VLA-COSMOS 3 GHz Large Project

Novak, Mladen

Doctoral thesis / Disertacija

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:195872

Rights / Prava: In copyright/Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2025-03-14**

University of Zagreb

Faculty of Science

Department of Physics

Mladen Novak

VLA-COSMOS 3 GHz Large Project:

Star formation through cosmic time

DOCTORAL THESIS

Zagreb, 2017

University of Zagreb

Faculty of Science

Department of Physics

Mladen Novak

VLA-COSMOS 3 GHz Large Project:

Star formation through cosmic time

DOCTORAL THESIS

Supervisor:

assoc. prof. dr. sc. Vernesa Smolčić

Zagreb, 2017

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Fizički odsjek

Mladen Novak

VLA-COSMOS 3 GHz Large Project:

Stvaranje zvijezda kroz kozmičko vrijeme

DOKTORSKI RAD

Mentor:

izv. prof. dr. sc. Vernesa Smolčić

Zagreb, 2017

Supervisor information

Vernesa Smolčić was born February, 23$^{\text{rd}}$, 1980, and her main area of research is observational astrophysics. In 2004. she attained a Masters degree in physics from the Department of Physics, Faculty of Science, University of Zagreb (Zagreb, Croatia), and in 2007 a doctoral degree from the Ruperto-Carola University (Heidelberg, Germany). Her doctoral research was conducted at the Max-Planck Institut für Astronomie (Heidelberg, Germany). From 2008 to 2010 she worked as a CARMA Postdoctoral Scholar at the California Institute of Technology (Caltech, Pasadena, CA, USA), and from 2010 to 2013 as an ESO ALMA COFUND Fellow (Argelander Institute for Astronomy, Bonn, Germany; European Southern Observatory, Garching, Germany). She became an assistant professor at the Department of Physics, Faculty of Science, University of Zagreb in 2010, and was promoted to associate professor in 2015. She held stipends to conduct research at Australian universities and institutes (ICRAR visiting fellowship for senior women in astronomy, 2015, and Group of Eight Fellowship, 2012).

Vernesa Smolčić is (co-)author of 101 refereed publications in Current Content journals with over 12 000 citations (at 34, with over 1 150 citations, she is the only, leading or co-leading author; her H-index is 44 based on NASA ADS). She presented over 50 scientific talks at institutes and universities around the globe, international conferences, workshops and symposia (over 20 invited talks/institute colloquia). She is regularly refereeing for Current Content journals (MNRAS, ApJ, AJ) and national foundations, and has been a member of CARMA, Chandra, OPTICON, Gemini, LMT time allocation committees, and was chair/member of 10 conference scientific and local organizing committees. She is a member, and coordinator of the Croatian science foundation physics panel (since 2016), member of the Institute of Physics (Zagreb, Croatia) governing council (since 2016), member of the COSMOS project "scientific steering committee" (since 2012), member of the XXL project "scientific steering committee" (since 2016), head of the Scientific Section of the Croatian Physical Society (since 2014), head of the graduate studies module "Astrophysics" within "Atomic and molecular physics and astrophysics" at University of Zagreb (since 2014), and member of the Graduate studies board at the University of Zagreb (since 2014). She has mentored nine Masters and three PhD theses,

published over ten refereed papers in co-authorshiop with undergaduate, graduate students, and postdocs, and introduced two new undergraduate and one new graduate course at the University of Zagreb.

She has held the European Research Council Starting Grant (2014. – 2018., 1.5 M€), Marie Curie Career Integration Grant (2013. – 2017.; 100 k€) and Unity Through Knowledge Fund "Homeland Visit" grant (2009.; ∼10 k€). During her carrier she obtained the following awards: University of Zagreb Rector's award for best students research projects (2003), Ernst Patzer Award for best publications of young MPIA scientist (2006), Zagreb female person of the year awarded by the City of Zagreb (2014), Faculty of Science award for significant science contribution and raising the reputation of the Faculty of Science (2014), nomination for the person of the year by the newspaper Večernji list (2014 and 2015), COSMO scientist of the year awarded by the Croatian Cosmopolitan magazine (2015), nomination for the female person of the year by the magazine Zaposlena (2017).

Acknowledgements

I would like to express a sincere gratitude to my supervisor Vernesa Smolčić for her guidance throughout my PhD.

I would like to thank all members of the (VLA-)COSMOS collaboration who made this thesis possible. A special mention to Eva Schinnerer, Gianni Zamorani, Marco Bondi, Paolo Ciliegi, Alex Karim and Mark Sargent for all their help, ideas, comments and discussions. I am also grateful to Felipe Navarrete, Kunal Mooley, Chris Carilli, Steve Myers, Francesca Civano, Julie Banfield and Chris Hales for sharing their expertise and insights with me.

Special thanks to my friends and colleagues Nikola, Oskari, Jacinta, Ivan, Lana and Krešimir for making all the work fun (let's thank them all again).

Finally, I wish to thank my parents, brother, Vera and Sophie for their continuous support, care and warmth.

Abstract

Radio continuum emission from approximately ten thousand galaxies was observed at 3 GHz (10 cm) and used to trace galaxy evolution and the dust unbiased cosmic star formation rate history since the universe was only 1 billion years old. Observations lasting almost 400 hours were performed with the upgraded Karl G. Jansky Very Large Array (VLA) interferometer pointed toward the 2 square degree equatorial COSMOS field. This VLA-COSMOS 3 GHz Large Project represents the currently deepest radio continuum observations performed across such a large area of the sky. Radio sources were paired with optical/near-infrared counterparts and a statistical approach was used to determine luminosity functions of two main galaxy populations: star-forming galaxies and active galactic nuclei. Number counts of radio galaxies were obtained, which suggest that the future deep radio surveys will predominantly observe star-forming galaxies. The redshift dependent infrared-radio correlation was used to link radio emission with star formation processes. This research suggests that galaxies were the most productive when the universe was 2.5 billion years old (at a redshift of $z \sim 2.5$), only a fifth of its current age. During that period, about 25% of all newborn stars were being created in massive galaxies where star formation rates can reach thousands of solar masses per year. Up to 20% more star formation might be occurring in highly dust obscured galaxies in the early universe, compared to what was previously thought.

Keywords: galaxies, galaxy evolution, radio emission, interferometry, luminosity function, star formation, cosmic star formation history, active galactic nuclei, COSMOS field, Very Large Array

Prošireni sažetak

Cilj ovog rada je pomoću dubokih radio opažanja pratiti evoluciju galaksija te odrediti kozmičku povijest stvaranja zvijezda od trenutka kad je svemir bio star otprilike jednu milijardu godina. U tu su svrhu napravljena radio opažanja u trajanju od gotovo 400 sati pomoću *Karl G. Jansky Very Large Array* (VLA) radio interferometra uperenog u COSMOS ekvatorijalno polje na nebu veličine dva kvadratna stupnja. Ova opažanja su dio VLA-COMOS 3 GHz velikog projekta (engl. *Large Project*) te predstavljaju trenutno najdublja mjerenja radio kontinuuma učinjena preko površine veće od jednog kvadratnog stupnja. Dugo trajanje mjerenja i površina COSMOS polja omogućavaju opažanje dovoljno velikog uzorka galaksija potrebnog za ovakvu vrstu analize. Prednost radio opažanja leži u činjenici da radio valovi ne međudjeluju sa sredinom kroz koju prolaze te se dobivaju opažanja bez utjecaja prašine koja inače zatamnjuje svjetlost kraćih valnih duljina.

Ključne riječi: galaksije, evolucija galaksija, radio zračenje, interferometrija, funkcija luminoziteta, stvaranje zvijezda, kozmička povijest stvaranja zvijezda, aktivna galaktička jezgra, COSMOS polje, Very Large Array

Osnovna svojstva galaksija i njihova evolucija

Galaksije su veličanstveni objekti u svemiru i predstavljaju najveće nakupine vidljive materije. Sastavljene su od zvijezda, plina i prašine koji kruže oko supermasivne crne rupe pod utjecajem privlačne gravitacijske sile (Sparke & Gallagher, 2007). Iako je svaka galaksija različita, one dijele mnoga svojstva koja astronomi pokušavaju klasificirati, korelirati, razumjeti i konačno predvidjeti. U kratkom pogledu prema nebu može se vidjeti daleko u povijest svemira zbog konačne brzine svjetlosti i dugog putovanja fotona od udaljenog objekta do promatrača. Taj efekt nam daje uvid u događaje koji su davno prošli, ali su još uvijek utkani u svjetlosti koja je putovala milijarde godina da nas dosegne. Galaktički procesi, kao što su spajanje galaksija, stvaranje zvijezda i kemijska evolucija, mjere se u milijunima godina i ne mogu se promatrati

u stvarnom kontinuiranom vremenu (Mo et al., 2010). Također, može se opažati samo jednu dostupnu orijentaciju svake pojedinačne galaksije. Proučavanje evolucije galaksija svodi se na smisleno povezivanje presjeka različitih vremenskih epoha svemira. Konačni rezultat je dublje razumijevanje svemira i njegove povijesti.

Model svemira

Da bi se galaksije moglo smjestiti u njihovo pravo vrijeme i prostor potreban je model svemira. Trenutno prihvaćeni kozmološki model je svemir s Lambdom i hladnom tamnom materijom (ΛCDM, engl. *Lambda cold dark matter*) koji pretpostavlja da je gravitacija ispravno opisana općom relativnošću (Mo et al., 2010). U ovom modelu svemir je započeo Velikim praskom u stanju iznimno velike gustoće energije, temperature i tlaka, nakon koje slijedi eksponencijalna inflacija. Da bi se inflacija mogla odvijati, vakuumska energija, koja je označena s Λ, mora biti različita od nule. Tijekom razdoblja inflacije male (kvantne) toplinske fluktuacije predstavljaju sjeme iz kojeg su izrasle danas vidljive strukture u svemiru. Jedan od temelja ovog kozmološkog modela je pozadinsko mikrovalno zračenje (Penzias & Wilson, 1965) koje pokazuje da je svemir homogen i izotropan na velikim skalama, što upućuje na kauzalnu povezanost cijelog svemira u početku. Tamna materija je uvedena da bi se objasnili slučajevi u kojima vidljiva materija nije bila dovoljna da objasni potrebnu dubinu gravitacijskog potencijala. Mjera Dopplerovog crvenog pomaka predstavlja dob svemira i udaljenost od objekta unutar kozmološkog modela. Prema trenutnom modelu svemir je star 13,8 milijardi godina, ravan te sastavljen od 71.8% tamne energije, 23.6% hladne tamne tvari, i 4.6% vidljive materije (Hinshaw et al., 2013).

Galaksije

Edwin Hubble je među prvima pokušao klasificirati galaksije te se koristio morfološkim kriterijima poput oblika galaksije, izgleda diska i spirala (Hubble, 1936). Kategorije su bile prilično široke, što je omogućilo lako proširivanje sustava i njegov opstanak do danas. Primjećuje se i da su spiralne galaksije uglavnom plave boje, dok su eliptične uglavnom crvenije, odnosno postoji dualnost galaksija. Hubble se koristio slikama galaksija u optičkom valnom pojasu, međutim izgled galaksija se može značajno razlikovati u ovisnosti o promatranoj valnoj duljini.

Galaksije u ranom svemiru su jako udaljene te se ne mogu uvijek razlučiti detalji zbog tehničkih ograničenja teleskopa. U tom slučaju jedina preostala opcija je promatrati integrirano svjetlo cijele galaksije, tzv. fotometrija. Skoro svo zračenje galaksije dolazi od zvjezdane svjetlosti, bilo izravne ili zatamnjene prašinom, te uslijed akrecijskih procesa oko supermasivne crne rupe. Mlade masivne zvijezde dominiraju zračenjem u ultraljubičastom (UV, engl. *ultra-*

violet) i plavom pojasu elektromagnetskog spektra, dok će stara zvjezdana populacija i manje masivne zvijezde davati crveniju svjetlost. Prašina apsorbira vidljivu i UV svjetlost te energiju zrači nazad u infracrvenom (IR, engl. *infrared*) dijelu spektra. Mjerenjem zračenja u različitim valnim pojasevima dobiva se spektralna raspodjela energije (SED, engl. *spectral energy distribution*) galaksije (vidi Conroy, 2013). Galaksija u svom središtu ima supermasivnu crnu rupu čije gravitacijsko djelovanje snažno privlači i skuplja okolnu materiju. Plin i prašina u tom procesu doživljavaju veliku promjenu potencijalne energije koju moraju zračiti. To zračenje se klasificira kao aktivna galaktička jezgra (AGN, engl. *active galatcic nucleus*). Temperatura plina u tim uvjetima doseže milijune kelvina te dolazi do emisije X-zraka. Pod određenim uvjetima AGN može proizvesti kolimirani mlaz relativističkih čestica koje emitiraju radio valove dok se kreću kroz magnetsko polje galaksije. Proučavanje galaksija preko cijelog elektromagnetskog spektra, od gama zraka do radio valova, je nužno za otkrivanje dominantnih fizikalnih procesa. Sintetski ili opaženi predložci spektara mogu se prilagoditi na fotometriju galaksije iz čega se određuju njena svojstva poput crvenog pomaka, zvjezdane mase i količine prašine.

Ukupna gustoća zvjezdane mase svemira se udvostručila od prije 7 milijardi godina, ali rast je uglavnom zabilježen u pasivnim crvenim galaksijama, dok se u plavim galaksijama nije puno mijenjala (Borch et al., 2006), iako se upravo u njima stvaraju nove zvijezde. U trenutno prihvaćenoj evolucijskoj slici plave galaksije stvaraju zvijezde iz plina u ovisnosti o njihovoj masi (tzv. glavni niz galaksija), dok se okolina ne promijeni i ugasi taj proces, npr. smanjenjem dotoka kozmičkog plina u galaksiju ili interakcijom s drugom galaksijom (Schawinski et al., 2014). Nakon tog trenutka zvjezdana populacija stari i galaksija prelazi u crveni niz gdje raste daljnjim spajanjima. Plave galaksije se koriste za praćenje izgradnje ukupne zvjezdane mase svemira te se stoga i nazivaju galaksije koje stvaraju zvijezde (engl. *star-forming galaxies*). Opažanja pokazuju da su galaksije stvarale najviše zvijezda prije otprilike 10 milijardi godina (Madau & Dickinson, 2014). Procesi akrecije na supermasivnu crnu rupu mogu biti toliko dominantni da zasjene svo zračenje koje dolazi iz zvijezda. Takve galaksije se zovu AGN-ovi te njihovo proučavanje daje uvid u regulacijske mehanizme galaksije. Epoha najefikasnije akrecije AGN-ova se podudara s epohom najefikasnijeg stvaranja zvijezda (Delvecchio et al., 2014), što ukazuje na povezanost fizikalnih procesa u galaksijama.

Radio astronomija

Početak radio astronomije može se staviti u trenutak kada je Karl G. Jansky prilikom proučavanja šuma u prekooceanskim telefonskim komunikacijama otkrio izvanzemaljski radio signal iz smjera centra Mliječnog puta (Jansky, 1933). Narednih par desetljeća su pomoću radio va-

lova otkriveni egzotični objekti, npr. ostaci supernova, pulsari i kvazari. Radio teleskop je instrument dizajniran za prikupljanje elektromagnetskog zračenja dugih valnih duljina (većih od centimetra). Kutna razlučivost teleskopa je difrakcijski ograničena Rayleighevim kriterijem koji je proporcionalan s valnom duljinom i obrnuto proporcionalan veličinom aperture. Veliki iskorak u radio opažanjima je postignut upotrebom interferometrijskih principa za aperturnu sintezu, za što je Martin Ryle nagrađen Nobelovom nagradom 1974. godine. Najmanji funkcionalni element u tom postavu je jedan par antena, tzv. bazna linija (engl. *baseline*). Radio val s nebeskog izvora putuje dvama različitim putanjama do svake od antena i geometrijska razlika u putevima uzrokuje fazni pomak između signala, što rezultira njihovom interferencijom. Kutna razlučivost interferometra je obrnuto proporcionalna razmaku dviju antena. Za razliku od teleskopa s jednom antenom, interferometar opaža Fourierove komponente prostorne raspodjele svjetlosnih izvora na nebu. Kako se radi o konačno mnogo diskretnih komponenti, konačna mapa radio izvora sadržava artefakte zbog nepotpunog uzorkovanja Fourierovih komponenti. Proces analize, odnosno redukcije, interferometrijskih podataka obuhvaća uporabu različitih tehnika za kalibraciju, korekcije i obradu radio slike.

Uzroci radio zračenja

Emisija radio valova u galaksijama potječe iz dva različita fizikalna procesa. Prvi je termičko zakočno zračenje (njem. *Bremsstrahlung*) od elektrona koji se raspršuju na ionima te se javljaju u područjima ioniziranog vodika (HII), npr. oko mladih vrućih zvijezdi. Drugi je netermičko sinkrotronsko zračenje iz relativističkih elektrona koji se gibaju po spiralnim putanjama u magnetskim poljima (Condon & Ransom, 2016). Velika energija je potrebna za akceleriranje elektrona do potrebnih brzina koja se može osloboditi npr. prilikom eksplozije supernova. Kombinacija svih pojedinačnih elektronskih spektara daje ukupnu radio SED koja se često parametrizira s jednostavnim eksponentom, tzv. spektralnim indeksom α, na način $S_\nu \propto \nu^\alpha$ ili $\alpha = d\ln S_\nu / d\ln \nu$, gdje je S_ν gustoća toka emitiranog na frekvenciji ν. Za mjerenje stvaranja zvijezda važno je da radio zračenje doista uzrokuju zvjezdani procesi jer se prilikom nakupljanja materije na supermasivnu crnu rupu također oslobađa ogromna količina energije.

Iako se različiti zvjezdani procesi (kolaps plina, fotoionizacija tvari oko zvijezde, apsorpcija svjetlosti na zrncima prašine i njezina ponovna emisija, eksplozija supernove i difuzija kozmičkih zraka) odvijaju na raznim vremenskim skalama, uočena je čvrsta empirijska korelacija između radio i IR zračenja (Bell, 2003). Preko te korelacije, poznavajući uzroke IR zračenja, moguće je povezati radio luminozitet s brzinom stvaranja novih zvijezda u galaksiji. Ova korelacija vrijedi u galaksijama koje stvaraju zvijezde, dok odstupanja mogu upućivati na prisustvo AGN-a. Za razliku od optičkih i UV opažanja, zrnca prašine zbog svojih malih dimenzija ne

utječu na radio valove.

Pregledi neba i brojčana gustoća galaksija

Nebo se prostire na preko $40\,000$ deg^2 i svaki kvadratni stupanj može sadržavati stotine tisuća galaksija. Veliko područje neba treba pokriti u potrazi za rijetkim objektima, dok istraživanje mladog svemira znači opažanje udaljenih i slabih izvora. Pregledi neba su stoga organizirani u više razina na način da se pokriju sve relevantne skale (vidi Condon, 2015b). Važno je opažati dovoljno veliki volumen koji pokriva što više mogućih kozmičkih sredina jer samo tako uzorak galaksija može biti reprezentativan za cijeli svemir.

Početnu analizu novih opažanja je najbolje uraditi jednostavnim statističkim metodama. Jedna od metoda uključuje prebrojavanje izvora u ovisnosti o toku zračenja. U Euklidskom prostoru uniformno popunjenom galaksijama diferencijalni broj galaksija bio bi $n = dN/dS \propto S^{-5/2}$, gdje je S tok zračenja. Ovakva ovisnost se ne opaža iz razloga što je prostor opisan općom teorijom relativnosti i širi se, a gustoća galaksija se mijenja kroz kozmičko vrijeme zbog njihove evolucije. Naprednija metoda prebrojavanja koja pruža obilje informacija o evoluciji galaksija je funkcija luminoziteta. Za njezin izračun potrebne su udaljenosti do galaksija, tj. crveni pomaci. Funkcija luminoziteta mjeri kozmičku gustoću izvora (broj galaksija po kubičnom megaparseku - mpc^3, gdje je 1 pc $\approx 3.1 \times 10^{16}$ m) po logaritmu luminoziteta. Intrinzični luminozitet galaksije je obično koreliran sa zanimljivim svojstvima poput fizičke veličine, zvjezdanom masom ili AGN aktivnosti. Model evolucije galaksija trebao bi biti u stanju reproducirati opažene brojčane gustoće galaksija, ali i predvidjeti što će detektirati budući dublji radio pregledi neba. Konačno, razmatrajući funkcije svjetlosti dobivene na različitim valnim duljinama (npr. radio, IR, UV), moguće je povezati različite populacije galaksija u jedinstvenu koherentnu sliku.

Opažanje COSMOS polja Very Large Array teleskopom

Ova radnja usredotočena je na COSMOS pregled neba (engl. *Cosmic Evolution Survey*, Scoville et al. 2007) opažanog u radio valnom području s nedavno nadograđenim VLA interferometrom. Ta nadogradnja je povećala osjetljivost opažanja radio kontinuuma pet do dvadeset puta. COSMOS pregled je pankromatsko opažanje ekvatorijalnog nebeskog polja površine 2 kvadratna stupnja (oko 10 punih Mjeseci) centriranog na koordinatama 10:00:28.6 +02:12:21.0 (u J2000 epohi). Opažanja COSMOS polja su urađena mnogim suvremenim teleskopima i satelitima od radio područja do X-zraka te rezultiraju bogatim podacima za preko milijun objekata s precizno određenim crvenim pomacima (Laigle et al., 2016). Područje neba je dovoljno veliko da nema

poteškoća uslijed promatranja nereprezentativnog volumena svemira (Moster et al., 2011), a njegov položaj na nebu odabran je da bude bez sjajnih izvora. Istraživanje COSMOS polja je optimizirano za proučavanje slabih izvora, a trenutni podaci dosežu epohu kada je svemir bio star samo milijardu godina. Novi dodatak COSMOS pregledu, oko kojeg je ovaj rad izgrađen, je VLA-COSMOS 3 GHz Large Project koji predstavlja gotovo 400 novih radio sati opažanja na 3 GHz (10 cm). S preko 5 000 radio izvora po kvadratnom stupnju opaženih rezolucijom manjom od jedne lučne sekunde, ovi podaci su optimalni za praćenje kozmičke povijesti stvaranja zvijezda radio mjerenjima koja nisu zatamnjena prašinom.

Prvih 130 sati opažanja i studija crne rupe u mogućem procesu odboja

Opažanja VLA-COSMOS na 3 GHz provođena su od studenog 2012. do svibnja 2014. Kako su podaci pristizali, bilo je potrebno testirati različite kalibracije i algoritme za stvaranje radio slika s obzirom da je teleskop nedavno nadograđen te mjeri kontinuum od 2 do 4 GHz. Testiranje dolaznih podataka bilo je kombinirano sa znanstvenim ciljem istraživanja egzotične galaksije CID-42: supermasivne crne rupe u mogućem procesu odbijanja.

Odboj supermasivne crne rupe

Tijekom spajanja galaksija, središnje supermasivne crne rupe (SMBH, engl. *supermassive black hole*) koje se nalaze unutar galaksija formiraju vezani dvojni SMBH sustav (Volonteri et al., 2003). Prilikom daljnjeg srastanja emitira se jako i usmjereno gravitacijsko zračenje koje ovisi o spinu i omjeru masa crnih rupa. Zbog zakona očuvanja količine gibanja, novostvorena SMBH može doživjeti odboj (engl. *recoil*) i biti izbačena iz galaksije (Peres, 1962), prilikom čega povlači sa sobom i najbližu okolnu materiju (akrecijski disk).

    Chandra-COSMOS izvor CXOC J100043.1+020637 ($z = 0.359$, Civano et al. 2012a), poznat i kao CID-42, je kandidat za ovakav proces s indikacijama u optičkim mjerenjima i X-zrakama. U galaksiji CID-42 vide se dvije komponente odvojene za $0.5''$ (2.5 kpc), dok je samo jedna komponenta (jugo-istočna) zaslužna za skoro sve opažene X-zrake. U optičkom spektru CID-42 izmjerena je razlika u brzinama između široke i uske komponente Hβ linije od oko $1\,300\,\mathrm{km\,s^{-1}}$. Podaci su u skladu sa scenarijem odboja SMBH koji se dogodio prije otprilike 1 – 6 milijuna godina (Blecha et al., 2013), iako se ne može potpuno isključiti scenarij gdje je i druga komponenta u galaksiji skrivena SMBH.

Novi radio podaci

Prvih 130 sati opažanja COSMOS polja s VLA teleskopom je bilo dostupno u trenutku izrade ove analize. S obzirom da opažanja sadrže mjerenja od 2 do 4 GHz, podaci su obrađeni na dva načina. Prvi pristup se svodi na podjelu cijelog valnog pojasa u uže spektralne prozore (široke 128 MHz), čije se slike potom slažu jedna na drugu. Prednost je ta što unutar užeg pojasa izvor ima približno konstantan tok zračenja, međutim smanjuje se omjer signala i šuma (S/N, engl. *signal-to-noise ratio*) u pojedinom spektralnom prozoru. Također se rezolucija mjerenja mijenja s frekvencijom te je sve spektralne prozore potrebno uskladiti na zajedničku (najlošiju) rezoluciju koja odgovara najnižoj frekvenciji. Drugi pristup koristi višefrekvencijsku sintezu (Rau & Cornwell, 2011) gdje se cijeli valni pojas razvija u Taylorov red, što koristi veći obujam mjerenja i konačna rezolucija može biti bolja. Međutim, ova metoda je razvijena za nove generacije teleskopa i nije još u potpunosti provjerena. Obje metode koriste rutinu CLEAN iz programskog paketa CASA (engl. *Common Astronomy Software Applications*, McMullin et al. 2007). Osim CID-42, još osam radio izvora koji imaju S/N između 5 i 230 je analizirano te je utvrđeno da su obje metode dobivanja radio slike konzistentne međusobno, kao i s dostupnim radio podacima na nižim frekvencijama.

Diskusija radio spektra CID-42

Radio zračenje CID-42, koje je pouzdano (7σ, gdje je σ lokalni šum u radio mapi) se podudara s komponentom koja zrači u X-zrakama. Najviši površinski sjaj (engl. *peak surface brightness*) izmjeren u novim radio podacima iznosi $33.2 \pm 4.8\ \mu\mathrm{Jy\,beam}^{-1}$ u složenoj slici spektralnih prozora na rezoluciji od $0.9''$, te $32.5 \pm 4.6\ \mu\mathrm{Jy\,beam}^{-1}$ u slici dobivenoj iz frekvencijske sinteze na rezoluciji od $0.7''$. To odgovara luminozitetu u sustavu mirovanja od $L_{3\,\mathrm{GHz}} \approx 1 \times 10^{22}\ \mathrm{W\,Hz}^{-1}$, što upada u radio tihi režim. Unutar valnog pojasa od 2 do 4 GHz CID-42 ima ravan radio spektar, što je konzistentno s izostankom VLBA (engl. *Very Large Baseline Array*) detekcije na 1.4 GHz (Wrobel et al., 2014) koja je urađena na vrlo visokoj rezoluciji ($0.015''$), što se može interpretirati kao radio tihi AGN. Ekstrapolacijom i usporedbom s prošlim VLA opažanjima (Schinnerer et al., 2007), na frekvenciji od 1.4 GHz opaža se višak radio zračenja u iznosu od 105 μJy koji mora imati neobično strm spektar da bi iščeznuo do 3 GHz. Višak bi moglo uzrokovati šire područje populacije elektrona koji su već izračili većinu svoje energije (što rezultira strmim radio spektrom) ili šokovi uslijed tokova iz AGN-a. Ukoliko i druga komponenta CID-42 zrači u radio valovima, ono mora biti ispod $L_{3\,\mathrm{GHz}} < 5.6 \times 10^{21}\ \mathrm{W\,Hz}^{-1}$. Ova radio opažanja još uvijek ne mogu potvrditi radi li se zapravo o gravitacijskom odboju crne rupe, ili o dva AGN-a na malom razmaku, ali nova ograničenja su stavljena na sustav CID-42 koja će

pomoći budućim istraživanjima ovog specifičnog objekta.

VLA-COSMOS 3 GHz Large Project: podaci i katalog

Mjerenja, kalibracija i dobivanje radio mape

Ukupno 384 sata opažanja COSMOS polja je urađeno VLA teleskopom u S-pojasu na 3 GHz s ukupnom širinom od 2048 MHz, podijeljenom u 16 spektralnih prozora širine 128 MHz. Mjerenja su rađena od studenog 2012. do siječnja 2013., od lipnja do kolovoza 2013. te od veljače do svibnja 2014. u A-konfiguraciji (324 sati) i C-konfiguraciji (60 sati). Za pokrivanje puna dva kvadratna stupnja COSMOS polja korišteno je ukupno 192 teleskopska usmjerenja (engl. *pointing*). Odziv VLA antene na 3 GHz pada za pola na polumjeru od $15'$. U svakoj opažačkoj sesiji izvor J1331+3030 je opažan oko 3 – 5 minuta za kalibraciju toka i valnog pojasa, 1024-0052 je opažan oko dvije minute svakih 30 minuta za kalibraciju faza, dok je J0713+4349 opažan 5 minuta za mjerenje propuštanja polarizacije. Uglavnom je korišteno 26 antena te su mjerenja u A-konfiguraciji rađena za dobrog vremena, dok su u C-konfiguraciji mjerenja djelomično bila pogođena ljetnim olujama.

Kalibracija podataka obavljena je u programskom paketu AIPSLite-u (Bourke et al., 2014) baziranog na AIPS-u (engl. *Astronomical Image Processing System*). Kalibracijski proces je prilagođen za potrebe VLA-COSMOS projekta iz pregleda neba Stripe 82 (Mooley et al., 2016) te uglavnom prati preceduru opisanu u AIPS kuharici. Drugi i treći spektralni prozor su u potpunosti uklonjeni zbog snažnog utjecaja radio interferencije. U preostalim kanalima odstupanje amplituda od srednje vrijednosti je oko 2 – 3%, uz iznimna tri spektralna prozora s najvišom frekvencijom gdje odstupanje prelazi 10%. Ukupno raspršenje je oko 5%, čime se osigurava dobra kalibracija toka zračenja.

Za konstrukciju radio mape iz interferometrijskih opažanja korišten je višeskalni, višefrekvencijski algoritam sinteze (MSMF, engl. *multiscale multifrequency synthesis*) implementiran u CASA-i (Rau & Cornwell, 2011). Ovaj pristup omogućava najbolju kombinaciju rezolucije, šuma i kvalitete slike. Rezultat redukcije podataka je mapa toka zračenja i mapa spektralnog indeksa (koja se dobije iz prilagodbe na radio spektar unutar pune širine valnog pojasa). U 44 od 192 teleskopska usmjerenja, na nebu je bio dovoljno snažan radio izvor koji je omogućio dodatnu kalibraciju njime. Ovaj proces je poboljšao kvalitetu slike i smanjio artefakte oko snažnih izvora. Prilikom redukcije, podaci su uzimani s takvom težinom da u konačnoj slici rezolucijski element bude cirkularan s punom širinom na pola maksimuma u iznosu od $0.75''$. Razina šuma u svakom od 192 usmjerenja je iznosila oko $4 – 5\,\mu\text{Jy beam}^{-1}$. Sve 192 radio mape su spojene u

mozaik na način da je svaka ušla u ukupnu sumu s težinskim faktorom proporcionalnim obrnutom kvadratu šuma. U konačnom MSMF mozaiku medijan (centralna vrijednost) šuma preko dva kvadratna stupnja COSMOS polja iznosi 2.3 μJy beam^{-1}.

Katalog radio izvora

Kako bi se izvukle komponente radio izvora iz VLA-COSMOS mozaika i katalogizirala njihova svojstva, korišten je program BLOBCAT (Hales et al., 2012). Proces se oslanja na algoritam „poplavljivanja" (engl. *flood fill*) koji popisuje nakupine piksela iznad neke S/N granice. Vrh se tada iščitava iz prilagodbe 2D parabole, dok ukupni tok zračenja dolazi iz sume piksela uz dodatne korekcije. Ovim procesom je u radio mapi pronađeno 10 899 komponenti značajnijih od 5σ. Kako bi se odredilo jesu li komponente razlučene (tj. veće od sintetskog rezolucijskog elementa, engl. *beam*), koristi se omjer ukupnog toka zračenja S_t i maksimalnog površinskog sjaja S_p. Taj omjer iznosi točno jedan u slučaju nerazlučenog izvora u odsustvu šuma. Omjer je veći od jedan kada izvor okupira više od jednog rezolucijskog elementa. Prilagodbom na raspršenje tog omjera u odnosu na pouzdanost detekcije (Bondi et al., 2008) dolazi se do izraza $S_t/S_p = 1 + 6 \times (\text{S/N})^{-1.44}$ koji razdvaja razlučene od nerazlučenih izvora u prisustvu šuma. Radio izvori difuzne strukture, poput radio galaksija ili razlučenih diskova u kojima se stvaraju zvijezde, mogu biti razdvojeni u više komponenti ako nisu spojene dovoljno jakim zračenjem. Vizualnom inspekcijom preko 2 500 komponenti pronađeno je 67 višekomponentnih izvora kojima je ručno izračunat tok zračenja i oni su posebno označeni u katalogu. Usporedbom koordinata izvora s kompaktnim AGN-ovima opaženih pomoću VLBA interferometra (Herrera Ruiz et al., 2017), nalazi se astrometrijska preciznost od 0.01″ kod snažnih izvora. Konačni katalog i radio mapa dostupni su na internetu preko COSMOS IRSA arhive.

Uptrebom MSMF algoritma, osim toka zračenja na središnjoj frekvenciji, dobiva se i spektralni radio indeks za svaki izvor. Uvidom u snažne izvore unutar preklapajućih teleskopskih usmjerenja dolazi se do zaključka da postoji sistematska greška u spektralnom indeksu koja raste linearno s udaljenošću od faznog centra. Do ovog efekta dolazi zbog nesavršene korekcije na odziv antene koji raste linearno s valnom duljinom. Drugi način za računanje spektralnog indeksa svodi se na omjer toka zračenja na dvije različite radio frekvencije, a COSMOS polje je već opažano na 1.4 GHz (Schinnerer et al., 2010). Međutim, analiza nije moguća za sve izvore zbog razlike u osjetljivosti između ova dva pregleda COSMOS polja. Srednji spektralni indeks izvora opaženih na obje frekvencije iznosi α = −0.84, uz standardnu devijaciju uzorka od σ = 0.35. Usporedba ovako izračunatih spektralnih indeksa i onih iz MSMF metode daje veliko raspršenje čak i kod vrlo svijetlih izvora. Detaljnija analiza korištenjem „metode opstanka" (engl. *survival analysis*), koja uzima u obzir granice od izvora koji nisu opaženi na 1.4 GHz

(ukupno 75%), daje rezultat $\alpha = -0.68$ koji je u skladu s postojećom literaturom (Kimball & Ivezić, 2008).

Broj izvora u ovisnosti o toku zračenja i statističke korekcije

Prije nego što se može napraviti kvalitetna statistička analiza nad dobivenim katalogom, potrebno je procijeniti koliki udio stvarnih izvora nedostaje i koliki udio unosa u katalogu vjerojatno nije stvarni izvor. Uzroci ovih poteškoća su kombinacija efekata poput nehomogenog šuma preko cijele mape, tendencija da vrhovi šuma budu označeni kao izvor, ili činjenica da interferometar nije jednako osjetljiv na sve kutne veličine. Sistematika koja proizlazi iz ovih efekata može se procijeniti Monte Carlo simulacijama. Umjetni izvori se ubace u mapu iz koje se potom izvuku pomoću BLOBCAT-a na isti način kao i sa stvarnim podacima. Kako je radio mapa na rezoluciji od $0.75''$ s pet tisuća izvora po kvadratnom stupnju uglavnom prazna, nije teško postići da se umjetni izvori ubacuju u postojeću mapu pritom izbjegavajući već popisane radio komponente. Usporedbom kataloga početnih i izvučenih simuliranih izvora dolazi se do korekcije na kompletnost.

Prilikom izrade kataloga umjetnih izvora potrebno je znati koliko ih ima u nekom intervalu gustoće zračenja te koja im je kutna dimenzija. Prva informacija je preuzeta iz Vernstrom et al. (2014), dok druga još uvijek nije dobro određena. Za kutnu veličinu izvora θ se pretpostavlja da raste s gustoćom toka S na način $\theta \propto S^n$, gdje se n dobiva iz usporedbe sa stvarnim podacima. Stvarni izvori mogu imati kompleksnu morfologiju koja se ovdje iz jednostavnosti simulira pomoću 2D Gaussiana u dva ekstrema: i) izvor je eliptičan - jedna os je izdužena, ii) izvor je kružan - obje osi su jednako izdužene. Još je bilo potrebno uvesti rez na najmanju moguću kutnu veličinu da se izbjegne stvaranje previše kompaktnih izvora. Rezultat ove analize je krivulja kompletnosti u ovisnosti o gustoći toka koja pokazuje da je katalog 60% potpun ispod 20 μJy, te oko 95% potpun iznad 40 μJy. Za procjenu koliko je izvora u radio katalogu samo slučajni šum, izvori su izvučeni iz negativne radio mape (pomnožene s -1). Rezultat je da 24% izvora u katalogu sa S/N između 5.0 i 5.1 vjerojatno nije stvarno. Taj postotak je manji od 3% iznad S/N>5.5, te iznosi ukupno 2.7% za cijeli 5σ katalog.

Ove dvije korekcije, od kojih prva statistički povećava broj izvora, a druga smanjuje, množe brojeve izvora u intervalima gustoće toka. Kada se broj izvora normira na Euklidsku geometriju (što se radi iz povijesnih razloga), dobiva se da diferencijalni broj izvora u intervalima gustoće toka ($S^{2.5}dN/dS$) opada od 10 mJy do 0.3 mJy, nakon čega se izravnava te nastavlja opadati ispod 30 μJy. Ovaj rezultat se slaže s Condon et al. (2012) rezultatom gdje je korištena tzv. $P(D)$ analiza koja procjenjuje raspodjelu izvora mjerenjem njihovog pozitivnog doprinosa ukupnom šumu u mapi koja je ograničena konfuzijom (rezolucija mora biti dovoljno mala da se izvori

često spajaju u jedan). Rezultati su također uspoređeni s brojnom literaturom gdje razlike mogu nastati zbog kozmičke varijance (opažanje isječaka svemira različitih gustoća), ili drukčijeg pristupa procjeni kompletnosti. Usporedbe s poluempirijskim modelima (npr. Wilman et al., 2008) omogućavaju njihovo bolje ugađanje.

Funkcije luminoziteta galaksija

Da bi se mogla proučavati evolucija galaksija kroz kozmičko vrijeme, svakom radio izvoru potrebno je odrediti crveni pomak iz kojeg se unutar kozmološkog modela određuje kojoj vremenskoj epohi galaksija pripada.

Povezivanje radio i optičkih kataloga

Katalog 10 830 radio izvora na 3 GHz je uparen s već postojećim COSMOS2015 katalogom sačinjenim od opažanja u preko 30 valnih pojasa u optičkom, IR i UV području (Laigle et al., 2016). U njemu postoji otprilike 800 000 izvora na 1.77 deg^2 s pouzdano izračunatim fotometrijskim crvenim pomacima dobivenima iz prilagodbi SED (Ilbert et al., 2013). S obzirom na to da optički i radio podaci imaju rezoluciju ispod lučne sekunde, uparivanje radio i optičkih izvora urađeno je metodom traženja najbližeg susjeda unutar polumjera od 0.8″. Kako su galaksije koje zrače u radio području često među najmasivnijim i najsvjetlijim galaksijama opaženim u vidljivom svjetlu, provedene su dodatne simulacije u kojima se utvrđuje vjerojatnost krivog uparivanja (vidi Smolčić et al., 2017b). Kada se izuzmu dijelovi neba kontaminirani zvijezdama iz Mliječog puta, preostaje 7 729 radio galaksija koje su uparene s galaksijom iz COSMOS2015 kataloga. Korištenjem starijeg kataloga izvučenog iz opažanja u *i*-pojasu (IR područje, Capak et al. 2007), ukupni broj uparenih izvora je 7 826. Crveni pomaci su određeni spektroskopskim mjerenjima za 35% galaksija. Ovo predstavlja uzorak galaksije koji se dalje analizira. Radio spektralni indeksi su određeni za optrilike četvrtinu svih galaksija u tom uzorku korištenjem 1.4 GHz kataloga (Schinnerer et al., 2010).

Klasifikacija galaksija

Da bi se kvantificiralo koliki udio radio zračenja galaksije dolazi od AGN-a, a koliki od procesa stvaranja zvijezda, uspoređuje se radio i IR zračenje na sljedeći način (preuzeto iz Delvecchio et al., 2017). Radi se prilagodba na SED koja uključuje dodatni predložak za AGN. Doprinos AGN-a u IR valnom pojasu se oduzima od ukupne SED te preostali dio bolometrijskog IR svjetla dolazi od prašine koju griju uglavnom mlade vruće zvijezde, što se može povezati s brzinom

stvaranja zvijezdi (SFR, engl. *star formation rate*) preko Kennicutt (1998) kalibracije. Ukoliko radio zračenje također dolazi od procesa stvaranja zvijezda, tada će ono biti korelirano sa SFR. Ako je u galaksiji prisutan AGN, postojat će višak radio zračenja. Višak koji je značajniji od 3σ ukazuje na to da je radio zračenje galaksije dominirano AGN-om, što je zadovoljeno kad je $\log\left(L_{1.4\,\mathrm{GHz}}[\mathrm{W\,Hz}^{-1}]/\mathrm{SFR_{IR}}[\mathrm{M_\odot\,god^{-1}}]\right) > 22 \times (1+z)^{0.013}$, gdje IR indeks označava valni pojas korišten za računanje SFR veličine i masa Sunca iznosi $\mathrm{M_\odot} \approx 2 \times 10^{30}$ kg. Ovakav kriterij odabire 1 814 (23%) galaksija s viškom radio zračenja koje se klasificiraju kao AGN. Za preostale galaksije se smatra da stvaraju zvijezde. S obzirom na to da su galaksije uglavnom hibridi gdje radio zračenje dolazi od oba procesa (zvijezde i AGN), ne postoji oštra granica između dva tipa galaksija. U ovom slučaju galaksije koje stvaraju zvijezde mogu biti kontaminirane AGN zračenjem, ali ono nije dominantan doprinos.

Računanje funkcije luminoziteta i njene evolucije

Funkcija luminoziteta (Φ, LF, engl. *luminosity function*) opisuje gustoću galaksija određenog luminoziteta neke kozmičke epohe. Radio luminozitet se određuje iz izmjerene gustoće toka i izračunatog crvenog pomaka, uz pretpostavku da se radio SED može opisati jednostavnim spektralnim indeksom. Prvo je potrebno podijeliti galaksije u intervale koji predstavljaju kozmičku epohu, kao i intervale luminoziteta. Prilikom računanja gustoće u epohi, za svaku galaksiju potrebno je izračunati maksimalni volumen V_{max} unutar kojeg galaksija određenog luminoziteta L može biti opažena (vidi Schmidt, 1968). To se svodi na sumu tipa $\Phi(L,z) = 1/\Delta\log L \cdot \sum_{i=1}^{N} 1/V_{\mathrm{max},i}$. U račun volumena ulaze faktori koji uzimaju u obzir kozmologiju, površinu opaženog dijela neba i korekcije na kompletnost radio opažanja. Dobivene točke u intervalima luminoziteta se prilagođavaju na analitičku funkciju koja ovisi o populaciji galaksije koja se proučava, a uglavnom se svodi na kombinaciju potencije, eksponencijalne i log-normalne krivulje. Evolucija kroz kozmičko vrijeme se može opisati preko lokalne LF $\Phi_0 = \Phi(z=0)$ koja se translatira po osi gustoće ili luminoziteta ($\log\Phi$, $\log L$). To se može zapisati na način $\Phi(L,z,\alpha_L,\beta_L,\alpha_D,\beta_D) = (1+z)^{\alpha_D+z\cdot\beta_D} \times \Phi_0\left[L/(1+z)^{\alpha_L+z\cdot\beta_L}\right]$, gdje parametri α i β (indeks D označava gustoću, a L luminozitet) opisuju evoluciju LF kroz vrijeme (z).

Populacija galaksija koje stvaraju zvijezde

Korištenjem nekoliko postojećih pregleda neba u kojima su galaksije opažene u preko pet dekada luminoziteta izračunata je lokalna LF za galaksije koje stvaraju zvijezde. Ona se može prilagoditi sa $\Phi_0^{\mathrm{SF}}(L) = \Phi_\star (L/L_\star)^{1-\alpha} \exp\left[-1/(2\sigma^2)\log^2(1+L/L_\star)\right]$, gdje su $\Phi_\star = 3.55 \times 10^{-3}$ $\mathrm{Mpc^{-3}dex^{-1}}$, $L_\star = 1.85 \times 10^{21}$ $\mathrm{W\,Hz^{-1}}$, $\alpha = 1.22$, $\sigma = 0.63$. Korištenjem 5 915 galaksija

iz VLA 3 GHz uzorka izračunate su radio LF za ovaj tip galaksija. Korištenjem navedene lokalne LF, prilagodbom se dobivaju evolucijski parametri $\alpha_L = 3.16 \pm 0.04$ i $\beta_L = -0.32 \pm 0.01$. Parametri za evoluciju gustoće su $\alpha_D = 0$ i $\beta_D = 0$ jer bi inače gustoće galaksija bile tako velike, kako nisu opažene u svemiru. Ova opažanja najbolje mjere svijetli kraj LF-a, iza pregiba definiranog s $L_\star$, te vrijede do $z \sim 5.5$.

Radi provjere robusnosti radio LF-a i stvaranja konzistentne slike preko šireg elektromagnetskog spektra, radio LF su uspoređene s drugim publiciranim mjerenjima. Poklapanje je dobro s rezultatima iz Smolčić et al. (2009a) koji su dobiveni na prošlom VLA-COSMOS pregledu neba. Slaganje je također dobro s IR podacima (Gruppioni et al., 2013) do $z \sim 2$ nakon čega su radio točke sistemski niže, vjerojatno zbog AGN doprinosa u IR valnom području. Usporedba s UV podacima ukazuje na to da trenutne korekcije na prašinu koje se koriste u UV području vjerojatno nisu dovoljne za nadoknaditi broj neopaženih galaksija s velikim udjelom prašine koje se lako opažaju u radio području na $z \gtrsim 4$. Također, postoji indikacija da je evolucijski model radio LF-a prejednostavan te predviđa premalo galaksija slabijeg sjaja (ispod $L_\star$).

AGN populacija

Lokalna LF za AGN-ove je preuzeta iz Mauch & Sadler (2007) te se opisuje s $\Phi_0^{\mathrm{AGN}}(L) = \Phi^* / \left[(L^*/L)^\alpha + (L^*/L)^\beta \right]$, gdje su $\Phi^* = \frac{1}{0.4} 10^{-5.5}$ Mpc^{-3} dex^{-1}, $L^* = 10^{24.59}$ W Hz^{-1}, $\alpha = -1.27$ i $\beta = -0.49$. Uzorak AGN-ova iz VLA 3 GHz kataloga sadržava 1 814 galaksija s viškom radio zračenja u odnosu na IR zračenje koje dolazi od prašine zagrijane zvijezdama. Izračunata je LF ovog uzorka, a prilagodba na evoluciju daje parametre $\alpha_L = 2.88 \pm 0.17$, $\beta_L = -0.83 \pm 0.07$ za čistu evoluciju luminoziteta ($\alpha_D = \beta_D = 0$), koja vrijedi do $z \sim 5.5$; te $\alpha_D = 2.00 \pm 0.01$, $\beta_D = -0.60 \pm 0.02$ za čistu evoluciju gustoće ($\alpha_L = \beta_L = 0$). Slaganje ovih radio LF-a s Donoso et al. (2009) rezultatima u epohi $0.4 < z < 0.7$ je odlična. Usporedba sa rezultatima iz Smolčić et al. (2009b) daje dobro slaganje s prošlim VLA-COSMOS mjerenjima, dok usporedba s Padovani et al. (2015) daje nešto veće raspršenje zbog manjeg broja galaksija opažanih u tom radu.

Moguća sistematika

Najveći izvor greške prilikom izračunavanja luminoziteta neke galaksije leži u nesavršenom poznavanju njenog spektralnog indeksa, ili općenitije, njene radio SED. Srednji spektralni indeks iznosi $\alpha = -0.7$, međutim postoji raspršenje od otprilike 0.4 (Kimball & Ivezić, 2008). Iako greška od $\Delta\alpha = 0.1$ uzrokuje promjenu luminoziteta od 0.1 dex za galaksiju na $z \sim 5$, s obzirom da je raspršenje α simetrično, ti efekti se uglavnom ponište u velikom uzorku. Izračuni

LF-a su ponovljeni korištenjem identičnog spektralnog indeksa za sve galaksije u kojem slučaju analitička prilagodba ostaje praktički nepromijenjena, ali ukupni χ^2 je manji (točke su manje raspršene oko prilagodbe).

Sistematski selekcijski efekt vezan za evoluciju AGN-ova dolazi od opažačke činjenice da postoje dvije vrste AGN-ova koji evoluiraju kroz kozmičko vrijeme na različit način (Hardcastle et al., 2007). Za ovu analizu treba koristiti dodatne AGN indikatore poput X-zraka. Klasifikacija koja se temelji na višku radio zračenja u odnosu na IR pretvoren u SFR, ne razaznaje ove dvije potkategorije AGN-ova.

Ukupna radio fukncija luminoziteta

Izračunata je i ukupna radio LF u koju ulaze svi opaženi radio izvori koji su upareni s galaksijama opaženim u optičkom i IR području. Ovaj korak je koristan za provjeru prije izračunatih LF-a jer se ovdje ne klasificiraju galaksije u prvom koraku. Detalji pojedine populacije galaksija ulaze u račun preko oblika LF-a koji se koristi za analitičku prilagodbu. Ukupna radio LF je stoga sagrađena od sume dvije LF, od kojih jedna predstavlja galaksije koje stvaraju zvijezde, dok druga opisuje AGN-ove. Izračunate točke LF-a se prilagođavaju uz 4 slobodna parametra $\Phi(L, z, \alpha_L^{\text{SF}}, \beta_L^{\text{SF}}, \alpha_L^{\text{AGN}}, \beta_L^{\text{AGN}}) = \Phi_0^{\text{SF}} \left[L/(1+z)^{\alpha_L^{\text{SF}} + z \cdot \beta_L^{\text{SF}}} \right] + \Phi_0^{\text{AGN}} \left[L/(1+z)^{\alpha_L^{\text{AGN}} + z \cdot \beta_L^{\text{AGN}}} \right]$. Za ovaj proces prilagodbe koristi se algoritam Monte Carlo Markovljevih lanaca (MCMC, engl. *Markov chain Monte Carlo*). Dobiveni rezultati su u skladu (unutar 3σ) s LF-ima izračunatim na pojedinim populacijama galaksija. Korištenjem različitih kombinacija LF-a može se dobiti bolji uvid u efekte ekstrapolacije LF izvan opaženog područja luminoziteta.

Brojčana gustoća radio galaksija u nano-Jy režimu

Jedan od načina proučavanja kako različite populacije galaksija doprinose ukupnom radio nebu je putem jednostavnog prebrojavanja galaksija određene gustoće toka sadržanih u nekom prostornom kutu neba (Padovani, 2016). Ova tehnika se u početku koristila za pokušaj analize geometrije svemira, ali je s vremenom prerasla u metodu praćenja evolucije galaksija. Ako se opazi više galaksija nego što trenutno prihvaćeni model predviđa, postoji mogućnost da je otkrivena nova populacija radio galaksija (Rowan-Robinson et al., 1993). Broj galaksija kao funkcija gustoće toka je prirodna posljedica oblika i evolucije ishodišnih LF-a.

Iz prije dobivene ukupne radio LF izračunati su brojevi izvora ekstrapolirani ispod direktno opaženih luminoziteta. Rezultati su uspoređeni s poluanalitičnim modelima (Wilman et al., 2008), kao i s radio opažanjima ograničenih konfuzijom izvora (Vernstrom et al., 2014) iz kojih

se može modelirati ukupna gustoća toka pozadinske raspodjele izvora (koja može biti i sakrivena u šumu). Dobro slaganje ovih rezultata potvrđuje robusnost LF-a i njihove ekstrapolacije. Omjeri brojeva dviju populacija galaksija ukazuju na to da su galaksije s opaženom gustoćom toka između 100 nJy i 10 μJy na 1.4 GHz u $90 - 95\%$ slučajeva galaksije koje stvaraju zvijezde, odnosno da je kontaminacija AGN-ovima manja od 10%. Modeli predviđaju da će galaksije opažene u tom intervalu gustoće toka uglavnom pripadati epohi na $z \sim 2$.

Kozmička povijest stvaranja zvijezda

Radio zračenje se može iskoristiti za računanje brzine stvaranja zvijezda, pružajući time mjerenja bez utjecaja prašine. Ovaj proces se oslanja na empirijsku činjenicu da je zračenje u termičkom IR i netermičkom radio području korelirano, a IR spektar se može povezati sa udjelom mladih zvijezda (Bell, 2003). Ova IR-radio korelacija se označava s q_{TIR}, (TIR označava ukupno IR svjetlo $8 - 1\,000\,\mu$m, engl. *total infrared*) koji je definiran preko omjera luminoziteta kao $q_{\text{TIR}} = \log\left(L_{\text{TIR}}/[3.75 \times 10^{12}\,\text{W}]\right) - \log\left(L_{1.4\,\text{GHz}}/[\text{W}\,\text{Hz}^{-1}]\right)$. Bilo je uobičajeno uzimati da je q_{TIR} konstantan, međutim nova istraživanja su pokazala da se mijenja kao funkcija kozmičkog vremena (Delhaize et al., 2017). Na uzorku galaksija koje stvaraju zvijezde, korištenjem metode opstanka koja uzima u obzir ograničenja dobivena iz galaksija detektiranih u radio i IR području ili u samo jednom od ta dva, dobivena je relacija $q_{\text{TIR}}(z) = (2.78 \pm 0.02) \times (1+z)^{-0.14 \pm 0.01}$. Konačni izraz kojim se dobiva brzina stvaranja zvijezda galaksije (u jedinicama mase zvijezdi stvorenih u godini) iz radio luminoziteta je $\text{SFR}[\text{M}_\odot\,\text{god}^{-1}] = f_{\text{IMF}} \times 10^{-24} 10^{q_{\text{TIR}}(z)} L_{1.4\,\text{GHz}}[\text{W}\,\text{Hz}^{-1}]$, gdje f_{IMF} označava normalizaciju inicijalne funkcije mase (IMF, engl. *initial mass function*) te iznosi 1 u slučaju korištenja Chabrier (2003) IMF-a. Kozmička povijest stvaranja zvijezda može se izračunati integriranjem ukupnog radio zračenja od svih galaksija iste populacije (zapisano u LF-u) neke epohe na način $\int \Phi(L,z) \times \text{SFR}(L)\,d\log L$, što je urađeno do epohe $z \sim 5.5$ kad je svemir bio star tek oko jednu milijardu godina.

Rezultati reproduciraju eksponencijalni pad stvorenih zvijezda od $z \sim 1$ te su u skladu s velikom kompilacijom objavljenih mjerenja (Madau & Dickinson, 2014) do $z \sim 2$, dok se u kasnijim epohama pronalazi nešto veće vrijednosti. Najefikasnija epoha stvaranja zvijezda bila je između $2 < z < 3$ (starost svemira oko dvije do tri milijarde godina), kada galaksije pretvaraju u prosjeku 0.1 $\text{M}_\odot$ plina i prašine po kubnom megaparseku u nove zvijezde svake godine. U istoj epohi, oko četvrtina svih zvijezda se stvara u galaksijama koje mogu svake godine proizvesti preko 100 $\text{M}_\odot$ zvijezda godišnje. radio mjerenjima dobiva $15 - 20\%$ više stvorenih zvijezdi u epohi $z \sim 4 - 5$. Radio mjerenja ukazuju na $15 - 20\%$ više stvorenih zvijezdi u

epohi $z \sim 4 - 5$ nego što se dobiva iz UV podataka koji direktno mjere zračenje mladih vrućih zvijezda, ali su pritom zatamnjeni prašinom. Ovi rezultati predstavljaju najdublja i najpreciznija radio mjerenja napravljena u svrhu dobivanja kozmičke povijesti stvaranja zvijezda te su time i glavni znanstveni doprinos ovog rada.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Galaxies are among the most magnificent objects in the universe and represent the largest accumulations of visible matter. They are composed of stars, gas and dust circling around a supermassive black hole (SMBH) under the influence of the gravitational force (Sparke & Gallagher, 2007). Although each galaxy is different, they share many properties which astronomers try to classify, correlate, understand and ultimately predict. By looking at the sky for an instant one can observe a long history of the universe due to the finite speed of the light and the photons' long voyage from a distant object to the observer. This effect grants us the vision of events long gone, yet still encoded in the light which traveled billions of years to reach us. Galactic processes such as galaxy merging, star formation, and chemical evolution are measured in no less than millions of years and cannot be observed in real continuous time (see Mo et al., 2010). Also, astronomers interested in a specific galaxy are able to observe it from only one angle. The goal of galaxy evolution studies is tying snapshots of the universe at different time epochs in a meaningful way. The final result is a deeper understanding of the universe and its history. This thesis focuses on galaxy evolution and star formation since the universe was only 1 billion years old by using radio wavelength observations.

1.1   Lambda cold dark matter universe

To put galaxies in their proper time and space, a framework of the universe is required. The most successful and currently accepted cosmological model of the universe is the Lambda cold dark matter (ΛCDM) universe which assumes that gravity is correctly described by the general relativity (Mo et al., 2010). In this model the universe started with the Big Bang, a state of extremely high energy density, temperature and pressure. This was followed by the inflation period (see also Bonometto et al., 2002) where an exponential expansion resulted in a rapid
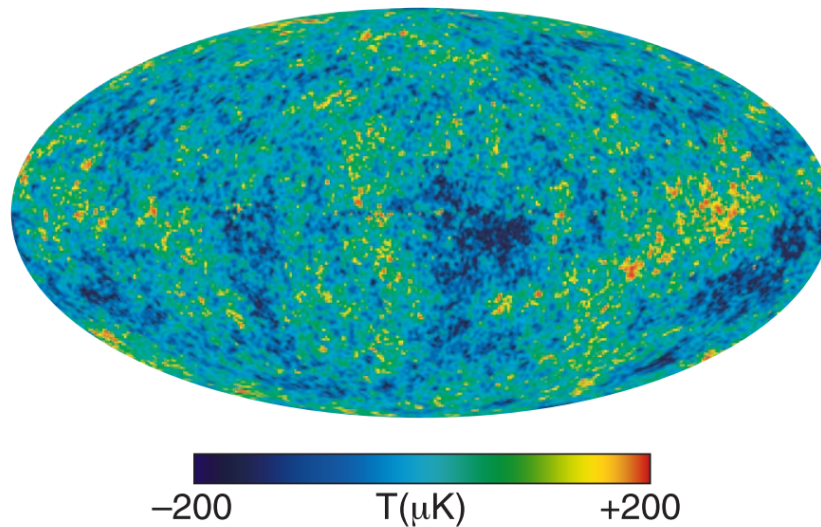
Figure 1.1: The cosmic microwave background from the nine-year Wilkinson microwave anisotropy probe (WMAP) observations (reproduced from Bennett et al., 2013). Colors represent temperature fluctuations.

temperature and density decrease allowing the first elementary particles to form from the quark-gluon plasma. The solution to general relativity equations which yielded an expanding universe needed a non zero vacuum energy, denoted by Λ. It represents the dark energy in empty space acting against the pull of gravity. During the inflation period any small scale (quantum) thermal fluctuations would become the seeds for the structure in the universe seen today. As the cooling progressed, protons and electrons recombined into neutral hydrogen atoms. The photons were no longer coupled to matter particles in thermal equilibrium and the universe became transparent to photons. The photons from this epoch are still observable today, Doppler redshifted to microwave wavelengths due to the universe expansion. This emission is called the cosmic microwave background (CMB, see Penzias & Wilson 1965), also shown in Fig. 1.1, and it is one of the cornerstones of the cosmological model. It demonstrates that the universe is homogeneous and isotropic on large scales, which requires a causal connection in its beginning. However, it also shows fluctuations in the temperature of the order of tens of microkelvins which results in density fluctuations and the final galaxy distributions of today. The Big Bang paradigm also explains the abundances of primordial nucleosynthesis isotopes: hydrogen, helium and lithium (e.g. Amsler et al., 2008).

After the recombination there were no new sources of light and the universe entered the so called "Dark Ages". Some time during this dark epoch the first stars must have been formed, as well as the first galaxies. With the onset of fusion in the young stars, new energetic photons

were being produced which participated in the reionization of the hydrogen in the universe. The details of this process are still unknown and it happened sometime when the age of the universe was $0.3 - 1$ billion years. The universe became plasma again, but remained transparent due to lower densities achieved by the expansion. From this point onward galaxies can be observed at all time epochs and the study of galaxy evolution can commence. Probing all the way to the epoch of reionization is the goal of many future sky surveys. Since the universe is expanding at an accelerating rate, meaning that more distant objects move faster away from us, the measure of Doppler shift, or redshift, represents both the age of the universe and the distance to an object within a cosmological model.

The final ingredient of the ΛCDM model is the so called dark matter. It is used to explain flat rotation curves of spiral galaxies (e.g. Rubin & Ford, 1970), galaxy clustering (e.g Springel et al., 2005), and gravitational lensing (e.g. Massey et al., 2010). In all cases the visible matter was not enough to account for the strength of the gravitational potential well. Therefore, cold dark matter was introduced as a non baryonic constituent which interacts only through gravity, has small velocities (compared to the speed of light) and is collisionless.

The success of the cosmological model is measured in the amounts of observables that it can reproduce while trying to keep the parameter space to a minimum. The ΛCDM six-parameter model (see Hinshaw et al., 2013) states that the universe is 13.8 billion years old, flat (slight spatial curvature allowed within the errors), and composed of 71.8% dark energy, 23.6% cold dark matter, and 4.6% baryonic (visible) matter.

1.2   Galaxies and their evolution

Establishing the model of the universe provides a consistent way of measuring distances and sizes of galaxies across all cosmic times. The redshift of a galaxy provides the age of the universe, allowing to properly populate the time axis. In order to build a picture of galaxy evolution, which is a goal of this thesis, common properties must first be identified and tied together: from the shapes of the galaxies to their multiwavelength features (see also Mo et al., 2010). Then, under the assumption of available evolutionary paths, different time slices can be connected together. The theory of galaxy evolution should be able to reproduce the universe seen today while respecting the constraints observed from the younger universe (e.g. Kereš et al., 2009).
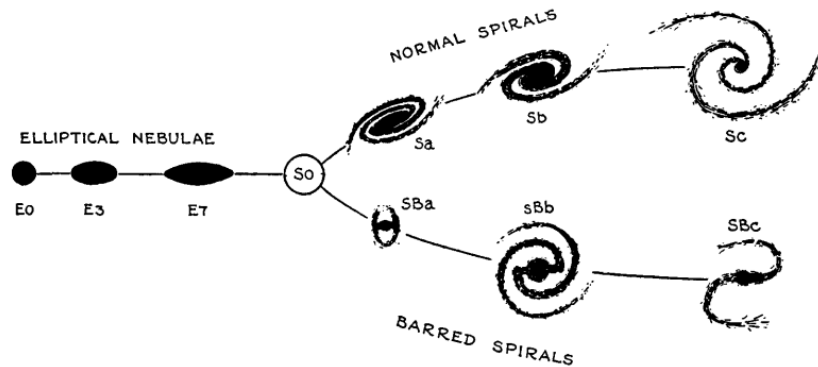
Figure 1.2: Hubble's galaxy classification based on visual morphology (reproduced from Hubble, 1936).

1.2.1 Morphology classes

One of the first attempts at galaxy classification was made by Edwin Hubble using visual inspection of galaxy morphology. He employed criteria such as the galaxy overall shape (elliptical or disk/spiral), presence of a bar, relative luminosities of the central bulge and spiral arms, tightness of the spiral arms winding. His classification scheme is shown in Fig. 1.2. The categories were rather broad which enabled the system to be easily expanded, e.g. to include irregular galaxies or the presence of rings (see Sandage, 1961; de Vaucouleurs et al., 1991), thus helping it survive to the present day despite its simplicity. The shape of the diagram, if read from left to right, resembles a time evolutionary sequence, giving birth to names such as "early types" for ellipticals and "late types" for spirals. Although the nomenclature will remain, the actual chronology is reversed as the ellipticals are the end product of galaxy mergers and growth (e.g. Faber et al., 2007). Despite an enormous variety of galaxies and difficulties in observing them, Hubble demonstrated that galaxies can be sorted into relatively few bins anticipating that these classes may reflect some dominant physical process. However, the danger in making conclusions from insufficient data is always present. Morphology is strongly dependent on the observing wavelength. For example, dust attenuates the optical light, but it is bright in the infrared (IR). Therefore, the same galaxy can appear significantly different depending on the observing wavelength. Subjectiveness of addressing morphology is another issue to bear in mind. Regardless, knowing the morphology of a galaxy remains a useful asset in its analysis, especially when its dynamics are concerned, i.e. spirals rotate coherently around one axis, while ellipticals are tri-axial ellipsoids.

The sheer number of galaxies and the requirement to visually inspect them presents a large time investment in order to produce a significant sample of galaxies. As telescopes become more advanced, survey speed is increased and the amount of data rapidly grows. When the

number of objects pending inspection grows to tens and hundreds of thousands, no single professional astronomer can go through them alone. By covering around a quarter of the whole sky, a sample of nearly a million objects was obtained from the Sloan Digital Sky Survey (SDSS) at optical and near-infrared (NIR) frequencies. An excellent example on how to approach the classification here is the crowd-sourcing performed for Galaxy Zoo[1] (Lintott et al., 2008). More than 100 000 volunteers were asked to classify sources using the internet interface into six distinct visual categories: elliptical galaxy, clockwise/Z-wise spiral galaxy, anticlockwise/S-wise spiral galaxy, other spiral galaxy (e.g. edge-on), star or do not know (e.g. artifact) and merger. The final product was a catalog with nearly 900 000 classifications, reliable due to the repeated and independent inputs for the same object (see Lintott et al., 2011). In the following years the Galaxy Zoo was expanded to include more morphological details (Hart et al., 2016) and also images from the *Hubble Space Telescope* (Willett et al., 2017), providing an ever growing database of galaxy morphologies and making possible the publication of more than fifty science papers so far. Another feasible approach to sorting galaxies is to use machine learning algorithms enabling computers to perform the classifications, however they still need to be tuned very carefully and require a reliable human selected training set (e.g. Banerji et al., 2010).

1.2.2   Emission across the electromagnetic spectrum

Galaxies in the early universe (i.e. at high redshift) appear smaller because of the large distance and so morphology classification is rendered impossible by insufficient telescope angular resolution. Most galaxies considered in this thesis fall into this category. In this case, the only remaining option is to observe the light integrated across the entire unresolved galaxy. Thankfully, all galaxies are real physical systems and they all obey the same physical rules. This means that many properties are correlated, thus making various classification approaches complementary to each other.

Color bimodality

Volunteers classifying morphologies in the Galaxy Zoo (explained in the previous section) would quickly notice that there is a tendency of disk/spiral late-type galaxies to appear bluer, while spheroidal early-type galaxies would appear redder. The color actually changes linearly from red to blue along the Hubble morphological sequence (see also Roberts & Haynes, 1994) and in the absence of resolved galaxy details, its color can be used as a proxy for its shape. Two main ingredients affect galaxy optical colors: stars and dust. Young massive stars will dominate

---

[1]https://www.galaxyzoo.org/

the emission in the ultraviolet (UV), hence provide blue light, while an old stellar population and less massive stars will provide red light. The dust attenuates more energetic light which will result in galaxy reddening. The bimodality of galaxies is apparent in the color versus absolute magnitude distribution (e.g. Strateva et al., 2001), where rest-frame $u - r$ color distribution could be explained with a sum of two Gaussian functions. This result strengthened the practice of classifying galaxies into the so called blue cloud and red sequence (e.g. Faber et al., 2007). Other properties of galaxies in these two categories are also different. For example, O and B stars are short lived and bright in the UV, and blue cloud galaxies must continuously form new stars in order to remain blue in color. On the other hand, galaxies in the red sequence are either very dusty, or have not formed new stars recently. They are also more strongly clustered (e.g. Budavári et al., 2003) providing clues on how environment affects the galaxy.

Spectral energy distribution

Astrophysicists are usually interested in the details of the underlying physical mechanisms, the so called true properties of the galaxy. Photons from radio to X-ray frequencies have energies spanning ten orders of magnitude, and if gamma-rays are also included, the range is extended even further. Different physical scenarios give rise to emission of varying energies, and a link between observed properties and theoretical physics must be established.

Telescopes can either try to capture a high resolution spectrum of a galaxy, or the average light emission integrated within some fixed wavelength bandwidth, yielding a single flux measurement point per photometric band. Both approaches have their strengths and weaknesses. Photometric surveys are cheaper and can quickly cover large portions of the sky, but lose details about narrow line features in the spectrum. Obtaining the full spectrum with good signal-to-noise ratio can be time consuming, but much more informative, especially since redshift can be best inferred from positions of known atomic and molecular lines. The usual practice is to perform a blind photometric imaging survey followed by spectroscopic observations of potentially interesting objects at coordinates inferred from the photometric imaging. Earth's atmosphere transparency provides limitations to which frequency windows can be observed from the ground, and where space telescopes are necessary. As the number of photometry points increased, correlations between different bands or colors became more apparent and it was easier to reach a consensus about galaxy properties. The goal of observations is to constrain the galaxy spectral energy distribution (SED) as best as possible. The understanding of physics encoded in the SED is obtained from the comparison with theoretical synthetic galaxy spectra.

If we disregard for the moment matter accretion onto a SMBH, the UV to IR emission of a galaxy is composed entirely from stellar light, either direct, or reprocessed by gas and
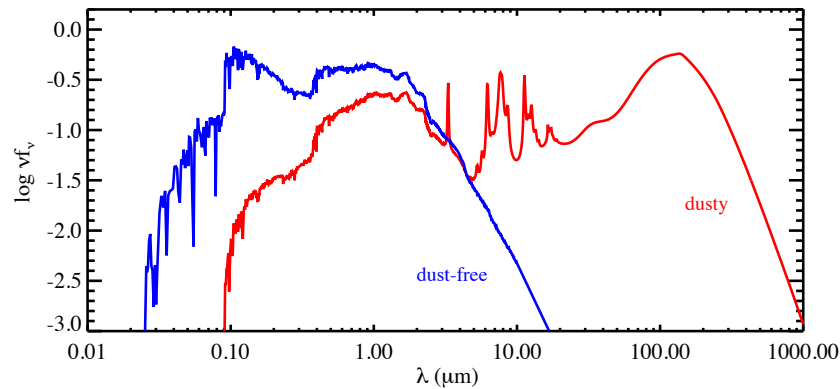
Figure 1.3: Galaxy spectral energy distribution example from near UV to mm wavelengths (reproduced from Conroy, 2013). The blue line represents a dust-free galaxy, while the the red one includes emission and attenuation by dust.

dust. To obtain the integrated spectra of all stars in the galaxy theoretically, a stellar population synthesis is applied (first attempted by Tinsley, 1972). Although the idea is straight forward, numerous assumptions must be made in the process (for a review see Conroy, 2013). First, a simple stellar population is generated using three ingredients: i) the initial mass function (IMF) which states the distribution of stars given their mass and provides the light-to-mass normalization, ii) an isochrone which selects a locus of stars with common age and metallicity in the Hertzsprung–Russell diagram, and iii) spectra for each stellar type. In the second step, the time dependent star formation rate and chemical evolution are implemented, resulting in a more complex stellar population with stars of different ages and metallicities. Then, effects of light reprocessing by dust are considered as the dust heats up by absorbing energetic UV photons and then re-emits the energy in the IR. Finally, all stellar populations and dust contributions can be integrated yielding a galaxy SED, an example is shown in Fig. 1.3. It is apparent from this figure that the SED shape can be explained with a combination of smooth continuum with the addition of narrow features and breaks. The featureless continuum is due to black-body radiation of stars and dust at different temperatures. Near UV and optical features are formed by emitting and absorbing stellar ions, atoms and molecules, while emission around 10 μm is dominated by heated dust, namely graphite (also polycyclic aromatic hydrocarbons) and silicate grains (e.g. Desert et al., 1990).

If the galaxy hosts a SMBH in its center, the surrounding matter will fall onto it due to gravity. As the matter is being accreted it undergoes a large change of gravitational potential energy which must be radiated away. This radiation exhibits pressure which slows down the accretion process (hydrostatic balance between the gravitation pull and the radiative pressure is achieved at the so called Eddington luminosity), but also heats up the dust. This dust can then

significantly contribute to the total galaxy SED with a broad and mostly featureless emission spectrum peaking in the IR (e.g Fritz et al., 2006; Ciesla et al., 2015). This addition usually smooths sharp SED features at μm wavelengths as the continuum is enhanced. Due to its physical shape and high temperatures (up to $\sim 1\,000$ K), this dust component is usually called the warm (dusty) torus, and the entire accretion process with subsequent emission is described as active galactic nuclei (AGN).

There are also interesting phenomena on both extreme ends of the galaxy spectrum. More energetic than UV, X-rays trace extremely hot environments with gas temperatures reaching millions of kelvins. Violent physical processes must occur to reach such kinetic energies, for example matter accretion onto a star (X-ray binary stars), supernova remnants, accretion onto a SMBH, i.e. AGN activity (e.g. Corral et al., 2011), hot gaseous coronae of galaxies (e.g. Bogdán et al., 2015), and shocks in galactic clusters (e.g. Markevitch et al., 1999). These processes, pushed to further extremes, can also produce even more energetic gamma-rays. The catastrophic release of energy from stellar mass objects gives rise to gamma-bursts, the most energetic form of radiation observed in the universe (e.g Piron, 2016).

On the opposite end of the spectrum, at very long wavelengths (centimeter and larger), more traces of already mentioned violent galaxy processes can be found by observing radio emission from relativistic electrons as they gyrate in magnetic fields (e.g. Burbidge, 1956). Additionally, the AGN activity can launch and sustain collimated kinetic jets of relativistic particles which emit strongly in the radio as they move through the interstellar and intergalactic media. Regarding non-violent processes, the most abundant element in the universe (Hydrogen) emits a radio line at 21 cm. The distinguishing feature of radio observations is that long wavelength waves are insensitive to microscopic dust in the galaxy, so no emission reprocessing occurs. Since this thesis explores the usage of radio emission in galaxy evolution studies, radio astronomy details will be discussed later in a separate section.

With an abundance of information available across the galaxy spectrum, it is obvious that a multiwavelength (panchromatic) approach is necessary in order to fully decipher galaxy properties and its dominant physical processes. Guided with this idea, libraries or templates of observed and synthetic galaxy spectra with various properties were compiled (e.g. Coleman et al., 1980; Bruzual & Charlot, 2003), to be used to fit the observed SED of a galaxy (e.g. Walcher et al., 2011). In the absence of an observed spectrum, photometry points can be used to estimate the redshift of a galaxy if left as a free parameter. Galaxy rest-frame templates can be shifted across wavelengths in the process of fitting the data points. If some prominent features are present in the observed filters such as the Lyman (912 Å) or the Balmer (3 646 Å) breaks, the photometric redshift estimate can be precise (see also Hildebrandt et al., 2010). In this thesis,

photometric redshifts are used extensively as they often represent the only way to estimate the distance to a galaxy. Under the assumption that the fitted template describes the galaxy well, it is possible to constrain its properties such as the stellar mass, the star formation history, the dust content and obscuration, the contribution from the SMBH accretion, extrapolations of the spectrum to unobserved frequencies and many others. A particularly interesting property that can be inferred from the SED fits, which is relevant for this thesis, is the integrated IR light emitted by heated dust.

1.2.3 Evolutionary paths

With some general morphological and spectral properties of observed galaxies explained in the previous sections, it is now possible to link them together through cosmic time. The discovery of galaxy bimodality in morphology (and color) posed a question about their origins (see also Baldry et al., 2004). This was presented as a "nature" versus "nurture" problem (see Thuan et al., 1992): do initial conditions set the galaxy on its evolutionary path, or do environment and interactions drive the galaxies into their final shape. In other words, are the two galaxy classes evolving independently in parallel or one becomes the other.

An important clue is that the total stellar mass density of the universe doubled since redshift $z \sim 1$ (roughly 7 billion years ago), but the growth was dominated by the stellar mass increase of passive galaxies on the red sequence, while the integrated stellar mass of blue galaxies remained mostly unchanged (Borch et al., 2006). This finding strongly supports the idea of hierarchical growth through galaxy mergers. In this model blue galaxies form stars until a merger with another galaxy destroys its disk and turns it into a spheroidal (e.g. Baugh et al., 1996). Merging disrupts normal star formation processes, but also increases the total mass of the galaxy.

Galaxy color and stellar mass distributions of galaxies are shown in the left panel of Fig. 1.4. Two peaks can be seen in this plot representing two galaxy populations: i) blue cloud late types and ii) red sequence early types. Judging from the uniform and low density of all galaxies in the green valley (color-mass parameter space between the two galaxy types), one might conclude that the process of evolution from the blue cloud to the red sequence is smooth and quick and that galaxies in the green valley represent an intermediate step in this evolution. Most late-type galaxies live on the main star-forming galaxy sequence, a tight (within a factor of 2) nearly linear correlation between star formation rates and stellar mass (e.g. Noeske et al., 2007). In order to remove galaxies from the main sequence, its star formation must be quenched and different scenarios have been proposed (see e.g. Faber et al., 2007). A merger event can trigger an intense star formation phase by bringing the gas to the center of the galaxy. Such levels of star formation cannot be sustained for long, and the galaxy quickly depletes its entire gas supply
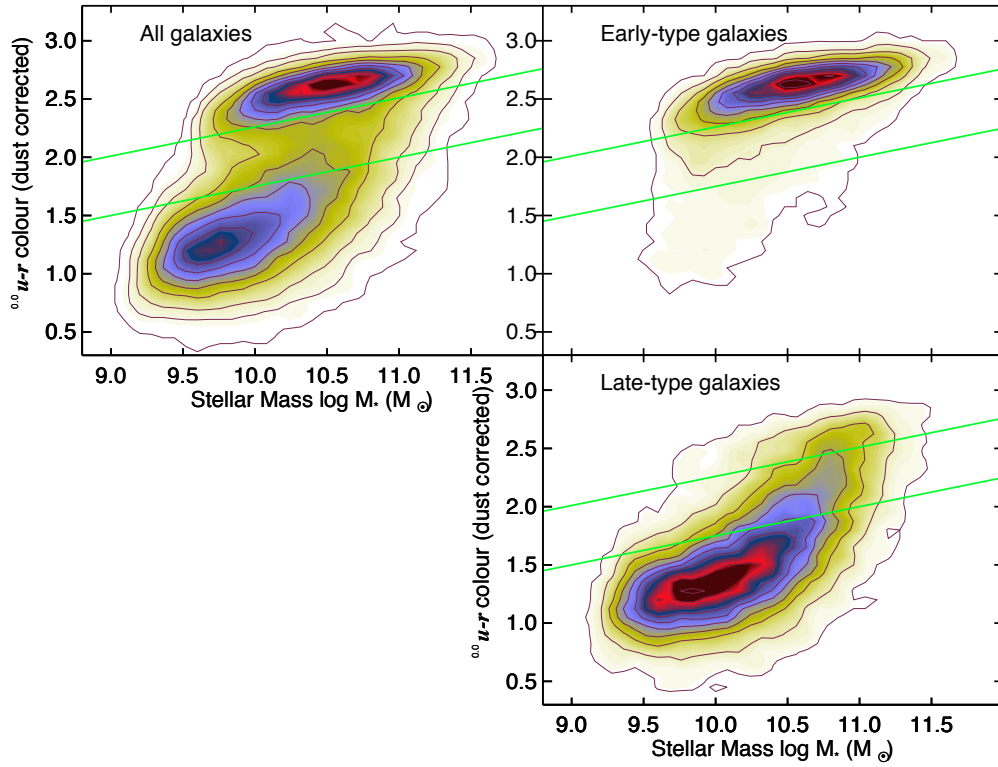
Figure 1.4: Galaxy color and stellar mass distributions (reproduced from Schawinski et al., 2014). Green lines indicate the green valley parameter space between late-type and early-type galaxies. Contours show the linear density.

and becomes quenched. In another scenario, AGN or supernova driven outflows can heat up the gas, or completely expel it thus preventing further star formation (e.g. Croton et al., 2016).

If a morphology cut is applied to galaxies, paired with color and stellar mass information, the evolution picture gets more complicated (two right panels in Fig. 1.4). Distribution of late types show a long tail toward the red and massive galaxies, implying slow processes ($\tau > 1$ Gyr), while there are very few early types with green and blue colors, implying fast processes ($\tau < 250$ Myr). The green valley now becomes a superposition of distribution tails from the two morphologically distinct galaxy populations. From this analysis it can be argued that there are two dominant evolutionary tracks (Schawinski et al., 2014): i) the slow one is triggered by blocking the inflow of cosmic gas into the galaxy causing it to slowly deplete the remaining gas and become a "red spiral", and ii) the fast track involving a major merger followed by the destruction of the disk and possible outflow episodes. In both cases the quenching event introduces a change in the blue galaxy environment that causes the galaxy to move off the main sequence. The actual path through the green valley and the time it takes to complete it is a strong function of galaxy morphology. In this thesis, the evolution of two main populations of galaxies

is traced through cosmic times. These populations are outlined in the next two sections.

1.2.4   Star-forming galaxies

The increase of stellar mass on the red sequence is mostly driven by galaxy mergers, but galaxies in the blue cloud are actively forming new stars from the cold gas (see Bell, 2004). By following the history of star formation it is possible to trace the cosmic buildup of stellar mass, and quantify evolution of galaxies across time. Since blue cloud galaxies are the dominant population in this regard, the nomenclature is changed from the observational property, its blue color, to a physical property: a star-forming (SF) galaxy. As mentioned before, SF galaxies exhibit a tight correlation between their star formation rates (SFR) and stellar mass - the galaxy main sequence. Galaxies with intense star formation are usually named starbursts (often SFR $>$ $100 \, M_\odot \, yr^{-1}$) and lie above the main sequence. They are rich in gas, often very dusty with perturbed morphology indicating recent interaction history. Abundant dust heated by many newborn stars will result in an SED dominated by IR emission ($L_{8-1\,000\,\mu m} > 10^{12} \, L_\odot$), and from an observational point of view they are also called ultra luminous infrared galaxies (ULIRGs). They will deplete the gas reservoir on timescales much smaller than the age of the galaxy, and therefore likely represent an intermediate state between the blue cloud and the red sequence (e.g. Sanders & Mirabel, 1996). Although they form stars $10 - 100$ times faster than normal galaxies, they are not as abundant and therefore contribute only a small part to the total cosmic production of new stars (see also Rodighiero et al., 2011). Galaxies with depleted gas reservoirs do not provide the universe with new stars and therefore lie below the main sequence.

For normal SF galaxies a simple gas regulation model can be used to understand the origin of the galaxy main sequence (Lilly et al., 2013). This model is based on continuous inflow and outflow rates and simple linear dependencies. In terms of star formation, the important constituents of a galaxy are its dark-matter halo, setting the overall gravitational well, and stellar and gas content. The gas is supplied into the halo from the environment at some rate, and a set fraction reaches the galaxy, enriching its gas supply. The gas is depleted by continuous star formation which is proportional to gas mass. Stellar evolution produces new metals which are returned to the galaxy at the end of the star's life. Finally, stellar winds, proportional to the SFR, drive a fraction of gas from the galaxy, and even its halo. Such a model can tie together different observed aspects of galaxy evolution, namely similar time evolution of specific star formation rate (SFR per unit mass) and halo growth, gas metal content in different galaxy populations and the ratio of stellar to dark matter mass.

Studies of time evolution of the main sequence has shown that the tight linear correlation holds up to high redshifts with only an increase in the normalization (e.g. Dutton et al., 2010).
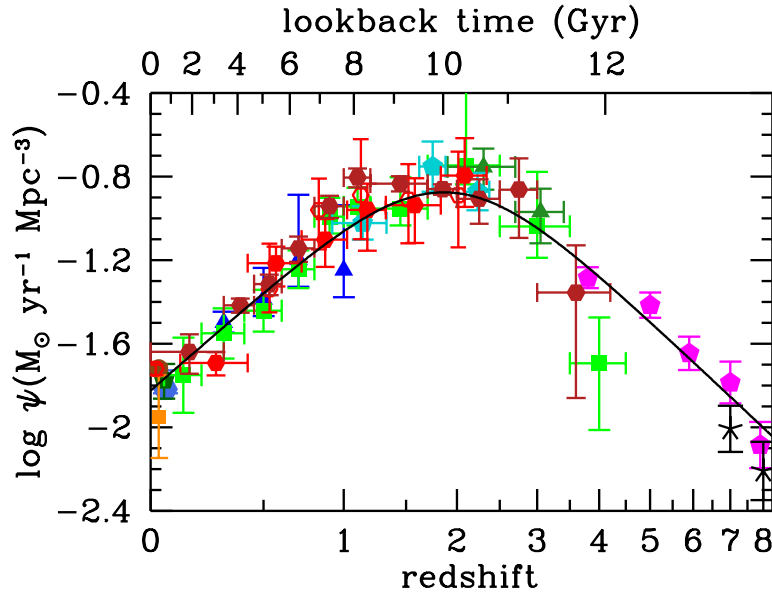
Figure 1.5: Cosmic star formation rate density history (reproduced from Madau & Dickinson, 2014). It shows the amount (mass) of stars created from gas per year and cubic mega-parsec as a function of cosmic time. Red and orange points are from IR measurements, while green, blue and purple points are from UV observations.

This increase means that galaxies of the same mass were forming stars more efficiently in the younger universe than today. A simple way to visualize the growth of new stars throughout cosmic time is presented in Fig. 1.5. It is obtained by integrating SFR contributions from all SF galaxies at a given time epoch in a given volume of space. This kind of information is a valuable input or a cross-check for semi-empirical simulations of galaxy evolution (see e.g. Behroozi et al., 2013). Several conclusions can be drawn from this figure: i) the cosmic star formation history has a peak at redshift $z \sim 2$, ii) it declines exponentially toward the present day with a total decrease of one order of magnitude, and iii) it also declines toward the high redshift universe. The right part of the diagram (early universe) is less constrained by having only UV measurements, which are heavily affected by dust attenuation (measured by e.g. Meurer et al., 1999). The main goal of this thesis is to use observed radio emission, which is not absorbed by dust (due to its long wavelength), to better constrain the high-redshift part of the history diagram. The cosmic SFR history encompasses the interplay between gas accretion, stellar feedback and the frequency of merger events. One aim of present and future surveys is to constrain the cosmic SFR as best as possible using a multiwavelength approach.

1.2.5 Active galactic nuclei

Beside stars, galaxies can have another powerful energy emitting source which is caused by accretion of material onto a SMBH. The amount of energy released from the small region (compared to the size of the galaxy) of the active nucleus during this process can outshine all other sources of light in the galaxy. Many galaxies are composites, where stellar and AGN light have varying contributions to the total emission. However, if accretion processes dominate, the galaxy is usually called an AGN. Furthermore, it is common to divide all galaxies into two populations: SF galaxies and AGN; a practice also employed in this thesis. This classification, unlike optical colors which divide galaxies into blue and red, takes into account a broader emission range, optimally from X-rays to radio. It is important to be able to distinguish the dominant origin of the radio emission since synchrotron radiation can also arise from supernova remnants. When radio luminosity is a proxy for tracing a specific physical process (such as star formation or AGN feedback), the AGN are usually considered as contamination of the SF galaxy sample, and vice versa. Different classification methods are used to separate galaxies into these two types, and some of them are discussed in this thesis.

Because AGNs are unresolved and do not have spherical symmetry, they can have different apparent properties depending on the viewing angle. When this was recognized, multiple types of objects were collected into a single unified model of AGN (Antonucci, 1993). In this model the central SMBH is surrounded by an accretion disk and a torus of dust. Material closer to the SMBH, if the view is not obscured by the torus, is dominated by broad spectral lines due to high temperatures ($10^3 - 10^4$ K) and therefore high velocity dispersion. If the dusty torus is in the line of sight, the broad line region is obscured. What remains are the narrow line features emitted from gas located at a larger distance from the SMBH at lower temperatures, along with the IR continuum from the torus.

All AGNs can be divided into two categories depending on the physics of the accretion (Heckman & Best, 2014). The Eddington luminosity or limit, $L_{\mathrm{Edd}} \approx 1.3 \times 10^{31} (M/M_\odot)$ W, sets the maximum possible emission obtainable in hydrostatic equilibrium and it depends on the mass of the SMBH. If the AGN luminosity is above about 1% L_{Edd} it is called radiatively efficient, resulting in mostly isotropic radiation (for more details see Heckman & Best, 2014). Such accretion is usually observed around less massive SMBHs in galaxies that can be found in the green valley (see e.g. Smolčić, 2009). On the other hand, radiatively inefficient AGN have luminosities below $\lesssim$ 1% L_{Edd}, and matter transport is dominated by advection, and energy is emitted in the form of bipolar collimated (highly nonisotropic) kinetic jets. These AGNs are usually hosted in massive elliptical galaxies which occupy the red sequence; devoid of ongoing star formation they would appear quiescent. However, radio observations sensitive to
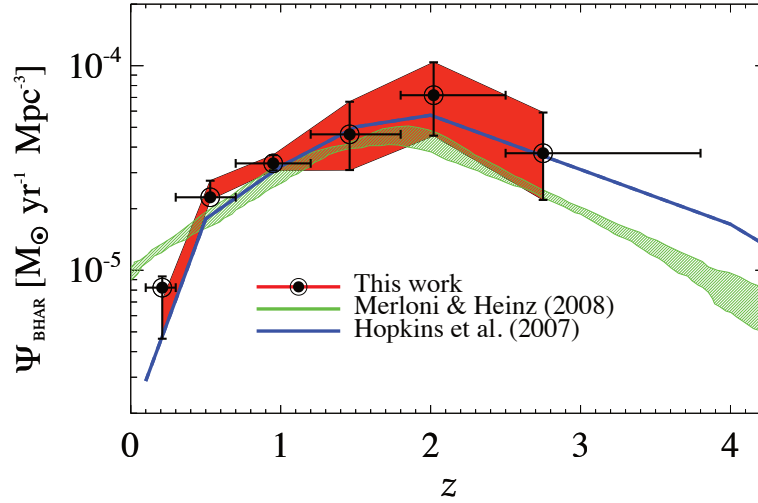
Figure 1.6: Supermassive black hole accretion rate density history (reproduced from Delvecchio et al., 2014).

synchrotron radiation from the jets show high activity in their nucleus.

What makes AGN very interesting in terms of galaxy evolution is the observed correlation between cosmic SMBH accretion rate history (see Fig. 1.6) and star formation history (see Fig. 1.5). They both share a similar shape, peak around $z \sim 2$ and decrease by an order of magnitude toward the present day. This correlation points to a concept of co-evolution between the SMBH and its host (e.g. Häring & Rix, 2004). The SMBH and galaxy grow in parallel, yielding an almost constant ratio between SMBH mass and stellar mass (Mullaney et al., 2012). The correlation holds irrespective of different AGN trigger mechanisms and galaxy host mass. When morphology is considered, in late types the most massive SMBHs are growing, and in early types the least massive SMBHs are preferentially growing (Schawinski et al., 2010). A complex interaction between gas inflow, supply and depletion; AGN and stellar feedback; and SMBH and stellar mass growth results in an intricate self-regulating mechanism of galaxies. Finding and connecting the underlying mechanisms is the goal of galaxy evolution studies.

1.3 Radio astrophysics

Observations at radio wavelengths are used in this thesis to probe star formation in galaxies. As a brief introduction into the field of radio astronomy, the following sections describe some aspects of radio observations such as its history, physical scenarios, technical realizations and challenges.

1.3.1 The beginning of radio astronomy

Radio astronomy emerged as a scientific field from studies of interference and noise in radio communications. Karl G. Jansky, an engineer working at Bell Telephone Laboratories, used a 14.6 m antenna to detect a faint and persistent signal which originated from the sky and had an oscillation period of exactly one sidereal day (Jansky, 1933). This implied that the source was stationary with respect to the stars and could not originate on Earth. He also noted that this radiation could not be of stellar origin, because the Sun was not a sufficiently strong emitter at this wavelength. The radio source was located in the constellation of Sagittarius and coincided with the center of the Milky Way. In the next decade, the first radio sky maps were constructed showing several local maxima in the radio emission (e.g. Reber, 1944), however the angular resolution of single dish radio telescopes was several degrees, insufficient to connect the radio emission with known objects observed at optical wavelengths. Different scenarios were proposed to explain this radio electromagnetic phenomena. If it were thermal in origin, such as Bremsstrahlung (braking radiation) from collisions between electrons an ions, the gas temperature would have to be extremely high, roughly $\sim 150\,000$ K. Radio stars were proposed to explain this mysterious emission (see Alfvén & Herlofson, 1950). The actual underlying physical mechanism are the cosmic-ray electrons gyrating in the magnetic field of the host galaxy (for further details see e.g. Mo et al., 2010; Burke & Graham-Smith, 2010) Application of interferometric principles, developed in the 1950s (discussed in the following section), enabled high resolution (arcsec and below) radio observations and therefore easier cross-correlation with optical sources. The strongest extragalactic radio source was connected to the optically faint and distant galaxy Cygnus A where, interestingly, the radio emission extended much further than the host galaxy (e.g Jennison & Das Gupta, 1953; Perley et al., 1984). Another radio source was shown to coincide with the Crab Nebula, a young supernova remnant in existence for only a thousand years (e.g Bolton et al., 1949; Velusamy et al., 1992). In the 1960s, peculiar objects were discovered that exhibited strong pulses at regular second like intervals. The theoreticians explained them as rapidly spinning neutron stars - so called pulsars (Hewish et al., 1968). Another important discovery was strong and compact radio sources with no obvious visual counterparts (e.g. Matthews & Sandage, 1963). They were named "quasi-stellar objects" or quasars and were a complete mystery at the time. We now know that these objects are actually AGN viewed from a specific angle. Radio astronomy thus became an excellent tool for studying previously unknown exotic objects in the universe, and from there it grew into an indispensable window for the studies of galaxies.

1.3.2 Radio interferometry

A radio telescope is an instrument designed to collect electromagnetic radiation with long wavelengths ($\gtrsim 1$ cm). In its simplest form it is a dipole antenna, followed by an amplifier connected to a detection system. A good telescope will at the same time try to optimize two properties: sensitivity and angular resolution. The sensitivity scales with the collecting area of the telescope and it benefits from all noise reduction measures (e.g. cooling of electronics or large distance from human made radio emission). The angular resolution θ of a telescope is diffraction limited by the Rayleigh criterion $\theta(\mathrm{rad}) = 1.22\lambda/D$ (for circular aperture), where λ is the observed wavelength and D is the diameter of the aperture. It is immediately apparent that large wavelength observations will suffer from much lower resolution than for example optical observations with the same aperture size. It is also challenging to build a large (> 100 m) and structurally stable dish, especially with steering capability.

A solution was provided by Martin Ryle at the Cambridge University who developed the aperture synthesis technique. For this contribution, he was awarded the Nobel Prize in 1974 (shared with Antony Hewish), the first one awarded in the field of astrophysics. The smallest functioning element producing an output signal in this approach is one pair of antennas. A radio wave traveling from a celestial source takes two different paths to reach each of the antennas, and the difference in geometrical paths taken produces a phase shift between the signals. If two antennas are separated by a vector **b** and the radio source is in the direction **s** (unit vector), the wave front generally does not reach both antennas at the same time, it reaches the more distant one after a time delay of $\tau = \mathbf{b} \cdot \mathbf{s}/c$, where c is the speed of light. This corresponds to a phase shift of $\phi = 2\pi\nu\tau$, where ν is the frequency. The recorded output is the cross-correlation of the two signals (multiplied and time averaged), which holds information about the amplitude and the time delay (for a more rigorous and mathematical explanation see e.g. Burke & Graham-Smith, 2010). As the radio source moves on the sky due to Earth's rotation, the phase shift changes, and the instrument records an interference pattern, thus earning the name radio interferometer. The vector **b** is tied to the Earth's coordinate system and is called a baseline, the building block of an interferometer. In modern systems the interference of all antenna pairs are done digitally by a specialized computer called the correlator. It is also possible to create an interferometer with only one antenna by using the Lloyd's mirror. One example is the antenna mounted on top of a Dover Heights cliff in Sydney overlooking the sea which acted as a large mirror (Bolton, 1982). The real signal from the sky and the reflected one from the sea would produce an interference pattern due to small differences in geometrical paths as the radio source rises with time.

One way to understand how the sky appears in interferometric observations is to imagine

that a baseline (antenna pair) projects an interference pattern on the sky with a fringe separation of $\theta \sim \lambda/|\mathbf{b}|$, where λ is the observed wavelength. If a radio source falls into the constructive interference fringe, the signal will be strong, if it falls into the destructive interference fringe, the signal will be absent. Since "the densest" interference pattern defines the best angular resolution achievable, by placing two antennas far apart (the separation can be thousands of kilometers) an interferometer can resolve details beyond any single aperture telescope. The true power of aperture synthesis is achieved when many antennas are used simultaneously (see Fig. 1.7). For N antennas, there are $N(N-1)/2$ baselines and each one "casts" a different interference pattern on the sky. An interferometer does not observe the sky as an ordinary telescope, it samples the Fourier transform of the sky brightness distribution. Formally, baselines are described with $\mathbf{b}/\lambda = (u, v, w)$, where the three coordinates correspond to Earth's east-west, north-south and up-down components respectively. The sky positions are defined as $\mathbf{s} = (l, m, \sqrt{1 - l^2 - m^2})$, where l and m coordinates correspond to the sky's east-west and north-south direction, while the third component ensures a unit length of this direction vector. By integrating over the entire sky, the interferometer records visibilities V as a function of u, v, w coordinates, for a simpler 2D case ($w = 0$) this can be written as

$$V(u,v) = \iint I(l,m) \cdot e^{-2\pi i(ul+vm)} \frac{dl\,dm}{\sqrt{1 - l^2 - m^2}}, \tag{1.1}$$

where $I(l,m)$ is the sky brightness distribution. The main difficulty lies in the fact that a finite amount of telescopes can never sample all Fourier components, even though the Earth's rotation causes the baselines to change with respect to the sky, thus filling the *uv* coordinates with time. When the Fourier inverse transform is applied to discretely and incompletely sampled visibilities, the obtained sky brightness map, which is the interesting physical quantity, will have remaining artifacts (see Fig. 1.8). The *uv* data can be gridded and weighted during the inversion, providing a tradeoff between sensitivity (more weight to abundant short baselines) and resolution (more weight to long baselines). The process of imaging the radio interferometric data includes usage of different techniques designed to mitigate known issues and produce a robust radio sky map. All steps performed to obtain the final map from the raw data, such as phase and flux calibration, removing corrupted data, imaging and cleaning the image, correction for antenna response and others, is called data reduction. Obtaining a tested and reliable radio map is one of the most important data products of radio continuum observations. These technical aspects of radio astronomy, depending on the volume of the data, may be extremely time consuming and are covered in detail in the next two chapters of this thesis.

The radio map at some frequency is a 2D surface brightness image with units of jansky

Figure 1.7: Karl G. Jansky Very Large Array radio interferometer in New Mexico, USA. Image courtesy of NRAO/AUI.
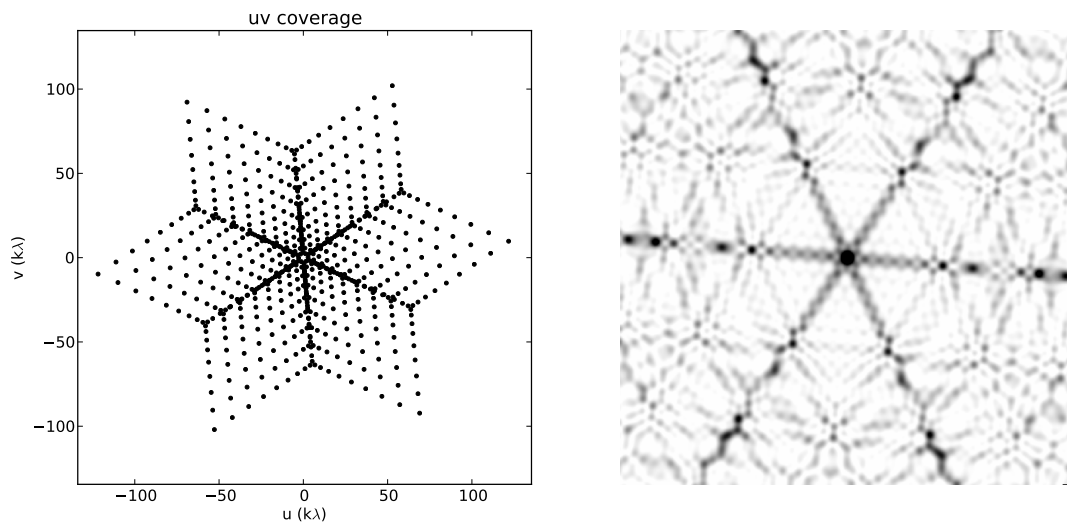
Figure 1.8: *Left panel:* *uv* coverage of a single snapshot at one frequency channel with the Karl G. Jansky Very Large Array. Each baseline provides two points (due to symmetry) and the Y-shape antenna configuration gives rise to the specific shape of the plot. With long observations, each point would trace an ellipse-like path in the *uv* plane due to Earth's rotation, thus sampling more of the Fourier components. *Right panel:* point spread function (PSF), also known as a dirty beam, obtained by Fourier inversion of the *uv* coverage. The central circle represents the beam size. Six strong sidelobes and additional artifacts are caused by an incomplete sampling of the *uv* plane. The radio interferometer "sees" the true radio sky brightness distribution convolved with the PSF.

per beam (Jy/beam), where $1\,\text{Jy} = 10^{-26}\,\text{W}\,\text{m}^{-2}\,\text{Hz}^{-1}$ is the flux density. The synthesized beam (in units of arcsec^2) is the resolution element obtained from the *uv* coverage. Details on angular scales below the beam size cannot be resolved and all emission within the beam is integrated and represented as a one pixel value. An interesting consequence of this is that the telescope sensitivity is not measured in terms of flux densities (magnitudes), but surface brightness (Jy/beam). Therefore, if two sources have the same flux density, the compact (point like) one is more likely to be detected than an extended one which spreads across multiple beams. The Rayleigh-Jeans approximation for the low energy black-body radiation can be used to assign an effective temperature for a radio source of specific surface brightness, thus removing the beam from further calculations.

1.3.3 Karl G. Jansky Very Large Array interferometer

The main instrument which provided around 400 hours of the raw radio data used in this thesis is the upgraded Karl G. Jansky Very Large Array (VLA). The telescope is situated on the Plains of San Agustin in the USA state of New Mexico at an altitude of approximately $2\,000$ m. It is a component of the National Radio Astronomy Observatory (NRAO[2]). There are 27 radio antennas, each 25 m wide, distributed on site in a Y-shaped configuration (see Fig. 1.7). The total collecting area of all antennas is equivalent to a 130 m dish. Aluminum panels constituting the dish surface are accurate up to 0.5 mm. Distances between the dishes can be increased or shortened by repositioning the antennas along the Y-shaped rails into different arrays. In its most extended A-array configuration, the maximum baseline of the VLA is 36 km long, thus the achievable angular resolution is that of a 36 km wide single dish. This separation can be shortened to just 1 km as the telescope is put into its most compact D-array configuration, where the minimum separation between antennas is just 35 m. Such a configuration is very sensitive to extended diffuse emission. The interferometer can thus accommodate multiple observation strategies in terms of resolution and sensitivity.

The original Very Large Array was built in the 1970s. Recently, it underwent a major upgrade that was emphasized by renaming it into the Karl G. Jansky Very Large Array. By increasing the bandwidth available for correlation to at least two GHz in each band, the instantaneous sensitivity was immensely improved providing 5 to 20 times better radio continuum sensitivity. With new electronics, receivers and the correlator (computer that interferes signals from antenna pairs), the VLA now provides a continuous coverage from 1 to 50 GHz (from 30 cm to 6 mm), with two additional narrower bands centered around 90 cm and 4 m.

---

[2]The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. (AUI).

Currently, the next upgrade is being planned, the so called Next Generation Very Large Array (ngVLA[3]). This upgrade would significantly enlarge the collecting area with new antennas spaced even further and closer apart (for improved resolution and sensitivity) and add new receivers probing up to 116 GHz.

1.3.4   Radio spectral energy distribution

Radio emission in galaxies originates in two distinct physical processes. The first one is the thermal Bremsstrahlung radiation from electrons scattering off ions which occurs in regions of ionized hydrogen (denoted as HII regions). This process is also called the free-free emission because the electrons are not captured in the process and particles remain free after the interaction. The second one is the non-thermal synchrotron radiation from relativistic electrons spiraling in magnetic fields (for more details see Condon & Ransom, 2016). Electron populations in the galaxy cover a wide range of parameters in terms of densities, opacities, kinetic energies, magnetic field strengths and orientations. A combination of all single electron spectra yields the total radio SED which is featureless and built from simple power laws with turnovers at critical frequencies. For this reason the radio SED is often described in terms of the spectral index α defined as $S_\nu \propto \nu^\alpha$ or $\alpha = d \ln S_\nu / d \ln \nu$, where S_ν is the flux density emitted at frequency ν.

The free-free spectrum at low frequencies is similar to a black-body ($\alpha \lesssim 2$) due to the opaqueness of the medium, and turns over to $\alpha \approx -0.1$ at higher frequencies due to a frequency dependent absorption coefficient. The synchrotron spectrum of an optically thin galaxy is a simple power law with $\alpha \approx -0.7$. This value is tightly correlated with the energy distribution $N(E)dE$ of cosmic ray (relativistic) electrons which is also a power law across at least two decades of energies. These radio spectra constituents are demonstrated in Fig. 1.9. The above spectral index values are obtained from simple models which are naturally only approximations (see Condon & Ransom, 2016). Understanding the radio SEDs of galaxies is necessary to properly infer their intrinsic rest-frame luminosities at different redshifts (needed for evolution studies). For this reason, potential biases from unconstrained radio spectral indices are discussed at several places in the thesis.

The radio emission traces astrophysical processes causing ionization of the Hydrogen clouds and high densities of cosmic ray electrons. Young bright stars photoionize the surrounding medium. After a lifetime measured in Myr these stars explode as supernovae. If they are massive enough ($M \gtrsim 8M_\odot$), they release a sufficient amount of energy in shocks to supply the galactic medium with cosmic rays (see also Condon, 1992). As relativistic electrons diffuse

---

[3]See also https://science.nrao.edu/futures/ngvla
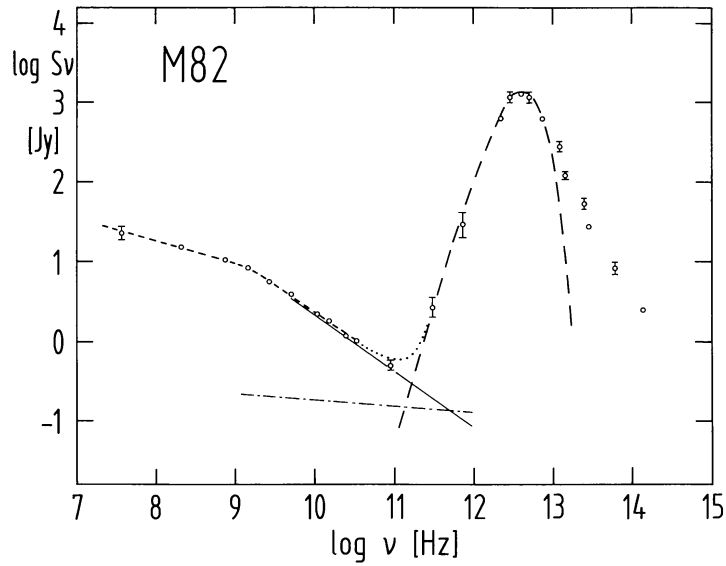
Figure 1.9: Spectral energy distribution (flux density as a function of frequency) of the starburst galaxy M82 from the radio to the far infrared (reproduced from Klein et al., 1988). The solid line shows the synchrotron emission, the dot-dashed line shows the free-free thermal emission and the dotted line is the sum of these two radio contributions. The long-dashed line above 100 GHz shows the dust thermal emission. The short dashed line demonstrates that the synchrotron radiation is absorbed by the emitting medium below 1 GHz thus reducing its intensity, this is also called synchrotron self-absorption. The emitted radio power is only a fraction of the bolometric power emitted from the radio to the IR.

throughout the galaxy they emit synchrotron radiation. Radio observations can therefore trace star formation by being sensitive to stellar deaths. A slight difficulty is that cosmic rays can live tens of millions of years, and given enough time, they can smear out details of spatial location of supernovae activities through diffusion across the galaxy. Free-free emission is a more direct indicator of young stars, however it contributes to only a fraction of the total emitted radio power and it is easily drowned by dominating synchrotron and dust emission (see Fig. 1.9).

To measure star formation in terms of gas mass converted to stars, it is important to make sure that the emission is indeed caused by stellar processes since accretion onto the SMBH also releases enormous amounts of energy. There are several indicators of this AGN activity visible in the radio. Jets are an obvious one if they can be resolved. The radio spectrum can also deviate strongly from the usual spectral index of $\alpha = -0.7$, for example it might be completely flat, or even convex with a peak at GHz wavelengths (e.g. Callingham et al., 2017). To produce such a spectrum emitting regions could be highly compact (less than 1 kpc) and inhomogeneous. As discussed in the next section, in this thesis the radio emission is compared with the IR light to gain insights into the dominant astrophysical process in galaxies.

1.3.5 The infrared-radio correlation

Given different timescales of various stellar processes such as gas collapse, photoionization of the surrounding medium by hot stars, light reprocessing by dust, life of a star before supernova explosion and cosmic ray diffusion; it is remarkable that an empirical correlation between radio and IR emission was found (e.g. van der Kruit, 1971; de Jong et al., 1985; Helou et al., 1985). This IR-radio correlation, which is almost linear, is valid across several orders of magnitude in luminosities. Although the IR-radio correlation is not fully understood from a theoretical perspective, it provides a powerful tool to consistently link radio luminosities with star formation. Infrared emission from dust heated by young, bright stars is easier to connect with star formation than the non-thermal synchrotron emission from diffusing relativistic electrons, which dominates the radio emission. The observed empirical relation between flux densities at these two wavelengths provides a way to estimate the amount of newly formed stars in relatively recent galactic history (~ 100 Myr) using only radio observations (see also Yun et al., 2001). The great strength of this approach is that radio waves, once emitted, practically do not interact with the medium they pass through (in terms of scattering or absorption leading to a decrease in signal strength). Therefore, no dust corrections are necessary, i.e. the radio measurement is dust-unbiased. However, without the IR-radio correlation, it is more difficult to derive star formation rates using the radio data alone (see also Murphy et al., 2011). Before using the IR-radio correlation, it is necessary to calibrate it as a function of cosmic time (e.g. Sargent et al., 2010) and radio spectra (e.g. Tabatabaei et al., 2017) using a well understood sample of galaxies. The IR-radio correlation is valid for SF galaxies where the radio emission is fueled by supernova remnants. Galaxies where the AGN significantly contributes to the total radio luminosity (e.g. radio jets) are outliers from the correlation, thus the IR-radio correlation can be used as an additional diagnostic for galaxy classification (see e.g. Delhaize et al., 2017). A similar approach in identifying AGN was used in this thesis by comparing the radio luminosity with the bolometric IR light integrated from the SED star-forming template. A radio excess can suggest an existence of unresolved radio jets, and thus AGN activity.

1.3.6 Survey approach to galaxy evolution

The sky spans more than $40\,000$ deg^2, for comparison, the Moon is only half a degree in diameter (area of 0.2 deg^2). Each square degree can contain hundreds of thousands of galaxies. The galactic plane covers more than a thousand square degrees and contains many stars and dust which obstruct the view toward extragalactic objects. With so many objects on the sky it is impossible to survey them all. Surveys at all frequencies are designed to optimize the number
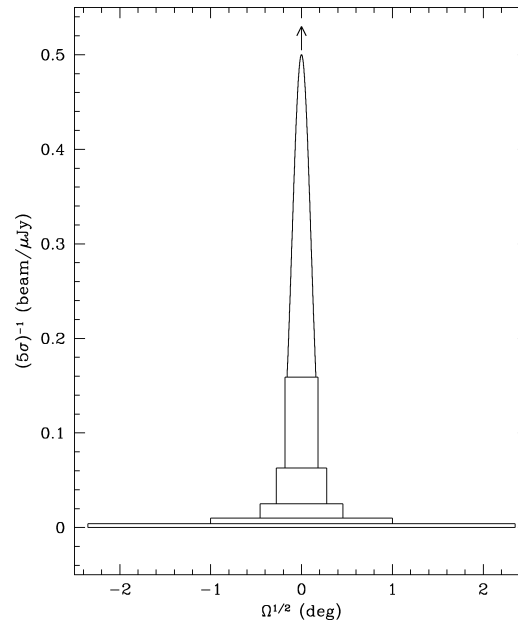
Figure 1.10: Multitiered observational setup commonly called the wedding cake (reproduced from Condon, 2015b). Inverse 5σ root mean square noise is on the y-axis (larger number implies better sensitivity) and the square root of the sky area covered is represented with the width on the x-axis. Due to telescope operating time constraints, an acceptable trade-off between covered area and sensitivity must be chosen: wide and shallow, or deep and narrow. A combination of different depths and areas can cover a broad range of scientific goals.

of observed environments using the so called wedding cake approach (see Fig. 1.10). A large sky area must be surveyed to find rare objects, while probing the early universe means detecting very distant and faint objects. Ultra deep (long exposure) surveys are essential for probing galaxy evolution and a small solid angle may still correspond to a large volume slice at high redshift. The main danger lies in the so called cosmic variance (see e.g. Moster et al., 2011). The usual assumption of an isotropic universe means that properties of one small sky patch are carried over to the full sky. If a survey falls into a randomly dense or sparse region of the universe, wrong conclusions might be drawn about galaxy populations given the non representative biased sample observed. For this reason it is important to observe multiple patches on the sky with variable area coverage. As the technology progresses, survey speed is constantly being increased. From early manual photographic plates, to computer data rates of several Terabytes per hour, astronomy is a never-ending quest for more data. The concept of deep and shallow surveys will probably remain, however the absolute sensitivity behind this nomenclature changes with each passing decade.
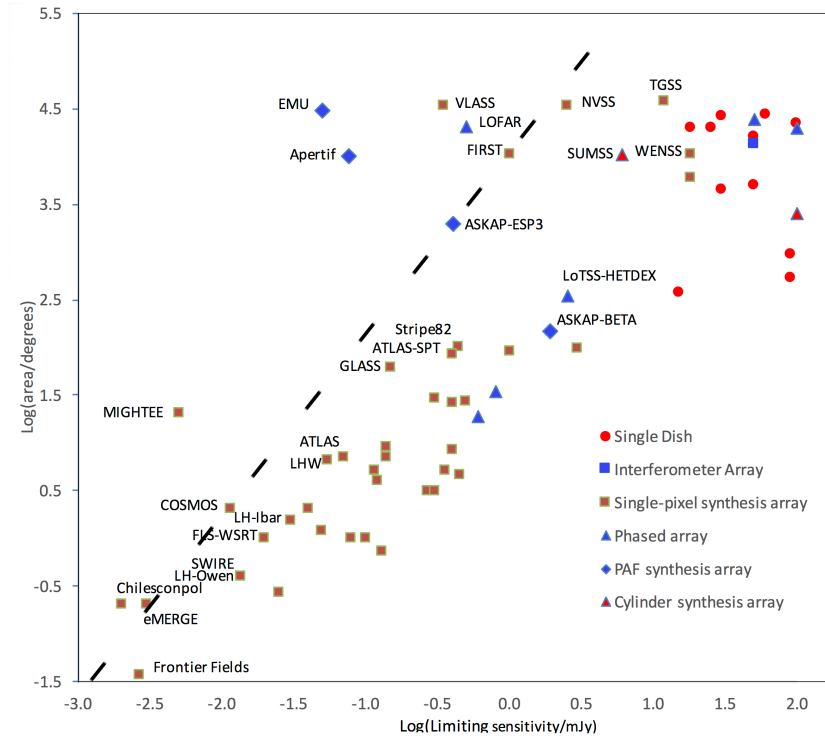
Figure 1.11: The sky area covered down to 5σ sensitivity of modern radio surveys (reproduced from Norris, 2017). The dashed line marks the boundary of existing surveys. The upper left corner of the plot is the frontier for new surveys representing large area covered with high sensitivity.

The COSMOS survey

This thesis focuses on the Cosmic Evolution Survey (COSMOS[4], Scoville et al. 2007) observed at radio wavelengths with the recently upgraded VLA (see Fig. 1.7) interferometer. The COSMOS survey is a panchromatic study of the equatorial 2 deg^2 (10 full Moons) field centered at $\alpha_{2000.0} = 10^h00^m28\overset{s}{.}6, \delta_{2000.0} = +2°12'21''0$. It has been observed with all leading ground-based and satellite facilities from radio to X-ray yielding a rich multiwavelength dataset of more than a million objects with accurate photometric redshifts (Laigle et al., 2016). The area is large enough not to suffer from severe cosmic variance issues (by observing an overly small and non representative volume of the universe) and its location on the sky was chosen to be devoid of bright sources. The COSMOS survey is therefore optimized for probing the faint sky and the current data reaches the time when the universe was only one billion years old.

Future radio surveys

An overview of different radio surveys is given in Fig. 1.11. The COSMOS survey lies on the edge of modern surveys and paves the way for future projects. One of the goals of this thesis is to use the COSMOS field to derive the numbers of radio sources down to faint flux densities (μJy and below), which will be observed by the next generation telescopes. Such constraints are extremely useful when deciding on the area and the sensitivity (time duration) of a survey, as the survey must yield a statistically significant sample of studied galaxies in order to justify its scientific goal.

The largest radio facility currently planned is the Square Kilometer Array (SKA), which would ultimately have a collecting area of approximately 1 km^2. Its first phase, named the SKA-1, although being just a fraction of the full telescope, will already provide an order of magnitude increase in sensitivity compared to current facilities. A number of SKA precursor surveys are about to start observing. Some of them are the Australian SKA Pathfinder (ASKAP), the International GHz Tiered Extragalactic Exploration with the Meer Karoo Array Telescope (Mightee/MeerKAT; see Jarvis et al. 2017) and the The Evolutionary Map of the Universe (EMU; for more details see Norris et al. 2013). These next generation instruments and surveys will revolutionize radio astronomy and provide key new insights into the evolution of galaxies, star formation, AGN, hydrogen gas, magnetic fields and many others from the earliest days of the universe. It is important to mention that by probing large volumes of space, there is always a chance of unexpected serendipitous discoveries as was in the beginning of radio astronomy.

1.3.7   Counting galaxies

With the survey observations finished and data reduced, it is instructive to perform the simplest statistical analysis of the data first (see e.g. Becker et al., 1995). This includes counting sources or binning them in terms of flux densities. If the universe were Euclidean space containing a uniform distribution of sources (galaxies of all luminosities are equally probable), the number of sources observed above some flux density S would be $N \propto S^{-3/2}$. This implies differential number counts of $n = dN/dS \propto S^{-5/2}$. However, this proportionality is not observed and radio number counts provided the first evidence for the non-steady state universe (see Ryle & Scheuer, 1955). The differential number counts generally follow a power law across multiple orders of magnitudes in flux densities, however they contain an inflection below 1 mJy at 1.4 GHz. This can be understood in terms of two main radio populations: SF galaxies and AGN, each dominating at different flux densities. However, this inflection point proved to be difficult to constrain

---

[4]http://cosmos.astro.caltech.edu/

due to the low brightness of these galaxies and technical limitations of previous generation interferometers (for more details see Padovani, 2016). With the new VLA-COSMOS observations presented in this thesis, the counts are constrained throughout the inflection point.

The luminosity of a galaxy usually correlates with interesting properties such as the its physical size, mass, the amount of star formation or AGN activity. A more advanced counting statistic which provides abundant information about galaxy evolution, and is intrinsically tied to the number counts, is the luminosity function (LF). The LF measures the cosmic density of sources (number of galaxies per cubic Mpc), per logarithm of luminosity. To construct it, distances are needed and they are estimated from redshifts assuming an underlying cosmological model. Radio continuum at 3 GHz alone cannot be used to calculate the redshift of a galaxy because the SED lacks any features (e.g. spectral lines), a redshifted power law is still a power law. In this thesis, radio sources are therefore paired with an ancillary COSMOS optical/NIR selected catalog containing redshifts, and radio LFs are constructed since $z \sim 5.5$. An immediate result obtained from such counting is that less luminous galaxies are more abundant compared to bright ones, but the overall density increases with redshift (not indefinitely).

Constraining the LFs of various galaxy populations can help answer questions such as when galaxy assembly proceeded the most efficiently, what evolutionary paths of galaxies are, whether one galaxy type evolves into another and when, how often galaxies with extreme star formation occur and many others. The local extragalactic universe ($z \approx 0$) is in close proximity which makes studies easier. Properties of the universe observed locally are often assumed to be also valid at earlier time epochs in the absence of high-redshift data. Local radio luminosity functions (measured by e.g Condon et al., 2002; Mauch & Sadler, 2007) are often the starting point for further evolution analysis across cosmic time. The evolution model should be able to reproduce the observed number counts, but also predict what future radio surveys will observe with their increased sensitivities. Finally, by considering LFs obtained from observations at different wavelengths (e.g. radio, IR, UV) it is possible to connect the evolution of various galaxy populations into a single coherent picture of galaxy evolution. Such comparison is used in this thesis to provide an estimate of the amount of star formation enshrouded in dust as a function of cosmic time.

1.4   Thesis outline

This thesis is built around the VLA-COSMOS 3 GHz Large Project, which provided new radio observations of the COSMOS field. The radio data was obtained using ~ 400 hours of observations with the VLA resulting in the deepest radio map currently available across an area larger

than a square degree. The interferometric nature of the data provides a sub-arcsecond angular resolution, even though the observed wavelength is 10 cm. The final map contains around 5 000 radio sources (above 5σ) detected per square degree as will be shown in the following chapters. Such a dataset is excellent for studies of SF galaxies as well as AGN activity out to high redshifts ($z \sim 5$) and it can be used to trace the cosmic history of star formation since the universe was approximately 1 billion years old. This data was preceded by the VLA-COSMOS 1.4 GHz Large Project (Schinnerer et al., 2007, 2010), enabling simpler cross-checks of the data and also the analysis of the radio spectral indices. One particular benefit that radio data can bring to extragalactic studies is its dust-unbiased nature, i.e. no corrections for dust extinction are necessary.

This thesis is organized as follows. Chapter 2 demonstrates how an incomplete radio dataset can be used to simultaneously test the incoming data and provide further constraints on the exotic object: a potentially recoiling black hole CID-42. Chapter 3 provides the full description of the VLA-COSMOS 3 GHz Large Project, from the observational setup and reduction of the full 384 hours data, to a catalog extraction and completeness corrections. Counting statistics is used in Chapter 4 to derive LFs of SF galaxies and AGN. Obtained LFs are used to model radio number counts, which provide constraints for future radio surveys, as presented in Chapter 5. Chapter 6 focuses on SF galaxies and provides the main results of this thesis: the cosmic star formation rate density history since $z \sim 5$. Chapter 7 summarizes this thesis by emphasizing important results.

The main results of the thesis work have been published in five papers, and presented in one still in submission phase; these papers are as follows:

1. **Novak, M.**, Smolčić, V., Civano, F., Bondi, M., Ciliegi, P., Wang, X., Loeb, A., Banfield, J., Bourke, S., Elvis, M., Hallinan, G., Intema, H. T., Klöckner, H.-R., Mooley, K., Navarrete, F., 2015, MNRAS, 447, 1282, *New insights from deep VLA data on the potentially recoiling black hole CID-42 in the COSMOS field*

2. Smolčić, V., **Novak, M.**, Bondi, M., Ciliegi, P., Mooley, K. P., Schinnerer, E., Zamorani, G., Navarrete, F., Bourke, S., Karim, A., Vardoulaki, E., Leslie, S., Delhaize, J., Carilli, C. L., Myers, S. T., Baran, N., Delvecchio, I., Miettinen, O., Banfield, J., Baloković, M., Bertoldi, F., Capak, P., Frail, D. A., Hallinan, G., Hao, H., Herrera Ruiz, N., Horesh, A., Ilbert, O., Intema, H., Jelić, V., Klöckner, H.-R., Krpan, J., Kulkarni, S. R., McCracken, H., Laigle, C., Middleberg, E., Murphy, E. J., Sargent, M., Scoville, N. Z., Sheth, K., 2017, A&A, 602, A1, *The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release*

3. Smolčić, V., Delvecchio, I., Zamorani, G., Baran, N., **Novak, M.**, Delhaize, J., Schin-

nerer, E., Berta, S., Bondi, M., Ciliegi, P., Capak, P., Civano, F., Karim, A., Le Fevre, O., Ilbert, O., Laigle, C., Marchesi, S., McCracken, H. J., Tasca, L., Salvato, M., Vardoulaki, E., 2017, A&A, 602, A2, *The VLA-COSMOS 3 GHz Large Project: Multiwavelength counterparts and the composition of the faint radio population*

4. **Novak, M.**, Smolčić, V., Delhaize, J., Delvecchio, I., Zamorani, G., Baran, N., Bondi, M., Capak, P., Carilli, C. L., Ciliegi, P., Civano, F., Ilbert, O., Karim, A., Laigle, C., Le Fèvre, O., Marchesi, S., McCracken, H., Miettinen, O., Salvato, M., Sargent, M., Schinnerer, E., Tasca, L., 2017, A&A, 602, A5, *The VLA-COSMOS 3 GHz Large Project: Cosmic star formation history since z∼5*

5. Smolčić, V., **Novak, M.**, Delvecchio, I., Ceraj, L., Bondi, M., Delhaize, J., Marchesi, S., Murphy, E., Schinnerer, E., Vardoulaki, E., Zamorani, G., 2017, A&A, 602, A6, *The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z∼5*

6. **Novak, M.**, Smolčić, V., Schinnerer, E., Zamorani, G., Delvecchio, I., Bondi, M., submitted to A&A, *Constraints on submicrojansky radio number counts based on evolving VLA-COSMOS luminosity functions*

Chapter 2

Initial 130 h of VLA-COSMOS 3 GHz Large Project observations and a case study of a potentially recoiling black hole CID-42

The VLA-COSMOS 3 GHz Large Project observations were performed from November 2012 to May 2014. As the data was being taken it was necessary to test it and justify that the full observations will indeed be able to reach the proposed sensitivity limit of several μJy beam^{-1}. It was also important to test different calibrations and radio imaging algorithms as the telescope was recently upgraded, which significantly expanded its instantaneousness sensitivity. The central frequency of observations is 3 GHz, but the continuum is observed from 2 to 4 GHz. Testing of the incoming data was paired with the scientific goal of investigating an exotic object: a potentially recoiling black hole CID-42. This chapter presents the data reduction performed on the first third of the data, which was used to constrain the radio spectrum of CID-42.

2.1   Recoiling black holes

During galaxy major mergers, the central SMBHs that reside within the merging galaxies will form a bound binary SMBH system that can further merge (e.g. Volonteri et al. 2003, Hopkins et al. 2008, Colpi et al. 2009). At the time of the SMBH binary coalescence, strong gravitational wave (GW) radiation is emitted an-isotropically, depending on the spin and mass ratio of the two SMBHs, and to conserve linear momentum, the newly formed single SMBH recoils (Peres 1962, Bekenstein 1973). Recoiling SMBHs are the direct products of processes in the strong

field regime of gravity and they are one of the key observable signatures of a SMBH binary merger. As the SMBH recoils from the center of the galaxy, the closest regions (disk and broad line regions) are carried with it and the more distant regions are left behind depending on the recoil velocity (Merritt et al., 2006, Loeb, 2007). Because GW recoil displaces (or ejects) SMBHs from the centers of galaxies, these events have the potential to influence the observed co-evolution of SMBHs with their host galaxies, as demonstrated by numerical simulations (Blecha et al. 2011; Guedes et al. 2011; Sijacki et al. 2011).

Only few serendipitous discoveries of recoiling candidates have been reported in the literature (Komossa et al. 2008; Shields et al. 2009; Jonker et al. 2010; Batcheldor et al. 2010; Robinson et al. 2010; Steinhardt et al. 2012; Bianchi et al. 2013; Koss et al. 2014) and systematic observational searches have resulted in no candidates so far (Bonning et al. 2007; Eracleous et al. 2012; Komossa 2012).

The *Chandra*-COSMOS source CXOC J100043.1+020637 ($z = 0.359$, Elvis et al. 2009, Civano et al. 2012b), also known as CID-42, is a candidate for being a GW recoiling SMBH with both imaging (in optical and X-ray) and spectroscopic recoil signatures (Civano et al. 2010, 2012a,b). CID-42 shows two components separated by $\approx 0.5''$ (≈ 2.5 kpc; see below for cosmology details[1]) in the *Hubble Space Telescope* Advanced Camera for Surveys (HST/ACS) image and embedded in the same galaxy. As presented in Civano et al. (2010) and Civano et al. (2012a), the South-Eastern (SE) optical source has a point-like morphology typical of a bright AGN and it is responsible for the entire (>97%) X-ray emission of this system. The North-Western (NW) optical source has a more extended profile in the optical band with a scale length of ≈ 0.5 kpc, and the upper limit measured for its X-ray emission is consistent with being produced by star formation. In the optical spectra of CID-42 (obtained with the Very Large Telescope - VLT, Magellan, SDSS and Deep Imaging Multi-Object Spectrograph - DEIMOS; see Civano et al. 2012a,b), a velocity offset of $\approx 1\,300$ km s^{-1} is measured between the broad and narrow components of the Hβ line (Figs. 5 and 6 of Civano et al. 2010).

Despite diverse scenarios being proposed to explain the nature of CID-42 (e.g., Comerford et al. 2009, Civano et al. 2010), the upper limit on the X-ray luminosity combined with the analysis of the multiwavelength spectral energy distribution favors the GW recoil scenario, although the presence of a very obscured SMBH in the NW component cannot be fully ruled out. The current data are consistent with a recoiling SMBH ejected approximately $1 - 6$ Myr ago, as shown by detailed modeling presented in Blecha et al. (2013).

Throughout this chapter, a *WMAP* seven-year cosmology (Spergel et al., 2007; Larson et al.,

---

[1]Conversions between arcseconds and kiloparsecs are done according to Wright (2006) using the assumed cosmology.

2011) with $H_0 = 71 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$ (Hubble's constant), $\Omega_M = 0.27$ (matter density), and $\Omega_\Lambda = 0.73$ (dark energy density) is assumed.

2.2 Very Large Array 3 GHz data

2.2.1 Observations and reduction

We make use of the observations of the first 130 hours of the COSMOS field with the VLA in S-band (3 GHz, 10 cm). The 2 GHz bandwidth is comprised of 16×128 MHz spectral windows (SPWs). We observed the 2 sq. deg. field with 110 hours in A-configuration and 20 hours in C-configuration between 2012 and 2013. In order to cover the entire field, we required a total of 64 pointings. The quasar J1331+3030 was used for flux and bandpass calibration and J1024-0052 for gain and phase calibration. Weather conditions were excellent during the A-array observations while the C-array occasionally suffered from Summer thunderstorms.

Calibration was done using the Astronomical Image Processing System (AIPS[2]) based data reduction pipeline AIPSLite pipeline (e.g. Bourke et al. 2014) originally developed for the VLA Stripe 82 survey (Mooley et al., 2016). The calibrated data was then exported to the Common Astronomy Software Applications (CASA[3], version 4.2.1), where flagging (clipping in amplitude) was done separately for the A- and C-configuration data. All the epochs were concatenated into a single measurement set used for the imaging stage described below.

2.2.2 Imaging

To study the radio properties of CID-42 at 3 GHz, we imaged eight pointings (P19, P22, P30, P31, P36, P38, P39, P44) individually and joined them together in a mosaic. CID-42 is located near the center of pointing P36. Due to the wide bandwidth of the data, imaging can be done in various ways. One approach involves dividing the entire bandwidth into narrower SPWs which are then imaged separately and finally stacked together in the image plane (for details see Condon et al. 2012). This method gives accurate flux densities in each SPW because the flux density is approximately constant inside a sufficiently small bandwidth. Downsides include i) loss of resolution as the lowest frequency SPW determines the resolution for the final stack,

---

[2]http://www.aips.nrao.edu/index.shtml; see Wells 1985

[3]CASA is developed by an international consortium of scientists based at the NRAO, the European Southern Observatory (ESO), the National Astronomical Observatory of Japan (NAOJ), the Commonwealth Scientific and Industrial Research Organisation Australia Telescope National Facility (CSIRO/ATNF), and the Netherlands Institute for Radio Astronomy (ASTRON) under the guidance of NRAO. http://casa.nrao.edu; McMullin et al. 2007

and ii) uncleaned sources which do not have a sufficiently large signal-to-noise ratio inside a single SPW. We describe the application of this method in Sect. 2.2.3.

Another imaging approach is the multiscale multifrequency (MSMF) method developed by Rau & Cornwell (2011) which allows the usage of the entire bandwidth at once to calculate the monochromatic flux density at a chosen frequency. It uses Taylor term expansion in frequencies along with multiple spatial scales to deconvolve the map and includes a map of spectral indices calculated from the Taylor series. We define spectral index α as $F_\nu \propto \nu^\alpha$, where F_ν is flux density at frequency ν. The final resolution is not limited to the lowest frequency beam size because all of the SPWs are used in the deconvolution. For more details on this algorithm see Rau & Cornwell (2011) and Rau et al. (2014). We describe this approach in Sect. 2.2.4. We applied both imaging methods using the CASA task CLEAN.

2.2.3   Spectral windows stack

To correctly stack maps, it is necessary that they all have the same resolution which is defined by the lowest frequency SPW. We used the robust parameter in the Briggs weighting scheme to achieve similar resolutions in different SPWs. With higher robust value (toward natural weighting) it is generally expected to get lower *rms* noise and a larger synthesized beam, but worse sidelobes. However, sidelobe contamination started to drastically degrade the map quality after a certain robust value. The final robust values we converged on are listed in Table 2.1. They give the optimum tradeoff between resolution, noise root-mean-square (*rms*) and sidelobes. After each SPW is cleaned, it is convolved to a common resolution of $0.9''$ using the CASA toolkit function CONVOLVE2D and then corrected for the primary beam (PB) response with the CASA task IMPBCOR.

To optimize the signal-to-noise ratio (S/N) in the final stack, the SPW maps have been stacked weighting each pixel by the inverse of its local noise squared. Thus, each pixel in the sum is assigned a weight of $w = (A(\rho)/\sigma)^2$, where $A(\rho)$ is the primary beam response, ρ is the distance from the pointing center and σ is the *rms*. Note that the primary beam response entering this term only scales the *rms* (σ drawn from the cleaned map prior to primary beam correction) as a function of distance from the pointing center. Because the full width at half maximum (FWHM) of the primary beam (and therefore noise) changes between SPWs, not all frequencies will contribute the same amount to each pixel in the final stack. To account for this effect, a map of effective frequencies is also created by averaging the SPW frequencies using corresponding noise weights for each SPW with the CASA task IMMATH. In total we produced three stacks (at a resolution of $0.9''$): one for the full bandwidth using all of the SPWs (SPWs 0 to 15) reaching an *rms* of 4.8 μJy beam^{-1}, one for the low frequency sideband (SPWs 0 to 8)

Table 2.1: Summary of robust parameters, flagged data and final *rms* values for each SPW. The *rms* noise is calculated after convolution to a common beam size but before primary beam correction. Last column shows which array data was entirely flagged due to corrupted data.

SPW	ν (GHz)	Robust	*rms* (μJy beam^{-1})	Flagged
0	2.05	0.0	25.6	-
1	-	-	-	A+C
2	-	-	-	A+C
3	2.44	0.0	20.2	-
4	2.56	0.5	16.5	-
5	2.69	0.5	17.4	-
6	2.82	0.5	17.1	-
7	2.95	0.5	16.0	-
8	3.05	0.5	16.0	-
9	3.18	0.5	14.4	-
10	3.31	1.0	13.0	-
11	3.44	1.0	12.8	-
12	3.56	1.0	12.8	-
13	3.69	1.5	19.5	C
14	3.82	1.5	30.9	C
15	3.95	1.5	32.0	C

reaching an *rms* of 6.9 μJy beam^{-1}, and one for the high frequency sideband (SPWs 9 to 15) reaching an *rms* of 6.5 μJy beam^{-1}. In Fig. 2.1 we show cutouts centered on CID-42 in each map we created. Resolution, *rms* and effective frequency of these maps are listed in Table 2.2.

2.2.4 Multiscale multifrequency

Imaging was independently performed with MSMF using the CASA task CLEAN that uses the entire 2 GHz bandwidth at once (Rau & Cornwell, 2011). Each pointing was cleaned individually. Two Taylor terms were used in the frequency expansion (NTERMS=2) along with three resolution scales. The final synthesised beam size was $0.7'' \times 0.6''$. We used Briggs weighting with a robust value of 0.5 to produce a map with low sidelobe contamination and good *rms*. After the deconvolution, wideband primary beam correction was applied using the CASA task WIDEBANDPBCOR. The resulting maps include flux densities at a reference frequency of 3 GHz
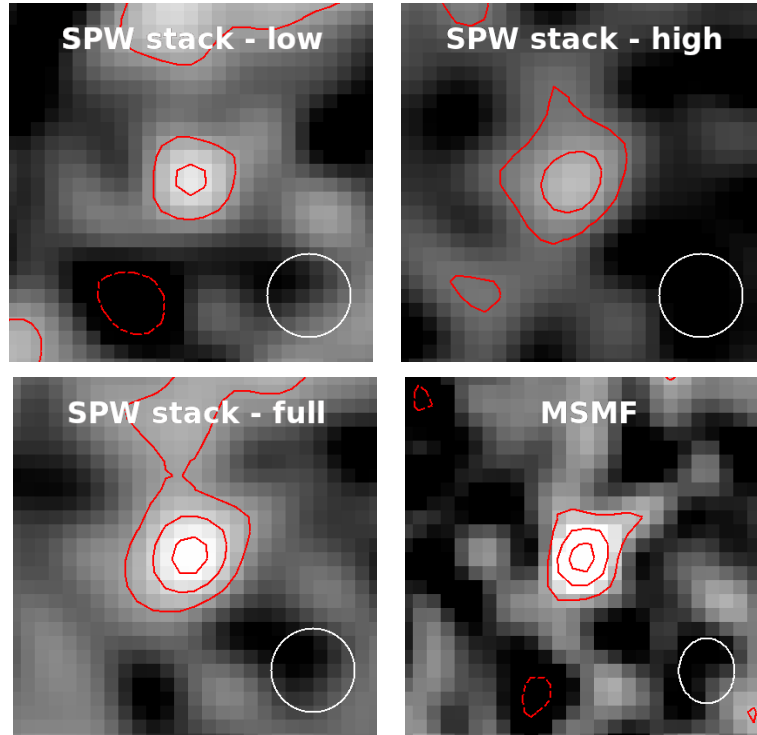
Figure 2.1: Cutouts of $4''$ on the side centered on CID-42: low frequency SPW stack, high frequency SPW stack, full SPW stack and MSMF map. All SPW stacks have the same resolution of $0.9''$, while the MSMF map has a resolution of $0.7''$. The synthesized beam size is shown in the lower right corner of each panel. Contours of $\pm2\sigma$, $\pm4\sigma$ and $\pm6\sigma$ are overplotted in red (negative contours are drawn with dashed lines). The local *rms* in each map is listed in Table 2.2.

along with the spectral index map (α map) all corrected for the PB response. The *rms* of the MSMF map is $4.6\,\mu\mathrm{Jy\,beam}^{-1}$ and a MFMS map stamp centered at CID-42 is shown in Fig. 2.1.

2.2.5 CID-42: Flux and size at 3 GHz

To properly extract the flux density of CID-42, information about its size is needed. A source is considered resolved if it is larger than the beam size, however, this is a function of signal-to-noise ratio as described in Bondi et al. (2003). Thus, to estimate whether CID-42 is resolved, we need to know how the integrated-to-peak flux density ratio $S_{\mathrm{int}}/S_{\mathrm{peak}}$ behaves as a function of S/N in our mosaic. For this purpose we use the AIPS task SAD (search and destroy) to extract a catalog of sources, their positions and flux densities over the mosaic. The catalog extraction was performed in the same way as for the other VLA-COSMOS surveys, described in detail in Schinnerer et al. (2007, 2010) and Smolčić et al. (2014). We have compared the peak surface brightness with the integrated flux density of CID-42, as well as its peak-to-integrated flux ratio

relative to those of other sources in the mosaic as a function of S/N. From this analysis we conclude that CID-42 is unresolved at a resolution of $0.9''$ and that it is marginally resolved at $0.7''$. We report the peak surface brightness of CID-42, drawn from the SPW stack map to be $33.2 \pm 4.8\,\mu$Jy and from the MSMF map $32.5 \pm 4.6\,\mu$Jy. Since CID-42 is only marginally resolved in the MSMF map, we report its integrated flux density of $39.7 \pm 9.5\,\mu$Jy, but we do not use this value in the spectral analysis. The error of the peak surface brightness is determined from the local *rms* obtained via fitting the histogram of the pixel flux density values with a Gaussian, and the error for the integrated flux density is drawn from the elliptical Gaussian fit.

As the resolution of our 3 GHz data is around a factor of 2 higher than that of the large VLA-COSMOS 1.4 GHz data (Schinnerer et al. 2007), to test whether we may be out-resolving a fraction of the flux density at 3 GHz we used two methods to match the resolution of the 3 GHz map to the 1.4 GHz map. Firstly, we convolved the 3 GHz map to a resolution of $1.5''$, matching that of the 1.4 GHz map. Peak surface brightness of CID-42 in this map is $36.8 \pm 7.3\,\mu$Jy beam^{-1}. Secondly, in the cleaning process, we weighted the visibilities with the Gaussian taper in the *uv*-plane. This approach gave a peak surface brightness of $34.6 \pm 6.3\,\mu$Jy beam^{-1}. Both methods result in flux densities consistent with the higher resolution map within the error bars.

2.3 Radio data analysis

2.3.1 Image analysis

Table 2.2 lists all the radio data used for the CID-42 analysis presented here, their resolutions and *rms* in the corresponding maps. In the top panel of Fig. 2.2 we show the HST/ACS image (Koekemoer et al., 2007) of CID-42 with contours from the 1.4 GHz (black) and 3 GHz (magenta and blue) radio data overlaid. The resolution of $1.5''$ at 1.4 GHz (Schinnerer et al. 2007) is not accurate enough to distinguish whether the radio emission is associated with the NW or the SE optical source (see also Civano et al. 2010). However, the 3 GHz data presented here at a resolution of $0.7''$ show that the 6σ contours are coincident with the SE component. The peak of the 3 GHz radio emission is located at $\alpha_{2000.0} = 10^{\text{h}}00^{\text{m}}43{.}^{\text{s}}148, \delta_{2000.0} = +2°06'37{.}''06$. The offset between the center of the SE component in the HST image and the center of the VLA 3 GHz emission is $0.08''$. This value is within the positional error of $0.1''$ estimated from the ratio of resolution and S/N which dominates the positional uncertainty of faint sources. We compared the positions of 30 sources measured with the Very Long Baseline Array (VLBA) by Herrera Ruiz et al. (2017) with our 3 GHz mosaic and found that positional errors in our

Table 2.2: Radio data used for the CID-42 flux density analysis along with the corresponding resolution and local *rms* noise inside the maps. CID-42 is resolved in the 1.4 GHz VLA map and marginally resolved in the 3 GHz MSMF map where integrated flux density is also reported in parentheses, for all other maps peak surface brightness is reported.

ν (GHz)	Instrument	Resolution	*rms* (μJy/b)	Flux density (μJy/b)	Author
0.32	VLA	$6''$	442	< 1326	Smolčić et al. (2014)
0.325	GMRT	$10.8''$	80	240 ± 80	Tisanić et al. (in prep.)
1.4	VLA	$1.5''$	10	85 ± 10	Schinnerer et al. (2007)
				$(138 \pm 38 \mu$Jy)	
1.5	VLBA	$0.015''$	13	< 78	Wrobel et al. (2014)
2.7	VLA (SPWs - low)	$0.9''$	6.9	31.2 ± 6.9	This work
3.4	VLA (SPWs - high)	$0.9''$	6.5	35.4 ± 6.5	This work
3.1	VLA (SPWs - full)	$0.9''$	4.8	33.2 ± 4.8	This work
3	VLA (MSMF)	$0.7''$	4.6	32.5 ± 4.6	This work
				$(39.7 \pm 9.5 \mu$Jy)	
9	VLA	$0.32''$	6	< 18	Wrobel et al. (2014)

mosaic are less than $0.03''$. This constrains the positional error of bright sources (S/N between 40 and 900) inside our mosaic. The low S/N sources positional error is therefore dominated by the mentioned term.

In the bottom panel of Fig. 2.2 we show that radio emission at 3 GHz is coincident with the X-ray emission (Civano et al. 2012b). As described in detail in Sect. 2.2.5, CID-42 is unresolved at a resolution of $0.9''$ meaning that the FWHM 3 GHz radio emission is constrained to a region < 4.5 kpc (at the source redshift) centered at the SE optical component. Since CID-42 is marginally resolved in the MSMF map at a resolution of $0.7''$ we calculated its deconvolved size. The deconvolved FWHM major axis is $0.6''$ and minor axis $0.1''$ (corresponding to 3 kpc and 0.5 kpc respectively) with a position angle of $-36°$ measured from North through East which could indicate an elongated shape of the source. However, better resolution and deeper data are needed to confirm this.

2.3.2 Spectral analysis

Fig. 2.3 shows spectral features of CID-42 across a frequency range of 320 MHz to 9 GHz. Schinnerer et al. (2007) observed CID-42 at 1.4 GHz and $1.5''$ resolution within the VLA-COSMOS survey. They find the source resolved at that resolution and report an integrated flux density of $138 \pm 38\ \mu$Jy. Wrobel et al. (2014) found no detection using a high resolution of
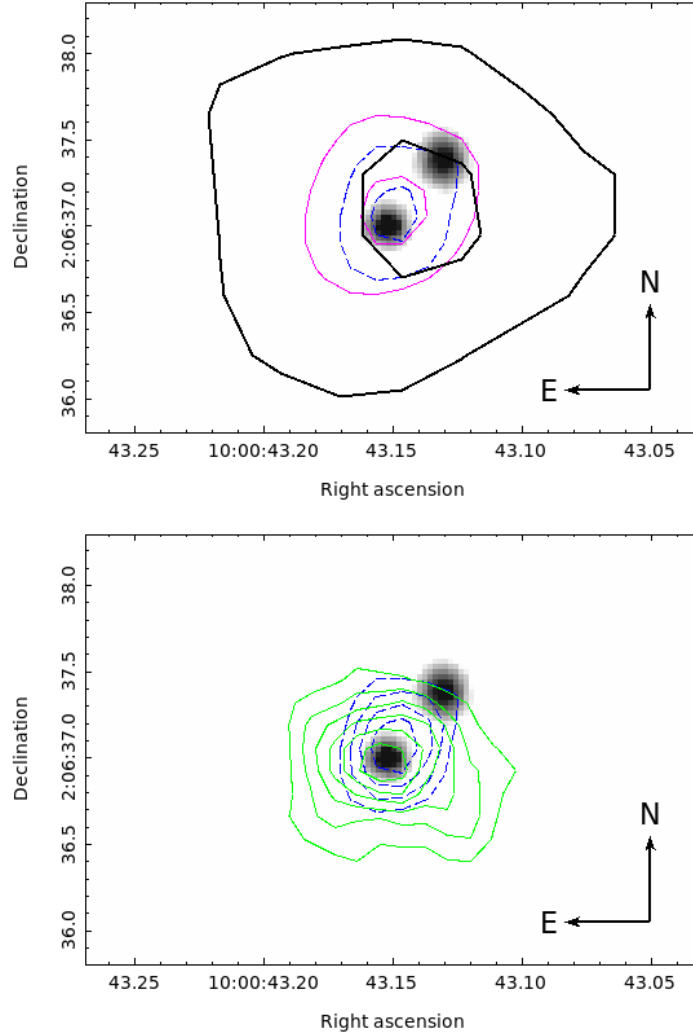
Figure 2.2: *Top panel*: HST/ACS gray-scale image showing the two optical components of CID-42 with contours from the radio data overlaid. For each radio map, two contour lines are shown that correspond to 3σ and 6σ. Thick black lines show data from the VLA 1.4 GHz large map from Schinnerer et al. (2007) with $\sigma = 10\,\mu\mathrm{Jy\,beam^{-1}}$. Thin magenta lines and dashed blue lines are from the VLA 3 GHz SPW stack map and MSMF map with the local *rms* of $4.8\,\mu\mathrm{Jy\,beam^{-1}}$ and $4.6\,\mu\mathrm{Jy\,beam^{-1}}$, respectively. *Bottom panel*: HST/ACS gray-scale image of CID-42 with contours from the adaptively smoothed X-ray *Chandra* High Resolution Camera image with a 3 pixel radius Gaussian kernel by Civano et al. (2012b) in green and VLA 3 GHz MSMF map with 1σ steps starting from the 3σ in blue.
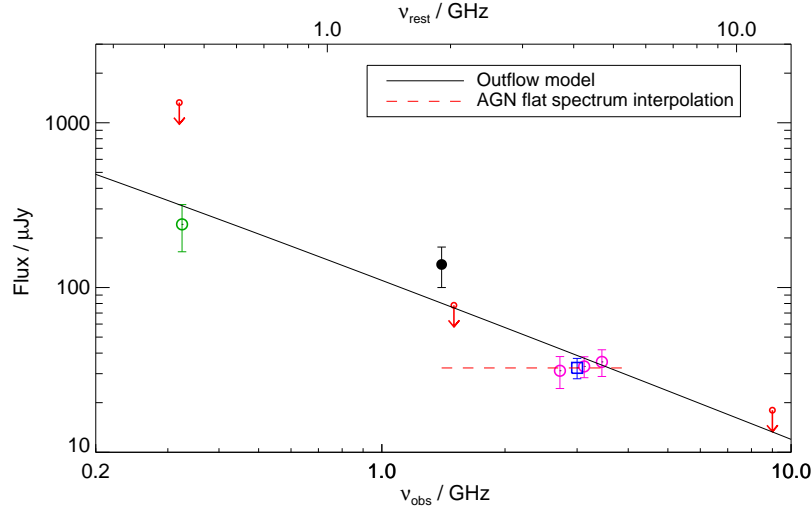
Figure 2.3: Radio spectrum of CID-42. Values are listed in Table 2.2. The black point marks the total integrated flux density from the 1.4 GHz VLA data (Schinnerer et al., 2007). Magenta points are peak surface brightnesses from SPWs stacks from our new VLA 3 GHz data. The flux density from the MSMF cleaned map is shown with a blue square. The green point is the peak surface brightness from the GMRT map (Tisanić et al., in prep). From left to right the red arrows show the upper 3σ limit from the 320 MHz VLA (Smolčić et al. 2014), the 3σ limit from the 9 GHz VLA data, and the 6σ limit from the 1.5GHz VLBA data (Wrobel et al. 2014). The flat spectrum interpolation for AGN is indicated with red dashed line. The black line corresponds to the radio spectrum due to shock generated by an outflow model (see text for details).

0.015″ VLBA at 1.5 GHz and report a 6σ upper limit of $78\,\mu$Jy beam^{-1}. Wrobel et al. (2014) also did not detect the source at the higher frequency of 9 GHz using the VLA with a 3σ upper limit of $18\,\mu$Jy beam^{-1}. In Fig. 2.3 we further plot the 325 MHz flux density of 240 ± 80 μJy beam^{-1} from the Giant Metrewave Radio Telescope (GMRT) data (Tisanić et al. in prep.). The GMRT gives a tentative 3σ detection for CID-42, but since we know the source location and we expect that it is unresolved at a resolution of $\approx 10''$ this may justify a peak surface brightness measurement. For this reason the spectrum shown in Fig. 2.3 is not strongly constrained by the 325 MHz measurement.

In the 2 GHz bandwidth centered at 3 GHz, the galaxy shows a slightly inverted spectra. This behaviour is evident in the flux densities extracted from the SPW stacks (see Fig. 2.3) and also from the spectral index calculated in the MSMF alpha map where $\alpha = 0.4 \pm 0.3$. For the sake of simplicity, we will treat this as a flat spectrum ($\alpha = 0$). Using only the MSMF value, the flux density of CID-42 at 3 GHz corresponds to rest frame spectral luminosity of $L_{3\,\mathrm{GHz}} \approx 1 \times 10^{22}$ W Hz^{-1}. Based on radio-loud vs. radio-quiet definitions in literature relying only on radio luminosity values (e.g. Miller et al. 1990, see also Baloković et al. 2012), this

falls into the radio-quiet regime.

To test the consistency between the two different wideband imaging methods and the robustness of our 3 GHz data in general and thus that for CID-42, we have closely examined the spectral behavior of 109 sources in the mosaic in different locations within the primary beam. In Fig. 2.4, we show eight representative sources over a broad S/N range (from 230 to 5). In the spectral plots we also show our 3 GHz flux densities with the 1.4 GHz Large Project catalog values (Schinnerer et al., 2007) and the 325 MHz GMRT data (Tisanić et al. in prep.). We find that both imaging methods (SPW stack and MSMF) give consistent flux densities, also consistent with the flux densities at lower frequencies. Thus, we conclude that our map is robust with no systematics for sources at either high or low S/N.

2.3.3 Discussion

The unresolved emission at $0.9''$ and flat spectrum we derived here for CID-42 at 3 GHz (see Fig. 2.3) suggests that the entire observed radio emission at 3 GHz (with 2 GHz bandwidth) may arise from an AGN within the SE component. The VLBA upper limit at 1.5 GHz by Wrobel et al. (2014) is consistent with the flat spectrum AGN picture, as the VLBA can observe scales up to 150 mas (750 pc) but it is not sensitive to star formation induced emission. The non detection at 9 GHz (Wrobel et al., 2014) could be explained by steepening of the spectrum due to synchrotron losses at higher frequencies. The 1.4 GHz detection in the VLA-COSMOS survey (Schinnerer et al., 2007) is resolved at a coarse resolution of $1.5''$ and exhibits significantly higher flux density than the 3 GHz and the VLBA flux densities. This may suggest that the 1.4 GHz flux density arises from a region more extended than $0.9''$ (4.5 kpc) observed at 3 GHz.

Assuming, as argued above, that the 3 GHz emission is dominated by a flat spectrum AGN, by extrapolating this flux density to 1.4 GHz and comparing it with the observed one at that frequency we find an excess of $\approx 105\,\mu$Jy (see Fig. 2.3). As tested in Sect. 2.2.5 we found no evidence that this may be a resolution effect. Thus, if this excess were due to star formation than it would need to have a spectral slope of $\alpha < -2$ to account for the observed 3 GHz emission. Such a steep slope is highly untypical for star formation as shown by e.g. Kimball & Ivezić (2008) who find that typical spectral slopes for star-forming galaxies are between -0.5 and -1.5 based on SDSS - NVSS[4] - FIRST[5] data. The steep spectrum may possibly reflect aging of the electrons by synchrotron radiation with a characteristic break frequency between 1.4 and 3 GHz (see van der Laan & Perola 1969 and Miley 1980) or could occur in radio jets. We have further tried to model the full observed spectrum shown in Fig. 2.3 assuming bow

---

[4]NRAO VLA Sky Survey (NVSS)
[5]Faint Images of the Radio Sky at Twenty-Centimeters (FIRST)
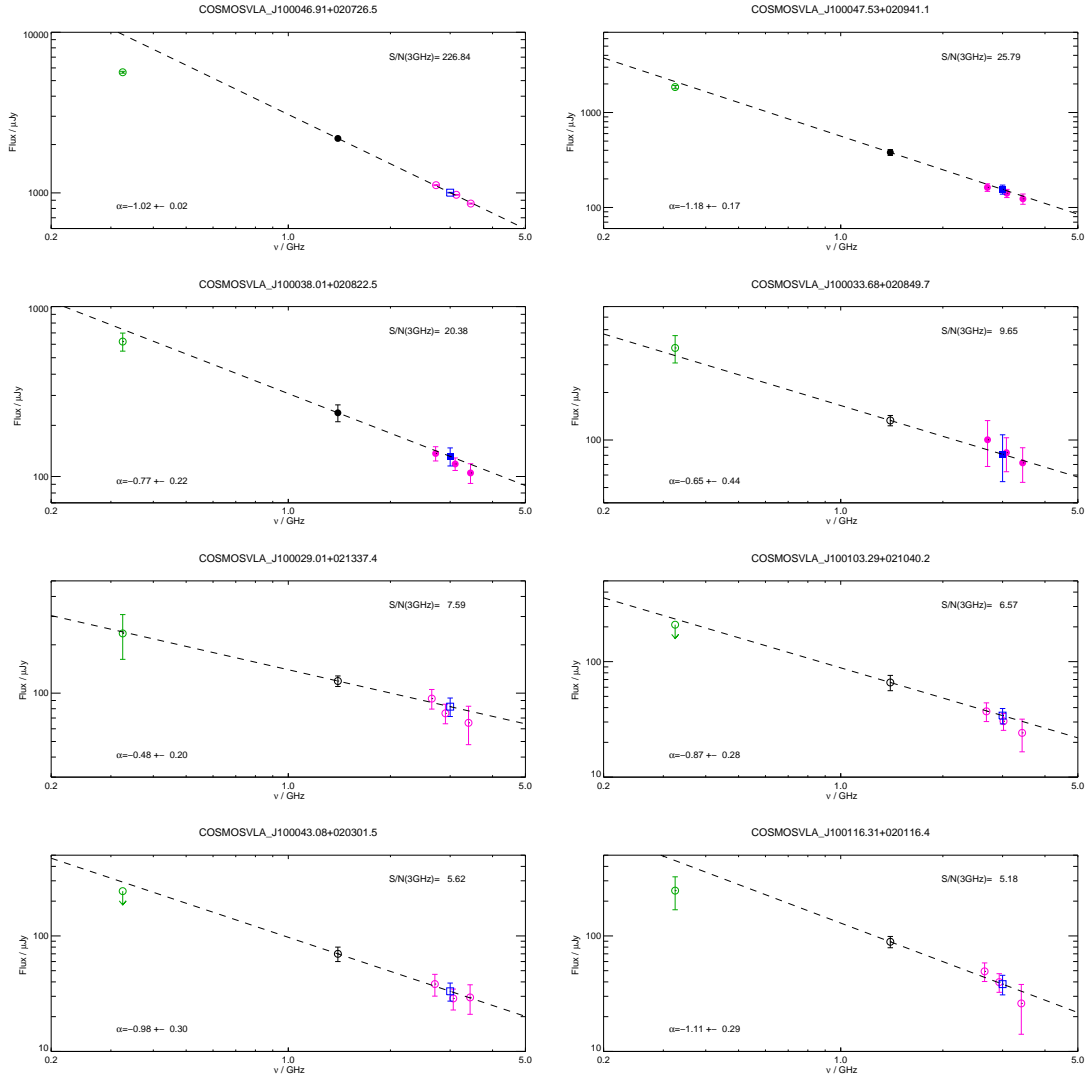
Figure 2.4: Flux densities for selected sources as a function of frequency inside the VLA-COSMOS pointing containing CID-42 at different S/N ratios. Filled symbols are used for integrated flux densities of resolved sources and empty symbols for peak surface brightness of unresolved sources. Black points mark the flux density from the 1.4 GHz large project catalogue (Schinnerer et al., 2007). Magenta points are flux densities from the low frequency, whole bandwidth, and high frequency SPW stack from the VLA 3 GHz data. Green points are peak surface brightnesses from the GMRT map (Tisanić et al., in prep). The flux density from the MSMF cleaned map is shown with the blue square. The black dashed line correspond to spectral index derived from the 1.4 GHz and 3 GHz point. Error bars represent the local *rms* for peak surface brightnesses and the fitting error for integrated flux densities.

shock and mass outflow models generated by an outflow from the SMBH (for details see Loeb 2007 and Wang & Loeb 2014). Both models are able to fit the data with a resulting spectral slope of $\alpha \approx -1$ in the range from 320 MHz to 9 GHz, yet the fits still cannot account for the excess at 1.4 GHz rendering the physical interpretation difficult (the outflow model is plotted in Fig. 2.3). It is possible that CID-42 is a variable source in radio. We can get an insight into the source variability by comparing the various radio COSMOS data. We find that CID-42 is not showing flux density variations in the 1.4 GHz VLA-COSMOS data taken between the years 2003 and 2006 (Schinnerer et al. 2004, 2007, 2010). If variability is the reason for the 1.4 GHz radio excess of CID-42, it would have had to occur between years 2006 and 2013 or on shorter timescales (day/month). To test whether CID-42 is variable on monthly scales we concatenated our 3 GHz A-array observations into two epochs lasting approximately 20 days each and imaged them separately using the MSMF algorithm. The peak flux density extracted from the first epoch (28 November 2012 to 19 December 2012) is $34.6 \pm 6.7 \, \mu$Jy and from the second epoch (20 December 2012 to 7 January 2013) is $24.8 \pm 6.3 \, \mu$Jy (errors represent local *rms* in the maps). Although the flux densities agree withing 2σ, their difference between these two time periods could indicate that CID-42 is variable on monthly time scales. Further observations are needed to confirm this.

In summary, the SE component of CID-42 appears to be a flat spectrum radio-quiet AGN with associated extended emission, the source of which is still unclear. On the other hand, the NW component is not detected in our 3 GHz data. If the NW component is an obscured AGN, as already suggested by in Civano et al. (2010) and Wrobel et al. (2014), then it is also radio-quiet with a 3σ upper limit rest frame spectral luminosity at 3 GHz of $L_{3\,\mathrm{GHz}} < 5.6 \times 10^{21} \ \mathrm{W\,Hz^{-1}}$ (assuming a spectral index of -0.8). The results presented here are not inconsistent with the recoiling black hole scenario, but still cannot rule out the presence of an obscured radio quiet SMBH in the NW component. Further spectroscopic observations would be able to spatially resolve the two nuclei in the optical spectrum.

2.4 Chapter summary

We used the first 130 hours of VLA-COSMOS 3 GHz data to analyze the radio synchrotron spectrum of CID-42, the best candidate recoiling black hole in the COSMOS field. Due to the large 2 GHz bandwidth, we imaged the data with two different methods: SPW stacking and MSMF. Both of these methods gave consistent flux densities for CID-42. Our 7σ detection shows that all of the 3 GHz radio emission is arising from the SE component of CID-42 and our new 3 GHz radio data confirm that the SE component is an unobscured radio-quiet AGN. These

data combined with other radio data from the literature (Schinnerer et al. 2007, Wrobel et al. 2014) suggest that the radio emission is composed of a flat spectrum AGN core and perhaps a more extended region of an aged electron synchrotron population or shocks generated by an outflow from the black hole. Only an upper limit of radio emission can be given for the NW component. There are indications that CID-42 could be variable on monthly scales based on the 2σ flux density variation between two different time epochs but further observations are needed to confirm this.

Chapter 3

VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release

This chapter explains all the steps performed with the full ~ 400 hours of VLA observations: from telescope setup to end products including a fully tested 3 GHz radio catalog and bias corrected number counts. A detailed technical description is necessary, because this radio catalog is the starting point for all subsequent analyses. Understanding biases, incompleteness, and systematics is crucial for a statistical study of large samples.

3.1 Radio surveys

One of the main quests in modern cosmology is understanding the formation of galaxies and their evolution through cosmic time. In the past decade it has been demonstrated that a panchromatic, X-ray to radio, observational approach is key to develop a consensus on galaxy formation and evolution (e.g., Dickinson et al. 2003; Scoville et al. 2007; Driver et al. 2009, 2011; Koekemoer et al. 2011; Grogin et al. 2011). In this context, the radio regime offers an indispensable window toward star formation and supermassive black hole properties of galaxies as radio continuum emission i) provides a dust-unbiased star formation tracer at high angular resolution (e.g., Condon 1992; Haarsma et al. 2000; Seymour et al. 2008; Smolčić et al. 2009a; Karim et al. 2011), and ii) directly probes those AGN that are hosted by the most massive quiescent galaxies and deemed crucial for massive galaxy formation (e.g., Croton et al. 2006; Bower et al. 2006; Best et al. 2006; Evans et al. 2006; Hardcastle et al. 2007; Smolčić et al. 2009b; Smolčić 2009; Smolčić & Riechers 2011; Smolčić et al. 2015).

In recent decades, radio interferometers, such as the VLA, GMRT and Australia Telescope Compact Array (ATCA), have surveyed fields of different sizes (ranging from tens of square

arcminutes to thousands of square degrees), depths (microjansky to Jansky), and multiwavelength coverage (e.g., Becker et al. 1995; Condon et al. 1998; Ciliegi et al. 1999; Georgakakis et al. 1999; Bock et al. 1999; Prandoni et al. 2001; Condon et al. 2003; Hopkins et al. 2003; Schinnerer et al. 2004; Bondi et al. 2003, 2007; Norris et al. 2005; Schinnerer et al. 2007, 2010; Afonso et al. 2005; Tasse et al. 2007; Smolčić et al. 2008; Owen & Morrison 2008; Miller et al. 2008, 2013; Owen et al. 2009; Condon et al. 2012; Smolčić et al. 2014; Hales et al. 2014). These past surveys have shown that deep observations at high angular resolution ($\lesssim 1''$) with exquisite panchromatic coverage are critical to comprehensively study the radio properties of the main galaxy populations, avoiding cosmic variance with large area coverage (e.g., Padovani et al. 2009; Padovani 2011; Smolčić et al. 2008, 2009a,b; Smolčić 2009; Smolčić & Riechers 2011; Seymour et al. 2008; Bonzini et al. 2012, 2013). In this context, large area surveys down to unprecedented depths are planned with new and upgraded facilities (e.g., VLA, Westerbork, ASKAP, MeerKAT, and SKA; e.g., Jarvis 2012; Norris et al. 2011, 2013, 2015; Prandoni & Seymour 2015). Fig. 3.1 shows the 1σ sensitivity of each survey as a function of the area covered for past, current, and future radio continuum surveys. The VLA-COSMOS 3 GHz Large Project bridges the gap between past and future radio continuum surveys by covering an area as large as two square degrees down to a sensitivity reached to date only for single pointing observations. This allows for individual detections of $> 10\,000$ radio sources, further building on the already extensive radio coverage of the COSMOS field at 1.4 GHz VLA (VLA-COSMOS Large, Deep and Joint projects; Schinnerer et al. 2004, 2007, 2010), 320 MHz VLA (Smolčić et al., 2014), 325 MHz and 616 MHz GMRT data (Tisanić et al., in prep.), 6 GHz VLA (principal investigator - PI: Steven Myers), and the deep multiwavelength X-ray to mm photometry (Scoville et al. 2007; Koekemoer et al. 2007; Hasinger et al. 2007; Capak et al. 2007; Sanders et al. 2007; Bertoldi et al. 2007; Elvis et al. 2009; Ilbert et al. 2013; McCracken et al. 2012; Scott et al. 2008; Aretxaga et al. 2011; Smolčić et al. 2012; Miettinen et al. 2015; Civano et al. 2016; Laigle et al. 2016, Capak et al., in prep.) and more than 97 000 optical spectroscopic redshifts (Salvato et al., in prep.; zCOSMOS, Lilly et al. 2007, 2009; Trump et al. 2007; Prescott et al. 2006; Le Fèvre et al. 2015; Aihara et al. 2011; Nagao et al., priv. comm.). This further makes the survey part of one of the richest multiwavelength data sets currently available.

Radio continuum surveys at 3 GHz with the upgraded VLA are still sparse in the literature. Condon et al. (2012) performed single-pointing observations targeting the Lockman hole for 50-hours on-source with the VLA in C-array configuration. The observations resulted in a confusion-limited map with an *rms* of 1 μJy beam^{-1}. Based on this they constrained the counts of discrete sources in the $1-10\,\mu$Jy range via a $P(D)$ analysis. A more complex $P(D)$ analysis using the same data was applied by Vernstrom et al. (2014) who probed the counts down to
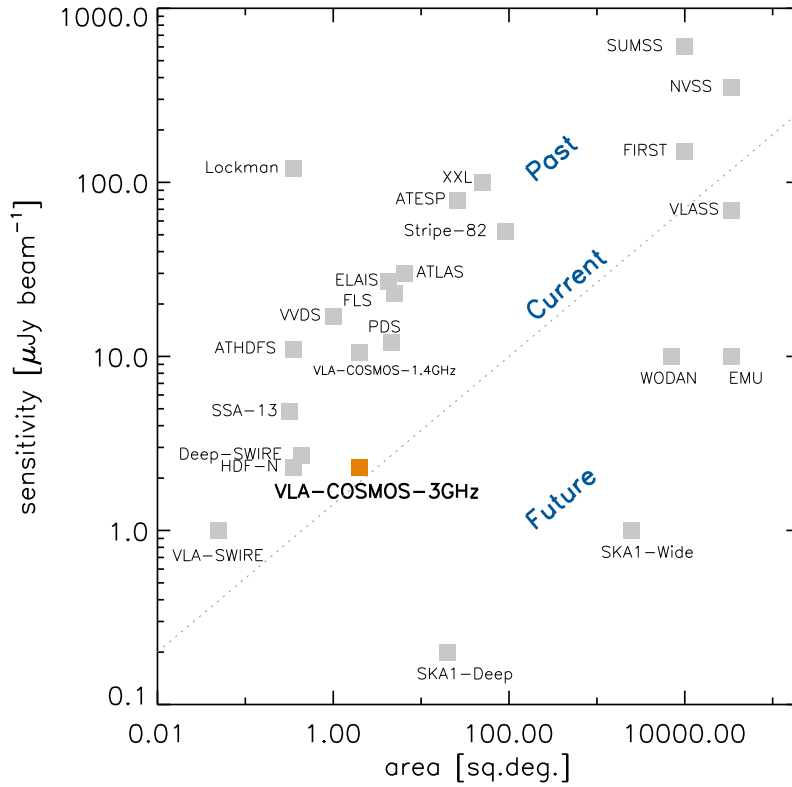
Figure 3.1: Sensitivity (at the observed frequency of the given survey) versus area for past, current, and future radio continuum surveys.

0.1 μJy. Both results are qualitatively in agreement with the already well-known flattening of the radio source counts (normalized to the $N(S) \propto S^{-3/2}$ of a static Euclidian space) below flux densities of $S_{1.4 \text{ GHz}} \approx 1$ mJy, and a further decrease of the counts with decreasing flux density below $S_{1.4 \text{ GHz}} \approx 60$ μJy. Such a shape of radio source counts is expected owing to the cosmic evolution of galaxy populations (e.g., Hopkins et al. 2000; Wilman et al. 2008; Béthermin et al. 2012), but this shape is contrary to that obtained based on i) the previous Lockman hole observations at 1.4 GHz (Owen & Morrison, 2008), and ii) a comparison of the sky brightness temperature measured by the Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission (ARCADE) 2 experiment (Fixsen et al., 2009) with that derived from the integral of the observed radio source counts (Vernstrom et al., 2011). The latter results instead point to a rise of the counts with decreasing flux density at these levels. To investigate this further, we here derive the radio source counts using our VLA-COSMOS 3 GHz Large project data, yielding the deepest radio counts derived to date based on direct source detections.

3.2 Observations and data reduction

3.2.1 Observations

A total of 384 hours of observations toward the COSMOS field were taken in S band using the *S 3s full width* setup covering a bandwidth of 2048 MHz centered at 3 GHz, and separated into 16 128 MHz-wide SPWs, with full polarization, and a 3s signal-averaging time. The observations were taken from November 2012 to January 2013, June to August 2013, and February to May 2014 in A-array (324 hours) and C-array configurations (60 hours; Legacy ID AS1163). Sixty-four pointings, separated by $10'$ in right ascension (RA) and declination (Dec), corresponding to two-thirds of the half power beam width (HPBW) at the central frequency of 3 GHz, were chosen to cover the full two square degree COSMOS field. Three sets of 64 pointings in such a grid were used to achieve a uniform *rms* over the field; this resulted in a total of 192 pointings (shown in Fig. 3.2). The first set of pointings is nominal, the second is shifted by $5'$ in RA and Dec, while the third set is shifted by $-5'$ in RA. Observing runs of 5 and 3 hours in length were conducted. In each observing run J1331+3030 was observed for flux and bandpass calibration for about $3 - 5$ minutes on-source (J0521+166 was used only for the first day of observations) at the end of every run, J1024-0052 was observed every 30 minutes for 1m 40s on-source for gain and phase calibration, while the source J0713+4349 was observed for 5 minutes on-source at the beginning of each run for polarization leakage calibration. During the five hour observing runs each pointing was visited twice, while the order of the pointing coverage blocks during the fixed 5-hour observing blocks was changed between the different observing runs to optimize the *uv* coverage. During the 3-hour observing blocks each pointing was visited once, and a good *uv* coverage was assured via dynamic scheduling. Typically, 26 antennas were used during each observing run. The A-array configuration observations were mostly conducted under good to excellent weather conditions. The C-array configuration observations were partially affected by poor weather conditions (Summer thunderstorms), yielding on some days up to 30% higher *rms* than expected based on the VLA exposure calculator.

3.2.2 Calibration

Calibration of the data was performed via the AIPSLite pipeline (Bourke et al., 2014) developed for the Stripe 82 Survey (Mooley et al., 2016). This pipeline was adapted for the VLA-COSMOS 3 GHz Large Project (as described below) and it follows, in general, the procedures outlined in Chapter E of the AIPS Cookbook[1].

---

[1]http://www.aips.nrao.edu/cook.html
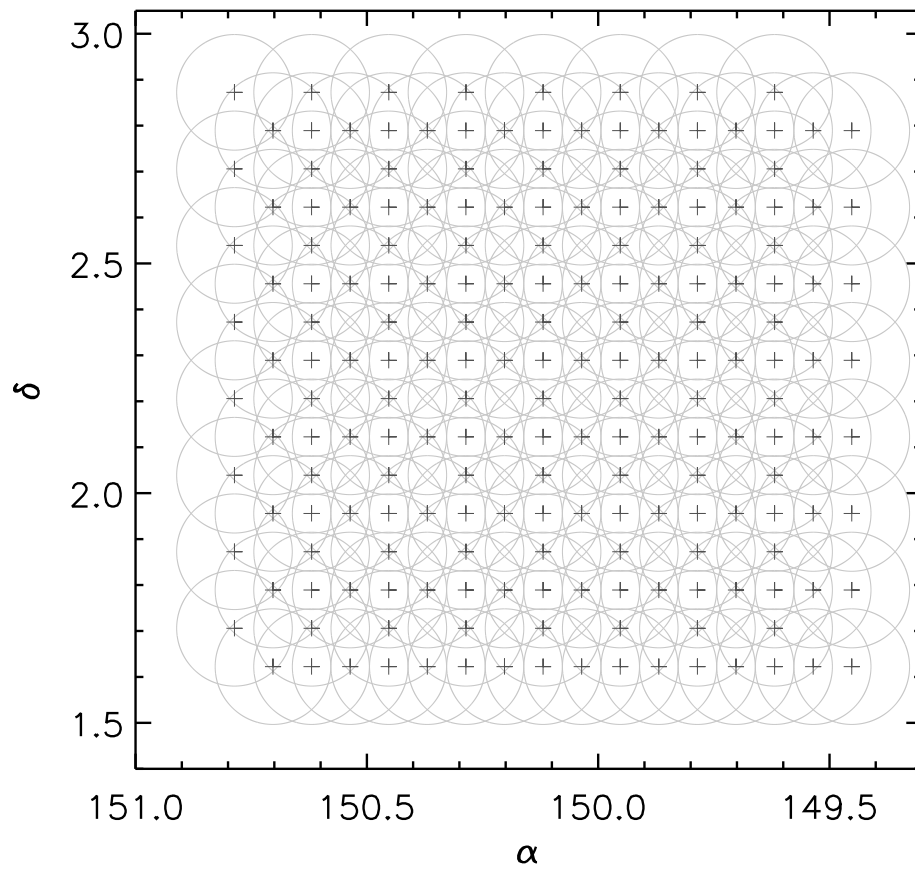
Figure 3.2: Pointing pattern used for the 3 GHz VLA-COSMOS Large Project. The centers of the 192 pointings are indicated by the plus signs. Circles indicate the primary beam of each pointing, represented here by the HPBW at 3 GHz ($15'$; the primary beam HPBW is a function of frequency and varies by a factor 2 between the lower and upper edge of the S band).

In brief, the data are first loaded with the Obit[2] task BDFIN. Band edges, and to a larger extent baseband edges, were then flagged with the task UVFLG. SPWs 2 and 3, found to be irreparably corrupted by radio frequency interference (RFI) in all observations (see Fig. 3.3), were entirely flagged using the task UVFLG. After flagging, FRING, BPASS, SETJY, CALIB, GETJY, and CLCAL tasks were used to derive the delay, bandpass, and complex gain solutions. Polarization calibration was performed using the tasks RLDLY, PCAL, and RLDIF as detailed in Sect. 7 of Chapter E in the AIPS Cookbook. The task RFLAG was used to flag all target pointings and the flags were applied using the UVCOP task. The derived calibration was applied and the calibrated data set was produced with the SPLAT task. Finally, the calibrated *uv* data were saved to disk using the task FITTP. During the pipeline process several diagnostic plots were generated to assess the quality of the calibration: bandpass solutions, antenna gains as a function of time, calibrated spectrum of the gain calibrator, and calibrated amplitude versus phase plots of the gain calibrator per pointing. In Table 3.1 we list the statistics for the amplitude of the phase calibrator in each SPW for all observing blocks. The average amplitude scatter around the mean is typically $2 - 3\%$, with the exception of the highest frequency SPWs, for which it is higher than 10%. The combined typical scatter around the mean is $\sim 5\%$. This assures a good flux calibration.

The highest frequency SPWs marked 14, 15, and 16 have low amplitude RFI and the phases are significantly affected for some observations. The C-array configuration data at the upper end of the S band are mostly unusable due to this RFI. These data have been manually flagged, and we additionally ran the RFLAG task on the rest of the C-array configuration data to further remove bad data and extend flags in frequency and time. The A-array configuration data for these SPWs are generally good. Our imaging tests show that the data from these SPWs generally improves the sensitivity, but limits the dynamic range for certain pointings. Looking at the overall imaging performance, we decided to retain these SPWs. Despite the data drop-outs, the median flux density values of the phase calibrator (J1024-0052; Table 3.1) are consistent with the spectral parameters inferred from the other SPWs. Through our tests we find that, in the majority of observations, RFI adversely affects the system temperature measurements, and hence we have left out the correction for the system temperature from the calibration process.

At this point the pipeline diverges in two directions to: i) image the target fields and ii) produce and export a calibrated data set in preparation for mosaicking. To image the target fields, they were split out with calibration applied (using the task SPLIT). The fields were then further auto-flagged (using the task RFLAG), imaged (using the task IMAGR), and exported (using the task FITTP) in parallel. The calibrated data set was generated by applying RFLAG and

---

[2]http://www.cv.nrao.edu/~bcotton/Obit.html
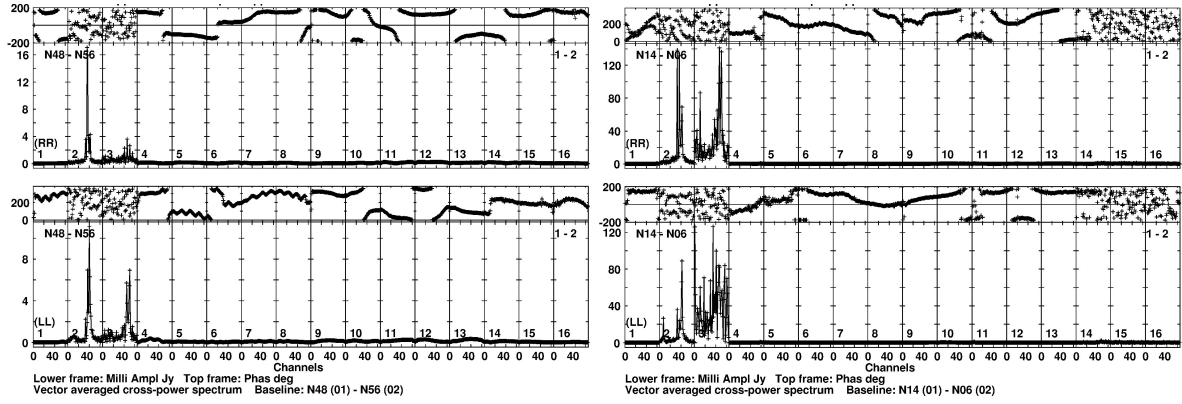
Figure 3.3: Raw spectra of the gain calibrator source, i.e., phase vs. channel, (top frame in each of the four panels) and amplitude vs. channel (bottom frame in each of the four panels) for the right-right and left-left polarizations (top panels and bottom panels, respectively). The panels to the left are for one night of observation in the A-array configuration and the panels to the right are for a C-array observation. No calibration was applied. All baselines and all pointings of the gain calibrator source were combined to produce these plots. We note the RFI in sub-bands 2 and 3.

Table 3.1: Amplitude of the phase calibrator (J1024-0052) in each SPW for all observing blocks

SPW	Frequency (GHz)	Mean flux density (Jy)	Median flux density (Jy)	Standard deviation (Jy)
1	2.060	0.739	0.735	0.029
4	2.444	0.707	0.704	0.025
5	2.572	0.700	0.696	0.023
6	2.700	0.684	0.680	0.023
7	2.828	0.668	0.665	0.024
8	2.956	0.652	0.648	0.026
9	3.084	0.645	0.642	0.020
10	3.212	0.635	0.632	0.020
11	3.340	0.625	0.622	0.021
12	3.468	0.615	0.611	0.020
13	3.596	0.603	0.600	0.021
14	3.724	0.539	0.579	0.129
15	3.852	0.525	0.569	0.153
16	3.980	0.535	0.566	0.124

imaging the target fields, including applying flags (using the task UVCOP), calibration (using the task SPLAT), and exporting the *uv* data and maps (using the task FITTP).

The pipeline performance and output were tested by i) manually reducing separate blocks of VLA-COSMOS observations and comparing the results with the pipeline output, and ii) comparing the output to the CASA-based NRAO reduction pipeline for randomly selected data taken in the A- and C-array configurations. No obvious differences were found. As the pipeline used here was tailored specifically to the COSMOS field (e.g., it includes polarization calibration), after this verification it was further applied to the remaining VLA-COSMOS data sets.

The calibrated *uv* data sets output by the pipeline for each observing block were first run through the AIPS task UVFIX to assure accurately computed positions. We note that applying UVFIX at the end of calibration has the same effect as applying it at the beginning of calibration. They were then further processed in CASA by clipping each calibrated *uv* data set in amplitude (above 0.4 Jy) using the task FLAGDATA[3], splitting the individual pointings using the task SPLIT, and concatenating all existing observations of the same pointing using the task CONCAT. The concatenated (u, v) data for each pointing were then imaged prior to being combined into the final mosaic, as described in detail in the next section.

3.2.3   Self-calibration, imaging, and mosaicking

To image our data we used the MSMF synthesis algorithm developed by Rau & Cornwell (2011) and implemented in CASA. This method uses the entire 2 GHz bandwidth at once to calculate the monochromatic flux density at 3 GHz and a spectral index between 2 and 4 GHz. After extensive testing of various imaging methods (see Chapter 2) we settled for the MSMF method as it allows for a combination of the best possible resolution, *rms*, and image quality. Because of the large data volume, joint deconvolution was not practical and we imaged each pointing individually and then combined them into a mosaic in the image plane.

We found sources that were bright enough (by having a peak surface brightness higher than 5 mJy beam^{-1}) to allow for self-calibration in 44 out of 192 pointings. To prevent artifacts affecting the model used for self-calibration small clean masks were centered around bright sources. An integration time of 3 min, which roughly corresponds to one scan length, was used to obtain phase-gain solutions for these pointings (i.e., only the phase part of the complex gain was solved for and applied). It was typically not possible to find a solution using self-calibration for 10% of the data with the fraction increasing to 20% for a few pointings, which was the maximum value we allowed. We applied gain solutions to the *uv* data but did not apply the flags

---

[3]In total, about $30 - 35\%$ of the data were flagged (using the tasks RFLAG and FLAGDATA).

calculated in the self-calibration process as that usually increased the noise in the map. For the remaining pointings we applied phase gains obtained by self-calibrating the phase-calibrator J1024-0052, as it further reduced artifacts and sidelobes around brighter sources as illustrated in Fig. 3.4.

We used the CLEAN task with Briggs weighting scheme for gridding of visibilities with a robust parameter of 0.5 to obtain the best compromise between the resolution and the noise. Two Taylor terms (NTERMS=2: TT0 and TT1) were used for multifrequency synthesis, which allows the reconstruction of the total intensities and spectral slopes (Rau & Cornwell, 2011). Each pointing was tapered with its own Gaussian to achieve a circular beam, where the difference between the major and the minor axis is 3% at maximum (see Fig. 3.5). Prior to this step the beam was slightly elliptical, but the position angle changed considerably between different pointings. A CYCLEFACTOR=3 was applied for a more robust deconvolution and to prevent artifacts in the map possibly caused by sidelobe intersections. Widefield imaging was necessary to produce correct astrometry far from the pointing center and we used 128 projection planes. We cleaned on three additional spatial scales corresponding to $2\times$, $5\times$, and $10\times$ the synthesized beam size to better handle extended sources such as radio jets and lobes. A GAIN=0.3 parameter was used to speed up this multiscale algorithm. Each pointing map was set to 8 000 pixels on-the-side with a pixel size of 0.2×0.2 arcsec^2. Cleaning was performed down to 5σ in the entire map and further down to 1.5σ using tight masks around sources. These masks were defined manually across the entire observed field by visually inspecting the mosaic[4]. Synthesized beam size variations between different pointings were about $0.03''$, which was small enough to allow restoration of every cleaned pointing to an average circular beam of $0.75''$. Finally, each pointing was corrected for the frequency-dependent primary beam response down to a value of 20% (corresponding to a radius of $10.5'$) using the WIDEBANDOBCOR task. The noise level in the phase center of an individual pointing was usually around $4 - 5\ \mu$Jy beam^{-1}.

To construct the mosaic of all pointings, we used our custom Interactive Data Language (IDL) procedure combined with the AIPS task FLATN to carry out noise weighted addition of the individually imaged pointings. Every pixel in the sum was weighted by the inverse square of the local *rms*, which was determined in the pointing itself via the AIPS task RMSD (see below). We mosaicked both Taylor terms individually using the noise weights calculated from the total intensity maps. The 3 GHz continuum mosaic is shown in Fig. 3.6, where we overplot Gaussian fits to the pixel surface brightness distributions across the mosaic. Cutouts of

---

[4]A preliminary mosaic was generated with pointings cleaned down to 5σ and then used to define cleaning masks. Masks were usually circles with $0.7''$ radius, but they were modified where necessary to accommodate larger (resolved) extended sources. It was not necessary to set clean boxes around known strong sources outside of the imaged area.
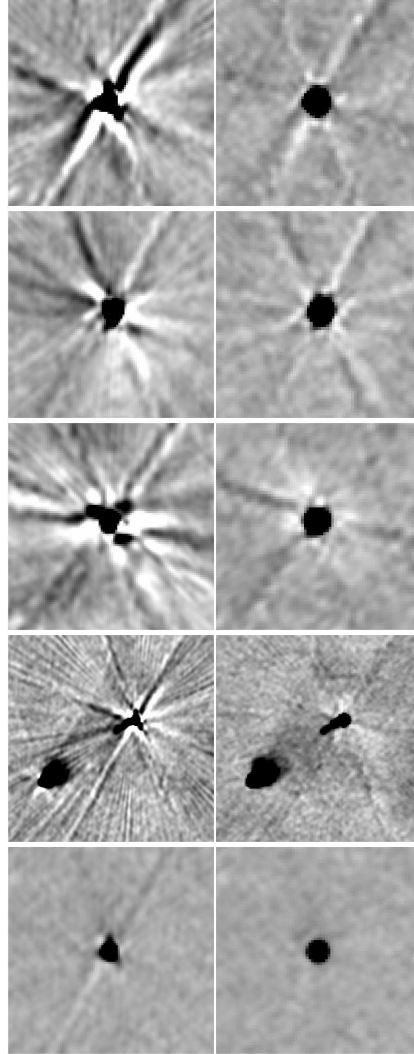
Figure 3.4: Artifacts around bright sources before (*left panels*) and after (*right panels*) applying self-calibration phase-gain solutions. The right column panels also have tapering applied that circularizes the beam shape. The top three rows show the same source with a peak surface brightness of around $S_p \approx 16$ mJy beam^{-1}, but located inside three different pointings that were observed during different time epochs. The fourth row shows artifacts around the brightest source in our data ($S_p \approx 18$ mJy beam^{-1}), which is also extended. The final row illustrates the improvement when applying self-calibration solutions only from the phase calibrator as this source with $S_p \approx$ 2 mJy beam^{-1} has insufficient S/N for self-calibration (see text for details).
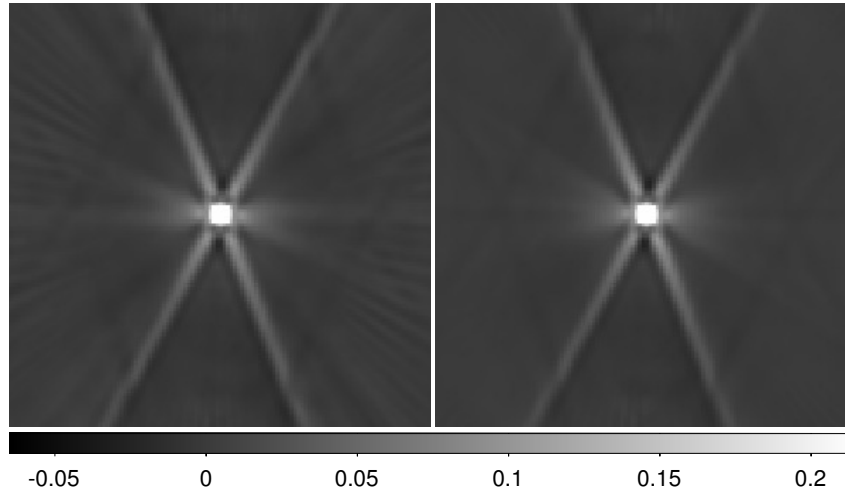
Figure 3.5: *Left panel*: final (A- plus C-array configuration combined) dirty beam of one pointing, after tapering. This beam was used in the cleaning of that pointing (see text for details). *Right panel*: mean stack of all 192 dirty beams. The contribution of radial sidelobes is 10% at maximum.

several extended sources and a mosaic zoom-in are presented in Fig 3.7. The visibility function showing the covered area at a given *rms* is presented in Fig. 3.8. In summary, the final mosaic has a resolution of $0.75''$, with a median *rms* of 2.3 μJy beam^{-1} over the COSMOS two square degrees.

3.3 Cataloging

3.3.1 Extracting source components

To extract source components from the VLA-COSMOS MSMF mosaic and catalog their properties we employed BLOBCAT developed by Hales et al. (2012). Extractor BLOBCAT uses the flood fill algorithm to find islands of pixels (blobs) above a certain S/N threshold. The local noise map used to evaluate the S/N at each pixel position was created from the total intensity mosaic with the AIPS task RMSD with a circular mesh size of 100 pixels. Once BLOBCAT locates islands, it measures the peak surface brightness (S_p) by fitting a two-dimensional (2D) parabola around the brightest pixel, while the total flux density (S_t) is obtained by summing up the pixel values inside the island and dividing the sum by the beam size in pixels. In the next step BLOBCAT takes into account a small positive peak surface brightness bias created by the presence of noise peaks in the map and also corrects for a negative integrated surface brightness bias caused by the finite island size used for integration. We used default parameters when running BLOBCAT (as Hales et al. 2012 ran extensive simulations to optimize them; see also Hales
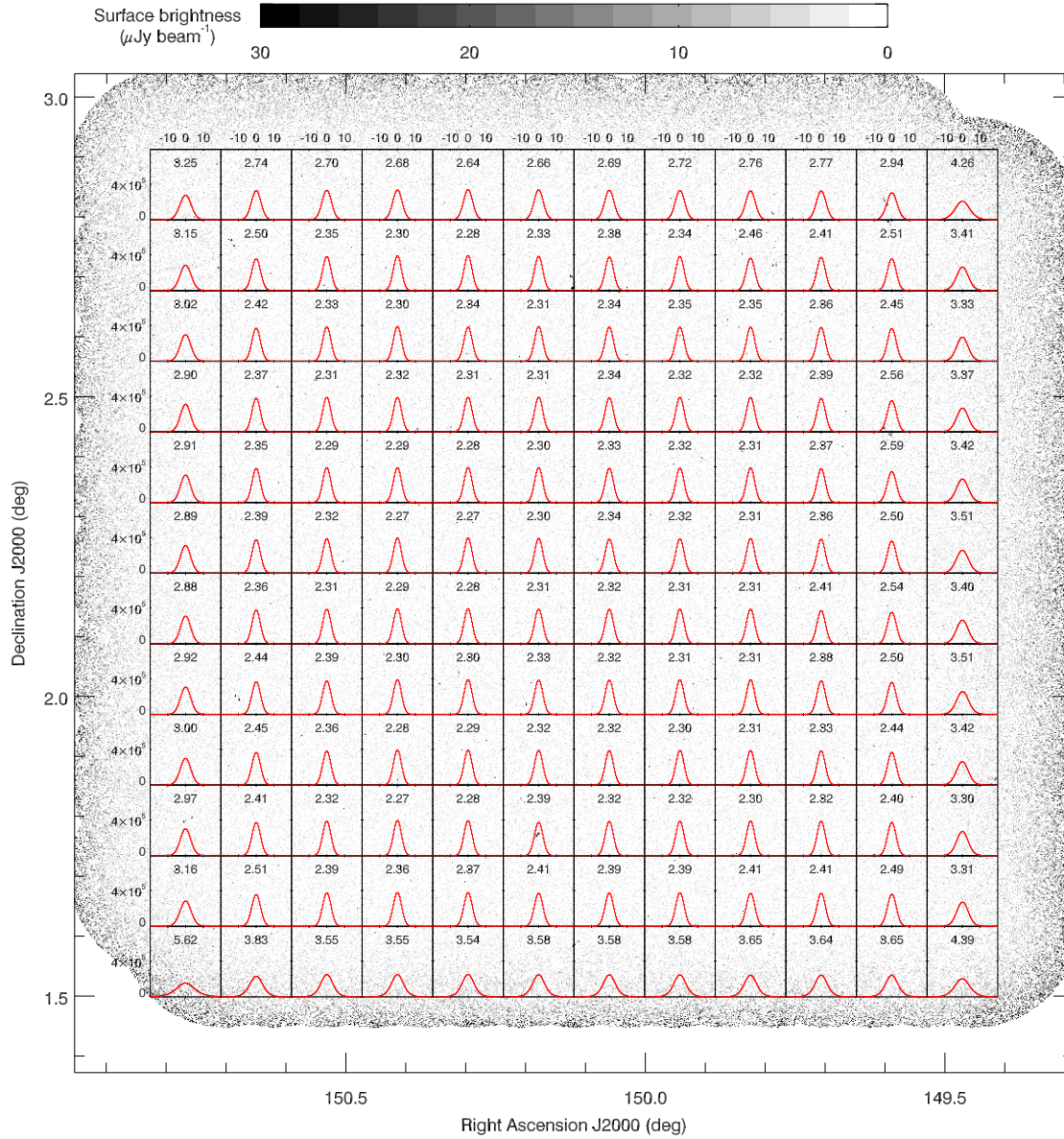
Figure 3.6: Final VLA-COSMOS 3 GHz MSMF mosaic with overlaid Gaussian fits to the pixel surface brightness distributions in various mosaic sectors. The *rms* obtained via the Gaussian fit (in units of μJy beam^{-1}) is indicated in each panel. The panels shown cover the full COSMOS two square degree field. The small-scale ($\sim 1'$) *rms* variations due to the pointing layout are less than 2%.
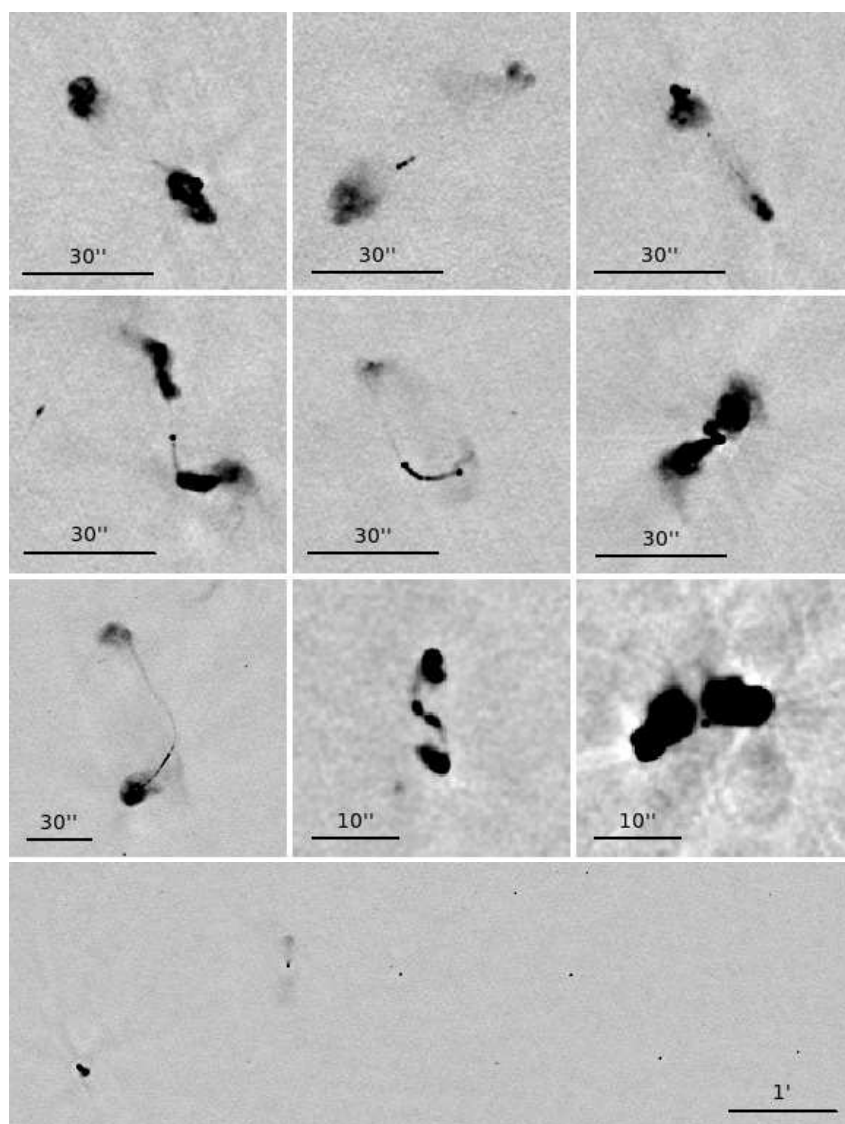
Figure 3.7: Stamps from the VLA-COSMOS 3 GHz continuum mosaic imaged with the MSMF algorithm showing examples of extended and compact radio sources.
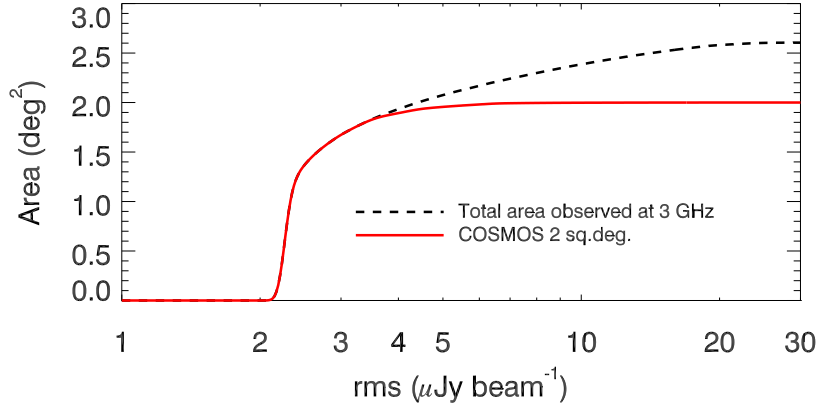
Figure 3.8: Visibility plot showing the total area covered down to a given noise level (black dashed line). Our data extend beyond the COSMOS two square degree field, which ensures more uniform noise inside it (red full line). The median noise level inside the COSMOS two square degrees is $\sigma = 2.3 \, \mu$Jy beam^{-1}.

et al. 2014), where the required size of a blob is at least 3 pixels in RA and 3 pixels in Dec. This was necessary to detect low S/N sources, which would have otherwise been missed owing to our relatively coarse pixel grid. With this setup we recovered 10 899 radio source components with local S/N greater or equal to 5 across the entire observed area. As detailed in Sect. 3.3.3, 67 components have been merged into unique, multicomponent sources resulting in a total of 10 830 radio sources.

3.3.2 Resolved versus unresolved sources

In order to determine whether our identified source components are resolved (i.e., extended, larger than the synthesized beam) we make use of the ratio between total flux density (S_t) and peak surface brightness (S_p) as this is a direct measure of the extension of a radio source. The flux densitites were computed by BLOBCAT as described in the previous section. For a perfect Gaussian unresolved source, the peak surface brightness in Jy beam^{-1} equals the integrated flux density in Jy or $S_t/S_p = 1$. The extension of a radio source increases its total flux density when compared to its peak surface brightness, however, background noise can lower the total flux density (see Bondi et al. 2003). Therefore, in Fig. 3.9 we plot the ratio between the total flux density and the peak surface brightness as a function of the S/N ($= S_p/rms$) for all 10 899 components in the catalog. To select the resolved components, we determined the lower envelope of the points in Fig. 3.9, which contains 95% of the components with $S_t < S_p$ and mirrored it above the $S_t/S_p = 1$ line (upper envelope in Fig. 3.9). The shape of the envelope was chosen following Bondi et al. (2008) and the fit to our data is given as $S_t/S_p = 1 + 6 \times (\text{S/N})^{-1.44}$. We consider the 3 975 components above the upper envelope as resolved. These resolved components were
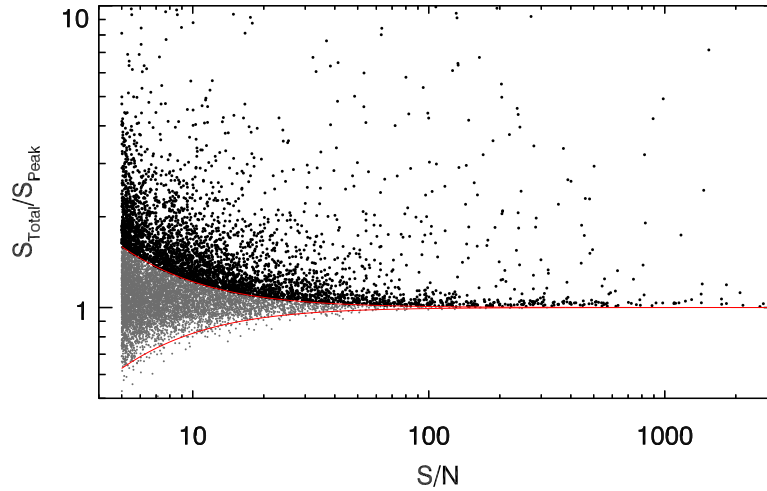
Figure 3.9: Ratio of total flux density to peak surface brightness as a function of S/N ratio. Components below the upper envelope (gray points) indicated by red solid lines are considered unresolved, while those above the upper envelope (black points) are considered resolved (see text for details).

flagged in the catalog. For the unresolved components the total flux density was set equal to the peak surface brightness in the catalog.

3.3.3 Multicomponent sources

Large sources with diffuse structures, such as radio galaxies (see Fig. 3.7) or resolved star-forming disks, can be listed in a component catalog as multiple entries. This can happen for example if there is no significant radio emission between the two radio lobes, or if the local *rms* noise is overestimated because of large-scale faint radio emission, which affects the ability of BLOBCAT to properly detect a contiguous blob. We identified 10 899 components in our mosaic, as described above. In order to generate a source catalog rather than a source component catalog, we aimed to identify such sources and convert the multiple entries into one entry that described the entire source, i.e., listing its proper total flux density and position. For this purpose we visually inspected over 2 500 components. The inspected sample was a combination of the i) brightest 2 500 components, ii) all known multicomponent sources that were identified and listed in the 1.4 GHz joint catalog (126 components), and iii) sources with $R_{EST} > 1 + 30 \times (S/N)^{-1}$ (351 components). The R_{EST} parameter is a size estimate reported by BLOBCAT, which can be used to find sources with non-Gaussian morphology; see Hales et al. (2012, 2014) for more details. Following the procedure already applied to the VLA-COSMOS 1.4 GHz survey sources (Schinnerer et al., 2007) these components were visually inspected with respect to the NIR images, i.e., the $z^{++}YJHK$ stacked maps (Laigle et al., 2016). In total,

we identifed 67 multicomponent sources. As for the previous VLA-COSMOS survey catalogs, we computed their total flux densities using the AIPS task TVSTAT in the area encompassed by 2σ contours, where σ is the local *rms* measured as the average *rms* from a $100 - 300$ pixel wide area around the source, ensuring that the *rms* is not biased by the influence of the strong sources. The source position is then taken to be the radio core or optical counterpart position (if identifiable) or the luminosity weighted mean. In our catalog we then excluded all the components combined into the multicomponent sources, and listed instead the multicomponent source with the above-defined position and total flux density, setting all other cataloged values to -99. A further column MULTI was added designating the multicomponent sources (MULTI=1 for a multicomponent source, and MULTI=0 for a single-component source). We note that the number of multicomponent sources is smaller than that identified in the shallower VLA-COSMOS 1.4 GHz survey. This is due to the higher frequency of breaking-up large sources into multiple components within the latter as it used the AIPS SAD source finding algorithm, when compared to the performance of the BLOBCAT algorithm used here. A full assessment of large sources in the survey will be presented by Vardoulaki et al. (in prep.).

3.3.4 Astrometric accuracy

To assess our astrometric accuracy at the bright end we have compared the positions of 443 sources at 3 GHz with S/N > 20, also detected in the VLBA 1.4 GHz survey (Herrera Ruiz et al., 2017). The results, shown in Fig. 3.10, yield an excellent agreement with a mean offset of $0.01''$ in Δ RA and $0.00''$ in Δ Dec and a standard deviation of $0.01''$ for both. We note that we did not correct the catalog entries for the $0.01''$ offset in Δ RA. We took the standard deviation value ($0.01''$) as the calibration error in RA and Dec to compute the positional uncertainties for our sources using the equations reported in Hales et al. (2012). We note that these are estimated to be accurate for point-sources, but likely underestimated for resolved sources (see Hales et al. 2012 and references therein for details).

3.3.5 Bandwidth smearing

Owing to the finite bandwidth of the antenna receiver, bandwidth smearing (BWS) occurs and radially smears peak surface brightness while conserving the integrated flux density. The effect is a function of distance from the phase center in a given pointing while it reaches a constant smearing value in the combined mosaic (see, e.g., Bondi et al. 2008). Although the bandwidth of the antenna receiver is large (~ 4 GHz), the relevant bandwidth for the smearing effect is only the 2 MHz channel width used to image the data.
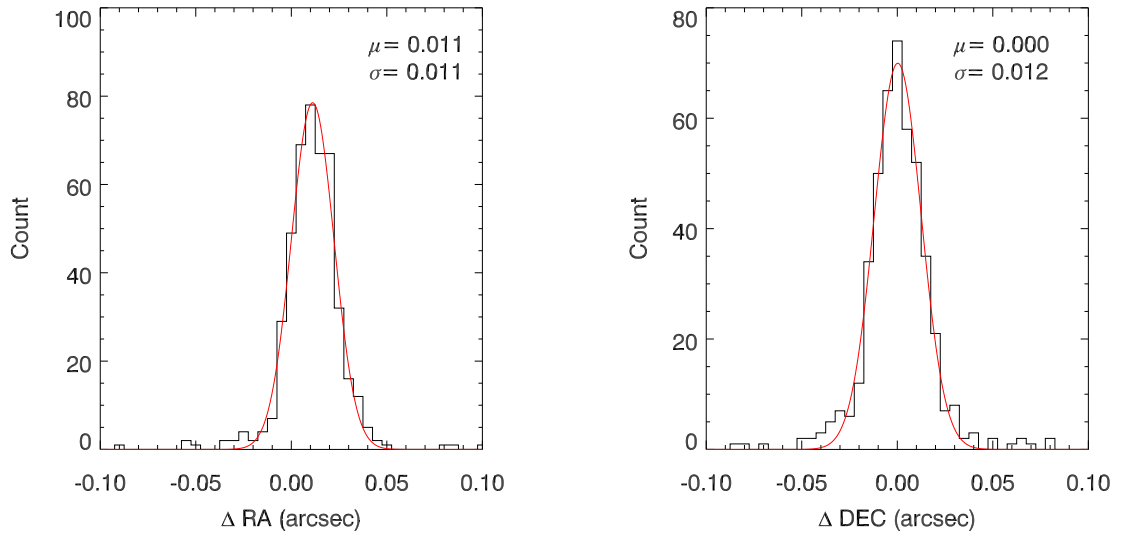
Figure 3.10: Astrometry comparison between 3 GHz VLA and 1.4 GHz VLBA data for 443 VLBA sources (Herrera Ruiz et al., 2017).

To empirically test BWS in our data, we selected 106 point-like ($0.9 \leq S_t/S_p \leq 1.1$) radio sources with S/N>200. Since each source can be observed in up to 11 neighboring pointings, we can compare the peak surface brightnesses obtained in various pointings (S_P) relative to the peak surface brightness retrieved from the mosaic (S_M) as a function of distance from the pointing center. If our data were affected by BWS, S_P/S_M would exhibit a declining trend with increasing distance from the pointing center. This surface brightness ratio, obtained by fitting an inverted parabola at the 106 bright source positions in the individual pointings, and the mosaic is shown in the top panel of Fig. 3.11. The median ratio stays constant ($S_P/S_M \approx 1$) across all distance ranges, with increasing scatter toward higher distances where the noise is amplified by the primary beam correction. This demonstrates that there are no empirical bandwidth-smearing issues. This is also in accordance with theoretical expectations. A theoretical prediction for BWS can be made using the Condon et al. (1998) equation [12] for the reduction of peak response $I/I_0 \approx 1/\sqrt{1 + 0.46\beta^2}$, where $\beta = (\Delta\nu/\nu_0) \times (\theta_0/\theta_{HPBW})$ equals fractional bandwidth times offset in synthesized beam-widths. Using the VLA channel width $\Delta\nu = 2$ MHz, central frequency $\nu_0 = 3$ GHz, distance from the phase center $\theta_0 = 300''$, and beam size of $\theta_{HPBW} = 0.75''$ the estimated peak reduction amounts to about 2%. The distance was chosen as a minimal distance between two different pointing centers. This is illustrated in Fig. 3.12 where we show the peak over total flux density for S/N > 200 sources in different pointings. An offset of $\sim 2.5\%$ is present in this diagram, however, it is not distance dependent, and thus unlikely to

be related to bandwidth smearing. Thus, for the reasons outlined above, we do not apply any corrections for the BWS effect.

3.3.6 The 3 GHz VLA-COSMOS Large Project catalog

A sample page of the catalog is shown in Table 3.5. For each source, we report its ID, 3 GHz name, RA and Dec position (weighted centroid) and error on the position, total flux density with relative error[5], 3 GHz *rms* calculated at the position of the source, S/N, number of pixels used in flux density integration, flag for resolved sources, and flag for multicomponent sources. The peak surface brightness of resolved sources can be obtained by multiplying the S/N with the *rms* value. The catalog as well as the radio map are available in electronic format in the COSMOS Infrared Science Archive (IRSA[6]).

3.4 Radio spectral indices

Given the wide bandwidth of our VLA-COSMOS 3 GHz survey and the existence of previous COSMOS radio surveys, we approached radio spectral index calculations in two ways. The first method uses the MSMF algorithm to construct spectral indices directly from our observed data by fitting a two-term Taylor polynomial to amplitudes between 2 and 4 GHz (MSMF-based spectral index or α_{MSMF} hereafter). The second method uses the cataloged monochromatic flux densities at 3 GHz in combination with the values taken from the 1.4 GHz joint catalog (Schinnerer et al., 2010) to calculate spectral indices between these two frequencies (1.4 – 3 GHz spectral index or $\alpha_{1.4-3\,\mathrm{GHz}}$ hereafter). In Sect. 3.4.1 we investigate systematics in the MSMF spectral index maps, and compare the differently derived spectral indices. In Sect. 3.4.2 we derive the 1.4 – 3 GHz spectral index distribution for the full sample of the 3 GHz sources.

3.4.1 MSMF-based versus 1.4 – 3 GHz spectral indices

We can calculate the MSMF-based spectral indices defined for each source using the wide bandwidth of our observations if the source has a sufficient S/N between 2 and 4 GHz. These spectral indices should be viable for point sources that have S/N>10, and for diffuse emission that has S/N>100. To do so, a mosaic of spectral indices (α-map) was generated by dividing the TT1 mosaic by the TT0 mosaic (see Rau & Cornwell, 2011). For each source, its spectral

---

[5] The flux errors reported do not depend on the number of pixels used for integration, but scale with the source brightness (see Hales et al. 2012, 2014).

[6] http://irsa.ipac.caltech.edu/frontpage/
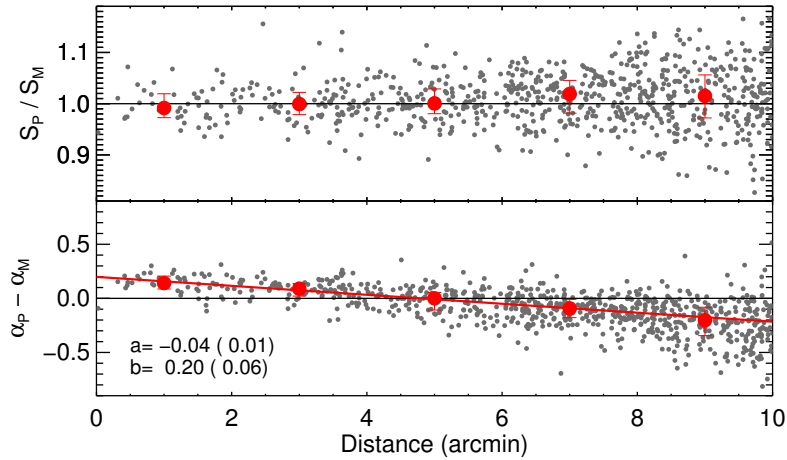
Figure 3.11: Comparison of peak surface brightnesses (*top panel*) and MSMF-based spectral indices (*bottom panel*) determined inside the mosaic (S_M, α_M) and individual pointings (S_P, α_P) as a function of distance from the pointing center for 106 bright, point-like sources ($0.9 \leq S_t/S_p \leq 1.1$, S/N >200) observed in up to 11 neighboring pointings at varying distances from the pointing center (gray points). In both panels the large red points and their corresponding errors indicate median values and interquartile ranges inside 5 equally spaced distance bins. In the bottom panel a linear fit is performed on the median values to obtain the needed correction of the systematic trend across all distances (see text for details).

index was extracted from the pixel in the α map that corresponds to the pixel containing the peak surface brightness in the total intensity mosaic.

To investigate possible systematics in the α map due to wideband primary beam corrections we utilized the 106 bright, point-like sources introduced in Sect. 3.3.5. We derived MSMF-based spectral indices both in the mosaic and individual pointings for these sources. In the bottom panel of Fig. 3.11 we show the difference between such derived spectral indices as a function of distance from the pointing center. The MSMF spectral indices show a systematic steepening with increasing distance, which likely arises due to an imperfect primary beam correction of TT1.[7] To correct for this effect a posteriori (as necessary here), we performed a linear fit to the trend. We then applied this distance-dependent correction to each α-map pointing prior to mosaicking to generate an α mosaic corrected for this effect.

In Fig. 3.13 we compare the corrected spectral indices from MSMF with those derived from the cataloged flux densities at 3 GHz and 1.4 GHz (joint catalog, Schinnerer et al. 2010). The catalogs were cross-matched using a search radius of 1.3″, which is half of the beam size of the lower resolution (1.4 GHz) survey. The sample was further limited to single-component

---

[7]The MSMF algorithm is still in active development and the upcoming software versions should correct for this.
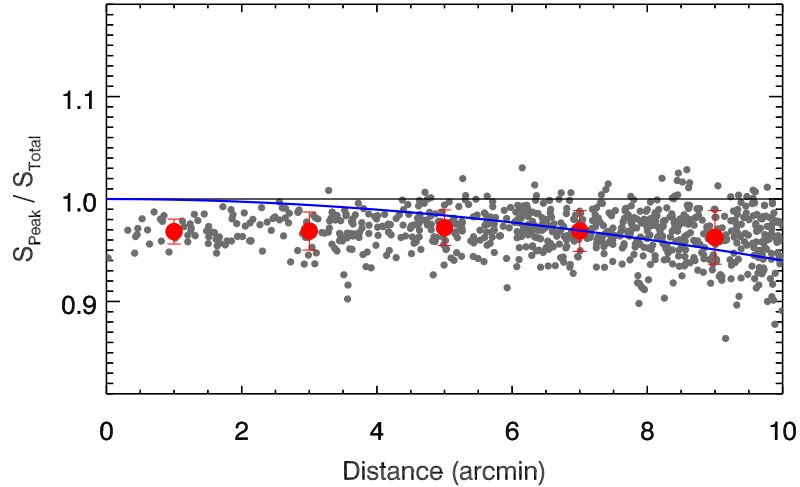
Figure 3.12: Comparison of peak surface brightnesses over total flux densities for 106 bright, point-like sources ($0.9 \leq S_t/S_p \leq 1.1$, S/N>200) observed in up to 11 neighboring pointings at varying distances from the pointing center (gray points). The large red points and their corresponding errors indicate median values and interquartile ranges inside 5 equally spaced distance bins. The theoretical prediction of the bandwidth smearing effect is shown by the blue curve (see text for details).

sources with S/N> 5 in the 1.4 GHz catalog yielding a total of 2 191 sources. Although there are no systematic offsets within the error margins, there is a rather large scatter between the spectral indices obtained with these two methods. A non-negligible portion of this spread is due to the large uncertainty on the in-band (i.e., MSMF derived) spectral indices; a point-like source with S/N ~ 50 and $\alpha = -0.7$ has an uncertainty of ~ 0.1 in its in-band spectral index (see Condon, 2015a). If the MSMF spectral indices had not been corrected, there would have been a systematic offset of -0.2 across the entire S/N range.

In summary, the MSMF-based spectral indices require further corrections after PB corrections are applied to the data. As a result of this, and the large scatter observed between the MSMF-based and $1.4 - 3$ GHz spectral indices, we do not include the MSMF-based spectral indices in the final catalog. New CASA software versions should intrinsically correct for this. For the further analysis of spectral indices presented here we have, therefore, used only the values based on flux density measurements at 3 and 1.4 GHz.

3.4.2   1.4 – 3 GHz spectral indices

A high percentage of 3 GHz sources do not have a counterpart in the 1.4 GHz survey because of the better sensitivity of our 3 GHz survey. We employed the survival analysis to properly constrain the distribution of spectral indices for our 3 GHz selected sample without introducing
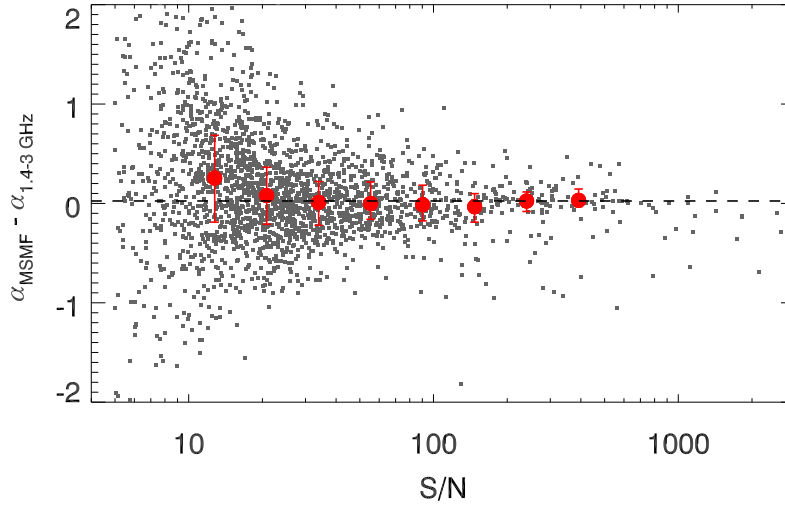
Figure 3.13: Comparison between MSMF-based and 1.4 – 3 GHz derived spectral indices, where the first were corrected for the observed systematic trend illustrated in Fig. 3.11. Red symbols and the corresponding errors denote median spectral indices and interquartile ranges, respectively, for sources in different S/N bins (10<S/N<500). The black dashed line indicates the median value of the red circles set at 0.02.

any bias due to neglecting sources not detected in one of the surveys. This is a statistical method that takes into account both direct detections as well as upper (or lower) limits (see Feigelson & Nelson 1985 and Schmitt 1985 for details).

We first cross-correlated and combined our 3 GHz catalog with the 1.4 GHz joint catalog (Schinnerer et al., 2010) with a maximum separation of $1.3''$, but also including sources without counterparts in one survey or the other. We then removed all sources that fell outside of the area observed at 1.4 GHz as the area observed at 3 GHz is larger (2.6 square degrees.). This was performed to ensure the same area for both surveys. We also removed 1.4 GHz multicomponent sources (80) and their 3 GHz counterparts. The final sample contains 10 523 entries out of which 23% were detected in both surveys, 74% were detected only at 3 GHz, and 3% were detected only at 1.4 GHz, as illustrated in Fig. 3.14 (top panel). If a source was not cataloged in one of the surveys we used five times the local *rms* value at the coordinates of the source as an upper limit on the flux density. Each nondetection at 1.4 GHz yielded one lower limit on spectral index, and similarly, each nondetection at 3 GHz yielded one upper limit.

A Gaussian fit to the distribution of spectral indices detected in both surveys (green line in Fig. 3.14, top panel) results in the peak at $\alpha = -0.84$ and a standard deviation of $\sigma = 0.35$. As this result is valid only for the subsample of 3 GHz sources also detected at 1.4 GHz, we employed the survival analysis to account for the full 3 GHz detected sample. We therefore ran the survival analysis on a single-censored data set that only included detections in both
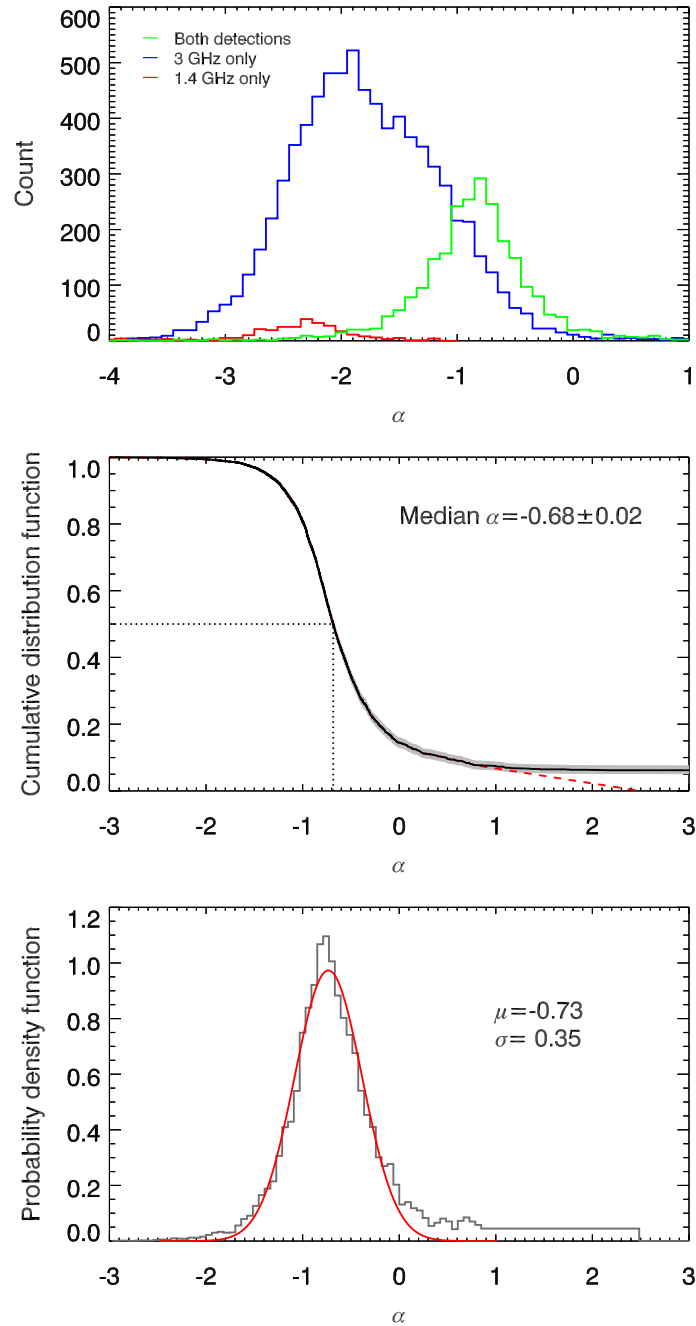
Figure 3.14: *Top panel*: distribution of 1.4 – 3 GHz spectral indices for sources detected at both frequencies (green line), and only at 3 GHz (lower limits, blue line) or 1.4 GHz (upper limits, red line). *Middle panel*: cumulative distribution function (CDF; black line) and error estimate (gray shaded area) of spectral indices calculated using the survival analysis also taking lower limits into account. The red dashed line shows a linear extrapolation of the distribution to zero assuming the maximum theoretically attainable spectral index of $\alpha = 2.5$ (see text for details). *Bottom panel*: probability density function for spectral indices calculated as a first derivative of the CDF extrapolated to 0 at high end (middle panel). A Gaussian fit to the distribution is also shown (red curve) and its mean and standard deviation are indicated in the panel.

surveys and lower limits. The method assumes that limits follow the same distribution as direct detections and generates a cumulative distribution for all sources in the sample. This is shown in the middle panel of Fig. 3.14. There was enough overlap between direct detections and lower limits enabling the survival analysis to properly constrain the median of the total distribution to $\alpha = -0.68 \pm 0.02$, even though there were three times more limits than detections. This method however cannot constrain all lower limits and the cumulative function does not converge to 0, yielding a total of 6% unconstrained sources. To constrain these (as needed to derive the spectral indices probability density function - PDF; see below) we employed a physical argument that a radio source exhibiting standard synchrotron self-absorption cannot have a spectral index higher than $\alpha_{max} = 2.5$ (Rees 1967; unless it is extremely rare and exotic; for example see Krishna et al. 2014). Our data can also constrain the distribution of spectral indices only up to value of $\alpha = 0.8$, since this interval contains 99.5% of sources detected at both 1.4 and 3 GHz. With these limits we can at best assume a flat probability that unconstrained sources have $0.8 < \alpha < 2.5$, and we can formally extrapolate the cumulative distribution function to 0 (red dashed line in Fig. 3.14, middle panel). Having constrained this, we then derived the PDF for spectral indices of our 3 GHz sources by calculating the first derivative of the cumulative distribution function extrapolated to 0. The PDF is shown in the bottom panel of Fig. 3.14. The best-fit Gaussian to the PDF yields a mean of $\alpha = -0.73$ and a standard deviation of $\sigma = 0.35$. Both the median of the distribution and the mean of the Gaussian fit agree very well with previous work carried out on spectral indices (e.g., Condon 1984; Lisenfeld & Völk 2000; Kimball & Ivezić 2008) and we conclude that our catalog flux densities do not show any significant systematics.

3.5 Radio source counts corrections

A well-established approach to estimate the combined effects of noise bias, source extraction and flux determination systematics, inhomogeneuos noise distribution over the imaged field, and resolution bias on the measured source counts (completeness and bias corrections hereafter) is to rely on mock samples of radio sources, as described in Sect. 3.5.1. As these corrections do not take into account the fraction of spurious sources (as the mock sources are always inserted into the same mosaic) in Sect. 3.5.2, we separately derive the false detection rate. The combination of the two corrections then yields the net radio source count corrections.

3.5.1 Completeness and bias corrections

We here describe the Monte Carlo simulations used to derive the completeness and bias corrections. Mock sources were injected over the imaged field and then recovered using the same technique adopted for the real radio sources. To properly construct mock sources, the flux density and size distributions have to be assumed. The results of the final simulations yield the completeness and bias corrections which should be applied to any subsequent statistical analysis.

Retrieval of mock radio sources

The procedure adopted to insert and retrieve mock sources in and from the mosaic is as follows. Since BLOBCAT does not produce a residual map, we inserted mock sources (Gaussian in shape) directly into the continuum map avoiding already cataloged components. This procedure was limited to the central two square degrees of the mosaic. For each mock source, a square shape with a width of $6\sigma + 21$ pixel on the side was required to be free of any cataloged emission (real or mock), where σ is the standard deviation along the Gaussian major axis. The positions were randomly chosen until this was satisfied. At a resolution of $0.75''$ the continuum map is mostly empty of sources and confusion is negligible. We did not observe any systematic clustering of mock sources toward the less populated parts of the mosaic (more noisy parts closer to the edge for example) by requiring no overlap between the components. After all mock sources were inserted, we ran BLOBCAT with the exact parameters as performed for the real sources. Since the extraction was carried out on a map containing both real and mock emission, all the 10 899 real components were always recovered and then removed from the extracted catalog, prior to further processing. To generate realistic mock catalogs of radio sources, however, we needed to assume i) a flux density distribution in (and below) the range tested by the observations and ii) an angular size distribution of the radio sources. This is described in detail in the following sections.

Flux density distribution

We simulated the flux density distribution using both a simple power law model (PL model) and a multinode power law (MPL model) that better reproduces the observed source counts below 500 μJy. In the former case we used the 1.4 GHz source counts from previous surveys scaled to 3 GHz (see Bondi et al. 2008). The multinode power law model is that derived by Vernstrom et al. (2014) (see their Table 4, Zone 1). For both models the mock catalogs were generated down to a total flux density of 5 μJy and contained more than 40 000 (65 000) objects

in the PL (MPL) model. This also allowed us to count sources with flux densities below the S/N threshold as positive noise fluctuations might lead these to have a measured peak surface brightness above our source detection threshold. As shown below, the results of our simulations do not yield differences between the two models, and we adopted the MPL model for our final simulations.

Angular size distribution

We needed to assign an intrinsic angular size to each mock source. Unfortunately, a satisfying description of the intrinsic source angular size distribution at sub-mJy flux density is still missing and we needed to rely on extrapolations from higher flux densities. Bondi et al. (2008) used a simple power law parametrization distribution of the angular sizes of the sources as a function of their total flux density. We followed the same method with some adaptations, as described below.

The angular size (θ) distribution was simulated, assuming a power law relation between angular size and flux density ($\theta \propto S^n$). This relation was normalized using the cumulative angular size distribution derived at ~ 1 mJy from the Visible MultiObject Spectrograph (VIMOS) VLT Deep Survey (VVDS) 1.4 GHz observations with a resolution of $6''$ (Bondi et al., 2003). The relatively low resolution of the VVDS survey allowed us to avoid bias against sources with angular sizes of up to $15''$ in our simulations (Bondi et al., 2003). We explored the range of n values between 0.3 and 1.0 in steps of 0.1. To infer the best n value, the angular size distribution of the sources from the catalog in a specific total flux density range was compared with the corresponding distribution derived from the mock samples with different n values. The value of n that gave the best match between the angular size distribution of observed and mock sources was then chosen as the best approximation for the intrinsic source size versus total flux density relation.

Since the observed source angular sizes are not provided by BLOBCAT, these were estimated using the relation between the ratio of the total flux density and peak surface brightness and angular sizes,

$$\frac{S_t}{S_p} = \frac{\sqrt{\theta_M^2 + \theta_b^2}\sqrt{\theta_m^2 + \theta_b^2}}{\theta_b^2},$$

(3.1)

where S_t is the total flux density, S_p is the peak surface brightness, θ_b is the FWHM of the circular beam ($0.75''$ in our observations), θ_M and θ_m are the major and minor FWHM intrinsic (deconvolved) angular sizes; see Bondi et al. 2008, where the same approach was used to derive a size estimate of sources affected by bandwidth smearing. In doing so we needed to make

some assumptions on the morphology of the sources and in particular on how the sources are, eventually, resolved. We considered two limiting cases as follows:

1. *Elongated geometry*: sources are resolved in only one direction. This implies that $\theta_m = 0$ and

$$\frac{S_t}{S_p} = \frac{\sqrt{\theta_M^2 + \theta_b^2}}{\theta_b^2}. \tag{3.2}$$

   The simulated mock sources were accordingly generated as sources extended in one direction and eq. 3.2 is the appropriate relation between S_t/S_p and the angular size.

2. *Circular geometry*: sources are uniformly resolved in all directions. This implies that $\theta_M = \theta_m$ and

$$\frac{S_t}{S_p} = \frac{\theta_M^2 + \theta_b^2}{\theta_b^2}. \tag{3.3}$$

   The simulated mock sources were accordingly generated as sources uniformly extended in all directions and eq. 3.3 is the appropriate relation between S_t/S_p and the angular size.

Mock catalogs were generated for each combination of the 2 source count models (PL and MPL), the 8 different n values (0.3 − 1.0 in steps of 0.1), and the 2 different source geometries ($\theta_m = 0$ or $\theta_m = \theta_M$). For each of these 32 combinations we generated and merged 10 different mock samples. Then, we derived for each of the 32 different mock catalogs the S_t/S_p distributions for sources with $S_t < 100\,\mu$Jy, splitting them into two subranges: $S_t \leq 40\,\mu$Jy, and $40 < S_t \leq 100\,\mu$Jy. This range in total flux density is more affected by the choice on the intrinsic source size distribution and therefore is the best suited for a comparison between the S_t/S_p distribution of the real cataloged sources and that derived from the mock samples reprocessed as the observed catalog. Using S_t/S_p as a proxy for the angular size of the radio sources has the advantage that we do not need to deal with upper limits in the measured source sizes because of sources classified as unresolved.

The results of this comparison can be summarized as follows. No significant differences were found using the PL or MPL distributions for the source counts. For this reason, we were able to adopt either of the two models in the following analysis and we decided to use the MPL model, which provides a more realistic and detailed description of the observed source counts. However, none of the 16 combinations of n value and source geometry provided a

satisfying match between the S_t/S_p distribution of the reprocessed mock catalog and that of the observed catalog, in the flux density range $S_t < 100\ \mu$Jy. While some combinations of parameters provided a reasonable match for sources with $S_t \gtrsim 40\ \mu$Jy, they all failed to reproduce the observed distribution of S_t/S_p below this threshold. In particular, the mock samples showed lower values of S_t/S_p than the catalog for $S_t \lesssim 40\ \mu$Jy. This is shown in Fig. 3.15 where we plot in the two panels the S_t/S_p distribution for sources with $S_t < 40\ \mu$Jy and sources with $40 < S_t \le 100\ \mu$Jy, respectively. Together with the observed distribution derived from the sources in the catalog we also plot the distribution obtained from our original simulation using $n = 0.6$ and elongated geometry. The two distributions are clearly shifted and this effect is found in all the simulations.

This result is not completely unexpected, The extrapolation to very low flux density of our power law relation between angular size and flux density, which has been previously tested only for sources more than one order of magnitude brighter, produces mock samples of radio sources dominated by extremely compact objects at the faint end of the total flux density distribution. For instance, for the simulations shown in Fig. 3.15, 45% of all the sources with $S_t \le 40\ \mu$Jy have $S_t/S_p < 1$. This result is at odds with the distribution of the observed catalog, where only 26% of the observed sources fainter than 40 μJy have $S_t/S_p < 1$.

The simplest way to decrease the number of extremely compact objects at the faint end of the flux distribution in our mock sources, without modifying the adopted power law relation between angular size and flux density, is to apply a minimum angular size to the faint mock sources. We tested the following expression for θ_{min}:

$$\theta_{min} = k_1 e^{-(S_t/k_2)^2}, \tag{3.4}$$

in which the exponential part is motivated by the fact that, on the basis of the analysis shown in the lower panel of Fig. 3.15, no θ_{min} is required at flux densities $\gtrsim 40$ microJy. We included the minimum angular size requirement in our procedure to generate the mock samples of radio sources and repeated the simulations, extraction process, and comparison of the S_t/S_p distributions. By varying the parameters k_1 and k_2 we found that the best value for the k_2 parameter is $k_2 = 40\ \mu$Jy, while for the normalization factor k_1 is equal to 0.3 (for the elongated geometry) and 0.2 (for the circular geometry). The different normalization is necessary because for a given intrinsic source size the area covered by a circular source is larger and derived from eq. 3.2 and 3.3. As shown in Fig. 3.15, this time we found a very good agreement between the observed and simulated distributions of S_t/S_p at low flux densities as well. In particular, for the simulation shown in Fig. 3.15, introducing a minimum angular size as a function of the total flux density reduces the percentage of faint sources ($S_t \le 40\ \mu$Jy) with $S_t/S_p < 1$ from 45% to 30% close to
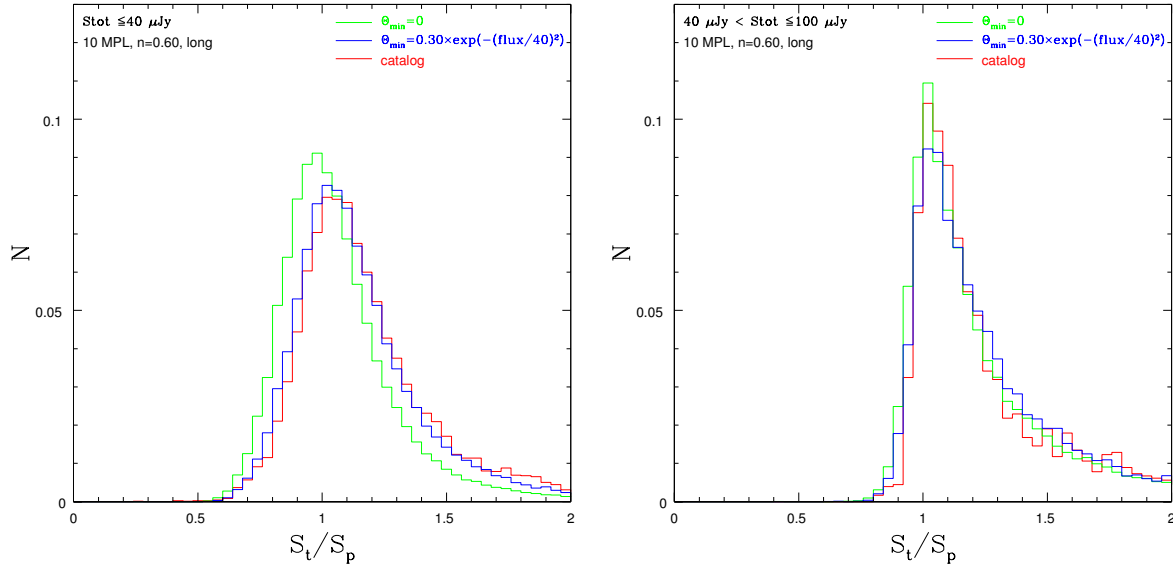
Figure 3.15: Total-to-peak flux density ratio distributions in two total flux density ranges: $S_t \leq 40\,\mu$Jy (*left panel*) and $40 < S_t \leq 100\,\mu$Jy (*right panel*). Each panel shows the distribution of the observed sources (red histogram) derived from the 10 sets of simulations using an elongated geometry and $n = 0.6$, no minimum angular size (green histogram) and with a minimum angular size $\theta_{min} = 0.3e^{-(S_t/40\,\mu Jy)^2}$ (blue histogram).

the observed value of 26%. Thus, we adopted the above-described size parametrization for our final simulations used to derive the completeness and bias corrections.

We further found that the geometry of the radio sources has some effects on the results we obtained. We obtained the best match between the S_t/S_p distributions for $n = 0.5 - 0.6$ for elongated geometry ($\theta_m = 0$), and we obtained the best match for $n = 0.6 - 0.7$ for circular geometry ($\theta_m = \theta_M$). We note that both the assumptions made on the source geometry are clearly simplistic and real sources will consist of a mix of elongated and circular sources. Thus, to compute our final completeness and bias corrections for the adopted MPL flux density distribution models we computed the completeness and bias correction factors using the average of those from the four best-matched simulations: i) elongated sources with $n = 0.5$, and $n = 0.6$, and ii) circular sources with $n = 0.6$, and $n = 0.7$, as described in more detail in the next section.

Derivation of completeness and bias corrections

We generated 60 mock catalogs using the parameterization as described above. The mock sources were inserted into the mosaic and retrieved with BLOBCAT. The retrieved mock sources were then cross-correlated with the input mock catalog and their measured total flux density
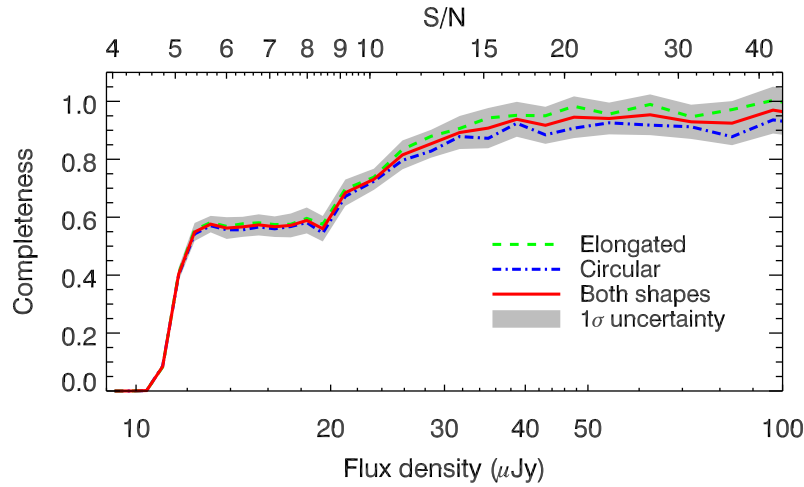
Figure 3.16: Completeness of our 3 GHz source catalog as a function of flux density and S/N. The mean completeness of all Monte Carlo runs (red line) and its standard deviation (gray shaded area) are shown. Also shown are the corrections when elongated (dash-dotted line) and circular (dashed line) geometries are assumed.

chosen to be either their integrated flux density if resolved, or peak surface brightness if unresolved. For this, the same S_t/S_p envelope was used as described in Sect. 3.3. Lastly, successfully extracted mock sources and original mock sources were binned separately in flux densities. The ratio of their numbers in each flux density bin represents the completeness and bias correction factor.

In Fig. 3.16 we show the net result of the above-described Monte Carlo simulations for the MPL model and best-matched simulations, i.e., i) elongated sources with $n = 0.5$ and 0.6, and ii) circular sources with $n = 0.6$ and 0.7. We take the average of these simulations as the completeness and bias correction with a confidence interval that takes into account the differences within the six sets of simulations. This is tabulated in Table 3.2. For reference, in Fig. 3.16 we also show the average completeness and bias corrections obtained using only the elongated and circular geometry approximations. The net curve yields values of about 55% at $12\,\mu$Jy (S/N = 5.2), and rather constant up to $20\,\mu$Jy (S/N < 9), beyond which they rise to a 94% completeness above $40\,\mu$Jy (S/N $\geq$ 16). The mean error of the completeness and bias corrections is 5%. The two (elongated and circular geometry) approximations are consistent up to $\sim 30\,\mu$Jy, beyond which they start diverging with the circular approximation being systematically lower at higher flux densities. However, beyond this limit both curves saturate at fairly constant values (~ 0.92 for circular and 0.98 for elongated morphology), implying average values of over 95% for fluxes higher than $\sim 40\,\mu$Jy.

Table 3.2: Completeness and bias correction factors for the VLA-COSMOS 3 GHz catalog as a function of flux density

Flux density (μJy)	Completeness and bias correction factor (C_{compl})	Error
< 10.4	0	-
11.0	0.08	0.01
11.6	0.40	0.02
12.3	0.55	0.03
13.0	0.58	0.03
13.8	0.56	0.04
14.6	0.57	0.03
15.5	0.57	0.04
16.4	0.57	0.04
17.3	0.57	0.04
18.4	0.59	0.04
19.4	0.56	0.04
21.1	0.68	0.05
23.3	0.73	0.04
25.8	0.82	0.05
28.6	0.85	0.05
31.7	0.89	0.06
35.1	0.91	0.07
38.8	0.94	0.06
43.0	0.92	0.06
47.6	0.95	0.07
53.9	0.94	0.05
62.4	0.95	0.07
72.2	0.93	0.06
83.5	0.92	0.08
96.7	0.97	0.08
> 100[a]	1.00[a]	0.05[a]

[a]Assumed corrections for fluxes > 100 μJy.

Biases addressed

There are several effects and biases that occur in the cataloging process that we addressed through our simulations. Firstly, an incompleteness in the extracted catalog exists as real sources on the sky are not be detected if i) their peak surface brightness falls below the chosen threshold of 5σ because of fluctuations in the local *rms*, or ii) they are extended enough for their peak surface brightness to fall below the detection threshold, even though their integrated flux density is well above it. Secondly, a contamination effect is also present. If a source is detected, its flux density might be wrongly computed because of the presence of a noise peak. Statistically, this effect is mostly symmetric around the mean flux density. However, when we set the total flux density of an unresolved source to its peak surface brightness we may introduce an asymmetric bias toward smaller flux densities. Some sources with $S_t > S_p$ within the envelope in Fig. 3.9 might truly be resolved, however noise variations do not allow us to determine this with sufficiently high accuracy leading to a potential bias. The final result is that a source can jump into a flux density bin where it does not belong, thus increasing its contamination. The combination of completeness, which always decreases with decreasing flux density, and the significant number of sources that move from their original flux density bin to another owing to errors in flux measurement at faint flux densities, produces the flat distribution of the completeness and bias correction factor seen at flux densitites of $\sim 12 - 20\,\mu$Jy in Fig. 3.16.

In summary, the simulations we performed account for both the fraction of nondetected sources (incompleteness), and also the redistribution of sources between various flux density bins. Thus, in principle, its value can be larger than one if the contamination is high. These corrections, however, do not take into account the fraction of spurious sources as a function of flux density, which are derived separately in the next section.

3.5.2 False detection rate

To assess the false detection rate of our source extraction we ran BLOBCAT on the inverted (i.e., multiplied by -1) continuum map with the same settings used for the main catalog. Since there is no negative emission on the sky, every source detected in the inverted map is per definition a noise peak (i.e., a false detection). The source extraction returned 414 negative detections with S/N ≥ 5 across the entire observed field, 95 of which were outside the central two square degrees.

The highest S/N negative detections were predominantly located around true bright sources as they suffer from artifacts; up to six negative components could be found around a single bright object due to the VLA synthesized beam shape (see also Sect. 7.1.1. in Vernstrom et al.

Table 3.3: False detection probability as a function of S/N and flux density in the COSMOS two square degree field

S/N	Fraction	Flux density (μJy)	Fraction ($F_{\text{false}-\text{det}}$)
5.05	0.24	10.75	0.40
5.15	0.15	11.00	0.38
5.25	0.11	11.25	0.27
5.35	0.09	11.50	0.21
5.45	0.06	11.75	0.15
5.55	0.03	12.00	0.13
5.65	0.02	12.25	0.09
5.75	0.03	12.50	0.07
5.85	0.01	12.75	0.06
5.95	0.02	13.00	0.06
6.05	0.01	13.25	0.03
6.15	0.01	13.50	0.04
6.25	0.02	13.75	0.03
6.35	0.01	14.00	0.05
6.45	0.00	14.25	0.03
6.55	0.00	14.50	0.03
6.65	0.01	14.75	0.08
6.75	0.01	15.00	0.08
6.85	0.00	15.25	0.00
6.95	0.01	15.50	0.01

2014 for an explanation of this effect). Since extraction of real emission does not exhibit this behavior, we removed all negative components that were less than $3''$ away from a real source with S/N > 100. This step removed further 40 components. We additionally removed four sources with catastrophic peak estimates, which increased their S/N by more than a factor of four owing to poor parabola fits. We note that there were no such sources in the catalog of real emission. The remaining 275 negative detections within the inner two square degrees were then classified into resolved and unresolved using the same envelope as was carried out for the real data. Finally, they were binned in S/N and flux densities alongside true detections to enable direct comparison. The results are shown in Fig 3.17 and also listed in Table 3.3. As expected, only the lowest S/N bins have a noticeable portion of false detections (24% for S/N = 5.0 − 5.1), which quickly decreases to less than 3% for any S/N bin at S/N > 5.5. The estimated fraction of spurious sources over the entire catalog above S/N > 5 (5.5) drawn from the inner two square degrees is only 2.7 (0.4)%.
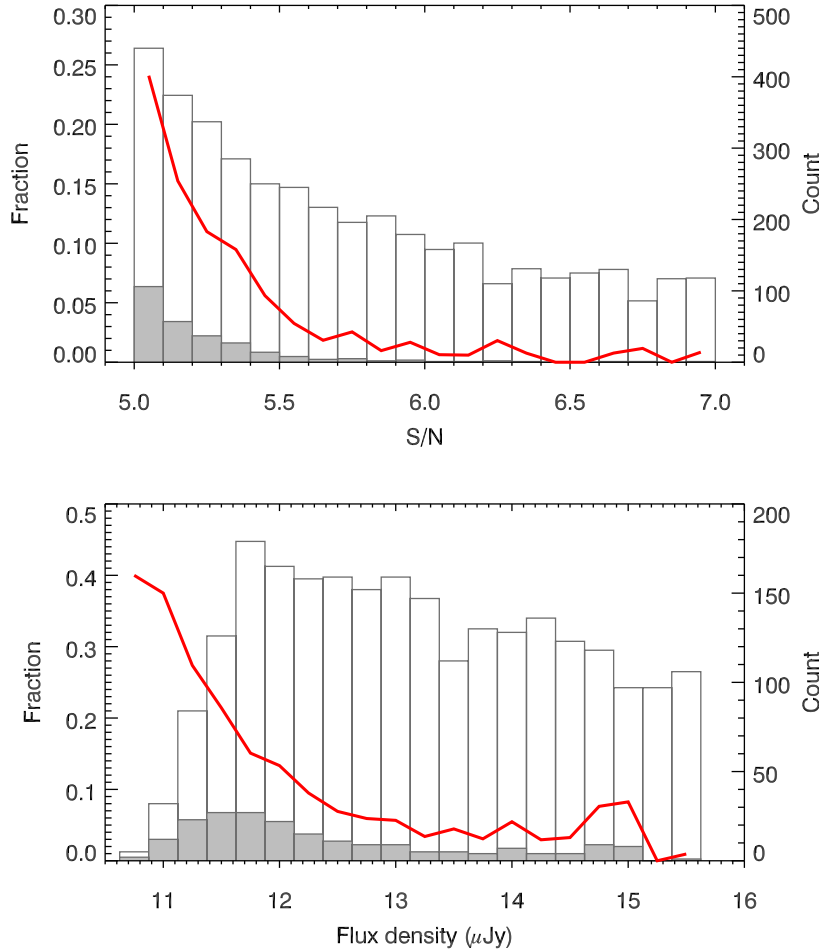
Figure 3.17: Fraction of false detections (red line) as a function of S/N (*top panel*) and flux density (*bottom panel*). The open (filled) histogram shows the number of components cataloged in the observed 3 GHz mosaic (detected in the inverted map), and limited to the central two square degrees. The data are also listed in Table 3.3.

3.6 Radio source counts

In this section we present our 3 GHz radio source counts (Sect. 3.6.1), and compare them to 3 GHz and 1.4 GHz counts available in the literature (Sects. 3.6.2 and 3.6.3, respectively).

3.6.1 VLA-COSMOS 3 GHz radio source counts

We present our 3 GHz radio source counts normalized to Euclidean geometry, both corrected and uncorrected, in the top panel of Fig. 3.18. In Table 3.4 we list the counts, errors, number of sources, and radio source count corrections (i.e., completeness and bias and false detection fraction corrections) in each flux density bin. The source count errors take into account both

the Poissonian errors as well as completeness and bias correction uncertainties. Most of our sources are located at low flux densities (below 0.5 mJy) with more than 500 sources in each flux density bin below 60 μJy resulting in small Poissonian errors. As evident from the plot, our source counts at 3 GHz exhibit a flattening at about 0.3 mJy that continues one order of magnitude in flux densities down to 30 μJy, steepening further at fainter flux densities.

3.6.2 Comparison with 3 GHz counts from the literature

In the middle panel of Fig. 3.18 we compare our 3 GHz source counts with other 3 GHz counts available in the literature (Condon et al., 2012; Vernstrom et al., 2014). Condon et al. (2012) performed a $P(D)$ analysis using 3 GHz confusion-limited data based on 50 hours of on-source C-array configuration observations of one VLA pointing targeting the Lockman hole and reaching an *rms* of 1 μJy beam^{-1}. Fitting single power law models to the data the analysis allowed these investigators to constrain the counts of discrete sources in the $1 - 10\,\mu$Jy range, also shown in Fig. 3.18. Vernstrom et al. (2014) performed a more complex $P(D)$ analysis on the same data fitting various (modified power law, and node-based) models to the data allowing them to probe the counts down to 0.1 μJy. In Fig. 3.18 we show the counts based on the fit of a phenomeno-logical parametric model of multiple joined power laws (their node-based model) applied to the inner circular area with a $5'$ radius (their Zone 1; see Vernstrom et al. 2014 for details).

The counts derived here are in very good agreement with those derived by Condon et al. (2012). Fitting the five faintest flux density bins using a power law, $dN/dS \propto S^{\gamma}$, we find that the slope $\gamma = -1.72$ is perfectly consistent with that inferred by Condon et al. (2012), while our normalization is slightly lower. Our comparison to the Vernstrom et al. (2014) results shows that the counts are in agreement down to $\sim 30\,\mu$Jy with a discrepancy at fainter flux densitites as our counts are systematically lower than theirs.

In general, the strength of the $P(D)$ analysis is the ability to probe counts below the nominal noise in the data, while avoiding resolution biases as it is applied on confusion-limited (thus, low resolution) data. However, as the $P(D)$ analyses discussed above were performed on a single VLA pointing, the resulting counts may be subject to cosmic variance due to the small area covered. This could potentially explain the observed discrepancy between the VLA-COSMOS 3 GHz Large Project counts based on a two square degree area and the Vernstrom et al. (2014) results based on a 0.022 square degree area (their Zone 1). To test this we subdivided the two square degree COSMOS field into 100 square and nonoverlapping subfields, each with an area of 0.020 deg^2 roughly corresponding to a circle with a radius of $5'$. In the middle panel of Fig. 3.18, we show the range of such obtained counts (corrected for completeness and bias, and false detection fractions, calculated on the full two square degrees and described in
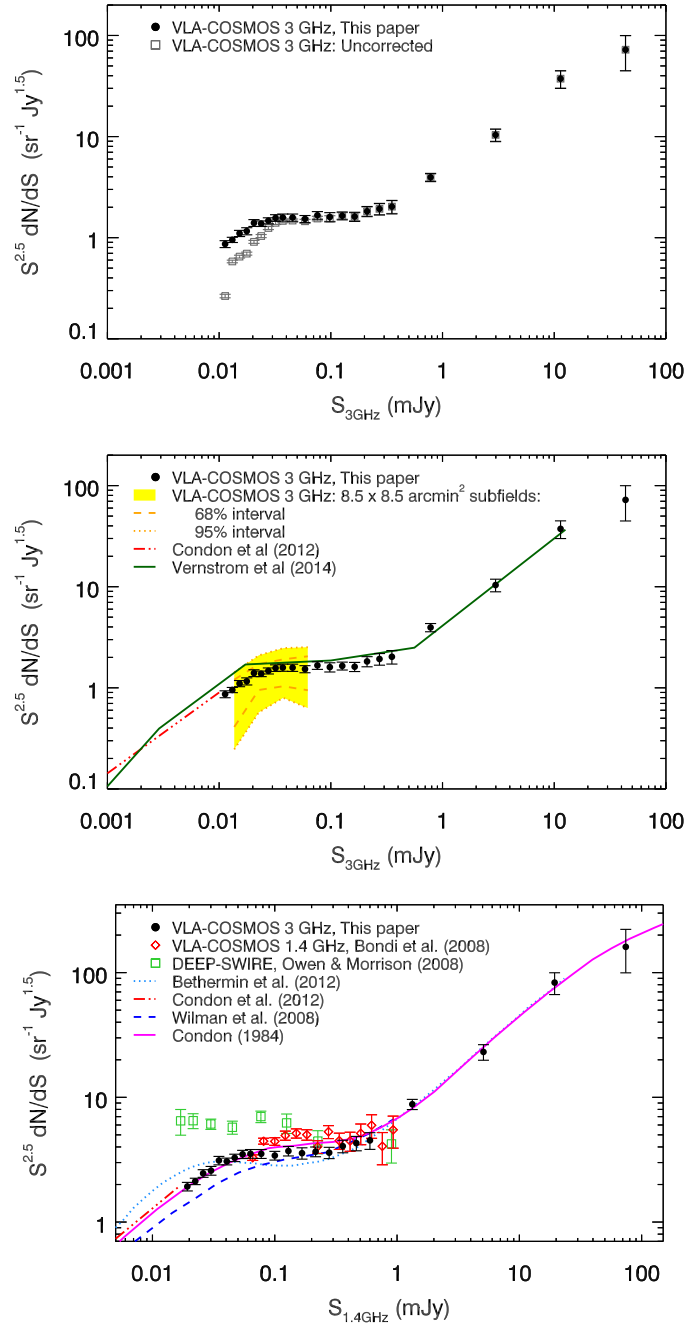
Figure 3.18: *Top panel:* VLA-COSMOS 3 GHz Euclidean-normalized radio source counts, corrected using the completeness and bias and false-detection correction factors (black filled points) and without corrections (gray squares). *Middle panel:* Same counts compared to Condon et al. (2012) $P(D)$ analysis with a single power law (dot-dashed red line) and Vernstrom et al. (2014) $P(D)$ analysis with multiple power laws (green line) at 3 GHz. The yellow shaded area contains 95% of different source counts obtained from 100 square and nonoverlapping (8.5×8.5 arcmin^2) subfields of the COSMOS field. *Bottom panel:* Same counts, but shifted to the 1.4 GHz observed frame using a spectral index of $\alpha = -0.7$ prior to binning (black filled points). A selection of existing 1.4 GHz source counts in the literature is also shown, as indicated in the legend.

Table 3.4: Radio source counts at 3 GHz within the COSMOS two square degree field, normalized to Euclidean geometry

Flux density (mJy)	Counts[a] $(Jy^{1.5}sr^{-1})$	Error[b] $(Jy^{1.5}sr^{-1})$	N	Correction factor[a]
0.011	0.866	0.068	631	3.27
0.013	0.952	0.056	1109	1.64
0.015	1.10	0.078	991	1.70
0.018	1.16	0.094	849	1.67
0.020	1.40	0.11	888	1.54
0.024	1.38	0.086	811	1.33
0.028	1.47	0.10	780	1.18
0.032	1.57	0.12	702	1.12
0.037	1.59	0.13	587	1.08
0.045	1.58	0.13	753	1.07
0.059	1.53	0.12	505	1.05
0.076	1.67	0.14	366	1.08
0.098	1.60	0.17	250	1.03
0.13	1.65	0.15	181	1.00
0.16	1.62	0.17	121	1.00
0.21	1.82	0.21	93	1.00
0.27	1.93	0.25	67	1.00
0.35	2.03	0.31	48	1.00
0.78	3.95	0.37	159	1.00
3.0	10.4	1.5	56	1.00
11	37.4	7.4	27	1.00
44	72.3	28	7	1.00

[a] The listed counts were corrected for completeness and bias (C_{compl}), as well as false detection fractions ($F_{false-det}$), by multiplying the raw counts by the correction factor given in the last column, and equal to $(1-F_{false-det})/C_{compl}$ (see Tables 3.2 and 3.3).
[b] The source count errors take into account only the Poissonian errors and completeness and bias correction uncertainties (see text for details).

Sect. 3.5). We find that sample variance that is quantified in this way can introduce a (1σ) scatter of $^{+0.1}_{-0.2}$ dex in the source counts. The distribution in counts in the 100 subfields are likely to be an underestimate of the true cosmic variance, which is dominated by cosmic large-scale structure, rather than sample variance, because these fields are likely not fully independent from each other. Thus, cosmic variance may explain the observed discrepancy.

3.6.3   Comparison with 1.4 GHz counts from the literature

To compare our result with more abundant 1.4 GHz observations and models (e.g., Condon 1984; Bondi et al. 2008; Owen & Morrison 2008; Wilman et al. 2008; de Zotti et al. 2010; Condon et al. 2012) we scale our flux densities to the 1.4 GHz observed frame using a spectral index of -0.7. This value, which is also in agreement with the spectral index survival analysis described in Sect. 3.4, is commonly used and provides the easiest comparison (e.g., Condon et al. 2012). We show the 1.4 GHz source count comparison in the bottom panel of Fig. 3.18.

The large spread of the 1.4 GHz source counts available in the literature at submillijansky levels (see, e.g., Fig. 1 in Smolčić et al. 2015) is usually attributed to a combination of i) cosmic variance as often the observed fields are rather small (see Fig. 3.1 and middle panel of Fig. 3.18), and ii) resolution bias leading to a loss of sources in radio continuum surveys conducted at intermediate to high ($\lesssim 2''$) angular resolution (as described in more detail in Sect. 3.5.1). The large, two square degree area of the COSMOS field minimizes the effect of cosmic variance, and in Sect. 3.5 we performed extensive Monte Carlo simulations to account for potential resolution biases. Our source counts agree well with those derived by Condon et al. (2012) based on the $P(D)$ analysis at the faint end (see previous section).

The counts derived here are in good agreement with those derived from the VLA-COSMOS 1.4 GHz Large Project (Schinnerer et al. 2007; Bondi et al. 2008; red diamonds in Fig. 3.18) at flux densities higher than ~ 200 μJy, but are slightly lower in the flux density range of 100 – 200 μJy. As this is the same field, cosmic variance cannot explain the discrepancy. The uncorrected counts from the two surveys are in very good agreement; the difference is the largest in the flux density range where the 1.4 GHz survey is the least complete (about 60%), and the corrections, thus, are the largest. In the same flux density range the corrections for the 3 GHz survey are not as severe given the much higher sensitivity of the 3 GHz survey. Further reasons that could explain part of the discrepancy are i) the effect of BWS on the radio source count corrections that are present in the 1.4 GHz data, but are not present in the 3 GHz data (see Bondi et al. 2008), and ii) a possibly overly simplistic scaling of the 3 GHz counts to 1.4 GHz using just one spectral index value. Source counts at 1.4 GHz depend on the steepness of the counts at 3 GHz and the spread of the spectral indices.

The largest discrepancy between the counts derived here and those in the literature is observed relative to the Owen & Morrison (2008) results. Owen & Morrison (2008) have observed the Lockman hole at 1.4 GHz in A-, B-, C-, and D-array configurations with the VLA reaching an angular resolution of $\sim 1.6''$ and $rms \approx 2.7\,\mu$Jy beam^{-1}. To correct for the resolution bias, they assumed a source size distribution with an extended tail at the high end (see their Fig. 8) and that distribution remains constant as a function of flux density. The source count corrections are significant under these assumptions and result in a flat source count distribution at flux densities fainter than $\sim 200\,\mu$Jy (green points in the bottom panel of Fig. 3.18). As already discussed by Condon et al. (2012) and Vernstrom et al. (2014), these corrections are most likely overestimated. In contrast, for the corrections applied to the data presented here we assumed a model for radio source sizes such that the radio size is a function of flux density with a limiting minimum size (see Sect. 3.5.1). The agreement between our source counts and those derived from confusion-limited data (Condon et al., 2012; Vernstrom et al., 2014) further strengthens the validity of this assumption.

In Fig. 3.18 we also compare our results with the models developed by Condon (1984), Wilman et al. (2008), and Béthermin et al. (2012). The faint end of our counts ($\lesssim 80\,\mu$Jy), combined with the results from Condon et al. (2012) that appear as an extrapolation of our data, agree the best with the Condon (1984) model. The model was constrained by source counts, redshift, and spectral-index distributions for various 400 MHz to 5 GHz flux-limited samples as well as the local 1.4 GHz luminosity function for two dominant, spiral and elliptical galaxy populations. The adopted model is not a unique solution and evolves all sources, i.e., ellipticals and spirals, steep- and flat-spectrum sources, in the same way. At flux densities above $\sim 80\,\mu$Jy the Condon (1984) model is slightly higher than our derived source counts and is consistent with the counts determined by Vernstrom et al. (2014).

Our derived source counts deviate from those predicted by the Wilman et al. (2008) and Béthermin et al. (2012) models. While they agree with the first down to $\sim 100\,\mu$Jy, they are systematically higher at fainter flux densities. On the other hand, the Béthermin et al. (2012) model underpredicts our counts in the flux density range of $\sim 50 - 300\,\mu$Jy, while it overpredicts the counts at flux densities $\lesssim 30\,\mu$Jy. The discrepancies may possibly be understood when considering how AGN and SF galaxies were implemented in the models. Béthermin et al. (2012) implement only models for X-ray selected AGN ($L_{2-10\,\text{keV}} \sim 10^{42} - 10^{44}$ erg s^{-1}; see Mullaney et al. 2011, 2012; Aird et al. 2012), and thus ignore the population of radio-loud AGN hosted by red, quiescent galaxies, regularly not identified as X-ray AGN, yet still substantial (e.g., Best et al. 2006; Smolčić et al. 2008; Bonzini et al. 2013; Smolčić et al. 2017b; Delvecchio et al. 2017). This could explain the lack of sources with flux densities in the range of $\sim$50 –

300 μJy in the model (see, e.g., Smolčić et al. 2008; Padovani et al. 2015; Smolčić et al. 2017b), compared to the observational results. On the other hand, Béthermin et al. (2012) model the SF galaxy population using the most recent results from Bouwens et al. 2007; Rodighiero et al. 2011; Magnelli et al. 2011; Karim et al. 2011; Sargent et al. 2012 by tracing the star-forming galaxy main sequence and the stellar mass function over cosmic time, while also taking into account main-sequence and starburst galaxy spectral energy distribution libraries. Therefore, the excess of Béthermin et al. (2012) model compared to that of Wilman et al. (2008) could suggest that the Wilman model carries potential for improvement in modeling the star-forming galaxy population.

3.7 Chapter summary

We presented the VLA-COSMOS 3 GHz Large Project based on 384 hours of observations with the Karl G. Jansky Very Large Array at 3 GHz (10 cm) toward the two square degree COSMOS field. Our final mosaic, imaged per pointing with the multiscale multifrequency algorithm and self-calibration, reaches a median *rms* of 2.3 μJy beam^{-1} over the two square degrees, at an angular resolution of $0.75''$. We further presented a catalog of 10 830 radio sources. Combining our data with the 1.4 GHz VLA-COSMOS Joint Project data using survival analysis, we found the expected median spectral index α of -0.7. Comparing the positions of our 3 GHz sources with those from the high-resolution VLBA imaging at 1.4 GHz, we estimated that the astrometry is accurate to $0.01''$ at the bright end. Radio source count corrections were calculated for the central two square degrees and used to infer radio source counts. The radio angular size parametrization adopted based on the comparison of mock versus real source total over peak flux density ratios suggests that the angular sizes of radio sources at these flux density levels can be modeled as a power law in flux density ($\theta \propto S^n$) with a minimal, flux-dependent size cutoff (eq. 3.4). Our corrected radio counts with direct detections down to 20 μJy (at 1.4 GHz) are consistent with those derived based on P(D) analyses (Condon et al., 2012), and agree best with the Condon (1984) model, while they are systematically higher than those predicted by the Square Kilometer Array Design Studies (SKADS) simulations (Wilman et al., 2008).

Table 3.5: Catalog sample page

ID	NAME	RA (J2000, deg)	RA_ERR (arcsec)	DEC (J2000, deg)	DEC_ERR (arcsec)	FLUX (μJy)	FLUX_ERR (μJy)	RMS (μJy beam^{-1})	SNR	NPIX	RES	MULTI
78	COSMOSVLA3 J095709.33+020940.7	149.288886	0.01	2.161331	0.01	13400.0	670.0	28.7	385.0	104	1	0
1110	COSMOSVLA3 J095709.83+015457.4	149.290996	0.021	1.915946	0.021	1190.0	64.0	22.1	29.0	102	1	0
5144	COSMOSVLA3 J095710.49+013644.7	149.29372	0.065	1.612418	0.065	145.0	19.0	17.4	8.32	20	0	0
3749	COSMOSVLA3 J095710.53+025132.2	149.29389	0.053	2.858966	0.053	345.0	32.0	26.8	10.4	31	1	0
10099	COSMOSVLA3 J095710.57+022657.0	149.294072	0.1	2.449174	0.1	92.0	18.0	17.4	5.3	11	0	0
4979	COSMOSVLA3 J095710.88+020929.8	149.295339	0.076	2.158297	0.076	2820.0	140.0	21.4	7.13	332	1	0
3366	COSMOSVLA3 J095711.09+023031.3	149.296231	0.048	2.508722	0.048	318.0	27.0	21.5	11.4	35	1	0
8753	COSMOSVLA3 J095711.15+021104.2	149.296459	0.095	2.184514	0.095	121.0	22.0	21.2	5.68	18	0	0
4046	COSMOSVLA3 J095711.64+021236.8	149.298529	0.055	2.210248	0.055	223.0	26.0	22.6	9.83	22	0	0
6546	COSMOSVLA3 J095711.67+021401.3	149.298655	0.077	2.233706	0.077	135.0	21.0	19.3	7.02	18	0	0
...												
10942	COSMOSVLA3 J100034.76+014635.7	150.144846	-99.0	1.776607	-99.0	374.0	-99.0	2.4	-99.0	828	1	1
9229	COSMOSVLA3 J100034.78+025027.4	150.144947	0.1	2.840947	0.1	25.6	2.7	2.4	5.18	25	1	0
8777	COSMOSVLA3 J100034.80+021421.1	150.145009	0.095	2.239195	0.095	13.0	2.4	2.27	5.7	14	0	0
489	COSMOSVLA3 J100034.81+025515.6	150.145042	0.013	2.921023	0.013	229.0	13.0	3.66	62.5	49	0	0
439	COSMOSVLA3 J100034.83+014247.1	150.145157	0.013	1.713091	0.013	162.0	9.0	2.27	71.3	51	0	0
2850	COSMOSVLA3 J100034.94+020234.9	150.145599	0.042	2.043054	0.042	29.9	2.8	2.29	13.1	29	0	0
2439	COSMOSVLA3 J100034.99+024524.5	150.145815	0.038	2.756821	0.038	32.7	2.8	2.22	14.7	29	0	0
2837	COSMOSVLA3 J100035.00+024614.3	150.145836	0.041	2.770643	0.041	39.6	2.9	2.18	13.3	41	1	0
4092	COSMOSVLA3 J100035.05+024154.6	150.146065	0.055	2.69852	0.055	28.3	2.6	2.23	9.93	31	1	0
4148	COSMOSVLA3 J100035.07+020350.5	150.14614	0.057	2.064044	0.057	23.4	2.7	2.43	9.64	23	0	0
...												
5351	COSMOSVLA3 J100345.81+015420.5	150.940886	0.067	1.905719	0.067	117.0	16.0	14.6	8.05	18	0	0
5225	COSMOSVLA3 J100346.48+023458.3	150.943703	0.067	2.582882	0.067	127.0	17.0	15.5	8.14	21	0	0
8447	COSMOSVLA3 J100346.52+022031.2	150.943867	0.11	2.342001	0.11	290.0	20.0	14.4	5.04	48	1	0
10677	COSMOSVLA3 J100346.56+015500.1	150.944027	0.11	1.916707	0.11	75.7	15.0	14.9	5.1	11	0	0
6533	COSMOSVLA3 J100346.63+022415.8	150.944306	0.076	2.404402	0.076	143.0	16.0	14.2	7.09	28	1	0
2195	COSMOSVLA3 J100347.12+022510.6	150.946345	0.036	2.419631	0.036	548.0	31.0	15.3	15.6	83	1	0
10084	COSMOSVLA3 J100347.27+020117.6	150.94699	0.1	2.021575	0.1	73.1	15.0	14.1	5.18	16	0	0
9695	COSMOSVLA3 J100347.51+023539.7	150.947979	0.1	2.594384	0.1	93.2	18.0	17.6	5.3	10	0	0
10652	COSMOSVLA3 J100347.80+024951.9	150.949175	0.11	2.831096	0.11	80.7	17.0	16.1	5.01	14	0	0
10484	COSMOSVLA3 J100348.53+021102.3	150.952242	0.11	2.183987	0.11	82.3	17.0	16.3	5.05	13	0	0

Chapter 4

Radio luminosity functions of star-forming galaxies and active galactic nuclei across cosmic times

The catalog of 3 GHz radio sources is paired with ancillary data from the rich multiwavelength coverage of COSMOS, enabling precise determination of redshifts and spectral energy distributions. This sample, comprised of almost 8 000 galaxies, is studied in a statistical manner by estimating radio luminosity functions. They represent the densities of galaxies at various luminosities and cosmic times and can provide insights into evolutionary trends of galaxy populations. From luminosity functions it is possible to obtain number counts as a function of flux densities providing constraints on the amount of galaxies that can be observed in a flux limited survey.

4.1 Data and galaxy samples

4.1.1 Radio data

The radio data were obtained with 384 hours of VLA A+C array observations in the S band (2 GHz bandwidth centered around 3 GHz) within the VLA-COSMOS 3 GHz Large Project survey. Details of the observational setup, calibration, imaging, and source extraction can be found in Chapter 3. Briefly, 192 pointings were used to obtain a map of the COSMOS 2 square degrees with a uniform *rms* noise equal to $2.3\,\mu\mathrm{Jy\,beam}^{-1}$ and an angular resolution of $0.75''$. Imaging was performed using the multiscale multifrequency synthesis (Rau & Cornwell, 2011) to ensure good deconvolution of both unresolved and extended sources using the entire avail-

able 2 GHz bandwidth at once. Self-calibration of pointings containing brighter sources was performed to improve the fidelity of the image. A catalog of source components with a S/N greater than 5 was extracted using the BLOBCAT software (Hales et al., 2012), which relies on a flood fill algorithm to detect contiguous blobs of emission. After visual inspection of multi-component sources, a final catalog of 10 830 radio sources was assembled, spanning the entire observed area of 2.6 square degrees (approximately 10 000 radio sources across the central COSMOS two square degrees). The astrometric accuracy is $0.01''$ at the bright end and around $0.1''$ for the faintest sources.

4.1.2 Optical and near-infrared counterparts

We use the auxiliary data from more than 30 bands in the optical, NIR, and near-ultraviolet (NUV) available in the COSMOS field from Ultra Visible and Infrared Survey Telescope for Astronomy (UltraVISTA) data release 2 (DR2), Subaru/Hyper-Suprime-Cam, and *Spitzer* Large Area Survey with Hyper-Suprime-Cam (SPLASH) legacy program collected in the multiwavelength COSMOS2015 catalog (Laigle et al., 2016). The catalog contains $\sim 800\,000$ sources with reliable photometry across an area of 1.77 deg^2 free of stellar contamination. Photometric redshifts were computed for all sources by SED fitting using the LePhare code (Arnouts et al., 1999; Ilbert et al., 2006) following methods described in Ilbert et al. (2013).

The counterpart matching method is fully described in Smolčić et al. (2017b), their Sect. 3, and briefly summarized below. Owing to high sub-arcsecond resolution of both optical and radio data, and the fact that the radio emission is usually linked with massive bright galaxies, a nearest-neighbor counterpart matching scheme was adopted in combination with a false match probability assignment using a well-constructed background model. Optical-NIR counterparts were assigned to radio sources within a $0.8''$ searching radius if they were deemed reliable. Estimates of false match probabilities were drawn from simulations using a background model that takes the $m_{3.6\,\mu m}$ magnitude distribution of radio counterparts into account. It was designed to consider the optical blocking effect, i.e., missing fainter optical-NIR sources in the COSMOS2015 catalog due to nearby presence of a bright radio counterpart. Given these choices, the percentage of spurious matches in the entire radio sample are negligible.

Out of 8 696 radio sources positioned in the unmasked optical-NIR area, 7 729 sources were assigned a COSMOS2015 counterpart. If additional counterpart candidates are considered from the *i*-band selected catalog (Capak et al., 2007), the sample increases to 7 826 galaxies, and if *Spitzer* Infrared Array Camera (IRAC) catalog (Sanders et al., 2007) is considered as well, the number of galaxies in the sample becomes 8 035 (see also Smolčić et al., 2017b). However, due to fewer photometry points, we do not use sources assigned only the IRAC counterpart, because

a reliable photometric redshift estimate is not available.

Approximately 11% (7 729/8 696) of radio sources positioned in the unmasked optical-NIR area was not assigned a COSMOS2015 counterpart. Half of those have $S/N < 6$ in the radio source catalog making them likely candidates for spurious sources. The false detection probability for radio sources can reach up to 24% for sources with $5 < S/N < 5.1$, and a total of $\sim 3\%$ of sources in the radio catalog can be considered spurious (see also Sect. 3.5.2). When discussing galaxy subpopulations, we limit the optical-NIR counterpart matching to the COSMOS2015 catalog for better consistency with work by Delvecchio et al. (2017) and Delhaize et al. (2017) where important calibrations used in this work were presented. By taking the fraction of spurious sources into account, we have an average $\sim 8\%$ incompleteness in our COSMOS2015 counterpart sample (reduced to 7% if *i*-band counterparts are also considered). It is likely that the redshift distribution of these unaccounted sources is not uniform across all redshifts, but skewed toward high redshift. This would imply somewhat larger densities of sources at high redshifts at a given luminosity than what the data suggests. However, in the absence of actual redshift distribution of these sources detected only in the radio we do not attempt to correct for this bias.

We use spectroscopic redshifts z_s from the internal COSMOS catalog (M. Salvato et al. in prep.) available for 35% of our radio sources. These redshifts were used only if the spectra were flagged as reliable and 90% of those are located at $z < 1.5$. Photometric redshifts z_p were used for the remainder of the sample. We estimate the accuracy of photometric redshifts of our radio sample by comparing them to the above mentioned spectroscopic catalog and find a median $\Delta z/(1+z_s) = 0.01$ at all redshifts, where $\Delta z = |z_p - z_s|$, and a 4% catastrophic failure rate, defined as $\Delta z/(1+z_s) > 0.15$. At redshifts $z > 1.5$ we find a slightly larger median $\Delta z/(1+z_s) = 0.04$ and a catastrophic failure rate of 12%. Laigle et al. (2016) report the photometric redshift normalized median absolute deviation of the entire COSMOS2015 catalog in $3 < z < 6$ to be $\sigma_{\Delta z(1+z_s)} = 0.021$ with a catastrophic failure rate of 13.2% (see their Table 5).

4.1.3   Classifying galaxies based on the radio excess

Unlike IR observations where the photometry can be used to trace a dusty torus in AGN (e.g., Donley et al., 2012), radio emission linked to star formation and AGN cannot be disentangled without assuming some correlation with emission at other wavelengths. In order to quantify AGN contribution in each galaxy of our sample, Delvecchio et al. (2017) in their Sect. 3 performed a three-component SED fit using the SED3FIT code[1] (Berta et al., 2013). These fits

---

[1]Publicly available at http://cosmos.astro.caltech.edu/page/other-tools
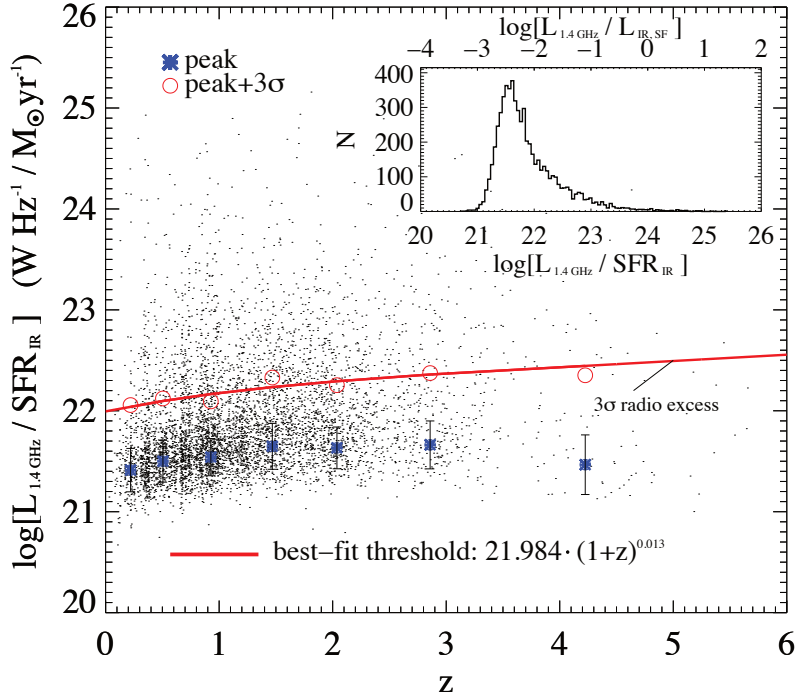
Figure 4.1: Distribution of ratio between radio luminosity $L_{1.4\,GHz}$ and IR-based star formation rate SFR_{IR} as a function of redshift (reproduced from Delvecchio et al., 2017). The inset panel shows the histogram across all redshifts.

were performed on the COSMOS2015 photometry using the best redshifts available and take the energy balance between the UV light absorbed by the dust and re-emitted in the IR into account, following the approach adopted in MAGPHYS (da Cunha et al., 2008). In addition, an AGN component including the continuum disk and the dusty torus emission was added to the fit, seeking a best-fit solution via χ^2 minimization (Berta et al., 2013). From these fits Delvecchio et al. (2017) estimated IR luminosity arising only from the star formation processes (constrained by the SED template) and calculated the IR-based ($8 - 1\,000$ μm) star formation rate SFR_{IR} using the Kennicutt (1998) relation and the Chabrier (2003) IMF. These SFR_{IR} values should be correlated with radio luminosities $L_{1.4\,GHz}$ given the existence of the IR-radio correlation. In order to quantify this correlation and find outliers, these investigators construct histograms of $r = \log(L_{1.4\,GHz}/SFR_{IR})$ in different redshift bins and fit a Gaussian distribution to the histogram (see Delvecchio et al. 2017, Sect. 4.2). The distributions of r peak at higher values with increasing redshift, and in each redshift bin they are skewed toward higher values, corresponding to higher radio luminosities (see Fig. 4.1). Delvecchio et al. (2017) define radio-excess sources when r deviates more than 3σ from the obtained peak of the distribution as a function of redshift, i.e.,

$$r = \log\left(\frac{L_{1.4\,\mathrm{GHz}}[\mathrm{W\,Hz}^{-1}]}{\mathrm{SFR}_{\mathrm{IR}}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}]}\right) > 22 \times (1+z)^{0.013}. \tag{4.1}$$

Such a cut enables us to discriminate 1 814 (23%) sources dominated by AGN emission in the radio. If we assume that the peak of the above r distribution at some redshift corresponds to the ideal correlation between the $L_{1.4\,\mathrm{GHz}}$ and the $\mathrm{SFR}_{\mathrm{IR}}$, and all values above the peak are due to the increasing AGN contribution, then we estimate that the above cut corresponds to at least 80% of the radio emission due to the radio AGN component. The choice of this cutoff is somewhat arbitrary and was chosen as a conservative limit by Delvecchio et al. (2017) to minimize contamination of their AGN sample by SF galaxies. Radio emission of galaxies below this 3σ threshold might still be partly contaminated by AGN emission, but not likely dominated by it. Possible biases of this selection criterion are further discussed in Sect. 6.4.2.

We are interested in measuring the amount of star formation in galaxies from radio observations, disregarding whether the galaxy is an AGN host or not. We are therefore not interested in removing all AGN host galaxies from our SF galaxy sample, but only those that show clear evidence of radio emission dominated by an AGN. Radio samples used throughout are as follows:

1. We consider 5 915 radio sources with COSMOS2015 counterparts and without radio excess as our SF galaxy sample studied in Sect. 4.3.

2. The complementary sample of 1 814 galaxies with radio excess constitute the AGN population discussed in Sect. 4.4.

3. For the full radio sample, apart from COSMOS2015 sources, we also include additional *i*-band selected counterparts yielding a total of 7 826 galaxies analyzed in Sect. 4.5.

The redshift, luminosity and stellar mass distributions of the first two samples are shown in Fig. 4.2. The samples reach out to a redshift of $z \sim 5.5$ and comprise of 1.4 GHz radio luminosities in the range of $10^{21} - 10^{27}$ W Hz^{-1} and typical stellar masses within $10^{9.5} - 10^{11.5}$ M$_\odot$. About 3% of the AGN have estimated stellar masses that are lower than $\sim 3 \times 10^9$ M$_\odot$. Although radio selection and X-ray selection preferentially select high-mass galaxies (e.g., Smolčić et al. 2009b; Smolčić 2009), this fraction is very similar to the fraction of low-mass galaxies in samples of X-ray selected AGN (see, for example, Fig. 1 in Bongiorno et al. 2016, where the host galaxy mass function of the X-ray Multi-Mirror Mission (XMM)-selected AGN in COSMOS is analyzed). The classification of these low-mass AGN as radio-excess sources seems to be statistically robust. Differential number counts of SF (no radio excess) and

AGN (radio excess) populations are shown in Fig. 4.3, where it can be seen that AGN population dominates the radio sky at high flux densities, and SF galaxies dominate the faint radio sky.

4.2   Estimating the luminosity function from the data

In this chapter we assume the flat concordance Lambda cold dark matter (ΛCDM) cosmology defined with Hubble constant $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, dark energy density $\Omega_\Lambda = 0.7$, and matter density $\Omega_\mathrm{m} = 0.3$. Radio luminosity functions at different cosmic epochs are used to measure the evolution of radio sources, while also providing constraints on galaxy evolution models. We first discuss methods of determining the LF from our detections, we then show how the data can be fitted with an analytical form.

Throughout this work we assume that radio sources exhibit a radio spectrum described as a simple power law $S_\nu \propto \nu^\alpha$, where S_ν is a monochromatic flux density at frequency ν and α is the spectral index. This leads to the standard radio K correction of $K(z) = (1+z)^{-(1+\alpha)}$. If not explicitly stated otherwise, $\alpha = -0.7$ is assumed. The final expression for the rest-frame radio luminosity L_{ν_1} at frequency ν_1 derived from the observed flux density S_{ν_2} at frequency ν_2, redshift z, and luminosity distance D_L is

$$L_{\nu_1} = \frac{4\pi D_L^2(z)}{(1+z)^{1+\alpha}} \left(\frac{\nu_1}{\nu_2} \right)^\alpha S_{\nu_2}. \tag{4.2}$$

Luminosities calculated at the rest-frame 1.4 GHz as a function of redshift are shown in the middle panel of Fig. 4.2. This frequency is chosen to simplify a comparison of our results with the literature, where 1.4 GHz observations and reported 1.4 GHz rest-frame luminosities are more common. For about $\sim 25\%$ of all radio sources we were able to derive the spectral index between 1.4 GHz (Schinnerer et al., 2010) and 3 GHz, while for the remaining sources we assumed the standard $\alpha = -0.7$, which is a valid median value to be expected for shock-accelerated cosmic ray electrons. In terms of subpopulations, the spectral index could be determined for $\sim 20\%$ of SF and $\sim 50\%$ of AGN galaxies. The latter percentage is higher because the AGN are generally brighter.

To compute the density of sources and subsequently the LF at different cosmic times (i.e., redshift bins), we employed the V_max method (Schmidt, 1968). This method uses the maximum observable volume of each source, while satisfying all selection criteria; it is not dependent on the shape of the LF and therefore reduces the sample and selection biases. The LF, $\Phi(L, z)$, gives the number of radio sources in a comoving volume per logarithm of luminosity and is
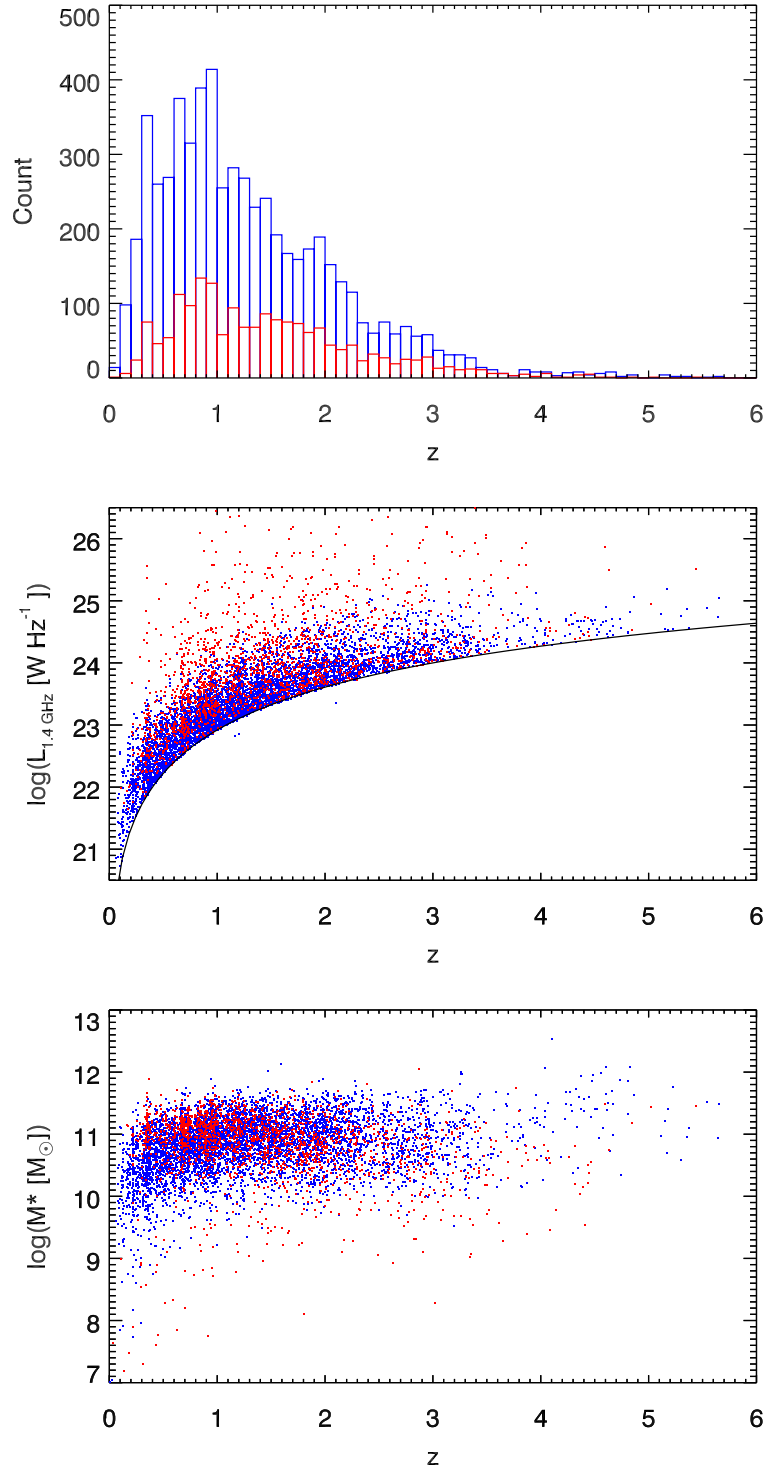
Figure 4.2: Number (*top panel*), rest-frame 1.4 GHz luminosity (*middle panel*) and stellar mass (*bottom panel*) distribution of our SF (blue) and radio-excess ANG (red) galaxies as a function of redshift. The black line indicates the 5σ sensitivity, where $\sigma = 2.3$ μJy beam^{-1} at 3 GHz and fixed $\alpha = -0.7$ is assumed.
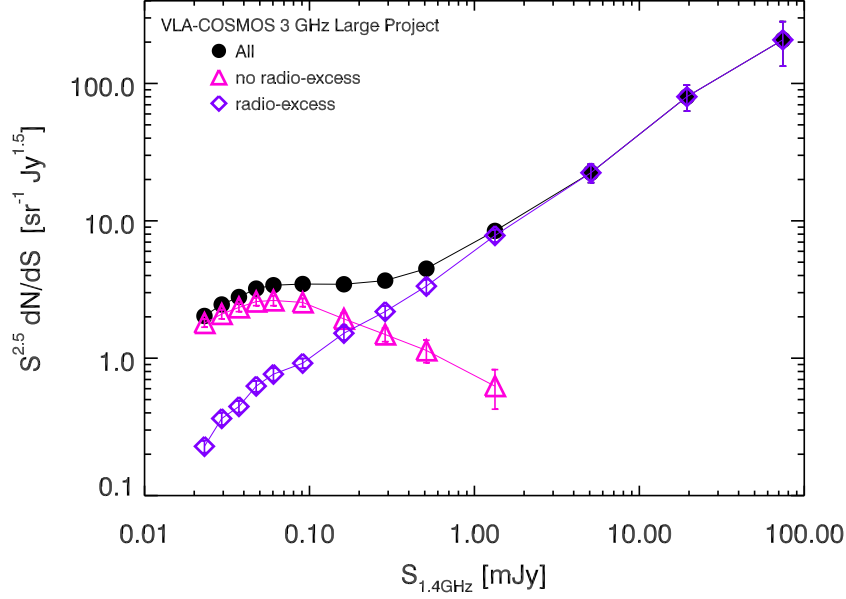
Figure 4.3: Radio differential number counts of the two galaxy populations normalized to Euclidean space.

obtained as

$$\Phi(L,z) = \frac{1}{\Delta \log L} \sum_{i=1}^{N} \frac{1}{V_{\text{max},i}},\tag{4.3}$$

where $V_{\text{max},i}$ is the maximum observable volume of the i-th source, $\Delta \log L$ is the width of the luminosity bin, and the sum goes over each source i in a given redshift and luminosity bin. To take into account different effects and biases, such as a luminosity limited sample or nonuniform noise in the radio map, which may lead to an incompleteness of the sample, we employed a very general form for calculating the maximum observable volume V_{max}, i.e.,

$$V_{\text{max},i} = \sum_{z=z_{\text{min}}}^{z_{\text{max}}} [V(z+\Delta z) - V(z)]C(z),\tag{4.4}$$

where the sum starts at the beginning of a chosen redshift bin and adds together comoving volume spherical shells $\Delta V = V(z+\Delta z) - V(z)$ in small redshift steps $\Delta z = 0.01$ until the end of the redshift bin is reached. The parameter $C(z)$ is the redshift-dependent geometrical and statistical correction factor that takes the observed area and sensitivity limit into account and further mitigates some of the other already mentioned completeness issues

$$C(z) = \frac{A_{\text{obs}}}{41\,253 \text{ deg}^2} \times C_{\text{radio}}[S_{3\text{ GHz}}(z)] \times C_{\text{opt}}(z),\tag{4.5}$$

where $A_{obs} = 1.77$ deg^2 corresponds to the effective unflagged area observed in the optical to NIR wavelengths, C_{radio} is the completeness of the radio catalog as a function of the flux density $S_{3\,GHz}$, and C_{opt} is the completeness owing to radio sources without assigned optical-NIR counterpart. The area observed in the radio encompasses the entire area observed in the optical-NIR and does not have flagged (cropped) regions, therefore NIR observations set the limit for the observed area. The C_{radio} factor depends on the redshift because a source with a given intrinsic luminosity changes its apparent flux density between z_{min} and z_{max} in V_{max} calculations; see Eq. (4.4).

Completeness corrections are shown and tabulated in Sect. 3.5; Fig. 3.16 and Table 3.2, respectively. These completeness corrections are linearly interpolated between tabulated values for any flux density below $S_{3\,GHz} < 100\,\mu$Jy. Simulations were not optimized to probe sources with flux densities above $100\,\mu$Jy and we assume a 100% completeness for such sources. Completeness corrections are based on Monte Carlo simulations of mock-source generation and extraction and take into account the nonuniform *rms*, proper derivation of flux densities for low S/N sources and the resolution bias (out-resolving and losing extended low-surface brightness radio emission). The last part, which was modeled by assuming the distribution of radio sizes, follows some functional form of flux densities, which reproduces the observed data. These corrections are a function of radio flux density only, meaning that all other physical properties are averaged out. For example, the presence of more resolved and extended (and also low-surface brightness) galaxies at lower redshift as a result of their closer proximity to us, which may introduce a redshift-dependent bias.

In Sect. 4.1.2 we mentioned that 11% of our radio sources were not assigned a counterpart. In Fig. 4.4 we show the completeness of our radio catalog C_{opt} due to matching with the COS-MOS2015 catalog as a function of redshift. It was obtained by considering an additional optical catalog selected in the $i-$band (Capak et al., 2007). The counterpart completeness was calculated as $1 - N_{i-band}/N_{i-band\,or\,COSMOS2015}$, where N_{i-band} is the number of new counterparts assigned only to i-band selected sources (1% of the total radio sample) and $N_{i-band\,or\,COSMOS2015}$ is the number of counterparts assigned to either optical catalog. As already mentioned, we only use the COSMOS2015 catalog when individual populations are studied for consistency reasons. Our LFs results are perfectly consistent between themselves whether the actual i-band counterparts or the C_{opt} correction curve is used. As shown in Fig. 4.4, the counterpart sample is complete up to $z \sim 1.5$, and $\sim 90\%$ complete at $z \sim 5.5$. The addition of this completeness correction, while also considering the 3% of spurious radio sources, leaves 7% of real radio sources unaccounted for. A reliable redshift distribution is not available for these sources. If the source follows a strong redshift trend, for example, all of these sources are located at $z > 3$, it
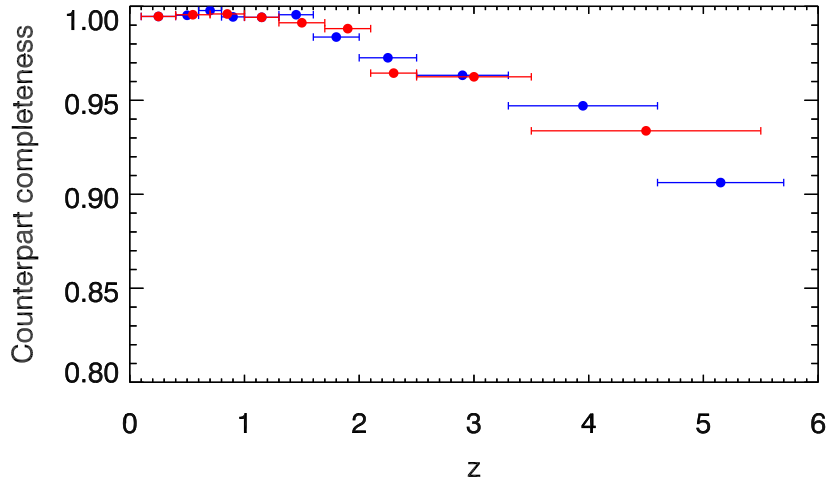
Figure 4.4: Optical-NIR counterpart completeness based on the amount of additional *i*-band sources that could be matched to our radio sources. Blue and red points correspond to SF and AGN galaxies respectively.

still might bias our high-redshift LFs low.

There might exist a small number of galaxies with a high radio flux density and a faint optical-NIR magnitude whose V_{max} would be determined by the optical-NIR limit. Since the optical-NIR catalog was selected on the χ^2 image, it does not have a well-defined magnitude limit and therefore we cannot apply a more precise correction on V_{max}. This may bias our high redshift LF low as the true V_{max} would be smaller than what we have used for such sources. However, we do not expect a significant effect since there are only ~ 10 sources in our sample that have K_s AB magnitudes fainter than 24.5 where the completeness of the optical-NIR catalog becomes an issue (see Laigle et al., 2016).

The *rms* error estimate of the LF in each redshift and luminosity bin is calculated as in Marshall (1985) by weighting each galaxy by its contribution to the sum

$$\sigma_\Phi(L,z) = \frac{1}{\Delta \log L} \sqrt{\sum_{i=1}^{N} \frac{1}{V_{\text{max},i}^2}}. \tag{4.6}$$

However, if there are ten sources or fewer in a luminosity bin we used the tabulated upper and lower 84% confidence intervals from Gehrels (1986). These intervals correspond to Gaussian 1σ errors so that $\sigma_\Phi = \Phi \times \sigma_N/N$, where σ_N is the small-number Poissonian statistical asymmetrical error on the measured number of sources. We do not add photometric uncertainties into the error budget, but the redshift bins are chosen to be large enough to mitigate possible

problems of sources falling into wrong bins. An additional contribution to the total error budget may arise from the imperfect radio SED (discussed in Sect. 4.3.4).

The procedure of binning sources into luminosities inherently introduces some biases due to averaging and the chosen bin sizes. To minimize possible completeness issues at the faint luminosity end within a redshift bin due to poorer sampling, all sources with luminosities below the observational luminosity limit (corresponding to $5\sigma = 11.5\,\mu$Jy at 3 GHz) at z_{max} of the redshift bin were put into single luminosity bin. All sources above this limit were distributed into equally wide luminosity bins spanning the observed luminosity range. The actual luminosity value of each point that we report is the median of all galaxies in a given luminosity bin, while horizontal error bars mark the bin width.

4.2.1　Evolving the radio luminosity function

Radio LFs across cosmic time are usually described by a local LF evolved in luminosity, or density, or both (e.g. Condon, 1984). The local LF, depending on the galaxy population and parametrization choice, is generally described as a power law at low luminosities, a turnover (knee), and another power law, or a log-normal, or an exponential decline for the bright end. We assume that the shape of the LF remains unchanged at all observed cosmic times and allow only the position of the turnover and normalization to change with redshift. This corresponds to the translation of the local LF in the $\log L - \log \Phi$ (luminosity - density) plane and can be divided into pure luminosity evolution (horizontal shift) and pure density evolution (vertical shift). Using a one parameter power law for each of these evolution cases, the form of the redshift evolved LF can be written as

$$\Phi(L, z, \alpha_{\mathrm{D}}, \alpha_{\mathrm{L}}) = (1+z)^{\alpha_{\mathrm{D}}} \Phi_0 \left[\frac{L}{(1+z)^{\alpha_{\mathrm{L}}}} \right], \qquad (4.7)$$

where α_{D} and α_{L} represent pure density evolution (PDE) and pure luminosity evolution (PLE) parameters, respectively, and $\Phi_0(L)$ is the local LF that depends on the galaxy population type and is discussed in the following sections. Since our data are more sensitive to the most luminous SF galaxy population above the knee of the LF, and less luminous AGN below the knee, these two evolution parameters may become degenerate preventing a precise estimate of the knee location, especially at higher redshifts.

The above description can be used for individual redshift bins, but a single parameter estimate might not be able to represent the evolution across larger redshift ranges. In order to create a simple continuous model for the evolution of LFs across the entire observed cosmic time, we can parametrize the redshift dependent evolution with two parameters for density (α_D, β_D) or

two parameters for luminosity evolution (α_L, β_L) to obtain

$$\Phi(L, z, \alpha_L, \beta_L, \alpha_D, \beta_D) = (1+z)^{\alpha_D + z \cdot \beta_D} \times \Phi_0 \left[\frac{L}{(1+z)^{\alpha_L + z \cdot \beta_L}} \right], \qquad (4.8)$$

where $\Phi_0(L)$ is the local LF. The linear form $\alpha + z \cdot \beta$ of the evolution parameter was chosen for its simplicity and its good fit to the data, however other multiparameter descriptions exist in the literature (see e.g. Padovani et al., 2015; Gruppioni et al., 2013). In the absence of data across enough luminosity decades, due to degeneracy of the fitted parameters, we can limit the parameter space by setting either density or luminosity parameter pairs to zero.

4.3 Radio luminosity function of star-forming galaxies

4.3.1 Local luminosity function of star-forming galaxies

Radio LFs of star-forming galaxies are usually described by four parameter analytical forms such as the power law plus log-normal distribution from Saunders et al. (1990)

$$\Phi_0^{\mathrm{SF}}(L) = \Phi_\star \left(\frac{L}{L_\star} \right)^{1-\alpha} \exp\left[-\frac{1}{2\sigma^2} \log^2 \left(1 + \frac{L}{L_\star} \right) \right], \qquad (4.9)$$

where the $L_\star$ parameter describes the position of the turnover of the distribution, $\Phi_\star$ is used for the normalization, α and σ are used to fit the faint and bright ends of the distribution, respectively.

Our deep COSMOS radio observations are best suited to study the high-luminosity end of the LF, especially at higher redshifts $(z > 1)$, where our data do not sample the faint end of the LF, but instead cover a large observed volume. If we are interested in the total amount of light emitted from SF galaxies at any redshift we must assume the shape of the LF that is not constrained by our data. These luminosities can be probed with wide and shallow low-resolution radio surveys of the local universe, such as the NVSS (Condon et al., 1998). There are a number of works related to the calculation of the local radio LF of SF galaxies (e.g., Condon, 1989; Condon et al., 2002; Sadler et al., 2002; Best et al., 2005; Mauch & Sadler, 2007) and they are all broadly consistent in the luminosity range of $21.5 < \log L_{1.4\,\mathrm{GHz}}[\mathrm{W\,Hz}^{-1}] < 23.5$.

To obtain a good estimate of the local luminosity function, we performed a fit on combined volume densities from Condon et al. (2002); Best et al. (2005); Mauch & Sadler (2007) using the form given in Eq. (4.9). By combining the data from both wide and deep surveys we can better constrain both the faint and bright end of the local LF. The data and the fit are shown in Fig. 4.5.
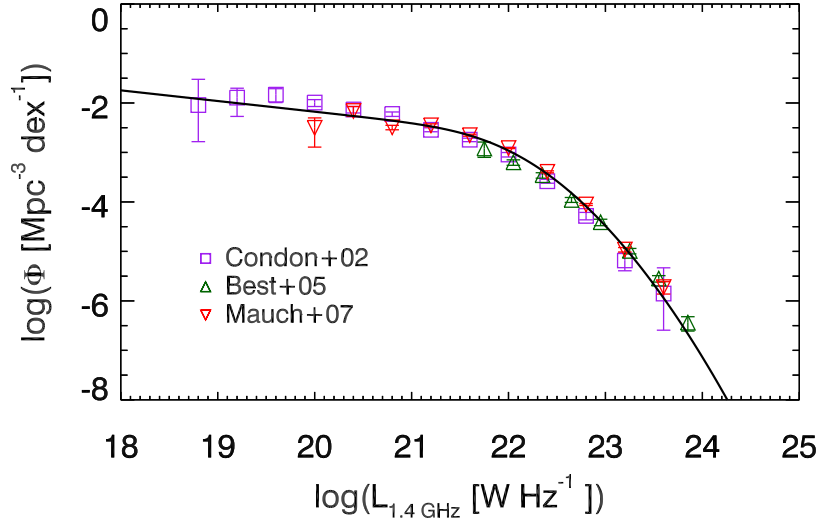
Figure 4.5: Local radio LF of SF galaxies from several surveys with different observed areas and sensitivities (colored data points) and our fit to the combined data (black line).

Obtained best-fit parameters are $\Phi_\star = 3.55 \times 10^{-3}\,\mathrm{Mpc^{-3}dex^{-1}}$, $L_\star = 1.85 \times 10^{21}\,\mathrm{W\,Hz^{-1}}$, and $\alpha = 1.22$, $\sigma = 0.63$.

4.3.2 Evolving the luminosity function of star-forming galaxies

We calculated the LFs of 5 915 star-forming galaxies (the first sample from in Sect. 4.1.3) using the V_{max} method. They are shown in Fig. 4.6 as black circles and tabulated in Table 4.1. Eleven redshift bins were chosen to be large enough to contain a statistically significant number of galaxies and to mitigate possible photometric uncertainties (i.e., sources falling into the wrong bin). Our data have small Poissonian error bars due to the relatively large number of sources in each bin and errors do not reflect all possible systematic effects, such as the imperfectly constrained radio K correction, the error on the completeness, or the sample contamination. A comparison with LFs derived by other authors at different wavelengths is discussed in the next section.

The data points in each redshift bin were first fitted with a two parameter evolved local LF given in Eq. (4.7). The redshift z that enters this expression is the median redshift of all galaxies in a given redshift bin. The best-fit evolution 1σ significant ranges obtained are shown as gray shaded area in Figs. 4.6 and 4.7. A χ^2 minimization was performed to obtain the best fit α_L and α_D parameters and errors on the parameters were estimated from the χ^2 statistics following Avni (1976). We derived the formal 1σ errors by projecting onto each parameter

axis (α_L and α_D) the 68% confidence contour around the minimum χ^2. Since the LF may have asymmetric errors in sparsely populated bins due to small number Poissonian statistics, an average value of the upper and lower errors on the LF was taken for the χ^2 computation. Evolution parameters are degenerate when either the faint or bright ends are not sampled well, yielding large uncertainties at high redshift, therefore a pure luminosity evolution ($\alpha_D = 0$) was computed as well. Pure density evolution would result in too large densities of fainter galaxies, making results inconsistent with known galaxy densities, therefore we do not consider it.

In order to create a single continuous model for the evolution of star-forming LF across the entire observed cosmic time, we simultaneously fit all LF points in all redshift bins with a two parameter pure luminosity evolution ($\alpha_D = 0$, $\beta_D = 0$) described in Eq. (4.8). From the χ^2 minimization fit we obtain the following values for parameters: $\alpha_L = 3.16 \pm 0.04$ and $\beta_L = -0.32 \pm 0.01$, also shown with blue line in Figs. 4.6 and 4.7.

4.3.3 Comparison with the literature

To check the robustness of our LF and also to create a consistent multiwavelength picture, we compare our LFs with work in the literature derived at radio, IR and UV wavelengths.

Radio and IR luminosity functions

In Fig. 4.6 we compare our results with the radio LFs by Smolčić et al. (2009a), which are based on the VLA-COSMOS 1.4 GHz survey (Schinnerer et al., 2007). These investigators constructed LFs up to $z < 1.3$ using a sample of 340 galaxies classified as star forming using optical rest-frame colors. The increase in sensitivity of the VLA-COSMOS 3 GHz survey along with a different selection method yielded ~ 10 times more detections of star-forming galaxies in the same redshift range. The two results generally agree with each other, although our LFs are slightly higher, most likely because of different selection criteria adopted.

We additionally plot LFs from the IR surveys to compare the validity of our results at higher redshifts as well. If the IR-radio correlation is linear, both IR and radio LFs should follow each other well. To convert the total infrared (TIR) to radio luminosity function, the redshift dependent IR-radio correlation q_{TIR} is required. This relation is described in Eq. (6.2) and (6.3) and will be discussed in more detail in the next chapter. Here, we use it as an empirical link that ties flux densities (luminosities) between two different wavelengths. We show the LFs by Magnelli et al. (2013) derived up to $z < 2.3$ from *Herschel* observations of the Great Observatories Origins Deep Survey-North/South (GOODS-N/S) deep and GOODS-S ultra-deep fields. We also show the LFs by Gruppioni et al. (2013), which were computed up to $z < 4.2$ and are based on
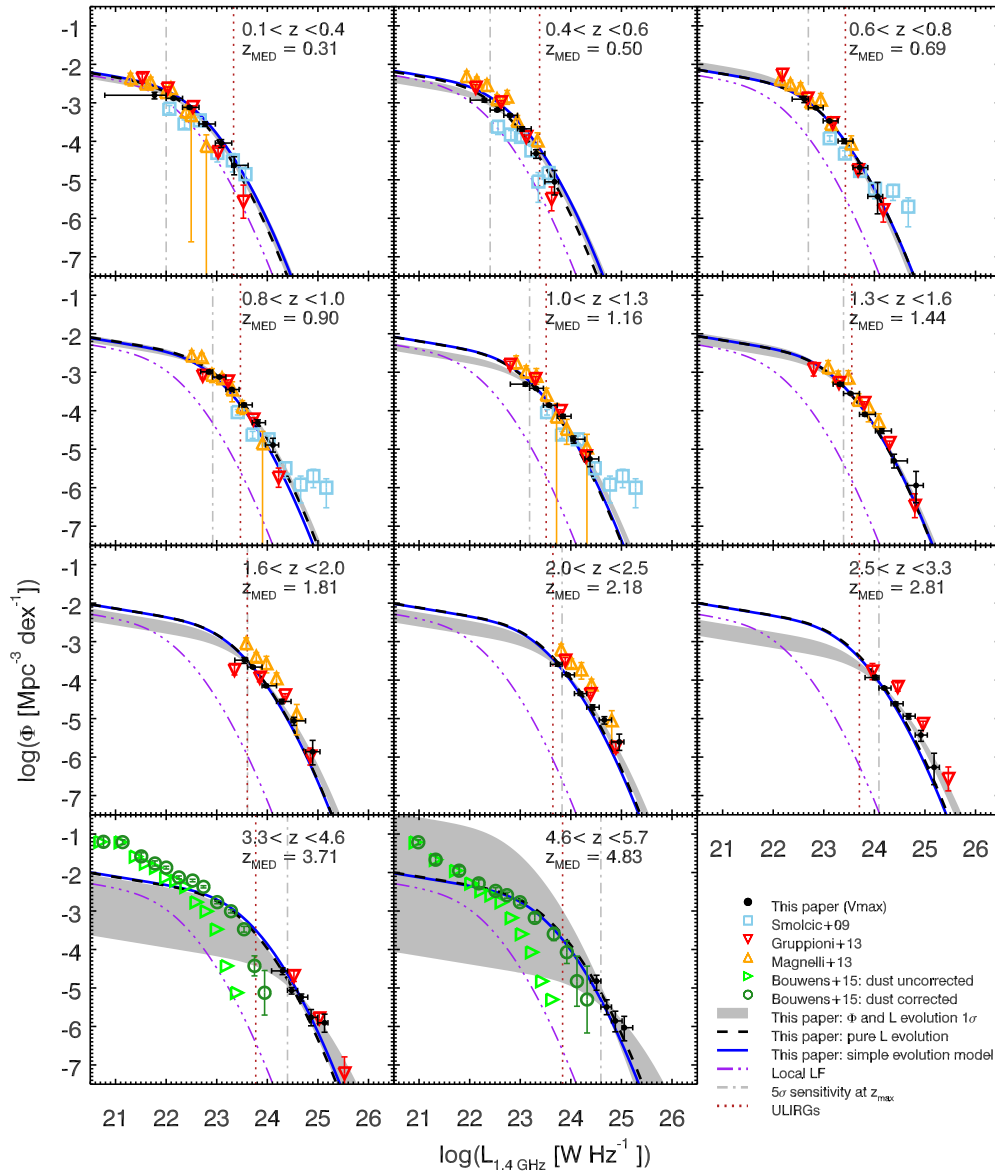
Figure 4.6: Radio luminosity functions of star-forming galaxies in different redshift bins (black filled circles). The best-fit pure luminosity evolved function in each redshift bin is shown with black dashed lines. Combined luminosity and density evolutions are shown by the gray shaded area (using a 68% confidence region in α_D, α_L parameter space around the minimum χ^2). The local radio function is shown for reference as a triple-dot-dashed purple line. The vertical dot-dashed line corresponds to the 5σ luminosity limit at the high redshift end of the bin ($1\sigma = 2.3\,\mu Jy\,beam^{-1}$ at 3 GHz) under the assumption of a fixed spectral index $\alpha = -0.7$. The vertical red dotted line defines the radio luminosity corresponding to ULIRGs under the assumption of a redshift evolving IR-radio correlation. The blue line shows the continuous PLE model described with Eq. (4.8), while all other fits are performed in individual redshift bins. All data shown for comparison are indicated in the legend; see Sect. 4.3.3 for details.

Table 4.1: Luminosity functions of the star-forming radio selected galaxy sample obtained with the V_{max} method. The luminosity $L_{1.4\,GHz}$ has units of $W\,Hz^{-1}$ and the density Φ has units of $Mpc^{-3}\,dex^{-1}$.

z	$\log L_{1.4\,GHz}$	$\log \Phi$	N	z	$\log L_{1.4\,GHz}$	$\log \Phi$	N
$0.1-0.4$	$21.77^{+0.23}_{-0.98}$	$-2.80^{+0.094}_{-0.077}$	189	$1.3-1.6$	$23.32^{+0.070}_{-0.14}$	$-3.31^{+0.043}_{-0.039}$	156
	$22.15^{+0.18}_{-0.15}$	$-2.88^{+0.031}_{-0.029}$	217		$23.53^{+0.18}_{-0.14}$	$-3.55^{+0.025}_{-0.024}$	323
	$22.46^{+0.19}_{-0.14}$	$-3.12^{+0.037}_{-0.034}$	149		$23.81^{+0.21}_{-0.10}$	$-4.10^{+0.041}_{-0.037}$	126
	$22.77^{+0.20}_{-0.12}$	$-3.55^{+0.063}_{-0.055}$	56		$24.15^{+0.19}_{-0.12}$	$-4.53^{+0.068}_{-0.059}$	48
	$23.09^{+0.21}_{-0.12}$	$-4.05^{+0.11}_{-0.090}$	19		$24.39^{+0.26}_{-0.053}$	$-5.30^{+0.18}_{-0.17}$	8
	$23.34^{+0.28}_{-0.048}$	$-4.63^{+0.25}_{-0.22}$	5		$24.82^{+0.14}_{-0.17}$	$-5.94^{+0.45}_{-0.37}$	2
$0.4-0.6$	$22.29^{+0.11}_{-0.28}$	$-2.93^{+0.052}_{-0.046}$	142	$1.6-2.0$	$23.55^{+0.065}_{-0.19}$	$-3.48^{+0.084}_{-0.070}$	156
	$22.54^{+0.13}_{-0.14}$	$-3.19^{+0.036}_{-0.033}$	160		$23.72^{+0.18}_{-0.11}$	$-3.66^{+0.025}_{-0.024}$	312
	$22.80^{+0.15}_{-0.12}$	$-3.33^{+0.038}_{-0.035}$	143		$23.98^{+0.21}_{-0.080}$	$-4.15^{+0.038}_{-0.035}$	141
	$23.04^{+0.18}_{-0.090}$	$-3.67^{+0.059}_{-0.052}$	65		$24.28^{+0.18}_{-0.10}$	$-4.56^{+0.062}_{-0.054}$	57
	$23.31^{+0.18}_{-0.096}$	$-4.32^{+0.12}_{-0.097}$	16		$24.53^{+0.23}_{-0.059}$	$-5.06^{+0.12}_{-0.092}$	18
	$23.68^{+0.081}_{-0.19}$	$-5.05^{+0.34}_{-0.30}$	3		$24.90^{+0.14}_{-0.14}$	$-5.86^{+0.34}_{-0.30}$	3
$0.6-0.8$	$22.61^{+0.080}_{-0.24}$	$-2.90^{+0.091}_{-0.075}$	179	$2.0-2.5$	$23.74^{+0.086}_{-0.14}$	$-3.59^{+0.050}_{-0.045}$	141
	$22.84^{+0.15}_{-0.15}$	$-3.13^{+0.027}_{-0.025}$	283		$23.94^{+0.13}_{-0.11}$	$-3.86^{+0.031}_{-0.029}$	219
	$23.11^{+0.17}_{-0.12}$	$-3.47^{+0.035}_{-0.033}$	165		$24.19^{+0.13}_{-0.12}$	$-4.35^{+0.046}_{-0.042}$	98
	$23.40^{+0.17}_{-0.12}$	$-3.99^{+0.066}_{-0.057}$	51		$24.43^{+0.13}_{-0.11}$	$-4.71^{+0.072}_{-0.061}$	44
	$23.71^{+0.16}_{-0.13}$	$-4.68^{+0.16}_{-0.11}$	11		$24.66^{+0.14}_{-0.10}$	$-5.04^{+0.11}_{-0.086}$	21
	$24.06^{+0.10}_{-0.19}$	$-5.43^{+0.45}_{-0.37}$	2		$24.96^{+0.092}_{-0.15}$	$-5.61^{+0.22}_{-0.20}$	6
$0.8-1.0$	$22.85^{+0.074}_{-0.16}$	$-2.99^{+0.046}_{-0.041}$	172	$2.5-3.3$	$24.01^{+0.079}_{-0.19}$	$-3.93^{+0.057}_{-0.051}$	128
	$23.05^{+0.13}_{-0.13}$	$-3.13^{+0.025}_{-0.024}$	312		$24.20^{+0.13}_{-0.11}$	$-4.21^{+0.037}_{-0.034}$	155
	$23.30^{+0.14}_{-0.12}$	$-3.45^{+0.032}_{-0.030}$	198		$24.42^{+0.15}_{-0.091}$	$-4.62^{+0.051}_{-0.046}$	81
	$23.54^{+0.16}_{-0.099}$	$-3.85^{+0.051}_{-0.046}$	82		$24.68^{+0.13}_{-0.11}$	$-4.94^{+0.076}_{-0.065}$	39
	$23.81^{+0.15}_{-0.11}$	$-4.31^{+0.088}_{-0.073}$	30		$24.92^{+0.13}_{-0.11}$	$-5.43^{+0.16}_{-0.12}$	11
	$24.11^{+0.11}_{-0.15}$	$-4.89^{+0.18}_{-0.17}$	8		$25.18^{+0.10}_{-0.13}$	$-6.27^{+0.45}_{-0.37}$	2
$1.0-1.3$	$23.10^{+0.081}_{-0.30}$	$-3.31^{+0.051}_{-0.046}$	218	$3.3-4.6$	$24.30^{+0.097}_{-0.22}$	$-4.56^{+0.096}_{-0.079}$	55
	$23.31^{+0.15}_{-0.12}$	$-3.42^{+0.025}_{-0.024}$	321		$24.48^{+0.12}_{-0.081}$	$-5.07^{+0.093}_{-0.077}$	27
	$23.57^{+0.16}_{-0.12}$	$-3.86^{+0.036}_{-0.034}$	156		$24.67^{+0.13}_{-0.074}$	$-5.24^{+0.10}_{-0.082}$	23
	$23.84^{+0.16}_{-0.11}$	$-4.15^{+0.052}_{-0.046}$	81		$24.86^{+0.13}_{-0.066}$	$-5.76^{+0.22}_{-0.20}$	6
	$24.06^{+0.22}_{-0.051}$	$-4.74^{+0.10}_{-0.084}$	22		$25.13^{+0.067}_{-0.13}$	$-5.91^{+0.25}_{-0.22}$	5
	$24.38^{+0.17}_{-0.10}$	$-5.25^{+0.20}_{-0.19}$	7	$4.6-5.7$	$24.51^{+0.085}_{-0.12}$	$-4.82^{+0.24}_{-0.15}$	11
					$24.71^{+0.095}_{-0.12}$	$-5.50^{+0.20}_{-0.19}$	7
					$24.88^{+0.13}_{-0.076}$	$-5.86^{+0.28}_{-0.25}$	4
					$25.06^{+0.17}_{-0.042}$	$-6.03^{+0.34}_{-0.30}$	3
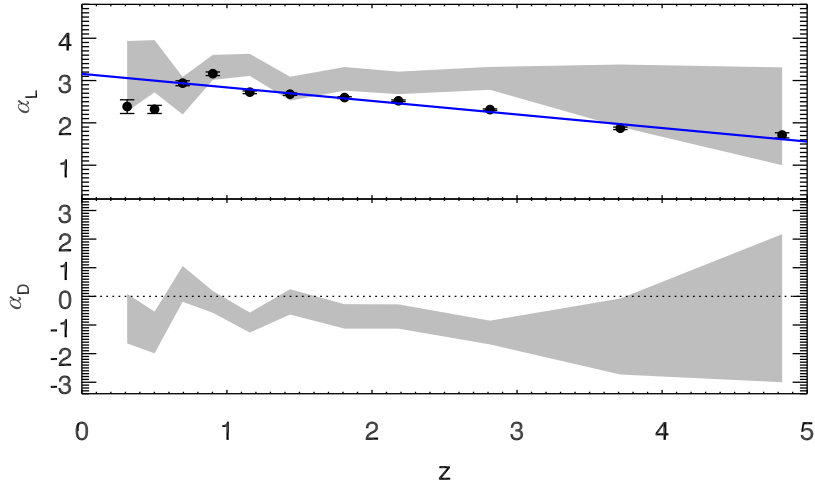
Figure 4.7: Best-fit parameters for the local LF evolution as a function of redshift. Filled black points correspond to a pure luminosity evolution in individual redshift bins ($\alpha_D = 0$). The gray shaded area shows the 68% confidence interval for a combined luminosity and density evolution. The large uncertainty in the combined fit is due to parameter degeneracy. The blue line shows the simple continuous pure luminosity evolution model described with Eq. (4.8) using $\alpha_D = 0$, $\beta_D = 0$.

the *Herschel* PEP/HerMES[2] data. To take into account the fact that the redshift bin ranges do not necessarily coincide, we evolved the LFs of other authors using the evolution parameters they reported in their work, if our median redshift value fell inside their redshift range. Small systematic offsets may arise when the mean redshift does not correspond to the median redshift. Our data agree well with these surveys both at low and intermediate redshifts. However, at redshift $z > 2$ our LFs are systematically slightly lower than those based on IR. Some of this offset might be attributed to a higher percentage of AGN in the IR selected sample at these redshifts (Gruppioni et al., 2013). They may constitute half of the sample above $z > 2.5$. However, since we start from a differently selected radio sample, exclude AGN identified with radio excess compared to IR emission, and we must rely on the redshift evolving $q_{TIR}(z)$, it makes the direct comparison difficult. If a constant $q_{TIR} = 2.64$ (Bell, 2003) is used for the conversion, instead of an evolving one, our observed radio LFs would actually be higher than implied by the observed IR-based LFs at high redshifts.

---

[2]Photoconductor Array Camera and Spectrometer (PACS); PACS Evolutionary Probe (PEP); Multi tiered Extragalactic Survey (HerMES)

UV luminosity functions

Our radio data are good tracers of highly star-forming and dusty galaxies (ULIRGs which have $L_{\mathrm{TIR},\,8-1\,000\,\mu m} > 10^{12}$ $L_\odot$ see Sanders & Mirabel 1996; and hyperluminous infrared galaxies - HyLIRGs with $L_{\mathrm{TIR},\,8-1\,000\,\mu m} > 10^{13}$ $L_\odot$), but lack the sensitivity to probe fainter sources at high redshifts. We make use of the work performed by Bouwens et al. (2015) in an attempt to constrain the faint end of the luminosity functions of SF galaxies with actual detections and to simultaneously test their dust corrections. Bouwens et al. (2015) utilize HST observations of more than ten deep and wide surveys covering $\sim 1\,000$ arcmin^2 to derive the rest-frame UV LFs between $4 < z < 10$ using a sample of more than $10\,000$ Lyman break galaxies (LBGs). The rest-frame UV light correlates strongly with the SFR, unless the galaxy is very dusty. Therefore we can make a broad comparison with our SF galaxy sample.

The SFR calibrations from Kennicutt (1998) are self-consistent, meaning that all tracers (radio, IR, and UV) should provide the same SFR estimate, thus enabling the link between radio and UV luminosities via the SFR. Although this correlation likely has a large scatter when applied to a specific galaxy, if used on larger samples as a statistical conversion factor, it should allow the conversion of UV magnitudes into radio luminosities. The conversion is needed to compare LFs at these two different wavelengths. The expression for this conversion using the Kennicutt (1998) calibration, Chabrier (2003) initial mass function, and the redshift-dependent IR-radio correlation from Eq. (6.3) is

$$\log \frac{L_{1.4\,\mathrm{GHz}}}{\mathrm{W\,Hz}^{-1}} = 16.556 - 0.4(M_{1\,600,AB} - A_{UV}) - q_{\mathrm{TIR}}(z), \qquad (4.10)$$

where $M_{1\,600,AB}$ is the rest-frame UV absolute magnitude reported in Bouwens et al. (2015) and A_{UV} is the extinction needed to calculate the dust-corrected magnitudes. The dust extinction, obtained from the the infrared excess (IRX)-β relationship (correlation between the ratio of FIR to UV fluxes with the UV spectral slope β; see Meurer et al. 1999), is given in the form of $A_{UV} = 4.43 - 1.99\beta$ and tabulated as a function of UV magnitudes in Bouwens et al. (2014a). The bottom panels in Fig. 4.6 show dust-uncorrected (green right triangles) and dust-corrected (dark green circles) LFs from Bouwens et al. (2015) for their $z \sim 4$ and $z \sim 5$ dropouts. Several results can be noted in the plots:

1. Significant dust corrections are needed at high luminosities and the rest-frame UV cannot be used to detect dustier ULIRGs and HyLIRGs, which are easily observed in the radio at high redshifts. Our radio detections, also available across more than $6\times$ larger area, can therefore provide an independent test for these dust corrections.

2. The bright end of the UV LF is lower than our radio LF, with the discrepancy being larger at $z \sim 4$ than at $z \sim 5$. Our result is also broadly consistent with the result of Heinis et al. (2013) in which they perform stacking of *Herschel* images of UV selected galaxies at $z \sim 1.5$. These authors found that the IR luminosity of bright galaxies ($L_{\mathrm{IR}} > 10^{11} \, \mathrm{L_\odot}$) obtained via stacking can recover as low as $52 - 65\%$ of the total star formation rate density derived from the IR-selected samples.

3. Even if we disregard the dust correction, the density of faint galaxies two decades in luminosity below our detection limit is very high. The dust-corrected UV LFs, are in broad agreement with our pure-luminosity fit extrapolation more than two decades below the lowest observed radio luminosity at $z \sim 5$. These arguments can be used to rule out most of the gray-shaded area in the bottom two panels of Fig. 4.6 arising from significant negative density evolution (see also lower panel of Fig. 4.7). Rest-frame UV observations are in favor of higher densities of galaxies than what would be obtained if a turnover in the radio LF is introduced immediately below the faintest observed radio luminosity.

4. There is a discrepancy between our pure luminosity evolution model and the UV LF at the faintest observed end. Since our radio data cannot constrain such low luminosities, pure luminosity evolution is probably not the best possible model for extrapolating our observed LFs below our detection limits. Indeed, a continuous steepening with redshift of the faint end slope of the LF has recently been proposed (e.g., Parsa et al., 2016).

We further discuss the UV data in the context of our radio estimates of star formation rate density in Sect. 6.3.4.

4.3.4 Potential bias from the unconstrained radio spectral indices

The highest uncertainty of the computed rest-frame 1.4 GHz luminosity, especially at high redshifts, lies in the insufficient knowledge of the radio K correction. For example, a rather large photometric error of $\Delta z = 0.3$ would result in a 0.05 dex error in luminosity at $z \sim 5$. However, an uncertainty in spectral index of just $\Delta \alpha = 0.1$ would produce an error of 0.1 dex in luminosity. It is known that a symmetric spread of the spectral indices around the mean value for SF galaxies is ~ 0.4 (e.g., Kimball & Ivezić, 2008). Even though this spread would produce a significant uncertainty on the estimated $L_{1.4\,\mathrm{GHz}}$ of each single galaxy, these variations cancel each other and give a valid average luminosity because of the symmetry of the distribution of spectral indices. It is widespread in the literature to assume a single spectral index for radio SED, where the usual values are $\alpha = -0.8$ or -0.7.
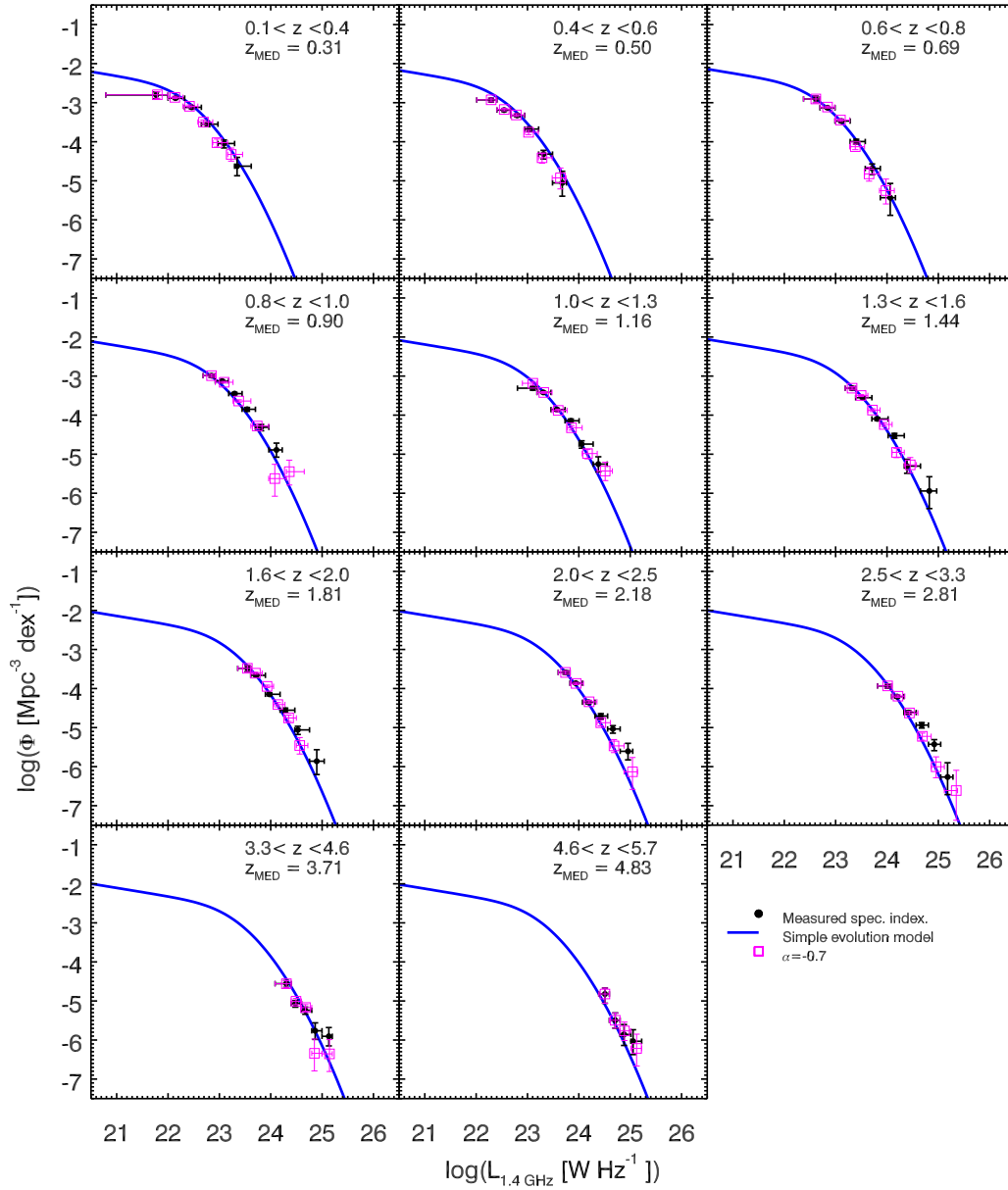
Figure 4.8: Radio luminosity functions of star-forming galaxies, same as Fig. 4.6. Magenta squares indicate recalculated LFs using the same spectral index of $\alpha = -0.7$ for all galaxies.

Approximately 75% of all our radio sources (80% of SF and 50% of AGN) were detected only at 3 GHz. The use of the measured spectral indices for the remaining can introduce a small bias toward steeper spectra (therefore higher luminosities), since our survey is currently the deepest radio survey of the COSMOS field. For example, a point source at the limit of our sensitivity ($rms = 2.3~\mu$Jy beam^{-1}) would have to have a spectral index steeper than -1.9 to be observed in the previous deep 1.4 GHz survey ($rms = 10~\mu$Jy beam^{-1}; Schinnerer et al. 2010). The median spectral index of sources detected in both surveys is $\alpha = -0.85$ (see also Sect. 3.4.2).

To assess the impact of the used spectral indices on our results, we repeated the LF calculations two times: the first time with the standard $\alpha = -0.7$ and the second time with $\alpha = -0.8$ for all sources regardless of the observed radio spectrum. When a single and identical spectral index is used for each source, the pure luminosity evolution of the chosen analytical local function from Eq. 4.9 better fits (smaller χ^2) the derived LFs at all luminosities and redshifts. Specifically, when $\alpha = -0.7$ is used (see Fig. 4.8), the best pure luminosity fit evolution remains essentially unchanged from that presented in Sect. 4.3, which is unsurprising given that 80% of the spectral indices remained unchanged as well. When $\alpha = -0.8$ is used, a stronger luminosity evolution is obtained, which is described by an increase of 0.16 in the evolution parameter α_L as given in Eq. (4.7) and previously shown in Fig. 4.7.

Finally, assuming a simple power law radio spectra might be an overly simplistic approach given the unknown contribution of free-free (Bremsstrahlung) emission to the total SED. Additional deep radio observations at higher frequencies are needed to properly model the radio SED and mitigate this limitation. Further biases are discussed in Sect. 6.4 in the context of cosmic star formation density history.

4.4　Radio luminosity function of AGN

The focus is now switched to the AGN population, which is the second sample described in Sect. 4.1.3, where radio emission in each of the 1 814 galaxies is dominated (>80%) by the central black hole accretion processes. Since the AGN sample is constituted of brighter sources than the SF galaxy population, more galaxies have 1.4 GHz cataloged detections (Schinnerer et al., 2010), yielding spectral index estimates for half of the sample. For the other half, we manually set $\alpha = -0.7$ as before. Spectral indices used in the further analysis are shown in Fig. 4.9. A trend toward steeper spectra can be seen with increasing redshift. Without deep observations at other radio frequencies we cannot state whether this trend is real (e.g., owing to spectral steepening of the radio spectrum at higher rest-frame frequencies, as sampled for our
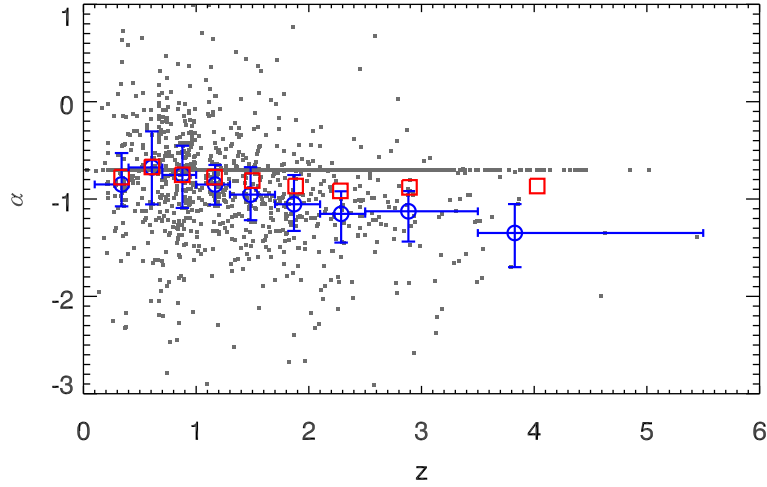
Figure 4.9: Spectral indices of our radio-excess AGN sample as a function of redshift (gray points). Median values with interquartile ranges for sources with measured $1.4 - 3$ GHz spectral indices in different redshift bins are shown with blue circles. Mean values for all sources (i.e., also undetected at 1.4 GHz for which we manually set $\alpha = -0.7$) in same redshift bins are shown with red squares.

high-redshift sources) or a bias due to our flux limited observations as steeper spectra yields higher flux densities at 1.4 GHz than flatter spectra.

4.4.1 Local luminosity function of AGN

The analytic form for the local radio AGN LF, adopted here, and also shown in Fig. 4.10, is taken from Mauch & Sadler (2007) and parametrized with two power laws

$$\Phi_0^{\mathrm{AGN}}(L) = \frac{\Phi^*}{(L^*/L)^\alpha + (L^*/L)^\beta}, \tag{4.11}$$

where the parameters are the normalization $\Phi^* = \frac{1}{0.4}10^{-5.5}$ Mpc^{-3} dex^{-1} (scaled to the base of $d\log L$), the knee position $L^* = 10^{24.59}$ W Hz^{-1}, and the bright and faint end slopes $\alpha = -1.27$, and $\beta = -0.49$, respectively. Mauch & Sadler (2007) derived their AGN LF using $2\,661$ detections in the Six-degree Field Galaxy Survey (6dFGS)–NVSS field with a median redshift of med$(z) = 0.073$ and a span of six decades in luminosities. With such a sample they were able to constrain well both the faint and bright end of the local AGN LF.
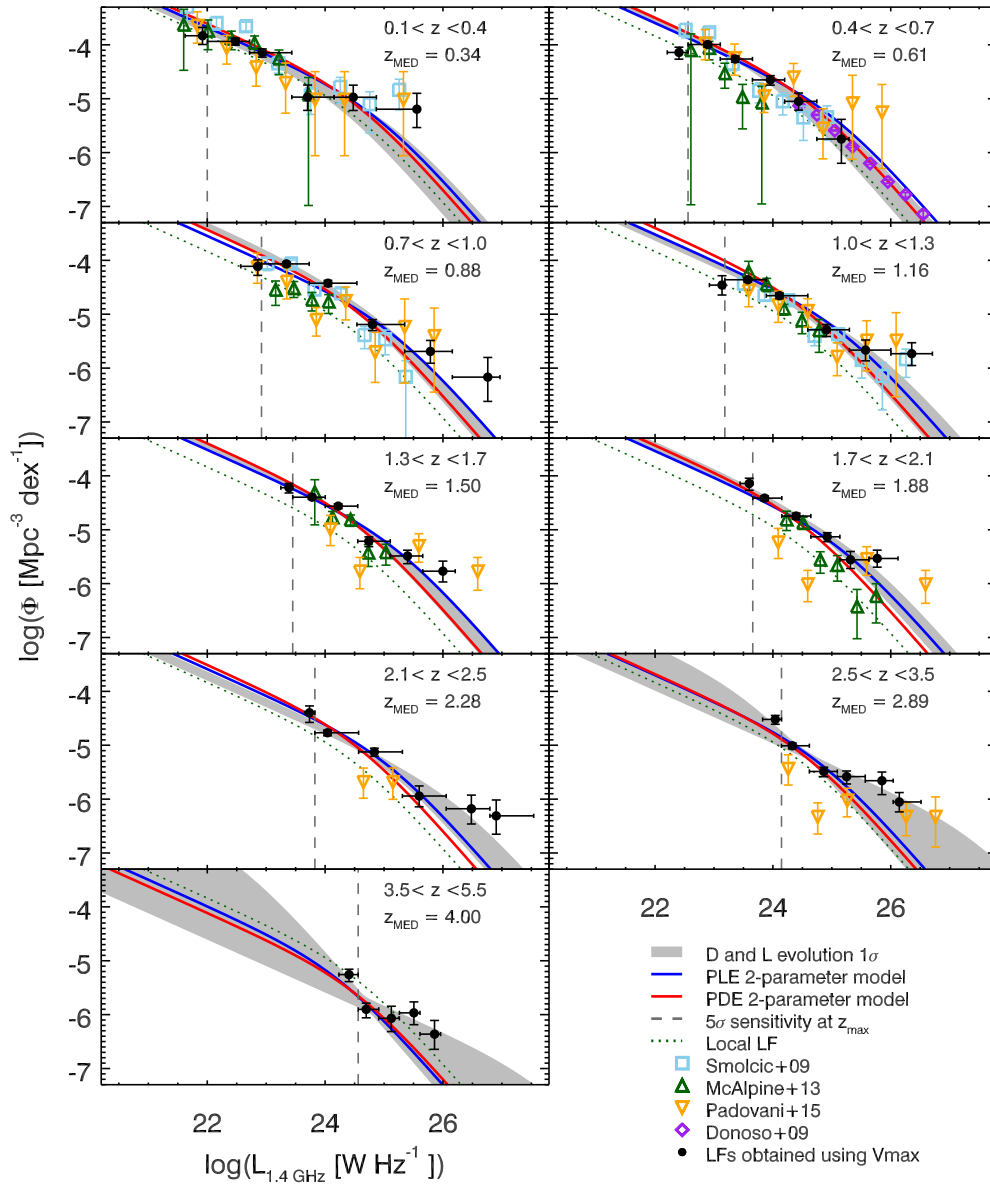
Figure 4.10: Radio AGN 1.4 GHz rest-frame luminosity functions in nine redshift bins out to $z \sim 5$ for radio-excess AGN in the COSMOS field. The volume densities derived from our radio sample with the V_{max} method are shown by the filled black dots. A 1σ confidence interval of the combined luminosity and density evolution in individual redshift bins is shown as the gray shaded area. Pure luminosity evolution and pure density evolution 2-parameter models are shown with blue and red lines, respectively (see the text for details). The dotted curve indicates the local luminosity function (Eq. 4.11). The vertical line in each panel shows the 5σ sensitivity limit (assuming $\alpha = -0.7$) at the high-redshift end of each redshift bin. Various results from the literature, as indicated in the legend, are also shown (each scaled to the median redshift in each bin using the evolution model as reported in the corresponding study).

Table 4.2: Luminosity functions of radio-excess AGN obtained with the $V_{\max}$ method. The listed luminosity values represent the median luminosity of the sources in the corresponding luminosity bin. The luminosity $L_{1.4\,\mathrm{GHz}}$ has units of $\mathrm{W\,Hz^{-1}}$ and the density Φ has units of $\mathrm{Mpc^{-3}\,dex^{-1}}$.

z	$\log L_{1.4\,\mathrm{GHz}}$	$\log \Phi$	N
$0.1-0.4$	$21.92^{+0.081}_{-0.30}$	$-3.83^{+0.16}_{-0.15}$	10
	$22.48^{+0.24}_{-0.48}$	$-3.94^{+0.067}_{-0.058}$	49
	$22.93^{+0.50}_{-0.22}$	$-4.15^{+0.083}_{-0.070}$	33
	$23.70^{+0.45}_{-0.27}$	$-4.97^{+0.25}_{-0.22}$	5
	$24.48^{+0.39}_{-0.33}$	$-4.97^{+0.25}_{-0.22}$	5
	$25.56^{+0.031}_{-0.69}$	$-5.19^{+0.34}_{-0.30}$	3
$0.4-0.7$	$22.41^{+0.15}_{-0.20}$	$-4.14^{+0.13}_{-0.097}$	19
	$22.89^{+0.21}_{-0.33}$	$-3.99^{+0.047}_{-0.042}$	97
	$23.36^{+0.29}_{-0.26}$	$-4.26^{+0.061}_{-0.053}$	59
	$23.96^{+0.24}_{-0.30}$	$-4.65^{+0.097}_{-0.079}$	25
	$24.44^{+0.31}_{-0.24}$	$-5.05^{+0.16}_{-0.15}$	10
	$25.16^{+0.13}_{-0.42}$	$-5.75^{+0.45}_{-0.37}$	2
$0.7-1.0$	$22.86^{+0.064}_{-0.29}$	$-4.11^{+0.17}_{-0.12}$	25
	$23.34^{+0.39}_{-0.42}$	$-4.07^{+0.035}_{-0.032}$	208
	$24.05^{+0.49}_{-0.32}$	$-4.42^{+0.059}_{-0.052}$	98
	$24.80^{+0.55}_{-0.26}$	$-5.19^{+0.11}_{-0.090}$	19
	$25.79^{+0.37}_{-0.44}$	$-5.69^{+0.22}_{-0.20}$	6
	$26.76^{+0.21}_{-0.60}$	$-6.17^{+0.45}_{-0.37}$	2
$1.0-1.3$	$23.14^{+0.048}_{-0.21}$	$-4.46^{+0.18}_{-0.17}$	8
	$23.57^{+0.31}_{-0.39}$	$-4.36^{+0.043}_{-0.039}$	114
	$24.11^{+0.48}_{-0.22}$	$-4.66^{+0.056}_{-0.049}$	69
	$24.91^{+0.39}_{-0.32}$	$-5.29^{+0.13}_{-0.097}$	16
	$25.57^{+0.43}_{-0.27}$	$-5.67^{+0.20}_{-0.19}$	7
	$26.36^{+0.35}_{-0.35}$	$-5.73^{+0.22}_{-0.20}$	6
$1.3-1.7$	$23.39^{+0.063}_{-0.14}$	$-4.22^{+0.097}_{-0.080}$	30
	$23.78^{+0.23}_{-0.33}$	$-4.40^{+0.040}_{-0.037}$	132
	$24.22^{+0.33}_{-0.22}$	$-4.56^{+0.046}_{-0.042}$	101
	$24.74^{+0.36}_{-0.19}$	$-5.21^{+0.10}_{-0.082}$	24
	$25.40^{+0.26}_{-0.29}$	$-5.49^{+0.14}_{-0.11}$	13
	$26.00^{+0.21}_{-0.34}$	$-5.77^{+0.20}_{-0.19}$	7

z	$\log L_{1.4\,\mathrm{GHz}}$	$\log \Phi$	N
$1.7-2.1$	$23.60^{+0.060}_{-0.16}$	$-4.14^{+0.13}_{-0.098}$	30
	$23.86^{+0.29}_{-0.20}$	$-4.41^{+0.045}_{-0.041}$	106
	$24.39^{+0.25}_{-0.24}$	$-4.75^{+0.061}_{-0.054}$	60
	$24.92^{+0.22}_{-0.28}$	$-5.13^{+0.091}_{-0.075}$	28
	$25.32^{+0.31}_{-0.18}$	$-5.56^{+0.16}_{-0.15}$	10
	$25.77^{+0.35}_{-0.14}$	$-5.53^{+0.16}_{-0.15}$	10
$2.1-2.5$	$23.74^{+0.091}_{-0.083}$	$-4.40^{+0.18}_{-0.13}$	13
	$24.04^{+0.53}_{-0.21}$	$-4.77^{+0.055}_{-0.049}$	74
	$24.84^{+0.48}_{-0.27}$	$-5.12^{+0.082}_{-0.069}$	37
	$25.60^{+0.46}_{-0.29}$	$-5.94^{+0.20}_{-0.19}$	7
	$26.48^{+0.32}_{-0.43}$	$-6.18^{+0.28}_{-0.25}$	4
	$26.90^{+0.64}_{-0.11}$	$-6.31^{+0.34}_{-0.30}$	3
$2.5-3.5$	$24.04^{+0.11}_{-0.21}$	$-4.52^{+0.090}_{-0.075}$	53
	$24.33^{+0.29}_{-0.18}$	$-5.01^{+0.057}_{-0.051}$	67
	$24.86^{+0.23}_{-0.24}$	$-5.49^{+0.097}_{-0.079}$	26
	$25.25^{+0.32}_{-0.16}$	$-5.58^{+0.14}_{-0.10}$	18
	$25.85^{+0.19}_{-0.28}$	$-5.66^{+0.26}_{-0.16}$	13
	$26.14^{+0.37}_{-0.10}$	$-6.05^{+0.18}_{-0.17}$	8
$3.5-5.5$	$24.41^{+0.16}_{-0.17}$	$-5.26^{+0.13}_{-0.10}$	18
	$24.70^{+0.21}_{-0.14}$	$-5.90^{+0.16}_{-0.11}$	11
	$25.12^{+0.14}_{-0.21}$	$-6.07^{+0.25}_{-0.22}$	5
	$25.51^{+0.10}_{-0.25}$	$-5.97^{+0.22}_{-0.20}$	6
	$25.86^{+0.10}_{-0.25}$	$-6.36^{+0.28}_{-0.25}$	4

Figure 4.11: Best-fit parameters for the evolution of the local LF with redshift. The best-fit pure luminosity (density) evolution in a given redshift bin is shown by the black points in the *top panel* (*bottom panel*). The blue (red) line in the top (bottom) panel shows the simple, continuous 2-parameter pure luminosity (density) evolution model (see text for details). Gray shaded areas in both panels correspond to the 68% confidence interval for a combined luminosity and density evolution. The large uncertainty in the combined fit is due to parameter degeneracy.

4.4.2 Evolving the luminosity function of AGN

The 1.4 GHz radio LFs for our radio AGN, separated into nine redshift ranges out to $z \sim 5$, are shown in Fig. 4.10, and tabulated in Table 4.2. We report LFs using the median luminosity in each luminosity bin, while the error bars show the width of the luminosity bin. In Fig. 4.11 we show the best-fit PDE ($\alpha_L = 0$) and PLE ($\alpha_D = 0$) for each redshift bin (see Eq. 4.7), which can be considered as the two extreme cases of evolution. For a conservative approach the outlying, lowest luminosity bins at $z < 1.3$ were ignored in the fitting process. To fit a simple, continuous model to the data, we added a redshift dependent term to the α_L, and α_D parameters as in Eq. 4.8. For pure luminosity evolution ($\alpha_D = \beta_D = 0$) the χ^2 minimization procedure yields best-fit parameter values of $\alpha_L = 2.88 \pm 0.17$, $\beta_L = -0.83 \pm 0.07$, while for pure density evolution ($\alpha_L = \beta_L = 0$) the best-fit parameters are $\alpha_D = 2.00 \pm 0.01$, $\beta_D = -0.60 \pm 0.02$. Both PDE and PLE models are similar, mostly because the shape of the LF does not deviate strongly from a simple power law at observed luminosities. Fitting for all four parameters simultaneously yields a strong degeneracy between the parameters and mainly differs from the two two-parameter fits at the high-luminosity end in high-redshift bins ($z > 2.1$).

4.4.3 Comparison with the literature

Based on the 2SLAQ[3] survey containing ~ 400 galaxies in a volume-limited sample at $0.4 < z < 0.7$, Sadler et al. (2007) have found that their radio AGN ($L_{1.4\,\text{GHz}} \approx 10^{24} - 10^{27}\,\text{W\,Hz}^{-1}$) undergo significant evolution since $z \sim 0.7$, parametrized with a pure luminosity evolution parameter $\alpha_L = 2.0 \pm 0.3$. Donoso et al. (2009) found fully consistent results, but with considerably smaller error bars as they used a sample of over $14\,000$ radio AGN. These derivations are in very good agreement with the luminosity function derived here (see $0.4 < z < 0.7$ bin in Fig. 4.10) and with the pure luminosity evolution we find for the VLA-COSMOS radio-excess AGN, detected at 3 GHz (see top panel of Fig. 4.11).

McAlpine et al. (2013) studied the evolution of faint radio AGN in the VISTA Deep Extragalactic Observations (VIDEO) Survey VIDEO-XMM3 field out to $z \sim 2.5$. By fitting a combined evolution of star-forming and AGN galaxies, they found a slightly weaker evolution than that inferred by Sadler et al. (2007, see Table 4 in McAlpine et al. 2013). The authors argue that this is because the evolution of the VIDEO-XMM3 radio AGN is primarily driven by the higher redshift range ($0.9 < z < 2.5$), not constrained by the 2SLAQ survey. The pure luminosity evolution inferred by McAlpine et al. (2013, $\alpha_L = 1.2 \pm 0.2$) is consistent with the average α_L derived here in the equivalent ($0.9 < z < 2.5$) redshift range (for example, for $z = 1.75$, $\alpha_L + z \cdot \beta_L = 1.4 \pm 0.2$ for our pure luminosity two-parameter fit; see top panel of Fig. 4.11). In each redshift range the VIDEO-XMM3-based volume densities of their AGN (shown in Fig. 4.10), which were identified during the photometric-redshift estimation process via template fitting (see their Sec. 6.2.1.), are slightly below those derived here, particularly at the high-luminosity end in each redshift bin (see Fig. 4.10). This underestimate of the number density of AGN at the high-luminosity end, and especially at $z > 1$, however, has already been observed and reported by McAlpine et al. (2013, see their Fig. 8).

Based on the VLA-COSMOS 1.4 GHz survey Smolčić et al. (2009b) have derived luminosity functions for their rest-frame color selected AGN out to $z = 1.3$, which is also shown in Fig. 4.10. Overall, there is good agreement between the derivations based on the (shallower) 1.4 GHz survey, and the 3 GHz survey used here. Smolčić et al. (2009b) inferred a pure luminosity (density) evolution out to $z = 1.3$ for their AGN of $\alpha_L = 0.8 \pm 0.1$ ($\alpha_D = 1.1 \pm 0.1$). While their pure density evolution is in agreement with that derived here, the pure luminosity evolution is here derived to be slightly higher; i.e., $\alpha_L + z \cdot \beta_L = 1.8 \pm 0.2$ for $z = 1.3$, and our pure luminosity two-parameter fit. The differences are due to a combination of factors, such as i) the ~ 3 times increased sensitivity of the VLA-COSMOS 3 GHz, compared to the 1.4 GHz Large

---

[3]Two-degree field (2dF) Survey of luminous red galaxies (LRGs) and quasi-stellar objects (QSO); 2dF-SDSS LRG and QSO (2SLAQ)

projects, ii) the different local luminosity functions used in the two studies; while the Mauch & Sadler 2007 local luminosity function is used here, the Sadler et al. 2002 local luminosity function was used by Smolčić et al. 2009b, and iii) differently selected AGN. Smolčić et al. (2009b) selected the radio AGN by requiring host galaxies with red rest-frame colors (Smolčić et al., 2008) that mimic the spectroscopically based identification (Baldwin et al., 1981) commonly used in the local universe, while we use the radio-excess criterion to identify our AGN, which includes both host galaxies with red rest-frame colors, but also AGN selected via X-ray, and mid-infrared (MIR) criteria within blue host galaxies (see Sect. 4.4.4, and Fig. 4.12 below; see also Delvecchio et al. 2017).

Bonzini et al. (2013) identified a sample of ($z \leq 4$) radio-loud AGN in the The Extended *Chandra* Deep Field South (ECDFS) field by requiring that at a given redshift, they lie below the 2σ deviation from the average observed 24 μm-to-1.4 GHz flux ratio obtained using the M82 galaxy template. This criterion is similar to that used here to select radio-excess AGN (see Sec. 4.3.2. in Delvecchio et al. 2017), and thus in Fig. 4.10 we also compare the luminosity functions derived here with those derived by Padovani et al. (2015) for the radio-loud AGN in the ECDFS field. Overall, the ECDFS-based luminosity functions agree well with those derived here out to $z = 1.3$. Beyond this redshift a higher degree of discrepancy is observed. This is likely related to small number statistics given the size of the ECDFS survey (0.3 square degrees) resulting in most of the ECDFS volume density values at $z > 1.3$ constrained by bins containing fewer than three sources (see Table 5 in Padovani et al. 2015 and Table 4.2 here).

4.4.4 Discussion of potential biases

Selection bias: Multiple AGN populations

Past studies have shown that radio AGN evolve differentially, consistent with cosmic downsizing; low radio-luminosity sources evolve less strongly than high radio-luminosity sources and the number density peak occurs at higher redshift for higher luminosity radio AGN (Dunlop & Peacock, 1990; Willott et al., 2001; Waddington et al., 2001; Rigby et al., 2011). Consistent with the evolution of optically and X-ray selected quasars (e.g., Schmidt et al., 1995; Silverman et al., 2008; Brusa et al., 2009), high-luminosity radio AGN ($L_{1.4\,\mathrm{GHz}} > 2 \times 10^{26}\,\mathrm{W\,Hz^{-1}}$) show a strong positive density evolution with redshift out to $z \sim 2$, beyond which their comoving volume density starts decreasing (Dunlop & Peacock, 1990; Willott et al., 2001). Intermediate- and low-luminosity radio AGN ($L_{1.4\,\mathrm{GHz}} \sim 1.6 \times 10^{24} - 3 \times 10^{26}\,\mathrm{W\,Hz^{-1}}$) have been shown to evolve slower with a comoving volume density turnover occurring at a lower redshift ($z \sim 1 - 1.5$; Waddington et al. 2001; Clewley & Jarvis 2004; Sadler et al. 2007; Donoso et al. 2009;

Smolčić et al. 2009b).

Two radio-luminous AGN populations have distinct host galaxy and AGN properties (e.g. Hardcastle et al., 2007; Padovani, 2016). One population is consistent with the standard, unified model AGN picture, in which the accretion occurs in a radiatively efficient manner at high Eddington rates ($1 - 10\% \lesssim \lambda_{Edd}$). Evidence has been presented that this class is fueled by the cold intragalactic medium phase and that it is not too likely to launch collimated jets (observable at radio wavelengths). The second population, however, exhibits radiatively inefficient accretion related to low Eddington ratios ($\lambda_{Edd} \lesssim 1 - 10\%$) and may be fueled by the hot phase of the IGM. It has been shown that this population is highly efficient in collimated jet production. The observed difference in Eddington ratio between the two populations can be linked to the switch between the standard accretion flow model, i.e., radiatively efficient, geometrically thin (but optically thick) disk accretion flow (Shakura & Sunyaev, 1973), and a radiatively inefficient, geometrically thick (but optically thin) accretion flow (Esin, 1997; Narayan et al., 1998), occurring at accretion rates below a certain Eddington ratio ($1 - 10\%$; e.g. Rees et al. 1982; Fanidakis et al. 2011). Further differences have been found demonstrating that the population of radiatively efficient accretors dominates the radio-AGN number densitites at the bright end (e.g., $L_{1.4\,GHz} \gtrsim 10^{26}$ W Hz^{-1} for $z < 0.3$; Pracy et al. 2016) and that this population evolves more rapidly with cosmic time (e.g., Willott et al., 2001; Pracy et al., 2016).

The radio AGN studied here have been selected based on an excess of their radio emission relative to that expected from the (IR-based) star formation rates in their host galaxies, thus assuring that $\gtrsim 80\%$ of the radio emission arises from the AGN core, jet, and lobe component. In Fig. 4.12 we plot the absolute and fractional contributions of X-ray-, MIR-selected AGN, and the remaining AGN (not identified via X-ray or MIR emission), those hosted by red, quiescent galaxies, and those hosted by galaxies with blue or green rest-frame colors, implying substantial star formation activity in the hosts (see Smolčić et al. 2017b for details). The X-ray and MIR regimes provide an efficient approach to identify radiatively efficient AGN, while red, quiescent host galaxies of radio AGN are shown to contain AGN with systematically lower radiative AGN luminosities (see, e.g., Fig. 7 in Smolčić et al. 2017b); further, this selection can be used to trace radiatively inefficient AGN (at least in the local universe; e.g., Smolčić et al. 2009b).

As seen in Fig. 4.12, at $z \lesssim 1$ our radio-excess AGN are composed of similar fractions of i) red, quiescent galaxies and ii) X-ray, MIR, and those hosted by star-forming galaxies, but not identified via X-ray or MIR emission. Beyond $z = 1$ the fraction of red, quiescent galaxies decreases steeply to a minimal fraction by $z \sim 2$. Hence, we can conclude that AGN with observational signatures of radiatively efficient accretion flows comprise a non-negligible fraction of our radio-excess AGN ($30 - 40\%$ X-ray- and MIR-selected AGN). If only a fraction of
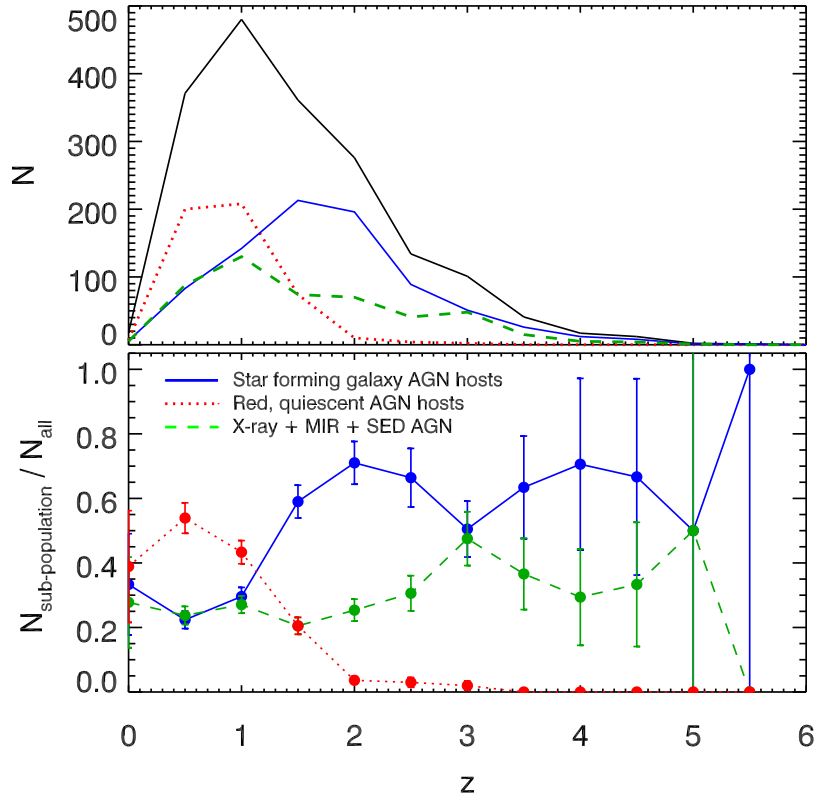
Figure 4.12: Absolute (*top panel*) and fractional (*bottom panel*) contributions of various subpopulations as indicated in the bottom panel to the full radio-excess AGN sample (black curve) as a function of redshift.

the radio-excess AGN hosted by blue or green, star-forming galaxies also contained radiatively efficient AGN, which could be the case given the expected cold gas supply in such, actively star-forming galaxies (e.g., Vito et al., 2014), the fraction of radiatively efficient AGN in our radio-excess sample would rise even further. This implies that our radio-excess AGN are likely a mix of radiatively efficient and inefficient black hole accretion flows, shown in other studies to evolve differently with cosmic time (e.g., Willott et al., 2001; Best et al., 2014; Pracy et al., 2016).

Radio spectrum bias

The volume densities derived here for our radio AGN are, at the high-luminosity end in every redshift range, somewhat higher than our best-fit pure luminosity/density two-parameter models (see Fig. 4.10). This could potentially be attributed to the rise given the faster evolution of
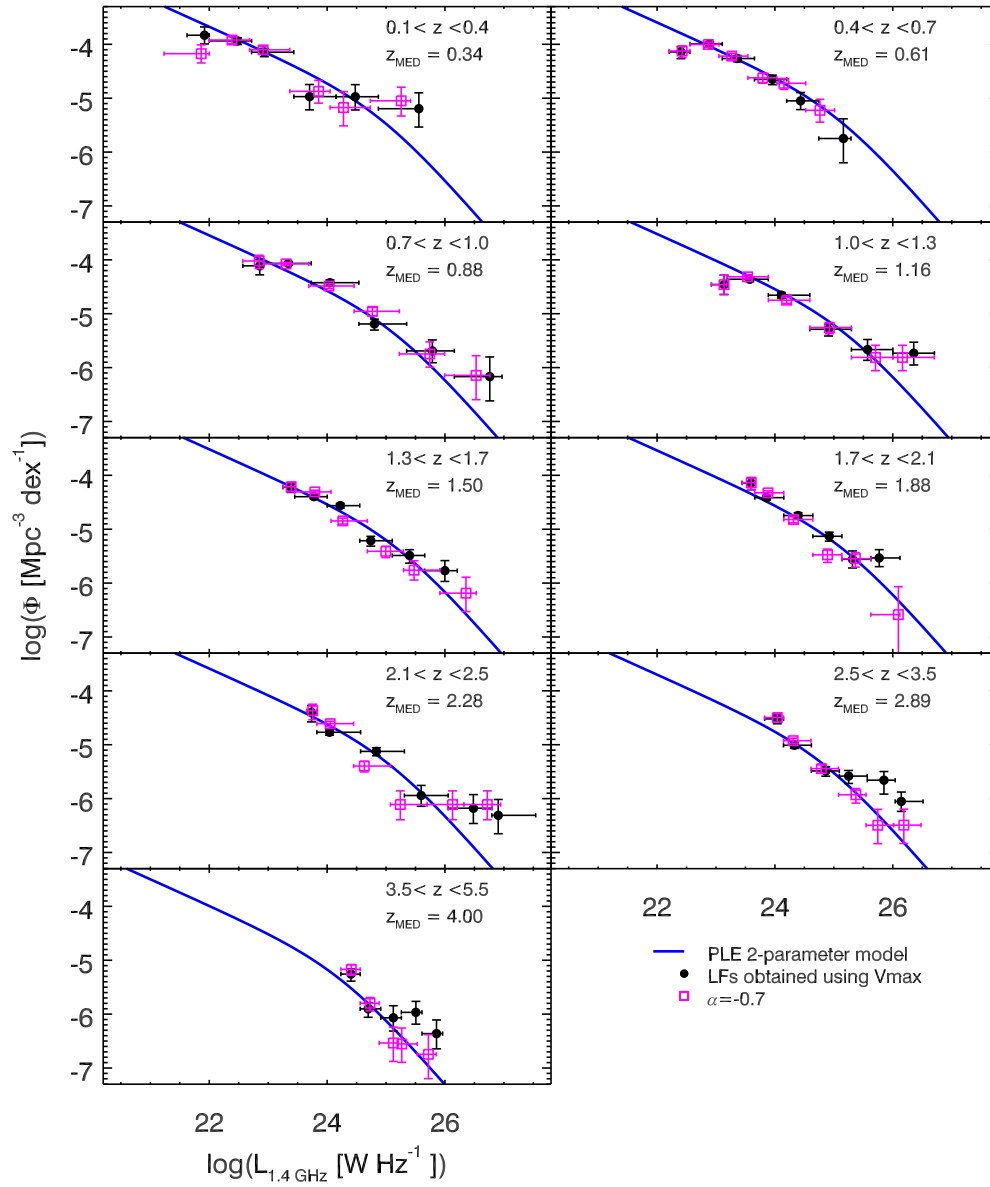
Figure 4.13: Radio luminosity functions of AGN, same as Fig. 4.10. Magenta squares indicate recalculated LFs using the same spectral index of $\alpha = -0.7$ for all galaxies.

the high-luminosity radio AGN that are consistent with the radiatively efficient accretion flow population. However, these particular, high-luminosity bins at each redshift are the most sensitive to the assumption of a simple synchrotron power law for the conversion from the observed 3 GHz flux density to rest-frame 1.4 GHz luminosity. For example, assuming a spectral index of $\alpha = -0.7$ for all our AGN instead of using the observed value, if available, which steepens toward high redshift (see Fig. 4.9), this would have very little effect on the low-luminosity volume densities in the high-redshift bins, but the high-luminosity bin values would decrease to be consistent with the pure luminosity (density) two-parameter models (see Fig. 4.13).

If the typical radio AGN spectrum were to steepen toward the high-frequency end owing to synchrotron energy losses (e.g., Miley 1980), as sampled by the observed frequency of 3 GHz at $z > 2$ (corresponding to rest-frame frequencies of > 9 GHz) then, for example, a broken power law, rather than a simple, single power law assumption would be more appropriate for the K correction. On the other hand, in this case, the simple, single power law assumption of $\alpha = -0.7$ could result in a more correct 1.4 GHz rest-frame luminosity value than the directly observed higher spectral index values.

4.5 Total radio luminosity function

The final sample for which we construct luminosity functions is the largest obtainable with the available data. By temporarily disregarding any galaxy classification we have a sample of 7 826 radio detected galaxies with COSMOS2015 and *i*-band counterparts and therefore good redshift estimates (the third sample in Sect. 4.1.3). The total radio LFs are presented in Fig. 4.14 and also listed in Table 4.3. We did not apply any statistical corrections for radio sources without assigned optical counterparts as the redshift distribution is unknown, however we expect these effects to be less than 10% (see also Sect. 4.1.2). In Fig. 4.14 we also show continuous pure luminosity evolution models obtained on individual populations presented in Sections 4.3 and 4.4 as blue and red areas, respectively.

4.5.1 Fitting the total radio luminosity function

Following the idea from McAlpine et al. (2013) we attempted to fit the total radio LF with a combination of different LFs. We constructed a four parameter redshift dependent pure luminosity evolution model, with two parameters for both SF and AGN populations. The total LF is
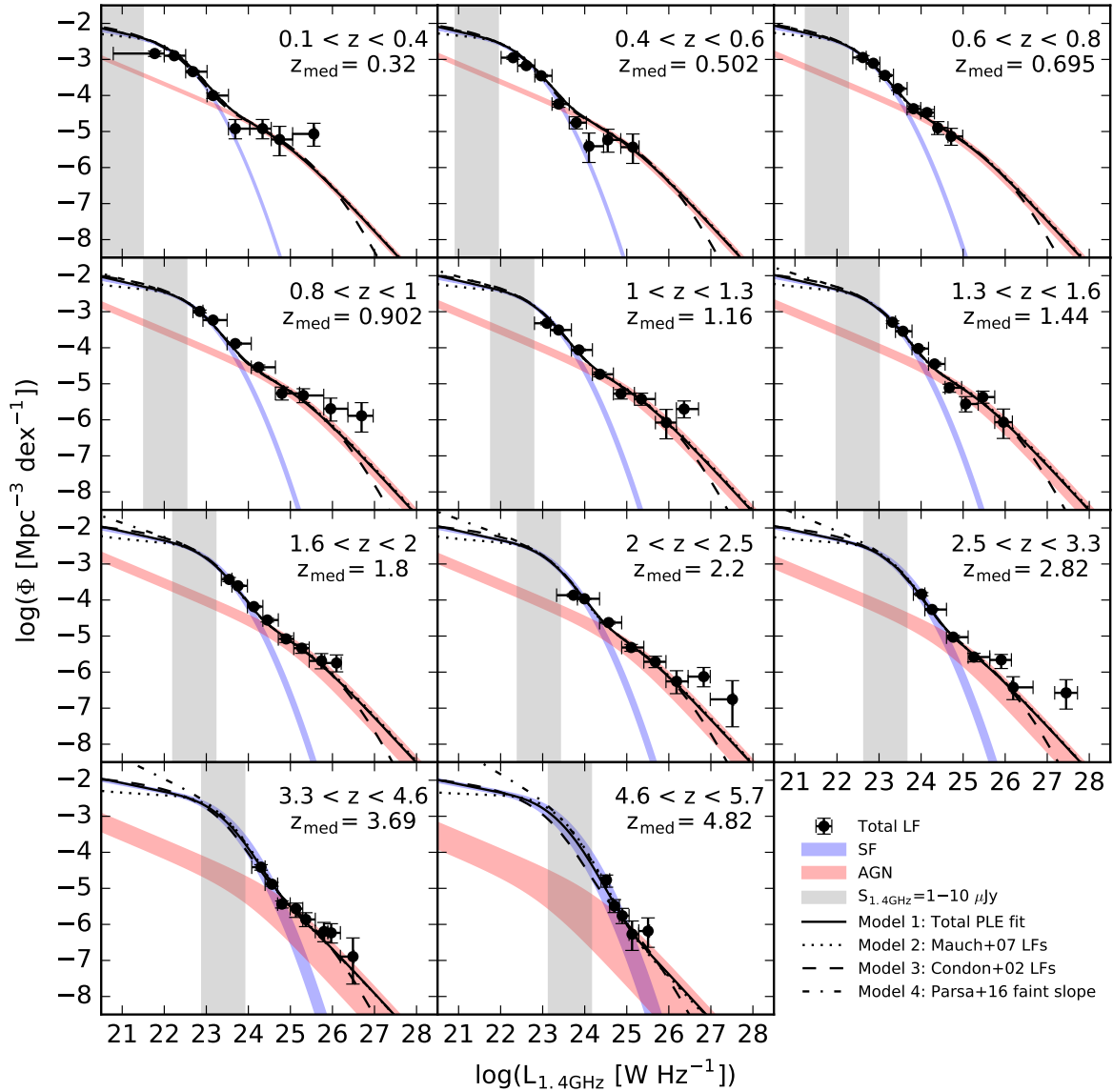
Figure 4.14: Total radio luminosity functions at different epochs. Black points show LFs derived with V_{max}, blue and red shaded area correspond to 3σ significant evolution fits of individual SF and AGN populations, respectively, the black line shows the fit on the total radio LF explained in Sect. 4.5.1, other models are also indicated in the legend. The vertical grey shaded area shows the luminosity decade that contributes to the radio source counts between 1 and 10 μJy.

Table 4.3: Luminosity functions of the total radio selected sample. The luminosity $L_{1.4\,\mathrm{GHz}}$ has units of $\mathrm{W\,Hz^{-1}}$ and the density Φ has units of $\mathrm{Mpc^{-3}\,dex^{-1}}$.

z	$\log L_{1.4\,\mathrm{GHz}}$	$\log \Phi$	N	z	$\log L_{1.4\,\mathrm{GHz}}$	$\log \Phi$	N
$0.1-0.4$	$21.77^{+0.23}_{-0.99}$	$-2.84^{+0.080}_{-0.068}$	202	$1.3-1.6$	$23.32^{+0.068}_{-0.14}$	$-3.29^{+0.040}_{-0.037}$	173
	$22.24^{+0.27}_{-0.24}$	$-2.90^{+0.024}_{-0.023}$	352		$23.57^{+0.22}_{-0.18}$	$-3.54^{+0.022}_{-0.021}$	432
	$22.68^{+0.34}_{-0.17}$	$-3.34^{+0.038}_{-0.035}$	145		$23.94^{+0.24}_{-0.15}$	$-4.02^{+0.033}_{-0.031}$	188
	$23.16^{+0.37}_{-0.14}$	$-4.00^{+0.083}_{-0.070}$	33		$24.32^{+0.25}_{-0.14}$	$-4.45^{+0.055}_{-0.049}$	72
	$23.69^{+0.35}_{-0.16}$	$-4.92^{+0.28}_{-0.25}$	4		$24.67^{+0.29}_{-0.10}$	$-5.11^{+0.13}_{-0.098}$	16
	$24.34^{+0.21}_{-0.30}$	$-4.92^{+0.28}_{-0.25}$	4		$25.06^{+0.30}_{-0.095}$	$-5.56^{+0.22}_{-0.20}$	6
	$24.74^{+0.31}_{-0.20}$	$-5.22^{+0.45}_{-0.37}$	2		$25.47^{+0.28}_{-0.11}$	$-5.37^{+0.17}_{-0.16}$	9
	$25.56^{+0.030}_{-0.50}$	$-5.07^{+0.34}_{-0.30}$	3		$25.96^{+0.21}_{-0.21}$	$-6.07^{+0.45}_{-0.37}$	2
$0.4-0.6$	$22.30^{+0.11}_{-0.29}$	$-2.95^{+0.048}_{-0.043}$	151	$1.6-2.0$	$23.54^{+0.067}_{-0.19}$	$-3.43^{+0.061}_{-0.053}$	192
	$22.61^{+0.20}_{-0.21}$	$-3.17^{+0.027}_{-0.026}$	271		$23.75^{+0.22}_{-0.14}$	$-3.61^{+0.020}_{-0.019}$	494
	$22.96^{+0.26}_{-0.15}$	$-3.45^{+0.036}_{-0.033}$	165		$24.13^{+0.21}_{-0.15}$	$-4.18^{+0.035}_{-0.032}$	173
	$23.38^{+0.25}_{-0.16}$	$-4.24^{+0.089}_{-0.074}$	29		$24.45^{+0.26}_{-0.11}$	$-4.55^{+0.055}_{-0.049}$	73
	$23.80^{+0.24}_{-0.17}$	$-4.76^{+0.17}_{-0.16}$	9		$24.90^{+0.18}_{-0.19}$	$-5.08^{+0.099}_{-0.081}$	24
	$24.10^{+0.35}_{-0.060}$	$-5.41^{+0.45}_{-0.37}$	2		$25.27^{+0.17}_{-0.19}$	$-5.34^{+0.14}_{-0.11}$	13
	$24.55^{+0.31}_{-0.099}$	$-5.23^{+0.34}_{-0.30}$	3		$25.74^{+0.075}_{-0.29}$	$-5.69^{+0.22}_{-0.20}$	6
	$25.14^{+0.15}_{-0.29}$	$-5.43^{+0.45}_{-0.37}$	2		$26.10^{+0.11}_{-0.28}$	$-5.75^{+0.25}_{-0.22}$	5
$0.6-0.8$	$22.61^{+0.080}_{-0.24}$	$-2.95^{+0.062}_{-0.054}$	190	$2.0-2.5$	$23.73^{+0.093}_{-0.40}$	$-3.87^{+0.044}_{-0.040}$	168
	$22.86^{+0.15}_{-0.17}$	$-3.11^{+0.024}_{-0.023}$	354		$23.99^{+0.36}_{-0.17}$	$-3.97^{+0.022}_{-0.021}$	410
	$23.14^{+0.20}_{-0.12}$	$-3.44^{+0.032}_{-0.030}$	205		$24.57^{+0.31}_{-0.21}$	$-4.63^{+0.044}_{-0.040}$	112
	$23.45^{+0.22}_{-0.10}$	$-3.81^{+0.048}_{-0.043}$	92		$25.10^{+0.30}_{-0.23}$	$-5.32^{+0.11}_{-0.085}$	23
	$23.82^{+0.17}_{-0.15}$	$-4.37^{+0.096}_{-0.079}$	26		$25.68^{+0.25}_{-0.27}$	$-5.71^{+0.16}_{-0.15}$	10
	$24.14^{+0.18}_{-0.15}$	$-4.48^{+0.11}_{-0.086}$	21		$26.18^{+0.28}_{-0.25}$	$-6.26^{+0.34}_{-0.30}$	3
	$24.40^{+0.24}_{-0.080}$	$-4.90^{+0.18}_{-0.17}$	8		$26.83^{+0.16}_{-0.37}$	$-6.13^{+0.28}_{-0.25}$	4
	$24.71^{+0.28}_{-0.074}$	$-5.14^{+0.25}_{-0.22}$	5		$27.51^{+0.028}_{-0.53}$	$-6.76^{+0.76}_{-0.52}$	1
$0.8-1.0$	$22.85^{+0.072}_{-0.16}$	$-2.99^{+0.042}_{-0.038}$	184	$2.5-3.3$	$24.01^{+0.076}_{-0.20}$	$-3.84^{+0.048}_{-0.043}$	172
	$23.16^{+0.33}_{-0.24}$	$-3.23^{+0.018}_{-0.017}$	647		$24.26^{+0.34}_{-0.17}$	$-4.26^{+0.025}_{-0.024}$	327
	$23.69^{+0.38}_{-0.19}$	$-3.89^{+0.035}_{-0.032}$	177		$24.76^{+0.36}_{-0.16}$	$-5.03^{+0.058}_{-0.051}$	67
	$24.24^{+0.40}_{-0.17}$	$-4.54^{+0.073}_{-0.062}$	42		$25.26^{+0.38}_{-0.14}$	$-5.58^{+0.13}_{-0.10}$	17
	$24.80^{+0.42}_{-0.15}$	$-5.27^{+0.18}_{-0.17}$	8		$25.91^{+0.24}_{-0.27}$	$-5.66^{+0.24}_{-0.15}$	12
	$25.31^{+0.48}_{-0.090}$	$-5.32^{+0.20}_{-0.19}$	7		$26.19^{+0.47}_{-0.042}$	$-6.42^{+0.34}_{-0.30}$	3
	$25.96^{+0.41}_{-0.16}$	$-5.69^{+0.34}_{-0.30}$	3		$27.45^{+0.27}_{-0.27}$	$-6.58^{+0.45}_{-0.37}$	2
	$26.69^{+0.28}_{-0.33}$	$-5.89^{+0.45}_{-0.37}$	2	$3.3-4.6$	$24.30^{+0.097}_{-0.22}$	$-4.41^{+0.084}_{-0.070}$	75
$1.0-1.3$	$23.10^{+0.080}_{-0.30}$	$-3.32^{+0.045}_{-0.041}$	228		$24.56^{+0.14}_{-0.16}$	$-4.89^{+0.058}_{-0.051}$	66
	$23.38^{+0.30}_{-0.19}$	$-3.51^{+0.019}_{-0.019}$	542		$24.80^{+0.20}_{-0.10}$	$-5.44^{+0.11}_{-0.087}$	21
	$23.86^{+0.32}_{-0.18}$	$-4.06^{+0.033}_{-0.031}$	190		$25.13^{+0.16}_{-0.14}$	$-5.56^{+0.24}_{-0.15}$	11
	$24.36^{+0.32}_{-0.18}$	$-4.73^{+0.072}_{-0.062}$	43		$25.37^{+0.23}_{-0.072}$	$-5.86^{+0.20}_{-0.19}$	7
	$24.86^{+0.32}_{-0.18}$	$-5.27^{+0.15}_{-0.11}$	12		$25.80^{+0.092}_{-0.21}$	$-6.20^{+0.28}_{-0.25}$	4
	$25.35^{+0.33}_{-0.17}$	$-5.42^{+0.17}_{-0.16}$	9		$25.97^{+0.22}_{-0.080}$	$-6.23^{+0.28}_{-0.25}$	4
	$25.94^{+0.24}_{-0.26}$	$-6.08^{+0.45}_{-0.37}$	2		$26.49^{+0.026}_{-0.30}$	$-6.89^{+0.76}_{-0.52}$	1
	$26.36^{+0.34}_{-0.18}$	$-5.70^{+0.25}_{-0.22}$	5	$4.6-5.7$	$24.52^{+0.077}_{-0.13}$	$-4.77^{+0.20}_{-0.14}$	14
					$24.71^{+0.12}_{-0.11}$	$-5.49^{+0.18}_{-0.17}$	8
					$24.89^{+0.16}_{-0.069}$	$-5.76^{+0.22}_{-0.20}$	6
					$25.13^{+0.16}_{-0.072}$	$-6.27^{+0.45}_{-0.37}$	2
					$25.51^{+0.035}_{-0.22}$	$-6.19^{+0.45}_{-0.37}$	2

defined as the sum

$$\Phi(L, z, \alpha_L^{SF}, \beta_L^{SF}, \alpha_L^{AGN}, \beta_L^{AGN}) = \Phi_0^{SF} \left[\frac{L}{(1+z)^{\alpha_L^{SF} + z \cdot \beta_L^{SF}}} \right] + \Phi_0^{AGN} \left[\frac{L}{(1+z)^{\alpha_L^{AGN} + z \cdot \beta_L^{AGN}}} \right],$$
$$(4.12)$$

where Φ_0^{SF} and Φ_0^{AGN} are the local LFs for the two galaxy populations. It was noted in Sect. 4.3.2 that the PDE of SF galaxies would push the densities to inconsistently high numbers. Our data can constrain only the bright log-normal part of the SF LF. For AGN it was shown in Sect. 4.4.2 that both PDE and PLE models are similar. Considering the above reasoning, while also trying to keep the parameter space degeneracy to a minimum, we decided to use the PLE only for our models. Both PLE and PDE models are common in the literature (e.g. Condon, 1984; Sadler et al., 2002; Gruppioni et al., 2013), however the true evolution might be a combination of both of these extremes (see e.g. Yuan et al., 2016), with a possible luminosity dependent evolution as well (see e.g. Fotopoulou et al., 2016).

We used the Markov chain Monte Carlo (MCMC) algorithm, available in the Python package EMCEE (Foreman-Mackey et al., 2013), to perform a multivariate fit to the data in Table 4.3 with a model described in Eq. 4.12, where the local LFs are defined in Eq. 4.9 and 4.11. Covariance maps of four fitted parameters are shown in Fig. 4.15. It can be seen from this plot that the total fit shows a slight degeneracy between SF and AGN populations, however the intersection of AGN and SF luminosity functions is well constrained by our data (see Fig. 4.14) and we therefore consider our fit robust despite this small degeneracy. For an individual population, a higher α parameter correlates strongly with a steeper (more negative) value of β. The redshift dependence of the total evolution parameter $(\alpha + z \cdot \beta)$ is necessary to describe the observations at all redshifts. The best-fit parameter values are listed in Table 4.4 as model 1. The total radio LF is also shown with a black line in Fig. 4.14.

4.5.2   Discussing the fitted model

One advantage of performing a fit to the total radio LF is that results are no longer sensitive to the galaxy classification method. However, it is important to assume a proper shape of the LFs for distinct galaxy populations in order to obtain a meaningful evolution fit. For this reason the shape of the subpopulation LF is chosen from the local one which is usually constrained across at least five orders of magnitudes in luminosities. The best-fit evolution parameters obtained from the MCMC on the total radio LF are consistent within the 3σ errors with χ^2 fits done on individual populations (see Fig. 4.14) and suggest slightly higher AGN densities. This implies that regardless of the specifics of the galaxy classification methods, if local LFs for SF and AGN
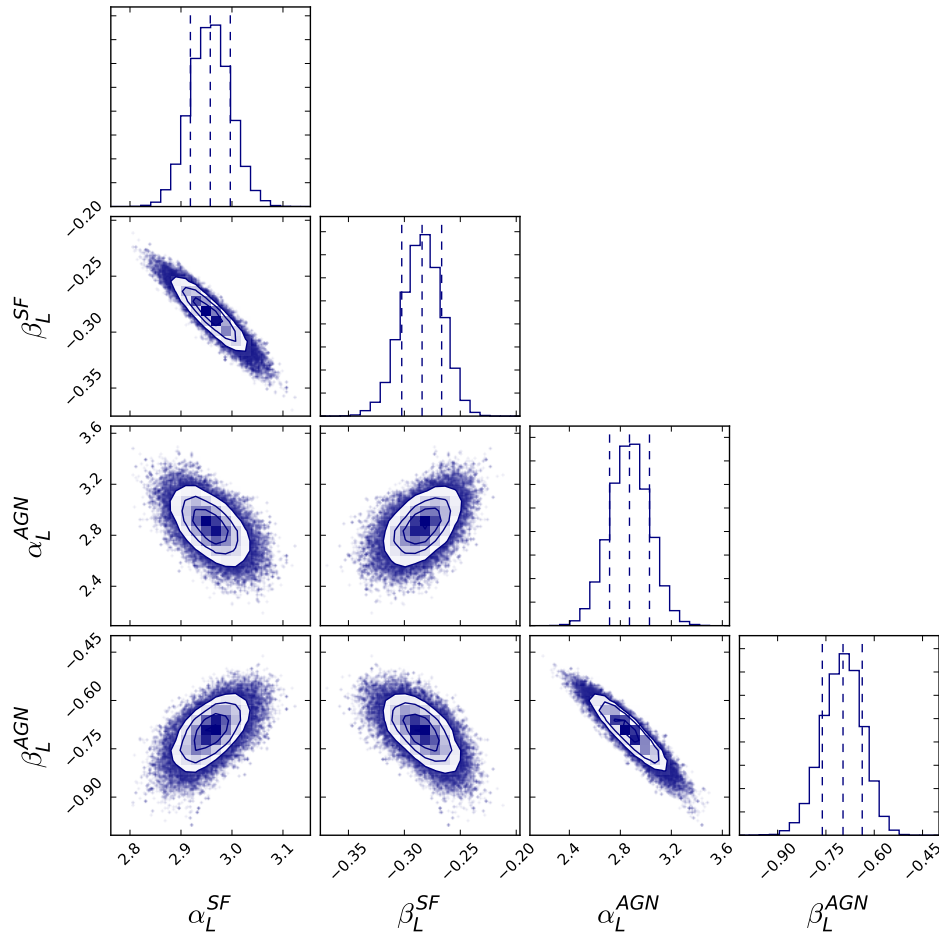
Figure 4.15: Covariance maps of the the four fitting parameters used to describe the total radio LF evolution. Probability distribution histograms are also shown for each parameter with 16th, 50th and 84th percentiles marked with vertical dashed lines.

galaxies are well described with Eq. 4.9 and 4.11, our PLE parameters are robust.

Given the dichotomy of the AGN galaxy population, the AGN LF should be additionally decomposed into two separate LFs, each one evolving differently (e.g. Pracy et al., 2016). Here we do not consider this additional classification for the following reasons. First, we are interested in separating galaxies based on the dominant process which gives rise to the observed radio emission, because the radio power can be further scaled to estimate for example the star formation rate or the AGN kinetic power. In this scenario it is not relevant if a galaxy hosts an AGN if almost all of its radio emission is due to supernova remnants. Second, given that the VLA-COSMOS is a two square degrees deep survey, with our data it is not possible to constrain well the AGN LF turnover as only the power law like faint end of the AGN population is sampled well. A wide (and shallow) survey is needed to fill the bright part of the LF with the rarest

Table 4.4: Best fit parameters obtained from the multivariate fitting of the total radio LF with two populations.

Model	Description	α_L^{SF}	β_L^{SF}	α_L^{AGN}	β_L^{AGN}
0	Subpopulations	3.16 ± 0.04	-0.32 ± 0.01	2.88 ± 0.17	-0.83 ± 0.07
1	PLE fit to the total LF	2.96 ± 0.04	-0.28 ± 0.02	2.87 ± 0.16	-0.70 ± 0.06
2	Mauch & Sadler (2007) local LFs	2.78 ± 0.04	-0.23 ± 0.02	3.05 ± 0.15	-0.75 ± 0.06
3	Condon et al. (2002) local LFs	3.58 ± 0.04	-0.44 ± 0.02	2.48 ± 0.16	-0.60 ± 0.06
4	Parsa et al. (2016) faint slope	3.05 ± 0.04	-0.19 ± 0.02	2.92 ± 0.15	-0.71 ± 0.06

objects. Furthermore, the uncertainty of the radio SED shape affects mostly the bright end of the LF which presents problems in disentangling the evolution trends from possible radio spectral index systematics (see also Sections 4.3.4 and 4.4.4 for more detailed discussion on this), possibly explaining the outlying most luminous LF points in Fig. 4.14.

The radio excess criteria from Eq. 4.1 was designed to select AGN with high purity, by reducing the threshold more galaxy composites would be reclassified as AGN. The SF population would also benefit from a further decomposition into differently evolving normal/disk galaxies, starbursts/mergers and galaxies with low level AGN activity (see e.g. Hopkins et al., 2010), however such a classification is beyond the scope of this work and would introduce too many degenerate parameters into our models. Given that the faint luminosity end is dominated by disk galaxies and a clear cut between different populations does not exist, we assume that the sum of two differently evolving LFs are enough to describe global trends from which we can constrain the faint radio sky. The fit is performed on the total radio LF obtained from V_{max} and therefore all subpopulation intricacies are encoded in the shapes of the fitting functions.

Since the shape of the LF is defined analytically, it extends beyond the observed data. The extrapolation toward the faint end (e.g. below $L_{1.4\,GHz} < 10^{23}$ W Hz^{-1} at $z > 1$) is particularly important because the densities of galaxies are the highest in this range. A change in the faint end slope of the LF can easily change relative fractions of galaxy populations at faint flux densities. The only way to mitigate this extrapolation issue is to obtain deeper data. However, in the absence of such observations we tried to estimate the effect of extrapolations by using different LFs. We performed additional MCMC fits using SF and AGN LFs from Condon et al. (2002), where both LFs are hyperbolic in shape, Mauch & Sadler (2007), and the faint power law slope for SF galaxies derived by Parsa et al. (2016). The best-fit evolution parameters are listed in Table 4.4. The effects these LF choices have on the final radio number counts are discussed in the next section.

4.6 Chapter summary

We used the VLA-COSMOS 3 GHz Large Project radio data to derive the radio luminosity functions of different galaxy populations out to redshifts of $z \lesssim 5.5$ using the $1/V_{max}$ method. Radio sources with a robust NIR/optical counterpart in the COSMOS2015 catalog (Laigle et al., 2016) with a valid photometric or spectroscopic redshift were categorized into two main galaxy populations based on radio excess of that expected from the IR-based star formation rates of their hosts (derived from SED fits, see Delvecchio et al. 2017). A $> 3\sigma$ excess assures that at least 80% of the galaxy radio emission is due to the AGN component, yielding 1 814 sources classified as AGN across an area of 1.77 deg^2. The remaining 5 915 galaxies without a 3σ excess were classified as star-forming galaxies. For the total radio sample, additional $i-$band selected counterparts were also included yielding a sample of 7 826 radio selected galaxies. Radio luminosities of the samples are in the range of $10^{21} - 10^{27}$ W Hz^{-1} with typical stellar masses within $10^{9.5} - 10^{11.5}$ M$_\odot$.

By comparing star-forming luminosity functions with those derived using IR and UV selected samples we checked their robustness and found that our radio SF LF can be very well described by a local LF with a power law plus log-normal form evolved only in luminosity as $L_{1.4\,GHz} \propto (1+z)^{(3.16\pm0.04)-(0.32\pm0.01)z}$. However, we do not observe the faint end of the LF at all redshifts to properly constrain a more complex evolution. The difference between radio and UV LFs suggests an underestimation of dust corrections obtained from UV slopes by Bouwens et al. (2014a). We constrained the evolution of AGN population with continuous models of pure density and pure luminosity evolutions finding best-fit parameters of $\Phi^* \propto (1+z)^{(2.00\pm0.01)-(0.60\pm0.02)z}$ and $L^* \propto (1+z)^{(2.88\pm0.17)-(0.83\pm0.07)z}$, respectively.

Chapter 5

Constraints on nano-Jy radio number counts based on evolving VLA-COSMOS luminosity functions

In this chapter we aim to verify that the luminosity functions of SF galaxies and AGN calculated in Sections 4.3 and 4.4, respectively, can reproduce number counts consistent with the literature. We also use different models of evolving LFs to estimate radio number counts, relative fractions of different galaxy sub-populations and redshift distributions of SF galaxies down to nano-Jy flux density levels. These results provide further constraints for future radio surveys.

5.1   Radio number counts

One way to constrain how different galaxy populations add up to the total radio sky is through source number counts, i.e. the number of galaxies contained in a solid angle of the sky with a given flux density (for a review from a radio perspective see e.g. Padovani, 2016). First used as a tool for studying the geometry of the universe, source number counts became a practical method for tracing galaxy evolution which also provides a statistical way to search for and describe new populations of galaxies (e.g. Longair, 1966; Franzen et al., 2014). Flattening of Euclidean normalized number counts at sub-mJy flux densities at 1.4 GHz indicated a surfacing of a new galaxy population in the radio (e.g. Rowan-Robinson et al., 1993; Seymour et al., 2004; Padovani et al., 2009), however the scatter was significant (e.g. Hopkins et al., 2003; Heywood et al., 2013). There were also some indications that counts should remain flat (e.g. Owen & Morrison, 2008). See also Fixsen et al. (2011) and Vernstrom et al. (2011) for a discussion about radio excess from ARCADE2 sky brightness measurements, which pointed toward a

possible existence of an abundant, but faint new radio population (possible diffuse emission from clusters or halos or dark matter annihilation). However, recent deep VLA observations show evidence of further steepening of number counts below $\sim 50\ \mu$Jy at 1.4 GHz consistent with the idea of SF galaxies dominating the faint radio sky (Vernstrom et al., 2016b, see also Sect. 3.6.1).

There are several methods applicable for calculating source number counts. The traditional one relies on counting discrete sources in flux density bins (see de Zotti et al. 2010 for a large compilation). It is the most straightforward approach, but one must be careful to take care of different biases such as non-uniform *rms* across the radio map, or resolution bias, and correct for such completeness issues (e.g. Hales et al., 2014; Vernstrom et al., 2016a, also see Sect. 3.5). A second approach is a blind probability of deflection analysis $P(D)$ (Scheuer, 1957) done on a confusion-limited survey as was recently utilized on deep VLA data by Condon et al. (2012); Vernstrom et al. (2014). The advantage of $P(D)$ is that it relies on a small number of assumptions and it can model source counts well below the usual 5σ sensitivity limit, because the method measures the bias in the noise. However, due to its blind nature it yields no information on specific sub-populations of galaxies. When the resolution is high enough so that the radio map is no longer dominated by confusion noise, one possible option is to stack the radio map at known positions using a catalog of priors obtained on a map with higher sensitivity (for a Bayesian approach see e.g. Mitchell-Wynne et al., 2014; Zwart et al., 2015b). This method also allows probing the radio map below the nominal sensitivity threshold, however careful assessment of biases is required (see also Zwart et al., 2015a).

If we assume a cosmological model for the universe, the shape of the number counts are a natural consequence of the shapes and the evolutions of underlying luminosity functions (e.g. Condon, 1989). Analytical forms of LFs can be used to extrapolate results below the sensitivity limit and obtain number counts down to nano-Jy flux density levels. This limit was chosen to encompass future deep radio surveys, as well as to provide insights into the number counts trends originating in uncertain extrapolations. Such constraints are useful for future radio surveys such as the SKA, ASKAP, EMU and MeerKAT (see also Norris et al., 2013). We show that the future radio observations will provide a valuable and simple diagnostic that can be used to better classify galaxies as SF or AGN.

5.2   Obtaining number counts from luminosity functions

The number of galaxies ΔN in a given spherical shell volume ΔV and a luminosity decade $\Delta \log L$ is given by the definition of the luminosity function Φ as

$$\Delta N = \Phi(L, z) \, \Delta V \, \Delta \log L. \tag{5.1}$$

To obtain the differential number counts n which are Euclidean normalized, i.e. weighted by flux density to the power of $5/2$ (obtained from a uniform distribution of sources in a flat universe), one must sum ΔN contributions from all redshifts and luminosities which would result in flux density range of ΔS around some S, formally

$$n(S) = \frac{1}{4\pi\Delta S} S^{2.5} \sum_{S}^{S+\Delta S} \Delta N(S). \tag{5.2}$$

We numerically integrated all LF models described with Eq. 4.12 and listed in Table 4.4 within a redshift range $0 < z < 6$, which is constrained by our radio data. We discuss the results in the following sections.

5.3   Individual subpopulation fits

We first focus on the 1.4 GHz number counts obtained from LFs of individual populations reported in Sections 4.3 and 4.4. Number counts are shown in Fig. 5.1 along with several other works from the literature.

We compare these number counts with results from semi-analytical models done by Wilman et al. (2008) and Béthermin et al. (2012) and they all agree around $100\,\mu$Jy. Our number counts are systematically slightly higher at fainter flux densities when compared to the Wilman et al. (2008) model, an offset probably caused by the choice of the LF and its evolution. Below $10\,\mu$Jy our number counts are increasingly higher than the Béthermin et al. (2012) model, however, the agreement with their simplified model is excellent in this range. One of the refinements that the authors implemented in their main model was the dust attenuation, which in practice flattened the faint end of the IR LF and probably causes this discrepancy. Because of the uncertainty of dust obscuration and its dependence on redshift, the dust-unbiased aspect of radio observations provides a valuable addition to galaxy evolution studies. We also compare our number counts with the work performed by Mancuso et al. (2015), where the authors use the evolving cosmic SFR function obtained from the far-infrared (FIR), UV and Hα data to estimate radio (syn-
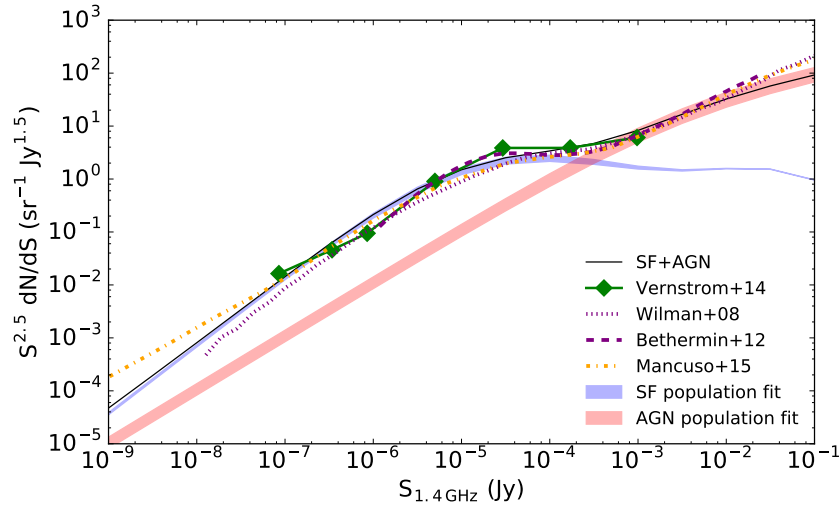
Figure 5.1: Euclidean normalized radio source number counts drawn from LFs described in Sections 4.3 and 4.4 compared to the literature values indicated in the legend. Shaded areas encompass the 3σ errors from the χ^2 fits done on individual populations.

chrotron and free-free) LFs. We find an excellent agreement with their work down to 0.1 μJy. Below this flux density there is an increasing disagreement between the counts, however in Sect. 5.4 we use a model with a changing faint end slope for SF galaxies, which yields higher densities of SF galaxies at nJy levels.

An important result stems from the comparison of our counts with those obtained by Vernstrom et al. (2014). These investigators modeled the confusion amplitude distribution $P(D)$ from the deep VLA 3 GHz observations of the Lockman Hole. Their analysis uses the noise distribution of the radio map to constrain the underlying radio number counts down to $S_{3\,\mathrm{GHz}} \sim$ 50 nJy. A great advantage of their approach is that it is not biased by incompleteness issues from direct counting and counterpart cross-correlations, and can be used to probe the radio sky properties well below the nominal sensitivity limit for source detection. The excellent agreement between results from their blind method and our counts suggests that the evolving LFs described in Sections 4.3 and 4.4 provide plausible cosmic densities of galaxies, even though they rely on uncertain extrapolations toward the faint end. There is no need for a potential third population of radio sources (e.g. dwarf galaxies) at $0.1 - 1$ μJy levels and these LFs can reproduce the observed radio sky background.

Our number counts at the bright end exhibit a turnover at somewhat lower flux densities when compared to semi-analytical models (derived from abundant observations at $S_{1.4\,\mathrm{GHz}} >$ 1 mJy, see de Zotti et al. 2010), probably due to an overly simplified treatment of the brightest AGN. However, we emphasize that our work is optimized to probe the faint radio sky and SF

galaxies, and extrapolations toward the faint end are discussed next.

5.4   The faint radio sky: What will the SKA see?

We consider the number counts obtained from the MCMC fit described in Sect. 4.5.1 (model 1) as our main result and list them in Table 5.1. To quantify possible issues due to the faint end slope of the LF, we tried evolving different LFs available in the literature. All models are shown in Fig. 5.2. They are consistent between themselves, with expected variations below 1 μJy. Due to the almost flat faint power law slope of the SF LF reported in Mauch & Sadler (2007), and a steep one for the AGN, this model predicts the fewest SF galaxies at nano-Jy flux densities. This is not the case with Condon et al. (2002) local LFs where the slopes for both populations are similar leading to a roughly constant ratio of SF to AGN number counts below 1 μJy.

In Sect. 4.3.3 we discussed that radio LF might have a too flat faint end slope (below $L_{1.4\,\mathrm{GHz}} < 10^{22}$) when compared to UV LFs at high redshift ($z > 3$). To obtain some insight into this effect we have constructed a fourth model which uses a redshift dependent faint end slope, derived by Parsa et al. (2016) to describe UV LFs. This model is the same as our first model, with the exception that the faint slope parameter for SF galaxies (α in Eq. 4.9) now changes linearly with redshift as $\alpha = 0.106 \cdot z + 1.187$. The result of this LF choice is an increased density of faint galaxies, which is then reflected in an order of magnitude larger number counts at 1 nJy than the values obtained from the first model. In this case the SF population dominates the radio number counts down to nJy levels, unless the AGN faint slope also steepens considerably.

To better illustrate the effect of how relative abundances of galaxies change with the telescope sensitivity, in Fig. 5.3 we show the fraction of each population as a function of flux density. Several conclusions can be drawn from this plot. Above 1 mJy the majority of radio emission is due to AGN, as has been well studied in the past. At around 200 μJy at 1.4 GHz both populations are equally present in the radio sky and at fainter flux densities SF galaxies become the dominant population. All LF models agree down to 1 μJy, however at 1 – 100 nJy they diverge due to different faint end slopes of the LFs. This divergence also causes the AGN percentage to rise in some models at faintest flux densities. The model based solely on Mauch & Sadler (2007) LFs predicts that AGN and SF galaxies will again be equally abundant at 1 nJy, however this is a product of extrapolating the LFs with significantly different faint end slopes (the SF LF slope is flat, while the AGN LF slope is steep). Observations at UV wavelengths do not support such a flat faint end slope of the SF LF (e.g. Bouwens et al., 2015). The most important result is that between 100 nJy and 10 μJy, where the 5σ sensitivity limits for future

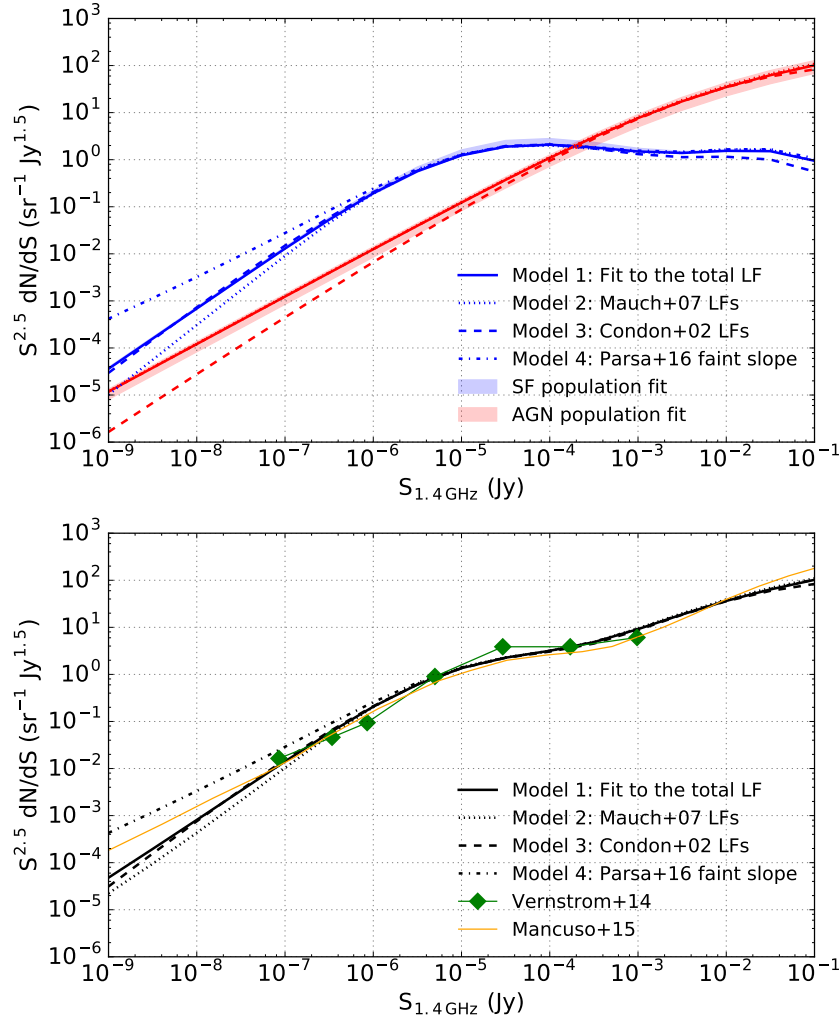Figure 5.2: *Top panel:* Euclidean normalized radio source number counts obtained by fitting the total radio LF using different evolving analytical LFs with the MCMC algorithm. Red lines (AGN) have the same line styles as blue lines (SF) which are indicated in the legend. Shaded areas are the same as in Fig. 5.1. *Bottom panel:* Total number counts for different models obtained by adding SF and AGN contributions.

Table 5.1: Euclidean normalized differential radio number counts n [sr^{-1} Jy$^{1.5}$] at 1.4 GHz as a function of flux density S [Jy].

$\log(S_{1.4\,\text{GHz}})$	$\log(n^{\text{SF}})$	$\log(n^{\text{AGN}})$	$\log(n^{\text{Total}})$
−9.0	−4.44	−4.93	−4.32
−8.5	−3.80	−4.42	−3.71
−8.0	−3.16	−3.92	−3.09
−7.5	−2.52	−3.41	−2.47
−7.0	−1.89	−2.91	−1.85
−6.5	−1.27	−2.40	−1.24
−6.0	−0.713	−1.90	−0.686
−5.5	−0.245	−1.40	−0.216
−5.0	0.0924	−0.911	0.133
−4.5	0.277	−0.428	0.355
−4.0	0.319	0.0387	0.502
−3.5	0.259	0.480	0.685
−3.0	0.174	0.884	0.961
−2.5	0.143	1.24	1.27
−2.0	0.188	1.54	1.56
−1.5	0.181	1.79	1.81
−1.0	−0.0232	2.00	2.01

SKA1 (corresponding to the 10% of the full planned SKA capacity) surveys at 1.4 GHz are located (see Prandoni & Seymour, 2015), the relative fraction is roughly constant with around 90 – 95% of radio sources being SF galaxies. The implication of this result is that a simple flux density cut at $S_{1.4\,\text{GHz}} < 10\,\mu$Jy will select star-forming galaxies yielding a sample with less than 10% contamination by AGN hosts. Applying such a cut on a radio selected sample observed from 1 mJy down to 1 μJy (0.1 μJy) would yield 88% (97%) completeness of the SF sample, with the remaining galaxies having $S_{1.4\,\text{GHz}} > 10\,\mu$Jy, based on the integrated source number counts. The AGN component of a galaxy is compact and star formation is usually smeared across the disk, therefore the above statement assumes that the radio telescope can observe the total integrated emission, which in reality depends on the beam size and the galaxy surface brightness.

We further analyzed our different models of evolving radio LFs and derived the redshift distribution of SF galaxies in different flux density ranges, which we present in Fig. 5.4. Different models can yield similar number counts, but with different intrinsic redshift distributions. The VLA-COSMOS 3 GHz Large Project survey has a median 5σ sensitivity of 11.5 μJy beam^{-1} and around 90% of all radio sources with optical-NIR counterparts fall inside the 10 – 100 μJy
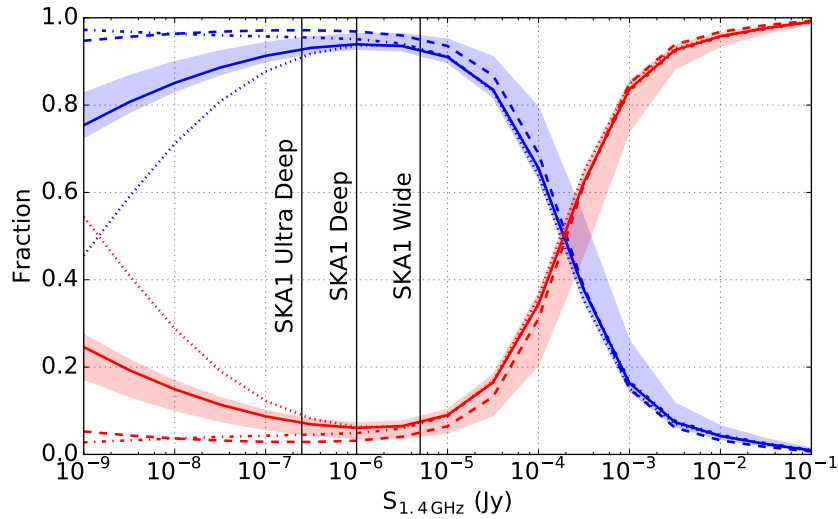
Figure 5.3: Percentages of AGN (red) and SF galaxies (blue) as a function of 1.4 GHz flux density. The legend is the same as in Fig. 5.2.

flux density bin. As can be seen from the figure, a large percentage of our observed galaxies are distributed around the redshift of $z \sim 1$. All our models predict that, with an increase in sensitivity, the bulk of all observed galaxies will shift toward the redshift of $z \sim 2$. Models 1 and 2 have a flatter faint end LF slope for SF galaxies and suggest that, even with an increase of several decades in sensitivity, the peak number of all observed galaxies will remain around $z \sim 2$. On the other hand, if there is a significant trend which favors much larger densities of the faintest galaxies at higher redshifts, such as the model 4, then the peak of the galaxy number distribution also shifts toward higher redshifts with each increase in sensitivity. Regardless of the peak position, at sensitivities below 10 μJy all distributions have extended tails toward higher redshifts and significantly higher overall numbers of galaxies. To give an example, future deep radio observations will yield as many redshift $z \sim 4$ SF galaxies as current surveys find at $z \sim 1$.

5.5 Future surveys across radio bands

Euclidean normalized number counts n at 1.4 GHz reported in Table 5.1 can be converted to any other radio frequency ν under an assumption of a simple power law radio SED with a spectral index α (usually $\alpha = -0.7$) using the following procedure. First, define a conversion factor $f = (1.4/\nu \, [\mathrm{GHz}])^{\alpha}$, then $S_{\nu} = S_{1.4\,\mathrm{GHz}}/f$ and $n_{\nu} = n_{1.4\,\mathrm{GHz}}/f^{1.5}$. While this conversion is probably correct on average for most situations, it might produce biased results in the case when the average radio SED of a galaxy population exhibits a significant contribution of free-free or self-absorption features. However, more deep surveys across the wide radio bandwidth
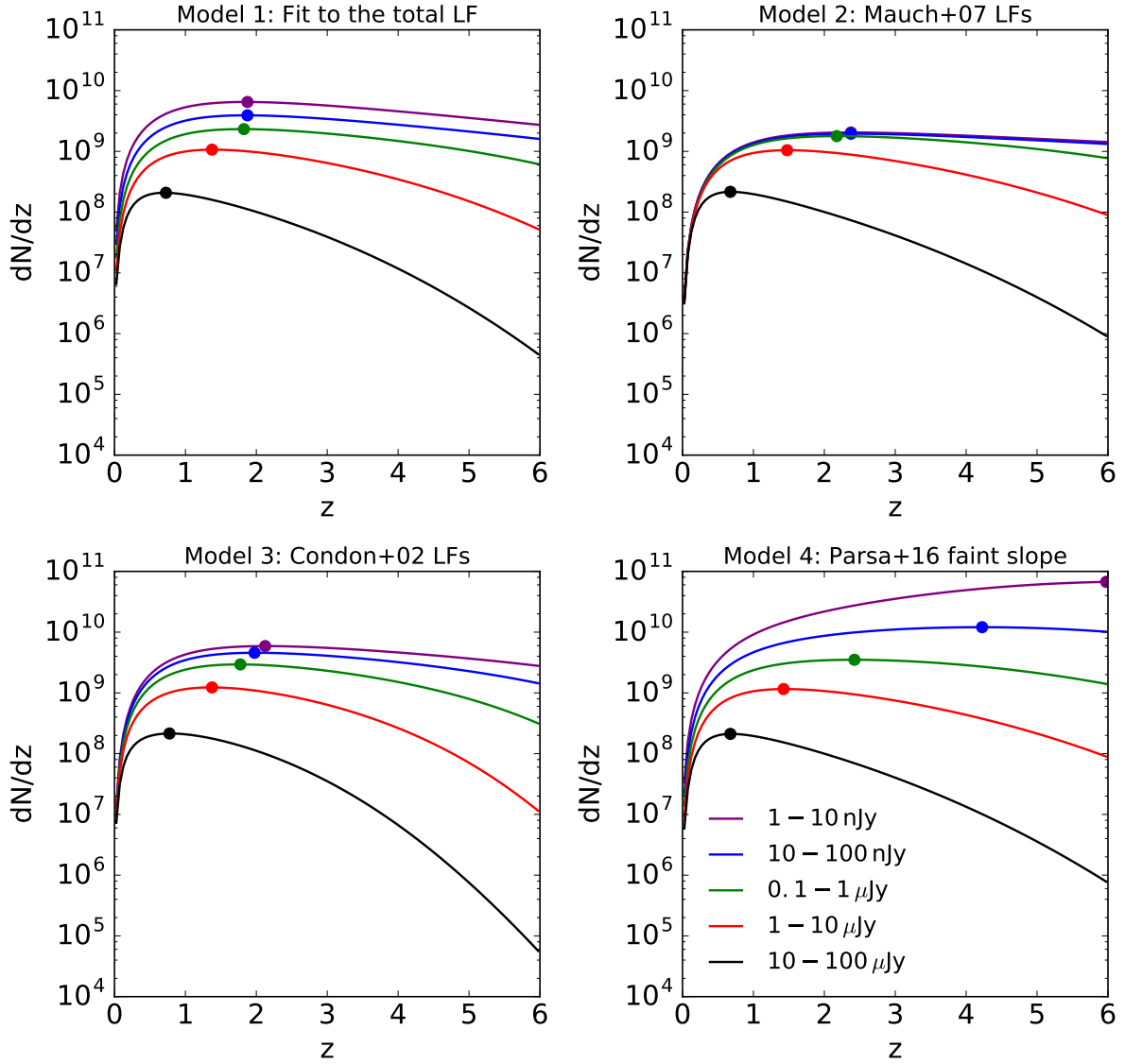
Figure 5.4: Redshift distribution (differential numbers across the full sky, i.e. 4π solid angle) of SF galaxies in different flux density decades based on evolving radio LFs at 1.4 GHz. Filled circles show the peak of each distribution.
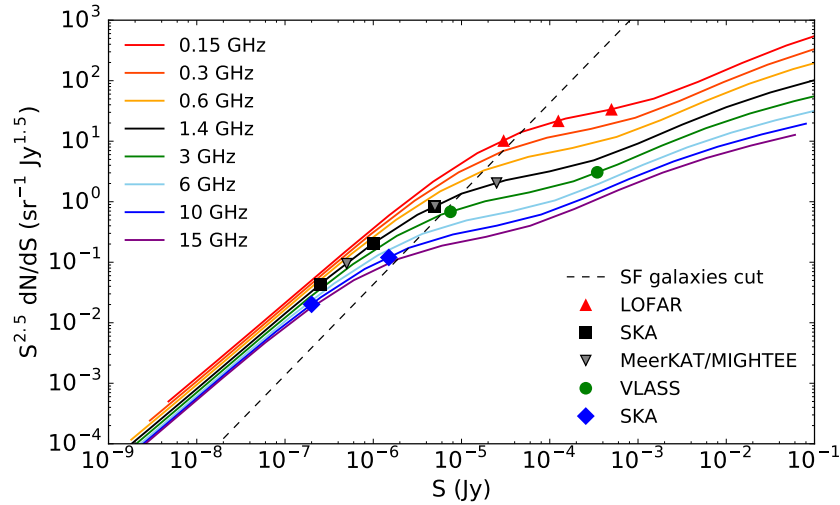
Figure 5.5: Euclidean normalized radio number counts at different frequencies scaled from 1.4 GHz with a simple power law radio SED (see text for details). The dashed line marks the $S_{1.4\,\mathrm{GHz}} = 10\,\mu$Jy cut, below which SF galaxies dominate the number counts. Filled symbols represent the 5σ sensitivity limits for proposed multitiered (ultra deep, deep, wide) future radio surveys.

are needed to properly constrain the average radio SED.

    We plot number counts form our model 1 scaled to different frequencies in Fig. 5.5, where we also show several planned radio continuum surveys according to their proposed 5σ sensitivity limits (see also Norris et al., 2013; Prandoni & Seymour, 2015). All of these surveys have multiple tiers corresponding to a wedding cake observing approach: from wide and shallow observations spanning thousands of square degrees to deep and ultra deep observations, which provide an order of magnitude better sensitivity, but are performed across a small area (usually less than 1 square degree). Near future observations such as the Low-Frequency Array (LO-FAR) survey (see Morganti et al. 2010 and also Shimwell et al. 2017) or the Very Large Array Sky Survey (VLASS[1]) will be able to provide radio based selection of SF galaxies only at their sensitivity limits of the deep observations. The flux density cut for selecting SF galaxies can of course be increased at the cost of the sample purity, for example a $S_{1.4\,\mathrm{GHz}} < 45\,\mu$Jy cut will yield up to 20% contamination by AGN dominated galaxies. The SKA, and also its precursors, will provide a reliable way to classify galaxies as star-forming in a similar manner that the VLBA can confirm AGNs, only across a much larger area of the sky.

---

[1] https://science.nrao.edu/science/surveys/vlass

5.6 Chapter summary

We found that the analytical forms of star-forming and AGN LFs derived in the previous chapter are consistent with Markov chain Monte Carlo fits done on the total LF. We calculated radio number counts from evolution models down to nJy levels. Our LFs can reproduce the number counts obtained from a blind probability of deflection $P(D)$ (Vernstrom et al., 2014) analysis implying that our extrapolations toward the faint end of the LFs are plausible even though it is not constrained directly by the data. Finally, we showed that the planned surveys with SKA will almost exclusively probe star-forming galaxy population at their sensitivity limits and most of the galaxies will be found around $z \sim 2$. Our results suggest that the radio sky between 0.1 and $10\,\mu$Jy at 1.4 GHz is dominated by star formation processes with up to 10% contamination from the AGN population, thus enabling the usage of radio flux densities as an additional diagnostic when classifying galaxies.

Chapter 6

Cosmic star formation history since $z \sim 5$

In this chapter luminosity functions of star-forming galaxies derived in Sect. 4.3 are scaled and integrated to obtain star formation rate densities (SFRDs) using a redshift dependent IR-radio correlation. Dust unbiased cosmic SFRD history estimated from the radio observations is the main result of this thesis. Results are compared with those based on different wavelengths to get an estimate of the amount of star formation enshrouded in dust and undetectable at rest-frame ultraviolet frequencies.

6.1   Tracing the buildup of stellar mass

One of the best methods to follow the buildup of stellar mass through cosmic times relies on inferring the cosmic star formation rate density history (for a review, see Madau & Dickinson, 2014). A consensus is achieved regarding recent history, where an exponential decline in SFRD by one order of magnitude from redshift $z \sim 2$ to the present day is inferred (e.g., Madau et al., 1996; Haarsma et al., 2000; Hopkins et al., 2006). On the other hand, with an increasing number of ultra-deep surveys the detection threshold is continually being pushed to higher redshifts (up to $z \sim 10$) slowly reaching the epoch of reionization (e.g., Bouwens et al., 2014b, 2015). The light of the early galaxies is a major factor in the process of reionization (e.g., Bouwens, 2016), and so accurate SFRD measurements are needed to better understand this epoch.

Although the wealth of observations has increased dramatically in the last decade, we still do not understand the core mechanism that governs SFR histories of individual galaxies. This is because of our inability to actually follow these galaxies throughout their evolution. We observe galaxy populations at different cosmic epochs and try to link them in a consistent way. A picture has emerged from this method in which blue SF galaxies evolve into red quiescent galaxies through ways of quenching, such as rapid gas reservoir depletion after major merger interac-

tions or AGN feedback (e.g., Bell, 2004; Schawinski et al., 2014). On the other hand, Bouché et al. (2010) presented a quenching-free model based on the cosmological decrease of accretion rates with time, which is able to reproduce the observed SFRD. Another model has also been proposed that uses simple mathematical log-normal forms for SFRD and individual SFR history to reproduce a wide range of observed relations (e.g., Gladders et al., 2013; Abramson et al., 2016). When the SFRD history is estimated with sufficient precision it can be used to further constrain semianalytical models of galaxy evolution, thereby deepening our understanding of the underlying physics.

Different SFR tracers can be used over the full electromagnetic spectrum, each with its own benefits and shortcomings (e.g., Kennicutt, 1998). The most direct tracer measures UV light from young massive stars and can be linked with the amount of star formation in the galaxy (e.g., Buat et al., 1989). The rest-frame UV emission is redshifted to optical and IR wavelengths for the most distant galaxies; this enables the usage of very sensitive instruments, such as the HST, to probe this epoch (e.g., Finkelstein et al., 2015). Currently, the SFRD in the earliest cosmic times (age of the universe less than 1 Gyr) is constrained almost exclusively with these kinds of observations (see also Behroozi et al., 2013). However, when measuring the rest-frame UV emission one must correct for dust extinction, which drastically diminishes the UV light. Well-constrained attenuation curves are needed to correct for this effect (e.g., Bouwens et al., 2009).

When dust grains absorb UV light they re-emit it at IR wavelengths. Therefore, FIR and sub-mm traces SFR best when the dust content is high, yielding a large optical depth. These observations can suffer from poor resolution and source blending, although this was mitigated with observations with the *Herschel Space Observatory*. Current observations allow IR surveys to constrain the dust content and SFRs up to redshift $z < 4$ (e.g., Caputi et al., 2005; Rodighiero et al., 2010; Reddy et al., 2012; Gruppioni et al., 2013). Ultraviolet and IR observations can be combined to obtain a more robust hybrid SFR estimator (e.g., Wuyts et al., 2011; Boquien et al., 2016). With the high-resolution sub-mm window opened by the Atacama Large Millimeter/submillimeter Array (ALMA), these wavelengths can be used to probe dusty submillimeter galaxies (SMGs) and their high star formation rates (e.g., Swinbank et al., 2014; Dunlop et al., 2017).

When massive stars undergo supernova explosions, the expanding remnants can accelerate the cosmic ray electrons and give rise to synchrotron radiation, which dominates the radio emission at rest-frame frequencies of $< 30\,\text{GHz}$. The observed nonthermal radio emission offers a dust-unbiased view at sub-arcsecond resolution of star formation processes inside the galaxy, and thus eliminates obscuration, while the high resolution assists counterpart matching (e.g.,

Seymour et al., 2008; Smolčić et al., 2009a). However, it relies heavily on multiwavelength data to provide galaxy redshift and classification due to the featureless shape of the radio spectrum (e.g Condon, 1984). Furthermore, the SFR calibration for radio luminosities is based on the empirical IR-radio correlation to link nonthermal emission with thermal emission (e.g., Helou et al., 1985; Yun et al., 2001; Bell, 2003). This correlation continues to be valid across more than five orders of magnitudes in luminosities and holds at least up to redshift of $z < 2$, albeit with some redshift evolution (e.g., Sargent et al., 2010; Magnelli et al., 2015), and it is likely to be valid even at earlier times up to $z \lesssim 5$ (Delhaize et al., 2017).

From observations and evolutionary models along with evidence presented in previous chapters, we know that SF galaxies dominate the faint end of the radio counts and have strongly evolving luminosity functions, therefore deep surveys are needed to probe this population at early cosmic epochs. However, deep surveys have to sacrifice area in order to be feasible, which makes them more susceptible to cosmic over- and underdensities. This cosmic variance can have a strong redshift-dependent impact to any counting statistic employed (e.g., Moster et al., 2011). Our radio data best traces high-mass ($M_\star > 10^{10}$ M$_\odot$) and highly SF galaxies (SFR > 100 M$_\odot$ yr^{-1}), which would also be classified as ULIRGs and even brighter HyLIRGs. To derive the total SFRD history of the entire radio population in the entire observed redshift range we must rely on extrapolations to lower luminosities below the sensitivity limit.

6.2 Cosmic star formation rate density history

6.2.1 From radio luminosity to star formation rate

Radio emission can be used as an extinction-free tracer of star formation rate when linked to other more direct (thermal) tracers such as the IR light. The first assumption is that the UV photons of massive young stars are absorbed by the dust and re-emitted in the IR so that the total IR emission of a galaxy correlates well with its SFR, which is valid for optically thick galaxies. The conversion factor relies on estimating mass from light and was calibrated by Kennicutt (1998) assuming the Salpeter (1955) IMF ($dN/dM \propto M^{-2.35}$) from 0.1 to 100 M$_\odot$ and is given by

$$\frac{\text{SFR}}{\text{M}_\odot \, \text{yr}^{-1}} = 4.5 \times 10^{-37} \frac{L_{\text{TIR}}}{\text{W}}, \tag{6.1}$$

where L_{TIR} contains the total integrated IR luminosity of a galaxy between $8 - 1\,000 \, \mu m$. This IMF produces more low-mass stars than are supported by observations that favor a turnover below 1 M$_\odot$. Since low-mass stars do not contribute significantly to the total light of the galaxy, only the mass-to-light ratio is changed when the Chabrier (2003) IMF is adopted instead. This

leads to a decrease in SFR by a factor of 1.7 (see also Pozzetti et al., 2007) because of there are fewer low-mass stars created. The calibration itself usually leads to a 0.3 dex scatter on a galaxy basis (see also Condon, 1992; Murphy et al., 2011; Kennicutt & Evans, 2012).

Radio observations can trace recent star formation of galaxies, and can trace these observations on timescales of up to 100 Myr (Condon, 1992). Estimation of a galaxy SFR from the radio data relies heavily on the observational IR-radio correlation that is known to span at least five orders of magnitudes (Helou et al., 1985; Yun et al., 2001). The IR-radio correlation links the radio luminosity to the TIR luminosity via the q_{TIR} parameter defined as

$$q_{\mathrm{TIR}} = \log\left(\frac{L_{\mathrm{TIR}}}{3.75 \times 10^{12}\,\mathrm{W}}\right) - \log\left(\frac{L_{1.4\,\mathrm{GHz}}}{\mathrm{W\,Hz}^{-1}}\right). \tag{6.2}$$

Usually, the q_{TIR} parameter is taken to be a constant value derived for local galaxies. However, recent works suggest that the q_{TIR} value might change with redshift (e.g., Sargent et al., 2010; Magnelli et al., 2015). In this work, we used methods from Delhaize et al. (2017) who constrained the median of the q_{TIR} as a function of redshift using a doubly censored survival analysis for a joint 3 GHz radio and IR-selected sample. They find a decrease of q_{TIR} with redshift that can be parameterized with a simple power law. To be self-consistent, we ran the survival analysis on the same SF galaxy sample as utilized in this work, while also taking into account limits for IR-detected galaxies without a 5σ significant radio emission, because in their work Delhaize et al. (2017) originally used a different sample selection criteria for excluding AGN. The obtained evolution of the IR-radio correlation for our sample can be written as

$$q_{\mathrm{TIR}}(z) = (2.78 \pm 0.02) \times (1+z)^{-0.14 \pm 0.01}. \tag{6.3}$$

The main idea behind the IR-radio correlation is that a linear relation exists between radio and IR luminosities for SF galaxies. There is a possibility that the decreasing $q_{\mathrm{TIR}}(z)$ actually mimics some complexities of the radio SED at high redshifts such as varying degrees of free-free contribution and inverse Compton losses. Inverse Compton losses off the CMB lead to suppression of nonthermal radio continuum emission, which would in turn increase q_{TIR} with redshift (Murphy, 2009), but the opposite trend was observed by Delhaize et al. (2017); see also Ivison et al. (2010); Magnelli et al. (2013). In the case of a more complicated radio SED, a simple power law K correction is not a valid assumption anymore. However, the use of a redshift-dependent $q_{\mathrm{TIR}}(z)$ parameter when calculating SFR should account for these intrinsic observational limitations under the assumption of a linear IR-radio correlation as explained in more detail in Sect. 6.4.4.

Finally, the expression that converts radio luminosity into SFR obtained from the steps de-

scribed above can be written as

$$\frac{\text{SFR}}{\text{M}_\odot \, \text{yr}^{-1}} = f_{\text{IMF}} \times 10^{-24} 10^{q_{\text{TIR}}(z)} \frac{L_{1.4\,\text{GHz}}}{\text{W}\,\text{Hz}^{-1}}, \tag{6.4}$$

where $f_{\text{IMF}} = 1$ for a Chabrier IMF and $f_{\text{IMF}} = 1.7$ for a Salpeter IMF.

6.2.2 Star formation rate density across cosmic times

Integrating below the LF first multiplied by the luminosity we can obtain the total 1.4 GHz radio luminosity density (W Hz^{-1} Mpc^{-3}) in a chosen redshift range. Similarly, if the radio luminosity is converted to SFR as given in Eq. (6.4) before integration, the integral yields the SFRD of a given epoch

$$\text{SFRD} = \int_{L_{\text{min}}}^{L_{\text{max}}} \Phi(L, z, \alpha_{\text{D}}, \alpha_{\text{L}}) \times \text{SFR}(L) \, \text{d}\log L. \tag{6.5}$$

We numerically integrated the above expression by taking the analytical form of the LF in each redshift bin and using the best-fit evolution parameters shown in Fig. 4.7. The integral was calculated in different luminosity ranges, which are listed below (results shown in Fig. 6.1 and also listed in Table 6.1):

1. *Entire luminosity range:* This formally means setting $L_{\text{min}} = 0$ and $L_{\text{max}} \to +\infty$. The integral converges and the major contribution to the SFRD arises from galaxies with luminosities around the turnover of the LF. The entire radio emission is recovered and if the LF shape and evolution is well constrained the SFRD estimate will be as well (within the SFR calibration errors). This is not the case at higher redshifts ($z > 2.5$), where only the bright end of the LF is observed, therefore extrapolation to the faint end can be substantial (see Fig. 4.6).

2. *Data constrained limits:* L_{min} and L_{max} correspond to the lowest and highest value of the observed luminosity function. By choosing integration limits that correspond to the actual data range, any bias due to LF extrapolation toward higher or lower luminosities is removed. The shape of the local LF also does not affect this result within the fitting errors. Numbers obtained from this integration range are a very conservative lower limit on the SFRD.

3. *ULIRGs:* Limits that correspond to galaxies with IR luminosity of $10^{12} \, \text{L}_\odot < L_{\text{TIR}} < 10^{13} \, \text{L}_\odot$ trace ULIRGs. The radio luminosity limits were obtained using an evolving q_{TIR}
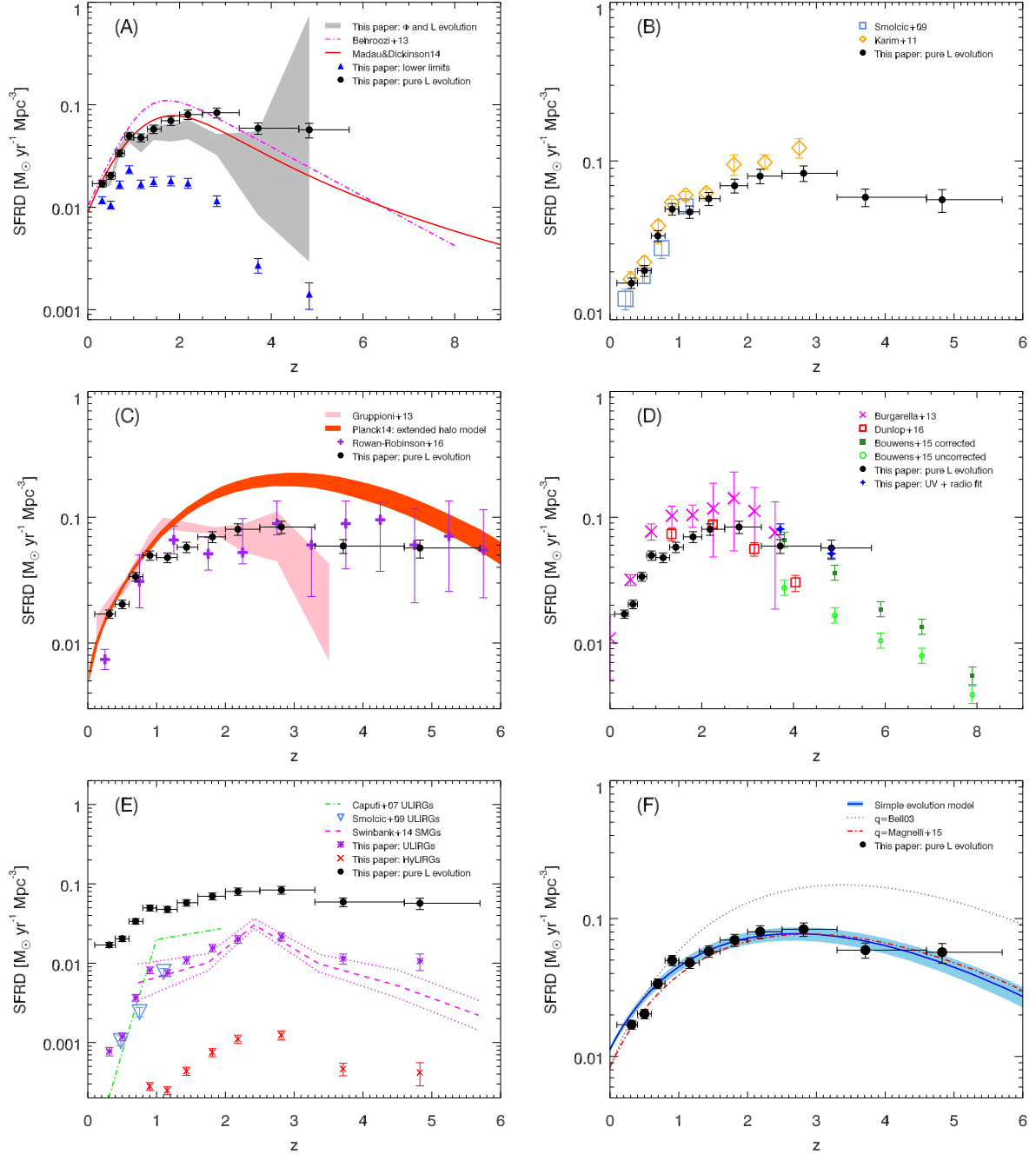
Figure 6.1: Cosmic star formation rate density history. Our total SFRD values estimated from the pure luminosity evolution in separate redshift bins are shown as filled black circles in all panels. All data shown for comparison are indicated in the legend of each panel; see text for details.

Table 6.1: Cosmic SFR density history obtained by integrating the analytical form of the best-fit LF in different redshift bins. All SFRD estimates except the last column refer to pure luminosity evolution. For a combined density and luminosity evolution only the 68% confidence interval is reported.

Redshift	Total	Lower limit	ULIRGs	HyLIRGs	Φ and L evolution
z	$\log(\mathrm{SFRD}[\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}}])$				
$0.312^{+0.088}_{-0.21}$	$-1.77^{+0.032}_{-0.034}$	$-1.93^{+0.033}_{-0.036}$	$-3.11^{+0.052}_{-0.055}$	$-5.05^{+0.069}_{-0.073}$	$[-1.82, -1.74]$
$0.501^{+0.099}_{-0.10}$	$-1.69^{+0.033}_{-0.036}$	$-1.98^{+0.036}_{-0.040}$	$-2.93^{+0.047}_{-0.053}$	$-4.78^{+0.061}_{-0.068}$	$[-1.79, -1.69]$
$0.695^{+0.11}_{-0.095}$	$-1.47^{+0.033}_{-0.036}$	$-1.78^{+0.036}_{-0.040}$	$-2.43^{+0.041}_{-0.045}$	$-4.07^{+0.050}_{-0.056}$	$[-1.50, -1.39]$
$0.903^{+0.097}_{-0.10}$	$-1.30^{+0.035}_{-0.038}$	$-1.63^{+0.037}_{-0.041}$	$-2.09^{+0.040}_{-0.044}$	$-3.55^{+0.046}_{-0.052}$	$[-1.37, -1.27]$
$1.16^{+0.14}_{-0.16}$	$-1.32^{+0.037}_{-0.041}$	$-1.77^{+0.039}_{-0.044}$	$-2.12^{+0.040}_{-0.046}$	$-3.60^{+0.045}_{-0.054}$	$[-1.47, -1.36]$
$1.44^{+0.16}_{-0.14}$	$-1.24^{+0.039}_{-0.043}$	$-1.75^{+0.042}_{-0.046}$	$-1.96^{+0.043}_{-0.047}$	$-3.36^{+0.048}_{-0.053}$	$[-1.34, -1.19]$
$1.81^{+0.19}_{-0.21}$	$-1.16^{+0.041}_{-0.046}$	$-1.74^{+0.043}_{-0.049}$	$-1.81^{+0.043}_{-0.049}$	$-3.12^{+0.047}_{-0.055}$	$[-1.36, -1.19]$
$2.18^{+0.32}_{-0.18}$	$-1.10^{+0.043}_{-0.048}$	$-1.76^{+0.046}_{-0.051}$	$-1.70^{+0.046}_{-0.051}$	$-2.96^{+0.050}_{-0.055}$	$[-1.34, -1.15]$
$2.81^{+0.49}_{-0.31}$	$-1.08^{+0.046}_{-0.052}$	$-1.94^{+0.049}_{-0.057}$	$-1.67^{+0.048}_{-0.055}$	$-2.91^{+0.051}_{-0.061}$	$[-1.49, -1.29]$
$3.71^{+0.89}_{-0.41}$	$-1.23^{+0.051}_{-0.059}$	$-2.57^{+0.065}_{-0.077}$	$-1.94^{+0.059}_{-0.070}$	$-3.33^{+0.071}_{-0.085}$	$[-2.08, -1.26]$
$4.83^{+0.87}_{-0.23}$	$-1.24^{+0.063}_{-0.080}$	$-2.85^{+0.11}_{-0.15}$	$-1.97^{+0.089}_{-0.12}$	$-3.38^{+0.12}_{-0.17}$	$[-2.53, -0.128]$

parameter from Eq. (6.3). The integral with such a range traces SFRD of galaxies that form stars very efficiently (SFR $100 - 1\,000\ \mathrm{M_\odot\,yr^{-1}}$) while also being well constrained by our observational data in $0.5 < z < 3$ range (see also the red dotted vertical line in Fig. 4.6).

4. *HyLIRGs:* Similarly, by integrating over galaxies with radio luminosities that translate into $L_{\mathrm{TIR}} > 10^{13}\ \mathrm{L_\odot}$, we trace HyLIRGs that have extreme star formation, namely SFR $> 1\,000\ \mathrm{M_\odot\,yr^{-1}}$.

Our errors are inferred from the LF fitting parameters uncertainties and added in quadrature with $q_{\mathrm{TIR}}(z)$ parameter errors and do not represent the entire error budget due to LF extrapolations.

6.3 Comparison with the literature

We compare our SFRD with the work in the literature derived at different wavelengths. All SFR estimates were rescaled to a Chabrier IMF where necessary.

6.3.1 Total SFRD estimates

Throughout all panels in Fig. 6.1 we show our total SFRD derived by integrating the pure luminosity evolved LF (from Sect. 4.3) in individual redshift bins as black filled circles. In panel A we compare our SFRD with the curve from the review by Madau & Dickinson (2014) where the fit was performed on a collection of previously published UV and IR data (red line). Below $z < 2$ our data agree well with this compilation of data, but show a turnover at higher redshift ($z \sim 2.5$) with a shallower decline yielding up to $2 - 3$ times larger SFRD at $z \gtrsim 4$. We also plot a slightly different Behroozi et al. (2013) fit to the data compilation in the same panel.

If we allow for both luminosity and density evolution there is a degeneracy of parameters leading to large uncertainties in the total SFRD estimate; the gray shaded area in panel A of Fig. 6.1 is obtained with fit parameters taken from the 1σ significant region in α_D and α_L parameter space. We do not fit a pure density evolution because it would increase the normalization of the LF to very high densities. The SFRD estimates would consequently be significantly higher, making our data completely inconsistent with other works in the literature at intermediate and high redshifts. In the same panel we also show very strict lower limits constrained by the data with blue triangles that demonstrate the amount of extrapolation needed to obtain the total SFRD. Although the extrapolation is significant, especially at higher redshifts, we note that the UV LFs support the need for such large extrapolations.

6.3.2 Comparison with previous radio SFRD

In panel B of Fig. 6.1 we show two radio estimates based on the VLA-COSMOS 20 cm survey (Schinnerer et al., 2007, 2010). Smolčić et al. (2009a) calculated the SFRD by integrating the pure luminosity evolution fit of a local LF taken from Sadler et al. (2002) and their results are shown with blue squares. Also, these estimates were obtained in the COSMOS field and therefore they represent a good consistency check at low redshift. A different approach was taken by Karim et al. (2011) who performed stacking on mass selected galaxies (shown as orange diamonds). They obtained a monotonous rise in the SFRD up to $z \sim 3$. Although the field is the same, their method of estimating SFRD is significantly different from ours since it depends on stacking individually undetected sources. Our estimates are slightly lower than

theirs, with the difference increasing with redshift. This offset is primarily due to a different IR-radio correlation used. They adopt a calibration from Bell (2003), which yields higher SFRD at higher redshifts.

6.3.3 Comparison with IR SFRD

In panel C of Fig. 6.1 the pink shaded area shows the 1σ uncertainty for the SFRD derived from the integrated total IR LF by Gruppioni et al. (2013). This LF has a rising trend up to $z \sim 1.1$ and then flattens out. The highest redshift estimate should be considered as a lower limit because the PEP selection might miss high-z sources. Our results are in broad agreement with theirs. Discrepancies at some redshifts might be attributed to different sample selection, since we are excluding AGN host galaxies classified as such only in the radio (see Sect. 6.4.2). Additionally, the agreement between the IR and the radio SFRD is better at $z > 2$ than at $z \sim 1.5$, while the opposite is true when comparing IR to radio LF (see Fig. 4.6). The reason for this effect is that the normalization of the Gruppioni et al. (2013) IR LF is slightly higher than ours. However, because of the significant negative density evolution and the unchanging faint end slope, this higher normalization is being progressively compensated in their SFRD integral by decreased densities of the faint end. Differing contributions of the faint and the bright end to the total SFRD as a function of redshift lead to apparent agreement between IR and radio SFRD estimates, even though the actual observed LFs do not match perfectly.

We also show the results from the recent work by Rowan-Robinson et al. (2016) as purple plus signs in the plot. Using SED fitting on $\sim 3\,000$ *Herschel* sources from 20.3 square degrees of the sky they derive an IR-based SFRD since $z \sim 6$. Even though their result has large uncertainties, the finding supports a much flatter SFRD trend at high redshifts. It is still consistent, however, with our findings within the error bars. In the same panel we additionally show SFRD results of an extended halo model estimated by Planck Collaboration et al. (2014) from the measured power spectra of the cosmic IR background anisotropies as orange-red shaded area. They report several possible reasons for such high values of SFRD at $z > 2$, and it is important to note that all measurements rely on some form of extrapolation toward fainter galaxies. These results might therefore be considered as upper limits.

6.3.4 UV addition to SFRD estimates

In panel D of Fig. 6.1 we show the recent rest-frame UV estimates from Bouwens et al. (2015) as dark-green filled squares (dust-corrected) and green open circles (dust-uncorrected). The SFRD is scaled to Chabrier IMF and Kennicutt (1998) calibration. Ultraviolet observations like

these are well suited to study the early universe owing to the ability to probe exceptionally high redshifts $z \sim 10$, as also reviewed in Madau & Dickinson (2014).

To simultaneously model both the faint and bright ends of the SF LFs at high redshifts in an attempt to better constrain the SFRD of that epoch we use the dust-corrected UV LFs from Bouwens et al. (2015) along with our own radio LFs and perform a fit on the combined data as explained below. The UV dust corrections are more severe at high luminosities and the Lyman break galaxies selection criteria can easily miss the most massive and dusty galaxies with significant SFRs. On the other hand, the radio emission is an excellent tracer of such SF galaxies. Therefore, we disregard the three most luminous UV LF points at redshifts $z \sim 4$ and $z \sim 5$ and fit an analytical form given in Eq. (4.9) to the remaining UV points combined with all of our radio LFs at the same redshift. The combined data span more than four decades in luminosities. Our results are shown in Fig. 6.2, where we show the SFR on the x-axis instead of the usual luminosity. Ultraviolet luminosities were scaled to SFR according to Kennicutt (1998) relation, while our radio luminosities were scaled using the redshift-dependent q_{TIR} given in Eq. (6.3). It is not our intention to obtain the best SF LF at these redshifts, but rather to calculate an estimate of the missed SFRD in the LBG sample from the radio perspective. Still, for completeness we report here the best-fit parameters obtained. They are $\Phi_\star = 9.35 \times 10^{-3}$ Mpc^{-3}dex^{-1}, $L_\star = 1.81 \times 10^{22}$ W Hz^{-1}, $\alpha = 1.62$, and $\sigma = 0.83$ at $z \sim 4$; and $\Phi_\star = 1.23 \times 10^{-3}$ Mpc^{-3}dex^{-1}, $L_\star = 1.26 \times 10^{23}$ W Hz^{-1}, $\alpha = 1.76$, and $\sigma = 0.67$ at $z \sim 5$.

The SFRD integral of the best LF fit of the combined dust-corrected UV and radio data is 0.08 dex higher at $z \sim 4$ and 0.06 dex higher at $z \sim 5$ than the values obtained from the UV data alone in the same luminosity range. These integrated values are also plotted as blue diamonds in panel D of Fig. 6.1. Assuming the dust corrections calculated by Bouwens et al. (2014a) are correct and start to become significant only at higher luminosities and SFRs (for details, see also Wang & Heckman, 1996), this suggests a 15 – 20% underestimation of highly obscured SFR estimated from the rest-frame UV observations. Since our radio LFs are slightly lower than the IR LFs at $z \sim 4$ (see Fig. 4.6), this underestimation could be considered a lower limit. Also, our pure radio SFRD estimate is likely underestimated at $z \sim 4$ due to a rather flat faint end slope, while at $z \sim 5$ it is actually higher than the combined UV plus radio estimate owing to a higher normalization of the pure evolution fit.

Our radio LFs are in very good agreement with the work done by Mancuso et al. (2016). These authors used the continuity equation approach with the main sequence star formation timescales to conclude that the number density of SF galaxies at high redshifts ($z \lesssim 7$) cannot be reliably estimated from the UV data alone, even when corrected for dust extinction. Their results also imply the existence of a high-redshift heavily dust-obscured galaxy population with
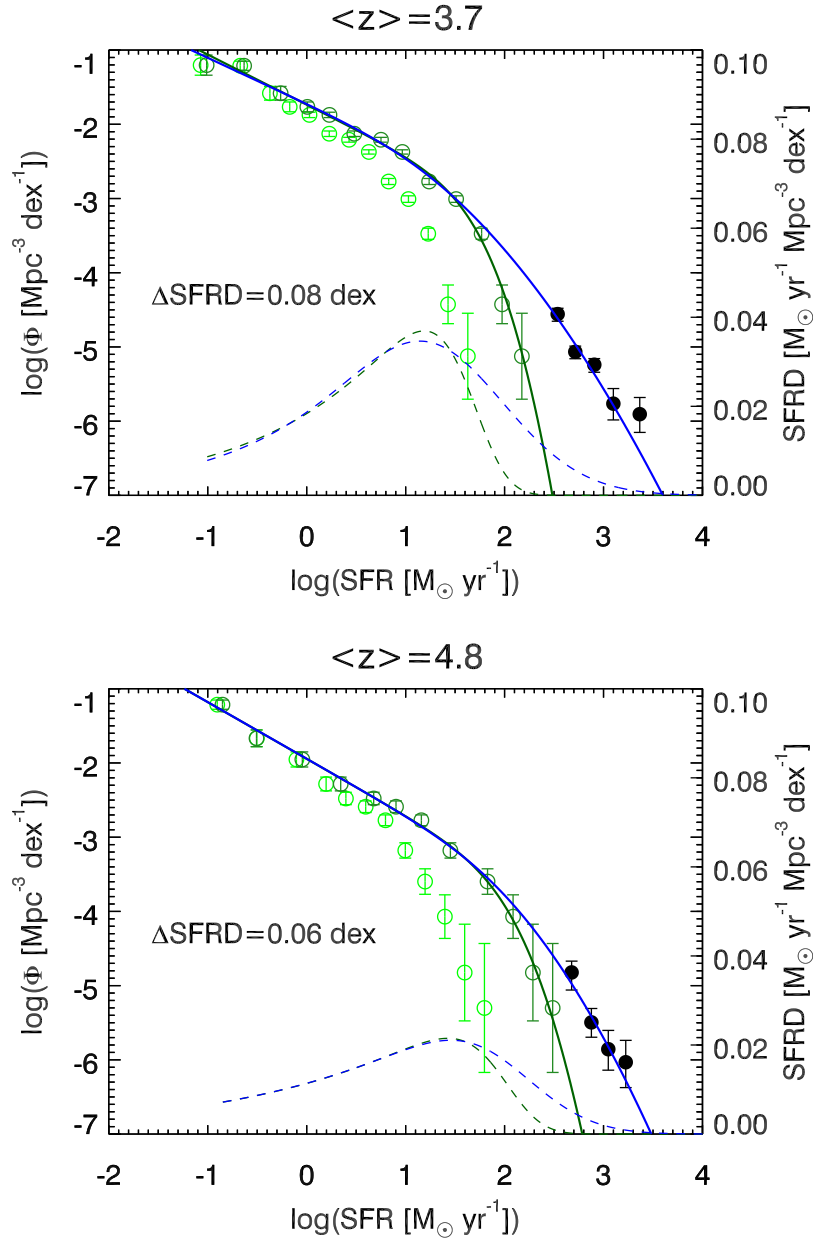
Figure 6.2: Number density of UV (Bouwens et al., 2015) and our radio SF galaxies as a function of SFR in the two highest redshift bins. Dust-corrected (uncorrected) UV data are shown with dark (light) green open circles, and our radio data are shown with filled black circles. A fit with the functional form given in Eq. (4.9) is performed on the UV data only (green full line) and the radio plus faint UV data (blue full line). Dashed lines show the SFRD contribution with the scale given on the right axis. See text for details.

SFRs larger than 100 $M_\odot \, yr^{-1}$.

In their work, Burgarella et al. (2013) attempted to constrain the SFRD by taking into account dust obscuration using combined IR and UV LFs reported in Gruppioni et al. (2013) and Cucciati et al. (2012), respectively. We show their results as magenta crosses in panel D of Fig. 6.1. It is interesting to note a good agreement in SFRD at $z \simeq 4$ between substantially different approaches such as the pure UV-based data, IR plus UV data, and the radio plus UV data. They are all consistent within $\sim 20\%$, but at the same time higher than previously reported SFRD fits (Madau & Dickinson, 2014; Behroozi et al., 2013). Work carried out by Dunlop et al. (2017) is another example that aims at a complete dust-obscured and unobscured (UV + FIR) SFRD census at high redshifts utilizing ALMA observations of the *Hubble* Ultra Deep Field (HUDF) at 1.3 mm. These investigators estimate UV contribution to the total SFR from evolving luminosity functions given in Parsa et al. (2016) and Bouwens et al. (2015). Dunlop et al. (2017) find SFRD (shown as red squares in panel D of Fig. 6.1) consistent with Behroozi et al. (2013) in the redshift range $2.5 < z < 4.5$. They also find a transition in the dominant SF population from dust obscured to dust unobscured at $z \gtrsim 4$. Given the wide distribution and uncertainties of calculated SFRD arising in insufficient knowledge of dust corrections, we believe that the inclusion of radio observations as a dust-unbiased tracer can help achieve a better consensus.

6.3.5 ULIRGs and HyLIRGs

In panel E of Fig. 6.1 we decompose our total SFRD to focus on galaxies that form stars efficiently (SFR $> 100 \, M_\odot \, yr^{-1}$). Our SFRD estimates for ULIRGs and HyLIRGs are shown as purple asterisks and red crosses, respectively. As previously, the first consistency check is to compare our SFRD for ULIRGs with those estimated by Smolčić et al. (2009a), shown as blue downward triangles. Our values are slightly higher than theirs, as was the case with the LFs, in the redshift range sampled by Smolčić et al. (2009a). Ultraluminous IR galaxies are well constrained by our data in the redshift range $0.6 < z < 3.3$, with little to no extrapolation needed. The data show that ULIRGs contribute above 16% to the total SFRD at $z > 1$ with a peak of $\sim 25\%$ at a redshift of $z \sim 2.5$, while HyLIRGs contribute an additional $\lesssim 2\%$ in the entire observed range.

Caputi et al. (2007) inferred the bolometric IR (5 – 1 000 μm) luminosity density up to $z \sim 2$ using *Spitzer* 24 μm selected galaxies in the GOODS fields. We show their SFRD results for ULIRGs only as green dot-dashed line. The agreement is good up to $z \sim 1$, but it gets worse at higher redshifts where their estimates are significantly higher than ours. Discrepancies may be caused by a different star-forming galaxy sample selection as mentioned in the previous

section where we compared IR- and radio-based SFRD. Additionally, the luminosity integration limits needed to calculate contributions from ULIRGs are directly scaled (see Sect. 6.2.2) by the redshift-dependent q_{TIR} parameter. The total scaling effect of the $q_{TIR}(z)$ on the SFRD integral is further discussed in 6.4.4.

Additionally, in the same panel, we show SFRD based on ~ 100 ALESS[1] submillimeter galaxies ($S_{870\mu m} > 1$ mJy) by Swinbank et al. (2014) as the magenta dashed line and 1σ errors as dotted lines. These SMGs are highly dust obscured and have large SFR. Since our lower integration limit for ULIRGs ($100\,M_\odot\,yr^{-1}$) is slightly higher than theirs ($80\,M_\odot\,yr^{-1}$), 0.11 dex should be added to our ULIRG SFRD values prior to comparison with Swinbank et al. (2014) results. There is a broad agreement within the error bars between these two results. However, there are some additional complications in comparing these results because their observations are less sensitive to hotter than average dust temperatures, and they report up to a factor of 2 uncertainty due to missing these ULIRGs. Therefore, their results represent lower limits. Also, both the Swinbank et al. (2014) results and our results rely on non-negligible extrapolations to fainter flux densities.

6.4 Potential biases

Here we summarize some critical assumptions and associated possible systematic effects on our results. While the biggest uncertainties arise from extrapolations, there are a number of additional redshift-dependent and independent effects that may scale our SFRD history.

6.4.1 Radio spectral index

The issues of unconstrained radio SED was already discussed in Sect. 4.3.4. Analysis of LFs and SFRD is repeated two times using $\alpha = -0.7$ and $\alpha = -0.8$ for all sources, Before deriving the SFRD values, the q_{TIR} parameter was recalculated using different spectral indices for consistency reasons. Newly obtained IR-radio correlation is within the 1.5σ of the nominal value given in Eq. (6.3). Derived total SFRDs are within the uncertainties of the nominal sample in both cases, which strengthens the robustness of our results.

---

[1]Atacama Pathfinder Experiment (APEX); Large APEX Bolometer Camera (LABOCA); LABOCA ECDFS Submillimeter Survey (LESS); ALMA LESS (ALESS)

6.4.2 AGN contamination

In this work we adopted an IR-radio-based discrimination of galaxy populations since our goal was to estimate LFs of star-forming galaxies and the total SFRD history from the radio perspective. We assume that the IR is a good tracer of SFR in our radio detected galaxies and that SF galaxies follow a IR-radio correlation with some intrinsic scatter. Because the observed scatter is nonsymmetrical, i.e., there is a tail of sources with large radio fluxes compared to IR measurements, we conclude that AGN contribution to the radio emission is large in such galaxies. The radio-excess method described in Sect. 4.1.3 is therefore good at selecting galaxies dominated by AGN emission in the radio band. The 3σ cut given in Eq. (4.1) ensures that only $\sim 0.15\%$ of removed galaxies are SF, giving us a high level of completeness of our SF sample. On the other hand, by counting sources below the 3σ radio-excess limit and above the best-fitting symmetric profile in all redshift bins (see Sect. 4.1.3) we estimate that the integrated radio emission can be contaminated by some AGN contribution for around 1 000 sources (17% of the sample), and this contribution is limited to a maximum of 80%. This potential AGN contribution is mitigated when calculating the SFRD integral by using a properly calibrated q_{TIR} relation (see Sect. 6.4.4). When AGN enter the sample, they increase the density in the LF, but at the same time lower the q_{TIR} parameter (see Delhaize et al., 2017). If a smaller, less than 3σ, cut were used, then more and more SF galaxies would be removed from the sample trading completeness for purity.

In an attempt to obtain a clean SF sample, free of AGN hosts, we employed a different selection method explained in Smolčić et al. (2017b). We start from the full radio sample with optical-NIR counterparts. The first step in removing AGN includes the use of a cutoff in the X-ray full band (rest-frame $0.5 - 8$ keV) luminosity (see Szokoly et al., 2004). In the second step, a warm dusty torus signature around the supermassive black hole is found in the MIR using a cut in the four IRAC bands as prescribed in Donley et al. (2012). The third step uses SED fits with AGN templates (da Cunha et al., 2008; Berta et al., 2013) to exclude galaxies with significant AGN contribution (see also Delvecchio et al., 2017). These three criteria remove moderate-to-high radiative luminosity AGN from the sample. The next step uses rest-frame optical colors $M_{NUV} - M_r$ corrected for internal dust extinction to select red quiescent galaxies ($M_{NUV} - M_r > 3.5$, Ilbert et al. 2010) that may host an AGN detected in the radio. If galaxies with such colors do not have a 5σ detection in the *Herschel* image, they are then classified as low-to-moderate radiative luminosity AGN and excluded from our sample. The remaining 4 555 green and blue galaxies ($M_{NUV} - M_r \leq 3.5$) without a 3σ radio excess (see Sect. 4.1.3) are considered a clean sample of SF galaxies based on available AGN diagnostics. Since this sample does not account for the star formation in AGN hosts, it represents a conservative lower

limit to SF LFs and SFRD. When the analysis is repeated with this clean SF sample, we find a median decrease of 0.12 dex in number densities across all observed epochs without significant redshift trends. For consistency, the q_{TIR} was recalculated for the clean SF galaxy sample. It gives slightly higher values at all redshifts, but agrees with that given in Eq. (6.3) within 1.5σ. The total SFRD integral median decrease is 0.035 dex, which is within the uncertainties of our nominal sample. The SFRD median decrease is not as significant as the LF median density decrease because we are still fitting the pure luminosity evolution to newly derived LFs, meaning that the faint end remains mostly unchanged; we also recalibrated the q_{TIR} parameter to match the clean SF sample.

6.4.3   The choice of the local LF

The choice of the analytical shape of the star-forming LF can have a significant effect on the total SFRD due to extrapolation toward unobserved luminosities. A compilation of local LF is shown in the top panel Fig. 6.3. Where necessary the LF was corrected for the change of cosmology by scaling $L(z) \propto d_m^2(z) \propto h^{-2}$ and $\Phi(z) \propto d_m^{-3}(z) \propto h^3$, according to Ascasibar et al. (2002). The bottom panel of Fig. 6.3 shows the contribution to the SFRDs as a function of luminosity for the various LFs; although all the LFs show a peak at a similar radio luminosity, the positions of these peaks can differ by up to ~ 0.3 dex. There is also no physical argument for the shape of the LF being fix and evolving in redshift by simple translation. However, our data cannot constrain the full luminosity range required to obtain the most significant bulk of the SFRD integral at all redshifts, so this way of extrapolation was chosen for its simplicity. For a more complex handling of the LF evolution, see for example Fotopoulou et al. (2016), where they used a Bayesian approach to model and constrain the shape of the AGN LF as a function of redshift.

6.4.4   IR-radio correlation

The most significant factor in our SFRD estimates is the q_{TIR} parameter since it directly scales our integrated radio luminosities as a function of redshift. Throughout this work we used $q_{\mathrm{TIR}}(z)$ estimated on the same SF galaxy sample with the methods from Delhaize et al. (2017). The following are a few underlying assumptions when using such an evolving $q_{\mathrm{TIR}}(z)$ value:

1. The IR emission is an accurate tracer of SFR at all redshifts and radio emission originates mostly in SF processes. Extragalactic radio observations can properly trace emission from SF processes in a galaxy when cosmic ray electrons are not allowed to escape it. The escape scenario is possible for small sized galaxies with $L_{1.4\,\mathrm{GHz}} \lesssim 2 \times 10^{21}\,\mathrm{W\,Hz^{-1}}$
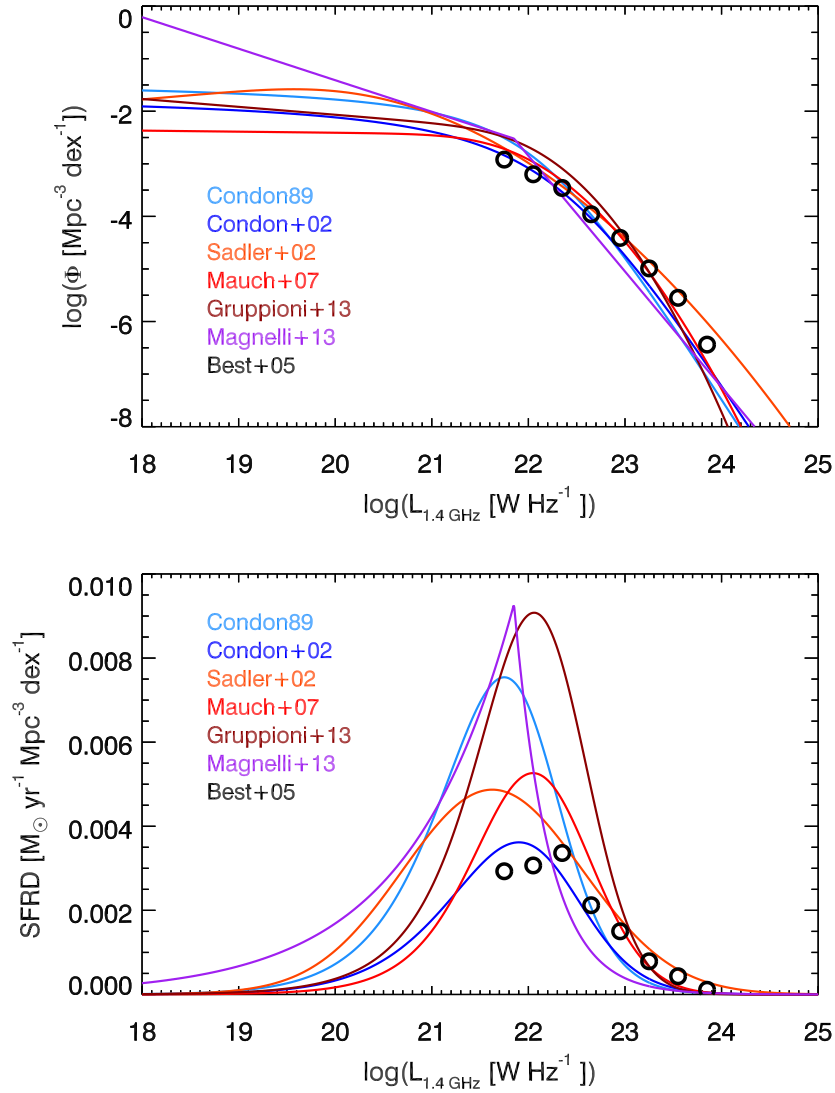
Figure 6.3: *Top panel*: Local radio and IR LF at 1.4 GHz from various authors as indicated in the legend. Red and dark red lines correspond to IR data while all other lines are derived from radio data. Best et al. (2005) did not attempt to fit an analytical form so we show their points as black circles. Functional forms are either broken power law (purple), hyperbolic form (two blue colors), or power law plus log-normal (three red colors). *Bottom panel*: LFs converted to SFRD per logarithm of luminosity using Eq. (6.4) and a local q_{TIR}=2.64 value from Bell (2003).

(e.g., Bell, 2003), which is far below our observational limit. However, the nonthermal radio emission needs a proxy to derive the actual SFR value and the assumption is that the IR emission is a good proxy.

2. Infrared-radio correlation is linear, meaning that it can be represented as a single line with a slope of one in the log-log plot of radio and IR luminosities.

3. Radio spectrum is a simple power law in frequency. This is a widely used approximation and is often taken for granted because of insufficient radio data, however, it plays an important role, especially at high redshifts.

Within the framework of these assumptions it is correct to use an evolving $q_{\mathrm{TIR}}(z)$ when calculating the SFR of a galaxy from radio emission. Even if the second or the third assumption was not correct, for example, because of various free-free contributions in the radio spectrum or the luminosity dependence of the IR-radio correlation, which might cause a difference between the IR and radio LF evolution, the $q_{\mathrm{TIR}}(z)$ evolution takes these wrong assumptions into account and produces a correct SFR value on average because it was calibrated using both the radio and IR data.

To demonstrate the scaling effect of the q_{TIR} parameter on our SFRDs we integrate our continuous simple evolution model from Sect. 4.3.2 and show the results with a blue line in panel F of Fig. 6.1, while the shaded area corresponds to the 1σ uncertainty owing to the errors on the fit parameters added in quadrature with the $q_{\mathrm{TIR}}(z)$ uncertainty. If we instead take the standard constant local value of $q_{\mathrm{TIR}} = 2.64$ from Bell (2003) and apply it to our simple LF evolution model, we would obtain three times larger SFRD estimates at $z \sim 4$ (see gray dotted line in the same panel). Observations however do not favor this choice. Another analysis of the IR-radio correlation was performed through stacking by Magnelli et al. (2015). They obtained $q_{\mathrm{FIR}}(z) = 2.35 \times (1+z)^{-0.12}$. This relation can be scaled as $\log(L_{\mathrm{FIR}}) = \log(L_{\mathrm{TIR}}) - \log(2)$ to obtain the $q_{\mathrm{TIR}}(z)$ needed for our conversion, which is valid in terms of median statistics; see also Delhaize et al. (2017). The SFRD obtained from this expression is shown as a red dot-dashed line in the same panel and is similar to ours. To summarize, the trend in the cosmic SFRD history that we obtain from our simple LF evolution model is linked with the trend in the q_{TIR} and it is important for this value to be well constrained at all observed redshifts.

6.5 Chapter summary

We studied a radio-selected sample of 5 915 galaxies star-forming galaxies from deep VLA-COSMOS 3 GHz observations, where the radio emission is not dominated by an AGN. We con-

verted our radio luminosities to SFR using a redshift-dependent IR-radio correlation where the q_{TIR} parameter decreases with increasing redshift (Delhaize et al., 2017). An accurate constraint on this parameter is the most important factor for estimating SFR from radio observations in the early universe. Our data suggest that the peak of the total SFRD history occurs at $2 < z < 3$. We find that the total SFRD estimates using only LBG galaxies (e.g., Bouwens et al., 2015), even if corrected for dust extinction, are still likely to miss up to $15 - 20\%$ of SFR in highly obscured galaxies at $z \gtrsim 4$. By integrating LF fits in various luminosity limits we estimated SFRDs of the total SF sample and the subpopulations of the sample, such as ULIRGs and HyLIRGs. We find that ULIRGs contribute at maximum up to $\sim 25\%$ of the total SFRD at $z \sim 2.5$, where this population of galaxies is well constrained by our data. Even though HyLIRGs can have very large SFRs (several $1\,000\,\text{M}_\odot\,\text{yr}^{-1}$), we find that they contribute less than 2% to the total SFRD at all redshifts owing to their low volume density.

Chapter 7

Thesis summary

Statistical galaxy evolution studies using radio astronomy were presented in this thesis. Radio waves (wavelengths larger than 1 cm) are a combination of synchrotron emission from cosmic electrons gyrating in galactic magnetic fields and braking radiation (Bremsstrahlung or free-free) from electrons deflecting off protons in interstellar regions of ionized gas. The basic picture of galaxy evolution and ways to study it using radio astronomy were described in Chapter 1. To observe radio emission from a statistically significant number of galaxies the VLA-COSMOS 3 GHz Large Project was approved and awarded 384 hours of observations with the recently upgraded Karl G. Jansky Very Large Array interferometer.

When one third of the observations were complete, initial tests of interferometric data reduction were performed as presented in Chapter 2. Different algorithms for imaging a 2 GHz bandwidth radio continuum dataset were tested and compared in terms of image quality, sensitivity and resolution. This data was used to analyze an interesting galaxy CID-42, which is a candidate for a recoiling black hole. The theory of general relativity and galaxy dynamics allow for a specific event to occur when two galaxies merge. As black holes residing in centers of merging galaxies coalesce, they emit gravitational waves anisotropically. Obeying the linear momentum conservation law, the energy emitted in one direction can provide enough force to eject the newly formed supermassive black hole out of the host galaxy due to recoil. Even with the new radio data the recoil hypothesis for CID-42 could not be confirmed, nor rejected. Nevertheless, new constraints were provided on the radio spectral energy distribution of CID-42. The results suggest that in the recoiling scenario the total radio emission might be composed of an active galactic nucleus surrounded by an aged electron population or outflows.

The full observed dataset contained 192 pointings covering approximately 2.6 square degrees centered at the COSMOS field and was presented in Chapter 3. The data was calibrated and each pointing was imaged with the multifrequency synthesis algorithm yielding two maps,

one for monochromatic flux densities and one for spectral indices. Additional steps were taken to ensure good quality of images, such as further flagging of poor data and self-calibration where sufficiently bright sources were present. The final mosaic has a sensitivity of 2.3 μJy beam^{-1}, an excellent resolution of $0.75''$ and an astrometric accuracy of $0.01''$ for bright sources. A blind search was performed to catalog radio components by finding blobs of pixels 5σ above the noise and integrating the flux density using a flood fill algorithm. Manual inspection of some components was required to join non-connected blobs (e.g. core plus two radio lobes) into 10 830 radio sources which constitute the final 3 GHz radio catalog. Completeness and bias corrections were performed. These are essential for any subsequent statistical analysis of the radio sample. They were based on false detection rate estimates and Monte Carlo simulations where mock sources (generated in a meaningful way) were inserted into the map and then re-extracted.

The final radio catalog is the starting point for galaxy evolution studies as shown in Chapter 4. First, the radio data was paired with the COSMOS ancillary data providing redshift information and rich radio to X-ray photometry that can be used for galaxy classification. The galaxies were distributed into two categories. The first one is the active galactic nuclei, where galaxies exhibit a 3σ significant radio excess compared to their infrared bolometric luminosity. Such an excess implies that at least 80% of the radio emission is due to the central supermassive black hole accretion. The complementary population is composed of star-forming galaxies, where the radio emission is mostly fueled by supernova remnants. Luminosity functions were computed for these galaxy populations providing constraints on their cosmic evolution since redshift of $z \sim 5$. Luminosity functions across cosmic time were then fitted with redshift dependent pure evolution models that allow for the extrapolation of data outside the observed range. The data suggests strong evolution of both radio populations, resulting in peak number densities of star-forming galaxies around $z \sim 2.5$ and somewhat later for active galactic nuclei, around $z \sim 1.5$. The total radio luminosity function was also fitted with analytical expressions from different models and then used to obtain number counts down to nano-Jansky flux densities.

The results suggest that, in more than 90% of the cases, galaxies with 1.4 GHz radio flux densities below 10 μJy are star-forming galaxies. These results can be used to help constrain semi-empirical models of galaxy evolution and to predict numbers of galaxies that will be observed by future surveys. Next generation deep and ultra-deep radio surveys with the Square Kilometer Array will dominantly probe star formation properties of galaxies.

The radio emission provides a high resolution (sub-arcsecond) and dust-unbiased (insensitive to dust extinction) view of the galaxies, which can be linked with star formation rates if the active galactic nuclei component is not contributing significantly to the total radio luminosity. Following the idea that the bolometric infrared luminosity traces recent star formation,

and using the empirical evidence of an existing infrared-radio correlation, integrated radio luminosities of star-forming galaxies were scaled in Chapter 6 to obtain the cosmic star formation rate density history since $z \sim 5$. The results suggest that the peak of cosmic star-forming activity was occurring between $z \sim 2 - 3$, roughly 10 billion years ago. Around 25% of new stars were being formed in ultra luminous infrared galaxies at this epoch. A comparison with the rest-frame ultraviolet data, which traces young bright stars but is severely affected by extinction, suggests that up to 20% of star formation is not recovered from such dust-obscured observations at redshifts $z \sim 4$.

Luminosity functions, number counts, and the cosmic star formation rate density history can enter cosmological simulations as parameters, or provide observational constraints against which the results of simulations can be compared. Simulations can test different physical scenarios, while the observations distinguish which are the most probable ones and closest to reality. Radio astronomy thus provides a unique window into the hidden physics of the galaxies and one aspect of its application was covered in this thesis. It is through the synergy of theoretical simulations and empirical observations that the advancement of our understanding of the cosmos is deepened.

References

Abramson L. E., Gladders M. D., Dressler A., Oemler Jr. A., Poggianti B., Vulcani B., 2016, ApJ, 832, 7, *Return to [Log-]Normalcy: Rethinking Quenching, The Star Formation Main Sequence, and Perhaps Much More*

Afonso J., Georgakakis A., Almeida C., Hopkins A. M., Cram L. E., Mobasher B., Sullivan M., 2005, ApJ, 624, 135, *The Phoenix Deep Survey: Spectroscopic Catalog*

Aihara H., et al., 2011, ApJS, 193, 29, *The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III*

Aird J., et al., 2012, ApJ, 746, 90, *PRIMUS: The Dependence of AGN Accretion on Host Stellar Mass and Color*

Alfvén H., Herlofson N., 1950, Physical Review, 78, 616, *Cosmic Radiation and Radio Stars*

Amsler C., et al., 2008, Physics Letters B, 667, 1, *Review of Particle Physics*

Antonucci R., 1993, ARA&A, 31, 473, *Unified models for active galactic nuclei and quasars*

Aretxaga I., et al., 2011, MNRAS, 415, 3831, *AzTEC millimetre survey of the COSMOS field - III. Source catalogue over 0.72 deg^2 and plausible boosting by large-scale structure*

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540, *Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North*

Ascasibar Y., Yepes G., Gottlöber S., Müller V., 2002, A&A, 387, 396, *Numerical simulations of the cosmic star formation history*

Avni Y., 1976, ApJ, 210, 642, *Energy spectra of X-ray clusters of galaxies*

Baldry I. K., Balogh M. L., Bower R., Glazebrook K., Nichol R. C., 2004, in Allen R. E., Nanopoulos D. V., Pope C. N., eds, American Institute of Physics Conference Series Vol. 743, The New Cosmology: Conference on Strings and Cosmology. pp 106–119, doi:10.1063/1.1848322

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5, *Classification parameters for the emission-line spectra of extragalactic objects*

Baloković M., Smolčić V., Ivezić Ž., Zamorani G., Schinnerer E., Kelly B. C., 2012, ApJ, 759, 30, *Disclosing the Radio Loudness Distribution Dichotomy in Quasars: An Unbiased Monte Carlo Approach Applied to the SDSS-FIRST Quasar Sample*

Banerji M., et al., 2010, MNRAS, 406, 342, *Galaxy Zoo: reproducing galaxy morphologies via machine learning*

Batcheldor D., Robinson A., Axon D. J., Perlman E. S., Merritt D., 2010, ApJ, 717, L6, *A Displaced Supermassive Black Hole in M87*

Baugh C. M., Cole S., Frenk C. S., 1996, MNRAS, 283, 1361, *Evolution of the Hubble sequence in hierarchical models for galaxy formation*

Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559, *The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters*

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57, *The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8*

Bekenstein J. D., 1973, ApJ, 183, 657, *Gravitational-Radiation Recoil and Runaway Black Holes*

Bell E. F., 2003, ApJ, 586, 794, *Estimating Star Formation Rates from Infrared and Radio Luminosities: The Origin of the Radio-Infrared Correlation*

Bell E. F., 2004, ArXiv Astrophysics e-prints, *Galaxy Assembly*

Bennett C. L., et al., 2013, ApJS, 208, 20, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*

Berta S., et al., 2013, A&A, 551, A100, *Panchromatic spectral energy distributions of Herschel sources*

Bertoldi F., et al., 2007, ApJS, 172, 132, *COSBO: The MAMBO 1.2 Millimeter Imaging Survey of the COSMOS Field*

Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9, *A sample of radio-loud active galactic nuclei in the Sloan Digital Sky Survey*

Best P. N., Kaiser C. R., Heckman T. M., Kauffmann G., 2006, MNRAS, 368, L67, *AGN-controlled cooling in elliptical galaxies*

Best P. N., Ker L. M., Simpson C., Rigby E. E., Sabater J., 2014, MNRAS, 445, 955, *The cosmic evolution of radio-AGN feedback to z = 1*

Béthermin M., et al., 2012, ApJ, 757, L23, *A Unified Empirical Model for Infrared Galaxy Counts Based on the Observed Physical Evolution of Distant Galaxies*

Bianchi S., Piconcelli E., Pérez-Torres M. Á., Fiore F., La Franca F., Mathur S., Matt G., 2013, MNRAS, 435, 2335, *The NGC 3341 minor merger: a panchromatic view of the active galactic nucleus in a dwarf companion*

Blecha L., Cox T. J., Loeb A., Hernquist L., 2011, MNRAS, 412, 2154, *Recoiling black holes in merging galaxies: relationship to active galactic nucleus lifetimes, starbursts and the M_{BH}-σ_* relation*

Blecha L., Civano F., Elvis M., Loeb A., 2013, MNRAS, 428, 1341, *Constraints on the nature of CID-42: recoil kick or supermassive black hole pair?*

Bock D. C.-J., Large M. I., Sadler E. M., 1999, AJ, 117, 1578, *SUMSS: A Wide-Field Radio Imaging Survey of the Southern Sky. I. Science Goals, Survey Design, and Instrumentation*

Bogdán Á., et al., 2015, ApJ, 804, 72, *Hot Gaseous Coronae around Spiral Galaxies: Probing the Illustris Simulation*

Bolton J. G., 1982, Proceedings of the Astronomical Society of Australia, 4, 349, *Radio astronomy at Dover Heights*

Bolton J. G., Stanley G. J., Slee O. B., 1949, Nature, 164, 101, *Positions of Three Discrete Sources of Galactic Radio-Frequency Radiation*

Bondi M., et al., 2003, A&A, 403, 857, *The VLA-VIRMOS Deep Field. I. Radio observations probing the mu Jy source population*

Bondi M., et al., 2007, A&A, 463, 519, *The VVDS-VLA deep field. III. GMRT observations at 610 MHz and the radio spectral index properties of the sub-mJy population*

Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, ApJ, 681, 1129, *The VLA-COSMOS Survey. III. Further Catalog Analysis and the Radio Source Counts*

Bongiorno A., et al., 2016, A&A, 588, A78, *AGN host galaxy mass function in COSMOS. Is AGN feedback responsible for the mass-quenching of galaxies?*

Bonning E. W., Shields G. A., Salviander S., 2007, ApJ, 666, L13, *Recoiling Black Holes in Quasars*

Bonometto S., Gorini V., Moschella U., eds, 2002, *Modern cosmology* Institute of Physics Publishing Bristol and Philadelphia

Bonzini M., et al., 2012, ApJS, 203, 15, *The Sub-mJy Radio Population of the E-CDFS: Optical and Infrared Counterpart Identification*

Bonzini M., Padovani P., Mainieri V., Kellermann K. I., Miller N., Rosati P., Tozzi P., Vattakunnel S., 2013, MNRAS, 436, 3759, *The sub-mJy radio sky in the Extended Chandra Deep Field-South: source population*

Boquien M., et al., 2016, A&A, 591, A6, *Towards universal hybrid star formation rate estimators*

Borch A., et al., 2006, A&A, 453, 869, *The stellar masses of 25 000 galaxies at $0.2 \leq z \leq 1.0$ estimated by the COMBO-17 survey*

Bouché N., et al., 2010, ApJ, 718, 1001, *The Impact of Cold Gas Accretion Above a Mass Floor on Galaxy Scaling Relations*

Bourke S., Mooley K., Hallinan G., 2014, in Manset N., Forshay P., eds, Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII. p. 367

Bouwens R., 2016, in Mesinger A., ed., Astrophysics and Space Science Library Vol. 423, Astrophysics and Space Science Library. p. 111, doi:10.1007/978-3-319-21957-8_4

Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928, *UV Luminosity Functions at z~4, 5, and 6 from the Hubble Ultra Deep Field and Other Deep Hubble Space Telescope ACS Fields: Evolution and Star Formation History*

Bouwens R. J., et al., 2009, ApJ, 705, 936, *UV Continuum Slope and Dust Obscuration from z ~ 6 to z ~ 2: The Star Formation Rate Density at High Redshift*

Bouwens R. J., et al., 2014a, ApJ, 793, 115, *UV-continuum Slopes of >4000 z ~ 4-8 Galaxies from the HUDF/XDF, HUDF09, ERS, CANDELS-South, and CANDELS-North Fields*

Bouwens R. J., et al., 2014b, ApJ, 795, 126, *A Census of Star-forming Galaxies in the Z ~ 9-10 Universe based on HST+Spitzer Observations over 19 Clash Clusters: Three Candidate Z ~ 9-10 Galaxies and Improved Constraints on the Star Formation Rate Density at Z ~ 9.2*

Bouwens R. J., et al., 2015, ApJ, 803, 34, *UV Luminosity Functions at Redshifts z ~ 4 to z ~ 10: 10,000 Galaxies from HST Legacy Fields*

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645, *Breaking the hierarchy of galaxy formation*

Brusa M., et al., 2009, ApJ, 693, 8, *High-Redshift Quasars in the COSMOS Survey: The Space Density of z > 3 X-Ray Selected QSOs*

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000, *Stellar population synthesis at the resolution of 2003*

Buat V., Donas J., Milliard B., 1989, Ap&SS, 157, 197, *Ultraviolet (200 nm) observations of the nearby cluster of galaxies Abell 1367*

Budavári T., et al., 2003, ApJ, 595, 59, *Angular Clustering with Photometric Redshifts in the Sloan Digital Sky Survey: Bimodality in the Clustering Properties of Galaxies*

Burbidge G. R., 1956, ApJ, 124, 416, *On Synchrotron Radiation from Messier 87.*

Burgarella D., et al., 2013, A&A, 554, A70, *Herschel PEP/HerMES: the redshift evolution ($0 \leq z \leq 4$) of dust attenuation and of the total (UV+IR) star formation rate density*

Burke B. F., Graham-Smith F., 2010, *An Introduction to Radio Astronomy*, 3rd edn. Cambridge University Press

Callingham J. R., et al., 2017, ApJ, 836, 174, *Extragalactic Peaked-spectrum Radio Sources at Low Frequencies*

Capak P., et al., 2007, ApJS, 172, 99, *The First Release COSMOS Optical and Near-IR Data and Catalog*

Caputi K. I., Dunlop J. S., McLure R. J., Roche N. D., 2005, MNRAS, 361, 607, *The evolution of K_s-selected galaxies in the GOODS/CDFS deep ISAAC field*

Caputi K. I., et al., 2007, ApJ, 660, 97, *The Infrared Luminosity Function of Galaxies at Redshifts z = 1 and z ~ 2 in the GOODS Fields*

Chabrier G., 2003, PASP, 115, 763, *Galactic Stellar and Substellar Initial Mass Function*

Ciesla L., et al., 2015, A&A, 576, A10, *Constraining the properties of AGN host galaxies with spectral energy distribution modelling*

Ciliegi P., et al., 1999, MNRAS, 302, 222, *A deep VLA survey at 20 CM of the ISO ELAIS survey regions*

Civano F., et al., 2010, ApJ, 717, 209, *A Runaway Black Hole in COSMOS: Gravitational Wave or Slingshot Recoil?*

Civano F., et al., 2012a, ApJS, 201, 30, *The Chandra COSMOS Survey. III. Optical and Infrared Identification of X-Ray Point Sources*

Civano F., et al., 2012b, ApJ, 752, 49, *Chandra High-resolution observations of CID-42, a Candidate Recoiling Supermassive Black Hole*

Civano F., et al., 2016, ApJ, 819, 62, *The Chandra Cosmos Legacy Survey: Overview and Point Source Catalog*

Clewley L., Jarvis M. J., 2004, MNRAS, 352, 909, *The cosmic evolution of low-luminosity radio sources from the Sloan Digital Sky Survey Data Release 1*

Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393, *Colors and magnitudes predicted for high redshift galaxies*

Colpi M., Callegari S., Dotti M., Mayer L., 2009, Classical and Quantum Gravity, 26, 094029, *Massive black hole binary evolution in gas-rich mergers*

Comerford J. M., Griffith R. L., Gerke B. F., Cooper M. C., Newman J. A., Davis M., Stern D., 2009, ApJ, 702, L82, *1.75 h $^{-1}$ kpc Separation Dual Active Galactic Nuclei at z = 0.36 in the Cosmos Field*

Condon J. J., 1984, ApJ, 287, 461, *Cosmological evolution of radio sources*

Condon J. J., 1989, ApJ, 338, 13, *The 1.4 gigahertz luminosity function and its evolution*

Condon J. J., 1992, ARA&A, 30, 575, *Radio emission from normal galaxies*

Condon J. J., 2015a, preprint, *An Analysis of the VLASS Proposal*

Condon J., 2015b, in The Many Facets of Extragalactic Radio Surveys: Towards New Scientific Challenges. p. 4

Condon J. J., Ransom S. M., 2016, *Essential Radio Astronomy*. Princeton University Press

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693, *The NRAO VLA Sky Survey*

Condon J. J., Cotton W. D., Broderick J. J., 2002, AJ, 124, 675, *Radio Sources and Star Formation in the Local Universe*

Condon J. J., Cotton W. D., Yin Q. F., Shupe D. L., Storrie-Lombardi L. J., Helou G., Soifer B. T., Werner M. W., 2003, AJ, 125, 2411, *The SIRTF First-Look Survey. I. VLA Image and Source Catalog*

Condon J. J., et al., 2012, ApJ, 758, 23, *Resolving the Radio Source Background: Deeper Understanding through Confusion*

Conroy C., 2013, ARA&A, 51, 393, *Modeling the Panchromatic Spectral Energy Distributions of Galaxies*

Corral A., Della Ceca R., Caccianiga A., Severgnini P., Brunner H., Carrera F. J., Page M. J., Schwope A. D., 2011, A&A, 530, A42, *The X-ray spectral properties of the AGN population in the XMM-Newton bright serendipitous survey*

Croton D. J., et al., 2006, MNRAS, 365, 11, *The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies*

Croton D. J., et al., 2016, ApJS, 222, 22, *Semi-Analytic Galaxy Evolution (SAGE): Model Calibration and Basic Results*

Cucciati O., et al., 2012, A&A, 539, A31, *The star formation rate density and dust attenuation evolution over 12 Gyr with the VVDS surveys*

Delhaize J., et al., 2017, A&A, 602, A4, *The VLA-COSMOS 3 GHz Large Project: The infrared-radio correlation of star-forming galaxies and AGN to z $\lesssim$ 6*

Delvecchio I., et al., 2014, MNRAS, 439, 2736, *Tracing the cosmic growth of supermassive black holes to z $\sim$ 3 with Herschel*

Delvecchio I., et al., 2017, A&A, 602, A3, *The VLA-COSMOS 3 GHz Large Project: AGN and host-galaxy properties out to z $\lesssim$ 6*

Desert F.-X., Boulanger F., Puget J. L., 1990, A&A, 237, 215, *Interstellar dust models for extinction and emission*

Dickinson M., Giavalisco M., GOODS Team 2003, in Bender R., Renzini A., eds, The Mass of Galaxies at Low and High Redshift. p. 324, doi:10.1007/10899892_78

Donley J. L., et al., 2012, ApJ, 748, 142, *Identifying Luminous Active Galactic Nuclei in Deep Surveys: Revised IRAC Selection Criteria*

Donoso E., Best P. N., Kauffmann G., 2009, MNRAS, 392, 617, *Evolution of the radio-loud galaxy population*

Driver S. P., et al., 2009, Astronomy and Geophysics, 50, 12, *GAMA: towards a physical understanding of galaxy formation*

Driver S. P., et al., 2011, MNRAS, 413, 971, *Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release*

Dunlop J. S., Peacock J. A., 1990, MNRAS, 247, 19, *The Redshift Cut-Off in the Luminosity Function of Radio Galaxies and Quasars*

Dunlop J. S., et al., 2017, MNRAS, 466, 861, *A deep ALMA image of the Hubble Ultra Deep Field*

Dutton A. A., van den Bosch F. C., Dekel A., 2010, MNRAS, 405, 1690, *On the origin of the galaxy star-formation-rate sequence: evolution and scatter*

Elvis M., et al., 2009, ApJS, 184, 158, *The Chandra COSMOS Survey. I. Overview and Point Source Catalog*

Eracleous M., Boroson T. A., Halpern J. P., Liu J., 2012, ApJS, 201, 23, *A Large Systematic Search for Close Supermassive Binary and Rapidly Recoiling Black Holes*

Esin A. A., 1997, ApJ, 482, 400, *Heating and Cooling of Hot Accretion Flows by Nonlocal Radiation*

Evans D. A., Worrall D. M., Hardcastle M. J., Kraft R. P., Birkinshaw M., 2006, ApJ, 642, 96, *Chandra and XMM-Newton Observations of a Sample of Low-Redshift FR I and FR II Radio Galaxy Nuclei*

Faber S. M., et al., 2007, ApJ, 665, 265, *Galaxy Luminosity Functions to z∼1 from DEEP2 and COMBO-17: Implications for Red Galaxy Formation*

Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53, *Grand unification of AGN activity in the ΛCDM cosmology*

Feigelson E. D., Nelson P. I., 1985, ApJ, 293, 192, *Statistical methods for astronomical data with upper limits. I - Univariate distributions*

Finkelstein S. L., et al., 2015, ApJ, 810, 71, *The Evolution of the Galaxy Rest-frame Ultraviolet Luminosity Function over the First Two Billion Years*

Fixsen D. J., et al., 2009, ArXiv e-prints, *ARCADE 2 Measurement of the Extra-Galactic Sky Temperature at 3-90 GHz*

Fixsen D. J., et al., 2011, ApJ, 734, 5, *ARCADE 2 Measurement of the Absolute Sky Brightness at 3-90 GHz*

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306, *emcee: The MCMC Hammer*

Fotopoulou S., et al., 2016, A&A, 587, A142, *The 5-10 keV AGN luminosity function at 0.01 < z < 4.0*

Franzen T. M. O., et al., 2014, MNRAS, 439, 1212, *Deep 20-GHz survey of the Chandra Deep Field South and SDSS Stripe 82: source catalogue and spectral properties*

Fritz J., Franceschini A., Hatziminaoglou E., 2006, MNRAS, 366, 767, *Revisiting the infrared spectra of active galactic nuclei with a new torus emission model*

Gehrels N., 1986, ApJ, 303, 336, *Confidence limits for small numbers of events in astrophysical data*

Georgakakis A., Mobasher B., Cram L., Hopkins A., Lidman C., Rowan-Robinson M., 1999, MNRAS, 306, 708, *The Phoenix Survey: optical and near-infrared observations of faint radio sources*

Gladders M. D., Oemler A., Dressler A., Poggianti B., Vulcani B., Abramson L., 2013, ApJ, 770, 64, *The IMACS Cluster Building Survey. IV. The Log-normal Star Formation History of Galaxies*

Grogin N. A., et al., 2011, ApJS, 197, 35, *CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey*

Gruppioni C., et al., 2013, MNRAS, 432, 23, *The Herschel PEP/HerMES luminosity function - I. Probing the evolution of PACS selected Galaxies to z ≃ 4*

Guedes J., Madau P., Mayer L., Callegari S., 2011, ApJ, 729, 125, *Recoiling Massive Black Holes in Gas-rich Galaxy Mergers*

Haarsma D. B., Partridge R. B., Windhorst R. A., Richards E. A., 2000, ApJ, 544, 641, *Faint Radio Sources and Star Formation History*

Hales C. A., Murphy T., Curran J. R., Middelberg E., Gaensler B. M., Norris R. P., 2012, MNRAS, 425, 979, *BLOBCAT: software to catalogue flood-filled blobs in radio images of total intensity and linear polarization*

Hales C. A., et al., 2014, MNRAS, 441, 2555, *ATLAS 1.4 GHz Data Release 2 - I. Observations of the CDF-S and ELAIS-S1 fields and methods for constructing differential number counts*

Hardcastle M., Evans D., Croston J., 2007, MNRAS, 376, 1849, *Hot and cold gas accretion and feedback in radio-loud active galaxies*

Häring N., Rix H.-W., 2004, ApJ, 604, L89, *On the Black Hole Mass-Bulge Mass Relation*

Hart R. E., et al., 2016, MNRAS, 461, 3663, *Galaxy Zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias*

Hasinger G., et al., 2007, ApJS, 172, 29, *The XMM-Newton Wide-Field Survey in the COSMOS Field. I. Survey Description*

Heckman T. M., Best P. N., 2014, ARA&A, 52, 589, *The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe*

Heinis S., et al., 2013, MNRAS, 429, 1113, *HERMES: unveiling obscured star formation - the far-infrared luminosity function of ultraviolet-selected galaxies at $z \sim 1.5$*

Helou G., Soifer B. T., Rowan-Robinson M., 1985, ApJ, 298, L7, *Thermal infrared and nonthermal radio - Remarkable correlation in disks of galaxies*

Herrera Ruiz N., et al., 2017, preprint, *The faint radio sky: VLBA observations of the COSMOS field*

Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F., Collins R. A., 1968, Nature, 217, 709, *Observation of a Rapidly Pulsating Radio Source*

Heywood I., Jarvis M. J., Condon J. J., 2013, MNRAS, 432, 2625, *Sample variance, source clustering and their influence on the counts of faint radio sources*

Hildebrandt H., et al., 2010, A&A, 523, A31, *PHAT: PHoto-z Accuracy Testing*

Hinshaw G., et al., 2013, ApJS, 208, 19, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*

Hopkins A., Georgakakis A., Cram L., Afonso J., Mobasher B., 2000, ApJS, 128, 469, *Microjansky Radio Sources in DC 0107-46 (Abell 2877)*

Hopkins A. M., Afonso J., Chan B., Cram L. E., Georgakakis A., Mobasher B., 2003, AJ, 125, 465, *The Phoenix Deep Survey: The 1.4 GHz Microjansky Catalog*

Hopkins P. F., Somerville R. S., Hernquist L., Cox T. J., Robertson B., Li Y., 2006, ApJ, 652, 864, *The Relation between Quasar and Merging Galaxy Luminosity Functions and the Merger-driven Star Formation History of the Universe*

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356, *A Cosmological Framework for the Co-Evolution of Quasars, Supermassive Black Holes, and Elliptical Galaxies. I. Galaxy Mergers and Quasar Activity*

Hopkins P. F., Younger J. D., Hayward C. C., Narayanan D., Hernquist L., 2010, MNRAS, 402, 1693, *Mergers, active galactic nuclei and 'normal' galaxies: contributions to the distribution of star formation rates and infrared luminosity functions*

Hubble E. P., 1936, *Realm of the Nebulae*. New Haven: Yale University Press

Ilbert O., et al., 2006, A&A, 457, 841, *Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey*

Ilbert O., et al., 2010, ApJ, 709, 644, *Galaxy Stellar Mass Assembly Between $0.2 < z < 2$ from the S-COSMOS Survey*

Ilbert O., et al., 2013, A&A, 556, A55, *Mass assembly in quiescent and star-forming galaxies since $z \simeq 4$ from UltraVISTA*

Ivison R. J., et al., 2010, MNRAS, 402, 245, *BLAST: the far-infrared/radio correlation in distant galaxies*

Jansky K. G., 1933, Popular Astronomy, 41, 548, *Electrical phenomena that apparently are of interstellar origin*

Jarvis M. J., 2012, African Skies, 16, 44, *Multi-wavelength Extragalactic Surveys and the Role of MeerKAT and SALT*

Jarvis M. J., et al., 2017, preprint, *The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey*

Jennison R. C., Das Gupta M. K., 1953, Nature, 172, 996, *Fine Structure of the Extra-terrestrial Radio Source Cygnus I*

Jonker P. G., Torres M. A. P., Fabian A. C., Heida M., Miniutti G., Pooley D., 2010, MNRAS, 407, 645, *A bright off-nuclear X-ray source: a type IIn supernova, a bright ULX or a recoiling supermassive black hole in CXOJ122518.6+144545*

Karim A., et al., 2011, ApJ, 730, 61, *The Star Formation History of Mass-selected Galaxies in the COSMOS Field*

Kennicutt Jr. R. C., 1998, ARA&A, 36, 189, *Star Formation in Galaxies Along the Hubble Sequence*

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531, *Star Formation in the Milky Way and Nearby Galaxies*

Kereš D., Katz N., Davé R., Fardal M., Weinberg D. H., 2009, MNRAS, 396, 2332, *Galaxies in a simulated ΛCDM universe - II. Observable properties and constraints on feedback*

Kimball A. E., Ivezić Ž., 2008, AJ, 136, 684, *A Unified Catalog of Radio Objects Detected by NVSS, First, WENSS, GB6, and SDSS*

Klein U., Wielebinski R., Morsi H. W., 1988, A&A, 190, 41, *Radio continuum observations of M82*

Koekemoer A. M., et al., 2007, ApJS, 172, 196, *The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing*

Koekemoer A. M., et al., 2011, ApJS, 197, 36, *CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey-The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics*

Komossa S., 2012, Advances in Astronomy, 2012, *Recoiling Black Holes: Electromagnetic Signatures, Candidates, and Astrophysical Implications*

Komossa S., Zhou H., Lu H., 2008, ApJ, 678, L81, *A Recoiling Supermassive Black Hole in the Quasar SDSS J092712.65+294344.0?*

Koss M., et al., 2014, preprint, *SDSS1133: An Unusually Persistent Transient in a Nearby Dwarf Galaxy*

Krishna G., Sirothia S. K., Mhaskey M., Ranadive P., Wiita P. J., Goyal A., Kantharia N. G., Ishwara-Chandra C. H., 2014, MNRAS, 443, 2824, *Extragalactic radio sources with sharply inverted spectrum at metre wavelengths*

Laigle C., et al., 2016, ApJS, 224, 24, *The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies*

Larson D., et al., 2011, ApJS, 192, 16, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters*

Le Fèvre O., et al., 2015, A&A, 576, A79, *The VIMOS Ultra-Deep Survey: ∼10 000 galaxies with spectroscopic redshifts to study galaxy assembly at early epochs $2 < z \simeq 6$*

Lilly S. J., et al., 2007, ApJS, 172, 70, *zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field*

Lilly S. J., et al., 2009, ApJS, 184, 218, *The zCOSMOS 10k-Bright Spectroscopic Sample*

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119, *Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos*

Lintott C. J., et al., 2008, MNRAS, 389, 1179, *Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey*

Lintott C., et al., 2011, MNRAS, 410, 166, *Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies*

Lisenfeld U., Völk H. J., 2000, A&A, 354, 423, *On the radio spectral index of galaxies*

Loeb A., 2007, Physical Review Letters, 99, 041103, *Observable Signatures of a Black Hole Ejected by Gravitational-Radiation Recoil in a Galaxy Merger*

Longair M. S., 1966, MNRAS, 133, 421, *On the interpretation of radio source counts*

Madau P., Dickinson M., 2014, ARA&A, 52, 415, *Cosmic Star-Formation History*

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388, *High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to z∼4*

Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35, *Evolution of the dusty infrared luminosity function from $z = 0$ to $z = 2.3$ using observations from Spitzer*

Magnelli B., et al., 2013, A&A, 553, A132, *The deepest Herschel-PACS far-infrared survey: number counts and infrared luminosity functions from combined PEP/GOODS-H observations*

Magnelli B., et al., 2015, A&A, 573, A45, *The far-infrared/radio correlation and radio spectral index of galaxies in the SFR-M$_*$ plane up to z~2*

Mancuso C., et al., 2015, ApJ, 810, 72, *Predictions for Ultra-deep Radio Counts of Star-forming Galaxies*

Mancuso C., Lapi A., Shi J., Gonzalez-Nuevo J., Aversa R., Danese L., 2016, ApJ, 823, 128, *The Quest for Dusty Star-forming Galaxies at High Redshift z $\gtrsim$ 4*

Markevitch M., Sarazin C. L., Vikhlinin A., 1999, ApJ, 521, 526, *Physics of the Merging Clusters Cygnus A, A3667, and A2065*

Marshall H. L., 1985, ApJ, 299, 109, *The evolution of optically selected quasars with Z less than 2.2 and B less than 20*

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901, *The dark matter of gravitational lensing*

Matthews T. A., Sandage A. R., 1963, ApJ, 138, 30, *Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects.*

Mauch T., Sadler E. M., 2007, MNRAS, 375, 931, *Radio sources in the 6dFGS: local luminosity functions at 1.4GHz for star-forming galaxies and radio-loud AGN*

McAlpine K., Jarvis M. J., Bonfield D. G., 2013, MNRAS, 436, 1084, *Evolution of faint radio sources in the VIDEO-XMM3 field*

McCracken H. J., et al., 2012, A&A, 544, A156, *UltraVISTA: a new ultra-deep near-infrared survey in COSMOS*

McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI. p. 127

Merritt D., Storchi-Bergmann T., Robinson A., Batcheldor D., Axon D., Cid Fernandes R., 2006, MNRAS, 367, 1746, *The nature of the HE0450-2958 system*

Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64, *Dust Absorption and the Ultraviolet Luminosity Density at z $\sim$ 3 as Calibrated by Local Starburst Galaxies*

Miettinen O., et al., 2015, A&A, 577, A29, *(Sub)millimetre interferometric imaging of a sample of COSMOS/AzTEC submillimetre galaxies. I. Multiwavelength identifications and redshift distribution*

Miley G., 1980, ARA&A, 18, 165, *The structure of extended extragalactic radio sources*

Miller L., Peacock J. A., Mead A. R. G., 1990, MNRAS, 244, 207, *The bimodal radio luminosity function of quasars*

Miller N. A., Fomalont E. B., Kellermann K. I., Mainieri V., Norman C., Padovani P., Rosati P., Tozzi P., 2008, ApJS, 179, 114, *The VLA 1.4 GHz Survey of the Extended Chandra Deep Field-South: First Data Release*

Miller N. A., et al., 2013, ApJS, 205, 13, *The Very Large Array 1.4 GHz Survey of the Extended Chandra Deep Field South: Second Data Release*

Mitchell-Wynne K., Santos M. G., Afonso J., Jarvis M. J., 2014, MNRAS, 437, 2270, *Beyond stacking: a maximum-likelihood method to constrain radio source counts below the detection threshold*

Mo H., van den Bosch F. C., White S., 2010, *Galaxy Formation and Evolution*. Cambridge University Press

Mooley K. P., et al., 2016, ApJ, 818, 105, *The Caltech-NRAO Stripe 82 Survey (CNSS). I. The Pilot Radio Transient Survey In 50 deg^2*

Morganti R., et al., 2010, preprint, *Continuum surveys with LOFAR and synergy with future large surveys in the 1-2 GHz band*

Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, ApJ, 731, 113, *A Cosmic Variance Cookbook*

Mullaney J. R., Alexander D. M., Goulding A. D., Hickox R. C., 2011, MNRAS, 414, 1082, *Defining the intrinsic AGN infrared spectral energy distribution and measuring its contribution to the infrared output of composite galaxies*

Mullaney J. R., et al., 2012, ApJ, 753, L30, *The Hidden "AGN Main Sequence": Evidence for a Universal Black Hole Accretion to Star Formation Rate Ratio since z $\sim$ 2 Producing an M$_{BH}$-M$_*$ Relation*

Murphy E. J., 2009, ApJ, 706, 482, *The Far-Infrared-Radio Correlation at High Redshifts: Physical Considerations and Prospects for the Square Kilometer Array*

Murphy E. J., et al., 2011, ApJ, 737, 67, *Calibrating Extinction-free Star Formation Rate Diagnostics with 33 GHz Free-free Emission in NGC 6946*

Narayan R., Mahadevan R., Grindlay J. E., Popham R. G., Gammie C., 1998, ApJ, 492, 554, *Advection-dominated accretion model of Sagittarius A*: evidence for a black hole at the Galactic center.*

Noeske K. G., et al., 2007, ApJ, 660, L43, *Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies*

Norris R. P., 2017, Nature Astronomy, 1, 233, *Extragalactic radio continuum surveys and the transformation of radio astronomy*

Norris R. P., et al., 2005, AJ, 130, 1358, *Radio Observations of the Hubble Deep Field-South Region. I. Survey Description and Initial Results*

Norris R. P., et al., 2011, PASA, 28, 215, *EMU: Evolutionary Map of the Universe*

Norris R. P., et al., 2013, PASA, 30, 20, *Radio Continuum Surveys with Square Kilometre Array Pathfinders*

Norris R., Basu K., Brown M., Carretti E., Kapinska A. D., Prandoni I., Rudnick L., Seymour N., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 86, *The SKA Mid-frequency All-sky Continuum Survey: Discovering the unexpected and transforming radio-astronomy*

Novak M., et al., 2015, MNRAS, 447, 1282, *New insights from deep VLA data on the potentially recoiling black hole CID-42 in the COSMOS field*

Novak M., et al., 2017, A&A, 602, A5, *The VLA-COSMOS 3 GHz Large Project: Cosmic star formation history since z∼5*

Owen F. N., Morrison G. E., 2008, AJ, 136, 1889, *The Deep Swire Field. I. 20 cm Continuum Radio Observations: A Crowded Sky*

Owen F. N., Morrison G. E., Klimek M. D., Greisen E. W., 2009, AJ, 137, 4846, *The Deep SWIRE Field. II. 90 cm Continuum Observations and 20 cm-90 cm Spectra*

Padovani P., 2011, MNRAS, 411, 1547, *The microjansky and nanojansky radio sky: source population and multi-wavelength properties*

Padovani P., 2016, A&A Rev., 24, 13, *The faint radio sky: radio astronomy becomes mainstream*

Padovani P., Mainieri V., Tozzi P., Kellermann K. I., Fomalont E. B., Miller N., Rosati P., Shaver P., 2009, ApJ, 694, 235, *The Very Large Array Survey of the Chandra Deep Field South. IV. Source Population*

Padovani P., Bonzini M., Kellermann K. I., Miller N., Mainieri V., Tozzi P., 2015, MNRAS, 452, 1263, *Radio-faint AGN: a tale of two populations*

Parsa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, MNRAS, 456, 3194, *The galaxy UV luminosity function at z ≃ 2-4; new results on faint-end slope and the evolution of luminosity density*

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*

Peres A., 1962, Phys. Rev., 128, 2471, *Classical radiation recoil*

Perley R. A., Dreher J. W., Cowan J. J., 1984, ApJ, 285, L35, *The jet and filaments in Cygnus A*

Piron F., 2016, Comptes Rendus Physique, 17, 617, *Gamma-ray bursts at high and very high energies*

Planck Collaboration et al., 2014, A&A, 571, A30, *Planck 2013 results. XXX. Cosmic infrared background measurements and implications for star formation*

Pozzetti L., et al., 2007, A&A, 474, 443, *The VIMOS VLT Deep Survey. The assembly history of the stellar mass in galaxies: from the young to the old universe*

Pracy M. B., et al., 2016, MNRAS, 460, 2, *GAMA/WiggleZ: the 1.4 GHz radio luminosity functions of high- and low-excitation radio galaxies and their redshift evolution to z = 0.75*

Prandoni I., Seymour N., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 67, *Revealing the Physics and Evolution of Galaxies and Galaxy Clusters with SKA Continuum Surveys*

Prandoni I., Gregorini L., Parma P., de Ruiter H. R., Vettolani G., Wieringa M. H., Ekers R. D., 2001, A&A, 365, 392, *The ATESP radio survey. III. Source counts*

Prescott M. K. M., Impey C. D., Cool R. J., Scoville N. Z., 2006, ApJ, 644, 100, *Quasars in the COSMOS Field*

Rau U., Cornwell T. J., 2011, A&A, 532, A71, *A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry*

Rau U., Bhatnagar S., Owen F. N., 2014, preprint, *Wideband Mosaic Imaging with the VLA - quantifying faint source imaging accuracy*

Reber G., 1944, ApJ, 100, 279, *Cosmic Static.*

Reddy N. A., Pettini M., Steidel C. C., Shapley A. E., Erb D. K., Law D. R., 2012, ApJ, 754, 25, *The Characteristic Star Formation Histories of Galaxies at Redshifts z ∼ 2-7*

Rees M. J., 1967, MNRAS, 136, 279, *Studies in radio structure-II. The relaxation of relativistic electron spectra in self-absorbed radio sources*

Rees M. J., Begelman M. C., Blandford R. D., Phinney E. S., 1982, Nature, 295, 17, *Ion-supported tori and the origin of radio jets*

Rigby E. E., Best P. N., Brookes M. H., Peacock J. A., Dunlop J. S., Röttgering H. J. A., Wall J. V., Ker L., 2011, MNRAS, 416, 1900, *The luminosity-dependent high-redshift turnover in the steep spectrum radio luminosity function: clear evidence for downsizing in the radio-AGN population*

Roberts M. S., Haynes M. P., 1994, ARA&A, 32, 115, *Physical Parameters along the Hubble Sequence*

Robinson A., Young S., Axon D. J., Kharb P., Smith J. E., 2010, ApJ, 717, L122, *Spectropolarimetric Evidence for a Kicked Supermassive Black Hole in the Quasar E1821+643*

Rodighiero G., et al., 2010, A&A, 518, L25, *The first Herschel view of the mass-SFR link in high-z galaxies*

Rodighiero G., et al., 2011, ApJ, 739, L40, *The Lesser Role of Starbursts in Star Formation at z = 2*

Rowan-Robinson M., Benn C. R., Lawrence A., McMahon R. G., Broadhurst T. J., 1993, MNRAS, 263, 123, *The evolution of faint radio sources*

Rowan-Robinson M., et al., 2016, MNRAS, 461, 1100, *The star formation rate density from z= 1 to 6*

Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*

Ryle M., Scheuer P. A. G., 1955, Proceedings of the Royal Society of London Series A, 230, 448, *The Spatial Distribution and the Nature of Radio Stars*

Sadler E. M., et al., 2002, MNRAS, 329, 227, *Radio sources in the 2dF Galaxy Redshift Survey - II. Local radio luminosity functions for AGN and star-forming galaxies at 1.4 GHz*

Sadler E. M., et al., 2007, MNRAS, 381, 211, *Radio galaxies in the 2SLAQ Luminous Red Galaxy Survey - I. The evolution of low-power radio galaxies to z ∼ 0.7*

Salpeter E. E., 1955, ApJ, 121, 161, *The Luminosity Function and Stellar Evolution.*

Sandage A., 1961, *The Hubble Atlas of Galaxies.* Washington: Carnegie Institution

Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749, *Luminous Infrared Galaxies*

Sanders D. B., et al., 2007, ApJS, 172, 86, *S-COSMOS: The Spitzer Legacy Survey of the Hubble Space Telescope ACS 2 deg² COSMOS Field I: Survey Strategy and First Analysis*

Sargent M. T., et al., 2010, ApJS, 186, 341, *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*

Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, ApJ, 747, L31, *The Contribution of Starbursts and Normal Galaxies to Infrared Luminosity Functions at z < 2*

Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, MNRAS, 242, 318, *The 60-micron and far-infrared luminosity functions of IRAS galaxies*

Schawinski K., et al., 2010, ApJ, 711, 284, *Galaxy Zoo: The Fundamentally Different Co-Evolution of Supermassive Black Holes and Their Early- and Late-Type Host Galaxies*

Schawinski K., et al., 2014, MNRAS, 440, 889, *The green valley is a red herring: Galaxy Zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies*

Scheuer P. A. G., 1957, Proceedings of the Cambridge Philosophical Society, 53, 764, *A statistical method for analysing observations of faint radio stars*

Schinnerer E., et al., 2004, AJ, 128, 1974, *The VLA-COSMOS Survey. I. Radio Identifications from the Pilot Project*

Schinnerer E., et al., 2007, ApJS, 172, 46, *The VLA-COSMOS Survey. II. Source Catalog of the Large Project*

Schinnerer E., et al., 2010, ApJS, 188, 384, *The VLA-COSMOS Survey. IV. Deep Data and Joint Catalog*

Schmidt M., 1968, ApJ, 151, 393, *Space Distribution and Luminosity Functions of Quasi-Stellar Radio Sources*

Schmidt M., Schneider D. P., Gunn J. E., 1995, AJ, 110, 68, *Spectrscopic CCD Surveys for Quasars at Large Redshift.IV.Evolution of the Luminosity Function from Quasars Detected by Their Lyman-Alpha Emission*

Schmitt J. H. M. M., 1985, ApJ, 293, 178, *Statistical analysis of astronomical data containing upper bounds - General methods and examples drawn from X-ray astronomy*

Scott K. S., et al., 2008, MNRAS, 385, 2225, *AzTEC millimetre survey of the COSMOS field - I. Data reduction and source catalogue*

Scoville N., et al., 2007, ApJS, 172, 1, *The Cosmic Evolution Survey (COSMOS): Overview*

Seymour N., McHardy I. M., Gunn K. F., 2004, MNRAS, 352, 131, *Radio observations of the 13^hXMM-Newton/ROSAT Deep X-ray Survey Area*

Seymour N., et al., 2008, MNRAS, 386, 1695, *The star formation history of the Universe as revealed by deep radio observations*

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337, *Black holes in binary systems. Observational appearance.*

Shields G. A., et al., 2009, ApJ, 707, 936, *The Quasar SDSS J105041.35+345631.3: Black Hole Recoil or Extreme Double-Peaked Emitter?*

Shimwell T. W., et al., 2017, A&A, 598, A104, *The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release*

Sijacki D., Springel V., Haehnelt M. G., 2011, MNRAS, 414, 3656, *Gravitational recoils of supermassive black holes in hydrodynamical simulations of gas-rich galaxies*

Silverman J. D., et al., 2008, ApJ, 679, 118, *The Luminosity Function of X-Ray-selected Active Galactic Nuclei: Evolution of Supermassive Black Holes at High Redshift*

Smolčić V., 2009, ApJ, 699, L43, *The Radio AGN Population Dichotomy: Green Valley Seyferts Versus Red Sequence Low-Excitation Active Galactic Nuclei*

Smolčić V., Riechers D. A., 2011, ApJ, 730, 64, *The Molecular Gas Content of $z < 0.1$ Radio Galaxies: Linking the Active Galactic Nucleus Accretion Mode to Host Galaxy Properties*

Smolčić V., et al., 2008, ApJS, 177, 14, *A New Method to Separate Star-forming from AGN Galaxies at Intermediate Redshift: The Submillijansky Radio Population in the VLA-COSMOS Survey*

Smolčić V., et al., 2009a, ApJ, 690, 610, *The Dust-Unbiased Cosmic Star-Formation History from the 20 CM VLA-COSMOS Survey*

Smolčić V., et al., 2009b, ApJ, 696, 24, *Cosmic Evolution of Radio Selected Active Galactic Nuclei in the Cosmos Field*

Smolčić V., et al., 2012, A&A, 548, A4, *Millimeter imaging of submillimeter galaxies in the COSMOS field: redshift distribution*

Smolčić V., et al., 2014, MNRAS, 443, 2590, *The VLA-COSMOS Survey - V. 324 MHz continuum observations*

Smolčić V., et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 69, *Exploring AGN Activity over Cosmic Time with the SKA*

Smolčić V., et al., 2017a, A&A, 602, A1, *The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release*

Smolčić V., et al., 2017b, A&A, 602, A2, *The VLA-COSMOS 3 GHz Large Project: Multiwavelength counterparts and the composition of the faint radio population*

Smolčić V., et al., 2017c, A&A, 602, A6, *The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z∼5*

Sparke L. S., Gallagher III J. S., 2007, *Galaxies in the Universe: An Introduction*, 2nd edn. Cambridge University Press

Spergel D. N., et al., 2007, ApJS, 170, 377, *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology*

Springel V., et al., 2005, Nature, 435, 629, *Simulations of the formation, evolution and clustering of galaxies and quasars*

Steinhardt C. L., et al., 2012, ApJ, 759, 24, *SDSS 0956+5128: A Broad-line Quasar with Extreme Velocity Offsets*

Strateva I., et al., 2001, AJ, 122, 1861, *Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data*

Swinbank A. M., et al., 2014, MNRAS, 438, 1267, *An ALMA survey of sub-millimetre Galaxies in the Extended Chandra Deep Field South: the far-infrared properties of SMGs*

Szokoly G. P., et al., 2004, ApJS, 155, 271, *The Chandra Deep Field-South: Optical Spectroscopy. I.*

Tabatabaei F. S., et al., 2017, ApJ, 836, 185, *The Radio Spectral Energy Distribution and Star-formation Rate Calibration in Galaxies*

Tasse C., Röttgering H. J. A., Best P. N., Cohen A. S., Pierre M., Wilman R., 2007, A&A, 471, 1105, *GMRT observations of the XMM large scale structure survey field*

Thuan T. X., Balkowski C., Tran Thanh Van J., eds, 1992, *Physics of nearby galaxies: Nature or nurture?*

Tinsley B. M., 1972, A&A, 20, 383, *Galactic Evolution*

Trump J. R., et al., 2007, ApJS, 172, 383, *Magellan Spectroscopy of AGN Candidates in the COSMOS Field*

Velusamy T., Roshi D., Venugopal V. R., 1992, MNRAS, 255, 210, *Multifrequency VLA observations of the Crab Nebula at different epochs - Expansion and radio spectra of filaments*

Vernstrom T., Scott D., Wall J. V., 2011, MNRAS, 415, 3641, *Contribution to the diffuse radio background from extragalactic radio sources*

Vernstrom T., et al., 2014, MNRAS, 440, 2791, *Deep 3 GHz number counts from a P(D) fluctuation analysis*

Vernstrom T., Scott D., Wall J. V., Condon J. J., Cotton W. D., Perley R. A., 2016a, MNRAS, 461, 2879, *Deep 3-GHz observations of the Lockman Hole North with the Very Large Array - I. Source extraction and uncertainty analysis*

Vernstrom T., Scott D., Wall J. V., Condon J. J., Cotton W. D., Kellermann K. I., Perley R. A., 2016b, MNRAS, 462, 2934, *Deep 3-GHz observations of the Lockman Hole North with the Very Large Array - II. Catalogue and µJy source properties*

Vito F., et al., 2014, MNRAS, 441, 1059, *Black hole accretion preferentially occurs in gas-rich galaxies**

Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559, *The Assembly and Merging History of Supermassive Black Holes in Hierarchical Models of Galaxy Formation*

Waddington I., Dunlop J. S., Peacock J. A., Windhorst R. A., 2001, MNRAS, 328, 882, *The LBDS Hercules sample of mJy radio sources at 1.4 GHz - II. Redshift distribution, radio luminosity function, and the high-redshift cut-off*

Walcher J., Groves B., Budavári T., Dale D., 2011, Ap&SS, 331, 1, *Fitting the integrated spectral energy distributions of galaxies*

Wang B., Heckman T. M., 1996, ApJ, 457, 645, *Internal Absorption and the Luminosity of Disk Galaxies*

Wang X., Loeb A., 2014, MNRAS, 441, 809, *Detecting floating black holes as they traverse the gas disc of the Milky Way*

Wells D. C., 1985, in di Gesu V., Scarsi L., Crane P., Friedman J. H., Levialdi S., eds, Data Analysis in Astronomy. p. 195

Willett K. W., et al., 2017, MNRAS, 464, 4176, *Galaxy Zoo: morphological classifications for 120 000 galaxies in HST legacy imaging*

Willott C. J., Rawlings S., Blundell K. M., Lacy M., Eales S. A., 2001, MNRAS, 322, 536, *The radio luminosity function from the low-frequency 3CRR, 6CE and 7CRS complete samples*

Wilman R. J., et al., 2008, MNRAS, 388, 1335, *A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes*

Wright E. L., 2006, PASP, 118, 1711, *A Cosmology Calculator for the World Wide Web*

Wrobel J. M., Comerford J. M., Middelberg E., 2014, ApJ, 782, 116, *Constraints on Two Active Galactic Nuclei in the Merger Remnant COSMOS J100043.15+020637.2*

Wuyts S., et al., 2011, ApJ, 738, 106, *On Star Formation Rates and Star Formation Histories of Galaxies Out to z ∼ 3*

Yuan Z., Wang J., Zhou M., Mao J., 2016, ApJ, 820, 65, *A Mixture Evolution Scenario of the AGN Radio Luminosity Function*

Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803, *Radio Properties of Infrared-selected Galaxies in the IRAS 2 Jy Sample*

Zwart J., et al., 2015a, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 172, *Astronomy Below the Survey Threshold in the SKA Era*

Zwart J. T. L., Santos M., Jarvis M. J., 2015b, MNRAS, 453, 1740, *Far beyond stacking: fully Bayesian constraints on sub-µJy radio source populations over the XMM-LSS-VIDEO field*

da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595, *A simple model to interpret the ultraviolet, optical and infrared emission from galaxies*

de Jong T., Klein U., Wielebinski R., Wunderlich E., 1985, A&A, 147, L6, *Radio continuum and far-infrared emission from spiral galaxies - A close correlation*

de Vaucouleurs G., de Vaucouleurs A., Corwin Jr. H. G., Buta R. J., Paturel G., Fouqué P., 1991, *Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h. Volume III: Data for galaxies between 12^h and 24^h*. Springer, New York, NY (USA)

de Zotti G., Massardi M., Negrello M., Wall J., 2010, A&A Rev., 18, 1, *Radio and millimeter continuum surveys and their astrophysical implications*

van der Kruit P. C., 1971, A&A, 15, 110, *Observations of core sources in Seyfert and normal galaxies with the Westerbork synthesis radio telescope at 1415 MHz.*

van der Laan H., Perola G. C., 1969, A&A, 3, 468, *Aspects of Radio Galaxy Evolution*

Appendix A

List of acronyms

ΛCDM: Lambda cold dark matter

2dF: Two-degree Field

6dFGS: Six-degree Field Galaxy Survey

ACS: Advanced Camera for Surveys

AGN: active galactic nuclei

AIPS: Astronomical Image Processing System

ALESS: ALMA LESS

ALMA: Atacama Large Millimeter/submillimeter Array

APEX: Atacama Pathfinder Experiment

ARCADE: Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission

ASKAP: Australian Square Kilometer Array Pathfinder

ASTRON: Netherlands Institute for Radio Astronomy

ATCA: Australia Telescope Compact Array

ATNF: Australia Telescope National Facility

AUI: Associated Universities, Inc.

BWS: bandwidth smearing

CASA: Common Astronomy Software Applications

CDF: cumulative distribution function

CMB: cosmic microwave background

COSMOS: Cosmic Evolution Survey

CSIRO: Commonwealth Scientific and Industrial Research Organisation

Dec: declination

DEIMOS: Deep Imaging Multi-Object Spectrograph

DR: data release

ECDFS: The Extended Chandra Deep Field South

EMU: Evolutionary Map of the Universe

ERC: European Research Council

ESO: European Southern Observatory

FIR: far-infrared

FIRST: Faint Images of the Radio Sky at Twenty-Centimeters

FWHM: full width at half maximum

GMRT: Giant Metrewave Radio Telescope

GOODS-N/S: Great Observatories Origins Deep Survey-North/South

GW: gravitational wave

HerMES: Herschel Multi tiered Extragalactic Survey

HPBW: half power beam width

HST: Hubble Space Telescope

HUDF: Hubble Ultra Deep Field

HyLIRG: hyperluminous infrared galaxy

IDL: Interactive Data Language

IMF: initial mass function

IR: infrared

IRAC: Infrared Array Camera

IRSA: Infrared Science Archive

IRX: infrared excess

LABOCA: Large APEX Bolometer Camera

LBG: Lyman break galaxy

LESS: LABOCA ECDFS Submillimeter Survey

LF: luminosity function

LOFAR: Low-Frequency Array

LRG: Luminous Red Galaxy

MCMC: Markov chain Monte Carlo

MeerKAT: Meer Karoo Array Telescope

MIGHTEE: MeerKAT International GHz Tiered Extragalactic Exploration

MIR: mid-infrared

MPL: multinode power law

MSMF: multiscale multifrequency

NAOJ: National Astronomical Observatory of Japan

ngVLA: Next Generation Very Large Array

NIR: near-infrared

NRAO: National Radio Astronomy Observatory

NUV: near-ultraviolet

NVSS: NRAO VLA Sky Survey

NW: North-West

PACS: Photoconductor Array Camera and Spectrometer

PB: primary beam

PDE: pure density evolution

PDF: probability density function

PEP: PACS Evolutionary Probe

PI: principal investigator

PL: power law

PLE: pure luminosity evolution

PSF: point spread function

QSO: quasi-stellar object or quasar

RA: right ascension

RFI: radio-frequency interference

rms: root-mean-square

S/N: signal-to-noise ratio

SAD: search and destroy

SDSS: Sloan Digital Sky Survey

SE: South-East

SED: spectral energy distribution

SF: star-forming

SFR: star formation rate

SFRD: star formation rate density

SKA: Square Kilometer Array

SKADS: Square Kilometer Array Design Studies

SMBH: supermassive black hole

SMG: submillimeter galaxy

SPLASH: Spitzer Large Area Survey with Hyper-Suprime-Cam

SPW: spectral window

TIR: total infrared

TT: Taylor Term

ULIRG: ultra luminous infrared galaxies

UV: ultraviolet

VIDEO: VISTA Deep Extragalactic Observations

VIMOS: Visible MultiObject Spectrograph

VISTA: Visible and Infrared Survey Telescope for Astronomy

VLA: Karl G. Jansky Very Large Array

VLASS: VLA Sky Survey

VLBA: Very Long Baseline Array

VLT: Very Large Telescope

VVDS: VIMOS VLT Deep Survey

WMAP: Wilkinson microwave anisotropy probe

XMM: X-ray Multi-Mirror Mission

Curriculum vitae

Mladen Novak was born in 1989 in Rijeka (Croatia), where he completed primary and high school. He then moved to Zagreb, where he obtained the Master's degree in Physics at the Department of Physics, Faculty of Science, University of Zagreb, on the subject of *Creation of hydrogen molecules on the surfaces of cosmic dust* under the supervision of assoc. prof. dr. sc. Goranka Bilalbegović in 2013. He remained at the same institution to start working in the astrophysical group led by assoc. prof. dr. sc. Vernesa Smolčić, funded by the ERC (European Research Council) Starting Grant *Constraining Stellar Mass and Supermassive Black Hole Growth through Cosmic Times: Paving the way for the next generation sky surveys* (project ID: 337595). At the same time he enrolled in doctoral studies in the field of atomic and molecular physics, module astrophysics. During his PhD studies he presented his work by giving more than ten talks at international astrophysical conferences.

Refereed publications

Submitted

22 Hale, C. L., Jarvis, M. J., Delvecchio, I., Hatfield, P. W., **Novak, M.**, et al., submitted to MNRAS, *The clustering and bias of radio-selected AGN and star-forming galaxies in the COSMOS field*

21 **Novak, M.**, Smolčić, V., Schinnerer, E., et al., submitted to A&A, *Constraints on submicrojansky radio number counts based on evolving VLA-COSMOS luminosity functions*

20 Molnar, D. Cs., Sargent, M. T., Delhaize, J., Delvecchio, I., Smolčić, **Novak, M.**, et al., submitted to MNRAS, *The infrared-radio correlation of spheroid- and disk-dominated star-forming galaxies to z ∼ 1.5 in the COSMOS field*

19 Lemaux, B. C., inter alios, **Novak, M.**, et al., submitted to A&A, *The VIMOS Ultra Deep Survey: Emerging from the Dark, a Massive Proto-Cluster at z∼4.57*

18 Daddi, E., inter alios, **Novak, M.**, et al., submitted to ApJ Letters, *Radio selection of the most distant galaxy clusters*

17 Butler, A., Huynh, M., Delhaize, J., Smolčić, V., Milaković, D., **Novak, M.**, et al., accepted for publication in A&A, arXiv:1703.10296, *The XXL Survey XVIII. ATCA 2.1 GHz radio source catalogue and source counts for the XXL-South field*

Published

16 Miettinen, O., inter alios, **Novak, M.**, et al., 2017, A&A, 606, A17, *An ALMA survey of submillimetre galaxies in the COSMOS field: Physical properties derived from energy balance spectral energy distribution modelling*

15 Smolčić, V., **Novak, M.**, Delvecchio, I., et al., 2017, A&A, 602, A6, *The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z∼5*

14 **Novak, M.**, Smolčić, V., Delhaize, J., Delvecchio, I., Zamorani, G., et al., 2017, A&A, 602, A5, *The VLA-COSMOS 3 GHz Large Project: The cosmic star formation history since z∼5*

13 Delhaize, J., Smolčić, V., Delvecchio, I., **Novak, M.**, et al., 2017, A&A, 602, A4, *VLA-COSMOS 3 GHz: The infrared-radio correlation of star-forming galaxies and AGN to z ≲ 6*

12 Delvecchio, I., inter alios, **Novak, M.**, et al., 2017, A&A, 602, A3, *The VLA-COSMOS 3 GHz Large Project: AGN and host-galaxy properties out to z≲6*

11 Smolčić, V., Delvecchio, I., Zamorani, G., Baran, N., **Novak, M.**, et al., 2017, A&A, 602, A2, *The VLA-COSMOS 3 GHz Large Project: Multiwavelength counterparts and the composition of radio sources*

10 Smolčić, V., **Novak, M.**, Bondi, M., et al., 2017, A&A, 602, A1, *The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release*

9 Miettinen, O., **Novak, M.**, Smolčić, V., et al., 2017, A&A, 602, A54, *An ALMA survey of submillimetre galaxies in the COSMOS field: The extent of the radio-emitting region revealed by 3 GHz imaging with the Very Large Array*

8 Miettinen, O., Delvecchio, I., Smolčić, V., **Novak, M.**, et al., 2017, A&A, 597, A5, *(Sub)millimetre interferometric imaging of a sample of COSMOS/AzTEC submillimetre galaxies. IV. Physical properties derived from spectral energy distributions*

7 Smolčić, V., Delhaize, J., Huynh, M., Bondi, M., Ciliegi, P., **Novak, M.**, et al., 2016, A&A, 592, A10, *The XXL Survey. XI. ATCA 2.1 GHz continuum observations*

6 Baran, N., Smolčić, V., Milaković, D., **Novak, M.**, et al., 2016, A&A, 592, A8, *The XXL Survey. IX. Optical overdensity and radio continuum analysis of a supercluster at z = 0.43*

5 Pierre, M., inter alios, **Novak, M.**, et al., 2016, A&A, 592, A1, *The XXL Survey. I. Scientific motivations - XMM-Newton observing plan - Follow-up observations and simulation programme*

4 Miettinen, O., **Novak, M.**, Smolčić, V., et al., 2015, A&A, 584, A32, *(Sub)millimetre interferometric imaging of a sample of COSMOS/AzTEC submillimetre galaxies. II. The spatial extent of the radio-emitting regions*

3 Miettinen, O., Smolčić, V., **Novak, M.**, et al., 2015, A&A, 577, A29, *(Sub)millimetre inter-ferometric imaging of a sample of COSMOS/AzTEC submillimetre galaxies. I. Multiwavelength identifications and redshift distribution*

2 Smolčić, V., Karim, A., Miettinen, O., **Novak, M.**, et al., 2015, A&A, 576, A127, *Physical properties of z>4 submillimeter galaxies in the COSMOS field*

1 **Novak, M.**, Smolčić, V., Civano, F., et al., 2015, MNRAS, 447, 1282, *New insights from deep VLA data on the potentially recoiling black hole CID-42 in the COSMOS field*

Non-refereed publications

2 Pierre, M., inter alios, **Novak, M.** et al., proceeding of the XMM Next Decade Workshop, arXiv:1610.01781, *The XXL survey: first results and future*

1 Baran, N., Smolčić, V., Delvecchio, I., **Novak, M.**, et al., proceeding of the conference "The many facets of extragalactic radio surveys", arXiv:1603.05996, *Star-forming galaxies versus low- and high-excitation radio AGN in the VLA-COSMOS 3GHz Large Project*