Technische Universität München

TECHNISCHE UNIVERSITÄT MÜNCHEN
Max-Planck Institut für Plasmaphysik

Halo Occupation of AGN in deep X-ray surveys

Viola Allevato

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. H. Friedrich
Prüfer der Dissertation:	1. Hon.-Prof. Dr. G. Hasinger
	2. Univ.-Prof. Dr. St. Paul

Die Dissertation wurde am 20.08.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 24.09.2012 angenommen.

a mio padre

a mia madre

voce e anima

del mio Universo

Abstract

Measurement of the spatial distribution (described by the correlation function, CF) of Active Galactic Nuclei (AGN) in the Universe provides a unique way to study the typical environment in which AGN preferentially reside (through the connection with their host dark matter halos, DMHs) and to address which physical processes are triggering AGN activity. This is important for understanding not only the origin and evolution of black holes (BHs) but also the origin and evolution of galaxies. The halo model approach reduces the problem of discussing the spatial distribution of AGN to studying how they populate their host halos. The CF amplitude reflects, through the bias parameter, the typical mass of DMHs in which AGN reside, while the halo occupation distribution (HOD) describes the physical relation between AGN and DMHs at the level of individual halos.

In this thesis I use the large sample of X-ray selected AGN and galaxy groups constructed using XMM and *Chandra* data in the two square degree COSMOS survey to investigate, through the halo model approach, the clustering properties of X-ray AGN as a function of redshift and the occupation of galaxy groups by X-ray AGN. In the first part of the thesis I focus on the redshift evolution of the auto-CF of 593 X-ray selected AGN, extracted from the 0.5-2 keV X-ray mosaic of the 2.13 deg^2 XMM-COSMOS. I find evidence of a redshift evolution of the AGN bias with a DMH mass consistent with being constant ($\sim 10^{13}$ h$^{-1}M_{\odot}$) up to $z < 2$, i.e. *X-ray AGN preferentially reside in dense environments typical of galaxy groups.* The theoretical models which assume a quasar phase triggered by major mergers, cannot reproduce the high bias factors and DMH masses found for XMM-COSMOS AGN. *These results extend up to $z \sim 2.2$ the statement that, for moderate-luminosity X-ray AGN, the contribution from major mergers is outnumbered by other processes, possibly secular ones, such as tidal disruptions or disk instabilities.*

In the second part of the thesis, I present the first direct measurement of the mean HOD of X-ray selected AGN in the COSMOS field at $z \leq 1$, based on the association of 41 XMM and 17 *Chandra*-COSMOS AGN with member galaxies of 189 X-ray detected galaxy groups from XMM and *Chandra* data. Separating the contribution to the occupation of galaxy groups by AGN in satellite and central galaxies, I find the average number of AGN among central galaxies to be modelled by a softened step function at logM $>$ logM_{min} $[M_{\odot}]$= 12.75(12.10, 12.95) while the satellite AGN HOD increases with the halo mass (following a power-law with slope $\alpha < 0.6$) slower than a simple linear proportion ($\alpha = 1$), i.e. the average number of AGN in satellite galaxies is not only triggered by the halo mass, as observed for satellite galaxies.

Zusammenfassung

Die Untersuchung der räumlichen Verteilung von Aktiven Galaktischen Kernen (engl. "active galactic nuclei", AGN) im Universum, beschrieben durch die Korrelationsfunktion (engl., "correlation function", CF), bietet die einzigartige Möglichkeit, sowohl die typische Umgebung zu untersuchen in welcher sich AGN vorzugsweise befinden (mittels der Verknüpfung der AGN mit ihren übergeordneten Dunkle Materie Halos, engl. "dark matter halos", DMHs), als auch zu verstehen welche physikalischen Prozesse einen AGN berhaupt aktivieren.

Die Untersuchung der räumlichen Verteilung von Aktiven Galaktischen Kernen (engl. "active galactic nuclei", AGN) im Universum, beschrieben durch die Korrelationsfunktion (engl., "correlation function", CF), bietet die einzigartige Möglichkeit, sowohl die typische Umgebung zu untersuchen in welcher sich AGN vorzugsweise befinden (mittels der Verknüpfung der AGN mit ihren übergeordneten Dunkle Materie Halos, engl. "dark matter halos", DMHs), als auch zu verstehen welche physikalischen Prozesse einen AGN berhaupt aktivieren. Das Wissen über diese Bedingungen ist nicht nur für das Verständnis der Aktivierung von Schwarzen Löchern (engl., "black holes", BH) wichtig, sondern hilft auch, die Entstehung und Evolution von Galaxien besser zu begreifen. Im Halo-Modell-Ansatz wird des Problem der räumlichen Verteilung auf die Besetzung in den jeweiligen DMHs reduziert. Die Amplitude der CF reflektiert mit Hilfe des Bias-Parameters die typische Masse der DMHs in denen AGN zu finden sind, während die Halo-Besetzungsverteilung (engl. "halo occupation distribution", HOD) die physikalische Beziehung zwischen den AGN und DMHs auf der Ebene der einzelnen Halos beschreibt.

In der vorliegenden Dissertation verwende ich umfangreiche Stichproben von AGN und Galaxiengruppen, ausgewählt aus den XMM und Chandra Röntgendaten im $2\,\mathrm{deg}^2$ COSMOS Feld, um mit Hilfe des Halo-Modell-Ansatzes die Eigenschaften der räumlichen Verteilung von Röntgen-AGN als Funktion der Rotverschiebung, sowie deren Besetzung von Galaxiengruppen, zu untersuchen. Im ersten Teil der Arbeit konzentriere ich mich auf die Rotverschiebungsentwicklung der Auto-CF von 593 Röntgen-selektierten AGN mit Rotverschiebung $z < 4$ aus dem 2.13 Quadratgrad großen $0.5-2\,\mathrm{keV}$ XMM-COSMOS Mosaik. Meine Untersuchungen deuten auf eine Evolution des AGN Bias mit der Rotverschiebung bis $z < 2$ hin, während gleichzeitig die DMH Masse konstant ($\sim 10^{13}\ \mathrm{h}^{-1}M_{\odot}$) bleibt, d.h., *Röntgen-AGN befinden sich bevorzugt in dichten Umgebungen typisch für Galaxiengruppen.* Die theoretischen Modelle in denen eine Quasar-Phase durch große Galaxienverschmelzungen ausgelöst wird können die für XMM-COSMOS AGN gefundenen hohen Bias-Faktoren und DMH Massen nicht reproduzieren. *Diese Ergebnisse erweitern bis zu $z \sim 2.2$ die Aussage, dass für moderat leuchtkräftige Röntgen-AGN Verschmelzungen von ähnlich großen Galaxien nur einen geringen Beitrag leisten und möglicherweise sekulare Prozesse, wie Gezeiten-Störungen oder Akkretionsscheibeninstabilitäten, bedeutender sind.*

Im zweiten Teil meiner Arbeit präsentiere ich die erste direkte Messung der mittleren HOD von röntgen-selektieren AGN im COSMOS-Feld bei $z \leq 1$, basierend auf der Assoziation von 41 XMM und 17 Chandra-COSMOS AGN mit Galaxien in 189 XMM und Chandra-detektierten Galaxiengruppen. Durch die Unterteilung der Galaxiengruppenbesetzungsbeiträge von AGN in Satelliten- und Zentralgalaxien zeige ich, dass die mittlere HOD von AGN in Zentralgalaxien durch eine geglättete Stufenfunktion bei logM>logM_{min} $[M_{\odot}]$= 12.75(12.10, 12.95) modelliert werden kann, während die AGN HOD der Satellitengalaxien langsamer (mit einem Potenzgesetz der Steigung $\alpha < 0.6$) mit der Halomasse ansteigt als der lineare Anteil ($\alpha = 1$). Dies bedeutet, das die mittlere Anzahl von AGN in den Satellitengalaxien nicht nur von der Halomasse abhängt, wie es auch für Satellitengalaxien beobachtet wurde.

Contents

1 Introduction **1**
1.1 Scientific Rationale . . . 1
1.2 Active Galactic Nuclei . . . 3
1.3 X-ray AGN Host Galaxies . . . 5
1.4 X-ray Groups of Galaxies . . . 6
1.5 AGN in Dense Environment . . . 8
1.6 COSMOS Survey . . . 11
1.7 Overview of the Thesis . . . 12

2 Clustering of AGN **13**
2.1 Large scale structure of the Universe . . . 13
2.2 The Two-Point Correlation Function . . . 15
2.3 Angular Clustering . . . 16
2.4 Projected two-point CF . . . 18
2.5 Bias Factor . . . 20
2.6 Clustering of X-ray AGN: Overview . . . 21

3 Halo Model Approach **29**
3.1 The standard cosmological paradigm . . . 29
3.2 Perturbation Density Field . . . 32
3.3 Dark Matter Halo Properties . . . 34
3.3.1 Halo Mass Function and Halo Bias . . . 35
3.3.2 Halo density profile . . . 36
3.4 Biased Galaxy Formation . . . 37
3.5 Halo Occupation Distribution . . . 38
3.6 Interpreting the two-point CF with the halo model . . . 41

4 Redshift Evolution of the AGN Bias in the COSMOS field **43**
4.1 Introduction . . . 43
4.2 AGN Catalog . . . 45
4.3 Random Catalog . . . 46
4.4 Two-point Statistics . . . 49

4.5 ACF: Standard Approach 51
4.6 Solving for Sample Variance using HOD 54
4.7 Measurements 58
4.8 Redshift Evolution of the AGN Bias 60
4.9 Discussion 62
4.9.1 Which DM halos host X-ray AGN? 62
4.9.2 Optically selected vs X-ray selected AGN 64
4.9.3 External vs Internal Triggering 66
4.10 Conclusions 70

5 Halo Occupation Distribution of AGN in the COSMOS field 73
5.1 Introduction 73
5.2 X-ray Galaxy Groups in the COSMOS Field 77
5.3 X-ray AGN in the COSMOS Field 79
5.4 AGN in galaxy groups 80
5.5 Halo Mass Function and AGN HOD 82
5.6 Two-point Statistics 89
5.7 Galaxy Groups ACF 90
5.8 X-ray AGN-Galaxy Groups Cross-Correlation 92
5.9 The bias factor in the HOD model 92
5.10 Discussion 94
5.11 Conclusions 96

Summary 99

Acknowledgments 101

Bibliography 103

Chapter 1

Introduction

1.1 Scientific Rationale

In the past decade it has become clear that every galaxy with a bulge-like component hosts a super massive black hole (SMBH, $M_{BH} > 10^6 M_\odot$) at its centre and that black holes reach high masses via one or more phases of intense accretion activity shining as Active Galactic Nuclei (AGN). This means that most galaxies have likely had one or more brief AGN periods, shining mostly from the power emitted by a thin, viscous, accretion disk orbiting the central SMBH (Shakura & Sunyaev, 1976). Such a disk produces a high amount of X-rays both from its hot inner regions (as far as the soft X-ray emission is concerned) and from a non thermal source which is supposed to be the primary source of X-rays (both soft and hard).

The co-evolution AGN-galaxy is motivated by the observed correlation between the mass of the central SMBH and the stellar velocity dispersion in the bulge (Gebhardt et al., 2000; Ferrarese & Ford, 2005), lending strong evidence to an interaction or feedback mechanism between the SMBH and the host galaxy. The specific form of the feedback mechanism, as well as the details of the AGN triggering, accretion, and fueling mechanisms, remains unclear. Given the fundamental link between SMBHs and bulge growth as well as the importance of AGN feedback for the evolution of the host galaxy, a key unresolved issue in astrophysics is determining the source of gas that fuels the growth of SMBHs and resultant nuclear activity, and what triggers the gas inflow. As well as the previously described merger paradigm, secular processes including bar-driven gas inflows (Kormendy & Kennicutt, 2004; Jogee et al., 2006), disk instabilities, stochastic collisions with molecular clouds (Hopkins et al., 2006), stellar winds from evolved stars (Ciotti & Ostriker, 1997) and tidal disruption of stars by a dormant BH (e.g., Hills, 1975; Rees, 1988), have all been proposed as means to supply gas onto SMBHs and trigger their activity.

Different cosmological simulations address possible scenarios for the co-evolution of AGN and their host galaxies (e.g., Kauffmann et al., 2000; Di Matteo et al., 2005; Cattaneo et al., 2006). Large volume, high-resolution simulations that include physical prescriptions for galaxy evolution and AGN feedback make predictions for the spatial clustering and large-scale environments of AGN and galaxies (Springel, 2005; Colberg & Di Matteo, 2008; Bonoli et al., 2009). Observed clustering measurements of AGN can be used to test these theoretical models, put constraints on the feedback mechanisms, identify the properties of the AGN host galaxies, and understand the accretion processes onto SMBHs and their fueling mechanism.

X-ray surveys allow us to identify AGN activity without contamination from the emission of the host galaxy, i.e., therefore efficiently detecting even low luminosity AGN. Moreover, X-ray surveys are exceptionally powerful tools for studying the evolution of black holes and their host galaxies, by detecting large numbers of AGN over a wide range of redshifts and cosmic environments from voids to groups and clusters. As an example, deep XMM-Newton and Chandra surveys detected >1500 AGN deg^{-2} against the <500 AGN deg^{-2} of optical surveys (Brandt & Hasinger, 2005).

In the past years the limited sample size of X-ray selected AGN prevented clustering analyses at a comparable detail as for optically selected objects. In particular, the lack of dedicated optical follow-up programs of X-ray sources providing large samples with spectroscopic measurements, has not allowed accurate estimates of the spatial clustering of X-ray selected AGN, limiting most studies to angular clustering. Now a large number of X-ray surveys provide larger samples of sources over wide sky areas and with different limiting fluxes. They allow studies of clustering of AGN in different redshift and luminosity regimes. A few examples of such surveys are X-Bootes (Murray et al., 2005), XMM-LSS (Pierce et al., 2007), Extended CDFS (Lehmer et al., 2005), AEGIS (Nandra et al., 2005) and XMM-COSMOS (Hasinger et al., 2007). One of these samples, the XMM survey in the two square degree COSMOS field, has been specifically designed to study with the best statistics the clustering of X-ray selected AGN.

Clusters and groups of galaxies host a wide diversity of galaxy populations, they are therefore perfect laboratories to study the AGN/galaxy co-evolution. In particular, the distribution of AGN within galaxy groups can be translated in how AGN occupy dark matter halos with different masses. The total mass of galaxy groups can be estimated via gravitational lensing and the distribution of AGN within halos can be investigated in groups by means of the distribution of the AGN host galaxies. The separation of the contribution to the occupation of galaxy groups from AGN in satellite or central galaxies can advance our understanding of the AGN evolution and is related to the mechanism of AGN activation.

Moreover, the study of AGN within galaxy groups provides additional information about how galaxies and BH co-evolve in dense environments. In fact the

Figure 1.1 : *Chandra* image of NGC 4151, a spiral galaxy with an actively growing SMBH at its center. This composite image contains X-rays (blue), optical data (yellow), and radio emission (red). [Credits: *Chandra* Archive]

physical processes that drive galaxy evolution, such as the available cold gas to fuel star formation and the BH growth, are substantially different in groups and clusters compared to the field. Many studies over the past decade have presented an evidence that AGN at z $\sim$ 1 are more frequently found in groups compared to the field (Georgakakis et al., 2009; Arnold et al., 2009). In addiction, X-ray observations reveal that a significant fraction of high-z clusters show overdensities of AGN in their outskirts (Cappi et al., 2001; Ruderman et al., 2005; Cappelluti et al., 2005).

1.2 Active Galactic Nuclei

A primary complication for understanding the nature of AGN and their impact on galaxy evolution is the observed zoo of AGN populations and classes, for which the emission from accretion may be visible over wavelengths from radio to X-rays and occurs over a wide range of Eddington ratios and in different host galaxies. Although many aspects remain poorly understood, a unified scheme (Urry & Padovani, 1995) has been developed which can explain many properties of these diverse classes of AGN and spectral properties in the different wavelength bands.

In this scheme, AGN derive their extraordinary luminosities through energy

release by matter accreting towards and falling into a central SMBH with M $\sim 10^{8-9} M_{\odot}$; this process release energy as UV, X-ray and γ-ray emission. Fig. 1.1 shows a composite image (from X-ray to radio emission) of NGC 4151, a spiral galaxy with an actively growing SMBH at its center. The material orbiting closer to the BH is photoionised, producing the broadened emission lines characteristic of the broad line region (BLR). At this distance from the black hole orbital speeds are several thousand kilometres per second. The clouds are fully exposed to the intense radiation from the engine and are heated to a high temperature.

Broad lines are not seen in every AGN. The general belief among astronomers is that every AGN has a broad-line region, but in some cases our view of the clouds is obscured by a dust torus, so broad lines do not appear in the spectrum. According to the model, in some subclasses of AGN, relativistic jets of material are ejected in the poleward direction and may escape into intergalactic space as the classical double lobed radio sources. Also lying generally in the poleward directions, but at a distance of up to several kpc, are found the narrow line regions (NLR) excited either by photoionisation from the UV continuum of the central source or by shock excitation related to the jets. The model places the NLR much further out from the central engine where orbital speeds are lower (10^{2-3} km s^{-1}). An important consequence of the NLR being outside the dust torus is that they are always in view, so narrow lines will be seen even if the broad-line emitting gas is obscured.

AGN are therefore generally classified by differences in their optical spectra. Type 1 or broad-line AGN have broad ($\nu_{FWHM} \gtrsim 1000$ km s^{-1}) emission lines superimposed on continua in the UV/optical and are the most luminous sources in the sky. Type 2 or narrow-line AGN lack broad emission lines and have weaker continua (frequently dominated by their host galaxies), but have strong narrow emission lines, especially from forbidden transitions. Historically, Type 2 AGN have been described as obscured version of Type 1 AGN, with the broad emission line region hidden behind the partially opaque torus. However, recent observations have revealed several serious limitations of a simple unified model based solely on geometric obscuration.

Selection of AGN via X-ray emission is in principle the cleanest and reliable approach, at least for $L_X \gtrsim 10^{42}$ erg s^{-1} where the identification is unambiguous. However, at lower luminosities, other plausible sources of X-ray emission such as populations of low-mass X-ray binaries, thermal emission from a hot halo of diffuse gas, or even star-formation must be considered and ruled out.

The detection of Compton thick AGN in the X-ray band is very difficult. These objects are supposed to be X-ray emitting SMBH surrounded by obscuring dust with a column density $nH > \sigma_T^{-1}$; yielding an optical depth for Compton scattering $\tau_C = 1$. This causes most of the light below 5 keV to be completely absorbed, making the detection of these objects very difficult in the energy range of focusing X-ray telescopes. Being very faint in the 0.5-10 keV energy band, at the flux limit

of the modern surveys, the fraction of Compton thick AGN observed up to now is of the order of 5-6%.

The hard X-ray emission is directly associated with accretion close in to the black hole (10-100 gravitational radii). It is produced in the hot corona that surrounds the black hole by the inverse Compton scattering of ultraviolet photons emitted by the accretion disk. In the case of infrared AGN this X-ray radiation is absorbed by the surrounding gas and dust torus, heating it up to $\sim$ 500-1500 K and re-radiating in the mid-infrared. The origin of radio-loud AGN appear somewhat different, the radio emission being synchnotron radiation from AGN-powered jets, rathen than being associated directly to the accretion disk.

Mid-infrared selection identifies dust-obscured AGN, but requires sufficient emission to outshine the overall continuum from the galaxy host and has the difficulty in distinguishing dust heated by the central AGN as opposed to star-formation. The mid-infrared spectral energy distributions (SEDs) of AGN typically show a featureless power-law continuum over $\sim$ 1-3- μm that rises with wavelength, whereas star- formation is characterised by emission predominately in the far-infrared, plus spectral features at 6-20 μm.

1.3 X-ray AGN Host Galaxies

Studying the physical properties of the AGN host galaxies provides complementary clues as to their accretion processes. The stellar masses, bulge masses, or stellar velocity dispersions can be used to estimate the mass of the central SMBH via the known tight correlations, from which the Eddington ratio can be derived. The galaxy colors can be used as a proxy for star formation history, and hence the availability of gas. Finally, galaxy morphologies provide constraints on the stage of any ongoing merging event or the presence of bars or disks required for the related secular processes.

Hickox et al. (2009) found that while radio AGN are hosted mainly by massive, red sequence galaxies, the X-ray and infrared-selected AGN are instead both found in $\sim L^*$ galaxies, with the X-ray population being preferentially green valley objects, while IR AGN are slightly bluer. Haggard et al. (2010) also found that X-ray AGN are much more likely to be located within the blue cloud or green valley than the red sequence. Georgakakis et al. (2009) suggest that the color distribution for X-ray AGN has not evolved between z $\sim$ 0.8 and the present day.

Moreover several works show that the morphologies of the AGN host galaxies do not present a preference for merging systems. Cisternas et al. (2011) found that <15% of X-ray AGN in the HST-COSMOS field (with $L_X \sim 10^{43.5}$ erg s^{-1}, z $\sim$ 0.3-1.0) showed any signs of distortions indicative of recent mergers, and indeed found no statistical difference in the distortion fractions between X-ray AGN and inactive galaxies, indicating that major mergers are not the most relevant mechanism for

the triggering of X-ray AGN at z $\sim$ 1. Instead they found that over 55% of the X-ray AGN are hosted by disk galaxies, and suggest that the bulk of black hole accretion occurs through internal secular fuelling processes and minor mergers.

Similarly, Griffin & Stern (2010) found while radio-loud AGN are mostly hosted by early-type galaxies, X-ray AGN are mostly either disk-dominated (31% - 46%) or unresolved point sources (31% - 61%) with few (9% - 21%) hosted by bulge-dominated systems. In the local universe, Koss et al. (2010) found that X-ray AGN are $\sim$5-10 times more likely to be hosted in spirals ($\sim$40%) or mergers ($\sim$20%) than inactive galaxies of the same stellar masses.

Silverman et al. (2011) used a sample of kinematic pairs identified from the zCOSMOS 20k bright catalog and *Chandra* observations that indicate those harboring AGN. They observed an enhancement of AGN in close pairs, which is in broad agreement with merger-driven models (Hopkins et al., 2008) of black hole growth. But they found that only $\sim$18% of such AGN are the result of interactions. Ellison et al. (2011) did show that some of the nuclear activity is triggered by interactions, showing that the AGN fraction increases by up to $\sim$ 2.5 times for galaxies in close pairs with projected separations <10 kpc, and that this enhancement in nuclear activity is greatest for equal-mass galaxy pairings.

At high redshift (z $\sim$2) recent findings of Schlegel et al. (2001); Rosario et al. (2011), who examined a smaller sample of AGN in the ERS-II region of the GOODS-South field, inferred that late-type morphologies are prevalent among the AGN hosts. The role that major galaxy mergers play in triggering AGN activity at $1.5 < z < 2.5$ was also studied in the Chandra Deep Field South (CDF-S). At z = 1.5-3, Schawinski et al. (2011) showed that for X-ray selected AGN in the CDF-S and with typical luminosities of 10^{42} erg s^{-1}< L_X < 10^{44} erg s^{-1} the majority (80%) of the host galaxies of these AGN have low Sersic indices indicative of disk-dominated light profiles, suggesting that secular processes govern a significant fraction of the cosmic growth of black holes. Later, Kocevski et al. (2011) found that X-ray selected AGN at z $\sim$ 2 do not exhibit a significant excess of distorted morphologies while a large fraction reside in late-type galaxies. They also suggested that these late-type galaxies are fueled by the stochastic accretion of cold gas, possibly triggered by a disk instability or minor interaction. Moreover, it is believed that major mergers dominate at high redshifts and bright luminosities, while minor interactions or bar instabilities or minor tidal disruptions are important at low redshifts ($z < 1$) and low luminosities ($L_{BOL} < 10^{44}$ erg s^{-1}) (see discussion in Hasinger et al. 2008, Hopkins & Henquist 2009).

1.4 X-ray Groups of Galaxies

Galaxy clusters are detected as bright, extended sources in the X-ray sky. This strong emission is due to thermal bremmsstrahlung (and some line emission) of

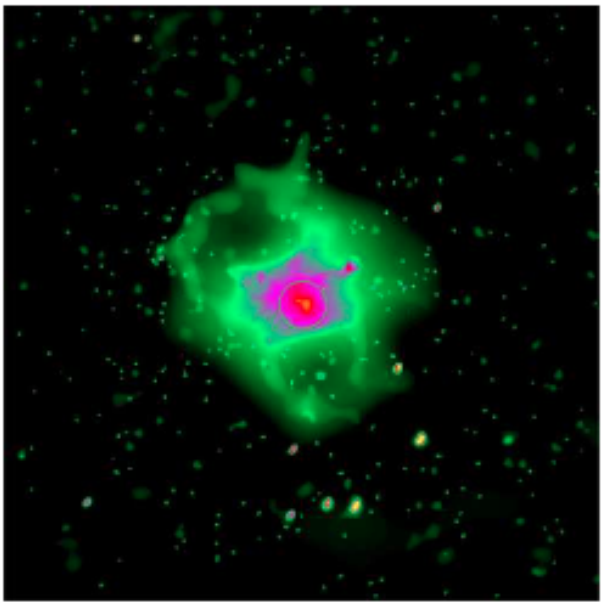

Figure 1.2 : *Chandra* image of HCG 62, a galaxy group with $L_X \sim 10^{43}$ erg s^{-1}. [Credit: Chandra Archive]

the gas in the potential well of the system, with temperature of 10^{7-8} K. The X-ray hot gas can reach up to 15% of the total mass of the system, with a typical X-ray luminosity of $\sim 10^{44} - 10^{45}$ erg s^{-1}. In the 1990s, thanks to the launch of two important X-ray telescopes, ROSAT and ASCA, it has been established that also many less massive groups of galaxies emit in the X-ray band. Since groups have a lower velocity dispersion, thus a lower temperature of gas, the abundant elements are not fully ionized and part of the flux is due to line emission. The typical X-ray luminosity of galaxy groups is $\sim 10^{41} - 10^{43}$ erg s^{-1}, while the typical mass is approximately 10^{13}M$_\odot$ (one hundreds time less massive than a cluster like Coma). This mass is directly related to the gas temperature. In particular, when the condition of hydrostatic equilibrium and spherical symmetry are met, the total mass at any radius is a simple function of the gas temperature.

Self-similar models predict simple scaling relation between basic cluster properties and the total mass. Three important correlations are the X-ray luminosity–temperature (L_X-T), mass-temperature (M_Δ-T) and entropy-temperature (S-T) relations. These relations are of great importance for the investigation of groups and cluster of galaxies. In particular the M-T relation constraints the scale of a system and it gives a direct measurement of the system mass when its temperature is known.

The identification of galaxy groups is more difficult than for rich galaxy clusters.

The reason is that groups are fainter at all wavelengths and contain a lower number of galaxies than rich clusters, therefore they are less well defined in the sky, and more affected by contamination from fore/background galaxies. Therefore deeper galaxy surveys are required to identify groups, especially at non local redshifts. The largest optical catalogue of galaxy groups is the one compiled from the Sloan Digital Sky Survey (SDSS; Yang et al. 2007), which contains ∼8000 groups with more than 3 galaxies at $z < 0.2$.

However, the most robust way to identify gravitationally bound groups is via detection of their extended X-ray emission. X-ray information enables us to easily obtain an estimate of the total mass of the system, and thus to define a physically motivated radius to characterise a cluster/group (the virial radius). However, since X-ray groups have a luminosity which is ∼10-100 times lower than that of rich clusters and that X-ray flux suffers from a large dimming effect with redshift, a robust identification of groups requires a deep X-ray survey on an area which is representative of the large scale structure of the Universe at different redshifts (100 h^{-1}Mpc; 1 degree at z∼1 corresponds to ∼40 h^{-1}Mpc). Currently, the largest catalogues of X-ray selected systems at masses lower than 10^{14} $M_{\odot}$ do not contain more than a few tens of systems (e.g. RASSCALS Mahdavi et al. 2000; Heldson & Ponman 2000). The number decreases drastically at high redshift (e.g. AEGIS Jeltema et al. 2007; Gonzalez et al. 2007; CNOC2 Finoguenov et al. 2009).

The only survey, up to now, which combines a large area, deep X-ray data, deep multiwavelength information and high resolution imaging is the COSMOS survey (Scoville et al., 2007). Fig. 1.2 shows the *Chandr*a image of the compact galaxy group known as HCG 62, with remarkable detail and complexity in the central region. The range of X-ray surface brightness is represented in this image by various colors: green depicts the lower-brightness regions while purple and reddish indicate increasing X-ray intensity.

1.5 AGN in Dense Environment

The environments and AGN host galaxies can provide important clues to understanding both the accretion processes powering the AGN (e.g., whether it is fuelled by the accretion of hot or cold gas), and its subsequent impact on the galaxy host in terms of building up its bulge or quenching its star formation. The distribution of AGN in galaxy clusters and groups provides a fundamental test for those accretion processes which require a ready supply of cold gas in the host galaxy, such as secular bar/disk instabilities (Hopkins et al., 2006) or the merger of two gas-rich galaxies.

In the case of the merger paradigm, nuclear activity should be strongly suppressed within rich clusters. Here the encounter velocities of galaxies are much greater than their internal velocity dispersions, preventing their coalescence, in

spite of the high galaxy densities (Aarseth & Fall, 1980). Gas-rich mergers should instead be most frequent in galaxy groups (Hopkins et al., 2008) and in the cluster outskirts, where many galaxies (including gas-rich ones), under the influence of the clusters tidal field, are part of a convergent flow resulting in enhanced interactions between neighbors (van den Weygaert & Babul, 1994). Since these galaxies are falling into the cluster for the first time, their gas contents have yet to be affected by their passage through the dense intracluster medium (ICM). Galaxy harassment due to frequent high-speed fly-by interactions has also been proposed as a means of triggering nuclear activity, by driving dynamical instabilities that efficiently channel gas onto the SMBHs (Moore et al., 1996) although again this requires the host galaxy to contain a gas reservoir. There is, however, also another important effect: the ram-pressure stripping of gas from galaxies falling into clusters and groups. When the density of the hot gas is large enough, the cold gas in the infalling galaxies is stripped out, thus quenching any cold-gas accretion. Then accretion can only happen in presence of hot gas.

The trends predicted by associating X-ray AGN to gas-rich galaxies could be diluted, however, or even reversed, by the tendency of X-ray AGN (above a given L_X) to be hosted by the most luminous galaxies (Sivakoff et al., 2008; Tasse et al., 2011) a population which is most centrally concentrated within clusters (Lin et al., 2004; Thomas & Katgert, 2006). Both Sivakoff et al. (2008) and Haggard et al. (2010) show the X-ray AGN fraction to increase by an order of magnitude from low-mass populations (M $\sim 10^{10} M_{\odot}$) to the most massive galaxies (M $\sim 10^{11} M_{\odot}$) in both cluster and field populations. These trends act as physical selection effect, as the most luminous galaxies are more likely bulge-dominated and hence have higher black hole masses. Thus, for the same accretion rate relative to the Eddington limit, the higher mass galaxy will likely have a higher X-ray luminosity, and be more likely detected above a fixed L_X limit.

The earliest works identified AGN via their optical emission lines (Gisler , 1978; Dressler et al., 1985) finding them to be much rarer among cluster members than in field samples. Finoguenov et al. (2004) showed in the Coma cluster that the X-ray activity is suppressed with respect to the field by a factor of 5.6, indicating a lower level of X-ray emission for a given stellar mass. Dressler et al. (2005) measured a fraction of AGN in clusters of the order of <1% which is a factor 5 less that what observed in the field. The comparison with studies in other wavelengths gives confusing results. Branchesi et al. (2007) pointed out that radio galaxies are a factor 2 more frequent in clusters than in the field. Ledlow & Owen (1996), making use of the bivariate radio/optical luminosity function, suggest that the fraction and the evolution of radio-galaxies is independent from the environment. More recent studies based on the SDSS have obtained rather conflicting results, with some confirming the previous decline in the AGN fraction with galaxy density (Kauffmann et al., 2004; Mahajan et al., 2010; Hwang et al., 2012), while others

find no apparent trend with environment (Miller et al., 2003; Haines et al., 2009).

The launches of the *Chandra* and XMM X-Ray Observatories have opened up an efficient way of identifying X-ray AGN in and around groups and clusters of galaxies. Statistical analyses of X-ray point sources in cluster fields have reported overdensities of X-ray sources with respect to non-cluster fields (e.g., Cappi et al., 2001; Molnar et al., 2002; Ruderman et al., 2005; Cappelluti et al., 2005; Branchesi et al., 2007). Gilmour et al. (2009) found an excess of ∼1.5 X-ray point sources per cluster within 1 Mpc for a sample of 148 clusters at $0.1 < z < 0.9$. They found the radial distribution of these excess sources to be consistent with a flat radial distribution within 1 Mpc, although they could also be consistent with being drawn from the general cluster galaxy population.

Ruderman et al. (2005) studied the spatial distribution of a sample of 508 soft X-ray sources detected around 51 *Chandra* massive clusters in the redshift range z= 0.3-0.7. The X-ray source surface density in their sample is higher in the inner projected 3.5 Mpc than in the field. They also measured the AGN density profile in the cluster and, surprisingly, detected features in the curve. They detected a strong excess in the inner 0.5 Mpc mainly attributable to the central cD galaxies, a depletion zone around 1.5 Mpc and a secondary excess above 3 Mpc from the cluster center. The secondary excess has been explained with merger-driven accretion onto SMBH at the edge of the cluster. This kind of triggering occurs most likely in low energy collisions favoured in the cluster-field transition region. The depletion zone is explained by the low probability of galaxy merging in regions of high velocity dispertion. Koulouridis & Plionis (2010) instead found that the overdensity of X-ray point sources in 16 clusters at $0.07 < z < 0.28$ to be a factor ∼4 less than that of bright optical galaxies, and concluded that the triggering of luminous ($L_X > 10^{42}$ erg s^{-1}) X-ray AGN to be strongly suppressed in rich clusters.

Unfortunately, all the X-ray results mentioned above suffer from a lack of spectroscopic counterparts, they therefore need more detailed optical/IR follow-up campaigns to clarify the origin of the overdensities. From these analyses it remains unclear whether the X-ray AGN observed in galaxy clusters should be considered to be virialized or an infalling population.

Haines et al. (2012) studied the distribution and host properties of 48 X-ray AGN identified from *Chandra* imaging of 26 massive clusters at $0.15 < z < 0.30$. They found that AGN are clearly dynamically identified with an infalling population. This is manifest by their preferential location along the cluster caustics, complete avoidance of the caustic phase space with low relative velocities and cluster-centric radii, and their high-velocity dispersion and non-Gaussian velocity distribution. The optically selected Type I Seyferts/QSOs in their cluster sample show the same kinematical signatures. These provide a strong observational constraints that X-ray AGN and optically selected Type I AGN/QSOs found in massive clusters are not a virialized population, and few if any can have resided within the

dense ICM for a significant length of time.

1.6 COSMOS Survey

COSMOS is the largest survey ever made using the Hubble Space Telescope (HST, Scoville et al. 2007). It covers 2 square degrees (a square field 1.4°on a side), with an exposure in the I band obtained by a single orbit down to $I_{AB} = 27$ mag. The whole survey consists of data from 640 orbits, over a period of 2 years. The coverage of such a large area enables the sampling of the large scale structure of the Universe, and reduces the cosmic variance as a source of systematic error. Indeed large scale structures as voids, filaments, groups or galaxy clusters occur on scales up to 100 Mpc comoving, and the COSMOS field can adequately map the galaxy evolution for a large range of environments. The COSMOS survey samples $\sim 2 \times 10^6$ galaxies with $I_{AB} < 27$ mag. The field is centred at at $(RA, DEC)_{J2000}$ = (150.1083, 2.210). Here galactic extinction is exceptionally low and uniform (< 20% variation; Sanders et al. 2007), but the infrared background is higher than in dark fields like the Lockman Hole which are not equatorial. On the other hand, it can be observed by telescopes located both in the northern and in the southern hemisphere.

The power of COSMOS resides in coupling the unique imaging resolution of HST (0.05 arcsec) with a multiwavelength coverage from both ground and space based facilities. In particular it guarantees a full spectral coverage, with X-ray (Chandra & XMM-Newton), UV (GALEX), optical (SUBARU), near-infrared (CFHT), mid-infrared (Spitzer), sub-millimetric (MAMBO) and radio (VLA) imaging.
The XMM-Newton wide-field survey in the COSMOS field (hereinafter XMM-COSMOS; Hasinger et al. 2007) is an important step forward in addressing the topics described above. The ∼2 deg^2 area of the HST has been surveyed with XMM-Newton for a total of ∼1.55 Ms during AO3, AO4, and AO6 cycles of XMM observations (Cappelluti et al., 2007, 2009). XMM-COSMOS provides an unprecedented large sample of point-like X-ray sources (∼ 1800), detected over a large, contiguous area, with complete ultraviolet to mid-infrared (including Spitzer data) and radio coverage, and extensive spectroscopic follow-up granted through the zCOSMOS (Lilly et al., 2007, 2009), Magellan/ IMACS (Trump et al., 2007, 2009) and DEIMOS/Keck (Capak et al. 2007) projects. The excellent multi-band photometry available in this area allows a robust photometric redshift estimates (Salvato et al., 2009, 2011) for the faint sources not reachable by optical spectroscopy, thus allowing a virtually complete sample of X-ray sources. The XMM-COSMOS project is described in Hasinger et al. (2007), while the X-ray point source catalog and counts from the complete XMM-COSMOS survey is presented in a companion paper Cappelluti et al. (2009).

1.7 Overview of the Thesis

In this Thesis I focus on the clustering analysis, interpreted through the halo model, of X-ray selected AGN up to z∼2, and on the occupation of X-ray selected galaxy groups by AGN at $z < 1$ in the COSMOS field.

- In Chapter 2 I describe the two-point correlation function as commonly used estimator of the spatial distribution of galaxies in the Universe, giving a detailed overview of previous results on clustering of X-ray selected AGN.

- In Chapter 3 I focus on the Halo Model Approach as powerful method to interpret the clustering of galaxies in terms of relation between galaxies and dark matter halos in which they reside. I describe the halo model formalism and all the theoretical ingredients that describe the dark matter properties and the galaxy bias.

- In Chapter 4 I present the clustering properties of X-ray selected AGN in the COSMOS field as a function of redshift up to z∼2. In particular, I focus on the redshift evolution of the AGN bias and on the corresponding mass of the hosting dark matter halos. I interpret the results in a context of AGN triggering mechanism and in comparison with previous studies.

- In Chapter 5 I perform the first direct measurement of the occupation of X-ray selected galaxy groups by X-ray AGN at $z < 1$ in the COSMOS field, providing a description of how AGN are distributed within dark matter halos with different masses, separating the contribution of AGN among satellite and central galaxies. Together with a comparison with previous results in literature on the halo occupation distribution of AGN, I discuss the implications of my results in terms of AGN formation and evolution. The correlation function of galaxy groups and the cross-correlation function AGN-groups are also presented and modelled following the results on the AGN halo occupation.

Chapter 2

Clustering of AGN

2.1 Large scale structure of the Universe

Galaxies are not uniformly distributed in space. On large scales the Universe displays coherent structures, with galaxies residing in groups and clusters on scales of ∼1-3 h^{-1}Mpc, which lie at the intersections of long filaments of galaxies that are >10 h^{-1}Mpc in length. Vast regions of relatively empty space, known as voids, contain very few galaxies and span the volume in between these structures. This observed large scale structure depends both on cosmological parameters and on the formation and evolution of galaxies. The idea of whether galaxies are distributed uniformly in space can be traced to Edwin Hubble, who used his catalog of 400 "extragalactic nebulae" to test the homogeneity of the Universe (Hubble, 1926), finding it to be generally uniform on large scales. In 1934 using the larger Shapley-Ames sample (Shapley & Ames, 1932), Hubble noted that on angular scales less than ∼10° there is an excess in the number counts of galaxies above what would be expected for a random Poisson distribution, though the sample follows a Gaussian distribution on larger scales (Hubble, 1934). Hence, while the Universe appears to be homogeneous on the largest scales, on smaller scales it is clearly clumpy.

Measurements of large scale structure took a major leap forward with the Lick galaxy catalog produced by Shane & Wirtanen (1967), which contained information on roughly a million galaxies obtained using photographic plates at the 0.5m refractor at Lick Observatory. The statistical spatial distribution of galaxies from this catalog and that of Zwicky et al. (1968) was analyzed by Jim Peebles and collaborators in a series of papers (e.g., Peebles, 1975) that showed that the angular two-point correlation function (defined below) roughly follows a power law distribution over angular scales of ∼0.1°- 5°. These results in part spurred the first large scale redshift surveys, which obtained optical spectra of individual galaxies in order to measure the redshifts and spatial distributions of large galaxy samples. Pioneering work by Gregory & Thompson (1978) mapped the three-dimensional

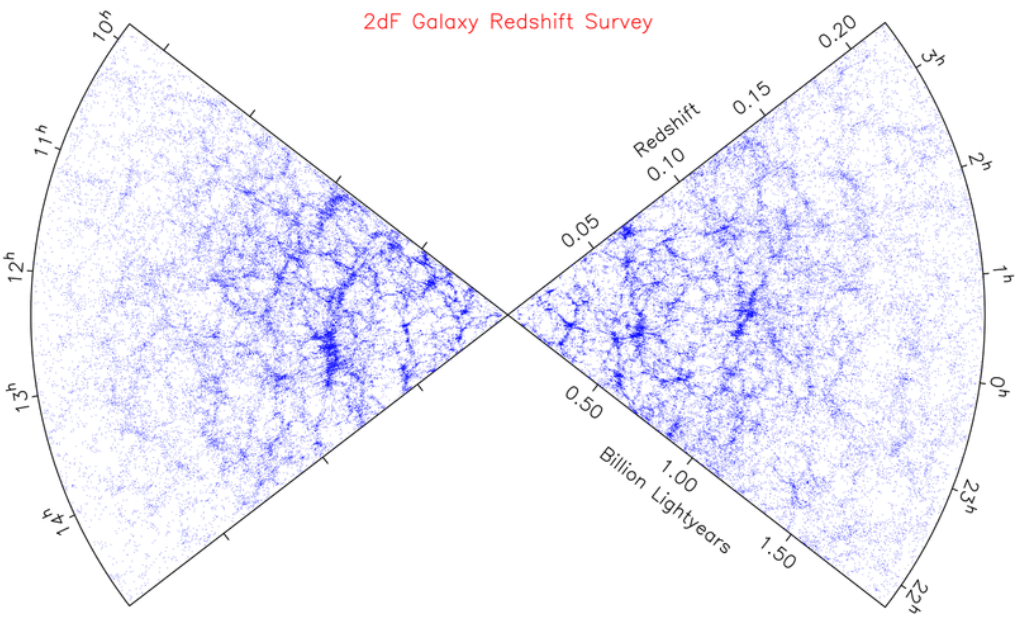

Figure 2.1 : The spatial distribution of galaxies as a function of redshift and right ascension (projected through 3° in declination) from the 2dF Galaxy Redshift Survey. [Credits: Colless et al. 2004]

spatial distribution of 238 galaxies around and towards the Coma/Abell 1367 supercluster. In addition, to surveying the galaxies in the supercluster, they found that in the foreground at lower redshift there were large regions (> 20 Mpc) with no galaxies, which they termed 'voids'.

Redshift surveys have rapidly progressed with the development of multi-object spectrographs, which allow simultaneous observations of hundreds of galaxies, and larger telescopes, which allow deeper surveys of both lower luminosity nearby galaxies and more distant, luminous galaxies. At present the largest redshift surveys of galaxies at low redshift are the Two Degree Field Galaxy Redshift Survey (2dF-GRS, Colless et al., 2001) and the Sloan Digital Sky Survey (SDSS, York et al., 2000), which cover volumes of $\sim 4 \times 10^7$ $h^3 Mpc^{-3}$ and $\sim 2 \times 10^8$ $h^3 Mpc^{-3}$ with spectroscopic redshifts for ∼220,000 and a million galaxies, respectively. These surveys provide the best current maps of large scale structure in the Universe today (see Fig. 2.1), revealing a sponge-like pattern to the distribution of galaxies (Gott et al., 1986). Voids of ∼10 Mpc are clearly seen, containing very few galaxies. Filaments stretching greater than 10 Mpc surround the voids and intersect at the locations of galaxy groups and clusters.

The prevailing theoretical paradigm regarding the existence of large scale structure is that the initial fluctuations in the energy density of the early Universe, seen

as temperature deviations in the cosmic microwave background, grow through gravitational instability into the structure seen today in the galaxy density field. The details of large scale structure – the sizes, densities, and distribution of the observed structure – depend both on cosmological parameters such as the matter density and dark energy as well as on the physics of galaxy formation and evolution. Measurements of large scale structure can therefore constrain both cosmology and galaxy evolution physics.

2.2 The Two-Point Correlation Function

The most commonly used quantitative measure of large scale structure is the two-point correlation function, $\xi(r)$, which traces the amplitude of clustering of objects in the Universe as a function of scale. $\xi(r)$ is defined as a measure of the excess probability dP, above the expected level for an unclustered random Poisson distribution, of finding a pair with an object in the volume dV_1 and another in the volume dV_2, separated by a distance r so that:

$$dP = n^2[1 + \xi(r)]dV_1 dV_2, \tag{2.1}$$

where n is the mean number density of the sample in question (Peebles, 1980). Measurements of $\xi(r)$ are generally performed in comoving space, with r having units of h^{-1}Mpc. The Fourier transform of the two-point correlation function is the power spectrum, which is often used to describe density fluctuations observed in the cosmic microwave background.

To measure $\xi(r)$ of a sample of AGN, one counts pairs as a function of separation and divides by what is expected for an unclustered distribution. To do this one must construct a "random catalog" that has the identical three dimensional coverage as the data – including the same sky coverage and smoothed redshift distribution – but is populated with randomly-distribution points. The ratio of pairs of AGN observed in the data relative to pairs of points in the random catalog is then used to estimate $\xi(r)$. Several different estimators for $\xi(r)$ have been proposed and tested. An early estimator that was widely used is from Davis & Peebles (1983):

$$\xi = \frac{n_R}{n_D}\frac{DD}{DR} - 1, \tag{2.2}$$

where DD and DR are counts of pairs of AGN (in bins of separation) in the data catalog and between the data and random catalogs, and n_D and n_R are the mean number densities of galaxies in the data and random catalogs. Hamilton (1993) later introduced an estimator with smaller statistical errors,

$$\xi = \frac{DD\ RR}{(DR)^2} - 1, \tag{2.3}$$

where RR is the count of pairs of AGN as a function of separation in the random catalog. The most commonly-used estimator is from Landy & Szalay (1993):

$$\xi = \frac{1}{RR}\left[DD\left(\frac{n_R}{n_D}\right)^2 - 2DR\left(\frac{n_R}{n_D}\right) + RR\right]. \tag{2.4}$$

where DD, DR and RR are the normalized number of data-data, data-random, random-random AGN pairs, respectively. This estimator has been shown to perform as well as the Hamilton estimator (Eq. 2.3), and while it requires more computational time it is less sensitive to the size of the random catalog and handles edge corrections well, which can affect clustering measurements on large scales (Kerscher et al., 2000).

As can be seen from the form of the estimators given above, measuring $\xi(r)$ depends sensitively on having a random catalog which accurately reflects the various spatial and redshift selection affects in the data. Several observational biases must be taken into account when generating a random sample of objects in a X-ray flux limited survey. In particular, in order to reproduce the selection function of the survey, one has to carefully reproduce the space and flux distributions of the sources, since the sensitivity in X-ray surveys is not homogeneous on the detector and therefore on the sky. Moreover in several case optical follow-up of the X-ray source is not 100% complete, therefore one must carefully reproduce the mask effect. What is usually done is that to create random samples in 3D, sources are placed at the same angular position of the real sources and redshift are randomly drawn from a smoothed redshift distribution of the real sources. If instead the spectral completeness is close to 100% then the right procedure is to occupy the survey volume with random sources drawn from a L-z dependent luminosity function and check if they would be observable using a sensitivity map.

If one is measuring a full three-dimensional correlation function (discussed below) then the random catalog must also accurately include the redshift selection of the data. The random catalog should also be large enough to not introduce Poisson error in the estimator. This can be checked by ensuring that the RR pair counts in the smallest bin are high enough such that Poisson errors are subdominant.

2.3 Angular Clustering

The spatial distribution of AGN can be measured either in two dimensions as projected onto the plane of the sky or in three dimensions using the redshift of each AGN. As it can be observationally expensive to obtain spectra for large samples of (particularly faint) AGN, redshift information is not always available for a given sample. One can then measure the two-dimensional projected angular correlation function $\omega(\theta)$, defined as the probability above Poisson of finding two AGN with

an angular separation θ:

$$dP = N[1 + \omega(\theta)]d\Omega \tag{2.5}$$

where N is the mean number of AGN per steradian and dΩ is the solid angle of a second AGN at a separation θ from a randomly chosen AGN.

The relation between the 2D (angular) auto-correlation function (ACF) and the 3D ACF is expressed by the Limber equation (e.g., Peebles, 1980). Under the assumption that the scale length of the clustering is much smaller than the distance to the object, this reduces to:

$$w(\theta)N^2 = \int \left(\frac{dN}{dZ}\right)^2 \int \xi\left(\sqrt{[d_A(z)\theta^2) + l^2(1+z]}\right)\left(\frac{dl}{dz}\right)^{-1} dl\ dz, \tag{2.6}$$

where $dA(z)$ is the angular distance, N is the total number of sources, and dN/dz is the redshift distribution (per z) of the sources. The redshift evolution of the 3D correlation function is customarily expressed by:

$$\xi(r, z) = \left(\frac{r}{r_0}\right)^{-\gamma}(1+z)^{-3-\epsilon+\gamma}, \tag{2.7}$$

where ϵ = -3 and $\epsilon = \gamma$ - 3 correspond to the case where the correlation length is constant in physical and comoving coordinates, respectively. In these notations, the zero-redshift 3D correlation length r_0 can be related to the angular correlation length θ_0 by:

$$\begin{aligned} r_0^\gamma &= (N^2/S)\theta_0^{\gamma-1}, \\ S &= H_\gamma \int \left(\frac{dN}{dZ}\right)^2 \left[\frac{c\ d\tau(z)}{dz}\right]^{-1} d_A^{1-\gamma}(1+z)^{-3-\epsilon}dz, \\ H_\gamma &= \frac{\Gamma[(\gamma-1)/2]\Gamma(1/2)}{]\Gamma(1/2)}, \end{aligned} \tag{2.8}$$

where τ(z) is the look-back time. We also define the comoving correlation length

$$r_0(z_{\bar{eff}}) = r_0(1 + z_{\bar{eff}})^{-3-\epsilon+\gamma}, \tag{2.9}$$

at the effective redshift $z_{\bar{eff}}$, which is the median redshift of the contribution to the angular correlation (the integrand of the second term).

If the redshift distribution of sources, dN/dz, is well known, then the amplitude of $\omega(\theta)$ can be predicted for a given power-law model of $\xi(r)$, such that measurements of $\omega(\theta)$ can be used to place constraints on the evolution of $\xi(r)$.

Interpreting angular clustering results can be difficult, however, as there is a degeneracy between the inherent clustering amplitude and the redshift distribution of the sources in the sample. For example, an observed weakly clustered signal

projected on the plane of the sky could be due either to the AGN population being intrinsically weakly clustered and projected over a relatively short distance along the line of sight, or it could result from an inherently strongly clustered distribution integrated over a long distance, which would wash out the signal. The uncertainty on the redshift distribution is therefore often the dominant error in analyses using angular clustering measurements. The assumed AGN redshift distribution (dN/dz) has varied widely in different studies, such that similar observed angular clustering results have led to widely different conclusions. A further complication is that each sample usually spans a large range of redshifts and luminosities, such that the mean intrinsic luminosity of the AGN is changing with redshift within a sample, which can hinder interpretation of the evolution of clustering measured in $\omega(\theta)$ studies.

Because of the sensitivity of the inferred value of r_0 on the redshift distribution of sources, it is preferable to measure the three dimensional correlation function. While it is much easier to interpret three dimensional clustering measurements, in cases where it is still not feasible to obtain redshifts for a large fraction of galaxies in the sample, angular clustering measurements are still employed (Elyiv et al., 2012).

2.4 Projected two-point CF

Measurements of the two-point correlation function use the redshift of an AGN, not its distance, to infer its location along the line of sight. This introduces two complications: one is that a cosmological model has to be assumed to convert measured redshifts to inferred distances, and the other is that peculiar velocities introduce redshift space distortions in ξ parallel to the line of sight (Sargent & Turner, 1977). On the first point, errors on the assumed cosmology are generally subdominant, so that while in theory one could assume different cosmological parameters and check which results are consistent with the assumed values, that is generally not necessary. On the second point, redshift space distortions can be measured to constrain cosmological parameters, and they can also be integrated over to recover the underlying real space correlation function.

On small spatial scales ($\lesssim 1$ h^{-1}Mpc), within collapsed virialized overdensities such as groups and clusters, galaxies have large random motions relative to each other. Therefore while all of the galaxies in the group or cluster have a similar physical distance from the observer, they have somewhat different redshifts. This causes an elongation in redshift space maps along the line of sight within overdense regions, which is referred to as "Fingers of God". The result is that groups and clusters appear to be radially extended along the line of sight towards the observer. Redshift space distortions are also seen on larger scales ($\gtrsim 1$ h^{-1}Mpc) due to streaming motions of galaxies that are infalling onto structures that are still collapsing. Adjacent galaxies will all be moving in the same direction, which leads

to coherent motion and causes an apparent *contraction* of structure along the line of sight in redshift space (Kaiser, 1987), in the opposite sense as the "Fingers of God".

To separate the effects of redshift distortions, the spatial correlation function is measured in two dimensions r_p and π, where r_p and π are the projected comoving separations between the considered objects in the directions perpendicular and parallel, respectively, to the mean line-of-sight between the two sources. Following Davis & Peebles (1983), r_1 and r_2 are the redshift positions of a pair of objects, s is the redshift-space separation $s = (r_1 - r_2)$, and $l = \frac{1}{2}(r_1 + r_2)$ is the mean distance to the pair. The separations between the two considered objects across r_p and π are defined as:

$$\pi = \frac{\mathbf{s} \cdot \mathbf{l}}{|\mathbf{l}|} \tag{2.10}$$

$$r_p = \sqrt{(\mathbf{s} \cdot \mathbf{s} - \pi^2)} \tag{2.11}$$

Redshift space distortions only affect the correlation function along the line of sight, so we estimate the so-called projected correlation function $w_p(r_p)$ (Davis & Peebles, 1983):

$$w_p(r_p) = 2 \int_0^{\pi_{max}} \xi(r_p, \pi) d\pi \tag{2.12}$$

where $\xi(r_p, \pi)$ is the two-point correlation function in term of r_p and π, measured using the Landy & Szalay (1993, LS) estimator.

An example of a measurement of $\xi(r_p, \pi)$ is shown in Fig. 2.2. Plotted is ξ as a function of separation r_p (defined in this figure to be σ) across and π along the line of sight. What is usually shown is the upper right quadrant of this figure, which here is reflected about both axes to emphasize the distortions. Contours of constant ξ follow the color-coding, where yellow corresponds to large ξ values and green to low values. On small scales across the line of sight (r_p or $\sigma \lesssim 2$ h^{-1}Mpc) the contours are clearly elongated in the π direction; this reflects the "Fingers of God" from galaxies in virialized overdensities. On large scales across the line of sight (r_p or $\sigma \gtrsim 10$ h^{-1}Mpc) the contours are flattened along the line of sight, due to "the Kaiser effect". This indicates that galaxies on these linear scales are coherently streaming onto structures that are still collapsing.

If $\xi(r)$ is modeled as a power-law, $\xi(r) = (r/r_0)^{-\gamma}$, then r_0 and γ can be readily extracted from the projected correlation function, $w_p(r_p)$, using an analytic solution to Eq. 2.12:

$$w_p(r_p) = r_p \left(\frac{r_0}{r_p}\right)^{\gamma} \frac{\Gamma(\frac{1}{2})\Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})}, \tag{2.13}$$

where Γ is the usual gamma function. A power-law fit to $w_p(r_p)$ will then recover r_0 and γ for the real-space correlation function, $\xi(r)$. In practice, Eq. 2.12 is not

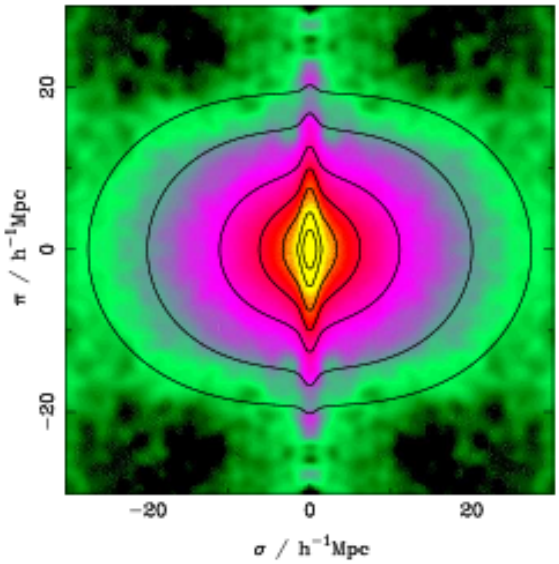

Figure 2.2 : The two-dimensional redshift space correlation function from 2dFGRS (Peacock et al., 2001). Shown is $\xi(r_p, \pi)$ (in the figure σ is used instead of r_p), the correlation function as a function of separation across (σ or r_p) and along (π) the line of sight. Contours show lines of constant ξ at ξ =10, 5, 2, 1, 0.5, 0.2, 0.1. Data from the first quadrant (upper right) are reflected about both the σ and π axes, to emphasize deviations from circular symmetry due to redshift space distortions.

integrated to infinite separations. Often values of $\pi_{\rm max}$ for galaxies and AGN are ∼40-80 h^{-1}Mpc, which includes most correlated pairs.

2.5 Bias Factor

It was realized decades ago that the spatial clustering of observable galaxies and then AGN need not precisely mirror the clustering of the bulk of the matter in the Universe. This "bias" – the relationship between the spatial distribution of galaxies and the underlying matter density field (which is dominated by dark matter) – is a result of the varied physics of galaxy formation which can cause the spatial distribution of baryons to differ from that of dark matter. The concept of galaxies being a biased tracer of the underlying total mass field was introduced by Kaiser (1984) in an attempt to reconcile the different clustering scale lengths of galaxies and rich clusters, which could not both be unbiased tracers of mass. Kaiser (1984) show that clusters of galaxies would naturally have a large bias as a result of being rare objects which formed at the highest density peaks of the

mass distribution, above some critical threshold. This idea is further developed analytically by Bardeen et al. (1986) for galaxies, who show that for a Gaussian distribution of initial mass density fluctuations, the peaks which first collapse to form galaxies will be more clustered than the underlying mass distribution.

The standard approaches used to estimate the bias factor of AGN are based on the power-law fit parameters of the AGN correlation function. This method assumes that the projected correlation function is well fitted by a power-law and the bias factors are derived from the best fit parameters r_0 and γ of the clustering signal at large scale. Using the power-law fit one can estimate the AGN bias factor using the power-law best fit parameters:

$$b_{PL} = \sigma_{8,AGN}(z)/\sigma_{DM}(z) \tag{2.14}$$

where $\sigma_{8,AGN}(z)$ is the rms fluctuations of the density distribution over the sphere with a comoving radius of 8 h^{-1}Mpc, $\sigma_{DM}(z)$ is the dark matter correlation function evaluated at 8 h^{-1}Mpc, normalized to a value of $\sigma_{DM}(z = 0) = 0.8$. For a power-law correlation function this value can be calculated by (Peebles, 1980):

$$(\sigma_{8,AGN})^2 = J_2(\gamma)(\frac{r_0}{8Mpc/h})^\gamma \tag{2.15}$$

where $J_2(\gamma) = 72/[(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma]$. Power-law fit bias measurements commonly use smaller scales ($< 1-2$ h^{-1}Mpc) that are in non-linear regime in order to increase the statistical significance. On the other hand if power-law fits are restricted only to larger scales, the method suffers from the problem that the lowest scale, where the linear biasing scheme can still be applied, varies from sample to sample and remains ambiguous.

As I discuss in Chapter 3, the AGN bias can also be estimated from the data directly by comparing the observed clustering of AGN with the clustering of dark matter measured in a cosmological simulation. Moreover, the halo occupation model is currently the optimal method to establish the AGN bias, allowing the use of the full range of scales.

2.6 Clustering of X-ray AGN: Overview

As I already mentioned, AGN are powerful X-ray emitters. The discovery of the intense cosmic X-ray background (e.g., Giacconi et al., 1972) in the early 1960s opened up a privileged window for the study of the energetic phenomena associated with accretion onto black holes. Due to the relative weakness of X-ray emission from stars and stellar remnants (magnetically active stars, cataclysmic variables and X-ray binaries), the X-ray sky is almost completely dominated by the evolving SMBH population, at least down to the faintest fluxes probed by current X-ray

focusing telescopes. The goal of reaching a complete census of evolving AGN, and thus of the accretion power released by SMBH in the history of the universe has therefore been intertwined with that of fully resolving the CXB into individual sources.

Since its discovery, the nature of the CXB has been strongly debated, but soon the community converged into interpreting most of the CXB as the integrated emission of AGN across the cosmic time. While the discrete nature of the CXB has been proposed (Bergamini Londrillo & Setti, 1967) and rapidly unveiled by experiments like *Einstein* (Giacconi et al., 1979) and ROSAT (see e.g. Hasinger et al., 1993), little cosmological information has been obtained from samples of AGN because of the scarce number of detected sources in the X-ray band. AGN as phase of the galactic evolution, is a quite rare phenomenon in the Universe as the space density of these objects is about 1/100 - 1/1000 lower than that of galaxies. This means that AGN surveys require large field of view and/or deep exposure to provide statistically significant samples. Clustering measurements of X-ray point sources have been made whenever new large-scale surveys have been completed and recent *Chandra* and XMM-*Newton* surveys allowed a study of the clustering of AGN with a precision comparable to that achievable with redshift galaxy surveys.

Barcons & Fabian (1988) measured with *Einstein* a clustering signal of the CXB on scales $\leq 5'$ corresponding to an angular correlation length $\theta_0 \sim 4'$. They have shown the importance of studying the angular structure of the CXB by pointing out that a large fraction of the CXB could have been attributed to sources with a redshift distribution similar to optical QSOs. In addition, the first prediction was not consistent with the hypothesis that the CXB was also partly produced by a diffuse hot Intergalactic Medium (IGM) component. It was also proposed that these sources were actually clustered on comoving scales of the order of $\sim$10 h^{-1}Mpc. Carrera & Barcons (1992), Georgantopoulos et al. (1993) and Soltan & Hasinger (1994) observed that the CXB was highly isotropic on scales of the order of 2°-25°.

The first attempt of measuring the clustering of X-ray selected AGN was performed by Boyle & Mo (1993), that measured a barely significant signal by using a sample of 183 EMSS sources, mostly local AGN (z<0.2). These evidences have brought the attention to the study of the clustering of the CXB down to the arcminute scale. The first significant upward turn for the measurement of AGN clustering in the X-ray band has been brought to light by ROSAT. By using a set of ROSAT-PSPC pointing on an area of $\sim$40 deg^2, Vikhlinin & Forman (1995) measured, for the first time, an angular correlation signal of faint (ROSAT) X-ray sources on scales $<10'$. By using the Limber equation (Peebles, 1980) they have de-projected their angular correlation function into a real space correlation function and found that, under the assumption that the redshift distribution of the sources was the same as that of optical QSOs, the spatial correlation length was

in the range 6-10 h^{-1}Mpc. With such a result, they confirmed the hypothesis that the CXB was mostly produced by sources with a redshift distribution comparable to that of optically selected QSO, though with almost double source density. By using the results of Vikhlinin & Forman (1995) and Akylas et al. (2000, who obtained similar results), Barcons et al. (2001) has shown for the first time that X-ray selected AGN are highly biased tracers of the underlying LSS at z<1 by showing a redshift evolving bias factor as large as b$\sim$2.

However, it is worth to consider that the deprojection of the angular correlation function into a 3D correlation relies on several assumptions, like the model dependent expected redshift distribution, which may lead to a biased estimate of the real space clustering. It is however worth noticing that angular correlation can be very useful to provide a first overview in the early phase of surveys, when optical identifications are not available, especially sampling new part of the parameter space of sources, like i.e. new unexplored luminosity/flux limits and therefore source classes.

The first firm detection of 3D spatial clustering of X-ray selected AGN has been claimed by Mullis et al. (2004) by using data of the ROSAT-NEP survey. They detected on an area of $\sim$81 deg^2 a 3σ significant signal in the redshift space auto-correlation function of soft X-ray selected sources at $\langle z \rangle \sim 0.22$. They have shown that, at that redshift, AGN cluster with a typical correlation length $r_0 = 7.4 \pm 1.9$ h^{-1}Mpc. Their results suggest that the population of AGN in such a sample is consistent with an unbiased population with respect to the underlying matter. Their result suggested that at that redshift AGN were hosted in DMHs of mass of the order of 10^{13} $h^{-1} M_{\odot}$.

With the development of *Chandra* and XMM-*Newton* surveys providing both the depth and large sky coverage, current survey have achieved both high source surface densities (i.e. $>$ 400-1000 deg^{-2}) and our capabilities in tracing the LSS has dramatically increased. One of the first evidences that AGN are highly correlated with the underlying LSS has been pointed out by Cappi et al. (2001) and Cappelluti et al. (2005) and references therein, who showed that around massive high-z galaxy clusters the source surface density of *Chandra* point sources is significantly, up to 2 times, higher than that of the background. More recently, Koulouridis & Plionis (2010) showed that although the X-ray source surface density of AGN around galaxy clusters is larger than in the background, the amplitude of their overdensities is about 4 times lower than that of galaxies in the same fields. This has been interpreted as a clear indication of an environmental influence on the AGN activity. Silverman et al. (2011) in the COSMOS field and Koss et al. (2010) in the *Swift*-BAT all-sky survey have shown that the AGN fraction in galaxy pairs is higher relative to isolated galaxies of similar stellar mass providing an additional evidence of the influence of the environment on AGN activity. *Chandra* and XMM-*Newton* performed several blanck sky extragalactic surveys, and most of them dedicated

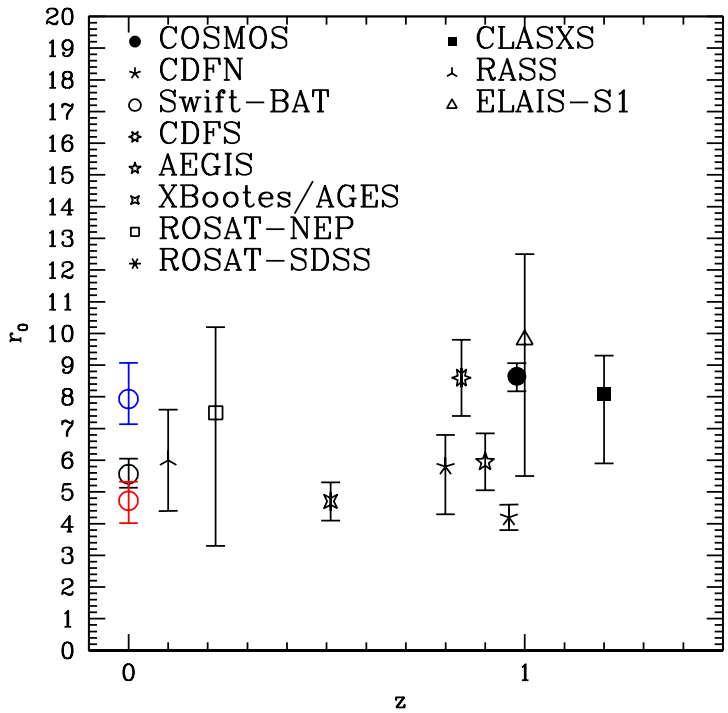

Figure 2.3 : Redshift evolution of the correlation length r_0 as estimated in different X-ray surveys (COSMOS, Gilli et al. (2009); CDFN, Gilli et al. (2005); Yang et al. (2006); Swift-BAT, Cappelluti et al. (2010); CDFS, Gilli et al. (2005); AEGIS, Coil et al. (2009); AGES, Hickox et al. (2009); ROSAT-NEP, Mullis et al. (2004); ROSAT-SDSS, Krumpe et al. (2010); CLASXS, Yang et al. (2006); RASS, Akylas et al. (2000); ELAIS-S1, Puccetti et al. (2006)). [Credits: Cappelluti, Allevato & Finoguenov (2012)]

part of their efforts in the study of the LSS traced by AGN to unveil their co-evolution. Basilakos et al. (2004, 2005) by using data of the XMM-*Newton* 2dF-survey have measured an unexpected high correlation length both in the angular ($\theta_0 \sim 10''$) and, by projection, in the real space ($r_0 \sim$16 h^{-1}Mpc). With the same technique, Gandhi et al. (2006) obtained a marginal 2 - 3σ detection of angular clustering in the XMM-LSS survey and obtained $\theta_0 = 6.3(42)''\pm3(^{+7}_{-13})$ in the 0.5-2 (2-10) keV bands and a slope $\gamma \sim 2.2$. Puccetti et al. (2006) measured the clustering of X-ray sources in the XMM-*Newton* ELAIS-S1 survey in the soft and hard energy bands with a sample of 448 sources. They obtained $\theta_0 = 5.2\pm3.8''$ and $\theta_0 = 12.8\pm7.8''$ in the two bands respectively. These measurements have been deprojected with the Limber's inversion in the real space and obtained $r_0 =$ 9.8-12.8 h^{-1}Mpc and $r_0 =$ 13.4-17.9 h^{-1}Mpc in the two bands, respectively. In the *Chandra* era, Gilli et al. (2005) measured the real space auto-correlation function of point sources in the CDFS-CDFN. They have measured in the CDFS $r_0 = 8.6\pm1.2$ h^{-1}Mpc at z=0.73, while in the CDFN they obtained $r_0 = 4.2\pm0.4$ h^{-1}Mpc. The discrepancy of these measurements has been explained with variance introduced by the relatively small field of view and the consequent random sampling of LSSs in the field. In the CLASXS survey Yang et al. (2006) obtained a measurement of the clustering at z = 0.94 with $r_0 = 8.1^{+1.2}_{-2.2}$ h^{-1}Mpc which proposes that AGN are hosted by DMH of mass of $10^{12.1}$ $h^{-1}M_{\odot}$. In addition they proposed that AGN clustering evolves with luminosity and they found that the bias factor evolves with the redshift. Such a behavior is similar to that found in optically selected quasars. The XMM-*Newton* (Hasinger et al., 2007; Cappelluti et al., 2007, 2009) and *Chandra* (Elvis et al., 2007; Puccetti et al., 2009) survey of the COSMOS field have provided a leap forward to the field of X-ray AGN clustering by surveying a 2 deg^2 field of view. The key to the success of this project is a redshift survey *zCOSMOS* (Lilly et al., 2007) performed simultaneously with the X-ray survey, together with observations in more than 30 energy bands from radio to X-ray, that allowed to measure either the spectroscopic or the photometric redshift (Salvato et al., 2009, 2011) of every source. In the X-ray band, the survey covers 2 deg^2 with XMM-*Newton* with a depth of $\sim$60 ks with the addition of a central 0.9 deg^2 observed by *Chandra* with $\sim$150 ks exposure.

The first sample of $\sim$1500 X-ray sources (Cappelluti et al., 2007) in the COSMOS field, has been used by Miyaji et al. (2007) to determine their angular correlation function, without knowing their distance, and just assuming a theoretical redshift distribution for the purpose of Limber's deprojection. Significant positive signals have been detected in the 0.5-2 keV band, in the angular range of 0.5′- 24′, while the positive signals were at the $\sim 2\sigma$ and 3σ levels in the 2-4.5 and 4.5-10 keV bands, respectively. With power-law fits to the ACFs without the integral constraint term, they have found correlation lengths of $\theta_0 = 1.9\pm0.3''$, $0.8^{+0.5}_{-0.4}''$, and $6\pm2''$ for the three bands, respectively, for a fixed slope γ=1.8. The inferred

Table 2.1 : Bias Factors and DMH masses as estimated in previous works

Survey	Band keV	N_{obj}	z	θ_0 arcsec	r_0 h^{-1}Mpc	γ	b(z)[a]	Log(M_h)[b] $h^{-1}M_\odot$
EMSS	0.5-2	183	<0.2	X	<10	X	X	X
RASS	0.1-2.4	2158	1-1.5	∼10	<10	1.7±0.3	X	X
RASS	0.1-2.4	2096	0.1	∼3.7	6.0±1.6	1.9±0.31	X	X
ROSAT-NEP	0.1-2.4	220	0.22	X	$7.5^{+2.7}_{-4.2}$	$1.85^{+1.90}_{-0.80}$	$1.83^{+1.88}_{-0.61}$	$13.51^{+0.91}_{-0.79}$
AXIS[1]	0.5-2	31288	0.96	22.9±2.0	6.54±0.12	1.12±0.04	2.48±0.07	$13.20^{+0.11}_{-0.12}$
AXIS[1]	2-10	9188	0.94	$29.2^{+5.1}_{-5.7}$	9.9±2.4	$2.33^{+0.10}_{?0.11}$	2.38±0.51	$13.14^{+0.28}_{-0.41}$
AXIS[1]	5-10	1259	0.77	$40.9^{+19.6}_{-29.3}$	5.1±4.1	$1.47^{+0.43}_{?0.57}$	2.14±1.88	$13.17^{+0.84}_{-2.44}$
ELAIS-S1	0.5-2	392	0.4	5.2±3.8	$9.8^{+2.7}_{-4.3}$	1.8	X	X
ELAIS-S1	2-10	205	0.4	12.8±7.8	$13.4\ ^{+2.7}_{-4.3}$	1.8	X	X
CDFS	0.5-2	97	0.84	X	8.6±1.2	1.33±0.11	$2.64^{+0.29}_{-0.30}$	$13.41^{+0.55}_{-0.18}$
CDFN[2]	0.5-2	164	0.96	X	4.2±0.4	1.42±0.07	$1.87^{+0.14}_{-0.16}$	$12.73^{+0.12}_{-0.17}$
XMM-2dF[3]	0.5-2	432	1.2	10.8±1.9	∼16	1.8	1.9-2.7	12.5-13.1
XMM-LSS	0.5-2	1130	0.7	6.3±3	6±3	2.2±0.2	X	X
XMM-LSS	2-10	413	0.7	42^{+7}_{-13}	6±3	$3.1^{+1.1}_{-0.5}$	X	X
CLASXS	0.5-8	233	1.2	X	$8.1^{+1.2}_{-2.2}$	2.1±0.5	$3.58^{+2.49}_{-1.38}$	$12.86^{+0.61}_{-0.16}$
CDFN [4]	0.5-8	252	0.8	X	$5.8^{+1.0}_{-1.5}$	$1.38^{+0.12}_{-0.14}$	$1.77^{+0.80}_{-0.15}$	$13.53^{+0.63}_{-0.71}$
XMM-COSMOS[5]	0.5-2	1037	1.1	2.9±0.6	11.8±1.1,	1.8	3.7±0.3	13.6±0.1
XMM-COSMOS[5]	2-4.5	545	0.9	$1.2^{+1.1}_{-0.9}$	$6.9^{+2.2}_{-3.1}$,	1.8	$2.5^{+0.7}_{-1.0}$	$13.3^{+0.3}_{-0.7}$
XMM-COSMOS[5]	4.5-10	151	0.6	$6.5^{+3.0}_{-2.7}$	$12.7^{+2.3}_{-2.7}$	1.8	$3.8^{+0.6}_{-0.8}$	13.9±0.2
XMM-COSMOS[6]	0.5-2	538	0.98	X	$8.65^{+0.41}_{-0.48}$	$1.88^{+0.06}_{-0.07}$	3.08±0.14	$13.51^{+0.05}_{-0.07}$
SWIFT-BAT	15-55	199	0.045	X	$5.56^{+0.49}_{-0.43}$	$1.64^{+0.07}_{-0.08}$	$1.21^{+0.06}_{-0.07}$	$13.15^{+0.09}_{-0.13}$
AEGIS	0.5-2	113	0.9	X	5.95±0.90	1.66±0.22	$1.97^{+0.26}_{-0.25}$	$13.0^{+0.1}_{-0.4}$
AGES	0.5-2	362	0.51	X	4.5±0.6	1.6±0.1	$1.35\ ^{+0.06}_{-0.07}$	$12.60^{+0.1}_{-0.1}$
ROSAT+SDSS	0.1-2.4	1552	0.27	X	$4.28^{+0.44}_{-0.54}$	$1.67^{+0.13}_{-0.12}$	$1.11^{+0.10}_{-0.12}$	$12.58^{+0.20}_{-0.33}$
XMM-LSS	0.5-2	4360	1.1	3.2±0.5	7.2±0.8	1.93±0.03	2.7±0.3	13.2±0.3
XMM-LSS	2-10	1712	1.0	9.9±0.4	10.1±0.9	1.98±0.04	3.3±0.3	13.7±0.3

X:Unconstrained or undetermined

[a]: Bias factors converted to a common cosmology ($\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\sigma_8 = 0.8$)

[b]: DMH masses estimated using van den Bosch (2002) and Sheth et al. (2001)

[1]: Ebrero et al. (2009) , fit ID=2, assuming no redshift evolution of the correlation length

[2]: Gilli et al. (2005)

[3]: Basilakos et al. (2005) , using the LDDE model

[4]: Yang et al. (2006)

[5]: Miyaji et al. (2007), fit ID=6 with integral constrain, assuming redshift evolution of the correlation length

[6]: Gilli et al. (2009)

comoving correlation lengths were $r_0 = 9.8 \pm 0.7$, $5.8^{+1.4}_{-1.7}$, and 12 ± 2 h^{-1}Mpc at the effective redshifts of z = 1.1, 0.9, and 0.6, respectively. Comparing the inferred rms fluctuations of the spatial distribution of AGN $\sigma_{8,AGN}$, with those of the underlying dark matter, the bias parameters of the X-ray source clustering at these effective redshifts were found in the range b = 1.5 - 4. Such a result lead to the conclusion that the typical mass of the DMH hosting an AGN is of the order $M_h \sim 10^{13}$ $h^{-1} M_\odot$. Similar results have been found by Ebrero et al. (2009) using the angular correlation function of 30000 X-ray sources in the AXIS survey.

In the XMM-LSS survey Elyiv et al. (2012) measured the clustering of ∼5000 AGN and computed via Limber's deprojection the obtained $r_0 = 7.2 \pm 0.8$ h^{-1}Mpc and $r_0 = 10.1 \pm 0.8$ h^{-1}Mpcand $\gamma \sim 2$ in the 0.5-2 keV and 2-10 keV energy bands, respectively. In the XMM-COSMOS field Gilli et al. (2009) measured the clustering of 562 X-ray selected and spectroscopically confirmed AGN. They have obtained that the correlation length of these source, $r_0 = 8.6 \pm 0.5$ h^{-1}Mpc and slope of $\gamma = 1.88 \pm 0.07$. They also found that if source in redshift spikes are removed the correlation length decreases to about 5-6 h^{-1}Mpc. Even if not conclusively, they also showed that narrow line AGN and broad line AGN cluster in the same way, indicating that both class of sources share the same environment, an argument in favor of the unified AGN model which predicts that obscuration, and therefore the Type-I/Type II dichotomy is simply a geometrical problem. However it is worth noticing that such a procedure may artificially reduce the clustering signal and the effects of such a cut in the sample, may lead to an unreliable estimate of the clustering signal.

On the same line, Coil et al. (2009) measured the clustering of non-quasar X-ray active galactic nuclei at z = 0.7-1.4 in the AEGIS field. Using the cross-correlation of *Chandra*-selected AGN with 5000 DEEP2 galaxies they measured a correlation length of $r_0 = 5.95 \pm 0.90$ h^{-1}Mpc and slope $\gamma = 1.66 \pm 0.22$. They also concluded that X-ray AGN have a similar clustering amplitude as red, quiescent and "green" transition galaxies at z ∼ 1 and are significantly more clustered than blue, star-forming galaxies. In addition they proposed a "sequence" of X-ray AGN clustering, where its strength is primarily determined by the host galaxy color; AGN in red host galaxies are significantly more clustered than AGN in blue host galaxies, with a relative bias that is similar to that of red to blue DEEP2 galaxies. They did not observe any dependence of clustering on optical brightness, X-ray luminosity, or hardness ratio. In addition they obtained evidence that galaxies hosting X-ray AGN are more likely to reside in groups and more massive DMHs than galaxies of the same color and luminosity without an X-ray AGN. Coil et al. (2009); Mountrichas & Georgakakis (2012) concluded that DEEP2 X-ray AGN at z ∼ 1 are more clustered than optically selected quasars (with a 2.6σ significance) and therefore may reside in more massive DMHs. In an evolutionary picture their results are consistent with galaxies undergoing a quasar phase while in the blue

cloud before settling on the red sequence with a lower-luminosity X-ray AGN, if they are similar objects at different evolutionary stages (Hickox et al., 2009). At lower redshift, Krumpe et al. (2010) confirmed the results of Coil et al. (2009).

Various recent works have been presented indications and/or evidences, of varying significance, regarding a correlation between the X-ray Luminosity and the AGN clustering amplitude, based either on the spatial (Yang et al., 2006; Gilli et al., 2009; Coil et al., 2009; Cappelluti et al., 2010; Krumpe et al., 2010, 2012), or the angular (Plionis et al., 2008) correlation function. Note that luminosity dependent clustering is one of the key features of merger triggered AGN activity and is one of the prime motivations for AGN clustering analyses. Low L_X AGN have been found to cluster in a similar way as blue star forming galaxies while high L_X AGN cluster like red passive galaxies. Such a result has been confirmed by Cappelluti et al. (2010) using the Swift-BAT all sky survey at z $\sim$ 0. They detected both a L_X dependence of AGN clustering amplitude and a larger clustering of Type I AGN than that of Type II AGN. Krumpe et al. (2010, 2012) confirm the weak dependence of the clustering strength on AGN X-ray luminosity at a 2σ level for $z < 0.5$.

Table 2.1 summarizes all the discussed results on the clustering of AGN in X-ray surveys with bias factors converted to a common cosmology ($\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\sigma_8 = 0.8$) in the EMSS, Boyle & Mo (1993); RASS, Vikhlinin & Forman (1995); Akylas et al. (2000); ROSAT-NEP, Mullis et al. (2004); AXIS, Ebrero et al. (2009); ELAIS-S1, Puccetti et al. (2006); CDFS, Gilli et al. (2005); CDFN, Gilli et al. (2005); Yang et al. (2006); XMM-2dF, Basilakos et al. (2005); XMM-LSS, Gandhi et al. (2006); CLASXS, Yang et al. (2006); COSMOS, Gilli et al. (2009); Swift-BAT, Cappelluti et al. (2010); AEGIS, Coil et al. (2009); AGES, Hickox et al. (2009); ROSAT-SDSS, Krumpe et al. (2010). Fig. 2.3 shows the redshift evolution of the correlation length r_0 as estimated in previous works, according to the legend.

Chapter 3

Halo Model Approach

3.1 The standard cosmological paradigm

The matter content of the Universe is constituted only for a small fraction by the visible objects that we can observe, such as gas, stars, galaxies, groups and clusters. The mass dominant component is indeed in the form of non-baryonic, dark matter (DM), which interacts only gravitationally. Although the constituent particles of the DM are still unknown, the scenario that best explains the observed Universe seems to require this matter component to be cold (CDM, cold dark matter), i.e. its particles were nonrelativistic at the moment of decoupling between matter and radiation, during the early stages of the evolution of the Universe. Nonetheless, the baryonic and dark matter cannot be the only cosmic components, since they would imply the Universe to stop expanding and eventually recollapse or the expansion rate to become remarkably low. On the contrary, observational studies of distant supernovae in the late 1990s (Riess et al., 1998; Permutter et al., 1999) showed that the Universe is expanding with a rate which is higher than in the past. This requires the introduction of an additional contribution to the cosmic energy content, which must be indeed dominant and have a repulsive effect.

The characteristics and nature of this component are still unknown and therefore this is usually referred to as dark energy. Despite the fact that several theoretical models have been developed in order to interpret the dark energy (e.g., Weinber & Kamionkowski, 2003; Bartelmann et al., 2006), it is commonly represented with the Cosmological Constant term, Λ, originally introduced by Einstein into the field equations describing the evolution of the Universe. The so-called ΛCDM paradigm (Bryan & Norman, 2010) represents the standard cosmological model describing the Universe. Recent observational estimates obtained from the seven years WMAP (Komatsu et al., 2011) data provides $\Omega_{0,\Lambda} = 0.728$, $\Omega_{0,m} = 0.272$ and $\Omega_{0,b} = 0.045$, for the density parameter of dark energy, dark matter and baryonic matter, respectively, and $H_0 = 70.4$ km s^{-1}Mpc^{-1} for the Hubble

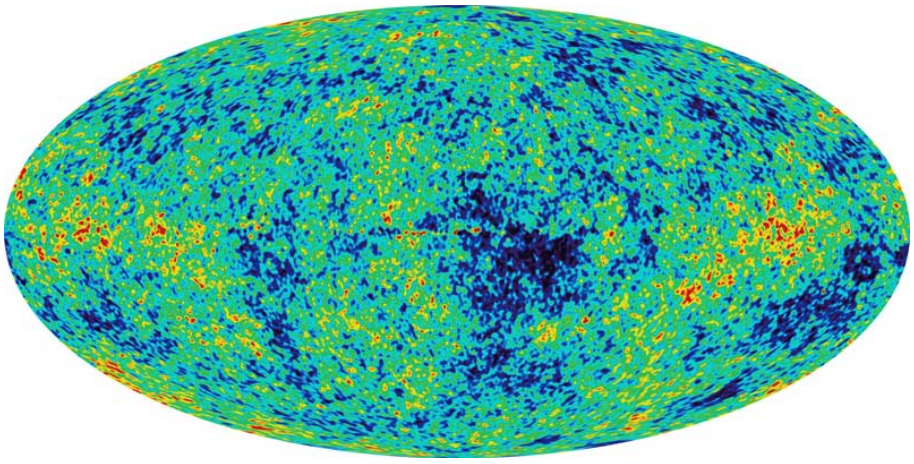

Figure 3.1 : Image of the CMB anisotropy map, obtained from the WMAP satellite. Temperature fluctuations with respect to the mean are color-coded from blue (colder) to red (warmer) and represent the perturbations that have eventually grown into the present-epoch cosmic structures. [Credits: NASA, WMAP Science Team]

constant.

The cosmic microwave background (CMB) radiation, serendipitously discovered in 1965 (for which Penzias and Wilson were later awarded with the Nobel Prize), is the thermal radiation permeating almost uniformly the whole Universe, theoretically expected to be left as a relic of a primordial, very hot phase of the cosmic evolution. CMB radiation has been proven to have an almost perfect black-body spectrum at a temperature of 2.725 K, and a low degree of anisotropy, confirming indeed the large-scale isotropy of the Universe. Nonetheless, sophisticated measurements obtained with the COBE and WMAP satellites (see Fig. 3.1) provide evidences for primordial inhomogeneities in the density/energy field, visible, for instance, as fluctuations of order 10^{-5} in the CMB temperature map. The challenge for models of structure formation consists of explaining in detail how these initial seeds have grown into the present-epoch structures, such as galaxies and clusters. The standard scenario assumes the perturbations observed in the CMB to originate from the amplification of quantum fluctuations during a phase of accelerated expansion of the early Universe (referred to as inflation).

Even though several complex theories about the primordial distribution have been proposed and debated during the last decades (e.g. non-Gaussian initial conditions), the probability distribution function for the primordial fluctuations of the cosmic density field is usually assumed to be Gaussian. The growth of these initial density perturbations can be studied analytically through the linear Jeans's

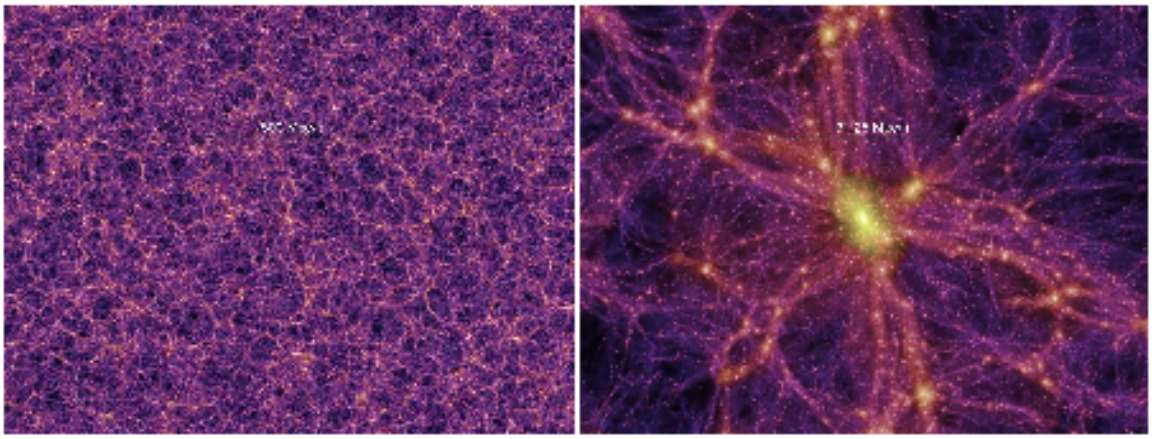

Figure 3.2 : Slices through the density field of the DM Millennium Simulation Run, at redshift z = 0. The righthandside panel shows a zoom in by a factor of 16 with respect to the lefthandside panel. Reference scales, in comoving units, are reported on the maps. It is evident from these snapshots of the simulation the filamentary structure which constitute the cosmic web, where small haloes are. [Credits: Springel et al. 2005]

theory of gravitational instability as long as the density contrast with respect to the underlying background is smaller than unity. The collapse into DM halos, however, implies the overdensities to enter in the nonlinear regime and eventually grow into highly nonlinear structures, for which the Jeans's theory provides no more a valid solution and a semianalytical or numerical approach is then required.

Starting from the seventies of the last century, numerical simulations have been a very powerful tool to confirm the reliability of the CDM scenario for structure formation (e.g., Davis et al., 1985; Frenk et al., 1985, 1988; White et al., 1987) and help constraining the standard cosmological paradigm, in remarkable agreement to the observed large-scale structure of the Universe (e.g. in the Millennium Simulation Run by Springel et al., 2005, see Fig. 3.2). Since pioneering theoretical studies in the 1970s (Press & Schechter, 1974; White & Rees, 1978) the CDM scenario emerged as the most suited to fit observations. This commonly accepted scenario for structure formation implies a "bottom-up" hierarchy of cosmic structures, according to which smaller halos of dark matter collapse first and consequently merge to form bigger structures. Clusters of galaxies form from the evolution and collapse of the rare high peaks in the primordial Gaussian density fluctuations, being therefore the largest structures in the Universe that have had enough time to reach virial equilibrium.

3.2 Perturbation Density Field

In studying how matter in the expanding Universe responds to its own self-gravity, a linear solution for how the matter behaves can be found by expanding the equations of motion in terms of a dimensionless density perturbation δ:

$$\delta \equiv \frac{\rho(\boldsymbol{x}) - \bar{\rho}}{\bar{\rho}} \tag{3.1}$$

where $\boldsymbol{x}$ is the comoving coordinate, $\rho(\boldsymbol{x})$ is the matter density field, which is a continuous function of the position vector $\boldsymbol{x}$, $\bar{\rho}$ ($\bar{\rho} \equiv \rho_0$) is its average density computed over a sufficiently large (representative) volume of the Universe. According to this definition, the spatial average at a given redshift requires $\langle \delta(\boldsymbol{x}, z) \rangle = 0$. The next momentum of the matter density probability distribution, the variance, is:

$$\langle \delta(\boldsymbol{x}_1, z)\, \delta(\boldsymbol{x}_2, z) \rangle \;=\; \xi(\boldsymbol{x}_1, \boldsymbol{x}_2, z) \;=\; \xi(|\boldsymbol{x}_1 - \boldsymbol{x}_2|, z) \;=\; \xi(r, z) \tag{3.2}$$

where we have defined the density auto-correlation function as the excess probability of finding masses at separation r. The function ξ depends only on the distance between the coordinates $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, because we have assumed the Universe to be statistically isotropic. If the initial density field is described with a Gaussian distribution, as derived from a class of standard inflation models, the mean and the variance alone completely describe the matter density distribution.

The density distribution of matter can be defined also in the Fourier space, where it can be described by the superposition of planes waves, which evolve independent one of each other during linear evolution. The Fourier Transform of δ is

$$\tilde{\delta}(\boldsymbol{k}, z) \;=\; \int d^3x\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \delta(\boldsymbol{x}, z)\,. \tag{3.3}$$

Then the variance of $\delta(\boldsymbol{x}, z)$ is given by:

$$\sigma^2(z) \;=\; \langle \delta^2\, \tilde{\delta}^*(\boldsymbol{k}', z) \rangle \;=\; \frac{1}{2\pi^2} \int_0^\infty P(k, z)\, k^2\, dk \tag{3.4}$$

where $P(k, z)$ is the *power spectrum* of density fluctuations as a function of redshift. Similar to $\xi(r, z)$, which is only a function of the modulus of the distance vector, $P(k, z)$ depends only on the absolute value of the wave-number, $k = |\boldsymbol{k}|$.

We want to study a class of observable structures of mass M which arise from the collapse of initial perturbations of size R, thus it is useful to introduce the smoothed density field:

$$\delta_R(\boldsymbol{x}, z) \;=\; \delta_M(\boldsymbol{x}, z) = \int \delta(\boldsymbol{y}, z)\, W_R(|\boldsymbol{x} - \boldsymbol{y}|)\, d\boldsymbol{y} \tag{3.5}$$

where W_R is the *window function* which weights the contribution from different spatial points to the density distribution. By convolving $\delta(\boldsymbol{k}, z)$ with W_R, we smooth out all the fluctuations at scale below R. Therefore, the variance of the fluctuation field at scale R is defined as:

$$\sigma_R^2(z) \;=\; \langle \delta_R^2 \rangle \;=\; \sigma_M^2 = \frac{1}{2\pi^2} \int dk\, k^2\, P(k,z)\, \tilde{W}_R^2(k) \tag{3.6}$$

where $\tilde{W}_R(k)$ is the Fourier transformation of W_R. The shape of the window function used in this Thesis is the top-hat one:

$$\tilde{W}_R(k) \;=\; \frac{3\,[\sin(kR) - k\,R\cos(kR)]}{(k\,R)^3} \tag{3.7}$$

for which the relation between mass scale and smoothing scale is $M = \frac{4\pi}{3} R^3 \bar{\rho}$. The normalization of the power spectrum is parametrized through the quantity σ_8, which is defined as the variance computed for a top-hat window having comoving radius $R = 8$ h^{-1}Mpc $\equiv R_8$. The choice of R_8 was motivated by results of early galaxy surveys (Davis & Peebles, 1983) finding $\delta_{gal}(R_8) \simeq \delta N_{gal}/N_{gal} \simeq \delta M/M \simeq 1$, i.e. the variance of galaxy number density in spatial bins of radius R_8 is about unity. A top-hat sphere with such radius contains a mass $M \sim 5 \times 10^{14}\ M_\odot$, which is the typical mass of a moderately rich galaxy cluster. Hence, σ_8 determines the height of density peaks and consequently the object abundance. Moreover, the mass variance of order of unity on the R_8 scale marks the transition region from the linear regime with $\delta \lesssim 1$ at $R > 8$ h^{-1}Mpc, to the non-linear regime at $R \ll 8$ h^{-1}Mpc.

The power spectrum is usually described by a power law $P(k) \propto k^{n_s}$. Thus, the second important parameter related to the power spectrum is the index n_s of the initial power spectrum. This index is observationally confirmed (Spergel et al., 2007) to be very close to the scale-free Harrison-Zel'dovich spectrum of primordial adiabatic fluctuation (i.e. $n_s = 1$), which is predicted by inflationary models. Mukhanov's inflation thoery requires the index to be deviant from 1, suggesting the value 0.96. Under the linear evolution assumption in the cosmic structure formation model, each mode $\tilde{\delta}(\boldsymbol{k}, t)$ evolves independently and hence the evolution density is a linear function of the initial conditions. If the initial conditions were Gaussian, then so are the evolved quantities, although with a different power spectrum. The growing-mode solution is expected to dominate after sufficient evolution, hence the power spectrum at a given redshift is described by:

$$P(k,z) = T^2(k)\, D^2(z)\, P_{in}(k) \tag{3.8}$$

where $D(z)$ is the linear growth mode of perturbations and $P_{in}(k)$ is the initial power spectrum described by $P(k) = A\, k^{n_s}$. The amplitude A of the power spectrum at $z = 0$ is directly linked to σ_8. The quantity $T(k)$ is the transfer function (Eisenstein & Hu, 1998).

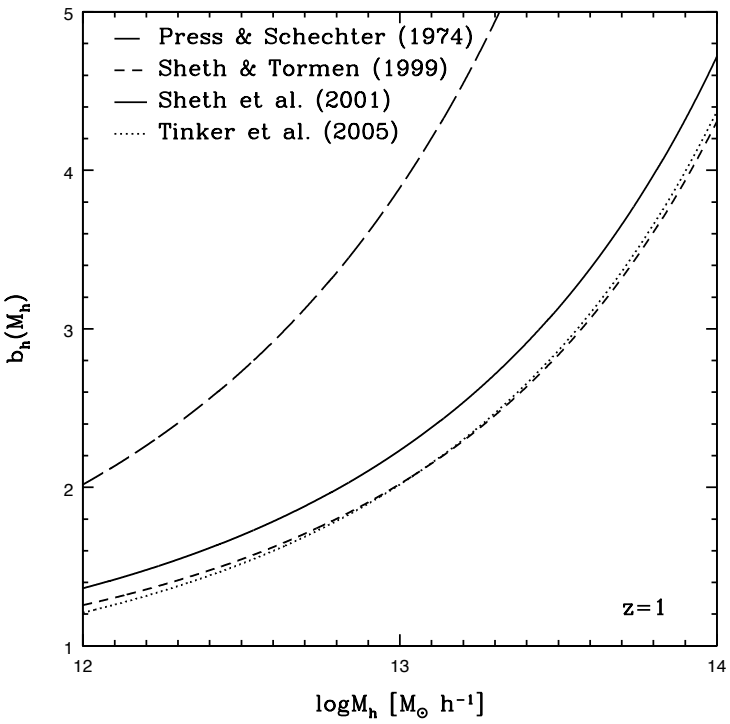

Figure 3.3 : Halo bias as a function of halo mass for redshift z=1 and the corresponding predictions of Press & Schechter (1974) (*long-dashed line*), Sheth & Tormen (1999, *dashed line*) , Sheth et al. (2001, *solid line*) and Tinker et al. (2005, *dotted line*). [Credits: Cappelluti, Allevato & Finoguenov (2012)]

3.3 Dark Matter Halo Properties

A dark matter halo is a roughly spherical, virialized object. An overdense region in the early universe grows until the density is high enough that self-gravity overcomes the local expansion of the universe. At this point, the overdensity ceases expanding and begins to collapse in on itself. Using this simple picture of spherical collapse, Gunn & Gott (1972) calculated that the structure would stop collapsing and virialize at an overdensity 178 times the background density. This picture is simplified in that it describes the formation of a halo as the monolithic collapse from a single overdensity. N-body simulations have revealed a much more chaotic picture of halo formation, in which structures form hierarchically; smaller halos collapse and merge to form larger halos, which in turn merge again. In this hierarchical structure paradigm, the basic result of Gunn & Gott holds true-halos are roughly spherical, mostly virialized, and to a good approximation have a mean interior density $\sim$ 200 times the background.

The accurate description of the dark matter clustering requires three major ingredients: the mass function, the mass-dependent biasing factor and the density profile of halos.

3.3.1 Halo Mass Function and Halo Bias

The mass function, defined as the number density of halos of mass M, is a consequence of the power spectrum of density fluctuations. High-amplitude fluctuations at large scales result in more massive halos than a power spectrum with less large-scale power. The mass function has been the subject of many analytic studies, beginning with the pioneering work of Press & Schechter (1974, PS).

In the spherical collapse model of the formation of DMHs, halos collapse when the linear overdensity associated with a peak in the density field crosses a critical barrier δ_c independent of halo mass. In this approach, it is assumed that all density perturbations continue to grow according to the linear growth rate $D(z)$, even when their amplitudes become non-linear. When perturbations are treated in this way, their variance on mass scale M as a function of redshift can be expressed by Eq. 3.6. According to the PS approach, under the assumption of initial Gaussian density field perturbations, the number density of virialized halos with mass in the range $[M, M + dM]$, found at redshift z in a comoving volume element is given by:

$$\frac{dn(M,z)}{dM} = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma_M(z)} \left| \frac{d\ln\sigma_M(z)}{d\ln M} \right| \exp\left(-\frac{\delta_c^2}{2\sigma_M(z)^2} \right) \tag{3.9}$$

Press & Schechter (1974) used this model to derive an expression for the halo bias. In fact halos themselves are biased tracers of the underlying matter distribution. Fluctuations in the number density of halos on large scales are observed to be more pronounced than the fluctuations of the underlying matter density. This means that, the fractional deviation of $dn(M)/dM$ from its mean value, within a given volume of the Universe, is observed to be larger than $\delta\rho/\rho$ in that same volume. The ratio, $b(M)$, between the perturbation in the number density of halos of mass M and the perturbation amplitude of the matter density is called the bias parameter of DMHs and can be define as

$$b^2(M,k) = \frac{P_h(k)}{P_{\rm lin}(k)}\,, \tag{3.10}$$

that is the ratio of the halo power spectrum to the linear matter power spectrum.

According to the predictions based on the spherical collapse model for the formation of DMHs from Gaussian perturbations, Press & Schechter derived a bias relation of the form:

$$b^{PS} = 1 + \frac{\nu^2 - 1}{\delta_c} \tag{3.11}$$

$\nu = \delta_c/\sigma(M)$, where δ_c is the critical density for collapse and σ is defined in Eq. 3.6. However, the PS mass function fails to reproduce the DMH mass function found in simulations (e.g., Sheth & Tormen, 1999; Jenkins et al., 2001). It is well known that it overpredicts (underpredicts) the halo abundance at the low (high) mass end. Motivated by the fact that halo collapses are generally triaxial, Sheth & Tormen (1999), proposed the ellipsoidal collapse model with a moving barrier where the collapse threshold also depends on mass. By imposing this mass dependence on a collapse barrier, the abundance of small halos is suppressed and fitting formulae for the halo mass function are obtained, which agree with N-body simulations much better than the spherical collapse model predictions. The scale-indipendent bias proposed by Sheth & Tormen (1999) is given by:

$$b^{ST} = 1 + \frac{a\nu^2 - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (a\nu^2)^p} \tag{3.12}$$

where $a = 0.707$ and $p = 0.3$ or the ellipsoidal collapse formula of Sheth et al. (2001):

$$b^{SMT} = 1 + \frac{1}{\sqrt{a}\delta_c}[\sqrt{a}(a\nu^2) + \sqrt{a}b(a\nu^2)^{1-c} - \frac{(a\nu^2)^c}{(a\nu^2)^c} + b(1-c)(1-c/2)] \tag{3.13}$$

where $a = 0.707$, $b = 0.5$, $c = 0.6$ or the recalibrated parameters $a = 0.707$, $b = 0.35$, $c = 0.8$ of Tinker et al. (2005). The ν parameter can be estimated following the Appendix of van den Bosch (2002). Fig. 3.3 shows the bias as function of the halo mass M_h at z=1, following the predictions of Press & Schechter (1974); Sheth & Tormen (1999); Sheth et al. (2001); Tinker et al. (2005).

3.3.2 Halo density profile

The models of spherical collapse suggest that the density profile around the center of a collapsed halo depends on the initial density distribution of the region which is collapsed. If halos are identified as peaks in the initial density field, then massive halos correspond to higher peaks in the initial fluctuation field. The density run around a high peak is shallower than the run around a smaller peak: high peaks are less centrally concentrated. Therefore one might reasonably expect massive virialized halos to be also less centrally concentrated than low mass halos. Such a trend is indeed found. Functions of the form:

$$\rho(r) = \frac{\rho_s}{(r/r_s)^\alpha(1 + r/r_s)^\beta} \tag{3.14}$$

(ρ_s and r_s denote the scaling density and radius) have been extensively studied as models of elliptical galaxies. $(\alpha, \beta) = (1, 2)$ in the expression gives the universal

halo density profile proposed by Navarro, Frenk & White (1995, 1996, 1997) (hereafter NFW). The profile is parameterized by r_s and ρ_s, which define a scale radius and the density at that radius, respectively. Although they appear to provide a two-parameter fit, in practice, one finds an object of given mass m and radius r_{vir} in the simulations, and then finds that r_s which provides the best fit to the density run. This is because the edge of the object is its virial radius r_{vir}, while the combination of r_s and the mass determines the characteristic density, ρ_s, following:

$$m = \int_0^{r_{vir}} dr 4\pi r^2 \rho(r) \tag{3.15}$$

For the NFW profile,

$$m = 4\pi\rho_s r_s^3 \left[ln(1+c) - \frac{c}{1+c} \right] \tag{3.16}$$

where $c = r_{vir}/r_s$ is known as the concentration parameter. Note that this expression explicitly assumes that the halo profile is truncated at r_{vir}, even though formally, the NFW profile extends to infinity. Because this profile falls as r^{-3} at large radii, the mass within them diverges logarithmically.

Although various density profiles have been advocated from numerical simulations (see, e.g., Moore et al. 1999), these variants have relatively little influence on galaxy clustering. The main conclusion of NFW that is of importance for this work, namely that the halo density profile is universal, has survived its challenges.

3.4 Biased Galaxy Formation

The theory of dark matter has revolutionized our understanding of structure formation, and in the last fifty years it has been refined to the point where its successes far outweigh the questions that it creates. Because dark matter is pressureless, the perturbations in the dark matter distribution grow via gravitational instability only and eventually create non-linear, collapsed structures. The creation of structure in dark matter potential wells also draws in the surrounding baryons, where they can cool, collapse to significantly higher densities, thereby forming stars and galaxies.

However, this picture of structure formation underscores the difficulty in comparing observations of the large-scale structure traced out by galaxies to predictions of the dark matter distribution. Dark matter is a smooth, continuous distribution, while galaxies are discrete objects. If a galaxy of a given type forms in a region with local density ρ, it does not follow that a region with density $\rho/2$ contains half a galaxy. The pathway from baryon blob to glowing galaxy necessarily places galaxies in special regions. The galaxy distribution is therefore *biased* relative to

the matter distribution. Bias is quantified by the ratio of clustering amplitudes of the objects in question (galaxies, AGN, clusters, etc) to the underlying matter distribution. In linear theory bias is a constant value with no dependence on scale, but linear theory only holds on large scales. At non-linear scales bias has a complex spatial dependence.

Early attempts to predict where the galaxies might lie in a given cosmological model focused on parameterizing the local number density of galaxies by the local dark matter density, a parameterization which breaks down at small scales. The innovation of the halo occupation distribution (HOD) is that galaxy bias is no longer calculated directly as a function of the dark matter distribution; the HOD makes the assumption that galaxies are only located in collapsed dark matter structures referred to as halos. Because each galaxy resides in a halo, quantifying galaxy clustering requires only that we specify the number of galaxies that reside in each halo. The question of bias is transferred to the halo population, an advantageous proposition because halos are the product of gravitational forces only, and not subject to complex gas physics that dominate the baryons. The properties of halos can be determined through the use of N-body simulations or analytic calculations. Once the bias of dark halos is known, galaxy bias comes directly from the HOD.

3.5 Halo Occupation Distribution

At this point, we have formulae for the abundance and spatial distribution of halos, as well as for the typical density run around a halo. This means that we can fill DMHs with objects based on a statistical HOD, allowing one to model the clustering of galaxies within halos (and thus at non-linear scales) while providing a self-consistent determination of the bias at linear scales. Similarly the problem of discussing the abundance and spatial distribution of AGN can be reduced to studying how they populate their host halos. In fact, the HOD formalism describes the bias relation between AGN and matter at the level of individual virialized DMHs, whose distribution and properties can be readily predicted by numerical simulations or analytic models given a cosmological model.

The key ingredient of this formalism is the probability distribution $P(N, M)$ that a halo of mass M contains N galaxies. In the most general case, $P(N, M)$ is entirely specified by all its moments which, in principle, could be observationally determined by studying galaxy clustering at any order. For the purpose of modeling the two-point CF, we only need the description of the first two moments, the lowest-order moment of $P(N, M)$

$$\langle N \rangle(M) = \sum_N P(N, M) \tag{3.17}$$

which is often called halo occupation number and the second moment $\langle N(N -$

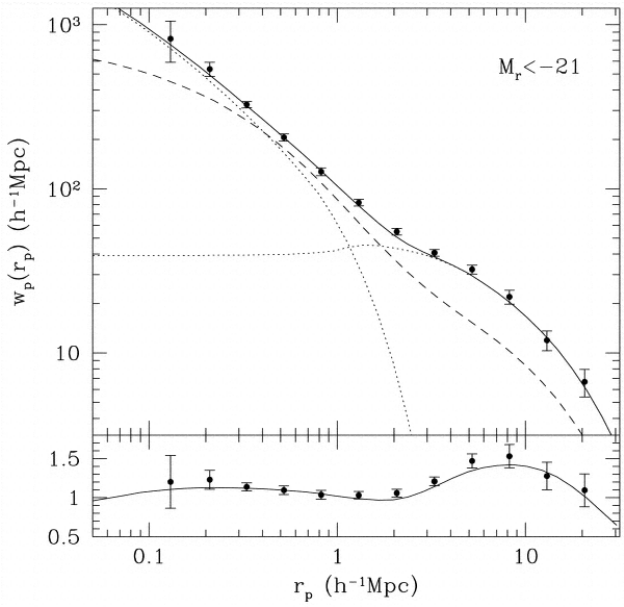

Figure 3.4 : Projected correlation function for a sample of SDSS galaxies with absolute magnitude $M_r < 21$, together with that for the best-fit HOD model as presented in Zehavi et al. (2004). The dotted curves show the one- and two-halo contributions to $w_p(r_p)$, and the dashed curve shows the projected correlation function for the matter computed from the linear power spectrum of Smith et al. (2003) .

1)$\rangle(M)$. This prescription assumes that the HOD depends on halo mass alone, i.e., the galaxy content of halos at a given mass is statistically independent of the large-scale environments within which those halos reside (e.g., Bond et al. 1991; Lemson & Kauffmann 1999). In particular, it is convenient to describe $\langle N\rangle(M)$ in terms of a few parameters whose values will then be constrained by the data. This term, along with the halo mass function $n(M)$ and the halo bias $b(M)$, provides some useful quantities; the mean comoving galaxy number density:

$$\bar{n} = \int n(M)\langle N\rangle(M)dM \tag{3.18}$$

and the large-scale bias:

$$b = \frac{\int n(M)\langle N\rangle(M)b(M)dM}{\int n(M)\langle N\rangle(M)dM} \tag{3.19}$$

In literature, the mean HOD $\langle N\rangle(M)$ is parametrized by separating the contribution from central and satellite galaxies

$$\langle N\rangle(M) = \langle N_{cen}\rangle(M) + \langle N_{sat}\rangle(M) \tag{3.20}$$

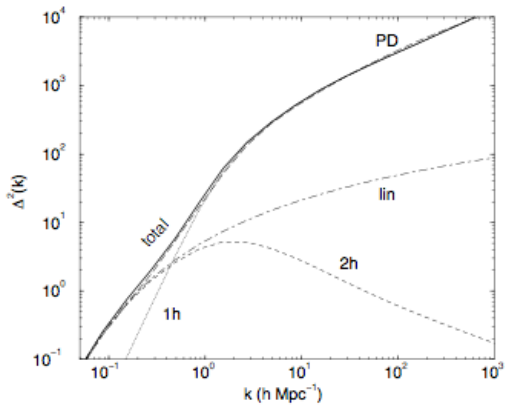

Figure 3.5 : Power spectrum of the dark matter density field at the present time expressed in terms of $\Delta^2(k) = k^3P(k)/2\pi^2$. Curve labeled PD shows the fitting formula of Peacock et al. (1996). Dot dashed curve labeled lin shows the linear $P_{lin}(k)$. Dotted and short dashed curves show the two terms which sum to give the total power (solid line) in the halo model.

where:

$$\langle N_{cen}\rangle(M_h) = erf\left(\frac{logM_h - logM_{min}}{\sigma_{logM}}\right) \tag{3.21}$$

$$\langle N_{sat}\rangle(M_h) = \left(\frac{M_h}{M_1}\right)^{\alpha_s} exp(-M_{cut}/M_h) \tag{3.22}$$

i.e. the central galaxy occupation follows a softened step function and the satellite occupation a rolling-off power law (e.g., Kravtsov et al., 2004; Zheng et al., 2005; Zehavi et al., 2005; Tinker et al., 2005; Chatterjee et al., 2012; Richardson et al., 2012).

The model admits five free parameters: M_{min}, the characteristic mass scale below which the mean HOD is zero; σ_{logM} the characteristic transition width of the softened step function; M_1, the approximate mass scale at which halos host, on average, one satellite galaxy; α, the power law index of the mean satellite occupation function; and M_{cut}, the mass scale below which the satellite mean occupation decays exponentially. This model implicitly assumes that the halo occupations of central and satellite galaxies are independent. That is, for a given halo, the occupation of satellite galaxies does not depend on whether there is a central quasar in the halo.

3.6 Interpreting the two-point CF with the halo model

The correlation function is defined as the probability above random of there being a pair of objects at separation r. In the HOD context, a pair of galaxies can reside within a single halo or come from two distinct halos. These two contributions are computed separately and combined to get the full correlation function, i.e.

$$\xi(r) = [1 + \xi_{\rm 1h}(r)] + \xi_{\rm 2h}(r). \tag{3.23}$$

The "1+" arises because it is the pair counts, proportional to $1 + \xi_{1h}$ and $1 + \xi_{2h}$, that sum to give the total pair counts, proportional to $1 + \xi$. The one-halo term is calculated in real space through (Berlind & Weinberg , 2002):

$$1 + \xi_{1h}(r) = \frac{1}{2\pi r^2 \bar{n}_g^2} \int dM \frac{dn}{dM} \frac{\langle N(N-1)\rangle_M}{2} \frac{1}{2R_{\rm vir}(M)} F'\left(\frac{r}{2R_{\rm vir}}\right), \tag{3.24}$$

where $\bar{n}_g$ is the mean number density of galaxies, dn/dM is the halo mass function and $\langle N(N-1)\rangle_M/2$ is the average number of pairs in a halo of mass M. The function $F(x)$ is the average fraction of galaxy pairs in a halo of mass M (or virial radius $R_{\rm vir}$) that have separation less than r, which is related to the halo density profile, $\rho_m(r)$, and $F'(x)$ is its derivative. In practice, $F(x)$ must be treated differently for central-satellite galaxy pairs and satellite-satellite pairs. In the former, the pair distribution is proportional to the volume-weighted density profile, $F'(x) \propto \rho_m(r)r^2$, normalized to one. For the latter it is derived from the halo profile convolved with itself, a calculation that can be done analytically for an NFW profile (Sheth et al., 2001). The average number of one-halo pairs in the range $(x, x + dx)$ in halos of mass M can be written explicitly as

$$\frac{\langle N(N-1)\rangle_M}{2} F'(x)\, dx = \langle N_{\rm cen} N_{\rm sat}\rangle_M F'_{\rm cs}(x)\, dx + \frac{\langle N_{\rm sat}(N_{\rm sat}-1)\rangle_M}{2} F'_{\rm ss}(x)\, dx, \tag{3.25}$$

where the subscripts cs and ss refer to central-satellite pairs and satellite-satellite pairs respectively. For a Poisson distribution of satellite occupation, $\langle N_{\rm sat}(N_{\rm sat} - 1)\rangle = \langle N_{\rm sat}\rangle^2$.

The one-halo term dominates $\xi_g(r)$ at small scales, while the two-halo term fully accounts for all galaxy pairs at separations $\gtrsim 1 - 2$ h^{-1}Mpc. It is this one-halo to two-halo transition that causes $w_p(r_p)$ to deviate from a power-law at scales near 1 h^{-1}Mpc (Zehavi et al. (2004), Fig. 3.4); as r increases, $\xi_{\rm 1h}$ drops rapidly, while the rise in $\xi_{\rm 2h}$ is regulated by halo exclusion. For brighter galaxies, which preferentially occupy high-mass halos, the rise in $\xi_{\rm 2h}$ occurs at a larger r, making the deviation from a power-law greater for brighter galaxy samples.

Since the radial distribution of galaxies within halos must be accounted for in the calculation of the two-halo term, the calculation itself is done in Fourier space, where the convolutions with the halo density profile become multiplications instead. The calculation of the two-halo term in Fourier space is:

$$P_{\rm gg}^{\rm 2h}(k,r) = P_{lin}(k)\left[\frac{1}{\bar{n}_g'}\int_0^{M_{\rm lim}} dM \frac{dn}{dM}\langle N\rangle_M b_h(M,r) y_g(k,M)\right]^2, \tag{3.26}$$

where $y_g(k,M)$ is the Fourier transform of the halo density profile, $b_h(M,r)$ is the halo bias at separation r, $\bar{n}_g$ is the average number density of galaxies and P_{lin} is the linear power spectrum.

At a given r, Eq. 3.26 is solved for all k, then converted to real space by

$$\xi_{2h}(r) = \frac{1}{2\pi^2}\int_0^\infty P_{\rm gg}^{\rm 2h}(k,r) k^2 \frac{\sin kr}{kr} dk, \tag{3.27}$$

Fig. 3.5 shows the power spectrum of the dark matter density field at the present time (z = 0). Dotted and short dashed lines show the contributions to the power from the one and two halo terms. Their sum (solid) is compared to the power spectrum measured in numerical simulations, represented here by the dashed curve labeled PD.

Chapter 4

Redshift Evolution of the AGN Bias in the COSMOS field

4.1 Introduction

As discussed in Chapter 2, the clustering analysis of AGN allow us to put tight constraints on how AGN are triggered and fueled, to identify the properties of the AGN host galaxies, and to understand how galaxies and AGN co-evolve. In addition, in the framework of the CDM structure formation scenario, clustering properties or the bias of AGN, may be related to the typical mass of DM halos in which they reside (Mo & White, 1996; Sheth & Tormen, 1999; Sheth et al., 2001; Tinker et al., 2005) and allow various types of AGN to be placed in a cosmological context.

Recently, several studies have been made, employing spectroscopic redshifts to measure the three dimensional correlation function of X-ray AGN. The majority of the X-ray surveys agree with a picture where X-ray AGN are typically hosted in DM halos with mass of the order of $12.5 < logM_h[\mathrm{h}^{-1}M_\odot] < 13.5$, at low ($z < 0.4$) and high ($z \sim 1$) redshift (Gilli et al., 2005; Yang et al., 2006; Gilli et al., 2009; Hickox et al., 2009; Coil et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010). This implies that X-ray AGN more likely reside in massive DM halos and preferentially inhabit dense environment typical of galaxy groups.

There have been attempts to detect X-ray luminosity dependence of the clustering. At $z \sim 1$, neither Gilli et al. (2009) nor Coil et al. (2009) found significant dependence of the clustering amplitudes on the optical luminosity, X-ray luminosity or hardness ratio, partially due to the larger statistical errors. Recent works by Krumpe et al. (2010), Cappelluti et al. (2010) found, however, that high X-ray luminosity AGN cluster more strongly than low X-ray luminosity ones at 2σ level for $z \sim 0.3$ and $z \sim 0$, respectively.

Until recently, the clustering of AGN has been studied mainly in optical, par-

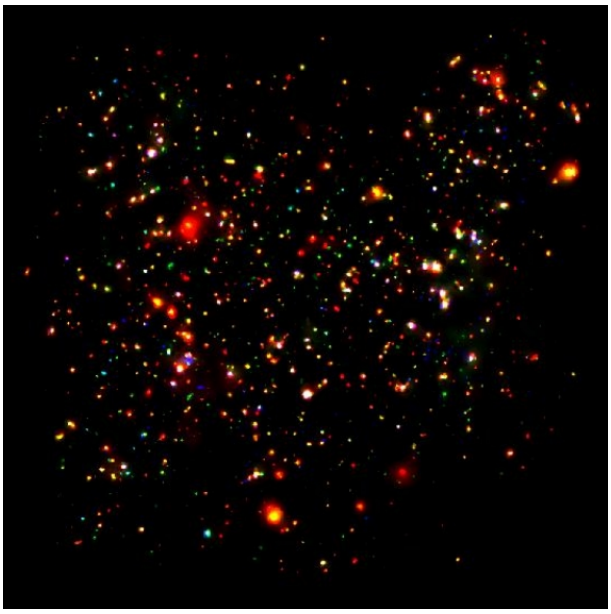

Figure 4.1 : False colour X-ray image of the XMM-COSMOS field: red, green and blue colours represent the 0.5-2 keV, 2-4.5 keV and 4.5-10 keV energy bands, respectively. [Credits: Cappelluti et al. (2009)]

ticularly in large area surveys such as 2dF (2QZ, Croom et al., 2005; Porciani & Norberg, 2006) and Sloan Digital Sky Survey (SDSS, Li et al., 2006; Shen et al., 2009; Ross et al., 2009). Croom et al. (2005) analysed the clustering of 2QZ QSO as a function of redshift finding a strong evolution of QSO bias, with $b_Q(z = 0.53) = 1.13 \pm 0.18$ at low redshift and $b_Q(z = 2.48) = 4.24 \pm 0.53$ at high redshift, as also observed in Porciani & Norberg (2006). The evidence of an evolution over time of the bias factor for SDSS quasars has been found in Shen et al. (2009), with bias values ranging from $b_Q(z = 0.50) = 1.32 \pm 0.17$ to $b_Q(z = 3.17) = 7.76 \pm 1.44$. The results from these surveys have also shown that the bias evolution of optically selected quasars is consistent with an approximately constant mass of the hosting DM halo in the range $logM_h[\mathrm{h}^{-1}M_\odot] \sim 12.5 - 13$ up to z $\sim$ 3.

Moreover, models of major mergers between gas-rich galaxies appear to naturally produce the bias of quasars as a function of L and z (Hopkins et al., 2008; Shen, 2009; Shankar et al., 2009, 2010; Shankar, 2010; Bonoli et al., 2009), supporting the observations that bright quasars are triggered by merging galaxies. It is still to be verified if the results from optical surveys can be extended to the whole AGN population and in particular to the X-ray selected AGN.

In this Chapter, I concentrate on the study of the redshift evolution of the bias

factor up to $z \sim 2$, using X-ray selected AGN in the XMM-COSMOS survey. The bias is estimated by using the power-law fitting of the ACF (which is the standard method, see §4.5) and with the two halo term (see §4.6). The XMM-COSMOS survey is particularly well suited to address AGN evolution in the context of the Large Scale Structure in which they reside and to investigate if AGN are biased tracers of the cosmic web. Throughout the Chapter, all distances are measured in comoving coordinates and are given in units of h^{-1}Mpc, where $h = H_0/100$ km/s. We use a ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b = 0.045$, $\sigma_8 = 0.8$. The symbol *log* signifies a base-10 logarithm.

This work is part of COSMOS (Cosmic evolution survey) project based on an ACS HST treasury program. The survey involves more than 100 scientists all over the world and makes use of most advanced multiwavelength observing facilities. X-ray observations have been first performed with the XMM-*Newton* and later integrated by *Chandra*. The contents of this Chapter have been already published in The Astrophysical Journal in August 2011 (Allevato et al., 2011).

4.2 AGN Catalog

In this Chapter I make use of the XMM-COSMOS multiwavelenght catalog (Brusa et al., 2010) which includes $\sim$ 1797 X-ray point-like sources detected in at least one of the soft (0.5-2 keV), hard (2-10 keV), or ultra-hard (5-10 keV) bands down to nominal limiting fluxes of $\sim 5 \times 10^{-16}$, $\sim 3 \times 10^{-15}$, and $\sim 7 \times 10^{-15}$ erg cm^{-2} s^{-1}, respectively. Fig. 4.1 shows an X-ray image of the COSMOS field: red, green and blue colours represent the 0.5-2 keV, 2-4.5 keV and 4.5-10 keV energy bands, respectively. In this catalog the X-ray positions are associated to the optical ones (I band) as well as the photometry at different wavelengths. The restriction to objects detected in the soft band (1465) guarantees the largest sample of XMM sources in the COSMOS field compared to the detection in the hard or ultra-hard band. This large sample has a spectroscopic completeness of $\sim$ 53% (780/1465). Starting from this sample of 780 objects with known spectroscopic redshift, I selected 593 AGN (I removed normal galaxies and ambiguous sources) with $I_{AB} < 23$ and $z < 4$. This magnitude cut increases the spectroscopic completeness to about 65%. The redshift distribution of the AGN sample (Fig. 4.2) shows prominent peaks at various redshifts, $z \sim 0.12$, $z \sim 0.36$, $z \sim 0.73$, $z \sim 0.95$, $z \sim 1.2$, $z \sim 2.1$. In particular, the structure at $z \sim 0.36$ was also observed at other wavelengths in COSMOS (Lilly et al., 2007) and already discussed (Gilli et al., 2009). The median redshift of the sample is $\langle z \rangle = 1.22$.

The sources have been classified in broad optical line AGN (BL AGN, 354), non-broad optical line AGN (NL AGN, 239) using a combination of X-ray and optical criteria (see Brusa et al., 2010, for more details), motivated by the fact that both obscured and unobscured AGN can be misclassified in spectroscopic studies,

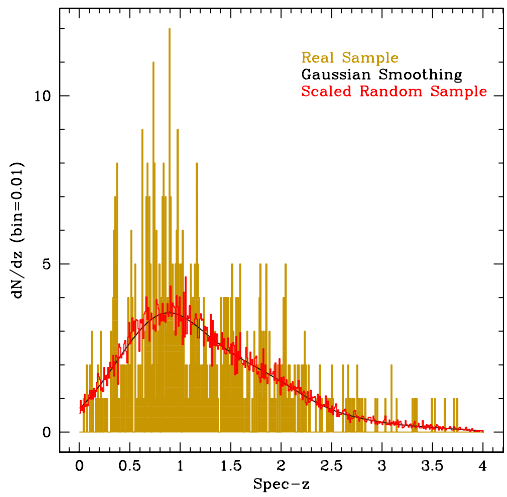

Figure 4.2 : Redshift distribution of 593 AGN (gold filled histogram) in bins of $\Delta z = 0.01$, with median $\langle z \rangle = 1.22$. The solid black curve is the Gaussian smoothing of the AGN redshift distribution with $\sigma_z = 0.3$, used to generate the random sample (red empty histogram).

given that the host galaxy light may over shine the nuclear emission. Fig. 4.3 (*Upper Panel*) shows the redshift distribution of BL AGN with $\langle z \rangle = 1.55$ and NL AGN with $\langle z \rangle = 0.74$.

In addition, I studied the clustering properties of X-ray unobscured and obscured AGN derived on the basis of the observed X-ray hardness ratio and corrected to take into account the redshifts effects. In particular I used the hard X-ray band (2-10 keV, which allows to sample the obscured AGN population) to select a subset of 184 X-ray unobscured sources (X-unobs hereafter) with $logN_H < 22\ cm^{-2}$ and 218 X-ray obscured (X-obs hereafter) sources with $logN_H \geq 22\ cm^{-2}$. The median redshifts of the two sub-samples are $\langle z \rangle = 1.12$ and $\langle z \rangle = 1.30$, respectively (see Fig. 4.3, *Lower Panel*). The 47% (40%) of BL (NL) AGN have been also observed in the hard band and classified as X-unobs (X-obs) AGN.

4.3 Random Catalog

The measurements of two-point correlation function requires the construction of a random catalog with the same selection criteria and observational effects as the data, to serve as an unclustered distribution to which to compare. XMM-*Newton* observations have varying sensitivity over the COSMOS field. In order to create an AGN random sample, which takes the inhomogeneity of the sensitivity over the field

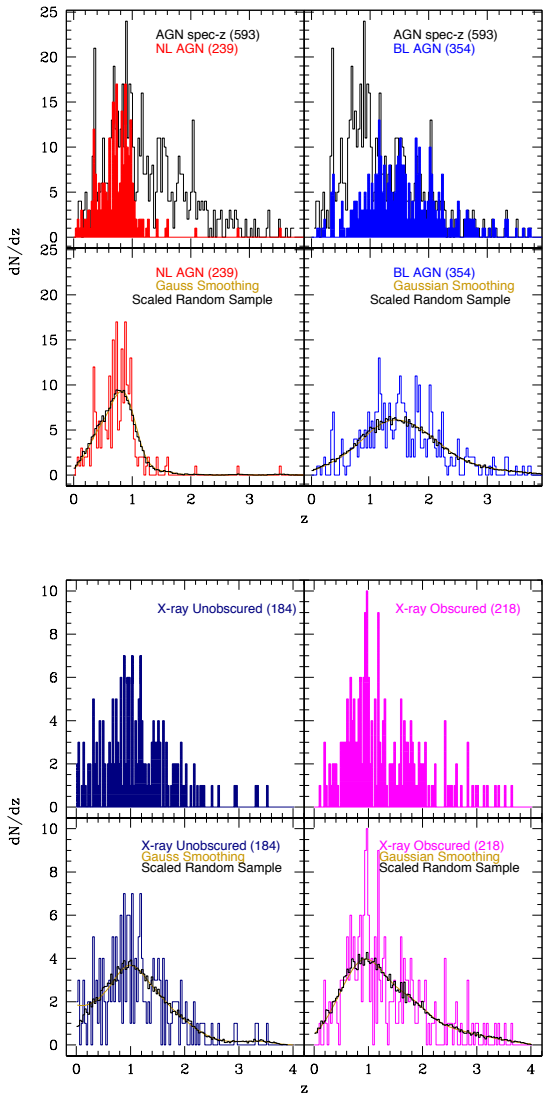

Figure 4.3 *Upper Panel*: Redshift distribution of XMM-COSMOS AGN (open histogram) selected in the soft band, compared with the redshift distribution of BL AGN (blue histogram, *upper right quadrant*) and NL AGN (red, *upper left quadrant*). Lower quadrants show the redshift distribution of the random catalogs (open black histograms) for both the AGN sub-samples, obtained using a Gaussian smoothing (gold lines) of the redshift distribution of the real samples. *Lower Panel*: Redshift distribution of unobscured (dark blue histogram) and obscured (magenta histogram) AGN selected in the hard band according with the column density (*upper quadrants*). Lower quadrants show the redshift distribution of the random catalogs (open black histograms) for both the AGN sub-samples, obtained using a Gaussian smoothing (gold lines) of the redshift distribution of the real samples.

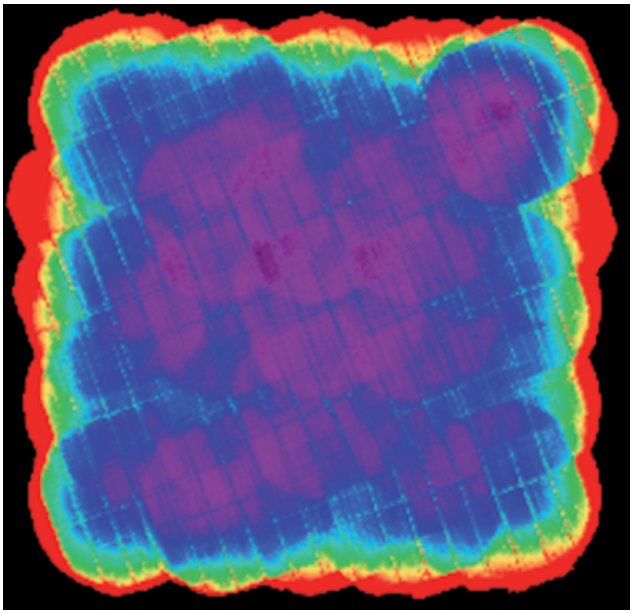

Figure 4.4 : The 0.5-2 keV sensitivity map of XMM-COSMOS in erg cm^{-2} s^{-1}. The map is plotted in colour coded scale from 5 $\times 10^{-16}$erg cm^2 s^1 (magenta) to 3 $\times$ 10^{-15} erg cm^2 s^1 (red). [Credits: Cappelluti et al. (2009)]

into account, each simulated source is placed at random position in the sky, with flux randomly extracted from the catalog of real sources fluxes (I verified that such flux selection produces the same results as if extracting the simulated sources from a reference input logN-logS). The simulated source is kept in the random sample if its flux is above the sensitivity map value at that position (see Fig. 4.4). Placing these sources at random position in the XMM-COSMOS field has the advantage of not removing the contribution to the signal due to angular clustering. On the other hand, this procedure does not take into account possible positional biases related to the optical follow-up program. Gilli et al. (2009), who instead decided to extract the coordinates of the random sources from the coordinate ensemble of the read sample, showed that there is a difference of only 15% in the correlation lengths measured with the two procedures.

The corresponding redshift for a random object is assigned based on the smoothed redshift distribution of the AGN sample. As in Gilli et al. (2009), I assumed a Gaussian smoothing length $\sigma_z = 0.3$. This is a good compromise between scales that are either too small, thus affected by local density variations or too large and thus oversmooth the distribution (our results do not change significantly using $\sigma_z = 0.2 - 0.4$). Fig. 4.2 shows the redshift distribution of 593 XMM-COSMOS AGN and the scaled random sample ($\sim$ 41000 random sources) which follows the

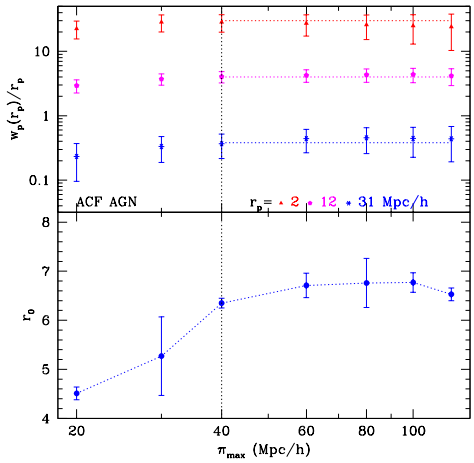

Figure 4.5 Projected AGN correlation function $w_p(r_p)$ computed at different r_p scale (see label) as a function of the integral radius π_{max}. Horizontal lines show that the ACF saturates for $\pi_{max} > 40$ h^{-1}Mpc which is also the minimum π_{max} at which $w_p(r_p)$ converges and returns the smaller error on the best-fit correlation parameter r_0, with γ fixed at 1.8.

red solid curve obtained by Gaussian smoothing.

4.4 Two-point Statistics

In §2.4 I introduced the two-point correlation function $\xi(r)$ as a technique for measuring the spatial clustering of a class of objects and I showed that, with a sample of AGN with known redshift, we cannot directly measure $\xi(r)$ in physical space, because peculiar motions of galaxies distort the line-of-sight distances inferred from redshift.

The solution is to measure the spatial correlation function in two dimensions r_p and π, where r_p and π are the projected comoving separations between the considered objects in the directions perpendicular and parallel, respectively. Since the redshift space distortions only affect the correlation function along the line of sight, one can define the so-called projected correlation function $w_p(r_p)$ (Davis & Peebles, 1983):

$$w_p(r_p) = 2 \int_0^{\pi_{max}} \xi(r_p, \pi) d\pi \tag{4.1}$$

where $\xi(r_p, \pi)$ is the two-point correlation function in term of r_p and π, measured using the Landy & Szalay (1993, LS) estimator (see §2.4 for more details).

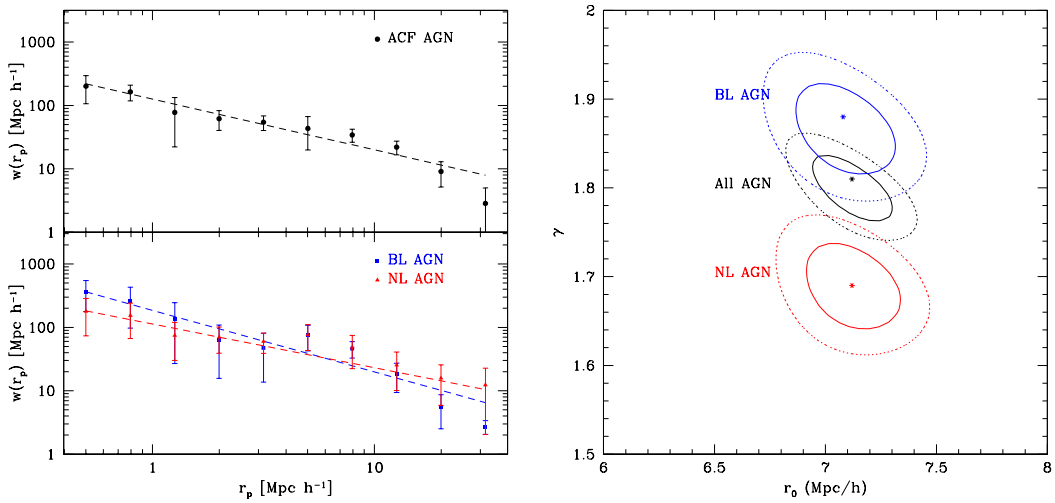

Figure 4.6 *Left panel*: Projected AGN ACF (black circles) compared to the auto-correlation of BL AGN (blue squares) and NL AGN (red triangles). The data points are fitted with a power-law model using the χ^2 minimization technique; the errors are computed with a bootstrap resampling method. *Right panel*: The confidence contours of the power-law best-fit parameters r_0 and γ, for the whole AGN sample (black), for the BL AGN (blue) and NL AGN (red) sub-samples. The contours mark the 68.3% and 95.4% confidence levels (respectively corresponding to $\Delta\chi^2 = 2.3$ and 6.17) which are plotted as continuous and dotted lines.

The LS estimator has been used to measure correlations in a number of surveys such as SDSS (Zehavi et al., 2005; Li et al., 2006), DEEP2 (Coil et al. , 2007; Coil et al., 2009), AGES (Hickox et al., 2009), COSMOS (Gilli et al., 2009). In practice, one truncates the integral at a finite π_{max} value, to maximize the correlation signal. One should avoid values of π_{max} too large since they would add noise to the estimate of $w_p(r_p)$; if instead, π_{max} is too small one would not recover all the signal. To determine the appropriate π_{max} values for the XMM-COSMOS AGN correlation function, I estimated $w_p(r_p)$ for different values of π_{max} in the range 20-120 h^{-1}Mpc. Besides, I determined the correlation length r_0 for this set of π_{max} values, by fitting $w_p(r_p)$ with a fixed γ=1.8 over r_p in the range 0.5-40 h^{-1}Mpc. In Fig. 4.5 I show the increase of the projected AGN auto-correlation $w_p(r_p)$ as a function of the integration radius π_{max}. The $w_p(r_p)$ values appear to converge for $\pi_{max} > 40$ h^{-1}Mpc. Therefore we adopt π_{max}= 40 h^{-1}Mpc in the following analysis, which is the minimum π_{max} at which the correlation function converges. Such π_{max} selection returns the smallest error on the best-fit correlation parameter r_0.

4.5 ACF: Standard Approach

To estimate the AGN auto-correlation function $\xi(r_p, \pi)$ using the LS formula, I created a grid with r_p and π in the range 0.1-100 h^{-1}Mpc, in logarithmic bins $\Delta log(r_p, \pi) = 0.2$ and we projected $\xi(r_p, \pi)$ on r_p using Eq. 4.1. In literature, several methods are adopted for error estimates in two-point statistics and no one has been proved to be the most precise. It is known that Poisson estimators generally underestimate the variance because they do not account for the fact that the points are not statistically independent, i.e. the same objects appear in more than one pair. In this work I computed the errors on $w_p(r_p)$ with a bootstrap resampling technique (Coil et al., 2009; Hickox et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010).

The standard approach used to evaluate the power of the clustering signal is to fit $w_p(r_p)$ with a power-law model (Coil et al., 2009; Hickox et al., 2009; Gilli et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010) of the form given in Eq. 2.13, using a χ^2 minimization technique, with γ and r_0 as free parameters. Fig. 4.6 (*left panel, upper quadrant*) shows the projected AGN ACF, evaluated in the projected separation range r_p= 0.5-40 h^{-1}Mpc. The best-fit correlation length and slope and the corresponding 1σ errors, are found to be $r_0 = 7.12^{+0.28}_{-0.18}$ h^{-1}Mpc and $\gamma = 1.81^{+0.04}_{-0.03}$. The projected correlation function of BL and NL AGN in the range $r_p = 0.5 - 40$ h^{-1}Mpc, are shown in Fig. 4.6 (*left panel, lower quadrant*). For BL AGN I found a correlation length $r_0 = 7.08^{+0.30}_{-0.28}$ h^{-1}Mpc and $\gamma = 1.88^{+0.04}_{-0.06}$, while for NL AGN I measured $r_0 = 7.12^{+0.22}_{-0.20}$ h^{-1}Mpc and a flatter slope $\gamma = 1.69^{+0.05}_{-0.05}$. Fig. 4.6 (*right panel*) shows the power-law best-fit parameters for the different AGN samples with the 1σ and 2σ confidence intervals for a two parameter fit, which correspond to $\chi^2 = \chi^2_{min} + 2.3$ and $\chi^2 = \chi^2_{min} + 6.17$.

As discussed in §2.6, the AGN bias factor can be estimated by using the power-law best fit parameters:

$$b_{PL} = \sigma_{8,AGN}(z)/\sigma_{DM}(z) \tag{4.2}$$

where $\sigma_{8,AGN}(z)$ is rms fluctuations of the density distribution over the sphere with a comoving radius of 8 h^{-1}Mpc, $\sigma_{DM}(z)$ is the DM correlation function evaluated at 8 h^{-1}Mpc, normalized to a value of $\sigma_{DM}(z = 0) = 0.8$. For a power-law correlation function this value can be calculated by (Peebles, 1980):

$$(\sigma_{8,AGN})^2 = J_2(\gamma)(\frac{r_0}{8Mpc/h})^\gamma \tag{4.3}$$

where $J_2(\gamma) = 72/[(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma]$. As the linear regime of the structure formation is verified only at large scales, the best-fit parameters r_0 and γ are estimated fitting the projected correlation function on $r_p = 1 - 40$ h^{-1}Mpc. The 1σ uncertainty of $\sigma_{8,AGN}$ is computed from the r_0 vs. γ confidence contour of the two-parameter fit corresponding to $\chi^2 = \chi^2_{min} + 2.3$.

Table 4.1 Bias Factors and hosting DM halo masses

(1) *AGN* Sample	(2) $\langle z \rangle$[a]	(3) b_{PL} Eq. 4.2	(4) b_{2h} Eq. 5.13	(5) $log\overline{M}_h$[b] $h^{-1}M_\odot$
Total (593)	1.22	$2.80^{+0.22}_{-0.90}$	2.98 ± 0.13	13.23 ± 0.06
BL (354)	1.55	$3.11^{+0.30}_{-1.22}$	3.43 ± 0.17	13.14 ± 0.07
NL (239)	0.74	$2.78^{+0.45}_{-1.07}$	2.70 ± 0.22	13.54 ± 0.10
X-unobs (184)	1.12	$2.98^{+0.34}_{-0.37}$	3.01 ± 0.21	13.33 ± 0.08
X-obs (218)	1.30	$1.66^{+0.31}_{-0.32}$	1.80 ± 0.15	12.30 ± 0.15
Subsample at $z < 1$				
BL (70)	0.57	$2.18^{+0.95}_{-1.02}$	2.32 ± 0.26	13.50 ± 0.11
NL (137)	0.53	$1.68^{+0.45}_{-0.57}$	1.40 ± 0.15	12.65 ± 0.18

[a]Median redshift of the sample.
[b]Typical DM halo mass based on Sheth et al. (2001) and van den Bosch (2002).

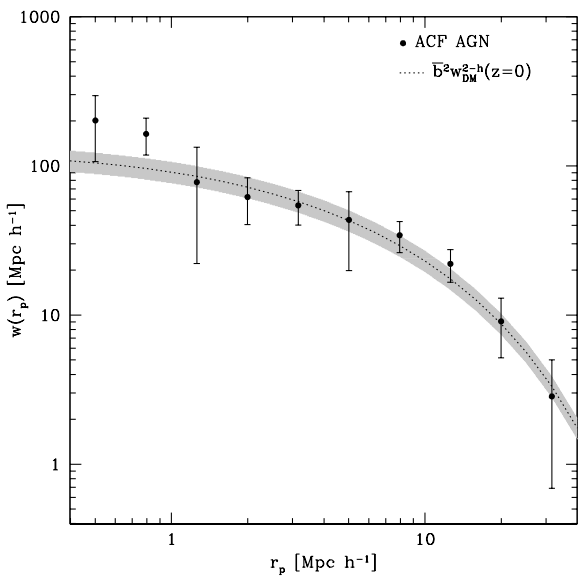

Figure 4.7 : Projected AGN ACF (black circles) compared to $\overline{b}^2 w^{2h}_{DM}(r_p, z = 0)$ (dotted line), where the weighed bias $\overline{b}$ is defined in Eq. 4.12. The shaded region shows the projected DM 2-halo term scaled by $(\overline{b} \pm \delta\overline{b})^2$.

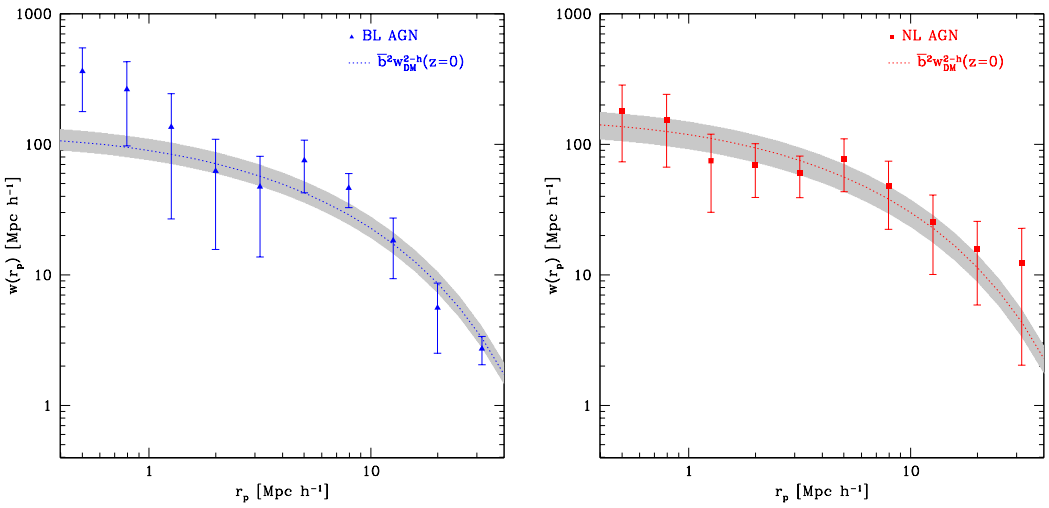

Figure 4.8 : Projected ACF of BL AGN (blue triangles, *left panel*) and NL AGN (red squares, *right panel*), compared to $\bar{b}^2 w_{DM}^{2h}(r_p, z=0)$ (dotted line), where the weighed bias $\bar{b}$ is defined in Eq. 4.12. The shaded region shows the projected DM 2-halo term scaled by $(\bar{b} \pm \delta\bar{b})^2$.

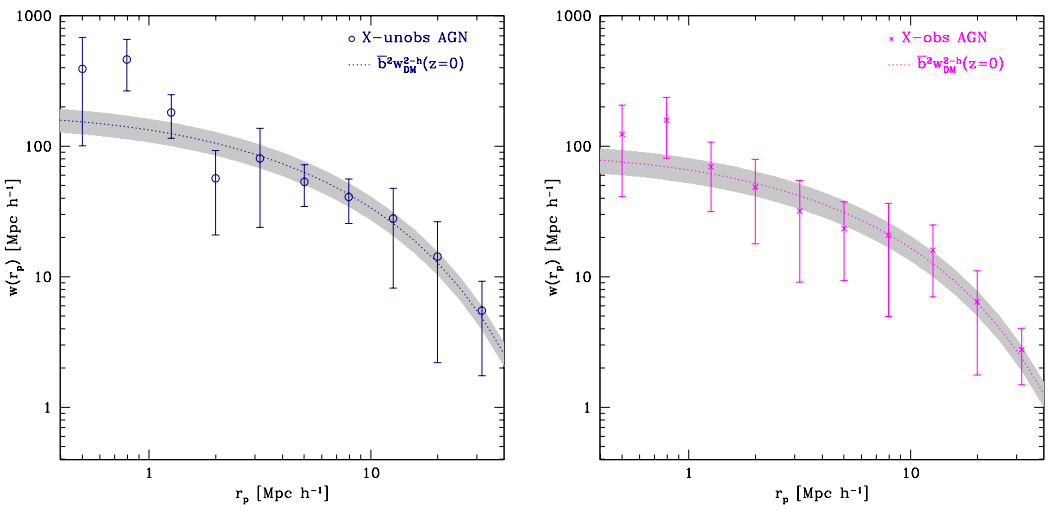

Figure 4.9 : Projected ACF of X-unobs AGN (darkblue open circles, *left panel*) and X-obs AGN (magenta diagonal crosses, *right panel*), compared to $\bar{b}^2 w_{DM}^{2h}(r_p, z=0)$ (dotted line), where the weighed bias $\bar{b}$ is defined in Eq. 4.12. The shaded region shows the projected DM 2-halo term scaled by $(\bar{b} \pm \delta\bar{b})^2$.

In the halo model approach, the clustering signal can be modelled as the sum of two contributions of pairs from the same DM halo (1-halo term) and those from different DM halos (2-halo term). In Fourier space, the 2-halo term can be explicitly written as (Seljak 2000, Cooray & Sheth 2002):

$$P_{2h} \approx b^2 P_{lin}(k, z) \tag{4.4}$$

where $P_{lin}(k, z)$ is the linear power spectrum and b is the bias factor of the sample (see §3.6 for more details). Then the AGN two-point correlation function at large scales is given by:

$$w_{p,2h}(r_p) = b^2_{AGN} \int \frac{k}{2\pi} P_{lin}(k) J_0(kr_p) dk \tag{4.5}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind. Following this model, the AGN bias defines the relation between the 2-halo term of DM and AGN clustering signal:

$$b^2_{AGN}(r_p) = \frac{w_{p,2h}(r_p)}{w^{2h}_{DM}(r_p, z = 0)} \tag{4.6}$$

According to this equation, I estimated the average bias factor in the range $r_p = 1 - 40$ h^{-1}Mpc. Table 4.1, column 4 shows the AGN bias factors using this method, compared with the ones based on the power-law fits of the ACF (column 3) for the different AGN subsets. The error on b_{AGN} corresponds to $\Delta\chi^2 = 1$ using a χ^2 minimization technique with 1 free parameter. The two sets of bias values from the different approaches are consistent within 1σ, but the errors on b_{PL} are bigger consistently with the fact that the AGN ACF is not well described by a power-law.

4.6 Solving for Sample Variance using HOD

The standard approaches used in previous works on clustering of X-ray AGN (Mullis et al., 2004; Yang et al., 2006; Gilli et al., 2005; Coil et al., 2009; Hickox et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010) to estimate the bias factors from the projected AGN ACF are based on the power-law fit parameters (method 1). This method assumes that the projected correlation function is well fitted by a power-law and the bias factors are derived from the best fit parameters r_0 and γ of the clustering signal at large scale.

Most of the authors (Hickox et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010) used an analytical expression (as the one described in Sheth & Tormen, 1999; Sheth et al., 2001; Tinker et al., 2005) to assign a characteristic DM halo mass to the hosting halos. The incongruity of this approach is that the bias used is the average bias of a given sample at a given redshift. In fact, following this approach

one can not take into account that the average bias is sensitive to the entirety of the halo mass distribution; different mass distributions with different average masses can give rise to the same average bias. On the contrary, by using the halo model, the average bias and the average mass of the sample properly account for the shape of the mass distribution: the average bias depends on the halo mass function and bias and on the AGN HOD, integrated over the total mass range.

At non-linear scales, the HOD defines how AGN can fill DMHs, allowing to model the distribution of AGN within halos (see §3.5). In literature, the common model used for the AGN HOD is a softened step function for the HOD of central AGN and a truncated power-law satellite HOD (introduced by Zehavi et al., 2005, for galaxies). Here I decided to assumed that all the AGN reside in central galaxies and that the AGN HOD follows a simple parametric form (more complicated models will be discussed in Chapter 5). This assumption is supported by Starikova et al. (2011). They found that X-ray AGN are predominantly located in the central galaxies of the host DM halos and tend to avoid satellite galaxies, fixing the limit to the fraction of AGN in non-central galaxies to be less than 10%. The same fraction of satellite galaxies hosting AGN is suggested in Shen (2009). Shankar et al. (2010) modelled the measurements of quasar clustering derived in the SDSS (Shen et al., 2009) and they verified that the predicted bias factors and the correlation functions are not altered including subhalos as quasar hosts. A further consideration is that there is in practice no distinction between central and satellite AGN in the 2-halo term that I used to estimate the AGN bias factor.

I assumed an AGN halo occupation $\langle N_{AGN}\rangle(M_h)$ described by a delta function:

$$\langle N_{AGN}\rangle(M_h) = f_A\delta(M_h - M_0) \tag{4.7}$$

where f_A is the AGN duty cycle. It is clear that I am not considering the full HOD model, but I am assigning to all the AGN the same average mass of the hosting halos. The motivation is that X-ray AGN mainly reside in massive halos with a narrow distribution of the hosting halo masses. It is clear that this assumption is specific to AGN and e.g. is not applicable to galaxies.

The δ-function is the simplest possible assumption in the treatment of the sample variance, which is due to the variation in the amplitude of source counts distribution. It has been shown in Faltenbacher et al. (2010) that the variation in the density field, which is responsible for the sample variance, can be replaced by the variation of the halo mass function. In terms of halo model, the bias factor as a function of the fluctuations Δ in the density field is expressed by:

$$b_A(\Delta) = \frac{\int_{M_h}\langle N_{AGN}\rangle(M_h)b_h(M_h)n(M_h,\Delta)dM_h}{\int_{M_h}\langle N_{AGN}\rangle(M_h)n(M_h,\Delta)dM_h} \tag{4.8}$$

where $\langle N_{AGN}\rangle$ is the AGN HOD, $b_h(M_h)$ is the halo bias and $n(M_h,\Delta)$ is the halo mass function, which depends on the density field. On the other hand the sample

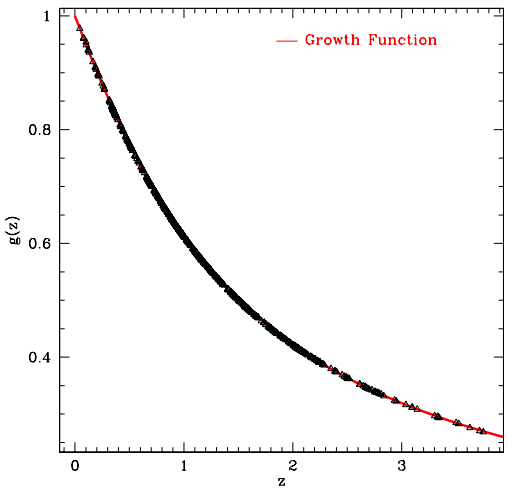

Figure 4.10 : Factor g as defined in Eq. 4.11, estimated at the redshift of each AGN (black triangles), compared to the growth function $D_1(z)/D_1(z=0)$ (red line, see Eq. (10) in Eisenstein & Hu 1999 and references therein). The bias of each AGN is weighted by this factor according to the redshift z of the source.

variance does not affect the AGN halo occupation. When we assume that all AGN reside in DM halos with the same mass, Eq. 4.8 becomes simpler:

$$\frac{\int_{M_h} \delta(M_h - M_0) b_h(M_h) n(M_h, \Delta) dM_h}{\int_{M_h} \delta(M_h - M_0) n(M_h, \Delta) dM_h} = b(M_0) \tag{4.9}$$

The equation shows that when the AGN HOD is close to a δ-function, the variations in the density field only change the AGN number density and put more weight on AGN bias at the redshift of large scale structure, but do not change the bias of AGN inside the structure. Our claim differs from the results presented in Gilli et al. (2005, 2009). They found that excluding sources located within a large-scale structures, the correlation length and then the bias factor strongly reduces. Such bias behaviour can be used to constrain more complicated shapes of the AGN HOD than a δ-function type distribution. However, even in the case of a δ-function HOD, two effects which are often omitted in the clustering analysis need to be taken into account: the LSS growth and the evolution of the bias factor with z. Ignoring these effects can by itself lead to a difference in the results reported for the different AGN samples.

The bias factor depends on the redshift as the structures grow over time, associated with the large redshift interval of XMM-COSMOS AGN. For the ith source at

redshift z_i, I considered the bias factor corresponding to the halo mass $M_h = M_0$:

$$b_i = b(M_0, z_i) \tag{4.10}$$

where $b(M_0, z)$ is evaluated using van den Bosch (2002) and Sheth et al. (2001). For each AGN at redshift z I estimated the factor $g(z)$ defined as the square root of the projected DM 2-halo term at redshift z normalized to the projected DM 2-halo term evaluated at $z = 0$:

$$g(z) = \sqrt{\frac{w_{DM}(z, r_p)}{w_{DM}(z = 0, r_p)}} \tag{4.11}$$

averaged over the scales $r_p = 1 - 40$ h^{-1}Mpc. As the amplitude of the projected DM 2-halo term decreases with increasing redshift, g is a decreasing function of z (see fig. 4.10), well described by the term $D_1(z)/D_1(z = 0)$, where $D_1(z)$ is the growth function (see Eq. (10) in Eisenstein & Hu (1998) and references therein).
By accounting for the fact that the linear regime of the structure formation is verified only at large scales, I estimated the AGN bias considering only the pairs which contribute to the AGN clustering signal at $r_p = 1 - 40$ h^{-1}Mpc. Hence, the *weighted* bias factor of the sample is defined as:

$$\bar{b}(M_0) = \sqrt{\frac{\sum_{i,j} b_i b_j g_i g_j}{N_{pair}}} \tag{4.12}$$

where $b_i b_j$ is the bias factor of the pair $i - j$, $g_i g_j$ is the g factor of the pair and N_{pair} is the total number of pairs in the range $r_p = 1 - 40$ h^{-1}Mpc.
Similarly, I defined a weighted average redshift of the AGN sample, weighting the redshift of each pair for the g factor and the bias of the pair ($b_i b_j$):

$$\bar{z} = \frac{\sum_{i,j} b_i b_j g_i g_j z_{pair}}{\sum_{i,j} b_i b_j g_i g_j} \tag{4.13}$$

where $z_{pair} = (z_i + z_j)/2$. Following this approach I estimated the value of M_0 that satisfies:

$$b_1 = \bar{b}(M_0)$$

where b_1 is the square root of the observed AGN ACF normalized to the projected DM 2-halo term at $z = 0$:

$$b_1 = \sqrt{\frac{w_{AGN}(r_p)}{w_{DM}(z = 0, r_p)}} \tag{4.14}$$

averaged over the scale $r_p = 1 - 40$ h^{-1}Mpc.

4.7 Measurements

The weighted bias factors $\overline{b}$ and redshifts $\overline{z}$, and the corresponding DM halo masses M_0 are shown in Table 4.2 for the different AGN sub-sample. Fig. 4.7, 4.8 and 4.9 show the ACF of the AGN, BL/NL AGN and X-unobs/obs AGN samples, compared to the term $\overline{b}^2 w^{2h}_{DM}(r_p, z = 0)$ (dotted line), where the weighed bias $\overline{b}$ is defined in Eq. 4.12. The shaded region shows the projected DM 2-halo term scaled by $(\overline{b} \pm \delta\overline{b})^2$. The AGN bias estimates indicate that XMM-COSMOS AGN reside in halos with average mass $logM_0[\mathrm{h}^{-1}M_\odot] = 13.01 \pm 0.09$, characteristic of moderate-size poor groups. This result is consistent with previous works on X-ray selected AGN that indicate that the typical mass of AGN hsoting halos is in the range $12.5 \lesssim logM_h[\mathrm{h}^{-1}M_\odot] \lesssim 13.5$. On the other hand, I found that BL and NL AGN (at $\overline{z} = 1.53$ and $\overline{z} = 0.82$, respectively), present bias factors which correspond to average DMH masses $logM_0[\mathrm{h}^{-1}M_\odot] = 13.24 \pm 0.06$ and 13.01 ± 0.08, respectively.

As described in Brusa et al. (2010), only a small fraction of the objects classified as NL AGN are located at $z > 1$, to be compared with 350 in the BL AGN sample. This is mostly due to the fact that high-redshift NL AGN are optically faint (typically $I \sim 23-24$) and have not been targeted yet with dedicated spectroscopic campaigns. Then these results might be affected by the limitations in the obscured AGN classification, considering that some models on the evolution of the obscured AGN fraction predict an increase of the fraction with the redshift (Hasinger et al., 2008).

In order to avoid the problem of different redshift distribution in comparing BL/NL AGN clustering amplitude, I selected for each sample a subset (70 BL AGN and 137 NL AGN) at $\overline{z} \sim 0.6$. At the same redshift I found that BL and NL AGN have a bias factor $\overline{b}_{BL} = 1.62 \pm 0.26$ and $\overline{b}_{NL} = 1.56 \pm 0.15$, which correspond to average halo masses $logM_0$ $[\mathrm{h}^{-1}M_\odot]$=13.27±0.10 and 12.97 ± 0.07, respectively. Similar results have been obtained using X-unobs and X-obs AGN samples; unobscured AGN at $\overline{z} = 1.16$ inhabit halos with average mass $logM_0$ $[\mathrm{h}^{-1}M_\odot] = 13.30 \pm 0.10$ which is higher at 2.5 σ level than the halo mass hosting obscured AGN ($logM_0$ $[\mathrm{h}^{-1}M_\odot]$ =12.97 ±0.08), at similar redshift.

In order to compare these results with previous works on the bias of X-ray selected AGN, I evaluated the bias factors corresponding to the halo mass M_0 at $\overline{z}$ using Sheth et al. (2001) as shown in Table 4.2, col (5). These results support the picture that at a given redshift, X-ray selected BL/X-unobs AGN reside in more massive halos compared to X-ray selected NL/X-obs AGN, that would be expected if the two classes of AGN correspond to different phases of the AGN evolution sequence (Hopkins et al., 2006, 2008; Hickox et al., 2009).

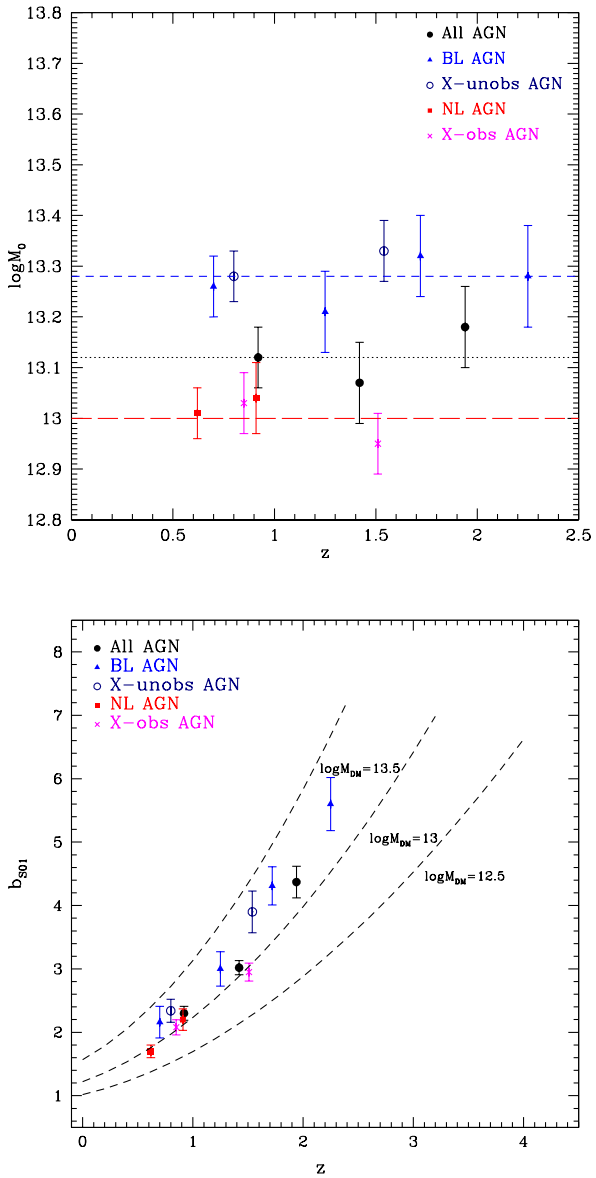

Figure 4.11 *Upper Panel*: DM halo mass M_0 as a function of z for different AGN sub-samples (see legend). The horizontal lines show the mean value of M_0 for BL/X-unobs AGN (dashed-blue), NL/X-obs AGN (long dashed-red) and for the whole AGN sample (dotted-black). The masses are given in units of $\mathrm{h}^{-1}M_{\odot}$. *Lower Panel*: Redshift evolution of the bias parameter b_{S01} of different AGN sub-samples. The dashed lines show the expected $b(z)$ of typical DM halo masses M_h based on Sheth et al. (2001). BL/X-unobs AGN present a strong bias evolution with redshift with a constant DM halo mass $\log M_0$ $[\mathrm{h}^{-1}M_{\odot}] = 13.28 \pm 0.07$ up to $\overline{z} \sim 2.4$. NL/X-obs AGN reside in less massive halos with $\log M_0$ $[\mathrm{h}^{-1}M_{\odot}] = 13.00 \pm 0.06$, constant at $\overline{z} < 1.5$.

Table 4.2 Weighted Bias factors and hosting DM halo masses

(1) *AGN* Sample	(2) $\overline{b}$ Eq. 4.12	(3) $\overline{z}$ Eq. 4.13	(4) $logM_0$ $\mathrm{h}^{-1}M_\odot$	(5) b_{S01}[a]
Total (593)	1.91 ± 0.13	1.21	13.10 ± 0.06	2.71 ± 0.14
BL (354)	1.74 ± 0.17	1.53	13.24 ± 0.06	3.68 ± 0.27
NL (239)	1.80 ± 0.22	0.82	13.01 ± 0.08	2.00 ± 0.12
X-unobs (184)	1.95 ± 0.21	1.16	13.30 ± 0.10	3.01 ± 0.26
X-obs (218)	1.37 ± 0.15	1.02	12.97 ± 0.08	2.23 ± 0.13
Subsample at $z < 1$				
BL (70)	1.62 ± 0.26	0.63	13.27 ± 0.10	1.95 ± 0.17
NL (137)	1.56 ± 0.15	0.60	12.97 ± 0.07	1.62 ± 0.15

[a]Bias estimated from M_0 using Sheth et al. (2001).

4.8 Redshift Evolution of the AGN Bias

In order to investigate the redshift evolution of the bias factor, I split the XMM-COSMOS AGN sample in three redshift bins. The sizes of the redshift bins have been determined such that there are more or less the same number of objects in each bin. The values of $\overline{b}$, $\overline{z}$ and M_0 for the total AGN sample are shown in Table 4.3. The meaning of the table columns are: (1) sample; (2) number of sources; (3) bias parameter from the projected DM 2-halo term, evaluated at the median $\langle z \rangle$ of the sample; (4) typical halo mass using van den Bosch (2002) and Sheth et al. (2001); (5) weighted bias of the sample; (6) weighted redshift of the sample; (7) Average DM halo mass; (8) Bias factor from M_0 estimated using Sheth et al. (2001). I observed an increase of the AGN bias factor with redshift, from $\overline{b}(\overline{z} = 0.92) = 1.80 \pm 0.19$ to $\overline{b}(\overline{z} = 1.94) = 2.63 \pm 0.21$ with a DM halo mass consistent with being constant at $logM_0[\mathrm{h}^{-1}M_\odot] \sim 13.1$ in each bin. These results support the picture that the bias of XMM-COSMOS AGN evolves with time according to a constant halo mass track at any redshift $z < 2$.

This conclusion, based on the analysis of the global XMM-COSMOS AGN sample, can however be affected by the fact that the relative proportions of BL and NL AGN are a strong function of redshift. In fact, since the XMM-COSMOS AGN sample is a flux limited sample, more luminous AGN are selected at high redshift and, also because of our magnitude limit, high-z sources in our sample are mainly BL AGN (see §4.3). For this reason BL AGN sample can be analysed up

Table 4.3. Bias Evolution

(1) $\langle z \rangle$[a]	(2) **N**	(3) b_{2-h} Eq. 5.13	(4) $log\overline{M}_{DM}$[b] $h^{-1}M_{\odot}$	(5) $\overline{b}$	(6) $\overline{z}$	(7) $logM_0$ Eq. 5.13	(8) b_{S01}[c] $h^{-1}M_{\odot}$
				All AGN			
0.80	190	2.70 ± 0.19	13.48 ± 0.10	1.80 ± 0.19	0.92	13.12 ± 0.06	2.30 ± 0.11
1.30	220	3.10 ± 0.18	13.21 ± 0.10	2.14 ± 0.18	1.42	13.07 ± 0.08	3.02 ± 0.11
2.07	183	5.18 ± 0.21	13.30 ± 0.11	2.63 ± 0.21	1.94	13.18 ± 0.08	4.37 ± 0.27
				BL AGN			
0.67	70	2.62 ± 0.20	13.57 ± 0.10	1.52 ± 0.20	0.70	13.26 ± 0.06	2.16 ± 0.25
1.25	108	3.06 ± 0.23	13.24 ± 0.08	2.02 ± 0.23	1.25	13.21 ± 0.08	3.00 ± 0.27
1.71	92	5.37 ± 0.28	13.60 ± 0.08	3.57 ± 0.28	1.72	13.32 ± 0.08	4.31 ± 0.30
2.46	85	6.82 ± 0.27	13.41 ± 0.10	4.02 ± 0.27	2.25	13.28 ± 0.10	5.60 ± 0.42
				X-unobscured AGN			
0.65	98	2.46 ± 0.17	13.51 ± 0.11	1.62 ± 0.17	0.80	13.28 ± 0.05	2.34 ± 0.18
1.66	86	4.85 ± 0.18	13.51 ± 0.10	2.10 ± 0.18	1.54	13.33 ± 0.06	3.90 ± 0.33
				NL AGN			
0.53	137	1.40 ± 0.13	12.65 ± 0.12	1.59 ± 0.13	0.62	13.01 ± 0.05	1.70 ± 0.10
1.02	102	2.11 ± 0.19	12.88 ± 0.15	1.87 ± 0.19	0.91	13.04 ± 0.07	2.20 ± 0.17
				X-obscured AGN			
0.73	106	1.80 ± 0.14	13.01 ± 0.11	1.51 ± 0.14	0.85	13.03 ± 0.06	2.08 ± 0.12
1.84	112	3.51 ± 0.16	12.94 ± 0.13	1.96 ± 0.16	1.51	12.95 ± 0.06	2.95 ± 0.14

[a]Median redshift of the sample.

[b]Typical DM halo mass based on Sheth et al. (2001) and van den Bosch (2002).

[c]Bias estimated from M_0 using Sheth et al. (2001).

to $z \sim 2.25$, while the maximum average redshift for NL AGN is $z \sim 0.91$.

I found evidence of a strong increase of the BL AGN bias factor in four redshift bins (see Table 4.3), with a DM halo mass constant at $\log M_0[\mathrm{h}^{-1}M_\odot] \sim 13.28$ at all redshifts $z < 2.25$. For NL AGN I estimated $\bar{b}(\overline{z} = 0.62) = 1.59 \pm 0.13$ and $\bar{b}(\overline{z} = 0.91) = 1.87 \pm 0.19$, which correspond to a constant halo mass $\log M_0[\mathrm{h}^{-1}M_\odot] \sim 13.02$. Moreover, I split the X-unobs and X-obs AGN samples in two redshift bins up to $\overline{z} \simeq 1.5$ and I found that the bias of X-unobs AGN (X-obs AGN) evolves according to a constant halo mass consistent with the mass of BL AGN (NL AGN) hosting halos.

Fig. 4.11 (*Upper Panel*) shows the redshift evolution of the average DM halo mass M_0 for all the AGN subsets. The horizontal lines represent the mean value of M_0 for BL/X-unobs AGN (dashed-blue), NL/X-obs AGN (long dashed-red) and for the whole AGN sample (dotted-black). Fig. 4.11 (*Lower Panel*) shows the redshift evolution of the bias factors b_{S01} (Table 4.3, col 7) for different AGN sub-samples. The dashed lines show the expected $b(z)$ associated to the typical DM halo mass based on Sheth et al. (2001).

These results show that X-ray selected BL/X-unobs AGN reside in more massive DM halos compared to X-ray selected NL/X-obs AGN at $\sim 3\sigma$ level. This suggests that the AGN activity is a mass triggered phenomenon and that different AGN phases are associated with the DM halo mass, irrespective of redshift z.

4.9 Discussion

4.9.1 Which DM halos host X-ray AGN?

I introduced a new method that uses the 2-halo term to estimate the AGN bias factor and that properly accounts for the growth of the structures over time associated with our use of AGN in a broad redshift interval. Using this approach I estimated an average mass of the XMM-COSMOS AGN hosting halos equal to $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.10 \pm 0.06$ which differs at $\sim 1.6\sigma$ level from the typical halo mass M_h based on Sheth et al. (2001) using the method 2 (see §4.6). The difference between the standard method and our own method is also clear for the mass of BL and NL AGN hosting halos. We have found that BL AGN inhabit DM halos with average mass $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.24 \pm 0.06$ at $\overline{z} = 1.53$ while halos hosting NL AGN have average mass $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.01 \pm 0.08$. BL AGN reside in more massive halos than NL AGN also selecting two subsamples that peak at the same median redshift $\overline{z} \sim 0.6$. I obtained similar results using X-ray unobscured AGN at $\overline{z} = 1.16$ and X-ray obscured AGN at $\overline{z} = 1.02$ ($\log M_0[\mathrm{h}^{-1}M_\odot] = 13.30 \pm 0.10$ and $\log M_0[\mathrm{h}^{-1}M_\odot] = 12.97 \pm 0.08$, respectively).

Instead the typical halo mass based on Sheth et al. (2001) using the AGN bias estimated with the method 2, strongly depends on the median redshift of the

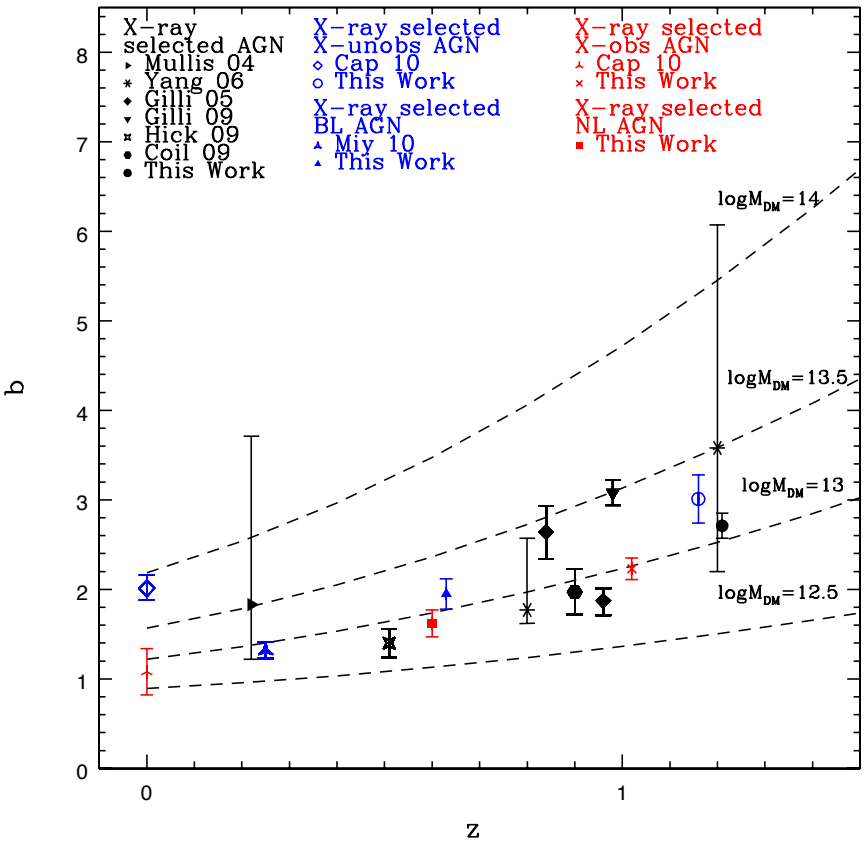

Figure 4.12 : Bias parameter as a function of redshift for various X-ray selected AGN (black data points), X-ray selected BL/X-unobs AGN (blue data points) and X-ray selected NL/X-obs AGN (red data points) as estimated in previous studies and in this work according to the legend. Our results refer to the bias factor b_{S01} shown in Table 5.2 col (5). The dashed lines show the expected $b(z)$ of typical DM halo masses M_h based on Sheth et al. (2001) . The masses are given in $logM_h$ in units of $\mathrm{h}^{-1}M_{\odot}$.

sample. According to the method 2, BL AGN at $\langle z \rangle = 1.55$ reside in less massive halos compared to NL AGN at $\langle z \rangle = 0.74$, while the result is different selecting two samples of BL and NL AGN at the same $\langle z \rangle \sim 0.5$. These results agrees with the majority of the recent studies of X-ray surveys which suggest a picture in which X-ray AGN are typically hosted in DM halos with mass in the range $12.5 < logM_h[\mathrm{h}^{-1}M_\odot] < 13.5$, at low ($< 0.4$) and high ($\sim 1$) redshift.

Fig. 4.12 shows the bias factors of X-ray selected AGN (black), BL/X-unobs AGN (blue) and NL/X-obs AGN (red) as estimated in different surveys (according to the legend). The dashed lines show the expected $b(z)$ assuming a constant typical DM halo mass M_{DM}, based on Sheth et al. (2001). Starikova et al. (2011) found that *Chandra/Bootes* AGN are located at the center of DM halos with $M > M_{min} = 4 \times 10^{12}$ $\mathrm{h}^{-1}M_\odot$. This mass estimate represents a threshold value, since they are assuming a halo occupation described by a step function (zero AGN per halo/subhalo below M_{min} and one above it). The previous studies of Gilli et al. (2005) for the CDFN, Gilli et al. (2009), Mullis et al. (2004), Yang et al. (2006) for CLASXS AGN suggest the scenario in which the typical DM halo mass hosting X-ray selected AGN is $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 13.5$. The bias values measured in Gilli et al. (2005) on CDFS, in Hickox et al. (2009), Coil et al. (2009) and Yang et al. (2006) and in this work, correspond to a lower halo mass ($\log M_h[\mathrm{h}^{-1}M_\odot] \sim 13$). A possible explanation could be that at fixed redshift, the bias and then the mass of the hosting halo, depends on the luminosity of the sample. The same explanation might be applied to the results on BL/X-unobs AGN. The bias estimates at $z < 1$ for NL/X-obs AGN in Cappelluti et al. (2010) and in this work, seem to indicate that the mass of NL/X-obs AGN hosting halos is $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 13$.

4.9.2 Optically selected vs X-ray selected AGN

I first found evidence of a redshift evolution of the bias factor of X-ray selected BL/ X-unobs AGN (fig. 4.13, blue data points) and NL/X-obs AGN (red data points). The bias evolves with redshift at constant average halo mass $\log M_0$ $[\mathrm{h}^{-1}M_\odot] \sim 13.3$ for BL/X-unobs AGN and $\log M_0[\mathrm{h}^{-1}M_\odot] \sim 13$ for NL/X-obs AGN at $z < 2.25$ and $z < 1.5$, respectively. The observed bias evolution suggests an average halo mass of the hosting halos, constant over time in the range $\log M_h[\mathrm{h}^{-1}M_\odot]$ = 13-13.5, instead of an evolution of the bias in a model in which objects are formed at a fixed time and their distribution evolves under the influence of gravity (Fry et al., 1996).

There have been several studies of the bias evolution of optical quasars with the redshift as shown in Fig. 4.13 (grey data points), based on large survey samples such as 2QZ and SDSS (Croom et al., 2005; Porciani & Norberg, 2006; Shen et al., 2009; Ross et al., 2009). Since the quasar samples used in these clustering analysis are defined as spectroscopically identified quasars with at least one broad

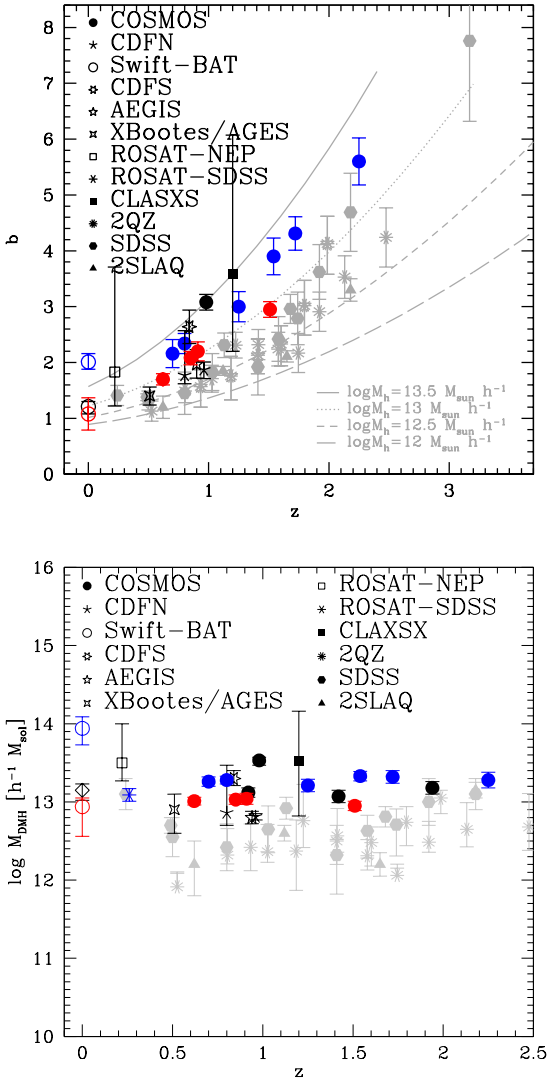

Figure 4.13 : Bias factor (*Upper Panel*) and mass of AGN hosting halos (*Lower Panel*) as a function of redshift for X-ray selected AGN (black data points), X-ray selected Type 1 AGN (blue data points) and X-ray selected Type 2 AGN (red data points) as estimated in different surveys (COSMOS, Gilli et al. (2009); Allevato et al. (2011); CDFN, Gilli et al. (2005); Yang et al. (2006); Swift-BAT, Cappelluti et al. (2010); CDFS, Gilli et al. (2005); AEGIS, Coil et al. (2009); AGES, Hickox et al. (2009); ROSAT-NEP, Mullis et al. (2004); ROSAT-SDSS, Krumpe et al. (2010); CLASXS, Yang et al. (2006)). The dashed lines show the expected $b(z)$ of DMHs with different masses according to the legend, based on Sheth et al. (2001). The grey points show results from quasar ACF measurements using spectroscopic samples from SDSS (Ross et al., 2009; Shen et al., 2009), 2QZ (Croom et al., 2005; Porciani & Norberg, 2006) and 2SLAQ (da Ângela et al., 2008).

(FWHM>1000 kms^{-1}) emission line, I refer to them as optically selected BL AGN.

All the previous studies infer the picture that the quasar bias evolves with redshift following a constant mass evolution, with the average mass that can vary in the range $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 12.5 - 13$, may be depending on the AGN sample luminosity as already suggested for X-ray selected AGN. The simplest interpretation according to the observed redshift evolution of the bias factors is that 1) X-ray selected AGN whether BL/X-unobs or NL/X-obs AGN inhabit DM halos with mass higher than the mass of optically selected quasar hosting halos in the range $z = 0.5 - 2.25$; 2) X-ray selected BL/X-unobs AGN reside in more massive halos compared to NL/X-obs AGN for $z = 0.6 - 1.6$ and the discrepancy between the bias factors of the two samples increases with z; 3) the AGN activity is a mass triggered phenomena and the different AGN evolutionary phases are associated with just the DM halo mass, irrespective of the redshift z.

4.9.3 External vs Internal Triggering

The major merger of galaxies is one of the promising mechanisms suggested to be responsible for fuelling quasars and in particular to be dominant for bright quasars at high redshift. Models of major mergers appear to naturally produce many observed properties of quasars, as the quasar luminosity density, the shape and the evolution of the quasar luminosity function and the large-scale quasar clustering as a function of L and z (Hopkins et al., 2008; Shen, 2009; Shankar et al., 2009, 2010; Shankar, 2010; Bonoli et al., 2009).

Clear evidence for higher incidence of mergers is seen among quasars (Serber et al., 2006; Hopkins et al., 2006; Veilleux et al., 2009). Additionally a large fraction of luminous quasars at low redshift are associated with either morphologically disturbed objects (Canalizo & Stockton, 2001; Guyon et al., 2006), or early-type hosts with fine structure in their optical light distribution, indicative of past interactions (Canalizo et al., 2007; Bennert et al., 2008). In the local Universe, for instance, the study of the environment of Swift BAT Seyfert galaxies (Koss et al., 2010) find a larger fraction of BAT AGN with disturbed morphologies or in close physical pairs (<30 kpc) compared to matched control galaxies. The high rate of apparent mergers (25%) suggests that AGN activity and merging are critically linked for the moderate luminosity AGN in the BAT sample. Moreover it is believed that major merger dominates at high redshift and bright luminosities (Hasinger et al., 2008; Hopkins et al., 2006), while minor interaction or bar instabilities or minor tidal disruptions are important at low redshift ($z \lesssim 1$) and low luminosities ($L_{BOL} \lesssim 10^{44}$ erg s^{-1} Hopkins & Henquist, 2009).

The results presented in this Chapter on the bias evolution of X-ray selected BL/X-unobs AGN infer that these objects with $L_{BOL} \sim 2 \times 10^{45}$ erg s^{-1} reside in massive DM halos $M_h \sim 2 \times 10^{13}$ h$^{-1}M_\odot$. Besides studies on BL AGN in the

COSMOS field (Merloni et al., 2010; Trump et al., 2011) suggest that our sample is characterized by BH masses in the range $M_{BH} = 10^7 - 10^9 M_\odot$ and Eddington ratio $\lambda > 0.01$. Optically selected quasars from large survey samples such as 2QZ and SDSS are high-luminosity quasars $L_{BOL} \gtrsim 10^{46}$ erg s^{-1} with BH masses in the range $M_{BH} = 10^8 - 10^{10} M_\odot$ and $\lambda > 0.01$. Clustering analysis of optical quasars have shown that they reside in DM halos with $M_h \sim 10^{12}$ h$^{-1} M_\odot$.

Fig. 4.14 shows the predicted bias as a function of luminosity computed according to Shen (2009) at $z = 2$. The theoretical model which assumes a quasar phase triggered by major mergers predicts an increasing bias with luminosity and reproduces the previous results obtained for optical quasars at $1.8 < z < 2.2$ (Croom et al. (2005, green-crosses), Porciani & Norberg (2006, green-star), Shen et al. (2009, green-open square), da Ângela et al. (2008, green-circles), Myers et al. (2007, green-squares)). On the other hand the model can not reproduce the high bias factor found for X-ray selected COSMOS BL AGN (blue triangle) and then can not explain why optically selected quasars characterized by higher bolometric luminosity compared to X-ray selected COSMOS BL/X-unobs AGN, are found in less massive halos. These differences suggest a switch to a different dominant mechanism for the AGN triggering.

Hopkins & Hernquist (2006) introduced a model for the fueling of low-luminosity AGN (Seyferts, with $L_{BOL} \lesssim 10^{44} - 10^{45}$ erg s^{-1} and $M_{BH} \lesssim 10^7 M_\odot$), which proposes AGN triggered by random accretion of gas via internal, secular processes. The stochastic accretion model and the merger-driven activity are fundamentally different, the former being determined by stochastic encounters with a cold gas supply in a quiescent system, the latter by the violent torquing of cold gas throughout entire galaxies into the galaxy center in major mergers. Accretion of cold gas in quiescent systems can account for low luminosity Seyferts but can not explain the higher luminosities and the larger BH masses observed for XMM-COSMOS BL AGN. The high Eddington ratios at masses in the range $M_{BH} \sim 10^8 - 10^9 M_\odot$ can not be maintained through this mode of accretion.

Furthermore, this fueling mechanism predicts lower bias factors compared to the major merger picture for bright quasars, which is completely in disagreement with our results. Fueling by stellar winds or hot gas accretion may represent yet a third qualitatively distinct mode of fueling. Ciotti & Ostriker (1997); Ciotti & Ostriker (2001) investigated the episodic AGN activity model in early-type galaxies, assuming at their center the presence of a massive BH growing with the accretion of matter and affecting the inflow through feedback. The duration of the single accretion event are extremely short but the maximum luminosities reached during the accretion events can be of the order of $L_{BOL} \sim 10^{46} - 10^{47}$ erg s^{-1}, depending on the input parameters of the model. The central BH grows by episodic accretion up to a mass in the observed range ($M \sim 10^{8.5} - 10^{9.5} M_\odot$) in all giant ellipticals. On the other hand the observational consequence of this model is

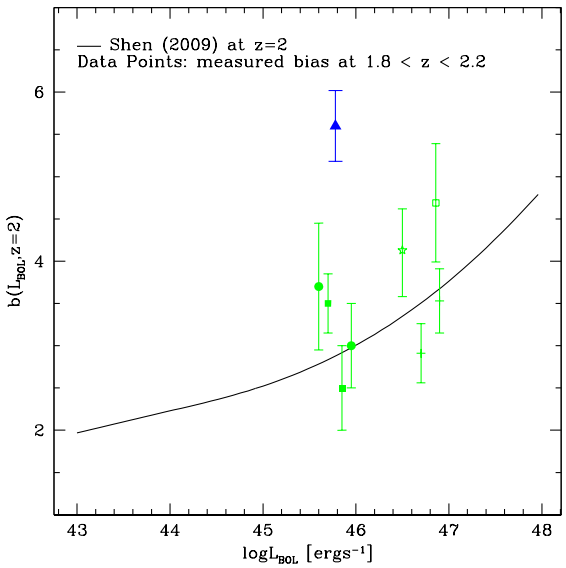

Figure 4.14 : Predicted bias as a function of luminosity, computed according to Shen (2009) fixing $z = 2$, compared to previous bias estimates at $1.8 < z < 2.2$, for optically selected BL AGN and for XMM-COSMOS BL AGN. Points are measurements from Croom et al. (2005, green-crosses), Porciani & Norberg (2006, green-star), Shen et al. (2009, green-open square), da Ângela et al. (2008, green-circles), Myers et al. (2007, green-squares) and my results (blue triangle). For ease of comparison, all luminosities are converted to bolometric luminosities using the corrections from Hopkins et al. (2007).

that the duty cycle is very low, typically of the order of $10^{-2} - 10^{-3}$. This result implies a small fraction of giant ellipticals observed in an AGN phase, too low compared to the observed 10% of X-ray AGN residing in massive galaxies. In the AGN evolutionary model described in Hickox et al. (2009), optically bright quasars are hosted by ongoing disk galaxy mergers and immediately precede an optically faint X-ray AGN phase, which evolves into an early-type galaxy. Following this evolutionary sequence, NL/X-obs AGN should be triggered in the first initial phase of vigorous star formation and obscured accretion which supports the scheme of NL AGN inhabiting halos with low typical masses $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 12.5$. An X-ray AGN phase immediately follows the quasar phase. Since DM halos grow and accumulate mass over time, X-ray AGN reside in more massive DM halos with typical mass $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 13 - 13.5$. This model predicts that X-ray AGN reside in more massive halos than QSO, but assumes a decline of the BH accretion rate from its peak in the quasar phase to $\dot{M} \lesssim 10^{-2}\dot{M}_{Edd}$ or lower, which is in

disagreement with *the high Eddington ratios* found for XMM-COSMOS BL AGN (Merloni et al., 2010; Trump et al., 2011).

A plausible scenario requires that high-luminosity quasars ($L_{BOL} > 10^{46}$ erg s^{-1}) are triggered by external processes such as major mergers between gas-rich galaxies with masses of the order of $M_* \sim 10^{10} M_\odot$. Instead for BL AGN with $L_{BOL} \sim 2 \times 10^{45}$ erg s^{-1}, internal mechanisms such as tidal disruptions or disk instabilities in massive galaxies ($M_* \sim 10^{11} M_\odot$) might play a dominant role.

The morphology of the AGN hosts galaxies provides an important clue into the mechanism that triggers their current AGN activity. It was observed that many AGN are not fueled by major mergers and only a small fraction of AGN are associated with morphologically disturbed galaxies. Cisternas et al. (2011) analysed a sample of X-ray selected AGN host galaxies and a matched control sample of inactive galaxies in the COSMOS field. They found that mergers and interactions involving AGN hosts are not dominant and occur no more frequently than for inactive galaxies. Over 55% of the studied AGN sample which is characterized by $L_{BOL} \sim 10^{45}$ erg s^{-1} and by mass of the host galaxies $M_* \gtrsim 10^{10} M_\odot$ are hosted by disk-dominated galaxies. This high disk fraction means that the lack of disturbed morphologies observed among the AGN hosts can not simply be due to a time lag between merger activity and X-ray visibility and suggests that secular fueling mechanisms can be high efficient.

It was also suggested by Georgakakis et al. (2009) that bar instabilities and minor interactions are more efficient in producing luminous AGN at $z \lesssim 1$ and not only Seyfert galaxies and low-luminosity AGN as the Hopkins & Hernquist (2006); Hopkins & Henquist (2009) models predict. Moreover several works on the AGN host galaxies (Dunlop et al., 2003; Grogin et al., 2005; Pierce et al., 2007; Gabor et al., 2009; Reichard et al., 2009; Tal et al., 2009) show that the morphologies of the AGN host galaxies do not present a preference for merging systems.

At the redshift of our interest, recent findings of Schlegel et al. (2001) and Rosario et al. (2011), who examined a smaller sample of AGN at $z \sim 2$ in the ERS-II region of the GOODS-South field, inferred that late-type morphologies are prevalent among the AGN hosts. The role that major galaxy mergers play in triggering AGN activity at $1.5 < z < 2.5$ was also studied in the CDF-S. Kocevski et al. (2011) found that X-ray selected AGN at $z \sim 2$ do not exhibit a significant excess of distorted morphologies while a large fraction reside in late-type galaxies. They also suggest that these late-type galaxies are fueled by the stochastic accretion of cold gas, possibly triggered by a disk instability or minor interaction.

I want to stress that these results by no means infer that mergers make no role in the AGN triggering. On the contrary, high luminosity AGN and probably a fraction of moderate luminosity AGN in the XMM-COSMOS AGN sample might be fuelled by mergers. In fact, given the complexity of AGN triggering, a proper selection of an AGN sub-sample, using for instance the luminosity, can help to

test a particular model boosting the fraction of AGN host galaxies associated with morphologically disturbed galaxies.

The work presented in this Chapter might extend the statement that for moderate luminosity X-ray selected BL AGN secular processes might play a much larger role than major mergers up to $z \sim 2.2$, compared to the previous $z \lesssim 1$, even during the epoch of peak merger-driven accretion.

4.10 Conclusions

In this Chapter I have studied the redshift evolution of the bias factor of 593 XMM-COSMOS AGN with spectroscopic redshifts $z < 4$, extracted from the 0.5-2 keV X-ray image of the 2deg^2 XMM-COSMOS field. A new method to estimate the bias factor and the associated DM halo mass, which accounts for the growth of the structures over time has been presented. Key results can be summarized as follows:

1. I estimated the AGN bias factor $b_{S01} = 2.71 \pm 0.14$ at $\overline{z} = 1.21$ which corresponds to a mass of AGN hosting DMHs equal to $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.10$ ± 0.10.

2. I split the AGN sample in broad optical emission lines AGN (BL) and AGN without optical broad emission lines (NL) and for each of them I considered a subset with $\overline{z} = 0.6$ and I found that BL and NL AGN present $b_{S01} = 1.95 \pm 0.17$ and $b_{S01} = 1.62 \pm 0.15$, which correspond to masses equal to $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.27 \pm 0.10$ and 12.97 ± 0.07, respectively.

3. I selected in the hard band a sample of X-ray unobscured and X-ray obscured AGN according to the column density and I found that X-ray unobscured (X-ray obscured) AGN inhabit DM halos with the same mass compared to BL (NL) AGN with $\log M_0[\mathrm{h}^{-1}M_\odot] = 13.30 \pm 0.10$ ($\log M_0[\mathrm{h}^{-1}M_\odot] = 12.97$ ± 0.08).

4. I found evidence of a redshift evolution of the bias factors for the different AGN subsets, corresponding to a constant DM halo mass threshold which differs for each sample. XMM-COSMOS AGN are hosted by DM halos with mass $\log M_0[\mathrm{h}^{-1}M_\odot]$=13.12 ± 0.07 constant at all $z < 2$, BL/X-ray unobscured AGN reside in halos with mass $\log M_0$ $[\mathrm{h}^{-1}M_\odot] = 13.28 \pm 0.07$ for $\overline{z} < 2.25$ while XMM-COSMOS NL/X-ray obscured AGN inhabit less massive halos $\log M_0$ $[\mathrm{h}^{-1}M_\odot] = 13.00 \pm 0.06$, constant at all $\overline{z} < 1.5$.

5. The observed bias evolution for XMM-COSMOS BL and NL AGN suggests that the AGN activity is a mass triggered phenomenon and that different

AGN evolutionary phases are associated with just the DM halo mass, irrespective of the redshift z.

6. The bias evolution of X-ray selected BL/X-ray unobscured AGN corresponds to halo masses in the range $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 13-13.5$ typical of poor galaxy groups at all redshifts. Optically selected BL AGN instead reside in lower density environment with constant halo masses in the range $\log M_h[\mathrm{h}^{-1}M_\odot] \sim 12.5-13$ at all redshifts. This indicates that X-ray and optically selected AGN do not inhabit the same DM halos.

7. The theoretical models which assume a quasar phase triggered by major mergers can not reproduce the high bias factors and DM halo masses found for X-ray selected BL AGN up to $z \sim 2.2$. This result might extend the statement that, for moderate luminosity X-ray selected BL AGN, secular processes might play a much larger role than major mergers up to $z \sim 2.2$, compared to the previous $z \lesssim 1$, even during the epoch of peak merger-driven accretion.

Chapter 5

Halo Occupation Distribution of AGN in the COSMOS field

5.1 Introduction

In Chapter 4, I showed that the clustering analysis can powerfully test theoretical model predictions and address which physical processes trigger AGN activity. The current picture for optical quasars, based on survey such as 2QZ and SDSS (Croom et al., 2005; Porciani & Norberg, 2006; Shen et al., 2009; Ross et al., 2009) is that bright AGN live in DMHs with logM_h [h$^{-1}M_\odot$] $\sim$ 12.5 up to z$\sim$3, i.e. halo masses similar to group scales, where the combination of low velocity dispersion and moderate galaxy space density yields to the highest probability of a close encounter (Hopkins et al., 2008; McIntosh et al., 2009). Models of major mergers between gas-rich galaxies appear to naturally produce many observed properties of quasars, such as the shape and the evolution of the quasar luminosity function and the large-scale quasar clustering as a function of luminosity and redshift (Hopkins et al., 2007, 2008; Shen, 2009; Shankar et al., 2009, 2010; Bonoli et al., 2009), supporting the scenario in which major mergers dominate the bright quasar populations.

On the contrary, I showed that the majority of the results on the clustering of X-ray selected AGN, suggest a picture where moderate-luminosity AGN live in massive DMHs ($12.5 < logM_h$ [h$^{-1}M_\odot$] < 13.5) up to $z \sim 2$ (Gilli et al., 2005; Yang et al., 2006; Gilli et al., 2009; Hickox et al., 2009; Coil et al., 2009; Krumpe et al., 2010; Cappelluti et al., 2010; Krumpe et al., 2010; Allevato et al., 2011; Miyaji et al., 2011), i.e. X-ray selected AGN samples appear to cluster more strongly than bright quasars (see §4.9). The reason for this is not completely clear but the results presented in the previous Chapter suggest that these large bias and DMH masses could suggest a different AGN triggering mechanism with respect to bright quasars characterized by galaxy merger-induced fueling. On the other hand,

several studies on the morphology of the AGN host galaxies have demonstrated that major mergers of galaxies are not likely to be the single dominant mechanism responsible for triggering AGN activity at low ($z \sim 1$) (Georgakakis et al., 2007; Silverman et al., 2009; Georgakakis et al., 2009; Dunlop et al., 2003; Grogin et al., 2005; Pierce et al., 2007; Gabor et al., 2009; Reichard et al., 2009; Tal et al., 2009; Cisternas et al., 2011; Silverman et al., 2011) and high redshift ($z \sim 2$) (Rosario et al., 2011; Kocevski et al., 2011; Schlegel et al., 2001).

In this Chapter I present how the clustering of AGN can be greatly enhanced through the Halo Occupation Distribution (HOD) framework, which describes the physical relation between AGN and DMHs at the level of individual halos. In this framework, the virialised DMHs with typical overdensities of Δ_{200} (defined w.r.t. the mean density) are described in terms of the probability $P(N, M)$ of a halo of given mass M of having N galaxies. A simple way to model the complicated shape of $N(M)$ is by assuming the existence of two separate galaxy populations within halos, central and satellite galaxies. This method has been used extensively to interpret galaxy CFs (Hamana et al., 2004; Tinker et al., 2005; Phleps et al., 2006; Zheng et al., 2007; Zehavi et al., 2010) to constrain how various galaxy samples are distributed among DMH as well as whether these galaxies occupy the centers of the DMHs or are satellite galaxies (Kravtsov et al., 2004; Zheng et al., 2005; Zehavi et al., 2005; Richardson et al., 2012). These two populations can be modeled with an HOD described by a step function above a halo mass limit for central galaxies, and a power-law for satellite galaxies (see Chapter 3 for more details).

Similarly, the problem of discussing the abundance and spatial distribution of AGN can be reduced to studying how they populate their host halos. In fact the observed departure of the AGN CF from a power law on small scales (1-2 h^{-1}Mpc) can be physically interpreted in the language of the halo model, as the transition between two scales - from small scales lying within the DMHs (1-halo term) to those larger than the halo (2-halo term). The 1-halo term constrains the HOD of satellite AGN and gives us the average profile of pairs of AGN in groups and clusters of galaxies. The 2-halo term reflects the large-scale AGN bias driven by the typical mass of the hosting halos.

Due to the low number density of AGN, there have been few results in the literature studying the AGN correlation function using HOD modeling. Previous works of Padmanabhan et al. (2009) at $z < 0.6$ and Shen et al. (2010) at $z = 3 - 4$ on QSO using the HOD modeling found that >25% and $\geqslant$10% of their QSOs, respectively, are satellites. Miyaji et al. (2011) described for the first time the shape of the HOD of X-ray selected AGN. By using the cross correlation function of ROSAT-RASS AGN with SDSS galaxies, they modelled the mean AGN occupation of DMHs suggesting that the satellite AGN fraction increases slow (or may even decrease) with M_h, in contrast with the satellite HOD of luminosity-limited samples of galaxies. Cosmological hydrodynamic simulations have been performed

in Chatterjee et al. (2012) to study the mean occupation function of low-luminosity AGN as function of redshift and luminosity. They used a softened step function for the central component plus a rolling-off power-law for the satellite component with $\alpha = 0.3 - 1.4$ depending on the redshift and AGN luminosity. Their results suggest a strong evolution of the AGN occupancy in the redshift range $z = 1 - 3$ estimated at three different luminosities $L_{BOL} \geq 10^{38}, 10^{40}, 10^{42}$ erg s^{-1}. Richardson et al. (2012) modelled the HOD of SDSS quasars at $z \sim 1.4$ following this parametrization. They found that the satellite occupation becomes significant at mass $\sim 10^{14}$ h$^{-1}M_{\odot}$, i.e. only the most massive halos host multiple quasars at this redshift and only a small satellite fraction ($f_{sat} = 7.4 \pm 1.3 \times 10^{-4}$) of SDSS quasars is required to fit the clustering signal at small scales. Moreover, they measured that the quasar HOD steepens considerably going from z=1.4 to 3.2 over halo mass scales $10^{13} - 10^{14}$ h$^{-1}M_{\odot}$ and that the characteristic halo mass increases with z for central quasars.

Despite the diverse methods for studying the HOD, counting the number of AGN within galaxy groups can constrain quite directly the average AGN number within a halo as a function of halo mass. The total mass of galaxy groups can be estimated via gravitational lensing and the distribution of AGN within halos can be investigated in groups by means of the distribution of the AGN host galaxies. Separating the contribution to the occupation of halos from AGN in satellite or central galaxies can advance our understanding of the co-evolution AGN/galaxy and is strictly related to the mechanism of AGN activation.

On the other hand, the cross-correlation (CCF) of AGN with galaxy groups provides additional information about how galaxies and BH co-evolve in dense environments. In fact the physical processes that drive galaxy evolution, such as the available cold gas to fuel star formation and the BH growth, are substantially different in groups and clusters compared to the field. Many studies over the past decade have presented an evidence that AGN at $z \sim 1$ are more frequently found in groups compared to galaxies (Georgakakis et al. 2008, Arnold et al. 2009). X-ray observations reveal that a significant fraction of high-z clusters show overdensities of AGNs in their outskirts (Henry et al. 1991, Cappi et al. 2001, Ruderman et al. 2005, Cappelluti et al. 2005).

In this Chapter, I present the first direct measurement of the mean halo occupation of X-ray AGN and projected cross-correlation function of AGN with galaxy groups, using a sample of X-ray selected AGN and galaxy groups in the COSMOS field at $z \leq 1$, from XMM and *Chandra* data. I use a ΛCDM cosmology with $\Omega_M = 0.28$, $\Omega_\Lambda = 0.72$, $\Omega_b = 0.045$, $\sigma_8 = 0.8$. For comparison with previous measurements I refer to correlation lengths and distances in units of h^{-1}Mpc comoving, where $H_0 = 100$ h km s^{-1} Mpc^{-1}. AGN luminosities and galaxy groups masses are calculated using $h = 0.72$. The contents of this Chapter have been accepted for publication in The Astrophysical Journal in August 2012, Allevato et al. 2012.

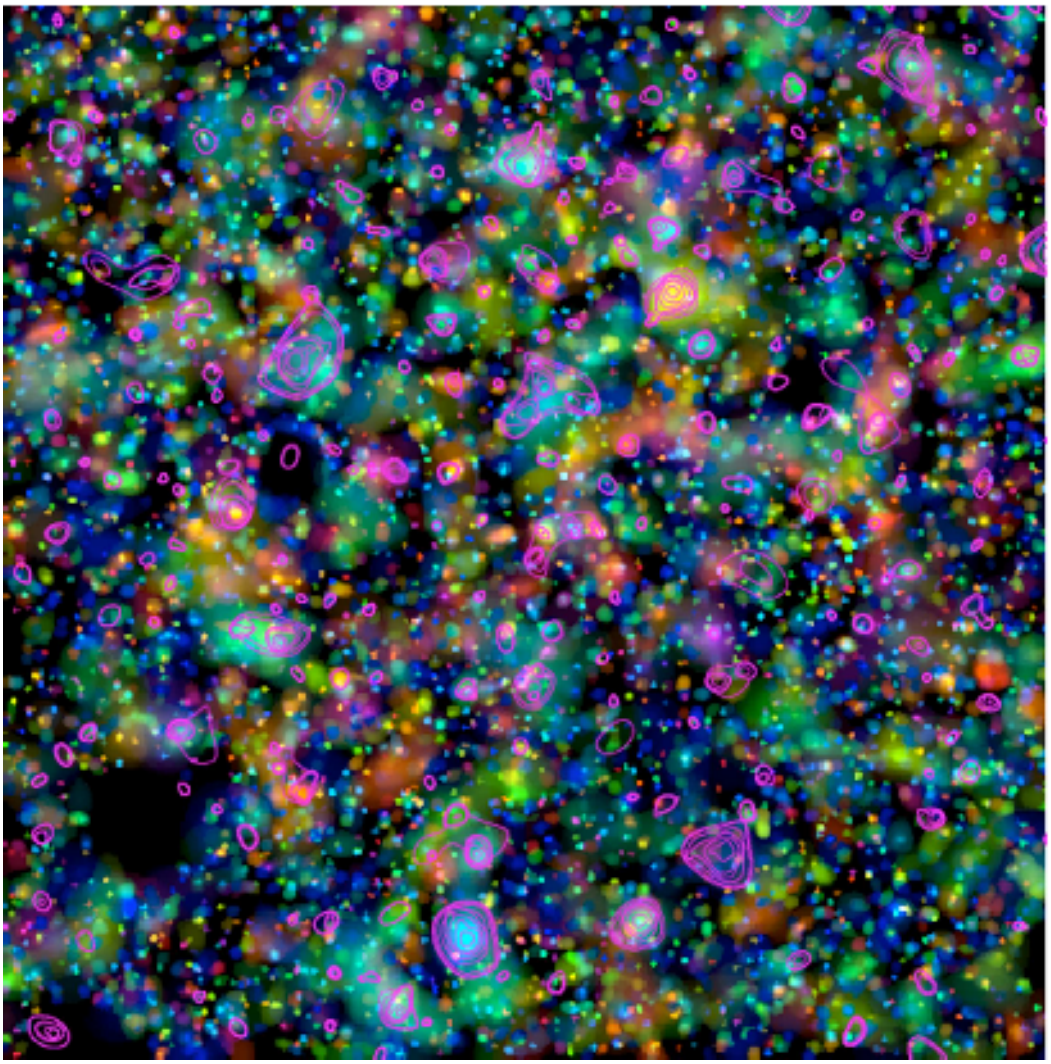

Figure 5.1 : The wavelet reconstruction of the early-type galaxy concentrations searched in the photometric redshift catalogue is color coded according to the average redshift: blue, 0.2; cyan, 0.4; green, 0.6; yellow, 0.8; and red, 1.0. The magenta contours outline the area of the X-ray emission associated with ∼300 extended source. The image is 1.5° on a side. [Credit: A. Finoguenov]

5.2 X-ray Galaxy Groups in the COSMOS Field

The galaxy groups catalog used in this Chapter is constructed from both XMM and *Chandra* data. It contains 270 objects in the redshift range $0 < z < 1.9$, spanning the rest-frame 0.1-2.4 keV luminosity range $10^{41} \leq L_X \leq 10^{44}$ s^{-1}erg (see Fig. 5.1). The general data reduction process is described in Finoguenov et al. (2007) and details regarding improvements and modification to the initial catalog are given in Leauthaud et al. (2010) and George et al. (2011). The identification of the groups has been done using the red sequence technique, and spectroscopic identification of groups has been achieved through zCOSMOS-BRIGHT program (Lilly et al., 2009), targeted follow-up using IMACS/Magellan and FORS2/VLT (George et al., 2011), Gemini GMOS-S (Balogh et al., 2011) as well as through secondary targets on Keck runs by COSMOS collaboration. At $z \leq 1$ the spectroscopic completeness has achieved 90% (A. Finoguenov 2012, in preparation).

As described in Leauthaud et al. (2010), the total X-ray fluxes have been obtained from the measured fluxes by assuming a beta profile and by removing the flux that is due to embedded AGN point sources. Some of the faint Chandra AGN could not be removed from the group fluxes, with their contribution to the total flux being $< 10\%$. The rest-frame luminosities have been computed in Finoguenov et al. (2007) and Leauthaud et al. (2010) from the total flux following $L_{0.1-2.4keV} = 4\pi d_L^2 K(z,T) C_\beta(z,T) F_d$, where $K(z,T)$ is the K-correction and $C_\beta(z,T)$ is an iterative correction factor, while to estimate the temperature of each group they used the $L_X - T$ relation of Markevitch (1998).

A quality flag (hereafter 'XFLAG') is assigned to the reliability of the optical counterpart, with flags 1 and 2 indicating a secure association, and higher flags indicating potential problems due to projections with other sources or bad photometry due to bright stars in the foreground. In detail, XFLAG=1,2 are assigned to groups with a confident spectroscopic association, while systems with only the red sequence identification have *XFLAG=3* (see Leauthaud et al., 2010, for more details).

The line-of-sight position of the group is assigned to be the centroid of the X-ray emission (the accuracy of the determination of the X-ray center is higher for *XFLAG=1*). If the X-ray centroid is not precise enough to be used directly, the Most Massive Central Galaxy (MMCG) located near the peak of the X-ray emission has been used to trace the center of the DM halos of groups (see Leauthaud et al., 2010; George et al., 2011, for more details). Group masses M_{200} are assigned from an empirical mass-luminosity relation, described in Leauthaud et al. (2010),

$$\begin{aligned} log_{10}(M_{200,c}) = p_0 - log_{10}E(z) + log_{10}(M_0) \\ +p_1[log_{10}(L_x/E(z)) - log_{10}(L_0)] \end{aligned} \tag{5.1}$$

where M_{200} is the mass within the radius containing the density of matter 200 times the critical density, in units of $M_\odot$. $\{p_0, p_1\} = \{0.729538, 0.561657\}$ are the fitting

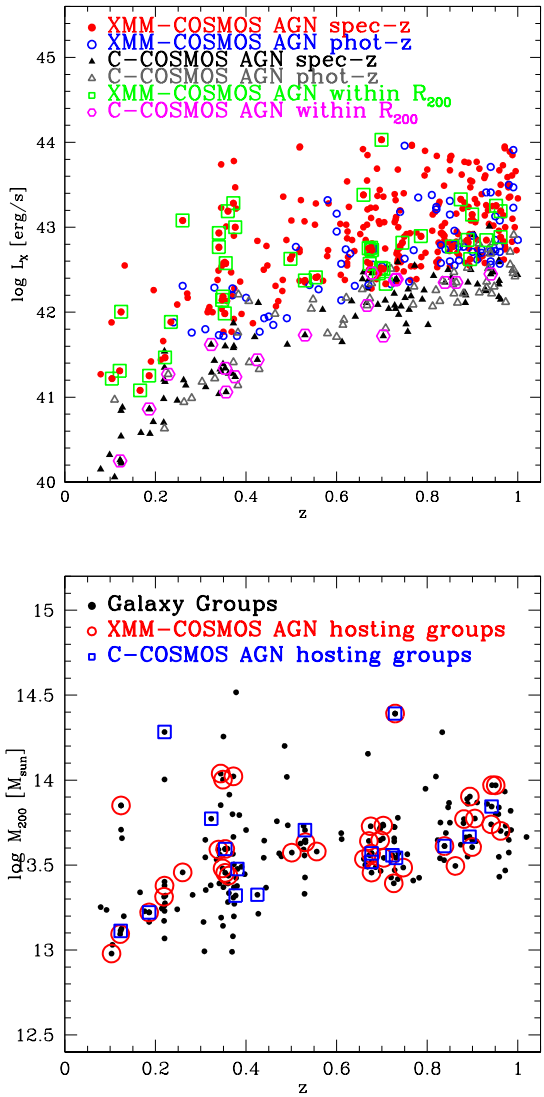

Figure 5.2 *Upper panel*: X-ray rest-frame soft luminosity L_X as a function of redshift for XMM-AGN (circles) and C-COSMOS selected AGN (triangle) with know spectroscopic (filled) and or photometric (open) redshifts. The open green squares indicate XMM-AGN within galaxy groups, while the open magenta hexagons represent 17 additional C-COSMOS selected AGN in galaxy groups. *Lower panel*: Galaxy group masses M_{200} as a function of redshift for the whole galaxy group sample (black circles) while open circles (squares) indicate XMM-AGN (C-COSMOS AGN) hosting groups. The mass estimates are defined with respect to 200 times the mean density, in units of $M_{\odot}$.

parameters, $\{logM_0, logL_0\} = \{13, 42.5\}$ are the calibration parameters and $E(z)$ is the correction for redshift evolution of scaling relations, which has been shown in Leauthaud et al. (2010) to reproduce well the $L_X - M$ relation of COSMOS groups.

In order to be consistent in comparing these mass values with the ones obtained studying the clustering properties of groups, I accounted for the difference between the mass defined with respect to 200 times the critical density and with respect to 200 times the mean density (hereafter M_{200} refers to masses obtained using the definition with respect to the mean density). In fact, the absolute bias which I am going to derive from the DM correlation function is based on the shape of the DM mass function defined with respect to the mean density. Starting from the Navarro-Frenk-White (NFW) profile with a concentration parameter $c = 5$, I derived the relation between the two mass definitions, $M_{200,m} = M_{200,c} \times \Omega(z)^{-0.134}$.

In particular, I decide to focus on galaxy groups with $z \leq 1$, where the spectroscopic completeness has achieved 90% (A. Finoguenov 2012, in preparation). Moreover, I only make use of objects with identification flag $XFLAG \leq 3$ which removes problematic identification cases, obtaining a catalog of 189 X-ray galaxy groups over 1.64 deg^2 with a rest-frame 0.1-2.4 keV luminosity range of $41.3 < logL_X[s^{-1}\mathrm{erg}] < 44.1$, and mass range of $13 < logM_{200}[M_\odot] < 14.5$ (see Fig. 5.2, *Lower Panel*).

5.3 X-ray AGN in the COSMOS Field

In §4.2 I introduced the XMM-COSMOS catalog of X-ray sources with optical/near-infrared identification, multiwavelength properties and redshift information, described in detail in Brusa et al. (2010). Starting from this catalog, I restricted the analysis to a sample of X-ray AGN (I removed normal galaxies and ambiguous sources) detected in the soft band which guarantees the largest sample of X-ray AGN in the COSMOS field, compared to AGN observed in the hard or ultra-hard band. Moreover, the selection in only one band allows a more simple treatment of the AGN X-ray luminosity function used to correct the AGN HOD (§5.5) and of the XMM and C-COSMOS sensitivity maps used to generate the AGN random catalog (§5.8).

Specifically, I selected a sample of 280 and 83 soft XMM-COSMOS AGN with spectroscopic and photometric redshift (Salvato et al., 2011) $z \leq 1$, respectively. The use of photometric redshifts for group membership assignment has been successfully demonstrated in George et al. (2011). Note that 184/363 sources are also *Chandra* detected AGN. In order to test if the AGN halo occupation significantly changes including sources from the C-COSMOS catalog that are only *Chandra* detected (hereafter C-COSMOS AGN), I included in the analysis a sample of 107 and 61 AGN, detected in the soft band, with known spectroscopic or photometric

Table 5.1 Properties of the Group and AGN Samples

(1) Sample	(2) N	(3) $\langle z \rangle^{a}$	(4) $\langle L_X \rangle^{b}$
XMM-AGN	363	0.66	$10^{42.8}$
C-COSMOS AGN	168	0.56	$10^{41.8}$
XMM-AGN in R_{200}	41	0.55	$10^{42.6}$
C-COSMOS AGN in R_{200}	17	0.53	$10^{41.7}$
XMM-AGN in the field[d]	253	0.67 [e]	$10^{42.8}$
XMM+C-COSMOS AGN in R_{200}	58	0.55	$10^{42.3}$
- Satellites	36	0.56	$10^{42.2}$
- Centrals	22	0.50	$10^{42.4}$
			$\langle M_{200} \rangle^{c}$
Galaxy Groups	189	0.56	$10^{13.60}$
XMM+C-COSMOS AGN host groups	52	0.55	$10^{13.62}$

[a]Median redshift of the sample.
[b]In units of h_{70}^2 erg s^{-1}
[c]Mass defined respect to 200 times the mean density known with a 20% error, in units of $M_\odot$.
[d]AGN sample used in estimating the CCF.
[e]Only spectroscopic redshifts.

redshifts $z \leq 1$, respectively. The rest-frame soft X-ray luminosity as a function of redshift is shown in Fig. 5.2 (*Upper Panel*) for XMM-COSMOS AGN (circles) and C-COSMOS AGN (triangles).

5.4 AGN in galaxy groups

I define group members as AGN located within $< 3\sigma$ and $< R_{200}$ from the group centers, where σ is the group line-of-sight velocity dispersion and R_{200} is the virial radius of a group within which the mean density is 200 times the mean density of the Universe at the group redshift.

In this analysis I used the sample of 363 XMM-COSMOS AGN with $z \leq 1$ described in §5.3 and I found 41 sources (35/41 are also *Chandra* detected) in galaxy groups with median $\langle z \rangle = 0.55$ and median $\langle L_X \rangle = 10^{42.6}$ erg s^{-1}. When I include in the analysis the sample of 168 C-COSMOS AGN, I found 17 additional AGN in galaxy groups, with known spectroscopic or photometric $z \leq 1$. As expected C-COSMOS AGN have lower soft fluxes respect to XMM-COSMOS AGN at any redshift, with a median $L_X = 10^{41.7}$ erg s^{-1}. In particular I found 2 galaxy groups at $z \sim 0.1$ and 2 groups at $z \sim 0.7 - 0.8$ with 2 AGN per halo and 1 group

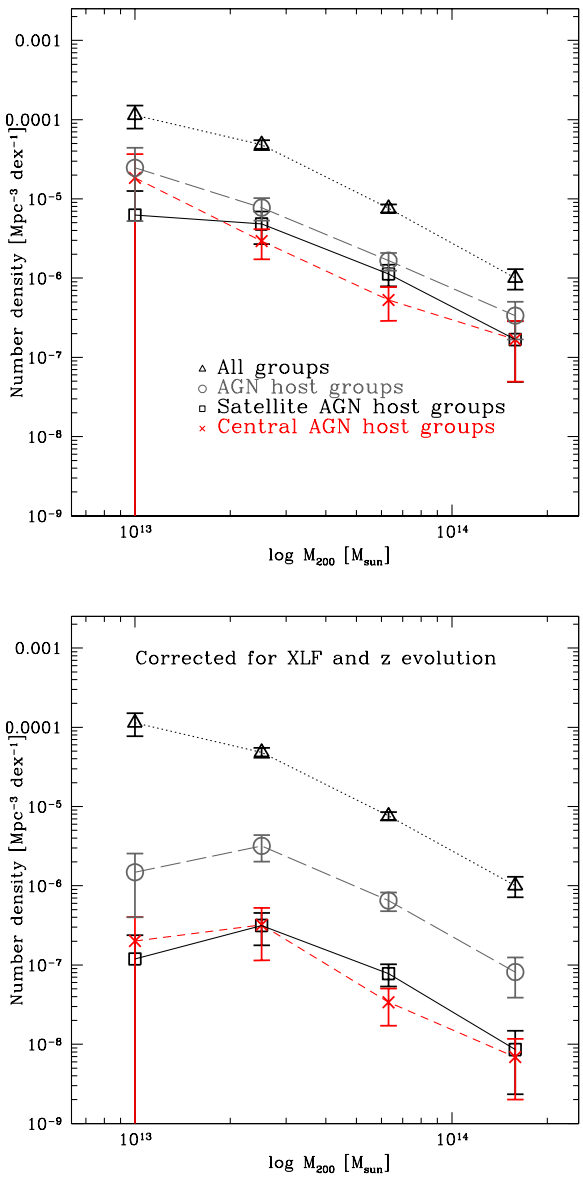

Figure 5.3 *Upper Panel*: Mass function of X-ray galaxy groups (triangles) and AGN host groups (circles). The black squares (red crosses) show the mass function of groups hosting an AGN in satellite (central) galaxies. *Lower Panel*: Same as left panel when correcting for the AGN soft XLF and the redshift evolution.

with 3 AGN at $z \sim 0.35$. All the properties of AGN and galaxy group samples are summarized in Table 5.1, while Table 5.2 shows the catalog of 58 XMM + C-COSMOS AGN in galaxy groups. For the sources with known spectroscopic redshifts (47/58) we know the classification in BL or non-BL AGN as described in Brusa et al. (2010) and Civano et al. (2012) for XMM and C-COSMOS AGN, respectively. In detail, I found 43 non-BL and 4 BL XMM + C-COSMOS AGN. Fig. 5.2 (*Upper Panel*) shows the rest-frame soft X-ray luminosity as a function of redshift for the subsamples of XMM (green squares) and C-COSMOS selected AGN (magenta hexagon) within R_{200}.

Since the errors on the AGN and galaxy groups photometric redshifts are much larger than the galaxy groups velocity dispersion, I checked the AGN-group associations by cross-matching the sample of AGN in groups with a galaxy membership catalog (Leauthaud et al., 2007; George et al., 2011). I verified that AGN classified as group members based on our method, have host galaxies associated with the same galaxy groups. I divided the sample of 58 AGN in groups in two subsets, according to their association with BCGs. In detail I found that 22/58 (16/41) AGN are in central galaxies, while 36/58 (25/41) are in satellites.

5.5 Halo Mass Function and AGN HOD

Fig. 5.3 (*Upper Panel*) shows the mass function of all X-ray galaxy groups and those marked by AGN presence, showing separately the contributions of groups hosting an AGN in central or satellite galaxies. I calculated the mass function by using the standard $1/V_{max}$ method (Schmidt 1968) and I counted twice galaxy groups with 2 AGN. Hence, in the ith mass bin, the comoving space density (n_i) and its corresponding error (σ_i) are computed by (see Bondi et al. 2008):

$$n_i = \sum_j \frac{1}{V_{max}^j} \qquad \sigma_i = \sum_j \sqrt{\left(\frac{1}{V_{max}^j}\right)^2} \tag{5.2}$$

In estimating the average number of AGN occupying a halo of mass M_{200}, some major effects need to be taken into consideration. The sample of AGN in R_{200} is a flux-limited sample and brighter AGN are detected at higher redshift. Similarly, one observes galaxy groups with small halo mass only at low redshift. However the relatively small number of COSMOS AGN in galaxy groups does not allow to select a volume complete subsample by using a cut in luminosity at $\log L_X$[erg s^{-1}] = 42.4. I have therefore made a correction for the effect of changes in the AGN density as a function of redshift and limiting luminosity, by using the AGN X-ray luminosity function (XLF). It has been shown in different works that the Luminosity Dependent Density Evolution (LDDE) model for the XLF provides the best framework that describes the evolutionary properties of AGN, both in

the soft (Miyaji et al., 2000; Hasinger et al., 2005) and hard X-rays (Ueda et al., 2003; La Franca et al., 2005). In this work I modelled the AGN soft XLF with the LDDE XLF described in Ebrero et al. (2009), because they modelled the soft XLF including in the analysis both type 1 and type 2 AGN, but I verified that using different best-fit parameters of the global XLF expression, the resulting mean AGN occupation stays within the error bars.

Then for each AGN redshift I defined two weights: w to correct for the fact that we are including in the analysis AGN with $\log L_X$[erg s^{-1}]$<$ 42.4 and w_0 to correct for this effect plus the redshift evolution of the AGN density:

$$w(z) = \frac{\int_{42.4}^{\infty} \phi(z, L_X) dL_X}{\int_{L_{lim}(z)}^{\infty} \phi(z, L_X) dL_X} \tag{5.3}$$

$$w_0(z) = \frac{\int_{42.4}^{\infty} \phi(z = 0, L_X) dL_X}{\int_{L_{lim}(z)}^{\infty} \phi(z, L_X) dL_X} \tag{5.4}$$

$\phi(z, L_X)$ is the soft XLF proposed in Ebrero et al. (2009) and L_{lim} depends on the survey flux limit and is a function of redshift. In detail, following Equations 9, 11 and 12 and Table 2 in Ebrero et al. (2009), I described the shape of the present-day luminosity function with slopes $\gamma_1 = 0.72 \pm 0.02$ and $\gamma_2 = 2.04 \pm 0.04$, $\log L_0 = 43.65 \pm 0.05$ h$_{70}$ erg s^{-1} which is the value of the luminosity where the change of slope occurs and normalization $A = 3.76 \pm 0.38 \times 10^{-6}$ h$_{70}^3$Mpc^{-3}. I estimated the evolution factor assuming $p_1 = 3.38 \pm 0.09$, $p_2 = -1.5$, $z_c = 1.42$, $\log L_a = 44.6$ h$_{70}$ erg s^{-1} and $\alpha = 0.100 \pm 0.005$.

Based on Eq. 5.3 and 5.4, the comoving space density n_i of AGN hosting groups corrected for the XLF is given by:

$$n_i = \sum_j \frac{w^j(z, L_X)}{V_{max,j}} \tag{5.5}$$

while the n_i corrected for both the XLF and the z evolution is estimated by using:

$$n_i = \sum_j \frac{w_0^j(z = 0, L_X)}{V_{max,j}} \tag{5.6}$$

Fig. 5.3 (*Lower Panel*) shows the mass function when all these effects are corrected.

The ratio between the mass function of X-ray groups hosting AGN within R_{200} by that of all X-ray galaxy groups generates the mean AGN HOD, which describes the occupation of DM halos by AGN. Fig. 5.4 (*Upper Panel*) shows the observed average number of XMM-COSMOS AGN in a halo of given mass as a function of $\log M_{200}$ (grey triangles) and the average number after correcting for the XLF (green open circles) and for both the XLF and the z evolution (black filled circles).

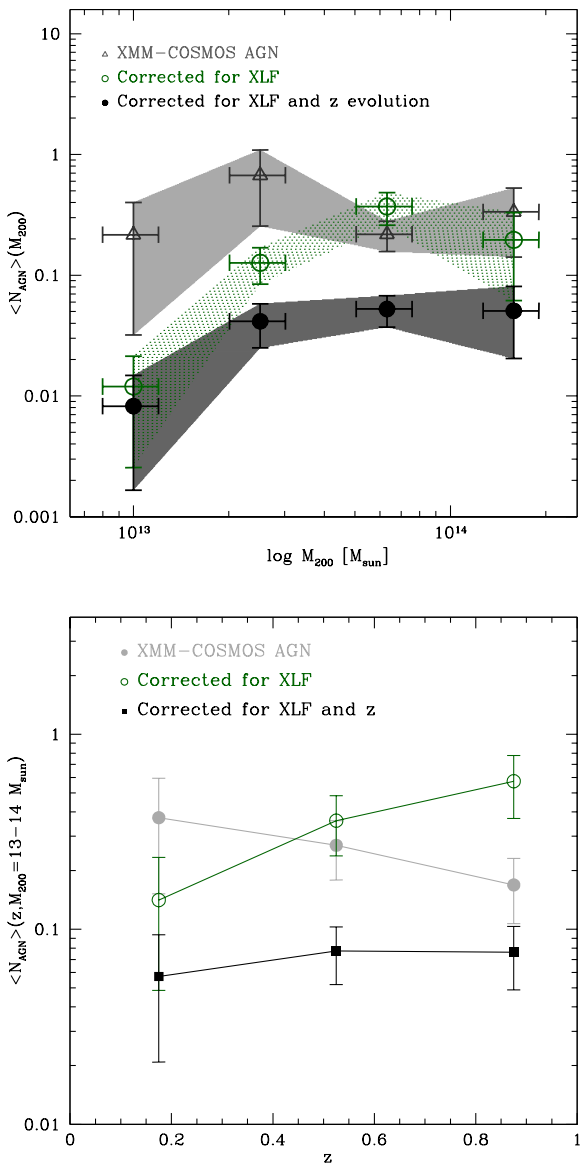

Figure 5.4 *Upper Panel*: Observed occupation of galaxy groups by XMM-COSMOS AGN as a function of the halo mass (open grey triangles) and after correcting for the soft XLF (green open circles) and for both the XLF and the AGN redshift evolution. *Lower Panel*: Mean AGN occupation as a function of redshift for AGN hosting groups with $\log M_{200}[M_{\odot}]$=13-14 (colors with same meaning as upper panel.)

The redshift evolution of the AGN fraction is shown in Fig. 5.3 (*Lower Panel*). The grey circles represent the mean AGN occupation as a function of redshift for halos with $M_{200} = 13 - 14 M_{\odot}$ while the green circles and the black squares show the AGN fraction when we correct for the XLF only and for both the XLF and the z evolution effects, respectively. When we correct for the XLF, the mean occupation increases with z since in the AGN luminosity range of our interest, the AGN density increases with z, while the redshift correction removes this trend producing a constant AGN fraction.

I fitted the total AGN HOD assuming the model with a rolling-off power-law:

$$\langle N_{AGN} \rangle (M_h) = f_a \left(\frac{M_h}{M_1} \right)^{\alpha} exp \left(\frac{M_{cut}}{M_h} \right) \tag{5.7}$$

where f_a is the normalization, M_1 is the halo mass at which the number of central AGN is equal to that of satellite AGN, M_{cut} is a cut-off mass scale. With our data alone, I cannot make meaningful constraints on M_1, therefore I fixed $\log M_1 = 13.8 M_{\odot}$ following the results of Miyaji et al. (2011). I verified that the result does not change for $\log M_1 = 13 - 14.2 M_{\odot}$, due to the fact that the fraction of AGN among central and satellite galaxies are comparable in this mass range. For the modelling of the rolling-off power-law, I assumed $\log M_{cut} = 13.4 M_{\odot}$ which corresponds to the mass below which our data points decay exponentially.

I obtained constraints in the (α, f_a) - space and I found as best fit parameters of the mean AGN HOD, $\alpha = 0.06(-0.22; +0.45)$ and $f_a = 0.06(0.04; 0.08)$, where the 68% confidence interval for a combined two parameter fit ($\Delta\chi^2 = 2.3$) is given in the brackets. Fig. 5.5 (*Upper Panel*) shows the mean occupation of XMM-COSMOS AGN with the best fit parameters (solid black line), and 1σ confidence interval (shaded grey region), compared to the mean AGN HOD including C-COSMOS AGN in the analysis. The solid red line corresponds to the best fit model for the mean occupation of XMM+C-COSMOS AGN, with $\alpha = 0.06(-0.22, 0.36)$ and $f_a = 0.05(0.04, 0.06)$, while the shaded region is the 1σ confidence interval (see Fig. 5.5, *Lower Panel*).

Moreover by dividing the mass function of satellite (central) AGN host groups by that of all X-ray galaxy groups, I provide the fraction of AGN among satellite (central) galaxies as a function of halo mass (see Fig. 5.6). I model the mean AGN occupation function in halos by decomposing it into the central and satellite contribution $\langle N_{AGN} \rangle (M_h) = \langle N_{cen} \rangle (M_h) + \langle N_{sat} \rangle (M_h)$:

$$\langle N_{cen} \rangle (M_h) = f_a \times erf \left(\frac{log M_h - log M_{min}}{\sigma_{logM}} \right) \tag{5.8}$$

$$\langle N_{sat} \rangle (M_h) = f_a \times \left(\frac{M_h}{M_1} \right)^{\alpha_s} exp(-M_{cut}/M_h) \tag{5.9}$$

where the central AGN occupation follows a softened step function and the satellite occupation, a rolling-off power law (e.g. Kravtsov et a. 2004, Zheng et al. 2005,

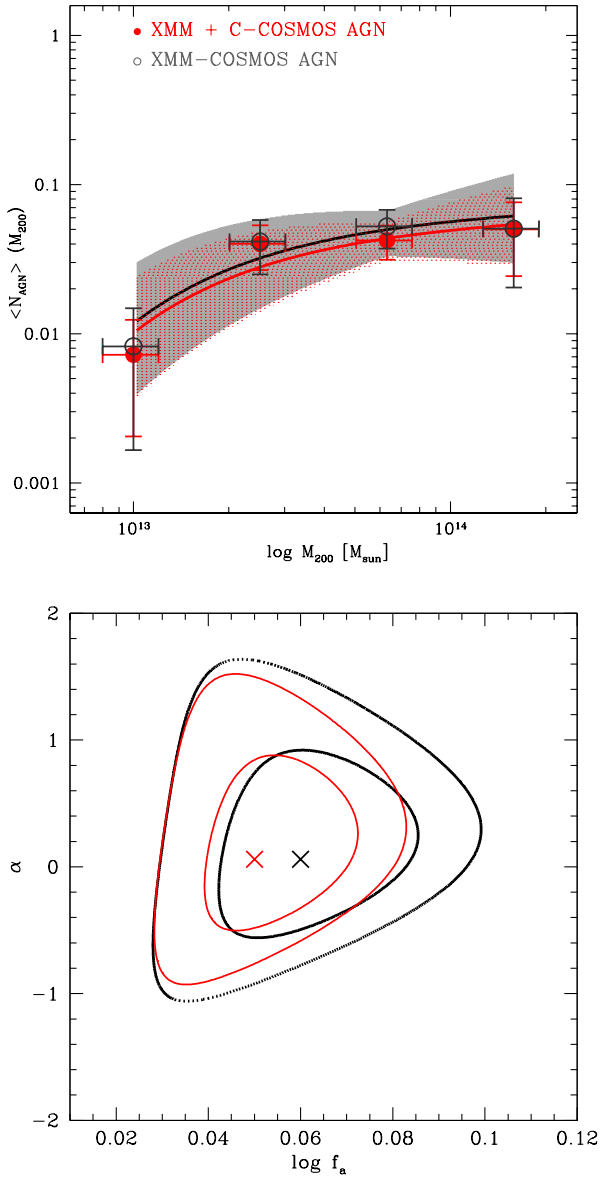

Figure 5.5 *Upper Panel*: Occupation of galaxy groups by 41 XMM-COSMOS AGN (black open circles) and 58 XMM + C-COSMOS AGN (red filled circles) as a function of the halo mass, when correcting for the XLF and for the redshift evolution of the AGN density. The fit assuming a rolling-off power-law dependence of the HOD is shown as the solid black lines (best fit) and shaded regions (1σ confidence interval, $\Delta\chi^2 = 2.3$). *Lower Panel*: the confidence contours of the power-law best-fit parameters α and f_a, for XMM-COSMOS AGN (black) and for XMM + C-COSMOS AGN (red) in galaxy groups. The contours mark the 68.3% and 95.4% confidence levels (respectively corresponding to $\Delta\chi^2 = 2.3$ and 6.17).

Zehavi et al. 2005, Tinker et al. 2005, Conroy et al. 2006, Chatterjee et al. 2011, Richardson et al. 2012). In this formalism there are four free parameters f_a, M_{min}, σ_{logM} and α_s, where M_{min} is the minimum mass where the occupation of central AGN is zero.

As shown in Fig. 5.6, the HOD of central AGN is described by a softened step function with $logM_{min}[M_\odot] = 12.75(12.10, 12.95)$ and $\sigma_{logM} = 1.46(0.4, 4.0)$ where the errors are the 1σ confidence intervals estimated by using the χ^2 minimization technique with 2 free parameters. On the other hand, the satellite AGN HOD suggests a picture in which the average number of satellite AGN increases with M_h with $\alpha_s = 0.22(-0.07, 0.63)$ and $f_a = 0.034(0.022, 0.046)$, slower than the satellite HOD of samples of galaxies ($\propto M_h^{\alpha_s=1-1.2}$).

Miyaji et al. (2011) first used a sample of ROSAT-RASS AGN and SDSS galaxies at $z \sim 0.3$ to study the occupancy of X-ray AGN in DM halos. They investigate three models: the first (model a) assumes that AGN only reside in satellite galaxies while the others explore the effects of centrals (model b and c). In particular, in model b, the HOD of central AGN is constant and the satellite HOD has a power-law form at halo masses above M_{min}. Model c has the same form except that only less massive DM halos contain central AGN. My results clearly show that only a model which includes AGN in both satellite and central galaxies at any halo mass $M > M_{min}$ can reproduce the observed AGN HOD.

In agreement with my results, they found that the upper limit of the power-law index of the satellite HOD is below unity, with $\alpha_s \leq 0.95$. In detail, their model b with best fit parameters $\alpha_s = 0.55$, $\log M_{min}[\mathrm{h}^{-1}M_\odot] = 12.5$ and $f_a = 2.9 \times 10^{-2}$ is in perfect agreement with my results, showing that the luminosity and redshift evolution of the mean AGN HOD is not strong in the luminosity and redshift ranges of our interest. In fact, I used an AGN sample with $\langle L_X \rangle = 10^{42.3}$ erg s^{-1}, while the results in Miyaji et al. (2011) provide the mean AGN HOD at $\langle z \rangle \sim 0.3$ for more luminous AGN with $\langle L \rangle_{0.1-2.4} \sim 10^{44.2}$ erg s^{-1} without any correction for the z evolution. Then the two models suggest a similar positive α_s range, but negative values of the slope are not rejected in their model ($\Delta\chi^2 < 2.3$). Note that while in my work the AGN fraction is a free parameter, they constrained it by normalizing the AGN HOD to the observed AGN number density. The two models suggest a similar α_s range, while the lower fraction of AGN might suggest a luminosity-dependent effect.

An assumption in the corrections applied by using Eq. 5.3 and 5.4, is that the shape of the XLF remains the same between AGN in groups and the general AGN population at $\log L_\mathrm{x}[\mathrm{erg\ s^{-1}}] < 42.4$. We still do not have sufficient observational basis to estimate the effects of possible difference in the XLF shapes between group environment and the field. At higher luminosities, Krumpe et al. (2010; 2012) found that AGN at $\log\langle L_X \rangle[\mathrm{erg\ s^{-1}}] \approx 44.6$ have higher bias values than those at ≈ 43.9, corresponding to the mean host halo mass of $\log\langle M_\mathrm{h} \rangle[M_\odot] = 13.0$ and 13.3

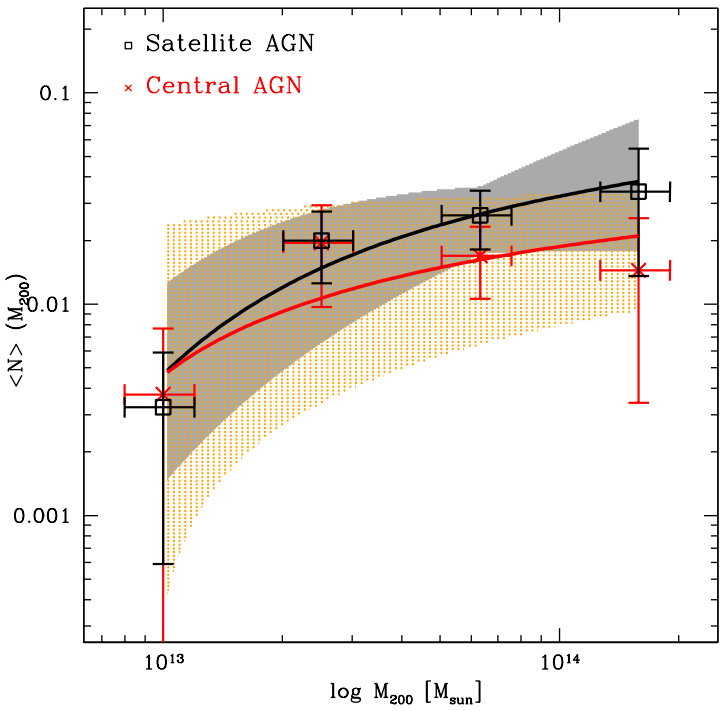

Figure 5.6 : Occupation of galaxy groups by satellite AGN (black squares) and central AGN (red crosses) as a function of the halo mass when correcting for the redshift evolution of the AGN density and for the soft XLF. The fit assuming a rolling-off power-law dependence is shown as solid black line (best fit) and dashed grey region (1σ confidence interval) for AGN among satellite galaxies. The HOD for central AGN has been modelled with a softened step function (solid red line) where the shaded orange region marks the 68.3% confidence level.

respectively (Miyaji et al. 2011). Viewing this trend from the $M_{\rm h}$ dependence of the XLF shape, the XLF in high $M_{\rm h}$ environments have XLF more biased towards higher $L_{\rm X}$. If the positive correlations between $L_{\rm X}$ and the host halo mass $M_{\rm h}$ extends to luminosities in the range $41.5 < \log L_X$ [erg s^{-1}] < 42.4, i.e. the range of $\log L_{\rm lim}(z)$ (see Fig. 1, *Upper Panel*), the correction factors using the overall XLF in Eq. 3 and 4 are overestimated at higher $L_{\rm X}$ and underestimated at lower $L_{\rm X}$. A quantitative assessment of this effect needs an estimate of the bivariate X-ray luminosity-host DMH mass function, which is far from being available. I comment that, if the higher L_X AGN preferentially occupy $\log M_{\rm h} > 13$ DMHs than the field, my estimate of $\alpha_{\rm s}$ should be corrected to a lower value.

5.6 Two-point Statistics

In §2.4 I presented the ACF as commonly used estimator of the spatial distribution of AGN in the Universe. Moreover I showed that a known effect when measuring pair separations is that the peculiar velocities combined with the Hubble flow may cause a biased estimate of the distance when using the spectroscopic redshift. To avoid this effect it is usually computed the projected ACF (Davis & Peebles, 1983) defined by:

$$w_p(r_p) = 2\int_0^{\pi_{max}} \xi(r_p, \pi) d\pi \tag{5.10}$$

where r_p is the distance component perpendicular and π parallel to the line of sight (Fisher et al., 1994).

An accurate estimate of the distribution function of the random samples is crucial in order to obtain a reliable estimate of $\xi(r_p, \pi)$. Several observational biases must be taken into account when generating a random sample of AGN in a X-ray flux limited survey. In particular, in order to reproduce the selection function of the survey, one has to carefully reproduce the space and flux distributions of AGN, since the sensitivity in X-ray surveys is not homogeneous on the detector and therefore on the sky.

I created an AGN random sample where each simulated source is placed at a random position in the sky, with flux randomly extracted from the catalog of real source fluxes. The simulated source is kept in the random sample if its flux is above the sensitivity map value at that position (Miyaji et al., 2007; Cappelluti et al., 2009). The corresponding redshift for a random object is assigned based on the smoothed redshift distribution of the AGN sample.

Similarly, an unclustered catalog of galaxy groups mimicking the selection function of the survey must be employed to quantify the degree to which the groups preferentially locate themselves in one another's neighborhood. A random group

catalog has been generated[1] by calculating at each area of a given sensitivity, the probability of observing a group of a given mass and redshift. Then Monte-Carlo simulations of the group positional assignment have been used, finally producing a catalog of hundred thousand objects. Also the X-ray surface brightness sensitivity map is non-uniform in depth and consequently the probability of detecting groups of a particular mass is variable with redshift; in particular the minimum mass below which a group will be detected is an increasing function of z.

5.7 Galaxy Groups ACF

I measured the projected ACF of galaxy groups in the range $r_p = 0.1 - 40$ $\mathrm{h^{-1}Mpc}$ by using Eq. 5.10, with $\pi_{max} = 80$ $\mathrm{h^{-1}Mpc}$ (see Fig. 5.7). The errors have been estimated using bootstrap resampling of the data, which consists of computing the variance of $w_p(r_p)$ in N_{real} bootstrap realizations of the sample. Each realization is obtained by randomly selecting a subset of groups from the data sample allowing for repetitions.

In the halo model approach, the clustering signal can be modelled as the sum of two contributions of pairs from the same DM halo (1-halo term) and those from different DM halos (2-halo term) (see §3.6). In Fourier space, the 2-halo term can be explicitly written as (Seljak , 2000; Cooray & Sheth, 2002):

$$P_{2-h} \approx b^2 P_{lin}(k, z) \tag{5.11}$$

where $P_{lin}(k, z)$ is the linear power spectrum and b is the bias factor of the sample. Then the galaxy group two-point correlation function at large scales is given by:

$$w_{p,2-h}(r_p) = b^2_{group} \int \frac{k}{2\pi} P_{lin}(k) J_0(kr_p) dk \tag{5.12}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind. Following this model, the galaxy group bias defines the relation between the 2-halo term of DM and groups clustering signal:

$$b^2_{group,obs}(r_p) = \frac{w_{p,2-h}(r_p)}{w^{2-h}_{DM}(r_p, z = 0)} \tag{5.13}$$

According to this equation, I estimated an average bias factor in the range $r_p = 1 - 40$ $\mathrm{h^{-1}Mpc}$ equal to $b_{group,obs} = 2.20 \pm 0.12$ where the error corresponds to $\Delta\chi^2 = 1$ using a χ^2 minimization technique with 1 free parameter. Following the bias mass relation $b(M_h, z)$ described in van den Bosch (2002) and Sheth et al. (2001), this observed bias, with respect to the DM distribution at $z = 0$ corresponds to a typical mass $\log M_{typ}$ $[\mathrm{h^{-1}}M_\odot] = 13.65^{+0.07}_{-0.08}$.

[1]The generation of the galaxy group random catalog has been performed by A. Finoguenov

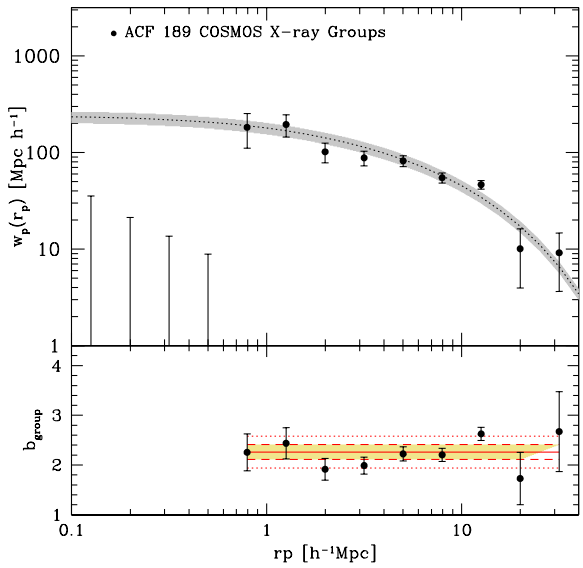

Figure 5.7 :*Upper Panel*: Projected auto-correlation function of galaxy groups with $z \leq 1$. Following the halo model approach, the clustering signal at large scale ($r_p > 1-2$ h^{-1}Mpc) is produced by separate DM halos (so called 2-halo term) and it can be expressed as $b^2_{group} w^{2-h}_{DM}(r_p, z=0)$ (dotted line), where the bias factor is defined in Equation 5.14. *Lower Panel*: Bias factor versus r_p. The solid line shows the best-fitting constant value. The shaded regions indicate the values of the bias for which $\Delta\chi^2 = 1$ (dashed line) and $\Delta\chi^2 = 4$ (dotted lines).

On the other hand the average bias of galaxy groups can be estimated starting from the known masses M_{200}. Since usually the bias factor is related to the average DM halo mass expressed in units of h$^{-1}M_\odot$, hereafter we refer to M_h to indicate the galaxy groups masses in these units. Since I know the mass estimates, I can predict the bias factor of this sample of galaxy groups. By accounting for the fact that the linear regime of the structure formation is verified only at large scales, I estimated the average bias of the sample, including only pairs which contribute to the clustering signal at $r_p = 1-40$ h^{-1}Mpc. Following the procedure described Chapter 4, I measured the average bias factor of the sample as:

$$b_{group}(M_h) = \sqrt{\frac{\sum_{i,j} b_{group,i} b_{group,j} D_i D_j}{N_{pairs}}} \tag{5.14}$$

where $b_{gg,i} b_{gg,j}$ is the bias factor of the group pair $i-j$ and N_{pairs} is the total number of group pairs in the range $r_p = 1-40$ h^{-1}Mpc. The D factor is defined by $D_1(z)/D_1(z=0)$, where $D_1(z)$ is the growth function (see eq. (10) in Eisenstein

& Hu (1998) and references therein) and takes into account that the amplitude of the DM 2-halo term decreases with increasing redshift.
By using this approach I obtained $b_{group} = 2.21^{+0.13}_{-0.14}$ where the errors have been estimated assuming a 20% error on the galaxy groups masses. This value is in perfect agreement with the bias obtained from the ACF, with $b_{group}/b_{group,obs} = 1.00 \pm 0.05$.

5.8 X-ray AGN-Galaxy Groups Cross-Correlation

Measurements of the cross-correlation function between AGN and groups use a version of the estimator proposed by Landy & Szalay (1993):

$$\xi = \frac{1}{R_g R_A}(D_g D_A - D_g R_A - D_A R_g + R_g R_A) \tag{5.15}$$

where each data sample, with pair counts D_i has an associated random catalog, with pair counts R_i normalized by its number density.

I estimated the cross-correlation function of 189 galaxy groups and a subset of 253 XMM-COSMOS AGN with known spectroscopic redshift ≤ 1, obtained excluding those within R_{200} (see Fig. 5.8). Following Eq. 5.13, the linear bias factor of the projected CCF of AGN with galaxy groups can be approximated by using the 2-halo term:

$$b^2_{CCF,obs}(r_p) = b_{AGN}(r_p) \times b_{group}(r_p) = \frac{w^{CCF}_{p,2-h}(r_p)}{w_{DM}(r_p, z=0)} \tag{5.16}$$

where b_{AGN} and b_{group} are the bias factor of AGN and galaxy groups, respectively and w_{DM} is the projected dark matter CF. Fig. 5.8 (*Lower Panel*) shows the linear bias b^2_{CCF} as a function of r_p over the scales $r_p \sim 1 - 40$ h^{-1}Mpc. I fitted the data points with a constant by using the χ^2 minimization technique and we found $b^2_{CCF,obs} = 3.90 \pm 0.28$. The shaded regions show the bias values for which $\Delta\chi^2 = 1$ and 4 (68% and 99% confidence levels for one parameter).

5.9 The bias factor in the HOD model

Based on the halo model approach, the AGN bias factor depends on the mean AGN HOD $\langle N_{AGN}\rangle(M_h)$:

$$b_{AGN}(z) = \frac{\int_{M_{min}}^{\infty} b_h(M_h, z)\langle N_{AGN}\rangle(M_h) n_h(M_h, z) dM_h}{\int_{M_{min}}^{\infty} \langle N_{AGN}\rangle(M_h) n_h(M_h, z) dM_h} \tag{5.17}$$

where M_{min} is the minimum mass below which the AGN HOD is zero, $n_h(M_h)$ and $b_h(M_h)$ are the halo mass function and the halo bias given by Sheth et al.

(2001). Note that the large-scale bias factor does not depend on the normalization $\log M_1[M_\odot] = 10^{13.8}$ and on f_a. Since I am estimating the CCF excluding from the XMM-COSMOS sample 41 AGN in galaxy groups, the clustering signal at large scale is due to AGN that live in halos with $M_{200} < 10^{13} M_\odot$ or with masses that we can not observe at a given redshift z. This implies that the AGN bias can be written as:

$$b_{AGN}(z) = \frac{\int_{M_{min}}^{M_x(z)} b_h(M_h, z)\langle N_{AGN}\rangle(M_h) n_h(M_h, z) dM_h}{\int_{M_{min}}^{M_x(z)} \langle N_{AGN}\rangle(M_h) n_h(M_h, z) dM_h} \tag{5.18}$$

where $M_x(z)$ is the minimum mass one can observe for a group at redshift z (at $z \sim 1$ one only detects luminous and then massive groups). Note that I am assuming the separability of the mass and redshift dependence of the $\langle N_{AGN}\rangle$, i.e. $\langle N_{AGN}\rangle(M_h, z) = \langle N_{AGN}\rangle(M_h, z = 0) \times \langle N_{AGN}\rangle(z)$.

Following this method, the bias factor b^2_{CCF} is defined by:

$$b^2_{CCF} = \frac{\sum_{i,j} b_{AGN,i} b_{group,j} D_i D_j}{N_{pairs}} \tag{5.19}$$

where the sum is over the pairs i, j contributing to the clustering signal at large scale and $b_{AGN}(z_i)$ is the AGN bias (Eq. 5.18) assuming a rolling-off power-law HOD based on my results. $b_{group,j}$ is the bias associated to the galaxy group masses and redshift following Sheth et al. (2001) and N_{pairs} is the total number of AGN-group pairs in the range $r_p = 1 - 40$ h^{-1}Mpc. I found $b^2_{CCF} = 3.97^{+0.10}_{-0.05}$, which is in perfect agreement with the observed bias factor defined in Eq. 5.16. The errors are due to the 1σ errors on the power-law index α, M_{min} and on b_{group}.

Similarly, I can define the average halo mass $\langle M\rangle$ corresponding to the observed CCF signal, i.e.:

$$\langle M\rangle = \frac{\sum_{i,j} b_{AGN,i} b_{group,j} D_i D_j M_h^{i,j}}{\sum_{i,j} b_{AGN,i} b_{group,j} D_i D_j} \tag{5.20}$$

where $M_h^i(z_i)$ is the average AGN host halo mass at redshift z_i:

$$M_h^i(z_i) = \frac{\int_{M_{min}}^{M_x(z)} \langle N_{AGN}\rangle(M_h) n_h(M_h, z_i) M_h dM_h}{\int_{M_{min}}^{M_x(z)} \langle N_{AGN}\rangle n_h(M_h, z_i) dM_h} \tag{5.21}$$

while $M_h^j(z_j)$ is the galaxy group mass, both in units of h$^{-1}M_\odot$. Fixing the values of α and M_{min}, I found the average mass of XMM-COSMOS AGN to be log $\langle M\rangle$ [h$^{-1}M_\odot$]= 13.20(13.10;13.25), where the errors correspond to the 68% confidence region in the $\log M_{min} - \alpha$ space.

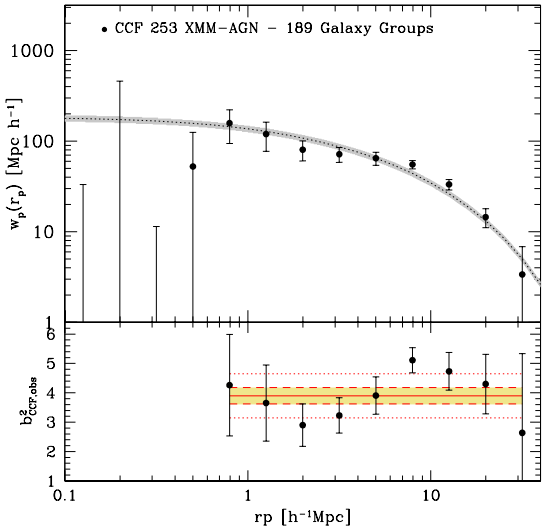

Figure 5.8 : *Upper Panel*: Projected cross-correlation function of 253 XMM-COSMOS AGN with $z_{spec} \leq 1$ and 189 galaxy groups, excluding from the analysis AGN that are within galaxy groups. The clustering signal at large scale $r_p > 1 - 2$ h^{-1}Mpc is due to AGN residing in different DM halos and is described by $b^2_{CCF} w^{2-h}_{DM}(r_p, z = 0)$ (dotted line), where the bias factor is given in Eq. 5.16. On the contrary the 1-halo term is zero since I removed AGN with R_{200}. *Lower Panel*: Bias factor versus r_p. The solid line shows the best-fitting constant value. The shaded regions indicate the values of the bias for which $\Delta\chi^2 = 1$ (dashed line) and $\Delta\chi^2 = 4$ (dotted lines).

5.10 Discussion

Using a sample of X-ray selected AGN and galaxy groups in the COSMOS field at $z \leq 1$, I performed the first direct measurement of the AGN HOD in the mass range $\log M_{200}[M_\odot]$ = 13-14.5, based on the mass function of galaxy groups hosting AGN. In contrast to previous works using the clustering signal of the sample, I directly counted the number of AGN within galaxy groups and I found 58 AGN in groups, associated to 22 central and 36 satellite galaxies. This allowed to put constrains on both the mean occupation of AGN among satellite and central galaxies as function of the halo mass, which provides information on the AGN triggering mechanism. Starikova et al. (2011) studied the halo occupation properties of AGN detected by the *Chandra* X-ray Observatory in the Bootes field over a redshift interval from z=0.17-3, showing that X-ray AGN are predominantly located at the centers of DMHs with $M_h > 4.1 \times 10^{12}$ h$^{-1} M_\odot$, with an upper limit of the satellite fraction of 0.1 ($\Delta\chi^2 < 2.3$). The central locations of the quasar host

galaxies are expected in major merger models because mergers of equally sized galaxies preferentially occur at the centers of DMHs (Hopkins et al., 2008). On the contrary Padmanabhan et al. (2009) observed the presence of the one-halo term in the cross-correlation function of optically selected $z < 0.6$ quasars and luminous red galaxies, and use this to conclude that a large fraction of the AGN is hosted by satellite galaxies.

My results show that the average number of AGN in satellite galaxies in the halo mass range $\log M_h$ $[M_\odot]$ = 13-14.5, and for an average AGN luminosity $\log\langle L_X\rangle$ erg s^{-1} = 42.3 might be comparable or even larger that the average number of AGN in central galaxies, i.e. X-ray AGN do not avoid satellite galaxies. A high fraction of AGN in satellite galaxies is expected in a picture where other phenomena like secular processes, might become dominant in the AGN activation. Milosavljevic et al. (2006), Hopkins et al. (2006) and Hopkins & Henquist (2009) showed that low-luminosity AGN could be triggered in more common non merger events, like stochastic encounters of the black holes and molecular clouds, tidal disruption or disk instability. This leads to the expectation of a characteristic transition to merger-induced fueling around the traditional quasar-Seyfert luminosity division.

Moreover I found the power-law slope, which defines the evolution of the mean satellite HOD with halo mass, to be $\alpha_s \sim 0 - 0.6$, suggesting a picture in which the average number of satellite AGN per halo increases with the halo mass. On the other hand, Miyaji et al. (2011) obtained $\alpha_s < 0.95$, but negative values of the slope are not rejected ($\Delta\chi^2 < 2.3$).

It is interesting to compare this result with HOD analyses of galaxies. Previous HOD analyses of galaxies found $\alpha_s \sim 1 - 1.2$ for a wide range of absolute magnitudes and redshifts at least up to $z \sim 1.2$ (Zehavi et al., 2005; Zheng et al., 2007; Zehavi et al., 2010), implying a simple proportionality between halo mass and satellite number, $\langle N_{sat}\rangle \propto M_h$. My results suggest that the mean HOD of satellite AGN might increase slower ($\alpha_s < 0.63$) with the halo mass respect to the linear proportion ($\alpha_s = 1$) in the satellite galaxy HOD, i.e. the AGN is not only triggered by the halo mass. On the contrary, a decreasing AGN fraction with the halo mass would be consistent with previous observations that the AGN fraction is smaller in clusters than in groups in the nearby universe.

In order to fully understand the growth history of SMBHs as well as the physical processes responsible for the AGN activity one needs to explore the AGN HOD at different redshifts, luminosities, and AGN types. The luminosity distribution of AGN that reside in halos of a given mass provides a tool to examine the distribution of halo mass for a given luminosity and study luminosity dependent clustering. While this formalism has been widely used in modelling galaxy clustering, it is still not applicable to AGN. In fact, it is also important to have larger numbers of AGN in galaxy groups which will enable stronger constraints on the shape of the satellite and central AGN HOD, hopefully as function of AGN luminosity and redshift.

5.11 Conclusions

I have performed the first direct measurement of the mean halo occupation distribution of X-ray AGN as function of halo mass, by directly counting the number of AGN within X-ray galaxy groups with masses $\log M_{200}[M_\odot]$ = 13-14.5, in the COSMOS field at $z \leq 1$. My findings are summarized as follows.

1. I identified 41 XMM-COSMOS AGN within galaxy groups, defined as AGN located within 3 times the group line-of-sight velocity dispersion and within R_{200} and 17 additional sources including in the analysis C-COSMOS only selected AGN.

2. I measured the mean AGN occupancy of galaxy groups as function of halo mass in the range $\log M_{200}[M_\odot]$ = 13- 14.5 and I modelled the data points with a rolling-off power-law with the best fit index $\alpha = 0.06(0.22, 0.36)$ and normalization parameter $f_a = 0.05(0.04, 0.06)$.

3. Using a galaxy membership catalog, I associated 22/58 and 36/58 AGN to central and satellites galaxies, respectively. I constrained that the mean AGN occupation function among central galaxies is described by a softened step function above $\log M_{min}[M_\odot]$=12.75(12.10, 12.95) while the satellite AGN HOD increases with the halo mass ($\alpha_s < 0.63$) slower than the satellite HOD of sample of galaxies ($\alpha_s = 1 - 1.2$).

4. I presented an estimate of the projected ACF of galaxy groups over the range r_p = 0.1-40 h^{-1}Mpc at $\langle z \rangle = 0.5$. I verified that the bias factor and the corresponding typical halo mass estimated from the observed galaxy group ACF, are in perfect agreement with the values b_{group} and $\langle M_h \rangle$ obtained by using the galaxy group mass estimates. In particular I found $b_{group} = 2.21^{+0.13}_{-0.14}$ and $log\langle M_h \rangle$ $[\mathrm{h}^{-1} M_\odot] = 13.61^{+0.09}_{-0.10}$.

Table 5.2. AGN in galaxy groups

ID Groups	Ra deg	Dec deg	z Groups	$\log M_{200}$[a] $[M_\odot]$	ID XMM[b]	AGN Ra deg	AGN Dec deg	ID $Chandra$[b]	z[c] AGN	$logL_X$[d]	Flag[e]
11	150.1898	1.65725	0.22	14.28	-99	150.19864	1.671924	12390	0.228	41.27	0
17	149.96413	1.68033	0.372	14.02	30744	149.96396	1.6805983	-99	0.372	43.28	1
19	150.37282	1.60944	0.103	12.98	2021	150.37256	1.6093965	1678	0.104	41.22	1
20	150.32494	1.60313	0.22	13.38	2186	150.33601	1.6012201	1671	0.234	41.88	0
35	150.2066	1.82327	0.53	13.70	-99	150.20702	1.823398	1292	0.529	41.73	1
39	149.82381	1.8252701	0.531	13.63	5502	149.81248	1.8238187	219	0.529	42.37	0
52	150.44704	1.8828501	0.671	13.64	30681	150.44731	1.8832886	690	0.670	42.75	1
69	150.42012	1.9708	0.862	13.49	5519	150.42473	1.9693011	1185	0.854	42.77	0
78	150.27461	1.98884	0.838	13.61	-99	150.27823	1.990302	685	0.838	42.35	1
78	150.27461	1.9888	0.838	13.61	206	150.27434	1.9884306	1100	0.673	42.57	0
79	150.44728	2.05392	0.323	13.77	-99	150.44765	2.053961	2605	0.323	41.62	1
87	150.51109	2.0269899	0.899	13.60	2387	150.51036	2.0293686	496	0.899	43.15	1
93	149.6692	2.0740	0.338	13.59	417	149.6692	2.073962	168	0.34	42.76	0
110	150.17979	2.1103699	0.361	13.43	6	150.17978	2.1101542	42	0.360	43.19	1
118	149.63463	2.1357	0.962	13.70	411	149.63828	2.1494889	323	0.952	43.25	0
124	150.05656	2.2085	0.186	13.22	302	150.0571	2.2063098	1297	0.186	41.25	0
124	150.05656	2.20854	0.186	13.22	-99	150.09554	2.2202671	1221	0.186	40.86	0
128	150.58435	2.1811	0.556	13.58	2855	150.57634	2.1812408	1634	0.554	42.41	0
127	150.44104	2.15873	0.377	13.32	-99	150.44202	2.1551671	2313	0.376	41.24	1
130	150.02382	2.2032	0.942	13.84	-99	150.02303	2.206567	2413	0.94	42.45	0
136	150.17493	2.2170	0.676	13.56	-99	150.17592	2.2156539	877	0.667	42.08	0
137	149.96271	2.2102399	0.425	13.32	-99	149.96216	2.2102211	12011	0.425	41.44	1
138	149.50874	2.2614	0.943	13.97	5539	149.50914	2.261276	-99	0.962	43.19	1
142	150.28798	2.2769	0.122	13.09	317	150.25331	2.2779887	3718	0.165	41.08	0
143	150.21454	2.2801	0.880	13.77	116	150.20604	2.2857771	22	0.874	43.33	0
149	150.41566	2.4302	0.124	13.85	45	150.33598	2.4335926	658	0.121	41.31	0
149	150.41566	2.4302001	0.124	13.85	73	150.416	2.4299896	634	0.124	42.00	1
161	149.95262	2.3418	0.941	13.74	430	149.95949	2.3560882	497	0.889	42.85	0
171	149.66328	2.2677	0.676	13.56	5411	149.66243	2.2693584	1646	0.676	42.76	0
173	150.05804	2.38045	0.347	13.48	241	150.05794	2.3805208	898	0.347	42.14	1
174	149.63988	2.3491	0.950	13.97	53937	149.64459	2.3587193	-99	0.962	42.88	0
175	150.24123	2.34835	0.723	13.55	-99	150.24109	2.3483839	977	0.703	41.72	0
186	150.21748	2.4003	0.905	13.77	53303	150.21141	2.4022212	975	0.894	42.62	0
194	149.69957	2.4028	0.354	13.45	135	149.70084	2.4025679	417	0.376	43.00	0

Table 5.2 (continued)

ID Groups	Ra deg	Dec deg	z Groups	$\log M_{200}$[a] $[M_\odot]$	ID XMM[b]	AGN Ra deg	AGN Dec deg	ID $Chandra$[b]	z[c] AGN	$logL_X$ [d]	Flag [e]
196	150.27898	2.4192	0.123	13.11	-99	150.27975	2.4222209	1315	0.122	40.25	0
216	150.06664	2.6474	0.696	13.64	5113	150.06633	2.642817	143	0.693	42.45	0
217	150.00713	2.4534299	0.731	13.54	-99	150.00188	2.4606521	1243	0.732	42.39	1
219	150.27148	2.5134399	0.704	13.54	158	150.27411	2.5117824	138	0.703	42.51	1
220	149.92343	2.5249	0.729	14.39	486	149.92007	2.5143571	562	0.698	42.52	0
220	149.92343	2.5249901	0.729	14.39	-99	149.91588	2.52139	1310	0.729	42.38	0
231	150.05421	2.58885	0.675	13.73	8	150.05383	2.5896702	142	0.699	44.03	1
234	150.15816	2.6082	0.893	13.66	-99	150.15796	2.6113441	727	0.863	42.35	0
237	150.11774	2.6842501	0.349	14.00	5118	150.11783	2.6840661	-99	0.349	42.18	1
259	149.65717	2.8195	0.703	13.73	60270	149.65953	2.8270493	-99	0.708	42.33	0
262	149.60007	2.8211	0.344	14.03	5320	149.63142	2.8174951	-99	0.34	42.93	0
275	149.83878	2.6750801	0.259	13.45	5112	149.83847	2.6750875	1608	0.259	43.08	1
277	150.00462	2.63275	0.677	13.45	5091	150.00452	2.6328416	616	0.678	42.73	1
289	150.11256	2.5560	0.501	13.57	142	150.10342	2.5504889	148	0.498	42.62	0
292	150.03307	2.5524	0.747	13.48	398	150.02638	2.5620575	991	0.745	42.82	0
296	149.55516	2.0020	0.894	13.90	10732	149.56131	2.0087938	3549	0.930	42.85	0
298	149.78191	2.1390	0.354	13.59	63	149.78223	2.1387713	313	0.355	42.57	0
298	149.78191	2.13906	0.354	13.59	392	149.79369	2.1256437	679	0.353	41.98	1
298	149.78191	2.13906	0.354	13.59	-99	149.77357	2.141633	1498	0.354	41.34	0
300	149.72893	2.2373	0.381	13.47	-99	149.74792	2.253087	2876	0.356	41.06	0
303	149.99364	2.2585399	0.660	13.53	19	149.99367	2.2585886	450	0.659	43.38	1
322	150.2254	2.26872	0.677	13.51	-99	150.228	2.2698331	24	0.678	42.46	0
324	150.02414	2.36050	0.726	13.39	254	150.03079	2.358371	533	0.786	42.89	0
333	150.0423	2.6949	0.219	13.31	5075	150.04155	2.6945302	623	0.221	41.46	1

[a]Mass defined respect to 200 times the mean density, with h =0.72.

[c]= -99 means NOT detected.

[c]Photometric or spectroscopic redshift.

[d]In units of $[h_{70}^{-2}$ erg $s^{-1}]$

[e]0: AGN among satellite galaxies; 1: AGN among central galaxies.

Summary

I this Thesis I investigated the relation between Active Galactic Nuclei (AGN) and the typical environment in which they preferentially reside, using: their spatial distribution in the Universe, which defines the connection with their hosting dark matter halos (DMHs); and the halo occupation distribution (HOD), which describes how AGN populate DMHs, at the level of individual halos.

The projected two-point correlation function has been estimated for X-ray selected AGN extracted from the 0.5-2 keV X-ray mosaic of the 2.13 deg^2 XMM-COSMOS survey and interpreted through the halo model. By using this approach I studied the redshift evolution of the amplitude of the clustering signal, which reflects the AGN bias and the typical mass of the AGN host halos. I found evidence of a redshift evolution of the AGN bias with a DMH mass consistent with being constant ($\sim 10^{13}$ h$^{-1}M_{\odot}$) up to z$\sim$ 2, i.e. X-ray AGN preferentially reside in dense environments typical of galaxy groups.

A difference has been observed in the redshift evolution of the AGN bias for broad-line (BL) and non-broad line (NL) AGN up to z<1.5 (BL AGN reside in more massive halos respect to NL AGN). I ascribe this difference to the fact that the two classes of AGN might correspond to different phases of the AGN evolution sequence.

Models of major mergers between gas-rich galaxies appear to naturally produce many observed properties of quasars, such as the shape and the evolution of the quasar luminosity function and the large-scale quasar clustering as a function of luminosity and redshift, supporting the scenario in which major mergers dominate the bright quasar populations. On the contrary, I showed that my results extend up to $z \sim 2$ the statement that, for moderate-luminosity X-ray AGN, secular processes such as tidal disruptions or disk instabilities play a much larger role than major mergers. This result is in line with several studies on the morphology of the AGN host galaxies showing that major mergers of galaxies are not likely to be the single dominant mechanism responsible for triggering the AGN activity at low ($z \sim 1$) and high redshift ($z \sim 2$).

I connected my finding of X-ray AGN in massive DMHs to a detailed study of the halo occupation of AGN in massive galaxy groups. Groups of galaxies host a wide diversity of galaxy populations and are therefore perfect labs to study the AGN evolution and provide additional information about how galaxies and black holes co-evolve in dense environments. The information of the distribution of AGN within galaxy groups can be translated in how AGN occupy DMHs with different

mass. The total mass of galaxy groups can be estimated via gravitational lensing and the distribution of AGN within halos can be investigated in groups by means of the distribution of AGN host galaxies.

A sample of X-ray selected AGN and galaxy groups in the COSMOS field at $z \leq 1$, constructed with XMM and *Chandra* data has been used. I presented the first direct measurement of the mean halo occupation of X-ray selected AGN based on the association of 41 XMM and additional 17 AGN from deeper *Chandra* dataset with member galaxies of 189 X-ray detected galaxy groups. Separating the contribution to the occupation of galaxy groups by AGN in satellite and central galaxies, I found the average number of AGN among central galaxies to be modelled by a softened step function at logM > logM_{min} $[M_{\odot}]$ = 12.75(12.10, 12.95) while the satellite AGN HOD increases with the halo mass, following a power-law with slope $\alpha < 0.6$.

In terms of AGN triggering mechanisms, these results show that the average number of AGN in satellite galaxies in the halo mass range log$M_h[M_{\odot}]$ = 13-14.5, might be comparable or even larger than the average number of AGN in central galaxies, i.e. X-ray AGN are not preferentially located in central galaxies. A high fraction of AGN in satellite galaxies is expected in a picture where other phenomena like secular processes, might become dominant in the AGN activation. It has been found that the satellite AGN HOD increases with the halo mass slower ($\alpha < 0.6$) than a simple linear proportion ($\alpha = 1$), i.e. the average number of AGN in satellite galaxies is not only triggered by the halo mass, as observed for satellite galaxies.

Acknowledgments

I would like to thank all the people that made this thesis possible.
First of all thanks to my supervisor Prof. Günther Hasinger, for your invaluable advices and scientific support. I felt continuously encouraged during these three years. Many thanks to Alexis Finoguenov, you encouraged and believed to me from the first time we met. Every time I entered your office I knew I could find not only a good scientist, but also a remarkable person. Thank you very much for your patient supervision, your teachings and your friendship. I hope to continue this friendly and stimulating scientific and personal collaboration.

This work would not be possible without the help of Nico Cappelluti, who with infinite patience thought me so many things about astronomy and science politics. These three years would not be the same without the precious support of Angela Bongiorno. Many thanks for your unconditional help, "training", encouragement. You are a good scientist and a good friend. A special thanks goes to Mara Salvato, in these three years I always could rely on you and your precious advices. Thank you very much to Prof. Stephan Paul, who with extraordinary patience, helpfulness and kindness made the exam possible.

In these three years I met so many wonderful persons which I would like to thank, Michele, Barbara, Hanna, Alessandro and Annalisa. Thank you very much for all the nice moments, coffee breaks, dinners and friendship. *Grazie per avermi aiutato a sopportare il freddo, il cibo tedesco e la lontananza da casa. Vi portero' sempre con me, ovunque mi porti la vita.* Thanks to my officemate Sotiria, for tolerating my mess and stress and to Alex, for the stimulating scientific discussions and for the several translations from German into English. Thanks to Arne and Rene for helping me in translating the abstract into German.

Very special thanks go to my parents. Many of the things I have done in my life would have not been possible without your help, support and encouragement. Thank you for giving me the possibility to pursue my dreams with unquestioning support and love. *Questa tesi, e' dedicata a voi, perche' siete sempre stati la mia guida, perche' vi ho sempre portato con me in ogni lungo viaggio, perche' avete rinunciato ad avermi vicino, anche nel momento del bisogno, pur di continuare a sostenere la vostra piccola scienziata.* Special thanks to Rosamaria, my sister. I love you so much. *So che questo lavoro mi portera' a vivere lontana da te, ma io ci saro' sempre nella tua vita. Sempre.* Many thanks also to Marta, Mariateresa, zio Tonino, zia Tanina, zio Dino e zia Maria Antonietta. You are my family as well. *Vi ho sempre immaginato in prima fila davanti ai miei talks, a fare il tifo per me. Un grazie particolare a zia Maria Antonietta, "la prima scienziata della casa".*

And last, but not least, special thanks to Giancarlo, for just being what you are, for the joy and the serenity you gave me. I would not be where I am without having met you. Thank you for sharing with me the most important and beautiful moments during these years and for always being present in the difficult ones.

Bibliography

Aarseth S. J., Fall S. M., 1980, ApJ, 236, 42

Akylas, A., Georgantopoulos, I., & Plionis, M. 2000, MNRAS, 318, 1036

Allevato, V., et al. 2011, ApJ, 736, 99

Arnold, T. J., Martini, P., Mulchaey, J. S., Berti, A., & Jeltema, T. E. 2009, ApJ, 707, 1691

Balogh, M. L., et al. 2011, MNRAS, 412, 2303

Barcons, X., & Fabian, A. C. 1988, MNRAS, 230, 189

Barcons, X., Carrera, F. J., Ceballos, M. T., & Mateos, S. 2001, X-ray Astronomy: Stellar Endpoints, AGN, and the Diffuse X-ray Background, 599, 3

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Bartelmann, M., Doran, M., & Wetterich, C. 2006), A&A, 454, 2736

Basilakos, S., Plionis, M., Georgakakis, A., et al. 2004, MNRAS, 351, 989

Basilakos, S., Plionis, M., Georgakakis, A., Georgantopoulos, I., et al. 2005, MNRAS, 356, 183

Bennert, N., Canalizo, G., Jungwiert, B. et al. 2008, ApJ, 677, 846

Bergamini, R., Londrillo, P., & Setti, G., 1967, Nuovo Cimento B Serie, 52, 495

Berlind, A. A., Weinberg, D. H., 2002, ApJ, 575, 587

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440460

Bonoli, S., Marulli, F., Springel, V., et al. 2009, MNRAS, 396, 423

Boyle, B. J., & Mo, H. J. 1993, MNRAS, 260, 925

Branchesi M., Gioia I. M., Fanti C., Fanti R., Cappelluti N., 2007, A&A, 462, 449

Brandt, W. N., & Hasinger, G. 2005, ARA&A , 43, 827

Brusa, M., Civano, F., Comastri, A., et al. 2010, ApJ, 716, 348

Bryan, G. L. & Norman, M. L. 1998, ApJ, 495, 80

Canalizo, G., Stockton, A., 2001, ApJ, 555, 719

Canalizo, G., Bennert, N., Jungwiert, B. et al. 2007, ApJ, 669, 801

Capak, P., Aussel, H., Ajiki, M., 2007, ApJ, 172, 99

Cappelluti, N., et al. 2005, ApJ, 430, 39

Cappelluti, N., et al. 2007, ApJ, 172, 341

Cappelluti, N., et al., 2009, A&A, 497, 635

Cappelluti, N., Aiello M., Burlon D., et al., 2010, ApJ, 716, 209

Cappelluti, N., Allevato, V., Finoguenov, A., Advances in Astronomy, vol. 2012, id. 853701

Cappi, M., et al., 2001, ApJ, 548, 624

Carrera, F. J., & Barcons, X., 1992, MNRAS, 257, 507

Cattaneo, A., Dekel, A.; Devriendt, J., Guiderdoni, B., Blaizot, J., 2006, MNRAS, 370, 1651

Chatterjee S., et al. 2012, MNRAS, 419, 2657

Ciotti, L., Ostriker, J. P., 1997, ApJ, 487, 105

Ciotti, L., Ostriker, J. P., 2001, ApJ, 551, 131

Cisterans, M., et al. 2011, ApJ, 726, 57

Civano, F., et al. 2012, ApJ, 752, 49

Coil A., et al. 2007, ApJ, 654, 115

Coil, A. L., Georgakakis, A., Newman, J. A., et al. 2009, ApJ 701 1484

Colberg, J. M. & Di Matteo, T., 2008, MNRAS, 387, 1163

Colless, M., et al. 2001, MNRAS, 328, 1039

Cooray, A., Sheth, R., 2002, PhR, 372, 1

Croom, Scott M., Boyle, B. J., Shanks, T., Smith, R. J., et al. 2005, MNRAS, 356, 415

da Ângela, J., Shanks, T., Croom, S. M., et al. 2008, MNRAS, 383, 565

Davis, M., Peebles, P. J. E., 1983, ApJ, 267, 465

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371394

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Dressler A., Thompson I. B., Shectman S. A., 1985, ApJ, 288, 481

Dunlop J. S., McLure R. J., Kukula, M. J., et al. 2003, MNRAS, 340, 1095

Ebrero, J., Mateos, S., Stewart, G. C., Carrera, F. J., & Watson, M. G., 2009, A&A, 500, 749

Elvis, M., Chandra-COSMOS Team, 2007, in Bullettin of the American Astronomical Society, Vol. 39, p.899

Elyiv, A., Clerc, N., Plionis, M., et al. 2012, A&A, 537A, 131

Eisenstein, Daniel J., Hu, Wayne., 1998, ApJ, 511, 5

Faltenbacher, A.; Finoguenov, A.; Drory, N., 2010, ApJ, 712, 484

Ferrarese L., Ford H., 2005, S. Sci. Rev., 116, 523

Finoguenov, A., et al. 2004, A&A., 419, 47

Finoguenov, A., et al. 2007, ApJS, 172, 182

Finoguenov, A., et al. 2009, ApJ, 704, 564

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A. & Huchra, J., 1994, MNRAS, 266, 50

Frenk, C. S., White, S. D. M., Efstathiou, G., & Davis, M. 1985, Nature, 317, 595597

Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. 1988, ApJ, 327, 507525

Fry, J. N., 1996, ApJ, 461, 65

Gabor, J. M., et al. 2009, ApJ, 691, 705

Gandhi, P., et al. 2006, A&A, 457, 393

Gebhardt K., et al., 2000, ApJ, 539, L13

Georgakakis, A., Nandra, K., Laird, E. S., et al. 2007, ApJ, 660, 15

Georgakakis, A., Coil, A. L., Laird, E. S., et al. 2009, MNRAS, 397, 623

Georgantopoulos, I., Stewart, G. C., Shanks, T., Griffiths, R. E., & Boyle, B. J., 1993, MNRAS, 262, 619

George, Matthew R., et al. 2011, ApJ, 742, 125

George, Matthew R., et al. 2012, submitted

Giacconi, R., et al., 1979, ApJL, 234, 1

Giacconi, R., Gursky, H., Paolini, F. R., Rossi, B. B., 1972, Physical Review Letters, vol. 9, Issue 11, 439, 443

Gilli, R., Daddi, E., Zamorani, G., et al. 2005, A&A, 430, 811

Gilli, R., et al. 2009, A&A, 494, 33

Gilmour R., Best P., Almaini O., 2009, MNRAS, 392, 1509

Gisler G. A., 1978, MNRAS, 183, 633

Gonzalez, A. H., Zaritsky, D., & Zabludoff, A. I. 2007, ApJ, 666, 147

Gott, III, J. R., Dickinson, M., & Melott, A. L. 1986, ApJ, 306, 341

Gregory, S. A., & Thompson, L. A. 1978, ApJ, 222, 784

Griffith R. L., Stern D., 2010, ApJ, 140, 533

Grogin, N. A., et al. 2005, ApJ, 627, 97

Gunn, J. E. & Gott, III, J. R. 1972, ApJ, 176, 1

Guyon, O., Sanders, D. B., Stockton, A., 2006, ApJS, 166, 89

Haggard D., Green P. J., Anderson S. F., Constantin A., Aldcroft T. L., Kim D.-W., Barkhouse W. A., 2010, ApJ, 723, 1447

Haines C. P. et al. 2009, ApJ, 704, 126

Haines C. P. et al. 2012, ApJ, 754, 97

—. 1993, ApJ, 417, 19

Hamana, T., Ouchi, M., Shimasaku, K., Kayo, I., Suto, Y., 2004, MNRAS, 347, 813

Hasinger, G., Burg, R., Giacconi, R., Hartne G., Schmidt, M., Trumper, J., & Zamorani, G., 1993, A&A, 275, 1

Hasinger, G., Hasinger, G., Schmidt, M., 2005, A&A, 441, 417

Hasinger, G., Cappelluti, N., Brunnen, H. et al. 2007, ApJS, 172, 29

Hasinger, G., et al. 2008, A&A, 490, 905

Hickox, R. C., Jones, C., Forman, W. R., 2009, ApJ, 696, 891

Hickox, R. C., et al. 2011, ApJ, 731, 117

Hills, J. G., 1975, Nature, 254, 295

Hopkins, P.F., Hernquist, L., Cox, T.J., Di Matteo, T., Robertson, B. Springel, V., 2006, ApJ, 163, 1

Hopkins, P.F., Hernquist, L., 2006, ApJS, 166, 1

Hopkins, P.F., et al. 2007, ApJ, 662, 110

Hopkins, P.F., Hernquist, L., Cox, T.J., Keres, D., 2008, ApJ, 175, 365

Hopkins, P. F., Cox, T. J., Keres, D., Hernquist, L. 2008, ApJS, 175, 390

Hopkins, P.F., Hernquist, L., 2009, ApJ, 694, 599

Hubble, E. 1934, ApJ, 79, 8

Hubble, E. P. 1926, ApJ, 64, 321

Hwang H. S., Park C., Elbaz, D., Choi Y.-Y., 2012, A&A, 538, 15

Jeltema, T. E., Mulchaey, J. S., Lubin, L. M., Fassnacht, C. D., 2007, ApJ, 658, 865

Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001, MNRAS, 321, 372384

Jogee S., 2006, in Physics of Active Galactic Nuclei at all Scales, ed. D. Alloin, R. Johnson & P. Lira (Lecture Notes in Physics Vol. 693; Berlin: Springer), p. 143

Kaiser, N. 1984, ApJL, 284, L9

—. 1987, MNRAS, 227, 1

Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576

Kauffmann G. et al., 2004, MNRAS, 353, 713

Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, ApJL, 535, L13

Kocevski, D. D., et al 2011, arXiv1109.2588K

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Kormendy J., Kennicutt R. C. Jr., 2004, ARA&A, 42, 603

Koss, M., Mushotzky, R., Veilleux, S., Winter, L., 2010, ApJ, 716, 125

Koulouridis, E., & Plionis, M. 2010, ApJL, 714, L181

Kravtsov, A. V., et al. 2004, ApJ, 609, 35

Krumpe, M., Miyaji, T., Coil, A. L. 2010, ApJ, 713, 558

Krumpe, M., Miyaji, T., Coil, A. L. & Aceves, H., 2012, ApJ, 746, 1

La Franca, F., et al. 2005, ApJ, 635, 864

Landy, S. D., & Szalay A. S., 1993, ApJ, 412, 64

Leauthaud, A., et al. 2007, ApJS, 172, 219

Leauthaud, A., Finoguenov, A., Kneib, J. P., et al 2010, ApJ, 709, 97

Ledlow, M. J., & Owen, F. N. 1996, AJ , 112, 9

Lehmer, B. D., et al. 2005, ApJS, 161, 21

Li, C., Kauffmann, G., Wang, L., et al., 2006, MNRAS, 373, 457

Lilly, S. J., Le Févre, O, Renzini, A, et al., 2007, ApJS, 172, 70

Lilly, S. J., Le Brun, V., Mayer, C., et al., 2009, ApJS, 184, 218

Lin Y.-T., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745

Mahajan S. M., et al. 2010, MNRAS, 404, 1745

Mahdavi, A., Böhringer, H., Geller, M. J., Ramella, M., 2000, ApJ, 534, 114

Markevitch, M., 1998, ApJ, 504, 27

McIntosh, D. H., Guo, Y., Mo, H. J., van den Bosch, F., & Yang, X. 2009, Bulletin of the American Astronomical Society, 41, #423.09

Martini, P., Sivakoff, G. R., Mulchaey, J. S., 2010, ApJ, 701, 66

Merloni, A., Bongiorno, A., Bolzonella, M., Brusa, M, et al., 2010, ApJ, 708, 137

Miller, C. J., Nichol, R. C., Gómez, P. L., Hopkins, A. M., Bernardi, M., 2003, ApJ, 597, 142

Milosavljevic, M., Merritt, D., & Ho, L. C. 2006, ApJ, 652, 120

Miyaji, T., Hasinger, G., Schmidt, M., 2000, A&A, 353, 25

Miyaji, T., Zamorani, G., Cappelluti, N., et al., 2007, ApJS, 172, 396

Miyaji, T., Krumpe, M., Coil, A. L., et al. 2011, ApJ, 726, 83

Mo H. J., & White, S. D. M. 1996, MNRAS, 282, 347

Molnar S. M., Hughes J. P., Donahue M., Marshall J., 2002, ApJ, 573, L91

Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613

Mountrichas, G., & Georgakakis, A., 2012, MNRAS, 420, 514

Myers, A. D., Brunner, R. J., Richards, G. T., et al. 2007, ApJ, 658, 99

Mullis, C. R., Henry, J. P., Gioia, I. M., et al., 2004, ApJ, 617, 192

Murray, S. S., et al. 2005, ApJS, 161, 1

Nandra, K., et al. 2005, MNRAS, 356, 568

Navarro, J. F., Frenk, C. S., White, S. D. M., 1995, ApJ, 275, 720740

Navarro, J. F., Frenk, C. S., White, S. D. M., 1996, ApJ, 462, 563

Navarro, J. F., Frenk, C. S., White, S. D. M., 1997, ApJ, 490, 493

Padmanabhan, N., White, M., Norberg, P., Porciani, C., 2009, MNRAS, 397, 1862

Peacock, J.A., Dodds, S.J. 1996, MNRAS, 280, L19.

Peacock, J. A., et al. 2001, Nature, 410, 169

Peebles, P. J. E. 1975, ApJ, 196, 647

—. 1980, The Large-Scale Structure of the Universe (Princeton, N.J., Princeton Univ. Press)

Phleps, S., Peacock, J. A., Meisenheimer, K., Wolf, C., 2006, A&A, 457, 145

Pierre, M., et al. 2007, MNRAS, 382, 279

Peebles P. J. E., 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565586

Pierce, C. M., et al. 2007, ApJ, 669, 19

Plionis, M., Rovilos, M., Basilakos, S., Georgantopoulos, I., & Bauer, F., 2008, ApJL, 674, L5

Porciani, C., Norberg, P., 2006, MNRAS, 371, 1824

Press, W. H., Schechter, P., 1974, ApJ, 187, 425

Puccetti, S., et al. 2006, A&A, 457, 501

Puccetti, S., et al. 2009, ApJS, 185, 586

Reichard, T. A., Heckmas, T. M., Rudnick, G., et al. 2009, ApJ, 691, 1005

Rees, M. J., 1988, Nature, 333, 523

Richardson, J., Zheng, Z., Chatterjee, S., Nagai, D., and Shen, Y., 2012, arXiv:1104.3550v2

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 1009, 1038

Rosario, D. J., McGurk, R. C., Max, C. E., et al. 2011, 2011arXiv1102.1733R

Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, ApJ, 697, 1634

Ruderman, J. T., Ebeling, H., 2005, ApJ, 623, 81

Salvato, M., et al. 2009, ApJ, 690, 1250

Salvato, M., et al. 2011, ApJ, 742, 61

Sargent, W. L. W., & Turner, E. L. 1977, ApJL, 212, L3

Schawinski, K,, Treister, E.,Urry, C. M., et al. 2011, ApJ, 727, 31

Schlegel, D.J., Abdalla, F., Abraham, T., et al. 2011, arXiv:1106.1706

Scoville, N., Abraham, R. G., Aussel, H., et al., 2007, ApJS, 172, 38

Seljak, 2000, MNRAS, 318, 2035

Shakura, N. I., & Sunyaev, R. A. 1976, MNRAS, 175, 613

Shane, C. D., & Wirtanen, C. A. 1967, in Pub. Lick Obs. Vol. 22, part 1

Shapley, H., & Ames, A. 1932, Annals of Harvard College Observatory, 88, 41

Serber, W., Bahcall, N., Menard, B., & Richards, G. 2006, ApJ, 643, 68

Shankar, F., Weinberg D. H., et al. 2009, ApJ, 690, 20

Shankar F., et al., 2010, ApJ, 718, 231

Shankar F., 2010, IAUS, 267, 248

Shen Y., Strauss, M. A., Ross, N. P., Hall, P. B., et al. 2009, ApJ, 697, 1656

Shen Y., 2009, ApJ, 704, 89

Shen Y., 2010, ApJ, 719, 1693

Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119

Sheth R. K., Mo H. J., Tormen G. 2001, MNRAS, 323, 1

Silverman, J. D.; Kovač, K., Knobel, C., 2009, ApJ, 695, 171

Silverman, J. D., et al. 2011, ApJ, 743, 2

Sivakoff G. R., Martini P., Zabludoff A. I., Kelson D. D., Mulchaey J. S., 2008, ApJ, 682, 803

Smith, R. E., et al. 2003, MNRAS, 341, 1311

Soltan, A., & Hasinger, G. 1994, A&A, 288, 77

Spergel, D. N., Bean, R., Dore, O., et al. 2007, ApJS, 170, 377408

Springel, V., 2005, MNRAS, 364, 1105

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629636

Starikova, S. et al., 2011, ApJ, 741, 15

Tal, T., van Dokkum P. G., Nelan, J., et al. 2009, ApJ, 138, 1417

Tasse C., Röttering H., Best P. N., 2011, A&A, 515, 127

Thomas T., Katgert P., 2006, A&A, 446, 31

Tinker, J. L., Weinberg, D. H., Zheng, Z., Zehavi, I., 2005, ApJ, 631, 41

Trump, J. R., Impey, C. D., McCarthy, P. J., et al. 2007, ApJS, 172, 383

Trump, J. R., Impey, C. D., Elvis, M., et al. 2009, ApJS, 696, 1195

Trump, J. R., et al. 2011, ApJ, 733, 60

Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T., 2003, ApJ,598, 886

Urry, C. M., Padovani, P., 1995, Publications of the Astronomical Society of the Pacific, v.107, p.803

van den Bosch, F. C., 2002, MNRAS, 331, 98

van de Weygaert R., Babul A., 1994, ApJ, 425, L59

Veilleux, S., et al. 2009, ApJ, 701, 587

Vikhlinin, A., & Forman, W. 1995, ApJL, 455, L109

Weinberg, N. N. & Kamionkowski, M. 2003), MNRAS, 341, 251262

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341358

White, S. D. M., Davis, M., Efstathiou, G., & Frenk, C. S. 1987, Nature, 330, 451453

Yang, Y., Mushotzky, R. F., Barger, A. J., & Cowie, L. L. 2006, ApJ, 645, 68

York, D. G., et al. 2000, AJ, 120, 1579

Zehavi, I., et al. 2004, ApJ, 608, 16

Zehavi, I., et al. 2005, ApJ, 621, 22

Zehavi, I., et al. 2010, arXiv:1005.2413

Zheng, Z., et al. 2005, ApJ, 633, 791

Zheng, Z., Coil, A. L., Zehavi, I., 2007, 667, 760

Zwicky, F., Herzog, E., & Wild, P. 1968, Catalogue of galaxies and of clusters of galaxies (Pasadena: California Institute of Technology, 1961-1968)