TECHNISCHE UNIVERSITÄT MÜNCHEN
Max-Planck-Institut für Plasma Physik
Max-Planck Institut für Extraterrestrische Physik

Active Galactic Nuclei Luminosity Function and the Lockman Hole Deep Field

Sotiria Fotopoulou

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Harald Friedrich
Prüfer der Dissertation:	1.	Hon.-Prof. Dr. Günther Hasinger
	2.	Univ.-Prof. Dr. Stefan Schönert

Die Dissertation wurde am 29.08.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 20.09.2012 angenommen.

ABSTRACT

Active Galactic Nuclei (AGN) are powered by gas accretion onto supermassive black holes (SMBH) located at the centers of galaxies. They are suspected to be an important phase in galaxy evolution, as suggested by the $M_{BH} - \sigma$ relation. In order to test this hypothesis we need to understand how AGN evolve in time. This thesis is focused on the X-ray Luminosity Function (XLF) of AGN in the 5-10 keV and 2-10 keV energy bands. Since a large sample is needed to fully study the AGN evolution, we first produced reliable redshifts for the Lockman Hole (LH) field. Consequently, assembling a sizable sample of $\sim$ 500 AGN from many surveys we were able to constrain the AGN evolution with redshift in the 5-10 keV band. The result is in good agreement with a unified 2-10 keV XLF, produced for the first time as a meta-analysis of previous works in the literature.

The LH is the field with the least galactic hydrogen column density along the line of sight. This provides an unobscured view of the soft X-ray energy range, while at the same time there is minimal galactic cirrus emission in the infra-red. Combining very deep ultra-violet, optical, and infra-red data, we created the first homogeneous multiwavelength photometric catalog in the LH for $\sim$ 187600 sources. Using this catalog, we computed photometric redshifts with special treatment for the AGN identified through their X-ray emission. The achieved accuracy of the photometric redshifts is a factor of two improvement over those previously available for this field.

As the next step we combined several X-ray fields, including LH, creating a sizable sample which allows for the first time the estimation of the XLF and its evolution with redshift at 5-10 keV. This selection ensures that the measured flux is affected very little from photoelectric absorption by the material surrounding the SMBH, a common issue at lower X-ray energy ranges. The derivation of the XLF is achieved using traditional methods such as the 1/Vmax, and Maximum Likelihood Estimation and compared to the more elaborate Bayesian analysis. We tested several evolutionary models proposed in the literature and identified the Luminosity Dependent Density Evolution as the best model to the current dataset using both the Akaike Information Criterion and Bayesian model selection.

Several scenarios have been used in the literature with inconsistent physical motivations to describe the AGN evolution through the modeling of the 2-10 keV XLF. Comparing these XLF estimates, discrepancies appear among them at low luminosities ($\log \mathrm{L}_{2-10\,\mathrm{keV}} < 43$) and higher redshifts ($z > 1.5$). We attribute the discrepancies to large uncertainties in the XLF determination. In order to unify the 2-10 keV XLF estimations in a single most up-to-date model independent result, we perform for the first time a meta-analysis of 2-10 keV XLF literature results. With the benefit of incorporating all previous evolutionary models, and consequently datasets, used in the literature we provide proper credible intervals and motivate the use of this result in further analysis when a 2-10 keV AGN XLF estimate is needed.

ZUSAMMENFASSUNG

Aktive Galaktische Kerne (Active Galactic Nuclei, AGN) werden durch die Akkretion von Gas auf die supermassereichen schwarzen Löcher (SMBH) im Zentrum der Galaxien angetrieben. Es wird vermutet, dass AGN eine wichtige Phase in der Galaxienentwicklung darstellen, wie die $M_{BH} - \sigma$ Relation suggeriert. Um diese Hypothese zu testen ist es notwendig die Entwicklung der AGN zu verstehen. Der Fokus der vorliegenden Arbeit ist die Röntgen - Leuchtkraftfunktion (XLF) der AGN, in den Energiebändern von 5-10 keV und 2-10 keV. Da eine große Stichprobe erforderlich ist, um die AGN Evolution vollständig zu studieren, wurden zuverlässige Rotverschiebungen für das Lockman Hole (LH) Feld erzeugt und eine beträchtliche Stichprobe von $\sim$ 500 AGN aus verschiedenen Studien zusammengestellt. Das Ergebnis ist eine Charakterisierung der Entwicklung von AGN im 5-10 keV Band, die auch in guter Übereinstimmung mit der 2-10 keV XLF ist, die zum ersten Mal als Meta-Analyse früherer Werke zusammengefasst wurde.

Das LH ist das Feld, das am wenigsten von galaktischem Wasserstoff entlang der Sichtlinie beeinträchtigt ist. Dadurch ist ein ungehinderter Blick auf den niedrigenergetischen Röntgenbereich möglich, während gleichzeitig der Infrarot-Cirrus minimal ist. Durch die Kombination von Katalogen aus ultravioletten, optischen und Infrarot-Daten wurde der erste homogene photometrischen Katalog im LH für $\sim$ 187600 Quellen erzeugt, der mehrere Wellenlängenbereiche abdeckt. Mithilfe dieses Katalog berechneten wir photometrische Rotverschiebungen, wobei die AGN die durch ihre Röntgenstrahlung identifiziert wurden, gesondert behandelt wurden. Die Genauigkeit dieser photometrischen Rotverschiebungen ist doppelt so gut, wie die zuvor Verfügbaren.

Im nächsten Schritt wurden mehrerer Röntgenbeobachtungen einschließlich dem LH zu einer großen Stichprobe zusammengeführt, die zum ersten Mal die Charakterisierung der XLF und seiner Entwicklung über die Zeit bei 5-10 keV ermöglicht. Diese Auswahl stellt im Gegensatz zu niedrigeren Röntgenenergien sicher, dass der gemessene Fluss von Absorption durch den photoelektrischen Effekt im Durchgang durch das Material rund um die SMBH unbeeinträchtigt bleibt. Die Bestimmung der XLF wird sowohl mit traditionellen Methoden wie 1/Vmax und der Maximum-Likelihood-Methode als auch mit einer ausführlicheren Bayesischen Analyse erreicht. Mehrere in der Literatur vorgeschlagene evolutionäre Modelle wurden geprüft, wobei mittels des Akaike Informationskriteriums und Bayesischer Modellwahl der aktuelle Datenbestand "Luminosity Dependent Density Evolution" als bestes Modell bevorzugt wird.

Mehrere Szenarien der AGN-Entwicklung wurden, mit inkonsistenten physikalischen Begründungen, in der Literatur für die Modellierung der 2-10 keV XLF vorgeschlagen und verwendet. Die gefundenen XLF unterscheiden sich insbesondere bei niedrigen Leuchtkräften ($\log \mathrm{L}_{2-10\,\mathrm{keV}} < 43$) und hoher Rotverschiebung ($z > 1.5$). Diese Diskrepanzen schreiben wir den großen Unsicherheiten bei der Bestimmung der XLF zu. Um die 2-10 keV XLF zu einem modellunabhängigen Ergebnis zu vereinheitlichen, führen wir zum ersten Mal eine Meta-Analyse der veröffentlichten 2-10 keV XLF durch. Mit dem Vorteil der Einbeziehung aller vorherigen evolutionäre Modelle und Datensätze, wie sie in der Literatur verwendet werden, veröffentlichen wir Vertrauensintervalle und schlagen vor, dieses Resultat in zukünftigen Analysen, welche die 2-10 keV AGN XLF benötigen, zu verwenden.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

_e_ROSITA	extended ROentgen Survey with an Imaging Telescope Array
AEGIS	All-wavelength Extended Groth strip International Survey
AGN	Active Galactic Nuclei
AIC	Akaike Information Criterion
APEMoST	Automated Parameter Estimation and Model Selection Toolkit
AU	Astronomical Unit
BH	Black Hole
BLG	Broad Line Region
BLRG	Broad Line Radio Galaxies
CDFN	Chandra Deep Field North
CDFS	Chandra Deep Field South
CFHTLS	Canadian-France-Hawaii Telescope Legacy Survey
COSMOS	Cosmic evolution Survey
CXRB	Cosmic X-ray Background
DEIMOS	DEep Imaging Multi-Object Spectrograph
ECDFS	Extended Chandra Deep Field South
EXTNV	Extended and Non-Varying
FDF	FORS Deep Field
FERMI-LAT	Fermi - Large Area Telescope
FUV	Far Ultra-Violet
FWHM	Full Width Half Maximum
GALEX	Galaxy Evolution Explorer
GSC	Gas Slit Camera

HBSS	Hard-Bright Serendipitous Survey
HR	Hardness Ratio
HST	Hubble Space Telescope
ILDE	Independent Luminosity and Density Evolution
IPAC	Infrared Processing and Analysis Center
IRAC	Infrared Array Camera
IRAF	Image Reduction and Analysis Facility
LADE	Luminosity And Density Evolution
LBT	Large Binocular Telescope
LDDE	Luminosity Dependent Density Evolution
Le Phare	Photometric Analysis for Redshift Estimate
LH	Lockman Hole
LINER	Low Ionization Nuclear Emission Line Regions
MAXI	Monitor of All-sky X-ray Image mission
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
NASA	National Aeronautics and Space Administration
NED	NASA/IPAC Extragalactic Database
NLG	Narrow Line Region
NLRG	Narrow Line Radio Galaxies
NUV	Near Ultra-Violet
PDE	Pure Density Evolution
PLE	Pure Luminosity Evolution
PSF	Point Spread Function
QSO	Quasi-Stellar Object
QSOV	Point-like or Varying
QUASARS	QUAsi-Stellar Radio Sources
ROSAT	Röntgensatellit
SDSS	Sloan Digital Sky Survey
SED	Spectral Enenrgy Distribution

SExtractor	Source Extractor
SSC	Solid-state Slit Camera
STILTS	Starlink Tables Infrastructure Library Tool Set
SWIRE	Spitzer Wide-Area Infrared Extragalactic Survey
TOPCAT	Tool for OPerations on Catalogues And Tables
UKIDSS	UKIRT Infrared Deep Sky Survey
UKIRT	United Kingdom Infrared Telescope
XBONG	X-ray Bright Optically Normal Galaxies
XLF	X-ray Luminosity Function
XMM	X-ray Multi-Mirror Mission

CHAPTER 1

Active Galactic Nuclei

Among the exiting discoveries in Extragalactic Astronomy is the realization that a few percent of the galaxies appear to have a tremendous energy release coming from their centers, which cannot be attributed to stellar emission. Collectively named *Active Galactic Nuclei* (AGN), these objects are thought to host a supermassive black hole at their centers which produces and sustains this energy release. Immediately after the discovery of AGN, it is was clear that these objects evolve strongly with time, contrary to what it was observed for normal galaxies and, efforts to determine the exact evolution have been taking place ever since.

Recent observations suggest not only that there might be a co-evolution between the AGN and the host galaxy, but additionally that AGN might be an *important phase* in galaxy evolution, as supermassive black holes have been recently detected at the centers of galaxies which show no signs of nuclear activity. Nevertheless, studies of inactive black holes, due to technological constrains, are limited even for nearby galaxies thus hampering any evolutionary studies. Therefore, before creating a link between the evolution of galaxies, it is crucial to understand exactly how AGN are evolving with time, which is the focus of this thesis.

In this introductory chapter, after presenting the concept of *redshift* (§1.1), used to determine the distances of galaxies, we will see how AGN entered the scene of extragalactic astronomy (§1.2) and how the combination of information across the electromagnetic spectrum led to the creation of a *Unified Model* to describe the central engine of AGN (§1.3). Additionally, we will show why X-ray selection is the best way in creating a census of AGN (§1.4) and how this census can be used to study the AGN evolution with time (§1.5.1 and §1.5.2). Lastly, we will give an overview of current research topics connecting AGN and galaxy evolution, demonstrating the need for accurate determination of the AGN *luminosity function* and, how with this work we are addressing common challenges associated to this computation (§1.6).

1.1 Cosmological Redshift

The year 1925 marks the birth of Extragalactic Astronomy. The first widely accepted distance determination of the Andromeda galaxy (M31) ended the several centuries old dispute on the origin of spiral nebula, establishing their extragalactic nature (Hubble, 1925). A few years later, the linear relation between radial velocity and distance of extra-galactic nebulae was put forward (Hubble, 1929). Hubble's relation between the radial velocity, u, of a galaxy and the distance, d, between the galaxy and the observer is:

$$u = H_0 \cdot d \tag{1.1}$$

where H_0 is a constant called, Hubble's constant (measured in $\mathrm{km\,s^{-1}\,Mpc^{-3}}$). This relation was considered as proof of the theoretically proposed expansion of the Universe as a solution to Einstein's equations (Lemaître, 1927).

For small radial velocities, this relation can be interpreted similarly to the Doppler effect. Let us assume for simplicity that a source is emitting a single emission line at wavelength λ_e. Thus, each wavefront of the radiation would be received by an observer at time intervals, Δt, given by:

$$\Delta t = \frac{\lambda_e}{c} \Longrightarrow \lambda_e = c \cdot \Delta t \tag{1.2}$$

where c is the speed of light. If this line emitting source is recessing relative to the observer at speed u, in the same time interval Δt the observed wavelength λ_o will be:

$$\lambda_o = (c + u) \cdot \Delta t \tag{1.3}$$

combining eq. (1.2) and (1.3) we get:

$$\frac{\lambda_o}{\lambda_e} = 1 + \frac{u}{c} \equiv 1 + z \tag{1.4}$$

where z describes the shift of the observed emission line with respect to the rest frame emission line. Since u is the relative velocity between the source and the observer, z can be either positive or negative. The heroic efforts of Slipher (1917) to obtain galaxy spectra showed that the majority of the objects displayed a positive shift in their emission lines, hence towards the red part of the optical electromagnetic spectrum. The term *redshift* is widely used to describe the recession of galaxies, even when a galaxy shows a negative shift[1].Thus, the combination of eq. (1.1) and (1.4), allows to determine distances of galaxies simply by measuring the redshift of known emission lines in the spectra of galaxies.

The equivalence of redshift and Doppler shift, is only valid when $u \ll c$. Larger redshifts, are interpreted as the combination of the galaxy's radial velocity and its recession due to the expansion of the Universe. Without going into the details here[2], it can be shown that assuming an expanding Universe with scale factor[3], R, Hubble's law becomes:

$$H = \frac{\dot{R}(t)}{R(t)} \tag{1.5}$$

where $\dot{R}$ is the rate of change of the scale factor R. Then, the redshift due to the expansion of the Universe is given by:

$$\frac{\lambda_o}{\lambda_e} = 1 + z = \frac{R(t_o)}{R(t_e)} \tag{1.6}$$

where $R(t_o)$ the scale factor of the Universe at time, t_o, when the photon was observed and equivalently $R(t_e)$ is the scale factor of the Universe at time t_e when the photon was emitted. Therefore, observations at different redshifts are snapshots of earlier epochs of the Universe, giving us the opportunity to perform evolutionary studies of galaxy populations with time.

---

[1] The term *blueshift* is also used. Interestingly, the Andromeda galaxy, the closest neighbor to our Milky Way has a negative shift of z=-0.001 since the two galaxies are approaching each other.

[2] More details can be found in Peacock (1999).

[3] The scale factor is the distance between two points that follows the expansion of the Universe

1.2 Discovery of AGN

Seyfert Galaxies

The first discovery of a peculiar class of galaxies came through spectroscopical observations of spiral galaxies with unusually bright nuclei. Seyfert (1943) showed that the spectra of what came to be known thereafter as *Seyfert galaxies*, exhibited strong hydrogen, helium, neon and oxygen emission lines which were understood to originate from high ionization regions in those galaxies. The observed widths of these lines correspond to a velocity of a few thousand kilometers per second, which means that the gas in which they are produced is moving fast. Further spectroscopical studies showed that Seyfert galaxies could be separated in two subcategories, Seyfert 1 and Seyfert 2 (Khachikian and Weedman, 1971, 1974). Type 1 Seyfert galaxies, show strong non-thermal continuum, broad lines of permitted transitions (Hα, Hβ, He I, width $\sim 10^3$km/sec) and narrow lines of permitted (He II, Hβ, Hα) and forbidden transitions[4] ([Ne V], [O II], [O III], etc, width$\sim 10^2$km/sec). Type 2 Seyfert galaxies show weak continuum emission and narrow emission lines.

Later on, Osterbrock (1977, 1981) showed that there exist also intermediate types of Seyfert galaxies, recognizing the categories Seyfert 1.9, 1.8, and 1.5. These intermediate classes have spectra of mixed Seyfert 1 and Seyfert 2 types. For example Seyfert 1.9 galaxies show weak continuum emission and narrow line emission placing them close to Seyfert 2 galaxies, but at the same time a faint broad Hα line is also observed. Penston and Perez (1984) showed that the broad lines and the strong continuum that had been observed in NGC 4151 in 1974, characteristics of a Seyfert 1 spectrum, were completely absent from a 1984 spectrum of the same source typical of Seyfert 2 spectrum. The authors, with an evolutionary scenario in mind between Seyfert 1 and Seyfert 2 galaxies, attributed this behavior to a temporary ‘switching-off’ of the central engine, or variability in the absorbing material.

Radio Galaxies

With the technological advancements of the 20th the possibility to study the Universe in wavelength ranges other than just the visible part of the spectrum became possible. The radio band was the first to be explored thanks to the pioneering studies of Karl Jansky, who was searching at the time for natural sources of radio emission which could interfere with radio transmissions. The first astronomical catalog of radio sources detected at 81.5MHz were compiled in Cambridge. The first (1C) and second (2C) version of the catalogs were found to contain spurious sources. The third version of the catalog (3C) contained sources detected at 159MHz, while the revised version (3CR, Bennett, 1962) contained sources detected at 178MHz. This catalog included, apart from supernova remnants within our Milky Way, mostly galaxies. Comparison to optical catalogs showed that the radio emission is associated mostly to elliptical galaxies and also at same cases to Seyfert galaxies.

In addition to these known galaxies, a puzzling population emerged from radio surveys. *Quasi-stellar radio sources* (quasars), named after their star-like appearance on photographic plates, were among the brightest objects in the radio sky but their nature remained unresolved. The mystery of the nature of these sources was resolved by Schmidt (1963), who identified the radio source 3C 273 as the nucleus of a galaxy with redshift of z=0.158 - the largest redshift measurement to that day. Sensitive optical observations showed a ‘fuzzy’ structure around quasars which we now identify as the host galaxy and collimated jets of material originating from the centers of the quasars. The optical spectra of quasars resemble those of Seyfert 1

---

[4] Forbidden transitions are low probability transitions, which can be observed only in low density gas.

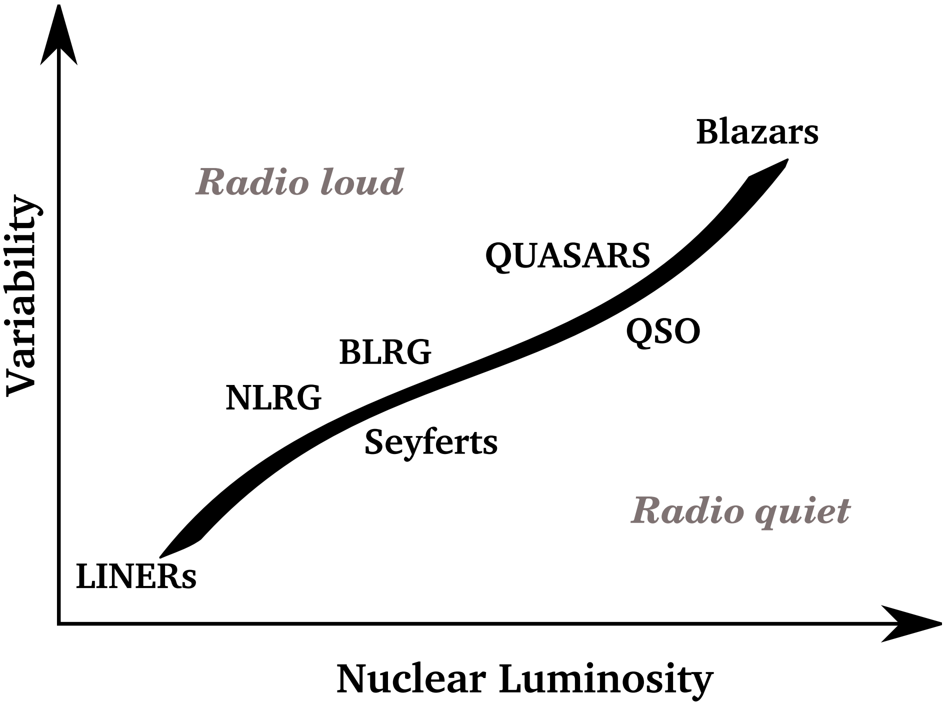


Figure 1.1: AGN classification scheme. Radio quiet objects typically have a few orders of magnitude smaller radio flux than radio loud objects, but they are 10 - 100 times more numerous.

galaxies, with an additional strong power law continuum and strong broad lines (line width $\sim 10^4$km/sec).

Nowadays, we know that AGN emitting in the radio band include apart from very bright quasars also less luminous *radio galaxies*, which also show jets of relativistic particles. They are further subdivided according to their optical spectra in *Broad Line Radio Galaxies* (BLRG, line widths $\sim 10^3$km/sec) containing both broad and narrow lines and *Narrow Line Radio Galaxies* (NLRG, line widths $\sim 10^4$km/sec) containing only narrow lines analogously to the distinction between Seyfert 1 and Seyfert 2 objects, but in both cases with broader lines compared to Seyferts.

QSOs

Sandage (1965), provided evidence that most of the blue star-like objects in color surveys at the visual part of the electromagnetic spectrum were actually the radio-quiet counterpart of quasars, named *Quasi-Stellar Objects* (QSO). QSOs were 10-100 times more numerous than quasars and exhibited similar emission lines and flux variability as observed in quasar and radio galaxy spectra. This evidence shaped the idea that there is a continuous sequence of galactic nuclei activity, ranging from the low activity systems like the Seyfert galaxies to the most energetic systems, the quasars (Fig. 1.1). Even thought these objects are named radio-*quiet*, is does not mean that they are lacking radio emission, rather the name suggests a lower radio emission compared to radio-*loud* objects (by a few orders of magnitude).

LINERs and Blazars

Two more classes of AGN have been identified at the two extreme ends of nuclei luminosity. On the low luminosity end lie sources with *Low Ionization Nuclear Emission line Regions* (LINERS, Heckman, 1980), which are found primarily in spiral galaxies and can be thought as low luminosity Seyfert 2 galaxies. At the other end of galactic activity, two radio sources serve as prototypes for very energetic, extremely variable AGN, commonly called blazars, which show almost a featureless power law continuum. Schmitt (1968) pointed out that the "irregular" variable star *BL Lacertae* was associated to the radio source VRO 42.44.01, while Strittmatter et al. (1972) suggested that BL Lac objects are a sub-class of quasars. Furthermore, the radio source 3C 279, serves a prototype for the sub-class of *Optically Violent Variable Quasar* (OVV), which shows dramatic flux variations with time in optical, radio and X-ray bands (respectively, Pica et al., 1988, Aller et al., 1985, Zamorani et al., 1984).

X-ray AGN

In the middle of the 20th century, the X-ray Universe became accessible to astronomers. Early pioneering rocket flights performing astronomical X-ray observations, revealed a very bright source in the constellation of Scorpius, Sco X-1 (Giacconi et al., 1962). Additionally, a uniformly distributed background radiation was observed in the X-rays, called the Cosmic X-ray Background (CXRB) (Gursky et al., 1963). With the increasing sophistication of rocket flights more and more X-ray sources were observed successfully, including the radio sources 3C 273 and M87. These early experiments demonstrated also the fact that X-ray sources can be highly variable and soon dedicated satellites were launched in space, since the Earth's atmosphere is opaque to X-ray radiation. The Uhuru X-ray observatory (NASA) was the first satellite launched and performed an all sky survey in the 2-20 keV band. First results, demonstrated the wealth of objects emitting X-rays including extragalactic sources such as quasars and Seyfert galaxies, and sources within our Galaxy such as stars, neutron stars, galactic black holes and, supernova remnants (Giacconi et al., 1971). With the advent of ROSAT (Röntegensattelit), the CXRB in the $0.5-2\,\mathrm{keV}$ energy band was largely resolved into discrete sources, mainly AGN (Hasinger et al., 1998). This fact was also confirmed in the $2-10\,\mathrm{keV}$ energy bands by the X-ray observatories Chandra and XMM-Newton (Moretti et al., 2003, Bauer et al., 2004, Worsley et al., 2004).

AGN X-ray spectra show a power law continuum in the $0.5-100\,\mathrm{keV}$ energy range with an exponential cut-off at 100 keV. The dichotomy between type 1 and type 2 objects is also used in the X-ray regime with type 1 being objects that do not show any signs of absorption in their spectra and type 2 objects that are at least partially absorbed showing a characteristic turn-over at lower X-ray energies. Additional characteristics of the X-ray spectra include the so called *soft-excess* which is increased flux below 1 keV with respect to the power law continuum, a reflection component at $\sim 30\,\mathrm{keV}$ and an iron line (Fe Ka 6.4 keV). X-ray spectra show rapid variability compared other energy bands of the AGN spectra, suggesting that the production of the X-ray radiation takes place in an area very close the black hole. It is worth mentioning that selecting AGN in the X-ray band is more efficient than in the optical. AGN have been found that emit X-rays but appear optically as normal galaxies (X-ray Bright Optically Normal Galaxies (XBONG), i.e. Elvis et al., 1981, Mushotzky et al., 2000, Barger et al., 2001). These objects are thought to have a 'hidden' AGN in the optical due to absorption.

The Central Engine

Theoretical understanding of the energy production mechanism responsible for the remarkable energy release which outshines in some cases the rest of the galaxy was needed. Salpeter (1964) and Lynden-Bell (1969) showed that the observed energy can be thought of as the transformation of kinetic energy, of incoming matter falling onto a central object, to gravitational potential energy. The rapid variability of the observed spectrum sets a limit on the physical size of the source. Let us assume that the central object has radius R. If the energy is produced at the center of this spherical source, then the time needed for the energy to reach the surface is equal to the timescale of the variation, $\Delta t \approx 1\,\mathrm{h}$. The energy can be transferred to the surface of the central engine at maximum with the speed of light, c. Then, the radius of the central source is roughly estimated to be:

$$R \approx c\Delta t = 1.1\,10^{14}\mathrm{cm} = 1.1\,10^{9}\mathrm{km} = 7\mathrm{A.U.} \tag{1.7}$$

At the same time, the maximum luminosity reached in the case of spherical gas accretion onto a central object, can be estimated assuming equilibrium between the gravitational force on the gas and the radiation pressure force. This limit is called the *Eddington luminosity, L_{edd}* and is given by:

$$L_{edd} = \frac{4\pi G c m_p}{\sigma_T} M_{obj} \tag{1.8}$$

where G is the gravitational constant, c the speed of light, m_p the proton mass, σ_T the Thomson scattering cross-section, and M_{obj} the mass of the central object. Quasar luminosities are typically L $\approx 10^{46}\mathrm{erg\,s^{-1}}$. Solving eq. (1.8) for mass we have that in order to produce the observed quasar luminosity the central object must have mass $M_{obj} \approx 8 \cdot 10^{7} M_{Sun}$. The concentration of this amount of mass confined in such a small area, rules out the possibility for the central object to be anything different than a supermassive black hole.

1.3 Unified Model

The striking similarities in emission lines, non-thermal continuum and variability among the several subclasses of AGN, point to the idea that the observed differences can be attributed to orientation effects and not to intrinsic differences. This idea originated from the fact that radio-loud and radio-quiet objects have indistinguishable optical spectral properties and the fact that polarization measurements of type 2 objects revealed broad emission lines similar to type 1 objects (Antonucci, 1982), suggesting that type 2 objects share the same central engine as type 1 objects, but it remains hidden.

Figure 1.2 shows a sketch of the *unified model* of AGN (Antonucci, 1993, Urry and Padovani, 1995). Moving outwards from the center, the most basic characteristic of an AGN is the presence of a supermassive black hole. The optical and ultra-violet continuum originates from a geometrically thin but optically thick accretion disk around the black hole. It is also believed that the accretion disk might contribute to the observed soft excess in the X-ray regime (Done et al., 2012). Above the accretion disk, ionized gas clouds irradiated from the disk are producing the optical broad emission lines (broad line region, BLR). In this area also the bulk of the X-ray production takes place. A population of thermal electrons up-scatters ultra-violet photons from the disk through Inverse Compton Scattering producing the X-ray radiation, while reflection of the X-ray radiation on the accretion disk gives rise to the 6.4 keV Fe Ka line. This system is surrounded by a optically opaque torus which is responsible for the absorption of the

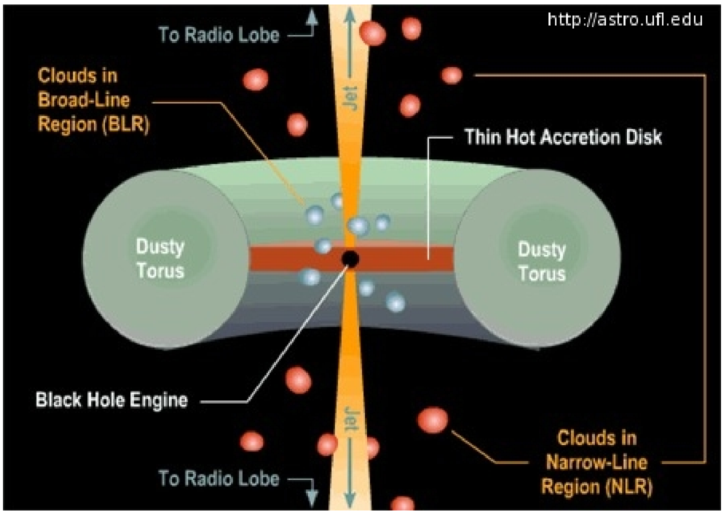


Figure 1.2: Sketch of the unified model of AGN. Orientation effects create the classification of AGN. Figure from http://astro.ufl.edu.

optical radiation which re-emitted in the infrared since the torus consists of gas and dust. Additionally, the obscuring torus is thought to produce the X-ray reflection component through repeat Compton Scattering of the X-ray spectrum produced by the thermal electrons. Above the obscuring torus, ionized gas clouds are the sources of narrow optical emission lines (narrow line region, NLG). Finally, the radio emission, is synchrotron radiation from relativistic particles in the jet.

The relative orientation between the AGN and the observer produces the dichotomy of type 1 and type 2 objects. In general, as type 1 are classified the sources that have both broad and narrow emission lines and do not show signs of obscuration. In this case, we are observing the AGN "face on". Contrary, type 2 objects have no broad lines and might show signs of obscuration also in other wavelengths, such as the X-rays. Then we believe that the viewing angle is "edge on" and the obscuring torus is blocking part of the radiation, including the optical broad lines. If a jet is present, the radio loudness of the source depends also on the orientation. If the axis of the jet is very close the line of sight, the source is radio-loud. If the jet points directly towards the observer, the relativistic beaming effects make the source appear very luminous in the radio band and often extremely variable. Then the source is identified as a blazar. Lastly, when the jet is not pointing towards the observed, the source is radio-quiet. The latter case includes also sources where a jet is completely absent.

Nevertheless, assuming - as the unified model implies - that the sole difference among AGN is the orientation, might introduce biases to inferred properties of the whole AGN population such as the fraction of Compton Thick objects (Elitzur, 2012, , and references therein). The now established clumpiness of the torus (Nenkova et al., 2008) suggests that the separation between a type 1 and type 2 object can be explained as the transition of an obscuring cloud along the line of sight. Such behavior, would naturally explain the transition between Seyfert 1 and Seyfert 2 (e.g. in NGC 4151), without assuming any changes on the central engine. The combination of multiwavelength detailed observations and numerical modeling will give us a more detailed view of AGN in the near future.

1.4 X-ray AGN Selection

X-ray selection of AGN is the most efficient method in selecting *unbiased and least contaminated* samples of AGN. With this selection, the majority of the AGN population is retrieved since X-ray radiation escapes the system almost intact in all but the most extreme cases of obscuring material (Brandt and Hasinger, 2005). In the following we will briefly see why that is.

1.4.1 X-ray obscuration

The obscuration of the X-ray radiation can be either intrinsic to the source or induced by the intergalactic and interstellar medium. Essentially, X-ray photon interactions relevant to obscuration can be described as an energy exchange between photons and electrons. The outcome of these interactions depend on the photon's energy, the electron's energy, the nature and density of the absorbing material. The latter is characterized through the *hydrogen column density*, N_H (measured in particles/cm^2).

In the limit where the photon's energy, $\hbar\omega$, is much smaller compared to the electron's rest mass energy, mc^2, the scattering between photons and free electrons is elastic and there is no energy transfer between the two interacting parties. This is called *Thomson scattering* and the optical depth, τ_e, of this process is given by:

$$\tau_e = N_e \sigma_T l \tag{1.9}$$

where, N_e, the electron density (measured in particles/cm^3), σ_T, the Thomson scattering cross-section and, l, the size of the region occupied by electrons. In this energy regime, photons loose energy primarily through *photoelectric absorption*, which is the interaction between a photon and a bound electron. In the case of X-ray radiation, this would be a K-shell electron which is completely expelled from the atom while the photon is absorbed. This absorption mechanism is dominant mostly in the lower X-ray regime ($\lesssim 1\,\text{keV}$).

On the other hand, if the photon's energy, $\hbar\omega$, is much larger than the electron's rest mass energy, mc^2, the dominant mechanism is *Compton scattering*. This scattering refers to the interaction between a photon and a free or valence electron. The Compton scattering becomes the dominant mechanism when the Thomson optical depth is much larger than unity ($\tau_e > 1$). From eq. (1.9) we see that this corresponds roughly to:

$$N_e \sigma_T l > 1 \Longrightarrow N_H \sigma_T > 1 \Longrightarrow N_H > 10^{24} cm^{-2} \tag{1.10}$$

where we use the approximation $N_e = N_H l$ and the fact that the value of the Thomson cross-section is $\sigma_T = 0.66\,10^{-24} cm^{-2}$. The resulting spectrum is the outcome of repeated scattering of the incident photons on the electrons and depends on the temperature of the electrons

Therefore, depending on the sensitivity of the observations and the redshift of the source[5], AGN with large intrinsic hydrogen column density ($N_H > 10^{23} cm^{-2}$), could be missed from X-ray surveys ($0.5 - 2\,\text{keV}$) due to photoelectric absorption, while sources with ($N_H > 10^{24} cm^{-2}$) are completely absorbed by Compton scattering and they are called *Compton Thick*.

---

[5] Due to the redshift of the spectrum, our observing band evolves as energy*(1+z), so at higher redshifts we observe higher rest frame energies which are affected less from the absorption.

1.4.2 Sample contamination

In addition to identifying the vast majority of the AGN population, X-ray AGN selection is the *least contaminated* by non-AGN sources. That is because the X-ray radiation by normal galaxies is very low compared to AGN emission, and only in the deepest X-ray observations non-AGN sources start to have a contribution in the observed population. In the currently available deepest X-ray observations (Chandra X-ray observatory, 4Ms, Xue et al., 2011) no more than 35% of the detected sources are identified as non-AGN.

A common threshold used for separating AGN from normal galaxies is selecting sources with luminosity $L_x > 10^{42}\,\mathrm{erg\,sec}^{-1}$, since only AGN are able to radiate at so high luminosities. At lower luminosities, the extragalactic X-ray sky includes both low luminosity AGN and star-forming galaxies. This is the selection criterion we will use to identify our AGN sample in Chapter 4, to study the evolution of AGN.

It has also been found that a combined selection criterion of X-ray and optical flux, namely:

$$-1 < \log(\frac{F_x}{F_{opt}}) < 1 \tag{1.11}$$

can distinguish efficiently AGN from normal galaxies, star-forming galaxies and, stars (Maccacaro et al., 1988, Hornschemeier et al., 2003). In section 3.1.1, we show that this criterion is in agreement with the selection of point-like or varying X-ray detected sources, which we adopt in Chapter 3 to separate between AGN dominated sources and sources for which the optical radiation is a combination of the host galaxy and the AGN component.

1.5 AGN Evolution

1.5.1 Evidence for Evolution

The subject of galaxy evolution already concerned astronomers from the early days of extragalactic astronomy. Hubble using galaxy counts “in successive spheres” (Fig. 1.3, Hubble, 1936) claimed that the observed distribution of galaxies is uniform for bright and nearby galaxies but it has discrepancies from the expected Euclidean counts as fainter and thus more distant galaxies are included in the sample. Nevertheless, since galaxies are evolving passively in the sense that the change of luminosity depends primarily on the evolution of the stellar population with time, number counts of normal galaxies reveal more information on the geometry of the Universe rather than the evolution of the population itself.

Another way to test deviation from the uniform distribution of sources in space is the V/V_{max} test, proposed by Schmidt (1968) who demonstrated that quasars are evolving with redshift. In order to determine the source number density independently of the expansion of the Universe, the co-moving volume is used. The comoving volume is the volume in which the number of a non-evolving population remains constant as the Universe expands, computed as:

$$dV_c(z) = \frac{c}{H_0}\frac{D_M^2}{E(z)}d\Omega dz \tag{1.12}$$

where c is the speed of light, H_0 Hubble’s constant, and Ω the observed area on the sky.
The quantities D_M and $E(z)$ are given by:

$$D_M = \frac{c}{H_0}\int_0^z \frac{dz'}{E(z')} \tag{1.13}$$

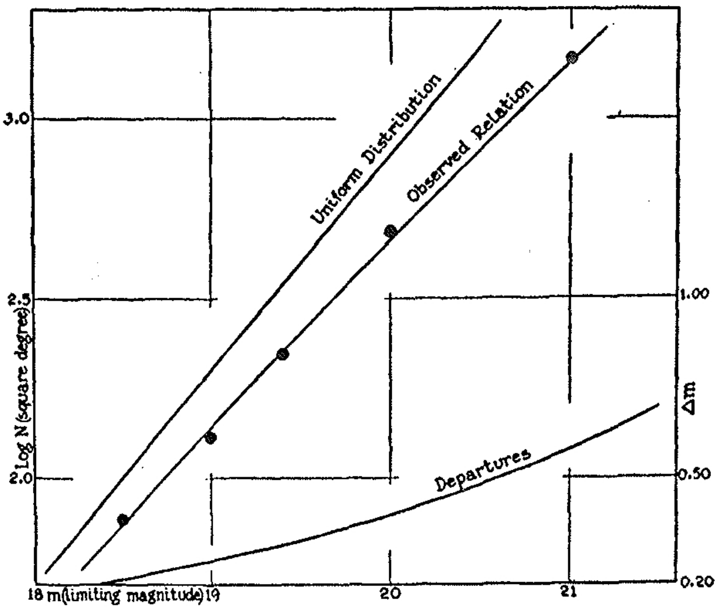


Figure 1.3: Logarithm of cumulative galaxy number counts, N, as a function of apparent magnitude from Hubble (1936). The author claimed that bright sources can be thought to originate from a uniform distribution in space, while fainter and thus more distance sources deviate from this distribution. This deviation is attributed to evolution of the galaxy population.

and

$$E(z) = \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda} \tag{1.14}$$

where Ω_M and Ω_Λ are the matter density and energy density of the Universe.

This test is applied to flux limited samples when redshift information is available for all the sources in the sample. With this test, for each source located at redshift z_i, we divide the comoving volume up to redshift z_i, V_i with the maximum volume $V_{max,i}$ up to which the source would still be above the flux limit, and thus detected in the sample. If this ratio is 0.5 it means that there is no evolution in the population. If the ratio in greater than 0.5, then the number of objects increases with redshift, while if it is less than 0.5, means that the number of objects decreases with redshift.

Schmidt (1968) using quasars from the *Revised 3C catalog* of radio sources[6] detected at 178 MHz (Bennett, 1962), showed that the mean value of V/V_{max} is 0.7 which indicates that the number density of quasars changes with redshift. Very recent studies of Chandra Deep Field South (CDFS) radio sources detected at 1.4 GHz with a flux limit of 43 μJy show, using the same method, that the sub-populations of the radio sky have a mixture of evolutionary behaviors (Padovani et al., 2011). According to this study, radio-quiet AGN show strong evolution with $V/V_{max} = 0.73 \pm 0.05$. On the other hand, radio-loud AGN with low power ($P < 24.5\,\mathrm{W\,Hz^{-1}}$) do not appear to evolve ($V/V_{max} = 0.43 \pm 0.04$) but high-power radio-loud AGN ($P > 24.5\,\mathrm{W\,Hz^{-1}}$) show negative evolution ($V/V_{max} = 0.28 \pm 0.05$).

---

[6] Northern sky ($\delta > 0.5°$), flux limit $9 \cdot 10^{-26} Wm^{-2}(c/s)^{-1} = 9Jy$

1.5.2 Luminosity Function

The *luminosity function* ,Φ, describes the number of objects, N, of given population per unit comoving volume, V_c (eq. 1.12), per unit luminosity, L. When computed in certain redshift bins, it reveals whether the population is evolving with redshift, luminosity or both.

$$\Phi(L,z) = \frac{d^2N}{dL\,dV_c} \tag{1.15}$$

The most straightforward method to estimate the luminosity function is the $1/V_{max}$ method, developed and applied by Schmidt (1968) on radio AGN. Creating logarithm-luminosity and comoving volume bins or equivalently redshift bins, the luminosity function is given by:

$$\frac{d\Phi(L,z)}{d\log L} = \frac{1}{\Delta \log L}\sum_{i=1}^{n}\frac{1}{V_{max,i}} \tag{1.16}$$

where V_{max} is either the volume defined by the upper limit of the redshift bin, or in case the bin is truncated by the flux limit of the survey, it corresponds to the maximum volume sampled according to the flux limit. This method is powerful in describing the shape of the luminosity function and its evolution with redshift with the assumption being the cosmological framework, which is typically specified from other experiments. The weak points lie in the limited available number of observed AGN, and the presence of flux limits in surveys which can truncate the luminosity - redshift bin, thus providing a biased estimation of the lowest luminosity bins.

Application of this method in radio (Laing et al., 1983), optical (Schmidt and Green, 1983), and X-ray (Maccacaro et al., 1991) selected AGN, showed that the AGN population is evolving with redshift. Due to the limited number of objects available until recently, there was no definite answer to whether the evolution should be described as *Pure Luminosity Evolution* (PLE, Schmidt, 1968), or as *Pure Density Evolution* (PDE, Schmidt, 1968). The PLE scenario would be possible in the case where the number density of AGN remains constant with redshift, but their luminosities are gradually declining from high redshift ($z \approx 2$) towards lower redshift ($z \approx 0$). The PDE scenario would be possible, if active AGN at redshift $z \approx 2$ would "turn off" as they evolve in order to match the lower number density observed in local universe.

Parametric methods have also been used to determine the evolution of the luminosity function, notably the Maximum Likelihood estimation (MLE, Marshall et al., 1983). Assuming a certain evolutionary model, the likelihood of detecting the observed sample at certain ($\mathrm{L_i}, \mathrm{z_i}$) pairs is constructed. Since the detection of n sources out of the N sources present in the Universe can be approximated as a Poisson processes (for small values of n), the likelihood is given by:

$$\ln \mathcal{L}(L,z) = \sum_{i}^{n} \ln \Phi(z_i, LogL_i)\frac{dV_c}{dz} - \int\int \Phi(z, LogL)\Omega(z, LogL)\frac{dV_c}{dz}dzdLogL \tag{1.17}$$

wide where Φ the model describing the luminosity function, dV_c/dz the differential comoving volume given by eq. (1.12), Ω the detection efficiency of the survey. The best parameters that describe the dataset are retrieved through the minimization of the quantity $S = -\ln\mathcal{L}$. In what follows, we summarize recent results on the X-ray luminosity function and its evolution.

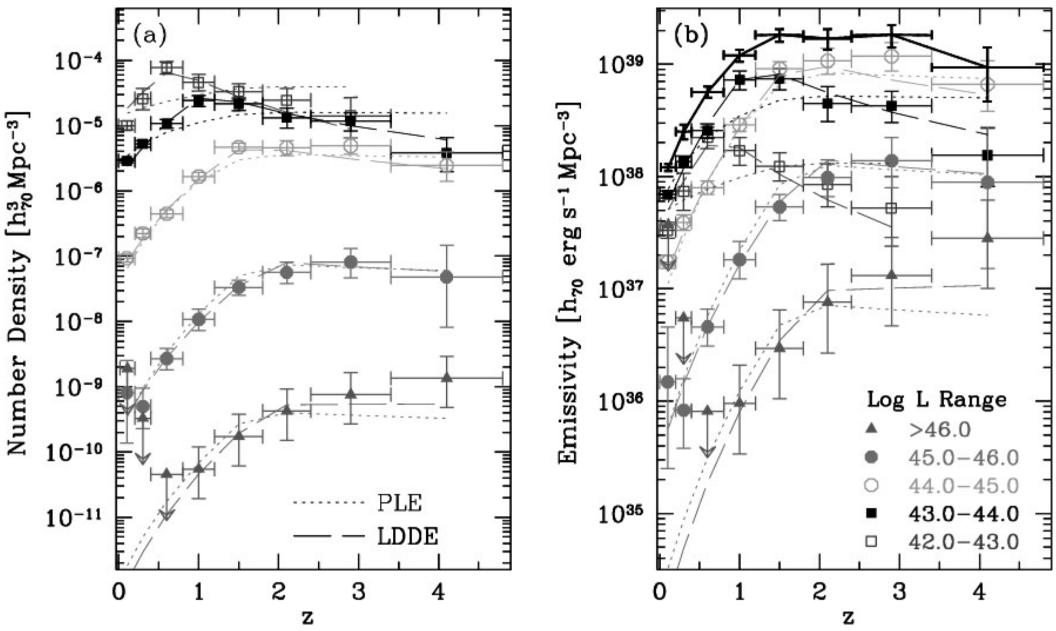


Figure 1.4: Type I AGN luminosity function in the 0.5-2 keV band. Left panel: AGN number density as a function of redshift for several luminosity bins. High luminosity ($\log L_x > 44.0$) AGN show a peak in number density around redshift 1-2. Right panel: volume emissivity of AGN as a function of redshift for several luminosity bins. Plot from Hasinger et al. (2005)

Recent results

Miyaji et al. (2000) and more recently Hasinger et al. (2005), using a numerous sample of ~ 700 and ~ 1000 AGN, respectively, detected in the soft X-ray band ($0.5-2$ keV), demonstrated that the evolution of the AGN luminosity function is described best as a *Luminosity Dependent Density Evolution* (LDDE, Fig. 1.4). Similar results showing a more complicated evolution than the initially proposed simple PLE and PDE models, are reached also with optically selected AGN samples (Bongiorno et al., 2007). Subsequent works on X-ray luminosity functions, strive to determine the evolution also in hard X-ray bands, such as the $2-10$ keV and $5-10$ keV bands.

The challenges associated to the determination of the luminosity function include the number of available AGN in the sample, the completeness in redshift of the sample, the exact knowledge of the spectral energy distribution of each source and, the intrinsic absorption of the X-ray radiation induced by the environment around the black hole. Due to the aforementioned challenges, AGN evolution studies in energy ranges above 2 keV have not been conclusive. Works suggest that the same LDDE behavior describes best the data (Ueda et al. (2003), La Franca et al. (2005), Silverman et al. (2008), Ebrero et al. (2009), Yencho et al. (2009)) compared to the simpler PLE evolution. Recent work by Aird et al. (2010) proposed that the evolution of AGN in the $2-10$ keV can be described as *Luminosity and Density Evolution* (LADE). The difference compared to the LDDE model, is that the evolution of AGN with redshift is described to happen simultaneously on luminosity and number density but without any dependence on the luminosity.

Moreover, the above works seem to reach contradictory conclusions regarding the shape

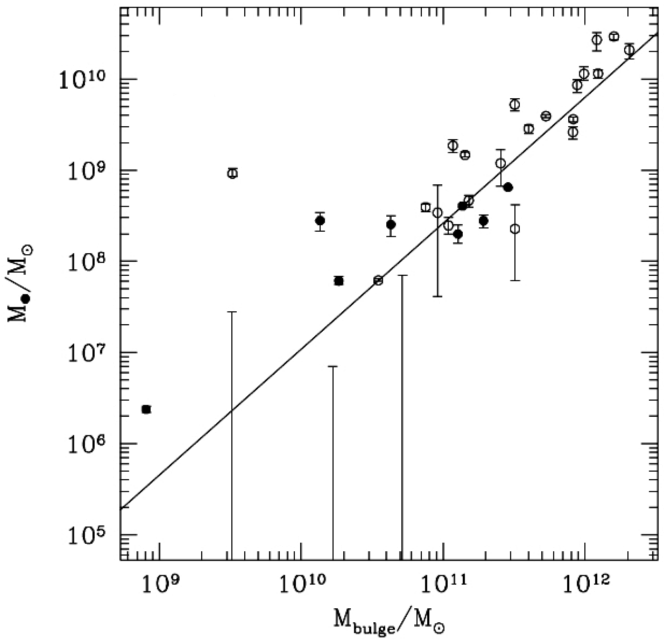


Figure 1.5: Proposed correlation between the black hole mass and the bulge mass for a sample of bright nearby galaxies. Open and filled circles refer to the light profile of the bulge (power-law and core respectively). Figure adopted from Magorrian et al. (1998).

of the $2 - 10$ keV luminosity function. Even when the same evolutionary model is used, different estimations of the luminosity function show large discrepancies, especially at low luminosities. This effect is attributed mostly to incomplete redshift information, and poor redshift determination. Nevertheless, as we will show in Chapter 5, this discrepancy is lifted once the uncertainties of the model parameters are taken into account properly, and each model is used within its range of validity.

Additionally, the exact behavior of the luminosity function at $z > 3$ is still under debate. Schmidt et al. (1987) using optical observations of quasars suggested that there is a strong number density decline at redshift $z > 3$. Recently, additionally evidence started to emerge showing a strong decline in the AGN number density at these high redshifts both in the X-rays (Brusa et al., 2009, Civano et al., 2011) and optical (Ikeda et al., 2011, Glikman et al., 2010, 2011, Masters et al., 2012). This behavior was not captured in the evolution determined in previous studies for the simple reason that very limited data were available at higher redshifts. As we will see in Chapter 4, our sample extends up to redshift 4, which enables the comparison of our modeling of the luminosity function to the results obtained from high redshift studies.

1.6 AGN – Galaxy Co-evolution?

Recent developments in galaxy studies suggest that galaxies with a spheroidal component (bulge) host a massive black hole in their centers, and that correlations exist between the mass of the black hole and properties of the bulge such as the mass of the bulge and the velocity dispersion. Magorrian et al. (1998) studying a sample of 32 nearby normal galaxies reached the

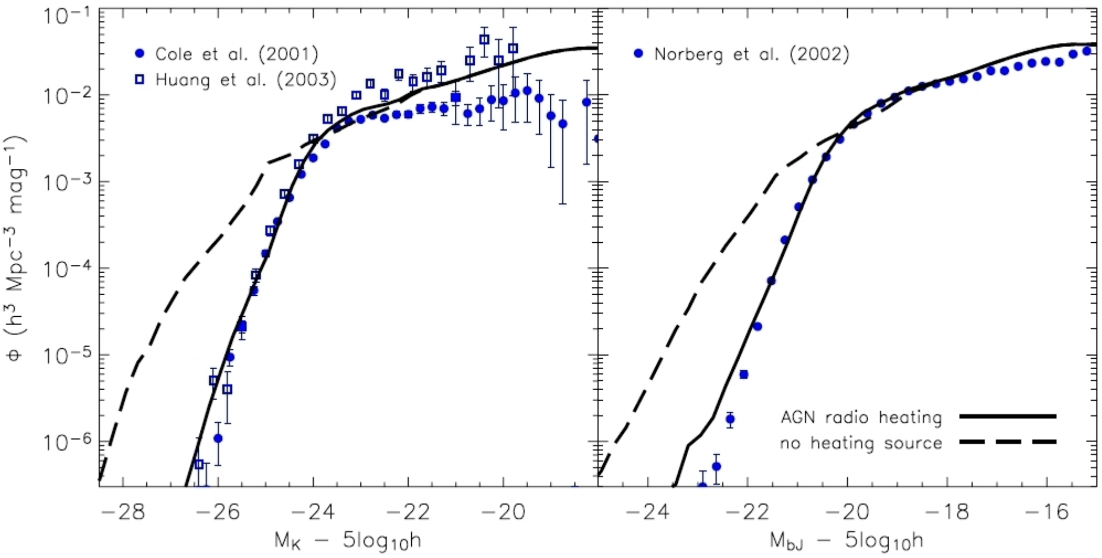


Figure 1.6: Normal galaxy luminosity functions in the K and B_J bands. The blue points show observations from Cole et al. (2001), Norberg et al. (2002), Huang et al. (2003). The dashed line is the predicted luminosity function from the semi-analytical models when no AGN feedback is assumed. The solid line depicts the model in which AGN feedback is included and is able to represent the observed data adequately. Figure adopted from Croton et al. (2006).

conclusion that there is a linear relation between the logarithm of mass of the ”massive dark object” - as they call it - and the logarithm of the mass of the galaxy bulge (Fig. 1.5).

Further studies (Ferrarese and Merritt, 2000, Gebhardt et al., 2000) showed that there exists also a correlation between the mass of the central black hole (M_{BH}) and the dispersion of the velocities of the stars in the bulge (σ), widely known at the $M - \sigma$ relation. The importance of the $M - \sigma$ relation does not lie primarily in the fact that the correlation merely exists but rather in the fact that the scatter in the $M - \sigma$ relation is very small. As a result, this relation provides a test of galaxy formation theories, while at the same time, allows the estimation of black hole mass in normal galaxies using the more easily observable quantity, the star’s velocity dispersion in the bulge.

The evidence for the existence of black holes in the centers of the majority of galaxies, at least the more massive ones, poses several questions directly linked to the nature of AGN and also to their role in galaxy evolution. The first question is whether AGN are a phase in galaxy evolution, possibly reoccurring. To support this hypothesis, the detection of evidence of past activity associated to a currently inactive black hole is required. Such a task, is very challenging to be achieved with current observational means. Nevertheless, recent FERMI-LAT γ-ray observations in our Milky Way suggest that two bubbles of relativistic electrons exist perpendicular to the Galactic center (Su et al., 2010) that can be interpreted as sign of past accretion event in our now quiescent Galaxy (Zubovas and Nayakshin, 2012).

Whether or not galaxies have experienced one or more AGN phases, is still matter of debate as is the triggering mechanism for AGN activity. Current observations suggest that there are two main mechanisms that might ignite an AGN (Hasinger, 2008). The first, is triggering through merging of galaxies typically called the major-merger scenario (Hopkins et al., 2006). The second mechanism, is through secular processes, whereby the black hole increases its mass from the inflow of gas from the galaxy (Kormendy and Kennicutt, 2004, Hopkins and Hernquist, 2006). Even though it is clear that both mechanisms appear in nature, evidence that low luminosity AGN grow primarily through secular processes have started to accumulate (Georgakakis et al., 2009, Cisternas et al., 2011, Orban de Xivry et al., 2011, Allevato et al., 2011).

The possibility that each galaxy could in some point host an AGN, triggered a new topic of research, aimed to link the AGN phase to galaxy evolution and cosmological studies. For example, a possible interplay between the evolution of the black hole and the evolution of the galaxy through feedback mechanisms is being addressed. The two main feedback mechanisms studied are the "quasar-mode" feedback whereby the black hole self-regulates its growth by blowing out the infalling material with a powerful wind, and "radio-mode" feedback where energy and matter are transferred back to the galaxy and the large scale environment through collimated outflows. Semi-analytical models of galaxy evolution combined with cosmological simulations of dark matter show that the inclusion of radio mode feedback from AGN at the centers of galaxy clusters is necessary to reproduce the exponential cut-off at the bright end of the galaxy luminosity function (Fig. 1.6, Croton et al., 2006).

Furthermore, a link between the AGN and the star-formation in the galaxy is being studied. Since both AGN and star formation share the same fuel, namely gas, it is interesting to see if a wind from the AGN, especially in the case of quasars, would expel enough gas from its vicinity in order to halt star formation (Croton et al., 2006). Additionally, when gas is transported towards the inner area of the galaxy fragmentation and gravitational collapse with the subsequent formation of stars would prevent the gas from reaching the black hole. Interestingly, the star-formation history of galaxies seems to peak about $z \sim 2$ at the same redshift as the AGN activity for $L \sim 10^{43}$, and additionally shows a rapid decline at higher redshifts (Bouwens et al., 2007, and references therein).

Understanding how the AGN population evolves with time is an important step in answering the previously mentioned outstanding questions. The luminosity function is the best approach to constrain the dependency of the AGN evolution on redshift and/or luminosity. Since the X-ray band provides the most unbiased and least contaminated AGN selection, a census of X-ray selected AGN is needed.

The creation of a complete sample of AGN, is hampered by issues such as the partial counterpart identification of the X-ray sources. Since X-ray telescopes do not reach the high angular resolution achieved in optical telescopes, the counterpart identification process often relies on a probabilistic assignment of the best optical counterpart selecting among a few possible sources. Additionally, accurate redshift determination is a challenging task. Spectroscopic observations provide a very accurate determination of redshift, but the process is very time consuming with the additional barrier that high quality spectra are rarely obtained for faint sources. Therefore, often the redshift has to be estimated less accurately through photometry whereby the observed Spectral Energy Distribution (SED) is compared to theoretical and empirical SED templates, in order to select the template-redshift combination which describes best the observed SED. Moreover, the photometric redshift determination is particularly challenging when applied for X-ray sources, since the SED of each source is a unique combination of radiation from the host galaxy and the AGN. Finally, corrections for intrinsic absorption of the X-ray radiation must be taken into account. The absorption can be estimated from the X-ray spectra of AGN, but in the case when spectra are not available, a rough estimation is performed from flux measurements in two or three X-ray energy bands. Even though this estimation can give valuable insight on the physics of the source, it remains a rough estimation with large uncertainties associated.

With this work, by using recent high quality multiwavelength surveys we create a sizable sample of AGN detected in the $5-10$ keV band. Due to the available multiwavelength information, the computation of accurate photometric redshifts was possible for all of the fields utilized here. Furthermore, the $5-10$ keV band selection renders the absorption corrections unnecessary. Lastly, applying rigorous analysis methods we determine the evolution parameters and select the model which describes best our observations.

1.7 Outline of this Dissertation

Modern astronomical surveys typically consist of multiwavelength observations, in order to fully understand every possible physical aspect of the sources under investigation. Usually, extragalactic observations have to be corrected for galactic extinction imposed by the intergalactic medium. In this work, we study the field with the least galactic hydrogen column density along the line of the sight, the Lockman Hole Deep Field. Due to this physical characteristic this field is a window to the extragalactic Universe. After homogenizing deep optical images, we compile the first extensive photometric catalog consisting of 21 broad photometric bands for the Lockman Hole Deep Field ranging from the ultra-violet to the infrared. Details of the creation of this catalog are discussed in Chapter 2.

Using the photometric catalog we compiled for Lockman Hole, we determine accurate photometric redshifts for all the sources in the field, discussed in Chapter 3. The achieved photometric redshift accuracy is $\sigma \sim 0.03$, with a fraction of 10% outliers. Since we are interested in using this field to determine the X-ray luminosity function of AGN, we pay special attention to the reliable determination of photometric redshifts for the X-ray detected sources, and we achieve the best possible accuracy for this sources ($\sigma \sim 0.07$, with 19% outliers). Both the accuracy and the fraction of outliers is a factor of two improvement over previously available photometric redshifts, enabling the use of this field in statistical studies of galaxy and AGN populations. The photometry and photometric redshift catalogs along with the data reduction process have been published in Fotopoulou et al. (2012). Additionally, the photometric redshifts have aided in the detection of one of the highest redshift X-ray detected galaxy clusters (Henry et al., 2010).

We identified optical counterparts to the X-ray sources in the Lockman Hole (published in Rovilos et al., 2011) and combined this very deep X-ray detections with similarly deep X-ray observations from the Chandra Deep Field South (CDFS), thus probing high redshift and low luminosity AGN. With the additional use of more extensive fields such as the Monitor of All-sky X-ray Image (MAXI), Cosmological Evolution Survey (COSMOS), and Hard Bright Serendipitous Survey (HBSS), we create a sizable sample of 500 AGN detected in the $5-10$ keV band. This is the first compilation of high quality data in this energy range which enables the accurate determination of the luminosity function and its evolution with redshift. We show that the $5-10$ keV energy band, is largely unaffected by photoelectric absorption, and thus we do not utilize crude assumptions in order to correct for this effect. Additionally, the use of well studied fields with multiwavelength information and therefore accurate photometric redshift determination, enables the creation of our sample with 98% redshift completeness. As we discuss in Chapter 4, we deploy several statistical analysis methods in order to determine the luminosity function ranging from the $1/V_{max}$, to Maximum Likelihood Estimation and the more elaborate Bayesian Analysis. We show using the Akaike Information Criterion and Bayesian Model Selection that the Luminosity Dependent Density Evolution model, describes best the evolution of the $5-10$ keV AGN luminosity function. For the first time, our model encapsulates the decline of AGN number density with redshift observed for higher redshift samples.

Furthermore, we show in Chapter 5 that the uncertainties of the model parameters used to describe the luminosity function are correlated. Therefore, we motivate the inclusion of the normalization of the luminosity function as an unknown variable during the parameter estimation, since the correlation of the parameters leads to reduced uncertainties in the luminosity function. Identifying the fact that the discrepancies among previous luminosity function estimates in the $2-10\,\mathrm{keV}$ band are less prominent when the uncertainties on the model parameters are properly taken into account, we create the first unified luminosity function. This *meta*-luminosity function is the combination of all available results in the literature, and has the benefit of incorporating all previously used evolutionary models and datasets. Our result in the 5-10 keV is in very good agreement with this unified luminosity function when transformed in the 2-10 keV band. Lastly, in Chapter 6 we summarize our work along with the main results and provide an outlook for future work.

CHAPTER 2

Lockman Hole Deep Field – Photometry

The Lockman Hole is the area on the sky with the lowest galactic hydrogen column density along the line of sight ($N_H \approx 5.7\,10^{19}\mathrm{cm}^{-2}$, Lockman et al. 1986, Schlegel et al. 1998). This physical characteristic provides the opportunity to perform extragalactic observations without significant absorption of the radiation in the soft X-rays and the ultra-violet and with minimal galactic cirrus emission in the infrared. For this reason the field has been observed in all energy bands from the Radio to the X-rays. A detailed overview of the various observations is given in Rovilos et al. (2009).

Up to now the spectroscopic follow up was dedicated to sources detected in specific bands (X-ray, infrared, radio). In addition, the photometry of the field was lacking crucial bands such as J and K, hampering the possibility to compute accurate photometric redshifts. The only public photometric redshifts for approximately half of the area of the field examined in this work are available via the SWIRE survey (Rowan-Robinson et al., 2008), computed using photometry in 4 broad band optical filters (U', g', r', i') and $3.6\mu m$ and $4.5\mu m$ mid-infrared filters from *Spitzer/IRAC*. Furthermore, the photometric redshifts were lacking proper treatment for the AGN.

In this chapter, we discuss the photometric observations and their analysis. We first cross-calibrate and then combine publicly available and private photometry from the ultra-violet to the mid infrared wavelengths, creating an extensive catalog, which we further use to compute photometric redshifts, discussed in the next chapter. We also provide additional information on the variability and morphology which will be of value during the computation of photometric redshift for the X-ray detected sources.

2.1 Data Set

In this section we describe the data that we used in each energy band, separated by instrument/satellite. Figure 2.1 shows the filter transmission curves of the available observations. In Figure 2.2 we give an overview of the coverage of Lockman Hole area in each energy band. It is evident from this figure that the best quality results are expected in the area covered by XMM-Newton (black circle).

The main part of this chapter has been published by Fotopoulou et. al (2012), *The Astrophysical Journal Supplement Series, 198, 1*, entitled *Photometry and photometric redshift catalogs for the Lockman Hole deep field.*

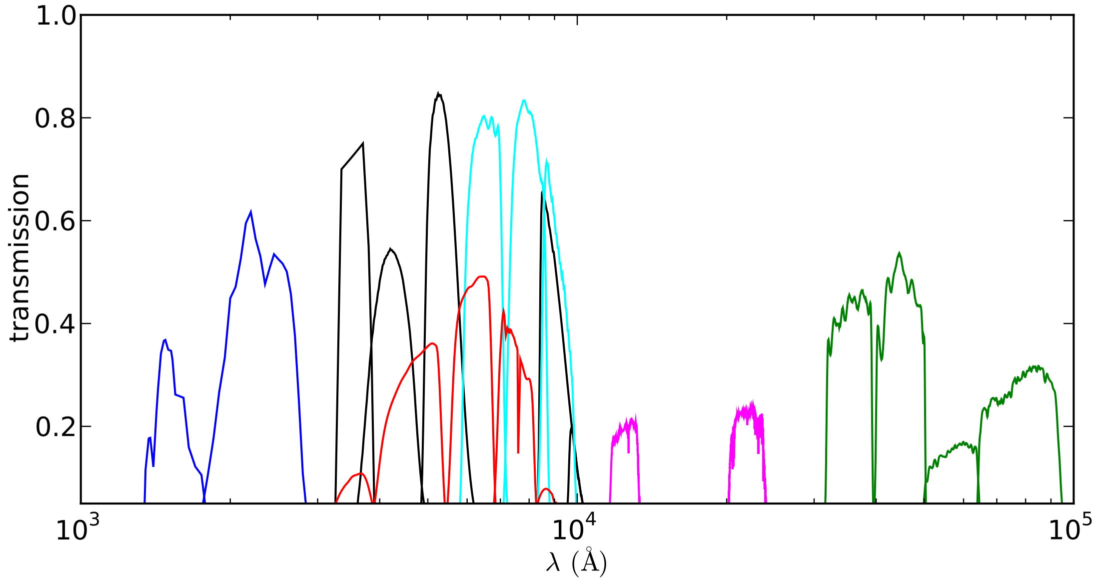


Figure 2.1: The filter coverage of the Lockman Hole Deep Field. Blue: GALEX (FUV, NUV), black: LBT (U, B, V, Y, z′), cyan: Subaru (R_c, I_c, z′), red: SDSS (u′, g′, r′, i′, z′), magenta: UKIDSS (J, K), green: Spitzer (3.6μm, 4.5μm, 5.8μm, 8.0¯m).

Table 2.1: Properties of recent X-ray fields sorted in decreasing area

Field	Area	Exposure	No.	Flux limit		Catalog
	(deg^2)	(ks)	sources	0.5 – 2 keV	2 – 8 keV	
XMM-COSMOS	2.13	60	1887	$1.7 \cdot 10^{-15}$	$9.3 \cdot 10^{-15b}$	Cappelluti et al. (2007)
Chandra-COSMOS	0.90	200	1761	$1.9 \cdot 10^{-16}$	$7.3 \cdot 10^{-16}$	Elvis et al. (2009)
AEGIS	0.67	200	1325	$5.3 \cdot 10^{-17}$	$3.8 \cdot 10^{-16}$	Laird et al. (2009)
ECDFS	0.30	250	762	$1.1 \cdot 10^{-16}$	$6.7 \cdot 10^{-16}$	Lehmer et al. (2005)
Lockman Hole	**0.20**	**560**	**409**	$\mathbf{1.9 \cdot 10^{-16}}$	$\mathbf{9.0 \cdot 10^{-16b}}$	**Brunner et al. (2008)**
CDFN	0.12	1000^a	503	$2.7 \cdot 10^{-17}$	$1.4 \cdot 10^{-16}$	Alexander et al. (2003)
CDFS	0.13	2000^a	740	$9.1 \cdot 10^{-18}$	$5.5 \cdot 10^{-17}$	Xue et al. (2011)

[a] 50% of the field has higher exposure than the quoted number.

[b] Quoted flux in 2-10 keV.

2.1.1 X-ray data

The field has been observed with XMM-Newton in the time period between April 2000 and December 2002, with a total raw exposure of 1.30 Ms for the detectors on board the satellite (EPIC MOS and EPIC pn). The field is centered at $+10^{h}52^{m}43^{s}$, $+57^{\circ}28'48''$ and has a radius of $15'$, thus covering an area of $\sim 0.2\,\mathrm{deg}^2$. The limiting flux is $F_{0.5-2.0\,\mathrm{keV}} = 1.9 \cdot 10^{-16}\mathrm{erg\,cm^{-2}s^{-1}}$ in the soft band, $F_{2.0-10.0\,\mathrm{keV}} = 9 \cdot 10^{-16}\mathrm{erg\,cm^{-2}s^{-1}}$, in the hard band and in the ultra hard band $F_{5.0-10.0\,\mathrm{keV}} = 1.8 \cdot 10^{-15}\mathrm{erg\,cm^{-2}s^{-1}}$.

The depth of the X-ray observations in combination with the size of the area place the field close to the Extended CDFS (ECDFS), between the very deep pencil-beam fields (e.g. CDFN and CDFS) and the more extended but shallower fields (e.g. XMM-COSMOS). In particular, the X-ray observations in the Lockman Hole Field reach one order of magnitude fainter flux limit than XMM-COSMOS in all detection bands (Table 2.1). The catalog of the X-ray detected sources is presented in Brunner et al. (2008) and contains a total of 409 sources with likelihood of being real detections greater than 10 (3.9σ). The catalog of optical counterparts of the X-ray sources is presented in Rovilos et al. (2011) (see also §2.2.3). In summary, the counterparts were assigned on the basis of a Likelihood Ratio (LR) technique applied to our optical and near-infrared catalogs, which reach a depth of $R_{\mathrm{c,lim}}$ = 26.0 mag and $[3.6\mu\mathrm{m}]_{lim}$ = 24.6 mag.

2.1.2 Ultra-violet data

In the far ultra-violet (FUV) and the near ultra-violet (NUV) the Lockman Hole area was observed by the Galaxy Evolution Explorer (GALEX). Here, we consider the magnitude in $3''$ diameter aperture available in the General Release 4/5 (GR4/5)[1]. To correct the aperture photometry to total, we followed the recipe of Morrissey et al. (2007). The authors presented the growth curve analysis using as targets white dwarfs and we use the same correction factors through out the whole catalog. The Lockman Hole is part of the GALEX Deep Imaging Survey and the third quartile of the magnitude distribution reaches $\mathrm{FUV_{lim}} = 24.5\,\mathrm{mag}$ and $\mathrm{NUV_{lim}} = 24.5\,\mathrm{mag}$.

2.1.3 Optical data

In the optical wavelengths, our dataset consists of a compilation of observations from LBT, Subaru and SDSS. Our images exhibit very good seeing[2] of the order of $0.9'' - 1''$, such that we could retrieve almost 100% of the flux for point-like sources within $3''$ aperture (see §2.2.2 for details). Thus, we are using the $3''$ aperture magnitudes without applying any corrections to total. In Fig. 2.3, we present the completeness analysis for all optical filters. We plot the ratio of detected over true number of simulated point-like sources, versus magnitude. Details on the area observed, used filters (λ_{mean} and FWHM), the total exposure time, the PSF full width half maximum (FWHM), the 50% detection magnitude limit for point-like sources (5σ, AB) and the AB to Vega correction factor for each filter, are presented in Table 2.2.

Large Binocular Telescope

We observed the Lockman Hole in the period between February 2007 and March 2009 in 5 bands (U, B, V, Y, z′) using the Large Binocular Telescope (LBT) covering an area of about $0.25\,\mathrm{deg}^2$. The reduction of the data and the number counts of the very deep observations in the U, B and V filters have been published in Rovilos et al. (2009). The two telescopes of LBT

---

[1] http://galex.stsci.edu/GR4/

[2] Calculated from the mean PSF FWHM of 30 stars.

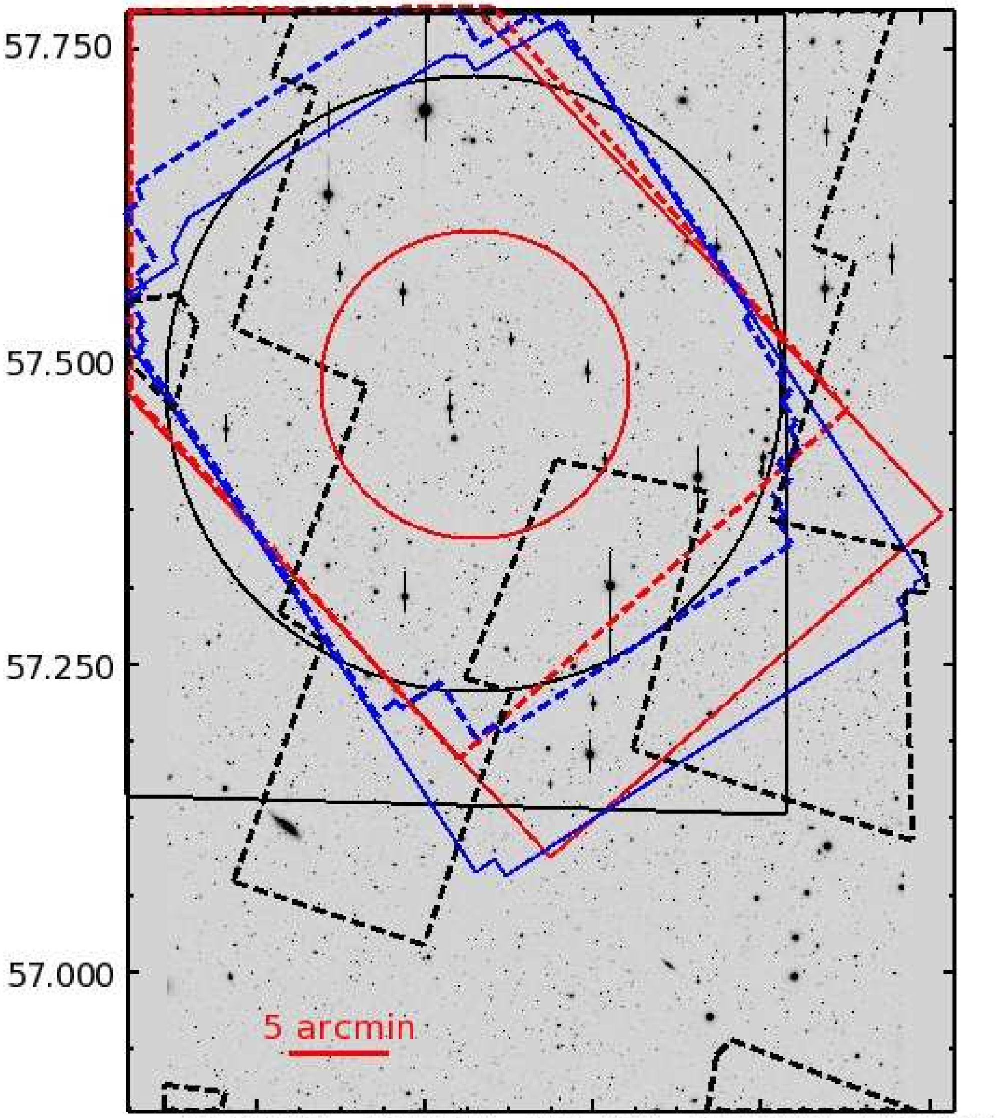


Figure 2.2: The area coverage of the Lockman Hole Deep Field. The image denotes the area observed in the R_c, I_c, z' filters with Subaru. The marked regions represent: U and B LBT filters (blue solid line), V, Y and z′ LBT filters (blue dashed line), 3.6μm and 5.8μm IRAC (red solid line), 4.5μm and 8μm IRAC (red dashed line), J and K UKIRT filters (black solid line). Even though the whole area is covered by SDSS, only the area enclosed in the black dashed outline has good photometry (flag=3 in the SDSS catalog). The black circle represents the area targeted by XMM-Newton and the red circle represents roughly the area targeted by the Hubble Space Telescope.

Table 2.2: Observational facts for the ground-based observations

	Large Binocular Telescope					Subaru		
	U	B	V	z′	Y	Rc	Ic	z′
Area (deg^2)	0.26	0.25	0.19	0.20	0.19	0.53	0.53	0.53
Filter λ_{eff}(Å)	3573	4249	5405	9050	9880	6518	7957	9064
Filter FWHM (Å)	540	916	845	1053	465	1167	1381	1154
Exposure (sec)	49680	19972	9540	14400	10980	3920	6235	10400
PSF FWHM (″)	1.06	0.90	0.95	1.06	0.60	0.90	0.98	0.96
Lim. Mag.[a]	26.7	27.0	26.7	24.2	23.5	26.1	25.5	24.8
AB correction[b]	0.964	-0.046	-0.005	0.528	0.558	0.207	0.436	0.521

[a] 5σ, AB system

[b] As computed by LePhare: M_{AB}=M_{Vega}+AB_{corr}

have slightly different cameras, one optimized for bluer and the other optimized for redder bands. In Rovilos et al. (2009) the published V band is the combination of two separate images ($V_{\rm blue}$, $V_{\rm red}$). However, here we are using only the $V_{\rm red}$ image, as the differences in the two filter curves and the low quality of the $V_{\rm blue}$ image, make the stacking not an option for the purposes of this study. Nevertheless, the number counts presented in Rovilos et al. (2009) are still accurate. In addition to U, B and V photometry, in this work we also include the shallower observations in the Y and z′ filters which were reduced later in a similar manner.

Subaru

Complementary to our own photometry, we made use of data from the Institute for Astronomy (Hawaii) Deep Survey. Observations with the Subaru telescope between November 2001 and April 2002 provided imaging in an area of 0.53 deg^2. The observations were carried out in the Rc, Ic and z′ filters. Details about the survey and the observations on the field can be found in Barris et al. (2004), while details on the analysis of the data can be found in §2.2.2.

Sloan Digital Sky Survey - SDSS

Unfortunately SDSS does not cover the Lockman Hole field with uniform photometric quality. However, we include SDSS (DR7, Abazajian et al. 2009) u′, g′, r′, i′, z′ photometry (fiber magnitude) when available, if $r'_{\rm AB} < 22$ mag and photometric quality flag = 3, provided in the SDSS catalog (Fig. 2.2 black dashed line). The first reason to use the SDSS catalog is that the observations from LBT and Subaru are very deep and the brightest sources ($R_c \approx 18$ mag) are saturated. Thus the shallower SDSS data can provide a solution to saturation problems. For example, in the R_c filter, 677 sources are flagged as saturated and photometry is not available. However, for 170 of these sources SDSS photometry is available and can be used to recover the SEDs and thus the photometric redshifts.

SDSS photometry is especially important for X-ray detected sources, for which we want to trace optical variability. This intrinsic property of AGN dominated systems, can affect the computation of photometric redshifts, when the photometry is obtained non-simultaneously and through multi-epoch observations. As we explain further in §2.2.3, we use the z′ band to

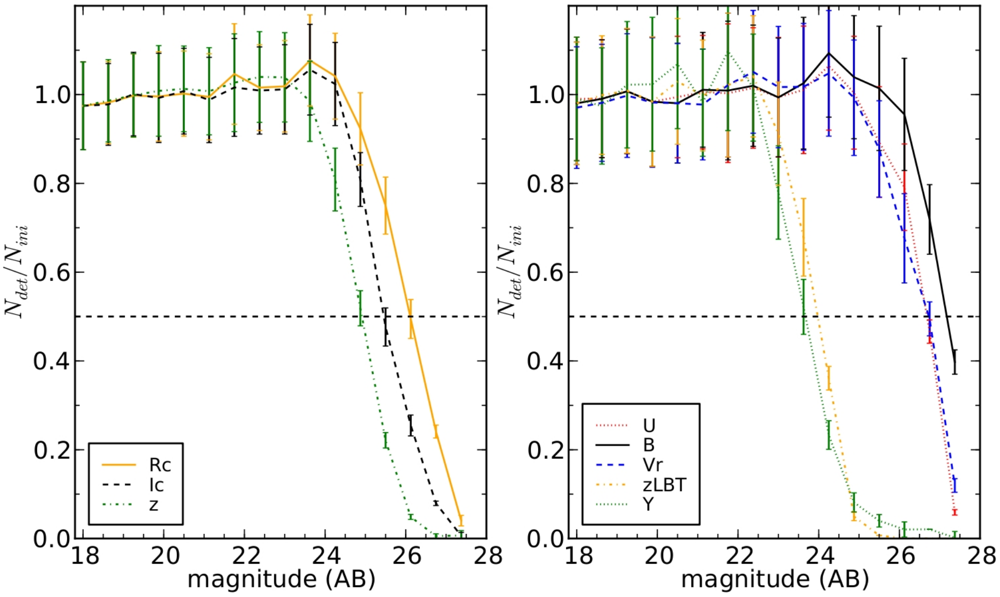


Figure 2.3: Completeness curves for the Subaru (left panel) and LBT (right panel) images, computed from simulated sources. The horizontal dashed line marks the 50% completeness limit.

detect the variability of a source, since in this filter we have observations from three telescopes at different epochs.

2.1.4 Near and Mid infrared data

UKIDSS data

The near-infrared view of our sources is provided by the data release 7 of the UKIRT Infrared Deep Sky Survey. The UKIDSS project is defined in Lawrence et al. (2007). UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali et al. 2007). The photometric system is described in Hewett et al. (2006) and the photometric calibration is explained in Hodgkin et al. (2009). The Lockman Hole is part of the Deep Extragalactic Survey to be observed in the J, H and K bands. Up to now, only the J and K bands are available with limiting magnitudes of $J_{lim} = 23.4$ mag and $K_{lim} = 22.9$ mag (5σ, point source) respectively. We use the 2.8″ aperture magnitude provided in the catalog which is already corrected to total for point-like sources. The aperture corrections were determined from the growth curve analysis of bright stars as described in Hodgkin et al. (2009).

Spitzer data

For the near to mid infrared wavelengths we are using observations at 3.6μm, 4.5μm, 5.8μm, 8.0μm from the Infrared Array Camera (IRAC). Although the IRAC images are the same as the images used in Pérez-González et al. (2008), we created our own catalog. As explained in Rovilos et al. (2011), we extracted the IRAC photometry using SExtractor in dual mode using the 3.6μm image to detect the sources in all other IRAC filters. Following Surace et al. (2005), we used the 2.8″ diameter aperture magnitude and applied the same corrections calculated for point-like sources. The aperture corrections have been determined performing a growth curve analysis of a composite PSF consisting of 10-20 stars.

The 50% efficiency limiting magnitude of the IRAC observations is 24.6 mag for the 3.6μm band (2σ). We verified our photometry against the publicly available SWIRE catalog[3] in the Lockman Hole, which is shallower and complementary to our observations. Using approximately 2400 sources in the overlapping 0.25deg^2 we find that the orthogonal distance regression gives slope of 0.995 ± 0.002, intercept 0.19 ± 0.05 and correlation length 0.73, in agreement with what is found by Pérez-González et al. (2008). The offset between our photometry and the SWIRE photometry is explainable in view of the different, continuously improved pipeline used by the IRAC team in reducing the data.

2.2 Catalog Assembly

The photometric catalog that we release in electronic version[4], covers the area observed by Subaru (Fig. 2.2) which is the widest in terms of area ($\sim 0.5\,\mathrm{deg}^2$). Only 42% of the Subaru observations are covered by the XMM-Newton observations, and for the remaining field we do not have information on the presence of an AGN. Thus, we decided to flag the sources outside the XMM-Newton area.

In this section we describe the re-processing and the photometric reduction of the LBT and Subaru images and the compilation of the catalogs.

2.2.1 Image re-processing

We re-processed the LBT (U, B, V, Y, z′) and Subaru (R$_c$, I$_c$, z′) images to make them as uniform as possible. First, in order to bring all the optical images to a common grid, we registered the LBT and Subaru images to the SWIRE astrometry using the 'Geomap' and 'Geotran' tasks in IRAF with a 6th order polynomial transformation to correct for any residual distortions. The final relative astrometric accuracy is of the order of 0.2″, equal to the pixel scale of the images (Rovilos et al., 2011).

Furthermore, in order to retrieve homogeneous photometry we convolved all the images to the largest seeing (PSF FWHM=1.06″). Assuming that the PSF of an image is approximated well by a gaussian with standard deviation σ, then by definition FWHM=2.35σ. Furthermore, since the convolution of two gaussian functions with standard deviations σ_1 and σ_2 is also gaussian with standard deviation $\sigma_{conv}^2 = \sigma_1^2 + \sigma_2^2$, the correction factor for each image is given by the formula:

$$\sigma_{corr} = \sqrt{\sigma_{max}^2 - \sigma_i^2} \tag{2.1}$$

where σ_{max} the standard deviation of PSF of the image with the largest seeing (images U and z$_{\mathrm{LBT}}$), and σ_i the standard deviation of the PSF of each other image. We performed this procedure using the appropriate Gaussian kernel for each image in the task 'Gauss' in IRAF.

2.2.2 Catalog compilation

In the next subsections, we describe the procedure we followed to obtain the optical photometry and as well as the combination of the catalogs at all wavelengths. Additionally, we describe

---

[3] http://irsa.ipac.caltech.edu/data/SPITZER/SWIRE/

[4] and will be periodically updated at http://www.rzg.mpg.de/~sotiriaf/surveys/LH/

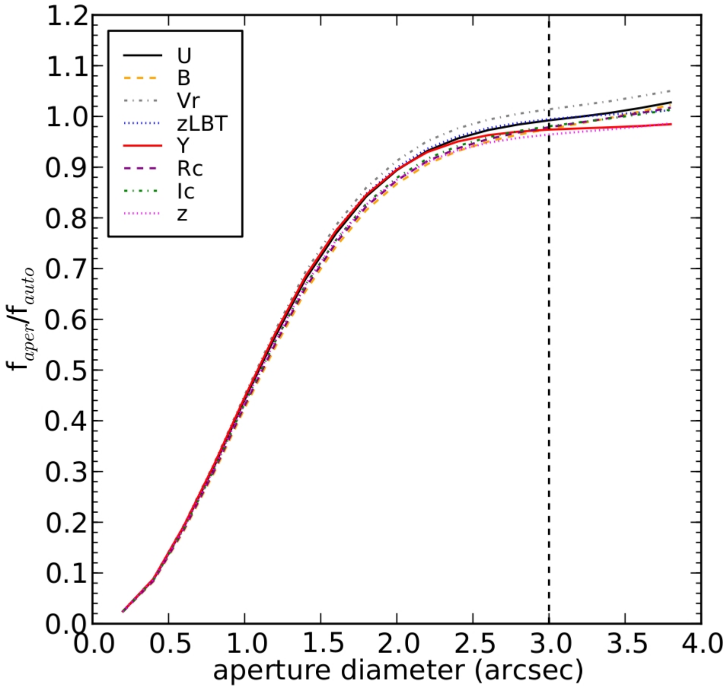


Figure 2.4: Mean growth curves for simulated point sources on the LBT and Subaru images. At least 96% of the total flux is recovered at 3″ aperture diameter.

the subsample of sources for which spectroscopic redshifts are available. The compilation and assembling of the final catalogs is performed using the publicly available codes for data mining in astrophysics, STILTS[5] (Taylor, 2006) and TOPCAT[6] (Taylor, 2005).

Optical Catalog

For all optical images, we calculate the 3″ diameter aperture magnitude using SExtractor (Bertin and Arnouts, 1996) in dual mode, using as reference for the source detection the R_c, z′, B images. The dual mode approach, guarantees that the photometry is measured on every image in an aperture centered at the same pixel position as the detection image.

If two or more sources are close in comparison to the angular resolution of the image, quantified by the PSF FWHM, then the detection algorithm might mistakenly assign one position to two partially overlapping sources. This is commonly referred to as *blending*. In such a case, not only the position of the source is wrong but also the measured photometry is very likely that will be affected. As seen Fig. 2.5 (right hand side), the red ellipses mark the detected position of sources when no filter is used. The objects marked with the numbers 1-3 are affected by blending issues, since in all cases, both the separation among the sources was not successful, but also it is clear that the photometry assigned to the detected object is a mixture of the light coming from the two, not separated sources (particularly true for case 1). To reduce such issues, we are using during the detection the 'Mexican Hat' filter, recommended for crowded fields. With this filter, the algorithm creates a light profile for each source and separates them

---

[5] http://www.starlink.ac.uk/stilts/

[6] http://www.starlink.ac.uk/topcat/

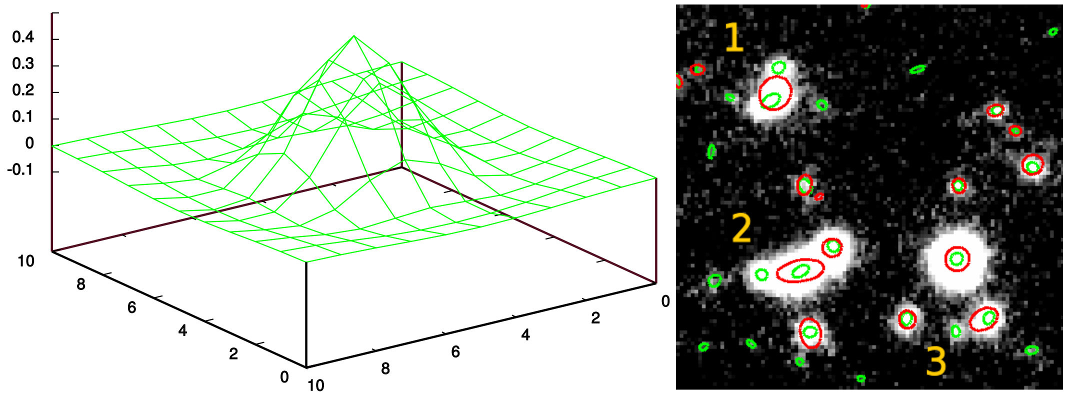


Figure 2.5: Left: The Mexican Hat filter, used during the source detection process to ensure proper deblending of close by sources. Right: Showcase of successful deblending (numbers 1-3). Red ellipses mark the position of source detections when no filter is applied, the green ellipses are the new positions after the Mexican Hat filter is applied.

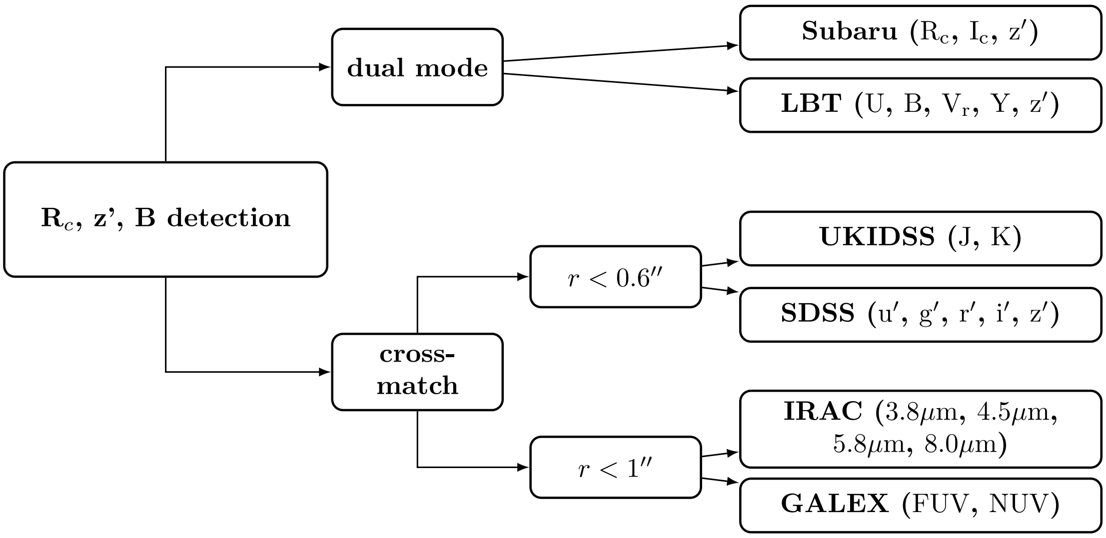


Figure 2.6: Schematic diagram of the merging procedure of the various catalogs. Using as base images the R$_c$, z$'$, B images, we calculate in dual mode the photometry from the LBT and Subaru images. We perform positional matching with the independently obtained photometric catalogs from UKDISS, SDSS, IRAC, and GALEX. The X-ray catalog is then matched to the multi-wavelength optical catalog using Likelihood Ratio matching as described in Rovilos et al. (2011). A detailed description is given in §2.2.2.

successfully assigning the correct positions determined from the peak of each light profile. The green ellipses show the successful separation of the sources in all 3 cases shown here.

In order to determine the flux lost when using a fixed aperture ($3''$), we include on the images, at random positions, simulated point sources with a PSF of $1.06''$ using the task 'mkobjects' in IRAF. In Fig. 2.4 we present the growth curve for all the images. The flux lost is of the order of 0.05 in all cases. This is of the order of our accuracy and we choose not to add any corrections in the optical photometry.

The final photometric catalog is obtained by merging the detection on three images and at two detection thresholds, 5σ and 3σ. The detection images are:

- Priority 1 – the Subaru Rc image. It is the best image in terms of seeing ($0.9''$), depth ($R_{c,lim} = 26.1$ mag), and area coverage ($0.53\,deg^2$). We retrieve 160633 sources at the 5σ detection level and 257352 sources at the 3σ level.
- Priority 2 – the Subaru z′ image. There are sources detected in the z′ band which fall below the detection limit of the Rc band, due to their intrinsic SED. These objects could be high redshift galaxies and/or galaxies that have large amounts of dust. We choose to perform the detection on the Subaru over the LBT z′ image because it covers a larger area, it is deeper and it has better seeing. We retrieve 127362 sources at the 5σ detection level and 250102 sources at the 3σ level.
- Priority 3 – the LBT B image. With LBT the field has been observed with a sequence of short exposures which allowed to reach the same depth of Subaru but with less saturated sources. The number of sources we retrieve is 68107 at 5σ, while at 3σ we retrieve 105658 sources.

In order to create the optical catalogs, we keep the entire Rc catalog and include from the z′ - and B - based catalogs, sources that are not present in the Rc catalog within $0.5''$ from the Rc - sources. This is done for the 5σ and 3σ detection thresholds separately.

As a consistency check of the photometric calibration and as cross calibration between different bands/telescopes, we compared theoretical colors of stars with the colors of star candidates from our catalog (Fig. 2.7). In this context, stars are those bright sources ($R_c < 19$) that SExtractor defines point-like ($R_{class} > 0.98$) by comparison with the PSF of the image. A more refined star/galaxy separation, on the basis of the SED fitting will be performed later on, in the final compilation of the catalog (see section 3.2.3).

After the comparison of the stellar photometry with the stellar tracks, a mean correction of the order of $\sim$0.1 mag is applied to the photometry with exception the Y band. As seen in Fig. 2.7, the Y band required a zero point offset of ~ 0.8. This was expected, as the zero point was only approximate, since no standard stars were observed during the Y band observations.

Combining Optical with Other Wavelengths

In order to merge our optical catalogs with the GALEX, SDSS, UKIDSS and IRAC catalogs, we proceed as follows (Fig. 2.6).

We associated all optical sources with the closest GALEX source within $1''$. As the PSF of the GALEX images is much larger ($5''$), a larger search radius should be adopted. However, as the publicly available GALEX catalog does not account for blended sources, a small radius reduces the number of false matches (see also Arnouts et al., 2005, for optical to GALEX associations). We retrieved ~ 3700 and ~ 10000 couterparts in FUV and NUV, respectively.

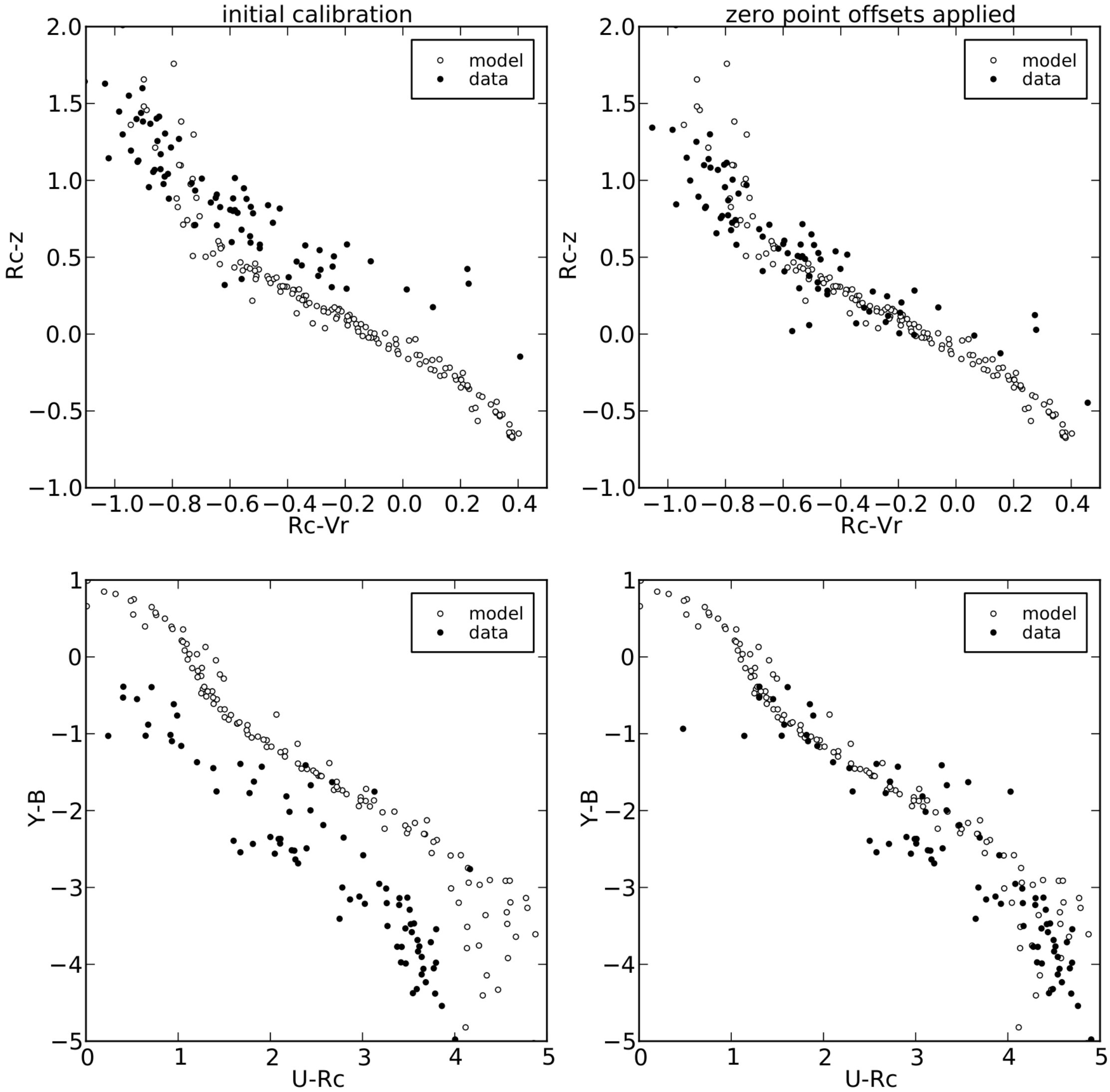


Figure 2.7: We used the star templates of Pickles (1998) to check for systematic offsets in the photometric calibration. Two examples of color-color plots of stars for models (open dots) and data (filled dots), before (left panels) and after (right panel) photometric correction are presented here.

For the Sloan Catalog, the matching radius is reduced to $0.6''$. This corresponds to 3 pixels on our optical images. We keep only the sources for which $r' < 22\,\mathrm{mag}$, similarly to the selection criterion of Oyaizu et al. (2008). We find approximately 5700 matches with the SDSS catalog.

For consistency, as we registered the optical images to the SWIRE coordinates, thus we registered the UKIDSS catalog to our grid. We found a systematic relative offset of about $0.2''$, which has been corrected by changing the astrometry of the UKIDSS catalog using Aladin[7] (Bonnarel et al., 2000). We then matched the optical and UKIDSS catalogs keeping the closest infrared source to the optical coordinates within a $0.6''$ radius to compensate for any further astrometric differences. We found approximately 25000 matches between the optical and the UKIDSS sources. From these matches, 98% lie inside a radius of $0.5''$ from the optical position. Additionally 650 matches are located within a distance between $0.5'' - 0.6''$ which corresponds to a physical distance of 2.5-3 pixels on our images.

Finally, we incorporated the Spitzer catalogs using positional matching within $1''$ from the optical position. We are using a large matching radius for the same reasons as in the case of GALEX. Approximately 26000 sources are matched to an optical source.

In order to create the final merged catalog, we keep all the sources detected in the 5σ level, and the sources detected in the 3σ level which i) have a UKIDSS counterpart and ii) are not present within $0.5''$ from the sources in the 5σ catalog. This choice is a compromise between sources of lower significance and false detections. In the catalog, we provide a flag denoting which catalog each source originates from. In addition, in order to account for problems originated by blending, we flag sources that are not isolated. The final catalog consists of 187611 sources and is described in detail in §2.2.5.

2.2.3 X-ray Sources and their properties

In the final catalog, sources that are identified as counterparts to the X-ray detections are flagged. The association in optical/near-infrared/mid-infrared bands using the maximum likelihood ratio technique is described in detail in (Rovilos et al., 2011). We include in the photometric catalog 388 out of the 409 X-ray detected sources in the field. The 20 sources not present are either too faint to be detected even at the 3σ threshold or are associated with stars. For the sources flagged as X-ray detections we provide additional information that will be of crucial importance for the computation of photometric redshifts. These are the morphology and variability analysis.

Morphology Recently, HST images became available for the central area of the XMM region (covering in total 0.034 deg^2, red circle in Fig. 2.2, PI: Somerville). We used those images to classify the counterparts of X-ray sources in i) point-like and ii) extended, by visual inspection. We also performed morphological classification by visual inspection of the B and Rc bands. The final morphology is assigned based primarily on the HST images, complemented by the ground based images for the sources with no HST coverage. We classified 134 sources as point-like and 140 as extended. From the remaining sources 103 are too faint to be classified and 12 are associated with saturated sources (see Table 2.4).

Variability AGN vary on time scales from days to years. This intrinsic property complicates the computation of photometric redshifts. The photometry is most of the times gathered over many years and a change in flux can be misinterpreted as an emission line. The variability in

---

[7] http://aladin.u-strasbg.fr/aladin.gml

COMBO 17 (Wolf et al., 2004) and in COSMOS (Salvato et al., 2009), was quantified through observations of the same energy band repeated over the years. The comparison of these repeated observations allowed the identification of the variable sources and the correction of the photometry. In the Lockman Hole, we can detect variability but we cannot correct for it. Thus, we limit the analysis in flagging the variable sources. The variability flag will serve as warning for potentially unreliable photometric redshifts.

We use as proxy of the variability the flux variation in the z′ band, for which we have observations from the Subaru telescope, from LBT and from SDSS. The variability is expected to be more significant for bright sources, as for fainter sources the photometric errors are large. Taking into account only bright sources (18 < z < 22), we compute the variability ($Z_{var,k}$, $k = 1, 2, 3$) and the associated error ($\delta Z_{var,k}$) for all three pairs of available z′ filters:

$$Z_{var,k} = z'_i - z'_j \tag{2.2}$$

and

$$\delta Z_{var,k} = \sqrt{(z'_{err,i})^2 + (z'_{err,j})^2} \tag{2.3}$$

Using theoretical magnitudes of galaxies (see §3.1.1 for details), we verified that the expected deviation in the photometry, because of the slightly differing z′ filters, is always less than 0.2 magnitudes for redshifts up to 5.6, for all types of galaxies. Therefore, we flag a source as variable if $|Z_{var,k} \pm \delta Z_{var,k}| > 0.2$ in at least one pair of the observations. We flag 102 sources as varying and 58 sources as non-varying. The remaining 229 sources, are either faint (z > 22) or they are lacking z-band photometry and they are treated as non-varying.

2.2.4 Spectroscopic sample

Compared to other fields, the number of spectroscopic redshift available in Lockman Hole is limited and focused mostly on AGN and infrared galaxies. From all the available catalogs (see Table 2.3 for a complete list), we extract the spectroscopic redshifts with the highest confidence. The quality is provided either by the authors (and in these cases we cannot verify the quality of the redshift estimate) or, when the spectrum is available, it is assessed by us. We define a redshift reliable when more than 1 feature (either in emission and/or in absorption) is present in the spectrum.

We are using 10 redshifts from Schmidt et al. (1998) ($0.245 < z < 2.144$) and 50 redshifts from Lehmann et al. (2001) ($0.074 < z < 4.45$), where the sources were observed as possible counterparts of the X-ray sources detected by ROSAT. Also, during the same observing run normal galaxies where observed as secondary targets. These observations were published in the PhD thesis of I. Lehmann (Universität Potsdam, 2000). In this way we recovered additional 29 reliable spectroscopic redshifts ($0.045 < z < 0.903$).

Furthermore, we include the observations by Zappacosta et al. (2005) ($0.085 < z < 1.13$), who studied the presence of a superstructure in the field. Their catalog contains 48 high quality spectroscopic redshifts, with the superstructure being at redshift ~ 0.8. We include 42 redshifts from the SDSS catalog ($0.00 < z < 3.269$) and 44 redshifts retrieved from NASA/IPAC Extragalactic Database [8] (NED) ($0.073 < z < 3.036$). We also include one spectroscopic redshift from Henry et al. (2010) who studied one of the most distant X-ray selected clusters, with the brightest cluster galaxy located at redshift 1.753.

Our group has also observed sources in the Lockman Hole Field. Focusing on the counterparts of XMM detected sources two observing runs were performed with KECK/DEIMOS

---

[8] The NASA/IPAC Extragalactic Database (NED) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Table 2.3: Spectroscopic redshift reference

Flag	reference	Flag	reference
1	Lehmann et al. (2001)	14	Mainieri et al. (2002)
2	Lehmann, PhD Thesis, Potsdam (2000)	15	Rodighiero et al. (2005)
3	unpublished KECK/DEIMOS (2004-2007)	16	Oyabu et al. (2005)
4	unpublished KECK/DEIMOS (2010)	17	Mateos et al. (2005)
5	SDSS DR2	18	Hasinger et al. (1998)
6	Zappacosta et al. (2005)	19	Ishisaki et al. (2001)
7	Chapman et al. (2005)	20	Smail et al. (2004)
8	Schmidt et al. (1998)	21	Ciliegi et al. (2003)
9	Swinbank et al. (2004)	22	Hashimoto et al. (2005)
10	Chapman et al. (2004)	23	Ivison et al. (2005)
11	Hainline et al. (2009)	24	Fadda et al. (2002)
12	Barris et al. (2004)	25	Huang et al. (in preparation)
13	Stevens et al. (2003)	26	Afonso et al. (2011)

Table 2.4: Morphology's flag description

Description	Flag	Number of Sources		Final
		HST	Ground based	morphology
Too faint / unresolved source	-2	44	86	103
Photometry blended	-1	5	10	12
Extended source	0	20	151	140
Point-like source	1	23	123	134

between 2004 and 2007 (38 high quality redshift, $0.029 < z < 3.408$) and in 2010 (20 sources, $0.353 < z < 1.302$). Furthermore, spectroscopic follow up of the SWIRE field provided 321 high quality spectroscopic redshifts (Huang et al. in preparation) ($0.018 < z < 3.471$).

In total, we have 602 high quality spectroscopic redshifts, out of which 487 are redshifts of normal galaxies with median redshift of $z = 0.42$ ($0.000 < z < 3.471$) with 253 galaxies being inside the XMM area. Furthermore, 115 correspond to X-ray detected sources with median redshift of $z = 0.79$ ($0.024 < z < 4.45$).

2.2.5 Description of the Photometric Catalog

In the following we give a brief description of each column in the catalog. We adopt the value "-99" for null fields inside the catalog.

- ID – unique identification number for the final catalog.
- Optical coordinates (J2000) from the detection image – in addition we provide the coordinates of the counterparts in the GALEX, UKIDSS, SDSS, and IRAC catalogs.

Table 2.5: Photometry's flag description

Flag	Description
-99	99 value produced by SExtractor either for the magnitude or the error, or source outside of the field
-5	flag in the UKIDSS catalog marking noise in the JHK bands
-4	saturation or incomplete/corrupted data produced by SExtractor
-3	source inside a stripe, as marked by optical inspection of the images
-2	magnitude error was negative
-1	magnitude error was greater than 1
0	everything is OK
1	FWHM in the detection band is zero
2	source inside wings of stars, potentially fake
3	magnitude greater than the detection limit

- AB Magnitudes and errors – aperture photometry corrected to total (only for GALEX and IRAC photometry) for point-like sources and associated error for the filters: FUV, NUV, U, B, V, z'_{LBT}, Y, R_c, I_c, z'_{Subaru}, u', g', r', i', z', J, K, 3.6μm, 4.5μm, 5.8μm, 8μm.

- Detection flags – the meaning of the flags is: 1 - Rc detection (5σ), 2 - z' detection (5σ), 3 - B detection (5σ), 4 - Rc detection (3σ), 3 - z' detection (3σ), 6 - B detection (3σ).

- Photometry flag – complementary to the flag provided by SExtractor indicating saturation, we masked problematic regions on the images which include bad pixels and problematic areas close to stars, using 'Weight Watcher' (Marmo and Bertin, 2008). In Table 2.5 we describe the flag values which we provide for each band. For more details, refer to the associated description file of the catalog.

- Neighbor flag – sources that have a neighbor within 1.5$''$ carry a flag 1, while sources without close by neighbors carry a flag 0.

- Star/galaxy classification – the classification provided by SExtractor (1 = star, 0 = galaxy), measured on the corresponding detection image.

- Spectroscopic redshift and reference – we use only the high quality spectroscopic redshifts from the original catalogs according to Table 2.3.

- Variability – sources with $|Z_{var,k} \pm \delta Z_{var,k}| > 0.2$ carry a flag 1, while the rest carry a flag 0.

- X-ray detection – sources detected in the X-rays are flagged as 1, sources inside the XMM area without X-ray detection are flagged as 0, while sources outside the XMM area are flagged as -99.

- X-ray ID – the sources identified as X-ray counterparts will have the corresponding X-ray ID number from the XMM catalog (Brunner et al., 2008). The non X-ray detected sources have a value of XID=-99.

- Morphology – as discussed in detail in §2.2.3, this column gives the merged morphology classification information from the HST and ground based images, provided only for the counterparts of the X-ray sources. The flags used and their meaning are presented in detail in Table 2.4.

2.3 Conclusions

The Lockman Hole Deep Field, is one of the deepest X-ray fields observed with XMM-Newton. The particullar characteristic of low galactic hydrogen column density makes this field ideal for extragalactic observations.

We produced an extensive photometric catalog consisting of 21 broadband filters for sources detected in the Lockman Hole area. This is the first public catalog containing homogeneously reduced and cross-calibrated deep multiwavelength photometry for this area. It contains 187611 objects out of which 388 sources are associated with X-ray sources. The 50% detection limits (5σ) for the photometry are $R_c = 26.1$ mag, $z' = 24.8$ mag and $B = 27.0$ mag.

This catalog can be used in further multiwavelength studies of galaxy populations in the Lockman Hole. Populations which benefit from the low galactic hydrogen column density along the line of sight, can be X-ray, radio and infrared selected sources.

A suggested future improvement of this work is the inclusion of intermediate and narrow band photometry, in order to unveil more details of the galaxy SEDs. We believe that the catalog presented here, and updated versions of it, will be combined with future observations from the majority of the planned telescopes to deliver an even more complete view of the galaxy population and motivate further interesting studies.

CHAPTER 3

Lockman Hole Deep Field – Photometric Redshift

Multiwavelength surveys provide the most successful observational strategy towards the understanding of galaxies. The multiwavelength coverage provides the opportunity to compute photometric redshifts and to study the Spectral Energy Distributions (SED) for a large number of galaxies and simultaneously study the intrinsic properties of the sources such as luminosities and determine the evolution with redshift of other fundamental parameters such as stellar masses, star formation rates, etc. Furthermore, distance measurements enable the use of galaxies as cosmological probes, for example through their clustering properties.

The distance measurement for galaxies is based on the determination of the redshift of the source (defined in §1.1), which is specified very accurately through spectroscopy. However, for large samples of sources either in deep pencil beam fields, or shallower but more extended fields, the most efficient way to compute the distance is via photometric redshifts (e.g. Budavári et al., 2000), although their accuracy is strongly depending on i) the number and the type of filters (broad-band versus intermediate-band), ii) the redshift range and type of galaxies of interest (passive, versus starforming or active galactic nuclei (AGN)).

For fields where extensive photometric datasets are available the photometric redshift technique has been employed with reliable results. For example in CFHTLS[1] (Ilbert et al., 2006), in the UKIDSS ultra-deep survey (Williams et al., 2009), in AEGIS[2] (Barro et al., 2011, Whitaker et al., 2011), in COSMOS[3] (Ilbert et al., 2009, Salvato et al., 2009, Whitaker et al., 2011), in FDF[4] (Bender et al., 2001), in the CDFN[5] (Barger et al., 2003) and in the CDFS[6] (Wolf et al., 2004, Wuyts et al., 2008, Luo et al., 2010, Cardamone et al., 2010) the photometric redshifts reached an accuracy $\Delta(z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm spec}) < 0.03$, thus allowing (among other applications) a 3D mapping of the Dark Matter (Massey et al., 2007), studies of the evolution of luminosity functions of normal galaxies (e.g. Ilbert et al., 2006, Gabasch et al., 2004, 2006, Caputi et al., 2006) and AGN (e.g. Hasinger et al., 2005, Barger et al., 2005, Ebrero et al., 2009, Aird et al., 2010), the first determination of the high-z (z>3) logN-logS and space density from X-ray selected AGN

The main part of this chapter has been published by Fotopoulou et. al (2012), *The Astrophysical Journal Supplement Series, 198, 1*, entitled *Photometry and photometric redshift catalogs for the Lockman Hole deep field.*

[1] Canadian-France-Hawaii Telescope Legacy Survey

[2] All-wavelength Extended Groth strip International Survey

[3] Cosmic Evolution Survey

[4] FORS Deep Field

[5] Chandra Deep Field North

[6] Chandra Deep Field South

(Brusa et al., 2009, Civano et al., 2011), Compton thick objects (e.g. Fiore et al., 2009, Luo et al., 2011), etc. Moreover, photometric redshifts are also used for the study of groups and clusters of galaxies (Giodini et al., 2009, Tanaka et al., 2010, Papovich et al., 2010, Henry et al., 2010, Geach et al., 2011) and much more.

Yet, not all deep and wide fields have photometric redshifts of comparable accuracy, and above all, very few have tuned photometric redshifts for AGN. Having the same photometric redshift accuracy and the same treatment for AGN, the surveys could be merged such that the galaxy and AGN samples can uniformly extend from faint (from deep surveys) to bright sources (from the wide surveys). Keeping this is in mind, here we present optically based extensive photometry and photometric redshift catalogs for the Lockman Hole Deep Field.

3.1 Photometric Redshifts

Photometric redshift estimation is an alternative method to the very accurate, but time consuming, spectroscopic redshift determination. Instead of identifying emission lines on the spectrum of the source, we recreate the continuum emission using as many photometric bands as possible. Then, comparing the observed magnitudes to a library which includes a set of representative galaxy templates created for selected values of redshift and intrinsic absorption we find the template that describes best our data set using the least squares method.

In this section we describe the configuration of the photometric redshift computation. This includes the SED templates, the extinction and redshift grid, the second order correction applied in the photometry as well as the special treatment of the X-ray sample. We also define the quantities we use later on in the discussion to assess the quality of the results.

3.1.1 LePhare Setup

We compute photometric redshifts for all the sources in the field, using the publicly available code, LePhare[7]. The code performs least squares minimization to retrieve the best fitting template to the photometric data. In order to achieve the optimum result for both normal galaxies and AGN, we treated the samples separately and with different sets of templates and priors. In the following we describe the templates and the priors used in the two cases.

Templates

We decided to adopt the same templates as used by Ilbert et al. (2009), Salvato et al. (2009) in the COSMOS field. Being well tested with a large spectroscopic sample, the templates are now included by default in the distribution of the LePhare. We briefly summarize here their major characteristics.

For the normal galaxies the template set consists of elliptical (templ. no. 1-7), spiral (templ. no. 8-19) and starburst templates (templ. no. 20-31) (Ilbert et al. (2009), Fig. 1). Extinction according to the Small Magellanic Cloud (SMC) law (Prevot et al., 1984) is applied for templates Sb-SB3, while for the templates SB4-SB11 Calzetti (Calzetti et al., 2000) and modified Calzetti laws are applied. No additional extinction is applied for templates redder than Sb. The intrinsic galactic absorption is computed with values E(B-V) = 0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.50. Finally, the templates are calculated at redshifts 0-6 with step $\Delta z = 0.01$, and at redshifts 6-7 with step $\Delta z = 0.2$. Emission lines are added to the templates as this has been proven to give better results even in the case of broadband photometry (Ilbert et al., 2009).

---

[7] http://www.cfht.hawaii.edu/ arnouts/LEPHARE/lephare.html

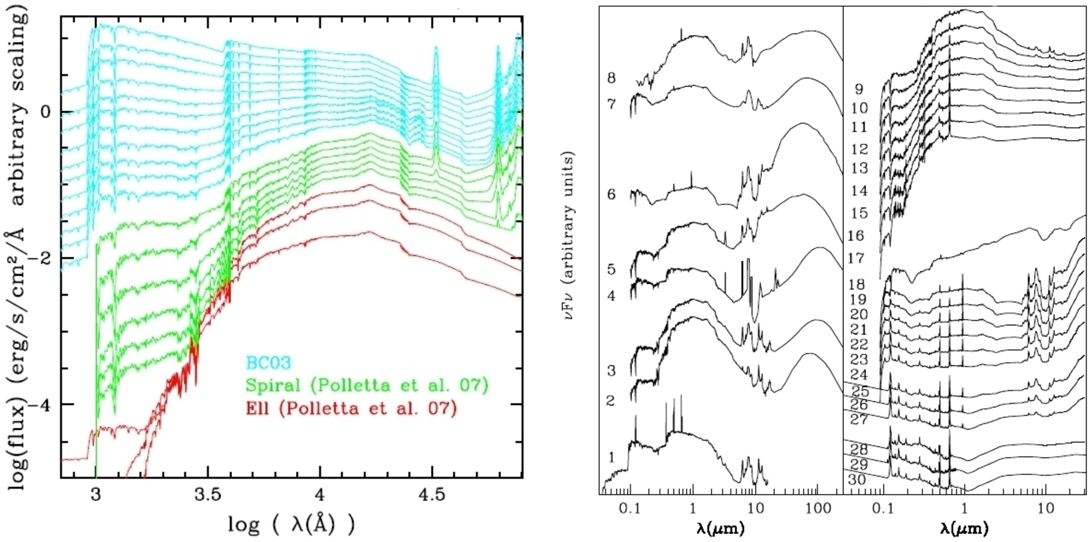


Figure 3.1: SED templates used for this work. Left: Normal galaxies templates (plot adopted from Ilbert et al. (2009)). Right: Compilation of normal galaxy, hybrid, and AGN templates (plot adopted from Salvato et al. (2009)).

For the X-ray detected sample, we are using the same library used in Salvato et al. (2009, 2011) for computing the photometric redshifts of XMM-COSMOS and *Chandra*-COSMOS. The library includes normal galaxies, local AGN, and hybrid templates. The templates forming the library where chosen (and, in the case of the hybrids created) to represent the spectroscopic sample available for the XMM-COSMOS survey, as documented in Salvato et al. (2009). The SEDs of galaxies hosting an AGN and pure galaxy differ mostly in the ultra-violet where the contribution of the accretion disk is expected and in the infrared where the reprocessed AGN radiation by the torus surrounding the central black hole is dominant. The templates of pure AGN dominated sources are taken from the library of Polletta et al. (2007) and are characterized by a typical power-law spectrum. To the pure type 1 AGN, the power-law is extended to the UV beyond the Lyα. The reason for it is that the templates are empirical and thus diminished in the UV by the absorbers along the line of sight. As LePhare accounts by default for this absorption, without the addition of the power-law, the templates would be absorbed twice. More details on the construction of the templates can be found in Salvato et al. (2009). We apply only the SMC extinction law with E(B-V) values of 0.00 - 0.5, in steps of 0.05. The templates are calculated with the same redshift steps as in the case of normal galaxies.

All the sources of our catalog are also fit with stellar SEDs of F-K dwarfs and G-K giant stars (Pickles, 1998), white dwarfs (Bohlin et al., 1995), low mass stars (Chabrier et al., 2000) and sub-dwarfs Bixler et al. (1991).

Residual zero-point offsets

Systematic differences can be found between templates and observed SEDs, due to uncertainties in the templates and second order calibration problems in photometry. The impact of calculating and using these offsets on the accuracy of the result is demonstrated in Ilbert et al. (2006). With the option 'AUTO_ADAPT' in LePhare we compute the average difference in magnitude

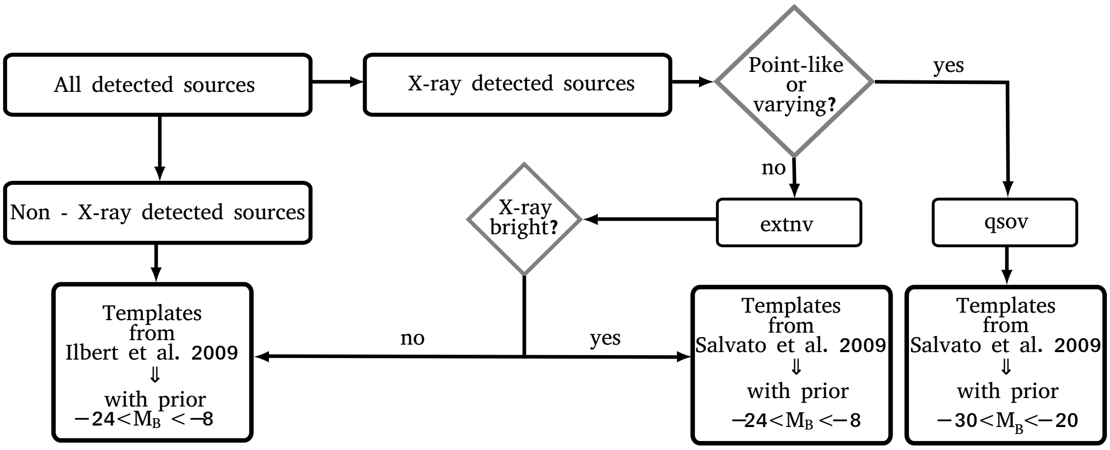


Figure 3.2: The flow chart describes the decision procedure for the optimum combination of templates and priors during the photometric redshift computation. The 'extnv' sample contains the extended and the non-varying sources, while the 'qsov' sample contains the point-like and varying sources. For this work, we adopt $F_{0.5-2\,\mathrm{keV}} = 8 \cdot 10^{-15} \mathrm{erg\,s^{-1}\,cm^{-2}}$ as flux threshold to separate between X-ray bright and faint sources.

between the photometry in a given band and the photometry for the best SED fitting at the fixed spectroscopic redshift of a sample of normal galaxies ($18 < R_c < 24$, ~260 sources). Iteratively, we then search for the best set of corrections that minimize the offsets. Once found, the offset is applied with the option 'APPLY_SYSSHIFT' to the SEDs when computing photometric redshift for the entire catalog. The same offsets are also applied when computing the photometric redshift for the X-ray sources. The offsets are presented in Table 3.1, and they are not included in the photometry presented in the catalog. We do not calculate any offsets for the 5.8μm and 8.0μm filters of IRAC, due to the large uncertainties in the theoretical models in these wavelengths.

We also include an additional factor to the photometric errors in quadrature (optical 0.02 mag, ultra-violet and infrared 0.2 mag). This factor compensates for the underestimated errors provided by SExtractor (Becker et al., 2007, McCracken et al., 2001).

Photometric Redshift Computation

The separate treatment of the sources based on the X-ray detection and emission, allows a pre-selection of templates and luminosity priors that reduces the possible parametric space of the solutions lifting degeneracies between the templates and therefore reducing the number of wrong photometric redshift solutions. In Fig. 3.2 we present the flow chart for the proper separation of the sample, and the template - prior combination we adopt for this work.

First of all, the sample is separated in non - X-ray detected and X-ray detected sources. For the non - X-ray detected sample we use the templates for normal galaxies presented in Ilbert et al. (2009) with priors $-24 < M_B < -8$, which is the typical range of the absolute B magnitude for normal galaxies. For the X-ray detected sources, a further separation is required between point-like or varying sources (QSOV) and extended and non-varying (EXTNV) sources. The QSOV sample contains the AGN dominated sources and for them we are using the AGN templates of Salvato et al. (2009) with the same priors $-30 < M_B < -20$.

Table 3.1: Offsets between theoretical templates and observations

Filter	Offset[a]	Filter	Offset[a]
FUV (GALEX)	-0.165	z′ (Subaru)	-0.004
NUV (GALEX)	-0.356	u′ (SDSS)	0.271
U (LBT)	0.125	g′ (SDSS)	-0.115
B (LBT)	-0.026	r′ (SDSS)	0.006
V (LBT)	-0.038	i′ (SDSS)	0.104
z′ (LBT)	-0.054	z′ (SDSS)	0.083
Y (LBT)	0.141	J (UKIRT)	0.249
R_c (Subaru)	-0.037	K (UKIRT)	0.294
I_c (Subaru)	-0.020	3.6¯m (IRAC)	0.203
		4.5¯m (IRAC)	0.346

[a] Included in the computation of the photometric redshifts, but not in the released catalog.

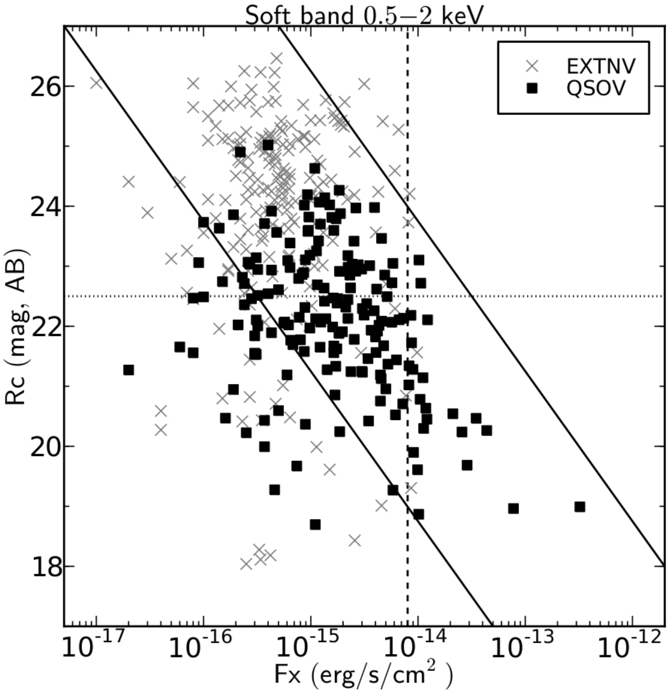


Figure 3.3: Rc magnitude versus soft X-ray flux $(0.5-2\,\mathrm{keV})$. The crosses represent all the EXTNV sample and the squares the QSOV sample. The black solid lines correspond to $\log(f_X/f_R) \pm 1$ where the majority of the AGNs lie. Non-active galaxies will be in the lower left region of the diagram. The vertical dashed line denotes the X-ray threshold we apply to the EXTNV sample for adopting different templates and priors. The horizontal line marks the bright subsample ($R_c < 22.5$).

The EXTNV sample contains moderately AGN dominated sources, starburst and normal galaxies. As demonstrated in Salvato et al. (2011), AGN that are low X-ray emitters, extended and not varying are better fit by normal galaxy templates, even if the X-ray luminosity is above $10^{42}\mathrm{erg}\cdot\mathrm{sec}^{-1}$. For this reason, we treat differently sources that above and below a threshold set at $\mathrm{F}_{0.5-2\,\mathrm{keV}} > 8\cdot 10^{-15}\mathrm{erg\,cm^{-2}\,s^{-1}}$. Above this value, the sources are fitted by AGN templates. The library of normal galaxies defined by Ilbert et al. (2009) is used otherwise. In both cases, the luminosity prior $-24 < M_B < -8$ is adopted. It is important to stress that the threshold was defined using a sample of 700 sources with spectra, belonging to the EXTNV group, in the COSMOS field.

In Fig. 3.3 we plot the optical magnitude ($\mathrm{R_c}$) versus the soft X-ray flux ($0.5-2\,\mathrm{keV}$) for all the X-ray detected sources. The squares indicate sources in the QSOV sample and the crosses indicate sources in the EXTNV sample. The two black solid lines mark the area in which the majority of the AGN-dominated sources is found and they are defined as:

$$log(\frac{f_x}{f_{Rc}}) = \pm 1 \tag{3.1}$$

where,

$$log(\frac{f_x}{f_{Rc}}) = log f_x + \frac{Rc}{2.5} + 5.5 \tag{3.2}$$

see also Maccacaro et al. (1988) Hornschemeier et al. (2003) and Brusa et al. (2005).

The separation of the sources in QSOV and EXTNV is purely based on morphological and variability considerations. Examining sources for which we can distinguish between point-like and extended morphology ($R_c < 24$) we find that 76.4% of the QSOV sources are found in the area between the two solid black lines in Fig. 3.3, while 53.5% of the EXTNV sources are found in the same area, justifying our original assumption that QSOV sources are AGN-dominated sources and need to be treated with appropriate templates and luminosity priors.

Focusing on the bright ($Rc < 22.5$, dotted line), extended, non varying sources (crosses), we see that the above mentioned empirical flux threshold in the X-rays (vertical dashed line at $\mathrm{F}_{0.5-2\,\mathrm{keV}} = 8\cdot 10^{-15}\mathrm{erg\,s^{-1}\,cm^{-2}}$), separates efficiently the AGN dominated systems from starbursts and normal galaxies, which populate the lower left corner of this diagram (Hornschemeier et al., 2003). The positive improvement of the X-ray threshold when computing photometric redshifts for X-ray sources, is only marginal for this work, as the brighter sources are rare and the Lockman Hole is a small field. However, the threshold resulted in a noticeable improvement on the XMM-COSMOS surveys (Salvato et al., 2011) and we decided to adopt the same strategy for consistency.

3.1.2 Outlier-accuracy definition

Outliers Since the photometric redshift computation is a multivariate problem with many degeneracies, is it bound to produce wrong solutions for some objects (see for discussion Richards et al. (2001)).

Comparing the photometric redshift solutions to the spectroscopic redshifts, we quantify the fraction of outliers η as the ratio $100\cdot N_{out}/N_{total}$ where N_{total} is the number of sources with spectroscopic redshift and N_{out} the number of sources with:

$$\frac{|z_{phot} - z_{spec}|}{1 + z_{spec}} > 0.15 \tag{3.3}$$

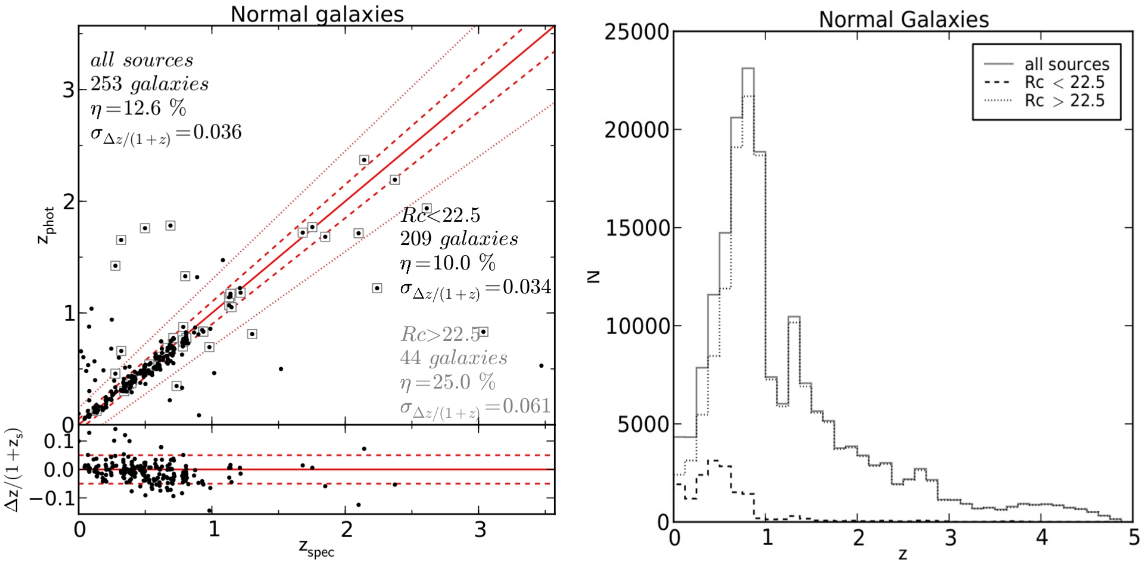


Figure 3.4: Left: Photometric redshift versus spectroscopic redshift for the normal galaxies. With gray squares we denote the faint sources (Rc > 22.5 mag). The solid line is the $z_{\rm phot} = z_{\rm spec}$ relation. The dashed lines are $z_{\rm phot} = 0.05 \pm (1 + z_{\rm spec})$. The dotted lines are $z_{\rm phot} = 0.15 \pm (1 + z_{\rm spec})$. Sources that lie outside the dotted lines are defined as outliers. Right: Redshift distribution for normal galaxies, using spectroscopic redshift when available. Bright galaxies ($R_c < 22.5$ mag) lie mostly at redshift $z < 1$.

Accuracy The accuracy of the photometric redshifts is usually quantified by the direct comparison of the photometric redshift solution against the spectroscopic value. In the literature there are many ways of computing the accuracy, from the pure root mean square (rms) of the $\Delta z = (z_{phot} - z_{spec})/(1 + z_{spec})$, to the rms after 1 or 3 sigma clipping, depending on the authors (Wolf et al., 2004, Mobasher et al., 2004, Margoniner and Wittman, 2008, Dahlen et al., 2010).

Another more robust way to quantify the accuracy of the sample is to consider the median of the deviations from the true (spectroscopic) value. In this way the accuracy accounts also for the outliers. We adopted this approach, described in detail in Hoaglin et al. (1983) and used e.g. by Brammer et al. (2008), Wuyts et al. (2008), Ilbert et al. (2009), Salvato et al. (2009), Cardamone et al. (2010). The Normalized Median Absolute Deviation (NMAD) is defined as:

$$\sigma_{NMAD} = 1.48 \cdot median|\Delta z| \tag{3.4}$$

where $\Delta z = (z_{phot} - z_{spec})/(1 + z_{spec})$. This quantity is expected to remain close to zero for all redshifts.

3.2 Results

3.2.1 Photometric Redshifts of normal galaxies

We retrieve photometric redshift for 184643 non - X-ray detected sources in the Lockman Hole area with 45% of them computed with at least six photometric bands. The remaining 1978

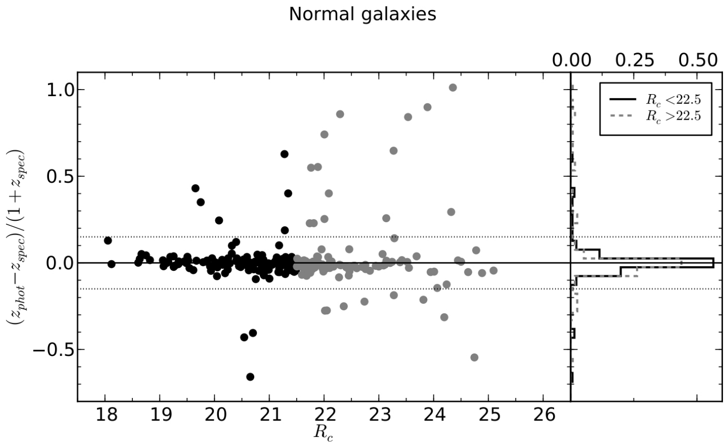


Figure 3.5: The ratio $(z_{\mathrm{phot}} - z_{\mathrm{spec}})/(1 + z_{\mathrm{spec}})$ versus optical magnitude R_c for normal galaxies and corresponding histogram on the right panel. The photometric redshifts of bright sources ($R_c < 22.5$, black filled circles and black solid line) are mostly confined between $|(z_{\mathrm{phot}} - z_{\mathrm{spec}})/(1 + z_{\mathrm{spec}})| < 0.15$ (dotted lines, marking the outlier region), while the outliers with the largest discrepancies are faint sources ($R_c > 22.5$ mag, gray open circles and gray dashed line).

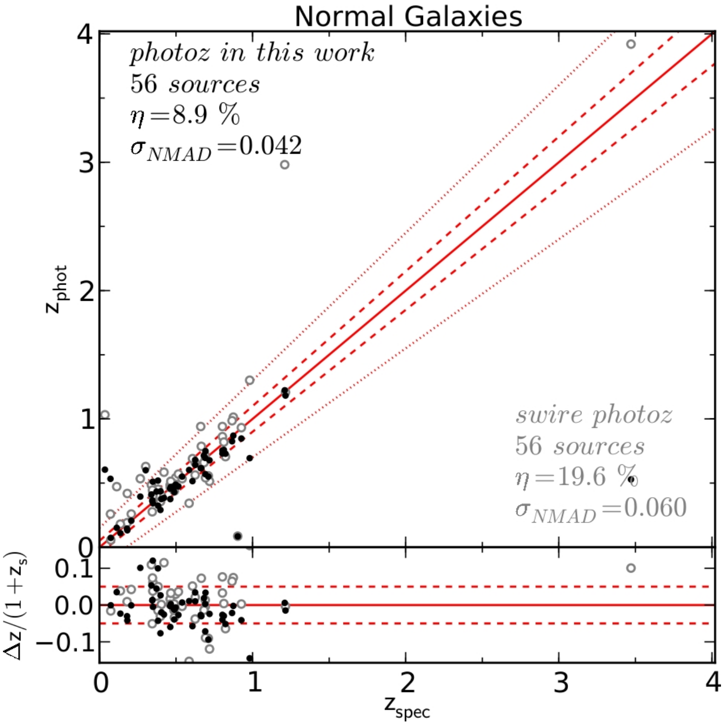


Figure 3.6: Comparing the quality of new photometric redshifts, using sources in common in the present work (black filled dots) and the SWIRE catalog (gray open circles). There is a dramatic decrease in the fraction of outliers and the accuracy is improved.

Table 3.2: Photometric redshift accuracy

	Non X-ray sources			X-ray sources		
	N(zspec)	σ	$\eta(\%)$	N(zspec)	σ	$\eta(\%)$
all sources	253	0.036	12.6	115	0.069	18.3
$R_c < 22.5$	209	0.034	10.0	90	0.069	18.9
$R_c > 22.5$	44	0.061	25.0	25	0.070	16.0
$0 < z < 1$	230	0.034	9.6	67	0.066	16.4
$1.0 < z < 5$	23	0.107	43.5	48	0.078	20.8

sources lack a photometric redshift solution since they have only two photometric bands available. In Fig. 3.4 we compare the photometric versus the spectroscopic redshifts. For the bright subsample (Rc $<$ 22.5 mag) the accuracy is $\sigma_{NMAD} = 0.034$ (eq. 3.4) and the fraction of outliers $\eta = 10.0\%$.

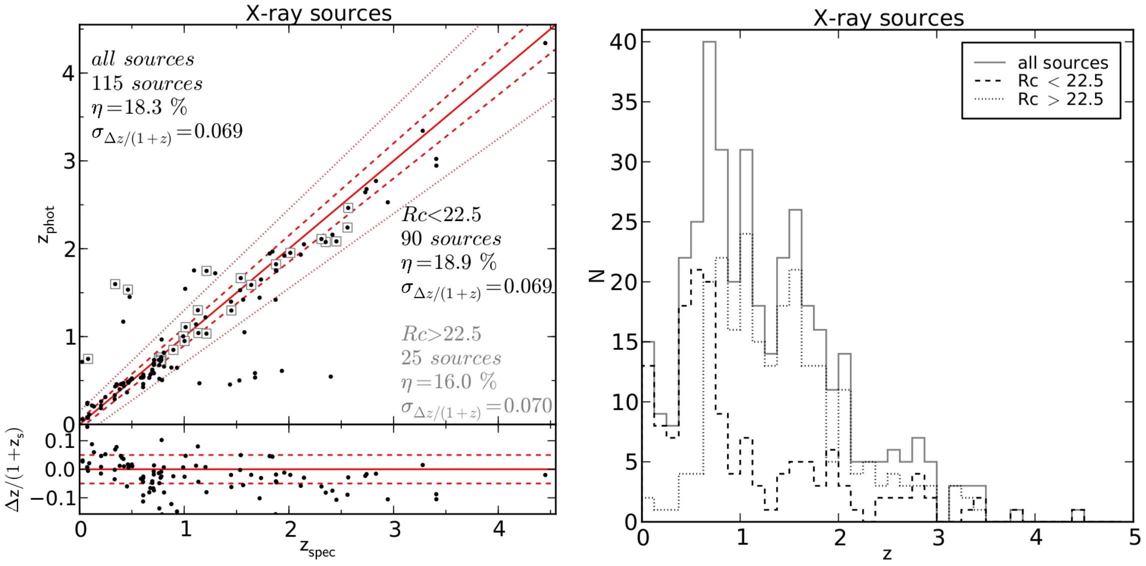
Even though our photometric catalog does not contain medium or narrow band photometry, our result does not suffer from systematic biases. This is also clear in Fig. 3.5, where the ratio $(z_{phot} - z_{spec})/(1 + z_{spec})$ is plotted as a function of optical magnitude. The distribution of the ratio $(z_{phot} - z_{spec})/(1 + z_{spec})$ is centered at zero (histogram, right panel) both for bright ($R_c < 22.5$, black solid line) and faint sources ($R_c > 22.5$, gray dashed line). For the latter, however, the fraction of outliers is higher and consequently the accuracy decreases. This is a well known trend, discussed already by many other authors (e.g., Cardamone et al., 2010, Barro et al., 2011, Ilbert et al., 2009, Salvato et al., 2009); at fainter magnitudes the spectral energy distribution is less tightly constrained, and only by upper limits in some bands, or has large statistical uncertainties associated with the photometry.

In Table 3.2 we summarize the accuracy and the percentage of the outliers for various subsamples. The majority of the sources with spectra available have redshift $z_{\rm spec}$ <1, and therefore we can characterize the quality of our redshift solutions, in a statistically robust way, only for low redshift sources. For $z_{\rm spec}$ <1, the accuracy is $\sigma_{NMAD} = 0.034$ (eq. 3.4) and the fraction of outliers is $\eta = 9.6\%$. Fig. 3.4 (b) shows the redshift distribution of the normal galaxies (solid line) separated in optically bright (Rc $<$ 22.5 mag, dashed line) and optically faint (Rc $>$ 22.5 mag, dotted line), where spectroscopic redshifts substitute photometric redshifts when possible.

In Fig. 3.6 we also compare the photometric redshifts from this work with the previous available photometric redshifts available from (Rowan-Robinson et al., 2008). We consider only the sources in common to the two samples for which spectroscopic redshifts are available (56 sources). Due to the increased number of bands used in this work, we achieve an accuracy improved up to a factor of 1.5 for normal galaxies and 2 times less outliers.

3.2.2 Photometric redshifts of X-ray detected sources

For the counterparts of the X-ray sources, mostly AGN, the accuracy of the bright subsample is $\sigma_{NMAD} = 0.069$, while the fraction of outliers is $\eta = 18.9\%$ (Fig. 3.7). This is similar to the accuracy reached by XMM-COSMOS and *Chandra*-COSMOS when the same bands and depths available for Lockman Hole are considered, without correcting for variability. The impact of variability in our photometric redshift computation is shown in Fig. 3.8. Even thought the



(a) (b)

Figure 3.7: (a) Photometric redshift versus spectroscopic redshift for the X-ray detected sources. The symbols and lines are the same as in Fig. 3.4. (b) Redshift distribution of X-ray sources, using spectroscopic redshift when available.

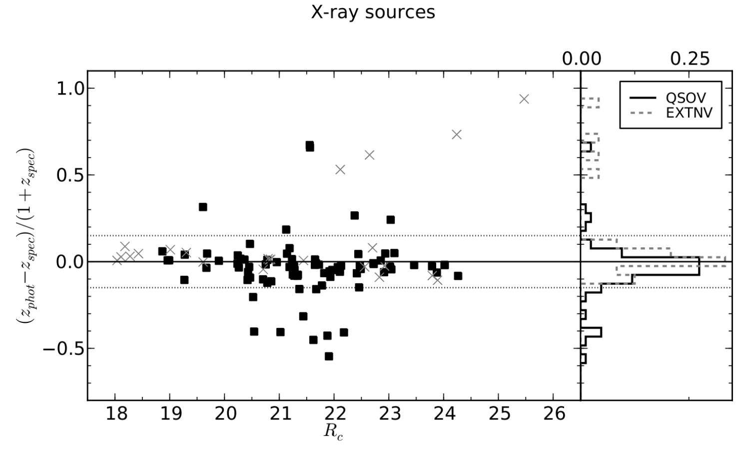


Figure 3.8: The ratio $(z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm spec})$ versus optical magnitude R_c for X-ray sources and corresponding histogram on the right panel. The squares denote the QSOV sample (point-like or varying sources) and the crosses denote the EXTNV sample (extended and not varying sources). The variability is a key factor in the photometric redshift computation, as 11 out of the 16 outliers belonging in the QSOV sample are identified as varying sources. As in the case of the normal galaxies, the largest discrepancies are found for faint sources ($R_c > 22.5$ mag).

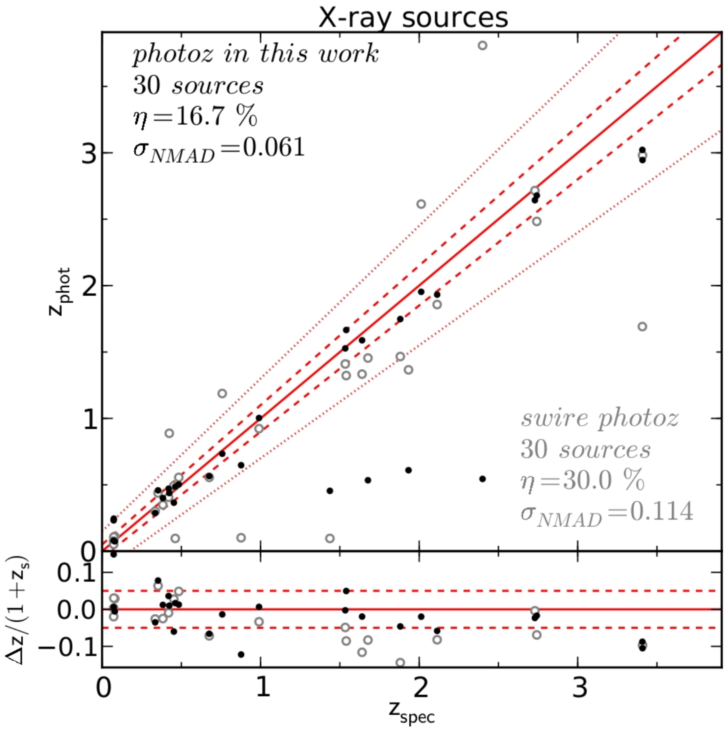


Figure 3.9: Comparison between previous photometric redshifts and results from this work for X-ray detected sources. The symbols are the same as in Fig. 3.6. As in the case of normal galaxies, the outliers are reduced significantly and the accuracy is improved.

Table 3.3: Photometric redshift accuracy for the X-ray detected sample

	QSOV			EXTNV		
	N(zspec)	σ	η(%)	N(zspec)	σ	η(%)
all sources	85	0.071	18.8	30	0.056	16.7
$R_c < 22.5$	70	0.084	21.4	20	0.036	10.0
$R_c > 22.5$	15	0.042	6.7	10	0.126	30.0

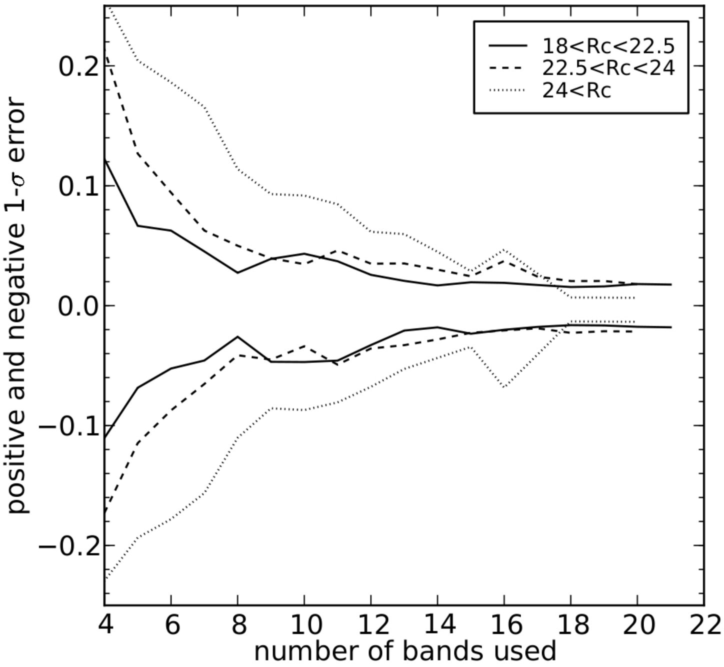


Figure 3.10: Mean error of photometric redshifts versus the number of bands used during the fitting. The bright sample (solid line) has a consistent behavior when six or more number of bands are used. The faint sample (dashed line) approaches the accuracy of the bright sample in cases when eight bands or more are available.

distribution of $(z_{\rm phot} - z_{\rm spec}/(1+z_{spec})$ is peaked around zero, there are some sources in the QSOV subsample (squares) that show quite large deviations. Indeed, 11 out of the 16 outliers are flagged as varying sources indicating once more the care that should be taken in planning photometric observations of AGN. In Table 3.3 we summarize our results separating them on the basis of the classification of the sources (either EXTNV or QSOV) and brightness . In Fig. 3.8 we present the histogram of the redshifts for the X-ray detected sources (solid line) separated in optically bright (Rc $<$ 22.5 mag, dashed line) and optically faint (Rc $>$ 22.5 mag, dotted line).

In Table 3.2 we give the detailed evaluation of our results, in direct comparison with the non X-ray detected sample. As it is expected, the photometric redshifts of the brightest sources are more accurate than the redshifts for the fainter sources. The highest accuracy is reached at $z < 1$ with $\sigma_{NMAD} = 0.066$ and fraction of outliers $\eta = 16.4\%$.

In Fig. 3.9 we compare our results to the photometric redshifts of Rowan-Robinson et al. (2008). Apart from the increased number of used bands an additional reason for the improvement is that, contrary to Rowan-Robinson et al. (2008), our work was tuned to this kind of sources. Rather than adding AGN-dominated templates to the library of normal galaxies, we limited the degeneracies by using only AGN-dominated templates and appropriate priors, thus achieving more accurate results by a factor of 1.8 both in terms of accuracy and outliers.

As already discussed, the accuracy of photometric redshifts correlates with the faintness of the sources. As a consequence, when the source is faint, less photometry is also available. In Fig. 3.10 we plot the median positive and negative $1-\sigma$ errors, defined as $(z_{best68,high} -$

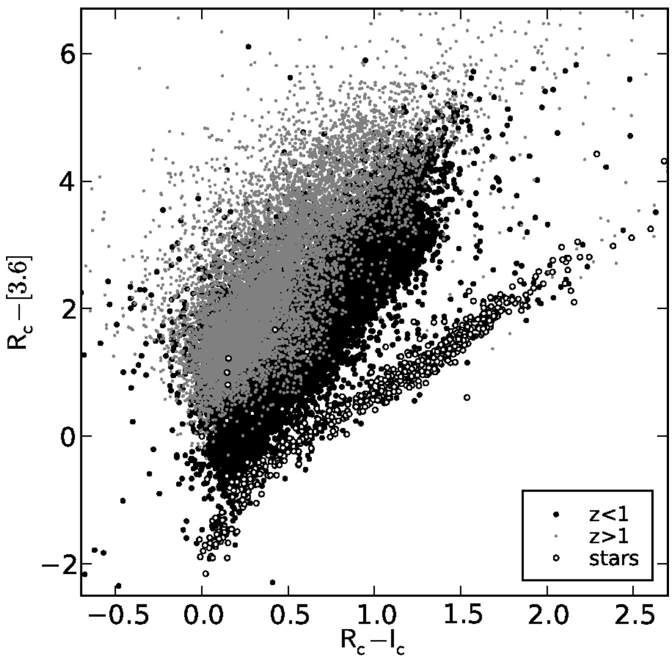


Figure 3.11: Color - color plot demonstrating the star - galaxy separation. Gray dots indicate high redshift galaxies, black circles low redshift galaxies and open circles indicate the stars. Most of the identified stars through the SED fitting, lie in the expected locus for stars.

z_{best}) and ($z_{best} - z_{best68,low}$), per number of bands used in the SED fitting for all sources (non X-ray and X-ray detected sources). As expected, the bright sources show small errors even when using only a few bands, as bright sources have usually small errors associated with their photometry, thus making the least squares fitting more precise. We underline that narrow errors of photometric redshifts does not mean that the solution is the correct one, rather that the probability distribution has a narrow peak around a given value.

3.2.3 Star/galaxy separation

Rather than distinguishing between stars and galaxies, SExtractor separates the objects in point-like and extended. This is performed by comparison of the measured FWHM of an object with the PSF of the image (given as input), as a result the code is unable to distinguish between stars and unresolved galaxies.

To assess the limitations of SExtractor we compare the number of stars defined by the code, with the expected number of stars defined by stellar population synthesis [8] models of galactic stars in the area covered by Lockman Hole. In the case of bright objects ($R_c < 19$) the number of stars detected by SExtractor (90 sources) agrees well with the expected number (103 sources). Inversely, in the case of less bright objects ($20 < R_c < 22$) SExtractor classifies $\sim$ 1500 sources as stars, almost three times more the number predicted by the simulation ($\sim$ 440 stars), as it confuses point-like sources with unresolved ones.

At the same time, SED fitting can misclassify objects as stars especially in the case of ellip-

---

[8] http://model.obs-besancon.fr/

tical galaxies that are lacking infrared photometry. To compensate between the two effects, we flag as stars sources that have been identified as stars both by SExtractor (Classification > 0.95) and by the SED fitting ($2 \cdot \chi^2_{star} < \chi^2_{best}$) using star templates from Pickles (1998), Bohlin et al. (1995), Chabrier et al. (2000), Bixler et al. (1991). With this conservative approach we flag ~ 700 sources as stars , excluding most of the false identifications. Indeed, according to the simulation we expect $\sim$ 800 stars in the Lockman Hole area having the same magnitude distribution as the the sources identified as stars from the previous criterion.

Fig. 3.11 shows a color-color plot of stars (open circles), low redshift (black) and high redshift (gray) galaxies. With the combined criterion of morphology and SED fitting, we retrieve stars mainly in the expected locus according to theoretical templates (see also Ilbert et al. (2009)).

3.2.4 Description of the Photometric Redshift Catalog

In the following we give a description of the photometric redshift catalog, which we provide separately from the photometry catalog. The photometric redshift catalog includes:

- ID – corresponding identification number from the photometric catalog.
- z_{best} – the best fitted solution for the photometric redshift.
- $z_{best68,low}$ – the lowest redshift at 68% significance.
- $z_{best68,high}$ – the highest redshift at 68% significance.
- $z_{best90,low}$ – the lowest redshift at 90% significance.
- $z_{best90,high}$ – the highest redshift at 90% significance.
- χ^2_{best} – the lowest χ^2 value for the best fitted galaxy model.
- PDZ_{best} – the probability that z_{best} is the correct photometric redshift.
- $model_{best}$ – the number corresponding to the model best fitting the SED. 100+($1, \ldots, 31$) from Ilbert et al. (2009); $1, \ldots, 30$ from Salvato et al. (2009).
- $Ext - law_{best}$ – the extinction-law applied for computing z_{best}.
- $E(B-V)_{best}$ – the absorption applied for computing z_{best}.
- $Nband_{best}$ – number of bands used in the SED fitting.

  Similarly for the second photometric redshift solution, when available:

- z_{sec} – the second best solution for the photometric redshift.
- χ^2_{sec} – the corresponding χ^2 value .
- PDZ_{sec} – the probability that the second z_{best} is the correct photometric redshift.
- $model_{sec}$ – the number corresponding to the model best fitting the SED.
- $E(B-V)_{sec}$ – the absorption applied for computing the second z_{best}:
- χ^2_{star} – the lowest χ^2 value for the best fitted star model.
- $Flag_{star}$ – we flag stars (1 = star, 0 = galaxy), as discussed in 3.2.3

3.3 Conclusions

Extensive photometric observations have made possible a panchromatic view of the Universe. Combining observations in a plethora of filters, we are able to reconstruct the continuum emission of a large number of sources. Additionally, through the comparison between theoretical and/or empirical templates of galaxy spectra, we are able to determine physical quantities of the observed sources.

Utilizing the extensive photometric catalog presented in the previous chapter, we produced an equally extensive catalog with photometric redshift information for all the sources. Depending on the nature of the sources (non X-ray and X-ray detected) we used different templates and priors, allowing a final accuracy for the bright subsample ($R_c < 22.5\,\mathrm{mag}$) of $\sigma_{NMAD} = 0.034$ with a fraction of 10% outliers for normal galaxies. Similarly for the X-ray detected sources an accuracy of $\sigma_{NMAD} = 0.069$ is reached with an 18.9% fraction of outliers.

This accuracy and fraction of outliers is in agreement with the estimated accuracy in the COSMOS field, when only broadband filters are used, and no correction for variability is applied. Further improvement can be achieved by the inclusion of intermediate and narrow band photometry and as well as from the increase of the spectroscopic sample.

CHAPTER 4

AGN X-ray Evolution

The X-ray luminosity function (XLF) of AGN and its evolution provides a view of the black hole (BH) growth across the cosmic time. Several studies use the XLF to constrain models of black hole evolution through simulations and semi-analytic models (e.g. Mahmood et al., 2005, Hopkins et al., 2005a) and to investigate the possible galaxy - black hole coevolution (e.g. Hopkins et al., 2007, Marulli et al., 2008, Zheng et al., 2009, Fanidakis et al., 2011). The XLF also is used to constrain the properties of the AGN population for example creating population synthesis models that describe the Cosmic X-ray Background (CXRB) and thus inferring the fraction of Compton Thick AGN (e.g. Gilli et al., 2007, Draper and Ballantyne, 2009). Additionally, the XLF is used to test the still open question of AGN triggering: mergers vs secular processes (Draper and Ballantyne, 2012).

Combining the XLF with luminosity functions in other wavelengths the multivariate changes of the AGN phase are studied (e.g. Han et al., 2012, Hopkins et al., 2005b). For example, the connection of X-ray and infrared radiation from AGN has been studied by means of reconciling the CXRB to the infrared background through the corresponding luminosity functions (Ballantyne and Papovich, 2007).

Early X-ray surveys showed that the local number density of AGN follows a broken power law distribution and it was proposed that the XLF evolves with redshift according to the Pure Luminosity Evolution model (PLE, Maccacaro et al., 1983, 1984), while subsequent studies showed that the evolution stops, or dramatically slows down after a critical redshift value (i.e. Boyle et al., 1994, Page et al., 1996, Jones et al., 1997). Recent works in the Soft X-ray regime ($0.5-2\,\mathrm{keV}$) support a Luminosity Dependent Density Evolution (LDDE) over the simple PLE with the number density of AGN peaking in redshift $z=1-2$ (Miyaji et al., 2000, Hasinger et al., 2005, Ebrero et al., 2009). Similarly in the hard X-ray band ($2-10\,\mathrm{keV}$) some works support LDDE over PLE (Ueda et al., 2003, Ebrero et al., 2009), while other studies of the same energy band also tested simultaneous variations in luminosity and density, namely Independent Luminosity and Density Evolution (ILDE, Yencho et al., 2009) and Luminosity and Density Evolution (LADE, Aird et al., 2010).

According to the unified model (Antonucci, 1993, Urry and Padovani, 1995) a supermassive black hole is found at the center of each AGN, surrounded by an accretion disk and a torus of gas and dust. The current accepted view of the radiation processes mechanism includes X-ray production in the vicinity of the black hole from a population of thermal electrons (see also §1.2). The torus is responsible for the obscuration of the X-rays due to photoelectric absorption and emission in the infrared, the latter being reprocessed optical radiation by the dust. For the luminosity function in the X-ray energy ranges $0.5 - 2\,\mathrm{keV}$ and $2 - 10\,\mathrm{keV}$ a correction factor must be applied depending on the absorption power of the obscuring torus in each source which is either calculated from the spectrum or roughly estimated from the observed flux in at least two X-ray energy bands. An additional correction factor is often applied to account for redshift incompleteness in the sample under investigation.

Studying higher X-ray energies ($5 - 10\,\mathrm{keV}$) we avoid the absorbed part of the spectrum. With the combination of recent multiwavelength surveys we are able to create a sizable sample of ~ 500 sources, while spectroscopic redshifts combined with accurate photometric redshifts provide a 98% redshift complete sample, ideal to probe AGN evolution. We compute for the first time the XLF and its evolution in the $5 - 10\,\mathrm{keV}$ testing the different evolutionary models. We also employ computational methods such as i) the classical $1/V_{max}$ method which provides binned estimates without any assumption on the functional form of the luminosity function, ii) Maximum Likelihood Estimation (MLE) to find the best fit parameters for each model and finally iii) we perform Bayesian analysis to investigate in detail the probability distribution function for each parameter that describes the luminosity function. The latter approach gives an accurate view of the parameters without any assumptions on the form of the distribution around the best value.

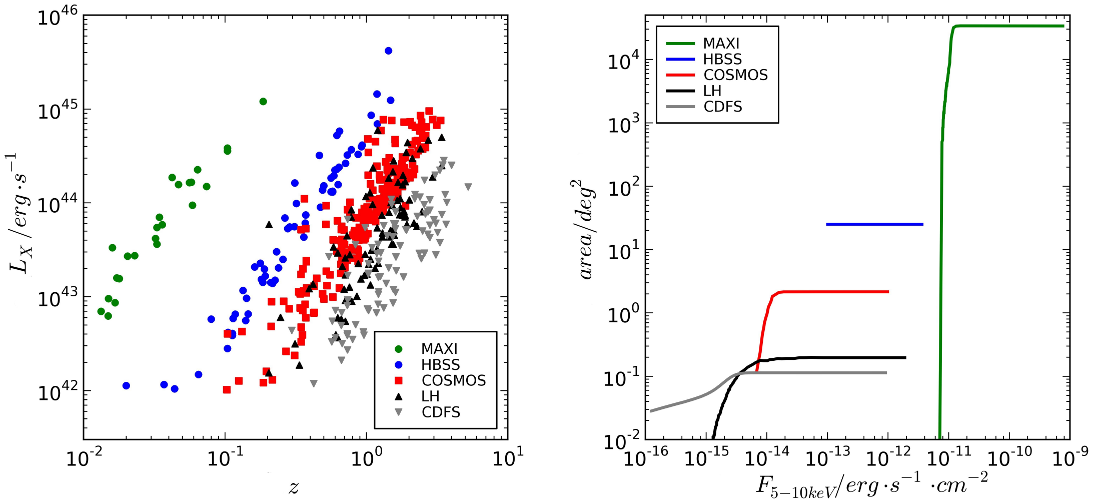
After presenting the dataset used for this work (§4.1.1), we are going to demonstrate the fact that the $5 - 10\,\mathrm{keV}$ band is practically unaffected by absorption (§4.1.2). In §4.2 we describe the models we used in computing the luminosity function. In §4.3 we describe the methods employed and present the preferred model, according to our selection criteria. In §4.4 we discuss the number density of AGN as a function of redshift. As we will show, our results are in agreement with the decrease found for the high redshift sample in the COSMOS survey (Brusa et al., 2009, Civano et al., 2011) and also with the color pre-selected high redshift sample of Aird et al. (2010), above the flux limit of our sample. On the contrary, large discrepancies appear below our flux limit which need further investigation. We adopt the cosmological parameters $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\Omega_k = 1 - \Omega_\Lambda - \Omega_m$ and $\mathrm{H}_0 = 70\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$.

4.1 Ultra - Hard AGN Selection

We combine wide angle surveys (MAXI: Hiroi et al. (2011), Hard Bright Serendipitous Survey (HBSS): Della Ceca et al. (2004), COSMOS: Cappelluti et al. (2009)) with pencil beam X-ray fields (Lockman Hole (LH): Brunner et al. (2008), Chandra Deep Field South (CDFS): Xue et al. (2011)) to create a sample of 5-10 keV detected AGN. The combined sample consists of 499 AGN, with luminosities between 10^{42}-10^{46}erg/sec, in the redshift range $0 < z < 4$. We exclude from this sample known stars and galaxy clusters.

Good coverage of the luminosity - redshift plane (Fig. 4.1 (a)) was possible due to the wide range of sky coverage and X-ray depth reached with the combination of these fields (Fig. 4.1 (b)). The number counts from the fields used in this work are presented in Fig. 4.2. A steep decrease is evident, for fluxes lower than $10^{-14}\mathrm{erg\,sec}^{-1}\,\mathrm{cm}^{-2}$, in agreement with observations from the $0.5 - 2\,\mathrm{keV}$ and $2 - 10\,\mathrm{keV}$ bands (Cappelluti et al., 2007, gives a recent compilation of X-ray number counts).

The sample is 98% complete in redshift, with 79% of the redshifts being spectroscopically determined. The remaining 21% of the sample consists of sources from the fields COSMOS, Lockman Hole, and Chandra Deep Field South, where the multiwavelength coverage of the fields allowed the computation of accurate photometric redshifts (Fig. 4.3). The accuracy in all fields is better than 7%, with a small fraction of outliers (defined in §3.1.2) (Salvato et al., 2009,



(a) (b)

Figure 4.1: (a) X-ray Luminosity - redshift plane. For the first time in the Ultra Hard band a good coverage is achieved. (b) Sky coverage as a function of X-ray flux for each survey.

2011, Luo et al., 2010, Cardamone et al., 2010, Rafferty et al., 2011, Fotopoulou et al., 2012).

In the following we present the characteristics per field in detail and demonstrate that for common values of hydrogen column densities N_H, the $5-10\,$keV band is affected only very little by photoelectric absorption, lifting the need to apply ambiguous corrections.

4.1.1 Dataset

MAXI Extragalactic Survey The Monitor of All-sky X-ray Image (MAXI) mission, on board the International Space Station (ISS) observes the entire sky every 92 minutes with two instantaneous fields of view each $160^\circ \times 3^\circ$. MAXI consists of two cameras, the Gas Slit Camera (GSC: Sugizaki et al. 2011, Mihara et al. 2011) sensitive in the $2-30\,$keV energy band and the Solid-state Slit Camera (SSC: Tsunemi et al. 2010, Tomida et al. 2011) sensitive in the $0.5-12\,$keV energy band.

Hiroi et al. (2011) presented the first MAXI/GSC 7-month data catalog detected in the $4-10\,$keV band and at high Galactic latitudes ($|b| > 10^\circ$, $34,000\mathrm{deg}^2$). Ueda et al. (2011) used 37 AGN from this catalog to compute the local AGN luminosity function. Here we are using the 24 AGN out of the 37 AGN presented in Ueda et al. (2011) which have spectroscopic redshift $\mathrm{z} > 0.01$. The flux limit of the sample is $\mathrm{f_x} = 1.5\,10^{-11}\mathrm{erg\,cm^2\,s^{-1}}$ covering a redshift range $0.0133 < z < 0.186$, with median redshift 0.0434.

Hard Bright Serendipitous Survey (HBSS) The XMM-Newton Bright Serendipitous Survey covers 25 deg^2 (Della Ceca et al., 2004) and provides two flux limited samples in the $0.5-4.5\,$keV and $4.5-7.5\,$keV band. In this work we are using the Hard sample $4.5-7.5\,$keV with a flux limit of $\mathrm{f_x} = 7\,10^{-14}\mathrm{erg\,cm^2\,s^{-1}}$. The optical counterparts and spectroscopic redshifts are presented in Caccianiga et al. (2008). The sources cover the redshift range $0.02 < z < 1.482$, with median redshift 0.32. This survey covers a very large area introducing to the sample

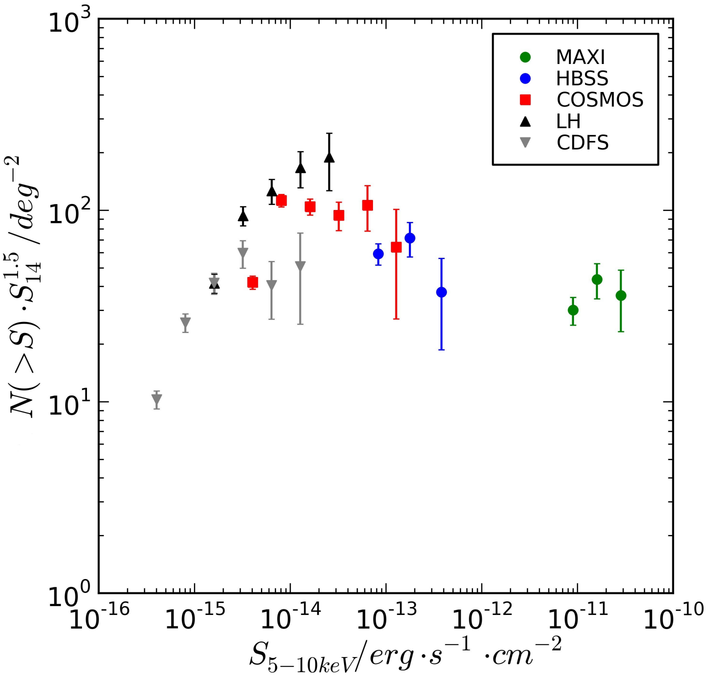


Figure 4.2: Number counts from all the surveys used in this work.

rare bright objects.

COSMOS The COSMOS field is one of the widest contiguous XMM field covering 2 deg^2. Having a good balance between depth and sky coverage, the XMM-COSMOS field contains 245 sources detected in the Ultra Hard Band (Cappelluti et al., 2009). The unprecedented multiwavelength coverage of this field provides optical (Capak et al., 2007), ultra-violet (Zamojski et al., 2007), near-infrared (McCracken et al., 2010), mid-infrared (Sanders et al., 2007, Ilbert et al., 2009, Frayer et al., 2009) counterparts and spectroscopic redshift (Trump et al., 2009, Lilly et al., 2007, 2009) (see Brusa et al., 2007, 2010, for a summary). For 79.2% of the sample spectroscopic redshifts are available, while the rest 18.3% of the sample has high quality photometric redshifts (Salvato et al., 2009, 2011) reaching a total of 97.5% redshift completeness. The COSMOS sources span a redshift range $0.04 < z < 3.34$, with a median redshift of 1.07.

Lockman Hole (LH) The Lockman Hole is one of the deepest XMM-Newton fields. The X-ray catalog presented in Brunner et al. (2008) contains 92 sources detected in the Ultra Hard Band and the optical counterparts are presented in Rovilos et al. (2011). Photometric redshifts and previously unpublished spectroscopic redshifts are presented in Fotopoulou et al. (2012) reaching 98.8% completeness. The Lockman Hole data set provides the faintest of the sources detected with XMM, allowing to probe the faint end of the luminosity function at redshift above 0.5. This dataset covers a redshift range of $0.118 < z < 3.41$, with median redshift 1.192.

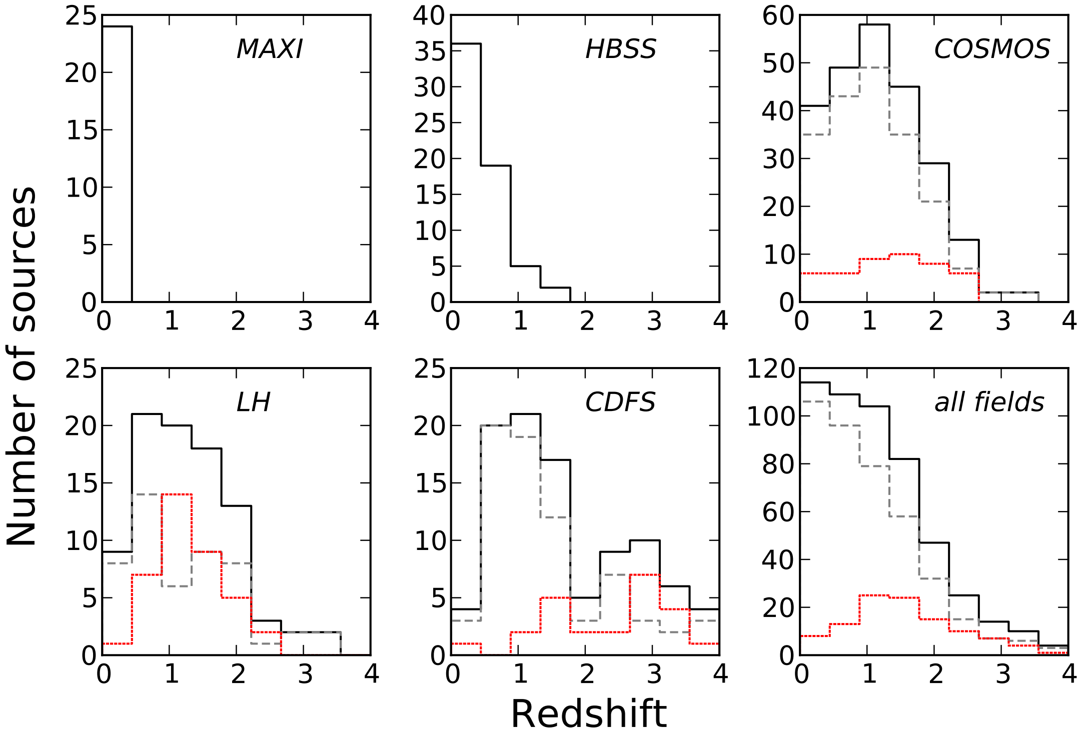


Figure 4.3: Redshift histograms per field and for the combined sample. Black solid line: all available redshifts, dashed gray line: spectroscopic redshifts, red dotted line: photometric redshifts.

Chandra Deep Field South (CDFS) We created the $5-10$ keV catalog in the CDFS using the latest *Chandra* observations of 4Ms (Xue et al., 2011)[1]. We cross-correlated the 5-10 keV detected sources with the available catalog in Xue et al. (2011) and found a counterpart for all the 97 detections within 2″. We adopt from the Xue et al. (2011) X-ray catalog the counterparts and their available redshifts keeping their 'best' redshift column which is spectroscopic or photometric for 100% of the sources. The inclusion of the CDFS data aids in populating the low luminosity and high redshift range and at the same time combining it with LH - the other deep field - we reduce the dependency of our results from cosmic variance. The redshift range covered by CDFS is $0.12 < z < 5.25$, with median redshift 1.5.

4.1.2 5-10 keV Benefit

X-ray spectra above 1 keV can be described adequately as a power law distribution:

$$F_x \propto \nu^{-\Gamma} \tag{4.1}$$

with $\Gamma= 1.4-1.6$ for radio loud AGN and $\Gamma= 1.8-2.0$ for radio-quiet AGN (Nandra and Pounds, 1994, Reeves and Turner, 2000, Piconcelli et al., 2005, Page et al., 2005, Mateos et al., 2005). Typically, absorption in the X-ray spectrum appears as a turn-over of the continuum

---

[1] The catalog was created by I. Georgantopoulos during a scientific visit at the University of Bologna

Table 4.1: Field statistics summary

	MAXI	HBSS	COSMOS	LH	CDFS	total
Detections	37	67	246	92	97	502
Stars	0	2	1	2	0	5
zspec	37	62	194	50	72	378
zphot	0	0	45	38	25	108
no z	0	0	5	0	0	5
no identification	0	2	1	1	0	4
clusters	0	1	0	1	0	2
completeness	100%	96.8%	97.5%	98.8%	100%	98%
redshift range	0.003-0.186	0.02-1.482	0.04-3.34	0.118-3.41	0.12-5.25	0.02-5.25
z median	0.0172	0.32	1.07	1.192	1.5	1.05
Included in sample[a]	24	62	233	86	94	499

[a] Sources with $42 < L_x < 46$ and $0.01 < z < 4.0$

with the lower energy range being more absorbed (details on the photoelectric effect are given in §1.4). In the case where X-ray spectra are not available, the count rates in the the soft ($0.5 - 2\,\mathrm{keV}$) and the hard ($2 - 10\,\mathrm{keV}$) energy bands are used to quantify the absorption. The ratio of the difference between these two energy bands over the total count rate is called the *hardness ratio*, HR:

$$HR = \frac{C(2-10\,keV) - C(0.5-2\,keV)}{C(2-10\,keV) + C(0.5-2\,keV)} \tag{4.2}$$

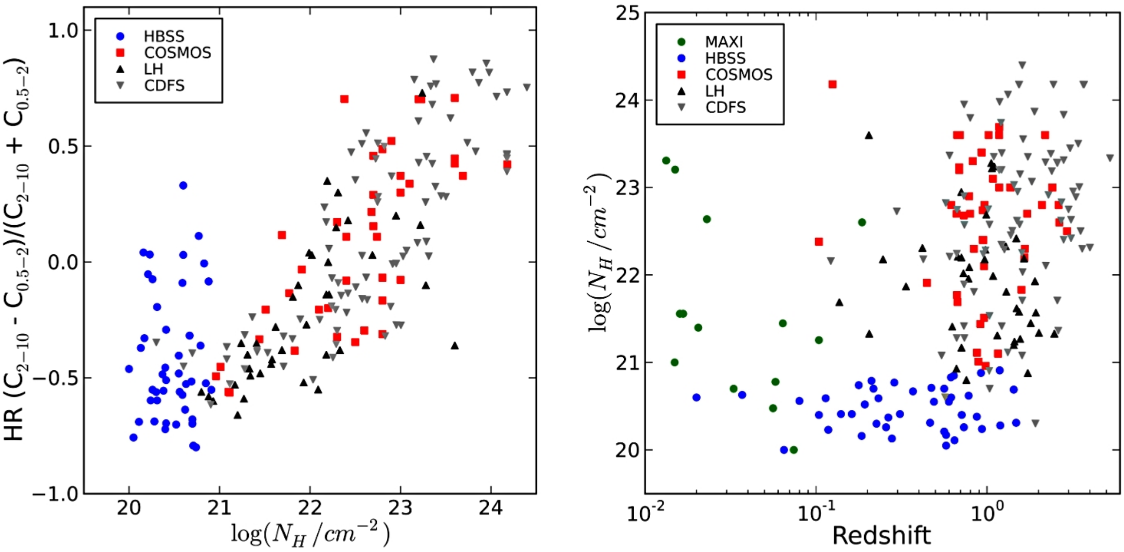
The HR serves only as a rough estimation of the intrinsic absorption of the source. In Fig. 4.4 (a), we show the comparison between HR and hydrogen column density $\mathrm{N_H}$ for the sources in our sample that have an estimated $\mathrm{N_H}$ value from X-ray spectra (MAXI:Ueda et al. (2011) and references therein, HBSS: Della Ceca et al. (2004), COSMOS: Mainieri et al. (2007, 2011), LH: Mateos et al. (2005), CDFS: Tozzi et al. (2006)). Even though there is a rough correlation between the two quantities, the scatter is large and is produced by the photon index (Γ) distribution of the sources. In addition we see, how the HR can cause an erroneous estimation of $\mathrm{N_H}$ value. For example, a few of the sources in the HBSS sample (blue points in Fig. 4.4 (a)) which all exhibit $20 < \log \mathrm{N_H} < 21$ would have been assigned $\log \mathrm{N_H}$ values as high as 23, if we would use the rough correlation between HR and $\mathrm{N_H}$ present in the plot.

In Fig. 4.4 (b), we give the $\mathrm{N_H}$ as a function of redshift, for the sources for which this information is available. We see that low redshift sources ($z < 0.1$) show small $\mathrm{N_H}$ values, while sources with $z \sim 1$ show a wide distribution of $\mathrm{N_H}$ values ($21 < \log \mathrm{N_H} < 24.5$). The true fraction of absorbed sources in the Universe is still an open question, and as we shall see in what follows, the observed distribution in the sample can be explained to some extent from selection effects.

Let us assume that the spectrum of an AGN can be described as a power law with photon index $\Gamma = 1.9$ absorbed by photoelectric absorption. Using the models *zphabs*cutoffpl* in PyXspec[2] we can calculate theoretical flux and count rate values for different $\mathrm{N_H}$ and redshift

---

[2] PyXspec is a Python interface to Xspec (Arnaud, 1996)



(a) (b)

Figure 4.4: (a) HR (HR versus hydrogen column density (N_H)). There is a large scatter in the correlation between the two quantities. (b) Hydrogen column density (N_H) as a function of redshift. Selection effects are responsible for the lack of objects with high N_H at low redshifts.

values. In Fig. 4.5 we plot the expected HR for several N_H values and redshift (blue lines, left hand side axis). The points in this figure correspond to the four fields used in this work (filled circles: HBSS, gray squares: XMM-COSMOS, empty circles: LH, triangles: CDFS). We exclude the MAXI sample from this plot since the soft band flux is not available. Fortunately, for this survey, all the N_H values are known and only two sources have column density higher than $10^{23} cm^{-2}$ (Fig. 4.4, (b)) . We immediately notice that with this crude estimation, the majority of the sources has N_H below $10^{23} cm^{-2}$ at any redshift.

On the other hand, if we plot the fraction of the observed flux over the intrinsic flux in the $5-10$ keV band for $N_H = 10^{23}, 5\,10^{23}, 10^{24} cm^{-2}$ (red lines, right hand side axis), we clearly see that even for $N_H = 10^{23} cm^{-2}$ the flux lost from absorption is never more that 20%. Moreover, for sources with higher hydrogen column density (e.g. $5\,10^{23} cm^{-2}$), the flux lost is less than 20% at redshifts higher than ≈ 1.2, while for the extreme case of $N_H = 10^{24} cm^{-2}$ the flux lost is less than 20% at redshifts higher than ≈ 1.6.

If we do not correct for absorption, the affected objects will be those lying at z=0-1.6 with $N_H > 10^{23} cm^{-2}$ (shaded area), which are only 30 out of 499 objects[3] (6.4%). This small fraction of objects will not affect the result of the luminosity function, as at those redshift bins the majority of the objects are unabsorbed ($HR < -0.3$). Thus, in order to avoid the ambiguity of the N_H determination from HRs, we do not to correct for the absorption and assume that the spectrum of an AGN can be described adequately as a power law with photon index Γ = 1.9.

---

[3] Including the two MAXI sources.

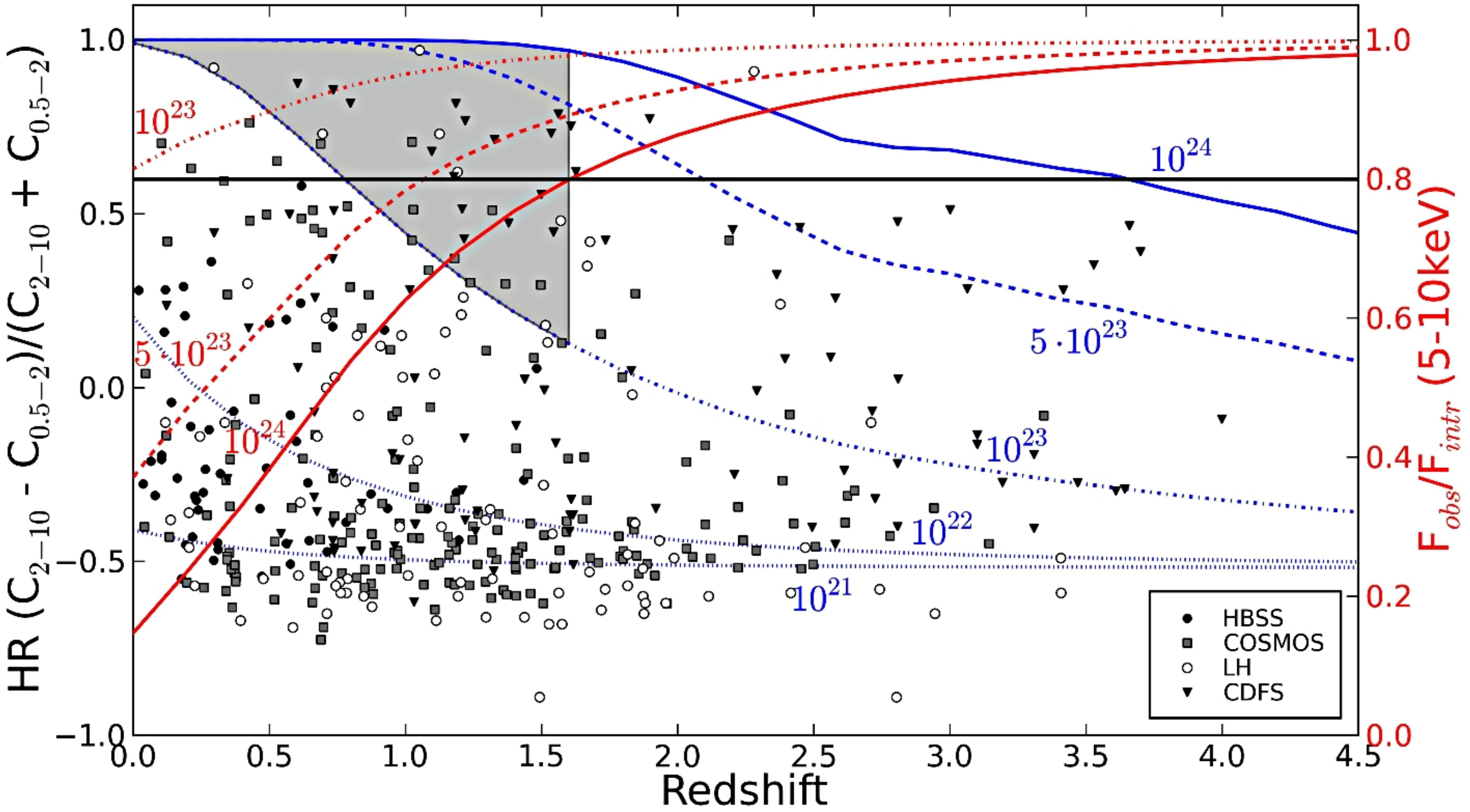


Figure 4.5: Left Axis (points and blue lines): the HR of AGN selected in the $5-10\,$keV band versus redshift (points). The blue lines show the expected HR of a power law spectrum absorbed by photoelectric absorption for different N_H values. The majority of the data used in this work have $N_H < 10^{23} cm^{-2}$. Right Axis (red lines): the fraction of the observed flux in the 5-10 keV band versus the intrinsic flux. Only a handful of objects (shaded area) will be affected by the photoelectric absorption, with the effect being smaller at higher redshifts. The horizontal black line marks the limit of 80% flux retrieval in the 5-10 keV band. The theoretical lines have been computed with PyXspec, the python interface to Xspec (Arnaud, 1996).

4.1.3 5-10 keV Limitation

The current X-ray observatories, namely XMM-Newton and Chandra, are less sensitive to energies above 5 keV. This fact means that only bright sources are detected in the $5-10\,\mathrm{keV}$ energy band since typically the detection limit in the 5-10 keV is an order of magnitude higher compared to the $0.5-2\,\mathrm{keV}$ energy band. Consequently, in this work we do not include in this AGN evolution study very faint sources. Works on the deepest X-ray fields show that the faint population has the same characteristics in the X-rays as the brighter population. Even though there are interesting cases on an individual source basis, we do not expect differences in a study of the global population as it has been demonstrated that the $5-10\,\mathrm{keV}$ does not select a special AGN population (Della Ceca et al., 2008). Nevertheless, the lack of very faint objects sampling high redshift and low luminosity bins might introduce bias in the determination of the faint end of the luminosity function. We will return to this point during in the discussion (§4.4).

4.2 Modeling the Luminosity Function

Many models have been proposed and used over the years to describe the evolution of the luminosity function[4] with redshift with gradually increasing complexity. Early modeling attempts of the X-ray luminosity function identified the local AGN sample (z $\sim$ 0) to follow a power law distribution with a turn over at low luminosities (Maccacaro et al., 1983, 1984):

$$\frac{d\Phi(L_x, z=0)}{d\log L_x} = \frac{A}{(\frac{L_x}{L_\star})^{\gamma_1} + (\frac{L_x}{L_\star})^{\gamma_2}} \tag{4.3}$$

where, L_x, is the luminosity at which the 'break' occurs and, γ_1 and γ_2 the slopes of the power law distributions below and above L_x.

As more numerous samples of AGN became available including fainter sources, it was evident that the AGN luminosity function was evolving strongly with redshift up to z $\sim$ 1.6 and then evolving less prominently (i.e. Boyle et al., 1994, Page et al., 1996, Jones et al., 1997). The behavior of the evolution was unclear and the models most commonly tested were the *Pure Luminosity Evolution* (PLE) which assumes that only the luminosity of AGN is changing with time and the *Pure Density Evolution* (PDE) which assumes that only the number density of AGN is changing with time. With the aid of modern X-ray observatories (ROSAT, XMM-Newton, Chandra), the luminosity function of AGN seems to be more complicated than the simple PLE and PDE models, with the evolution after some critical redshift not only stopping but potentially becoming negative.

Nowadays, the model used to describe the evolution is mainly the *Luminosity Dependent Density Evolution* (Miyaji et al., 2000), which describes the changes in the number density of AGN allowing also for separate evolution between faint and bright objects. A combination of PLE and PDE, namely the *Independent Luminosity and Density Evolution* (ILDE) was used in Yencho et al. (2009) to describe the evolution of AGN below redshift z=1.2, therefore excluding in the formalization the critical redshift, z_c, after which the evolution changes dramatically. A more general version of the ILDE, the *Luminosity And Density Evolution* (LADE) was introduced in Aird et al. (2010), which - as the name suggests - allows for evolution in luminosity and number density simultaneously. This model includes a critical redshift (z_c) value after which the rate of the evolution changes.

---

[4] The luminosity function, introduced in §1.5.2, gives the number of objects per unit comoving volume per unit luminosity.

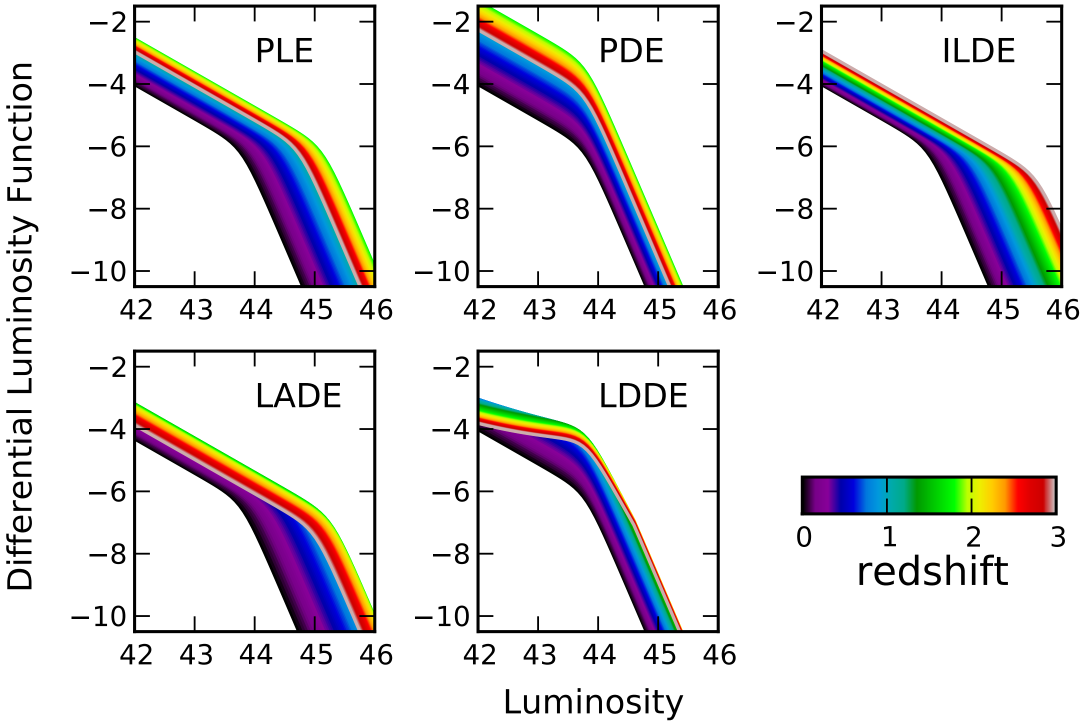


Figure 4.6: Comparison of common AGN luminosity function evolutionary models. Models computed using fictitious evolutionary parameters to demonstrate the qualitative difference in evolution. The color scale shows the evolution of each model with redshift ranging from z=0 (black) to z=3 (red). The critical redshift is chosen $z_c = 1.8$ (bright green) for all models.

Fig. 4.6 shows a visual overview of the evolutionary models PLE, PDE, ILDE, LADE and LDDE. In these plots, we give the differential luminosity function versus luminosity computed at several redshifts given by the color scale. The models were computed using fictitious parameters to demonstrate the qualitative difference in the evolution. The critical redshift is chosen to be $z_c = 1.8$ for all models. In the rest of this section, we give the mathematical formula describing each model and a brief description connecting the evolutionary schemes to the AGN physics.

Pure Luminosity Evolution – PLE

The PLE model was one of the first examined (Schmidt, 1968). The physical interpretation of this model is that AGN luminosities change with time, but the total number of AGN remains constant. As seen from Figure 4.6, the PLE model is apparent as a shift of the luminosity function from higher to lower luminosities when moving from higher to lower redshift. Since the shape of the luminosity function is assumed to remain the same, this would be interpreted as a change in luminosity of the global AGN population. The evolution is most commonly expressed as:

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L/e(z), z=0)}{d\log L_x} \tag{4.4}$$

with,

$$e(z) = \begin{cases} (1+z)^{p_1} & z \leq z_c \\ (1+z_c)^{p_1} \cdot (\frac{1+z}{1+z_c})^{p_2} & z \geq z_c \end{cases} \tag{4.5}$$

where z_c the redshift after which the evolution, $e(z)$, changes behavior and also follows a broken power law distribution with slopes dependent on p_1 and p_2.

Pure Density Evolution – PDE

The PDE model was also examined in the very early studies of luminosity function evolution (Schmidt, 1968). The physical interpretation of this model is that AGN change in numbers, but their luminosities remain constant. This would be possible if the transition from active to inactive phase and vice versa would be rapid and thus, hardly observable. This evolution would appear as a change in the normalization of the luminosity function (see Fig. 4.6). PDE is expressed as:

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L,z=0)}{d\log L_x} \cdot e(z) \tag{4.6}$$

with, the evolutionary factor, $e(z)$ given by eq. 4.5.

Independent Luminosity Density Evolution – ILDE

The Independent Luminosity Density Evolution was used by Yencho et al. (2009) to describe the evolution of AGN for redshifts below z=1.2. This model postulates that there is a simultaneous change of the luminosity and number of AGN. Since this model is confined below redshift z<1.2, no critical redshift value was introduced.

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L/e_L(z),z=0)}{d\log L_x} e_D(z) \tag{4.7}$$

with,

$$e_L(z) = (1+z)^{p_L} \quad and \quad e_d(z) = (1+z)^{p_D} \tag{4.8}$$

similarly to the PLE (eq. 4.5) and PDE (eq. 4.6) models below z_c.

Luminosity And Density Evolution – LADE

The Luminosity and Density Evolution was introduced by (Aird et al., 2010). This model allows for evolution in the luminosity function both in luminosity and number density, but with the inclusion of a critical redshift value.

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L/e_L(z),z=0)}{d\log L_x} \cdot e_d(z) \tag{4.9}$$

with the evolution in redshift and number density respectively,

$$e_L(z) = (\frac{1+z_c}{1+z})^{p_1} + (\frac{1+z_c}{1+z})^{p_2} \quad and \quad e_d(z) = A \cdot 10^{d(1+z)} \tag{4.10}$$

We point out that in the formalization given in the Aird et al. (2010) there was an inconsistency. The term describing the luminosity evolution, $e_L(z)$ was not properly normalized at z=0. This might result in a misinterpretation of the z=0 luminosity function parameters. Typically the models used in the literature are normalized at z=0, so that the parameters deduced for L_x, γ_1, γ_2 and, A can be inserted in eq. 4.3 and the z=0 luminosity function is reproduced. The correct computation of the LADE model at z=0 requires that either eq. 4.9 is used, or that the parameter L_x in eq. 4.3 is multiplied by the factor $(1 + z_c)^{p_1} + (1 + z_c)^{p}_{2}$.

Luminosity Dependent Density Evolution – LDDE

The Luminosity Dependent Density Evolution was introduced by Miyaji et al. (2000) to describe the Soft X-ray luminosity function of type 1 (unabsorbed) AGN and has been used extensively ever since. This more complex model encapsulates the physics that the number density of AGN changes, but since the evolution of bright and low luminosity AGN exhibits different timescales, the critical redshift, z_c, depends on the luminosity.

Hopkins et al. (2006) interpreted the AGN luminosity function in terms of quasar lifetimes and found good agreement of their simulations with the LDDE modeling and relevant observations. They claim that the observed break in the luminosity function corresponds to the maximum of the peak luminosity distributions of quasars at a certain redshift. The bright end of the luminosity function is populated by those quasars that emit at their peak luminosity, while the faint end is populated by the quasars that emit at lower luminosities. In their modeling, quasars spend the majority of their lifetime below their peak luminosity while at the same time, more luminous objects will transit to a less luminous stage faster than the objects with lower peak luminosity. This implies that the slope of the faint end luminosity function is flatter at higher redshift.

In this work we are using the formalization introduced by Ueda et al. (2003):

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L,z=0)}{d\log L_x} \cdot e(L,z) \tag{4.11}$$

with,

$$e(L,z) = \begin{cases} (1+z)^{p_1} & z \leq z_c(L) \\ (1+z_c)^{p_1} \cdot (\frac{1+z}{1+z_c})^{p_2} & z \geq z_c(L) \end{cases} \tag{4.12}$$

and

$$z_c(L) = \begin{cases} z_c^* & L \geq L_a \\ z_c^* \cdot (\frac{L}{L_a})^a & L < L_a \end{cases} \tag{4.13}$$

We express the evolution factor $e(L,z)$ of eq. 4.12 as:

$$e(z,L_x) = \frac{(1+zc)^{p1} + (1+zc)^{p2}}{(\frac{1+z}{1+zc})^{-p1} + (\frac{1+z}{1+zc})^{-p2}} \tag{4.14}$$

with z_c defined as in eq. 4.13. This formula is equivalent to the Ueda et al. (2003) evolution factor creating a smooth transition of the XLF before and after the critical redshift. Additionally, it is normalized correctly at redshift zero.

4.3 The 5-10 keV Luminosity Function

In this section, we present our results from the computation of the luminosity function. First we show binned estimates of the luminosity function based on the $1/V_{max}$ method, which is a model independent estimation of the luminosity function and its evolution. Then, we test all of the aforementioned models and select the model that describes best our dataset based on the Akaike Information Criterion and Bayesian model selection. Both of the selection criteria, identify the same model as the best parametrization of our dataset, namely the LDDE.

4.3.1 1/V$_{\rm max}$

The $1/V_{max}$ method (Schmidt, 1968) has been used widely in the literature to estimate the luminosity function of AGN (described in detail in §1.5.2). The advantage of this method is that the estimation is independent of any assumption for the underlying model. When a sizable sample of AGN is used, the computation of the luminosity function in thin redshift bins is powerful in revealing the presence of evolution and roughly its shape. Nevertheless, the method as introduced by Schmidt (1968), might overestimate the maximum accessible volume in the luminosity-redshift bins truncated by the sample's flux limit. Page and Carrera (2000) showed that a better estimation of the luminosity function is given by:

$$\frac{d\Phi(L,z)}{d\log L} = \frac{n}{\int_{\log L_{min}}^{\log L_{max}} \int_{z_{min}}^{z_{max}(L)} \frac{dV}{dz} dz\, d\log L} \tag{4.15}$$

where, n, is the number of AGN in the bin $[L_{min}, L_{max}]$, and $[z_{min}, z_{zmax}(L)]$. The value $z_{max}(L)$ corresponds to the maximum redshift up to which the n sources would be still present in the sample and it is either the maximum redshift of the redshift bin or it is given by the flux limit.

Fig. 4.7 shows the result of the $1/V_{max}$ method (points) while the lines plotted serve only as guides to the eye. Actual model estimations will be presented in subsequent sections. The main outcome of this method is that the luminosity function in the 5-10 keV follows a broken power law distribution at low redshift (e.g z=0.20, dark blue points). As we trace the evolution at higher redshifts, we see that the luminosity function has a complicated behavior beyond redshift z=1.5. There is a sharp change in the evolution which at these redshifts is neither consistent with the PLE behavior, nor with the PDE behavior which would appear as a shift along the x and y axes respectively.

4.3.2 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) is a model parameter estimation method commonly used in the literature to specify the best fit parameters of models, among which is the luminosity function. The probability of observing an AGN with luminosity between [L, L+dL] and redshift between [z, z+dz] is in general given by:

$$p(L,z) = \frac{1}{N}\phi(L,z)\frac{dV}{dz} \tag{4.16}$$

where, N, the total number of AGN in the Universe, $\phi(L,z)$ the luminosity function and, dV/dz the comoving volume. Treating the observation of n sources out of the available N AGN in the Universe as a Poisson process (applicable in the limit $n/N \ll$) (Marshall et al., 1983) the likelihood of observing n sources is given by:

$$\begin{aligned} \ln \mathcal{L}(L,z) = &\sum_{i}^{n} \ln \int\int \phi(z_i, LogL_i)\frac{dV}{dz} p(\log L_x)p(z)d\log L_x dz \\ &- \int\int \phi(z, LogL)\Omega(z, LogL)\frac{dV_c}{dz} dz dLogL \end{aligned} \tag{4.17}$$

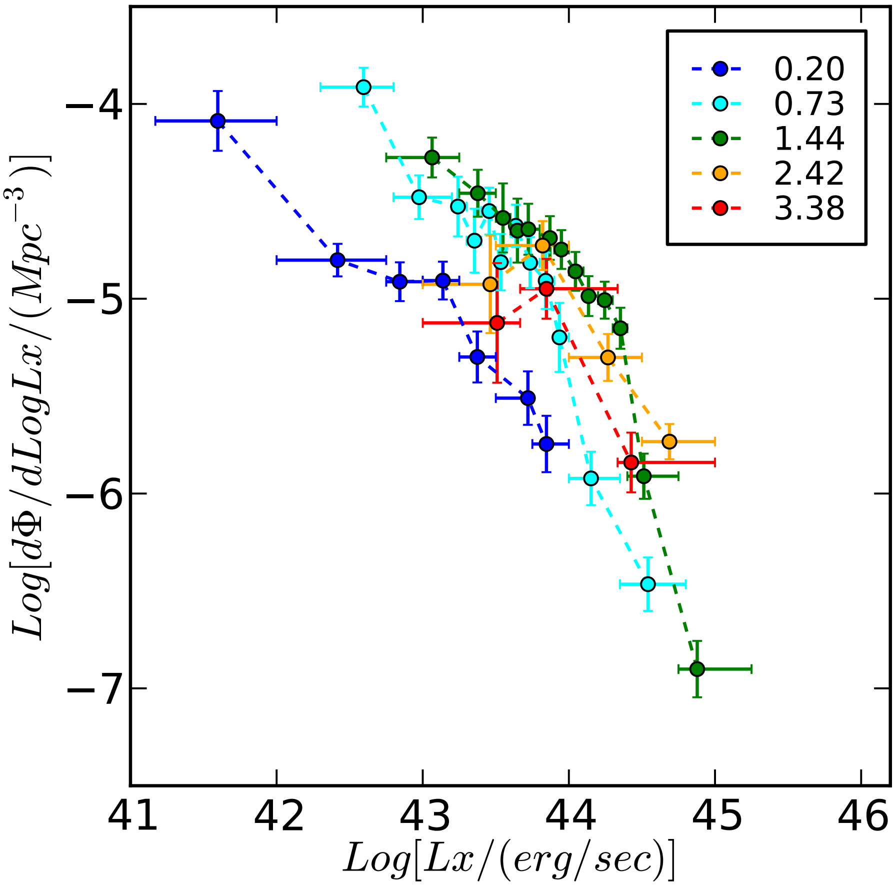


Figure 4.7: Binned estimates of the AGN luminosity function and its evolution. The $1/V_{max}$ method shows that the luminosity function at low redshifts is to a good approximation a broken power law, while at higher redshifts a complicated evolution is present. The lines are not fit to the data, but only serve as guides to the eye.

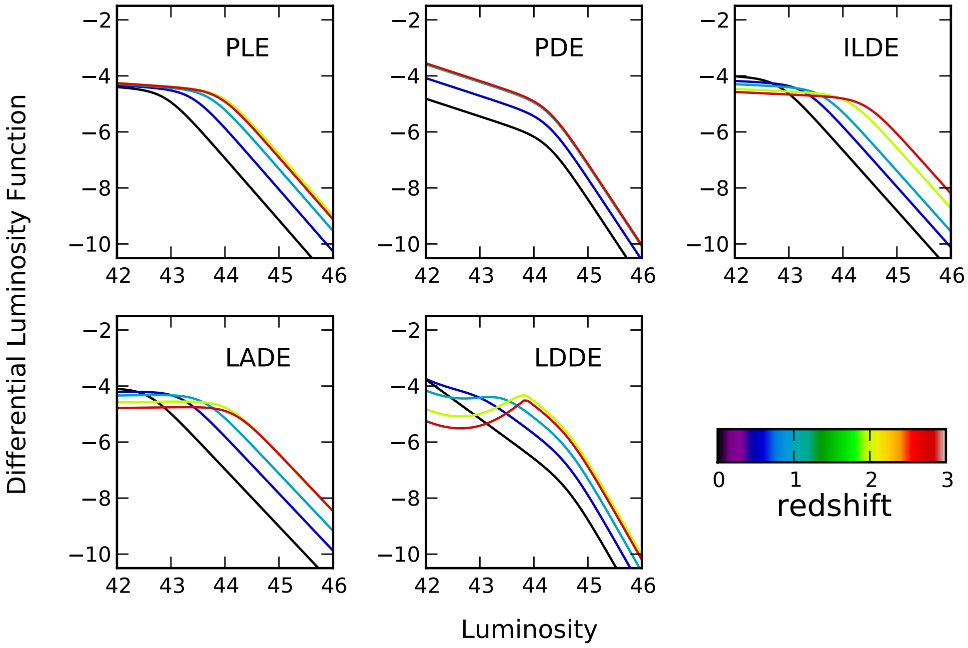


Figure 4.8: MLE results for all the models fitted to the 5-10 keV AGN dataset. The model parameters are given in Table 4.2

where $p(\log L_X)$, $p(z)$ the uncertainties on the flux and redshift, respectively.

It is computationally more attractive to find the minimum of the logarithm of the likelihood. Therefore, is usual to determine the parameters that minimize the function:

$$\mathcal{S} = -2 \ln \mathcal{L} \tag{4.18}$$

The factor "2" is introduced to create an equivalence of the uncertainty estimates to those calculated from least squares method.

In this work, we assume that the error in the X-ray flux is described by a gaussian[5]. For the uncertainties in spectroscopic redshifts, since we do not have an error estimation for the majority of the sources, we adopt a conservative value for the uncertainty of 0.01. For the photometric redshifts for the COSMOS and the Lockman Hole samples, we use the probability distribution function calculated with LePhare during the photometric redshift estimation.

To determine the parameters which minimize eq. (4.18) and hence maximize the likelihood, we use the code MINUIT (James and Roos, 1975) widely used in high energy physics, through the python interface, *pyMinuit*[6]. In Table 4.2 we give the maximum likelihood parameters for all models and their 1-σ errors as determined in MINUIT. We also give the estimated distance to the minimum (edm), which provides an estimation of the quality of the fit[7].

A first glance of the fitted models (Fig. 4.8) might lead to the conclusion that the models show large discrepancies. As we show in Appendix A, comparing the Bayesian results to the model-independent $1/V_{max}$ method, this is not true. Nevertheless, there are a few features

---

[5] None of our sources have less than 15 counts

[6] http://code.google.com/p/pyminuit/

[7] Large edm value points to untrustworthy fit.

Table 4.2: Maximum Likelihood Parameter Estimation

Parameter	Model				
	PLE	PDE	ILDE	LADE	LDDE
$\log L_0$	42.97±0.10	44.26±0.07	42.81±0.12	44.42±0.10	44.63±0.13
γ_1	0.13±0.10	0.63±0.05	0.09±0.11	-0.03±0.13	1.40±0.07
γ_2	2.21±0.11	2.92±0.20	2.16±0.10	2.04±0.10	3.4±0.4
p_1	2.75±0.13	4.07±0.22	2.74±0.20	-1.3±0.9	4.89±0.28
p_2	-0.15±0.26	0.04±0.23	-1.21±0.26	3.56±0.25	-3.8±0.6
z_c	1.526±0.006	1.031±0.008	.....	2.3±0.3	2.03±0.14
$\log L_\alpha$	.....	.....	.....	.....	43.85±0.06
α	.....	.....	.....	.....	0.55±0.06
d	.....	.....	.....	-0.23±0.05	.....
$\log A$	-4.51±0.10	-6.24±0.11	-4.08±0.14	-3.83±0.16	-7.45±0.29
edm[a]	1.53e-07	1.28e-06	1.58e-07	2.93e-06	2.28e-07
AIC$_\mathrm{i}$-AIC$_{LDDE}$	129	286	189	110	0

[a] Estimated Distance to the Minimum

[a] Covariance matrices for all the models are found in Appendix A

that should be pointed out. First, the ILDE model, even though it does not contain a z_c value after which the evolution changes, can effectively mimic this behavior with the combination of positive luminosity evolution and simultaneous negative density evolution. Additionally, the LDDE model shows at high redshift a curvature in the faint end power law. This shape is due to the dependence of the z_c on the luminosity, which is defined to be a sigmoidal function (eq. 4.13). We will return to the discussion of the results in §4.4.

AIC Model Selection Even though the maximization of the likelihood does not provide us with absolute goodness of fit estimation, we can make and educated guess on which is the best fit model using the Akaike Information Criterion (AIC, Akaike, 1974).

The AIC, selects the model with the least information loss, penalizing the model with the higher number of free parameters. Therefore, it incorporates the Occam's razor principle according to which among two equally plausible explanations, the simplest one is preferred. The AIC is given by:

$$AIC = 2k - 2\ln L \tag{4.19}$$

where, k, is the number of parameters present in the model and, L, the maximum likelihood value. According to this criterion the preferred model having the least AIC value is the LDDE. In Table 4.2, we give the difference between the AIC value of each model and the LDDE, AIC$_i$-AIC$_{LDDE}$. Since all other models show large $AIC_i - AIC_{LDDE}$ values, the only acceptable model is the LDDE according to the AIC criterion.

4.3.3 Bayesian Analysis

An alternative approach in performing inference, is Bayesian analysis. According to Bayes' theorem the probability, $p(\theta|D, M)$, of the model parameters θ, given the observed data, D, and the model M, is proportional to the probability, $p(\theta|M)$, of prior knowledge on the model parameters θ, times the probability, $p(D|\theta, M)$, to observe the data, D, under the given model, M, and set of model parameters θ. The term $p(\theta|D, M)$, is called the *posterior* and gives us the probability distribution of the model parameters. Expressing Bayes' theorem in mathematical terms:

$$p(\theta|D, M) = \frac{p(\theta|M)p(D|\theta, M)}{\int p(\theta|M)p(D|\theta, M)d\theta} \tag{4.20}$$

where $p(D|\theta, M)$ is also called the *likelihood function* for a given model. The denominator of eq. 4.20, is independent on the model parameters, and is called the *evidence*. This term is used to select the model that describes best the dataset. We point the interested reader to Kelly et al. (2008) for in depth discussion on Bayesian analysis of the luminosity function.

Combining eq. 4.17, which describes the likelihood to observe a sample of n AGN with luminosities, $\{L_i\}$, and redshifts, $\{z_i\}$, with prior information on the model parameters, we can perform Bayesian analysis. It is often argued that Bayesian analysis is not objective due to the presence of priors. Nevertheless, if enough information is present in the data and the likelihood, the prior information is of minor importance. This can be seen in the cases where the final probability distribution of a model parameter is significantly different from the initially assumed prior. In this work we are using flat (non-informative) distributions of priors for all parameters. The reason for this choice is that the expected range for each parameter is known by previous works of luminosity functions in the $2-10\,\mathrm{keV}$ band, but we do not wish to assume any shape for the distribution of the parameters.

The method we used to explore the posterior is the *Markov Chain Monte Carlo* (MCMC). Very briefly, the MCMC algorithm is sampling the parameter space of the posterior preferentially towards values of larger likelihood. If we start from an initial value, θ_0, then a random set of θ_i is proposed according to the prior information. If this new set of parameters corresponds to higher likelihood value, then this proposal is accepted and the procedure continues with a new proposed value θ_i. If the new set of parameters corresponds to lower likelihood value, the proposal is accepted with probability 30%. This happens in order to allow the chain to explore the posterior less strictly, and discover features such as additional modes, in case they are present. The algorithm continues indefinitely, until a stopping criterion is matched. In our case, the stopping criterion is the maximum number of iterations. Using PyMultinest[8] (Feroz et al., 2009) we verified that the posterior is unimodal. It is then justified to used one long chain of MCMC samples and assume that if the chain is long enough, that the procedure will converge to the maximum.

To perform Bayesian analysis we used APEMoST, a FORTRAN Markov Chain Monte Carlo (MCMC) code which has available a Python wrapper[8].As starting point for the MCMC we used the MLE best parameters. Setting the prior ranges for the parameters, we allowed the code to calibrate the step-width in order to reach acceptance rate of the order 0.3. It is important to allow the Markov Chain to perform enough steps, so as to reach roughly the correct parameter space close to the peak of the likelihood of the posterior. Therefore, is typical to ignore the first part of the iteration procedure, called *burn in* iterations. After 10^4 burn in iterations, we obtain 10^5 draws from the posterior. These draws can be viewed as a random, representative sample of the posterior around its maximum. Thus we can use them to infer the properties of the posterior.

Bayesian Model Selection As discussed in the previous section, the denominator of eq. 4.20, is often called the *evidence*. Higher value of Bayesian evidence is the criterion which gives the

---

[8] https://github.com/JohannesBuchner/PyMultiNest

Table 4.3: Bayesian Evidence

	Model				
	PLE	PDE	ILDE	LADE	LDDE
Δ Evidence	66.09	114.42	96.78	57.46	0.0

Table 4.4: LDDE - MCMC Parameter Estimation

Parameter	Prior Interval	Mean	Standard Deviation	Min – Max
L_0	40.0 – 46.0	44.58	0.09	44.34 – 44.83
γ_1	-1.0 – 3.0	1.38	0.05	1.23 – 1.53
γ_2	1.5 – 6.0	3.30	0.29	2.54 – 4.94
p_1	3.0 – 7.0	4.84	0.19	4.15 – 5.51
p_2	-6.0 – -1.0	-3.8	0.5	-5.95 – -2.25
z_c	0.5 – 3.5	2.05	0.10	1.68 – 2.43
L_a	40.0 – 46.0	43.85	0.05	43.68 – 44.05
α	0.0 – 1.0	0.55	0.04	0.41 – 0.73
A	-9.0 – -4.0	-7.33	0.18	-7.87 – -6.89

preferred model that describes best the dataset. The code used here, APEMoST, provides the Bayesian evidence for each of the models tested. The LDDE model shows the highest evidence and in Table 4.3 we provide the difference of the evidence between the LDDE and the other models. Given the large difference in the evidence, we accept the LDDE model as the one describing best our dataset, under the assumed priors. Note that in the MCMC, we use eq. 4.17, hence the factor of two difference between the Bayesian evidence and the AIC results.

In Table 4.4 we summarize the MCMC result for the LDDE model. We report for each parameter the prior interval, the mean value, the standrard deviation, and the minimum and maximum values. In Fig. 4.9 we plot the marginal distribution for each parameter, comparing the result obtained with MLE and with the Bayesian approach. The gray shaded area depicts the probability function, the solid vertical line marks the mean parameter value and the black bell shaped line shows the best gaussian representation for each parameter. Similarly, the blue gaussian is the estimated MLE parameter value, also shown as an empty circle above the distribution with the 1-σ error. It is evident from this plot that the MLE and Bayesian analysis give consistent results for the best fit parameters. Nevertheless, using MCMC to explore the posterior we have the added benefit of determining the shape of the probability distribution around the best fit value, without having to make assumptions on its shape. For the sample used in this work the uncertainties on the model parameters do not follow a gaussian distribution for the break luminosity, L_0, and the normalization, A. It is also interesting to see, that for this particular model, the MLE values of L_0, fl_1, fl_2 and A, were local minima.

In Fig. 4.10 we plot the differential luminosity function for several redshift bins. The black points are the result of the $1/V_{max}$ method. Since it is an independent method, it serves as a

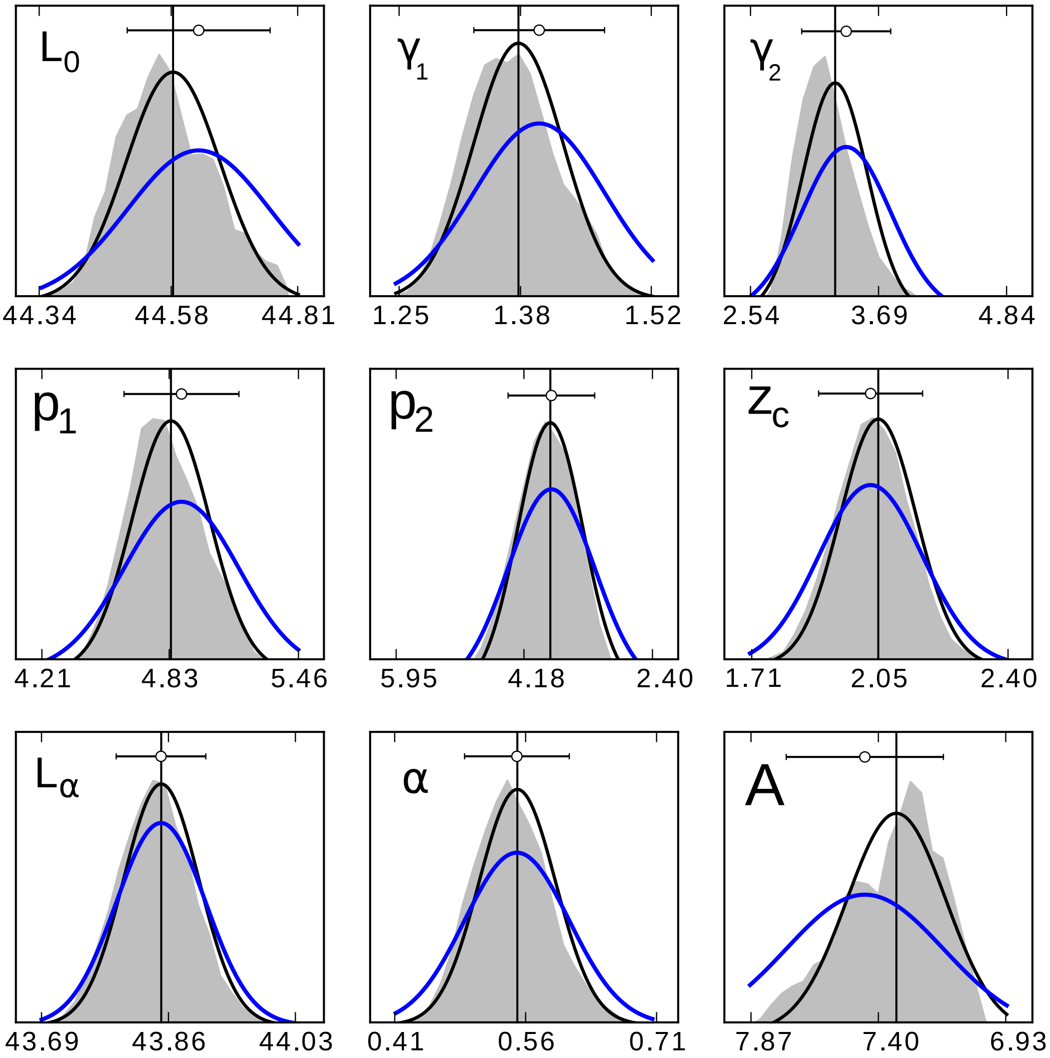


Figure 4.9: Marginal probability distribution function for LDDE parameters. The solid vertical line is the mean value from the MCMC simulation. The bell shaped line is the best gaussian representation of the distribution, black: bayesian analysis, blue: MLE result. The MLE result is also represented as an empty circle along with the 1-σ error above the distribution. The mean and standard deviation of the bayesian result are given in Table 4.4

sanity check for our fitting result. The dashed black line is the luminosity function computed at redshift zero, using the most probable parameters. The solid black line is drawn using the most probable parameters at the median redshift of the dataset in each redshift bin, while the gray shaded area encloses 99% probability of the differential luminosity function, $d\phi/dlogL_x$. To determine this area, we compute $d\phi/dlogL_x$ in each redshift bin for 40 values of the luminosity, L_x, for all the 10^5 draws from the posterior. In this way, we incorporate naturally the true shape of the uncertainties for all parameters.

4.4 Discussion

4.4.1 Model Comparison

In Appendix A, we present in detail the results from the parameter estimation procedures of the models discussed here. More specifically, we give the best parameters according to MLE and their covariance matrix, a summary of the Bayesian parameter estimation similar to Table 4.4, and we also plot similarly to Fig. 4.9 and Fig. 4.10 the marginal distribution of the parameters and the differential luminosity function for nine redshift bins in comparison with the $1/V_{max}$ results. These plots reveal that the seemingly large discrepancies we pointed out in §4.3.2, are consistent with the binned estimates of the $1/V_{max}$ method. Thus, it is hard to distinguish the model that describes best the evolution simply using the $1/V_{max}$ method. Elaborate selection criteria must be used, such as the AIC and Bayesian model selection that were used here.

Our modeling shows that the 5-10 keV AGN luminosity function is best represented by the *Luminosity Dependent Density Evolution* - LDDE model. This result is in agreement with previous studies in the Soft X-ray energy band $0.5-2$ keV (Miyaji et al., 2000, Hasinger et al., 2005, Ebrero et al., 2009) and studies in the $2-10$ keV (Ueda et al., 2003, La Franca et al., 2005, Silverman et al., 2008, Ebrero et al., 2009, Yencho et al., 2009). Aird et al. (2010) used the PLE, LDDE and, LADE models to describe their $2-10$ keV AGN sample and found only little evidence to prefer the more complicated LDDE model over the simpler LADE model, thus concluding that the LADE model is adequate to describe the AGN evolution. On the contrary, in this work we find strong evidence to prefer the LDDE model over every other simpler model tested here. Thus we conclude that the best description of the luminosity function is achieved with the LDDE model, although a caveat in the interpretation must be considered (see following subsection).

4.4.2 Malmquist Bias and Cosmic Variance

In §4.3.2, we noted the curved shape of the faint end of the luminosity function, present in the LDDE model. This shape is due to eq. 4.13 and is present in the luminosity function for certain combination of parameters. Since this model has the most degrees of freedom, it adjusts more easily even to subtle changes in the dataset. The drop of the luminosity function at low luminosities and high redshifts, drives the model to change the slope of the faint end dramatically and thus, creating this curved feature.

The reason for the change in faint end slope, can be either physical (see also §4.2) describing the intrinsic AGN peak emission luminosity of AGN at this redshifts, or it can either be an artifact in the dataset. Such an artifact can be introduced in cases were the faint AGN population is underrepresented in the sample due to the fact that fainter objects (virtually corresponding to high redshift and low luminosity) are more difficult to be detected, also known as *Malmquist Bias*. Ideally, since the luminosity function is a volume weighted function, this bias should not

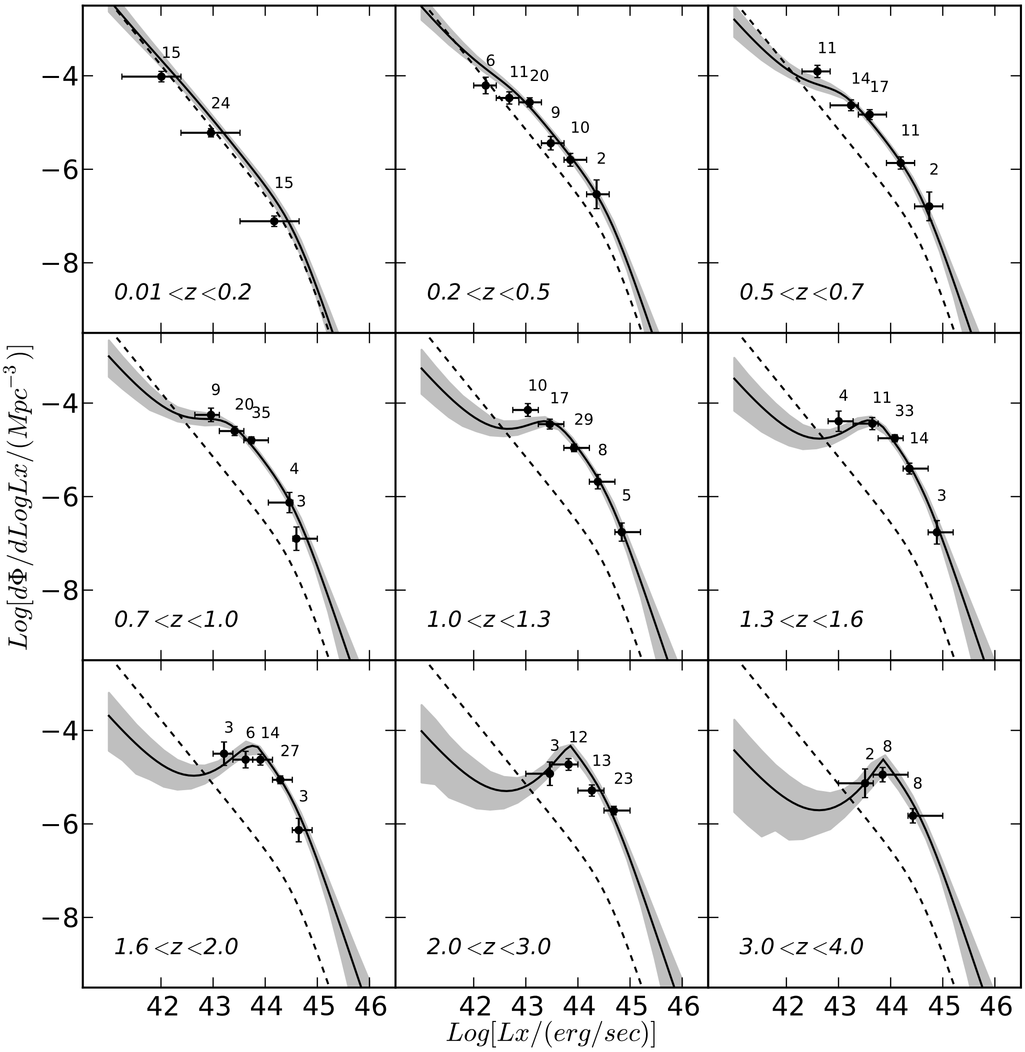


Figure 4.10: Differential luminosity function versus luminosity for different redshift bins. The number of sources in each bin is given on the top-right of each point.

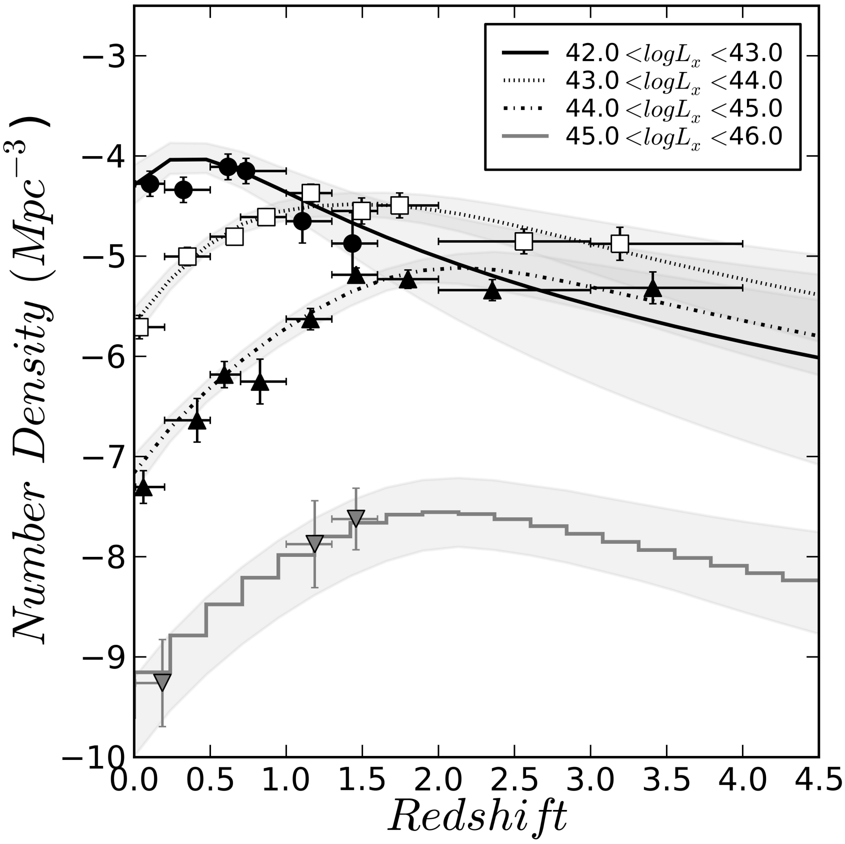


Figure 4.11: Number density as a function of redshift, for different luminosity bins. We find a strong decrease in number density for sources with $42 < \log \mathrm{Lx} < 43$ above $z \sim 0.7$.

affect the result. The other possibility for a sudden change in the dataset, is *Cosmic Variance*. The lowest luminosity bins at redshifts above 2, include only sources from our deepest available field, the CDFS. It is possible, that the objects in our sample are not representing the global AGN population properly, since they are observed in a very small area on the sky. The latter effect can be corrected if we populated this area of the luminosity - redshift plane better, with the inclusion of an equally deep field. The last possibility that might affect the result, is that at those redshift bins, the photometric redshift estimations are less accurate and therefore, there might be catastrophic outliers[9] present which bias our result, even though we used the full probability distribution of the photometric redshifts to avoid this possibility.

4.4.3 AGN number density

One of the important results of the estimation of the luminosity function is the determination of the number density of AGN across the cosmic time for given luminosity bins (Fig. 4.11). As previous works have shown (e.g. Miyaji et al., 2000, Ueda et al., 2003, Hasinger et al., 2005,

---

[9] Catastrophic outliers are called sources with wrong photometric redshift estimations. See also §3.1.2

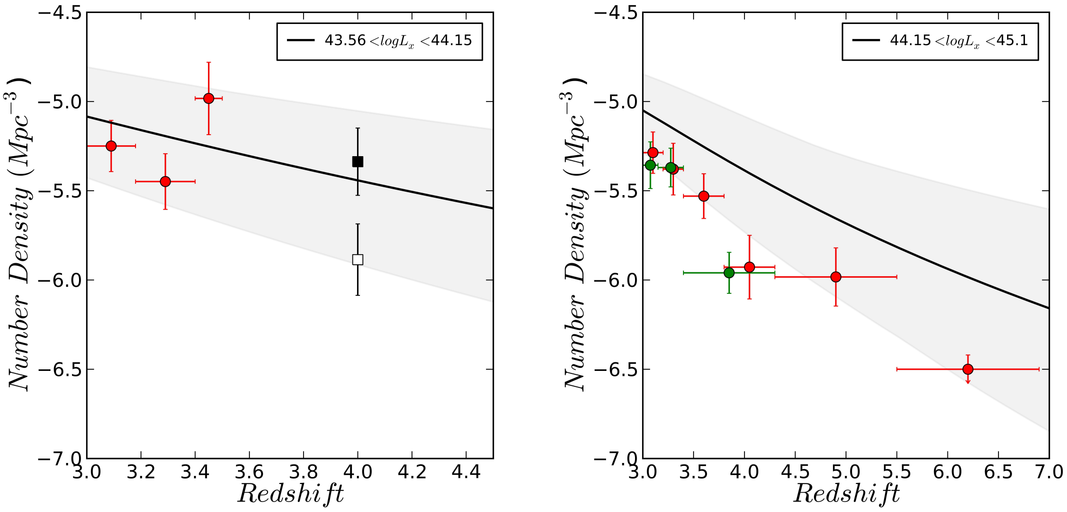


Figure 4.12: High-z comparison between number density of AGN in the 2-10 keV band from Civano et al. (2011) (points) and expected number density from our best fit in the 5-10 keV band, transformed to 2-10 keV (gray area, black line). Red points correspond to Chandra sources, green points to XMM sources (Brusa et al., 2009) (for details see Civano et al., 2011, points courtesy of F. Civano.)

Ebrero et al., 2009, Aird et al., 2010), we all also find that the number density of more luminous AGN peaks at earlier epochs, demonstrating the cosmic downsizing of supermassive black holes.

Additionally, we find a strong decrease in number density for sources with $42 < \log \mathrm{Lx} < 43$ above $z \sim 0.7$, in accordance with the modeling of quasar lights curves of Hopkins et al. (2006). Usually, hard selected samples do not include sources within this redshift range. These sources are introduced in our sample by the inclusion of the MAXI and HBSS sources. The only previous work that contains sources in this luminosity and redshift range is Yencho et al. (2009), which report a shallower decrease of the number density with redshift.

COSMOS high-z X-ray sample

Recent works on the highest redshift X-ray detected AGN (Brusa et al., 2009, Civano et al., 2011), demonstrate that there is a decrease in the number density of AGN beyond redshift 3 for $\log \mathrm{Lx} > 43$. This is behavior is present in our modeling as our sample extends up to redshift four, at luminosities up to $\log \mathrm{Lx} \sim 44$ (Fig. 4.12). Nevertheless, this is the area of uncertainty in our sample. As discussed in the previous section, this area might be biased due to Cosmic Variance. This result is worth investigating further in future work, with the inclusion of additional deep X-ray field and/or additional AGN selection criteria.

Extrapolating our result at higher luminosities and higher redshifts, there is an offset between the observed number density in the COSMOS field and our modeling but they agree within the uncertainties. As we will see in Chapter 5, the extrapolation of the luminosity function beyond its range of validity can lead to discrepancies and should be interpreted carefully.

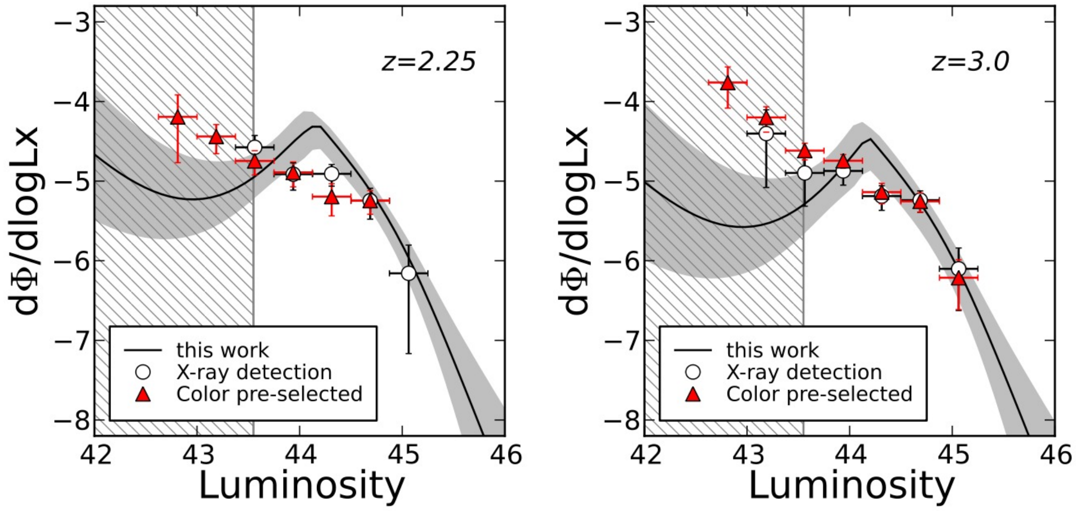


Figure 4.13: Comparison between the current fit presented in this chapter and the high-z pre-selected sample of Aird et al. (2010). The red triangles are the color selected sample used in Aird et al. (2010) to constrain the high-z AGN luminosity function. The open circles represent the X-ray detected sources that have not been used for their fitting. The solid line is the result from our fit, using the values from Table 4.4. The gray shaded area encompasses 99% credible interval. The hatched area for luminosities below $\log L_x < 43.6$ represents the flux limit of our sample at redshift ($z > 2$). Points courtesy of J. Aird.

High-z color pre-selected sample

In Aird et al. (2010), the authors try to estimate the 2-10 keV luminosity function from a collection of X-ray fields. The recognize that the photometric redshifts available to them, suffer from systematic problems. In order to identify a clean sample and constrain the high 2-10 keV luminosity function at high redshifts ($z > 2$), they create a pre-selection criterion based on optical colors which they then apply to the X-detected sample.

With this pre-selection, they aim to keep only objects in the sample that are truly at high redshift, incorporating at the same time their selection criterion into the modeling of the luminosity function. During their modeling, they tested the PLE, LADE, and LDDE evolutionary models. The evidence for the LDDE are only marginally higher that the simpler LADE and they conclude that there is no need for a more complicated evolution to be introduced.

In Figure 4.13 we give a comparison between the high redshift color pre-selected sample used in Aird et al. (2010) with our best model calculated at redshifts 2.25 and 3, respectively. The gray shaded area is the result of our modeling of the 5-10 keV luminosity function. The empty circles, denote the purely X-ray selected sample from Aird et al. (2010), while the red triangles are the color pre-selected X-ray sources which were used to constrain the evolution of the luminosity function at high redshift. The hatched area represents the flux limit of our sample. Focusing at the bright end of the luminosity function, we see that our result agrees well with their X-ray selected sample (empty circles), while small discrepancies are present comparing to the color pre-selected sample (red triangles). Even though, the break luminosity in our modeling shows a sharp edge introduced probably by Cosmic Variance, the binned estimates from Aird et al. (2010) still fall within the 99% credible interval of our result, above our flux limit. Below our flux limit, where we have to extrapolate our result, there is large disagreement which is due to the fact that our 5-10 keV selected sample, does not contain sources as faint as in the 2-10 keV and 0.2-5 keV X-ray bands, used to create the sample of Aird et al. (2010).

4.5 Conclusions

Using the most recent X-ray observations from MAXI, HBSS, COSMOS, LH and CDFS in the Ultra Hard X-ray band $5-10\,\mathrm{keV}$, we compile a sample of 499 AGN with 98% redshift completeness. Our sample contains 79% spectroscopic redshifts, while the rest 21% are very accurate photometric redshifts estimations from the fields COSMOS, LH and, CDFS. Studying this energy range, we avoid the potentially absorbed part of the spectrum for common N_H values, effectively avoiding any assumptions on otherwise necessary corrections to retrieve intrinsic X-ray luminosities.

Deriving the best fit parameters for the luminosity function using several different methods ($1/\mathrm{V}_{\mathrm{max}}$, Maximum Likelihood Estimation, Bayesian analysis), we exclude the possibility that the different results obtained by various authors are due to the chosen parameter estimation procedure. We select the best model using two best model selection methods, namely the Akaike Infromation Criterion (AIC) and Bayesian evidence.

Given the models considered here and the current dataset, both selection methods identify as the preferred model the Luminosity Dependent Density Evolution (LDDE). The preferred solution, demonstrates a decrease in AGN number density with redshift after the peak number density for each luminosity bin. This behavior is steeper to what was previously observed in luminosity function estimations in the literature at higher redshifts and it is similar to number density estimations in the COSMOS high-z sample. We find that given the large uncertainties in luminosity function estimations, and especially when the fit is extrapolated beyond its range of validity, our results is in reasonable agreement with previous estimations at the high redshift Universe.

In order to improve our estimation and lift possible biases, future work regarding our result, will include an additional deep X-ray field, namely the Chandra Deep Field North (CDFN). More general, future work on the luminosity function should most definitely include the uncertainty distribution of the sources both on the redshifts and X-ray fluxes, especially when photometric redshifts are used. We find the faint end of the luminosity function at high redshift in the $2-10\,\mathrm{keV}$ is still uncertain, even though efforts have been taken towards this estimation (i.e. Yencho et al., 2009, Brusa et al., 2009, Aird et al., 2008, 2010, Civano et al., 2011). Future surveys planned, such as the extended coverage of the COSMOS field with the Chandra X-ray observatory and of course the all sky survey by eROSITA, will provide an unprecedented number of sources thus allowing a better determination of the AGN X-ray luminosity function.

CHAPTER 5

$2 - 10$ keV *meta*-Luminosity Function

Estimating the X-ray Luminosity Function of Active Galactic Nuclei (AGN) has been a long standing goal for many decades. First observations revealed a strong change of the luminosity function of AGN up to some critical redshift value, z_c, after which the evolution has an elusive behavior. Nowadays, the dependence of the number density of AGN on both luminosity and redshift is generally accepted, but the details of the dependence remain under debate. There are many challenges associated to the estimation of a luminosity function which we encountered in the previous Chapter: i) unidentified X-ray sources, ii) incomplete and/or inaccurate redshift estimations which are particularly important at $1 < z < 2$ where the spectroscopic redshift determination is primarily possible from near infra-red observations while at the same redshift range photometric redshift estimations suffer from confusion due to degeneracies potentially leading to wrong redshift solutions, iii) flux absorption corrections and so forth. Usually, the most up to date samples contain fainter X-ray sources, thus reaching high redshift and low luminosity sources, and more accurate redshift estimations. Additionally, correction factors are estimated and applied to cure redshift incompleteness and intrinsic X-ray absorption.

In §4.2, we gave a detailed description of the models commonly used in the literature to describe the AGN X-ray luminosity function. Briefly, the *Pure Luminosity Evolution* (PLE, Schmidt, 1968) and *Pure Density Evolution* (PDE, Schmidt, 1968) are the simplest models considered, allowing for changes either with respect to luminosity or with respect to the normalization of the luminosity function. The more elaborate *Luminosity Dependent Density Evolution* (LDDE, Miyaji et al., 2000) allows for evolution in the normalization of the luminosity function, dependent on the luminosity. More recently, two models proposed namely the *Independent Luminosity and Density Evolution* (ILDE, Yencho et al., 2009) and the *Luminosity And Density Evolution* (LADE, Aird et al., 2010), incorporate a simultaneous evolution in luminosity and in normalization of the luminosity function. The difference between the two models is that the former does not include a critical redshift value z_c after which the normalization is changing, since it was used to describe low redshift AGN (for a graphical overview, see Fig. 4.6).

There are two main approaches that can be used to distinguish between these models when used to describe the same dataset. The first approach is to use model selection criteria, such as the AIC and Bayesian model selection discussed in §4.3. With the exception of Aird et al. (2010), all previous works in the literature use as model selection criteria 1-d and 2-d Kolmogorov-

---

The work presented in this Chapter has been done in close collaboration with Johannes Buchner at the Max Planck Institute for Extraterrestrial Physics.

Schmirnov tests between the data and the model. Nevertheless, a Kolmogorov-Schmirnov test cannot be applied in this situation, since the model parameters have been estimated from the same dataset to which the validity of the model is tested against. The second approach in distinguishing between two models, is to test the prediction power of each one in an area where the predictions give significantly different estimations. A common test of the parameter estimation is, under the assumption of an AGN population with some properties, to predict the Cosmic X-ray Background Radiation (CXRB). With this approach it has been shown that the CXRB can be explained up to 80-90% as the summed emission of AGN in the $0.5 - 2\,\mathrm{keV}$ energy band(Miyaji et al., 2000). Nevertheless, this estimation requires copious modeling and free parameters can be adjusted in order to reach the desired CXRB level.

With this work, instead of selecting the best model, we combine all luminosity function estimations to create a credible interval within which the true luminosity function must lie. First we show typical differences in the luminosity function models available in the literature and then discuss the range of validity of a luminosity function model (§5.1). In section §5.2 we demonstrate that the uncertainties on the errors of the models are correlated, and this fact reduces significantly the uncertainty estimation of the luminosity function. In section §5.3, we combine all the available luminosity function models from the literature to create the first *meta*-luminosity function, using each model explicitly within its range of validity. We show that this result is the most up-to-date estimation of the luminosity function, encompassing virtually all datasets used in the literature. This result serves also as a consensus of the scientific community on which is the luminosity function of AGN, since we treat all models as having the same chance of being the true one. In section §5.3.3 we compare the *meta*-luminosity function to the $5 - 10\,\mathrm{keV}$ luminosity function result obtained in Chapter 4. We show that the bright end agrees well with the unified luminosity function. Additionally, the change of the faint end slope which we discussed in §4.4.2, seems to be present in other works close to their flux limit which points to the conclusion that indeed this dramatic change in slope is an artifact possibly associated to sample incompleteness and/or Cosmic Variance. The Cosmic Variance could originate from the fact that the deepest X-ray observations typically cover only a small area on the sky ($< 0.2\,\mathrm{deg}^2$) which might introduce a non-representative population of faint AGN in the sample. We close this Chapter, discussing directions for future work regarding the *meta*-analysis in section §5.4.

5.1 Motivation

X-ray luminosity function estimations have been performed since the early X-ray telescopes have been launched to space (Einstein Observatory, Maccacaro et al., 1983). Ever since, improvements and extensions of the X-ray AGN datasets, motivated updated estimations of the X-ray luminosity function. A turning point in the luminosity function research area was reached with the advent of the ROSAT satellite (short for Röntgensatellit). In an influential paper, Miyaji, Hasinger and Schmidt (2000) using a sample of $\sim$700 AGN detected in the $0.5 - 2\,\mathrm{keV}$ energy band, showed that the faint end of the luminosity function significantly flattens at higher redshifts and motivated the description of the evolution of the luminosity function as *Luminosity Dependent Density Evolution* (LDDE).

Subsequent works using data from the X-ray observatories HEAO-1, ASCA, Chandra, and XMM-Newton, have tried to estimate the luminosity function in the $2 - 10\,\mathrm{keV}$ energy band. Table 5.1 gathers all the majors works published in the years 2003 - 2010, in chronological order. If an updated version of the luminosity function was performed by the same group of authors, we keep only their most recent work. Table 5.1 includes basic information for each sample such as the number of sources used, the bright and faint X-ray flux limits, the redshift range covered by the dataset, correction for redshift incompleteness and X-ray absorption. Additionally, we quote which luminosity function models were considered, and which X-ray fields were combined to created each dataset.

Even though the redshift range covered by the datasets remains largely the same ($0 < z < 3$), in more recent works, the number of sources and has increased not only due to deeper X-ray observations, but also due to extensive multiwavelength follow-up which enables the accurate determination of redshift also for fainter sources. The identified AGN up to redshift 3 now reach one order of magnitude fainter flux limit compared to when the first deep X-ray field was published (Chandra Deep Field North (CDFN), Brandt et al., 2001).

Even though large efforts have been concentrated to create bigger and better samples of X-ray AGN, the exact behavior of the luminosity function at high redshift remains under discussion. This is mainly due to the fact that the faint end of the luminosity function at high redshift is still scarcely populated. In Fig. 5.1, we plot in the upper panels, the differential luminosity function versus luminosity for three redshift bins, using only the point estimation of the model parameters. The difference in the faint end of the luminosity function is striking, even though in this figure we are plotting the same evolutionary model (LDDE). In the lower panels, we plot the same results, but this time we include uncertainty estimates (90% probability), using the quoted 1-σ uncertainty estimation of the parameters as provided by the authors. It is evident that the models give somewhat consistent results within the large uncertainties.

In Fig. 5.2, we give an overview of the LDDE model parameters obtained in the $2-10\,\mathrm{keV}$ band from the works quoted in Table 5.1. The black points are the estimated parameters with their associated uncertainties, while the empty squares denote parameters that were kept constant during the Maximum Likelihood Estimation (MLE). On the x-axis we place the reference of the work for each set of parameters, in increasing chronological order. Luminosity function estimations from the year 2003 (Ueda et al., 2003) up to 2009 (Ebrero et al., 2009), always kept at least some of the evolutionary parameters fixed in a pre-defined value. Only the latest works by Yencho et al. (2009), Aird et al. (2010) attempt to estimate all the parameters describing the evolution of the luminosity function. Since the parameters do not show so large deviations, it is virtually impossible to judge from the evolutionary parameters alone, the deviation of the faint of the luminosity function and attribute it to an 'erroneously' specified parameter.

As we already noted in §4.3.3, it is impossible to judge which model fits the dataset better, simply by comparing the MLE result to the $1/V_{max}$ by eye. To highlight this, in Fig. 5.1 we plot the $1/V_{max}$ result and the MLE model estimations[1] for the $5-10\,\mathrm{keV}$ luminosity function for two redshift bins. The empty circles denote bins which are populated by less than 5 AGN. The different lines show several luminosity function models tested. It is clear, that all models are reasonable representations of the binned estimates, within the luminosity - redshift region covered by the data. On the other hand, large discrepancies appear quickly when extrapolating the luminosity function. Thus, in the following, we define as *range of validity* of a luminosity function, the region of the luminosity - redshift plane covered by the respective dataset.

A *meta*-luminosity function

From Fig. 5.1, it is clear that the luminosity function estimations disagree mainly at the region of high redshift and low luminosity. This specific region is very hard to populate with a large number of sources due to current technical limitations, such as the long integration times needed to detect faint X-ray sources and with the additional challenge of determining accurately redshifts. Alternative selection techniques have been motivated in order to increase the

---

[1] This MLE includes only the fields HBSS, COSMOS, and Lockman Hole.

Table 5.1: $2 - 10$ keV Luminosity Function Estimates

	No. AGN	Flux range	Redshift range	Abs. Cor.	Incompl. Cor.	LF models	X-ray Fields
[A]	247	$10^{-10.0} - 10^{-14.42}$	0.0 – 3.0	yes	yes	PLE PDE LDDE	HEAO 1 A-2 HEAO 1 MC-CLASS AMSSn AMSSs ALSS ASCA Lockman ASCA Lynx CDFN
[B]	508	$10^{-10.1} - 10^{-15.0}$	0.0 – 4.0	yes	yes	PLE LDDE	HELLAS2XMM (1 deg^2) HELLAS2XMM (0.5 deg^2) HEAO 1 A-2 AMSSn HBS28 Lockman Hole CDFN CDFS
[C]	682	$10^{-11.5} - 10^{-15.2}$ $10^{-13.3} - 10^{-16.0}$	0.0 – 3.0 3.0 – 4.0	yes	yes	PLE mod-PLE LDDE	ChaMP CLASXS CDFN CDFS Lockman Hole (NW) AMSSn
[D]	435	$10^{-10.67} - 10^{-15.347}$	0.0 – 4.0	yes	yes	PLE LDDE	AMSS XMS CDFS
[E]	638	$10^{-12.29} - 10^{-15.93}$ $10^{-13.53} - 10^{-16.71}$	0.0 – 3.0 3.0 – 5.0	no	no	ILDE LDDE	ASCA CLASXS CLANS CDFN CDFS SWIFT/BAT[b]
[F]	852	$10^{-10.96} - 10^{-15.85}$	0.0 – 4.0	no	yes	PLE LADE LDDE	CDFS CDFN AEGIS-X ALSS AMSS

[a] [A] Ueda et al. (2003), [B] La Franca et al. (2005), [C] Silverman et al. (2008) [D] Ebrero et al. (2009), [E] Yencho et al. (2009), [F] Aird et al. (2010)

[b] 9 month sample

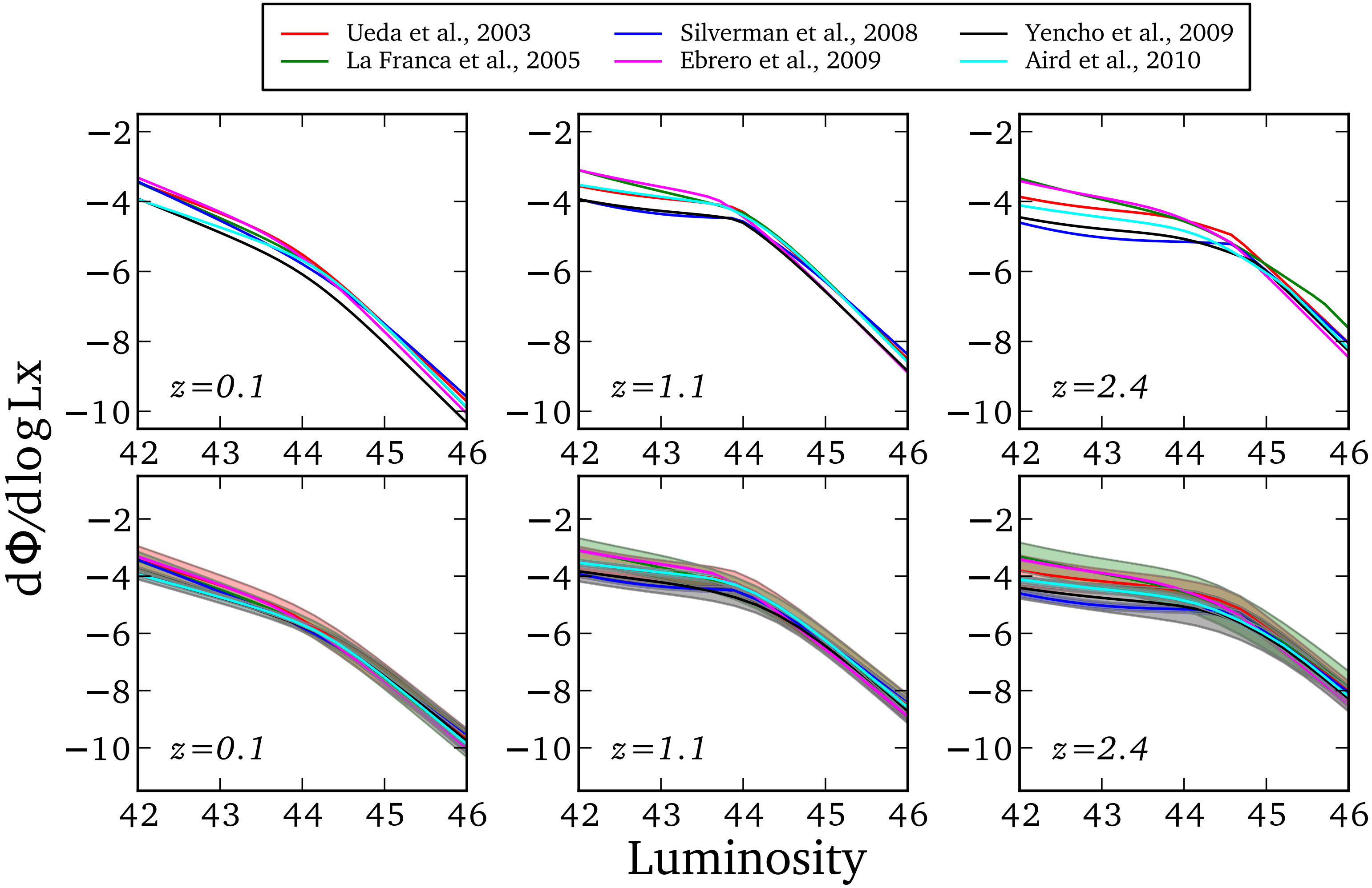


Figure 5.1: Comparison of 2-10 keV XLF estimates in three redshift bins. Upper panels: best fit estimates. Lower panels: best fit solution with uncertainty estimates. The bright end is generally in agreement among the models, while the faint end, particularly at higher redshifts is still highly uncertain.

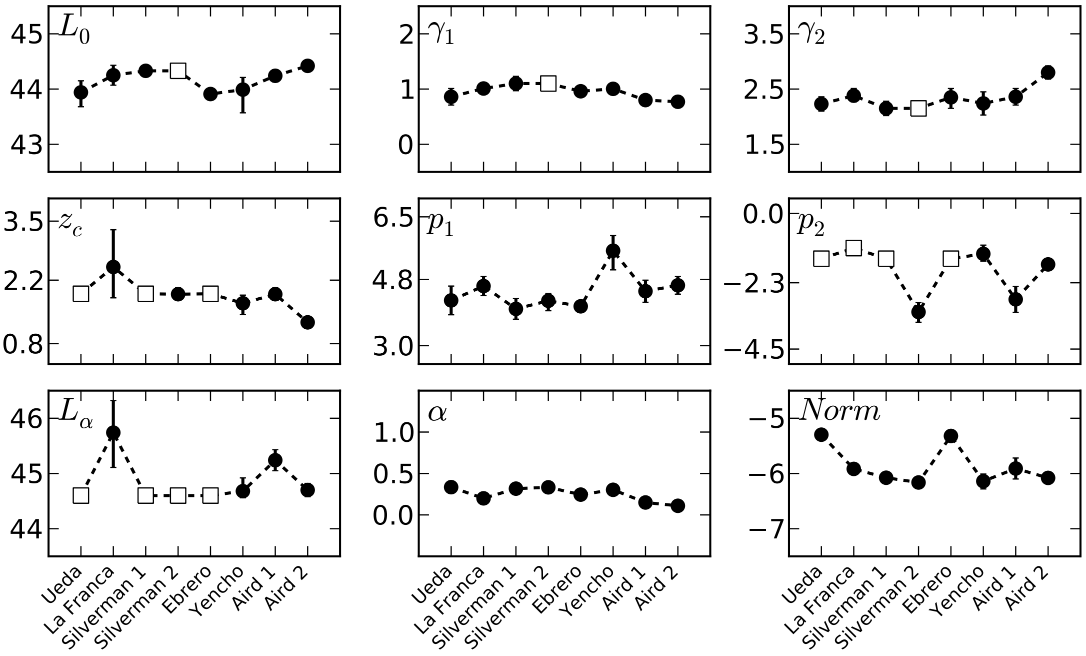


Figure 5.2: Best fit parameters for the LDDE models from Ueda et al. (2003), La Franca et al. (2005), Silverman et al. (2008), Ebrero et al. (2009), Yencho et al. (2009), Aird et al. (2010). The parameters fixed during the best parameter estimation are plotted with an empty square.

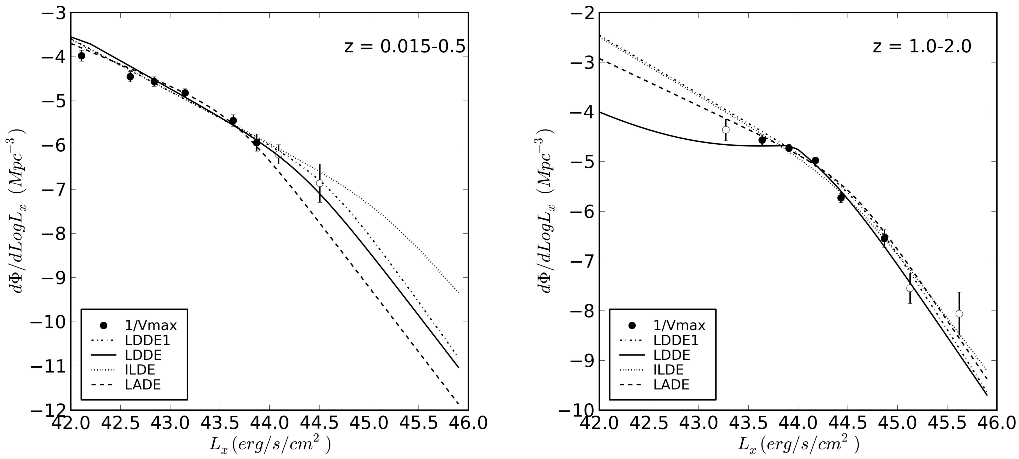


Figure 5.3: Comparison between LF evolutionary models. Discrepancies arise beyond the range of validity defined from the area covered by the data. The data shown here are from the fields HBSS, COSMOS and Lockman Hole, defined in 5-10 keV sample used in §4.

sample, and they have been used to constrain the faint end of the X-ray luminosity function at redshift ~ 3 (Aird et al., 2008, Brusa et al., 2009, Yencho et al., 2009, Aird et al., 2010, Civano et al., 2011) with some success. Nevertheless, even such interesting selections also suffer from biases and identification incompleteness which must be taken into account appropriately.

Since the luminosity functions in the literature appear to have discrepancies not only when different evolutionary scenaria are examined but also when the same evolutionary model is used, we motivate a different approach to determine the true luminosity function in a semi model-independent way. Under the premise that each model, within its range of validity, is capturing at least a glimpse of the true luminosity function we combine all the works present in Table 5.1, to create the first *meta*-luminosity function.

In good faith, we treat all models as being equally good, always within their range of validity. Even though there are overlapping X-ray fields used in the literature, they typically show differences in the association of the X-ray detected sources to their optical counterparts and the quality of the redshift determinations. For this first study, we will assume that all the works are independent and we will refer to future work for a more detailed analysis (see also the discussion in §5.4).

5.2 Parameter Correlations

In section §4.3 we discussed two parameter estimation methods, namely the Maximum Likelihood Estimation (MLE) and Bayesian analysis - realized through Markov Chain Monte Carlo (MCMC) in this work. Identifying the optimal model parameters that represent best the dataset at hand, reduces to the search of a minimum or a maximum of a function. In our case, it was the minimum of the data likelihood (eq. 4.18) or the maximum of the posterior (eq. 4.20). In the following we will refer only to a function minimum, but the same discussion applies also for maxima.

It is convenient to approximate the minimum of a function as gaussian distribution, at least in an interval very close to the minimum. In the case of an n-dimensional function, an n-dimensional gaussian is assumed:

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}|\mathbf{\Sigma}|^{\mathbf{1/2}}} e^{-\frac{1}{2}(x-\boldsymbol{\mu})^T \mathbf{\Sigma^{-1}}(\mathbf{x}-\boldsymbol{\mu})} \tag{5.1}$$

where $x = [x_1, \dots, x_n]$ a random vector of parameters, $\boldsymbol{\mu}$ the n-dimensional mean vector, $\mathbf{\Sigma}$ the covariance matrix and $|\mathbf{\Sigma}|$ its determinant. The covariance matrix is given by:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \rho_{n2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix} \tag{5.2}$$

where the square root of the diagonal elements gives the uncertainty estimates of the optimal parameters $\boldsymbol{\mu} = [\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_n]$ (identified through MLE or Bayesian analysis), while the off-diagonal elements give the correlations, ρ_{ij}, between the uncertainties of the parameters. Minimization routines, such as the package MINUIT - largely used to estimate the luminosity function parameters - often result to this approximation in order to compute the uncertainties of the estimated model parameters. Therefore, it is a reasonable assumption to accept that the luminosity function parameters of the literature follow normal distributions, with mean value the point estimate of the parameter and standard deviation, the quoted error estimation.

In Fig. 5.4 we show the 99% credible interval of the LDDE model of the luminosity function in two redshift bins. The credible interval has been computed using three approaches. The dark blue lines correspond to the 99% credible interval calculated using the covariance matrix of the model parameters. We created 10^5 random draws from a 9-dimensional[2] gaussian including the covariance matrix of the parameters. We then calculated the luminosity function for each set of parameters at a certain luminosity and redshift, and identified the area which enclosed 99% of the resulting distribution. We repeated this process until we covered the luminosity - redshift area under consideration. The black lines, correspond to the 99% credible interval, estimated treating the MCMC draws as a random sample of the luminosity function parameter distribution. Then we follow the same procedure as before calculating the corresponding values of the luminosity function until we cover the luminosity - redshift plane of our choice. The agreement between the two methods is remarkable. Lastly, we show the comparison to the 'naive' sampling of the luminosity function, assuming all the uncertainties of the parameters uncorrelated. Sampling 10^5 values from independent gaussians for each parameter, we calculated the credible interval as described above. The 99% credible interval of the luminosity function is overestimated by more than an order of magnitude.

It is interesting to note that, in the case of the MCMC, the correlations on the parameters uncertainties are naturally incorporated in the MCMC draws, lifting the need to supply a covariance matrix. Additionally, the uncertainty estimate mirrors the number of data present at a certain luminosity-redshift range. This means that the credible interval is narrower when a large number of data is available at a certain luminosity-redshift region, for example $44 < \log L_x < 44.5$ and $3.0 < z < 4.0$, and wider when extrapolating beyond the region covered by the dataset, for example at $\log L_x < 43.5$ and $3.0 < z < 4.0$. This effect is more prominent at high redshift bin and low luminosities, when the number of available sources is small.

5.2.1 Example Covariance Matrix

In Fig. 5.5 we provide a visualization of the covariance matrix (eq. 5.2), showing the marginal 2-dimensional joint distribution of parameter pairs for the LDDE model, computed using function 4.18. We notice a range in the behavior of the parameter correlations. From the tight correlation between the break luminosity, L_0, and the normalization, $Norm$, to the seemingly very loose correlation between the parameters p_1 and L_α.

The visualization of the covariance matrix, sheds light to the large differences observed in the estimated credible intervals in Fig. 5.4. The tight correlations reduce dramatically the uncertainties. Additionally, it is an argument to motivate the treatment of the normalization of the luminosity function as a model parameter, estimated during the MLE or Bayesian analysis. As we will discuss in the following section, it has been assumed in the literature that the normalization can be calculated after the best parameters have been estimated, by setting the integral of the luminosity function over luminosity and redshift, equal to the number of sources in the sample. Even though this approach will give a correct estimation of the normalization, the correlation of the parameters' uncertainties is lost and therefore the credible interval of the luminosity function is largely overestimated.

---

[2] The dimensions of the gaussian are equal to the model parameters, here we test the LDDE model which has 9 free parameters

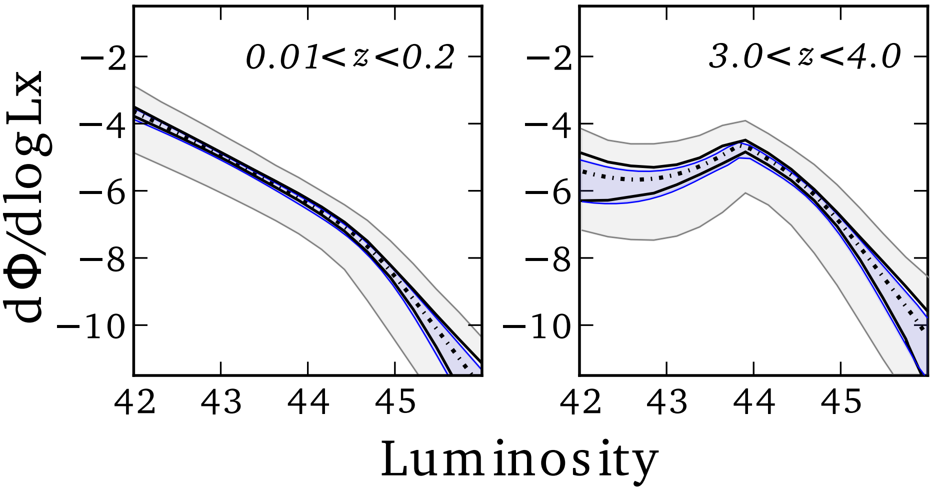


Figure 5.4: LDDE model for the $5-10\,\mathrm{keV}$ luminosity function computed in Chapter 4. Blue lines: 99% probability of the luminosity function from the MLE result. The covariance matrix of the parameters has been used to draw random values from the Luminosity Function. Black lines: 99% probability as defined from the Bayesian analysis. Using the visited points of the Markov Chain, the correlation of the parameters is taken into account. Light gray area: 99% probability of the luminosity function, using independent random draws on the parameters estimated from MLE. The uncertainties are assumed to the follow a gaussian distribution.

5.2.2 Scaling the Covariance Matrix

We have shown thus far, that the credible interval of a luminosity function model, encompasses significant information, since the interval is narrower within the range of validity of the dataset, and wider when extrapolating. We have also shown that in order to avoid overestimating largely this credible interval the covariance matrix of the models' parameters must be used.

Since the correlations of the uncertainty estimations have been largely ignored in the literature regarding the X-ray luminosity function, the covariance matrices are not published with the best parameter estimations for the respective models. We assume for simplicity that the correlations of the uncertainty estimations included in the covariance matrix are dependent primarily on the used model, and secondarily on the dataset. Therefore, in order to create credible intervals for the literature luminosity function estimations, we scale the covariance matrix obtained in Chapter 4.

As seen from eq. 5.2, the correlations among the parameters' uncertainties are the off-diagonal elements of the covariance matrix. They can be summarized in the *correlation matrix* given by:

$$\mathbf{P} = \sigma^{\mathbf{T}}\mathbf{\Sigma}\sigma = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{bmatrix} \tag{5.3}$$

where $\sigma = [\sigma_1, \dots, \sigma_n]$ the standard deviations of the parameters.

Since the majority of the authors use MLE and the algorithm suite MINUIT in order to define the best parameter estimation of their luminosity function, we are confident that the

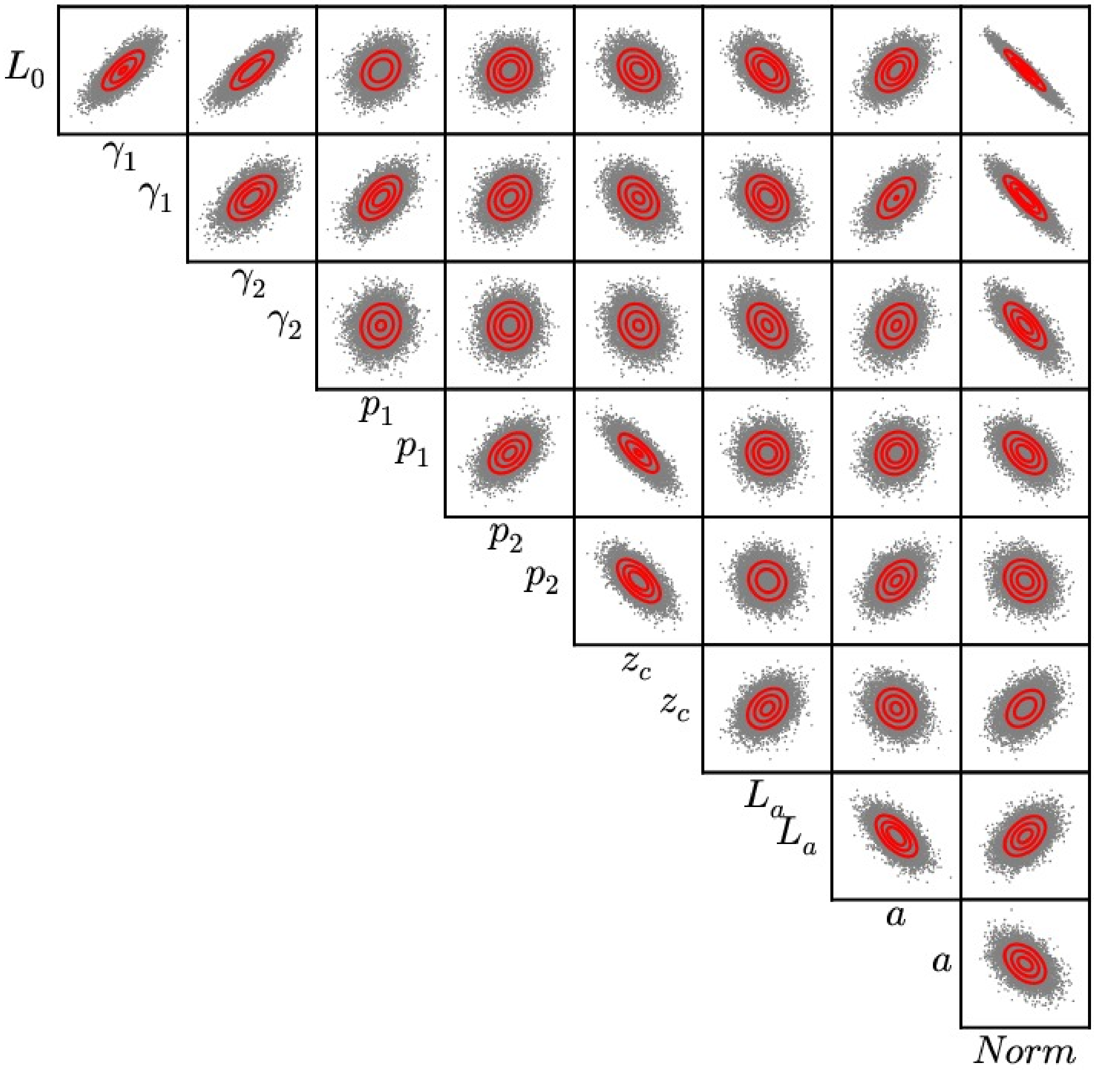


Figure 5.5: Covariance matrix visualization for the LDDE model. They gray points correspond to the random values of the LDDE parameters drawn as described in §5.3.1 while the red lines are isodensity lines, drawn as a guide to the eye. The tight correlation of the normalization and the parameters describing the local luminosity function (L_0, γ_1,γ_2) is prominent.

uncertainties on an individual parameter basis are properly estimated. Therefore, we use the following equation to scale the correlation matrix accordingly for each work:

$$\mathbf{\Sigma} = \boldsymbol{\sigma}^{\mathbf{T}} \mathbf{P} \boldsymbol{\sigma} \tag{5.4}$$

where $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_n]$ the uncertainties of the parameters quoted in each work.

As demonstrated in Fig. 5.2, in several cases the authors fix certain evolutionary parameters during the MLE. Roughly speaking, if we fix parameter "i" this corresponds to the removal of row "i" and column "i" from the covariance matrix. In order to have an accurate representation of the correlation matrix, we used our $5 - 10$ keV dataset and MINUIT to fit the models used in the literature fixing the corresponding parameters each time. We find that the new covariance matrix, after fixing some of the parameters, shows small deviations from the original matrix computed without fixing any of the parameters. After the appropriate scaling this procedure resulted in having a representative covariance matrix for each work.

5.3 2-10 keV *meta*-Luminosity Function

5.3.1 Sampling the Luminosity Functions

In order to create credible intervals for each luminosity function in the literature, we proceed as follows. We perform 10^5 random draws from a multidimensional gaussian (eq. 5.1), supplying the vector $\boldsymbol{\mu}$ comprised of the model parameter estimates from each work and the respective covariance matrix $\mathbf{\Sigma}$, scaled appropriately as described in 5.2.2. Whenever the authors state that the normalization is calculated after the estimation of the rest of the model parameters, we draw from the normalization independently assuming an one-dimensional gaussian function.

We then calculate the luminosity function for each of the 10^5 sets of model parameters in a luminosity (L) and redshift (z) 20×20 grid. We flag the (L, z) regions at which a luminosity function is outside its range of validity, both in terms of flux limits and in terms of redshift. Lastly, we calculate the mean and the standard deviation of the luminosity function at each (L,z), which is well approximated by a gaussian. In this way, we can define credible intervals at any (L,z) point at the desired level of significance.

5.3.2 The *meta*-luminosity function

We performed the aforementioned procedure for all the model estimations in the literature quoted in Tab. 5.1 (15 models in total). We then create a mean luminosity function for different model combinations[3].

In order to study differences in the evolution of the luminosity function, we first group the works according to the model when possible. In Fig. 5.6, we show the resulting evolution of the luminosity function (68% level) in four redshift bins. The red band, corresponds to the combination of all LDDE models (7 in total). The blue corresponds to the combination of all PLE models (6 in total). For the other models plotted, there is only one estimation in the literature. The green band shows the LADE model, the cyan band the PDE model, while the gray band shows the ILDE model.

A few conclusions can be drawn from this first comparison. All models show the same behavior having larger uncertainty with increasing redshift, as we would expect since there are less data at higher redshift, thus the models are less constrained. At the lowest redshift bin,

---

[3] The model combinations presented here were performed by Johannes Buchner at the Max Planck Institute for Extraterrestrial Physics.

all models agree remarkably well, with the exception of the bright end of the LADE model. The best agreement among all models appears at redshift ∼1. The critical redshift above which the evolution changes dramatically, is placed usually around redshift $1 < z < 2$. Therefore, beyond redshift 2 we expect the largest disagreement among the models. Fig. 5.6 shows that at redshifts 2.5 and 3.5, the uncertainties of the models are so large, that even at the 68% level the credible intervals of all models overlap.

In Fig. 5.7, we present the 90% (dark gray) and 99% (light gray) credible intervals of the $2-10$ keV *meta*-luminosity function computed from the combination of all the models present in Table 5.1. It is the most accurate estimation of the *true* AGN luminosity function at different redshift bins. We see that the uncertainties remain large even after the combination of all works, mirroring the discrepancies in the models.

This result, should be interpreted only as an interval estimation. According to the current knowledge of the luminosity function, given by the works included in this analysis, there is 90% probability that the true luminosity function will be enclosed in the dark gray shaded area of Fig. 5.7.

5.3.3 5-10 keV Comparison

In Fig. 5.8, we compare the $5-10$ keV luminosity function - transformed to the $2-10$ keV energy band - to the *meta*-luminosity function including only the models that correct for photoelectric absorption (excluding Yencho et al., 2009, Aird et al., 2010). We see that the 99% credible interval of the $5-10$ keV luminosity function is always enclosed in the 90% credible interval of $2-10$ keV result (with the exception of $\log \mathrm{Lx} > 45.5$).

We also notice that the flattening of the faint end of the $5-10$ keV luminosity function is probably due to cosmic variance (discussed also in §4.4.2). This is visible in the 'zig-zag' of the combined luminosity function, and in the per-model combination of Fig. 5.6. Every time a sample reaches its flux limit, the small number of sources in that particular luminosity and redshift region forces the model to underestimate the luminosity function. But, with the combination of other works, using deep X-ray data, this drop is cured. Notice that this drop is also present at the lowest luminosity of the $2.0 < z < 3.0$ luminosity function.

5.4 Future Prospects

In this Chapter we discussed the first *meta*-analysis performed on the AGN X-ray luminosity function. With this first study, we showed the feasibility of this project. More detailed accounts of the models, and the used works will be used in further studies with the hope to create a narrower credible interval.

Improvements of this work include the use of the luminosity function models, even in regions were they are extrapolated, essentially utilizing their predictive power. Such a task requires appropriate weighting during the combination of the models, since a model based on a dataset which covers that specific luminosity redshift region, must be a more accurate representation of the true luminosity function.

Assigning a weight on each model even within their range of validity, is motivated by the different number of AGN present in each sample and redshift completeness of the sample. An additional factor which could motivate a weighting factor, is the use of the same X-ray observations is several works, summarized in Fig. 5.9. Nevertheless, it is difficult to separate the X-ray fields within the luminosity function determination, since the quality of the optical

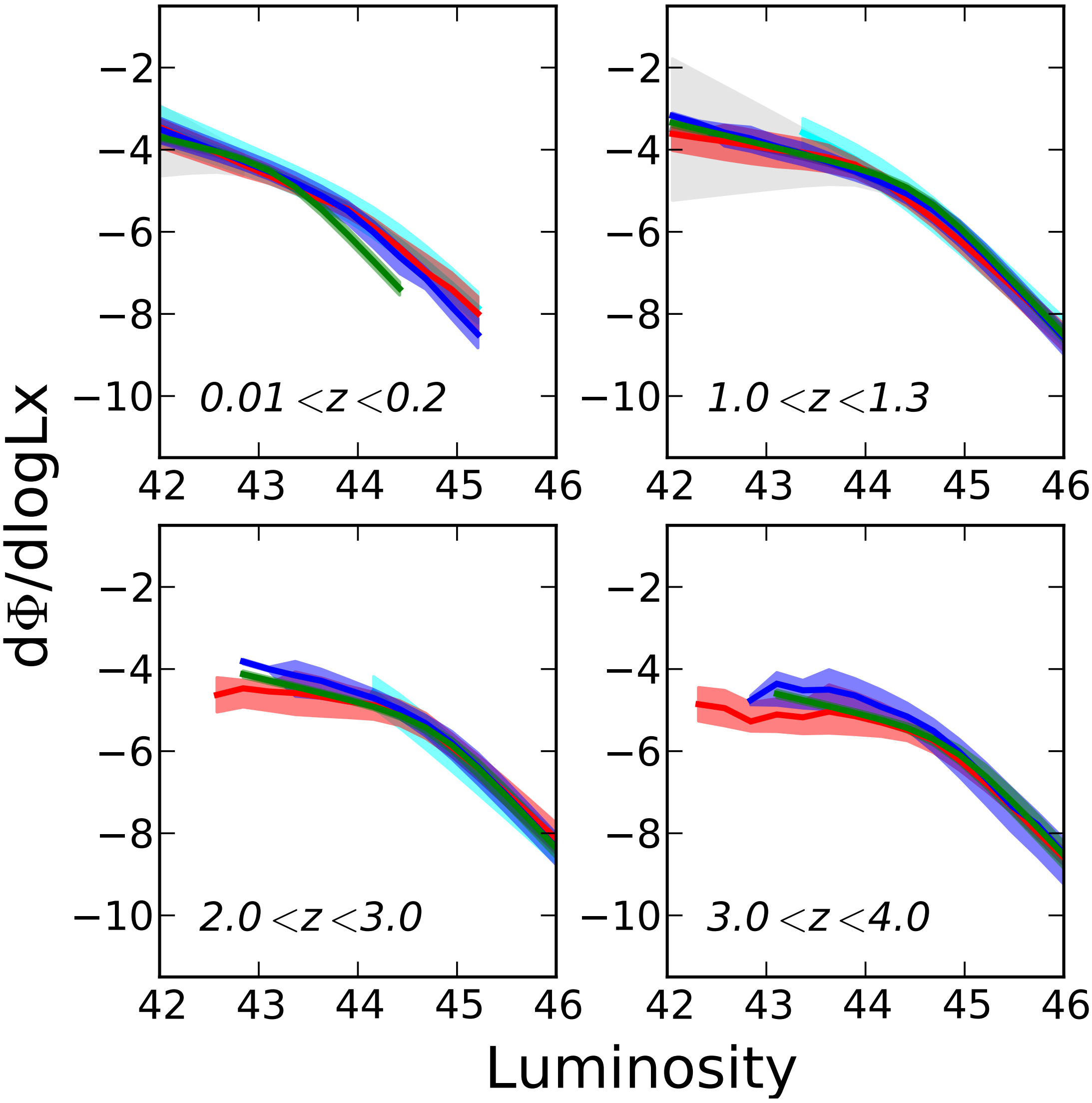


Figure 5.6: Comparison of evolutionary models, used only within their range of validity. Red area: LDDE meta-analysis, blue: PLE meta-analysis, green: LADE Aird et al. (2010), cyan: PDE Ueda et al. (2003), gray: ILDE Yencho et al. (2009). All models are shown with the corresponding 68% credible interval.

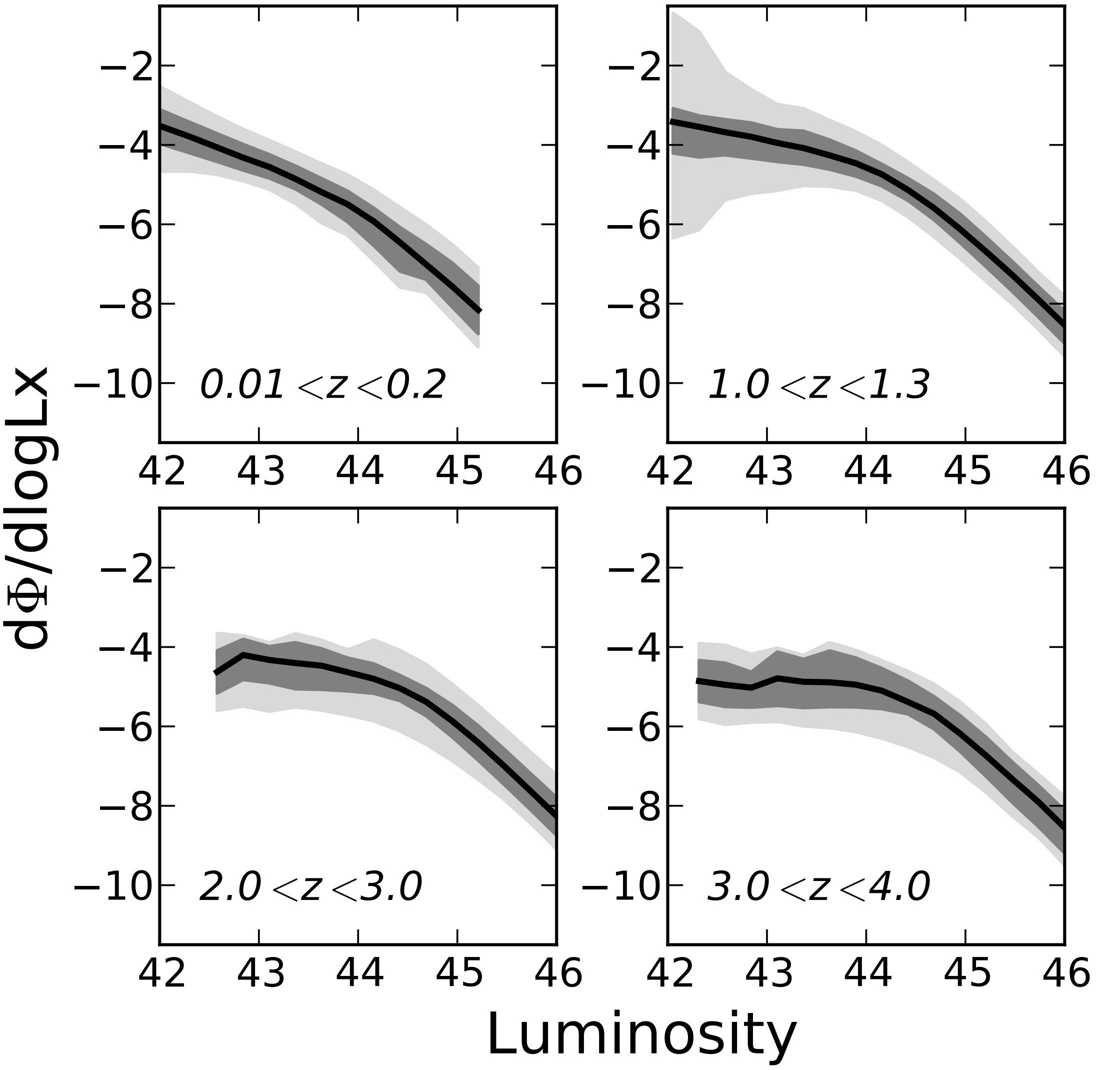


Figure 5.7: Meta Luminosity Function including all models available in the literature. Each model has been used within its range of validity. (dark gray 90%, light gray 99% credible intervals)

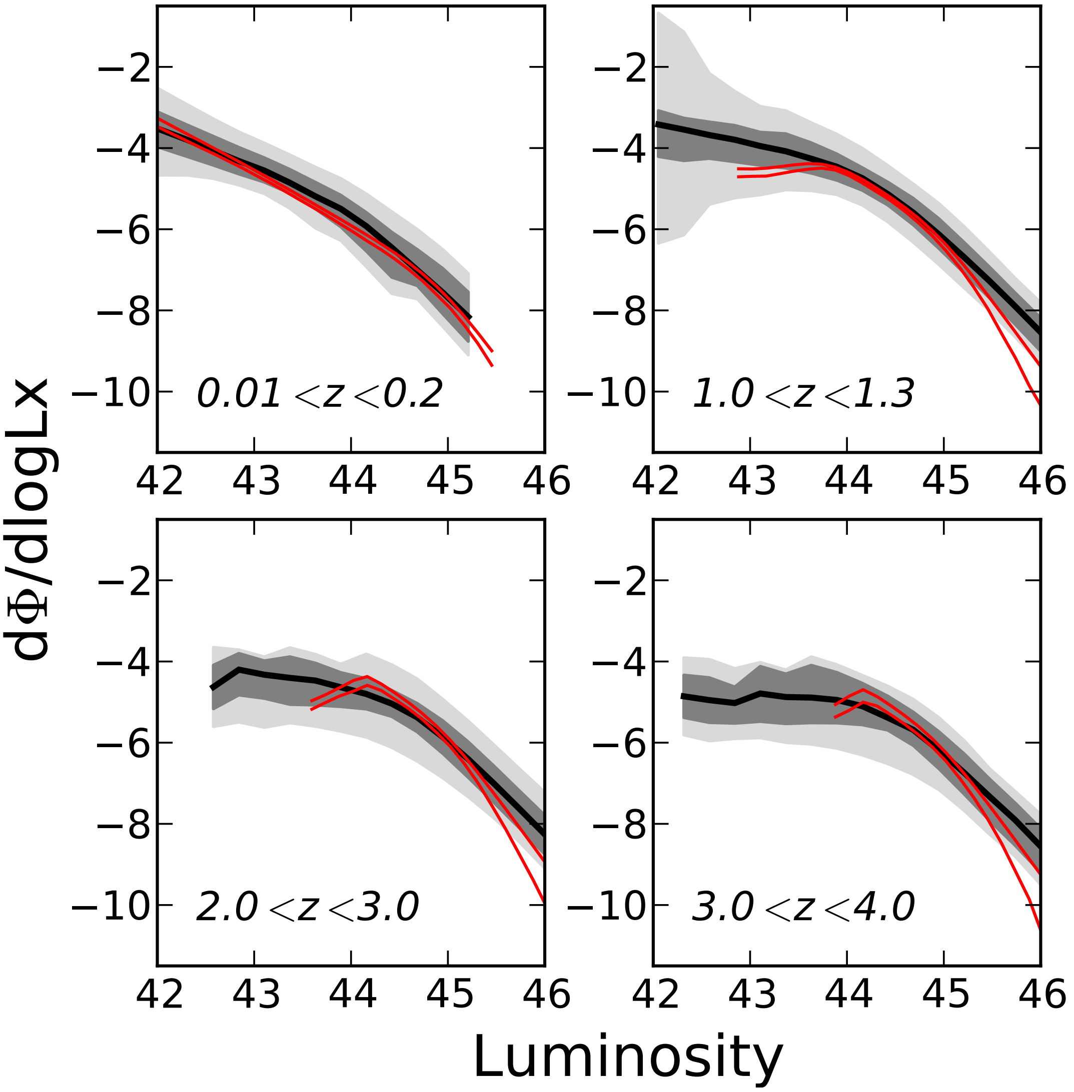


Figure 5.8: Gray shaded areas same as Figure 5.7. The red lines overplotted show the 5-10 keV result from the previous Chapter, transformed to the 2-10 keV. The two results are in good agreement.

counterparts and redshifts (both spectroscopic and photometric) is generally different from work to work and tends to be higher in more recent works.

5.5 Conclusions

In this Chapter, we showed that the AGN X-ray luminosity functions available in the literature show discrepancies, especially at the faint end of the luminosity function. This effect is present not only when comparing different evolutionary models, but also when comparing different determinations of the same evolutionary model. With the motivation to identify a credible interval which encloses the true luminosity function of AGN, we performed the first *meta*-analysis of the $2-10\,\mathrm{keV}$ AGN luminosity function models available in the literature in the years 2003 - 2010.

Our approach is based on the assumptions that:

1. Every model encompasses a glimpse of the true luminosity function of AGN.
2. All models in the literature are estimated equally well, within their range of validity.

where we define as range of validity, the luminosity - redshift region covered by the dataset used for the model estimation.

We show that the per-model unified estimations give overlapping predictions in the 68% credible interval. We motivate the use of this *meta*-luminosity function when an accurate representation of the AGN luminosity function is needed. The credible interval estimations of *meta*-luminosity function are available on request.

The approach presented in this Chapter can be easily updated as new luminosity function estimations will become available in the next years. In particular, the coming X-ray telescope *eROSITA* with an all sky survey in the $2-10\,\mathrm{keV}$ energy band, will provide us with the unique opportunity to determine the bright end of the luminosity function in local Universe, where the rarity of bright object still allows for statistical uncertainties in the models.

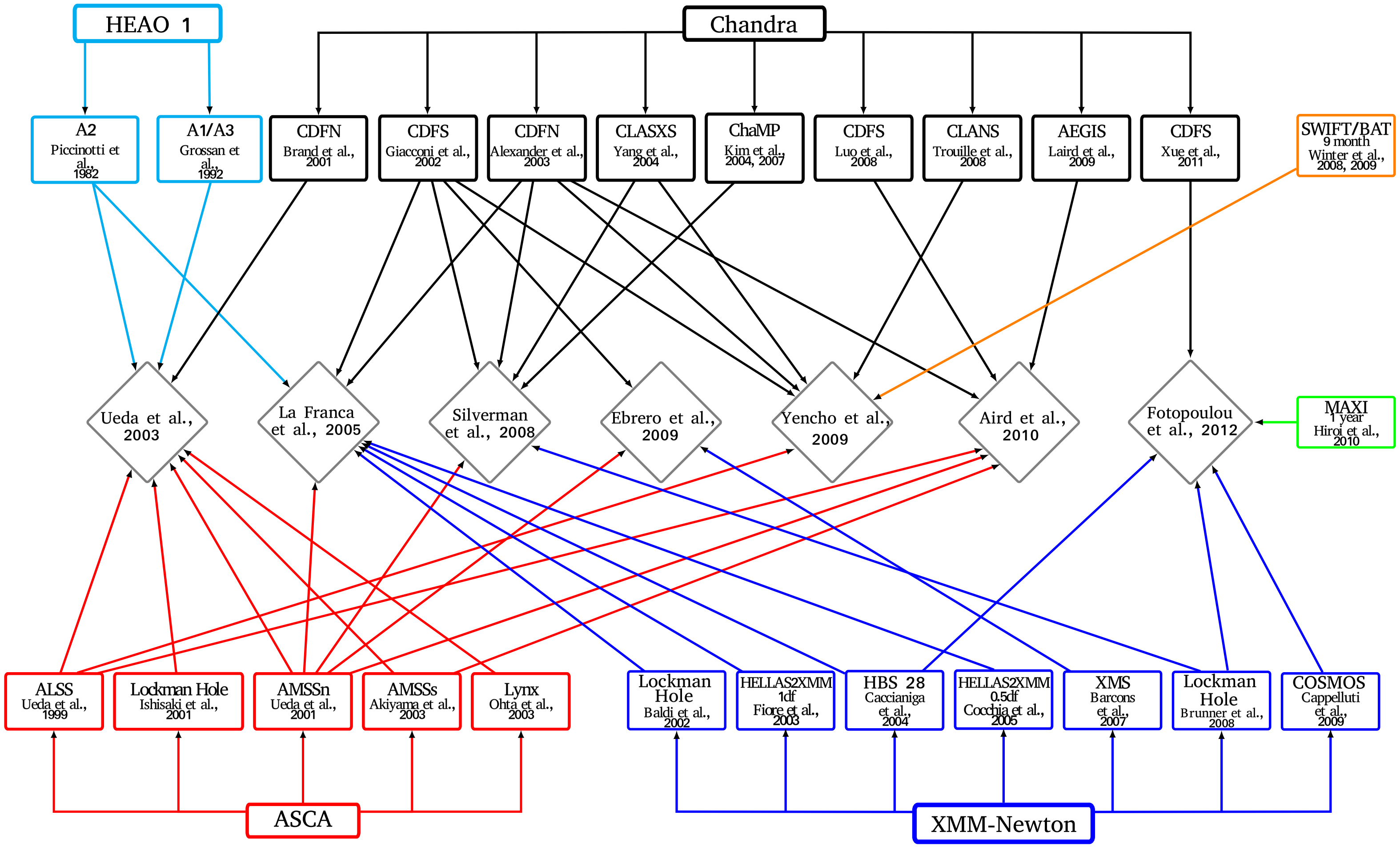


Figure 5.9: X-ray fields used in Luminosity Function estimations. Even though it appears as if the same field has been used several times, the accuracy of optical counterpart identification and redshift estimations have improved greatly with time.

CHAPTER 6

Epilogue

Recent observations suggest that Active Galactic Nuclei are an important phase in galaxy evolution. With the possibility that all galaxies undergo an active phase, this thesis was driven by the question of AGN evolution with time in the X-ray energy band. Evolutionary studies of an astrophysical population can only be addressed in a statistical manner. In order to create a representative sample of the AGN population, we have to combine extended but shallow X-ray fields, which contribute the brightest but rare sources in the sample with narrow but very deep fields, which reach the faintest sources known to date.

The currently available deepest X-ray field observed with XMM-Newton is the Lockman Hole deep field. A wealth of observations has been gathered over the years from the collective effort of dedicated astronomers. At the moment, imaging in more than 20 broad-band filters is available. Putting these efforts in good use, we collected, homogenized and analyzed the available photometric data on the Lockman Hole. This resulted in the first extensive photometric catalog for this area of the sky, ranging from the ultra-violet to the infra-red.

This catalog was then used to determine photometric redshifts for the normal galaxies in the field ($\sim$ 187000 sources). With a particular interest in AGN, we determined photometric redshift for X-ray selected sources (389 sources), following the combined criteria of morphology, variability and flux based source pre-selection procedure which has been shown to work well for the COSMOS field. The results from this work are a factor of two improvement over previously available photometric redshifts, both in accuracy and number of outliers, for normal galaxies and AGN alike. The photometric catalog and the corresponding photometric redshifts were published in Fotopoulou et al. (2012). Among other applications, the photometric redshifts aided in the discovery of one of the most distant X-ray detected galaxy clusters, presented in Henry et al. (2010).

Combining our results from the Lockman Hole field, with other X-rays fields such as MAXI, HBSS, COSMOS, CDFS we create a sizable sample of 500 AGN detected in the 5-10 keV band. The Lockman Hole field, bridges nicely the gap in flux between the 2deg^2 of COSMOS and the very deep CDFS, increasing the statistics of faint sources. The 5-10keV band, has the advantage of being largely unaffected by photoelectric absorption and we show that even though other works in lower energy ranges need to correct for this effect, our sample is free from this potential bias. Using a Poisson likelihood to describe the detection of AGN we perform Maximum Likelihood Estimation and Bayesian analysis to determine the best parameters for the luminosity function models. We test all evolutionary models available in the literature and we selected the preferred model according to the Akaike Information Criterion and Bayesian Model Selection. Both selection criteria identify the *Luminosity Dependent Density Evolution* as the best model for our dataset in accordance with previous observations in the Soft (0.5-2 keV) and Hard (2-10 keV) X-ray bands. Independent comparisons with previous works at high redshift, show that our modeling retrieves the decline in number density of AGN observed in the COSMOS survey, but it is possible that this agreement is driven by Cosmic Variance in our sample, as the deepest X-ray fields cover typically a small area ($< 0.2\,deg^2$) which might introduce a non-representative population of faint AGN in the sample.

A fair comparison between previous works in the 2-10 keV can only be achieved when the uncertainties in the model parameters are used. We show that the uncertainties are correlated and that the covariance matrix is of paramount importance when using the best parameter estimations. Especially we point out that the normalization of the luminosity function should be included as model parameter during the parameter estimation procedure, because the uncertainties between the normalization and the break luminosity show strong dependence. We find that luminosity functions should not be extrapolated far from their range of validity identified by the data coverage of the luminosity redshift plane. In order to incorporate all previous works into one single result which encloses the *true* luminosity function, we combine the all model estimates to create the first AGN *meta*-luminosity function. We show that with in the uncertainties the previous works are consistent. This meta-luminosity function is the best representation of the *true* luminosity function of AGN and its evolution since it incorporates all the models used by the various authors and should be used whenever an accurate estimation of the AGN luminosity function is needed.

The future of astrophysics at the moment seems very promising. Bigger and better instruments are currently scheduled to become available to the community in the following years. Addressing specifically the Lockman Hole Field, the current needs I identify are primarily a large spectroscopic program, deep near infra-red photometric observations (at the moment only shallow J and K bands are available) as well as intermediate- and narrow-band optical photometry. This will aid in determining photometric redshifts with grater accuracy, and also enable studies of the physical parameters such as star-formation rate and galaxy mass.

Connecting the AGN evolution to the galaxy-black hole coevolution problem, I believe that the effect of the AGN activity can be seen mainly in two areas. First, in the greater environment of the galaxy, where the energetic output of the jet created by the AGN has been found to shape the gas properties. The second area is very close to the black hole, studying the dynamics of stars and gas in the vicinity of the black hole. Even though the first case is relatively easy to observe, the number of AGN with jet is very low hampering the statistical interpretation. On the contrary, even if we accept that all galaxies with a bulge host a massive black hole in their centers, it is currently a technically challenging task to detect black holes based on the effect they have on the environment in their vicinity, let alone to study their evolution with time. This leaves us with the third pillar of modern astronomy which we did not address during this thesis, simulations. Numerical simulations have proven to be a powerful approach to understand nature. Identifying the proper AGN feedback recipe is a two way process between simulations and observations. This problem is closely linked to galaxy and AGN evolution, clearly mapped on the luminosity function and it is a promising way to understand the physical procedure responsible for the observed evolution of the AGN number density with time.

Therefore, a detailed study of the X-ray AGN luminosity function is currently a very promising approach in understanding the AGN phenomenon and linking it to the fate of galaxies.

Concerning the AGN evolution, it is my personal belief that the two pillars addressed in this work should be further expanded. That is the creation of the AGN sample and the statistical methods used to determine the evolution. As we showed during the meta-analysis, the differences in the literature arise in the low flux limit. Even though the upcoming X-ray satellite, eROSITA, will provide millions of AGN as a true successor of the ROSAT satellite, they will all be bright sources giving a definite answer on the bright end of the luminosity function. At the moment, there is no program scheduled to detect even fainter sources in numbers large enough to increase the statistics in the range of dispute. I believe that the Lockman Hole is an ideal target to reach ever fainter flux limit in combination to the already deep X-ray observations.

On the other hand, using elaborate analysis methods, we can still learn valuable lessons from the existing observations. My personal view of the way forward in the determination of the evolution of the luminosity function is using model independent methods, which require no binning and are going to reveal the true shape of the luminosity function. Only then we can discuss with certainty about which is the physical interpretation of the evolution of the luminosity function which we are still lacking.

APPENDIX A

Model Estimation

In the following, we give first the marginal distribution of the model parameters and then differential luminosity function for several redshift bins for all models tested in this work (see Chapter 4). The black points are the result from the $1/V_{max}$ method. The dashed black line is the luminosity function computed at redshift z=0, using the respective best fit parameters. The solid black line is plotted using the best fit parameters at the median redshift of the dataset in each redshift bin, while the gray shaded area is the 99% probability interval of the differential luminosity function, $d\phi/dlogL_x$. To determine this area, we compute $d\phi/dlogL_x$ in each redshift bin for 40 values of the luminosity, L_x, for all the 10^5 draws from the posterior. In this way, we incorporate naturally the true shape of the uncertainties for all parameters.

All models use a double power law distribution to describe the local luminosity function:

$$\frac{d\phi(L, z=0)}{d\log L_x} = \frac{A}{(\frac{L}{L_0})^{\gamma_1} + (\frac{L}{L_0})^{\gamma_2}} \tag{A.1}$$

For each model we present the formula which describes the evolution, a table with summary information on the best parameters determined by the MCMC and a table with the results from the maximum likelihood estimation and the covariance matrix. The visited points of the posterior needed for further statistical analysis are available electronically upon request.

A.1 Pure Luminosity Evolution - PLE

The Pure Luminosity Evolution (PLE), describes differences only in the luminosity of the objects, and not in their number. The evolution of the luminosity function is given by:

$$\frac{d\phi(L, z)}{d\log L_x} = \frac{d\phi(L/e(z), z=0)}{d\log L_x} \tag{A.2}$$

with,

$$e(z) = \begin{cases} (1+z)^{p_1} & z \leq z_c \\ (1+z_c)^{p_1} \cdot (\frac{1+z}{1+z_c})^{p_2} & z \geq z_c \end{cases} \tag{A.3}$$

The most appropriate parameters for the current dataset are given in Tables A.2, and A.1.

Table A.1: PLE – MCMC Parameter Estimation

Parameter	Prior Interval	Mean	Std	Min – Max
L_0	40.0 — 46.0	42.97	0.11	42.63 — 43.28
γ_1	-1.0 — 3.0	0.12	0.11	-0.31 — 0.47
γ_2	1.0 — 6.0	2.22	0.11	1.88 — 2.71
p_1	1.0 — 7.0	2.74	0.15	2.19 — 3.33
p_2	-6.0 — 1.0	-0.22	0.34	-1.77 — 0.88
z_c	0.5 — 3.5	1.55	0.11	1.11 — 2.07
A	-9.0 — -2.0	-4.51	0.11	-4.86 — -4.17

Table A.2: PLE – MLE Parameter Estimation

Maximum Likehihood Parameters

L_0	γ_1	γ_2	p_1	p_2	z_c	A
43.0	0.13	2.2	2.7	-0.15	1.5	-4.5

Covariance Matrix

	L_0	γ_1	γ_2	p_1	p_2	z_c	A
L_0	0.010	0.009	0.009	-0.008	0.0018	0.0000014	-0.010
γ_1	—	0.012	0.007	-0.004	-0.0023	-0.0000022	-0.010
γ_2	—	—	0.012	-0.006	0.0024	-0.0000006	-0.008
p_1	—	—	—	0.016	-0.014	-0.000024	0.005
p_2	—	—	—	—	0.07	-0.00007	0.0010
z_c	—	—	—	—	—	0.00004	0.0000023
A	—	—	—	—	—	—	0.010

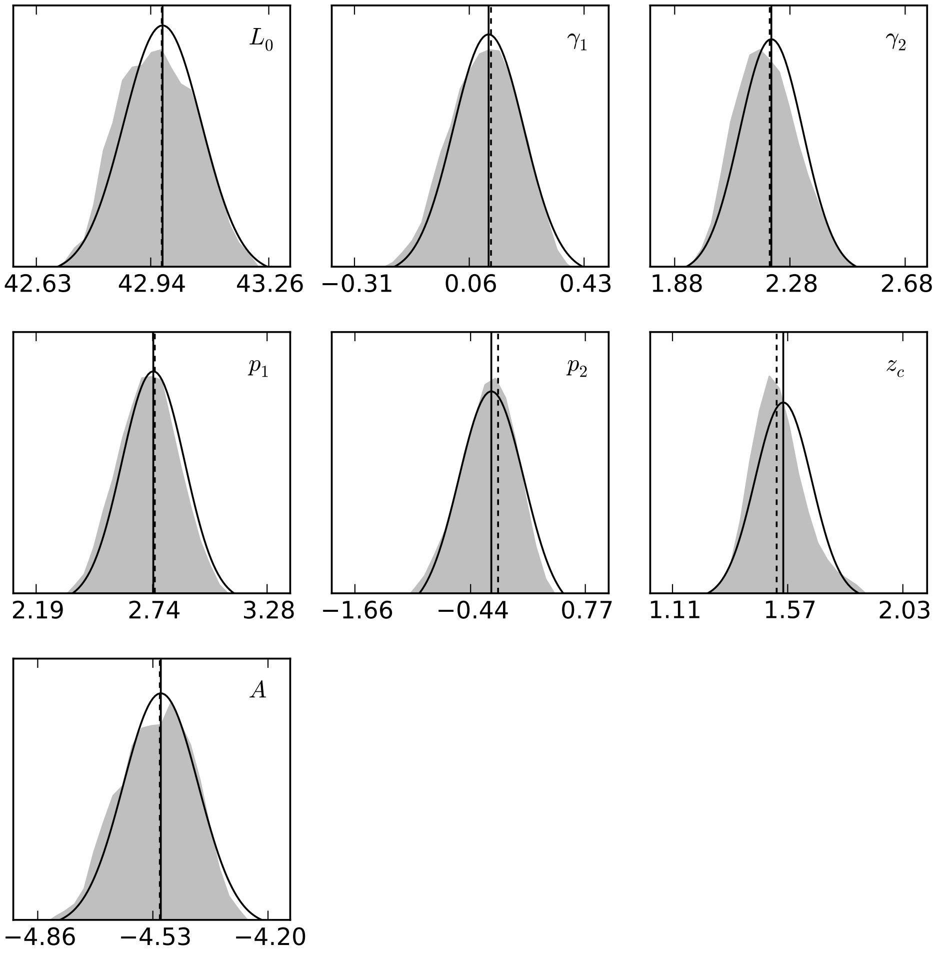


Figure A.1: Marginals for Pure Luminosity Evolution (PLE). The gray area encloses the 99% probability area. The bell-shaped black line, shows the respective gaussian for the mean and standatd deviation in Table A.1. The vertical solid line shows the mean value from the MCMC result, while the dashed line shows the MLE result (Table A.2, see also §4.3.3).

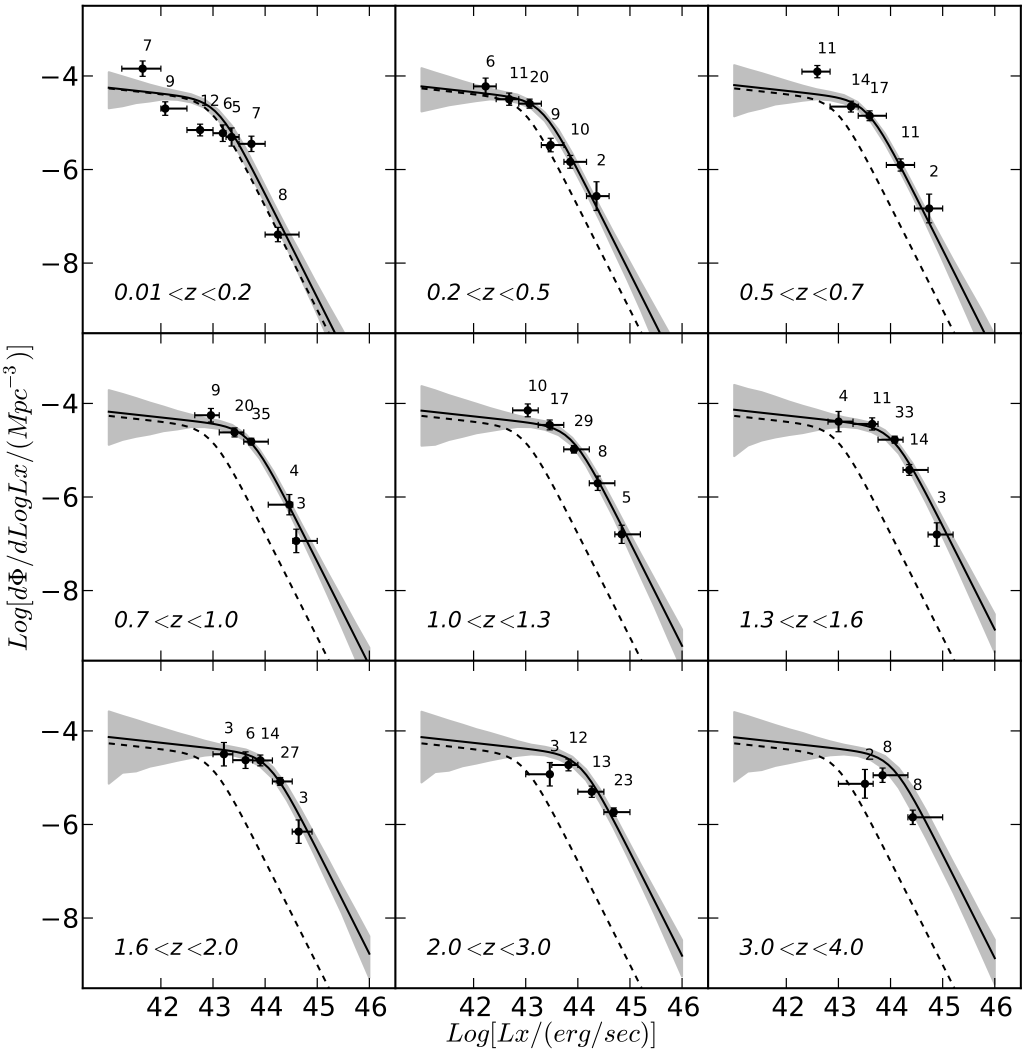


Figure A.2: Differential Luminosity Function for Pure Luminosity Evolution (PLE). The gray area encloses 99% of the probability. The dashed line is the extrapolated luminosity function at z=0. The solid line is computed using the mean values from Table A.1. The points are the result of the $1/V_{max}$ method. The numbers above each point indicate the number of AGN included in each bin. Lack of points significies extrapolation of the best solution.

A.2 Pure Density Evolution - PDE

The Pure Density Evolution (PDE), describes differences only in the number density of the objects, and not in their luminosity. The evolution of the luminosity function is given by:

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L,z=0)}{d\log L_x} \cdot e(z) \tag{A.4}$$

with, e(z) given by eq. A.3 The best parameters for the current dataset are given in Tables A.4, and A.3.

Table A.3: PDE – MCMC Parameter Estimation

Parameter	Prior Interval	Mean	Std	Min – Max
L_0	40.0 — 46.0	44.25	0.07	44.04 — 44.51
γ_1	-1.0 — 3.0	0.62	0.05	0.42 — 0.84
γ_2	1.0 — 6.0	2.92	0.2	2.32 — 3.91
p_1	2.0 — 7.0	3.78	0.33	2.85 — 4.98
p_2	-6.0 — 2.0	-0.42	0.53	-2.76 — 1.07
z_c	0.5 — 3.5	1.22	0.18	0.65 — 1.88
A	-9.0 — -4.0	-6.2	0.12	-6.68 — -5.83

Table A.4: PDE – MLE Parameter Estimation

Maximum Likehihood Parameters

L_0	γ_1	γ_2	p_1	p_2	z_c	A
44.26	0.63	2.92	4.1	0.03	1.03	-6.2

Covariance Matrix

	L_0	γ_1	γ_2	p_1	p_2	z_c	A
L_0	0.004	0.0027	0.010	-0.00014	0.00017	-0.0000047	-0.0064
γ_1	—	0.0027	0.0053	-0.00078	0.00016	-0.0000025	-0.0044
γ_2	—	—	0.039	-0.0018	0.0013	-0.000012	-0.0135
p_1	—	—	—	0.047	-0.024	-0.00015	-0.008
p_2	—	—	—	—	0.053	-0.00014	0.002
z_c	—	—	—	—	—	0.00007	0.00002
A	—	—	—	—	—	—	0.012

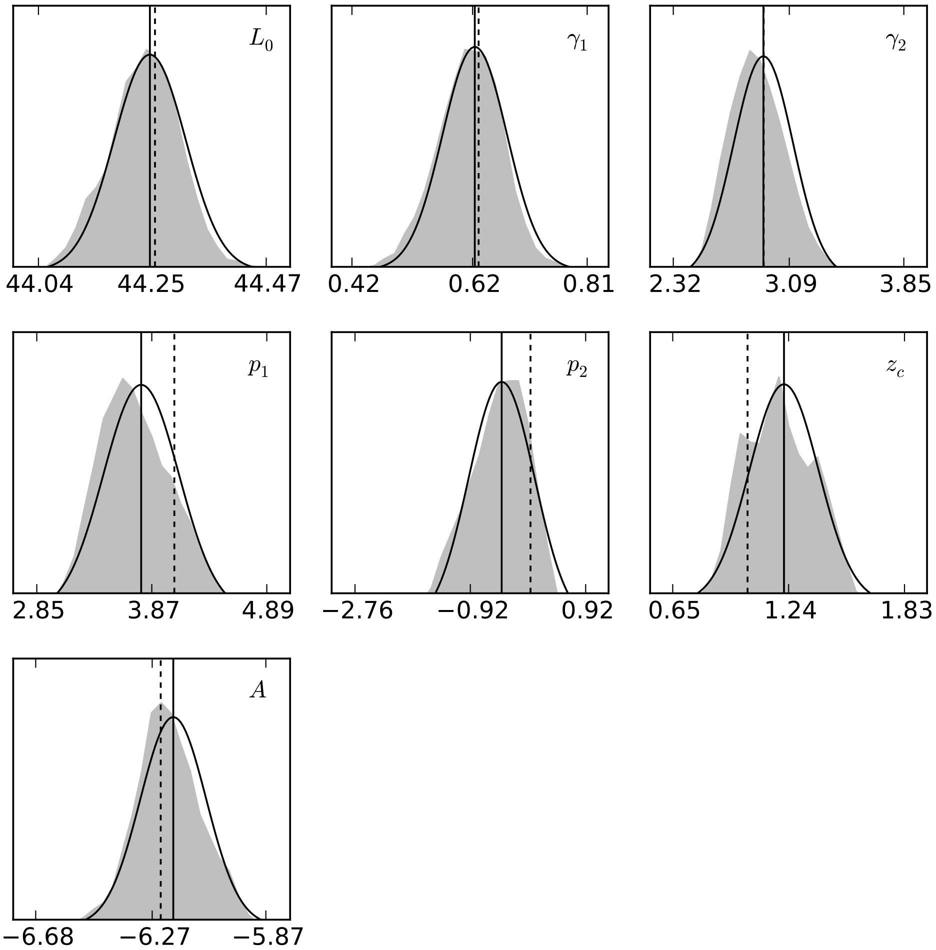


Figure A.3: Marginals for Pure Density Evolution (PDE). The gray area encloses the 99% probability area. The bell-shaped black line, shows the respective gaussian for the mean and standatd deviation in Table A.3. The vertical solid line shows the mean value from the MCMC result, while the dashed line shows the MLE result (Table A.4).

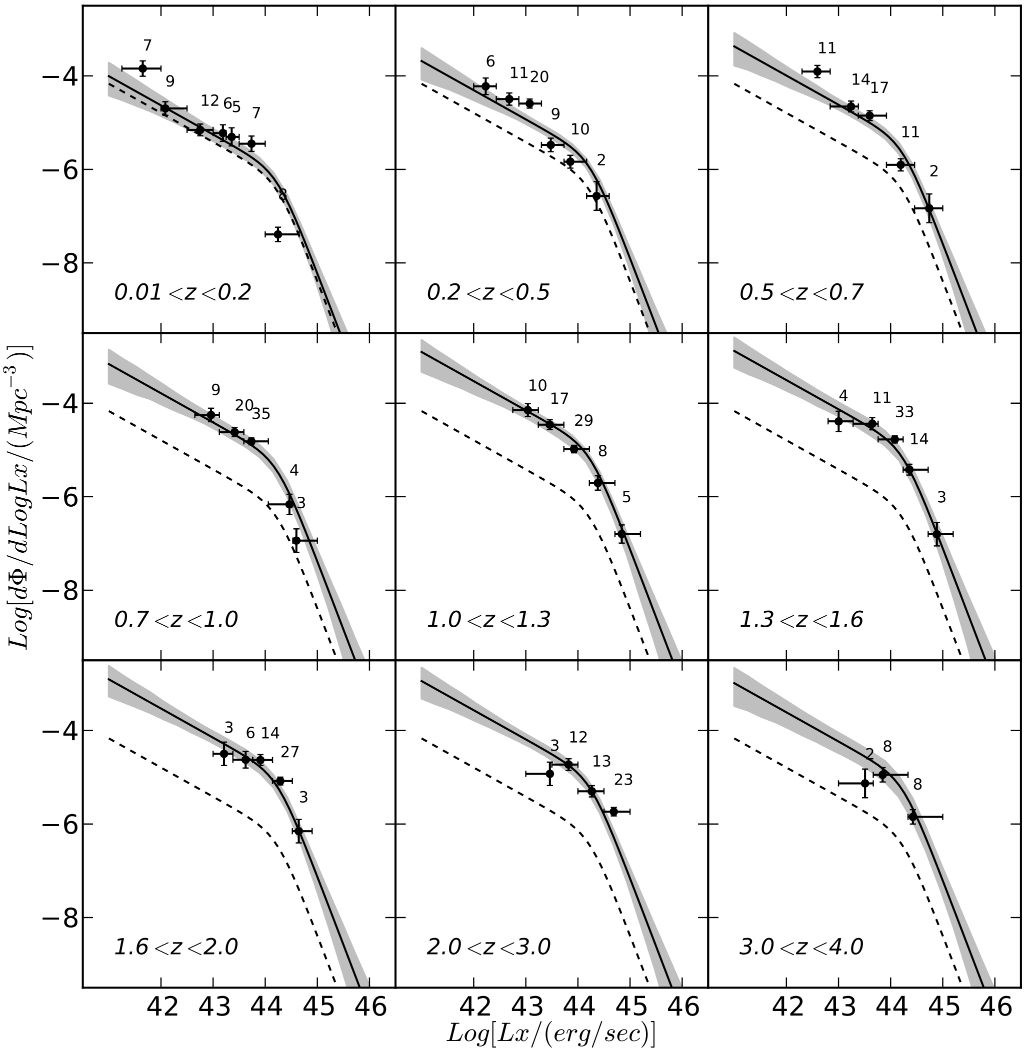


Figure A.4: Differential Luminosity Function for Pure Density Evolution (PDE). The gray area encloses 99% of the probability. The dashed line is the extrapolated luminosity function at z=0. The solid line is computed using the mean values from Table A.3. The points are the result of the $1/V_{max}$ method. Lack of points significies extrapolation of the best solution.

A.3 Independent Luminosity and Density Evolution - ILDE

As the name of the model suggests, the evolution of the luminosity function is free to change both in luminosity and redshift. A point of notice is that this model does not assume a characteristic redshift after which the evolution changes. The model is :

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L/e(z), z=0)}{d\log L_x} \tag{A.5}$$

with,

$$e(z) = \begin{cases} (1+z)^{p_1} & z \leq z_c \\ (1+z_c)^{p_1} \cdot (\frac{1+z}{1+z_c})^{p_2} & z \geq z_c \end{cases} \tag{A.6}$$

The most appropriate parameters for the current dataset are given in Tables A.6, and A.5

Table A.5: ILDE – MCMC Parameter Estimation

Parameter	Prior Interval	Mean	Std	Min – Max
L_0	40.0 – 46.0	42.83	0.12	42.47 – 43.21
γ_1	-1.0 – 3.0	0.09	0.11	-0.35 – 0.5
γ_2	0.0 – 7.0	2.17	0.11	1.85 – 2.63
p_1	0.0 – 7.0	2.73	0.19	2.06 – 3.45
p_2	-6.0 – 1.0	-1.2	0.25	-2.07 – -0.38
A	-7.0 – -2.0	-4.1	0.14	-4.55 – -3.68

Table A.6: ILDE – MLE Parameter Estimation

Maximum Likehihood Parameters

	L_0	γ_1	γ_2	p_1	p_2	A
	42.8	0.08	2.15	2.73	-1.20	-4.08

Covariance Matrix

	L_0	γ_1	γ_2	p_1	p_2	A
L_0	0.015	0.009	0.010	-0.019	0.022	-0.016
γ_1	—	0.011	0.006	-0.0045	0.003	-0.010
γ_2	—	—	0.010	-0.010	0.011	-0.011
p_1	—	—	—	0.040	-0.049	0.022
p_2	—	—	—	—	0.068	-0.028
A	—	—	—	—	—	0.020

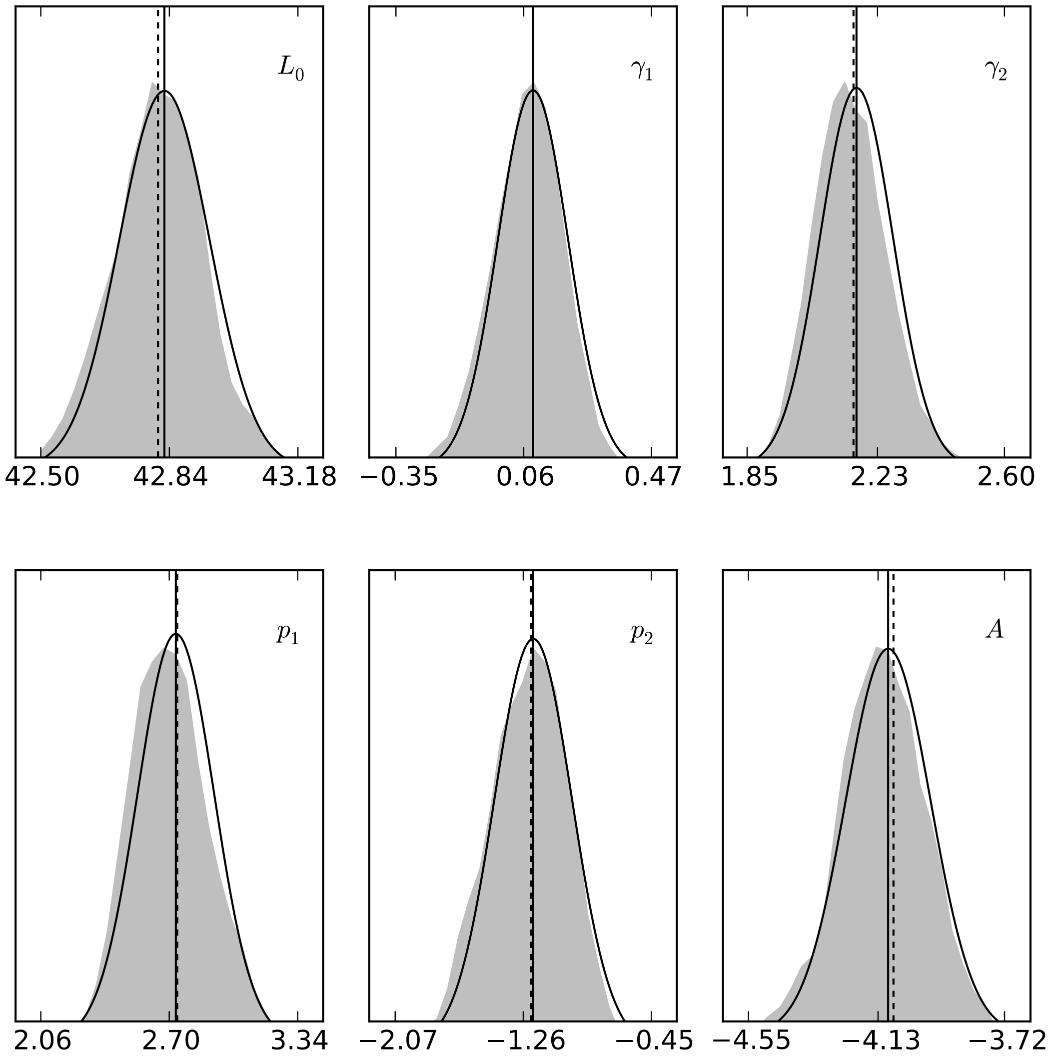


Figure A.5: Marginals for Independent Luminosity Density Evolution (ILDE). The gray area encloses the 99% probability area. The bell-shaped black line, shows the respective gaussian for the mean and standatd deviation in Table A.5. The vertical solid line shows the mean value from the MCMC result, while the dashed line shows the MLE result (Table A.6).

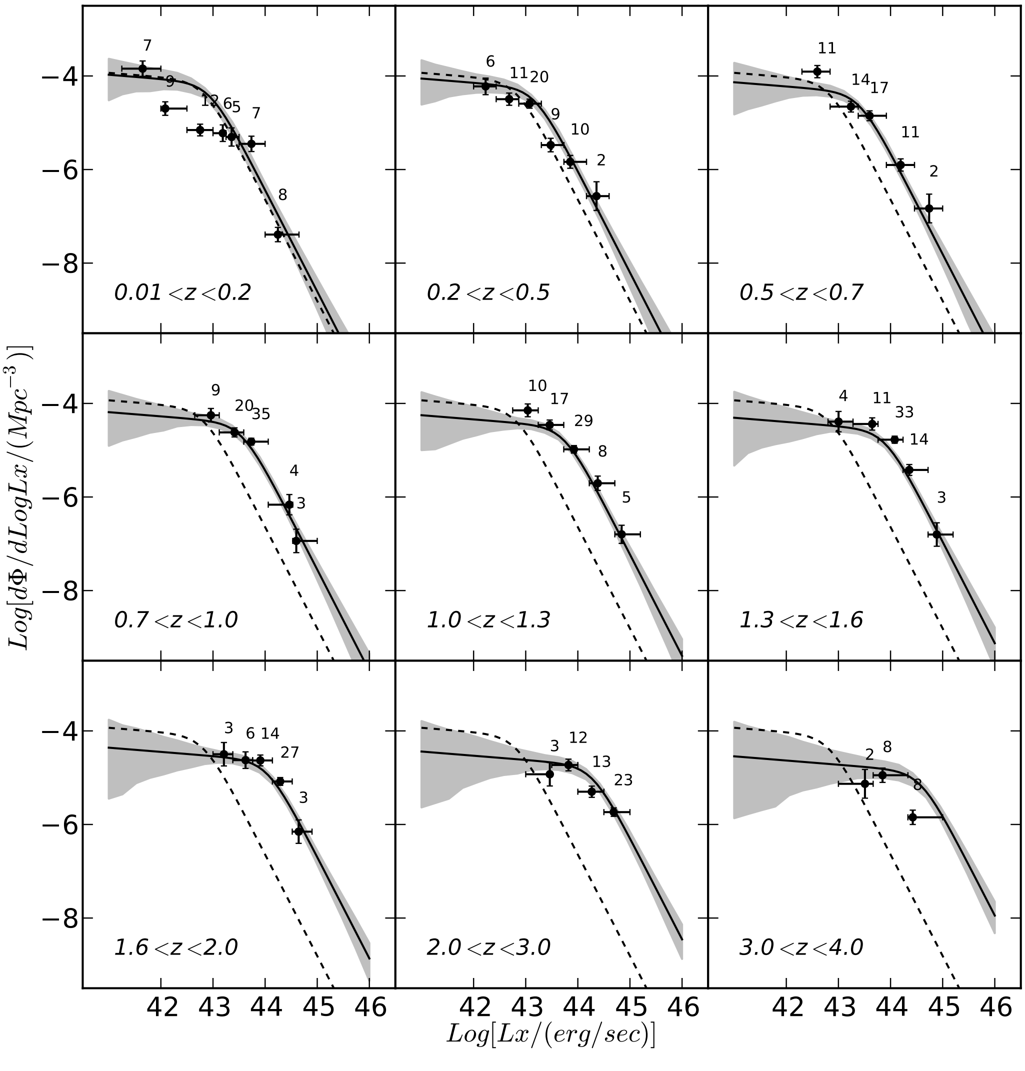


Figure A.6: Differential Luminosity Function for Independent Luminosity Density Evolution (ILDE). The gray area encloses 99% of the probability. The dashed line is the extrapolated luminosity function at z=0. The solid line is computed using the mean values from Table A.5. The points are the result of the $1/V_{max}$ method and the numbers indicated the number of AGN present in each bin. Lack of points significies extrapolation of the best solution.

A.4 Luminosity And Density Evolution - LADE

Another model describing evolution in both luminosity and density is the LADE model, which stands for Luminosity and Density Evolution. The evolution is given by:

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L/e(z), z=0)}{d\log L_x} \cdot e_d(z) \tag{A.7}$$

with the evolution in redshift and number density respectively,

$$e(z) = (\frac{1+z_c}{1+z})^{p_1} + (\frac{1+z_c}{1+z})^{p_2} \quad and \quad e_d(z) = A \cdot 10^{d(1+z)} \tag{A.8}$$

Table A.7: LADE – MCMC Parameter Estimation

Parameter	Prior Interval	Mean	Std	Min – Max
L_0	40.0 – 46.0	44.43	0.12	44.03 – 44.76
γ_1	-1.0 – 3.0	-0.05	0.13	-0.6 – 0.4
γ_2	1.5 – 6.0	2.04	0.09	1.74 – 2.45
p_1	1.0 – 6.0	3.55	0.28	2.71 – 4.7
p_2	-6.0 – 1.0	-1.46	0.97	-5.82 – 1.0
z_c	0.5 – 3.5	2.35	0.33	1.22 – 3.35
d	-1.0 – 1.0	-0.24	0.05	-0.4 – -0.09
A	-7.0 – -2.0	-3.81	0.17	-4.34 – -3.41

Table A.8: LADE – MLE Parameter Estimation

Maximum Likehihood Parameters								
	L_0	γ_1	γ_2	p_1	p_2	z_c	d	A
	44.4	-0.03	2.04	-1.27	3.56	2.26	-0.23	-3.83
Covariance Matrix								
	L_0	γ_1	γ_2	p_1	p_2	z_c	d	A
L_0	0.01	0.008	0.0054	-0.045	-0.013	0.02	-0.0007	-0.0049
γ_1	—	0.016	0.0074	-0.0036	-0.008	-0.000020	0.0011	-0.013
γ_2	—	—	0.0085	-0.0069	-0.010	0.0016	0.0015	-0.010
p_1	—	—	—	0.8	0.13	-0.20	-0.0038	0.012
p_2	—	—	—	—	0.06	-0.05	-0.0051	0.019
z_c	—	—	—	—	—	0.097	-0.0028	0.006
d	—	—	—	—	—	—	0.002	-0.0060
A	—	—	—	—	—	—	—	0.025

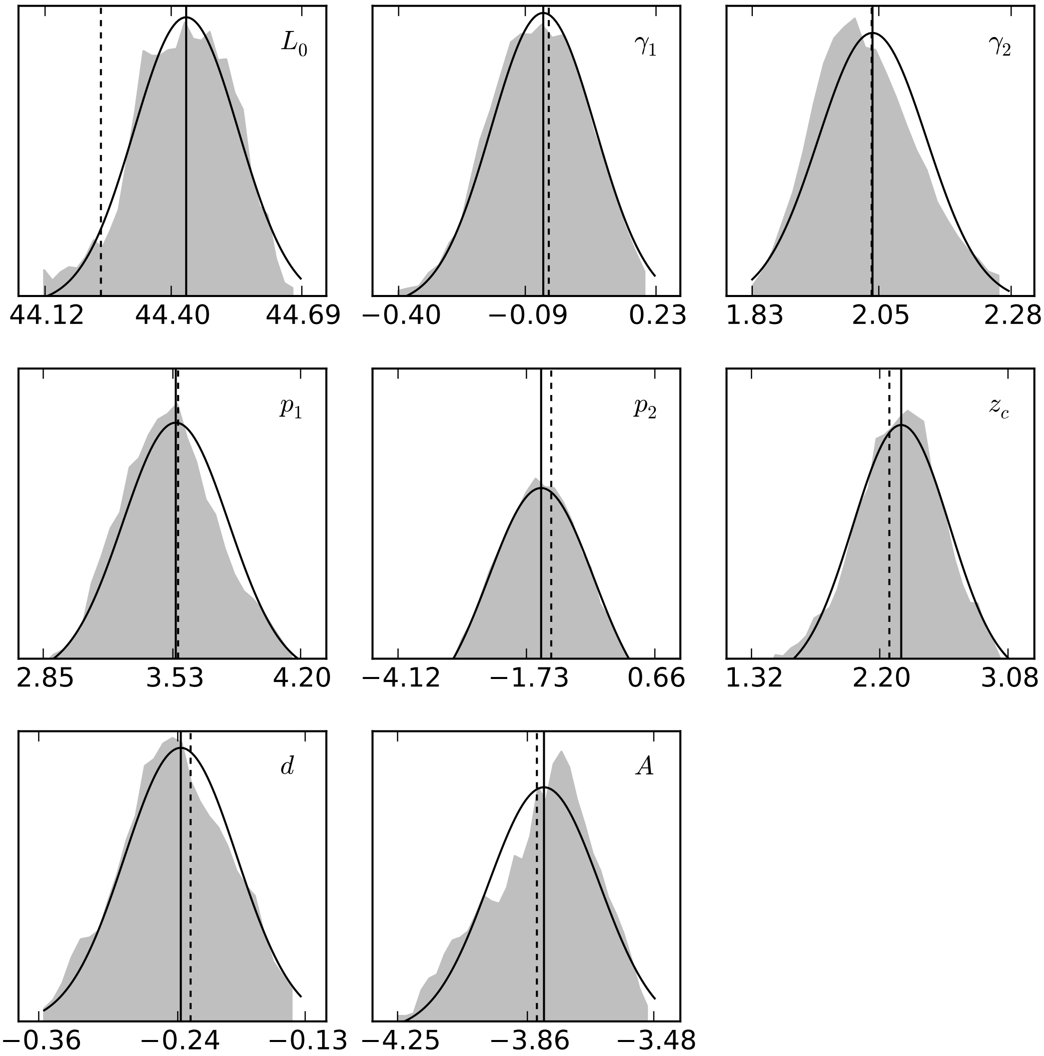


Figure A.7: Marginals for Luminosity and Density Evolution (LADE). The gray area encloses the 99% probability area. The bell-shaped black line, shows the respective gaussian for the mean and standatd deviation in Table A.3. The vertical solid line shows the mean value from the MCMC result, while the dashed line shows the MLE result (Table A.4).

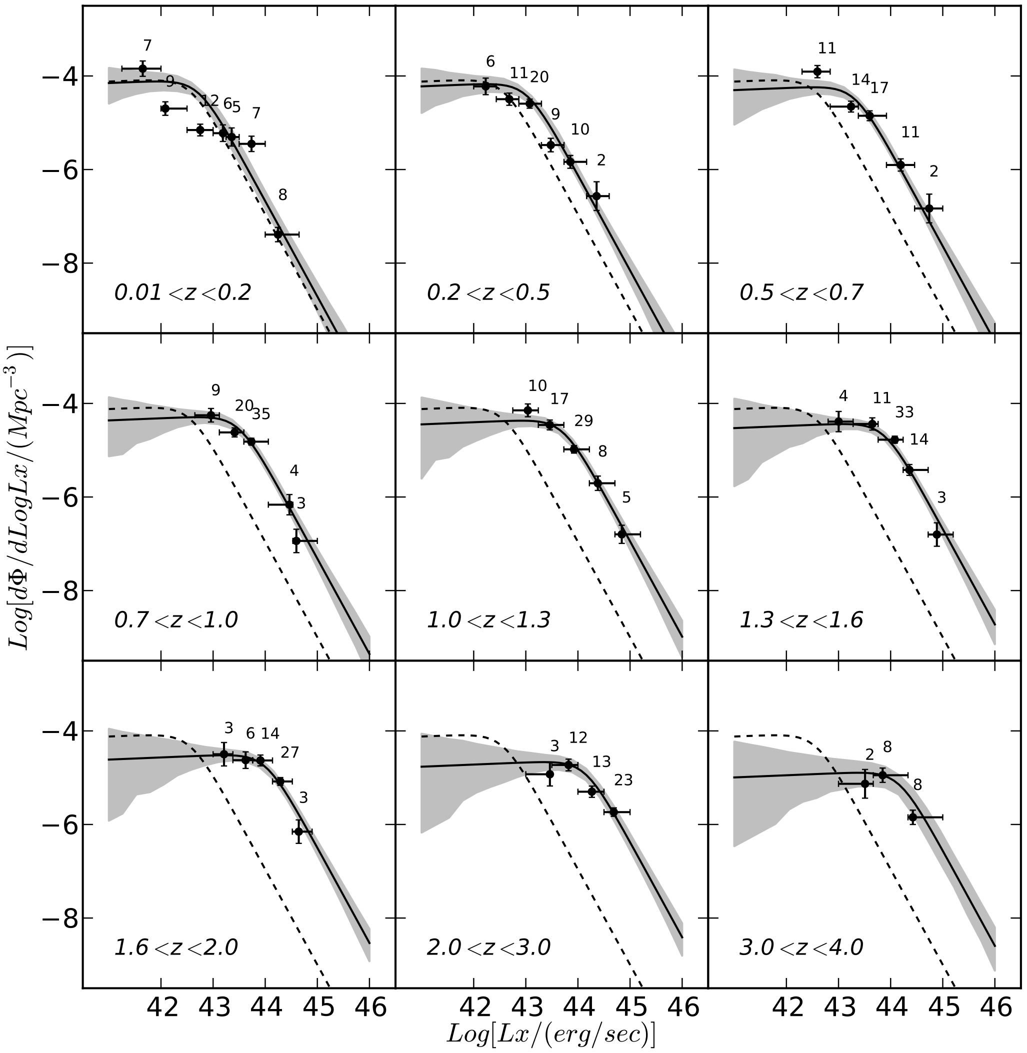


Figure A.8: Differential Luminosity Function for Luminosity and Density Evolution (LADE). The gray area encloses 99% of the probability. The dashed line is the extrapolated luminosity function at z=0. The solid line is computed using the mean values from Table A.3. The points are the result of the $1/V_{max}$ method. The numbers above each point indicate the number of AGN included in each bin. Lack of points significies extrapolation of the best solution.

A.5 Luminosity Dependent Density Evolution - LDDE

Lastly, we have the model that allows the most complicated evolution, with the density evolution being dependent on luminosity.

$$\frac{d\phi(L,z)}{d\log L_x} = \frac{d\phi(L,z=0)}{d\log L_x} \cdot e(L,z) \tag{A.9}$$

with,

$$e(L,z) = \frac{(1+z_c)^{p_1} + (1+z_c)^{p}_{2}}{(\frac{1+z}{1+z_c})^{-p_1} + (\frac{1+z}{1+z_c})^{-p_2}} \quad and \quad z_c(L) = \begin{cases} z_c^* & L \geq L_a \\ z_c^* \cdot (L/L_a)^a & L < L_a \end{cases} \tag{A.10}$$

This formalism is based on the formula used by Ueda et al. (2003), with the change in eq. A.10 to account for a soft transition before and after the critical redshift.

Table A.9: LDDE – MCMC Parameter Estimation

Parameter	Prior Interval			Mean	Std	Min – Max		
L_0	40.0	–	46.0	44.58	0.09	44.34	–	44.83
γ_1	-1.0	–	3.0	1.38	0.05	1.23	–	1.53
γ_2	1.5	–	6.0	3.3	0.29	2.54	–	4.94
p_1	3.0	–	7.0	4.84	0.19	4.15	–	5.51
p_2	-6.0	–	-1.0	-3.81	0.45	-5.95	–	-2.25
z_c	0.5	–	3.5	2.05	0.1	1.68	–	2.43
L_a	40.0	–	46.0	43.85	0.054	3.68	–	44.05
a	0.0	–	1.0	0.55	0.04	0.41	–	0.73
A	-9.0	–	-4.0	-7.33	0.18	-7.87	–	-6.89

Table A.10: LDDE – MLE Parameter Estimation

	Maximum Likehihood Parameters								
	L_0	γ_1	γ_2	p_1	p_2	z_c	L_a	α	A
	44.6	1.4	3.4	4.9	-3.8	2.03	43.8	0.55	-7.5
	Covariance Matrix								
	L_0	γ_1	γ_2	p_1	p_2	z_c	L_a	α	A
L_0	0.017	0.007	0.05	0.009	0.008	-0.005	-0.0038	0.003	-0.04
γ_1	—	0.005	0.017	0.010	0.011	-0.003	-0.0014	0.0024	-0.019
γ_2	—	—	0.20	0.012	0.013	-0.011	-0.010	0.007	-0.09
p_1	—	—	—	0.08	0.07	-0.029	-0.0010	0.0015	-0.04
p_2	—	—	—	—	0.4	-0.05	-0.003	0.014	-0.03
z_c	—	—	—	—	—	0.019	0.003	-0.0014	0.015
L_a	—	—	—	—	—	—	0.0034	-0.0020	0.007
α	—	—	—	—	—	—	—	0.003	-0.007
A	—	—	—	—	—	—	—	—	0.08

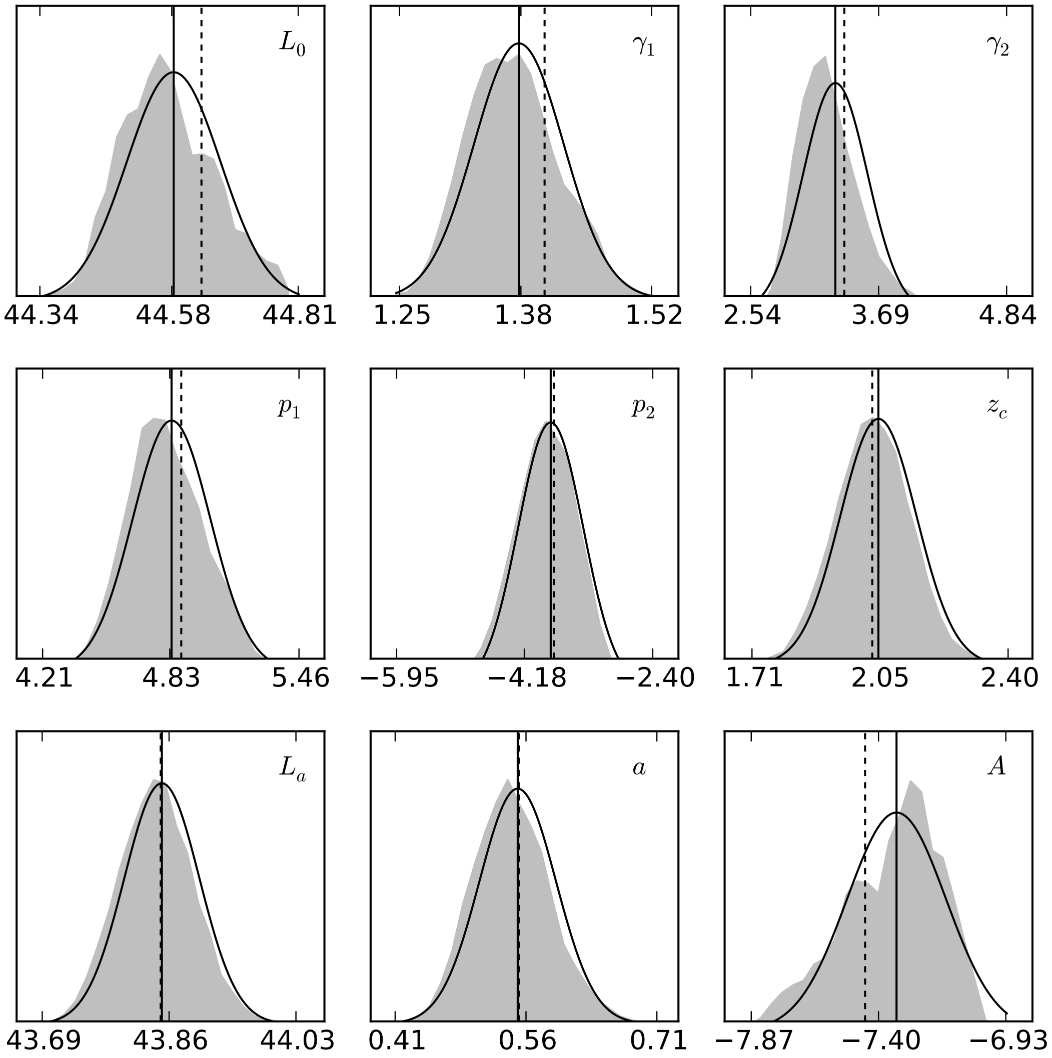


Figure A.9: Marginals for Luminosity and Density Evolution (LDDE). The gray area encloses the 99% probability area. The bell-shaped black line, shows the respective gaussian for the mean and standatd deviation in Table A.3. The vertical solid line shows the mean value from the MCMC result, while the dashed line shows the MLE result (Table A.4).

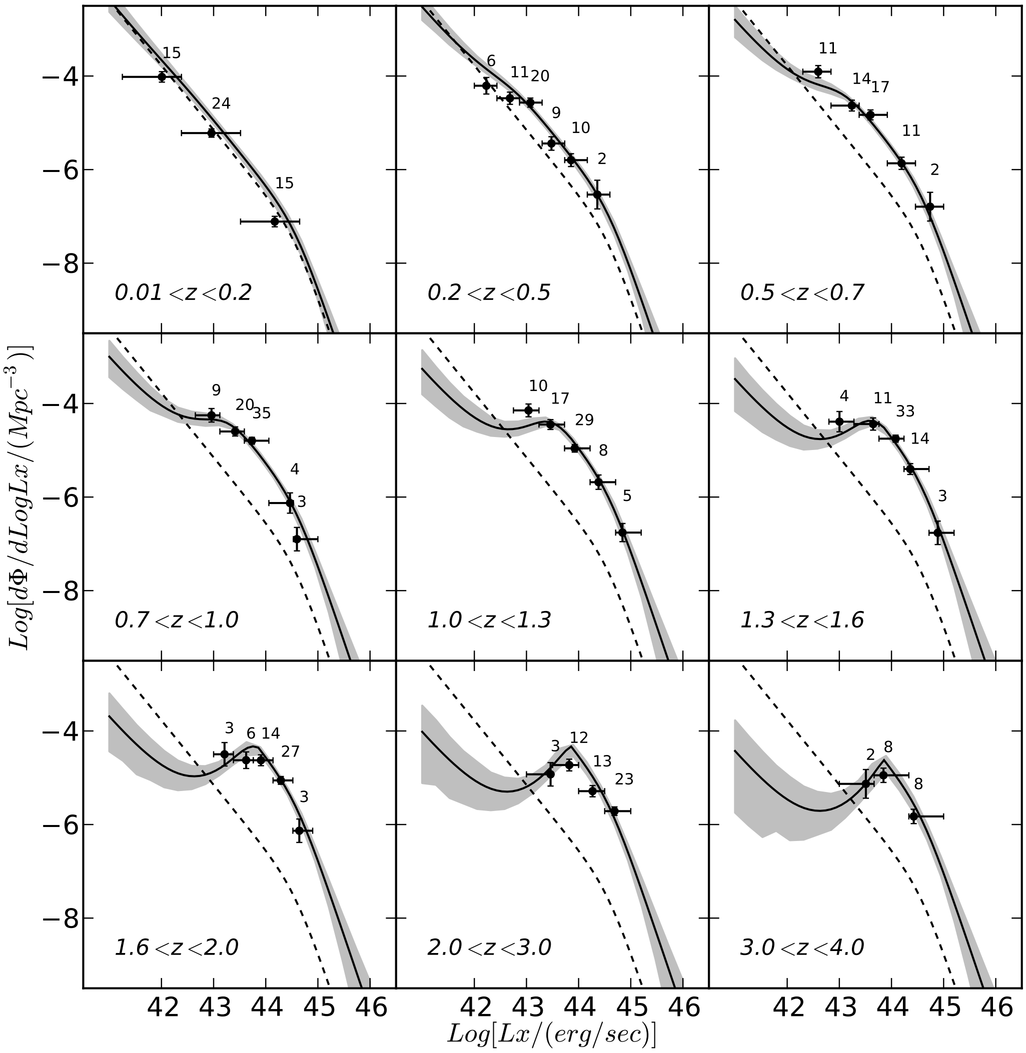


Figure A.10: Differential luminosity function Luminosity Dependent Density Evolution (LDDE).The gray area encloses 99% of the probability. The dashed line is the extrapolated luminosity function at z=0. The solid line is computed using the mean values from Table A.3. The points are the result of the $1/V_{max}$ method. The numbers above each point indicate the number of AGN included in each bin. Lack of points significies extrapolation of the best solution.

Acknowledgments

With this journey coming to an end, it is my pleasure to thank all the people who influenced and shaped my everyday life during the last three years. First and foremost, I thank Prof. Hasinger for giving me the opportunity to come to Garching and experience this unique scientific environment. His keen eye and recommendations were always a guide for my thought process. It is with great pleasure that I thank my daily supervisor and mentor, Mara Salvato. Her ethos and pathos shaped me as a scientist, while her constant support and encouragement were my day to day driving force. Thank you Mara, I have learned a lot from you.

My life here in Munich would have never been as interesting without the wonderful people I met during these years. My fellow IMPRS students always added fun to the otherwise gray weather. I particularly thank my closest friends Felicia Ziparo, Gilles Orban de Xivry, David Gruber, and Emmanuel Aller. I thank my office-mates, Viola Allevato and Jaydeep Belapure with both of which we shared more than an office during this time. Next to my closest friends I would like to thank a newer friend which turned out to be a valuable collaborator, Johannes Buchner. I hope we will continue our interesting discussions and collaboration.

I would like also to thank my family who has been a constant support in this endeavor, my parents Alexandra and Nikos and my sisters Dionysia and Theoni. Last but not least I thank, him who has been my rock for the past seven years, my fiancé Konstantinos Raftopoulos. His support is beyond words.

I only hope that my future holds as interesting experiences as those of my years in Munich.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A. et al. (2009): ApJS**182**: 543–558.

Afonso, J., Bizzocchi, L., Ibar, E. et al. (2011): ApJ**743**: 122.

Aird, J., Nandra, K., Georgakakis, A. et al. (2008): MNRAS**387**: 883–896.

Aird, J., Nandra, K., Laird, E. S. et al. (2010): MNRAS**401**: 2531–2551.

Akaike, H. (1974): IEEE Transactions on Automatic Control **19**: 716–723.

Alexander, D. M., Bauer, F. E., Brandt, W. N. et al. (2003): AJ**126**: 539–574.

Aller, H. D., Aller, M. F., Latimer, G. E. and Hodge, P. E. (1985): ApJS**59**: 513–768.

Allevato, V., Finoguenov, A., Cappelluti, N. et al. (2011): ApJ**736**: 99.

Antonucci, R. (1993): ARA&A**31**: 473–521.

Antonucci, R. R. J. (1982): Nature**299**: 605.

Arnaud, K. A. (1996): in G. H. Jacoby and J. Barnes (eds), Astronomical Data Analysis Software and Systems V, Bd. 101 von Astronomical Society of the Pacific Conference Series, S. 17.

Arnouts, S., Schiminovich, D., Ilbert, O. et al. (2005): ApJ**619**: L43–L46.

Ballantyne, D. R. and Papovich, C. (2007): ApJ**660**: 988–994.

Barger, A. J., Cowie, L. L., Bautz, M. W. et al. (2001): AJ**122**: 2177–2189.

Barger, A. J., Cowie, L. L., Capak, P. et al. (2003): AJ**126**: 632–665.

Barger, A. J., Cowie, L. L., Mushotzky, R. F. et al. (2005): AJ**129**: 578–609.

Barris, B. J., Tonry, J. L., Blondin, S. et al. (2004): ApJ**602**: 571–594.

Barro, G., Pérez-González, P. G., Gallego, J. et al. (2011): ApJS**193**: 30.

Bauer, F. E., Alexander, D. M., Brandt, W. N. et al. (2004): AJ**128**: 2048–2065.

Becker, A. C., Silvestri, N. M., Owen, R. E. et al. (2007): PASP**119**: 1462–1482.

Bender, R., Appenzeller, I., Böhm, A. et al. (2001): The FORS Deep Field: Photometric Data and Photometric Redshifts, in S. Cristiani, A. Renzini and R. E. Williams (eds), Deep Fields, S. 96.

Bennett, A. S. (1962): MNRAS**125**: 75.

Bertin, E. and Arnouts, S. (1996): A&AS**117**: 393–404.

Bixler, J. V., Bowyer, S. and Laget, M. (1991): A&A**250**: 370–388.

Bohlin, R. C., Colina, L. and Finley, D. S. (1995): AJ**110**: 1316.

Bongiorno, A., Zamorani, G., Gavignaud, I. et al. (2007): A&A**472**: 443–454.

Bonnarel, F., Fernique, P., Bienaymé, O. et al. (2000): A&AS**143**: 33–40.

Bouwens, R. J., Illingworth, G. D., Franx, M. and Ford, H. (2007): ApJ**670**: 928–958.

Boyle, B. J., Shanks, T., Georgantopoulos, I. et al. (1994): MNRAS**271**: 639.

Brammer, G. B., van Dokkum, P. G. and Coppi, P. (2008): ApJ**686**: 1503–1513.

Brandt, W. N., Alexander, D. M., Hornschemeier, A. E. et al. (2001): AJ**122**: 2810–2832.

Brandt, W. N. and Hasinger, G. (2005): Deep Extragalactic X-Ray Surveys, ARA&A**43**: 827–859.

Brunner, H., Cappelluti, N., Hasinger, G. et al. (2008): A&A**479**: 283–300.

Brusa, M., Civano, F., Comastri, A. et al. (2010): ApJ**716**: 348–369.

Brusa, M., Comastri, A., Daddi, E. et al. (2005): A&A**432**: 69–81.

Brusa, M., Comastri, A., Gilli, R. et al. (2009): ApJ**693**: 8–22.

Brusa, M., Zamorani, G., Comastri, A. et al. (2007): ApJS**172**: 353–367.

Budavári, T., Szalay, A. S., Connolly, A. J. et al. (2000): AJ**120**: 1588–1598.

Caccianiga, A., Severgnini, P., Della Ceca, R. et al. (2008): A&A**477**: 735–746.

Calzetti, D., Armus, L., Bohlin, R. C. et al. (2000): ApJ**533**: 682–695.

Capak, P., Aussel, H., Ajiki, M. et al. (2007): ApJS**172**: 99–116.

Cappelluti, N., Brusa, M., Hasinger, G. et al. (2009): A&A**497**: 635–648.

Cappelluti, N., Hasinger, G., Brusa, M. et al. (2007): ApJS**172**: 341–352.

Caputi, K. I., McLure, R. J., Dunlop, J. S. et al. (2006): MNRAS**366**: 609–623.

Cardamone, C. N., van Dokkum, P. G., Urry, C. M. et al. (2010): ApJS**189**: 270–285.

Casali, M., Adamson, A., Alves de Oliveira, C. et al. (2007): A&A**467**: 777–784.

Chabrier, G., Baraffe, I., Allard, F. et al. (2000): ApJ**542**: 464–472.

Chapman, S. C., Blain, A. W., Smail, I. et al. (2005): ApJ**622**: 772–796.

Chapman, S. C., Smail, I., Blain, A. W. et al. (2004): ApJ**614**: 671–678.

Ciliegi, P., Zamorani, G., Hasinger, G. et al. (2003): A&A**398**: 901–918.

Cisternas, M., Jahnke, K., Inskip, K. J. et al. (2011): ApJ**726**: 57.

Civano, F., Brusa, M., Comastri, A. et al. (2011): ApJ**741**: 91.

Cole, S., Norberg, P., Baugh, C. M. et al. (2001): MNRAS**326**: 255–273.

Croton, D. J., Springel, V., White, S. D. M. et al. (2006): MNRAS**365**: 11–28.

Dahlen, T., Mobasher, B., Dickinson, M. et al. (2010): ApJ**724**: 425–447.

Della Ceca, R., Caccianiga, A., Severgnini, P. et al. (2008): A&A**487**: 119–130.

Della Ceca, R., Maccacaro, T., Caccianiga, A. et al. (2004): A&A**428**: 383–399.

Done, C., Davis, S. W., Jin, C. et al. (2012): MNRAS**420**: 1848–1860.

Draper, A. R. and Ballantyne, D. R. (2009): ApJ**707**: 778–786.

Draper, A. R. and Ballantyne, D. R. (2012): ApJ**751**: 72.

Ebrero, J., Carrera, F. J., Page, M. J. et al. (2009): A&A**493**: 55–69.

Elitzur, M. (2012): ApJ**747**: L33.

Elvis, M., Civano, F., Vignali, C. et al. (2009): ApJS**184**: 158–171.

Elvis, M., Schreier, E. J., Tonry, J. et al. (1981): ApJ**246**: 20–27.

Fadda, D., Flores, H., Hasinger, G. et al. (2002): A&A**383**: 838–853.

Fanidakis, N., Baugh, C. M., Benson, A. J. et al. (2011): MNRAS**410**: 53–74.

Feroz, F., Hobson, M. P. and Bridges, M. (2009): MNRAS**398**: 1601–1614.

Ferrarese, L. and Merritt, D. (2000): ApJ**539**: L9–L12.

Fiore, F., Puccetti, S., Brusa, M. et al. (2009): ApJ**693**: 447–462.

Fotopoulou, S., Salvato, M., Hasinger, G. et al. (2012): ApJS**198**: 1.

Frayer, D. T., Sanders, D. B., Surace, J. A. et al. (2009): AJ**138**: 1261–1270.

Gabasch, A., Bender, R., Seitz, S. et al. (2004): A&A**421**: 41–58.

Gabasch, A., Hopp, U., Feulner, G. et al. (2006): A&A**448**: 101–121.

Geach, J. E., Murphy, D. N. A. and Bower, R. G. (2011): MNRAS**413**: 3059–3067.

Gebhardt, K., Bender, R., Bower, G. et al. (2000): ApJ**539**: L13–L16.

Georgakakis, A., Coil, A. L., Laird, E. S. et al. (2009): MNRAS**397**: 623–633.

Giacconi, R., Gursky, H., Paolini, F. R. and Rossi, B. B. (1962): Physical Review Letters **9**: 439–443.

Giacconi, R., Kellogg, E., Gorenstein, P. et al. (1971): ApJ**165**: L27.

Gilli, R., Comastri, A. and Hasinger, G. (2007): A&A**463**: 79–96.

Giodini, S., Pierini, D., Finoguenov, A. et al. (2009): ApJ**703**: 982–993.

Glikman, E., Bogosavljević, M., Djorgovski, S. G. et al. (2010): ApJ**710**: 1498–1514.

Glikman, E., Djorgovski, S. G., Stern, D. et al. (2011): ApJ**728**: L26.

Gursky, H., Giacconi, R., Paolini, F. R. and Rossi, B. B. (1963): Physical Review Letters **11**: 530–535.

Hainline, L. J., Blain, A. W., Smail, I. et al. (2009): ApJ**699**: 1610–1632.

Han, Y., Dai, B., Wang, B. et al. (2012): MNRAS**423**: 464–477.

Hashimoto, Y., Henry, J. P., Hasinger, G. et al. (2005): A&A**439**: 29–33.

Hasinger, G. (2008): A&A**490**: 905–922.

Hasinger, G., Giacconi, R., Gunn, J. E. et al. (1998): A&A**340**: L27–L30.

Hasinger, G., Miyaji, T. and Schmidt, M. (2005): A&A**441**: 417–434.

Heckman, T. M. (1980): A&A**87**: 152–164.

Henry, J. P., Salvato, M., Finoguenov, A. et al. (2010): ApJ**725**: 615–624.

Hewett, P. C., Warren, S. J., Leggett, S. K. et al. (2006): MNRAS**367**: 454–468.

Hiroi, K., Ueda, Y., Isobe, N. et al. (2011): PASJ**63**: 677.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W. (1983).

Hodgkin, S. T., Irwin, M. J., Hewett, P. C. et al. (2009): MNRAS**394**: 675–692.

Hopkins, P. F., Bundy, K., Hernquist, L. et al. (2007): ApJ**659**: 976–996.

Hopkins, P. F. and Hernquist, L. (2006): ApJS**166**: 1–36.

Hopkins, P. F., Hernquist, L., Cox, T. J. et al. (2005a): ApJ**632**: 81–91.

Hopkins, P. F., Hernquist, L., Cox, T. J. et al. (2005b): ApJ**630**: 716–720.

Hopkins, P. F., Hernquist, L., Cox, T. J. et al. (2006): ApJ**639**: 700–709.

Hornschemeier, A. E., Bauer, F. E., Alexander, D. M. et al. (2003): AJ**126**: 575–595.

Huang, J.-S., Glazebrook, K., Cowie, L. L. and Tinney, C. (2003): ApJ**584**: 203–209.

Hubble, E. (1929): Proceedings of the National Academy of Science **15**: 168–173.

Hubble, E. P. (1925): Popular Astronomy **33**: 252.

Hubble, E. P. (1936): Realm of the Nebulae.

Ikeda, H., Nagao, T., Matsuoka, K. et al. (2011): ApJ**728**: L25.

Ilbert, O., Arnouts, S., McCracken, H. J. et al. (2006): A&A**457**: 841–856.

Ilbert, O., Capak, P., Salvato, M. et al. (2009): ApJ**690**: 1236–1249.

Ishisaki, Y., Ueda, Y., Yamashita, A. et al. (2001): PASJ**53**: 445–457.

Ivison, R. J., Smail, I., Dunlop, J. S. et al. (2005): MNRAS**364**: 1025–1040.

James, F. and Roos, M. (1975): Computer Physics Communications **10**: 343–367.

Jones, L. R., McHardy, I. M., Merrifield, M. R. et al. (1997): MNRAS**285**: 547–560.

Kelly, B. C., Fan, X. and Vestergaard, M. (2008): ApJ**682**: 874–895.

Khachikian, E. E. and Weedman, D. W. (1971): Astrofizika **7**: 389–406.

Khachikian, E. Y. and Weedman, D. W. (1974): ApJ**192**: 581–589.

Kormendy, J. and Kennicutt, Jr., R. C. (2004): ARA&A**42**: 603–683.

La Franca, F., Fiore, F., Comastri, A. et al. (2005): ApJ**635**: 864–879.

Laing, R. A., Riley, J. M. and Longair, M. S. (1983): MNRAS**204**: 151–187.

Laird, E. S., Nandra, K., Georgakakis, A. et al. (2009): ApJS**180**: 102–116.

Lawrence, A., Warren, S. J., Almaini, O. et al. (2007): MNRAS**379**: 1599–1617.

Lehmann, I., Hasinger, G., Schmidt, M. et al. (2001): A&A**371**: 833–857.

Lehmer, B. D., Brandt, W. N., Alexander, D. M. et al. (2005): ApJS**161**: 21–40.

Lemaître, G. (1927): Annales de la Societe Scietifique de Bruxelles **47**: 49–59.

Lilly, S. J., Le Brun, V., Maier, C. et al. (2009): ApJS**184**: 218–229.

Lilly, S. J., Le Fèvre, O., Renzini, A. et al. (2007): ApJS**172**: 70–85.

Lockman, F. J., Jahoda, K. and McCammon, D. (1986): ApJ**302**: 432–449.

Luo, B., Brandt, W. N., Xue, Y. Q. et al. (2010): ApJS**187**: 560–580.

Luo, B., Brandt, W. N., Xue, Y. Q. et al. (2011): ApJ**740**: 37.

Lynden-Bell, D. (1969): Nature**223**: 690–694.

Maccacaro, T., della Ceca, R., Gioia, I. M. et al. (1991): ApJ**374**: 117–133.

Maccacaro, T., Gioia, I. M., Avni, Y. et al. (1983): ApJ**266**: L73–L77.

Maccacaro, T., Gioia, I. M. and Stocke, J. T. (1984): ApJ**283**: 486–494.

Maccacaro, T., Gioia, I. M., Wolter, A. et al. (1988): ApJ**326**: 680–690.

Magorrian, J., Tremaine, S., Richstone, D. et al. (1998): AJ**115**: 2285–2305.

Mahmood, A., Devriendt, J. E. G. and Silk, J. (2005): MNRAS**359**: 1363–1378.

Mainieri, V., Bergeron, J., Hasinger, G. et al. (2002): A&A**393**: 425–438.

Mainieri, V., Bongiorno, A., Merloni, A. et al. (2011): A&A**535**: A80.

Mainieri, V., Hasinger, G., Cappelluti, N. et al. (2007): ApJS**172**: 368–382.

Margoniner, V. E. and Wittman, D. M. (2008): ApJ**679**: 31–51.

Marmo, C. and Bertin, E. (2008): in R. W. Argyle, P. S. Bunclark and J. R. Lewis (eds), Astronomical Data Analysis Software and Systems XVII, Bd. 394 von Astronomical Society of the Pacific Conference Series, S. 619.

Marshall, H. L., Tananbaum, H., Avni, Y. et al. (1983): ApJ**269**: 35–41.

Marulli, F., Bonoli, S., Branchini, E. et al. (2008): MNRAS**385**: 1846–1858.

Massey, R., Rhodes, J., Ellis, R. et al. (2007): Nature**445**: 286–290.

Masters, D., Capak, P., Salvato, M. et al. (2012): ArXiv e-prints .

Mateos, S., Barcons, X., Carrera, F. J. et al. (2005): A&A**444**: 79–99.

McCracken, H. J., Capak, P., Salvato, M. et al. (2010): ApJ**708**: 202–217.

McCracken, H. J., Le Fèvre, O., Brodwin, M. et al. (2001): A&A**376**: 756–774.

Mihara, T., Nakajima, M., Sugizaki, M. et al. (2011): PASJ**63**: 623.

Miyaji, T., Hasinger, G. and Schmidt, M. (2000): A&A**353**: 25–40.

Mobasher, B., Idzi, R., Benítez, N. et al. (2004): ApJ**600**: L167–L170.

Moretti, A., Campana, S., Lazzati, D. and Tagliaferri, G. (2003): ApJ**588**: 696–703.

Morrissey, P., Conrow, T., Barlow, T. A. et al. (2007): ApJS**173**: 682–697.

Mushotzky, R. F., Cowie, L. L., Barger, A. J. and Arnaud, K. A. (2000): Nature**404**: 459–464.

Nandra, K. and Pounds, K. A. (1994): MNRAS**268**: 405.

Nenkova, M., Sirocky, M. M., Nikutta, R. et al. (2008): ApJ**685**: 160–180.

Norberg, P., Cole, S., Baugh, C. M. et al. (2002): MNRAS**336**: 907–931.

Orban de Xivry, G., Davies, R., Schartmann, M. et al. (2011): MNRAS**417**: 2721–2736.

Osterbrock, D. E. (1977): ApJ**215**: 733–745.

Osterbrock, D. E. (1981): ApJ**249**: 462–470.

Oyabu, S., Yun, M. S., Murayama, T. et al. (2005): AJ**130**: 2019–2042.

Oyaizu, H., Lima, M., Cunha, C. E. et al. (2008): ApJ**674**: 768–783.

Padovani, P., Miller, N., Kellermann, K. I. et al. (2011): ApJ**740**: 20.

Page, K. L., Reeves, J. N., O'Brien, P. T. and Turner, M. J. L. (2005): MNRAS**364**: 195–207.

Page, M. J. and Carrera, F. J. (2000): MNRAS**311**: 433–440.

Page, M. J., Carrera, F. J., Hasinger et al. (1996): MNRAS**281**: 579–590.

Papovich, C., Momcheva, I., Willmer, C. N. A. et al. (2010): ApJ**716**: 1503–1513.

Peacock, J. A. (1999): Cosmological Physics.

Penston, M. V. and Perez, E. (1984): MNRAS**211**: 33P–39P.

Pérez-González, P. G., Rieke, G. H., Villar, V. et al. (2008): ApJ**675**: 234–261.

Pica, A. J., Smith, A. G., Webb, J. R. et al. (1988): AJ**96**: 1215–1226.

Pickles, A. J. (1998): PASP**110**: 863–878.

Piconcelli, E., Jimenez-Bailón, E., Guainazzi, M. et al. (2005): A&A**432**: 15–30.

Polletta, M., Tajer, M., Maraschi, L. et al. (2007): ApJ**663**: 81–102.

Prevot, M. L., Lequeux, J., Prevot, L. et al. (1984): A&A**132**: 389–392.

Rafferty, D. A., Brandt, W. N., Alexander, D. M. et al. (2011): ApJ**742**: 3.

Reeves, J. N. and Turner, M. J. L. (2000): MNRAS**316**: 234–248.

Richards, G. T., Weinstein, M. A., Schneider, D. P. et al. (2001): AJ**122**: 1151–1162.

Rodighiero, G., Fadda, D., Franceschini, A. et al. (2005): MNRAS**357**: 449–460.

Rovilos, E., Burwitz, V., Szokoly, G. et al. (2009): A&A**507**: 195–208.

Rovilos, E., Fotopoulou, S., Salvato, M. et al. (2011): A&A**529**: A135.

Rowan-Robinson, M., Babbedge, T., Oliver, S. et al. (2008): MNRAS**386**: 697–714.

Salpeter, E. E. (1964): ApJ**140**: 796–800.

Salvato, M., Hasinger, G., Ilbert, O. et al. (2009): ApJ**690**: 1250–1263.

Salvato, M., Ilbert, O., Hasinger, G. et al. (2011): ApJ**742**: 61.

Sandage, A. (1965): ApJ**141**: 1560.

Sanders, D. B., Salvato, M., Aussel, H. et al. (2007): ApJS**172**: 86–98.

Schlegel, D. J., Finkbeiner, D. P. and Davis, M. (1998): ApJ**500**: 525.

Schmidt, M. (1963): Nature**197**: 1040.

Schmidt, M. (1968): ApJ**151**: 393.

Schmidt, M. and Green, R. F. (1983): ApJ**269**: 352–374.

Schmidt, M., Hasinger, G., Gunn, J. et al. (1998): A&A**329**: 495–503.

Schmidt, M., Schneider, D. P. and Gunn, J. E. (1987): ApJ**316**: L1–L3.

Schmitt, J. L. (1968): Nature**218**: 663.

Seyfert, C. K. (1943): ApJ**97**: 28.

Silverman, J. D., Green, P. J., Barkhouse, W. A. et al. (2008): ApJ**679**: 118–139.

Slipher, V. M. (1917): Proceedings of the American Philosophical Society **56**: 403–409.

Smail, I., Chapman, S. C., Blain, A. W. et al. (2004): ApJ**616**: 71–85.

Stevens, J. A., Page, M. J., Ivison, R. J. et al. (2003): MNRAS**342**: 249–258.

Strittmatter, P. A., Serkowski, K., Carswell, R. et al. (1972): ApJ**175**: L7.

Su, M., Slatyer, T. R. and Finkbeiner, D. P. (2010): ApJ**724**: 1044–1082.

Sugizaki, M., Mihara, T., Serino, M. et al. (2011): PASJ**63**: 635.

Surace, J. A., Shupe, D. L., Fang, F. et al. (2005): American Astronomical Society Meeting Abstracts, Bd. 37 von Bulletin of the American Astronomical Society, S. 1246.

Swinbank, A. M., Smail, I., Chapman, S. C. et al. (2004): ApJ**617**: 64–80.

Tanaka, M., Finoguenov, A. and Ueda, Y. (2010): ApJ**716**: L152–L156.

Taylor, M. B. (2005): in P. Shopbell, M. Britton and R. Ebert (eds), Astronomical Data Analysis Software and Systems XIV, Bd. 347 von Astronomical Society of the Pacific Conference Series, S. 29.

Taylor, M. B. (2006): in C. Gabriel, C. Arviset, D. Ponz and S. Enrique (eds), Astronomical Data Analysis Software and Systems XV, Bd. 351 von Astronomical Society of the Pacific Conference Series, S. 666.

Tomida, H., Tsunemi, H., Kimura, M. et al. (2011): PASJ**63**: 397–405.

Tozzi, P., Gilli, R., Mainieri, V. et al. (2006): A&A**451**: 457–474.

Trump, J. R., Impey, C. D., Elvis, M. et al. (2009): ApJ**696**: 1195–1212.

Tsunemi, H., Tomida, H., Katayama, H. et al. (2010): PASJ**62**: 1371–.

Ueda, Y., Akiyama, M., Ohta, K. et al. (2003): ApJ**598**: 886–908.

Ueda, Y., Hiroi, K., Isobe, N. et al. (2011): PASJ**63**: 937.

Urry, C. M. and Padovani, P. (1995): PASP**107**: 803.

Whitaker, K. E., Labbé, I., van Dokkum, P. G. et al. (2011): ApJ**735**: 86.

Williams, R. J., Quadri, R. F., Franx, M. et al. (2009): ApJ**691**: 1879–1895.

Wolf, C., Meisenheimer, K., Kleinheinrich, M. et al. (2004): A&A**421**: 913–936.

Worsley, M. A., Fabian, A. C., Barcons, X. et al. (2004): MNRAS**352**: L28–L33.

Wuyts, S., Labbé, I., Schreiber, N. M. F. et al. (2008): ApJ**682**: 985–1003.

Xue, Y. Q., Luo, B., Brandt, W. N. et al. (2011): ApJS**195**: 10.

Yencho, B., Barger, A. J., Trouille, L. et al. (2009): ApJ**698**: 380–396.

Zamojski, M. A., Schiminovich, D., Rich, R. M. et al. (2007): ApJS**172**: 468–493.

Zamorani, G., Maccacaro, T., Giommi, P. and Tananbaum, H. (1984): ApJ**278**: 28–36.

Zappacosta, L., Maiolino, R., Finoguenov, A. et al. (2005): A&A**434**: 801–809.

Zheng, X. Z., Bell, E. F., Somerville, R. S. et al. (2009): ApJ**707**: 1566–1577.

Zubovas, K. and Nayakshin, S. (2012): MNRAS**424**: 666–683.