original LOFAR detection centroid. The same distance formula was applied, and among all SuperDeblended sources matched within the mask, the one closest to the LOFAR centroid was chosen as the central source.

After identifying the central source, the flux densities of all other SuperDeblended sources within the mask were summed and subtracted from the originally measured LOFAR flux density. This subtraction ensures that the final corrected flux density reflects only the emission of the central component, without contamination from nearby or overlapping secondary sources. To avoid unphysical negative fluxes, a lower limit of zero was imposed on the corrected values.

This procedure was also systematically applied to the masks located within GRGs, where extended structures contain multiple regions with complex source blending. By removing the contribution of secondary deblended sources inside these masks as well, we ensured a consistent and reliable correction of the flux densities across the entire sample, including the particularly extended and complex GRG systems.

Through this procedure, a more accurate measurement of the intrinsic radio emission for each source was obtained, which is essential for reliable further analysis, such as constructing spectral energy distributions and studying the physical properties of the radio galaxies (Algorithm 6).

2.5.2 Error estimation in position and flux density, SNR and RMS measurements

In any astronomical measurement, it is crucial not only to estimate the central value of a quantity, but also to quantify the associated uncertainty. Error estimation refers to the systematic method by which uncertainties in individual measurements are transferred to derived quantities. In our case, the positions (RA and Dec) and the flux densities of the radio sources are the key measurements carrying uncertainties. Accurately propagating these errors through our calculations allows us to assess the reliability of the final results and ensures that comparisons between different sources and physical interpretations are statistically meaningful. Incorporating error estimates is therefore an essential step for a robust and credible scientific analysis (see also Algorithm 7).

2.5.2.1 RA & DEC error estimation & SNR - RMS

To properly account for uncertainties in the positional measurements of our radio sources, we implemented a systematic calculation of the errors associated with the RA and Dec coordinates. In addition, the local signal-to-noise ratio (SNR) and root mean square (RMS) noise were estimated for each source, providing a comprehensive quantification of the reliability of the detections.

The starting point of the procedure was the analysis of the LOFAR radio image from which the sources were extracted. For each radio source, a white mask defining its spatial extent was available, along with the centroid coordinates expressed in pixel units. These centroids were mapped onto the radio image to extract the peak surface brightness corresponding to each source.

To characterize the local noise environment around each source, a rectangular box was defined, extending slightly beyond the mask boundaries (by 10 pixels on each side), and the standard deviation of the pixel values inside this box was computed. This standard deviation was adopted as the local RMS noise for the source. In order to avoid unrealistically small values that could lead to numerical instabilities, a minimum RMS floor was imposed at 10^{-7} Jy/beam.

The signal-to-noise ratio (SNR) for each source was then calculated as:

$$SNR = \frac{\text{Peak Surface Brightness}}{RMS} \quad (2.10)$$

where the peak surface brightness is the value of the radio image at the centroid position.

Given the beam shape of the LOFAR observations (with major and minor axes of 15 and 6 arcseconds, respectively), the positional uncertainties in RA and Dec were estimated using the following relations:

$$\sigma(RA) = \frac{\text{Beam major axis}}{SNR \cdot 3600} \quad (2.11) \quad \sigma(DEC) = \frac{\text{Beam minor axis}}{SNR \cdot 3600} \quad (2.12)$$

where Beam major axis and Beam minor axis are expressed in arcseconds and the factor of 3600 converts the result into degrees. The formulas reflect the general principle that positional accuracy improves with increasing signal-to-noise ratio and is limited by instrumental resolution (beam size).

For sources whose centroid coordinates fell outside the boundaries of the radio image, or for cases where the centroid location corresponded to invalid pixels, the positional uncertainties, SNR, and RMS values were assigned as NaN to indicate undefined measurements.

Finally, the computed values were incorporated into the final catalog, adding columns for the RA and Dec uncertainties (RA_error and DEC_error), as well as the corresponding SNR and RMS measurements for each source.

This procedure ensured that the positional information used in subsequent scientific analysis is accompanied by realistic and robust uncertainty estimates, considering both the instrumental characteristics and the local noise properties of the radio image.

2.5.2.2 Flux Density (Jy) error estimation

To account for the statistical and systematic uncertainties associated with the radio flux density measurements, we applied a conservative 15% relative error on all flux values. This choice is motivated

by the flux scale uncertainty reported in the LOFAR Two-metre Sky Survey DR2, where (Shimwell, et al., 2022) quote a typical calibration uncertainty of approximately 10%. However, this value does not include additional sources of uncertainty related to local noise variations, beam effect, or deblending complexity, particularly in regions of low surface brightness and high source density. Therefore, the adoption of a 15% uncertainty allows for the inclusion of both the instrumental and observational scatter that is not fully captured by the catalogued errors, providing a more robust propagation of uncertainties throughout our analysis.

2.5.3 Matching process cat masks & deblended, counterpart and pybdsf

To generate the final LOFAR-based catalogue of radio sources, we developed and implemented a robust and hierarchical matching framework (Algorithm 9). This process aimed to associate each of our internally defined LOFAR-based regions (masks) with the most physically meaningful and accurate entries from three key external catalogues:

- the VLA 3GHz **Counterpart catalogue**,
- the **SuperDeblended LOFAR-FIR-radio** catalogue, and
- the **PyBDSF** source extraction catalogue.

This matching was carried out on a **region-by-region basis**, using their sky positions (RA, Dec), and was guided by a clearly defined set of rules, priorities, and fallback strategies to ensure the highest confidence in source association. The entire methodology was supported by **automated logic** and **visual inspection validation**.

Definition of Primary Regions via Masks

The process began by identifying individual source regions (masks) from the LOFAR map using image-based segmentation techniques. Each mask corresponds to a distinct radio-emitting region and contains metadata such as:

- pixel-level coordinates of the source,
- its geometric centroid (based on peak flux), and
- its integrated flux within the region.

These masks served as spatial anchors onto which we projected all external catalogue entries.

Step 1: Matching with the Counterpart Catalogue

The first external catalogue matched against the masks was the **VLA 3GHz Counterpart catalogue** (Smolcic, et al., 2017). For each region:

- **If exactly one** counterpart entry was located inside the mask boundary, it was directly matched with that region.
- If **multiple** entries were present, we selected the one **closest to the region's centroid**, using the **Euclidean distance formula**:

$$D = \sqrt{(RA(catalogue) - RA(mask))^2 + (DEC(catalogue) - DEC(mask))^2} \quad (2.13)$$

This ensured that the selected counterpart was most representative of the true center of the emission.

- If the matched counterpart source was included in the **FR (Fanaroff–Riley)** radio galaxy catalogue (Vardoulaki et al. 2023), we **overrode all other matches** and performed a **forced match** to that FR entry. In such cases, we adopted the **RA and Dec of the FR source** as the true position of the central galaxy, as these entries have undergone dedicated classification and curation.

This step was critical for ensuring that **core-dominated AGNs and SFGs** were matched using the most reliable coordinates available. A prominent example is the forced match to ID 10913 (counterpart ID), manually validated through radio morphology.

Step 2: Matching with the SuperDeblended Catalogue

Once counterpart matching was completed, we proceeded to associate each region with entries from the **SuperDeblended LOFAR-FIR-radio catalogue**.

- If exactly one source was present inside the mask, it was matched directly.
- If multiple sources were inside the same mask, we computed distances **not relative to the mask centroid**, but rather relative to the **RA/Dec of the previously matched Counterpart source**. This choice reflects the higher astrometric precision of the VLA 3GHz catalogue and its proximity to the actual galactic center.

If **NO Counterpart** had been matched, then we fell back to using the original mask-derived coordinates of the region.

- The closest SuperDeblended source to the reference RA/Dec was selected for inclusion in the final table.

This logic ensured consistency: the same anchor point was used when resolving ambiguous matches in both the counterpart and deblended cases.

Additionally, for **specific cases of complex morphology or known association**, we introduced **hard-coded forced matches**. For instance, LOFAR region ID 377 was linked explicitly to SuperDeblended ID 20002251, bypassing all counterpart associations due to prior classification confidence. The same was done for ID 625.

Step 3: Matching with the PyBDSF Catalogue

Matching to the **PyBDSF catalogue** was treated as a supplementary step. PyBDSF contains source components extracted directly from the 3GHz VLA map using a source-finding algorithm.

For regions that already had a matched Counterpart and/or SuperDeblended entry, we attempted to match the closest PyBDSF entry **within the same mask**.

- The reference coordinates for distance calculation were again taken from the matched CP/DB source, if available.
- Otherwise, the LOFAR centroid was used.
- In rare cases where multiple PyBDSF entries were found, the closest one to the reference position was selected.

Some regions, such as IDs 641 and 919, were **explicitly excluded** from PyBDSF matching due to confusion or poor quality in the radio morphology.

Matching Priority, Integration and Filtering

The three-stage matching logic was carefully designed with a **prioritization hierarchy**:

1. **Counterpart catalogue** was always evaluated first, with FR classification taking precedence.
2. **Deblended catalogue** relied on the CP/FR reference coordinates for robust matching.
3. **PyBDSF** matching was conditional and used only to supplement the above matches.

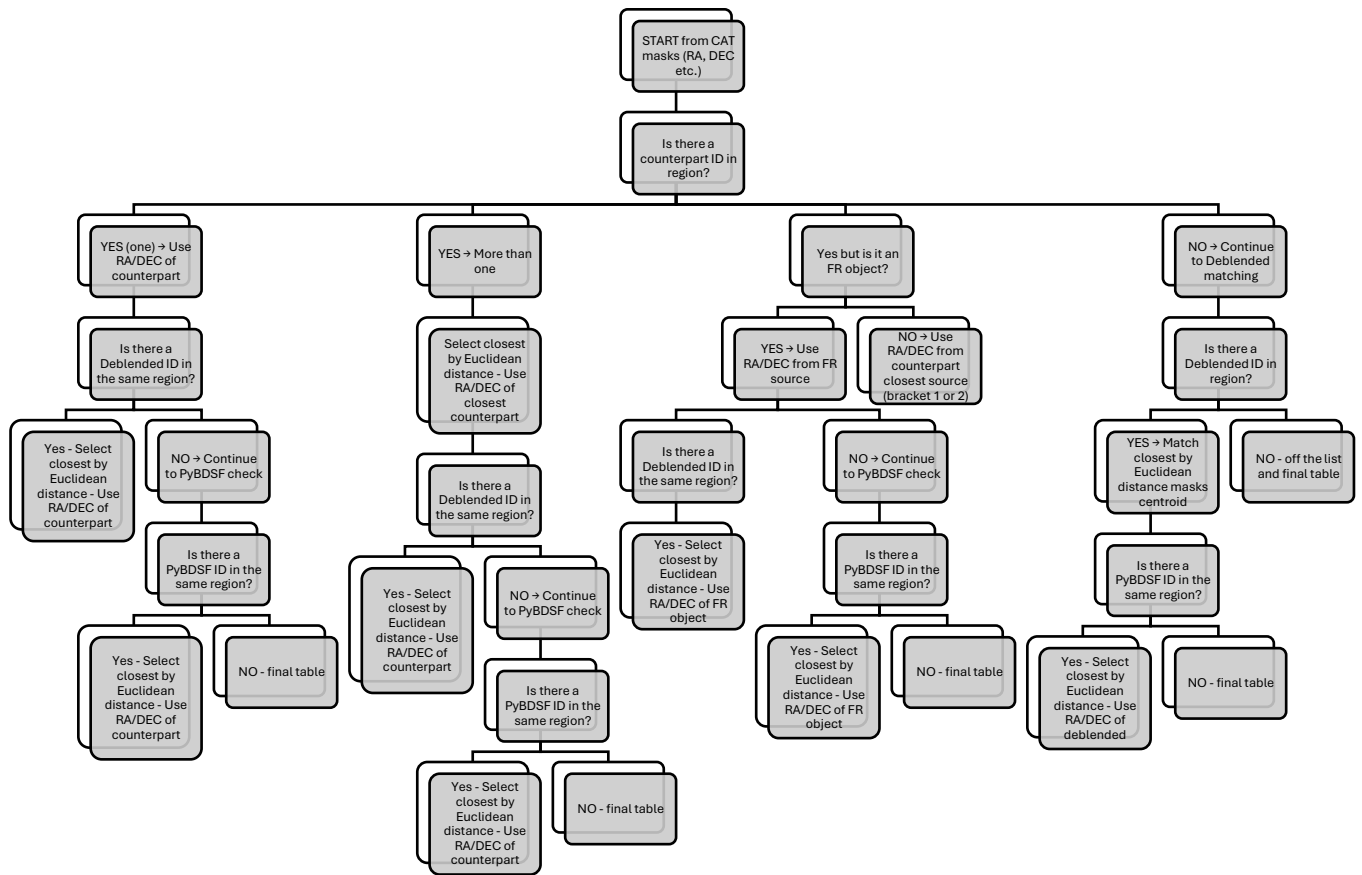
Once the associations were established, we constructed the final catalogue by:

- merging all relevant fields from each matched catalogue into the LOFAR base table,
- converting flux units from μJy to Jy for consistency,
- removing entries with invalid or negative flux values,
- and applying a series of **manual source-merging operations** (e.g., combining entries representing different components of GRG lobes into a single source by summing fluxes and pixels).

A final filtering step removed a list of visually inspected and confirmed spurious or duplicate IDs.

Visual Validation and Manual Refinement

While the bulk of the matching was automated, **visual inspection in DS9** played a key role in verifying edge cases, extended sources, and ambiguous matches. This was especially important for **large radio galaxies** (e.g. GRGs) with complex morphologies, for which the automated match would not suffice. In such cases, we implemented manual overrides and source merges based on visual assessment.



2.6 Final Catalogue/dataset

The final catalogue is the outcome of the cross-matching strategy described in the previous sections, which systematically linked LOFAR-detected radio sources with their counterparts in multi-wavelength and radio deblended datasets. Following the application of this matching logic, filtering rules, and forced assignments where necessary, we constructed a comprehensive table that consolidates the most reliable information available for each radio-emitting source in our field.

Each row in the final catalogue corresponds to a single radio source detected in the LOFAR image. The source has been spatially cross-identified, wherever possible, with entries from three complementary catalogues: the **VLA 3GHz counterpart catalogue**, the **SuperDeblend LOFAR-FIR-radio catalogue**, and the **PyBDSF source list**. The resulting table integrates multi-wavelength data, deblended flux measurements, morphological information, and astrophysical classifications, providing a highly detailed and versatile dataset for further analysis.

The catalogue includes a wide array of columns, organized into thematic groups:

- **LOFAR radio source properties:**

- o Unique source ID, centroid position (RA, Dec), and associated positional errors (RA_error, DEC_error), derived from the pixel coordinates of the peak within each mask.
- o Corrected total flux density at 144 MHz (FluxJy_LOFAR_masks), expressed in Jy, after subtracting contributions from overlapping sources within the same mask.
- o Associated uncertainty on the corrected flux (error_FluxJy_LOFAR_masks), assuming a conservative 15% fractional error.
- o The source area in number of pixels (NumPixels), providing a direct estimate of angular extent within the mask.
- o Deconvolved angular area estimate (deconvolved_area_pybdsf_cat), derived from the PyBDSF catalog major/minor axes and corrected for beam convolution, where available ($A = (\pi / (4 \times \ln 2)) \times \theta_{\text{maj}} \times \theta_{\text{min}}$ where both axes are in arcseconds).
- o Additional source resolved flags and local background estimates (e.g., SNR, RMS).
- **Counterpart catalogue entries:**
 - o Accurate position from VLA data, redshift (Z_BEST), infrared star formation rates, and total IR luminosity (L_TIR_SF).
 - o AGN diagnostics across multiple wavelengths, including X-ray (Xray_AGN), MIR (MIR_AGN), and SED-based classifications (SED_AGN, HLAGN, MLAGN, Quiescent_MLAGN, etc.).
 - o SFG flags, including Clean_SFG, useful for isolating pure SFG populations.
- **Deblended FIR-radio measurements:**
 - o Multi-frequency flux densities at wavelengths such as 10 cm, 20 cm, 23 cm, 50 cm, 90 cm, and 2 meters, all corrected for blending (e.g., DB_f10cm, DB_f2m_lofar).
 - o Corresponding uncertainties, redshift (z_spectroscopic and z_photometric) and IR-derived properties like photometric redshifts (z_FIR) and IR-based SFRs (SFR_IR).
- **PyBDSF radio morphological data:**
 - o Parameters such as major and minor axes, peak and total fluxes, beam position angles, and deconvolved source sizes, contributing to the morphological classification of extended sources.

The catalogue is structured to facilitate both **statistical population studies** and **individual source analysis**. For example, one can directly isolate clean star-forming galaxies with known redshifts, extract sources with radio excess indicative of AGN activity, or study extended structures via PyBDSF morphology. Furthermore, since the flux densities have been corrected through the mask-based subtraction of secondary components, the final values represent the most physically meaningful measurements for intrinsic radio emission.

In summary, this final catalogue offers a multi-dimensional and scientifically robust view of the LOFAR-detected radio sources in the COSMOS field. It is designed to support further exploration of galaxy evolution, AGN feedback, star formation rates, and large-scale radio structures, serving as the foundation for all subsequent analyses in this thesis.

2.7 Data Visualization and Classification

The use of diagrams and histograms is a fundamental tool in the visual representation and interpretation of astronomical data. By constructing histograms and scatter plots for key catalog

parameters, we are able to directly observe the distribution, range, and structure of our sample, as well as identify outliers, selection effects, and potential subpopulations. Such visualizations facilitate a deeper understanding of the underlying properties of the radio sources, allow the comparison of different source classes, and help reveal trends and correlations that might otherwise remain hidden in tabular data. In the following sections, we present a series of scatter plots for crucial parameters (e.g., flux density, SNR, number of pixels), interpreting their astrophysical significance and discussing the implications for the composition and completeness of the sample.

Total sources	1085
Matched with Counterpart	861
Matched with Deblended	1065
Matched with PyBDSF	861
GRG count	2
NumPixels min	1
NumPixels max	2693
NumPixels mean	86,43686636
NumPixels std	190,6683403
SNR min	0,036190776
SNR max	12,14124411
SNR mean	4,787110716
SNR std	1,508037616
RMS min	9,95679E-05
RMS max	0,044500098
RMS mean	0,00067416
RMS std	0,0022681
Area min	20,21975824
Area max	2362,887311
Area mean	188,5606413
Area std	152,478133
Resolved	585
Unresolved	498

Table 1: Summary of the main properties for all 1085 sources, including cross-matches with Counterpart, Deblended, and PyBDSF catalogs, the number of GRGs, and basic statistics (min, max, mean, std) for NumPixels, SNR, RMS, and deconvolved area. The last rows list the number of resolved and unresolved sources in the sample.

Based on the 861 sources from the counterpart catalog and the information provided for each galaxy type, we constructed several plots to better understand our sample and to gain a more comprehensive overview of our dataset. These plots are purely statistical, with colors corresponding to the classification of each galaxy as indicated in the counterpart catalog. Specifically, the sources are flagged as Clean SFG, Moderate-Luminosity AGN (MLAGN), and High-Luminosity AGN (HLAGN), as well as all possible combinations of these flags, including radio-excess sources. This categorization enables us to visually distinguish the various physical processes contributing to the observed radio emission. By examining the distribution and relative abundance of each class—individually and in combination—we can assess the dominant populations within our sample and identify sources where mixed mechanisms (such as simultaneous star formation and AGN activity) may play a significant role. These visualizations thus provide an essential first step in interpreting the physical nature and diversity of the radio sources in our field.

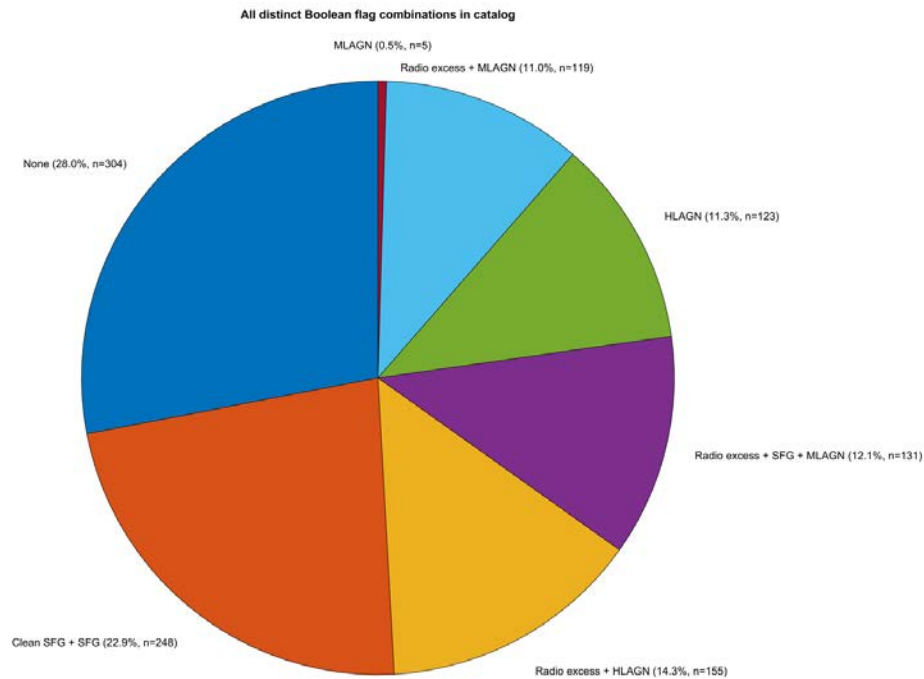


Figure 18: This pie chart illustrates the distribution of sources in the final catalog according to the distinct combinations of Boolean flag categories. Each color represents a different group of sources as classified by the flags: *dark blue* corresponds to sources with no specific category (“None”), *red* to clean star-forming galaxies together with SFG (“Clean SFG + SFG”), *yellow* to sources with both radio excess and HLAGN, *purple* to sources with radio excess, SFG, and MLAGN, *green* to HLAGN, *light blue* to sources with radio excess and MLAGN, and *maroon* (dark red) to sources classified exclusively as MLAGN. The percentage and the absolute number of sources are displayed next to each sector of the chart.

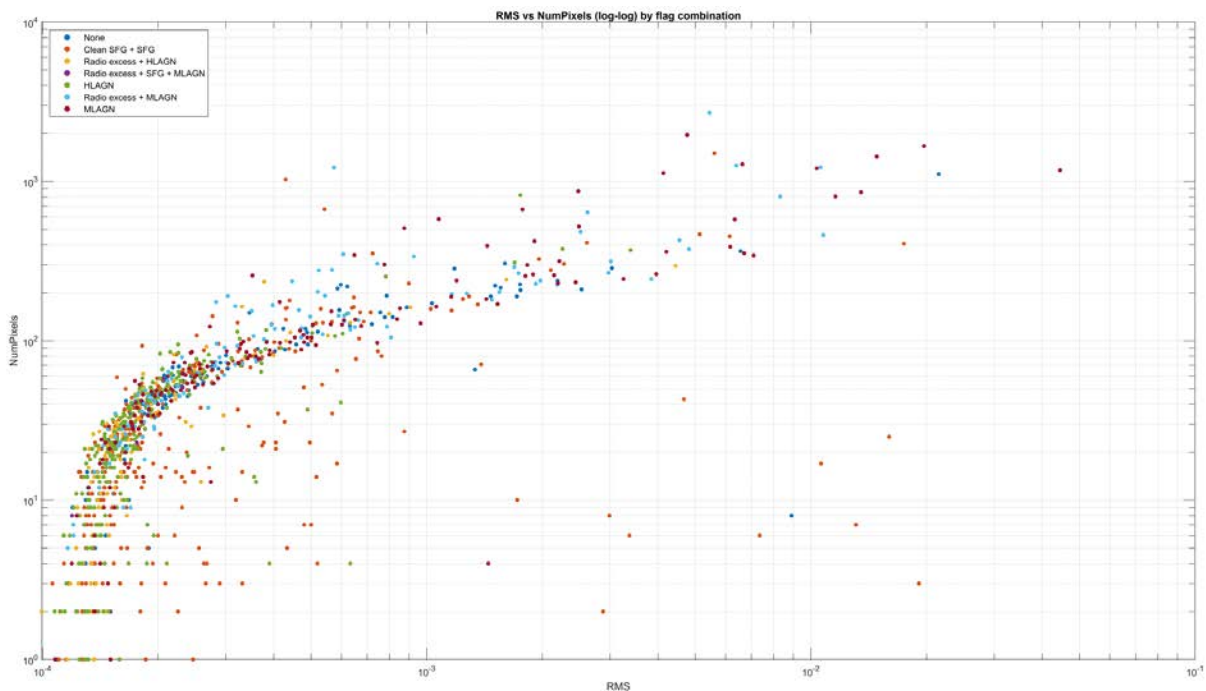


Figure 19: Distribution of the number of pixels (NumPixels) as a function of sensitivity (RMS) for all sources in the catalog, plotted on a log-log scale and separated by Boolean flag combination. Different source categories are shown in different colors, as indicated in the legend. Sources with lower RMS values (higher sensitivity) tend to have fewer pixels, while sources with higher RMS exhibit a broader range of sizes.

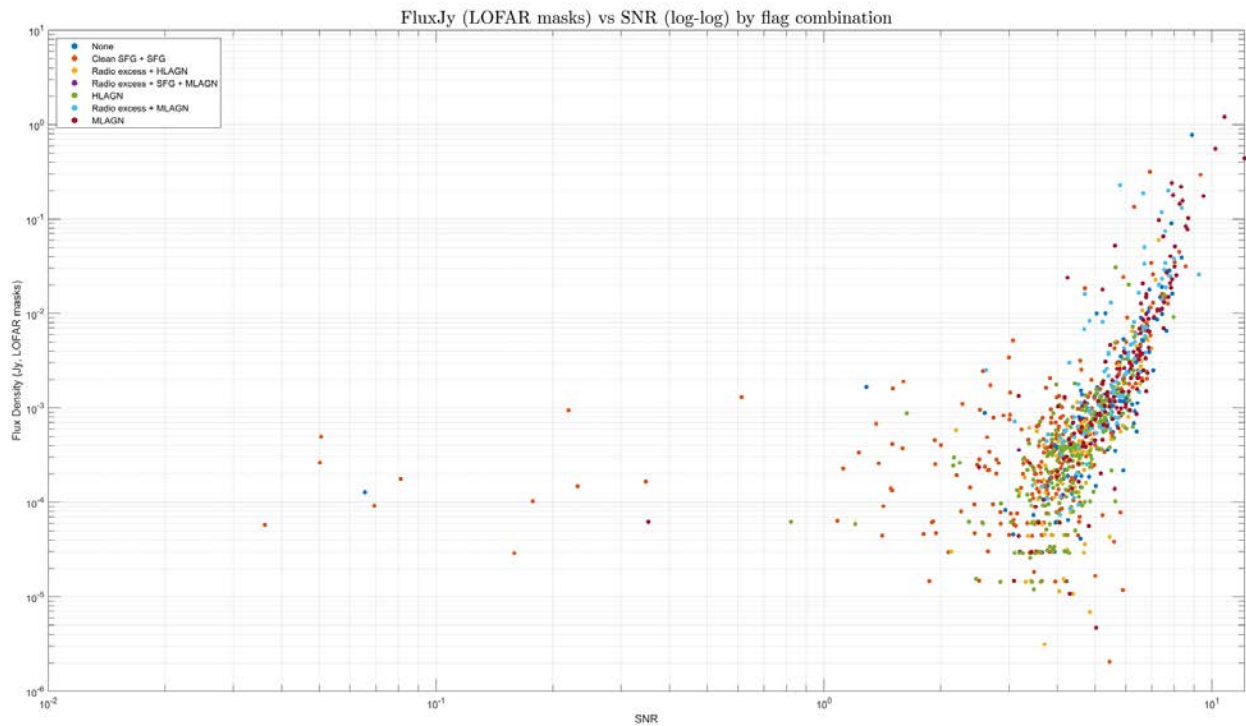


Figure 20: Scatter plot of LOFAR mask-based flux density (Jy) versus signal-to-noise ratio (SNR) for all sources in the catalog, color-coded by Boolean flag combination. The diagram highlights the dependence of measured flux density on SNR across different radio source categories.

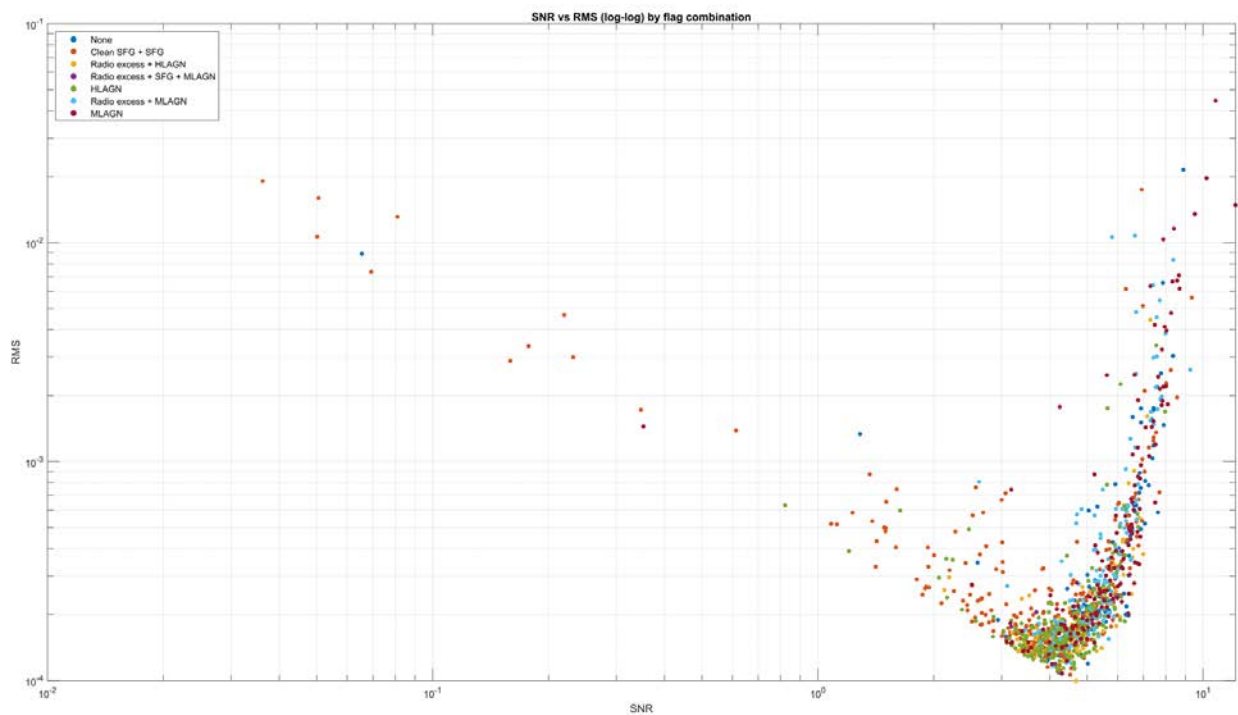


Figure 21: Scatter plot of SNR versus RMS (root mean square noise) for all sources, with both axes in logarithmic scale and points color-coded by Boolean flag combination. The plot visualizes the relationship between measured noise and signal quality across different source categories in the LOFAR masks catalog.

Chapter 3: Analysis & Results

3.1 Description of the Sample

In order to study the radio spectral energy distributions (RSEDs) of extragalactic sources with precision and physical relevance, it is essential to define a carefully curated sample that reflects the scientific objectives of this work. Given the wide variety of radio-emitting sources—ranging from star-forming galaxies to powerful radio-loud AGN—a uniform treatment of the entire catalogue would risk obscuring the physical processes specific to each population.

Therefore, in this chapter, we construct and describe two distinct subsamples drawn from the final catalogue presented in Chapter 2. Each subsample is selected based on clearly defined criteria that reflect the nature of the sources under consideration and the requirements of radio SED analysis. The two main populations we investigate are:

1. **Active Galactic Nuclei**, classified as RxAGN, which are characterized by the presence of compact or extended non-thermal emission driven by accreting SMBH, and whose radio emission is substantially greater than what would be produced by star formation alone, and
2. **Highly star-forming galaxies (SFGs)**, with HSFR, whose radio emission primarily traces intense recent star formation activity via synchrotron radiation from supernova remnants and thermal bremsstrahlung.

The analysis of each subsample is treated separately in the sections that follow, in order to account for the different physical mechanisms dominating the radio spectra in each case. For both samples, we explore the selection logic, filtering criteria, and final composition of the datasets used for SED fitting and interpretation.

3.1.1 Subsample for AGNs

To construct a suitable subsample of radio-excess AGNs for our analysis, we followed a multi-step cross-matching approach inspired by the methodology of Tisanić et al. (2020), who presented a sample of 744 AGNs selected from multi-frequency radio data in the COSMOS field. While the specific identifiers of the sources included in their AGN classification were not available, we were provided with the full flux catalogues at 3 GHz, 1.4 GHz, 610 MHz, and 325 MHz. This allowed us to follow a comparable data integration methodology, though not to exactly reproduce their sample.

Our process began by internally matching the VLA-COSMOS 1.4 GHz (VLA1.4JP) and 3 GHz (VLA3LP) catalogues using the unique ID_VLA field. This step alone yielded more than 2,000 sources detected at both frequencies. We then matched this intermediate list with the GMRT 610 MHz catalogue, again using ID_VLA, and retained all resulting sources. This produced a working catalogue of 800 sources that had flux measurements at all three frequencies: 1.4 GHz, 3 GHz, and 610 MHz.

Subsequently, we matched this 800-source catalogue with the GMRT 325 MHz catalogue using ID_VLA. For sources not found in the 325 MHz dataset, the corresponding flux values were left empty to allow upper limits to be assigned later using local RMS measurements.

We then matched the combined multi-frequency catalogue with our LOFAR-based source list at 144 MHz. Since the LOFAR catalogue was already associated with the VLA-COSMOS counterpart catalogue (see Chapter 2), the cross-match was again performed via the common ID_VLA. For unmatched sources at this frequency, upper limits were similarly assigned using local RMS information ($5 \times \text{RMS}$).

From the 800 sources in the assembled multi-band catalogue, 447 were flagged as radio-excess AGNs according to the `Radio_excess == true` criterion defined in the counterpart catalogue. To further constrain the sample and focus on radio-luminous AGNs in a physically meaningful redshift range, we applied two additional filters: (a) photometric redshifts in the range $0 \leq z \leq 4$, and (b) logarithmic rest-frame radio luminosities at 1.4 GHz in the range $24 \leq \log L_{1.4 \text{ GHz}} \leq 26$. These luminosity bounds are chosen to focus on powerful radio AGNs while excluding both low-luminosity star-forming galaxies and extremely rare, high-luminosity outliers. After applying these filters, 314 sources remained.

Among these, 228 were successfully matched with the full multi-frequency catalogue (including 144 MHz) and retained for our final AGN sample (for the remaining 86 sources, upper limits were calculated). This subset was used for constructing the rest-frame RSEDs and for the subsequent spectral fitting analysis.

3.1.2 Subsample for SFGs

In parallel with the AGN sample, we constructed a second, well-defined subsample consisting of galaxies whose radio emission is primarily driven by star formation processes. These SFGs act as a benchmark sample for investigating non-thermal synchrotron emission produced by cosmic-ray electrons accelerated in supernova remnants, in the absence of contamination from active galactic nuclei. Notably, due to sensitivity limitations at lower frequencies, only 17% and 9% of the original HSFG sample were detected at 610 MHz and 325 MHz, respectively, as illustrated in Figure 22.

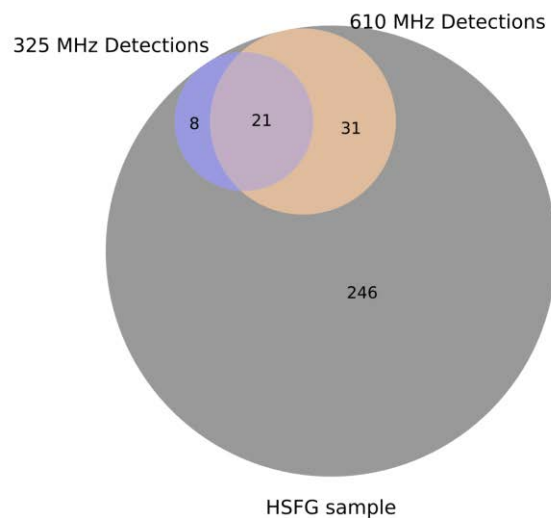


Figure 22: Detection statistics of HSFG sources in the GMRT catalogues. Sources not detected in the GMRT images are shown in gray, while detections at 325 MHz and 610 MHz are illustrated in blue and orange, respectively. (Tisanić, et al., 2019)

To this end, we utilized the original sample of 306 highly star-forming galaxies presented in Tisanić et al. (2020), constructed by selecting galaxies from the COSMOS field with robust detections at both 1.4 GHz and 3 GHz, and applying a conservative lower limit on the infrared-derived star formation rate ($\text{SFR} > 100 \text{ M}_{\odot} \text{ yr}^{-1}$). This selection ensures completeness over the redshift range $0.3 \lesssim z \lesssim 4$, while AGN were removed through a combination of multi-wavelength rejection criteria.

From the initial catalogue of 218 Clean SFGs identified in our multifrequency masks dataset, a subset of 54 sources overlapped with the 306 HSFGs of Tisanić et al. (2020) and were matched to our LOFAR-based source list at 144 MHz. All matched sources were explicitly flagged as `Clean_SFG ==`

true, signifying that they satisfied all AGN rejection diagnostics and are fully consistent with star formation-driven radio emission.

To increase the statistical weight of the sample, we additionally included 252 unmatched HSFGs with upper limits at 144 MHz (5*RMS). This yielded a final sample of 306 LOFAR-based SFGs contributing to our RSED analysis. This extended sample serves as the foundation for our investigation of the radio continuum properties of star-forming galaxies at low frequencies, complementing existing constraints from higher-frequency VLA data.

3.2 Normalization Procedures for Radio SED Construction

3.2.1 General Normalization Procedure for both Subsamples

In this section, we outline the general methodology employed to normalize the radio flux density measurements for all sources in our study, irrespective of their physical classification. The goal is to establish a unified framework that allows for a direct, unbiased comparison of RSEDs across heterogeneous extragalactic populations—namely, AGNs and SFGs. The normalization procedure transforms observed flux densities into dimensionless, rest-frame quantities, and addresses the treatment of censored data (upper limits), binning strategies, and the statistical modeling of the resulting distributions. Our approach closely follows the normalization and statistical framework described in Tisanić et al. (2019, 2020), ensuring full consistency with the established methodology for this analysis. Specific choices or methodological variations pertinent to each subsample (such as the selection of reference frequency) are detailed in Sections 3.2.2 (AGN subsample) and 3.2.3 (SFG subsample).

I. Rest-Frame Frequency Calculation

For each source, rest-frame frequencies were calculated using:

$$\nu_{(rest)} = \nu_{(obs)} \times (1 + z) \quad (3.1)$$

where $\nu_{(obs)}$ is the observed frequency in the observer's frame, z is the redshift of the source and $\nu_{(rest)}$ is the corresponding rest-frame frequency, i.e., the intrinsic emission frequency in the galaxy's own frame of reference. This transformation is essential for constructing rest-frame SEDs, as it enables the interpretation of radio flux measurements in terms of the physical emission processes within each source, independent of their cosmological redshift. By shifting the observed frequencies to the rest frame, we ensure that sources across different redshifts can be directly compared on equal physical grounds. This provides a consistent foundation for modeling and analyzing the radio spectral behaviour of AGNs throughout cosmic time.

II. Interpolation and Normalization to a Reference Frequency

For each source, flux densities were normalized to a common rest-frame frequency $\nu_{(ref)}$, even when direct measurements at this frequency were unavailable. To estimate the flux at $\nu_{(ref)}$, we performed a log–log linear interpolation between the two nearest available flux measurements in frequency space:

- The spectral index (α) is computed as:

$$\alpha = \frac{\log F(3GHz) - \log F(1.4GHz)}{\log \nu_{(rest, 3.0)} - \log \nu_{(rest, 1.4)}} \quad (3.2)$$

- The flux at the reference frequency is then:

$$\log F(v(\text{ref})) = a * \log(v(\text{ref})) + b \quad (3.3)$$

- where the intercept b is defined as:

$$b = \log F(3\text{GHz}) - a * \log v(\text{rest}, 3.0) \quad (3.4)$$

- The resulting normalization flux is:

$$F_{\text{norm}} = 10^{\log F v(\text{ref})} \quad (3.5)$$

III. Flux Density Normalization

Each measured flux at rest-frame frequency ν_i was divided by the normalization flux:

$$F_i' = \frac{F_i}{F_{\text{norm}}} \quad (3.6)$$

and subsequently transformed to logarithmic space:

$$\log_{10}(F_i') \quad (3.7)$$

This procedure yields dimensionless, log-normalized fluxes, which form the basis for subsequent stacking and statistical analysis.

IV. Filtering and Quality Control

Only sources with finite, positive flux densities at the relevant frequencies were retained. Further, normalized fluxes that were unphysically low (e.g., dominated by noise or extreme outliers) were removed to ensure the reliability of the final sample.

V. Handling Upper Limits: Survival Analysis

A substantial number of sources, particularly at low frequencies, are undetected and thus provide only upper limits (censored data). To ensure an unbiased statistical treatment of both detections and upper limits in the construction of average SEDs, we applied survival analysis techniques following the methodology of Feigelson & Nelson (1985) and Tisanić et al. (2019, 2020):

Kaplan–Meier Estimator

- As a non-parametric maximum likelihood technique, the Kaplan–Meier estimator is employed to determine the distribution function of datasets that include censored observations.
- It enables the calculation of unbiased means and variances in bins containing both detections and upper limits.
- Assumptions: Censored (upper limit) and detected values are independent and drawn from the same underlying distribution. This is justified here since normalization is performed to a common reference frequency, ensuring comparability across sources.
- For each frequency bin, the estimator constructs the empirical survival function and derives statistics (such as mean normalized flux) that fully incorporate the information from censored data.

Weibull Distribution Fit

- In bins with a significant fraction of censored data, the survival function was also fit with a **Weibull distribution**. This distribution is flexible and allows the survival function to be analytically linearized, facilitating the calculation of means and standard deviations for the log-normalized fluxes.
- This approach improves the stability of statistical estimates when the number of upper limits is high, as it avoids bias that would arise from simply ignoring non-detections.
- The use of the Weibull fit follows the justification and methodology described by Tisanić et al. (2020).

VI. Binning and Statistical Aggregation

Normalized log-fluxes (detections and upper limits) were grouped into bins in log(rest-frame frequency) space (typically 20 bins, empirically chosen to balance sample size and frequency resolution).

- **For bins without censored data:** The mean and standard deviation were calculated directly from detections.
- **For bins with censored data:** The Kaplan–Meier estimator or Weibull distribution was used to derive unbiased means and uncertainties.
- This dual approach ensures that the final binned SED is statistically robust across the full frequency range, even where upper limits dominate.

VII. Parametric Modeling and Regression

The binned, log-normalized fluxes (including both detections and upper limits, appropriately modeled via survival analysis) serve as the input for parametric model fitting.

- **Model fitting** (such as broken power-law models) was performed in log–log space, with parameter estimation via Markov Chain Monte Carlo (MCMC) techniques.
- For best-fit parameter estimation, we adopted **orthogonal distance regression (ODR)** rather than ordinary least-squares, as ODR explicitly accounts for uncertainties in both the independent (log-frequency) and dependent (log-flux) axes.
 - ODR propagates measurement errors in both axes through the fit, yielding more robust and accurate confidence intervals for model parameters.
 - The covariance matrix from the ODR fit was used to estimate parameter uncertainties, as in Tisanić et al. (2020).

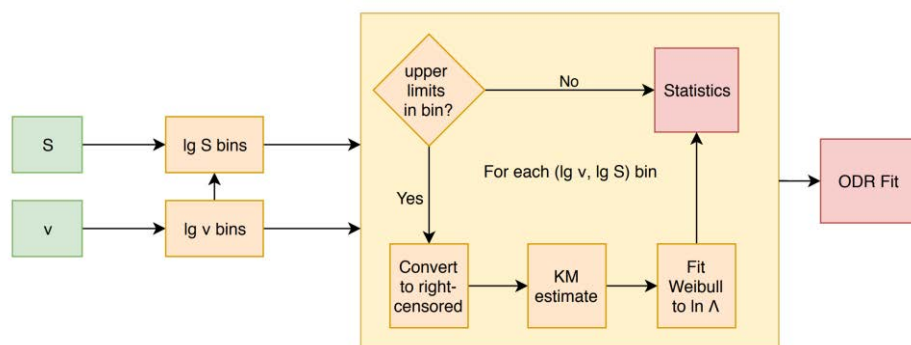


Figure 23: Flowchart summarizing the radio SED normalization and fitting (Tisanić, et al., 2019).

3.2.2 AGN Subsample: Normalization Procedure (Specific Implementation)

For the **radio-excess AGN** subsample, the normalization methodology described in Section 3.2.1 was applied with the following specific choices:

❖ **Reference Frequency Selection:**

The normalization of AGN flux densities was anchored at a rest-frame frequency of **4.0 GHz**, chosen for its physical significance as the typical spectral break frequency in AGN radio SEDs (see Tisanić et al. 2020) and for its location within the range robustly covered by our data. Where direct measurements at this frequency were unavailable, normalization was performed via log–log linear interpolation between the nearest measured bands (typically 1.4 and 3.0 GHz), as in Eq. (3.2)–(3.5).

❖ **Quality Control and Filtering:**

Following the general approach of Section 3.2.1, only AGN sources with valid, positive, and physically meaningful normalized fluxes were retained for further analysis. Cases where the interpolation at 4.0 GHz produced unphysical or non-positive values were excluded to ensure the integrity of the final AGN sample.

❖ **Statistical Treatment:**

All subsequent steps, including the treatment of upper limits, survival analysis (Kaplan–Meier estimator and Weibull fits), binning, and parametric modeling, were performed exactly as described in Section 3.2.1. No further methodological modifications specific to AGNs were required.

This implementation ensures the construction of a consistent, statistically robust, rest-frame radio SED for the AGN population, directly comparable to the SFG subsample and to literature results.

3.2.3 SFG Subsample: Normalization Procedure (Specific Implementation)

The normalization procedure for the SFG subsample follows the general methodology outlined in Section 3.2.1, ensuring consistency in the treatment of both SFGs and AGNs throughout the analysis. Here, we focus on the aspects that are either specific to SFGs or differ from the AGN implementation.

❖ **Normalization Reference Frequency**

For SFGs, all observed rest-frame flux densities are normalized to a reference frequency of **5.7 GHz**. This value represents the logarithmic mean of the 1.4 and 3.0 GHz rest-frame frequencies, as adopted by Tisanić et al. (2019), and is chosen to minimize extrapolation uncertainties and ensure the normalization occurs within a frequency range of high data reliability for the SFG sample. The normalization flux is computed for each source via log–log linear interpolation between the available rest-frame fluxes at 1.4 GHz and 3.0 GHz. Only sources with positive, finite fluxes at both frequencies are retained for normalization. This approach ensures that sources across a wide range of redshifts contribute equally to the construction of the average SED and eliminates potential biases due to uneven spectral coverage.

❖ **Inclusion of Upper Limits (Censored Data)**

Compared to the AGN sample, the SFG dataset contains a larger fraction of upper limits, particularly at lower frequencies, resulting from observational sensitivity limitations. To achieve an unbiased statistical representation of the SFG population, we incorporate censored data using survival analysis techniques, specifically the Kaplan–Meier estimator and, where appropriate, a Weibull

distribution fit (Feigelson & Nelson, 1985; Tisanić et al., 2019). As discussed in Section 3.2.1, this approach allows for a robust estimate of mean normalized fluxes and confidence intervals in each frequency bin, even when a significant portion of data points are upper limits. The validity of this method is further supported by the normalization of all sources to a common reference frequency, which ensures the mutual independence of detections and upper limits (Tisanić et al., 2020).

❖ Binning and Statistical Aggregation

The normalized logarithmic fluxes are grouped into 20 logarithmically spaced bins in rest-frame frequency, matching the procedure described for AGNs (Section 3.2.1). For each bin, the mean and standard deviation of the normalized log-fluxes are derived directly from detections or, when censored data are present, via the Kaplan–Meier or Weibull survival analysis as appropriate. The number of bins (20) is chosen empirically to balance frequency resolution with sufficient sample size in each bin, following the conventions of Tisanić et al. (2020). This dual approach ensures the statistical robustness of the average SED across the entire frequency range.

❖ Parametric Modeling and Final Output

The binned, log-normalized fluxes are used for parametric modeling of the average SED. We fit a broken power-law model to the binned mean fluxes, fixing the break frequency at 4.3 GHz rest-frame in line with prior studies (Tisanić et al., 2019; 2020). The two spectral slopes and normalization constant are estimated via MCMC techniques, while the fit itself is performed using ODR, which properly accounts for uncertainties in both axes. This ensures a reliable determination of the best-fit model parameters and their confidence intervals.

The final output of this procedure for the SFG subsample consists of:

- Individual normalized fluxes for each galaxy across all available frequency bands,
- Log-frequency binned mean values with associated confidence intervals,
- A best-fit broken power-law model with 1σ error margins.

This output provides the basis for quantitative investigation of the average spectral properties of star-forming galaxies and enables direct comparison with the corresponding results for AGNs, as presented in subsequent sections.

3.3 Data Representation

The RSED diagrams that follow aim primarily to compare our results with those of previous studies. Notably, these plots include newly derived LOFAR-based measurements, which were computed specifically as part of this work. In the following analysis, we assess whether the inclusion of our data affects the spectral slopes previously established in the literature, particularly in the context of a broken power-law model.

To ensure that these comparisons remain meaningful and consistent, we adopted the same normalization strategy as that employed in the studies by Tisanić et al. (2019) and Tisanić et al. (2020). The details of this procedure were outlined in the previous section.

3.3.1 AGN - RSED

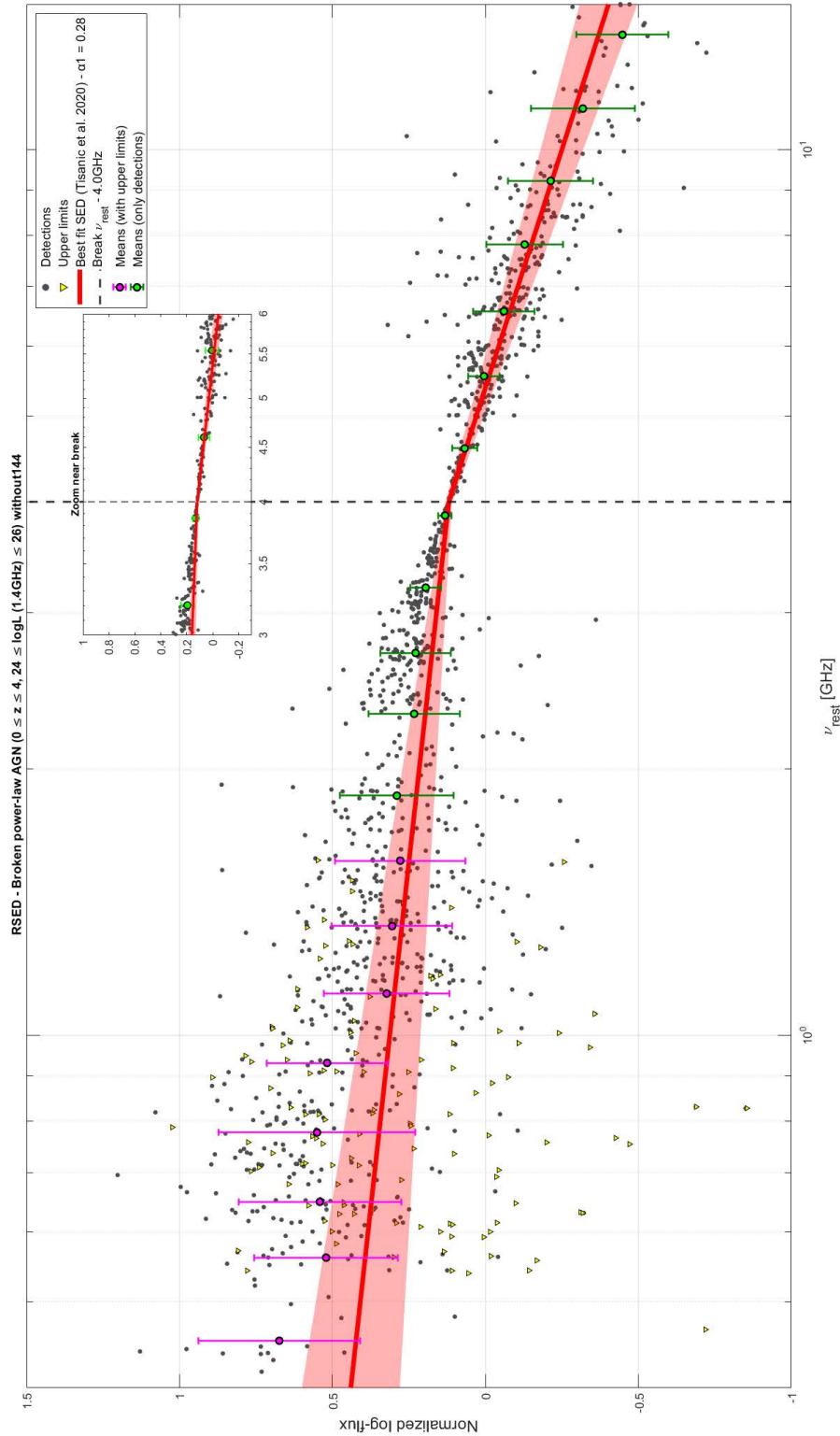


Figure 24: Reproduction of the average radio SED of radio-excess AGN. This plot presents the normalized log-flux densities as a function of rest-frame frequency for a sample of radio-excess AGN with redshifts $0 \leq z \leq 4$ and radio luminosities $24 \leq \log L_{1.4\text{GHz}} \leq 26$. Gray dots indicate individual detections, while yellow downward-pointing triangles represent upper limits. Magenta circles show the mean flux values computed using Kaplan–Meier survival analysis in bins that include upper limits, and green circles represent means from bins containing only detections. The solid red line corresponds to the best-fitting broken power-law (BPL) model from Tisanić et al. (2020), featuring a fixed spectral break at $\nu_{\text{rest}}=4.0$. The shaded red area illustrates the 1σ confidence interval of the fit. The inset panel provides a zoomed-in view around the break frequency to highlight the spectral break.

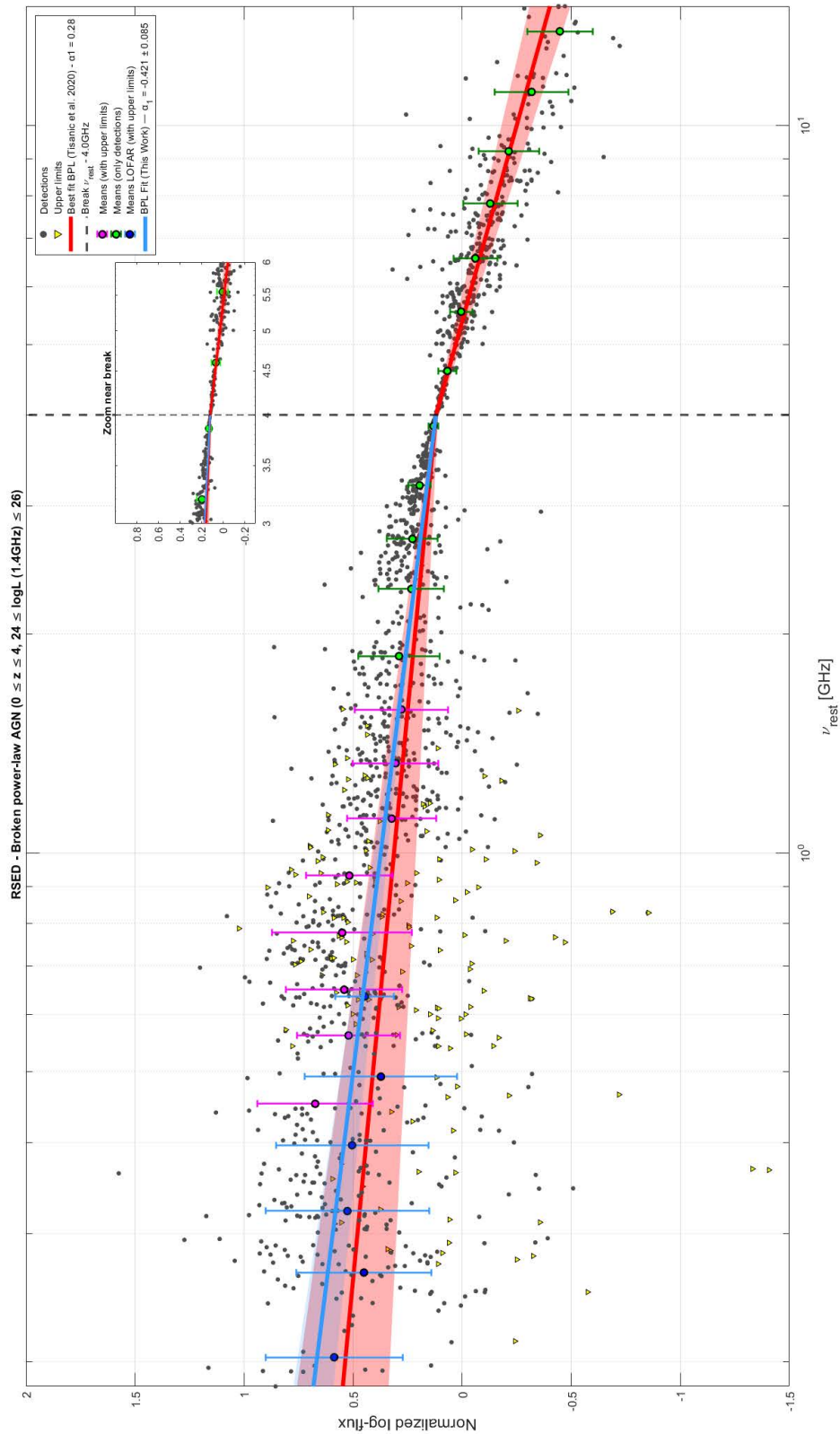


Figure 25: Reproduction of the average radio SED of radio-excess AGN with additional LOFAR data. The plot shows normalized log-flux densities as a function of rest-frame frequency for a sample of radio-excess AGN, extending the work of Tisanić et al. (2020) through the inclusion of low-frequency data at 144 MHz from LOFAR. Detections are shown as gray dots and upper limits as yellow downward arrows. Magenta circles represent mean values calculated using Kaplan–Meier survival analysis in bins containing upper limits, while green circles indicate mean values from bins with only detections. The red solid line corresponds to the original best-fitting broken power-law (BPL) model with a fixed break frequency at 4.0 GHz, and the shaded area shows its 1σ confidence interval. The blue curve represents the updated BPL fit derived from the extended dataset including LOFAR data, as developed in this thesis, with its own 1σ confidence interval.

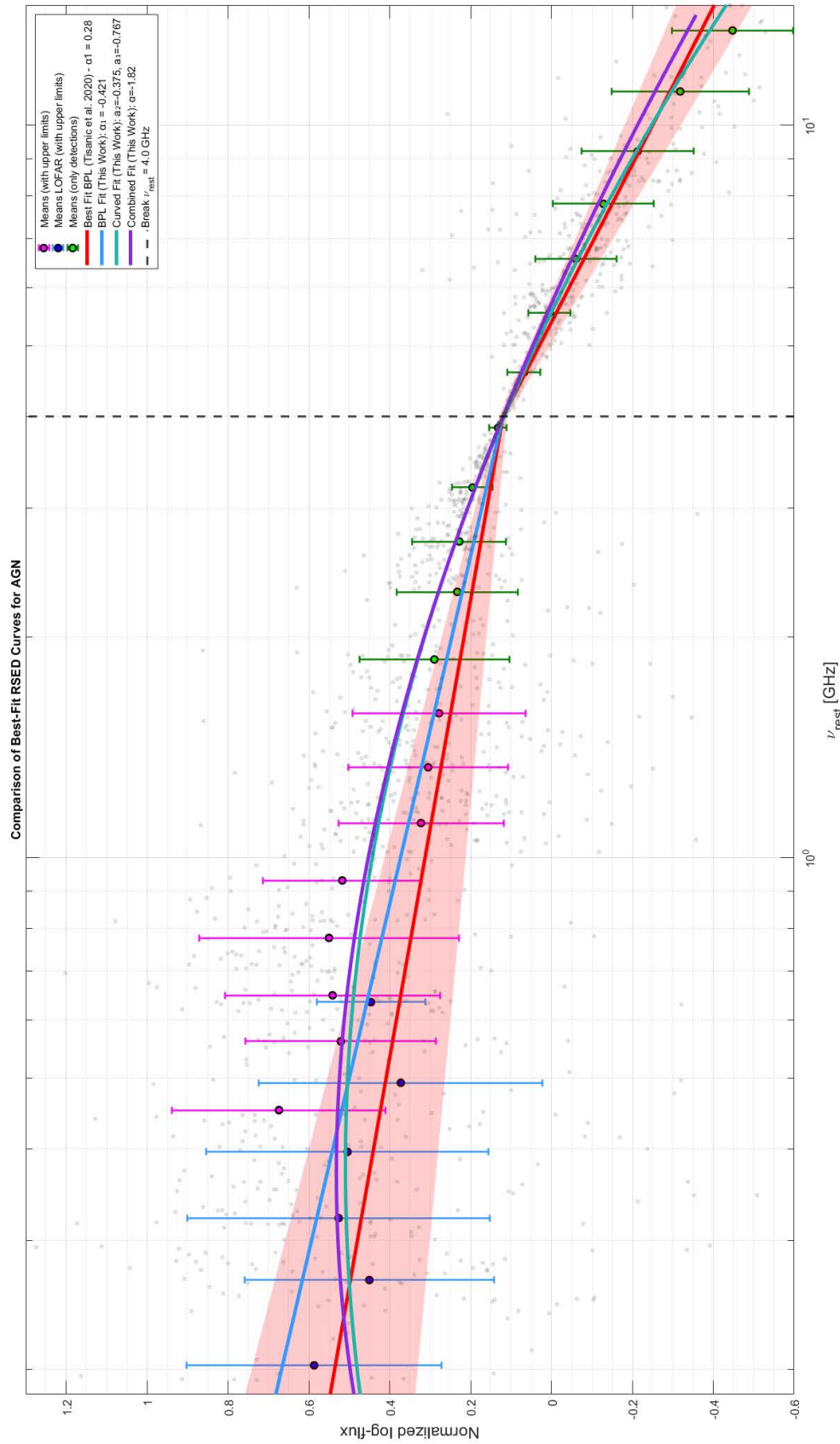


Figure 26: Comparison of best-fit RSED curves for AGN. Gray points show individual sources. Magenta means include upper limits (survival analysis), blue means are from LOFAR (with upper limits), and green means use only detections. Vertical error bars indicate the standard error on the mean. The red line shows the best-fit broken power-law (BPL) from Tisanić et al. (2020). The blue (BPL fit), olive (curved fit), and purple (combined fit) lines are derived in this work. The shaded red area is the confidence region for the Tisanić et al. BPLs fits. The dashed line marks the break at $\nu_{\text{rest}} = 4$ GHz.

3.3.2 SFG - RSED

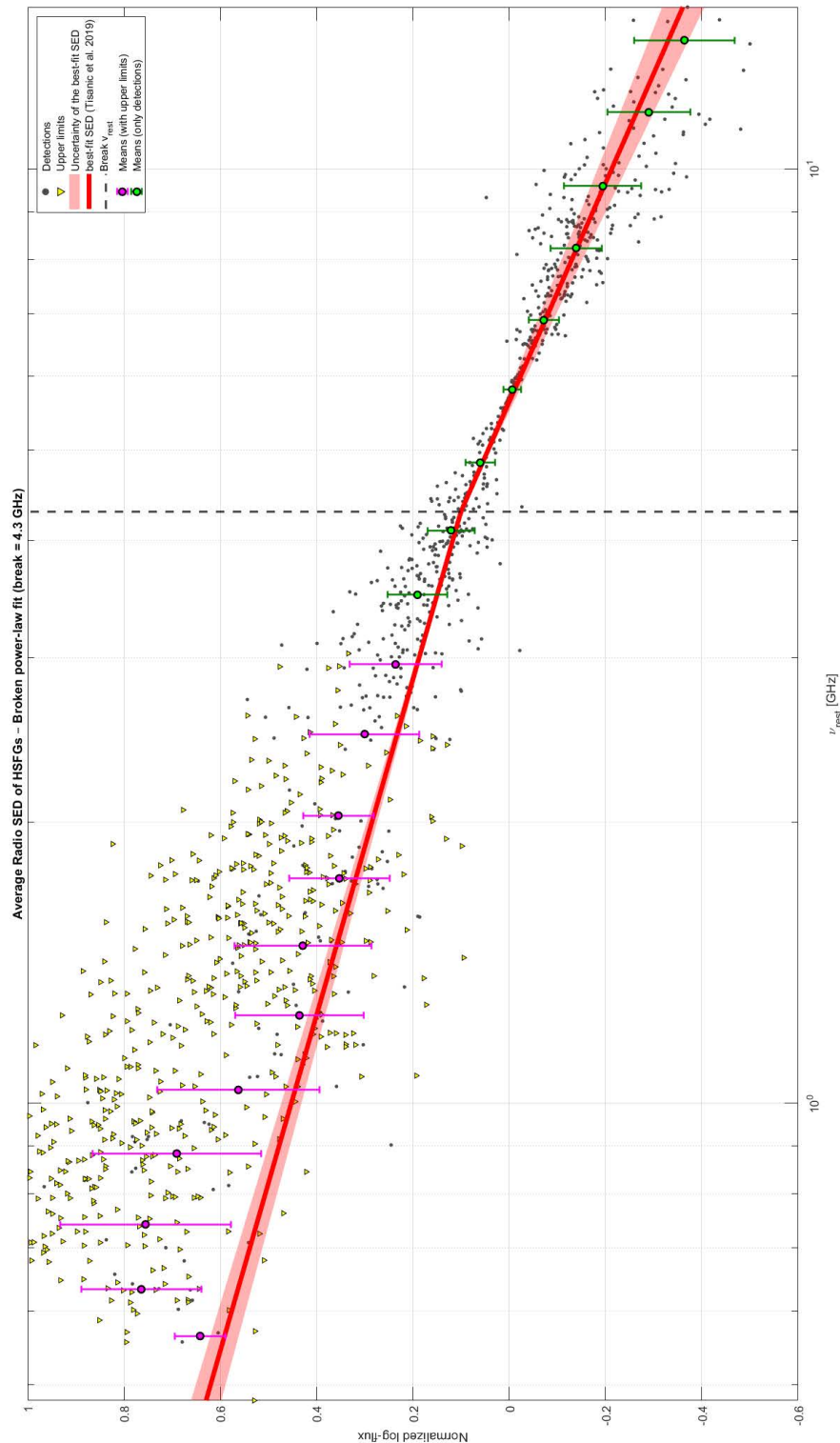


Figure 27: Reproduction of the average radio SED of highly star-forming galaxies (HSFGs) based on the methodology and dataset presented in Tisanić et al. (2020). The plot displays normalized log-flux densities as a function of rest-frame frequency, with detections shown as gray dots and upper limits as yellow downward arrows. Green circles represent mean flux values in bins containing only detections, while magenta circles denote mean values derived via Kaplan–Meier survival analysis in bins that include upper limits. The red solid curve corresponds to the best-fitting broken power-law (BPL) model with a fixed break frequency at 4.3 GHz, and the shaded area indicates the 1σ confidence interval of the fit. This visualization highlights the characteristic spectral curvature of HSFGs, with a flattening below the break frequency and steepening above it.

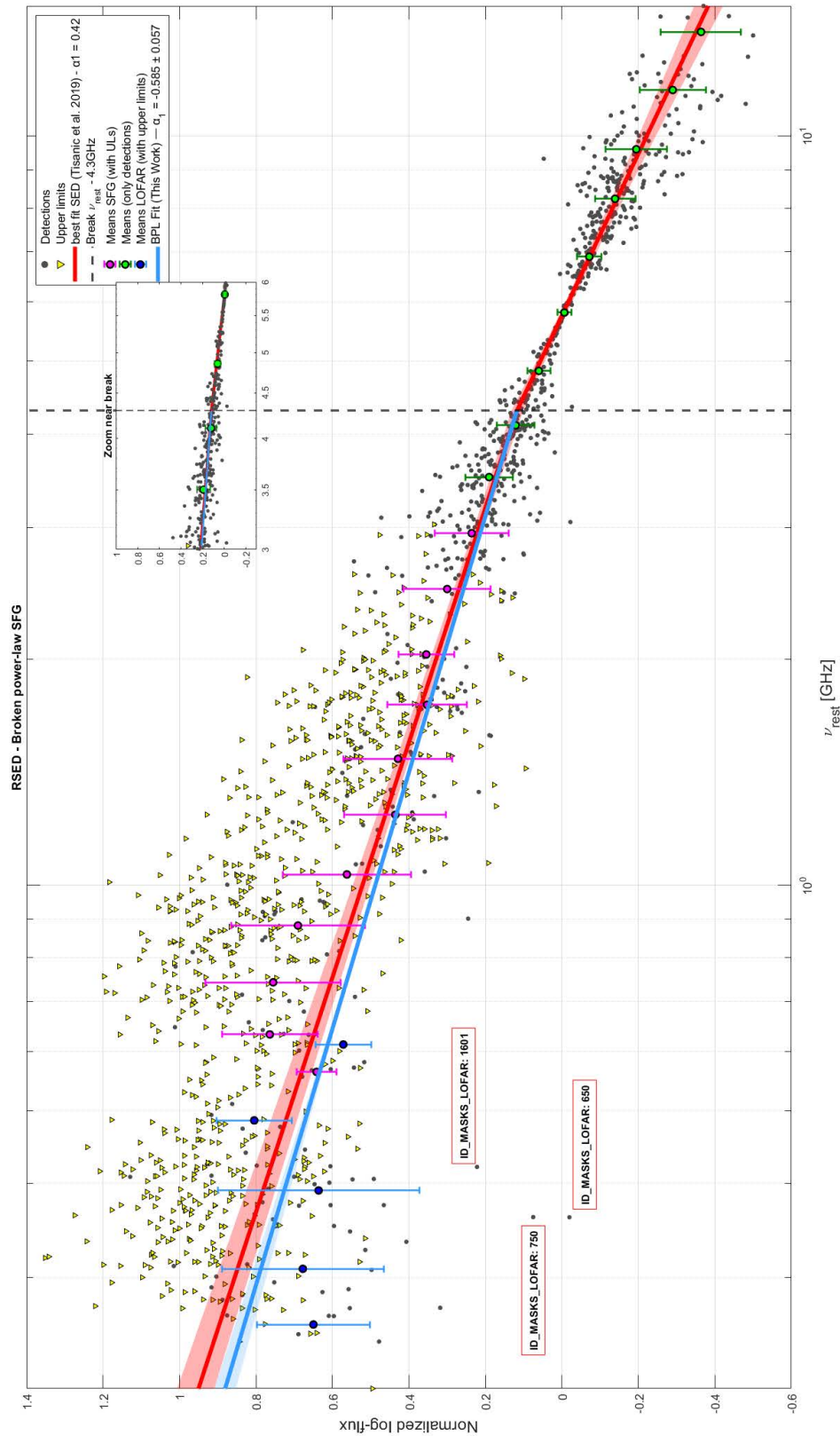


Figure 28: Refined average radio SED of star-forming galaxies (SFGs), following the methodology of Tisanić et al. (2019), with additional LOFAR data at 144 MHz. The plot displays normalized log-flux densities as a function of rest-frame frequency. Detections are shown as gray dots, and upper limits as yellow downward arrows. Magenta circles indicate mean values derived via Kaplan–Meier survival analysis (bins containing both detections and upper limits), while green circles represent mean fluxes computed from detections only. The red solid line with shaded region corresponds to the original broken power-law (BPL) fit of Tisanić et al., with a fixed break at 4.3 GHz. The newly introduced blue solid line represents a broken power-law fit applied to the updated dataset, incorporating additional detections and upper limits from LOFAR (144 MHz), while maintaining the same uncertainty envelope.

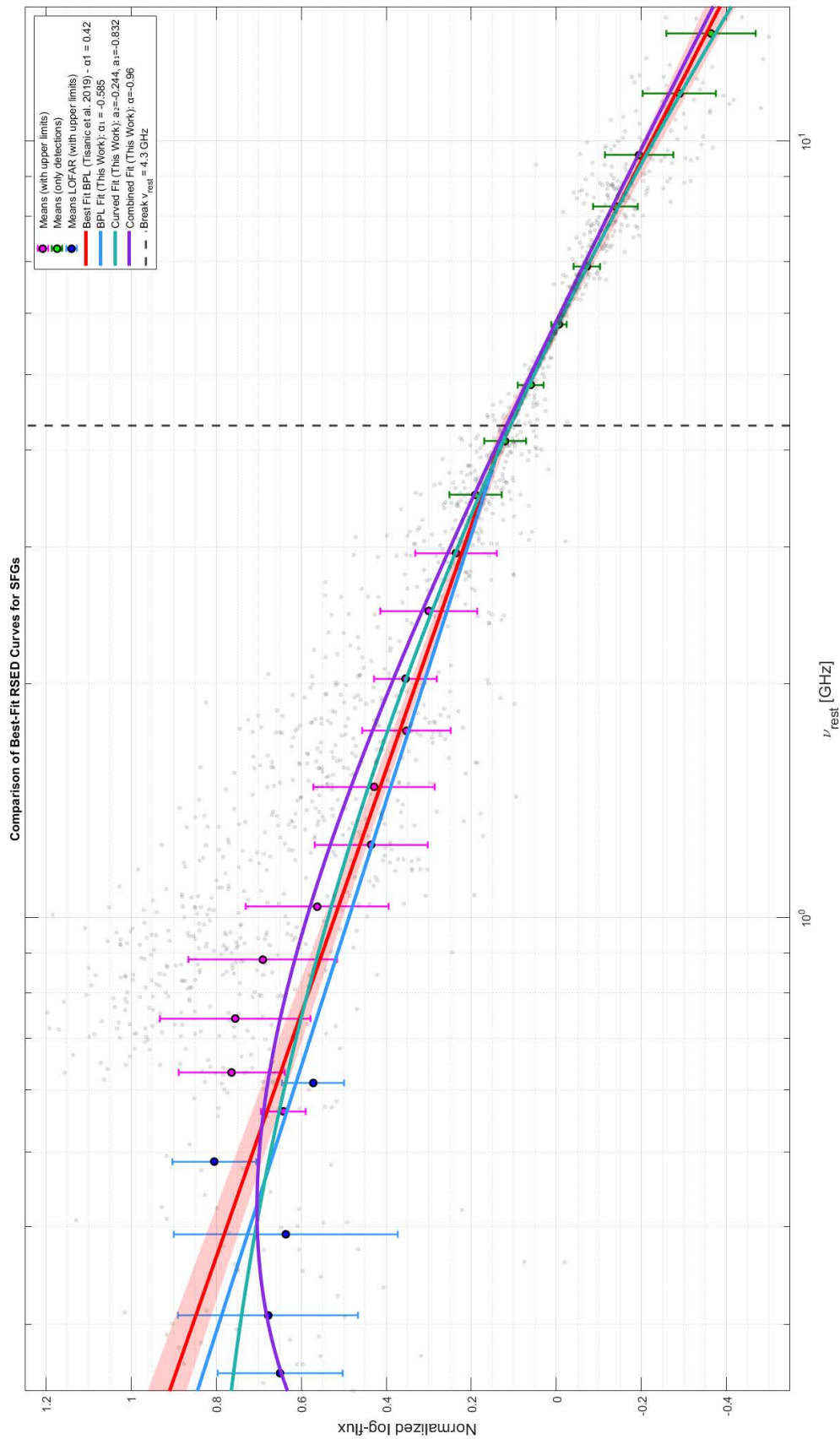


Figure 29: Comparison of best-fit radio SED curves for SFGs. Gray points show individual sources. Magenta means include upper limits (survival analysis), green means use only detections, and blue means are from LOFAR (with upper limits). Vertical error bars show the standard error on the mean. The red line is the best-fit broken power-law (BPL) from Tisanić et al. (2020). The blue (BPL fit), olive (curved fit), and purple (combined fit) lines are derived in this work. The shaded red area shows the confidence region for the Tisanić et al. BPLs fits. The dashed line marks the rest-frame break at 4.3 GHz.

3.3.3 SFG Individual Cases

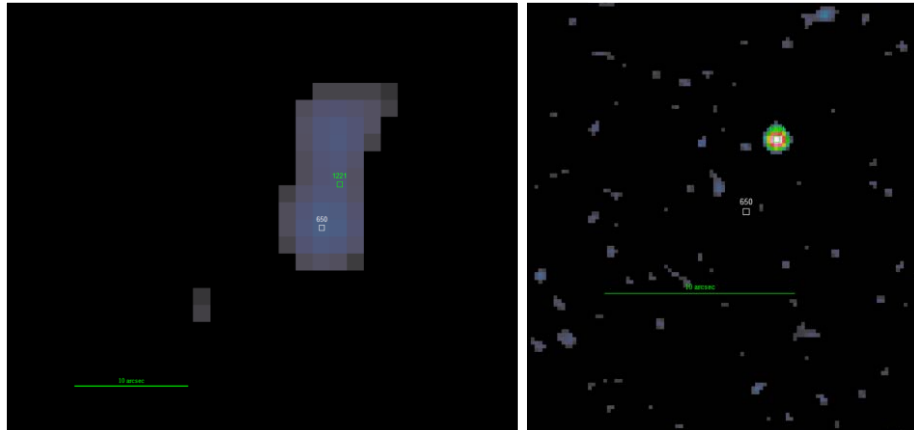


Figure 30: The white box shows the mask ID (650, max pixel position), and the green box shows the VLA ID (1221). The green line represents 10 arcsec. Left: 144 MHz & Right: 3 GHz.

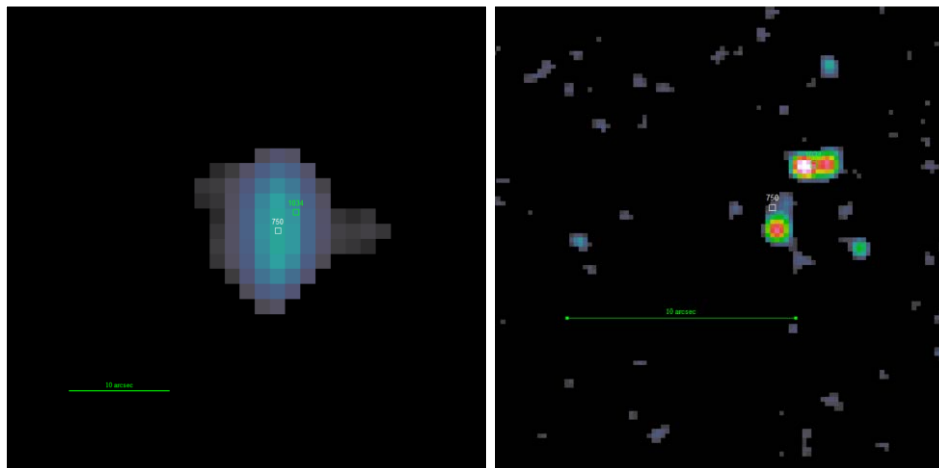


Figure 31: The white box shows the mask ID (750, max pixel position), and the green box shows the VLA ID (1034). The green line represents 10 arcsec. Left: 144 MHz & Right: 3 GHz.

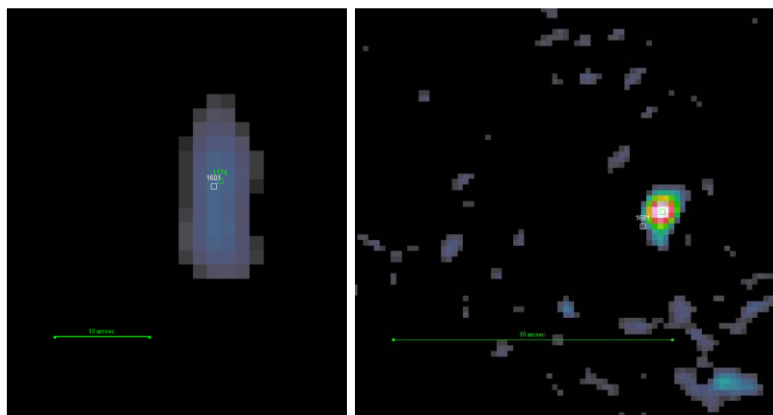


Figure 32: The white box shows the mask ID (1601, max pixel position), and the green box shows the VLA ID (1174). The green line represents 10 arcsec. Left: 144 MHz & Right: 3 GHz.

3.4 Numerical Results

3.4.1 Quantitative Analysis of the AGN Sample

Bin	Freq_GHz	N_total	N_det	N_ulim	MeanLogFlux	StdLogFlux
1	0,201797719	12	12	0	0,586968047	0,314951717
2	0,263735867	55	50	5	0,467935227	0,316033674
3	0,322007507	64	56	8	0,569975094	0,363344613
4	0,397744992	51	40	11	0,559733014	0,342912988
5	0,496976647	18	14	4	0,41771902	0,347786586
6	0,634818356	13	11	2	0,440614429	0,141971074
7	0,451656627	13	12	1	0,673927652	0,264151534
8	0,560811819	38	25	13	0,520867722	0,235389901
9	0,647853308	78	53	25	0,541302851	0,265489803
10	0,77566552	90	63	27	0,550353293	0,320926448
11	0,929703655	62	37	25	0,517496532	0,196925509
12	1,113175729	101	91	10	0,322866434	0,204110154
13	1,326686025	94	84	10	0,305004766	0,196843774
14	1,572328435	73	69	4	0,278336139	0,213563508
15	1,862681442	51	51	0	0,289720611	0,185728308
16	2,302449451	54	54	0	0,233095104	0,150077082
17	2,699142968	82	82	0	0,228190635	0,115941567
18	3,196628627	86	86	0	0,195338675	0,050522588
19	3,855971919	67	67	0	0,132029876	0,02163532
20	4,588835999	47	47	0	0,068016666	0,040916654
21	5,539120192	90	90	0	0,004875983	0,052023431
22	6,558417659	88	88	0	-0,060310217	0,100528863
23	7,798293208	65	65	0	-0,128364364	0,124590908
24	9,201714197	40	40	0	-0,213185421	0,138461852
25	11,10442788	21	21	0	-0,318258779	0,169953203
26	13,44780019	14	14	0	-0,447621031	0,150260176

Table 2: Statistical Properties per Frequency Bin for the AGN Sample. Table presents the numerical statistics of the AGN sample across 26 logarithmic rest-frame frequency bins. Columns include: the bin index, the mean rest-frame frequency (Freq_GHz), the total number of sources within each bin (N_total), the number of detections (N_det), and upper limits (N_ulim). The last two columns show the mean and standard deviation of the normalized logarithmic flux values, computed exclusively from detections in each bin.

Model	Alpha1	Alpha2	NuBreak_GHz
SPL (Tisanic et al. 2020)	0,64		
BPL (Tisanic et al. 2020)	0,28	1,16	4
BPL (This Work)	0,420866399	1,16	4
Curved Fit (This Work)	0,7671744	0,374985201	
Absorp. Fit (This Work)	1,824626109		

Table 3: Best-Fit Model Parameters for the AGN Sample. Table presents the spectral indices and break frequencies derived from different models fitted to the average radio SED of AGNs. Models from the literature (Tisanić et al. 2020) include a single power-law (SPL) and a broken power-law (BPL), while our work introduces three alternative fits: a BPL based on updated binned data, a second-order polynomial (Curved Fit), and a composite fit incorporating synchrotron self-absorption (Absorp. Fit). For the latter, the curvature is applied below the fixed break frequency of 4.0 GHz, ensuring consistent comparison across models.

3.4.2 Quantitative Analysis of the SFG Sample

Bin	Freq_GHz	N_total	N_det	N_ulim	MeanLogFlux	StdLogFlux
1	0,259343149	12	5	7	0,649767914	0,147000609
2	0,307748088	80	15	65	0,67755199	0,211542108
3	0,391488579	132	23	109	0,636311796	0,26398446
4	0,485552982	60	5	55	0,804576671	0,098784297
5	0,612991703	19	3	16	0,571722395	0,072863614
6	0,563580058	5	2	3	0,641724625	0,05270106
7	0,63232779	35	7	28	0,763936323	0,124935568
8	0,742030845	71	5	66	0,755181304	0,177366147
9	0,883507123	94	12	82	0,690310677	0,175355402
10	1,033616238	74	7	67	0,562386622	0,167826203
11	1,241879686	79	11	68	0,435592581	0,133068021
12	1,474766726	88	11	77	0,428645393	0,142179119
13	1,74071362	90	18	72	0,352665876	0,104496619
14	2,031876064	49	8	41	0,354610005	0,073401442
15	2,484187577	39	18	21	0,300119674	0,113846602
16	2,948972144	65	60	5	0,23593522	0,095976453
17	3,500495322	77	77	0	0,190539434	0,061604119
18	4,103485288	90	90	0	0,120889377	0,049188871
19	4,848186246	44	44	0	0,060244428	0,030492243
20	5,803865461	57	57	0	-0,006271316	0,017686778
21	6,888174615	74	74	0	-0,071572184	0,03151129
22	8,228738757	100	100	0	-0,138850203	0,052823613
23	9,583793068	56	56	0	-0,194428262	0,080210137
24	11,49582015	28	28	0	-0,289706693	0,086314306
25	13,73210651	8	8	0	-0,363754099	0,104719734

Table 4: Statistical Properties per Frequency Bin for the SFG Sample. Table summarizes the statistical properties of the Star-Forming Galaxy (SFG) sample across 25 logarithmic bins in rest-frame frequency. The columns include: bin number, mean rest-frame frequency in GHz (Freq_GHz), total number of sources in the bin (N_total), number of detections (N_det), and number of upper limits (N_ulim). The last two columns provide the mean and standard deviation of the normalized logarithmic fluxes, calculated using detections only.

Model	Alpha1	Alpha2	NuBreak_GHz
SPL (Tisanic et al. 2019)	0,67		
BPL (Tisanic et al. 2019)	0,42	0,94	4,3
BPL (This Work)	0,584546761	0,94	4,3
Curved Fit (This Work)	0,831500921	0,243928915	
Absorp. Fit (This Work)	0,958369659		

Table 5: Best-Fit Model Parameters for the SFG Sample. This table summarizes the parameters of five different models used to fit the radio spectral energy distribution (RSED) of star-forming galaxies (SFGs) in this study. The models include: (1) a single power-law (SPL) from Tisanić et al. (2019), (2) a broken power-law (BPL) from the same study, (3) the best-fit BPL derived in this work assuming a fixed break at 4.3 GHz, (4) a second-degree polynomial (curved) fit applied to data below the break, and (5) an empirical synchrotron self-absorption (SSA) model also fitted below the break. “Alpha1” corresponds to the spectral slope below the break frequency, and “Alpha2” to the slope above the break (when defined). All models use a fixed break frequency at 4.3 GHz, except for the SSA model, where this value represents the characteristic turnover frequency.

Chapter 4: Discussion

4.1 Interpretation and Scientific Assessment of Results

Following the presentation of the results in Chapter 3, we now proceed with the interpretation and scientific assessment of our findings. In this chapter, we evaluate the implications of our analysis for both AGN and SFGs, situating our results within the broader astrophysical context. Special emphasis is placed on a direct comparison of the derived RSED slopes and best-fit parameters with the corresponding results from Tisanić papers, allowing us to highlight both similarities and differences in the radio spectral properties of AGN and SFG populations. We further consider what these findings reveal about the physical processes governing galaxy evolution, address the principal limitations and challenges encountered throughout this work, and conclude by outlining promising directions for future research.

4.1.1 AGN Radio SEDs: Best-Fit Approaches and Results

The analysis of AGN results in this study is strongly linked to the methodologies and diagnostic diagrams developed by (Tisanić, et al., 2020), which we have aimed to reproduce and extend with our own sample. For the construction of the radio SEDs of AGN, we incorporated flux density measurements at 1.4 GHz, 3 GHz, 610 MHz, and 325 MHz, as well as newly included 144 MHz data from LOFAR. As described in Chapter 3, the normalization procedure for all AGN was carefully adapted to follow the prescriptions outlined by Tisanić et al., ensuring a consistent and direct comparison of both the methodology and the resulting fits.

A central result of Tisanić et al. was the demonstration that a simple power-law does not provide an adequate fit to the average AGN radio SED. Instead, their analysis established that a BPL offers a significantly better description, with a well-defined spectral break around 4 GHz. The red curve in Figure 24 represents this reference BPL model as derived in their paper, using the published best-fit parameters. The choice of the 4 GHz break point in the AGN RSED is grounded in a detailed statistical approach: Tisanić et al. (2020) initially treated the break frequency as a free parameter in their broken power-law fits, employing MCMC methods and ODR to robustly estimate model parameters. The fitting consistently converged on a break frequency just above 4 GHz (4.1 ± 0.2 GHz - 4GHz break frequency), with this value being well constrained by the data and not significantly affected by upper limits at higher frequencies. This statistical determination is also supported by the observed spectral curvature—interpreted as synchrotron aging and cooling of the electron population—making the 4 GHz break both an empirically justified and physically meaningful feature of the average AGN radio SED. In our own sample, we adopted the same break frequency at 4 GHz, thus allowing for a direct and meaningful comparison.

The major innovation in our analysis is the extension of the AGN sample to lower frequencies by including LOFAR measurements at 144 MHz. The addition of this low-frequency data provides a more complete view of the SED and extends the dynamic range of the analysis. Upon incorporating these data points, we recalculated the best-fit BPL parameters for our sample, yielding the blue line in Figure 25. It should be noted that this new best-fit (blue line) is plotted only for the low-frequency segment below the break point at 4 GHz, as the newly incorporated data extend exclusively toward lower frequencies; above 4 GHz, the spectral slope remains unchanged and coincides with the original (red) BPL model. Notably, the updated fit remains consistent, within statistical uncertainties, with the BPL

model originally derived by Tisanić et al. This agreement reinforces the robustness of the broken power-law description, at least for the AGN population and frequency coverage probed by our study. However, it should be emphasized that this result specifically pertains to the present dataset; a more extensive survey or the inclusion of a larger AGN sample might well modify the best-fit parameters, especially at low frequencies, given the intrinsic diversity of this population properties.

A notable feature of the AGN RSED is the substantial scatter observed across the data points. This large dispersion (scattering) is indicative of the intrinsic diversity and complexity within this population. Unlike more homogeneous galaxy samples, AGN are inherently mixed populations, spanning a range of physical properties, evolutionary stages, and host galaxy environments. In our case, the selection of radio-excess AGN is intended to isolate sources dominated by non-thermal activity; nevertheless, contamination from star-forming components within the host galaxies cannot be ruled out, especially for composite sources. This admixture likely contributes to the observed scatter, as does the range of accretion modes, jet properties, and environments associated with different subtypes. The central core of an AGN, powered by accretion onto a SMBH, is itself subject to complex and often poorly constrained processes, including variations in jet orientation, magnetic field configuration, and radiative efficiency. As a result, the ensemble-averaged SED of AGN is shaped not only by the intrinsic emission mechanisms but also by the stochastic and heterogeneous nature of the population. This complexity underscores the challenges inherent in interpreting the mean properties of AGN from unresolved radio observations.

Moving beyond the broken power-law analysis, Figure 26 presents a comparative view of the various models fitted to our AGN RSED. Four distinct curves are shown: the BPL model from Tisanić et al. (2020) (with best-fit parameters $\alpha_1 = 0.28$, $\alpha_2 = 1.16$, break frequency $\nu_b = 4$ GHz), our own BPL fit (blue curve; $\alpha_1 = 0.42$, $\alpha_2 = 1.16$, $\nu_b = 4$ GHz), a polynomial (curved) fit, characterized by spectral indices $\alpha_1 = 0.76$ and $\alpha_2 = 0.37$, which reflects higher-order deviations from a simple BPL, and an absorption model with $\alpha_1 = 1.82$, designed to capture potential low-frequency flattening or turnover due to synchrotron self-absorption. The parameters of all models are provided in Table 3 immediately following the plots.

A comparative analysis of these fits reveals that both the curved fit and the absorption model in our work predict a clear downturn in the SED at the lowest frequencies. In contrast, the original self-absorption model of (Tisanić, et al., 2020) - as shown in Figure 4 (upper panel) of their paper - does not exhibit such a pronounced downturn in this frequency regime. However, it is important to note that the Tisanić model does not exclude the possibility of a downturn; rather, the available data in their study did not extend to sufficiently low frequencies to confirm or reject such spectral behavior. This behavior, if confirmed by future observations, could signal the onset of absorption effects or a fundamental change in the dominant emission mechanism in AGN at low radio frequencies. At the same time, the overall similarity of the BPL fits - between our extended sample and the original Tisanić model - supports the notion that the broken power-law remains a robust statistical descriptor of the mean AGN radio SED over the range 144 MHz to several GHz. Nevertheless, the persistence of large scatter and the prospect of new trends emerging at even lower frequencies (such as 60 MHz) both motivate further studies with deeper, wider-area surveys.

In summary, the analysis of AGN in this work confirms the necessity of a broken power-law description for the average radio SED, as initially demonstrated by Tisanić et al., and shows that the inclusion of LOFAR 144 MHz data does not fundamentally alter this conclusion for the present sample. The substantial scatter in the SED, attributable to the mixed nature of the AGN population and the underlying physical diversity, remains an important factor in interpreting ensemble results. Future work - particularly with the addition of even lower-frequency data and refined AGN classifications - will be essential for uncovering the full complexity of radio emission processes in active galactic nuclei.

4.1.2 SFG Radio SEDs: Best-Fit Approaches and Results

The analysis of the RSEDs for HSFGs in the COSMOS field, as constructed in this work, leverages both the legacy of (Tisanić, et al., 2019) and the new observational power provided by LOFAR 144 MHz data. Our approach follows closely the normalization procedure detailed in the Tisanić et al. study, ensuring a robust and fair comparison with previous literature, and allowing us to test and extend the established paradigms regarding the form of the HSFG radio SED.

In this study, our sample consists of galaxies classified as HSFGs. These are systems with extremely elevated star formation rates, typically exceeding 100 solar masses per year. Such galaxies are rare in the local Universe but become increasingly prevalent at higher redshifts, often corresponding to massive, gas-rich systems undergoing intense episodes of stellar birth. The inclusion of these objects allows us to probe the RSED in regimes dominated by vigorous star formation, offering unique insight into the interplay between cosmic ray electron injection, magnetic field strength, and ISM conditions in extreme star-forming environments.

The dataset utilized in our analysis consists of flux measurements at five key frequencies: LOFAR at 144 MHz, GMRT at 325 and 610 MHz, and VLA at 1.4 and 3 GHz. This coverage provides a broad and continuous spectral baseline, critical for capturing both low-frequency curvature and high-frequency flattening in the SED. As shown in Figure 27 of this thesis, all data points were normalized according to the methodology of Tisanić et al., with the normalization pivot set to 4.3 GHz, consistent with the break frequency (ν_{break}) identified in their work.

A crucial outcome of the Tisanić et al. (2019) analysis was the demonstration that the average HSFG radio SED is not well described by a simple, single power law across the full range of frequencies, but instead requires a BPL to account for the observed curvature. This conclusion was physically motivated by synchrotron aging and the increasing contribution of thermal (free-free) emission at higher frequencies, and is statistically substantiated by the much better fit of the BPL compared to a simple power law. The red line in Figure 27 represents the original Tisanić et al. best-fit BPL model, which was calculated using survival analysis techniques to incorporate the information from both detections and upper limits. In the context of our expanded COSMOS HSFG sample, we observe that the Tisanić BPL model continues to provide an excellent description of the data, validating their conclusion that HSFGs exhibit a spectral steepening below the break and a flattening above.

In our analysis, the break point (ν_{break}) at 4.3 GHz is of particular importance. In the Tisanić et al. paper, this break was chosen as it marks the frequency at which the physical mechanisms underlying the radio emission transition: below the break, synchrotron emission dominates, subject to energy-dependent losses (i.e., synchrotron aging), while above the break, thermal free-free emission becomes increasingly significant, leading to a flattening of the SED. The selection of this break point is empirically motivated by the inflection observed in the stacked HSFG SEDs, as well as by theoretical models predicting such curvature for typical star-forming galaxies. By retaining this ν_{break} in our modeling, we maintain direct comparability with previous studies and ensure that our new fits can be interpreted in the same physical framework.

The major advancement presented in this thesis is the extension of the HSFG sample to lower frequencies, enabled by the inclusion of 144 MHz LOFAR data. This addition not only

increases the statistical power of the dataset but also provides critical leverage on the low-frequency regime, where curvature from synchrotron aging and absorption effects are expected to manifest most strongly. Upon fitting a new broken power-law model to the expanded data (shown as the blue line in Figure 28), we observe that our best-fit parameters have shifted relative to the original Tisanić et al. values - reflecting the influence of the newly included 144 MHz data on the low-frequency slope. However, this shift remains within the statistical uncertainties of the previous fit, indicating that while our extended sample does affect the shape of the broken power law, it does not introduce any significant deviation. This outcome underscores the robustness of the BPL description, while also highlighting the sensitivity of the low-frequency SED shape to the addition of new data points at the lower end of the spectrum.

The comparison of best-fit models is presented in Figure 29, where four distinct curves are shown: (1) the original BPL from Tisanić et al. (red), (2) our new BPL fit (blue), (3) a polynomial (curved) fit (olive), and (4) an absorption model (purple). The precise parameter values for each model are listed in Table 5 (see after figures), for direct quantitative comparison. The Tisanić BPL fit is characterized by $\alpha_1 = 0.42$, $\alpha_2 = 0.94$, and $\nu_{\text{break}} = 4.3$ GHz. Our updated BPL fit yields $\alpha_1 = 0.58$, $\alpha_2 = 0.94$, and $\nu_{\text{break}} = 4.3$ GHz (see Table 5). It is important to note that our new BPL fit (blue line) is constrained by the inclusion of 144 MHz data only at frequencies below the break, i.e., for $\nu < 4.3$ GHz. For $\nu > 4.3$ GHz, the high-frequency slope is identical to that of Tisanić et al., since no additional data points were introduced in this regime, and thus the fit naturally follows the original model. The curved polynomial fit, characterized by coefficients $\alpha_1 = 0.83$ and $\alpha_2 = 0.24$, is essentially a second-order polyfit designed to flexibly capture any additional curvature. The absorption model, with a low-frequency spectral index $\alpha_1 = 0.95$, introduces a turnover at the lowest frequencies consistent with synchrotron self-absorption, with parameters determined from the fit accordingly.

Examining the plot, it is evident that both the curved fit and absorption model display a clear downturn at the lowest frequencies, qualitatively matching the expected impact of absorption or aging processes and consistent with models 1, 2, and 4 from (as shown in Fig. C.1 of (Tisanić, et al., 2019)). This behavior, also seen in our data, suggests that such curvature is a generic property of HSFG populations, further validating the necessity of a BPL or higher-order model for accurately representing the mean SED.

A notable aspect of our analysis concerns the treatment of upper limits. As with the AGN sample, a significant fraction of SFG measurements at the lowest frequencies fall below the detection threshold, resulting in many upper limits that must be properly incorporated into the fit. Survival analysis, as described by Tisanić et al. (2019) and implemented here, enables us to statistically include these censored data points, ensuring unbiased parameter estimates. The presence of many upper limits is partly a reflection of the survey depth, but also an astrophysical signal: as shown by Tisanić, low-frequency suppression can be physical, due to absorption and aging effects, rather than merely an observational artifact. Our results confirm this, with the fitted models naturally bending down to accommodate the non-detections at 144 and 325 MHz.

During the analysis, we identified three sources whose radio SEDs at 144 MHz exhibited unusually low flux values compared to the main sample trend (see Figures 30, 31, 32). Although these objects could in principle be cases of underestimated flux or unaccounted-for absorption, we did not have conclusive evidence to classify them as outliers or to remove them.

Thus, they were retained in the analysis, in accordance with rigorous statistical treatment and to avoid introducing bias. This approach follows the best practice outlined in Tisanić's thesis, where all objects are included unless robust grounds for exclusion exist.

To summarize, the main findings for the SFG sample can be encapsulated as follows: (1) the mean radio SED is best described by a broken power law with a break at 4.3 GHz, in agreement with Tisanić et al. (2019) the inclusion of LOFAR 144 MHz data extends and refines our understanding of the low-frequency SED, leading to a new fit that reveals subtle but physically meaningful changes in the spectral shape - while remaining statistically consistent with previous results; (3) alternative models (curved and absorption) provide additional perspective on the underlying emission and absorption processes, with the extended frequency coverage of our sample enabling a more stringent test of these models than was previously possible, even though the BPL continues to offer a robust statistical description within the current data set; (4) the statistical treatment of upper limits is essential, as low-frequency suppression is both a data reality and a physical signature; and (5) the overall SED curvature observed here supports a scenario where synchrotron aging and absorption shape the low-frequency spectrum of SFGs.

4.2 Limitations and Challenges

Despite the significant advancements enabled by the expanded low-frequency coverage and the application of advanced statistical modeling, several limitations and challenges affect the robustness and generalization of our results. These issues arise both from the intrinsic properties of the COSMOS field and from the technical characteristics of the datasets used. A summary of the main limitations follows:

➤ *Spatial Resolution Limitations:*

- The LOFAR and GMRT radio maps, particularly at 144, 325, and 610 MHz, have relatively poor angular resolution compared to VLA 3 GHz data (with beam sizes ranging from ~6–15 arcsec for LOFAR to ~10×9.5 arcsec for GMRT 325 MHz).
- This limited resolution increases source blending, especially in crowded regions, which can bias flux measurements and source identification.

➤ *Sensitivity Constraints:*

- The available integration times for the low-frequency maps are relatively short, leading to higher noise levels (e.g., median RMS of 97 μ Jy/beam at 325 MHz and 39 μ Jy/beam at 610 MHz for GMRT). As a result, only a small fraction of the HSFG sample is detected at the lowest frequencies (see Tisanić, Sect. 2.2.1 & 4.1), reducing the statistical significance of the results in these bins.

➤ *Elongated and Non-Circular Beams:*

- The synthesized beams at LOFAR and GMRT frequencies are elongated and highly variable in orientation, which complicates accurate deconvolution and introduces additional uncertainty into both flux and astrometric measurements.

➤ *Field Coverage and Selection Effects:*

- The COSMOS field, though one of the best-studied extragalactic fields, covers only 2-square-degrees and is subject to edge effects and variable depth across the field, leading to spatial inhomogeneities in sensitivity and completeness.

➤ *Potential for Misassociation During Source Matching:*

- The process of cross-matching radio sources across different frequency catalogs is especially vulnerable to misidentification at 144 MHz, where the spatial resolution of LOFAR is limited. This constraint increases the likelihood that a given radio source may be incorrectly associated with the wrong counterpart. These mismatches can bias the physical interpretation of individual galaxies and introduce additional uncertainties into the statistical analysis of the population.

➤ *Calibration and Data Reduction Challenges:*

- Low-frequency radio data are particularly sensitive to systematic calibration errors (e.g., due to ionospheric effects) and require sophisticated algorithms for direction-dependent calibration (see Tisanic, Sect. 2.2.1).
- Small errors in flux scale or astrometric alignment can propagate into significant biases in the constructed SEDs.

➤ *Underlying Star-Forming Galaxy Contribution in AGN Sample:*

- The presence of a substantial star-forming component within AGN-dominated sources can lead to an overestimation of the radio continuum emission attributed solely to AGN activity. This blending complicates efforts to derive accurate AGN luminosities, spectral indices, and evolutionary trends, potentially biasing interpretations of AGN energetics and feedback processes within galaxies.

4.3 Future Work and Prospects

Looking forward, several promising directions arise from the present study, both in terms of observational capabilities and methodological advances. The scientific impact of this work can be further expanded through targeted improvements and the exploitation of upcoming facilities and techniques. Key prospects and recommendations for future research include:

✓ *Extension to Lower Frequencies:*

The anticipated inclusion of 60 MHz data will be transformative. Sampling this regime is critical for directly testing the presence and shape of low-frequency turnovers observed in our current absorption fits (see Figures 26 and 29). These new data will allow us to:

- Distinguish between synchrotron self-absorption, free-free absorption, or alternative mechanisms as the drivers of SED curvature.
- Confirm whether the observed downturn is universal among AGN and HSFGs or dependent on physical/environmental conditions.

✓ *Expansion of Survey Area and Sample Size:*

Increasing the survey footprint and depth—through deeper integrations or participation in larger-area programs—will:

- Grow the AGN and HSFG samples, enhancing statistical power.
- Improve the detection of rare or extreme sources, thus refining constraints on the shape and scatter of the radio SED.
- Enable a more thorough exploration of population mixtures, including composite and ambiguous objects

✓ *Adoption of Next-Generation Radio Facilities:*

The deployment of advanced telescopes such as the SKA (SKA low 50-350MHz, SKA mid 350MHz-15.4Ghz with a goal of 24 GHz), will:

- Offer higher angular resolution and improved sensitivity, mitigating limitations from source confusion.
- Enable better calibration and dynamic range, facilitating more accurate SED reconstruction across the full frequency range.

✓ *Application of Machine Learning and Advanced Statistics:*

Future work should harness machine learning for:

- Automated source classification and matching across catalogs.
- Sophisticated SED fitting, feature extraction, and uncertainty quantification, especially for large heterogeneous datasets.

These developments will directly address the principal limitations outlined in this thesis, including issues of resolution, sample size, matching accuracy, and frequency coverage.

Ultimately, the findings of this thesis—particularly regarding low-frequency spectral curvature (combined fit -SSA) and the comparison of competing fit models—open new avenues for understanding galaxy evolution. As radio SED studies expand to cover a wider range of frequencies and more diverse galaxy samples, they are set to become a key tool for exploring the processes that shape galaxies and for advancing our understanding of cosmic evolution.

Chapter 5: Conclusion & Outlook

5.1 Summary of Key Findings and Major Conclusions

This thesis set out to constrain the average radio spectral energy distributions (RSEDs) of active galactic nuclei (AGN) and star-forming galaxies (SFGs) in the COSMOS field, utilizing multi-frequency radio data spanning from LOFAR 144 MHz up to VLA 3 GHz (LOFAR 144MHz (median rms: 100 μ Jy/beam, PI: Vardoulaki; project DDT19_002), GMRT 610MHz (median rms: 39 μ Jy/beam, PI: H. R. Klöckner; project 11HRK01), GMRT 325 MHz (median rms: 97 μ Jy/beam, PI: S. Croft; project 07SCB01), VLA-COSMOS 1.4 GHz Joint Project (median rms: 12 μ Jy/beam, Schinnerer et al. 2010) and VLA-COSMOS 3 GHz Large Project (median rms: 2.3 μ Jy/beam, Smolcic et al. 2017)), complemented by ancillary multi-wavelength datasets. The primary goal was to assess how the inclusion of new low-frequency observations refines the characterization of the spectral index, spectral curvature, and the physical processes governing radio emission in extragalactic sources. Advanced statistical techniques, including survival analysis for censored data, were systematically applied to construct unbiased average SEDs and to robustly distinguish between AGN- and SFG-dominated radio populations.

Key findings and major conclusions are summarized as follows:

- Catalogue Development and Methodological Advances: A detailed and reproducible methodology was established for the construction of the LOFAR radio source catalogue, encompassing all steps from source identification and flux measurement to source matching across multiple datasets. This approach enhances the reliability of the derived physical parameters and facilitates direct comparison with legacy catalogues and future surveys.
- Comprehensive Multi-Catalogue Source Matching and Deblending: The analysis implemented a systematic cross-matching procedure across multiple radio source catalogues (masks_lofar, counterpart, SuperDeblended, PyBDSF), incorporating advanced deblending and masking techniques to accurately identify primary and secondary radio sources, including the robust separation of blended and extended emission components.
- Extension of RSEDs to Lower Frequencies: For the first time, this work incorporates LOFAR 144 MHz measurements into the statistical construction of average RSEDs for COSMOS galaxies, significantly improving the constraints on the low-frequency spectral slope for both AGN and SFGs.
- Spectral Curvature and Model Fitting: The results confirm that a simple power-law is insufficient to describe the radio SEDs of both populations. Instead, a BPL model, as established in previous literature, provides a significantly better fit, capturing both the flattening at low frequencies and the steepening at higher frequencies—signatures of synchrotron aging and the interplay between non-thermal and thermal emission processes.
- Distinct Spectral Properties for AGN and SFGs: AGN-dominated sources display pronounced spectral curvature, with a flatter spectral index below the break and a much steeper index above it, while SFGs exhibit more moderate curvature, in line with their younger cosmic-ray electron populations and higher star formation rates. These results align with, and refine, recent studies, but extend their validity into the low-frequency regime.

- *Automated and Reproducible Data Pipeline*: The entire catalogue generation, data cleaning, and RSED construction process was implemented as an automated, reproducible pipeline-integrating custom MATLAB code, detailed documentation, and validation steps - which sets a transparent foundation for future large-sample studies and cross-comparisons with external datasets.

In summary, the integration of new low-frequency radio data and advanced analysis techniques has extended the empirical characterization of the average radio SEDs to lower frequencies, enabling a deeper understanding of their behavior and the physics underlying these phenomena for both AGN and SFGs in the COSMOS field. These results establish a new reference point for future studies of galaxy populations in the era of next-generation radio surveys.

5.2 Implications and Significance

The results of this thesis provide critical insights into the physical mechanisms that shape galaxy evolution, demonstrating the importance of accurately characterizing the RSEDs of both AGN and SFGs across cosmic time. By extending the statistical RSED analysis down to 144 MHz, this work enables a more robust determination of the spectral curvature at low frequencies—revealing the role of synchrotron aging and energy losses in cosmic ray electron populations, and offering a direct observational constraint on the non-thermal emission processes at work in different galaxy populations.

The precise measurement of the average RSEDs and their broken power-law (BPL) characterization directly advances our ability to calculate K-corrections for radio luminosities, a cornerstone for deriving intrinsic properties such as star formation rates (SFR) and for comparing galaxy populations across different epochs. The spectral index values determined for each segment of the BPL allow for accurate luminosity estimates in the rest-frame, through the standard relation (1.13). For SFGs, the empirically derived BPL slopes in this work provide essential inputs for reliable luminosity and SFR calculations, particularly at high redshifts where assumptions of a fixed spectral index would lead to systematic biases. However, for the AGN population, the calculation of K-corrections is inherently more complex, as AGN represent a heterogeneous class affected by numerous factors such as host galaxy type and inclination. In this study, apart from the radio excess criterion and the characteristics of LERG AGN within $0 \leq z \leq 4$ and $24 \leq \log L(1.4 \text{ GHz}) \leq 26$, no further subdivision of the AGN sample was performed to disentangle additional behaviors. As a result, future work should focus on a more detailed treatment of AGN, leveraging multiwavelength catalogs to achieve robust AGN classification and a better understanding of their RSED properties.

Furthermore, these results challenge the universality of the linear infrared-radio correlation (the so-called q-parameter). While this correlation has long been used to infer star formation rates and trace galaxy evolution, the findings here support recent studies demonstrating that q is not a constant, but instead varies with redshift, galaxy type, and physical conditions. This variation reflects the changing balance between infrared and radio emission mechanisms over cosmic history and underscores the necessity of population-dependent SED modeling—especially as SFR peaks at redshift $z \sim 2$, corresponding to the "cosmic noon" of star formation in the Universe.

By mapping the evolution of the RSEDs for both AGN and SFGs as a function of redshift, this thesis provides direct constraints on the changing nature of radio emission across cosmic time. Since redshift acts as a proxy for cosmic time (with higher redshifts corresponding to the earlier Universe), the population-level RSEDs traced here enable a deeper understanding of how star formation and black hole activity co-evolve and how their radio signatures are shaped by both internal and external (environmental) factors.

Importantly, the broken power-law parameterizations established here are not merely descriptive, but constitute a powerful tool for the wider community. They can be adopted in future work as population templates, allowing researchers to apply accurate K-corrections, test evolutionary models, and analyze large samples in ongoing and next-generation radio surveys. As such, the methodology and results of this thesis move the field forward—not only by delivering new measurements for the COSMOS field, but also by providing a practical framework for interpreting the complex radio emission processes that govern galaxy evolution on cosmological scales.

5.3 Outlook for Future Research

The results and methodology presented in this thesis lay the groundwork for a wide range of future research directions in the study of extragalactic radio sources. One immediate prospect involves the extension of the catalogue construction and matching techniques developed here to even larger samples or additional survey fields, further testing their robustness and adaptability. As new radio data from ongoing and planned surveys in the COSMOS field become available, these approaches can be readily applied or modified to incorporate increased depth, improved resolution, or additional frequency coverage. Applying these techniques to independent sky regions and different cosmic environments will also allow for a more thorough assessment of their universality and limitations.

A particularly promising direction is the systematic inclusion of ancillary data—such as detailed morphological classifications, higher-precision redshift estimates, and multi-wavelength photometric measurements—which will help to disentangle the complex nature of radio emission in both AGN and SFGs. Further methodological improvements, including more sophisticated error estimation, improved handling of blended or extended sources, and the automation of source classification, are expected to yield even more robust and homogeneous physical parameters. The cross-comparison of results across different fields and redshift ranges, as well as the integration of homogeneous spectroscopic and infrared data, will enable a more nuanced understanding of the population properties and their evolution.

Looking ahead, it will be important to expand the analysis to cover a broader range of radio frequencies, and to harness the potential of new statistical and computational tools, such as machine learning methods, for more effective source identification and SED modeling. By combining comprehensive datasets of higher sensitivity—achieved through additional observing hours with LOFAR, as well as the application of LOFAR VLBI techniques to obtain a circular beam and higher angular resolution, thereby enabling more accurate cross-matching with other catalogs—with advanced analysis techniques, future studies will be able to address outstanding questions about the drivers of radio emission in galaxies, test the generality of the trends revealed here, and provide deeper insights into the co-evolution of star formation and active galactic nuclei over cosmic time.

Chapter 6: References

- Blundell, K. M., & Rawlings, S. (2000, March). The Spectra and Energies of Classical Double Radio Lobes. *The Astronomical Journal*, 119, 12. doi:10.1086/301254
- Blundell, K. M., & Rawlings, S. (2002). Spectral ageing: a new age perspective. (R. A. Blundell, Ed.) *Particles and Fields in Radio Galaxies Conference*, 250, 9. doi:https://arxiv.org/abs/astro-ph/0209372
- Condon, J. (1992, September). Radio Emission from Normal Galaxies. *Annual Review of Astronomy and Astrophysics*, 30, 575-611. doi:https://doi.org/10.1146/annurev.aa.30.090192.003043
- Condon, J., Cotton, W., & Broderick, J. (2002, August). Radio Sources and Star Formation in the Local Universe. *The Astronomical Journal*, 124, 675-689. doi:10.1086/341650
- Darvish, B., Mobasher, B., Martin, D. C., Sobral, D., Scoville, N., Stroe, A., . . . Kartaltepe, J. (2017, March). Cosmic Web of Galaxies in the COSMOS Field: Public Catalog and Different Quenching for Centrals and Satellites. *Astrophysical Journal*, 837, 16. doi:https://doi.org/10.3847/1538-4357/837/1/16
- Darvish, B., Mobasher, B., Sobral, D., Rettura, A., Scoville, N., Faisst, A., & Capak, P. (2016, July). THE EFFECTS OF THE LOCAL ENVIRONMENT AND STELLAR MASS ON GALAXY QUENCHING TO $z \sim 3$. *The Astrophysical Journal*, 825, 113. doi:10.3847/0004-637X/825/2/113
- Delhaize, J., Heywood, I., Prescott, M., Jarvis, M., Delvecchio, I., Whittam, I., . . . Vaccari, M. (2021, March). MIGHTEE: are giant radio galaxies more common than we thought? *Monthly Notices of the Royal Astronomical Society*, 501, 3833–3845. doi:https://doi.org/10.1093/mnras/staa3837
- Delhaize, J., Smolčić, V., Delvecchio, I., Novak, M., Sargent, M., Baran, N., . . . Tasca, L. (2017, June). The VLA-COSMOS 3 GHz Large Project: The infrared-radio correlation of star-forming galaxies and AGN to $z \leq 6$. *Astronomy & Astrophysics*, 602, 17. doi:https://doi.org/10.1051/0004-6361/201629430
- Delvecchio, I., Smolčić, V., Zamorani, G., Del P. Lagos, C., Berta, S., Delhaize, J., . . . Tasca, L. (2017, June). The VLA-COSMOS 3 GHz Large Project: AGN and host-galaxy properties out to $z \leq 6$. *Astronomy & Astrophysics*, 602, 22. doi:https://doi.org/10.1051/0004-6361/201629367
- Fanaroff, B., & Riley, J. (1974, May). The morphology of extragalactic radio sources of high and low luminosity. *Monthly Notices of the Royal Astronomical Society*, 167, 31P-36P. doi:https://doi.org/10.1093/mnras/167.1.31P
- Heckman, T. M., & Best, P. N. (2014, June). The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe. *Annual Review of Astronomy and Astrophysics*, 52. doi:https://doi.org/10.1146/annurev-astro-081913-035722
- Helou, G., Soifer, B., & Rowan-Robinson, M. (1985, November). Thermal infrared and nonthermal radio: Remarkable correlation in disks of galaxies. *The Astrophysical Journal*, 298:2, L7-L11. doi:https://doi.org/10.1086/184556
- Jiménez-Andrade, E., Magnelli, B., Karim, A., Zamorani, G., Bondi, M., Schinnerer, E., . . . Koekemoer, A. (2019, May). Radio continuum size evolution of star-forming galaxies over $0.35 < z < 2.25^*$. *Astronomy & Astrophysics*, 625, 20. doi:https://doi.org/10.1051/0004-6361/201935178
- Jin, S., Daddi, E., Liu, D., Smolčić, V., Schinnerer, E., Calabrò, A., . . . Capak, P. (2018, September). "Super-deblended" Dust Emission in Galaxies. II. Far-IR to (Sub)millimeter Photometry and

- High-redshift Galaxy Candidates in the Full COSMOS Field. *The Astrophysical Journal*, 864:56, 32. doi:<https://doi.org/10.3847/1538-4357/aad4af>
- Kennicutt, R. C., & Evans, N. J. (2012, September). Star Formation in the Milky Way and Nearby Galaxies. *Annual Review of Astronomy and Astrophysics*, 50, 531–608. doi:<https://doi.org/10.1146/annurev-astro-081811-125610>
- Koekemoer, A., Aussel, H., Calzetti, D., Capak, P., Giavalisco, M., Kneib, J.-P., . . . Shopbell, P. (2007, September). The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing. *The Astrophysical Journal Supplement Series*, 172, 7. doi:<https://doi.org/10.1086/520086>
- Kormendy, J., & Ho, L. C. (2013, August). Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *Annual Review of Astronomy and Astrophysics*, 51, 136. doi:<https://doi.org/10.1146/annurev-astro-082708-101811>
- Laigle, C., McCracken, H., Ilbert, O., Hsieh, B., Davidzon, I., Capak, P., . . . Zabl, J. (2016, June). THE COSMOS2015 CATALOG: EXPLORING THE $1 < Z < 6$ UNIVERSE WITH HALF A MILLION GALAXIES. *The Astrophysical Journal Supplement Series*, 224, 26. doi:[10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- Lilly, S., Le Fevre, O., Renzini, A., Zamorani, G., Scodreggio, M., Contini, T., . . . Zucca, E. (2007, September). zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field. *The Astrophysical Journal Supplement Series*, 172, 16. doi:[10.1086/516589](https://doi.org/10.1086/516589)
- Machalski, J., Jamrozy, M., & Zola, S. (2001, May). The new sample of Giant radio sources: I. Radio imaging, optical identification and spectroscopy of selected candidates. *Astronomy & Astrophysics*, 371, 25. doi:<https://doi.org/10.1051/0004-6361:20010352>
- Magnelli, B., Ivison, R., Lutz, D., Valtchanov, I., Farrah, D., Berta, S., . . . Wuyts, S. (2015, January). The far-infrared/radio correlation and radio spectral index of galaxies in the SFR–M* plane up to $z \sim 2$. *Astronomy & Astrophysics*, 573, 18. doi:<https://doi.org/10.1051/0004-6361/201424937>
- Mohan, N., & Rafferty, D. (2015). PyBDSF: Python Blob Detection and Source Finder. *Astrophysics Source Code Library*, record ascl:1502.007. doi:<https://github.com/lofar-astron/PyBDSF>
- Murphy, E. J., Condon, J. J., Alberdi, A., Barcos-Muñozarcos, L., Beswick, R. J., Brinks, E., . . . Turner, J. L. (2018). Radio Continuum Emission from Galaxies: An Accounting of Energetic Processes. In E. Murphy (Ed.), *ASP Conference Series, Monograph 7*. 517, p. 421. San Francisco: Astronomical Society of the Pacific. doi:<https://ui.adsabs.harvard.edu/abs/2018ASPC..517..421M>
- Murphy, E., Condon, J., Schinnerer, E., Kennicutt, R., Calzetti, D., Armus, L., . . . Smith, J.-D. (2011, August). Calibrating Extinction-free Star Formation Rate Diagnostics with 33 GHz Free-free Emission in NGC 6946. *The Astrophysical Journal*, 737, 18. doi:<https://doi.org/10.1088/0004-637X/737/2/67>
- Rybicki, G., & Lightman, A. (1979). *Radiative processes in astrophysics*. A Wiley-Interscience Publication. Retrieved from <https://ui.adsabs.harvard.edu/abs/1979rpa..book.....R>
- Santoro, F., Tadhunter, C., Baron, D., Morganti, R., & Holt, J. (2020, December). AGN-driven outflows and the AGN feedback efficiency in young radio galaxies. *Astronomy & Astrophysics*, 644, 38. doi:<https://doi.org/10.1051/0004-6361/202039077>
- Sargent, M., Schinnerer, E., Murphy, E., Aussel, H., Le Floch, E., Frayer, D., . . . Scoville, N. (2010, January). The VLA-COSMOS Perspective on the Infrared-Radio Relation. I. New Constraints on Selection Biases and the Non-Evolution of the Infrared/Radio Properties of Star-Forming and Active Galactic Nucleus Galaxies at Intermediate and High Redshift. *The Astrophysical Journal Supplement Series*, 186, 341–377. doi:[10.1088/0067-0049/186/2/341](https://doi.org/10.1088/0067-0049/186/2/341)

- Schinnerer, E., Sargent, M., Bondi, M., Smolcic, V., Datta, A., Carilli, C., . . . Scoville, N. (2010, May). The VLA-COSMOS Survey. IV. Deep Data and Joint Catalog. *The Astrophysical Journal Supplement Series*, 188, 384–404. doi:<https://doi.org/10.1088/0067-0049/188/2/384>
- Scoville, N., Abraham, R., Aussel, H., Barnes, J., Benson, A., Blain, A., . . . Yan, L. (2007, September). COSMOS : Hubble Space Telescope Observations. *The Astrophysical Journal Supplement Series*, 172, 38–45. doi:<https://doi.org/10.1086/516580>
- Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C., Elvis, M., . . . Tyson, N. (2007, September). The Cosmic Evolution Survey (COSMOS): Overview. *The Astrophysical Journal Supplement Series*, 172, 1-8. doi:[10.1086/516585](https://doi.org/10.1086/516585)
- Shimwell, T., Hardcastle, M., Tasse, C., Best, P., Röttgering, H., Williams, W., . . . Zheng, J. (2022, March). The LOFAR Two-metre Sky Survey V. Second data release. *Astronomy & Astrophysics*, 659, 27. doi:<https://doi.org/10.1051/0004-6361/202142484>
- Smolcic, V., Delvecchio, I., Zamorani, G., Baran, N., Novak, M., Delhaize, J., . . . Vardoulaki, E. (2017, June). The VLA-COSMOS 3 GHz Large Project: Multiwavelength counterparts and the composition of the faint radio population. *Astronomy & Astrophysics*, 602, 21. doi:<https://doi.org/10.1051/0004-6361/201630223>
- Smolcic, V., Novak, M., Bondi, M., Ciliegi, P., Mooley, K., Schinnerer, E., . . . Sheth, K. (2017, June). The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release. *Astronomy & Astrophysics*, 602, 19. doi:<https://doi.org/10.1051/0004-6361/201628704>
- Smolcic, V., Novak, M., Delvecchio, I., Ceraj, L., Bondi, M., Delhaize, J., . . . Zamorani, G. (2017, June). The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since $z \sim 5$. *Astronomy & Astrophysics*, 602, 13. doi:<https://doi.org/10.1051/0004-6361/201730685>
- Tasse, C., Shimwell, T., Hardcastle, M., O’Sullivan, S., van Weeren, R., Best, P., . . . Wiaux, Y. (2021, April). The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. I. Direction-dependent calibration and imaging. *Astronomy & Astrophysics*, 648, 18. doi:<https://doi.org/10.1051/0004-6361/202038804>
- Tisanic, K., Smolčić, V., Delhaize, J., Novak, M., Intema, H., Delvecchio, I., . . . Vardoulaki, E. (2019, January). The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of highly star-forming galaxies*. *Astronomy & Astrophysics*, 621, 19. doi:<https://doi.org/10.1051/0004-6361/201834002>
- Tisanic, K., Smolčić, V., Imbrišak, M., Bondi, M., Zamorani, G., Ceraj, L., . . . Delhaize, J. (2020, November). The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of active galactic nuclei. *Astronomy & Astrophysics*, 643, 11. doi:<https://doi.org/10.1051/0004-6361/201937114>
- Toni, G., Maturi, M., Finoguenov, A., Moscardini, L., & Castignani, G. (2024, June). AMICO-COSMOS galaxy cluster and group catalogue up to $z = 2$: Sample properties and X-ray counterparts. *Astronomy & Astrophysics*, 687, 15. doi:<https://doi.org/10.1051/0004-6361/202348832>
- Vardoulaki, E., Gozaliasl, G., Finoguenov, A., Novak, M., & Khosroshahi, H. (2025, January). The evolution of the radio luminosity function of group galaxies in COSMOS. *Astronomy & Astrophysics*, 693, 14. doi:<https://doi.org/10.1051/0004-6361/202243713>
- Vardoulaki, E., Jiménez Andrade, E., Delvecchio, I., Smolčić, V., Schinnerer, E., Sargent, M., . . . Harrington, K. (2021, April). FR-type radio sources at 3 GHz VLA-COSMOS: Relation to physical properties and large-scale environment*. *Astronomy & Astrophysics*, 648, 33. doi:<https://doi.org/10.1051/0004-6361/202039488>
- Vardoulaki, E., Jiménez Andrade, E., Karim, A., Novak, M., Leslie, S., Tisanic, K., . . . Ciliegi, P. (2019, July). A closer look at the deep radio sky: Multi-component radio sources at 3 GHz

VLA-COSMOS. *Astronomy & Astrophysics*, 627, 31. doi:<https://doi.org/10.1051/0004-6361/201832982>

6.1 Online Reference Links

<https://cosmos.astro.caltech.edu/page/public> (Wallpaper)

<https://webbtelescope.org/contents/media/images/2021/042/01FD7GZBGRTJ7H8AVYGK3CP7ZV>
(Figure 6)

<https://webbtelescope.org/contents/media/images/2024/109/01HRD5S0V963RHPMPTC3MZ584N?news=true> (Figure 8)

<https://astronomy.swin.edu.au/cosmos/a/Active+Galactic+Nuclei> (Figure 9)

<https://science.nasa.gov/asset/hubble/compass-and-scale-image-of-hercules-a/> (Figure 10)

<https://cosmos.astro.caltech.edu/>

https://cosmos.astro.caltech.edu/page/ecole_lofar

<https://herscheltelescope.org.uk/results/cosmos-field/>

<https://www.sarao.ac.za/news/gigantic-galaxies-discovered-with-the-meerkat-telescope/>

<https://science.nrao.edu/ska>

<https://www.skao.int/en/explore/telescopes/ska-mid>

<https://www.skao.int/en/explore/telescopes/ska-low>

<https://www2.mpia-hd.mpg.de/COSMOS/skywalker/#>

<https://irsa.ipac.caltech.edu/frontpage/>

Chapter 7: Appendices

7.1 List of figures

Figure 1: Observed radio to far-infrared (FIR) spectrum of the starburst galaxy M82	5
Figure 2: Radio SED of radio-excess AGN (RxAGN) as a function of rest-frame frequency (ν_{rest}) for sources with $\log L_{1.4} \in [24.0, 26.0]$ and redshift $z \in [0, 4]$, based on the VLA-COSMOS 3 GHz Large Project (Tisanic, et al., 2020)	7
Figure 3: Radio SED of highly star-forming galaxies (HSFGs) as a function of rest-frame frequency (ν_{rest}), constructed from VLA and GMRT data in the COSMOS field (Tisanic, et al., 2019)	8
Figure 4: Comparison of the effective noise levels (converted to 1.4 GHz assuming a spectral index of $\alpha = 0.7$) versus surveyed area for existing and upcoming radio continuum surveys	10
Figure 5: Status of the LOFAR Two-metre Sky Survey (LoTSS) observations as of April 2021, showing the approximate current sensitivity coverage overlaid on the Haslam et al. (1982) 408 MHz all-sky map	11
Figure 6: The COSMOS Hubble ACS Field (NASA, ESA, Anton M. Koekemoer (STScI), Nick Scoville (Caltech))	14
Figure 7: Postage-stamp cutouts from the VLA-COSMOS 3 GHz continuum mosaic, imaged using the MSMF (Multi-Scale Multi-Frequency) algorithm	15
Figure 8: Multi-wavelength view of the starburst galaxy M82 (the “Cigar Galaxy”), combining Hubble (left) and James Webb (right) observations	18
Figure 9: Radio galaxy Cygnus A. (Image courtesy of NRAO/AUI; R. Perley, C. Carilli & J. Dreher)	21
Figure 10: Hubble Space Telescope view of the giant elliptical galaxy Hercules A (3C 348)	24
Figure 11: 610 MHz mosaic to the central 2 deg ² COSMOS field	33
Figure 12: 325 MHz mosaic to the central 2 deg ² COSMOS area	33
Figure 13: Composite multi-frequency images of GRG1	42
Figure 14: GRG-1 with the polygonal region across the field	43
Figure 15: Composite multi-frequency images of GRG-2	44

Figure 16: GRG-2 with the polygonal region across the field	45
Figure 17: Masks ID 1930 - sample of multiple sources are found within the radio island	46
Figure 18: This pie chart illustrates the distribution of sources in the final catalog according to the distinct combinations of Boolean flag categories	55
Figure 19: Distribution of the number of pixels (NumPixels) as a function of sensitivity (RMS) for all sources in the catalog, plotted on a log-log scale and separated by Boolean flag combination	55
Figure 20: Scatter plot of LOFAR mask-based flux density (Jy) versus signal-to-noise ratio (SNR) for all sources in the catalog, color-coded by Boolean flag combination	56
Figure 21: Scatter plot of SNR versus RMS (root mean square noise) for all sources, with both axes in logarithmic scale and points color-coded by Boolean flag combination	56
Figure 22: Detection statistics of HSFG sources in the GMRT catalogues	58
Figure 23: Flowchart summarizing the radio SED normalization and fitting (Tisanić, et al., 2019).	61
Figure 24: Reproduction of the average radio SED of radio-excess AGN	64
Figure 25: Reproduction of the average radio SED of radio-excess AGN with additional LOFAR data	65
Figure 26: Comparison of best-fit RSED curves for AGN	66
Figure 27: Reproduction of the average radio SED of highly star-forming galaxies (HSFGs) based on the methodology and dataset presented in Tisanić et al. (2020)	67
Figure 28: Refined average radio SED of star-forming galaxies (SFGs), following the methodology of Tisanić et al. (2019), with additional LOFAR data at 144 MHz	68
Figure 29: Comparison of best-fit radio SED curves for SFGs	69
Figure 30: Mask ID 650 - VLA ID 1221	70
Figure 31: Mask ID 750 - VLA ID 1034	70
Figure 32: Mask ID 1601 - VLA ID 1174	70

7.2 List of tables

<i>Table 1: Summary of the main properties for all 1085 sources</i>	54
<i>Table 2: Statistical Properties per Frequency Bin for the AGN Sample</i>	71
<i>Table 3: Best-Fit Model Parameters for the AGN Sample</i>	71
<i>Table 4: Statistical Properties per Frequency Bin for the SFG Sample</i>	72
<i>Table 5: Best-Fit Model Parameters for the SFG Sample</i>	72

7.3 MATLAB Pseudo-Codes

```
function Detect_And_Export_White_Regions()
imageData = load_FITS('image_fullband.int.restored_3GHz.fits')
threshold = 450e-6
binaryImage = imageData > threshold
[labels, numRegions] = label_connected_regions(binaryImage)
catalog = []
for i = 1 to numRegions
[rowIdx, colIdx] = find(labels == i)
area = length(rowIdx)
peakIdx = index_of_max(imageData at [rowIdx, colIdx])
centroid = (rowIdx[peakIdx], colIdx[peakIdx])
add_to_catalog(catalog, i, rowIdx, colIdx, area, centroid)
end for
save_catalog_to_mat(catalog, 'white_regions_catalog.mat')
fid = open_file('white_regions_catalog.reg')
write_DS9_header(fid)
for i = 1 to numRegions
centroid = catalog[i].centroid
id = catalog[i].ID
write_DS9_point(fid, centroid, id)
end for
close_file(fid)
display_region_summary(catalog)
visualize_regions(labels)
end function
```

Algorithm 1: Pseudocode for detection and cataloguing of binary regions above a flux threshold in a FITS radio image.

```
function Calculate_Region_Flux()
Load FITS image and region catalog
For each region in catalog:
Retrieve pixel coordinates
Sum flux values at those pixels
```

```

    Save all region flux sums to .mat file
end function

```

Algorithm 2:Pseudocode for calculating the total flux of each region from pixel values in a FITS image.

```

function Calculate_and_Save_Clean_Total_Flux()
    Load region flux values and region catalog
    For each region:
        Compute area in pixels
        Convert area to arcsec^2
        Subtract beam area to get corrected region area
        Calculate pixels per beam
        Compute total flux in Jy, correcting for beam area
    Save all results in .mat file
    Create results table and save as .csv
end function

```

Algorithm 3:Pseudocode for calculation and export of corrected total flux and area for extended sources.

```

function Build_Final_LOFAR_Catalogue_With_Radec()
    Load FITS image and region catalog
    For each region:
        Retrieve pixel coordinates
        Calculate total flux from pixels
        Get centroid pixel position
        Read WCS header info (CRPIX, CDELTA, CRVAL)
        Convert centroid pixel to RA/DEC using WCS projection
        Store ID, total flux, RA, DEC for each region
    Save RA/DEC results to intermediate file
    Load total flux and corrected area for regions
    For each region:
        Retrieve ID, RA, DEC, area, and total flux
        Set resolved/unresolved flag based on corrected area
        Store all fields in final catalogue
    Save final catalogue as .mat and .csv files
end function

```

Algorithm 4:Pseudocode for extracting total flux and celestial coordinates for each region, then assembling the final catalogue with resolved/unresolved classification.

```

function Match_Sources_To_Regions()
    Load mask region catalog and radio source catalog (with RA/DEC)
    Read WCS information from FITS header
    Convert all source RA/DEC to pixel coordinates (apply Y-axis flip)
    For each region:
        Find all sources whose pixel positions fall within the region's mask
        Store region ID, area, number of matches, and matched source IDs
    Export results to CSV file
    % Note: This process is repeated separately for the Counterpart, SuperDeblended,
    % and PyBDSF catalogs, to identify which and how many sources fall inside each mask.
end function

```

Algorithm 5:Step-by-step pseudocode for cross-matching mask regions with sources from different catalogs (Counterpart, SuperDeblended, PyBDSF) using WCS-based pixel conversion, with results exported to CSV for each run.

```

function Correct_Flux_With_Matching()
    Load final region catalogue and matched source tables
    Load SuperDeblend, FR, and Counterpart catalogues
    For each region:
        Get all matched source IDs (deblended, counterpart, FR)

```

```

    If multiple matched sources:
        If region matches FR source:
            Use FR position to select nearest SuperDeblend as central source
        Else if region has counterparts:
            Use nearest counterpart position as reference to select central SuperDeblend
        Else:
            Use LOFAR centroid to select nearest SuperDeblend as central
            Subtract fluxes of non-central SuperDeblend sources from original flux
        Else:
            Keep original flux
    Store corrected flux for each region
    Add corrected flux column to catalogue and export updated table as CSV
end function

```

Algorithm 6: Pseudocode for correcting region flux by subtracting secondary matched SuperDeblend sources, using hierarchical position matching.

```

function Calculate_Position_Errors_and_SNR()
    Load FITS image, region catalog, and flux table
    For each region:
        Find centroid pixel coordinates
        Compute local rms in small box around the source
        Calculate SNR as peak/rms
        Estimate RA and DEC error based on beam and SNR
    Add new columns for RA/DEC error, SNR, and RMS to table
    Save updated table as CSV
end function

```

Algorithm 7: Pseudocode for calculating RA/DEC uncertainties, SNR, and RMS for each region, then exporting all results to a new catalogue.

```

function Update_Catalogue_For_GRGs()
    Load final catalogue with errors
    For each GRG (e.g., GRG1, GRG2):
        Define center source and secondary member IDs
        For the center:
            Add flux and area from all secondary sources
            Change ResolvedFlag to GRG label
            Remove SNR and RMS values
        Remove secondary sources from the catalogue
    Save updated catalogue to new CSV
end function

```

Algorithm 8: Pseudocode describing the process of combining GRG member fluxes and pixel counts into their core region, updating relevant columns, and exporting the modified catalogue.

```

function Construct_Clean_Matched_Catalogue()
    Load main LOFAR region catalogue and multiple external radio source catalogs (Counterpart, SuperDeblend, PyBDSF, FR)
    Prepare and rename all variable fields to avoid naming conflicts
    For each region in the LOFAR catalogue:
        Apply custom mask-matching logic to associate each region with matching sources from all catalogs
        - Forced matches and exclusions for specific region IDs
        - For Counterpart matches: Prefer FR sources, otherwise pick nearest counterpart
        - For Deblended matches: Use counterpart position as reference if available, otherwise use LOFAR centroid
        - For PyBDSF matches: Use reference position from counterpart or deblended source
    Copy all relevant columns from matched external sources into the main catalogue row

```

```

    Mark region as matched and to be retained
    Remove all unmatched regions and unwanted IDs
    Merge selected multi-component sources into single entries, summing flux and area as needed
    Rescale flux columns to common units
    Save final cleaned and merged catalogue as CSV
    Identify and report sources with near-zero or negative deblended flux
end function

```

Algorithm 9: Pseudocode for cross-matching LOFAR regions to external catalogs with hierarchical rules, handling special cases and merges, and producing a cleaned, merged catalogue.

```

function Construct_AGN_Radio_SED_and_Fits()
    Load main AGN sample
    Filter: Radio_excess_1_1 == true,  $0 \leq z \leq 4$ ,  $24 \leq \log L(1.4\text{GHz}) \leq 26$ 
    Extract fluxes and rms (325, 610, 1400, 3000 MHz)
    Set upper limits:  $5 \times \text{RMS}$  where flux is missing
    Compute rest-frame frequencies
    Normalize all fluxes to 4.1 GHz (interpolation between 1.4 & 3 GHz)
    Discard sources with invalid normalization
    Calculate log-normalized flux for all data
    Mark detections and upper limits
    Concatenate all frequency-log(flux) pairs
    Bin data by frequency, calculate means and errors (detections & survival analysis)
    Add 144 MHz LOFAR data (normalize as above)
    Assign FR source fluxes where needed
    Calculate 144 MHz upper limits using  $5 \times \text{RMS}$ 
    Merge all binned statistics (main bands & LOFAR)
    Fit broken power law (BPL) to binned means
    Fit curved polynomial to all binned means (incl. 144 MHz)
    Fit combined absorption model
    Plot results: raw points, binned means, all fit curves, confidence regions, inset near break
    Export statistics and fit parameters to .xlsx
end function

```

Algorithm 10: Step-by-step construction of the normalized Radio SED for AGN.

```

function Construct_SFG_Normalized_RSED()
    Load main SFG dataset (Tsanic et al. cross-matched array)
    Load LOFAR SFG sample with fluxes at 144 MHz and other frequencies
    Filter both datasets to keep only finite, positive flux values
    For LOFAR SFGs:
        Set normalization reference frequency (e.g., 5.7 GHz)
    For each source:
        Normalize all available flux densities to reference frequency
    For main SFGs:
        Assign upper limits at 325 and 610 MHz based on RMS columns
    For each source **not matched** to a LOFAR SFG:
        If detected at 325 MHz:
            Extrapolate to 144 MHz using power-law slope
            Assign as upper limit at 144 MHz
    Combine all fluxes, frequencies, and upper limit flags from main and LOFAR SFGs
    Bin data in logspace along  $\nu_{\text{rest}}$  axis
    For each bin:
        Calculate mean and standard deviation (detections only)
        Apply survival analysis (e.g., Kaplan-Meier/Weibull) to include upper limits
    Fit broken power-law (BPL) to binned means (break fixed at 4.3 GHz)
    Fit curved polynomial model to all binned means

```

```
Fit absorption (combined) model to all binned means
Save main bin statistics (counts, means, upper limits) as table
Save best-fit model parameters as table
Generate and export final SFG RSED plots (with all model curves and confidence regions)
end function
```

Algorithm 11: Pseudocode for the construction of the normalized SFG Radio SED.

